Can we predict keystones?

Lisa Röttjers and Karoline Faust

In their recent Opinion article, Banerjee, Schlaeppi and van der Heijden claim that microbial networks can identify keystones; that is, taxa with a high impact on the structure and functioning of ecosystems (Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **16**, 567–576 (2018))[1]. Although we agree with the authors on the importance of keystones, we doubt that a highly connected taxon in a microbial network (a hub) is necessarily a keystone, and we therefore want to moderate their claim that over 200 microbial keystones have been identified.

Keystones are of particular interest because they have a greater impact on the ecosystem than other taxa. The classic experimental validation of keystones involves comparing the effects of keystone removal and/or addition with the removal and/or addition of other community members. As such experiments are difficult to carry out, only few microbial taxa have been experimentally confirmed as keystones[2–5]. If keystones could be accurately predicted from microbial networks, experiments would no longer be required. However, edges in microbial networks usually do not represent known ecological interactions, but rather statistically significant co-occurrences or mutual exclusions of taxa in sequencing data.

The question is how accurately inferred microbial networks can identify keystones. Weiss and colleagues did not validate keystones, but showed that the prediction accuracy for ecological interactions in inferred microbial networks is low[6]. We found that the prediction accuracy for synthetic hub taxa is also low and, furthermore, that the prediction of ecological interactions is hampered by underlying environmental gradients[7]. By contrast, Berry and Widder did report a high prediction accuracy for synthetic keystones[8]. As few microbial keystones are known, all three evaluations have been carried out on noise-free synthetic data that were generated in a manner that does not necessarily reflect the processes in real ecosystems. In summary, the evidence for accurate keystone prediction from inferred networks is mixed at best.

The large number of keystones reported by Banerjee and colleagues[1] is surprising. However, upon closer inspection of the 29 references in supplementary table 1 that predict keystones from networks, only two report experimental validation. Moreover, in some of these, a network is built from samples taken in different conditions, so that it is not clear whether a hub is connected to many taxa because it influences them or because the taxa change together in response to the altered condition. Of the remaining eight references, three demonstrated that the addition of keystones had an effect on community structure but did not show that they were also hub taxa in microbial networks. When only counting experimentally validated keystones, we arrive at a total of 7 instead of 200 microbial keystones.

Metabolic networks constructed from metagenomic data are a promising alternative to co-occurrence networks. The links in metabolic networks do not need to be inferred, as they represent known biochemical reactions. Keystones can be identified as taxa that provide 'bottleneck' reactions in the metabolic network[9].

In conclusion, we think that better network inference tools and more validation experiments are needed before hub taxa in inferred networks can be classified as keystones.

There is a reply to this letter by Banerjee, S., Schlaeppi, K. & van der Heijden, M. G. A. *Nat. Rev. Microbiol.* https://doi.org/10.1038/s41579-018-0133-x (2018).

Lisa Röttjers[1] *and Karoline Faust*[1]*
KU Leuven Department of Microbiology and Immunology, Rega Institute, Laboratory of Molecular Bacteriology, Leuven, Belgium.
**e-mail: karoline.faust@kuleuven.be*

https://doi.org/10.1038/s41579-018-0132-y

1. Banerjee, S., Schlaeppi, K. & Heijden, M. G. A. Keystone taxa as drivers of microbiome structure and functioning. *Nat. Rev. Microbiol.* **16**, 567–576 (2018).
2. Garrett, W. S. et al. Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. *Cell Host Microbe* **8**, 292–300 (2010).
3. Hajishengallis, G. et al. Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host Microbe* **10**, 497–506 (2011).
4. Ze, X. et al. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. *ISME J.* **6**, 1535 (2012).
5. Agler, M. T. et al. Microbial hub taxa link host and abiotic factors to plant microbiome variation. *PLOS Biol.* **14**, e1002352 (2016).
6. Weiss, S. et al. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J.* **10**, 1669–1681 (2016).
7. Röttjers, L. & Faust, K. From hairballs to hypotheses–biological insights from microbial networks. *FEMS Microbiol. Rev.* **42**, 761–780 (2018).
8. Berry, D. & Widder, S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front. Microbiol.* **5**, 219 (2014).
9. Roume, H. et al. Comparative integrated omics: identification of key functionalities in microbial community-wide metabolic networks. *NPJ Biofilms Microbiomes* **1**, 15007 (2015).

Competing interests
The authors declare no competing interests.