Metaproteomics to Decipher CF Host-Microbiota Interactions: Overview, Challenges and Future Perspectives

Pauline Hardouin 1,2, Raphael Chiron 3, Hélène Marchandin 4, Jean Armengaud 2 and Lucia Grenga 2,*  

1 Laboratoire Innovations technologiques pour la Détectio... 30207 Bagnols-sur-Cèze, France; jean.armengaud@cea.fr  
2 Département Médicaments et Technologies pour la Santé (DMTS), Université Paris-Saclay, CEA, INRAE, SPI, 30200 Bagnols-sur-Cèze, France; jean.armengaud@cea.fr  
3 HydroSciences Montpellier, CNRS, IRD, Université de Montpellier, Centre de Ressources et de Compétences de la Mucoviscidose, CHU de Montpellier, 34093 Montpellier, France; r-chiron@chu-montpellier.fr  
4 HydroSciences Montpellier, CNRS, IRD, Université de Montpellier, Service de Microbiologie et Hygiène Hospitalière, CHU Nîmes, 34093 Nîmes, France; helene.marchandin@umontpellier.fr  
* Correspondence: lucia.grenga@cea.fr

Abstract: Cystic fibrosis (CF) is a hereditary disease caused by mutations in the CF transmembrane conductance regulator (CFTR) gene, triggering dysfunction of the anion channel in several organs including the lung and gut. The main cause of morbidity and mortality is chronic infection. The microbiota is now included among the additional factors that could contribute to the exacerbation of patient symptoms, to treatment outcome, and more generally to the phenotypic variability observed in CF patients. In recent years, various omics tools have started to shed new light on microbial communities associated with CF and host–microbiota interactions. In this context, proteomics targets the key effectors of the responses from organisms, and thus their phenotypes. Recent advances are promising in terms of gaining insights into the CF microbiota and its relation with the host. This review provides an overview of the contributions made by proteomics and metaproteomics to our knowledge of the complex host–microbiota partnership in CF. Considering the strengths and weaknesses of proteomics-based approaches in profiling the microbiota in the context of other diseases, we illustrate their potential and discuss possible strategies to overcome their limitations in monitoring both the respiratory and intestinal microbiota in sample from patients with CF.

Keywords: microbiota; cystic fibrosis; proteomics; mass spectrometry

1. Introduction

The genetic mechanism underlying cystic fibrosis (CF) is linked to a combination of genetic mutations in the gene coding for the cystic fibrosis transmembrane conductance regulator (CFTR). Over 2000 known CFTR variants are now associated with a wide range of biological and functional consequences [1] resulting in poor hydration of mucus and impaired function in respiratory, digestive, and reproductive organs. Although recent advances in the molecular understanding of the functional impact of specific CFTR variants have revolutionised drug development, phenotypic differences between CF patients expressing the same genotypic CFTR variant remain to be addressed. The microbiota is now included among the additional factors that could contribute to the clinical presentation, pulmonary exacerbations, and heterogeneity of patient responses to treatments.

Culture-independent profiling of CF airways revealed that (i) the CF respiratory microbiome is a complex community dominated by a more diverse core group of taxa than previously thought [2], and that (ii) stability and resilience are overarching features of the highly variable CF microbiota [3,4]. The CF microbiota includes Achromobacter, Burkholderia, Pseudomonas, Staphylococcus, Stenotrophomonas, and Streptococcus spp. alongside of obligate anaerobes such as Prevotella and Veillonella, and alongside an increased prevalence of non-tuberculous mycobacteria (NTM), viruses, fungi of the Aspergillus, Candida or Malassezia...
genera and archaeal phyla in relatively smaller percentages [5]. Overall, even if every individual CF patient harbour a unique and evolving microbial community, an emerging ecological pattern in CF highlights the relationship between decreasing microbiota diversity and worsening lung function [2,6]. Interactions between members of the microbial community cause CF-related lung disease to progress, dictate the response to therapy, and affect clinical outcomes [7,8]. In concert, environmental factors such as diet and lifestyle as well as host genetics define the chemical and physical landscape inhabited by the microbiome, and shape its composition [9,10]. For a comprehensive overview of the current understanding of the structure, composition, and dynamics of CF microbiota, readers are referred to several recent reviews [7,11–13].

Conditions that could lead to dysbiosis also exist in the gut of CF patients, where reduced bicarbonate secretion from the pancreas, intestines, and biliary tree are linked to the primary CFTR defect. Regular use of antibiotics due to recurrent pulmonary infections, increased amounts of malabsorbed luminal content, and impaired innate immunity may further contribute to the development of microbial dysbiosis [14]. A significant reduction in gut microbial richness and diversity in samples from CF patients was reported in a number of studies [14–18]. Specifically, the gut microbial communities within the CF population are characterised by reduced abundances of specific bacterial taxa including members of the genera \textit{Clostridium}, \textit{Eubacterium}, \textit{Faecalibacterium}, and \textit{Bacteroides}, as well as the order \textit{Lactobacillales}. Emerging evidence associates this dysbiosis within the CF gut with seeding of opportunistic pathogens such as potentially pathogenic \textit{Escherichia coli} and \textit{Eubacterium biforme} [19]. In addition, reduced microbial diversity correlates with a shorter time to CF exacerbation and respiratory colonisation by \textit{Pseudomonas aeruginosa} [20]. Some reports have investigated the links between inflammation and bacterial dysbiosis [15,21], but the mechanisms linking microbiome composition to gastrointestinal manifestations of CF remain largely unexplored, and would merit further research [22,23].

The CF microbiome has mainly been studied using nucleic acid sequencing technologies, particularly based on sequencing of specific marker amplicons (e.g., 16S rRNA bacterial gene). Due to the lack of universal gene markers for fungi, archaea, and viruses, use of these techniques has led the overwhelming majority of investigations of CF microbiota to focus on bacteria. Recently, shotgun metagenomics emerged in this arena and revealed a great number of insights into other components of the microbiota, thus advancing our knowledge of its composition. Nevertheless, to obtain further insights into microbial function and the interactions between the various constituents, complementary approaches should be applied [24,25].

In this mini-review, taking into account the state-of-the-art CF microbiota research and constant advances in proteomics strategies, we explore how and to what extent proteomics-based approaches could contribute to advancing our knowledge of both respiratory and intestinal microbiota in a CF context, and of host–microbiota crosstalk.

2. Can the Current Understanding of the CF Microbiota Shed Light on Genotype–Phenotype Associations?

Microbiomes contribute substantially to the biology of their host and often expand the physiological capacity of the partner organism. In certain cases, these connections have an experimentally identified biochemical basis, whereas in others, these relationships are poorly defined [26,27]. The human microbiome is particularly large and displays significant diversity and complexity, but the lack of appropriate biological models has limited the possibilities for experimentation.

In the context of CF, a link between \textit{CFTR} gene mutations and modifications to the microbiota was first demonstrated by Cox et al. [28]. These authors reported specific \textit{CFTR} mutations, known to influence the airway environment, and found them to be strongly associated with distinct pathogen profiles. These profiles could potentially explain the varying degrees of severity of pulmonary symptoms commonly observed with various \textit{CFTR} genotypes. Similarly, a link was established also between \textit{CFTR} variants and alterations to faecal microbiota [19], opening the way to studies that perceive CF as a systemic disease,
and linking the lung and the gut in a single axis. Despite these pioneering works, both the causes and consequences of dysbiosis for diseases like CF remain difficult to determine. A large number of confounding variables such as antibiotic use, diet, and underlying genetic or environmental factors may significantly contribute to the creation of a distinct niche, and influence the resident microbial community. To this end, the relationships between CFTR dysfunction, microbes, and intestinal health were recently tested in a germ-free CF mouse model [29]. Not only were CFTR gene mutations sufficient to alter the microbiome in the gastrointestinal (GI) tract, but they also modulated the host’s adaptive immune response. Concomitantly, host gene–microbiome interactions were identified in the colon of CF patients [30]. Despite this intensive research focus, the complexity of the issue means many questions remain open. For example, how are CF host genetics, microbiome, and host phenotypes linked? What are the relative contributions of the host genome and microbiome to the CF disease phenotype? The highly integrated host physiologies and their associated microbes have increased the popularity of the concept of the holobiont, the consortium consisting of the host plus its associated microbes evolving to function as an inextricable single unit [31]. In CF, genotype–phenotype associations derived from the analysis of both human genetics and the human microbiome are not yet available, and this type of data is rare even in the context of other diseases [32].

3. Meta-Omics in CF Microbiota Research

The PubMed database was searched for articles published between January 2007 and 2021 using the terms “metagenomics”, “metatranscriptomics”, “metaproteomics”, “metabolomics” and “cystic fibrosis”. The results revealed that in recent years, several omics techniques were employed to answer a wealth of broad questions linked to CF (Table 1). The majority of these studies focused on determining the components of the CF respiratory microbiome. Only recently, investigations have also begun to explore the CF microbiome in other organs, including the GI tract. As gut dysfunction is another prominent feature of CF, and a growing body of evidence supports the existence of a gut–lung axis [33], findings from these studies may have important clinical implications. Strikingly, transcriptomics and metaproteomics studies have rarely been applied in this field, due to the intrinsic difficulty of the functionally oriented approaches.

Table 1. Survey of CF-related “omics” publications *.

| Molecular Level (Methodology) | Biological Questions Addressed                                                                 | CF Associated Records |
|------------------------------|------------------------------------------------------------------------------------------------|-----------------------|
| DNA (metagenomics)           | What is the ecology of the CF lung microbiome and the ecological patterns of CF microbiota?     | 108                   |
| mRNAs (transcriptomics)      | Which genes are expressed? Which components of the microbiota are active?                      | 3                     |
| Proteins (metaproteomics)    | What are the key players in the CF lung? Could such proteins be biomarkers of exacerbation?    | 2 **                  |
| Metabolites (metabolomics)   | How does the gut microbiota affect host metabolism?                                            | 128                   |
| Lipids (lipidomics)          | Is there a lipid signature associated with CF progression?                                      | 22                    |

* PubMed search results. This is not a systematic review. ** [34,35].

3.1. Metagenomics Profiling of CF Microbiota

To date, the majority of culture-independent CF microbiome studies have used 16S ribosomal RNA profiling (metabarcoding targeting prokaryotes), which provides information on the bacteria and some archaea present in a given sample. The speed of analysis and cost effectiveness of this approach mean it is compatible with the analysis of large
numbers of samples. However, it has a limited taxonomic resolution and is subject to biases associated with the DNA extraction and PCR steps required to amplify the marker genes [36].

Untargeted sequencing of DNA extracted from microbial communities, or shotgun metagenomics, enhances the taxonomic and functional resolution of amplicon sequencing and has been shown to surpass the 16S rRNA-based approach when characterising a microbial community [25,37]. Within the CF lung habitat, metagenomics studies reveal a broad repertoire of viruses, moulds, fungi, archaea, and bacteria [38–40]. These data allowed the classification of a larger proportion of microbiome inhabitants at the species level and prominent CF pathogens at the strain level, as well as the detection of multiple subpopulations of specific pathogens [39,41–43] (Figure 1). Recently, Bacci et al. [42] and Dmitrijeva et al. [40] used shotgun metagenomics approaches to track the temporal dynamics of the sputum microbiome. Their results open up the possibility that such methods could be used to help monitor disease progression by providing routine high-resolution characterisation of the CF lung microbiome. Shotgun metagenomic analysis was also used to systematically characterise the CF faecal microbiome [21], and revealed marked CF-associated taxonomic dysbiosis and functional imbalance. Despite these promising results, the method may be significantly affected by biases in the metagenomic output, similar to those observed during characterisation of the CF respiratory tract biome, because of the presence of a high proportion of human DNA [25,37]. To overcome the attendant limitations and improve metagenomics results, a culture-enriched metagenomic sequencing strategy was investigated by Whelan et al. [44]. Culture enrichment allowed the identification of 63% more operational taxonomic units than direct sequencing, proving that culture-enriched molecular profiling is a valuable strategy for the in-depth characterisation of CF microbiota.

Figure 1. Meta-omics profiling of CF microbiota. For each “omics” technology employed to characterise the respiratory or intestinal microbiota associated with CF, the aspects investigated and key contributions to the field are indicated. The colour of the dots next to each finding listed indicates the technique by which it was obtained.
3.2. Metatranscriptomics Profiling of CF Microbiota

Metatranscriptomics reveals the transcripts present within a microbiome at a given point, and thus in time provides a snapshot of its active members. Although the implementation of next-generation sequencing (NGS) technologies to RNA significantly increased the scope for applications of this technique [45], applications to CF microbial communities remain limited to less than a handful of reports (Table 1, Figure 1).

Metatranscriptomics applied to fresh CF sputum revealed similar metabolic profiles between patients highlighting the existence of a shared pool of metabolic genes required for survival in the CF environment [37]. The much more extensive interpatient variation observed in terms of taxonomic composition of the microbiome as compared to functional abundances was corroborated by the comprehensive view of the major physiological processes activated by microbes in the CF lung [46]. Although the technique is biased towards organisms with higher transcriptional rates, the metatranscriptomic profile of the CF microbiome may inform the course of action for treatment and help guide the development of novel therapies through manipulation of the CF environment. This potential for exploitation of metaproteomics data was also illustrated when it was combined with other omics approaches to determine which members of the microbial communities were responsible for a fatal exacerbation [47].

However, several challenges remain to be overcome when attempt to apply metatranscriptomics to host-associated microbes. For instance, messenger RNAs (mRNAs) account for only ~5% of total cellular RNA. While various rRNA depletion methods have been developed to enrich the mRNA content of samples, some are not suitable for use with samples containing large amounts of eukaryotic mRNA, such as the CF sputum [37]. In addition, the lack of standardised methods for RNA preparation and storage, or even data analysis further limits the application of the technique when profiling the CF microbiome.

3.3. Metabolomics Profiling of CF Microbiota

Metabolomics involves the study of metabolites in a biological sample at a given time. Although it remains difficult to determine whether the molecules are of human or microbial origin, a few studies applying metabolomics to CF microbiota have been published (Table 1, Figure 1). Parallel breath metabolomics and shotgun metagenomics revealed that 2,3-butanedione, a fermentation product potentially produced by *Streptococcus* species, the concentration of which varies with antibiotic therapy, is abundant in CF patients and can serve as a cue to neighbouring microbes [48]. These findings, in combination with volatile metabolite profiles associated with common CF pathogens and the host’s immune response to infection, suggest that changes to the volatile metabolome during an exacerbation correlate with changes to the microbiome [49]. Similarly, a study combining metagenomics with two types of untargeted metabolomics analyses applied to sputum samples indicated that the temporary and moderate changes to the lung microbiome observed after initiating the CFTR modulating treatment Ivacaftor/Lumacaftor are mainly characterised by a reduction in the relative abundance of *P. aeruginosa* [50]. A branch of metabolomics dedicated to lipid analysis [51] (lipidomics) studies the lipids present in a sample, arising from host cells and tissues/lipid substrates [52]. Interestingly, when this technique was applied to the CF airway, it raised the question as to the possible involvement of some components of the microbiota in the synthesis and metabolism of some of the lipids identified. This data fuelled what remains a fertile area of investigation [52].

With increasing evidence showing a link between gastrointestinal microbiota and the progression of lung disease in CF [53,54], metabolomics was also employed to investigate the association between an altered functional environment and gut dysbiosis [55,56]. Although recent studies have shown that supplementation with probiotics like *Lactobacillus* could decrease pulmonary exacerbations [15,57], the temporal changes and clinical significance of gut dysbiosis and its effects on the disease process remain poorly understood [54]. In this context, targeted investigation and quantification of metabolites could improve our understanding of how the respective metabolic activity of the lung and gut micro-
Genes 2021, 12, 892

biota affect respiratory health [58]. Furthermore, by establishing distinct microbiome and metabolome signatures, a more personalised approach to patient diagnosis and treatment can be proposed [13,47].

4. (Meta) Proteomics-Based Approaches Applied to CF Microbiota: Where Are We Today?
4.1. Metaproteomics Profiling of CF Microbiota

Proteomics is a multi-faceted omics approach because of the myriad of options it provides to study proteins from various perspectives: abundance, post-translational modifications, quaternary structure, interactions, subcellular localisation, tissue-specific expression, and half-life. Discovery proteomics has extensively contributed to CF research, from the understanding of CFTR biology [59] to the discovery of biomarker for CF diagnosis [60–62]. Proteomics-based CF research has used a wide range of techniques including two-dimensional gel electrophoresis (2-DE), liquid chromatography, mass spectrometry (MS), and antibody/protein microarrays. These techniques have been applied to analyse secretions, cells, and whole tissues from in vitro or in vivo disease models, human subjects, and infecting microorganisms, as well as, to a lesser extent, to profile the CF microbiota. Recent progress in MS has made the identification and quantification of thousands of proteins possible in a single shotgun proteomics study. In this approach, the proteins are proteolyzed with trypsin and the resulting peptides are resolved by reverse phase chromatography before their identification by tandem mass spectrometry (MS/MS) [63,64]. This approach is compatible with differential comparative studies and can be used to identifying the key players in various conditions.

Over the past decade, metaproteomics [64], defined as the large-scale profiling of the whole protein complement produced by a complex microbial ecosystem, has been used to analyse human microbiota from distinct anatomical sites. These type of analyses are gradually gaining momentum as they unravel the functionality of the complex microbial consortium in situ. Metagenomics provides insights into the taxonomic composition, but is restricted to a prediction of the functional potential of the samples analysed. In contrast, metaproteomics can reveal the actual functional profile and provide an in-depth view of the interplay between microorganisms and their host or environment [65,66]. Moreover, having provided information on the role of microbiota in healthy individuals, the field has progressed to explore the functional profiles of disease-related dysbiosis [67]. These data contribute to our understanding of the underlying pathophysiology and pave the way for targeted approaches to improve health.

Notwithstanding its potential, to date, proteomics-based CF microbiota profiling is limited to the work of Debyser et al. [34,35] (Table 1, Figure 1). Applying LC-MS to comparatively analyse stool protein extracts from CF patients and their healthy siblings as controls, these authors extended previous knowledge on gut dysbiosis and inflammation observed in CF patients. Thus, they reported CF to be associated with lower numbers of butyrate-producing bacteria and increased numbers of Enterobacteriaceae, Ruminococcus gnavus, and some clostridia. These species could therefore be viewed as markers of intestinal dysbiosis in patients with CF. In addition to these elements, their metaproteomics approach revealed a list of host proteins associated with inflammation and mucus production that, in combination with bacterial proteins, could serve as LC-MS biomarkers. These biomarkers could be used to monitor gut health in patients with CF, particularly to track responses to probiotic or chemotherapeutic interventions.

Although it has not yet been applied to characterise respiratory microbiota in CF patients, the metaproteomics approach has been used in the context of other respiratory diseases. These studies demonstrated that it was possible to identify microbial peptides, bacterial taxa, and associated biological functions activated by the microbial community in the distal lung [68]. Recently, Pathak and colleagues [69] used metaproteomics to identify changes in pathogen and host gene expression in patients with ventilator-associated pneumonia (VAP). They identified 62 unique pathogen-produced peptides in VAP patients compared to controls and more than 3000 human proteins, many of which are associated
with innate and adaptive immunity. These findings provide mechanistic insights into host-pathogen interactions associated with VAP and could therefore contribute to facilitating diagnosis and treatment.

4.2. Methodological Considerations When Profiling CF Microbiota by Metaproteomics: Data Acquisition

Proteomics data are typically generated by MS following a workflow such as that reviewed by Lin et al. [70]. In metaproteomics, proteins are commonly measured using a bottom-up approach involving extraction, isolation, and digestion of proteins to produce peptides. These peptides are then separated and analysed using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). The resulting MS/MS spectra are typically matched against in silico generated spectra derived from a protein sequence database to identify the peptides detected and infer the identities of the original proteins. The inferred proteins are then used to determine which taxa are active in the community, their functions, and the relative gene expression levels. Each of the steps in this process can potentially influence the outcome of a metaproteomics analysis, as well as providing its own specific benefits and challenges [71]. Overall, two major challenges still limit the routine application of metaproteomics: (i) sample preparation due to high sample complexity and extensive contamination, and (ii) data analysis due to the computational efforts required to treat large datasets, a lack of appropriate annotated protein sequence databases, and difficulties with protein inference leading to ambiguous protein annotation [72].

Among the factors hindering the use of metaproteomics to profile the CF microbiota, particularly the respiratory one, we can include the nature of the CF samples. Compared to the microbiota inhabiting other anatomical sites, the respiratory microbiota is a low biomass community, extensively contaminated with proteins from epithelial cells or cells associated with the acute immune response triggered by diseases such as CF. Because of the high percentage of host proteins, sample pre-treatment steps must be applied to increase the recovery of microbial relative to human biomass. Several methods to enrich and purify microbiota cells, and their potential biases, were recently reviewed [73]. Based on this review, freshly collected samples should be preferred to avoid uncontrolled microbial lysis caused by freezing and thawing of samples. This lysis potentially negatively affects recovery of the proteins of interest, as does the complexity of the very thick and sticky CF mucus [74], and the use of antibiotics [75]. In addition, as suggested by Jagtap et al. [68], coverage of lower-abundance proteins should be improved by depleting known medium- and high-abundance proteins (Figure 2). Similarly, the need for a step to enrich microbial proteins to allow comprehensive proteomic analysis of faecal samples from CF patients was recently addressed through the development and comparison of two shotgun metaproteomic workflows [34]. The results of this comparison revealed the huge influence of the protein extraction method used on the results obtained, underlying the need for a standardised and tailored sample preparation for metaproteomics analysis.

4.3. Methodological Considerations When Profiling CF Microbiota by Metaproteomics: Data Interpretation

The analysis of metaproteomics datasets presents several computational challenges that must be dealt with to ensure reliable peptide and protein identification and support the interpretation of results with meaningful protein annotations. The unique challenges and pitfalls facing researchers at each step of the analysis have been reviewed by several authors. Among others, Starr et al. [73], Heyer et al. [72], and more recently Schiebenhoefer et al. [76] provide comprehensive, detailed, and flexible guides to analysing metaproteomics data. Despite the challenges outlined, the added value of metaproteomics compared to other omics approaches has already been demonstrated numerous times. Thus, in the environmental and medical fields, it has provided unprecedented insights into the functional activity of microbial communities [64]. Recent advances in interpretation of metaproteomics data, such as trough application of the phylopeptidomics concept to allow precise estimation of the biomass contribution of each microorganism identified in a given sam-
ple, add additional value to what can be obtained from each sample [77]. Although complementary to the more commonly used nucleic acid sequencing-based methods, the information provided by metaproteomics inherently differs from the approximations of cell or genome copy counts [66]. Furthermore, culturomics and proteotyping, a proteomics-derived methodology for rapid identification of microbial isolates, could provide a wealth of novel information on the CF microflora [78]. Because of the direct link between phenotypic traits and proteome characteristics, insights into the physiological state and functions of microbial populations could be obtained by proteotyping microorganisms grown under differing culture conditions.

**Figure 2.** Schematic workflow employed in metaproteomics experiments. For each one of the three main steps in the metaproteomic analysis, critical aspects that should be taken into account or carefully evaluated are highlighted.

5. **Perspectives for Applications of Proteomics-Based Approaches to CF Microbiota**

How can proteomics improve our understanding of CF phenotypic variability and more generally of CF disease? Figure 3 shows the set of questions that we believe can now by developing new aspects such as relative biomass estimation and improved functional integration. Like for other microbiota, the contribution of both the airway and the faecal microbiota to CF can be considered to be only partially explored when the taxonomy of their members is considered in isolation. Today, the degree of functional diversity of the microbiota associated with extensive interpatient variability in CF remains underexplored, and inferences between changes to microbial composition and disease phenotypes have been based on associations and correlations, with little attempt to explain the underlying mechanisms [79]. In this scenario, metaproteomics profiling of the CF microbiota (discovery proteomics), in addition to allowing analysis of the composition of the microbiota by quantifying the biomass contributions of individual species, could provide a parallel assessment of the specific functional roles of microbes in a defined microbial community. It thus has the potential to reveal their relevance to the clinical course of CF (Figure 3).
In parallel to characterisation of the microbiota, thanks to the analysis of proteins of all origins, metaproteomics provides insights into the molecular functions attributed to host proteins. These functions can then infer the host–microbiome interactions underlying the disease. Metaproteomics assessment of CF samples also holds great promise for the integration of microbiota information as part of CF infection management. Host proteins could inform on changes to the inflammatory state, the host immune response, and more generally on how therapies affect disease trajectory. Inflammation of the airways is a common biological response to damage or infection and, though generally beneficial, it can significantly influence the composition of the airway microbiome [80]. In CF airways, increased secretion of proinflammatory mediators (e.g., interleukin (IL)-6 and IL-8) by potentially defective airway epithelium and immune cells causes a disproportionate influx of neutrophils. The subsequent release of proteases, such as neutrophil elastase, can damage structural lung proteins, leading to an inevitable decline in pulmonary function. Thus, the development of a range of microbiome-based molecular biomarkers could be useful in monitoring the onset or progression of airway disease, giving clinicians advance warning of changing conditions, and guiding their therapeutic interventions [13] (Figure 3). With therapies that correct the underlying defects in CFTR proteins either available or in the pipeline, there has been great interest in profiling their effect on microbiota. In this context, proteomics approaches alone and in association with metabolomics could provide insights into the correlation between taxonomic and functional dynamics in the microbiota and the efficacy of CFTR modulating therapies. In addition, it could identify potential

---

**Figure 3.** Questions that could be answered using proteomics-based approaches. Sample questions that could currently be addressed by analysing airways and intestinal microbiota from CF patients using proteomics-based approaches. For the discovery proteomics, for each question described (left side), the corresponding output of the metaproteomic analysis is summarised (right side). The main steps in the experimental workflow and the widely used approaches are also highlighted. For the translational proteomics, the main approaches are indicated alongside with their output.
biomarkers (microbes and/or key proteins) to predict the efficacy of a treatment and/or serve as new therapeutic targets.

In addition to expanding our understanding of CF microbial communities, proteomics, particularly its subdiscipline “targeted proteomics”, supports the translation to clinical application of candidate biomarkers identified during the discovery phase. Currently, MS techniques such as SRM/MRM (selected reaction monitoring or multiple reaction monitoring) and SWATH (sequential window acquisition of all theoretical mass spectra) can be used to quantify proteins present in biofluids, tissues, cells, and faeces. This quantification can contribute to biomarker discovery and validation, patient stratification, monitoring of clinical trials, and investigation of proteogenomic correlations (Figure 3). Although they have not yet been applied in the context of CF, the successful use of translational proteomics methods in several proteomics and microbiome studies [70,71,73] underlines their considerable potential.

The recent development of proteomics and metaproteomics technologies should provide additional valuable insights into how microbiota contribute to the progression of CF disease, and its interaction with the host. Moving away from taxonomic inventories towards a better understanding of how the CF microbiota functions could open a new avenue for the identification of novel biomarkers of CF progression and targets for antimicrobial therapy or microbiota-based interventions. However, the true value of proteomics approaches is revealed through integration with other -omics technologies, including genomics, epigenomics, lipidomics, and metabolomics. Ultimately, only by combining the knowledge gained in each of these fields will we be able to truly elucidate the complex interplay between the microbiome and its host in CF.

Author Contributions: P.H. and L.G. prepared the first draft to which R.C., H.M. and J.A. contributed. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: P.H. thanks the “Région Occitanie” for supporting her Ph.D. fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Ahmadi, S.; Bozoky, Z.; Di Paola, M.; Xia, S.; Li, C.; Wong, A.P.; Wellhauser, L.; Molinski, S.V.; Ip, W.; Ouyang, H.; et al. Phenotypic profiling of CFTR modulators in patient-derived respiratory epithelia. *NPJ Genom. Med.* 2017, 2, 12. [CrossRef]
2. Cuthbertson, L.; Walker, A.W.; Oliver, A.E.; Rogers, G.B.; Rivett, D.W.; Hampton, T.H.; Ashare, A.; Elborn, J.S.; De Soyza, A.; Carroll, M.P.; et al. Lung function and microbiota diversity in cystic fibrosis. *Microbiome* 2020, 8, 45. [CrossRef] [PubMed]
3. Cuthbertson, L.; Rogers, G.B.; Walker, A.W.; Oliver, A.; Green, L.E.; Daniels, T.W.; Carroll, M.P.; Parkhill, J.; Bruce, K.D.; van der Gast, C.J. Respiratory microbiota resistance and resilience to pulmonary exacerbation and subsequent antimicrobial intervention. *ISME J.* 2016, 10, 1081–1091. [CrossRef] [PubMed]
4. Allison, S.D.; Martiny, J.B. Colloquium paper: Resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. USA* 2008, 105 (Suppl. 1), 11512–11519. [CrossRef] [PubMed]
5. Koskinen, K.; Pausan, M.R.; Perras, A.K.; Beck, M.; Bang, C.; Mora, M.; Schillhabel, A.; Schmitz, R.; Moissl-Eichinger, C. First Insights into the Diverse Human Archaeome: Specific Detection of Archaea in the Gastrointestinal Tract, Lung, and Nose and on Skin. *mBio* 2017, 8. [CrossRef]
6. van der Gast, C.J.; Walker, A.W.; Stressmann, F.A.; Rogers, G.B.; Scott, P.; Daniels, T.W.; Carroll, M.P.; Parkhill, J.; Bruce, K.D. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. *ISME J.* 2011, 5, 780–791. [CrossRef] [PubMed]
7. Bevivino, A.; Bacci, G.; Drevinek, P.; Nelson, M.T.; Hoffman, L.; Mengoni, A. Deciphering the Ecology of Cystic Fibrosis Bacterial Communities: Towards Systems-Level Integration. *Trends Mol. Med.* 2019, 25, 1110–1122. [CrossRef]
8. Rogers, G.B.; Bruce, K.D.; Hoffman, L.R. How can the cystic fibrosis respiratory microbiome influence our clinical decision-making? *Curr. Opin. Pulm. Med.* 2017, 23, 536–543. [CrossRef]
Genetic Determinants of the Gut Microbiome in UK Twins. Cell Host Microbe 2016, 19, 731–743. [CrossRef] [PubMed]

Héry-Arnaud, G.; Boutin, S.; Cuthbertson, L.; Elborn, S.J.; Tunney, M.M. The lung and gut microbiome: What has to be taken into consideration for cystic fibrosis? J. Cyst. Fibros. 2019, 18, 13–21. [CrossRef]

Francoise, A.; Héry-Arnaud, G. The Microbiome in Cystic Fibrosis Pulmonary Disease. Genes 2020, 11, 536. [CrossRef]

Lee, A.J.; Einarrsson, G.G.; Gilpin, D.F.; Tunney, M.M. Multi-Omics Approaches: The Key to Improving Respiratory Health in People with Cystic Fibrosis? Front. Pharmacol. 2020, 11, 569821. [CrossRef] [PubMed]

Nielsen, S.; Needham, B.; Leach, S.T.; Day, A.S.; Jaffe, A.; Thomas, T.; Ooi, C.Y. Disrupted progression of the intestinal microbiota with age in children with cystic fibrosis. Sci. Rep. 2016, 6, 24857. [CrossRef]

Bruzzone, E.; Callegari, M.L.; Raia, V.; Viscoivo, S.; Scotto, R.; Ferrari, S.; Morelli, L.; Buccigrossi, V.; Lo Vecchio, A.; Ruberto, E.; et al. Disrupted intestinal microbiota and intestinal inflammation in children with cystic fibrosis and its restoration with Lactobacillus GG: A randomised clinical trial. PLoS ONE 2014, 9, e87796. [CrossRef]

Duytschaever, G.; Huys, G.; Bekaert, M.; Boulanger, L.; De Boeck, K.; Vandamme, P. Dysbiosis of bifidobacteria and Clostridium cluster XIVa in the cystic fibrosis fibrobiota. J. Cyst. Fibros. 2013, 12, 206–215. [CrossRef]

Duytschaever, G.; Huys, G.; Bekaert, M.; Lourleur, S.; De Boeck, K.; Vandamme, P. Cross-sectional and longitudinal comparisons of the predominant fecal microbiota compositions of a group of pediatric patients with cystic fibrosis and their healthy siblings. Appl. Environ. Microbiol. 2011, 77, 8013–8024. [CrossRef] [PubMed]

Lynch, S.V.; Goldfarb, K.C.; Wild, Y.K.; Kong, W.; De Lisle, R.C.; Brodie, E.L. Cystic fibrosis transmembrane conductance regulator knockout mouse exhibit aberrant gastrointestinal microbiota. Gut Microbes 2013, 4, 41–47. [CrossRef] [PubMed]

Schippa, S.; lebbas, V.; Santangelo, F.; Gagliardi, A.; De Biase, R.V.; Bertasi, S.; Lucarelli, M.; Conte, M.P.; Quattrucci, S. Cystic fibrosis transmembrane conductance regulator (CFTR) allelic variants relate to shifts in faecal microbiota of cystic fibrosis patients. PLoS ONE 2013, 8, e61176. [CrossRef]

Hoen, A.G.; Li, J.; Moulton, L.A.; O’Toole, G.A.; Housman, M.L.; Koestler, D.C.; Guill, M.F.; Moore, J.H.; Hibberd, P.L.; Morrison, H.G.; et al. Associations between Gut Microbial Colonization in Early Life and Respiratory Outcomes in Cystic Fibrosis. J. Pediatrics 2015, 167, 138–147.e133. [CrossRef]

Manor, O.; Levy, R.; Pope, C.E.; Hayden, H.S.; Brittnacher, M.J.; Carr, R.; Radey, M.C.; Hager, K.R.; Heitsel, S.L.; Ramsey, B.W.; et al. Metagenomic evidence for taxonomic dysbiosis and functional imbalance in the gastrointestinal tracts of children with cystic fibrosis. Sci. Rep. 2016, 6, 22493. [CrossRef]

Ooi, C.Y.; Durie, P.R. Cystic fibrosis from the gastroenterologist’s perspective. Nat. Rev. Gastroenterol. Hepatol. 2016, 13, 175–185. [CrossRef]

Garg, M.; Ooi, C.Y. The Enigmatic Gut in Cystic Fibrosis: Linking Inflammation, Dysbiosis, and the Increased Risk of Malignancy. Curr. Gastroenterol. Rep. 2017, 19, 6. [CrossRef]

Hahn, A.; Sanyal, A.; Perez, G.F.; Colberg-Poley, A.M.; Campos, J.; Rose, M.C.; Perez-Losada, M. Different next generation sequencing platforms produce different microbial profiles and diversity in cystic fibrosis sputum. J. Microbiol. Methods 2016, 130, 95–99. [CrossRef]

Budden, K.F.; Shukla, S.D.; Rehman, S.F.; Bowerman, K.L.; Keely, S.; Hugenholtz, P.; Armstrong-James, D.P.H.; Adcock, I.M.; Chotirmall, S.H.; Chung, K.F.; et al. Functional effects of the microbiota in chronic respiratory disease. Lancet Respir. Med. 2019, 7, 907–920. [CrossRef]

Lynch, J.B.; Hsiao, E.Y. Microbiomes as sources of emergent host phenotypes. Science 2019, 365, 1405–1409. [CrossRef] [PubMed]

Brussow, H. Parkinson disease, levodopa and the gut microbiota—when microbiology meets pharmacology. Environ. Microbiol. 2020, 22, 808–812. [CrossRef]

Cox, M.J.; Allgaier, M.; Taylor, B.; Baek, M.S.; Huang, Y.J.; Daly, R.A.; Karaoz, U.; Andersen, G.L.; Brown, R.; Fujimura, K.E.; et al. Airway microbiota and pathogen abundance in age-stratified cystic fibrosis patients. PLoS ONE 2010, 5, e11044. [CrossRef] [PubMed]

Meecker, S.M.; Mears, K.S.; Sangwan, N.; Britttnacher, M.J.; Weiss, E.J.; Treuting, P.M.; Tolley, N.; Pope, C.E.; Hager, K.R.; Vo, A.T.; et al. CFTR dysregulation drives active selection of the gut microbiome. PLoS Pathog. 2020, 16, e1008251. [CrossRef] [PubMed]

Dayama, G.; Priya, S.; Niccum, D.E.; Khorsut, A.; Spielman, P. Interactions between the gut microbiome and host gene regulation in cystic fibrosis. Genome Med. 2020, 12, 12. [CrossRef] [PubMed]

Bordenstein, S.R.; Theis, K.R. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. PLoS Biol. 2015, 13, e1002226. [CrossRef] [PubMed]

Brussow, H. The relationship between the host genome, microbiome, and host phenotype. Environ. Microbiol. 2020, 22, 1170–1173. [CrossRef] [PubMed]

Enaud, R.; Prevot, R.; Ciarlo, E.; Beaufils, F.; Wieser, G.; Guery, B.; Delhaes, L. The Gut-Lung Axis in Health and Respiratory Diseases: A Place for Inter-Organ and Inter-Kingdom Crosstalks. Front. Cell. Infect. Microbiol. 2020, 10, 9. [CrossRef] [PubMed]
58. Lee-Sarwar, K.A.; Lasky-Su, J.; Kelly, R.S.; Litonjua, A.A.; Weiss, S.T. Gut Microbial-Derived Metabolomics of Asthma. *Metabolites* 2020, 10, 97. [CrossRef]

59. Liessi, N.; Fedemonte, N.; Armiriotti, A.; Braccia, C. Proteomics and Metabolomics for Cystic Fibrosis Research. *Int. J. Mol. Sci.* 2020, 21, 5439. [CrossRef] [PubMed]

60. Pattison, S.H.; Gibson, D.S.; Johnston, E.; Peacock, S.; Rivera, K.; Tunney, M.M.; Pappin, D.J.; Elborn, J.S. Proteomic profile of cystic fibrosis sputum cells in adults chronically infected with *Pseudomonas aeruginosa*. *Eur. Respir. J.* 2017, 50. [CrossRef]

61. Dong, K.; Moon, K.M.; Chen, V.; Ng, R.; Foster, L.J.; Tebbutt, S.J.; Quon, B.S. Identification of novel blood biomarkers of treatment response in cystic fibrosis pulmonary exacerbations by label-free quantitative proteomics. *Sci. Rep.* 2019, 9, 17126. [CrossRef]

62. Benabdellah, H.; Alamri, H.; Okla, M.; Masood, A.; Abdel Jabar, M.; Alanazi, I.O.; Alfadda, A.A.; Dasouki, M.; Abdel Rahman, A.M. Serum-Based Proteomics Profiling in Adult Patients with Cystic Fibrosis. *Int. J. Mol. Sci.* 2020, 21, 7415. [CrossRef]

63. Gouveia, D.; Grenga, L.; Pible, O.; Armengaud, J. Quick microbial molecular phenotyping by differential shotgun proteomics. *Environ. Microbiol.* 2020, 22, 2996–3004. [CrossRef]

64. Wilmes, P.; Bond, P.L. The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ. Microbiol.* 2004, 6, 911–920. [CrossRef] [PubMed]

65. Kleiner, M. Metaproteomics: Much More than Measuring Gene Expression in Microbial Communities. *mSystems* 2019, 4. [CrossRef]

66. Salvato, F.; Hettich, R.L.; Kleiner, M. Five key aspects of metaproteomics as a tool to understand functional interactions in host-associated microorganisms. *PloS Pathog.* 2021, 17, e1009245. [CrossRef]

67. Lee, P.Y.; Chin, S.F.; Neoh, H.M.; Jamal, R. Metaproteomic analysis of human gut microbiota: Where are we heading? *J. Biomed. Sci.* 2017, 24, 36. [CrossRef] [PubMed]

68. Jagtap, P.D.; Viken, K.J.; Johnson, J.; McGowan, T.; Pendleton, K.M.; Griffin, T.J.; Hunter, R.C.; Rudney, J.D.; Bhargava, M. BAL Fluid Metaproteome in Acute Respiratory Failure. *Am. J. Respir. Cell Mol. Biol.* 2018, 59, 648–652. [CrossRef]

69. Pathak, K.V.; McGilvrey, M.I.; Hu, C.K.; Garcia-Mansfield, K.; Lewandoski, K.; Eftekhari, Z.; Yuan, Y.C.; Zenhausern, F.; Menashi, E.; Pirrotte, P. Molecular Profiling of Innate Immune Response Mechanisms in Ventilator-associated Pneumonia. *Mol. Cell. Proteom.* 2020, 19, 1688–1705. [CrossRef]

70. Lin, H.; He, Q.Y.; Shi, L.; Sleeman, M.; Baker, M.S.; Nice, E.C. Proteomics and the microbiome: Pitfalls and potential. *Expert Rev. Proteom.* 2019, 16, 501–511. [CrossRef] [PubMed]

71. Van Den Bossche, T.; Kunath, B.; Schallert, K.; Schäpe, S.; Abraham, P.; Armengaud, J.; Arntzen, M.; Bassignani, A.; Benndorf, D.; Fuchs, S.; et al. Critical Assessment of Metaproteome Investigation (CAMPI): A Multi-Lab Comparison of Established Workflows. Front. Microbiol. 2021. [CrossRef]

72. Heyer, R.; Schallert, K.; Budel, A.; Zoun, R.; Dorl, S.; Behne, A.; Kohrs, F.; Puttker, S.; Siewert, C.; et al. A Robust and Critical Assessment of Metaproteome Investigation (CAMPI): A Multi-Lab Comparison of Established Workflows. *Clin. Mass Spectrom.* 2019, 10, 9–17. [CrossRef]

73. Schiebenhöfer, H.; Schallert, K.; Renard, B.Y.; Trappe, K.; Schmid, E.; Benndorf, D.; Riedel, K.; Muth, T.; Fuchs, S. A complete and flexible workflow for metaproteomics data analysis based on MetaProteomeAnalyzer and Prophane. *Anal. Chem.* 2021, 93, 907–916. [CrossRef]

74. Gouveia, D.; Grenga, L.; Pible, O.; Armengaud, J. Quick microbial molecular phenotyping by differential shotgun proteomics. *Front. Microbiol.* 2019, 10, 5439. [CrossRef] [PubMed]

75. Pible, O.; Allain, F.; Jouffret, V.; Culotta, K.; Miotello, G.; Armengaud, J. Estimating relative biomasses of organisms in microbiota using “phylopeptidomics”. *Microbiome* 2020, 8, 30. [CrossRef]

76. Schiebenhoefer, H.; Schallert, K.; Renard, B.Y.; Trappe, K.; Schmid, E.; Benndorf, D.; Riedel, K.; Muth, T.; Fuchs, S. A complete and flexible workflow for metaproteomics data analysis based on MetaProteomeAnalyzer and Prophane. *Nat. Protoc.* 2020, 15, 3212–3239. [CrossRef] [PubMed]

77. Pible, O.; Allain, F.; Jouffret, V.; Culotta, K.; Miotello, G.; Armengaud, J. Estimating relative biomasses of organisms in microbiota using “phylopeptidomics”. *Microbiome* 2020, 8, 30. [CrossRef]

78. Grenga, L.; Pible, O.; Armengaud, J. Pathogen protoype: A rapidly developing application of mass spectrometry to address clinical concerns. *Clin. Mass Spectrom.* 2019, 1–17. [CrossRef]

79. Bacci, G.; Mengoni, A.; Fiscarelli, E.; Segata, N.; Taccetti, G.; Dolce, D.; Paganin, P.; Morelli, P.; Tuccio, V.; De Alessandri, A.; et al. A Different Microbiome Gene Repertoire in the Airways of Cystic Fibrosis Patients with Severe Lung Disease. *Int. J. Mol. Sci.* 2017, 18, 1654. [CrossRef]

80. Huffnagle, G.B.; Dickson, R.P.; Lukacs, N.W. The respiratory tract microbiome and lung inflammation: A two-way street. *Mucosal Immunol.* 2017, 10, 299–306. [CrossRef] [PubMed]