

ABSTRACT

A Molecular Genetic Study of the International Cocoa Genebank, Trinidad
towards Efficient Conservation and Utilisation

Lambert Anthony Motilal

Cacao (*Theobroma cacao* L.) is an important neotropical tree crop from which cocoa butter and chocolate is derived. Cacao germplasm is maintained as trees in field genebanks like the International Cocoa Genebank, Trinidad (ICG,T) which is the largest global public domain collection of cacao. The efficient conservation and utilisation of germplasm is hampered by mislabelling in the collection which may have led to inaccurate descriptions of phenotypic diversity, genetic diversity and genetic relationships. Knowledge of genetic relationships can be used in designing better breeding programs. However, since substantial acreage and time is needed to obtain improved tree varieties, identifying molecular markers for use at the seedling stage could accelerate breeding programs.

This research project, involving molecular fingerprinting, population and conservation genetics; morphology; numerical taxonomy; and plant pathology, aimed to achieve more efficient management of the germplasm, to improve use of the collection and to identify molecular markers for traits of interest.

Detection of mislabelling in the genebank was accelerated by finding a set of nine microsatellite (SSR) loci that gave better identity verification than the conventional set of 15 loci. The ICG,T had mislabelled trees in 25-39% of the

plots, with an estimated 27% homonymous and 7.5% synonymous error rates. A pedigree approach revealed 56 full-sib families within 189 half-sib families, and 16 centres of interconnectivity indicating a high redundancy in the genebank.

Criollo germplasm from Belize, allowed for the first detection of low Criollo ancestry in the Trinitario accessions and in the genebank as a whole. Clustering of accession groups into genetic groups were supported by seven genetic differentiation measures and existence of private alleles based on microsatellite (SSR) and single nucleotide polymorphisms (SNPs). The Refractario genetic cluster had two main subgroups although private alleles from SSRs but not SNPs were present. Phenotypic diversity from 51 traits (leaves, flowers, fruits, seeds and plant disease reactions), including two novel quantitative fruit traits (length and diameter of ovary cavity) were described. Phenotypic data corroborated the separation of genetic clusters. Core collections compiled from SSR, SNP and phenotypic data, were complementary in nature, and thematic core collections were recommended.

An association mapping approach discovered 26 SSR and 61 SNP markers that were significantly, and reliably, linked to fruit and seed traits, as well as, resistance to witches' broom and black pod disease. The markers, superior individuals and trait correlations in this study were used to suggest an association mapping strategy to obtain more productive cultivars.

Keywords: association mapping; cacao; core collection; disease; fingerprinting; genetic diversity; germplasm; phenotypic diversity; seed.