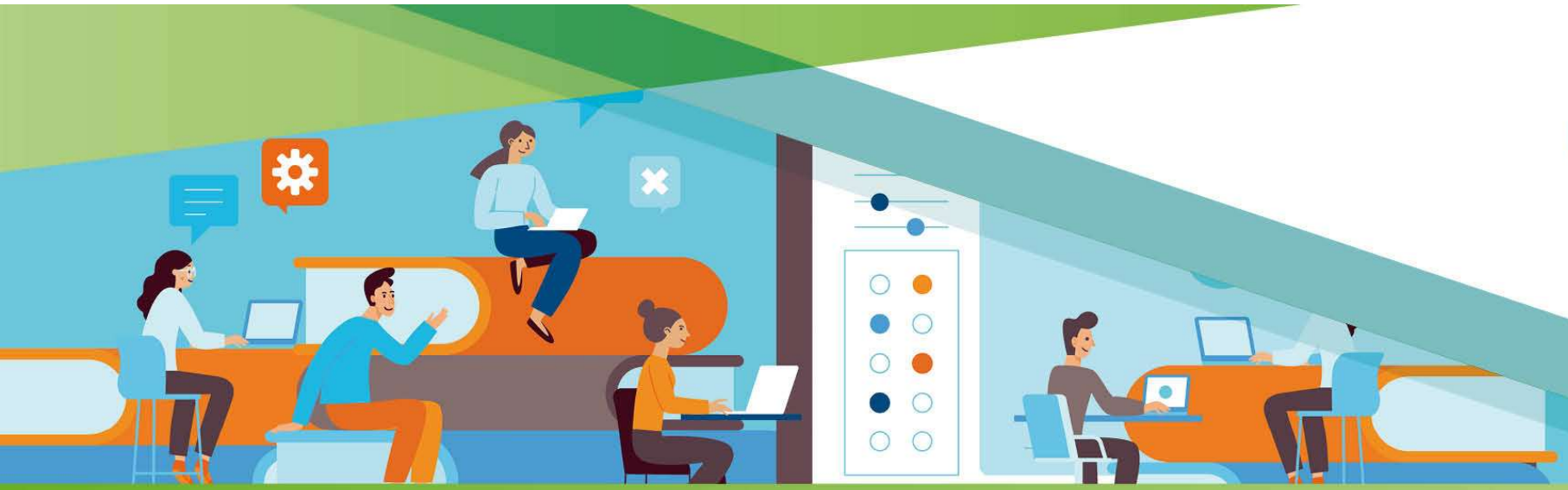


Methods for whole genome sequencing and the importance of metadata

GenEpi-BioTrain Virtual Training 29 - Genomic epidemiology of mpox virus (MPXV)



5 October 2026

Dr. David F. Nieuwenhuijse

Aim and learning objectives

Aim

Familiarize yourself with whole genome sequence analysis and data interpretation

Specific learning objectives

At the end of this session, you will be able to:

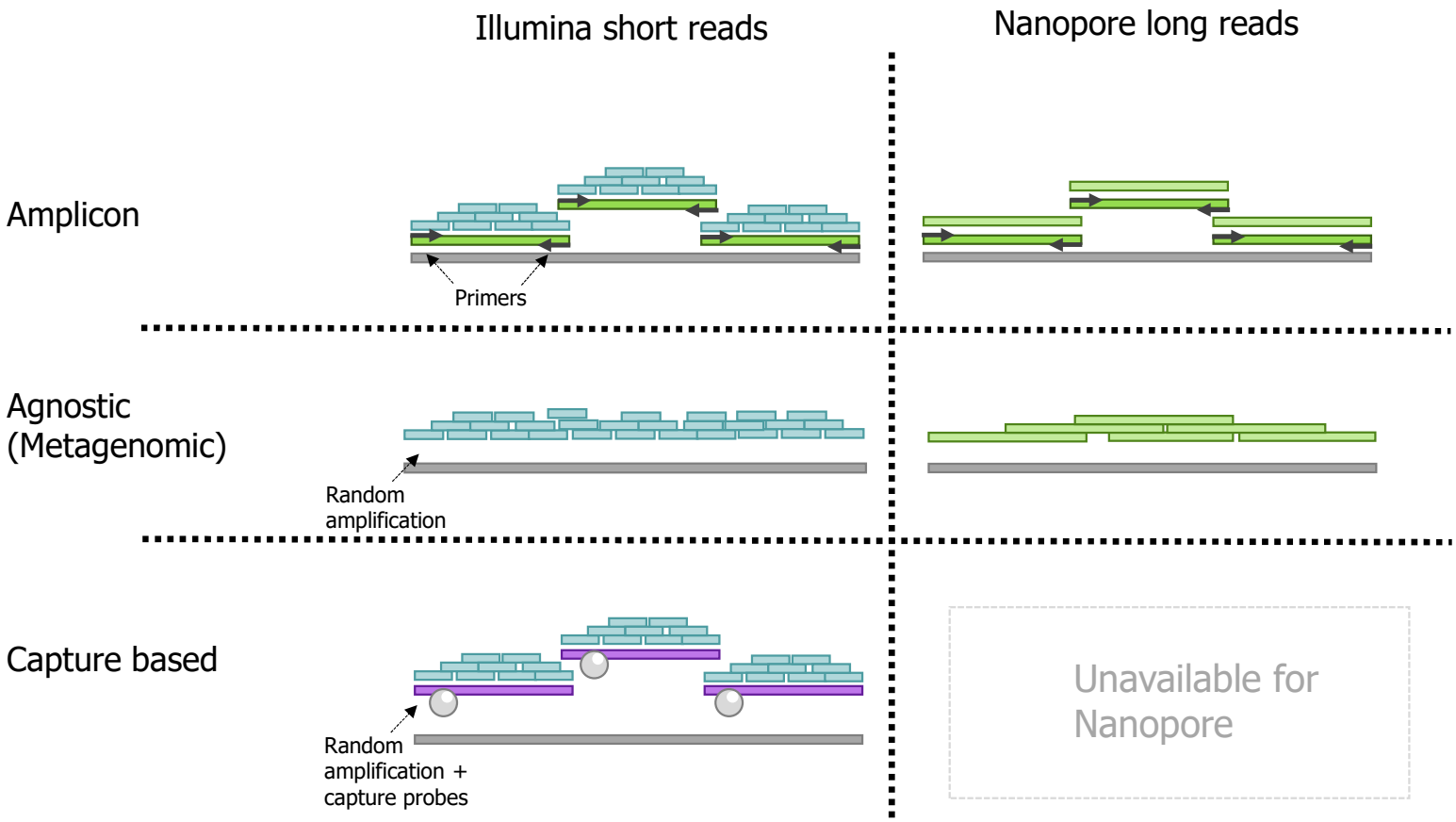
1. List different methods of whole genome sequencing and discuss their advantages and disadvantages
2. Explain the bioinformatic steps needed to generate a consensus sequence
3. Interpret sequencing errors and understand how to investigate and solve them

Outline

This session consists of the following elements

1. Introduction of different whole genome sequencing methods
2. Step by step explanation of consensus generation
3. Learning how to spot consensus generation errors

Several different approaches can be used for whole genome sequencing



Classified as ECDC NORMAL

Do not edit
How to change the design



What WGS technology have you personally used before?

ⓘ The Slido app must be installed on every computer you're presenting from

slido

Comparing metagenomic sequencing to amplicon sequencing

Four viruses

- *ZIKV*
- *USUV*
- *WNV*
- *YFV*



Three dilutions:

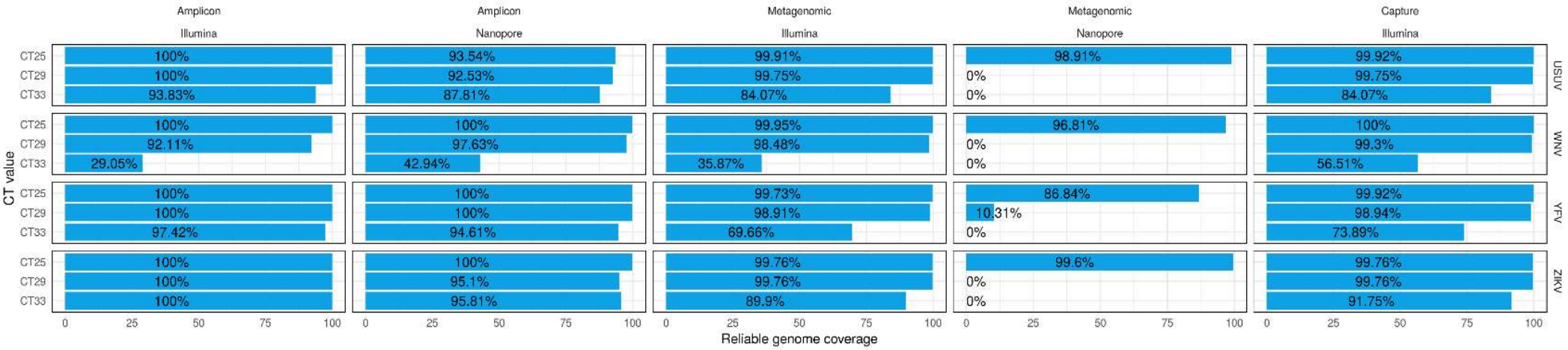
- *Ct 25 > 29 > 33*
  High Low

- Amplicon Illumina
- Amplicon Nanopore
- Metagenomic Illumina
- Metagenomic Nanopore
- Capture based Illumina

Nieuwenhuijse, D.F., van der Linden, A., Kohl, R.H.G. *et al.* **Towards reliable whole genome sequencing for outbreak preparedness and response.** *BMC Genomics* **23**, 569 (2022). https://doi.org/10.1186/s12864-022-08749-5

Classified as ECDC NORMAL

Most strategies lead to almost complete genome coverage above



Classified as ECDC NORMAL

Other parameters to consider for the different sequencing methods

Sequence method	Amplicon	Amplicon	Metagenomic	Metagenomic	Capture-based
Platform	Illumina	Nanopore	Illumina	Nanopore	Illumina
Time to result	3 days	<24 hours	3 days	<24 hours	8 days
Costs of the instrument*	€200.000	€1.000	€200.000	€1.000	€200.000
Costs per sample	€200	€50	€200	€50	€300
Virus specific PCR required	Yes	Yes	No	No	No
Ct threshold for WGS	Ct33	Ct29	Ct29	Ct25	Ct33
Ct threshold for detection	Ct33	Ct33	Ct33	Ct33	Ct33

* Estimated costs are laboratory and/or country dependent

Classified as ECDC NORMAL

Long read (Nanopore) amplicon-based sequencing



scientific reports

Explore content · About the journal · Publish with us

nature › scientific reports › articles › article

Article | Open Access | Published: 18 February 2020

Genomic monitoring to understand the emergence and spread of Usutu virus in the Netherlands, 2016–2018

B. Bas Oude Munnink, E. Münger, D. F. Nieuwenhuijse, R. Kohl, A. van der Linden, C. M. E. Schapendonk, H. van der Jeugd, M. Kik, J. M. Rijks, C. B. E. M. Reusken & M. Koopmans

Scientific Reports **10**, Article number: 2798 (2020) | Cite this article

3136 Accesses | **18** Citations | **15** Altmetric | Metrics



Eurosurveillance Europe's journal on infectious disease surveillance, epidemio

Current · Archives · Print Editions · Collections · About Us · Editorial Policies

Home / Eurosurveillance / Volume 25, Issue 40, 08/Oct/2020 / Article

Rapid communication — Open Access

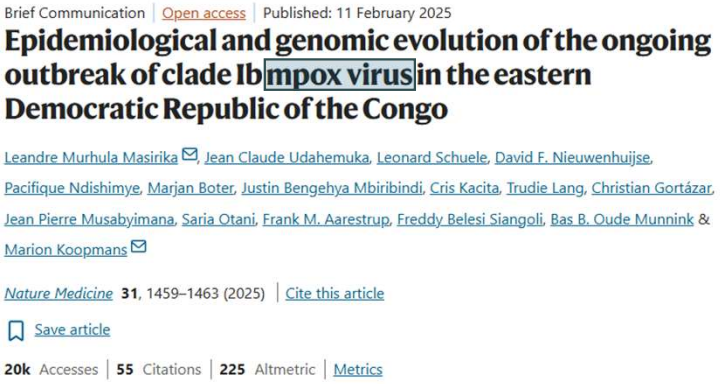
Detection of West Nile virus in a common whitethroat (*Curruca communis*) and *Culex* mosquitoes in the Netherlands, 2020

Like 0 · Download · Check for updates

Reina S Sikkema[1], Maarten Schrama[2], Tijs van den Berg[3], Jolien Morren[3], Emmanuelle Munger[1], Louie Krol[2,4], Jordy G van der Beek[4], Rody Blom[7], Irina Chestakova[1], Anne van der Linden[1], Marjan Boter[1], Tjomme van Mastrigt[3,8,9], Richard Molenkamp[1], Constantianus JM Koenraadt[7], Judith MA van den Brand[5,6], Bas B Oude Munnink[1], Marion PG Koopmans[1], Henk van der Jeugd[3]

View Affiliations

View Citation



Brief Communication | Open access | Published: 11 February 2025

Epidemiological and genomic evolution of the ongoing outbreak of clade Ib mpox virus in the eastern Democratic Republic of the Congo

Leandre Murhula Masirika, Jean Claude Udahemuka, Leonard Schuele, David F. Nieuwenhuijse, Pacifique Ndishimye, Marjan Boter, Justin Bengehya Mbiribindi, Cris Kacita, Trudie Lang, Christian Gortázar, Jean Pierre Musabyimana, Saria Otani, Frank M. Aarestrup, Freddy Belesi Siangoli, Bas B. Oude Munnink & Marion Koopmans

Nature Medicine **31**, 1459–1463 (2025) | Cite this article

Save article

20k Accesses | **55** Citations | **225** Altmetric | Metrics



nature medicine

Explore content · About the journal · Publish with us

nature › nature medicine › letters › article

Letter | Published: 16 July 2020

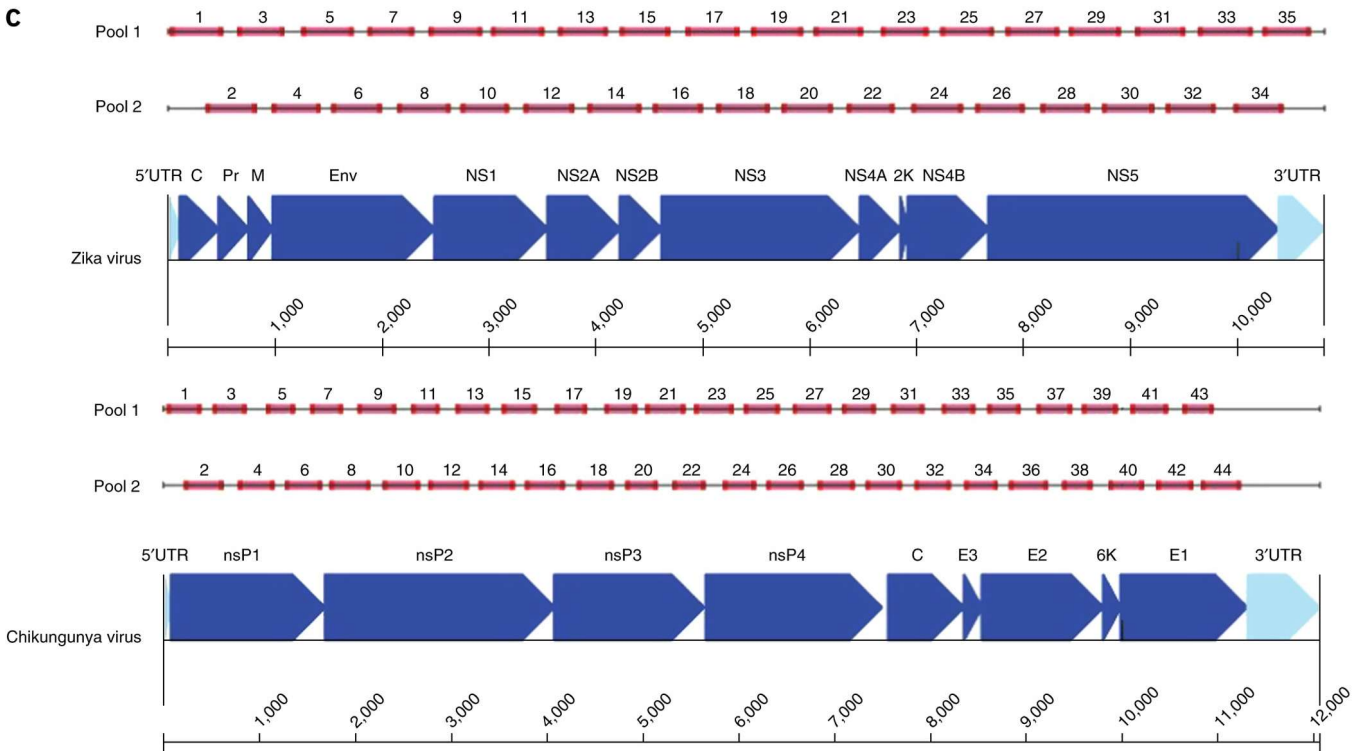
Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands

Bas B. Oude Munnink, David F. Nieuwenhuijse, … The Dutch-Covid-19 response team

Classified as ECDC NORMAL

Each new virus needs another specifically developed primer set

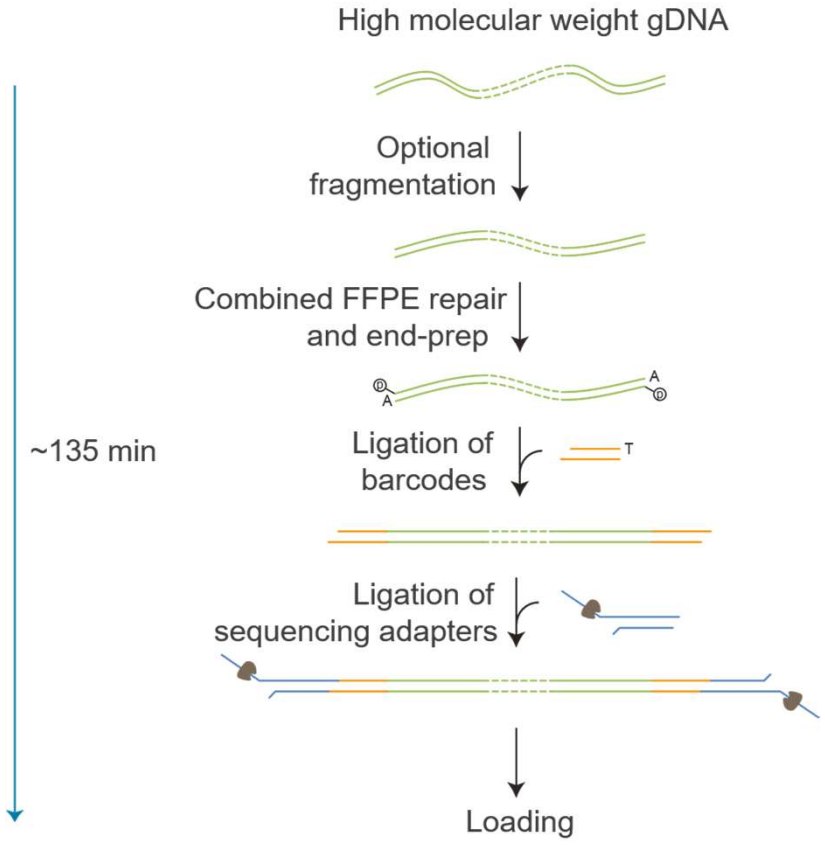
- Design specific amplicon primers
- PCR reaction using two separate primer pools



Quick, J., Grubaugh, N., Pullan, S. *et al.* **Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples**. *Nat Protoc* **12**, 1261–1276 (2017). https://doi.org/10.1038/nprot.2017.066

Classified as ECDC NORMAL

Barcode ligation, pooling and sequencing



- Possibility to multiplex 96 samples on a flowcell
- Takes around 1 full day in the lab due to wash steps and Qubit measurements

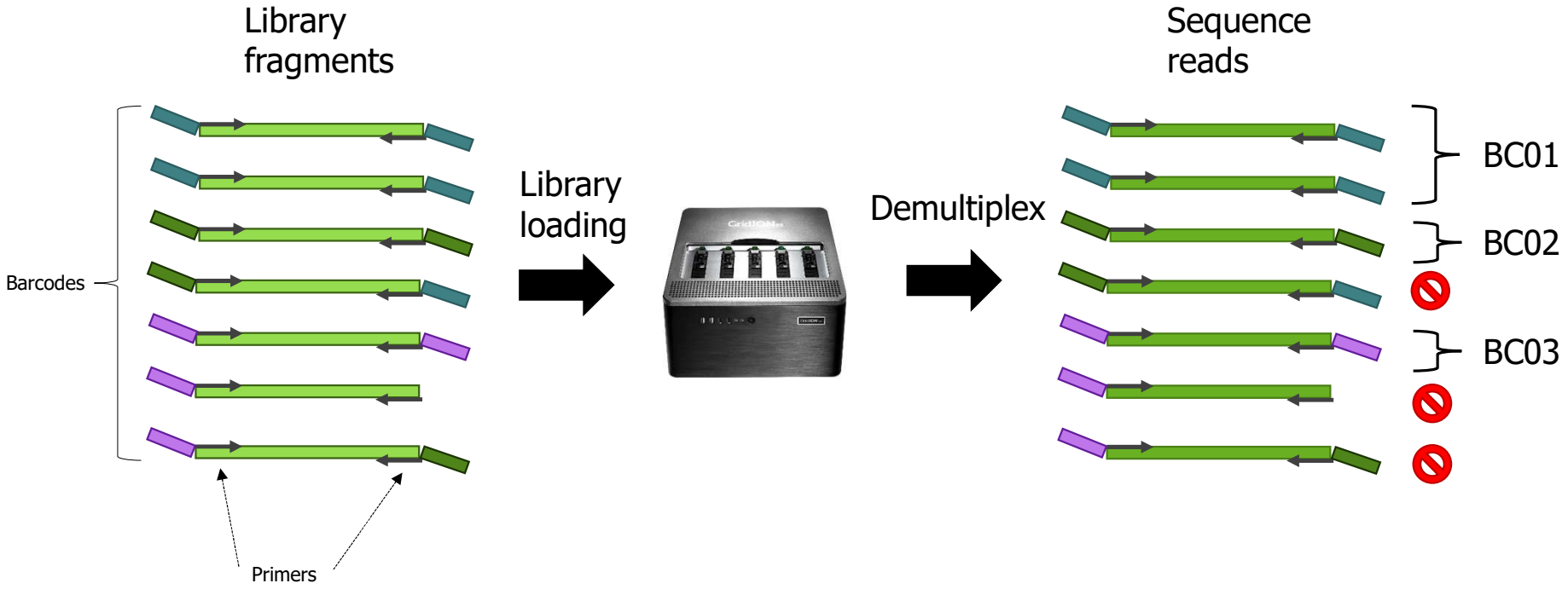
Classified as ECDC NORMAL

Bioinformatics: Generating a consensus sequence

1. Demultiplexing
2. Data Quality Control
3. Primer trimming
4. Read mapping
5. Variant calling
6. Consensus generation

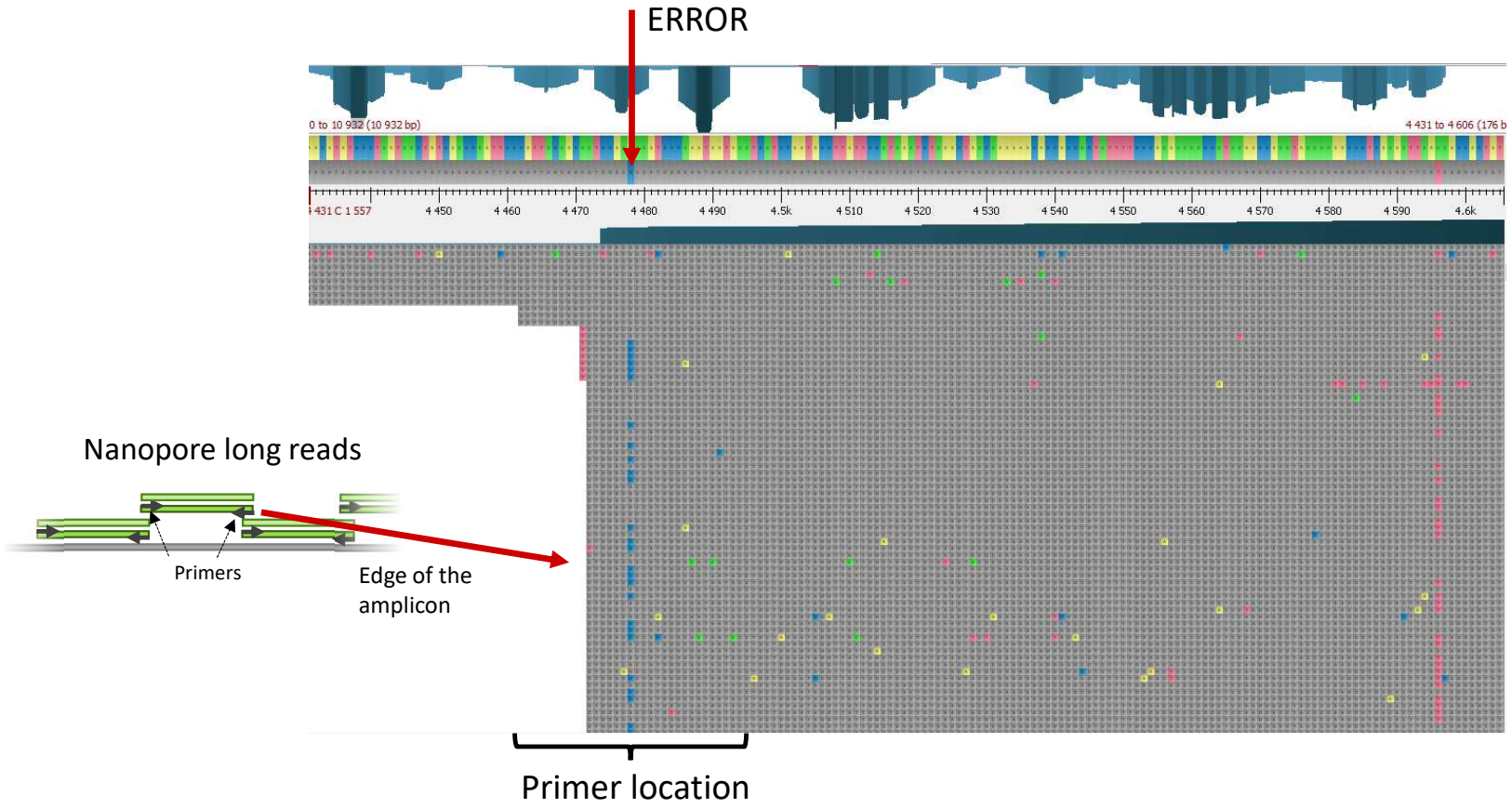
Classified as ECDC NORMAL

Demultiplexing raw reads



Classified as ECDC NORMAL

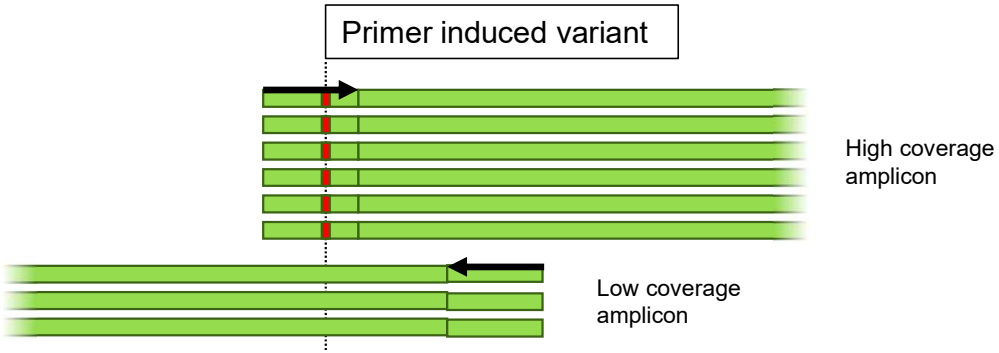
Amplicon primer trimming



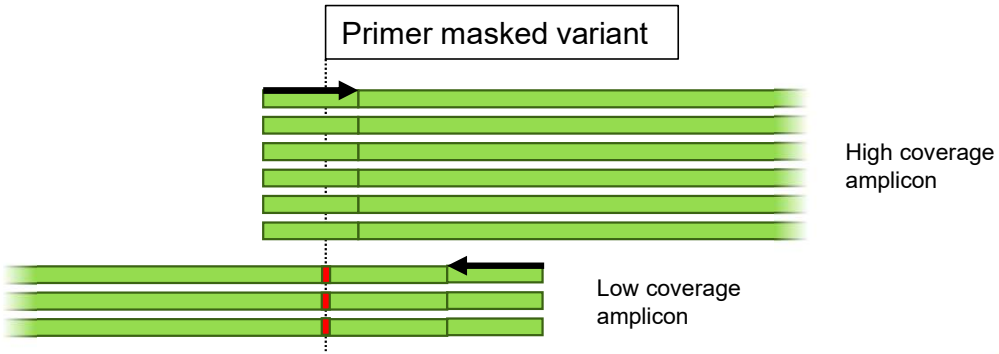
Classified as ECDC NORMAL

What causes these errors?

1. Primer SNP in stead of actual sequence

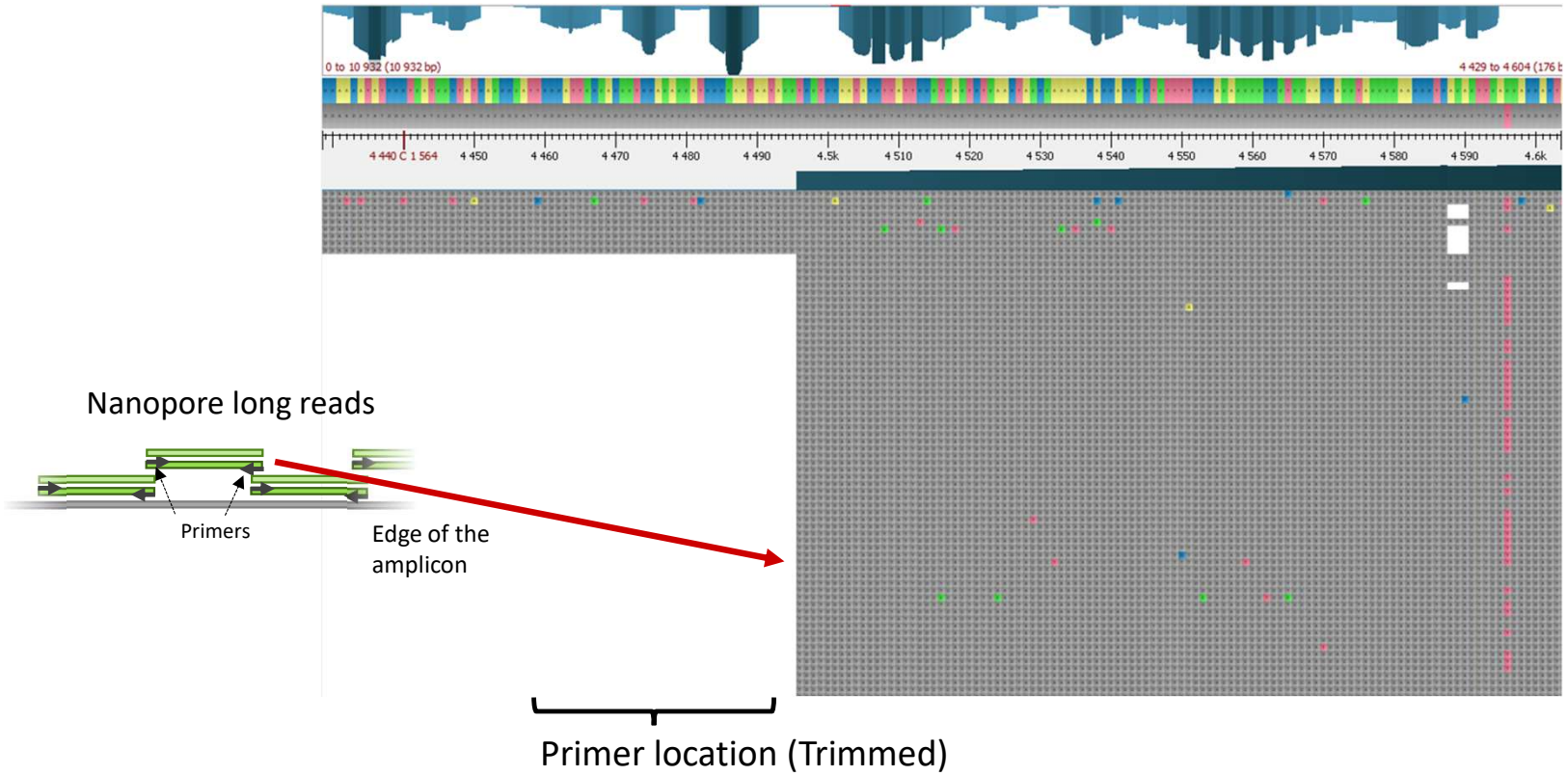


2. Primer sequence in stead of SNP



Classified as ECDC NORMAL

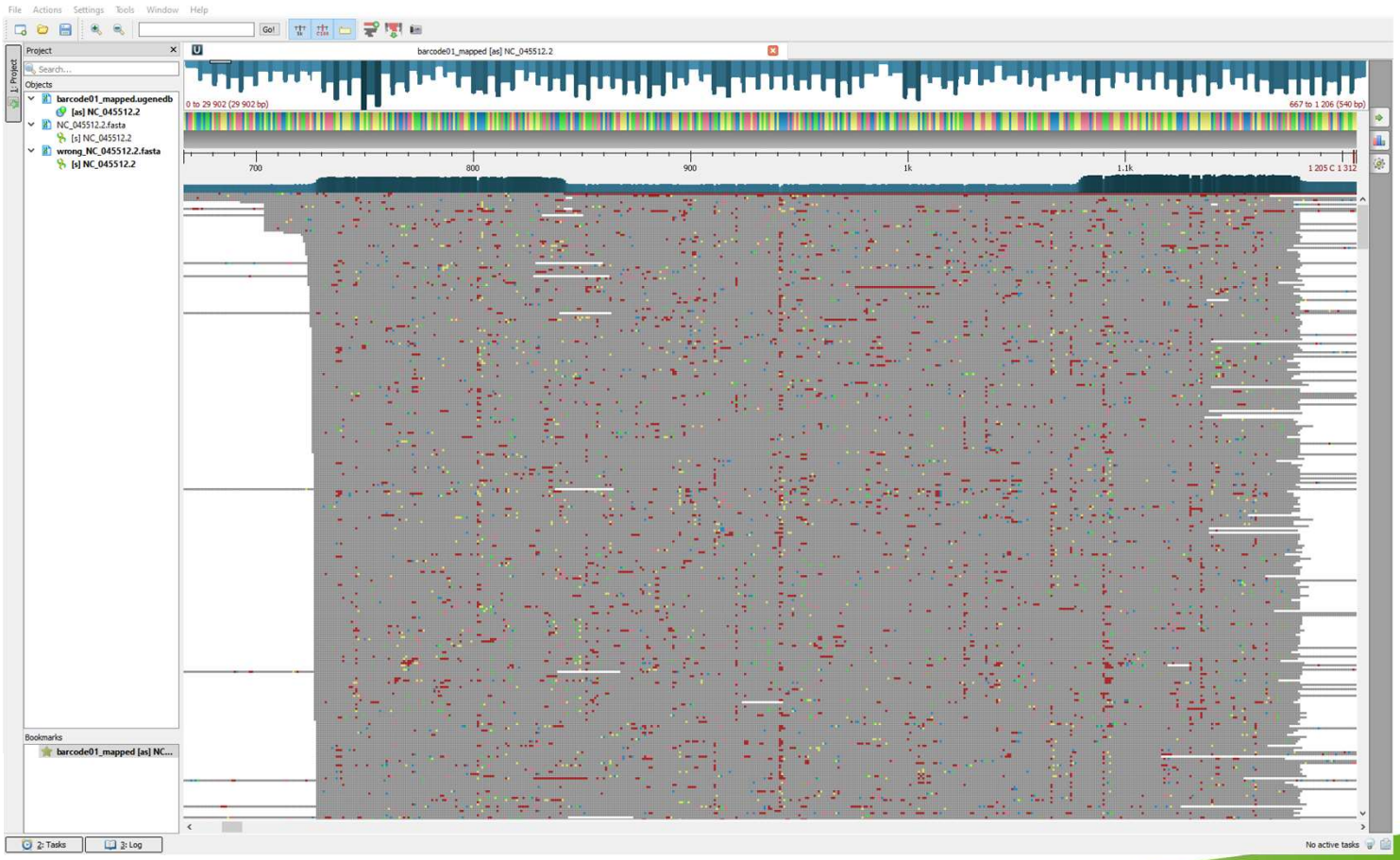
Trimming primer completely removes SNP



Classified as ECDC NORMAL

Minimap2: read mapping

- Map reads against close reference
- Nanopore:
  - High error rate
  - Many random SNPs
  - Many small deletions

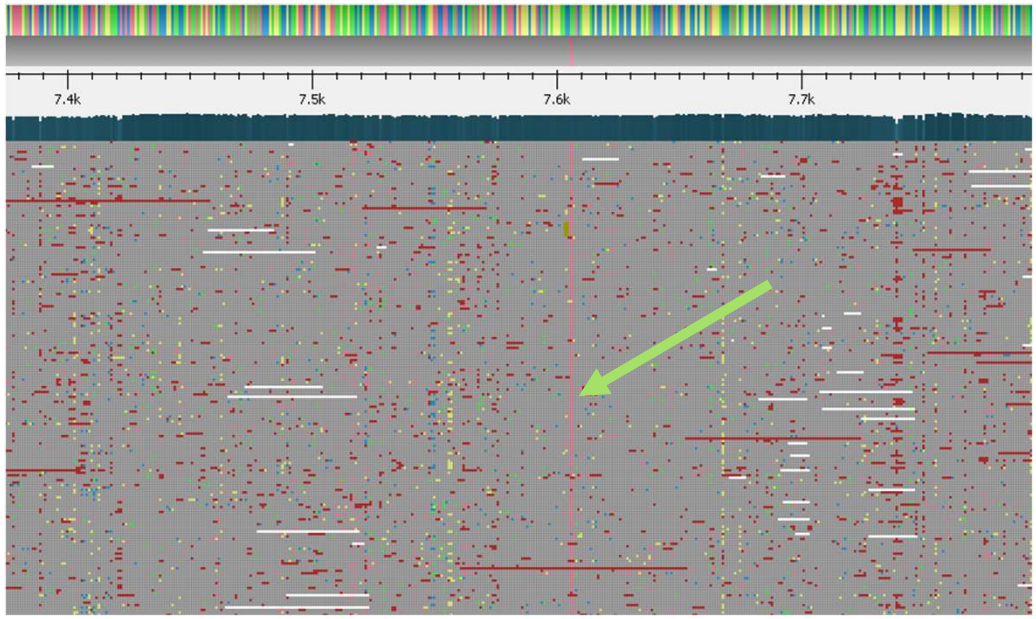


Unipro UGENE

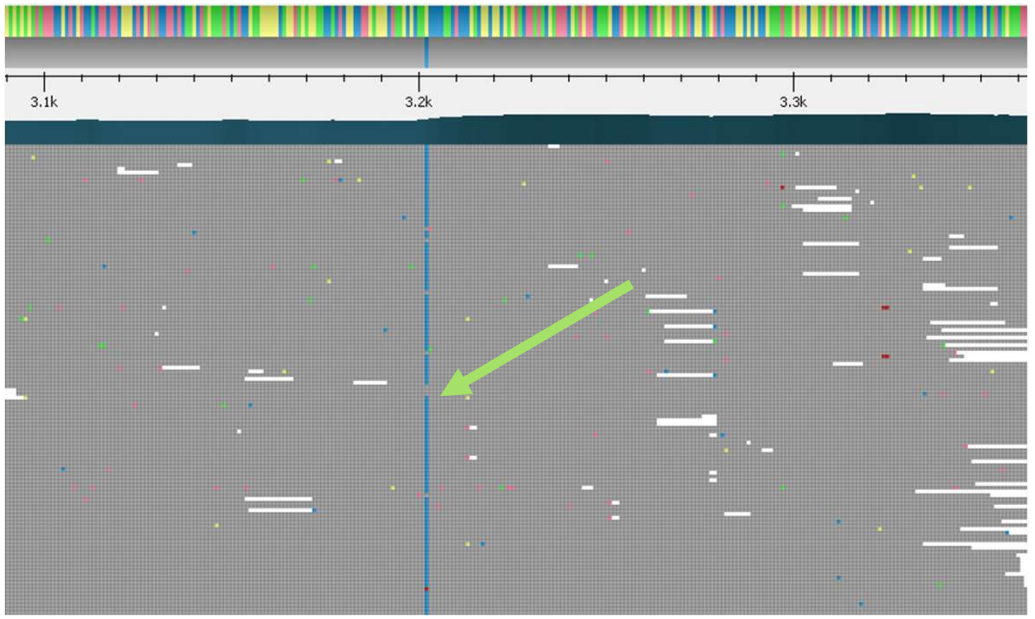
Classified as ECDC NORMAL

Variant calling

Nanopore sequencing data



Illumina sequencing data



Classified as ECDC NORMAL

Major vs minor variants

Variant calling can be complicated with Nanopore errors

Easy to call variants:

Major



>95% alternative nucleotide

Minor



<<50% alternative nucleotide

Complex to call variants:

Minor (ambiguous)



~50% alternative nucleotide

Major (non-error)

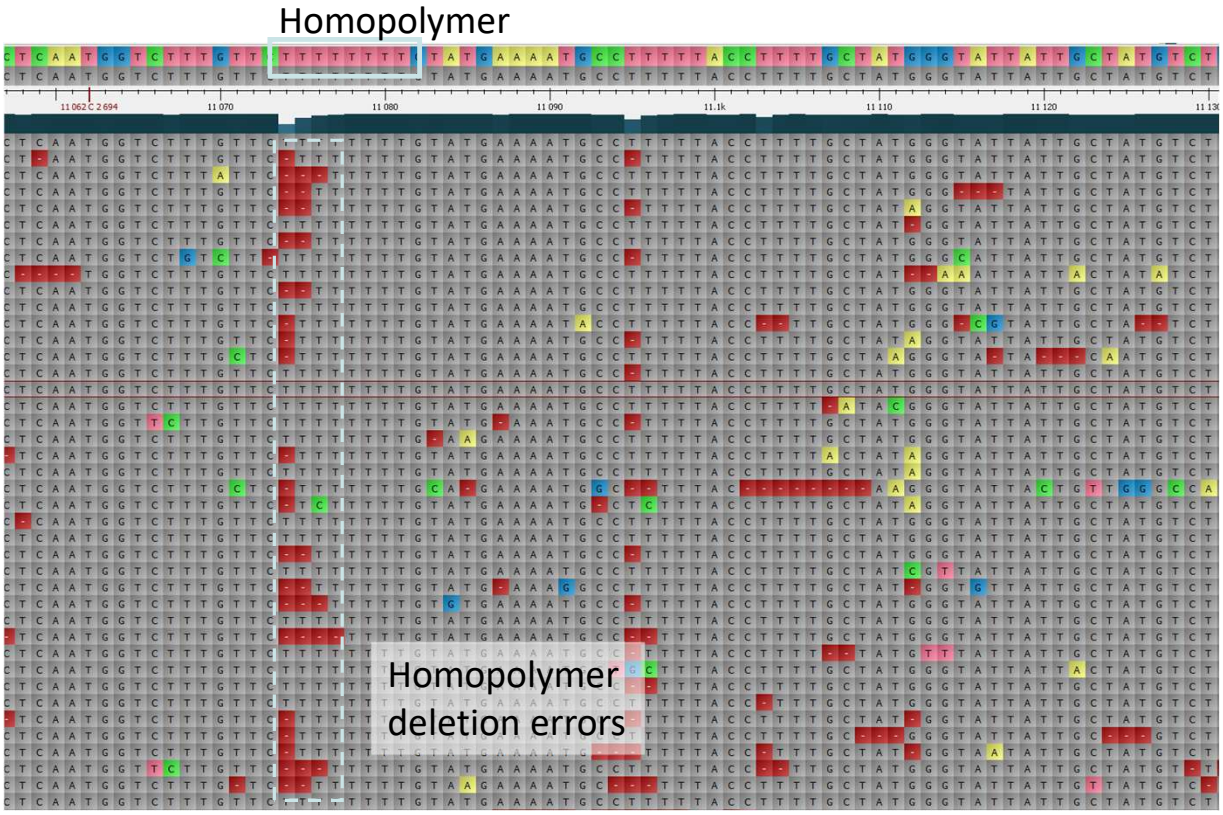


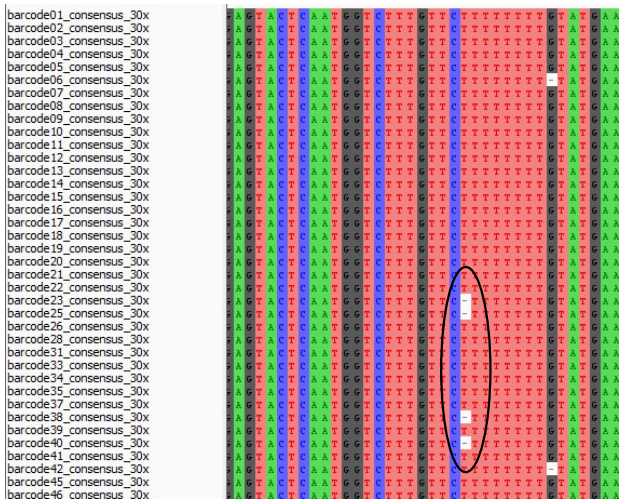
~33% alternative nucleotide
~66% error

Ensure you have enough coverage (30x)

Classified as ECDC NORMAL

Homopolymer errors





Result in multiple sequence alignment

Do not edit
How to change the design



Which of these consensus generation issues were you aware of?

ⓘ The Slido app must be installed on every computer you're presenting from

slido

Consensus generation

- Incorporate major consensus SNPs, insertions, and deletions
- Mask low coverage regions (below 30x) with “N”
- Manually curate (homopolymer) errors

Classified as ECDC NORMAL