YEPIE ORO FLORA, doctoral student in Genetics and Improvement of Plant Production. I do part of the Biotechnology Axis led by the University Nangui Abrogoua as part of the Program National Research Institute on Cashew Nuts in Côte d'Ivoire.

Tel (WhatsApp): +225 0787227787

CICC Scientific Webinar

Theme: SNP diversity in potentially promising cashew trees (*Anacardium occidentale* L.) identified by participatory selection in Ivory Coast

NAME: YEPIE ORO FLORA

STRUCTURE: Nangui Abrogoua University

Date: 17/10/2024

EXHIBITION PLAN

- Introduction
- Materials and methods
- Results and discussion
- Conclusion and perspectives

INTRODUCTION

❖CONTEXT

- Global demand for cashew nuts is booming.
- Demand response → adoption of strategies and approaches to increase production
- In Ivory Coast, efforts made by the cashew industry and their technical partners → focused on studies focused on the agronomic evaluation of 209 cashew trees by the CNRA (Kouakou *et al* ., 2018)

INTRODUCTION

❖CONTEXT

- Molecular efforts ➡ based on genomic data, included the characterization of 209 cashew trees using microsatellites (Kouakou *et al* ., 2020)
- Efforts are at their beginning.

INTRODUCTION

❖ISSUES

- Ivory Coast, 1,225,935 t in 2023 (FAOSTAT, 2024) has no improved variety.
- Yield remains low in orchards
- 560 kg/ha compared to 1 to 2 t/ha (Brazil, India, Vietnam)
- Aging orchards, unselected genotypes, diverse origins, heterogeneous characteristics, → vulnerability to diseases and insect pests

INTRODUCTION

❖ACTIONS TAKEN

- Since 2019, mass selection in rural areas with the participation of producers
- Four regions belonging to three contrasting ecological zones
- Bafing, Bounkani, Indenié-Djuablin and Folon
- Bounkani and Folon between the Sudanese and sub-Sudanese savannah
- Bafing in the forest-savannah mosaic

INTRODUCTION

❖ACTIONS TAKEN

- Indenié-Djuablin in the semi-deciduous forest
- Two to three locations visited per region
- 40 orchards prospected with 39 producers
- Producer selection criterion: ability to describe the behavior of genotypes
- 121 genotypes selected, evaluated 16 agro-morphological variables

INTRODUCTION

❖ACTIONS TAKEN

- Duration of assessments = four production cycles (2019-2022)
- 42 genotypes, best agronomic performances, declared promising
- Five agro-morphological characteristics:
  - tree height
  - number of hermaphrodite flowers per panicle
  - sex ratio

INTRODUCTION

❖ACTIONS TAKEN

- number of fruits per panicle
- fruiting episode discrimination → 121 genotypes into two pools, regardless of geographic origin

- First pool 44 ▬ genotypes including 20 promising genotypes, best performances
- Second pool 77 ▬ genotypes including 22 promising genotypes, low performance

INTRODUCTION

❖ACTIONS TAKEN

- Better understanding of the revealed structure requires the use of markers with high discriminatory power (SNP).
- To plan future *in situ mass selection efforts* and establishment of *ex situ working collections* , genome-level analysis using molecular genotyping strategies is necessary.

INTRODUCTION

❖RESEARCH OBJECTIVES

Objective 1

Identifying highly discriminatory genetic markers (SNPs) across the cashew genome by genotyping

Objective 2

Using these markers, establish the genetic structure of the 121 genotypes with reference to agronomic performance and geographical origin.

INTRODUCTION

❖RESEARCH OBJECTIVES

Objective 3

Characterize the genetic differentiation between and within the resulting groups

MATERIALS AND METHODS

MATERIALS AND METHODS

❖Definitions

DNA: DNA is the carrier of genetic information in living organisms, especially plants. From a computer point of view, a DNA molecule is seen as a text built on a four-letter alphabet (A, T, G and C).

Genome: The genome of an organism is the set of these letters (DNA) which are found identically in each of these cells.

MATERIALS AND METHODS

❖Definitions

Genotyping by sequencing: Genotyping by sequencing is a process that reveals the arrangement of the different letters A, T, C and G in an individual's genome.

SNP: SNPs are point substitutions of one letter for another that occur at specific positions in the genome.

MATERIALS AND METHODS

❖Definitions

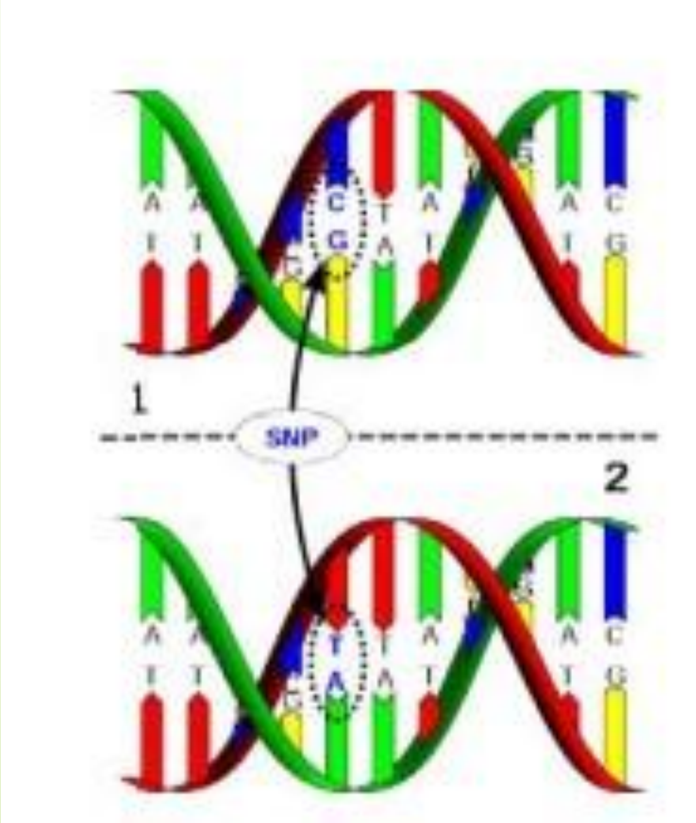

Figure 1. Illustration of a single point substitution (SNP) between DNA sequences 1 and 2 (Nickerson *et al* ., 1997)

MATERIALS AND METHODS

❖Definitions

Allele: An allele is one of the alternative forms of a gene that is located at a specific position in the genome.

MATERIALS AND METHODS

❖Plant material

- 121 samples of mature leaves
- Origin: 121 cashew trees identified by participatory selection
- Origin: Bafing, Bounkani , Indenié-Djuablin and Folon

MATERIALS AND METHODS

❖Methodology

- ➢ DNA Extraction

- ○ Genomic DNA, extracted from ground tissue (Extraction Kit)

- ○ Evaluation of DNA quality on agarose gel

- ○ Quantification and normalization to 3ng/ ul

- ○ 15 DNA extracts removed from the panel

MATERIALS AND METHODS

❖Methodology

- Genotyping by sequencing of the 106 selected extracts

- Genotyping approach: random shearing of DNA by restriction enzymes followed by barcoding (Elshire *et al* ., 2011)

- Three restriction enzymes (PstI , NsI , and ApeKI) were selected for testing.

MATERIALS AND METHODS

❖Methodology

➢ Filtering generated SNPs

- SNPs with a call rate above 80% were retained.
- All monomorphic sites have been deleted.
- Exporting retained data in VCF format using TASSEL software version 5.2.6

MATERIALS AND METHODS

Methodology

- Statistical analyses and genetic differentiation tests
  - Genotypic summary of SNPs identified using TASSEL software version 5.2.6
    - Minor allele frequency (MAF)
    - Observed heterozygosity (Ho)
    - Polymorphism Information Content (PIC)

MATERIALS AND METHODS

❖Methodology

➢ Statistical analyses and genetic differentiation tests

- Observed nucleotide diversity (π)

- Expected nucleotide diversity (ϴ)

- Distance from Tajima (D)

MATERIALS AND METHODS

❖Methodology

➢ Statistical analyses and genetic differentiation tests

○ Data quality control (PLINK and VCF Tools software)

- Genotypes with more than 20% missing data were eliminated from the panel.
- SNPs were filtered to retain only bi-allelic sites.
- SNPs with minor allele frequency < 5% were deleted (Luo *et al* ., 2019; Chander *et al* ., 2021) .

MATERIALS AND METHODS

❖Methodology

- Statistical analyses and genetic differentiation tests

- Genetic structuring of cashew trees

- 99 genotypes and 1146 SNPs saved
- multivariate
- Hierarchical classification (phylogenetic tree)
- Discriminant Principal Components Analysis
- ADMIXTURE → analyzes structuring inference and identifies hybrid genotypes (R software)

MATERIALS AND METHODS

Methodology

- Statistical analyses and genetic differentiation tests
- Genetic structuring of cashew trees
- 99 genotypes and 1146 SNPs saved
- Molecular Analysis of Variance (MAV) → relationship between the clustering mode of genotypes and their geographic origin

MATERIALS AND METHODS

❖Methodology

➢ Statistical analyses and genetic differentiation tests

- Structure and level of differentiation of genotypic groups
  - Molecular Analysis of Variance (AMOVA) ➡ quality of the genetic structure of the groups that have been defined
  - Genetic diversity indices ➡ characterize each group
    - ✓ Consanguinity coefficient (Fis)
    - ✓ Shannon Diversity Index (I)
    - ✓ Effective allele number (Ne)

MATERIALS AND METHODS

❖Methodology

➢ Statistical analyses and genetic differentiation tests

- Structure and level of differentiation of genotypic groups
  - ✓ Different number of alleles (Na)
  - ✓ Number of private alleles (Ap)
  - ✓ Observed heterozygosity (Ho)
  - ✓ Expected heterozygosity (He)

RESULTS AND DISCUSSION

RESULTS AND DISCUSSION

➢ Characteristics of SNPs identified on the cashew genome

○ Results

- 193,855 SNPs identified following genotyping
- Distance from Tajima (- 0.015)
- Proportion of Heterozygosity (Ho) between 0.01 and 0.071
- Weaker (Ho) F42 ➡ genotype
- Average proportions (Ho) 35 ➡ genotypes

RESULTS AND DISCUSSION

- Characteristics of SNPs identified on the cashew genome
  - Results
    - Stronger (Ho) → genotypes A13 and E21
  - Discussions
    - Tajima's D statistic to → test the neutral equilibrium model and infer the action of natural selection based on different characteristics of sequence data
    - Negative Tajima value positive → selection at the level of genes essential for survival and reproduction

RESULTS AND DISCUSSION

- Characteristics of SNPs identified on the cashew genome
  - Discussions
    - Low proportions of heterozygotes, observed in 193,855 identified SNPs
    - Low proportions of heterozygotes confirm the positive selection phenomenon that occurred at the loci.
    - Selection distorts the frequency of genetic variants by inducing the segregation of an excess of rare variants relative to the equilibrium model (Stephen and Brandon, 2005) .

RESULTS AND DISCUSSION

- Characteristics of SNPs identified on the cashew genome
  - Discussions
    - Positive selection drives a beneficial allele to high prevalence or fixation rapidly within a population. When a beneficial allele and surrounding variants on the same haplotype together reach high prevalence, they lead to a reduction in heterozygotes (Vitti *et al* ., 2013) .
    - There is genetic stability at the level of the 106 genotypes

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes

- Results

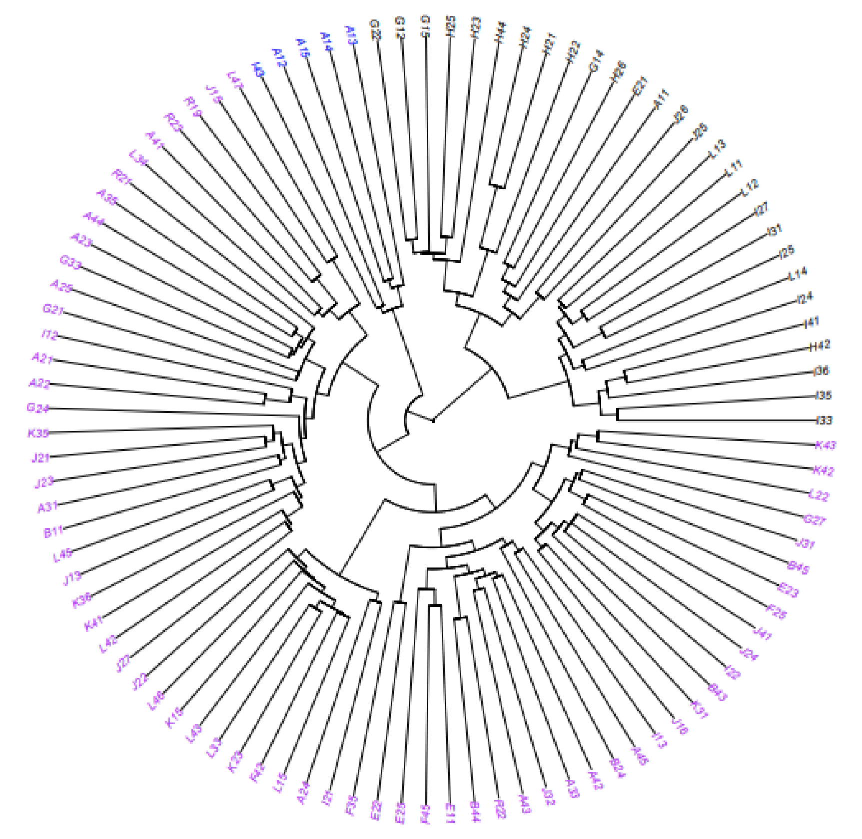

Figure 2. Phylogenetic tree of the 99 genotypes, group 1 (black), group 2 (blue), group 3 (purple)

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes
  - Results
    - First group: 28.28% with 28 genotypes (six Bafing; one Bounkani ; 20 Folon and one Indenié-Djuablin)
    - Second group: 5.05% with five genotypes (four Bounkani and a Folon)
    - Third group: 66.67% with 66 genotypes (28 Bafing; 22 Bounkani ; eight Folon and eight Indenié-Djuablin)

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes

- Results

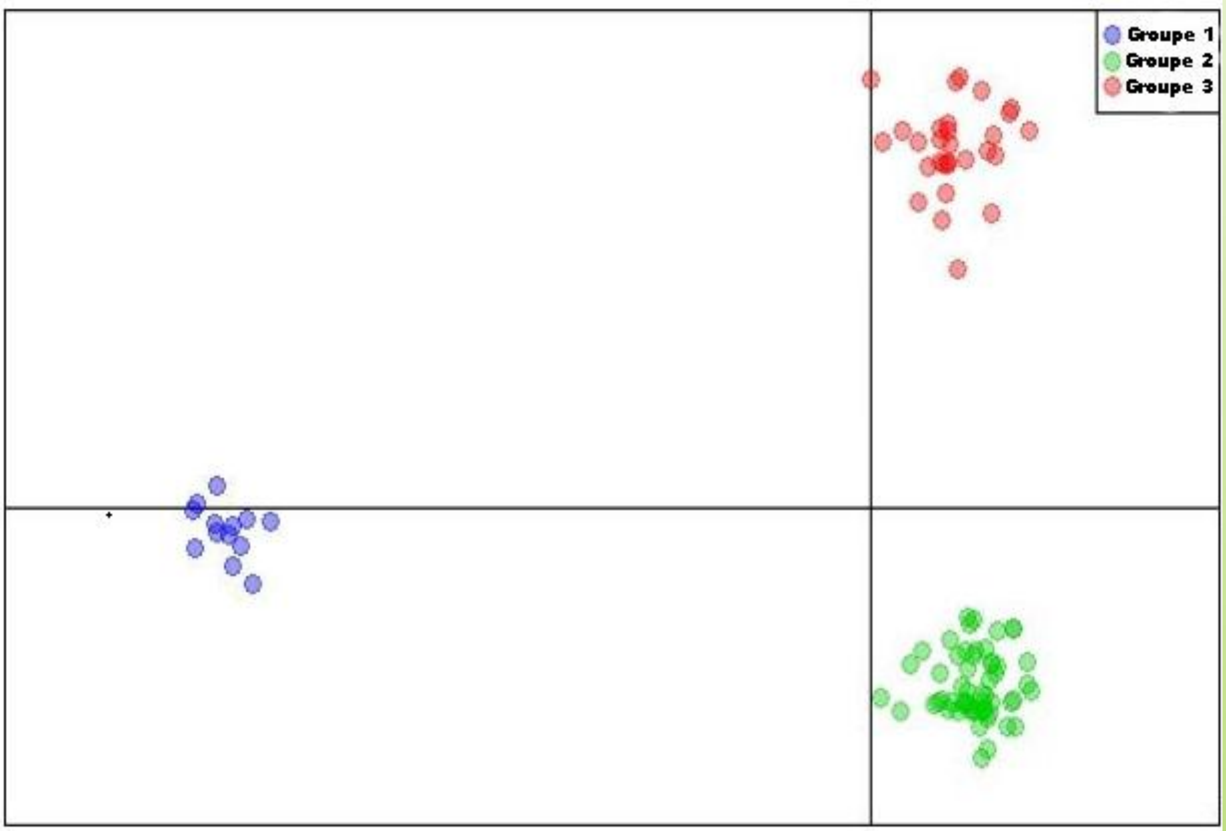

Figure 3. ADCP of the 99 genotypes: group 1 (blue), group 2 (green), group 3 (red)

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes

- Results
  - First group: 13.13% with 13 genotypes (three Bafing; two Bounkani ; two Folon and six Indenié-Djuablin)
  - Second group: 57.58% with 57 genotypes (23 Bafing; 25 Bounkani ; seven Folon and two Indenié-Djuablin)
  - Third group: 29.29% with 29 genotypes (eight Bafing; 20 Folon and one Indenié-Djuablin)

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes

- Results

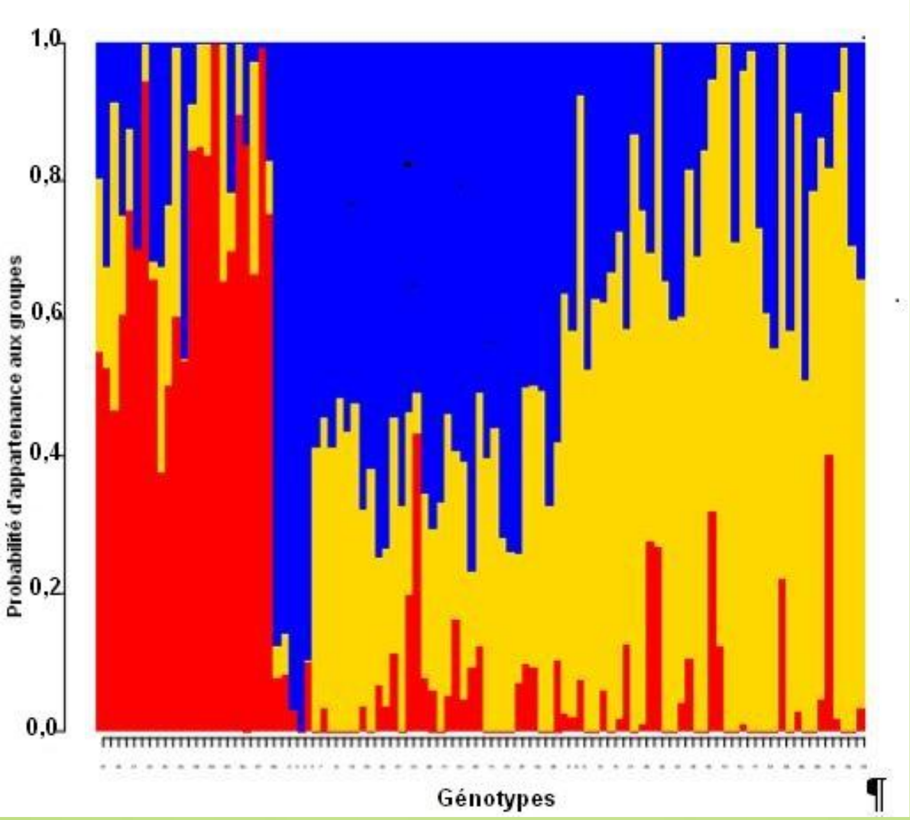

Figure 4. Affiliation diagram (Admixture Analysis): group 1 (red), group 2 (gold), group 3 (blue)

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes
  - Results
    - However, presence of hybrid genotypes Group ➡ 4
    - First group: 33.33% contained 33 genotypes with 15 promising genotypes
    - Second group: 35.35% contained 35 genotypes with 11 promising genotypes
    - Third group: 20.20% contained 20 genotypes with 7 promising genotypes

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes
  - Results
    - Fourth hybrid group: 11.11% contained 11 genotypes with 5 promising genotypes
    - AMOVA (*P* =0.664), assuming a dependence of the clustering of genotypes on geographical origin, not significant

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes

- Discussion

  - Genetic relatedness, ADCP and ADMIXTURE used as complementary approaches to ➡ classify genotypes into three distinct gene pools , regardless of geographic origin
  - Result confirmed by AMOVA
  - This assignment of genotypes is the consequence of a genetic flow through exchanges of material between regions.

RESULTS AND DISCUSSION

- Genetic relationships among 99 cashew genotypes
  - Discussion

- Admixture is a hybridization that occurs naturally between populations of the same species. This phenomenon improves plant performance and leads to new adaptive genotypes (Charlesworth and Willis, 2009; Verhoeven *et al* ., 2011; Shi *et al* ., 2018) .
- 11 hybrid genotypes have good adaptive capacity to environmental variations.

RESULTS AND DISCUSSION

- Characterization of genetic variability between 99 genotypes
- Results
  - Genetic variance 12% recorded between four groups
  - 80% genetic variance recorded within the four groups
- Discussions
  - Low variability results from common parental backgrounds, especially since all Ivorian cashew trees are from the Jumbo variety.

RESULTS AND DISCUSSION

➢ Characterization of genetic variability between 99 genotypes

○ Discussions

- It is also due to the use of seeds from mass selection when creating orchards.
- High variability within groups is linked to the allogamous reproduction mode of the cashew tree.
- Allogamous species maintain high genetic variation within populations and low variation between populations (Hamrick and Godt, 1997) .

RESULTS AND DISCUSSION

- Characterization of genetic variability between 99 genotypes

- Result
  - Private alleles, discovered in certain genotypes, particularly F35 (484 private alleles)
- Discussion
  - Presence of private alleles indicates unique genetic variability at certain loci (Brondani *et al* ., 2006; Salem and Sallam , 2016) .
  - F35 is a genotype with exclusive genetic variability

CONCLUSION AND PERSPECTIVES

CONCLUSION AND PERSPECTIVES

➢ Conclusion

- 1146 highly discriminating markers (SNPs) were identified in cashew (*Anacardium occidentale* L.) for the first time in Ivory Coast.
- A genetic base of 106 cashew trees identified through participatory selection has been established. This genetic base is stable.
- The genetic basis of 15 cashew trees, including four promising ones, has not been established.

CONCLUSION AND PERSPECTIVES

Conclusion

- 99 cashew trees including 38 promising ones were structured into four genetic groups.
- Fourth genetic group consisted of cashew trees hybrids including five promising ones .
- 22 cashew trees, including four promising ones, were not assigned to the identified genetic groups.

CONCLUSION AND PERSPECTIVES

Conclusion

- Distribution of cashew trees in groups is not linked to geographical origin.
- Promising cashew trees that had been classified into two groups phenotypically are actually classified into four genetic groups.
- Cashew (F35) with exclusive genetic variability has been identified in Ivory Coast.

CONCLUSION AND PERSPECTIVES

- Recommendations

- Cashew industry

  - Make promising cashew clones available to producers for the renewal and creation of orchards.

- Producers

  - Use of clones from different genetic groups for the renewal and creation of orchards.

CONCLUSION AND PERSPECTIVES

- Recommendations
  - Selectors
    - Use this additional resource to strengthen the cashew germplasm established by the CNRA.
    - Integrate promising cashew trees, particularly hybrid cashew trees, into the breeding program initiated by the CNRA.
    - Conserve the F35 genotype *in vitro* for future uses

CONCLUSION AND PERSPECTIVES

- Perspectives

- Undertake further work that will be devoted to the implementation of a crop-assisted selection program (KAPS), involving the most promising cashew trees in order to combine interesting complementary traits.
- Research QTLs associated with traits of economic interest (nut yield and CNSL production potential) in order to improve the varietal selection program initiated in Ivory Coast

THANKS !!!

Contact Us:

Executive Secretariat

27 BP 604 Abidjan 27, Ivory Coast Plateau, Avenue Franchet d' Esperey , OLLO Building, Floor 1, Pavilion B.

Phone: +2252720272289 Email : secretariat@ciccashew.org web: www.ciccashew.org