**Tao Zhang**
**Combining DNA fingerprints to dissect microbial communities**
**Dr. James Foster**
**May 5th, 2014**

## Summary

- It is hard to dissect microbial communities.
- Human is one out of 2 million named species. There are ten times more bacterial than human cells in your body. Up to 1 Billion species, there's only about 5,000 are known, and approximate only 3% of bacteria can be grown in the lab.
- Dr. James pointed out that there is one technique can distinguish one species from another, called "fingerprint". They developed a technique for combining DNA sequnce data from multiple fingerprints, while the current techniques can only use one single fingerprint. He talked about high throughput fingerprinting:
  - Get "every" DNA molecule in a sample: break cells up, wash, filter
  - Isolate fingerprint regions from all bacteria
  - Sequence them all

  There will be 10-20million fingerprint sequences. Then infer how many of which species were there.
- To interpret fingerprint data:
  - Compute similarity between fingerprint sequences
  - Cluster, call a cluster a "species"
  - Number of clusters is species richness
  - Size of clusters species abundance