

Introduction to Sequencing for Global Health

UChicago Center in Paris

Paris, France

January 2024

What is sequencing?

What is sequencing?

- Sequencing determines the order of the four bases (A, T, G and C) that make up DNA
- The first DNA sequences were obtained in the early 1970s
- Requires both laboratory and bioinformatics skills!

The Human Genome Project

- Began October 1990 and completed April 2003
- Researchers aimed to decipher the human genome in three major ways:
 - Determining the sequence of all the bases in the human genome
 - Mapping the locations of genes for major sections of our chromosomes
 - Producing linkage maps, so inherited traits (e.g. for genetic disease) can be tracked over generations
- The sequence is derived from the DNA of several volunteers
- Cost \$2.7 billion!
- Mainly used Sanger sequencing technology

What is sequencing?

Sanger



Illumina
(short-read)

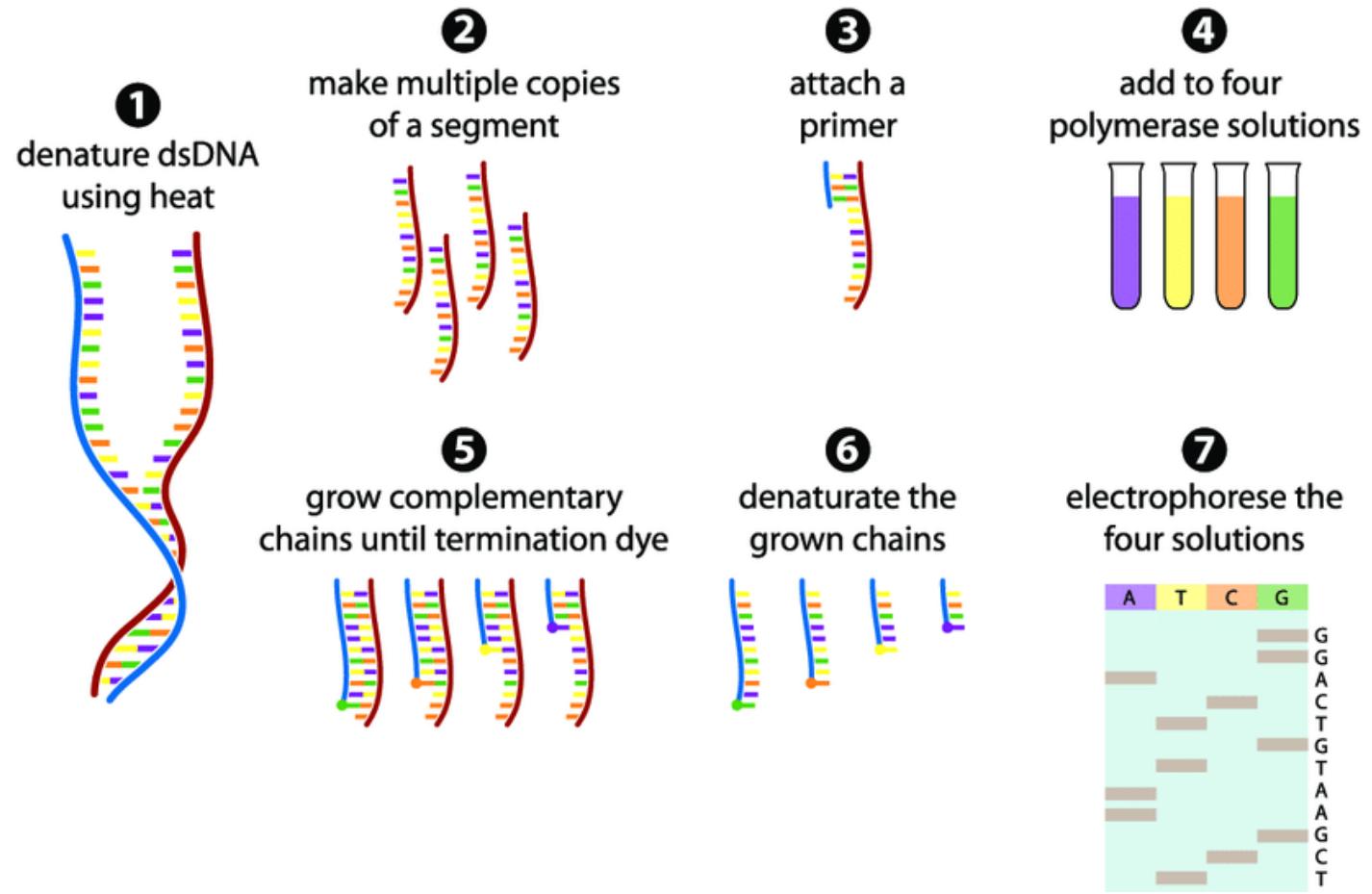


Nanopore
(long-read)



Sanger Sequencing

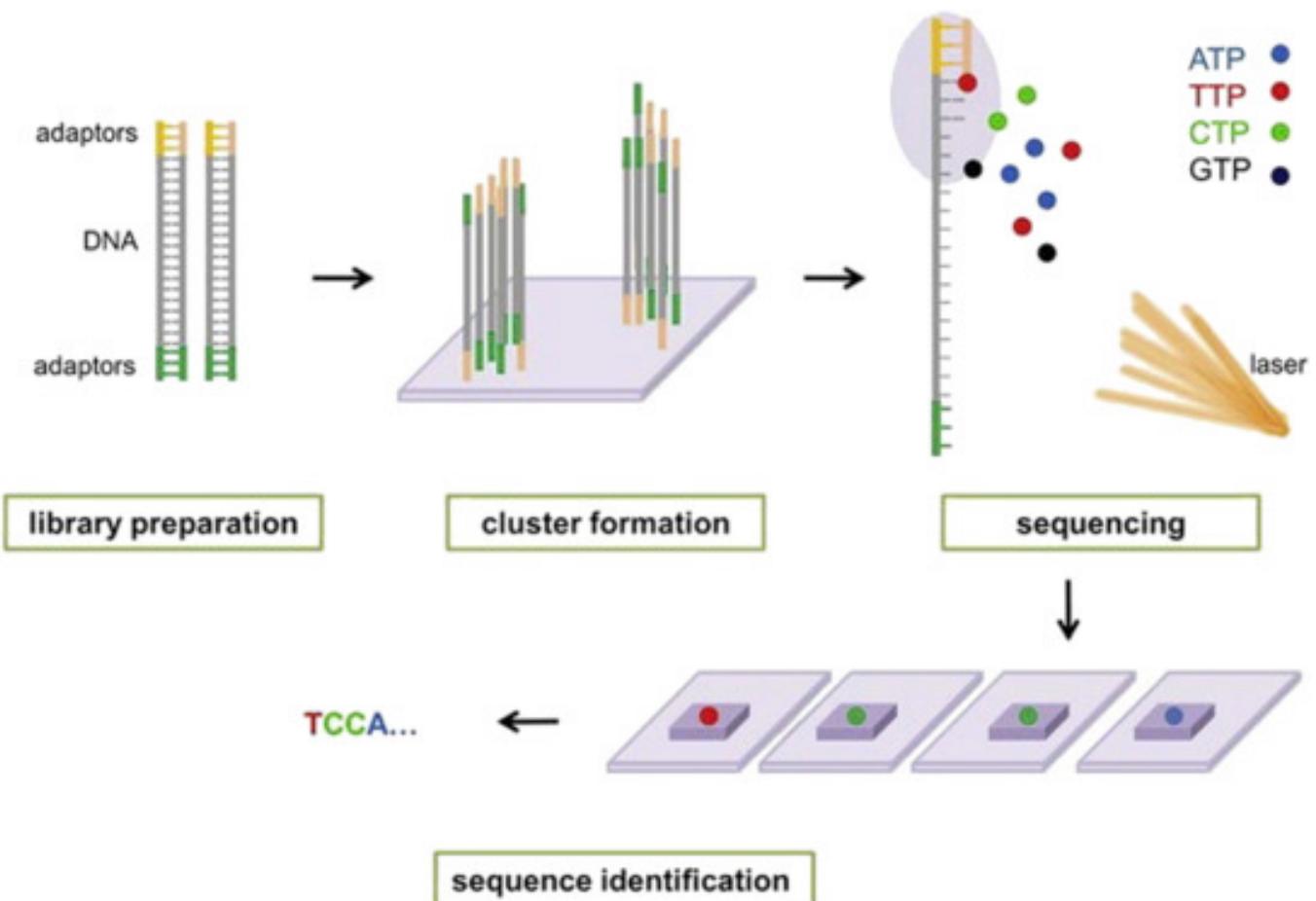
- The DNA sample is divided into four separate reactions, containing all four of the standard deoxynucleotides (dATP, dGTP, dCTP and dTTP) and a DNA polymerase (which attaches the dNTPs)
 - To each reaction is added only one dideoxynucleotide (ddATP, ddGTP, ddCTP, or ddTTP)
 - Four separate reactions are needed in this process to test all four ddNTPs
 - The ddNTP stops the DNA polymerase when it comes to a base of that type (e.g. A, T, G, C)
 - The fragments are then run on a gel. The smallest move through the gel furthest and the ‘ladder’ shows the sequence of the DNA



<https://www.youtube.com/watch?v=FvHRio1yyhQ&t=171s>

Illumina (short read) Sequencing

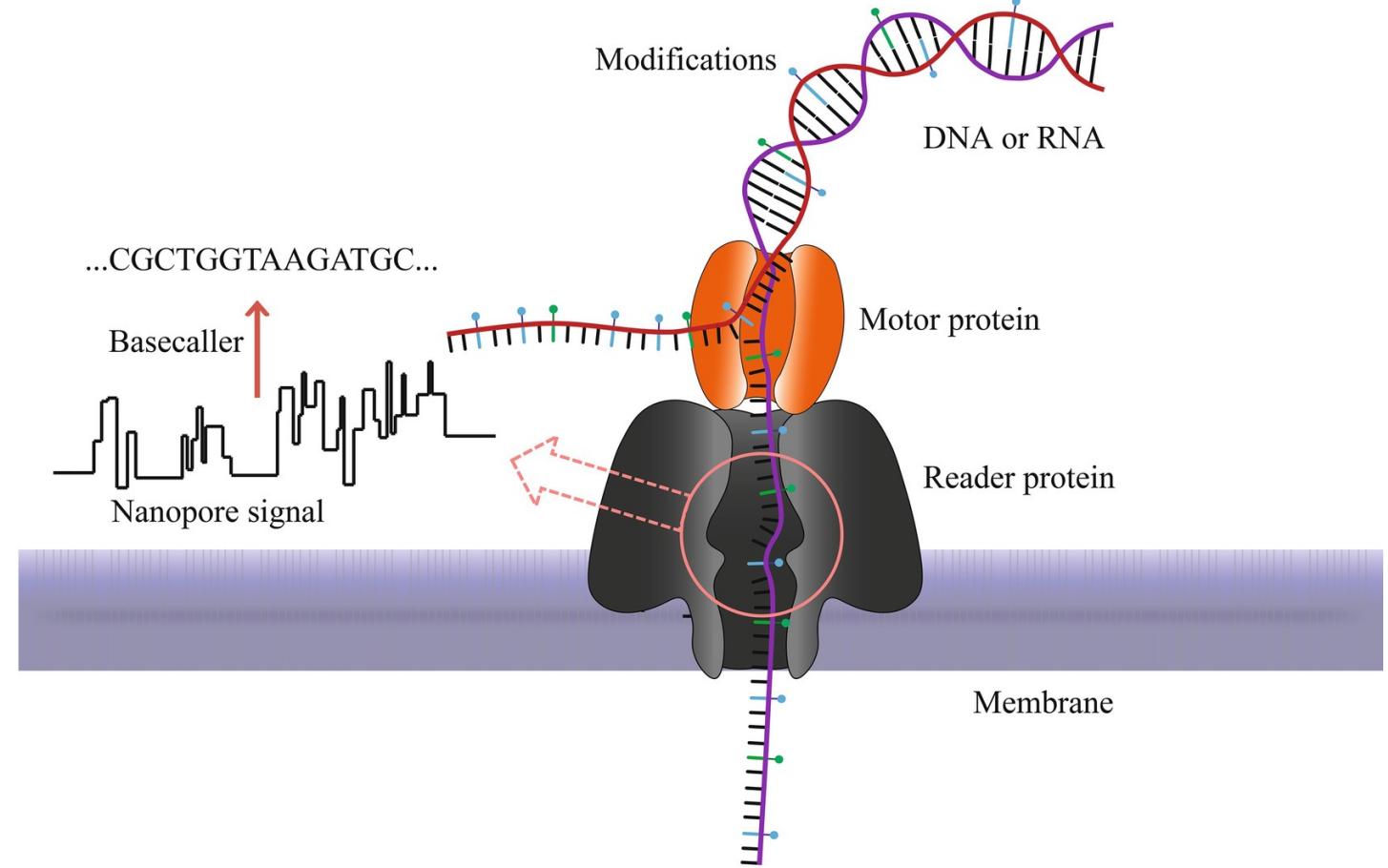
- Sequencing by synthesis
- ‘Short read’ technology – DNA is cut up into 200-600 bp chunks
- The DNA is amplified, so there are lots of copies of the chunks
- They are denatured, then fluorescent complimentary bases are attached
- These fluoresce different colours, which is recorded, and the sequence is identified



<https://www.youtube.com/watch?v=womKfikWIxM>

Nanopore (long read) Sequencing

- 'Long read' technology
- A DNA library is prepared (proteins are added)
- Nucleic acids are passed through a protein nanopore
- As the different bases move through the nanopore, it creates a different electrical signal
- These resulting changes in the electrical signal is decoded to provide the specific DNA or RNA sequence

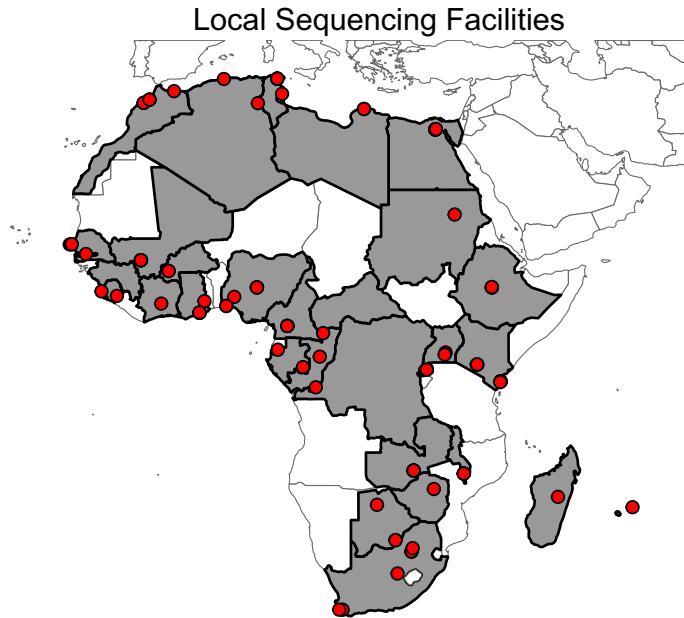


<https://www.youtube.com/watch?v=RcP85JHLmnl&t=18s>

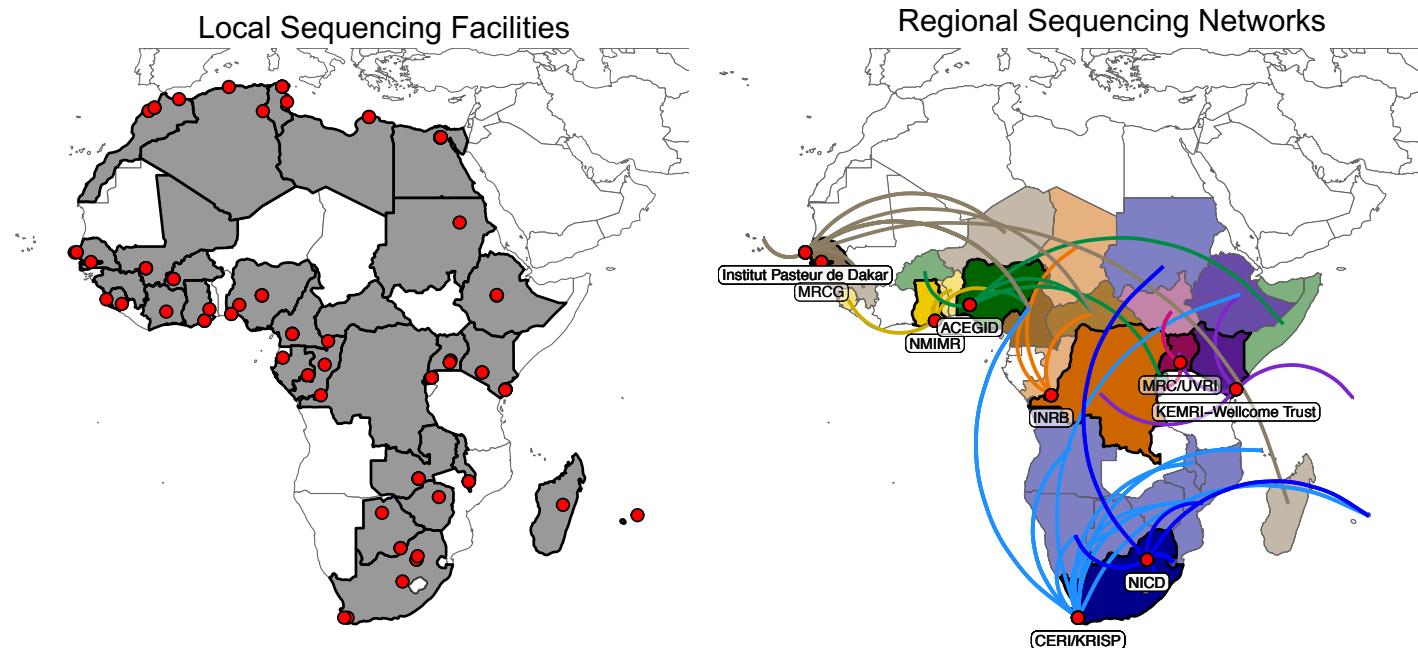
Where do these sequences go once we have them?

- [Nextclade](#): genome quality and curation
- [GISAID](#): Global Initiative on Sharing All Influenza Data
- [NCBI](#): National Center for Biotechnology Information

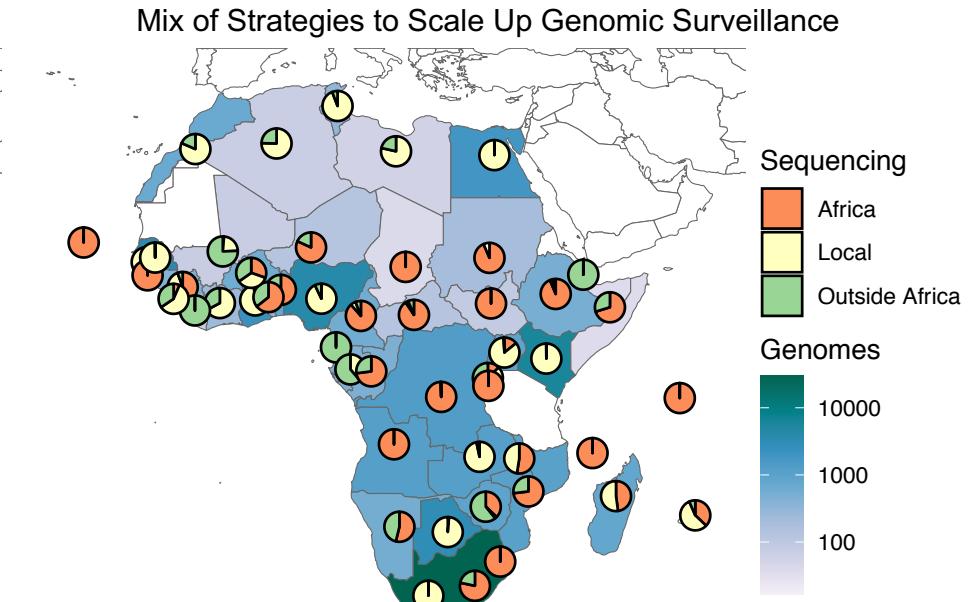
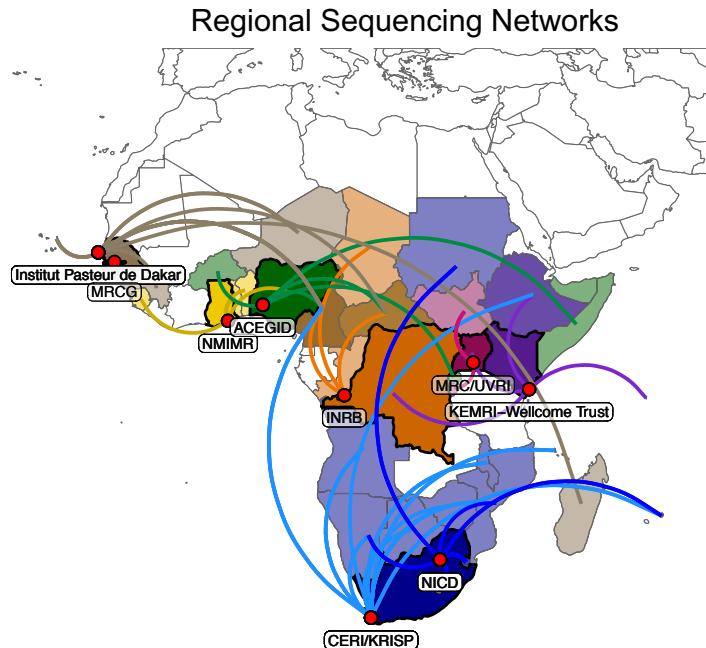
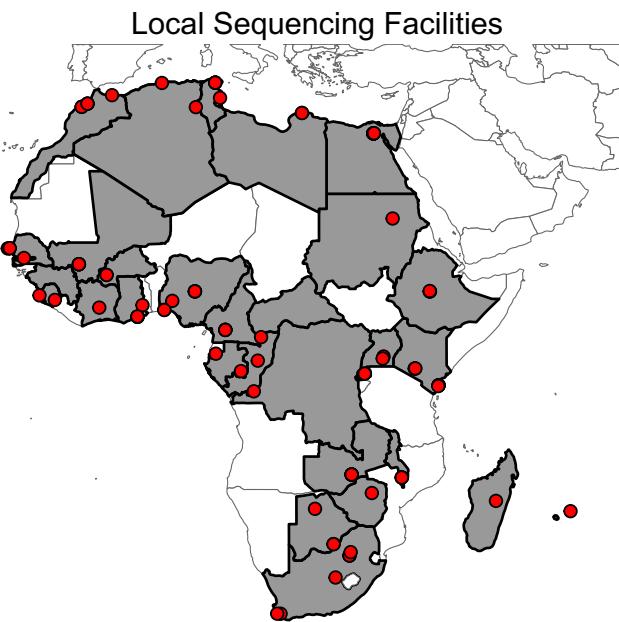
COVID-19 has hastened the
expansion of pathogen genomic
sequencing in Africa



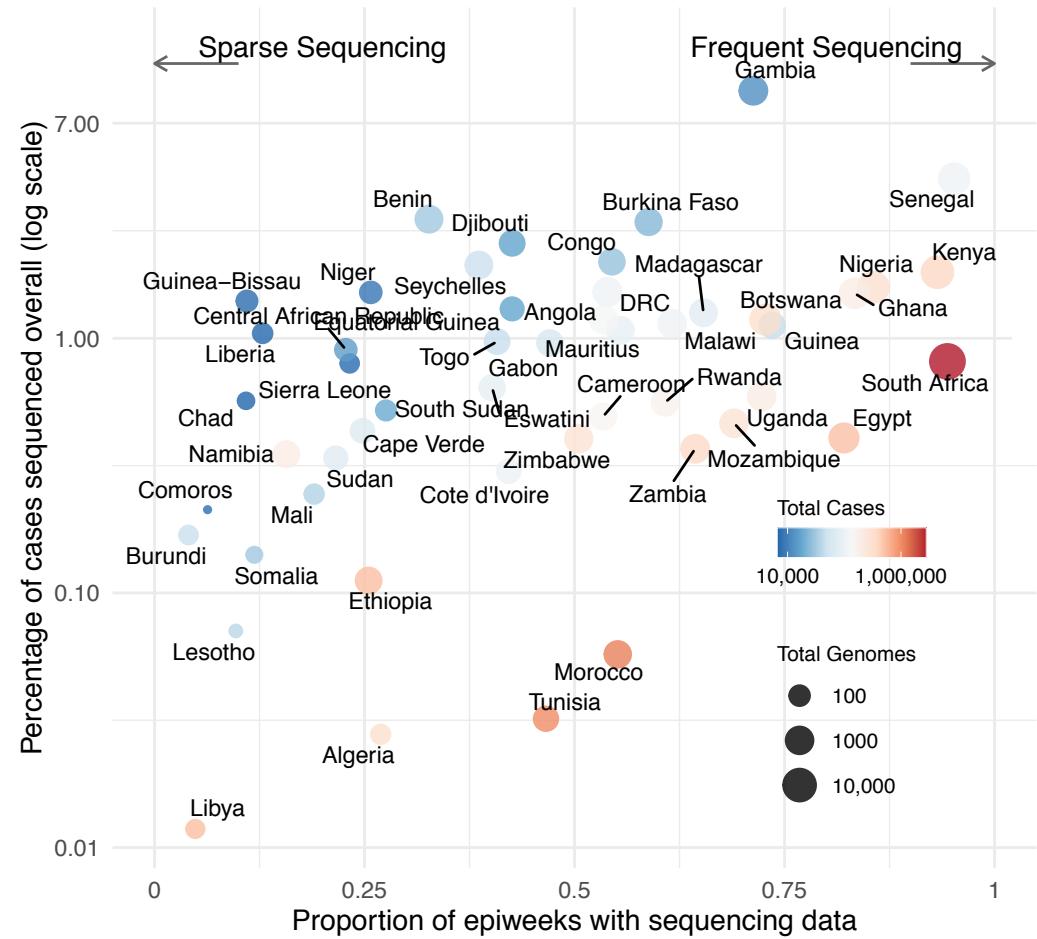
COVID-19 has hastened the expansion of pathogen genomic sequencing in Africa



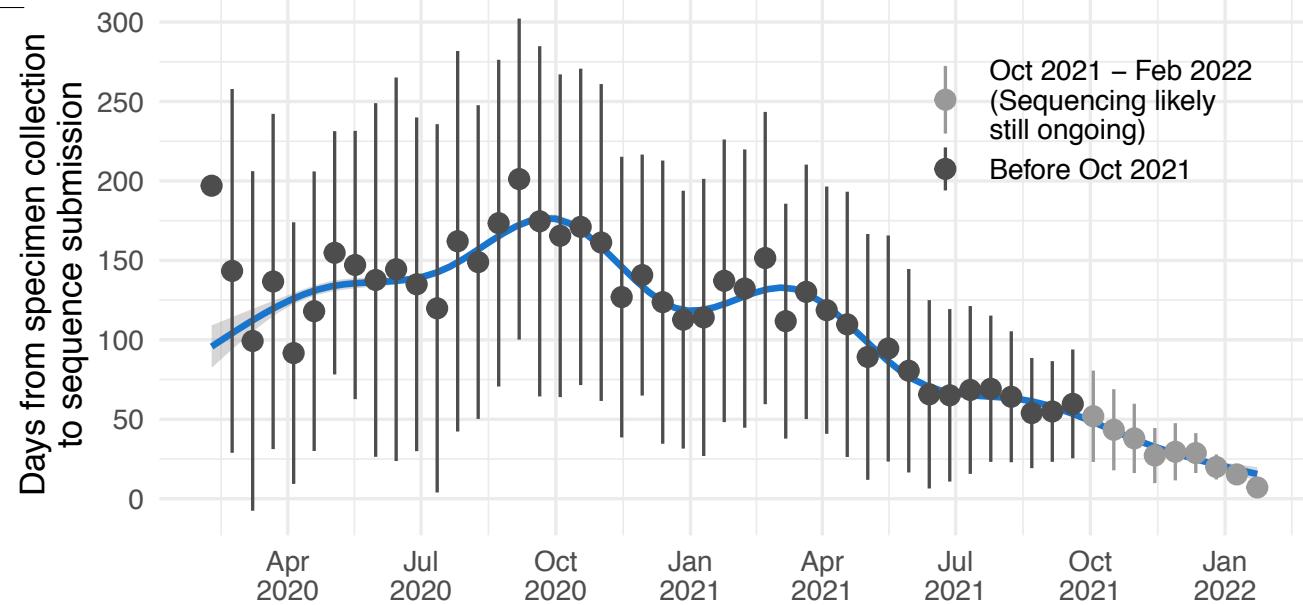
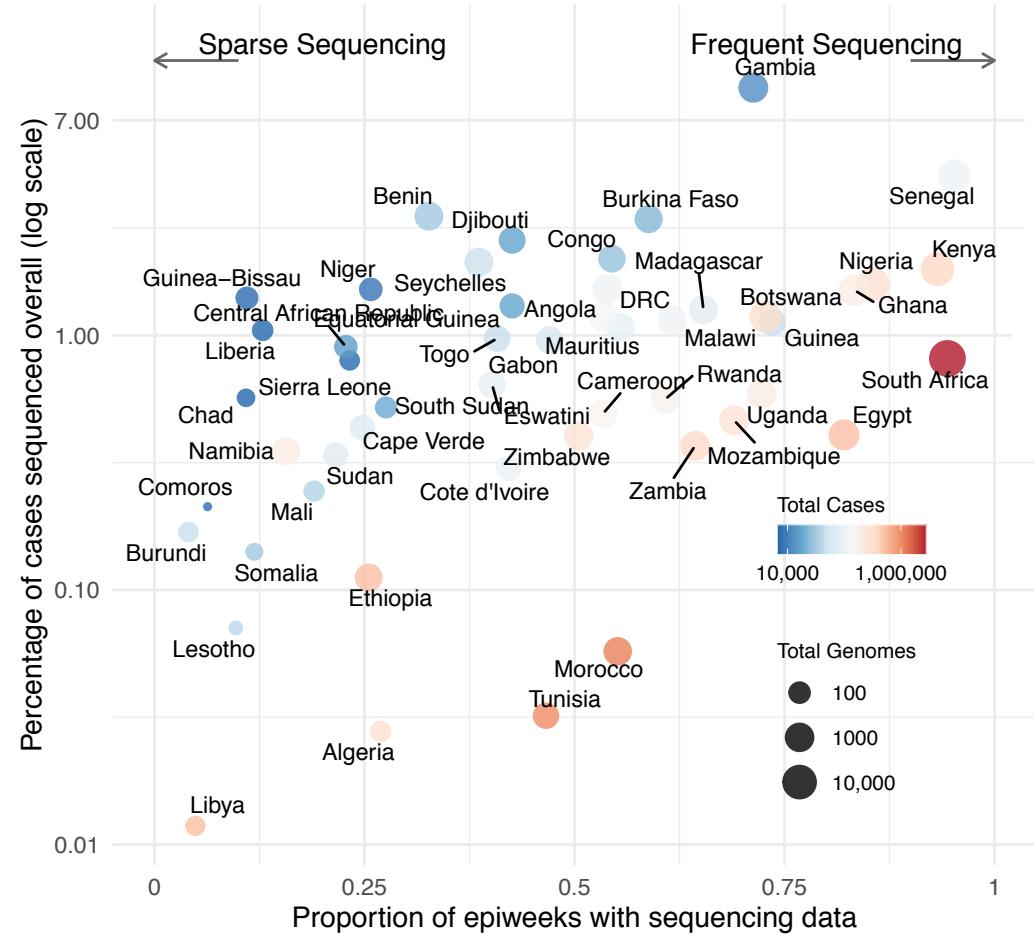
COVID-19 has hastened the expansion of pathogen genomic sequencing in Africa



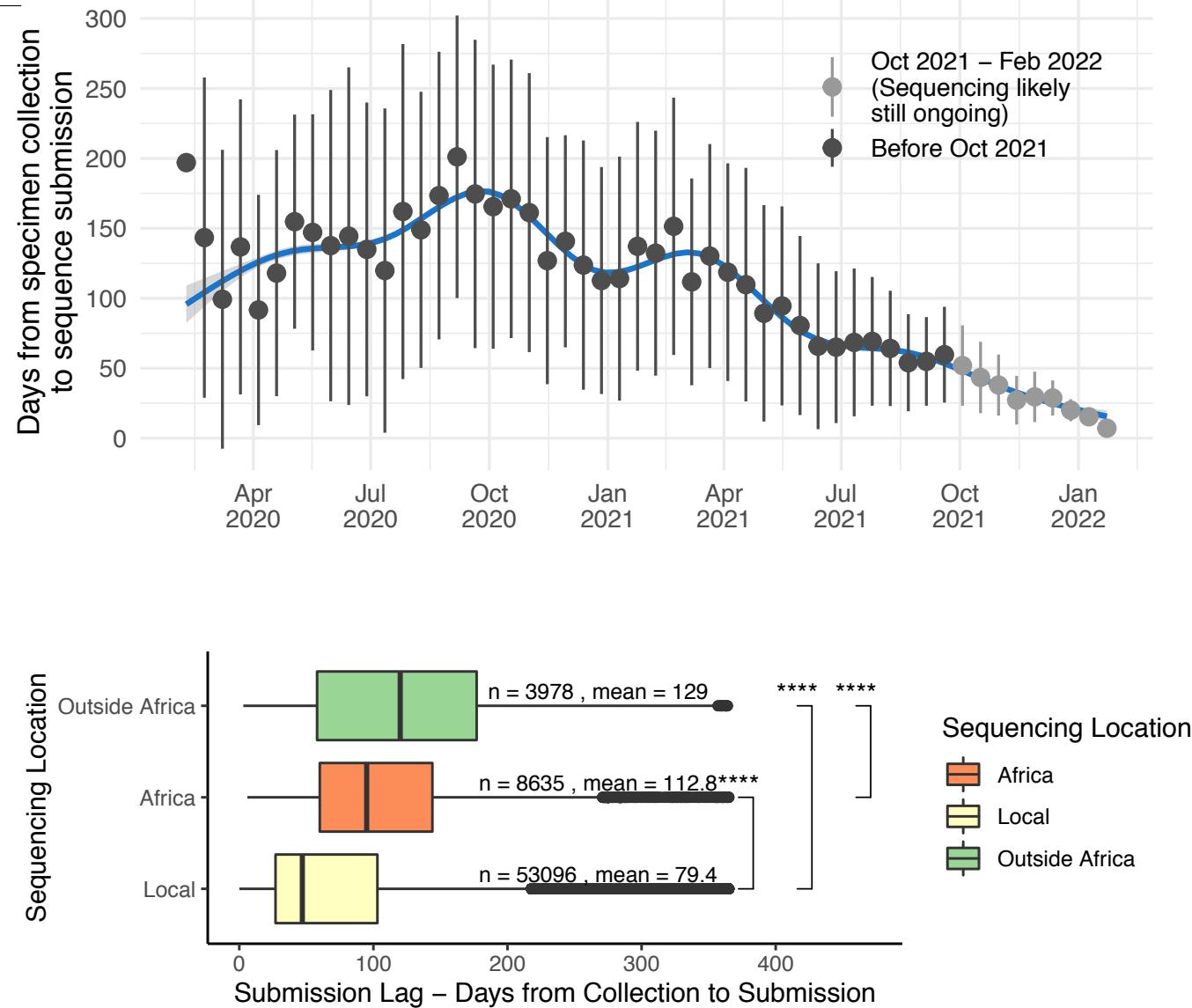
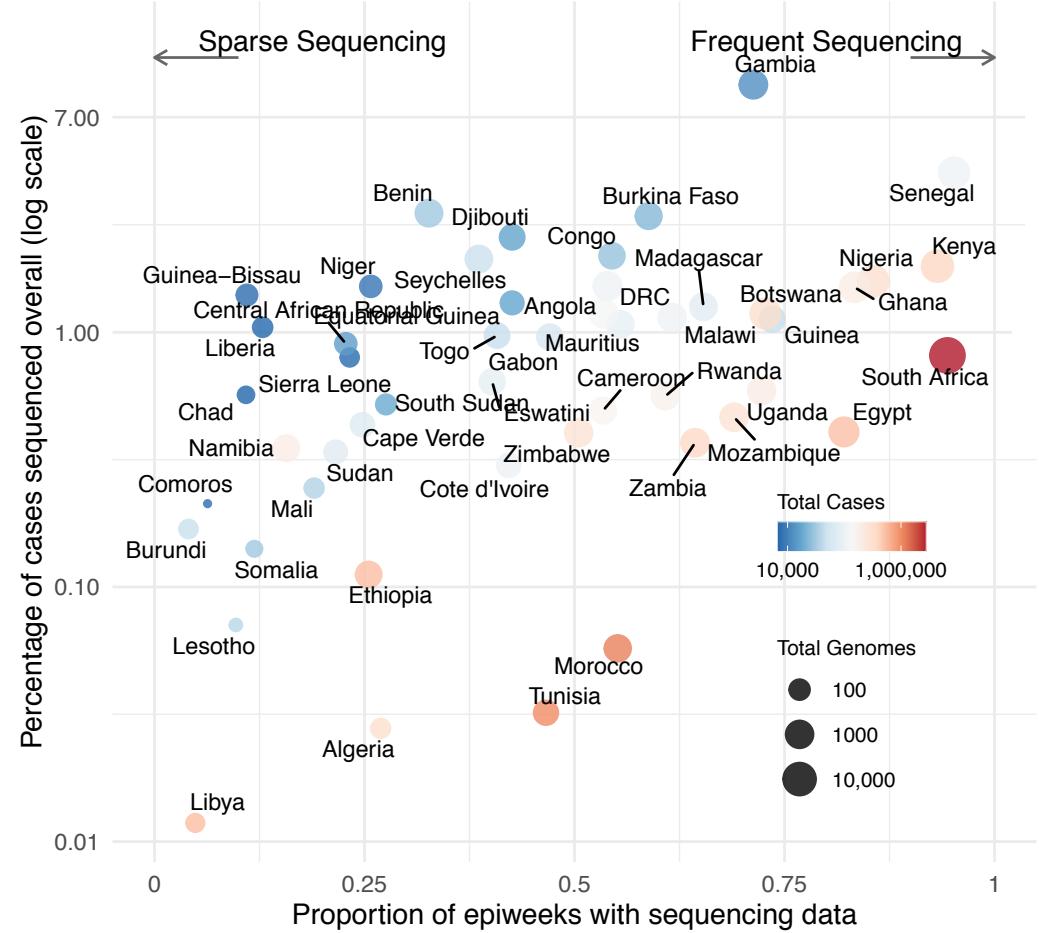
COVID-19 has hastened the expansion of pathogen genomic sequencing in Africa



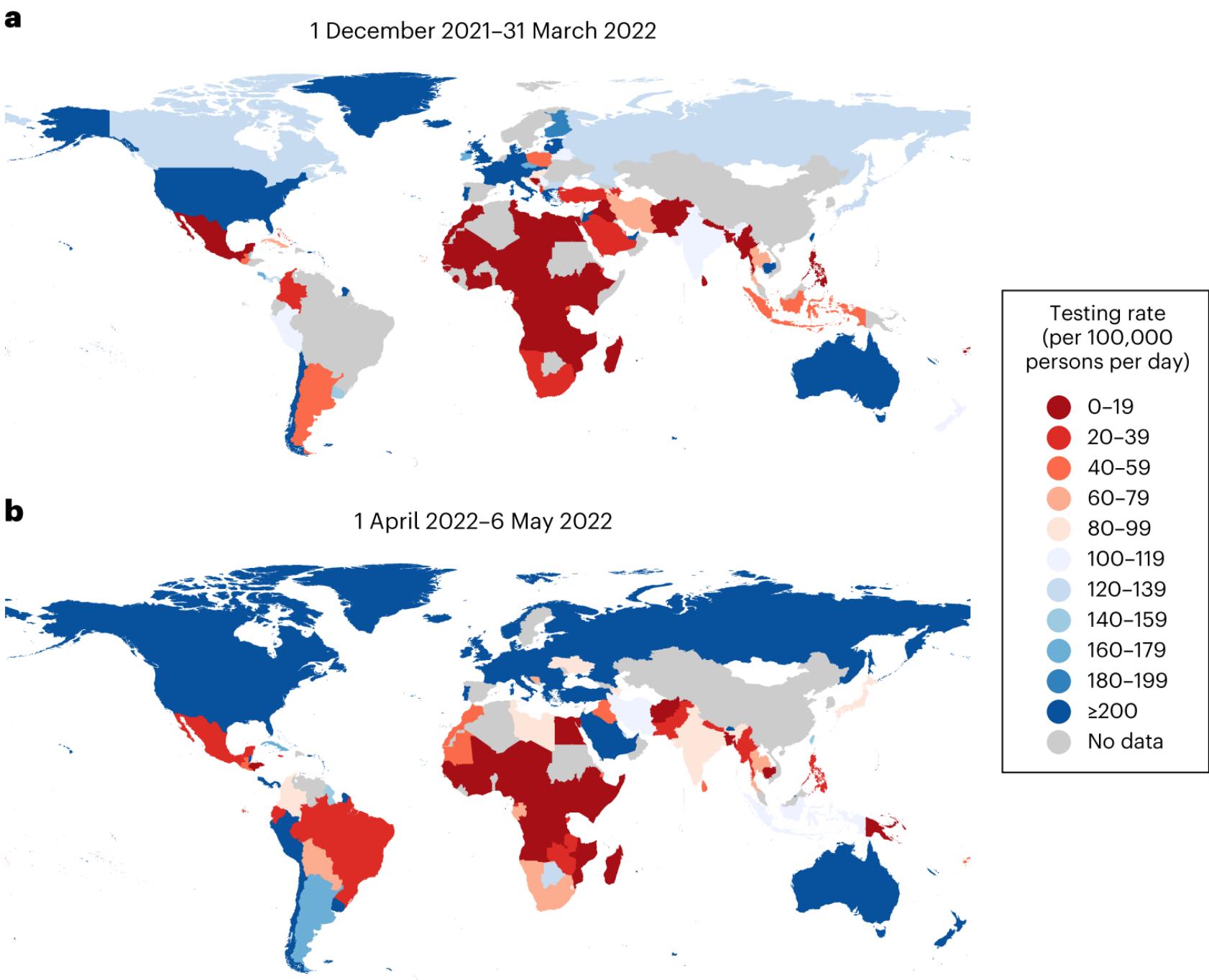
COVID-19 has hastened the expansion of pathogen genomic sequencing in Africa



COVID-19 has hastened the expansion of pathogen genomic sequencing in Africa



But there are still
many, many gains to
be made!



Proportion of positive specimens to sample for sequencing per day (%)

