International Journal of Multidisciplinary Education and Research
www.educationjournal.in
ISSN: 2455-4588
Received: 30-08-2021, Accepted: 15-09-2021, Published: 01-10-2021
Volume 6, Issue 4, 2021, Page No. 1-7

Differential gene expression analysis by single cell RNA sequencing reveals insight to genetic diseases

Sudheer Menon
Department of Bioinformatics, Bharathiar University, Coimbatore, Tamil Nadu, India

Abstract
World's population is increasing at a highest rate ever in the history. In the past people lived a very short lifespan due to prevailed diseases. Now a days, people are enjoying modern health facilities and living a long life. It has become possible due to the services of science for mankind. Single cell RNA sequencing has led us to revolutionize the field of medicine by making it possible to study the role of genetic expression in causing different diseases. Gene expression has unveiled the possible causes of genetic and infectious diseases. The prime objective of the study was to devise new ways and methods to explicit the role of gene expression to reveal the insights of several genetic diseases. Total 12 rats were used as model organism and were reared under laboratory conditions. Rats were slaughtered and cell culture was obtained. Laser Capture Microdissection was used to extract single cell from cell culture. 6 pieces of 12 micrometer size were obtained and further processed to study the gene expression. The results showed that there was a significant involvement of genes expression in causing genetic diseases like cancer, phenylketonuria, sickle cell anemia and.

Keywords: RNA sequencing, genetic diseases

Introduction

The world's population is expected to age rapidly in most of the regions and people are living a longer life. This situation is similar in Korea and the country became an aging society (elderly population 7% of the total population) in 2000 (United Nations, 2018). In 2018, Korea became an aged society (defined as an elderly population 14% of the total population), and by 2026 will be a super-aged society (Greco *et al*., 2019). In elderly populations, there exists a high prevalence of many chronic diseases including osteoporosis and sarcopenia. Although osteoporosis-related studies involving the elderly population have reported clinical outcomes and established treatment guidelines in the last decades; in recent years, increased attention has been given to sarcopenia and related studies (Lo *et al*., 2020). In addition, sarcopenia has been the International Classification of Disease, Tenth Revision, Clinical Modification (ICD-10-CM) as M62.84 since October 1, 2016 (Wang *et al*., 2017) [3]. The identification of genes expressed in different tissues leads to a better understanding of molecular mechanisms that are relevant for animal production. Ribonucleic acid sequencing (RNA-seq) allows a wider gene expression analysis because the total RNA from a given tissue can be sequenced (Wang at al., 2009) [4]. Different statistical methods are available for performing RNA-seq differential gene expression analysis in a range of biological experiments (Robinson *et al*., 2010a). Over the past two decades, vision researchers have employed transcriptomics to study both normal physiology and disease within ocular tissues. Early experiments using serial analysis of gene expression (SAGE) (Bowes Rickman *et al*., 2006) and microarrays (Wagner *et al*., 2013) [10] catalogued RNA expression in the human retina and retinal pigment epithelium (RPE)/choroid. As sequencing technology advanced, whole transcriptome expression analysis in these tissues was achieved with RNA sequencing, which captured sequence-level information including splicing and other transcriptional events (Yoshida *et al*., 2002). More recently, single-cell RNA sequencing has allowed for the transcriptomic profiling of individual cells, improving the ability to study gene expression in many cell types comprising complex ocular tissues. However, these technological advancements in sequencing have been coupled with increasing data analysis complexity. Numerous bioinformatic tools aid in mapping, clustering, and visualizing single-cell data, yet accessing results can be challenging g, particularly for those without a bioinformatics background (Luecken and Theis, 2019).

Oncogenesis is a chronic, multistep pathogenesis that involves a number of genetic and epigenetic alterations. Growing evidence suggests that genomic alterations acquired during the oncogenesis help the ability of tumor cells to escape immune surveillance (Zhang *et al*., 2021) [17]. Our previous studies performed whole-exome sequencing (WES) in a larger number of patien ts with GBC and identified that ErbB signaling is the most extensively mutated pathway that mediates anti-tumor immunity, and elevated ErbB mutation levels are correlated with a worse outcome of GBC. However, these studies generally relied on the data from bulk tumor profiling, limiting their ability to accurately decipher the tumor heterogeneity mechanism in which the genomic variations of individual cells expectedly contribute to tumor transformation and poor prognosis. Single-cell RNA sequencing (scRNA-seq) brings single-cell level resolution to the analysis of transcriptomics. The technique has been applied in many areas, such as novel cell population discovery, cell heterogeneity dissection, and cell lineage construction (Hwang *et al*., 2008). There are two main quantification schemes for scRNA-seq: read count and unique molecular

identifier (UMI) count (Serena and Trapnell, 2018). The UMI count has the advantage of avoiding application biases introduced by sequencing library construction, which can be approximated by a negative binomial model. As with other high-dimensional data, accounting for the batch effects in an analysis is critical for revealing the real biological effects (Klein *et al*., 2015). There are several techniques devised for the study of genome or transcriptome for a cell at single cell stage. Single cell RNA sequencing is one of the most widely used techniques around the world. This technique is used to examine the cell behavior and other properties of a cell based on modern needs. Microarray analyses suggest that Tang *et al* firstly used this modern technique and he is regarded as a pioneer for this technique. Because of its wide implications and accuracy, this method is being efficiently used in determining of the developmental processes in humans. This technique is implicated to study the immune responses and disease controlling behavior of a cell inside a human body. This technique has led us to strengthen our mindset regarding pathogenicity and controlling the expression of a gene at a single cell stage. This helps us to understand the regulation of a network that controls embryonic development of humans.

Study of Batch Effect

Several methods have been proposed to account for known batch effects in DE analysis in scRNA-seq data by incorporating batch variables as covariates in a regression model (Cole *et al*., 2019) [7]. Other approaches have been developed to directly output a batch corrected matrix for downstream analysis, mostly for visualization/ clustering (Tung *et al*., 2017) [22]. Although several methods (Com-Bat, MNN Correct, Zinbwave, and scMerge (Soneson and Robinson, 2017). Robinson achieved relatively good performance in a limited comparison with others, their performance in DE analysis has not been systematically evaluated. he aggregation-based methods pool all cells from a batch to produce a pseudo-bulk sample and then analyze the pooled data by using approaches designed for bulk RNA-seq. Nested fixed-effect methods treat the batch effects as fixed effects nested within each group and then test the group effects for each gene. Alternatively, the batch effects can be modeled in mixed effects models, in which all cells from each batch share a random effect. Although the nested fixed effect models and nested mixed-effect models were designed for scRNA-seq, they belong to the single-gene based methods, which ignore potential common information shared among all genes, which in turn might result in a loss of power (Haghverdi *et al*., 2018; Risso *et al*., 2018) [30, 31, 38]. Most scRNA-seq platforms produce either read count or UMI count based gene expression matrices. Although a high abundance of zeroes in the expression matrix is common with both schemes, we have shown that the UMI count can be modeled by simpler models. Moreover, the negative binomial model is a good approximation model and zero-inflated models are not needed for UMI counts (Lin *et al*., 2019). The advent of single-cell RNA sequencing (scRNA-seq) technology revolutionized transcriptomics through generating gene expression data at the single cell resolution level. It has numerous advantages over bulk RNA-seq technology, which only characterize the global expression dynamics of genes in a tissue sample, while ignoring the inherent cell-cell heterogeneity. Thus, it is pertinent to assess the variability that exists among the cells in a tissue sample as this is crucial to understand the complexity and dynamics of biological processes such as embryogenesis, cance, etc. Through scRNA-seq technology, expression is quantified by mapping reads to a reference genome followed by counting the number of reads mapped to each gene.

Here, individual transcript molecules are attached with a Unique Molecular Identifier (UMI) tag; subsequently, counting the UMIs usually yields the number of transcripts for each gene in a cell. Further, huge amounts of UMI count data are generated for several thousand(s) of genes across thousand(s) of cells and subsequently deposited in public domain databases by researchers across the globe. Hence, it is necessary to develop new, and innovative statistical approaches and tools for such data analysis to harness the potential of this new technology.

This technique utilizes RNA to determine the role of a single cell in development. This starts with the isolation of a healthy and intact RNA fragment. Three approaches are utilized in this regard. These include RNA fixation, Optimal cutting temperature approach and study of immediately frozen tissues. One of the major outcomes associated with this technique is, it produces a high quality and intact RNA. One thing should be considered very important that is the storage of tissues which should be done at −70°C followed by various protocols.

Small amounts of the mRNA molecules and imperfect procedures for capturing them in individual cells lead to dropout events, i.e., genes show zero or very low expression, even though they are expressed in cells. Further, it is well established that the capture rates vary between cells for a given scRNA-seq protocol, and this is a major source of unwanted technical variation that adds to the dropout events. Addition of UMIs during the library preparation step reduces the amplification bias but has no effects on dropout events. Further, the dropout events add more zeros to the output data and can be categorized as either true/biological zeros (gene is not expressed in the cell), or false/technical zeros (gene is expressed but not detected).

Next generation sequencing technology RNA sequencing has become an important technology in the comprehensive analysis of disease transcriptomes and enabling researchers to investigate living systems on an unprecedented scale (. However, the gene expression profiles following over-expressed ZNF750 in OSCC were not studied yet. Therefore, in this context, RNA sequencing was used to investigate gene expression profiles of over-expressed ZNF750 in OSCC cell line CAL-27 cell, to understand the mechanism of anti-tumor effect of ZNF750 on OSCC (Kumar *et al*., 2017) [37]. presence of higher proportions of zeros and technical noise in scRNA-seq data can severely affect the performance of downstream Differential Expression (DE) analysis. Bulk RNA-seq DE methods such as edgeR, and DESeq2 have been used extensively for DE analysis of scRNA-seq data.

Materials and Methods

Cell Culture

Single cell RNA sequencing starts with the promulgation of cell culture. A cell culture is developed by using the tissues obtained from the model organism. In most cases, rat is used as a model organism. The fresh and healthy tissues are obtained after sacrificing the rat by subjecting to anesthesia.

In this process, the brain of rat is removed to prevent from further complications. Finally, the tissues are then processed to obtain a single cell.

Single Cell Isolation

Presently, there are various techniques available to obtain a single cell from a tissue i.e. Magnetic Activated Cell Sorting, Microfluidics and Laser Capture Microdissection. The most used method is Laser Capture Microdissection. These methods are categorized into two categories based on the method of cell isolation. These categories are both high and low throughput methods for the cell extraction from the tissues. Laser Capture Microdissection belongs to low-throughput category. It utilizes a microscope to isolate a single cell followed by visual examination. It clearly reveals the morphology of a cell stained with a specific dye. This technique requires specialized skills. If an unexperienced person performs this experiment, it will spoil the quality of the specimen. This technique is implied for the cutting and extraction of a single cell from the already fixed tissues of the sample with formaldehyde. There are several advantages associated with this technique. Firstly, it ensures the accuracy and maintains the versatility. Secondly, it is a fast speed method requiring minimal time. One of the disadvantages associated with this technique is the isolated cell's RNA may be damaged. To prevent from damaging, the RNA sequences are fixed by using various chemicals. This not only ensures health but also increases the viability of these RNA sequences.

RNA Extraction

Six sections of RNA having a size of 12 μm were extracted from the cell from each sample. It was further processed, and the processed material was placed in a centrifuge for centrifugation. Centrifugation yielded healthy and active RNA segments for further experimentation. This whole experimentation was done under controlled conditions to prevent from any complications that can spoil the quality of RNA. These RNA sections were then stored at a temperature of −80°C.

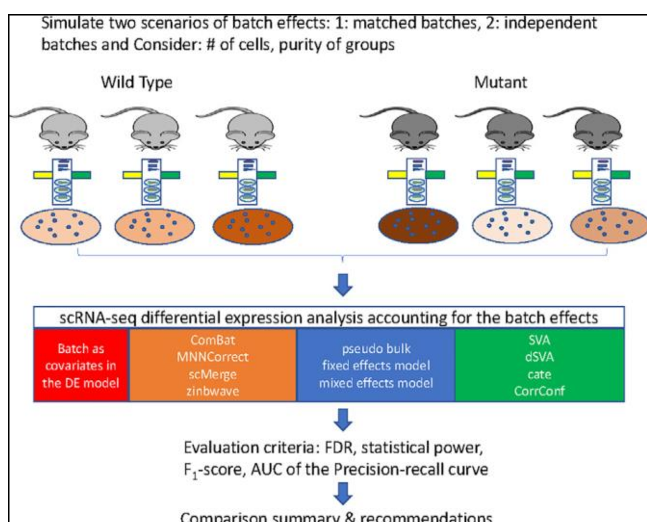

Fig 1

Tumor-specific expression in cancer based on single-cell RNA sequencing

Head and neck squamous cell carcinoma (HNSCC) have been widely reported and considered as one of the most threatening diseases to human health. Derived from complicated tissue subtypes, HNSCC has diverse symptoms and pathogenesis. They make the identification of the core carcinogenic factors of such diseases at the multi-cell level difficult. With the development of single-cell sequencing technologies, the effects of non-malignant cells on traditional bulk sequencing data can be eliminated directly. On the basis of fresh single-cell RNA-seq data, we set up a computational filtering strategy for tumor cell identification in an expression rule manner. This strategy can reveal the accurate expression distinction between tumor cells and adjacent tumor microenvironment, which are all supported by literature reports.

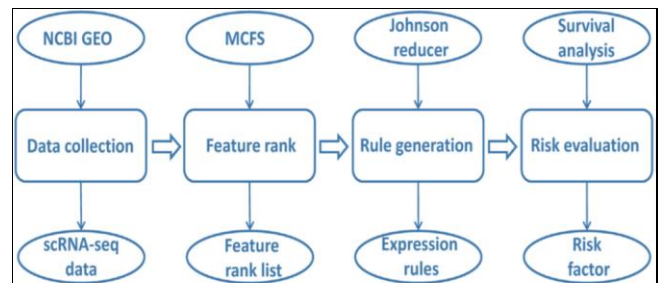

Fig 2

Monte Carlo feature selection (MCFS)

MCFS is a feature selection method used to find the distinguishing features for supervised classifiers. The algorithm is based on the expansion of the decision tree; it trains multiple decision trees on bootstrap samples, and feature subsets with m features are randomly selected from the original M features (m < <M). For each feature subset, p decision trees are grown and evaluated on the respective training set of the bootstrap dataset. The above process is repeated t times for obtaining t feature subsets and p × t decision trees. MCFS calculates a score called relative importance (RI) for each feature on the basis of its performance in the above p × t decision tree classifiers (Chen *et al*., 2019; Pan *et al*., 2019) [54, 53].

Rule learning

Rule learning is a term in machine learning that aims to generate readable rules for interpreting the decision (e.g., the predicted cell type) on the basis of available data (Cai *et al*., 2018) [55]. It is characterized by identifying and utilizing a set of relational rules that collectively represent the knowledge captured by the learning system. Rules are used to describe the association between the genes (i.e., the left side of the rule) and the cell types (i.e., the right side of the rule). In short, a classification rule is expressed as an if–then relationship on the basis of gene expression values. The interpretable classification rules are generated by the Johnson reducer algorithm in the MCFS package in this work (Zeng and Liu, 2010) [56].

Results and discussion

Single cell RNA Seq Overlap Analysis Provides Coverage

Independent Determination of Ploidy or Cell Number Single-cell analysis has proved particularly useful as a tool to find new cell types or intermediate cell states and reconstruct differentiation trajectories. However, we currently lack tools to unambiguously prove that a

transcriptional profile originates from a single cell.

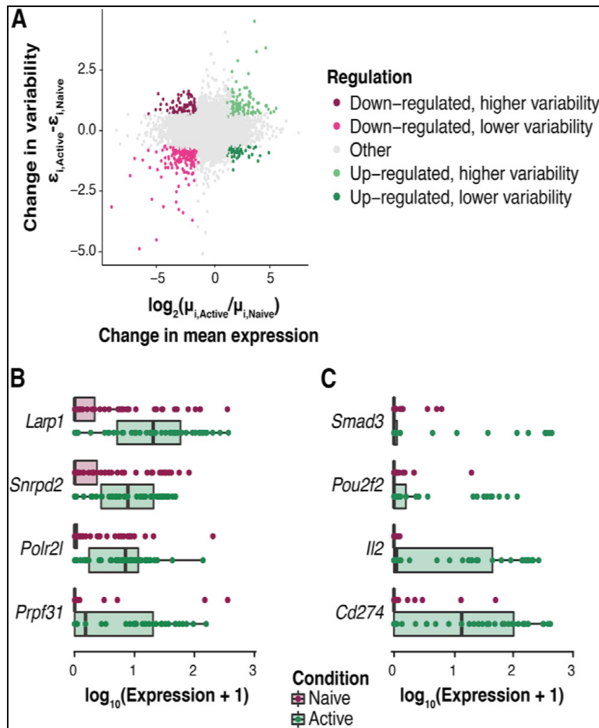

Fig 3

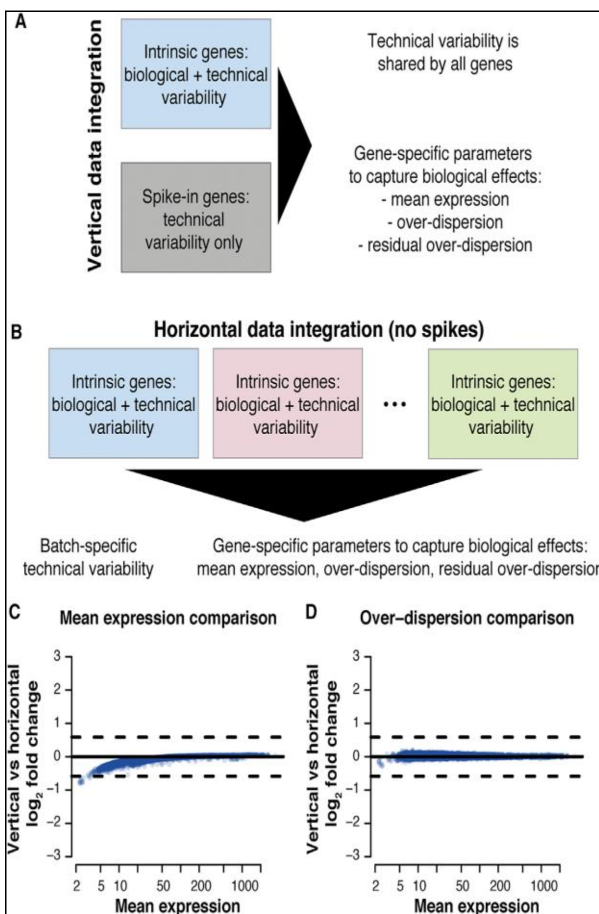

Fig 4

Doublets of cells from two different cell types appear as intermediate transcriptional profiles and cannot be distinguished from a ''true'' intermediate cell. The direct tagmentation procedure of DNTR-seq, however, offers an orthogonal type of data that can directly determine whether the genetic material corresponds to a single genome (2n) or two genomes (4n). After direct tagmentation, the number of overlapping fragments at a specific genomic site (its physical genomic coverage) is restricted by the ploidy at that site. Each chromosome is directly cut into non-overlapping fragments by the transposome with no prior amplification. For a diploid (2n) cell, this means that a coverage of more than two is not possible, whereas for a tetraploid cell or cell doublet (4n) the maximum coverage is four. When using ultra-low coverage sequencing, the probability of observing >23 coverage is exceedingly low even for 4n genomes found in doublets. However, the difference in scarcity of template between a diploid cell compared to a tetraploid also affects the probability of observing 13 coverage compared to 23 coverage in a predictable manner (illustrated at high a simulated high depth of coverage in Figure 4B). We can use this to calculate standardized coverage probabilities (SCPs), where SCP[k] is the probability of a genomic position being covered by k fragments at a standardized sequencing depth. Rather than sequencing the libraries at a specified depth, we estimate the SCP by extrapolating from existing data (Zachariadis and Cheng, 2020) [57].

Approaches to Analyze the Sc-RNA Sequences:

Principal component analyses:

The ROR1e LECs were dissociated immediately prior to processing for scRNA-seq. RNA was extracted from the cultured LECs and processed using a 10X Genomics single cell 3% mRNA- prep kit. Sequencing was performed using an Illumina Next Seq 500 sequencer. The scRNA-seq data can be accessed via Array Express accession E-MTAB-9178. Analysis using the Seurat guided clustering suite yielded 1979 cells, with a total of 17,944 genes expressed. Assessment via prin- cipal component analysis (PCA) revealed three distinct cell clusters.

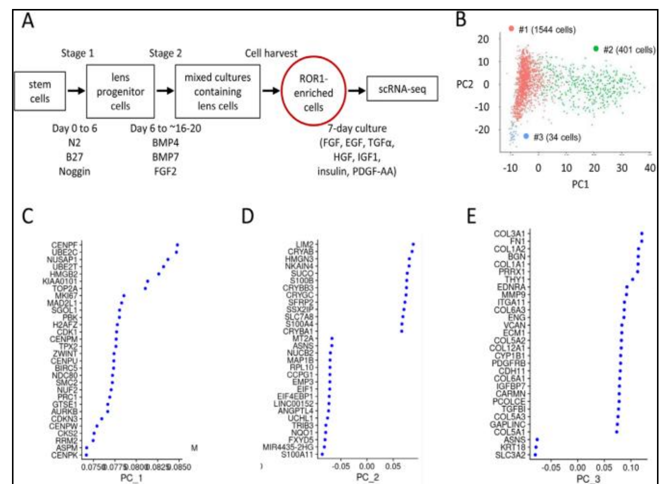

Fig 5

Heat map analysis of critical lens genes

Heat map analyses can be used to investigate expression of critical lens genes in each of the three cell clusters. The initial analysis could be examined expression of 65 genes known to be required for lens development. This could include growth factor signaling genes, crystallin genes and their regula- tors, proliferation and cell survival genes, and various transcription factors (Cvek and Zhang, 2007).

Analysis of cluster 1 marker genes

Heat maps and GO analyses were then used to assess the 17 genes identified as marker genes. The heat map analysis showed cluster marker genes can be more similarly expressed in cluster 2 than in cluster 3 suggesting cluster 1 is similar to cluster 2. No significant GO terms were associated with the cluster 1 marker genes. Comparison with published lens transcriptomic datasets revealed all the cluster 1 marker genes are expressed in adult (Kakrana *et al*., 2018) human lenses and/or mouse lenses including protein-coding and non-coding genes required for normal lenses (e.g., BEX2, TKT, SNHG8, MALAT1).

Role of Single-cell RNA sequencing in cancer

Cancer has become one of the greatest threats to human health, and new technologies are urgently needed to further clarify the mechanisms of cancer so that better detection and treatment strategies can be developed. At present, extensive genomic analysis and testing of clinical specimens shape the insights into carcinoma.

Nevertheless, carcinoma of humans is a complex ecosystem of cells, including carcinoma cells and immunoselected and stroma-related subsets, with accurate characteristics obscured by extensive genome-related approaches. A growing body of research shows that sequencing of single-cell RNA (scRNA-seq) is emerging to be an effective way for dissecting human tumor tissue at single-cell resolution, presenting one prominent way for explaining carcinoma biology. This review summarizes the research progress of scRNA-seq in the field of tumors, focusing on the application of scRNA-seq in tumor circulating cells, tumor stem cells, tumor drug resistance, the tumor microenvironment, and so on, which provides a new perspective for tumor research (Heyer *et al*., 2010) [60].

During the last three decades, carcinoma studies have largely discussed somatic gene-centric mutations (so-called oncogenes) that target their functional characteristics and biochemical activity.1,2 Thus far, multiple targeted therapies have been approved to treat multiple tumors, and more are in development or in early clinical trials (Ma *et al*., 2015).

However, with the extensive use of targeted therapies, common themes of treating relapse and drug resistance have been proposed.3,4 Carcinoma refers to a selectively proliferating, invasive somatic mutant phenotype. There is an implicit concept that following the evolution path of carcinoma, genetically complex groups of different individual carcinoma cells may develop and interact in a dynamic manner with each other. Therefore, studying this potential intratumoral genetic heterogeneity is of great significance for the selection of anti-target treatment methods (Lito *et al*., 2013) [62].

Emerging Trends in Single Cell RNA sequencing in gene expression

Indeed, as next-generation sequencing (NGS) technology is emerging, sequencing of large amounts of DNA and RNA retrieved from carcinoma tissue can be conducted in depth, fine-grained studies of intra-tumor genetic heterogeneity can be carried out, and computational inferences of subclonality can be achieved.

However, as impacted by practical and technical limitations, deep sequencing alone is insufficient to fully explore the genomic and transcriptomic heterogeneity of carcinoma. Individual cells are the basic substrates for the mechanisms of mutation and selection at work to evolve complex structures known as tumor blocks. For this reason, gaining insights into individual carcinoma cells under the individual condition and, overall, into dynamically related systems of interaction (Carcinoma microenvironment) will indeed further clarify therapy related resistance and biology of tumors generally (Garziera *et al*., 2019) [63].

The quickly advancing method of single-cell RNA sequencing (scRNAseq), by exhibiting its ability for characterizing the epigenome, transcriptome, and genome of one individual cell, profoundly presents insights into genetics and tumor biology, and it will enable us to understand the changes in the various stages of tumor progression to advanced metastatic disease (Tang *et al*., 2009). In addition, the clinical application of sc-RNA-seq may profoundly alter the way we treat carcinoma. Much information has been gathered using the mentioned techniques, and, although numerous difficulties still exist, it is considered that the capacity exhibited by scRNA-seq will continuously drive innovation and yield methods to solve current issues that profoundly increase our understanding of the disease. In this review, we discuss the topic of scRNA-seq in the carcinoma field and highlight the use of emerging scRNA-seq approaches to circulating tumor cells (CTCs), tumor drug resistance, and the tumor microenvironment (TME).

Conclusions

Single cell RNA sequencing is relatively a new technology and can be effectively used to elaborate the role of gene expressions in causing genetic diseases. The human population is increasing day by day, so there is a damn need to better understand the causes of genetic diseases. It can bring fruitful results to better understand the controlling of such diseases and making all the strategies beneficial to uproot genetic diseases all over the world. There is a future brightening potential in this technology, and it will surely enhance the living standards of the modern communities by completely elaborating the roles of those faulty genes expression causing genetic diseases.

References

1. Greco EA, Pietschmann P, Migliaccio S. Osteoporosis and sarcopenia increase frailty syndrome in the elderly. Front Endocrinol, 2019, 10.
2. Lo JH, U KP, Yiu T, Ong MT, Lee WY. Sarcopenia: current treatments and new regenerative therapeutic approaches. J Orthop Transl,2020:23:38-52.
3. Wang J, Leung K-S, Chow SK-H, Cheung W-H. Inflammation and age-associated skeletal muscle deterioration (sarcopaenia). J Orthop Transl,2017:10:94-101.
4. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. Nat. Rev. Genet,2009:10:57-63.
5. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics,2010a:26:139-140.
6. Sudheer Menon. "Preparation and computational analysis of Bisulphite sequencing in Germfree Mice" *International Journal for Science and Advance Research In Technology,*2020:6(9):557-565.

7. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal. “Computational identification of promoter regions in prokaryotes and Eukaryotes” EPRA International Journal of Agriculture and Rural Economic Research (ARER),2021:9(7):21-28.
8. Sudheer Menon. “Bioinformatics approaches to understand gene looping in human genome” EPRA International Journal of Research & Development (IJRD),2021:6(7):170-173.
9. Sudheer Menon. “Insilico analysis of terpenoids in Saccharomyces Cerevisiae”international Journal of Engineering Applied Sciences and Technology ISSN No. 2455-2143,2021:6(1):43-52.
10. Wagner AH, Anand VN, Wang WH, Chatterton JE, Sun D, Shepard AR *et al*. Exon-level expression profiling of ocular tissues. Exp. Eye Res,2013:111:105-111.
11. Yoshida S, Yashar BM, Hiriyanna S, Swaroop A. Microarray analysis of gene expression in the aging human retina. Investigative ophthalmology & visual science,2002:43:2554-2560.
12. Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol. Syst. Biol. 15, e8746.
13. Sudheer Menon. “Computational analysis of Histone modification and TFBs that mediates gene looping” Bioinformatics, Pharmaceutical, and Chemical Sciences (RJLBPCS),2021:7(3):53-70.
14. Sudheer Menon Shanmughavel piramanayakam, Gopal Prasad Agarwal. “FPMD-Fungal promoter motif database: A database for the Promoter motifs regions in fungal genomes” EPRA International Journal of Multidisciplinary research,2021:7(7):620-623.
15. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal. Computational Identification of promoter regions in fungal genomes, International Journal of Advance Research, Ideas and Innovations in Technology,2021:7(4):908-914.
16. Sudheer Menon, Vincent Chi Hang Lui, Paul Kwong Hang Tam. methods for identifying hirschsprung disease genes, International Journal for Research in Applied Science & Engineering Technology (IJRASET),2021:9(VII):2974-2978.
17. Zhang Y, Chunman Zuo Liguo Liu, Yunping Hu Bo Yang. Single-cell RNA-sequencing atlas reveals an MDK-dependent immunosuppressive environment in ErbB pathway-mutated gallbladder cancer. J. Hepatology, 2021, 187-200.
18. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med,2018:50:96.
19. Serena L, Trapnell Cole. Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000Res, 2016, 5.
20. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V *et al*. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell,2015:161:1187-201.
21. Hicks SC, Townes FW, Teng M, Irizarry RA. Missing data and technical variability in single-cell RNA-sequencing experiments. Biostatistics,2018:19:562-78.
22. Tung PY, Blischak JD, Hsiao CJ, Knowles DA, Burnett JE, Pritchard JK *et al*. Batch effects and the effective design of single-cell gene expression studies. Sci Rep,2017:7:39921.
23. Sudheer Menon. Bioinformatics approaches to understand the role of African genetic diversity in disease, International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET),2021:4(8):1707-1713.
24. Sudheer Menon. Comparison of High-Throughput Next generation sequencing data processing pipelines, International Research Journal of Modernization in Engineering Technology and Science (IRJMETS),2021:3(8):125-136.
25. Sudheer Menon. Evolutionary analysis of SARS-CoV-2 genome and protein insights the origin of the virus, Wuhan, International Journal of Creative Research Thoughts (IJCRT),2021:9(8):PP b696-b704.
26. Sudheer Menon, Vincent Chi Hang Lui, Paul Kwong Hang Tam. A step-by-step work flow of Single Cell RNA sequencing data analysis, International Journal for Scientific Research and Development (IJSRD),2021:9(6):1-13.
27. Cole MB, Risso D, Wagner A, DeTomaso D, Ngai J, Purdom E *et al*. Performance assessment and selection of normalization procedures for single-cell RNA-Seq. Cell Syst,2019:8(315-328):e318.
28. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. Nat Methods,2018:15:255-61.
29. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK *et al*. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol,2015:16:278.
30. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNAsequencing data are corrected by matching mutual nearest neighbors. Nat Biotechnol,2018:36:421-7.
31. 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,2018:9:284.
32. Sudheer Menon. Computational characterization of Transcription End sites in Human Genome, International Journal of All Research Education and Scientific Methods (IJRESM),2021:9(8):1043-1048.
33. Sudheer Sivasankaran Menon, Shanmughavel Piramanayakam. Insilico prediction of gyr A and gyr B in *Escherichia coli* insights the DNA-Protein interaction in prokaryotes, International Journal of Multidisciplinary Research and Growth Evaluation, (IJMRD),2021:2(4):709-714.
34. Sudheer Menon, Vincent Chi Hang Lui, Paul Kwong Hang Tam. Bioinformatics tools and methods to analyze single cell RNA sequencing data, International Journal of Innovative Science and Research Technology, (IJISRT),2021:6(8):282-288.
35. Sudheer Menon. Computational genome analysis for identifying Biliary Atresia genes, International Journal of Biotechnology and Microbiology, (IJBM),2021:3(2):29-33.
36. Lin Y, Ghazanfar S, Wang KYX, Gagnon-Bartsch JA, Lo KK, Su X *et al*. scMerge leverages factor analysis, stable expression, and pseudoreplication to merge multiple single-cell RNA-seq datasets. Proc Natl Acad Sci USA,2019:116:9775-84.

37. Kumar, Kankainen M, Parsons A, Kallioniemi O, Mattila P, Heckman CA. The impact of RNA sequence library construction protocols on transcriptomic profiling of leukemia, BMC Genomics,2017:18(1):629.
38. K. Van den Berge, Perraudeau F, Soneson C, Love MI, Risso D *et al*. Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications, Genome Biol,2018:19:24:1406-4.
39. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics,2010:26:139-140.
40. Anders S, Huber W. Differential expression analysis for sequence count data, Genome Biol,2010:11:R106. https://doi.org/10.1186/gb-2010-11-10-r106.
41. Love MI, Anders S, Huber W. Differential analysis of count data - the DESeq2 package, Genome Biol. (2014.
42. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts, Genome Biol,2014:15:R29.
43. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al*. Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res, 2015, 43. https://doi.org/10.1093/nar/gkv007 e47.
44. Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data, Bioinformatics, 2009.
45. Fujita K, Iwaki M, Yanagida T. Transcriptional bursting is intrinsically caused by interplay between RNA polymerases on DNA, Nat. Commun, 2016.
46. Sudheer Menon. Recent Insilco advancements in genome analysis and characteristics of SARS-Cov2. International Journal of Biology Research, (IJBR),2021:6(3):50-54.
47. Sudheer Menon. Bioinformatics methods for identifying Human disease genes, International Journal of Biology Sciences, (IJBR),2021:3(2):1-5.
48. Sudheer Menon. SARS-CoV-2 Genome structure and protein interaction map, insights to drug discovery, International Journal of Recent Scientific Research, (IJRSR),2021:12(8):42659-42665.
49. Sudheer Menon. Insilico Insights to Mutational and Evolutionary aspects of SARS-Cov2, International Journal of Multidisciplinary Research and Development, (IJMRD),2021:8(8):167-172.
50. Wang J, Huang, Torre E, Dueck H, Shaffer S, Murray J. *et al*. Gene expression distribution deconvolution in single-cell RNA sequencing, Proc. Natl. Acad. Sci. U. S. A, 2018. https://doi.org/10.1073/pnas.1721085115.
51. Miao Z, Deng K, Wang X, Zhang X. DE single for detecting three types of differential expression in single-cell RNA-seq data. Berger B, editor, Bioinformatics,2018:34:3223-3224.
52. Xiangtian Yua, Zhenjia Wangb, Tao Zengc. Essential gene expression pattern of head and neck squamous cell carcinoma revealed by tumor-specific expression rule based on single-cell RNA sequencing. Molecular Basis of Diseases. 1866. 2020. 165791.
53. Chen L, Zeng T, Pan X, Zhang YH, Huang T, Cai YD. Identifying methylation pattern and genes associated with breast cancer subtypes, Int J Mol Sci, 2019, 20.
54. Pan X, Hu X, Zhang Y-H, Chen L, Zhu L, Wan S *et al*. Identification of the copy number variant biomarkers for breast cancer subtypes, Mol. Gen. Genomics. 294, 2019, 95-110.
55. Y-D Cai, Zhang S, Zhang Y-H, Pan X, Feng K, Chen L, Huang T, Kong X. Identification of the gene expression rules that define the subtypes in glioma, J. Clin. Med,2018:7:350.
56. Zeng T, Liu J. Mixture classification model based on clinical markers for breast cancer prognosis, Artif. Intell. Med,2010:48:129-137.
57. Zachariadis V, Cheng H. Nathanael Andrews, Martin Enge.A Highly Scalable Method for Joint Whole-Genome Sequencing and Gene-Expression Profiling of Single Cells. Molecular Cells,2020:2(3):550-580.
58. Cvek A, Zhang X. Signaling and gene regulatory networks in mammalian lens development, Trends Genet,2017:33:677-702.
59. Kakrana A, Yang A, Anand D, Djordjevic D, Ramachandruni D, Singh A *et al*. iSyTE 2.0: a database for expression-based gene discovery in the eye, Nucleic Acids Res,2018:46:D875-D885.
60. Heyer J, Kwong LN, Lowe SW, Chin L. Non-germline genetically engineered mouse models for translational cancer research. Nat. Rev. Cancer,2010:10:470-480.
61. Ma CX, Reinert T, Chmielewska I, Ellis MJ. Mechanisms of aromatase inhibitor resistance. Nat. Rev. Cancer,2015:15:261-275.
62. Lito P, Rosen N, Solit DB. Tumor adaptation and resistance to RAF inhibitors. Nat. Med,2013:19:1401-1409.
63. Garziera M, Roncato R, Montico M, Mattia De, Gagno E, Poletto S *et al*. New challenges in tumor mutation heterogeneity in advanced ovarian cancer by a targeted next-generation sequencing (NGS) approach. Cells,2019:8:584.
64. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N *et al*. mRNA-seq whole-transcriptome analysis of a single cell. Nat. Methods,2009:6:377-382.
65. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J *et al*. Tumour evolution inferred by single-cell sequencing. Nature,2011:472:90-94.