PEBP4 gene expression in lung squamous cell carcinoma: A meta-analysis-based study of the molecular pathways involved

GUIPING YU1*, NING ZHONG2*, BIN HUANG1 and YEDONG MI1

1Department of Cardiothoracic Surgery, The Affiliated Jiangyin Hospital of Southeast University Medical College, Jiangyin, Jiangsu 214400; 2Department of Cardiothoracic Surgery, The First People's Hospital of Kunshan, Kunshan, Jiangsu 215300, P.R. China

Received July 23, 2018; Accepted May 17, 2019

DOI: 10.3892/ol.2020.11386

Abstract. Previous studies have suggested increased activity of phosphatidylethanolamine binding protein 4 (PEBP4) may be associated with the prognosis of non-small cell lung cancer. However, to the best of our knowledge, no direct association between PEBP4 and lung squamous cell carcinoma (LSCC) has been reported. In the present study, a systematic review and meta-analysis was performed to examine the gene expression activity of PEBP4 in LSCC. A total of 10 out of 131 gene expression datasets from the Gene Expression Omnibus (GEO) were selected, including 574 samples (319 patients with LSCC and 255 healthy controls). Subsequently, multiple linear regression (MLR) was employed to study three potential influencing factors: Sample size, population region and study date. A literature-based pathway analysis was then conducted to examine the potential mechanisms through which PEBP4 may exert influence on LSCC. The results of a meta-analysis indicated that, in LSCC, PEBP4 exhibited significantly low expression levels (P<0.033), with mildly increased gene expression levels observed in three studies (log fold-change: 0.072-2.13). However, a significant between-study variance was observed from the heterogeneity analysis. MLR indicated that population region was a significant factor (P<0.0065), whereas sample size and study age were not (P>0.46). Eight functional pathways were subsequently identified, through which PEBP4 may influence the prognosis of LSCC and its response to treatment. The results of the present study suggested that the effects of PEBP4 on LSCC can be neglected in most cases of LSCC, where PEBP4 demonstrated decreased expression levels. However, in the case of PEBP4 overexpression, it may contribute to the progression of LSCC and lead to the development of drug resistance.

Introduction

Lung cancer is a leading cause of cancer-associated mortality worldwide, with 85-90% of cases classified as non-small cell lung cancer (NSCLC) (1). As one type of NSCLC, the lung squamous cell carcinoma (LSCC) subtype accounts for 25-30% of all lung cancer cases (2). Despite advances in the targeted treatment of patients with NSCLC, patients with LSCC do not often benefit from them. For example, the epidermal growth factor receptor (EGFR) has been reported as a treatment target for NSCLC (3); however, LSCC rarely responds to EGFR kinase inhibitors (4). In pure LSCC, EGFR mutations do not occur; however, they do appear in mixed adenosquamous carcinoma (5). Therefore, further studies of the genetic etiology of LSCC are required.

Mutations in phosphatidylethanolamine-binding protein 4 (PEBP4) have frequently been reported in numerous types of cancer (6), and PEBP4 has been suggested as an important treatment target for ovarian (7), prostate (6) and rectal tumors (8). Our previous study observed a possible role for PEBP4 in NSCLC progression through the PI3K/Akt/mTOR signaling pathway (9). However, to the best of our knowledge, no studies have reported a direct association between PEBP4 and LSCC. To address this issue, the present study conducted a systematic review and meta-analysis to examine the gene expression changes of PEBP4 in LSCC. The results were subsequently integrated with a literature-based pathway analysis to examine possible functional pathways through which PEBP4 may exert effects on LSCC. The aim of this study was to gain comprehensive knowledge of the variations in the gene expression levels of PEBP4 in LSCC, and to understand the influence of its expression variance on LSCC using functional pathway analysis.

Materials and methods

Data selection. A systematic search was conducted on expression datasets from the Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/geo). Fig. 1 shows the workflow for...
expression data selection for the meta-analysis. In total, 157 studies were identified based on a keyword search using ‘lung squamous cell carcinoma’. A total of 10 out of these 157 studies satisfied the selection criteria of this study and were included in the meta-analysis, as presented in Table I (10-19). The selection criteria were as follows: i) The data organism was *Homo sapiens*; ii) the data type was RNA expression detected by array; iii) the study design was limited to LSCC vs. healthy cases; and iv) the data included gene expression of PEBP4. For the 10 studies included, there were 574 samples in total, comprising 255 LSCC cases and 319 controls. Despite no date limitation in the systematic review, all data collected were between 1 and 10 years old (2008-2017), as determined using the following formula: Current year-collection date + 1.

The selection of the data covered all LSCC expression array datasets from the GEO, which is owned by the National Institute of Health. The datasets are publicly available, and no permission or confirmation is required for their use. In addition, the dataset extraction had no selection bias in terms of publication journals, owner affiliations and authors. All authors agreed on the data selection criteria to avoid any subjective bias. In addition, the original data were used, rather than the processed results of each dataset, to perform the analysis, in order to avoid any potential noise caused by individual data processing.

**Meta-analysis models.** The fixed-effect and random-effects models (20) were employed to study the effect size of PEBP4 in LSCC. For each expression dataset, the log fold-change (LFC) was calculated for the LSCC samples and used as the index of effect size in the meta-analysis. The expression data were normalized and log2-transformed, if not ready done so in the original dataset. Results from both models were reported and compared. The heterogeneity of the meta-analysis was analyzed to study the variance within and between the different studies, where between-study variance [Tau-squared ($\tau^2$)] was calculated. All analysis was conducted by an individually-developed MATLAB (version R2017a; https://www.mathworks.com/products/matlab.html) meta-analysis package. The additional detailed results of the meta-analysis are available online (http://gousinfo.com/database/Data_Genetic/LSCC_PEBP4.xlss).

**Multiple linear regression (MLR) analysis.** MLR analysis was employed to study the possible influence of three factors on the gene expression alterations in LSCC: Sample size, population region and study date. P-values and 95% confidence interval (CI) were reported for each of the factors. The analysis was done in MATLAB (R 2017a) using the ‘regress’ statistical analysis package.

**Pathway analysis.** Literature-based functional pathway analysis was conducted using Pathway Studio (version 12.1.0.9; www.pathwaystudio.com) to study the potential functional pathways associated with PEBP4 in LSCC. The results were presented as a network graph with the corresponding supporting association list of references.

**Results**

**Meta-analysis results.** The effect sizes and associated statistics from the ten studies, and the meta-analysis results for the PEBP4 gene are presented in Fig. 2 and Table II. The results indicated that the weights from the random-effects model and the fixed-effect model were different for the PEBP4 gene (Table II), which suggested that between-study variance existed and the random-model should be used for the analysis.

Heterogeneity analysis indicated that the between-study variance ($\tau^2$) was calculated as 4.56, indicating a significant between-study variance. The total variance (Q) was 354.43, with an expected variance ‘df’ (under the assumption that all studies have the same effect size) of 9. This resulted in an ISq of 97.46, indicating that >97.46% of variances were due to between-study variance; and $P<1x10^{-20}$ for the hypothesis that Q was from within-study variances only. These results suggested a significant between-study variance of the effect size (LFC). Therefore, a random-effects model was indicated to be more appropriate for this study, which estimates the mean of effect sizes from different studies. The following discussion focused on the results from the random-effects model only.

The LFC from the meta-analysis was -1.80 [95% CI: (-2.80, -0.80); P=0.03; Fig 2]. These results suggested that, on average, PEBP4 presented significantly decreased expression levels in cases of LSCC in the ten studies involved. However, there were significant between-study variances (P<1x10^{-24}; see LSCC_PEBP4+ Ref for pathway analysis), with three studies exhibiting increased expression levels, including LFC=2.13, 0.080 and 0.072, calculated from datasets GSE67061 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67061), GSE84784 (https://www.ncbi.nlm.nih.gov/geo/query/acc.
variability in results requires an analysis to study the influence of potential factors affecting the results. For more results, please refer to LSCC_PEBP4→sumResults.

MLR analysis. Results from the MLR models indicated that population region was a significant influencing factor for the expression fold-change of PEBP4 (P=0.0064), as presented in Table III. Conversely, the sample size and study date indicated no significant influence (P>0.35).

Pathway analysis results. Pathway analysis using Pathway Studio (www.pathwaystudio.com) was conducted to identify possible pathways through which PEBP4 may exert an effect on LSCC. As shown in Fig. 3, there were various pathways that linked PEBP4 with LSCC through different entities, including proteinsgenes, cell processes and functional class. Five potential genetic pathways were identified, as indicated in Fig. 3. Among these five pathways, three suggested an LSCC-promoting effect of PEBP4, including PEBP4→BCL2L1→LSCC, PEBP4→AKT1→LSCC and PEBP4→prostaglandin-endoperoxide synthase 2 (PTGS2)→LSCC. In addition, two genetic pathways indicated a treatment-resistant effect of PEBP4, including PEBP4→MAPK3→LSCC and PEBP4→MAPK1→LSCC. Three cellular processing pathways were also identified, which were associated with cell proliferation, apoptosis and epithelial-mesenchymal transition (Fig. 3). Each relationship (edge) in Fig. 3 had ≥1 supporting reference, assisting in understanding the potential mechanism.

Table I. The 10 studies used in the present meta-analysis.

<table>
<thead>
<tr>
<th>Study name</th>
<th>Dataset GEO ID</th>
<th>Control (n)</th>
<th>Case (n)</th>
<th>Country</th>
<th>Data type</th>
<th>(Refs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nazarov et al</td>
<td>GSE84784</td>
<td>9</td>
<td>9</td>
<td>Luxembourg</td>
<td>Expression by array</td>
<td>(10)</td>
</tr>
<tr>
<td>Tong et al</td>
<td>GSE67061</td>
<td>8</td>
<td>69</td>
<td>China</td>
<td>Expression by array</td>
<td>-</td>
</tr>
<tr>
<td>Mascaux et al</td>
<td>GSE33479</td>
<td>27</td>
<td>14</td>
<td>USA</td>
<td>Expression by array</td>
<td>-</td>
</tr>
<tr>
<td>Rousseaux et al</td>
<td>GSE30219</td>
<td>14</td>
<td>61</td>
<td>France</td>
<td>Expression by array</td>
<td>(11)</td>
</tr>
<tr>
<td>Girard et al</td>
<td>GSE32036</td>
<td>59</td>
<td>12</td>
<td>USA</td>
<td>Expression by array</td>
<td>(12,13)</td>
</tr>
<tr>
<td>Kuner et al</td>
<td>GSE27489</td>
<td>10</td>
<td>10</td>
<td>Germany</td>
<td>Expression by array</td>
<td>(14)</td>
</tr>
<tr>
<td>Philipsen et al</td>
<td>GSE19188</td>
<td>65</td>
<td>27</td>
<td>Netherlands</td>
<td>Expression by array</td>
<td>(15)</td>
</tr>
<tr>
<td>Ishikawa et al</td>
<td>GSE2088</td>
<td>30</td>
<td>48</td>
<td>Japan</td>
<td>Expression by array</td>
<td>(16)</td>
</tr>
<tr>
<td>Takahashi et al</td>
<td>GSE11969</td>
<td>5</td>
<td>35</td>
<td>Japan</td>
<td>Expression by array</td>
<td>(17,18)</td>
</tr>
<tr>
<td>Boelens et al</td>
<td>GSE12428</td>
<td>28</td>
<td>34</td>
<td>Netherlands</td>
<td>Expression by array</td>
<td>(19)</td>
</tr>
</tbody>
</table>

GEO; Gene Expression Omnibus.

Figure 2. Meta-analysis results using a random-effects model for the PEBP4 gene. (A) Mean of effect size (red diamonds) and 95% CI (blue lines) of each dataset. (B) Bar plot of weights for each dataset. The brighter the color (i.e. green), the bigger the weight. CI, confidence interval; PEBP4, phosphatidylethanolamine binding protein 4.
underlying the effects of PEBP4 on the pathogenesis of LSCC. For example, it has been reported that increased expression levels of PEBP4 can inhibit the activity of MAPK3 (7), which is already recognized as a therapeutic target for LSCC (6); notably, a PEBP4 → MAPK3 → LSCC pathway was present in Fig. 3. These findings suggested an anti-drug-effect pathway of PEBP4 in the treatment of LSCC. In total, there were eight pathways composed of 16 relationships (edges), and these relationships were supported by 163 references. The full list of these relations and the corresponding supporting references are presented in LSCC_PEBP4→Ref for pathway analysis, which is available online (http://gousinfo.com/database/Data_Genetic/LSCC_PEBP4.xlsx).

Discussion

During the past few years, PEBP4 has been identified as a contributor to the development of numerous types of cancer (6-9,20,21). Our previous study indicated that overexpression of PEBP4 enhances NSCLC cell proliferation and the invasive ability of cancer cells, while inhibiting apoptosis (9). Another study indicated that PEBP4 promotes the epithelial-to-mesenchymal transition by activating the sonic hedgehog signaling pathway in NSCLC (22).

Table II. The effects of the two models for phosphatidylethanolamine binding protein 4.

<table>
<thead>
<tr>
<th>Study name</th>
<th>Effect size</th>
<th>Lower limit of 95% CI</th>
<th>Upper limit of 95% CI</th>
<th>Z-value</th>
<th>P-value</th>
<th>Weight of fixed-effect model</th>
<th>Weight of random-effects model</th>
<th>(Refs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nazarov et al, 2017</td>
<td>0.08</td>
<td>-0.77</td>
<td>0.93</td>
<td>0.18</td>
<td>0.43</td>
<td>5.26</td>
<td>0.21</td>
<td>(10)</td>
</tr>
<tr>
<td>Tong et al, 2016</td>
<td>2.13</td>
<td>-1.58</td>
<td>5.84</td>
<td>1.12</td>
<td>0.13</td>
<td>0.28</td>
<td>0.12</td>
<td>-</td>
</tr>
<tr>
<td>Mascaux et al, 2014</td>
<td>-4.63</td>
<td>-5.06</td>
<td>-4.20</td>
<td>-21.00</td>
<td>&lt;0.001</td>
<td>21.10</td>
<td>0.22</td>
<td>-</td>
</tr>
<tr>
<td>Rousseaux et al, 2014</td>
<td>-3.88</td>
<td>-7.15</td>
<td>-0.60</td>
<td>-2.30</td>
<td>0.01</td>
<td>0.36</td>
<td>0.14</td>
<td>(11)</td>
</tr>
<tr>
<td>Girard et al, 2012</td>
<td>-0.07</td>
<td>-0.66</td>
<td>0.51</td>
<td>-0.30</td>
<td>0.40</td>
<td>11.20</td>
<td>0.22</td>
<td>(12,13)</td>
</tr>
<tr>
<td>Kuner et al, 2011</td>
<td>-0.25</td>
<td>-2.78</td>
<td>2.27</td>
<td>-0.20</td>
<td>0.42</td>
<td>0.60</td>
<td>0.16</td>
<td>(14)</td>
</tr>
<tr>
<td>Philipse et al, 2010</td>
<td>-3.76</td>
<td>-6.54</td>
<td>-1.00</td>
<td>-2.70</td>
<td>&lt;0.001</td>
<td>0.50</td>
<td>0.15</td>
<td>(15)</td>
</tr>
<tr>
<td>Ishikawa et al, 2009</td>
<td>-2.20</td>
<td>-3.90</td>
<td>-0.50</td>
<td>-2.50</td>
<td>0.01</td>
<td>1.32</td>
<td>0.19</td>
<td>(16)</td>
</tr>
<tr>
<td>Takahashi et al, 2009</td>
<td>-0.97</td>
<td>-1.41</td>
<td>-0.50</td>
<td>-4.20</td>
<td>&lt;0.001</td>
<td>19.20</td>
<td>0.22</td>
<td>(17,18)</td>
</tr>
<tr>
<td>Boelens et al, 2008</td>
<td>0.072</td>
<td>-0.22</td>
<td>0.37</td>
<td>0.48</td>
<td>0.32</td>
<td>43.90</td>
<td>0.22</td>
<td>(19)</td>
</tr>
<tr>
<td>Fixed model</td>
<td>-1.15</td>
<td>-1.34</td>
<td>-1.00</td>
<td>-12.00</td>
<td>&lt;0.001</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Random-effects model</td>
<td>-1.36</td>
<td>-2.80</td>
<td>0.09</td>
<td>-1.80</td>
<td>0.03</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

CI, confidence interval.

Table III. Multiple linear regression analysis results.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sample size</th>
<th>Population region</th>
<th>Study date</th>
<th>Beta 0.002</th>
<th>LowLimit -0.600</th>
<th>UpLimit 0.063</th>
<th>P-value 0.46</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>0.0020</td>
<td>0.88</td>
<td>0.066</td>
<td>0.0020</td>
<td>0.020</td>
<td>0.57</td>
<td>0.46</td>
</tr>
</tbody>
</table>

To the best of our knowledge, no study has reported a direct association between LSCC and PEBP4. In the present study, a meta-analysis and literature-based functional pathway analysis was conducted, to examine the possible influence of PEBP4 on LSCC and the potential underlying mechanisms.

The present meta-analysis used gene expression datasets available in the GEO (www.ncbi.nlm.nih.gov/geo), which passed data selection criteria. A total of 10 human gene expression datasets were included with a design of LSCC vs. control. These datasets had a publication-age gap of ≤10 years (2008-2017), and were from different regions, including France, USA, Netherlands, Japan, Germany, Luxembourg and China.

Meta-analysis indicated that there was a decreased activity of PEBP4 in terms of LFC, with a significant between-study variance (LSCC_PEBP4→Reference Table). One possible reason may lie in the data generation of different studies, as they used different platforms, including GPL570 for GSE30219 and GPL6884 for GSE32036, and sample sources, including lung samples in GSE67061 and bronchial biopsy samples in GSE33479 (LSCC_PEBP4→Data). Other explanations could be the existence of multiple influential factors that may lead to increased or decreased expression of PEBP4 under different circumstances. MLR analysis confirmed this and indicated that the sample population region (country) was a significant influential factor. MLR results also indicated that the sample size and study ages had no significant influence on the expression variance of PEBP4 in the patients in the 10 LSCC groups included in this study; however, the sample sizes varied from 9 to 69, and the publication years were between 2008 and 2017. This suggested that, under similar circumstances (e.g., same population region), the expression of PEBP4 may not be significantly changed in LSCC cases from the past 10 years.
Despite between-study differences in terms of PEBP4 expression levels, the present meta-analysis integrated information from independent but related studies by using a random-effects model, thus assisting in improving the reliability of the results. The meta-analysis results indicated that increased and decreased activity of PEBP4 may occur in LSCC. Therefore, it is necessary to examine the possible consequences in the case of varied PEBP4 activity. To address this issue, a pathway analysis was conducted using the shortest path functionality of Pathway Studio (www.pathwaystudio.com). This function identifies directed pathways through which PEBP4 may exert functional influence on LSCC. Each relationship (edge) had one or more supporting references (LSCC_PEBP4→Ref for pathway analysis).

Multiple potential pathways were identified, through which PEBP4 may promote the pathological development of LSCC. Notably, overexpression of PEBP4 promotes the activity of PTGS2, AKT1 and BCL2L1 (23,24), and these three genes serve an important role in the development of LSCC (25,26). These pathways indicated that the overexpression of PEBP4 may promote the development of LSCC. In addition, PEBP4 inhibits MAPK3 and MAPK1 (23), whereas the activation of these two genes has been suggested as therapeutic targets for the treatment of LSCC (20).

At a cellular processing level, three potential LSCC-promoting pathways were identified. Firstly, the prognosis of LSCC is associated with LSCC cancer cell proliferation, apoptosis and epithelial-to-mesenchymal transition (27). Secondly, it has been reported that PEBP4 may enhance NSCLC cancer cell proliferation (28), inhibit apoptosis of numerous cancer cells (8), and promote the epithelial-to-mesenchymal transition in multiple cancer cells (6,28). These pathways may assist in understanding the potential associations between PEBP4 and LSCC.

There are numerous limitations of this study that require further investigation. Firstly, due to limited meta-data, only three potential factors were studied. Considering the significant between-study variance of PEBP4 expression levels in LSCC, more influential factors are expected and should be examined, including age, sex and the presence of co-morbidities. Secondly, the functional pathways analysis was literature-based. Although supported by previous studies, specific experiments should be conducted to test these pathways. Finally, in the meta-analysis, only expression by array datasets from the GEO were used. In the GEO database, besides expression by array, there are multiple other types of expression data, including expression profiling by single nucleotide polymorphism array and by high throughput sequencing. However, to avoid noise brought by the use of different data types, this study only used expression profiling by array data. Meta-analysis using other types of data could be conducted in the future to confirm the results of this study.

In the present study, PEBP4 demonstrated overall decreased expression levels in the meta-analysis, which was in accordance with most of the studies (7 out of 10) employed. Integrating results from the pathway analysis, this study indicated that in the majority of LSCC cases, the influence of PEBP4 on LSCC can be neglected. However, the expression levels of PEBP4 demonstrated strong heterogeneity, due to numerous influential factors, including population and region, with possible increased PEBP4 activity occurring in patients with LSCC. In such cases, the potential influence of PEBP4 on LSCC should be considered as it may promote the development of LSCC and lead to drug resistance.

Acknowledgements

Not applicable.
Funding
This study was partly supported by the 2017 Fifth Provincial ‘333 Project’ Research Project; by non-small cell lung cancer research funding, Jiangsu Provincial Planning Commission (grant no. H201657); by clinical diagnosis and treatment of small lung lesions and normative research, Wuxi City Health Planning Commission (grant no. MS201625); and by Thirteen Five Science and Education in Jiangsu Province.

Availability of data and materials
All data analyzed during this study is included in this published article. The results of the meta-analysis are available online at http://gousinfo.com/database/Data_Genetic/LSCC_PEBP4.xlsx. The GEO datasets are stated in LSCC_PEBP4-GEO datasets, including the study name, GEO ID and URL for the datasets.

Authors' contributions
GY and NZ contributed to the data collection, analysis and manuscript writing. BH and YM contributed to the data collection and analysis. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References


