How much has been sequenced?

Number of known sequences

Environmental sequencing

First bacterial genome

100 bacterial genomes

1,000 bacterial genomes
How much will be sequenced?

- Everybody in Hawaii
- Everybody in USA
- All cultured Bacteria
- 100 people
- Most major microbial environments
- One genome from every species

Year:
- 1/1/83
- 12/29/92
- 12/27/02
- 12/24/12
- 12/22/22
- 12/19/32
- 12/17/42
Why Metagenomics?

• What is there?
• How many are there?
• What are they doing?
• Experimental manipulations?
• Real time metagenomics?
Uncultured Viruses

200 liters water
5-500 g fresh fecal matter

Concentrate and purify viruses or bacteria

Extract nucleic acids

DNA/RNA LASL

Sequence

Epifluorescent Microscopy
How do you sequence the environment?

- Extract DNA
  - Soil extraction kit
  - Water extraction kit

- Create library
  - LASLs
  - fosmids

- Sequence fragments
Linker-Amplified Shotgun Libraries (LASLs)

↓ Soil Extraction Kit

↓ Hydroshear

↓ Blunt-ending

↓ Addition of Linkers

↓ Amplification of Fragments

→ Clone and Transform

This method produces high coverage libraries of over 1 million clones from as little as 1 ng DNA

David Mead - Lucigen™

Breitbart (2002) PNAS
Early Attempts at a Metagenomics Platform

- http://phage.sdsu.edu/~rob/cgi-bin/remoteblast.cgi

- Submit BLAST to local and remote databases
  - Local (as fast as possible)
  - NCBI (one search every 3 seconds)

- Many concurrent searches
  - One search versus 1,000 searches

- Parse data into tables for Excel
  - Access to taxonomy etc
Human-associated viruses

- More bacteria than somatic cells by at least an order of magnitude
- More phages than bacteria by an order of magnitude
- Sample the bacteria in the intestine by sampling their phage
Most Viral DNA Sequences in Adult Human Feces are Unknown Phages

- Known: 40%
- Unknown: 60%

Most Human RNA Viruses are Known

- Known: 92%
- Unknown: 8%

Other Plant Viruses: 9%
Pepper Mild Mottle Virus: 65%

Other: 26%

Pepper Mild Mottle Virus (PMMV)

- ssRNA virus; ≈6 kb genome
- Related to Tobacco Mosaic Virus
- Infects members of *Capsicum* family
- Widely distributed – spread through seeds
- Fruits are small, malformed, mottled
- Rod-shaped virions

Viral particles in fecal sample

TOBACCO MOSAIC VIRUS http://www.rothamsted.bbsrc.ac.uk/pplinks/virusems/
PMMV is common in Human Feces

Fecal samples

→

Extract total RNA

→

RT-PCR for PMMV

San Diego: 78% people are positive
Singapore: 67% people are positive

10-50 fold increase in feces compared to food

$10^6$-$10^9$ PMMV copies per gram dry weight of feces
Which Foods Contain PMMV?

- Chili powder
- Chili sauces
- NOT FOUND IN FRESH PEPPERS

Indian curry
Pork noodle red chili
Chicken rice
Chinese food
Hong Kong chili sauce
Hong Kong green chili
Vegetarian chili
Koch’s Postulates
Human microbial metagenome is more important than human genome

NIH started the human microbiome project last year

- Nasal
- Oral
- Skin
- Gastro-intestinal
- Urogenital
Artists impression: not all machines explode!
Uncultured Viruses

200 liters water
5-500 g fresh fecal matter

Concentrate and purify
viruses or bacteria

Extract nucleic
acids

DNA/RNA
LASL

Pyrosequencing

Sequence

Epifluorescent
Microscopy
BACK!