TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells

Taiwen Li[1,2], **Jingyu Fan**[3], **Binbin Wang**[3], **Nicole Traugh**[2], **Qianming Chen**[1], **Jun S Liu**[4], **Bo Li**[2,4,*], and **X. Shirley Liu**[2,3,4,*]

[1]State Key Laboratory of Oral Diseases, West China Hospital of Stomatology, Sichuan University, Chengdu, China

[2]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA, USA

[3]School of Life Science and Technology, Tongji University, Shanghai, China

[4]Department of Statistics, Harvard University, Boston, MA, USA

Abstract

Recent clinical successes of cancer immunotherapy necessitate the investigation of the interaction between malignant cells and the host immune system. However, elucidation of complex tumor-immune interactions presents major computational and experimental challenges. Here we present Tumor Immune Estimation Resource (TIMER, cistrome.shinyapps.io/timer) to comprehensively investigate molecular characterization of tumor-immune interactions. Levels of six tumor-infiltrating immune subsets are pre-calculated for 10,897 tumors from 32 cancer types. TIMER provides 6 major analytic modules that allow users to interactively explore the associations between immune infiltrates and a wide-spectrum of factors, including gene expression, clinical outcomes, somatic mutations, and somatic copy number alterations. TIMER provides a user-friendly web interface for dynamic analysis and visualization of these associations, which will be of broad utilities to cancer researchers.

Keywords

web resource; immune infiltrates; tumor-immune interaction; cancer genomics; cancer immunology

Introduction

Recent advances in immunotherapy, especially checkpoint blockade, have resulted in clinical success in treating late-stage cancers (1). Immune infiltrates obtained from different patients

[*]Corresponding Authors: X. Shirley Liu, Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, 450 Brookline Ave., Boston, MA 02215, USA. Tel: +1(617)632-2472. Fax: +1(617)632-2444. xsliu@jimmy.harvard.edu. Bo Li, Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, 450 Brookline Ave., Boston, MA 02215, USA. Tel: +1(734)546-7574. Fax: +1(617)632-2444. bli@jimmy.harvard.edu.

Disclosure of Potential Conflicts of Interest

The authors declare no potential conflicts of interest.

with the same tumor type are heterogeneous and may impact clinical outcome (2). Due to the complexity of tumor genomes and the plasticity of the host immune system, it remains challenging to characterize the interaction between cancer cells and immune infiltrates (3). Recently, a variety of computational methods to infer immune infiltrates have been developed to investigate tumor immunology and develop new effective immunotherapies (4–7). One seminal study measures effector cell cytolytic activity using the transcript levels of two genes, *GZMA* and *PRF1*, to elucidate possible mechanisms of immune evasion (8). Immune signature genes have also been used to characterize immune infiltrates and predict clinical outcome (4). One influential deconvolution method, CIBERSORT, utilizes microarray data with a pre-defined immune signature matrix to estimate the fraction of 22 tumor-infiltrating immune cells (TIICs) within a given sample (5). Its results have been presented as online resources, such as PRECOG (6) and TCIA (7), to analyze the impact of TIICs and individual gene expression levels on clinical outcome. However, CIBERSORT estimations are potentially affected by statistical multicollinearity due to the inclusion of highly correlated immune cell types, leading to high estimations of uncertainty (Supplementary Material). In addition, PRECOG and TCIA have limited analysis and visualization functions.

Despite these efforts, a comprehensive computational tool for cancer researchers to conveniently explore and visualize tumor immunological and genomics data is still lacking. Therefore, we developed an interactive web application, TIMER, as a public resource to enable cancer biologists to analyze and visualize the abundance of TIICs in a comprehensive and flexible manner. TIMER applies a deconvolution method we previously published (9) to infer the abundance of TIICs from gene expression profiles. In this work, we re-analyze gene expression data, which includes 10,897 samples across 32 cancer types from The Cancer Genome Atlas (TCGA) to estimate the abundance of six TIIC subsets (B cells, CD4 T cells, CD8 T cells, macrophages, neutrophils, and dendritic cells). The workflow is based on our previous work of statistical deconvolution of immune infiltrates (9). In brief, we select informative genes which are negatively correlated with tumor purity (percentage of malignant cells in a tumor tissue) for each cancer type (10) and apply constrained least squares fitting on the expression of selected genes to predict the abundance of six TIIC subsets.

Results

To dynamically explore and visualize the interactions between TIICs and important tumoral genomic changes, we introduce TIMER, a user-friendly web interface (cistrome.shinyapps.io/timer). TIMER consists of six functional modules, including the association of TIIC abundance with gene expression (Gene), overall survival (Survival), somatic mutations (Mutation), and DNA somatic copy number alterations (SCNA), as well as analysis of differential gene expression (DiffExp) and gene-gene correlations (Correlation) (Fig. 1). Each module allows users to input function-specific parameters, with resulting figures dynamically generated and displayed at the bottom of the TIMER webpage. Detailed descriptions of each module are described below.

Gene Module

Malignant cells and TIICs interact through multiple gene products and pathways during tumor development and progression. The "Gene" module allows users to input a gene symbol, cancer types, and immune cell types of interest. A scatterplot displaying gene expression levels against tumor purity is first displayed. This is because the tumor microenvironment is a complex milieu of non-cancerous cells mainly consisting of immune cells around tumor cells, and genes highly expressed in cells in the microenvironment are expected to have negative associations with tumor purity, while the opposite is expected for genes highly expressed in the tumor cells. Therefore, this panel is helpful to identify the cancer type and specific TIIC subsets that are associated with the expression of a certain gene of interest. For example, if users input the granzyme A gene (*GZMA)*, breast cancer (BRCA), and all six TIIC subsets, a negative correlation between *GZMA* expression and tumor purity is observed because granzymes are produced in the tumor microenvironment. Six panels, showing the purity-corrected partial Spearman's correlation and statistical significance between *GZMA* expression and each TIIC subset, would also be displayed. To facilitate easy comparisons across different cancer types, TIMER allows users to choose multiple cancer types and simultaneously displays all correlations.

Survival Module

Many studies have reported a prominent impact of TIICs on clinical outcome of cancer patients (11). The "Survival" module allows users to explore the clinical relevance of one or more TIIC subsets, with the flexibility to correct for multiple covariates in a multivariable Cox proportional hazard model. Clinical covariates include patient age, gender, ethnicity, and tumor stages where T, N, and M delineate primary, lymph node spreading, and metastatic tumors respectively. Users are also allowed to include a list of gene expression values into the Cox model to investigate the clinical relevance of genes as well as TIICs. TIMER outputs the Cox regression results including hazard ratios and statistical significance. Users can easily explore the clinical impact of a gene of interest and correct for potential confounding factors. For example, applying this module, we could find that CD40 ligand (CD40LG) independently predicts better outcome (HR=0.79, p=0.003) in lung adenocarcinoma corrected for patient age, stage and TIICs. This result is consistent with the established role of CD40/CD40LG in regulating immune cell function (12). For each survival analysis, TIMER outputs Kaplan-Meier plots for TIICs and genes to visualize the survival differences between the upper and lower x percentile of patients, where x can be adjusted using a slider. A log-rank p-value is also calculated and displayed for each Kaplan-Meier plot.

Genetic Aberration Modules

Tumor development and progression are associated with multiple genomic aberrations which in turn may influence TIICs. The "Mutation" and "SCNV" modules are designed to facilitate the investigation of such associations. In the "Mutation" module, users are able to compare the abundance of TIICs with or without the presence of a given somatic mutation. To identify potential neoantigens or genes that are related to immune infiltration, we only

include non-synonymous mutations that can cause amino acid changes in the protein sequence. Therefore, we chose the top 50 genes with the most frequent non-synonymous mutations as options in the "Mutated Gene" field for each cancer type. Box plots are generated for each TIIC subset, to compare the distribution of the abundance of TIICs with different gene mutation status, with statistical significance estimated using two-sided Wilcoxon Rank Sum Test. For example, when head and neck cancer and TP53 mutation are selected, the returned box plots show significantly lower infiltration for most immune cells in tumors with TP53 mutations. This result is potentially related to the important role that p53 plays in regulating the innate immune response (13). The "SCNA" module provides the comparison of the abundance of TIICs among tumors with different somatic copy number aberrations for a given gene. SCNAs are defined by GISTIC 2.0 (14), including deep deletion, shallow deletion, diploid/normal, low-level gain, and high amplification. Box plots are presented to show the distributions of each TIIC subset for each copy number status in selected cancer types using the same statistical tests as in the "Mutation" module.

Gene Comparison Modules

TIMER provides two additional modules, "DiffExp" and "Correlation", for researchers to explore interesting genes expressed in tumor and adjacent normal tissue or gene-gene associations that may be related to cancer immunity. The "DiffExp" module allows users to study the differential expression between tumor and adjacent normal tissues for any gene of interest across all TCGA tumors. Distributions of gene expression levels are displayed using box plots, with statistical significance of differential expression evaluated using Wilcoxon test. Users can identify genes that are up- or down- regulated in the tumors compared to normal tissues for each cancer type, as displayed in gray columns when normal data are available. The "Correlation" module plots the expression scatterplots between a pair of user-defined genes in a given cancer type, together with the Spearman's correlation and estimated statistical significance, which could be adjusted by tumor purity or age optionally. A brief video tutorial is provided in the supplementary material (See Video 1) and the main page of TIMER website.

Discussion

The increasing number and complexity of cancer immunotherapy studies in recent years pose new challenges to analyze and visualize the rapidly accumulating clinical and cancer genomics data. Characterization of the immune infiltration landscape is necessary for the investigation of tumor-immune interactions, which requires innovative computational methods to integrate and deconvolve multi-dimensional datasets. Our study provides a comprehensive analytical web tool for cancer researchers to conveniently access the tumor immunological, clinical, and genomic features. Intuitive inputs and outputs of the six functional modules in TIMER simplify the analysis of the interactions between genes of interest and immune cells in multiple cancer types. To the best of our knowledge, TIMER is the first method that allows users to perform integrative analysis of tumor immunological, clinical, and genomics data. The development and maintenance of the TIMER web server will be carried out by our laboratory for a minimum of five years. With the increasing numbers of studies of cancer immunology, we anticipate TIMER to attract broad interest

from the cancer research community and aid the discovery of novel therapeutic targets and new immune evasion mechanisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer. 2012; 12:252–64. [PubMed: 22437870]

2. Li B, Li T, Pignon J-C, Wang B, Wang J, Shukla SA, et al. Landscape of tumor-infiltrating T cell repertoire of human cancers. Nature Genetics. 2016; 48:725–32. [PubMed: 27240091]

3. Hackl H, Charoentong P, Finotello F, Trajanoski Z. Computational genomics tools for dissecting tumour-immune cell interactions. Nature reviews Genetics. 2016; 17:441–58.

4. Iglesia MD, Parker JS, Hoadley KA, Serody JS, Perou CM, Vincent BG. Genomic Analysis of Immune Cell Infiltrates Across 11 Tumor Types. Journal of the National Cancer Institute. 2016; 108:djw144.

5. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nature Methods. 2015; 12:453–7. [PubMed: 25822800]

6. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nature Medicine. 2015; 21:938–45.

7. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell reports. 2017; 18:248–62. [PubMed: 28052254]

8. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. CELL. 2015; 160:48–61. [PubMed: 25594174]

9. Li B, Severson E, Pignon J-C, Zhao H, Li T, Novak J, et al. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome biology. 2016; 17:174. [PubMed: 27549193]

10. Li B, Li JZ. A general framework for analyzing tumor subclonality using SNP array and DNA sequencing data. Genome biology. 2014; 15:473. [PubMed: 25253082]

11. Fridman WH, Pages F, Sautes-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. Nature Reviews Cancer. 2012; 12:298–306. [PubMed: 22419253]

12. Kikuchi T, Crystal RG. Anti-tumor immunity induced by in vivo adenovirus vector-mediated expression of CD40 ligand in tumor cells. Hum Gene Ther. 1999; 10:1375–87. [PubMed: 10365667]

13. Menendez D, Shatz M, Resnick MA. Interactions between the tumor suppressor p53 and immune responses. Curr Opin Oncol. 2013; 25:85–92. [PubMed: 23150340]

14. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2. 0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome biology. 2011; 12:R41. [PubMed: 21527027]
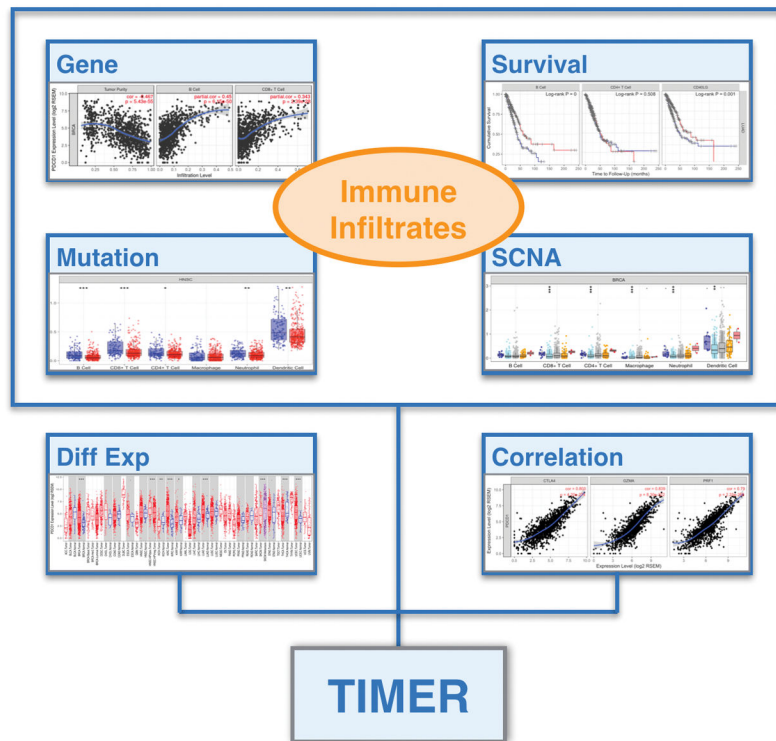
Figure 1.
Overview of TIMER modules on the website. TIMER comprises of six modules. Four modules in the top box are to explore the associations of TIICs with gene expression (Gene), overall survival (Survival), somatic mutation (Mutation) and somatic copy number alteration (SCNA), as well as analysis of differential gene expression (DiffExp) and correlation between two groups of genes (Correlation). Examples visualization for each module are displayed in the corresponding text boxes.