

**H3ABioNet**

Pan African Bioinformatics Network for H3Africa

CONNECTING
SCIENCEADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

Next Generation Sequencing Bioinformatics Course 2021

Module 2: Introduction to NGS Technologies

General Introduction to NGS sequencing and technologies

Fatma Guerfali

Learning Objectives

- Recognize the importance of the Human Genome Project (sequencing & databases)
- Recognize that sequencing greatly impacts today's research and its applications
- Summarize the evolution of sequencing technologies
- Describe the principles of NGS technologies



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING SCIENCE
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Session Plan

01

General Overview

The HGP and the Impact on
of Sequencing Technologies

02

Why HT Sequencing

The importance of HT Sequencing
and its applications

03

The Generations of Sequencing Technologies

From Sanger Sequencing to
NGS

04

NGS Sequencing Technologies

The Different NGS Sequencing
Technologies



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING SCIENCE
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Session Plan

01

General Overview

The HGP and the Evolution
of Sequencing Technologies

02

Why HT Sequencing

The importance of HT Sequencing
and its applications

03

The Generations of Sequencing Technologies

From Sanger Sequencing to
NGS

04

NGS Sequencing Technologies

The Different NGS Sequencing
Technologies



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

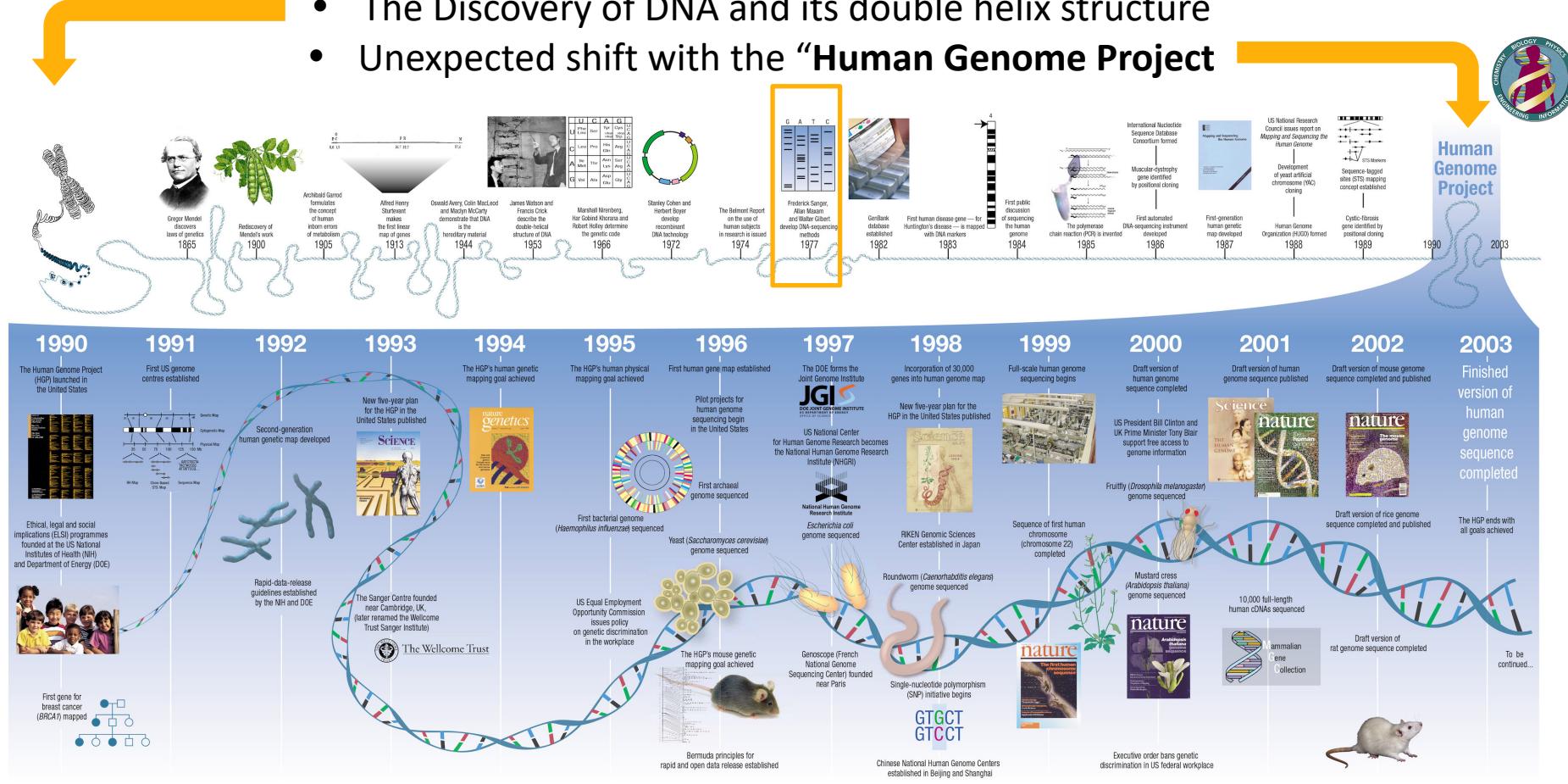


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

General Overview

1st & 2nd Revolutions in Genomics

- The Discovery of DNA and its double helix structure
- Unexpected shift with the “Human Genome Project



Credit: Darryl Leja, NHGRI
<https://www.genome.gov>

Next Generation Sequencing Bioinformatics
 Trainer Name: Fatma Guerfali



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
 CONNECTING
 ADVANCED
 COURSES +
 SCIENTIFIC
 CONFERENCES





HGP : The legacy

Goals



Sequencing the nucleotides that constitute human DNA

vs

Unexpected Outputs



Identify and map all human genes at both a physical and a functional levels



High-level of repetitive regions and segmental duplication

Identify the genetic variants that are related to diseases



Around 22,300 protein-coding genes (same range in other mammals)



Diseases cannot be explained only by protein-coding genes



H3ABioNet

Pan African Bioinformatics Network for H3Africa

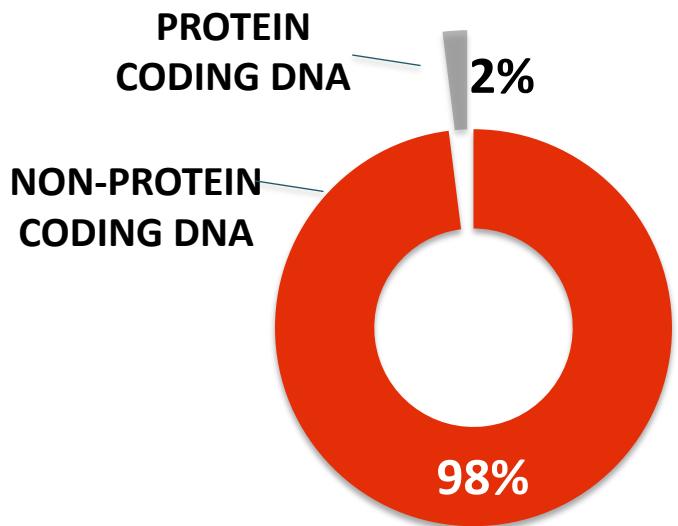
WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali



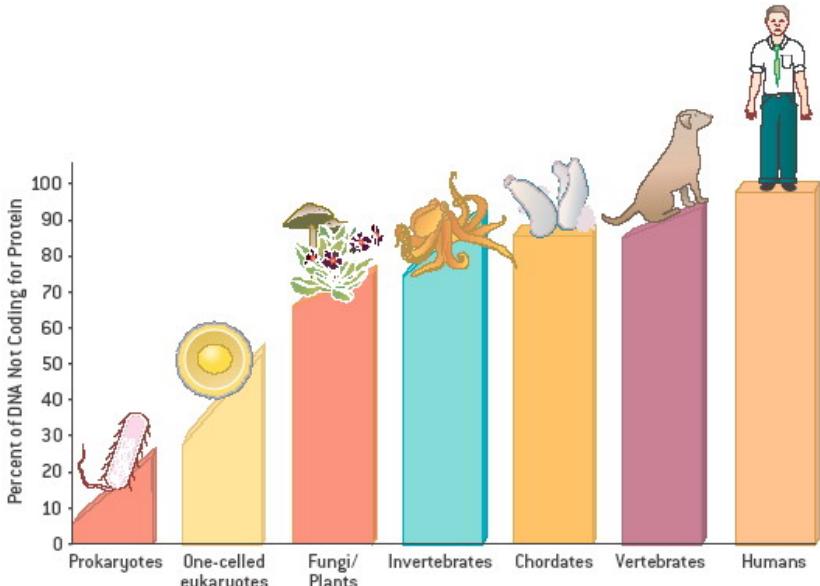
HGP : The legacy



$\approx 22,300$

The human genome contains only about 22,300 protein-coding genes : sequence alone is not enough to explain the whole complexity !

The proportion of non-coding DNA increases with organism complexity



(Mattick, 2011)

<http://www.yourgenome.org/stories/how-is-the-completed-human-genome-sequence-being-used>



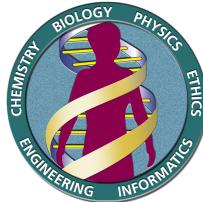
H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali



HGP : The legacy

● New Challenges !

- **Exploitation:** Large amount of information -> major impact in medical science
- **Global Analysis Technologies:** information on the sequence of an entire genome does not answer all of our questions - need a full understanding of the function of genes and related regulation from other regions of the genome
- **Output:** discover how sets of genes and their products work together under normal and “abnormal” conditions (diseases).

→ One of the main requirements of these studies is the appropriate development / use of High-Throughput (HT) Sequencing technologies



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Session Plan

01

General Overview

The HGP and the Impact on
of Sequencing Technologies

02

Why HT Sequencing

The importance of HT Sequencing
and its applications

03

The Generations of Sequencing Technologies

From Sanger Sequencing to
NGS

04

NGS Sequencing Technologies

The Different NGS Sequencing
Technologies



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING SCIENCE
ADVANCED COURSES +
SCIENTIFIC CONFERENCES

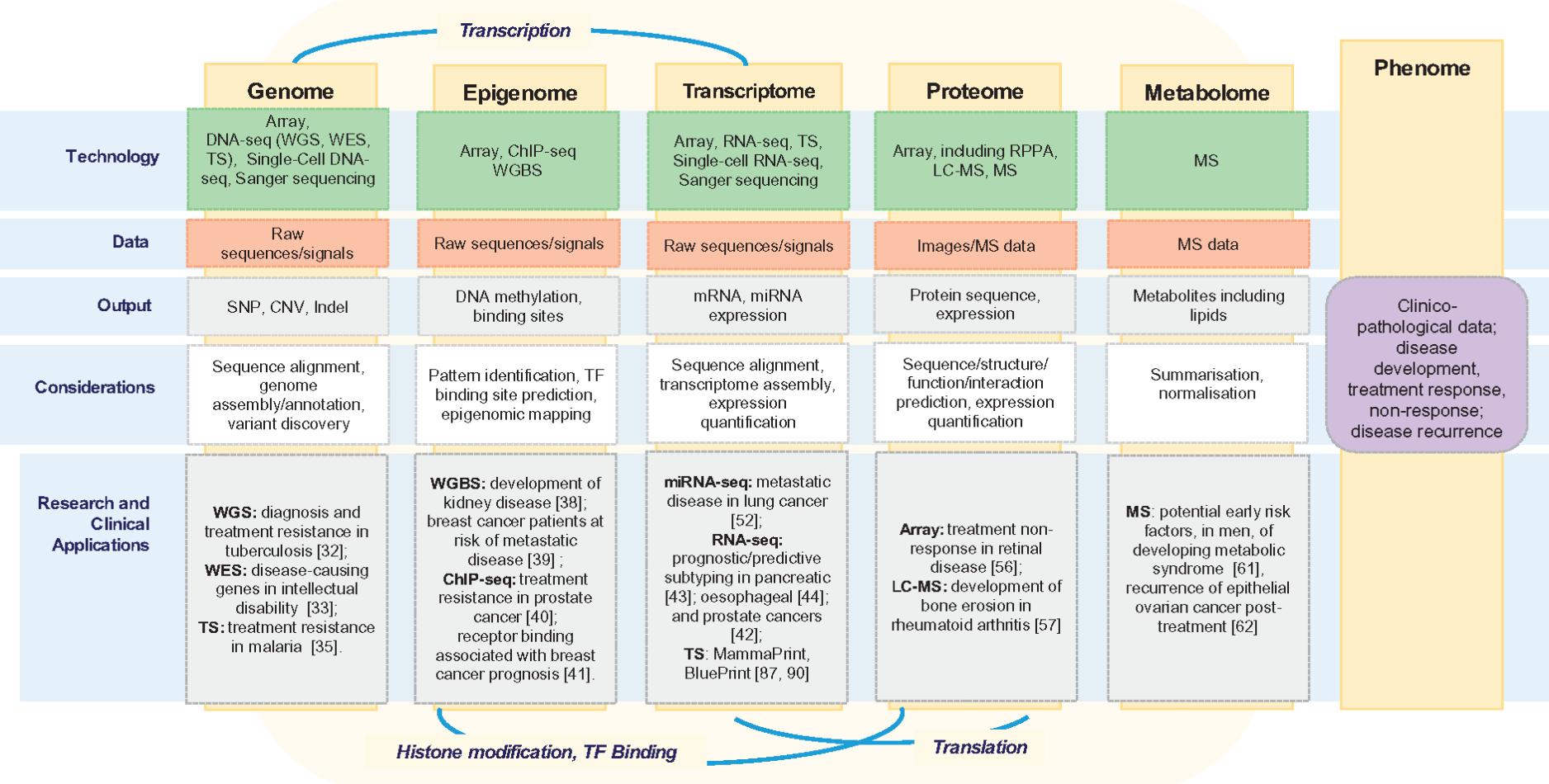


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

The importance of HT Sequencing

- Thanks to **advances in Sequencing technologies**, we can now routinely use High-Throughput Sequencing as opposed to individual gene-based approaches (conventional Sanger Sequencing)
- **Advances include:**
 - reduction of sequencing costs
 - adoption of well-established sequencing workflows
 - continuous improvement of platforms
 - sequencing performance (output size and rapidity)
- These technologies have been central to create **an exponential growth in:**
 - Capacity to analyse genomes (1990-2003: HGP, 2016: days, 2019: hours)
 - Capacity to analyze many other « Omes » → « **Omics** »
 - **Performance and output** of sequencing technologies (omics data) with full genome sequencing now producing gigabases of reads on a daily basis.

The Applications of HT Sequencing



Session Plan

01

General Overview

The HGP and the Impact on
of Sequencing Technologies

02

Why HT Sequencing

The importance of HT Sequencing
and its applications

03

The Generations of Sequencing Technologies

From Sanger Sequencing to
NGS

04

NGS Sequencing Technologies

The Different NGS Sequencing
Technologies



H3ABioNet

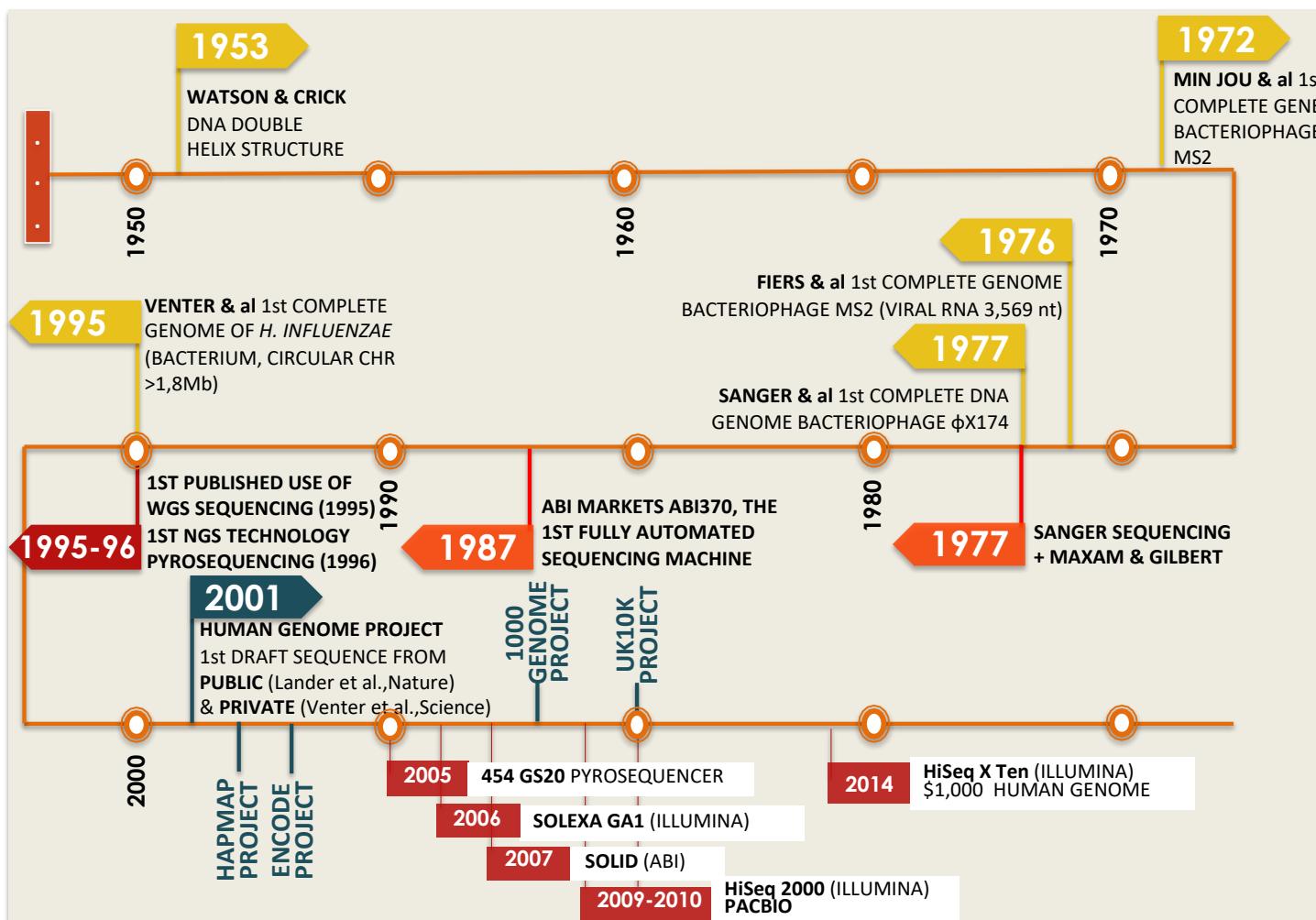
Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Today



H3ABioNet

Pan African Bioinformatics Network for H3Africa

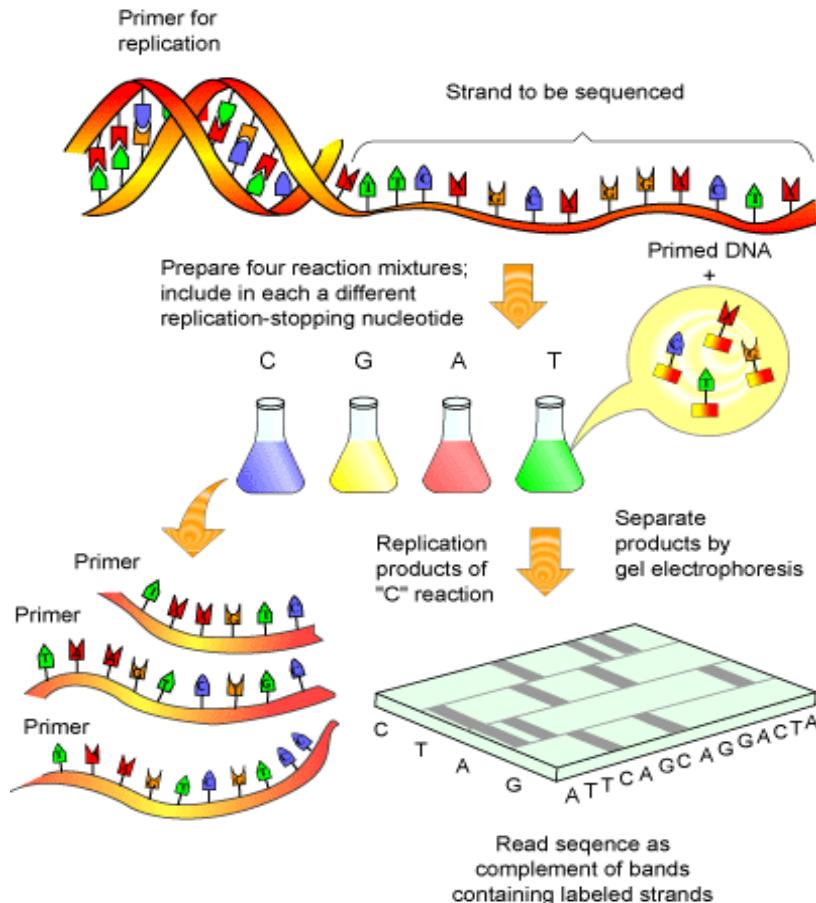
WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

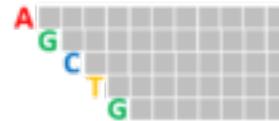
First Generation

- Sanger Sequencing

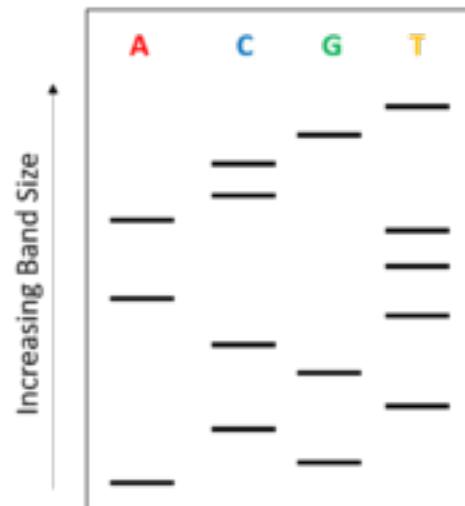


Sequence: **A G C T G C T A T T A C C G T**

Chain-terminating
PCR:



Gel Electrophoresis:



scq.ubc.ca

<https://www.pacb.com/blog/the-evolution-of-dna-sequencing-tools/>



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

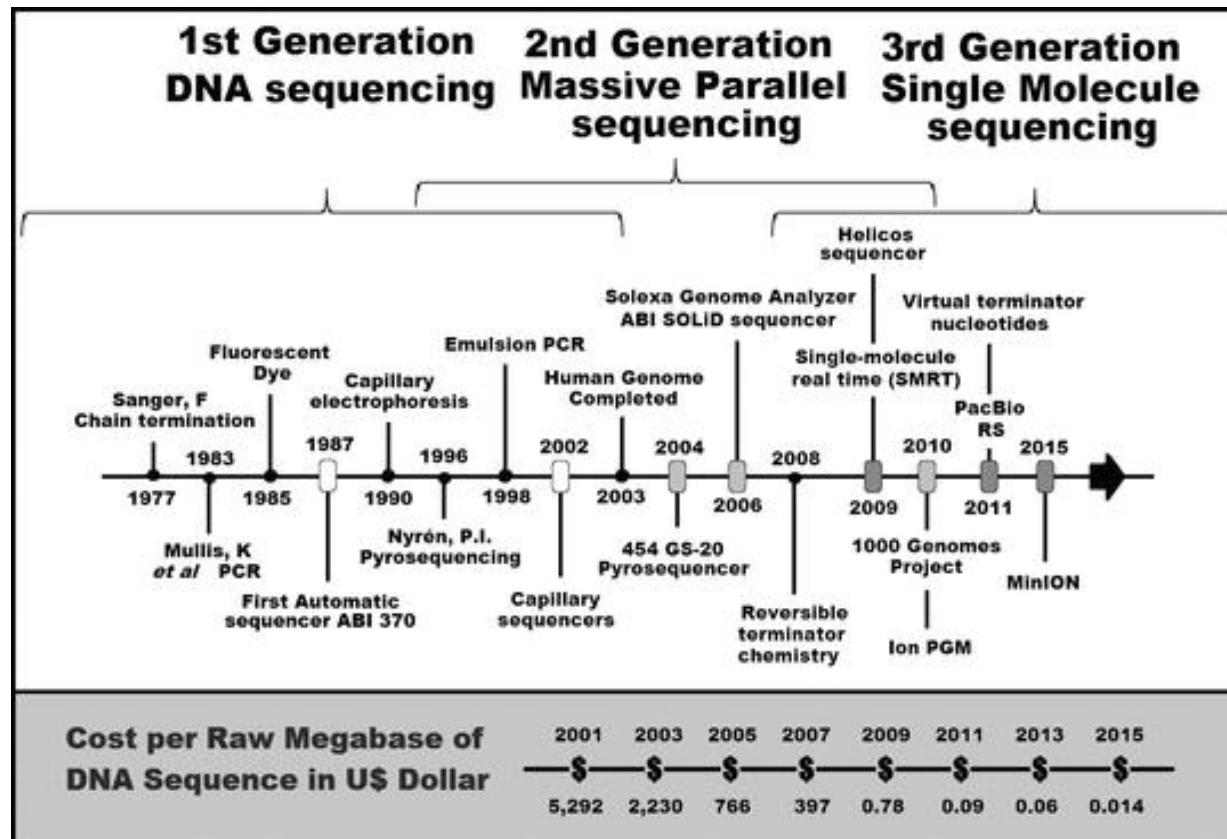


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Generations of Sequencing Technologies

NGS: 3rd Revolution in Genomics

- The technological shift with the development of **NGS**
“Next generation Sequencing” technologies



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Session Plan

01

General Overview

The HGP and the Impact on
of Sequencing Technologies

02

Why HT Sequencing

The importance of HT Sequencing
and its applications

03

The Generations of Sequencing Technologies

From Sanger Sequencing to
NGS

04

NGS Sequencing Technologies

The Different NGS Sequencing
Technologies



H3ABioNet

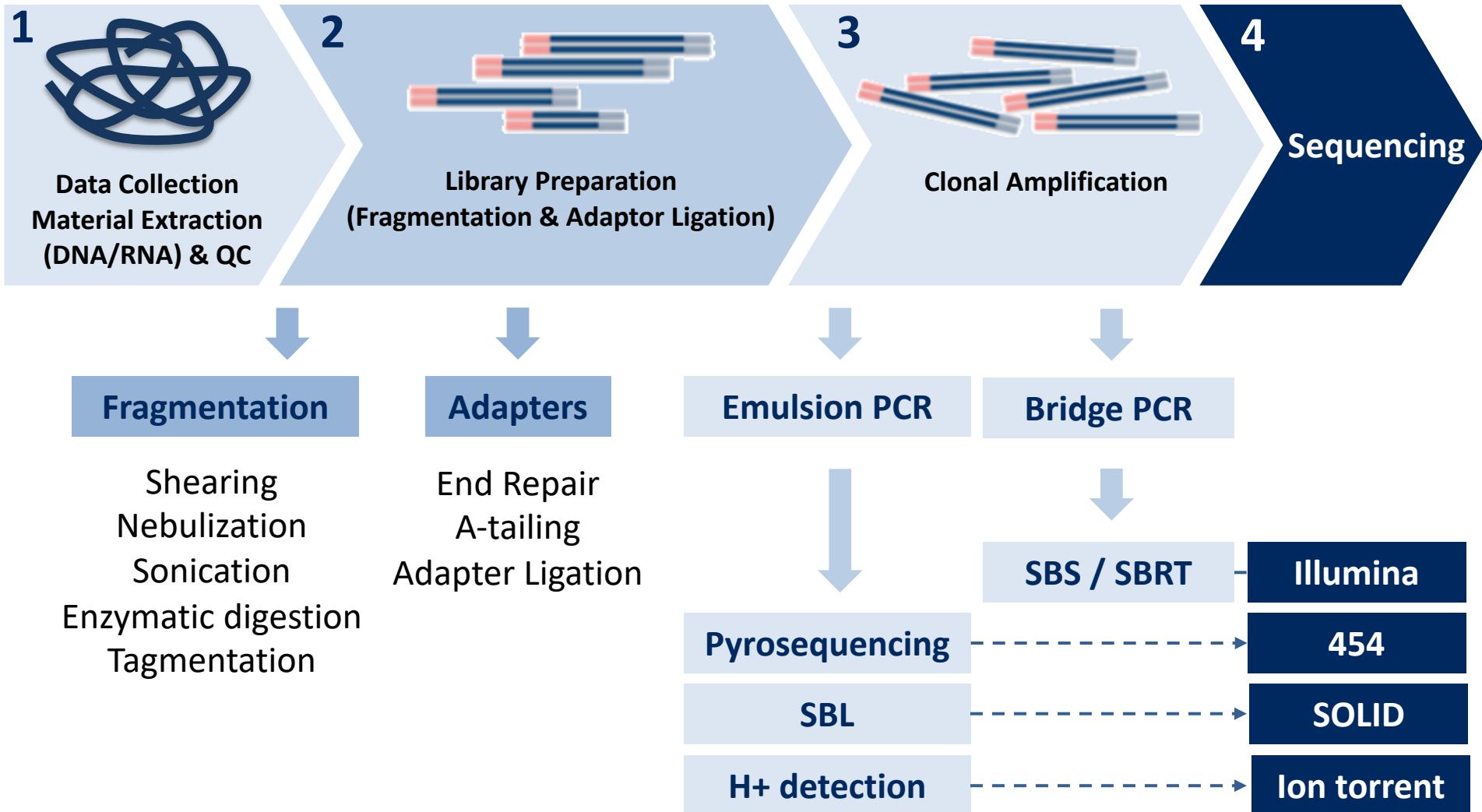
Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING SCIENCE
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

General Principle of NGS (2nd)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second and Third generation

Technology	Company	Support	Chemistry
Massively Parallel Sequencing			
Solexa 454 SOLiD Ion Torrent	Illumina Roche Applied Science Thermo Fisher Thermo Fisher	Bridge PCR on flowcell emPCR on beads emPCR on beads emPCR on beads	SBS / SBRT Pyrosequencing Seq-By-Ligation Proton detection
Single Molecule Sequencing			
PacBio SMRT Nanopore	Pacific Biosciences Oxford Nanopore Tech/McNally	Pol performance Translocation	Real-time-Seq NA



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second and Third generation

Technology	Company	Support	Chemistry
Massively Parallel Sequencing			
Solexa	Illumina	Bridge PCR on flowcell	SBS / SBRT
454	Roche Applied Science	emPCR on beads	Pyrosequencing
SOLiD	Thermo Fisher	emPCR on beads	Seq-By-Ligation
Ion Torrent	Thermo Fisher	emPCR on beads	Proton detection

SBL

SBS



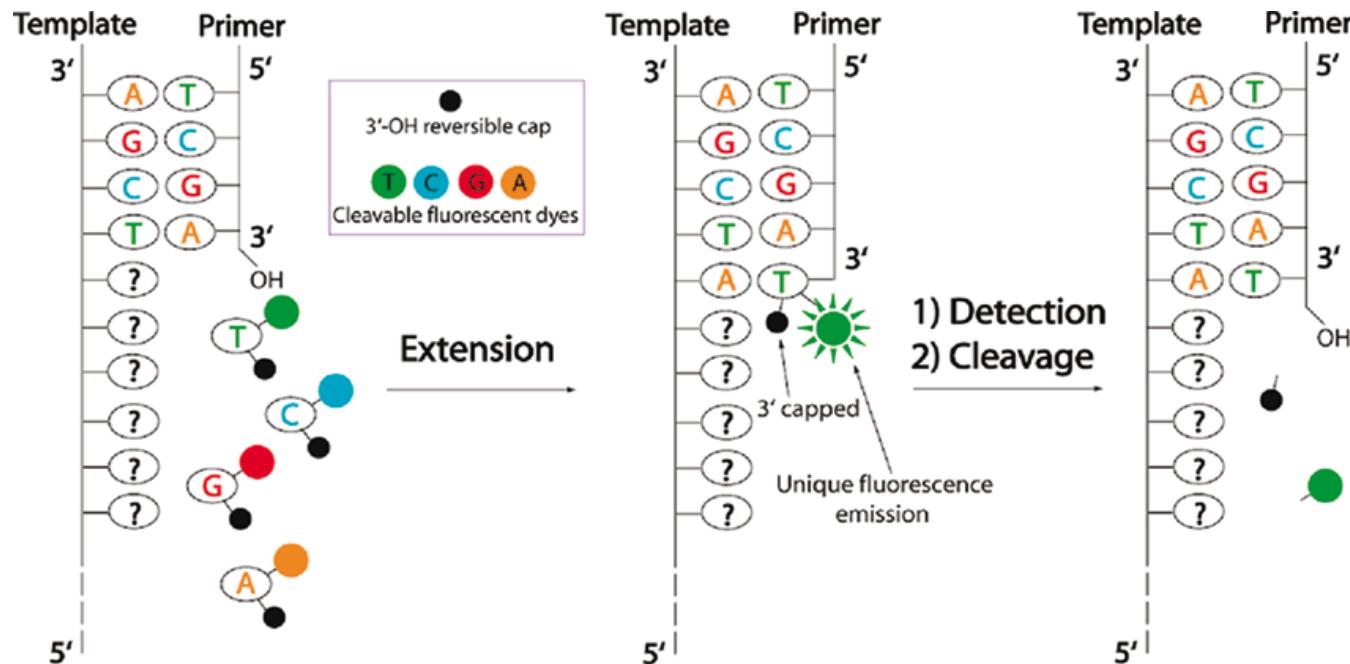
Pyrosequencing
Sequencing by reversible termination
Sequencing by detection of hydrogen ions

Second Generation

Principle of " Sequencing By Synthesis " (commonly SBS)

= tracking the addition of fluorescently labeled nucleotides as the DNA chain is copied.

- The DNA template is immobilized.
- Solutions of A, nucleotides C , G and T added and removed.
- Fluorescence is emitted when a nucleotide complements the unpaired base.
- Chemiluminescent signal detected to determine the sequence.



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



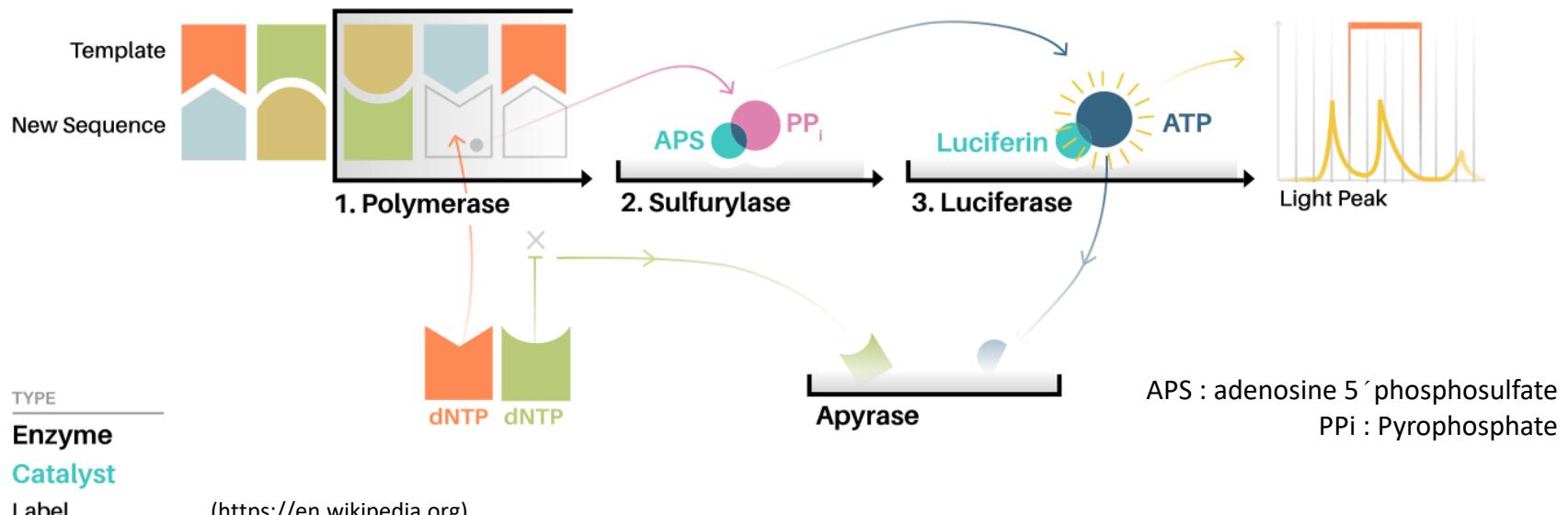
Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second Generation

Principle of " Pyrosequencing"

= the DNA sequence is determined by the light (luciferase activity) emitted during the incorporation of nucleotides.

- The DNA template is immobilized.
- Solutions of A, nucleotides C , G and T added **sequentially** and removed.
- Incorporation of a dNTPs → releases pyrophosphate (PP_i) → ATP sulfurylase converts PP_i to ATP in the presence of APS → ATP = substrate for the luciferase-mediated conversion of luciferin to oxyluciferin → generates visible light



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second and Third generation

Technology	Company	Support	Chemistry
Massively Parallel Sequencing			
Solexa	Illumina	Bridge PCR on flowcell	SBS / SBRT
454	Roche Applied Science	emPCR on beads	Pyrosequencing
SOLiD	Thermo Fisher	emPCR on beads	Seq-By-Ligation
Ion Torrent	Thermo Fisher	emPCR on beads	Proton detection

SBL

SBS

→ Pyrosequencing
Sequencing by reversible termination
Sequencing by detection of hydrogen ions



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING SCIENCE
ADVANCED COURSES + SCIENTIFIC CONFERENCES

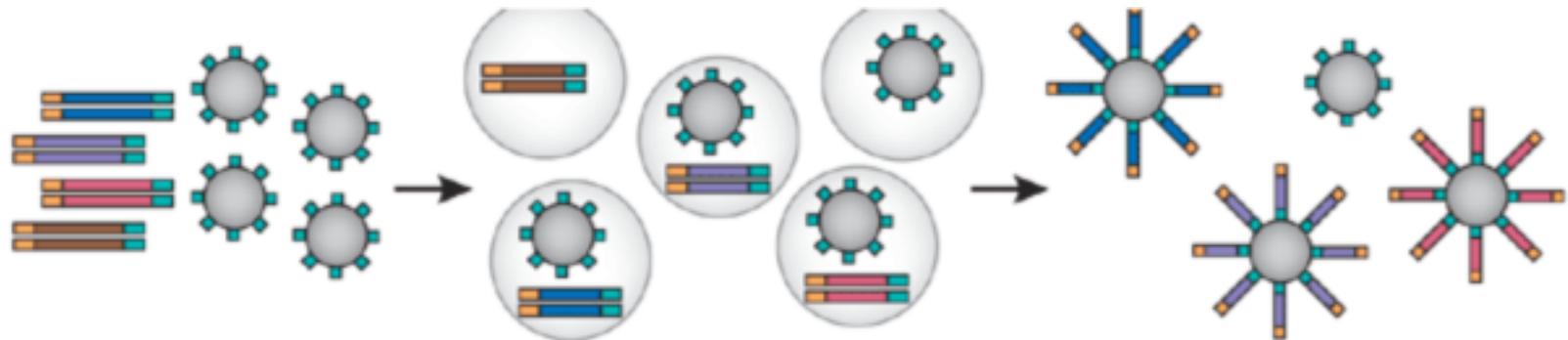


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second Generation

Principle of "emPCR"

- An adaptor-flanked shotgun library is PCR amplified in the context of a water-in-oil emulsion.
- PCR primer is 5'-attached on micron-scale beads.
- 1 bead-containing compartments = 0 or 1 template DNA.
- PCR amplicons are captured to the surface of the bead.
- 1 clonally amplified bead = PCR products corresponding to amplification of a single molecule from the library.



(Shendure & Ji, 2008)



H3ABioNet
Pan African Bioinformatics Network for H3Africa

WELLCOMBE GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

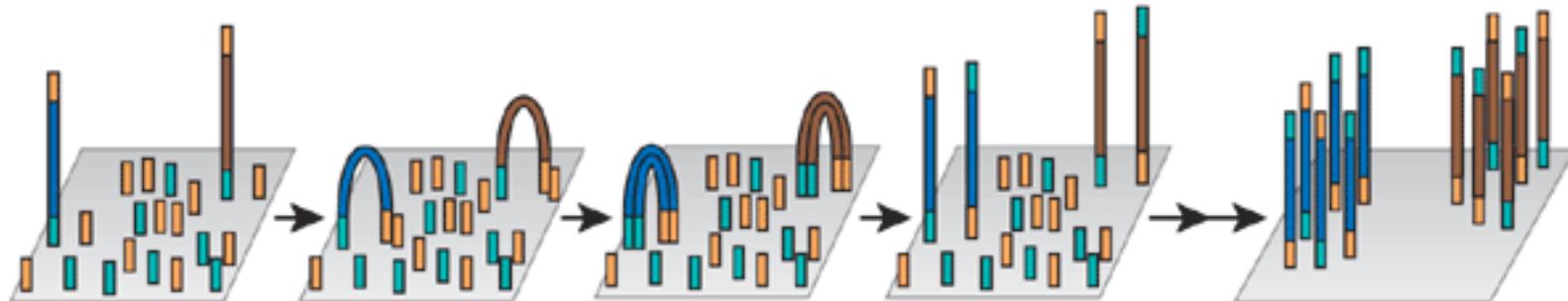


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second Generation

Principle of " Bridge PCR"

- An adaptor-flanked shotgun library is PCR amplified on a flow cell
- both primers coat the surface of a solid substrate
- Amplification products from any given member of the library remain locally fixed near the point of origin = cluster
- The PCR produces clonal clusters that contains copies of a single DNA.



(Shendure & Ji, 2008)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOMBE GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

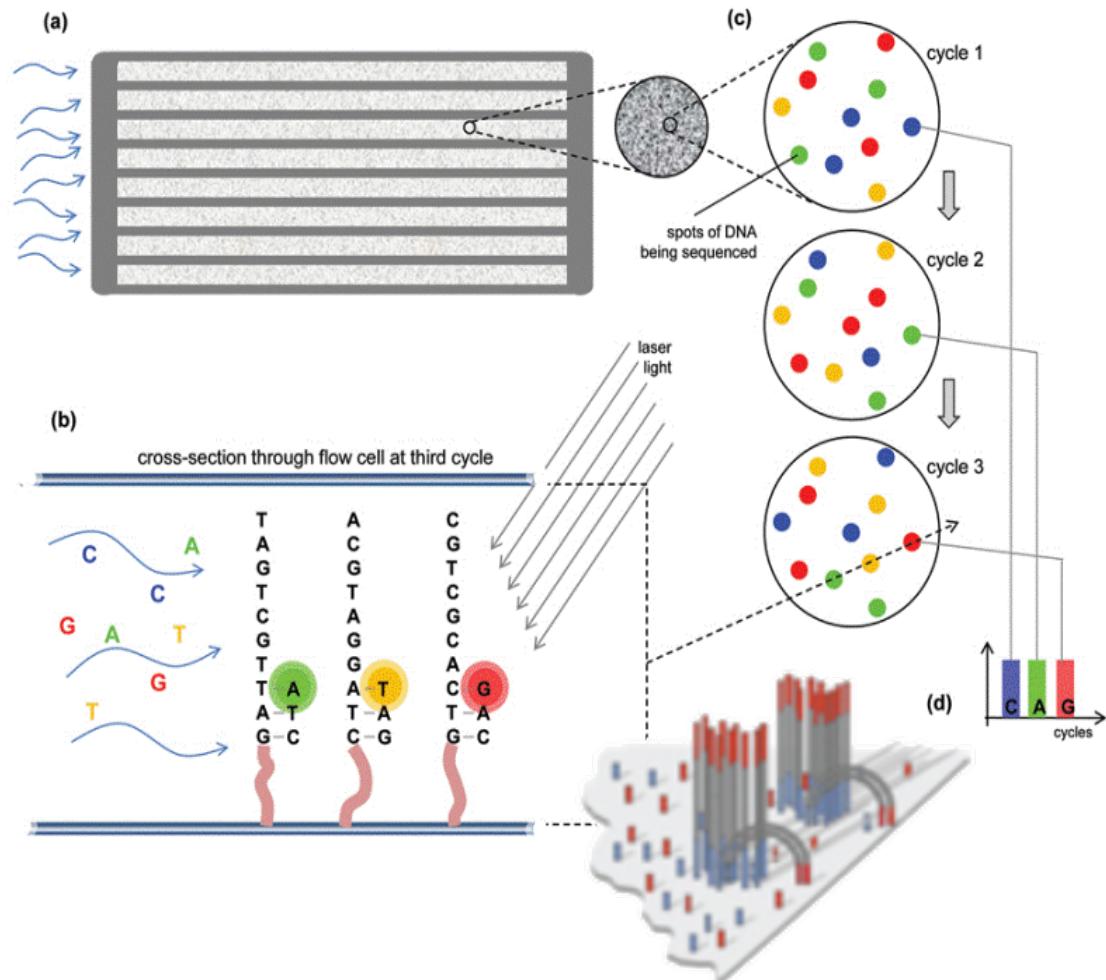


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second Generation

► Solexa (Illumina)

- The input sample must be cleaved into short sections.
- Fragments are ligated to adaptors and annealed to the slide using the adaptors.
- Fragments are separated into single strands to be sequenced.
- Nucleotides are modified so that each emits a different coloured light when excited by a laser.
- they have a terminator, so that only one base is added at a time.
- PCR, process repeated in cycles, images analyzed (SBS).



(Adapted from Gilchrist, 2010)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

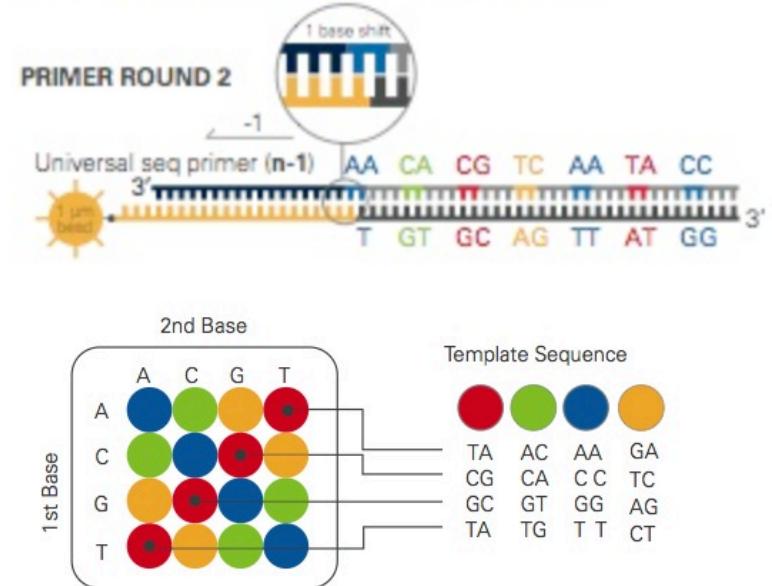


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second Generation

SOLID

- SOLiD (Sequencing by Oligonucleotide Ligation and Detection)
- The strategy uses a two-base encoding scheme (each base is effectively sequenced twice) in a **Sequencing-By-Ligation** process



<http://seqanswers.com/forums/showthread.php?t=10>



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

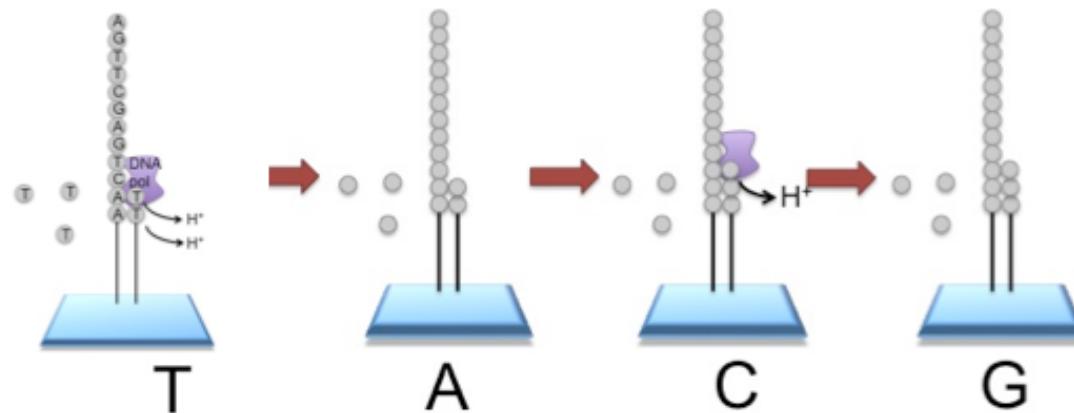


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second Generation

► Ion torrent (Thermo Fisher)

- As in other kinds of NGS, the input DNA is fragmented.
- Unlike other methods, Ion Torrent do not use optical signaling.
- Adaptors are added and one molecule is placed onto a bead.
- Amplification on the bead by emulsion PCR. Each bead is placed into 1 well of a slide.
- The dNTPs are washed away, and the process is repeated in cycles.
- **The pH is detected in each of the wells, as each H⁺ ion released will decrease the pH.** The changes in pH allow us to determine if that base, and how many thereof, was added to the sequence read.



<http://seqanswers.com/forums/showthread.php?t=10>



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Second and Third generation

Technology	Company	Support	Chemistry
Massively Parallel Sequencing			
Solexa 454 SOLiD Ion Torrent	Illumina Roche Applied Science Thermo Fisher Thermo Fisher	Bridge PCR on flowcell emPCR on beads emPCR on beads emPCR on beads	Seq-By-Synthesis Pyrosequencing Seq-By-Ligation Proton detection
Single Molecule Sequencing			
PacBio SMRT Nanopore	Pacific Biosciences Oxford Nanopore Tech/McNally	Pol performance Translocation	Real-time-Seq NA



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOMBE GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

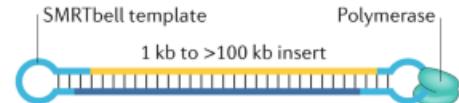
Third Generation

PacBio

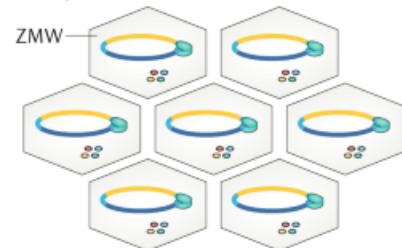
- DNA is fragmented and ligated to hairpin adapters (light blue) to form a topologically circular molecule known as a SMRTbell.
- SMRTbell is bound by a DNA polymerase and loaded onto a SMRT Cell for sequencing.
- Each SMRT Cell contain zero-mode waveguides (ZMWs), which are chambers that hold picolitre volumes.
- Fluorescent dNTP are added and detected briefly (the fluorophore is linked to phosphate that will be naturally cleaved → prevent fluorescence interference).

a PacBio SMRT sequencing

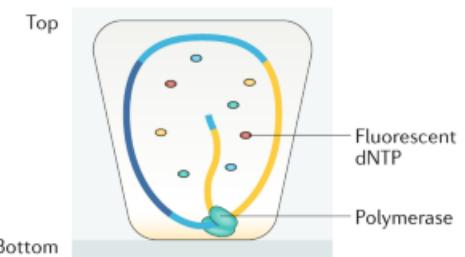
Template topology



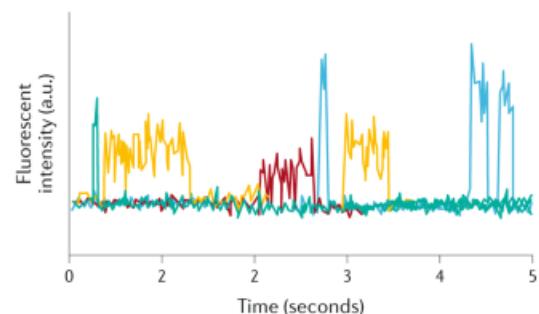
Flow cell (top view)



Single ZMW
(cross section)



Readout



(Logsdon et al., 2020)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

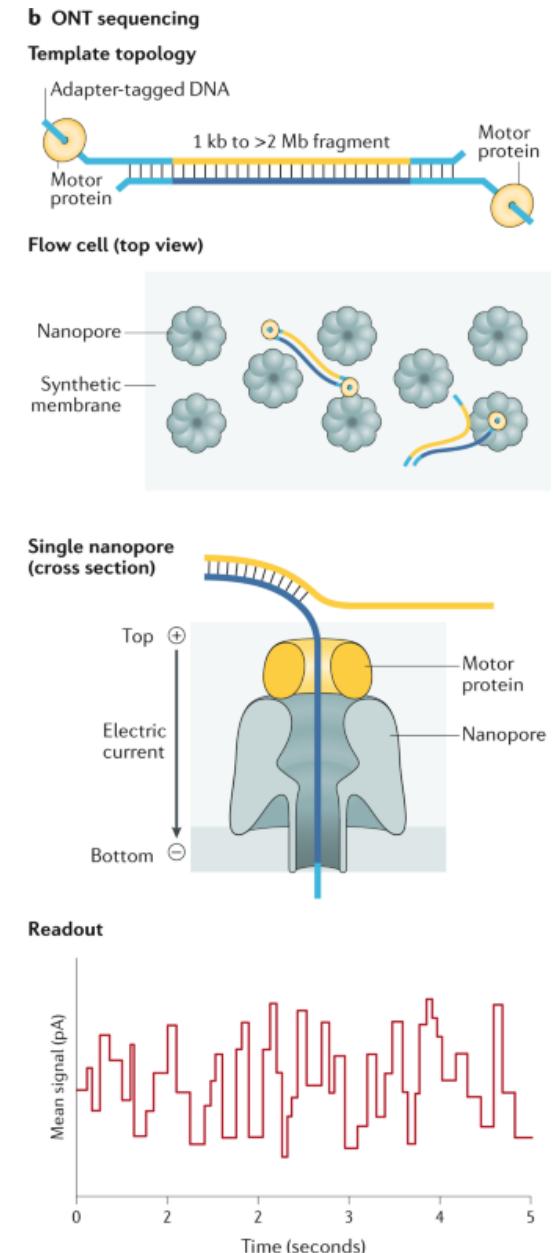


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Third Generation

► Oxford Nanopore (ONT)

- DNA is tagged with sequencing adapters preloaded with a motor protein.
- The flow cell contains hundreds to thousands of protein nanopores in a synthetic membrane
- Sequencing adapter inserts into the nanopore, and the motor protein unwinds the DNA.
- An electric current is applied driving the negatively charged DNA through the pore
- As the DNA moves through the pore, it causes characteristic disruptions to the current, generating a readout known as a 'squiggle'.



(Logsdon et al., 2020)



H3ABioNet

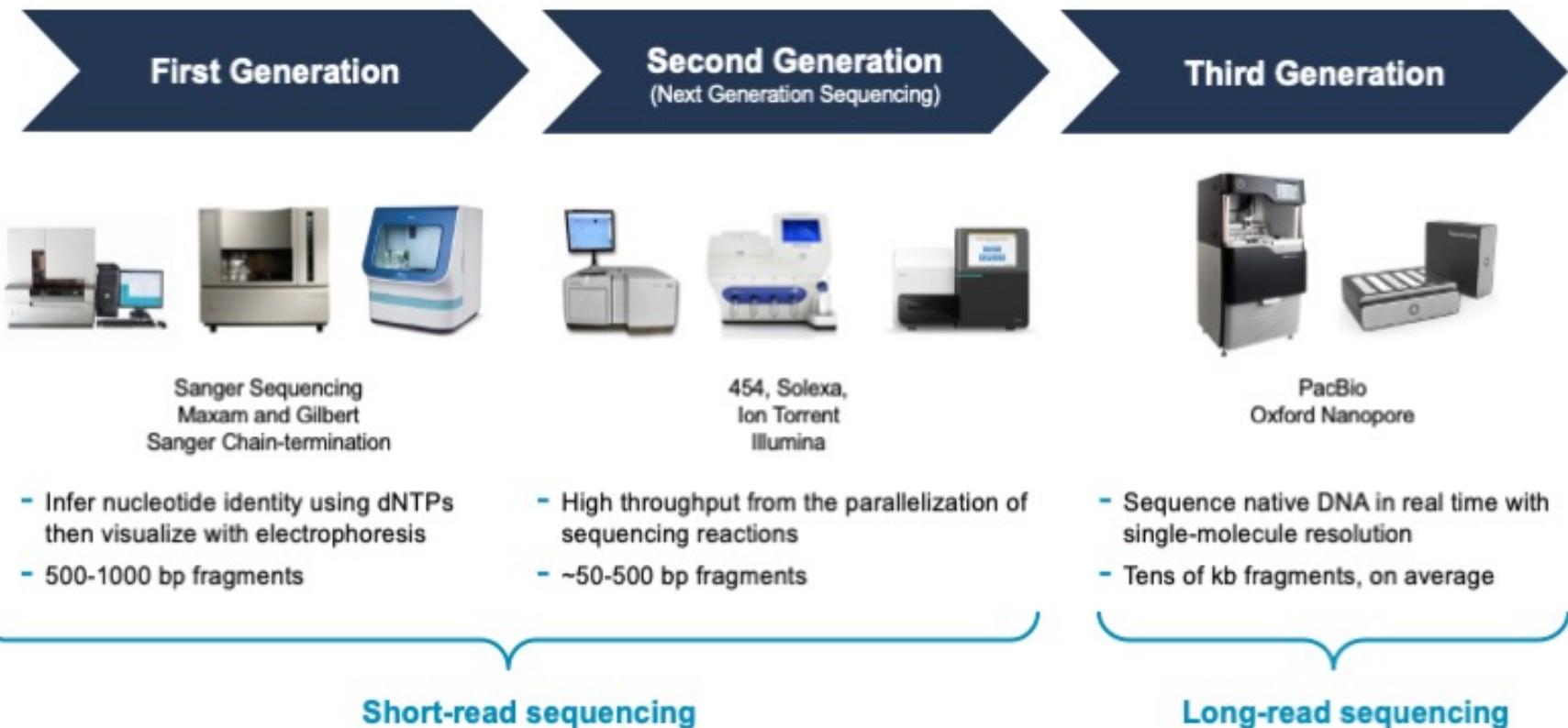
Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Generations of Sequencing Technologies



<https://www.pacb.com/blog/the-evolution-of-dna-sequencing-tools/>



H3ABioNet

Pan African Bioinformatics Network for H3Africa

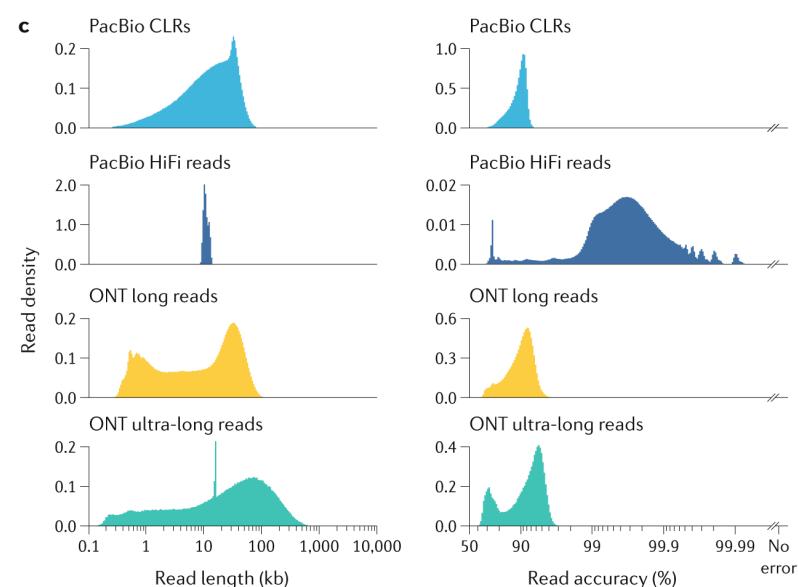
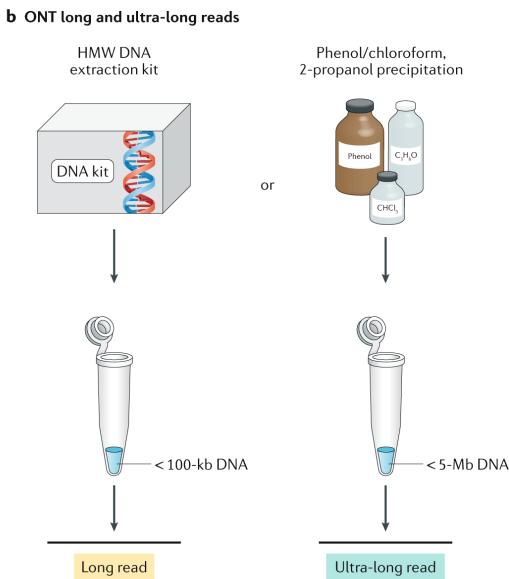
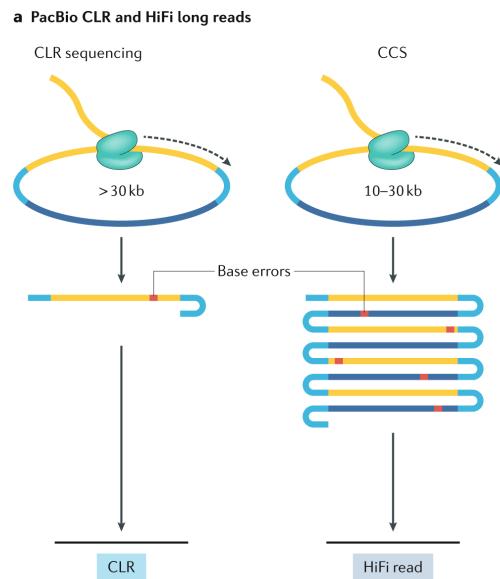
WELLCOMBE GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Long reads

- PacBio and ONT both are considered long-read sequencing technologies
- They generate different types of long reads that differ both in length and accuracy



(Logsdon et al., 2020)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

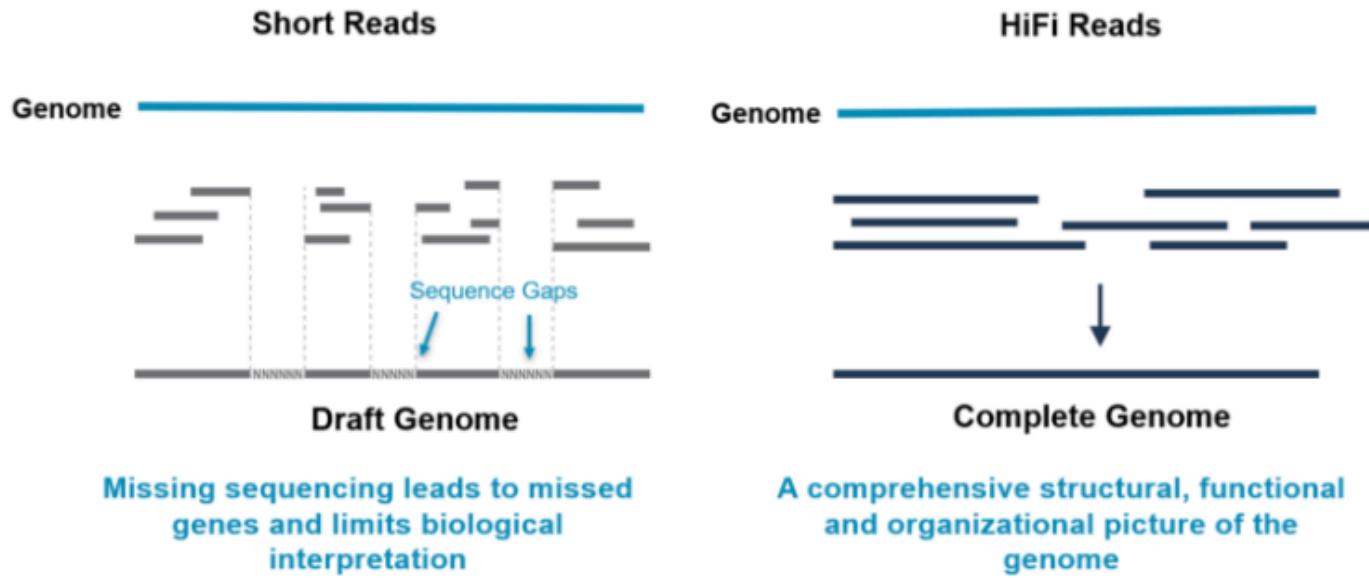
WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Long reads: improvement and Impact

- The sequencing of « short reads » produces reads of 50-500 bp
- The sequencing of « long reads » produces reads of 10s of kb in length



<https://www.pacb.com/blog/the-evolution-of-dna-sequencing-tools/>



H3ABioNet

Pan African Bioinformatics Network for H3Africa

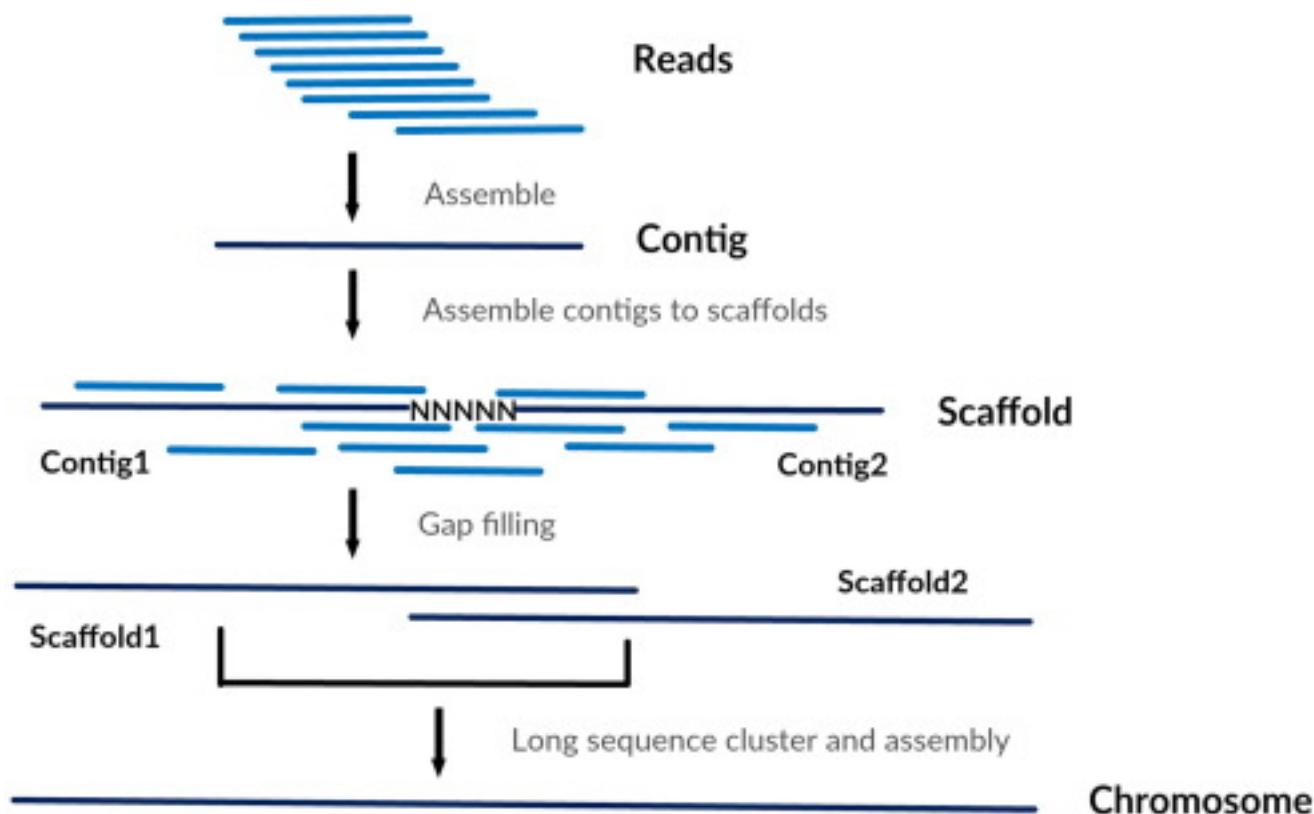
WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Long reads

- Long reads are used to improve genome assembly



(Guo et al., 2017)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

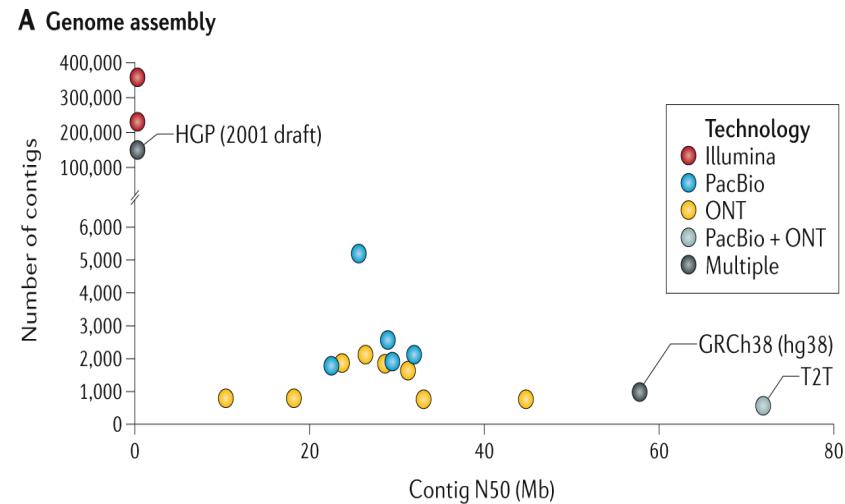
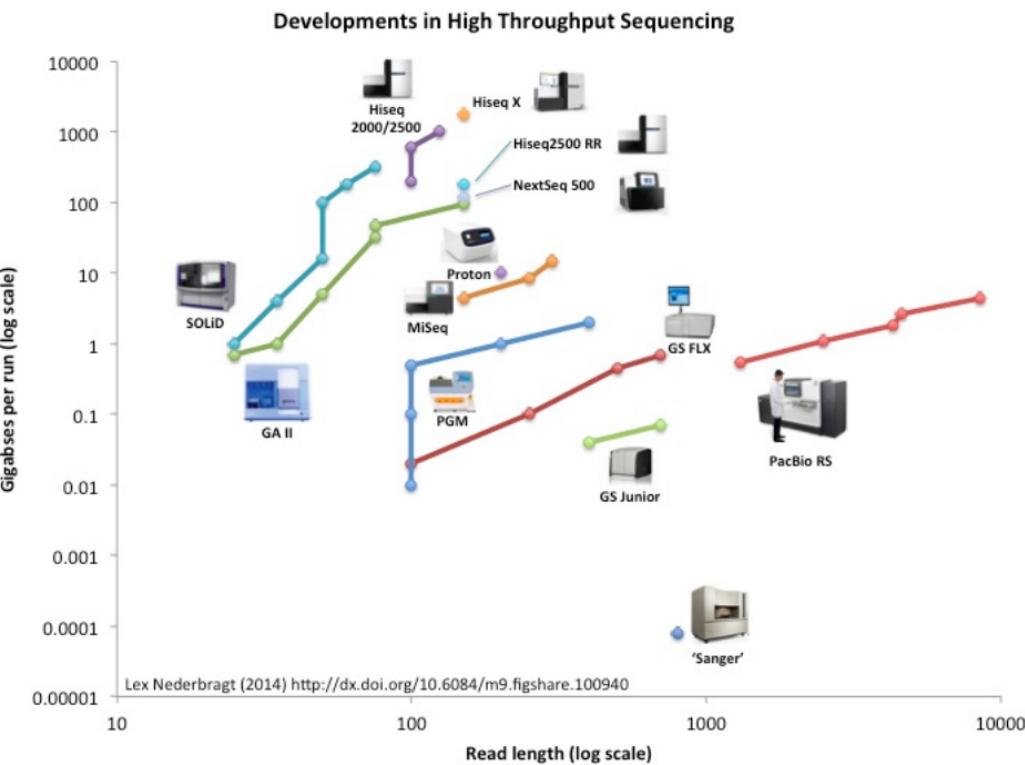
WELLCOMBE GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Long reads

- Long reads are used to improve genome assembly



(Logsdon et al., 2020)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Short vs Long reads

- Short & Long reads can complement each other

Sequencing lengths available

NGS Single End (50–300, Illumina)

NGS Paired End (2*75–300, total 150–600bp, Illumina)

Long Read (>10 000, no fixed upper limit)

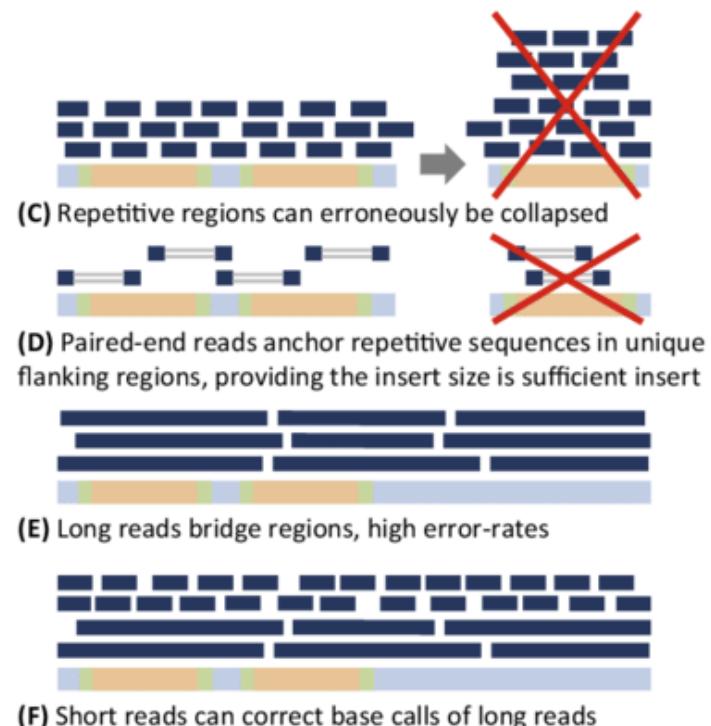
Assembling genome *de novo*



(A) Correct assembly with short reads



(B) Uneven coverage results in missing regions



Trends in Parasitology

(Wit and Gillard., 2017)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOMBE GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

The Challenges of HT Sequencing

Data Collection

- Extensive use of technical and human resources (Data size, Data security, etc)
- Sampling accuracy (collection or heterogeneity can bias later analytical stages)

Data Analysis and Interpretation

- Development and management of tools, analysis workflows, etc
- Multidisciplinary environment is crucial (biology, biostatistics, bioinformatics)

Analytical/computational challenges

- HT Sequence data are both high-dimensional and complex in structure → both algorithmically and computationally challenging to integrate it with other data sets, platforms or technologies, to obtain a complete disease profile..
- We need specific algorithmic approaches including:
 - meta-omics (integration of independent data sets at the same omics level)
 - poly-omics (integration of different omics types)

OXFORD

UNIVERSITY PRESS

Lightbody et al., 2019. <https://doi.org/10.1093/bib/bby051>



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOMBE GENOME CAMPUS
CONNECTING
SCIENCE
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Today

Projects

- ENCODE: Research consortium investigating functional elements in human and model organism DNA
- GENCODE: identify and classify all gene features in the human and mouse genomes (accurate annotations)
- HAPMAP: a tool to find genes and genetic variations that affect health and disease
- 1,000 GENOMES: International research effort on global reference for Human genetic variation
- UK10K : Genome-wide study of thousands of DNA seq to explore rare diseases variants

Public Databases

- Resources and tools to centralize information
- NCBI / Ensembl
- HUGO: Human Genome Organization
- HGNC: HUGO Gene Nomenclature Committee to set the standard for Human Gene Nomenclature
- HPO: human Phenotype Ontology
- dbSNP / dbGAP / dbVar → ClinVar
- African Genome Variation Project

Sequencing technologies

- 2nd Generation
- 3rd Generation



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED
COURSES +
SCIENTIFIC
CONFERENCES

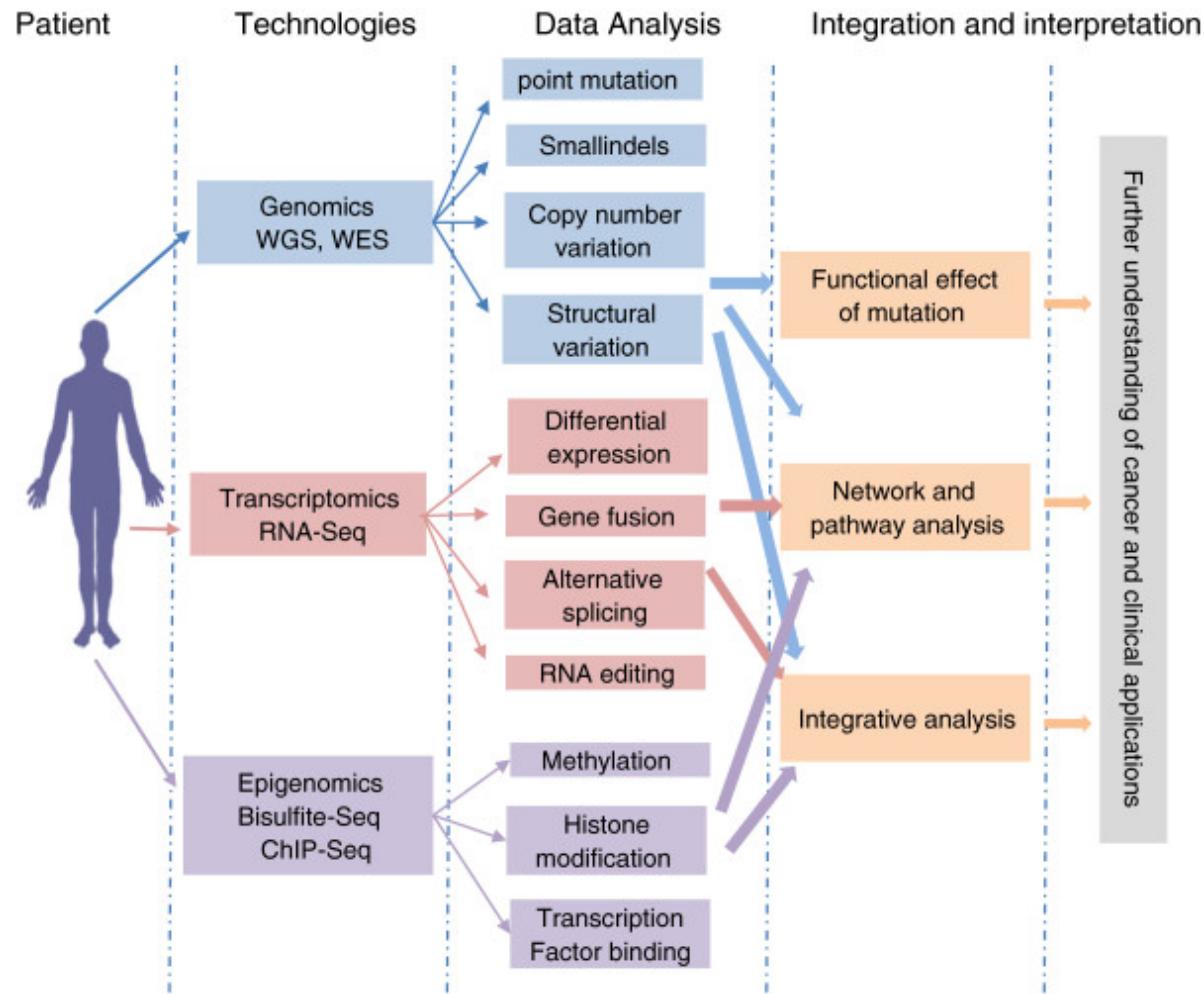


Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Take-home messages

Today, HT Sequencing offers a broad range of applications

- Discover variants
- Population studies
- Analyze clinical isolates
- Study Rare diseases
- Characterize the Metagenome
- Characterize DNA-protein (TF) interactions



(Shyr and Liu, 2013)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali

Take-home messages

- There is no single method better than another one, all depends on your question of interest !
- One can envision a powerful symbiosis between old and new technologies (microarrays vs NGS, etc)
- or a combination of short vs long reads sequencing !

(Shyr and Liu, 2013)



H3ABioNet

Pan African Bioinformatics Network for H3Africa

WELLCOME GENOME CAMPUS
CONNECTING SCIENCE
ADVANCED COURSES +
SCIENTIFIC CONFERENCES



Next Generation Sequencing Bioinformatics
Trainer Name: Fatma Guerfali