Project Proposal

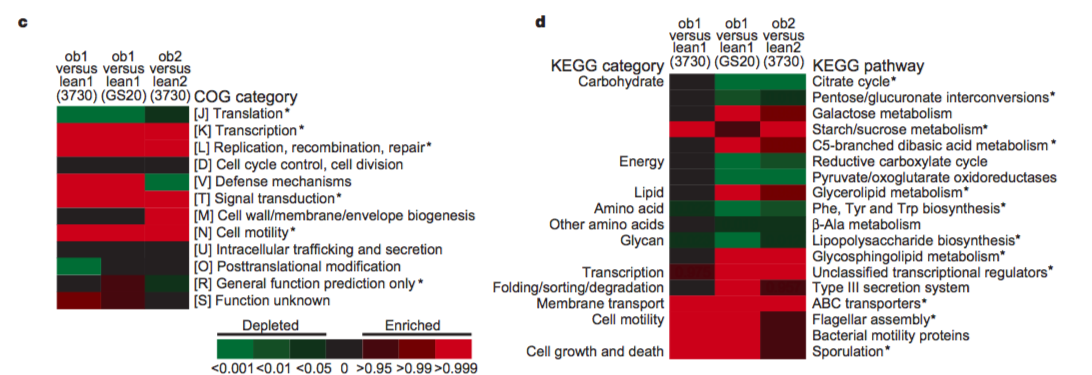
Dustin Michels

27 Oct 2017

**Scientific Question**

I am interested in comparing “types” of proteins found in different parts of the ocean, following somewhat in the vein of the “Structure and function of the global ocean microbiome” paper we read for the annotated bibliography assignment. (Sunagawa et al., 2015). It seems that the authors of this paper compared functional diversity in parts of the ocean, after functionally annotating all the reads from the TARA dataset, using the eggNOG and KEGG databases. I wonder if I can use those annotations or generate my own to replicate some of the functional analyses between groups we saw elsewhere?

For instance, in the journal club paper that investigated the distal gut microbiomes of mice, the authors quantified the number of genes in the COG category and KEGG pathway for obese mice vs. lean mice. Perhaps I could produce a similar graphic to the one below, but instead of comparing obese mice to lean mice I could compare two different depths or geographic regions, and see if certain COG or KEGG pathways are consistently enriched or depleted between the two? (Figure 1, Turnbaugh et al., 2006).

****

**Methods**

We have used interproscan to annotate the genes of our contigs, and we have looked up the KEGG pathway online, but I don’t think we have used tools to functionally annotate contigs. Perhaps I could find such a tool (KEGG mapper?[[1]](#footnote-1)) or perhaps I could simply use the results published by Sunagawa et al. If I wanted to compare two depths, perhaps I could generate or reference annotations for all the contigs for all the class data at those two zones, then simply count the number of contigs mapped to each functional group. I could explore various data visualization tools for generating an insightful graphic, like the one in the Turnbaugh et al paper.

This is a rough idea that surely needs some guidance and refining. I’ll just mention that I’m interested in using BioPython / writing some useful and reusable Python script.

**References**

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D.R., Alberti, A., et al. (2015). Structure and function of the global ocean microbiome. Science *348*, 1261359.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature *444*, 1027–1131.

1. <http://www.kegg.jp/kegg/tool/annotate_sequence.html> [↑](#footnote-ref-1)