Proceedings of the Pakistan Academy of Sciences:
B. Life and Environmental Sciences 58(3): 19-28 (2021)

ISSN (Print): 2518-4261; ISSN (Online): 2518-427X
http://doi.org/10.53560/PPASB(58-3)663

Pakistan Academy of Sciences

Review Article

Single Cell RNA Sequencing (scRNA-Seq) as an Emerging Technology in Cancer Research

Atta Ur Rehman[1*], Abdur Rashid[2], and Ijaz Anwar[3]

[1]Department of Biomedical Sciences, Pak-Austria Fachhochschule:
Institute of Applied Sciences and Technology, Khanpur Road, Mang Haripur, Pakistan
[2]Government of Khyber Pakhtunkhwa, Department of Higher Education Archives and Libraries, Peshawar, Pakistan
[3]Centre for Applied Molecular Biology, University of the Punjab, Lahore, Pakistan

Abstract: RNA sequencing (RNA-seq) has revolutionized basic biomedical research by studying the transcriptome at high resolution, and thus it has been proved to be very successful for understanding the molecular mechanisms of cancers. For example, RNA-seq has facilitated a comprehensive and multidimensional mapping of the key genomic changes that lead to various types of cancers. Nevertheless, the heterogeneous nature of cancer tissues has always been a problem. To overcome this challenge, single-cell RNA sequencing (scRNA-seq) has emerged as the most powerful tool to characterize cancer tissues by enhancing our knowledge of transcriptome at a single-cell resolution. In addition to disentangling the heterogeneity problem, scRNA-seq has other applications such as determining the molecular mechanisms of cellular differentiation, characterizing gene expression levels, and determining rare cell types found within cancer tissue. scRNA-Seq is used, as an emerging diagnostic tool, in tertiary healthcare settings with diverse clinical applications. Thus, the utility of scRNA-Seq in a healthcare system not only provides compelling evidence about understanding cancer biology but also points towards the development of therapeutic options in the future. The purpose of this review is to educate readers about the applications of scRNA-seq in cancer research in a wider context..

Keywords: Soybean; Cancer, Gene Expression, Single Cell RNA Sequencing, Transcriptome Profiling.

1. INTRODUCTION

RNA sequencing (RNA-seq) is a powerful and important technology for understanding the underlying molecular mechanisms of cancer, thus revolutionizing basic biomedical research. By profiling gene expression levels across the entire transcriptome, RNA-seq allows cancer researchers to detect diverse genetic changes such as translocations, deletions, insertions, alternative splicing, and gene fusions. As cancer tissues are mostly characterized by different kinds of genomic changes, RNA-Seq has proven valuable for diagnosing tumors, and for further characterizing them [1]. Besides, RNA-Seq has also revolutionized other fields of biology such as research involving vision, autoimmune diseases, cardiovascular diseases, developmental/evolutionary relationships among brain cells by delineating biological processes at an unprecedented scale [2-6].

The Cancer Genome Atlas (TCGA) has been a notable example of RNA-seq application in cancer research [7]. By using genome sequencing, TCGA has quantified gene expression levels in 33 types of cancers and generated a comprehensive and multidimensional map of the key genomic changes that drive these cancers [7]. Similarly, The Encyclopaedia of DNA Elements (ENCODE) is another example of genome analysis based on RNA sequencing. ENCODE aims to identify all functional elements in human and mouse genomes through various functional genomic assays by studying transcription factors and their target genes,

Received: June 2021; Accepted: September 2021
*Corresponding Author: Atta Ur Rehman <atta.rehman@hu.edu.pk>

chromatin structures, and histone modifications [8].

These two datasets strongly improved and accelerated our understanding of genomic and molecular characteristics of cancers, which further helps to improve the prevention, diagnosis, and treatment of cancer. However, until recently, all transcriptomic studies were typically conducted on a bulk level, averaging the variable transcriptomes from millions of cells [9], even though cancer tissues are highly heterogeneous. In addition to cancer cells themselves, cancer contains other cell types including immune cells and other kinds of stromal cells. Furthermore, due to a variety of factors including genomic instability, epigenetic alterations, environmental differences, even individual cancer cells within a single tumor may be highly plastic in their molecular signatures during the development of tumor progression.

Given the extreme heterogeneity within a cancer tissue, it is not surprising that cancer can develop resistance to cancer treatments. Thus, understanding the underlying mechanisms that drive heterogeneity in cancer tissues is becoming more urgent. The development of sing-cell RNA sequencing (scRNA-seq) is a leap towards this challenge. With scRNA-seq, cancer researchers could further dissect the diversity of tumors into different molecular states [10], and once fully explored, scRNA-seq will finally help to tailor cancer treatments and to develop personalized medicine [11].

2. RNA-SEQ TOOLS IN CANCER RESEARCH

Several analytical tools have been developed to further enhance the contribution of RNA-seq to cancer research and to simplify the massive data that are being produced in RNA sequencing (Table 1). These tools have been promising in the field of cancer research as intra-tumor heterogeneity in a single cancer sample has always been a problem. Among them, CIBERSORT [12] is well-known and considered to be the most popular tool used by various cancer research groups. Essentially, CIBERSORT stands out from the other tools in being highly accurate, resistant to noise, and can characterize the cellular composition in a cancer sample. Its predecessor, ESTIMATE [13], as its name indicate, can estimate the fraction of stromal and immune cells in tumor tissue. This fraction is important in determining the tumor purity. But as time passed by, other tools such as MCP-Counter [14], xCELL [15], and most recently ICeD-T [16], were introduced to answer other questions that were previously unresolved by CIBERSORT and ESTIMATE. MCP-Counter, for example, is a tool used mainly to identify multiple immune cells in a tumor microenvironment that are involved in the prognosis of certain cancers, such as adenocarcinoma and breast cancer [14]. Similarly, xCELL is used to paint a picture of the tumor microenvironment of the disease. This includes analyzing 64 different cell types based solely on their gene signatures [15]. Lastly, ICeD-T uses tumor purity information to effectively discriminate between the aberrant group of genes from the consistent group [16].

In addition to tools that use immune cells as the quantitative parameter, there are tools designed to target or evaluate the outcome of certain treatments. For example, QuanTISeq works by quantifying the tumor immune contexture. Contexture used in this analysis is the absolute fraction based on five immune cells. The information can then be derived to determine the optimum pharmacological therapies which include a combination of pharmacological therapies [17]. Other examples include TIMER, which is an analytical tool that is based on the estimation of abundance of six tumor-infiltrating immune cells. These parameters can then be derived to determine the antitumor therapy on the disease [18]. Altogether, these analytical tools have undeniably elevated the contribution of RNA-seq in the field of cancer. Among the challenges, however, include predicting cell types that have a very low fraction and determining the gene expression of specific cell types. To this end, scRNA-seq emerged as a promising tool thus enabling scientists and clinicians to analyze transcriptomic data at a single-cell resolution.

3. APPLICATIONS OF scRNA-SEQ IN THE FIELD OF CANCER RESEARCH

A range of applications of scRNA-Seq in the field of cancer research is summarized in Figure 1 and discussed below.

3.1 Intra-tumour Heterogeneity in Cancer

Cancer is generally associated with genetic modifications (also called “mutations”) such as

Table 1. Various RNA-seq analytical tools

Analytical Tool	Application	Output Data	Advantages	References
CIBERSORT	Characterization of cell heterogeneity using RNA mixtures from nearly any tissue	Cell composition in complex tissues	High accuracy in analysis of mixture with noise and closely related cell types	[12]
ESTIMATE	Estimate the fraction of stromal and immune cells in tumor tissues	Determine the fraction of infiltrating stromal and immune cell in tumor tissues Predicting tumor purity	Able to determine the stromal score and immune score separately Has the potential to be used in other type of cancer which were not validated in the study	[13]
MCP-Counter	Robust quantification of the absolute abundance of multiple immune and stromal cell populations in the heterogeneous tissues	Quantitative analysis of the microenvironment of normal and cancer tissues via eight immune and two stromal cell populations Prognostic value associated with MCP-counter estimation	Able to quantify more than two immune cell populations Quantitatively validated Has been used in studies involving lung adenocarcinoma, colorectal, and breast cancer prognosis	[14]
TIMER	Explorations of the disease-specific clinical impact of different immune infiltrates in the tumor microenvironment	Estimation of the abundance of six tumor infiltrating immune cell types The impact of antitumor immunity on cancer immunotherapies	Do not suffer from biased estimations due to statistical co-linearity Able to analyze data derived from TCGA RNA-seq data Has been used in studies involving 23 different types of cancer	[19]
EPIC	Robust approach to determine the immune/cancer cell profile in bulk tumors	The proportion of immune and cancer cells	Applicable to most solid tumors	[20]
QuanTISeq	Quantification of the tumor immune contexture	Quantitative analysis of immune cell proportion in the form of absolute fractions based on five immune cells Provides mechanistic rationale in designing combination of pharmacological therapies	Provide prognostic values from immunoscore Has been validated in three different types of cancer Potential to be used in inflammatory and infectious diseases	[17]
xCELL	Cell type enrichment analysis from gene expression data for multiple immune and stroma cell types	Portrays the tumor microenvironment through gene signature-based method	Analyze gene signatures for 64 cell types (adaptive and innate immunity cell, hematopoietic progenitors, epithelial cells, and extracellular matrix cells) Does not discriminate between closely related cell types	[15]
Iced-T	Asist users with immune cell expression deconvolution within tumor tissues Identifies the genes whose expression in tumor samples inconsistent with reference profiles	Quantitative visualization of immune cell expression within tumor tissues Estimates cell types abundance through the use of tumor purity information	Able to perform deconvolution o the linear-scale and incorporate the beneficial properties of the log-transformation at the same time Able to differentiate genes into two different groups: aberrant group and consistent group	[16]
Slingshot	Identifies multiple lineage trajectories based on noisy sir	To model branching lineage trajectories using scRNA-seq data	Slingshot is flexible to the type dimensionality reduction, normalization, and clustering procedures	[21]

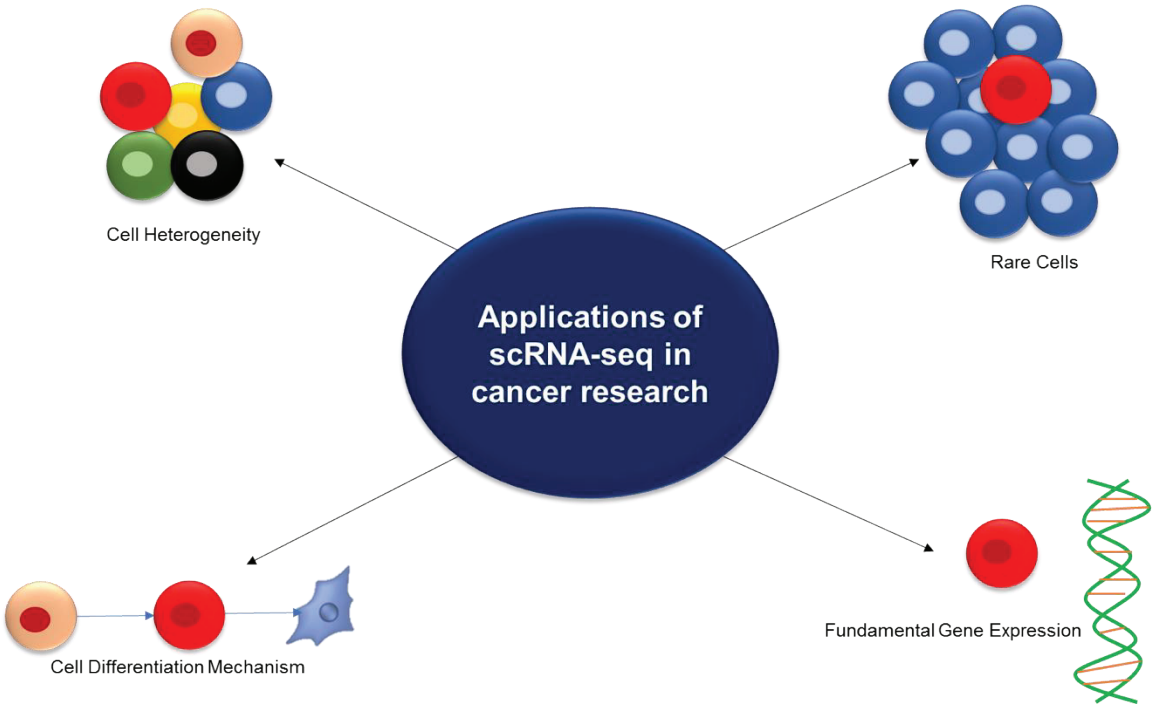

Fig.1. Applications of scRNA-seq in cancer research

the alteration of chromosome number and/or structural chromosomal abnormalities [22]. Such modifications of the genome can arise in different organs and lead to different forms of cancers, although the known mechanisms (or hallmarks) of cancers seem to be recurrent [23]. To add complexity to the situation, tumors are not uniform masses of cells presenting the same characteristics in all situations. Rather, their composition varies from one patient to another [24], and thus comprises an array of cellular types, many of them potentially contributing to the progression of cancer (Pietras and Ostman, 2010) [25]. This intratumoral heterogeneity is a real challenge from a research standpoint because it raises the possibility that every cell type could, in principle, behave differently and express specific genes. It is not clear yet how the interaction between these cell types affects cancer progression, however stromal cells, for example, seem to be involved in the process [26]. It was also proposed that some normal cells could be recruited to provide a micro-environment for cancer cells [27].

Analyses of the tumors transcriptome are usually performed in bulk. Bulk analysis performs well to differentiate through different types of tumors, such methods are, however, not informative enough keeping in view the intratumoral heterogeneity [28]. This is because the observed signal is only the compilation of profiles from all cell types, meaning that the transcription of a specific cell type is lost. Thanks to the recent advances in single-cell RNA sequencing that have allowed a very fine detection of gene expression profiles at a single-cell resolution. With scRNA-seq, it is now possible to characterize gene expression levels of virtually all cell types, with an exception for those genes whose expression rates are extremely low. Many variants of the scRNA-seq have been developed to better study cellular subpopulations. Most of these variants have been compiled in Table 2. An example of the potential application of these methods is the work performed by Klein and colleagues [29], who developed the droplet barcoding method as a way of identifying the transcriptome of individual mouse embryonic stem cells. With this method, the cells are initially captured individually in a droplet followed by barcoding and sequencing. The barcoding allows for the easy association of a cDNA to a cell in particular.

3.2. Identification of Rare Cell Types in Cancer

Among the variety of cellular types in cancer tissue, particular cells play an important role concerning cancer though their number is limited. These include rare cell types such as cancer stem cells from which

the tumor originates [30]. These cells are essential for the proliferation and the maintenance of the tumor and compared to thousands of other cellular types that could not lead to the formation of a tumor, the number of these rare tumorigenic cells remains to be in hundreds [31]. As long as specific markers for detection of these rare cell types are integrated into the analysis, traditional bulk sequencing remains to be inadequate in this scenario because the signal of a handful of rare cells would most likely be lost among the ones from all other cells types. Even if the signal was noticeable for the tumor as a whole, it would still be impossible to trace their particular lineage.

Instead, scRNA-seq could effectively discriminate among rare cell types based upon their transcriptional activity. Since these rare cells are anticipated to be most challenging in terms of treatment, understanding their correct transcriptomic profiling and their underlying molecular mechanisms would possibly pave the way for the development of cancer therapy in the future. As an example, Cao and colleagues have successfully developed single-cell combinatorial indexing RNA sequencing (sci-RNA-seq, Figure 1), a technique that allowed them to cluster all the cell types in *C. elegans*, a multicellular organism [32]. Among these cell types, Cao and colleagues recovered rare neuronal cells in the L2 larval stage. Thus, the application of this method for early detection of rare tumorigenic cell types such as stem cells are very promising and could lead to a better understanding of the initiation of cancers.

3.3 Cellular Differentiation and Developmental Mechanisms

Many tissues continuously grow and regenerate through a coordinated differentiation activity of stem cells. Yet, understanding the molecular mechanisms that control cellular differentiation and fate determination in stem cells is still one of the fundamental questions in developmental and stem cell biology [33]. To answer this question, it is imperative not only to trace the lineage of a cell or a group of cells that underlie these processes but to fully dissect all the intermediate stages as well as the endpoints along the path of differentiation. In other words, understanding relationships among these lineages will illuminate the fundamental mechanisms underlying normal development, and can provide insight into the development of different pathologies including cancer. In general, lineage relationships are experimentally revealed through fate mapping methods, and once fate mapping is carried out at a single-cell resolution, it is known as lineage tracing [34]. Traditionally, this goal was achieved by: (1) direct examination of dividing cells in a transparent embryo, (2) through staining cells with dyes in opaque embryos, and (3) via distinct pigmentation patterns or cellular appearance [33]. Though versatile, these methods had certain inherent limitations. For instance, these methods were providing sufficient information about fate maps of the labeled cells but failed to explain the lineage trajectories, transient intermediates, and lineage branch points during the whole differentiation process.

With the advent of high-throughput sequencing approaches, it is now possible to define the molecular status of cells and to fine map their path along a differentiation axis. One such technique is scRNA-seq that allows for discrimination cellular heterogeneity. Based on single-cell transcriptomic data, scRNA-seq could be used efficiently in stem cells research to trace the lineage trajectories, cellular intermediates, and branch points along the axis of differentiation [35-38]. To achieve these goals, the only requirement for researchers would be to collect cell samples at different time points. Interestingly, by integrating scRNA-seq with more recent approaches like CRISPR- or transposon-mediated modification of DNA barcodes, it is possible to fine map the lineage hierarchies and to identify endpoint cells based on their transcriptomic profiles without previous knowledge about cell types [39-42]. Taking advantage of a single cell resolution, this integrative approach now provides a robust method for deconstructing how individual stem cells maintain tissues. Finally, by integrating single-cell RNA sequencing with clonal lineage tracing, it is possible to delineate the molecular mechanisms that control the differentiation of stem cells into tissues, organs, and ultimately an organism.

3.4 Fundamental Characteristics of Gene Expression

Growing evidence suggests that gene transcription

Table 2. Various techniques developed for scRNA-seq

Protocols	Method	Throughput	Cost	Reference
Indrop	Droplet microfluidic-based	High throughput (4,000-12,000s cells per hour) Low noise Low cell bias	High cost Low efficiency in cells of genes with transcript abundance lower than 20-50 transcripts Low detection rate	[29, 49]
Drop-seq	Droplet-based	Relatively high throughput (10,000 cells per day for 6.5 cents per cell) High efficiency	Venerable to impurities in the cell isolation process Low detection rate	[49]
CytoSeq	Microwell-based	High throughput (10,000 - 100,000s cells per day) High detection rate	High cost Restricted to pre-defined set of genes	[50, 51]
Sci-RNA-Seq	Combinatorial indexing	High detection rate High throughput (50,000 cells) Reduced batch effect	Relatively high cost Limited to copy number variant	[32]
Smartseq2	PCR-plate based (Full-length RNA sequencing)	Low cost	Relatively low throughput (100s cells) Labor intensive	[49, 52]
CEL-seq and CEL-seq2	Tag-based sequencing	Strand specificity High barcoding efficiency Low cost	Relatively low throughput Low sensitivity for low expressed transcripts	[53]
MARS-seq	Tag-based sequencing	Molecular counting High degree of multiplexing Low cost	Requires enrichment to detect rare cell population	[54]
Fluidigm C1	Microfluidic chips	High detection rates Highly-accessible and affordable	High cost-efficiency Low throughput (96 cells) Sub-optimal performance on primary cells	[55]

does not follow a continuous path. Rather, it completes in several short steps called "transcriptional bursts". Under the control of various stochastic processes, these transcriptional bursts finally lead to extreme transcriptional heterogeneity as seen between cells. Recent studies also suggest that cells preferentially express a single allele or a single splice isoform resulting in a process known as differential gene expression. Additionally, a single genetic locus can undergo diverse splicing patterns resulting in multiple isoforms, each having different transcriptional start sites or polyadenylation sites [43]. Moreover, processes such as dosage compensation or X-chromosome inactivation in most diploid female animals ensure expression of only one allele at a time, either maternal or paternal [44]. Historically, much of our understanding of the transcriptome is based on bulk studies conducted on cell populations. It is now, however, well established that the homogeneous cell populations *in vitro* or *in vivo* are quite heterogeneous in terms of expression patterns owing to both intrinsic and extrinsic factors [45].

Understanding the gene expression pattern at a single cell level is important for two main reasons. First, recent findings suggest that the expression of a transcript in all individual cells of a particular tissue or organ is only facultative rather than obligatory. This was previously not anticipated. Similarly, the adjoining cells that share the same microenvironment are deemed to differentially

express a transcript for unknown reasons. Studying gene expression patterns at a single-cell resolution is therefore important for the identification of co-regulated gene modules and to define gene-regulatory networks [46]. Though several methods have been previously used to achieve these goals, scRNA-seq has been exceptional in this respect. This method allows genome-wide profiling of both coding and non-coding cellular transcriptome with higher efficiency. Thus, even minor changes in gene expression could be detected with greater sensitivity. Moreover, scRNA-seq has the potential to identify previously uncharacterized transcript isoforms, novel exons, SNPs, or mutations [11]. This method has specific applications when allele-specific expression of either autosomal or X-linked genes is required [47]. Finally, Single-cell RNA sequencing has greatly revolutionized the fields of medicine and biology by delineating biological processes at an unprecedented scale and resolution in the recent past.

4. TECHNIQUES DEVELOPED FOR scRNA-SEQ

With such a wide range of applications and impact on the field of cancer research, various methods and protocols for scRNA-seq have been developed (Table 2). These techniques are used to serve a range of purposes: from understanding the tumor architecture to identifying specific cell types in the tumor microenvironment. These methods range from a high throughput method, which can easily sequence 10,000 cells to more affordable methods which sequence 100 cells in a single run. Some can complete a run in a matter of hours, while others might take a whole day. Also, some of the methods need a few days to prepare while others which has a higher cost might have a less hands-on approach. A recent study has suggested Drop-seq as the most cost-efficient method among its contemporaries for quantification of the transcriptome at a large scale [48]. With so many options already in place, it is intriguing to see how the field of cancer research is going to grow soon, especially with the help of scRNA-seq.

5. UTILITY OF scRNA-SEQ IN HEALTHCARE SETTINGS

scRNA-Seq is used, as an emerging diagnostic tool, in tertiary healthcare settings with diverse clinical applications. For example, scRNA-Seq has been crucial in the discovery of differentially expressed genes in different cancer types [56], and has helped track the intra-tumoral heterogeneity and in visualizing tumor microenvironment [57]. In addition, studies suggest that scRNA-Seq has been key to the detection of many cellular and molecular therapeutic targets [58], and for the identification of biological features representing as biomarkers of clinical outcomes [59]. Indeed, scRNA-Seq has also enhanced our understanding of the immune cell heterogeneity and the infiltrating T Cells [60-61]. Thus, the utility of scRNA-Seq in a healthcare system not only provides compelling evidence about understanding cancer biology but also points towards the development of therapeutic options in the future.

6. LIMITATIONS OF scRNA-SEQ

scRNA-Seq has numerous limitations, and they are beyond the scope of this review. However, a few of them include excessive noise and complexity due to technical or biological reasons, transcript coverage bias, also known as high dropouts or low capture efficiency, high variation in data output, and computationally challenging data analysis [62]. For an in-depth insight on the subject, Lähnemann and colleagues [63] have recently highlighted eleven major challenges encountered while dealing with scRNA-Seq data.

7. CONCLUSION

In summary, big advancements such as high throughput transcriptome profiling coupled with the availability of robust analytical tools have greatly revolutionized the field of cancer research in the recent past. The sole purpose was (1) to fully characterize cellular components in a tumor microenvironment (2) to understand the molecular pathophysiology of cancer, and (3) to mitigate cancer-related morbidity and mortality through the development of novel therapeutics. Among them, scRNA-seq have been developed to suit any scale of research in cancer, whether it is in clinical settings, institutional or private research.

8. CONFLICT OF INTEREST

The authors declared no conflict of interest.

9. REFERENCES

1. M. Wan, J. Wang, X. Gao, and J. Sklar. RNA Sequencing and its Applications in Cancer Diagnosis and Targeted Therapy. *North American Journal of Medicine and Science* 7(4): (2014).
2. A.P. Voigt, N.K. Mullin, E.M. Stone, B.A. Tucker, T.E. Scheetz, and R.F. Mullins. Single-cell RNA sequencing in vision research: Insights into human retinal health and disease. *Progress in Retinal and Eye Research* 83: 100934 (2021).
3. M. Zhao, J. Jiang, M. Zhao, C. Chang, H. Wu, and Q. Lu. The Application of Single-Cell RNA Sequencing in Studies of Autoimmune Diseases: a Comprehensive Review. *Clinical Reviews in Allergy and Immunology* 60(1): 68-86 (2021).
4. A.M. Ranzoni, A. Tangherloni, I. Berest, S.G. Riva, B. Myers, P.M. Strzelecka, J. Xu, E. Panada, I. Mohorianu, J. B. Zaugg, and A. Cvejic. Integrative Single-Cell RNA-Seq and ATAC-Seq Analysis of Human Developmental Hematopoiesis. *Cell Stem Cell* 28(3): 472-487 (2021).
5. F. Iqbal, A. Lupieri, M. Aikawa, and E. Aikawa. Harnessing Single-Cell RNA Sequencing to Better Understand How Diseased Cells Behave the Way They Do in Cardiovascular Disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* 41(2): 585-600 (2021).
6. E.J. Armand, J. Li, F. Xie, C. Luo, and E.A. Mukamel. Single-Cell Sequencing of Brain Cell Transcriptomes and Epigenomes. *Neuron* 109(1): 11-26 (2021).
7. K. Tomczak, P. Czerwińska, and M. Wiznerowicz. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary oncology* 19(1): A68 (2015).
8. E.P. Consortium. The ENCODE (ENCyclopedia of DNA elements) project. *Science* 306 (5696): 636-640 (2004).
9. S. Zhu, T. Qing, Y. Zheng, L. Jin, and L. Shi. Advances in single-cell RNA sequencing and its applications in cancer research. *Oncotarget* 8(32): 53763-53779 (2017).
10. P. Angerer, L. Simon, S. Tritschler, F.A. Wolf, D. Fischer, and F.J. Theis. Single cells make big data: New challenges and opportunities in transcriptomics. *Current Opinion in System Biology* 4: 485–91 (2017).
11. S. Picelli. Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biology* 14(5): 637-650 (2017).
12. A.M. Newman, C.L. Liu, M.R. Green, A.J. Gentles, W. Feng, Y. Xu, C.D. Hoang, M. Diehn, and A.A. Alizadeh. Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods* 12(5): 453-7 (2015).
13. K. Yoshihara, M. Shahmoradgoli, E. Martínez, R. Vegesna, H. Kim, W. Torres-Garcia, V. Treviño, H. Shen, P.W. Laird, D.A. Levine, and S.L. Carter. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature Communications* 4(1): 1-11 (2013).
14. E. Becht, N.A. Giraldo, L. Lacroix, B. Buttard, N. Elarouci, F. Petitprez, J. Selves, P. Laurent-Puig, C. Sautès-Fridman, W.H. Fridman, and A. De Reyniès. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biology* 17(1): 1-20 (2016).
15. D. Aran, Z. Hu, and A.J. Butte. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biology* 18(1): 1-14 (2017).
16. D.R. Wilson, C. Jin, J.G. Ibrahim, and W. Sun. Iced-t provides accurate estimates of immune cell abundance in tumor samples by allowing for aberrant gene expression patterns. *Journal of the American Statistical Association* 115(531): 1055-1065 (2020).
17. P. Charoentong, F. Finotello, M. Angelova, C. Mayer, M. Efremova, D. Rieder, H. Hackl, and Z. Trajanoski. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell reports* 18(1): 248-262 (2017).
18. B. Li, E. Severson, J.C. Pignon, H. Zhao, T. Li, J. Novak, P. Jiang, H. Shen, J.C. Aster, S. Rodig, and S. Signoretti. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome biology* 17(1): 1-16 (2016).
19. M. Barron and J. Li. Identifying and removing the cell-cycle effect from single-cell RNA-Sequencing data. S*cientific reports* 6(1): 1-10 (2016).
20. J. Racle, K. de Jonge, P. Baumgaertner, D.E. Speiser, and D. Gfeller. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife* 6(13): (2017).
21. K. Street, D. Risso, R.B. Fletcher, D. Das, J. Ngai, N. Yosef, E. Purdom, and S. Dudoit. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC genomics* 19(1): 1-16 (2018).
22. C. Lengauer, K.W. Kinzler, and B. Vogelstein. Genetic instabilities in human cancers. *Nature*

396(6712): 643-9 (1998).
23. D. Hanahan and R.A. Weinberg. Hallmarks of cancer: the next generation. *Cell* 144 (5): 646-74 (2011).
24. R.A. Burrell, N. McGranahan, J. Bartek, and C. Swanton. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501(7467): 338-45 (2013).
25. K. Pietras and A. Ostman. Hallmarks of cancer: interactions with the tumor stroma. *Experimental Cell Research* 316(8): 1324-31 (2010).
26. M.I. Koukourakis, A. Giatromanolaki, A.L. Harris, and E. Sivridis. Comparison of metabolic pathways between cancer cells and stromal cells in colorectal carcinomas: a metabolic survival role for tumor-associated stroma. *Cancer Research* 66(2): 632-637 (2006).
27. C. Feig, A. Gopinathan, A. Neesse, D.S. Chan, N. Cook, and D.A. Tuveson. The pancreas cancer microenvironment. *Clinical Cancer Research* 18(16): 4266-76 (2012).
28. A.P. Patel, I. Tirosh, J.J. Trombetta, A.K. Shalek, S.M. Gillespie, H. Wakimoto, D.P. Cahill, B.V. Nahed, W.T. Curry, R.L. Martuza, and D.N. Louis. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190): 1396-1401 (2014).
29. A.M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D.A. Weitz, and M.W. Kirschner. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 161(5): 1187-1201 (2015).
30. S.K. Singh, I.D. Clarke, M. Terasaki, V.E. Bonn, C. Hawkins, J. Squire, and P.B. Dirks. Identification of a cancer stem cell in human brain tumors. *Cancer research* 63(18): 5821-5828 (2003).
31. M. Al-Hajj, M.S. Wicha, A. Benito-Hernandez, S.J. Morrison, and M.F. Clarke. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 100(7): 3983-8 (2003).
32. J. Cao, J.S. Packer, V. Ramani, D.A. Cusanovich, C. Huynh, R. Daza, X. Qiu, C. Lee, S.N. Furlan, F.J. Steemers, and A. Adey. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357(6352): 661-667 (2017).
33. K. Kretzschmar, and F.M. Watt. Lineage tracing. Cell 148(1-2): 33-45 (2012).
34. M.B. Woodworth, K.M. Girskis, and C.A. Walsh. Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nature Review Genetics* 18(4): 230-244 (2017).
35. R.B. Fletcher, D. Das, and J. Ngai. Creating Lineage Trajectory Maps Via Integration of Single-Cell RNA-Sequencing and Lineage Tracing: Integrating transgenic lineage tracing and single-cell RNA-sequencing is a robust approach for mapping developmental lineage trajectories and cell fate changes. *Bioessays* 40(8): e1800056 (2018).
36. L. Gadye, D. Das, M.A. Sanchez, K. Street, A. Baudhuin, A. Wagner, M.B. Cole, Y.G. Choi, N. Yosef, E. Purdomand, and S. Dudoit. Injury activates transient olfactory stem cell states with diverse lineage capacities. *Cell Stem Cell* 21(6): 775-790 (2017).
37. T.J. Nowakowski, A. Bhaduri, A.A. Pollen, B. Alvarado, M.A. Mostajo-Radji, E. Di Lullo, M. Haeussler, C. Sandoval-Espinosa, S.J. Liu, D. Velmeshev, and J.R. Ounadjela. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* 358(6368): 1318-1323 (2017).
38. I. Stevant, Y. Neirijnck, C. Borel, J. Escoffier, L.B. Smith, S.E. Antonarakis, E.T. Dermitzakis, and S. Nef. Deciphering cell lineage specification during male sex determination with single-cell RNA sequencing. *Cell reports* 22(6): 1589-1599 (2018).
39. B. Raj, D.E. Wagner, A. McKenna, S. Pandey, A.M. Klein, J. Shendure, J.A. Gagnon, and A.F. Schier. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nature biotechnology* 36(5): 442-450 (2018).
40. B. Spanjaard, B. Hu, N. Mitic, P. Olivares-Chauvet, S. Janjuha, N. Ninov, and J.P. Junker. Simultaneous lineage tracing and cell-type identification using CRISPR–Cas9-induced genetic scars. *Nature biotechnology* 36(5): 469-473 (2018).
41. D.E. Wagner, C. Weinreb, Z.M. Collins, J.A. Briggs, S.G. Megason, and A.M. Klein. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* 360(6392): 981-987 (2018).
42. A. Alemany, M. Florescu, C.S. Baron, J. Peterson-Maduro, and A. van Oudenaarden. Whole-organism clone tracing using single-cell sequencing. *Nature* 556(7699): 108-112 (2018).
43. A.K. Shalek, R. Satija, X. Adiconis, R.S. Gertner, J.T. Gaublomme, R. Raychowdhury, S. Schwartz, N. Yosef, C. Malboeuf, D. Lu, and J.J. Trombetta. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 498(7453): 236-240 (2013).

44. Q. Deng, D. Ramskold, B. Reinius, and R. Sandberg. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* 343(6167): 193-6 (2014).
45. M. Bengtsson, A. Stahlberg, P. Rorsman, and M. Kubista. Gene expression profiling in single cells from the pancreatic islets of Langerhans reveals lognormal distribution of mRNA levels. *Genome Research* 15(10): 1388-92 (2005).
46. A. Haque, J. Engel, S.A. Teichmann, and T. Lonnberg. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine* 9(1): 1-12 (2017).
47. J.A. Benitez, S. Cheng, and Q. Deng. Revealing allele-specific gene expression by single-cell transcriptomics. *The International Journal of Biochemistry & Cell Biology* 90: 155-160 (2017).
48. C. Ziegenhain, B. Vieth, S. Parekh, B. Reinius, A. Guillaumet-Adkins, M. Smets, H. Leonhardt, H. Heyn, I. Hellmann, and W. Enard. Comparative analysis of single-cell RNA sequencing methods. *Molecular cell* 65(4): 631-643 (2017).
49. T.S. Andrews and M. Hemberg. Identifying cell populations with scRNASeq. *Molecular Aspects of Medicine* 59: 114-122 (2018).
50. X. Fan, X. Zhang, X. Wu, H. Guo, Y. Hu, F. Tang, and Y. Huang. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome biology* 16(1): 1-17 (2015).
51. A.K. Bjorklund, M. Forkel, S. Picelli, V. Konya, J. Theorell, D. Friberg, R. Sandberg, and J. Mjösberg. The heterogeneity of human CD127+ innate lymphoid cells revealed by single-cell RNA sequencing. *Nature immunology* 17(4): 451-460 (2016).
52. S. Picelli, A.K. Bjorklund, O.R. Faridani, S. Sagasser, G. Winberg, and R. Sandberg. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nature Methods 10(11): 1096-8 (2013).
53. T. Hashimshony, F. Wagner, N. Sher, and I. Yanai. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Reports* 2(3): 666-73 (2012).
54. D.A. Jaitin, E. Kenigsberg, H. Keren-Shaul, N. Elefant, F. Paul, I. Zaretsky, A. Mildner, N. Cohen, S. Jung, A. Tanay, and I. Amit. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 343(6172): 776-779 (2014).
55. V. Svensson, K.N. Natarajan, L.H. Ly, R.J. Miragaia, C. Labalette, I.C. Macaulay, A. Cvejic, and S.A. Teichmann. Power analysis of single-cell RNA-sequencing experiments. *Nature methods* 14(4): 381-387 (2017).
56. Z. Chen, M. Zhao, M. Li, Q. Sui, Y. Bian, J. Liang, Z. Hu, Y. Zheng, T. Lu, Y. Huang, and C. Zhan. Identification of differentially expressed genes in lung adenocarcinoma cells using single-cell RNA sequencing not detected using traditional RNA sequencing and microarray. *Laboratory Investigation* 100(10):1318-1329 (2020).
57. M.B. Lopes and S. Vinga. Tracking intratumoral heterogeneity in glioblastoma via regularized classification of single-cell RNA-Seq data. *BMC Bioinformatics* 21(1): 1-12 (2020).
58. H.W. Lee, W. Chung, H.O. Lee, D.E. Jeong, A. Jo, J.E. Lim, J.H. Hong, D.H. Nam, B.C. Jeong, S.H. Park, and K.M. Joo. Single-cell RNA sequencing reveals the tumor microenvironment and facilitates strategic choices to circumvent treatment failure in a chemorefractory bladder cancer patient. *Genome medicine* 12: 1-21 (2020).
59. A. Maynard, C.E. McCoach, J.K. Rotow, L. Harris, F. Haderk, D.L. Kerr, A.Y. Elizabeth, E.L. Schenk, W. Tan, A. Zee, and M. Tan. Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. *Cell* 182(5): 1232-1251 (2020).
60. E. Papalexi and R. Satija. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nature Review Immunology* 18(1): 35-45 (2018).
61. C. Zheng, L. Zheng, J.K. Yoo, H. Guo, Y. Zhang, X. Guo, B. Kang, R. Hu, J.Y. Huang, Q. Zhang, and Z. Liu. Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. *Cell* 169(7): 1342-1356 (2017).
62. G. Chen, B. Ning, and T. Shi. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Frontiers in Genetics* 10: 317 (2019).
63. D. Lahnemann, J. Köster, E. Szczurek, D.J. McCarthy, S.C. Hicks, M.D. Robinson, C.A. Vallejos, K.R. Campbell, N. Beerenwinkel, A. Mahfouz, and L. Pinello. Eleven grand challenges in single-cell data science. *Genome biology* 21(1): 1-35 (2020).