



## OPEN ACCESS

## EDITED BY

Pavel Banerjee,  
University of Michigan, United States

## REVIEWED BY

Fabio Grizzi,  
Humanitas Research Hospital, Italy  
Tanmay Chatterjee,  
University of Michigan, United States  
Liuhan Dai,  
Harvard University, United States  
Kartik Chandra Guchhait,  
Debra Thana Sahid Kshudiram Smriti  
Mahavidyalaya, India

## \*CORRESPONDENCE

Ioana Berindan-Neagoe

✉ ioana.neagoe@umfcluj.ro

Andreas Bender

✉ andreas.bender@umfcluj.ro

RECEIVED 27 August 2025

REVISED 05 December 2025

ACCEPTED 08 December 2025

PUBLISHED 06 January 2026

CORRECTED 14 January 2026

## CITATION

Bica C, Zanoaga O, Pop L, Ciocan C,  
Raduly L, Nuțu A, Berindan-Neagoe I and  
Bender A (2026) Tumor heterogeneity  
assessment using single-cell RNA sequencing  
(scRNA-seq): applications in lung cancer  
for diagnosis and treatment.  
*Front. Immunol.* 16:1693784.  
doi: 10.3389/fimmu.2025.1693784

## COPYRIGHT

© 2026 Bica, Zanoaga, Pop, Ciocan, Raduly,  
Nuțu, Berindan-Neagoe and Bender. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication  
in this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Tumor heterogeneity assessment using single-cell RNA sequencing (scRNA-seq): applications in lung cancer for diagnosis and treatment

Cecilia Bica<sup>1</sup>, Oana Zanoaga<sup>1</sup>, Laura Pop<sup>1</sup>, Cristina Ciocan<sup>1</sup>,  
Lajos Raduly<sup>1</sup>, Andreea Nuțu<sup>1</sup>, Ioana Berindan-Neagoe<sup>1,2,3\*</sup>  
and Andreas Bender<sup>1,4,5,6\*</sup>

<sup>1</sup>Department of Genomics, MEDFUTURE Institute for Biomedical Research, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania, <sup>2</sup>Doctoral School, Iuliu Hatieganu University of Medicine and Pharmacy, Cluj-Napoca, Romania, <sup>3</sup>Romanian Academy of Medical Sciences, Bucharest, Romania, <sup>4</sup>College of Medicine and Health Sciences, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates, <sup>5</sup>Centre for Molecular Informatics, Department of Chemistry, University of Cambridge, Cambridge, United Kingdom, <sup>6</sup>STAR-UBB Institute, Babeş-Bolyai University, Cluj-Napoca, Romania

Recent progress in single-cell RNA sequencing has led to mechanistic and clinically actionable insight into genetic heterogeneity and tumor progression *via* transcriptome profiling at single-cell level, applied to identification of cell types, gene expression patterns, and signaling pathways involved in cancer development. In this work, we review the use of single-cell RNA sequencing (scRNA-seq) applications to gain insights into tumor molecular and cellular characteristics, such as cellular heterogeneity, rare cell populations, characteristic pathogenic cell populations, cells of the immune tumor microenvironment, and information regarding clonal evolution, none of which can be observed using bulk RNA-seq. We describe how this set of methods facilitates a better understanding of tumor heterogeneity, interactions between the tumor cells and the cells of the tumor microenvironment (TME), and can elucidate potential therapeutic targets. From the applied clinical perspective, we summarise the ability of scRNA-seq data to identify molecular indicators for diagnosis, outcome, and prediction of response to therapy. This is particularly relevant due to the low response rate to therapy of non-small cell lung cancer (NSCLC) and acquired resistance, including in immunotherapy.

## KEYWORDS

lung cancer, molecular profile, single-cell RNA sequencing (scRNA-seq), tumor heterogeneity, tumor microenvironment (TME)

## Introduction

Cancer contributes considerably to the global disease burden. According to the most recent statistics, reports, and forecasts, the annual number of global cancer deaths is increasing—a trend expected to continue for at least two decades (1–5). One reason for this situation is related to the rise in the incidence of cancer following lifestyle changes, e.g. a growing middle class in developing countries which consumes more meat, alcohol, and cigarettes, etc. (6), although a decrease in mortality rates in cancer is observed in some developed countries, as according to a study investigating the UK population (37% decline in men and 33% decline in women aged 35–69 years) (7). Hence, cancer remains of enormous medical relevance to society.

One of the main determinants of therapy resistance in cancer is tumor heterogeneity. Evidence suggests that tumor heterogeneity is not solely caused by genomic instability but by cooperation with other elements, such as clonal evolution and selection, to maintain and promote tumor development and progression (8), and hence understanding those aspects in an integrated manner is key for both diagnosis and treatment. The hallmarks of therapy resistance are either inherited through the genetic and epigenetic makeup of cancer cells or acquired following treatment. The acquired form results from therapeutic interventions that foster the evolution of treatment-resistant cellular subpopulations (9, 10). The evaluation of tumor heterogeneity is therefore critical, distinct in every patient regarding spatial distribution and temporal heterogeneity, with associated (different) effects on tumor progression, recurrence, and resistance to therapy (11, 12). In this respect, the complex tumor microenvironment (TME), characterized as a highly structured ecosystem of elements, such as immune and stromal cells, blood vessels, and the extracellular matrix plays a vital role in cancer cell survival, invasion, and metastasis (13). The dynamics within the tumor need to be understood *spatially* and *over time*, as the tumor configuration includes various cell types – cancer cells, endothelial cells, fibroblasts, and immune cells, which incorporate multiple distinct cell populations (14). Different cell populations

possess distinct molecular signatures, which confer varying levels of treatment sensitivity and ultimately shape the overall response to therapy (8). Such insights into tumor heterogeneity and dynamics needed for disease understanding and to drive treatment are currently being facilitated by novel techniques to generate and analyze single-cell ‘-omics’ data.

According to the International Agency for Research on Cancer, lung cancer remains the number one cause of cancer-related deaths, with 9.7 million deaths in 2022 (15, 16). The current state of knowledge is that actionable genomic alterations encountered in non-small cell lung cancer (NSCLC) tumors are present mainly in lung adenocarcinoma tumors, as opposed to squamous cell tumors (17). Targeting these driver alterations using suitable targeted treatment options improves overall survival (18). Still, in terms of absolute numbers, in lung adenocarcinoma (LUAD) tumors, oncogenic driver mutations were observed in only 27% to 41% of tumors, with variations related to the histologic subtype (19). It was noted that although complex subclonal heterogeneity characterizes recurrent lung cancer, the dominant oncogenic driver that was clinically relevant at the time of diagnosis (and initial surgical treatment) often still exists at the time of recurrence, indicating their role in conferring sensitivity to targeted therapy (19). Still, the majority of NSCLC patients do not harbor (currently known) driver mutations/alterations, thus indicating both a lack of biological understanding and our inability to select suitable treatment options currently. In particular, driver mutations are rarely present (20) in squamous cell carcinoma, and higher intratumor heterogeneity was observed in these tumors compared to lung adenocarcinomas (21, 22). The majority (about 70%) of NSCLC patients diagnosed and undergoing surgical resection in the early stages will develop recurrent metastatic disease (23). Thus, NSCLC is currently poorly treated due to a lack of molecular testing on early-stage specimens, as treatment strategies and prognosis may differ within the same stage due to its broad classification (24), and we still need to understand both disease biology and actionable signals in said biology much better. In lung cancer, heterogeneity can occur not only as interpatient heterogeneity, which will result in different treatment behavior in clinic, but also intratumor heterogeneity, which arises not solely from mutations in multiple genes, but also in terms of cell populations with different phenotypic features (25, 26). Stratifying NSCLC patients based on their molecular characteristics has improved with the development of single-cell sequencing technologies as our knowledge of the mechanisms underlying lung cancer has deepened. A high-resolution view of the cells in the TME in early and advanced stages of LUAD and offering insights regarding cellular and molecular network dynamics during tumor progression possible using single-cell sequencing technologies provides the fundamentals for future discoveries of molecular therapeutic targets (27). Moreover, the mechanism of resistance sensitivity to therapy can be identified by identifying rare cell populations/subpopulations and their role in regulating crucial biological pathways and patterns related to immune cell infiltration in lung tumors (28, 29). Considering the highly heterogeneous nature of NSCLC, identifying driver gene alterations specific to each lung

**Abbreviations:** AT2, type II alveolar cells; ATNKGS, APC/T/NK cells-related gene signature; CAF, cancer-associated fibroblasts; cMAP, Connectivity Map; CRD, circadian rhythm disruption; DEGs, differentially expressed genes; Gln, glutamine; ICD, immunogenic cell death; LINCS, Library of Integrated Network-based Cellular Signatures; LPRI, Lung Cancer Prognostic Regulon Index; LUAD, lung adenocarcinoma LUSC, lung squamous cell carcinoma NETosis, Neutrophil Extracellular Traps process; NK, natural killer; NSCLC, non-small cell lung cancer; OS, overall survival; PCA, principal component analysis; PD/NR, progressive disease/nonresponse; PD-1, programmed death receptor 1 PD-L1, programmed death ligand 1 PFS, progression-free survival; scRNA-seq, single-cell RNA sequencing; scGSEA, Single-cell gene set enrichment analysis; t-SNE, t-distributed stochastic neighbour embedding; TAM, tumor associated macrophage; TCMGrisk, T-cell marker genes risk score; TKIs, tyrosine kinase inhibitors; TME, tumor microenvironment; TIME, tumor immune microenvironment; TT, tumor tissue; UMAP, uniform manifold approximation and projection; UMIs, unique molecular identifiers.

TABLE 1 Comparative table summarizing strengths/weaknesses of major scRNA-seq platforms.

Sequencing platform/reference	Type/typical use	Strengths	Weaknesses
Illumina (46, 47)	Short reads/High-throughput gene quantification	-High throughput and accuracy -low error rate -well validated	-limited isoform resolution due to short reads -
Pacific Biosciences (48)	Long reads/Isoform characterization and full-length transcript profiling	-extremely accurate long reads -direct RNA sequencing	-lower throughput -high cost per base
Oxford Nanopore (48)	Long reads/Full-length isoform detection	-ultra long reads -direct RNA sequencing -detections of base modifications -portable devices	- higher raw read error rate in homopolymer regions - low capture efficiency
DNBSeq (49)	Short reads/High-throughput gene quantification	-low costs -high throughput -low duplication rates	-library compatibility still expanding -high technical noise

cancer subtype can be done by adequately exploiting data generated using scRNA-seq (22, 30).

In addition to heterogeneity, practical reasons make tumor characterization in the clinical context difficult. According to the National Comprehensive Cancer Network (NCCN) (31) (as well as in the authors' experience), in recurrent NSCLC, it is often difficult to obtain a suitable tissue sample for molecular profiling due to small sample sizes or lack of access to the tumor location. This situation creates disadvantages in accessing important molecular information regarding the histological profile, biomarkers, and other resolutions for molecular testing (32).

Given the progress of the field, the purpose of this review is to describe the use of scRNA-seq in exploring tumor heterogeneity with a particular focus on lung cancer and immune cell infiltration, and covering both diagnosis and treatment response.

## From RNA sequencing to single-cell RNAseq - methods

Single-cell and spatial RNA sequencing are analytical methods of significant current interest used to gain insights into tumor molecular and cellular characteristics in complex tumor heterogeneity (33, 34). Although in the last two decades, bulk RNA-seq was extensively used to characterize cancer biology, it enabled the clinical translation of only a few gene panels into clinical practice. The reasons for this were partially intrinsic to the data obtained since low-resolution data cannot characterize heterogeneous tumor biology (35, 36). More precisely, rare cell populations, characteristic pathogenic cell populations, cells of the immune tumor microenvironment, and information regarding clonal evolution cannot be observed using bulk RNA-seq (37). However, single-cell RNA-seq (scRNA-seq) and spatial transcriptomics methods developed in recent years contribute in obtaining a finer-grained picture of tumor heterogeneity and better understanding of interactions between tumor cells and cells of the TME and to assist the identification of novel therapeutic targets (38, 39). The reader is referred to recent reviews for a more detailed overview of single-cell and spatial transcriptomics sequencing (40–

43) technologies. Different scRNA-seq platforms were developed, and they use different strategies for transcriptome profiling, which translates in differences of sensitivity, capacity, and reproducibility. Several studies were conducted comparing the sequencing methods/platforms to indicate which one should be used considering the sample type or size (44, 45). In the table below, we summarized the major sequencing platforms used for scRNA-seq, highlighting the strengths and weaknesses (Table 1).

## Software and data analysis

Progress in data generation required the development of computational tools to analyze scRNA-sequencing data, yet the continuous expansion of these tools makes the establishment of best-practice workflows difficult. Currently various analysis pipelines for scRNA-sequencing exist, the most commonly used being Seurat (50, 51), Scanpy (52), and Bioconductor (53). For a more detailed review, see (41).

The general scRNA data analysis workflow is illustrated in Figure 1. The initial steps require raw data processing (specific to the individual sequencing technology), including the mRNA sequence reads are mapped *via* cell barcodes or unique molecular identifiers (UMIs) to a reference genome (Figure 1.1). The resulting count matrices need to be processed with additional filtering steps for quality control, such as to filter for doublets, low-quality (e.g. damaged or stressed) and dying cells (Figure 1.2) (43, 54). Next, data normalization is required to overcome differences in gene expression counts that are generated by sampling effects of cells, to adjust systematic variations – the so-called “uninteresting” variation, which in scRNA-seq is generated by the effects of cell cycle on the transcriptome – and thereby to obtain accurate relative gene abundances between cells (Figure 1.3) (55). As the dimensions of the expression matrices are still high, visualization (Figure 1.4) requires reduction techniques, such as principal component analysis (PCA) (56), uniform manifold approximation and projection (UMAP) (57), and t-distributed stochastic neighbour embedding (t-SNE) (58). These techniques are often used in the manual investigation of the dataset related to clustering and cell

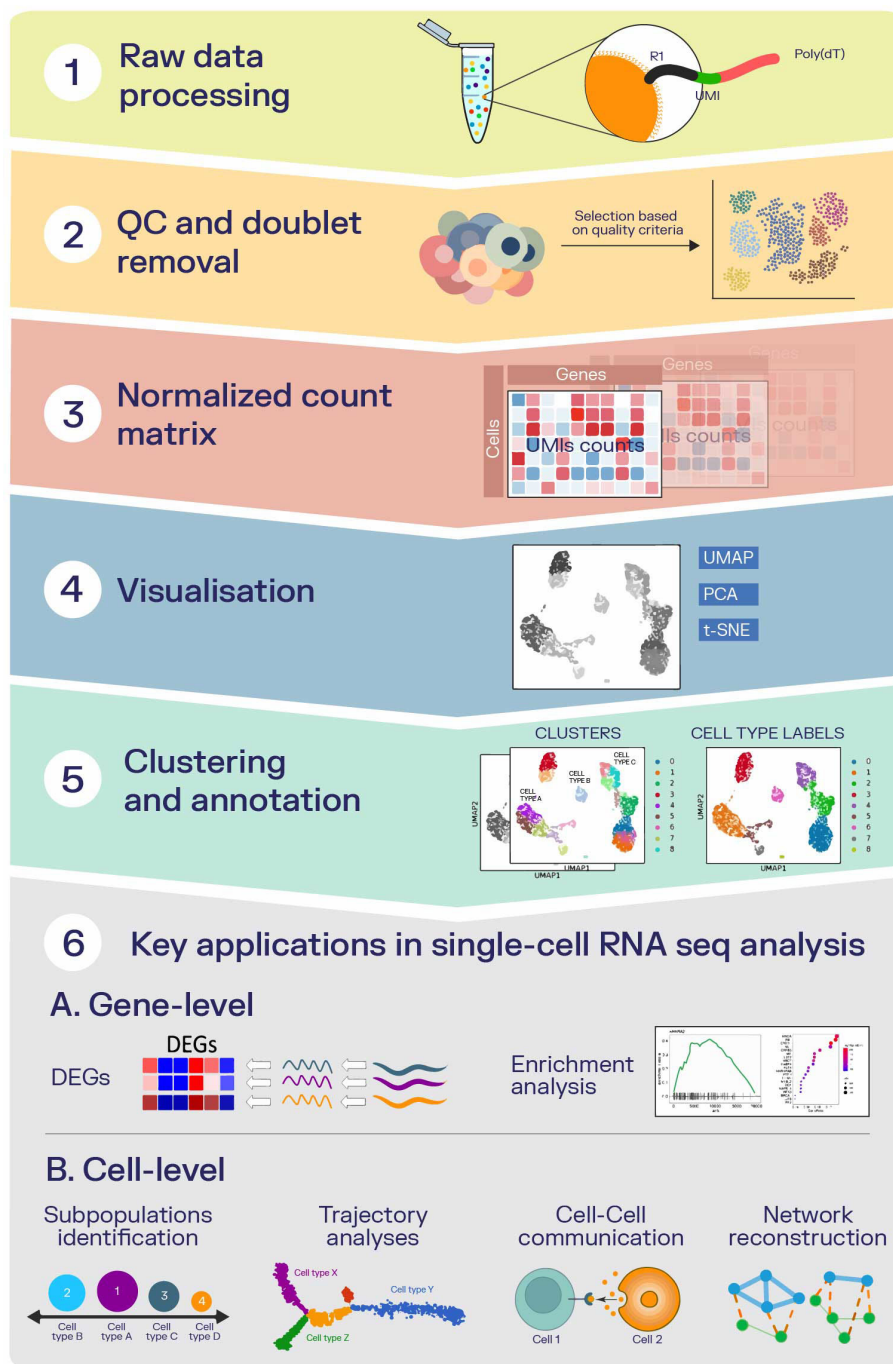


FIGURE 1

Analysis steps and applications using scRNA data which can generally be performed, with a focus in this review particularly on cancer-related applications. **1.** The initial steps require raw data processing - the mRNA sequence reads are mapped via cell barcodes or UMIs to the reference genome. **2.** Count matrices are processed with other filtering steps for QC (filter for doublets, low-quality, and dying cells). **3.** Data normalization - required to obtain accurate relative gene abundances between cells. **4.** Reduction techniques used in the manual investigation of the dataset related to clustering and cell type annotations - PCA, UMAP, and t-SNE. **5.** Clustering - the assignment of clusters is generally based on their typical classification markers. **6.** Exploratory analyses: **(A)** Gene-level - Identification of DEGs and selection of key genes that can represent biomarkers for cancer progression and enrichment analysis - aggregation of the entities in a DEGs list into pathways using GSEA. **(B)** Cellular-level - Subpopulations identification - to observe the abundance of different subpopulations, such as tumor, stromal, and immune cells, as well as rare subpopulations of cells; Trajectory analyses - enables the study of dynamic changes that are related to gene expression; Cell-cell interactions - influence multiple biological processes related to cancer, such as cellular growth and division, differentiation and progression; Network analysis - for understanding, predicting, or optimizing the structure and behavior of complex cancer systems.

type annotations (59). Clustering allows identifying immune cell populations and their abundance in experiments centered on solid cancer tissues. In studies centered on specific cell types in different cancer types using scRNA-seq, the assignment of clusters is generally based on their typical classification markers (Figure 1.5) (60), with more recent approaches being based on machine learning (e.g. deep learning) approaches (61); however, it needs to be kept in mind that in particular non-linear projection methods can be misleading and hence care needs to be taken in their interpretation (62, 63). Beyond visualization and cell type assignment, gene-level exploratory analyses (Figure 1.6A) can cover the expression data to identify differentially expressed genes, allowing the construction of regulatory networks and the establishment of a clonality tree (64). Identification of differentially expressed genes (DEGs) in cancer represents one key application in scRNA-seq analysis and it is used for detecting key genes that can represent biomarkers for cancer progression. When performing the analysis of DEGs in scRNA-seq, the biomarkers are detected for individual cell types, and hence on a much finer resolution, unlike in bulk RNA-seq where DEGs are identified in tumor vs. non-tumor/case vs. control/treated vs. untreated (65). The tools used to determine DEGs in scRNA-seq experiments are further outlined in a recent review (2019) (66), while another comprehensive review summarizes methods for gene selection using gene expression data (65). Gene set enrichment analysis (GSEA) is the most used tool for enrichment analysis. This tool aggregates the entities in a DEGs list into pathways (67).

In addition, intra-tumor heterogeneity can be explored in cellular-level applications (Figure 1.6B) to observe the abundance of different subpopulations, such as tumor, stromal, and immune cells, as well as rare subpopulations of cells, and explore their role in regulating cancer pathogenesis, angiogenesis and in mediating the immune response (68). DEG data can also provide input for various secondary analyses, such as the analysis of transcriptionally regulated pathways, gene sets or network analysis (69). Moreover, cell-cell interactions and information on the cell cycle can be explored using scRNA-seq data, as these interactions influence multiple biological processes related to cancer, such as cellular growth and division, differentiation and progression (70, 71). ScRNA-seq data can be used for cell trajectory analyses, as it enables the study of dynamic changes that are related to gene expression (34, 72). In the process of trajectory inference, genes that are associated with lineage of the trajectory can be identified, as well as genes that are differentially expressed between lineages (73) (Table 2).

## Applications of scRNA-seq in lung cancer

In lung cancer research, scRNA-seq technology has been broadly used in applications such as biomarker discovery, the study of tumor heterogeneity, patient stratification based on the immune cell populations present within tumors, and in target and drug discovery (Figure 2A). Experimental scRNA-seq discoveries can be translated into clinical practice, following translational

research, which include biomarker discovery, validation of drug targets, and development of therapeutic models, which lead to clinical implementation of these discoveries in clinical trials (82) (Figure 2B). These applications are also summarized in Table 2, and they will be described in more detail in the sections that follow.

## Biomarker discovery using scRNA data in lung cancer

*Diagnosis:* Considering the poor response rate of lung cancer patients and frequent acquired resistance, also in immunotherapy, early diagnostic biomarkers are required to detect and confirm the existence of this condition and are used to improve the classification, optimization of treatment options, and survival of these patients. A set of six mitophagy-related genes with diagnostic potential in NSCLC was developed by Yu et al. (83). The AUC for this model was 0.925, respectively 0.966, indicating its high predictive accuracy for the occurrence of NSCLC, in two different datasets (83). Another study identified that their Tumor Immune Dysfunction and Exclusion (TIDE) risk model had an AUC of 0.688 for the diagnosis of LUAD patients (84) (Table 3A). The ongoing investigations indicate that due to the highly heterogeneous nature of lung tumors, combined biomarkers represent a more valuable diagnostic power than single markers, as already reported in other types of cancer, such as ovarian (85), leading to an increase in the detection rate (86).

*Prognostic biomarkers:* Prognostic biomarkers aim to estimate the likelihood of a future clinical event of patients and they usually accompany clinical markers, such as TNM stage, histologic subtype, lymph node involvement, and the presence of metastases (87). Overall, predictive biomarkers for targeted therapy (e.g., EGFR, ALK) and immunotherapy (e.g., PD-1/PD-L1) have improved outcomes for many NSCLC patients. However, new therapeutic predictors are still needed to ensure better outcomes, particularly for those with the LUSC histotype (88). For LUAD and LUSC tissues, by combining scRNA-seq data with deconvoluted bulk RNA-seq data, Zhang et al. (89) intended to identify cellular subtypes for each lung cancer histotype. They were able to observe using the log-rank four main groups based on cellular composition for each histotypes, and concluded that increased percentages of type II alveolar cells ( $p=0.0043$ ) and basal cells ( $p=0.0038$ ) in LUSC, and increased percentages of fibroblasts ( $p=0.016$ ) in LUAD predicted poor survival in these patients, indicating the contribution of cellular composition to assessing the prognosis of lung cancer patients (89). A conceptually similar approach was used by Song et al. (90) to uncover a prognostic gene signature based on natural killer cells markers in LUAD patients. A specific seven-gene prognostic signature expressed in various cell clusters was established in this study. When patients were categorized in low-/high-risk based on the median risk score to assess survival status and the prognostic signature was validated in clinical subgroups based on age, smoking status, gender, stage indicating its predictive stability (HR: 2.227, 95% CI: 1.782–2.784,  $p<0.0001$ ) (90). To emphasize the importance of T-cell marker

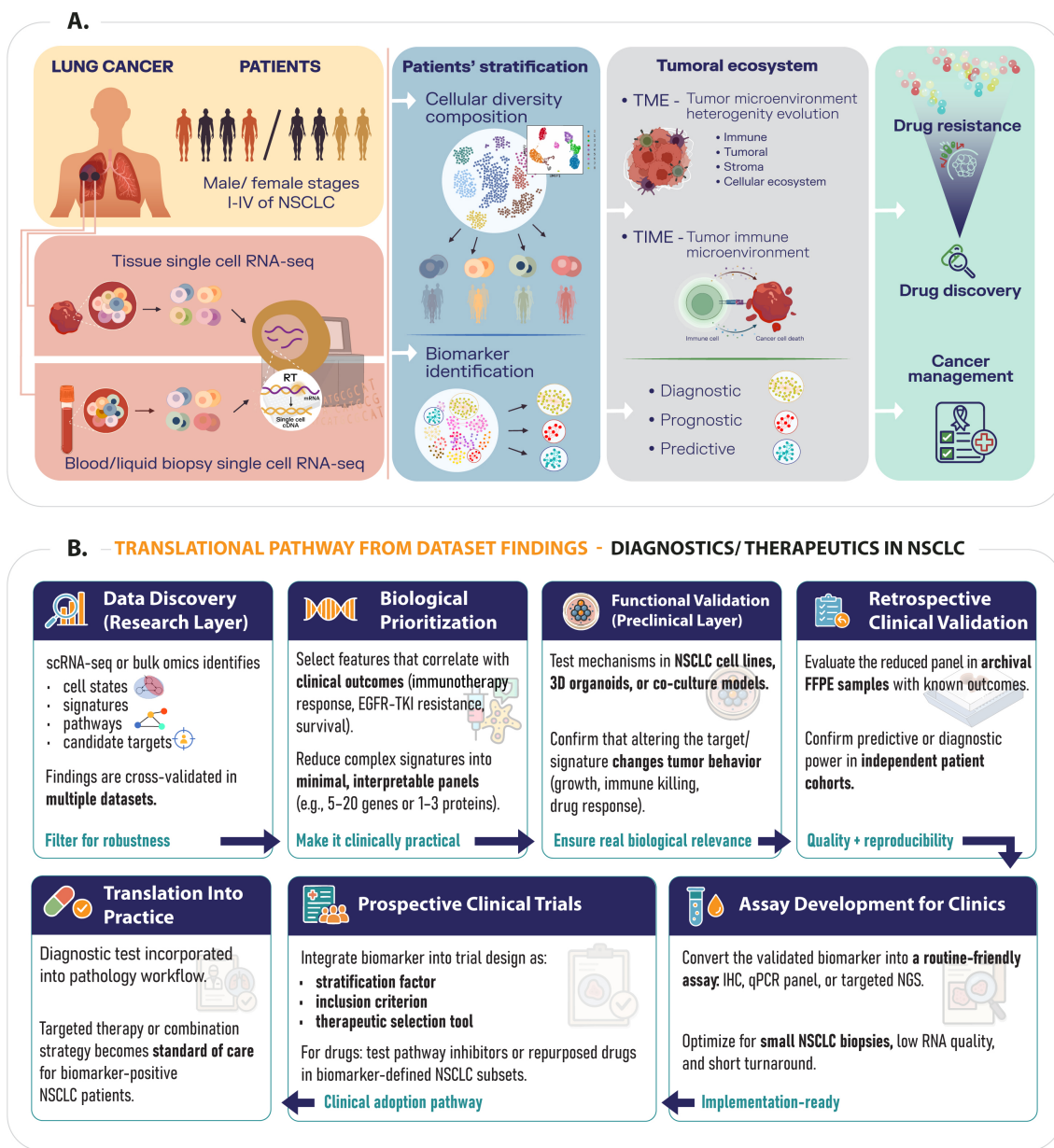
TABLE 2 Types of analyses carried out using scRNA-seq data in cancer studies.

Analysis type	Example objectives	Example study/outcome
Subpopulation identification	Understanding the dynamics of cell evolutionary processes during disease progression and treatment – treatment-resistant cells Identification of cell populations with the ability to initiate new tumors and induce metastasis	-scRNA-seq data from untreated/chemotherapy treated lung cancer patients identified cell types (epithelial, endothelial, fibroblasts, and immune cells), and it was observed that the proportion of immune cells in tumor tissue was higher in treated patients than in the untreated ones, indicating that chemotherapy influences the architecture of TME and this may serve as a guide for making therapy more effective in these patients (74). - in small cell lung cancer (SCLC), a pro-metastatic subpopulation of cancer cells, characterized by high expression of <i>PLCG2</i> was identified, and patients in which this subpopulation represents more than 75% of the total cancer cells, had significantly worse OS compared to others (75)
Differentially expressed genes (DEGs)	Identification of key genes that can represent biomarkers for cancer progression	By comparing the DEGs in NSCLC subtypes, it was revealed that NK cells, Gran, and CD8 cells shared more common altered genes between LUAD and LUSC (76).
Cell-cell communication	Determination of intercellular communication pathways via aggregating ligand and receptor expression data for groups of cells to determine the groups of cells that are likely to interact with one another	Interaction between macrophages and neutrophils in the TME was observed in LUAD patients, indicating that TAMs communicate with neutrophils through the ICAM1 family, while neutrophils are recruiting TAMs via the CXCL family (77).
Trajectory analyses	Inferring the developmental or differentiation trajectories of cells Facilitating the reconstruction on evolutionary trajectories of different types of cell populations This assessment can be done spatially resolved (e.g. to understand the connection between different tumor sites or primary tumors vs. metastatic tumors) and temporally resolved (pre-treatment vs. post-treatment)	Single-cell transcriptomics data from head and neck tumors used to build trajectories of tumor-targeting T-cells from pre-nodal metastases in primary tumors using the Monocle 2.0 algorithm and discovered that these cells move in two different directions, the 'naïve-dysfunction trajectory' and the 'memory-dysfunction trajectory'. This information can then be used for the identification of modulators for T-cell dysfunction in cancer (78, 79).
Network reconstruction	For revealing regulatory interactions between active transcription factors and target genes to provide insights for the identification of causal regulatory factors in biological processes	Possible immunotherapy options for bladder cancer patients were investigated using an interaction network which was constructed using 16 differentially expressed inflammatory genes, which indicated four altered pathways related to T-cell activation and leukocyte cell adhesion (80).
Enrichment analysis	Aggregates the entities in a DEGs list into pathways Scoring coordinated gene activity in cells in order to identify active pathways in each cell	After applying GSEA on different colorectal cancer patients RNA expression profile datasets, it was observed that in patients indexed as low-risk, based on a pyroptosis-related risk model, the pathways enriched were G2M checkpoint, <i>IF<math>\alpha</math>/<math>\gamma</math></i> response, DNA repair, and E2F targets; in high risk patients, the enriched pathways were hypoxia, coagulation, EMT, myogenesis, and downregulated response to UV (81).

EMT, epithelial to mesenchymal transition; IF, interferon; OS, Overall Survival; TAM, tumor associated macrophages; UV, ultraviolet.

genes in LUSC patients, a prognostic signature consisting of five T-cell marker genes was developed in a different study and validated on two different cohorts (91). Given those relatively low AUC values it can be seen that the model doesn't capture all biological complexity in LUSC at this stage. In another study a supervised machine algorithm was used to identify 12 key genes in a prediction model using scRNA-seq data available in Gene Expression Omnibus (GEO) with prognostic potential in NSCLC. By applying Cox-proportional hazards regression model on two different NSCLC cohorts it was observed that survival probability in low risk vs. high risk group decreased over time (HR = 1.85 (1.36-2.53),  $p < 0.0001$ ) (92). Another risk score calculated based on Neutrophil Extracellular Traps process (NETosis) related genes in LUAD scRNA-seq data was shown to be rather predictive in assessing NSCLC patient survival. Using nine different NSCLC cohorts, significantly worse prognosis was observed in patients

with high vs. low score patients (Hazard Ratios (HRs) values fall within 1.56-3.22 range,  $p < 0.01$  for all 9 cohorts) (93). Also solute carriers seem to be suitable as prognostic markers, the role of *SLC7A11* as an independent prognostic indicator in LUAD patients was revealed using scRNA-seq data combined with bulk RNA-seq data (94). It was observed that elevated expression of this gene was observed in clusters that were labeled as epithelial cells, and increased levels of *SLC7A11* indicated poor OS for these patients (94), likely due to differences in oxidative stress response (95). As illustrated in Table 3B, using scRNA-seq data, genes/markers of prognosis can be distinguished for specific pathways or markers of the immune microenvironment encountered altered in cancer. Most of the studies evaluated the stability of these markers in predicting the outcome of NSCLC patients for 1-, 3-, and 5-years. It can be noted that often relatively similar Hazard Ratios are obtained in studies such as the above, which is likely due to the underlying



**FIGURE 2** (A) Applications of scRNA-seq in translational research in lung cancer; (B) Translational pathway for operating dataset-derived discoveries into diagnostics and therapeutics for NSCLC.

correlation structure and relatively low dimensionality of biological readout space, which often gives relatively similar predictive value, largely independent of the variables employed in models (96). Hence, the understanding and prediction of prognostic biomarkers still require significant future progress in both research and clinical translation.

**Drug Response Prediction:** Predictive biomarkers are used to predict the likelihood of a patient to respond to a specific treatment, as well as in stratification of these patients based of their probability of response (97). The risk score calculated based on the gene signature for prediction of natural killer cell markers developed by Song et al. (90) for LUAD was valuable for predicting response to

PD-L1 therapy with 76.1% accuracy (90). Another study that focused on genes enriched in circadian-related signaling pathways using scRNA-seq data of LUAD patients established a ‘disruption score’ to assess the response of patients to tyrosine kinase inhibitors (TKIs) (98). Also in the immunotherapy area a study analyzing glutamine metabolism in LUAD, a risk model including 10 genes was constructed that divided patients into low-risk and high-risk groups. Based on this score, it was observed that not only in LUAD patients, but also in other types of cancer (melanoma and urothelial cancer), the low risk group was associated with lower non-response rates (of 5%) to anti-PD1 and anti-CTLA4 immunotherapy compared to those in the high risk group (where the value was

**TABLE 3** Studies investigating biomarkers for the diagnosis, prognosis, and prediction of response to therapy in lung cancer using ex-vivo scRNA-seq data.

Lung cancer type	Scoring model	Signature	Outcome
<b>A. Diagnosis</b>			
NSCLC (83)	SVM and RF algorithms	6 mitophagy-related genes ( <i>SRC</i> , <i>UBB</i> , <i>PINK1</i> , <i>FUNDC1</i> , <i>MAP1LC3B</i> , and <i>CSNK2A1</i> )	-high expression of <i>SRC</i> , <i>FUNDC1</i> , and <i>CSNK2A1</i> and low expression of <i>UBB</i> , <i>PINK1</i> , and <i>MAP1LC3B</i> associated with increased infiltration of naive B cells, M0/M1 macrophages, NK cells, plasma cells, T helper cells, and regulatory T cells in TT
LUAD (84)	Risk model	8 CAFs marker genes ( <i>TIMP1</i> , <i>TPM2</i> , <i>NR2F2</i> , <i>MFAP4</i> , <i>SOD3</i> , <i>CAV1</i> , <i>SERPINH1</i> , and <i>FMO2</i> )	-risk model (AUC = 0.688) for LUAD diagnosis was developed -a positive correlation between the selected CAFs and EMT and Wnt/ $\beta$ -catenin signaling was found
<b>B. Prognosis</b>			
NSCLC (92)	SVM prediction model	12 key genes ( <i>MS4A1</i> , <i>CCL5</i> , <i>GZMB</i> , <i>HBA1</i> , <i>TYROBP</i> , <i>CD74</i> , <i>LTB</i> , <i>EEF1A1</i> , <i>RPLP1</i> , <i>RPS18</i> , <i>GADPH</i> , and <i>ACTB</i> )	-survival probability in high risk vs. low risk group decreased over time - HR = 1.85 (1.36-2.53), p= 9.766e-05
NSCLC (93)	NETosis risk model	18 NETosis-Related Genes ( <i>ALDH2</i> , <i>ALOX5AP</i> , <i>CCT6A</i> , <i>CD69</i> , <i>CKAP4</i> , <i>DDIT4</i> , <i>DOCK4</i> , <i>ERO1A</i> , <i>FBP1</i> , <i>FKBP4</i> , <i>KRT8</i> , <i>LDHA</i> , <i>MS4A1</i> , <i>S100P</i> , <i>SEC14L4</i> , <i>SLC16A3</i> , <i>SNX30</i> , and <i>UBE2S</i> )	-in nine different NSCLC cohorts, significantly worse prognosis was observed in patients with high vs. low score patients. HR values were in 1.56-3.22 range, p $\leq$ 0.006 for all cohorts
NSCLC (101)	Risk model	5 CAF marker genes ( <i>FBLIM1</i> , <i>LAMC2</i> , <i>NDN</i> , <i>SEMA3C</i> , and <i>SLC39A14</i> )	-worse OS in low-score vs. high-score patients - HR = 2.31 (1.59–3.34), p<0.0001
LUAD (90)	Risk model	7 NK cell marker genes ( <i>GCSAML</i> , <i>ACTG1</i> , <i>ACOT7</i> , <i>SELENOK</i> , <i>PEBP1</i> , <i>BIRC3</i> , <i>ACAP1</i> )	-in six independent cohorts, high-risk patients had inferior prognosis than the low-risk patients - pooled HR = 2.22(1.78-2.78), p<0.0001
LUAD (103)	Risk model	7 exosome-related genes ( <i>CCL20</i> , <i>MAP3K8</i> , <i>SEC61G</i> , <i>SLC34A2</i> , <i>CD79A</i> , <i>BIRC3</i> , and <i>RBM39</i> )	-patients in high risk group had poorer OS, with AUC ranges from 0.722 to 0.827 for 1-year, 0.636-0.706 for 3-year, 0.606-0.746 for 5-year, 0.5-0.726 for 7-year, and 0.621-0.708 for 10-year in different datasets
LUAD (104)	LPRI	15 prognostic-associated regulons ( <i>HDAC2</i> , <i>FOSL1</i> , <i>MAFK</i> , <i>KLF16</i> , <i>MAFF</i> , <i>XBP1</i> , <i>HLF</i> , <i>PRDM16</i> , <i>PPARG</i> , <i>WT1</i> , <i>DLX2</i> , <i>NPAS2</i> , <i>E2F7</i> , <i>HOXC9</i> , and <i>TEAD4</i> )	-patients with high LPRI had a lower OS probability, with AUC ranges from 0.7 to 0.96 for 1-year, 0.57–0.79 for 3-year, and 0.58–0.81 for 5-year in different datasets
LUAD (100)	Risk model	4 CAF-related genes ( <i>FAP</i> , <i>FDGFRB</i> , <i>ACTA2</i> , and <i>NOTCH3</i> )	-AUCs of high risk group range from 0.64 to 0.69 for 1-year, 0.66-0.68 for 3-year, 0.64-0.84 for 5-year in different datasets.
LUAD (105)	Risk model	5 NK cell markers ( <i>ADRB2</i> , <i>IDH2</i> , <i>SFTPC</i> , <i>CCDC69</i> , and <i>CCND2</i> )	-AUCs of high risk group were 0.703 and 0.658 for 1-year in two datasets.
LUAD (106)	Risk model	3 necroptotic anoikis-related genes ( <i>KRT6A</i> , <i>HMMR</i> , and <i>FAM83A</i> )	-higher expression of genes correlated with lower OS rates; AUCs of high risk group were 0.720 for 1-year, 0.733 for 3-years, 0.669 for 5-years
LUAD (107)	Risk model	9 disulfidptosis associated genes ( <i>TLE1</i> , <i>LDHA</i> , <i>SHC1</i> , <i>EMC6</i> , <i>HTATIP2</i> , <i>JAG1</i> , <i>EPHX1</i> , <i>MYO6</i> , and <i>HERPUD1</i> )	-AUCs of high risk groups ranges from 0.565 to 0.92 for 1-year, 0.65-0.77 for 3-year, 0.62-0.83 for 5-year in different datasets
LUAD (108)	Risk model	5 disulfidptosis associated genes ( <i>ERO1A</i> , <i>KRT18</i> , <i>GALNT2</i> , <i>PPIA</i> , and <i>CAPN12</i> )	-AUCs of high risk group were 0.77 and 0.65 for 1-year, 0.73 and 0.68 for 3-year, 0.65 and 0.66 for 5-year in two different datasets
LUAD (109)	Risk model	8 cuproptosis-related stemness genes ( <i>KLF4</i> , <i>SCGB3A1</i> , <i>COL1A1</i> , <i>SPPI1</i> , <i>CABPA</i> , <i>TSPAN7</i> , <i>CAV2</i> , and <i>CTHRC1</i> )	-AUCs of high risk group were 0.704 for 1-year, 0.704 for 3-years, 0.683 for 5-years
LUAD (110)	Risk model	16 ATNKGS ( <i>CD28</i> , <i>CTSL</i> , <i>HLA-DOB</i> , <i>HSPA2</i> , <i>HSPA6</i> , <i>IL2</i> , <i>MAP2K1</i> , <i>NCR3</i> , <i>NRAS</i> , <i>PAK2</i> , <i>PTPN6</i> , <i>RAET1E</i> , <i>RFX5</i> , <i>RFXAP</i> , <i>SHC1</i> , and <i>TAP2</i> )	-patients in high risk group had poorer OS outcomes in six different datasets – pooled HR = 2.55 (2.05–3.17), p<0.0001
LUAD (111)	Risk model	4 muscle failure-related genes ( <i>BDNF</i> , <i>FNDC5</i> , <i>IL15</i> , and <i>MSTN</i> )	-in two different cohorts, this model proved to be a prognostic marker for OS in LUAD patients: HR=4.383 (3.056–6.285), p<0.0001 HR=1.337 (1.142–1.566), p<0.0001

(Continued)

TABLE 3 Continued

Lung cancer type	Scoring model	Signature	Outcome
<b>B. Prognosis</b>			
LUAD (112)	Risk model	12 angiogenesis-related genes ( <i>HPGD, IRX2, SFTPB, CHIA, HOXD1, HSD17B6, MUC16, S100P, Clorf116, KRT16, EGLN3, and ALC2A1</i> )	-AUCs of high risk group were 0.71, 0.71, 0.68 and 0.70 for 3-years; HR = 3.12(2.36–4.12), p<0.001
LUAD (84)	Risk model	8 CAFs marker genes ( <i>TIMP1, TPM2, NR2F2, MFAP4, SOD3, CAV1, SERPINH1, and FMO2</i> )	-AUCs of high risk group range from 0.649 to 0.876 for 1-year, 0.572-0.709 for 3-year, 0.584-0.638 for 5-year in different datasets
LUAD (113)	Adenosine-related prognostic signature	3 adenosine-signaling-related genes ( <i>CD4, FGR, and MNDA</i> )	-the signature is an independent prognostic indicator of OS in LUAD - HR = 3.032(1.339–6.865), p=0.008
LUSC (91)	TCMGrisk	5 T-cell marker genes ( <i>BTG1, JUND, IER3, ZNF331 and PSAP</i> )	-AUCs of high risk group range from 0.669 and 0.661 for 1-year, 0.603 and 0.628 for 3-years, 0.645 and 0.590 for 5-years in two different datasets
LUSC (114)	Risk model	4 apoptosis-related genes ( <i>BMP2, GPX3, JUN, and AIFM3</i> )	-AUCs of high risk group were 0.682 for 3-years, 0.652 for 5-years, and 0.679 for 7-years
<b>C. Response Prediction</b>			
NSCLC (101)	Risk model	5 CAF marker genes ( <i>FBLIM1, LAMC2, NDN, SEMA3C, and SLC39A14</i> )	-high risk patients had poor response (14%) to immunotherapy compared to low-risk patients (45%)
LUAD (90)	Risk model	7 NK cell marker genes ( <i>GCSAML, ACTG1, ACOT7, SELENOK, PEBP1, BIRC3, ACAP1</i> )	-low risk group associated with increased response to PD-L1 therapy -risk score models could predict anti-PD-L1 response with 76.1% accuracy.
LUAD (100)	Risk model	4 CAF-related genes ( <i>FAP, FDGFRB, ACTA2, and NOTCH3</i> )	-in low-risk group, the proportion of stable disease/progressive disease patients was lower (0.31/0.64) than that in high-risk group (0.15/0.38) in two different datasets
LUAD (98)	CRD score model	23 circadian-related genes ( <i>ADA, C1QTNF7, CACNA2D2, CELF2, CLOCK, DIXDC1, EPHX1, FAT1, GIMAP5, HERPUD1, ICOS, KCNMA1, MICU3, NAALAD2, NAP1L5, NRXN1, OPHN1, PBXIP1, RBM38, RPS6KA5, SEC61G, SFTPC, and SLC25A42</i> )	-patients in high score group had higher response rate to TKIs, with AUCs of 0.927, 0.882, 0.829, 0.782, 0.836, and 0.853 between PD/NR, and RD patients
LUAD (99)	Gln risk model	10 Gln metabolism related genes ( <i>EPHB2, CAV2, RTN2, SCPEP1, UNC5D, FURIN, PITPNC1, CH25H, RGS20, and TSPAN11</i> )	-low risk group associated with higher response rates (95%) to anti-PD1 and anti-CTLA4 immunotherapy compared to those in high risk group (88%)
LUAD (94)	Expression level	<i>SLC7A11</i>	-patients with low <i>SLC7A11</i> expression levels were more likely to benefit from immunotherapy
LUAD (113)	Adenosine-related prognostic signature	3 adenosine signaling-related genes ( <i>CD4, FGR, and MNDA</i> )	-patients with high adenosine-related prognostic score had a poor PFS after immunotherapy

AT2, type II alveolar cells; ATNKGS, APC/T/NK cells-related gene signature; CAF, cancer-associated fibroblasts; CRD, circadian rhythm disruption; Gln, glutamine; ICD, immunogenic cell death; LPRI, Lung Cancer Prognostic Regulon Index; NETosis, Neutrophil Extracellular Traps process; NK, natural killer; OS, overall survival; PD/NR, progressive disease/nonresponse; PFS, progression-free survival; RD, response; TCMGrisk, T-cell marker genes risk score; TKIs, tyrosine kinase inhibitors; TT, tumor tissue.

12%) (99). Using a risk score based on a signature consisting of 4 CAF-related genes, Ren et al. (2023) demonstrated that low-risk patients were more likely to have a partial or complete response to PD-L1 blockade immunotherapy, as the proportion of stable disease/progressive disease patients was lower (0.31/0.64) in the high-risk groups compared to low-risk groups (0.15/0.38) in two different datasets (100). Similarly, another risk score calculated using the TIDE algorithm including a set of five CAF marker genes indicated that high-score patients had reduced CD8 and elevated CAF-markers expression, indicating these patients had a poor response to immunotherapy (101) (Table 3C). Resistance to

immunotherapy was investigated in a study in which NSCLC patients were categorized as those who experienced major pathologic response and those with non-major pathologic response. Differences among the two groups were observed in cell populations and highly expressed *ACTN4, ATF3, BRD2, CDKN1A, and CHMP4B* in epithelial cells of non-major response pathologic response patients represent potential biomarkers associated with resistance to immunotherapy in NSCLC (102). All these signatures indicate a large number of molecular indicators for diagnosis, outcome, and prediction of response to therapy, which need further exploration and validation in a clinical setting.

## Tumor heterogeneity characterization of the TME using scRNA-seq

Tumor heterogeneity, a major challenge in oncology and a primary cause of drug resistance (115), arises from differences among cancer cells in their transcriptional profiles, morphology, and metabolism. These variations occur both within individual tumors and between different tumors. In addition, recent data indicated that tumor heterogeneity also includes characterization of the TME, as well as cell-cell interactions and their dynamics as these factors significantly influence treatment efficiency in cancer patients (116). Drug resistance originating from tumor heterogeneity is encountered in all cancer types and it covers all therapeutic modes – chemotherapy, radiotherapy, immunotherapy, and targeted therapy (117), thus highlighting the relevance of characterizing the temporal and spatial evolution in tumor development and progression to unravel mechanisms of drug resistance. Using scRNA-seq enables heterogeneity assessment in tumor and adjacent non-tumor tissue and peripheral blood, to reveal knowledge regarding new cell populations with characteristic gene expression profiles, cell-cell communication, cellular differentiation, and gene regulation networks in the TME (118–120). The high resolution of individual cells assisted by scRNA-seq allows for an improved understanding of both intra- and inter-tumoral heterogeneity (121), where the characterization of the tumor-infiltrating immune cells are able to assist in predicting responsiveness to immunotherapies. The characterization of the tumor immune microenvironment (TIME) can be done considering the different populations and total number of immune cells present within a tumor, but also considering the spatial distribution of these cells to understand aspects related to recruitment and activation of immune cells, as means to develop new immunotherapies (38). Considering the observations indicating that sex-related differences in the prognosis of NSCLC patients are primarily due to immune response variation (122), a recent investigation (123) evaluated differences in cellular compositions and gene expression profiles in immune cells of the TME among men and women with NSCLC. The dissimilarities identified were in DEGs in tumor associated macrophages (TAMs), where *CIQC*, a complement-related gene is highly expressed in female-derived TAMs and associated with poor prognosis, while *FN1* and *SPP1* were highly expressed in male-derived TAMs. This suggests that male patients with higher expression of *SPP1* in TAMs are more likely to benefit from adjuvant immunotherapy, as this gene was reported as potential target for this type of treatment, improving its efficacy (123). When analyzing NSCLC subtypes, it was observed (76) that CD8, NK, and Gran cells had a common pattern of altered genes in the two main subtypes, LUAD and LUSC, while dendritic cells, T-regs, CD4 cells, and macrophages had fewer overlapping DEGs. A more detailed analysis of the macrophages showed that these cells had more immune functions, such as shaping the TME, in the progression in LUSC than in LUAD, with different dominant subclusters for the two subtypes, namely *FABP4*-macrophages in LUAD and *SPP1*-

macrophages in LUSC, the latter being reported as promoting to lung fibrosis. This evaluation indicates significant differences among LUAD and LUSC TIME, which can potentially be exploited for therapy (76, 124, 125). Another recent study indicated that analyzing NSCLC patients that were treated with chemotherapy combined with neoadjuvant immunotherapy based on TIME features can indicate patterns related to response to therapy, as the expansion of different immune cell types expresses the ability to target malignant cells. An immune signature of exhausted T cells was determined that was efficient in identifying patients at risk of recurrence that were categorized as non-major response (126). Although SCLC has long been considered a relatively homogeneous neuroendocrine carcinoma, recent single-cell RNA sequencing studies have demonstrated extensive intratumoral heterogeneity and striking lineage plasticity (75, 127–129). Tumor heterogeneity in SCLC early-stage tumors was investigated and it was observed that malignant cells had DEGs, such as downregulated genes involved in the immune response - *CD74*, *IDO1*, and *ISG15* – highlighting the existing (destructive) synergy between the immune and tumor cells (130). Moreover, the intertumor heterogeneity of SCLC tumors was reflected in differentially expressed transcription factors specific for SCLC subtypes - *ASCL1*, *NEUROD1*, and *POU2F3* (130). DEGs from circulating tumor cell (CTC) and CTC-derived xenografts from SCLC patients using single-cell profiling, indicated variable expression of *DLL3*, which is an inhibitory NOTCH ligand, a therapeutic target for which several targeted therapies are currently in clinical trials (chimeric antigen receptor (CAR) T cell therapy and CAR-NK). On the other hand, rovalpituzumab tesirine - an antibody-drug conjugate targeting *DLL3* was recently discontinued as it failed as a third-line treatment for SCLC-, as *DLL3* it was reported as having a higher expression level in SCLC, indicating limited efficacy on its target (131, 132). This implies that expression heterogeneity of *DLL3* may contribute to therapy resistance, as its expression was dynamic and in some cases faded after chemotherapy, suggesting a potential mechanism for adaptive resistance in targeted therapy and the importance of timing in administration of targeted therapy. Analyzing CTC collected at different time points of treatment advance, the lowest amount of CTCs was present in samples during treatment response, while samples collected after relapse had the highest number of CTCs (132). These findings and their relevance according to different lung cancer histotypes are synthesized in Table 4, considering the contribution of scRNA-seq analyses in characterizing tumor cells, lineage and evolution, immune microenvironment, stromal cells, and therapeutic implications (Table 4). In SCLC, further scRNA-seq investigations are needed to offer a broad image of intratumoral heterogeneity and its implications in resistance to immunotherapy, although the current limitations are widely acknowledged – rare primary human SCLC tumor samples (129), longitudinal data remain scarce (128), and the tumor immune landscape is incompletely characterized (75). Assessing the immune environment of tumors and identifying different subclasses based on their immune landscape has hence overall become essential for

TABLE 4 Relevance of scRNA-seq findings across lung cancer subtypes.

Category	LUAD	LUSC	SCLC
Tumor cell characterization	Identified multiple transcriptional tumor cell states (AT2-like, EMT-like, hypoxic, proliferative) that are not visible by bulk RNA or histology (133, 134).	Revealed squamous lineage states (basal vs differentiating) and stress-response programs, showing greater cellular heterogeneity than expected from morphology (89).	Discovered coexistence of neuroendocrine and non-neuroendocrine tumor states and dynamic switching between them, especially after treatment (128).
Lineage and evolution	Pseudotime analysis confirmed AT2 cell origin and mapped progression from pre-malignant to invasive states (133).	scRNA-seq traced lineage to airway basal stem cells and showed limited lineage plasticity compared to LUAD (89).	scRNA-seq demonstrated subtype plasticity over time, defining “state switching” as a mechanism of treatment resistance (75).
Immune microenvironment	Showed immune-inflamed TME with exhausted CD8 <sup>+</sup> T cells, plasma cells, and diverse macrophage populations not detected in bulk profiles (135).	Revealed that T cells are present but spatially excluded and suppressed by neutrophils and macrophages—information missed by single-cell-free platforms (136).	Showed immune desert features with almost complete absence of effector T cells and dominance of suppressive stromal/myeloid cells (130).
Stromal insights	Identified distinct CAF subtypes (inflammatory CAFs, ECM-rich CAFs) with defined signaling to immune cells (137).	scRNA-seq mapped dense CAF-immune interactions, especially TGF- $\beta$ and matrix programs driving immune exclusion (125).	Identified stromal signaling pathways that support neuroendocrine-to-non-neuroendocrine transitions during relapse (128).
Therapeutic implications	Mapping of resistant subclones (e.g., EGFR-TKI-persistent cells) and immune phenotypes linked to immunotherapy response (98).	Showed stromal/myeloid pathways (e.g., TGF- $\beta$ , CD36) as potential therapeutic targets rather than focusing only on checkpoint blockade (136).	Supported therapeutic strategies targeting cell-state transitions (e.g., NOTCH, MYC, YAP) rather than only immune-based treatment (132).

AT2, type II alveolar cells; CAF, cancer-associated fibroblasts; ECM, extracellular matrix.

further developments in immunotherapy, where scRNA-seq provides the data for resolving mechanisms related to immune-modulating therapies.

## Patient stratification based on T(I)ME architecture using scRNA-seq

Patient stratification can significantly increase the likelihood of clinical success (138) and hence there is hope that scRNA-seq data provides further opportunities in the future. This involves, in particular, the dynamics between immune and cancer cells, which enables us to explore the mechanisms underlying resistance to immunotherapy with TME involvement.

The majority of scRNA-seq studies that explored the TME in lung tumors were focused on characterizing T cell populations, as recent observations indicate that resistance to immune checkpoint inhibitors is associated with exhausted T cells phenotype (139). In this regard, in one study (140) that sequenced T-cells isolated from tumor and normal adjacent tissues and peripheral blood from NSCLC patients, the authors ascertained that based on the patterns of TILs patients could be divided into two groups, one enriched for pre-exhausted CD8<sup>+</sup> T cells, non-activated T-regs and activated CD4<sup>+</sup> cells, and the other one enriched for exhausted T cells and activated T-regs, with patients in the first group showing an improved outcome in terms of OS (140). Using both tumor tissue and adjacent non-tumor tissues from 7 untreated early-stage LUAD patients, the same clusters of T cells were found in both tumor and non-tumor tissues. Yet, a more detailed analysis revealed T cells in tumor tissues highly expressed exhaustion and regulatory markers, such as *TIGIT*, *LAYN*, *FOXP3*, and *CTLA4*. T cells from non-tumor tissues had a higher expression of markers for naïve and effector T-cell markers, indicating an immunosuppressive TME in

early-stage LUAD which was characterized by T cells that differentiate into exhausted and regulatory subtypes, indicating the importance of the TME profile and tumor-TME interactions for future drug discoveries in LUAD (141). In NSCLC, exploration of tumor tissues before and after immunotherapy combined with chemotherapy was investigated using scRNA-seq from 3 pre-treatment and 12 post-treatment patients. Analysis on how the TME of the immune system and cancer cells dynamics change in response to neoadjuvant PD-1 blockade combined with chemotherapy was performed. It was observed that in patients with major pathologic response, defined as less than 10% residual viable tumor cells present in HE (haematoxylin and eosin) staining following treatment (142), MHC-II genes were highly expressed, indicating that the MHC-II pathway is a possible mechanism used by the TME to boost response to immunotherapy. Moreover, the tumor microenvironment (TME) differed between patients with distinct pathologic responses to immunotherapy, distinguishing good from poor responders. In good responders, cytotoxic T-cells were activated and recruited, while immunosuppressive cells—such as T-regs, CCL3<sup>+</sup> neutrophils, and SPP1<sup>+</sup> TAMs—were less abundant. In contrast, poor responders showed cytotoxic cell activation only at the onset of therapy. In contrast, the immunosuppressive cells were more abundant in the TME. Hence, these observations suggest that different pathological responses in patients are reflected in significant differences in the TME remodeling following therapy (143). Another stratification of patients considering the cellular composition of the TME was performed by Bischoff et al. (2021) (135) in LUAD samples. The patients/tumors were divided into (1) those with normal-like myofibroblasts, conventional T cells, NK cells, non-inflammatory monocyte-derived macrophages, and myeloid dendritic cells in one group, and (2) those characterized by cancer-associated myofibroblasts, exhausted CD8<sup>+</sup> T cells, proinflammatory

monocyte-derived macrophages, and plasmacytoid dendritic cells into the other group. Based on this signature, the first group was characterized as possessing a TME inert pattern, while the latter showed a TME activated pattern. However, this separation did not precisely correspond to the separation based on histological grades of LUAD tumors, although a signature of 20 genes based on this TME pattern indicated that patients with inert TME pattern had a better OS than those with activated TME pattern (HR = 0.5, 95% CI 0.35-0.71,  $p = 0.0001$ ). These findings indicate that characterizing the TME at single-cell level represents a potential tool in assisting therapeutical decision making (135).

## Recent advances in drug and target discovery in lung cancer using scRNA-seq

Drug discovery in cancer remains difficult and expensive, not least due to an insufficient understanding of tumor biology, disease-related mechanisms, heterogeneity in disease response, and identification of efficient actionable therapeutic targets (82). One approach to match patient transcriptomics to compound treatment has been facilitated by the development of Connectivity Map (CMap) which, combined with integration of scRNA-seq data and cancer cell lines drug response analyses from the Library of Integrated Network-based Cellular Signatures (LINCS) has shown ways how to use cellular gene expression data for drug repurposing and target discovery of successful drugs able to target specific cell subpopulations (144, 145). However, the data in Cmap has been derived from cell culture systems, hence while it considers different cell types, it does not represent primary tumors, and also not tumor heterogeneity in the assay system used.

This section compiles recent examples of therapeutic agents and targets explored in lung cancer in *in vitro* and *in vivo* studies, which are listed in Table 5.

After assessing the transcriptomic heterogeneity of several LUAD cell lines, gene expression alterations in response to vandetanib were evaluated. It was found that even though vandetanib directly targets EGFR and RET, no changes in their expression patterns were observed after treatment. Thus, more in-depth determinations are needed to underlie the mechanisms of acquiring drug resistance in these cells, such as transcriptomic variances between individual cells using more cell types and different environmental conditions (147).

One significant difficulty in achieving better outcomes in patients with EGFR-mutant lung cancer is related to the persistence of drug-tolerant cancer cells. Therefore, to investigate the role of different cell populations in tumor recurrence and to trace the origins of these drug-tolerant persistent cells, tyrosine kinase inhibitor treatment in the EGFR-mutant lung cancer PDX model was applied in one study (148) which inhibited tumor growth and induced drug-tolerant persistent cells. Moreover, the treatment significantly stimulated these pathways and produced alterations in the CAF population. All these observations indicate that NF- $\kappa$ B and IL-6/JAK/STAT3 pathways represent potential targets to be further investigated to overcome recurrence in these patients (148).

In a study which aimed to evaluate the immune landscape of NSCLC to underline its importance in developing efficient immunotherapy approaches and for the prediction of clinical responses, scRNA-seq on 72,475 immune cells from 19 NSCLC patients was employed to determine their immune cell transcriptome atlas (76). After outlining the immune heterogeneity of LUAD and LUSC in terms of M $\phi$  and lymphocytes, a novel lymphocyte-related subcluster SELENOP-M $\phi$  characterized by increased expressions of *CD3D*, *FOLR2*, *IL32*, and *LTC4S*, which might play an antitumor role in LUAD, SELENOP-M $\phi$  is involved in the progression of LUAD via regulating peptide metabolism, protein transport and cytokine secretion. Considering the impact that immunotherapy might have on NSCLC, this study is essential since it allows us to understand how different immune responses can be explained by investigating and identifying subtype differences in the immune microenvironment of these patients (76). A four-gene prognostic signature was identified through scRNA-seq in a TCGA-LUAD cohort that can be independently used to predict outcomes (HR = 1.925; 95% CI: 1.405-2.63) and immune landscapes for LUAD patients. Furthermore, patients could be stratified considering three different immune patterns based on macrophage-related genes (149). A predictive model consisting of nine hub genes associated with the progression of LUAD and correlated with macrophages, B cells, CD8 + T cells, neutrophils, CD4 + T cells and dendritic cells was proposed by Zhong et al. (2021) to be correlated with immune infiltration and providing new potential clinical strategies for LUAD patients (150). While assessing the immune landscape and its role in the initiation and development of LUAD, another study focused on tumor-associated macrophages, it was observed that VEGFA, TIMP1 and SPP1 are critical regulatory molecules of macrophages in LUAD, being overexpressed in immunosuppression-associated macrophages and significantly associated with patient's survival. Moreover, an alteration in macrophage functions and a connection between the infiltration of macrophages in the immune microenvironment and tumorigenesis have been observed (151). Wang et al. demonstrated by integrating scRNA-seq and bulk RNA sequencing data that CD36 participates in regulating the differentiation of macrophages, and high expression of CD36 is correlated with poor prognosis of LUSC patients. Also, the study has performed a screening from the CTD database of small molecular compounds (e.g. estradiol, alitretinoin, dexamethasone, retinol acetate and cholesterol) that can decrease the CD36 expression. Their observations indicated these small molecules efficiently reduce CD36 expression, demonstrating their potential as a targeted treatment for LUSC (136). Hence, we can conclude that scRNA-seq technology is applied to make observations regarding some of the key endeavors (82) related to drug discovery – target identification (136, 150, 151), preclinical model selection (147, 148), and drug response and disease monitoring.

## Challenges

While providing a wealth of information, scRNA-seq also comes with several challenges, which will be only briefly mentioned here (for a more detailed review, see (152–154)). The most obvious challenge is

TABLE 5 Therapeutic effect investigated using scRNA-seq in different subtypes of lung cancer.

Drug/target	Cancer type	Study type/model	Effects	References
<b>Therapeutic agents</b>				
AKR1B1 knockdown using siRNA	NSCLC	<i>In vitro</i> - Human H2009 and SW480 cell lines Translational/tissue samples	Inhibition of cell proliferation Stimulation of cell apoptosis Identification of a subpopulation of mixed-lineage tumor cells	(146)
Vandetanib	NSCLC	<i>In vitro</i> - LC2/ad, LC2/ad (rep), LC2/ad-RPC-9 and VMRC-LCD cell lines	Diverse expression patterns of individual cells	(147)
Erlotinib	LUAD	PDX model	Drug-tolerant persisters (DTPs) induction	(148)
SELENOP-Mφ	NSCLC	Translational/40 samples of tumor and normal tissue	Inhibition of tumor progression	(76)
<b>Therapeutic targets</b>				
PTPRC, CCR2, SLAMF1, and HLA-DQA2	LUAD	Data from TCGA-LUAD cohort	Prediction of outcome and immune landscape	(149)
CBFA2T3, CR2, SEL1L3, TM6SF1, TSPAN32, ITGA6, MAPK11, RASA3, and TLR6	LUAD	Data from TCGA-LUAD cohort	Outcome prediction	(150)
VEGFA, TIMP1, and SPP1	LUAD	Data from GEO database	Association with poor prognosis	(151)
CD36	LUSC	Data from GEO Database	Regulation of immune and energy metabolism signals	(136)

represented by the high costs of both sample preparation and machines, which limit the number of samples profiled in many studies, often requiring the integration of the bulk RNA-seq data available in the public databases to deepen our insights in tumor biology (155). The existence of different sequencing platforms impacts cell assessment in scRNA-seq, as shortcomings may appear due to inefficient cell capture, different protocols may fail to generate a satisfactory number of reads for protein-coding genes, thus manipulating the generation of accurate gene expression profiles which affects reproducibility across studies (45). Although plenty of studies assess biomarkers with prognostic/diagnostic significance (as outlined above), most of these models are constructed using data from public datasets (111, 156) and still need to be validated in clinical studies. Since scRNA-seq is susceptible to sample quality, poor handling and suboptimal preservation often lead to confounding results (155). An alternative to using fresh tissue samples is using FFPE samples, which can be stored long-term, thus enabling large scale retrospective studies, although RNA degradation and fragmentation in FFPE samples can hinder data quality and interpretation (157). There are many studies assessing intratumoral heterogeneity dealing with shortcomings regarding the archived tissue samples as there are challenges to accurately compare cellular populations derived from different tumour sites, circulating tumour cells, and patient-derived xenografts (132). Although scRNA-seq is a comprehensive technology, no information about protein abundance or post-translational modifications are generated, indicating that an exhaustive characterization of single cells should also include proteomics and epigenomics (ATAC-seq) analyses (158). Moreover, tumor heterogeneity characterization, especially in the TME is affected by the loss of spatial information caused by tissue dissociation, thus

underlining the necessity of combining scRNA-seq with spatial transcriptomics technology to assess tumor heterogeneity and tumor infiltration levels, to offer important information, particularly for drug discovery in cancer (159). In addition, this technique allows only a limited number of cells to have their transcriptomic profile measured, and in these circumstances, lowly expressed genes may remain overlooked (155). Another restraint is encountered in rare malignancies, where there is only incomplete or insufficient reference data to annotate all cell types identified (160). In longitudinal studies, where therapy efficiency or resistance mechanisms are evaluated, clinically actionable conclusions are difficult to obtain, among other reasons, because paired samples are not always collected from the same patient, and insufficient clinical data are available for these patients (161). Finally, different ways of analysing data often lead to somewhat different results, partly due to the inconsistent implementation of algorithms in other software packages (162). Ultimately, analyzing data is only a starting point for a testable hypothesis, and experiments in a relevant biological system must follow up.

## Conclusion and future developments

In this work, we illustrated the use of scRNA-seq techniques in translational research, particularly in lung cancer. This holds great potential for both mechanistic and fundamental insights and practical applications to disease understanding and drug discovery. Some challenges concerning sample acquisition, data analysis, interpretation, and actionability remain, which we have briefly summarized and will hopefully be alleviated with further developments.

Improving the transparency and reproducibility of scRNA-seq studies starts with making the data truly usable for others—sharing the raw files, the processed matrices, and clear explanations of what was done at each step. Providing the analysis code, rather than only the final figures, helps others understand exactly how the results came together. Using established analysis pipelines can prevent avoidable errors, and being open about why certain filters or parameters were chosen makes the workflow easier to follow. Cell-type labels should be supported by straightforward marker evidence instead of relying entirely on automated tools. Just as importantly, the main conclusions should be tested in independent datasets to make sure they are not specific to a single experiment, especially given how much patient-to-patient variation exists. When data are easy to access, methods are easy to understand, and results can be checked elsewhere, single-cell studies become far more reliable and far more useful for researchers trying to turn these insights into real clinical advances. Patient diversity has a meaningful impact on what single-cell studies reveal, as tumors are widely different. Ethnicity, environmental exposures, as well as lung cancer subtype can all influence tumor cells behavior within the tumor. Future studies should deliberately include a broader range of patients and clearly describe groups represented in the analysis. By acknowledging and reporting diversity in this way, single-cell studies become more trustworthy, more inclusive, and better aligned with the real variability seen in clinical populations.

## Author contributions

CB: Conceptualization, Writing – original draft, Writing – review & editing. OZ: Writing – original draft. LP: Writing – original draft. CC: Writing – original draft. RL: Writing – original draft. AN: Visualization, Writing – original draft. IB-N: Conceptualization, Project administration, Writing – review & editing. AB: Conceptualization, Project administration, Writing – review & editing.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was granted by

## References

1. Foreman KJ, Marquez N, Dolgert A, Fukutaki K, Fullman N, McGaughey M, et al. Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: reference and alternative scenarios for 2016–40 for 195 countries and territories. *Lancet*. (2018) 392:2052–90. doi: 10.1016/S0140-6736(18)31694-5
2. Bray F, Jemal A, Grey N, Ferlay J, Forman D. Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. *Lancet Oncol*. (2012) 13:790–801. doi: 10.1016/S1470-2045(12)70211-5
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin*. (2021) 71:209–49. doi: 10.3322/caac.21660
4. Global Burden of Disease Cancer C, Kocarnik JM, Compton K, Dean FE, Fu W, Gaw BL, et al. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-

PNRR contract no. 760066/23.05.2023, cod 83/15.11.2022 entitled “Lung squamous cell carcinoma therapeutic targets using systems-level machine learning based on single cell RNA sequencing”.

## Conflict of interest

The authors declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

## Correction note

This article has been corrected with minor changes. These changes do not impact the scientific content of the article.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by *Frontiers* with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

adjusted life years for 29 cancer groups from 2010 to 2019: A systematic analysis for the global burden of disease study 2019. *JAMA Oncol*. (2022) 8:420–44. doi: 10.1001/jamaoncol.2021.6987

5. Roser M, Ritchie H, Spooner F. *Burden of disease*. Our World in Data (2021). Available online at: <https://ourworldindata.org/burden-of-disease>.

6. Khaltayev N, Axelrod S. Global lung cancer mortality trends and lifestyle modifications: preliminary analysis. *Chin Med J*. (2020) 133:1526–32. doi: 10.1097/CM9.0000000000000918

7. Shelton J, Zotow E, Smith L, Johnson SA, Thomson CS, Ahmad A, et al. 25 year trends in cancer incidence and mortality among adults aged 35–69 years in the UK, 1993–2018: retrospective secondary analysis. *BMJ*. (2024) 384:e076962. doi: 10.1136/bmj-2023-076962

8. Dagogo-Jack I, Shaw AT. Tumour heterogeneity and resistance to cancer therapies. *Nat Rev Clin Oncol*. (2018) 15:81–94. doi: 10.1038/nrclinonc.2017.166

9. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov.* (2022) 12:31–46. doi: 10.1158/2159-8290.CD-21-1059
10. Tufail M, Hu JJ, Liang J, He CY, Wan WD, Huang YQ, et al. Hallmarks of cancer resistance. *iScience.* (2024) 27:109979. doi: 10.1016/j.isci.2024.109979
11. Uhlen M, Zhang C, Lee S, Sjostedt E, Fagerberg L, Bidkhorji G, et al. A pathology atlas of the human cancer transcriptome. *Science.* (2017) 357. doi: 10.1126/science.aan2507
12. Stanta G, Bonin S. Overview on clinical relevance of intra-tumor heterogeneity. *Front Med (Lausanne).* (2018) 5:85. doi: 10.3389/fmed.2018.00085
13. Anderson NM, Simon MC. The tumor microenvironment. *Curr Biol.* (2020) 30:R921–5. doi: 10.1016/j.cub.2020.06.081
14. Geiger T. Tackling tumor complexity with single-cell proteomics. *Nat Methods.* (2023) 20:324–6. doi: 10.1038/s41592-023-01784-4
15. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Pineros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. *Int J Cancer.* (2021) 149:778–789. doi: 10.1002/ijc.33588
16. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin.* (2024) 74:229–63. doi: 10.3322/caac.21834
17. Li T, Kung HJ, Mack PC, Gandara DR. Genotyping and genomic profiling of non-small-cell lung cancer: implications for current and future therapies. *J Clin Oncol.* (2013) 31:1039–49. doi: 10.1200/JCO.2012.45.3753
18. Kris MG, Johnson BE, Berry LD, Kwiatkowski DJ, Iafrate AJ, Wistuba II, et al. Using multiplexed assays of oncogenic drivers in lung cancers to select targeted drugs. *Jama.* (2014) 311:1998–2006. doi: 10.1001/jama.2014.3741
19. Lengel HB, Connolly JG, Jones GD, Caso R, Zhou J, Sanchez-Vega F, et al. The emerging importance of tumor genomics in operable non-small cell lung cancer. *Cancers.* (2021) 13:3656. doi: 10.3390/cancers13153656
20. Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature.* (2012) 489:519–25. doi: 10.1038/nature11404
21. Santos ES, Rodriguez E. Treatment considerations for patients with advanced squamous cell carcinoma of the lung. *Clin Lung Cancer.* (2022) 23:457–66. doi: 10.1016/j.clcc.2022.06.002
22. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. *Nat Commun.* (2021) 12:2540. doi: 10.1038/s41467-021-22801-0
23. McGuire AL, McConechy MK, Melosky BL, English JC, Choi JJ, Peng D, et al. The clinically actionable molecular profile of early versus late-stage non-small cell lung cancer, an individual age and sex propensity-matched pair analysis. *Curr Oncol.* (2022) 29:2630–43. doi: 10.3390/curroncol29040215
24. Asmara OD, Hardavella G, Ramella S, Petersen RH, Tietzova I, Boerma EC, et al. Stage III NSCLC treatment options: too many choices. *Breathe (Sheff).* (2024) 20:240047. doi: 10.1183/20734735.0047-2024
25. de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, Yates L, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science.* (2014) 346:251–6. doi: 10.1126/science.1253462
26. Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science.* (2014) 346:256–9. doi: 10.1126/science.1256930
27. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun.* (2020) 11:2285. doi: 10.1038/s41467-020-16164-1
28. Salcher S, Sturm G, Horvath L, Untergasser G, Kuempers C, Fotakis G, et al. High-resolution single-cell atlas reveals diversity and plasticity of tissue-resident neutrophils in non-small cell lung cancer. *Cancer Cell.* (2022) 40:1503–1520.e1508. doi: 10.1016/j.ccell.2022.10.008
29. Leader AM, Grout JA, Maier BB, Nabet BY, Park MD, Tabachnikova A, et al. Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification. *Cancer Cell.* (2021) 39:1594–1609.e1512. doi: 10.1016/j.ccell.2021.10.009
30. Goveia J, Rohlenova K, Taverna F, Treps L, Conradi LC, Pircher A, et al. An integrated gene expression landscape profiling approach to identify lung tumor endothelial cell heterogeneity and angiogenic candidates. *Cancer Cell.* (2020) 37:21–36.e13. doi: 10.1016/j.ccell.2019.12.001
31. *Non-small cell lung cancer* (2025). Available online at: [https://www.nccn.org/professionals/physician\\_gls/pdf/nscl.pdf](https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf) (Accessed May 9, 2024).
32. *Non-small cell lung cancer NCCN clinical practice guidelines in oncology* (2025).
33. Ho DW, Tsui YM, Chan LK, Sze KM, Zhang X, Cheu JW, et al. Single-cell RNA sequencing shows the immunosuppressive landscape and tumor heterogeneity of HBV-associated hepatocellular carcinoma. *Nat Commun.* (2021) 12:3684. doi: 10.1038/s41467-021-24010-1
34. Puram SV, Tirosch I, Parikh 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) 171:1611–1624.e1624. doi: 10.1016/j.cell.2017.10.044
35. Li X, Wang CY. From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci.* (2021) 13:36. doi: 10.1038/s41368-021-00146-0
36. Gyanchandani R, Lin Y, Lin HM, Cooper K, Normolle DP, Brufsky A, et al. Intratumor heterogeneity affects gene expression profile test prognostic risk stratification in early breast cancer. *Clin Cancer Res.* (2016) 22:5362–9. doi: 10.1158/1078-0432.CCR-15-2889
37. Liu C, Pu M, Ma Y, Wang C, Kong L, Zhang S, et al. Intra-tumor heterogeneity and prognostic risk signature for hepatocellular carcinoma based on single-cell analysis. *Exp Biol Med (Maywood).* (2022) 247:1741–51. doi: 10.1177/15353702221110659
38. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* (2018) 24:541–50. doi: 10.1038/s41591-018-0014-x
39. Shen M, Kang Y. Complex interplay between tumor microenvironment and cancer therapy. *Front Med.* (2018) 12:426–39. doi: 10.1007/s11684-018-0663-7
40. Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med.* (2018) 50:1–14. doi: 10.1038/s12276-018-0071-8
41. Slovin S, Carissimo A, Panariello F, Grimaldi A, Bouche V, Gambardella G, et al. Single-cell RNA sequencing analysis: A step-by-step overview. *Methods Mol Biol.* (2021) 2284:343–65. doi: 10.1007/978-1-0716-1307-8\_19
42. Zhang Y, Wang D, Peng M, Tang L, Ouyang J, Xiong F, et al. Single-cell RNA sequencing in cancer research. *J Exp Clin Cancer Res: CR.* (2021) 40:81. doi: 10.1186/s13046-021-01874-1
43. Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, et al. Best practices for single-cell analysis across modalities. *Nat Rev Genet.* (2023) 24:550–72. doi: 10.1038/s41576-023-00586-w
44. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, et al. Comparative analysis of single-cell RNA sequencing methods. *Mol Cell.* (2017) 65:631–643.e634. doi: 10.1016/j.molcel.2017.01.023
45. Ashton JM, Rehrauer H, Myers J, Myers J, Zanche M, Balys M, et al. Comparative analysis of single-cell RNA sequencing platforms and methods. *J Biomol Tech.* (2021) 32:3fcl1f5fe.3eccea01. doi: 10.7171/3fcl1f5fe.3eccea01
46. Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc.* (2018) 13:599–604. doi: 10.1038/nprot.2017.149
47. Slatko BE, Gardner AF, Ausubel FM. Overview of next-generation sequencing technologies. *Curr Protoc Mol Biol.* (2018) 122:e59. doi: 10.1002/cpmb.59
48. Udaondo Z, Sittikankaew K, Uengwetwanit T, Wongsurawat T, Sonthirod C, Jenjaroenpun P, et al. Comparative analysis of pacBio and oxford nanopore sequencing technologies for transcriptomic landscape identification of penaeus monodon. *Life (Basel).* (2021) 11:862. doi: 10.3390/life11080862
49. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y. Single-cell RNA sequencing technologies and applications: A brief overview. *Clin Transl Med.* (2022) 12:e694. doi: 10.1002/ctm2.694
50. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol.* (2018) 36:411–20. doi: 10.1038/nbt.4096
51. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol.* (2023) 42:293–304. doi: 10.1038/s41587-023-01767-y
52. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* (2018) 19:15. doi: 10.1186/s13059-017-1382-0
53. Amezquita RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with Bioconductor. *Nat Methods.* (2020) 17:137–45. doi: 10.1038/s41592-019-0654-x
54. Su M, Pan T, Chen QZ, Zhou WW, Gong Y, Xu G, et al. Data analysis guidelines for single-cell RNA-seq in biomedical studies and clinical applications. *Mil Med Res.* (2022) 9:68. doi: 10.1186/s40779-022-00434-8
55. Lause J, Berens P, Kobak D. Analytic Pearson residuals for normalization of single-cell RNA-seq UMI data. *Genome Biol.* (2021) 22:258. doi: 10.1186/s13059-021-02451-7
56. Ringner M. What is principal component analysis? *Nat Biotechnol.* (2008) 26:303–4. doi: 10.1038/nbt0308-303
57. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol.* (2018) 37:38–44. doi: 10.1038/nbt.4314
58. Kobak D, Berens P. The art of using t-SNE for single-cell transcriptomics. *Nat Commun.* (2019) 10:5416. doi: 10.1038/s41467-019-13056-x
59. Sun S, Zhu J, Ma Y, Zhou X. Accuracy, robustness and scalability of dimensionality reduction methods for single-cell RNA-seq analysis. *Genome Biol.* (2019) 20:269. doi: 10.1186/s13059-019-1898-6
60. Zhang J, Lu T, Lu S, Ma S, Han D, Zhang K, et al. Single-cell analysis of multiple cancer types reveals differences in endothelial cells between tumors and normal tissues. *Comput Struct Biotechnol J.* (2023) 21:665–76. doi: 10.1016/j.csbj.2022.12.049
61. Dong S, Deng K, Huang X. Single-cell type annotation with deep learning in 265 cell types for humans. *Bioinform Adv.* (2024) 4:vbae054. doi: 10.1093/bioadv/vbae054
62. Yousuff M, Babu R, Rathinam A. Nonlinear dimensionality reduction based visualization of single-cell RNA sequencing data. *J Anal Sci Technol.* (2024) 15:1. doi: 10.1186/s40543-023-00414-0

63. Verma A, Engelhardt BE. A robust nonlinear low-dimensional manifold for single cell RNA-seq data. *BMC Bioinf.* (2020) 21:324. doi: 10.1186/s12859-020-03625-z
64. Lei Y, Tang R, Xu J, Wang W, Zhang B, Liu J, et al. Applications of single-cell sequencing in cancer research: progress and perspectives. *J Hematol Oncol.* (2021) 14:91. doi: 10.1186/s13045-021-01105-2
65. Das S, Rai SN. Statistical approach for biologically relevant gene selection from high-throughput gene expression data. *Entropy (Basel).* (2020) 22:1205. doi: 10.3390/e21111205
66. Wang T, Li B, Nelson CE, Nabavi S. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinf.* (2019) 20:40. doi: 10.1186/s12859-019-2599-6
67. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci United States America.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
68. Jindal A, Gupta P, Jayadeva, Sengupta D. Discovery of rare cells from voluminous single cell expression data. *Nat Commun.* (2018) 9:4719. doi: 10.1038/s41467-018-07234-6
69. Dal Molin A, Baruzzo G, Di Camillo B. Single-cell RNA-sequencing: assessment of differential expression analysis methods. *Front Genet.* (2017) 8:62. doi: 10.3389/fgene.2017.00062
70. Ma F, Zhang S, Song L, Wang B, Wei L, Zhang F. Applications and analytical tools of cell communication based on ligand-receptor interactions at single cell level. *Cell Biosci.* (2021) 11:121. doi: 10.1186/s13578-021-00635-z
71. Wang X, Almet AA, Nie Q. The promising application of cell-cell interaction analysis in cancer from single-cell and spatial transcriptomics. *Semin Cancer Biol.* (2023) 95:42–51. doi: 10.1016/j.semcancer.2023.07.001
72. Yost KE, Satpathy AT, Wells DK, Qi Y, Wang C, Kageyama R, et al. Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat Med.* (2019) 25:1251–9. doi: 10.1038/s41591-019-0522-3
73. Van den Berge K, Roux de Bezieux H, Street K, Saelens W, Cannoodt R, Saeyns Y, et al. Trajectory-based differential expression analysis for single-cell sequencing data. *Nat Commun.* (2020) 11:1201. doi: 10.1038/s41467-020-14766-3
74. Cheng C, Zhu G, Li Y, Wang H, Wang S, Li C, et al. Single-cell transcriptome analysis reveals cellular and molecular alterations in small cell lung cancer tumors following chemotherapy. *Int J Cancer.* (2023) 153:1273–86. doi: 10.1002/ijc.34629
75. Chan JM, Quintanal-Villalonga A, Gao VR, Xie Y, Allaj V, Chaudhary O, et al. Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell.* (2021) 39:1479–1496.e1418. doi: 10.1016/j.ccell.2021.09.008
76. Wang C, Yu Q, Song T, Wang Z, Song L, Yang Y, et al. The heterogeneous immune landscape between lung adenocarcinoma and squamous carcinoma revealed by single-cell RNA sequencing. *Signal Transduction Targeted Ther.* (2022) 7:289. doi: 10.1038/s41392-022-01130-8
77. Wang Y, Zhu Z, Luo R, Chen W. Single-cell transcriptome analysis reveals heterogeneity of neutrophils in non-small cell lung cancer. *J Gene Med.* (2024) 26:e3690. doi: 10.1002/jgm.3690
78. Good CR, Aznar MA, Kuramitsu S, Samareh P, Agarwal S, Donahue G, et al. An NK-like CAR T cell transition in CAR T cell dysfunction. *Cell.* (2011) 184:6081–6100.e6026. doi: 10.1016/j.cell.2021.11.016
79. Quah HS, Cao EY, Suteja L, Li CH, Leong HS, Chong FT, et al. Single cell analysis in head and neck cancer reveals potential immune evasion mechanisms during early metastasis. *Nat Commun.* (2023) 14:1680. doi: 10.1038/s41467-023-37379-y
80. Tan Z, Chen X, Zuo J, Fu S, Wang H, Wang J. Comprehensive analysis of scRNA-Seq and bulk RNA-Seq reveals dynamic changes in the tumor immune microenvironment of bladder cancer and establishes a prognostic model. *J Trans Med.* (2023) 21:223. doi: 10.1186/s12967-023-04056-z
81. Ding C, Yang X, Li S, Zhang E, Fan X, Huang L, et al. Exploring the role of pyroptosis in shaping the tumor microenvironment of colorectal cancer by bulk and single-cell RNA sequencing. *Cancer Cell Int.* (2023) 23:95. doi: 10.1186/s12935-023-02897-8
82. Van de Sande B, Lee JS, Mutasa-Gottgens E, Naughton B, Bacon W, Manning J, et al. Applications of single-cell RNA sequencing in drug discovery and development. *Nat Rev Drug Discov.* (2023) 22:496–520. doi: 10.1038/s41573-023-00688-4
83. Yu H, Liu Q, Jin M, Huang G, Cai Q. Comprehensive analysis of mitophagy-related genes in NSCLC diagnosis and immune scenery: based on bulk and single-cell RNA sequencing data. *Front Immunol.* (2023) 14:1276074. doi: 10.3389/fimmu.2023.1276074
84. Zhang H, Wang Y, Wang K, Ding Y, Li X, Zhao S, et al. Prognostic analysis of lung adenocarcinoma based on cancer-associated fibroblasts genes using scRNA-sequencing. *Aging.* (2023) 15:6774–97. doi: 10.18632/aging.204838
85. Kang KN, Koh EY, Jang JY, Kim CW. Multiple biomarkers are more accurate than a combination of carbohydrate antigen 125 and human epididymis protein 4 for ovarian cancer screening. *Obstet Gynecol Sci.* (2022) 65:346–54. doi: 10.5468/ogs.22017
86. Li X, Lu J, Ren H, Chen T, Gao L, Di L, et al. Combining multiple serum biomarkers in tumor diagnosis: A clinical assessment. *Mol Clin Oncol.* (2013) 1:153–60. doi: 10.3892/mco.2012.23
87. Holdenrieder S, Nagel D, Stieber P. Estimation of prognosis by circulating biomarkers in patients with non-small cell lung cancer. *Cancer Biomarkers: Sect A Dis Markers.* (2010) 6:179–90. doi: 10.3233/CBM-2009-0128
88. Sutic M, Vukic A, Baranasic J, Forstl A, Dzubur F, Samarzija M, et al. Diagnostic, predictive, and prognostic biomarkers in non-small cell lung cancer (NSCLC) management. *J Pers Med.* (2021) 11:1102. doi: 10.3390/jpm11111102
89. Zhang L, Zhang Y, Wang C, Yang Y, Ni Y, Wang Z, et al. Integrated single-cell RNA sequencing analysis reveals distinct cellular and transcriptional modules associated with survival in lung cancer. *Signal Transduction Targeted Ther.* (2022) 7:9. doi: 10.1038/s41392-021-00824-9
90. Song P, Li W, Guo L, Ying J, Gao S, He J. Identification and validation of a novel signature based on NK cell marker genes to predict prognosis and immunotherapy response in lung adenocarcinoma by integrated analysis of single-cell and bulk RNA-sequencing. *Front Immunol.* (2022) 13:850745. doi: 10.3389/fimmu.2022.850745
91. Shi X, Dong A, Jia X, Zheng G, Wang N, Wang Y, et al. Integrated analysis of single-cell and bulk RNA-sequencing identifies a signature based on T-cell marker genes to predict prognosis and therapeutic response in lung squamous cell carcinoma. *Front Immunol.* (2022) 13:992990. doi: 10.3389/fimmu.2022.992990
92. Sultana A, Alam MS, Liu X, Sharma R, Singla RK, Gundamaraju R, et al. Single-cell RNA-seq analysis to identify potential biomarkers for diagnosis, and prognosis of non-small cell lung cancer by using comprehensive bioinformatics approaches. *Transl Oncol.* (2023) 27:101571. doi: 10.1016/j.tranon.2022.101571
93. Zhang L, Zhang X, Guan M, Yu F, Lai F. In-depth single-cell and bulk-RNA sequencing developed a NETosis-related gene signature affects non-small-cell lung cancer prognosis and tumor microenvironment: results from over 3,000 patients. *Front Oncol.* (2023) 13:1282335. doi: 10.3389/fonc.2023.1282335
94. Wu X, Wang S, Chen K. Bulk RNA-seq and scRNA-seq reveal SLC7A11, a key regulatory molecule of ferroptosis, is a prognostic-related biomarker and highly related to the immune system in lung adenocarcinoma. *Medicine.* (2023) 102:e34876. doi: 10.1097/MD.00000000000034876
95. Yan Y, Teng H, Hang Q, Kondiparthi L, Lei G, Horbath A, et al. SLC7A11 expression level dictates differential responses to oxidative stress in cancer cells. *Nat Commun.* (2023) 14:3673. doi: 10.1038/s41467-023-39401-9
96. Cortes-Ciriano I, van Westen GJ, Bouvier G, Nilges M, Overington JP, Bender A, et al. Improved large-scale prediction of growth inhibition patterns using the NCI60 cancer cell line panel. *Bioinformatics.* (2016) 32:85–95. doi: 10.1093/bioinformatics/btv529
97. Ballman KV. Biomarker: predictive or prognostic? *J Clin Oncol.* (2015) 33:3968–71. doi: 10.1200/JCO.2015.63.3651
98. He L, Fan Y, Zhang Y, Tu T, Zhang Q, Yuan F, et al. Single-cell transcriptomic analysis reveals circadian rhythm disruption associated with poor prognosis and drug-resistance in lung adenocarcinoma. *J Pineal Res.* (2022) 73:e12803. doi: 10.1111/jpi.12803
99. Liu J, Shen H, Gu W, Zheng H, Wang Y, Ma G, et al. Prediction of prognosis, immunogenicity and efficacy of immunotherapy based on glutamine metabolism in lung adenocarcinoma. *Front Immunol.* (2022) 13:960738. doi: 10.3389/fimmu.2022.960738
100. Ren Q, Zhang P, Lin H, Feng Y, Chi H, Zhang X, et al. A novel signature predicts prognosis and immunotherapy in lung adenocarcinoma based on cancer-associated fibroblasts. *Front Immunol.* (2023) 14:1201573. doi: 10.3389/fimmu.2023.1201573
101. Wang S, Fan G, Li L, He Y, Lou N, Xie T, et al. Integrative analyses of bulk and single-cell RNA-seq identified cancer-associated fibroblasts-related signature as a prognostic factor for immunotherapy in NSCLC. *Cancer Immunol Immunother: CII.* (2023) 72:2423–42. doi: 10.1007/s00262-023-03428-0
102. Hu X, Wu Y, Wang L, Yang F, Ye L, Chen X, et al. Single-cell RNA sequencing reveals microenvironmental infiltration in non-small cell lung cancer with different responses to immunotherapy. *J Gene Med.* (2024) 26:e3736. doi: 10.1002/jgm.3736
103. Lin S, Zhou S, Han X, Yang Y, Zhou H, Chang X, et al. Single-cell analysis reveals exosome-associated biomarkers for prognostic prediction and immunotherapy in lung adenocarcinoma. *Aging.* (2023) 15:11508–31. doi: 10.18632/aging.205140
104. Xiong Y, Zhang Y, Liu N, Li Y, Liu H, Yang Q, et al. Integration of single-cell regulon atlas and multi-omics data for prognostic stratification and personalized treatment prediction in human lung adenocarcinoma. *J Trans Med.* (2023) 21:499. doi: 10.1186/s12967-023-04331-z
105. Jin X, Hu Z, Sui Q, Zhao M, Liang J, Liao Z, et al. A novel prognostic signature revealed the interaction of immune cells in tumor microenvironment based on single-cell RNA sequencing for lung adenocarcinoma. *J Immunol Res.* (2022) 2022:6555810. doi: 10.1155/2022/6555810
106. Sui P, Liu X, Zhong C, Sha Z. Integrated single-cell and bulk RNA-Seq analysis enhances prognostic accuracy of PD-1/PD-L1 immunotherapy response in lung adenocarcinoma through necroptotic anoikis gene signatures. *Sci Rep.* (2024) 14:10873. doi: 10.1038/s41598-024-61629-8
107. He D, Tang H, Yang X, Liu X, Zhang Y, Shi J. Elaboration and validation of a prognostic signature associated with disulfidoptosis in lung adenocarcinoma, consolidated with integration of single-cell RNA sequencing and bulk RNA sequencing techniques. *Front Immunol.* (2023) 14:1278496. doi: 10.3389/fimmu.2023.1278496

108. Wang H, Zhu X, Zhao F, Guo P, Li J, Du J, et al. Integrative analysis of single-cell and bulk RNA-sequencing data revealed disulfidptosis genes-based molecular subtypes and a prognostic signature in lung adenocarcinoma. *Aging*. (2024) 16:2753–73. doi: 10.18632/aging.205509
109. Yang J, Liu K, Yang L, Ji J, Qin J, Deng H, et al. Identification and validation of a novel cuproptosis-related stemness signature to predict prognosis and immune landscape in lung adenocarcinoma by integrating single-cell and bulk RNA-sequencing. *Front Immunol*. (2023) 14:1174762. doi: 10.3389/fimmu.2023.1174762
110. Huang L, Lou N, Xie T, Tang L, Han X, Shi Y. Identification of an antigen-presenting cells/T/NK cells-related gene signature to predict prognosis and CTSL to predict immunotherapeutic response for lung adenocarcinoma: an integrated analysis of bulk and single-cell RNA sequencing. *Cancer Immunol Immunother: CII*. (2023) 72:3259–77. doi: 10.1007/s00262-023-03485-5
111. Gu X, Cai L, Luo Z, Shi L, Peng Z, Sun Y, et al. Identification and validation of a muscle failure index to predict prognosis and immunotherapy in lung adenocarcinoma through integrated analysis of bulk and single-cell RNA sequencing data. *Front Immunol*. (2022) 13:1057088. doi: 10.3389/fimmu.2022.1057088
112. Tang L, Chen Z, Yang J, Li Q, Wang S, Mo T, et al. Single-cell and Bulk RNA-Seq reveal angiogenic heterogeneity and microenvironmental features to evaluate prognosis and therapeutic response in lung adenocarcinoma. *Front Immunol*. (2024) 15:1352893. doi: 10.3389/fimmu.2024.1352893
113. Han T, Wu J, Liu Y, Zhou J, Miao R, Guo J, et al. Integrating bulk-RNA sequencing and single-cell sequencing analyses to characterize adenosine-enriched tumor microenvironment landscape and develop an adenosine-related prognostic signature predicting immunotherapy in lung adenocarcinoma. *Funct Integr Genomics*. (2024) 24:19. doi: 10.1007/s10142-023-01281-z
114. Zhu P, Yang W, Wang B, Zeng T, Hu Z, Zhang D, et al. Systematic analysis of apoptosis-related genes in the prognosis of lung squamous cell carcinoma: a combined single-cell RNA sequencing study. *J Thorac Dis*. (2023) 15:6946–66. doi: 10.21037/jtd-23-1712
115. Vasan N, Baselga J, Hyman DM. A view on drug resistance in cancer. *Nature*. (2019) 575:299–309. doi: 10.1038/s41586-019-1730-1
116. Proietto M, Crippa M, Damiani C, Pasquale V, Sacco E, Vanoni M, et al. Tumor heterogeneity: preclinical models, emerging technologies, and future applications. *Front Oncol*. (2023) 13:1164535. doi: 10.3389/fonc.2023.1164535
117. Zhang A, Miao K, Sun H, Deng CX. Tumor heterogeneity reshapes the tumor microenvironment to influence drug resistance. *Int J Biol Sci*. (2022) 18:3019–33. doi: 10.7150/ijbs.72534
118. Aibar S, Gonzalez-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. (2017) 14:1083–6. doi: 10.1038/nmeth.4463
119. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc*. (2020) 15:1484–506. doi: 10.1038/s41596-020-0292-x
120. Sun B, Xun Z, Zhang N, Liu K, Chen X, Zhao H. Single-cell RNA sequencing in cancer research: discovering novel biomarkers and therapeutic targets for immune checkpoint blockade. *Cancer Cell Int*. (2023) 23:313. doi: 10.1186/s12935-023-03158-4
121. Huang T, Song X, Xu D, Tiek D, Goenka A, Wu B, et al. Stem cell programs in cancer initiation, progression, and therapy resistance. *Theranostics*. (2020) 10:8721–43. doi: 10.7150/thno.41648
122. Araujo JM, Prado A, Cardenas NK, Zaharia M, Dyer R, Doimi F, et al. Repeated observation of immune gene sets enrichment in women with non-small cell lung cancer. *Oncotarget*. (2016) 7:20282–92. doi: 10.18632/oncotarget.7943
123. Yang Q, Zhang H, Wei T, Lin A, Sun Y, Luo P, et al. Single-cell RNA sequencing reveals the heterogeneity of tumor-associated macrophage in non-small cell lung cancer and differences between sexes. *Front Immunol*. (2021) 12:756722. doi: 10.3389/fimmu.2021.756722
124. Morse C, Tabib T, Sembrat J, Buschur KL, Bittar HT, Valenzi E, et al. Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur Respir J*. (2019) 54:1802441. doi: 10.1183/13993003.02441-2018
125. Tong Y, Wang Y, Chen Y, Fan Y, Li H. Decoding the tumor immune microenvironment in lung squamous cell carcinoma: characteristics, regulatory mechanisms, and future directions in immunotherapy. *Trans Lung Cancer Res*. (2025) 14:4112–30. doi: 10.21037/tlcr-2025-350
126. Liu Z, Yang Z, Wu J, Zhang W, Sun Y, Zhang C, et al. A single-cell atlas reveals immune heterogeneity in anti-PD-1-treated non-small cell lung cancer. *Cell*. (2025) 188:3081–3096.e3019. doi: 10.1016/j.cell.2025.03.018
127. Gay CM, Stewart CA, Park EM, Diao L, Groves SM, Heeke S, et al. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. *Cancer Cell*. (2021) 39:346–360.e347. doi: 10.1016/j.ccell.2020.12.014
128. Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, et al. MYC drives temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate. *Cancer Cell*. (2020) 38:60–78.e12. doi: 10.1016/j.ccell.2020.05.001
129. Rudin CM, Poirier JT, Byers LA, Dive C, Dowlati A, George J, et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat Rev Cancer*. (2019) 19:289–97. doi: 10.1038/s41568-019-0133-9
130. Tian Y, Li Q, Yang Z, Zhang S, Xu J, Wang Z, et al. Single-cell transcriptomic profiling reveals the tumor heterogeneity of small-cell lung cancer. *Signal Transduction Targeted Ther*. (2022) 7:346. doi: 10.1038/s41392-022-01150-4
131. Rudin CM, Pietanza MC, Bauer TM, Ready N, Morgensztern D, Glisson BS, et al. Rovalpituzumab tesirine, a DLL3-targeted antibody-drug conjugate, in recurrent small-cell lung cancer: a first-in-human, first-in-class, open-label, phase 1 study. *Lancet Oncol*. (2017) 18:42–51. doi: 10.1016/S1470-2045(16)30565-4
132. Stewart CA, Gay CM, Xi Y, Sivajothi S, Sivakamasundari V, Fujimoto J, et al. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. *Nat Cancer*. (2020) 1:423–36. doi: 10.1038/s43018-019-0020-z
133. Han G, Sinjab A, Rahal Z, Lynch AM, Treekitkarnmongkol W, Liu Y, et al. An atlas of epithelial cell states and plasticity in lung adenocarcinoma. *Nature*. (2024) 627:656–63. doi: 10.1038/s41586-024-07113-9
134. Ran X, Tong L, Chenghao W, Qi L, Bo P, Jiaying Z, et al. Single-cell data analysis of Malignant epithelial cell heterogeneity in lung adenocarcinoma for patient classification and prognosis prediction. *Heliyon*. (2023) 9:e20164. doi: 10.1016/j.heliyon.2023.e20164
135. Bischoff P, Trinks A, Obermayer B, Pett JP, Wiederspahn J, Uhlitz F, et al. Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. *Oncogene*. (2021) 40:6748–58. doi: 10.1038/s41388-021-02054-3
136. Wang H, Pang J, Zhang S, Yu Q, Chen Y, Wang L, et al. Single-cell and bulk RNA-sequencing analysis to predict the role and clinical value of CD36 in lung squamous cell carcinoma. *Heliyon*. (2023) 9:e22201. doi: 10.1016/j.heliyon.2023.e22201
137. Wu Z, Luo M, Hu H, Jiang Z, Lu Y, Xiao ZJ. Hidden forces: the impact of cancer-associated fibroblasts on non-small cell lung cancer development and therapy. *J Trans Med*. (2025) 23:830. doi: 10.1186/s12967-025-06791-x
138. Wong CH, Siah KW, Lo AW. Estimation of clinical trial success rates and related parameters. *Biostatistics*. (2019) 20:273–86. doi: 10.1093/biostatistics/kxx069
139. Sen DR, Kaminski J, Barnitz RA, Kurachi M, Gerdemann U, Yates KB, et al. The epigenetic landscape of T cell exhaustion. *Science*. (2016) 354:1165–9. doi: 10.1126/science.aac0491
140. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med*. (2018) 24:978–85. doi: 10.1038/s41591-018-0045-3
141. He D, Wang D, Lu P, Yang N, Xue Z, Zhu X, et al. Single-cell RNA sequencing reveals heterogeneous tumor and immune cell populations in early-stage lung adenocarcinomas harboring EGFR mutations. *Oncogene*. (2021) 40:355–68. doi: 10.1038/s41388-020-01528-0
142. Travis WD, Dacic S, Wistuba I, Sholl L, Adusumilli P, Bubendorf L, et al. IASLC multidisciplinary recommendations for pathologic assessment of lung cancer resection specimens after neoadjuvant therapy. *J Thorac Oncol*. (2020) 15:709–40. doi: 10.1016/j.jtho.2020.01.005
143. Hu J, Zhang L, Xia H, Yan Y, Zhu X, Sun F, et al. Tumor microenvironment remodeling after neoadjuvant immunotherapy in non-small cell lung cancer revealed by single-cell RNA sequencing. *Genome Med*. (2023) 15:14. doi: 10.1186/s13073-023-01164-9
144. Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell*. (2017) 171:1437–1452.e1417. doi: 10.1016/j.cell.2017.10.049
145. Pruteanu LL, Bender A. Using transcriptomics and cell morphology data in drug discovery: the long road to practice. *ACS Med Chem Lett*. (2023) 14:386–95. doi: 10.1021/acscmedchemlett.3c00015
146. Li Q, Wang R, Yang Z, Li W, Yang J, Wang Z, et al. Molecular profiling of human non-small cell lung cancer by single-cell RNA-seq. *Genome Med*. (2022) 14:87. doi: 10.1186/s13073-022-01089-9
147. Suzuki A, Matsushima K, Makinoshima H, Sugano S, Kohno T, Tsuchihara K, et al. Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment. *Genome Biol*. (2015) 16:66. doi: 10.1186/s13059-015-0636-y
148. Moghal N, Li Q, Stewart EL, Navab R, Mikubo M, D’Arcangelo E, et al. Single-cell analysis reveals transcriptomic features of drug-tolerant persisters and stromal adaptation in a patient-derived EGFR-mutated lung adenocarcinoma xenograft model. *J Thorac Oncol*. (2023) 18:499–515. doi: 10.1016/j.jtho.2022.12.003
149. Li R, Tong R, Zhang Z, Deng M, Wang T, Hou G. Single-cell sequencing analysis and transcriptome analysis constructed the macrophage related gene-related signature in lung adenocarcinoma and verified by an independent cohort. *Genomics*. (2022) 114:110520. doi: 10.1016/j.ygeno.2022.110520
150. Zhong H, Wang J, Zhu Y, Shen Y. Comprehensive analysis of a nine-gene signature related to tumor microenvironment in lung adenocarcinoma. *Front Cell Dev Biol*. (2021) 9:700607. doi: 10.3389/fcell.2021.700607
151. Yu D, Zhang S, Liu Z, Xu L, Chen L, Xie L. Single-cell RNA sequencing analysis of gene regulatory network changes in the development of lung adenocarcinoma. *Biomolecules*. (2023) 13:671. doi: 10.3390/biom13040671
152. Huang D, Ma N, Li X, Gou Y, Duan Y, Liu B, et al. Advances in single-cell RNA sequencing and its applications in cancer research. *J Hematol Oncol*. (2023) 16:98. doi: 10.1186/s13045-023-01494-6

153. Chong ZX, Ho WY, Yeap SK, Wang ML, Chien Y, Verusingam ND, et al. Single-cell RNA sequencing in human lung cancer: Applications, challenges, and pathway towards personalized therapy. *J Chin Med Assoc: JCMA*. (2021) 84:563–76. doi: 10.1097/JCMA.0000000000000535
154. Aran D. Single-cell RNA sequencing for studying human cancers. *Annu Rev BioMed Data Sci*. (2023) 6:1–22. doi: 10.1146/annurev-biodatasci-020722-091857
155. Suva ML, Tirosh I. Single-cell RNA sequencing in cancer: lessons learned and emerging challenges. *Mol Cell*. (2019) 75:7–12. doi: 10.1016/j.molcel.2019.05.003
156. Guo S, Liu X, Zhang J, Huang Z, Ye P, Shi J, et al. Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels T cell-related prognostic risk model and tumor immune microenvironment modulation in triple-negative breast cancer. *Comput Biol Med*. (2023) 161:107066. doi: 10.1016/j.compbiomed.2023.107066
157. Gonzalez-Martinez S, Palacios J, Carretero-Barrio I, Lanza VF, Garcia-Cosio Piqueras M, Caniego-Casas T, et al. Single-cell RNA sequencing on formalin-fixed and paraffin-embedded (FFPE) tissue identified multi-ciliary cells in breast cancer. *Cells*. (2025) 14:197. doi: 10.3390/cells14030197
158. Sun G, Li Z, Rong D, Zhang H, Shi X, Yang W, et al. Single-cell RNA sequencing in cancer: Applications, advances, and emerging challenges. *Mol Ther Oncolytics*. (2021) 21:183–206. doi: 10.1016/j.omto.2021.04.001
159. Zhang L, Yang Y, Tan J. Applications and emerging challenges of single-cell RNA sequencing technology in tumor drug discovery. *Drug Discov Today*. (2025) 30:104290. doi: 10.1016/j.drudis.2025.104290
160. Yu X, Xie L, Ge J, Li H, Zhong S, Liu X. Integrating single-cell RNA-seq and spatial transcriptomics reveals MDK-NCL dependent immunosuppressive environment in endometrial carcinoma. *Front Immunol*. (2023) 14:1145300. doi: 10.3389/fimmu.2023.1145300
161. Zhang K, Erkan EP, Jamalzadeh S, Dai J, Andersson N, Kaipio K, et al. Longitudinal single-cell RNA-seq analysis reveals stress-promoted chemoresistance in metastatic ovarian cancer. *Sci Adv*. (2022) 8:eabm1831. doi: 10.1126/sciadv.abm1831
162. Rich JM, Moses L, Einarsson PH, Jackson K, Luebbert L, Booeshaghi AS, et al. The impact of package selection and versioning on single-cell RNA-seq analysis. *bioRxiv*. (2024). doi: 10.1101/2024.04.04.588111