



# Metaproteomics—An Advantageous Option in Studies of Host–Microbiota Interaction

Oleg Karaduta <sup>1,\*</sup>, Zeljko Dvanajscak <sup>2</sup> and Boris Zybailov <sup>1</sup>

<sup>1</sup> Department of Biochemistry and Molecular Biology, UAMS, Little Rock, AR 72205, USA; BLZybailov@uams.edu

<sup>2</sup> Arkana Laboratories, Little Rock, AR 72211, USA; zeljko.dvanajscak@arkanalabs.com

\* Correspondence: okkaraduta@uams.edu; Tel.: +1-501-251-5381

**Abstract:** Gut microbiome contributes to host health by maintaining homeostasis, increasing digestive efficiency, and facilitating the development of the immune system. Manipulating gut microbiota is being recognized as a therapeutic target to manage various chronic diseases. The therapeutic manipulation of the intestinal microbiome is achieved through diet modification, the administration of prebiotics, probiotics, or antibiotics, and more recently, fecal microbiome transplantation (FMT). In this opinion paper, we give a perspective on the current status of application of multi-omics technologies in the analysis of host–microbiota interactions. The aim of this paper was to highlight the strengths of metaproteomics, which integrates with and often relies on other approaches.

**Keywords:** metaproteomics; metagenomics; microbiome; microbiota; dietary change

**Citation:** Karaduta, O.; Dvanajscak, Z.; Zybailov, B. Metaproteomics—An Advantageous Option in Studies of Host–Microbiota Interaction. *Microorganisms* **2021**, *9*, 980. <https://doi.org/10.3390/microorganisms9050980>

Academic Editor:  
Sharon M. Donovan

Received: 31 March 2021

Accepted: 30 April 2021

Published: 30 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The role of the gut microbiome in human health and disease is becoming clearer thanks to high throughput sequencing technologies (HTS). The development and application of fast and low-cost DNA sequencing methods has been revolutionary. HTS has been widely used to examine the complexity of the gut microbiome due to the speed, scale and precision of information yielded. For compositional analysis, the 16S rRNA gene has been most frequently targeted due to its presence in all prokaryotes and the existence of variable domains that allow different taxa to be distinguished. While 16S rRNA studies provide data in relation to the microbial composition of an ecosystem, these do not provide direct information regarding the microbial viability or the functional potential of the populations present.

Metagenomic (or shotgun sequencing) studies go beyond the 16S rRNA gene to characterize the full genetic content of a community, thereby providing an insight into the potential functional capacity of the microbes present [1–3]. Regardless of the approach taken, it is important to note that these sequencing technologies require detailed bioinformatic analyses to deal with the large volumes of data generated [4,5].

Although these gene-centric approaches have provided much information regarding the content of the gut, we also need to understand the activity of these genes and their impact on the metabolic networks within the gut. To further determine specific microbial activity, it is necessary to analyze gene expression (metatranscriptomics), protein products (metaproteomics) and metabolic profiles (metabolomics). These techniques are complex, and to different extents, are still somewhat in their infancy. To date, metatranscriptomics, based on the large-scale sequencing of 16S rRNA transcripts, has been used to look at the composition of the active microbiota in healthy individuals and has revealed that the transcriptional profile across individuals is more similar than indicated by the associated taxonomic diversity [6]. Functionality between the individuals is more similar than the taxonomy, also when analyzing metaproteomics [7].

The reason that metaproteomics has become an important tool can be partially explained by the “quantum leaps” in enabling technologies that have made it more feasible and affordable. Technological advances have mostly occurred in the realm of liquid chromatography (LC) enabling the separation of highly complex peptide mixtures, and high-resolution mass spectrometry (MS) instrumentation enabling the acquisition of large numbers of accurate mass spectra, and computational tools for data processing and analyses.

## 2. Tasks That Can Be Solved Through Metaproteomics

In its basic form, metaproteomics allows us to study the presence and abundances of proteins in any microbial community. With a well-curated protein sequence database in hand, we can assign these proteins to individual species or higher taxa and understand the functional roles and interactions of individual members in the community. Since proteins convey structure and activities to cells, knowing their abundances provides a picture of the cellular phenotypes at the molecular level. Pure culture studies from a diversity of organisms have shown that the abundance of a protein indicates its relevance and activity under a given condition. Applying this at the community level can lead to critical insights and significant progress. For example, in a marine worm that relies on its five bacterial symbionts for all nutrition, we did not know which environmental energy sources the symbionts were using for carbon fixation. A metaproteomic study revealed that some of the symbionts abundantly expressed carbon monoxide dehydrogenases, which suggested they were using carbon monoxide as an energy source [8].

Differential metaproteomics can be used to identify changes in the expression of individual genes, for example, after actively manipulating a microbial community by changing substrates or if environmental conditions naturally fluctuate. In a recent publication, studies have used this approach to show that hydrogen transfer is the basis of the symbiosis between a *Breviatea* protist and its *Arcobacter* symbiont by comparing the metaproteomes from both partners to those of only one partner [9,10].

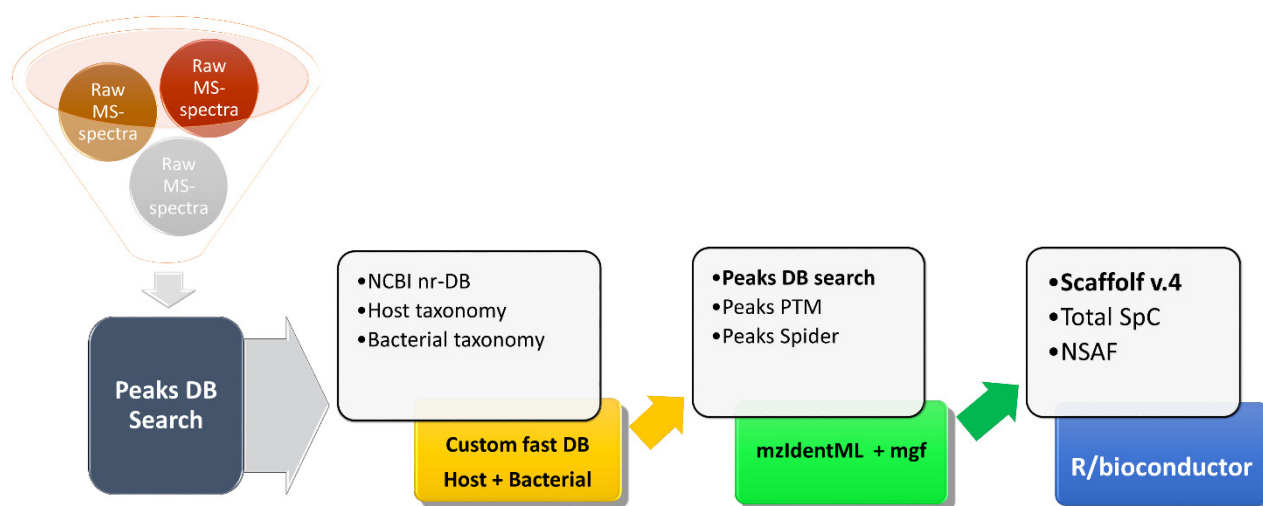
## 3. Novel Software Tools in Metaproteomics Analysis

Most of the reported metaproteomics studies use shotgun proteomics for the identification of bacterial proteins, which may be influenced by complexity and sensitivity issues, such that the detection of low abundant proteins may be challenging. However, the complexity and dynamic range issues may be addressed using the emerging spectral library-based methods such as data-independent acquisition (DIA), which provides a library-based, broader coverage for peptide/protein detection [11–15]. While such an approach has been increasingly applied in the quantitative analysis of the human proteome, its applications in metaproteomics has lagged behind, in part due to the complexity involved in bioinformatics. Reference databases are constantly integrating mass spec analysis of cultured bacterial and pathogen species to generate reference databases for improved species identification [16,17]. Recently developed software tools specifically designed for metaproteomic data analysis have become available as well. MetaLab uses spectral clustering to improve peptide identification speeds [18]. Others have shown improved protein identification in metaproteomics by employing de Bruijn graph assembly to predict protein sequences from metagenomics sequence data and generate a reference database [19,20]. The taxon-specific classification of peptide sequences can be performed using UniPept (<http://unipept.ugent.be>) which uses shotgun proteomics data from UniProt KB with identification noise filtering to provide enhanced biodiversity analysis [21,22]. Another option for the analysis of metaproteomics data is MetaPro-IQ, which is ideal for fecal samples since the gut microbiome gene catalog was curated from fecal studies which negates the need for matched metagenomics data but makes it less applicable for other types of microbiome samples [21].

#### 4. Metaproteomics Pipeline

In recent years, progress in developing a novel technique enabling the species-level resolution of gut microbiota has been shown. Corresponding data analyses usually consist of numerous manual steps that must be closely synchronized. The general approach usually consists of getting raw mass spectrometry data, processing it with sequential steps of database construction, peptide identification and quantification, taxonomy and function profile construction. A variety of methods and platforms are utilized to ease this process, such as iMetaLab [23], MetaProteomeAnalyzer and Prophan [24].

In our laboratory, we developed metaproteomic methods to study animal- and human-associated microbiota with a focus on understanding the critical interactions between all partners, including the host (Figure 1). Such studies are challenging for multiple reasons, including the presence of fecal or soil-derived substances that interfere with sample preparation and the high abundances of host-associated proteins that potentially swamp out the microbial signal. This particular multi-step database search strategy was used in PEAKS Studio (Bioinformatics Solutions) to arrive at the final list of identified proteins [25].



**Figure 1.** Experimental and analytical workflow. The metaproteomics data-analysis shows the software tools used to arrive at the identified proteins list.

At the first step, PEAKS generates de novo peptide sequence tags from the raw mass spectra without the use of any sequence database. For the initial quality assessment of metaproteomics samples, the list of de novo peptide sequences was supplied to UniPept, an online metaproteomics tool [22,26].

At this initial step, one can assess the quality of sample preparation by comparing ratios between host and bacterial proteins, with more bacterial proteins indicating higher quality. Outlier samples, where the bacterial proteins are under-represented or missing, can be removed at this step, to save time in the downstream analysis. At the second step, the PEAKS-generated de novo tags are searched against large sequence databases, which include both host and bacterial proteins. These databases have millions of entries, and after the list of identified proteins is obtained, their sequences are further used to create a custom, smaller database, against which the lists of de novo peptide tags are searched again, to arrive at a final protein list.

For protein quantification, Scaffold v.4 with the quantitation module (Proteome Software) was used. Data and the custom FASTA database were exported from PEAKS into Scaffold as mzIdentML and mascot generic format files. Scaffold-derived normalized spectrum abundance factors (NSAF) values and Total Unique spectral counts were used for the statistical analysis in R/bioconductor [27]. Taxonomic information from the NCBI-

nr and Uniprot database was used to derive taxonomic units. The sum of spectral counts matching a given taxonomic unit, weighted by a total spectral count per a given replicate, was used as a measure of abundance for that taxonomic unit. Moderated t-test (limma package, Bioconductor) was used to establish the differential abundance of taxonomic units. Bonferroni–Hochberg correction for multiple testing was enforced for both protein and taxonomic unit quantification [28].

This method might improve our understanding of the host–bacteria interaction with further progress not only in disease management; it might improve and deepen our understanding of how neonatal diet alters fecal microbiota profiles. Even though a massive study to investigate fecal microbiota and metabolites at different ages in infants who were breastfed, received dairy-based milk formula, or received soy-based formula was performed, authors were not able to capture species-level differences due to the limitations of 16S rRNA sequencing [29].

This method could be also be a part of a multi-omics approach in studies aiming to elucidate longitudinal dynamics in the microbiome structure [30] or in studies of modulating gut microbiota composition in a preclinical neonatal model [31].

## 5. Technical Challenges and Future Perspectives

Despite metaproteomics having been around for fifteen years, it is still in its infancy and there is great potential for its continued development to address otherwise intractable questions. The appearance of new technologies, training additional personnel on metaproteomics approaches, validating and standardizing approaches, developing custom computational tools, and establishing dedicated metaproteomics meetings are the keys that will drive this method forwards.

Predictions about the future of metaproteomics need to anticipate its future applications. Foreseeable trends are an increase in MS resolution and therefore more data will be acquired. Although metaproteomics is a very powerful method, issues within the bioinformatics evaluation impede its success. There will be more studies of gut microbial protein expression in response to dietary patterns that would lead to new heights in understanding microbiome behavior [32].

Protein identification is difficult if the taxonomic composition is unknown or protein entries are missing from protein databases. For example, the UniProt/TrEMBL database only contains proteins from 1,240,268 species (<http://www.ebi.ac.uk/uniprot/TrEMBLstats>, status, accessed on 29 March 2021), but the number of microbial species on Earth is estimated to be up to one trillion [33]. Thereby, already small changes in the protein sequence between related microorganisms have a big impact on protein identification. One mutation in every tenth amino acid leads to completely different tryptic peptides which hinder the identification of any peptide for the investigated protein.

Another often neglected aspect is the reproducibility of results using different metaproteomics software tools. Thus far, only Tanca et al. [34] tested their complete metaproteomics workflow for a defined mixed culture of nine different microorganisms. A comparison where multiple research groups evaluate an identical sample would also be desirable. Additionally, it is evident that sample complexity and inter-individual variation in gut microbiota are extensive [35]. It is also important to note that host and microbiota interactions involve delicate interplay between factors such as age, genetics, immunity and dietary habits, which could act as potential confounding variables. Establishing an integrated multi-omics protocol will also be beneficial for the development of metaproteomics as either a stand-alone approach or as part of a complex strategy [30].

## 6. Conclusions

With the growing interest in understanding the link between gut microbiota in health and diseases, metaproteomics analysis is instrumental to characterize the activity and functional pathways of the microbial community. It represents a powerful tool for the taxonomic and functional characterization of complex microbial communities from all the

variety of samples. In the future, it has the potential to become a valuable tool for routine diagnostics, e.g., the analysis of human feces. Although challenging, further advances in sample preparation methods, the development of more sophisticated analytical tools in addition to the availability of relevant software and databases are expected to facilitate the progress of metaproteomics in the coming years. Metaproteomics studies will benefit from software supporting the taxonomic and functional interpretation of results. Even if it is obvious, the close cooperation of bioinformaticians and biologists should also be considered during software development.

By harnessing the power of the emerging technologies, it is anticipated that more details on the microbial functionality and their connection with a human host will be uncovered in the future.

**Author Contributions:** Conceptualization, original draft preparation: O.K.; proofreading, correction of manuscript: Z.D.; final revision of the manuscript: B.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This publication was supported by the Center for Translational Pediatric Research (CPTR) NIH Center of Biomedical Research Excellence Award P20GM121293 (awarded to Alan Tackett).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank the Department of Biochemistry and Molecular Biology, UAMS; Arkana Laboratories, Little Rock, AR, USA.

**Conflicts of Interest:** The authors declare that there are no conflicts of interest regarding the publication of this article.

## References

1. Kurokawa, K.; Itoh, T.; Kuwahara, T.; Oshima, K.; Toh, H.; Toyoda, A.; Takami, H.; Morita, H.; Sharma, V.K.; Srivastava, T.P.; et al. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res* **2007**, *14*, 169–181, doi:10.1093/dnares/dsm018.
2. Qin, J.; Li, R.; Raes, J.; Arumugam, M.; Burgdorf, K.S.; Manichanh, C.; Nielsen, T.; Pons, N.; Levenez, F.; Yamada, T.; et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **2010**, *464*, 59–65, doi:10.1038/nature08821.
3. Turnbaugh, P.J.; Ridaura, V.K.; Faith, J.J.; Rey, F.E.; Knight, R.; Gordon, J.I. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* **2009**, *1*, 6ra14, doi:10.1126/scitranslmed.3000322.
4. Kuczynski, J.; Lauber, C.L.; Walters, W.A.; Parfrey, L.W.; Clemente, J.C.; Gevers, D.; Knight, R. Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* **2011**, *13*, 47–58, doi:10.1038/nrg3129.
5. Hickl, O.; Heintz-Buschart, A.; Trautwein-Schult, A.; Hercog, R.; Bork, P.; Wilmes, P.; Becher, D. Sample Preservation and Storage Significantly Impact Taxonomic and Functional Profiles in Metaproteomics Studies of the Human Gut Microbiome. *Microorganisms* **2019**, *7*, doi:10.3390/microorganisms7090367.
6. Pérez-Cobas, A.E.; Gosalbes, M.J.; Friedrichs, A.; Knecht, H.; Artacho, A.; Eismann, K.; Otto, W.; Rojo, D.; Bargiela, R.; von Bergen, M.; et al. Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut* **2013**, *62*, 1591–1601, doi:10.1136/gutjnl-2012-303184.
7. Visconti, A.; Le Roy, C.I.; Rosa, F.; Rossi, N.; Martin, T.C.; Mohn, R.P.; Li, W.; de Rinaldis, E.; Bell, J.T.; Venter, J.C.; et al. Interplay between the human gut microbiome and host metabolism. *Nature Communications* **2019**, *10*, 4505, doi:10.1038/s41467-019-12476-z.
8. Kleiner, M. Metaproteomics: Much More than Measuring Gene Expression in Microbial Communities. *mSystems* **2019**, *4*, doi:10.1128/mSystems.00115-19.
9. Hamann, E.; Gruber-Vodicka, H.; Kleiner, M.; Tegetmeyer, H.; Riedel, D.; Littmann, S.; Chen, J.; Milucka, J.; Viehweger, B.; Becker, K.; et al. Environmental Breviatea harbour mutualistic *Arcobacter* epibionts. *Nature* **2016**, *534*, 254–+, doi:10.1038/nature18297.
10. Lohmann, P.; Benk, S.; Gleixner, G.; Potthast, K.; Michalzik, B.; Jehmlich, N.; Bergen, M.v. Seasonal Patterns of Dominant Microbes Involved in Central Nutrient Cycles in the Subsurface. *Microorganisms* **2020**, *8*, 1694.
11. Chapman, J.D.; Goodlett, D.R.; Masselon, C.D. Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrom Rev* **2014**, *33*, 452–470, doi:10.1002/mas.21400.

12. Nigjeh, E.N.; Chen, R.; Allen-Tamura, Y.; Brand, R.E.; Brentnall, T.A.; Pan, S. Spectral library-based glycopeptide analysis-detection of circulating galectin-3 binding protein in pancreatic cancer. *Proteomics Clin Appl* **2017**, *11*, doi:10.1002/prca.201700064.
13. Rosenberger, G.; Liu, Y.; Röst, H.L.; Ludwig, C.; Buil, A.; Bensimon, A.; Soste, M.; Spector, T.D.; Dermitzakis, E.T.; Collins, B.C.; et al. Inference and quantification of peptidoforms in large sample cohorts by SWATH-MS. *Nat Biotechnol* **2017**, *35*, 781–788, doi:10.1038/nbt.3908.
14. Oleg, K.; Loutfouz, Z. Shk-9: A new tool in approach of glycoprotein annotation. **2018**, *7*, 302–303, doi:doi.org/10.1016/j.softx.2018.08.004.
15. Aakko, J.; Pietilä, S.; Suomi, T.; Mahmoudian, M.; Toivonen, R.; Kouvonen, P.; Rokka, A.; Hänninen, A.; Elo, L.L. Data-Independent Acquisition Mass Spectrometry in Metaproteomics of Gut Microbiota—Implementation and Computational Analysis. *Journal of Proteome Research* **2020**, *19*, 432–436, doi:10.1021/acs.jproteome.9b00606.
16. Alispahic, M.; Hummel, K.; Jandreski-Cvetkovic, D.; Nöbauer, K.; Razzazi-Fazeli, E.; Hess, M.; Hess, C. Species-specific identification and differentiation of *Arcobacter*, *Helicobacter* and *Campylobacter* by full-spectral matrix-associated laser desorption/ionization time of flight mass spectrometry analysis. *J Med Microbiol* **2010**, *59*, 295–301, doi:10.1099/jmm.0.016576-0.
17. Schape, S.; Krause, J.; Engelmann, B.; Fritz-Wallace, K.; Schattenberg, F.; Liu, Z.; Muller, S.; Jehmlich, N.; Rolle-Kampczyk, U.; Herberth, G.; et al. The Simplified Human Intestinal Microbiota (SIHUMIx) Shows High Structural and Functional Resistance against Changing Transit Times in In Vitro Bioreactors. *Microorganisms* **2019**, *7*, doi:10.3390/microorganisms7120641.
18. Cheng, K.; Ning, Z.; Zhang, X.; Li, L.; Liao, B.; Mayne, J.; Stintzi, A.; Figeys, D. MetaLab: an automated pipeline for metaproteomic data analysis. *Microbiome* **2017**, *5*, 157, doi:10.1186/s40168-017-0375-2.
19. Tang, H.; Li, S.; Ye, Y. A Graph-Centric Approach for Metagenome-Guided Peptide and Protein Identification in Metaproteomics. *PLoS Comput Biol* **2016**, *12*, e1005224, doi:10.1371/journal.pcbi.1005224.
20. Karaduta, O.K.; Deon, A.F.; Menyaev, Y.A. Designing the Uniform Stochastic Photomatrix Therapeutic Systems. *Algorithms* **2020**, *13*, 41.
21. Mesuere, B.; Devreese, B.; Debyser, G.; Aerts, M.; Vandamme, P.; Dawyndt, P. Unipept: tryptic peptide-based biodiversity analysis of metaproteome samples. *J Proteome Res* **2012**, *11*, 5773–5780, doi:10.1021/pr300576s.
22. Mesuere, B.; Van der Jeugt, F.; Willems, T.; Naessens, T.; Devreese, B.; Martens, L.; Dawyndt, P. High-throughput metaproteomics data analysis with Unipept: A tutorial. *J Proteomics* **2018**, *171*, 11–22, doi:10.1016/j.jpro.2017.05.022.
23. Liao, B.; Ning, Z.; Cheng, K.; Zhang, X.; Li, L.; Mayne, J.; Figeys, D. iMetaLab 1.0: a web platform for metaproteomics data analysis. *Bioinformatics* **2018**, *34*, 3954–3956, doi:10.1093/bioinformatics/bty466.
24. 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 Prophan. *Nat Protoc* **2020**, *15*, 3212–3239, doi:10.1038/s41596-020-0368-7.
25. Karaduta, O.; Glazko, G.; Dvanajscak, Z.; Arthur, J.; Mackintosh, S.; Orr, L.; Rahmatallah, Y.; Yeruva, L.; Tackett, A.; Zybailov, B. Resistant starch slows the progression of CKD in the 5/6 nephrectomy mouse model. *Physiol Rep* **2020**, *8*, e14610, doi:10.14814/phy2.14610.
26. Gurdeep Singh, R.; Tanca, A.; Palomba, A.; Van der Jeugt, F.; Verschaffelt, P.; Uzzau, S.; Martens, L.; Dawyndt, P.; Mesuere, B. Unipept 4.0: Functional Analysis of Metaproteome Data. *J Proteome Res* **2019**, *18*, 606–615, doi:10.1021/acs.jproteome.8b00716.
27. Huber, W.; Carey, V.J.; Gentleman, R.; Anders, S.; Carlson, M.; Carvalho, B.S.; Bravo, H.C.; Davis, S.; Gatto, L.; Girke, T.; et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* **2015**, *12*, 115–121, doi:10.1038/nmeth.3252.
28. Zybailov, B.L.; Glazko, G.V.; Rahmatallah, Y.; Andreyev, D.S.; McElroy, T.; Karaduta, O.; Byrum, S.D.; Orr, L.; Tackett, A.J.; Mackintosh, S.G.; et al. Metaproteomics reveals potential mechanisms by which dietary resistant starch supplementation attenuates chronic kidney disease progression in rats. *PLoS One* **2019**, *14*, e0199274, doi:10.1371/journal.pone.0199274.
29. Brink, L.R.; Mercer, K.E.; Piccolo, B.D.; Chintapalli, S.V.; Elolimy, A.; Bowlin, A.K.; Matazel, K.S.; Pack, L.; Adams, S.H.; Shankar, K.; et al. Neonatal diet alters fecal microbiota and metabolome profiles at different ages in infants fed breast milk or formula. *The American Journal of Clinical Nutrition* **2020**, *111*, 1190–1202, doi:10.1093/ajcn/nqaa076.
30. Gierse, L.C.; Meene, A.; Schultz, D.; Schwaiger, T.; Karte, C.; Schröder, C.; Wang, H.; Wünsche, C.; Methling, K.; Kreikemeyer, B.; et al. A Multi-Omics Protocol for Swine Feces to Elucidate Longitudinal Dynamics in Microbiome Structure and Function. *Microorganisms* **2020**, *8*, 1887.
31. Wang, M.; Monaco, M.H.; Hauser, J.; Yan, J.; Dilger, R.N.; Donovan, S.M. Bovine Milk Oligosaccharides and Human Milk Oligosaccharides Modulate the Gut Microbiota Composition and Volatile Fatty Acid Concentrations in a Preclinical Neonatal Model. *Microorganisms* **2021**, *9*, 884.

- 
32. Pan, S.; Hullar, M.A.J.; Lai, L.A.; Peng, H.; May, D.H.; Noble, W.S.; Raftery, D.; Navarro, S.L.; Neuhouser, M.L.; Lampe, P.D.; et al. Gut Microbial Protein Expression in Response to Dietary Patterns in a Controlled Feeding Study: A Metaproteomic Approach. *Microorganisms* **2020**, *8*, 379.
  33. Locey, K.J.; Lennon, J.T. Scaling laws predict global microbial diversity. *Proc Natl Acad Sci U S A* **2016**, *113*, 5970–5975, doi:10.1073/pnas.1521291113.
  34. Tanca, A.; Palomba, A.; Deligios, M.; Cubeddu, T.; Fraumene, C.; Biossa, G.; Pagnozzi, D.; Addis, M.F.; Uzzau, S. Evaluating the impact of different sequence databases on metaproteome analysis: insights from a lab-assembled microbial mixture. *PLoS One* **2013**, *8*, e82981, doi:10.1371/journal.pone.0082981.
  35. Li, X.; LeBlanc, J.; Truong, A.; Vuthoori, R.; Chen, S.S.; Lustgarten, J.L.; Roth, B.; Allard, J.; Ippoliti, A.; Presley, L.L.; et al. A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PLoS One* **2011**, *6*, e26542, doi:10.1371/journal.pone.0026542.