# Introduction to Sequencing

PRESENTED BY: PRAISSY ZEFI J (DTU)

# LEARNING OBJECTIVES

- Understand the principles of sequencing technologies available
- Compare and contrast the NGS technologies – Illumina vs ONT
- Breakdown of the protocol of the NGS technologies
- Understand the necessity of following key steps in NGS
- How to select the right method for your research?

# SEQUENCING TECHNIQUES

## FIRST GENERATION

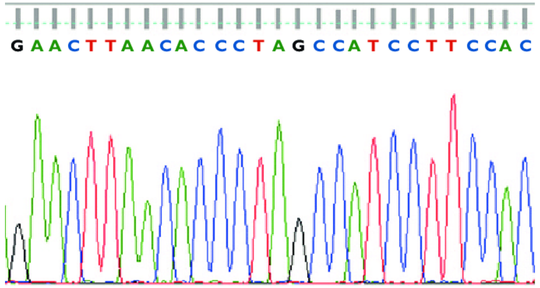
- Sanger Sequencing

## SECOND GENERATION

- Roche/454 Pyrosequencing
- IonTorrent sequencing
- Illumina sequencing
- SoLiD sequencing

## THIRD GENERATION

- Pacific Biosciences
- Oxford Nanopore technology



*Collinson, Jon Martin. (2018). CSI: Birding – DNA-based identification of birds. British Birds. 110.*

# SEQUENCING TECHNIQUES

## FIRST GENERATION

**Sanger Sequencing**

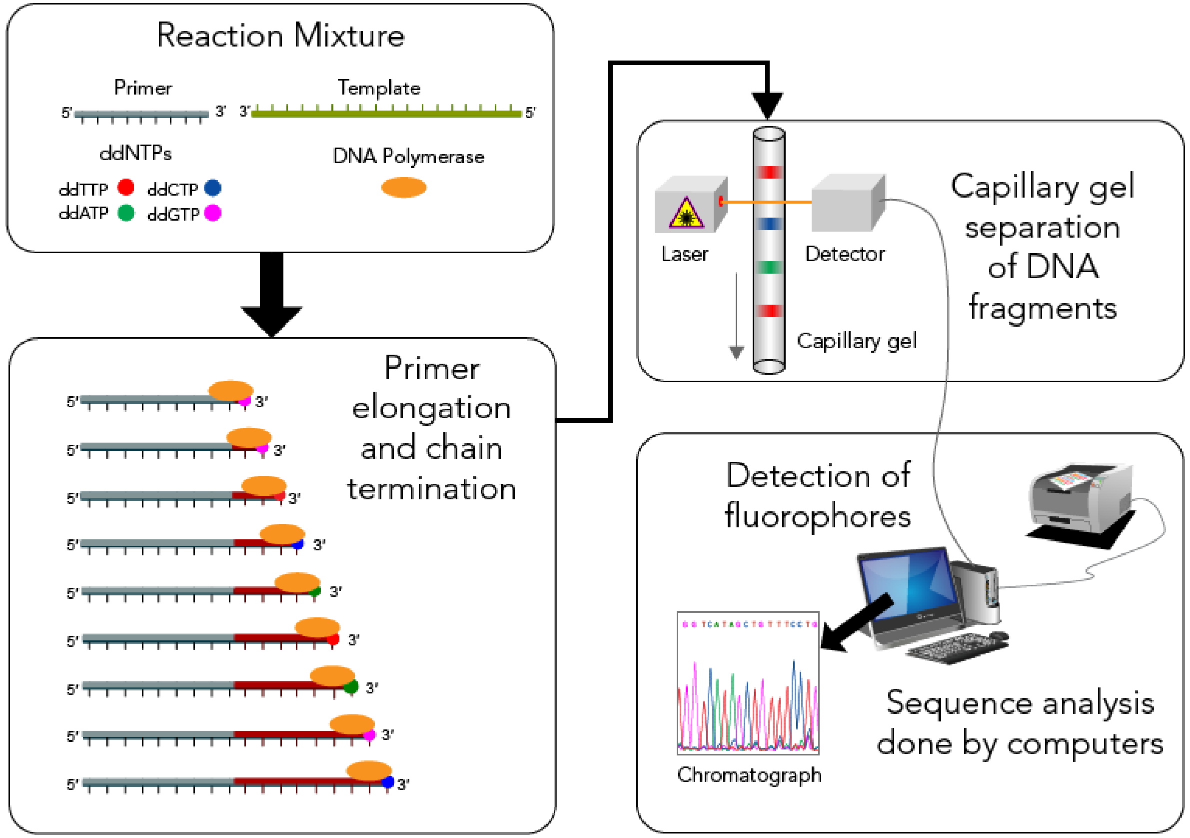
- **“chain termination method” or “dideoxy termination sequencing”**
- **selective incorporation of chain-terminating dideoxynucleotides (ddNTP) by DNA polymerase**
- **99.99% accuracy – “gold standard”**

**Requirements:**

**DNA, primer, DNA polymerase, DNA nucleotides(dNTP) , ddNTP**

**3 main steps :**

1) **DNA sequence for chain termination**
2) **Size separation by gel electrophoresis**
3) **Gel analysis and determination of DNA sequence**



**Advantage: Highly accurate**

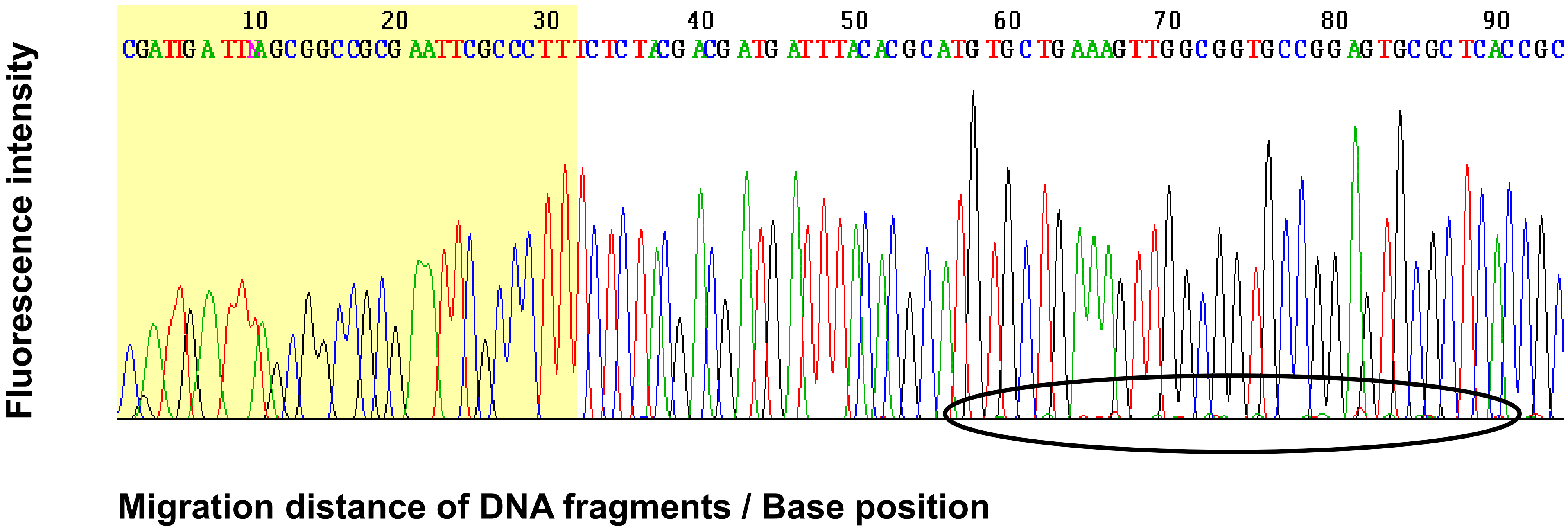
**Disadvantage: Low throughput**

**Cost: ~$0.50–$5 per reaction**

# SEQUENCING TECHNIQUES

## FIRST GENERATION

- Specific point mutations → clinical diagnostics
- Targeted sequencing
- Validation of NGS results



SnapGene

https://sites.biology.unt.edu/~jajohnson/Chromatogram_Interpretation

# SEQUENCING TECHNIQUES

## SECOND GENERATION


MiniSeq System
Power and simplicity for targeted sequencing.


MiSeq Series
Small genome and targeted sequencing.

Illumina Sequencing → "Sequencing by synthesis"

Utilize fluorescently labelled nucleotides that possess reversible terminators

3 main steps
- Library preparation
- Sequencing
- Data analysis


NextSeq Series
Everyday genome, exome transcriptome sequencing, and more.


HiSeq Series
Production-scale genome, exome, transcriptome sequencing, and more.

**Advantages:**
- ✓ High throughput
- ✓ High accuracy
- ✓ cost effective

**Disadvantages:**
- ✓ Limited real-time analysis
- ✓ Requires PCR

**Cost:~$100–$1,000 for whole genome (human)**

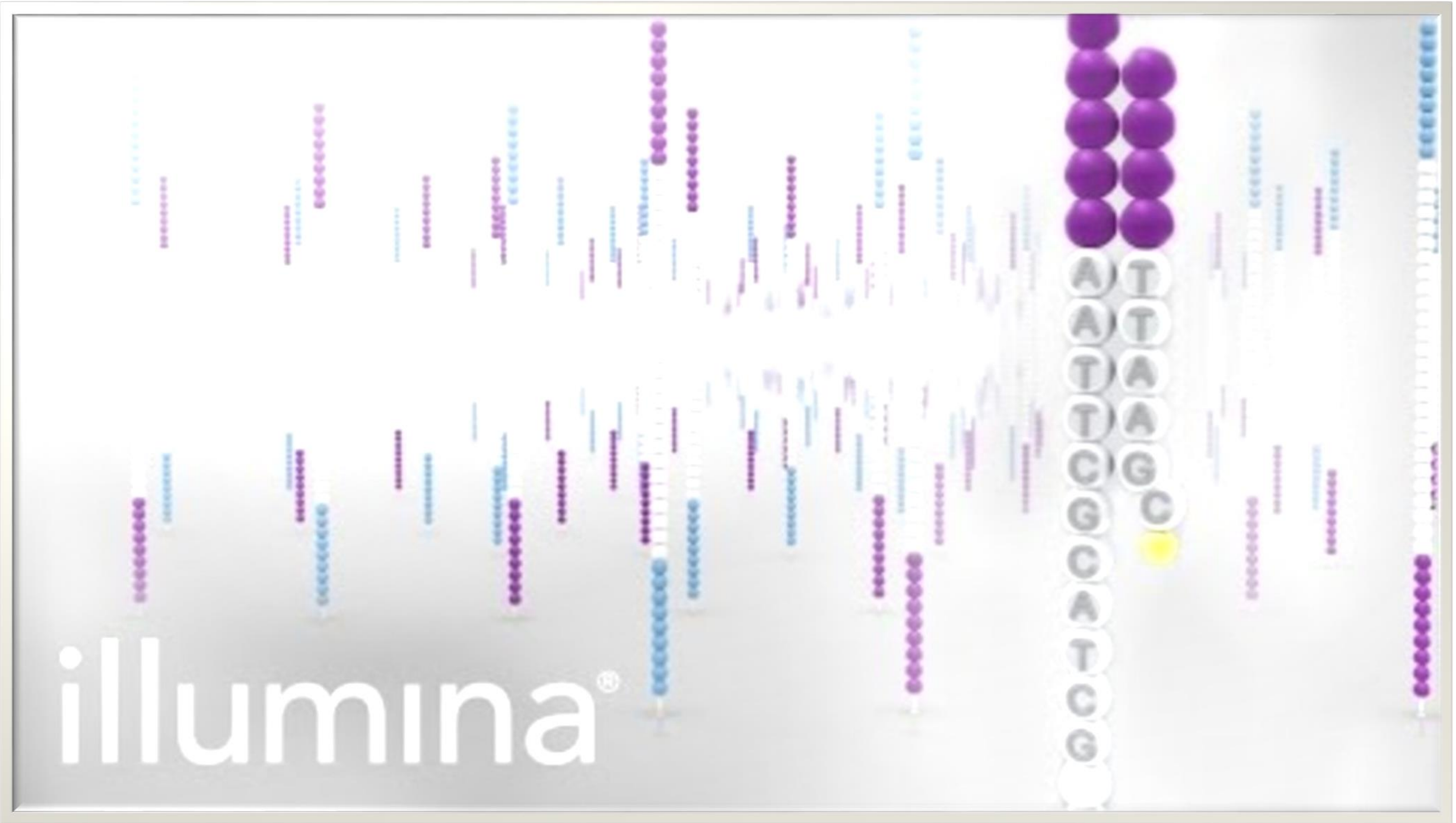
**Sequencing yield varies per sequencer: anywhere between 1.2 Gb to 540 Gb**

**Sequencing time varies per sequencer: anywhere between 4 to 44 hours**


HiSeq X Series
Population- and production-scale human whole-genome sequencing.


NovaSeq Series
Population- and production-scale genome, exome, transcriptome sequencing, and more.

# ILLUMINA SEQUENCING TECHNOLOGY



https://www.youtube.com/watch?v=fCd6B5HRaZ8

# SEQUENCING TECHNIQUES

## THIRD GENERATION

- vague definition
- Reading long fragments of DNA without amplification (PCR)
- “massive parallel sequencing”
- Nanopore = 1 nm diameter

**Advantages:**

✓ No PCR, real time sequencing, complex genome analysis

**Disadvantages:**

✓ High error, high cost

Cost:~$500–$2,000+ per sample (depends on coverage and tech)

# OXFORD NANOPORE TECHNOLOGY

https://www.youtube.com/watch?v=E9-Rm5AoZGw&t=5s

# GENERAL STRATEGY - MASSIVE PARALLEL SEQUENCING

SAMPLE COLLECTION

DNA EXTRACTION

DNA QUALITY CONTROL

***Illumina***

DNA BARCODING*

LIBRARY PREPARATION

LIBRARY QC

PCR

SEQUENCING

***ONT***

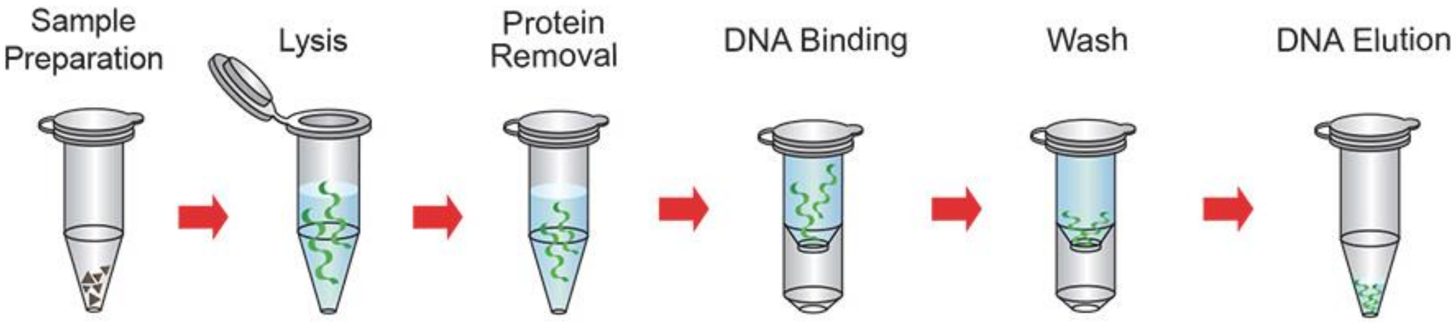
DNA BARCODING*

LIBRARY PREPARATION

SEQUENCING

# DNA EXTRACTION – importance

- Purity and quality of DNA
- Concentration
- Preservation of DNA integrity
- High throughput and Efficiency
- Facilitates analysis

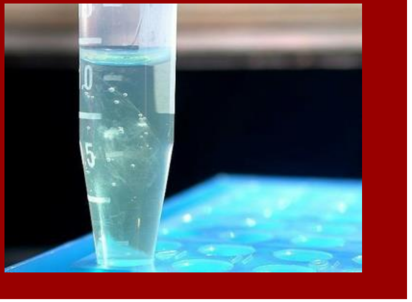


**PRINCIPLE:**

**lysis of cellular and nuclear membranes → DNA separation from impurities, proteins, and other substances**

**General Stages of DNA Extraction are:**

- **Cell dissolution → lysis of cell and the nucleus to extract DNA into the buffer**
- **Precipitation → removing impurities and proteins from sample**
- **Purification → purify DNA sample ready to be used**



# DNA extraction kits

Qiagen DNeasy kit

Spin column

&

96-well high throughput formats

Promega Wizard Genomic DNA kit

- Lysis
- RNase digestion
- Salt precipitation
- Genomic DNA extraction

Invitrogen MagMAX

Lysis → heat + chemical + mechanical disruption

Zymo Research Quick DNA Fungal/Bacterial kit

Recovers genomic DNA >= 40 kb

Sigma-Aldrich GenElute Bacterial Genomic DNA Kit

Gram +ve → lysozyme

Silica-based system + microspin format

# DNA extraction kits – for HMW DNA

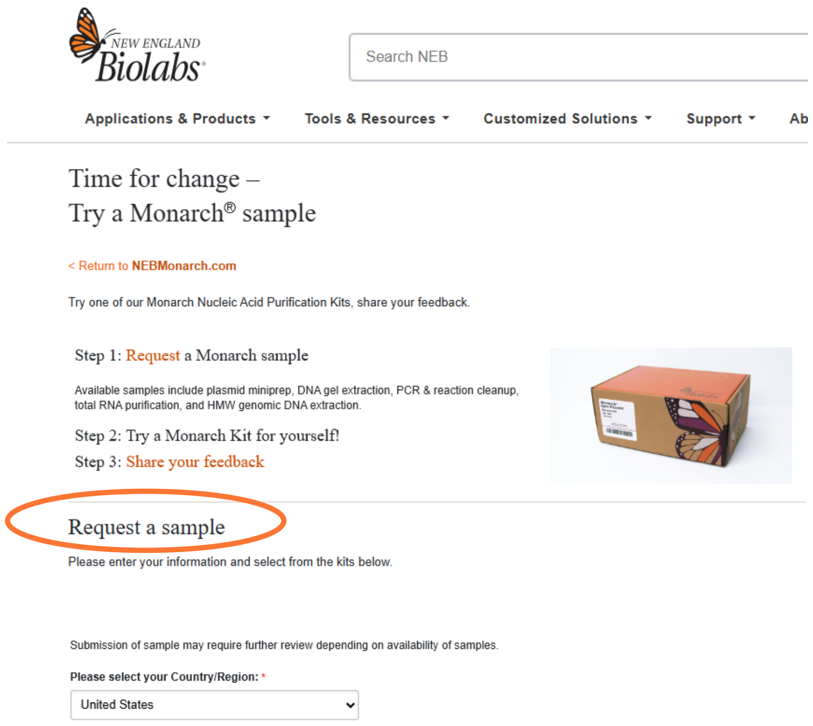
Monarch DNA extraction kit

Monarch Genomic DNA purification kit → silica spin column-based process

Monarch HMW DNA extraction kit → glass bead-based process

https://www.neb.com/en/products/t3050-monarch-hmw-dna-extraction-kit-for-cells-and-blood

- Gentle cell lysis + tunable fragment length generation → precipitation of extracted DNA
- Standard protocol : DNA size = 50 – 250 kb
- High yield of purified HMW DNA



Time for change –
Try a Monarch® sample

< Return to NEBMonarch.com

Try one of our Monarch Nucleic Acid Purification Kits, share your feedback.

Step 1: Request a Monarch sample

Available samples include plasmid miniprep, DNA gel extraction, PCR & reaction cleanup, total RNA purification, and HMW genomic DNA extraction.

Step 2: Try a Monarch Kit for yourself!

Step 3: Share your feedback

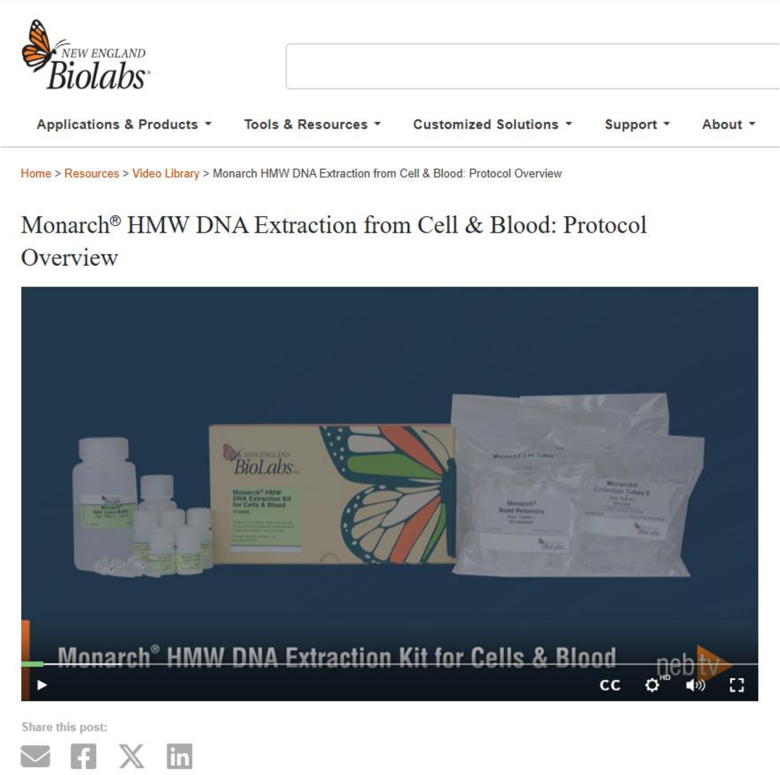
Request a sample

Please enter your information and select from the kits below.

Submission of sample may require further review depending on availability of samples.

Please select your Country/Region: *

United States



Monarch® HMW DNA Extraction from Cell & Blood: Protocol Overview

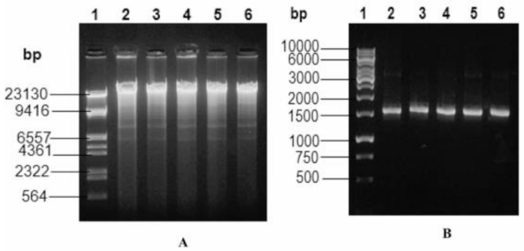
# QUALITY CHECK OF EXTRACTED DNA

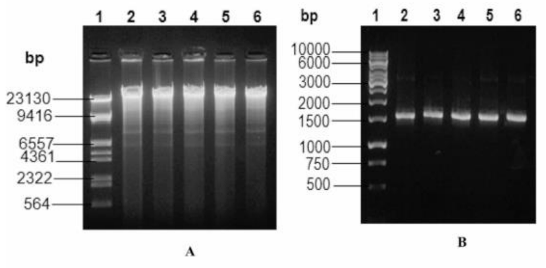
**Reasons**:

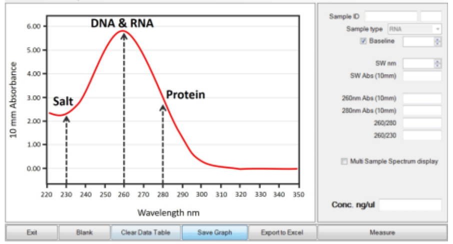
- Accurate downstream results
- Detect contamination
- Quantify DNA correctly
- Ensure DNA integrity
- Avoid wasting data and money

**MAIN METHODS**

- Gel electrophoresis
- NanoDrop (Spectrophotometry)
- Qubit Fluorometer (Flurometric measurement)







Sample ID
Sample type RNA
☑ Baseline
SW nm
SW Abs (10mm)
260nm Abs (10mm)
280nm Abs (10mm)
260/280
260/230
☐ Multi Sample Spectrum display
Conc. ng/ul

Gel Electrophoresis

Nanodrop spectrophotometer

Qubit fluorometer

## DNA INTEGRITY

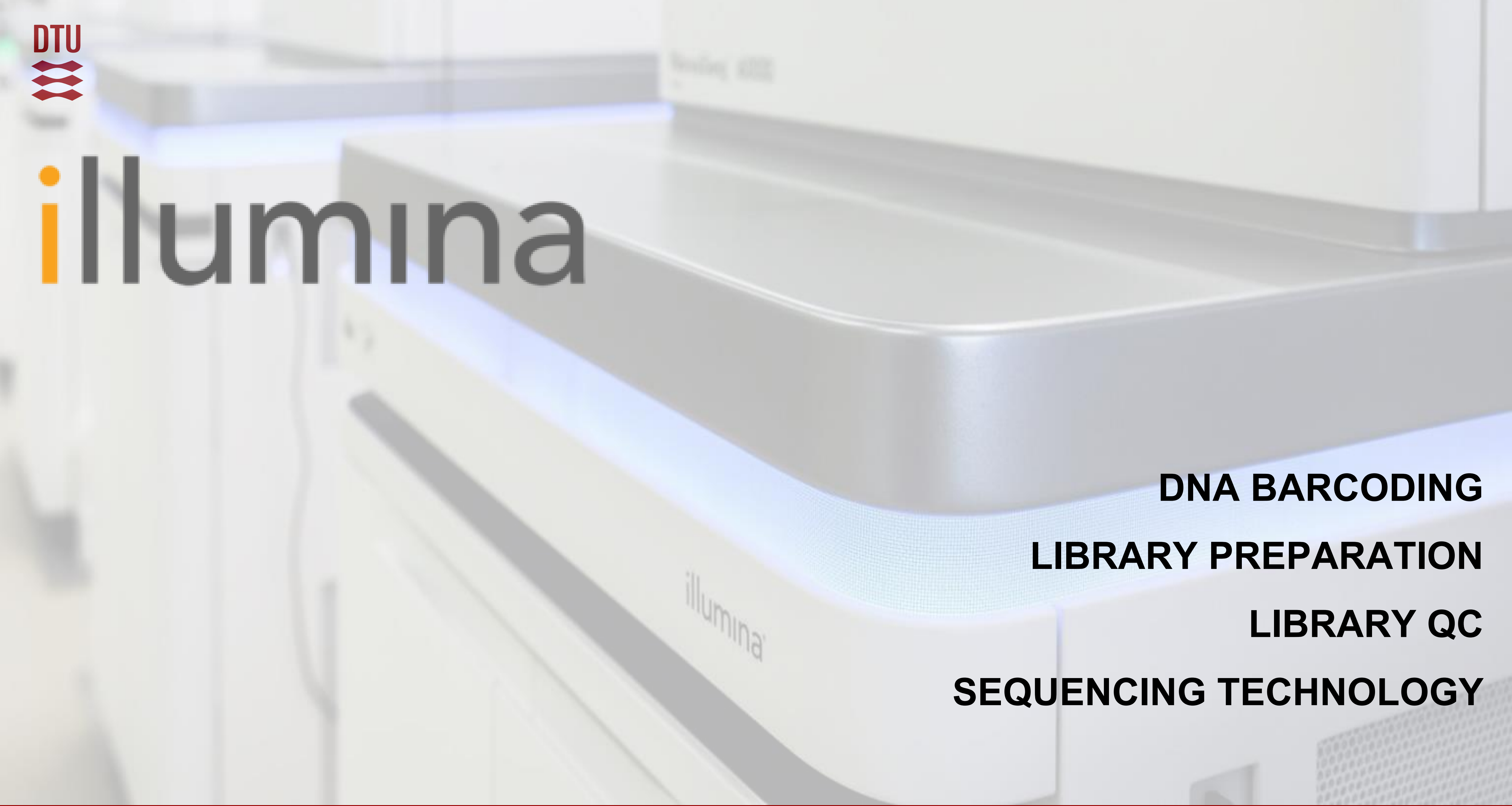
- **Used for High molecular weight DNA → ONT**
- **Size and degradation**

## DNA QUALITY

- **Screens for contamination**
- **260/280 ratio ~ 1.8 – 2.0**
- **260/230 ratio ~ 2.0 – 2.2**

## DNA QUANTIFICATION

- **Fluorescence based quantification instrument**
- **Libraries for DNA sequencing**
- **Dyes → dsDNA**
- **Fluorescence intensity is proportional to amount of DNA present**
- **Accurate and sensitive**



**DNA BARCODING**

**LIBRARY PREPARATION**

**LIBRARY QC**

**SEQUENCING TECHNOLOGY**

# DNA BARCODING – ILLUMINA SEQUENCING

**Barcoding (aka “indexing”) → allows multiple samples to be pooled and sequenced in the same run**

**→ then later demultiplexed (separated) bioinformatically**

**Illumina**

**short DNA sequences (6–12 bp)**

**Each sample receives a unique combination of:**

- **i7 index (Index 1)**
- **i5 index (Index 2)**

NEBNext Multiplex oligos for Illumina

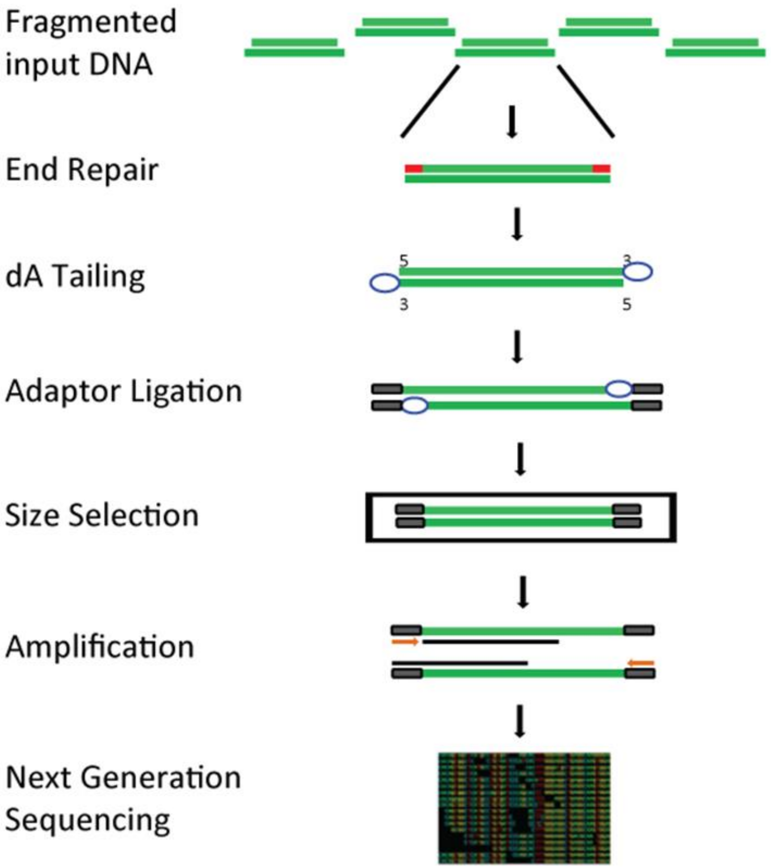
NEXTERA XT Index kit

# LIBRARY PREPARATION – ILLUMINA SEQUENCING

**converting nucleic acid samples into a library of uniformly sized, adapter-ligated DNA fragments**

## IMPORTANCE :

- ✓ Ensures High-Quality Sequencing Data
- ✓ Confirms Proper Fragmentation
- ✓ Verifies Adapter Ligation Efficiency
- ✓ Detects Contamination or Over-Representation of Certain Sequences
- ✓ Quantifies Library Concentration



**Workflow**

- Bacterial DNA
- fragmented DNA
- Quality control DNA
- End repaired DNA
- dA tailing
- Adapter-ligated fragments
- Size selected fragment library
- PCR amplified library

# COMMON LIBRARY PREPARATION KITS

Nextera XT

- **Tagmentation-based (enzymatic fragmentation + adapter tagging in one step)**
- **Low input (~ 1ng)→ Fast**
- **Ideal for bacterial genomes, plasmids, small genomes**

Nextera DNA Flex

- **Combined DNA fragmentation and adapter ligation into a single 15 minutes reaction**
- **No need for library quantification**
- **1-500 ng input DNA required (small genomes)**
- **100-500 ng input DNA required (large genomes)**

# LIBRARY PREPARATION – QUALITY CONTROL

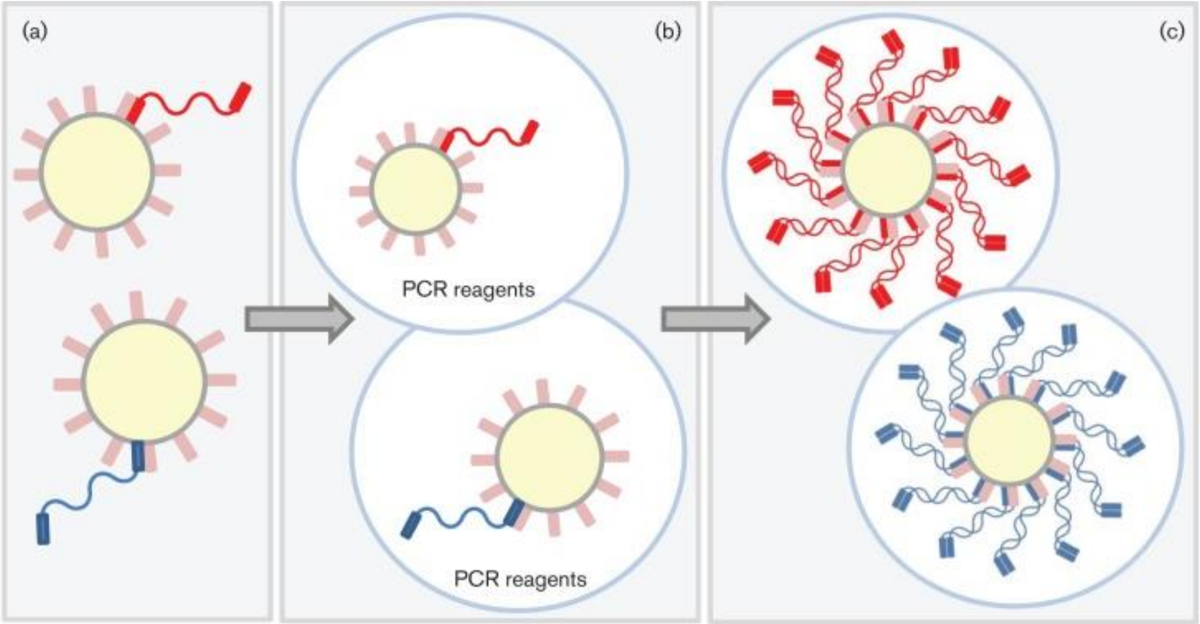
## GENERAL CHALLENGES

- Sample quality
- Biases and artifacts
- Adapter dimer formation
- Automation

## SUCCESSFUL EXECUTION

i. Optimize input amounts
ii. Maintain clean workspaces
iii. Follow protocols carefully
iv. Include controls
v. Stay updated

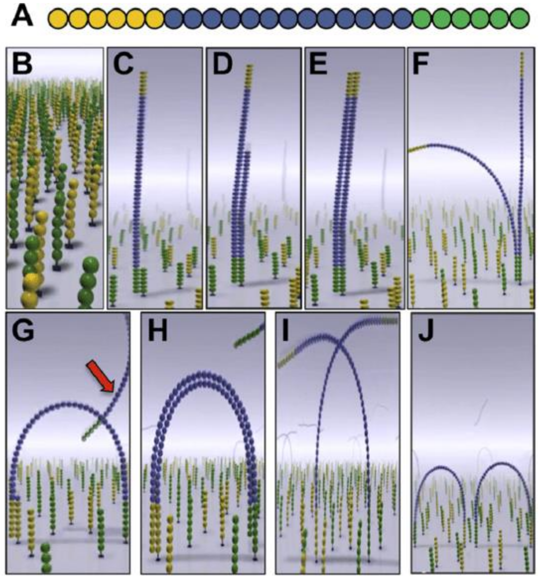
# PCR AMPLIFICATION METHODS



## EMULSION PCR

**Roche 454, Ion Torrent, some early NGS platforms**

A. Adapters are used to capture single molecules of template onto microbeads by primer hybridization.
B. Beads are incorporated into a carefully controlled emulsion, in which each bubble constitutes a microreactor containing DNA template, primer and reagents for PCR.
C. Following amplification, each bead is coated with clonally amplified molecules.



## BRIDGE PCR

**Ilumina platforms (e.g., HiSeq, NextSeq, NovaSeq, MiSeq)**

A. DNA construct with adaptors in yellow and green; sequence in blue.
B. dense lawn of adaptor complementary sequences (primers)
C. DNA constructs bind randomly to the primers attached to the flow cell.
*D. and* E ) DNA elongates along the bound DNA template.
F. The de novo synthesized sequence dissociates from the template strand.
G. The de novo synthesized sequence binds to an adjacent primer, generating a bridge, and the template strand is removed ( red arrow).
H. Bridge amplification takes place.
I. Both strands dissociate.
J. Non-attached extremes bind to new primers and a new bridge amplification cycle starts

# ILLUMINA – SEQUENCING TECHNOLOGY

**MiniSeq**

7.5 Gb
~ 5- 24 hr
2 × 150 bp

**MiSeq**

15 Gb
~ 5.5 - 56 hr
2 × 300 bp

**NextSeq**

30 Gb
~ 4 - 15 hr
2 × 300 bp

**HiSeq**

120 Gb
~ 4 - 15 hr
2 × 300 bp

**NovaSeq 6000**

3 Tb
~ 13 - 44 hr
2 × 250 bp

**iSeq 100**

1.2 Gb
~ 9.5 – 19 hr
2 × 150 bp

**DNA BARCODING**

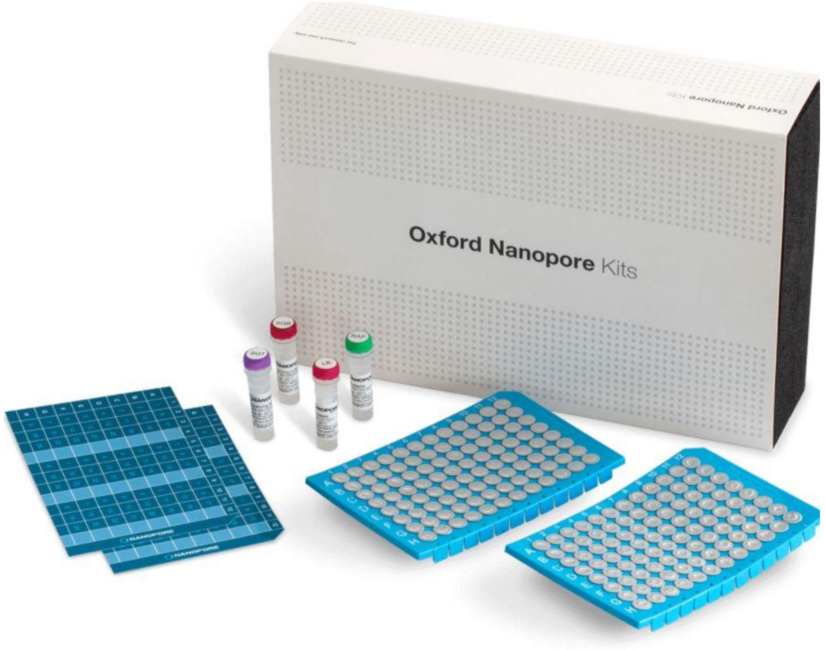
**LIBRARY PREPARATION**

**SEQUENCING**

# DNA BARCODING - ONT

**ONT**

- **barcodes are longer (~20–30 bp) → ligated or tagmented onto DNA fragments.**
- **part of special adapter sequences used during library prep**
- **Demultiplexing → Guppy, qcat, or Porechop**

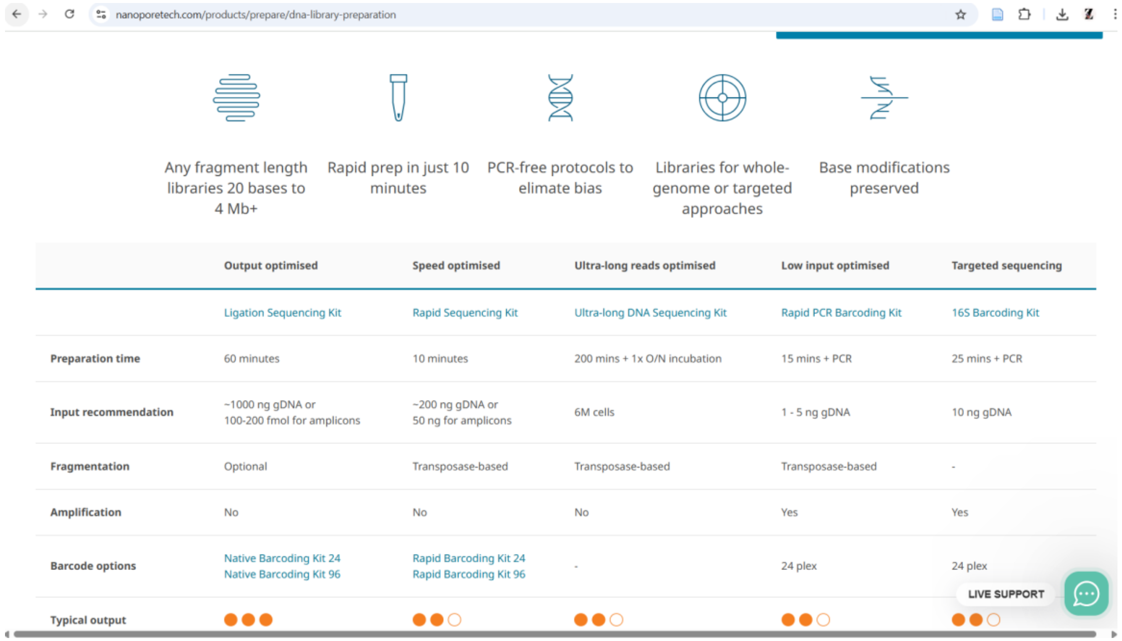


SQK-RBK1114.24

- want to do 16S sequencing
- interested in genus-level bacterial identification
- a method of amplifying the entire ~1,500 bp 16S rRNA gene from extracted gDNA

# LIBRARY PREPARATION – ONT

https://store.nanoporetech.com/us/sample-prep.html



Any fragment length libraries 20 bases to 4 Mb+

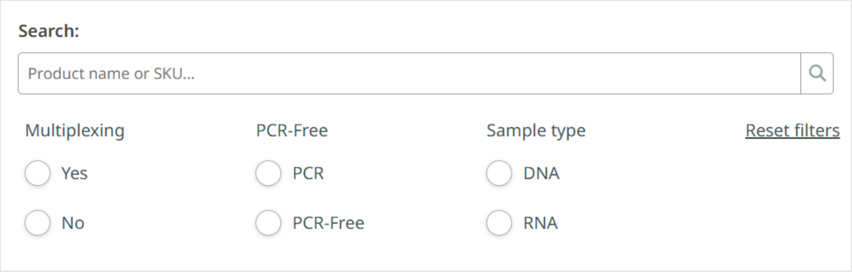
Rapid prep in just 10 minutes

PCR-free protocols to elimate bias

Libraries for whole-genome or targeted approaches

Base modifications preserved

| | Output optimised | Speed optimised | Ultra-long reads optimised | Low input optimised | Targeted sequencing |
|---|---|---|---|---|---|
| | Ligation Sequencing Kit | Rapid Sequencing Kit | Ultra-long DNA Sequencing Kit | Rapid PCR Barcoding Kit | 16S Barcoding Kit |
| **Preparation time** | 60 minutes | 10 minutes | 200 mins + 1x O/N incubation | 15 mins + PCR | 25 mins + PCR |
| **Input recommendation** | ~1000 ng gDNA or<br>100-200 fmol for amplicons | ~200 ng gDNA or<br>50 ng for amplicons | 6M cells | 1 - 5 ng gDNA | 10 ng gDNA |
| **Fragmentation** | Optional | Transposase-based | Transposase-based | Transposase-based | - |
| **Amplification** | No | No | No | Yes | Yes |
| **Barcode options** | Native Barcoding Kit 24<br>Native Barcoding Kit 96 | Rapid Barcoding Kit 24<br>Rapid Barcoding Kit 96 | - | 24 plex | 24 plex |
| **Typical output** | ●●● | ●●○ | ●●○ | ●●○ | ●●○ |



Search:

Product name or SKU...

| Multiplexing | PCR-Free | Sample type | Reset filters |
|---|---|---|---|
| ○ Yes | ○ PCR | ○ DNA | |
| ○ No | ○ PCR-Free | ○ RNA | |

**Ligation Sequencing Kit V14 (SQK-LSK114)**

**Ligation Sequencing Kit XL V14 SQK-LSK114-XL**

**Native Barcoding Kit 24 V14 SQK-NBD114.24**

**Ultra-Long DNA Sequencing Kit V14 SQK-ULK114**

**Rapid Barcoding Kit 24 V14 SQK-RBK114.24**

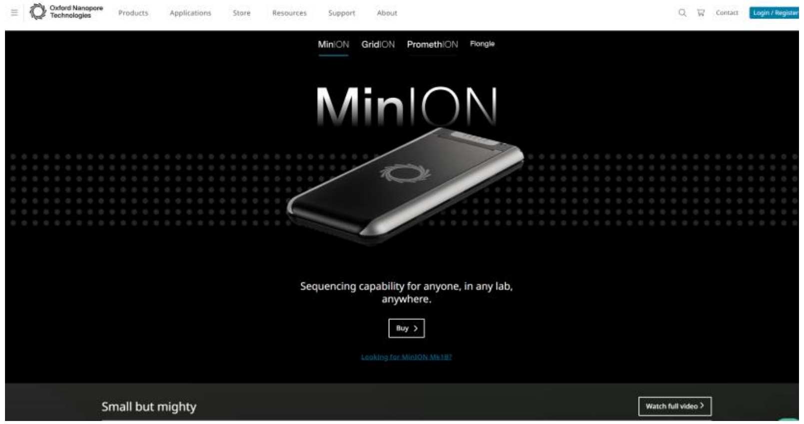
**Rapid Sequencing Kit V14 SQK-RAD114**

# SEQUENCING TECHNOLOGIES - ONT

**Signals translated into DNA sequence using base-calling algorithms**

**Error correction to improve accuracy**

**Sequence data is aligned to a reference genome → genome assembly!**



Sequencing capability for anyone, in any lab, anywhere.

Small but mighty

**MinION**

→ short to ultra long reads

→ Connects to computer USB 3.0 cable

~10 – 30Gb

**GridION**

→ runs 5 MinION flow cells

→ Short to ultra long reads

→ Handles large number of samples

→ ~50 – 150Gb

**PromethION**

→ 24 – 48 parallel flow cells

→ Small-scale to population studies

~ 290+ Gb

**Flongle**

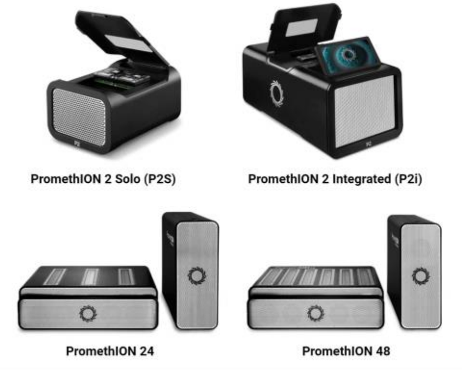
→ Smaller sequencing experiments

→ Run single samples

MinION Products from Oxford Nanopore Technologies

GridION from Oxford Nanopore Technologies

PromethION products from Oxford Nanopore Technologies



Flongle from Oxford Nanopore Technologies

# ILLUMINA

# OXFORD NANOPORE

## ADVANTAGES

- Parallel sequencing → high throughput
- Highly accurate
- Minimum errors relatively
- Cost-effective

## ADVANTAGES

- NO PCR REQUIRED
- provides real-time data
- Compact and portable (eg:minION)
- rapid library preparation

## DISADVANTAGES

- Short reads → assembling complex genomes difficult
- Large volume requires higher computational power
- Overclustering → compromise quality and accuracy

## DISADVANTAGES

- Higher error rate
- Large amounts of input required
- Advanced and complex bioinformatics tools required
- Sensitive to sample preparation and contamination

# THANKYOU FOR LISTENINING, ANY QUESTIONS?

*Kindly drop your email in the chat to receive the certificates*


DTU