DNA Sequencing

DNA sequence determination is a tremendous human achievement. DNA sequencing includes several methods and technologies in use for determining the order of the nucleotide bases (adenine, guanine, cytosine, and thymine) in a molecule of DNA. Knowledge of DNA sequences has become indispensable for basic biological research, other research branches utilizing DNA sequencing, and in numerous applied fields such as diagnostic, biotechnology, forensic biology and biological systematics.

The advent of DNA sequencing has significantly accelerated biological research and discovery. Rapid sequencing, the result of modern DNA sequencing technology, is instrumental in the sequencing of the human genome for the Human Genome Project. Related projects, often by scientific collaboration across continents, have generated complete DNA sequences of humans as well as numerous animals, plants and microbial genomes.

DNA sequencing methods currently under development include labeling DNA polymerase and reading the sequence as a DNA strand transits through nanopores. Additional methods include microscopy-based techniques such as atomic force microscopy or transmission electron microscopy that are used to identify the positions of individual nucleotides within long DNA fragments (>5000 bp) by nucleotide labeling with heavier elements (e.g., halogens) for visual detection and recording. Third generation technologies aim to increase throughput and decrease the time to result and cost by eliminating the need for excessive reagents and harnessing the processivity of DNA polymerase.

In 1977 Sanger and colleagues published an enzymatic method for DNA sequencing using dideoxynucleotides as base-specific chain-terminators (Sanger et al., 1977). Within a few years, this approach became the preferred method of choice, mainly owing to ease of handling and generation of data and limited exposure to hazardous chemicals, especially compared to the chemical cleavage method of DNA sequencing (Maxam and Gilbert, 1977).

Today the chemical cleavage approach is generally restricted to those cases where it may be necessary to sequence short chemically synthesized oligonucleotides, which cannot be analysed using chain terminator chemistry. Many oligonucleotides are too small for a sequencing primer to be used and the sequence at the priming site cannot be deduced. There have been enormous advances in DNA sequencing technology, starting with a range of different DNA polymerases:

Klenow (Sanger et al., 1977), T4 DNA polymerase (Cammeron-Mils, 1988) and AMV reverse transcriptase, finally culminating in Taq DNA polymerase (Innis et al., 1988), one of its modified derivatives or other thermostable DNA polymerases. These thermostable polymerases facilitated the development of cycle sequencing (Lee, 1991), essentially a one-sided PCR that requires less starting template, can overcome potential secondary structure problems, and is highly suited to the direct sequencing of double-stranded PCR products or plasmids.

In all these cases of enzymic sequencing, the basic theme is to have a reference point (the sequencing primer) that is extended by the DNA polymerase in the presence of deoxyribonucleoside triphosphates (dNTP=building block) and a dideoxyribonucleoside triphosphate (ddNTP=chain-terminator).

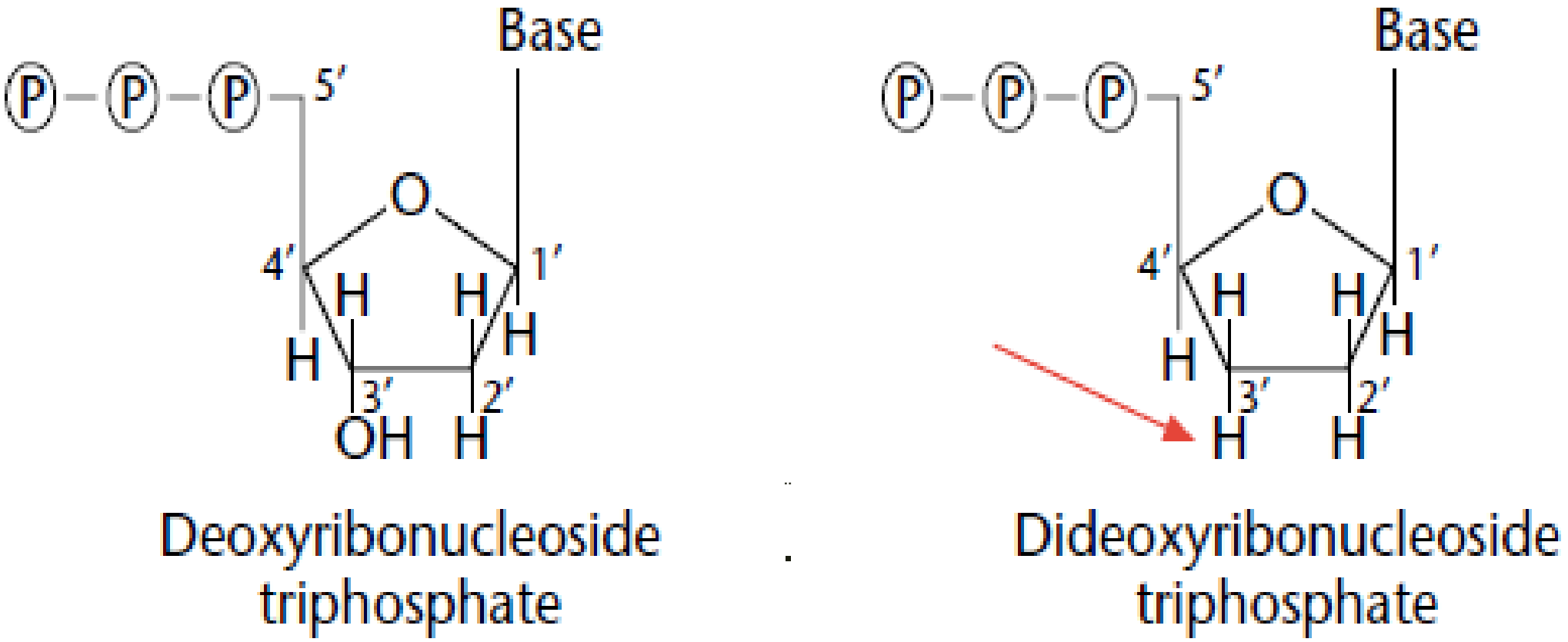
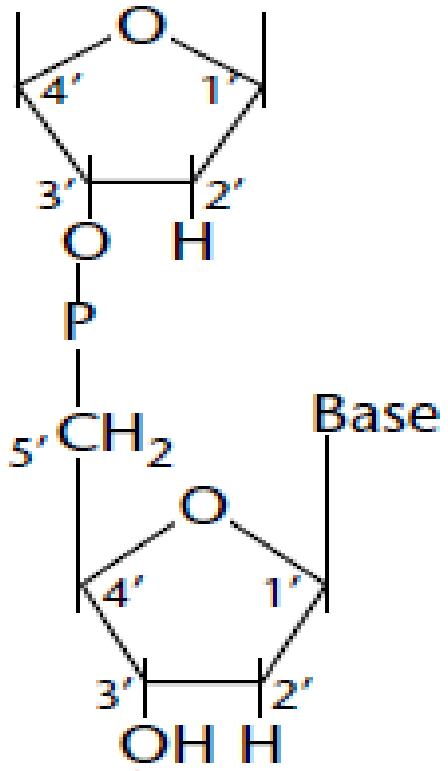
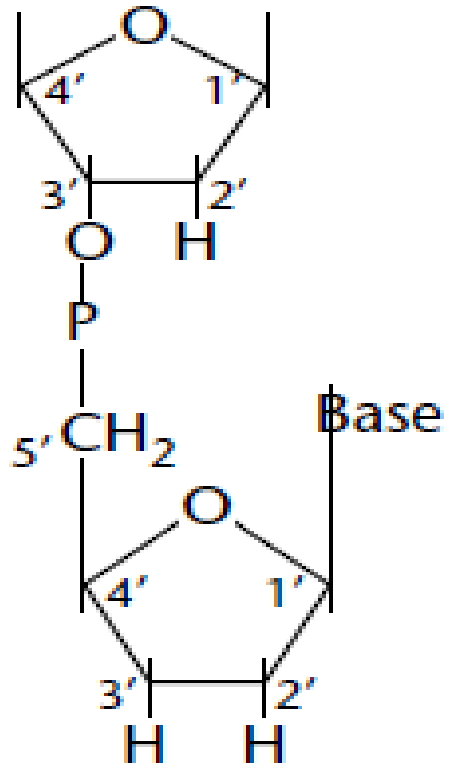


Figure 1 Deoxynucleoside triphosphate and its analogue the dideoxynucleoside triphosphate.



Incoming base incorporated at the 3′OH group (chain extension)



No 3′OH group for incorporation of the next base (chain termination)

Figure 2 Chain termination upon the incorporation of a dideoxynucleoside triphosphate.

A set of sequencing reactions consists of four separate reactions, each with its own chain terminator:

. A stops owing to the incorporation of ddATP.
. C stops owing to the incorporation of ddCTP.
. G stops owing to the incorporation of ddGTP.
. T stops owing to the incorporation of ddTTP.

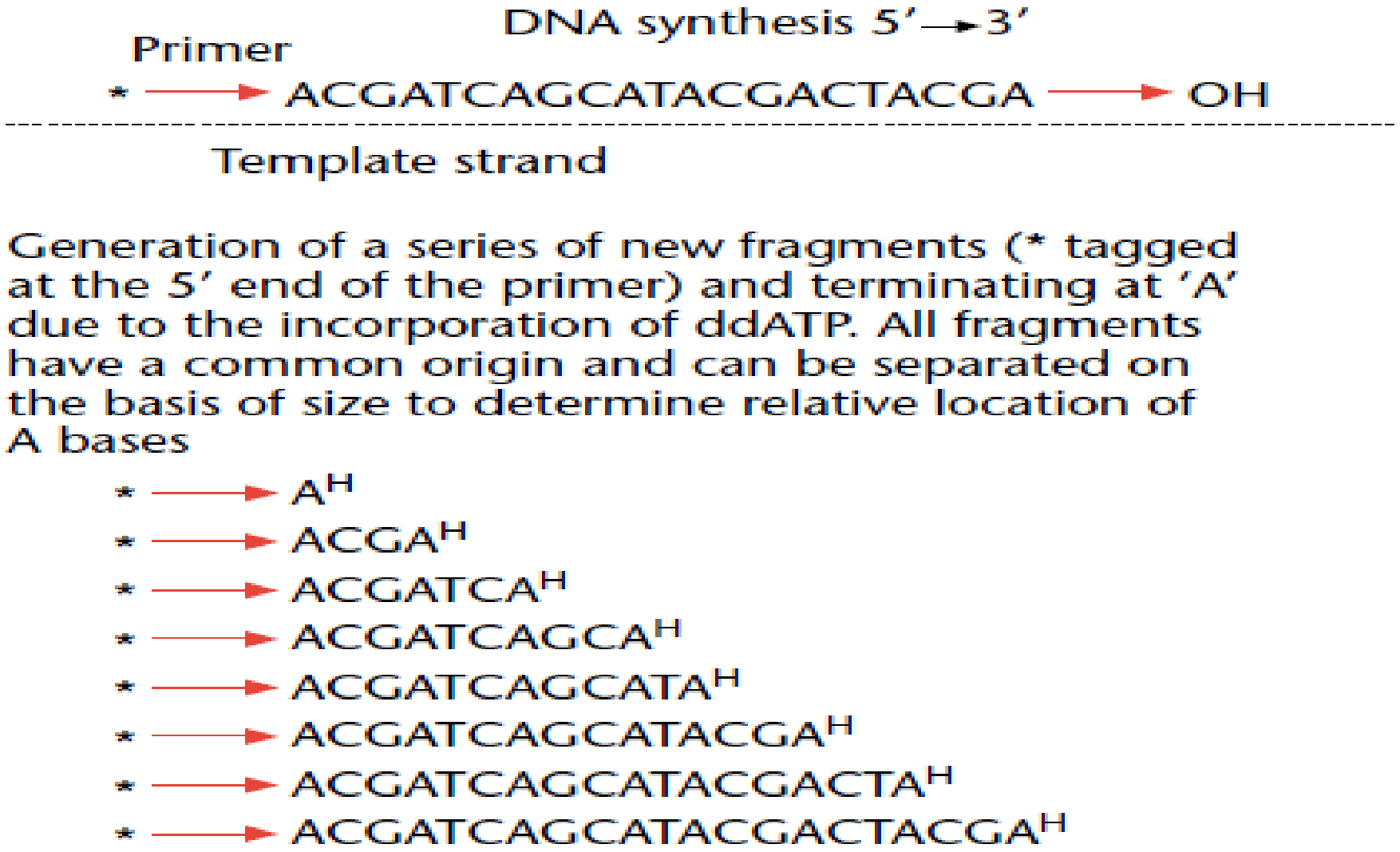


Figure 3 5′-terminal labelling of DNA sequencing fragments via the primer.

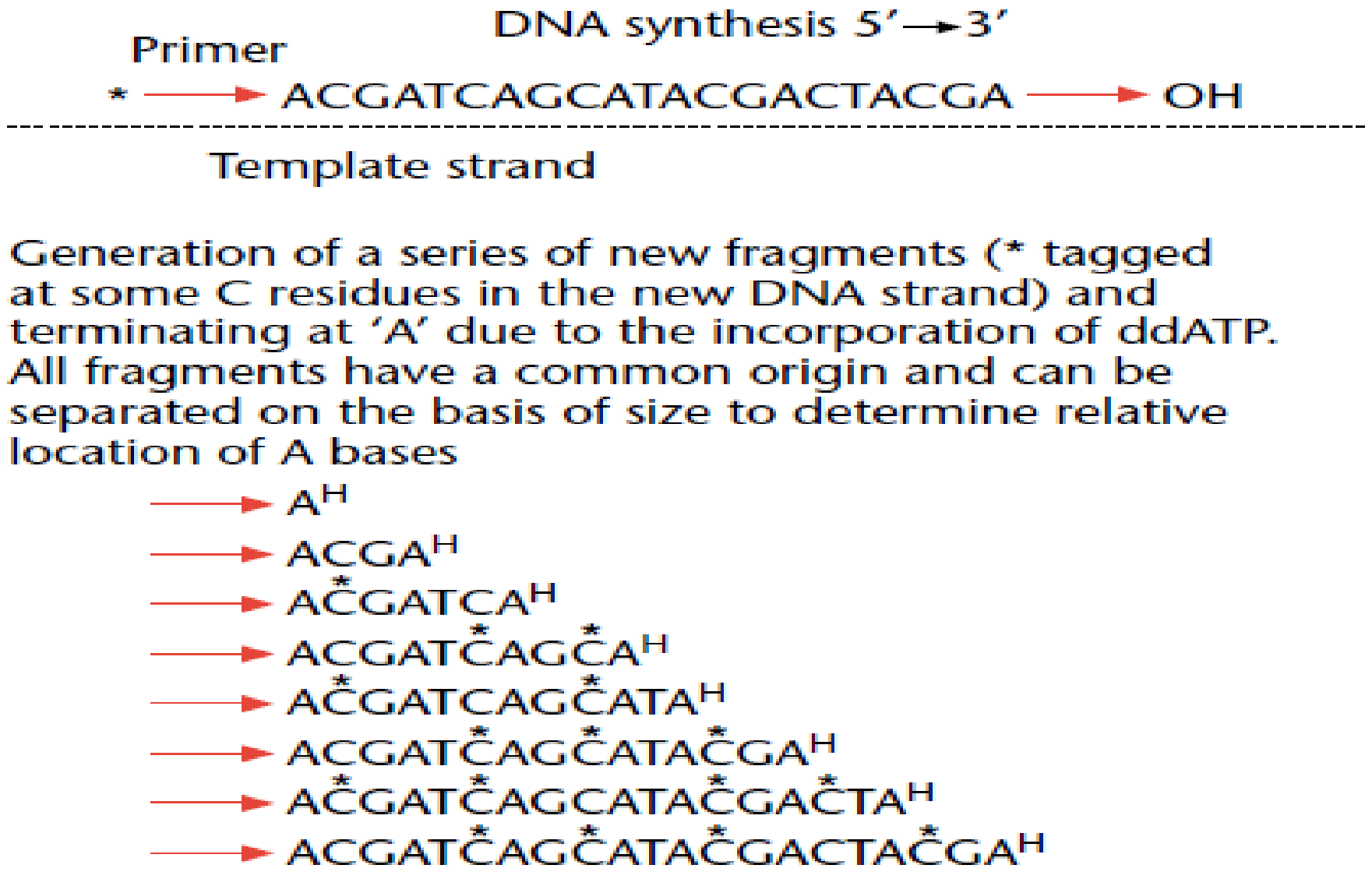
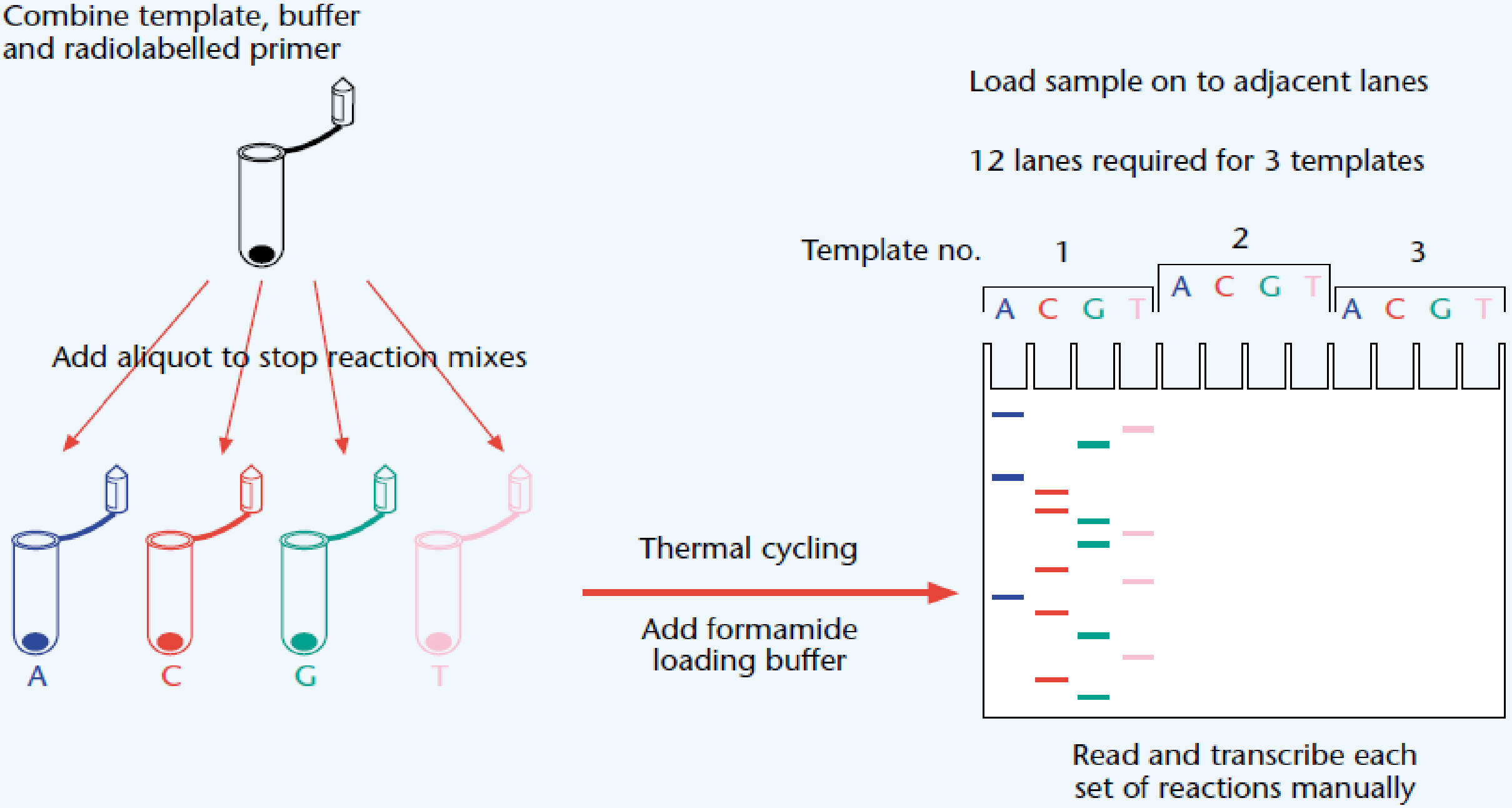


Figure 4 Incorporation of α-radiolabel into the sugar phosphate backbone.


Combine template, buffer
and radiolabelled primer
Add aliquot to stop reaction mixes
A
C
G
T
Thermal cycling
Add formamide
loading buffer
Load sample on to adjacent lanes
12 lanes required for 3 templates
Template no.
1
2
3
A C G T
A C G T
A C G T
Read and transcribe each
set of reactions manually

Step 1: Equipment and Solutions

Reagents

- 7-deaza dGTP
- Acetic acid
- Acrylamide
- Ammonium persulfate
- Bis-acrylamide
- Boric acid
- Bromophenol blue
- dATP
- dCTP
- ddATP
- ddCTP
- ddGTP
- ddTTP
- Dithiothreitol (DTT)
- dTTP
- EDTA
- Formamide
- $MgCl_2$
- Methanol
- Mixed Bed Resin (BioRad Mixed Bed Resin AG 501-X8(D) or equivalent)
- NaOH
- Spermidine
- Template preparation kits (QIAGEN's QIAprep Spin Plasmid Kit, Promega's Wizard Minipreps DNA Purification Systems or equivalent)
- TEMED

Step 2: Chain Termination Reactions –Primer Labelling
Step 3: Thermal Cycling Conditions
Step 4: Gel Fractionation of the Sequencing Products – Introduction
Step 5: Electrophoresis Setup
Step 6: Fractionation of the Sequencing Reactions
Step 7: Gel Manipulations
Step 8: Fixing and Drying the Gel
Step 9: Reading the Sequence

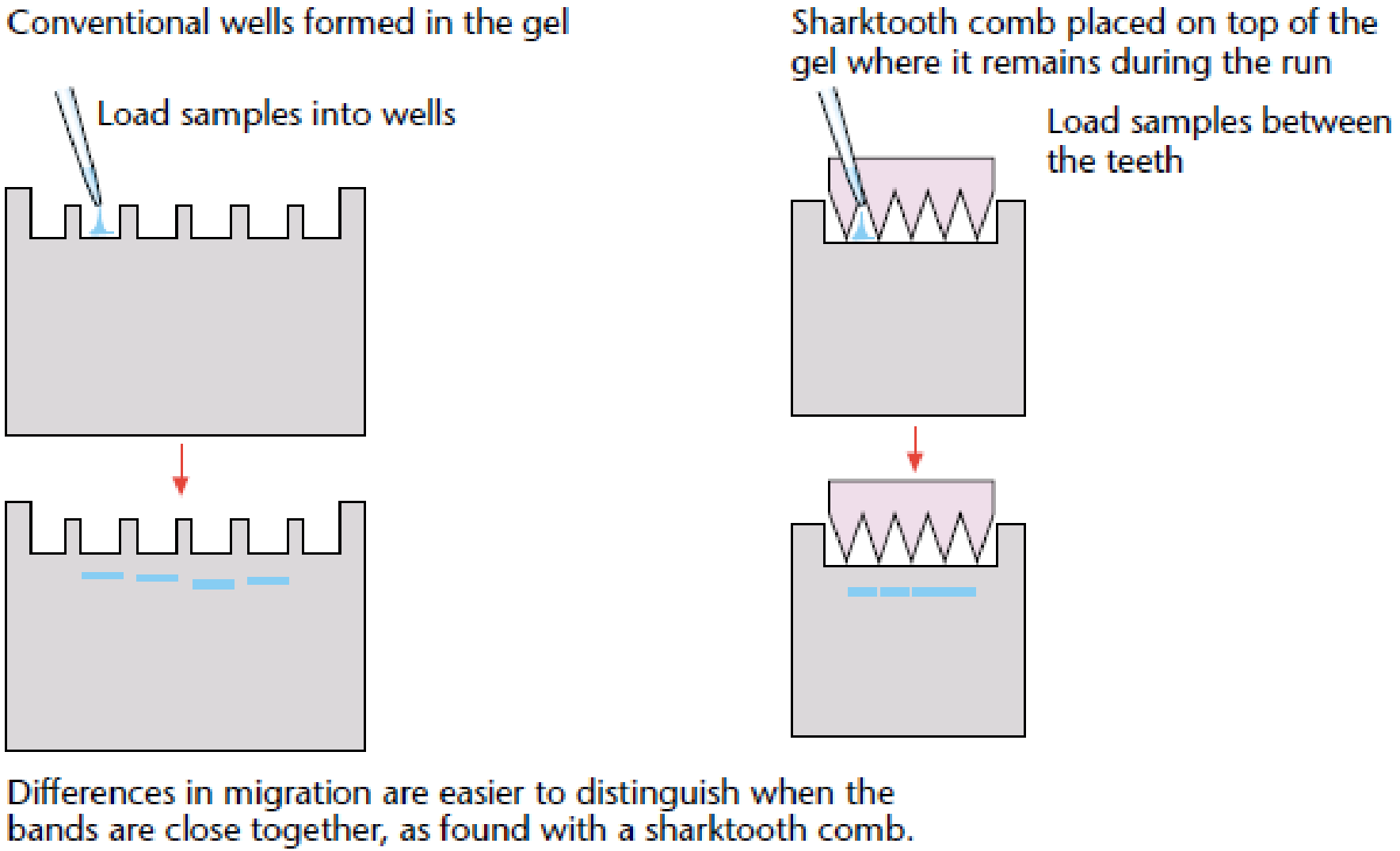


Figure 6 Comparison of well-forming and sharktooth combs for sequencing gels.

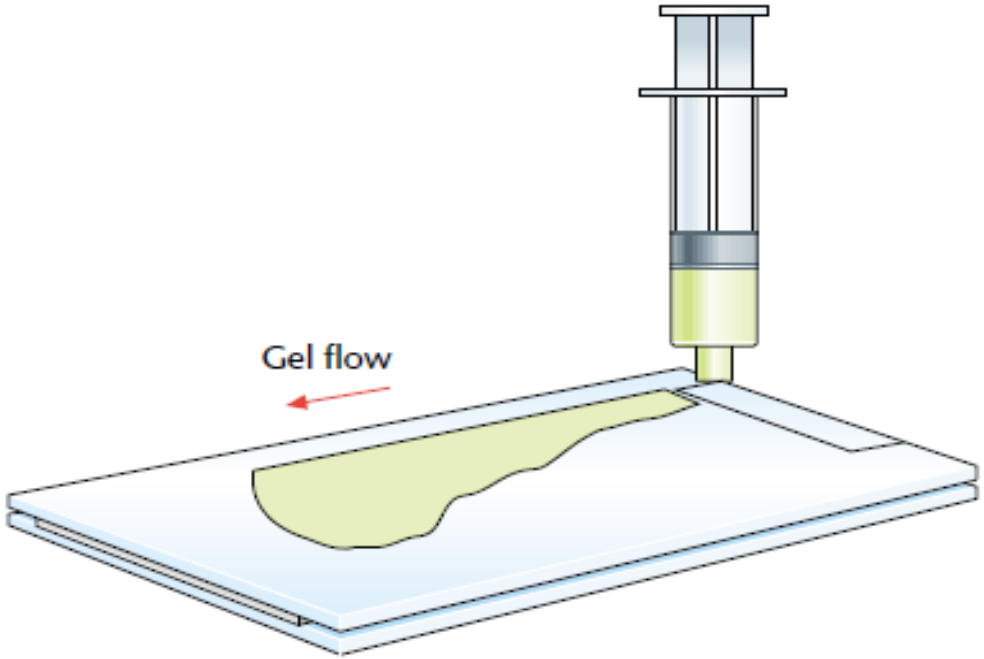


Figure 7(a) Inserting the gel into the gel apparatus.

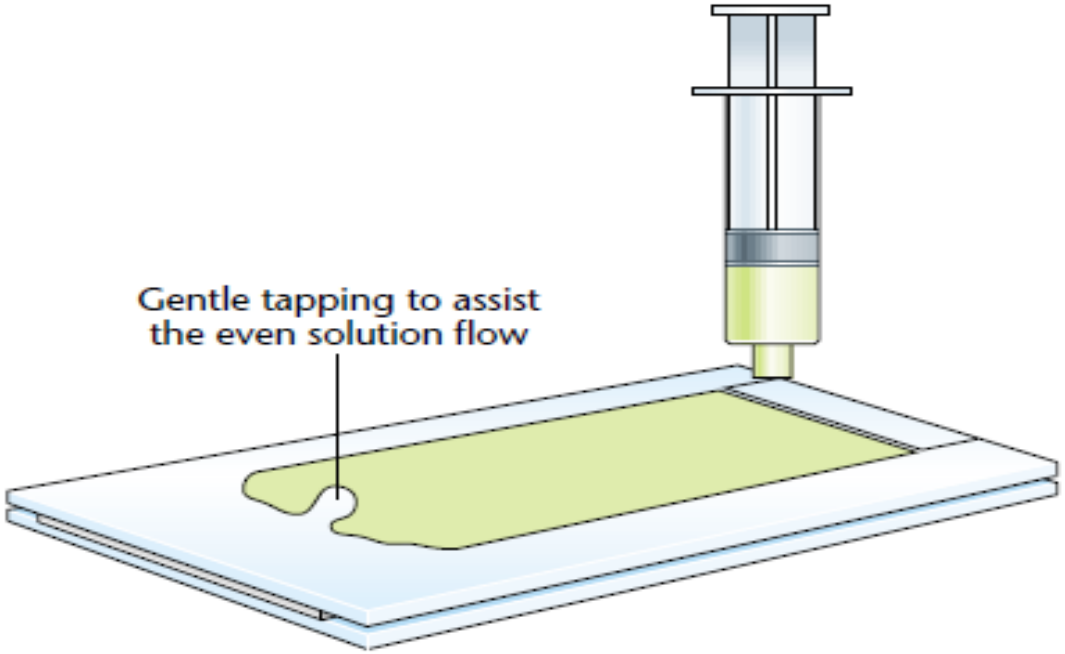


Figure 7(b) Pouring acrylamide mix into the gel cassette.

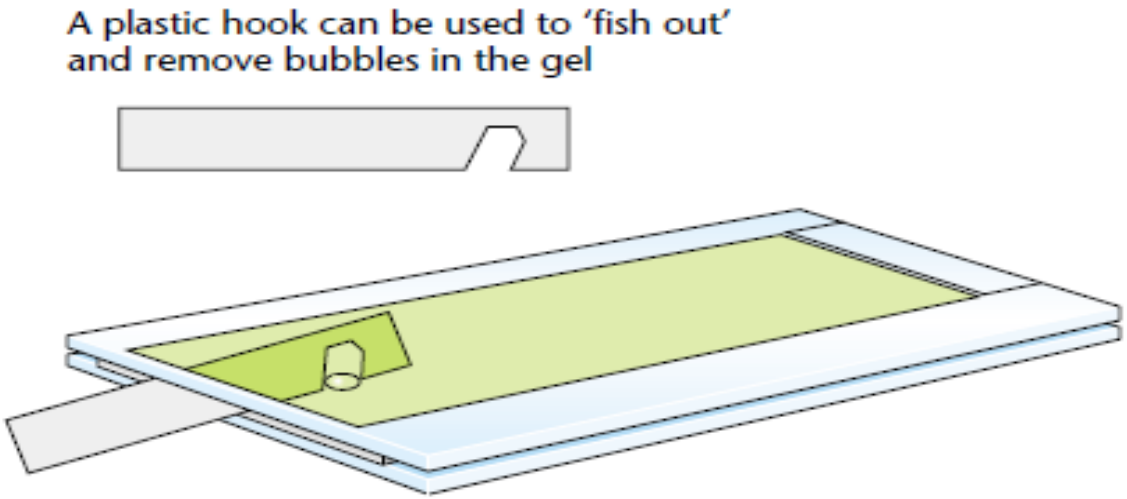


Figure 7(c) Bubble formation during gel pouring.

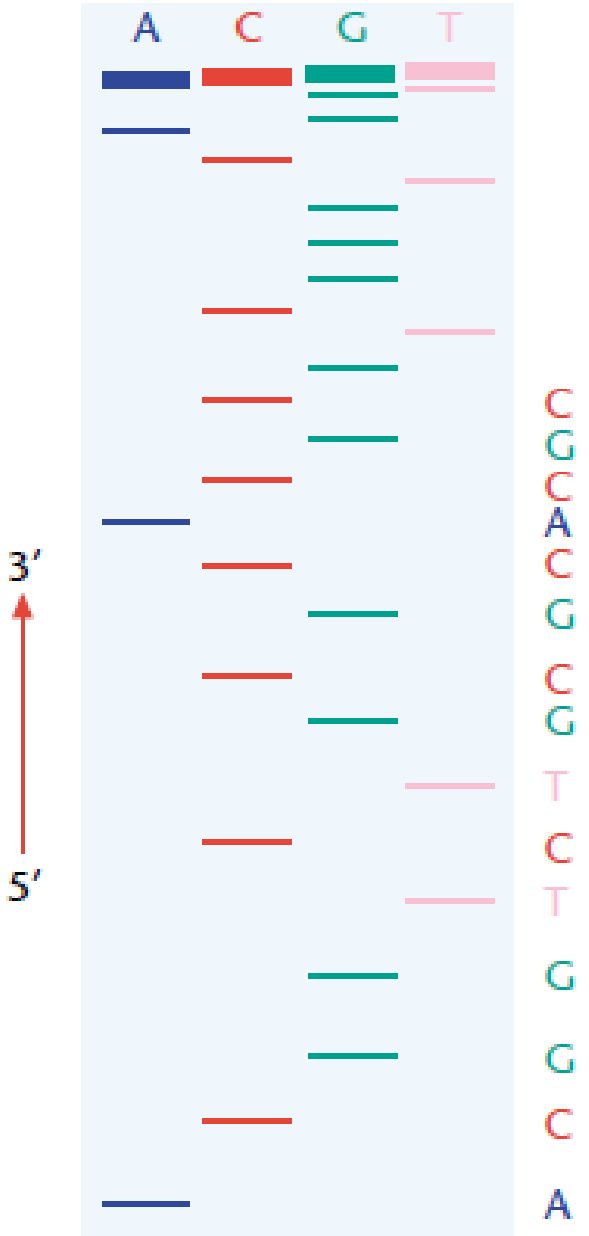


Figure 8 Dideoxy sequencing ladder.

Bands become too close to call near the top of the gel

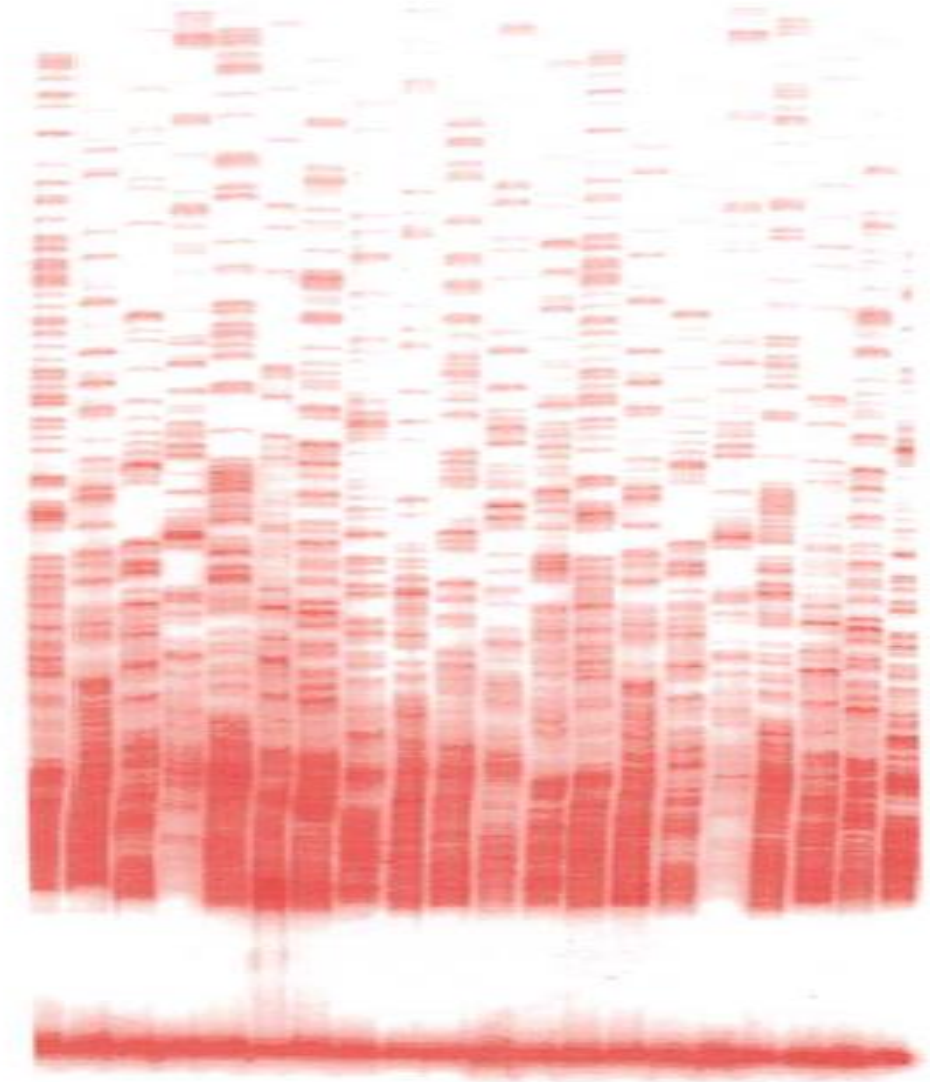


Figure 9 Autoradiograph.

The Illumina method was invented by the American company Solexa in 2005 and is now part of the Illumina Company. This method is also known as Solexa sequencing. This method is divided into 4 groups based on which sequencing device is used to determine the sequence of parts:

1- Galle

2-Hiscan SQ

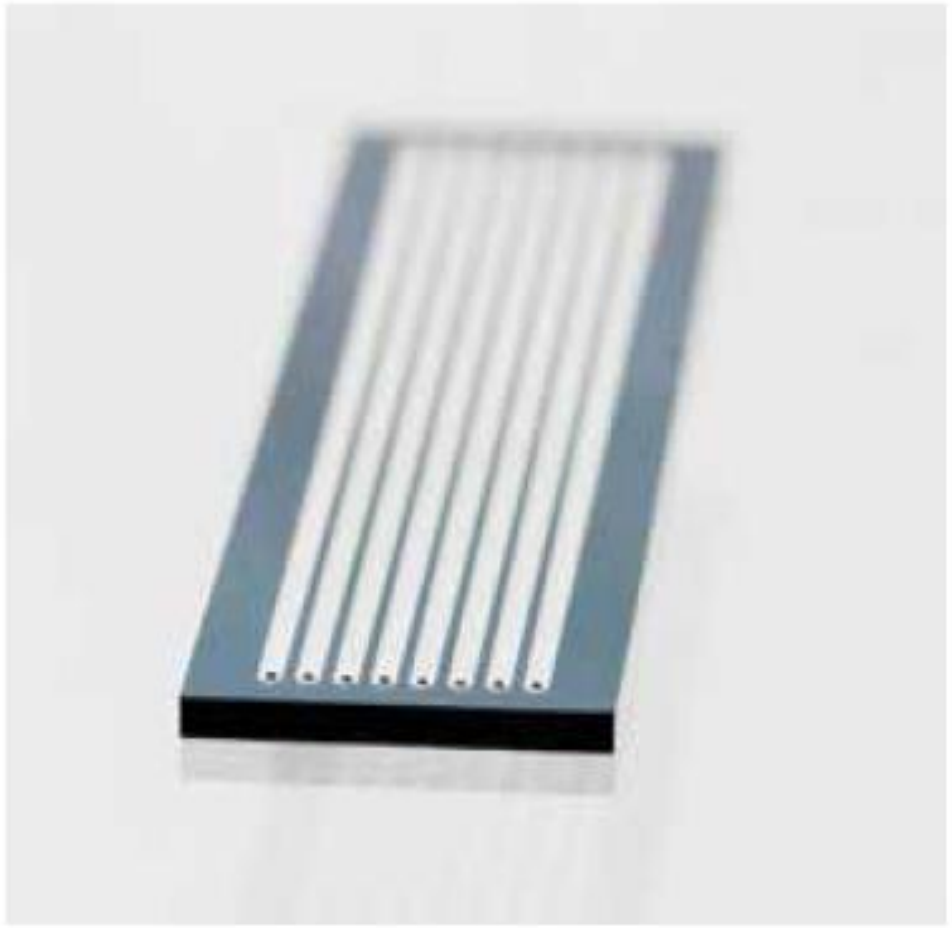
3-GallX

4-Hiseq2000

In all these methods, the 4 main steps of gene library preparation, gene cluster creation, sequencing and data analysis are common to them.

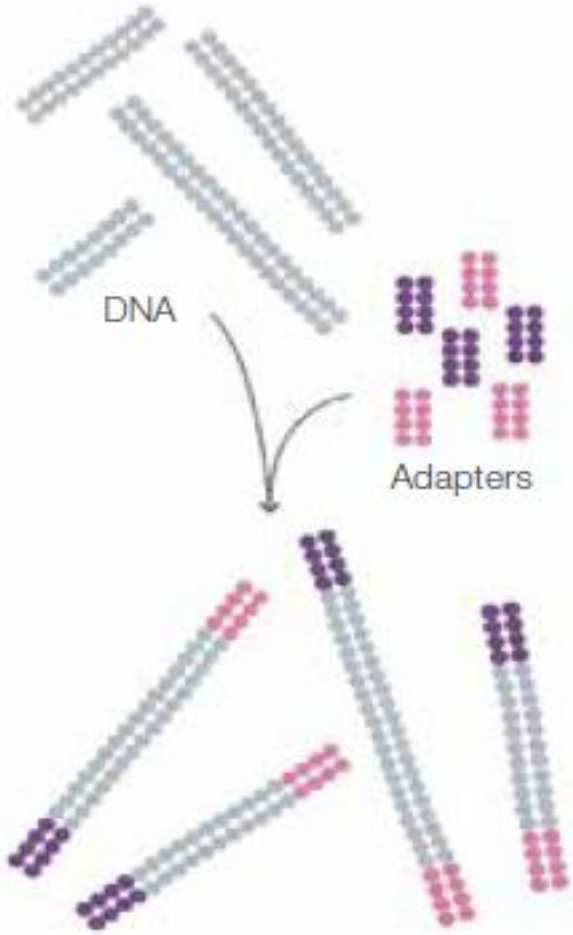
The first archive preparation instructions include random fragments of RNA or DNA samples, gel-based sampling, binding of platform-specific oligonucleotides, expansion of the polymerase chain reaction, and several other steps. In general, a library or sequence archive is prepared by randomly shredding a genomic sample and connecting 3’ or 5’ adapters. Similarly, the signal addition reaction involves a combination of coupling and fractional reactions. These steps dramatically increase the efficiency of the archive preparation process. In general, the steps required to generate a gene library can be considered in four main steps: a) DNA fragmentation or cdNAB b) repair of end ridges and addition of adenine nucleotides to the end of the chain c) Adapto-release binding d) Selection of liganded genomes.

Figure 1: Illumina Flow Cell



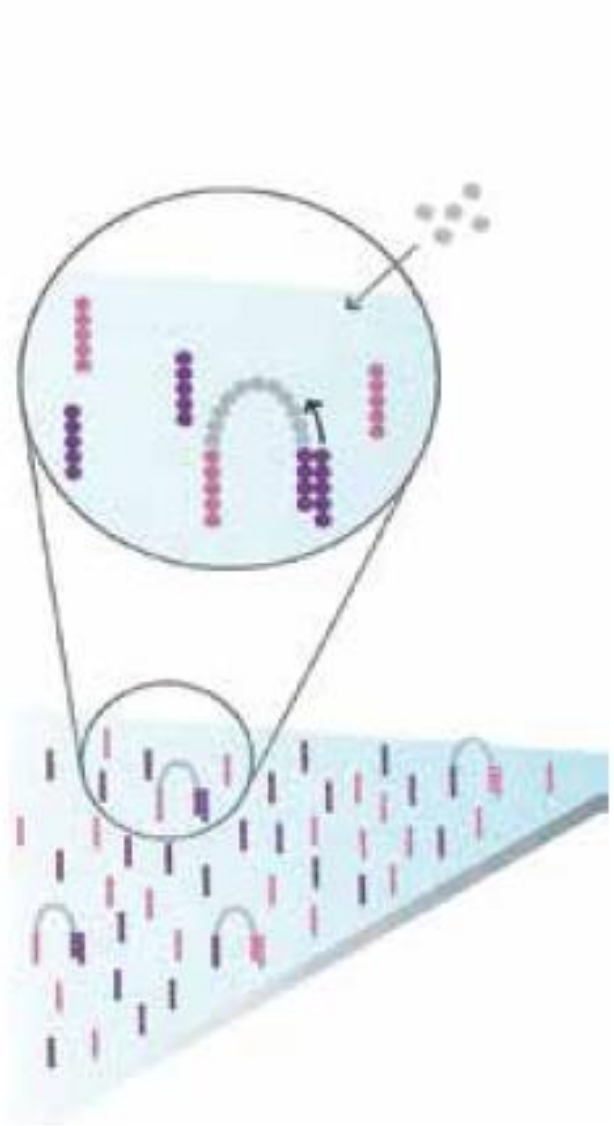
Several samples can be loaded onto the eight-lane flow cell for simultaneous analysis on an Illumina Sequencing System.

Figure 2: Prepare Genomic DNA Sample



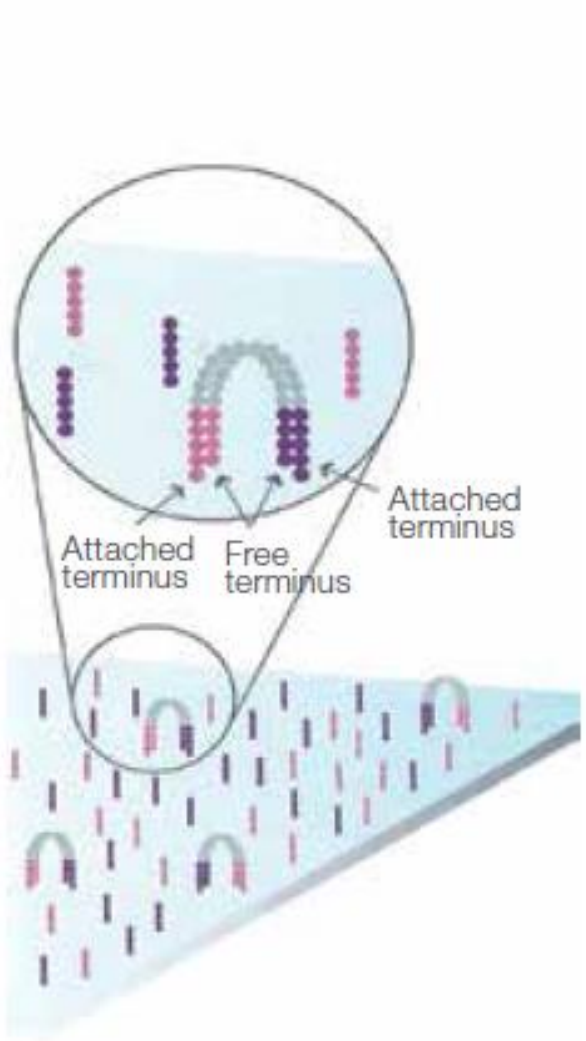
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

Figure 4: Bridge Amplification



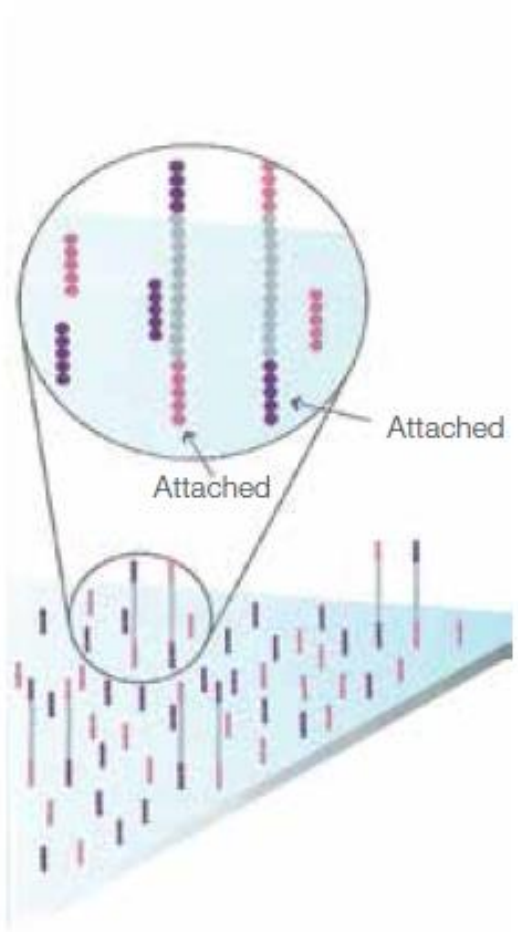
Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.

Figure 5: Fragments Become Double Stranded



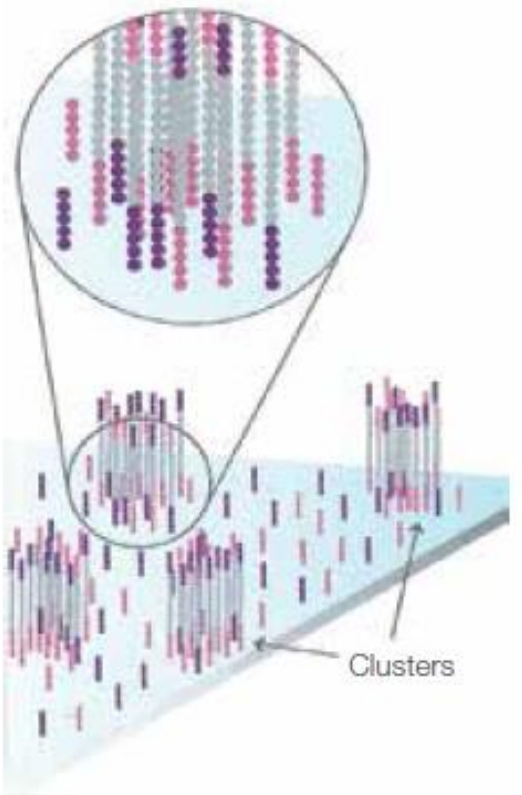
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

Figure 6: Denature the Double-Standed Molecules



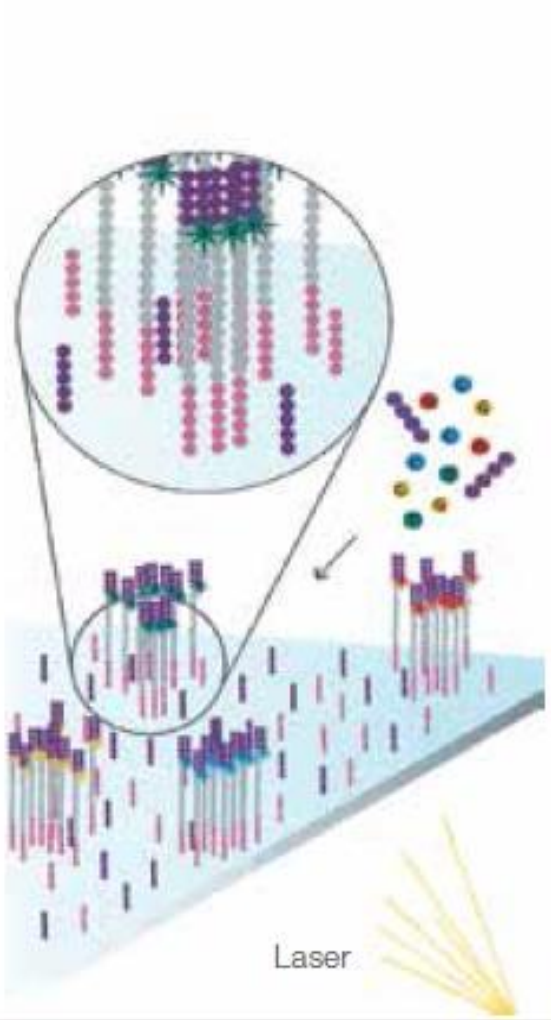
Denaturation leaves single-stranded templates anchored to the substrate.

Figure 7: Complete Amplification



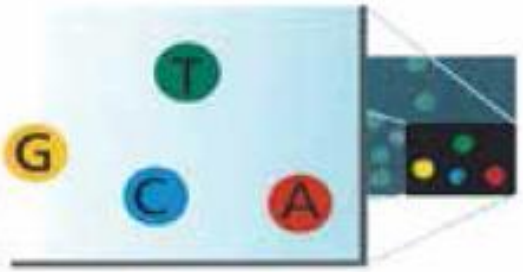
Several million dense clusters of double-stranded DNA are generated in each channel of the flow cell.

Figure 8: Determine First Base



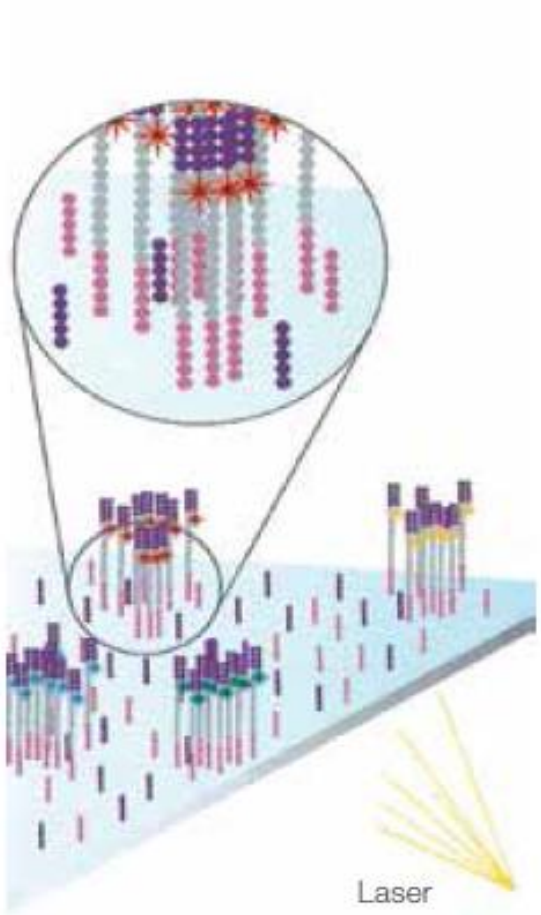
The first sequencing cycle begins by adding four labeled reversible terminators, primers, and DNA polymerase.

Figure 9: Image First Base



After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.

Figure 10: Determine Second Base

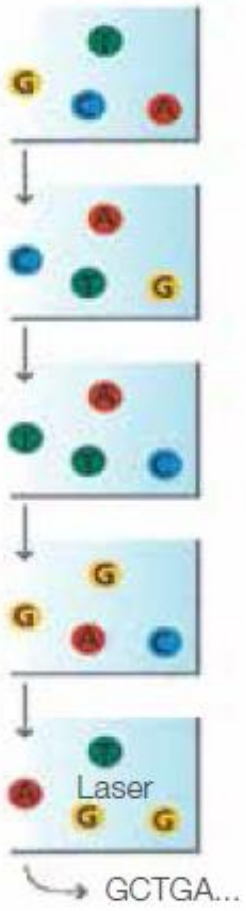


The next cycle repeats the incorporation of four labeled reversible terminators, primers, and DNA polymerase.

Figure 11: Image Second Chemistry Cycle

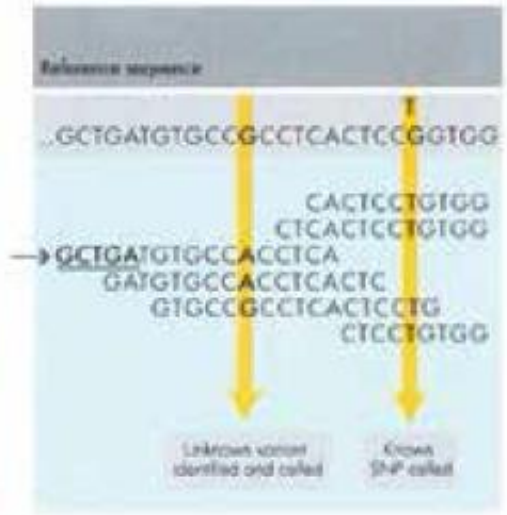
After laser excitation, the image is captured as before, and the identity of the second base is recorded.

Figure 12: Sequencing Over Multiple Chemistry Cycles

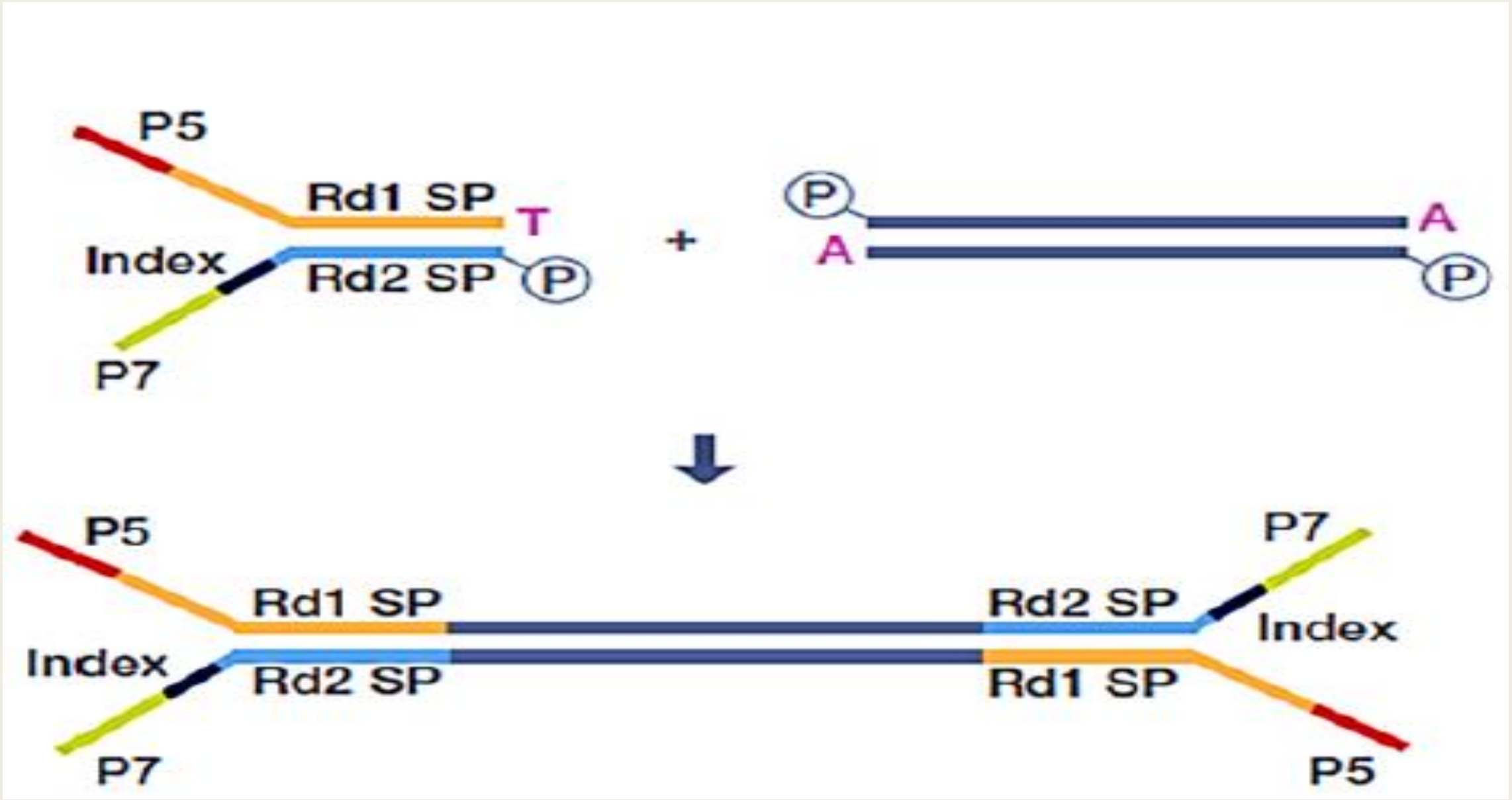


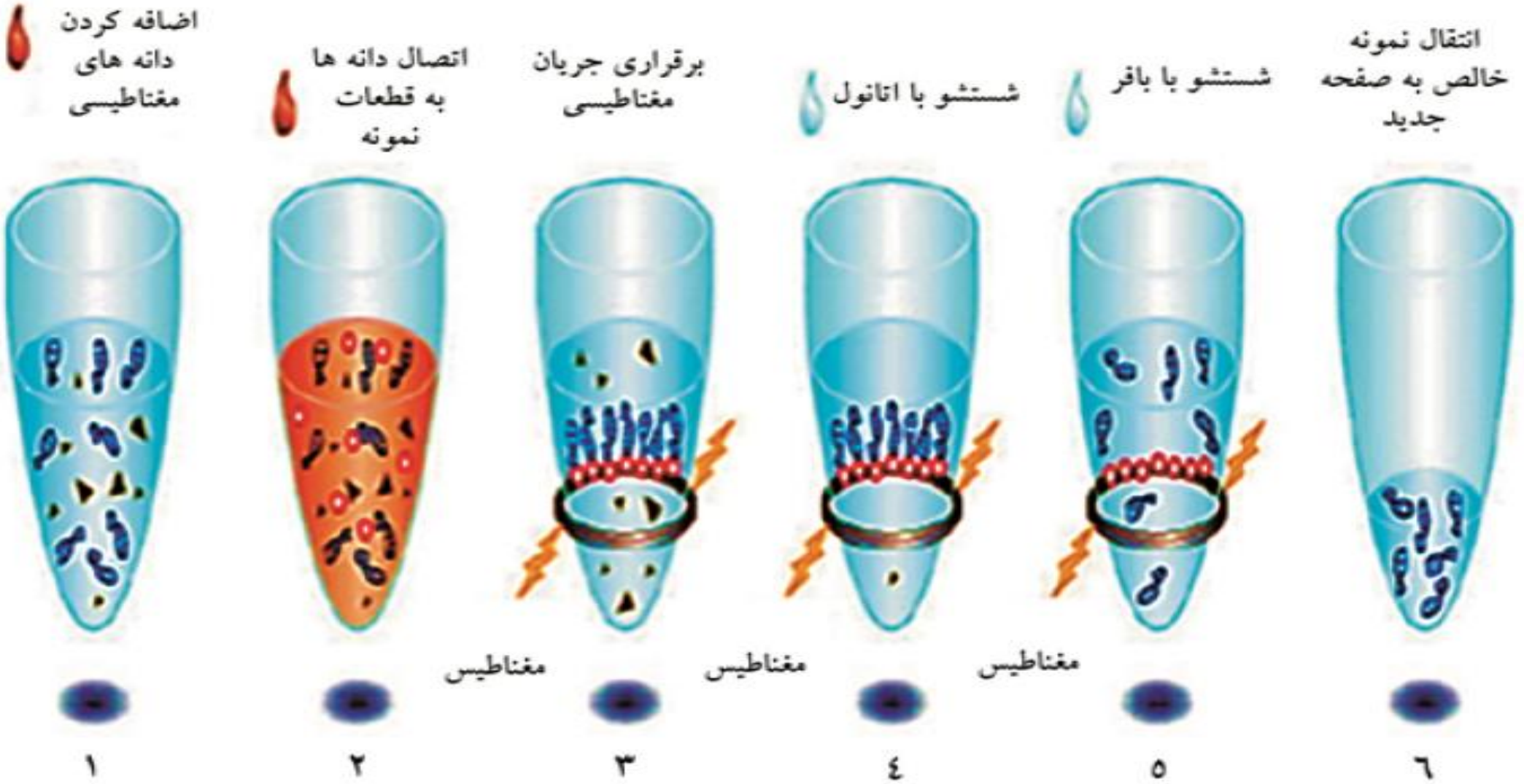
The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time.

Figure 13: Align Data



The data are aligned and compared to a reference, and sequencing differences are identified.


P5
Rd1 SP
T
Index
Rd2 SP
P
P7
+
P
A
A
P
P5
Rd1 SP
Index
Rd2 SP
P7
Rd2 SP
Index
Rd1 SP
P5



شکل ۳) مراحل طی‌شده برای خالص‌سازی قطعات نمونه ژنی