An Anaplastic Lymphoma Kinase Pathway Signature is associated with Cell De-differentiation, Neoadjuvant Response and Recurrence Risk in Breast Cancer

CURRENT STATUS: POSTED

Yong Wu  yongwu@cdrewu.edu
Charles Drew University of Medicine and Science
Corresponding Author
ORCiD: 0000-0002-1698-0897

Dingxie Liu
Bluewater Biotech LLC

DOI: 10.21203/rs.2.12140/v1

SUBJECT AREAS
Cancer Biology  Oncology

KEYWORDS
Anaplastic Lymphoma Kinase, Gene signature, Breast cancer, Differentiation, Prognosis prediction, Neoadjuvant Response, Disease free survival
Abstract

The role of ALK signaling in the pathogenesis of breast cancer (BC) is not clear. Previously we generated a gene signature for ALK pathway based on the difference of gene expression profiles between tumor cells with activated and inactive ALK pathway. Here, this signature was used to compute ALK pathway activity in BC samples from 42 microarray datasets, and the associations between ALK pathway score and the clinical outcome were examined by logistic regression and survival analysis. Our results indicated that high ALK pathway activity was a significant risk factor for the presence of higher-grade breast cancer with loss of ER and PR expression in the 42 datasets (n= 6381). ALK pathway activity was also positively associated with pathological complete response (pCR) in 15 datasets annotated with patient’s neoadjuvant response information (n= 2093, overall OR 1.67, p=2.00E-12), and with recurrence risk in 30 datasets annotated with patient’s survival information (n= 4678, overall HR 1.21, p=3.31E-08). The associations of ALK pathway activity with pCR and recurrence were more significant in HER2 negative and grade 1&2 tumors than in HER2 positive and grade 3 tumors. Notably, the association between ALK and tumor recurrence was statistically significant in patient with age>50 but not ≤50 years old, in patient with positive but not negative lymph node and in patients with residual disease but not pCR following neoadjuvant chemotherapy. These data indicate that ALK may be involved in BC tumorigenesis and ALK pathway signature represents a novel biomarker for BC clinical management.

Background

Breast cancer (BC) is the most common cancer among women, accounting for approximately a quarter of all new cancer cases diagnosed in women worldwide [1]. Recent genomic studies demonstrate that BCs are highly heterogeneous in their molecular
biology. In addition to those well-known genetically altered oncogenes and tumor suppressor genes, including BRCA1, EGFR, FGFR, MET, PI3K, TP53 and RB1, many lesser-known or lesser-characterized genes in the case of BC, such as anaplastic lymphoma kinase (ALK), were found to be mutated or amplified in BC [2, 3].

ALK is a transmembrane tyrosine kinase receptor belonging to the insulin receptor superfamily. Aberrant activation of ALK is involved in the tumorigenesis of a subset of haematopoietic, epithelial and mesenchymal neoplasms such as anaplastic large cell lymphoma, lung cancer, inflammatory myofibroblastic tumors and neuroblastoma [4–7].

The role of ALK signaling in the pathogenesis of BC is not clear. Conventionally, pathway activation is assessed by methods including immunohistochemistry that detects the levels of pathway-related protein expression or fluorescence in situ hybridization and quantitative PCR that detects amplification/overexpression of related oncogenes. The disadvantage of these methods is that they may sometimes be unreliable because most pathways can be activated at multiple points. The latest advances in high-throughput genomic technologies provide alternative strategies for semi-quantifying pathway activity through analyzing the expression profile of a pathway-specific gene signature by using approaches such as Bayesian binary regression (BinReg) that was developed by Nevins group [8, 9]. We previously generated a gene signature for ALK pathway based on the difference of gene expression profiles between tumor cells with activated and inactive ALK pathway [10]. In the current study, ALK pathway activity in BC samples from 42 microarray datasets was computed and its associations with the clinical outcome were examined. Our results indicate that ALK pathway is associated with cell de-differentiation, neoadjuvant response and recurrence risk in BC.

Methods
2.1 Microarray datasets and patient cohorts

Microarray datasets from Affymetrix U133 GeneChip microarray platforms (including HU133A, HU133A 2.0 and HU133 Plus 2.0) were analyzed since ALK pathway signature used in this study was generated from datasets of these platforms. To avoid selection bias, only datasets with more than 50 samples were chosen in this study. When performing this study, we found 42 publicly available datasets annotated with neoadjuvant response or cancer recurrence information meet the above-mentioned requirements. Among the 42 microarray datasets, 15 datasets have patient’s neoadjuvant response information and 30 datasets have cancer recurrence information (3 datasets contain both information). The clinical characteristics of BC patients from these datasets were described in Supplementary Table 1.

The raw CEL files of datasets were downloaded from Gene Expression Omnibus (GEO), except for dataset MDA133 that was downloaded from MD Anderson (https://bioinformatics.mdanderson.org/public-datasets/). The CEL files were normalized using Microarray Suite 5.0 (MAS5.0) and Robust Multi-array Average (RMA) approaches in R environment, respectively. Array quality was assessed by R simpleaffy package [11]. Batch effects for arrays from different hybridization dates were estimated using principal component analysis. ComBat program was used to eliminate batch effects when multiple datasets were merged or apparent batch effects were observed in a single dataset [12]. For datasets GSE3494, GSE2990, GSE6532 GSE7390, the patient survival data were updated and duplicate samples were removed according to the curated clinical data made by Dr. Jonas Bergh’s that is available in GEO database (accession number: GSE83232).

2.2 Pathway activity prediction by BinReg

Using the BinReg approach to generate pathway signatures and predict pathway activities of individual samples has been described in detail before [8, 9]. Briefly, the gene
expression patterns of two sets of samples (with one pathway being ‘on’ and ‘off’ respectively) were analyzed, and the pathway-specific informative genes (signature genes) were identified. Principal components were then used to compute weights for each of signature genes, such that the weighted average of expression levels showed a clear ability to distinguish the pathway “on” and “off” group. By applying binary regression on the principal components to the gene expression dataset of an unknown sample, a probability score of pathway activity for that sample was produced.

2.3 ALK Pathway signature

Two microarray datasets were used to generate ALK pathway signature as we described previously [10]. Briefly, the gene expression data of anaplastic largecell lymphoma cell line TS treated with or without ALK inhibitors A2 or A3 (GSE6184) was used as training set to generate signature. This signature was then validated by the gene expression data of TS cells with or without knock-down of ALK (GSE6184) and the expression data of lung cancer cell line NCI-H2228 treated with or without ALK inhibitor CH5424802 (GSE2511817).

2.4 Statistics

Statistics was performed by using R packages including Metafor [13], Survival [14], and Survminer. Odds-ratios for the associations of neoadjuvant response with risk scores were calculated using logistic regression. Kaplan-Meier survival curves with log-rank test and cox proportional hazards regression were used to analyze the association between disease-free survival and risk scores. Disease free survival (DFS) was defined as the time from surgery to the first confirmed relapse or metastasis. In this study distant metastasis-free survival (DMFS) was preferred to be used in survival analysis, while when the data is not available, relapse-free survival (RFS) data was used. The overall hazard ratio (HR) of a variable of interest was calculated using a random-effects model. The
significance of the overall effects across multiple datasets was estimated by Z test.

To examine the association of ALK pathway with neoadjuvant response and cancer recurrence in different subgroups of BC, the 42 microarray datasets were merged into 4 cohorts. Cohort 1 was merged from the 15 datasets (MDA133, GSE16446, GSE18728, GSE18864, GSE20194, GSE20271, GSE25055, GSE25065, GSE26639, GSE32646, GSE37946, GSE41998, GSE42822, GSE50948 and GSE66305) that has patient neoadjuvant response information. Cohort 2 contains 16 datasets (GSE11121, GSE12276, GSE2034, GSE17705, GSE2603, GSE20685, GSE2990, GSE26971, GSE3494, GSE45255, GSE58812, GSE6532, GSE65194, GSE7390, GSE9195, GSE88770) in which DMFS information is available and the longest follow-up time is > 8 years. Cohort 2 comprises 8 datasets (GSE12093, GSE20711, GSE21653, GSE31519, GSE42568, GSE1456, GSE7378, GSE71258) in which only RFS information is available and the longest follow-up time is > 8 years. The remaining 6 datasets (GSE16391, GSE16446, GSE17907, GSE19615, GSE25055 and GSE25065) that have either RFS or DMFS information were merged as cohort 4 since these datasets have a short follow-up time. Combat program was used to remove the batch effects in the cohort. It is worth noting that 3 datasets (GSE16446, GSE25055 and GSE25065) were present in both Cohort 1 and Cohort 4, therefore when both these 2 datasets were used for same statistical analysis, these 3 datasets were removed from cohort1. The disease free survival time was censored at 8 years in cohort 2 and cohort 3 while censored at 5 years in cohort 4.

The datasets used in this study were generated from different research groups and contain different clinical variables. The sample size will be too small if all the variables were put together for adjustment in multivariate regression analysis. Therefore, in this study, ALK pathway score was tested with adjustment for one variable each time in multivariate regression analysis.
All statistical analyses were two-sided and considered significant when \( p < 0.05 \).

Results

3.1 ALK pathway signature score is associated with de-differentiation of BC

The association between ALK pathway signature score and seven clinical characteristics of BC were analyzed in 42 affymatrix datasets individually by univariate regression analysis. Figure 1 showed that in most of the datasets, ALK pathway signature score was a significant unfavorable factor for ER-positive (overall OR 0.30, 95% CI, 0.26-0.35, \( p=3.02\times10^{-58} \)) and PR-positive (overall OR, 0.45, 95% CI 0.39-0.51, \( p=5.43\times10^{-32} \)) tumors, while it was a significant favorable factor for grade 3 tumors (overall OR 2.35, 95% CI 2.01-2.75, \( p=9.69\times10^{-27} \)). By contrast, no apparent associations of ALK pathway scores with age, HER2 status, lymph node status and tumor size were observed in most of the datasets (Fig. 2).

The 42 datasets were further merged into four cohorts, and similar associations of ALK pathway scores with the seven clinical characteristics were observed in the four merged cohorts (Supplementary Fig. 1). Multivariate logistic regression was also performed in these merged cohorts to assess whether the association of ALK pathway scores with estrogen receptor (ER) status, progesterone receptor (PR) status and tumor grade remain significant after adjustment for other covariates. Overall, the ORs for association of ALK pathway score with ER status stay constant when the pathway score was tested separately with correction for age, tumor grade, and tumor size, status of PR, HER2 or lymph node (Fig. 3A). Similar trend was also observed for the associations of ALK pathway scores with tumor grade or PR status, except that the association with PR status was almost abolished when adjusted for ER status (Fig. 3B and C). These results indicate that ALK pathway score is an independent risk factor for the presence of ER-negative and high grade BC, or in the other words, ALK pathway score is an independent factor associated with de-
differentiation of BC since loss of ER expression and high pathological grade are the features of de-differentiated BC [15].

3.2 ALK pathway signature score is associated with pathological complete response (pCR) of BC

Logistic regression analysis showed that ALK pathway score was positively associated with pCR rate of BC in 15 individual datasets that contain patient’s neoadjuvant response information (overall OR 1.67, 95% CI 1.45-1.93, p=2.00E-12) (Fig. 4A). This association reached statistical significance (p<0.05) in 8 of the 15 datasets, and approached the borderline of significance (p=0.08) in 2 of the datasets (Fig. 4A). As to the remaining 5 datasets with p > 0.08, two (GSE16446, and GSE18864) contain only ER-negative BC samples, and another two (GSE37946 and GSE50948) are mainly composed of HER2-positive samples (Supplementary Table 1), suggesting that ER and HER2 status affect the association between ALK pathway score and pCR rate. This effect was also observed in the cohort merged from the 15 datasets. As shown in Fig. 4B, an increase of one standard deviation (SD) of ALK pathway score is associated with 62% and 85% increase of pCR rate in ER-positive (95% CI 1.28-2.04, p=5.31E-5) and HER2-negative (95% CI 1.63-2.10, p=1.80E-21) BC respectively, while only with 17% and 28% increase in ER-negative (95% CI 1.02-1.34, p=0.03) and HER-positive (95% CI 1.03-1.58, p=0.02) BC respectively. In addition, ALK pathway score is also more strongly associated with pCR rate in grade 1&2 tumors (OR 2.16, 95% CI 1.63-2.87, p=8.09E-08) than in grade 3 tumors (OR 1.36, 95% CI 1.17-1.59, p=9.59E-05).

Multivariate logistic regression analysis on the merged cohort showed that the association of ALK pathway scores with pCR was hardly affected after adjustment for other clinical variables, except that the OR values decreased moderately when adjusted by ER or PR status (Fig. 4C).
3.3 ALK pathway signature score is associated with recurrence risk of BC

Logistic regression analysis showed that ALK pathway score was positively associated with recurrence risk of BC in 30 individual datasets in which patient’s survival information is available (overall HR 1.21, 95% CI 1.13-1.29, p=3.31E-08) (Fig. 5A). Among the 30 datasets, GSE16446, GSE31519, GSE58812 contain only ER-negative samples and GSE17907 is mainly composed of Her2-positive samples (Supplementary Table 1). ALK pathway score failed to achieve positive association significantly with recurrence risk in all of these 4 datasets. Four additional datasets (GSE 20711, GSE2603, GSE2990 and GSE7378) have sample size <100. Among the remaining 22 datasets, the association of ALK pathway with recurrence risk reached statistical significance in 11 datasets (Fig. 5A). The positive association of ALK pathway score with recurrence risk remain statistically significant in three cohorts that were merged from the 30 datasets in Cox regression analysis (p=7.29E-07, 8.98E-05, 2.54E-03, respectively) (Fig. 5B-D) and Kaplan-Meier survival curve analysis (p=1.29E-06, 1.79E-03, 2.50E-03, respectively) (Fig. 6A-C).

Like the association with pCR, the association of ALK pathway score with recurrence risk was also more significant in HER2 negative (p=9.00E-03, 1.25E-03, 7.85E-05, respectively) and grade 1&2 (p=2.42E-03, 9.81E-03, 0.07, respectively) tumors than in HER2 positive (p=0.14, 0.11, 0.10, respectively) and grade 3 (p=0.36, 0.34, 0.38, respectively) tumors in the three cohorts (Fig. 5B-D). In addition, the ALK pathway score was also more significantly associated with cancer recurrence in patient with age>50 (p=2.31E-06, 0.03, 5.96E-03, respectively) than with age ≤50 years old (p=0.49, 0.24, 0.10, respectively) and in patient with positive lymph node (p=9.52E-06, 3.42E-04, 7.52E-03, respectively) than with negative lymph node (p=0.17, 0.18, 0.21, respectively) in the three cohorts (Fig. 5B-D).

3.4 ALK pathway signature score is associated with recurrence risk only in
patients with residual disease (RD) following neoadjuvant chemotherapy

Our data showed that ALK pathway score is positively associated with both pCR rate and recurrence risk of BC. These results seem to be controversial since BC patients with pCR generally have better prognosis than patients with RD [16]. A potential explanation is that ALK pathway signature score is associated with recurrence risk only in RD but not in pCR patients. To test this hypothesis, samples in cohort 4 that were annotated with both pCR and survival information were tested. Kaplan-Meier analysis showed that patients with pCR have much lower relapse rate than patients with RD (black vs grey, 5-year DFS rate: 0.93 vs 0.69, p=1.96E-05) (Fig. 6D). When RD patients were further stratified into two subgroups based on ALK pathway scores, the subgroup with low ALK pathway score had better DFS rate than that with high ALK score (blue vs violet, 5-year DFS rate: 0.60 vs 0.77, p=1.83E-04) (Fig. 6D). By contrast, the two subgroups formed by further stratifying the pCR patients based on ALK pathway scores showed similar DFS rates (green vs red, p = 0.67) (Fig. 6D).

Discussion

Abnormal activation of ALK signaling through gene rearrangement, amplification, mutation or protein overexpression has been found in various human cancers, including neuroblastoma, glioblastoma, lung cancer, rhabdomyosarcoma, ovarian cancer, melanoma, colorectal carcinoma, astrocytoma, Ewing’s sarcoma and retinoblastoma [4–7]. It is controversial whether alterations of ALK gene exist in BC. While Perez-Pinera et al. [17] reported that activated ALK strongly expressed in different histological subtypes of BC, Fukuyoshi et al. [18] and Lerebours et al. [19] failed to observe such abnormality in BC patients. Recently, Kim et al. [20] demonstrated that ALK copy number significantly increased in inflammatory BC (IBC) and was associated with higher recurrence risk in IBC patients. In addition, TCGA (The Cancer Genome Atlas Network) genomic analysis showed
amplification of ALK gene in 43 of 476 BC samples [3]. In this study, we found that higher ALK pathway score was associated with cell de-differentiation and higher recurrence risk in BC. These results support the notion that active ALK signaling has a functional role in the pathogenesis of BC.

ALK activates multiple pathways, including phospholipase C γ (PLCγ), Janus kinase -signal transducer and activator of transcription 3 (JAK/STAT3), Phosphoinositide 3-kinase–AKT (PI3K/AKT), mitogen-activated protein kinase (MAPK), sonic hedgehog, JUNB and RAPGEF1/RAP1 GTPase signaling cascades [7]. Aberrant activation of some of these pathways, such as JAK/STAT3, PI3K/AKT and sonic hedgehog, has been found to promote tumor metastasis and be associated with cell de-differentiation in BC [21–23]. Which of these downstream pathways mediate the oncogenic effects of ALK signaling in BC need to be further clarified.

An interesting finding in this study is that the association between high ALK and tumor recurrence differs by age group (>50 vs. ≤50 years), with ALK significantly associated with risk of recurrence only in the middle-aged and elderly patients. This characteristic of ALK has not been reported in other cancer studies. The reason for this is unclear, although it might be due to the complex hormonal environment of young women. A retrospective analysis performed on 260 elderly and 294 middle-aged patients with primary BC shows that negative lymph node status, small tumor size and positive ER status are favorable indicators of survival in both the elderly and the middle-aged patients [24]. In addition, another study shows that ER/PR status and HER2 gene amplification or overexpression are prognostic factors in elderly patients with BC [25]. ALK, ER/PR and HER2 share common downstream activation pathways, such as Ras/MEK/ERK and PI3K/Akt/mTOR, which ultimately lead to increased transcription, cell proliferation, growth and survival [26–30]. This also may be used to explain that they have similar prognostic effects on the same
type of patients.

Another interesting finding in this study is that the relationship between ALK and tumor recurrence exists in patient with lymph node metastasis, but not in the node negative patients. A potential explanation is that the growth/survival of BC with metastastic capabilities (such as BC with positive lymph node) will benefit more from active ALK signaling than that without the capability (such as BC with negative lymph node). To establish a distant tumor, a metastatic cancer cell needs to survive and adapt to a new environment, rebuild cell-matrix interactions, from a micrometastasis and finally restart an unlimited growth process. Metabolic rewiring is a key factor for tumor cells to finish this metastatic cascade, and several pathways including TGF-β and hypoxic signaling are involved in this rewiring processes [31]. Recent studies revealed that the upregulation of hypoxia-inducible factors under hypoxic conditions, as well as the induction of VEGF secretion induced by TGF-β, were both in an ALK-dependent manner in different tumor cells [32, 33], suggesting a potential connection between ALK signaling and metabolic rewiring. Probably through this connection, ALK signaling plays more active roles in the molecular pathogenesis of metastatic cells than the primary tumor cells that may not benefit from metabolic rewiring.

A number of prognostic gene signatures have been developed for prediction of neoadjuvant response or recurrence risk in BC [34-40]. In this study, an ALK pathway gene signature was found to well predicted both the neoadjuvant response and recurrence risk in multiple datasets encompassing >5000 BC cases, suggesting it as a potentially promising biomarker for BC prognosis and management. Compared with those gene signatures reported previously, the unique feature of ALK pathway signature is that it is specifically associated with recurrence in BC patients with age >50, with lymph node metastasis, or with RD after neoadjuvant chemotherapy, which indicates that these
patients may particularly benefit from using this signature for risk estimation. Further validation of ALK pathway signature in additional independent and clinically relevant analyses will be necessary before entering clinical trials.

In summary, this study highlights that ALK pathway gene signature represents a potentially promising biomarker for guiding clinical management of BC. Our results also support the notion that ALK signaling may have an oncogenic role in the pathogenesis of BC and therefore may be a potential molecular target for BC therapy.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69:7–34.
2. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. Nature. 2012;486:400–4.
3. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature. 2012;490:61–70.
4. Morris SW, Kirstein MN, Valentine MB, Dittmer KG, Shapiro DN, Saltman DL, et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin’s lymphoma. Science. 1994;263:1281–4.
5. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. 2007;448:561–6.
6. Rikova K, Guo A, Zeng Q, Possemato A, Yu J, Haack H, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell. 2007;131:1190–203.
7. Della Corte CM, Viscardi G, Di Liello R, Fasano M, Martinelli E, Troiani T, et al. Role and targeting of anaplastic lymphoma kinase in cancer. Molecular Cancer.
8. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006;439:353-7.

9. Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, et al. A pathway-based classification of human breast cancer. Proc Natl Acad Sci USA. 2010;107:6994-9.

10. Liu D, Liu X, Xing M. Activities of multiple cancer-related pathways are associated with BRAF mutation and predict the resistance to BRAF/MEK inhibitors in melanoma cells. Cell Cycle. 2014;13:208-19.

11. Wilson CL, Miller CJ. Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. Bioinformatics. 2005;21:3683-5.

12. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118-27.

13. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010;36:1-48.

14. Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer-Verlag; 2000. https://www.springer.com/us/book/9780387987842. Accessed 17 Mar 2019.

15. Fisher ER, Redmond CK, Liu H, Rockette H, Fisher B. Correlation of estrogen receptor and pathologic characteristics of invasive breast cancer. Cancer. 1980;45:349-53.

16. Bear HD, Anderson S, Brown A, Smith R, Mamounas EP, Fisher B, et al. The effect on tumor response of adding sequential preoperative docetaxel to preoperative doxorubicin and cyclophosphamide: preliminary results from National Surgical Adjuvant Breast and Bowel Project Protocol B-27. J Clin Oncol. 2003;21:4165-74.

17. Perez-Pinera P, Chang Y, Astudillo A, Mortimer J, Deuel TF. Anaplastic lymphoma...
kinase is expressed in different subtypes of human breast cancer. Biochem Biophys Res Commun. 2007;358:399-403.

18. Fukuyoshi Y, Inoue H, Kita Y, Utsunomiya T, Ishida T, Mori M. EML4-ALK fusion transcript is not found in gastrointestinal and breast cancers. Br J Cancer. 2008;98:1536-9.

19. Lerebours F, Callens C, Vacher S, Hatem R, Guinebretière J-M, Bièche I. Rare overexpression of anaplastic lymphoma kinase gene in inflammatory and non-inflammatory breast cancer. Eur J Cancer. 2013;49:2774-6.

20. Kim MH, Lee S, Koo JS, Jung KH, Park IH, Jeong J, et al. Anaplastic lymphoma kinase gene copy number gain in inflammatory breast cancer (IBC): prevalence, clinicopathologic features and prognostic implication. PLoS ONE. 2015;10:e0120320.

21. Banerjee K, Resat H. Constitutive activation of STAT3 in breast cancer cells: A review. Int J Cancer. 2016;138:2570-8.

22. Kenna MM, McGarrigle S, Pidgeon GP. The next generation of PI3K-Akt-mTOR pathway inhibitors in breast cancer cohorts. Biochim Biophys Acta Rev Cancer. 2018;1870:185-97.

23. Riobo-Del Galdo NA, Lara Montero Á, Wertheimer EV. Role of Hedgehog Signaling in Breast Cancer: Pathogenesis and Therapeutics. Cells. 2019;8.

24. Alberts AS, Falkson G, van der Merwe R. Factors influencing prognosis in elderly patients with primary breast cancer. S Afr J Surg. 1991;29:8-10.

25. Cappellani A, Vita MD, Zanghì A, Cavallaro A, Piccolo G, Majorana M, et al. Prognostic factors in elderly patients with breast cancer. BMC Surg. 2013;13 Suppl 2:S2.

26. Hrustanovic G, Bivona TG. RAS-MAPK in ALK targeted therapy resistance. Cell Cycle. 2015;14:3661-2.

27. Moore NF, Azarova AM, Bhatnagar N, Ross KN, Drake LE, Frumm S, et al. Molecular
rationale for the use of PI3K/AKT/mTOR pathway inhibitors in combination with crizotinib in ALK-mutated neuroblastoma. Oncotarget. 2014;5:8737–49.

28. Ciruelos Gil EM. Targeting the PI3K/AKT/mTOR pathway in estrogen receptor-positive breast cancer. Cancer Treat Rev. 2014;40:862–71.

29. Díaz-Serrano A, Angulo B, Domínguez C, Pazo-Cid R, Salud A, Jiménez-Fonseca P, et al. Genomic Profiling of HER2-Positive Gastric Cancer: PI3K/Akt/mTOR Pathway as Predictor of Outcomes in HER2-Positive Advanced Gastric Cancer Treated with Trastuzumab. Oncologist. 2018;23:1092–102.

30. Paplomata E, O’Regan R. The PI3K/AKT/mTOR pathway in breast cancer: targets, trials and biomarkers. Ther Adv Med Oncol. 2014;6:154–66.

31. Elia I, Doglioni G, Fendt S-M. Metabolic Hallmarks of Metastasis Formation. Trends in Cell Biology. 2018;28:673–84.

32. Martinengo C, Poggio T, Menotti M, Scalzo MS, Mastini C, Ambrogio C, et al. ALK-dependent control of hypoxia-inducible factors mediates tumor growth and metastasis. Cancer Res. 2014;74:6094–106.

33. Seystahl K, Tritschler I, Szabo E, Tabatabai G, Weller M. Differential regulation of TGF-β-induced, ALK-5-mediated VEGF release by SMAD2/3 versus SMAD1/5/8 signaling in glioblastoma. Neuro-oncology. 2015;17:254–65.

34. van ’t Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AAM, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415:530–6.

35. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, et al. A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. New England Journal of Medicine. 2004;351:2817–26.

36. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, et al. Gene expression profiling
in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst. 2006;98:262–72.

37. Ma X-J, Salunga R, Dahiya S, Wang W, Carney E, Durbecq V, et al. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. Clin Cancer Res. 2008;14:2601–8.

38. Bild AH, Parker JS, Gustafson AM, Acharya CR, Hoadley KA, Anders C, et al. An integration of complementary strategies for gene-expression analysis to reveal novel therapeutic opportunities for breast cancer. Breast Cancer Research. 2009;11:R55.

39. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol. 2009;27:1160–7.

40. Filipits M, Rudas M, Jakesz R, Dubsky P, Fitzal F, Singer CF, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res. 2011;17:6012–20.

Declarations

Acknowledgments

The results here are based upon the breast cancer microarray data from public repositories. We are very grateful to the investigators who contributed these microarray datasets to the public domain. This work was supported in part by NIH/NCI SC1CA200517 and Accelerating Excellence in Translational Science Pilot Grant G0812D05 to Y. Wu.

Conflict of interest statement

Liu has equity interest in Bluewater Biotech LLC.

Figures
Figure 1

Associations of ALK pathway signature score with tumor grade and status of ER and PR in 42 individual breast cancer datasets. A) ER status (ER positive as event and ER negative as non-event); B) PR status (PR positive as event and PR negative as non-event); C) Tumor grade (grade 3 as event and grade 1&2 as non-event). The OR (per one standard deviation increment) with the increase of ALK pathway signature score (used as continuous variable) was analyzed using univariate logistic regression in 42 individual breast cancer datasets. The overall effect was calculated through a random-effects (RE) model. The ORs were shown in forest plots, in which the squares and horizontal lines represent the OR and 95% CI for the individual cohorts, while the diamonds represent the OR and 95% CI for the overall estimate. Only datasets with event case >5, non-event case >5 and total case >30 were included for regression analysis, therefore different number of datasets may be analyzed for different dependent variables.
Figure 2

Associations of ALK pathway signature score with age, tumor size and status of HER2 and node in 42 individual breast cancer datasets. A) Patient age (>50 as event and ≤50 as non-event); B) HER2 status (HER2 positive as event and HER2 negative as non-event); C) Lymph node status (node positive as event and node negative as non-event); D) Tumor size (>2cm as event and ≤20 as non-event).

The OR (per one standard deviation increment) with the increase of ALK pathway signature score (used as continuous variable) was analyzed using univariate logistic regression in 42 individual breast cancer datasets. Only datasets with event case >5, non-event case >5 and total case >30 were included for regression analysis.
Associations of ALK pathway signature score with tumor grade and status of ER and PR in four merged breast cancer cohorts. A) ER status (ER positive as event and ER negative as non-event); B) PR status (PR positive as event and PR negative as non-event); C) Tumor grade (grade 3 as event and grade 1&2 as non-event). ALK pathway score was used as continuous variable and tested with adjustment for one variable each time in multivariate regression analysis. The OR [95%CI] is presented per one-SD increment. The four cohorts were merged from the 42 individual breast cancer datasets as described in Material and Methods. The three datasets (GSE16446, GSE25055 and GSE25065) that were present in both cohort 1 and cohort 4 were removed from cohort 1 in multivariate regression analysis.
Figure 4

Associations of ALK pathway signature score with pCR in breast cancer. ALK pathway score was used as continuous variable and the OR [95%CI] is presented per one-SD increment. A) Univariate logistic regression analysis of associations of ALK pathway score with pCR in 15 individual breast cancer datasets. B) Comparison of the associations of ALK pathway score with pCR in different subgroups of breast cancer. The 15 individual breast cancer datasets with pCR information were merged into a single cohort and Combat program was used to remove the batch effects in the cohort. Logistic regression analysis was performed in different subgroups of breast cancer patients as indicated. C) Multivariate logistic regression analysis of associations of ALK pathway score with pCR. The cohort merged from 15 individual breast cancer datasets was analyzed. ALK pathway score was tested with adjustment for one variable each time in
multivariate regression analysis.

Figure 5

Associations of ALK pathway signature score with recurrence risk in breast cancer. The associations of ALK pathway score with recurrence risk were analyzed by univariate logistic regression in 30 individual breast cancer datasets (A), and were also compared among different subgroups of breast cancer in three

23
independent cohorts that were merged from the 30 datasets, including cohort 2 (B), cohort 3 (C) and cohort 4 (D). The detail of these 3 breast cancer cohorts are described in Material and Methods. ALK pathway score was used as continuous variable and recurrence risk (indicated by HR) is presented per one-SD increment.

Figure 6
Kaplan-Meier analysis of the associations of ALK pathway signature score with recurrence risk in breast cancer. The patients were stratified into three groups based on ALK pathway signature score (high: ≥ 2/3 percentile; low ≤ 1/3 percentile; intermediate: < 2/3 percentile and >1/3 percentile) in cohort 2 (A)[], cohort 3 (B) and cohort 4 (C) respectively[and Kaplan-Meier analysis was performed to compare the probabilities of DFS among the three groups. To compare the associations of ALK pathway score with recurrence risk in breast cancer patients with pCR or RD following neoadjuvant chemotherapy, the samples
in cohort 4 that were annotated with both pCR and survival information were analyzed. The patients were first stratified into pCR and RD groups, and these two groups were further stratified into two subgroups based on the medium value of ALK pathway signature score, respectively. Kaplan-Meier analysis was then performed to compare the DFS probabilities between pCR and RD patients, between pCR patients with high and low ALK pathway scores, and between RD patients with high and low ALK pathway scores, respectively.

Supplementary Files

This is a list of supplementary files associated with the primary manuscript. Click to download.

Supplementary results.docx