Obtaining DNA from degraded samples for NGS sequencing

A brief overview of Alexander (Sasha) Mikheyev's lecture at USC 03/13/14

Presented by Jacqueline Robinson
04/23/2014

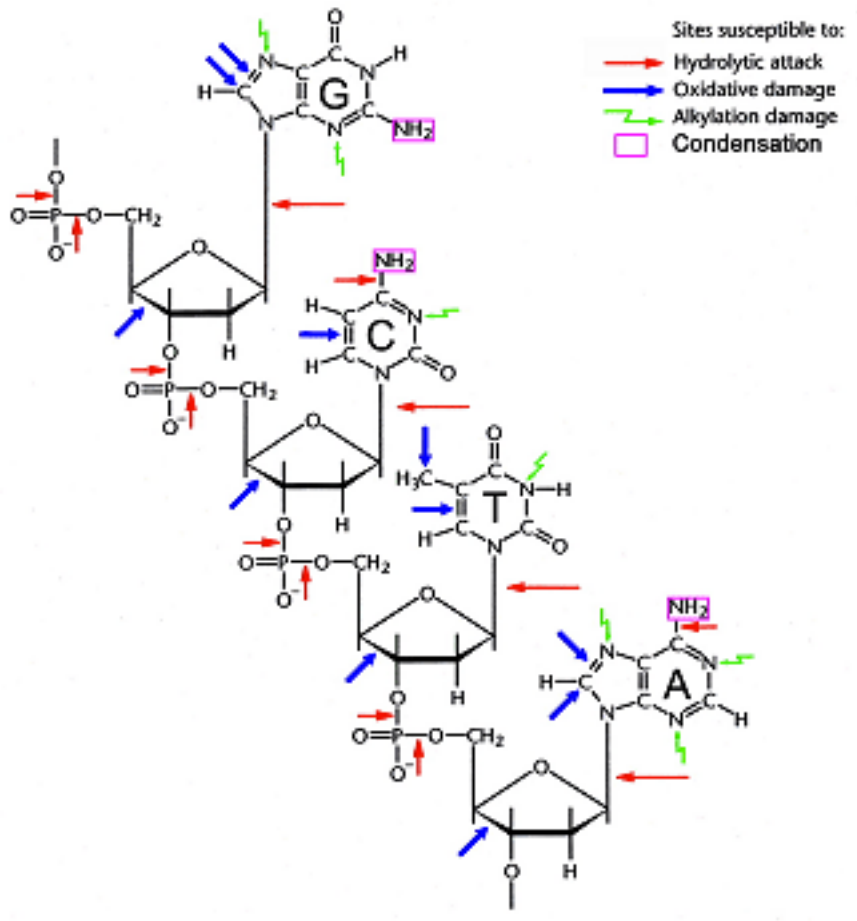
NGS is great, but...

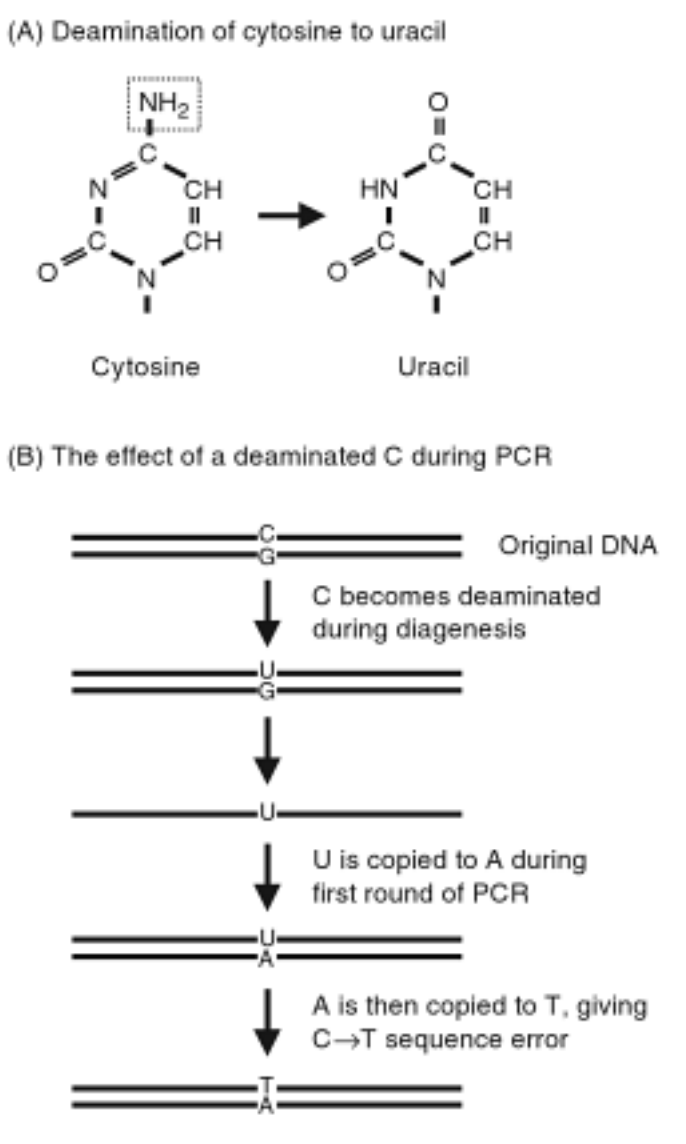
- Standard protocols require large amounts (~1ug) of input DNA that is of high quality
- This presents challenges for samples that are...
  - Old – ancient samples, museum samples, DNA from long-term storage freezers, etc.
  - Maybe not old, but poor quality – inadequately preserved, treated with chemicals (formalin-fixed, paraffin-embedded), badly extracted, etc.
  - Low in quantity – many reasons

Degraded DNA Challenges

- Degraded DNA is **highly fragmented**, tends to be **single stranded** (ssDNA), and contains many **altered nucleotides**
- Additionally, samples tend to be from museum or archaeological specimens which are rare, precious, and **may not be destroyed or damaged**
- Samples are often **highly contaminated** with microbial and other environmental DNA

Molecular Alterations





How does DNA get degraded?

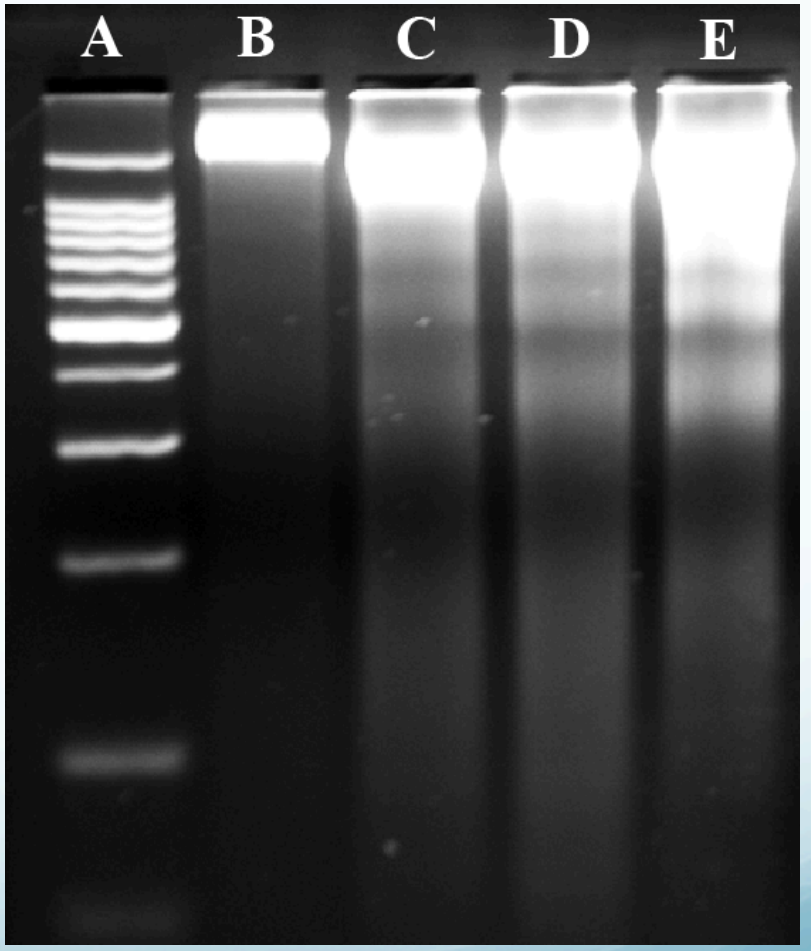
- Enzymatic reactions (autolysis)
- Microbial attack
- Hydrolytic reactions with water
- Oxidative reactions with oxygen
- Heat and various other forms of radiation

How “ancient” can we go?

- Record for oldest authentic ancient DNA sequence: ~ half a million years
- Theoretically, under ideal preservation conditions (such as storage in permafrost), DNA that is **less than 1.5 million years old** could be recovered and sequenced
- DNA half-life is estimated to be several hundred years*, depending on preservation conditions (cold, dark, and dry is best)
  - *http://rspb.royalsocietypublishing.org/content/279/1748/4724

How do you know if your DNA sample is degraded?

- Run the DNA through a gel (1-2% agarose, 45-60 mins)
- High quality DNA will appear as a bright, thin band at the top of the lane with little spreading (B)
- Low quality DNA will appear as a smear through the lane (C-E)



If you don't have extracted DNA yet...

- Use **Adam McCoy's magic buffer** to non-destructively obtain DNA from your sample
- Protocol involves soaking the sample overnight in a buffer that contains some nasty chemicals, followed by silica bead extraction
- Extraction time/difficulty not much worse than simple Qiagen extraction, and is less labor-intensive than phenol-chloroform extraction
- Seems cheap (beads might be expensive)
- Can be used for any DNA extraction, actually

What's in the Magic Buffer?

- For 100ml:
  - 50g **guanidine isothiocyanate** *(inactivates DNAses)*
  - 50ml **ultrapure water**
  - 5.3ml 1M **Tris-Cl**, pH 7.5 *(permeabilize membranes, buffer pH)*
  - 5.3ml 0.2 M **EDTA** *(chelating agent, reduces DNAse activity)*
  - Stir until dissolved
  - Add 10.6ml 20% **Sarkosyl** *(surfactant)* and 1ml **2-mercaptoethanol** *(DNA antioxidant)*
  - optional: 200 μg/ml **proteinase K** – may improve yield

Before making libraries

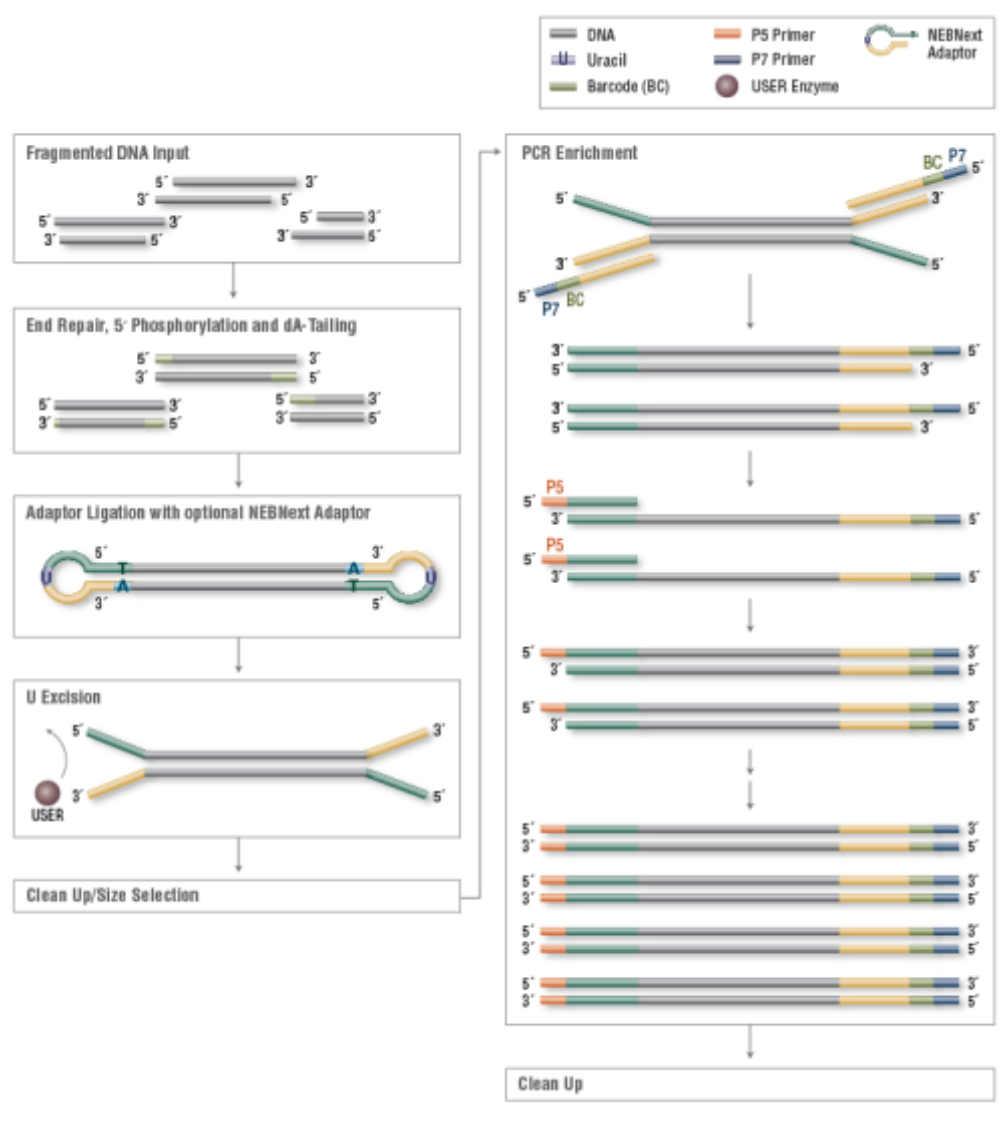
- Always: NanoDrop to check approximate DNA concentration and 260/280, 260/230 ratios
  - 260/280 should be 1.8 – 2.0 (pure DNA: 1.8, pure RNA: 2.0)
  - 260/230 should be higher than the 260/280 ratio, and ideally 2.0 – 2.2
    - Low ratios indicate contamination, but are also seen with very low DNA concentration
- Use Qubit quantification to get an estimate of dsDNA concentration
- Highly recommend alcohol (ethanol or isopropanol) precipitation if there is sufficient DNA quantity
- Lastly, no need to shear DNA that is highly fragmented, can just do a size selection (<250bp) with a gel excision

Making a library from degraded DNA

- Most library prep. protocols involve PCR
  - PCR incorporates more nucleotide errors (both *spontaneous* and *systematic*)
  - Without careful cycle optimization, results in PCR duplicates
  - PCR is necessary when there is a very low amount of starting material (currently, <200ng)
- PCR-free approach incorporates fewer errors, but requires a higher amount of starting material

Standard Library Protocol (NEB)

- Input DNA can be very low – even as low as 5ng – with standard PCR approach
- Requires dsDNA
- Kits are expensive
- General workflow: shear DNA, end repair/blunting, add A's, ligate (indexed) adapters, PCR, elution
  - Bead cleanups in between every step!



Protocol with PCR

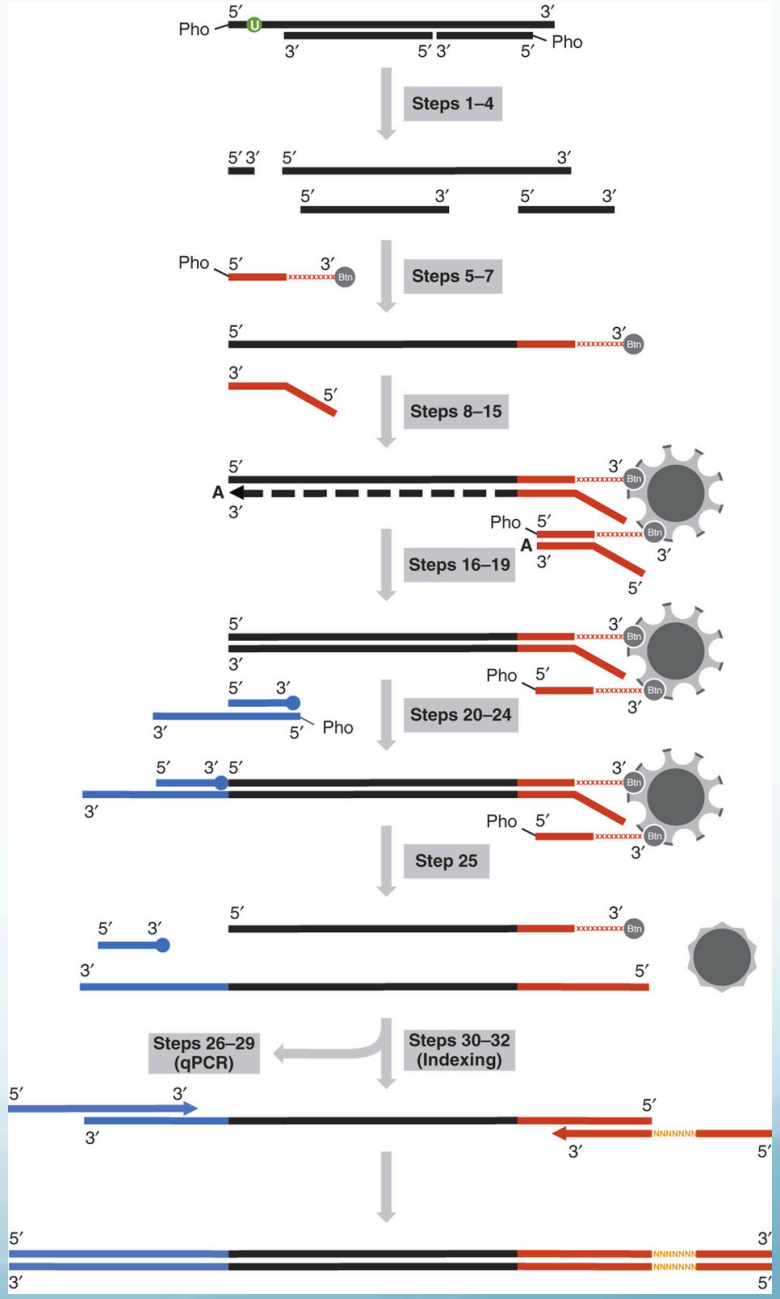
1. Repair 3' ends (dephosphorylate, add –OH)
2. Add –GGGGG to 3' end with TdT (special polymerase that does not require a template)
3. Synthesize complementary strand from template strand with exo-Klenow fragment (polymerase that lacks exonuclease activity which would otherwise destroy ssDNA)
   - Reaction incorporates –CCCCC oligonucleotides that are complementrary to –GGGGG ends, forming double-stranded binding sites for polymerase
4. Blunt ends with T4 DNA polymerase
5. Column purification with Qiagen MinElute
6. Blunt-end ligation of adapters with T4 DNA ligase
7. PCR with high-fidelity Phusion polymerase
8. Bead purification
9. Quantify, pool, purify, run Bioanalyzer, quantify with qPCR, dilute

Protocol without PCR

1. Repair 3' ends (dephosphorylate, add –OH)
2. Add –GGGGG to 3' end with TdT (special polymerase that does not require a template)
3. Ligation with T4 DNA polymerase
   - adapter has C's on one end that are complementary to –GGGGG end
4. Synthesis of complementary strand with high-fidelity Phusion polymerase
5. Dephosphorylation with calf intestinal phosphatase (prevents circularization)
6. Column purification with Qiagen MinElute
7. Blunt-end ligation of adapters with T4 DNA ligase
8. Bead purification
9. Run Bioanalzyer, quantify with qPCR, pool, run Bioanalyzer, dilute

Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA

- Marie-Theres Gansauge & Matthias Meyer
  - Max Planck Institute
  - Nature Protocols, 2013
- Single-stranded amplification of DNA bound to biotinylated adapters
- Requires 15pg-15ng DNA
- Used to sequence Denisovan autosome to 30X coverage



Now you can sequence!

- Paired or single-end Illumina short read sequencing
- Deep coverage recommended, but really depends on your application
- No special requirements for the sequencing – just hand over your libraries to the professionals
- Post-sequencing notes:
  - Check quality and PCR duplication rate with FastQC
  - Must map to a reference: currently, no way to construct a *de novo* genome using only short fragments
  - Estimate contamination using mtDNA

Questions?

- Thank you, A. Mikheyev & Google
- Resources:
  - A. Mikheyev page: https://groups.oist.jp/evolution
    - Protcols: http://ecoevo.unit.oist.jp/lab/
  - All about Illumina adapters: http://tucf-genomics.tufts.edu/documents/protocols/TUCF_Understanding_Illumina_TruSeq_Adapters.pdf
  - Ancient DNA library prep protocol from Nature Protocols: http://www.nature.com/nprot/journal/v8/n4/pdf/nprot.2013.038.pdf