Utilizing gene expression profiles to characterize tumor infiltrating lymphocytes in cancers

Ching-Hsuan Chen, Tzu-Pin Lu

Institute of Epidemiology and Preventive Medicine, Department of Public Health, National Taiwan University, Taipei

Correspondence to: Tzu-Pin Lu. Institute of Epidemiology and Preventive Medicine, Department of Public Health, National Taiwan University, Taipei. Email: tplu@ntu.edu.tw.

With the advancement in immunotherapy and high-throughput technologies, tumor infiltrating lymphocytes (TILs) have become an important research topic in several diseases, especially different cancers. TILs are lymphoid cells, including T-cells and B-cells, that can infiltrate solid tumor cells and trigger immune responses from the individual. Unsurprisingly, the prognostic effects that TILs showed in cancer patients can be traced back to the 1980s (1). At that time, identification of TILs relied on examining the pathology slides of tumor tissues resected from patients; this is still a standard approach to analyze TILs in clinics. However, the labor-intensive nature and associated high costs of this approach make it impractical to use on a large-scale investigation with a large number of patients.

In recent decades, high-throughput genomic technologies, such as microarray and next-generation sequencing (NGS), allow researchers to simultaneously examine the expression level of the whole transcriptome in one patient within a short timeframe and at a relatively low cost. Consequently, many algorithms and statistical models have been developed to use the expression profiles of selected genes to calculate the expression levels of TILs (2-4). A detailed discussion and review of different algorithms was described previously (5). Since most transcription profiles in the public domain were still from microarray, two popular methods based on microarray, including CIBERSORT (2) and microenvironment cell populations-counter (MCP-counter) (3), are described in this editorial.

The CIBERSORT algorithm was published in Nature Methods in 2015 and has been widely used to analyze the expression levels of different kinds of TILs in solid tumor tissues. In their original paper on CIBERSORT, the authors proposed a gene signature matrix named LM22 that contains 547 genes. LM22 can be used to distinguish 22 different hematopoietic cell phenotypes, including T-cells and B-cells, among others. The CIBERSORT algorithm uses a support vector regression (SVR) method to deconvolve the expression profiles of the 547 selected genes into the intensity levels of 22 different TILs. More than 3,000 human tissue specimens and cell lines were validated in the original paper. Sensitivity and specificity values were both greater than 0.94, and the area under the receiver operator characteristic (ROC) curve (AUC) to identify a sample with at least one TIL out of the 22 TILs was greater than 0.98. For one specific TIL, the authors performed flow cytometry to evaluate the performance of the CIBERSORT algorithm in 10 TIL subsets by using 27 human adult samples. Most of the TIL subsets were successfully predicted, which demonstrate the algorithm's potential to be used in large-scale investigations.

MCP-counter is another popular algorithm to analyze gene expression profiles to estimate the expression levels of multiple TILs. In contrast to the SVR method in the CIBERSORT algorithm, the MCP-counter algorithm considers the variations of the expression level of one gene in a specific TIL type and retains the genes showing the lowest variation within the TIL type. Consequently, the MCP-counter algorithm identifies

genes for distinguishing 10 different TILs within one sample. Notably, the most important difference between CIBERSORT and MCP-counter is that CIBERSORT investigates the expression levels of TILs within one sample of immune cell proportions, i.e., intra-sample, whereas MCP-counter reports the abundance of TILs for inter-sample comparisons, at the cost of presenting the TIL in arbitrary units. Similar to the CIBERSORT algorithm, the MCP-counter algorithm's performance is validated by using the polymerase chain reaction (PCR) experiments of prepared RNA mixtures from the peripheral blood specimens. Significant associations were reported between the predicted values and the prepared concentrations; thus, the MCP-counter has also been widely utilized in recent studies.

Some limitations exist in using these two algorithms to analyze gene expression profiles directly to obtain the expression levels of TILs. The rapid growth of NGS technology has facilitated the accumulation of significant amounts of RNA-seq data. However, the two algorithms were both developed from microarray datasets. The dynamic ranges of the gene expression profiles obtained from the microarray datasets are not the same as what we can obtain from the real RNA-seq data. Therefore, severe systematic biases will occur if we apply these two algorithms to directly RNA-seq data. To address this issue, a normalization method, such as spline regression, must be performed before using either of these algorithms to directly analyze RNA-seq data.

Furthermore, it is well-known that different gene expression microarray platforms have different probe designs. Some genes that are required for analysis by the CIBERSORT and the MCP-counter algorithms are not included in the corresponding microarray design. Therefore, the estimated expression levels of TILs will contain some biases due to the microarray design, if multiple datasets are analyzed simultaneously. Lastly, since selected genes/biomarkers are used to estimate the expression levels of TILs, the collinearity among these selected genes must be taken into consideration. Ignoring the collinearity may result in overestimation of the prognostic effects of multiple TILs with similar expression patterns. In conclusion, analyzing multiple datasets, especially from different microarray designs, at the same time to do inter-sample comparisons is not recommended.

In the last portion of this editorial, we use real studies to demonstrate how TILs play important roles in different cancers, including breast cancer, colorectal cancer, head and neck cancer, and hepatocellular carcinoma. Previous studies have shown that the prognosis of breast cancer patients is determined by age, stage, tumor grading, and molecular subtypes. In addition to these factors, TILs have been reported to be associated with a good prognosis in human epidermal growth factor receptor 2 (HER2) enriched and triple negative subtypes, while inconsistent associations between TILs and prognosis were observed in luminal patients (6-10). For example, in the ShortHER trial, a 9-week trastuzumab adjuvant therapy was linked to increased relapse risk [hazard ratio (HR): 1.75, 95% confidence interval (CI): 1.09–2.80] in HER2+ patients with low TILs but not in patients with high TILs (HR: 0.23, 95% CI: 0.05–1.09) (11). In the NSABP B-31 trial, stromal TILs served as an independent prognostic factor but showed no statistical interaction with the addition of trastuzumab to chemotherapy in HER2+ breast cancer patients (12). One possible explanation of this variation was that different TIL compositions and immune cell responses exist in the patients. Notably, B-cells and CD4+ T-cells can facilitate CD8+ T-cell cytotoxic effect and were associated with reduced risk of death in cancer patients (6-9). In addition to lymphocytes, myeloid lineage cells participate in immune regulation, and the polarized M2 macrophage was associated with poor prognosis in early-stage breast cancer patients (6,7). Notably, since most of the current studies were based on targeted pathology staining, a comprehensive analysis of TILs is warranted to initiate systematic and large-scale investigations.

TILs can serve as prognostic factors in colorectal cancer and their expression levels were correlated to the microsatellite instability (MSI). The expressions of CD3+, CD8+, FOXP3+ and CD45RO+ lymphocytes were associated with favorable prognosis in colorectal cancer patients (13,14). MSI-high colorectal cancer exhibited higher TILs and M1 macrophages, which might partially explain why MSI was identified as an important prognostic factor (14,15). TILs also affect systemic therapy in colorectal cancer and vice versa. Patients with TILs benefit more from 5-FU adjuvant therapy then those without TILs (HR: 0.22 vs. 0.84) (16). A cetuximab treatment clinical trial in 46 colorectal cancer patients showed higher expression levels of CD4+ T-cells, CD8+ T-cells, type 1 helper T-cells, regulatory T-cells and myeloid-derived suppressor cells; this sheds light on the possibility of utilizing TILs as an indicator for further treatments (17).

Human papillomavirus (HPV) is an emerging topic in head and neck cancer. Positive HPV status is associated

with increased expression levels of CD8+ T-cells, B cells and a better prognosis in head and neck cancer (18,19). Intriguingly, a linear model showed the collinearity between HPV status and CD8+ T-cells in head and neck cancer patients, suggesting the immune cells can account for the prognostic findings in patients with these cancer types (19). Among operable tongue cancer patients with all etiology, Chen *et al.* reported three TILs, high CD4/FOXP3 ratio, high CD8/FOXP3 ratio and low FOXP3, were associated with better clinical outcomes (20). In contrast, Badoual *et al.* reported high expressions of CD3 and CD8 levels were associated with longer overall survival and progression-free survival in head and neck cancer patients with definite chemoradiotherapy, but high expressions of CD4 or FOXP3 levels were not (21). Therefore, these results imply that replication and validation studies are prerequisite before applying the findings of TILs in real clinics.

Hepatitis B virus (HBV) and hepatitis C virus (HCV) are major causes and subtypes of hepatocellular carcinoma. In contrast to HPV in head and neck cancer, patients with positive HBV or HCV infections were reported to have lower TILs in liver cancer tissues than in the adjacent normal liver tissues (22,23). In several studies, among the immune cell subsets, CD8+ cells were consistently lower in the tumor tissue, but their expression was still reported to be a biomarker for indicating a better regional disease control in liver cancer. The depletion of CD8+ cells was linked to a down-regulation of CD28 and an up-regulation of programmed cell death 1 receptor (PD-1), which is the most popular target of immunotherapy nowadays (24,25). In conclusion, TILs may serve as not only the prognostic biomarkers for multiple cancers but also as predictive biomarkers for advanced treatments, including chemoradiotherapy and immunotherapy in the future.

Acknowledgments

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Hiratsuka H, Imamura M, Ishii Y, et al. Immunohistologic detection of lymphocyte subpopulations infiltrating in human oral cancer with special reference to its clinical significance. Cancer 1984;53:2456-66.
2. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12:453-7.
3. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016;17:218.
4. Kang HW, Kim WJ, Yun SJ. The role of the tumor microenvironment in bladder cancer development and progression. Transl Cancer Res 2017;6:S744-58.
5. Finotello F, Trajanoski Z. Quantifying tumor-infiltrating immune cells from transcriptomics data. Cancer Immunol Immunother 2018;67:1031-40.
6. Chen CH, Lu YS, Cheng AL, et al. Disparity in tumor immune microenvironment of breast cancer and prognostic impact: Asian versus Western populations. Oncologist 2019. [Epub ahead of print].
7. Ali HR, Chlon L, Pharoah PD, et al. Patterns of immune infiltration in breast cancer and their clinical implications: a gene-expression-based retrospective study. PLoS Med 2016;13:e1002194.
8. Ali HR, Provenzano E, Dawson SJ, et al. Association between CD8+ T-cell infiltration and breast cancer survival in 12,439 patients. Ann Oncol 2014;25:1536-43.
9. Heimes AS, Madjar K, Edlund K, et al. Subtype-specific prognostic impact of different immune signatures in node-negative breast cancer. Breast Cancer Res Treat 2017;165:293-300.
10. Denkert C, von Minckwitz G, Darb-Esfahani S, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. Lancet Oncol 2018;19:40-50.
11. Dieci MV, Conte P, Bisagni G, et al. Association of tumor-infiltrating lymphocytes with distant disease-free survival in the ShortHER randomized adjuvant trial for patients with early HER2+ breast cancer. Ann Oncol

2019;30:418-23.

12. Kim RS, Song N, Gavin PG, et al. Stromal tumor-infiltrating lymphocytes in NRG oncology/NSABP B-31 adjuvant trial for early-stage HER2-positive breast cancer. J Natl Cancer Inst 2019. [Epub ahead of print].

13. Kong JC, Guerra GR, Pham T, et al. Prognostic impact of tumor-infiltrating lymphocytes in primary and metastatic colorectal cancer: a systematic review and meta-analysis. Dis Colon Rectum 2019;62:498-508.

14. Pagès F, Mlecnik B, Marliot F, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. Lancet 2018;391:2128-39.

15. Narayanan S, Kawaguchi T, Peng X, et al. Tumor infiltrating lymphocytes and macrophages improve survival in microsatellite unstable colorectal cancer. Sci Rep 2019;9:13455.

16. Morris M, Platell C, Iacopetta B. Tumor-infiltrating lymphocytes and perforation in colon cancer predict positive response to 5-fluorouracil chemotherapy. Clin Cancer Res 2008;14:1413-7.

17. Woolston A, Khan K, Spain G, et al. Genomic and transcriptomic determinants of therapy resistance and immune landscape evolution during anti-EGFR treatment in colorectal cancer. Cancer Cell 2019;36:35-50.e9.

18. Sivars L, Landin D, Grün N, et al. Validation of Human Papillomavirus as a Favourable Prognostic Marker and Analysis of CD8+ Tumour-infiltrating Lymphocytes and Other Biomarkers in Cancer of Unknown Primary in the Head and Neck Region. Anticancer Res 2017;37:665-73.

19. Carrero I, Liu HC, Sikora AG, et al. Histoepigenetic analysis of HPV- and tobacco-associated head and neck cancer identifies both subtype-specific and common therapeutic targets despite divergent microenvironments. Oncogene 2019;38:3551-68.

20. Chen WY, Wu CT, Wang CW, et al. Prognostic significance of tumor-infiltrating lymphocytes in patients with operable tongue cancer. Radiat Oncol 2018;13:157.

21. Badoual C, Hans S, Rodriguez J, et al. Prognostic value of tumor-infiltrating CD4+ T-cell subpopulations in head and neck cancers. Clin Cancer Res 2006;12:465-72.

22. Hsiao YW, Chiu LT, Chen CH, et al. Tumor-infiltrating leukocyte composition and prognostic power in hepatitis B- and hepatitis C-related hepatocellular carcinomas. Genes (Basel) 2019. doi: 10.3390/genes10080630.

23. Wang Q, Luan W, Warren L, et al. Prognostic role of immune cells in hepatitis B-associated hepatocellular carcinoma following surgical resection depends on their localization and tumor size. J Immunother 2016;39:36-44.

24. Maki A, Matsuda M, Asakawa M, et al. Decreased expression of CD28 coincides with the down-modulation of CD3zeta and augmentation of caspase-3 activity in T cells from hepatocellular carcinoma-bearing patients and hepatitis C virus-infected patients. J Gastroenterol Hepatol 2004;19:1348-56.

25. Hsu PN, Yang TC, Kao JT, et al. Increased PD-1 and decreased CD28 expression in chronic hepatitis B patients with advanced hepatocellular carcinoma. Liver Int 2010;30:1379-86.