Multi-omics to study chronic respiratory diseases and viral infections

Sobia Idrees[1], Hao Chen[1], Tayyaba Sadaf[1], Saima Firdous Rehman[1], Matt D. Johansen[1], Keshav Raj Paudel[1], Gang Liu[2], Yuting Wang[3], Malte D. Luecken[3,4], Elinor Hortle[1], Ashleigh S. Philp[5], Kurtis F. Budden[6], Matthew O'Rourke[1], Gerard E. Kaiko[6], Sionne E.M. Lucas[7], Joanne L. Dickinson[7], Peter C. Allen[8,9], Joseph E. Powell[8,10], Lai-Ying Zhang[11], Daniel C. Chambers[11], Tamera Corte[11], Gaetano Caramori[12], Maor Sauler[13], Peter A. Wark[14], Janine Gote-Schniering[15], Mareike Lehmann[4,16,17], Thomas M. Conlon[4], Theodore S. Kapellos[4,18], Ali Önder Yildirim[4,18], Rosa Faner[19], Shyamali C. Dharmage[20], Craig E. Wheelock[21], Maarten van den Berge[22], Martijn C. Nawijn[23], Francesca Polverino[24], Gabrielle T. Belz[25], Sanjay H. Chotirmall[26,27], Leopoldo N. Segal[28], Alen Faiz[22,29,30] and Philip M. Hansbro[1,6,30]

[1]Centre for Inflammation, Centenary Institute and University of Technology Sydney, Faculty of Science, School of Life Sciences, Sydney, Australia. [2]School of Life Sciences, Faculty of Science, University of Technology Sydney, Sydney, Australia. [3]Institute of Computational Biology, Helmholtz Munich, Neuherberg, Germany. [4]Institute of Lung Health and Immunity (LHI), Helmholtz Munich, Comprehensive Pneumology Center (CPC), Member of the German Center for Lung Research (DZL), Munich, Germany. [5]School of Clinical Medicine, UNSW Medicine and Health, St Vincent's Healthcare Clinical Campus, Sydney, Australia. [6]Hunter Medical Research Institute and University of Newcastle, Newcastle, Australia. [7]Menzies Institute for Medical Research, University of Tasmania, Tasmania, Australia. [8]Translational Genomics Program, Garvan Institute of Medical Research, Sydney, Australia. [9]Faculty of Medicine and Health, University of New South Wales, Sydney, Australia. [10]UNSW Cellular Genomics Futures Institute, University of New South Wales, Sydney, Australia. [11]School of Medicine, The University of Queensland and Queensland Lung Transplant Service, The Prince Charles Hospital, Brisbane, Australia. [12]Pulmonology, Department of Medicine and Surgery, University of Parma, Parma, Italy. [13]Division of Pulmonary, Critical Care, and Sleep Medicine, Yale School of Medicine, New Haven, CT, USA. [14]Respiratory Medicine, Central Clinical School, Monash University and AIRMED, Alfred Hospital, Melbourne, Australia. [15]Department of Rheumatology and Immunology, Department of Pulmonary Medicine, Allergology and Clinical Immunology, Inselspital, Bern University Hospital and Lung Precision Medicine, Department for BioMedical Research, University of Bern, Bern, Switzerland. [16]Institute for Lung Research, Philipps-University Marburg, Member of the German Center for Lung Research (DZL), Marburg, Germany. [17]Institute for Lung Health (ILH), Member of the German Center for Lung Research (DZL), Giessen, Germany. [18]Institute of Experimental Pneumology, Comprehensive Pneumology Center (CPC-M), Ludwig-Maximilians-University (LMU) Hospital, Member of the German Center for Lung Research (DZL), Munich, Germany. [19]Centro Investigación Biomédica en Red Enfermedades Respiratorias (CIBERES), Fundació Clinic Recerca Biomedica-IDIBAPS, University of Barcelona, Barcelona, Spain. [20]Allergy and Lung Health Unit, School of Population and Global Health, The University of Melbourne, Melbourne, Australia. [21]Unit of Integrative Metabolomics, Institute of Environmental Medicine, Karolinska Institutet and Department of Respiratory Medicine and Allergy, Karolinska University Hospital, Stockholm, Sweden. [22]University of Groningen, University Medical Center Groningen, Department of Pulmonary Diseases, Groningen, The Netherlands. [23]University of Groningen, University Medical Center Groningen, Department of Pathology and Medical Biology, Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, The Netherlands. [24]Pulmonary and Critical Care Division, Department of Medicine, Baylor College of Medicine, Houston, TX, USA. [25]The University of Queensland Frazer Institute, Faculty of Health, Medicine, and Behavioural Sciences, The University of Queensland, Woolloongabba, Australia. [26]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore. [27]Department of Respiratory and Critical Care Medicine, Tan Tock Seng Hospital, Singapore, Singapore. [28]Division of Pulmonary and Critical Care Medicine, NYU Grossman School of Medicine, New York, NY, USA. [29]RBMB Group, University of Technology Sydney, Faculty of Science, School of Life Sciences, Sydney, Australia. [30]Both authors contributed equally.

Corresponding author: Philip M. Hansbro (philip.hansbro@uts.edu.au)

Shareable abstract (@ERSpublications)
Integrating omics layers reveals complex biological systems, unveiling vital insights into chronic respiratory diseases and COVID-19. This holistic approach is crucial for developing more effective treatments and understanding intricate relationships. https://bit.ly/3VKFzJH

Cite this article as: Idrees S, Chen H, Sadaf T, *et al*. Multi-omics to study chronic respiratory diseases and viral infections. *Eur Respir Rev* 2026; 35: 240286 [DOI: 10.1183/16000617.0286-2024].

Received: 22 Dec 2024
Accepted: 30 Aug 2025

Abstract

Despite recent advances, the underlying mechanisms of the development and progression of many chronic respiratory diseases remain to be elucidated. Factors such as heterogeneity and complexity of human diseases and difficulty interpreting large datasets hinder research into chronic respiratory diseases. Omics assesses the changes in specific biological entities, such as mRNA expression, epigenetics/epigenomics, genomics, proteomics, metagenomics and metabolomics, and provides valuable insights into the roles of these processes in chronic respiratory diseases. High-throughput omics at bulk, single-cell and spatial levels empower the exploration of disease-related changes through untargeted data-driven statistical methods. Multi-omics is the exploration and integration of multiple biological processes, which compared to a single-omics, can provide a substantially greater and more holistic overview of the pathogenic mechanisms that underpin complex diseases. Multi-omics analysis can comprehensively characterise the mechanisms that drive chronic respiratory diseases, capturing unique biological signatures and cellular interactions at different omics levels. Use of these methods has begun to identify key factors and biomarkers in chronic respiratory diseases. Here, we review current omics approaches and highlight recent advances in respiratory research achieved using multi-omics and integrative methods. Our review provides a valuable resource for researchers and clinicians in this area.

Introduction

Recent advances in high-throughput omics methods, including transcriptomics, epigenomics, proteomics, microbiomes, and metabolomics have produced novel insights into the molecular and cellular mechanisms underlying chronic respiratory diseases, including bronchial asthma, COPD, idiopathic pulmonary fibrosis (IPF), and respiratory viral infections like coronavirus disease 2019 (COVID-19), rhinovirus and respiratory syncytial virus (RSV) and influenza viruses. The application of different omics to well-characterised clinical cohorts [1] has led to the development of methods to integrate each of these levels, termed multi-omics, which involves generating, analysing and integrating different modalities to comprehensively understand biological processes driving disease (figure 1) [1]. Multi-omics has advanced respiratory research by analysing big data collected from microarrays, bulk, single-cell and now spatial RNA sequencing (RNA-seq), proteomics, epigenetics/epigenomics, metagenomics, meta-transcriptomics and metabolomics, among others [2]. Multi-omics studies of chronic respiratory diseases have revealed key alterations in immune signalling, epithelial remodelling and metabolic pathways. These findings are crucial for understanding the biological changes underlying disease onset and progression and the response to therapeutic interventions [3]. Many datasets of the same omics level are publicly available and can be combined, resulting in greater statistical power, more reproducible findings and enhanced interpretation. Thus, based on single-omics profiles of patients and optimal animal models of disease [4, 5], more effective treatments can be developed [6–8].

Multi-omics is proposed to overcome the inherent limitations of each single omics. This is greatly improved when samples are matched from the same patient or use assays that generate multiple layers of

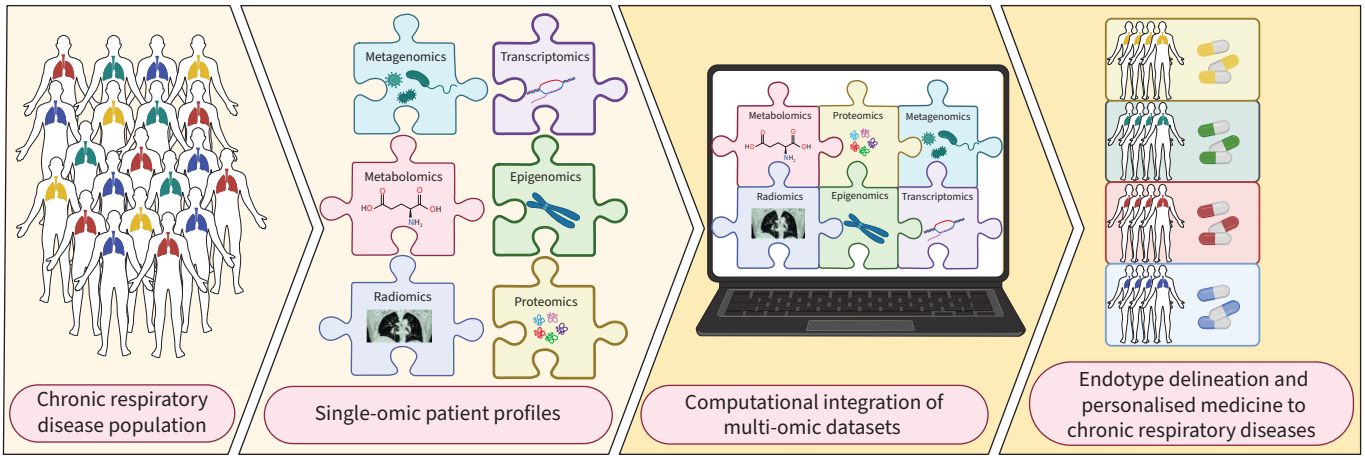

FIGURE 1 General overview of multi-omics layers including transcriptomics, proteomics, metagenomics, metabolomics, radiomics and epigenomics. Multi-omics level integration provides a comprehensive understanding of chronic respiratory diseases, can interconnect data sources, and guide tailored treatment strategies, improving precision and personalised medicine.

omics data from the same sample. This is further enhanced when sophisticated computational biology is applied [3, 9]. Such integration of technologies produces a more comprehensive biological picture, providing answers that single omics may not. However, while multi-omics has shown promise in identifying disease endotypes and uncovering biological diversity, translating these molecular insights into effective treatments remains challenging. The high cost, technical complexity, and need for extensive validation in clinical settings hinder the widespread application of these technologies. Nevertheless, continued advances in computational methods, standardised protocols and large-scale clinical studies could help overcome these obstacles, paving the way for multi-omics to become a powerful tool in precision medicine [10].

Current single-omic techniques

Epigenetics/epigenomics

Gene expression is regulated by complex and reversible modifications to DNA and histones, which influence transcriptional activity without altering the underlying DNA sequence. These modifications, studied through epigenomics, have critical roles in cellular identity and disease progression. Among them, methylation quantitative trait loci are single-nucleotide polymorphisms associated with variations in DNA methylation levels [11]. To analyse DNA methylation, several bulk techniques have been widely adopted. These include methylation arrays and sequencing-based approaches such as whole-genome bisulfite sequencing (WGBS) and reduced representation bisulfite sequencing [12–14]. Chromatin accessibility, another key epigenetic feature, is commonly assessed using DNase-seq (to detect DNase I hypersensitive sites), ATAC-seq, and FAIRE-seq (formaldehyde-assisted isolation of regulatory elements) [15]. Histone modifications are typically profiled using chromatin immunoprecipitation followed by sequencing (ChIP-seq) [16].

Recent advances in single-cell (sc) epigenetic profiling have enabled unprecedented resolution in understanding cellular heterogeneity. Techniques like scATAC-seq now employ combinatorial indexing and automated workflows to enhance throughput and sensitivity. Integrative platforms combining scRNA-seq with scATAC-seq provide comprehensive insights into gene regulation and cellular states, supporting scalable multi-omics analysis. Moreover, emerging methods such as scM&T-seq (single-cell methylome and transcriptome sequencing), snmCT-seq (profiles transcriptome, DNA methylome and chromatin accessibility of single nuclei using molecular partitioning), and scTrio-seq (captures genome, methylome and transcriptome data simultaneously) enable parallel profiling of multiple epigenetic layers. Additional assays like scNOMe-seq (nucleosome occupancy and methylome sequencing) and scCOOL-seq (integrating chromatin accessibility, nucleosome positioning and DNA methylation) offer high-resolution chromatin mapping. Other single-cell techniques such as bisulfite-sequencing and scWGBS provide enhanced coverage and accuracy for DNA methylation analysis [17, 18].

Thus, epigenetic changes may offer potential biomarkers for disease susceptibility and progression. However, the reversibility of epigenetic modifications suggests that therapies targeting epigenetic regulation could have profound implications, though the full therapeutic potential for chronic respiratory diseases remains underexplored.

Transcriptomics

Transcriptomics measures mRNA expression levels, which is the most used omics approach and, along with other omics techniques, has helped better understand chronic respiratory diseases (table 1). The early version of this technique was microarray-based [19], where >10 000 distinct DNA probe libraries are printed onto a chip that binds corresponding sample fragments that are detected using labelled molecules that bind to specific RNA sequences [20]. It assesses mRNA expression on a large scale as well as long noncoding (lnc) and micro-(mi)RNAs [21–24]. However, arrays capture only a portion of the transcriptome that is known and have a small dynamic range compared to more recent RNA-seq [25]. RNA-seq is the predominant method for quantifying large mRNA, mi-RNA and lncRNA (appropriate extraction and library preparation are needed) [26, 27]. It enables studies of entire transcriptomes and their dynamics and depending on the method of library preparation can provide detailed information on alternative splicing, and RNA editing. This improves understanding of gene regulation, protein diversity and the functionality of RNA processing molecules. This technology has been used to show how diverse RNA molecules actively contribute to cellular function and regulation, significantly advancing insights into their roles in health and disease. More recently, scRNA-seq has enabled assessment of cellular heterogeneity and dynamics in complex tissues and diseases [28, 29]. It achieves remarkably detailed characterisation of structural and infiltrating immune cells in healthy and diseased respiratory tracts [30]. This identifies previously unrecognised cells and their roles in chronic respiratory diseases [28, 31–36]. *In vitro* air–liquid interface cultures have provided further evidence to identify rare cell types, including ionocytes in primary

TABLE 1 Key features of different omics techniques

	Description	Pros	Cons	Applications
Epigenetics				
DNA methylation	Whole-genome bisulfite sequencing to detect DNA methylation patterns	Nongenetic inheritance Reversibility (epigenetic changes are reversible)	Epigenetic changes are reversible Targeting specific changes is difficult Epigenetic mechanisms are intricate and complex to understand Nongenetic inheritance needs further exploration	Study gene regulatory networks Cancer research Personalised medicine
Epigenomics				
ATAC-seq	Technique to detect open chromatin regions, offering more efficient processing compared to ChIP-seq, and can be applied at the single-cell level	Detection of open chromatin More efficient processing compared to ChIP-seq No antibody required Can do at single-cell level	Lower resolution compared to ChIP-seq Data quality can be affected by sample preparation Not as effective for studying protein–DNA interactions	Chromatin accessibility study Transcriptional regulation
ChIP-seq	Measures protein–DNA interactions by identifying histone modifications and transcription factors	Identify protein–DNA interactions Higher resolution compared to ATAC-seq	Labour-intensive Crosslinking artefacts can affect data accuracy Data quality dependent on antibody quality Higher resolution compared to ATAC-seq	Protein–DNA interaction study Gene and epigenetic regulation
Transcriptomics				
Microarray	A high-throughput technique used to measure the expression of thousands of genes simultaneously	High-throughput analysis of gene expression Established technology with widespread use	Limited to known probes; does not capture novel or rare transcripts Limited dynamic range Prone to cross-hybridisation	Biomarker discovery Pathway analysis
RNA-seq	Sequencing-based method to quantify gene expression by sequencing RNA transcripts	High sensitivity and broad dynamic range Detects alternative splicing and novel transcripts	Requires high computational resources for data analysis Library preparation biases can affect results	Biomarker discovery Pathway analysis
Single-cell RNA-seq	Allows RNA sequencing at single-cell resolution to measure gene expression heterogeneity	Provides insights into cellular heterogeneity Enables identification of rare cell populations High reproducibility and low input requirements	Limited sensitivity for low-input samples Higher cost per cell compared to bulk RNA-seq	Single-cell genomics Cell lineage tracing Differentiation studies
Spatial transcriptomics	Combines transcriptomic profiling with spatial information to study gene expression in tissue sections	Captures spatial gene expression within tissue context Allows mapping of gene expression patterns *in situ*	Requires high-quality tissue preservation; data analysis can be challenging due to high dimensionality Sensitivity may vary depending on the platform	Spatial transcriptomics studies Understanding tissue heterogeneity
Proteomics				
MS-based proteomics	Expensive, requires high technical expertise; challenges with data interpretation	High sensitivity and specificity High coverage	Expensive, requires high technical expertise; challenges with data interpretation	Biomarker discovery Drug discovery Disease mechanisms

Continued

https://doi.org/10.1183/16000617.0286-2024

TABLE 1 Continued

	Description	Pros	Cons	Applications
Spatial proteomics	Combines MS with imaging to map protein distribution across tissue sections	Provides a comprehensive view of protein distribution and localisation High resolution and high sensitivity	Requires high-resolution tissue sections; complex data analysis	Provides insights into cellular diagnosis with applications in disease diagnosis and treatment
Metabolomics				
MS-based metabolomics	MS used to analyse small molecules (metabolites) in biological samples	Can be compatible with imaging to make a spatial map of metabolites	Ionisation bias Complex sample preparation	Typically used for broad metabolite coverage
HPLC	Separation technique used for metabolite profiling, often combined with UV or MS detection	Compatible with MS Separation power	More expensive Sample matrix effect and interference from co-eluting compounds can affect data quality	Typically used for sample clean-up and fraction before MS analysis
NMR spectroscopy	Noninvasive technique used to identify and quantify metabolites based on their nuclear magnetic properties	Minimal sample preparation required Nondestructive process Provides quantitative value of metabolites without standards	Lower sensitivity Not high-throughput	Used when samples are needed for other processes after metabolomic profiling
Capillary electrophoresis	A separation technique that uses an electric field to separate metabolites based on their size and charge	Requires only small sample volumes Compatible with MS	Use limited to certain settings	Typically used for separation of small and charged molecules
Microbiome				
16S rRNA gene sequencing	Sequencing of the 16S ribosomal RNA gene to profile microbial communities in various samples	Cost-effective Targeted approach Broad taxonomic coverage	Provides limited resolution at species and strain levels due to short length of 16S rRNA Subject to PCR amplification biases and primer selection Cannot detect nonbacterial microbes	Typically used for taxonomic profiling of microbial communities
Metagenomics, metatranscriptomics, metaproteomics	Techniques for profiling the entire genomic, transcriptomic and proteomic content of microbial communities	Unbiased approach Strain-level detection Potential to identify novel molecules	Expensive Reconstruction of microbial genomes/transcriptomes/proteomes can be challenging, especially for low-abundance microbes and complex communities	Typically used for host–microbiome interactions

ATAC: assay for transposase-accessible chromatin; seq: sequencing; ChIP: chromatin immunoprecipitation; MS: mass spectrometry; HPLC: high-performance liquid chromatography; NMR: nuclear magnetic resonance; UV: ultraviolet.

https://doi.org/10.1183/16000617.0286-2024

bronchoepithelial cells [37]. In parallel, scRNA-seq has revealed additional cellular diversity across other lung compartments. One study subclassified pulmonary capillary endothelial cells into distinct populations, including aerocytes and general capillary endothelial cells [38]. Atypical basaloid cells were identified that co-express markers of basal epithelial, mesenchymal, senescence and development and are situated at the periphery of myofibroblast foci in IPF lungs [35]. Following the emergence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), scRNA-seq has been used on COVID-19 patient samples to reveal distinct alveolar macrophage and monocyte populations associated with disease severity [39, 40] and mechanisms of increased susceptibility to infection in COPD [28]. Nevertheless, standard scRNA-seq methods often introduce dissociation-induced biases. These arise during tissue processing, where enzymatic digestion can disproportionately capture more robust, dissociation-tolerant cells (typically immune cells) while under-representing fragile or tightly adherent cell types such as epithelial or stromal cells. Additionally, transcriptomic stress responses may be artificially induced during cell isolation. Single-nucleus (sn)RNA-seq addresses many of these limitations. By isolating nuclei instead of whole cells, snRNA-seq reduces dissociation bias, minimises the over-representation of immune cells, and enables the capture of more delicate or structurally embedded cell types, improving representation of the full cellular landscape of tissues [41].

Overlay of computational methods maximises the interpretation of single-cell data and is being extended to multi-omic analysis. Cell–cell communication (CCC) can be inferred using reference databases of known ligand/receptor pairs and then overlaying cell type (*e.g.* CellPhoneDB [42]) or condition (*e.g.* NicheNet [43]). NicheNet infers CCC by predicting how ligands secreted by one cell type influence gene expression in a target cell. It uses ligand–receptor interaction databases and gene expression data to identify and rank potential signalling pathways and key regulatory genes involved in cellular communication [43].

Trajectory inference methods analyse scRNA-seq data to reconstruct the dynamic developmental paths that cells follow over time. By modelling cellular progression through various states, these methods capture how cells transition from one state to another, such as from progenitor to differentiated cell types. With ∼70 methods currently available, approaches range from graph-based techniques to tree-based models, which include multifurcation and bifurcation structures as well as linear and cyclic trajectories, each suited to capturing different developmental or differentiation patterns. These methods provide insights into complex cellular processes, enabling a deeper understanding of their development and disease processes [44].

Where scRNA-seq data from identical samples is unavailable, leveraging publicly accessible reference atlases like the Human Lung Cell Atlas is a viable alternative [30]. Reference datasets for mice (and accurate mouse models [5, 45]) and humans are crucial foundations for annotating, inferring and interpreting multi-omics data in relevant contexts. As more scRNA-seq data are generated, tools are emerging to integrate datasets across laboratories, increasing statistical power to detect subtle changes in RNA and enabling more in-depth analysis [46, 47].

Proteomics

Proteomics explores protein interactions, composition, structures and cellular activity [2, 48, 49]. It uses tissue array and mass spectrometry (MS) (matrix-assisted laser desorption/ionisation (MALDI)-time-of-flight (TOF)/MS, liquid chromatography (LC)-MS/MS, surface-enhanced laser desorption/ionisation (SELDI)-TOF/MS, and two-dimensional electrophoresis (2-DE)/MS) to identify and verify proteins [50, 51]. Tissue array is rapid, but has target selection challenges, while MALDI-TOF/MS is favoured for accuracy and is untargeted. Untargeted LC-MS/MS is high-throughput with limitations in sensitivity, sample preparation requirements, variability and cost. ELISA and immunoblot are used for protein verification and quantification, with ELISA preferable for rapidity, sensitivity and accuracy [50, 52]. MS-based techniques are utilised to generate comprehensive data in chronic respiratory diseases [2, 53]. Their use has revealed disease-associated biomarkers (*e.g.* TPP1, CTSD, DPYSL2, TGM2) and therapeutic targets (RNA biosynthesis and binding proteins, S100A1, mitochondrial dysfunction, oxidative stress) [53–55] in experimental and human COPD, asthma, IPF and bronchiectasis [56]. Among the emerging high-throughput proteomics platforms, Olink has gained prominence for its ability to achieve multiplex protein biomarker analysis with high sensitivity and specificity. It uses a proprietary proximity extension assay, where pairs of oligonucleotide-labelled antibodies bind to specific target proteins. Upon binding, the proximity of the antibody pairs allows their attached DNA strands to hybridise and be extended by DNA polymerase, generating unique DNA barcodes. These barcodes are then quantified using quantitative PCR or next-generation sequencing, enabling precise protein detection across hundreds of targets from minimal sample volumes [57]. Complementing this, the SomaScan assay, an aptamer-based proteomics platform, has been employed recently to identify plasma biomarkers associated with COPD and to predict emphysema

risk [58]. Unlike antibody-dependent methods, SomaScan uses chemically modified DNA aptamers (SOMAmers) that bind selectively to thousands of proteins, providing extensive proteomic coverage with high specificity in complex biological samples [58].

Integrating RNA-seq with proteomics provides a more comprehensive understanding of chronic respiratory diseases. RNA-seq delivers high-resolution insights into gene expression at the mRNA level, and proteomics complements this by confirming protein-level changes and uncovering post-transcriptional regulatory mechanisms that are not captured by transcriptomic analysis alone [59].

Metabolomics

Metabolomics studies small-molecule metabolites involved in cellular metabolism [60, 61]. The most common methods used are nuclear magnetic resonance, MS, Raman spectroscopy and Fourier-transformed ion cyclotron resonance MS. MS-based techniques employ differential separation of molecules using gas or LC, or capillary electrophoresis, depending on the physicochemical properties of the molecules being assessed. A major challenge is the accurate annotation of metabolites, and the method used to confirm their identity must be transparent. Targeted approaches measure a predefined set of metabolites employing standard chemical libraries for higher certainty of annotation [62]. Untargeted approaches employ high-resolution MS for discovery, where metabolites are annotated using spectral databases [63] and/or molecular networking [64]. Cell-based metabolomics uses fluorescence-activated cell-sorting and immunomagnetic separation to isolate specific cell subsets, enabling targeted metabolite analysis [65]. The metabolic coverage of a method depends on the chemistry selected for the analysis, and issues of coverage and incorrect annotation can bias pathway mapping [66]. Several studies applied these techniques to chronic respiratory diseases and identified biomarkers in COPD [67–72]. One study showed that a childhood asthma subtype marked by early onset and heightened airway resistance is distinguished by diminished sphingolipid levels linked to 17q21 genetic variations and activity of serine palmitoyltransferase [73]. Another identified sphingomyelins and glycosphingolipids (including ceramide d18.1.N16.0, ganglioside GM3 d18.1.N16.0 and sphingomyelin d18.1.N16.0) were associated with COPD features [74]. Various traditional metabolome profiling approaches, particularly mass spectrometry techniques, have been applied to COPD. In the large SPIROMICS cohort, analysis of sputum supernatant metabolites revealed that sialic acid, hypoxanthine, xanthine, methylthioadenosine, adenine and glutathione were strongly associated with COPD patients experiencing exacerbations compared to those without. These findings suggest that metabolome-based biomarkers can differentiate between disease states and predict exacerbation risk. However, the predictive power of these biomarkers in terms of sensitivity, specificity and their integration into clinical practice remains an area of active investigation. Further studies are needed to validate these findings in independent cohorts and to assess their utility in guiding personalised treatment strategies or predicting long-term outcomes in COPD [75].

Challenges persist in conducting metabolomics studies in chronic respiratory diseases, primarily due to the lack of standardised sampling methods. For example, collecting airway epithelial lining fluids *via* bronchoscopy using normal saline lavage often results in a dilution of these fluids up to 100-fold. This varies between samples and is frequently overlooked in metabolomics studies, leading to inconsistencies.

With increasing recognition of the pivotal roles of metabolites in health and disease, spatial metabolomics is now increasingly being used to profile metabolites, lipids, drugs and other small molecules in cells, tissues and organisms [76].

Radiomics

Radiomics is a quantitative medical imaging approach that uses advanced mathematical analysis to extract textural information, offering additional information over visual interpretation. It enhances clinical decision-making and is being applied to early tumour detection, survival prediction, and assessing therapeutic response. There are technical challenges, a lack of standard criteria, and limited real-world experience that impact clinical outcomes [77]. Radiomics has been applied to chronic respiratory diseases, including nodules, lung cancer, obstructive and restrictive conditions, and interstitial lung diseases/IPF [78–81]. Radio-transcriptomics, integrating radiomics with transcriptomics, is a novel strategy offering insights into the dynamic interplay between tumour cell populations and the microenvironment. Demonstrating efficacy in predicting thrombosis and death in COVID-19 patients and responses to dexamethasone, radio-transcriptomics shows promise in developing noninvasive imaging biomarkers. The incorporation of single-cell or single-nuclei RNA-seq data may enhance the representation of disease heterogeneity. Ongoing advances in imaging and RNA-seq technologies, coupled with *in silico* methods, hold potential for constructing advanced models with clinical and translational value [82].

Metagenomics

Understanding the composition and function of respiratory microbial communities is achieved using 16S rRNA and metagenomic approaches. They define the phylogenetic, taxonomic and genetic characteristics of taxa and decipher the roles of the microbiome in health and disease [83]. Sequencing of bacterial 16S rRNA amplicons enables cost-effective and rapid taxonomic classification using PCR amplification with primers targeting variable regions of the bacterial 16S ribosomal gene. This requires only modest sequencing depth as it assesses a specific bacterial genome region [84]. Shotgun metagenomics involves whole DNA sequencing of the genetic content within samples, enabling the assessment of bacteria, DNA viruses and eukaryotic DNA (including fungi). It produces an unbiased and more in-depth taxonomic analysis by sequencing diverse regions of microbial genomes [85].

Now, microbiome data is being integrated with other omics to generate a holistic view of chronic respiratory diseases [86] in discovery and clinical investigations [87]. This suggests the potential of integration to identify distinct molecular subphenotypes, enabling more targeted and personalised treatment. It emphasises the importance of integrating data from various molecular levels and anatomical locations for a comprehensive understanding of chronic respiratory diseases.

Spatial omics

scRNA-seq, proteomics and other omics have unveiled many valuable biomolecular changes in chronic respiratory diseases, but lack spatial context due to the need for tissue dissociation. Spatial transcriptomics has emerged that preserves cell location and reveals spatially resolved cell interactions in chronic respiratory disease pathogenesis [87]. It involves sequencing- and imaging-based methods that are targeted or untargeted (*e.g.* whole transcriptome *versus* specific panels) [87, 88]. Platforms like Visium and GeoMx employ whole-transcriptome RNA-seq with trade-offs in RNA capture efficiency and single-cell resolution. SCRINSHOT, Xenium and CosMx use single-molecule fluorescence *in situ* hybridisation and imaging with machine learning for targeted single-cell resolution [89, 90]. Platform selection involves balancing whole transcriptome coverage and single-cell resolution. One study used Visium to spatially resolve transcriptomes and cell types in healthy human lungs, providing a reference study for chronic respiratory disease research [91]. Another used GeoMx to reveal interactions between cytotoxic lymphocytes and inflammatory macrophages in COVID-19 [92]. GeoMx was also used to define transcriptional distinctions between fibroblastic foci, fibrous regions and normal areas in IPF [93] and identify new fibrogenic biomarkers (*i.e.* MMPs, CDK4/6 inhibitors). SCRINSHOT was used to generate a reference spatially resolved map of developing human lung tissue [94]. scRNA-seq data can be integrated with spatial information computationally, quantifying the proportions of various cell types in given spatial regions [95]. Applying scRNA-seq and spatial transcriptomics to the same biological samples from a single experiment gives utmost precision in cell type identification. Where this is not possible, reference atlases are used [30]. These combinations can be applied to any omics.

Approximately 50% of proteins have multiple subcellular locations, contributing to intricate spatial proteomic landscapes [96]. High-throughput imaging and quantitative MS decipher these subcellular protein networks. Untargeted MS proteomics and imaging (MSI) provides complementary spatial proteomics data applicable to serial tissue sections [97, 98]. Understanding the dynamic spatial proteome is challenging, but yields valuable insights into cellular diagnosis, with applications in disease diagnosis and treatment [99].

Serial-section analysis is often necessary, particularly for untargeted studies, because single-omics methods are typically destructive and cannot preserve the sample for further analysis. However, some spatial platforms overcome this limitation by enabling simultaneous detection of targeted proteins, including 10x Genomics, NanoString, SM-Omics, spatial-CITE-seq and SPOTS [100–102]. Computational integration of spatial transcriptome and proteome data is challenging. It is typical to define the concomitant presence of transcripts and proteomes and perform pairwise correlations. One investigation involving bulk RNA-seq and untargeted proteomics in human lung cells revealed that only ∼40% of mRNA–protein pairs had matched expression, and correlations were relatively low (Spearman rank coefficient ∼0.4) [103]. This indicates nonlinearity between different biomolecular layers in the lungs. Thus, when working with multiple types of data, finding common features that exist across them and using them as integration anchors helps ensure that integration is meaningful. It may be possible to integrate RNA-seq and proteomics with metabolites, lipids and other omics to further understand chronic respiratory diseases.

Recently, spatial metabolomics was used to identify glycogen as a target in IPF [104]. Accompanying MSI generates extensive hyperspectral imaging data, prompting the development of specialised computational methods to evaluate metabolomics in image analysis [76]. Another study used this to find blunted N-linked

glycan in fibrotic compared to wild-type mice showing that lysosomal utilisation of glycogen is required for the progression of pulmonary fibrosis [104]. Finally, desorption electrospray ionisation-multiple-reaction-monitoring mass spectrometry was recently developed to perform spatially resolved quantification of oxylipins in lung tissue [105].

Characterisation of biological processes

Multi-omics and computational methods have been developed recently to infer biological processes from data, identifying signalling pathways and potential protein–protein interactions (PPIs) [106–110], *e.g.* BioGRID [111], IntAct [112] and STRING [113]. PPIs are essential in all cellular processes [114], and aberrant ones are involved in various chronic respiratory diseases. However, the text-mining approaches used have limitations in accuracy and context-dependent information. Additionally, reliance on pre-existing knowledge may bias the source, and lead to unreliable predictions [106]. Targeting PPIs is an important strategy for developing new drugs. Computational methods can virtually screen libraries of compounds against protein active sites to predict their likelihood of binding and identify PPI modulators that may be drugs [115]. Such techniques were used to develop PPI modulators of MDM2/p53 [116] and TCF/β-catenin [117].

Disruption in cell signalling pathways can lead to improper decoding of cellular messages and cause disease. Thus, identifying and characterising intercellular signalling pathways is another area of intense research [118]. Construction of networks wherein genes exhibiting co-expression patterns are interconnected can identify essential disease pathways and signatures and factors that control them. Computational methods like weighted gene co-expression network analysis [119], co-expression centrality analysis [120], and gene whole co-expression network analysis [121] are employed to construct these networks (figure 2). One study developed a multiplex network approach to study 20 million gene relationships and identified distinct phenotypic modules in 3771 rare diseases [122]. Thus, these methods provide a holistic understanding of the collective involvement of groups of genes in specific biological processes or pathways associated with disease.

This is achieved using computational methods, such as transcription factor enrichment analysis, which detects positional motif enrichments related to transcription [123, 124]. Other methods include ChEA3 [125] and VIPER [126]. Other studies used network and enrichment tools to infer new transcription factors

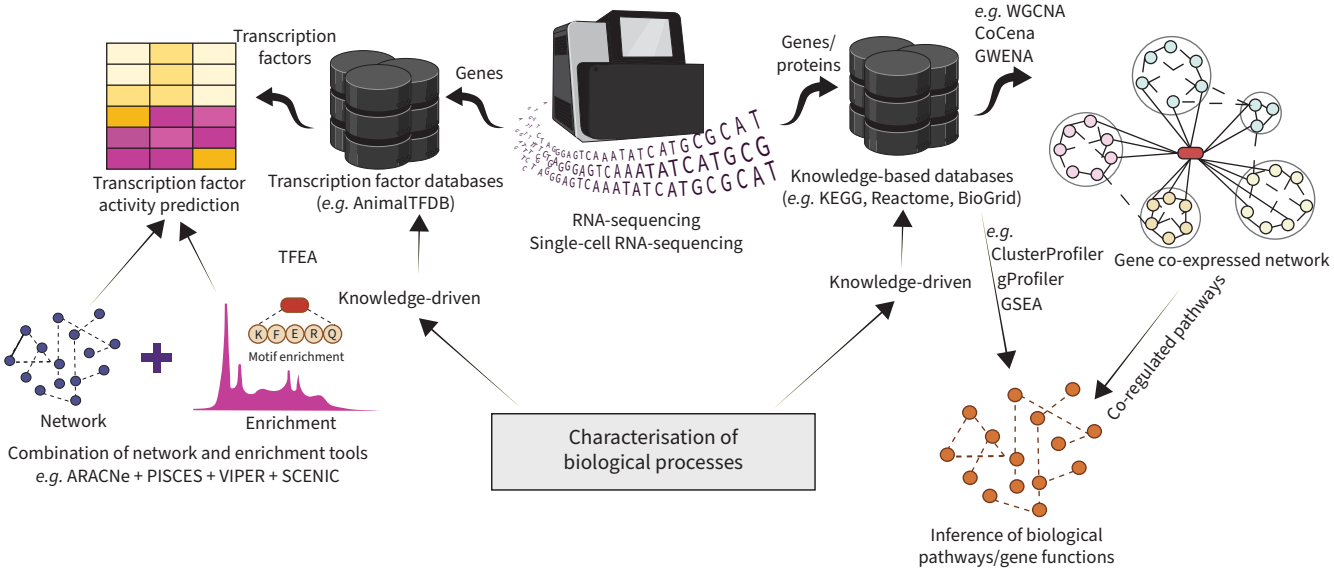

FIGURE 2 Methods used to characterise biological processes. Knowledge-based databases (*e.g.* Kyoto Encyclopedia of Genes and Genomes (KEGG)/Reactome) are used to predict biological pathways using gene and protein information and/or to construct co-expressed networks and infer co-regulated pathways. Moreover, transcription factor databases (*e.g.* Animal transcription factor database (AnimalTFDB)) are used to infer transcription factor activity using various methods such as transcription factor enrichment analysis (TFEA) or a combination of network and enrichment methods (*e.g.* ARACNe, VIPER, PISCES, SCENIC). WGCNA: weighted gene co-expression network analysis; CoCena: co-expression centrality analysis; GWENA: gene whole co-expression network analysis; GSEA: gene set enrichment analysis. Figure created using BioRender.com.

from scRNA-seq data [127] and ARACNe [128], PISCES, VIPER [129] and SCENIC [130] to infer transcription factor activity and enrichment in different cells and diseases [129] (figure 2). Regulating gene expression involves complex transcriptional programmes driven by transcription factors and epigenetic events, but defining this coordinated control of cell-type-specific gene expression remains challenging. New computational methods (*e.g.* AnnoMiner) are being developed to annotate and integrate transcription factor and epigenetic data that can help assess gene regulation and identify functional regulatory elements, which in turn can enhance the understanding of how transcription factors and epigenetic modifications control gene expression [131]. In short, these methods have substantially improved our understanding of biological processes by identifying key molecules driving diseases. However, findings are predictions and require validation *in vivo* (animal, human) and *in situ* to ensure accuracy and reliability.

Multi-omics integration

Multi-omics data can be measured at different resolutions: at the sample level (bulk) or the cellular level (single-cell). The choice of multi-omics data integration methods must be tailored to the resolution and the type of data that is measured. Thus, different methods are required for different data types (figure 3) [132].

Integrating multi-omics data at the sample level

Bulk multi-omics data provides an average profile across cells within a sample. Thus, bulk multi-omics data integration is used to understand variation at the tissue or sample level. In disease research bulk multi-omics integration can be used to discover biomarkers (such as transcripts, DNA, genes, proteins and metabolites that indicate whether a biological activity is normal or abnormal [133]) or stratify patients. Bulk data integration methods can be classified into five categories: regression/association-based methods, multiple kernel learning, Bayesian-based methods, network-based methods and matrix factorisation-based methods.

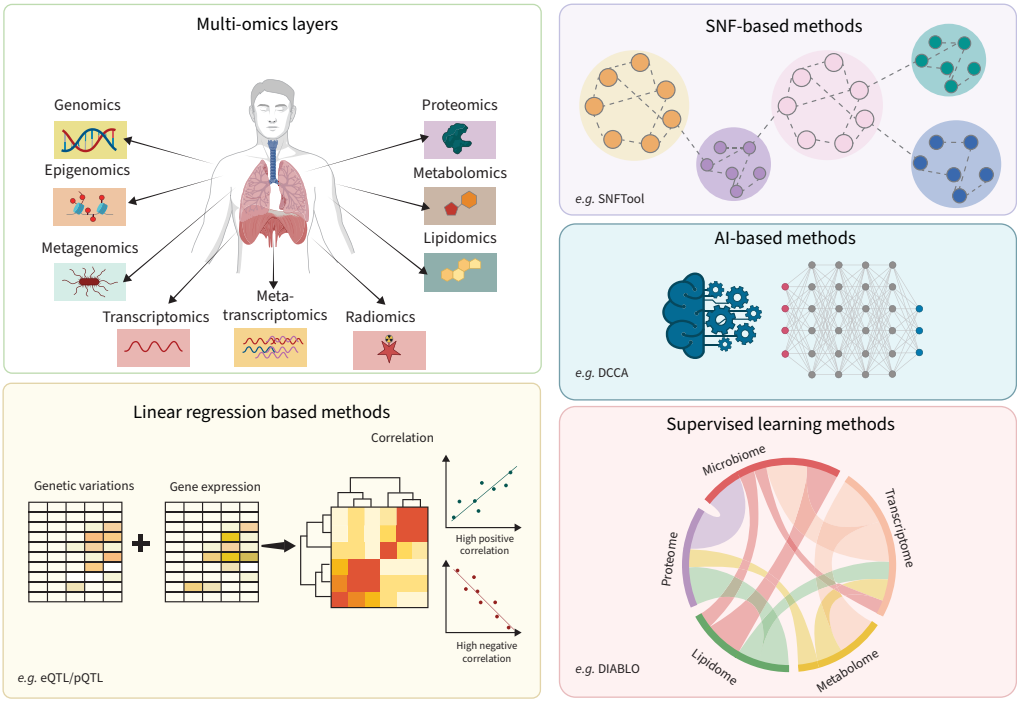

FIGURE 3 Various multi-omic layers including genomics, epigenomics, metagenomics, transcriptomics, meta-transcriptomics, radiomics, proteomics, metabolomics and lipidomics are currently being used to study respiratory diseases. Current multi-omics methods include correlation approaches for two-layered integration, similarity network fusion (SNF) that involves constructing similarity networks independently for each data type and then fusing these networks to create an integrated similarity network. Artificial intelligence (AI) (*e.g.* machine learning/neural network)-based multi-omics methods that can generate missing data based on the learned latent representation of other omics layers (*i.e.* transcriptomics, epigenetics), and supervised learning methods are being adapted such as mixOmics, Anvi'o, and integrOmics. eQTL: expression quantitative trait loci; pQTL: protein quantitative trait loci. Figure created using BioRender.com.

Regression/association-based methods [134, 135] contribute to biomarker discovery by identifying associations/correlations between different omics layers [136]. Association-based approaches [135, 137] have been used to link single-nucleotide polymorphisms to quantitative trait loci (QTLs) for gene expression (eQTLs). These associations help to identify genetic factors responsible for certain diseases and uncover the pathways impacted by these factors [137, 138]. Multiple-kernel learning methods [139] are used to integrate different types of omics data (*e.g.* genomics, proteomics, transcriptomics) by creating separate "kernels" for each data type and then combining them into a single, unified kernel matrix for analysis. In the weight-boosted kernel [140], features significantly upregulated or downregulated in samples are given higher weights, because they reflect key disease-related molecular changes. This ensures these biologically relevant features have greater influence in constructing kernels for each omics data type. These features are assigned higher weights, meaning they have more influence when constructing the kernels for each omics data type [141]. Separate kernels are integrated into a unified matrix, which is used for clustering to identify disease subtypes. This approach helps to prioritise the most relevant and informative data sources, improving the accuracy and robustness of biomarker identification and improving disease classification [140]. Thus, this method enhances the potential for more personalised and effective therapeutic strategies [139].

Integrating between five and seven omics data-blocks has proven effective in diagnosing COPD by providing comprehensive maps of the molecular alterations associated with the disease. This approach enables the identification of complex biomarkers and molecular signatures that are more specific and sensitive than traditional diagnostic methods, facilitating a deeper understanding of disease heterogeneity and enabling the differentiation of COPD subtypes. While spirometry remains a standard diagnostic tool, multi-omics provides insights into disease mechanisms, progression and potential therapeutic targets that cannot be achieved by spirometry alone. These insights can inform personalised treatment strategies, ultimately improving patient outcomes. However, the clinical utility of multi-omics is currently limited due to high costs, technical complexity, and the need for further validation in larger cohorts. Despite these challenges, the potential to enhance disease understanding and guide more targeted therapies justifies investment in these technologies, especially as advances in computational methods and reduced costs increase accessibility and clinical applicability in the future [86]. For patient stratification, Bayesian-based methods [142] formulate probabilistic models that capture the distribution and correlation of various omics datasets. Probabilistic stratification uses Dirichlet process (infinite mixture) models to integrate copy number variation and gene expression data, enabling joint classification of patients into disease subgroups [142]. While initially applied to cancer, similar Bayesian approaches have been extended to complex chronic respiratory diseases. Shaddox *et al.* [143] applied a Bayesian hierarchical graphical modelling framework to gene expression and metabolite data from subjects with varying COPD severity, capturing complex inter-omics relationships pertinent to disease progression. These applications underscore the versatility of such models in integrating heterogeneous omics data to elucidate disease mechanisms and support precision medicine in chronic respiratory diseases. Network-based methods [144] construct networks where nodes represent biological entities (*e.g.* genes, proteins, cells) and edges represent interactions or associations between them. These methods capture complex, nonlinear dependencies in biological systems. In COPD, one study used a network-based approach to integrate transcriptomics, proteomics, and PPI data, improving the accuracy of disease classification. Key genes like CXCL11 and TLR2 were identified, and pathways such as glycosaminoglycan signalling were enriched, providing insights into molecular mechanisms [145]. PARADIGM [144], a pathway-based multi-omics integration method, was used to successfully identify consistent pathway-level activities in subsets of patients with glioblastoma multiforme, that are often overlooked when analysing genes individually. While initially developed for cancer, this approach can also be applied to chronic respiratory diseases where understanding pathway perturbations may help uncover molecular endotypes. Matrix factorisation methods [146] project the variations among different omics layers onto a dimension-reduced space. iCluster+ [146] has been used to integrate DNA methylation with other omics data, revealing two distinct subsets in lung adenocarcinoma. This highlighted the role of chromatin maintenance in the disease and provided insights into its molecular subgroups [147].

Integrating RNA-seq and proteomics provides a comprehensive view of gene expression by linking transcriptomic data to protein abundance, but it faces several methodological challenges. RNA-seq measures mRNA levels using high-throughput sequencing, while proteomics employs MS to quantify proteins. Correlation analysis, pathway enrichment and machine learning are commonly used to connect transcript and protein data, enabling insights into gene regulation and post-translational modifications [148]. However, this integration is complicated by differences in sensitivity and dynamic range between RNA-seq and MS, as well as the fact that mRNA levels do not always correlate directly with protein abundance due to translational control and protein degradation. Moreover, proteomics may fail to detect low-abundance proteins or those with low ionisation efficiency, while RNA-seq might miss short-lived or

rare transcripts [149]. While bulk data does not achieve the resolution of single-cell data, bulk datasets typically exist at a far larger scale of sample sizes. This can be highly beneficial, as integration across many samples often results in more generalisable predictions, which is crucial to discovering robust biomarkers and patient subgroups.

Integrating multi-omics data at the cell level

Single-cell approaches benefit from analysing cellular heterogeneity, whereas traditional bulk methods assess data at the sample level and may require deconvolution for cellular insights [145]. Multi-omics integration methods model variations at both the cellular and feature levels, which is particularly important for chronic respiratory diseases like asthma and COPD [150]. In asthma, single-cell multi-omics identified distinct subtypes by analysing cellular interactions and inflammatory pathways, improving the understanding of disease heterogeneity and potential therapeutic targets [151]. In another study, COPD was primarily associated with molecular alterations in monocytes, which induced pro-inflammatory effects on alveolar epithelial cells. Additionally, aged monocytes and club cells contributed to COPD development, while smoking-related macrophage dysfunction disrupted the sphingolipid balance [152]. These analyses rely on joint embeddings produced by multi-omics data integration methods. Such methods, known as global integration strategies [153], are designed to detect larger-scale patterns of covariation across different modalities, enabling the unsupervised identification of global shifts in cellular states. Global data integration methods can be categorised into three primary approaches: neural network-based, matrix factorisation-based, and network-based. Neural network-based methods [154] use artificial neural networks to model complex relationships and patterns in data. These methods, which include architectures like variational autoencoders, deep canonical correlation analysis, and other deep learning models, are designed to learn low-dimensional representations of data through multiple layers of nonlinear transformations. DeepSpiro, a deep learning based model, was developed to predict future COPD risk using volume–flow spirometry data. By incorporating modules for smoothing, encoding key variability patterns, integrating heterogeneous data, and explaining predictions, DeepSpiro successfully predicted undiagnosed high-risk cases with long-term accuracy, demonstrating its effectiveness for early COPD detection and risk stratification [155]. Matrix factorisation based methods [156–158] decompose original matrices into two low-dimensional submatrices using linear approaches. One submatrix captures signature information (*e.g.* genes), while the other records cell information [159]. These methods help reduce the complexity of large datasets while retaining the essential features that describe the biological processes. Network-based methods [95] construct networks where nodes represent biological entities (*e.g.* genes, proteins, cells) and edges represent interactions or associations between these entities. These methods capture complex, nonlinear dependencies in biological systems. In COPD research, a study used a network-based approach to integrate transcriptomics, proteomics and PPI data, improving the accuracy of disease classification. Key genes like CXCL11 and TLR2 were identified, and pathways such as glycosaminoglycan signalling were enriched, providing insights into molecular mechanisms [160]. Feature-level variation can be understood by inferring gene-regulatory relationships, a process known as local integration [153]. This approach identifies potential interactions between specific features across diverse molecular layers. While many local integration methods rely on regression models [161], distinguishing true feature interactions from false associations is a key challenge that can occur due to confounding factors or upstream influences [153]. To address this, linear mixed models are widely used, incorporating fixed and random effects into the analysis [162, 163]. Furthermore, mapping eQTLs using single-cell expression profiles enables the identification of cell-type-specific eQTLs in rare cell populations [164]. Tools such as CellOracle [165] and SCENIC+ [166] are also available for inferring gene regulatory networks. The aforementioned single-cell multi-omics integration methods rely on jointly measured multi-omics profiles per cell. However, most available single-cell data across omics modalities have been measured unimodally. To leverage less expensive unimodal technologies for multimodal data integration, the different modalities must first be aligned. This can be achieved using manifold alignment methods [167–170] to infer a lower-dimensional structure within multiple complex datasets. However, this integration task is complicated by the incomplete understanding of the properties of latent biological manifolds [153]. In this case, matched multimodal assays can serve as gold-standard datasets to address these challenges and benchmark integration strategies [171].

Multi-omics data integration across resolutions

As single-cell datasets expand to profile hundreds of patients, studies can correlate single-cell gene expression with data obtained from bulk profiling [172]. Furthermore, integrating spatial and single-cell transcriptomics is crucial for understanding cell-to-cell communication [173]. In this context, we treat bulk RNA and single-cell RNA as separate modalities.

Integrating single-cell gene expression with bulk data modalities can help to identify the cell-type composition of disease-relevant tissues [174]. This can be achieved by deconvolution methods. These

methods often use regression-based techniques, such as non-negative least squares [175] and support vector regression [176], to map bulk gene expression profiles to specific cell types. This allows the identification of cellular targets for treatment and offers a better understanding of disease mechanisms [174]. Furthermore, to better understand interindividual variation, in the Human Lung Cell Atlas, demographic covariates (such as age, sex and body mass index) are modelled on single-cell gene expression with generalised linear mixed models [30]. While scRNA-seq identifies cell subpopulations within the tissue, it does not capture spatial distribution or local networks of intercellular communication [177]. Spatial techniques that localise RNA within tissue, including next-generation sequencing based methods, *in situ* sequencing (high-plex RNA imaging) and *in situ* hybridisation-based methods, do not yet offer single-cell resolution. Thus, integrating single-cell and spatial data is necessary to achieve maximal resolution in tissue. Two primary approaches for integrating scRNA-seq and spatial data are deconvolution and mapping. Like bulk deconvolution described earlier, spatial deconvolution identifies cell subpopulations by separating mRNA transcript mixtures at each capture spot. Methods [178–182] typically require an annotated scRNA-seq dataset with known cell-type markers for deconvolution [172]. Regression-based deconvolution methods merge scRNA-seq data with capture spot data to generate a matrix that profiles capture spots with overlaid scRNA-seq cell subtypes. This matrix is used to identify the scRNA-seq subtype profiles that most accurately explain the composition of each capture spot mixture [177]. Probabilistic distribution models are used to characterise each capture spot by assessing how well each cell type's model fits the transcript distribution of the spot. Recent methods, such as BayesTME [88] and SpatialGlue [183] perform deconvolution without reference to scRNA-seq data. For high-plex RNA imaging data, cluster-based mapping methods [47, 184, 185] integrate scRNA-seq and spatial data into a shared low-dimensional space. Meanwhile, other methods [186–188] create joint embeddings that can be used for downstream analysis, such as identifying spatially variable genes, defining spatial domains and analysing spatial heterogeneity. These methods further enhance the understanding of communication between adjacent cells and position-specific states [189].

Multi-omics to understand chronic respiratory diseases and viral infections

Multi-omics is being used to make significant advances in understanding chronic respiratory diseases by uncovering detailed molecular mechanisms through integrating data from single omics [1]. While multi-omics integration has identified potential biomarkers for chronic respiratory diseases, translating these findings into clinical interventions remains challenging. Variability in patient populations, the complexity of chronic respiratory diseases, and the need for rigorous clinical validation hinder the widespread application of multi-omic-based therapies. However, opportunities exist to refine patient stratification and develop more targeted treatments. Here, we review progress in COPD, asthma, IPF and viral infection including COVID-19, rhinovirus, RSV and influenza viruses (table 2).

COPD

Multi-omics approaches, including transcriptomics, metagenomics, metabolomics, lipidomics and proteomics, have been extensively applied to both pre-clinical COPD animal models and human COPD samples, providing important insights into disease mechanisms and defining therapeutic targets [28, 53, 195, 215–221]. Lipidomic and metabolomics have identified early-stage COPD biomarkers (*e.g.* sialic acid, hypoxanthine, xanthine, methylthioadenosine, adenine, glutathione) enabling personalised medicine strategies such as chronotherapy and improving treatment efficacy [222].

Epigenetic profiling of blood and lung tissues unveiled the effects of smoking in patients with mild-to-moderate COPD. In severe COPD, marked demethylation was identified, showing associations with increased expression of specific transcripts [223]. A combination of ATAC-seq and ChIP was used to elucidate the epigenetic regulation of monocyte extravasation to the lung in COPD by initially identifying the target of interest, PRMT7, applying ATAC-seq and then further validating using the PRMT7 promoter using ChIP-seq data [224].

Proteomics and metabolomics studies associated proteins (*e.g.* fibulin-1, tripeptidyl-peptidase 2) and metabolites (*e.g.* theophylline, hypoxanthine, cadherin-5) with advanced COPD diagnosis, highlighting diagnostic targets and therapeutic potential [195, 225].

Metagenomics and metabolomics have revolutionised our understanding of the lung and gut microbiome in COPD [226–230]. We discovered a disease-associated network in patients with COPD connecting *Streptococcus parasanguinis_B* with N-acetylglutamate and its analogue N-carbamoylglutamate [195]. Others found that hypoxanthine, theophylline, palmitoylethanolamide and CDH5 could identify patients with advanced COPD [231].

TABLE 2 Key examples of multi-omics applications in chronic respiratory diseases and viral infections

First author (year) [reference]	Key findings	Key omics approaches	Relevance
COPD			
Skerrett-Byrne (2021) [53]	Significant shifts in RNA biosynthesis involving HNRNPC and MSI2 identified, with inflammatory modulators like S100A1 validated in human COPD.	Proteomics, transcriptomics	Provides insights into conserved molecular changes in experimental models and human COPD
Schwartz (2023) [190]	Integrated DNA methylation and RNA-seq analysis uncovered novel transcriptomic and epigenetic signatures associated with COPD onset, including enriched binding sites for transcription factors TCF21 and FOSL2/FRA2. Analysed samples from 3 non-COPD individuals, 3 with mild COPD (GOLD stage I), and 5 with moderate to severe COPD (GOLD stage II–IV). Spearman correlation was performed to define associations between gene expression and differentially methylated regions.	Epigenomics, transcriptomics	Highlights the role of epigenetic regulation in COPD onset and progression
Chotirmall (2022) [191]	Azithromycin enhances anti-inflammatory bacterial metabolite production, linking respiratory microbiota to COPD.	Metagenomics, proteomics, metabolomics	Links microbiome changes to inflammation and COPD progression, suggesting therapeutic targets
Lynch (2015) [192]	Multi-omics identified emphysema- and airway-predominant COPD subtypes.	CT imaging, transcriptomics, proteomics	Helps refine COPD classification and potential subtypes for targeted therapies
Li (2025) [193]	Identified 248 early COPD-associated proteins, many involved in inflammation. Found 137 early COPD-associated metabolites indicating early metabolic disturbances.	Proteomics, metabolomics	Highlights early molecular changes in COPD Supports use of proteomic and metabolomic biomarkers for early diagnosis and intervention in chronic respiratory diseases
Yan (2022) [194]	Microbial dysbiosis in COPD with decreased *Prevotella* and tryptophan-derived metabolite IAA. Reduced IAA impairs IL-22 signalling. IAA restoration *via Lactobacillus* reduces inflammation in mice.	Metagenomics, metabolomics, transcriptomics, proteomics	Demonstrates how host–microbe interactions drive COPD pathology Identified IAA and microbial modulation as potential therapeutic avenues
Bowerman (2020) [195]	Identified bacterial species, including *Streptococcus* spp. and members of Lachnospiraceae, correlated with reduced lung function. Untargeted metabolomics revealed a COPD-associated signature comprising lipids, xenobiotics and amino acid related metabolites.	Metagenomics, metabolomics	Highlights the association between gut microbiome/metabolome alterations and COPD, suggesting potential biomarkers and therapeutic targets *via* the gut–lung axis
Budden (2024) [196]	Used proteomics, metabolomics and microbiome analysis to show that faecal transfers and complex carbohydrate supplementation improved COPD by modulating glucose and starch metabolism and altering bacterial abundance (Muribaculaceae, Desulfovibrionaceae, Lachnospiraceae).	Metagenomics, proteomics, metabolomics	Demonstrates the contribution of the gut microbiome to COPD pathogenesis and its potential as a therapeutic target through dietary interventions
Asthma			
Guang (2023) [127]	Integration of clinical, plasma and metagenomic/metabolomic data reveals clinical heterogeneity in allergic asthmatic children.	Metagenomics, metabolomics	Identifies biomarkers for stratified treatment approaches
Abdel-Aziz (2022) [197]	Elevated eicosanoids (11-dehydrothromboxane-B2, prostaglandin-E2) linked to microbiome–inflammation axis, correlating with PTGS2 gene expression.	Metabolomics, transcriptomics, proteomics	Suggests microbiome–inflammation interactions as therapeutic targets
Sunata (2024) [198]	IL-4 stimulation induces IL1RL1-high eosinophils with upregulated cysteinyl leukotriene metabolism (*e.g.* increased LTC4S, GGT5 expression), leading to elevated leukotriene D4 production *via* the STAT6 pathway.	Transcriptomics, proteomics, lipidomics	Identifies a unique eosinophil phenotype linked to type 2 inflammation in asthma and chronic rhinosinusitis, providing novel insights into IL-4-mediated inflammatory pathways and potential therapeutic targets
IPF			
Allen (2017) [199]	Epigenetic drivers in lung fibroblasts identified, including twist transcription factor-1 and forkhead box protein-A1 expression.	Transcriptomics, epigenomics	Sheds light on pathogenic roles of epigenetic modifications in IPF progression

Continued

https://doi.org/10.1183/16000617.0286-2024

TABLE 2 Continued

First author (year) [reference]	Key findings	Key omics approaches	Relevance
Wen (2025) [200]	A multi-omics approach identified 6 genes (ANO9, BRCA1, CCDC200, EZH1, FAM13A, SFR1) as potential therapeutic targets in IPF.	Genomics, transcriptomics	Key therapeutic targets identified for further validation and potential therapy development
Zhang (2024) [201]	Role of Mo-AMs in IPF highlighted through transcriptional and epigenetic profiling.	scRNA-seq, ATAC-seq	Identifies macrophages as targets for fibrotic treatments
COVID-19			
Ng (2021) [202]	Plasma proteomics and metabolomics identified dysregulated complement pathways, suggesting therapeutic potential for complement-targeting interventions.	Proteomics, metabolomics	Reveals complement system dysregulation as a therapeutic target
Ma (2022) [203]	Integration of multi-omics with next-generation sequencing identified Sigma1/Sigma2 modulators and an Nsp16 mutation as potential therapeutic candidates.	Multi-omics, next-generation sequencing	Unveils new antiviral drug targets and potential vaccine candidates
Su (2020) [204]	Blood-based multi-omics revealed immune dynamics, with plasma lipids and cytokines linked to disease severity.	Proteomics, metabolomics	Identifies biomarkers for COVID-19 severity and early intervention
Barh (2020) [205]	Integrated omics identified potential drug repurposing candidates (betamethasone, statins, vitamin D, zinc) for COVID-19.	Transcriptomics, proteomics, host genetics	Provides a framework for repurposing existing drugs and improving treatment strategies
Rhinoviruses			
Djeddi (2024) [206]	Rhinovirus infections exacerbate asthma by triggering immune responses that lead to airway inflammation and hyperresponsiveness.	Transcriptomics, GWAS	Highlights the role of nonciliated epithelial cells in rhinovirus-induced asthma exacerbations
Raita (2021) [207]	Identified endotypes in infants with rhinovirus bronchiolitis, characterised by specific viral profiles, microbiome compositions, and immune responses. Certain endotypes were associated with an increased risk of developing childhood asthma.	Metagenomics, transcriptomics, metabolomics	Highlights the heterogeneity in bronchiolitis and the potential for early identification of infants at high risk for asthma development through integrated omics approaches
RSV			
Huang (2023) [208]	Elevated serum collagen levels suggest airway remodelling, disruptions in lipid metabolism and complement system activation.	Proteomics, metabolomics	Highlights potential therapeutic strategies for managing RSV-induced airway remodelling
Raita (2021) [209]	Identified 4 endotypes of RSV bronchiolitis based on immune responses, microbial and metabolic profiles.	Genomics, transcriptomics, microbiome, metabolomics	Identifies high-risk infants for asthma development, enabling early identification for targeted interventions
Influenza viruses			
Zhu (2023) [210]	Built *H2Flu*, a curated database of human host genes/proteins associated with 14 IAV subtypes plus influenza B/C, a resource for influenza research.	Genomics, transcriptomics, proteomics data	Enables efficient identification and prioritisation of influenza host factors
Riese (2022) [211]	Identified molecular and immune pathway signatures that differentiate responders *versus* nonresponders to seasonal flu vaccines.	Transcriptomics, proteomics, immune profiling	Guides personalised vaccine strategies and better formulation design
Martin-Sancho (2021) [212]	Discovered that IAV M2 protein interacts with TBC1D5 to avoid lysosomal degradation revealing viral autophagy evasion.	Genetic CRISPR screens, transcriptomics, proteomic	Elucidates mechanisms of viral persistence and potential antiviral targets
Platt (2022) [213]	Showed IAV binding to *Streptococcus pneumoniae* alters bacterial physiology and virulence, offering insight into viral–bacterial synergy in pneumonia.	Transcriptomics, proteomics, phenotypic assays	Co-infection with IAV and bacterial pathogens is a major driver of exacerbations in chronic respiratory diseases such as COPD and asthma; this study enhances understanding of mechanisms behind infection-triggered disease worsening
Hu (2022) [214]	Found specific shifts in oropharyngeal microbiota and metabolites linked to disease severity in children with IAV pneumonia.	Microbiome profiling (16S/ shotgun), metabolomics	Potential for noninvasive biomarkers and therapeutic targeting

IPF: idiopathic pulmonary fibrosis; COVID-19: coronavirus disease 2019; RSV: respiratory syncytial virus; GOLD: Global Initiative for Chronic Obstructive Lung Disease; CT: computed tomography; IAA: indole-3-acetic acid; IL: interleukin; IL1RL: IL-1 receptor-like; Mo-AM: monocyte-derived alveolar macrophage; scRNA-seq: single-cell RNA sequencing; ATAC-seq: assay for transposase-accessible chromatin sequencing; GWAS: genome-wide association study; IAV: influenza A virus.

Airway colonisation by *Staphylococcus aureus* was linked to worsened lung function by activating homocysteine through the AKT1-S100A8/A9 axis. In a mouse model, bacteriophage elimination of *S. aureus* restored lung function, suggesting microbiome-modifying therapies as novel strategies to suppress COPD progression [232]. Dysbiosis of the lung microbiome has also been associated with disease progression and exacerbations, with specific microbial signatures associated with neutrophilic or eosinophilic endotypes [75, 187]. Cigarette smoke and COPD were shown to cause pathological changes in the gut that also impacted the lung. The authors then defined the gut microbiome and metabolome differences in human and experimental COPD [195, 196]. Dysbiosis of the lung microbiome has also been linked to disease progression and exacerbations, with specific microbial signatures associated with neutrophilic or eosinophilic endotypes [195, 233]. Faecal microbiome transfers alleviated experimental COPD and associated gut pathology. Cigarette smoke-associated microbiota induced lung inflammation and were associated with disease, and proteomics and metabolomics showed that they had reduced glucose and starch metabolism. Dietary complex carbohydrates improved COPD in mice and patients. Lactobacilli depletion and reduced indole-3-acetic acid levels disrupted interleukin-22 pathways, leading to epithelial cell apoptosis [194]. Furthermore, *S. parasanguinis* B was linked to distinct metabolomic profiles in COPD patients, indicating the potential of targeting both lung and gut microbiomes for therapy [53, 195].

Multi-omics studies have also shed light on adaptive immune dysregulation in COPD. Integrated analyses of large datasets identified immune cell deficiencies, particularly $CD4^+$ T-cells, as predictors of acute exacerbations, suggesting that immune monitoring could improve management [234]. Spatial transcriptomics and computed tomography imaging revealed increased B-cell-related gene expression in severe emphysema, with abnormal B-cell activity associated with emphysema severity [235]. Patients with Global Initiative for Chronic Obstructive Lung Disease 1–2 COPD showed distinct B-cell activation signatures, linking B-cell maturation and antibody production to emphysema development. In addition, blood-based multi-omics identified biomarkers and predictive models combining clinical and protein data, enabling early diagnosis and potential therapeutic targeting of emphysema in COPD [235, 236].

These findings show the value of multi-omics in understanding molecular mechanisms, identifying biomarkers and guiding therapeutic development in COPD. By integrating diverse omics layers, these approaches offer a holistic understanding of COPD pathogenesis, progression, and potential interventions.

Asthma

Asthma has been extensively explored using multi-omics approaches, including transcriptomics, proteomics, genomics and metabolomics, providing insights into its complexity and clinical heterogeneity [215, 282]. Studies like U-BIOPRED and the Severe Asthma Research Program have advanced our understanding of genetic predisposition, disease heterogeneity and treatment responses [120]. A large-scale meta-analysis of global asthma genome-wide association study (GWAS) identified five novel genetic loci and two new associations at known loci, enhancing our understanding of asthma susceptibility and paving the way for personalised therapies [121]. Epigenomic investigations revealed cell lineage specific differences and asthma-associated DNA methylation changes, such as in lymphotoxin-α and tumour necrosis factor genes, suggesting that epigenetic mechanisms may be potential therapeutic targets [122]. A recent study combined eicosanoids, transcriptomics and proteomics to show higher levels of 11-dehydrothromboxane-B2 and prostaglandin-E2 in microbiome-driven asthma, correlating with upregulated PTGS2 gene expression in sputum. This suggests that a microbiome–inflammation axis could guide targeted anti-inflammatory interventions [197]. These findings collectively provide actionable insights into asthma mechanisms, inform precision medicine strategies and identify novel therapeutic opportunities.

IPF

Multi-omics approaches have transformed our understanding of the pathogenesis and therapeutic opportunities in IPF by uncovering novel mechanisms and cellular processes. Genomic studies identified rare and common genetic variants, shedding light on disrupted biological pathways in inherited forms of pulmonary fibrosis, such as telomere biology, surfactant metabolism and immune function, which suggest genetic predispositions that could guide personalised treatments [237, 238].

scRNA-seq has revealed the cellular heterogeneity in IPF lungs, uncovering previously unrecognised fibrotic and immune cell populations, including a $KRT5^-/KRT17^+$ pathogenic epithelial population that contributes to extracellular matrix production, highlighting its potential as a therapeutic target [35, 239]. Novel fibrosis-specific immune, stromal, endothelial and epithelial cell states have been identified, which persist in human fibrosis and transiently appear during lung regeneration in mice, suggesting their crucial and potentially targetable role in disease progression [34, 240–242].

An integrative IPF cell atlas compiling published scRNA-seq datasets has improved our understanding of molecular and cellular relationships in IPF [243]. Combining transcriptomics, epigenomics and proteomics characterised differentially methylated regions, lncRNAs, and their regulation of key mediators such as matrix metalloproteinase-7, revealing epigenetic drivers of disease progression [244].

Integrative studies have identified potential tissue biomarkers, such as butyrophilin-like-9 and plasmolipin, that distinguish clinically relevant molecular subtypes [245]. Blood-based multi-omics from the IPF-PRO Registry revealed two molecular subtypes with distinct prognostic implications; subtype-1 was associated with more severe disease and shorter progression time, providing a framework for risk stratification and personalised care [246]. Multi-modal integration of scRNA-seq and proteomics from bronchoalveolar lavage fluid and plasma identified peripheral protein biomarker signatures reflecting lung cell state changes, which predicted disease status and functional decline [247].

Lastly, combining whole-exome sequencing, bulk RNA-seq and scRNA-seq provides unique insights into genomic and transcriptomic interactions, offering a holistic understanding of IPF [248]. Multi-omics approaches have also facilitated the discovery of numerous therapeutic targets, which have been validated in cells and tissues from IPF patients, supporting the development of targeted therapies. A recent study employed a comprehensive multi-omics approach, integrating genomic and transcriptomic data, to uncover six genes ANO9, BRCA1, CCDC200, EZH1, FAM13A and SFR1 as potential therapeutic targets for IPF [200, 201]. Moreover, a recent study identified TRAF2- and NCK-interacting kinase as a target for fibrosis treatment and used artificial intelligence (AI)-driven drug discovery to identify INS018_055, a small-molecule inhibitor, as a potential therapeutic. INS018_055 was then used to demonstrate significant antifibrotic and anti-inflammatory effects in pre-clinical models and showed promising safety and pharmacokinetic profiles in phase I clinical trials, positioning it as a potential therapy for fibrotic diseases [249].

These findings collectively advance the field by linking molecular discoveries to clinical outcomes, enabling biomarker development, molecular subtyping and the identification of actionable therapeutic targets.

COVID-19

With the COVID-19 endemic and the emergence of long COVID, the development of effective therapies is urgently needed. Omics studies have provided critical insights into how SARS-CoV-2 hijacks host cellular machinery to induce pathogenesis, uncovering pathways and potential therapeutic targets [28, 250]. scRNA-seq revealed that differentiated COPD bronchoepithelial cells exhibit higher susceptibility to infection due to increased viral co-receptor expression and impaired antiviral responses, and exacerbating COPD by disrupting protease balance. This identified co-receptor inhibitors as potential therapies for patients with both COPD and COVID [28].

While single-omics studies have provided valuable insights, they often fail to capture the complexity of COVID-19. The evolving dynamics of the disease, shaped by widespread immunity and the virus's endemicity, demand a multi-omics approach for deeper understanding [251]. Multi-omics analyses of lower airway microbiomes from critically ill COVID-19 patients revealed that high viral loads and impaired adaptive immunity were linked to poor outcomes [203]. They also identified microbial and transcriptomic signatures that could serve as prognostic markers or therapeutic targets. Metagenomics and metatranscriptomics of bronchoscopy samples showed that elevated levels of *Mycoplasma salivarium* and distinct airway transcriptomic profiles were associated with worse clinical outcomes, emphasising the critical role of microbial–immune interactions in severe disease [252].

Integration of multi-omics with next-generation sequencing and reverse genetics has advanced the identification of novel therapeutic targets. Studies revealed Sigma1/Sigma2 modulators as potential antiviral drug targets and identified an Nsp16 point mutation that attenuates SARS-CoV-2 while eliciting protective immunity, suggesting its potential as a vaccine candidate [203, 253].

Blood-based multi-omics investigations have elucidated immune dynamics during disease progression, linking decreases in plasma lipids and amino acids, elevated pro-inflammatory cytokines, and exhausted $CD4^+$ T-cells to severe disease [204]. Circulating cytokines, lipids and metabolites have been identified as biomarkers for COVID-19 severity and hospitalisation risk, enabling better risk stratification and early intervention [254].

By linking these findings to actionable outcomes, such as identifying biomarkers, guiding therapeutic interventions and advancing vaccine development, multi-omics approaches provide a holistic framework to improve the management and treatment of COVID-19.

Rhinoviruses

Rhinoviruses, members of the Picornaviridae family, are small, non-enveloped viruses with single-stranded RNA genomes and are the leading cause of the common cold. These highly contagious viruses primarily infect the upper respiratory tract and are responsible for ∼90% of acute respiratory tract infections, ranging from mild colds to severe lower respiratory illnesses such as pneumonia [255]. Rhinoviruses are common causes of exacerbating conditions like COPD [256] and asthma [257]. They do so by triggering immune responses that lead to airway inflammation and hyperresponsiveness. Airway epithelial cells are central to initiating these responses, and recent studies suggest that modulating epithelial pathways may offer therapeutic potential in managing rhinovirus-induced asthma exacerbations [257]. The application of multi-omics approaches, integrating genomics, transcriptomics, proteomics and metabolomics has significantly advanced our understanding of rhinovirus biology and its interactions with host systems. Notably, a recent study integrating GWAS with transcriptomic profiling identified that rhinovirus infections interact with asthma-associated genetic risk factors predominantly through nonciliated airway epithelial cells. These findings highlight a critical mechanism in childhood-onset asthma and point to nonciliated epithelial cells as potential targets for mitigating virus-induced exacerbations [206].

RSV

RSV is highly contagious and primarily affects the respiratory tract. It is a leading cause of lower respiratory infections, including bronchiolitis and pneumonia, especially in infants, the elderly and immunocompromised individuals. RSV is also linked to an increased risk of wheezing and the development of asthma later in life. Due to its significant global disease burden, particularly in children aged <5 years, RSV is a critical target for vaccine and therapeutic development [258]. This highlights the importance of multi-omics (genomics, transcriptomics, proteomics, metabolomics) in advancing predictive, preventive and personalised medicine for RSV. By providing a comprehensive view of host–pathogen interactions and individual variability, multi-omics can enhance early risk assessment, enable targeted interventions, and inform patient-specific therapies, ultimately improving outcomes and reducing healthcare burdens [259]. A study employing integrated proteomic and metabolomic analyses explored the host responses in children with RSV pneumonia. By analysing serum samples from affected children, healthy controls and those with other infections, the researchers identified several differentially expressed metabolites, with six showing promise as diagnostic biomarkers. Notably, neuromedin N and histidylproline diketopiperazine were validated in an independent cohort. The study also revealed elevated serum collagen levels, suggesting potential airway remodelling, and highlighted disruptions in lipid metabolism and complement system activation. These findings provide valuable insights into RSV pathogenesis and could inform future diagnostic and therapeutic strategies [208]. In another study, a comprehensive multi-omics approach including clinical data, viral profiling, airway microbiome analysis, transcriptomics, and metabolomics was used to identify biologically distinct endotypes in infants hospitalised with RSV bronchiolitis. The authors identified four endotypes, each defined by unique microbial compositions, immune responses and metabolic profiles. One endotype, marked by high interferon responses, IgE sensitisation and co-infection with rhinovirus, exhibited a significantly increased risk of developing asthma by the age of 5 years. These findings emphasise the role of multi-omics (genomics, transcriptomics, proteomics, metabolomics) in supporting predictive, preventive and personalised (3P) medicine in RSV infection management. By capturing a comprehensive picture of host–pathogen interactions and individual variability, multi-omics tools can guide early risk assessment, targeted interventions and patient-tailored therapies, ultimately improving outcomes and reducing healthcare burdens. In addition, they highlight the potential of integrated omics in early identification of infants at high risk for asthma and suggest that targeted preventive strategies could be developed based on these insights [209].

Influenza viruses

Influenza viruses, particularly influenza A and influenza B, are the primary causative agents of seasonal flu in humans. While influenza A viruses (IAV) are responsible for both seasonal epidemics and global pandemics due to their high mutation rate and ability to infect multiple species, influenza B viruses are largely restricted to humans and tend to cause milder, seasonal outbreaks [260, 261]. IAVs are highly contagious enveloped RNA viruses that pose a significant public health challenge. A recent study used a comprehensive multi-omics approach integrating genome-wide CRISPR screens, transcriptomics and proteomics to uncover host factors that restrict IAV replication. This identified TBC1D5, a key regulator of autophagy, as a potent antiviral factor that promotes lysosomal degradation of viral components. Notably, the viral M2 protein was found to bind TBC1D5, disrupting its interaction with Rab7 and thereby evading autophagic clearance, facilitating viral survival and replication. These findings reveal a novel immune evasion mechanism and suggest TBC1D5 as a target for host-directed antiviral therapy [212]. Another multi-omics study examined the basis of reduced vaccine responsiveness in the elderly population by integrating transcriptomics, proteomics, cytokine profiling and immune phenotyping. It identified distinct baseline immune and molecular signatures that predicted vaccine responsiveness. Individuals who

responded well exhibited upregulation of genes involved in B-cell activation, interferon signalling and cellular metabolism, whereas nonresponders had elevated expression of inflammatory and stress pathways. These insights highlight the heterogeneity of immune responses in ageing populations and offer potential biomarkers and therapeutic targets for enhancing influenza vaccine efficacy [211].

Future directions for the application of multi-omics to chronic respiratory diseases and viral diseases

There are several promising areas that should be explored to improve patient outcomes. First, integrating multi-omics data from chronic respiratory disease patients could help capture dynamic disease progression over time, allowing for more precise identification of biomarkers and therapeutic targets. This approach could reveal subtle molecular changes that single-omics studies may miss, aiding earlier diagnosis and more personalised treatment. For viral diseases, multi-omics can offer a deeper understanding of viral–host interactions, immune responses and long-term consequences [262]. Moreover, advancing longitudinal multi-omics studies can provide a more comprehensive picture of disease progression over time, especially for conditions like long COVID, and other post-viral syndromes, by integrating genomic, transcriptomic and microbiome data. These efforts are crucial to identify biomarkers that predict disease outcomes, enabling the development of targeted therapeutic strategies [263]. By combining genomic, transcriptomic, proteomic and microbiome data, researchers can uncover mechanisms of immune evasion, viral replication and host susceptibility. This would help to develop more effective treatments, but also identify patients at higher risk of severe outcomes. Combining multi-omics with clinical and environmental data can further refine precision medicine approaches, facilitating better diagnosis and more effective treatments. As computational tools and validation methods improve, these multi-omics approaches will play an increasingly vital role in developing personalised therapies for both chronic respiratory diseases and viral diseases [30, 264, 265].

Summary and challenges

Human studies face key limitations: they often rely on samples taken at a single time point, use lung tissue from end-stage disease (making it difficult to study how the disease develops and progresses), and analyse peripheral fluids (like blood or sputum) that do not fully capture what is occurring in the lungs [266]. These limitations are alleviated by the use of animal models that replicate key features of human disease, enabling longitudinal assessment of the development, progression, exacerbations and treatment responses using comparable exposure methods and systemic outcomes [5, 45, 53, 267]. Findings from animal models should be validated and translated in complementary human *ex vivo* lung models, including organoids and precision-cut lung slices [268–273]. They should be combined with spatial omics technologies to enable longitudinal assessment of disease-related changes. In our recent work, we used scRNA-seq and spatial proteomics to identify five neutrophil subtypes in COPD, of which only one was disease-associated [218]. Using a validated mouse model, we traced the emergence of these pathogenic neutrophils from the bone marrow to the bloodstream and into the lung, revealing a systemically driven disease mechanism with therapeutic potential. These insights cannot be obtained through *ex vivo* or *in vitro* models [274].

Challenges remain in handling heterogeneous clinical metadata, variable sample and file sizes, missing data, batch effects and insufficient sequencing depth. Effective integration of different omics layers requires detailed characterisation of patient populations to address variability in disease phenotypes. Variability across independent cohorts due to differences in patient demographics, disease severity and data processing protocols further complicates data harmonisation. In addition to these technical and biological sources of variability, the extent and quality of sample and metadata collection play a crucial role in the utility of multi-omics data. Many studies lack standardised clinical annotations, longitudinal sampling, or representation across diverse demographic groups. Enhanced collection of detailed metadata including comorbidities, treatment history, environmental exposures and disease progression along with the inclusion of a broader range of biological samples (*e.g.* bronchoalveolar lavage fluid, biopsies, nasal swabs and time-series samples) will greatly enhance translational potential. Establishing standardised protocols and harmonised metadata frameworks across studies is essential to enable meaningful integration and interpretation of datasets in clinical research. However, this variability can also contribute to improved generalisability by reflecting a wider range of disease manifestations, thus enhancing the relevance of findings across diverse patient populations. Disparities in omics layers and sample sizes can limit statistical power and generalisability, while high resource demands for data integration and potential cohort selection bias may restrict the applicability of findings to the broader chronic respiratory disease population [86, 275]. Much of this will be resolved by continued enhanced high-quality data collection.

Data cleaning, filtration, quality control and benchmarking before integrating is also essential. Improved analysis will be facilitated by acquiring more comprehensive and accurately curated patient metadata. Omics data also originates from various sources; therefore, benchmarking and establishment of gold standard methods of data collection, terminology and formatting are critical. To address these issues,

universal platforms are being developed to enable access and integration of data from different groups (*e.g.* Human Lung Cell Atlas) [30, 264, 265], which reduces the need to run new experiments. Moreover, long-read sequencing may be a game-changer in genetic analysis, reducing costs while capturing a more complete genomic picture enhancing our understanding of genetic heterogeneity [276].

Various AI and statistical methods have been developed to address the challenges of integrating multi-omics datasets. These approaches require high-performance computing, advanced software, and rapidly evolving machine learning and AI technologies [277]. Different programming languages, such as Python, R and Perl, are used to create new integration tools [277, 278]. However, no gold-standard workflows have yet emerged, making it difficult to achieve consistent and reproducible analyses [279]. Data protection and legal restrictions further complicate cross-jurisdiction studies [280]. Swarm learning has emerged as a solution, integrating data while maintaining confidentiality through cutting-edge computing and blockchain-based peer-to-peer networking for onsite analysis [281].

In summary, multi-omics provide new insights into the molecular mechanisms of chronic respiratory diseases, improving our understanding of disease heterogeneity and revealing potential therapeutic targets. Techniques like transcriptomics, proteomics, epigenomics, metagenomics and metabolomics produce detailed maps of transcriptional, protein, microbiome and metabolic changes. Longitudinal studies, combined with multi-omics integration are essential to identify biomarkers and therapeutic targets. Despite these advances, translating molecular insights into effective treatments remains challenging due to the need for further validation, the high costs of multi-omic analyses, and the lack of robust computational tools. These obstacles, along with the complexity of chronic respiratory diseases, hinder the optimal use of multi-omics data in clinical settings. However, with improvements in computational tools and continued validation, multi-omics holds the potential for more personalised and targeted treatments in the future [263].

Advancing multi-omics: what is next?

Opportunities in multi-omics research

- There is a substantial opportunity to improve the reliability and effectiveness of integrative multi-omics methods. By refining computational tools and analytical frameworks, researchers can enhance the robustness and clinical relevance of multi-omics analyses.
- Addressing batch effects and data heterogeneity remains a key challenge, but also an opportunity. Emerging computational solutions have promise, and continued development of advanced integration techniques will significantly improve data harmonisation and interpretation across studies.
- Reducing the costs and logistical barriers of multi-omics research will enable larger, more inclusive studies. This will increase statistical power, broaden the diversity of study populations, and enhance the generalisability and clinical impact of findings.
- Improving the collection of comprehensive metadata including comorbidities, treatment history, environmental exposures and disease progression alongside more diverse biological samples, represents a major opportunity to boost the translational potential of multi-omics data and bridge the gap between bench and bedside.

Strategic recommendations for advancing multi-omics research

- Promote open data sharing: depositing multi-omics data into globally accessible platforms is essential for enabling large-scale, cross-disciplinary analyses. This practice fosters international collaboration and accelerates the discovery of deeper, more comprehensive insights into complex diseases including chronic respiratory diseases.
- Establish standardised protocols: standardising data collection protocols and harmonising metadata frameworks across studies is crucial to improve dataset integration. It will reduce variability, enhance reproducibility and ensure consistent interpretation of results in clinical and translational research.
- Foster multidisciplinary collaboration for endotype discovery: understanding and classifying disease endotypes is vital for developing accurate diagnostics and targeted therapies. Achieving this requires close collaboration between scientific, clinical, and industry experts. The Horizon Europe MSCA Doctoral Network “RESPIRE-EXCEL” aims to train future leaders in translational medicine. It focuses on identifying disease mechanisms, developing predictive biomarkers, detecting individual biomarker activity, designing endotype-specific treatments, and advancing precision medicine strategies for sustained disease remission.

Acknowledgements: We acknowledge COPD-iNET (www.copd-inet.com) for fostering scientific exchange and collaboration in this review.

Provenance: Submitted article, peer reviewed.

Author contributions: S. Idrees, P.M. Hansbro and A. Faiz designed the work. S. Idrees, H. Chen, T. Sadaf, M.D. Johansen, K.R. Paudel, S.F. Rehman, M. O'Rourke, Y. Wang, M.D. Luecken and P.M. Hansbro wrote the first draft. S. Idrees, H. Chen, T. Sadaf, M.D. Johansen, K.R. Paudel, S.F. Rehman, E. Hortle, A.S. Philp, G. Liu, K.F. Budden, G.E. Kaiko, S.E.M. Lucas, J.L. Dickinson, P.C. Allen, J.E. Powell, L-Y. Zhang, D.C. Chambers, T. Corte, G. Caramori, M. Sauler, P.A. Wark, J. Gote-Schniering, M.D. Luecken, T.M. Conlon, T.S. Kapellos, A.Ö. Yildirim, M. Lehman, R. Faner, S.C. Dharmage, C.E. Wheelock, M. van den Berge, M.C. Nawijn, F. Polverino, G.T. Belz, S.H. Chotirmall, L.N. Segal, M. O'Rourke, A. Faiz and P.M. Hansbro edited and/or wrote the manuscript. A. Faiz and P.M. Hansbro supervised the study. All authors reviewed or revised the manuscript and approved the final draft for submission.

Conflict of interest: M.D. Luecken contracted for the Chan Zuckerberg Initiative, received speaker fees from Pfizer and Janssen Pharmaceuticals, and consults for CatalYm GmbH. The other authors declare no competing interests according to the ERS conflicts of interest policy.

Support statement: G. Liu was supported by the Al and Val Rosenstrauss Fellowship from the Rebecca L. Cooper Medical Research Foundation. P.M. Hansbro is funded by a Fellowship and grants from the National Health and Medical Research Council (NHMRC) of Australia (1175134, 2010287, 2008937, 2015613), Australian Research Council (200101058, 230101156) and by UTS. Funding information for this article has been deposited with the Open Funder Registry.

References

1 Kan M, Shumyatcher M, Himes BE. Using omics approaches to understand pulmonary diseases. *Respir Res* 2017; 18: 149.
2 Altelaar AF, Munoz J, Heck AJ. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nat Rev Genet* 2013; 14: 35–48.
3 Sheppard D. Aspen Lung Conference 2010: systems biology of lung diseases – progress in the omics era. *Proc Am Thorac Soc* 2011; 8: 199–202.
4 Sidhaye VK, Nishida K, Martinez FJ. Precision medicine in COPD: where are we and where do we need to go? *Eur Respir Rev* 2018; 27: 180022.
5 Kim RY, Horvat JC, Pinkerton JW, *et al.* MicroRNA-21 drives severe, steroid-insensitive experimental asthma by amplifying phosphoinositide 3-kinase-mediated suppression of histone deacetylase 2. *J Allergy Clin Immunol* 2017; 139: 519–532.
6 Collins DC, Sundar R, Lim JSJ, *et al.* Towards precision medicine in the clinic: from biomarker discovery to novel therapeutics. *Trends Pharmacol Sci* 2017; 38: 25–40.
7 Nimmesgern E, Benediktsson I, Norstedt I. Personalized medicine in Europe. *Clin Transl Sci* 2017; 10: 61–63.
8 Agusti A, Faner R, Celli B, *et al.* Precision medicine in COPD exacerbations. *Lancet Respir Med* 2018; 6: 657–659.
9 Kaminsky DA, Irvin CG, Sterk PJ. Complex systems in pulmonary medicine: a systems biology approach to lung disease. *J Appl Physiol* 2011; 110: 1716–1722.
10 Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell* 2008; 134: 714–717.
11 Shang L, Zhao W, Wang YZ, *et al.* meQTL mapping in the GENOA study reveals genetic determinants of DNA methylation in African Americans. *Nat Commun* 2023; 14: 2711.
12 Eriksson Ström J, Kebede Merid S, Pourazar J, *et al.* Chronic obstructive pulmonary disease is associated with epigenome-wide differential methylation in BAL lung cells. *Am J Respir Cell Mol Biol* 2022; 66: 638–647.
13 Mao Y, Huang P, Wang Y, *et al.* Genome-wide methylation and expression analyses reveal the epigenetic landscape of immune-related diseases for tobacco smoking. *Clin Epigenetics* 2021; 13: 215.
14 Nkongolo K, Michael P. Reduced representation bisulfite sequencing (RRBS) analysis reveals variation in distribution and levels of DNA methylation in white birch (*Betula papyrifera*) exposed to nickel. *Genome* 2024; 67: 351–367.
15 Giresi PG, Kim J, McDaniell RM, *et al.* FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* 2007; 17: 877–885.
16 Zhang Z, Pugh BF. High-resolution genome-wide mapping of the primary structure of chromatin. *Cell* 2011; 144: 175–186.
17 Vandereyken K, Sifrim A, Thienpont B, *et al.* Methods and applications for single-cell and spatial multi-omics. *Nat Rev Genet* 2023; 24: 494–515.

18 Benincasa G, DeMeo DL, Glass K, *et al.* Epigenetics and pulmonary diseases in the horizon of precision medicine: a review. *Eur Respir J* 2021; 57: 2003406.
19 Duggan DJ, Bittner M, Chen Y, *et al.* Expression profiling using cDNA microarrays. *Nat Genet* 1999; 21: Suppl. 1, 10–14.
20 Konishi H, Ichikawa D, Arita T, *et al.* Microarray technology and its applications for detecting plasma microRNA biomarkers in digestive tract cancers. *Methods Mol Biol* 2016; 1368: 99–109.
21 Allison DB, Cui X, Page GP, *et al.* Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 2006; 7: 55–65.
22 Spira A, Beane J, Shah V, *et al.* Effects of cigarette smoke on the human airway epithelial cell transcriptome. *Proc Natl Acad Sci USA* 2004; 101: 10143–10148.
23 Zhang X, Sebastiani P, Liu G, *et al.* Similarities and differences between smoking-related gene expression in nasal and bronchial epithelium. *Physiol Genomics* 2010; 41: 1–8.
24 Spira A, Beane JE, Shah V, *et al.* Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 2007; 13: 361–366.
25 Ditz B, Boekhoudt JG, Aliee H, *et al.* Comparison of genome-wide gene expression profiling by RNA sequencing *versus* microarray in bronchial biopsies of COPD patients before and after inhaled corticosteroid treatment: does it provide new insights? *ERJ Open Res* 2021; 7: 00104-2021.
26 Zhou C, Chen H, Han L, *et al.* Screening of genes related to lung cancer caused by smoking with RNA-seq. *Eur Rev Med Pharmacol Sci* 2014; 18: 117–125.
27 Yick CY, Zwinderman AH, Kunst PW, *et al.* Transcriptome sequencing (RNA-seq) of human endobronchial biopsies: asthma *versus* controls. *Eur Respir J* 2013; 42: 662–670.
28 Johansen MD, Mahbub RM, Idrees S, *et al.* Increased SARS-CoV-2 infection, protease, and inflammatory responses in chronic obstructive pulmonary disease primary bronchial epithelial cells defined with single-cell RNA sequencing. *Am J Respir Crit Care Med* 2022; 206: 712–729.
29 Stoeckius M, Hafemeister C, Stephenson W, *et al.* Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* 2017; 14: 865–868.
30 Sikkema L, Ramírez-Suástegui C, Strobl DC, *et al.* An integrated cell atlas of the lung in health and disease. *Nat Med* 2023; 29: 1563–1577.
31 Jackson ND, Everman JL, Chioccioli M, *et al.* Single-cell and population transcriptomics reveal pan-epithelial remodeling in type 2-high asthma. *Cell Rep* 2020; 32: 107872.
32 Schulte-Schrepping J, Reusch N, Paclik D, *et al.* Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell* 2020; 182: 1419–1440.
33 Lukassen S, Chua RL, Trefzer T, *et al.* SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *EMBO J* 2020; 39: e105114.
34 Reyfman PA, Walter JM, Joshi N, *et al.* Single-cell transcriptomic analysis of human lung provides insights into the pathobiology of pulmonary fibrosis. *Am J Respir Crit Care Med* 2019; 199: 1517–1536.
35 Adams TS, Schupp JC, Poli S, *et al.* Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci Adv* 2020; 6: eaba1983.
36 Kelly JN, Laloli L, V'kovski P, *et al.* Comprehensive single cell analysis of pandemic influenza A virus infection in the human airways uncovers cell-type specific host transcriptional signatures relevant for disease progression and pathogenesis. *Front Immunol* 2022; 13: 978824.
37 Plasschaert LW, Žilionis R, Choo-Wing R, *et al.* A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* 2018; 560: 377–381.
38 Schupp JC, Adams TS, Cosme C Jr, *et al.* Integrated single-cell atlas of endothelial cells of the human lung. *Circulation* 2021; 144: 286–302.
39 Morse C, Tabib T, Sembrat J, *et al.* Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur Respir J* 2019; 54: 1802441.
40 Liao M, Liu Y, Yuan J, *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat Med* 2020; 26: 842–844.
41 Fischer J, Ayers T. Single nucleus RNA-sequencing: how it's done, applications and limitations. *Emerg Top Life Sci* 2021; 5: 687–690.
42 Efremova M, Vento-Tormo M, Teichmann SA, *et al.* CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 2020; 15: 1484–1506.
43 Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* 2020; 17: 159–162.
44 Saelens W, Cannoodt R, Todorov H, *et al.* A comparison of single-cell trajectory inference methods. *Nat Biotechnol* 2019; 37: 547–554.
45 Beckett EL, Stevens RL, Jarnicki AG, *et al.* A new short-term mouse model of chronic obstructive pulmonary disease identifies a role for mast cell tryptase in pathogenesis. *J Allergy Clin Immunol* 2013; 131: 752–762.
46 Xu C, Prete M, Webb S, *et al.* Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell* 2023; 186: 5876–5891.

47 Korsunsky I, Millard N, Fan J, *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods* 2019; 16: 1289–1296.
48 Al-Amrani S, Al-Jabri Z, Al-Zaabi A, *et al.* Proteomics: concepts and applications in human medicine. *World J Biol Chem* 2021; 12: 57–69.
49 Yu X, Schneiderhan-Marra N, Joos TO. Protein microarrays for personalized medicine. *Clin Chem* 2010; 56: 376–387.
50 Thomas SN, French D, Jannetto PJ, *et al.* Liquid chromatography-tandem mass spectrometry for clinical diagnostics. *Nat Rev Methods Primers* 2022; 2: 96.
51 Sung JY, Hwang Y, Shin MH, *et al.* Utility of conventional culture and MALDI-TOF MS for identification of microbial communities in bronchoalveolar lavage fluid in comparison with the GS Junior next generation sequencing system. *Ann Lab Med* 2018; 38: 110–118.
52 López-Campos JL, Arellano E, Calero C, *et al.* Determination of inflammatory biomarkers in patients with COPD: a comparison of different assays. *BMC Med Res Methodol* 2012; 12: 40.
53 Skerrett-Byrne DA, Bromfield EG, Murray HC, *et al.* Time-resolved proteomic profiling of cigarette smoke-induced experimental chronic obstructive pulmonary disease. *Respirology* 2021; 26: 960–973.
54 Gharib SA, Nguyen EV, Lai Y, *et al.* Induced sputum proteome in healthy subjects and asthmatic patients. *J Allergy Clin Immunol* 2011; 128: 1176–1184.
55 Ohlmeier S, Nieminen P, Gao J, *et al.* Lung tissue proteomics identifies elevated transglutaminase 2 levels in stable chronic obstructive pulmonary disease. *Am J Physiol Lung Cell Mol Physiol* 2016; 310: L1155–L1165.
56 Huang JT, Cant E, Keir HR, *et al.* Endotyping chronic obstructive pulmonary disease, bronchiectasis, and the "chronic obstructive pulmonary disease-bronchiectasis association". *Am J Respir Crit Care Med* 2022; 206: 417–426.
57 Woldhuis RR, Heijink IH, van den Berge M, *et al.* COPD-derived fibroblasts secrete higher levels of senescence-associated secretory phenotype proteins. *Thorax* 2021; 76: 508–511.
58 Serban KA, Pratte KA, Strange C, *et al.* Unique and shared systemic biomarkers for emphysema in alpha-1 antitrypsin deficiency and chronic obstructive pulmonary disease. *EBioMedicine* 2022; 84: 104262.
59 Brandsma CA, Guryev V, Timens W, *et al.* Integrated proteogenomic approach identifying a protein signature of COPD and a new splice variant of SORBS1. *Thorax* 2020; 75: 180–183.
60 Chen L, Lu W, Wang L, *et al.* Metabolite discovery through global annotation of untargeted metabolomics data. *Nat Methods* 2021; 18: 1377–1385.
61 Wishart DS. Metabolomics for investigating physiological and pathophysiological processes. *Physiol Rev* 2019; 99: 1819–1875.
62 Broadhurst D, Goodacre R, Reinke SN, *et al.* Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 2018; 14: 72.
63 Chaleckis R, Meister I, Zhang P, *et al.* Challenges, progress and promises of metabolite annotation for LC-MS-based metabolomics. *Curr Opin Biotechnol* 2019; 55: 44–50.
64 Wang M, Carver JJ, Phelan VV, *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* 2016; 34: 828–837.
65 Schönberger K, Mitterer M, Buescher JM, *et al.* Targeted LC-MS/MS-based metabolomics and lipidomics on limited hematopoietic stem cell numbers. *STAR Protoc* 2022; 3: 101408.
66 Wieder C, Frainay C, Poupin N, *et al.* Pathway analysis in metabolomics: recommendations for the use of over-representation analysis. *PLoS Comput Biol* 2021; 17: e1009105.
67 Kachroo P, Stewart ID, Kelly RS, *et al.* Metabolomic profiling reveals extensive adrenal suppression due to inhaled corticosteroid therapy in asthma. *Nat Med* 2022; 28: 814–822.
68 Kelly RS, Mendez KM, Huang M, *et al.* Metabo-endotypes of asthma reveal differences in lung function: discovery and validation in two TOPMed cohorts. *Am J Respir Crit Care Med* 2022; 205: 288–299.
69 Reinke SN, Naz S, Chaleckis R, *et al.* Urinary metabotype of severe asthma evidences decreased carnitine metabolism independent of oral corticosteroid treatment in the U-BIOPRED study. *Eur Respir J* 2022; 59: 2101733.
70 Fujimura KE, Sitarik AR, Havstad S, *et al.* Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med* 2016; 22: 1187–1191.
71 Meier C, Freiburghaus K, Bovet C, *et al.* Serum metabolites as biomarkers in systemic sclerosis-associated interstitial lung disease. *Sci Rep* 2020; 10: 21912.
72 Rindlisbacher B, Schmid C, Geiser T, *et al.* Serum metabolic profiling identified a distinct metabolic signature in patients with idiopathic pulmonary fibrosis – a potential biomarker role for LysoPC. *Respir Res* 2018; 19: 7.
73 Rago D, Pedersen CT, Huang M, *et al.* Characteristics and mechanisms of a sphingolipid-associated childhood asthma endotype. *Am J Respir Crit Care Med* 2021; 203: 853–863.
74 Bowler RP, Jacobson S, Cruickshank C, *et al.* Plasma sphingolipids associated with chronic obstructive pulmonary disease phenotypes. *Am J Respir Crit Care Med* 2015; 191: 275–284.
75 Esther CR Jr, O'Neal WK, Anderson WH, *et al.* Identification of sputum biomarkers predictive of pulmonary exacerbations in COPD. *Chest* 2022; 161: 1239–1249.

76 Alexandrov T. Spatial metabolomics and imaging mass spectrometry in the age of artificial intelligence. *Annu Rev Biomed Data Sci* 2020; 3: 61–87.
77 van Timmeren JE, Cester D, Tanadini-Lang S, *et al.* Radiomics in medical imaging –"how-to" guide and critical reflection. *Insights Imaging* 2020; 11: 91.
78 Frix AN, Cousin F, Refaee T, *et al.* Radiomics in lung diseases imaging: state-of-the-art for clinicians. *J Pers Med* 2021; 11: 602.
79 Ginsburg SB, Lynch DA, Bowler RP, *et al.* Automated texture-based quantification of centrilobular nodularity and centrilobular emphysema in chest CT images. *Acad Radiol* 2012; 19: 1241–1251.
80 Lauer D, Magnin CY, Kolly LR, *et al*. Radioproteomics stratifies molecular response to antifibrotic treatment in pulmonary fibrosis. *JCI Insight* 2024; 9.
81 Schniering J, Maciukiewicz M, Gabrys HS, *et al*. Computed tomography-based radiomics decodes prognostic and molecular differences in interstitial lung disease related to systemic sclerosis. *Eur Respir J* 2022; 59: 2004503.
82 Logotheti S, Georgakilas AG. More than meets the eye: integration of radiomics with transcriptomics for reconstructing the tumor microenvironment and predicting response to therapy. *Cancers* 2023; 15: 1634.
83 Lee S-W, Kuan C-S, Wu LS-H, *et al.* Metagenome and metatranscriptome profiling of moderate and severe COPD sputum in Taiwanese Han males. *PLoS One* 2016; 11: e0159066.
84 Janda JM, Abbott SL. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 2007; 45: 2761–2764.
85 Knight R, Vrbanac A, Taylor BC, *et al.* Best practices for analysing microbiomes. *Nat Rev Microbiol* 2018; 16: 410–422.
86 Li CX, Wheelock CE, Sköld CM, *et al.* Integration of multi-omics datasets enables molecular classification of COPD. *Eur Respir J* 2018; 51: 1701930.
87 Williams CG, Lee HJ, Asatsuma T, *et al.* An introduction to spatial transcriptomics for biomedical research. *Genome Med* 2022; 14: 68.
88 Zhang H, Hunter MV, Chou J, *et al*. BayesTME: an end-to-end method for multiscale spatial transcriptional profiling of the tissue microenvironment. *Cell Systems* 2023; 14: 605–619.e7.
89 Marx V. Method of the year: spatially resolved transcriptomics. *Nat Methods* 2021; 18: 9–14.
90 Sountoulidis A, Liontos A, Nguyen HP, *et al.* SCRINSHOT enables spatial mapping of cell states in tissue sections with single-cell resolution. *PLoS Biol* 2020; 18: e3000675.
91 Madissoon E, Oliver AJ, Kleshchevnikov V, *et al.* A spatially resolved atlas of the human lung characterizes a gland-associated immune niche. *Nat Genet* 2023; 55: 66–77.
92 Cross AR, de Andrea CE, Villalba-Esparza M, *et al.* Spatial transcriptomic characterization of COVID-19 pneumonitis identifies immune circuits related to tissue injury. *JCI Insight* 2023; 8: e157837.
93 Keow J, Cecchini MJ, Jayawardena N, *et al.* Digital quantification of p16-positive foci in fibrotic interstitial lung disease is associated with a phenotype of idiopathic pulmonary fibrosis with reduced survival. *Respir Res* 2022; 23: 147.
94 Sountoulidis A, Marco Salas S, Braun E, *et al.* A topographic atlas defines developmental origins of cell heterogeneity in the human embryonic lung. *Nat Cell Biol* 2023; 25: 351–365.
95 Hao Y, Hao S, Andersen-Nissen E, *et al.* Integrated analysis of multimodal single-cell data. *Cell* 2021; 184: 3573–3587.
96 Sariyar S, Sountoulidis A, Hansen JN, *et al.* High-parametric protein maps reveal the spatial organization in early-developing human lung. *Nat Commun* 2024; 15: 9381.
97 O'Rourke MB, Roediger BR, Jolly CJ, *et al.* Viral biomarker detection and validation using MALDI mass spectrometry imaging (MSI). *Proteomes* 2022; 10: 33.
98 Rosenberger FA, Thielert M, Strauss MT, *et al.* Spatial single-cell mass spectrometry defines zonation of the hepatocyte proteome. *Nat Methods* 2023; 20: 1530–1536.
99 Borner GHH. Spatial proteomics: a gateway to understanding cell biology. *Proteomics* 2020; 20: e1900328.
100 Vickovic S, Lötstedt B, Klughammer J, *et al.* SM-Omics is an automated platform for high-throughput spatial multi-omics. *Nat Commun* 2022; 13: 795.
101 Liu Y, DiStasio M, Su G, *et al.* High-plex protein and whole transcriptome co-mapping at cellular resolution with spatial CITE-seq. *Nat Biotechnol* 2023; 41: 1405–1409.
102 Ben-Chetrit N, Niu X, Swett AD, *et al.* Integration of whole transcriptome spatial profiling with protein markers. *Nat Biotechnol* 2023; 41: 788–793.
103 Du Y, Clair GC, Al Alam D, *et al.* Integration of transcriptomic and proteomic data identifies biological functions in cell populations from human infant lung. *Am J Physiol Lung Cell Mol Physiol* 2019; 317: L347–L360.
104 Conroy LR, Clarke HA, Allison DB, *et al.* Spatial metabolomics reveals glycogen as an actionable target for pulmonary fibrosis. *Nat Commun* 2023; 14: 2759.
105 Smith MJ, Nie M, Adner M, *et al.* Development of a desorption electrospray ionization-multiple-reaction-monitoring mass spectrometry (DESI-MRM) workflow for spatially mapping oxylipins in pulmonary tissue. *Anal Chem* 2024; 96: 17950–17959.

106 Hou J, Acharya L, Zhu D, *et al.* An overview of bioinformatics methods for modeling biological pathways in yeast. *Brief Funct Genomics* 2016; 15: 95–108.
107 Wu Z, Liao Q, Liu B. A comprehensive review and evaluation of computational methods for identifying protein complexes from protein-protein interaction networks. *Brief Bioinform* 2020; 21: 1531–1548.
108 Idrees S, Paudel KR. Proteome-wide assessment of human interactome as a source of capturing domain-motif and domain-domain interactions. *J Cell Commun Signal* 2024; 18: e12014.
109 Idrees S, Paudel KR, Sadaf T, *et al.* Uncovering domain motif interactions using high-throughput protein-protein interaction detection methods. *FEBS Lett* 2024; 598: 725–742.
110 Idrees S, Chen H, Panth N, *et al.* Exploring viral-host protein interactions as antiviral therapies: a computational perspective. *Microorganisms* 2024; 12: 630.
111 Chatr-Aryamontri A, Breitkreutz BJ, Oughtred R, *et al.* The BioGRID interaction database: 2015 update. *Nucleic Acids Res* 2015; 43: D470–D478.
112 Del Toro N, Shrivastava A, Ragueneau E, *et al.* The IntAct database: efficient access to fine-grained molecular interaction data. *Nucleic Acids Res* 2022; 50: D648–D653.
113 Szklarczyk D, Gable AL, Lyon D, *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; 47: D607–D613.
114 Idrees S, Paudel KR, Sadaf T, *et al.* How different viruses perturb host cellular machinery *via* short linear motifs. *EXCLI J* 2023; 22: 1113–1128.
115 Lu H, Zhou Q, He J, *et al.* Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct Target Ther* 2020; 5: 213.
116 Lawrence HR, Li Z, Yip ML, *et al.* Identification of a disruptor of the MDM2-p53 protein-protein interaction facilitated by high-throughput *in silico* docking. *Bioorg Med Chem Lett* 2009; 19: 3756–3759.
117 Tian W, Han X, Yan M, *et al.* Structure-based discovery of a novel inhibitor targeting the β-catenin/Tcf4 interaction. *Biochemistry* 2012; 51: 724–731.
118 Almet AA, Cang Z, Jin S, *et al.* The landscape of cell-cell communication through single-cell transcriptomics. *Curr Opin Syst Biol* 2021; 26: 12–23.
119 Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008; 9: 559.
120 Oestreich M, Holsten L, Agrawal S, *et al.* hCoCena: horizontal integration and analysis of transcriptomics datasets. *Bioinformatics* 2022; 38: 4727–4734.
121 Lemoine GG, Scott-Boyer MP, Ambroise B, *et al.* GWENA: gene co-expression networks analysis and extended modules characterization in a single Bioconductor package. *BMC Bioinformatics* 2021; 22: 267.
122 Buphamalai P, Kokotovic T, Nagy V, *et al.* Network analysis reveals rare disease signatures across multiple levels of biological organization. *Nat Commun* 2021; 12: 6306.
123 Cheng C, Alexander R, Min R, *et al.* Understanding transcriptional regulation by integrative analysis of transcription factor binding data. *Genome Res* 2012; 22: 1658–1667.
124 Rubin JD, Stanley JT, Sigauke RF, *et al.* Transcription factor enrichment analysis (TFEA) quantifies the activity of multiple transcription factors from a single experiment. *Commun Biol* 2021; 4: 661.
125 Keenan AB, Torre D, Lachmann A, *et al.* ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res* 2019; 47: W212–W224.
126 Alvarez MJ, Shen Y, Giorgi FM, *et al.* Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet* 2016; 48: 838–847.
127 Guang Z, Min Z, Jun-Tan L, *et al.* Single-cell protein activity analysis reveals a novel subpopulation of chondrocytes and the corresponding key master regulator proteins associated with anti-senescence and OA progression. *Front Immunol* 2023; 14: 1077003.
128 Margolin AA, Nemenman I, Basso K, *et al.* ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 2006; 7: Suppl. 1, S7.
129 Holland CH, Tanevski J, Perales-Patón J, *et al.* Robustness and applicability of transcription factor and pathway analysis tools on single-cell RNA-seq data. *Genome Biol* 2020; 21: 36.
130 Aibar S, González-Blas CB, Moerman T, *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017; 14: 1083–1086.
131 Meiler A, Marchiano F, Haering M, *et al.* AnnoMiner is a new web-tool to integrate epigenetics, transcription factor occupancy and transcriptomics data to predict transcriptional regulators. *Sci Rep* 2021; 11: 15463.
132 Palsson B, Zengler K. The challenges of integrating multi-omic data sets. *Nat Chem Biol* 2010; 6: 787–789.
133 Dhillon A, Singh A, Bhalla VK. A systematic review on biomarker identification for cancer diagnosis and prognosis in multi-omics: from computational needs to machine learning and deep learning. *Arch Comput Methods Eng* 2023; 30: 917–949.
134 Meng C, Kuster B, Culhane AC, *et al.* A multivariate approach to the integration of multi-omics datasets. *BMC Bioinformatics* 2014; 15: 162.

135 Broman KW, Gatti DM, Simecek P, *et al.* R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics* 2019; 211: 495–502.
136 Vahabi N, Michailidis G. Unsupervised multi-omics data integration methods: a comprehensive review. *Front Genet* 2022; 13: 854752.
137 Shabalin AA. Matrix eQTL: ultra fast eQTL analysis *via* large matrix operations. *Bioinformatics* 2012; 28: 1353–1358.
138 Qi J, Asl HF, Björkegren J, *et al.* kruX: matrix-based non-parametric eQTL discovery. *BMC Bioinformatics* 2014; 15: 11.
139 Speicher NK, Pfeifer N. Integrating different data types by regularized unsupervised multiple kernel learning with application to cancer subtype discovery. *Bioinformatics* 2015; 31: i268–i275.
140 Cao H, Jia C, Li Z, *et al.* wMKL: multi-omics data integration enables novel cancer subtype identification *via* weight-boosted multi-kernel learning. *Br J Cancer* 2024; 130: 1001–1012.
141 Ruan P, Wang Y, Shen R, *et al.* Using association signal annotations to boost similarity network fusion. *Bioinformatics* 2019; 35: 3718–3726.
142 Yuan Y, Savage RS, Markowetz F. Patient-specific data fusion defines prognostic cancer subtypes. *PLoS Comput Biol* 2011; 7: e1002227.
143 Shaddox E, Stingo FC, Peterson CB, *et al.* A Bayesian approach for learning gene networks underlying disease severity in COPD. *Stat Biosci* 2018; 10: 59–85.
144 Vaske CJ, Benz SC, Sanborn JZ, *et al.* Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 2010; 26: i237–i245.
145 Li Y, Ma L, Wu D, *et al.* Advances in bulk and single-cell multi-omics approaches for systems biology and precision medicine. *Brief Bioinform* 2021; 22: bbab024.
146 Mo Q, Wang S, Seshan VE, *et al.* Pattern discovery and cancer gene identification in integrated cancer genomic data. *Proc Natl Acad Sci USA* 2013; 110: 4245–4250.
147 Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014; 511: 543–550.
148 Greenbaum D, Colangelo C, Williams K, *et al.* Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol* 2003; 4: 117.
149 Liu Y, Beyer A, Aebersold R. On the dependency of cellular protein levels on mRNA abundance. *Cell* 2016; 165: 535–550.
150 Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol* 2019; 15: e8746.
151 Zhang W, Zhang Y, Li L, *et al.* Unraveling heterogeneity and treatment of asthma through integrating multi-omics data. *Front Allergy* 2024; 5: 1496392.
152 Huang Q, Wang Y, Zhang L, *et al.* Single-cell transcriptomics highlights immunological dysregulations of monocytes in the pathobiology of COPD. *Respir Res* 2022; 23: 367.
153 Argelaguet R, Cuomo ASE, Stegle O, *et al.* Computational principles and challenges in single-cell data integration. *Nat Biotechnol* 2021; 39: 1202–1215.
154 Zuo C, Dai H, Chen L. Deep cross-omics cycle attention model for joint analysis of single-cell multi-omics data. *Bioinformatics* 2021; 37: 4091–4099.
155 Mei S, Li X, Zhou Y, *et al.* Deep learning for detecting and early predicting chronic obstructive pulmonary disease from spirogram time series. *NPJ Syst Biol Appl* 2025; 11: 18.
156 Argelaguet R, Velten B, Arnol D, *et al.* Multi-Omics Factor Analysis-a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol* 2018; 14: e8124.
157 Argelaguet R, Arnol D, Bredikhin D, *et al.* MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol* 2020; 21: 111.
158 Jin S, Zhang L, Nie Q. scAI: an unsupervised approach for the integrative analysis of parallel single-cell transcriptomic and epigenomic profiles. *Genome Biol* 2020; 21: 25.
159 Ma A, McDermaid A, Xu J, *et al.* Integrative methods and practical challenges for single-cell multi-omics. *Trends Biotechnol* 2020; 38: 1007–1022.
160 Zhuang Y, Xing F, Ghosh D, *et al.* Deep learning on graphs for multi-omics classification of COPD. *PLoS One* 2023; 18: e0284563.
161 Packer J, Trapnell C. Single-cell multi-omics: an engine for new quantitative models of gene regulation. *Trends Genet* 2018; 34: 653–665.
162 Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 2012; 44: 821–824.
163 Lippert C, Listgarten J, Liu Y, *et al.* FaST linear mixed models for genome-wide association studies. *Nat Methods* 2011; 8: 833–835.
164 Cuomo ASE, Nathan A, Raychaudhuri S, *et al.* Single-cell genomics meets human genetics. *Nat Rev Genet* 2023; 24: 535–549.
165 Kamimoto K, Stringa B, Hoffmann CM, *et al.* Dissecting cell identity *via* network inference and *in silico* gene perturbation. *Nature* 2023; 614: 742–751.

166 Bravo González-Blas C, De Winter S, Hulselmans G, *et al.* SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat Methods* 2023; 20: 1355–1367.
167 Welch JD, Hartemink AJ, Prins JF. MATCHER: manifold alignment reveals correspondence between single cell transcriptome and epigenome dynamics. *Genome Biol* 2017; 18: 138.
168 Cao K, Hong Y, Wan L. Manifold alignment for heterogeneous single-cell multi-omics data integration using Pamona. *Bioinformatics* 2021; 38: 211–219.
169 Cao K, Bai X, Hong Y, *et al.* Unsupervised topological alignment for single-cell multi-omics integration. *Bioinformatics* 2020; 36: Suppl. 1, i48–i56.
170 Hao Y, Stuart T, Kowalski MH, *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* 2024; 42: 293–304.
171 Wang C, Sun D, Huang X, *et al.* Integrative analyses of single-cell transcriptome and regulome using MAESTRO. *Genome Biol* 2020; 21: 198.
172 Curion F, Theis FJ. Machine learning integrative approaches to advance computational immunology. *Genome Med* 2024; 16: 80.
173 Wan X, Xiao J, Tam SST, *et al.* Integrating spatial and single-cell transcriptomics data using deep generative models with SpatialScope. *Nat Commun* 2023; 14: 7848.
174 Dong M, Thennavan A, Urrutia E, *et al.* SCDC: bulk gene expression deconvolution by multiple single-cell RNA sequencing references. *Brief Bioinform* 2021; 22: 416–427.
175 Fan J, Lyu Y, Zhang Q, *et al.* MuSiC2: cell-type deconvolution for multi-condition bulk RNA-seq data. *Brief Bioinform* 2022; 23: bbac430.
176 Frishberg A, Peshes-Yaloz N, Cohn O, *et al.* Cell composition analysis of bulk genomics using single-cell data. *Nat Methods* 2019; 16: 327–332.
177 Longo SK, Guo MG, Ji AL, *et al.* Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet* 2021; 22: 627–644.
178 Kleshchevnikov V, Shmatko A, Dann E, *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat Biotechnol* 2022; 40: 661–671.
179 Cable DM, Murray E, Zou LS, *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat Biotechnol* 2022; 40: 517–526.
180 Dong R, Yuan GC. SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol* 2021; 22: 145.
181 Elosua-Bayes M, Nieto P, Mereu E, *et al.* SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res* 2021; 49: e50.
182 Ma Y, Zhou X. Spatially informed cell-type deconvolution for spatial transcriptomics. *Nat Biotechnol* 2022; 40: 1349–1359.
183 Long Y, Ang KS, Sethi R, *et al.* Deciphering spatial domains from spatial multi-omics with SpatialGlue. *Nat Methods* 2024; 21: 1658–1667.
184 Qian X, Harris KD, Hauling T, *et al.* Probabilistic cell typing enables fine mapping of closely related cell types *in situ*. *Nat Methods* 2020; 17: 101–106.
185 Welch JD, Kozareva V, Ferreira A, *et al.* Single-cell multi-omic integration compares and contrasts features of brain cell identity. *Cell* 2019; 177: 1873–1887.
186 Haviv D, Remšík J, Gatie M, *et al.* The covariance environment defines cellular niches for spatial inference. *Nat Biotechnol* 2025; 43: 269–280.
187 Schaar AC, Tejada-Lapuerta A, Palla G, *et al.* Nicheformer: a foundation model for single-cell and spatial omics. *bioRxiv* 2024; preprint [https://doi.org/10.1101/2024.04.15.589472].
188 Partel G, Wählby C. Spage2vec: unsupervised representation of localized spatial gene expression signatures. *FEBS J* 2021; 288: 1859–1870.
189 Svensson V, Teichmann SA, Stegle O. SpatialDE: identification of spatially variable genes. *Nat Methods* 2018; 15: 343–346.
190 Schwartz U, Llamazares Prada M, Pohl ST, *et al.* High-resolution transcriptomic and epigenetic profiling identifies novel regulators of COPD. *EMBO J* 2023; 42: e111272.
191 Chotirmall SH, Bogaert D, Chalmers JD, *et al.* Therapeutic targeting of the respiratory microbiome. *Am J Respir Crit Care Med* 2022; 206: 535–544.
192 Lynch DA, Austin JH, Hogg JC, *et al.* CT-definable subtypes of chronic obstructive pulmonary disease: a statement of the Fleischner Society. *Radiology* 2015; 277: 192–205.
193 Li B, Liu J, Cao Y, *et al.* Multi-omics characterization of early chronic obstructive pulmonary disease. *Respir Res* 2025; 26: 167.
194 Yan Z, Chen B, Yang Y, *et al.* Multi-omics analyses of airway host-microbe interactions in chronic obstructive pulmonary disease identify potential therapeutic interventions. *Nat Microbiol* 2022; 7: 1361–1375.
195 Bowerman KL, Rehman SF, Vaughan A, *et al.* Disease-associated gut microbiome and metabolome changes in patients with chronic obstructive pulmonary disease. *Nat Commun* 2020; 11: 5886.
196 Budden KF, Shukla SD, Bowerman KL, *et al.* Faecal microbial transfer and complex carbohydrates mediate protection against COPD. *Gut* 2024; 73: 751–769.

197 Abdel-Aziz MI, Vijverberg SJH, Neerincx AH, *et al.* A multi-omics approach to delineate sputum microbiome-associated asthma inflammatory phenotypes. *Eur Respir J* 2022; 59: 2102603.
198 Sunata K, Miyata J, Kawashima Y, *et al.* Multiomics analysis identified IL-4-induced IL1RL1^{high} eosinophils characterized by prominent cysteinyl leukotriene metabolism. *J Allergy Clin Immunol* 2024; 154: 1277–1288.
199 Allen RJ, Porte J, Braybrooke R, *et al.* Genetic variants associated with susceptibility to idiopathic pulmonary fibrosis in people of European ancestry: a genome-wide association study. *Lancet Respir Med* 2017; 5: 869–880.
200 Wen Z, Liang W, Yang Z, *et al.* Genetic insights into idiopathic pulmonary fibrosis: a multi-omics approach to identify potential therapeutic targets. *J Transl Med* 2025; 23: 337.
201 Zhang M, Zhang J, Hu H, *et al.* Multiomic analysis of monocyte-derived alveolar macrophages in idiopathic pulmonary fibrosis. *J Transl Med* 2024; 22: 598.
202 Ng N, Powell CA. Targeting the complement cascade in the pathophysiology of COVID-19 disease. *J Clin Med* 2021; 10: 2188.
203 Ma J, Deng Y, Zhang M, *et al.* The role of multi-omics in the diagnosis of COVID-19 and the prediction of new therapeutic targets. *Virulence* 2022; 13: 1101–1110.
204 Su Y, Chen D, Yuan D, *et al.* Multi-omics resolves a sharp disease-state shift between mild and moderate COVID-19. *Cell* 2020; 183: 1479–1495.
205 Barh D, Tiwari S, Weener ME, *et al.* Multi-omics-based identification of SARS-CoV-2 infection biology and candidate drugs against COVID-19. *Comput Biol Med* 2020; 126: 104051.
206 Djeddi S, Fernandez-Salinas D, Huang GX, *et al.* Rhinovirus infection of airway epithelial cells uncovers the non-ciliated subset as a likely driver of genetic risk to childhood-onset asthma. *Cell Genom* 2024; 4: 100636.
207 Raita Y, Camargo CA Jr, Bochkov YA, *et al.* Integrated-omics endotyping of infants with rhinovirus bronchiolitis and risk of childhood asthma. *J Allergy Clin Immunol* 2021; 147: 2108–2117.
208 Huang X, Li F, Wang Y, *et al.* Multi-omics analysis reveals underlying host responses in pediatric respiratory syncytial virus pneumonia. *iScience* 2023; 26: 106329.
209 Raita Y, Pérez-Losada M, Freishtat RJ, *et al.* Integrated omics endotyping of infants with respiratory syncytial virus bronchiolitis and risk of childhood asthma. *Nat Commun* 2021; 12: 3601.
210 Zhu Z, You R, Li H, *et al.* Multi-omics data integration reveals the complexity and diversity of host factors associated with influenza virus infection. *PeerJ* 2023; 11: e16194.
211 Riese P, Trittel S, Akmatov MK, *et al.* Distinct immunological and molecular signatures underpinning influenza vaccine responsiveness in the elderly. *Nat Commun* 2022; 13: 6894.
212 Martin-Sancho L, Tripathi S, Rodriguez-Frandsen A, *et al.* Restriction factor compendium for influenza A virus reveals a mechanism for evasion of autophagy. *Nat Microbiol* 2021; 6: 1319–1333.
213 Platt MP, Lin YH, Penix T, *et al.* A multiomics analysis of direct interkingdom dynamics between influenza A virus and *Streptococcus pneumoniae* uncovers host-independent changes to bacterial virulence fitness. *PLoS Pathog* 2022; 18: e1011020.
214 Hu Q, Liu B, Fan Y, *et al.* Multi-omics association analysis reveals interactions between the oropharyngeal microbiome and the metabolome in pediatric patients with influenza A virus pneumonia. *Front Cell Infect Microbiol* 2022; 12: 1011254.
215 Tu X, Kim RY, Brown AC, *et al.* Airway and parenchymal transcriptomics in a novel model of asthma and COPD overlap. *J Allergy Clin Immunol* 2022; 150: 817–829.
216 Hansbro P. Omics technologies to study virus infection and chronic lung diseases. *Respirology* 2023; 28: 403.
217 Sulaiman I, Wu BG, Chung M, *et al.* Lower airway dysbiosis augments lung inflammatory injury in mild-to-moderate chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2023; 208: 1101–1114.
218 Kapellos TS, Bassler K, Fujii W, *et al.* Systemic alterations in neutrophils and their precursors in early-stage chronic obstructive pulmonary disease. *Cell Rep* 2023; 42: 112525.
219 Liu G, Haw TJ, Starkey MR, *et al.* TLR7 promotes smoke-induced experimental lung damage through the activity of mast cell tryptase. *Nat Commun* 2023; 14: 7349.
220 Liu G, Jarnicki AG, Paudel KR, *et al.* Adverse roles of mast cell chymase-1 in COPD. *Eur Respir J* 2022; 60: 2101431.
221 Thapa R, Magar AT, Shrestha J, *et al.* Influence of gut and lung dysbiosis on lung cancer progression and their modulation as promising therapeutic targets: a comprehensive review. *MedComm* 2024; 5: e70018.
222 Paudel KR, Jha SK, Allam V, *et al.* Recent advances in chronotherapy targeting respiratory diseases. *Pharmaceutics* 2021; 13: 2008.
223 Casas-Recasens S, Noell G, Mendoza N, *et al.* Lung DNA methylation in chronic obstructive pulmonary disease: relationship with smoking status and airflow limitation severity. *Am J Respir Crit Care Med* 2021; 203: 129–134.
224 Günes Günsel G, Conlon TM, Jeridi A, *et al.* The arginine methyltransferase PRMT7 promotes extravasation of monocytes resulting in tissue injury in COPD. *Nat Commun* 2022; 13: 1303.
225 Koba T, Takeda Y, Narumi R, *et al.* Proteomics of serum extracellular vesicles identifies a novel COPD biomarker, fibulin-3 from elastic fibres. *ERJ Open Res* 2021; 7: 00658-2020.

226 Budden KF, Shukla SD, Rehman SF, *et al.* Functional effects of the microbiota in chronic respiratory disease. *Lancet Respir Med* 2019; 7: 907–920.
227 Budden KF, Gellatly SL, Wood DLA, *et al.* Emerging pathogenic links between microbiota and the gut–lung axis. *Nat Rev Microbiol* 2017; 15: 55–63.
228 Shukla SD, Budden KF, Neal R, *et al.* Microbiome effects on immunity, health and disease in the lung. *Clin Transl Immunology* 2017; 6: e133.
229 Donovan C, Liu G, Shen S, *et al.* The role of the microbiome and the NLRP3 inflammasome in the gut and lung. *J Leukoc Biol* 2020; 108: 925–935.
230 Madapoosi SS, Cruickshank-Quinn C, Opron K, *et al*. Lung microbiota and metabolites collectively associate with clinical outcomes in milder stage chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2022; 206: 427–439.
231 Zhang Z, Wang J, Li Y, *et al.* Proteomics and metabolomics profiling reveal panels of circulating diagnostic biomarkers and molecular subtypes in stable COPD. *Respir Res* 2023; 24: 73.
232 Liang W, Yang Y, Gong S, *et al.* Airway dysbiosis accelerates lung function decline in chronic obstructive pulmonary disease. *Cell Host Microbe* 2023; 31: 1054–1070.
233 Alemao CA, Budden KF, Gomez HM, *et al.* Impact of diet and the bacterial microbiome on the mucous barrier and immune disorders. *Allergy* 2021; 76: 714–734.
234 Suryadevara R, Gregory A, Lu R, *et al.* Blood-based transcriptomic and proteomic biomarkers of emphysema. *Am J Respir Crit Care Med* 2024; 209: 273–287.
235 Rojas-Quintero J, Ochsner SA, New F, *et al.* Spatial transcriptomics resolve an emphysema-specific lymphoid follicle B cell signature in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2024; 209: 48–58.
236 Ryu MH, Yun JH, Morrow JD, *et al.* Blood gene expression and immune cell subtypes associated with chronic obstructive pulmonary disease exacerbations. *Am J Respir Crit Care Med* 2023; 208: 247–255.
237 Silhan LL, Shah PD, Chambers DC, *et al.* Lung transplantation in telomerase mutation carriers with pulmonary fibrosis. *Eur Respir J* 2014; 44: 178–187.
238 Lucas SEM, Raspin K, Mackintosh J, *et al.* Preclinical interstitial lung disease in relatives of familial pulmonary fibrosis patients. *Pulmonology* 2023; 29: 257–260.
239 Habermann AC, Gutierrez AJ, Bui LT, *et al.* Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Sci Adv* 2020; 6: eaba1972.
240 Kobayashi Y, Tata A, Konkimalla A, *et al.* Persistence of a regeneration-associated, transitional alveolar epithelial cell state in pulmonary fibrosis. *Nat Cell Biol* 2020; 22: 934–946.
241 Choi J, Park JE, Tsagkogeorga G, *et al.* Inflammatory signals induce AT2 cell-derived damage-associated transient progenitors that mediate alveolar regeneration. *Cell Stem Cell* 2020; 27: 366–382.
242 Raslan AA, Pham TX, Lee J, *et al*. Lung injury-induced activated endothelial cell states persist in aging-associated progressive fibrosis. *Nat Commun* 2024; 15: 433.
243 Neumark N, Cosme C Jr, Rose KA, *et al.* The idiopathic pulmonary fibrosis cell atlas. *Am J Physiol Lung Cell Mol Physiol* 2020; 319: L887–L893.
244 Konigsberg IR, Borie R, Walts AD, *et al.* Molecular signatures of idiopathic pulmonary fibrosis. *Am J Respir Cell Mol Biol* 2021; 65: 430–441.
245 Zheng P, Sun S, Wang J, *et al.* Integrative omics analysis identifies biomarkers of idiopathic pulmonary fibrosis. *Cell Mol Life Sci* 2022; 79: 66.
246 Ruan P, Todd JL, Zhao H, *et al.* Integrative multi-omics analysis reveals novel idiopathic pulmonary fibrosis endotypes associated with disease progression. *Respir Res* 2023; 24: 141.
247 Mayr CH, Simon LM, Leuschner G, *et al.* Integrative analysis of cell state changes in lung fibrosis with peripheral protein biomarkers. *EMBO Mol Med* 2021; 13: e12871.
248 Guo S, Dong Y, Wang C, *et al.* Integrative analysis reveals the recurrent genetic etiologies in idiopathic pulmonary fibrosis. *QJM* 2023; 116: 983–992.
249 Ren F, Aliper A, Chen J, *et al.* A small-molecule TNIK inhibitor targets fibrosis in preclinical and clinical models. *Nat Biotechnol* 2025; 43: 63–75.
250 Murugan C, Ramamoorthy S, Kuppuswamy G, *et al.* COVID-19: a review of newly formed viral clades, pathophysiology, therapeutic strategies and current vaccination tasks. *Int J Biol Macromol* 2021; 193: 1165–1200.
251 Sameh M, Khalaf HM, Anwar AM, *et al.* Integrated multiomics analysis to infer COVID-19 biological insights. *Sci Rep* 2023; 13: 1802.
252 Sulaiman I, Chung M, Angel L, *et al.* Microbial signatures in the lower airways of mechanically ventilated COVID-19 patients associated with poor clinical outcome. *Nat Microbiol* 2021; 6: 1245–1258.
253 Nogales A, Martínez-Sobrido L. Reverse genetics approaches for the development of influenza vaccines. *Int J Mol Sci* 2016; 18: 20.
254 Byeon SK, Madugundu AK, Garapati K, *et al.* Development of a multiomics model for identification of predictive biomarkers for COVID-19 severity: a retrospective cohort study. *Lancet Digit Health* 2022; 4: e632–e645.

255 Wang X, Ulloa L. Editorial: Omics in respiratory virus infectious diseases: integrating multi-omics to reveal data characteristics and mechanisms for the diagnosis and treatment of disease. *Front Med* 2023; 10: 1273662.

256 Owuor N, Nalamala N, Gimenes JA Jr, *et al.* Rhinovirus and COPD airway epithelium. *Pulm Crit Care Med* 2017; 2: 10.15761/PCCM.1000139.

257 Steinke JW, Borish L. Immune responses in rhinovirus-induced asthma exacerbations. *Curr Allergy Asthma Rep* 2016; 16: 78.

258 Langedijk AC, Bont LJ. Respiratory syncytial virus infection and novel interventions. *Nat Rev Microbiol* 2023; 21: 734–749.

259 Bajinka O, Ouedraogo SY, Li N, *et al.* Multiomics as instrument to promote 3P medical approaches for the overall management of respiratory syncytial viral infections. *EPMA J* 2025; 16: 217–238.

260 Paules CI, Sullivan SG, Subbarao K, *et al.* Chasing seasonal influenza – the need for a universal influenza vaccine. *N Engl J Med* 2018; 378: 7–9.

261 de Vries RD, Altenburg AF, Rimmelzwaan GF. Universal influenza vaccines: a realistic option? *Clin Microbiol Infect* 2016; 22: Suppl. 5, S120–S124.

262 Ota M, Fujio K. Multi-omics approach to precision medicine for immune-mediated diseases. *Inflamm Regen* 2021; 41: 23.

263 Li CL, Liu SF. Exploring molecular mechanisms and biomarkers in COPD: an overview of current advancements and perspectives. *Int J Mol Sci* 2024; 25: 7347.

264 Clark C, Rabl M, Dayon L, *et al.* The promise of multi-omics approaches to discover biological alterations with clinical relevance in Alzheimer's disease. *Front Aging Neurosci* 2022; 14: 1065904.

265 Doran S, Arif M, Lam S, *et al.* Multi-omics approaches for revealing the complexity of cardiovascular disease. *Brief Bioinform* 2021; 22: bbab061.

266 Colombo SAP, Brown SL, Hepworth MR, *et al.* Comparative phenotype of circulating *versus* tissue immune cells in human lung and blood compartments during health and disease. *Discov Immunol* 2023; 2: kyad009.

267 Kim RY, Sunkara KP, Bracke KR, *et al.* A microRNA-21-mediated SATB1/S100A9/NF-κB axis promotes chronic obstructive pulmonary disease pathogenesis. *Sci Transl Med* 2021; 13: eaav7223.

268 Uhl FE, Vierkotten S, Wagner DE, *et al.* Preclinical validation and imaging of Wnt-induced repair in human 3D lung tissue cultures. *Eur Respir J* 2015; 46: 1150–1166.

269 Conlon TM, John-Schuster G, Heide D, *et al.* Inhibition of LTβR signalling activates WNT-induced regeneration in lung. *Nature* 2020; 588: 151–156.

270 Wu X, Bos IST, Conlon TM, *et al.* A transcriptomics-guided drug target discovery strategy identifies receptor ligands for lung regeneration. *Sci Adv* 2022; 8: eabj9949.

271 Costa R, Wagner DE, Doryab A, *et al.* A drug screen with approved compounds identifies amlexanox as a novel Wnt/β-catenin activator inducing lung epithelial organoid formation. *Br J Pharmacol* 2021; 178: 4026–4041.

272 Alsafadi HN, Staab-Weijnitz CA, Lehmann M, *et al.* An *ex vivo* model to induce early fibrosis-like changes in human precision-cut lung slices. *Am J Physiol Lung Cell Mol Physiol* 2017; 312: L896–L902.

273 Lehmann M, Buhl L, Alsafadi HN, *et al.* Differential effects of nintedanib and pirfenidone on lung alveolar epithelial cell function in *ex vivo* murine and human lung tissue cultures of pulmonary fibrosis. *Respir Res* 2018; 19: 175.

274 Lang NJ, Gote-Schniering J, Porras-Gonzalez D, *et al. Ex vivo* tissue perturbations coupled to single-cell RNA-seq reveal multilineage cell circuit dynamics in human lung fibrogenesis. *Sci Transl Med* 2023; 15: eadh0908.

275 Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. *Genome Biol* 2017; 18: 83.

276 Amarasinghe SL, Su S, Dong X, *et al.* Opportunities and challenges in long-read sequencing data analysis. *Genome Biol* 2020; 21: 30.

277 Dar MA, Arafah A, Bhat KA, *et al.* Multiomics technologies: role in disease biomarker discoveries and therapeutics. *Brief Funct Genomics* 2023; 22: 76–96.

278 Bottini S, Emmert-Streib F, Franco L. Editorial: AI and multi-omics for rare diseases: challenges, advances and perspectives. *Front Mol Biosci* 2021; 8: 719978.

279 Yugi K, Kubota H, Hatano A, *et al.* Trans-omics: how to reconstruct biochemical networks across multiple 'omic' layers. *Trends Biotechnol* 2016; 34: 276–290.

280 Scheibner J, Ienca M, Kechagia S, *et al.* Data protection and ethics requirements for multisite research with health data: a comparative examination of legislative governance frameworks and the role of data protection technologies. *J Law Biosci* 2020; 7: lsaa010.

281 Warnat-Herresthal S, Schultze H, Shastry KL, *et al.* Swarm Learning for decentralized and confidential clinical machine learning. *Nature* 2021; 594: 265–270.

282 Wilson NG, Hernandez-Leyva A, Schwartz DJ, *et al*. The gut metagenome harbors metabolic and antibiotic resistance signatures of moderate-to-severe asthma. *FEMS Microbes* 2024; 5: 517.