

The linearized virtualization of the pan-genomic graph paths (colored bars) vs. the sequence or topology of the graph generated using Optimized Dynamic Genome Graph Implementation visualization (ODGI viz) from the pan-genomic graph for the CLCA2 gene.

## The challenge

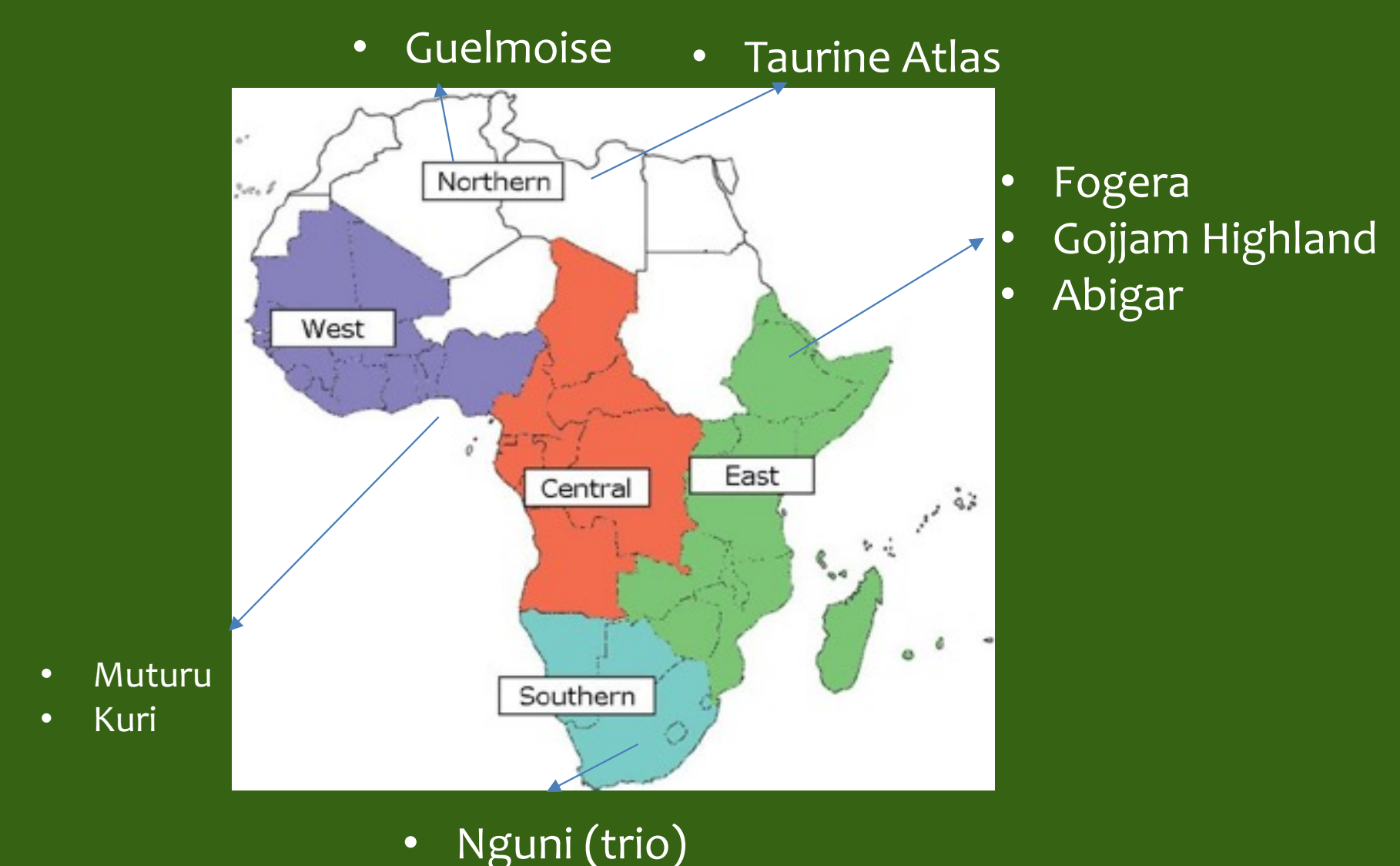
- The existing cattle reference genomes are **European or Asian** breed-specific and not representative of **African** cattle breeds.
- Genetic control of many **African unique adaptive phenotypes** remain unknown.
- Our project aims to overcome the cattle reference genomes bias and shortcoming.
- The pan-genome approach may reveal hidden **African cattle specific genetic variations and control**

## Our innovative approach

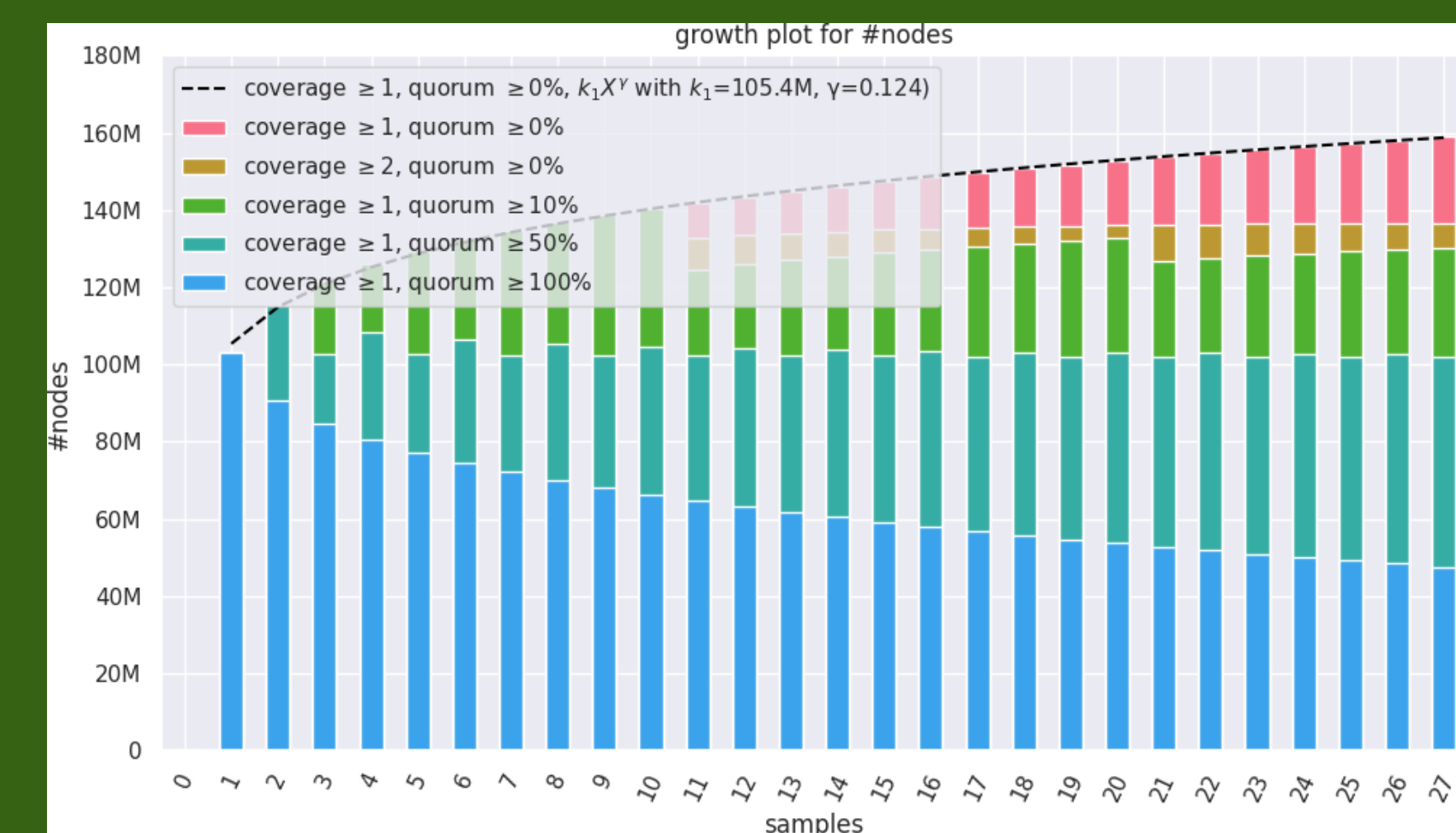
- De novo** genome sequencing and annotation of representative African cattle (breed types, geography).
- PacBio and Oxford Nanopore long-read sequencing data assembled and aligned = **pan-genome**.
- Investigation of **large structural variants (SV)** linked to environmental adaptation (climate).



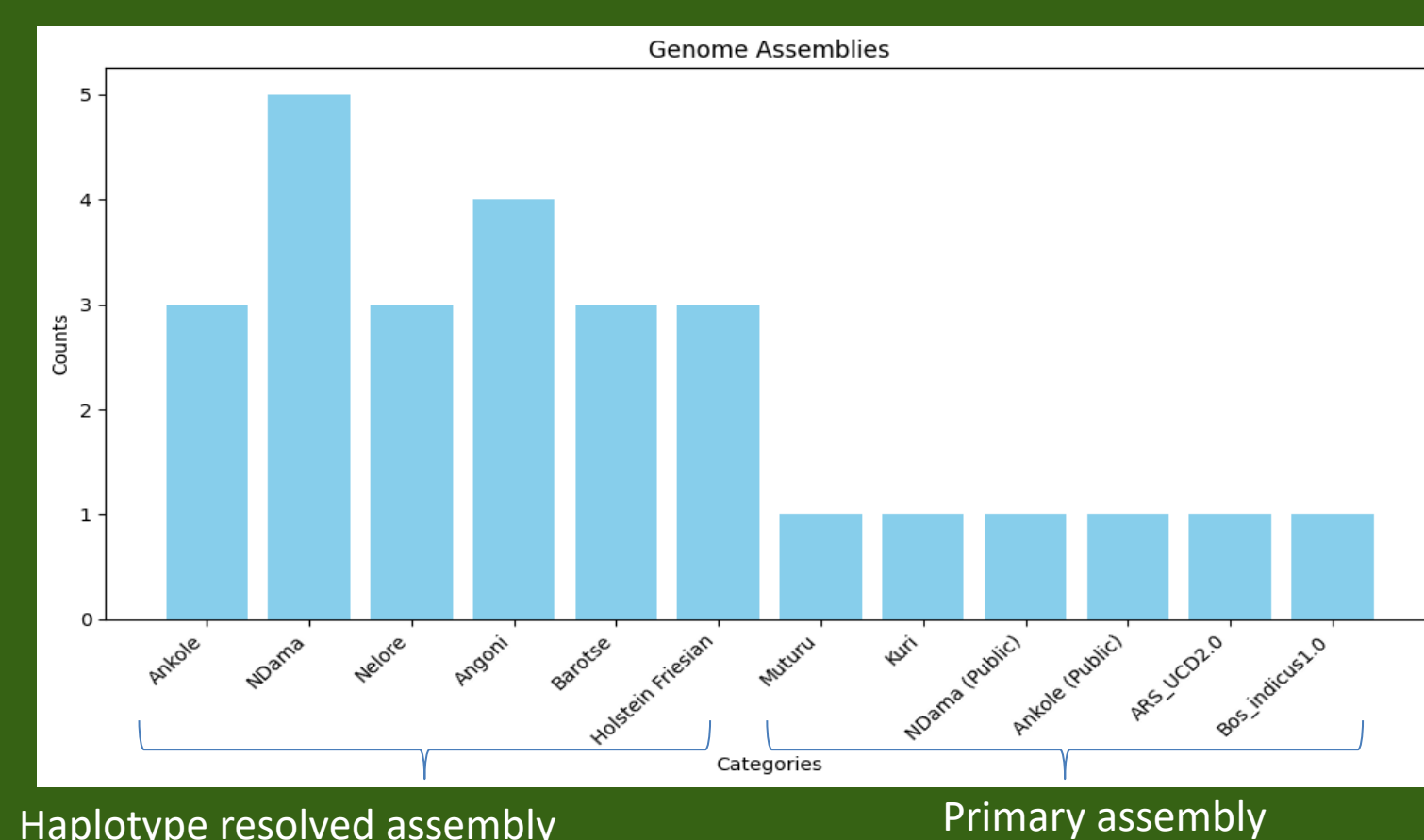
# Pan-genome a new approach for understanding the genomic diversity of African cattle



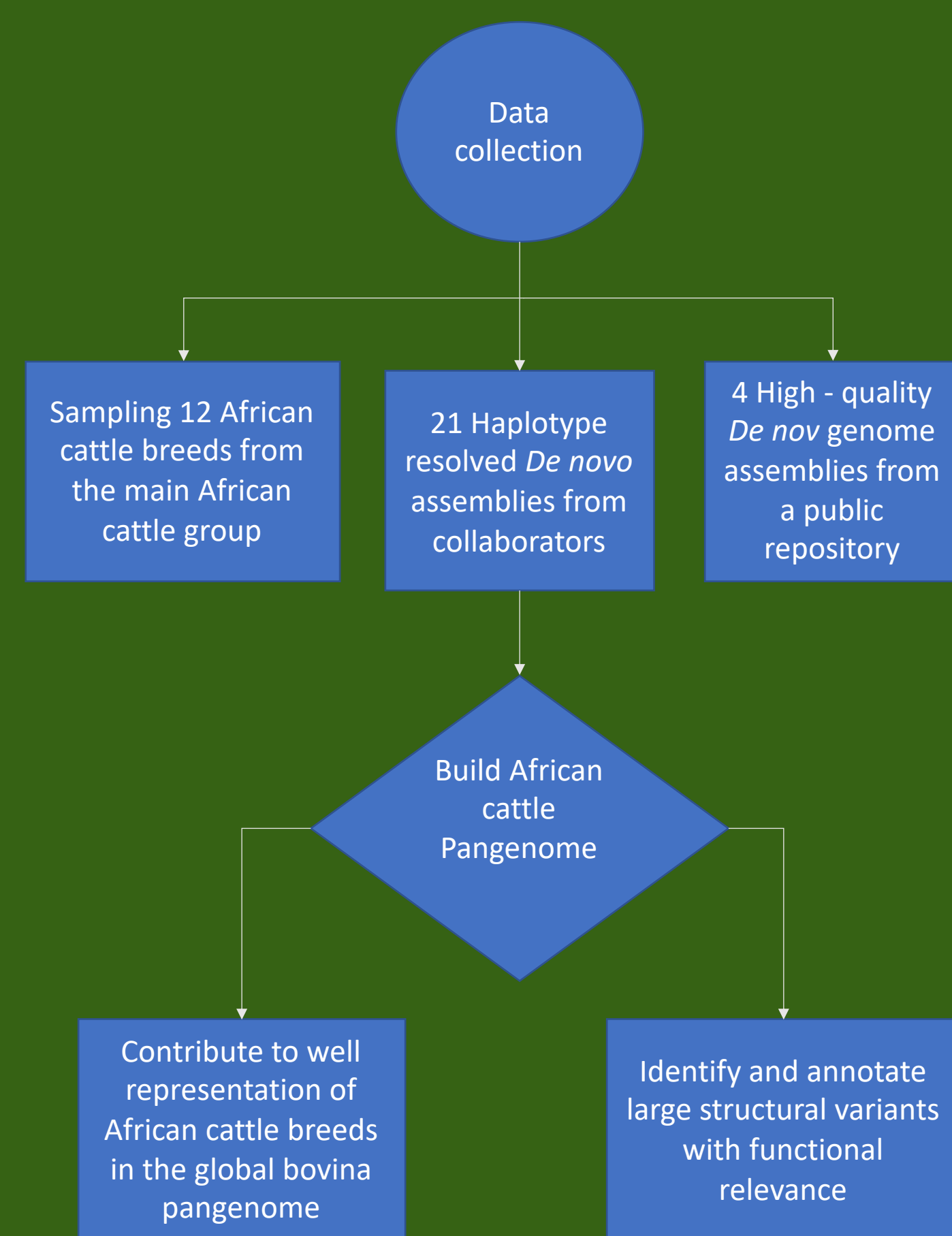
Samples representative of major African cattle breeds and geography collected for long-read sequencing



The pan-genome growth plot indicates the number of nodes on the graph in millions on the y-axis in relation to the number of incremental *De novo* genome samples on the x-axis. The color coding identifies the nodes based on the pangenome coverage and quorum.



*De novo* genome used to build the African Pan-genome. The x-axis breeds, Y-axis *De novo* genome count

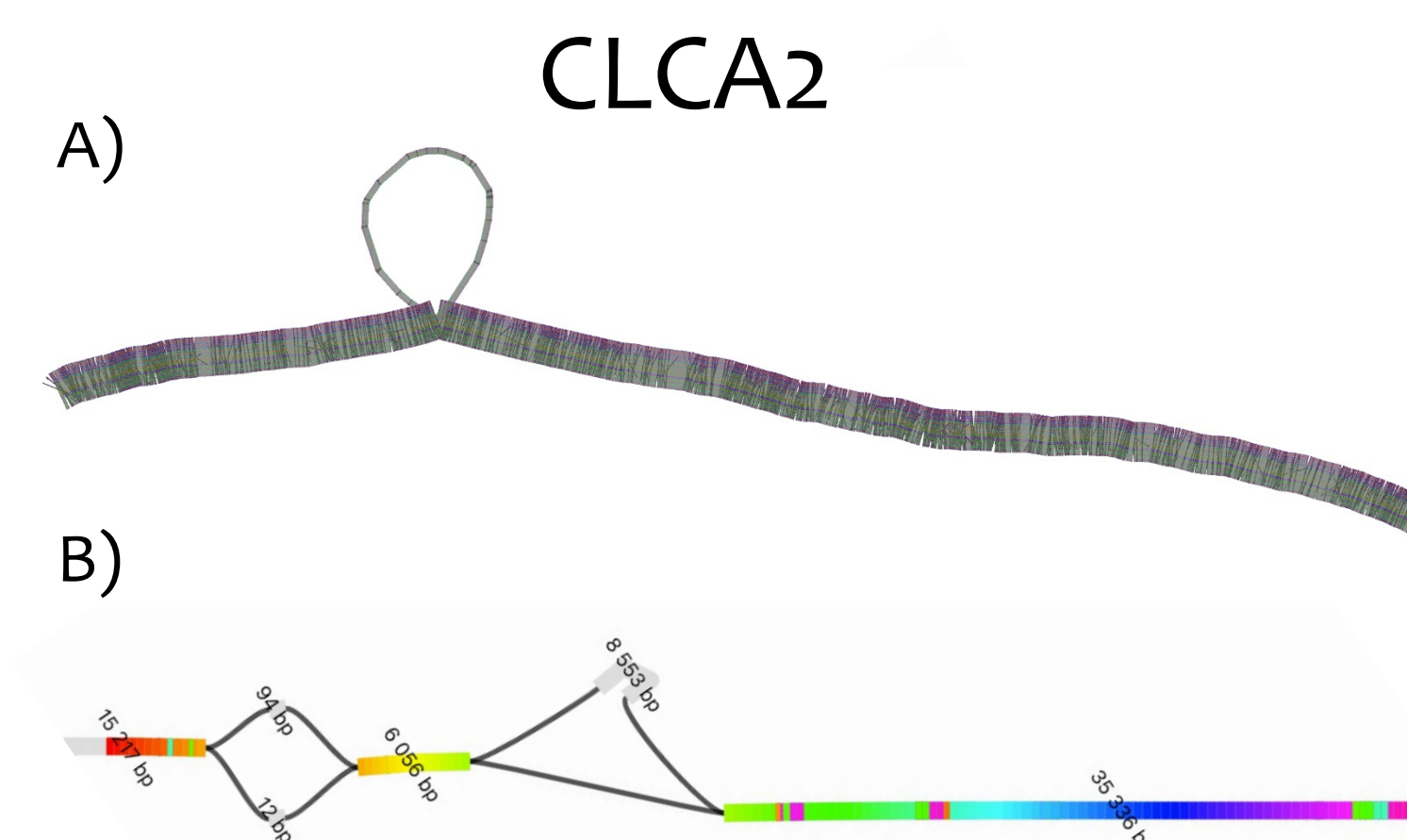


Nigatu Ayele  
Postdoctoral fellow  
n.ayele@cigar.org

## Results

### Example

- CLCA2 gene** is associated with **high-altitude adaptation** in cattle, having a role in renin secretion and response to hypoxia
- N'Dama-specific SV** at the intronic part of CLCA2



A) Visualization of the graph pan-genome for the CLCA2 gene and its alternative sequence(bubble) generated with ODGI draw. B) The bandage visualization for the N'Dama-specific non-referenced 8553 bp insertion on the CLCA2 gene

## Next steps

- De novo* genome assemblies from **all main African cattle lineages**.
- Genome-wide identification of **structural variants** and **functional annotation**
- Aurochs introgression** analysis in African cattle as a shortcut approach to identify genome regions linked to **environmental adaptation**

## Partners



The International Livestock Research Institute thanks all donors & organizations which globally support its work through their contributions to the CGIAR Trust Fund.  
cgiar.org/funders

This document is licensed for use under the Creative Commons Attribution 4.0 International Licence. July 2024

