SOFTWARE Open Access

CHARTS: a web application for characterizing and comparing tumor subpopulations in publicly available single-cell RNA-seq data sets



Matthew N. Bernstein¹, Zijian Ni², Michael Collins¹, Mark E. Burkard^{3,4}, Christina Kendziorski^{5*} and Ron Stewart^{1*}

*Correspondence: kendzior@biostat.wisc.edu; rstewart@morgridge.org ¹ Morgridge Institute for Research, Madison, WI 53715, USA⁵ Department of Biostatistics and Medical Informatics, University of Wisconsin - Madison, Madison, WI 53792, USA Full list of author information is available at the end of the article

Abstract

Background: Single-cell RNA-seq (scRNA-seq) enables the profiling of genome-wide gene expression at the single-cell level and in so doing facilitates insight into and information about cellular heterogeneity within a tissue. This is especially important in cancer, where tumor and tumor microenvironment heterogeneity directly impact development, maintenance, and progression of disease. While publicly available scRNA-seq cancer data sets offer unprecedented opportunity to better understand the mechanisms underlying tumor progression, metastasis, drug resistance, and immune evasion, much of the available information has been underutilized, in part, due to the lack of tools available for aggregating and analysing these data.

Results: We present CHARacterizing Tumor Subpopulations (CHARTS), a web application for exploring publicly available scRNA-seq cancer data sets in the NCBI's Gene Expression Omnibus. More specifically, CHARTS enables the exploration of individual gene expression, cell type, malignancy-status, differentially expressed genes, and gene set enrichment results in subpopulations of cells across tumors and data sets. Along with the web application, we also make available the backend computational pipeline that was used to produce the analyses that are available for exploration in the web application.

Conclusion: CHARTS is an easy to use, comprehensive platform for exploring single-cell subpopulations within tumors across the ever-growing collection of public scRNA-seq cancer data sets. CHARTS is freely available at charts.morgridge.org.

Background

Over the past three decades, the cancer research community has amassed large quantities of gene expression data from tumors. The premier example of such data was generated by The Cancer Genome Atlas [4], which generated bulk RNA-seq and microarray data from thousands of tumors across dozens of cancer types. These data have enabled a greater understanding into the molecular biology of cancer and have revealed great heterogeneity not only between cancer types, but also between tumors of the same



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 2 of 9

cancer type [2]. Unfortunately, investigations utilizing this resource are limited by the fact that gene expression was profiled using bulk methods, which measure gene expression on average across thousands, or tens of thousands, of cells in a sample. With the advent of single-cell RNA-seq (scRNA-seq), investigators are now able to measure gene expression at the single-cell level thereby gaining access to the substantial heterogeneity of cells within a tumor and the tumor microenvironment [9]. Publicly available scRNA-seq cancer data sets offer unprecedented opportunity to better understand the mechanisms of tumor progression, metastasis, drug resistance, and immune evasion. However, analyzing these data in the aggregate is challenging, especially for those without strong computational skills. To this end, easy-to-use web-based tools are important for enabling the broader research community to perform integrative analyses and, in doing so, to increase their ability to leverage their knowledge and comprehensively examine scientific and/or clinically relevant hypotheses in multiple data sets.

While a few web-based tools for analyzing scRNA-seq data are available, they are not designed specifically for cancer research or do not easily enable exploration of existing public data sets. For example, recent tools such as Alona [7] and Granatum [29] enable scRNA-seq analysis in the web browser; however, these tools are not cancer-specific and therefore do not enable important cancer-specific tasks such as classifying cells as being either transformed malignant cells or untransformed cells of the tumor microenvironment. Furthermore, these tools do not enable exploration of preprocessed, publicly available scRNA-seq data sets. Another tool, GREIN [16], enables exploration of public gene expression data, but it is neither single-cell specific nor cancer-specific and, consequently, does not implement features necessary for single-cell analysis such as cell type identification, clustering, or gene set enrichment, nor does it implement cancer-specific analyses such as malignancy classification. CancerSEA [27] enables exploration of gene set enrichment scores for gene sets pertaining to cancer-related processes, but does not enable visualization, differential expression, or cell type identification. In short, while web-based tools exist for exploring expression data, most do not allow for detailed analysis of scRNA-seq data across diverse tumors and data sets.

To address this gap, we present CHARacterizing Tumor Subpopulations (CHARTS), a web application and associated computational pipeline for analyzing and characterizing publicly available cancer scRNA-seq data sets. As described in detail below, for each tumor in its database, CHARTS identifies clusters and enables exploration via interactive dimension-reduction methods. Derived clusters are annotated with cell types from the Cell Ontology [1] via CellO [3], with information provided on the probability of the specific cell type as well as its ancestors. For example, the data may provide substantial evidence to classify cells within a cluster as T cells, but less evidence may be available to classify cells into more specific functional groups (e.g. helper or memory T cells). In addition, for each cluster within each tumor, enrichment of genes involved in biological processes and pathways is provided. Genes that are differentially expressed between the cluster and others are also available. Finally, CHARTS can be used to distinguish malignant vs. non-malignant cells allowing for precise exploration into the interactions between cell subpopulations within the tumor microenvironment. CHARTS currently enables exploration of 198 tumors across 15 cancer types, and data is being continually added. CHARTS is freely available at charts.morgridge.org.

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 3 of 9

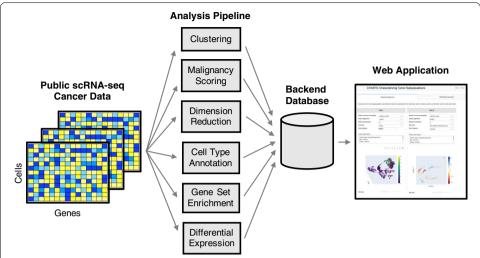


Fig. 1 Overview. A schematic diagram of the CHARTS pipeline. Public scRNA-seq data sets are collected and analyzed with a custom pipeline. This pipeline computes clusters, malignancy scores, dimension reduction transformations, cell type annotations, gene set enrichment scores, and differentially expressed genes for each cluster. Results are stored in a backend database and are accessed from the frontend web application

Implementation

Data set selection and preprocessing

We used the curated database describing single-cell RNA-seq data sets by Svensson et al. [22] to identify single-cell cancer data sets that are publicly available in the Gene Expression Omnibus (GEO) [6]. We selected all studies that sequenced primary tumor samples (i.e., non-cell line and non-xenograft samples) that consist of cells from the full tumor microenvironment (i.e., that do not select for a specific cell type). For each study, we wrote a custom script for separating cells by their tumor and normalizing the data by estimating the transcripts per million (TPM) for each gene and then computing log(TPM+1). We note that for droplet-based assays that sequence transcripts only from their end, such as Chromium 10x, an estimate of each gene's TPM in a cell can be obtained by dividing each gene's UMI count by the total UMI count in the cell and multiplying by one million. By estimating TPM, we measure gene expression using consistent units across assays. We exclude samples for which their corresponding GEO entry does not include enough information to estimate TPMs. Specifically, we exclude data sets originating from full length assays (i.e., where reads can originate from anywhere along the full length of the transcript), such as Smart-seq2 [21], if their GEO entry either does not include estimates of TPM, or includes counts, but does not include an estimate of each gene's expected length, which is required for estimating TPM. This process selected 198 tumors across 18 studies comprising a total of 259,488 cells.

These datasets were then processed with an offline computational pipeline. This pipeline implements a number of analyses in order to enable comprehensive characterization and comparison of tumor subpopulations within and between tumors (Fig. 1). All analysis output is stored in a backend database, which is quickly and easily accessible to a user through a frontend web application.

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 4 of 9

Dimension reduction

A user may construct interactive dimension-reduction scatterplots using two or three-dimensional UMAP [17] or PHATE [18] results. Each cell can be colored by the expression of a user-specified gene, cluster, malignancy score, cell type, or gene set enrichment. Two scatterplots are placed side-by-side enabling users to compare two characteristics (e.g. two different genes' expression values or a gene's expression value and the predicted cell types) within the same tumor or to compare a single characteristic (e.g. a single gene's expression values) between two tumors. For computational efficiency, we perform UMAP on the first 50 principal components after running principal components analysis (PCA).

Clustering

Clustering is performed using the Leiden community detection algorithm [25], as implemented in Python's Scanpy library [26]. Determining the optimal clustering is a challenging open problem in single-cell RNA-seq analysis [11]. If clustering is too coarse, multiple cell types may be erroneously combined into a single cluster. Similarly, if clustering is too fine, homogenous populations of cells may be broken into multiple clusters. Rather than choose a single clustering granularity for all tumors, CHARTS provides clusterings at three levels of granularity via three values for Leiden's resolution parameter (0.5, 1.0, and 2.0). Clustering is performed on the first 50 principal components after running PCA. The web interface enables a user to select a level of clustering granularity for exploring each tumor.

Cell type annotation

For each clustering granularity, each cluster is annotated with cell types from the Cell Ontology [1] via CellO [3]. The Cell Ontology is a hierarchically structured knowledge-base of known cell types. Specifically, the Cell Ontology forms a directed acyclic graph (DAG) where edges in the graph represent "is a" relationships. Because of this DAG structure, each cell is assigned to a specific cell type as well as all ancestors of this specific cell type within the DAG. CellO was executed using the isotonic regression correction algorithm. CHARTS exposes both CellO's binary cell type decisions for each cell type as well as CellO's estimated probability that each cell is of a given type.

Gene set enrichment

Each cluster's mean gene expression profile is scored for enrichment of gene sets describing molecular processes. Specifically, CHARTS uses GSVA [10] to score each cluster for enrichment of gene sets in the hallmark gene set collection from MSigDB [14] and the gene set collection used by CancerSEA. Specifically, we treat the TPM expression measurements as being distributed according to a log-normal distribution [8] and use the Gaussian kernel in the GSVA algorithm.

Malignancy status

Each cell is assigned a malignancy score that describes the likelihood that the cell is malignant. The malignancy scoring approach builds upon the approaches used by Tirosh Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 5 of 9

et al. [24] and Couturier et al. [5] for classifying cells as either transformed, malignant cells or untransformed cells within the tumor microenvironment (Additional file 1: Supplementary Methods).

Differential expression and cluster comparison

For each cluster within each tumor, CHARTS uses a Wilcoxon rank-sum test, as implemented in Scanpy, to compute the set of genes differentially expressed in the given cluster versus cells outside the cluster within the given tumor. CHARTS presents the top 50 genes ranked according to their significance. In addition to differential expression analysis, CHARTS also presents boxplots and violin plots for comparing the distribution of gene expression for a user-selected gene between clusters.

Results

Two case studies demonstrate how CHARTS can be used, both to examine and generate new hypotheses.

Case study: dysfunctional CD8 + T cells in lung adenocarcinoma

Investigators have recently reported a dysfunctional population of CD8+T cells in lung cancer [23] and melanoma [15] that express genes associated with immune suppression. In some melanoma samples, this population was also found to be highly proliferative [15]. We used CHARTS to explore whether this dysfunctional state was common across the majority of CD8+T cells, and to evaluate whether dysfunctional CD8+T cells were also highly proliferative. We used both CellO's classification results and CD8 expression to find the CD8+T cell population. We found that in the majority of lung adenocarcinomas, only a subset of CD8+T cells express marker genes for this dysfunctional state. Two adenocarcinomas from [13] are shown in Fig. 2. Using the gene set enrichment feature of CHARTS, we further found that dysfunctional cells are enriched for cell cycle genes, which may indicate that these dysfunctional CD8+T cells are highly proliferative in lung adenocarcinoma, as has been recently observed in melanoma.

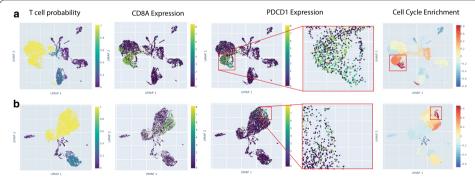


Fig. 2 Dysfunctional CD8+T Cells in Lung Adenocarcinoma. For lung adenocarcinoma tumors LX682 (**a**) and LX676 (**b**), we used CHARTS to visualize the probability that each cell is a T cell as well as expression of CD8A, expression of the dysfunctional CD8+T cell marker PDCD1, and each cell's enrichment score for genes in CancerSEA's cell cycle gene set as produced by GSVA

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 6 of 9

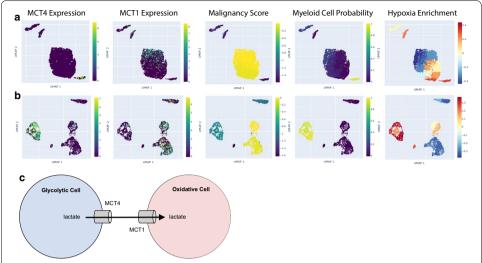


Fig. 3 Monocarboxylate Transporter Expression in Glioblastoma. For glioblastoma tumors PJ025 from Yuan et al. [28] (a) and MGH125 from Neftel et al. [19] (b), we used CHARTS to visualize the expression of MCT4 (gene symbol "SLC16A3" in the CHARTS application), the expression of MCT1 (gene symbol "SLC16A1" in the CHARTS application), malignancy score, the probability that each cell is a myeloid cell, and each cell's enrichment score for genes in the Hallmark hypoxia gene set as produced by GSVA. (**C**) A schematic illustration of the metabolic symbiosis between hypoxic, glycolytic tumor cells expressing MCT4 and oxidative tumor cells expressing MCT1

Case study: monocarboxylate transporters in glioblastoma

We investigated the expression of MCT4, a prognostic biomarker of glioblastoma aggression [12, 30]. Using CHARTS, we found that MCT4 tended to be expressed in the myeloid tumor-infiltrating immune cells. Two tumors, from Yuan et al. [28] and Neftel et al. [19], are shown in Fig. 3a and b respectively. While MCT4 is known to be involved in a metabolic symbiosis between hypoxic tumor cells, which express MCT4 to expel lactate, and oxidative tumor cells, which express MCT1 to intake lactate [20] (Fig. 3c), the specific cell types expressing MCT4 in glioblastoma have not been well characterized. We used CHARTS to determine which cells express MCT1 in glioblastoma and found that this gene was primarily expressed in cells with high malignancy scores (Fig. 3a, b). Using the gene set enrichment feature of CHARTS, we observed that cells expressing MCT1 tended to express genes enriched for hypoxia, whereas cells expressing MCT4 tended to express genes that were less enriched for hypoxia (Fig. 3a, b). This observation indicates a possible metabolic symbiosis between malignant cells and myeloid cells in the tumor microenvironment of glioblastoma, which to the best of our knowledge, has not been well characterized.

Conclusion

In this work, we present CHARTS: a comprehensive framework for exploring single-cell subpopulations within tumors and the tumor microenvironment across ever-growing data sets. CHARTS can be used to develop and explore new hypotheses underlying tumor progression, drug resistance, and immune evasion. Specifically, CHARTS exposes the results of an analysis pipeline executed on publicly available data from GEO.

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 7 of 9

There are a number of avenues that would benefit from further investigation. First, we note that CHARTS presents the results of a static analysis pipeline that was executed on diverse single cell datasets. Thus, individual datasets may benefit from tweaking the pipeline's parameters on an individual basis (e.g., parameters for clustering or dimensionality reduction). Future work will entail devising methods for fine-tuning the parameters used to process each dataset.

We note that this analysis pipeline was executed on published gene expression matrices and did not involve processing the raw sequencing reads. Thus, there may be variability between data sets due to the varying methods employed by the data submitters for alignment, gene expression quantification, and quality control. Future work will seek to remove some of this variability by starting with the raw reads from each dataset, rather than the expression matrices, and uniformly quantifying expression and filtering cells across datasets.

Availability and requirements

Project name: CHARTS

Project home page: https://charts.morgridge.org

Operating systems(s): The web application is platform independent. The backend

pipeline has been tested only on Linux and MacOS.

Programming language: Python

Other requirements: The backend pipeline requires Python 3.6 or higher.

License: MIT

Any restrictions to use by non-academics: None

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12859-021-04021-x.

Additional file 1. Supplemental Methods. Description and benchmarking of malignancy score.

Abbreviations

scRNA-seq: Single-cell RNA sequencing; GEO: Gene Expression Omnibus; DAG: Directed acyclic graph; TPM: Transcripts per million.

Acknowledgements

Not applicable.

Authors' contributions

MNB led the conceptual design and implementation of the project and also wrote the manuscript. MNB and ZN developed the malignancy scoring technique. MC provided technical support and set up the web server. MEB provided input as well as biological and medical expertise. CK and RS supervised the project. All authors read and approved the final manuscript.

Funding

M.N.B. acknowledges support of a postdoctoral fellowship provided by the Morgridge Institute for Research. Z.N. and C.K. were funded by the NIHGM102756. M.E.B. is funded by CA234904. Z.N. also acknowledges support provided by the Morgridge Institute for Research. R.S. acknowledges a grant from Marv Conney. The funders had no role in study design, data collection and analysis, or preparation of the manuscript.

Availability of data and materials

Code implementing the web application and offline data analysis pipeline is available at https://github.com/stewart-lab/CHARTS.

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 8 of 9

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Morgridge Institute for Research, Madison, WI 53715, USA. ² Department of Statistics, University of Wisconsin - Madison, Madison, WI 53706, USA. ³ Department of Medicine, Hematology/Oncology, University of Wisconsin - Madison, MI 53705, USA. ⁴ University of Wisconsin Carbone Cancer Center, Madison, WI 53705, USA. ⁵ Department of Biostatistics and Medical Informatics, University of Wisconsin - Madison, Madison, WI 53792, USA.

Received: 5 October 2020 Accepted: 11 February 2021

Published online: 23 February 2021

References

- Bard J, Rhee SY, Ashburner M. An ontology for cell types. Genome Biol. 2005. https://doi.org/10.1186/ ab-2005-6-2-r21
- 2. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. Nature. 2013;501(7467):355–64.
- 3. Bernstein MN, Ma Z, Gleicher M, Dewey CN. CellO: comprehensive and hierarchical cell type classification of human cells with the cell ontology. iScience. 2020;24(1):101913.
- Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Mills Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The cancer genome atlas pan-cancer analysis project. Nat Genet. 2013;45(10):1113–20.
- Couturier CP, Ayyadhury S, Le PU, Nadaf J, Monlong J, Riva G, Allache R, et al. Single-cell RNA-Seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. Nat Commun. 2020;11(1):3406.
- Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207–10.
- 7. Franzén O, Björkegren JLM. Alona: a web server for single-cell RNA-Seq analysis. Bioinformatics. 2020;36(12):3910–2.
- 8. Gierliński M, Cole C, Schofield P, Schurch NJ, Sherstnev A, Singh V, Wrobel N, et al. Statistical models for RNA-Seq data derived from a two-condition 48-replicate experiment. Bioinformatics. 2015;31(22):3625–30.
- González-Silva L, Quevedo L, Varela I. Tumor functional heterogeneity unraveled by scRNA-Seq technologies. Trends Cancer Res. 2020;6(1):13–9.
- Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinform. 2013. https://doi.org/10.1186/1471-2105-14-7.
- 11. Kiselev VY, Andrews TS, Hemberg M. Challenges in unsupervised clustering of single-cell RNA-Seq data. Nat Rev Genet. 2019;20(5):273–82.
- 12. Lai S-W, Lin H-J, Liu Y-S, Yang L-Y, Lu D-Y. Monocarboxylate transporter 4 regulates glioblastoma motility and monocyte binding ability. Cancers. 2020. https://doi.org/10.3390/cancers12020380.
- 13. Laughney AM, Hu J, Campbell NR, Bakhoum SF, Setty M, Lavallée V-P, Xie Y, et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. Nat Med. 2020;26(2):259–69.
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417–25.
- Li H, van der Leun AM, Yofe I, Lubling Y, Gelbard-Solodkin D, van Akkooi ACJ, van den Braber M, et al. Dysfunctional CD8 T cells form a proliferative, dynamically regulated compartment within human melanoma. Cell. 2019:176(4):775-89.e18.
- 16. Mahi NA, Najafabadi MF, Pilarczyk M, Kouril M, Medvedovic M. GREIN: an interactive web platform for re-analyzing GEO RNA-Seq data. Sci Rep. 2019;9(1):7580.
- 17. McInnes, L., J. Healy, and J. Melville. 2015. UMAP: uniform manifold approximation and projection for dimension reduction. arXiv:1802.03426.
- 18. Moon KR, van Dijk D, Wang Z, Gigante S, Burkhardt DB, Chen WS, Yim K, et al. Visualizing structure and transitions in high-dimensional biological data. Nat Biotechnol. 2019;37(12):1482–92.
- 19. Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ, Richman AR, et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. Cell. 2019;178(4):835-49.e21.
- Payen VL, Mina E, Van Hée VF, Porporato PE, Sonveaux P. Monocarboxylate transporters in cancer. Mol Metab. 2020;33(March):48–66.
- 21. Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-Seq from single cells using Smart-seq2. Nat Protoc. 2014;9(1):171–81.
- Svensson V, da Veiga Beltrame E, Pachter L. A curated database reveals trends in single-cell transcriptomics. Database J Biol Databases Curation. 2020. https://doi.org/10.1093/database/baaa073.
- Thommen DS, Koelzer VH, Herzig P, Roller A, Trefny M, Dimeloe S, Kiialainen A, et al. A transcriptionally and functionally distinct PD-1 CD8T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. Nat Med. 2018;24(7):994–1004.
- 24. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-Seq. Science. 2016;352(6282):189–96.

Bernstein et al. BMC Bioinformatics (2021) 22:83 Page 9 of 9

- Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019. https://doi.org/10.1038/s41598-019-41695-z.
- Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15.
- 27. Yuan H, Yan M, Zhang G, Liu W, Deng C, Liao G, Xu L, et al. CancerSEA: a cancer single-cell state atlas. Nucleic Acids Res. 2019;47(D1):D900-908.
- 28. Yuan J, Levitin HM, Frattini V, Bush EC, Boyett DM, Samanamud J, Ceccarelli M, et al. Single-cell transcriptome analysis of lineage diversity in high-grade glioma. Genome Med. 2018;10(1):57.
- 29. Zhu X, Wolfgruber TK, Tasato A, Arisdakessian C, Garmire DG, Garmire LX. Granatum: a graphical single-Cell RNA-Seq analysis pipeline for genomics scientists. Genome Med. 2017;9(1):108.
- 30. Zuo S, Zhang X, Wang L. A RNA sequencing-based six-gene signature for survival prediction in patients with glioblastoma. Sci Rep. 2019;9(1):2615.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

