Computational and Structural Biotechnology Journal 21 (2023) 4933–4943

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Mini-review

The application of multi-omics in the respiratory microbiome: Progresses, challenges and promises

Jingyuan Gao, Xinzhu Yi, Zhang Wang *

Institute of Ecological Sciences, School of Life Sciences, South China Normal University, Guangzhou, Guangdong Province, China

ARTICLE INFO

Keywords:
Multi-Omics
Respiratory microbiome
Microbiome-host interaction
Respiratory diseases

ABSTRACT

The study of the respiratory microbiome has entered a multi-omic era. Through integrating different omic data types such as metagenome, metatranscriptome, metaproteome, metabolome, culturome and radiome surveyed from respiratory specimens, holistic insights can be gained on the lung microbiome and its interaction with host immunity and inflammation in respiratory diseases. The power of multi-omics have moved the field forward from associative assessment of microbiome alterations to causative understanding of the lung microbiome in the pathogenesis of chronic, acute and other types of respiratory diseases. However, the application of multi-omics in respiratory microbiome remains with unique challenges from sample processing, data integration, and downstream validation. In this review, we first introduce the respiratory sample types and omic data types applicable to studying the respiratory microbiome. We next describe approaches for multi-omic integration, focusing on dimensionality reduction, multi-omic association and prediction. We then summarize progresses in the application of multi-omics to studying the microbiome in respiratory diseases. We finally discuss current challenges and share our thoughts on future promises in the field.

1. Introduction

The human lungs harbor a consortium of microorganisms including bacteria, fungi, and viruses that collectively are known as the respiratory or lung microbiome [1,2]. The lung microbiome alters in respiratory diseases and could be involved in disease pathogenesis, and has received increasing attentions over the past decade [3,4]. Sequencing of bacterial 16S rRNA gene and fungal internal transcribed spacer (ITS) allows the taxonomic classification and is routinely used to assess the composition of the bacterial and fungal components of the lung microbiome. Shotgun metagenomic sequencing provides a more in depth characterization of microbial composition to the species and possibly strain levels [5], with unique insights into microbial functional capabilities. The advancement of next-generation sequencing and mass spectrometry technology has enables the characterization of additional multi-omic features along the axis of microbial-host interaction, including metatranscriptome, metaproteome, and metabolome, together enabling a holistic view of microbiome metabolic activities and its interaction with host activities [6]. Novel, mathematical-based approaches have been applied to studying the microbiome, with a clinical and translational focus [7]. The emergence of additional, clinical relevant omic data types such as radiome has allowed further insights into the clinical manifestation and pathophysiology of respiratory diseases [8]. The study of the respiratory microbiome has reached a multi-omic era. However, the advent of the multi-omics, with the generation of massive and complex datasets, is met with unique challenges in particular in the capability of multi-omic data integration to understand microbial-host interaction and disease biology. Here, we review the current progress of multi-omics in studying the microbiome in respiratory diseases (Fig. 1). We begin by introducing the respiratory sample types and the omic data types relevant for respiratory microbiome studies. We then summarized the current state-of-the-art methods for integrating multi-omic data, their advances and limitations, followed by a description on the applications of the multi-omics to chronic, acute and other types of respiratory diseases. We finally share our views on current challenges and future promises for the use of multi-omics toward a more in-depth and comprehensive view of the microbiome in respiratory diseases.

2. The respiratory sample types for multi-omics

In comparison to the gut microbiome where fecal samples have been routinely investigated, there is no 'gold standard' specimen for sampling

* Corresponding author.
E-mail address: wangz@m.scnu.edu.cn (Z. Wang).

https://doi.org/10.1016/j.csbj.2023.10.016
Received 13 June 2023; Received in revised form 10 October 2023; Accepted 10 October 2023
Available online 12 October 2023

the lung microbiome. Naso- and oropharyngeal swabs are convenient ways in sampling the upper respiratory tract, while sampling the lower respiratory tract can be more difficult. Sputum has been commonly used as a proxy to study the lung microbiome. As a non-invasive approach, sputum has its unique advantage in its ability for large-scale and serial sampling of the lung microbiome, in particular for healthy individuals. For individuals unable to produce sputum spontaneously, sputum induction can be performed using nebulized saline, which is a clinically safe and routine procedure [9]. However, sputum represents an inherent admixture of upper, lower airways and oral materials, and its distinction from the lower airways or the lung should always be considered [10]. Additional more invasive sampling approaches, such as bronchoalveolar lavage (BAL), bronchial brushing, and tracheal aspirate, can be performed for patients with such clinical needs. Essentially, these sample types should more accurately represent the lower airway environment with less contaminations from upper airways and oral cavity. Lung tissue is generally impractical to obtain unless surgically justified (lung resection or transplantation), but can provide the possibly most direct insight into the local environment with unique merit in capturing topographic distribution of the microbiome [11].

3. The multi-omics for the lung microbiome

3.1. Amplicon-based microbiome

Amplicon-based sequencing has been widely applied in characterizing the bacterial and fungal members of the lung microbiome, by targeting 16S rRNA gene or ITS region, respectively. The 16S rRNA gene comprises nine hypervariable regions (V1-V9) and one or a few of them were often sequenced by next-generation sequencing, enabling the characterization of bacterial taxa in general to the genus level. The full-length 16S rRNA gene sequencing using long-read sequencing in PacBio or Nanopore platform has revealed additional variations at the species or strain level in the lung microbiome that is otherwise unseen by regular short-read sequencing [12,13]. The sequence of V4 or V3V4 region has been the most common choice for lung microbiome studies, although a recent analysis by us suggests that the V4 region could be sub-optimal in capturing the diversity of the airway bacterial community [13]. For fungal community, often a portion of ITS1 or ITS2 sequence was characterized. By amplification of the ITS1–5.8S-ITS2 region followed by fragmentation and sequencing using a metagenomic-like approach, Mac Aogain et al. obtained the full-length ITS sequences from sputum samples, which led to the identification of *Aspergillus* species in association with clinical characteristics in bronchiectasis [14]. In light of the unique characteristics of the microbiome data (i.e. compositionality, hierarchical taxonomic classification) compared with the other omic data types, specialized algorithms and approaches have been developed to analyzing the microbiome data (i.e. Bray-Curtis dissimilarity matrix for ordination, LEfSe for differential analysis, SparCC for network analysis, etc).

3.2. Metagenome

Metagenomic sequencing, by shotgun sequencing of the DNA fragments in the microbial community, has its advantage in in-depth taxonomic characterization as well as identification of microbial functional genes, thereby enabling the answering of not only 'who are there' but also 'what they can do' [15]. However, the application of metagenomics

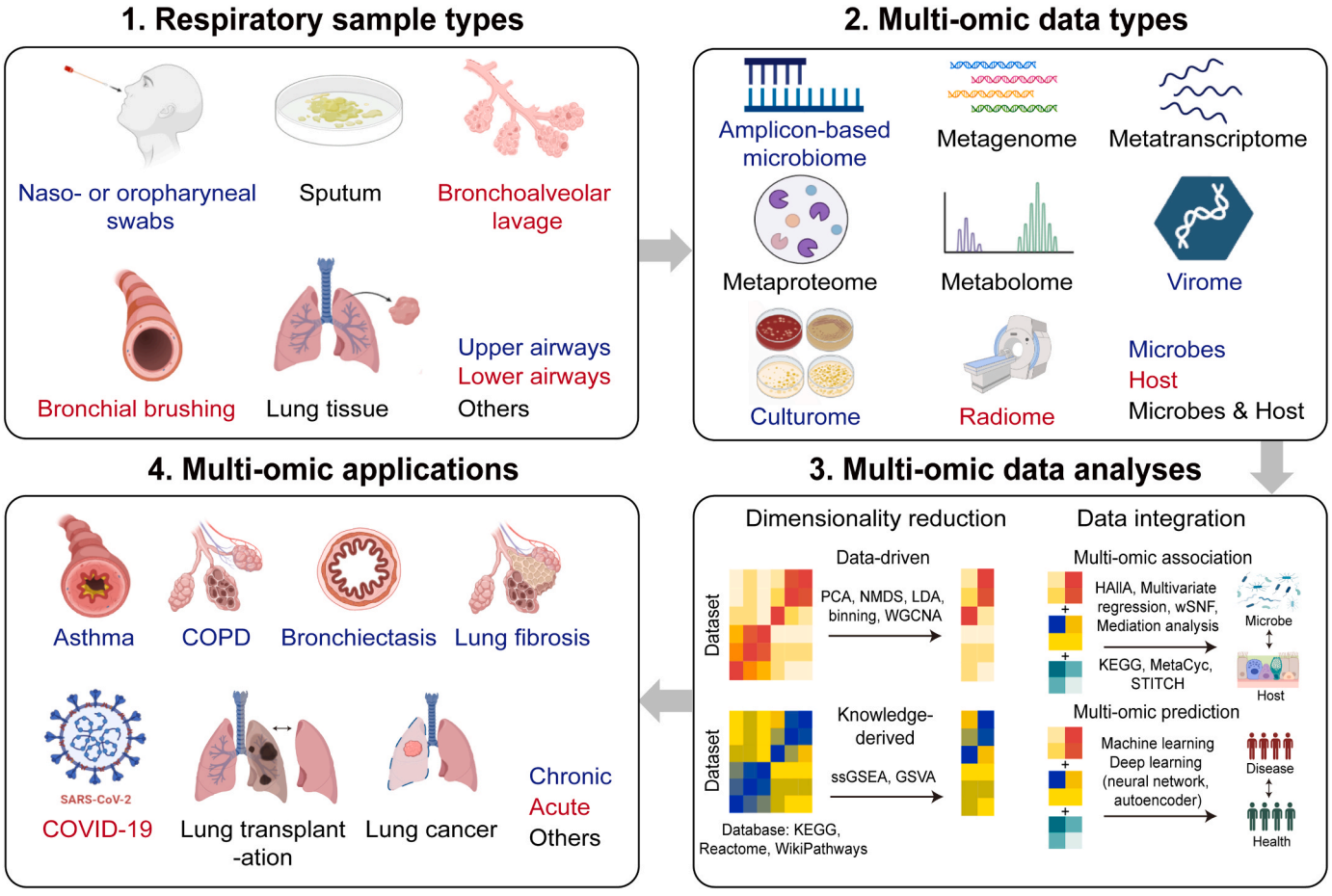

Fig. 1. A summary for the use of multi-omics in studying respiratory microbiome, including 1) respiratory sample types, 2) multi-omic data types, 3) multi-omic data analyses, and 4) multi-omic applications to respiratory diseases.

in the lung microbiome remains sporadic, which is largely limited by the low microbial biomass and high host-to-microbe ratio in most airway samples. As host sequences often comprise the bulk of the metagenomic reads for airway samples (~90%), an ultra-deep sequencing is necessary to achieve sufficient microbial coverage. Host cell depletion prior to sequencing has been applied in airway samples but with varied performance and risk of simultaneous modification of the microbial communities [16–18].

3.3. Metatranscriptome

Metatranscriptome has been employed to sequencing the total RNA in the microbial community, thereby moving one step further from 'what they can do' to 'what they are doing' [19]. Metatranscriptome provides a more refined landscape on the microbiome by profiling transcriptionally active microbial taxa and functions. However, as RNA can degrade more rapidly than DNA, airway sample collection and preservation approaches could be crucial for adequate metatranscriptomic readouts. Metatranscriptomics suffers from the same limitation as metagenomics in terms of the predominance of host reads. However, it could have the unique benefit in simultaneously characterizing gene expression for both microbiome and host, which can provide valuable information for biologically active microbial-host interaction [20]. Sulaiman et al. characterized the functional features of the airway microbiome through metatranscriptome and showed that it efficiently captured transient active microbial metabolism [21]. The same team further analyzed the airway microbiome in COVID-19 patients through combined metagenomics and metatranscriptomics and found specific microbiome and host transcriptome profiles as associated with poor clinical outcome [22]. In a recent study, they further investigated airway microbiome in chronic obstructive pulmonary disease (COPD) through 16S rRNA gene sequencing, whole genome, RNA metatranscriptome and host RNA transcriptome, and showed that lower airway dysbiosis contributes to inflammatory injury early in COPD [23].

3.4. Metaproteome

Downstream to transcription, protein biosynthesis is essential for the functionality of microbiome and host, which could be characterized by metaproteomic approaches via a range of techniques such as two-dimensional gene electrophoresis, liquid chromatography mass spectrometry, and antibody/protein microarrays [24,25]. In principle, metaproteome can be employed to simultaneously determine the protein composition of microbiome and host. However, in practice, this can be challenging due to the high microbial species diversity and high sequence homology at the protein level in particular between closely related species [26]. The lack of appropriate databases for proteome analysis aggravates the challenge. Due to these technical limitations, metaproteome has been applied only on a limited basis to the lung microbiome studies [27]. However, microarray-based assays (i.e. SOMAscan) have been employed to characterize a set of host proteins to understand microbiome-host interactions in respiratory microbiome studies [28–30]. Employing microbiome and host proteome, Dicker et al. identified associations between Proteobacteria dominance and neutrophil activation in COPD sputum [31]. Keir et al. identified neutrophil extracellular trap (NET) as a key marker of bronchiectasis severity and treatment response [32]. Hull et al. identified protein markers associated with severity and treatment response in non-tuberculous mycobacterial lung disease [33].

3.5. Metabolome

The airway microbiome communicates with host through producing metabolites that interact with host ligands and influence downstream signaling processes [34]. The assessment of this process requires the measurement of metabolites in biological samples through metabolomics. Nuclear magnetic resonance and high-resolution mass spectrometry are two main analytical techniques for metabolomics, with the latter coupled with liquid or gas chromatography. Given the central role of metabolites in bridging the activities of microbiome and host, metabolomics have been increasingly applied to the lung microbiome studies in recent years [29,35,36], where numerous microbial metabolites were identified to interact with host immunity and influence disease phenotypes. Of note, it is generally difficult to differentiate the exact microbial or host origin of the metabolites if they can be produced by both microbiome and host. However, prediction of microbial origin of metabolites may be possible via integrating microbiome and metabolome data through bioinformatic approaches [37,38].

3.6. Virome

In addition to bacteria and fungi, viruses are a crucial player in respiratory health and diseases. Although specific viral pathogens that cause clinical lung infection have been well characterized through PCR-based approaches, much less is understood for the composition of commensal virus communities or the virome in the lung [39]. This is largely because of the small proportion of viruses compared to bacteria and host cells in the lung, the difficulty in isolating viral particles from airway specimens, and the instability of viral DNA or RNA after sample collection. Viral reads can be identified from metagenomics or metatranscriptomics; however, the proportion of viral reads from these data is remarkably low and can be insufficient to achieve adequate viral coverage. By using a viral purification and enrichment approach, recent studies have begun to uncover the virome in the airways. Li et al. characterized the respiratory virome associated with recurrent acute respiratory tract infections (ARTIs) in children and found specific bacteriophages could be a predictor for ARTIs [40]. Choi et al. characterized the virome profiles in sputum from asthma and healthy individuals, and found that viral taxa had greater associations with asthma severity and exacerbations [41]. Mac Aogain et al. characterized the virome in bronchiectasis patients and found that the integrated multi-biome, including bacterial, fungal and viral communities, better associated with key clinical characteristics than a single microbial group [42].

3.7. Culturome

Although the lung bacterial taxonomy and functions can be well characterized through sequencing-based approaches, being able to obtain pure bacterial culture is essential for downstream functional and mechanistic studies [43]. Currently it remains incompletely clear how much proportion of the airway microbes can be cultured. Using a combination of multiple culture conditions and media, Whelan et al. showed that an average of 82.13% of bacterial OTUs from sputum of cystic fibrosis patients can be cultured [44]. Through a comprehensive cultivation effort, Muggeo et al. found an enrichment of *Enterobacterales* was associated with deteriorated clinical symptom in COPD patients [45]. Sun et al. have characterized the lung and oral microbiota using culturomics and 16S rRNA gene sequencing and found site and pathology-dependent alterations in patients with lung cancer [46]. Compared with gut microbiome for which culturomics have been extensively applied, there remains a lack of specific media and conditions optimized for the cultivation of microbes in the airways.

3.8. Radiome

Radiological changes (i.e. in computed tomography or CT) are a prominent manifestation or diagnostic criteria for many lung diseases such as COPD, bronchiectasis, and lung cancer [47]. The radiological features are potentially associated with the alteration of the lung ecology. As a high-throughput quantitative imaging approach, radiomics aim to extract high-dimensional features from medical images in CT scans, thereby yielding data in similar settings to other omics data [48].

Zhou et al. created a radiogenomic map linking radiological features with transcriptomic data, and identified non-invasive features associated with molecular pathways in patients with non-small cell lung cancer [49]. By integrating radiomic features and sputum bacterial and fungal microbiota, Wang et al. identified microbial taxa that are associated with key radiological features including emphysema and airway structural lesions in COPD, revealing radiological markers that can potentially reflect changes in airway microbiome [50].

4. Strategies for integrative multi-omic analyses

A central question for the multi-omic studies is how to integrate the multi-omic datasets to answer the biological or clinical questions regarding the microbiome, i.e. to test a biological hypothesis for microbial-host interaction, or to develop a clinical diagnostic model with the identification of disease biomarkers. The high-dimensionality of the multi-omic data poses a critical challenge to identify the significant associations between different omic data, as direct correlations between hundreds to thousands of omic features are often prone to false-positive observations. It is therefore paramount that the massive multi-omic associations be adjusted for multiple comparison (i.e. using Bonferroni or Benjamin-Hochberg procedure). A grand sample size is often required to achieve sufficient statistical power, if we aim to identify significant associations directly from high dimensional multi-omic data after multiple comparison adjustment [51]. Therefore, dimensionality reduction (DR) is often an important first step for a statistically rigorous analysis of the multi-omic data. Here we introduce the commonly used approaches for multi-omic dimensionality reduction, which are mainly categorized into data-driven clustering methods or methods utilizing prior biological knowledge (Table 1).

4.1. Multi-omic dimensionality reduction

4.1.1. Data-driven approaches

Unsupervised approaches were often employed to achieve data-driven DR. For instance, for microbiome and host multi-omic data, principle component analysis (PCA, or principle coordinate analysis [PCoA] for microbiome data) can be performed to reduce the dimension space from features to a couple of principle components (PCs) that capture the majority of data variance. Other DR approaches such as non-metric multidimensional scaling (NMDS) and linear discriminant analysis (LDA) can be applied for similar purpose. For metabolomic and transcriptomic data, co-abundance clustering approaches such as weighted correlation network analysis (WGCNA) can be performed to collapse the metabolites and host genes into a set of co-abundant clusters [52]. For microbiome, using algorithms that combine information such as sequence composition and coverage, metagenomic species or genomic bins can be constructed from the assembled metagenomic data, essentially reducing the dimensionality from millions of microbial genes into bacterial genomes [53]. Different combinations of DR approaches can be applied to the various omic data types, according to the biological questions and statistical properties of the data at hand.

Table 1
Examples for approaches, tools, databases and applications for multi-omic dimensionality reduction, including data-driven and knowledge-derived dimensionality reduction.

Multi-omic data analysis	Approaches/ Tools	Databases	Applications
Data-driven dimensionality reduction	PCA, PCoA, NMDS, LDA, WGCNA[52], metagenomic binning (i.e. MetaWRAP [113])	Not available	NMDS and LDA in microbiome[51], metagenomic binning in microbiome[29, 114], WGCNA in metabolome[29, 115], PCA and LDA in transcriptome [51], WGCNA in transcriptome [29], PCA in transcriptome and proteome[28]
Knowledge-derived dimensionality reduction	Single-sample GSEA (ssGSEA) [62], GSVA[63]	KEGG[54], CAZy [55], CARD[56], VFDB[57], mobileOG-db[58], environmental allergens[59], Reactome[60], WikiPathways[61]	ssGSEA in microbiome[29], ssGSEA and GSVA in metabolome [116], ssGSEA in transcriptome [117], GSVA in transcriptome [118,119], GSVA in proteome[120]

4.1.2. Knowledge-derived approaches

In comparison to the unsupervised data-driven DR, knowledge-derived DR can be achieved through annotation of the omic data according to existing domain knowledge for the specific type of data in a supervised manner. For instance, the massive microbial genes in the metagenomic data can be annotated to gene families in KEGG database (KEGG orthologs or KOs) [54], which can be further aggregated to the module or pathway level. Other more specialized databases focusing on a specific type of functions can also be employed, such as CAZy for carbohydrate metabolism [55], CARD for antibiotic resistant genes [56], and VFDB for bacterial virulence factors [57], mobileOG-db for mobile genetic elements [58], and sequences for environmental allergens [59]. For human transcriptome and proteome, gene sets can be obtained according to the pathway information from additional databases such as KEGG [54], Reactome [60] and WikiPathways [61], and can be utilized to obtain gene-set-level expression profiles using approaches such as single-sample gene set enrichment analysis (GSEA) [62] or gene set variation analysis (GSVA) [63]. Custom gene sets can also be established through literature mining. While the main advantage of data-driven DR lies in its capability in capturing the variance of the original high dimensional data, the benefit of knowledge-derived DR is the biological interpretability of the dimensionality reduced data. In practice, both types of approaches can complement each other to achieve DR of multi-omic data.

4.2. Multi-omic data integration

Following DR, the dimension-reduced multi-omic data can be further analyzed and integrated to tackle specific biological or clinical questions (Table 2). For example, statistical analysis (i.e. Wilcoxon rank-sum test, general linear model) can be performed on each type of dimension-reduced omic data features, to identify differentially abundant omic features between disease and control groups and specific disease subgroups of interest. Correlation analysis (i.e. Spearman's rank correlation) can also be performed between dimension-reduced multi-omic features and specific demographic and clinical features with a continuous distribution. In regard to these analyses, the dimension reduction provides the unique benefit in achieving adequate statistical power by reducing the number of features subject to the penalty of statistical correction for multiple hypothesis testing [51].

In addition to single-omic association, multi-omic integration can be further performed to generate potential hypotheses of microbial-host interactions. Correlation-based univariate analysis can be performed for paired omic features. Multivariate regression models can be established to assess combinatory associations of multiple omic features of interest. Population confounders such as age and gender often need to be accounted for in the multivariate model. Alternatively, a regression can be initially performed between each omic feature and the confounders, to obtain the feature residues for downstream analyses [64]. Other more sophisticated approaches, such as mediation analysis, can be employed

Table 2
Examples for approaches, tools, databases and applications for multi-omic integration, including multi-omic association and multi-omic prediction.

Multi-omic integration	Approaches/Tools	Databases	Applications
Multi-omic association	Correlation analysis (Pearson correlation, Spearman's rank correlation, HAllA[67]), Multivariate regression (mixOmics[70], MaAsLin2[121]), Sequential mediation analysis[29], Network-based approaches (wSNF[42], CoNet[122], SPEIC-EASI [123], SparCC[124])	KEGG[54], MetaCyc[65], STITCH[66]	Multi-omic association with clinical characteristics[36,64], Microbiome-host interactions[28,29,36,51], Microbial-microbial interactions[42,125]
Multi-omic prediction	Machine learning approaches, including random forest (R caret [126], randomforest[127] packages), support vector machine (R caret package[126]), etc., Deep learning approaches, including neural network (MOGONET[128], AutoGGN[129]), auto-encoder (MOVE[130], AIME[131], MAE[132]), etc, Network-based approaches (wSNF[42]), Cox proportional hazards regression (R survival package[133])	Not available	Disease diagnosis[96], Prediction of disease progression[36,106], response to therapies[134,135], and prognosis[101,103,136], Disease phenotyping[30,42,84,105,137], Biomarker identification [30,42,85,137]

for three or more types of omic data, to explore potential causal associations between them. We recently proposed a sequential mediation analysis approach, by assessing the multi-omic associations in a stepwise manner along the microbial-host axis, from metagenome, metabolome, to host transcriptome and proteome [29]. The associated multi-omic modules or features can be further mined to identify microbiome-metabolite-host interactions, leveraging knowledge from existing databases (i.e. KEGG [54], MetaCyc [65], STITCH [66]). As direct correlation of the massive features in the omic data suffers from high penalty from multiple comparison adjustment of P-values, a hierarchical framework named HAllA has been developed for structured association testing, which effectively increases statistical power for associations between paired high-dimensional datasets [67].

In addition, multi-omic prediction can be performed to identify biomarkers for disease diagnosis and phenotyping using machine learning approaches (i.e. random forest, support vector machine). The cutting-edge deep learning algorithms such as neural networks and auto-encoders, which better account for heterogeneity and high dimensionality of the multi-omic data, have also been proposed for multi-omic biomarker discovery [68,69]. Some of the above mentioned methods have been implemented in an R package called mixOmics, which provides a wide range of multivariate methods for multi-omic integration [70]. In addition, network-based approaches such as weighted similarity network fusion (wSNF) have also been applied in respiratory microbiome studies for patient stratification and biomarker identification [42,71,72].

5. Application of multi-omics to microbiome studies in respiratory diseases

With its unique power in enabling a holistic perspective for the microbiome, host and their interactions, integrated multi-omics have been increasingly applied in microbiome studies in respiratory health and diseases. Here we summarize the current literature in application of multi-omics in studying the respiratory microbiome, focusing on chronic, acute and other types of lung diseases (Table 3).

5.1. Asthma

Asthma is a chronic respiratory disease characterized by wheezing, shortness of breath, chest tightness, and cough, affecting individuals of all ages [73]. The lung microbiome alters in asthma [74–76], and multi-omics have been applied to characterizing the lung microbial-microbial, and microbial-host interactions in relation to asthma phenotypes [77,78]. Chiu et al. identified specific connections between airway microbiome and circulating metabolites in mite-sensitized pediatric asthmatic children through an integrated analysis of airway metagenome and serum metabolome [79]. Through a co-profiling of bacterial and fungal microbiota in paired endobronchial brush and BAL samples from 39 asthma patients and 19 healthy controls, Sharma et al. characterized the multi-kingdom microbiota in asthma associated phenotypes, where *Aspergillus* in BAL was associated with T2-high and lung function, and *Penicillium* in EB was associated with atopy. A distinct inter-relationship between bacterial and fungal microbiome was observed for different sample types [80]. By characterizing the sputum microbiome, metabolome (eicosanoids), transcriptome, and proteome in 97 severe, 46 mild and moderate asthma patients and 47 healthy controls, Abdel-Aziz et al. identified two clusters of patients defined by the microbiota with distinct multi-omic profiles [30]. In particular, the microbiome-driven cluster 2 had reduced microbial diversity, enrichment of *Haemophilus influenzae* and *Moraxella catarrhalis*, elevated levels of 11-dehydrothromboxane B2 and prostaglandin E2, downregulated pathways for cell growth, proliferation, metabolism and DNA repair, and worse clinical outcomes, suggesting that this cluster might be amenable to microbiome-targeted therapy [30]. Multi-omics have been applied to characterize the molecular signatures of childhood asthma [81,82]. In an integrative analysis of the nasal and bronchial microbiome and transcriptome in 27 healthy and 27 asthmatic children, Chun et al. identified site-specific microbial features associated with genes in ciliary function and inflammation, providing a window for assessing host-microbiome associations in childhood asthma [83]. Raita et al. identified four distinct biological endotypes for 221 infants with respiratory syncytial virus bronchiolitis through airway microbiome, transcriptome, and metabolome, with one specific subtype showing a high risk for developing asthma [84]. Wang et al. systematically benchmarked the methods in combining six omic data type (GWAS, miRNA, mRNA, microbiome, metabolome, DNA methylation) of 748 child participants, and showed that specific omic combinations can reach the optimal prediction of childhood asthma development [85].

5.2. COPD

COPD is chronic lung disease characterized by irreversible lung function decline, airway inflammation and emphysema, and is a leading cause of morbidity and mortality worldwide [86,87]. An altered airway microbiome is implicated in COPD [13,28,76], associates with its phenotypes and endotypes [88–90], and could mechanistically contribute to host inflammation and disease progression [29,36,91]. Using an approach of multi-omic meta-analysis that integrates publicly available airway microbiome and host transcriptomic data, we previously established potential 'microbiome-metabolite-host' interaction links for COPD and found that about 69.9% of these links can be validated in a pilot cohort. Building upon this analysis, we conducted a more detailed multi-omic characterization on COPD patients and healthy individuals. A combination of data-driven (i.e. WGCNA) and knowledge-derived (i.e. ssGSEA) dimensionality reduction approaches were applied to the multi-omic data. Sequential mediation analysis were then employed to identify microbiome-host interaction from the multi-omic data. Together with in vivo animal models and in vitro cellular assays, we

Table 3
Summary of key studies on the application of multi-omics to studying the respiratory microbiome in chronic, acute, and other types of lung diseases.

Disease	Study	Specimens	Design and sample size	Omic data types	Key findings
Asthma[a]	Sharma et al. 2019 [80]	BAL, endobronchial brush	39 asthma patients, 19 healthy controls	16 S rRNA gene microbiome, ITS microbiome	• Fungal diversity decreased in asthma patients with T2-high inflammation. • Clear differences in bacterial and fungal microbioa in asthma-associated phenotypes.
Asthma[a]	Abdel-Aziz et al. 2022[30]	Sputum	97 severe, 46 mild asthma patients, 47 healthy controls	16 S rRNA gene microbiome, eicosanoids, transcriptome, proteome	• Two patient clusters were identified based on sputum microbiome (C1 and C2). • C2 had airway dysbiosis, increased neutrophilia, downregulated cell growth and proliferation pathways, and worse clinical outcomes.
Asthma[a]	Chun et al. 2020 [83]	Bronchoscopy, nasal brushing	27 asthma patients, 27 healthy controls	16 S rRNA gene microbiome, transcriptome	• *Moraxella* and *Alloiococcus* are hub genera in the nasal microbiome of asthmatic children. • Asthmatic children express more nasal genes for ciliary function with more nasal *Streptococcus*.
COPD[a]	Wang et al. 2020 [138]	Sputum	1666 microbiome samples,1340 human transcriptome samples	16 S rRNA gene microbiome, metagenome, transcriptome	• 29.6% of differentially expressed human pathways were predicted to be targeted by microbial metabolites in COPD. • Butyrate, homocysteine, and palmitate were the microbial metabolites showing strongest interactions with COPD-associated host genes.
COPD[a]	Yan et al. 2022[29]	Sputum	99 COPD patients, 36 healthy controls	Metagenome, transcriptome, proteome, metabolome	• Airway microbiome-derived IAA mitigates neutrophilic inflammation, apoptosis, emphysema and lung function decline. • Intranasal inoculation of two airway lactobacilli restored IAA and recapitulated its protective effects in mice.
COPD[a]	Liang et al. 2023 [36]	Sputum	181 COPD patients (UK), 61 COPD patients (Guangazhou, China)	Metagenome, metabolome, transcriptome	• Airway dysbiosis at baseline associates with accelerated lung function decline in COPD patients. • *Staphylococcus aureus* promotes lung function decline through homocysteine-AKT1-S100A8/A9 axis.
COPD[a]	Madapoosi et al. 2022[92]	BAL	137 COPD patients	16 S rRNA gene microbiome, metabolome	• Lower lung function and COPD diagnosis associated positively with *Streptococcus, Neisseria,* and *Veillonella.* • *Prevotella*, together with metabolites such as sialic acid and glutathione, associated with better lung function or less symptoms.
COPD[a]	Dicker et al. 2021 [31]	Sputum	253 COPD patients	16 S rRNA gene microbiome, proteome	• Proteobacteria predominance and lower diversity associated with increased COPD severity and mortality. • A significant correlation was found between *Haemophilus* and neutrophil activation pathway in sputum.
COPD[a]	Sulaiman et al. 2023[23]	BAL	26 COPD patients, 31 smoker controls	16 S rRNA gene sequencing, metagenome, metatranscriptome, host transcriptome	• COPD lower airways were enriched with common oral commensals with differences in markers of inflammation and tumorigenesis. • Lower airway dysbiosis augments the inflammatory injury in a COPD pre-clinical murine model.
Bronchiectasis[a]	Mac Aogain et al. 2021[42]	Sputum	217 bronchiectasis patients	16 S rRNA gene microbiome, ITS microbiome, metagenome	• Patients at greatest risk of exacerbation have less complex microbial co-occurrence networks. • *Pseudomonas* interactome networks, rather than abundance alone, are associated with exacerbation risk.
Bronchiectasis[a]	Li et al. 2022[35]	Sputum, lung tissue (mouse), serum (mouse)	225 bronchiectasis patients	16 S rRNA gene microbiome, metagenome, transcriptome, metabolipidome	• Bronchiectasis bacteriomes defined by the presence of *Neisseria* spp. associate with poor clinical outcomes. • The culturable species *N. subflava* weakens barrier integrity and induces inflammation.
Bronchiectasis[a]	Narayana et al. 2023[95]	Sputum, stool	57 bronchiectasis patients	16 S rRNA gene microbiome, ITS microbiome, metagenome	• Microbial communities in stable bronchiectasis demonstrate a significant gut-lung interaction. • A high gut-lung interaction cluster is associated with increased exacerbations and greater disease severity.
Lung fibrosis[a]	Molyneaux et al. 2017[96]	BAL, blood	60 IPF patients, 20 healthy controls	16 S rRNA gene microbiome, transcriptome	• Two gene modules strongly associated with IPF, BAL bacterial burden, and specific microbial taxa and neutrophilia. • The modules involve genes in host defense response and antimicrobial peptides.
Lung fibrosis[a]	Huang et al. 2017 [97]	BAL, blood	68 IPF patients	16 S rRNA gene microbiome, transcriptome	• Down-regulation of immune response pathways was associated with worse progression-free survival.

(*continued on next page*)

Table 3 (*continued*)

Disease	Study	Specimens	Design and sample size	Omic data types	Key findings
Lung fibrosis[a]	O'Dwyer et al. 2019[98]	BALF	68 IPF patients	16 S rRNA gene microbiome, Proteome (cytokine measurements)	• Increased abundance of *Streptococcus* correlated with increased NOD-like receptor signaling. • Disruption of the lung microbiome predicts disease progression, and correlates with local host inflammation. • Lung dysbiosis precedes peak lung injury and is persistent.
COVID-19[b]	Sulaiman et al. 2021[101]	Bronchoscopy	142 COVID-19 patients	Metagenome, metatranscriptome	• Poor clinical outcome was associated with lower airway enrichment with *Mycoplasma salivarium*. • Increased SARS-CoV-2 abundance and a distinct airway transcriptomic profile is predictive of mortality.
COVID-19[b]	Zhong et al. 2021 [102]	Sputum, nasal swab, throat swab, anal swab and feces	23 COVID-19 patients	Metatranscriptome	• Co-detection of other human respiratory viruses was demonstrated in 30.8% of the severely ill patients. • The predominant respiratory microbial taxa of severely ill patients are *Burkholderia cepacia complex, Staphylococcus epidermidis* and *Mycoplasma* spp.
COVID-19[b]	Ren et al. 2021 [103]	Oropharyngeal swab	192 COVID-19 patients, 95 healthy controls	Metatranscriptome	• The upper airway microbiota in patients with COVID-19 differed from that in healthy controls. • The abundance of *S. parasanguinis* on admission was correlated with prognosis in nonsevere patients.
COVID-19[b]	Lloréns-Rico et al. 2021	Nasopharyngeal swabs, BAL	58 patients (upper airway), 35 patients (lower airway)	16 S rRNA gene microbiome, single-cell transcriptome	• Patient confounders such as ICU stay and type of oxygen support could explain the microbiome variation. • Mechanistic ventilation is linked to altered microbiome and shifts in COVID-19-associated oral taxa.
Lung transplantation[c]	Das et al. 2021 [105]	BAL	64 lung transplant recipients	16 S rRNA gene microbiome, culturome	• The lung microbiota post-transplant can be categorized into four distinct 'pneumotypes'. • The predominant pneumotype is characterized by a diverse bacterial community, moderate viral loads, and host gene expression in immune tolerance.
Lung transplantation[c]	Watzenbock et al. 2022[106]	BAL	78 lung transplant recipients and donors	16 S rRNA gene microbiome, metabolome and lipidome	• Recipient-specific and environmental factors, rather than donor microbiome, shpe the long-term recipient lung microbiome. • Multi-omic data, in particular microbial profiles can predict future changes in FEV1.
Lung cancer[c]	Greathouse et al. 2018[109]	Lung tissue	143 lung cancer patients, 33 controls	16 S rRNA gene microbiome, transcriptome	• A lower alpha diversity in normal lung as compared to non-tumor adjacent or tumor tissue. • *Acidovorax* exhibit higher abundance among the subset of squamous cell carcinoma cases with TP53 mutations.
Lung cancer[c]	Tsay et al. 2021 [110]	Bronchoscopy	83 lung cancer patients	16 S rRNA gene microbiome, transcriptome	• *Veillonella parvula* drives the association between lower airway dysbiosis and upregulation of the IL17, PI3K, MAPK,and ERK pathways. • *V. parvula* led to decreased survival, increased tumor burden, IL-17 inflammation, and activation of checkpoint inhibitor markers in a lung cancer mouse model.

[a] Chronic lung diseases,
[b] Acute lung diseases,
[c] Other lung diseases

showed that airway lactobacilli may ameliorate COPD inflammation and epithelial apoptosis through producing indole-acetic acid that interacts with host IL-22 signaling [29]. Through multi-continental cohorts, multi-omic analysis and murine and cellular experiments, we recently found that airway dysbiosis with enrichment of *Staphylococcus aureus* could accelerate COPD lung function decline through homocysteine-AKT1-S100A8/A9 axis [36]. Through integrated analysis of microbiome (16 S rRNA gene sequencing) and metabolomics, Madapoosi et al. identified the combination of the multi-omic features in relation to clinical outcomes in milder COPD [92]. In a recent study, Sulaiman et al. characterized the 16 S rRNA gene microbiome, metagenome, metatranscriptome and host transcriptome of 26 COPD patients and 31 smoker controls and found that COPD lower airways were enriched with oral commensals and had differences in markers of inflammation and tumorigenesis in host transcriptome [23]. In conjunction with a pre-clinical model, they showed that lower airway dysbiosis augments the inflammatory injury early in COPD. Together, these studies demonstrate the power of multi-omics in elucidating the mechanistic role of the airway microbiome in COPD.

5.3. Bronchiectasis

Bronchiectasis is another primary chronic lung disease characterized by progressive and irreversible dilation of the airways resulting from damages to the airway wall [93]. Recurrent airway microbial infection with a perturbation in the airway microbiome is implicated in bronchiectasis [94]. By using an integrative microbiomic approach, Mac Aogain et al. characterized the bacterial, viral, and fungal communities in the airways of bronchiectasis patients [42]. They present an approach to integrate these microbial multi-omic data through weighted similarity network fusion (wSNF) and found that the interactome better associated with exacerbation risk than the use of abundance alone, thereby demonstrating the power of multi-omics in capturing microbial interactions in association with clinical characteristics. Through further integrating microbiome, transcriptome, metabolome and lipidome data together with murine assays, Li et al. identified *Neisseria* species as a key airway pathobiont invoking host inflammatory responses and the loss of epithelial integrity in bronchiectasis [35]. By integrating microbial multi-omics from concurrent stool and sputum samples through wSNF, Narayana et al. further showed a dysregulated gut-lung axis, driven by lung *Pseudomonas*, was associated with a worse clinical outcome in bronchiectasis [95].

5.4. Lung fibrosis

The multi-omics have been utilized to characterize the airway dysbiosis in fibrotic lung diseases such as idiopathic pulmonary fibrosis (IPF). By performing dimensionality reduction of host transcriptome using WGCNA, followed by integration with 16 S rRNA gene-based microbiome data in 60 IPF patients and 20 controls, Molyneaux et al. identified two co-expression gene modules associated with IPF diagnosis, bacterial burden, neutrophilic inflammation and specific microbial taxa, providing evidence for a host response to the alteration of microbiota in IPF [96]. Huang et al. obtained paired microarray gene expression and microbiome data in the COMET-IPF study [97]. By using WGCNA for dimensionality reduction of the host transcriptome data, followed by a network analysis to integrate the multi-omic data, they identified host-microbiome interactions that could influence progression-free survival and fibroblast responsiveness. By examining the lung microbiota and cytokines in 68 IPF patients along with a germ-free murine model study, O'Dwyer et al. found that the airway dysbiosis was correlated with IPF symptoms, high levels of surfactant protein-D and lactate dehydrogenase in the serum and elevated alveolar profibrotic cytokines [98].

5.5. COVID-19

COVID-19, caused by the SARS-CoV-2 virus infection, has resulted in a public health emergency worldwide and remains to be a global epidemic [99]. Acute respiratory distress syndrome and respiratory failure are the main clinical manifestations associated with mortality for severe COVID-19 patients [100], where an airway dysbiosis could be implicated. Sulaiman characterized the lower airway metagenomics, metatranscriptomics, and host immune response profiling in 142 COVID-19 patients [101]. Through multi-omic integration using multi-scale embedded co-expression network analysis, followed by prediction analysis using Cox proportion hazards regression model, they showed an association between an oral commensal *Mycoplasma salivarium* and poor clinical outcome, and an increased SARS-CoV-2 abundance and a distinct airway transcriptomic profile were predictive of mortality. By a co-profiling of lung microbiome and virome through a metatranscriptomic sequencing in 27 COVID-19 patients, Zhong et al. found evidence for the enrichment of antibiotic-resistant pathogens in COVID-19 patients in relation to disease severity, which underscored the need of detection, tracking and prevention of antimicrobial resistance for patient management [102]. By an in-depth metatranscriptomic sequencing on 588 oropharyngeal swabs from 192 COVID-19 patients and 95 healthy controls, Ren et al. found the upper airway dysbiosis, in particular the loss of *S. parasanguinis*, to be associated with increased mortality, suggesting the potential use of the upper airway microbiota as a biomarker for COVID-19 prognosis [103]. By analyzing the upper and lower airway microbiome in COVID-19 patients in combination with host immunoprofiling, Lloréns-Rico et al. showed that potential clinical confounders, in particular ICU stay and type of oxygen support, may explain the variation of the microbiome in the patients [104]. Despite these progresses, a mechanistic link remains to be established between airway dysbiosis and COVID-19.

5.6. Lung transplantation

Lung transplantation is the ultimate treatment for patients with end-stage respiratory diseases. A characterization of the lung microbial-host multi-omics may help understand the pathogenesis leading to complications post lung transplantation. Das et al. characterized the microbial communities through a combined 16 S rRNA gene amplicon sequencing and culturome in 234 BAL samples from 64 lung transplant recipients and established their links to host transcriptome data [105]. They identified distinct 'pneumotypes' through an unsupervised machine learning algorithm and established links to viral loads, host gene expression, lung function, and transplant health through a combination of classification and regression models. By analyzing the temporal dynamics of the 16 S rRNA gene sequencing-based microbiome, metabolome, lipidome, and the cellular composition using flow-cytometry in 78 patients post lung transplantation, Watzenbock et al. showed that the integrative multi-omic data could predict future changes in the lung function [106]. Given the need of the recipients in undergoing frequent surveillance of lung infection and post-transplantation complications, lung transplantation presents a unique opportunity to study the longitudinal dynamics of lung microbiome and microbiome-host interaction [107].

5.7. Lung cancer

Lung cancer is the leading cause of cancer incidence and mortality worldwide [108]. The local microbiota may play a role in the pathogenesis of tumor development and could be a novel marker in the initiation, progression and treatment outcome of lung cancer. Through a 16 S rRNA gene sequencing on the lung tissue from lung cancer patients along with the RNA-sequencing data from the Cancer Genome Atlas as validation, Greathouse et al. demonstrated an association between *Acidovorax* and lung squamous cell carcinoma with TP53 mutations

[109]. Through an integration of the airway microbiome and host transcriptome, together with lung cancer murine models, Tsay et al. showed that oral commensal members *Veillonella parvula* was causally associated with decreased survival, increased tumor burden, IL-17 inflammation and activation of checkpoint inhibitor markers [110]. These studies demonstrate a role of the lung microbial-immune crosstalk in lung cancer development.

6. Current challenges and future directions in respiratory multi-omics

Despite advances in utilizing multi-omics at the lung microbiome-host interface, the field is still in its infancy. There remain critical challenges from sample preparation, to data analysis and validation. These challenges are mainly caused by the limitations in physically sampling the airway environment and the inherent complexity of the multi-omic data generated from heterogeneous methods and platforms. With respect to sample preparation, the specimens (i.e. sputum, BAL fluid) collected from the airways often need to be processed in fresh (i.e. by liquefaction, separating cell pellets and supernatants), to generate concurrent samples for the multiple types of omics (metagenomics, metatranscriptomics, metabolomics, and metaproteomics). A standardized approach in processing the airway samples for multi-omics is lacking. With respect to data analysis, most of the multi-omic studies for the respiratory microbiome are correlation-based and are therefore descriptive. Approaches have been developed to further mine the evidence of biological interaction from the correlations leveraging public databases. However, as these databases are built upon existing domain knowledge (i.e. on microbial genetics and metabolite-host interactions), they could inherently limit the possibility on identifying novel hypotheses and pathways for microbiome-host interactions. With respect to validation, compared to the gut microbiome where approaches such as fecal microbiome transplantation are routinely applied in animal models, there remain a lack of approach to precisely manipulate the lung microbiome. Although transplantation of the microbiome in mouse BAL fluid has been applied in a few recent studies [91,111,112], a standardized procedure for effective lung microbiome manipulation is lacking.

In light of these challenges, future studies are warranted to establish a standardized methodology in sampling, processing and analyzing the respiratory multi-omic data, to develop novel approaches leveraging machine learning and the state-of-the-art artificial intelligence algorithms to mine the multi-omic associations without the limitation of existing databases, and to leverage cutting-edge experimental approaches to understand the causative and mechanistic role of the lung microbial ecosystems. Ultimately, through efforts in overcoming these challenges, there are foreseeable promises for using multi-omics toward a better understanding the lung microbiome and its potential as a biomarker or therapeutic target for respiratory diseases.

Author contributions

Jingyuan Gao: Data curation, Writing - original draft. **Xinzhu Yi:** Data curation, Writing - original draft. **Zhang Wang:** Conceptualization, Funding acquisition, Supervision, Writing - review & editing.

Declaration of Generative AI and AI-assisted technologies in the writing process

No generative AI or AI-assisted technologies were used in the writing process.

Competing interests

The authors declare no competing interests.

Acknowledgements

This work was supported by the National Key R&D Program of China (2022YFA1304300), and the National Natural Science Foundation of China (31970112, 32170109, 41907211).

References

[1] Moffatt MF, Cookson WO. The lung microbiome in health and disease. Clin Med (Lond) 2017;17:525–9. https://doi.org/10.7861/clinmedicine.17-6-525.
[2] Yi X, Gao J, Wang Z. The human lung microbiome—a hidden link between microbes and human health and diseases. iMeta 2022;e33. https://doi.org/10.1002/imt2.33.
[3] Whiteside SA, McGinniss JE, Collman RG. The lung microbiome: progress and promise. J Clin Invest 2021;131. https://doi.org/10.1172/JCI150473.
[4] Dickson RP, Erb-Downward JR, Huffnagle GB. The role of the bacterial microbiome in lung disease. Expert Rev Respir Med 2013;7:245–57. https://doi.org/10.1586/ers.13.24.
[5] Blanco-Miguez A, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol 2023. https://doi.org/10.1038/s41587-023-01688-w.
[6] Jansson JK, Baker ES. A multi-omic future for microbiome studies. Nat Microbiol 2016;1:16049. https://doi.org/10.1038/nmicrobiol.2016.49.
[7] Narayana JK, et al. Mathematical-based microbiome analytics for clinical translation. Comput Struct Biotechnol J 2021;19:6272–81. https://doi.org/10.1016/j.csbj.2021.11.029.
[8] Frix AN, et al. Radiomics in lung diseases imaging: state-of-the-art for clinicians. J Pers Med 2021;11. https://doi.org/10.3390/jpm11070602.
[9] Paggiaro PL, et al. Sputum induction. Eur Respir J Suppl 2002;37:3s–8s. https://doi.org/10.1183/09031936.02.00000302.
[10] An SQ, Warris A, Turner S. Microbiome characteristics of induced sputum compared to bronchial fluid and upper airway samples. Pedia Pulmonol 2018;53:921–8. https://doi.org/10.1002/ppul.24037.
[11] Charlson ES, et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med 2011;184:957–63. https://doi.org/10.1164/rccm.201104-0655OC.
[12] Matsuo Y, et al. Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION nanopore sequencing confers species-level resolution. BMC Microbiol 2021;21:35. https://doi.org/10.1186/s12866-021-02094-5.
[13] Wang Z, et al. A refined view of airway microbiome in chronic obstructive pulmonary disease at species and strain-levels. Front Microbiol 2020;11:1758. https://doi.org/10.3389/fmicb.2020.01758.
[14] Mac Aogain M, et al. Immunological corollary of the pulmonary mycobiome in bronchiectasis: the CAMEB study. Eur Respir J 2018;52. https://doi.org/10.1183/13993003.00766-2018.
[15] Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis. Nat Biotechnol 2017;35:833–44. https://doi.org/10.1038/nbt.3935.
[16] Nelson MT, et al. Human and extracellular DNA depletion for metagenomic analysis of complex clinical infection samples yields optimized viable microbiome profiles. Cell Rep 2019;26(2227–2240):e2225. https://doi.org/10.1016/j.celrep.2019.01.091.
[17] Marotz CA, et al. Improving saliva shotgun metagenomics by chemical host DNA depletion. Microbiome 2018;6:42. https://doi.org/10.1186/s40168-018-0426-3.
[18] Charalampous T, et al. Nanopore metagenomics enables rapid clinical diagnosis of bacterial lower respiratory infection. Nat Biotechnol 2019;37:783–92. https://doi.org/10.1038/s41587-019-0156-5.
[19] Shakya M, Lo CC, Chain PSG. Advances and challenges in metatranscriptomic analysis. Front Genet 2019;10:904. https://doi.org/10.3389/fgene.2019.00904.
[20] Ren L, et al. Transcriptionally active lung microbiome and its association with bacterial biomass and host inflammatory status. mSystems 2018;3. https://doi.org/10.1128/mSystems.00199-18.
[21] Sulaiman I, et al. Functional lower airways genomic profiling of the microbiome to capture active microbial metabolism. Eur Respir J 2021;58. https://doi.org/10.1183/13993003.03434-2020.
[22] Sulaiman I, et al. Microbial signatures in the lower airways of mechanically ventilated COVID19 patients associated with poor clinical outcome. Res Sq 2021. https://doi.org/10.21203/rs.3.rs-266050/v1.
[23] Sulaiman I, et al. Lower airway dysbiosis augments lung inflammatory injury in mild-to-moderate COPD. Am J Respir Crit Care Med 2023. https://doi.org/10.1164/rccm.202210-1865OC.
[24] Kleiner M. Metaproteomics: much more than measuring gene expression in microbial communities. mSystems 2019;4. https://doi.org/10.1128/mSystems.00115-19.
[25] Maron PA, Ranjard L, Mougel C, Lemanceau P. Metaproteomics: a new approach for studying functional microbial ecology. Micro Ecol 2007;53:486–93. https://doi.org/10.1007/s00248-006-9196-8.
[26] Heyer R, et al. Challenges and perspectives of metaproteomic data analysis. J Biotechnol 2017;261:24–36. https://doi.org/10.1016/j.jbiotec.2017.06.1201.
[27] Finch S, et al. Pregnancy zone protein is associated with airway infection, neutrophil extracellular trap formation, and disease severity in bronchiectasis. Am J Respir Crit Care Med 2019;200:992–1001. https://doi.org/10.1164/rccm.201812-2351OC.

[28] Wang Z, et al. Airway host-microbiome interactions in chronic obstructive pulmonary disease. Respir Res 2019;20:113. https://doi.org/10.1186/s12931-019-1085-z.

[29] Yan Z, et al. Multi-omics analyses of airway host-microbe interactions in chronic obstructive pulmonary disease identify potential therapeutic interventions. Nat Microbiol 2022;7:1361–75. https://doi.org/10.1038/s41564-022-01196-8.

[30] Abdel-Aziz MI, et al. A multi-omics approach to delineate sputum microbiome-associated asthma inflammatory phenotypes. Eur Respir J 2022;59. https://doi.org/10.1183/13993003.02603-2021.

[31] Dicker AJ, et al. The sputum microbiome, airway inflammation and mortality in chronic obstructive pulmonary disease. J Allergy Clin Immunol 2020. https://doi.org/10.1016/j.jaci.2020.02.040.

[32] Keir HR, et al. Neutrophil extracellular traps, disease severity, and antibiotic response in bronchiectasis: an international, observational, multicohort study. Lancet Respir Med 2021;9:873–84. https://doi.org/10.1016/s2213-2600(20)30504-x.

[33] Hull RC, et al. Sputum proteomics in nontuberculous mycobacterial lung disease. Chest 2022;161:1180–91. https://doi.org/10.1016/j.chest.2021.11.014.

[34] Jansma J, El Aidy S. Understanding the host-microbe interactions using metabolic modeling. Microbiome 2021;9:16. https://doi.org/10.1186/s40168-020-00955-1.

[35] Li L, et al. Neisseria species as pathobionts in bronchiectasis. Cell Host Microbe 2022;30(1311–1327):e1318. https://doi.org/10.1016/j.chom.2022.08.005.

[36] Liang W, et al. Airway dysbiosis accelerates lung function decline in chronic obstructive pulmonary disease. Cell Host Microbe 2023. https://doi.org/10.1016/j.chom.2023.04.018.

[37] Shaffer M, et al. AMON: annotation of metabolite origins via networks to integrate microbiome and metabolome data. BMC Bioinforma 2019;20:614. https://doi.org/10.1186/s12859-019-3176-8.

[38] Yu G, Xu C, Zhang D, Ju F, Ni Y. MetOrigin: discriminating the origins of microbial metabolites for integrative analysis of the gut microbiome and metabolome. iMeta 2022;1:e10. https://doi.org/10.1002/imt2.10.

[39] Wylie KM. The virome of the human respiratory tract. Clin Chest Med 2017;38:11–9. https://doi.org/10.1016/j.ccm.2016.11.001.

[40] Li Y, et al. Altered respiratory virome and serum cytokine profile associated with recurrent respiratory tract infections in children. Nat Commun 2019;10:2288. https://doi.org/10.1038/s41467-019-10294-x.

[41] Choi S, et al. Lung virome: new potential biomarkers for asthma severity and exacerbation. J Allergy Clin Immunol 2021;148(1007–1015):e1009. https://doi.org/10.1016/j.jaci.2021.03.017.

[42] Mac Aogain M, et al. Integrative microbiomics in bronchiectasis exacerbations. Nat Med 2021;27:688–99. https://doi.org/10.1038/s41591-021-01289-7.

[43] Lagier JC, et al. Culturing the human microbiota and culturomics. Nat Rev Microbiol 2018;16:540–50. https://doi.org/10.1038/s41579-018-0041-0.

[44] Whelan FJ, et al. Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. Nat Microbiol 2020;5:379–90. https://doi.org/10.1038/s41564-019-0643-y.

[45] Muggeo A, et al. Extended bacteria culture-based clustering identifies a phenotype associating increased cough and enterobacterales in stable chronic obstructive pulmonary disease. Front Microbiol 2021;12:781797. https://doi.org/10.3389/fmicb.2021.781797.

[46] Sun Y, et al. Characterization of lung and oral microbiomes in lung cancer patients using culturomics and 16S rRNA gene sequencing. Microbiol Spectr 2023;e0031423. https://doi.org/10.1128/spectrum.00314-23.

[47] Raju S, Ghosh S, Mehta AC. Chest CT signs in pulmonary disease: a pictorial review. Chest 2017;151:1356–74. https://doi.org/10.1016/j.chest.2016.12.033.

[48] Rogers W, et al. Radiomics: from qualitative to quantitative imaging. Br J Radio 2020;93:20190948. https://doi.org/10.1259/bjr.20190948.

[49] Zhou M, et al. Non-small cell lung cancer radiogenomics map identifies relationships between molecular and imaging phenotypes with prognostic implications. Radiology 2018;286:307–15. https://doi.org/10.1148/radiol.2017161845.

[50] Wang R, et al. Respiratory microbiota and radiomics features in the stable COPD patients. Respir Res 2023;24:131. https://doi.org/10.1186/s12931-023-02434-1.

[51] Morgan XC, et al. Associations between host gene expression, the mucosal microbiome, and clinical outcome in the pelvic pouch of patients with inflammatory bowel disease. Genome Biol 2015;16:67. https://doi.org/10.1186/s13059-015-0637-x.

[52] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinforma 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.

[53] Alneberg J, et al. Binning metagenomic contigs by coverage and composition. Nat Methods 2014;11:1144–6. https://doi.org/10.1038/nmeth.3103.

[54] Ogata H, et al. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 1999;27:29–34.

[55] Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 2014;42:D490–5. https://doi.org/10.1093/nar/gkt1178.

[56] McArthur AG, et al. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 2013;57:3348–57. https://doi.org/10.1128/AAC.00419-13.

[57] Chen L, et al. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 2005;33:D325–8. https://doi.org/10.1093/nar/gki008.

[58] Brown CL, et al. mobileOG-db: a manually curated database of protein families mediating the life cycle of bacterial mobile genetic elements. Appl Environ Microbiol 2022;88:e0099122. https://doi.org/10.1128/aem.00991-22.

[59] Tiew PY, et al. Environmental fungal sensitisation associates with poorer clinical outcomes in COPD. Eur Respir J 2020;56. https://doi.org/10.1183/13993003.00418-2020.

[60] Croft D, et al. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res 2011;39:D691–7. https://doi.org/10.1093/nar/gkq1018.

[61] Martens M, et al. WikiPathways: connecting communities. Nucleic Acids Res 2021;49:D613–21. https://doi.org/10.1093/nar/gkaa1024.

[62] Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545–50. https://doi.org/10.1073/pnas.0506580102.

[63] Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinforma 2013;14:7. https://doi.org/10.1186/1471-2105-14-7.

[64] Lloyd-Price J, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019;569:655–62. https://doi.org/10.1038/s41586-019-1237-9.

[65] Karp PD, Riley M, Paley SM, Pellegrini-Toole A. The MetaCyc Database. Nucleic Acids Res 2002;30:59–61.

[66] Szklarczyk D, et al. STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. Nucleic Acids Res 2016;44:D380–4. https://doi.org/10.1093/nar/gkv1277.

[67] Ghazi AR, et al. High-sensitivity pattern discovery in large, paired multiomic datasets. Bioinformatics 2022;38:i378–85. https://doi.org/10.1093/bioinformatics/btac232.

[68] Yu T. AIME: Autoencoder-based integrative multi-omics data embedding that allows for confounder adjustments. PLoS Comput Biol 2022;18:e1009826. https://doi.org/10.1371/journal.pcbi.1009826.

[69] Benkirane H, Pradat Y, Michiels S, Cournede PH. CustOmics: a versatile deep-learning based strategy for multi-omics integration. PLoS Comput Biol 2023;19:e1010921. https://doi.org/10.1371/journal.pcbi.1010921.

[70] Rohart F, Gautier B, Singh A, Le Cao K. A. mixOmics: an R package for 'omics feature selection and multiple data integration. PLoS Comput Biol 2017;13:e1005752. https://doi.org/10.1371/journal.pcbi.1005752.

[71] Narayana JK, Mac Aogain M, Ali N, Tsaneva-Atanasova K, Chotirmall SH. Similarity network fusion for the integration of multi-omics and microbiomes in respiratory disease. Eur Respir J 2021;58. https://doi.org/10.1183/13993003.01016-2021.

[72] Li CX, Wheelock CE, Skold CM, Wheelock AM. Integration of multi-omics datasets enables molecular classification of COPD. Eur Respir J 2018;51. https://doi.org/10.1183/13993003.01930-2017.

[73] Mathew J, Aronow WS, Chandy D. Therapeutic options for severe asthma. Arch Med Sci 2012;8:589–97. https://doi.org/10.5114/aoms.2012.30280.

[74] Loverdos K, et al. Lung Microbiome in Asthma: Current Perspectives. J Clin Med 2019;8. https://doi.org/10.3390/jcm8111967.

[75] Barcik W, Boutin RCT, Sokolowska M, Finlay BB. The Role of Lung and Gut Microbiota in the Pathology of Asthma. Immunity 2020;52:241–55. https://doi.org/10.1016/j.immuni.2020.01.007.

[76] Hilty M, et al. Disordered microbial communities in asthmatic airways. PLoS One 2010;5:e8578. https://doi.org/10.1371/journal.pone.0008578.

[77] Gautam Y, Johansson E, Mersha TB. Multi-omics profiling approach to asthma: an evolving paradigm. J Pers Med 2022;12. https://doi.org/10.3390/jpm12010066.

[78] Logotheti M, Agioutantis P, Katsaounou P, Loutrari H. Microbiome research and multi-omics integration for personalized medicine in asthma. J Pers Med 2021;11. https://doi.org/10.3390/jpm11121299.

[79] Chiu CY, et al. Integration of metagenomics-metabolomics reveals specific signatures and functions of airway microbiota in mite-sensitized childhood asthma. Allergy 2020;75:2846–57. https://doi.org/10.1111/all.14438.

[80] Sharma A, et al. Associations between fungal and bacterial microbiota of airways and asthma endotypes. J Allergy Clin Immunol 2019;144(1214–1227):e1217. https://doi.org/10.1016/j.jaci.2019.06.025.

[81] Forno E, et al. A multiomics approach to identify genes associated with childhood asthma risk and morbidity. Am J Respir Cell Mol Biol 2017;57:439–47. https://doi.org/10.1165/rcmb.2017-0002OC.

[82] Soliai MM, et al. Multi-omics colocalization with genome-wide association studies reveals a context-specific genetic mechanism at a childhood onset asthma risk locus. Genome Med 2021;13:157. https://doi.org/10.1186/s13073-021-00967-y.

[83] Chun Y, et al. Integrative study of the upper and lower airway microbiome and transcriptome in asthma. JCI Insight 2020;5. https://doi.org/10.1172/jci.insight.133707.

[84] Raita Y, et al. Integrated omics endotyping of infants with respiratory syncytial virus bronchiolitis and risk of childhood asthma. Nat Commun 2021;12:3601. https://doi.org/10.1038/s41467-021-23859-6.

[85] Wang XW, et al. Benchmarking omics-based prediction of asthma development in children. Respir Res 2023;24:63. https://doi.org/10.1186/s12931-023-02368-8.

[86] Lopez AD, et al. Chronic obstructive pulmonary disease: current burden and future projections. Eur Respir J 2006;27:397–412. https://doi.org/10.1183/09031936.06.00025805.

[87] Franklin W, Lowell FC, Michelson AL, Schiller IW. Chronic obstructive pulmonary emphysema; a disease of smokers. Ann Intern Med 1956;45:268–74.

[88] Wang Z, et al. Lung microbiome dynamics in COPD exacerbations. Eur Respir J 2016;47:1082–92. https://doi.org/10.1183/13993003.01406-2015.

[89] Wang Z, et al. Sputum microbiome temporal variability and dysbiosis in chronic obstructive pulmonary disease exacerbations: an analysis of the COPDMAP study. Thorax 2018;73:331–8. https://doi.org/10.1136/thoraxjnl-2017-210741.

[90] Wang Z, et al. Inflammatory endotype-associated airway microbiome in chronic obstructive pulmonary disease clinical stability and exacerbations: a multicohort longitudinal analysis. Am J Respir Crit Care Med 2021;203:1488–502. https://doi.org/10.1164/rccm.202009-3448OC.
[91] Yadava K, et al. Microbiota promotes chronic pulmonary inflammation by enhancing IL-17A and autoantibodies. Am J Respir Crit Care Med 2016;193:975–87. https://doi.org/10.1164/rccm.201504-0779OC.
[92] Madapoosi SS, 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;206:427–39. https://doi.org/10.1164/rccm.202110-2241OC.
[93] Chalmers JD, Chang AB, Chotirmall SH, Dhar R, McShane PJ. Bronchiectasis. Nat Rev Dis Prim 2018;4:45. https://doi.org/10.1038/s41572-018-0042-3.
[94] Flume PA, Chalmers JD, Olivier KN. Advances in bronchiectasis: endotyping, genetics, microbiome, and disease heterogeneity. Lancet 2018;392:880–90. https://doi.org/10.1016/S0140-6736(18)31767-7.
[95] Narayana JK, et al. Microbial dysregulation of the gut-lung axis in bronchiectasis. Am J Respir Crit Care Med 2023;207:908–20. https://doi.org/10.1164/rccm.202205-0893OC.
[96] Molyneaux PL, et al. Host-microbial interactions in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2017;195:1640–50. https://doi.org/10.1164/rccm.201607-1408OC.
[97] Huang Y, et al. Microbes are associated with host innate immune response in idiopathic PULMONARY FIbrosis. Am J Respir Crit Care Med 2017;196:208–19. https://doi.org/10.1164/rccm.201607-1525OC.
[98] O'Dwyer DN, et al. Lung microbiota contribute to pulmonary inflammation and disease progression in pulmonary fibrosis. Am J Respir Crit Care Med 2019;199:1127–38. https://doi.org/10.1164/rccm.201809-1650OC.
[99] Guan WJ, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;382:1708–20. https://doi.org/10.1056/NEJMoa2002032.
[100] Aslan A, Aslan C, Zolbanin NM, Jafari R. Acute respiratory distress syndrome in COVID-19: possible mechanisms and therapeutic management. Pneumonia, 13. Nathan,; 2021. p. 14. https://doi.org/10.1186/s41479-021-00092-9.
[101] Sulaiman I, et al. Microbial signatures in the lower airways of mechanically ventilated COVID19 patients associated with poor clinical outcome. medRxiv 2021. https://doi.org/10.1101/2021.02.23.21252221.
[102] Zhong H, et al. Characterization of respiratory microbial dysbiosis in hospitalized COVID-19 patients. Cell Discov 2021;7:23. https://doi.org/10.1038/s41421-021-00257-2.
[103] Ren L, et al. Dynamics of the upper respiratory tract microbiota and its association with mortality in COVID-19. Am J Respir Crit Care Med 2021;204:1379–90. https://doi.org/10.1164/rccm.202103-0814OC.
[104] Llorens-Rico V, et al. Clinical practices underlie COVID-19 patient respiratory microbiome composition and its interactions with the host. Nat Commun 2021;12:6243. https://doi.org/10.1038/s41467-021-26500-8.
[105] Das S, et al. A prevalent and culturable microbiota links ecological balance to clinical stability of the human lung after transplantation. Nat Commun 2021;12:2126. https://doi.org/10.1038/s41467-021-22344-4.
[106] Watzenboeck ML, et al. Multi-omics profiling predicts allograft function after lung transplantation. Eur Respir J 2022;59. https://doi.org/10.1183/13993003.03292-2020.
[107] Natalini JG, Singh S, Segal LN. The dynamic lung microbiome in health and disease. Nat Rev Microbiol 2023;21:222–35. https://doi.org/10.1038/s41579-022-00821-x.
[108] Global Burden of Disease Cancer C, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study. JAMA Oncol 2019;5:1749–68. https://doi.org/10.1001/jamaoncol.2019.2996.
[109] Greathouse KL, et al. Interaction between the microbiome and TP53 in human lung cancer. Genome Biol 2018;19:123. https://doi.org/10.1186/s13059-018-1501-6.
[110] Tsay JJ, et al. Lower airway dysbiosis affects lung cancer progression. Cancer Discov 2021;11:293–307. https://doi.org/10.1158/2159-8290.CD-20-0263.
[111] Hosang L, et al. The lung microbiome regulates brain autoimmunity. Nature 2022;603:138–44. https://doi.org/10.1038/s41586-022-04427-4.
[112] Wang S, Zhou Q, Tian Y, Hu X. The lung microbiota affects pulmonary inflammation and oxidative stress induced by pm(2.5) exposure. Environ Sci Technol 2022;56:12368–79. https://doi.org/10.1021/acs.est.1c08888.
[113] Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome 2018;6:158. https://doi.org/10.1186/s40168-018-0541-1.
[114] Zhao L, et al. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. Science 2018;359:1151–6. https://doi.org/10.1126/science.aao5774.
[115] Pedersen HK, et al. Human gut microbes impact host serum metabolome and insulin sensitivity. Nature 2016;535:376–81. https://doi.org/10.1038/nature18646.
[116] Wieder C, Lai RPJ, Ebbels TMD. Single sample pathway analysis in metabolomics: performance evaluation and application. BMC Bioinforma 2022;23:481. https://doi.org/10.1186/s12859-022-05005-1.
[117] Wang L, Wang L, He P. Comprehensive analysis of immune-related gene signature based on ssGSEA algorithms in the prognosis and immune landscape of hepatocellular carcinoma. Front Genet 2022;13:1064432. https://doi.org/10.3389/fgene.2022.1064432.
[118] Hekking PP, et al. Transcriptomic gene signatures associated with persistent airflow limitation in patients with severe asthma. Eur Respir J 2017;50. https://doi.org/10.1183/13993003.02298-2016.
[119] Kuo CS, et al. T-helper cell type 2 (Th2) and non-Th2 molecular phenotypes of asthma using sputum transcriptomics in U-BIOPRED. Eur Respir J 2017;49. https://doi.org/10.1183/13993003.02135-2016.
[120] Li N, et al. Differential proteomic patterns of plasma extracellular vesicles show potential to discriminate β-thalassemia subtypes. iScience 2023;26:106048. https://doi.org/10.1016/j.isci.2023.106048.
[121] Mallick H, et al. Multivariable association discovery in population-scale meta-omics studies. PLOS Comput Biol 2021;17:e1009442. https://doi.org/10.1371/journal.pcbi.1009442.
[122] Faust K, Raes J. CoNet app: inference of biological association networks using Cytoscape. F1000Res 2016;5:1519. https://doi.org/10.12688/f1000research.9050.2.
[123] Kurtz ZD, et al. Sparse and compositionally robust inference of microbial ecological networks. PLoS Comput Biol 2015;11:e1004226. https://doi.org/10.1371/journal.pcbi.1004226.
[124] Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. PLoS Comput Biol 2012;8:e1002687. https://doi.org/10.1371/journal.pcbi.1002687.
[125] Lin L, et al. The airway microbiome mediates the interaction between environmental exposure and respiratory health in humans. Nat Med 2023. https://doi.org/10.1038/s41591-023-02424-2.
[126] Kuhn M. Building Predictive Models in R Using the caret Package. J Stat Softw 2008;28:1–26.
[127] Liaw A, Wiener M. Classification and Regression by randomForest. R N 2002;2:18–22.
[128] Wang T, et al. MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. Nat Commun 2021;12:3445. https://doi.org/10.1038/s41467-021-23774-w.
[129] Zhang L, et al. AutoGGN: a gene graph network AutoML tool for multi-omics research. Artif Intell Life Sci 2021;1:100019. https://doi.org/10.1016/j.ailsci.2021.100019.
[130] Allesøe RL, et al. Discovery of drug–omics associations in type 2 diabetes with generative deep-learning models. Nat Biotechnol 2023;41:399–408. https://doi.org/10.1038/s41587-022-01520-x.
[131] Yu T. AIME: Autoencoder-based integrative multi-omics data embedding that allows for confounder adjustments. PLOS Comput Biol 2022;18:e1009826. https://doi.org/10.1371/journal.pcbi.1009826.
[132] Ma T, Zhang A. Integrate multi-omics data with biological interaction networks using Multi-view Factorization AutoEncoder (MAE. BMC Genom 2019;20:944. https://doi.org/10.1186/s12864-019-6285-x.
[133] Therneau, T.M. A Package for Survival Analysis in R. (2023).
[134] Liu H, et al. Association of sputum microbiome with clinical outcome of initial antibiotic treatment in hospitalized patients with acute exacerbations of COPD. Pharm Res 2020;160:105095. https://doi.org/10.1016/j.phrs.2020.105095.
[135] Durack J, et al. Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. J Allergy Clin Immunol 2017;140:63–75. https://doi.org/10.1016/j.jaci.2016.08.055.
[136] Leitao Filho FS, et al. Sputum microbiome is associated with 1-year mortality following COPD hospitalizations. Am J Respir Crit Care Med 2018. https://doi.org/10.1164/rccm.201806-1135OC.
[137] Li CX, Wheelock CE, Sköld CM, Wheelock Å, M. Integration of multi-omics datasets enables molecular classification of COPD. Eur Respir J 2018;51. https://doi.org/10.1183/13993003.01930-2017.
[138] Wang Z, et al. Multi-omic meta-analysis identifies functional signatures of airway microbiome in chronic obstructive pulmonary disease. ISME J 2020;14:2748–65. https://doi.org/10.1038/s41396-020-0727-y.