Identifying synthetic microbial communities using flow cytometry and machine learning

Peter Rubbens[1*], Cristina Garcia-Timermans[2*], Ruben Props[2], Nico Boon[2], Willem Waegeman[1]

[1]KERMIT, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University
[2]Center for Microbial Technology and Ecology (CMET), Ghent University
Ghent, Belgium

$*$ Contributed equally

Flow cytometry, machine learning, microbial community dynamics, synthetic microbiology.

Abstract

Synthetic ecology studies are on the rise [1]. As synthetic microbial communities allow us to validate specific microbial hypotheses, they have become exceedingly more popular over the past few years [2–4]. However, we still lack an unbiased method for quantifying the community composition in function of certain ecological questions. Recently, our group has been able to retrieve the community composition of low-complexity synthetic bacterial communities using a combination of flow cytometry and supervised machine learning methods [5]. Flow cytometry is a high-throughput technology, able to give a multivariate phenotypic description of a single cell based on scatter and fluorescence information. As flow cytometry is capable of measuring more than thousands of cells per second, this high-throughput method lends itself to the use of sophisticated predictive models in order to discriminate between bacterial populations.

We have shown that bacterial cultures of varying taxonomic identity can be identified phenotypically using flow cytometry [5]. For low-complexity communities, single cells can be identified with an accuracy higher than 90% for more than half of the communities on a low-end cytometer. A recent study reported that the identification capacity is even better when biological replicates were analyzed on a high-end cytometer [6]. These properties are independent from the microbial community at hand, which means that they can be further exploited

for future experiments. These results motivate that flow cytometry has not yet reached its full potential for synthetic microbiology. Therefore, we propose a data-driven approach to develop and improve identification methods for microbial flow cytometric analyses. This will allow us to use flow cytometry to its full capacity and to characterize synthetic microbial communities up to high precision.

References

[1] Michael T. Mee and Harris H. Wang. Engineering ecosystems and synthetic ecologies. *Mol Biosyst*, 8(10):2470–2483, 2012. doi:{10.1039/c2mb25133g}.

[2] Tobias Grosskopf and Orkun S. Soyer. Synthetic microbial communities. *Current Opinion in Microbiology*, 18:72–77, APR 2014. doi:{10.1016/j.mib.2014.02.002}.

[3] Karen De Roy, Massimo Marzorati, Pieter Van den Abbeele, Tom Van de Wiele, and Nico Boon. Synthetic microbial ecosystems: an exciting tool to understand and apply microbial communities. *Environmental Microbiology*, 16(6, SI):1472–1481, JUN 2014. doi:{10.1111/1462-2920.12343}.

[4] Stefanie Widder, Rosalind J. Allen, Thomas Pfeiffer, Thomas P. Curtis, Carsten Wiuf, William T. Sloan, Otto X. Cordero, Sam P. Brown, Babak Momeni, Wenying Shou, Helen Kettle, Harry J. Flint, Andreas F. Haas, Beatrice Laroche, Jan-Ulrich Kreft, Paul B. Rainey, Shiri Freilich, Stefan Schuster, Kim Milferstedt, Jan R. van der Meer, Tobias Grosskopf, Jef Huisman, Andrew Free, Cristian Picioreanu, Christopher Quince, Isaac Klapper, Simon Labarthe, Barth F. Smets, Harris Wang, Orkun S. Soyer, and Isaac Newton Inst Fellows. Challenges in microbial ecology: building predictive understanding of community function and dynamics. *The ISME Journal*, 10(11):2557–2568, NOV 2016. doi:{10.1038/ismej.2016.45}.

[5] P. Rubbens, R. Props, N. Boon, and W. Waegeman. Flow cytometric single-cell identification of populations in synthetic bacterial communities. *PLoS One*, 12(1):1–19, 2017. doi:10.1371/journal.pone.0169754.

[6] P. Rubbens, R. Props, C. Garcia-Timermans, N. Boon, and W. Waegeman. Stripping flow cytometry: how many detectors do we need for bacterial identification? *Under review*, 2017.