Speaker: Susan Holmes, Stanford Statistics

Title: Human microbiome: Using the data, all the data, without all the noise

Abstract:

Microbiome data are collected as noisy reads from particular genes. Some serve as the fingerprints of specific taxa. I will show that if the noise is ignored one can obtain very biased measures of diversity.

We have proposed dada2: a denoising mechanism that uses all the available data and allows a much cleaner picture of the variations in the microbiome. Using these denoised data, we are then able to leverage sparse methods to combine the known phylogenetic tree with other sources of information, providing useful decompositions of taxa variabilities across individuals and time. This has proved a useful approach for detecting markers of preterm birth and understanding the resilience of certain subjects to perturbations.

This is joint work with David Reman’s group at the Medical School as well as Ben Callahan, Kris Sankaran, Pratheepa Jeganathan, Julia Fukuyama, John Cherian, Lan Nguyen and Joey McMurdie.