Genetic Diversity and Population Structure of Shea Trees for Sustainable Conservation and Breeding

PAG 32, Leveraging Advances in Multi-Omics Technologies for Resilient Agri-Food Systems in Africa
Workshop

Affi Jean Paul Attikora

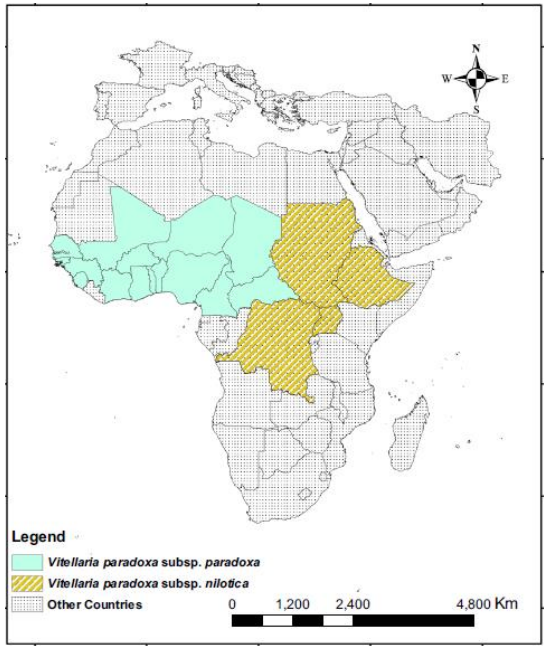
Plant Genetics lab, Uliege-Gembloux Agro Bio-Tech, Belgium

Plan of this presentation

- Context of the study
- Objectives
- Methods and Results
- Conclusions and perspectives

Context of the study

- ✓ Shea tree (*Vitellaria paradoxa):* African native plant species (2n=24)
- ✓ Two sub-species*:* subs*. nilotica and* subs. *paradoxa*
- ✓ Second largest oil producing plant in Africa
- ✓ Shea butter in cosmetics, pharmaceuticals and food industries
- ✓ Global shea butter market in 2024: $2.4 billions
- ✓ Contribution to women’s autonomy in its natural range
- ✓ Carbon sequestration and substantial carbon reserve



Context of the study

- ✓ Shea tree is facing many threats: human activities and climate change
- ✓ Overuse of shea tree as household material
- ✓ Lack of regeneration
- ✓ Consequences:
- ✓ Reduction of shea tree densities
- ✓ Loss of interesting genotypes
- ✓ Shea butter supply and sustainability problems

Context of the study

✓ Actions have been taken:

✓ Shea tree planting

✓ Natural assisted regeneration

✓ *In situ* collection establishment via superior shea trees identification

✓ Restoration, conservation and breeding purposes

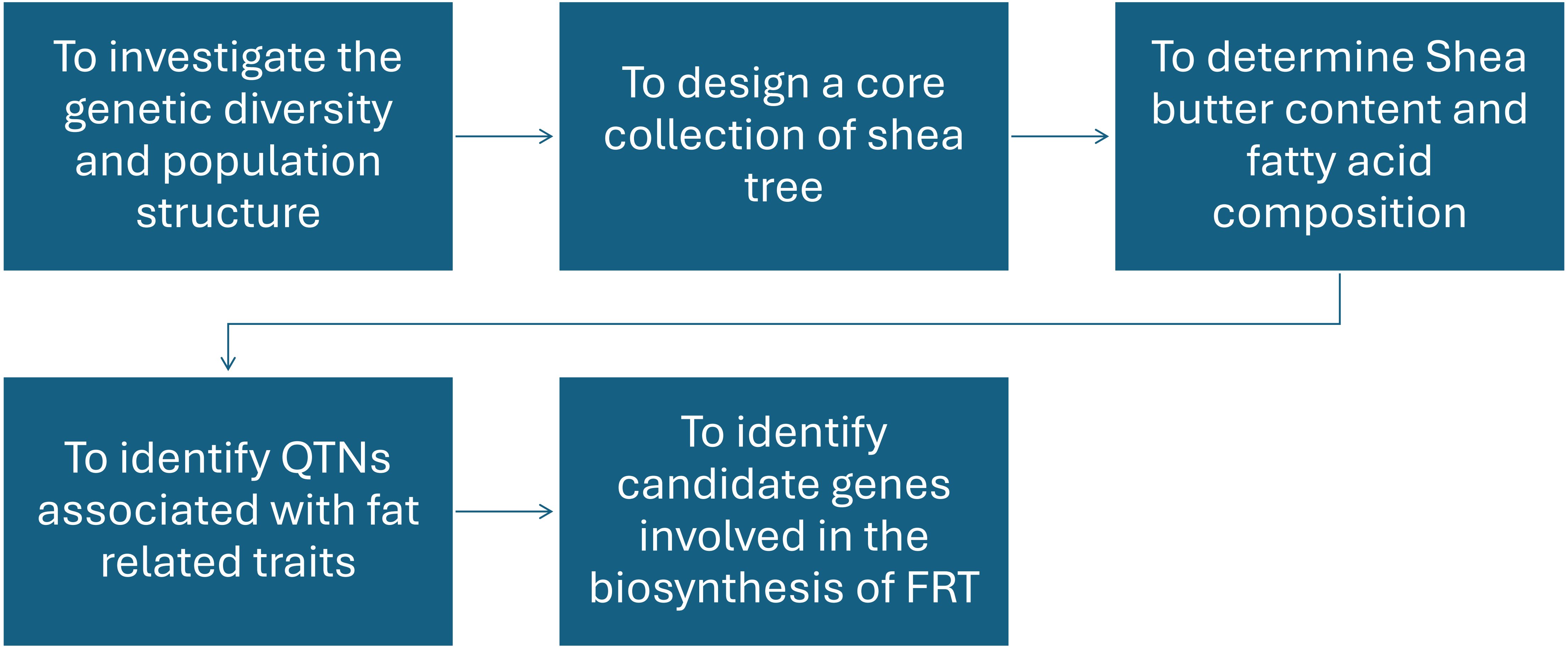
✓ But…

✓ Difficult and expensive to manage

✓ Shea trees are still threatened

Objectives

To provide solutions for the Long-term conservation and Sustainable use of Shea tree resources



To investigate the genetic diversity and population structure

To design a core collection of shea tree

To determine Shea butter content and fatty acid composition

To identify QTNs associated with fat related traits

To identify candidate genes involved in the biosynthesis of FRT

Genetic Diversity and Population Structure

- **Methodology**

✓ Sample: leaves of 333 shea trees were collected

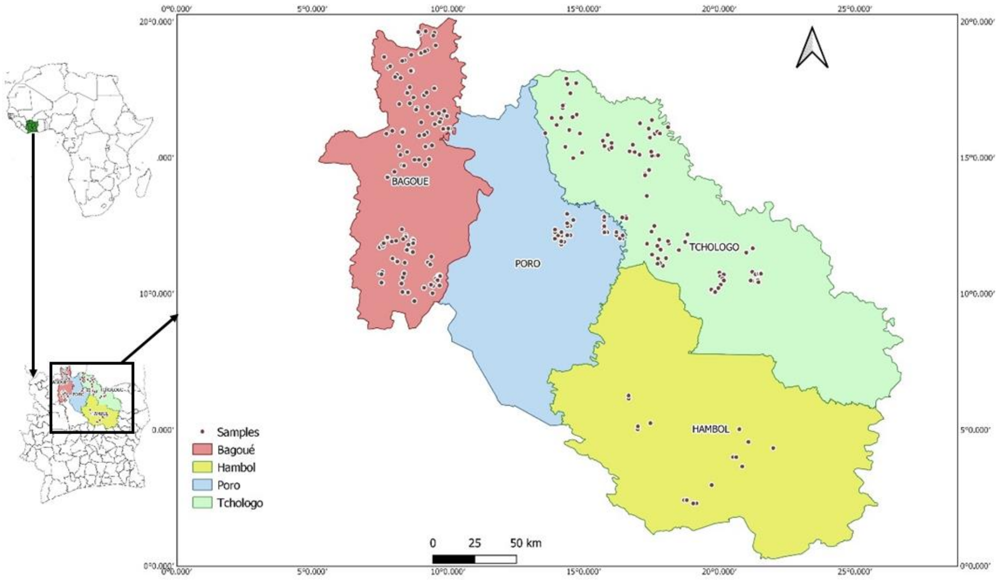
✓ DNA extraction: NucleoMag Plant Kit

✓ Sequencing: Diversity Array Technology Sequencing (DArTseq)

✓ SNP markers, aligned to the reference Genome to locate their chromosome positions

✓ 42736 SNP markers

✓ 7559 high-quality SNP markers after filtering



Genetic Diversity and Population Structure

- Genetic Diversity

✓ Ho: 0.17

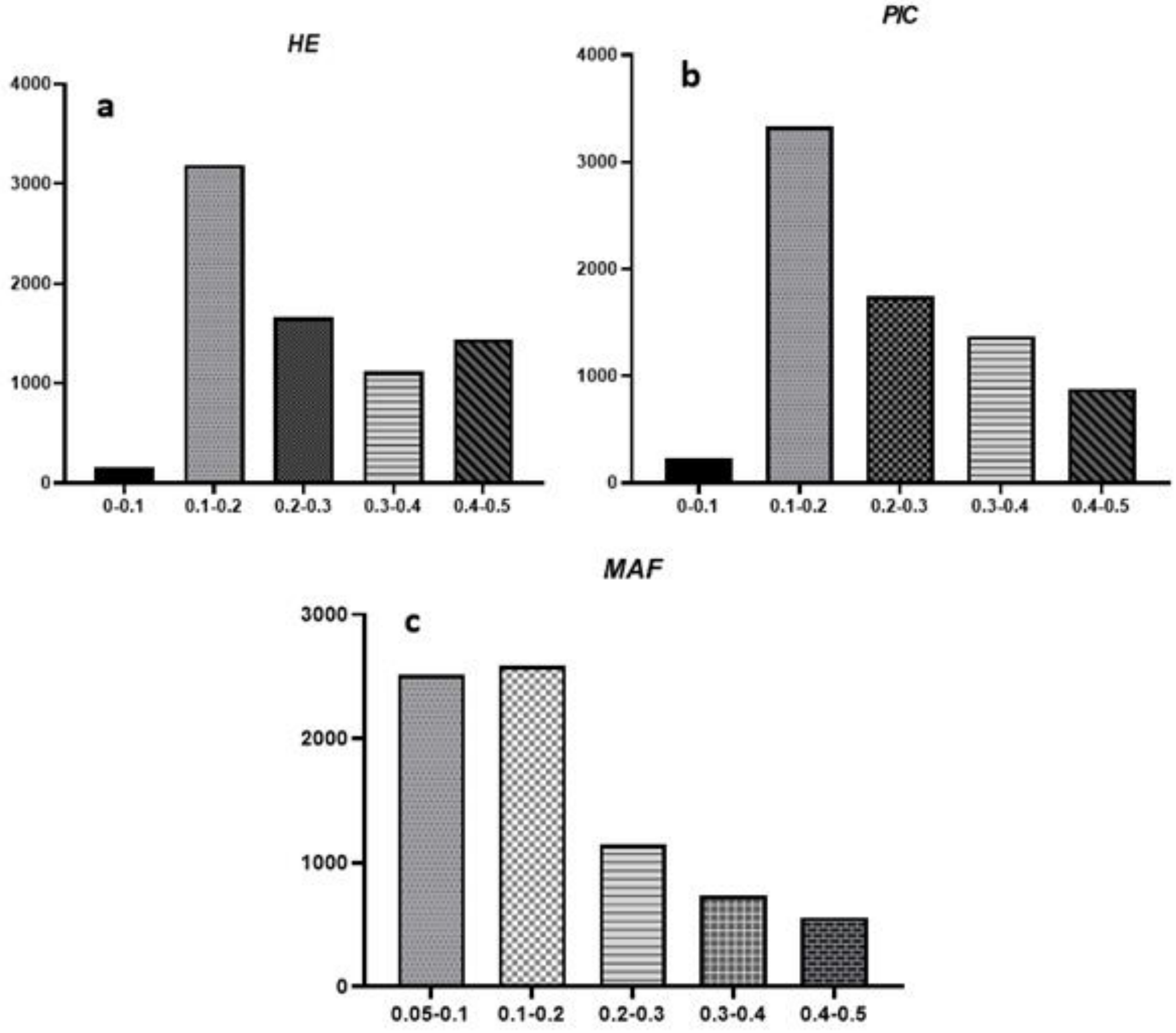
✓ He: 0.26

✓ PIC: 0.24

✓ MAF range: 0.05-0.5 (5,104 SNPs (67.5%) > 0.1)

➢ Moderate genetic diversity

➢ Adaptative Opportunity within the collected area

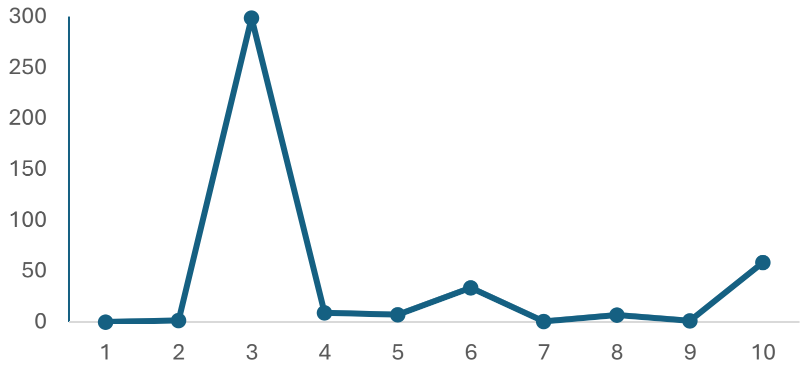


Genetic Diversity and Population Structure

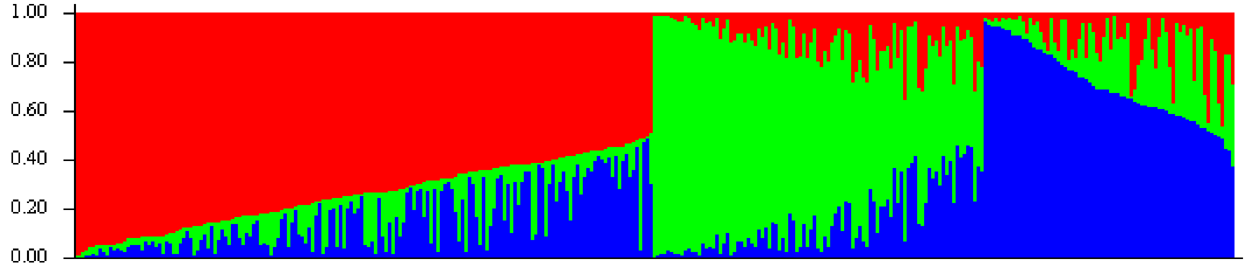
- **Population Structure and Genetic Differentiation**

❖ Structure v.2.3.4

✓ Three genetic groups: peak of Delta K value


DeltaK = mean(|L"(K)|) / sd(L(K))

✓ Structure bar plot: high admixtures proportion

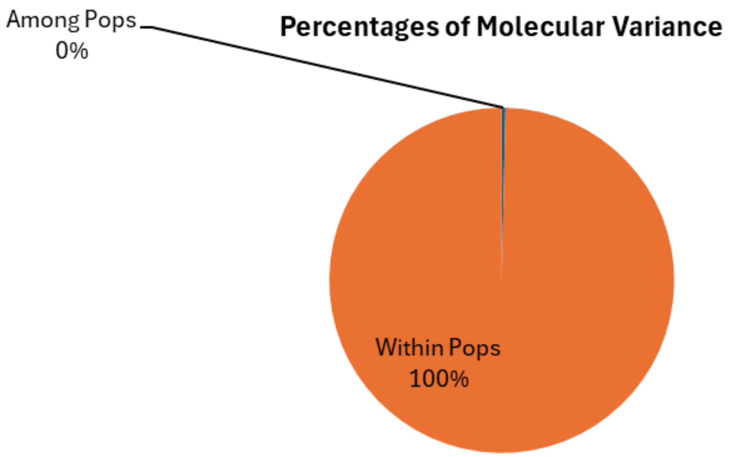


❖ AMOVA (GenAlex V.6.503)

✓ 100% of variance within populations

✓ Fixation index (Fst): 0.004

✓ Nm (Haploid): 59.02

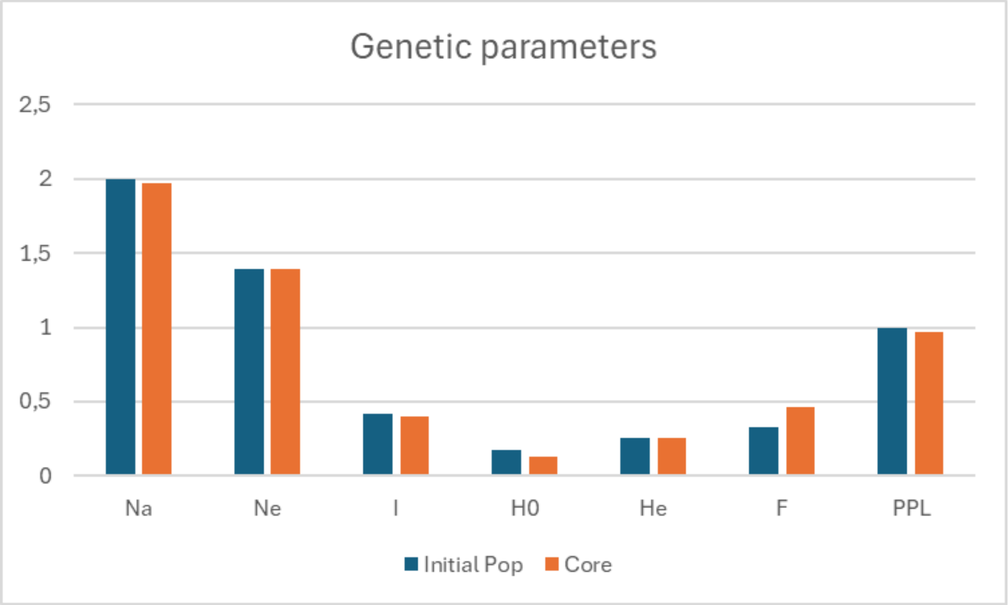

Percentages of Molecular Variance

➢ Low genetic differentiation: due to semi-domestication effects, out-crossing nature and the extensive gene flow between groups.

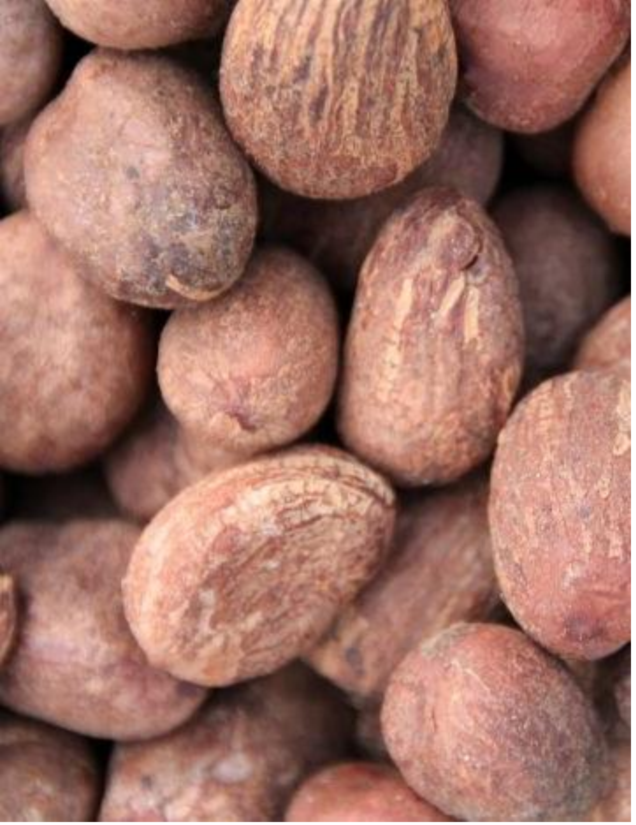
➢ Characteristic for perennial, insect-pollinated out-cross and widespread in a continuous range plant species

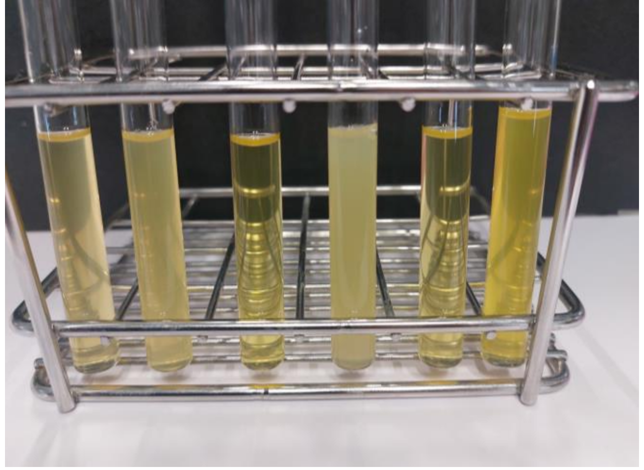
Design of the core collection

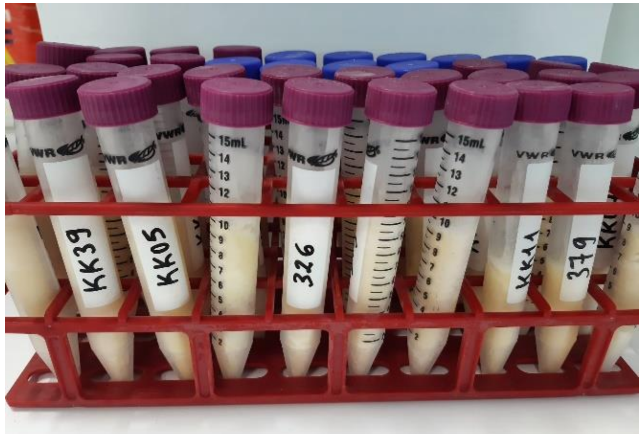
- ❖ DARwin v.6.0.21 (Max. length sub-tree)
- ✓ Core set: 100 shea trees, representing 30% of the full panel.
- ✓ Display similar genetic diversity parameters to the full panel
- ➢ The Core captured the full allelic diversity
- ➢ Robust for conservation, Genetic diversity, and breeding



Shea butter (SB) content and fatty acid profile







- **122 samples**
- **Extraction:**
- Maceration with n-Hexane
- **Fatty acid profile**:
- Transestherification with BF3
- Analysis: Gas Chromatography

Shea butter (SB) content and Fatty Acid profile

- **Shea butter variation**:

✓ Full: 36 to 58%

✓ HFC: 53.5 to 58%

✓ LFC: 36 to 46%

- **Main Fatty acids**

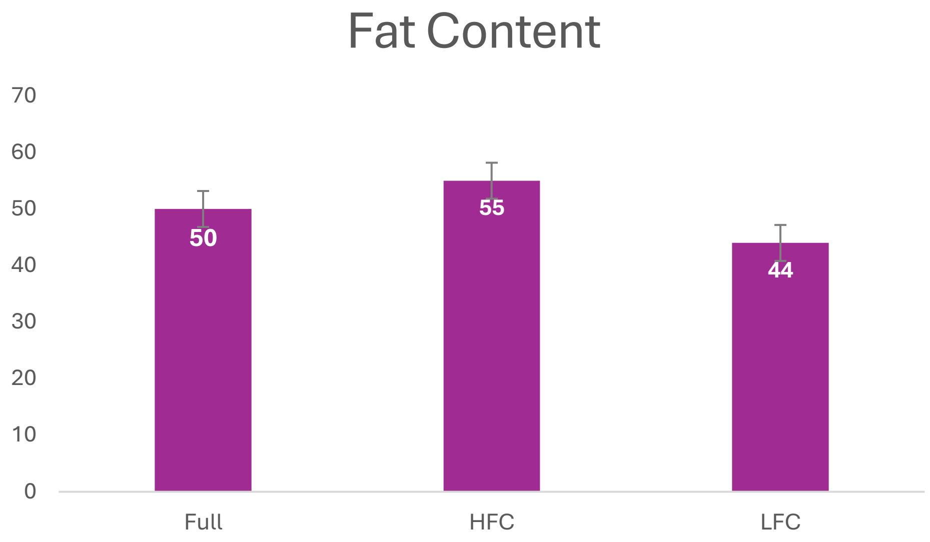
✓ Palmitic acid: 2 to 6.1%

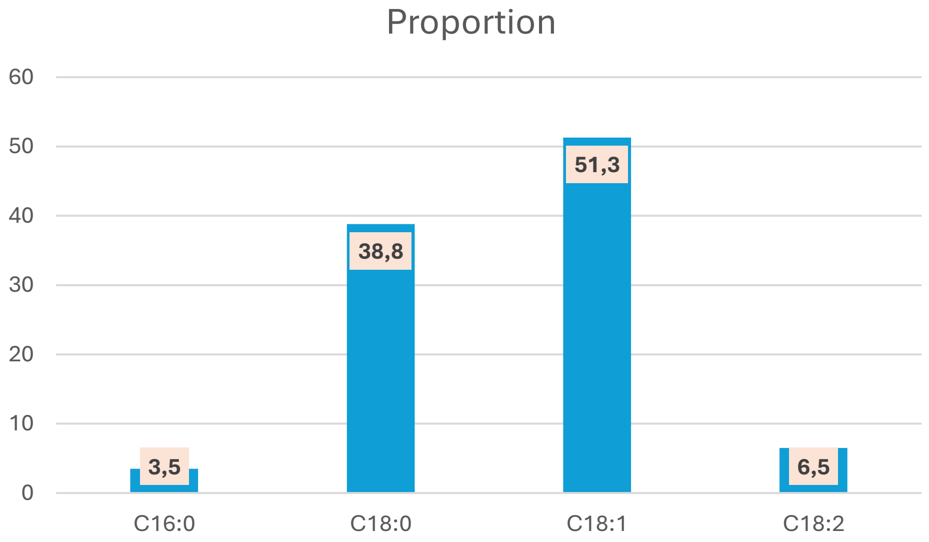
✓ Stearic acid: 22.3 to 50.5%

✓ Oleic acid: 40.3 to 65.7%

✓ Linoleic acid: 4.8 to 9%

➢ Oleic and stearic acids (90%) are known to hydrate and protect the skin. So, SB is highly used in cosmetic and personal care products.

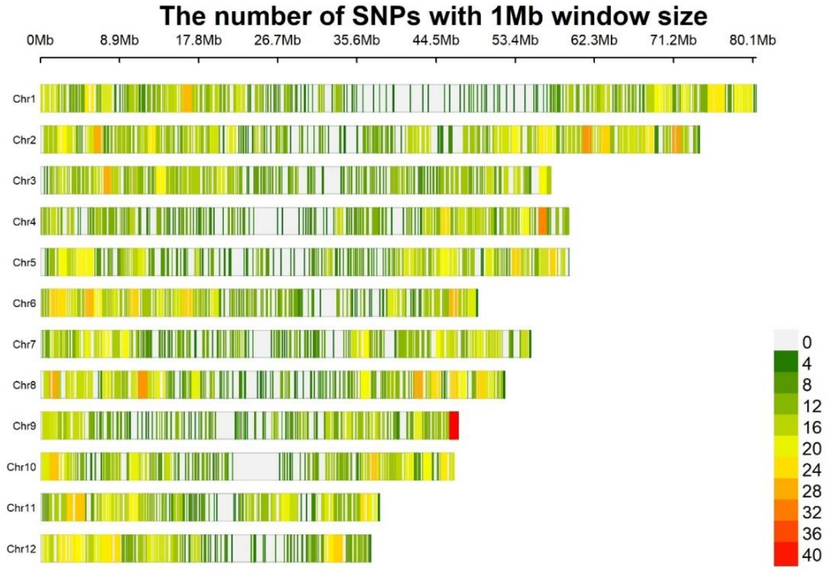




Genome-wide association study

- 7559 SNP markers
- Six Multi-locus GWAS methods:
- mrMLM, FASTmrMLM, FASTmrEMMA, pLARmEB, pKWmEB, and ISIS EM-BLASSO
- QTNs and superior alleles identification
- **Identification of potential candidate genes**

QTNs were searched from the *V. paradoxa* whole genome assembly and annotation

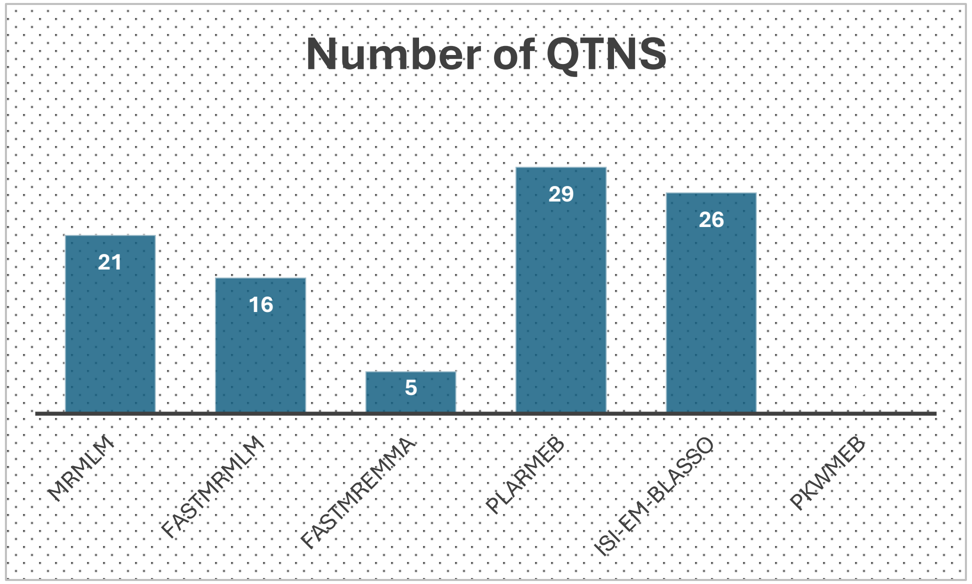


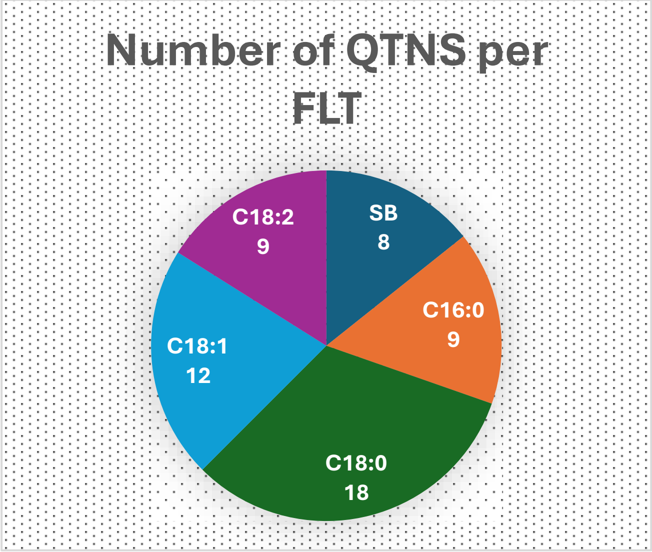
Genome-wide association study

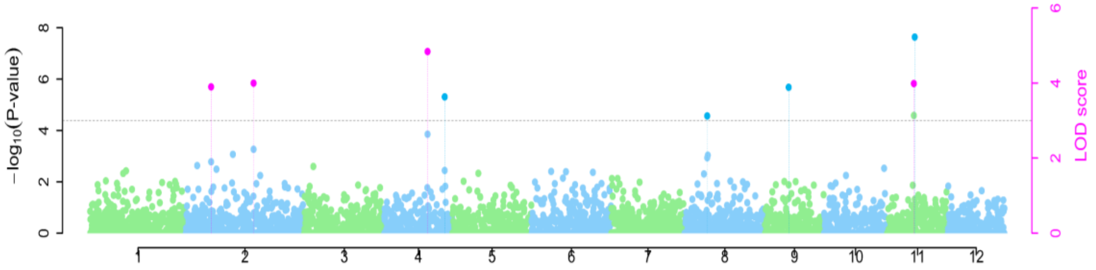
- **Identification of QTNs**

✓ 47 significant QTNs

✓ 25 common QTNs: 4 (C16:0), 4 (FC), 6 (C18:2), 9 (C18:0 and C18:1)







Genome-wide association study

- **Proportion of Superior Alleles**

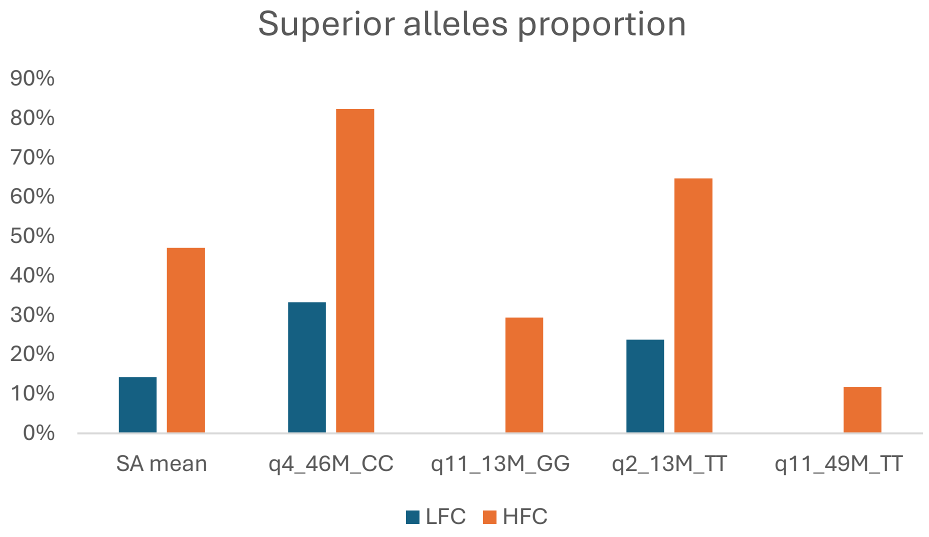
❑ FC QTNs superior allele proportion (SAP)

✓ SAP for LFC: 14.3%; 47.1% for HFC

✓ Two important common QTNs: q4_46M_CC and q2_13M_TT

✓ With 82.4% and 64.7% respectively in HFC

✓ Could be considered as strongly involved in the shea seed oil biosynthesis



Genome-wide association study

- **Potential Candidate genes**

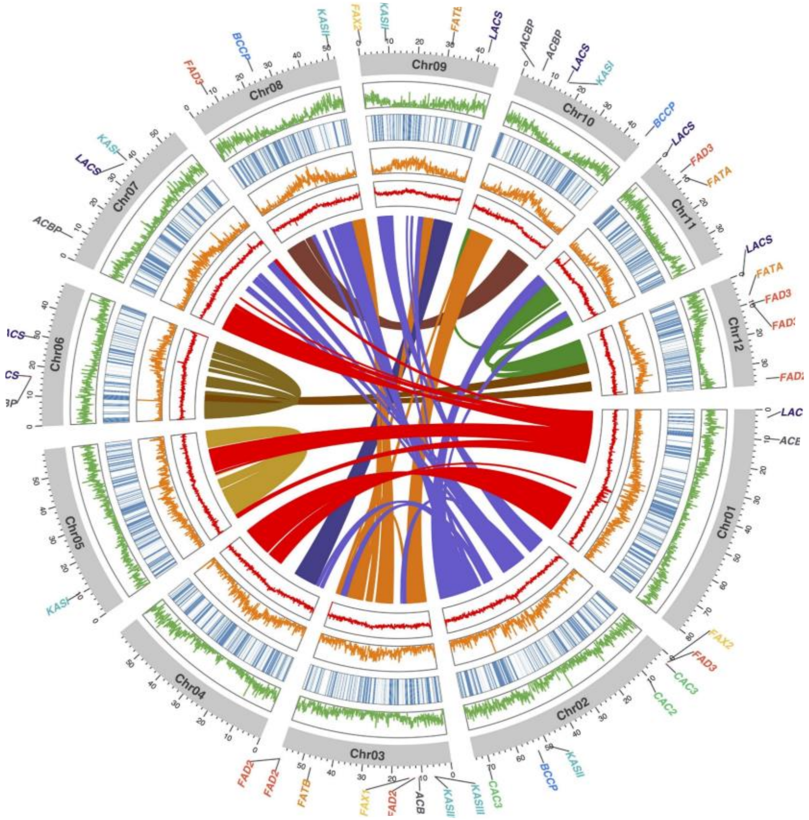
✓ 24 putative genes

✓ Associated with 8 gene/protein families

✓ Involved in FA biosynthesis of shea tree

✓ Fat Content: 3 gene families (*LACS, KINB2, KAS*)

✓ Fatty acids: 7 gene families (*ACBP, BCCP, CAC, FADs, FAX, KAS, LACS*)



Conclusion and perspectives

- ✓ The molecular study highlighted a moderate genetic diversity and AMOVA revealed a low genetic differentiation the three genetic groups
- ✓ A core collection of 100 trees was designed, capturing 30% of the population while preserving all allelic diversity.

❖ Propagation of the core collection via grafting and conserved in a secure *ex situ* site for sustainable conservation

- ✓ A high variability in shea butter content and FA composition was found and 17 shea trees HFC were identified.
- ✓ A GWAS identified 25 QTNs linked to shea butter production.
- ✓ Two QTNs were seemed to be strongly involved in Shea butter production
- ✓ Were associated with KAS and LACS gene families

Conclusion and perspectives

- Use the HFC samples as parental for breeding

- Creation of SSR markers for molecular selection based on the common QTNs

- Further characterization of superior shea trees should include **triglyceride profiles**, **melting points**, and **free fatty acid content**, which are critical parameters for food and cosmetics sectors.

- Establishing a centralized *ex situ* repository will ensure long-term preservation and enable controlled breeding programs for the genetic improvement of the shea tree.


THANK YOU
FOR YOUR
ATTENTION