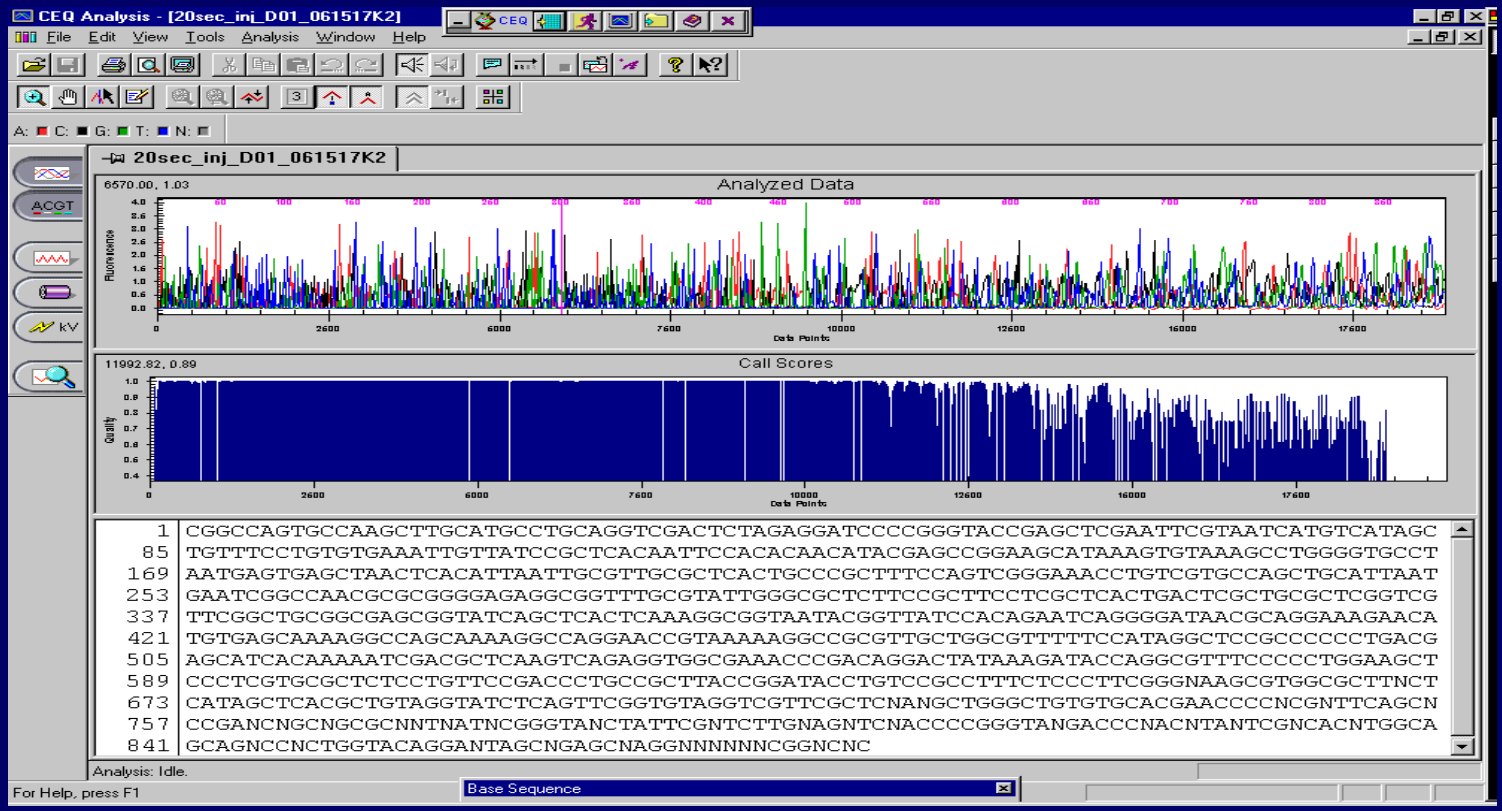
Genome sequencing

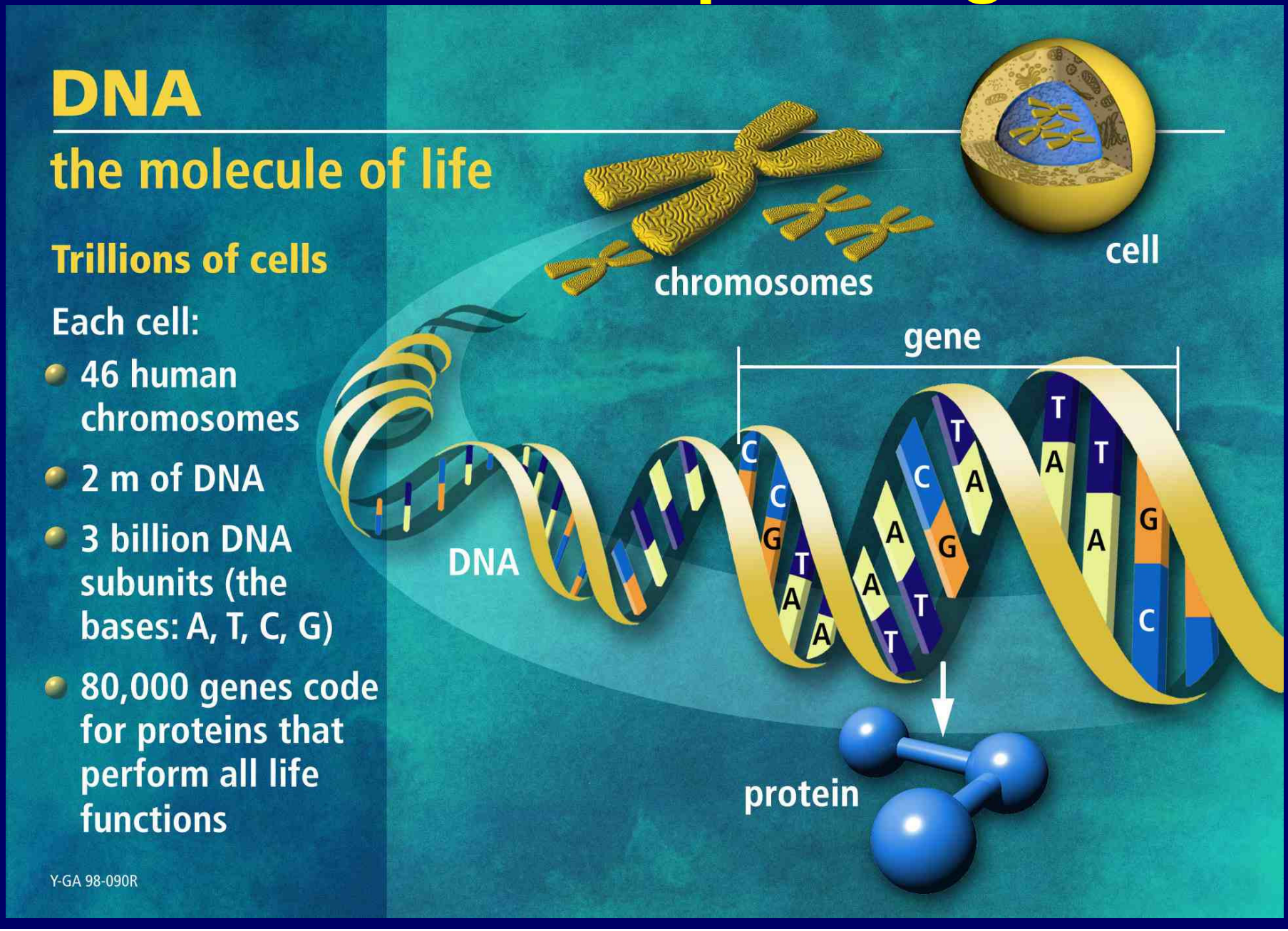


Prof. Dr. Hamdy M. El-Aref

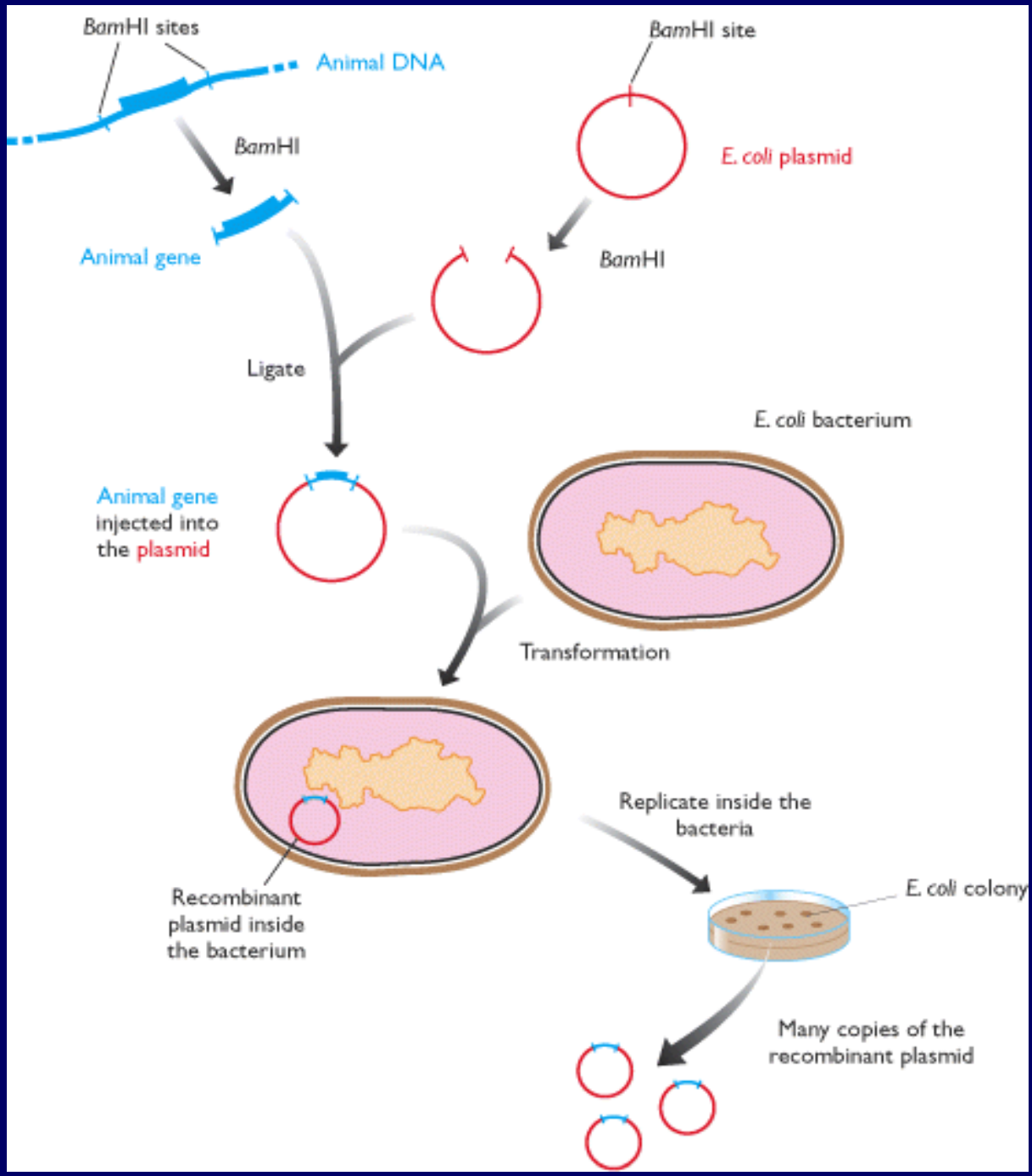
Assiut University, Fac. of Agriculture

Genetics Department

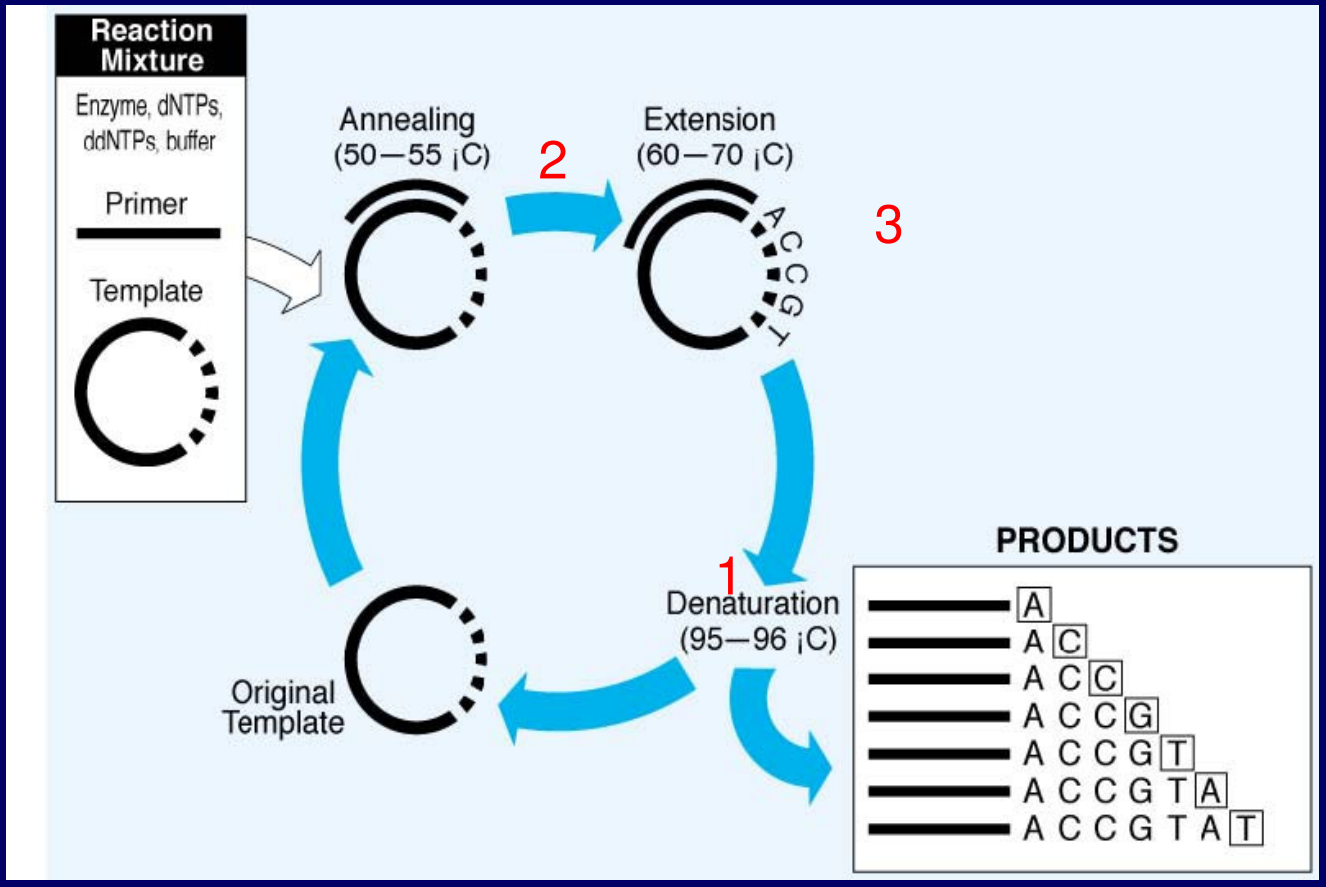
Genome sequencing

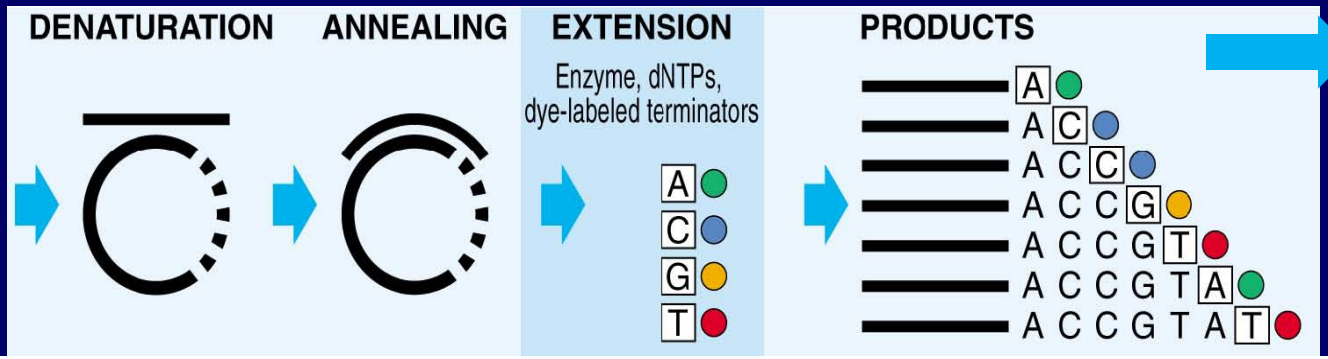


DNA Cloning Overview



Overview: Dideoxy (Sanger) Method

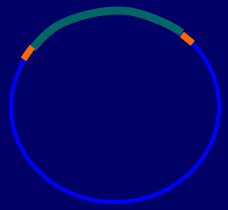




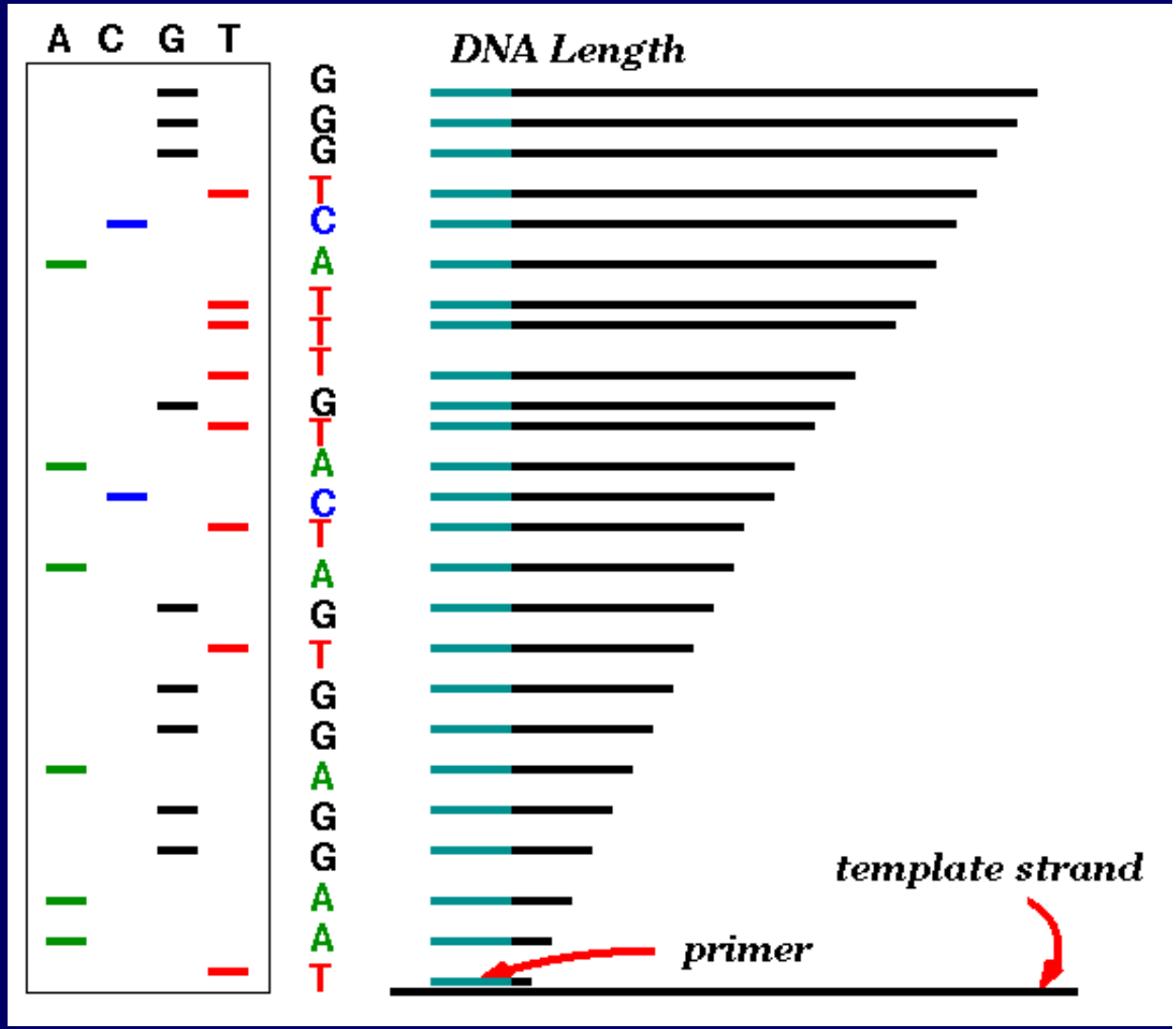
Gel electrophoresis

5

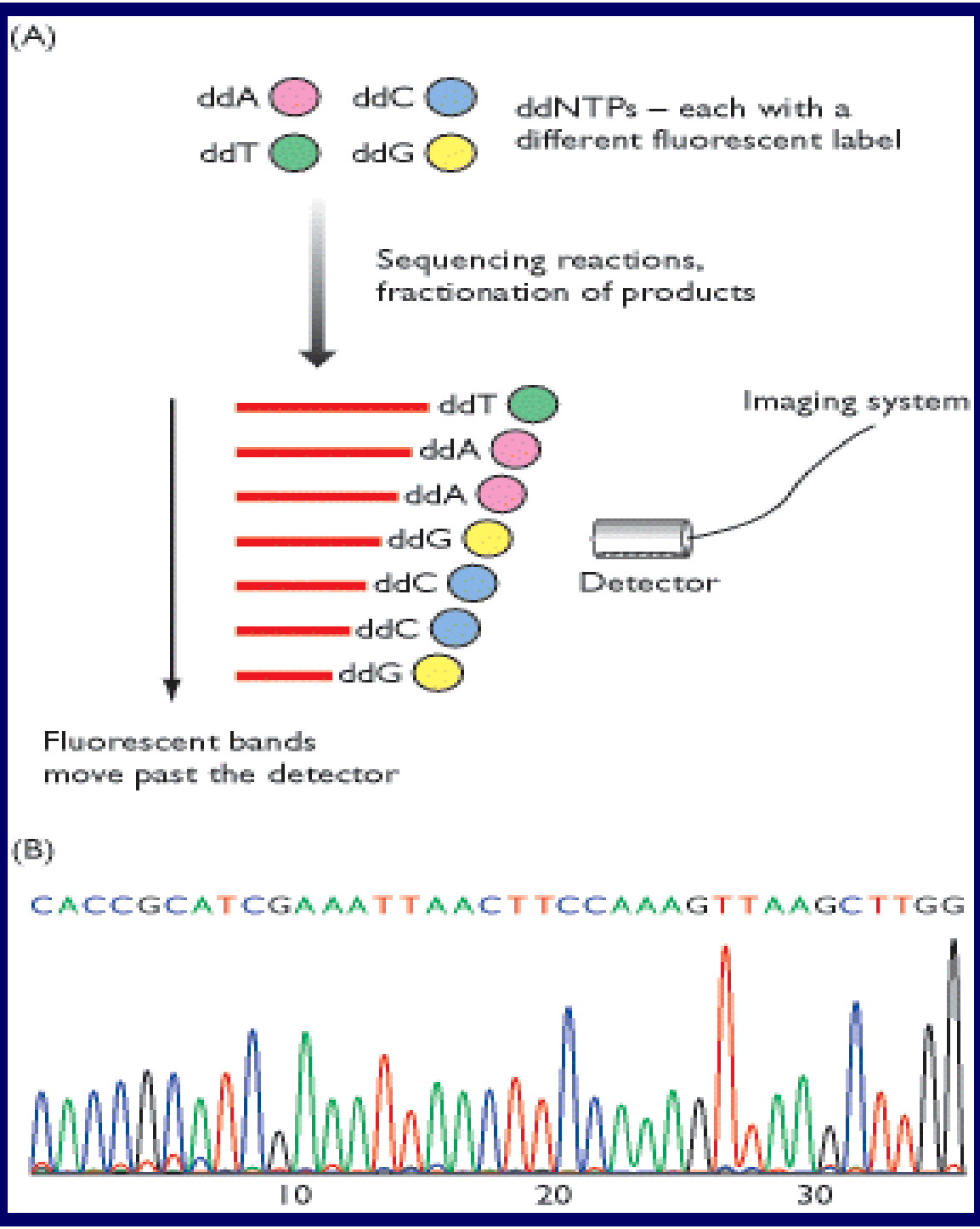
Sanger Method: Generating Read



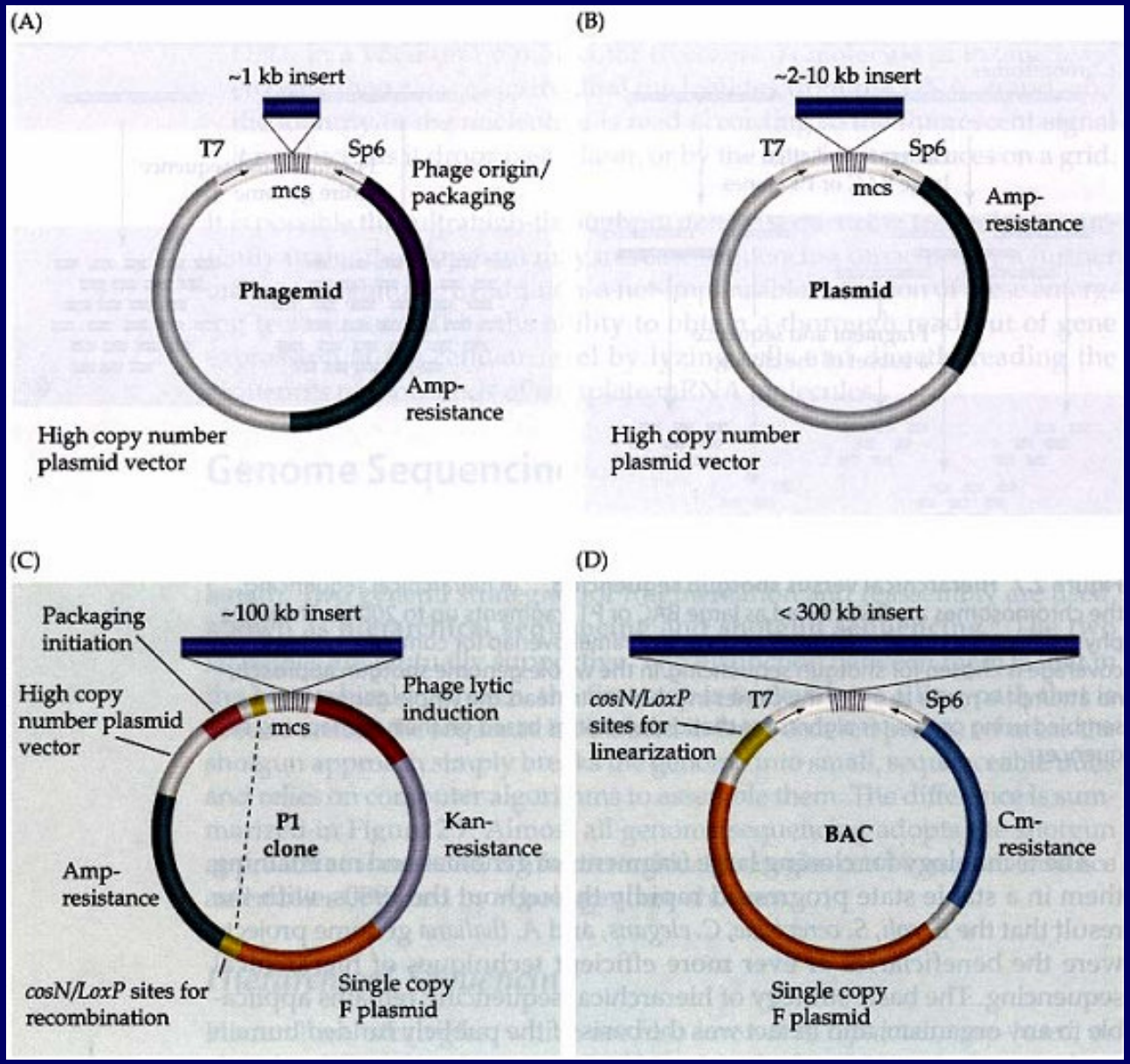
1. Start at primer (restriction site)
2. Grow DNA chain
3. Include ddNTPs
4. Stops reaction at all possible points
5. Separate products by length, using gel electrophoresis



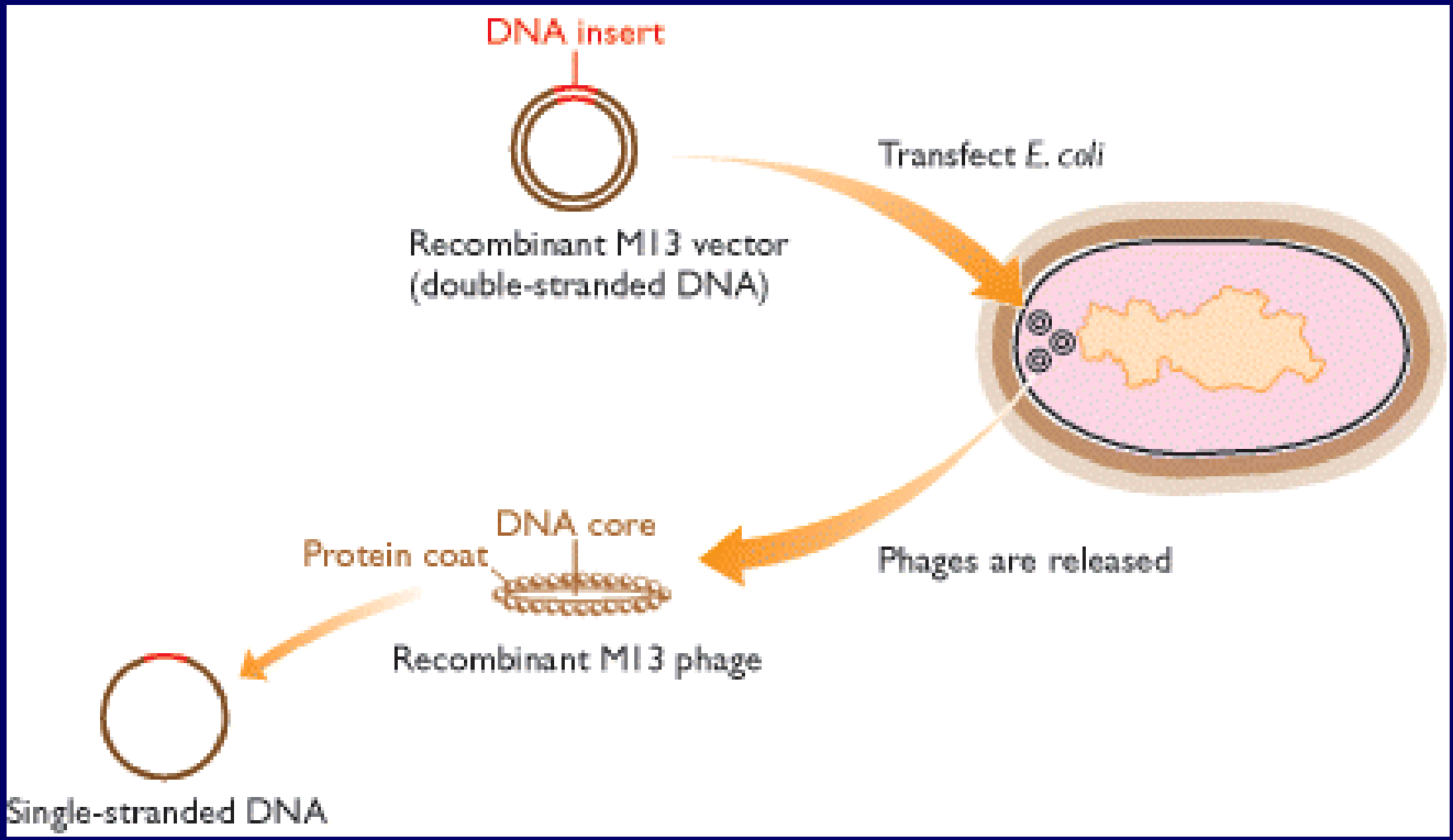
Automatic DNA sequencing



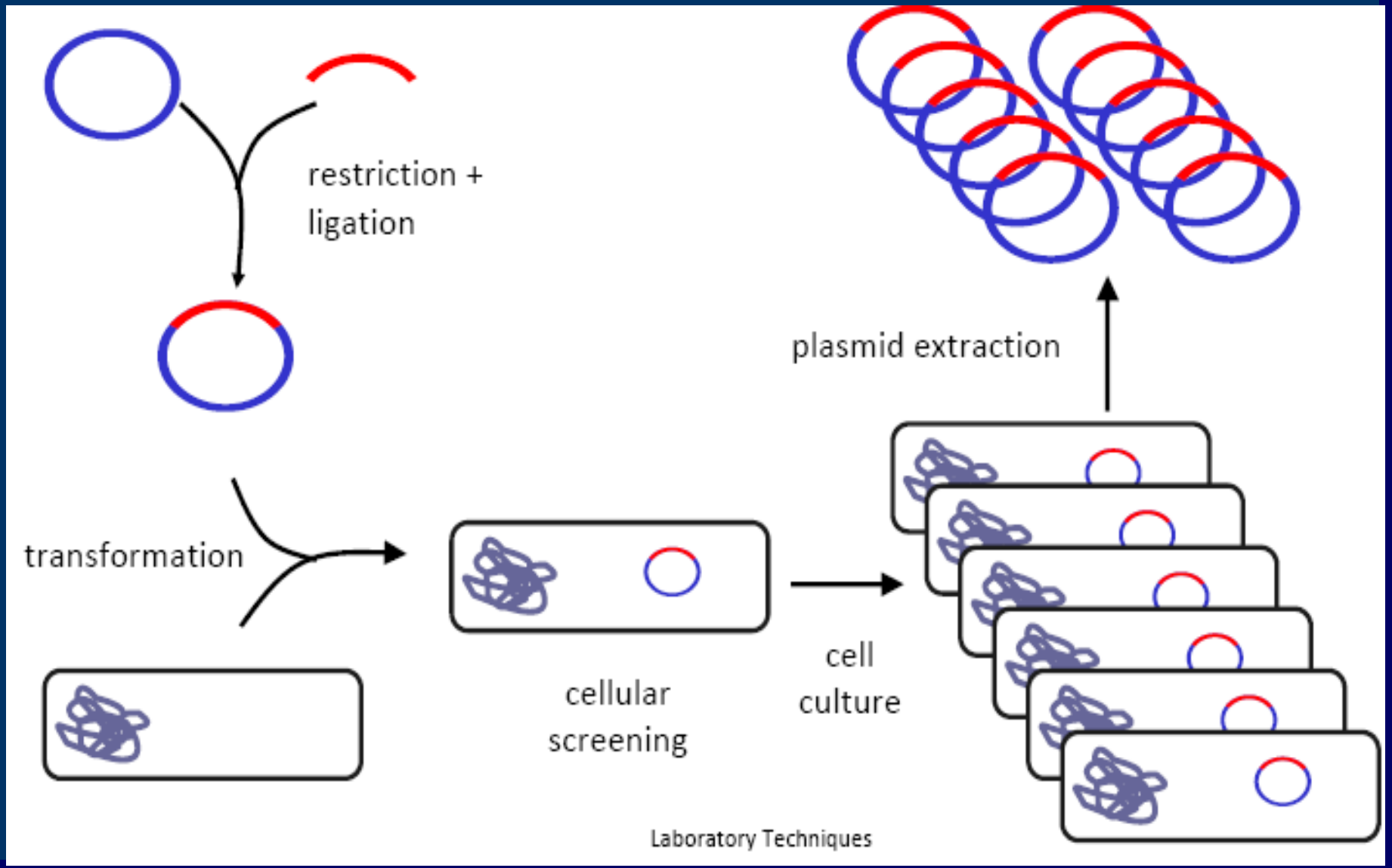
Cloning vectors for genome sequencing



Sequencing vector M13



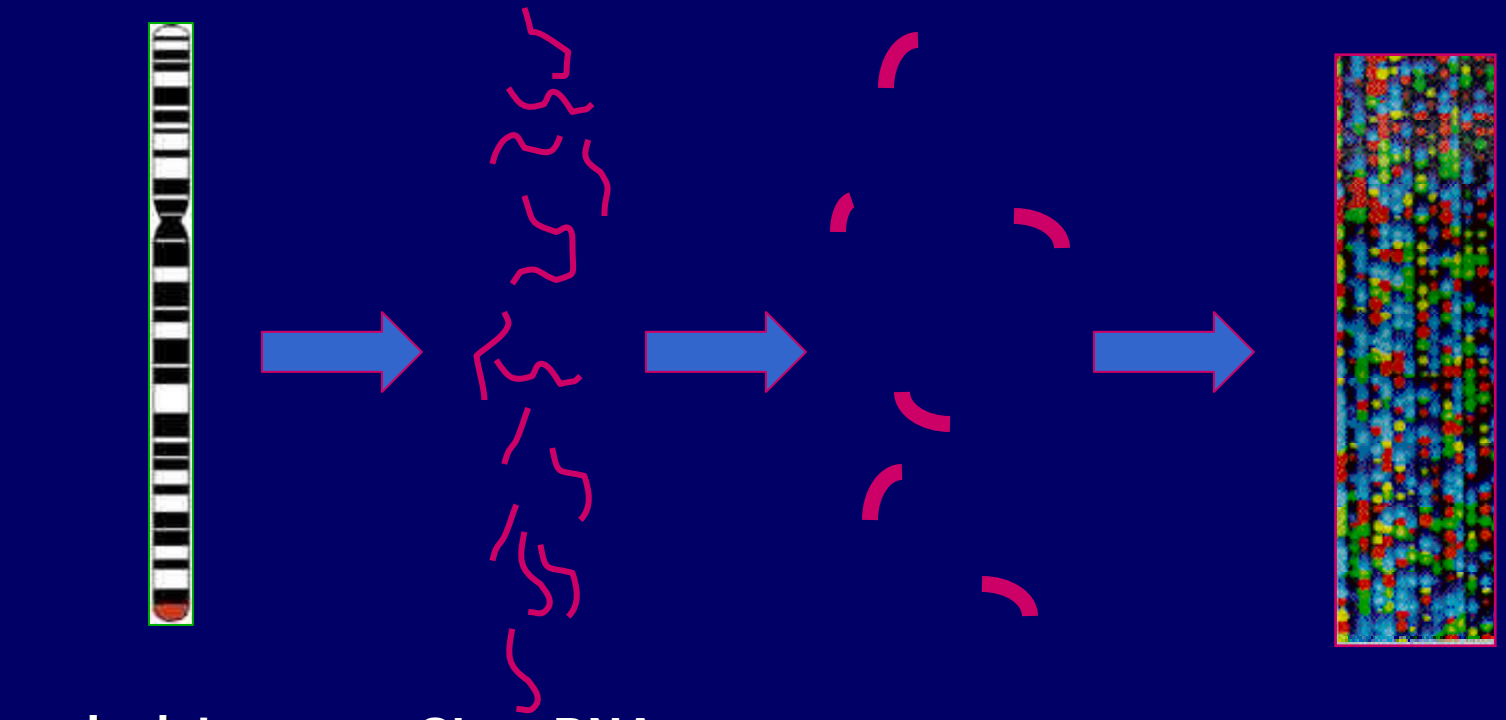
DNA Cloning and Sequencing with BACs



Sizes of inserts in sequencing vectors

Vector	Size (approx.)
P1	100 Kb
YAC	300 -1500 Kb
BAC	70 - 300 Kb)
Cosmid	~ 40 Kb
Plasmid	2 -10 Kb
M13 or Phagmid	~ 1 Kb

Genome Sequencing



Isolate Chromosome

ShearDNA into Fragments

Clone into Seq. Vectors

Sequence

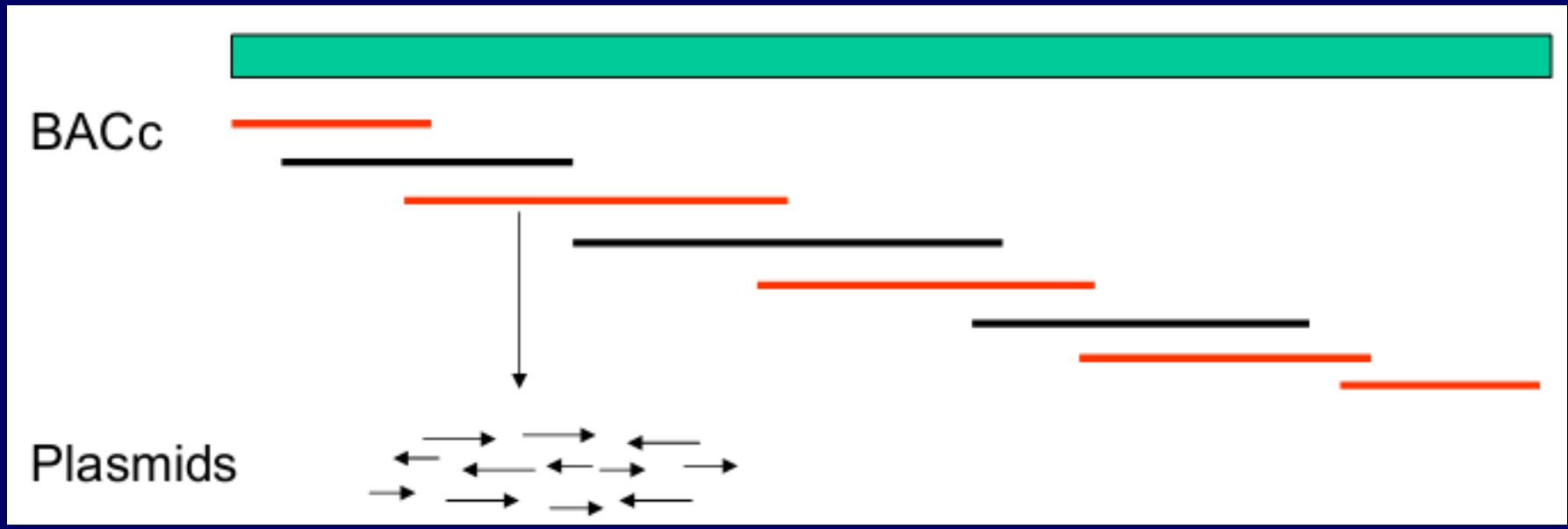
Genome Sequencing Strategies

1. Clone-by-clone or hierarchical sequencing

2. Shotgun sequencing

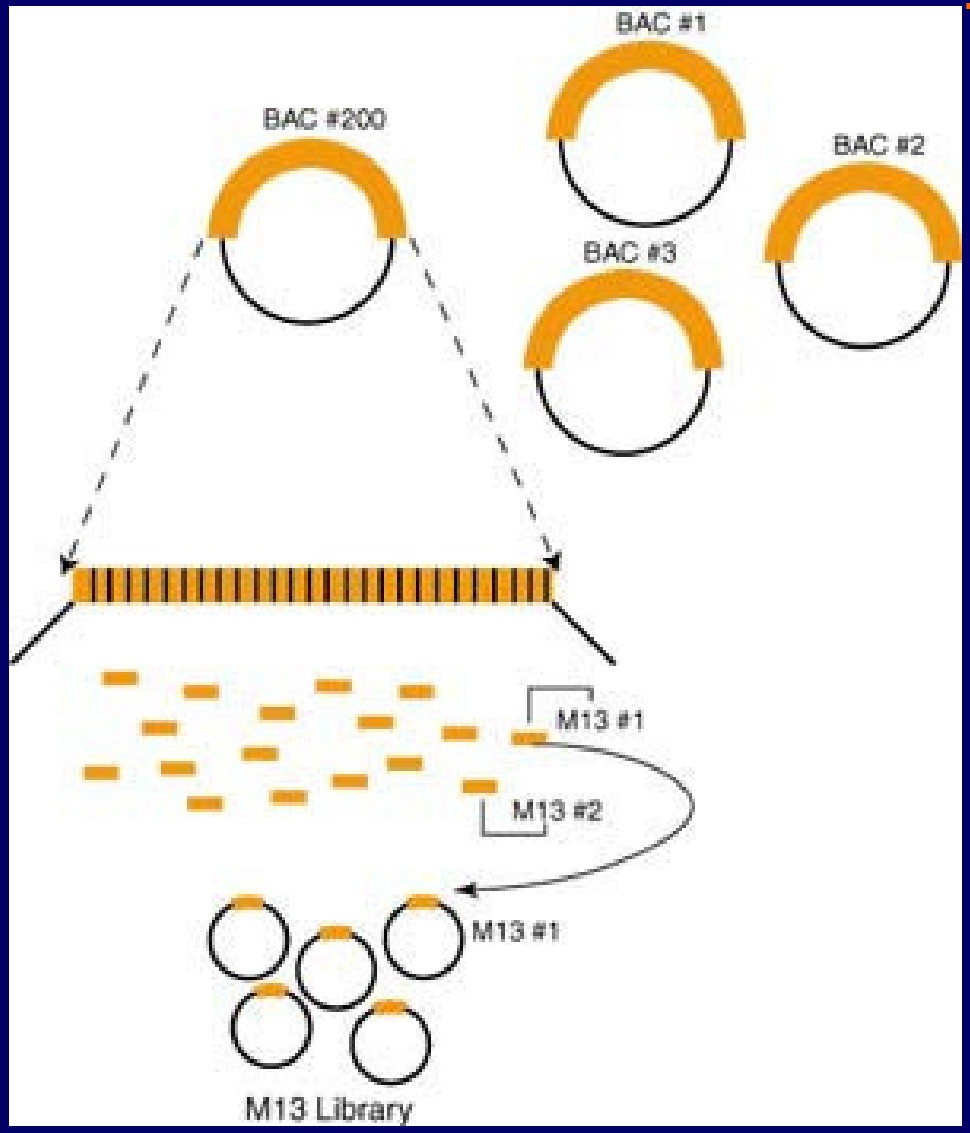
Clone-by-clone or hierarchical sequencing strategy

- After constructing a complete physical map, clone by clone sequencing can be started in any specific region.



BAC to BAC Sequencing Or BAC to Plasmid Sequencing

Whole Genome shotgun Sequencing



This step not needed in shotgun sequencing

Clone-by-clone or hierarchical sequencing strategy

- **Advantages:**

1. Ability to fill gap and re-sequence the regions.
2. Ability to distribute the clones to other labs.
3. Ability to check the produced sequence by restriction enzymes

- **Disadvantages:**

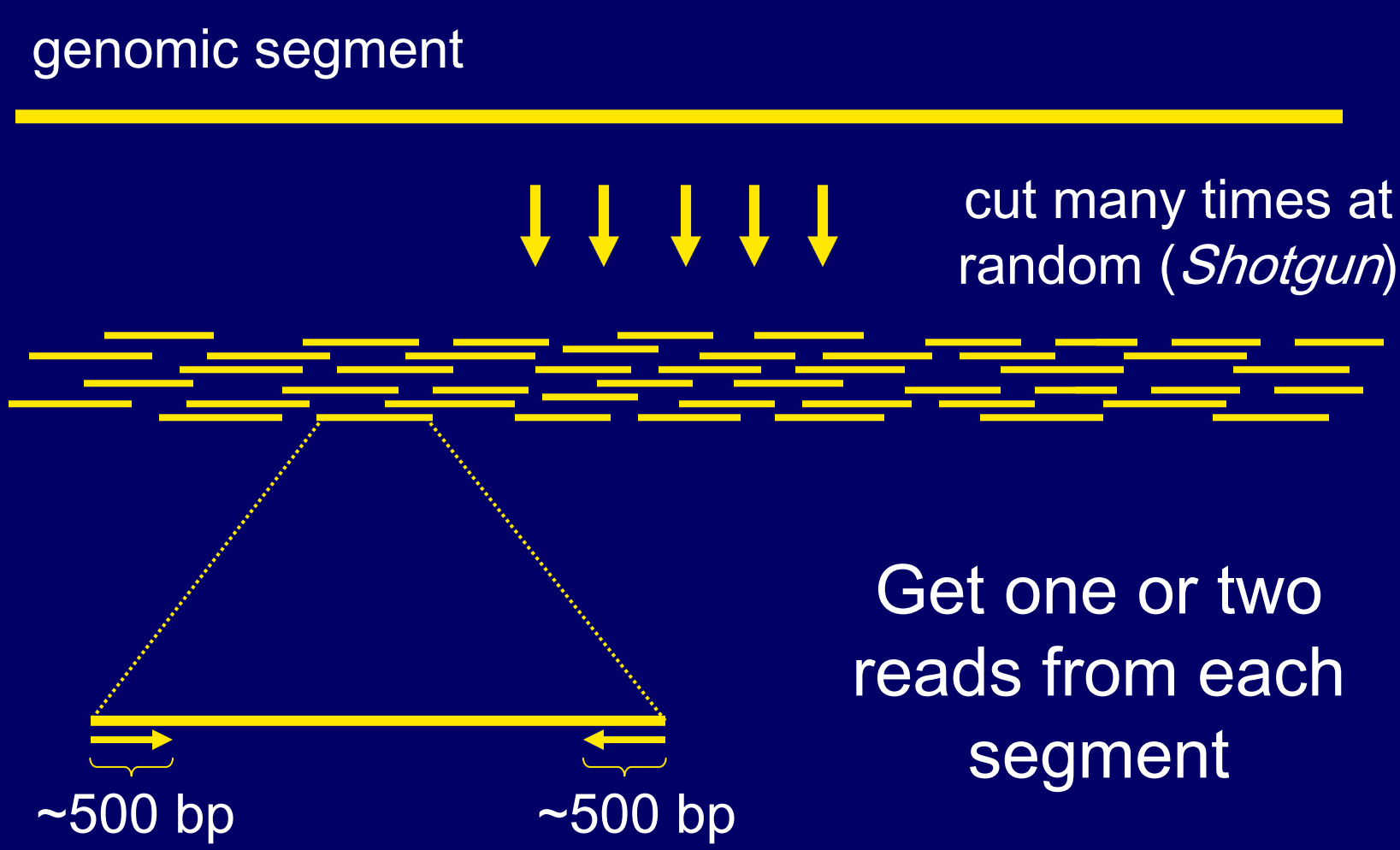
1. Expensive and time-consuming for construction of the physical map.
2. Experienced personnel are required.

VENTER'S SHOTGUN
Whole genome shotgun sequencing

- It is a much faster approach, and enabled researchers to speed up the timetable for sequencing enormously.

- The shotgun method was developed by J. Craig Venter et al., 1996.

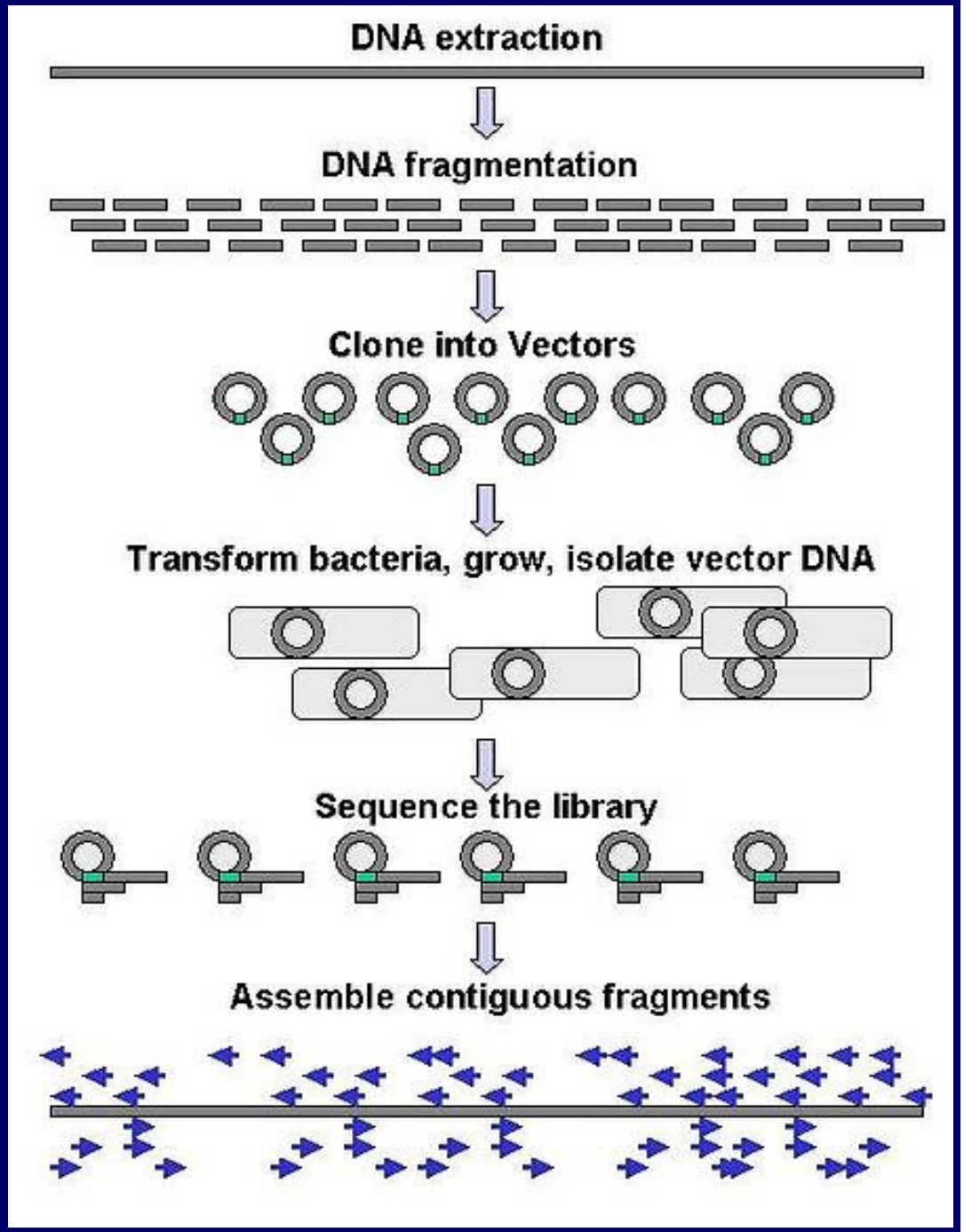
Shotgun Sequencing



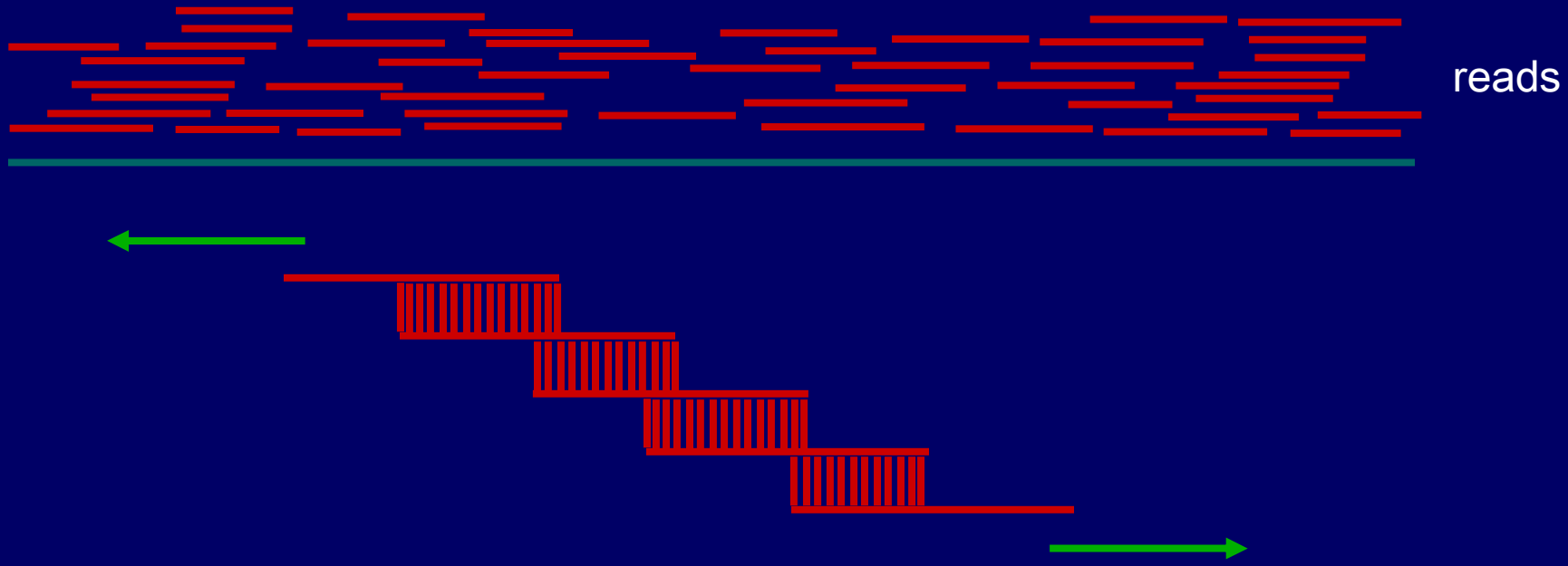
Get one or two reads from each segment

Shotgun Sequencing

- Used to sequence whole genomes
- Steps:
  - DNA is broken up randomly into smaller fragments
  - Dideoxy method produces reads
  - Look for overlap of reads

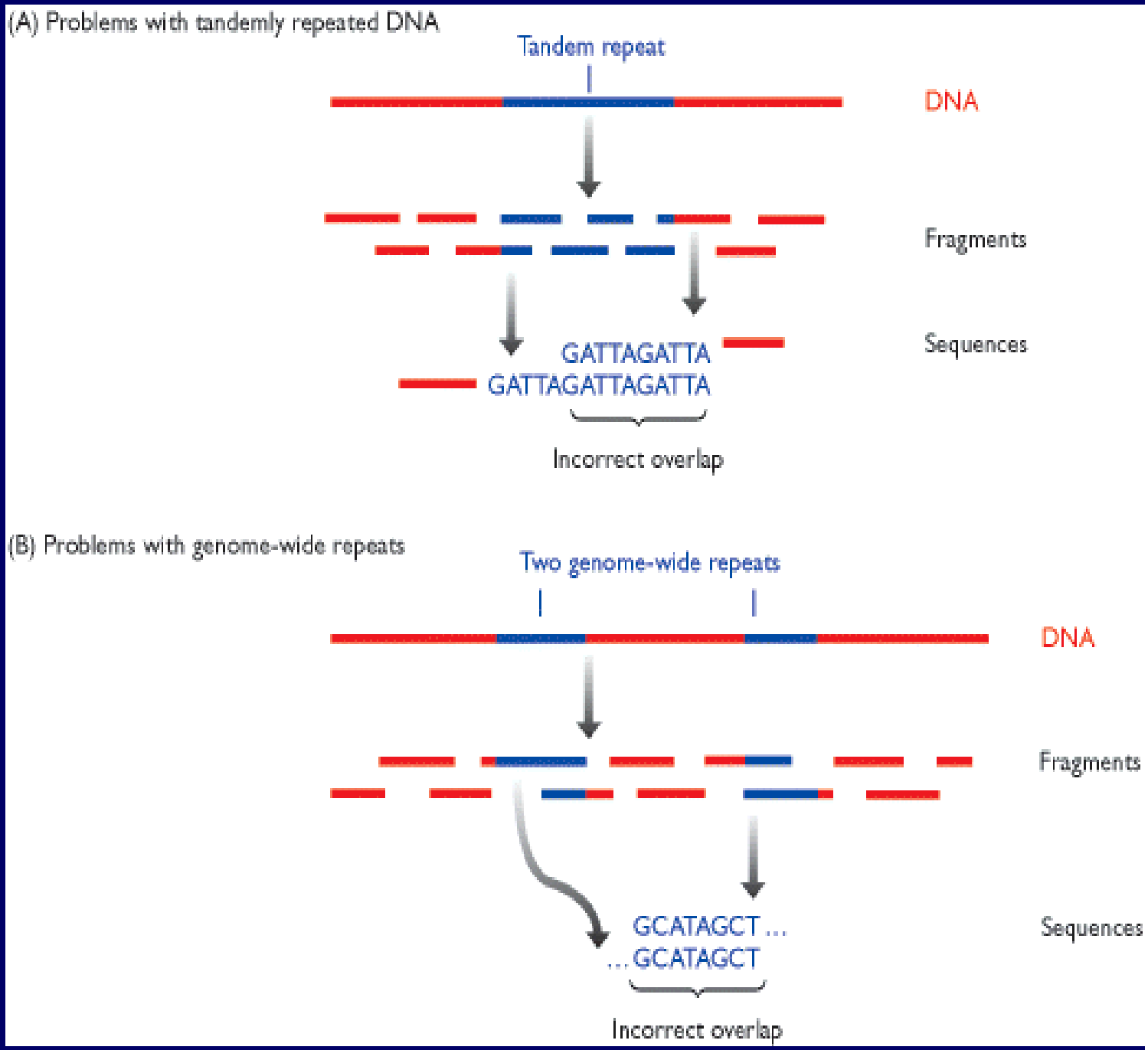


Fragment Assembly



Overlap reads and extend to reconstruct the original genomic region

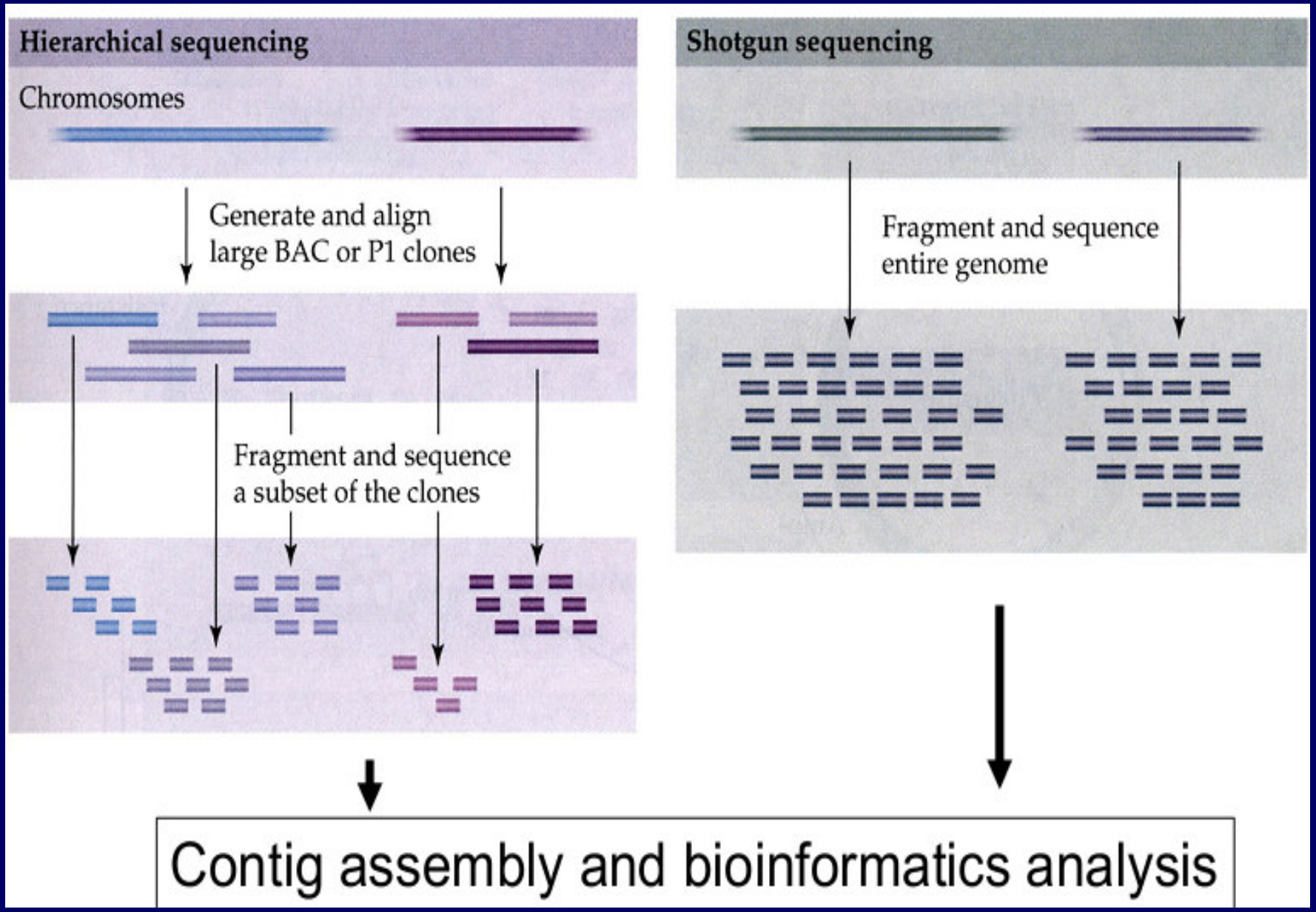
Problems with the shotgun approach



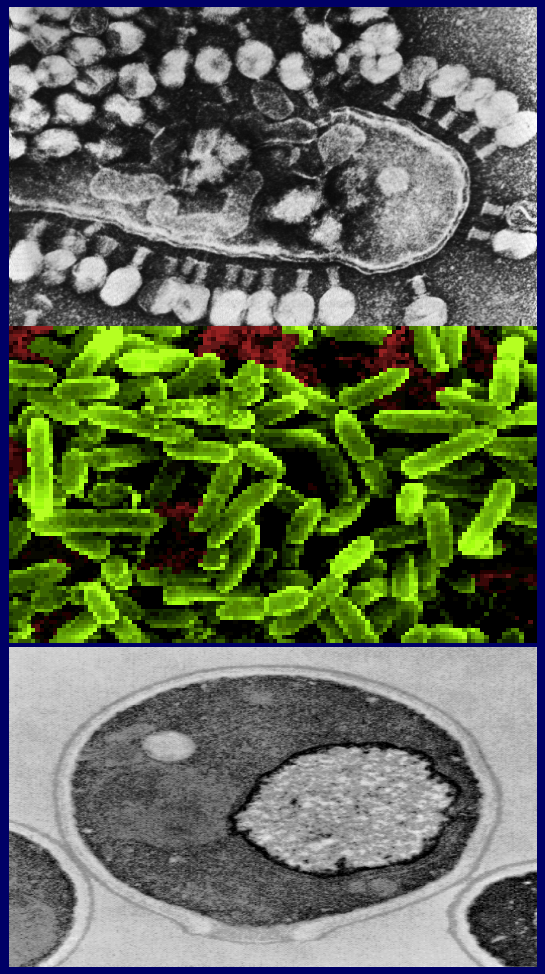
Shotgun Sequencing

- Very efficient process for small-scale (~10 kb) sequencing (preferred method)
- First applied to whole genome sequencing in 1995 (*H. influenzae*)
- Now standard for all prokaryotic genome sequencing projects
- Successfully applied to *D. melanogaster*
- Moderately successful for *H. sapiens and other genotypes …*

Comparison of two sequencing methods



Sequencing Successes



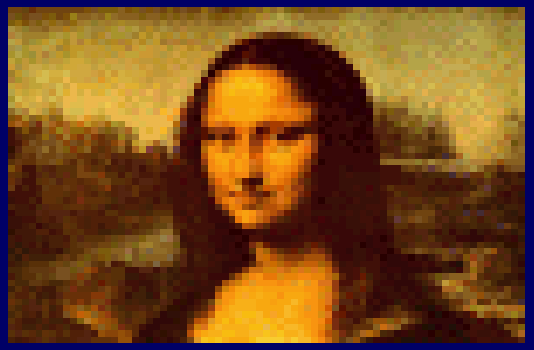
T7 bacteriophage
completed in 1983
39,937 bp, 59 coded proteins

Escherichia coli
completed in 1998
4,639,221 bp, 4293 ORFs

Sacchoromyces cerevisae
completed in 1996
12,069,252 bp, 5800 genes

Sequencing Successes

Drosophila melanogaster
completed in 2000
116,117,226 bp, 13,601 genes



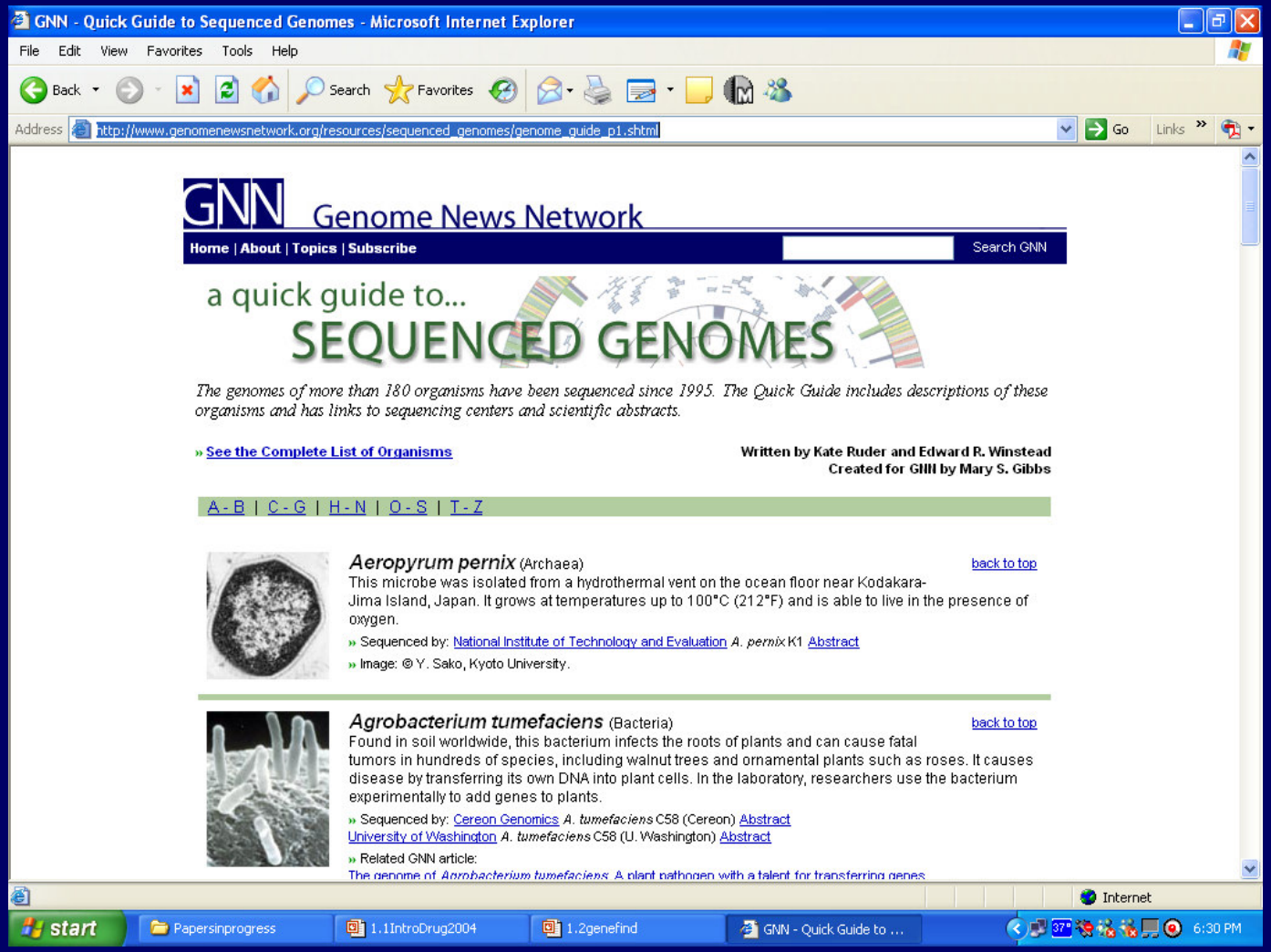
Homo sapiens
completed in 2003
3,201,762,515 bp, 31,780 genes

Oryza Sativa (Rice)
430 million base
60,000 genes

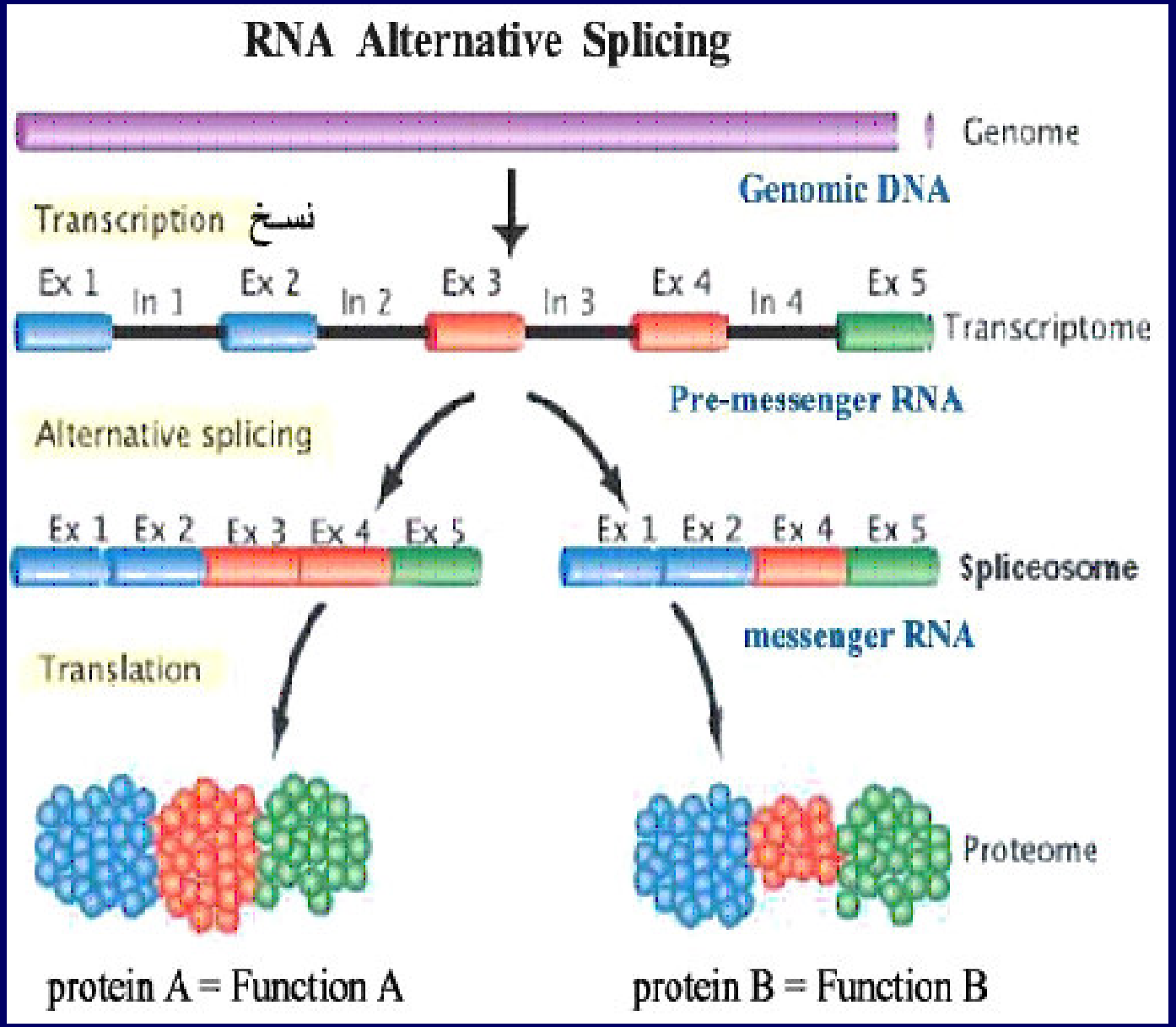
Genomes to Date

- 8 vertebrates (human, mouse, rat, fugu, zebrafish)
- 2 plants (arabadopsis, rice)
- 2 insects (fruit fly, mosquito)
- 2 nematodes (*C. elegans, C. briggsae*)
- 1 sea squirt
- 4 parasites (plasmodium, guillardia)
- 4 fungi (*S. cerevisae, S. pombe*)
- 200 bacteria and archebacteria
- 1900+ viruses

Sequenced Genomes

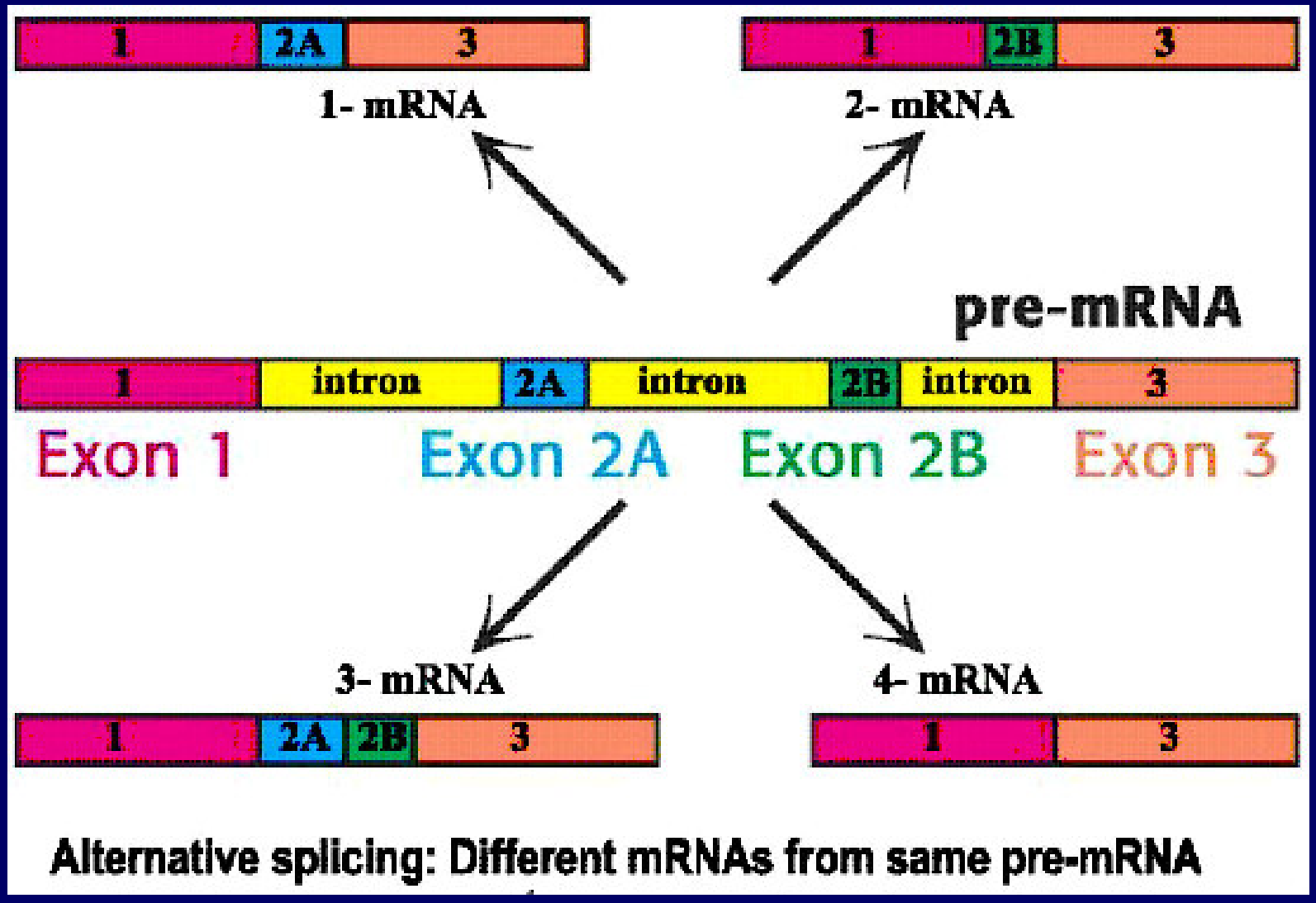


http://www.genomenewsnetwork.org/

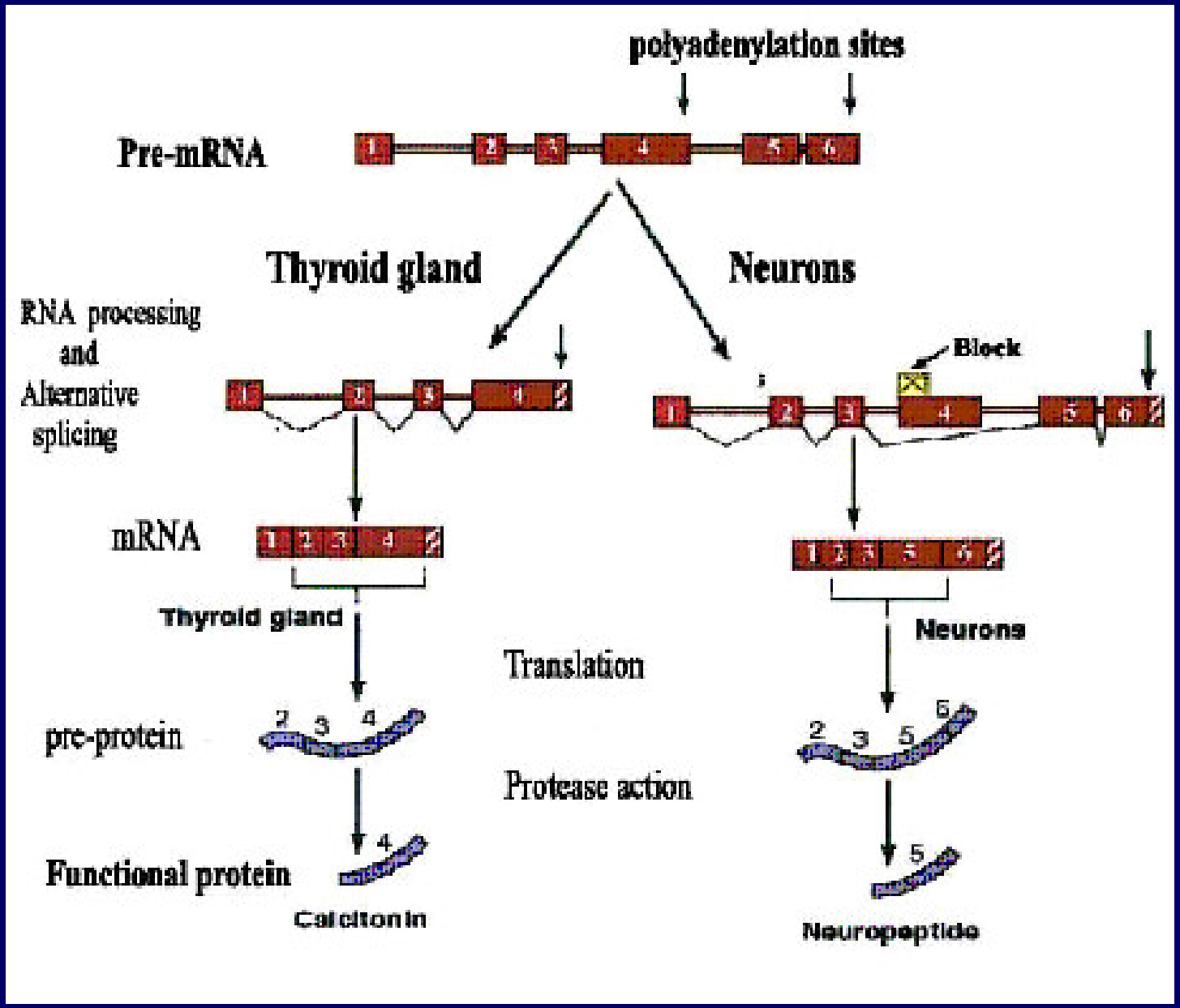


RNA Alternative Splicing and Gene Function in Eukaryoutes

Amylase = 2
Tropomiosine = 10
Troponin-T = 64



Alternative splicing: Different mRNAs from same pre-mRNA



calcitonin gene

Thank you

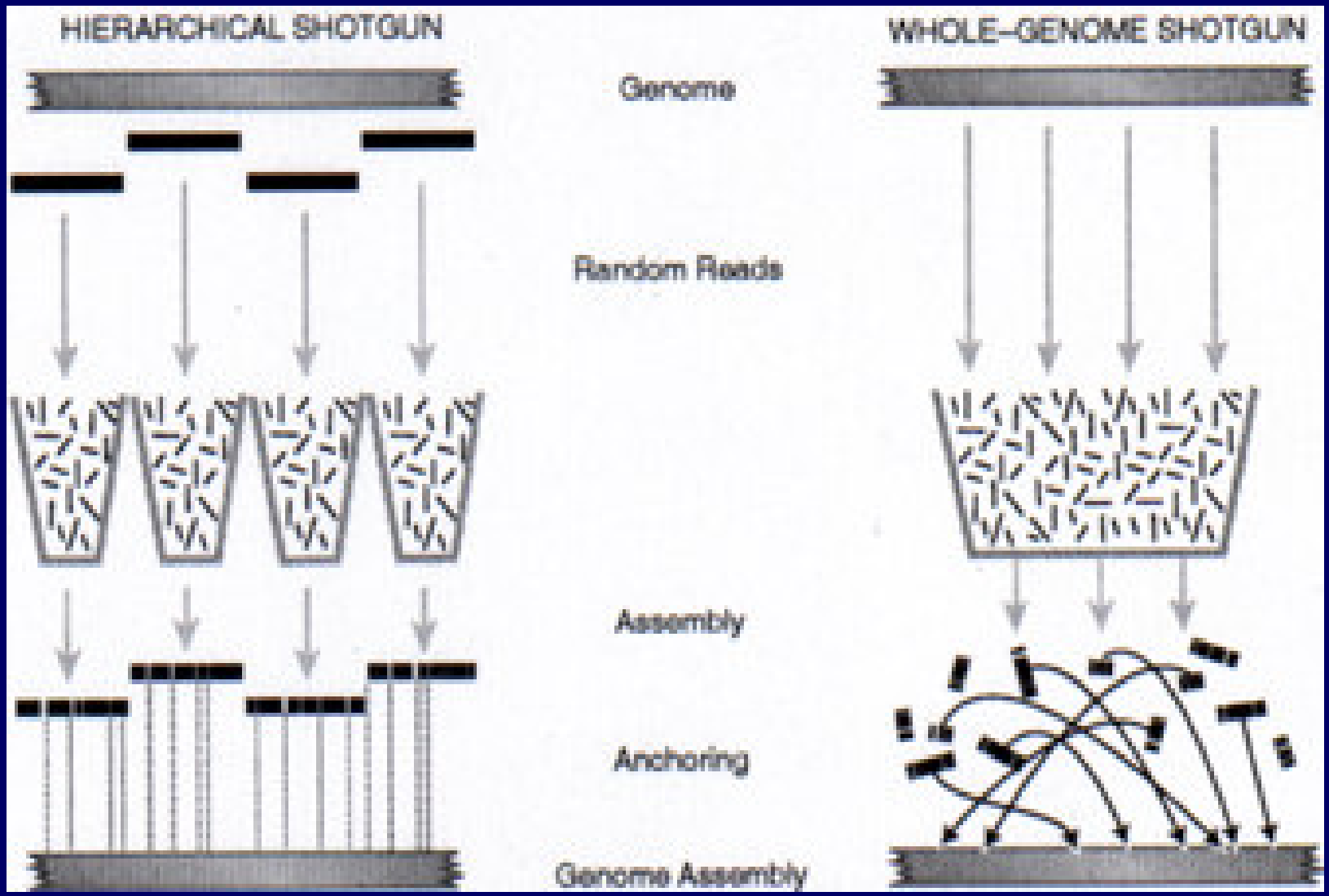
كلية الزراعة

السلام عليكم

Prof. Dr. Hamdy El-Aref

Sequencing strategies

- Whole genome



Rapid re-sequencing of human Ad1: Time trial

- Have sequence of Ad 1.
- In theory, have a minimally tiled set of PCR primers to cover entire 36,001 base genome.
- In theory, have a minimally tiled set of sequencing primers as well.
- Want draft sequence in a minimal time, including primer delivery from a vendor.
- In practice design two parallel sets of minimally tiled PCR primers and amplify two sets.
- In practice, assume 750 base reads--> 48 primers, one direction.
- Compare with consensus: Determine accuracy, timing and evaluate operation.