PUBLISHED: 26 JULY 2016 | ARTICLE NUMBER: 16101 | DOI: 10.1038/NMICROBIOL.2016.101

comment

Are multi-omics enough?

Cristina Vilanova and Manuel Porcar

Multi-omic techniques are often seen as the future of microbiome studies. We argue that recent strategies for simplifying complex omic-derived data will need to be combined with improved cultivation techniques to pave the way towards a more targeted approach for understanding microbial communities.

In 1665, Robert Hooke was the first human to observe a microorganism. Two centuries elapsed between these first sightings of microbial cells and the isolation of bacterial species in pure culture with artificial media. The identification and study of microbial species relied then on their cultivation and phenotypic characterization for centuries, even though it was soon discovered that only ~1% of the microorganisms present in environmental samples were cultivable[1]. In 2002, the birth of metagenomic sequencing facilitated an explosion of microbiome studies, enabling much of a microbial community to be identified in a single experiment. A few years later, other omic technologies arose (metatranscriptomics, metaproteomics, meta-metabolomics, etc.) to complement metagenomics, expanding the landscape of tools available for the high-throughput analysis of complex biomes.

Even though the integration of multi-omic data into the 'trans-omic' pipeline[2] is able to generate unprecedentedly complete results, the analysis of such datasets often ignores the search for ecologically relevant conclusions, and focuses, instead, on getting increasingly exhaustive catalogues of species, expressed genes, or metabolites. As a consequence, multi-omics has the risk of increasing the complexity it is supposed to address. As an example of 'trees hidden by the forest', biological interactions among members of a microbial community often remain buried beneath the massive multi-omic datasets. It would be wrong to assume that because multi-omics is used, relevant biological interactions will emerge. It is known that microorganisms are naturally assembled into interacting communities, and that these community structures are directly linked to microbial processes. Therefore, the identification of key players in a taxonomically complex sample is necessary to understand the ecology of a particular habitat. This is especially true when it comes to the study of biotechnologically relevant microbial consortia, such as those present in the biogas industry[3], where engineers tend to consider their fermenters as 'black boxes' that produce biogas. Omic approaches certainly help to shed light on the taxonomic or functional complexity of a fermenting biomass, but such a strategy might fail to identify the ecological and economical core of the process. In plain words, the challenge is to reduce complexity to improve understanding.

We argue that a core of ecological conclusions has to emerge beyond the combination of the complex information obtained through multi-omics studies: multi-omic analyses should yield more than the sum of their parts, as suggested by some ecologists[4]. Even though there is still a long way to go, the relatively recent birth of so-called reverse ecology might prove helpful for the prediction of interactions among species and for improving our understanding of metabolic networks in the context of their natural habitats[5]. Moreover, alternative approaches to predict ecologically relevant information have emerged in the last couple of years. For instance, artificial neural networks have been proposed for modelling microbial communities as functions of environmental parameters and intra-microbial interactions[6]; the dynamics and composition of microbial consortia have been unveiled by measuring temporal variations in interspecies metabolic interactions[7]; and new approaches for visualizing microbial consortia through mathematical modelling and multi-dimensional scaling have been recently reported[8]. Nevertheless, the integration of experimental multi-omic data with predictive mathematical models based on mechanistic understanding is still considered a missing link in microbial ecology[9].

We envision a bright future in microbial ecology, where multi-omic databases will soon be analysed with approaches that are able to condense the gigabytes of information into simpler, ecologically relevant, conclusions. Does this mean that we will never again need to culture microorganisms? We strongly believe that culturing is needed more today than ever before. In fact, the mere identification and characterization of bacterial species or consortia through multi-omics may not be enough when

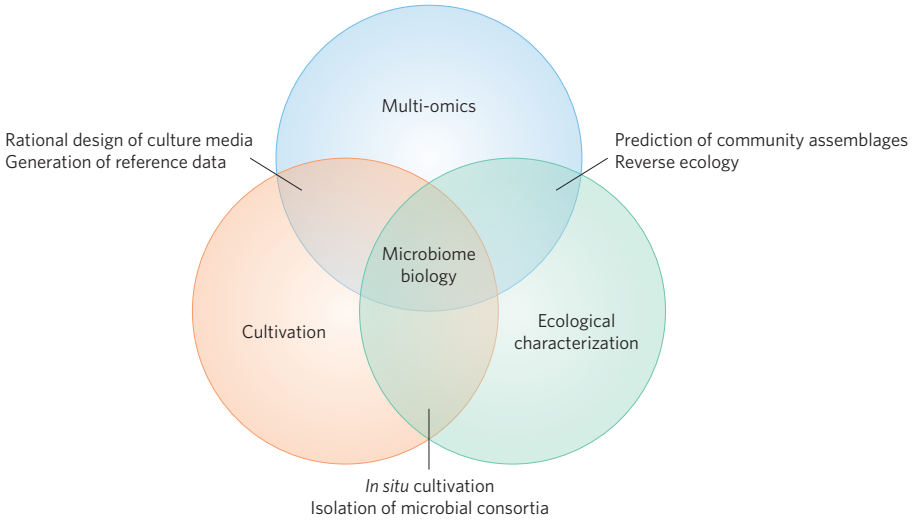

Figure 1 | Microbiome biology as the interphase among multi-omic data, ecological characterization, and cultivation-dependent techniques.

the interest of the study is focused on rare species, which still evade even high-coverage metagenomic sequencing, or when one pursues biotechnological applications. The development of single-cell genomic approaches[10], improved analysis of metagenomic data[11], or innovative sampling methods (like analysing the biomass passed through a ~0.2-μm filter[12]), have unveiled the extraordinary phylogenetic and functional diversity of many new microorganisms. Therefore, an approach able to fully characterize rare species, and to experimentally test the *in silico* predictions for particular microorganisms of biotechnological interest is needed. Such approaches do exist: culturing. In this sense, improved bioprospecting techniques aimed at the cultivation of hard-to-culture species have recently been developed. Among such techniques, *in situ* cultivation methods have proved successful to isolate bacterial species carrying novel gene sequences or producing new antibiotics[13]. These features cannot be predicted with multi-omic approaches, since they might only match with sequences of unknown function in the databases. Indeed, it is estimated that at least 7–60% of the sequences obtained through metagenomic sequencing cannot be properly classified due to the limiting number of reference annotated genomes in public databases[14]. Another obvious obstacle is when rare microorganisms are the 'ultra-small' bacteria, with incomplete metabolic networks, and so harbour an intrinsic difficulty for *in vitro* culture. Here, novel approaches for isolating natural microbial consortia[8], supported by efficient methods to predict microbial interactions, are essential. The isolation of rare, new-to-science species or microbial consortia allows not only their experimental characterization in the laboratory, but also the complete analysis of their genomes, which can then be used as new reference data and for improving our understanding of organismal and community biology. We argue that cultivation-dependent and cultivation-independent approaches not only complement each other, but in fact need each other. Multi-omics needs more reference genomes to better analyse new, complex, microbiomes, and microbial ecologists need multi-omics to know what else is out there, and thus what they can attempt to culture. To further support this complementary approach, multi-omic data analysed through metabolic modelling can be used to predict the essential nutrients required for the cultivation of hard-to-culture species on the basis of its metabolic network[15]. Therefore, we view microbiome biology not as a simple combination of multi-omic data, but as an emerging crossroad arising from the interphase between multi-omics, cultivation, and ecological characterization (Fig. 1).

A recent calculation based on scaling laws suggests that a trillion microbial species are yet to be discovered[16]. Studying those species is one of the greatest challenges of microbiology, and we argue that strategies transcending the information of multi-omics are essential for actually unveiling the composition and the ecology of such incredibly complex microbial communities. Simply combining layers of high-throughput biological data will result in improved databases and methodologies for the discovery of a myriad of unknown genes, microbial species, or metabolites, but their biology will remain obscure. Too often, biodiversity repositories are considered as biological enlightenment, rather than what they actually are: impressive yet raw sources of future knowledge, something not to be confused with knowledge itself. ❐

Cristina Vilanova and Manuel Porcar are at the Institute for Integrative Systems Biology (I2SysBio, University of Valencia-CSIC), the Cavanilles Institute of Biodiversity and Evolutionary Biology and Darwin Bioprospecting Excellence, SL, Universitat de València, postal code 22085, 46071 València, Spain.
e-mail: manuel.porcar@uv.es; cristina.vilanova@uv.es

References

1. Staley, J. T. & Konopka, A. *Annu. Rev. Microbiol.* **39,** 321–346 (1985).
2. Yugi, K., Kubota, H., Hatano, A. & Kuroda, S. *Trends Biotechnol.* **34,** 276–290 (2016).
3. Abendroth, C., Vilanova, C., Günther, T., Luschnig, O. & Porcar, M. *Biotechnol. Biofuels* **8,** 87 (2015).
4. Morales, S. E. & Holben, W. E. *FEMS Microbiol. Ecol.* **75,** 2–16 (2011).
5. Levy, R. & Borenstein, E. *Adv. Exp. Med. Biol.* **751,** 329–345 (2012).
6. Larsen, P., Dai, Y. & Collart, F. R. *Methods Mol. Biol.* **1260,** 33–43 (2015).
7. Zomorrodi, A. R., Islam, M. M. & Maranas, C. D. *ACS Synth. Biol.* **3,** 247–257 (2014).
8. Dorado-Morales, P., Vilanova, C., Garay, C. P., Martí, J. M. & Porcar, M. *Sci. Rep.* **5,** 18396 (2015).
9. Widder, S. *et al. ISME J.* http://doi.org/bkhv (2016).
10. Rinke, C. *et al. Nature* **499,** 431–437 (2013).
11. Albertsen, M. *et al. Nature Biotechnol.* **31,** 533–538 (2013).
12. Luef, B. *et al. Nature Commun.* **6,** 6372 (2015).
13. Ling, L. L. *et al. Nature* **517,** 455–459 (2015).
14. Prakash, T. & Taylor, T. D. *Brief. Bioinform.* **13,** 711–727 (2012).
15. Yus, E. *et al. Science* **326,** 1263–1268 (2009).
16. Locey, K. J. & Lennon, J. T. *Proc. Natl Acad. Sci. USA* **113,** 5970–5975 (2016).