Tumor Biology and Immunology

Cancer Research

TIP: A Web Server for Resolving Tumor Immunophenotype Profiling

Liwen Xu[1], Chunyu Deng[1], Bo Pang[1], Xinxin Zhang[1], Wei Liu[1], Gaoming Liao[1], Huating Yuan[1], Peng Cheng[1], Feng Li[1], Zhilin Long[1], Min Yan[1], Tingting Zhao[2], Yun Xiao[1], and Xia Li[1]

Abstract

Systematically tracking the tumor immunophenotype is required to understand the mechanisms of cancer immunity and improve clinical benefit of cancer immunotherapy. However, progress in current research is hindered by the lack of comprehensive immune activity resources and easy-to-use tools for biologists, clinicians, and researchers to conveniently evaluate immune activity during the "cancer-immunity cycle." We developed a user-friendly one-stop shop web tool called TIP to comprehensively resolve tumor immunophenotype. TIP has the capability to rapidly analyze and intuitively visualize the activity of anticancer immunity and the extent of tumor-infiltrating immune cells across the seven-step cancer-immunity cycle. Also, we precalculated the pan-cancer immunophenotype for 11,373 samples from 33 The Cancer Genome Atlas human cancers that allow users to obtain and compare immunophenotype of pan-cancer samples. We expect TIP to be useful in a large number of emerging cancer immunity studies and development of effective immunotherapy biomarkers. TIP is freely available for use at http://biocc.hrbmu.edu.cn/TIP/.

Significance: TIP is a one-stop shop platform that can help biologists, clinicians, and researchers conveniently evaluate anticancer immune activity with their own gene expression data. *Cancer Res; 78(23); 6575–80. ©2018 AACR.*

See related commentary by Hirano, p. 6536

Introduction

The emergence of cancer immunotherapy has revolutionized cancer treatment, whose success highly depends on the immune cell development and activation in the host microenvironment (1, 2). Anticancer immune response can be conceptualized as a series of stepwise events called the cancer-immunity cycle, including release of cancer cell antigens (step 1), cancer antigen presentation (step 2), priming and activation (step 3), trafficking of immune cells to tumors (step 4), infiltration of immune cells into tumors (step 5), recognition of cancer cells by T cells (step 6), and killing of cancer cells (step 7; ref. 3). Activity status of these seven-step anticancer immune responses and the extent of tumor-infiltrating immune cells form the complex tumor immunophenotype underlying the tumor microenvironment. Therefore, tracking tumor immunophenotype will be essential for the understanding of the mechanisms of cancer immunity and the development of biomarkers of response to immunotherapy.

Currently, there is a substantial lack of global characterization of the whole anticancer immunity, due to the lengthy collection of typical signature through literature curation and the requirement of additional expertise in computer programming and statistical inference. Although some methods were proposed for analyzing part of the cancer-immunity cycle, they had the limitation of long execution time, when dealing with a large sample size, and, generally, were short of the intuitive display of results that may be important for interpretation (4–7). Obviously, there is an urgent demand for an easy-to-use and one-stop shop tool for comprehensively resolving the activity status of the whole anticancer immunity.

To address this problem, we developed TIP (tracking tumor immunophenotype), a user-friendly web-based tool focusing on profiling the anticancer immune microenvironment based on the seven-step cancer-immunity cycle. TIP provides a one-stop shop for systematically tracking, analyzing, and visualizing of the activity status of anticancer immunity and the extent of tumor-infiltrating immune cells in user-defined samples (see Supplementary Video S1). Even without additional computer and programming expertise, biologists and clinicians can comprehensively investigate the tumor immunophenotype easily and conveniently. We manually collected 23 confirmed signature sets involved in seven-step anticancer immunity, allowing researchers to explore the immune status without time consumption for lengthy literature curation and data gathering (8, 9). In addition, we prebuilt the landscape of The Cancer Genome Atlas (TCGA) tumor immunophenotype for 11,373 patients across 33 TCGA human cancer types, so that users can analyze and compare the immune status across the tumor samples.

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, Heilongjiang, China. [2]Department of Neurology, The First Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

L. Xu, C. Deng, B. Pang, and X. Zhang contributed equally to this article.

Corresponding Authors: Xia Li, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China. E-mail: lixia@hrbmu.edu.cn; and Yun Xiao, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China. E-mail: xiaoyun@ems.hrbmu.edu.cn

doi: 10.1158/0008-5472.CAN-18-0689

Materials and Methods

Depict the status of anticancer immunity

To characterize the status of anticancer immunity, we manually collected 178 signature genes (including stimulatory and inhibitory genes, grouped into 23 sets) involved in the seven steps of the cancer-immunity cycle from published studies, with the keywords such as "checkpoints," "cytotoxic factors," "chemokines," and "MHC molecules." Then, the activity levels of these signature gene sets were calculated using single-sample Gene Set Enrichment Analysis (ssGSEA) based on gene expression of individual samples (4). Notably, the stimulatory (positive) and inhibitory (negative) gene sets in each step of the cancer-immunity cycle were calculated, respectively (i.e., $ES_{positiveSet}$ and $ES_{negativeSet}$). To make ssGSEA scores comparable between different expression platforms and different samples, the activity scores were normalized according to the z-score method after 100 permutations (i.e., $NES_{positiveSet}$ and $NES_{negativeSet}$). Finally, the activity score for each signature set was produced by calculating the difference between the normalized ssGSEA scores of positive set and negative set (i.e., $NES_{positiveSet} - NES_{negativeSet}$).

Depict the proportion of tumor-infiltrating immune cells

TIP is also designed to infer the proportion of various tumor-infiltrating immune cells, such as T cells, B cells, dendritic cells (DC), natural killer (NK) cells, macrophages, based on a widely used gene-expression deconvolution algorithm, CIBERSORT (6). CIBERSORT estimates the cell fractions using nu-support vector regression (ν-SVR) algorithm through considering a gene-signature matrix, which describes the expression patterns of different immune cell types. When users upload microarray expression profiling, TIP will estimate the fractions of infiltrated immune cell types for each sample by using the original leukocyte gene signature matrix (LM22) derived from CIBERSORT, which contains 547 genes that distinguish 22 human hematopoietic cell types. To infer the cell proportion for RNA-seq bulk expression data using CIBERSORT, we constructed a gene-signature expression matrix (LM14, involving seven T-cell types, B cells, CD14 and CD16 monocytes, DCs and plasmacytoid DC cells, plasma cells, and NK cells; Supplementary Table S1A) based on single-cell RNA-seq data (scRNA-seq), which is composed of 33,000 peripheral blood mononuclear cells (PBMC), 2,000 $CD4^+$ helper t cells, and 2,000 $CD4^+/CD25^+$ regulatory T cells from the 10× Genomics (https://support.10xgenomics.com/single-cell-gene-expression/datasets, Supplementary Table S1B). In detail, we preprocessed the raw count data of the scRNA-seq data by R package Seurat (10). The count values were converted to TPM using BSEQsc software (11). Moreover, we identified differentially expressed genes for each cell type (vs. others) using SCde package (Supplementary Table S1C; ref. 12). To generate an appropriate signature gene set for LM14, we merged these differentially expressed genes among the 14 cell types and signature genes from the LM22. Finally, the LM14 gene signature matrix, which is a TPM expression matrix of 973 signature genes (row) and 14 immune cell types (columns), was built to infer the proportion of tumor-infiltrating immune cells using CIBERSORT when users upload RNA-seq bulk expression data.

Software development

TIP computational functions were developed in R; Bootstrap framework, Struts2, and JavaScript were used to communicate between R and web interfaces; the web interfaces were achieved using JavaScript, tabular results were generated by DataTables, heat map, bar, pie, and scatter plots were generated by HighCharts, D3 provided power for line, box, radar, and circle plots.

Source code: https://github.com/dengchunyu/TIP.

Input

In the general analysis, TIP accepts two types of bulk tumor expression data (i.e., the whole tumor samples) as input. One is RNA-seq expression data (raw count or TPM), and the other is microarray expression data (log or non-log transformed). In the pan-cancer analysis, TIP provides the precalculated immunophenotype for 33 TCGA cancer types.

Output

The output allows users to interactively inspect the results from a global (Fig. 1) and individual perspective (Fig. 2). A detailed help page is provided to serve as a guide for users in understanding the usage of TIP and interpreting the output data.

From a global perspective, TIP displays the 23 immune activity scores reflecting the activity status of the seven-step cancer-immunity cycle for all samples (Fig. 1A). Below the activity scores, the relative proportion of tumor-infiltrating immune cells across all samples are provided, in which users can select and focus on immune cells of interest (Fig. 1B). The expression pattern and the principal component analysis of 178 step-specific signature genes are also presented (Fig. 1C). Note that an overall activity score for each sample is shown on the right of Fig. 1D, allowing to investigate immunophenotype from an individual perspective. When the user clicks a sample, he is provided with the overall information of immune activity, immune cell infiltration, and signature gene expression of this single sample (Fig. 2A and B). Importantly, TIP interactively displays the status of the seven-step cancer-immunity cycle, within the context of the reference tumor immunophenotype of the corresponding cancer type from TCGA pan-cancer data, by circle plots (Fig. 2C), which can help to evaluate the activity of anticancer immunity for this sample. Additionally, the reference tumor immunophenotypes for 11,373 patients across 33 TCGA human cancers are provided in PancancerAnalysis. All of the interactive images and the corresponding data can be downloaded via Download button for further analysis.

Results

Pan-cancer immunophenotype profiling

TIP provides the tumor immunophenotype profiling of 11,373 patients across 33 TCGA human cancers in the PancancerAnalysis tab, from which users can resolve anticancer immune activity status and the extent of immune cell infiltration for all patients.

Taking the BRCA (breast invasive carcinoma with 1,256 samples) as an example, we found that the signature genes involved in the cancer-immunity cycle showed distinct expression patterns between tumor and normal samples (Fig. 1C). Even among the tumor samples, an obvious difference was still observed. In the single sample TCGA-A2-A0ST-01A-12R-A084-07 with the highest overall activity, higher immune activity scores and more T-cell infiltration were observed when compared with most of the reference TCGA BRCA samples (Fig. 2C).

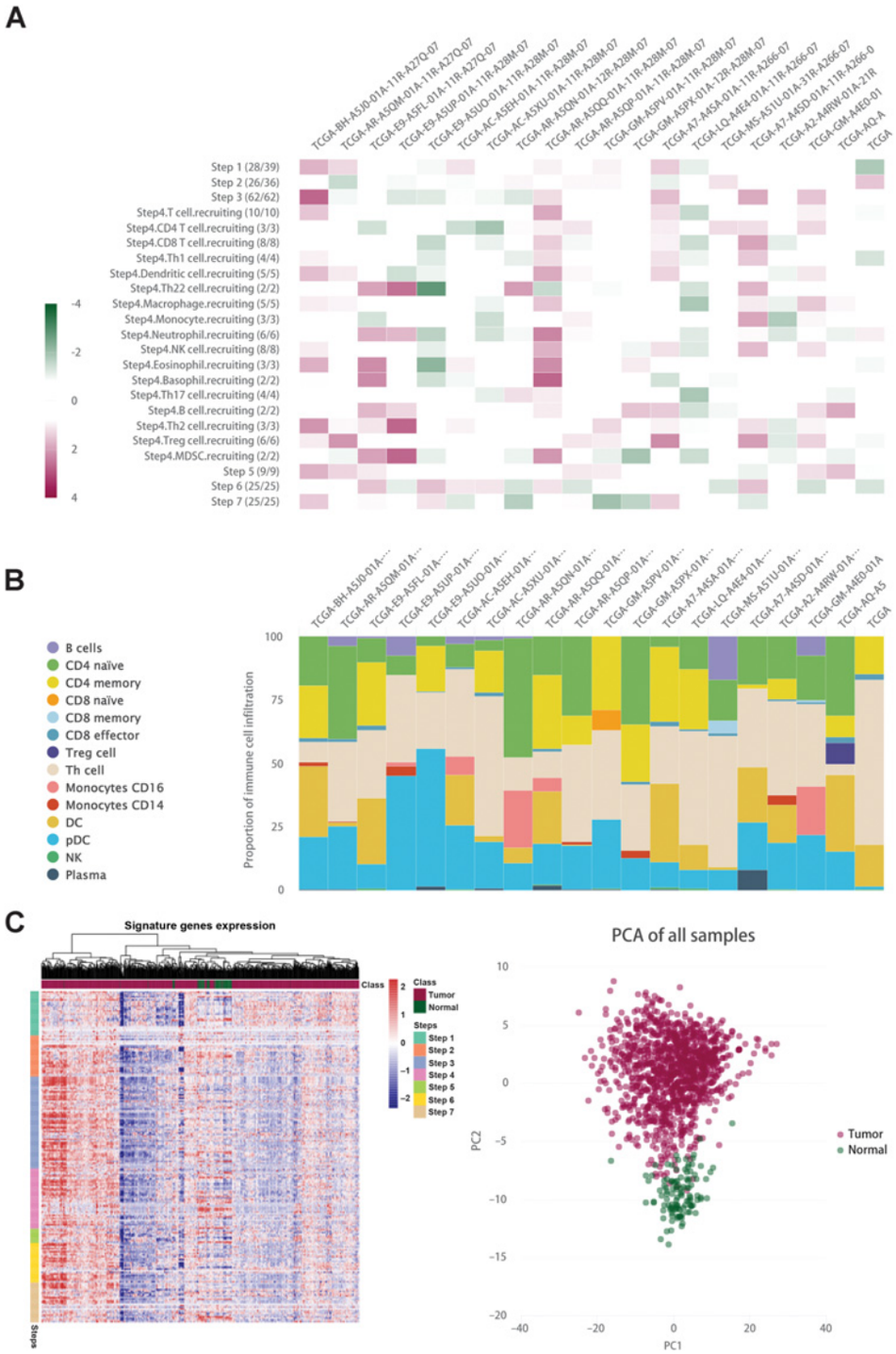

Figure 1.

Global visualization of the immunophenotype across 1,256 BRCA samples from TCGA. **A,** The 23 normalized immune activity scores. **B,** The relative proportion of tumor-infiltrating immune cells. **C,** Left, the expression pattern of 178 signature genes from the seven-step cancer-immunity cycle. Each row represents a single gene, and each column represents one sample. Right, the principal component analysis (PCA) of signature genes expression for all samples (red, tumor samples; green, normal samples). **D,** Overall activity scores of anticancer immunity (by summing up the 23 normalized scores) for each sample.

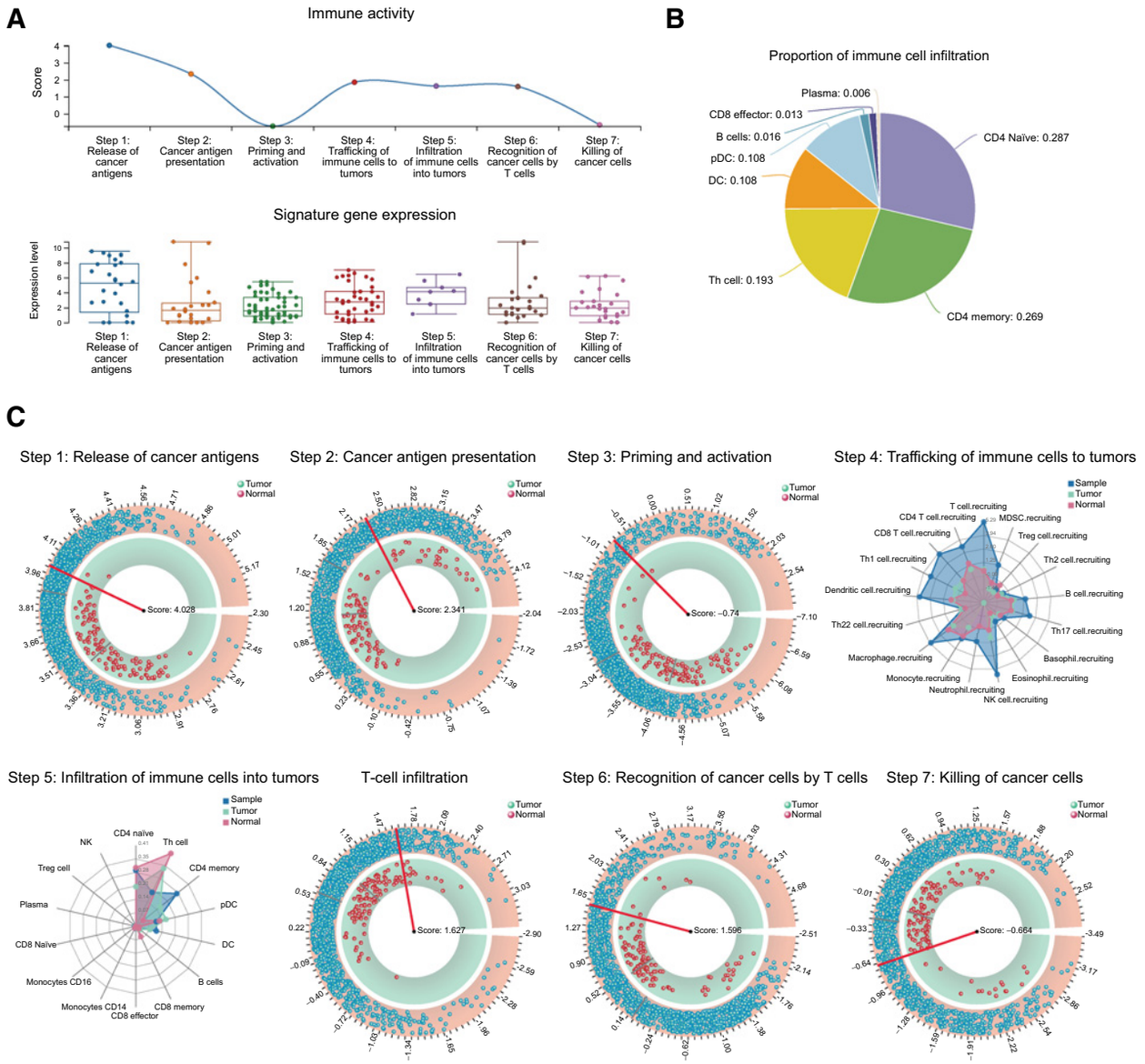

Figure 2.

Visualization of immunophenotype for one specific BRCA sample, TCGA-A2-A0ST-01A-12R-A084-07. **A,** The line graph shows activity scores of anticancer immunity in seven steps across the cancer-immunity cycle for this sample. The boxplot shows the expression distribution of signature genes from seven steps of the cancer-immunity cycle. **B,** Relative proportions of tumor-infiltrating immune cells. Only cells with detectable proportions (>0) are shown. **C,** Six circular plots show the comparison of activity scores of anticancer immunity for this specific sample (red pointer) with those for all tumor (green points in outer circle) and normal (red points in inner circle) samples from TCGA BRCA data. Two radar plots show the activity scores of "Trafficking of immune cells to tumors" and the relative proportions of tumor-infiltrating immune cells for different immune cell types in this sample (blue) and all tumor and normal samples (overall sample mean; green, tumor; red, normal) from TCGA BRCA data.

A case study for anti–PD-1 immunotherapy

RNA-seq data (Illumina HiSeq 2000, GSE78220) from Hugo and colleagues (13), including 28 samples of melanoma following anti–PD-1 immunotherapy (responding, $n = 15$; nonresponding, $n = 13$), were uploaded to TIP, with TCGA SKCM (Skin Cutaneous Melanoma) as the reference cancer type (all the results were available at JobID: 20180120084858Z4B95IHRQQZC1X). From the global perspective, we found that the nonresponding samples showed significantly higher activity scores of "Step1. release of cancer cell antigens" ($P = 0.041$, Mann–Whitney U test) and "Step4.MDSC.recruiting" ($P = 0.046$, Mann–Whitney U test) than the responding samples (Supplementary Fig. S1A).

From the individual perspective, we observed that the nonresponding patient Pt1 had substantial infiltration of T regulatory cells, but lacked the infiltration of CD8 effector T cells that have been found to be a predictor for anti–PD-1 therapy response (Supplementary Fig. S1B; ref. 14), despite the higher activity

scores of "step1.release of cancer cell antigens" than all reference TCGA SKCM samples.

Discussion

We have developed an open-access, user-friendly web-based tool, TIP, which provides a one-stop shop to comprehensively evaluate the tumor immunophenotype including activity status of the seven-step cancer-immunity cycle and immune cell infiltration. TIP enhances the display of results by multiple advanced and interactive visualizations and computes five to ten times faster than simple chaining together of ssGSEA (using R package "GSVA") and CIBERSORT. For 50 samples, total execution time would be reduced from more than 100 minutes to around 20 minutes, a saving of 80% (Supplementary Table S2). TIP manually collected 23 confirmed signature sets for assessing anticancer immune activity status and prebuilt TCGA tumor immunophenotype landscape for analysis and comparison across samples. In summary, TIP substantially expands the available repertoire of current softwares and facilitates tracking the tumor immunophenotype without lengthy literature curation, long execution time, and additional expertise in computer programming.

As a key component of TIP, a detailed and biologically relevant validation of immune cell fraction and anticancer immunity activity status was performed. We assessed TIP analysis using bulk microarray data set from SKCM tumor samples of TCGA ($n = 328$; ref. 15), 27 colon cancer primary tumors (GSE39582; ref. 16), the RNA-seq data set of bulk PBMC samples from two different donors (GSE64655; ref. 17), and microarray data set from PBMCs of 20 adult subjects (GSE65136; ref. 6), as well as the corresponding relative proportion of immune cell types determined by immunostaining and flow cytometry analysis. All of the above results suggested that TIP analysis was significantly consistent with corresponding experimental measurements (Supplementary Fig. S2A–S2D). In addition, to evaluate anticancer immunity status in TIP analysis, we analyzed the bulk RNA-seq data set of urothelial carcinoma samples from http://doi.org/10.5281/zenodo.546110, which measured corresponding PD-L1 tumor-infiltrating immune cell status by IHC analysis (18). The result demonstrated that the patients with the higher level of PD-L1 tumor-infiltrating immune cell status showed significantly lower anticancer immunity activity scores of "Step6. Recognition of cancer cells by T cells." In general, the agreement of TIP analysis and ground truth values indicated that the immune cell fraction and the anticancer immunity activity status can be quantitatively exhibited directly using TIP (Supplementary Fig. S2E).

Recently, DNA methylation patterns were found to be highly variable across different cell types and were able to precisely characterize patient phenotype (19–21). Thus, DNA methylation profiling, such as transcriptome profiling, may also be used to effectively resolve tumor immunophenotype. Indeed, DNA methylation patterns have been harnessed by recent studies to estimate proportions of infiltrating immune cell types. In a seminal paper, Houseman and colleagues (22) used the DNA methylation markers that are highly discriminative of five principal components of human white blood cells to accurately infer relative proportions of immune cell types in whole blood. This method was expanded by additional reference differentially methylated regions of more leukocyte subsets (23). Furthermore, focusing on the whole population of leukocytes, Li and colleagues and Thorsson and colleagues both identified DNA methylation loci with the greatest difference between pure leukocytes and normal tissues and then investigated the leukocyte fractions of overall tumor tissues based on different linear models (24, 25). Therefore, in the future, it is necessary to incorporate DNA methylation-based analysis into our platform to capture more comprehensive immunophenotype of patients. Also, we will combine the clinical phenotype information (such as survival and response to immunotherapy) and add tools for developing effective biomarkers. We expect that TIP will continue to serve as a powerful platform for assisting the anticancer immunity field and cancer immunotherapy community.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: T. Zhao, Y. Xiao, X. Li

Development of methodology: L. Xu, C. Deng, B. Pang, X. Zhang, W. Liu, T. Zhao

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L. Xu, C. Deng, B. Pang, X. Zhang, H. Yuan, P. Cheng, F. Li

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): L. Xu, C. Deng, B. Pang, X. Zhang, G. Liao, H. Yuan, M. Yan, Y. Xiao

Writing, review, and/or revision of the manuscript: L. Xu, C. Deng, B. Pang, X. Zhang, W. Liu, G. Liao, H. Yuan, P. Cheng, F. Li, Z. Long, M. Yan, T. Zhao, Y. Xiao, X. Li

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): L. Xu, C. Deng, B. Pang, X. Zhang, W. Liu, G. Liao, P. Cheng, F. Li, Z. Long, X. Li

Study supervision: T. Zhao, Y. Xiao, X. Li

Acknowledgments

We thank Dr. Etienne Becht from Singapore Immunology Network for generous support with IHC data. We thank Dr. Lin Pang from Harbin Medical University for the helpful suggestions and contributions.

This work was supported in part by the National High Technology Research and Development Program of China (863 Program, Grant No. 2014AA021102), the National Program on Key Basic Research Project (973 Program, Grant No. 2014CB910504), the National Natural Science Foundation of China (Grant Nos. 91439117, 61473106, and 61573122), the China Postdoctoral Science Foundation (2016M600260), Wu lien-teh youth science fund project of Harbin Medical University (Grant No. WLD-QN1407), Special funds for the construction of higher education in Heilongjiang Province (Grant No. UNPYSCT-2016049), and the Heilongjiang Postdoctoral Foundation (LBH-Z16098).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 8, 2018; revised July 20, 2018; accepted August 20, 2018; published first August 28, 2018.

References

1. Daud AI, Wolchok JD, Robert C, Hwu WJ, Weber JS, Ribas A, et al. Programmed death-ligand 1 expression and response to the anti-programmed death 1 antibody pembrolizumab in melanoma. J Clin Oncol 2016;34:4102–9.

2. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252–64.
3. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity 2013;39:1–10.
4. Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 2009;462:108–12.
5. Hackl H, Charoentong P, Finotello F, Trajanoski Z. Computational genomics tools for dissecting tumour-immune cell interactions. Nat Rev Genet 2016;17:441–58.
6. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12:453–7.
7. Yuan Y, Failmezger H, Rueda OM, Ali HR, Graf S, Chin SF, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. Sci Transl Med 2012;4:157ra43.
8. Harlin H, Meng Y, Peterson AC, Zha Y, Tretiakova M, Slingluff C, et al. Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment. Cancer Res 2009;69:3077–85.
9. Nagarsheth N, Wicha MS, Zou W. Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. Nat Rev Immunol 2017;17:559–72.
10. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 2015;33:495–502.
11. Baron M, Veres A, Wolock SL, Faust AL, Gaujoux R, Vetere A, et al. A single-cell transcriptomic map of the human and mouse pancreas reveals inter- and intra-cell population structure. Cell Syst 2016;3:346–60e4.
12. Fan J, Salathia N, Liu R, Kaeser GE, Yung YC, Herman JL, et al. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. Nat Methods 2016;13:241–4.
13. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. Cell 2017;168:542.
14. Speiser DE, Ho PC, Verdeil G. Regulatory circuits of T cell function in cancer. Nat Rev Immunol 2016;16:599–611.
15. Cancer Genome Atlas N. Genomic classification of cutaneous melanoma. Cell 2015;161:1681–96.
16. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016;17:218.
17. Hoek KL, Samir P, Howard LM, Niu X, Prasad N, Galassie A, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. PLoS One 2015;10:e0118528.
18. Snyder A, Nathanson T, Funt SA, Ahuja A, Buros Novik J, Hellmann MD, et al. Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: an exploratory multi-omic analysis. PLoS Med 2017;14:e1002309.
19. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. Nat Rev Genet 2018;19:129–47.
20. Montano CM, Irizarry RA, Kaufmann WE, Talbot K, Gur RE, Feinberg AP, et al. Measuring cell-type specific differential methylation in human brain tissue. Genome Biol 2013;14:R94.
21. Teschendorff AE, Breeze CE, Zheng SC, Beck S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. BMC Bioinformatics 2017;18:105.
22. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012;13:86.
23. Accomando WP, Wiencke JK, Houseman EA, Nelson HH, Kelsey KT. Quantitative reconstruction of leukocyte subsets using DNA methylation. Genome Biol 2014;15:R50.
24. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome Biol 2016;17:174.
25. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. Immunity 2018;48:812–30e14.