

Computational deconvolution of transcriptomic data for the study of tumor-infiltrating immune cells

Marco Bolis, Arianna Vallerga and Maddalena Fratelli 

The International Journal of Biological
Markers

2020, Vol. 35(S1) 20–22

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1724600820903317

journals.sagepub.com/home/ijbm



Abstract

Cancer is a complex disease characterized by a wide array of mutually interacting components constituting the tumor microenvironment (connective tissue, vascular system, immune cells), many of which are targeted therapeutically. In particular, immune checkpoint inhibitors have recently become an established part of the treatment of cancer. Despite great promise, only a portion of the patients display durable response. Current research efforts are concentrated on the determination of tumor-specific biomarkers predictive of response, such as tumor mutational burden, microsatellite instability, and neo-antigen presentation. However, it is clear that several additional characteristics pertaining to the tumor microenvironment play a critical role in the effectiveness of immunotherapy. Here we comment on the computational methods that are used for the analysis of the tumor microenvironment components from transcriptomic data, discuss the critical needs, and foresee potential evolutions in the field.

Keywords

Tumor immune contexture, computational methods, tumor microenvironment, cell-specific transcriptomics

Date received: 16 December 2019; accepted: 10 January 2020

Short Communication

One important source of information in the precision medicine armamentarium comes from the transcriptomic analysis of samples from tumor tissues (biopsies or surgical specimens). These are mixtures of several different cell types in addition to cancer cells, and include fibroblasts, adipocytes, endothelial, and immune cells. We refer to this complex ecosystem of interacting cells as the tumor microenvironment (TME). Changes in the cell composition of TME are associated with functional alterations. In particular, the immune contexture of solid tumors in humans is a recognized hallmark of cancer with prognostic potential.¹ The clinical outcome of multiple cancer types in patients was shown to depend on the pre-existing adaptive immunity.² A growing number of reports demonstrated an improved anti-tumor response to therapy when combined with checkpoint inhibitors^{3,4} and an immune-based rationale to guide the use of such therapeutic strategies has been advocated.⁵

Recently, several computational methods have been reported for predicting fractions of multiple cell types in bulk gene expression profiles of tissue samples. These methods are based on the consideration that the gene expression profile of a heterogeneous sample is the convolution of the gene expression levels of the different cell components. The quantitative estimation of the unknown cell fractions is therefore based on a signature matrix describing the cell-type-specific expression profiles, which should be known in advance. The expression of each gene in the heterogeneous sample will be the weighted sum of its expression values of each cell type present in the mixture. Several computational approaches have been developed to

Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milano, Italy

Corresponding author:

Maddalena Fratelli, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Via Mario Negri 2, Milan, 20156, Italy.

Email: maddalena.fratelli@marionegri.it



solve this deconvolution problem and have been extensively reviewed in recent publications.⁶⁻⁸

As the deconvolution process depends on both the mathematical analysis and the use of gene signatures specific for cell types, we would like to point out that the quality of the results will depend on the quality of such lists as much as on the efficacy of the deconvolution algorithms. Actually, Vallania et al.⁹ demonstrated that the basis matrix of gene signatures is the major determinant of deconvolution accuracy, and that virtually no computational method can overcome possible biological and technical bias present in a basis matrix. In fact, they tested different matrices and different methods and found that for a given basis matrix, all methods gave highly correlated proportions, while for a given method, the use of different basis matrices showed a lower correlation in the estimated proportions.

Therefore, the availability of accurate cell-type-specific gene signatures is of the utmost importance for the success of deconvolution approaches. Moreover, the gene expression profile of a single cell type can change depending on the interaction with other surrounding cells or on the presence of other external stimuli and can be substantially different in health or disease conditions. This will require additional condition-specific information to be incorporated in the cell-type-specific signatures. The first gene matrices used in the deconvolution approaches were derived from microarray profiling on sorted immune cells from healthy individuals—IRIS^{10,11} and LM22¹²—and may have limited use in the analysis of data obtained by the next-generation sequencing platforms. Recently, Monaco et al.¹³ characterized 29 healthy human immune cell types by RNA sequencing after a complex flow cytometry isolation procedure, and defined modules of specific, co-expressed, and housekeeping genes. This effort, together with the definition of an optimized normalization approach, allowed absolute deconvolution of human immune cell types.

The approaches described above have a major limitation in being restricted to circulating immune cells from healthy individuals and may not describe tumor-infiltrating immune cells and their heterogeneity. Moreover, limited data are available for other cell components of the TME, which may have a role in the regulation of tumor-immune interaction. However, with the advent of the single-cell RNA sequencing technology, it is now possible to overcome these limitations. Using this technology, Schelker et al.¹⁴ determined gene expression profiles for tumor-infiltrating immune cells, tumor-associated non-malignant cells, and individual tumor cells from the same solid tumor biopsy. These profiles were used to benchmark deconvolution, and the results were validated using independent data. Single cell sequencing involves very complex procedures and logistics, and it is difficult to envisage that it will be possible in the context of routine clinical

practice. However, we expect that more gene matrices will become available from single cell sequencing experiments and will be instrumental in improving the potential of deconvolution on bulk transcriptomic data, which are and will become increasingly available in clinical oncology.

One additional caveat on the reliability of gene matrices merits attention: all methods of cell isolations have limits. Flow-cytometry cell separation procedures, even if they are performed on circulating cells, involve several steps of cell-surface markers labeling and physical sorting that may influence gene expression. Single cell sequencing requires an aggressive tissue disaggregation procedure that can damage the cells and lead to a selective loss of the most susceptible ones. This will introduce a bias on the surviving cells, in addition to other possible gene expression changes in response to the chemical and physical insults in the surviving cells.

A groundbreaking and rapidly growing family of techniques is spatial sequencing, which allows spatial detection of transcripts in tissue sections.¹⁵⁻¹⁷ This approach is going to be very useful, also in this context, for two reasons. First, RNA sequencing from fresh frozen tissue slices with a spatial barcoding will avoid any isolation procedure that could introduce an expression or a cell selection bias. Second, the topography of tumor-associated immune cells is also relevant to define the immune contexture. In fact, some tumors show pronounced immune infiltrates in their core (“inflamed” or “hot” tumors) while in some tumors immune cells aggregate at the tumor boundaries (“immune-excluded” tumors).¹⁸ Currently, the number of transcripts that can be sequenced (a few hundred) or the spatial resolution (10–100 microns) is limited, but we can envisage that considerable progress will be made in the near future to allow for a complete spatial transcriptomic analysis at the single-cell resolution level. As in the case of single-cell sequencing, this technology may not be easily implemented into clinical practice, but will be able to provide more detailed information on the gene expression characteristics of the single cell types, in association with their tissue localization. For instance, cells in immune-excluded tumors will probably have different expression patterns from cells in “inflamed” ones. These patterns will be useful for the generation of more detailed gene matrices for deconvolution.

In conclusion, we envisage that the complementation of more accurate basis matrices of cell-type and condition-dependent gene expression signatures with the available deconvolution algorithms will greatly improve the possibility of describing the cell components of the tumor microenvironment. This information will add to the existing immunohistochemistry and digital pathology assays.


Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Maddalena Fratelli  <https://orcid.org/0000-0002-1769-3427>

References

1. Hanahan D and Weinberg RA. Hallmarks of cancer: the next generation. *Cell* 2011; 144: 646–674.
2. Fridman WH, Pagès F, Sautès-Fridman C, et al. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* 2012; 12: 298–306.
3. Demaria S, Coleman CN and Formenti SC. Radiotherapy: changing the game in immunotherapy. *Trends Cancer* 2016; 2: 286–294.
4. Zheng W, Skowron KB, Namm JP, et al. Combination of radiotherapy and vaccination overcomes checkpoint blockade resistance. *Oncotarget* 2016; 7: 43039–43051.
5. Galon J and Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discov* 2019; 18: 197–218.
6. Mohammadi S, Zuckerman N, Goldsmith A, et al. A critical survey of deconvolution methods for separating cell types in complex tissues. *Proceedings of the IEEE*. 2017; 105: 340–366.
7. Avila Cobos F, Vandesompele J, Mestdagh P, et al. Computational deconvolution of transcriptomics data from mixed cell populations. *Bioinformatics* 2018; 34: 1969–1979.
8. Finotello F and Trajanoski Z. Quantifying tumor-infiltrating immune cells from transcriptomics data. *Cancer Immunol Immunother* 2018; 67: 1031–1040.
9. Vallania F, Tam A, Lofgren S, et al. Leveraging heterogeneity across multiple datasets increases cell-mixture deconvolution accuracy and reduces biological and technical biases. *Nat Commun* 2018; 9: 4735.
10. Abbas AR, Baldwin D, Ma Y, et al. Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 2005; 6: 319–331.
11. Abbas AR, Wolslegel K, Seshasayee D, et al. Deconvolution of blood microarray data identifies cellular activation patterns in systemic lupus erythematosus. *PLoS ONE* 2009; 4: e6098.
12. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015; 12: 453–457.
13. Monaco G, Lee B, Xu W, et al. RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep* 2019; 26: 1627–1640.e7.
14. Schelker M, Feau S, Du J, et al. Estimation of immune cell content in tumour tissue using single-cell RNA-seq data. *Nat Commun* 2017; 8: 2032.
15. Eisenstein M. Companies seek slice of spatial imaging market. *Nat Biotechnol* 2019; 37: 490–491.
16. Rodrigues SG, Stickels RR, Goeva A, et al. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* 2019; 363: 1463–1467.
17. Maniatis S, Åijö T, Vickovic S, et al. Spatiotemporal dynamics of molecular pathology in amyotrophic lateral sclerosis. *Science* 2019; 364: 89–93.
18. Kather JN, Suarez-Carmona M, Charoentong P, et al. Topography of cancer-associated immune cells in human solid tumors. *eLife* 2018; 7: e36967.