



# An Innovative Protocol for Metaproteomic Analyses of Microbial Pathogens in Cystic Fibrosis Sputum

Alexander C. Graf<sup>1\*</sup>, Johanna Striesow<sup>2</sup>, Jan Pané-Farré<sup>3</sup>, Thomas Sura<sup>4</sup>, Martina Wurster<sup>5</sup>, Michael Lalk<sup>5</sup>, Dietmar H. Pieper<sup>6</sup>, Dörte Becher<sup>4</sup>, Barbara C. Kahl<sup>7</sup> and Katharina Riedel<sup>1</sup>

<sup>1</sup> Institute of Microbiology, Department of Microbial Physiology & Molecular Biology, University of Greifswald, Greifswald, Germany, <sup>2</sup> Research Group ZIK Plasmatis, Leibniz Institute for Plasma Science and Technology, Greifswald, Germany, <sup>3</sup> Center for Synthetic Microbiology, Department of Chemistry, Philipps-University Marburg, Marburg, Germany, <sup>4</sup> Institute of Microbiology, Department of Microbial Proteomics, University of Greifswald, Greifswald, Germany, <sup>5</sup> Institute of Biochemistry, Department of Cellular Biochemistry & Metabolomics, University of Greifswald, Greifswald, Germany, <sup>6</sup> Research Group Microbial Interactions and Processes, Helmholtz Centre for Infection Research, Braunschweig, Germany, <sup>7</sup> Institute of Medical Microbiology, University Hospital Münster, Münster, Germany

## OPEN ACCESS

### Edited by:

Chelsie Armbruster,  
University at Buffalo, United States

### Reviewed by:

Megan R. Kiedrowski,  
University of Alabama at Birmingham,  
United States  
Lucia Grenga,  
Commissariat à l'Energie Atomique et  
aux Energies Alternatives (CEA),  
France

### \*Correspondence:

Alexander C. Graf  
grafa@uni-greifswald.de

### Specialty section:

This article was submitted to  
Molecular Bacterial Pathogenesis,  
a section of the journal  
Frontiers in Cellular and  
Infection Microbiology

Received: 13 June 2021

Accepted: 11 August 2021

Published: 27 August 2021

### Citation:

Graf AC, Striesow J, Pané-Farré J, Sura T, Wurster M, Lalk M, Pieper DH, Becher D, Kahl BC and Riedel K (2021) An Innovative Protocol for Metaproteomic Analyses of Microbial Pathogens in Cystic Fibrosis Sputum. *Front. Cell. Infect. Microbiol.* 11:724569. doi: 10.3389/fcimb.2021.724569

Hallmarks of cystic fibrosis (CF) are increased viscosity of mucus and impaired mucociliary clearance within the airways due to mutations of the cystic fibrosis conductance regulator gene. This facilitates the colonization of the lung by microbial pathogens and the concomitant establishment of chronic infections leading to tissue damage, reduced lung function, and decreased life expectancy. Although the interplay between key CF pathogens plays a major role during disease progression, the pathophysiology of the microbial community in CF lungs remains poorly understood. Particular challenges in the analysis of the microbial population present in CF sputum is (I) the inhomogeneous, viscous, and slimy consistence of CF sputum, and (II) the high number of human proteins masking comparably low abundant microbial proteins. To address these challenges, we used 21 CF sputum samples to develop a reliable, reproducible and widely applicable protocol for sputum processing, microbial enrichment, cell disruption, protein extraction and subsequent metaproteomic analyses. As a proof of concept, we selected three sputum samples for detailed metaproteome analyses and complemented and validated metaproteome data by 16S sequencing, metabolomic as well as microscopic analyses. Applying our protocol, the number of bacterial proteins/protein groups increased from 199-425 to 392-868 in enriched samples compared to nonenriched controls. These early microbial metaproteome data suggest that the arginine deiminase pathway and multiple proteases and peptidases identified from various bacterial genera could so far be underappreciated in their contribution to the CF pathophysiology. By providing a standardized and effective protocol for sputum processing and microbial enrichment, our study represents an important basis for future studies investigating the physiology of microbial pathogens in CF *in vivo* – an important prerequisite for the development of novel antimicrobial therapies to combat chronic recurrent airway infection in CF.

**Keywords:** cystic fibrosis, sputum, microbial community, microbiome, 16S sequencing, metaproteomics, metabolomics, *in vivo*

## INTRODUCTION

Cystic fibrosis is the most common inherited monogenic disorder in Caucasian populations with an incidence of approx. one in 3,000 births (O'Sullivan and Freedman, 2009). The disease is caused by mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene, encoding an anion channel localized in epithelial cells e.g. of the respiratory and gastrointestinal tract (Chmiel and Davis, 2003). More than 1,500 mutations of the CFTR gene are described, which all lead to the CF phenotype. Most importantly, the CF phenotype is characterized by an impaired ion homeostasis, which, in consequence, leads to a sticky, dehydrated mucus within the respiratory tract and an impaired mucociliary clearance (Ratjen, 2009). Ultimately, these hallmarks of CF pave the way for the colonization by opportunistic microbial pathogens establishing chronic infections, which starts already early after birth, and is considered to be the main reason for mortality (Rogers et al., 2014). Typically, *Staphylococcus aureus* and *Haemophilus influenzae* represent early colonizers, which are followed by other bacterial pathogens including e.g. *Pseudomonas aeruginosa*, *Burkholderia cepacia* complex and *Stenotrophomonas maltophilia* but also fungal pathogens like *Aspergillus fumigatus* and *Candida albicans*, and viruses (e.g. influenza and respiratory syncytial virus) (Filkins and O'Toole, 2015). The polymicrobial communities within the CF lung are highly dynamic and differ greatly from patient to patient. In the past few years, culture independent diagnostic methods revealed even larger diversity of core genera, which are abundant in the majority of adult patients including *Streptococcus* and *Neisseria*, as well as obligate anaerobes like *Prevotella*, *Veillonella*, and *Catonella* (Rogers et al., 2014; Filkins et al., 2015).

Of note, CF airways are characterized by an inflammatory milieu, which can be attributed to the microbial colonization/infection eliciting a host immune response characterized by the dysregulation of epithelial innate immunity and airway leukocytes. Proteolytic and oxidative products derived from an exuberant immune response in combination with microbial virulence factors are the main reasons for lung tissue damage, which ultimately lead to respiratory failure and death (Cohen and Prince, 2012; Eiserich et al., 2012; Kamath et al., 2015). Thus, deeper insights into these complex polymicrobial infections, focusing on the (patho-)physiology of the microbial CF lung community as well as host-microbe interactions are of essential importance for a better understanding of the disease progression and the development of novel treatment strategies.

In the past, (meta-)proteomics approaches were used as a powerful tool to investigate the physiological alterations of lung tissues and body fluids (e.g. bronchoalveolar lavage, blood, feces, and sputum) in CF patients as well as CFTR post-translational modifications and CF biomarkers (Eiserich et al., 2012; Kamath et al., 2015; Debyser et al., 2016; Liessi et al., 2020). However, most of these studies were limited by focusing on the host perspective while overlooking the microbial side of infection. Studies characterizing the bacterial and fungal pathogens of CF lungs were typically performed *in vitro*, using lung isolates grown under lung-mimicking conditions (Kamath et al., 2015). Consequently, novel

approaches for the *in vivo* analyses of the microbial pathophysiology directly at the site of infection are urgently needed. Here, we present the first *in vivo* microbial metaproteome analysis, complemented by 16S sequencing, metabolomics, and microscopic analyses to study microbial communities and facultative microbial pathogens within CF sputum. To this end, we established an innovative sputum processing protocol, which overcomes major technical and analytical challenges of CF sputum including (I) limited sample volume, (II) challenging processability of CF sputum due to its viscous and slimy character, (III) extraction of nucleic acids, proteins and metabolites out of a single sputum sample, (IV) enormous dominance of human proteins (e.g. mucins, albumins, immunoglobulins) over microbial proteins of interest, and (V) high abundance of (neutrophil-derived) proteases unspecifically digesting microbial proteins of interest (Kamath et al., 2015). In this protocol, a combination of differential centrifugation and filtration is used as key elements for the enrichment of bacterial cells significantly increasing bacterial protein identification coverage.

Our study represents a fundamental basis for follow-up studies investigating the microbial metaproteome and bacterial pathophysiology in CF sputum, which is an essential prerequisite for the development of innovative antimicrobial treatment approaches.

## EXPERIMENTAL PROCEDURES

### Study Cohort, Ethics Statement, and Sputum Sampling

In total, 24 sputum samples derived from 20 different patients were collected. The study was approved by the institutional ethics review board Münster, Germany (2010-155-f-S). Of the 24 sputum samples, 21 were used as test samples to establish a reliable, reproducible and widely applicable protocol suitable for sputum processing, nucleic acid extraction, microbial enrichment and subsequent protein and metabolite extraction. As a proof of concept, three samples, which were derived from three individual patients (designated Patient A, Patient B, and Patient C, respectively) were selected for detailed 16S sequencing, metaproteome and metabolome analyses. Clinical data of these three patients are summarized in **Table 1**. Patient B carried the homozygous Phe508del CFTR genotype, while Patients A and C carried other CFTR-mutations. Importantly, antibiotic therapy of all three patients finished before the time point of sputum sample collection, reducing the risk of false functional analysis due to lysed and/or dead microbial cells. The three Patients A, B, and C were selected based on their differences in age, lung function, antibiotic therapy, disease progression, and microbial lung community structure in order to show the applicability of our sputum processing protocol over preferably diverse samples.

The freshly expectorated sputum samples were immediately chilled on ice and transported to the laboratory for further processing. Next, samples were transferred into 5 mL reaction tubes, three ceramic beads (diameter approx. 2 mm) were added and the samples were homogenized using a Retsch mill at 15 Hz for 120 s (Stokell et al., 2014). Aliquots of the homogenized sputum samples were stained according to the Gram procedure

**TABLE 1 |** Clinical Data of the three CF patients A, B, and C included in metaproteome and metabolome analyses.

	Patient A	Patient B	Patient C
Age	19	24	38
Sex	male	male	male
Exacerbation acc. to Fuchs <sup>a</sup>	0	0	1
FEV1% predicted <sup>b</sup>	81%	52%	28%
Antibiotic therapy	Cefaclor	Cefuroxim	Amoxicillin/Clavulanic acid, Meropenem
CFU/mL <i>S. aureus</i>	1.4 × 10 <sup>7</sup>	3.6 × 10 <sup>6</sup>	3.2 × 10 <sup>7</sup>
CFU/mL <i>P. aeruginosa</i>	–	–	1.5 × 10 <sup>8</sup>
Quantification Neutrophils <sup>c</sup>	3	2	2
Quantification Epithelial Cells <sup>c</sup>	1	2	2

<sup>a</sup>0 = no exacerbation; 1 = a minimum of 4 criteria acc. to Fuchs pertain (Fuchs et al., 1994).

<sup>b</sup>FEV = forced expiratory volume at 1 s.

<sup>c</sup>1 = 1 cell/field of view, 2 = up to 10 cells/field of view, 3 = up to 100 cells/field of view.

and numbers of neutrophils, epithelial cells and bacteria were semi-quantitatively evaluated according to standard diagnostic procedures for CF specimens (Gilligan, 2014). Samples were aliquoted and diluted using 250  $\mu$ L of homogenized samples and 250  $\mu$ L ice-cold 0.9% NaCl. Glycerol was added to a final concentration of 10% and samples were subsequently stored at -80 °C for further analyses.

## Community Composition Analysis by 16S Sequencing

Samples were gently mixed with an equivalent volume of Sputolysin (10%) and incubated for 30 min at 37°C on a ThermoMixer. RNA was extracted using the RNeasy kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, but including a mechanical lysis step (Schulz et al., 2018). After DNA digestion, first-strand complementary DNA was synthesized using the Superscript IV First-Strand Synthesis System (Invitrogen, Carlsbad, CA) and random primers, following the manufacturer's instructions. DNA was extracted from the samples using the FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH, USA) following the manufacturer's instructions (Camarinha-Silva et al., 2014). Amplicon libraries covering the V1-V2 region of the 16S rRNA gene were amplified in a two-step PCR as previously described (Rath et al., 2017) and sequenced on a MiSeq (2X250 bp, Illumina, Hayward, California, USA). Bioinformatic processing was performed as previously described. Raw reads were merged with the Ribosomal Database Project (RDP) assembler (Cole et al., 2013). Sequences were aligned within MOTHUR (gotoh algorithm using the SILVA reference database) and subjected to preclustering (diffs=2) (Schloss et al., 2009) yielding so-called phylotypes that were filtered for an average abundance of  $\geq 0.001\%$  and a sequence length  $\geq 250$  bp before analysis. Phylotypes were assigned to a taxonomic affiliation based on the naïve Bayesian classification (Wang et al., 2007) with a pseudo-bootstrap threshold of 80%. Phylotypes were then manually analyzed against the RDP database using the Seqmatch function. A species name was assigned to a phylotype when only 16S rRNA gene fragments of previously described isolates of that species showed a seqmatch score  $>0.95$ .

## Sputum Sample Processing and Microbial Enrichment

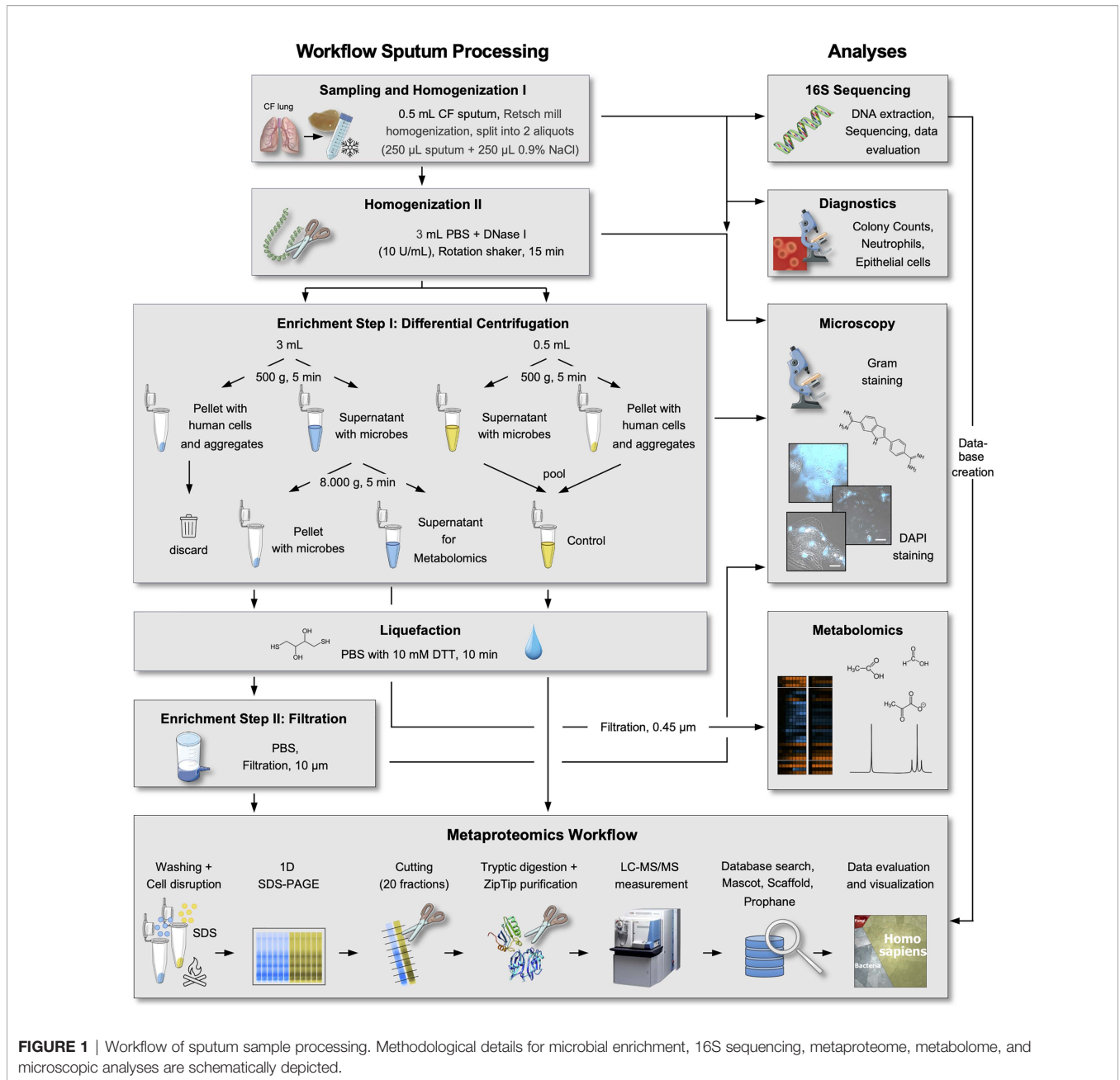
All sputum processing steps were carried out at 4 °C in order to minimize changes of the *in vivo* sputum metaproteome and the metabolome, respectively. The entire workflow is summarized in **Figure 1**. Homogenization of 500  $\mu$ L sputum samples (250  $\mu$ L Retsch mill treated sputum plus 250  $\mu$ L 0.9% NaCl) was performed by adding 3 mL ice-cold PBS<sub>EDTA/PIC</sub> (137 mM NaCl, 0.2 mM KCl, 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4, plus 10 mM EDTA, and 1 tablet protease inhibitor cocktail (PIC, cOmplete, Mini, Sigma-Aldrich) per 10 mL), which was additionally supplemented with DNase I (10 U/mL, ThermoFisher) to break down eDNA-based aggregates (Shak et al., 1990). The samples were subsequently incubated on a rotation shaker (Stuart, Cole-Parmer) at 20 rpm for 15 min. Success of the further homogenization and breakdown of eDNA-based aggregates was microscopically verified (see below). The homogenized sputum suspensions (3.5 mL) were split into a first sub-sample (3 mL) for further enrichment of microbial cells and a second sub-sample (500  $\mu$ L) to obtain a non-enriched control.

For enrichment of microbial cells, the first sub-sample was subjected to differential centrifugation as the first enrichment step of microbial cells. Here, the samples were centrifuged at 500 g for 5 min (keeping cell lysis as low as possible) and the pellet containing human cells and bigger aggregates was discarded. The supernatant, which contained microbial cells, was subsequently centrifuged at 8,000 g for 5 min. The resulting supernatant was filter-sterilized (0.45  $\mu$ m cut-off, Sarstedt) and used for metabolome analyses (see below). The pellet was resuspended in 500  $\mu$ L PBS<sub>EDTA/PIC</sub>, which was additionally supplemented with DTT (10 mM, Sigma-Aldrich) and incubated at 4 °C for 10 min on a rotation shaker (Stuart, Cole-Parmer) to further homogenize and liquify the sample. As the second enrichment step for microbial cells, the cell suspension was subsequently filtered (10  $\mu$ m cut-off, Merck) to remove remaining human cells/aggregates, the filter was washed with 1 mL ice-cold PBS<sub>EDTA/PIC</sub> and the filtrate was collected. The filtrate containing enriched microbial cells was then centrifuged (8,000 g, 5 min), and the pellet was washed twice using ice-cold TE<sub>PIC</sub>-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8, containing 1 tablet protease inhibitor cocktail (PIC, cOmplete, Mini, Sigma-Aldrich) per 10 mL) to further reduce contamination by human proteins. The washed pellet was resuspended in 200  $\mu$ L TE<sub>PIC</sub> and subjected to protein extraction for MS-analyses as described below.

In order to keep preparation protocols of the enriched sample and the non-enriched control as similar as possible, the non-enriched control sample was also subjected to differential centrifugation as described above (500 g, 5 min followed by 8,000 g, 5 min), pooled again and also subjected to liquefaction using 500  $\mu$ L PBS<sub>EDTA/PIC</sub> with DTT (10 mM, Sigma-Aldrich). The suspension was incubated and again centrifuged as described above. The pellets were resuspended using 200  $\mu$ L TE<sub>PIC</sub>, pooled again, and subjected to protein extraction for MS-analyses as described below.

## Protein Extraction

Suspensions of enriched microbial cells and the non-enriched control were subjected to mechanical cell disruption as



**FIGURE 1** | Workflow of sputum sample processing. Methodological details for microbial enrichment, 16S sequencing, metaproteome, metabolome, and microscopic analyses are schematically depicted.

previously described (Becher et al., 2009; Zühlke et al., 2016) since this method was shown to effectively disrupt one of the most robust cell types we expected in our samples – the gram-positive, spherical cocci of *S. aureus*. Due to limited sputum-sample volume and the concomitant small number of microbial cells in the enriched sample, the cell disruption described by (Becher et al., 2009; Zühlke et al., 2016) was downscaled using 200 µL of the respective suspensions and 150 mg glass beads (0.1 to 0.11 mm, Sartorius Stedim Biotech) in 0.5 mL cryotubes (Sarstedt) followed by 3 homogenization cycles at 6.5 m/s for 30 s with intermittent cooling on ice for 1 min in a FastPrep-24™ classic bead beating grinder and lysis system (MP Biomedicals). Subsequently, the samples were centrifuged (15,000 g, 4°C,

5 min) and the tube content (including glass beads and cell debris) was transferred into a fresh 1.5 mL reaction tube. 200 µL 2x extraction buffer (100 mM Tris-HCl, 0.3 M NaCl, 2 mM EDTA, 4% SDS, pH 8.5, adapted from (Chourey et al., 2010)) were added and the suspension was boiled at 95°C and 1,200 rpm for 10 min in a thermo-shaker (Eppendorf). Glass beads and cell debris were pelleted by centrifugation at 15,000 g at 4°C for 5 min and the supernatant, representing the protein extract, was collected. Proteins were concentrated (approx. 4-fold) in a vacuum centrifuge (Eppendorf AG) for 1 h followed by determination of the protein concentration using the BCA-assay microplate procedure (ThermoFisher) according to the manufacturer's instructions.



## MS Sample Preparation

40 µg protein per sample were mixed 3:1 with an SDS-sample buffer (15% glycerol, 5% 2-mercaptoethanol, 2.4% SDS, 0.8% Tris, 0.005% bromophenol blue), boiled for 10 min at 95°C and subsequently separated on a 4-12 % SDS-polyacrylamide gradient gel (Criterion, BioRad). The gel was fixed, washed and stained using Colloidal Coomassie Brilliant Blue G-250 as previously described (Laemmli, 1970; Neuhoff et al., 1988). After the staining procedure, excessive Coomassie stain was removed from the gel using water. Subsequently, gel lanes were fractionated into 20 gel pieces, cut into gel blocks of approx. 1 mm<sup>3</sup> and prepared for MS/MS analysis as described by (Lassek et al., 2015). Obtained peptides were resolved in 0.1% acetic acid and desalted using ZipTips (C18, Merck Millipore). The desalted peptide mixtures were again vacuum-dried and stored at -80°C until MS/MS analysis.

## MS/MS Analysis

Purified peptides were reconstituted with 0.1% acetic acid and analyzed by reversed phase liquid chromatography (LC) electrospray ionization (ESI) MS/MS using an Orbitrap Elite mass spectrometer (Thermo Fisher Scientific, Waltham, USA). Nano-reversed-phase-LC columns (20 cm length x 100 µm diameter) packed with 3.0 µm C18 particles (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) and heated to 45°C were used to separate the purified peptides with an EASY-nLC 1200 system (Thermo Fisher Scientific, Waltham, USA). The peptides were loaded with solvent A [0.1% acetic acid (v/v)] and subsequently eluted by a non-linear gradient from 2% to 99% solvent B (0.1% acetic acid (v/v), 95% acetonitrile) at a flow rate of 300 nl\*min<sup>-1</sup> over 91 min. A full scan was recorded in the Orbitrap with a resolution of 60,000 at m/z 400. The twenty most abundant precursor ions were consecutively isolated and fragmented *via* collision-induced dissociation (CID) with a normalized collision energy of 35. Singly charged ions and ions with unassigned charge state were rejected and lock mass correction as well as dynamic exclusion (fragmented precursors were excluded from fragmentation for 30 s) were enabled. Each sample was measured twice, creating two technical replicates per sample.

## Metaproteomics Data Base Assembly and Search

Three patient-specific databases were constructed based on the phylogenetic information derived from community composition analysis by 16S sequencing. In order to keep the databases and concomitant computational costs as small as possible, genera with a relative abundance of less than 0.1% according to sequencing results were not considered (Table S1). The following protein sequences were added: *Homo sapiens*, the most common (pathogenic) fungal genera in CF (*Aspergillus*, *Blumeria*, *Candida*, *Cladosporium*, *Cryptococcus*, *Exophiala*, *Rasamsonia*, *Rhodotorula*, *Saccharomyces*, *Scedosporium*, and *Sporobolomyces* according to (Chotirmall and McElvaney, 2014; Williams et al., 2016), common laboratory contaminants, and DNase I. For this purpose, FASTA protein sequences were downloaded from

UniProt on September 18, 2018 and redundant entries were removed using the Linux-implemented FASTA tool kit resulting in three patient-specific protein databases, which contained 5.546.037 (Patient A), 4.024.158 (Patient B), and 3.331.936 (Patient C) entries, respectively. Database search was performed using the Mascot software (version 2.6.2, Matrix Science, Boston, MA, USA) with the following settings: peptide tolerance of 10 ppm, MS/MS tolerance of 0.8 Da, up to two missed cleavages allowed, methionine oxidation set as a variable modification, and carbamidomethylation set as a fixed modification. A second database search was performed using Scaffold (version 4.8.7, Proteome Software, Portland, OR, USA) and the built-in X! Tandem search engine with the same settings as described above, as well as the following settings: protein probability = 95%, peptide probability = 99%, single peptide identifications allowed. Here, Mascot and Scaffold used the given databases (containing bacterial, fungal, and human protein sequences) for an *in silico* digestion calculating theoretical peptide sequences and creating theoretical spectra thereof. These theoretical spectra were then matched with experimentally achieved MS/MS spectra for protein identification (Schiebenhoefer et al., 2020). Protein quantification was based on normalized spectral abundance factors (NSAF) as previously described (Zybailov et al., 2006; Zhu et al., 2010). Taxonomic and functional assignment of identified protein groups was performed using ProPhane (Schiebenhoefer et al., 2020) (version 3.1.4) with the settings stated in Table S2. Here, ProPhane provides an automated bioinformatic platform enabling the taxonomic and functional annotation of metaproteome data by integrating various databases (e.g. NCBI, EggNOG, Pfams) and algorithms (e.g. diamond blastp, Hmmer) (Schiebenhoefer et al., 2020).

## Metaproteomics Data Analyses and Visualization

ProPhane output files were used to calculate the mean NSAF of both technical replicates and to create Voronoi treemaps (Bernhardt et al., 2009; Liebermeister et al., 2014) using the Paver software (version 2.1, DECODON GmbH, Greifswald, Germany). Here, Voronoi treemaps visualize taxonomic and functional diversity of sputum samples, respectively, according to relative NSAF-based abundances of different taxonomic genera. To this end, mean NSAF values of all protein/protein group were used, which were identified in at least one out of two technical replicates. Moreover, mean values of both technical replicates were used to visualize protein abundances in bar graphs according to NSAF-based relative abundance as well as absolute number of protein groups for each patient. The protein abundances of each patient were used to calculate mean values and to assess statistically significant differences of protein abundances between enriched and control samples by multiple unpaired t-tests. Enrichment factors of bacterial proteins/protein groups were calculated using three complementary approaches: (i) sum of all NSAFs in the enriched sample divided by the sum of all NSAFs in the control, (ii) absolute number of proteins/protein groups identified in the enriched sample divided by absolute number of proteins/protein groups identified in the control, (iii) percentage of proteins/protein groups identified in

the enriched sample divided by absolute number of proteins/protein groups identified in the control.

## Metabolome Analyses

Samples were lyophilized overnight (Christ, Germany). Dried samples were derivatized for 90 min at 37°C in 40  $\mu$ l Methoxyamine hydrochloride (MeOX) (20 mg/ml in pyridine) and afterwards for 30 min with 80  $\mu$ l of N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) at 37°C. Analytical GC-MS system consisted of an Agilent Technologies 7890B gas chromatograph and a mass selective detector (5977B Inert Plus Turbo MSD, Agilent Technologies). Injection was done with SSL (split/splitless) injector (G4513A, Agilent Technologies) (split 1:25 at 250°C, 1.0  $\mu$ l; carrier gas: Helium with a flow of 1.0 ml/min). The MS operated in the electron impact mode with an ionization energy of 70 eV. The oven program started with 1 min at 70°C, was increased up to 76°C with 1.5°C/min followed by heating up to 220°C with 5°C/min and heating up to 325°C with 20°C/min. The final temperature of 325°C was hold for 8 min. Mass spectra were acquired in scan mode from 50-500 m/z at a rate of 2.74 scans/s and with a solvent delay of 6.0 minutes. Chromatography was performed using a 30 m HP-5 column (Agilent Technologies) with 0.25 mm i.d. and 0.25  $\mu$ m film thickness. The detected compounds were identified by processing of the raw GC-MS data with MassHunter software Qual B.08.00 and comparing with NIST 2017 and Fiehn mass spectral databases and with retention times and mass spectra of standard compounds (inhouse database). The supplemented list contains compounds with scores to libraries of 70 or more (**Table S6**). Metabolites were relatively quantified among the three patient samples and depicted as circles, which areas correlate with metabolite abundance.

## Microscopic Analyses

25  $\mu$ l of the samples derived from different sputum processing steps as indicated in **Figure 1** (after homogenization, after the first differential centrifugation step, after filtration) were transferred into a 96-well microtiter plate and diluted to an OD of 5 using PBS. OD was measured at 500 nm in a microtiter plate reader (Synergy MX, BioTek Instruments, Winooski, USA). Samples were stained in the dark at room temperature for 15 min using DAPI (2  $\mu$ g/ml final concentration, Merck Millipore). 4  $\mu$ l of these samples were applied on a thin layer of 1.5 % agarose in 0.9 % NaCl, which was mounted on a microscope slide. Phase contrast and fluorescence microscopy images were acquired and processed using a Zeiss Imager M2 (Carl Zeiss, Jena, Germany) equipped with a 100x/NA 1.3 oil immersion objective, a filter for monitoring DAPI fluorescence (excitation at 358 nm, emission at 461 nm), and the ZEN 2011 software package (Carl Zeiss, Jena, Germany). The number of human cells, particles/aggregates of different sizes, and microbial cells were counted in 50 randomly selected fields-of-view per sample. The results were averaged among Patient A, B, and C, and statistically significant differences were assessed by multiple unpaired t-tests.

## RESULTS AND DISCUSSION

It is well described that microbial pathogens frequently establish infections in the airways of CF patients during infancy, which

may become chronic, cause severe tissue damage and ultimately lead to death due to respiratory failure (Lyczak et al., 2002) - the leading cause of CF mortality (Rogers et al., 2014). However, the molecular mechanisms underlying co-infection, microbial interplay, and disease progression are still poorly understood.

In order to address these critical open questions, we developed an *in vivo* approach with a specific focus on the metaproteomic analyses of the CF microbiome, driven by 16S sequencing community composition analyses. We complemented these results by metaproteomic data acquired by metabolic footprint analyses, and microscopic data. Since major technical challenges, related to sputum consistency and processability, have so far precluded such analyses, we first established a reliable, reproducible and widely applicable protocol for sputum sample processing and subsequent metaproteomic and metabolomic analyses with a focus on microbial pathogens.

## A Metaproteomic and Metabolomic Analyses Protocol Overcoming Major Technical Challenges Related to CF Sputum Processing

We established a straight-forward workflow allowing nucleic acid extraction, metabolome footprint analyses and microbial protein enrichment and analyses from a single CF sputum sample. The major technical challenges and the different steps of our protocol addressing these technical challenges (**Figure 1**) are presented and discussed in the following paragraphs:

### Limited Sample Volume

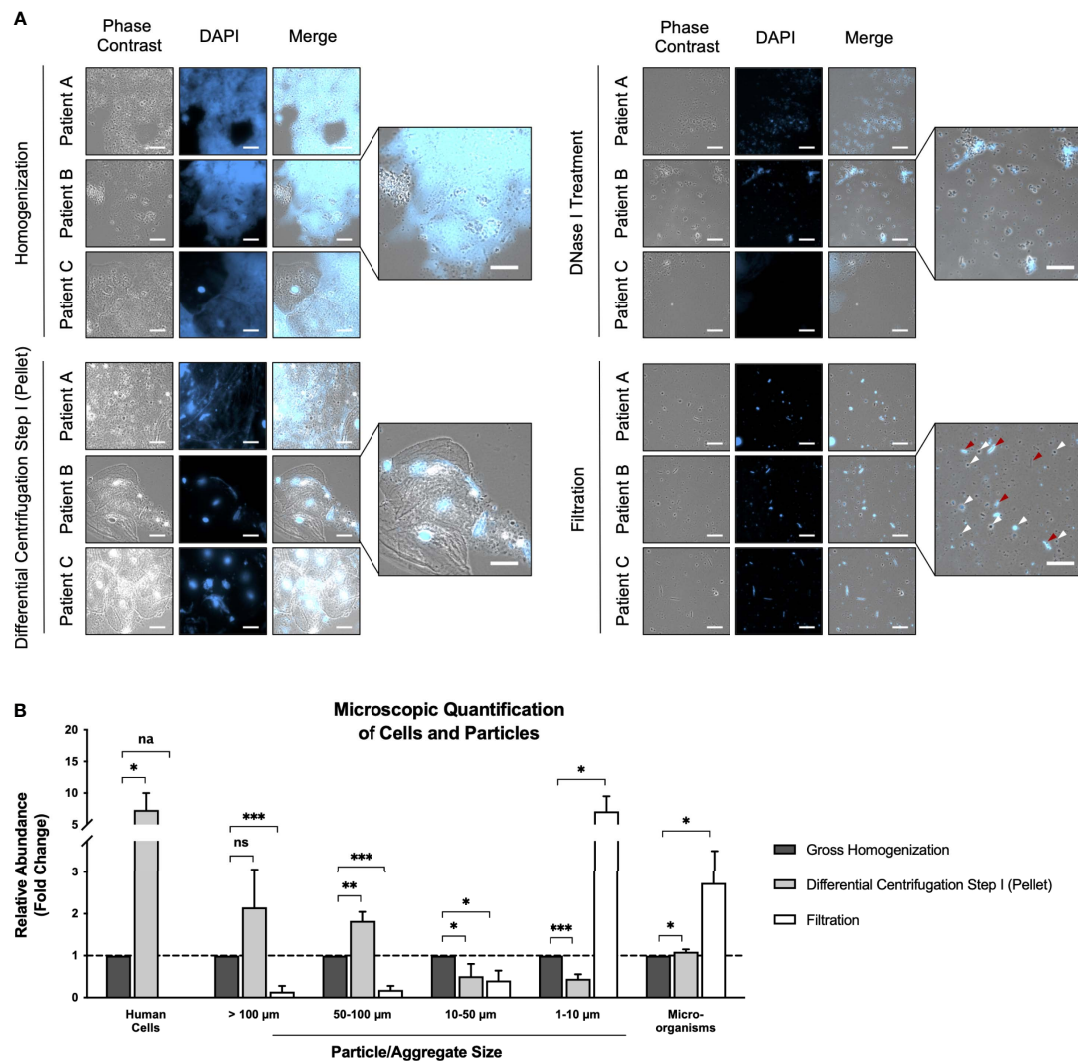
The amount of CF sputum sampled varies from patient to patient and rarely exceeds volumes of a few milliliters, which limits the biomass available for simultaneous nucleic acid, protein, and metabolite extraction. For this study, we collected 24 sputum samples derived from 20 different patients with a sample volume ranging from 0.3 ml to 2 ml (average = 0.7 ml, median = 0.6 ml). In order to establish a protocol, which is applicable to a great variety of different CF patients, we used a sputum volume of 0.5 ml as starting material. This amount was sufficient for simultaneous analyses of nucleic acids, proteins, and metabolites from a single sample.

### Homogenization and Digestion of eDNA-Based Aggregates

CF sputum represents a very viscous and slimy matrix due to macromolecules like eDNA and heavily glycosylated mucins (Stokell et al., 2014; Kamath et al., 2015), which complicates and prolongs downstream processing. Indeed, our microscopic analyses clearly showed massive cell clusters embedded in “clouds” of eDNA, which partially exceeded sizes of 500  $\mu$ m. A common source of this eDNA are NETs (neutrophil extracellular traps): networks of primarily neutrophil-derived eDNA loaded with proteins, which show antimicrobial activity and simultaneously protect the eDNA from degradation (Herzog et al., 2019). To make cells trapped in these eDNA “clouds” accessible, they needed to be broken down prior to further processing (**Figure 2A**). Available techniques for sputum homogenization and liquefaction include mechanical,

chemical or enzymatic treatment at room temperature, or 37 °C (Palmer et al., 2007; Son et al., 2007; Fu et al., 2012; Yang et al., 2012; Wu et al., 2019). However, for metaproteome and metabolome analyses sputum processing needs to be carried out quickly and at 4 °C in order to avoid changes in the composition of the metaproteome and metabolome. Consequently, all sample processing steps were performed at 4 °C. We started our microbial enrichment protocol by homogenizing the sputum samples using a Retsch mill followed by the addition of ice-cold PBS including DNase I and subsequent incubation on a rotary shaker at 4°C. This combination of mechanical and enzymatic treatment resulted in a very efficient homogenization of the sputum

samples. Following this treatment, samples can be pipetted easily and appear homogeneous with the naked eye. Fluorescence microscopy demonstrated that the aforementioned “clouds” of eDNA were successfully digested (**Figure 2A**). Importantly, DNase-treatment not only reduces viscosity of sputum samples (Shak et al., 1990), but also releases microbes trapped within eDNA “clouds” as we confirmed microscopically (**Figure 2A**). Furthermore, DNA digest also releases microbes from biofilms, which are frequently formed by microbial pathogens within the CF lung and contain eDNA as one of the major stabilizing components (Otto, 2008; Goerke and Wolz, 2010; Schwartbeck et al., 2016; Kovach et al., 2017). Thus, DNase-treatment at 4°C critically



**FIGURE 2** | Microscopic analyses during microbial enrichment procedure. **(A)** Representative images of all three samples selected for metaproteome and metabolome analyses acquired after gross homogenization, DNase I treatment, differential centrifugation step I (pellet), and filtration as phase contrast, DAPI, and merged images. One merged image of each enrichment step is magnified for an improved visualization. White arrows indicate cocci, red arrows indicate rod-shaped bacteria. The scale bar represents 10 μm. **(B)** Quantification of human cells, particles/aggregates of varying size, and microorganisms in 50 images acquired in randomly selected fields of view after each enrichment step. Results are depicted as relative abundance (mean fold changes of all three sputum samples) ± standard deviation compared to quantitative values of the gross homogenization sample, which were set to 1. Statistical significance was assessed by multiple unpaired t-tests. Na, not applicable; ns, not significant; \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.



improves microbial enrichment and protein identification coverage of CF sputum samples.

Since sputum processing time needs to be kept as short as possible in order to preserve the metaproteome and the metabolome, we used sputum test samples to gradually reduce DNase I incubation time from 30 min to 10 min, without observing a decrease in DNA digestion efficiency (**Figure S1**), resulting in a total sputum processing time of approx. 60 min.

### Avoiding Liquefaction and Homogenization Strategies Interfering With Metaproteome and Metabolome Analyses

Other methods for mechanical homogenization and eDNA breakdown like vortexing, intense shaking, and sonification, respectively, were intentionally avoided in order to keep human and microbial cells as much intact as possible, which is a prerequisite for unbiased metaproteome and metabolome quantification. Moreover, another commonly used method for sputum homogenization and liquefaction - the digestion of the sputum samples using proteases (Son et al., 2007; Wu et al., 2019) - was also avoided, because it would interfere with metaproteome analysis and would significantly decrease protein coverage. To further reduce unwanted protein degradation due to the high abundance of serine- and metalloproteases in CF sputum samples, a protease inhibitor cocktail was added to all processing steps for metaproteome preservation (Sloane et al., 2005; Kamath et al., 2015).

A further frequently applied strategy for sputum homogenization and liquefaction is the use of chemicals, primarily DTT (commercially available as Sputasol, Sputolysin, or Cleland's reagent) (Stokell et al., 2014). Stokell et al. even considered DTT treatment mandatory, since it is not possible to pipette sputum samples due to their high viscosity without DTT treatment (Stokell et al., 2014). However, we avoided DTT treatment in early steps of our protocol, since high amounts of DTT inhibit DNase I activity and also interfere with metabolome analyses by masking other analytes. Therefore, DTT was only added at a late step of our protocol (after DNase-treatment and sampling aliquots for metabolome analysis) for liquefying the remaining pellet allowing filtration (**Figure 1**). However, depending on the viscosity of the remaining pellet, it should be carefully assessed, if DTT should be used or not since DTT can assist bacterial cell lysis and therefore might have a negative impact on microbial cell recovery (Liu et al., 2018).

### Enrichment of Microbial Proteins Overcoming the Outnumbering Human Protein Abundance

The enormous abundance of human proteins (e.g. mucins, serine and metalloproteases, immunoglobulins, serum albumin) (Kamath et al., 2015) masks comparably low abundant microbial proteins during MS/MS analyses. Confirming this, we found high amounts of human proteins in the non-enriched controls (**Tables S3–S5**). In order not to increase this problem, we kept the first steps of sputum processing as mild as possible, to minimize lysis of human cells and a concomitant contamination of the microbial metaproteome (and metabolic footprint, see above). Thus, to further increase the amount of

microbial proteins compared to human proteins, microbial cells were enriched, while human cells were depleted, prior to MS/MS analyses. Several approaches were tested using the sputum test samples to reduce sample complexity and enrich microbial cells: differential centrifugation (Tanca et al., 2014; Tanca et al., 2017), filtration (Xiong et al., 2015; Schultz et al., 2020), as well as density gradient centrifugation (Hevia et al., 2016). Each procedure was investigated for the enrichment success microscopically (**Figure S2**). Since *S. aureus* is one of the most prevalent and important CF pathogens, we additionally tested an enrichment protocol combining *S. aureus* specific antibodies and magnetic beads adapted from (Bicart-See et al., 2016; Wei et al., 2016) (**Figure S2**). However, neither density gradient centrifugation, nor antibody/magnetic bead enrichment resulted in a reproducible and sufficient enrichment. However, differential centrifugation as well as filtration did result in reproducible but only small enrichment of microbial cells as observed microscopically (data not shown). Therefore, differential centrifugation and filtration were combined resulting in a successful enrichment of microbial cells and an efficient depletion of human cells, respectively (**Figure 2**).

Although our enrichment strategy markedly increased the concentration of recovered microbial cells, a significant amount of biomass was lost during this two-step enrichment procedure. This reasons the necessity to use a higher starting volume of the homogenized sputum for microbial enrichment (3 mL) than for the non-enriched control (0.5 mL) ensuring that the protein yield after enrichment is still sufficiently high. Equal protein amounts (40 µg) of both the enriched samples and the non-enriched controls were used for metaproteome analyses to account for the different input volumes.

### Increasing Protein Yield by Optimizing Cell Disruption and Protein Extraction

Moreover, it has been reported that the cell disruption method critically impacts the extraction efficiency and true species representation in various environmental samples (Starke et al., 2019). Therefore, the subsequent cell disruption and protein extraction process was optimized to increase protein yield. To this end, we used the sputum samples for protocol development and evaluated multiple cell disruption and protein extraction procedures. These included sonification, freeze and thaw cycles, boiling, bead beating, enzymatic treatment, harsh extraction buffers and different combinations of these procedures. Cell disruption and protein extraction efficiency was evaluated by measuring extracted protein concentrations and total protein amounts, respectively (data not shown). Based on these results, we decided to use a combination of downscaled beat-beating adapted from (Becher et al., 2009; Zühlke et al., 2016), which was shown to most effectively disrupt the enormously robust cells of the spheric, Gram-positive CF key pathogen *S. aureus* followed by subsequent boiling of the samples in a harsh SDS-based extraction buffer with a final SDS concentration of 1% adapted from (Chourey et al., 2010). Using this method, we were able to extract the highest protein amounts out of the enriched microbial fraction ranging from approximately 40 to 110 µg, which is sufficient for subsequent metaproteome analysis.



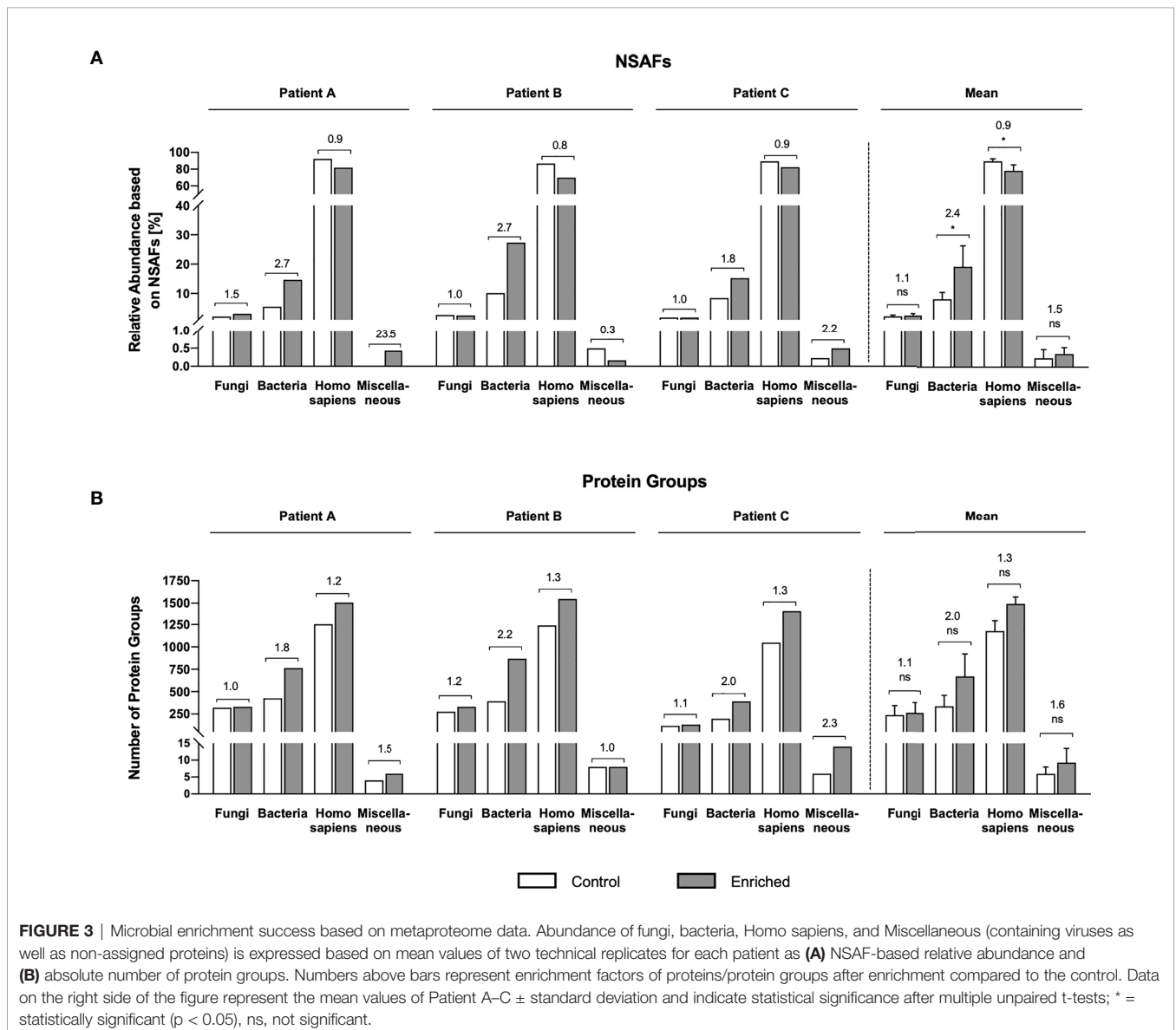
Taken together, we established and optimized a sputum-processing protocol for microbial enrichment characterized by the following major steps (i) mechanical and enzymatic homogenization, (ii) differential centrifugation as the first microbial enrichment step, (iii) liquefaction with DTT, (iv) filtration as the second microbial enrichment step, (v) optimized cell disruption and protein extraction by a combination of beat beating and boiling in SDS extraction buffer.

### Microbial Proteins Were Enriched by a Maximum Factor of 2.7

After we established a protocol for microbial enrichment, we selected three different patients, designated Patient A, Patient B, and Patient C, for detailed metaproteome analyses as a proof of concept. In order to assess the enrichment efficiency of these three sputum samples, we compared a non-enriched control and the enriched sample using a state-of-the-art metaproteomics workflow (Figure 1) and monitored

microbial cell count microscopically (Figure 2). For metaproteome analyses, we measured two technical replicates of each sample in LC-MS/MS experiments showing decent reproducibility (Figure S3). However, the overall percentage of assigned spectra is rather low compared to other metaproteomics datasets (Hinze et al., 2019), which most likely might be attributed to the typically high proteolysis rates within CF sputum caused by neutrophil-derived proteases (Sloane et al., 2005; Folkesson et al., 2012).

Two different complementary read-outs were used to evaluate microbial protein enrichment efficiency: relative protein abundance based on NSAFs, and the number of identified proteins/protein groups (= group of proteins sharing the same identified peptide(s)) (Figure 3). Briefly, NSAF-based quantification of proteomic data refers to a label-free quantification method relying on a spectral counting approach. More precisely, quantification of proteins is carried out by comparing the number of identified MS/MS spectra



of a specific protein over several LC-MS/MS experiments, since protein abundance correlates with the number of proteolytic peptides and thus with the number of total MS/MS spectra (spectral counts). Considering that large proteins naturally contribute a higher number of peptides/spectra compared to small proteins, spectral counts undergo normalization to create the NSAF. Therefore, the number of spectral counts (SC) of a specific protein is divided by the protein's length (L), divided by the sum of all SC/L values from the given experiment (Zhu et al., 2010) allowing relative quantification of proteins throughout samples.

Both, the NSAF as well as the protein group-based evaluation showed a clear trend of successful enrichment of bacterial proteins and depletion of human proteins in all three samples (Figure 3). In fact, the 250 most prominent human proteins (e.g. including mucins, albumins, immunoglobulins), which contribute to the total proteome mass by approx. 40%, were depleted by a mean factor of 1.6 fold (Tables S3–S5). Regarding the enrichment of bacterial proteins, NSAF-based enrichment factors range from 1.8-fold (Patient C) to 2.7-fold (Patient A and Patient B) (Figure 3A). Notably, these enrichment factors are also well reflected by our microscopic analyses. 50 randomly selected fields of view were acquired for the different steps of the protocol. Both qualitatively (Figure 2A) as well as quantitatively (Figure 2B) we observed a clear reduction of human (epithelial) cells and particles bigger than 50  $\mu\text{m}$  in our samples after the first step of differential centrifugation (Figure 1). Particles smaller than 10  $\mu\text{m}$  as well as microbial cells were obviously enriched after filtration (Figure 2). Notably, the mean enrichment factor of bacterial cells calculated from microscopic analyses of 2.7-fold is very close to the bacterial enrichment factors calculated from NSAFs.

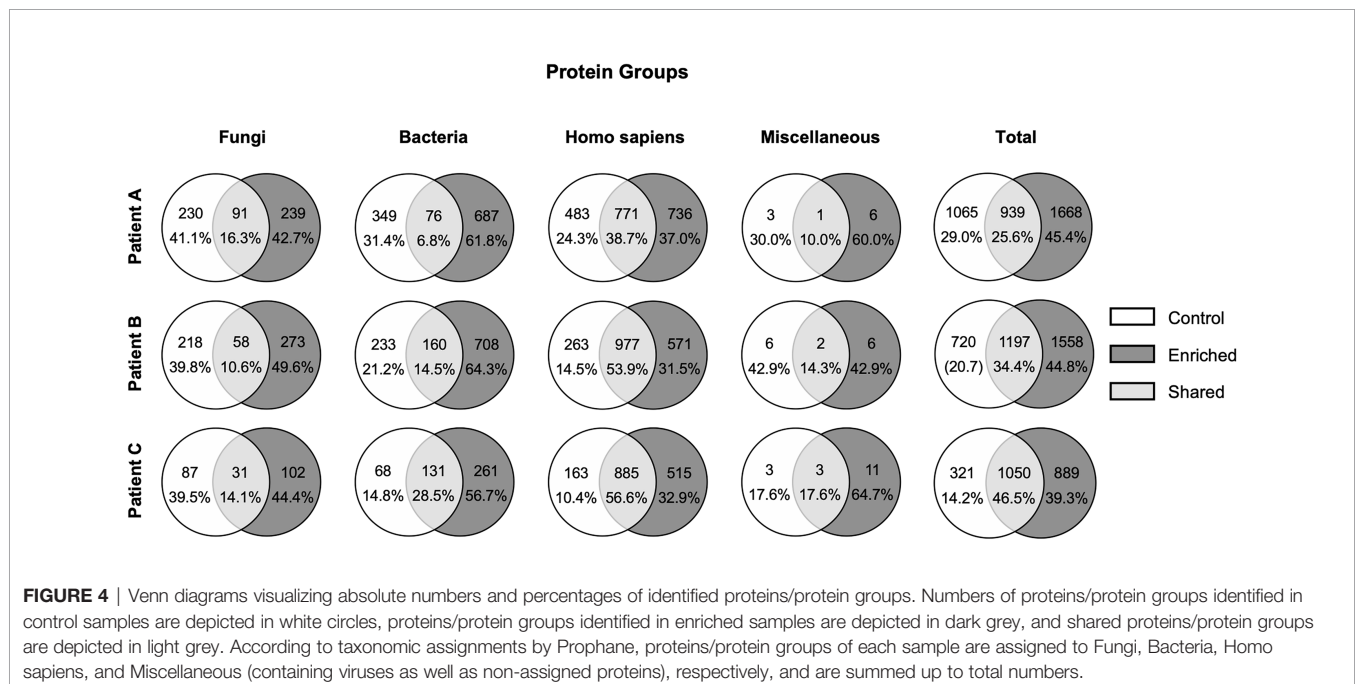
Protein group-based enrichment factors for bacterial proteins range from 1.8-fold (Patient A: 425 total protein groups in control, 763 after enrichment), to 2.0 fold (Patient C: 199 total

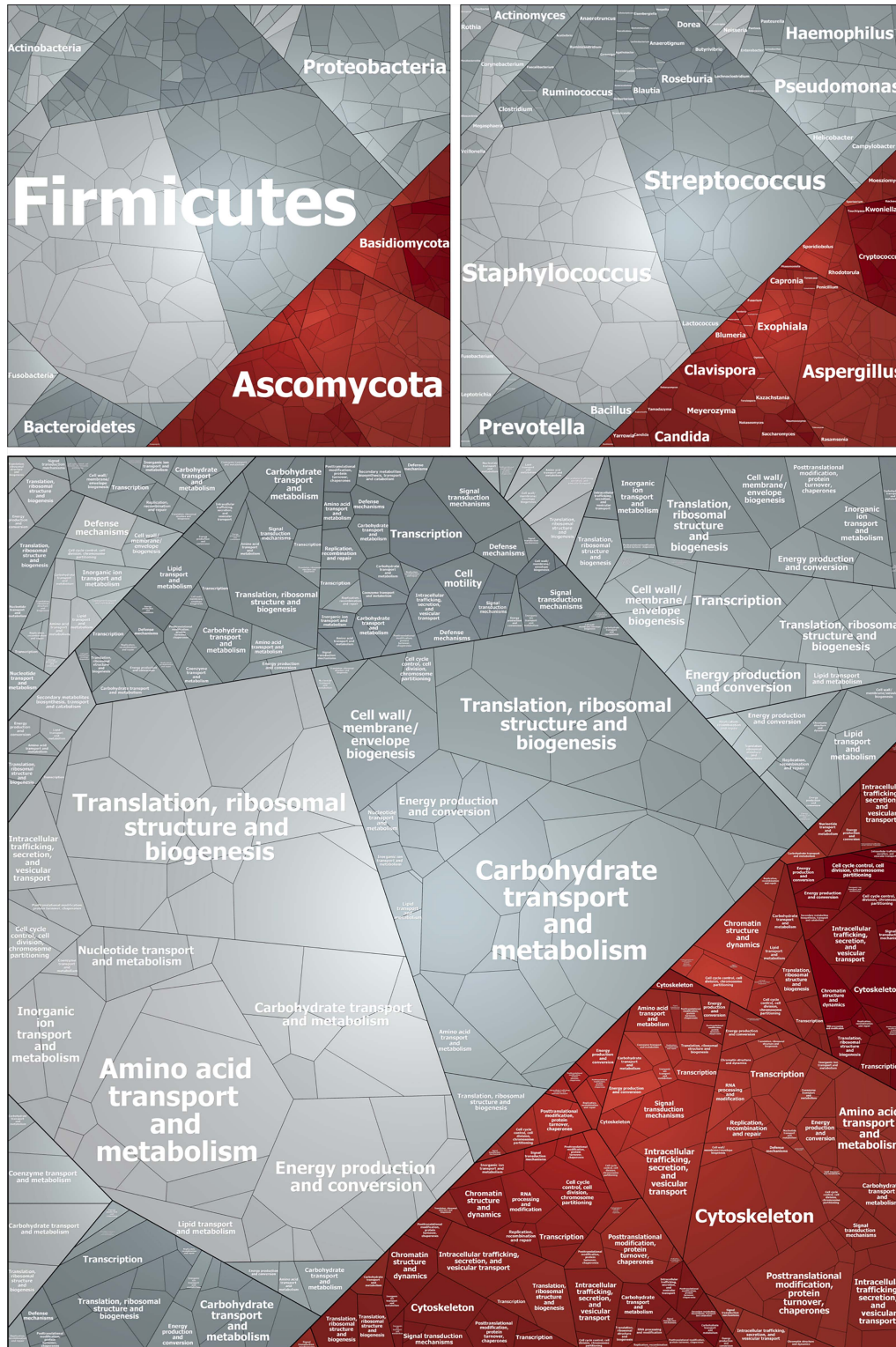
protein groups in control, 392 after enrichment), and 2.2-fold (Patient B: 393 total protein groups in control, 868 after enrichment) (Figures 3B, 4). However, considering that the total number of identified proteins/protein groups is overall higher after enrichment compared to the control, the enrichment factors need to be normalized accordingly. Doing so, normalized protein group-based enrichment factors range from 1.4-fold (Patient A and C) to 1.5-fold (Patient B).

Total numbers of protein groups depicted in Figure 4 indicate a rather small overlap between proteins/protein groups found in the control and the enriched fraction, respectively. This overlap ranges from 6.8% (Patient A), to 14.5% (Patient B) and 28.5% (Patient C). One explanation for this might be that a high proportion of bacterial proteins was lost during the enrichment process. Surprisingly, these potentially lost proteins are only partially annotated as extracellular proteins (e.g. nucleases and toxins, Tables S3–S5), which indicates that a great number of proteins in the extracellular sputum milieu are derived from cell lysis (e.g. proteins belonging to energy metabolism or DNA replication, ribosomal proteins, stress response proteins, Tables S3–S5).

### Evaluation of the Bias Introduced by Microbial Enrichment

In general, every enrichment process introduces a bias. E.g. the composition of microbial proteins in stool changes in the ratio of Firmicutes- and Bacteroidetes-derived proteins after differential centrifugation and in the proportion of extracellular and host proteins (Tanca et al., 2015). This emphasizes the relevance of the chosen processing protocol influencing metaproteome data acquisition. Here, we cannot exclude that we lost big multicellular aggregates/biofilms during the enrichment. To address this problem, it might be useful to analyze different fractions during sputum processing in order to increase bacterial protein





**FIGURE 5** | Voronoi treemap visualizing the taxonomic and functional affiliation of bacterial (grey) and fungal (red) protein/protein groups identified after enrichment in Patient A. Each cell represents a single protein/protein group, which size correlates with NSAF-based protein abundance. Proteins/protein groups are clustered according to Prophan results based on their taxonomic assignment on class level (upper left), genus level (upper right), and based on their functional assignment (lower panel). Proteins of unknown function are excluded from this visualization.



identification coverage. For instance, extracellular proteins can easily be enriched and extracted using Strata-Clean beads as described by (Bonn et al., 2014; Graf et al., 2019) (data not shown) and subsequently analyzed by metaproteomics. However, to exclude that a systematic error is inherent with our enrichment protocol and the abundancies of bacterial proteins are not excessively over- or underrepresenting any bacterial species, we compared the NSAF-based protein abundancies of different species with CFU counts revealing a reasonable relation (**Table 1, Figure 5 and Figures S4, S5**).

Interestingly, according to our metaproteome data, fungal cells/proteins were not enriched (**Figure 3**). This might be explained by the varying cells size of different fungal species in yeast or hyphae form e.g. ranging from 4 to 12  $\mu\text{m}$  in diameter for yeast cells, 1 to 3  $\mu\text{m}$  in diameter and several 100  $\mu\text{m}$  in length for hyphae, and 1 to 5  $\mu\text{m}$  for spores and conidia (Hickey and Read, 2009; Chotirmall and McElvaney, 2014; Thomson et al., 2016; Williams et al., 2016). This means, that small fungal cells would be enriched during our first enrichment step by differential centrifugation, but bigger fungal cells will be depleted during our second enrichment step by filtration (cut off 10  $\mu\text{m}$ ). However, those fungal genera, which are most frequently identified in CF in the literature, match the genera we identified as the most abundant by metaproteomics: namely, *Aspergillus* (prevalence up to 57%), *Candida*, *Blumeria*, *Exophiala*, *Clavispora*, and *Cryptococcus* (**Figure 5 and Figures S4, S5**) (Chotirmall and McElvaney, 2014; Williams et al., 2016; Tracy and Moss, 2018). The genus *Scedosporium* (prevalence ranging from 3.1% to 10.6% (Williams et al., 2016), however, plays a minor role according to our data (**Figure 5 and Figures S4 1, S5**).

## Important Bacterial (Patho-)Physiological Pathways Revealed by Metaproteome and Metabolome Analyses

The total number of identified proteins/protein groups after enrichment differs from patient to patient. In more detail, we identified 2607 proteins/protein groups for Patient A, 2755 for Patient B and the lowest number of 1939 for Patient C (**Figure 4**) – the patient whose microbial lung community is dominated by *P. aeruginosa* and who shows the lowest lung function (**Table 1 and Table S1, Figure S4**). Interestingly, these results are in line with the literature stating an exacerbation/reduced lung function due to *P. aeruginosa* infection, which is caused by the extensive recruitment of neutrophils and concomitant proteolytic digestion of lung tissue and proteins of bacterial pathogens (Sloane et al., 2005; Folkesson et al., 2012). This neutrophil-derived proteolytic digestion in consequence likely leads to a reduced protein identification coverage (and reduced percentage of assigned spectra) as observed in Patient C (**Figure 4 and Figure S3**).

The number of protein/protein groups assigned to the most prominent bacterial genera in CF like *Pseudomonas* (166 in Patient C), *Staphylococcus* (185 in Patient A), *Burkholderia* (408 in Patient B), *Haemophilus* (127 in Patient B), and *Streptococcus* (128 in Patient A) give a first insight into the physiology of these pathogens during CF infection. This means that there is still room for improvement of our enrichment protocol to ultimately increase bacterial protein identification coverage. Notably, the majority of

protein groups assigned to the aforementioned dominant bacterial pathogens are poorly or even uncharacterized, indicating that important host-adaptation strategies of the identified pathogens have so far not been addressed and uncovered experimentally. Protein groups of known function identify various physiological pathways and virulence factors, which are key for bacterial pathogens to establish chronic infections: host immune evasion, anaerobic metabolism, and virulence/antibiotic resistance (Folkesson et al., 2012). We would like to emphasize that out of the 69 proteins/protein groups related to the above mentioned traits (**Figure 6**) 46 proteins/protein groups were exclusively identified in the enriched samples. Only 4 proteins/protein groups were exclusively identified in the control samples, while 19 proteins/protein groups were shared by control and enriched samples. Out of those 19 proteins/protein groups two were slightly less abundant in the enriched samples. Together this underlines the value of our enrichment protocol.

## Oxidative Stress

During host immune response, large numbers of neutrophils are recruited, which fight pathogens by the production of reactive oxygen species (Folkesson et al., 2012; Kamath et al., 2015). Consistently, in Patient A we detected an alkyl hydroperoxide reductase (*Staphylococcus* sp.), a Glutathione S-transferase (*Haemophilus* sp.), an iron-sulfur-cluster repair protein (*Staphylococcus* sp.), glutathione-disulfide reductase (*Streptococcus* sp.), and the molecular chaperones DnaK (*Staphylococcus* sp.) and GroL (various species) as well as the protease ClpP (*Streptococcus* sp.), which all are involved in protein protection, repair, or degradation of proteins and inducible after various stress conditions like oxidative stress (Michta et al., 2014; Ezraty et al., 2017). Moreover, in Patient B *Burkholderia* sp. express the chaperone GroL, the protease ClpB and ClpX as well as catalase/oxidase. In this sample, we additionally identified thioredoxin, glutathione disulfide reductase, peroxiredoxin, and iron/manganese superoxide dismutase expressed by *Burkholderia* sp., *Haemophilus* sp., *Streptococcus* sp., and *Staphylococcus* sp., respectively. In Patient C, we identified the following (oxidative) stress proteins: thioredoxin, peroxiredoxin, thioredoxin-disulfide reductase, DnaK, and ClpB expressed by *Pseudomonas* sp. and *Staphylococcus* sp. (**Tables S3–S5**). Collectively, these data suggest that the most important CF pathogens within our samples might cope with oxidative stress during CF infection, which is in line with the literature (Treffon et al., 2018).

## Oxygen Limitation and pH Homeostasis

The lung environment and especially the CF lung is not considered to be entirely aerobic, due to the viscous character of mucus, oxygen consumption by colonizing microbes, and phagocytes. Rather, oxygen gradients ranging from hypoxic to even anoxic/anaerobic microenvironments characterize the CF-lung (Worlitzsch et al., 2002). Thus, even strict anaerobic bacteria are able to thrive in lungs of CF patients (Filkins and O'Toole, 2015). Consistent with hypoxic and anaerobic conditions, we identified marker-proteins of fermentative metabolism from *Staphylococcus* sp. including lactate

	Patient	TrEMBL Genus	Description	TIGRFAMs / PFAM	Identifier	Quantification (Mean NSAFs)	
						Control	Enriched
<b>(Oxidative) stress</b>	<b>A</b>	Staphylococcus	alkyl hydroperoxide reductase subunit F		TIGR03140	0,0E+00	1,3E-04
		Staphylococcus	iron-sulfur cluster repair di-iron protein		TIGR03652	0,0E+00	2,2E-04
		Staphylococcus	chaperonin GroL		TIGR02348	0,0E+00	4,0E-05
		Staphylococcus	chaperone protein DnaK		TIGR02350	0,0E+00	1,3E-04
		Pseudomonas	chaperonin GroL		TIGR02348	0,0E+00	3,9E-05
		Haemophilus	glutaredoxin-family domain		TIGR02190	0,0E+00	3,0E-04
	<b>B</b>	Streptococcus	chaperonin GroL		TIGR02348	0,0E+00	7,9E-05
		Streptococcus	glutathione-disulfide reductase		TIGR01421	1,0E-04	1,1E-04
		Streptococcus	chaperonin GroL		TIGR02348	1,1E-04	1,3E-03
		Burkholderia	chaperonin GroL		TIGR02348	2,7E-04	1,9E-03
		Burkholderia	ATP-dependent chaperone protein ClpB		TIGR03346	3,4E-05	3,7E-04
		Burkholderia	ATP-dependent Clp protease, ATP-binding subunit ClpX		TIGR00382	0,0E+00	3,2E-04
	<b>C</b>	Burkholderia	catalase/peroxidase HPI		TIGR00198	3,9E-05	4,1E-04
		Burkholderia	thioredoxin		TIGR01068	4,0E-04	3,7E-04
		Haemophilus	thioredoxin		TIGR01068	8,1E-04	1,0E-03
		Streptococcus	glutathione-disulfide reductase		TIGR01421	1,3E-04	0,0E+00
		Staphylococcus	peroxiredoxin		TIGR03137	0,0E+00	2,9E-04
		Staphylococcus	Iron/manganese superoxide dismutases		PF00081	4,4E-04	0,0E+00
<b>Oxygen limitation and pH homeostasis</b>	<b>A</b>	Staphylococcus	L-lactate dehydrogenase		TIGR01771	0,0E+00	7,6E-04
		Staphylococcus	arginine deiminase		TIGR01078	0,0E+00	2,4E-04
		Staphylococcus	ornithine carbamoyltransferase		TIGR00658	0,0E+00	3,0E-04
		Staphylococcus	carbamate kinase		TIGR00746	0,0E+00	2,3E-04
		Staphylococcus	formate acetyltransferase		TIGR01255	0,0E+00	9,0E-04
		Streptococcus	formate acetyltransferase		TIGR01255	0,0E+00	2,8E-05
	<b>B</b>	Streptococcus	acetate kinase		TIGR00016	0,0E+00	7,2E-05
		Streptococcus	L-lactate dehydrogenase		TIGR01771	0,0E+00	3,0E-04
		Staphylococcus	formate acetyltransferase		TIGR01255	0,0E+00	6,8E-05
		Burkholderia	acetate kinase		TIGR00016	7,2E-05	0,0E+00
		Burkholderia	nitrate reductase, alpha subunit		TIGR01580	0,0E+00	4,3E-05
		Haemophilus	ornithine carbamoyltransferase		TIGR00658	0,0E+00	2,4E-04
<b>C</b>	Haemophilus	urease accessory protein UreG		TIGR00101	0,0E+00	1,2E-04	
	Staphylococcus	ornithine carbamoyltransferase		TIGR00658	1,3E-04	6,0E-04	
	Pseudomonas	ornithine carbamoyltransferase		TIGR00658	0,0E+00	5,0E-04	
	Pseudomonas	carbamate kinase		TIGR00746	0,0E+00	1,1E-04	
	Pseudomonas	arginine deiminase		TIGR01078	5,0E-04	1,6E-04	
	<b>Nutrient limitation</b>	<b>A</b>	Staphylococcus	oligoendopeptidase F		TIGR00181	0,0E+00
Corynebacterium			ATP-dependent metallopeptidase HfIB		TIGR01241	0,0E+00	6,5E-05
<b>B</b>		Burkholderia	TonB-dependent siderophore receptor		TIGR01783	0,0E+00	5,1E-04
		Burkholderia	ATP-dependent metallopeptidase HfIB		TIGR01241	0,0E+00	8,5E-05
		Haemophilus	endopeptidase La		TIGR00763	0,0E+00	9,9E-05
		Haemophilus	aminopeptidase N		TIGR02414	0,0E+00	6,2E-05
<b>C</b>		Pseudomonas	signal peptide peptidase SppA, 36K type		TIGR00706	0,0E+00	2,1E-04
		Pseudomonas	ATP-dependent chaperone protein ClpB		TIGR03346	0,0E+00	2,4E-04
		Pseudomonas	Clp protease		PF00574	0,0E+00	1,4E-04
		Pseudomonas	TonB-dependent siderophore receptor		TIGR01783	0,0E+00	2,0E-03
		Pseudomonas	Siderophore-interacting protein		PF04954	0,0E+00	2,1E-04
		Pseudomonas	TonB-dependent heme/hemoglobin receptor family protein		TIGR01785	3,3E-04	7,5E-04
<b>Virulence</b>	<b>A</b>	Staphylococcus	Leukocidin/Hemolysin toxin family		PF07968	1,3E-04	1,8E-03
		Staphylococcus	staphylococcal accessory regulator family		TIGR01889	0,0E+00	4,0E-04
		Streptococcus	adhesin isopeptide-forming adherence domain		TIGR04228	0,0E+00	2,8E-05
	<b>B</b>	Staphylococcus	Leukocidin/Hemolysin toxin family		PF07968	0,0E+00	4,9E-04
		Burkholderia	type VI secretion system effector, Hcp1 family		TIGR03344	6,7E-04	4,6E-04
		Burkholderia	type VI secretion system Vgr family protein		TIGR03361	3,2E-05	0,0E+00
		Burkholderia	Putative type VI secretion system Rhs element Vgr		PF13296	0,0E+00	3,1E-05
		Burkholderia	type VI secretion ATPase, ClpV1 family		TIGR03345	0,0E+00	3,0E-05
		Burkholderia	efflux transporter, RND family, MFP subunit		TIGR01730	1,4E-04	9,6E-04
	<b>C</b>	Burkholderia	efflux transporter, outer membrane factor (OMF) lipoprotein		TIGR01845	1,1E-04	7,5E-04
		Haemophilus	Extended Signal Peptide of Type V secretion system		PF13018	7,8E-05	1,1E-04
		Haemophilus	efflux transporter, RND family, MFP subunit		TIGR01730	0,0E+00	1,4E-04
		Haemophilus	efflux transporter, outer membrane factor (OMF) lipoprotein		TIGR01845	0,0E+00	1,8E-04
		Staphylococcus	Leukocidin/Hemolysin toxin family		PF07968	5,3E-04	2,5E-03
		Haemophilus	efflux transporter, RND family, MFP subunit		TIGR01730	0,0E+00	8,5E-05
	<b>C</b>	Pseudomonas	Protein export membrane protein		PF02355	0,0E+00	1,9E-04
		Pseudomonas	efflux transporter, outer membrane factor (OMF) lipoprotein		TIGR01845	0,0E+00	5,5E-04

**FIGURE 6** | Selected Proteins/Protein Groups of important physiological pathways showing high abundances. Each row represents one protein/protein group assigned to (from left to right): a physiological pathway, a patient sample, a genus, a functional description, an identifier, a NSAF-based quantitative value (mean of two technical replicates) of the control and the enriched sample. Yellow bar graphs reflect NSAF values for an improved results visualization and eased results extraction for the reader.

dehydrogenase, formate acetyltransferase and acetate kinase in all three samples. In the samples of Patient A and Patient B we further identified formate acetyltransferase, L-lactate dehydrogenase and acetate kinase, assigned to the genera *Streptococcus*, *Staphylococcus*, and *Burkholderia*, respectively (**Figure 6** and **Tables S3–S5**). A further metabolic strategy to overcome oxygen limitation conserved in many pathogens is the fermentation of arginine via the arginine deiminase pathway (Lindgren et al., 2014). According to our metaproteome data, arginine deiminase, ornithine carbamoyltransferase, and carbamate kinase, are highly abundant and identified in multiple genera: *Staphylococcus* in Patient A and Patient C, *Haemophilus* in Patient B, and *Pseudomonas* in Patient C (**Figure 6** and **Tables S3–S5**). In support of our metaproteome data, we identified ornithine (Patient A-C) and citrulline (Patient B), key metabolites of the arginine deiminase pathway, by metabolic footprint analysis (**Figure S6**) (Lindgren et al., 2014). Clear induction of the arginine deiminase pathway suggests that not only ATP production, but also raise in pH due to production of ammonia is important for the pathogen to counteract acidification upon fermentation and likely upon phagocytosis by immune cells (Li et al., 2000; Beenken et al., 2004; Resch et al., 2005; Filkins and O'Toole, 2015). As evidence for anaerobic metabolism and acidification, we identified lactic acid as a fermentation product in all three samples (**Figure S6**). Moreover, we identified the urease accessory protein UreG from *Haemophilus* sp. in Patient B, which is a protein of the urease pathway and similarly to the arginine pathway is involved in pH homeostasis (**Figure 6** and **Table S4**) (Li et al., 2000). Notably, urea as the substrate of the urease pathway was also detected in Patient A and Patient B (**Figure S6**). Finally, we identified the alpha subunit of the nitrate reductase from *Burkholderia* sp. in Patient B as part of the nitrate respiration pathway (**Figure 6** and **Table S4**) (Filkins and O'Toole, 2015) utilizing nitrate as an alternative terminal electron acceptor.

### Nutrient Limitation

The competition for nutrients within the CF airways is an important selective pressure influencing the composition of the CF community. For example *Pseudomonads* and *Streptococci* are able to efficiently utilize amino acids, organic acids and alcohols (partly produced by other community members) leading to high growth rates within the lung (Yang et al., 2008; Henson et al., 2019; La Rosa and Molin, 2019). Notably, our metabolome data revealed various amino acids (and organic acids) within all three sputum samples (**Figure S6**), which supports that amino acids are the major carbon- and nitrogen source in CF sputum (Palmer et al., 2005; La Rosa and Molin, 2019). Moreover, we found evidence that these amino acids could result from the hydrolytic activity of a variety of different proteases and peptidases (Kamath et al., 2015; Quinn et al., 2019). Currently, proteases of human origin are believed to be the key players responsible for proteolytic digestion in the CF airways. Specifically, human neutrophil-derived elastase and cathepsins are considered to be the most abundant and potent proteases (Voynow et al., 2008). Although we identified these human proteases among the most abundant in our proteome

data, we additionally found a great number of proteases and peptidases of bacterial origin. For example, we identified a staphylococcal oligopeptidase in Patient A, the metallopeptidase HflB, peptidase Do, and protease HslVU from *Burkholderia* sp. and *Streptococcus* sp. as well as endopeptidase La, and aminopeptidase N from *Haemophilus* sp. in Patient B, and multiple Clp protease proteins in Patient C (**Figure 6** and **Table S3–S5**). In fact, by using NSAFs to calculate the contribution of all human and all bacterial proteases and peptidases to the total proteome mass, we found that the human proteases/peptidases in the non-enriched control accounts for 2.15% of the total proteome mass and bacterial proteases/peptidases in the enriched sample for 0.84% (**Figure 6** and **Tables S3–S5**). This suggests that the role of bacterial proteases and peptidases in the CF pathophysiology is larger than previously acknowledged (Kamath et al., 2015).

Another vital nutrient during CF infection is iron, which is needed by bacteria as a cofactor in essential metabolic enzymes (Reid et al., 2009). Although the iron concentration within the CF airway is relatively high compared to other human body sites, pathogens within the CF lung fight for iron by sequestering iron chelating siderophores and proteases degrading transferrin, lactoferrin, and heme-containing proteins like hemoglobin and myoglobin (Reid et al., 2009; Runyen-Janecky, 2013; Treffon et al., 2018). This fight for iron is well reflected by our metaproteome data, e.g. revealing multiple TonB-dependent siderophore receptors and TonB-dependent heme/hemoglobin receptor family proteins of *Burkholderia* sp. and *Pseudomonas* sp. in Patient B and Patient C, respectively (**Figure 6** and **Tables S3–S5**).

### Virulence Factors

Among the identified microbial proteins many important virulence factors were detected. For instance, staphylococcal leukocidin, an immune evasion protein that mediates lysis of leukocytes (Scherr et al., 2015) was very abundant in all three samples, underlining the ongoing battle between the host immune system and the pathogen within the CF-lung (**Figure 6** and **Tables S3–S5**). However, our proteomic data did not provide support for presence of microbial biofilms in the samples, which is considered a major microbial phenotype during infection (Folkesson et al., 2012; Filkins and O'Toole, 2015; Kovach et al., 2017). Only a streptococcal adhesin (Patient A), one adhesin of *Pseudomonas* sp. (Patient C), and the aforementioned staphylococcal leukocidins (Patient A-C), which have the potential to moonlight as a stabilizing extracellular matrix component under acidic conditions, were detected (**Figure 6** and **Tables S3–S5**) (Graf et al., 2019). We cannot entirely exclude that the lack of biofilm-related proteins in our sample is related to the enrichment procedure, as large cellular clusters were cleared from the sample during the first centrifugation and/or the filtration step.

Another well described concept of bacterial virulence are secretion systems, which transport effectors or DNA across membranes to manipulate the physiology of host cells or competing bacteria (Voth et al., 2012). Indeed, we identified



multiple proteins of a Type VI secretion system of *Burkholderia* sp. in Patient B. Furthermore, we found an Extended Signal Peptide of Type V secretion system protein in *Haemophilus* sp. (Patient B) (**Figure 6** and **Tables S3–S5**). This emphasizes the importance of secretion systems, especially for *Burkholderia* sp., during successful CF infection. Finally, we found multiple efflux transporter mediating antibiotic resistance (Li and Nikaido, 2009) in *Burkholderia* sp., *Haemophilus* sp., *Pseudomonas* sp. (Patient B and Patient C), which likely reflect a response towards the antimicrobial therapy that the three investigated patients underwent (**Figure 6**, **Table 1** and **Tables S4, S5**).

## Limitations of Our Study

The detailed analysis of three sputum samples provided novel insights into the microbial pathophysiology within the CF lung and revealed high expression of the arginine deiminase pathway and multiple proteases, demonstrating the applicability of our protocol. We are aware that larger sample numbers will be required to validate the significance of these findings. Future studies should not only include larger sample numbers but should also consider specific mutations of the CFTR gene and the individual patient treatment regimens to account for the high in-between variability of CF patients (Tanca et al., 2014). Additionally, larger numbers of biological and technical replicates will also increase protein coverage. Including a proteomic analysis of the supernatant of the differential centrifugation step II could provide further insight into the microbial secretome (Bonn et al., 2014; Graf et al., 2019) (**Figure 1**). An additional question that deserves further investigation is the small overlap of protein groups between enriched and non-enriched samples (**Figure 4**) (Starr et al., 2017). Moreover, to increase the coverage of our metabolome analysis (La Rosa and Molin, 2019), in addition to PBS-buffer based metabolite extraction combined with GC-MS analyses resulting in the identification of 52 metabolites, further metabolite extraction and analysis methods e.g. according to (Yang et al., 2012; Quinn et al., 2015) could be beneficial. Finally, measuring absolute metabolite concentrations will improve the comparison of metabolite levels between different studies.

## CONCLUSIONS

We established an innovative, reliable, and easy-to-handle sputum processing protocol for *in vivo* metaproteome analyses. With this protocol in hand, we provide the first *in vivo* study of microbial CF sputum communities combining metaproteomic and metabolomic analyses supported by 16S sequencing and microscopic data as a proof of concept. Our metaproteome data show that we were able to enrich bacterial proteins by a maximum factor of 2.7, thereby increasing protein identification coverage to a level, which provides novel valuable insights into bacterial CF-lung pathophysiology. Our early data, which are derived from just 3 sputum samples proof the applicability of our protocol but do lack statistical power. However, they still indicate that the infecting bacteria might be coping with oxygen and nutrient limitation as well as

oxidative stress and the human immune system, respectively. Our early data also provide evidence that the arginine deiminase pathway as well as bacterial proteases play an underappreciated role in CF pathophysiology.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available in NCBI (using accession number PRJNA741386 for 16S sequencing data) and within the ProteomeXchange consortium *via* the PRIDE partner repository (using dataset identifier PXD025134 for metaproteome data).

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional ethics review board Münster, Germany (2010-155-f-S). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AG, JP-F, and KR were responsible for the study conceptualization. BK carried out the prospective study including CF sputum sampling, collection of clinical data and microbiological work-up of sputum specimens. AG and JS performed the experiments to develop the sputum processing protocol. 16S sequencing analyses were performed by DP. AG, TS, and DB performed metaproteome analyses and ML and MW were responsible for metabolome analyses. DP analyzed 16S sequencing data, while AG analyzed metaproteome and metabolome data. AG, JP-F and KR wrote the manuscript, which was critically edited by all other co-authors. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was funded by the German research foundation (<https://www.dfg.de/en/>) Collaborative Research Center Transregio 34, subprojects A3 to KR, A8 to JP-F, C7 to BK, Z2 to DB, and Z4 to ML. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## ACKNOWLEDGMENTS

We are very grateful to Tobias Kroniger, Stephan Fuchs, Daniela Zühlke, and Dirk Albrecht for help with the MS analyses and database search, respectively.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2021.724569/full#supplementary-material>

## REFERENCES

- Becher, D., Hempel, K., Sievers, S., Zühlke, D., Pané-Farré, J., Otto, A., et al. (2009). A Proteomic View of an Important Human Pathogen – Towards the Quantification of the Entire *Staphylococcus Aureus* Proteome. *PLoS One* 4, e8176–e8112. doi: 10.1371/journal.pone.0008176
- Beenken, K. E., Dunman, P. M., McAleese, F., Macapagal, D., Murphy, E., Projan, S. J., et al. (2004). Global Gene Expression in *Staphylococcus Aureus* Biofilms. *J. Bacteriol.* 186, 4665–4684. doi: 10.1128/JB.186.14.4665-4684.2004
- Bernhardt, J., Funke, S., Hecker, M., and Siebourg, J. (2009). *Visualizing Gene Expression Data via Voronoi Treemaps*. 233–241.
- Bicart-See, A., Rottman, M., Cartwright, M., Seiler, B., Gamini, N., Rodas, M., et al. (2016). Rapid Isolation of *Staphylococcus Aureus* Pathogens From Infected Clinical Samples Using Magnetic Beads Coated With Fc-Mannose Binding Lectin. *PLoS One* 11, e0156287–12. doi: 10.1371/journal.pone.0156287
- Bonn, F., Bartel, J., Büttner, K., Hecker, M., Otto, A., and Becher, D. (2014). Picking Vanished Proteins From the Void: How to Collect and Ship/Share Extremely Dilute Proteins in a Reproducible and Highly Efficient Manner. *Anal. Chem.* 86, 7421–7427. doi: 10.1021/ac501189j
- Camarinha-Silva, A., Jauregui, R., Chaves-Moreno, D., Oxley, A. P. A., Schaumburg, F., Becker, K., et al. (2014). Comparing the Anterior Nare Bacterial Community of Two Discrete Human Populations Using Illumina Amplicon Sequencing. *Environ. Microbiol.* 16, 2939–2952. doi: 10.1111/1462-2920.12362
- Chmiel, J. F., and Davis, P. B. (2003). State of the Art: Why do the Lungs of Patients With Cystic Fibrosis Become Infected and Why Can't They Clear the Infection? *Respir. Res.* 4, 8. doi: 10.1186/1465-9921-4-8
- Chotirmall, S. H., and McElvaney, N. G. (2014). Fungi in the Cystic Fibrosis Lung: Bystanders or Pathogens? *Int. J. Biochem. Cell Biol.* 52, 161–173. doi: 10.1016/j.biocel.2014.03.001
- Chourey, K., Jansson, J., VerBerkmoes, N., Shah, M., Chavarria, K. L., Tom, L. M., et al. (2010). Direct Cellular Lysis/Protein Extraction Protocol for Soil Metaproteomics. *J. Proteome Res.* 9, 6615–6622. doi: 10.1021/pr100787q
- Cohen, T. S., and Prince, A. (2012). Cystic Fibrosis: A Mucosal Immunodeficiency Syndrome. *Nat. Med.* 18, 509–519. doi: 10.1038/nm.2715
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2013). Ribosomal Database Project: Data and Tools for High Throughput rRNA Analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244
- Debyser, G., Mesuere, B., Clement, L., Van de Weygaert, J., Van Hecke, P., Duytschaever, G., et al. (2016). Faecal Proteomics: A Tool to Investigate Dysbiosis and Inflammation in Patients With Cystic Fibrosis. *J. Cyst. Fibros.* 15, 242–250. doi: 10.1016/j.jcf.2015.08.003
- Eiserich, J. P., Yang, J., Morrissey, B. M., Hammock, B. D., and Cross, C. E. (2012). Omics Approaches in Cystic Fibrosis Research: A Focus on Oxylipin Profiling in Airway Secretions. *Ann. N. Y. Acad. Sci.* 1259, 1–9. doi: 10.1111/j.1749-6632.2012.06580.x
- Ezraty, B., Gennaris, A., Barras, F., and Collet, J.-F. (2017). Oxidative Stress, Protein Damage and Repair in Bacteria. *Nat. Rev. Microbiol.* 15, 1–12. doi: 10.1038/nrmicro.2017.26
- Filkins, L. M., Graber, J. A., Olson, D. G., Dolben, E. L., Lynd, L. R., Bhujii, S., et al. (2015). Coculture of *Staphylococcus Aureus* With *Pseudomonas Aeruginosa* Drives *S. Aureus* Towards Fermentative Metabolism and Reduced Viability in a Cystic Fibrosis Model. *J. Bacteriol.* 197, 2252–2264. doi: 10.1128/JB.00059-15
- Filkins, L. M., and O'Toole, G. A. (2015). Cystic Fibrosis Lung Infections: Polymicrobial, Complex, and Hard to Treat. *PLoS Pathog.* 11, e1005258–8. doi: 10.1371/journal.ppat.1005258
- Folkesson, A., Jelsbak, L., Yang, L., Johansen, H. K., Ciofu, O., and Molin, S. (2012). Adaptation of *Pseudomonas Aeruginosa* to the Cystic Fibrosis Airway: An Evolutionary Perspective. *Nat. Rev. Microbiol.* 10, 841–851. doi: 10.1038/nrmicro2907
- Fuchs, H. J., Borowitz, D. S., Christiansen, D. H., Morris, E. M., Nash, M. L., Ramsey, B. W., et al. (1994). Effect of Aerosolized Recombinant Human DNase on Exacerbations of Respiratory Symptoms and on Pulmonary Function in Patients With Cystic Fibrosis. The Pulmozyme Study Group. *N. Engl. J. Med.* 331, 637–642. doi: 10.1056/NEJM199409083311003
- Fu, Y. R., Yi, Z. J., Guan, S. Z., Zhang, S. Y., and Li, M. (2012). Proteomic Analysis of Sputum in Patients With Active Pulmonary Tuberculosis. *Clin. Microbiol. Infect.* 18, 1241–1247. doi: 10.1111/j.1469-0691.2012.03824.x
- Gilligan, P. H. (2014). Infections in Patients With Cystic Fibrosis. *Clin. Lab. Med.* 34, 197–217. doi: 10.1016/j.cll.2014.02.001
- Goerke, C., and Wolz, C. (2010). Adaptation of *Staphylococcus Aureus* to the Cystic Fibrosis Lung. *Int. J. Med. Microbiol.* 300, 520–525. doi: 10.1016/j.ijmm.2010.08.003
- Graf, A. C., Leonard, A., Schäuble, M., Rieckmann, L. M., Hoyer, J., Maaß, S., et al. (2019). Virulence Factors Produced by *Staphylococcus Aureus* Biofilms Have a Moonlighting Function Contributing to Biofilm Integrity. *Mol. Cell Proteomics* 18, 1036–1053. doi: 10.1074/mcp.RA118.001120
- Henson, M. A., Orazi, G., Phalak, P., and O'Toole, G. (2019). Metabolic Modeling of Cystic Fibrosis Airway Communities Predicts Mechanisms of Pathogen Dominance: Supplemental Tables 1–43. *mSystems* 4. doi: 10.1101/520619
- Herzog, S., Dach, F., de Buhr, N., Niemann, S., Schlagowski, J., Chaves-Moreno, D., et al. (2019). High Nuclease Activity of Long Persisting *Staphylococcus Aureus* Isolates Within the Airways of Cystic Fibrosis Patients Protects Against NET-Mediated Killing. *Front. Immunol.* 10, 2552. doi: 10.3389/fimmu.2019.02552
- Hevia, A., Delgado, S., Margolles, A., and Sánchez, B. (2016). Application of Density Gradient for the Isolation of the Fecal Microbial Stool Component and the Potential Use Thereof. *Sci. Rep.* 5, 16807–9. doi: 10.1038/srep16807
- Hickey, P. C., and Read, N. D. (2009). Imaging Living Cells of *Aspergillus In Vitro*. *Med. Mycol.* 47, S110–S119. doi: 10.1080/13693780802546541
- Hinzke, T., Kouris, A., Hughes, R.-A., Strous, M., and Kleiner, M. (2019). More Is Not Always Better: Evaluation of 1D and 2D-LC-MS/MS Methods for Metaproteomics. *Front. Microbiol.* 10, 238. doi: 10.3389/fmicb.2019.00238
- Kamath, K. S., Kumar, S. S., Kaur, J., Venkatakrishnan, V., Paulsen, I. T., Nevalainen, H., et al. (2015). Proteomics of Hosts and Pathogens in Cystic Fibrosis. *Prot. Clin. Appl.* 9, 134–146. doi: 10.1002/prca.201400122
- Kovach, K., Davis-Fields, M., Irie, Y., Jain, K., Doorwar, S., Vuong, K., et al. (2017). Evolutionary Adaptations of Biofilms Infecting Cystic Fibrosis Lungs Promote Mechanical Toughness by Adjusting Polysaccharide Production. *NPJ Biofilms Microbiomes*, 3, 1. doi: 10.1038/s41522-016-0007-9
- Laemmli, U. K. (1970). Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0
- La Rosa, J., and Molin, (2019). Adapting to the Airways: Metabolic Requirements of *Pseudomonas Aeruginosa* During the Infection of Cystic Fibrosis Patients. *Metabolites* 9, 234–215. doi: 10.3390/metabo9100234
- Lassek, C., Burghartz, M., Chaves-Moreno, D., Otto, A., Hentschker, C., Fuchs, S., et al. (2015). A Metaproteomics Approach to Elucidate Host and Pathogen Protein Expression During Catheter-Associated Urinary Tract Infections (CAUTIs). *Mol. Cell Proteomics* 14, 989–1008. doi: 10.1074/mcp.M114.043463
- Li, Y. H., Chen, Y. Y. M., and Burne, R. A. (2000). Regulation of Urease Gene Expression by *Streptococcus Salivarius* Growing in Biofilms. *Environ. Microbiol.* 2, 169–177. doi: 10.1046/j.1462-2920.2000.00088.x
- Liebermeister, W., Noor, E., Flamholz, A., Davidi, D., Bernhardt, J., and Milo, R. (2014). Visual Account of Protein Investment in Cellular Functions. *PNAS* 111, 8488–8493. doi: 10.1073/pnas.1314810111
- Liessi, N., Pedemonte, N., Armirotti, A., and Braccia, C. (2020). Proteomics and Metabolomics for Cystic Fibrosis Research. *IJMS* 21, 5439. doi: 10.3390/ijms21155439
- Lindgren, J. K., Thomas, V. C., Olson, M. E., Chaudhari, S. S., Nuxoll, A. S., Schaeffer, C. R., et al. (2014). Arginine Deiminase in *Staphylococcus Epidermidis* Functions to Augment Biofilm Maturation Through pH Homeostasis. *J. Bacteriol.* 196, 2277–2289. doi: 10.1128/JB.00051-14
- Li, X.-Z., and Nikaido, H. (2009). Efflux-Mediated Drug Resistance in Bacteria. *Drugs* 69, 1555–1623. doi: 10.2165/11317030-000000000-00000
- Liu, Y., Schulze-Makuch, D., de Vera, J.-P., Cockell, C., Leya, T., Baqué, M., et al. (2018). The Development of an Effective Bacterial Single-Cell Lysis Method Suitable for Whole Genome Amplification in Microfluidic Platforms. *Micromachines (Basel)* 9, 367. doi: 10.3390/mi9080367
- Lyczak, J. B., Cannon, C. L., and Pier, G. B. (2002). Lung Infections Associated With Cystic Fibrosis. *Clin. Microbiol. Rev.* 15, 194–222. doi: 10.1128/CMR.15.2.194-222.2002
- Michta, E., Ding, W., Zhu, S., Blin, K., Ruan, H., Wang, R., et al. (2014). Proteomic Approach to Reveal the Regulatory Function of Aconitase AcnA in Oxidative Stress Response in the Antibiotic Producer *Streptomyces Viridochromogenes* Tü494. *PLoS One* 9, e87905–e87910. doi: 10.1371/journal.pone.0087905
- Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988). Improved Staining of Proteins in Polyacrylamide Gels Including Isoelectric Focusing Gels With Clear Background at Nanogram Sensitivity Using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9, 255–262. doi: 10.1002/elps.1150090603
- O'Sullivan, B. P., and Freedman, S. D. (2009). Cystic Fibrosis. *Lancet* 373, 1891–1904. doi: 10.1016/S0140-6736(09)60327-5

- Otto, M. (2008). Staphylococcal Biofilms. *Curr. Top. Microbiol. Immunol.* 322, 207–228. doi: 10.1007/978-3-540-75418-3\_10
- Palmer, K. L., Aye, L. M., and Whiteley, M. (2007). Nutritional Cues Control Pseudomonas Aeruginosa Multicellular Behavior in Cystic Fibrosis Sputum. *J. Bacteriol.* 189, 8079–8087. doi: 10.1128/JB.01138-07
- Palmer, K. L., Mashburn, L. M., Singh, P. K., and Whiteley, M. (2005). Cystic Fibrosis Sputum Supports Growth and Cues Key Aspects of Pseudomonas Aeruginosa Physiology. *J. Bacteriol.* 187, 5267–5277. doi: 10.1128/JB.187.15.5267-5277.2005
- Quinn, R. A., Adem, S., Mills, R. H., Comstock, W., Goldasich, L. D., Humphrey, G., et al. (2019). Neutrophilic Proteolysis in the Cystic Fibrosis Lung Correlates With a Pathogenic Microbiome 1–13. *Microbiome* 7, 23. doi: 10.1186/s40168-019-0636-3
- Quinn, R. A., Phelan, V. V., Whiteson, K. L., Garg, N., Bailey, B. A., Lim, Y. W., et al. (2015). Microbial, Host and Xenobiotic Diversity in the Cystic Fibrosis Sputum Metabolome. *ISME J.* 10, 1–16. doi: 10.1038/ismej.2015.207
- Rath, S., Heidrich, B., Pieper, D. H., and Vital, M. (2017). Uncovering the Trimethylamine-Producing Bacteria of the Human Gut Microbiota *Microbiome* 5, 54–14. doi: 10.1186/s40168-017-0271-9
- Ratjen, F. A. (2009). Cystic Fibrosis: Pathogenesis and Future Treatment Strategies. *Respir. Care* 54, 595–605. doi: 10.4187/aarc0427
- Reid, D. W., Anderson, G. J., and Lamont, I. L. (2009). Role of Lung Iron in Determining the Bacterial and Host Struggle in Cystic Fibrosis. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 297, L795–L802. doi: 10.1152/ajplung.00132.2009
- Resch, A., Rosenstein, R., Nerz, C., and Götz, F. (2005). Differential Gene Expression Profiling of Staphylococcus Aureus Cultivated Under Biofilm and Planktonic Conditions. *Appl. Environ. Microbiol.* 71, 2663–2676. doi: 10.1128/AEM.71.5.2663-2676.2005
- Rogers, G. B., Carroll, M., Hoffman, L., Walker, A., Fine, D., and Bruce, K. (2014). Comparing the Microbiota of the Cystic Fibrosis Lung and Human Gut. *Gut Microbes* 1, 85–93. doi: 10.4161/gmic.1.2.11350
- Runyen-Janecky, L. J. (2013). Role and Regulation of Heme Iron Acquisition in Gram-Negative Pathogens. *Front Cell Infect Microbiol* 3, 55. doi: 10.3389/fcimb.2013.00055/abstract
- Scherr, T. D., Hanke, M. L., Huang, O., James, D. B. A., Horswill, A. R., Bayles, K. W., et al. (2015). Staphylococcus Aureus Biofilms Induce Macrophage Dysfunction Through Leukocidin AB and Alpha-Toxin. *mBio* 6, e01021–15–13. doi: 10.1128/mBio.01021-15
- Schiebhofer, H., Schallert, K., Renard, B. Y., Trappe, K., Schmid, E., Benndorf, D., et al. (2020). A Complete and Flexible Workflow for Metaproteomics Data Analysis Based on MetaProteomeAnalyzer and Prophan. *Nat. Protoc.* 15, 3212–3239. doi: 10.1038/s41596-020-0368-7
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing Mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09
- Schultz, D., Zühlke, D., Bernhardt, J., Francis, T. B., Albrecht, D., Hirschfeld, C., et al. (2020). An Optimized Metaproteomics Protocol for a Holistic Taxonomic and Functional Characterization of Microbial Communities From Marine Particles. *Environ. Microbiol. Rep.* 13, 290210. doi: 10.1111/1758-2229.12842
- Schulz, C., Schütte, K., Koch, N., Vilchez-Vargas, R., Wos-Oxley, M. L., Oxley, A. P. A., et al. (2018). The Active Bacterial Assemblages of the Upper GI Tract in Individuals With and Without Helicobacter Infection. *Gut* 67, 216–225. doi: 10.1136/gutjnl-2016-312904
- Schwartzbeck, B., Birtel, J., Treffon, J., Langhanki, L., Mellmann, A., Kale, D., et al. (2016). Dynamic *In Vivo* Mutations Within the Ica Operon During Persistence of Staphylococcus Aureus in the Airways of Cystic Fibrosis Patients. *PLoS Pathog.* 12, e1006024. doi: 10.1371/journal.ppat.1006024
- Shak, S., Capon, D. J., Hellmiss, R., Marsters, S. A., and Baker, C. L. (1990). Recombinant Human DNase I Reduces the Viscosity of Cystic Fibrosis Sputum. *PNAS* 87, 9188–9192. doi: 10.1073/pnas.87.23.9188
- Sloane, A. J., Lindner, R. A., Prasad, S. S., Sebastian, L. T., Pedersen, S. K., Robinson, M., et al. (2005). Proteomic Analysis of Sputum From Adults and Children With Cystic Fibrosis and From Control Subjects. *Am. J. Respir. Crit. Care Med.* 172, 1416–1426. doi: 10.1164/rccm.200409-1215OC
- Son, M. S., Matthews, W. J., Kang, Y., Nguyen, D. T., and Hoang, T. T. (2007). *In Vivo* Evidence of Pseudomonas Aeruginosa Nutrient Acquisition and Pathogenesis in the Lungs of Cystic Fibrosis Patients. *Infect. Immun.* 75, 5313–5324. doi: 10.1128/IAI.01807-06
- Starke, R., Jehmlich, N., Alfaro, T., Dohnalkova, A., Capek, P., Bell, S. L., et al. (2019). Incomplete Cell Disruption of Resistant Microbes. *Sci. Rep.* 9, 5618–5. doi: 10.1038/s41598-019-42188-9
- Starr, A. E., Deeke, S. A., Li, L., Zhang, X., Daoud, R., Ryan, J., et al. (2017). Proteomic and Metaproteomic Approaches to Understand Host–Microbe Interactions. *Anal. Chem.* 90, 86–109. doi: 10.1021/acs.analchem.7b04340
- Stokell, J. R., Khan, A., and Steck, T. R. (2014). Mechanical Homogenization Increases Bacterial Homogeneity in Sputum. *J. Clin. Microbiol.* 52, 2340–2345. doi: 10.1128/JCM.00487-14
- Tanca, A., Fraumene, C., Manghina, V., Palomba, A., Abbondio, M., Deligios, M., et al. (2017). Diversity and Functions of the Sheep Faecal Microbiota: A Multi-Omic Characterization. *Microb. Biotechnol.* 10, 541–554. doi: 10.1111/1751-7915.12462
- Tanca, A., Palomba, A., Pisanu, S., Addis, M. F., and Uzzau, S. (2015). Enrichment or Depletion? The Impact of Stool Pretreatment on Metaproteomic Characterization of the Human Gut Microbiota. *Proteomics* 15, 3474–3485. doi: 10.1002/pmic.201400573
- Tanca, A., Palomba, A., Pisanu, S., Deligios, M., Fraumene, C., Manghina, V., et al. (2014). A Straightforward and Efficient Analytical Pipeline for Metaproteome Characterization. *Microbiome* 2, 1–16. doi: 10.1186/s40168-014-0049-2
- Thomson, D. D., Berman, J., and Brand, A. C. (2016). High Frame-Rate Resolution of Cell Division During Candida Albicans Filamentation. *Fungal Genet. Biol.* 88, 54–58. doi: 10.1016/j.fgb.2016.02.001
- Tracy, M. C., and Moss, R. B. (2018). The Myriad Challenges of Respiratory Fungal Infection in Cystic Fibrosis. *Pediatr. Pulmonol.* 53, S75–S85. doi: 10.1002/ppul.24126
- Treffon, J., Block, D., Moche, M., Reiß, S., Fuchs, S., Engelmann, S., et al. (2018). Adaptation of Staphylococcus Aureus to Airway Environments in Patients With Cystic Fibrosis by Upregulation of Superoxide Dismutase M and Iron-Scavenging Proteins. *J. Infect. Dis.* 217, 1453–1461. doi: 10.1093/infdis/jiy012
- Voth, D. E., Broderdorf, L. J., and Graham, J. G. (2012). Bacterial Type IV Secretion Systems: Versatile Virulence Machines. *Future Microbiol.* 7, 241–257. doi: 10.2217/fmb.11.150
- Voynow, J., Fischer, B., and Zheng, S. (2008). Proteases and Cystic Fibrosis. *Int. J. Biochem. Cell Biol.* 40, 1238–1245. doi: 10.1016/j.biocel.2008.03.003
- Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences Into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07
- Wei, S., Park, B.-J., Seo, K.-H., and Oh, D.-H. (2016). Highly Efficient and Specific Separation of Staphylococcus Aureus From Lettuce and Milk Using Dynabeads Protein G Conjugates. *Food Sci. Biotechnol.* 25, 1501–1505. doi: 10.1007/s10068-016-0233-1
- Williams, C., Ranjendran, R., and Ramage, G. (2016). Pathogenesis of Fungal Infections in Cystic Fibrosis. *Curr. Fungal Infect. Rep.* 10, 163–169. doi: 10.1007/s12281-016-0268-z
- Worlitzsch, D., Tarran, R., Ulrich, M., Schwab, U., Cekici, A., Meyer, K. C., et al. (2002). Effects of Reduced Mucus Oxygen Concentration in Airway Pseudomonas Infections of Cystic Fibrosis Patients. *J. Clin. Invest.* 109, 317–325. doi: 10.1172/JCI13870
- Wu, X., Siehnel, R. J., Garudathri, J., Staudinger, B. J., Hisert, K. B., Ozer, E. A., et al. (2019). *In Vivo* Proteome of Pseudomonas Aeruginosa in Airways of Cystic Fibrosis Patients. *J. Proteome Res.* 18, 2601–2612. doi: 10.1021/acs.jproteome.9b00122
- Xiong, W., Giannone, R. J., Morowitz, M. J., Banfield, J. F., and Hettich, R. L. (2015). Development of an Enhanced Metaproteomic Approach for Deepening the Microbiome Characterization of the Human Infant Gut. *J. Proteome Res.* 14, 133–141. doi: 10.1021/pr500936p
- Yang, J., Eiserich, J. P., Cross, C. E., Morrissey, B. M., and Hammock, B. D. (2012). Metabolomic Profiling of Regulatory Lipid Mediators in Sputum From Adult Cystic Fibrosis Patients. *Free Radical Biol. Med.* 53, 160–171. doi: 10.1016/j.freeradbiomed.2012.05.001
- Yang, L., Haagensen, J. A. J., Jelsbak, L., Johansen, H. K., Sternberg, C., and Molin, S. (2008). *In Situ* Growth Rates and Biofilm Development of Pseudomonas Aeruginosa Populations in Chronic Lung Infections. *J. Bacteriol.* 190, 2767–2776. doi: 10.1128/JB.01581-07
- Zhu, W., Smith, J. W., and Huang, C.-M. (2010). Mass Spectrometry-Based Label-Free Quantitative Proteomics. *J. BioMed. Biotechnol.* 2010, 840518. doi: 10.1155/2010/840518
- Zühlke, D., Dörries, K., Bernhardt, J., Maaß, S., Muntel, J., Liebscher, V., et al. (2016). Costs of Life - Dynamics of the Protein Inventory of Staphylococcus Aureus During Anaerobiosis. *Sci. Rep.* 6, 1–13. doi: 10.1038/srep28172



Zybailov, B, Mosley, AL, Sardi, ME, Coleman, MK, Florens, L, and Washburn, MP. (2006). Statistical Analysis of Membrane Proteome Expression Changes in *Saccharomyces Cerevisiae*. *J Proteome Res* 5, 2339–2347.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of

the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2021 Graf, Striesow, Pané-Farré, Sura, Wurster, Lalk, Pieper, Becher, Kahl and Riedel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*