

Unravelling the genomic landscape of minor Root and Tuber crops, cassava and taro, using pangenome approach and population genomics

Description:

Taro and cassava exemplify convergent evolution in response to human selection, despite wide phylogenetic and genomic divergence.

Both plants have undergone convergent evolution to become starch-rich, clonally propagated crops suited to tropical agriculture. Indeed, both developed enlarged underground storage organs (corms in taro, tuberous roots in cassava), driven by human selection for carbohydrate yield. These two crops are primarily propagated vegetatively (taro via corms, cassava via stem cuttings), leading to clonality and reduced effective recombination. Domestication of those two crops led to changes in the regulation of starch biosynthesis genes, storage organ development genes, as well as changes in hormonal signalling (e.g., auxin and cytokinin). In addition, both crops show signs of domestication syndrome, with reduced toxicity (cyanogenic compounds in cassava; acidity in taro) and enhanced yield.

Despite these phenotypic similarities, both crops have many different life-history traits. While taro (Araceae) is a monocot with a haploid genome size of ~2.5 Gb, displaying different levels of ploidy ($2n=24$) (diploid or triploid), and with most likely two origins of domestication in South East Asia and in Oceania, cassava (Euphorbiaceae) is a eudicot with a haploid genome size of ~770 Mb diploid ($2n = 36$), domesticated in the Amazon basin- South America. Furthermore, taro is often sterile with limited sexual reproduction, while sexual reproduction is still more proliferative in cassava.

To unravel the molecular basis and consequences of selection in both taro and cassava, a pangenome approach will be applied, allowing a holistic comparison of the molecular processes underlying their convergent domestication. The advent of sequencing technologies, along with the development of powerful bioinformatic tools, has led to a sharp understanding of the molecular mechanisms underlying evolutionary processes, unveiling the significant impact of structural variants (copy number variations, chromosomal rearrangements, transposable elements, presence/absence variations ...) (even) at the population level. The pangenome approach bears several advantages compared to the classic comparative genomics approach, allowing to unravel, among other things, gene family size variation within a single species, identify genomic regions and features that are shared by a whole species or those that are specific to some populations and account for both single nucleotide and structural variant polymorphisms.

Using whole genome re-sequencing data of cassava (~ 100 accessions) and taro (~ 50 accessions) produced within the ROTATES project, the methodology will be as follow:

- **Pangenome Construction:** to identify within each species core vs. accessory genes, gene presence/absence variations, and structural rearrangements, pangenome graphs will be constructed for each species using at least two already published high quality genome assemblies (from different accessions) for each species. Those graphs will then be augmented with the reads obtained from the sequencing of the different accessions collected during ROTATES, as well as accessions already sequenced (cassava > 300, taro > 5).
- **Comparative Genomics:** To unravel the molecular basis of convergent evolution in cassava and taro, a phylogenomic approach will be applied using the target species (cassava and taro), their wild relatives, and sister groups, as well as outlier species. Those analyses will yield a wide phylogenetic reconstruction, the characterization of ortholog genes for the different clades and families, and the dynamics of gene family size variation and protein domains rearrangements.
- **Detection of Selection Analyses:** To detect footprint selection, genome-wide inter- and intraspecific selection tests will be performed. Interspecific selection tests will be conducted with the detection of selection in phylogenies. At the intraspecific level, several population genetic metrics will be calculated to account for different types of selection.

For all the approaches described above, the PhD student will adapt its workflow according to the literature and the development of new tools and algorithms, enabling state-of-the-art analyses to unravel the molecular basis underlying the convergent evolution of cassava and taro.

Fixed-term contract · PhD research · 36 months Bac+5 / Master's degree
CIRAD - AGAP - DEFI · Montpellier (France)

Start date:

2 March 2026

Application Procedure

Send an email to Dr Chair and Dr Fouks with your CV, cover letter and the contact details of two referees.

Deadline

15 December 2025

Contacts

Hanâ CHAIR

hana.chair@cirad.fr

Bertrand Fouks

bertrand.fouks@cirad.fr