Multi-omics profiling predicts allograft function after lung transplantation

Martin L. Watzenboeck[1,2,7], Anna-Dorothea Gorki[1,2,7], Federica Quattrone[1,2,7], Riem Gawish[1,2,7], Stefan Schwarz[3,7], Christopher Lambers[3], Peter Jaksch[3], Karin Lakovits[1,2], Sophie Zahalka[1,2], Nina Rahimi[1,3], Philipp Starkl[1,2], Dörte Symmank[1,2], Tyler Artner[1,2], Céline Pattaroni[4], Nikolaus Fortelny[2], Kristaps Klavins[2], Florian Frommlet[5], Benjamin J. Marsland[4], Konrad Hoetzenecker[3], Stefanie Widder[1,2,6,8] and Sylvia Knapp[1,2,8]

[1]Research Laboratory of Infection Biology, Dept of Medicine I, Medical University of Vienna, Vienna, Austria. [2]CeMM, Research Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria. [3]Division of Thoracic Surgery, Dept of Surgery, Medical University of Vienna, Vienna, Austria. [4]Dept of Immunology and Pathology, Monash University, Melbourne, Australia. [5]Institute of Medical Statistics, Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Vienna, Austria. [6]Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria. [7]These authors contributed equally. [8]S. Widder and S. Knapp contributed equally to this article as lead authors and supervised the work.

Corresponding author: Sylvia Knapp (Sylvia.knapp@meduniwien.ac.at)

Shareable abstract (@ERSpublications)
Broncho-alveolar microbiome, cellular composition, metabolome and lipidome show specific temporal dynamics after lung transplantation. The post-transplantation lung microbiome can predict future changes in lung function with high precision. https://bit.ly/3iFlp1u

Cite this article as: Watzenboeck ML, Gorki A-D, Quattrone F, *et al*. Multi-omics profiling predicts allograft function after lung transplantation. *Eur Respir J* 2022; 59: 2003292 [DOI: 10.1183/13993003.03292-2020].

Received: 27 Aug 2020
Accepted: 09 June 2021

Abstract

Rationale Lung transplantation is the ultimate treatment option for patients with end-stage respiratory diseases but bears the highest mortality rate among all solid organ transplantations due to chronic lung allograft dysfunction (CLAD). The mechanisms leading to CLAD remain elusive due to an insufficient understanding of the complex post-transplant adaptation processes.

Objectives To better understand these lung adaptation processes after transplantation and to investigate their association with future changes in allograft function.

Methods We performed an exploratory cohort study of bronchoalveolar lavage samples from 78 lung recipients and donors. We analysed the alveolar microbiome using 16S rRNA sequencing, the cellular composition using flow cytometry, as well as metabolome and lipidome profiling.

Measurements and main results We established distinct temporal dynamics for each of the analysed data sets. Comparing matched donor and recipient samples, we revealed that recipient-specific as well as environmental factors, rather than the donor microbiome, shape the long-term lung microbiome. We further discovered that the abundance of certain bacterial strains correlated with underlying lung diseases even after transplantation. A decline in forced expiratory volume during the first second (FEV_1) is a major characteristic of lung allograft dysfunction in transplant recipients. By using a machine learning approach, we could accurately predict future changes in FEV_1 from our multi-omics data, whereby microbial profiles showed a particularly high predictive power.

Conclusion Bronchoalveolar microbiome, cellular composition, metabolome and lipidome show specific temporal dynamics after lung transplantation. The lung microbiome can predict future changes in lung function with high precision.

Introduction

Lung transplantation is the ultimate, life-saving treatment for patients with end-stage respiratory failure. Despite substantial medical advances, the 5-year survival rate remains at only 54%, which is the lowest among all solid organ recipients [1]. This grim prognosis is a consequence of a range of pathologies that prevent the transplanted lung from maintaining normal function, which are summarised under the umbrella

term chronic lung allograft dysfunction (CLAD). The most frequent manifestation of CLAD is bronchiolitis obliterans syndrome (BOS), which is characterised by the development of airflow limitation caused by bronchiolitis obliterans [2]. The underlying disease mechanism has been linked to allo- and auto-immunity as well as microbial triggers [3], but the precise pathophysiology is not known and no specific treatment is available.

Once considered a sterile body site, the lungs are now known to harbour a unique microbiome that is critical for respiratory health and immune homeostasis [4]. Recent studies indicate that community-level activities can modulate microbiome pathogenicity [5], myeloid cell responses [6], pulmonary remodelling [7, 8] and CLAD development [9, 10] after lung transplantation. However, host–microbe interplay in the transplanted lung is insufficiently understood and it is unclear to what degree the donor lung microbiome contributes to the post-transplant microbiome and whether this bears relevance for CLAD development.

Innate immune responses are an emerging concept in understanding the events leading to chronic inflammation and CLAD [11]. As such, resident alveolar macrophages (AMs) seem to be instrumental in CLAD development [12], and the higher abundance of neutrophils in transplants is considered detrimental in graft acceptance and survival [11, 13]. These findings clearly suggest a role for myeloid cells in transplant adaptation and function, yet studies investigating the simultaneous coordination of cellular and pulmonary microbial changes are lacking.

Metabolites reflect local metabolic states and can mediate the crosstalk between immune responses and microbiota. In lung transplant recipients, the metabolome composition of bronchoalveolar lavage (BAL) fluid was used to reliably classify the degree of BOS severity [14] and metabolomics/lipidomics of donor lung *ex vivo* lung perfusion fluids were shown to predict primary graft dysfunction [15]. Considering these findings, we hypothesised that the pulmonary metabolome and lipidome might provide important clues to understanding multi-level adaptation after lung transplantation.

A better understanding of the contribution of the pulmonary microbiome, metabolome, lipidome and cellular profiles to the development of CLAD requires a holistic assessment of post-transplant dynamics. In this explorative cohort study, we characterised the reshaping of the pulmonary environment after transplantation with the primary objective to identify time-dependent factors of the process and the secondary objective to describe drivers of lung function decline. Variation of cell composition, microbial diversity, lipid and metabolite profiles, as well as the spirometry parameter FEV_1 (forced expiratory volume during the first second) served as statistical endpoints. We discovered that the microbial composition after lung transplantation was primarily driven by environmental and recipient-specific factors, independent of the donor microbiome, and identified selected microbial species that correlated with transplant indication. Using a computational model enabled us to predict lung function trajectories from multi-omics data sets. The explorative study design entails a descriptive interpretation of p-values and limits conclusions on negative results due to lack of power for small effects.

Ultimately, a more comprehensive knowledge of graft adaptation will enable us to identify novel therapeutic angles to prevent lung allograft dysfunction in the future.

Materials and methods

Detailed descriptions of materials and methods are provided as a supplement.

Patient collective and recruitment

We included all 78 patients who underwent lung transplantation at the Medical University of Vienna between February 2017 and November 2018. The cohort consisted of 76 double lung and two single lung recipients. Donor lung derived BAL samples of sufficiently high quality were available for 23 matched transplant recipients. To obtain a reliable number of donor samples for cross-sectional analyses we included 24 additional high-quality donor BALs without matching recipients. Detailed inclusion and exclusion criteria are stated in the supplementary information. All BAL samples were collected under sterile conditions during routine bronchoscopy, except for one sample that was collected due to acute respiratory symptoms. Routine bronchoscopy was performed before extubation, at four, eight, 12, 26 and 52 weeks after transplantation or whenever medically indicated. The collection of follow-up samples was extended to April 2019. Processing and analyses of BAL samples are described in the supplementary information. Collected patient variables included recipients' age, sex, transplant indication, immunosuppressive therapy, antibiotic and antifungal therapies, serum C-reactive protein (CRP) concentration, lung function (FEV_1) and time after transplantation at the time of sample collection. Detailed patient characteristics are presented in table 1. The study was approved by the ethics committee at

TABLE 1 Patient characteristics

	All	Subset 16S rRNA Seq	Subset FACS	Subset metabolomics/lipidomics
Patients n	78	73	54	36
Age in years at Tx; median (range)	55.5 (17–69)	55 (17–69)	56.5 (17–69)	58.5 (20–69)
Female patients; n (%)	32 (41)	30 (41)	20 (37)	11 (31)
Samples n	164	143	116	70
Donor samples (recipient-matched) n	47 (23)	35 (15)	24 (12)	9 (6)
Recipient samples n	117	108	92	61
Samples per time bin n				
0–7	22	18	18	16
7–30	18	16	15	10
30–75	28	28	24	10
75–400	49	46	35	25
Samples per patient; median (range)	2 (1–6)	1 (1–5)	2 (1–6)	2 (1–6)
Tx indication; samples n/patients n				
COPD	44/25	40/22	35/20	29/15
CF	20/20	19/19	10/9	6/6
IPF	14/9	12/9	12/7	10/5
A1ATD	14/5	14/5	13/5	7/3
PPH	6/6	5/6	5/3	1/1
Other	19/13	18/12	17/10	9/6
Immunosuppressive therapy[#]; samples n/patients n				
Prednisolone	117/56	108/54	92/43	61/33
Ciclosporin	4/2	3/1	4/2	1/1
Tacrolimus	113/55	105/33	88/41	60/32
Mycophenolate/mofetil	50/23	47/22	41/18	29/13
Everolimus	6/3	6/3	4/2	1/1
Antibiotic therapy[#]; samples n/patients n				
Trimethoprim/sulfamethoxazole	76/40	73/40	62/33	34/22
Gentamicin inhalation	26/16	21/16	23/16	20/13
Piperacilin/tazobactam	24/17	19/16	20/15	20/14
Tobramycin inhalation	9/7	9/7	4/3	2/2
Meropenem	9/8	8/8	7/6	2/2
Colistin	8/6	7/6	3/2	1/1
Ciprofloxacin	5/5	5/5	2/2	2/2
Other	32/23	29/20	21/15	15/11

[#]: number of specimens collected from patients while receiving respective drugs at the time of bronchoalveolar lavage. CF: cystic fibrosis; IPF: idiopathic pulmonary fibrosis; A1ATD: alpha-1 antitrypsin deficiency; FACS: fluorescence-activated cell sorting; 16S rRNA Seq: 16S rRNA gene sequencing; PPH: primary pulmonary hypertension; Tx: treatment.

the Medical University of Vienna (EK-Nr:1418/2018) and written informed consent was obtained from all patients.

16S rRNA gene sequencing

Bacterial DNA was isolated using the QIAamp DNA Microbiome kit (Qiagen) according to the manufacturer's protocol. We included sampling controls by flushing sterilised bronchoscopes and negative controls for DNA extraction. Isolated bacterial DNA was amplified using barcoded, adaptor-linked PCR primers targeting the V1–V2 16S rRNA gene region. PCR products were sequenced using Illumina MiSeq technology in the 2×250 bp configuration. Sequencing data were processed using the dada2.R package. A detailed description of bacterial DNA extraction, library preparation, sequencing and sequencing data processing is included in the supplementary information.

Lipidomics and metabolomics

Liquid chromatography–mass spectrometry (LC-MS) analysis was performed using a Vanquish ultra high performance liquid chromatography system combined with an Orbitrap Fusion Lumos Tribrid mass spectrometer for lipidomics or an Orbitrap Q Exactive mass spectrometer for metabolomics. Lipid separation was performed by reversed phase chromatography. The detailed description of lipid extraction, lipid- and metabolite LC-MS analysis, and data processing is included in the supplementary information.

Statistical analyses

Associations of microbial diversity and richness, cell populations, lipid species and metabolites with time after transplantation were analysed by permutational multivariate variance testing and linear mixed-effect models. Time intervals were entered as a factor in the permutational multivariate variance testing, and as a metric covariate together with patient identity as a random effect in the linear mixed-effect models. Factors explaining microbial variation were identified using distance-based redundancy analysis (dbRDA). The association of amplicon sequencing variant (ASV) abundance with time after transplantation and transplant indication was determined using log-transformed linear models on cumulative sum scaling-normalised ASV data (MaAsLin2 R package). Prediction of lung function from omics data sets was done with ridge regression, using a nested cross-validation scheme for hyperparameter tuning and estimation of predictive accuracy. Throughout the manuscript, when applicable, p-values are adjusted for multiple testing using the Benjamini–Hochberg procedure. Due to the exploratory nature of our research, all p-values are to be interpreted in a descriptive manner. Detailed information on the statistical analyses is provided as supplementary material.

Results

Sample stratification and inflammatory parameters

We stratified samples of 78 lung transplant patients into time bins spanning from 0 to 7, 7 to 30, 30 to 75 and >75 days after lung transplantation based on temporal changes in administered medications (immunosuppressive and antimicrobial therapies) (figure S1a) to detect nonlinear associations with time after transplantation. We observed strong associations of serum CRP and blood leukocyte counts ($p<0.001$ for both) with time, indicating decreased systemic inflammation with time after transplantation (figures S1b and c).

Pulmonary microbiome shows temporal dynamics after lung transplantation

By analysing the microbial taxonomic compositions based on 16S rRNA gene sequencing, we found a high diversity between samples at genus and phylum level (figure 1a, figure S2a). Hierarchical clustering of genus-level profiles instantaneously separated samples collected early (<30 days) and late (>30 days) after transplantation (figure 1a). Species-level microbial diversity (Shannon) and richness (Chao 1) (figure 1b, figure S2b) showed highly significant ($p<0.001$) associations with time after transplantation. Using principal coordinate analysis, we observed a tendency of samples to cluster according to time after transplantation (figure 1c, figure S2c and d). We confirmed these visual patterns by permutational multivariate analysis of variance (PERMANOVA; $p<0.05$ for all pairwise comparisons between time groups) (table S1). Donor samples exhibited a microbial profile that was most similar to samples collected within the first week after transplantation, but their composition differed significantly from later timepoints (PERMANOVA $p=0.057$ *versus* 0–7 days after transplantation; $p=0.001$ *versus* any other time bin) (table S1). These results show that lung microbial composition and diversity undergo significant changes from pre-implantation (donor) states to post-transplantation.

Antibiotic therapy and underlying disease explain variation in microbial composition

To identify the key factors that influence the lung microbiome after transplantation, we combined dbRDA with stepwise model building. We included 108 samples from 73 lung recipients that were collected over the entire study period (table 1). The primary objective was to assess dependent covariates and to identify the combination of covariates that best explained microbial variation (figure 1d, figure S2e).

We first analysed all recipient samples (all time bins) and found that the administration of trimethoprim/sulfamethoxazole, the transplant indication, as well as colistin, meropenem and (inhaled) tobramycin together could best explain inter-sample variation (figure 1d, top panel). Next, we separately analysed each time bin after transplantation and observed that only ciprofloxacin during the first week after transplantation accounted for a significant variation between microbiome samples, while none of the tested antibiotics explained inter-sample variation 7–30 or 30–75 days after transplantation (figure 1d, second to fourth panel). Interestingly, out of all tested factors, only transplant indication was significantly associated with microbial composition 30–75 days after transplantation. For samples collected later than 75 days, transplant indication best explained microbial variation, followed by some antibiotics, age and sex (figure 1d, last panel). We also tested the potential impact of immunosuppression on the recipients' lung microbiome, but the very homogenous immunosuppressive therapy scheme received by this cohort (figure S1a) prevented us from identifying clear associations (data not shown). These results suggest that microbial profiles in samples collected early after transplantation are mainly driven by antibiotic therapy, while patient-associated factors including transplant indication, age and sex show significant explanatory power for microbiome samples collected later.

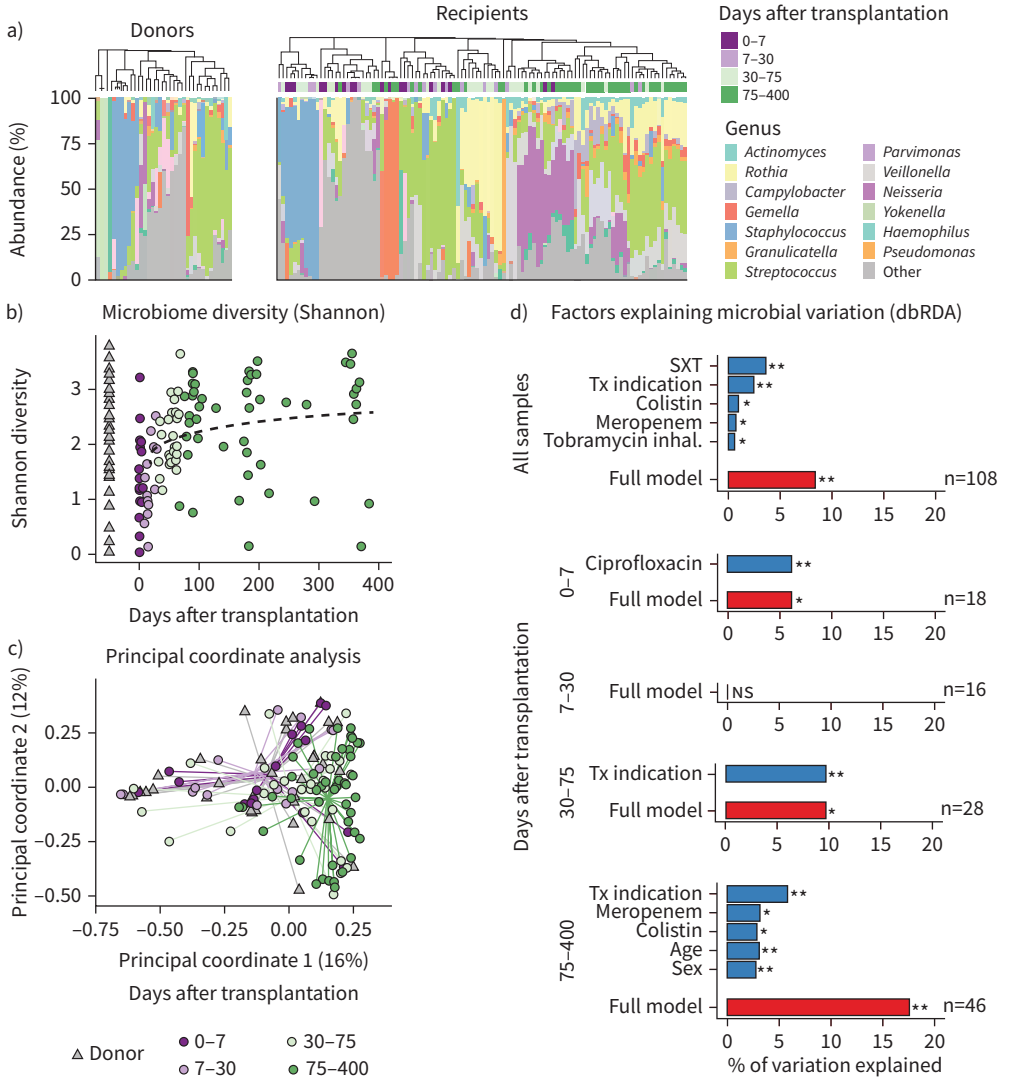

FIGURE 1 Pulmonary microbial dynamics after lung transplantation. a) Relative abundances (genus level) of 16S rRNA gene sequencing amplicons in donor and recipient bronchoalveolar lavage samples. Samples are clustered according to Bray–Curtis distance. Time after transplantation is annotated for recipient samples (heatmap). b) Change in Shannon diversity (species level) with time (linear mixed model p<0.001). Samples are coloured according to time after transplantation and shape indicates donor (triangle)/recipient (circle) origin. c) Principal coordinate analysis on Bray–Curtis distances of donor and recipient microbiomes (species level). Every triangle/circle represents a sample and is coloured according to time after transplantation, while shape indicates donor (triangle)/recipient (circle) origin. Lines connect samples to the cluster centroids p<0.001, permutational multivariate analysis of variance (PERMANOVA). d) Microbial variation explained by various factors analysed by distance-based redundancy analyses (Bray–Curtis distance). R^2 and significance for multivariate models are shown, whereby the best model was generated by selecting predictors in a stepwise manner, until no further increase in accuracy could be reached. *: p<0.05, **: p<0.01, PERMANOVA test. SXT: trimethoprim/sulfamethoxazole; dbRDA: distance-based redundancy analysis; Tx: treatment.

Selected strains associate with time and transplant indication

After identifying global associations of time and transplant indication with the pulmonary microbial profile, we tested if individual ASVs differentiated between these associations with time or transplant indication, respectively. In these analyses, we included 108 recipients' samples derived from 73 lung transplant patients with the following indications for lung transplantation: COPD (n=22), cystic fibrosis (CF) (n=19), idiopathic pulmonary fibrosis (IPF) (n=9), alpha-1 antitrypsin deficiency (A1ATD) (n=5), primary pulmonary hypertension (PPH) (n=6) and other (n=12) (table 1). The distribution of underlying diseases is representative of the patient population at our study site (based on 885 patients between 2010

and 2018; data not shown). We found that 33 individual ASVs significantly (false discovery rate (FDR) <0.1) associated with time after transplantation (figure 2a, table S2). Hierarchical clustering separated these strains into three different clusters (figure 2a). Meanwhile, six ASVs associated with transplant indication (figure 2b, table S2). Of these, we detected the strongest association for a strain of *Pseudomonas* (ASV_2, FDR=0.006), which was increased in CF patients, supporting previous reports that showed a predisposition

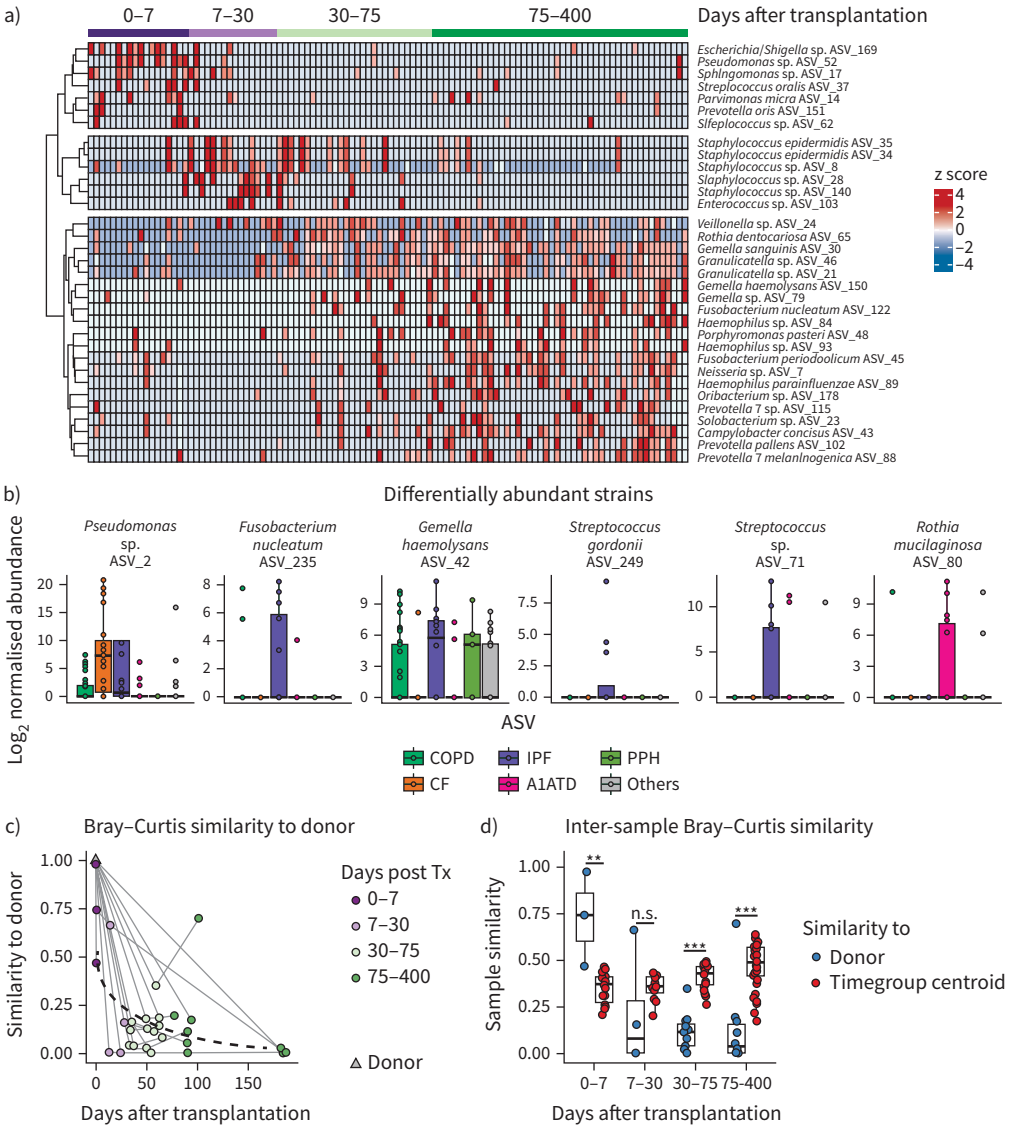

FIGURE 2 Time and underlying disease influence the recipient lung microbiome. a) Heatmap of amplicon sequencing variant (ASVs) showing significant (MaAsLin false discovery rate (FDR)<0.1) association with time after transplantation. Raw counts are cumulative-sum normalised, followed by log-transformation and standardisation (z-score transformation). Samples (columns) are ordered according to time after transplantation. ASVs (rows) are clustered using the "ward.D2" method. The first three branches of the dendrogram are separated for visualisation purposes. b) Boxplots of ASVs showing a significant (MaAsLin FDR<0.1) association with underlying disease. Raw counts are cumulative-sum normalised, followed by log-transformation. Boxplots are indicative of median, interquartile range (IQR) (boxes) and 1.5×IQR (whiskers). CF: cystic fibrosis; IPF: idiopathic pulmonary fibrosis; A1ATD: alpha-1 antitrypsin deficiency; PPH: primary pulmonary hypertension; sp.: species. c) Change of recipient–donor similarity (Bray–Curtis) over time (linear mixed model p<0.001). Connecting lines indicate patient identity; Tx, treatment. d) Comparison between donor–recipient similarity and intersample similarity for various timepoints after transplantation. Boxplots are indicative of median, IQR (boxes) and 1.5×IQR (whiskers). **: p<0.01, ***: p<0.001, Mann–Whitney U-test.

of CF patients to lung allograft colonisation with *Pseudomonas aeruginosa* [16–18]. Furthermore, three strains were almost exclusively found in IPF patients, of which *Fusobacterium nucleatum* (ASV_226) showed the strongest association with transplant indication (FDR=0.01). Another strain identified as *Gemella haemolysans* (ASV_42, FDR=0.03) was rarely encountered in patients with CF or A1ATD, but frequently found in patients with other transplant indications. In addition, *Rothia mucilaginosa* (ASV_80) was preferentially found in patients transplanted due to A1ATD (FDR=0.09). Our results show that transplant indications continuously influence the abundance of certain bacterial strains in the lung after transplantation, in particular in patients with CF, IPF or A1ATD.

Donor lung microbial profiles do not determine the post-transplant microbiome

To determine to what extent the lung allograft microbiome is influenced by the donor lung microbiome, we analysed the similarity between matched donor and recipient microbiomes over time. Bray–Curtis similarity between donors and recipients decreased after transplantation ($p<0.001$) (figure 2c). Similar trends were observed for weighted ($p<0.001$) and unweighted UniFrac similarities ($p<0.01$) (figure S3a and b), which incorporate phylogenetic distances between different organisms. In the first week after transplantation, we found recipient samples to be more similar to matched donor samples than to time-matched recipients (figure 2d). However, at all subsequent timepoints, recipient samples were more similar to samples collected within a comparable timeframe than to their respective donors (figure 2d, figure S3c and d). Thus, the donor microbiome influenced the recipient microbiome only shortly after transplantation, while recipient-associated factors became dominant in shaping the long-term post-transplant pulmonary microbiome.

Pulmonary cellular and molecular profiles change after lung transplantation

The abundance of AMs and inflammatory cells, such as neutrophils, in the BAL correlates with the lung inflammatory state [19]. To investigate alterations in pulmonary cellular profiles after lung transplantation, we conducted flow cytometry of BAL myeloid cell populations (figure S4a). We observed that the numbers of AMs and cells expressing major histocompatibility complex class II (MHCII) but not cluster of differentiation 206 (CD206) were significantly associated with time after transplantation (FDR<0.01 for AMs, FDR<0.001 for $MHCII^+$ $CD206^-$ cells, table S3). While the AM population consistently increased with time (figure 3a), $MHCII^+$ $CD206^-$ cells peaked between 30 and 75 days after transplantation (figure 3b). Concomitantly, myeloid cells ($CD11b^+$) that expressed neither MHCII nor CD206 nor the neutrophil marker CD66b expanded early after transplantation and then declined (FDR<0.05) (figure S4b, table S3). Neutrophils, on average, were surprisingly abundant at all timepoints, but showed high variability between patients (figure 3c). None of the remaining populations significantly associated with time after transplantation (figure S4c and d, table S3) and all investigated cell populations were independent of transplant indication (FDR>0.1 for all) (table S3). Together, we discovered a gradual increase in the proportion of AMs, indicating a transition towards a pattern similar to that observed in healthy lungs [19].

To investigate whether the dynamics of microbial and cellular compositions correlated with BAL metabolic profiles, we performed metabolome and lipidome profiling on a subset of 61 recipients and nine donor samples (table 1). Principal component analysis separated samples according to time after transplantation (figure 3d–e). Samples collected within the first week after transplantation differed most from later samples, suggesting that the most pronounced changes occur early after transplantation (PERMANOVA $p<0.05$ for 0–7 days after transplantation *versus* all other time bins) (tables S4 and S5). To determine whether pre-implantation profiles were more similar to profiles observed early or late after transplantation, we compared donor lavages to recipient lavages. Donor samples differed significantly (PERMANOVA $p<0.05$) from all time bins except the first (0–7 days after transplantation) for both lipids and metabolites (tables S4 and S5), suggesting that lung adaptation after transplantation, rather than the implantation procedure, was responsible for the observed early changes in lipid and metabolite profiles.

Having shown that the global pulmonary lipid and metabolite profile changed after lung transplantation, we next determined which lipid and metabolite species accounted for these changes. Controlling for transplant indication and patient identity, 28 lipid species and eight metabolites showed significant (FDR<0.1) associations with time after transplantation (figure 3f, tables S6 and S7). Of these, 13 lipid species and four metabolites positively correlated with time, whereas 15 lipid species and four metabolites showed a negative correlation with time (figure 3f, figure S4E). None of the tested lipid species were affected by transplant indication (FDR>0.1 for all) (table S6). Transplant indication was significantly associated with leucine concentrations (FDR<0.1), which were higher on average in patients with COPD or A1ATD than in patients receiving transplantation due to other diseases (figure S4f, table S7). These results show that alterations in pulmonary microbial profiles after transplantation were accompanied by changes in intra-alveolar cellular composition and molecular profiles.

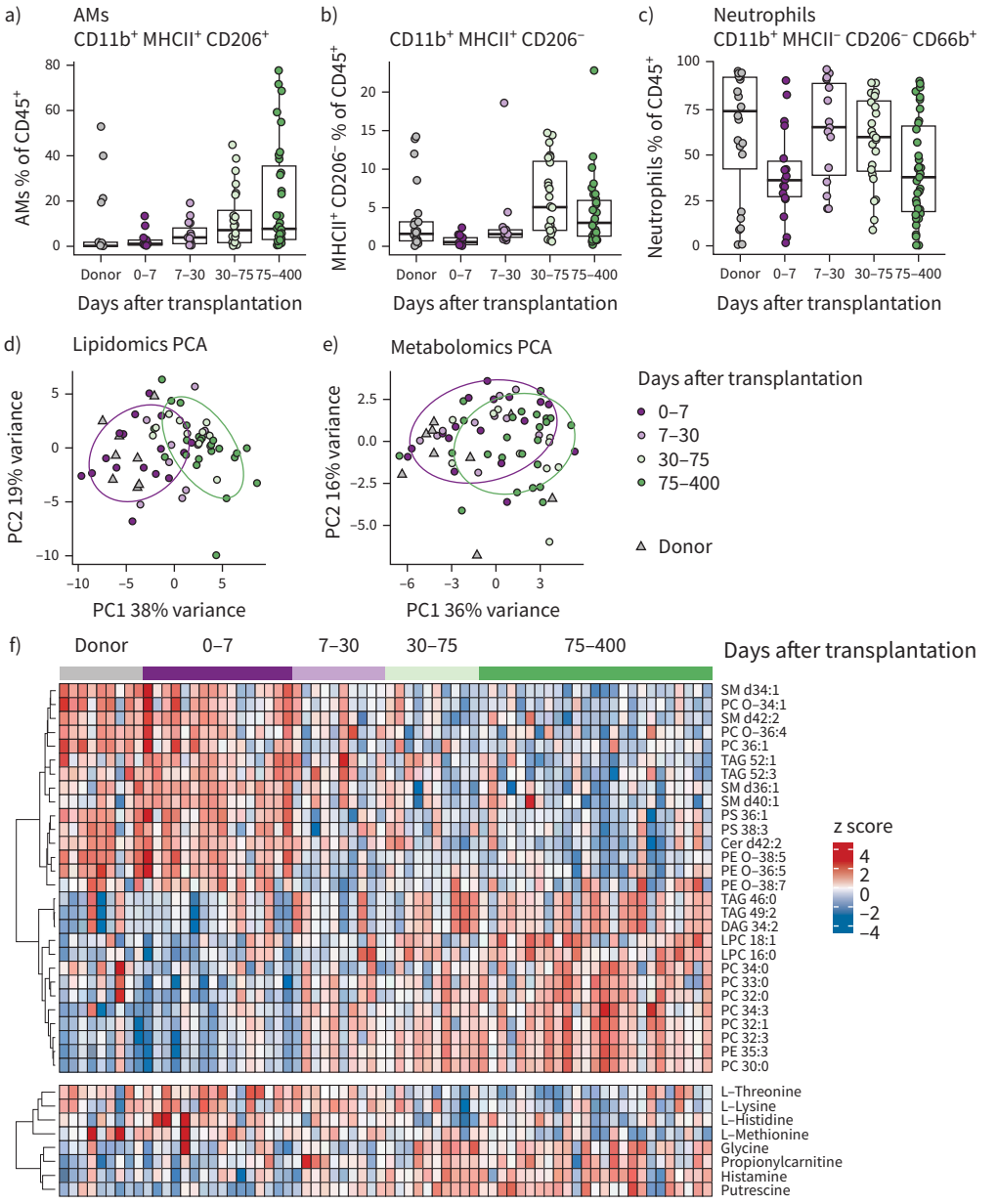

FIGURE 3 Changes in pulmonary cellular and molecular state after lung transplantation. a, b, c) Relative abundance of indicated cell populations in the bronchoalveolar lavage fluid after lung transplantation was assessed by flow cytometry. Boxplots are indicative of median, interquartile range (IQR) (boxes) and 1.5×IQR (whiskers). False discovery rate (FDR)<0.001 for (a), FDR<0.01 (b), FDR>0.1 for (c), linear mixed model. d, e) Principal component analysis (PCA) for lipidome (d) and metabolome (e) data sets. Every triangle/circle represents a sample and is coloured according to time after transplantation, while shape indicates donor (triangle)/recipient (circle) origin. Confidence ellipses (68%) are drawn around samples collected at the earliest (0–7 days after transplantation, purple) and latest (75–400 days after transplantation, green) timepoints. f) Heatmap of lipids and metabolites that show a significant association with time after transplantation (linear mixed model FDR<0.1, controlling for transplant indication and patient identity). AM: alveolar macrophage; MHCII: major histocompatibility complex class II; CD: cluster of differentiation; SM: sphingomyelin; PC: phosphatidylcholine; TAG: triacylglycerol; PS: phosphatidylserine; Cer: ceramide; PE: phosphatidyethanolamine; DAG: diacylglycerol; LPC: lysophosphatidylcholine.

TABLE 2 Characteristics of samples used for prediction of lung function data

	Data sets used for predictions		
	Prediction ΔFEV_1 within 30 days	Prediction ΔFEV_1 within 60 days	Prediction ΔFEV_1 within 90 days
Samples n	32	31	30
Patients n	19	18	17
Females n	8	8	8
Current FEV_1, L; median (range)	2.62 (1.09–4.50)	2.68 (1.09–4.50)	2.70 (1.09–4.50)
Percent change in FEV_1 within X days; median (range)	3 (−7–27)	5 (−18–32)	5 (−29–42)
Age, years; median (range)	59.4 (20.1–68.9)	59.8 (20.1–68.9)	59.4 (20.1–68.9)
CRP, $mg \cdot dL^{-1}$; median (range)	0.17 (0.03–7.46)	0.17 (0.03–7.46)	0.16 (0.03–7.46)

CRP: C-reactive protein; FEV_1: forced expiratory volume during the first second.

Machine learning predicts future lung changes in lung function

CLAD is defined as a persistent decline in FEV_1 of more than 20% compared to the mean of the two best postoperative values (obtained at least three weeks apart) in the absence of other explanations for declining lung function (*e.g.* acute cellular/antibody-mediated rejection, infection, airway stenosis, tracheomalacia) [2]. Several recent studies [3, 20–22] support the importance of the early diagnosis of a decline in lung function, as this knowledge might be utilised to prevent progression and CLAD development. To address this medical need, our secondary objective was to assess whether microbiome, lipidome and metabolome profiles exhibited any predictive value for future changes in FEV_1. We hypothesised that an earlier identification of patients at risk might help in preventing CLAD development and assessed whether the microbiome, lipidome and metabolome profiles exhibited any predictive value for future changes in FEV_1. We included 19 recipients where lung function tests were repeatedly performed over an interval of up to 90 days and all relevant data modalities (microbiome, lipidome, metabolome) and clinical meta-information were available for model building (tables 1 and 2).

Employing a machine learning approach, we trained ridge regression models using clinical metadata, microbiome, metabolome and lipidome data sets as predictors, and changes in FEV_1 at 30, 60 or 90 days after sample collection as response variables (figure 4a). Predictive performance was evaluated using samples collected from patients not included in the training cohort (supplemental methods). To prevent overfitting, we used a nested cross validation scheme by iteratively removing individual patients from the model.

We first tested this approach on individual data sets (*i.e.* microbiome, lipidome or metabolome, respectively) or clinical metadata (sex, age, time after transplantation, current FEV_1 and serum CRP concentration) alone. Doing so, we observed the most accurate short-term predictions of FEV_1 (30 days) when the algorithm was trained on microbiome data (Pearson $r=0.76$, $p<0.001$). For mid-term changes (60 days), microbiome data and clinical metadata performed similarly well (Pearson $r=0.47$, $p<0.01$ for microbiome, $r=0.49$, $p<0.01$ for clinical metadata) while long-term changes (90 days) were most accurately predicted by clinical metadata (Pearson $r=0.42$, $p<0.05$) (figure 4b). Lipidome data performed poorly in comparison to microbiome data, and metabolome data alone could not predict changes in FEV_1. When testing whether combining these individual omics data sets with clinical metadata would improve predictive performance, we found metabolome data in combination with clinical metadata more accurate than any other tested data set in predicting mid-term changes of FEV_1 values (Pearson $r=0.63$, $p<0.001$ for 60 days) (figure 4c). We furthermore evaluated the predictive performance of data sets generated by concatenating individual data modalities (figure 4d), as well as the effect of adding clinical metadata to the combined data sets (figure 4e), which collectively offered superior predictive accuracy over isolated data sets for short- and mid-term changes in FEV_1.

Next, we investigated which specific features contributed to the prediction of FEV_1. The best-performing model for short-term changes in lung function was built on microbiome and metabolome data (Pearson $r=0.78$, $p<0.001$) (figure 4d, f). Analysis of model coefficients showed that *Capnocytophaga gingivalis* received the most positive coefficient (indicating association with improved FEV_1), while *Veillonella dispar* and L-arginine received the most negative coefficients (figure 4f). We observed a tendency towards a higher absolute weight of microbiome *versus* metabolome features (figure S5A). Furthermore, a model built solely on microbiome data performed similarly well (Pearson $r=0.76$, $p<0.001$) (figure 4b, figure

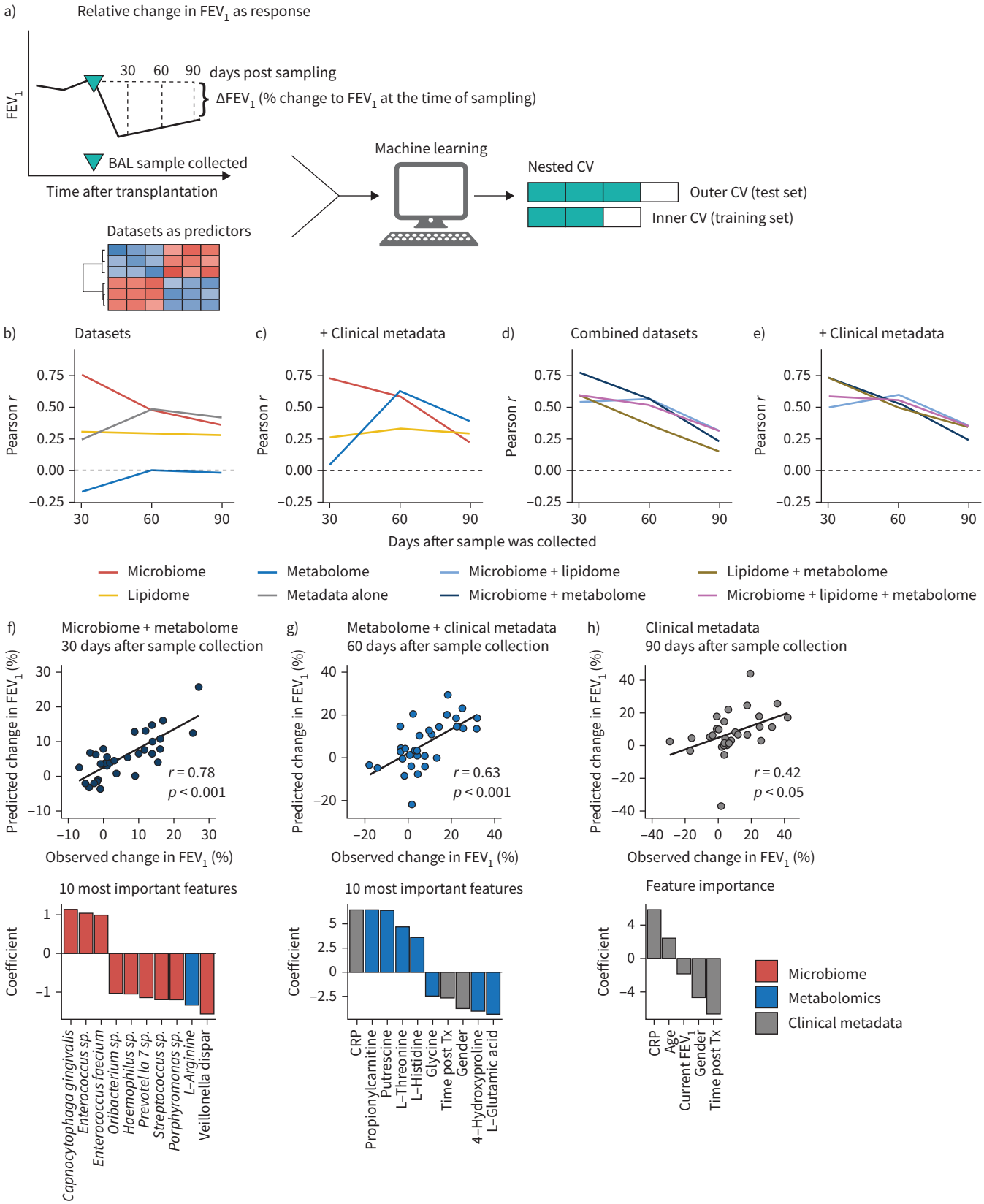

FIGURE 4 Machine learning enables prediction of changes in lung function from multi-omics data. a) Layout of data collection and preparation. Change in forced expiratory volume during the first second (FEV_1) 30, 60 and 90 days after sampling was expressed as percent change to FEV_1 values at the time of sampling. We used our omic data sets to predict these relative changes in FEV_1 using machine learning. To estimate the

performance of our machine learning models on samples which were not included in the data set used to train the model, we employed a nested cross-validation (CV) scheme. b–e) Prediction accuracy (Pearson r between predicted and observed values) for ridge regression models trained on different data sets for changes in FEV_1 at 30, 60 and 90 days after sample collection. Prediction accuracy was evaluated for models trained on specified data sets only (b), on specific data sets plus clinical metadata (age, sex, time after transplantation, serum C-reactive protein (CRP) concentration and current FEV_1) (c), as well as on combined data sets (d) or combined data sets plus clinical metadata (e). f–h) Scatter plots showing observed *versus* predicted changes in FEV_1 and barplots showing the 10 features receiving the highest absolute model coefficients for the best performing models for f) 30, g) 60 and h) 90 days after sample collection. Tx: treatment; sp.: species.

S5b). Intermediate-term changes in lung function were most accurately predicted by a combination of metabolome data and clinical metadata (Pearson r=0.63, p<0.001) with CRP exhibiting the most positive coefficient (figure 4c, g). Finally, for long-term changes in FEV_1, clinical metadata outperformed all other data sets (Pearson r=0.42, p<0.05), with CRP receiving the most positive coefficient, and time after transplantation the most negative (figure 4b, h).

In summary, our data reveal a high predictive power of the lung microbiome for changes in lung function manifesting 30 days after sampling. Surprisingly, clinical metadata including time after transplantation and CRP levels could not match this predictive power for short-term changes, suggesting that effects of the lung microbiome on lung function are independent of time (since transplantation) and systemic inflammation. Long-term changes in FEV_1 were more accurately predicted by clinical metadata, albeit with lower accuracy.

Discussion

Our study provides insight into the pulmonary microbial, cellular and metabolic dynamics after lung transplantation. The analysis of 117 BAL samples from 78 patients together with 47 donor samples makes this one of the largest studies on the human lower respiratory tract microbiome.

We demonstrated that richness and diversity of the microbiome increased rapidly in the first month after lung transplantation and stabilised thereafter, a finding reminiscent of the recently described increase in lung microbiome diversity after birth [23]. Early compositional shifts were linked to antibiotic therapy, in particular trimethoprim/sulfamethoxazole, which is routinely given to prevent infections with *Pneumocystis jirovecii*. Intriguingly, microbial variability at later time points (>30 days after transplantation) showed the strongest association with recipient-specific factors, most notably the transplant indication. We unexpectedly identified several microorganisms that appeared as signature taxa for these lung diseases. While *Pseudomonas* species are the typical pathogens associated with CF, we discovered three bacterial signature strains in IPF patients and one strain associated with A1ATD. Two strains linked to IPF, (*Fusobacterium nucleatum* and *Streptococcus gordonii*), and one strain linked to A1ATD (*Rothia mucilaginosa*) are known colonisers of the oral cavity and upper respiratory tract [24–26]. We hypothesise that these findings may be attributable to recolonisation from the oral flora, which might be altered in IPF or A1ATD patients. Alternatively, altered communication between the upper gastrointestinal tract and the lung (*e.g.* microaspiration) may be responsible, considering the potential association between IPF and reflux disease [27]. While further studies are needed to elucidate the link between IPF or A1ATD and the oral microbiome, we here show specific alterations present in the pulmonary microbiome in IPF and A1ATD patients after lung transplantation.

It is critical to investigate the contribution of the donor microbiome to allograft function, given the fact that non-sterile organs are transplanted to patients. Previous authors suggested that the established lung microbiome after transplantation represents an altered donor microbiome, modified by immigration of microbial species from extrapulmonary sites [28]. This donor microbiome has been proposed as a trigger of pulmonary immune activation, which may contribute to the development of CLAD [3]. In matched samples, we discovered that donor-recipient similarity rapidly decreased after transplantation and that the donor lung microbiome was replaced by a novel bacterial flora in transplanted lungs. These results emphasise that the post-transplantation lung microbiome was independent of the donor microbiome, and that instead a microbial community established that was conserved between patients. Based on these findings, we propose that constant immune activation by the donor microbiome after lung transplantation does not occur.

Early diagnosis of CLAD might enable treatment before irreversible fibrotic remodelling of the lung occurs. Unfortunately, the lack of biomarkers is currently hindering such efforts. Here, using a machine learning approach enabled us to accurately predict changes in FEV_1 based on microbial profiles. Despite

the relatively small-scale sample size used for machine learning, we achieved a highly significant correlation of 0.76 between predicted and observed outcome values. Interestingly, a recent study illustrated the prognostic role of lung microbiota in predicting clinical outcomes, as patients with higher bacterial burdens and the presence of gut-associated taxa were at higher risk of acute respiratory distress syndrome and poor intensive care unit outcome [29]. The authors suggested that translocation of intestinal bacteria to the lung might contribute to lung injury in critically ill patients. Interestingly, we found an abundance of species belonging to the genus *Enterococci*, which are common commensals of the intestinal tract, to be associated with subsequent improvements in FEV_1. This improvement could be linked to adjustments of antibiotic therapy occurring after collection of samples, thereby promoting the elimination of pathogenic bacteria and subsequently improved pulmonary function. Similarly, in predictive models utilising clinical metadata, we observed higher CRP levels to be associated with subsequent improvements in FEV_1, which might seem counterintuitive at first. Elevated serum CRP is a marker of (systemic) inflammation that requires close surveillance and instigates clinicians to initiate proactive treatment. We hypothesise that these activities assist the treatment of potential complications and accelerate resolution, which ultimately leads to improved lung function. Determining whether predictors in our models are beneficial or harmful for lung function is difficult because they are based on associations rather than causality. Beyond CRP, which is a well-understood biomarker, this is particularly true for less well-characterised features such as microbes or metabolites. Our study shares this limitation with any non-interventional approach, which is why clinical trials of targeted therapeutic approaches are required for causal interpretations of host–microbe interplay and lung function to ultimately develop innovative and prophylactic therapies to prevent CLAD.

Acknowledgements: We thank all participating patients, and the medical staff of the Division of Thoracic Surgery, Medical University of Vienna. We thank Thorina Boenke, Martin Senekowitsch, Michael Schuster and Thomas Penz (Biomedical Sequencing Facility, CeMM and Medical University Vienna) for assistance with 16S amplicon sequencing.

Author contributions: M.L. Watzenböck, A.-D. Gorki, F. Quattrone, S. Widder and S. Knapp conceived the project; S. Schwarz, C. Lambers, P. Jaksch, N. Rahimi and K. Hoetzenecker collected samples and clinical data, performed transplantations and managed patient follow-ups; K. Lakovits, F. Quattrone, A.-D. Gorki, R. Gawish, P. Starkl, M.L. Watzenböck, S. Zahalka, D. Symmank and T. Artner processed and analysed samples; K. Klavins performed metabolomic and lipidomic analyses; M.L. Watzenböck analysed data and performed bioinformatics analyses; R. Gawish analysed flow cytometry data; F. Frommlet provided statistical expertise to the study design; C. Pattaroni, N. Fortelny and B.J. Marsland provided technical and analytical support; S. Widder and S. Knapp supervised the project; M.L. Watzenböck, S. Widder and S. Knapp wrote the manuscript. All authors approved the final version of the manuscript.

Conflict of interest: M.L. Watzenböck has nothing to disclose. A.-D. Gorki has nothing to disclose. F. Quattrone has nothing to disclose. R. Gawish has nothing to disclose. S. Schwarz has nothing to disclose. C. Lambers has nothing to disclose. P. Jaksch has nothing to disclose. K. Lakovits has nothing to disclose. S. Zahalka has nothing to disclose. N. Rahimi has nothing to disclose. P. Starkl has nothing to disclose. D. Symmank has nothing to disclose. T. Artner has nothing to disclose. C. Pattaroni has nothing to disclose. N. Fortelny has nothing to disclose. K. Klavins has nothing to disclose. F. Frommlet has nothing to disclose. B.J. Marsland has nothing to disclose. K. Hoetzenecker has nothing to disclose. S. Widder reports grants from Austrian Science Fund (Elise Richter V585-B31), during the conduct of the study. S. Knapp reports grants from FWF, during the conduct of the study.

Support statement: Funding was provided by the Austrian Science Fund (FWF), DK Cell Communication in Health and Disease (W 1205-B09) and the Special Research Program Chromatin Landscapes (L-Mac: F 6104-B21) (to S. Knapp). S. Widder was supported by the Austrian Science Fund (FWF), Elise Richter V585-B31. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Yeung JC, Keshavjee S. Overview of clinical lung transplantation. *Cold Spring Harb Perspect Med* 2014; 4: a015628.
2 Verleden GM, Glanville AR, Lease ED, *et al.* Chronic lung allograft dysfunction: definition, diagnostic criteria, and approaches to treatment–A consensus report from the Pulmonary Council of the ISHLT. *J Heart Lung Transplant* 2019; 38: 493–503.
3 Kuehnel M, Maegel L, Vogel-Claussen J, *et al.* Airway remodelling in the transplanted lung. *Cell Tissue Res* 2017; 367: 663–675.
4 Wypych TP, Wickramasinghe LC, Marsland BJ. The influence of the microbiome on respiratory health. *Nat Immunol* 2019; 20: 1279–1290.

5 Dickson RP, Erb-Downward JR, Martinez FJ, *et al.* The microbiome and the respiratory tract. *Annu Rev Physiol* 2016; 78: 481–504.
6 Sharma NS, Wille KM, Athira S, *et al.* Distal airway microbiome is associated with immunoregulatory myeloid cell responses in lung transplant recipients. *J Heart Lung Transplant* 2018, 37: 206–216.
7 Bernasconi E, Pattaroni C, Koutsokera A, *et al.* Airway microbiota determines innate cell inflammatory or tissue remodeling profiles in lung transplantation. *Am J Respir Crit Care Med* 2016; 194: 1252–1263.
8 Mouraux S, Bernasconi E, Pattaroni C, *et al.* Airway microbiota signals anabolic and catabolic remodeling in the transplanted lung. *J Allergy Clin Immunol* 2018; 141: 718–729.e717.
9 Schott C, Weigt SS, Turturice BA, *et al.* Bronchiolitis obliterans syndrome susceptibility and the pulmonary microbiome. *J Heart Lung Transplant* 2018; 37: 1131–1140.
10 Willner DL, Hugenholtz P, Yerkovich ST, *et al.* Reestablishment of recipient-associated microbiota in the lung allograft is linked to reduced risk of bronchiolitis obliterans syndrome. *Am J Respir Crit Care Med* 2013; 187: 640–647.
11 Hsiao HM, Scozzi D, Gauthier JM, *et al.* Mechanisms of graft rejection after lung transplantation. *Curr Opin Organ Transplant* 2017; 22: 29–35.
12 Nayak DK, Zhou F, Xu M, *et al.* Long-term persistence of donor alveolar macrophages in human lung transplant recipients that influences donor-specific immune responses. *Am J Transplant* 2016; 16: 2300–2311.
13 Speck NE, Schuurmans MM, Benden C, *et al.* Plasma and bronchoalveolar lavage samples in acute lung allograft rejection: the potential role of cytokines as diagnostic markers. *Respir Res* 2017; 18: 151.
14 Ciaramelli C, Fumagalli M, Viglio S, *et al.* (1)H NMR to evaluate the metabolome of bronchoalveolar lavage fluid (BALf) in bronchiolitis obliterans syndrome (BOS): toward the development of a new approach for biomarker identification. *J Proteome Res* 2017; 16: 1669–1682.
15 Hsin MK, Zamel R, Cypel M, *et al.* Metabolic profile of *ex vivo* lung perfusate yields biomarkers for lung transplant outcomes. *Ann Surg* 2018; 267: 196–197.
16 Nunley DR, Grgurich W, Iacono AT, *et al.* Allograft colonization and infections with *Pseudomonas* in cystic fibrosis lung transplant recipients. *Chest* 1998; 113: 1235–1243.
17 Beaume M, Kohler T, Greub G, *et al.* Rapid adaptation drives invasion of airway donor microbiota by *Pseudomonas* after lung transplantation. *Sci Rep* 2017; 7: 40309.
18 Walter S, Gudowius P, Bosshammer J, *et al.* Epidemiology of chronic *Pseudomonas aeruginosa* infections in the airways of lung transplant recipients with cystic fibrosis. *Thorax* 1997; 52: 318–321.
19 Meyer KC, Raghu G, Baughman RP, *et al.* An official American Thoracic Society clinical practice guideline: the clinical utility of bronchoalveolar lavage cellular analysis in interstitial lung disease. *Am J Respir Crit Care Med* 2012; 185: 1004–1014.
20 Royer PJ, Olivera-Botello G, Koutsokera A, *et al.* Chronic lung allograft dysfunction: a systematic review of mechanisms. *Transplantation* 2016; 100: 1803–1814.
21 Kneidinger N, Milger K, Janitza S, *et al.* Lung volumes predict survival in patients with chronic lung allograft dysfunction. *Eur Respir J* 2017; 49: 1601315.
22 Todd JL, Neely ML, Finlen Copeland CA, *et al.* Prognostic significance of early pulmonary function changes after onset of chronic lung allograft dysfunction. *J Heart Lung Transplant* 2019; 38: 184–193.
23 Pattaroni C, Watzenboeck ML, Schneidegger S, *et al.* Early-life formation of the microbial and immunological environment of the human airways. *Cell Host Microbe* 2018; 24: 857–865.e854.
24 Han YW. *Fusobacterium nucleatum*: a commensal-turned pathogen. *Curr Opin Microbiol* 2015; 23: 141–147.
25 Krantz AM, Ratnaraj F, Velagapudi M, *et al. Streptococcus gordonii* empyema: a case report and review of empyema. *Cureus* 2017; 9: e1159.
26 Maraki S, Papadakis IS. *Rothia mucilaginosa* pneumonia: a literature review. *Infect Dis* 2015; 47: 125–129.
27 Ghisa M, Marinelli C, Savarino V, *et al.* Idiopathic pulmonary fibrosis and GERD: links and risks. *Ther Clin Risk Manag* 2019; 15: 1081–1093.
28 Mitchell AB. The lung microbiome and transplantation. *Curr Opin Organ Transplant* 2019; 24: 305–310.
29 Dickson RP, Schultz MJ, van der Poll T, *et al.* Lung microbiota predict clinical outcomes in critically ill patients. *Am J Respir Crit Care Med* 2020; 201: 555–563