

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-------------------------------------|--|
| n/a | Confirmed |
| ☐ | ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ A description of all covariates tested |
| ☐ | ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☐ | ☒ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Cellranger v2.02 (10X Genomics) was used to align and generate expression matrices for downstream analysis.
Data analysis	Cell hashing reads were aligned using the CITE-seq count tool v1.4 (https://github.com/Hoohm/CITE-seq-Count). Scanpy v1.4 was used to perform analysis of single cell data. scikit-learn v0.20.3 was used to perform regression analysis and assess classifier performance. CIBERSORT source code was obtained from the original publication (https://cibersort.stanford.edu/index.php) and used as is. MetaIntegrator v2.00 was used to performed meta-analysis. bedtools v2 and Homer v4.1 were used to analyze ATAC-seq data. edgeR v3.1 was used to perform differential expression and peak accessibility analysis.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

scRNA-seq data is available for visualization and secondary analysis through the Broad Institute Single Cell Portal (https://singlecell.broadinstitute.org/single_cell): SCP548 (patients PBMCs) and SCP550 (bone marrow stimulation). Expression matrices are also deposited in SCP. Raw sequencing data are available for download from the controlled-access data repository, Broad DUOS (<https://duos.broadinstitute.org/>)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculations were performed. Sample size was determined to be adequate based on the degree and consistency of differences between groups. 1. Single cell RNA-sequencing and cell state identification - A total of 65 patients were profiled. Cells from all study subjects were used to determine cell states. 2. Sorting and FACS Analysis of MS1 - a number of HC, UTI, and URO patients were ran with the MS1 panel based on the availability of aliquots of frozen samples. 3. Publicly available datasets - the number of datasets and samples included therein were used as outlined in Sweeney, et. al. 2017 and Sweeney, et. al. 2018.
Data exclusions	Datasets outlined in Sweeney, et. al. 2018 for which the processed gene expression data matrix was not publicly available were not analyzed.
Replication	No assays were repeated on the same patient sample due to limited availability of cells.
Randomization	Single cell RNA-sequencing - For the primary cohort, patient samples were batched such that each multiplex scRNA-seq has 1/2 samples from UTI or URO and 1/2 HC. For the secondary cohort, patient samples were processed in the order they were received.
Blinding	Single cell RNA-sequencing - cells from all patients were initially analyzed blind to which patient they originated from to define the 16 cell states.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
☐	☒ Antibodies	☒	☐ ChIP-seq
☒	☐ Eukaryotic cell lines	☐	☒ Flow cytometry
☒	☐ Palaeontology	☒	☐ MRI-based neuroimaging
☒	☐ Animals and other organisms		
☐	☒ Human research participants		
☒	☐ Clinical data		

Antibodies

Antibodies used	CD3-APC (HIT3a), CD19-APC (HIB19), CD20-APC (2H7), CD56-APC (5.1H11), CD14-FITC (M5E2), CD16-AF700 (B73.1), CD45-PE-Cy7 (HI30), HLA-DR-PE (L243), CD45-AF700 (HI30), HLA-DR-PE-Cy7 (L243) and cell hashing antibodies (TotalSeq-A) were obtained from BioLegend. IL1R2-PE (34141) was obtained from ThermoFisher Scientific
Validation	For FACS analysis, the antibodies were used as prescribed by the manufacturer (5 uL per 100 uL of cells). Antibodies used are validated with immunofluorescent staining with flow cytometry analysis as indicated by the manufacturer. For cell hashing and scRNA-seq, 0.2 ug of each HTO antibody was used per sample, the concentration was determined by titrating the antibody with scRNA-seq analysis of pooled PBMC samples.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	Median ages of the clinical cohorts range from 53-71. 49% of the the individuals profiled are male, and 86% are white, non-Hispanic. Further details on disease status, age, gender, and race of all participants in this study are provided in Supplementary Table 1.
----------------------------	--

Recruitment

Primary cohort- patients entering the Emergency Department at the Massachusetts General Hospital were screened for inclusion criteria pertinent to a urinary tract infection, with and without organ damage.

Secondary cohorts - all patients entering the Medical Intensive Care Unit at the Brigham and Women's Hospital were recruited in the study before a diagnosis. Patients in the hospital wards are screened for positive bacteraemia cultures.

Enrollment in the cohorts were based on voluntary participation or through surrogate consent, no selection biases were present.

Ethics oversight

Research on the samples was approved by the Institutional Review Boards at the Broad Institute of MIT and Partners Health Care, (MA, USA).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Samples were from cryopreserved PBMCs (Methods)

Instrument

All FACS analysis were performed on a Sony SH800

Software

Data were analyzed using FlowJo v10.2

Cell population abundance

The post-sort purity of the sorted CD45+ and dendritic cells were not measure via flow cytometry. Based on scRNA-seq profiles, the sort purities are > 90% for CD45+ cells and > 70% for dendritic cells.

Gating strategy

Gating strategies are described in Supplementary Figure 3, Figure 3f, and Figure 4a.

- ☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.