

Analysis and interpretation of metagenome data from microbiomes and complex microbial communities

Rohan B.H. Williams

Singapore Centre for Environmental Life Sciences Engineering (SCELSE)
Nanyang Technological University and National University of Singapore

The investigation of microbiomes and complex microbial communities is now a major research direction in biology, health and medicine. *The ability to understand the biology of such microbial consortia, and develop their translational potential, is underpinned by the computational analysis of metagenome data.*

In this tutorial we will cover the major approaches to analysing and interpreting metagenome data obtained from microbiomes and complex microbial communities.

I will commence with a short historical survey of the field, focusing on the limitations of culture techniques for isolating environmental bacteria, which drove the development of non-culture based methods such as 16S amplicon sequencing.

My major focus will be on analysis and interpretation of whole community shotgun metagenomics, obtained with short read sequencing technologies (Illumina). I will describe how these complex data can be analysed using different approaches, namely 1) read level analysis making use of reference sequence databases; 2) via direct use of reference genomes and 3) by using metagenome assembly methods as a way of recovering draft genomes of the member species of the community. I will discuss both taxonomic and functional analysis of these data. Emphasis will be placed on the strengths and limitations of each approach, and the practical issues related to their use and interpretation (including command line usage examples and review of specific workflows).

I will conclude by describing several new approaches that are gaining importance, including HiC-metagenomics, long read metagenomics and single cell based methods using microfluidics.

Recommended reading:

Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J., Segata, N. (2017), Shotgun metagenomics, from sampling to analysis, *Nature Biotechnology* **35**: 833-84.

Speaker biography

Rohan Williams is Head of the Integrative Analysis Unit at the Singapore Centre for Environmental Life Sciences Engineering (SCELSE), an autonomous research centre in microbial ecology and biofilm biology co-hosted by the National University of Singapore and Nanyang Technological University. Following undergraduate studies in Physics, Williams obtained PhD in Medicine from the University of New South Wales (Sydney, Australia) in 2003. From 2004-2007 he was an NHMRC Peter Doherty Fellow at UNSW and then a Group Leader at the Australian National University (2007-2011) prior to taking up his present appointment. His interests and expertise lie in statistical bioinformatics, design, analysis and interpretation of experiments using genomic technologies, systems microbiology and the analysis of complex microbial communities using multi-omics approaches.