Comprehensive evaluation of cell-type quantification methods for immuno-oncology

Gregor Sturm
Experimental Bioinformatics, Technical University of Munich
Pieris Pharmaceuticals GmbH, Freising

g.sturm@tum.de  github.com/grst  grst.github.io
Type and abundance of immune cells in the tumor microenvironment affect outcome.

|                | CD8⁺ T cells | TLS | Treg cell | M  | M1 | M2 |
|----------------|--------------|-----|-----------|----|----|----|
| Breast cancer  | ![Green](#)  | ![Green](#) | ![Green](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| Melanoma       | ![Green](#)  | ![Green](#) | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| Pancreatic cancer | ![Green](#) | ![Green](#) | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| NSCLC          | ![Green](#)  | ![Green](#) | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| Hepatocellular cancer | ![Green](#) | ![Green](#) | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| Ovarian cancer | ![Green](#)  | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| Head and Neck  | ![Green](#)  | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) | ![Red](#) |
| Bladder cancer | ![Green](#)  | ![Orange](#) | ![Red](#) | ![Red](#) | ![Red](#) | ![Red](#) |

Fridman, W. H., et al. (2017). Nature Reviews Clinical Oncology. doi:10.1038/nrclinonc.2017.101
Computational methods can estimate cell type abundance from bulk RNA-seq data.

Bulk RNA-seq data → Deconvolution → Cell type fractions

30%
50%
20%

Finotello et al. (2019). Genome Medicine, doi:10.1186/s13073-019-0638-6
But … which method should I use?
‘Deconvolution’, as opposed to ‘marker gene-based’ methods allow to compute cell fractions.

**Marker genes:** list of enriched genes for each cell type

**Deconvolution:** ‘inverse’ matrix multiplication with reference-profiles

---

Finotello et al. (2018). Cancer Immunology, Immunotherapy. doi:10.1007/s00262-018-2150-z
EPIC and quanTIseq are the only methods to compute cell fractions.

### Marker gene-based

| tool       | score      | between sample | between cell-type |
|------------|------------|----------------|------------------|
| MCP-counter| arbitrary units | ✓              | ✗                |
| xCell      | arbitrary units | ✓              | ✗                |

### Deconvolution-based

| tool           | score          | between sample | between cell-type |
|----------------|----------------|----------------|------------------|
| CIBERSORT      | immune cell fractions | ✗              | ✓                |
| CIBERSORT abs. | arbitrary units | ✓              | ✓                |
| EPIC           | cell fractions  | ✓              | ✓                |
| quanTIseq      | cell fractions  | ✓              | ✓                |
| TIMER          | arbitrary units | ✓              | ✗                |
FACS is a “gold standard” for comparing computational cell-type quantification methods.
FACS is a “gold standard” for comparing computational cell-type quantification methods.

Only 15 samples available!

Image credit: https://www.abcam.com/protocols/fluorescence-activated-cell-sorting-of-live-cells
Simulating bulk RNA-seq samples from single-cell data allows us to systematically assess the methods.

Schelker et al. (2017). Nature Communications, doi:10.1038/s41467-017-02289-3
Simulating bulk RNA-seq samples from single-cell data allows us to systematically assess the methods.

$11k$ single cells

Simulated bulk RNA-seq

$\text{correlation}$

$\text{background predictions}$

& detection limit

$\text{true}$

$\text{predicted}$

$\text{true}$

Schelker et al. (2017).
Nature Communications,
doi:10.1038/s41467-017-02289-3
Simulating bulk RNA-seq samples from single-cell data allows us to systematically assess the methods.

Schelker et al. (2017). Nature Communications, doi:10.1038/s41467-017-02289-3
Simulating bulk RNA-seq samples to assess...  

Correlation true vs. predicted
Simulating bulk RNA-seq samples to assess...

**Background predictions**

- S1
  - Dendritic cells
  - Macrophages
  - Melanoma cells

No CD4+ T cells!
Simulating bulk RNA-seq samples to assess...

**Background predictions**

- **S1**
  - Dendritic cells
  - Macrophages
  - Melanoma cells

- **No CD4+ T cells!**

**Minimal detection fraction**

- **S2, S3, S4, S5**

  - **CD4+ T cells**

  - **0% 5% 10%**

  - **true CD4+ T**

  - **predicted CD4+ T**
Simulating bulk RNA-seq samples to assess...

**Spillover**

S1  S2  S3  S4

- Macrophages
- CD4+ T cells
- CD8+ T cells
- Dendritic cells
Simulating bulk RNA-seq samples to assess...

**Spillover**

|   | S1 | S2 | S3 | S4 |
|---|----|----|----|----|
|   | ![Bar Chart](image) |

- **Macrophages**
- **CD4+ T cells**
- **CD8+ T cells**
- **Dendritic cells**

**Predicted cell-type**
We recommend EPIC and quanTIseq for general purpose deconvolution.
Beware of dendritic cell subtypes!

- m = myeloid
- d = monocyte-derived
- p = plasmacytoid
Background predictions are widespread among deconvolution-based approaches.
xCell is robust against background predictions (stat. enrichment test)

|      | B    | DC   | Mac/Mono | NK   | T CD4+ | T CD8+ | T CD4+ n.r. | T reg | CAF   | Endo |
|------|------|------|----------|------|--------|--------|-------------|-------|-------|-------|
| 0.0  | 0.0  | 0.0  | 0.0      | 0.0  | 0.0    | 0.0    | 0.0         | 0.0   | 0.0   | 0.0   |
| 0.1  | 0.1  | 0.1  | 0.1      | 0.1  | 0.1    | 0.1    | 0.1         | 0.1   | 0.1   | 0.1   |
| 0.2  | 0.2  | 0.2  | 0.2      | 0.2  | 0.2    | 0.2    | 0.2         | 0.2   | 0.2   | 0.2   |

**Diagram:**
- **Predicted** vs. **true**
- **Background predictions**
- **Minimal detection fraction**

**Graph:**
- **Fraction of spike-in cells**
- **Average estimate**
Removing genes with low cell-type specificity can reduce background predictions

**quanTIseq: Macrophage/Monocyte**

![Graph showing average estimate against fraction of spike-in cells](image)

**Mathematical formulation:**

\[
M = S_1 F_1 + S_2 F_2 + ... + S_N F_N
\]

for \( j = 1, ..., N \)
Removing genes with low cell-type specificity can reduce background predictions

quanTIseq: Macrophage/Monocyte

\[
M = \sum_{j=1}^{N} S_j F_{j1} + S_j F_{j2} + \ldots + S_j F_{jC} \quad \text{for } j = 1, \ldots, N
\]
Removing genes with low cell-type specificity can reduce background predictions

quanTIseq: Macrophage/Monocyte

\[ M_i = S_{ij}F_1 + S_{ij}F_2 + \ldots + S_{ij}F_{c} \quad \text{for} \quad j = 1, \ldots, N \]
Removing genes with low cell-type specificity can reduce background predictions.

Remove 5 non-specific genes.
Which method should I use?

- EPIC, quanTIseq (absolute scores, solid performance)
- MCP-counter (good for between-sample comparisons)
- xCell (no ‘background predictions’)

More observations

- We need signatures that address dendritic cell subtypes
- Background predictions can be addressed by identifying non-specific genes

Outlook

- More scRNA-seq data now available (200k+ cells)
- Cancer-type specific signatures?
Availability

- This talk ➔ grst.github.io/talks
- The paper ➔ Sturm et al. in the proceedings

Immunedeconv R package
Unified interface to methods

deconvolute(expr_mat, "epic")

github.com/grst/immunedeconv

Reproducible Pipeline
Reproduce entire benchmark

snakemake --use-conda

github.com/grst/immune_deconvolution_benchmark
Acknowledgements

- Florent Petitprez, Ligue contre le cancer, Paris
- Wolf H. Fridman, Cordeliers Research Centre, Paris
- Jitao David Zhang, Roche, Basel
- Jan Baumbach, Experimental Bioinformatics, TUM
We are hiring!

Pieris Pharmaceuticals
● zettl@pieris.com

Experimental Bioinformatics, TUM
● exbio@wzw.tum.de

Division of Bioinformatics, Medical University of Innsbruck
● francesca.fiotello@i-med.ac.at

Markus Zettl
Tatsiana Aneichyk
Markus List

Francesca Finotello
Supplementary Slides
### Table

#### Table C

|          | B     | DC    | Mono  | NK    | T     | T CD4+ | T CD8+ |
|----------|-------|-------|-------|-------|-------|--------|--------|
| CBA      | 0.89  | 0.25  | 0.58  | 0.93  | 0.77  | 0.74   | 0.32   |
| CBS      | 0.95  | 0.3   | 0.79  | 0.97  | 0.93  | 0.36   | 0.21   |
| EPC      | 0.95  | 0.87  | 0.97  | 0.98  | 0.72  | < 0    |        |
| MCP      | 0.83  | 0.88  | 0.57  | 0.98  | 0.95  | 0.72   | < 0    |
| QTS      | 0.93  | 0.45  | 0.7   | 0.96  | 0.97  | 0.75   | 0.48   |
| TMR      | 0.6   | 0.31  | 0.7   | < 0   | 0.93  | 0.57   |        |
| XCL      | 0.91  | 0.83  | 0.85  | 0.81  | 0.9   | 0.54   | 0.75   |
| All datasets (n=15) |       |       |       |       |       |        |        |

#### Table D

|          | B     | DC    | Mono  | NK    | T     |
|----------|-------|-------|-------|-------|-------|
| CBA      | 0.79  | n/a   | 0.58  | 0.98  | 0.77  |
| CBS      | 0.81  | n/a   | 0.79  | 0.99  | 0.93  |
| EPC      | 0.9   |       | 0.87  | 0.98  | 0.98  |
| MCP      | 0.89  | 0.87  | 0.57  | 0.99  | 0.95  |
| QTS      | 0.73  | 0.55  | 0.7   | 0.99  | 0.97  |
| TMR      | 0.73  | 0.71  |       | 0.99  | < 0   |
| XCL      | 0.85  | 0.94  | 0.85  | 0.99  | 0.9   |

### Notes

- **Table C** represents data from various datasets with the number of datasets specified as (n=15).
- **Table D** represents data from Hoek (PBMC, n=8) with the following observations:
  - CBA: 0.79
  - CBS: 0.81
  - EPC: 0.9
  - MCP: 0.89
  - QTS: 0.73
  - TMR: 0.73
  - XCL: 0.85
| Cell type     | Recommended methods | Overall perf. | Abs. score | No background predictions |
|---------------|---------------------|---------------|------------|---------------------------|
| B             | EPIC                | ++            | ++         | +                         |
|               | MCP-counter         | ++            | -          | -                         |
| T CD4+        | EPIC                | ++            | ++         | -                         |
|               | xCell               | ++            | -          | ++                        |
| T CD4+ n.r.   | quantITseq          | +             | ++         | +                         |
|               | xCell               | +             | -          | ++                        |
| T reg.        | quantITseq          | ++            | ++         | -                         |
|               | xCell               | ++            | -          | ++                        |
| T CD8+        | quantITseq          | ++            | ++         | -                         |
|               | EPIC                | ++            | ++         | -                         |
|               | MCP-counter         | ++            | -          | -                         |
|               | xCell               | +             | -          | ++                        |
| NK            | EPIC                | ++            | ++         | +                         |
|               | MCP-counter         | ++            | -          | -                         |
| Mac/Mono      | xCell               | -             | ++         |                           |
|               | EPIC                | +             | ++         | +                         |
|               | MCP-counter         | ++            | -          | -                         |
| CAF           | EPIC                | ++            | ++         | +                         |
|               | MCP-counter         | ++            | -          | -                         |
| Endo          | EPIC                | ++            | ++         | +                         |
|               | xCell               | ++            | -          | ++                        |
| DC            | None of the methods can be recommended to estimate overall DC content. MCP-counter and quantITseq can be used to profile myeloid DCs. |
| dataset/method | subtype                  | reference                                                                                   |
|---------------|-------------------------|---------------------------------------------------------------------------------------------|
| Schelker\(^1\) | plasmacytoid DC         | Identified in the single cell data using CD123 and CD303 marker genes\(^1\) which are pDC marker genes according to \(^6\). |
| Hoek\(^3\)    | myeloid DC              | primary human myeloid DC according to annotation on GSE64655                              |
| MCP-counter\(^9\) | myeloid DC              | signature explicitly annotated as myeloid DC                                                |
| CIBERSORT\(^10\) | monocyte-derived DC     | “Monocytes isolated as above were cultured in RPMI with 10% heat-inactivated FBS, 1 × Pen/Strep, 2 mM L-glutamine, 10 mM HEPES, 1 mM sodium pyruvate, then differentiated into dendritic cells by 17 ng/ml IL4, and 67 ng/ml GMCSF for 5 days at 5 × 10^6 cells/ml.” (GSE22886)\(^11\) |
| quantIseq\(^5\) | myeloid DC              | signatures derived from Hoek\(^3\) data                                                     |
| EPIC\(^4\)     | (no DC signature provided) |                                                                                             |
| TIMER\(^12\)   | monocyte-derived DC     | training data is a mix of various monocyte-derived DCs from HPCA (See table S8 of \(^12\)) |
| xCell\(^13\)   | myeloid DC              | uses a combination of various, mostly myeloid, DC samples (personal communication with authors) |
Certain cell-types are susceptible to spillover.
Certain cell-types are susceptible to spillover
Spillover occurs between NK and CD8+ T cells
Spillover occurs between CD4+ and CD8+ T cells
Spillover occurs between DCs and B cells
What causes spillover between DC and B cells?

|                          | Simulated sample (single cell) | Pure sample (FACS) |
|--------------------------|--------------------------------|--------------------|
| CD4+ T ↔ CD8+ T          | ✓                              | ✓                  |
| NK ↔ CD8+ T              | ✓                              | ✓                  |
| DC ↔ B                   | ✓                              | ✗                  |
