20.106J – Systems Microbiology
Lecture 11
Prof. DeLong

- Chapter 15 – Brock

- Genomics
  - DNA sequencing technology – things have really changed. There's a real race going on for who can develop the best technology
    - Human genome project: only around 30,000 genes in the human code.
    - The day is not at all far off when doctors will read people's genomes to discover what their inherent risks are.
    - The human genome project involved two main groups – one more commercially based (J. Craig Venter – Celera), and one more public, open source, with funding from NIH (Francis Collins – NHGRI). Also the Sanger Centre, Whitehead Institute…

- The human genome project drove innovation in biotechnology. Two major technological benefits:
  - Stimulated development of high throughput methods – the assembly line. It's not just the individual with a pipette any more – it's more like a factory approach (which matters for the social aspect of how science works).
    - However, this might work back in the other direction as efficient machines develop…
  - Reliance on computational tools for data mining and visualization of biological information
    - Biology is rapidly becoming informational science – bioinformatics and computational biology.

- DNA sequencing
  - Sanger's technique
    - Uses primer extension and DNA polymerase
    - Dideoxynucleotides halt the replication at particular base pairs. Then you run for length on a slab gel, and you can tell which base pairs are at which locations, reading off the sequence and recording them manually.
  - Later people realized that you can use fluorescent labels instead of radiolabels.
    - This meant that you didn't have to deal with radioactivity
    - It also meant that you could run them all in one lane.
    - Instead of a slab gel, people use a thin tube, with a fluorescence detector automatically reading the wavelengths as they come out the other end.
    - This method is fast and accurate

- Graph: the progress of the Human Genome Project
- Shotgun genome sequencing – what about the gaps?
  - Doing a complete sequence for a bacterial genome is still expensive – around 10 or 20 thousand dollars, but it's relatively straightforward.
- Small insert cloning/sequencing
- There are automated plate pickers now
- Some people thought that shotgun sequencing would work (the Celera group used this) but others thought that it wouldn't (the public consortium used a different method)
- BAC sequencing – larger segments, you know where they lie on the genome
- Both approaches have advantages: shotgun sequencing is much quicker, but BAC sequencing closes the gaps – the best method is to use some of both.
- 454 Sequencing – Commercial, massively parallel DNA sequencing technology:
  - Run PCR with beads, each of which attaches to on strand of DNA
  - Emulsion PCR
  - Lots of DNA comes from an individual strand.
  - You pour the reaction over a chip with tiny wells that each fit one bead.
  - 1.5 million reactions all at once
  - Thus the factory approach to sequencing is no longer necessary – it can all be done in little machines like this.
  - It's still pretty expensive, but it's getting closer, and it's changing things rapidly.
- The evolution of sequencing is moving towards nano-methods, and pretty soon there will probably be sequencing technologies in hospitals.

- Microbial genomes
  - The first complete, published microbial genome came in 1995.
  - Incomplete bacterial genomes in process is now approaching 2,500.
  - Archeal genomes have been studied less.
  - Most of the DNA sequencing is happening at select sequencing centers, and it will probably continue to consolidate in this way.
  - Archeal genomes are generally smaller.
  - The smaller genomes are typically associated with obligate parasites.
    - There are obligate symbionts in some insects that live inside the cell and just work as little amino acid factories.
      - Called Carsonella
      - They have extremely small genomes
      - It's also very very AT rich (very little GC content)
      - The genome basically only has replication mechanisms and the DNA for amino acid synthesis

        - These symbionts are on their way toward becoming essentially organelles
    - We’re figuring out what sort of genes scale with genome size.