microbial biotechnology

Minireview

Elucidation of complexity and prediction of interactions in microbial communities

Cristal Zuñiga, Livia Zaramela and Karsten Zengler*
Department of Pediatrics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0760, USA.

Summary

Microorganisms engage in complex interactions with other members of the microbial community, higher organisms as well as their environment. However, determining the exact nature of these interactions can be challenging due to the large number of members in these communities and the manifold of interactions they can engage in. Various omic data, such as 16S rRNA gene sequencing, shotgun metagenomics, metatranscriptomics, metaproteomics and metabolomics, have been deployed to unravel the community structure, interactions and resulting community dynamics *in situ*. Interpretation of these multi-omic data often requires advanced computational methods. Modelling approaches are powerful tools to integrate, contextualize and interpret experimental data, thus shedding light on the underlying processes shaping the microbiome. Here, we review current methods and approaches, both experimental and computational, to elucidate interactions in microbial communities and to predict their responses to perturbations.

Introduction

Microorganisms form complex and diverse communities, which inhabit almost every environment on Earth. These

microbial populations interact with each other in many different ways, creating complex and intertwined networks. The interaction between living cells and their environment forms the foundation for microbial diversity. In nature, microbial diversity is often accompanied by vast metabolic capability reflected in fundamental biogeochemical processes (McCalley and Sparks, 2009; Guidi *et al.*, 2016; Louca *et al.*, 2016). Microorganisms also interact directly with higher organisms. The recent boost in studies of host–microbe interactions has shed light on the crucial role the human microbiome plays in health and disease. The human microbiota shapes significant parts in the host physiology and provides essential functions, e.g. nutrient catabolism (Flint *et al.*, 2012), syntheses of vitamins (Degnan *et al.*, 2014) and modulation of the human immune system (Fischbach and Segre, 2016). Albeit the importance of microorganisms to human health and evolution, we currently possess limited knowledge about how microbes interact amongst each other and with their host and how these interactions contribute to health. Additionally, the response of microbial communities to perturbations (e.g. nutrient availability or response to environmental shifts) is poorly understood, thus limiting our ability to predict the responses of the microbiome to interventions.

For centuries, most of our knowledge that paved the way for our understanding of microbial cells has been gathered through approaches that focused on single microorganisms, i.e. pure cultures (Fig. 1). Even with the introduction of advanced molecular biology techniques, most work has been performed on cultures in isolation, studying single molecules, such as single genes, transcripts, proteins, metabolites (Fleischmann *et al.*, 1995; Lander *et al.*, 2001). While pioneers in microbial ecology described complex microbial systems early on, using e.g. microscopy and enrichment cultures (Cohn, 1875; Winogradsky, 1887; Beijerinck, 1895; Dobell, 1923), they were limited by the technical methods available at the time. Only recently have we been able to apply molecular methods to study complex microbial systems *in situ*, thereby gaining insight into the complexity of microbial networks (Fig. 1). In sync with the development of new experimental methods has been the advancement in
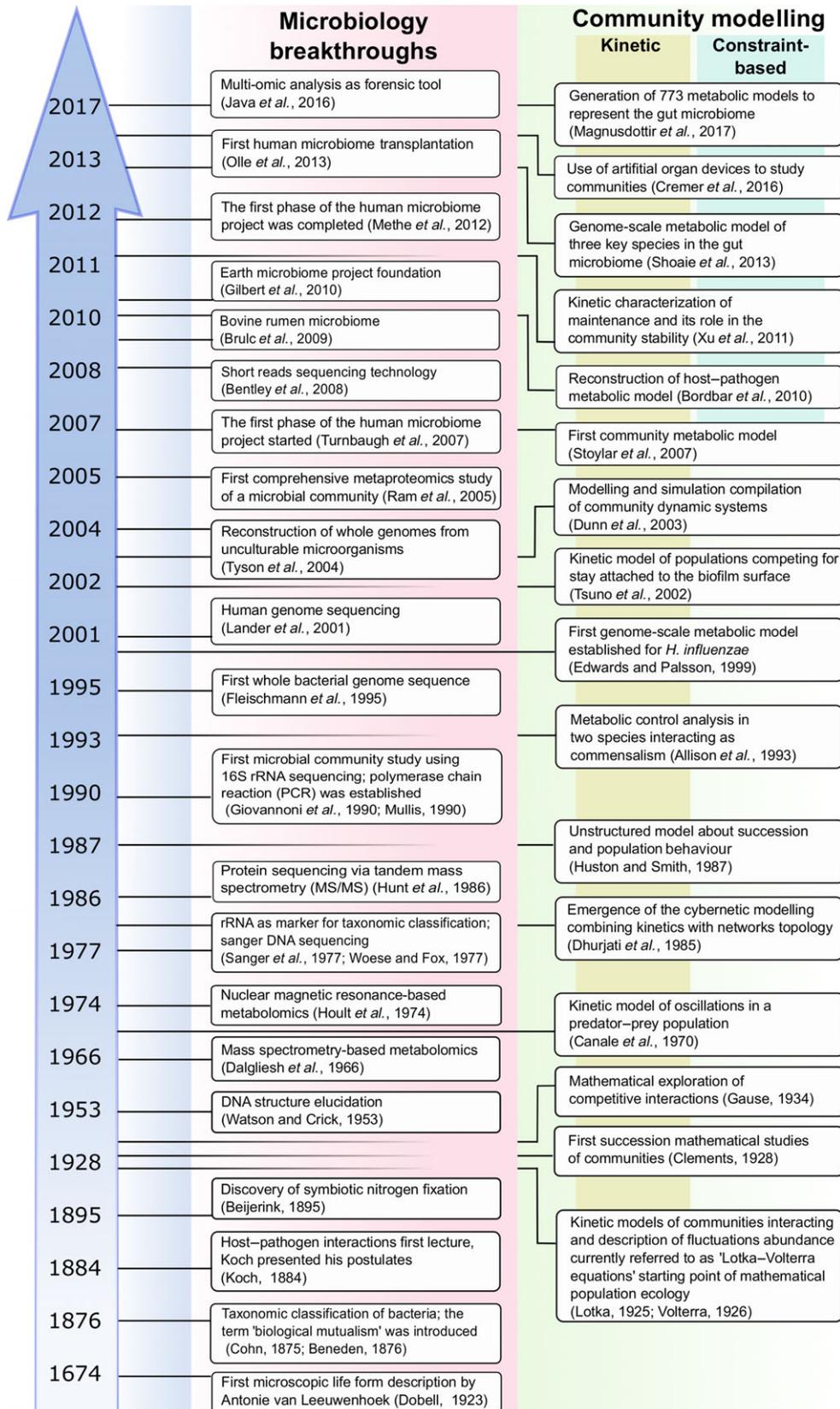
Fig. 1. Milestones in microbiology and computational modelling for the study of microbial communities.

computational methods, which enables us to contextualize complex data types. Systems biology, the computational modelling of biological systems deploying a holistic approach, emerged as an interdisciplinary field that enables the study of interactions between various components of a living cell (Bordbar *et al.*, 2014). While first applied to single organisms, systems biology approaches have been expanded to study the microbiome.

The comprehension of microbial communities in the context of systems biology beyond the description of individual parts of a living cell is a relatively new concept (Zengler and Palsson, 2012). In order to help unravel properties inherent to microbial interactions, systems biology approaches dynamically and simultaneously investigate the multiple operational components of the community. The emergence of culture-independent omic approaches, e.g. metagenomics, metatranscriptomics, metaproteomics and metabolomics, has increased our ability to probe microbial community composition and function *in situ.* These multi-omic approaches have changed our ability to investigate biological systems, providing essential experimental data to reconstruct and constrain computational models. At the same time, integrative data analysis methods, fundamental to elucidate mechanisms beyond individual interactions in microbial communities, have been developed and deployed (Embree *et al.*, 2015). Lastly, studying the microbiome using systems biology tools, such as predictive computational models, allows generating testable hypothesis on a community level (Magnúsdóttir *et al.*, 2016). While systems biology approaches have successfully been

deployed for single organisms to design and optimize the production of diverse products such as antibiotics, alcohols and amino acids, these approaches can potentially be extended to multispecies communities (Perez-Garcia *et al.*, 2016; von Kamp and Klamt, 2017). In this review, we discuss the current state of fundamental mechanisms elucidating and shaping the diversity and structure of microbial communities. We will highlight relevant biotechnological examples applying model-driven multi-omic analysis. Furthermore, we will provide an outlook on the development of predictive models for complex communities combining the use of omic tools in an iterative design–build–test–learn cycle.

Types of interactions

In nature, organisms rarely exist in isolation. Rather, they form complex relationships that are developed through evolution and adaptation to diverse ecosystems (Dunn *et al.*, 2003; Tan *et al.*, 2015). This complex network of interactions defines how a community is assembled and maintained spatially and temporally. Different types of interactions have been defined based on the benefit for an organism. Interactions thus can be beneficial, neutral or disadvantageous for the individual organism (Fig. 2). The best-studied type of interaction in microbiology is beneficial mutualism. A prime example for this type of interaction is syntrophy (Morris *et al.*, 2013), in which all community members are sustained by the activity of the other(s). One example of mutualism is the host–bacterial interaction in the human intestine, where the host gains
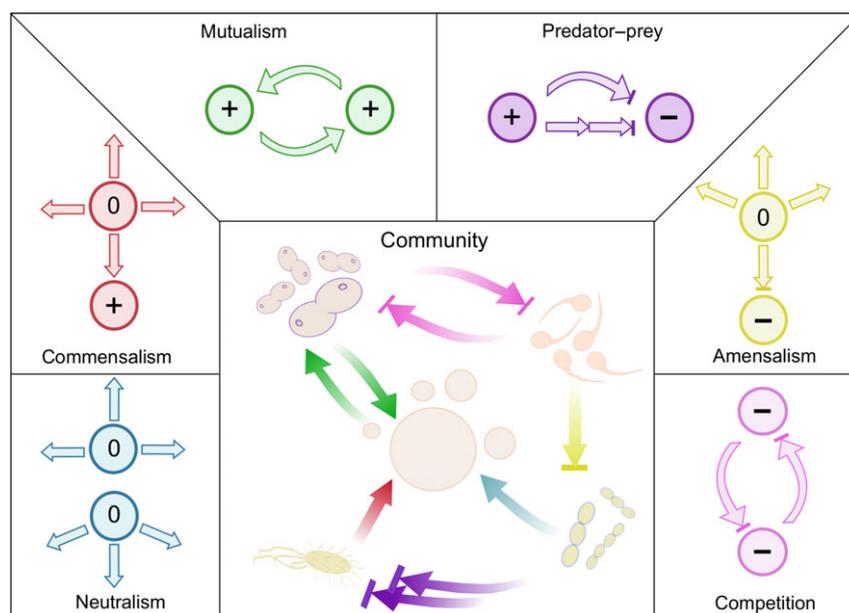

Fig. 2. Interactions amongst microorganisms. Within a complex community, the members of the population may engage in multiple interactions at the same time. 0 refers to neutral interaction, while + and – depicts a positive or a negative outcome of interaction.

energy from short-chain fatty acids released by the microbes through fermentation of glycans provided by the host (Backhed *et al.*, 2005). Another interaction type, in which one member benefits and the other neither benefits nor has a disadvantage, is known as commensalism. Examples of commensalism include, for example, nitrification and methanogenesis (Allison *et al.*, 1993). Commensalism and mutualism can also include interactions where one organism consumes specific metabolites that are inhibitory for another organism (Dunn *et al.*, 2003). Cross-feeding of inhibitory metabolites is a variation of mutualism and commensalism where the products of one organism are the substrate for another and the uptakes of those become beneficial to one member (Germerodt *et al.*, 2016). Amensalism on the other hand is a relationship in which one member is disadvantaged and the other neither benefits nor is harmed. In plant–microbe communities, amensalism is often referred to as allelopathy and can involve the natural secretion of secondary metabolites, which inhibits the growth of surrounding microorganisms (Chaparro *et al.*, 2014). When members of a local environment do not significantly affect each other, the interaction is termed neutralism (Dunn *et al.*, 2003). If one population member is disadvantaged and another benefits, the interaction is termed predator–prey or parasitism. This relationship has been extensively studied throughout all forms of life (i.e. eukaryotes including plants, bacteria and viruses). Apparently, predator–prey interactions can stabilize populations by creating a cyclic trend between species as well as preventing the depletion of resources by fast-growing species (Canale, 1970; Alhumazi and Ajbar, 2005; Chen *et al.*, 2011). Finally, the interaction type in which all members are disadvantaged by the presence of others is referred to as competition (Gause, 1934). A model case of competition is two populations of autotrophic and heterotrophic bacteria compete for the oxygen, thus promoting a spatial competition for the attachment to the surface (Tsuno *et al.*, 2002). Defining these basic forms of interactions in complex communities can be a challenge as interactions are prone to change over time due to changing physiological and physical parameters such as nutrient availability. Furthermore, interactions can be dependent on the spatial organization of members of a community.

Multi-omic approaches to study microbial communities

Microbiome research has greatly benefited from recent breakthroughs in sequencing technologies as well as in improvements in proteomics and metabolomics, and accompanying data analysis tools (Franzosa *et al.*,

2015; Jansson and Baker, 2016). Gene amplicon sequencing, metagenomics and metatranscriptomics are currently applied to elucidate microbial community composition and function and to determine potential interactions (Fig. 3). While these methods promoted recent advances in microbiome research, they can only provide limited mechanistic insight (by itself) into the determination of the mode of interaction. In combination with computational tools and modelling approaches, these experimental approaches can be used to unravel intertwined interactions, provide mechanistic knowledge and predict community behaviour. In this section, we will discuss advantages and disadvantages of different omic approaches for the study of microbial communities and their potential application for quantitative analysis by computational modelling.

Gene amplicon sequencing

The first studies interrogating microbial community composition emerged decades ago, through the comparative analysis of ribosomal RNA (Lane *et al.*, 1985; Giovannoni *et al.*, 1990; Delong, 1992; Fuhrman *et al.*, 1992; Fig. 1). Gene amplicon sequencing is currently the fastest assay to perform taxonomic identification and phylogenetic profiling of microbial communities. This technique consists of genomic DNA extraction, followed by amplification of a conserved genomic region and sequencing. Phylogenetic classification based on gene amplicon sequencing of small ribosomal subunits (SSUs) for bacteria, archaea and eukaryotes (16S or 18S rRNA gene) or internal transcribed spacer (ITS) region for fungi are routinely applied to describe operational taxonomic units (OTUs) in microbial ecology (Walters *et al.*, 2016). SSUs are highly conserved sequences intercalated by hypervariable regions that can be used as phylogenetic marker. In addition, sequencing of hypervariable regions instead of the full-length gene became very popular for profiling microbial communities (Yang *et al.*, 2016). This technique has been successfully employed to investigate the taxonomic composition and phylogenetic diversity for very different samples, from geothermal springs (Ward *et al.*, 2017) and oil reservoirs (Lewin *et al.*, 2014; Sierra-Garcia *et al.*, 2016) to the human microbiome and host–microbiome-related diseases (Boguniewicz and Leung, 2011; Halfvarson *et al.*, 2017). The success and low cost of amplicon sequencing encouraged major initiatives, such as the Human Microbiome Project (HMP), American Gut Project (AGP) and the Earth Microbiome Project (EMP; Turnbaugh *et al.*, 2007; Gilbert *et al.*, 2010; McDonald *et al.*, 2015). These large sequencing efforts have already generated taxonomic profiles for tens of thousands of samples from different sides of the
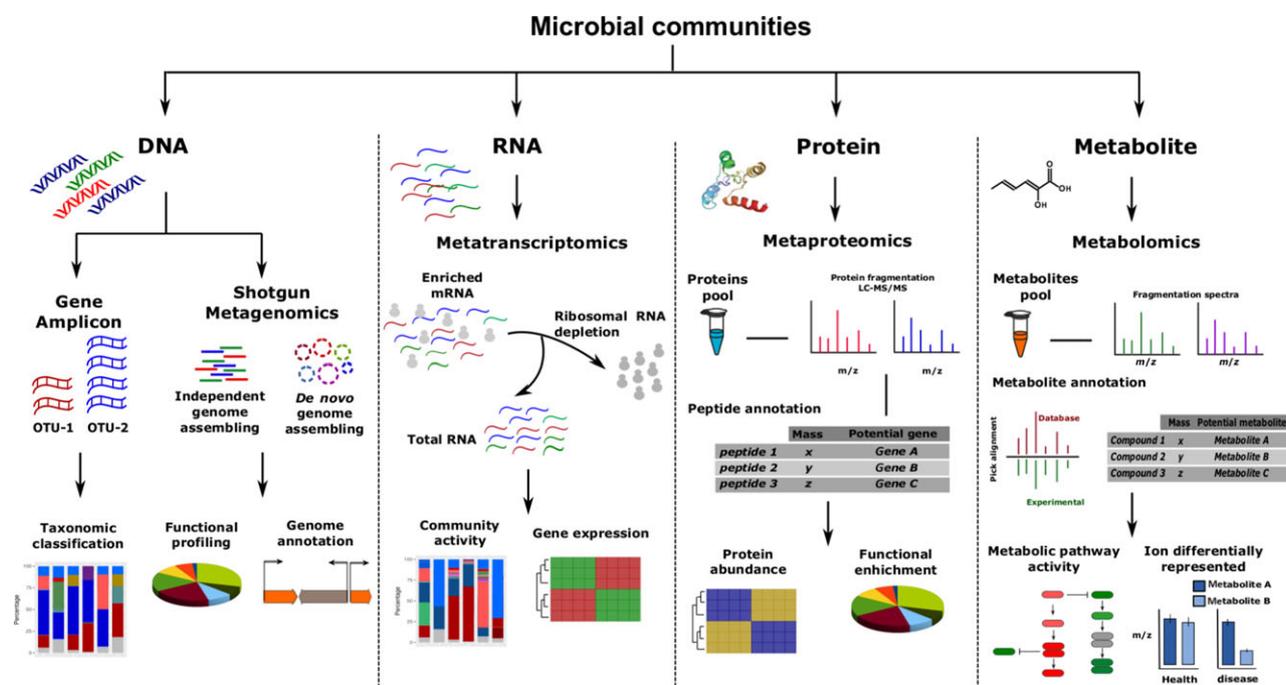
Fig. 3. Overview of different omics techniques. Different experimental approaches and computational strategies are applied for the study of microbial communities. During gene amplicon analysis, SSUs or ITS sequences are clustered into operational taxonomic units (OTUs) and the taxonomic identity is assigned for each OTU based on sequence homology against known sequences in a database. The resulting OTUs are used to calculate the relative abundance of each organism and quantify the population diversity between samples. In shotgun metagenomics, genomic DNA sequences can either be mapped to a reference database or used for *de novo* assembly of genomes. The recovered genomes can be used to assign phylogeny, calculate the relative abundance of the identified genome and assess the functional capability. In metatranscriptomics, messenger RNA (mRNA) is used to generate complementary DNA libraries that can either be mapped to reference genomes to generate gene expression profiles. These expression profiles are used to identify active pathways, genes and organisms. In metaproteomics, mass spectrometry and fragmentation are used to reveal the amino acid sequence of peptides. The identified peptides are associated with full-length proteins by sequence homology searches against a reference database. Similar to metatranscriptomic analysis, protein expression profiles can be used to identify active pathways as well as active organisms. In metabolomics, metabolites are separated using chromatography techniques and identified and quantified using mass spectrometry. Similar to metaproteomics, the comparison between fragmentation profiles and reference databases is used to annotate the metabolic compound. Enrichment and clustering analysis can be applied to reveal patterns between sets of samples or to identify condition-dependent compounds.

human body, the human gastrointestinal tract, as well as diverse samples from around the world.

The advantages of gene amplicon-based sequencing are low cost, easy sample preparation protocols and broad availability of bioinformatics tools (Caporaso *et al.*, 2010). Although the gene amplicon sequencing has provided unprecedented insight into microbiome composition, this technique can only provide community member profiling and indirect insights into functionality through comparison with reference genomes (Brooks *et al.*, 2015). While universal primers can readily be designed to cover bacterial and archaeal 16S rRNA genes, 18S rRNA and ITS regions in eukaryotes are more variable than 16S rRNA genes in bacteria and archaea, which makes universal primer design and thus eukaryotic diversity surveys challenging (Wang *et al.*, 2014). Additionally, shallow resolution based on amplicon size makes it difficult to identify closely related species

belonging to the lowest level of the taxonomic classification or in some cases to the same genus. While unambiguous identification at the strain level is not possible by 16S rRNA (Brooks *et al.*, 2015), this resolution is often critical for diagnostic in human health, e.g. to identify and discriminate a certain pathogen from a commensal strain. The importance of accurate identification at strain level has been highlighted in model-driven analysis of multiple bacteria strains, such as *Staphylococcus aureus* (Bosi *et al.*, 2016) and *Escherichia coli* (Monk *et al.*, 2013). These studies found significant differences in metabolic capabilities amongst different strains, such as amino acids and vitamin autotrophies, which can be related to pathogenesis. The correct strain identification is vital for the systematic understanding of microbial communities, which is the first step in the iterative cycle to unravel community interactions and dynamics (see Fig. 4).

Whole-metagenomic shotgun sequencing

Whole-metagenomic shotgun sequencing is a powerful metagenomics technique to study microbial communities beyond describing OTU composition of a sample. For this, genomic DNA from all cells in a community is extracted, sheared into short fragments and subsequently sequenced.

The short DNA sequences (i.e. reads), usually smaller than 250 bp, can be used to taxonomic classification by recovering the reads aligned to taxonomically informative genomic loci (16S or 18S rRNA, ITS or conserved proteins; Meyer *et al.*, 2008; Truong *et al.*, 2015), to estimate relative microbial abundance, to generate functional profiling and to reveal genomic content of unknown microbes
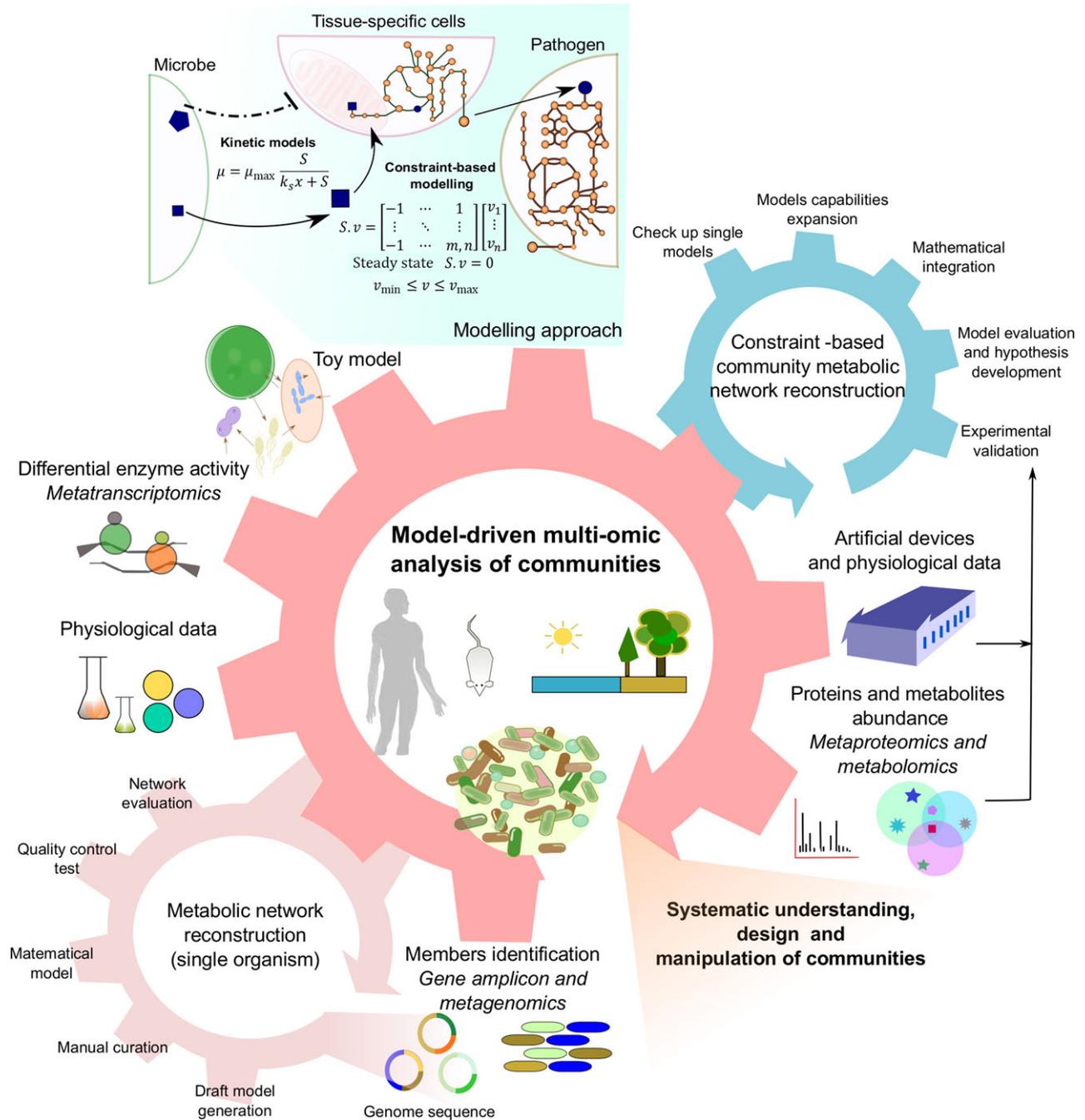
Fig. 4. Integration of experimental data and *in silico* analysis in a community systems biology approach. The first step in the iterative workflow encompasses the identification of individual community members, followed by the creation and validation of single metabolic networks. Subsequently, a predictive community model is constructed and validated. Various experimental and computational methods are interspersed along the design–build–test–learn cycle to unravel community interactions and predict community dynamics.

(Abubucker *et al.*, 2012; Hug *et al.*, 2016). The identification of unknown microorganisms usually depends on *de novo* genome assembling methods that combine overlapping reads to generate contiguous fragments of DNA sequence (i.e. contigs). Assembling methods can be time-consuming, require robust computational resources and deep-sequencing data sets, and rely on differential binning tools to group contigs into potential genomes (Jansson and Baker, 2016). Despite this, *de novo* genome assembling is fundamental to create compendium genomes representing not yet cultured microorganisms, to generate comprehensive organism-specific metabolic mapping (Sangwan *et al.*, 2016) and to detect larger complex genomic features such as polyketide synthase (PKS), cluster regulatory interspaced short palindromic repeats (CRISPR) and non-ribosomal peptide synthase (NRPS) gene clusters (Vollmers *et al.*, 2017).

Another, less time-consuming analysis method for metagenomics data is the reference-based mapping method, which facilitates comparisons amongst multiple samples (Abubucker *et al.*, 2012). This method usually involves mapping reads to reference genomes and interpreting the mapping results in terms of relative abundance and functional profiles of the sample. Hence, independent genome assembly outcomes depend heavily on genome reference databases and often present high false-positive rates due to regions of local homology (Carr and Borenstein, 2014). Both metagenomic analysis approaches have been applied to simultaneously investigate the metabolic requirement of individual members in a microbial community and to determine potential microbe–microbe and microbe–host interactions (Greenblum *et al.*, 2012; Embree *et al.*, 2015).

De novo genome assembling studies have been used for downstream community modelling (Biggs *et al.*, 2015; Embree *et al.*, 2015; Magnúsdóttir *et al.*, 2016). The metagenomic results were used as a genome platform to reconstruct and validate genome-scale metabolic networks for these communities. In turn, independent genome assembly methods have been extensively used to infer the metabolic repertoire of different microbiomes (Greenblum *et al.*, 2012; Levy and Borenstein, 2013). Usually, experimental data is analysed using a metabolic analysis network, in which the microbiome is treated as a single 'independent' biological system (Abubucker *et al.*, 2012; Greenblum *et al.*, 2012). This approach does not take into account the contribution of each organism in the community, ignoring the boundaries between cells and compartmentalization of various metabolites. Although simplified models are subject to inaccuracies and noises, they can provide valuable insight into the metabolic potential of the entire system (Ji and Nielsen, 2015). For example, topological analysis of metabolic networks, observing changes in connections and centrality of the nodes (genes or enzymes), can be helpful to identify genes/functions associated with different conditions (Tan *et al.*, 2015; Steffani-Vallejo *et al.*, 2017).

While metagenomic sequencing can provide relevant information related to the taxonomic profile and the metabolic capability of a community, the data do not provide information about viability of the microbial community members. It can be essential to complement sequencing with experimental methods that distinguish metabolically active from inactive cells. Several methods targeting active microbial cells have been discussed in recent reviews (Franzosa *et al.*, 2015; Singer *et al.*, 2017). *In situ* microbial activity methods can be potentially applied to study microbial population. These methods can differ dependent on the target molecule (DNA, RNA, protein and metabolites) and the biological process to be investigated (DNA replication, cell division, transcription, amino acid biosynthesis). Most of these techniques such as bromodeoxyuridine labelling, DNA-SIP (stable isotope probing), iREP (index of replication), RNA-SIP, protein-SIP are usually restricted to samples with low complexity or provide information about a limited number of community members. Furthermore, the methods are generally associated with high cost and require immediate processing of the sample. In addition, applying these techniques to studies in human can be challenging (Singer *et al.*, 2017).

Metatranscriptome: transcriptional expression in microbial communities

Conventional transcriptomic techniques measure messenger RNA (mRNA) levels in a simple biological sample (usually a single organism exposed to different conditions) and analyse the activity, i.e. differential expression, under these scenarios (Fig. 3). Metatranscriptomic approaches are an extension of this method and are measuring mRNA levels for all microbes in a given community. Accessing the expression level of the microbiome allows for determining its functional profile, inferring the biological role, and thus helps to decipher the type of interactions microbes are engaged in and to relate this information to metabolic pathways. This technique has been successfully applied to study diverse microbial communities. For example, metatranscriptomics has been used to determine the relationship between the oral and the gut community (Franzosa *et al.*, 2014) and to study human microbiome-related diseases, e.g. periodontitis (Jorth *et al.*, 2014), colorectal cancer (Dutilh *et al.*, 2013) and acne vulgaris (Kang *et al.*, 2015).

Transcriptomic data have been used extensity to constrain metabolic models of single cultures or co-cultures and to gain insight into metabolic activity and predict metabolic fluxes (Bordbar *et al.*, 2014; Zielinski *et al.*, 2015). These model-driven analyses provided a

mechanistic interpretation for the better understanding of the relationships between diet, microbiota and host (Shoaie *et al.*, 2013). Computational and mathematical approaches, as transcriptional networks, have emerged as potential alternative to elucidate microbe–microbe interaction in complex communities. These approaches were extended to investigate transcriptional interactions in the human gut microbiome (Plichta *et al.*, 2016). Community-wide transcriptional regulation is suggested to be a mechanism for reducing competition and for niche segregation (Zelezniak *et al.*, 2015). For example, germ-free mice colonized with two major bacteria of the human distal gut, *Eubacterium rectale* and *Bacteroides thetaiotaomicron*, were used to demonstrate how these bacteria change the whole-transcriptional profile to adapt to the presence of each other (Mahowald *et al.*, 2009). A similar approach was applied to analyse the metatranscriptomics data from 233 human faecal samples. Transcriptional interactions between pairs of coexisting gut microbes revealed the transcriptional changes, which led to a reduced expression of orthologous functions between interacting species pairs. Specific species–species transcriptional interactions were enriched for essential functions, such as butyrate biosynthesis, ATP-binding cassette (ABC) transporters, flagella assembly and bacterial chemotaxis (Plichta *et al.*, 2016).

Metatranscriptomic approaches are advantageous for the study of microbial communities *in situ*. Despite this, the approach has its pitfalls, one of them being the common practice to use mRNA expression levels as proxy for levels of corresponding proteins (Guimaraes *et al.*, 2014). Comparisons amongst transcriptional data and proteomics data from different biological samples showed that the correlation between mRNA levels and protein abundance can vary dramatically (0.2–0.9; Maier *et al.*, 2009). This nonlinear relationship can be explained by various reasons, such as half-life time of RNAs and proteins, post-transcriptional and translational efficiency and experimental error and noise. Ribosomal footprint profiling (RFP) approach that uses the ribosomal machinery genomic occupancy to provide a high-resolution quantitative profile of translation across the transcriptome can circumvent part of these obstacles (Latif *et al.*, 2015; Ingolia, 2016). However, up to now this method has not been applied to any microbiome study.

Metaproteomics: large-scale characterization of the protein content in microbial communities

Metaproteomics approaches aim to measure protein abundance in all cells in a microbial community. Usually, proteomic methods consist of three main steps: sample fractionation, protein separation and protein identification/quantification. Technological advances of molecular separation methods coupled with mass spectrometry enabled the study of the proteome of microbial communities. Such advances include two-dimensional polyacrylamide gels coupled with mass spectrometry (2D nano-LC/MS-MS), liquid chromatography coupled with mass spectrometry (LC-MS/MS) and matrix-assisted laser desorption/ionization (MALDI) coupled with time-of-flight mass spectrometer (TOF). Metaproteomics has been applied to examine different microbial communities from various environments (Ng *et al.*, 2010; Zampieri *et al.*, 2016), including the human gut, small intestine, colon, liver and adipose tissue (Mardinoglu *et al.*, 2015; Xiong *et al.*, 2015).

For example, metaproteomics was applied to understand the symbiotic relationship between human cells and the microbial community in the gastrointestinal tract (Young *et al.*, 2015). The proteome of human cells and microbes was simultaneously monitored over time throughout early development of the preterm infant microbiome. This study revealed that in response to early microbiome colonization, human cells synthesize an increased amount of proteins responsible for epithelial barrier function and antimicrobial activity. In addition, neutrophil-derived proteins were detected in high abundance, indicating activation of the human innate immune response (Young *et al.*, 2015). Similar to transcriptomic data, proteomic data have been successfully applied to constraint metabolic models and increase the prediction accuracy (O'Brien *et al.*, 2014), as well as to reconstruct tissue-specific models. Twenty-eight metabolic models were reconstructed for the small intestine, colon, liver and adipose tissues based on proteomics data. These model-driven proteomics analyses showed that the gut microbiota directly influences host amino acid and glutathione metabolism in mice (Mardinoglu *et al.*, 2015).

Although recent technological advances in mass spectrometry have allowed high-throughput proteome measurement in a fast and cost-effective way, this technique has technical limitations. One of them is the protein identification resolution. For example, the estimated number of unique proteins in human-associated microbiome samples is in the order of millions, but only a few thousand can be identified using current methods (Erickson *et al.*, 2012). The assessment of proteins with different features, e.g. different solubility or trans-membrane domain, requires specific protocols and multiple sample preparations to generate a complete proteome map. Various alternative methods to increase proteome coverage and resolution have become recently available (Wilmes *et al.*, 2015; Vincent and Postovit, 2017).

Metabolomics: tapping into the metabolic diversity

Metabolite exchange is often assumed to be a primary driver shaping microbial communities. While changes in transcript levels and protein abundances can usually be

associated with one or a few genes, often changes in metabolic fluxes and metabolite levels are the final result of complex interactions (Palsson and Zengler, 2010). Metabolites produced and secreted by living cells have tremendous physicochemical diversity, and can usually differ by multiple orders of magnitude in concentration. Due to the dynamic nature of these compounds, their concentrations and compositions can change rapidly, which makes detection and data analysis a challenging task (Ponomarova and Patil, 2015). Hence, the detection of all metabolites in a biological sample requires a combination of several sample preparation methods and often analysis on multiple platforms. Although thousands of metabolites can be detected in a single experiment, <2% of all compounds from an untargeted metabolomics experiment can currently be annotated (Silva *et al.*, 2014).

Improved tools to annotate and integrate metabolomics data sets have been recently developed. One of them is the GNPS (Global Natural Products Social Molecular Networking) that similar to BLAST (Basic Local Alignment Search Tool) for nucleotides and amino acid sequences, assists the storage, analysis and dissemination of public MS/MS data, hence contributing in the identification of known compounds (Wang *et al.*, 2016). Another tool is XCMS online that provides simplified metabolomics data processing and analysis (Gowda *et al.*, 2014) and more recently allows multi-omics analysis by integrating metabolic pathway information with metabolomics, metatranscriptomics and metaproteomics data (Huan *et al.*, 2017). Despite some of these limitations and others regarding technical and analytical requirements, metabolomics is the method of choice for the discovery of new antibiotics (Nothias *et al.*, 2016) and to evaluate the effect of bacterial metabolites on human health (Wang *et al.*, 2011). Host–microbe metabolomics is an emergent area in systems biology with a high potential impact on the discovery of novel drugs and biomarkers (Heinken and Thiele, 2015). The use of metabolomics has been fundamental in identifying new biomarkers for various human diseases, such as chronic kidney disease (Hocher and Adamski, 2017), coronary heart disease (Yao *et al.*, 2017) or Parkinson's disease progression (LeWitt *et al.*, 2017).

In addition to an untargeted approach, targeted metabolomics is often used to identify and/or quantify a certain metabolite of interest. Using targeted metabolomic approaches to determine metabolism intermediates, e.g. acetate, ethanol, butyrate, lactate, succinate or amino acids, are usually applied to validate metabolic model prediction for single organisms (Shoaie *et al.*, 2015), as well as for the development of novel constraint-based metabolic methods (Bordbar *et al.*, 2017). However, in the context of microbial communities, complex microbe–

microbe or microbe–host interactions can make the task of tracking the origin of a certain metabolite almost impossible, requiring pulse–chase labelling experiments. Thus, even when the metabolite can be identified, assigning this metabolite to a single organism (microbe or host) can be challenging. The combination of metabolomics with metagenomics and metatranscriptomics can contribute towards a comprehensive understanding of metabolomics data and guide the annotation process.

Community systems biology

The previous sections exemplified current approaches to elucidate microbial diversity, community composition, functionality and activity; all focused on unravelling the interactions that shape the microbiome. Despite all the recent advances in omics technologies, no single approach can paint a comprehensive picture of the complex interaction network in a microbial community. Therefore, interdisciplinary, multi-omic tools and analyses are necessary for elucidating the role of the different community members and to unravel mechanisms and interactions that drive community composition and dynamics. As it is increasingly easier and more cost-effective to generate large, multi-omic data sets, data integration and analysis have become a time-consuming and often challenging part in the investigation of microbial communities (Palsson and Zengler, 2010). Although still in its infancy, systems biology approaches applied to communities have attracted more and more attention over the last years (Levy and Borenstein, 2013). At the same time as systems biology approaches for single bacterial and eukaryotic microorganisms have matured, e.g. through the reconstruction of comprehensive predictive models (Islam *et al.*, 2014; Bosi *et al.*, 2016; Kerkhoven *et al.*, 2016; Levering *et al.*, 2016; Zuñiga *et al.*, 2016; Bartell *et al.*, 2017), a Community Systems Biology (CoSy Biology) approach has emerged (Zengler and Palsson, 2012). CoSy Biology builds on the success of systems biology approaches for single organisms. For example, multi-omic data, as described above, can be used in the context of computational models to gain insight into community interactions (Nagarajan *et al.*, 2013; Embree *et al.*, 2015). It should be noted that these models generally do not consider biological noise (i.e. variation of individual cells and their growth stage) but rather assume unified growth phenotypes of members within a community. In the following sections, we will describe current efforts in the construction of mathematical models for the systematic understanding of communities, and the development of robust community models guided by multi-omics data. In analogy to systems biology approaches for single organisms, we propose an iterative design–build–test–learn cycle (Nielsen and Olson, 2002; Paddon and

Keasling, 2014) that would aid the systematic understanding of communities by combining experimental and *in silico* tools in an iterative fashion (Fig. 4).

Kinetic modelling

High-throughput community studies, either nucleic acid-based, metabolome-based or phenotypical-based, greatly outnumber model-based analyses (Röling *et al.*, 2010). The absence of robust mathematical tools to systematically contextualize experimental information has limited our understanding of community interactions (Gonzalez *et al.*, 2011). However, significant progress has recently been made towards predicting various features of communities, such as community interactions and potential spatial organization, in addition to more traditional abundance predictions. Here, we recapitulate previous efforts to understand and elucidate the function of communities using mathematical models.

Since the introduction of the concept of biological mutualism by Van Beneden one and a half centuries ago (Beneden, 1876) (Fig. 1), different models emerged in the 1920s to quantitatively describe community interactions (Lotka, 1925; Volterra, 1926). The most extensively used approach has been kinetic modelling, which provides a dynamic picture of how the components of a studied system interact and respond to certain stimuli over time (Dunn *et al.*, 2003; Xu *et al.*, 2011). Kinetic models can be classified as unstructured or structured and stochastic or deterministic depending on their scope, the parameters considered and the way in which they are solved. To date, structured and unstructured kinetic models are the primary choice to simulate inhibition processes within communities (Tabiś *et al.*, 2014). To create structured or unstructured models, a well-characterized community is needed in order to proceed with model computational analysis, in which often the favoured interaction modelled is mutualism. The characterization involves accounting for a number of growth-related parameters such as experimentally determined yields, substrate uptake rates, and biomass, substrate and/or product concentrations, as well as reaction saturation and inhibition constants. Mathematically, these models involve solving a set of ordinary differential equations or partial differential equations, which represent the conservation of mass in internal and external cell processes like mass conversion or exchange. The experimental and estimated parameters included into the model will depend on the community complexity. When experimental measurements of all parameters are not available, which is often the case for single microorganisms and complex communities, approximation of parameters becomes time-consuming and computationally intensive (Bordbar *et al.*, 2014).

Unstructured kinetic models

Unstructured models simulate dynamics of biomass production by treating each individual organism as a 'black box'. These models are mainly based on Monod-type relationships. Unstructured kinetic models vary in complexity, from the simplest representation of a population's dynamic behaviour (basic growth principles; Huston and Smith, 1987; Dunn *et al.*, 2003) to models involving vast amount of parameters related to environmental conditions, fitness, sensing and adaptation (Cabello *et al.*, 2014). The most complex models describe the interplay between physical and chemical properties of the studied condition, as well as participant-specific biological parameters such as rates of maintenance, inhibition, biomass formation or cell death (Alhumazi and Ajbar, 2005; Wade *et al.*, 2016). A number of unstructured growth models for communities have been produced over the decades (Faust and Raes, 2012), the majority of which were applied to microbial ecology questions with emphasis on mutualistic systems (Wade *et al.*, 2016).

Unstructured models can be used to simulate various community interaction types such as inhibitory behaviours or cross-feeding. In 2016, Germerodt *et al.* (2016) studied the growth of six different *Escherichia coli* strains assuming cross-feeding, where each strain differed in its requirement to obtain certain amino acids. The Cellular Automaton of Bacterial Cross-feeding (CELL-ABC) tool was used for simulation purposes, using as input parameters growth rates, amino acid concentrations and amino acid diffusion rates. Over time, the different genotypes in the community became stable. The simulations allowed predicting the performance of species over time and the codependency between them (Germerodt *et al.*, 2016).

Furthermore, community models have also been created based on reconstructed devices that represent artificial organ systems, for example the 'minigut', which is a device that simulates the peristalsis of the human colon under controlled conditions (Cremer *et al.*, 2016). The minigut was deployed to study a commensal synthetic community of two *E. coli* strains, in which one strain was engineered to produce galactose and the other to consume this sugar (Fig. 4). Simulations predicted strain-specific abundances related to the position in the device over time, suggesting that flow rate and mixing play a critical role in shaping the microbiota of the colon.

Another intriguing example for unstructured model is the recently published model for a co-culture consisting of *Monoraphidium* and *Chlamydomonas* species that predicts the coexistence or exclusion (by abundance) of each community member. This work links the modern coexistence theory and the contemporary niche theory (Letten *et al.*, 2017). The modern coexistence theory is

based on the overlap of two mechanisms mediating the coexistence. The first is equalization in which the fitness between active species is considered, and the second is stabilization, which encompasses the environmental conditions surrounding the community. On the other hand, the contemporary niche theory considers two species competing for limited resources. The model considers empirical relationships between supply ratio, impact niche (weight of certain member on the environment) and requirement niche (nutrients demand) resulting in the prediction of the degree of stabilization and equalization of the community.

Structured kinetic models

The goal of structured models is to predict population behaviours based on each organism's capabilities. In structured modelling approaches, the biological system is broken down into specific compartments for each member of the population. Parameterization of these models may include such properties as the increase of an accumulated polymer (polyhydroxyalkanoates, lipids, etc.; Meeuwse *et al.*, 2012; Volova *et al.*, 2014), the mass transportation in and out of organelles (Cortassa *et al.*, 2009), the synthesis of outer membrane vesicles (Berleman and Auer, 2013) or the metabolite exchange amongst the community through the shared space (e.g. culture medium; Wang and Papenguth, 2001; Mee *et al.*, 2014). Modelling approaches follow basic principles such as mass and energy conservation (Resat *et al.*, 2011). In addition, structured models provide information about specific cellular components, such as metabolites, cellular structure and its organization. Examples of structured models include a model to simulate plasmid stability in a community containing wild type and recombinant microorganisms (Zhang *et al.*, 1997; Dunn *et al.*, 2003).

One of the primary challenges in kinetic modelling (for unstructured as well as structured kinetic models) of monocultures and microbial communities is the feasibility and accuracy of the reaction rate constant measurements such as Michaelis–Menten constant (Km), equilibrium constant (Keq) and turnover rate (Kcat) necessary to build and parameterize these models. Another challenge is the generalization of a workflow for parameter estimation (e.g. selection of automatic step size, or methods). Efforts are underway to address this limitation using machine learning-based methods to determine kinetic parameters (McGibbon and Pande, 2013; Du *et al.*, 2016), to elucidate chemical reactions (Kayala *et al.*, 2011) or to characterize complex reaction mechanisms (Blurock, 2003). However, these methods have primarily been applied to monocultures, and at large have not been expanded to include communities yet in

part due to the challenge to accurately measure required parameters in a community setting.

Constraint-based models

While simplified constraint-based models exist (e.g. the *E. coli* core model; Schellenberger *et al.*, 2011), the power of these models lies in the scope to represent biochemical and genetic features at genome scale. The vast majority of constraint-based models are genome-scale representations of the metabolism of an organism, rooted in the genome sequence of the organism and topology of its own metabolic network. There are two types of genome-scale models. One of which are the metabolic models (M-models) and the other one type is a more comprehensive version referred to as the ME-models. ME-models account for highly detailed protein synthesis and degradation considering transcription units, transcription factors, tRNAs and chaperones needed for translation and protein folding. These ME-models can also contain ion–metal enzymatic requirements for catalysis (Lerman *et al.*, 2012; O'Brien *et al.*, 2014) as well as translocation (Liu *et al.*, 2014). To date, ME-models have only been reconstructed for single species such as *Thermotoga maritima* (Lerman *et al.*, 2012) and *E. coli* (Liu *et al.*, 2014; O'Brien *et al.*, 2014).

M-models have been built for organisms for almost two decades (Edwards and Palsson, 2000), and the first community model was reconstructed a decade ago (Stolyar *et al.*, 2007). M-models are mathematically expressed by a numerical matrix of reaction stoichiometric coefficients assigned to their respective coordinates by reaction (columns) and involved metabolites (rows). The reaction stoichiometric coefficients, along with measured metabolite uptake and secretion rates, become the system constraints. Each constraint plays a critical role during simulation and prediction of growth rate depends on the accuracy of experimental data. High accuracy is especially important for determination of uptake and secretion rates, as these parameters are critical in the modelling. While these data are relatively easy to obtain for axenic cultures, those constraints are challenging to determine for individual members within a community (see Multi-omic approaches to study microbial communities). The lack of accurate measurements for individual members (e.g. uptake rate) is one of the greatest challenges in CoSy Biology. The lack of information can therefore require multiple assumptions to be made during the parameterization step. In the case of constraint-based community modelling, constraints such as uptake rates and the stoichiometric coefficients in the biomass reaction are parameters that often have to be obtained from the literature. The biomass reaction contains all biomass precursors (nucleotides, proteins, lipids,

carbohydrates), which means that the biomass reaction of microorganisms in pure culture is assumed to be identical to the composition of these organisms existing in communities (Nagarajan *et al.*, 2013). However, it is well known that different substrates or physical factors, such as temperature, can influence cellular composition, such as lipid content. The assumption that the biomass reaction is independent from environmental parameters can thus be misleading. However, most of the time organisms cannot be obtained as axenic cultures and thus the biomass reaction has to be obtained from closely related organisms (Zengler, 2008).

The reconstruction of a highly curated model for a single organism is a time-consuming process, while reconstruction of complex community models containing many members often relies on automatically generated models (Henry *et al.*, 2010). Auto-generated models usually contain the constraints and biomass objective function of *E. coli* independently of the metabolic features of the community member and its possible role in the community (Freilich *et al.*, 2011; Levy and Borenstein, 2013). Although auto-generated models represent a valuable resource, it is important to note that these models lack the high accuracy of manually curated models. Auto-generated models are subject of intensive gap-filling (i.e. adding reactions and metabolites without gene associations) generating redundant results for community models. Furthermore, gap-filling reduces representation of auxotrophies in the model, which have been shown to be critical for community assembly and maintenance (Embree *et al.*, 2015).

Once constraints are applied to the model, independent of the model reconstruction process, a convex optimization method called flux balance analysis (FBA) generates an optimal solution representing the flow of mass through the network. FBA identifies the optimal solution of a given objective function. Normally, this objective is the biomass reaction (i.e. growth) which encompasses all the metabolites necessary to produce biomass (Schellenberger *et al.*, 2011). FBA is based on the assumption of a steady state and is solved using linear programming. Constraint-based models successfully predict growth rates and gene essentiality, and can accurately represent the genotype–phenotype relationship.

As these metabolic models (M-models) are reconstructed at genome scale, they are well suited for the integration of high-throughput community data (e.g. metabolomics, metatranscriptomics), thus enabling organism-specific analysis and interpretation of multiomic data sets. Moreover, many microbial communities maintain relationships based on the exchange of metabolites, which can be assessed directly using constraint-based analysis, as these exchange fluxes are explicitly modelled under this framework. However, it is important to note that communities are assembled and maintained not only on metabolic exchanges alone and interactions can be manifold and intertwined depending on environmental conditions (Embree *et al.*, 2014; Cremer *et al.*, 2016; Hibberd *et al.*, 2017).

The COBRA (Constraint-Based Reconstruction and Analysis) toolbox (Schellenberger *et al.*, 2011) has been widely used to generate and solve single species and community constraint-based models (Bordbar *et al.*, 2010; Nagarajan *et al.*, 2013; Hamilton *et al.*, 2015). This tool allows the optimization of multiple objective functions at the same time, which is critical for communities that may be adapting to maximize the growth of multiple community members simultaneously. For example, direct interspecies electron transfer between two *Geobacter* species was simulated by optimizing the growth rate of both community members (Nagarajan *et al.*, 2013). In another example of multiple objectives, the main drivers in the interaction between the bacterium *Syntrophobacter fumaroxidans* and the methanogenic archaeon *Methanospirillum hungatei* were identified and attributed to the concurring exchange of H_2 and formate (Hamilton *et al.*, 2015). Embree *et al.* (2015) used individual models to quantify the simultaneous exchange of several metabolites between the five most abundant members in a complex methanogenic community and revealed the existence of an intertwined network of amino acid auxotrophies.

M-models can be used to predict potential competitive or cooperative interactions in microbial communities. Freilich *et al.* used large sets of auto-generated M-models and compared the probability of interaction between microorganisms from different ecosystems. This approach was based on similar principles of modern coexistence theory and the contemporary niche theory. The goal of the study was to use the available genome and metabolic information annotated for each identified bacteria to make predictions of community behaviour in lieu of determining strain-specific kinetic parameters (Freilich *et al.*, 2011).

Community M-models have also been reconstructed to gain insight into human host–pathogen interaction and to determine the effect the microbiome has as an important component of human health. The first community M-model describing host–pathogen interaction was the model of a human macrophage and the bacterium *Mycobacterium tuberculosis*. The model was able to represent three distinct pathological states and gave insights into the infection development (Bordbar *et al.*, 2014).

Other modelling efforts around communities for human health applications have been focused on the human gut and bacterial interactions. The first such work focused

on representative strains from the human gut, such as *Bacteroides thetaiotaomicron*, *Eubacterium rectale* and *Methanobrevibacter smithii* representing the phyla Bacteroidetes, Firmicutes and Euryarchaeota respectively. The resulting simulations of population densities amongst the species partially resemble the densities during colonization of each microorganism into germ-free mice in different ecosystems (Shoaie *et al.*, 2013). In a different approach, coexistence of community members was estimated by integrating metagenomic data with auto-generated reconstructions of 154 models of human gut species. The authors found that control over the health of the host does not change the observed microbial patterns, indicating that species interactions are not fully defined by the host (Levy and Borenstein, 2013). One recent achievement was the reconstruction of preliminary models for 773 human gut bacteria (Magnúsdóttir *et al.*, 2016). The authors used the AGORA (Assembly of Gut Organisms through Reconstruction and Analysis) toolbox to semi-automatically create 773 genome-scale metabolic models of the most representative microorganisms in the gut microbiome (see Fig. 1). These models are a resource and are also expected to facilitate the study of host–microbiome interactions as their nomenclature and structure are compatible with a reconstruction of human metabolism, Recon2 (Swainston *et al.*, 2016). While Recon2 is not a tissue-specific model and represents all human metabolism together, tissue-specific models have been reconstructed in the past and can be utilized for studying specific interactions of microbial communities in the human body (Ryu *et al.*, 2015; Schultz and Qutub, 2016).

Recently, various tools to increase the scope of community M-models have been developed, see for example a recent review by Biggs *et al.* (2015). While FBA-based modelling approaches assume a steady state, most communities are highly dynamic. To account for this, a dynamic flux balance analysis (dFBA) has been developed and extended to microbial communities to predict time-varying interactions between species and their effects on microbial composition. This approach is referred to as dynamic multispecies metabolic modelling (Zhuang *et al.*, 2011). The CASINO toolbox (Community And Systems-level INteractive Optimization) was created as a comprehensive platform to analyse gut microbial communities (Shoaie *et al.*, 2015). The authors used experimental data from 45 obese and overweight individuals to validate their predictions. The predictions quantitatively describe exchange of metabolites in response to diet intervention (Shoaie *et al.*, 2015).

The success of genome-scale models in predicting community interactions has been empowered by the high quality of single organism reconstructions used to generate community M-models. Presently, a challenge in constraint-based community modelling is the reconstruction of reliable models from either single-cell sequencing or from genomes obtained by differential binning of metagenomics data (Martinez-Garcia *et al.*, 2012). A current limitation of this approach is that microorganisms described solely by omics tools are typically poorly characterized, and thus modelling approaches rely on genome information only. The continuous contextualization of omic data will help to improve these models so that they can accurately predict how a microorganism or a metabolite benefits human health (see Fig. 4).

A hybrid modelling concept that combines kinetic and constraint-based modelling is cybernetic modelling (Dhurjati *et al.*, 1985). Here, the models account for metabolic regulation as in kinetic models, while considering a cellular objective function similar to constraint-based modelling. Cybernetic modelling has been extended to microbial community modelling in describing the interaction of the yeasts *S. cerevisiae*, *Pichia stipitis* and *Kluyveromyces marxianus* (Geng *et al.*, 2012). Those studies focused on the inference of enzyme kinetic parameters and reaction thermodynamics for the production of bioethanol within the metabolic pathways of interest in each member. Another hybrid model is known as whole-cell computational model (Karr *et al.*, 2012), which was used to predict certain stages of interactions between the human host and *Mycoplasma genitallium*. This model considers the main processes inside a cell, such as the synthesis of protein, RNA, DNA and metabolite levels. All processes are explained in the context of cell geometry, external stimulus, time and host–pathogen interaction (Karr *et al.*, 2012).

Computational tools for data enrichment and analysis

Both constraint-based and kinetic modelling approaches rely heavily on experimental data necessary to reconstruct and constrain these models. Advances in computational tool development have been going hand in hand with advances made in molecular biology. The latest advances in database development, network visualization, statistical approaches and other mathematical tools to produce and analyse high-throughput community data have played an essential role in CoSy Biology. For example, large databases including the human and the earth microbiome projects (Turnbaugh *et al.*, 2007; Gilbert *et al.*, 2010; Methé *et al.*, 2012) are now available. Furthermore, expanded metabolic databases such as KEGG (Kanehisa *et al.*, 2014), MetaCyc (Caspi *et al.*, 2014), GNPS (Wang *et al.*, 2016) and tools for identification and annotation of metabolites using Bayesian approaches (Silva *et al.*, 2014) became recently available online. In addition, new software and graphical

visualization tools have been generated to ease interpretation and presentation of big data (King *et al.*, 2015).

In addition to constraint-based and kinetic modelling, statistical models have also been employed to study community behaviour. One such variation of unstructured models is the various correlation approach (e.g. Pearson's product–moment coefficient, rank correlation coefficients, sensitivity and data distribution; Bersanelli *et al.*, 2016), which has been widely used for the analysis of omic experimental data. The main application of the correlation approach has been the analyses of species co-occurrence, diversity, as well as replication rates (Roling and van Bodegom, 2014; Brown *et al.*, 2016). Additionally, statistical correlations and other advanced statistical analyses (e.g. multivariable analysis) have been used to estimate community assembly, approximation of strain and expression abundance, taxa turnover and phylogenetic structure (Nemergut *et al.*, 2013). One key issue with the correlation-based approach is to establish methodologies oriented to avoid computational artefacts associated with the presence of low abundance organisms or low expressed genes within specific members of a community (Nemergut *et al.*, 2013).

Correlation-based methods require dimension reduction approaches. These approaches extract the linear relationships that best explain the correlation across omic data sets, and data variability including issues such as batch effects or outliers (Meng *et al.*, 2016). One example of such a method is the QIIME toolbox for the analysis of community sequence data. This toolbox explores data integration using principal component analysis (Caporaso *et al.*, 2010). A recent study used 16S rRNA gene sequence data analysed with QIIME to demonstrate the influence of pathogens on the microbiota of an arthropod (Abraham *et al.*, 2017). Arthropods are agents, which trigger human infections, e.g. granulocytic anaplasmosis. The study showed that external pathogens became active and manipulated the local environment (in this case, the microbiota of the arthropod) to facilitate pathogen infection.

Potential biotechnological applications

Insights into microbial communities are poised to change various aspects of our daily life, including human health care, animal husbandry, agriculture and fermented food industry (Gilbert *et al.*, 2010; Sekirov *et al.*, 2010; Fouts *et al.*, 2012; Ivey *et al.*, 2013; Shoaie *et al.*, 2013; McDonald *et al.*, 2016; Knight *et al.*, 2017). Multi-omic data generation and subsequent computational modelling are key to our understanding of the microbiome. The hope is that this knowledge will enable direct or indirect modulation of the microbiome (e.g. in the form of

pro-, post- or prebiotics) to benefit human life, improve biotechnological application or promote a healthier environment. Model-driven data analysis has already associated host–microbiome features to human wellness and health and has even been deployed for forensic purposes (Sekirov *et al.*, 2010; Shoaie *et al.*, 2015; Javan *et al.*, 2016a,b). Additionally, microbiomes of livestock and plants have also been exploited to promote increase in agricultural productivity (Andreote and Pereira e Silva, 2017; Thibodeau *et al.*, 2017). It has thus been a long-standing goal of the biotechnology and medical industry to understand and control the function of microbial communities. Besides bacteria and fungi, viruses represent an inherent part of the microbiome. The viral microbiome not only plays a major role in the environment (Zeigler Allen *et al.*, 2017) and human health (Cadwell, 2015) but also is of critical importance for the bioprocess industry (Marco *et al.*, 2012). Viral contamination of cell-based fermentations is responsible for hundreds of millions of dollars in lost revenue per year. Computational modelling of the virome in context of the microbiome could provide new strategies to combat viral infections (Maynard *et al.*, 2010). This section focuses on biotechnological applications enabled by new insights into microbial communities and the human microbiome (see Fig. 1).

Data-driven advances in human health

Human individuals have a unique microbiome that has a profound impact on their health. As such, personalized medicine has the potential to play an important role in guiding wellness and providing possible biotherapies targeting the microbiome (McDonald *et al.*, 2016). Microbiome-based therapies (McDonald *et al.*, 2016; Nature-Biotechnology, 2017) can be classified into the following categories: prebiotic, probiotic, post-biotic and modulation treatments. However, identification of what constitutes a healthy microbiome, which microorganism is beneficial for a defined outcome and determining the causal effects of a manipulated microbiome remain key challenges for the field.

Prebiotics are composed primarily of non-human digestible food or ingredients that promote the growth of beneficial microorganisms inside the host. Recent studies have shown promising results for the use of prebiotics to counter obesity in adults and improve wellness in infants (Lyon *et al.*, 2011; Koleva *et al.*, 2015; Nature-Biotechnology, 2017). Probiotics consists of individual beneficial microorganisms or combinations thereof (Gibson *et al.*, 2017). Probiotics can be either consumed or applied topically and have been shown to confer positive effects in a variety of areas of human health (Holz *et al.*, 2017). Natural probiotics have a long history of success in human gut health and have shown promise in treating

inflammatory bowel disease, necrotizing enterocolitis, allergies, urogenital problems and skin imbalances (Nakatsuji *et al.*, 2017). However, administration of probiotics does not necessary re-establish the microbiota to 'normal' levels (Petschow *et al.*, 2013; Holz *et al.*, 2017). Currently, there is still a lack of understanding about the mode of action and collateral effects of probiotic treatment on human health. Post-biotics consist of metabolic products of non-live probiotic microorganisms that have been shown to grant favourable effects onto the host health (Klemashevich *et al.*, 2014); especially, the short-chain fatty acids, acetate, propionate and butyrate have shown encouraging outcomes for the treatment of different diseases, such as colitis, arthritis, asthma, gout or pneumonia (Thomaz *et al.*, 2016).

Finally, microbiome transplants, which can be classified as modulation methods (Olle, 2013), involve the administration of live microorganisms from a donor into a recipient in order to re-establish the healthy microbiome of the recipient. These transplants have been attempted for a variety of different diseases. The first transplants were used to successfully treat urinary tract infections by implanting *Lactobacillus crispatus* (see Fig. 1), which promoted recolonization of the urinary tract and eventually led to the restoration of health of the recipient (Olle, 2013). Skin microbiome transplants have also shown positive effect for the treatment of atopic dermatitis (Myles *et al.*, 2016; Nakatsuji *et al.*, 2017). Faecal microbiome transplants have proven to be useful for the treatment of *Clostridium difficile* colitis (Gupta *et al.*, 2016) and various companies are currently developing probiotics as therapy for *C. diff* colitis. Preliminary results have also suggested potential use of faecal transplants for the treatment of inflammatory bowel disease, obesity, metabolic syndrome and functional gastrointestinal disorders (Thomaz *et al.*, 2016). There have been a number of attempts to define methods and apparatuses for faecal transplant, but robust and reliable methods are still under development (Gupta *et al.*, 2016; Kumar *et al.*, 2017). One of the main limitations of faecal microbiome transplants has been defining standards for effective microbiome profiles and appropriate donors.

There have been a number of other applications targeting the microbiome to improve human health, prolong wellness and for diagnostic medicine. Two recent microbiome studies have made progress towards drafting a generalized human diet in the context of ageing and human longevity (Heintz and Mair, 2014; Shoaie *et al.*, 2015). Additionally, multi-omic tools applied to the human microbiome have shown promise towards the discovery of biomarkers, such as saliva-based biomarkers useful for the detection of oral cancer, Crohn's disease, pancreatic cancer, chronic pancreatitis, periodontal disease, dental caries and obesity (Yoshizawa *et al.*, 2013).

Microbiomes in forensic analysis

Multi-omics data of deceased subjects have resulted in defining the thanatomicrobiome (Javan *et al.*, 2016a), in which insights into post-mortem microbial colonization (putrefaction) of external and internal parts of the human body establish the basis for forensic applications. Recently, studies have proposed that multi-omic microbiome tools may be a valuable resource to generate markers for deciphering human ante-mortem and post-mortem events (Javan *et al.*, 2016a,b). A longitudinal study performed on 12 female and 15 male cadavers demonstrated significant differences in microbiome colonization by organ (brain, buccal cavity, heart, liver and spleen) over the course of 240 h after death. The determination of microbiome profiles in each organ as putrefaction progresses lays the foundations for estimating post-mortem interval times (Javan *et al.*, 2016b). In addition, amplicon sequencing, shotgun metagenomics and metabolomics were proposed as tools for the contextualization of ante-mortem evidence (Metcalf *et al.*, 2017). The hope is that the origin of soil samples collected from cadavers could be geographically located using for example the Earth Microbiome Project database (Gilbert *et al.*, 2010). Also, the skin microbiome is highly diverse and can be perturbed by environmental conditions or by being in contact with objects before and after death. It has been hypothesized that skin microbiome features can provide clues of cohabitation patterns, lifestyle habits, medication and daily ante-mortem routines of the subject (Metcalf *et al.*, 2017). Standardized methods and protocols of the thanatomicrobiome are currently under development, and these multi-omic data studies could become a reliable tool in forensic analysis.

Microbiome methods for agriculture and livestock production

The microbiomes of plants and animals have recently become a primary interest for the agricultural industry to improve production. The microbiome of the rhizosphere, which is the narrow region of soil that is directly influenced by root secretions and the associated microorganisms, as well as different livestock animals, have been studied to determine inroads for improvement (Brulc *et al.*, 2009; Berendsen *et al.*, 2012; Isaacson and Kim, 2012). Plant-associated microbiomes vary substantially depending on species of plant, the location on the plant, e.g. leaves, branches or roots (Wagner *et al.*, 2016), as well as on the soil and its condition. Manipulating plant microbiomes has become a promising strategy for

agricultural plant growth management through increasing nutrient-use efficiency, abiotic stress tolerance and disease resistance in the short term (Gopal *et al.*, 2013; Busby *et al.*, 2017). Data-driven analysis has been focused on empowering the growth of plants as an alternative to reduce the use of synthetic fertilizers and pesticides. Increased production yields have been obtained by coating seeds with beneficial microbes (O'Callaghan, 2016). In addition, plant-probiotics supplementation into the soil has been proven to enhance the production of crops like soybean, wheat and corn (Rascovan *et al.*, 2016). Uncovering the nature and significance of microbial–plant interactions remains a serious effort for the field, and currently, plant–microbiome experimental studies greatly outnumber modelling studies of plants, due to the inherent complexity of plant genomes and the resulting reconstruction efforts. However, model-driven multi-omic analysis of plant microbiomes has the potential to improve agricultural productivity, especially in the face of human population growth and climate change (Andreote and Pereira e Silva, 2017).

Analogous to plant improvement efforts, the demand for efficiency continues to increase also for the livestock sector (Council, 2015). While health is a determining factor for livestock as well, efforts for improvement are targeting performance (e.g. weight and growth) as well as cost/benefit. Modelling and data-driven tools applied to livestock will likely contribute to the understanding and control of breeding, nutrition and animal health, in order to reach sufficient and improved, as well as biosustainable production. Multi-omic data analysis has provided insights into beneficial effects of probiotics on the chicken gut microbiome. The study showed that chickens fed with a selenium-yeast probiotic were more resistant to the colonization of the pathogen *Campylobacter jejuni* (Thibodeau *et al.*, 2017). Understanding the role of the microbiome in livestock remains a growing area of research.

Future directions and conclusions

The integration of experimental and computational approaches is a key driver in obtaining in-depth knowledge of microbial communities and learning about assembly, and maintenance of communities in order to predict how microbomes will react to specific perturbations. Generation of longitudinal data has shown to be highly informative to unravel community dynamics. Incorporation of large metadata will further improve our knowledge about what members are playing what roles in complex associations. Moving forward, it is necessary to update and advance algorithm and software development in order to balance experimental data production with fast and comprehensive data analysis tools. These

improvements require interdisciplinary research efforts encompassing different disciplines, such as microbiology, medicine, biotechnology, statistics, computer science and systems biology (Bersanelli *et al.*, 2016; Meng *et al.*, 2016).

The construction of models to characterize communities in response to stimuli has been initiated and has already yielded new insight into microbial interactions. Various models have been deployed to understand and manipulate microbial communities. In order to improve the accuracy of predictive community modelling, methods need to be developed to better parameterize kinetic models, which are often hampered by a lack of accurate measurements necessary to describe the entangled nonlinear biological processes. In the case of constraint-based modelling, it is important to improve on automated reconstruction methods, as well as to conduct well-designed experiments to determine exchange capabilities, as well as parameterize and validate these models. The experimental design should account for controls capturing as many details as possible, e.g. measurements of free compounds in the culture medium over time and thus delineate their origin (i.e. by secretion or lysis). Even though the genome provides evidence about transport capabilities, physiological and omic data can help to reveal activity of these transporters.

Furthermore, greater strides need to be made in determining and predicting the effect of spatial organization of cells within the community, effect of temperature, pH, osmolarity and ionic strength. The influence of these parameters is often not fully understood even for pure cultures, and as a result, their effects are rarely included in community models.

Future analysis of host–microbe interactions should be based on robust and reliable mathematical tools that can accurately elucidate different types of interactions and their underlying mechanisms. The first steps have been taken by adapting Koch's postulates using the latest experimental techniques to guide and evaluate the cause of multimicrobe diseases (see Byrd and Segre, 2016 and highlighted work therein). It is also possible to better predict drug side effects if host and microbiome are considered as actively interacting entities. Another application where knowledge about microbial interactions becomes crucial is the rational design of synthetic communities to be used in medicine and agriculture. Advances in the integration of multi-omics data to mathematical modelling will provide unprecedented insight into microbial communities and will further advance tools on how to manipulate them.

Conflict of interest

The authors declare no conflict of interests.

References

Abraham, N.M., Liu, L., Jutras, B.L., Yadav, A.K., Narasimhan, S., Gopalakrishnan, V., *et al.* (2017) Pathogen-mediated manipulation of arthropod microbiota to promote infection. *Proc Natl Acad Sci USA* **114:** E781–E790.

Abubucker, S., Segata, N., Goll, J., Schubert, A.M., Izard, J., Cantarel, B.L., *et al.* (2012) Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* **8:** e1002358.

Alhumazi, K., and Ajbar, A. (2005) Dynamics of predator-prey interactions in continuous cultures. *Eng Life Sci* **5:** 139–147.

Allison, S.M., Small, J.R., Kacser, H., and Prosser, J.I. (1993) Control analysis of microbial interactions in continuous culture: a simulation study. *J Gen Microbiol* **139:** 2309–2317.

Andreote, F.D. and Pereira e Silva, M.C. (2017) Microbial communities associated with plants: learning from nature to apply it in agriculture. *Curr Opin Microbiol* **37:** 29–34.

Backhed, F., Ley, R., Sonnenburg, J., Peterson, D. and Gordon, J. (2005) Host-bacterial mutualism in the human intestine. *Science* **307:** 1915–1920.

Bartell, J.A., Blazier, A.S., Yen, P., Thøgersen, J.C., Jelsbak, L., Goldberg, J.B., and Papin, J.A. (2017) Reconstruction of the metabolic network of *Pseudomonas aeruginosa* to interrogate virulence factor synthesis. *Nat Commun* **8:** 14631.

Beijerinck, M.W. (1895) Über Spirillum desulfuricans als ursache von sulfatreduktion. *Zentralblatt Bakteriol* **1:** 1–9.

Beneden, M. (1876) Mutualists. In *Animal Parasites and Messmates*. HS King & Company (eds). New York, NY Second edition: pp. 1–68.

Berendsen, R.L., Pieterse, C.M.J., and Bakker, P.A.H.M. (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478–486.

Berleman, J., and Auer, M. (2013) The role of bacterial outer membrane vesicles for intra- and interspecies delivery. *Environ Microbiol* **15:** 347–354.

Bersanelli, M., Mosca, E., Remondini, D., Giampieri, E., Sala, C., Castellani, G., and Milanesi, L. (2016) Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinformatics* **17:** S15.

Biggs, M.B., Medlock, G.L., Kolling, G.L., and Papin, J.A. (2015) Metabolic network modeling of microbial communities. *Wiley Interdisc Rev Syst Biol Med* **7:** 317–334.

Blurock, E.S. (2003) Characterizing complex reaction mechanisms using machine learning clustering techniques. *Int J Chem Kinet* **36:** 107–118.

Boguniewicz, M., and Leung, D.Y.M. (2011) Atopic dermatitis: a disease of altered skin barrier and immune dysregulation. *Immunol Rev* **242:** 233–246.

Bordbar, A., Lewis, N.E., Schellenberger, J., Palsson, B.Ø. and Jamshidi, N. (2010) Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol* **6:** 422–435.

Bordbar, A., Monk, J.M., King, Z.A., and Palsson, B.O. (2014) Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet* **15:** 107–120.

Bordbar, A., Yurkovich, J.T., Paglia, G., Rolfsson, O., and Palsson, B.O. (2017) Elucidating dynamic metabolic physiology through network integration of quantitative time-course metabolomics. *Sci Rep* **7:** 46249.

Bosi, E., Monk, J.M., Aziz, R.K., Fondi, M., Nizet, V., and Palsson, B.Ø. (2016) Comparative genome-scale modelling of *Staphylococcus aureus* strains identifies strain-specific metabolic capabilities linked to pathogenicity. *Proc Natl Acad Sci USA* **113:** E3801–E3809.

Brooks, J.P., Edwards, D.J., Harwich, M.D., Rivera, M.C., Fettweis, J.M., Serrano, M.G., *et al.* (2015) The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* **15:** 66.

Brown, C.T., Olm, M.R., Thomas, B.C., and Banfield, J.F. (2016) Measurement of bacterial replication rates in microbial communities. *Nat Biotechnol* **34:** 1256–1263.

Brulc, J.M., Antonopoulos, D.A., Berg Miller, M.E., Wilson, M.K., Yannarell, A.C., Dinsdale, E.A., *et al.* (2009) Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc Natl Acad Sci USA* **106:** 1948–1953.

Busby, P.E., Soman, C., Wagner, M.R., Friesen, M.L., Kremer, J., Bennett, A., *et al.* (2017) Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biol* **15:** e2001793.

Byrd, A.L. and Segre, J.A. (2016) Adapting Koch's postulates. *Science* **351:** 224–226.

Cabello, J., Morales, M., and Revah, S. (2014) Dynamic photosynthetic response of the microalga *Scenedesmus obtusiusculus* to light intensity perturbations. *Chem Eng J* **252:** 104–111.

Cadwell, K. (2015) The virome in host health and disease. *Immunity* **42:** 805–813.

Canale, R.P. (1970) An analysis of models describing predator-prey interaction. *Biotechnol Bioeng* **12:** 353–378.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7:** 335–336.

Carr, R., and Borenstein, E. (2014) Comparative analysis of functional metagenomic annotation and the mappability of short reads. *PLoS One* **9:** e105776.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., *et al.* (2014) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* **42:** 459–471.

Chaparro, J.M., Badri, D.V., and Vivanco, J.M. (2014) Rhizosphere microbiome assemblage is affected by plant development. *ISME J* 8: 790–803.

Chen, H., Athar, R., Zheng, G., and Williams, H.N. (2011) Prey bacteria shape the community structure of their predators. *ISME J* 5: 1314–1322.

Cohn, F. (1875) Untersuchungen über bakterien. *Beitr Biol Pflanzen* 1: 141–207.

Cortassa, S., O'Rourke, B., Winslow, R.L., and Aon, M.A. (2009) Control and regulation of mitochondrial energetics in an integrated model of cardiomyocyte function. *Biophys J* 96: 2466–2478.

Council, N.R. (2015) *Critical Role of Animal Science Research in Food Security and Sustainability*. Washington, DC: National Academies Press.

Cremer, J., Segota, I., Yang, C., Arnoldini, M., Sauls, J.T., Zhang, Z., et al. (2016) Effect of flow and peristaltic mixing on bacterial growth in a gut-like channel. *Proc Natl Acad Sci USA* 113: 11414–11419.

Degnan, P.H., Taga, M.E., and Goodman, A.L. (2014) Vitamin B12 as a modulator of gut microbial ecology. *Cell Metab* 20: 769–778.

Delong, E.F. (1992) Archaea in coastal marine environments. *Proc Natl Acad Sci USA* 89: 5685–5689.

Dhurjati, P., Ramkrishna, D., Flickinger, M.C., and Tsao, G.T. (1985) A cybernetic view of microbial growth: modeling of cells as optimal strategists. *Biotechnol Bioeng* 27: 1–9.

Dobell, C. (1923) A protozoological bicentenary: Antony van Leeuwenhoek (1632–1723) and Louis Joblot (1645–1723). *Parasitology* 15: 308.

Du, B., Zielinski, D.C., Kavvas, E.S., Dräger, A., Tan, J., Zhang, Z., et al. (2016) Evaluation of rate law approximations in bottom-up kinetic models of metabolism. *BMC Syst Biol* 10: 40.

Dunn, I.J., Heinzle, E., Ingham, J., Pfenosil, J.E., (2003) Biological reaction engineering. Principles, applications and modeling with PC simulation. Wiley-VCH GmbH & Co. KGaA. New York: Second, Completely Revised Edition, pp. 400–445.

Dutilh, B.E., Backus, L., Van Hijum, S.A.F.T., and Tjalsma, H. (2013) Screening metatranscriptomes for toxin genes as functional drivers of human colorectal cancer. *Best Pract Res Clin Gastroenterol* 27: 85–99.

Edwards, J.S., and Palsson, B.Ø. (2000) The *Escherichia coli* MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities. *Proc Natl Acad Sci USA* 97: 5528–5533.

Embree, M., Nagarajan, H., Movahedi, N., Chitsaz, H., and Zengler, K. (2014) Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. *ISME J* 8: 757–767.

Embree, M., Liu, J.K., Al-Bassam, M.M., and Zengler, K. (2015) Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc Natl Acad Sci USA* 112: 15450–15455.

Erickson, A.R., Cantarel, B.L., Lamendella, R., Darzi, Y., Mongodin, E.F., Pan, C., et al. (2012) Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS One* 7: e49138.

Faust, K., and Raes, J. (2012) Microbial interactions: from networks to models. *Nat Rev Microbiol* 10: 538–550.

Fischbach, M.A., and Segre, J.A. (2016) Signaling in host-associated microbial communities. *Cell* 164: 1288–1300.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269: 496–512.

Flint, H.J., Scott, K.P., Louis, P., and Duncan, S.H. (2012) The role of the gut microbiota in nutrition and health. *Nat Rev Gastroenterol Hepatol* 9: 577–589.

Fouts, D.E., Szpakowski, S., Purushe, J., Torralba, M., Waterman, R.C., MacNeil, M.D., et al. (2012) Next generation sequencing to define prokaryotic and fungal diversity in the bovine rumen. *PLoS One* 7: e48289.

Franzosa, E.A., Morgan, X.C., Segata, N., Waldron, L., Reyes, J., Earl, A.M., et al. (2014) Relating the metatranscriptome and metagenome of the human gut. *Proc Natl Acad Sci USA* 111: E2329–E2338.

Franzosa, E.A., Hsu, T., Sirota-Madi, A., Shafquat, A., Abu-Ali, G., Morgan, X.C., and Huttenhower, C. (2015) Sequencing and beyond: integrating molecular "omics" for microbial community profiling. *Nat Rev Microbiol* 13: 360–372.

Freilich, S., Zarecki, R., Eilam, O., Segal, E.S., Henry, C.S., Kupiec, M., et al. (2011) Competitive and cooperative metabolic interactions in bacterial communities. *Nat Commun* 2: 589.

Fuhrman, J.A., McCallum, K., and Davis, A.A. (1992) Novel major archaebacterial group from marine plankton. *Nature* 356: 148–149.

Gause, G.F. (1934) The struggle for existence. The Williams & Wilkins company 1934, (ed.). Baltimore, pp. 163–170.

Geng, J., Song, H.-S., Yuan, J., and Ramkrishna, D. (2012) On enhancing productivity of bioethanol with multiple species. *Biotechnol Bioeng* 109: 1508–1517.

Germerodt, S., Bohl, K., Lück, A., Pande, S., Schröter, A., Kaleta, C., et al. (2016) Pervasive selection for cooperative cross-feeding in bacterial communities. *PLoS Comput Biol* 12: e1004986.

Gibson, G.R., Hutkins, R., Sanders, M.E., Prescott, S.L., Reimer, R.A., Salminen, S.J., et al. (2017) Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat Rev Gastroenterol Hepatol* 14: 491–502.

Gilbert, J.A., Meyer, F., Antonopoulos, D., Balaji, P., Brown, C.T., Brown, C.T., et al. (2010) Meeting report: the terabase metagenomics workshop and the vision of an earth microbiome project. *Stand Genomic Sci* 3: 243–248.

Giovannoni, S.J., DeLong, E.F., Schmidt, T.M., and Pace, N.R. (1990) Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl Environ Microbiol* 56: 2572–2575.

Gonzalez, D.J., Haste, N.M., Hollands, A., Fleming, T.C., Hamby, M., Pogliano, K., et al. (2011) Microbial competition between *Bacillus subtilis* and *Staphylococcus aureus* monitored by imaging mass spectrometry. *Microbiology* 157: 2485–2492.

Gopal, M., Gupta, A. and Thomas, G.V. (2013) Bespoke microbiome therapy to manage plant diseases. *Front Microbiol* 4: 355–359.

Gowda, H., Ivanisevic, J., Johnson, C.H., Kurczy, M.E., Benton, H.P., Rinehart, D., *et al.* (2014) Interactive XCMS online: simplifying advanced metabolomic data processing and subsequent statistical analyses. *Anal Chem* **86:** 6931–6939.

Greenblum, S., Turnbaugh, P.J., and Borenstein, E. (2012) Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc Natl Acad Sci USA* **109:** 594–599.

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., *et al.* (2016) Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532:** 465–470.

Guimaraes, J.C., Rocha, M., and Arkin, A.P. (2014) Transcript level and sequence determinants of protein abundance and noise in *Escherichia coli. Nucleic Acids Res* **42:** 4791–4799.

Gupta, S., Allen-Vercoe, E., and Petrof, E.O. (2016) Fecal microbiota transplantation: in perspective. *Therap Adv Gastroenterol* **9:** 229–239.

Halfvarson, J., Brislawn, C.J., Lamendella, R., Vázquez-Baeza, Y., Walters, W.A., Bramer, L.M., *et al.* (2017) Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* **2:** 17004.

Hamilton, J.J., Calixto Contreras, M., and Reed, J.L. (2015) Thermodynamics and H_2 transfer in a methanogenic, syntrophic community. *PLoS Comput Biol* **11:** e1004364.

Heinken, A., and Thiele, I. (2015) Systems biology of host-microbe metabolomics. *Wiley Interdiscip Rev Syst Biol Med* **7:** 195–219.

Heintz, C., and Mair, W. (2014) You are what you host: microbiome modulation of the aging process. *Cell* **156:** 408–411.

Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M., Linsay, B., and Stevens, R.L. (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* **28:** 977–982.

Hibberd, M.C., Wu, M., Rodionov, D.A., Li, X., Cheng, J., Griffin, N.W., *et al.* (2017) The effects of micronutrient deficiencies on bacterial species from the human gut microbiota. *Sci Transl Med* **9:** eaal4069.

Hocher, B., and Adamski, J. (2017) Metabolomics for clinical use and research in chronic kidney disease. *Nat Rev Nephrol* **13:** 269–284.

Holz, C., Benning, J., Schaudt, M., Heilmann, A., Schultchen, J., Goelling, D., and Lang, C. (2017) Novel bioactive from *Lactobacillus brevis* DSM17250 to stimulate the growth of *Staphylococcus epidermidis*: a pilot study. *Benef Microbes* **8:** 121–131.

Huan, T., Forsberg, E.M., Rinehart, D., Johnson, C.H., Ivanisevic, J., Benton, H.P., *et al.* (2017) Systems biology guided by XCMS Online metabolomics. *Nat Methods* **14:** 461–462.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., *et al.* (2016) A new view of the tree of life. *Nat Microbiol* **1:** 16048.

Huston, M., and Smith, T. (1987) Plant succession: life history and competition. *Am Nat* **130:** 168–198.

Ingolia, N.T. (2016) Ribosome footprint profiling of translation throughout the genome. *Cell* **165:** 22–33.

Isaacson, R., and Kim, H.B. (2012) The intestinal microbiome of the pig. *Anim Health Res Rev* **13:** 100–109.

Islam, M.A., Waller, A.S., Hug, L.A., Provart, N.J., Edwards, E.A., and Mahadevan, R. (2014) New insights into *Dehalococcoides mccartyi* metabolism from a reconstructed metabolic network-based systems-level analysis of *D. mccartyi* transcriptomes. *PLoS One* **9:** e94808.

Ivey, M., Massel, M., and Phister, T.G. (2013) Microbial interactions in food fermentations. *Annu Rev Food Sci Technol* **4:** 141–162.

Jansson, J.K., and Baker, E.S. (2016) A multi-omic future for microbiome studies. *Nat Microbiol* **1:** 16049.

Javan, G.T., Finley, S.J., Abidin, Z. and Mulle, J.G. (2016a) The thanatomicrobiome: a missing piece of the microbial puzzle of death. *Front Microbiol* **7:** 225–232.

Javan, G.T., Finley, S.J., Can, I., Wilkinson, J.E., Hanson, J.D., and Tarone, A.M. (2016b) Human thanatomicrobiome succession and time since death. *Sci Rep* **6:** 29598.

Ji, B. and Nielsen, J. (2015) From next-generation sequencing to systematic modeling of the gut microbiome. *Front Genet* **6:** 219–228.

Jorth, P., Turner, K.H., Gumus, P., Nizam, N., Buduneli, N. and Whiteley, M. (2014) Metatranscriptomics of the human oral microbiome during health and disease. *MBio* **5:** e01012–e01014.

von Kamp, A., and Klamt, S. (2017) Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nat Commun* **8:** 15956.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42:** 199–205.

Kang, D., Shi, B., Erfe, M.C., Craft, N. and Li, H. (2015) Vitamin B12 modulates the transcriptome of the skin microbiota in acne pathogenesis. *Sci Transl Med* **7:** 293ra103.

Karr, J.R., Sanghvi, J.C., Macklin, D.N., Gutschow, M.V., Jacobs, J.M., Bolival, B., *et al.* (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* **150:** 389–401.

Kayala, M.A., Azencott, C.-A., Chen, J.H., and Baldi, P. (2011) Learning to predict chemical reactions. *J Chem Inf Model* **51:** 2209–2222.

Kerkhoven, E.J., Pomraning, K.R., Baker, S.E. and Nielsen, J. (2016) Regulation of amino-acid metabolism controls flux to lipid accumulation in *Yarrowia lipolytica. NPJ Syst Biol Appl* **2:** 16005.

King, Z.A., Dräger, A., Ebrahim, A., Sonnenschein, N., Lewis, N.E., and Palsson, B.O. (2015) Escher: a web application for building, sharing, and embedding data-rich visualizations of biological pathways. *PLoS Comput Biol* **11:** e1004321.

Klemashevich, C., Wu, C., Howsmon, D., Alaniz, R.C., Lee, K., and Jayaraman, A. (2014) Rational identification of diet-derived postbiotics for improving intestinal microbiota function. *Curr Opin Biotechnol* **26:** 85–90.

Knight, R., Callewaert, C., Marotz, C., Hyde, E.R., Debelius, J.W., McDonald, D. and Sogin, M.L. (2017) The microbiome and human biology. *Annu Rev Genomics Hum Genet* **18:** 1–22.

Koleva, P., Bridgman, S., and Kozyrskyj, A. (2015) The infant gut microbiome: evidence for obesity risk and dietary intervention. *Nutrients* **7:** 2237–2260.

Kumar, R., Yi, N., Zhi, D., Eipers, P., Goldsmith, K.T., Dixon, P., *et al.* (2017) Identification of donor microbe species that colonize and persist long term in the recipient after fecal transplant for recurrent *Clostridium difficile. NPJ Biofilms Microbiomes* **3:** 12.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.

Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L., and Pace, N.R. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci USA* **82:** 6955–6959.

Latif, H., Szubin, R., Tan, J., Brunk, E., Lechner, A., Zengler, K. and Palsson, B.O. (2015) A streamlined ribosome profiling protocol for the characterization of microorganisms. *Biotechniques* **58:** 329–332.

Lerman, J.A., Hyduke, D.R., Latif, H., Portnoy, V.A., Lewis, N.E., Orth, J.D., *et al.* (2012) In silico method for modelling metabolism and gene product expression at genome scale. *Nat Commun* **3:** 929.

Letten, A.D., Ke, P.-J. and Fukami, T. (2017) Linking modern coexistence theory and contemporary niche theory. *Ecol Monogr* **87:** 161–177.

Levering, J., Broddrick, J., Dupont, C.L., Peers, G., Beeri, K., Mayers, J., *et al.* (2016) Genome-scale model reveals metabolic basis of biomass partitioning in a model diatom. *PLoS One* **11:** 1–22.

Levy, R., and Borenstein, E. (2013) Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc Natl Acad Sci USA* **110:** 12804–12809.

Lewin, A., Johansen, J., Wentzel, A., Kotlar, H.K., Drabløs, F., and Valla, S. (2014) The microbial communities in two apparently physically separated deep subsurface oil reservoirs show extensive DNA sequence similarities. *Environ Microbiol* **16:** 545–558.

LeWitt, P.A., Li, J., Lu, M., Guo, L., and Auinger, P. (2017) Metabolomic biomarkers as strong correlates of Parkinson disease progression. *Neurology* **88:** 862–869.

Liu, J.K., O'Brien, E.J., Lerman, J.A., Zengler, K., Palsson, B.O., and Feist, A.M. (2014) Reconstruction and modeling protein translocation and compartmentalization in *Escherichia coli* at the genome-scale. *BMC Syst Biol* **8:** 110.

Lotka, A.J. (1925) Elements of physical biology reissued as elements of mathematical biology. Williams & Wilkins Company (Eds) Baltimore: MD, pp. 161-184.

Louca, S., Parfrey, L.W. and Doebeli, M. (2016) Decoupling function and taxonomy in the global ocean microbiome. *Science* **353:** 1272–1277.

Lyon, M., Wood, S., Pelletier, X., Donazzolo, Y., Gahler, R., and Bellisle, F. (2011) Effects of a 3-month supplementation with a novel soluble highly viscous polysaccharide on anthropometry and blood lipids in nondieting overweight or obese adults. *J Hum Nutr Diet* **24:** 351–359.

Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D.A., Bauer, E., Noronha, A., *et al.* (2016) Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol* **35:** 81–89.

Mahowald, M.A., Rey, F.E., Seedorf, H., Turnbaugh, P.J., Fulton, R.S., Wollam, A., *et al.* (2009) Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. *Proc Natl Acad Sci USA* **106:** 5859–5864.

Maier, T., Güell, M., and Serrano, L. (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Lett* **583:** 3966–3973.

Marco, M.B., Moineau, S., and Quiberoni, A. (2012) Bacteriophages and dairy fermentations. *Bacteriophage* **2:** 149–158.

Mardinoglu, A., Shoaie, S., Bergentall, M., Ghaffari, P., Zhang, C., Larsson, E., *et al.* (2015) The gut microbiota modulates host amino acid and glutathione metabolism in mice. *Mol Syst Biol* **11:** 834.

Martinez-Garcia, M., Swan, B.K., Poulton, N.J., Gomez, M.L., Masland, D., Sieracki, M.E., and Stepanauskas, R. (2012) High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *ISME J* **6:** 113–123.

Maynard, N.D., Gutschow, M.V., Birch, E.W., and Covert, M.W. (2010) The virus as metabolic engineer. *Biotechnol J* **5:** 686–694.

McCalley, C.K. and Sparks, J.P. (2009) Abiotic gas formation drives nitrogen loss from a desert ecosystem. *Science* **326:** 837–840.

McDonald, D., Birmingham, A., and Knight, R. (2015) Context and the human microbiome. *Microbiome* **3:** 52.

McDonald, D., Glusman, G., and Price, N.D. (2016) Personalized nutrition through big data. *Nat Biotechnol* **34:** 152–154.

McGibbon, R.T., and Pande, V.S. (2013) Learning kinetic distance metrics for markov state models of protein conformational dynamics. *J Chem Theory Comput* **9:** 2900–2906.

Mee, M.T., Collins, J.J., Church, G.M., and Wang, H.H. (2014) Syntrophic exchange in synthetic microbial communities. *Proc Natl Acad Sci USA* **111:** E2149–E2156.

Meeuwse, P., Akbari, P., Tramper, J., and Rinzema, A. (2012) Modeling growth, lipid accumulation and lipid turnover in submerged batch cultures of *Umbelopsis isabellina. Bioprocess Biosyst Eng* **35:** 591–603.

Meng, C., Zeleznik, O.A., Thallinger, G.G., Kuster, B., Gholami, A.M., and Culhane, A.C. (2016) Dimension reduction techniques for the integrative analysis of multi-omics data. *Brief Bioinform* **17:** 628–641.

Metcalf, J.L., Xu, Z.Z., Bouslimani, A., Dorrestein, P., Carter, D.O. and Knight, R. (2017) Microbiome tools for forensic science. *Trends Biotechnol* **35:** 814–823.

Methé, B.A., Nelson, K.E., Pop, M., Creasy, H.H., Giglio, M.G., Huttenhower, C., *et al.* (2012) A framework for human microbiome research. *Nature* **486:** 215–221.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E., Kubal, M., *et al.* (2008) The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9:** 386.

Monk, J.M., Charusanti, P., Aziz, R.K., Lerman, J.A., Premyodhin, N., Orth, J.D., *et al.* (2013) Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional

environments. *Proc Natl Acad Sci USA* **110:** 20338–20343.

Morris, B.E.L., Henneberger, R., Huber, H., and Moissl-Eichinger, C. (2013) Microbial syntrophy: interaction for the common good. *FEMS Microbiol Rev* **37:** 384–406.

Myles, I.A., Williams, K.W., Reckhow, J.D., Jammeh, M.L., Pincus, N.B., Sastalla, I., *et al.* (2016) Transplantation of human skin microbiota in models of atopic dermatitis. *JCI Insight* **1:** 1–10.

Nagarajan, H., Embree, M., Rotaru, A.-E., Shrestha, P.M., Feist, A.M., Palsson, B.Ø., *et al.* (2013) Characterization and modelling of interspecies electron transfer mechanisms and microbial community dynamics of a syntrophic association. *Nat Commun* **4:** 1–10.

Nakatsuji, T., Chen, T.H., Narala, S., Chun, K.A., Two, A.M., Yun, T., *et al.* (2017) Antimicrobials from human skin commensal bacteria protect against *Staphylococcus aureus* and are deficient in atopic dermatitis. *Sci Transl Med* **9:** 1–12.

Nature-Biotechnology (2017) Recent patents in microbiomes and microbial genomics. *Nat Biotechnol* **35:** 517.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., *et al.* (2013) Patterns and processes of microbial community assembly. *Microbiol Mol Biol Rev* **77:** 342–356.

Ng, C., DeMaere, M.Z., Williams, T.J., Lauro, F.M., Raftery, M., Gibson, J.A.E., *et al.* (2010) Metaproteogenomic analysis of a dominant green sulfur bacterium from Ace Lake, Antarctica. *ISME J* **4:** 1002–1019.

Nielsen, J., and Olson, L. (2002) An expanded role for microbial physiology in metabolic engineering and functional genomics: moving towards systems biology. *FEMS Yeast Res* **2:** 175–181.

Nothias, L.-F., Knight, R., and Dorrestein, P.C. (2016) Antibiotic discovery is a walk in the park. *Proc Natl Acad Sci USA* **113:** 201618221.

O'Brien, E.J., Lerman, J.A., Chang, R.L., Hyduke, D.R., and Palsson, B.O. (2014) Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol Syst Biol* **9:** 693.

O'Callaghan, M. (2016) Microbial inoculation of seed for improved crop performance: issues and opportunities. *Appl Microbiol Biotechnol* **100:** 5729–5746.

Olle, B. (2013) Medicines from microbiota. *Nat Biotechnol* **31:** 309–315.

Paddon, C.J., and Keasling, J.D. (2014) Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol* **12:** 355–367.

Palsson, B., and Zengler, K. (2010) The challenges of integrating multi-omic data sets. *Nat Chem Biol* **6:** 787–789.

Perez-Garcia, O., Lear, G. and Singhal, N. (2016) Metabolic network modeling of microbial interactions in natural and engineered environmental systems. *Front Microbiol* **7:** 673–703.

Petschow, B., Dore, J., Hibberd, P., Dinan, T., Reid, G., Blaser, M., *et al.* (2013) Probiotics, prebiotics, and the host microbiome: the science of translation. *Ann N Y Acad Sci* **1306:** 1–17.

Plichta, D.R., Juncker, A.S., Bertalan, M., Rettedal, E., Gautier, L., Varela, E., *et al.* (2016) Transcriptional

interactions suggest niche segregation among microorganisms in the human gut. *Nat Microbiol* **1:** 16152.

Ponomarova, O., and Patil, K.R. (2015) Metabolic interactions in microbial communities: untangling the Gordian knot. *Curr Opin Microbiol* **27:** 37–44.

Rascovan, N., Carbonetto, B., Perrig, D., Díaz, M., Canciani, W., Abalo, M., *et al.* (2016) Integrated analysis of root microbiomes of soybean and wheat from agricultural fields. *Sci Rep* **6:** 28084.

Resat, H., Costa, M.N., and Shankaran, H. (2011) Spatial aspects in biological system simulations. *Methods Enzymol* **487:** 485–511.

Roling, W.F.M. and van Bodegom, P.M. (2014) Toward quantitative understanding on microbial community structure and functioning: a modeling-centered approach using degradation of marine oil spills as example. *Front Microbiol* **5:** 125–137.

Röling, W.F., Ferrer, M., and Golyshin, P.N. (2010) Systems approaches to microbial communities and their functioning. *Curr Opin Biotechnol* **21:** 532–538.

Ryu, J.Y., Kim, H.U., and Lee, S.Y. (2015) Reconstruction of genome-scale human metabolic models using omics data. *Integr Biol* **7:** 859–868.

Sangwan, N., Xia, F., and Gilbert, J.A. (2016) Recovering complete and draft population genomes from metagenome datasets. *Microbiome* **4:** 8.

Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J.D., Feist, A.M., *et al.* (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc* **6:** 1290–1307.

Schultz, A., and Qutub, A.A. (2016) Reconstruction of tissue-specific metabolic networks using CORDA. *PLoS Comput Biol* **12:** e1004808.

Sekirov, I., Russell, S.L., Antunes, L.C.M., and Finlay, B.B. (2010) Gut microbiota in health and disease. *Physiol Rev* **90:** 859–904.

Shoaie, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S. and Nielsen, J. (2013) Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci Rep* **3:** 1–10.

Shoaie, S., Ghaffari, P., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot, E., *et al.* (2015) Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell Metab* **22:** 320–331.

Sierra-Garcia, I.N., Dellagnezze, B.M., Santos, V.P., Chaves B, M.R., Capilla, R., Santos Neto, E.V., *et al.* (2016) Microbial diversity in degraded and non-degraded petroleum samples and comparison across oil reservoirs at local and global scales. *Extremophiles* **21:** 211–229.

Silva, R.R., Jourdan, F., Salvanha, D.M., Letisse, F., Jamin, E.L., Guidetti-Gonzalez, S., *et al.* (2014) ProbMetab: an R package for Bayesian probabilistic annotation of LC-MS-based metabolomics. *Bioinformatics* **30:** 1336–1337.

Singer, E., Wagner, M. and Woyke, T. (2017) Capturing the genetic makeup of the active microbiome in situ. *ISME J* **11:** 1949–1963.

Steffani-Vallejo, J.L., Zuniga, C., Cruz-Morales, P., Lozano, L., Morales, M., Licona-Cassani, C., *et al.* (2017) Draft genome sequence of *Sphingobacterium* sp. CZ-UAM, isolated from a methanotrophic consortium. *Genome Announc* **5:** 1–2.

Stolyar, S., Van Dien, S., Hillesland, K.L., Pinel, N., Lie, T.J., Leigh, J.A. and Stahl, D.A. (2007) Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* **3:** 1–14.

Swainston, N., Smallbone, K., Hefzi, H., Dobson, P.D., Brewer, J., Hanscho, M., *et al.* (2016) Recon 2.2: from reconstruction to model of human metabolism. *Metabolomics* **12:** 109.

Tabiś, B., Skoneczny, S. and Stryjewski, W.S. (2014) Non-linear dynamics of a controlled stirred tank bioreactor with predator-prey relationship. *Chem Process Eng* **35:** 349–360.

Tan, J., Zuniga, C., and Zengler, K. (2015) Unraveling interactions in microbial communities - from co-cultures to microbiomes. *J Microbiol* **53:** 295–305.

Thibodeau, A., Letellier, A., Yergeau, E., Larriviere-Gauthier, G. and Fravalo, P. (2017) Lack of evidence that selenium-yeast improves chicken health and modulates the caecal microbiota in the context of colonization by *Campylobacter jejuni*. *Front Microbiol* **8:** 451–460.

Thomaz, A., Fukumori, C., and Ferreira, C.M. (2016) New insights into therapeutic strategies for gut microbiota modulation in inflammatory diseases. *Clin Transl Immunology* **5:** e87.

Truong, D.T., Franzosa, E.A., Tickle, T.L., Scholz, M., Weingart, G., Pasolli, E., *et al.* (2015) MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods* **12:** 902–903.

Tsuno, H., Hidaka, T., and Nishimura, F. (2002) A simple biofilm model of bacterial competition for attached surface. *Water Res* **36:** 996–1006.

Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., and Gordon, J.I. (2007) The human microbiome project. *Nature* **449:** 804–810.

Vincent, K.M. and Postovit, L.-M. (2017) Investigating the utility of human melanoma cell lines as tumour models. *Oncotarget* **8:** 10498–10509.

Vollmers, J., Wiegand, S., and Kaster, A.-K. (2017) Comparing and evaluating metagenome assembly tools from a microbiologist's perspective - not only size matters!. *PLoS One* **12:** e0169662.

Volova, T., Kiselev, E., Vinogradova, O., Nikolaeva, E., Chistyakov, A., Sukovatiy, A., and Shishatskaya, E. (2014) A glucose-utilizing strain, *Cupriavidus euthrophus* B-10646: growth kinetics, characterization and synthesis of multicomponent PHAs. *PLoS One* **9:** e87551.

Volterra, V. (1926) Fluctuations in the abundance of a species considered mathematically. *Nature* **118:** 558–560.

Wade, M.J., Harmand, J., Benyahia, B., Bouchez, T., Chaillou, S., Cloez, B., *et al.* (2016) Perspectives in mathematical modelling for microbial ecology. *Ecol Modell* **321:** 64–74.

Wagner, M.R., Lundberg, D.S., del Rio, T.G., Tringe, S.G., Dangl, J.L., and Mitchell-Olds, T. (2016) Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nat Commun* **7:** 12151.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., *et al.* (2016) Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* **1:** e00009–e00015.

Wang, Y., and Papenguth, H.W. (2001) Kinetic modeling of microbially-driven redox chemistry of radionuclides in subsurface environments: coupling transport, microbial metabolism and geochemistry. *J Contam Hydrol* **47:** 297–309.

Wang, Z., Klipfell, E., Bennett, B.J., Koeth, R., Levison, B.S., DuGar, B., *et al.* (2011) Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472:** 57–63.

Wang, Y., Tian, R.M., Gao, Z.M., Bougouffa, S. and Qian, P.Y. (2014) Optimal eukaryotic 18S and universal 16S/18S ribosomal RNA primers and their application in a study of symbiosis. *PLoS One* **9:** 1–11.

Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., *et al.* (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **34:** 828–837.

Ward, L., Taylor, M.W., Power, J.F., Scott, B.J., McDonald, I.R. and Stott, M.B. (2017) Microbial community dynamics in Inferno Crater Lake, a thermally fluctuating geothermal spring. *ISME J* **i:** 1–10.

Wilmes, P., Heintz-Buschart, A., and Bond, P.L. (2015) A decade of metaproteomics: where we stand and what the future holds. *Proteomics* **15:** 3409–3417.

Winogradsky, S. (1887) Ueber schwefelbacterien. *Bot Zeitung* **31:** 489–507.

Xiong, W., Abraham, P.E., Li, Z., Pan, C., and Hettich, R.L. (2015) Microbial metaproteomics for characterizing the range of metabolic functions and activities of human gut microbiota. *Proteomics* **15:** 3424–3438.

Xu, A., Dolfing, J., Curtis, T.P., Montague, G., and Martin, E. (2011) Maintenance affects the stability of a two-tiered microbial "food chain"? *J Theor Biol* **276:** 35–41.

Yang, B., Wang, Y., and Qian, P.-Y. (2016) Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* **17:** 135.

Yao, W., Gao, Y., and Wan, Z. (2017) Serum metabolomics profiling to identify biomarkers for unstable angina. *Biomed Res Int* **2017:** 1–8.

Yoshizawa, J.M., Schafer, C.A., Schafer, J.J., Farrell, J.J., Paster, B.J., and Wong, D.T.W. (2013) Salivary biomarkers: toward future clinical and diagnostic utilities. *Clin Microbiol Rev* **26:** 781–791.

Young, J.C., Pan, C., Adams, R.M., Brooks, B., Banfield, J.F., Morowitz, M.J., and Hettich, R.L. (2015) Metaproteomics reveals functional shifts in microbial and human proteins during a preterm infant gut colonization case. *Proteomics* **15:** 3463–3473.

Zampieri, E., Chiapello, M., Daghino, S., Bonfante, P., and Mello, A. (2016) Soil metaproteomics reveals an inter-kingdom stress response to the presence of black truffles. *Sci Rep* **6:** 25773.

Zeigler Allen, L., McCrow, J.P., Ininbergs, K., Dupont, C.L., Badger, J.H., Hoffman, J.M., *et al.* (2017) The Baltic Sea virome: diversity and transcriptional activity of DNA and RNA viruses. *mSystems* **2:** 1–16.

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D.R., Bork, P. and Patil, K.R. (2015) Correction for Zelezniak et al., Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc Natl Acad Sci USA* **112:** E7156.

Zengler, K. (ed) (2008) Accessing uncultivated microorganisms: from the environment to organisms and genomes and back. Washington, DC: ASM Press.

Zengler, K., and Palsson, B.O. (2012) A road map for the development of community systems (CoSy) biology. *Nat Rev Microbiol* **10:** 366–372.

Zhang, Z., Scharer, J., and Moo-Young, M. (1997) Protein production using recombinant yeast in an immobilized-cell-film airlift bioreactor. *Biotechnol Bioeng* **55:** 241–251.

Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., and Lovley, D.R. (2011) Genome-scale dynamic modeling of the competition between *Rhodoferax* and *Geobacter* in anoxic subsurface environments. *ISME J* **5:** 305–316.

Zielinski, D.C., Filipp, F.V., Bordbar, A., Jensen, K., Smith, J.W., Herrgard, M.J., *et al.* (2015) Pharmacogenomic and clinical data link non-pharmacokinetic metabolic dysregulation to drug side effect pathogenesis. *Nat Commun* **6:** 7101.

Zuñiga, C., Li, C.-T., Huelsman, T., Levering, J., Zielinski, D.C., McConnell, B.O., *et al.* (2016) Genome-scale metabolic model for the green alga *Chlorella vulgaris* UTEX 395 accurately predicts phenotypes under autotrophic, heterotrophic, and mixotrophic growth conditions. *Plant Physiol* **172:** 589–602.