DNA Sequencing lecture (no wet lab)
http://www.dnalc.org/ddnalc/resources/sangerseq.html
http://www.dnalc.org/ddnalc/resources/cycseq.html

On board: Explain Sanger sequencing method

- Based on ordinary DNA synthesis using a DNA polymerase (usually Klenow fragment of *E. coli* DNA pol)
- Components of the reaction:
    1. Single stranded template or target DNA
    2. Sequence-specific primer (specificity affected by temp, concentration, etc. like PCR)
    3. DNA polymerase
    4. **All 4 dNTPs** (IMPORTANT)
    5. **ONE** of the 4 dideoxynucleotide triphosphates (ddATP, or ddCTP, or ddTTP, or ddGTP)
- Incorporation of ddNTP is random; when it occurs, **chain termination**
- Because there are many copies of the template being read simultaneously, a big random mix of newly synthesized DNA products is made, **of all sizes,** each ending in the ddNTP of that reaction
- Traditionally, this mix of DNAs is run on a (denaturing) polyacrylamide gel which can resolve bands differing in size by as little as 1 nt
- If synthesized DNA was radioactively labeled, then each ddNTP reaction is run in a separate lane on the gel, and the sequence is read 5' to 3' from the bottom of the gel up.

Polyacrylamide gels:

- Poured as liquid with a catalyst added which **crosslinks** the acrylamide into polyacrylamide, irreversibly solidifying it (can't melt & re-pour like agarose)
- Very thin gels (1-2 mm), usually poured between two glass plates
- **Much higher resolution** of fragment sizes than agarose; can easily separate DNA fragments that differ in size by only one nucleotide
- (Fragment sizes separated in this way are small, from about 20-300 nucleotides)
- Polyacrylamide sequencing gels usually contain urea as a denaturing agent, **so DNA runs single-stranded**

Automated sequencing:

- Each ddNTP is fluorescently labeled with a different "color"
- All reactions can be run together, again generating a mix of all DNA fragment sizes, each ending in a ddNTP
- Products are separated by size (1 nt difference) and analyzed by a laser & detector.
- Data are presented as a **chromatogram** showing fluorescence intensity and color of each DNA fragment.
- Chromatogram sequence is read 5' to 3', from left to right.

Point out that relatively ***higher*** concentration of ddNTP in reaction will give sequence read relatively ***closer*** to the primer (as ddNTP incorporation is more likely earlier). Conversely, a relatively lower [ddNTP] gives longer reads (further away from primer).

Show animations:

http://dnalc02.cshl.edu/ddnalc/resources/sangerseq.html (sangerseq movie, downloaded)

http://dnalc02.cshl.edu/ddnalc/resources/cycseq.html (automated seq movie, downloaded)

Discuss **shotgun sequencing**

- A *strategy* for sequencing large DNAs (uses Sanger method)
- A single Sanger method reaction can give you about 200-500 nt of sequence
- To continue sequencing a longer DNA, you would need to then synthesize a sequence-specific primer at the end of the sequence you just read to start the next Sanger reaction
- If your DNA is a Mb-size chromosome, this would take forever.
- Shear your big DNA randomly into small (~2 kb) fragments
- Clone MANY of these fragments into plasmids (one fragment per plasmid) and transform them into bacteria to make a **library**
- Use two "universal" sequencing primers, which hybridize to the plasmid on either side of the insertion site, to sequence into the insert (your cloned small piece of the big DNA)
- If you sequence enough clones, you'll get sequence data for the whole original DNA
- It is important to **sequence both strands** (i.e., from both ends of the insert) to maximize accuracy as certain regions of ssDNA sequence poorly
- Because the DNA was originally broken up at random, many of the clones should overlap.
- Use computer to find regions of overlap and reassemble all the small sequence bits into one big long sequence for the original molecule. This reassembled puzzle is called a **contig**.
- If there are **gaps** in your contig (regions with no overlap), you can either choose more clones to sequence at random, or, design sequence-specific primers to sequence into the gap

Show animation (on the web):
http://smcg.cifn.unam.mx/enp-unam/03-EstructuraDelGenoma/animaciones/humanShot.swf