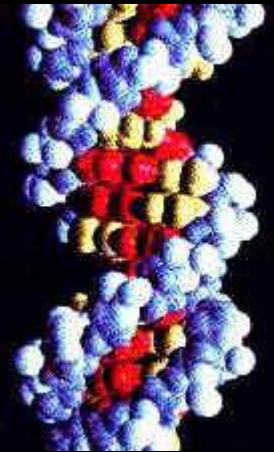


Genomics


Sequencing genomes



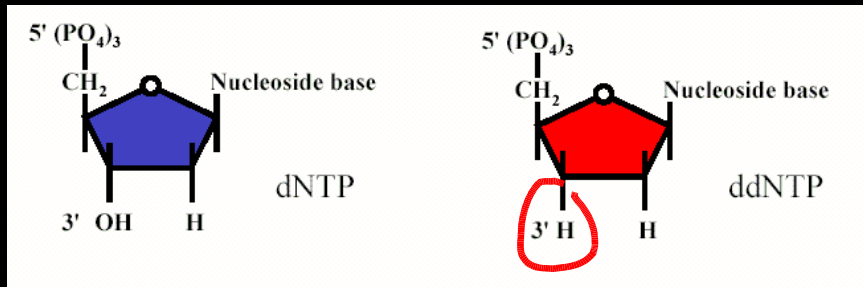


DNA Sequencing

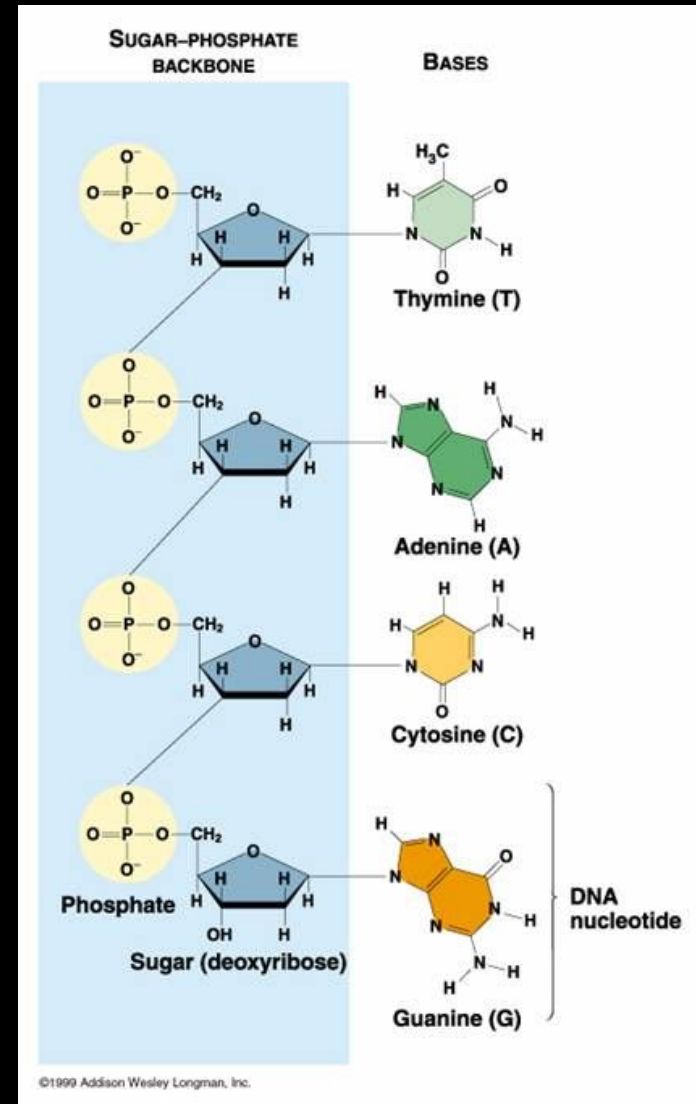
Chain Termination Method

- Sanger, 1977
 - single stranded DNA, ~800b
 - Method:
 - Electrophoresis can separate DNA molecules differing 1bp in length
 - Dideoxynucleotide (*ddNTP*) are used - which stop replication
- 

ddNucleotides



ddA, ddT, ddC, ddG
Each type marked with fluorescent dye
When incorporated into DNA chain – stops replication



Chain Termination Method, An Outline

Replication

- Obtaining ssDNA
- Add a (universal) primer

Start replication in a soup of A,T,C,G



Continuously add tiny amounts of ddA, ddT, ddC, ddG

- gradually stopping all the processes

Chain Termination Method, Reading the Sequence

Running through
electrophoresis gel

- Four types of ddNTP have four different fluorescent labels
- Automated reading

See: <http://www.dnalc.org/Shockwave/cycseq.html>



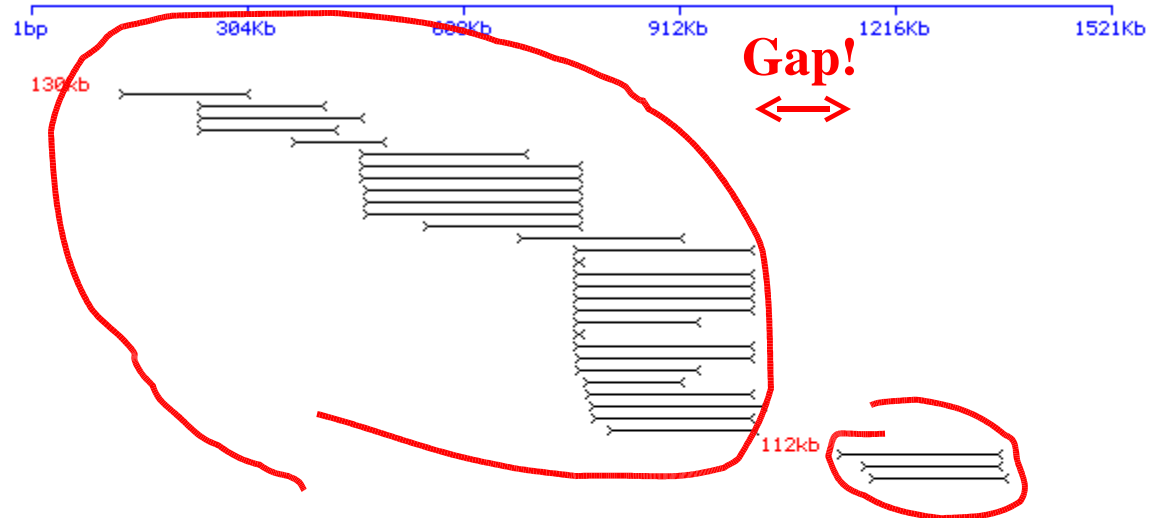
Shotgun Method - Overview

Cut genome into short fragments

Sequence DNA fragments

Create contigs

Contig - continuous set of overlapping sequences





Shotgun Method – Contig Construction

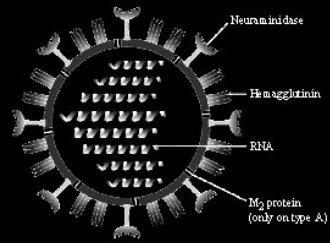
Two DNA sequences:
X=CTATCA
Y=AGTAT

How do they overlap?

X_____Y or Y_____X

Try to apply dynamic programming

Shotgun Method – Haemophilus Influenzae Sequencing



Extract DNA

Sonicate

Electrophoresis

DNA library

1.5-2kb

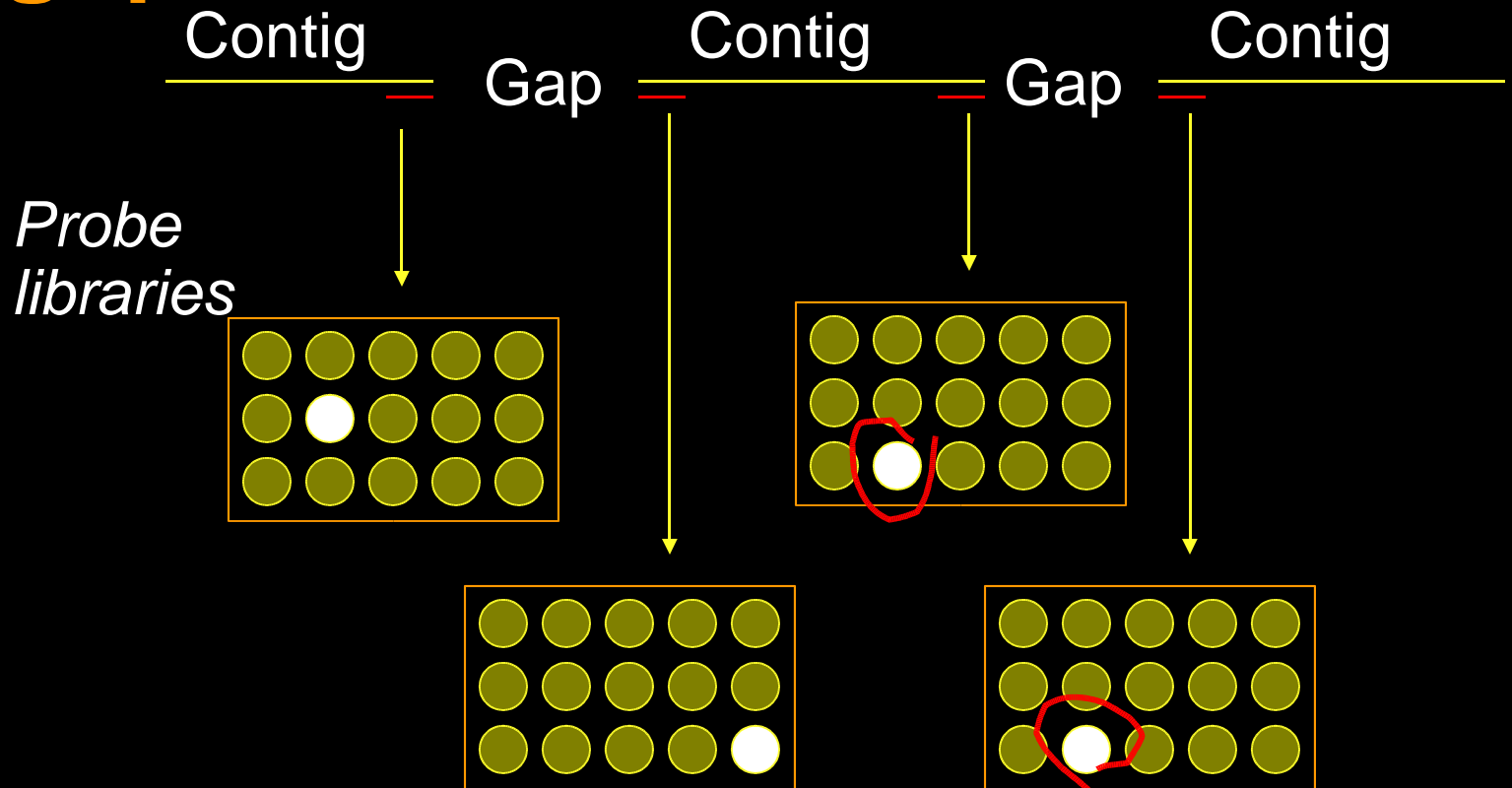
Sequence

Construct
contigs

Sequenced



Shotgun Method - Filling in gaps





Shotgun Method - Pros and Cons

Pros

- Human labour reduced to minimum

Cons

- Computationally demanding – $O(n^2)$ comparisons
- High error rate in contig construction
 - Repeats as the main problem

Shotgun vs. Hierarchical Method



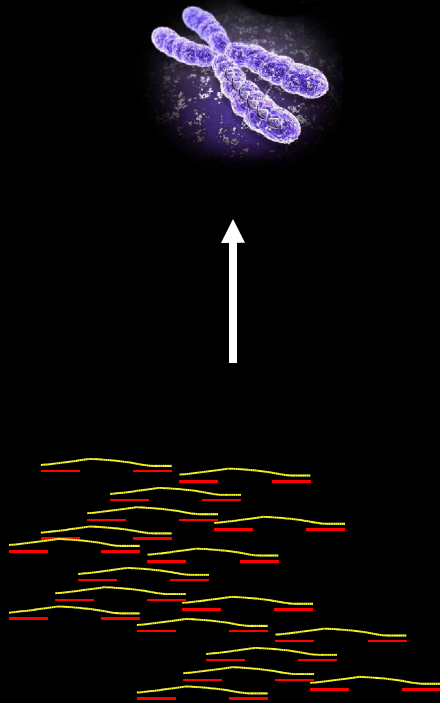
Celera vs. Human Genome Project

Hierarchical (top-down) assembly:

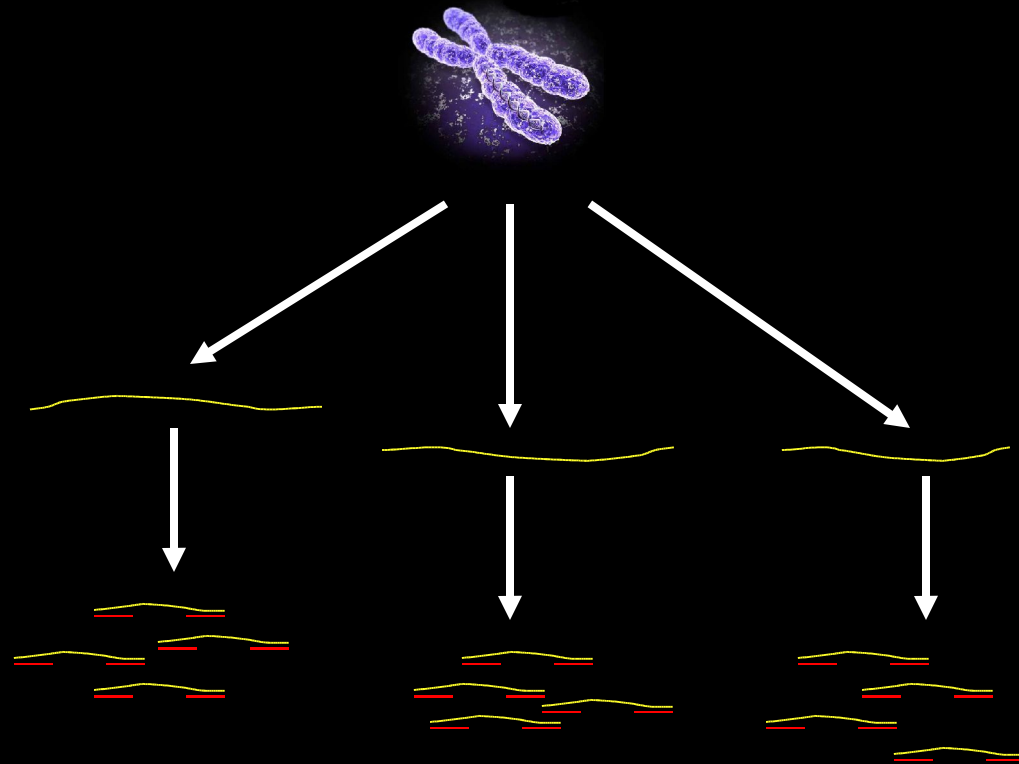
- The genome is carefully mapped
- “Shotgun” into large chunks of 150kb
 - Exact location of each chunk is known
- Each piece is again “shotgun” into 2kb and sequenced

Shotgun vs. Hierarchical Method

Shotgun
bottom-up



Hierarchical
top-down





New Sequencing Methods

Sequencing By Hybridization

- Check which from all possible fragments of length k (k -tuples) hybridize to the sequence

TAA

AAG

AGC

ATTCG

TAAGC