**Notes on UCI presentation slides**

Anders Dohlman

Slide 1. “…from time series data” ?

Slide 2. Good job of explaining the ubiquity the microbiome. I think its also important here to mention **why** its important to build a microbial interactome.

* We’ve observed the efficacy of fecal transplants, probiotics, and identified links between microbiota and disease, but *mechanism* remains murky
* Its difficult to make predictions about effect of introducing new species
* We’d like to be able to engineer precise bacterio-therapeutic interventions
* Systems level understanding will help explain the role of microbiome in health

Slide 3.

* I like the figure on the right -- did you make that yourself?
* Why just two groups, though? Need more than two to calculate correlations
* Careful about using “direction” here, as it may be confused with directed networks. Perhaps better to say “positively associated” or “negatively associated”?

Slide 4.

* Main idea: directed network allows to you untangle *causality*
  + ie. Which is the dependent variable?
* Why time series? Because a time delay can indicate causality
  + True causality can only be identified by experimentation
  + But some smart mathematicians (Granger) have identified quantitative definitions of causality using time delay

Slide 4.5 Help to set up the “problem” here…

Why can’t we use normal network detection methods, like generalized linear models developed for mapping gene networks? Because microbiome data is

* Compositional
* Highly sparse

Lots of people have designed methods to account for these characteristics!

Slide 5. Great job of articulating the problem and the motivations here

Slide 6. Good!

Slide 7.

* Aim1, measuring *prediction accuracy* ie. recovery of ecological interactions
* Aim 2, may need to introduce these ecological interactions before?
* Aim 3, idea is to use an unbiased time series dataset

Slide 8.

* Sampling frequency: eg. gut microbiomes fluctuate on the scale of hours or days, but soil fluctuates on the scale of months

Slide 9.

* Ecological interactions here are important re: causality
  + Parasitisim defines one species as the cause, another as the effect
* Perhaps include generalized Lotka Volterra equation here?
  + d*xi*/dt = *xi* ( r + A*x* )
    - x is a vector of species
    - r is a vector of inherent growth rates
    - A is the community matrix

Slide 10-12.

* We don’t need to go into huge detail here about the system… just say they have wastewater with bacteria that we can sample from
* Summarize experimental design:
  + sample every 2 weeks for 12 months
  + 2 technical replicates x 3 biological replicates per timepoint?
  + Plus RT-qPCR to calculate bacterial load
* I think fine to show a table for 16S sequencing costs