Genetic Diversity and Population Structure of Shea Tree For Sustainable Conservation and Breeding

Plant & Animal Genome Conference

PAG 32 / January 10-15, 2025,
San Diego, USA

Affi Jean Paul ATTIKORA[1], Nafan DIARRASSOUBA[2], Taofic ALABI[3], Enoch G. ACHIGAN-DAKO[4], Ludivine LASSOIS[1]

[1]Plant Genetics and Rhizosphere Processes lab. Gembloux Agro Bio-Tech, Gembloux, Belgium, Corresponding author: ajpattikora@uliege.be/affijean0121@gmail.com

[2]Department of Biochemistry-Genetics. University of Péléforo Gon Coulibaly (UPGC), Côte d'Ivoire

[3]Functional and Evolutive Entomology. Gembloux Agro Bio-Tech, Gembloux, Belgium

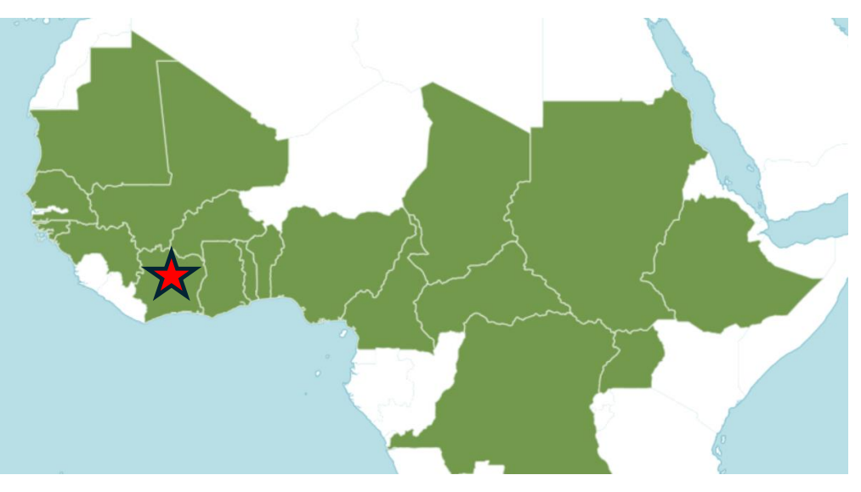
[4]Genetics, Biotechnology, and Seed Science Center, Cotonou, Benin.

Context and objectives

Shea tree (*Vitellaria paradoxa* Gaertn.) grows naturally in 21 African countries. It is diploid (2n=24) and has two subspecies: *paradoxa and nilotica*. It is the second largest oil-producing plant in Africa after the oil palm. Shea butter is used in cosmetics, pharmaceuticals and as Cocoa Butter Equivalent (CBE) in the chocolate industry. The global shea market is estimated to be worth $2.4 billion by 2024. It contributes to women's autonomy in its natural range. The species has a huge potential for carbon sequestration. However, the shea tree faces many threats, including the overuse of shea as a household material, lack of regeneration due to human activities and the effects of climate change. These threats are affecting shea tree population densities and contributing to the loss of interesting genotypes. Actions such as shea tree planting, natural assisted regeneration have been taken to mitigate these threats. In addition, superior shea trees have been identified to establish an *in-situ* collection to conserve the shea tree genetic resources of Côte d'Ivoire. However, *in situ* collections are difficult and expensive to manage, and as a result, shea trees are continually threatened by funding cuts. In this study, SNP markers were used to Assess the genetic diversity and population structure; Design a core collection; Determine shea butter content and fatty acid composition; and Identify QTNs and potential candidate genes involved in the biosynthesis of shea butter related traits. For the long-term conservation and sustainable use of the shea tree genetic resources.



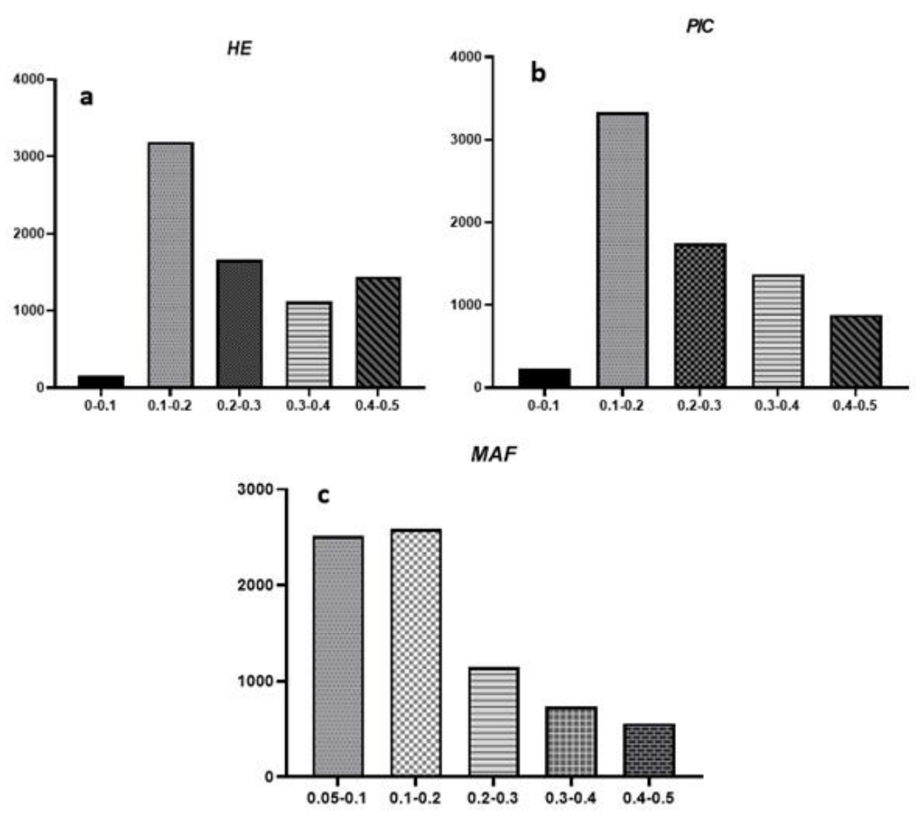






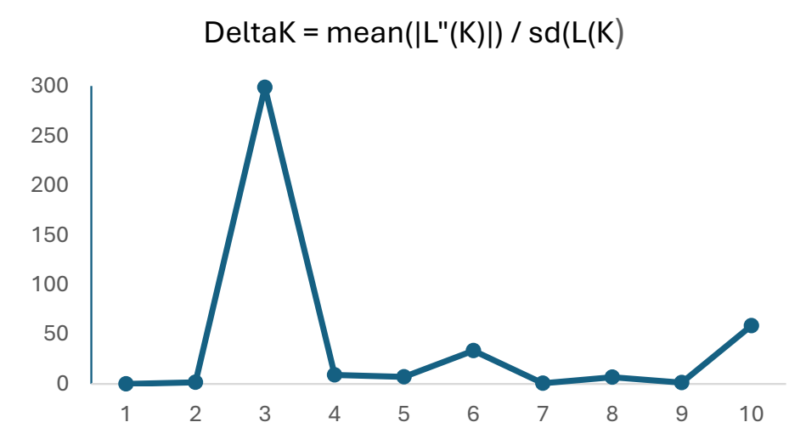
Results

Genetic diversity of Shea Tree using SNP markers

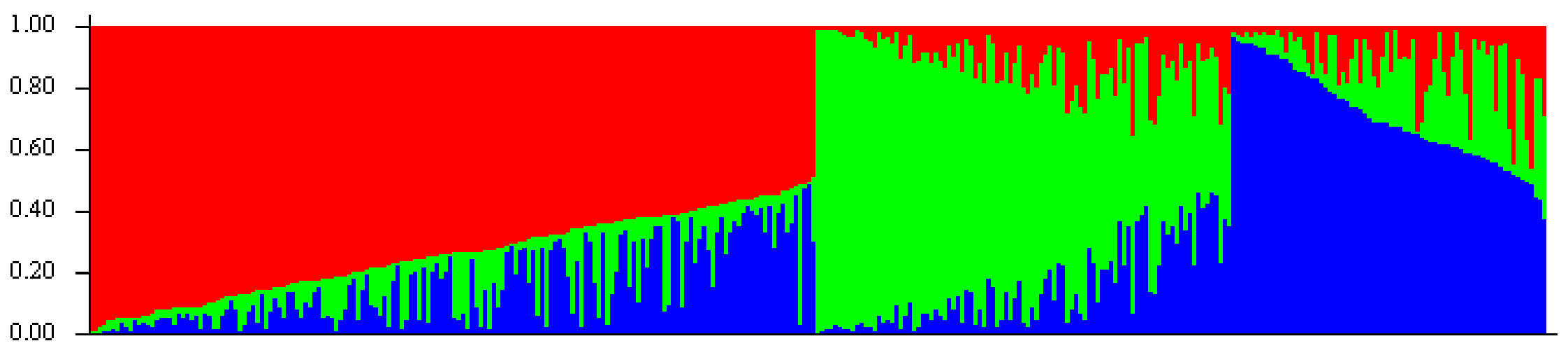


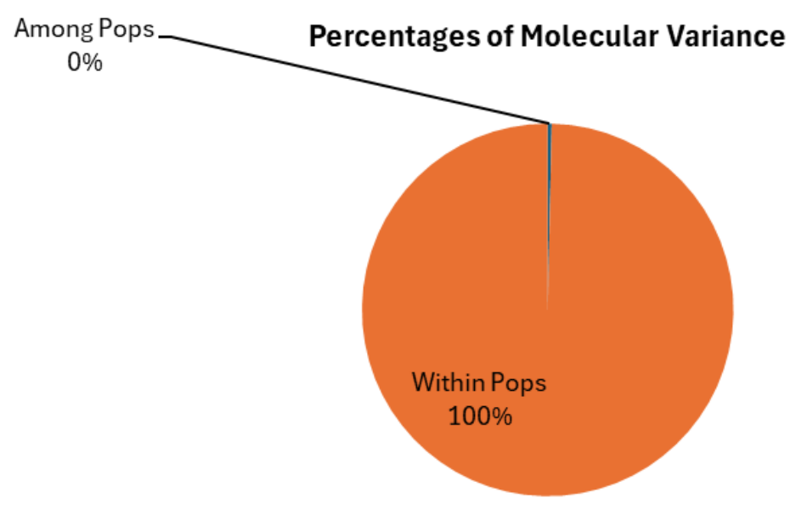
- ✓ Ho: 0.17; He: 0.26; PIC: 0.24
- ✓ MAF: 5,104 SNPs (67.5%) > 0.1
- ➢ Moderate genetic diversity and offers an adaptive opportunity

Population structure and Genetic Differentiation



- Three genetic groups
- Presence of admixtures



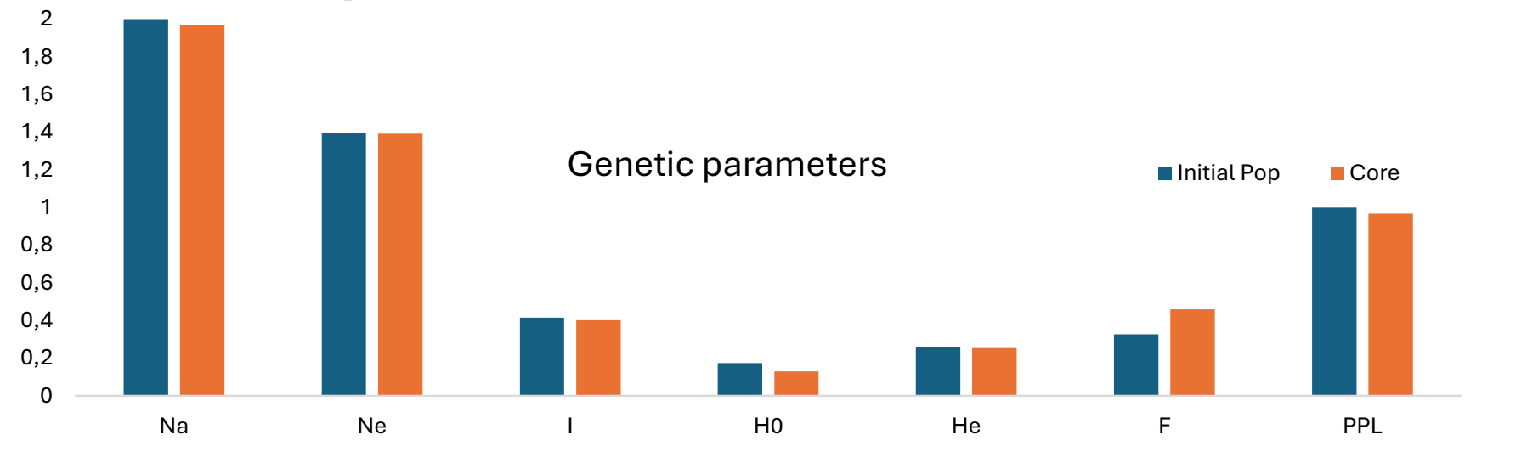


- ✓ 100% of variance within populations
- ✓ Fixation index (Fst): 0.004
- ✓ Nm (Haploid): 59.02

- ✓ These results suggested a low genetic populations differentiation: this is due to the semi-domestication effects, out-crossing nature and the extensive gene flow between the three genetic groups.
- ✓ Characteristic for perennial, insect-pollinated out-cross and widespread in a continuous range plant species

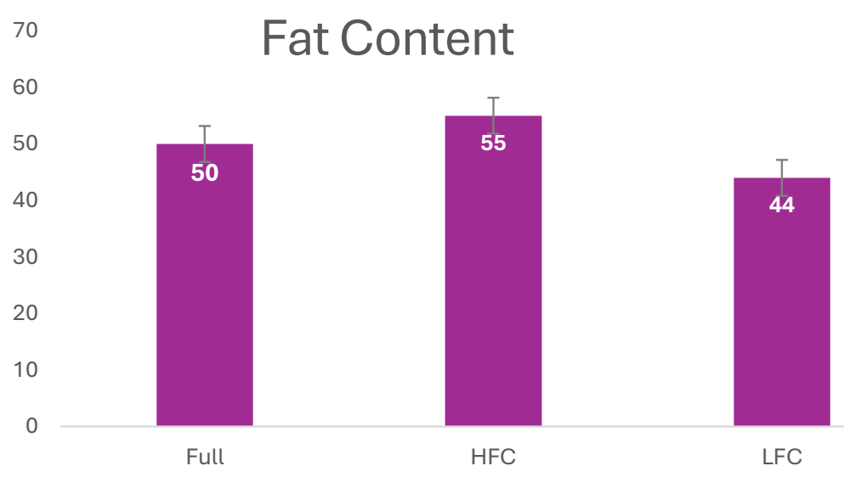
Construction of a Core Collection

- ✓ 100 Shea trees, representing 30% of the full panel, were selected as a core set. They displayed similar genetic diversity to the full panel



- ➢ The Core captured the full allelic diversity and is suitable for conservation, Genetic diversity, and breeding

Determination of Shea Butter Content



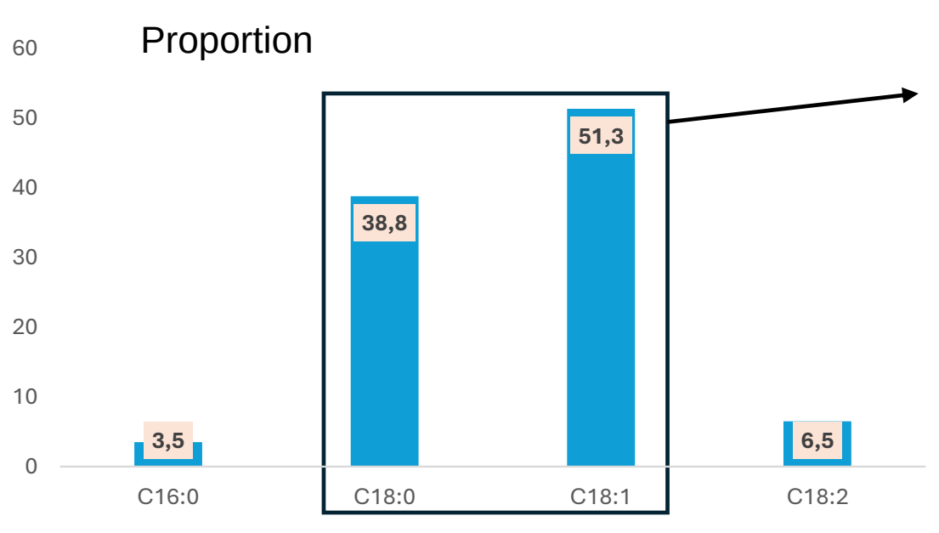
Shea butter variation:

- ✓ Full: 36 to 58%
- ✓ HFC: 53.5 to 58%
- ✓ LFC: 36 to 46%

Determination of Fatty Acid Profile

Principal 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%

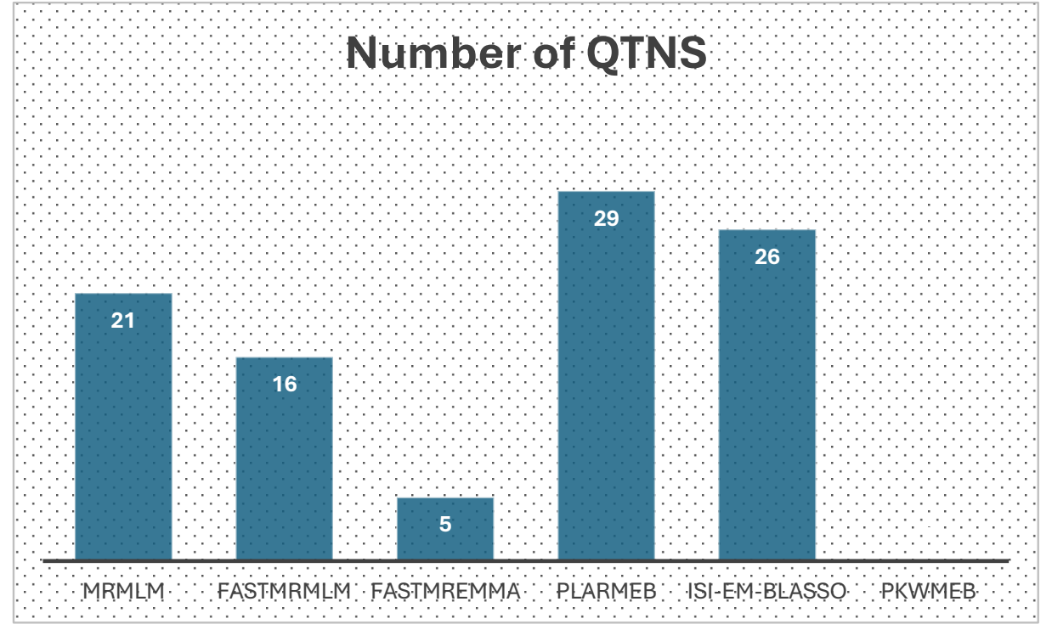


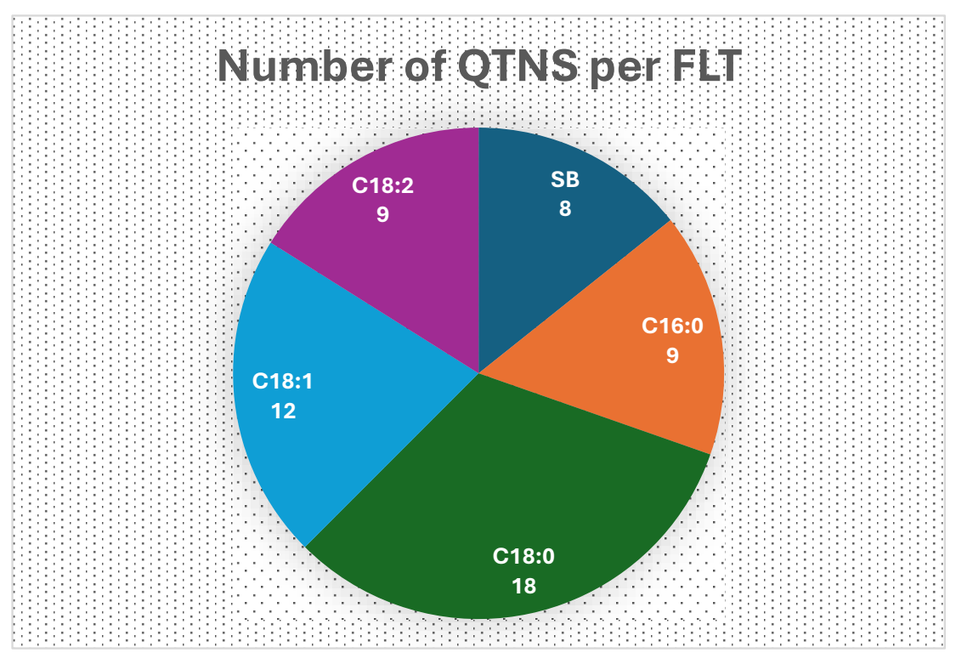
Major fatty acids accounting for 90% of total FA

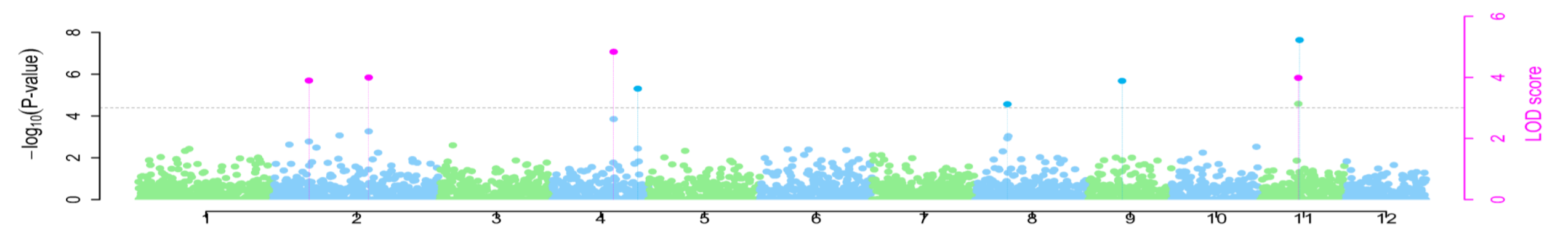
- ➢ Oleic and stearic acids are known to hydrate and protect the skin. So, SB is widely used in cosmetics and personal care products.

Genome Wide Association Study

- ✓ 47 significant QTNs

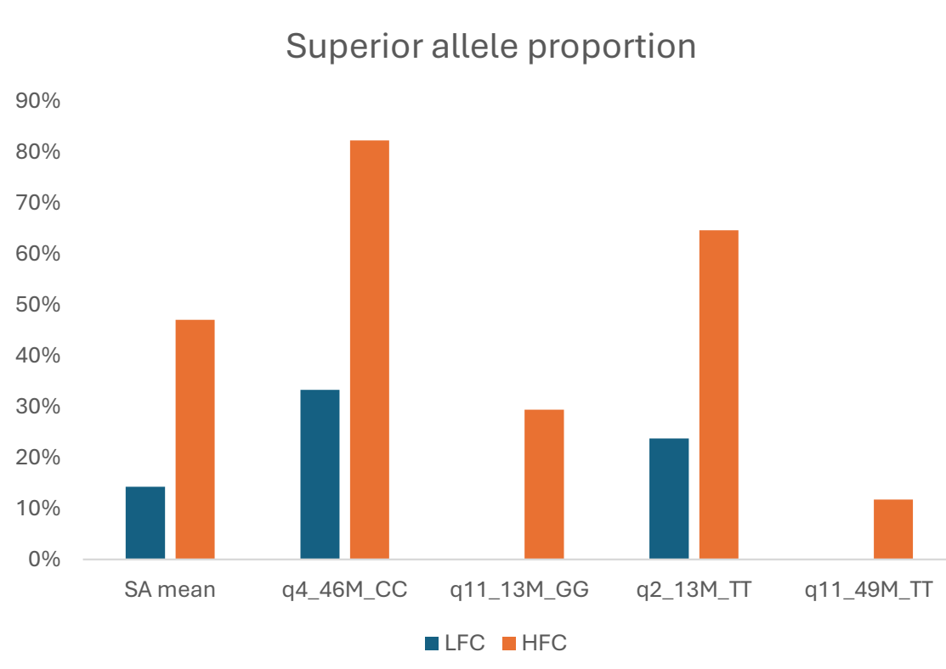






- ✓ 25 common QTNs: 4 (C16:0), 4 (FC), 6 (C18:2), 9 (C18:0 and C18:1)

Proportion of Superior Alleles

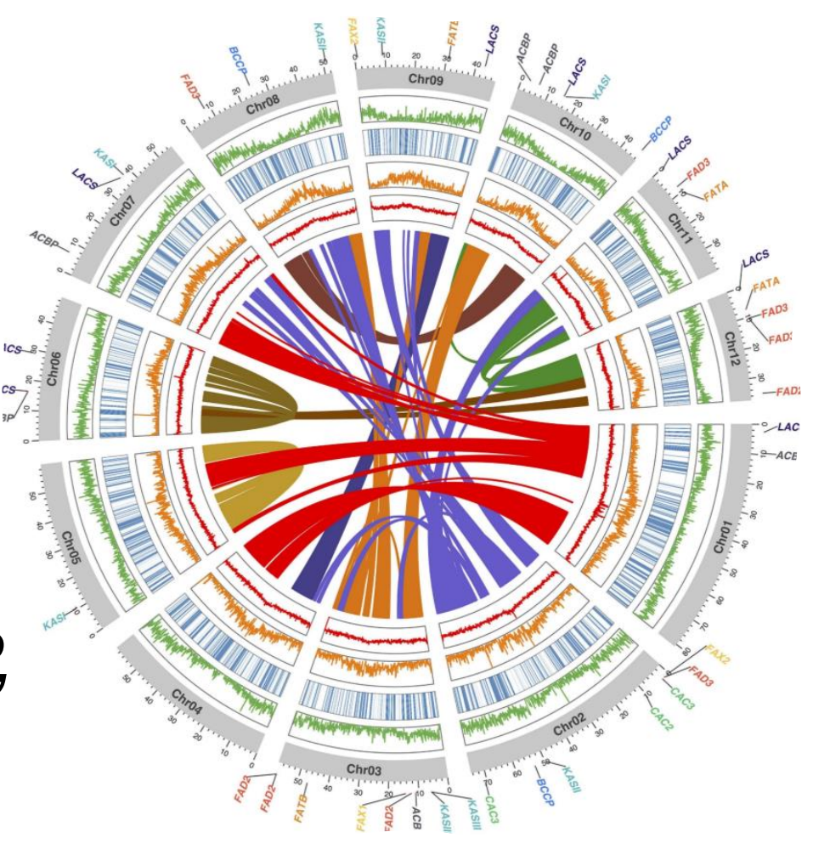


- ☐ FC QTNs superior allele proportion (SAP)
- ✓ SAP for LFC: 14.3%; 47.1% for HFC
- ✓ Two important common QTNs: q4_46M_CC and q2_13M_TT

- ✓ With 82.4% and 64.7% respectively in HFC
- ✓ Could be considered as strongly involved in the shea seed oil biosynthesis

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*)



Methods

- ☐ Leaf sample of 333 shea trees
- ☐ DNA extraction: NucleoMag Plant kit
- ☐ Genotyping: **DArTseq**
- ☐ Markers: **7559 SNPs**
- ☐ Clustering: **Structure Software (2.3.4)**
- ☐ AMOVA: **GenAlEx version 6.503**
- ☐ Core construction: **DARwin software version 6.0.21**
- ☐ Shea Butter Extraction: **Maceration with Hexane**
- ☐ Fatty Acid Composition: **GC gas Chromatograph**
- ☐ **ML-GWAS: R package mrMLM.GUI version 4.0.2**
- ☐ Six GWAS methods: ***mrMLM, FASTmrEMMA, FASTmrMLM, ISIS EM-BLASSO, pKWmEB, pLARmEB***
- ☐ Superior Allele: **QTN Effect value**
- ☐ Candidate gene: ***V. paradoxa* genome Assembly and Annotation**

Conclusions and Perspectives

- ✓ The molecular study, highlighted a moderate genetic diversity within the shea population and AMOVA revealed an intra-individual variance of 100%, indicating low genetic differentiation
- ✓ A core collection of 100 trees was designed, capturing 30% of the population while preserving all allelic diversity.
- ✓ A high variability in shea butter content and FA composition was found and 17 shea trees with high fat content were identified. The oleic and stearic acids, which make up 90% of the total fatty acids, are the determinants of the unique properties of Shea Butter.
- ✓ A genome-wide association study (GWAS) identified 25 significant quantitative trait nucleotides (QTNs) linked to shea butter production. Twenty-four copy genes corresponding to 8 protein families were identified to be involved in FA biosynthesis of shea tree. Two genes, LACS and KAS, were found to be strongly involved in shea butter production.
- ❖ 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.
- ❖ Propagation of the core collection and HFC trees via grafting and conserved in a secure *ex situ* site
- ➢ Establishing a centralized *ex situ* repository will ensure long-term preservation and enable controlled breeding programs for the genetic improvement of the shea tree.