JOURNAL OF PROTEOMICS 97 (2014) 3–16

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jprot

Review

Metaproteomics of our microbiome — Developing insight in function and activity in man and model systems☆

Carolin A. Kolmeder[a,*], *Willem M. de Vos*[a,b,c]

[a]*Department of Veterinary Biosciences, University of Helsinki, P.O. Box 66, FIN-00014 Helsinki, Finland*
[b]*Department of Bacteriology and Immunology, University of Helsinki, P.O. Box 21, FIN-00014 Helsinki, Finland*
[c]*Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands*

ARTICLE INFO

Available online 24 May 2013

Keywords:
Metaproteomics
Human microbiome

ABSTRACT

We are all colonized by a large microbiome, a complex set of microbes that have intimate associations with us. Culture-based approaches have provided insights in the complexity of the microbial communities living on surfaces inside and outside the body. However, the application of high-throughput sequencing technologies has identified large numbers of community members at both the phylogenetic and the (meta-)genome level. The latter allowed defining a reference set of several millions of mainly bacterial genes and provided the basis for developing approaches to target the activity and function of the human microbiome with proteomic techniques. Moreover, recent improvements in protein and peptide separation efficiencies and highly accurate mass spectrometers have promoted the field of metaproteomics, the study of the collective proteome of microbial communities. We here review the approaches that have been developed to study the human metaproteomes, focusing on intestinal tract and body fluids. Moreover, we complement these by considering metaproteomic studies in mouse and other model systems offering the option to study single species or simple consortia. Finally, we discuss present and future avenues that may be used to advance the application of metaproteomic approaches to further improve our understanding of the microbes inside and around our body.
This article is part of a Special Issue entitled: Trends in Microbial Proteomics.

Contents

☆ This article is part of a Special Issue entitled: Trends in Microbial Proteomics.

* *Corresponding author*. Tel.: +358 9 191 57013; fax: +358 9 191 57033.
E-mail address: carolin.kolmeder@helsinki.fi (C.A. Kolmeder).

1874-3919/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jprot.2013.05.018

1. Introduction

Our planet is estimated to contain over 10^{30} bacteria and archaea [1] that in many cases are adapted to harsh conditions [2]. However, one of the microbial habitats that is receiving intensive interest is our own body, notably as there is an increasing awareness of its impact on our health [3]. Considerable technological progress, particularly massive parallel sequencing developments, has opened up the analysis of the microbial communities in our body, also known as the human microbiome. These could previously not be studied as they were too complex, refractory to culture-based studies, and relatively unknown and inaccessible. The use of first and next generation technology sequencing has provided rapid insight into the composition of the microbial communities based on 16S rRNA sequence analysis [4]. However, the most considerable progress has been made by applying next generation technology sequencing in advanced metagenomics studies that revealed the enormous genetic potential of the human microbiome [5–8].

The human microbiome is in constant contact with its host and the surrounding environment. A recent comparative study highlighted the differences between the microbial communities in and on the body of over 200 subjects [9]. The results revealed the highest diversity in the oral and intestinal communities that, however, showed low and high individual variations, respectively. To date the deepest and most comprehensive study has provided a total of 3.3 Mio unique genes in colonic samples of 124 Europeans [7]. In addition, metagenomic inventories of other body sites have been reported, including the oral cavity, stomach and upper intestinal tract [10–12]. These metagenomic studies are being complemented by the ever-increasing information on the genomes of single microbial species and the largest one reported over 100 genomes with 30,000 new genes [13]. Most studies have dealt with the human intestinal microbiota as this is the bodies' most densely populated and most complex ecosystem. The intestinal ecosystem consists of trillions of bacteria, which are derived from several thousands of species or species-like taxa, most of which have not yet been cultured [14,15]. Moreover, the intestinal microbiome is known to play an important role in our health and over 20 different diseases have been associated with the microbes in that ecosystem [16].

The next step after identifying the composition and the coding capacity of the human microbiome is to study their activity and functionalities in their environment by high throughput functional metagenomics approaches [14]. This is an important step as phylogenetic and metagenomic approaches may reveal candidate species and genes that may be important in certain conditions or diseases. However, they do not provide evidence for the actual involvement of these species or genes. Hence, functional approaches are needed that aim to identify the active molecules and species. These also provide the basis for studying the ecological interactions between the human microbial species on the one hand and the host on the other hand. This is of further importance as functional studies are expected to be instrumental in revealing mechanistic insight and hence developing treatment strategies for diseases associated with the human microbiota. However, the experimental approaches needed to address the functions are more challenging than the simple collection of metagenomic or other genome sequence information. Transcriptomic [17], proteomic (see Table 1) and metabolomic approaches [18] at the community level and hence termed meta-omics, have all been applied to observe activities of the human microbiota but only in very limited number of studies with low throughput. Metatranscriptomics studies are difficult as the prokaryotic mRNA is highly instable and rapid processing of body samples may be challenging. Metabolomics often suffer from the fact that all body fluids are in contact with cells and hence the metabolites are often quickly absorbed — a possible exception is urine that however may not provide dynamic information of the ecosystem. Moreover, metabolite analysis is rather challenging and needs a series of high throughput instruments and large databases that are incomplete and hence need to be complemented by de novo identification. In contrast, metaproteomics is addressing proteins that in general are rather stable and have the advantage to relate directly to the genetic code, notably in prokaryotes that are known to have limited post-translational processing. As a result, metaproteomics is expected to provide a reliable and high throughput as well as comprehensive and stable picture of the function of microbial communities.

Rapid developments in efficient protein separation combined with highly accurate and high throughput mass spectrometry analysis have stimulated the field of metaproteomics. In addition, metaproteomic approaches have gained enormously from the rapidly growing metagenomics databases and computing power. Hence, fast and reliable high throughput identification of peptide masses has, in theory, become straightforward. However, metaproteomics studies are just emerging as there are