Cassava Reference Set Selection and Preparation for Association Mapping

Morag Ferguson[1], Dominique Dumet (PI)[1], Paula Hurtado[2], Myriam Cristina Duque[2], Jonathan Mkumbira[1] and Martin Fregene[2]

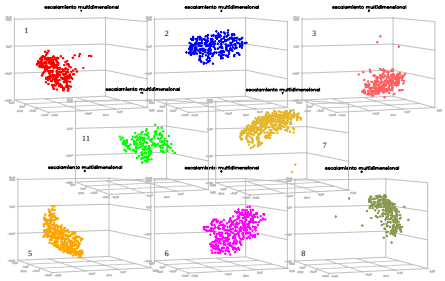
www.iita.org | iita@cgiar.org

Introduction

Studies of association genetics for drought tolerance require phenotypic evaluation in a number of different environments. Unfortunately the movement of cassava germplasm to different environments is difficult due to quarantine restrictions, both between South America and Africa, and within Africa. Where quarantine restrictions apply, all germplasm to be distributed must be certified disease-free. This generally means the plants must be distributed *in vitro*.

Under the Generation Challenge Program (GCP), IITA, CIAT and EMBRAPA completed the genotyping of 3000 cassava clones, using 36 SSR primers. This information has been used to develop a reference set of germplasm (from CIAT, IITA and EMBRAPA) for specific use in association mapping studies for drought tolerance. It is envisaged that the reference set may be used for other purposes such as allele mining across a broad genetic base, and association mapping for other traits, so the concept of a flexible reference set has been developed. The reference set consists of:

- A base reference set. This provides the variation for association mapping, with high diversity and minimum population structure.
- Supplementary trait-based germplasm consisting of germplasm with a range of responses to water stress.
- Supplementary interesting germplasm that consists of germplasm with other interesting traits, such as biotic and abiotic stress tolerance and nutritional traits.



Multidimensional scaling on SSR data of individuals was used to assign clusters used to develop the reference set.

Objectives

This project aims at preparing cassava germplasm for association mapping. This involves:

- Selecting a reference cassava set and known drought tolerant varieties
- Placing selected varieties *in-vitro* and certifying them disease-free
- Multiplying disease-free germplasm
- Distributing germplasm for multi-locational drought tolerance evaluations

Reference Set Selection

Methodology

- Multidimensional scaling was used to assign groupings to 2494 accessions genotyped at 22 SSR loci.
- The number of accessions needed to represent a cluster, to form a base reference set of 250, so that the number of accessions from each cluster is proportional to the logarithm of the cluster size was determined. This methodology, used by Iguarta *et al.* (1998), is thought to represent diversity yet minimise population structure.
- Within each grouping the required number of accessions were selected at random.
- Accessions were identified by IITA and CIAT breeders to form the supplementary 'trait-based' germplasm set, and the 'other interesting' germplasm.

Description of the flexible reference set

- The base reference set consists of 250 accessions, 102 from IITA, 144 from CIAT and 4 from EMBRAPA. It includes 82.5% of the alleles found in the core (175/212), and has representatives from 29 out of 46 countries.
- The supplementary trait-based set consists of 68 clones, 25 from CIAT and 43 from IITA.
- The supplementary interesting germplasm set consists of 201 accessions, 147 from CIAT and 54 from IITA.

Table. 1. Description of the base reference set in relation to the core	Cluster 1	Cluster 2	Cluster 3	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Cluster 11	No cluster
No. of accessions in cluster	267	298	265	337	362	418	287	178	82
No. of accessions in reference set	22	28	29	28	29	30	30	28	26
No. of alleles in core	148	165	165	177	192	153	150	156	117
% of total alleles in cluster present in core	69.81%	77.80%	77.80%	83.50%	90.60%	72.20%	70.80%	73.58%	55.19%
No. of alleles in reference set	126	134	135	142	140	123	119	122	102
% of total alleles present in reference set	59.43%	63.21%	63.68%	66.98%	66.04%	58.02%	56.13%	57.55%	48.11%
% of alleles present in core also present in reference set	85.14%	81.21%	81.82%	80.23%	72.92%	80.39%	79.33%	78.21%	87.18%
Maximum frequency of allele present in core but not in ref set	0.0021	0.0026	0.0019	0.0046	0.0031	0.0019	0.0027	0.0026	0.0020

In vitro propagation and Indexation Process

- Meristems, excised from 4 to 6 week old plants, are aseptically transferred on a Murashige and Skoog basic medium supplemented with NAA, BAP and GA3 and thermo-treated for 5 to 6 weeks (38oC day/ 25ºC night). Shoots are then transferred on MS medium supplemented with NAA and BAP (1). Six to eight weeks later they are sent for acclimatization (2).
- Indexation involves visual observation of acclimatized plants as well as ELISA and PCR tests on leaf samples. Diagnostics are performed for African cassava mosaic virus (ACMV) and East African cassava mosaic virus (EACMV).

Progress

Reference sets have been formed and all accessions have been placed in vitro. Plantlets are ready for virus indexing.

Reference

Igartua, E., M.P. Gracia, J.M. Lasa, B. Medina, J.L. Molina-Cano, J.L. Montoya, and I. Romagosa (1998). The Spanish barley core collection. Genetic Resources and Crop Evolution. 45: 475-481.

[1]International Institute of Tropical Agriculture (IITA)

[2]Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia