microbial biotechnology

Open Access

Crystal ball

Getting back to the nature of the microbial world: from the description and inductive reasoning to deductive study after 'meta-omics'

Yong Nie[1] and **Xiao-Lei Wu**[1,2,3,*]

[1]*College of Engineering, Peking University, Beijing 100871, China.*

[2]*Institute of Ocean Research, Peking University, Beijing 100871, China.*

[3]*Institute of Ecology, Peking University, Beijing 100871, China.*

On 25th September 2015, after near 1-year revision–rejection–revision cycles, my first *Molecular Microbiology* paper was finally accepted for publication. I was so proud of this paper, as I thought it was the best paper of my research career. However, if the impact factor (IF) is used to estimate the quality of my research, I wondered why other omics articles written by my group could be published in journals of much higher IF, even though they were not the same calibre as my last MM paper. Worse even, the IF of *Molecular Microbiology* keeps declining, for reasons only known to those who invented this neat and devilish 'performance indicator'.

Impact factors may reflect the quality of the journal, but only within a particular professional field. For comparison in a broader field, it reflects the hotspot of the academic activities, which may not only be related to the science itself, but also to how researchers behave. The fact that a lot of researchers are following the research hotspots without real scientific questions is exactly the same as my own ill-gotten behaviour when perpetually chasing hot stocks, only to lose my money again and again.

'Meta-omics' undoubtedly represent a major hotspot in the field of microbial ecology. Since the development of next-generation sequencing (NGS) at the beginning of this century (Metzker 2010), the IFs of journals publishing 'meta-omics' works have been increasing. More and more microbiologists and microbial ecologist changed their focus to 'meta-omics'-related fields based on NGS technologies. One important reason is the declining cost of sequencing and the mass of data generated/obtained. Sequencing a metagenome is now cheaper (about $7/Gb data) than the expression and purification of a protein, due to its higher labour cost (Goodwin *et al*., 2016).

Based on metagenomic technologies, we can describe in great detail the microbial compositions and their functional genes in any type of environment we are interested in. The scientists of microbiology and microbial ecology have investigated the microbiomes in the ocean (Sunagawa *et al*., 2015), in the soil (Delgado-Baquerizo *et al*., 2018), in lakes (Zorz*et al*., 2019), in the air (Adams *et al*., 2013), inside buildings (Lax *et al*., 2017) and we even know about the microbial communities living in space stations (Sielaff *et al*., 2019). Besides, using sequencing technologies such as DNA/RNA stable-isotope probing and sequencing (Chen and Murrell, 2010), and meta-transcriptome sequencing (Urich *et al*., 2008), we can now bridge the gap between function and taxonomy information. The emergent bioinformatic and mathematic tools also promote the study of 'meta-omics'. The binning tools enable us to construct the genomes of single strains from a stack of chaotic sequence data (Sangwan *et al*., 2016). We can predict the microbial interactions using network analysis (Layeghifard *et al*., 2017) and speculate on the assembly and succession process of complex microbial communities using inductive reasoning (Zhou and Ning, 2017). Although most of the studies investigating microbial communities are descriptive, we can still learn much, for example what the biogeographic distribution of microorganisms is on earth (Thompson *et al*., 2017), or how microorganisms are related to soil productivity (Fierer 2017), how gut microbiota affect human health (Shreiner *et al*., 2015) and what roles microorganisms in the biogeochemical cycles of earth play (Falkowski *et al*., 2008), apart from discovering more and more unculturable microorganisms hitherto unknown to humanity (Parks *et al*., 2017).

Received 27 September, 2020; accepted 30 September, 2020.
*For correspondence. E-mail xiaolei_wu@pku.edu.cn; Tel./Fax+86-10-62759047
Microbial Biotechnology (2021) **14**(1), 22–25
doi:10.1111/1751-7915.13681

'Meta-omics'-related research undoubtedly improves our understanding of the microbial world. However, because most of the 'meta-omics' studies use descriptive methods and inductive reasoning, what exactly do these studies achieve? Whether these hypotheses or so-called theories can be tested and applied in other environments remains unknown. Also, whether the scientific questions in the present 'meta-omics' studies are real scientific questions or merely technical problems, and whether they are meaningful, remains in doubt and thus needs to be discussed. Generally, the answer to a real scientific question can explain a phenomenon, and the scientific hypothesis is a proposed explanation for the phenomenon, an explanation that is testable using scientific methods.

What else can we learn from 'meta-omics' and NGS? Imagine a time when the cost of sequencing becomes negligibly small, when automatic tools for bioinformatic and mathematical analysis are developed (online automatic platforms are now emerging, allowing users without prior experience to upload their sequencing raw data and perform popular bioinformatic and embedded mathematical analysis), 'meta-omics' analysis based on NGS would become common technologies like PCR used in every biological laboratory today. Critically, once artificial intelligence is mature enough, will it be able to assist, or even replace us in the analysis of patterns emerging from 'meta-omics' data? Or even develop hypotheses that explain those patterns?

To find the answer to these questions, we need to return to the scientific questions and hypotheses themselves. In my opinion, future scientists should focus less on using new technologies to simply generate vast amounts of data and instead return to first developing scientific hypotheses aimed at real scientific questions based on these data and pay more attention to test the hypotheses in a deductive approach. Although the inductive and deductive approaches are different, they are also complementary. For a complete cycle of scientific research, investigators/researchers begin with a set of phenomena, then look for patterns, before exploring different theories to explain these patterns and developing a working hypothesis. Next, they design experiments that allow the testing of their hypothesis, before either verifying or amending their theories. Recently, using a similar approach, several studies on 'synthetic microbial community' or 'synthetic microbiome' have emerged (see Lawson *et al*., 2019 for references?). These studies are characterized by a 'design–build–test–learn' (DBTL) cycle. The cycle begins with a hypothesis from a microbial ecological theory, before a model is designed that embodies the hypothesis. Subsequently, microbiomes are built on the basis of the design and then tested for their performances, including how function changes over time, to determine whether the results are in line with the hypothesis, and to ultimately obtain a deeper understanding from the results. This approach has also been used for industrial applications, such as biosynthesis of valuable products or degradation of refractory pollutants, based on rational design and construction of microbiomes (Gilmore *et al*., 2019; Sgobba and Wendisch, 2020).

The basis of design is the hypothesis and assumptions supporting the hypothesis. Hypotheses are constructed not to simply fit the data, but to explain phenomena observed. For example, the statement 'the community changed with temperature' is not really a hypothesis, as it only describes a phenomenon of the community. It fails to regard the mechanisms and cannot explain how the community forms in a given environment. By contrast, 'higher temperatures should favour the slower-growing species in a bacterial community with microbial interactions' is a much better hypothesis (Lax *et al*., 2020), which considers the mechanistic assumption (i.e. the difference in growth rates between the two species decreases when the temperature is not optimal for both species), and can be tested by experiments.

On the basis of such a hypothesis, a quantitative model can actually be proposed, and experiments can be designed. For synthetic microbial ecology research, a quantitative mathematical model is needed to simulate the dynamics of microbiomes (Zomorrodi and Segre, 2016). The Lotka–Volterra model is one of the most commonly used population dynamic models that describes pairwise microbial interactions. Another useful model is the individual-based model (IBM), which treats each individual cell as a discrete independent entity for exploring microbial interactions in space. In addition, the genome-scale stoichiometric model based on flux balance analysis (FBA) quantitatively predicts the cellular metabolic fluxes to build a bridge between systems biology and microbial ecology. These models greatly simplify the complexity of microbiomes and identify the critical parameters suitable for experimental interrogation. Experiments can now be designed to test the prediction of the hypothesis and the quantitative models.

Synthetic microbiomes can be constructed using either a top-down or bottom-up approach to meet the assumptions of a hypothesis of synthetic microbiomes.

In contrast to the top-down approach, the bottom-up assembly begins with defined axenic strains or engineered strains. To test the fitness, the axenic strains are usually co-cultured to test the fitness to each other and used to explore the effects of microbial interactions and random processes on the stability of microbial communities (Friedman *et al*., 2017; Gokhale *et al*., 2018; Abreu *et al*., 2019). They are also used to study the trade-offs of different parameters such as temperature and growth

rate (Lax *et al*., 2020). Compared with the axenic strains, the engineered strains have a clearer genetic background, and the interactions between X and Y are under better control. Using genetic circuits, more and more types of microbial interactions have recently been constructed artificially by engineering strains. Using quorum-sensing regulatory modules, investigators/scientists/researchers can now construct consortia: is mutualism a consortium? how can 'competition' be 'constructed?' including mutualism, competition and predation interaction modes, and explore the critical factors on these interactions (Song *et al*., 2014). Moreover, interactions based on metabolite exchange such as cross-feeding can also be constructed between inter- or intraspecies: inter already means 'between/across', intra already means 'within' -> for both within and across species (Shou *et al*., 2007). However, the consortia constructed with engineered strains usually contain fewer members than those from axenic strains. They are mainly used to explore microbial interactions. Ultimately, no matter what approach is being used, each is driven by a scientific question or hypothesis.

One of the bottlenecks for efficiently studying synthetic microbiomes is the lack of high-throughput tests of the co-cultures. Although mathematical models point out the critical parameters for the construction of microbiomes and simplify the design of experiments, the number of possible co-cultures is extremely high, especially for synthetic microbiomes designed to contain multiple species. Moreover, for those from axenic strains, mass replicates are needed to ensure the stability of the results. A small number of high-throughput methods have recently been developed for constructing and testing the synthetic microbiomes, containing up to several thousands of co-cultures per day (Kehe *et al*., 2019). Another challenge to X is presented by Y the high-throughput investigation of microbial compositions of mass co-cultures. For co-cultures from a few species, the morphology of colonies can be used to discriminate different species (Celiker and Gore, 2014). Therefore, to match the need for synthetic microbiome research, it is critical to develop low-cost methods based on NGS and techniques of intelligent image recognition. The inductive methods are also important for the iterative process between model prediction and the testing of the hypothesis.

In summary, the development of low-cost/high-yield technologies greatly affects the way we do science. If publishing work solely based on new technologies, without actually asking a good scientific question becomes the norm, then our progress in understanding how nature works will slow considerably. New technologies might be the focus for tech-focused researchers, but it should never be the focus of an entire research discipline. 'Omics' studies have by now deposited massive amounts of data into the databases, and it is now time to return to the question as to what can we actually learn from them. Without testing, the hypotheses and theories from 'omics' data based on inductive reasoning are only in the air. It is now time to return to ask hypothesis-driven scientific questions in the field of microbial ecology and find approaches suitable to finding answers. Increased application of the deductive approach in synthetic microbial ecology and synthetic microbiome research will undoubtedly provide exciting new opportunities for advancing our understanding of microbial ecology.

Acknowledgements

This work was supported by National Key R&D Program of China (2018YFA0902100 and 2018YFA0902103) and National Natural Science Foundation of China (31761133006, 31770120 to XLW, 31770118 to YN).

References

Abreu C. I., Friedman J., Andersen Woltz V. L., and Gore J. (2019) Mortality causes universal changes in microbial community composition. *Nature Communications*, **10** (1): 2120. http://dx.doi.org/10.1038/s41467-019-09925-0

Adams, R.I., Miletto, M., Taylor, J.W., and Bruns, T.D. (2013) Dispersal in microbes: fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances (vol 7, 1262, 2013). *ISME J* **7:** 1460.

Celiker, H., and Gore, J. (2014) Clustering in community structure across replicate ecosystems following a long-term bacterial evolution experiment. *Nat Commun* **5:** 4643.

Chen, Y., and Murrell, J.C. (2010) When metagenomics meets stable-isotope probing: progress and perspectives. *Trends Microbiol* **18:** 157–163.

Delgado-Baquerizo, M., Oliverio, A.M., Brewer, T.E., Benavent-Gonzalez, A., Eldridge, D.J., Bardgett, R.D., *et al*. (2018) A global atlas of the dominant bacteria found in soil. *Science* **359:** 320–325.

Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320:** 1034–1039.

Fierer, N. (2017) Embracing the unknown: disentangling the complexities of the soil microbiome. *Nat Rev Microbiol* **15:** 579–590.

Friedman, J., Higgins, L. M., and Gore, J. (2017) Community structure follows simple assembly rules in microbial microcosms. *Nat Ecol Evol* **1:** 0109.

Gilmore, S.P., Lankiewicz, T.S., Wilken, S., Brown, J.L., Sexton, J.A., Henske, J.K., *et al*. (2019) Top-down enrichment guides in formation of synthetic microbial consortia for biomass degradation. *Acs Synthet Biol* **8:** 2174–2185.

Gokhale, S., Conwill, A., Ranjan, T., and Gore, J. (2018) Migration alters oscillatory dynamics and promotes survival in connected bacterial populations. *Nat Commun* **9:** 5273.

Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016) Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* **17:** 333–351.

Kehe, J., Kulesa, A., Ortiz, A., Ackerman, C.M., Thakku, S.G., Sellers, D., *et al.* (2019) Massively parallel screening of synthetic microbial communities. *Proc Natl Acad Sci USA* **116:** 12804–12809.

Lawson, C.E., Harcombe, W.R., Hatzenpichler, R., Lindemann, S.R., Loffler, F.E., O'Malley, M.A., *et al.* (2019) Common principles and best practices for engineering microbiomes. *Nat Rev Microbiol* **17:** 725–741.

Lax, S., Abreu, C.I., and Gore, J. (2020) Higher temperatures generically favour slower-growing bacterial species in multispecies communities. *Nat Ecol Evol* **4:** 560–567.

Lax, S., Sangwan, N., Smith, D., Larsen, P., Handley, K. M., Richardson, M., *et al.* (2017) Bacterial colonization and succession in a newly opened hospital. *Sci Transl Med* **9:** eaah6500.

Layeghifard, M., Hwang, D.M., and Guttman, D.S. (2017) Disentangling interactions in the microbiome: a network perspective. *Trends Microbiol.* **25:** 217–228.

Metzker, M.L. (2010) APPLICATIONS OF NEXT-GENERATION SEQUENCING Sequencing technologies - the next generation. *Nat Rev Genet* **11:** 31–46.

Parks, D.H., Rinke, C., Chuvochina, M., Chaumeil, P.A., Woodcroft, B.., Evans, P.N., *et al.* (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol* **2:** 1533–1542.

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

Sgobba, E., and Wendisch, V.F. (2020) Synthetic microbial consortia for small molecule production. *Curr Opin Biotechnol* **62:** 72–79.

Shou, W.Y., Ram, S., and Vilar, J.M.G. (2007) Synthetic cooperation in engineered yeast populations. *Proc Natl Acad Sci USA* **104:** 1877–1882.

Shreiner, A.B., Kao, J.Y., and Young, V.B. (2015) The gut microbiome in health and in disease. *Curr Opin Gastroenterol* **31:** 69–75.

Sielaff, A.C., Urbaniak, C., Mohan, G.B.M., Stepanov, V.G., Tran, Q., Wood, J.M., *et al.* (2019) Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. *Microbiome* **7:** 50.

Song, H., Ding, M.Z., Jia, X.Q., Ma, Q., and Yuan, Y.J. (2014) Synthetic microbial consortia: from systematic analysis to construction and applications. *Chem Soc Rev* **43:** 6954–6981.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., *et al.* (2015) Structure and function of the global ocean microbiome. *Science* **348.**

Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., *et al.* (2017) A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551:** 457–463.

Urich, T., Lanzén, A., Qi, J., Huson, D.H., Schleper, C., and Schuster, S.C. (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS One* **3:** e2527.

Zhou, J.Z., and Ning, D.L. (2017) Stochastic community assembly: does it matter in microbial ecology?, *Microbiol.Mol Biol Rev* **81:** e00002-17.

Zomorrodi, A.R., and Segre, D. (2016) Synthetic ecology of microbes: mathematical models and applications. *J Mol Biol* **428:** 837–861.

Zorz, J.K., Sharp, C., Kleiner, M., Gordon, P.M.K., Pon, R.T., Dong, X.L., and Strous, M. (2019) A shared core microbiome in soda lakes separated by large distances. *Nat Commun* **10:** 4230.

