

Alliance



# Virology & Crop Protection Team

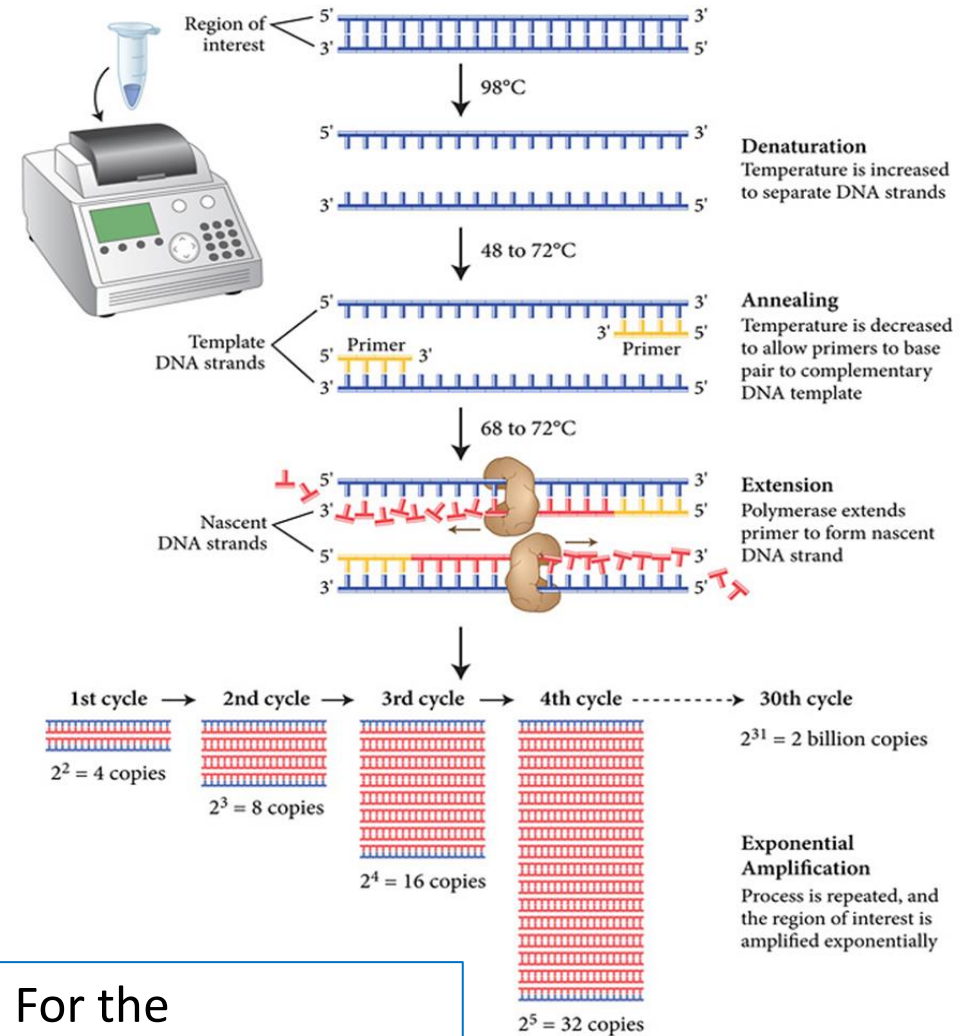
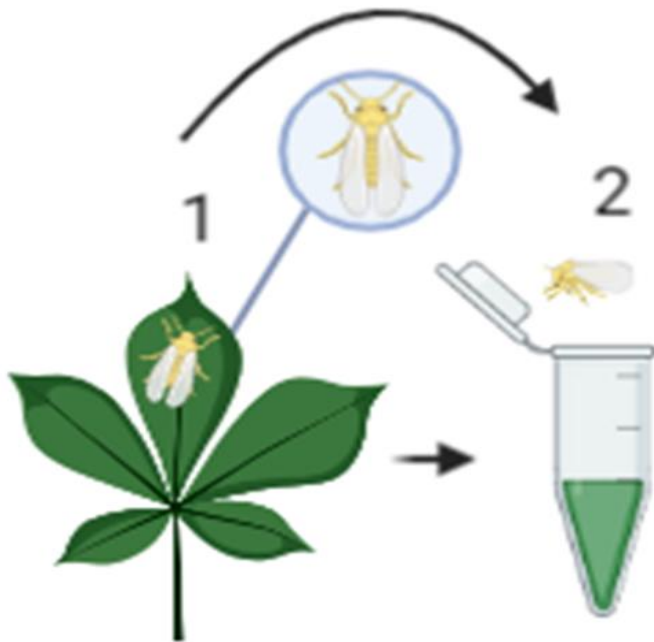
# Introduction to Nanopore sequencing



Bioversity International and the International Center for Tropical Agriculture (CIAT) are CGIAR Research Centers. CGIAR is a global research partnership for a food-secure future.

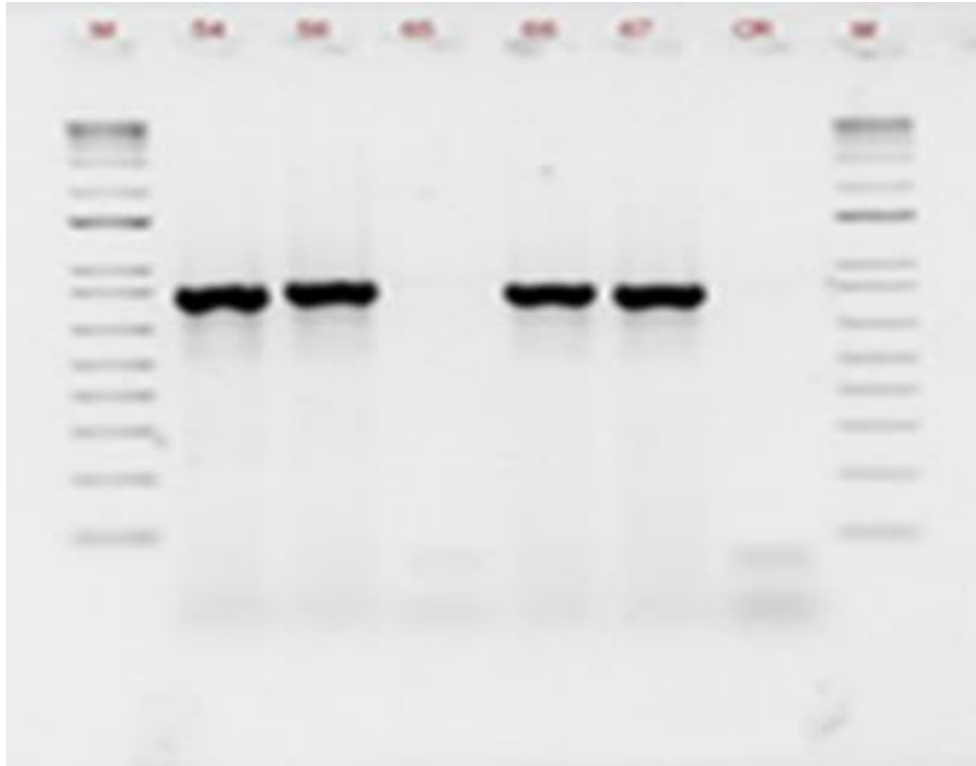
**Cassava Program  
Lever: Crops for Nutrition and Health  
Workshop Vientiane- Lao People's Democratic Republic August -  
2022**

# Polymerase chain reaction (PCR)

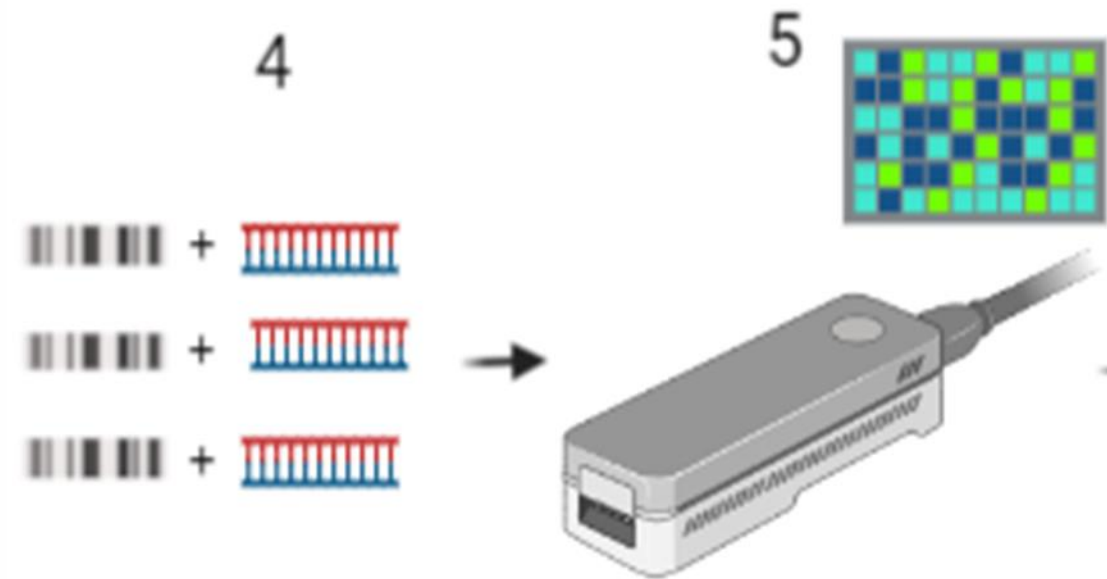
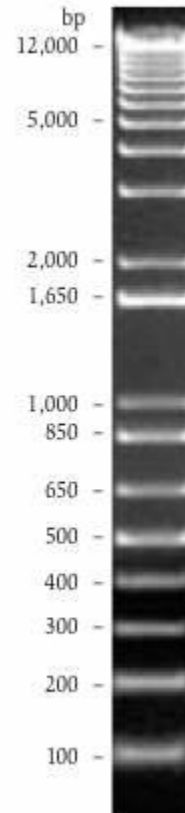


PCR using the Mitochondrial cytochrome c oxidase subunit I (COI). For the identification of critical species of *Bemisia tabaci*. (Asia II 1, Asia II 6)

# Electrophoresis, Library preparation and sequencing



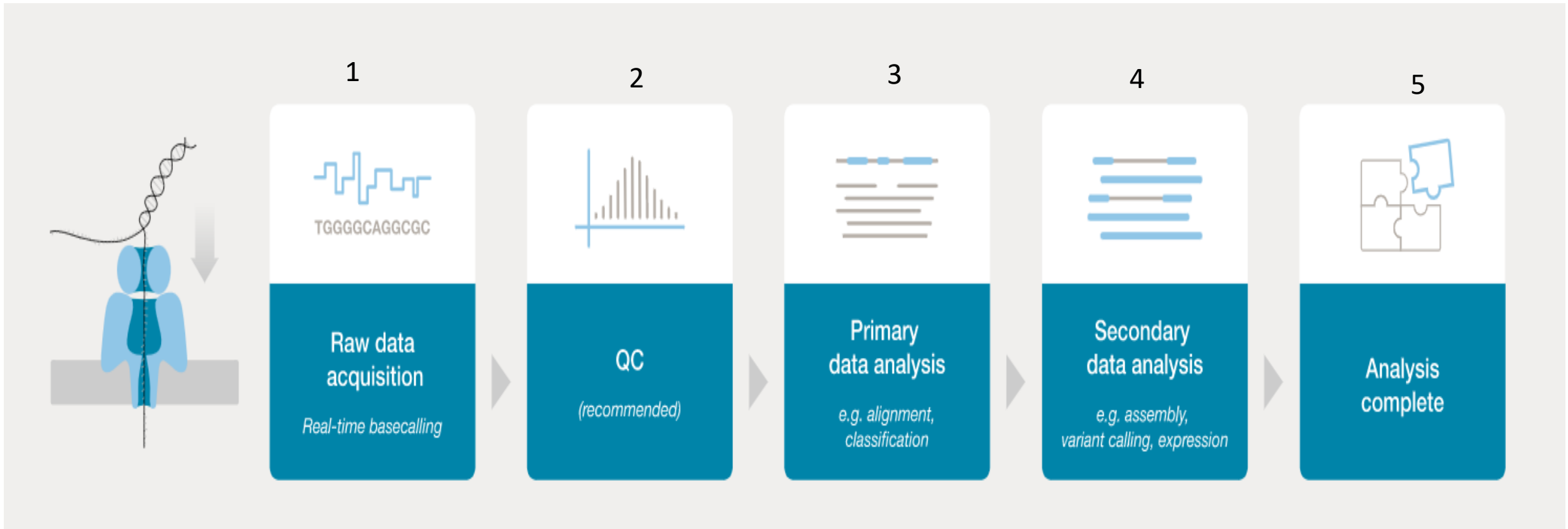
(867pb)



Library is a pool of DNA fragments with adapters attached. Adapters are designed to interact with a specific sequencing platform, either the flow-cell surface

# Analyzing Data using Bioinformatics

Bioinformatics is the use of computers to analyze biological information, especially genetics and genomics information.



Pipelines with five simple steps.

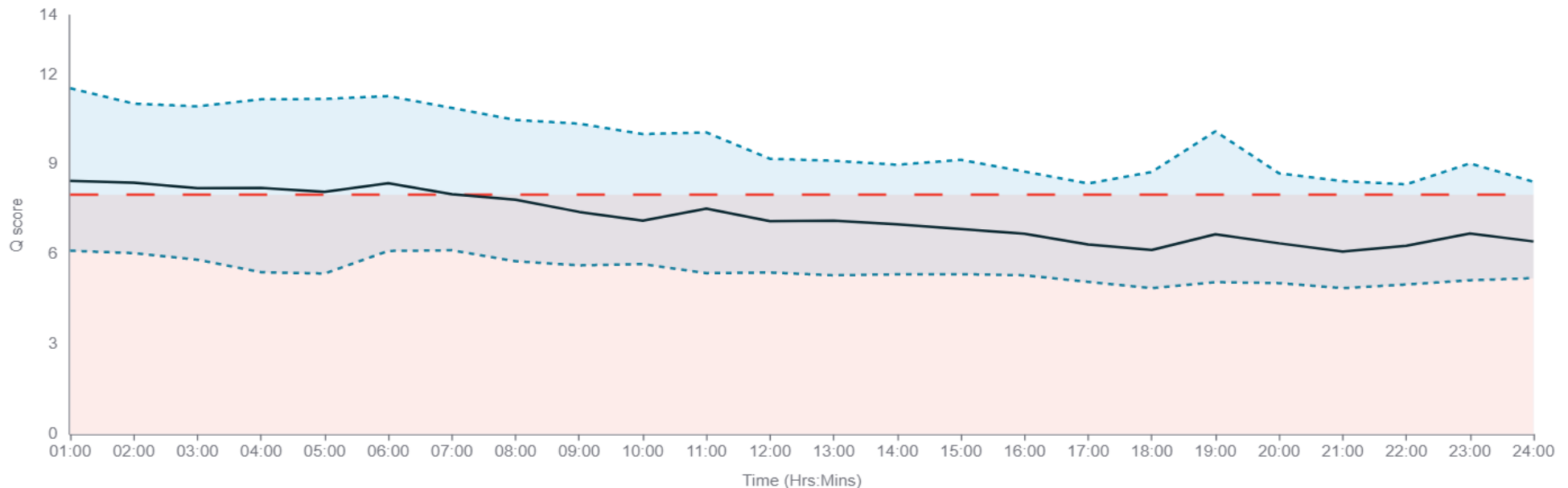
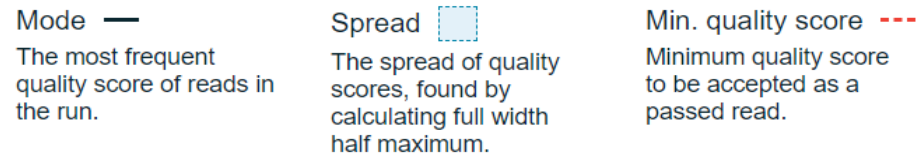
# Basecalling and QC control

MinKNOW produces FAST5 (HDF5) files and/or FASTQ files, according to your preference. FAST5 files contain raw signal data that can be used for basecalling and calling base modifications, such as methylation.

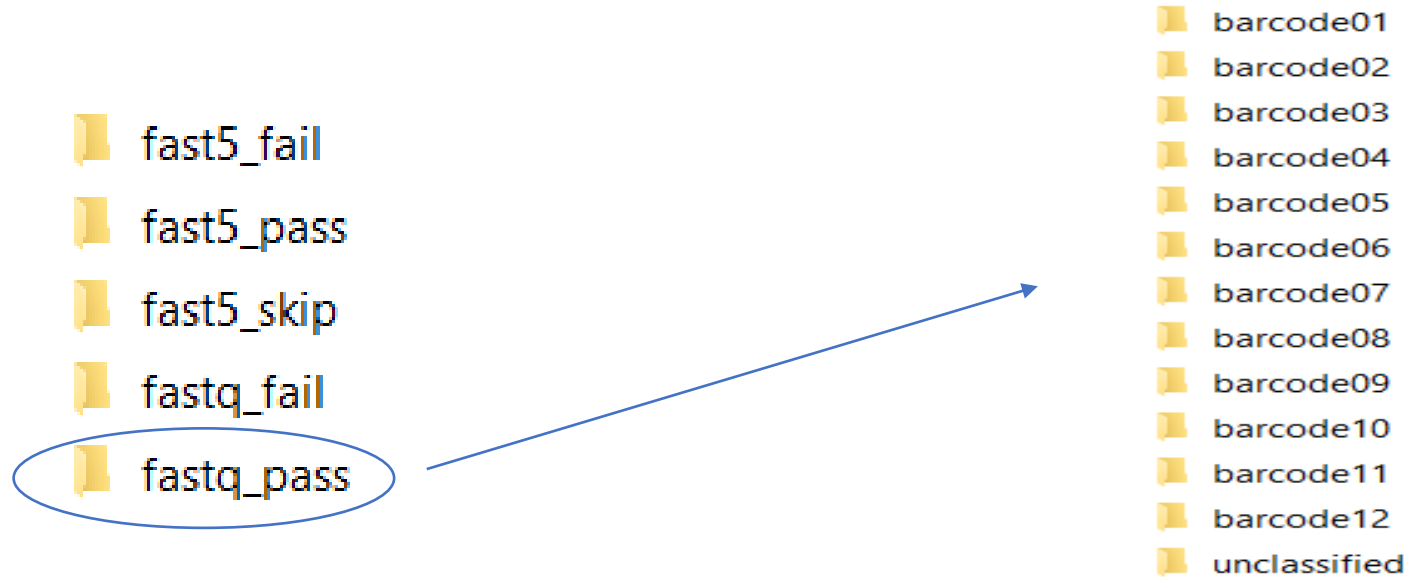
## ^ QUALITY SCORE

The quality score is calculated as basecalling is performed on your device. Quality score is measured on the logarithmic PHRED scale. Reads that fall below the minimum value of 8 will be classified as failed reads. You can alter the accepted minimum quality score in MinKNOW.

Legend



# Data Output



Instrument Name

Flowcell Info

x,y of cluster in tile

index #, member of pair

@H06HDADXX130110:2:2116:3345:91806/1

GTTAGGGTTAGGGTTGGGTTAGGGTTAGGGTTAGGGTTAGGGGTTAGGG...

Raw sequence letters

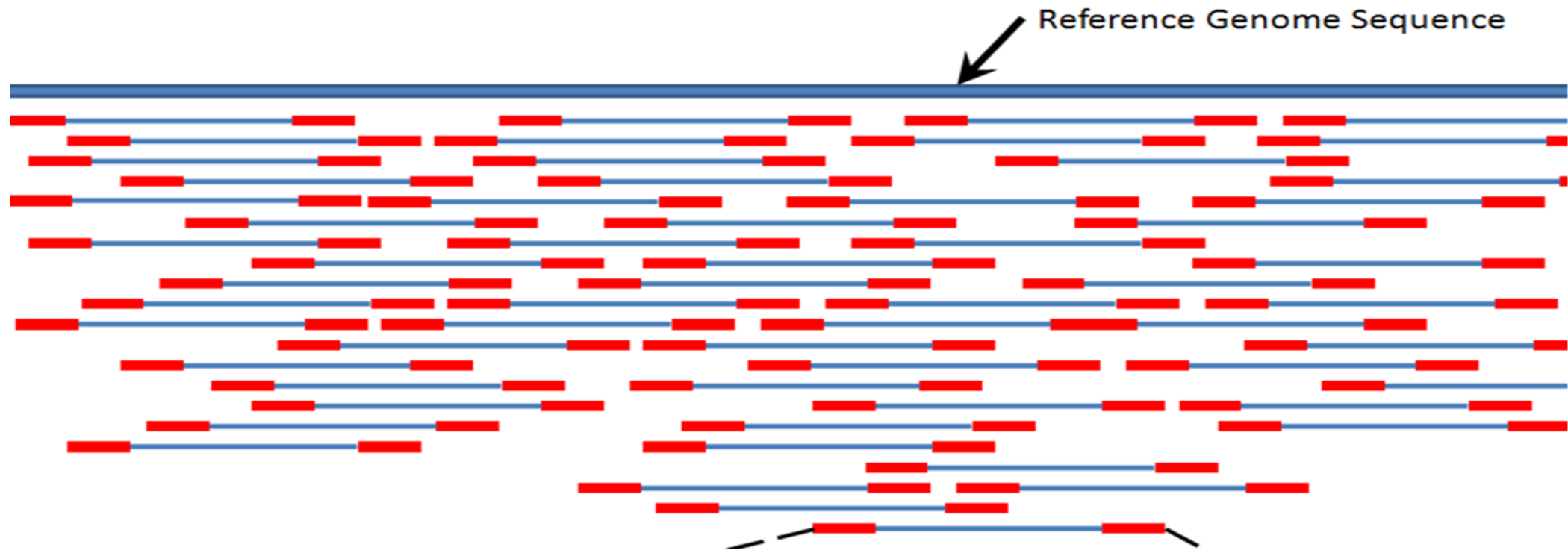
+

>=<=?>?>??=>>8<?><=2=<===1194<? ; :?>>?#3==>##...

Quality values of raw sequence letters

# Alignment data

The first phase (Alignment) involves aligning or mapping the reads to the reference genome. From the NCBI gen Bank



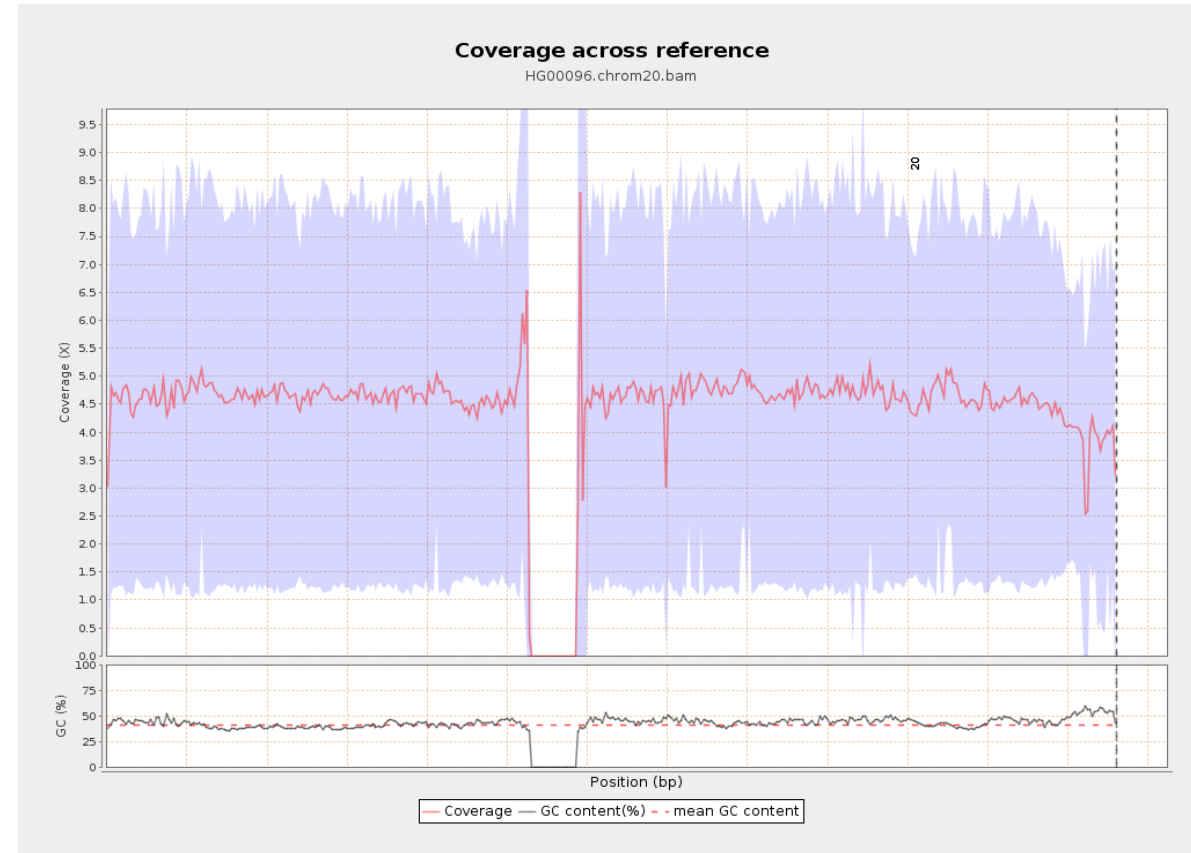
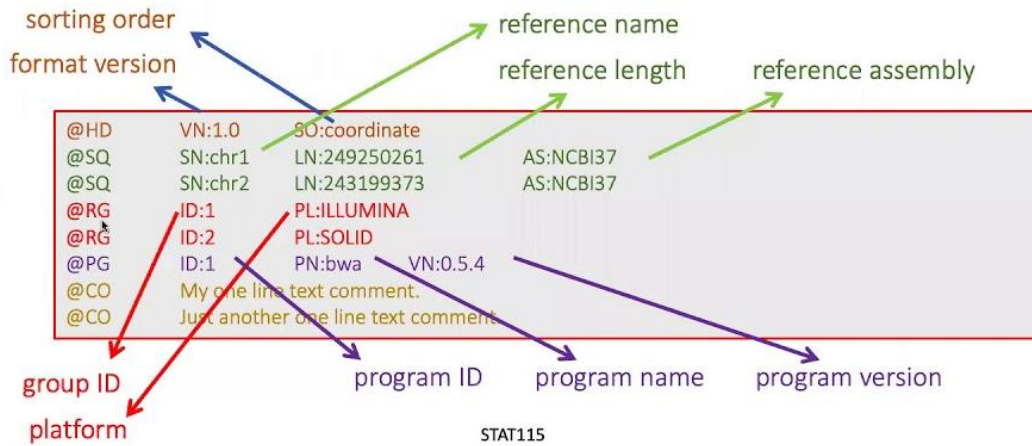
Program Minimap2

# Control to the assemble

- **Qualimap** using the files SAM and BAM.

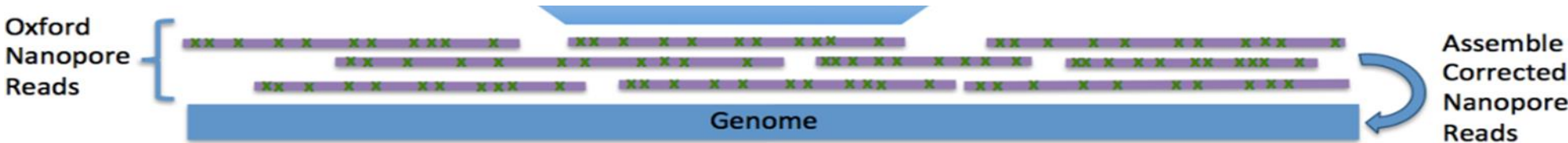
## SAM File - Header

- @HD – Header line.
- @SQ – Reference genome information.
- @RG – Read group information.
- @PG – Program (software) information.
- @CO – Commentary line.





# Assembly and polish the samples



Two program Pilon and Medaka



```

Header Sequence
● >VIT_201s0011g03530.1
● AATTAAGCATAAAATACTCACTCTTACCCCCTTATTTTCTTATCTCTCATCACTTTTGGTGCGAAG
● GACCATGAGAACAAGCTGCAATGGGTGTAGGGTCTTOGCAAGGCATGCAGCCAAGACTGCATCA
Header Sequence
● >VIT_201s0011g03540.1
● CAGGTAGCGTGAAGTTAAAOCCTAGCGCTTTAGACAAACAGCTGTAGTCAOCGCCACAAACACC
● AGCCTCTGAGACACCACCTCAAACCTTTCCACTTAAATACACATCCCTCACAOCTTTTCAATTC
Header Sequence
● >VIT_201s0011g03550.1
● CATGCAAAGCTGAACGCGATGCTGTGATTGGTGGTAAGTGGTAGTTGAGTAAATTTGACAGTGAA
● GCOGAAATGGTAAAAGACTAAGGCTAGAAGTAGAATAOCCACTGTTCTTCTCATCACGTGGGCCCA
  
```

# NCBI



The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

<https://www.ncbi.nlm.nih.gov/home/analyze/>

>Barcode12

```
GGTCATCCGGAAGTTTATGTTCTTATTTTACCAGGTTTTGGTATTGTTTCTCATTTAATTAGGAGTGAGGCTGGAAAGCTTGAAGTATTTGGC
AGATTAGGTATAATTTATGCTATAGTGACGATTGGCATTCTAGGTTTTATTGTGTGAGGTCATCATATTTACTGTTGGAATAGATGTTGATACT
CGGGCTTATTTTACTTCAGCTACTATGATTATTGCTGTTCCGACTGGAATCAAATTTTTAGGTGACTTGCTACTCTAGGTGGAATAAAGTCTA
ACATATTTAGTCCGCTTGGACTTTGGTTTGCTGGATTTCTTTTCTTATTTACTATGGGTGGATTAAGTGAATTATTCTTGGTAATTCTTCTGTT
GATGTCTGTTTACATGATACTTACTTTGTTGTTGCTCATTTTCATTATGTTTTATCTATAGGAATTATCTTTGCTATCGTGGGAGGTGTTATTTATT
GATTTCCAGTAATCTTGGGATTAACACTAAATAGTCATAGCCTGGTATCGCAGTTTTACATTATGTTTTTGGGAGTAAATTTAACGTTTTTTCCCA
CAGCATTTTCTTGGGCTGAGAGGTATACCTCGTCGCTATTCAGACTATCCTGACTGTTACCTAATATGAAATAAAATTTCTTCTGCGGGGAGA
ATTTTGAGCATTATTTCTGTTATTTATTTTTTATTATTGTCTTAGAGTCTTTACTTCTTTTGC GGTTAGTGGGTTTTAAGCTTGGTATAATTAGA
CATTTAGAATGGAAAATCAATAAGCCGGCTCTTAGGCATAGTTTTAAAGAGCTGTGTTTAA
```



**Colab** notebooks allow you to combine executable code and rich text in a single document, along with images, HTML, LaTeX and more. When you create your own **Colab**. Colab, or "Colaboratory", allows you to write and execute Python in your browser, with

- Zero configuration required
- Access to GPUs free of charge
- Easy sharing

Whether you're a **student**, a **data scientist** or an **AI researcher**, Colab can make your work easier.



[https://colab.research.google.com/?utm\\_source=scs-index#scrollTo=-Rh3-Vt9Nev9](https://colab.research.google.com/?utm_source=scs-index#scrollTo=-Rh3-Vt9Nev9)



Alliance



# Thanks!



Biodiversity International and the International Center for Tropical Agriculture (CIAT) are CGIAR Research Centers.  
CGIAR is a global research partnership for a food-secure future.



This Presentation contains imagens were taken via search engine from the internet