The principles of detection by sequencing

12:55-13:40

Iver Nordentoft, PhD

AARHUS
UNIVERSITY
DEPARTMENT OF CLINICAL MEDICINE

NEXT GENERATION SEQUENCING

Sequencing of DNA fragments in parallel

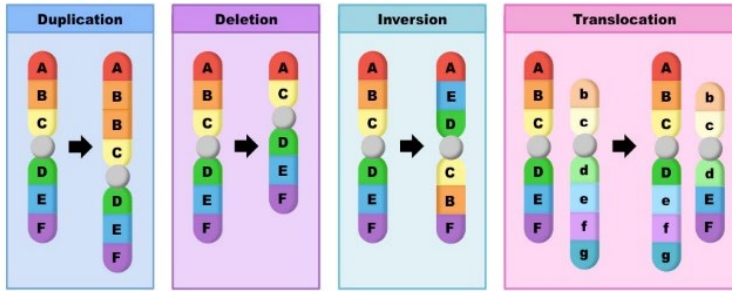
TARGETS

Genetic variations

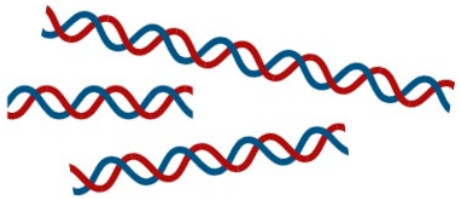
SNV
Single Nucleotide Variant

CNV
Copy Number Variant

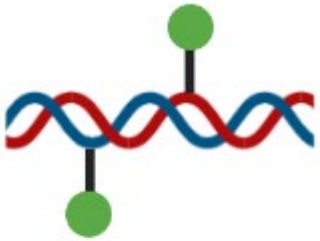
Chromosomal changes



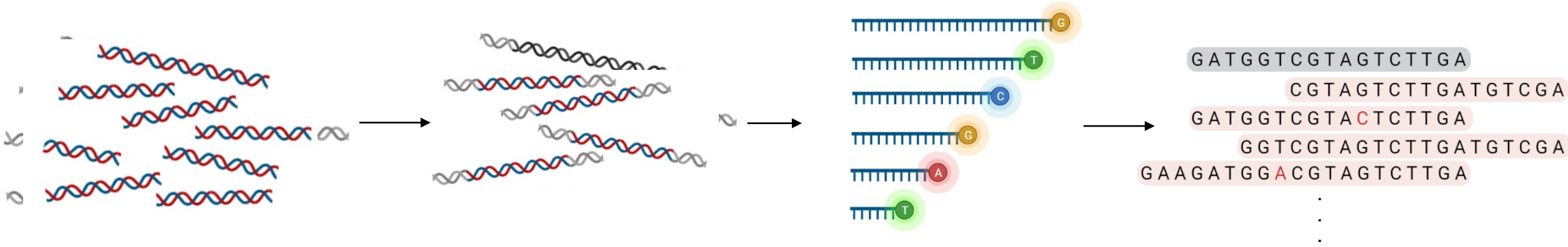
Fragment length patterns



DNA methylation



NGS WORKFLOW



Library preparation, 1 day

- Addition of sequencing adapters to fragmented DNA
- PCR to amplify DNA

(Capture), 2 days

- Focus analysis on regions of interest
- PCR to amplify DNA

Sequencing, 1-2 days

- Sequencing by synthesis
- Emits fluorescent signal

Data analysis

- Align sequence data to reference genome
- Identify mutations

LIBRARY PREPARATION

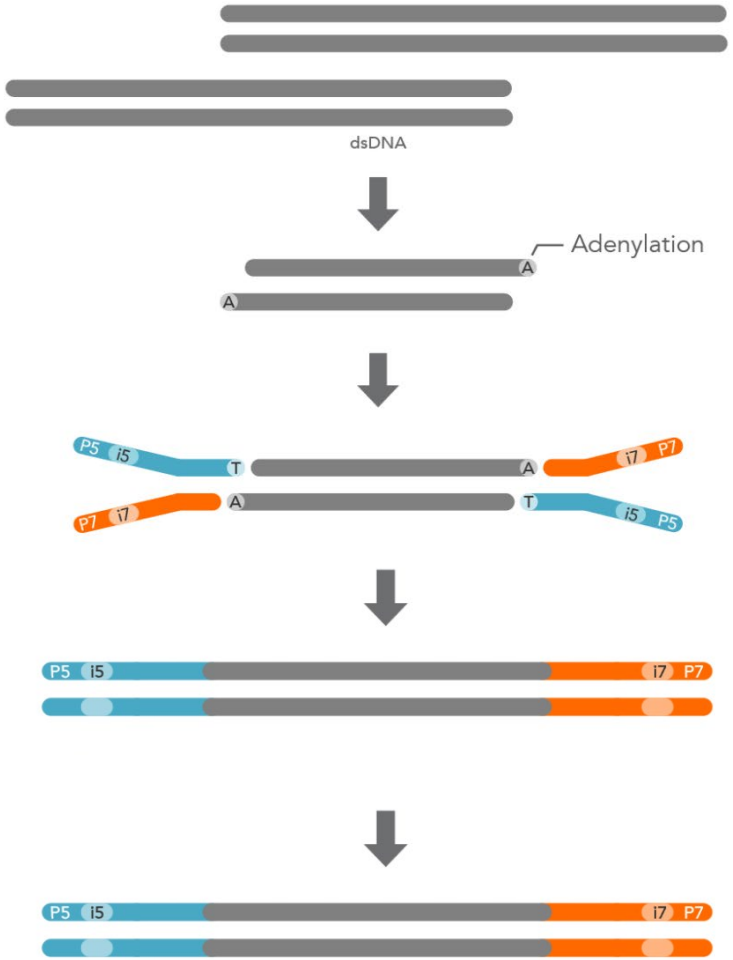
Fragmentation

End repair and A-tailing

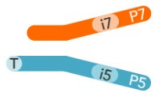
Ligation

PCR amplification

Library



Adapter structure



P5 + P7

Binding to the sequencing flowcell

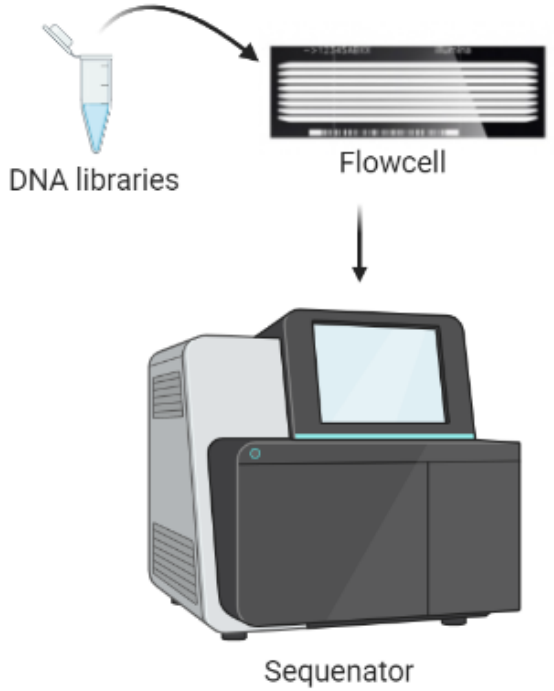
i5 + i7

Sample-specific index sequences

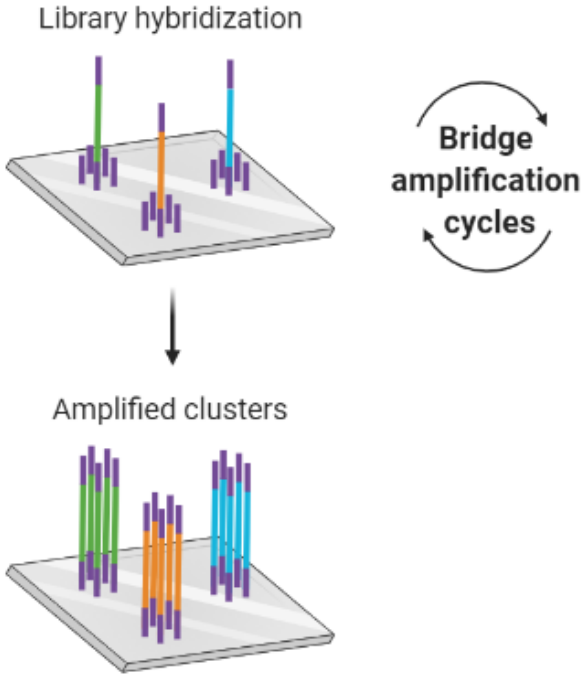
https://www.youtube.com/watch?v=fCd6B5HRaZ8&t=19s

ILLUMINA SEQUENCING

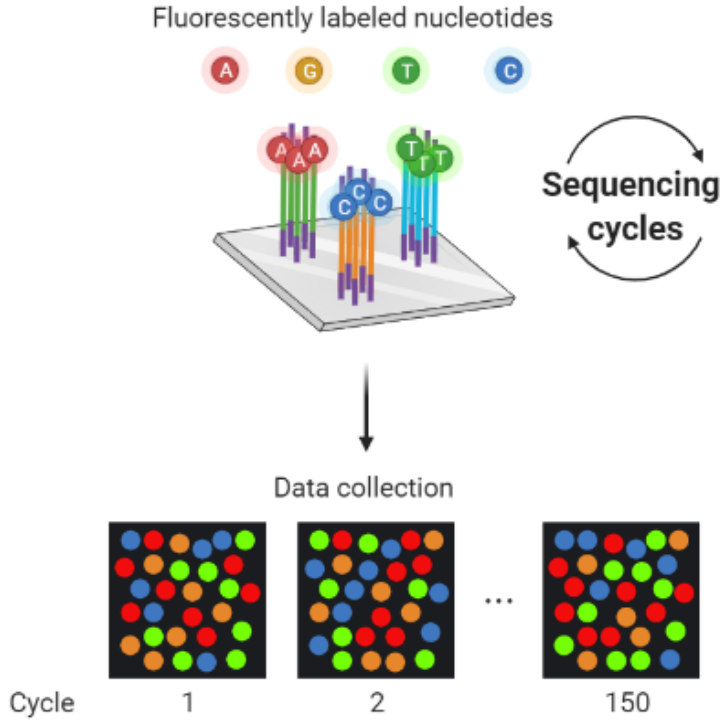
1 DNA libraries loaded onto flowcell



2 DNA library bridge amplification



3 DNA library sequencing

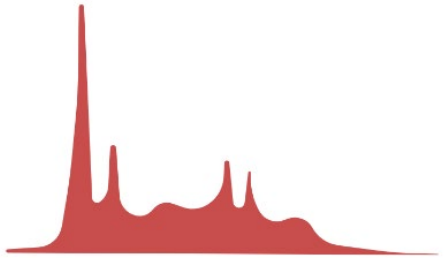


(CAPTURE) Select regions of interest to focus analyses and minimize data output

Whole genome sequencing (no capture)

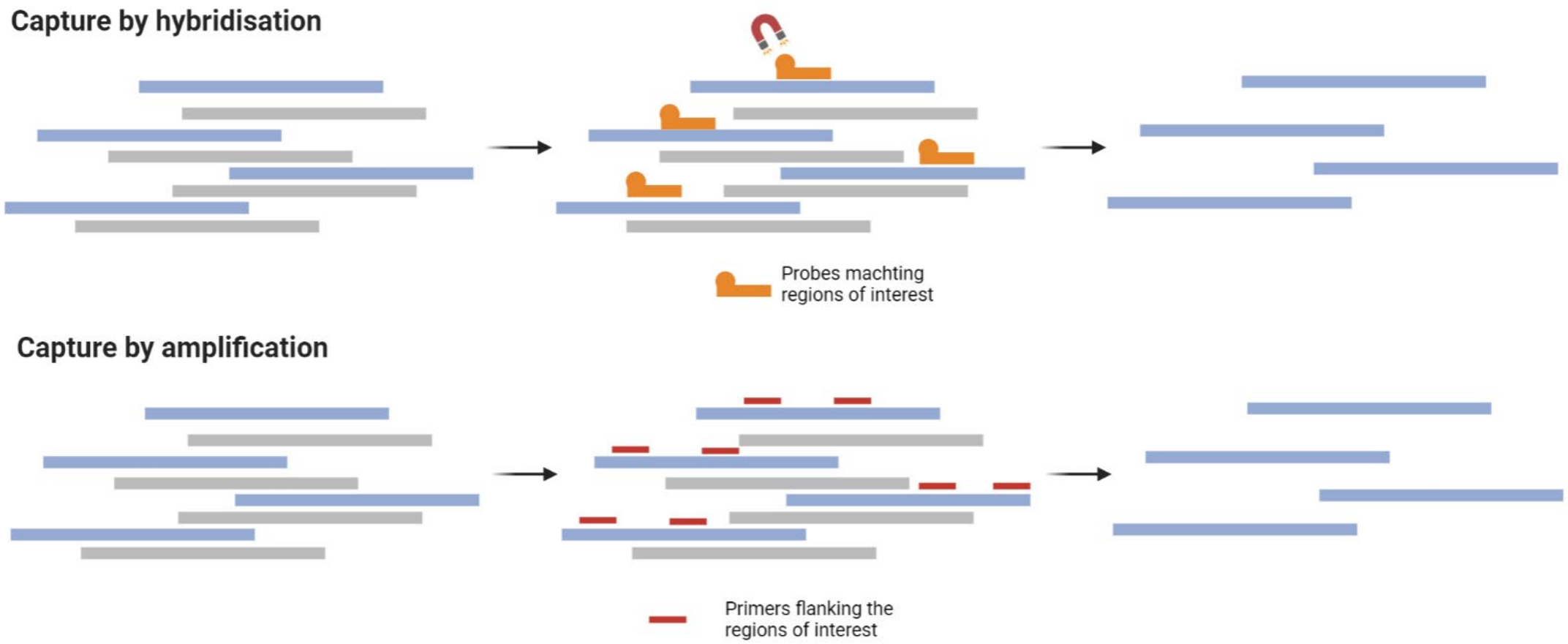
Sequence entire genome

Targeted sequencing (capture regions of interest)



Focus analysis on regions of interest

(CAPTURE) Select regions of interest to focus analyses and minimize data output



(CAPTURE) Select regions of interest to focus analyses and minimize data output

Capture allow for higher sequencing depth at a fixed price

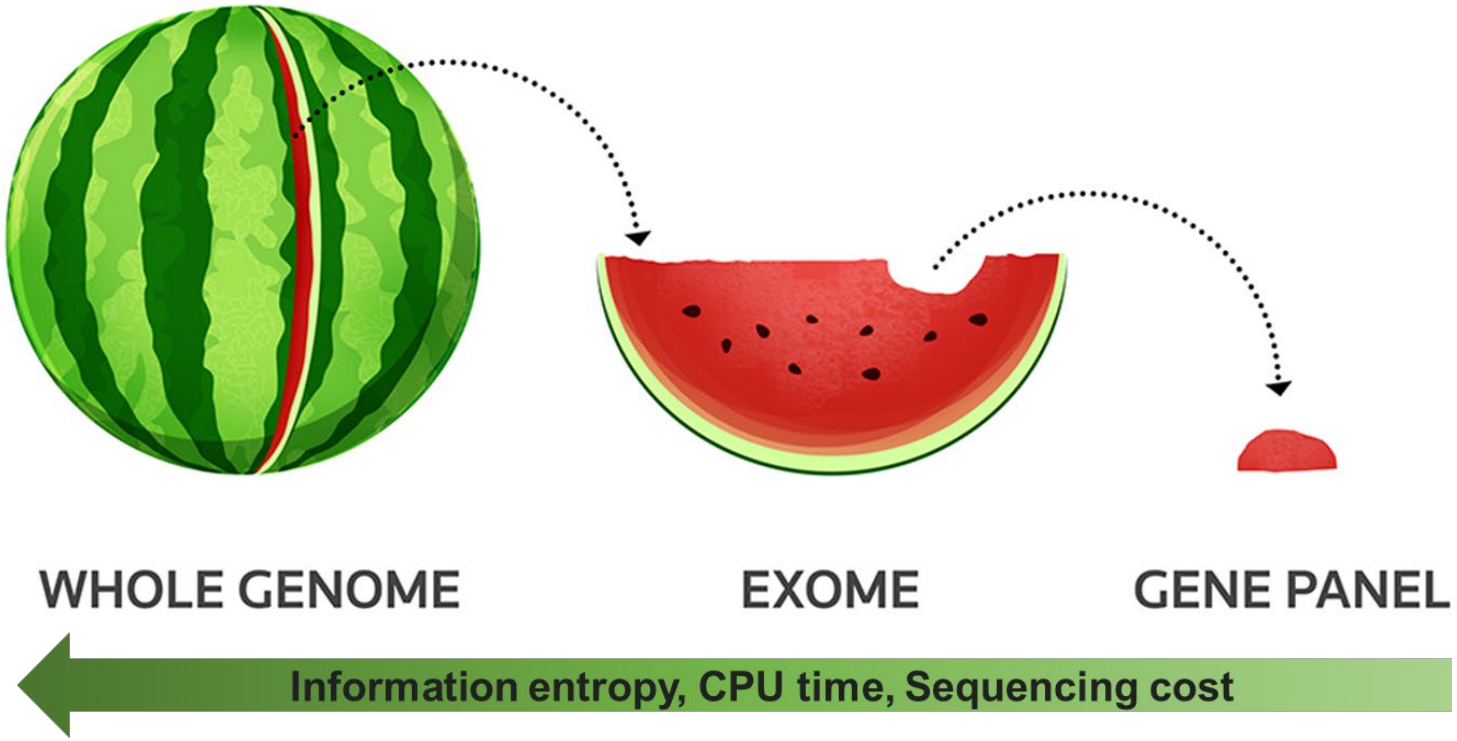
Whole genome sequencing at 1x:
3.3 Gb (100% of genome)

or

Whole exome sequecing at 100x:
33 Mb (1% of genome)

or

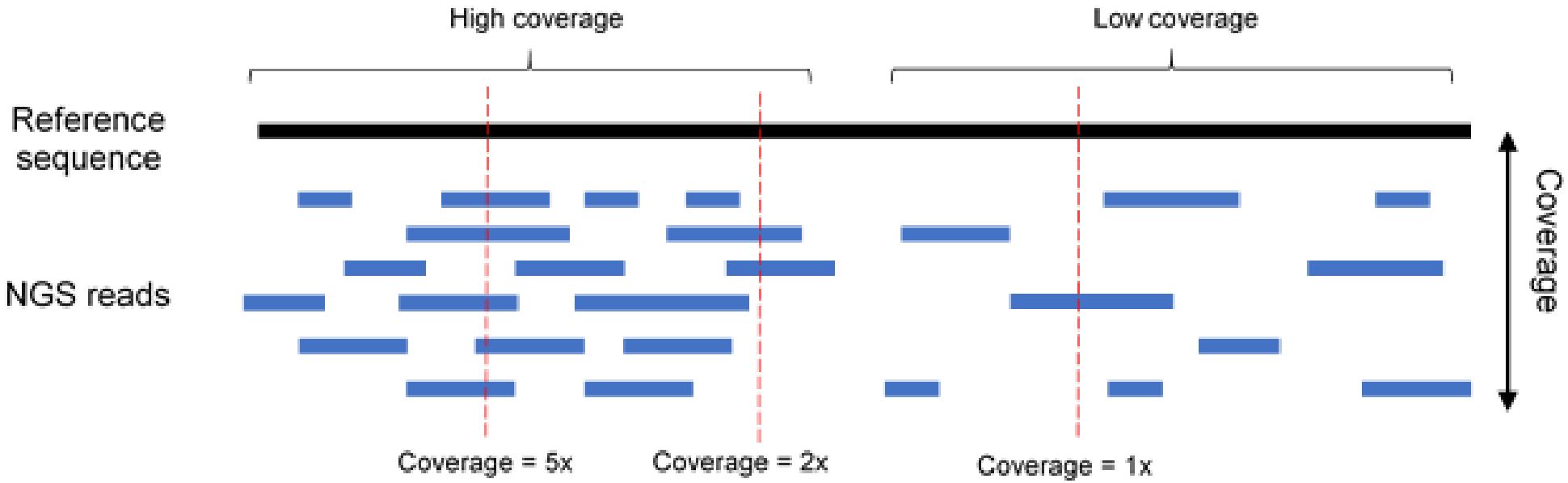
Targeted sequencing at 100.000x coverage :
30,000 – 300,000 bp (0.001- 0.01% of genome)



SEQUENCING COVERAGE

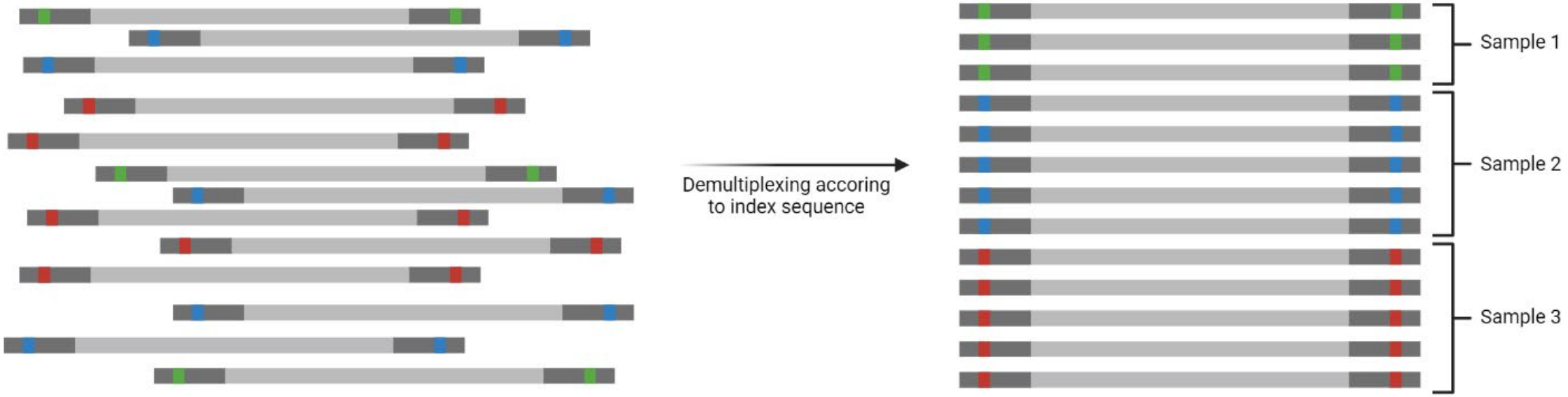
1000x coverage:

Possible to detect 1 in 1000 mutated reads

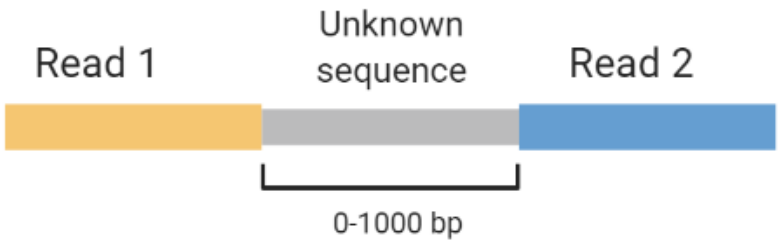


MULTIPLEXING SAMPLES

Adapter with
sample-specific index sequence



PAIRED-END SEQUENCING



- High quality alignment
- Determine fragment length

ERRORS DURING SEQUENCING

Errors can be introduced during:

- Samples storage
- Sample processing, e.g. during PCR
- Sequencing: PCR and optical errors
- Data analysis: e.g. alignment

Action point: "unique molecular identifier" to each DNA fragment

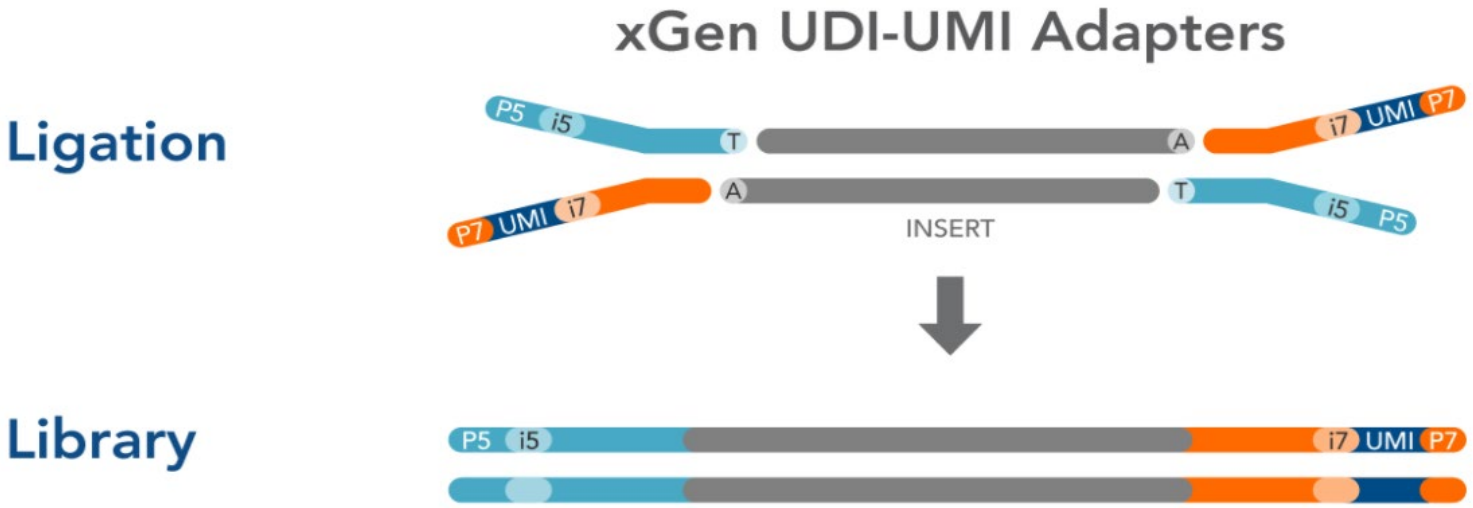
→ Difficult to differentiate errors from true mutations → similar frequency

RISK OF FALSE POSITIVE

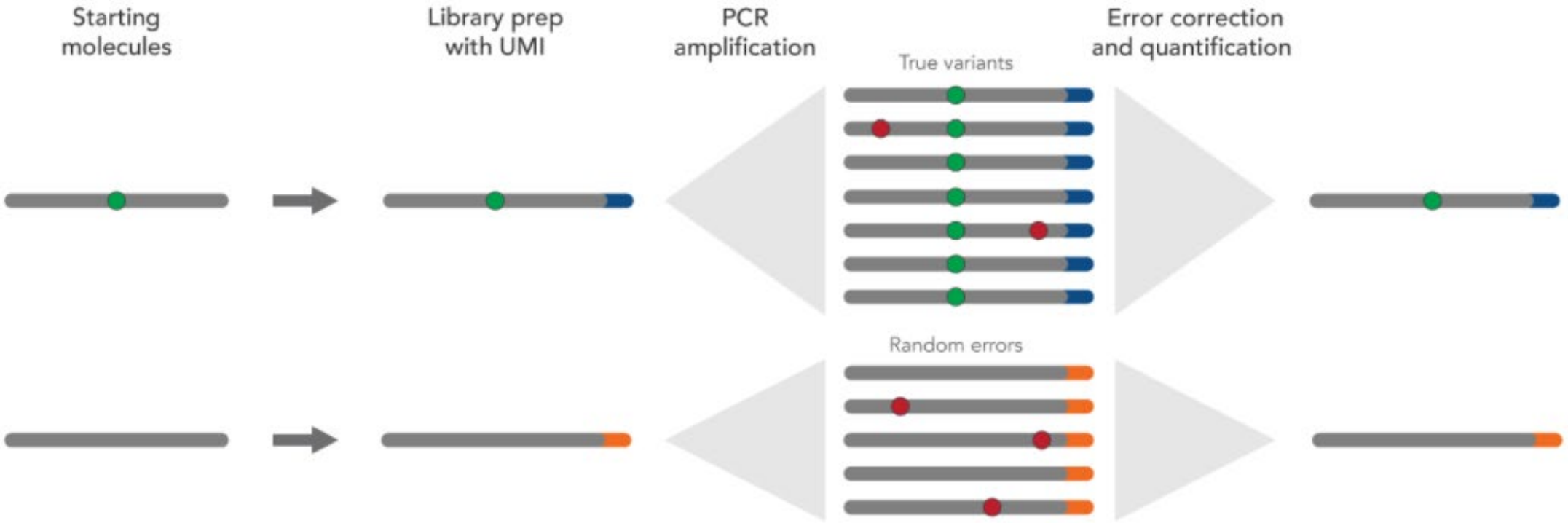
UMI BASED ERROR CORRECTION

UMI = unique molecular identifier

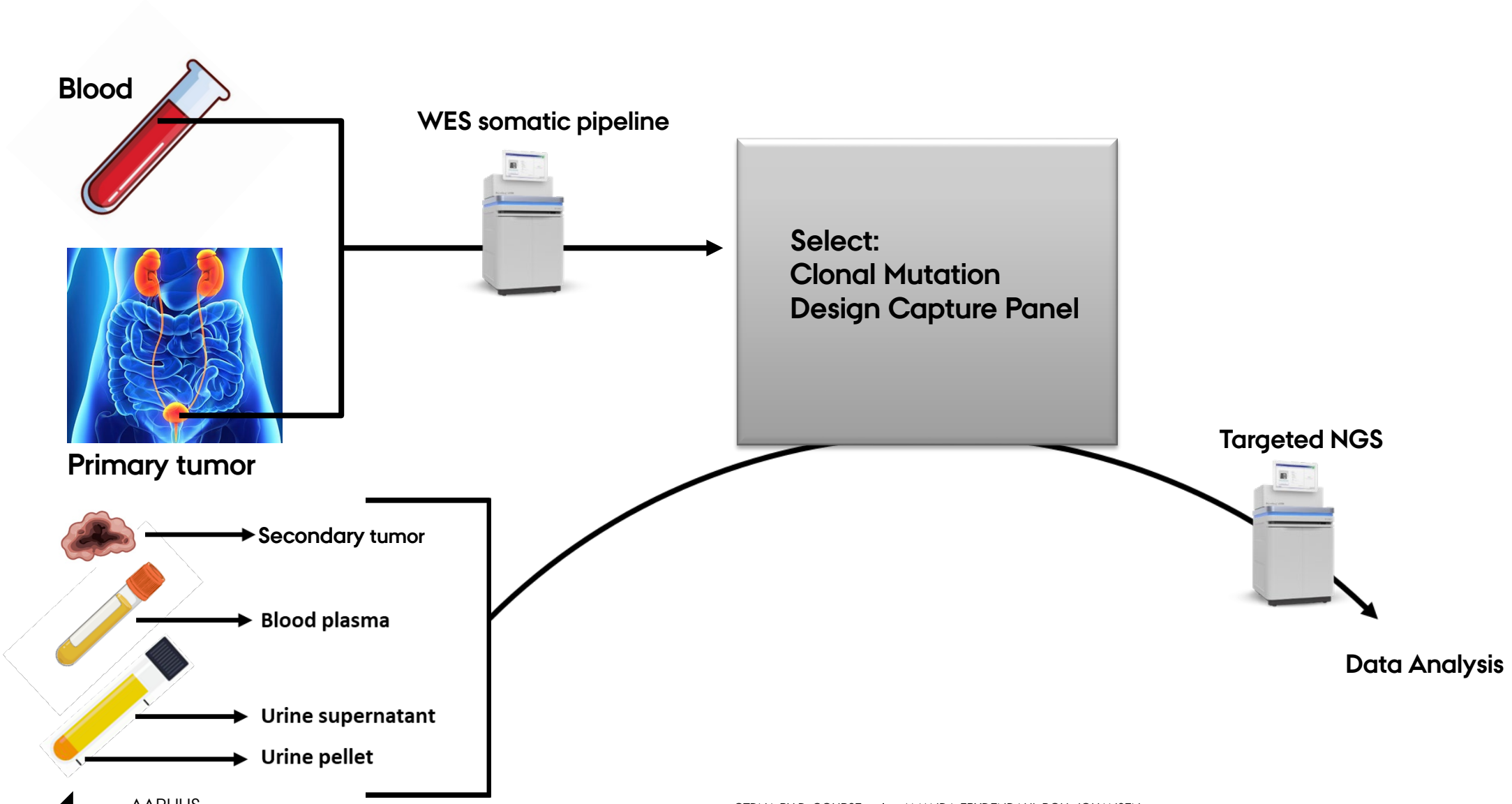
- Random sequence of 3-9 nucleotides
- DNA molecule specific



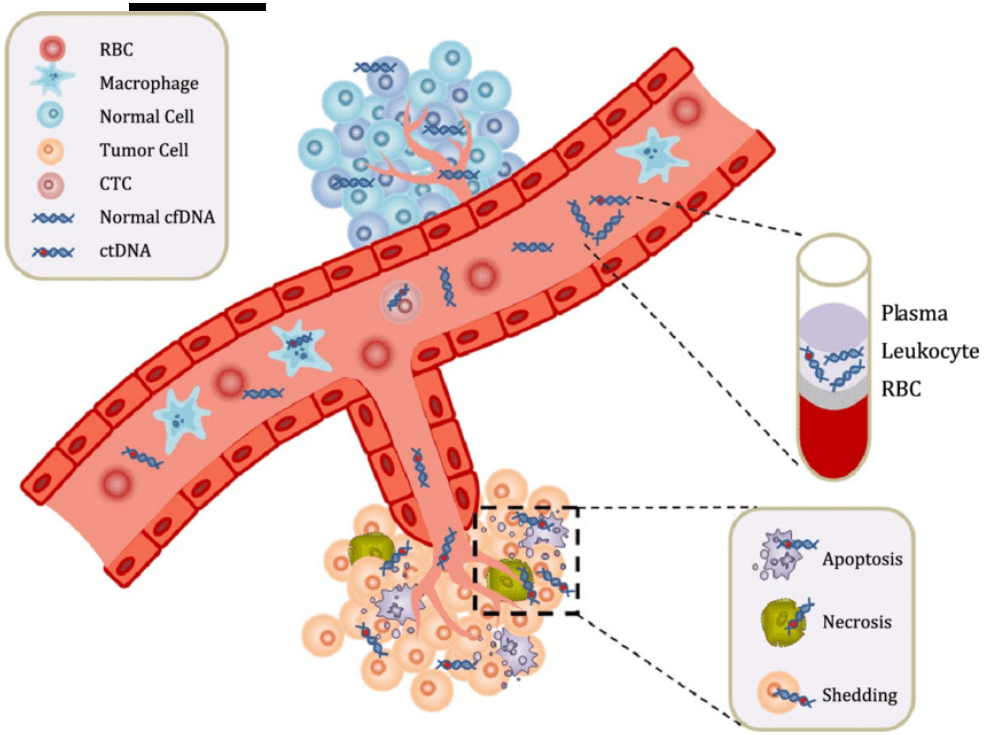
UMI BASED ERROR CORRECTION



WORKFLOW FOR TUMOR INFORMED PANEL SEQ FOR CANCER DISEASE MONITORING



DETECT GENOMIC ALTERATIONS FROM CTDNA IN CFDNA FROM PLASMA



Challenges

- Low amount of cfDNA (typically <50ng)
- The fraction of ctDNA in cfDNA are low.
- The frequency of tumor mutations in cfDNA can be very small
- Need ultra deep sequencing
- WGS and WES are costly

Solution

- Tumor informed ddPCR assays
- Design small custom tumor informed panels
- Design small hotspot panels for specific cancers
- Include UMI to correct for increased PCR and NGS instrument bias

WHEN TO SEQUENCE WHAT?

1. cfDNA from liquid of choice (plasma, urine ect.)
   - Self-explanatory: necessary for ctDNA detection
     - Whole genome sequencing / targeted sequencing

2. DNA from normal cells (PBMC, normal tissue)
   - Filter out patient-specific SNP's (VAF's around 50% or 100%, require coverage around 20-50x)
   - Filter out CHIP mutations (VAF's around ~0.1%, require high coverage)

3. DNA from tumor
   - Identify clonal targets for subsequent single-target analysis (e.g. ddPCR analysis)
   - Generate a patient-specific mutational compendium to search for ctDNA

GROUP DISCUSSION

- Is sequencing useful in your projects for ctDNA detection?
- Can you think of applications where unique molecular identifies (UMIs) are not relevant/necessary?
- If you are/wish to use NGS in your projects, is sequencing of the tumor and/or normal cells necessary? Why? Why not?

AARHUS
UNIVERSITY