


REVIEW

Open Access

What are the applications of single-cell RNA sequencing in cancer research: a systematic review



Lvyuan Li^{1,2}, Fang Xiong³, Yumin Wang^{2,3}, Shanshan Zhang³, Zhaojian Gong⁴, Xiayu Li⁵, Yi He¹, Lei Shi⁴, Fuyan Wang², Qianjin Liao¹, Bo Xiang^{1,2}, Ming Zhou^{1,2}, Xiaoling Li^{1,2}, Yong Li⁶, Guiyuan Li^{1,2}, Zhaoyang Zeng^{1,2}, Wei Xiong^{1,2*}  and Can Guo^{1,2*}

Abstract

Single-cell RNA sequencing (scRNA-seq) is a tool for studying gene expression at the single-cell level that has been widely used due to its unprecedented high resolution. In the present review, we outline the preparation process and sequencing platforms for the scRNA-seq analysis of solid tumor specimens and discuss the main steps and methods used during data analysis, including quality control, batch-effect correction, normalization, cell cycle phase assignment, clustering, cell trajectory and pseudo-time reconstruction, differential expression analysis and gene set enrichment analysis, as well as gene regulatory network inference. Traditional bulk RNA sequencing does not address the heterogeneity within and between tumors, and since the development of the first scRNA-seq technique, this approach has been widely used in cancer research to better understand cancer cell biology and pathogenetic mechanisms. ScRNA-seq has been of great significance for the development of targeted therapy and immunotherapy. In the second part of this review, we focus on the application of scRNA-seq in solid tumors, and summarize the findings and achievements in tumor research afforded by its use. ScRNA-seq holds promise for improving our understanding of the molecular characteristics of cancer, and potentially contributing to improved diagnosis, prognosis, and therapeutics.

Keywords: Single-cell RNA sequencing, Sequencing platform, Tumor, Specimen preparation, Data analysis

Background

Tumors are generally considered to be of monoclonal origin, with mutations facilitating the expansion of a malignant cell in the body to visible tumor tissue, as well as from carcinoma in situ to metastatic carcinoma. Notably, patient-derived tumor tissue includes cancer cells as well as other cell types, such as infiltrating immune cells and fibroblasts. Therefore, the use of traditional bulk RNA sequencing technology can only permit an

average qualitative characterization of such highly complex tissue, representing the average of cancer cells and non-cancer cells. Although bulk RNA approaches have made many invaluable contributions to medical science [1], such approaches have ignored the distinct phenotypic and functional traits of single cells within tumor samples.

Since the first single-cell RNA sequencing (scRNA-seq) study was published in 2009 [2], various commercial platforms and methods have been developed for scRNA-seq. ScRNA-seq is a technique allowing for the study of tissues at single-cell resolution. Many researchers within the life sciences employ scRNA-seq for the investigation of diverse biological functions. In

* Correspondence: xiongwei@csu.edu.cn; guocde@csu.edu.cn

¹NHC Key Laboratory of Carcinogenesis and Hunan Key Laboratory of Cancer Metabolism, Hunan Cancer Hospital and the Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China
Full list of author information is available at the end of the article



© 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.

particular, great achievements have been made in tumor research using scRNA-seq technology. The single-cell resolution afforded by scRNA-seq enables direct measurement of the transcriptional output of cells from tumor samples [3], comparison of differences between the transcriptomes of various cells, identification of rare cell subpopulations, such as heterogeneous tumor subpopulations [4], or the revelation of differences between stimulated dendritic cells [5], in turn providing unprecedented insights that have contributed to the development of cancer therapy.

In the present review, we introduce methods for the preparation of solid tumor samples, related scRNA-seq platforms, the analysis of scRNA-seq data, and achievements in tumor research facilitated by the use of scRNA-seq.

Single-cell RNA sequencing

To carry out scRNA-seq, solid tumor samples first need to be processed to effectively isolate viable single cells from the tissue of interest [6]. Thereafter, the single cells are lysed to obtain RNA, which is reverse transcribed into cDNA and then amplified to construct a sequencing library. The most suitable sequencing instrument must be selected based on the experimental scheme and research objectives. After sequencing, the data need to be correctly analyzed to reveal new findings.

Preparation of solid tumor specimens

Many published articles on the use of scRNA-seq in cancer research have detailed the preparation of solid tumor specimens. After cancer is diagnosed, tumor tissues are removed by biopsy and treated immediately to preserve single cells. Tumor tissue is usually cut into sections of approximately 1 mm³ and washed with PBS to remove fat, visible vessels, and surrounding necrotic areas [7]. Tissue is separated into single cells via grinding and filtration [8], using a Human Tumor Dissociation Kit [7, 9], or through other methods. Separated cells are then centrifuged and resuspended. All samples are stained with trypan blue to confirm viability, evaluate sample quality, and remove dead cells. To avoid introducing gene expression changes associated with processing, after washing and counting, cell suspensions are preserved in cold storage in preparation for cell sorting.

Selection of sequencing platform

Currently, major single-cell sequencing platforms include the 10X Genomics Chromium, Nadia (Dolomite Bio), Illumina Bio-Rad ddSEQ Single-Cell Isolator, BD Rhapsody Single-Cell Analysis System (BD), ICELL8 Single-cell System (Takara), Fluidigm C1, and others. Although new platforms for scRNA-seq are still being

developed, the most widely used platform remains the 10X Genomics Chromium.

Selecting the most suitable sequencing platform for your research project is a key step for achieving the desired research results. There are a number of published articles comparing performance, cost, and other aspects of existing platforms. In 2019, Zhang et al. [10] compared three widely used drop-based high-throughput scRNA-seq systems: inDrop, Drop-seq, and 10X Genomics Chromium, and found that the 10X Genomics Chromium had the highest sensitivity and the lowest technical noise. The 10X Genomics platform is therefore a more suitable choice for rare samples, such as human embryos, which require a more efficient cell capture platform. For abundant samples, Drop-seq may be more cost-effective, while InDrop is a better choice when a customized protocol is necessary. Natarajan et al. [11] compared the BGISEQ-500 platform with Illumina HiSeq and observed that these platforms were comparable with regard to sensitivity, accuracy, and repeatability. However, the sequencing cost of the former was lower. These comparative studies indicate that each sequencing platform has its own merits, and no sequencing platform is suitable for all research objectives. In summary, researchers have to choose the sequencing platform that can best meet their specific experimental needs.

Analysis of scRNA-seq data

Data generated during scRNA-seq usually gets processed via two analytical procedures: pre-processing (including quality control, batch-effect correction, and normalization) and downstream analysis (cell cycle phase assignment, clustering, reconstruction of cell trajectory and pseudo time, differential expression and gene set enrichment analysis, as well as gene regulatory network inference). The following section will briefly introduce the significance of and methods for these analytical procedures.

Quality control

Quality control should always be carried out as the first step for scRNA-seq data. Various methods for quality control have been developed [12–14]. Data obtained from poor quality cells (poor activity, containing degraded RNA, and zero expression levels of housekeeping genes such as GAPDH and ACTB) [6] should be excluded before subsequent bioinformatics analysis to mitigate its influence on downstream analysis results. The percentage of mitochondrial reads is a common quality control metric [15]. When there are a large number of mitochondrial transcripts, it means that the cells are in a state of stress [16], so a threshold is commonly applied to exclude data from cells with too many mitochondrial transcripts. Similarly, the proportion of

ribosomal reads is another commonly used quality control metric. Because scRNA-seq is mainly used to study functional (messenger) RNA, cells that have had their ribosomes removed and still have a high proportion of ribosome reads cannot be further analyzed [17]. In addition to the quality control methods for a single dataset, a method denoted ‘scRNABatchQC’ has been proposed that facilitates quality assessment across datasets to intuitively detect biases and outliers [18]. To aid researchers who are intimidated by scRNA-seq analysis, Etherington et al. [19] developed tools and training materials that can be used for scRNA-seq training and quality control.

Batch-effect correction

During the scRNA-seq experimental procedure, when cells subject to different conditions are cultured, captured, and sequenced separately, batch effects will be evident [20]. There are several methods available for batch-effect correction of scRNA-seq data, including Seurat 3 [21], MMD-ResNet [22], Harmony [23], Scanorama [24], Liger [25], scMerge [26], ZINB-WaVE [27], and others. Based on a variety of evaluation indicators, Harmony, Liger, and Seurat 3 are the recommended methods for dealing with batch effects, among which Harmony is the first choice due to its shorter run time [28]. Recently, a novel numerical algorithm for batch-effect correction of bulk and scRNA-seq data was proposed, denoted ‘scBatch’. This approach is not limited by the hypothesis of the batch-effect generation mechanism, and is superior to the benchmark batch-effect correction algorithms [29].

Normalization

Data normalization is essential for scRNA-seq to make gene expression comparable within and/or between samples. A number of methods have been developed for the normalization of RNA-seq data [30–33]. However, the majority of methods follow the same principle as bulk RNA-seq normalization and, thus, are not applicable to scRNA-seq data [30, 31]. Nevertheless, several methods have recently been devised to normalize scRNA-seq data, such as SCONE [34] and regularized negative binomial regression [35]. SCONE provides a flexible framework for users to choose appropriate normalization methods. Normalization using regularized negative binomial regression effectively eliminates technical differences due to different sequencing depth without inhibiting biological heterogeneity. A previous study [36] compared seven scRNA-seq data normalization methods with regard to reduction of noise or bias, and found that each of these methods was suitable to normalize specific types of data for further downstream analysis.

Cell cycle phase assignment

Determining the cell cycle phase of a single cell can facilitate understanding of biological processes such as tumorigenesis [37–41] and cell differentiation [42, 43], and avoid the confounding effects caused by the cell cycle phase prior to downstream analysis. Scialdone et al. [44] described and compared six supervised cell cycle prediction methods based on a cell transcriptome, of which the parameter-free PCA-based method and the custom predictor known as the “Pairs” method performed best in allocating cells to the correct cell cycle stage. Buettner et al. [45] proposed a calculation method, denoted ‘single-cell latent variable model’ (scLVM), which can be used to eliminate variations caused by cell cycle and other confounding factors before downstream analysis. Recently, Hsiao et al. [46] proposed a new method to characterize the progress of the cell cycle, which is different from the traditional classification of cells according to the standard of cell cycle stage (G1, S or G2/M phase), but can quantify the cell cycle progression of induced pluripotent stem cells on a continuum, which provides a basis for the characterization of the cell cycle in other cell types.

Cell clustering

One of the basic goals of scRNA-seq data analysis is to identify cell types from experimental samples to elucidate tissue complexity and heterogeneity. Due to the importance of cell type recognition, efforts have been made to develop new algorithms, including CountClust [47], CIDR [48], SIMLR [49], SAFE [50], and other advanced methods. A few studies [51–54] have compared and summarized diverse clustering algorithms for scRNA-seq data analysis. Unlike previous methods, Geddes et al. [55] proposed the first ensemble clustering framework based on autoencoder dimension reduction, which could be combined with different clustering algorithms to promote the accurate recognition of cell types. Since then, several new clustering methods have emerged, including DivBiclust [56], a biclustering-based framework, SAME [57], which extracts cluster solutions from multiple methods, and PARC [58], which is suitable for large-scale single-cell data. To overcome flaws associated with the manual labeling of cell types, Shao et al. [59] developed an automatic annotation toolkit, denoted ‘scCATCH’, based on clustering, which can accurately annotate cell types with acceptable repeatability.

Reconstruction of cell trajectory and pseudo-time

By reconstructing cell trajectory and pseudo-time based on scRNA-seq data, dynamic processes in cells can be calculated and simulated, which is of great significance for understanding the transition between cell states in cancer [60]. Currently available algorithms include

Monocle 2 [61], Monocle 3 [62], TSCAN [63], Slingshot [64], SLICE [65], LISA [66], p-creode [67], Waddington-OT [68] and others. Considering the exponential growth in the size of scRNA-seq data, Chen et al. [66] proposed an unsupervised method, denoted 'Lisa', for the reconstruction of cell trajectory and pseudo-time for a large number of scRNA-seq datasets. p-creode is another unsupervised algorithm that can predict cell state-transition trajectories. Waddington-OT uses the mathematical method of optimal transport (OT) to infer ancestor-descendant fate, and reconstruct cell trajectories. In addition to the three methods mentioned above, the algorithms Monocle 2 (the new version Monocle 3), TSCAN, and Slingshot have been shown to have good performance at reconstructing cell trajectory and pseudo-time [69].

Differential expression and gene set enrichment analysis

One of the most common uses of gene expression data is for the identification of differentially expressed (DE) genes under different experimental conditions (e.g., stimulated versus non-stimulated, mutant versus wild-type, or between different time points), and thus to determine the root cause of phenotypic differences observed under different conditions [70]. A zero value for a gene's expression level in scRNA-seq data may indicate two things. One is the "real" zero, caused by the changing characteristics of single-cell gene transcription, while the other is the "dropout" zero, caused by technical reasons, which often affects the validity of differential expression analysis. Miao et al. [71] developed the R package DESingle, which can accurately distinguish between the two types of zeros. DECENT is a DE gene analysis method based on UMI scRNA-seq data, and is used to analyze the pre-dropout distributions of inferred RNA molecules [72]. In addition to dropout zeros, another challenge in differential expression analysis of scRNA-seq data is multimodal data distribution. ZIAQ is the first approach to consider both dropout rates and the complex distributions of scRNA-seq data, which can be used to identify more DE genes [73].

We usually group DE genes according to their participation in common biological processes to facilitate the interpretation of results [74]. Existing gene set enrichment (GSE) analysis methods include DAVID [75], PAGE [76], CAMERA [77] and others, but almost all of these methods are more suitable for bulk RNA-seq analysis [78]. In addition, almost all existing GSE methods are used as a separate step after DE analysis. Considering the above shortcomings, Ma et al. [79] proposed IDEA, a computational method integrating DE analysis and GSE analysis for scRNA-seq, which could greatly improve the outcomes of both.

Gene regulatory network inference

The combination of active transcription factors and their target genes is usually described in gene regulatory networks (GRNs). Revealing these regulatory interactions is the goal of GRN inference methods, providing valuable insights for the identification of causal regulatory factors in biological processes [74]. A class of GRN inference methods are based on Boolean network models, such as SCNS toolkit [80] and BTR [81]. Another approach for the inference of regulatory networks is based on co-expression analysis, and example models include SINC ERA [82], which is specifically used for scRNA-seq data. In addition, there are algorithms based on ordinary differential equations, such as SCODE [83] and InferenceSnapshot [84]. Recently, Moerman et al. [85] proposed the GRNBoost2 and Arboreto frameworks, which can help researchers to deduce high-quality GRNs from large datasets in a reasonable amount of time.

Progress of single-cell RNA sequencing in tumors

Cancer patients may be unresponsive to therapy due to drug resistance and metastasis of single cells, both of which constitute major challenges in the treatment of malignant tumors. About 90% of available drugs are effective in less than half of patients [86]. Cancer is associated with the interaction of thousands of gene products, and genotype as well as interactions vary greatly within and between tumors [87], which is a key reason for the failure of some drugs. In contrast to bulk analysis, which does not account for the differences between cancer cells and their cancer-related counterparts, scRNA-seq provides unprecedented high resolution for the analysis of each individual malignant cell, stromal cell, endothelial cell, parenchymal cell, and immune cell, as well gene expression and pathway activation [88]. Thus, scRNA-seq provides insights that contribute to the development of strategies for cancer treatment and personalized medicine (see Additional file 1).

Highlighting intra- and inter-tumoral heterogeneity

The considerable heterogeneity of tumors and tumor tissue samples between different patients is an important reason for treatment failure [89]. Therefore, understanding the functional status of individual tumor cells and recognizing cell subset composition and characteristics is of great significance for cancer biology and treatment strategies (Table 1 (see Additional file 2)).

Glioblastoma (GBM) is the most common primary malignant brain tumor in adults [90]. It is the glioma with the highest degree of malignancy, and the most often seen in clinical practice, with poor prognosis and a lack of effective treatment regimens [91]. In 2014, Patel et al. [92] analyzed 430 cells of 5 primary GBMs (all IDH1/2 wild-type primary GBMs) using scRNA-seq, and

found that these cells differed in the expression of various programs related to carcinogenic signaling pathways, cell proliferation, the immune response, and hypoxic stress. In 2018, Yuan et al. [93] analyzed high-grade glioma (HGG) with large-scale parallel scRNA-seq using a high-density microwell system, and found that, similar to oligodendrocyte progenitors, glioma cells exhibited proliferative characteristics. In contrast, similar to astrocytes, neuroblasts, and oligodendrocytes, glioma cells exhibited an amitotic state in tumors.

Melanoma is a highly malignant skin cancer with four clinically distinguishable subtypes and is responsible for approximately half of skin cancer-related deaths in Japan [94]. Gerber et al. [95] used scRNA-seq to analyze transcription in cells from three different metastatic melanoma patients (BRAF/NRAS wild type, BRAF mutant/NRAS wild type, and BRAF wild type /NRAS mutant). BRAF/NRAS wild-type samples had a low-abundance subgroup with high expression of ABC transporters, while cells from the other two samples exhibited more homogeneous single-cell gene expression patterns.

Head and neck squamous cell carcinoma (HNSCC) encompasses a group of malignant tumors originating from the squamous epithelium of the oral cavity, oropharynx, larynx, and hypopharynx [96–98]. Puram et al. [9] generated scRNA-seq profiles of HNSCC in 18 patients and found that stromal and immune cells had the same expression program per patient, whereas the expression programs of malignant cells within and between tumors were different with regard to cell cycle, hypoxia, stress, and epithelial differentiation.

Lung cancer (LC) can be divided into two subtypes: small-cell lung carcinoma (SCLC) and non-small-cell lung carcinoma (NSCLC), among which NSCLC accounts for the majority of LC cases [99]. Lambrechts et al. [100] analyzed 92,948 NSCLC cells (84,341 stromal cells) at single-cell resolution, and identified 52 subtypes of stromal cells, providing a comprehensive list of stromal cell types. Based on these findings, the tumor microenvironment (TME) of lung cancer is more complex and heterogeneous than previously thought.

Acute myeloid leukemia (AML) is a highly heterogeneous hematological malignancy. Van Galen et al. [101] profiled the cells of 16 patients with AML and 5 healthy donors by combining scRNA-seq with single-cell genotyping, and then correctly classified the cell types by machine learning, which could correctly distinguish normal cells from cancer cells, and identified 6 malignant cell types, which comprehensively parsed the heterogeneous ecosystem of AML.

Discovery of invasion and metastasis mechanisms

The ability of single-cell gene expression profiling to identify specific patterns of gene expression allows for

the elucidation of mechanisms underlying tumor invasion and metastasis [102–105] that are critical for preventing the spread of cancer, which considerably complicates treatment (Table 2 (see Additional file 3)).

Pancreatic cancer is one of the leading causes of cancer-related death in developed countries, and it ranks 7th in cancer-related deaths among both men and women [106]. Ting et al. [107] obtained high-quality transcripts from 93 single mouse pancreatic circulating tumor cells (CTCs) and compared them with scRNA-seq data from pancreatic cancer patients. The expression of an extracellular matrix protein-encoding gene, SPARC, was abnormally high in both mouse and human pancreatic CTCs, and its knockdown decreased the ability of cancer cells to invade and metastasize, indicating that the gene is closely related to pancreatic cancer invasion and metastasis. In 2017, Puram et al. [9] identified partial epithelial-to-mesenchymal transition (p-EMT) in malignant HNSCC tumor cells through scRNA-seq. Cells exhibiting p-EMT activation were located at the outer area of primary tumors, promoting the invasion and metastasis of tumor cells.

Breast cancer is by far the most common malignancy in women [108] as well as the leading cause of cancer-related death in women [109]. Chung et al. [110] conducted transcriptomic analysis of 515 single cells from patients with different breast cancer subtypes, revealing the gene expression patterns of cancer cells and immune cells. It is worth noting that rare cell types exhibiting pronounced EMT and stemness phenotypes were identified in triple-negative breast cancer (TNBC), and these may be a driving force of tumor progression and metastasis. Multiple myeloma (MM) is the second most common hematologic malignancy and cannot be completely cured [111]. Geng et al. [112] performed single-cell transcriptomic analysis of cells taken from 21 MM patients, and found that the chemokine CXCL12 was abnormally upregulated in circulating plasma cells (cPCs). CXCL12 may prompt cPCs to escape bone marrow retention and migrate to the bloodstream, eventually forming an extramedullary plasmacytoma (EMP).

Study of drug resistance mechanisms

When anti-cancer drugs are used to target tumor cells, some cells develop resistance to drugs and contribute to disease relapse, rendering treatment ineffective. ScRNA-seq can be employed to distinguish between primary and recurrent tumor cells, providing insight into the cell subsets responsible for tumor recurrence, mutated genes, and multiple pathways that drive tumor growth. Thus, scRNA-seq can guide strategies for the treatment of tumors (Table 3 (see Additional file 4)).

In 2014, Lee et al. [113] employed scRNA-seq in metastatic breast cancer cells treated with the

chemotherapeutic agent paclitaxel. A small number of cells failed to respond to paclitaxel by producing specific RNA variants, but these transcripts were not detected in the two groups of cells that were not treated with chemotherapy or effective chemotherapy. In 2015, Kim et al. [114] conducted scRNA-seq on 34 patient-derived xenograft (PDX) cells isolated from a xenograft tumor of a lung adenocarcinoma (LUAD) patient, and identified a tumor cell subpopulation related to anti-cancer drug resistance that expressed KRAS^{G12D} through gene expression and mutation profiling. In 2016, Tirosh et al. [4] performed scRNA-seq on 4645 cells isolated from 19 patients with melanoma, and found that all tumors contained cells with two transcriptional states, namely “MITF-high” and “AXL-high”. After treatment with RAF/MEK inhibitors, the number of “AXL-high” cells gradually increased from their initially small population. However, the proportion of “AXL-high” cells in a lot of the tumors remained almost unchanged, indicating that the cells with high AXL expression were related to targeted therapy resistance. In 2018, Chen et al. [115] performed scRNA-seq in over 200 single cells (from patients with primary and relapsed glioblastoma multiforme), revealing relapse pathways in GBM. A trio of mutated genes, all involved in the RAS/GEF GTP-dependent signaling pathway, were identified in single cells from relapsed GBM multiforme, but not in those from primary tumors. The authors further confirmed the identified molecular pathway by meta-analysis of RNA-seq data from thousands of patients.

Medulloblastoma (MB) is one of the most common malignant brain tumors in children, with a higher incidence in children aged 3 to 4 years and 8 to 10 years, and is slightly more prevalent in boys [116], with no significant racial bias [117]. In 2019, Ocasio et al. [118] analyzed the response of Sonic Hedgehog (SHH)-driven mouse MB to the SHH-pathway inhibitor vismodegib by using scRNA-seq and lineage tracking. Vismodegib could reduce some proliferative cell subgroups and promote the differentiation of certain cells. However, specific cell types continued to proliferate in vismodegib-treated tumors, exhibiting either sustained SHH-pathway activation or stem cell characteristics.

Characterization of the tumor immune microenvironment

Various types of immune cells infiltrate tumor tissue [119–128]. The gene expression status of each myeloid cell and lymphocyte can be obtained via scRNA-seq, providing insights into the immune status of cancers. Such scRNA-seq data is a valuable asset for formulating effective immunotherapy approaches for cancer, alongside studying the basic characteristics of tumor-infiltrating immune cells (Table 4 (see Additional file 5)).

Hepatocellular carcinoma (HCC) usually occurs in the context of advanced chronic liver disease, and is primarily associated with hepatitis B or C virus infection and alcohol abuse, accounting for 70–85% of the total cases of liver cancer [129]. Zheng et al. [8] performed deep scRNA-seq on 5,063 T cells isolated from tumors, peripheral blood, and the adjacent normal tissue of 6 HCC patients. They observed that 11 T cell subsets were present in the HCC microenvironment. The gene *LYAN* was highly expressed in activated CD8⁺ T cell and Tregs, which could inhibit the function of CD8⁺ T cells. In 2017, Muller et al. [130] applied scRNA-seq to glioblastoma-derived myeloid cells, particularly tumor-associated macrophages (TAMs), for the first time. The results revealed that blood-derived TAMs and microglial TAMs exhibited different phenotypes and localization within the tumor. Compared with microglial TAMs, blood-derived TAMs could upregulate immunosuppressive cytokines, exhibited metabolic changes, and were enriched in perivascular and necrotic areas. Moreover, in 2017, researchers from another group [131] used paired single-cell analysis to study the innate immunity of early-stage LUAD, and found a large number of Treg and non-functional T cells, as well as NK cell depletion, in a stage I tumor, indicative of compromised anti-tumor immunity. Some of the immune features found in advanced LC were also present in early LC. For example, PD-1 was present on some of the CD4⁺ and CD8⁺ T cells at both stages. In 2018, Azizi et al. [132] performed scRNA-seq on 45,000 immune cells, mapped the various immune cell phenotypes in the breast TME, and identified most of the immune cell types in 8 patients. Although the expression profiles of immune cells were quite similar between tumor and normal tissue samples, specific phenotypic expansion was observed in tumor immune cells, and T cells were in a continuous state of activation, which did not conform to the traditional polarization distribution model of cancer macrophages.

Colorectal cancer (CRC) is a common gastrointestinal malignancy, and is the third most common cancer in the United States [133]. Zhang et al. [134] obtained the transcriptomes of 11,138 single T cells from 12 patients with CRC by scRNA-seq, and quantitatively analyzed the dynamic relationship between the development, function, and migration of 20 T cell subsets through TCR tracking.

Resolution of long disputed questions

ScRNA-seq can solve problems that previous technologies could not address, thus resolving debates and contributing to the development of oncology.

The relationship between human cytomegalovirus (HCMV) and GBM is an ongoing debate, as the detection of HCMV in GBM samples via molecular assays

may be influenced by cellular heterogeneity. In 2017, Johnson et al. [135] aligned the scRNA-seq reads from 5 GBM tumors and 2 cell lines to HCMV, and found that no complete transcripts of the HCMV virus were discovered in either tumor samples or cell lines.

Understanding the origin and evolution of tumors

ScRNA-seq can be used to better understand the occurrence and development of tumors at the cellular level, which enables to the exploration of new therapeutic methods (Table 5 (see Additional file 6)).

In 2016, Tirosh et al. [136] selected 4347 cells from 6 untreated grade II oligodendrogliomas (harboring IDH1 or IDH2 mutations) and analyzed them using scRNA-seq, and found that the cells with proliferative characteristics were highly enriched in a small number of undifferentiated rare subpopulations, suggesting that malignant proliferative cells within oligodendrogliomas could mainly originate from such a subgroup. In 2017, Venteicher et al. [137] combined scRNA-seq profiles from astrocytoma (IDH-A) and oligodendroglioma (IDH-O) tumors with 165 TCGA bulk RNA-seq profiles to analyze their genotype, phenotype, and the TME. The shared glial lineages and developmental hierarchy observed in both IDH-A and IDH-O tumors indicated that all IDH mutant gliomas had a common progenitor. In 2019, Saurty-Seerunghen et al. [138] analyzed public scRNA-seq data from surgically resected GBM samples and obtained a comprehensive view of metabolic pathways. ELOVL2 was closely related to the tumorigenicity of GBM cells, and knockdown of the ELOVL2 gene could inhibit GBM tumorigenicity, indicating that the overexpression of ELOVL2 promoted the growth of GBM. The authors further explored the mechanisms of ELOVL2 and found that the formation and release of extracellular vesicles is one of the pathways through which the regulation of which ELOVL2 controls GBM development.

In 2019, Hovestadt et al. [139] obtained 25 MB samples, including all molecular subgroups, and created the first cell map of all MB subgroups, revealing that group 4-MB probably originated from unipolar brush cells and glutamatergic cerebellar nuclei. In the same year, Weng et al. [140] identified Zfp3611 as a key regulator of oligodendrocyte-astrocyte lineage transition and gliomagenesis by lineage-targeted single-cell transcriptomics analysis. Through experiments in mouse and human glioma cells, Zfp3611 was demonstrated to control the genesis and growth of gliomas. Furthermore, gliomas with higher expression of Zfp3611 were more difficult to treat. In 2020, Chen et al. [141] collected a large cohort of histone H3.3 G34R/V (glycine 34 to arginine or valine) HGGs ($n = 95$), and proved that G34R/V HGGs originated from interneuron progenitors expressing GSX2/

DLX using scRNA-seq, and then confirmed this result using epigenomic profiles.

Applications in anticancer research design

At present, practice has proven the guiding significance of scRNA-seq in anticancer protocol design, with important clinical application value in personalized diagnosis, treatment, and prognosis, as well as the evaluation of treatment effect for highly heterogeneous tumors (Table 6 (see Additional file 7)).

About 30% of patients with renal cell carcinoma (RCC) are diagnosed with metastases, and metastatic renal cell carcinoma (MRCC) is one of the most drug-resistant malignancies [142]. In 2016, Kim et al. [143] used scRNA-seq to detect the activation of targeted pathways in patients with refractory MRCC at the single-cell level. On the basis of predicting the activation of multiple drug target pathways, the authors proposed a combinatorial therapeutic strategy for the treatment of metastatic cancer. Whether in vitro or in vivo, the effect of this combined strategy was significantly improved compared to monotherapy. In 2019, Zhang et al. [144] proposed a multilayer network biomarker (MNB) based on the scRNA-seq data of IDH-mutant astrocytoma samples. MNB links cancer cells and the TME, allowing for prognosis and prediction of the therapeutic response in glioma patients.

Identification of potential therapeutic targets

Modern cancer treatment approaches are often limited in their efficacy. In-depth analysis of malignant proliferative cells and immune cells at single-cell resolution is conducive to the identification of potential therapeutic targets, contributing to the development of new drug therapy and the improvement of patient survival rate (Table 7 (see Additional file 8)).

In 2017, Darmanis et al. [145] performed scRNA-seq on 3589 cells taken from the tumor core and surrounding tissue of 4 GBM patients. Groups of genes involved in size regulation, energy production, inhibition of apoptosis, regulation of intercellular adhesion, and central nervous system development were identified by analyzing the upregulated genes in infiltrating tumor cells, which could open up new avenues for treatment. In 2018, a study [146] conducted scRNA-seq on 3321 cells from 6 primary gliomas with histone H3 lysine27-to-methionine mutations (H3K27m-glioma) and matched models (PDX, gliomaspheres, and differentiated glioma cells), and found that H3K27m-glioma were mainly composed of cells similar to oligodendrocyte precursor cells (OPC-like). OPC-like cells exhibited greater proliferative ability and are maintained, at least in part, via PDGFRA signaling. In 2019, Wang et al. [147] extracted 16,128 tumor cells from epidermal growth factor

receptor (EGFR) and EGFRvIII GBM tumors for scRNA-seq. RAD51AP1 was for the first time identified as an oncogene in GBM, highly related to EGFRvIII. This study revealed a new possibility for treatment using temozolomide combined with RAD51AP1, which could enhance the therapeutic effect and prolong patient survival.

Nasopharyngeal carcinoma (NPC) is an epithelial carcinoma arising from the lining of nasopharyngeal mucosa [148–152]. Although there is considerable NPC incidence in east and southeast Asia, it is not common compared to other types of cancer. Zhao et al. [7] constructed the first transcriptome profiles of primary NPC malignant cells and infiltrating immune cells using scRNA-seq. A high proportion of B cells infiltrating into the TMEs of three NPC patients was identified, and the expression of cell cycle genes related to proliferation was significantly upregulated in EBV-positive NPC cells, suggesting that B cells and cell cycle gene expression may be potential treatment targets. In addition, LAG-3 and HAVCR-2 were the most expressed checkpoint genes in CD8+ T cells and may thus be potential new checkpoint inhibition targets for NPC immunotherapy.

Conclusions

High-throughput scRNA-seq technology is transforming biomedical research by establishing the transcriptome profiles of individual cells from different tissue samples [55, 153]. While there are various sequencing platforms and tools available for analyzing scRNA-seq data, each has its limitations. With ongoing technological developments, scRNA-seq will become more high-throughput, the cost of scRNA-seq and associated data analysis will decrease, the time required will be shorter, and various indicators such as accuracy and sensitivity will improve.

In time, scRNA-seq will be more widely used in the study of various tumor types, helping to develop better diagnostic and prognostic biomarkers, and design more accurate anticancer therapy to improve therapeutic efficacy and avoid tumor drug resistance. While the transcriptome of cells can be studied through scRNA-seq, there are some limitations to this method, particularly with regard to genomic and protein-level information. When CITE-seq was used in combination with scRNA-seq, the transcripts and surface marker proteins of cells could be detected simultaneously, and the phenotypes that could not be detected by scRNA-seq alone could be found [154]. The combination of scRNA-seq with genotyping can more accurately distinguish malignant cells from normal cells [101]. In the future, different omics technologies could combine with scRNA-seq technology to more comprehensively characterize individual cells. With deepening understanding of the cellular dynamics of cancer, the efficacy of personalized medicine will

improve, ultimately saving lives and reducing the global burden of cancer on healthcare systems.

Abbreviations

scRNA-seq: Single-cell RNA sequencing; scLVM: Single-cell latent variable model; OT: Optimal transport; DE: Differentially expressed; GSE: Gene set enrichment; GRN: Gene regulatory network; GBM: Glioblastoma; HGG: High-grade glioma; LC: Lung cancer; SCLC: Small-cell lung carcinoma; NSCLC: Non-small-cell lung carcinoma; TME: Tumor microenvironment; AML: Acute myeloid leukemia; CTC: Circulating tumor cell; p-EMT: Partial epithelial-to-mesenchymal transition; TNBC: Triple negative breast cancer; MM: Multiple myeloma; cPC: Circulating plasma cell; EMP: Extramedullary plasmacytoma; PDX: Patient-derived xenograft; LUAD: Lung adenocarcinoma; MB: Medulloblastoma; SHH: Sonic Hedgehog; HCC: Hepatocellular carcinoma; TAM: Tumor-associated macrophage; CRC: Colorectal cancer; HCMV: Human cytomegalovirus; G34R/V: Glycine 34 to arginine or valine; RCC: Renal cell carcinoma; MRCC: Metastatic renal cell carcinoma; MNB: Multilayer network biomarker; OPC: Oligodendrocyte precursor cell; NPC: Nasopharyngeal carcinoma; EGFR: Epidermal growth factor receptor

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13046-021-01955-1>.

Additional file 1 : Fig. 1. Application of scRNA-seq in tumor research. ScRNA-seq can be used to analyze the proportion of various immune cell types and immune repertoire characteristics, guiding choices related to immunotherapy. The analysis of single tumor cells can improve our understanding of drug resistance, invasion, and metastasis mechanisms, as well as cell origin and evolution, in turn aiding the choice of targeted drugs. T, T cell; B, B cell; NK, natural killer cell; TAM, tumor-associated macrophage; Th, helper T cell; Tc, cytotoxic T cell; Treg, regulatory T cell; Tm, memory T cell; Ts, suppressor T cell; B1, T cell independent B cell; B2, T cell dependent B cell.

Additional file 2 : Table 1. Overview of related studies using scRNA-seq.

Additional file 3 : Table 2. Overview of related articles using scRNA-seq.

Additional file 4 : Table 3. Overview of related studies using scRNA-seq.

Additional file 5 : Table 4. Overview of related articles using scRNA-seq.

Additional file 6 : Table 5. Overview of related articles using scRNA-seq.

Additional file 7 : Table 6. Overview of related articles using scRNA-seq.

Additional file 8 : Table 7. Overview of related articles using scRNA-seq.

Acknowledgements

Not applicable.

Authors' contributions

LLY, XF, WYM, ZSS, GZJ, LXY, HY, SL, WFY, LQJ, XB, ZM, LXL collected the related paper and drafted the manuscript. LY, LGY, ZZY, XW, GC participated in the design of the review and draft the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported partially by grants from the National Natural Science Foundation of China (81903138, 81972776, 81803025, 81772928, 81702907, 81772901, 81672993, 81672683), and the Natural Science Foundation of Hunan Province (2019JJ50778, 2018SK21210, 2018SK21211, 2018JJ3704, 2018JJ3815).

Availability of data and materials

Not applicable.

Declarations

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

¹NHC Key Laboratory of Carcinogenesis and Hunan Key Laboratory of Cancer Metabolism, Hunan Cancer Hospital and the Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Changsha, China. ²Key Laboratory of Carcinogenesis and Cancer Invasion of the Chinese Ministry of Education, Cancer Research Institute, Central South University, Changsha, China. ³Department of Stomatology, Xiangya Hospital, Central South University, Changsha, China. ⁴Department of Oral and Maxillofacial Surgery, The Second Xiangya Hospital, Central South University, Changsha, China. ⁵Hunan Key Laboratory of Nonresolving Inflammation and Cancer, Disease Genome Research Center, The Third Xiangya Hospital, Central South University, Changsha, China. ⁶Department of Medicine, Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX, USA.

Received: 26 January 2021 Accepted: 20 April 2021

Published online: 11 May 2021

References

- Williams MJ, Werner B, Heide T, et al. Quantification of subclonal selection in cancer from bulk sequencing data. *Nat Genet.* 2018;50(6):895–903. <https://doi.org/10.1038/s41588-018-0128-6>.
- Tang F, Barbacioru C, Wang Y, et al. mRNA-seq whole-transcriptome analysis of a single cell. *Nat Methods.* 2009;6(5):377–82. <https://doi.org/10.1038/nmeth.1315>.
- Suva ML, Tirosh I. Single-cell RNA sequencing in cancer: lessons learned and emerging challenges. *Mol Cell.* 2019;75(1):7–12. <https://doi.org/10.1016/j.molcel.2019.05.003>.
- Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352(6282):189–96. <https://doi.org/10.1126/science.aad0501>.
- Shalek AK, Satija R, Shuga J, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature.* 2014;510(7505):363–9. <https://doi.org/10.1038/nature13437>.
- Haque A, Engel J, Teichmann SA, Lonnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* 2019;9(1):75. <https://doi.org/10.1186/s13073-017-0467-4>.
- Zhao J, Guo C, Xiong F, Yu J, Ge J, Wang H, et al. Single cell RNA-seq reveals the landscape of tumor and infiltrating immune cells in nasopharyngeal carcinoma. *Cancer Lett.* 2020;477:131–43. <https://doi.org/10.1016/j.canlet.2020.02.010>.
- Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, et al. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell.* 2017;169(7):1342–56 e1316. <https://doi.org/10.1016/j.cell.2017.05.035>.
- Puram SV, Tirosh I, Parkh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell.* 2017;171(7):1611–24 e1624. <https://doi.org/10.1016/j.cell.2017.10.044>.
- Zhang X, Li T, Liu F, Chen Y, Yao J, Li Z, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. *Mol Cell.* 2019;73(1):130–42 e135. <https://doi.org/10.1016/j.molcel.2018.10.020>.
- Natarajan KN, Miao Z, Jiang M, Huang X, Zhou H, Xie J, et al. Comparative analysis of sequencing technologies for single-cell transcriptomics. *Genome Biol.* 2019;20(1):70. <https://doi.org/10.1186/s13059-019-1676-5>.
- Jiang P, Thomson JA, Stewart R. Quality control of single-cell RNA-seq by SinQC. *Bioinformatics.* 2016;32(16):2514–6. <https://doi.org/10.1093/bioinformatics/btw176>.
- Ilicic T, Kim JK, Kolodziejczyk AA, Bagger FO, McCarthy DJ, Marioni JC, et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.* 2016;17(1):29. <https://doi.org/10.1186/s13059-016-0888-1>.
- McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics.* 2017;33(8):1179–86. <https://doi.org/10.1093/bioinformatics/btw777>.
- Osorio D, Cai JJ. Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA sequencing data quality control. *Bioinformatics.* 2020. <https://doi.org/10.1093/bioinformatics/btaa751>.
- Zhao Q, Wang J, Levichkin IV, Stasinopoulos S, Ryan MT, Hoogenraad NJ. A mitochondrial specific stress response in mammalian cells. *EMBO J.* 2002;21(17):4411–9. <https://doi.org/10.1093/emboj/cdf445>.
- Li X, Nair A, Wang S, Wang L. Quality control of RNA-seq experiments. *Methods Mol Biol (Clifton, NJ).* 2015;1269:137–46. https://doi.org/10.1007/978-1-4939-2291-8_8.
- Liu Q, Sheng Q, Ping J, et al. scRNABatchQC: multi-samples quality control for single cell RNA-seq data. *Bioinformatics.* 2019;35(24):5306–8. <https://doi.org/10.1093/bioinformatics/btz601>.
- Etherington GJ, Soranzo N, Mohammed S, Haerty W, Davey RP, Palma FD. A Galaxy-based training resource for single-cell RNA-sequencing quality control and analyses. *Gigascience.* 2019;8(12):giz144. <https://doi.org/10.1093/gigascience/giz144>.
- Hicks SC, Townes FW, Teng M, Irizarry RA. Missing data and technical variability in single-cell RNA-sequencing experiments. *Biostatistics.* 2018;19(4):562–78. <https://doi.org/10.1093/biostatistics/kxx053>.
- Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. *Cell.* 2019;177(7):1888–1902.e1821. <https://doi.org/10.1016/j.cell.2019.05.031>.
- Shaham U, Stanton KP, Zhao J, et al. Removal of batch effects using distribution-matching residual networks. *Bioinformatics.* 2017;33(16):2539–46. <https://doi.org/10.1093/bioinformatics/btx196>.
- Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with harmony. *Nat Methods.* 2019;16(12):1289–96. <https://doi.org/10.1038/s41592-019-0619-0>.
- Hie B, Bryson B, Berger B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat Biotechnol.* 2019;37(6):685–91. <https://doi.org/10.1038/s41587-019-0113-3>.
- Welch J, Kozareva V, Ferreira A, Vanderburg C, Martin C, Macosko E. Integrative inference of brain cell similarities and differences from single-cell genomics. *bioRxiv.* 2018:459891. <https://doi.org/10.1101/459891>.
- Lin Y, Ghazanfar S, Wang KYX, et al. scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. *Proc Natl Acad Sci U S A.* 2019;116(20):9775–84. <https://doi.org/10.1073/pnas.1820061116>.
- Risso D, Perraudeau F, Gribkova S, Dudoit S, Vert JP. A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat Commun.* 2018;9(1):284. <https://doi.org/10.1038/s41467-017-02554-5>.
- Tran HTN, Ang KS, Chevrier M, et al. A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol.* 2020;21(1):12. <https://doi.org/10.1186/s13059-019-1850-9>.
- Fei T, Yu T. scBatch: batch effect correction of RNA-seq data through sample distance matrix adjustment. *Bioinformatics.* 2020;36(10):3115–23. <https://doi.org/10.1093/bioinformatics/btaa097>.
- Lun AT, Bach K, Marioni JC. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol.* 2016;17:75. <https://doi.org/10.1186/s13059-016-0947-7>.
- Vallejos CA, Marioni JC, Richardson S. BASiCS: Bayesian analysis of single-cell sequencing data. *PLoS Comput Biol.* 2015;11(6):e1004333. <https://doi.org/10.1371/journal.pcbi.1004333>.
- Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11(3):R25. <https://doi.org/10.1186/gb-2010-11-3-r25>.
- Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol.* 2010;11(10):R106. <https://doi.org/10.1186/gb-2010-11-10-r106>.
- Cole MB, Risso D, Wagner A, et al. Performance assessment and selection of normalization procedures for single-cell RNA-seq. *Cell Syst.* 2019;8(4):315–328.e318. <https://doi.org/10.1016/j.cels.2019.03.010>.
- Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* 2019;20(1):296. <https://doi.org/10.1186/s13059-019-1874-1>.
- Lytal N, Ran D, An L. Normalization methods on single-cell RNA-seq data: an empirical survey. *Front Genet.* 2020;11:41. <https://doi.org/10.3389/fgene.2020.00041>.

37. Kastan MB, Bartek J. Cell-cycle checkpoints and cancer. *Nature*. 2004; 432(7015):316–23. <https://doi.org/10.1038/nature03097>.
38. Bar-Joseph Z, Siegfried Z, Brandeis M, et al. Genome-wide transcriptional analysis of the human cell cycle identifies genes differentially regulated in normal and cancer cells. *Proc Natl Acad Sci U S A*. 2008;105(3):955–60. <https://doi.org/10.1073/pnas.0704723105>.
39. Wang D, Zeng Z, Zhang S, et al. Epstein-Barr virus-encoded miR-BART6-3p inhibits cancer cell proliferation through the LOC553103-STMN1 axis. *FASEB J*. 2020;34(6):8012–27. <https://doi.org/10.1096/fj.202000039r>.
40. Wu P, Mo Y, Peng M, Tang T, Zhong Y, Deng X, et al. Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA. *Mol Cancer*. 2020;19(1):22. <https://doi.org/10.1186/s12943-020-1147-3>.
41. Jin K, Wang S, Zhang Y, et al. Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis. *Cell Mol Life Sci*. 2019;76(21):4275–89. <https://doi.org/10.1007/s00018-019-03222-1>.
42. Pauklin S, Vallier L. The cell-cycle state of stem cells determines cell fate propensity. *Cell*. 2013;155(1):135–47. <https://doi.org/10.1016/j.cell.2013.08.031>.
43. Singh AM, Chappell J, Trost R, et al. Cell-cycle control of developmentally regulated transcription factors accounts for heterogeneity in human pluripotent cells. *Stem Cell Rep*. 2013;1(6):532–44. <https://doi.org/10.1016/j.stemcr.2013.10.009>.
44. Scialdone A, Natarajan KN, Saraiva LR, et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods (San Diego, Calif)*. 2015;85:54–61. <https://doi.org/10.1016/j.jymeth.2015.06.021>.
45. Buettner F, Natarajan KN, Casale FP, et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol*. 2015;33(2):155–60. <https://doi.org/10.1038/nbt.3102>.
46. Hsiao CJ, Tung P, Blischak JD, et al. Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis. *Genome Res*. 2020; 30(4):611–21. <https://doi.org/10.1101/gr.247759.118>.
47. Dey KK, Hsiao CJ, Stephens M. Visualizing the structure of RNA-seq expression data using grade of membership models. *PLoS Genet*. 2017; 13(3):e1006599. <https://doi.org/10.1371/journal.pgen.1006599>.
48. Lin P, Troup M, Ho JW. CIDR: ultrafast and accurate clustering through imputation for single-cell RNA-seq data. *Genome Biol*. 2017;18(1):59. <https://doi.org/10.1186/s13059-017-1188-0>.
49. Wang B, Zhu J, Pierson E, Ramazzotti D, Batzoglou S. Visualization and analysis of single-cell RNA-seq data by kernel-based similarity learning. *Nat Methods*. 2017;14(4):414–6. <https://doi.org/10.1038/nmeth.4207>.
50. Yang Y, Huh R, Culpepper HW, Lin Y, Love MI, Li Y. SAFE-clustering: single-cell aggregated (from ensemble) clustering for single-cell RNA-seq data. *Bioinformatics*. 2019;35(8):1269–77. <https://doi.org/10.1093/bioinformatics/bty793>.
51. Freytag S, Tian L, Lönnstedt I, Ng M, Bahlo M. Comparison of clustering tools in R for medium-sized 10x Genomics single-cell RNA-sequencing data. *F1000Research*. 2018;7:1297. <https://doi.org/10.12688/f1000research.15809.2>.
52. Duò A, Robinson MD, Soneson C. A systematic performance evaluation of clustering methods for single-cell RNA-seq data. *F1000Research*. 2018;7: 1141. <https://doi.org/10.12688/f1000research.15666.3>.
53. Kim T, Chen IR, Lin Y, Wang AY, Yang JYH, Yang P. Impact of similarity metrics on single-cell RNA-seq data clustering. *Brief Bioinform*. 2019;20(6): 2316–26. <https://doi.org/10.1093/bib/bby076>.
54. Peng L, Tian X, Tian G, et al. Single-cell RNA-seq clustering: datasets, models, and algorithms. *RNA Biol*. 2020;17(6):765–83. <https://doi.org/10.1080/15476286.2020.1728961>.
55. Geddes TA, Kim T, Nan L, Burchfield JG, Yang JYH, Tao D, et al. Autoencoder-based cluster ensembles for single-cell RNA-seq data analysis. *BMC Bioinformatics*. 2019;20(Suppl 19):660. <https://doi.org/10.1186/s12859-019-3179-5>.
56. Fang Q, Su D, Ng W, Feng J. An effective biclustering-based framework for identifying cell subpopulations from scRNA-seq data. *IEEE/ACM Trans Comput Biol Bioinform*. 2020. <https://doi.org/10.1109/tcbb.2020.2979717>.
57. Huh R, Yang Y, Jiang Y, Shen Y, Li Y. SAME-clustering: single-cell aggregated clustering via mixture model ensemble. *Nucleic Acids Res*. 2020;48(1):86–95. <https://doi.org/10.1093/nar/gkz959>.
58. Tsia KK, So HKH, Ho JWK, et al. PARC: ultrafast and accurate clustering of phenotypic data of millions of single cells. *Bioinformatics*. 2020;36(9):2778–86. <https://doi.org/10.1093/bioinformatics/btaa042>.
59. Shao X, Liao J, Lu X, Xue R, Ai N, Fan X. scCATCH: automatic annotation on cell types of clusters from single-cell RNA sequencing data. *iScience*. 2020; 23(3):100882. <https://doi.org/10.1016/j.isci.2020.100882>.
60. Chen G, Ning B, Shi T. Single-cell RNA-seq technologies and related computational data analysis. *Front Genet*. 2019;10:317. <https://doi.org/10.3389/fgene.2019.00317>.
61. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979–82. <https://doi.org/10.1038/nmeth.4402>.
62. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature*. 2019;566(7745):496–502. <https://doi.org/10.1038/s41586-019-0969-x>.
63. Ji Z, Ji H. TSCAN: Pseudo-time reconstruction and evaluation in single-cell RNA-seq analysis. *Nucleic Acids Res*. 2016;44(13):e117. <https://doi.org/10.1093/nar/gkw430>.
64. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*. 2018;19(1):477. <https://doi.org/10.1186/s12864-018-4772-0>.
65. Guo M, Bao EL, Wagner M, Whitsett JA, Xu Y. SLICE: determining cell differentiation and lineage based on single cell entropy. *Nucleic Acids Res*. 2017;45(7):e54. <https://doi.org/10.1093/nar/gkw1278>.
66. Chen Y, Zhang Y, Ouyang Z. LISA: accurate reconstruction of cell trajectory and pseudo-time for massive single cell RNA-seq data. *Pac Symp Biocomput*. 2019;24:338–49. https://doi.org/10.1142/9789813279827_0031.
67. Herring CA, Banerjee A, McKinley ET, et al. Unsupervised trajectory analysis of single-cell RNA-seq and imaging data reveals alternative tuft cell origins in the gut. *Cell Syst*. 2018;6(1):37–51.e39. <https://doi.org/10.1016/j.cels.2017.10.012>.
68. Schiebinger G, Shu J, Tabaka M, et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell*. 2019;176(4):928–943.e922. <https://doi.org/10.1016/j.cell.2019.01.006>.
69. Saelens W, Cannoodt R, Todorov H, Saeyns Y. A comparison of single-cell trajectory inference methods: towards more accurate and robust tools. *bioRxiv*. 2018:276907. <https://doi.org/10.1101/276907>.
70. Delmans M, Hemberg M. Discrete distributional differential expression (D3E)—a tool for gene expression analysis of single-cell RNA-seq data. *BMC Bioinformatics*. 2016;17:110. <https://doi.org/10.1186/s12859-016-0944-6>.
71. Miao Z, Deng K, Wang X, Zhang X. DEsingle for detecting three types of differential expression in single-cell RNA-seq data. *Bioinformatics*. 2018; 34(18):3223–4. <https://doi.org/10.1093/bioinformatics/bty332>.
72. Ye C, Speed TP, Salim A. DECENT: differential expression with capture efficiency adjustment for single-cell RNA-seq data. *Bioinformatics*. 2019; 35(24):5155–62. <https://doi.org/10.1093/bioinformatics/btz453>.
73. Zhang W, Wei Y, Zhang D, Xu EY. ZIAQ: a quantile regression method for differential expression analysis of single-cell RNA-seq data. *Bioinformatics*. 2020;36(10):3124–30. <https://doi.org/10.1093/bioinformatics/btaa098>.
74. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol*. 2019;15(6):e8746. <https://doi.org/10.15252/msb.20188746>.
75. Huang DW, Sherman BT, Tan Q, et al. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res*. 2007;35(Web Server issue): W169–75. <https://doi.org/10.1093/nar/gkm415>.
76. Kim SY, Volsky DJ. PAGE: parametric analysis of gene set enrichment. *BMC Bioinformatics*. 2005;6:144. <https://doi.org/10.1186/1471-2105-6-144>.
77. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res*. 2012;40(17):e133. <https://doi.org/10.1093/nar/gks461>.
78. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol*. 2012;8(2): e1002375. <https://doi.org/10.1371/journal.pcbi.1002375>.
79. Ma Y, Sun S, Shang X, Keller ET, Chen M, Zhou X. Integrative differential expression and gene set enrichment analysis using summary statistics for scRNA-seq studies. *Nat Commun*. 2020;11(1):1585. <https://doi.org/10.1038/s41467-020-15298-6>.
80. Woodhouse S, Piterman N, Wintersteiger CM, Göttgens B, Fisher J. SCNS: a graphical tool for reconstructing executable regulatory networks from single-cell genomic data. *BMC Syst Biol*. 2018;12(1):59. <https://doi.org/10.1186/s12918-018-0581-y>.
81. Lim CY, Wang H, Woodhouse S, et al. BTR: training asynchronous Boolean models using single-cell expression data. *BMC Bioinformatics*. 2016;17(1): 355. <https://doi.org/10.1186/s12859-016-1235-y>.
82. Guo M, Wang H, Potter SS, Whitsett JA, Xu Y. SINCERA: a pipeline for single-cell RNA-Seq profiling analysis. *PLoS Comput Biol*. 2015;11(11):e1004575. <https://doi.org/10.1371/journal.pcbi.1004575>.

83. Matsumoto H, Kiryu H, Furusawa C, et al. SCODE: an efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. *Bioinformatics*. 2017;33(15):2314–21. <https://doi.org/10.1093/bioinformatics/btx194>.
84. Ocone A, Haghverdi L, Mueller NS, Theis FJ. Reconstructing gene regulatory dynamics from high-dimensional single-cell snapshot data. *Bioinformatics*. 2015;31(12):i89–96. <https://doi.org/10.1093/bioinformatics/btv257>.
85. Moerman T, Aibar Santos S, Bravo González-Blas C, et al. GRNBoost2 and Arboreto: efficient and scalable inference of gene regulatory networks. *Bioinformatics*. 2019;35(12):2159–61. <https://doi.org/10.1093/bioinformatics/bty916>.
86. [No authors listed]. What happened to personalized medicine? *Nat Biotechnol*. 2012;30(1):1. <https://doi.org/10.1038/nbt.2096>.
87. Benson M. Clinical implications of omics and systems medicine: focus on predictive and individualized treatment. *J Intern Med*. 2016;279(3):229–40. <https://doi.org/10.1111/joim.12412>.
88. Shalek AK, Benson M. Single-cell analyses to tailor treatments. *Sci Transl Med*. 2017;9(408):eaan4730. <https://doi.org/10.1126/scitranslmed.aan4730>.
89. Wei J, Wu C, Meng H, et al. The biogenesis and roles of extrachromosomal oncogene involved in carcinogenesis and evolution. *Am J Cancer Res*. 2020; 10(11):3532–50.
90. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N Engl J Med*. 2005;352(10): 987–96. <https://doi.org/10.1056/nejmoa043330>.
91. Sottoriva A, Spiteri I, Piccirillo SG, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A*. 2013;110(10):4009–14. <https://doi.org/10.1073/pnas.1219747110>.
92. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. 2014;344(6190):1396–401. <https://doi.org/10.1126/science.1254257>.
93. Yuan J, Levitin HM, Frattini V, et al. Single-cell transcriptome analysis of lineage diversity in high-grade glioma. *Genome Med*. 2018;10(1):57. <https://doi.org/10.1186/s13073-018-0567-9>.
94. Namikawa K, Yamazaki N. Targeted therapy and immunotherapy for melanoma in Japan. *Curr Treat Options in Oncol*. 2019;20(1):7. <https://doi.org/10.1007/s11864-019-0607-8>.
95. Gerber T, Willscher E, Loeffler-Wirth H, et al. Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. *Oncotarget*. 2017;8(1):846–62. <https://doi.org/10.18632/oncotarget.13666>.
96. Fan C, Wang J, Tang Y, Zhang S, Xiong F, Guo C, et al. Upregulation of long non-coding RNA LOC284454 may serve as a new serum diagnostic biomarker for head and neck cancers. *BMC Cancer*. 2020;20(1):917. <https://doi.org/10.1186/s12885-020-07408-w>.
97. Ban Y, Tan P, Cai J, et al. LNCAROD is stabilized by m6A methylation and promotes cancer progression via forming a ternary complex with HSPA1A and YBX1 in head and neck squamous cell carcinoma. *Mol Oncol*. 2020; 14(6):1282–96. <https://doi.org/10.1002/1878-0261.12676>.
98. Yi M, Tan Y, Wang L, et al. TP63 links chromatin remodeling and enhancer reprogramming to epidermal differentiation and squamous cell carcinoma development. *Cell Mol Life Sci*. 2020;77(21):4325–46. <https://doi.org/10.1007/s00018-020-03539-2>.
99. Deng X, Xiong W, Jiang X, Zhang S, Li Z, Zhou Y, et al. LncRNA LINC00472 regulates cell stiffness and inhibits the migration and invasion of lung adenocarcinoma by binding to YBX1. *Cell Death Dis*. 2020;11(11):945. <https://doi.org/10.1038/s41419-020-03147-9>.
100. Lambrechts D, Wauters E, Boeckx B, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med*. 2018;24(8):1277–89. <https://doi.org/10.1038/s41591-018-0096-5>.
101. van Galen P, Hovestadt V, Wadsworth IJ, et al. Single-cell RNA-Seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. 2019;176(6):1265–1281 e1224. <https://doi.org/10.1016/j.cell.2019.01.031>.
102. Fan C, Qu H, Xiong F, Tang Y, Tang T, Zhang L, et al. CircARHGAP12 promotes nasopharyngeal carcinoma migration and invasion via ezrin-mediated cytoskeletal remodeling. *Cancer Lett*. 2021;496:41–56. <https://doi.org/10.1016/j.canlet.2020.09.006>.
103. Tang L, Xiong W, Zhang L, et al. circSETD3 regulates MAPRE1 through miR-615-5p and miR-1538 sponges to promote migration and invasion in nasopharyngeal carcinoma. *Oncogene*. 2021;40(2):307–21. <https://doi.org/10.1038/s41388-020-01531-5>.
104. Wu Y, Wang D, Wei F, et al. EBV-miR-BART12 accelerates migration and invasion in EBV-associated cancer cells by targeting tubulin polymerization-promoting protein 1. *FASEB J*. 2020;34(12):16205–23. <https://doi.org/10.1096/fj.202001508r>.
105. Tang T, Yang L, Cao Y, et al. LncRNA AATBC regulates Pinin to promote metastasis in nasopharyngeal carcinoma. *Mol Oncol*. 2020;14(9):2251–70. <https://doi.org/10.1002/1878-0261.12703>.
106. Wu C, Li M, Meng H, Liu Y, Niu W, Zhou Y, et al. Analysis of status and countermeasures of cancer incidence and mortality in China. *Sci China Life Sci*. 2019;62(5):640–7. <https://doi.org/10.1007/s11427-018-9461-5>.
107. Ting David T, Wittner Ben S, Ligorio M, et al. Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep*. 2014;8(6):1905–18. <https://doi.org/10.1016/j.celrep.2014.08.029>.
108. Mermer G, Turk M. Assessment of the effects of breast cancer training on women between the ages of 50 and 70 in Kemalpaşa, Turkey. *Asian Pac J Cancer Prev*. 2014;15(24):10749–55. <https://doi.org/10.7314/apjcp.2014.15.24.10749>.
109. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359–86. <https://doi.org/10.1002/ijc.29210>.
110. Chung W, Eum HH, Lee H-O, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun*. 2017;8(1):15081. <https://doi.org/10.1038/ncomms15081>.
111. Brigle K, Rogers B. Pathobiology and diagnosis of multiple myeloma. *Semin Oncol Nurs*. 2017;33(3):225–36. <https://doi.org/10.1016/j.soncn.2017.05.012>.
112. Geng S, Wang J, Zhang X, et al. Single-cell RNA sequencing reveals chemokine self-feeding of myeloma cells promotes extramedullary metastasis. *FEBS Lett*. 2019;594(3):452–65. <https://doi.org/10.1002/1873-3468.13623>.
113. Lee MC, Lopez-Diaz FJ, Khan SY, et al. Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proc Natl Acad Sci U S A*. 2014;111(44):E4726–35. <https://doi.org/10.1073/pnas.1404656111>.
114. Kim KT, Lee HW, Lee HO, et al. Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. *Genome Biol*. 2015;16:127. <https://doi.org/10.1186/s13059-015-0692-3>.
115. Chen X, Wen Q, Stucky A, et al. Relapse pathway of glioblastoma revealed by single-cell molecular analysis. *Carcinogenesis*. 2018;39(7):931–6. <https://doi.org/10.1093/carcin/bgy052>.
116. Millard NE, De Braganca KC. Medulloblastoma. *J Child Neurol*. 2016;31(12): 1341–53. <https://doi.org/10.1177/0883073815600866>.
117. Smoll NR, Drummond KJ. The incidence of medulloblastomas and primitive neuroectodermal tumours in adults and children. *J Clin Neurosci*. 2012;19(11): 1541–4. <https://doi.org/10.1016/j.jocn.2012.04.009>.
118. Ocasio J, Babcock B, Malawsky D, et al. scRNA-seq in medulloblastoma shows cellular heterogeneity and lineage expansion support resistance to SHH inhibitor therapy. *Nat Commun*. 2019;10(1):5829. <https://doi.org/10.1038/s41467-019-13657-6>.
119. Wei X, Chen Y, Jiang X, et al. Mechanisms of vasculogenic mimicry in hypoxic tumor microenvironments. *Mol Cancer*. 2021;20(1):7. <https://doi.org/10.1186/s12943-020-01288-1>.
120. Jiang X, Wang J, Deng X, Xiong F, Zhang S, Gong Z, et al. The role of microenvironment in tumor angiogenesis. *J Exp Clin Cancer Research*. 2020; 39(1):204. <https://doi.org/10.1186/s13046-020-01709-5>.
121. Fan C, Zhang S, Gong Z, et al. Emerging role of metabolic reprogramming in tumor immune evasion and immunotherapy. *Sci China Life Sci*. 2021; 64(4):534–547. <https://doi.org/10.1007/s11427-019-1735-4>.
122. Zhu K, Li P, Mo Y, Wang J, Jiang X, Ge J, et al. Neutrophils: accomplices in metastasis. *Cancer Lett*. 2020;492:11–20. <https://doi.org/10.1016/j.canlet.2020.07.028>.
123. Wei F, Wang D, Wei J, et al. Metabolic crosstalk in the tumor microenvironment regulates antitumor immunosuppression and immunotherapy resistance. *Cell Mol Life Sci*. 2021;78(1):173–93. <https://doi.org/10.1007/s00018-020-03581-0>.
124. Jiang X, Wang J, Deng X, Xiong F, Ge J, Xiang B, et al. Role of the tumor microenvironment in PD-L1/PD-1-mediated tumor immune escape. *Mol Cancer*. 2019;18(1):10. <https://doi.org/10.1186/s12943-018-0928-4>.
125. Wang YA, Li XL, Mo YZ, Fan CM, Tang L, Xiong F, et al. Effects of tumor metabolic microenvironment on regulatory T cells. *Mol Cancer*. 2018;17(1): 168. <https://doi.org/10.1186/s12943-018-0913-y>.

126. Duan S, Guo W, Xu Z, et al. Natural killer group 2D receptor and its ligands in cancer immune escape. *Mol Cancer*. 2019;18(1):29. <https://doi.org/10.1186/s12943-019-0956-8>.
127. Peng M, Mo Y, Wang Y, et al. Neoantigen vaccine: an emerging tumor immunotherapy. *Mol Cancer*. 2019;18(1):128. <https://doi.org/10.1186/s12943-019-1055-6>.
128. Ren D, Hua Y, Yu B, et al. Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy. *Mol Cancer*. 2020;19(1):19. <https://doi.org/10.1186/s12943-020-1144-6>.
129. Marengo A, Rosso C, Bugianesi E. Liver cancer: connections with obesity, fatty liver, and cirrhosis. *Annu Rev Med*. 2016;67(1):103–17. <https://doi.org/10.1146/annurev-med-090514-013832>.
130. Muller S, Kohanbash G, Liu SJ, et al. Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome Biol*. 2017;18(1):234. <https://doi.org/10.1186/s13059-017-1362-4>.
131. Lavin Y, Kobayashi S, Leader A, et al. Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. *Cell*. 2017;169(4):750–765 e717. <https://doi.org/10.1016/j.cell.2017.04.014>.
132. Azizi E, Carr AJ, Plitas G, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell*. 2018;174(5):1293–1308 e1236. <https://doi.org/10.1016/j.cell.2018.05.060>.
133. Bo H, Fan L, Li J, et al. High expression of lncRNA AFAP1-AS1 promotes the progression of colon cancer and predicts poor prognosis. *J Cancer*. 2018;9(24):4677–83. <https://doi.org/10.7150/jca.26461>.
134. Zhang L, Yu X, Zheng L, Zhang Y, Li Y, Fang Q, et al. Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature*. 2018;564(7735):268–72. <https://doi.org/10.1038/s41586-018-0694-x>.
135. Johnson TS, Abrams ZB, Mo X, Zhang Y, Huang K. Lack of human cytomegalovirus expression in single cells from glioblastoma tumors and cell lines. *J Neuro-Oncol*. 2017;23(5):671–8. <https://doi.org/10.1007/s13365-017-0543-y>.
136. Tirosh I, Venteicher AS, Hebert C, et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature*. 2016;539(7628):309–13. <https://doi.org/10.1038/nature20123>.
137. Venteicher AS, Tirosh I, Hebert C, et al. Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science*. 2017;355(6332):eaai8478. <https://doi.org/10.1126/science.aai8478>.
138. Saurty-Seerunghen MS, Bellenger L, El-Habr EA, et al. Capture at the single cell level of metabolic modules distinguishing aggressive and indolent glioblastoma cells. *Acta Neuropathol Commun*. 2019;7(1):155. <https://doi.org/10.1186/s40478-019-0819-y>.
139. Hovestadt V, Smith KS, Bihannic L, et al. Resolving medulloblastoma cellular architecture by single-cell genomics. *Nature*. 2019;572(7767):74–9. <https://doi.org/10.1038/s41586-019-1434-6>.
140. Weng Q, Wang J, Wang J, et al. Single-cell transcriptomics uncovers glial progenitor diversity and cell fate determinants during development and gliomagenesis. *Cell Stem Cell*. 2019;24(5):707–723 e708. <https://doi.org/10.1016/j.stem.2019.03.006>.
141. Chen CCL, Deshmukh S, Jessa S, Hadjadj D, Lisi V, Andrade AF, et al. Histone H3.3G34-mutant interneuron progenitors co-opt PDGFRA for gliomagenesis. *Cell*. 2020;183(6):1617–1633 e1622. <https://doi.org/10.1016/j.cell.2020.11.012>.
142. Gupta K, Miller JD, Li JZ, Russell MW, Charbonneau C. Epidemiologic and socioeconomic burden of metastatic renal cell carcinoma (mRCC): a literature review. *Cancer Treat Rev*. 2008;34(3):193–205. <https://doi.org/10.1016/j.ctrv.2007.12.001>.
143. Kim KT, Lee HW, Lee HO, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol*. 2016;17:80. <https://doi.org/10.1186/s13059-016-0945-9>.
144. Zhang J, Guan M, Wang Q, Zhang J, Zhou T, Sun X. Single-cell transcriptome-based multilayer network biomarker for predicting prognosis and therapeutic response of gliomas. *Brief Bioinform*. 2019;21(3):1080–1097. <https://doi.org/10.1093/bib/bbz040>.
145. Darmanis S, Sloan SA, Crote D, et al. Single-cell RNA-Seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. *Cell Rep*. 2017;21(5):1399–410. <https://doi.org/10.1016/j.celrep.2017.10.030>.
146. Filbin MG, Tirosh I, Hovestadt V, Shaw MKL, Escalante LE, Mathewson ND, et al. Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science*. 2018;360(6386):331–5. <https://doi.org/10.1126/science.aao4750>.
147. Wang Q, Tan Y, Fang C, et al. Single-cell RNA-seq reveals RAD51AP1 as a potent mediator of EGFRvIII in human glioblastomas. *Aging*. 2019;11(18):7707–22. <https://doi.org/10.18632/aging.102282>.
148. Cai J, Chen S, Yi M, et al. $\Delta Np63\alpha$ is a super enhancer-enriched master factor controlling the basal-to-luminal differentiation transcriptional program and gene regulatory networks in nasopharyngeal carcinoma. *Carcinogenesis*. 2020;41(9):1282–93. <https://doi.org/10.1093/carcin/bgz203>.
149. Chen S, Youhong T, Tan Y, et al. EGFR-PKM2 signaling promotes the metastatic potential of nasopharyngeal carcinoma through induction of FOSL1 and ANTXR2. *Carcinogenesis*. 2020;41(6):723–33. <https://doi.org/10.1093/carcin/bgz180>.
150. Fan C, Tu C, Qi P, et al. GPC6 promotes cell proliferation, migration, and invasion in nasopharyngeal carcinoma. *J Cancer*. 2019;10(17):3926–32. <https://doi.org/10.7150/jca.31345>.
151. Mo Y, Wang Y, Xiong F, Ge X, Li Z, Li X, et al. Proteomic analysis of the molecular mechanism of lovastatin inhibiting the growth of nasopharyngeal carcinoma cells. *J Cancer*. 2019;10(10):2342–9. <https://doi.org/10.7150/jca.30454>.
152. Xiong F, Deng S, Huang HB, et al. Effects and mechanisms of innate immune molecules on inhibiting nasopharyngeal carcinoma. *Chin Med J*. 2019;132(6):749–52. <https://doi.org/10.1097/cm9.000000000000132>.
153. Zhang Y, Wang D, Peng M, et al. Single-cell RNA sequencing in cancer research. *J Exp Clin Cancer Res*. 2021;40(1):81. <https://doi.org/10.1186/s13046-021-01874-1>.
154. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK, Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*. 2017;14(9):865–8. <https://doi.org/10.1038/nmeth.4380>.

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

