Validation of Diversity Arrays Technology (DArT) As A Platform for Whole Genome Profiling in Cassava

Prapit Wongtiem[1], Ling Xia[2], Martin Fregene[3], Carmen De Vicente[4], Jean-Christophe Glaszmann[5], Andrzej Kilian[2]

1. RYFCRC, Rayong , Thailand.
2. DArT P/L, Canberra , Australia..
3. CIAT, Cali, Colombia.
4. IPGRI, c/o CIAT , Cali, Colombia.
5. CIRAD-Agropolis, Montpellier, France.

Diversity array Technology (DArT)

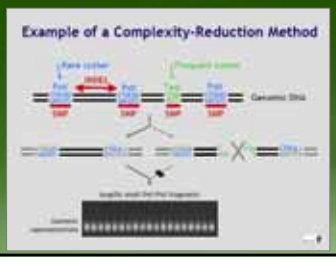
DArT is a high-throughput genotyping system based on a microarray platform

DArT technology consists of several steps:

1. Complexity reduction of the DNA of interest
2. Library creation
3. Microarraying libraries onto glass slides
4. Hybridisation of fluoro-labelled DNA onto slides
5. Scanning of slides for hybridisation signal
6. Data extraction and analysis

1. Complexity reduction

DArT works by reducing the complexity of a DNA sample to obtain a 'representation' of that sample. The method of complexity reduction relies on a combination of restriction enzyme digestion and adapter ligation, followed by amplification



2. Library creation

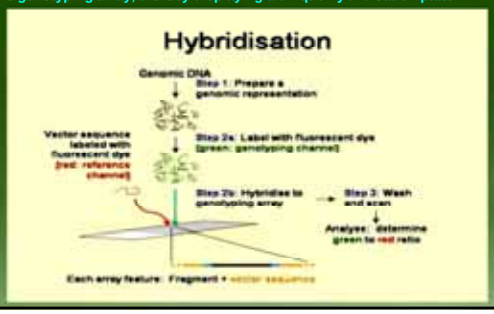
To create a library for any species :

1. A mixture of Genomic 'representations' from a pool of individuals covering the genetic diversity of the species is amplified.
2. These fragments are cloned into a vector that is introduced into *E. coli* to form a library.
3. Within the library, each colony contains one of the fragments from the genomic 'representation'.

3. Microarraying

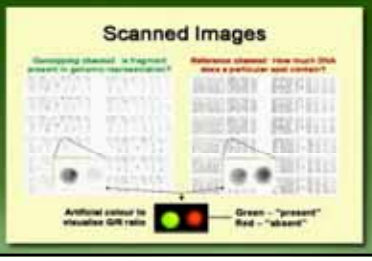
At present the high-throughput capability of DArT is based on a microarray platform. After library creation, a selection of clones from the library are arranged into a plate format (usually 384-well plates). The fragments within the library are amplified and spotted onto glass slides using a microarrayer to form a genotyping array.

4. **The genotyping arrays are hybridised with genomic 'representations' of individual DNA samples prepared using the same complexity reduction method. These individual 'representations' are labelled with one fluorescent label, while the vector fragment is labelled with another fluorescent label to act as a reference. Each individual 'representation' will only hybridise to matching fragments on the genotyping array, thereby displaying a unique hybridisation pattern.**



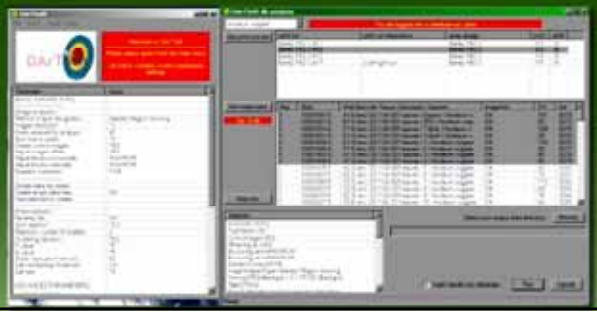
5. Scanning

The hybridised slides are first washed and processed to remove unbound labelled DNA. The slides are then scanned using a scanner to detect fluorescent signal emitted from the hybridised fragments. The result from each fluorescent channel is recorded and the resulting images are stored in tif format.



6. Data analysis

The data from the scanned images is extracted and analysed using the DArT soft software and the information is managed by the DArTdb Laboratory Information Management System.



DArTdb

DArTdb is a Laboratory Information Management System (LIMS) based on a relational database entirely built from Open Source components (using the so-called LAMP architecture: Linux, Apache, MySQL and Perl). Fully barcode-enabled, it has been developed to manage resources, or technology development and the provision of genotyping service.

Project objective

- Develop and validate a high-density array for cassava
- Genotyping of a collection of cassava accessions using high resolution DArT arrays
- Enable analysis of association between DArT marker and dry matter Content (DMC)

Material and methods

- Core collection of 114 cultivated cassava accessions
- 20 wild relative cassava accessions
- DArT representations
  - PstI/ TaqI genomic reduction
  - Slide of 6144 clones for each representation

Results

1. Rearraying of the clones polymorphic in previous analysis :

- Reanalysed data generated during the previous one.(2004)
- Proceeded to rearraying these clones into new "polymorphic clone" library.
- A total of 764 clones were selected and rearrayed using MicroGrid II robot.
- The polymorphic clones were used to print all arrays used in the project together with the new "discovery clones" prepared during library expansion

2. Library expansion

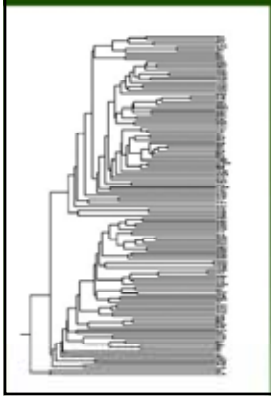
- **Library expansion was carried out using DNA samples from the 134 accessions**
- **134 of the samples were also used for the identification of polymorphic clones.**
- **A total of 5,500 clones were picked and inserts produced and purified as described above.**
- **All candidate and polymorphic clones were printed together on the arrays used for all hybridisation experiments.**

Table 1. Expansion of PstI/TaqI library.

Library			Varieties	Clones
Pst I / Taq I	(1)	(cultivated)	114	2688
Pst I / Taq I	(2)	(wild relative)	20	768
Pst I / Taq I	(3)	(different countries)	40	1920

3. Cultivated accessions

Cassava Diversity Analysis 424 DArT markers



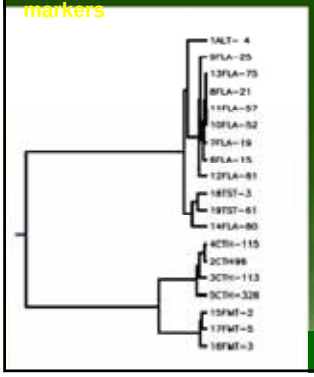
DArT analysis using expanded arrays.

1. **All samples were tested using the methods described and analysed using DArTsoft.**
2. **The percentage of polymorphic clones showing a binary distribution of relative florescence intensity among accessions was observed for 1208 Pst I/Taq clones, whic was equal to 20% of 6144 clones present on the array.**
3. **Using the scores for the best 424 markers and created a UPGMA dendrogram based on the Nei/ Li restriction fragment distance (Fig. 1).**
4. **The relationships among samples are indicative of their provenience,as samples from the same country/geograph region tend to cluster.**

Figure 1. Genetic relationship among cultivated accessions analysed. A UPGMA dendrogram based on the Nei/ Li restriction fragment distance matrix calculated from 424 highest quality Pst I/ Taq I marker data (Nei and Li 1979; Felsenstein 1989).

4. Wild relative

Cassava Diversity Analysis 1710 DArT markers



1. The same array was hybridised with targets prepared from 20 wild cassava accessions
2. The level of polymorphism was significantly higher(over 2000 markers).
3. After selecting 1710 best quality markers the relationship among the samples was represented by UPGMA dendrogram (Figure 2).
4. One sample (FMT 6) could not be Effectively analysed due to DNA quality problems,
5. The remaining 19 samples showed a clear pattern of diversity consistent with their botanical classification.
6. The only exception to this rule is grouping of *M. flabellifolia* #80 with the two accessions of *M tristis.*

Figure 2. Genetic relationship among a group of wild relative 19 accessions. A UPGMA dendrogram based on the Nei/ Li restriction fragment distance matrix (Nei and Li 1979; Felsenstein 1989) using 1711 best quality PstI/TaqI markers.

Conclusions

1. A cassava DArT genotyping array was developed containing over 1,000 polymorphic clones,
2. DArT can be an effective tool for exhaustive fingerprint of germplasm collections.
3. The total number of genomic clones tested (over 12,000) with the samples provided by CIAT provides sufficient density of data to initiate search for marker trait associations in cassava.