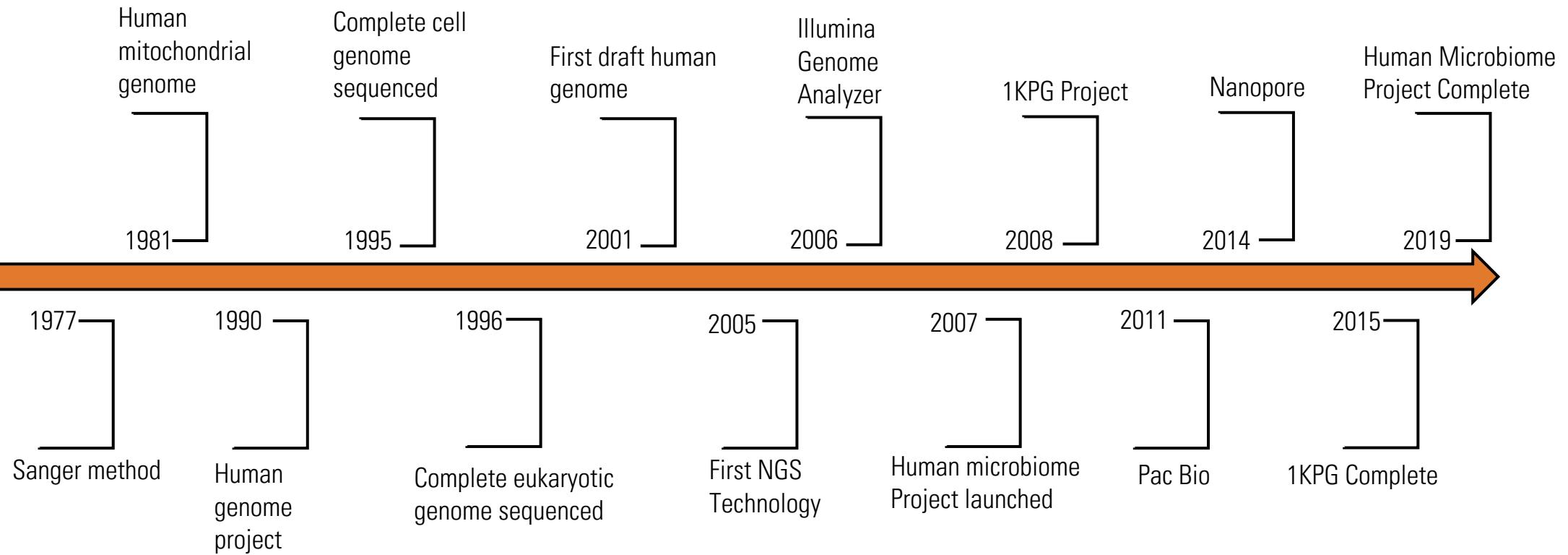


The background on the left side of the slide features a complex, abstract geometric pattern. It consists of numerous thin, parallel lines in various colors—blue, pink, orange, and black—arranged in a grid-like structure that creates a sense of depth and motion. The lines are slightly curved and overlap each other, forming a textured, woven appearance.

"KNOWLEDGE OF SEQUENCES COULD CONTRIBUTE
MUCH TO OUR UNDERSTANDING OF LIVING
MATTER" – FREDRICK SANGER

NEXT GENERATION SEQUENCING

FROM DNA TO SEQUENCING

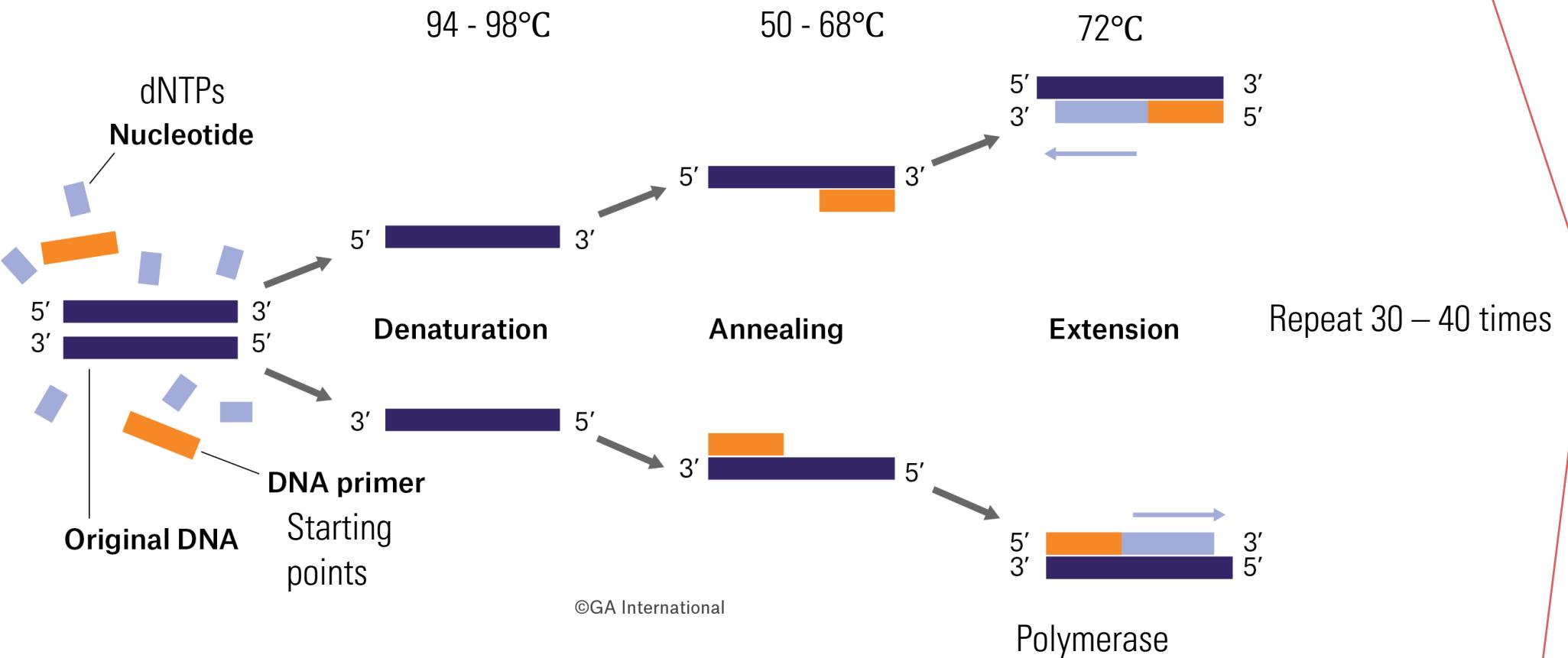


The background features a complex network of thin, red, intersecting lines forming a grid-like pattern, creating a sense of depth and geometric complexity.

POLYMERASE CHAIN REACTION

AMPLIFYING DNA SO WE HAVE ENOUGH!

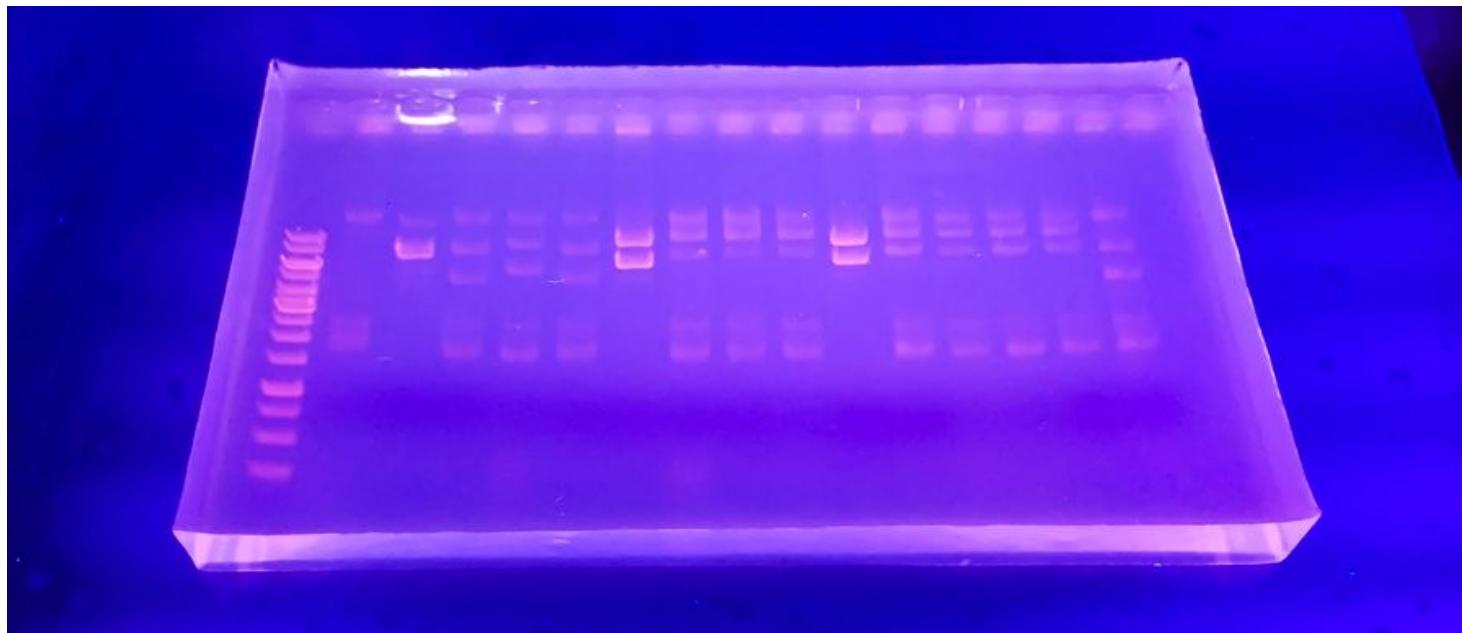
POLYMERASE CHAIN REACTION



POLYMERASE CHAIN REACTION

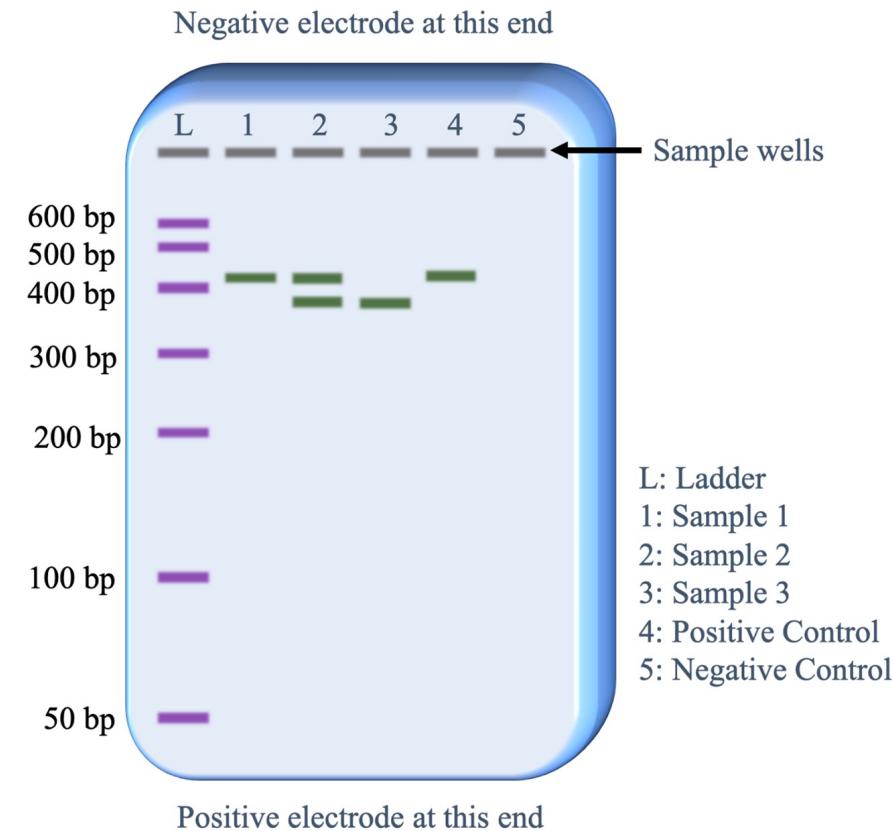


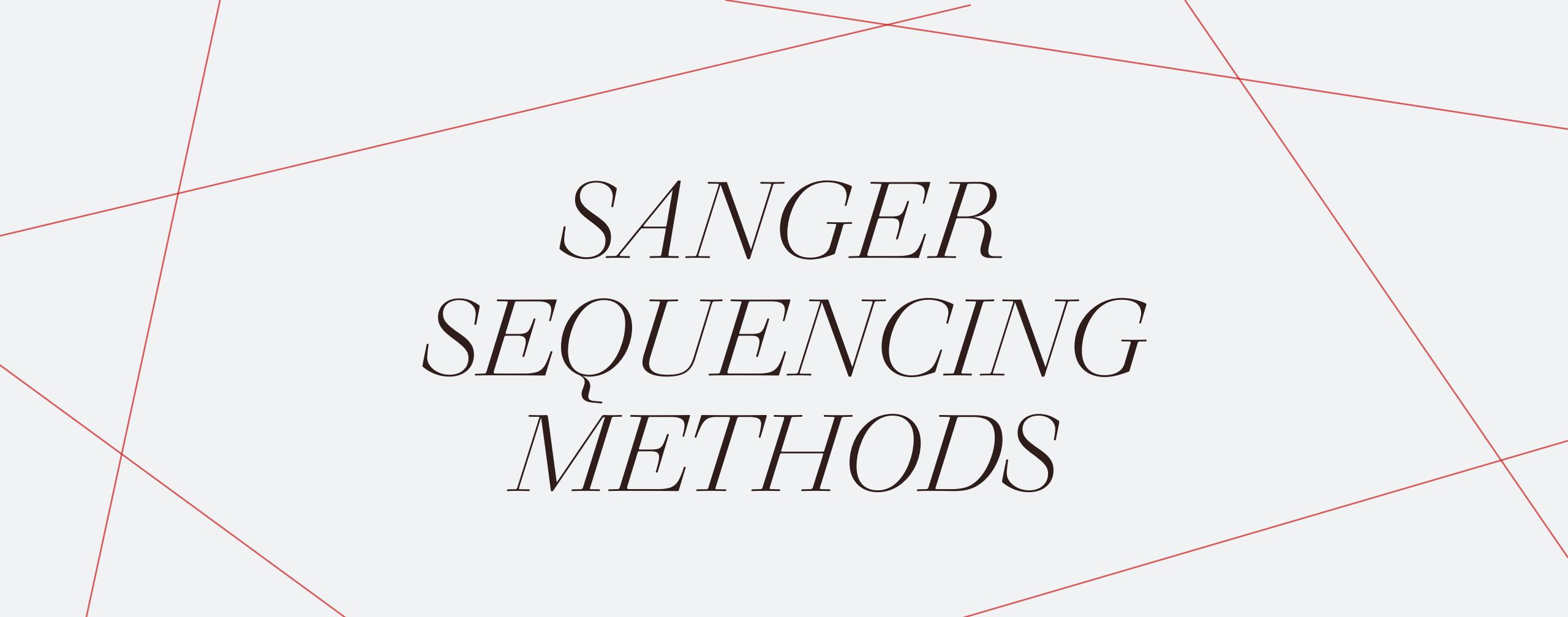
POLYMERASE CHAIN REACTION



GEL ELECTROPHORESIS

- Load samples into wells with dye in agarose gel
- Load ladder for a size metric
- Phosphate backbone of DNA is attracted to positive node
- Shorter strands will “run” faster along the gel
- Longer strands will “run” much slower

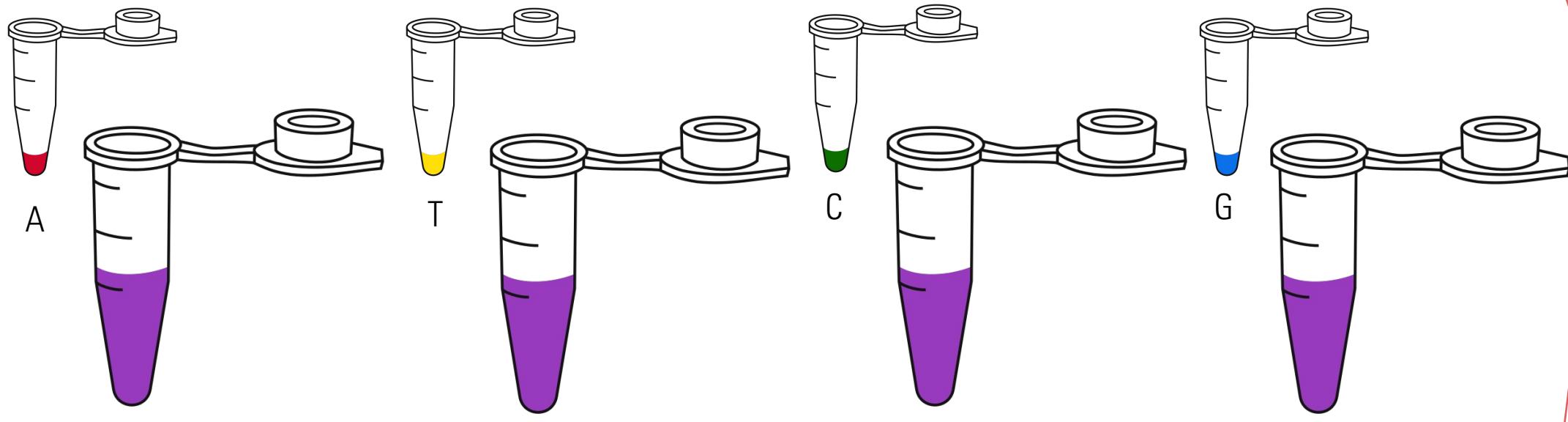




SANGER
SEQUENCING
METHODS

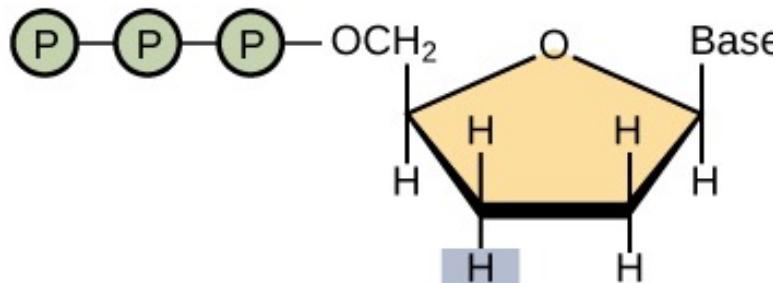
SANGER SEQUENCING

1. Primer
2. DNA Polymerase
3. dNTPS (A, C, G, T)
4. ddNTPs

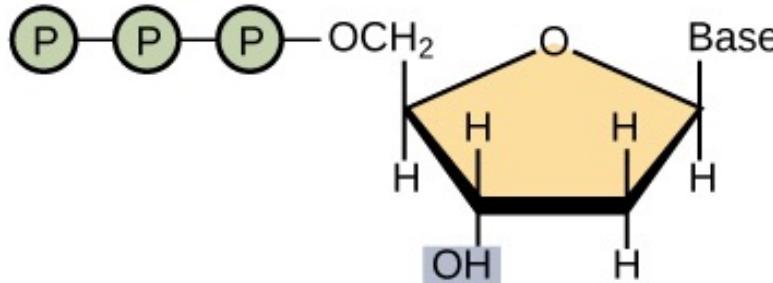


Samples 1 - 4

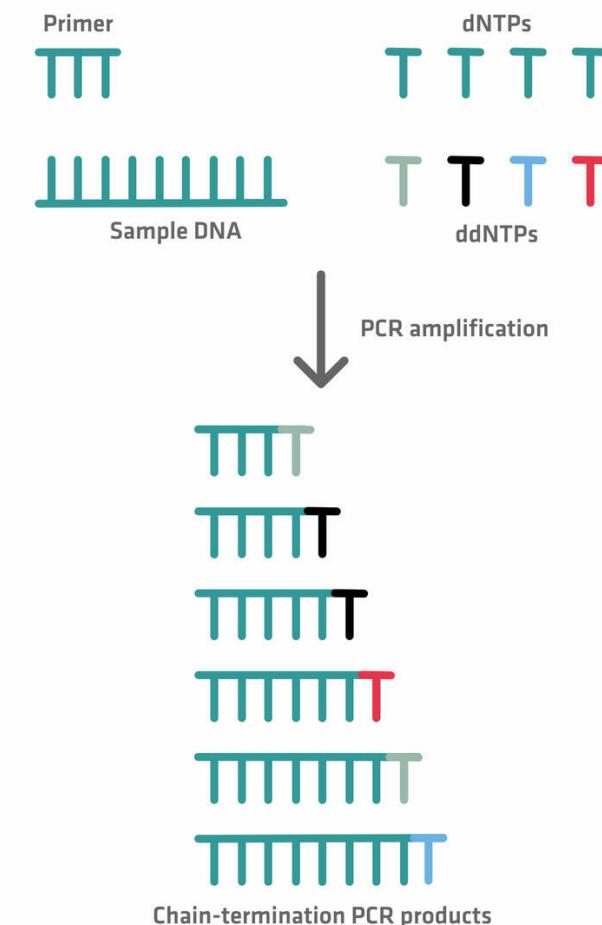
SANGER SEQUENCING



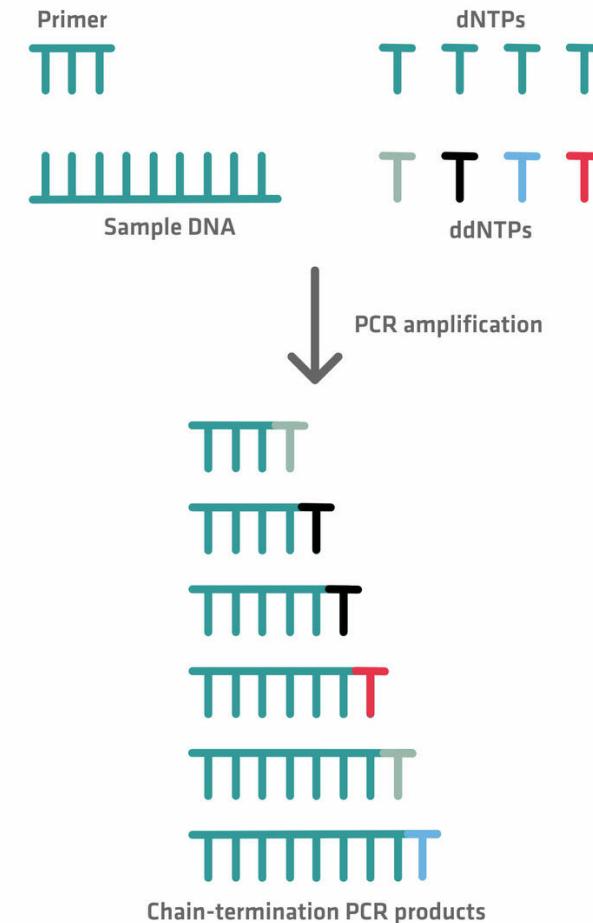
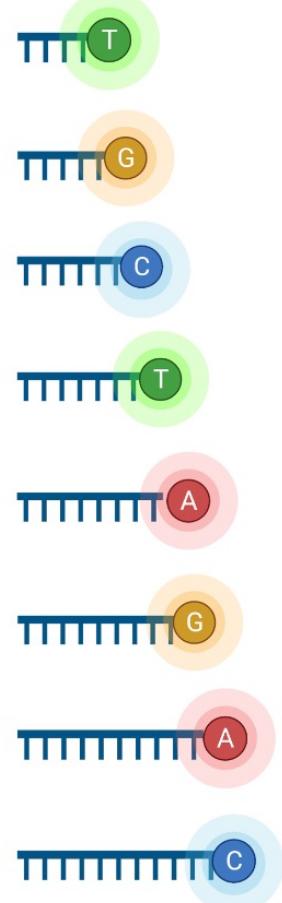
Dideoxynucleotide (ddNTP)



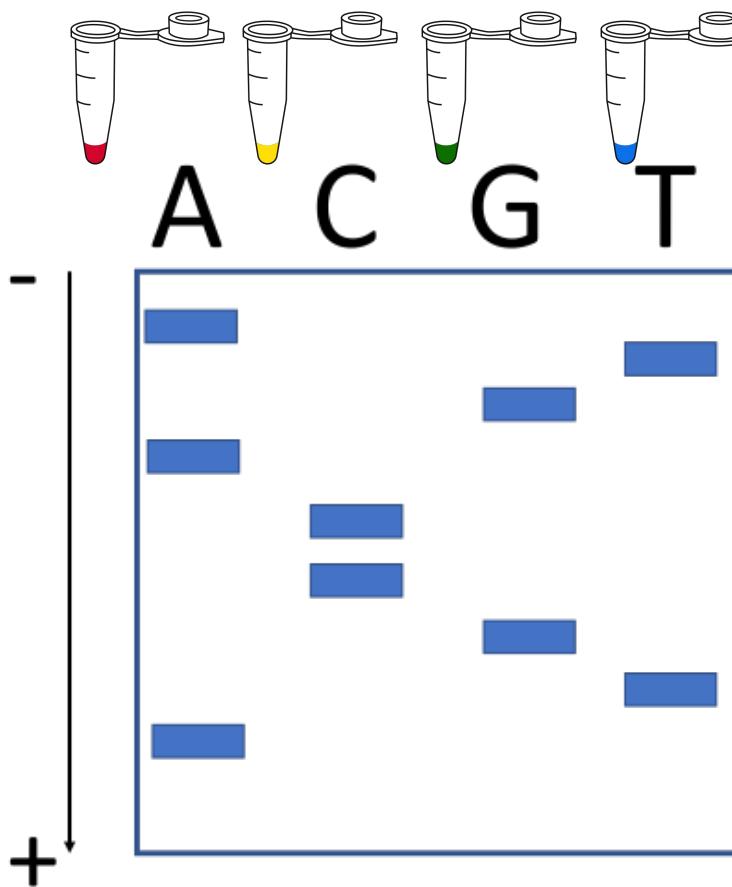
Deoxynucleotide (dNTP)



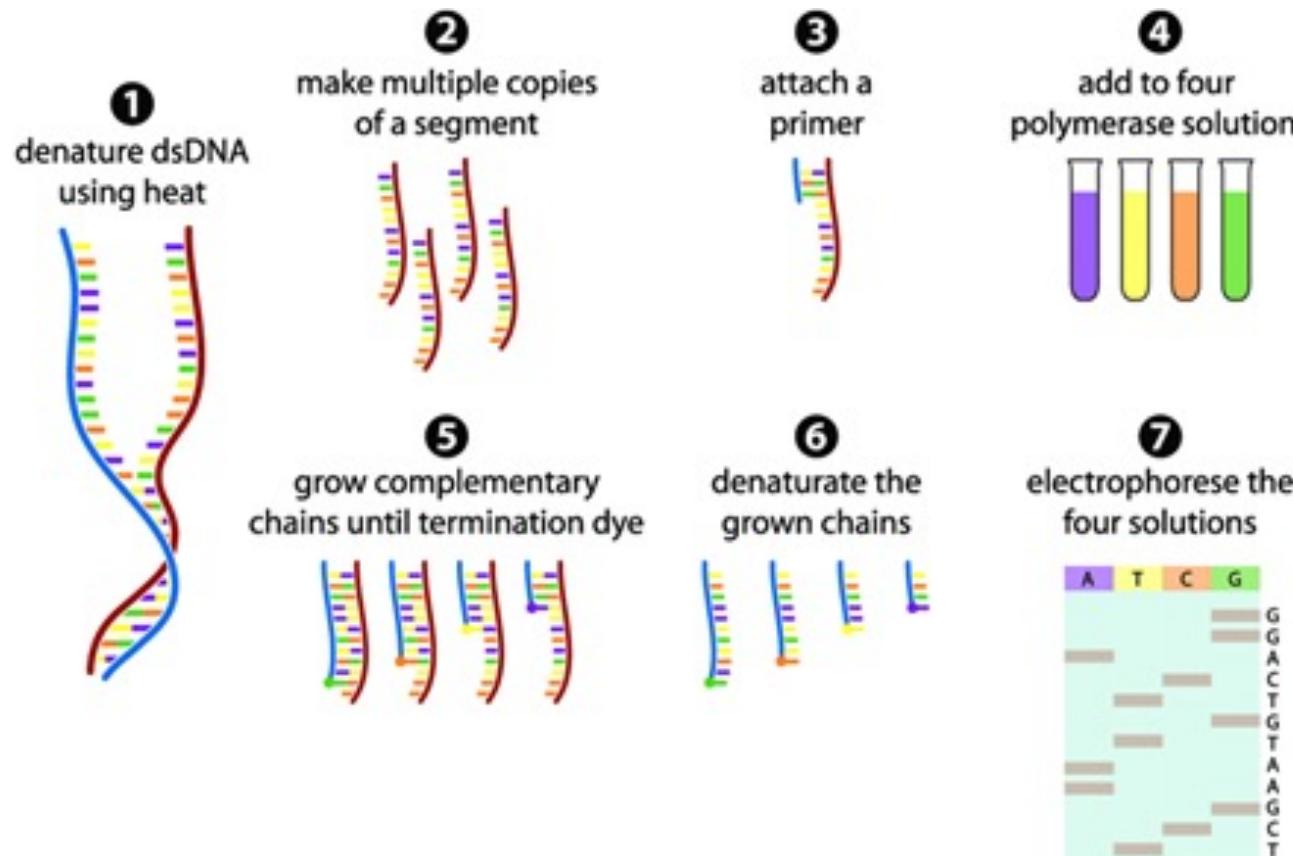
SANGER SEQUENCING



SANGER SEQUENCING



SANGER SEQUENCING

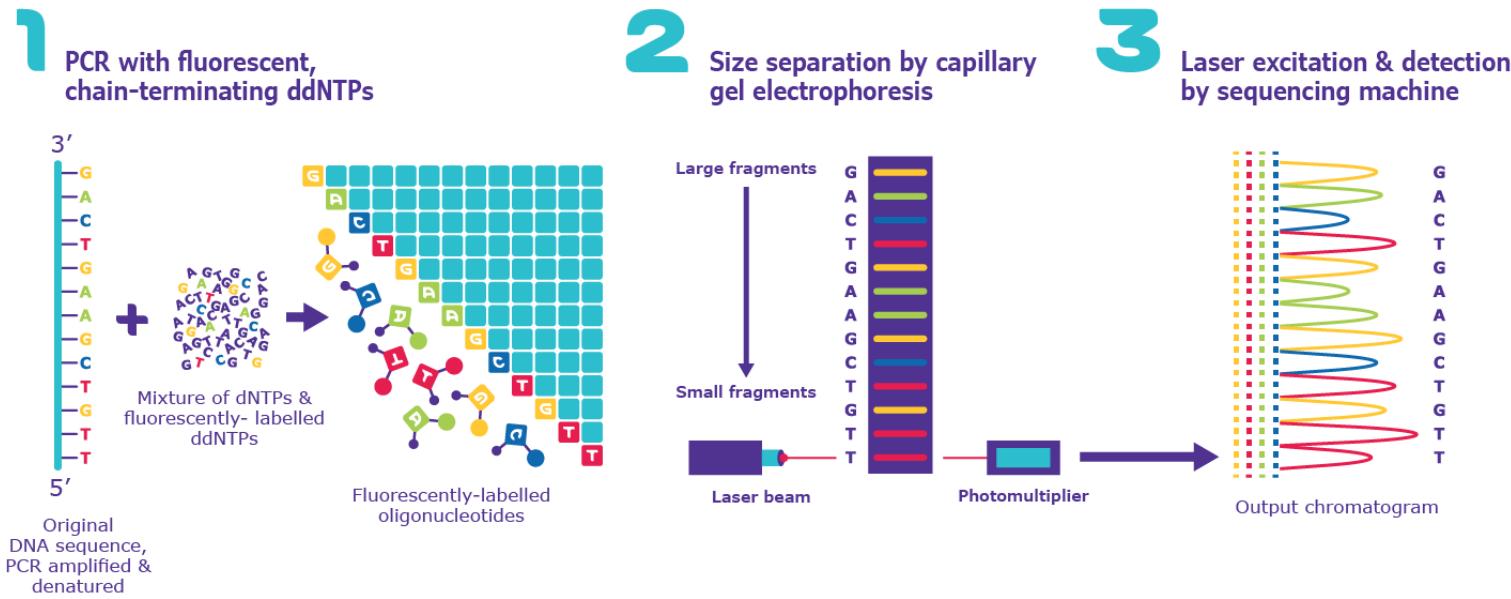


- Time consuming
- No automation

SANGER SEQUENCING - 1987

Leroy Hood and Michael Hunkapiller Improve Sanger Method

1. DNA labeled with fluorescent dyes instead of radioactive molecules
 2. Data acquisition/analysis made easier with computers



SANGER SEQUENCING - 1987



CYCLE SEQUENCING - 1989

Sanger Sequencing

1. DNA Polymerase
2. Much more DNA than Polymerase
3. A bunch of wasted DNA fragments

Cycle Sequencing

1. **Taq Polymerase – Vincent Murray**
2. Taq can withstand high temps
3. Can use our normal PCR process!



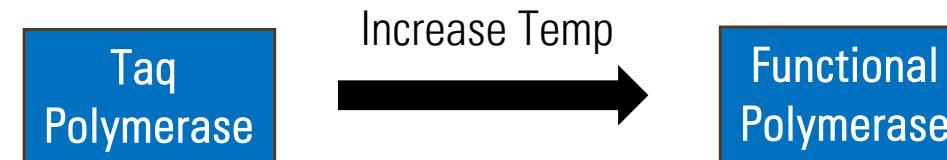
CYCLE SEQUENCING - 1989

Sanger Sequencing

1. DNA Polymerase
2. Much more DNA than Polymerase
3. A bunch of wasted DNA fragments

Cycle Sequencing

1. **Taq Polymerase – Vincent Murray**
2. Taq can withstand high temps
3. Can use our normal PCR process!



CYCLE SEQUENCING - 1989

Sanger Sequencing

1. DNA Polymerase
2. Much more DNA than Polymerase
3. A bunch of wasted DNA fragments

Cycle Sequencing

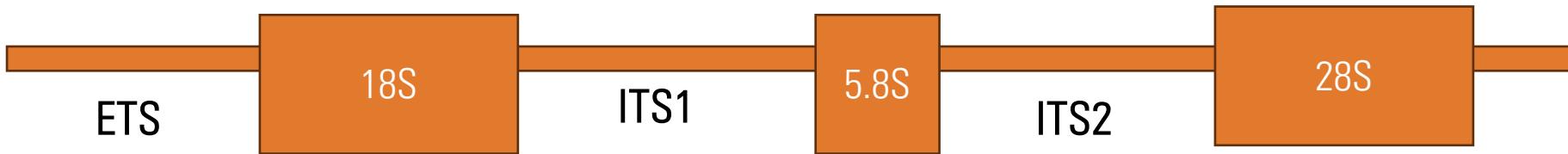
1. **Taq Polymerase – Vincent Murray**
2. Taq can withstand high temps
3. Can use our normal PCR process!

Two Major Improvements

1. Higher Fluorescent Signal
2. Less DNA needed per reaction

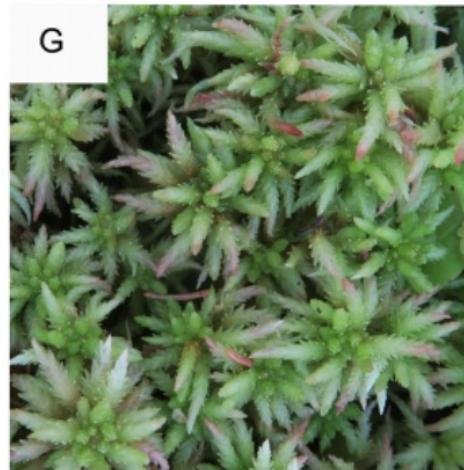
CYCLE SEQUENCING SCALE

1. Fingerprint regions of the genome
2. Low-copy nuclear genes
3. Target genes of interest

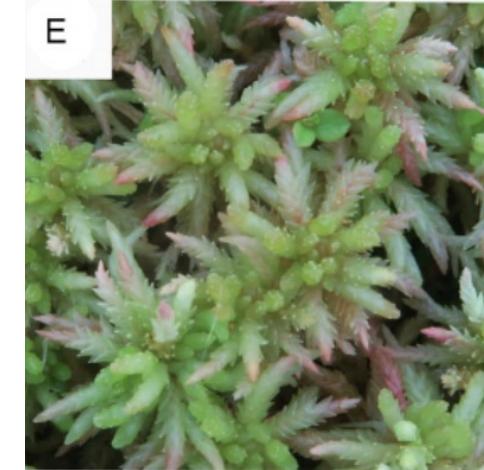


CYCLE SEQUENCING APPLICATIONS

1. Taxonomy (molecular + morphological)
2. Biogeography
3. Functional and diagnostic genetics
4. Cryptic species identification



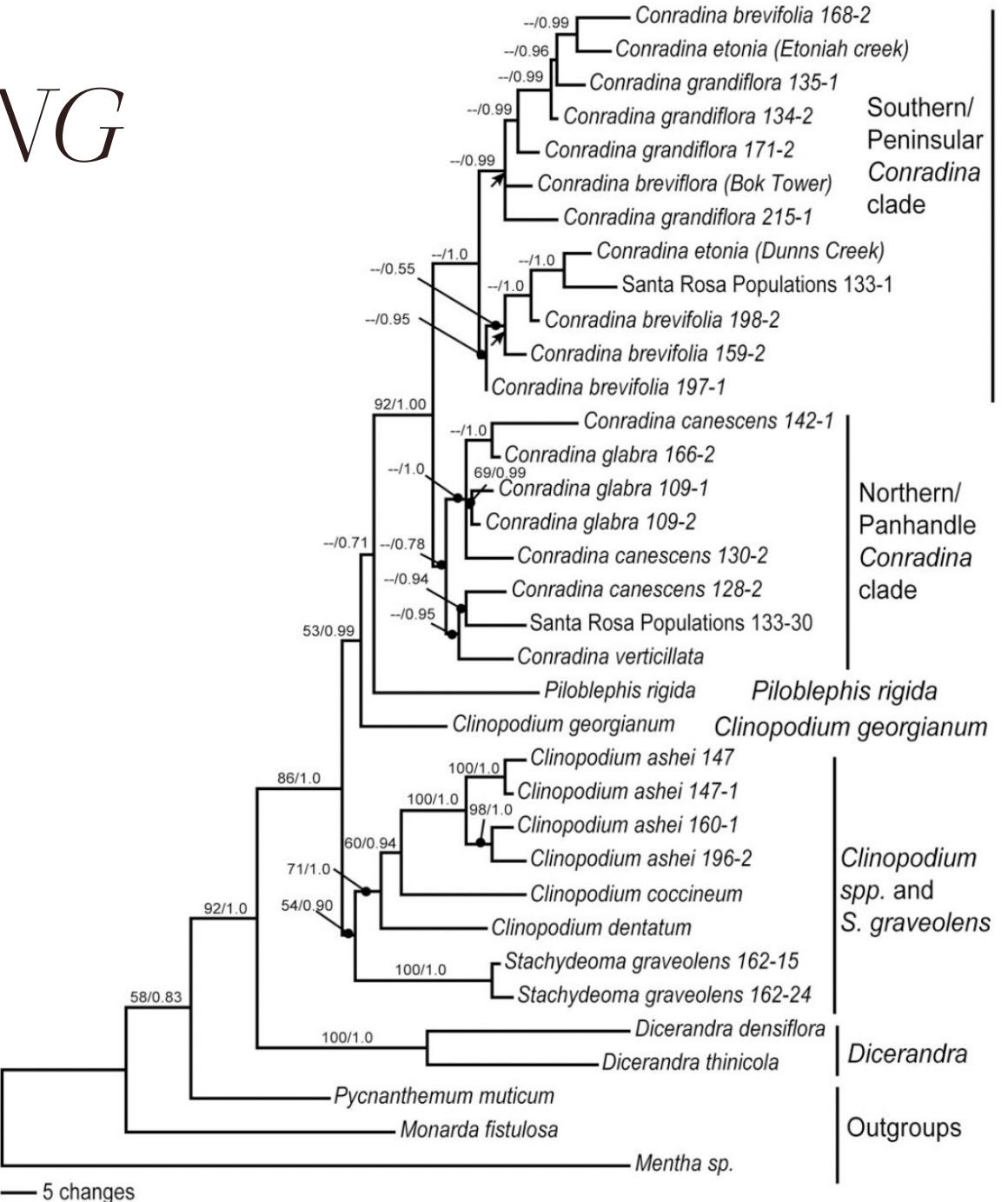
Sphagnum diabolicum



Sphagnum divinum

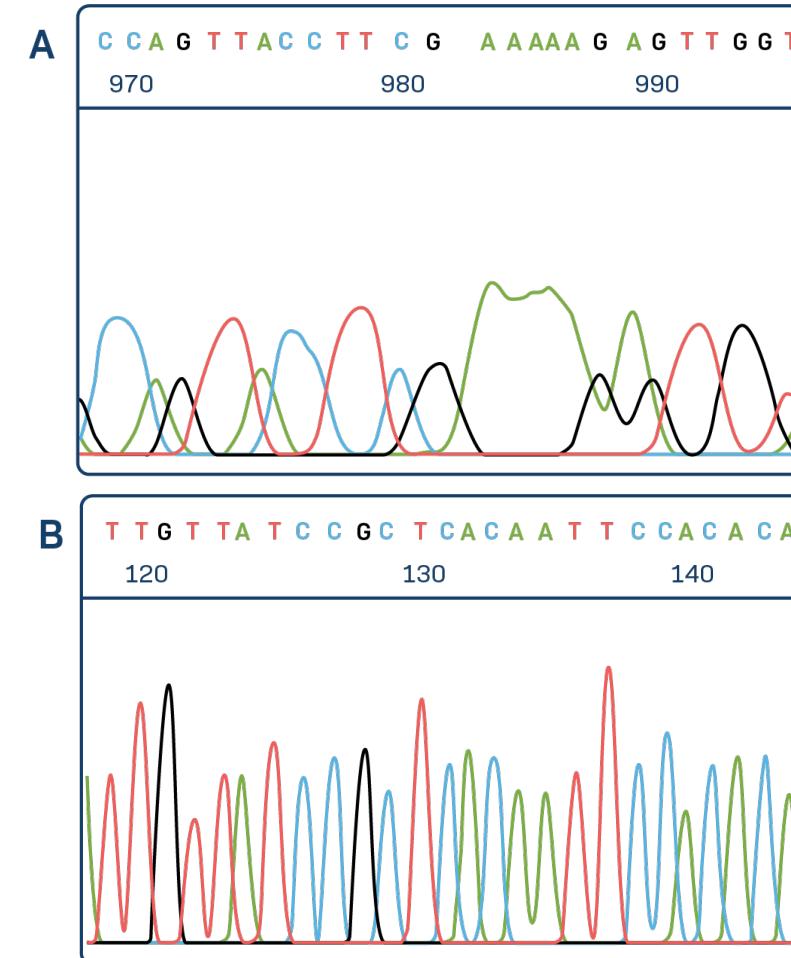
CYCLE SEQUENCING APPLICATIONS

1. Taxonomy (molecular + morphological)
2. Biogeography
3. Functional and diagnostic genetics
4. Cryptic species identification

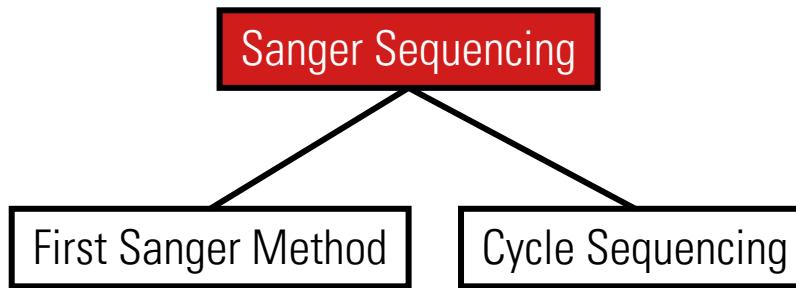


CYCLE SEQUENCING LIMITATIONS

1. Only targets genes/areas of interest
2. Short read length
3. Poor data quality at ends of reads
4. Difficulty dealing with repetitive regions
5. Cannot capture fine-scale differentiation among/within populations



SEQUENCING METHODS



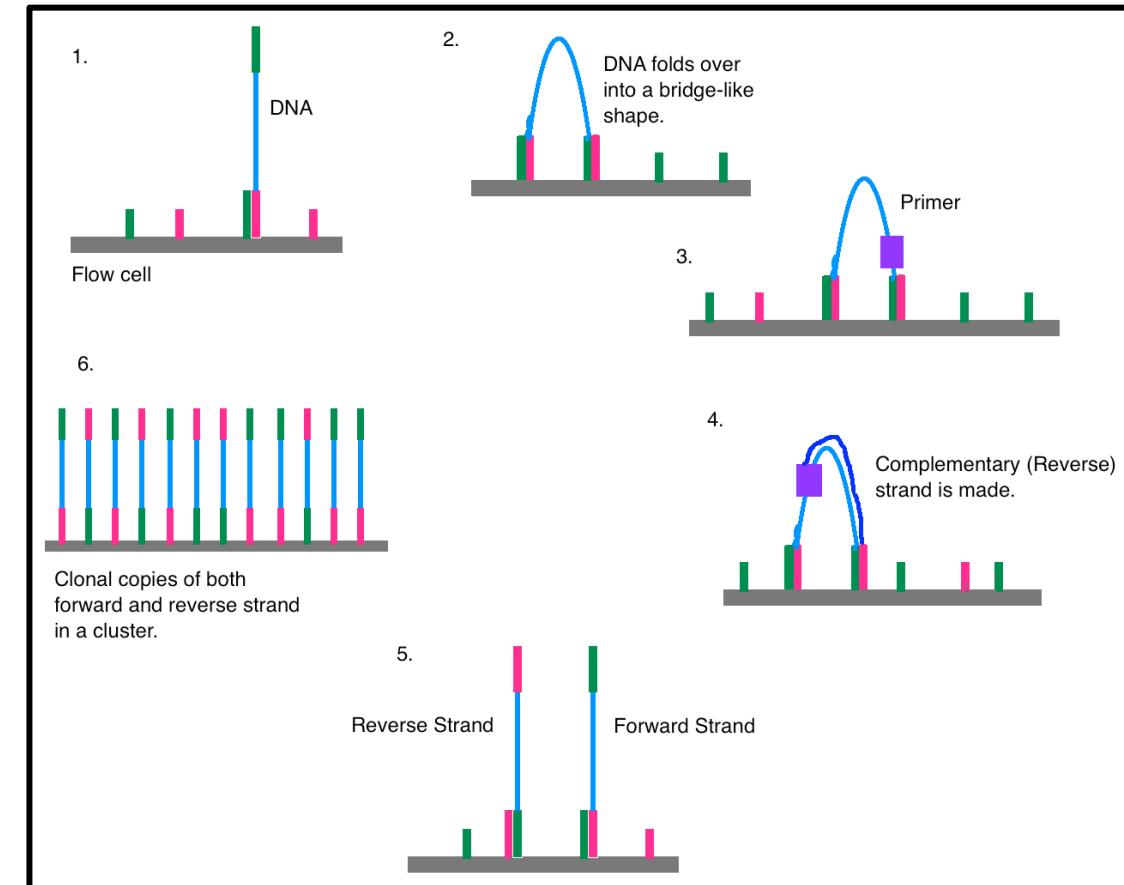
The background features a complex network of thin red lines forming various intersecting and parallel patterns across the white space.

2ND GENERATION

ILLUMINA

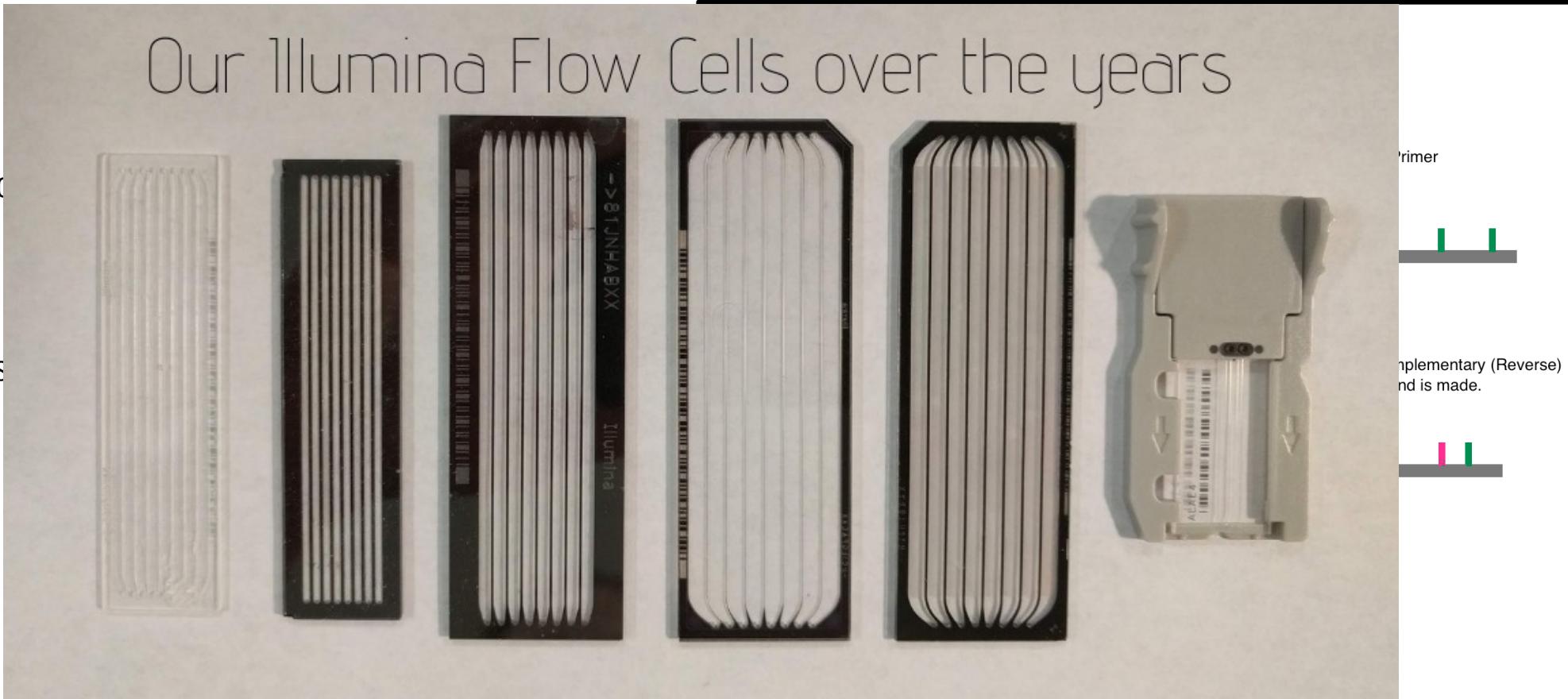
DYE SEQUENCING

- Follows core principles of Sanger Methods
 - Amplify > Sequence > Analyze
 - Sequence by Synthesis (SBS)
- Massive Parallel Sequencing



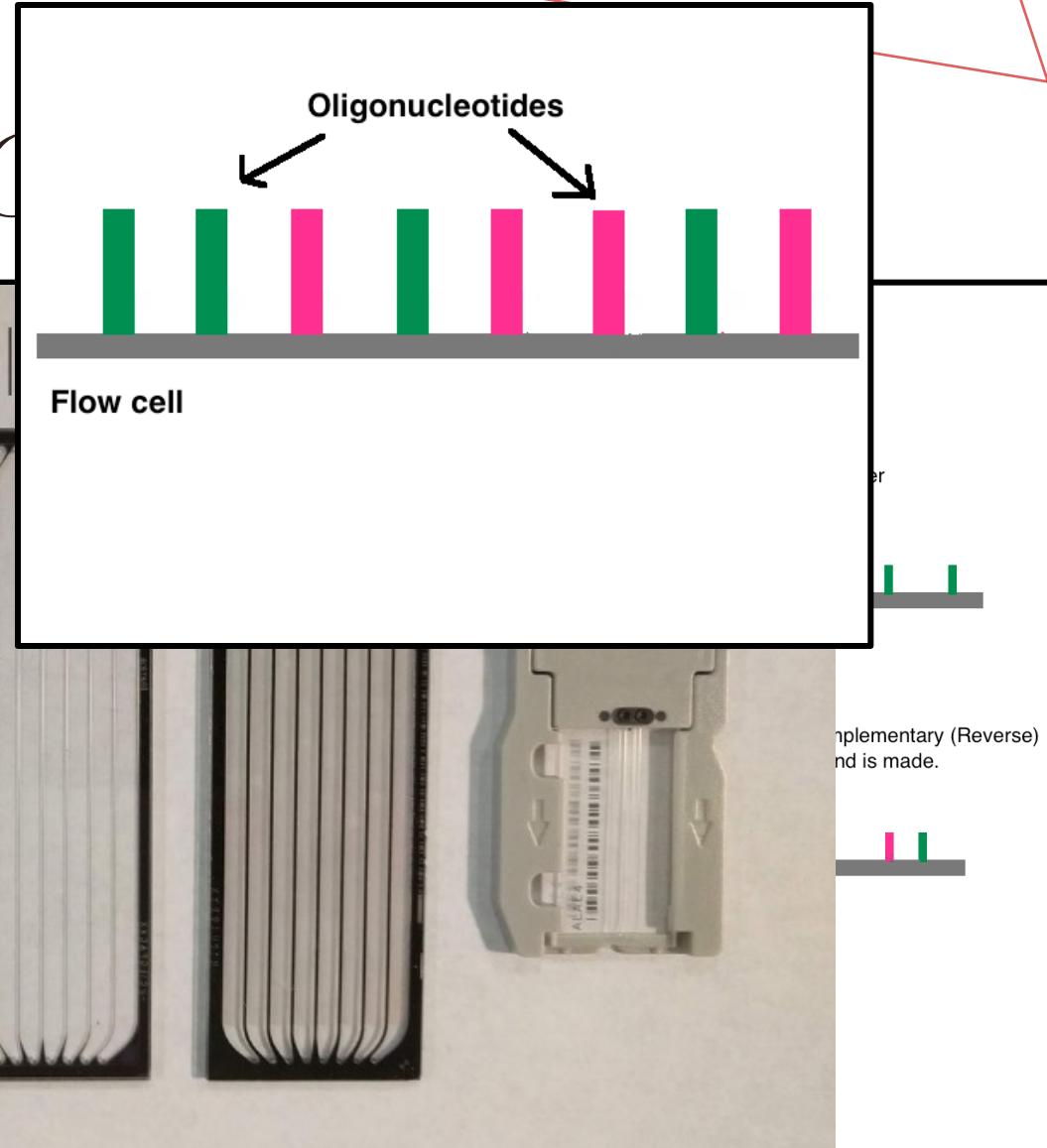
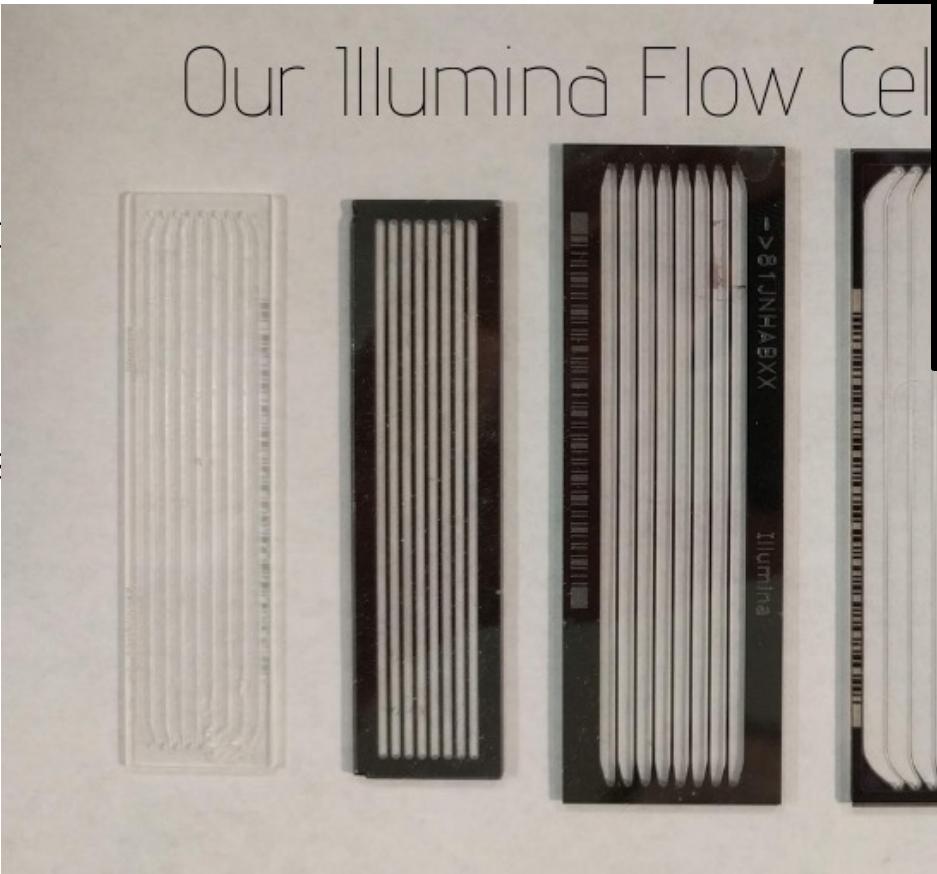
DYE SEQUENCING

- Follow
- •
- •
- Mass



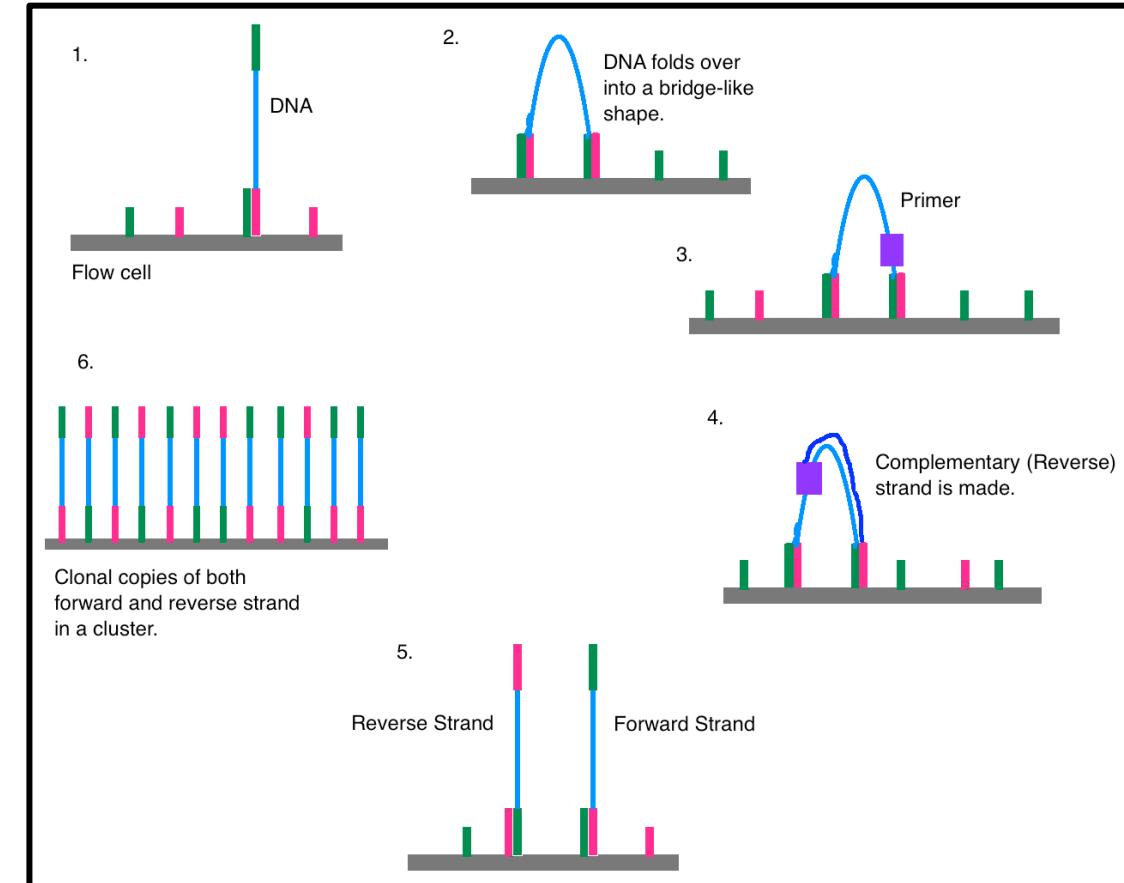
DYE SEQUENCING

- Follow
- •
- •
- Mass



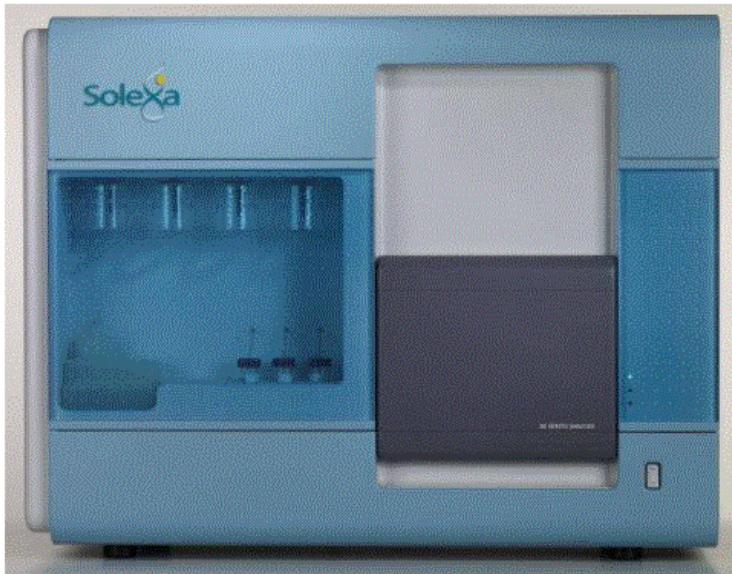
DYE SEQUENCING

- Follows core principles of Sanger Methods
 - Amplify > Sequence > Analyze
 - Sequence by Synthesis (SBS)
- Massive Parallel Sequencing
- **HUGE SPEED IMPROVEMENT**



DYE SEQUENCING

Solexa 1G Genome Analyzer



November 2006

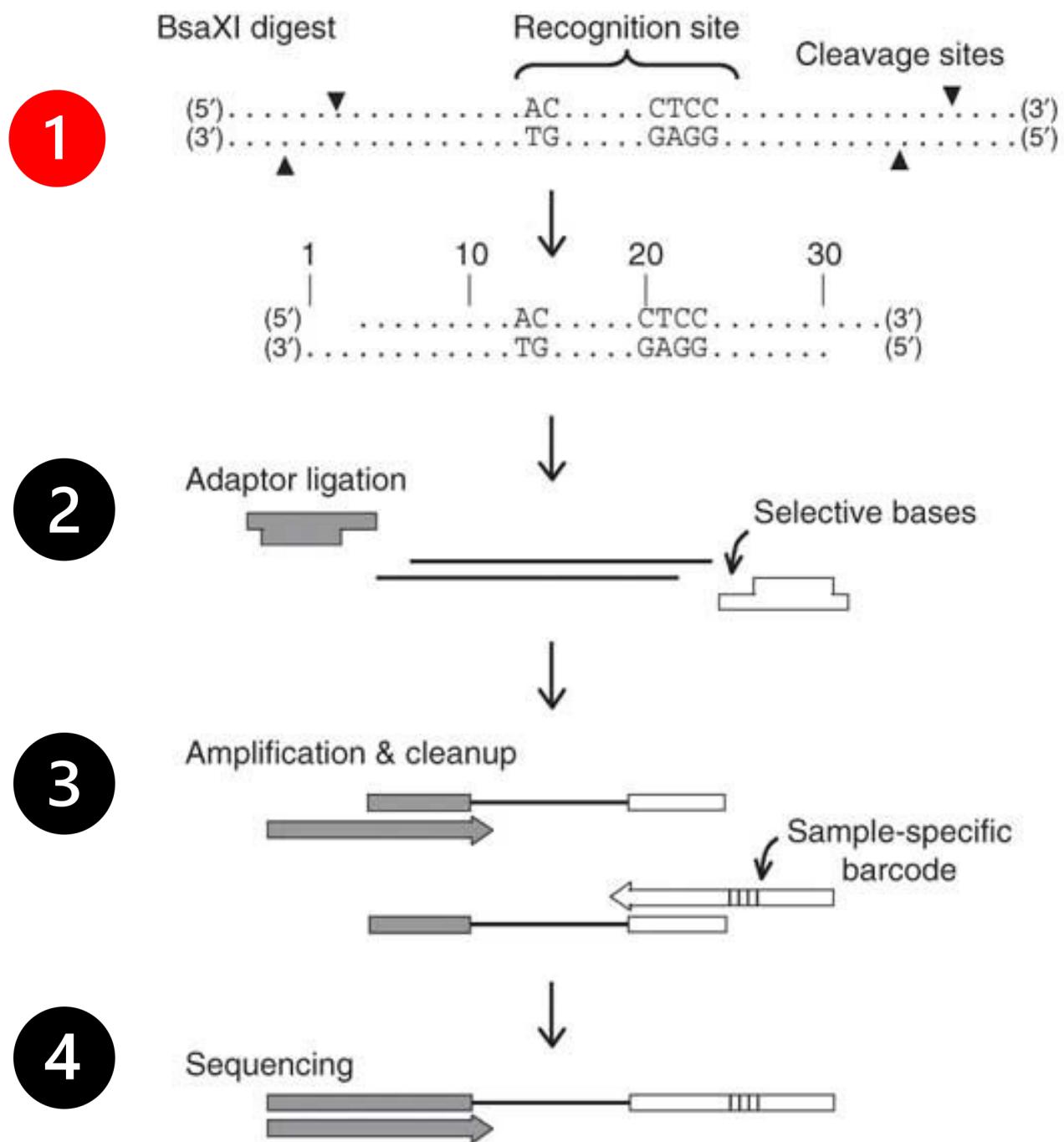


4



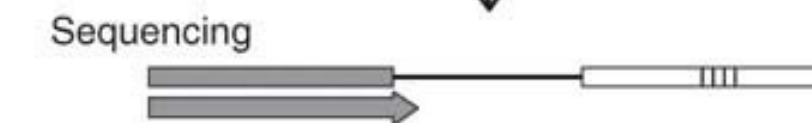
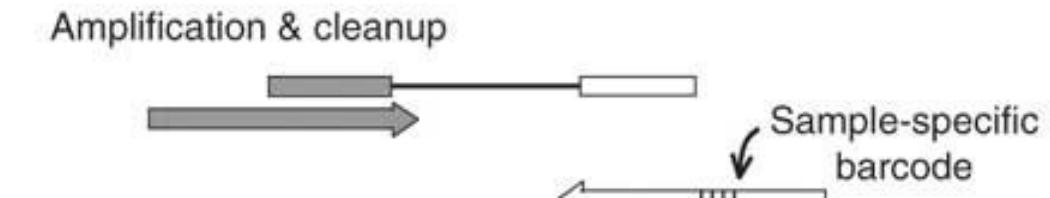
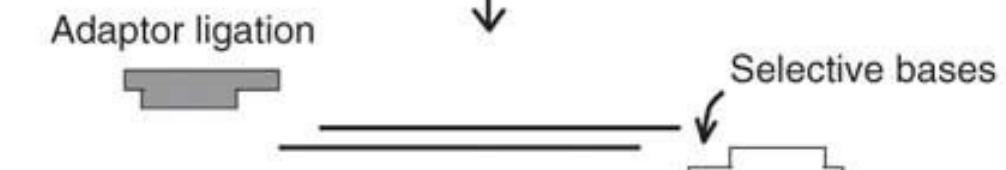
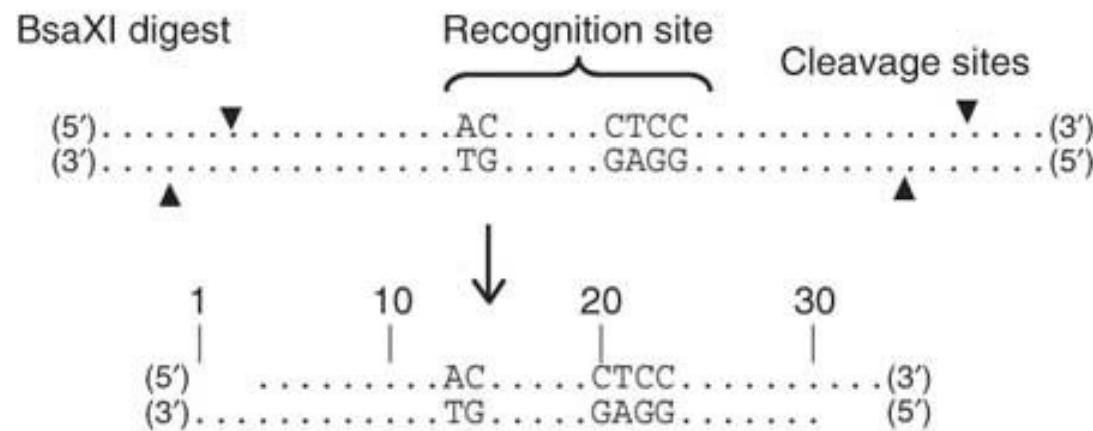
RAD-SEQ

Restriction enzymes cut at specific sites throughout the genome



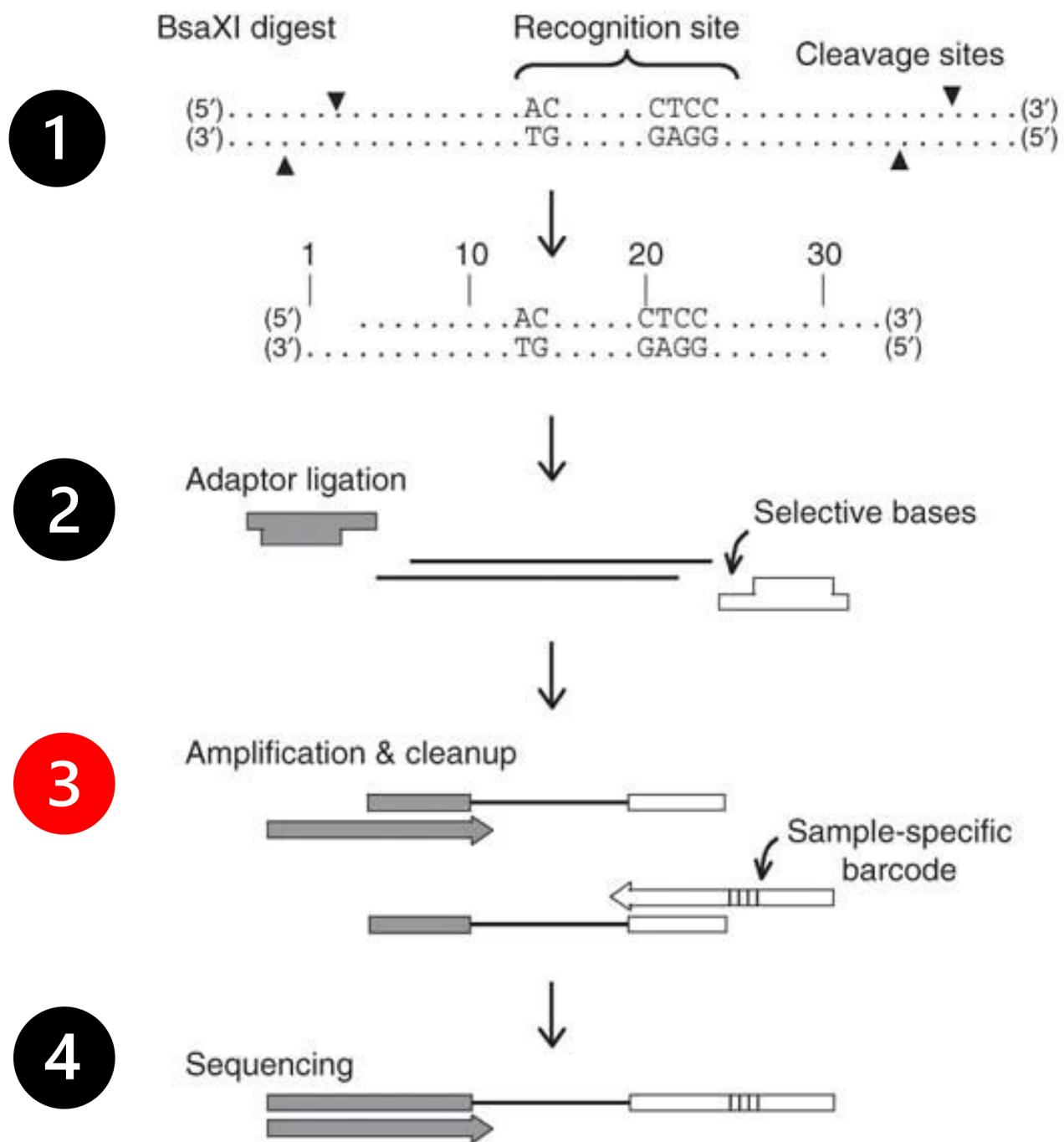
RAD-SEQ

Adapters differentiate samples which are sequenced together as "pools"



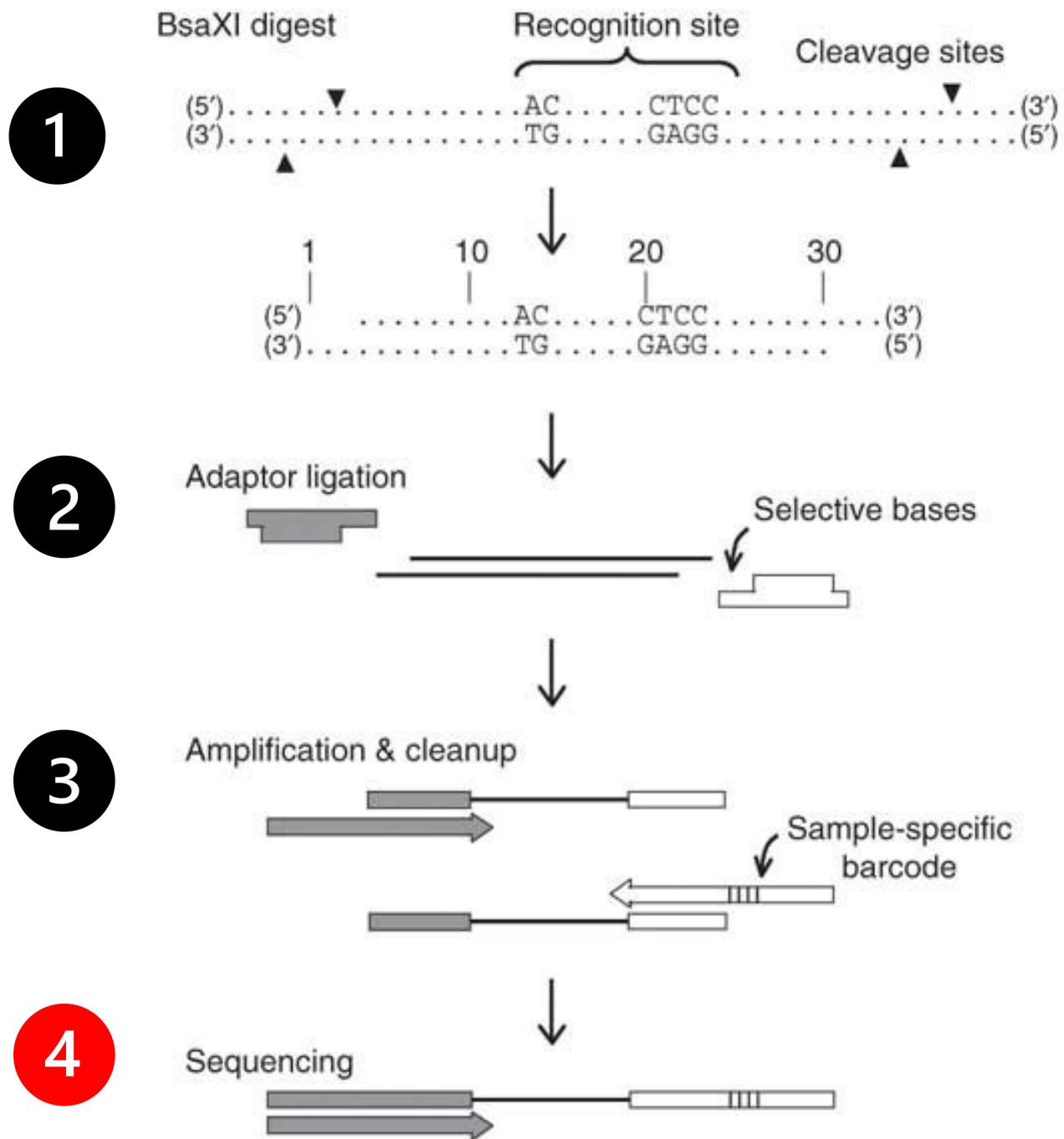
RAD-SEQ

Size selection of the fragment with the restriction enzyme



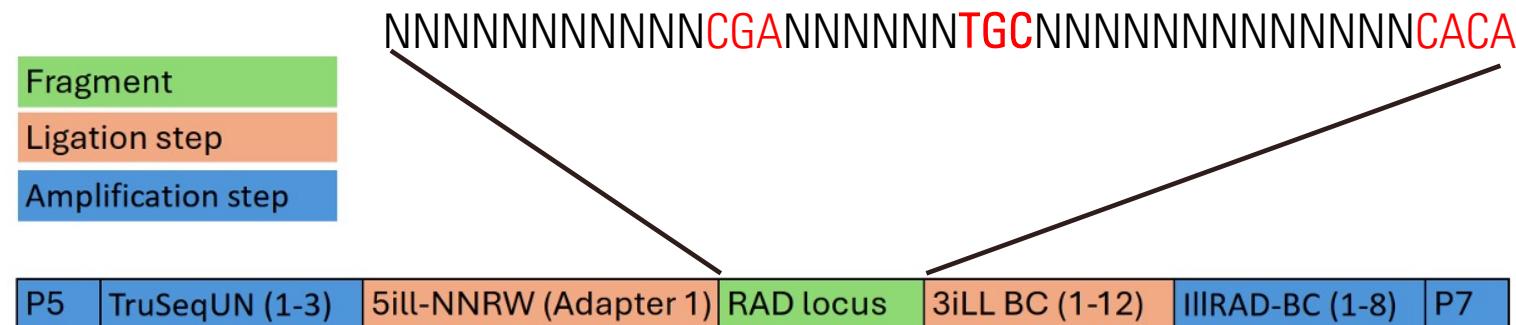
RAD-SEQ

Short fragments sequenced at high depth



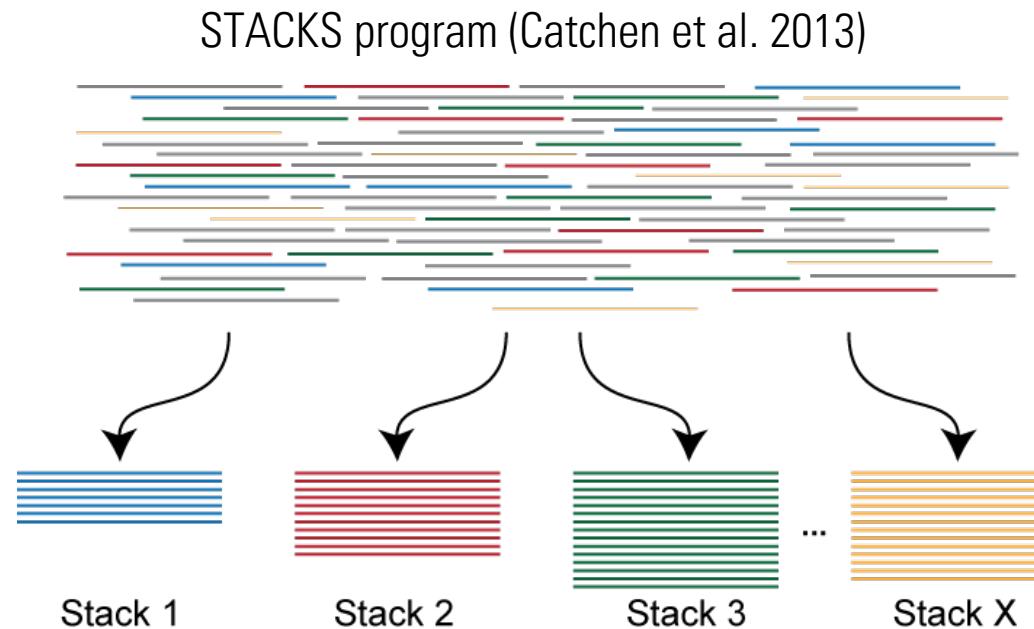
2B-RADSEQ APPROACH

1. Digestion with BgI restriction enzyme from *Bacillus coagulans*
 2. Cleaves up- and downstream of sites with a predictable pattern (36 base pair fragments)
 3. Ligation of column and row specific adapters



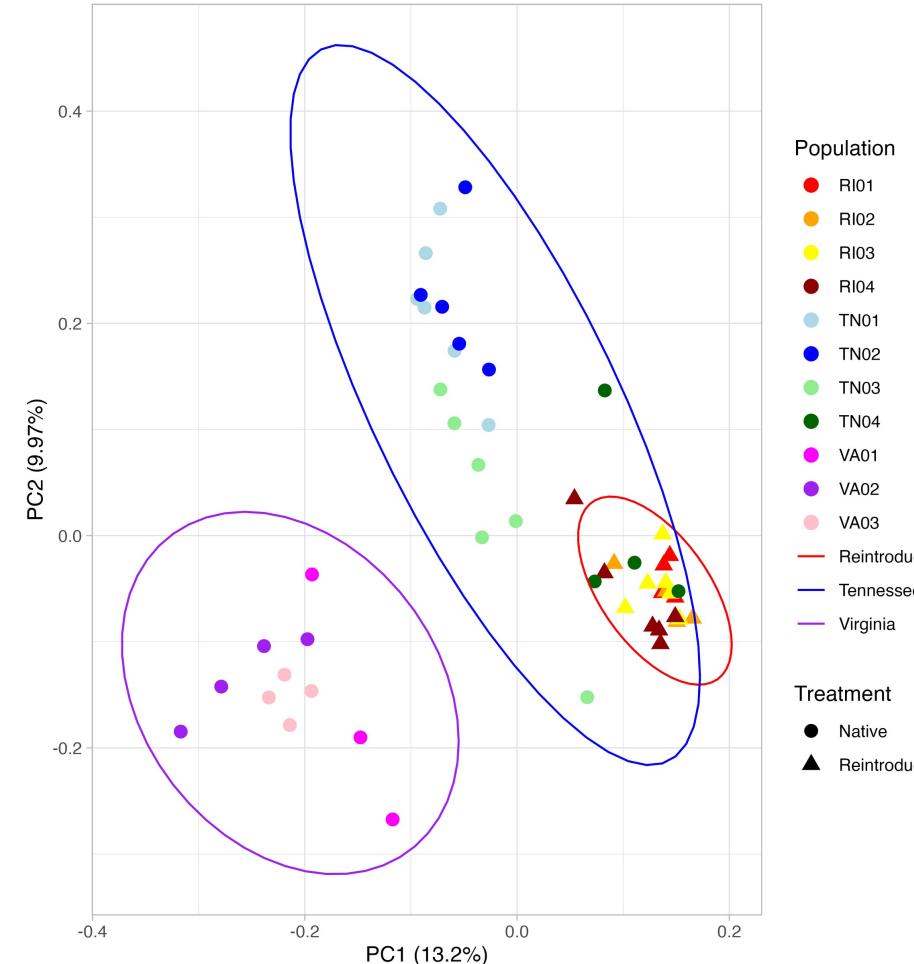
RAD-SEQ GENOME ASSEMBLY

1. Data assembled into “stacks”
2. Aligned to a reference genome or assembled *de novo*
3. Datasets composed of loci (36 base pairs) scattered throughout the genome and 1,000s of SNPs

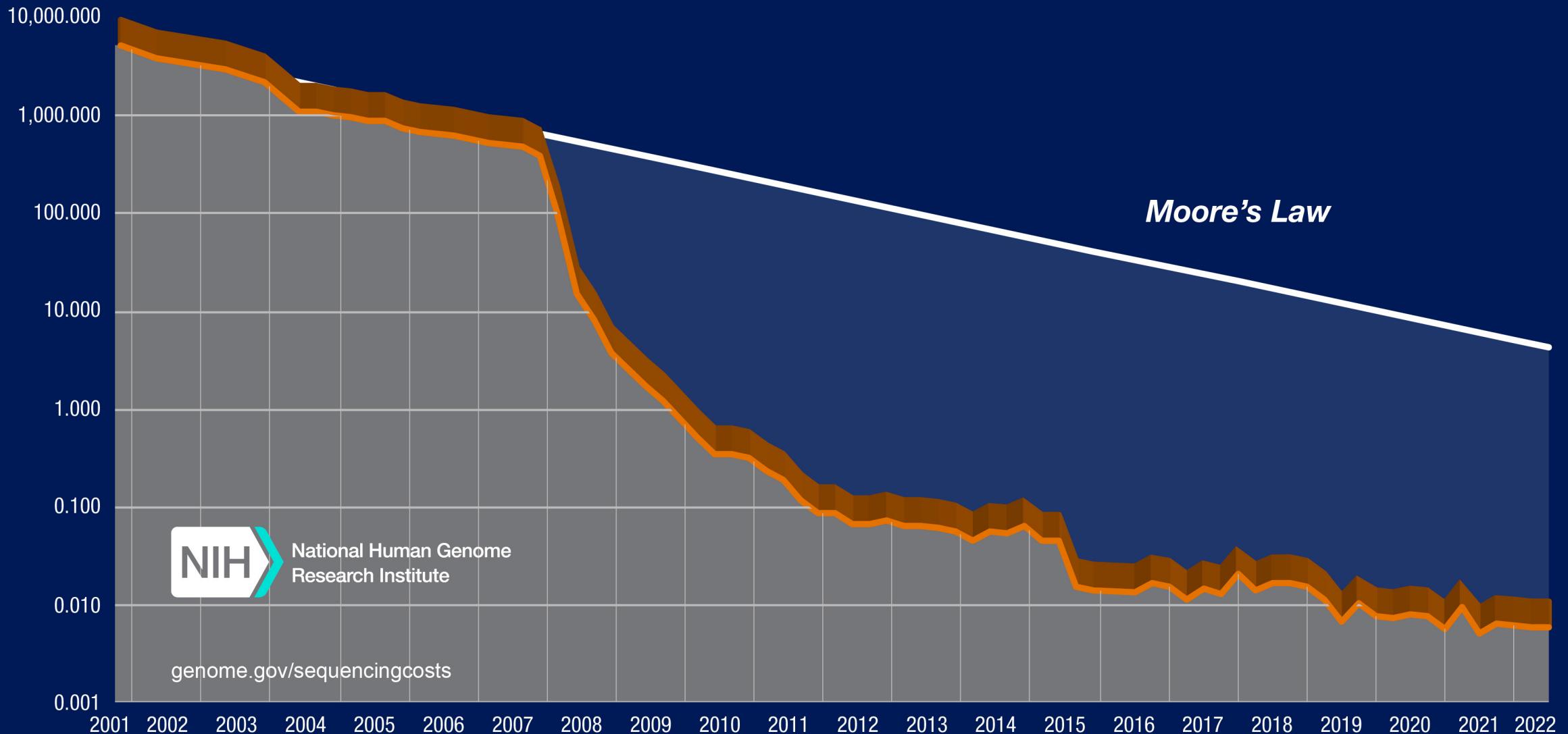


RAD-SEQ ADVANTAGES

- Genome-wide coverage at a lower cost
- De novo genome assembly without a reference genome for study of non-model organisms
- Able to detect fine-scale population structure
- Conservation applications



Cost per Raw Megabase of DNA Sequence



SEQUENCING METHODS



The background features a complex geometric pattern of red lines. It includes several straight lines forming a large 'X' shape, with additional lines radiating from the intersections and endpoints of these lines, creating a star-like or network-like effect.

3RD GENERATION AND ONWARDS

PACBIO AND OXFORD NANOPORE

WHAT MAKES 3RD GEN 3RD GEN?

- 1st and 2nd generation sequencing methods typically use the same framework
 - Amplify
 - Sequence
 - Analyze
- 2nd generation sequencing breaks DNA into small segments (50 – 400 bp) for sequencing
- Smaller chunks are easier to manage and result in fewer errors but take longer to sequence and analyze

WHAT MAKES 3RD GEN 3RD GEN?

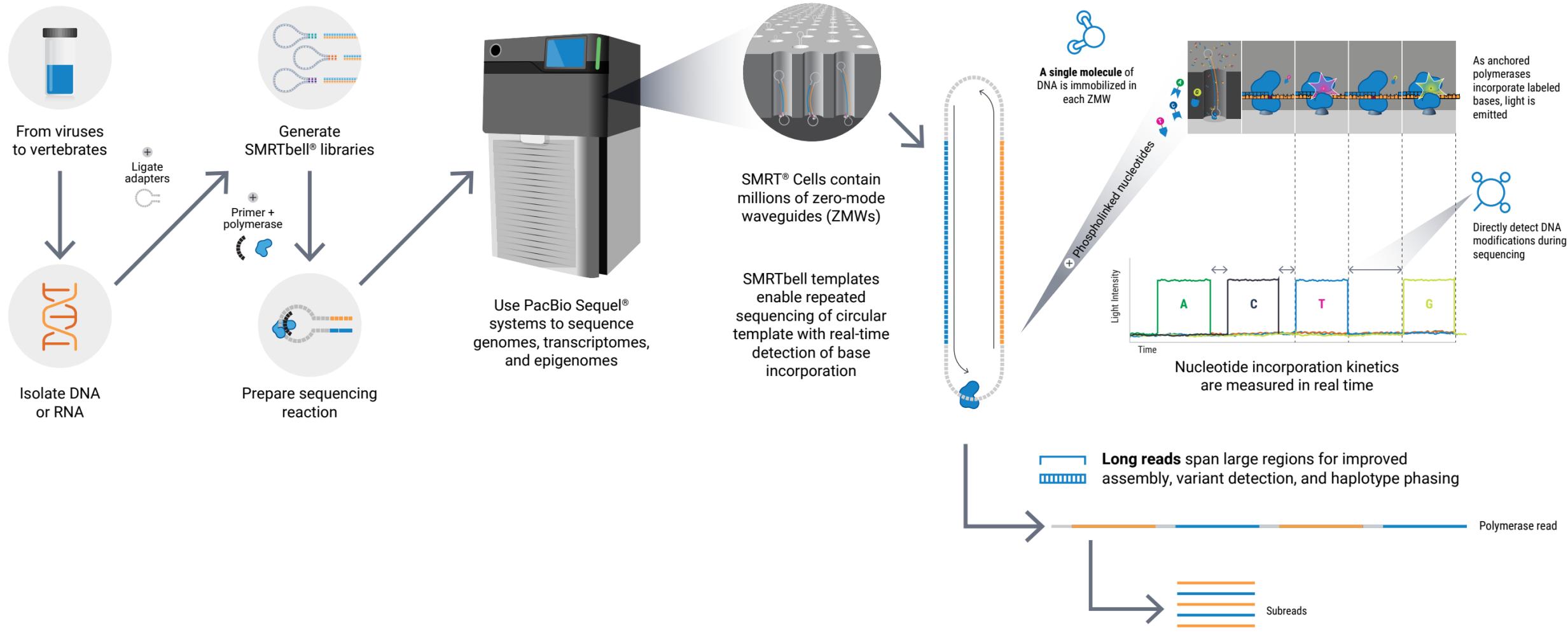
- 3rd generation makes two improvements over general 2nd generation methods
 1. Sequence without breaking DNA into smaller parts (short-reads) and sequence longer length of DNA (long reads, 1,500 – 100k bp)
 2. Skipping the amplification step in typical 1st and 2nd gen sequencing frameworks
- 3rd generation is **not strictly better** than 2nd generation but is **just different**
 - Long reads can tell us different information but is overall **less accurate** than 2nd generation

PacBio

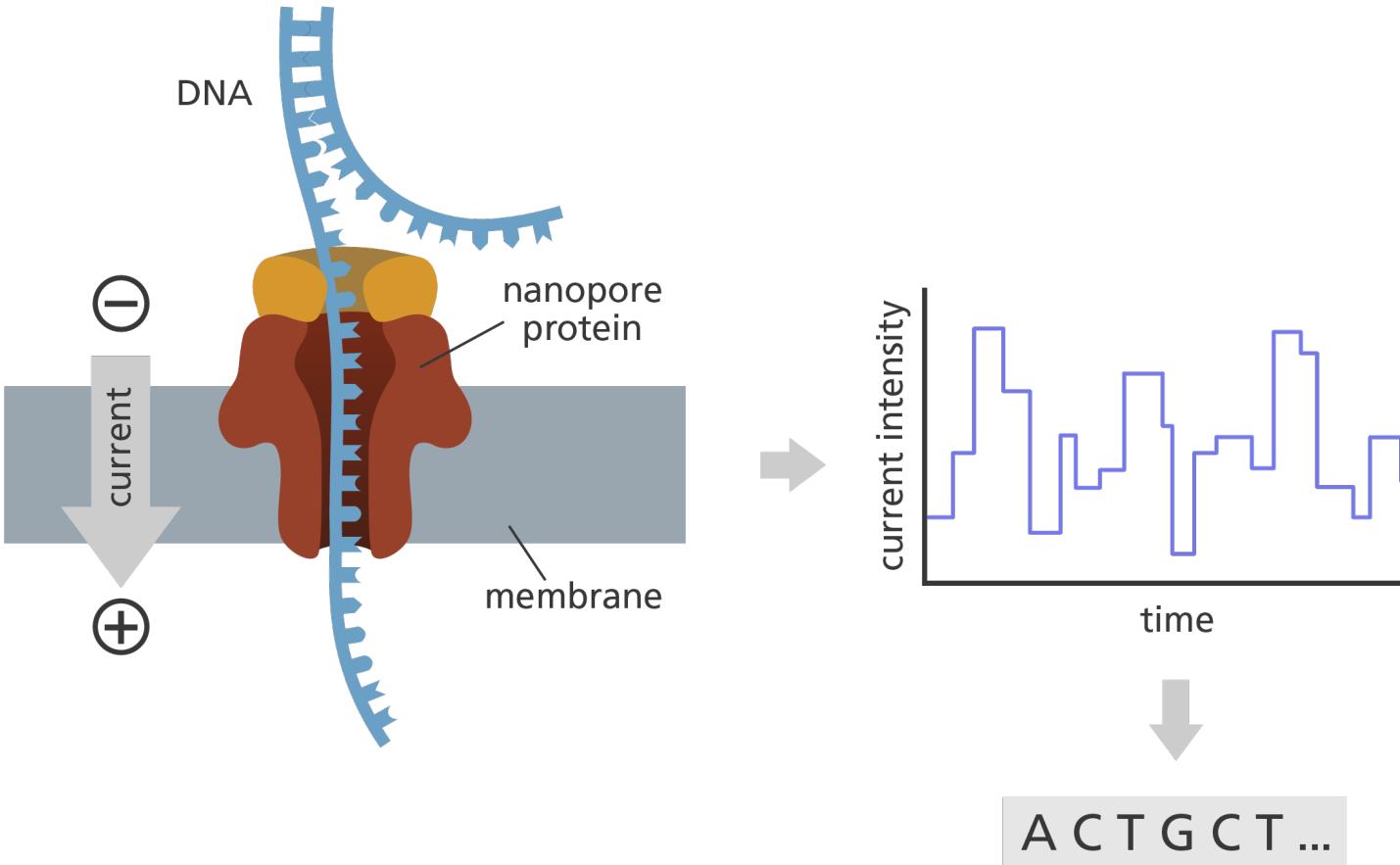
Oxford
NANOPORE
Technologies™

PACBIO SMRT SEQUENCING - 2011

Single Molecule, Real-Time (SMRT) Sequencing



NANOPORE



NANOPORE MINION



LONG READS VS SHORT READS

SHORT READS

- Shorter read length can't capture complex genomic structures
- Struggles to reconstruct entire genomes
- Very accurate
- Low DNA input requirement
- Inexpensive per base but can scale fast

LONG READS

- Long reads capture complex genomic structures
- Facilitates complete genome assembly
- Less accurate but improving
- Need more DNA input
- Higher operating cost

SEQUENCING METHODS

