insert amazing title here

Specific Aims

background and motivation, testable hypothes(is/es), two specific aims, and a summary of the proposed research. [1]

some bacteria that aren't taxonomically closely related can be capable of producing some of the same gene products. just like different charismatic megafauna can have the same niche in an ecosystem (predator X and Y both eat prey Z), different bacteria can cover the same niche. looking only at taxonomy may not be enough to understand the role and response of the gut microbiome in health and disease processes.

some studies use PiCrust in an ill-conceived attempt to reconstruct the possible functional potential of communities from 16S sequence. That's a dumb idea. Let's use metagenomics and metabolomics instead, since metagenomes tells us the actual genes capable of producing products, and metabolomics tells us the functions that the microbiome are performing at the time of sampling.

- techniques: - metagenomics: functional potential of a community. - 16S OTUs: taxonomic community composition. - lc-ms/ms: metabolites in the gut. - some metabolies are from bacteria, human, and diet. and there's overlap. - use AMON to identify known features as from bacteria vs human vs either. - limited to what we know in KEGG. - use GNPS networking to cluster unknown features with known ones. - "Another issue is that there is an overlap between metabolites that are externally derived, produced by host metabolism, and produced through microbial cometabolism, which adds to the difficulty in setting up these types of databases; DAS is one such metabolite." - Johnson 2016 - get taxonomic community composition and functional community structure - should get fcnl comm struct from metagenomes and/or metabolomes? - find one paper that classifies fcnl potential from metagenomes, another from metabolomes. use both techniques and compare? - "definition of functional redundancy indicating the mere ability of multiple distinct organisms to perform a specific function... is practical" [louca 2018]. - but if we have metabolomes, we can verify - are these more or less correlated in CRC compared to healthy? - wrt carcinogenic functions / genes / metabolites - comparisons of taxonomic composition, functional gene composition, and metabolomes: - within disease states - between disease states - taxonomic vs functional gene composition - metabolomic vs functional gene composition - TODO: look at notes from talk by Robert (lastname?) of MSU from last year. used GNPS and QIIME in gut microbiome study. - limitations: - temporal changes in metabolites sampled. this study is not longitudinal.

hypotheses: - Clustering communities based on functional potential improves classification into disease states compared to clustering based on taxonomic composition. - When there is not a significant difference in taxonomic composition of CRC and normal communities, there is a significant difference in their functional potential. and in their actual metabolites. - when CRC communities have different taxonomic composition, they have similar functional composition.

Aim 1. Assess concordance of functional redundancy and taxonomic composition of the microbiome in CRC

Hypothesis:

- is there correlation between taxonomic composition and functional potential (composition) of microbiome? - does the idea of functional redundancy hold up during CRC? - is concordance stronger/weaker in CRC vs normal? - some studies couldn't find strong signal in 16S OTUs to differentiate CRC from normal. [Weir 2013] - where OTU composition doesn't show a difference, does functional potential show a significant difference? - does taking fcnl redundancy into account improve classification of samples into CRC vs normal? even in studies that did find a difference?

Aim 2. metabolites concordance with metagenomes in CRC

- does functional redundancy from metagenomes correlate with actual metabolites? - some toxigenic gene products produced by bacteria are known. can we use GNPS to identify unknown features as

potentially toxigenic products based on their similarity to known toxigenic ones?
insert concluding/summary sentence here

References

[1] S. Bikel, A. Valdez-Lara, F. Cornejo-Granados, K. Rico, S. Canizales-Quinteros, X. Soberón, L. Del Pozo-Yauner, and A. Ochoa-Leyva. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Computational and Structural Biotechnology Journal*, 13:390–401, Jan. 2015.