RUOKAVIRASTO

Livsmedelsverket • Finnish Food Authority

RUOKAVIRASTO
Livsmedelsverket • Finnish Food Authority

What is HTS?

General presentation on the technology

Johanna Santala

Head of plant pest and grain section

Plant Analytics

5th EPPO Workshop for Heads of Plant Pest Diagnostic Laboratories

19.4.2023

Terms related to HTS

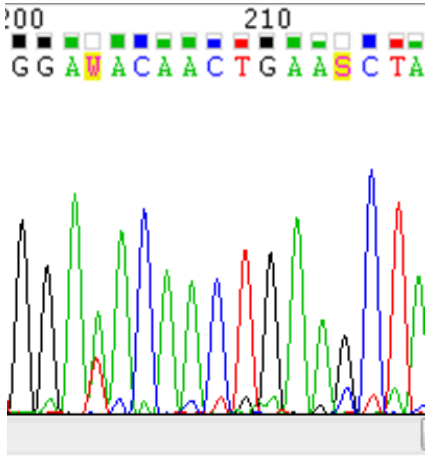
- High throughput sequencing (HTS) = Next generation gequencing (NGS) = deep sequencing
- Different platforms: Illumina dye sequencing, PacBio's SMRT, Ion Torrent, Oxford Nanopore, SOLiD…
- Different tecniques: amplicon sequencing (eg. metabarcoding), shotgun sequencing (eg. whole genome sequencing, siRNA sequencing)

Sanger sequencing vs HTS

- Sanger sequencing:
- Targeted
- Nucleic acids are prepared for sequencing by PCR-reaction that produces different lenght products ending with labeled nucleotides
- Sequencer separates the products by lenght and detects the labeled nucleotides
- One consensus sequence is produced

Extracted nucleic acids

Preaparation for sequencing

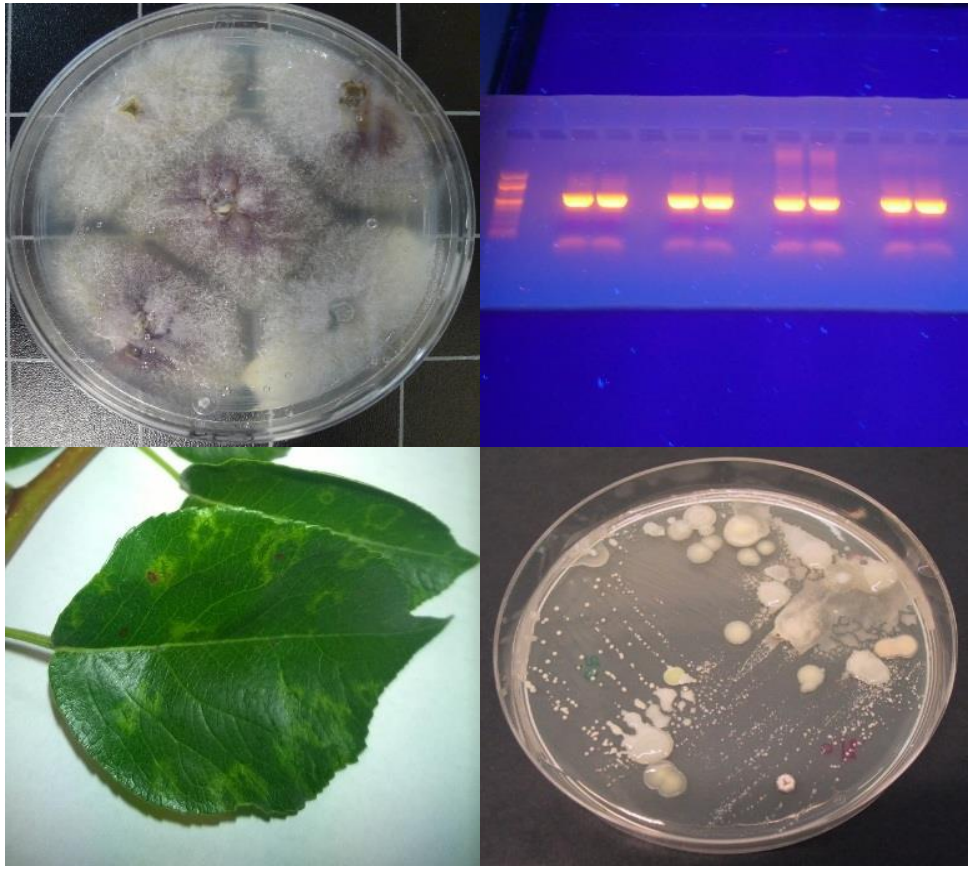


Output GGAAACAACTGAAGCTA
GGAAACAACTGAACCTA
GGATACAACTGAAGCTA

- HTS
- Targeted or non-targeted
- Nucleic acids are prepared for sequencing by library preparation to obtain right amount of right size nucleic acids with flanking adapters (depending on the platform)
- Sequencers detect the nucleic acid incorporation event during elongation or read the strand directly (depending on the platform)
- Individual sequences for each nucleic acid in the sample is produced

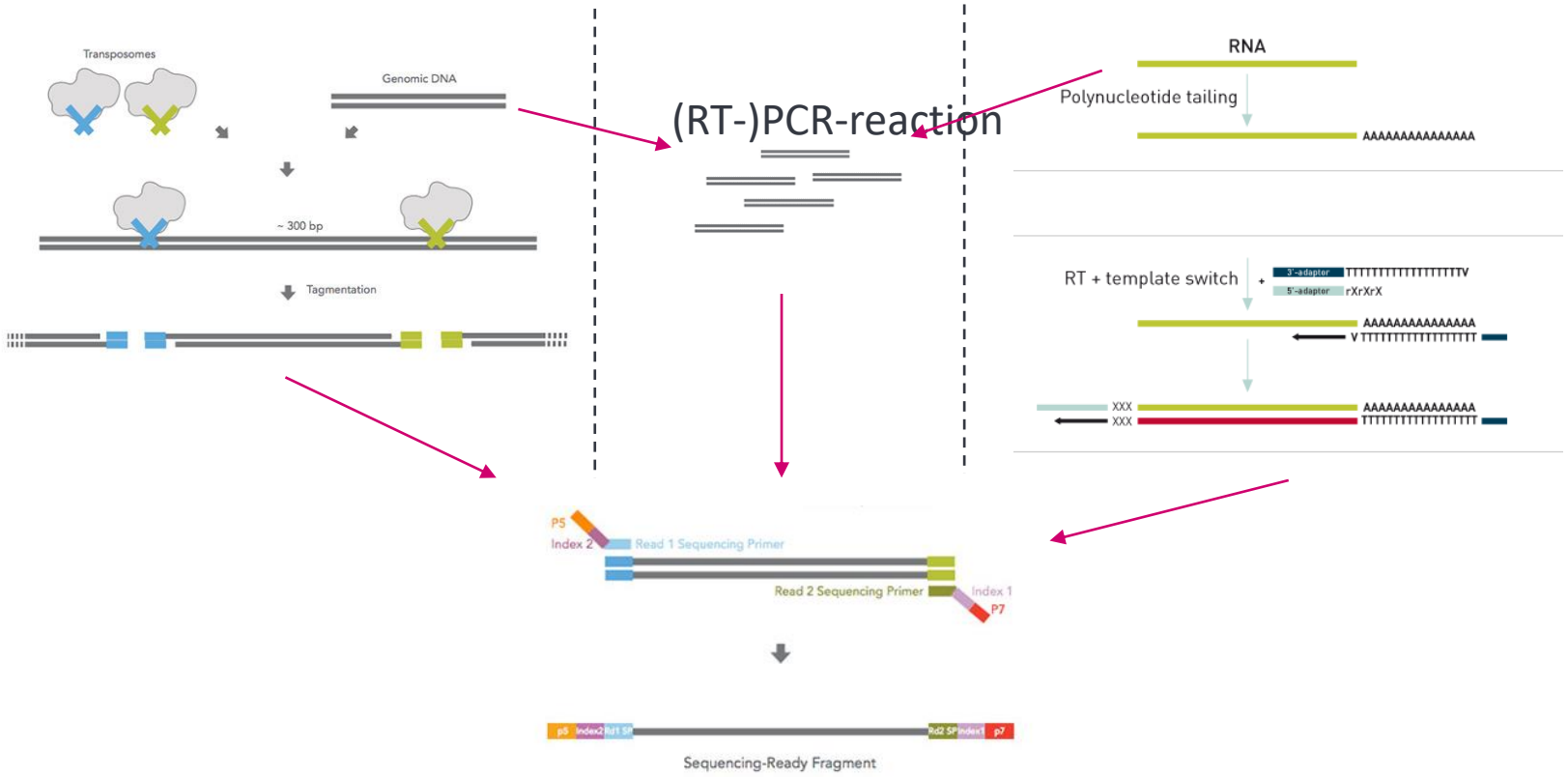
HTS procedure starts with (sampling and) nucleic acid extraction

- matrix with multiple organisms, e.g.
  - plant tissue containing microorganisms
  - environmental samples
  - spore traps
- isolated organisms e.g. microbial colonies on artificial media
- The nucleic acids can be
  - genomic DNA or RNA
  - total DNA or RNA
  - small RNAs
  - double-stranded RNAs



Library preparation

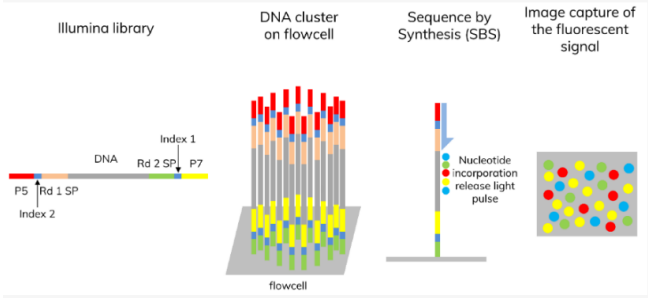
- The library preparation produces a sufficient amount of nucleic acids of appropriate size that are flanked with adapters (oligonucleotide sequences) required for sequencing
- Note, direct RNA sequencing also possible



Sequencing

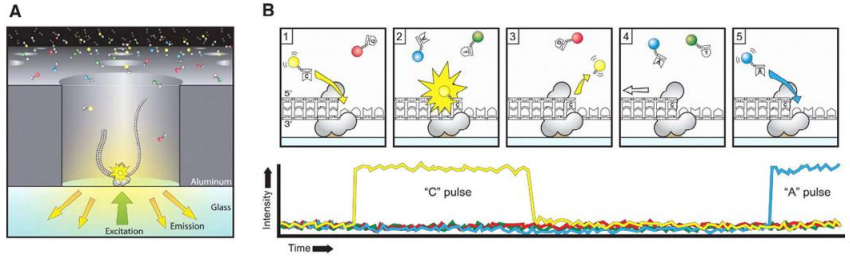
- This step produces the sequences and corresponding quality scores usually in FASTQ format

Illumina dye sequencing capturing fluorescent signals caused by incorporation of labeled nucleotides



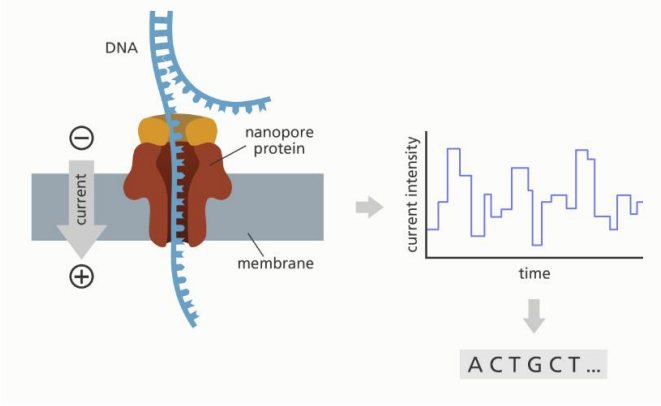
https://www.mdpi.com/2075-4418/13/3/373

PacBio Single molecule real-time sequencing measuring kinetic changes caused by nucleotide incorporation



https://www.cd-genomics.com/pacbio-smrt-system-single-molecule-real-time-sequencing.html

Nanopore sequencing monitoring changes to an electrical current as nucleic acids pass through a protein nanopore



https://www.yourgenome.org/facts/what-is-oxford-nanopore-technology-ont-sequencing/

Analysis of raw reads (including quality control)

- Bioinformatic analysis that may include:
  - Elimination of low quality reads
  - Demultiplexing for pooled samples
  - Primer/adapter/index removal
  - Dublicated or background reads removal
  - Pairing and merging reads
  - Artefacts removal

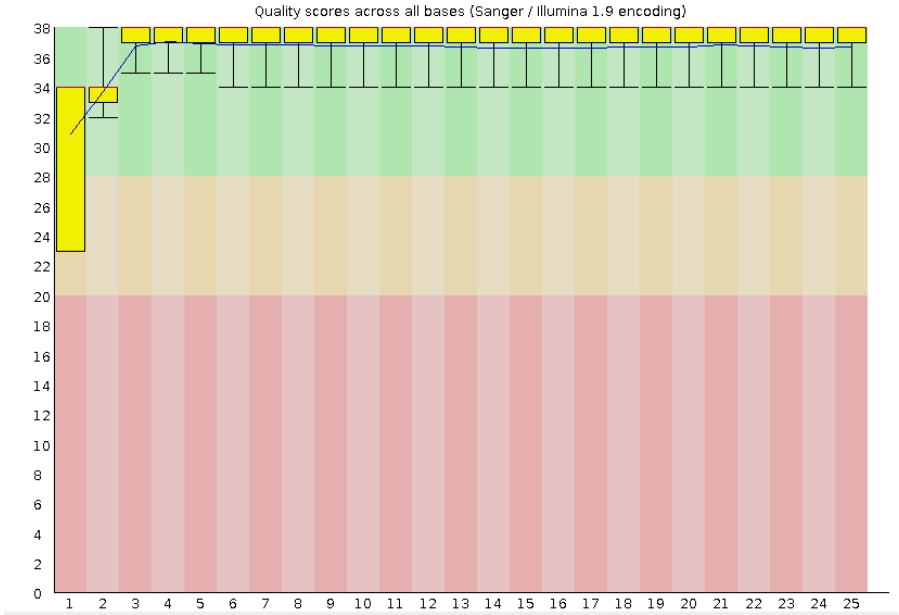
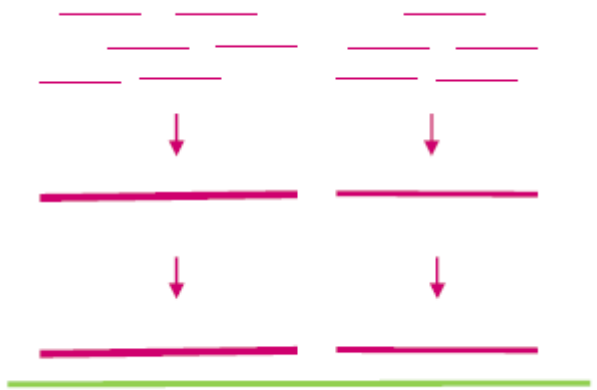


Table 1: Quality Scores and Base Calling Accuracy

Phred Quality Score	Probability of Incorrect Base Call	Base Call Accuracy
10	1 in 10	90%
20	1 in 100	99%
30	1 in 1,000	99.9%
40	1 in 10,000	99.99%
50	1 in 100,000	99.999%

Identification of targets

- Bioinformatic step including optional steps
  - Direct annotation of individual reads
  - De novo assembly of (short) reads to create longer contigs
  - Clustering of amplicon sequences based on similarity
  - Mapping of reads on reference sequence
- Leading to annotation of the sequences, e.g
  - Taxonomic classification
  - Functional annotation
  - Variant calling



Analysis of controls

- Bioinformatic step in which the controls used in the experiment are checked against predefined quality metrics and thresholds to
  - Evaluate contamination
  - Evaluate whether targets are found as expected
  - Find possible sequencing errors
  - Evaluate whether the results need to be confirmed

Applications in plant health

- Identification of pests causing novel diseases
- Development of specific tests (generating sequence data)
- Surveillance programmes
- Detection of pests in propagation material and nuclear stocks
- Post-entry quarantine testing
- Other rutine testing

RUOKAVIRASTO
Livsmedelsverket • Finnish Food Authority

Thank you

RUOKAVIRASTO

Livsmedelsverket • Finnish Food Authority