Taxonomic classifications of single- and paired-end sequences
Human microbiome

The 16S-based approach
- Extract DNA
- Amplify and sequence 16S rRNA
- Group similar sequences into OTUs
- Use database to identify OTUs

Community composition: Which organisms are present?
- Abundance
- Relative abundance of OTUs in community
- OTU phylogeny
- Variant sequences and SNPs
  - GATTACA
  - GATTACA
  - GATTACA
  - GATTACA

PICRUSt
- Predict metagenome functional content from marker gene

The shotgun metagenomic approach
- Extract DNA
- Sequence community DNA
- Compare sequences to reference genomes
- Community function: What can the community do?
- KEGG
- SEED
- BLAST
- Use database to identify sequences
- Abundance
- Functions
- Relative abundance of gene pathways in community
Functional analyses of the microbiome

Morgan Langille et al Nature Biotechnology (2013)
The 16S-based approach:
- Extract DNA
- Amplify and sequence 16S rRNA
- Group similar sequences into OTUs
- Use database to identify OTUs
- Community composition: Which organisms are present?
- OTU abundance
- Relative abundance of OTUs in community
- OTU phylogeny

The shotgun metagenomic approach:
- Extract DNA
- Sequence community DNA
- Compare sequences to reference genomes
- Community function: What can the community do?
- Use database to identify sequences
- Relative abundance of gene pathways in community
Functional analyses of the microbiome

Data cleaning
- Low quality removal
- Host sequences removal
  (HMP: BMTagger or Bowtie2)
Functional analyses of the microbiome

Extract DNA

The shotgun metagenomic approach

Compare sequences to reference genomes

Sequence community DNA

Community function: What can the community do?

Use database to identify sequences

Relative abundance of gene pathways in community

Abundance Functions

KEGG
SEED
BLAST

NCBI
Functional analyses of the microbiome

1. **Extract DNA**
2. **The shotgun metagenomic approach**
3. **Sequence community DNA**
4. **Compare sequences to reference genomes**
5. **KEGG, SEED, BLAST**
   - **KEGG**: Use database to identify sequences
   - **SEED, BLAST**: Relative abundance of gene pathways in community
6. **Community function**: What can the community do?
7. **Abundance**
8. **Functions**
9. **NCBI**
10. **blastx**: translated nucleotide \( \rightarrow \) protein
11. **USEARCH**: Ultra-fast sequence analysis
12. **UBLAST**

protein sequence database
The Microbiome study

Current techniques for human microbiome profiling

Which organisms are present?
- Extract DNA
- Amplify 16S rRNA gene by PCR and sequence
- Bin similar sequences into OTUs
- Compare OTUs to databases
- Identify OTUs in sample and relative frequencies

Phylogenetic view of community composition
- Identify microbial sequences, variants, and polymorphisms in sample
- GATTACA

What are the functions of the community?
- Sequence total DNA
- Compare sequences to reference genomes
- GATTACA
- GATTCTA

Identify genes, pathways, and relative frequencies in sample
- Functions

Emerging methods

How does the community respond to its environment?
- Extract RNA, protein, and small molecules
- Reverse transcribe
- Mass spectroscopy
- Metabolite profiling

Key:
- Regulation
- Flux
- Interaction

Xochitl Morgan et al. Trends in Genetics 2013
Baochen Shi
CNSI 4338, UCLA
biosbc@gmail.com