Computer Methods and Programs in Biomedicine 155 (2018) 179–187

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Immunohistochemical quantification of expression of a tight junction protein, claudin-7, in human lung cancer samples using digital image analysis method

Zhe Lu[a,1], Yi Liu[a,1], Junfeng Xu[b], Hongping Yin[a], Haiying Yuan[c], Jinjing Gu[a], Yan-hua Chen[d], Liyun Shi[e], Dan Chen[a], Bin Xie[b,*]

[a] School of Medicine, Hangzhou Normal University, Hangzhou, Zhejiang Province 311121, China
[b] School of Science, Hangzhou Normal University, Hangzhou, Zhejiang Province 311121, China
[c] Department of Clinical Laboratory, Women's Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province 310000, China
[d] Brody School of Medicine, East Carolina University, Greenville, NC 27834, USA
[e] School of Basic Medical Science, Nanjing University of Chinese Medicine, Nanjing, Jiangsu Province 210023, China

ARTICLE INFO

Article history:
Received 31 May 2017
Revised 11 November 2017
Accepted 11 December 2017

Keywords:
Lung cancer
Claudin-7
Image analysis
Intensity
Fragmentation
Immunohistochemistry

ABSTRACT

Background and objectives: Tight junction proteins are correlated with cancer development. As the pivotal proteins in epithelial cells, altered expression and distribution of different claudins have been reported in a wide variety of human malignancies. We have previously reported that claudin-7 was strongly expressed in benign bronchial epithelial cells at the cell-cell junction while expression of claudin-7 was either altered with discontinued weak expression or completely absent in lung cancers. Based on these results, we continued working on the expression pattern of claudin-7 and its relationship with lung cancer development. We herein proposed a new Digital Image Classification, Fragmentation index, Morphological analysis (DICFM) method for differentiating the normal lung tissues and lung cancer tissues based on the claudin-7 immunohistochemical staining.

Methods: Seventy-seven lung cancer samples were obtained from the Second Affiliated Hospital of Zhejiang University and claudin-7 immunohistochemical staining was performed. Based on C++ and Open Source Computer Vision Library (OpenCV, version 2.4.4), the DICFM processing module was developed. Intensity and fragmentation of claudin-7 expression, as well as the morphological parameters of nuclei were calculated. Evaluation of results was performed using Receiver Operator Characteristic (ROC) analysis.

Results: Agreement between these computational results and the results obtained by two pathologists was demonstrated. The intensity of claudin-7 expression was significantly decreased while the fragmentation was significantly increased in the lung cancer tissues compared to the normal lung tissues and the intensity was strongly positively associated with the differentiation of lung cancer cells. Moreover, the perimeters of the nuclei of lung cancer cells were significantly greater than that of the normal lung cells, while the parameters of area and circularity revealed no statistical significance.

Conclusions: Taken together, our DICFM approach may be applied as an appropriate approach to quantify the immunohistochemical staining of claudin-7 on the cell membrane and claudin-7 may serve as a marker for identification of lung cancer.

1. Introduction

Lung cancer is the leading cause of cancer death among both men and women. Adenocarcinoma (ADC) is the most common type of carcinoma followed by squamous cell carcinoma (SCC) and small cell lung carcinoma (SCLC) [1]. Tumor evaluation is historically critical in the pathological diagnosis of lung cancer as the treatment for each subtypes is very different. For example, thyroid transcrip-

* Corresponding author at: School of Science, Hangzhou Normal University, 2318th Yuhangtang Road, Yuhang District, Hangzhou, Zhejiang Province 311121, China.
E-mail address: xiebin@hznu.edu.cn (B. Xie).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.cmpb.2017.12.014

tion factor-1 (TTF1) is the most reliable marker for adenocarcinoma and p63 is a reliable squamous marker [2]. More and more new markers for tumor evaluation and classification are emerging, such as excision repair cross-complementation group 1 (ERCC1) [3], regulatory subunit of ribonucleotide reductase (RRM1) [4], and breast cancer 1 (BRCA1) [5].

Lungs are frequently exposed to pathogens and exogenous compounds, which can influence the normal function of bronchial airways and alveolar cells. In human airways, tight junctions (TJs) serve to separate the internal and external environment of the lungs, thus provide a line of defense against inhaled harmful substances. Claudins are one of the three TJs families. Claudin-1, -2, -3, -4, -5, and -7 are detected in normal human airways epithelial cells. Recent studies have shown that claudins are closely associated with lung tumorigenesis. Altered expression and distribution of different claudins have been reported in various human lung cancer tissues [6], indicating that they could be a useful prognostic predictor and potential drug treatment target for patients with lung cancer. Our previous immunohistochemical (IHC) staining showed that claudin-7 was strongly expressed in benign bronchial epithelial cells with a predominant cell-cell junction staining pattern while expression of claudin-7 was either altered with discontinued weak expression or completely absent in lung cancers [7]. However, for IHC, the interpretation relies solely on subjective visual estimation and yields only qualitative results even though a standardized protocol is set. Individual observers, experienced or naive, reveals different results. Even among experienced observers, variation occurs [8,9]. Thus, a methodology for quantitative analysis is demanded. Indeed, computational methods, such as SACAIA (Semiautomated Computer-Assisted Image Analysis) [10], D-HSCORE (Digital histological score) [8], DISBE (Digital Image Subtraction, Blue filter, Enhancement) [9], MIA (Membrane Isolation Algorithm) [11] are recently emerging. These methods provide more information from the immunostained specimens and lead to a more accurate elucidation of data obtained from different research groups.

In this study, we further want to investigate whether the expression pattern of claudin-7 can be automatically extracted and analyzed and whether it is correlated with different subtypes of lung cancers or differentiation of lung cancer cells. So that, claudin-7 may potentially be used as a marker for identifying the different subtypes of lung cancers or as a marker for prognosis. For these purposes, we developed an image analysis program to automatically analyze the expression level and pattern of claudin-7 in lung cancer cells, which were then applied to analyze the relationship between claudin-7 and lung cancers.

2. Materials and methods

2.1. Immunohistochemical staining

Seventy-seven lung cancer tissues were obtained from the Second Affiliated Hospital of Zhejiang University, 34 of them were squamous cell carcinoma and 33 of them were adenocarcinoma. Normal lung tissues were used as controls. All the tissues were from surgical resection. Immunohistochemical staining procedure was performed as previously described [7]. All procedures were conducted in accordance with university guidelines and approved by the ethical committee of Hangzhou Normal University.

2.2. Image acquisition

3,3′-diaminobenzidine (DAB)-labeled immunostain sections were imaged using brightfield optics on a Nikon eclipse 90i microscope (field of view: 22 mm). Individual images were acquired using a Nikon DS-Ri1 camera. NIS-Elements BR 3.2 was used to capture the images. Images were samples randomly throughout histological sections, but areas that contained preparation artifacts, cell debris, nonspecific staining, or edges were avoided. Acquired images (24-bit RGB, 8 bits/color) from tissue sections had a resolution of 3840×3072 pixels and were stored in JPEG format.

2.3. DICFM approach

2.3.1. Overview of computer-assisted image analysis scheme

A digital image processing approach was developed and called Digital Image Classification, Fragmentation index, Morphological analysis (DICFM). The workflow of the image analysis scheme is outlined in Fig. 1. Firstly, claudin-7 expression was separated by utilizing a multi-center minimum distance classification algorithm. Subsequently, intensity and fragmentation of claudin-7 expression were obtained through correcting image background and identifying contours. And then, nucleus were extracted to calculate morphology parameters including area, perimeter and circularity. Lastly, based on these data, we provided a gross evaluation of the relationship between claudin-7 expression and nucleus morphology and lung cancer progression. Meantime, each image was also analyzed by manual quantification exerted by two pathologists.

2.3.2. Intensity measurement of claudin-7 expression

2.3.2.1. Classification. In order to measure intensity, researchers have used image classification methods such as minimum distance [12], K-means [13], Bayes [14], decision tree [15], neural network [16], random forest [17] and support vector machines [18] to identify target object firstly. In this study, supervised classification of claudin-7 expression was adopted based on multi-center minimum distance classifier. First of all, multiple centers (RGB vector spaces) of claudin-7 expression training samples were established, and then the Euclidean distance of RGB vector between the input pixel and target class was calculated.

Define class i has center Cik, $k = 1,2,\ldots,n$. The minimum distance between pixel P and class i is:

$$d(P,i) = min[D(P,Cik),\ \ k = 1,2,\ldots,n] \tag{1}$$

Where Euclidean distance $D(P,Cik)$ was calculated as follows:

$$D(P,Cik) = \sqrt{(R(P)-R(Cik))^2 + (G(P)-G(Cik))^2 + (B(P)-B(Cik))^2} \tag{2}$$

Wherein, $R(P)$, $G(P)$, $B(P)$ are the RGB vectors of input pixel P, and $R(Cik)$, $G(Cik)$, $B(Cik)$ are the RGB vectors of center Cik. When $d(P,i)$ is less than the given distance threshold, then P belongs to class i. The classification results were post-processed by category merging and small patch removal, and the confusion matrix and Kappa coefficient [12] were calculated to evaluate the accuracy of the classification.

2.3.2.2. Background correction. Different environmental conditions or the photographers lead into different color tones between DAB digital images inevitably. In order to eliminate the influence, researchers have tried different methods to correct the background values of the images [8,9,19]. On the basis of Bernardo's method [9], we calculated the average value of each RGB channel of the background of the original image, and used this value to subtract each corresponding RGB channel of classification image of the claudin-7 expression, and then got the background correction image.

2.3.2.3. Intensity score. Currently, researchers have developed a lot of immunostaining intensity score methods [8,11]. On the basis of Zizzi's method of mean gray level value [20], we calculated