Title: DISCo: Design of an identifiable synthetic community

Authors: Dana L. Carper1, Travis J. Lawrence1, Alyssa A. Carrell2, Dale A. Pelletier1, Dave J. Weston1

Affiliations:

1 Biosciences Division, Oak Ridge National Laboratory, Oak Ridge TN, USA

2 Bredesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee, Knoxville, TN, USA

Abstract:

Background:

Multicellular eukaryotes are associated with complex communities of microorganisms. These host associated communities provide many beneficial functions for the host organism including, nutrient acquisition, pathogen defense, and moderate abiotic stresses. Although, DNA sequencing-based approaches have given insight into determining taxonomically who is present in the community and functional potential of that community but interpretation and reconciling of these results can be extremely difficult. For this reason, some researchers have shifted approaches to synthetic or constructed based microbial communities’ experiments to unravel host-microorganism interactions. This method involves adding known microbial strains to a host organism allowing for an easier interpretation of colonization, shifts in community structure, microbe-microbe interactions and with genomic information of the strain’s potential functions. However, there is no clear method for how to design a reproducible microbial community for these experiments.

Designing of a microbial community can be carried out by many methods. Some researchers, primarily in microbial engineering have focused on functional assets of the microbes in order to engineer specific microbial outputs. While this method works well for bio-engineering purposes it may not be as informative for microbiome studies. Host associated synthetic communities have primarily consisted a very few strains and confirmed through re-isolation, limiting the extrapolation and interpretation in more diverse communities. Designing a diverse community that is reproducible and can be distinguished through cheaper sequencing technologies has been extremely difficult without strong computational abilities.

In this paper, we describe an easy to use command-line software program, DISCo, for the creation of a reproducible diverse community of organisms that can be distinguished through DNA sequencing technology. Currently this program will take an input alignment of sequences and create the most diverse community possible at a specified sequence difference level. The program can either start with no previous knowledge or with an initial input community of organisms of interest. The program can also take into account the number of strains wanted in the final community or the proportions of taxonomic groups through a sub setting procedure. DISCo also has error correcting ability for DNA sequences for interpretation after synthetic community experimentation. While this program was initial designed for bacterial community construction, any DNA alignment can be used and thus users are not restricted to specific organismal group. This program is implemented in python and is available through github, conda or pip installation.

Implementation:

FIGURE 1 (Diagram of workflow)

Results and Discussion:

FIGURE 2 (Community diversity)

FIGURE 3 (Error correction accuracy)

References:

Availability and requirements

Project name: DISCo: Design of an identifiable synthetic community

Project home page: <https://github.com/dlcarper/DISCo-Design-of-an-identifiable-synthetic-community.git>

Operating system(s): Platform independent

Programming language: Python

Other requirements: None

License:

Abbreviations:

DNA: Deoxyribonucleic acid

FASTA: Fast-all (file format)