Bioinformatics, YYYY, 0–0
doi: 10.1093/bioinformatics/xxxxx
Advance Access Publication Date: DD Month YYYY
Manuscript Category

T Cell Exhaustion

Single-cell and bulk RNA sequencing of T cells for TOX for predicting immune checkpoint inhibition therapy response in cancer

Alexander Ying[1,*]

[1]Department of Biomedical Engineering, Columbia University, New York, NY, USA

*To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on 19 December 2022; revised on 19 December 2022; accepted on XXXXX

Abstract

Motivation: Immune checkpoint inhibition therapy is known to be an effective chemotherapy for multiple types of cancer. These treatments target protein receptors on the surface of exhausted T cells, but also induce a number of undesirable side effects. Thus, finding biomarkers for immune checkpoint inhibition therapy is an important step towards improving patient outcomes.

Results: *TOX* expression correlates to known markers of T cell exhaustion, and single-cell sequenced CD8+ T cell expression correlates with anti-PD-1 and combination treatment outcomes in melanoma patients. Bulk-RNA *TOX* expression does not predict anti-PD-1 and combination treatment outcomes in melanoma patients, but when outcomes are combined with time before recurrence or adverse effect, bulk-RNA *TOX* expression is associated with a lower risk for recurrence or adverse effects.

Contact: awy2110@columbia.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

T cell exhaustion is a condition that occurs across a broad spectrum of diseases. In this state, T cells express a greater number of checkpoint proteins such as *PD-1* and *CTLA-4* on their surface [1]. Immune checkpoint inhibition (ICI) therapies such as anti-PD-1 and anti-CTLA-4 are thus used across a wide range of cancers in order to target these proteins and prevent T cell deactivation [2]. Due to the risks of these treatments, determining a patient's suitability for ICI therapies and the chances for improvement are important to raising the standard of care for these diseases [3].

Although PD-1 and CTLA-4 are known markers for T-cell exhaustion, determining the mechanistic causes for exhaustion has uncovered a number of additional genes that can possibly play a role. One of these is TOX [4,5,6]. TOX is known to be both a regulator for T cell differentiation as well as a critical upstream regulator of main exhaustion processes [4,5,6]. For this reason, investigating the role TOX plays from a clinical perspective will both likely bear fruitful results, as well as help reinforce the findings from cell and epigenetic studies.

In this study, we closely followed the procedure laid out by Kim and Park *et al*., which focused on transcription factors associated with promoting T cell exhaustion and exploring their clinical significance as it relates to anti-PD-1 treatment outcomes in cancer [7]. Specifically, exploring single-cell RNA sequenced (scRNA-seq) data, to find transcription factors upregulated in exhausted CD8+ T cell populations in melanoma and non-small cell lung cancer (NSCLC) patients. After confirming the up-regulation of TOX in exhausted CD8+ T cells, we then extended the clinical analysis performed in Kim and Park *et al.* by applying it to a sc-RNA and bulk RNA-sequenced population of melanoma patients that were then treated with either anti-PD-1 monotherapy or anti-PD-1 and anti-CTLA-4 combination immune checkpoint inhibitor therapy [8].

2 Methods

2.1 Preprocessing and differential expression analysis

We analyzed single-cell transcriptome data taken from 3 different datasets, 2 datasets being analyzed previously [7] and 1 dataset in which this pipeline was reapplied [8]. Beginning with the previously analyzed dataset, expression data was available for 17 patients with melanoma (GSE72056) [9] and 14 patients with NSCLC (GSE99254) [10]. These datasets were normalized already, but using different methodologies. The melanoma dataset was normalized using the base-2 log1p of the TPM [9]. Cells were also filtered according to TPM of known marker genes or if fewer than 1700 genes were expressed [9]. The NSCLC samples were normalized using the scran method [10] and thus left in log2 space. Cells were then filtered against high mitochondrial gene expression, and the TPM of known marker genes. This was then transformed into linear space before processing with Scanpy [11]. This resulted in the correct total of 4645 melanoma cells and 11769 NSCLC cells [7].

In order to more closely match the expression values published [7], these expression counts were re-log-transformed using Scanpy's log1p function (this was done despite the melanoma data most likely being log-transformed already, although Scanpy did not detect previous log-transformation). Genes expressed in fewer than 5% of all cells were also filtered from the dataset. The melanoma dataset was then further to only include cells previously annotated as T cells, and filtered using Louvain clustering [12] on 16 principal components to isolate cells belonging to clusters with a high proportion of CD8+ cells. This obtained a subset of 1066 cells compared to the 1072 cells obtained in Kim and Park *et al.* using Seurat [13]. The NSCLC dataset was also further filtered to only include cells labeled as "TTC", indicating "tumor cytotoxic T cell" [7]. This obtained the expected 2123 NSCLC samples.

The cells in each dataset were then partitioned by *PD-1* expression. The median of *PD-1* expression in each dataset was determined and samples expressing *PD-1* higher than or equal to the median were labeled as "*PD-1* high" while samples expressing *PD-1* lower than the median were labeled "*PD-1* low". We then tested the gene expression of every gene between the two groups with the Wilcoxon rank-sum test using Scanpy's rank_genes_groups function with Benjamini-Hochberg correction. These genes were then filtered by significance: adjusted p-value < 0.05 and < 0.001 for the melanoma dataset and NSCLC dataset, respectively. The different p-values were chosen due to the large discrepancy between the identified number of differentially expressed genes using Scanpy, and the results reported in Kim and Park *et al*. The differentially expressed genes were then further filtered by the mean expression level in each subset (*PD-1* high and *PD-1* low). Genes with a mean expression level in both subsets less than 0.693 or 2 for the melanoma or NSCLC datasets, respectively, were discarded. This resulted in a total of 64 differentially expressed genes in the melanoma dataset and 195 differentially expressed genes in the NSCLC dataset.

Since Kim and Park *et al.* obtained 175 and 92 differentially expressed genes for the melanoma and NSCLC datasets, respectively, this process was repeated for the NSCLC dataset using Seurat [13], due to the difference in genes identified in the melanoma set being susceptible to different approaches in clustering. The same process of filtering cells and genes, and subsetting the dataset by PD-1 expression was repeated. Then, marker genes were again found by the Wilcoxon rank-sum test using the FindMarkers function with a threshold of 0.05. This procedure identified 87 marker genes.

The cells were then visualized using UMAP projection – as opposed to t-SNE used in Kim and Park *et al.* – and colored by the expression of *PD-1*, *HAVCR2*, *CTLA4*, *TIGIT*, and *TOX* [14]. We also used violin plots to demonstrate the distribution of expression in the latter 4 marker genes in the *PD-1* high and *PD-1* low groups.

2.2 Analysis of scRNA-sequenced cohort of melanoma patients

To determine the predictive value of *TOX* as a biomarker for anti-PD-1 or combination treatment outcome, this pipeline was reapplied towards a scRNA-seq dataset taken from 13 melanoma patients before treatment (GSE189125) that included clinical outcomes for these patients' response to anti-PD-1 or combination treatment [8]. As with other datasets, this dataset was filtered and normalized before analysis. The data was filtered against disproportionate mitochondrial gene read counts, and against cells expressing too few or too many genes [8]. The normalization was performed using Seurat's NormalizeData function, and only 2000 genes identified by FindVariableFeature were included in the dataset [8]. To validate the cell type annotations, expression of *CD4*, *CD8A*, and *CD8B* were compared by cell type annotation, and from this CD8+ T cells and T/NK T cells were selected for further analysis, resulting in 2915 cells total for analysis.

The clinical data included 3 outcomes: durable clinical benefit (DCB), no durable benefit (NDB), and not evaluable (NE) [8]. The latter two outcomes were grouped together as the terminal event for survival analysis using Kaplan-Meier curve fit and Cox proportional-hazard modeling [15]. Patients were stratified by mean *TOX* expression, with patients expressing *TOX* greater than the top 70 percentile *TOX* expression marked as "High *TOX*" and patients expressing lower *TOX* marked as "Low *TOX*" [7].

The single-cell data was then grouped by patient clinical response (DCB vs NDB and NE as "responder" vs "non-responder", respectively) to anti-PD-1 or combination treatment, and the expression of *TOX* was compared between the two groups by the Wilcoxon rank-sum test. The log-fold change was also compared to the log-fold change between *TOX* expression groups of all other genes.

2.3 Analysis of bulk-RNA sequenced cohort of melanoma patients

Bulk analysis gave access to a much larger cohort of 53 patients with both *TOX* expression data (GSE186143) and clinical outcome data [8]. Again, patients were stratified by *TOX* expression, again at the top 70 percentile of *TOX* expression within the bulk-RNA group. Clinical outcomes were also again grouped in the same way, and then fit to the Kaplan-Meier survival curve and Cox proportional-hazard model [15]. The Wilcoxon rank-sum test was also used to compare *TOX* expression between the responder and non-responder groups. Finally, *TOX* expression was used as the parameter for a Poisson-distributed linear model trained on half the clinical dataset, and the predictive value of *TOX* for clinical outcome was evaluated using the area under the receiver operating characteristics curve (AUROC) [15].

3 Results

3.1 TOX is differentially expressed between *PD-1* high and *PD-1* low T cells

Compared to 175 differentially expressed genes identified by Kim and Park *et al.* in the melanoma dataset, 64 differentially expressed genes were found by our recreation of the methodology, of which 59 genes are conserved between the two findings. Compared to 92 differentially expressed genes identified by Kim and Park *et al*. in the NSCLC dataset, 195 differentially expressed genes were found using Scanpy and 87 differentially expressed genes were found using Seurat. Notably, the only difference in differentially expressed genes between the three sets was by omission. That is, every differentially expressed gene found by Kim and Park *et al.* was found by the Scanpy methodology, and every differentially expressed gene found by the Seurat methodology was found by Kim and Park *et al*. It is assumed that the difference in NSCLC results may be due to differences in implementation of the Wilcoxon rank-sum test between Scanpy and Seurat, as Seurat seems to include a stochastic process to reduce computation time. This could introduce some randomness that needs to be accounted for, and indicates an area in which further interrogation of the respective packages would be required. In any case, *TOX* was found as a differentially expressed gene in all methodologies between *PD-1* high and *PD-1* low CD8+ T cells.

Visualization of the known T cell exhaustion marker genes shows that these genes were similarly distributed across both the melanoma and NSCLC datasets (Fig 1 B, C), though the similarity and localization is much stronger in the NSCLC dataset as opposed to the melanoma dataset. Similarly, the expression of *TOX* as a function of subset was much more pronounced in the NSCLC dataset, with a log fold change of 1.12 ($p < 0.0001$) as opposed to 0.46 ($p < 0.05$).

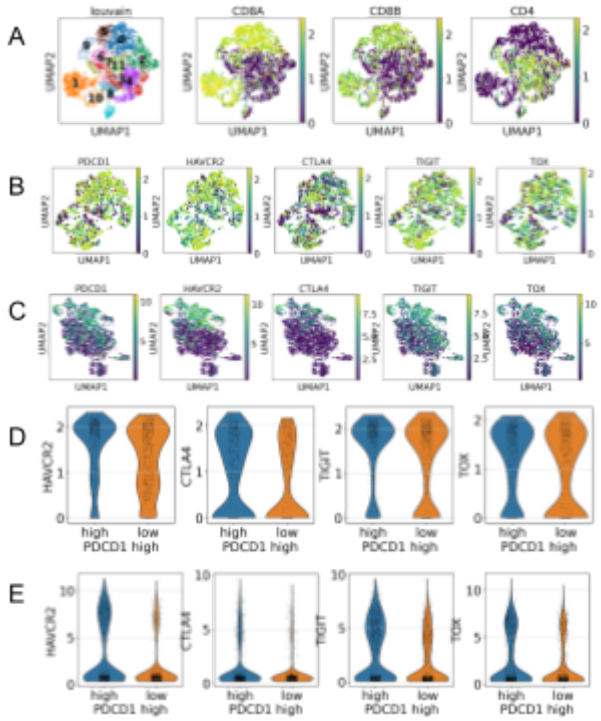

Fig. 1. Initial clustering and differential expression analysis results. a UMAP embedding of NSCLC cells colored by Louvain clusters, *CD8A* expression, *CD8B* expression, and *CD4* expression. **b-c** UMAP embedding of CD8+ T cells taken from melanoma and NSCLC patients, respectively, colored by marker gene expression. **d-e** Violin plots displaying expression distribution of T cell exhaustion marker genes in melanoma and NSCLC datasets, respectively.

3.2 *TOX* is differentially expressed between cells taken from responder and non-responder anti-PD-1 or combination treatment patients

Partitioning the single-cell RNA-seq data belonging to the 13 melanoma patients by clinical outcome did result in a significant difference in *TOX* expression (Fig 2.B). However, the magnitude of the log fold change was quite low (0.28), and 1423 genes were found to be significantly differentially expressed with a log fold change. Of these 1423 genes, *TOX* was ranked 455th, though this dataset included many spurious differentially expressed genes.

3.3 *TOX* expression did not a predict anti-PD-1 response in melanoma patients

Using the scRNA-seq data partitioning the patients by *TOX* expression did not result in a significant result when using the Cox proportional hazard model ($p = 0.70$). However, when using the bulk RNA-seq data, *TOX* expression does become a predictor for survival, with *TOX* high resulting in a 34% increase in risk for a remittance or a severe immune-related adverse effect. This suggests part of the reason for the non-significant survival result in the single-cell sequenced cohort is due to its low population.

Comparing patients solely on the basis of a remittance or immune related adverse effect, however, was less promising. Using the Wilcoxon rank-sum test between responders and non-responders resulted in no significant difference in *TOX* expression between groups ($p = 0.91$). Finally, the Poisson regression model failed to produce the promising results seen in Kim and Park *et al*., with the classifier achieving a 0.39 AUC (Fig 2 E). Overall, these results suggest that further work is needed to validate and improve *TOX* as a biomarker for immune checkpoint inhibitor success in patients, and that *TOX* alone does not predict successful outcomes.

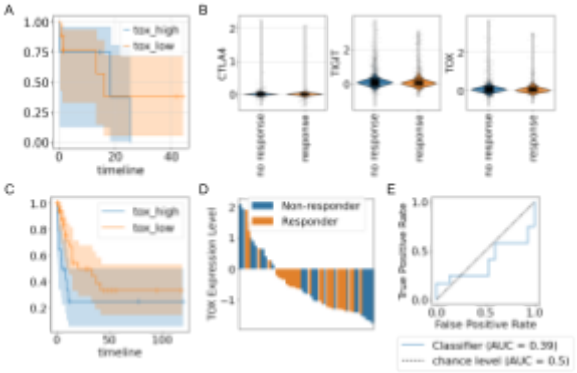

Fig. 2. Analysis of clinical outcomes in melanoma patients. a Kaplan-Meier survival plot of scRNA sequenced melanoma patients (n = 13) treated with anti-PD-1 or combination treatment, stratified by per-patient mean *TOX* expression in CD8+ T cells. The patients with mean *TOX* expression in the 70th percentile, and the rest were labeled as *TOX* low. **b** Violin plots displaying single-cell T cell exhaustion marker genes expression distribution in responding and non-responding patient populations. **c** Kaplain-Meier survival plot of bulk-RNA sequenced melanoma patients (n = 53) treated with anti-PD-1 or combination treatment, stratified by per-patient *TOX* expression. The patients in the 70th percentile of *TOX* expression were labeled as *TOX* high, and the rest were labeled as *TOX* low. **d** Waterfall plot depicting bulk-RNA *TOX* expression level colored by response to anti-PD-1 or combination therapy. Y-axis represents the Z-score of *TOX* expression, translated such that 0 is the 70th percentile of *TOX* expression. **e** Receiver operating characteristic curve with corresponding area under the curves for patient treatment response predictions from a Poisson linear regression model for bulk-sequenced melanoma patients (blue line) and random chance (dotted line). Note that model was trained on 50% of the bulk-sequenced melanoma patient population.

Full list of marker genes and respective p-values and log-fold changes can be found in the GitHub repository:
https://github.com/alexying2110/tcell_exh_melanoma_nsclc

Funding

This work has been supported by my roommate's coffee machine.

References

1. Wherry, E. J. & Kurachi, M. Molecular and cellular insights into T cell exhaustion. *Nat Rev Immunol* **15**, 486–499 (2015).
2. Khan, O. *et al.* TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. *Nature* **571**, 211–218 (2019).
3. Li, H., van der Merwe, P. A. & Sivakumar, S. Biomarkers of response to PD-1 pathway blockade. *Br J Cancer* **126**, 1663–1675 (2022).
4. Scott, A. C. *et al.* TOX is a critical regulator of tumour-specific T cell differentiation. *Nature* **571**, 270–274 (2019).
5. Alfei, F. *et al.* TOX reinforces the phenotype and longevity of exhausted T cells in chronic viral infection. *Nature* **571**, 265–269 (2019).
6. Khan, O. *et al.* TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. *Nature* **571**, 211–218 (2019).
7. Kim, K. *et al.* Single-cell transcriptome analysis reveals TOX as a promoting factor for T cell exhaustion and a predictor for anti-PD-1 responses in human cancer. *Genome Medicine* **12**, 22 (2020).
8. Lozano, A. X. *et al.* T cell characteristics associated with toxicity to immune checkpoint blockade in patients with melanoma. *Nat Med* **28**, 353–362 (2022).
9. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq | Science.
10. Guo, X. *et al.* Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* **24**, 978–985 (2018).
11. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15 (2018).
12. Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184–197 (2015).
13. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
14. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.
15. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).