Original Article
Combined expression of ZEB2 and ALDH1A1 is correlated with poor prognosis of breast cancer patients

Yuanyuan Yan1,2, Miao He1,2, Zhaojin Yu1,2, Mingli Sun1,2, Lin Zhao1,2, Haishan Zhao1,2, Weifan Yao1,2, Minjie Wei1,2

1Department of Pharmacology, School of Pharmacy, China Medical University, Shenyang, Liaoning Province, P.R. China; 2Liaoning Key Laboratory of Molecular Targeted Anti-tumor Drug Development and Evaluation, Shenyang, P.R. China

Received June 2, 2017; Accepted December 22, 2017; Epub March 15, 2018; Published March 30, 2018

Abstract: Aldehyde dehydrogenase 1 family member A1 (ALDH1A1) is an important marker of cancer stem cells (CSCs). Zinc finger E-box binding homebox 2 (ZEB2) is a key transcription factor in regulating epithelial-to-mesenchymal transition (EMT) process of cancer cells. The aim of this study was to investigate the association of ALDH1A1 and ZEB2 alone or combined expression with breast cancer progression and prognosis. Immunohistochemistry was performed to detect the expression of ALDH1A1 and ZEB2 in 274 breast cancer tissues. ZEB2 was positively related to ALDH1A1 in breast cancer tissues (P < 0.0001). High expression of ZEB2 or ALDH1A1 alone was associated with lymph node metastasis, TNM stage and worse clinical outcome. The combined expression of ZEB2 and ALDH1A1 in breast cancer patients was positively associated with lymph node metastasis (P < 0.001), advanced TNM stage (P < 0.001) and shorter OS and DFS (P < 0.01). ZEB2 and ALDH1A1 play important roles in breast tumor aggressiveness and prognosis, and ZEB2 and ALDH1A1 combined expression may be a more helpful indicator to identify high-risk breast cancer patients.

Keywords: Zinc finger E-box binding homebox 2, aldehyde dehydrogenase 1 family member A1, breast cancer stem cells, poor prognosis

Introduction
Breast cancer is the most common carcinoma in women worldwide [1]. In general, breast tumors are comprised of heterogeneous populations of cells. Of these, 1-2% cells are breast cancer stem cells (BCSCs) which are believed to initiate and promote tumorigenesis [2]. BCSCs have the ability to survive and proliferate in anchorage-independent conditions, and play a central role in tumor progression, spreading and relapse [3]. It will be of great help for diagnosis and therapy evaluation of breast cancer to find rational BCSCs biomarkers.

Several cell-surface markers have been used to characterize BCSCs, such as CD44+/CD24-/low, Lin, and functional markers like high aldehyde dehydrogenase (ALDH+) activity or the presence of an ABC transporter-dependent Hoechst side population [2]. Among them, ALDHs are a group of proteins that share highly conserved sequences essential for proliferation, differentiation, and survival, as well as the cellular response to oxidative stress [4]. It has been reported that ALDHs are involved in normal stem cells as well as cancer stem cells, including the ALDH1 family, ALDH2*2, ALDH3A1, ALDH4A1 and ALDH7A1 [5]. Particularly, ALDH1A1 has been suggested as a BCSC marker [6, 7]. Previous studies describe the association of ALDH1 expression with early metastasis and decreased survival [8, 9], however, the specific mechanisms how the ALDH1 phenotype contributes to malignant cell metastatic behavior and whether epithelial-to-mesenchymal transition (EMT) is involved in this process are yet to be established.

Epithelial-to-mesenchymal transition (EMT) is a cellular program that leads to a change from epithelial to mesenchymal phenotype, and involved in the migration and invasion of cancer cells [10]. There is increasing evidence showing that cancer cells, such as breast cancer [11],...
prostate cancer and others [12, 13], can be transformed into CSCs through EMT. Zinc finger E-box binding homebox 2 (ZEB2), belongs to a small family of transcriptional factors characterized by containing a homeodomain flanked by two separated zinc finger clusters [14]. Through interacting with different miRNAs, ZEB2 functions as a key player to regulate EMT and stem cell properties [15, 16]. However, there is no study about the association of ZEB2 and ALDH1A1 in stem phenotype and pathological characteristics. In our study, we found that the expression of ZEB2 and ALDH1A1 was positive correlation in breast cancer patients, and the
combined expression of ALDH1A1 and ZEB2 was positively correlated with lymph node metastasis, TNM stage and poor outcomes. ALDH1A1 and ZEB2 would be an independent prognostic marker in breast cancer.

Materials and methods

Clinical tissue samples

A total of 274 paraffin-embedded breast cancer tissues were from the First Affiliated Hospital of China Medical University between January 2006 and December 2008. All patients did not receive radiation therapy or chemotherapy prior to surgery. According to the pathological staining, breast cancer was diagnosed. The patient age, menopausal status, tumor type, tumor size, and lymph node metastasis were all obtained from clinical records. The histological grade was determined according to the World Health Organization classification system. Patients were staged based on tumor-node-metastasis (TNM) classification of the International Union Against Cancer [17]. This study was approved by the Medical Ethics Committee of China Medical University. The Medical Ethics Committee waived the need of written informed consents by patients because of the retrospective nature of this study.

Immunohistochemistry staining and evaluation

Immunohistochemistry was carried out as previously described [18]. In a word, after deparaffinizing, rehydrating and heating to retrieve antigen, the sections (4-µm) from paraffin-embedded tissue were incubated with primary antibodies against ALDH1A1 (Abcam, UK, 1:100), or ZEB2 (Novus Biologicals, 1:200) over night at 4°C. Negative control sections were incubated with PBS instead of the primary antibody. After washing in PBS, a biotin-marked secondary antibody was applied followed by streptavidin horseradish peroxidase (LSAB kit; Dako, Denmark). After the 3, 3-diaminobenzidine (DAB) staining, counterstaining with hematoxylin, dehydrating, and mounting, images were captured with Nikon Eclipse 80i microscope (Japan), a digital sight digital camera.
Two independent investigators evaluated these slides in a double-blinded manner. The immunoreactivity intensity was scored by negative (0), weak (1), moderate (2), and strong (3) staining. The percentage of stained cells was scored by using 5% increments (0, 5, 10, and 100%) as previously reported [19, 20]. The final score was calculated by multiplying intensity score with percentage of positively stained cells score, ranging from 0 to 300%. Depending on receiver operating characteristic (ROC) curves, final scores were used to determine the cut-off value for discriminating ZEB2 and ALDH1A1 positive or negative expression. The sensitivity and specificity for survival status (alive or dead) of breast cancer patients was plotted to generate ROC curves.

**Statistical analysis**

Statistical analyses were performed using SPSS 16.0 (SPSS, Inc., Chicago, IL, USA). Pearson’s Chi-square test or Fisher’s exact test was used for evaluating the association of ALDH1A1 and ZEB2 expression with pathological characteristics. Pearson’s rank correlation analysis was applied to assess the association of ALDH1A1 and ZEB2 expression. Mann-Whitney U test was used to analyze the expression of ZEB2 or ALDH1A1 in different lymph node metastasis or TNM stage. The Kaplan-Meier curves were plotted to estimate the survival difference, and the survival probabilities were assessed by a log-rank test. Analysis of predictive factors for overall survival (OS) and disease-free survival (DFS) was carried out with univariate and multivariate Cox proportional hazards regression models. P value < 0.05 was considered statistically significant.

**Results**

**Association of ZEB2 and ALDH1A1 expression with the pathological data of breast cancer patients**

The baseline characteristics of 274 breast cancer patients were summarized. The age of pa-
Expression of ZEB2 and ALDH1A1 in breast cancer patients

Patients ranged from 29 to 79 years, with an average age of 50.7 years. 146 (61.3%) cases had tumor sizes bigger than 2.0 cm and 92 (38.7%) cases had less than 2.0 cm. According to WHO histological grading criteria, 79.3% was grade II-III and 20.7% was grade I. According to the TNM staging system, 126 patients (71.2%) were classified as stage I-II, 51 (28.8%) as stage III-IV. 129 (47.6%) patients occurred lymph node metastasis.

Follow-up information included of 224 breast cancer patients from 25 to 77 months. 166 cases relapsed and 65 cases were found cancer-associated deaths. The average OS and DFS were 65.3 and 54.6 months, respectively. The 5-years survival was 82.1%.

We used immunohistochemistry to examine the expression of ZEB2 and ALDH1A1 in 274 samples from breast cancer patients. The positive staining of ALDH1A1 was mainly located in cytoplasm and ZEB2 in nucleus of breast cancer cells (Figure 1A). The cut-off scores of 172.5% and 155% were selected for determining positive expression of ALDH1A1 and ZEB2 by ROC curve analysis based on survival status, respectively (Figure 1B). The ZEB2 positive expression rate was 34.3% and the ALDH1A1 positive expression rate was 59.49% in 274 breast cancer samples.

The association of ZEB2 and ALDH1A1 expression with clinical pathological characteristics of breast cancer patients was shown in Table 1. ZEB2 high expression was associated with higher TNM stage (P<0.001), lymph node metastasis (P<0.001), and bigger tumor size (P=0.026). The ALDH1A1 high expression was associated with lymph node metastasis (P=0.001), higher TNM stage (P=0.005). Mann-Whitney U test also showed that there was higher expression of ZEB2 and ALDH1A1 in breast cancer patients with lymph node metastasis or higher TNM stage (Figure 2).

Association of the combined expression of ALDH1A1 and ZEB2 with lymph node metastasis and TNM stage in breast cancer

<table>
<thead>
<tr>
<th>Features</th>
<th>Node metastasis</th>
<th>TNM stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No (%)</td>
<td>Yes (%)</td>
</tr>
<tr>
<td>ALDH1A1 (high) and ZEB (high)</td>
<td>26 (32.1)</td>
<td>55 (67.9)</td>
</tr>
<tr>
<td>ALDH1A1 (high) and ZEB (low) or ALDH1A1 (low) and ZEB (high)</td>
<td>53 (55.8)</td>
<td>42 (44.2)</td>
</tr>
<tr>
<td>ALDH1A1 (low) and ZEB (low)</td>
<td>63 (66.3)</td>
<td>32 (33.7)</td>
</tr>
</tbody>
</table>

P-value obtained from Pearson’s Chi-square test or Fisher’s exact test.
Expression of ZEB2 and ALDH1A1 in breast cancer patients

Atlas (TCGA) database with 1168 breast cancer samples’ mRNA expression was downloaded. The significant positive correlation was found in mRNA expression of ALDH1A1 and ZEB2 by Pearson’s rank correlation analysis ($R^2 = 0.4646, P < 0.0001$, Figure 3A). We also found

Figure 4. Kaplan-Meier survival analysis of ZEB2 and ALDH1A1 expression in 224 breast cancer patients. A, B. Survival curves of ZEB2 positive/negative expression with overall survival and disease-free survival; C, D. Survival curves of ALDH1A1 positive/negative expression with overall survival and disease-free survival; E, F. Survival curves of the combined expression of ZEB2 and ALDH1A1 with overall survival and disease-free survival.
that ALDH1A1 protein expression was significantly positively correlated with ZEB2 protein expression in 274 breast cancer patients ($R^2=0.2368$, $P < 0.0001$, Figure 3B). These results suggested that the stem property of BCSCs was probably correlated with EMT process.

We further examined the association of the combined expression of ALDH1A1 and ZEB2 with lymph node metastasis and TNM stage (Table 2). Compared with tumors with ALDH1A1 low and ZEB2 high, ALDH1A1 high and ZEB2 low, or ALDH1A1 low and ZEB2 low expression, tumors with ALDH1A1 high and ZEB2 high expression were more associated with lymph node metastasis ($P < 0.01$) and higher TNM stage ($P < 0.001$).

**The combined expression of ALDH1A1 and ZEB2 is correlated with poor outcomes of breast cancer patients**

We then evaluated the association of ZEB2 expression alone, ALDH1A1 expression alone, and co-expression of ZEB2 and ALDH1A1 with the overall survival (OS) or disease-free survival (DFS) in breast cancer patients. The Kaplan-Meier analysis curves showed that patients with high ZEB2 expression had shorter OS ($P < 0.01$) and DFS ($P < 0.01$, Figure 4A and 4B). Patients with high ALDH1A1 expression had shorter OS ($P < 0.01$) and DFS ($P < 0.01$, Figure 4C and 4D). In addition, patients with the combined high ZEB2 and high ALDH1A1 expression had shorter OS ($P < 0.01$) and DFS ($P < 0.01$, Figure 4E and 4F) compared with the patients with the combined low ZEB2 and low ALDH1A1 expression, low ZEB2 and high ALDH1A1 expression, or high ZEB2 and low ALDH1A1 expression.

The Univariate and multivariate Cox regression analysis was used to evaluate the impact of ZEB2 and ALDH1A1 expression and pathological characteristics on OS and DFS in 224 breast cancer patients (seen in Table 3). The Univariate analysis showed that TNM stage (OS, $P < 0.001$; DFS, $P < 0.001$), lymph node metastasis (OS, $P < 0.001$; DFS, $P=0.007$), ZEB2 expression (OS, $P < 0.001$; DFS, $P < 0.001$) and ALDH1A1 expression (OS, $P < 0.001$; DFS, $P < 0.001$) were prognostic factors for breast cancer. Furthermore, in the multivariate Cox regression analysis, we found that TNM stage (OS, $P=0.001$; DFS, $P=0.001$) and ZEB2 expression (OS, $P=0.002$; DFS, $P < 0.001$) were independent prognostic factors for breast cancer.

**Discussion**

EMT process is a critical characteristic of breast cancer stem cells phenotype. ZEB2 is a key protein in regulating the process of EMT by acting as a repressor of E-cadherin. ZEB2 has been reported to mediate EMT and disease aggressiveness in ovarian and breast cancer [21]. Several studies indicated that the increased level of ZEB2 transcription was associated with invasion and metastasis in cancers with advanced stages [22]. It was reported that ZEB2 overexpression could be an independent prognostic marker in colorectal cancer [23]. Sayan also reported that ZEB2 overexpression was an independent prognostic factor in bladder cancer and positively correlated with a poor therapeutic outcome [24]. In this study, we found that ZEB2 expression had significantly positive association with lymph node metastasis and TNM advanced stage. In the Kaplan-Meier survival analysis, we found ZEB2 expression was significantly with shorter OS and DFS. All these results indicated that ZEB2 overexpression was associated with poor prognosis and clinical outcome.

ALDH1A1 is first indicated as a marker and a characteristic feature of primitive human hematopoietic stem cells (HSCs) isolated from bone marrow [25] and neural stem cells [26, 27]. It was reported that either mRNA or protein of ALDH1A1 high expression was correlated with poor OS and RFS (recurrence-free survival) in breast cancer patients [8, 28]. ALDH1A1 was the only ALDH1 isozyme capable of serving as a biomarker for predicting poor survival in breast cancer patients, and ALDH1A1 expression might be a good CSC marker and an important predictor of progression and poor survival [29]. In our study, we found ALDH1A1 expression was positively associated with lymph node metastasis and TNM stage. In addition, the Kaplan-Meier survival analysis revealed that the breast cancer patients with high ALDH1A1 expression had significantly shorter OS and DFS than those with low ALDH1A1 expression.

More importantly, we found that the ZEB2 expression was positively correlated with
Expression of ZEB2 and ALDH1A1 in breast cancer patients

Table 3. Univariate and Multivariate Cox regression analysis of the association of clinical pathological data with OS and DFS in breast cancer patients

<table>
<thead>
<tr>
<th>Features</th>
<th>OS Univariate analysis</th>
<th>DFS Univariate analysis</th>
<th>OS Multivariate analysis</th>
<th>DFS Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR (95% CI)</td>
<td>P</td>
<td>RR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>Age (years)</td>
<td>≤50/&gt;50</td>
<td>0.947 (0.576-1.5557)</td>
<td>0.83</td>
<td>0.743 (0.546-1.010)</td>
</tr>
<tr>
<td>Menopausal status</td>
<td>Pre/Post</td>
<td>0.831 (0.508-1.360)</td>
<td>0.461</td>
<td>0.745 (0.545-1.017)</td>
</tr>
<tr>
<td>Tumor size (cm)</td>
<td>≤2.0/&gt;2.0</td>
<td>0.656 (0.366-1.174)</td>
<td>0.156</td>
<td>0.679 (0.481-0.960)</td>
</tr>
<tr>
<td>Tumor type</td>
<td>Ductal/Lobular carcinoma</td>
<td>0.732 (0.356-1.507)</td>
<td>0.398</td>
<td>1.251 (0.752-2.081)</td>
</tr>
<tr>
<td>Histological grade</td>
<td>I/II-III</td>
<td>0.654 (0.321-1.332)</td>
<td>0.242</td>
<td>0.736 (0.487-1.112)</td>
</tr>
<tr>
<td>TNM stage</td>
<td>I-II-III</td>
<td>0.220 (0.126-0.384)</td>
<td>0.000</td>
<td>0.366 (0.252-0.531)</td>
</tr>
<tr>
<td>Lymph node metastasis</td>
<td>No/Yes</td>
<td>0.390 (0.230-0.662)</td>
<td>0.000</td>
<td>0.655 (0.481-0.891)</td>
</tr>
<tr>
<td>ALDH1A1 expression</td>
<td>Low/High</td>
<td>0.154 (0.070-0.338)</td>
<td>0.000</td>
<td>0.349 (0.249-0.490)</td>
</tr>
<tr>
<td>ZEB2 expression</td>
<td>Low/High</td>
<td>0.140 (0.080-0.244)</td>
<td>0.000</td>
<td>0.256 (0.184-0.356)</td>
</tr>
</tbody>
</table>

OS: overall survival; DFS: disease-free survival; RR: relative risk; 95% CI: 95% confidence interval.
Expression of ZEB2 and ALDH1A1 in breast cancer patients

ALDH1A1 expression in breast cancer tissues. The results may reveal the relationship of EMT process and the stem properties of breast cancer stem cells. Also, we found that co-expression of ZEB2 and ALDH1A1 was obviously associated with lymph node metastasis and advanced TNM stage. In addition, the Kaplan-Meier survival analysis showed that patients with ZEB2 high and ALDH1A1 high expression had much shorter OS and DFS. Hence, the combined expression of ZEB2 and ALDH1A1 was associated with worse prognosis and clinical outcome. In addition, the multivariate Cox mode analysis indicated that ZEB2 status was an independent factor for both prognosis indexes (OS and DFS).

In general, we demonstrated that ZEB2 and ALDH1A1 may play an important role in lymph node metastasis, tumor-advanced stage, and prognosis. They could work as a promising target for prognostic prediction in breast cancer. Determination of the combined ZEB2 and ALDH1A1 expression may be more helpful to identify high-risk breast cancer patients.

Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (No. 81373427, 81673475, 81572898), Key Laboratory Foundation from Shenyang S&T Projects (F16-094-1-00), and National Science and Technology Major Projects for “Significant New Drugs Development” (2014ZX092010-02-004).

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Minjie Wei, Department of Pharmacology, School of Pharmacy, China Medical University, 77 Puhe Road, Shenyang North New Area, Shenyang 110122, Liaoning Province, P.R. China. Tel: +86-24-31939448; Fax: +86-24-31939448; E-mail: weiminjiecnu@163.com

References

[5] Li XS, Xu Q, Fu XY and Luo WS. ALDH1A1 over-expression is associated with the progression and prognosis in gastric cancer. BMC Cancer 2014; 14: 705.
Expression of ZEB2 and ALDH1A1 in breast cancer patients


