Coconut in its environment

- A coastal palm tree, adapted to marginal environments
  - Fruit adapted to dissemination by currents
  - Tolerance to salt stress and poor soil
- Today, it is mostly cultivated
- Confronted to various biotic threats
  - Phytoplasma diseases (e.g., lethal yellowing)
  - Fungal diseases (e.g., phytophthora)
  - Nematodes (e.g., red ring)
  - Insects (e.g., rhinoceros beetle, red palm weevils)

Economic importance of coconut

- Stagnation of copra oil market
  - Used to be a major oil crop
  - Three major producers (India, Indonesia, Philippines)
  - Competition of oil palm
  - Ageing plantations
- Remains essential in the economy of producing countries
  - Household consumption and local market
  - Diversity of usages
- New markets are growing
  - Coconut water
  - Virgin coconut oil (VCO)
  - Sugar, vinegar, from sap.
Malayan Yellow Dwarf (self-pollinating)

Challenges for coconut genetic improvement
- Low prolificacy, long life cycle, low planting density
  - 5 isolation bag to 60 seed garden seednuts/mother palm/year (Thousands in oil palm)
  - 5-6 years from seed to seed, but 15 years minimum to assess performance at adult age
- 143 to 205 palms/ha Large experiments
- Assessing biotic stress resistance is difficult
  - Phytoplasma cannot be cultivated
- Early resistance tests generally unavailable
- Other factors
  - Seeds are expensive
  - Smallholder crop, Little investment from private industry
  - Coconuts are long lived little incentive to replant
  - Population hybrids are highly productive but not always adapted
  - Multiple selection objectives

Diversity of coconut (Structure)

Distribution of genetic diversity in coconut
Dual origin of cultivated coconut

STRUCTURE OF GENETIC DIVERSITY
AMOVA (10 SSR markers)

Benefits of genome studies in coconut
- Providing guidelines to design crosses
  - Objective:
    - Maximize heterosis
    - Maximize diversity of selected members in advanced generation of crosses between
      - Indo Atlantic group (Talls)
      - Pacific group (Tall and Dwarfs)
- Saving space
  - Selection in nursery
- Reducing breeding cycle duration
  - By skipping the observation period (5-6 year breeding generation)
- Identifying genes of interest
  - Identifying QTLs
    - Segregating populations.
    - GWAS: Genome-Wide Association Studies.
  - Candidate gene approach
    - Exploring results obtained in other species
    - Validating candidate genes in coconut

Inferred from the text:

- The image shows genetic diversity of coconut, with a focus on the dual origin of cultivated coconut.
- Challenges include low prolificacy, long life cycle, and difficulties in assessing biotic stress resistance.
- Benefits of genome studies are highlighted, including providing guidelines for designing crosses and reducing breeding cycles.

Inference Explanation:
- The text indicates the importance of understanding genetic diversity in coconut to improve crop quality and efficiency.
- Challenges such as low prolificacy and long life cycle highlight the need for effective breeding strategies.
- The benefits of genome studies are outlined, emphasizing the potential for increasing productivity and addressing specific genetic traits through modern techniques.
Transcriptomics

Studies genes expression in a given plant organs, possibly according to experimental conditions (e.g., abiotic or abiotic stress, tissue culture). Allows to identify genes of interest. Several studies already published. A recent paper published by CIRAD-AGAP, along with various transcriptomic resources.

Improved transcriptome

- Exploiting previously released transcriptomes
- Using the oil palm protein sequences as a guide to reduce redundancy and fragmentation
- When result is aligned onto rice gene sequences:
  - 1) More complete gene sequences
  - 2) Less incomplete fragments
  - Also, less redundant and chimeric genes.

Transcriptomic resources

Other functionalities:
- Retrieving coconut protein sequences
- Sequence Alignment (Blast)
- Metabolic pathways
- Downloads

Oil palm sequence
Oil palm genes
Coconut homologous genes

The coconut genome is large

- 1.3 times the oil palm genome and 14 times that of rice

Coconut genome sequencing

- Has become an active research field.
  - Chloroplast sequence in 2013
  - Mitochondrial sequence in 2016
  - Nuclear sequence Several announcements
    - January 2014: Saudi Arabia and China (CAAS). (Poster)
    - August 2017: Philippines. (Conference paper)
    - September 2017: India. (Newspaper article)
  - First publication in October 2017: China (CRI-CATAS and BGI), France (CIRAD), Côte d’Ivoire (CNRA)

A collaborative project: producing a reference genome sequence of coconut

- Aim:
  - producing a reference genome sequence of coconut
- Components:
  1. Genome sequencing of a Hainan Tall (CRI-CATAS and BGI – China)
  2. Scaffold assembly (CRI-CATAS – China)
  3. Production of a mapping population (CNRA – Côte d’Ivoire)
  4. Identity checking (CIRAD – France)
  5. GBS genotyping (CIRAD with Genotoul – France)
  6. Linkage map construction (CIRAD – France)
  7. “Pseudomolecule” assembly (CIRAD – France)
  8. Genome expression study in saline condition (CRI-CATAS – China and CIRAD - France)
  9. Phenotyping the mapping population (CNRA – Côte d’Ivoire)
Draft sequence: \textit{de novo} sequences of \textit{Cocos nucifera} genome

- More than 400 gigabases (Gb) sequences
- Illumina HiSeq 2000 platform
- paired-end and mate-pair libraries of various lengths
- 170 $\times$ read depth

Draft sequence: assembly

- Total scaffold length: 2.20 Gb
  - (N50 = 418 Kb),
  - 90.91\% of the genome.
- 28,039 predicted protein-coding genes. Less than in \textit{Phoenix dactylifera} and \textit{Elaeis guineensis}.
- BUSCO evaluation: scaffold sequences covered 90.8\% of the coconut genome, and the genome annotation was 74.1\% complete.
- 72.75\% of the coconut genome consisted of transposable elements (52.23\% of which are LTRs)
  - $\sim$ 50% in oil palm
  - Excluding TR, \textit{C. n.} and \textit{E. g.} genomes are about the same size.

A coconut mapping population

Expected disjunction 50\% DD, 50\% DT

Mapping strategy

- Perform GBS (Illumina)
- Align resulting sequences onto coconut scaffolds (TASSEL)
- Identify polymorphic SNPs with expected distribution
- Filter data (enough data per individual and per marker). Identify linkage groups (Joinmap)
- Order markers while preserving their positions on scaffolds (Scaffhunter)
- Mapping population was planted in Côte d'Ivoire and is under observation. QTLs for early development traits were identified.

Linkage map and coconut sequence anchoring

- A high density linkage map was produced with >8000 SNPs and 16 linkage groups.
- 46\% of the genome was anchored onto the map (representing 77\% of the genes)
- Collinearity between coconut and oil palm genomes
- The common ancestor of coconut and oil palm had 18 chromosomes\textsuperscript{2}.
- In both species, 2 chromosome fusions occurred independently, resulting in 16 chromosomes.

The next steps

- Paper in preparation
- Make sequence and further genomic resources available
  - As a tool for the coconut genomics and breeding community
  - Go on with the translational approach (with oil palm and other crops)
- Resequencing
  - Explore the species' genome diversity
    - Indo-atlantic vs Pacific
    - Dwarf vs Tall
  - Make the best of other sequencing efforts

The next steps

- Transcriptomes
  - In response to biotic and abiotic stresses
- Phenotyping
  - Mapping populations
    - Many traits can be observed in dedicated experiments.
  - Association studies
    - Especially regarding biotic stress
      (identify haplotypes associated with survival)

26

In brief

- Despite its agronomic importance,
  *C. nucifera* is under-studied.
- Yet, advances were made on genome studies
  (transcriptomics, sequencing).
- A published reference sequence will help making the
  best of these efforts.
- Combining phenotyping, resequencing, functional
  genomic and translational genomics will facilitate future
  functional genomics and molecular assisted breeding in
  coconut.

Thank you for your attention