Future of genetic study in selection of plants for commercializing

How genomics and marker-assisted selection are reshaping which plants we choose to commercialise

Dr Ansia van Coller
South African Medical Research Council. Genomics Platform
Ansia.vanColler@mrc.ac.za

A biodiversity superpower, still largely uncharacterised

South Africa holds extraordinary botanical wealth — yet little of it is backed by genomic data that could accelerate responsible commercialisation.

3rd

Most biologically diverse country on Earth, after Brazil and Indonesia.

~24,000

Indigenous plant species — including the entire Cape Floristic Region.

<20

Of the ~800 plant genomes sequenced globally are endemic to Africa.

African BioGenome Project (AfricaBP): building Africa's genomic foundation

A continent-wide effort to generate high-quality reference genomes for Africa's indigenous species — sequenced and stored on the continent, in service of conservation and improvement of African biodiversity.

Endemic

— plants, animals, fungi
Through SAMRC: focused on plants with ecological, economic or cultural importance

In Africa

sequencing and data storage kept on the continent, building local capacity.

Partners

linked to global efforts including the Earth BioGenome Project and the 10,000 Plant Genomes Project. 1KSA

The mission: expand genomic resources so African biological wealth can benefit Africans first.

KSA Decoding South Africa's Biodiversity

saMRC advancinglife

Endemic African genomes, assembled on African soil

Helichrysum odoratissimum

Imphepho — first chromosome-level genome assembled locally in Africa. A prized aromatic, antimicrobial medicinal plant.

Gambeya albida

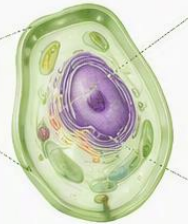
White Star Apple — chromosome-scale genome (2026) for a species central to West & Central African livelihoods.

What is Genetics?

Plants contain DNA, the genetic blueprint of life.
DNA is organised into **genes**, which contain instructions for building and regulating the plant.

PLANT

All plants contain cells.



CELL

Each cell contains DNA.

DNA

DNA is the genetic blueprint of the plant.

GENES

DNA is organised into genes.
Each gene contains instructions.

Genes are sections of DNA that contain instructions.

In short: DNA ⟶ Genes ⟶ Instructions ⟶ Plant traits and functions

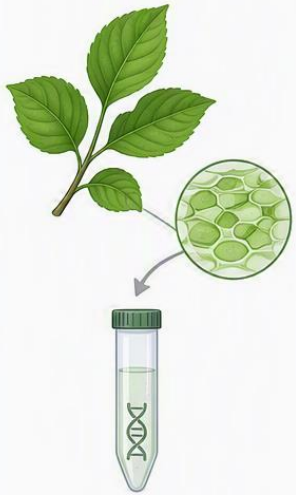
GENES CONTAIN INSTRUCTIONS THAT INFLUENCE PLANT TRAITS AND FUNCTIONS

- **Growth rate** — How fast the plant grows
- **Plant height** — How tall the plant grows
- **Flowering time** — When the plant flowers
- **Oil yield** — How much oil the plant produces
- **Essential oil composition** — Which compounds and how much
- **Disease and pest resistance** — Ability to resist threats
- **Drought tolerance** — Ability to withstand water stress

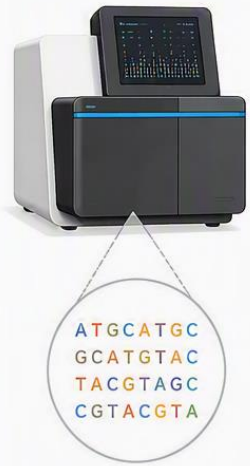
Genome Assembly: From Plant to Complete Genome

Putting millions of short DNA reads together to reconstruct the full genome.

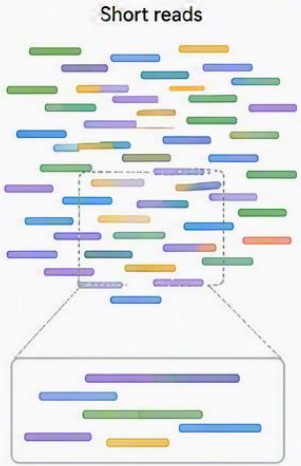
1 **Sample Collection & DNA Extraction** → 2 **Sequencing (Short Reads)** → 3 **Millions of Short, Overlapping Reads** → 4 **Assembly Using Software Tools** → 5 **Assembled Genome**



DNA is extracted from plant tissue (e.g., leaf).



The DNA is broken into small pieces and sequenced by next-generation sequencers.

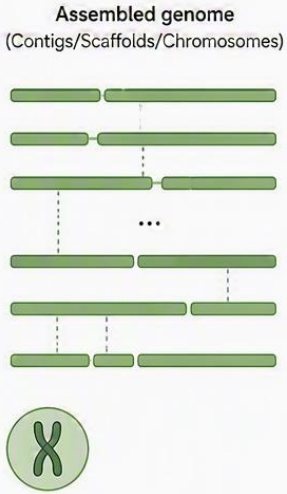


Millions of short reads are generated. Each read overlaps with many others.

Assembly algorithms

- ✓ Find overlaps
- ✓ Build contigs
- ✓ Resolve repeats
- ✓ Order and orient sequences

Software assembles the overlapping reads into longer sequences (contigs) and orders them.



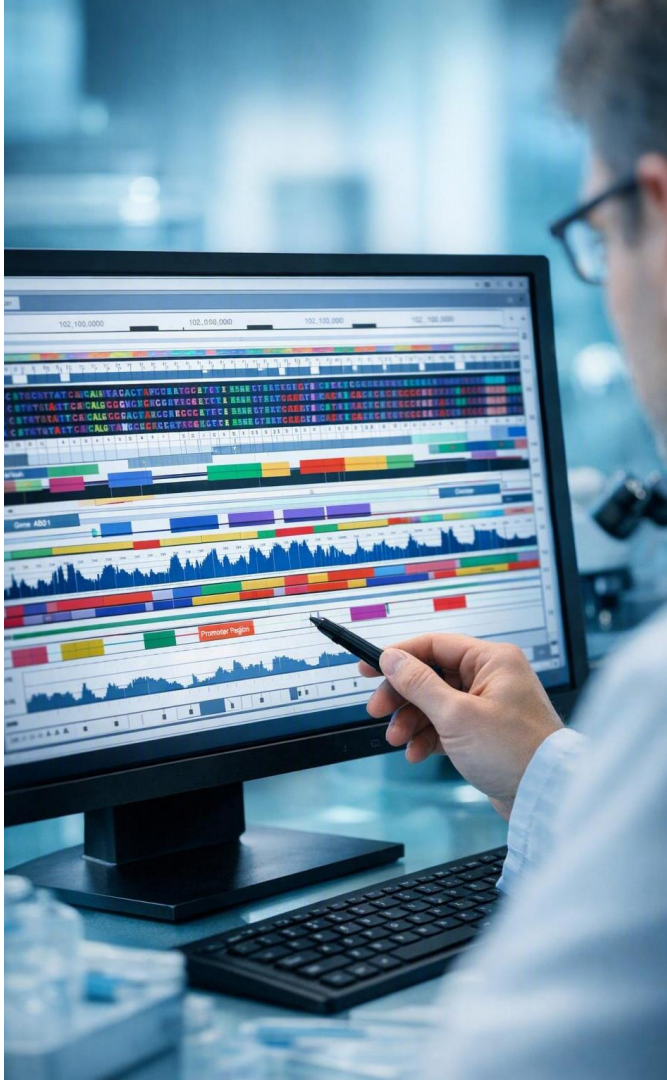
A complete or near-complete genome is reconstructed for downstream analysis and applications.

Genome assembly turns millions of short, overlapping reads into a complete picture of the plant's genome, enabling insights into genes, traits, and valuable compounds.

Reading the genome for valuable traits

Once assembled, the genome has to be interpreted — locating the genes and linking them to traits of interest.

- **De novo annotation.** Software and/or AI tools to locate genes and biosynthetic pathways across a newly sequenced genome from scratch.
- **Trait discovery.** Link those genes to what we value — oil yield, chemotype, potency and resilience.
- **Compare to known genes.** Match sequences against well-studied species to flag genes of interest fast.



From wild harvest to precision selection

Genomic tools let us select the right plants — and the right individuals — for quality, yield and consistency, instead of relying on wild collection and trial-and-error.

1

Reference genome

A high-quality genome maps every gene and biosynthetic pathway.

2

Trait discovery

Link DNA to oil yield, chemotype, potency and resilience.

3

Molecular markers

Screen seedlings by DNA — years before a plant matures.

4

Selected cultivar

Propagate elite, standardised, market-ready material.

Genomics approach: many commercial destinations

The same reference genomes and markers unlock value across the full spectrum of plant-derived products.

Medicinal & essential oils

Higher-potency actives, defined chemotypes.

Cosmetics

Botanical actives with consistent quality.

Crop plants

Faster breeding for yield and resilience.

Perfume & fragrance

Reliable aroma profiles at scale.

Pest control

Bio-based repellents and biopesticides.

Sustainable supply

Traceable, standardised value chains.

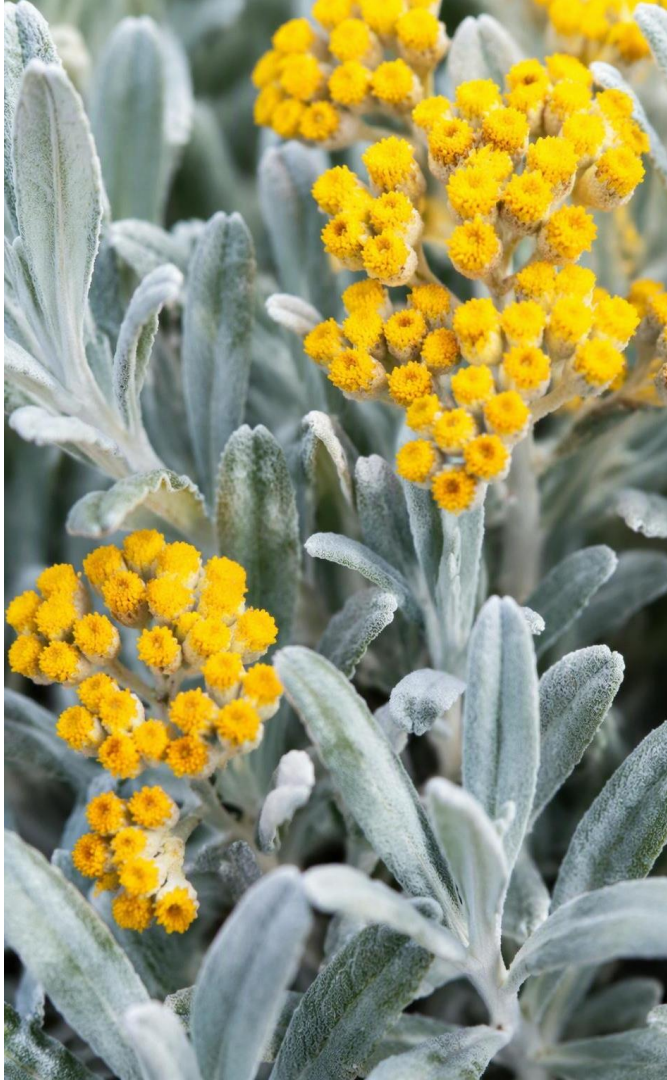
MEDICINAL PLANTS & ESSENTIAL OILS

Selecting for potency, not just presence

Value lives in secondary metabolites — the terpenes, phenolics and volatiles behind antimicrobial, anti-inflammatory and aromatic activity.

- Marker-assisted breeding raises target-metabolite content and defines stable chemotypes.
- Genomes reveal biosynthetic pathways — guiding which populations to cultivate or protect.
- Consistency turns a traditional remedy into a certifiable, exportable product.

Case in point: imphepho (Helichrysum odoratissimum) — recently sequenced and assembled to chromosome scale on African soil.



COSMETICS · PERFUME · FRAGRANCE

Reliable botanical identity for premium markets

Cosmetic and fragrance buyers pay for consistency and provenance. Genomics delivers both — authenticating species and locking in the aroma and actives that define a brand.

Cosmetics & natural actives

- DNA authentication guards against adulteration and mislabelling
- Select lines rich in antioxidants, oils and skin-active compounds
- Underpins clean-label and certified-natural claims

Perfume & fragrance

- Fix aroma chemotypes so scent profiles stay identical batch to batch
- Breed for higher essential-oil yield per hectare
- Protect signature indigenous notes as a regional asset

CROP PLANTS

Marker-assisted selection shortens the breeding cycle

For food and industrial crops, selecting on DNA rather than waiting for the mature phenotype compresses years of breeding into a single screening step.

Conventional breeding

- Select on visible traits at maturity
- Many seasons per generation
- Recessive & complex traits hard to track
- Slow response to market and climate shifts

Marker-assisted selection

- Screen seedlings directly by DNA
- Fewer cycles — years saved
- Stack yield, quality & disease-resistance genes
- Rapid, targeted, data-driven improvement

PEST CONTROL

Plants that defend themselves — and our crops

Essential oils and inherent plant traits offer bio-based alternatives to synthetic pesticides — safer residues, lower resistance risk, and strong appeal for organic markets.

- **Essential-oil biopesticides.** Terpene-rich oils repel and disrupt insects and pathogens.
- **Breed the trait in.** Markers let us select plants with natural pest resistance.
- **Cleaner, compliant produce.** Fewer synthetic residues meet tightening export standards.

SUSTAINABLE SUPPLY & QUALITY

Genomics anchors traceable, standardised value chains

Global buyers demand proof of quality, safety and origin. DNA identity and defined chemotypes provide an objective backbone from field to export.

Verified planting stock

DNA-authenticated, elite cultivars

→

Cultivation

Consistent chemotype in the field

→

Extraction & QC

Standardised, testable specifications

→

Certified export

Meets market standards

Standardisation is the unlock. Initiatives like national quality programmes and SAEOPA show that traceable genomic identity plus quality standard systems is what turns biodiversity into durable export revenue.

IN SUMMARY

Four points to carry forward — from South Africa's untapped biological wealth to the shared, sustainable value that genomic resources can unlock.

1

Untapped biodiversity

The 3rd most biodiverse nation, yet limited genomic resources

2

The decisive tool

Reference genomes and markers move us from wild harvest to precision selection.

3

Many destinations

One genomic foundation serves medicine, cosmetics, crops, pest control and supply.

4

African value

Local genomic resources turn biodiversity into shared, sustainable benefit.

AN OPEN INVITATION

Species of ecological, economic and cultural value?

If you're interested in joining the initiative — particularly around a species with ecological, economic or cultural importance — you're welcome to reach out to us.

If you'd like to join

- A species with ecological, economic or cultural importance
- Reach out if you're interested in joining the initiative
- Bioprospecting & access/sampling permits may be required

How benefits are shared

- Fair, equitable benefit-sharing built in from the start
- Aligned with the Nagoya Protocol on access & benefit-sharing
- Supports Kunming-Montreal Framework (KMGBF) Target 13
- African resources generating value for African communities

THANK YOU

African genomes, powering African economies

Genetic study is becoming a decisive tool in choosing which plants to commercialise — for medicine, cosmetics, crops, pest control and beyond.
By building indigenous genomic resources, we can turn biodiversity into shared, sustainable value.

Ansia.vanColler@mrc.ac.za