OUTPUT 12
Development and use of biotechnology tools for cassava improvement

Activity 12.1 Molecular Marker-Assisted Breeding for Resistance to the Cassava Mosaic Disease in Latin American Cassava Gene Pools

Collaborators:
Cesar Ospina, Jaime Marín, Edgar Barrera, Nelson Morante, Hernán Ceballos, Martin Fregene (CIAT)

Important Outputs
Molecular marker analysis of 1291 BC_2 progenies and identification of genotypes that combine resistance to CMD and CGM.
Testing of FTA paper leaf squashes as a replacement for the DNA isolation step in MAS.

Rationale
Perhaps the most powerful use of MAS in cassava breeding is in selection of recombinants that combine many genes, for example resistance to CMD and CGM, without the need for field trials. This could accelerate efforts to combine many traits into elite varieties.   During the second year of the implementation of MAS at CIAT, we attempted to combine CMD and CGM resistance, derived from a wild relative of cassava, into genotypes that can serve as parents for introgression of the two traits. This year we also sought to address some of the most labor-intensive steps of MAS, namely the tedious DNA isolation step and identification of markers that are useful in a wide range of germplasm.   A low cost method for MAS could be PCR amplification of leaf squashes on FTA paper (Whatmann PLC, UK). An attempt was also made to evaluate the feasibility of this approach for MAS on cassava.

Methodology
Four inter-specific F_1 hybrids with resistance to CGM were crossed extensively to MTAI8, SM1741-1, SM1669-5, SM121-9, CM3306-4, and SM1460-1, elite parents of CIAT's cassava gene pools. BC_1 progenies having resistance to CGM, from field evaluations, were crossed to CMD donor parents at CIAT, a total of 1490 BC_2 sexual seeds from into 43 families were obtained. The sexual seeds were germinated from embryo rescue and multiplied (Activity 12.18, this report). DNA isolation and molecular marker analysis was with a SCAR marker obtained from the RAPD marker RME1 located at less than 4cM from the CMD resistance gene (CIAT 2003) or SSR markers, NS74, NS217 and NS260, associated with CGM and 1 or 2 *in vitro* plantlets as described earlier (Mba et al. 2001).

To reduce the cost of marker genotyping, PCR amplification of leaf squashes on FTA paper (Whatmann Inc., UK) was tested. Fresh leaves were harvested from 15 *in vitro* plantlets, 15 plants in the screen house, and 15 plants from the field, all plants were 2-3 months old. Leaf squashes were made using 0.5-2g of leaf tissue on FTA paper and a 1mm disc excised using a FTA paper punch supplied by the manufacturer (Whatmann Inc., UK). The FTA paper disc was either used directly in the PCR or washed as follows: the disc was transferred to a 96-well PCR plate and 200ul of 70% isopropanol was added and mixed using a pipette, the wash was repeated with IX TE (Tris 10mM, EDTA 1mM) PCR amplifications were conducted with both SCAR and SSR markers as described above.

Results

A total of 1291 plants were successfully established by embryo rescue from the 1490 seeds. Molecular marker genotyping of 1291 progenies from 43 BC_2 families allowed the identification of 335 progenies that combine resistance to CMD and CGM. These progenies have been multiplied and shipped to Tanzania to serve as parents for molecular marker-assisted selection (MAS) introgression of CMD and CGM resistance into local cassava varieties (Activity 12.2, this report). The SCAR marker developed last year for CMD resistance revealed very good results in the analysis of the 1291 BC_2 progenies (Fig8.1).
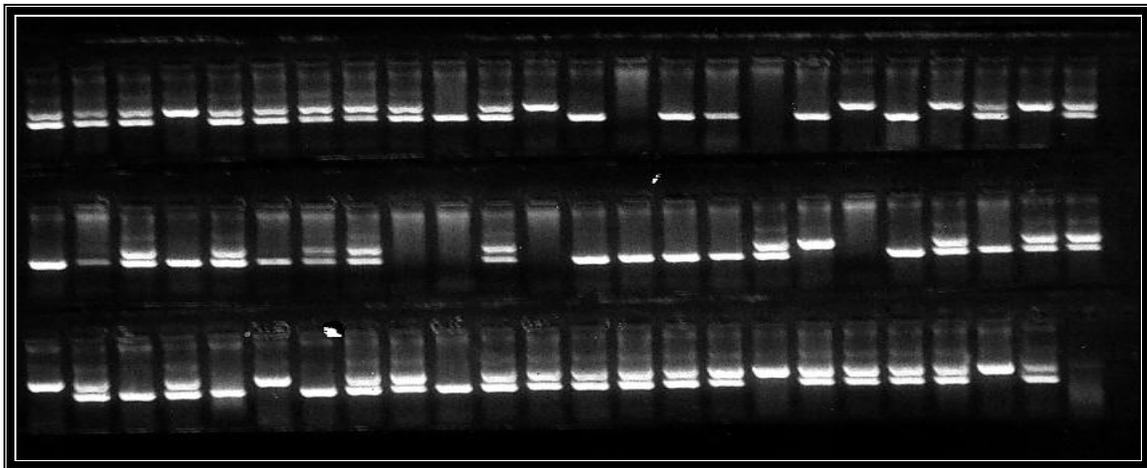
Figure 12.1. PCR amplification of BC_2 progenies using a SCAR marker developed from a RAPD marker RME1 located at less than 4cM from *CMD2*, the CMD resistance gene. The larger weight allele is associated with resistance

PCR amplification of FTA paper discs with leaf squashes from *in vitro* plants was 100% successful for both RAPD and PCR markers with or without the washing step. Leaf squashes using leaves from screen house or field plants were also 100% successful but only after inclusion of the washing step. Elimination of the washing step lead to a high number of failed PCR reactions suggesting that impurities from matured leaves was inhibiting the PCR reaction. This result suggests that FTA paper leaf squashes could replace cumbersome DNA isolation step.

Conclusion and Future Perspective

Molecular marker-assisted selection of 1291 BC_2 genotypes lead to the identification of 335 lines that combine resistance to CMD and CGM. A low cost method for MAS, PCR amplification of leaf squashes on FTA paper, was also evaluated and found to be feasible. We intend to extend these preliminary experiments to many more plants and also to see the effects on PCR amplification of the storage of FTA paper leaf squashes at room temperature for extended periods.

References

CIAT 2003. Annual Report IP3. Improved cassava for the developing world. Pp8-68 to 8-70

Mba R. E.C., P. Stephenson, K. Edwards, S. Melzer, J. Nkumbira, U. Gullberg, K. Apel, M. Gale, J. Tohme, and M. Fregene (2001) Simple Sequence Repeat (SSR) Markers Survey of the Cassava (*Manihot esculenta* Crantz) Genome: Towards a SSR-Based Molecular Genetic Map of Cassava. Theor and Appl Genet 102:21-31

Activity 12.2 Molecular Marker-Assisted and Farmer Participatory Improvement of Cassava Germplasm for Farmer/Market Preferred Traits in Tanzania

Collaborators:

Alois Kullaya, Julius Mugini (ARI Mikocheni, Tanzania), Heneriko Kulembeka (ARI, Ukiriguru , Tanzania), Kiddo Mtunda, Esther Masumba (ARI, Kibaha, Tanzania), Morag Ferguson (IITA-ICRISAT, Nairobi, Tanzania), Luis Fernando Cadavid (CLAYUCA), Jaime Marín, Cesar Ospina, Edgar Barrera, Nelson Morante, Hernán Ceballos, Reinhardt Howeler, Martin Fregene (CIAT).

Funding:

Important Outputs

1) Shipment to Tanzania of 191 genotypes with resistance to the cassava mosaic disease (CMD) and 335 genotypes that combine resistance to CMD and to the cassava green mites (CGM) (derived from a wild relative), up to 10 plants each per genotype. The plants were derived from embryo axes from mature sexual seeds
2) Molecular diagnostics of the introduced material for frog skin disease (FSD) and transfer of the introductions to the field in Tanzania
3) Collection and evaluation of local varieties in Tanzania for crosses to the introductions to introduce resistance to the cassava mosaic disease and green mites into the local varieties.

Rationale

The Tanzanian MAS project funded by the Rockefeller foundation seeks to massively improve farmer-preferred varieties for CMD and CGM resistance by crossing introductions resistant to the above disease and pest to local varieties collected all over the country. The resulting large sized segregating populations are reduced with markers and the best genotypes, from the point of view of the farmer and breeder are selected over 2 cycles of parallel evaluations at

the research stations and in farmer's field. We report here progress in the project this year. A total of 335 BC_2 progenies (AR lines) that combine resistance to CMD and to the cassava green mites (CGM), derived from a wild relative, and 207 genotypes (CR lines) obtained from crossing CIAT elite parents and CMD resistant lines were introduced from Colombia to Tanzania in three shipments this year. They were transferred to the screen house and while there evaluated for frog skin disease (FSD) and then transferred to the field. They will be evaluated later this season and no less than 60 genotypes selected based on evaluation of highly heritable agronomic traits to be crossed to 90 local varieties selected from all over the country. Molecular markers associated with CMD and CGM will be used to discard much of the resulting segregating populations so that the breeder and farmers can concentrate on a small number of progeny having resistance to the principal pest and disease and farmer/end-user preferred traits. The Tanzanian MAS project seeks to transfer useful variability from the crop's center of diversity of cassava to Africa. The concept is already being extended to additional NARs in Africa, the AR and CR lines have been shipped to Uganda and Nigeria already in preparation for crossing to local varieties. Several concept notes have been prepared to fund the above efforts this year and if successful should begin next year.

Methodology
Following a decision made by CIAT management to permit direct transfer of cassava germplasm to African NARs without going through a third party, a committee was set up at CIAT to draw up guidelines for the safe transfer of cassava germplasm to Africa. The following recommendations were made by the committee:
1)A request for cassava materials with information regarding the quarantine requirements of the receiving country.
2) Only seeds from mother plants that are free of cassava frogskin disease (FSD) will be used. The mother plants will be inspected for root symptoms in the field. A significant sample of the mother plants will be tested using a diagnostic method appropriate for CFSD.
3) A record of the results of testing will be kept, and one copy will be sent to ICA quarantine officer.
4) Only plants that are placed *in vitro* through somatic embryo rescue will be exported.
5) Permission to export plants must be obtained from ICA.
6) All seed shipments from CIAT are accompanied by a Material Transfer Agreement (MTA).
7) The receiving country will have a quarantine period before the release of these materials in the field. The quarantine facilities should be insect proof in order to be sure that no biological agents from the receiving country are introduced during the quarantine period.

Two different sets of germplasm were shipped: first 191 F_1 genotypes derived from crosses of elite CIAT lines to CMD resistant parents followed by MAS for CMD resistance (CR lines) and 335 BC_2 genotypes obtained crossing CMD resistant lines to BC_1 derivatives of a wild close relative of cassava introgressed with elite CIAT parents (AR Lines). This second set of genotypes combines resistance to CMD and cassava green mites (CGM). Between 5 and 10 plants per genotype were shipped. The tissue culture plantlets were shipped in three batches, November 15 (CR plants), March 23 (AR first batch), and April 29 (AR second batch), to avoid over-loading facilities at ARI-Kibaha where the plants were sent to. On arrival in Dar es Salaam the plants were received by plant quarantine officials from the Tropical Pesticide Research Institute (TPRI, Arusha) and transferred to the tissue culture growth room of ARI-Mikocheni. After 7 days to allow the plants recover, they were moved to the screen house at ARI-Kibaha for hardening according to standard methods laid down at CIAT (Roca et al.

1984). The plants were inspected after one month in the screen house by TPRI officials and at 2 months just before transfer to the field. After the second inspection and further molecular diagnostics, the plants were transferred to the field at the Alawi estate, a 4000ha sisal plantation owned by the Mohammed Enterprises who are now interested in producing cassava for starch.

In shipping germplasm to Tanzania, the conditions laid down by the CIAT committee on shipment of germplasm to Africa were strictly adhered to. Nevertheless molecular diagnostics for the presence of frog skin disease was carried out while the plants where in the green house to ensure that there was no escape in the germplasm. A molecular diagnostic method for the detection of CFSD based on hybridization of an FSD cDNA clone CFSV-S5 was used. The method is a modification at CIAT of the to dsRNA extraction the Morris and Dodds method (1993??). Briefly, three grams of young leaves or petioles were frozen with liquid nitrogen and homogenized with two volumes of extraction buffer (2X STE, 10% SDS, 1% bentonite, and 0.5% 2-mercaptoethanol) and 0.5 volumes of chloroform:pentanol (24:1). The extracts were centrifuged for 10 min at 8,000 G, and the aqueous phase was collected. Ethanol was added to a final volume of 16.5%, and for each gram of tissue 0.1 g CF-11 cellulose was added. The slurries were stirred for 30-60 minutes and poured into columns and drained completely. The columns were rinsed with 100 ml of 1X STE containing 16.5% ethanol. The column was rinsed with 0.1 ml of 1X STE, and the ds-RNAs were eluted using three 0.1 ml aliquots of 1X STE. The nucleic acids were precipitated with 2.5 volumes of absolute ethanol followed by centrifugation. The pellets were dried and then resuspended in sterile water.

The extracted products are run on 1% agarose gels using TAE 1X. The cDNA CFSD-S5 clone is run on the gels as the positive control. These are denatured in the gel by treatment with 0.05M NaOH and 0.15M, NaCl followed by neutralization in 0.1M Tris-HCl and 0.15M NaCl and transferred to nitrocellulose membranes using buffer 20X SSC. The labeling and detection was done with the Pierce chemiluminescent hybridization and detection kit with CSPD according to the manufacturer's instructions. The hybridization temperature is 42°C and highly stringent conditions are used to wash the filter (68°C), after which it was exposed to film for 15min to 1h.

Local cassava cultivars were this year collected from the principal growing regions by NARS partners in Tanzania. Germplasm collected from the Eastern zone, around Tanga, Kibaha and the coastal areas of Dar es Salaam were established at the Alawi estate in Kibaha. Collections from the south, Matwara, Lindi and Nachigwea districts were established at the Alawi estate in Kibaha, those from the Lake region around Geita, Musoma, Tarime, Muleba and Kasulu districts were established at ARI-Maruku. The experimental design of the trials was a random complete block design with 3 blocks and 10 plants per block. Farmers from the different regions will be invited during evaluations of these collections at harvest to determine the best local land races after their own criteria. Evaluations will be undertaken in September/October 2004. Selections will be planted in a crossing block early next year at the Alawi estate and ARI-Naliendele for genetic crossing.

Results
A total of 191 CR and 335 AR genotypes were successfully shipped to Tanzania. A description of the germplasm shipped and their parents are shown in Tables 12.1 and 12.2 More than 85% of all plants, and 100% of all genotypes were successfully established in the

screen house, a very high percentage of success. Molecular diagnostics carried out for all the introductions revealed that they were free of frog skin disease (Figure 12.2). Inspections by plant quarantine inspectors from TPRI also revealed an absence of pests and diseases in the plants growing in the screen house. The plantlets in the screen house were transferred to the field, the Alawi estate, in two batches, one set was moved in April 2004 and the second set was transferred in July 2004 (Figure 12.3). Some plants of 3 CR genotypes in the field showed some symptoms of purple/black discolorations (Figure 12.4) on the leaves but discussion with CIAT agronomist, Reinhardt Howeler and CLAYUCA agronomist Luis Fernando Cadavid revealed it might be a micronutrient deficiency due to boron or Iron. Application of liquid fertilizer with boron and zinc lead to the elimination of the trait.

Table 12.1. List of CR F_1 genotypes from with resistance to CGM and CMD resistance introduced to Tanzania this year

Code	Female Parent	Male Parent	No. of Genotypes
CR11	MCOL 2206	C-127	15
CR14	C-4	CM523-7	6
CR15	CM523-7	C-33	2
CR20	CM3306-4	C-33	5
CR21	CM3306-4	C-243	3
CR24	CM7951-5	C-18	3
CR25	CM7951-5	C-33	1
CR26	CM7951-5	C-39	1
CR27	CM7951-5	C-243	8
CR34	SM1741-1	C-18	5
CR35	SM1741-1	C-33	5
CR36	SM1741-1	C-39	2
CR37	C-4	CM4574-7	3
CR41	C-18	MCOL 2056	3
CR42	C-18	MCOL 2206	5
CR43	C-33	CM4574-7	13
CR44	C-39	CM3306-4	3
CR45	C-39	CM4574-7	9
CR46	C-39	SM1219-9	1
CR49	C-243	CM4574-7	6
CR51	C-243	OW280-1	5
CR52	C-243	SM1219-9	15
CR53	C-243	MCOL 2206	1
CR54	C-243	MTAI 8	8
CR55	MBRA 12	C-18	2
CR57	MCOL 2206	C-18	4
CR58	MMAL 66	C-18	2
CR59	MTAI 2	C-18	6
CR62	MTAI 8	C-39	3
CR8	C-4	MCOL 2206	13
CR9	C-4	MTAI 8	33
Total			**191**

Table 12.2. List of AR genotypes from BC_2 families that combine resistance to CGM and CMD resistance introduced to Tanzania this year

Code	Female Parent	Male Parent	No. of Genotypes
AR1-1	C-127	CW27-12	85
AR11-2	C-243	CW259-43	6
AR23-1	C-39	CW259-43	1
AR26-2	C-413	CW259-43	1
AR41-2	C-19	CW259-42	1
AR37-1	C-33	CW259-42	41
AR38-1	C-377	CW259-42	5
AR40-3	C-39	CW259-42	13
AR42-3	C-413	CW259-42	3
AR16-1	C-33	CW259-3	12
AR22-1	C-39	CW259-3	1
AR36-5	C-127	CW259-10	9
AR34-2	C-19	CW259-10	1
AR32-1	C-33	CW259-10	3
AR33-1	C-39	CW259-10	1
AR17-1	C-33	CW258-17	14
AR21-2	C-39	CW258-17	1
AR30-3	C-413	CW258-17	3
AR9-2	C-243	CW257-12	43
AR15-1	C-33	CW257-12	9
AR20-1	C-39	CW257-12	1
AR35-1	C-243	CW257-10	2
AR14-1	C-33	CW257-10	7
AR6-1	C-4	CW235-72	9
AR7-4	C-127	CW234-2	25
AR4-1	C-19	CW234-2	1
AR8-3	C-243	CW234-2	1
AR12-2	C-33	CW234-2	30
AR2-3	CW236-14	C-4	6
Total			**335**

The introductions will all be harvested in the March/April period and evaluated emphasis will be placed on high heritability traits like dry matter content, harvest index, plant architecture, and production of quality planting materials. About 60 genotypes will be selected for establishment in a crossing block for genetic crosses to local land races.

A total of 80 varieties were collected from the Eastern coastal region, 90 from the southern region and 120 from the Lake region. The cultivars from the Eastern and Southern region were established at the Alawi estate and have been evaluated for morphological characteristics. Collections from the Lake region were established at ARI-Maruku and are yet to be evaluated. Harvest at both sites will be conducted in March next year and the varieties evaluated for dry matter yield, harvest index, plant type, dry matter content, and culinary quality. Farmer groups will also be invited for the harvest to take into considerations their criteria. At least 90 genotypes will be selected for crossing to the improved introductions from CIAT. Two crossing blocks, using a polycross design, will be established at the Alawi estate and at ARI-Naliendele for genetic crosses.

Conclusions and Future Perspectives
One hundred and ninety one genotypes with resistance to the cassava mosaic disease (CMD) and 335 genotypes that combine resistance to CMD and to the cassava green mites (CGM) (derived from a wild relative), were shipped to Tanzania this year for the MAS breeding project. Molecular diagnostics of the introduced material for frog skin disease (FSD) revealed the absence of the disease and the introductions were transferred to the field in Tanzania Collection and evaluation of local varieties in Tanzania for crossing to the introductions were also carried out. Future activities include evaluation of the introductions and local varieties in March next year and genetic crosses between the two groups of germplasm.
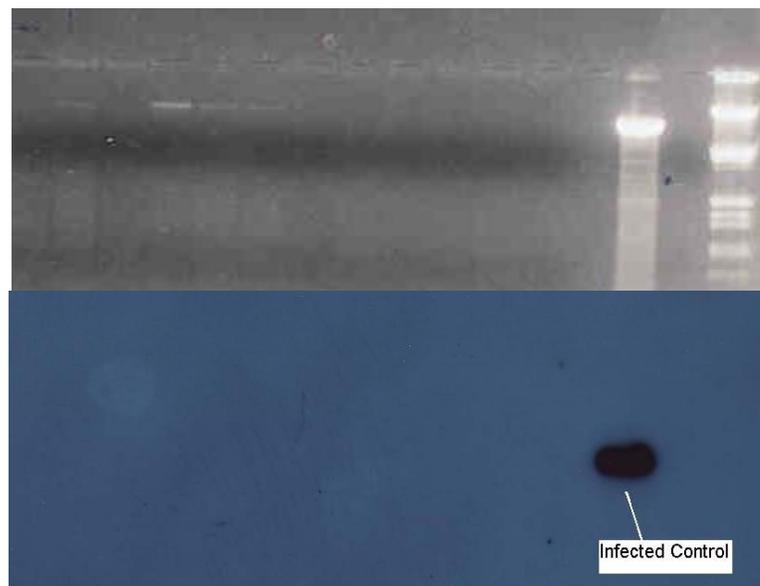
Figure 12.2. Agarose gel (upper picture) showing ethidium bromide stained double stranded RNA extraction from some AR plants (lanes 1 to 13), lanes 14 and 15 are extraction from frog skin disease (FSD) infected plants and the FSD cDNA probe (CFSD-S5). The lower picture is the Southern hybridization of CFSD-S5 hybridized to a Southern blot of the same gel.

Figure 12.3. CR (background) and AR (foreground) plants in the field at the Alawi estate in Tanzania. The CR plants are 6 months old while the AR plants are 1 month old.

Figure 12.4. Leaves from 3 genotypes in the field showing purple/black discolorations symptoms due to micronutrient deficiency.

References
ROCA, W.M., RODRÍGUEZ, J.A., MAFLA, G. Y ROA, J. (1984) Procedures for recovering cassava clones distributed *in vitro*. Centro Internacional de Agricultura Tropical (CIAT) 8 pp.

Activity 12.3. Genetic Mapping of Genes Involved in the Biosynthesis of Beta-carotene

Collaborators:
Ana-Maria Correa, Edgar Barrera, Wilson Castelblanco, Nelson Morante, Hernán Ceballos, Joe Tohme, Martin Fregene (CIAT)

Funding:
Harvest Plus Challenge Program

Important Output
1) Genetic mapping of the phytoene synthase gene using an S_1 mapping population (AM320) from the yellow variety MTAI8
2) There was no association between beta-carotene content and the phytoene synthase gene.

Rationale
The Harvest plus project in cassava seeks to improve, via conventional and genetic transformation methods, beta-carotene content in cassava and deploy these pro-vitamin A dense varieties in the fight against Vitamin A deficiency in the tropics. Naturally existing genetic variability for beta-carotene content in cassava is the basis for conventional improvement of beta-carotene content in cassava and knowledge of functional diversity provides for a more rational exploitation and faster progress in breeding. A study was initiated last year to identify markers and genes in the biosynthetic pathway associated with beta-carotene content as a first step to analysis of functional diversity and development of markers for conventional breeding. Three SSR markers, SSRY251, NS980, and SSRY240 were identified associated with high beta-carotene content in the S_1 family AM320 obtained from MTAI8, a yellow cassava variety. This year the study was extended to genetic mapping, and searching for associations, with pro-vitamin A content, of 2 known biosynthetic genes, phytoene synthase and phytoene destarurase.

Methodology
The mapping population AM320 comprised of 100 S_1 plants obtained from selfing MTAI8, an elite cassava cultivar developed by the CIAT-Thailand breeding program. This population is also being used for genetic mapping of cyanogenic glucosides and dry matter content, two traits that are high in MTAI8. Two cDNA clones each for phytoene synthase and 2 phytoene desaturase had earlier been obtained from a cDNA library of the cassava variety MNG2 (Andrea et al. unpublished data; CIAT 2002). Genetic mapping of the cDNA clones was as restriction fragment length polymorphism (RFLP). First, a parental survey of polymorphism was conducted using the restriction enzymes *Eco*RI, *Eco*RV, *Hae*III, *Hind*III, and *Dra*I. Parental survey filters were made using 10ug of cassava genomic from the MTAI8 parent and 4 S_1 progenies DNA digested with the enzymes mentioned above and separated on a 0.9% agarose gels as described earlier (Fregene et al. 1997). Progeny filters containing restricted DNA from the 100 S_1 plants, including DNA from the parent in the first lane, were prepared using the restriction enzyme that revealed polymorphism in the parental survey. The raw RFLP data was read as codominant markers and joined with 100SSR markers already evaluated in the S1 population. Linkage analysis and genetic mapping was as described earlier (Fregene et al. 1997) using a LOD score of 4.0 and a recombination fraction of 0.3.

Association between the markers and beta-carotene content, earlier evaluated in the S_1 cross was by single marker analysis using simple regression.

Results

Of the 2 genes used in the parental survey only phytoene synthase revealed 2 alleles in MTAI8 that segregated in the 4 S_1 progenies in the expected model with the restriction enzyme HindIII, phytoene desaturase was monomorphic (Figure 12.5). RFLP data from Progeny hybridization with the same enzyme permitted the mapping of phytoene synthase in a linkage group different from that with SSRY251 a cDNA-SSR marker that was earlier found to be associated with beta-carotene content in the same AM320 population. Incidentally SSRY251 shows very high homology to pyroxidine synthase, a gene known to be involved in the biosythesis of vitamin B6. Single point marker analysis by simple regression between the phytoene synthase gene, as independent variable, and beta-carotene content, as dependent variable revealed no association with the gene explaining 30% of phenotypic variance for the trait.
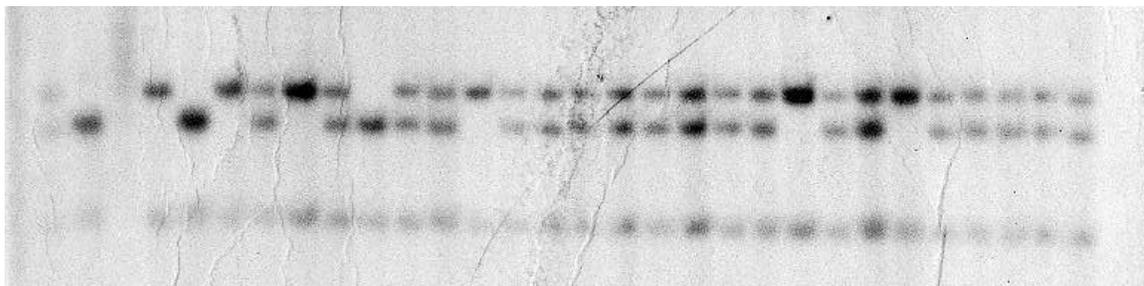
Figure 12.5. Southern hybridization of the phytoene synthase gene with *HindIII* digested DNA of progenies of AM320.

Conclusion and Perspectives

Genetic mapping of the phytoene synthase gene in an S_1 mapping population (AM320) from the yellow variety MTAI8 has been achieved. Mapping of another biosynthetic gene, phytoene desaturase, could not be achieved due to a lack of heterozygosity for this gene in MTAI8 with the five restriction enzymes employed. Current efforts are directed to search for polymorphisms using another panel of 5 restriction enzymes. The phytoene synthase gene was not associated with beta-carotene content in the AM320 cross and it explained 30% of phenotypic variance. The above results reveal that there are other genes that act to give the yellow color and need to be cloned for a complete understanding of the inheritance of beta-carotene content in cassava. Future activities include assessing SSR diversity of a collection of more than 200 yellow varieties and the combination of different alleles of the gene to assess the effect of combining different alleles of the gene.

References
CIAT 2003

Fregene, M.A., F. Angel, R. Gómez, F. Rodríguez, W. Roca, J. Tohme, and M. Bonierbale (1997). A molecular genetic map of cassava (*Manihot esculenta* Crantz). Theor. and Appl. Genet. TAG 95 (3) 431-441.

Activity 12.4. Progress in Genetic Mapping of Dry Matter Content (DMC) in Cassava

Collaborators:
Henry Ojulong, Nelson Morante, Jaime Marin, Edgar Barrera, Cesar Ospina, Martin Fregene (CIAT)

Funding:
Rockefeller Foundation, CIAT

Important Outputs:
1) Analysis of dry matter content (DMC) in 23 F_1 families from a diallel experiment over a 3-year period
2) Putative markers SSRY160 and SSRY150 found to be associated with dry matter content (DMC) in the GM313 family are also linked with the trait in other families
3) Discovery of a inter-specific hybrid family with a very wide segregation for dry matter content and initiation of bulked segregant analysis (BSA) for the identification of markers

Rationale
In 2002 a diallel experiment was initiated to provide information on the genetics of traits of agronomic interest in cassava (CIAT 2003). Based upon GCA estimates in the parents and high standard deviation of dry matter content (DMC), 23 families were selected for further evaluation of dry matter content (DMC). Two families, GM313 and GM312, of the 23 were also used for bulked segregant analysis of DMC and 2 SSR markers SSRY160 and SSR150 were found to be associated with the trait (CIAT 2003). The utility of these markers in the other 21 families was tested in the past year. Due to the small sizes of families in the original diallel experiment, 30-50 progenies, that are inadequate for QTL mapping, larger sized families for 9 of the pair-wise combinations of parents were also generated for QTL mapping. A seedling trial of more than 1,500 genotypes from these families is currently in the field this year. Also during the year, an inter-specific family CW208, a cross between MTAI8 and *Manihot tristis* was identified with a very wide segregation for DMC. This family is part of a large-scale evaluation of inter-specific between cassava and several wild *Manihot* species started in 2001. We describe here completion of evaluation of the 23 families of the diallel experiment, association of the SSR markers 160 and 150 with DMC in the other 21 families, and initiation of bulk segregant analysis (BSA) of CW208 using 600 SSR markers.

Methodology
Genotypes from 23 families selected based on large standard deviation for DMC in the 2002 harvest was planted in Santander de Quilichao in October 2002 and harvested in October 2003. The experiment consisted of six plants per genotype planted in a completely

randomized block design. Plants were harvested and bulked per genotype for measurement of percent dry matter content (DMC) using the standard CIAT procedure of weighing in air and water and calculating specific gravity. Plants were observed for incidence of frog skin disease (FSD) and rated as absent (0) or present (1), all genotypes showing any signs of frog skin infection were discarded. In November 2003 these families were re-planted in a trial with 4 x 5 plant plots and four replications in a randomized block design. Harvesting was done in August 2004, the central eight plants from each plot were harvested and bulked. Data was taken as described above, this time severity of frog skin disease (FSD) was rated on a 0-5 scale with 0 signifying no observed symptoms and 5 very severe. Data was subjected to ANOVA having as sources of variation genotypes and replication, a Generalized Linear Model (GLM) analysis was also carried out using the 3-year data, due to uneven number of observations across years; variation due to years was included as a source of variation.

Randomly amplified polymorphic DNA (RAPD) was employed in bulked segregant analysis in a search for more markers associated with DMC in the families GM313 and GM312. A total of 492 10-nucleotide random primers available at CIAT Cassava Genetics laboratory were evaluated in two sets of bulks, high and low for DMC respectively, from the families GM 312 and GM 313. RAPD analysis was as described earlier (CIAT 2003). PCR product was run on 1.5% agarose gels at 240 volts for about one hour, stained with ethidium bromide and photographed. Polymorphic markers in the bulks were analyzed in individuals of the bulks, and markers that remained polymorphic between individuals high and low in dry matter content were analyzed in all genotypes of the cross.

The SSR marker SSR 160 and SSR150 were also analyzed in the remaining 21 families, to assess the utility of these markers in different genetic background. DNA samples from all genotypes were extracted from 1g of oven dried (48h at 50°C) leaves using a mini-prep version of the Dellaporta *et al.*, (1983) protocol. Between 500μg to 1000μg of high quality DNA was obtained from each extraction and quantified using flourometer, the samples where then diluted to 10ng/μl for PCR amplification. SSR analysis was as described earlier by Mba et al. (2001). Bulk segregation analysis (BSA) of DMC in the inter-specific family CW208 was also carried out using SSR markers as described earlier (CIAT 2003).

Results

2003 harvest
During the 2003 harvest number of tubers per plant ranged from 0.30 to 31.00 in GM 283-8 and GM 267-5 respectively with an average of 9.49 (Table 12.3). They tended to be affected by FSD as infected plants had many undeveloped tubers. Fresh root yield greatly varied with the lowest, GM 306-21 having as low as 0.2 ton/ha and the highest CM 9642-26, 111.8 ton/ha. Percent dry matter content had values ranging from 16.09 in GM 284-23 and GM 309-9 and as high as 51.07% and 69.07% in GM 313-23 and GM 311-15 respectively, while dry yield values ranged from 0.08 ton/ha in GM 306-21 and GM 309-39 to 39.26 t/ha in CM 9642-26.

High incidence of frog skin disease was observed with 23% of the genotypes showing symptoms, simple correlation analysis in infected genotypes revealed significant negative estimates with all the yield parameters. Highest estimate was obtained with percentage dry matter (-0.33) followed by harvest index (-0.24), dry tuber yield and number of commercial roots (-0.23), fresh tuber yield and number of tubers (-0.20). Correlation between the 2002

and 2003 percent dry matter estimates revealed a low positive value of 0.12 suggesting that frog skin disease pressure had affected evaluations in 2003.

Table 12.3. Simple statistics of yield components estimated from clones harvested in October 2003 in Quilichao.

	Minimum	Maximum	Average	Std Deviation[a]
Stand count	1	6	4.32	1.73
FSD[b]	0.00	1.00	0.23	0.42
Foliage weight[c]	0.10	62.70	10.19	7.86
ComRt[d]	0.00	65.00	17.23	13.58
TbNo[e]	0.30	31.00	9.49	4.63
Harvest Index	0.02	0.92	0.51	0.16
Fyield[f]	0.20	111.80	21.96	16.74
DM[g]	16.30	69.07	31.10	5.06
Dyield[h]	0.08	39.26	7.06	5.73

[a]standard deviation; [b]Frog skin disease, 0=absent, 1=present; [c]Vegetative yield per plot(kg); [d]Number of commercial roots per plot; [e]Number of tubers per plant; [f]Fresh tuber yield (t/ha); [g]Percent dry matter content; [h]Dry tuber yield (t/ha)

2004 harvest
A high incidence of frog skin disease was again observed in the 2004 harvest, 36.9% of the genotypes showed symptoms although with low severity, most affected plants showed average severity of 1 or less. The low level of severity could be a result of discarding infected plants from the previous year therefore avoiding inoculum buildup. Fresh and dry root yield was significantly lower in the 2004 harvest (Table 12.4). Analysis of variance (ANOVA) showed differences amongst genotypes to be highly significant (P=0.001) for all the yield parameters (Table 12.5). Replication was highly significant (P=0.001) for all the parameters except DMC, which was significant at 0.05, indicating that DM is the most stable of yield parameters, this is supported by the low CV value (11.28%). GLM was performed on the three sets of data (2002, 2003 and 2004) and results are shown in Table 12.6. Genotype and year were highly significant (P=0.001) for dry matter yield, most likely due to the variable climatic conditions in Santander de Quilichao. Percent dry matter showed the lowest coefficient of variation (9.65%) implying that it is stable across years and that a single year data is sufficient for evaluation of DMC.

A total of 70 primers were polymorphic in one or both of the bulks (16 were polymorphic in both). When run on open bulks most of the primers were false positives. However, eight primers:- AB15, C18, H09, K10, O01, O14, H09, AH09 and AO14 continued to be polymorphic (Figure 12.6) and were tested on the whole populations. Simple regression of DM phenotypic data on the RAPD marker genotype classes produced very low values regression coefficients, 1.05 to 1.15%, eliminating the utility of these RAPD markers.

Table 12.4. Simple statistics of yield components estimated from clones harvested in August 2004 in Santander de Quilichao

	Minimum	Maximum	Average	Std Deviation[a]
PltHa[b]	1.00	8.00	4.9	1.66
FSD[c]	0.00	3.00	0.3	0.53
ComRt[d]	0.00	30.50	8.2	5.59
Tbplt[e]	0.29	12.00	3.5	1.63
FolWt	0.08	13.90	4.75	2.56
HI[f]	7.89	90.74	41.1	13.24
Fyield[g]	0.23	25.06	5.8	3.95
DM[h]	18.66	40.40	30.2	3.51
DYield[i]	0.04	8.82	1.8	1.32

[a]Standard deviation [b]Number of plants harvested [c]Frog skin disease scores on a scale of 1-5 (1 no symptoms detected, 5 very severe), [d]Number of commercial roots per plot, [e]Number of roots per plant, [f]Harvest index, [g]Fresh tuberous root yield (tons/ha), [h]Percent dry matter content, [i]Dry root yield (tons/ha).

Evaluation of the markers SSRY160 and SSRY150, earlier observed to be associated with DMC in the family GM313, in the other 21 families of the diallel experiment revealed association with the trait in several other families, for example all families that have SM1741-1 as parent showed a strong association and high regression coefficients between SSR marker 160 and DMC (Table 12.7). This suggests that this marker is associated with a favorable allele for DMC found in a specific genetic background. Efforts are now directed to evaluating this marker in larger sized families having SM1741-1 as one of the parents generated last year and planted in the field this year to obtain a more accurate value of regression coefficient in preparation of its use in marker assisted selection (MAS) of DMC.

Table 12.5. Analysis of variance (ANOVA) table of yield parameters evaluated at harvest in CIAT, Quilichao in August 2004.

Source of Variation		M	E	A	N	S	Q	U	A	R	E	S
	df[a]	FSD[b]	TbNo[c]	ComRt[d]	TbPlt[e]	HI[f]		Fyield[g]		DM[h]		Dyield[i]
Genotype	430	0.75	228.77	88.39	6.50	418.75		45.92		32.51		5.05
		(P<.0001)	(P<.0001)	(P<.0001)	(P<.0001)	(P<.0001)		(P<.0001)		(P<.0001)		(P<.0001)
Replication	3	1.86	7947.54	2313.45	60.92	1909.87		783.54		3.43		71.03
		(P<.001)	(P<.0001)	(P<.0001)	(P<.0001)	(P<.0001)		(P<.0001)		(P<.05)		(P<.0001)
Error	757	0.39	106.08	37.88	4.09	132.63		19.76		11.58		2.12
CV[j]			61.3	74.9	57.56	28.02		75.83		11.28		79.7

[a]Degree of freedom; [b]Frog skin disease; [c]Number of tubers per plot [d]Number of commercial roots per plot [e]Number of tubers per plant; [f]Harvest index [g]Fresh tuberous root yield (t/ha); [h]Percent dry matter content, [i]Dry tuberous root yield; [f]Coefficient of variation

Table 12.6. Analysis of variance (ANOVA) table of yield parameters taken over 3 years at CIAT, Quilichao.

Source of Variation	MEAN SQUARES					
	df	FSD	HI	DM	Fyield	Dyield
Genotype	345	0.11	0.02 (P<.0001)	19.70 (P<.0001)	336.42 (P<.0001)	46.21 (P<.0001)
Year	2	9.74 (P<.0001)	1.90 (P<.0001)	2795.22 (P<.0001)	165359.37 (P<.0001)	22200.14 (P<.0001)
Error	428	0.11	0.01	9.78	255.65	32.81
CV (%)		276.20	20.98	9.65	64.6	67.35

Segregation of dry matter content in the inter-specific family CW208 is maybe the highest found to date. Distribution of the trait in this family reveals a non-normal distribution that suggests large QTLs genes might be involved (Figure 12.7). Bulked segregant analysis with 600SSR markers is ongoing and should lead to identification of the responsible genes.
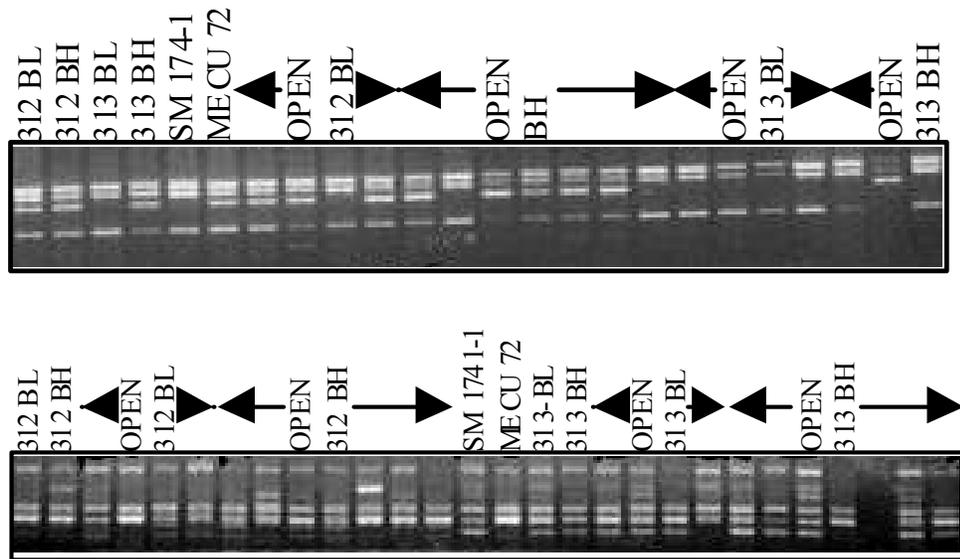
Figure 12.6.-Ethidium strained agarose gel showing PCR amplification of two RAPD primers AB 15 and K10 of parents, bulks and individuals constituting the respective bulks of family GM313 and GM312. BL and BH signify bulk high and bulk low respectively.

Fig 1 Distribution of percent dry matter in the wild cross CW 208
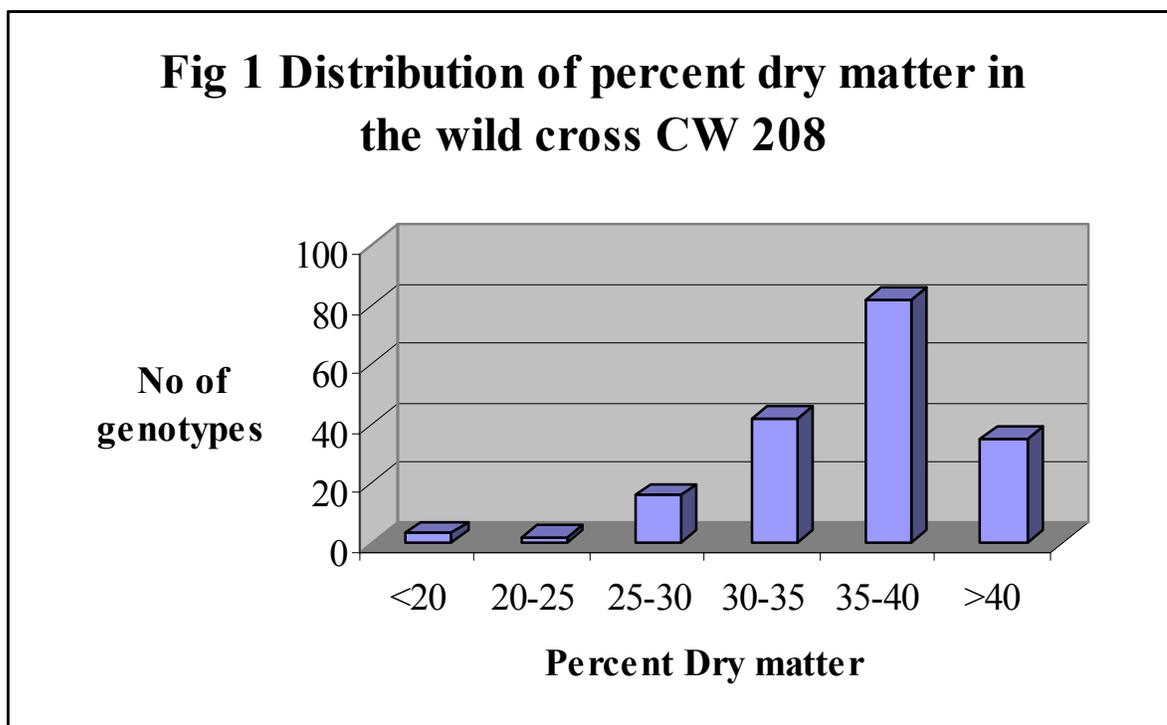
Figure 12.7. Frequency distribution of different classes of dry matter content in an inter-specific family CW208 obtained from a cross between MTAI8 and *M. tristis*

Conclusion and Future Perspectives

Analysis of dry matter content (DMC) in 23 F_1 families from a diallel experiment over a 3-year period revealed the profound effect of different seasons (years) on DMC, the effect of replication was of lesser importance. Putative markers SSRY160 and SSRY150 found earlier to be associated with dry matter content (DMC) in the GM313 family was also linked to the trait in other families. Discovery of markers for DMC have also been extended to a inter-specific hybrid family with a very wide segregation for dry matter content. Future perspectives are completion of bulked segregant analysis (BSA) in the inter-specific hybrid cross and evaluation of markers identified until now in larger sized families generated for QTL mapping.

Table 12.7. Simple regression coefficients of DMC of the SSR markers SSR150 and SSRY160 in 21 families obtained from a diallel experiment

Family	Mother	Father	R² (SSRY 150)	R² (SSRY 160)
GM 257	SM 1219- 9	SM 1636- 24	17.2	5.3
GM260	SM 1219- 9	SM 1673- 10	ND	11.98
GM265	SM 1219- 9	MPER 183	4.8	0.5
GM268	SM 1278- 2	SM1673- 10	4.2	12.6
GM269	SM 1278- 2	SM 1741- 1	3.2	25.7
GM283	SM 1636- 24	SM 1673- 10	1.5	ND
GM284	SM 1636- 24	SM 1741- 1	9.8	28.89
GM285	SM 1636- 24	HMC 1	ND	0.56
GM286	SM 1636- 24	MPER 183	12.08	11.72
GM293	SM 1673- 10	HMC 1	0.3	2.77
GM294	SM 1673- 10	MPER 183	0.1	5.3
GM306	MECU 72	MPER 183	0	8.07
GM309	MECU 72	SM 1219- 9	5.98	2.25
GM310	MECU 72	SM 1278- 2	0.8	11.36
GM313	MECU 72	SM 1741- 1	18.1	29.3
GM311	MECU 72	SM 1636- 24	0.6	0.4
GM314	MECU 72	HMC 1	0.4	10.23
CM9642	CM 6740- 7	MPER 183	6.29	1.84
CM977	HMC 1	MPER 183	36.56	0.46
CM9901	CM 6740- 7	SM 1219- 9	7.69	ND

ND: Not determined

References

CIAT, 2003

Dellaporta, S. L., J. Wood, J. B. Hicks. 1983. A plant DNA minipreparation: Version II. Plant Molecular Biology Reporter 1 (14): 19-21.

Mba R. E.C., P. Stephenson, K. Edwards, S. Melzer, J. Nkumbira, U. Gullberg, K. Apel, M. Gale, J. Tohme, and M. Fregene (2001) Simple Sequence Repeat (SSR) Markers Survey of the Cassava (*Manihot esculenta* Crantz) Genome: Towards a SSR-Based Molecular Genetic Map of Cassava. Theor and Appl Genet 102:21-31

Activity 12.5. QTL Mapping of Cyanogenic Glucoside Content in a S₁ Population derived from MTAI8 and Candidate Gene Mapping of Two Cytochrome P-450 Biosynthetic Genes (CYP79D1 Y CYP79D2)

Collaborators:
Ana Maria Correa Morales, Jaime Marín, Edgar Barrera, Cesar Ospina, Nelson Morante, Teresa Sánchez, Martin Fregene (CIAT), Elizabeth Kizito, Urban Gullberg (SLU, Uppsala, Sweden)

Important Outputs
1) Development of a partial map of the cassava genome using SSR and DArT markers and a S_1 population derived from MTAI8 and genetic mapping of a cytochrome P450 gene
2) Establishment of a replicated trial of 200 S_1 progenies to measure cyanogenic potential in leaves and roots.

Rationale:
The presence of cyanogenic glucosides in cassava roots is a nutritional deficiency and a potential health problem for human and animal consumers. There has therefore been a considerable amount of interest in understanding the biosynthesis of the two cyanogenic glucosides, Linamarin and Lotaustralin, produced in cassava and ways of reducing or eliminating them all together in the roots. In 2000, the enzyme that catalyzes the rate limiting the most important step in the biosynthesis of the cyanogenic glucosides, the conversion of amino acids to oxime, was cloned and identified as a cytochrome P-450 gene, 2 cDNAs (CYPD1 and D2) with about 80% homology were identified (Anderson et al 2000). In collaboration with the group of Prof Moller that cloned the CYP genes and a doctoral student from the Swedish Agricultural University (SLU), Uppsala, an attempt was made to associate the genes with cyanogenic glucoside content and also identify other QTLs controlling the trait in cassava. We describe here genetic mapping of the CYP genes, as RFLP markers, 70 SSR and 150 Diversity Array Technology (DArT) markers in an S_1 family derived from MTAI8, a Thai variety with high cyanogenic glucoside content. We also report a field experiment to measure the trait in the S_1 family. The identification of markers associated with cyanogenic potential (CNP) in cassava will provide tools to accurately identify the trait in an effort to breed for low CNP cassava varieties.

Methodology
The S_1 family (AM320) consisted initially of 104 individuals but was increased to 200 from new selfs made with MTAI8 this year. DNA was isolated from all 200 genotypes using a mini-prep method of the Dellaporta extraction protocol (1983). SSR markers for genetic mapping were 600 SSR markers developed earlier in cassava, they were screened in the MTAI8 parent and 5 other S_1 progenies as described earlier by Mba et al (2001). A previously constructed DArT chip of about 1000 polymorphic markers (Liu et al. 2004) was the source of DArT markers for evaluating the AM320 population. The cytochrome P-450 genes CYPD1 and D2 were evaluated in the MTAI8 parent along with 4 S_1 progenies for restriction length polymorphisms (RFLPs) using the following restriction enzymes: *EcoRI, EcoRV, HindIII, HaeIII,* and *DraI*. Preparation of parental and progeny filters, and Southern hybridization of the filters were as described by Fregene et al. (1997). Polymorphic SSR and RFLP markers were

evaluated in the entire S_1 progeny. A chi-square test at a confidence level of 0.05 and 0.01 was used to test goodness of fit of the segregation date with the expected model of 1:2:1 ratio for co-dominant markers and 3:1 ratio for dominant markers. Linkage analysis of the raw segregation marker data was done using the Mapmaker linkage analysis software (Lander et al. 1993) and a LOD of 5.0 and recombination fraction (theta) of 0.3 for the dominant markers and a LOD of 9.0 and theta of 0.2 for the dominant markers. Map distances were calculated by the Kosambi method that takes into account double-crossovers. Initial linkage analysis was carried out with the co-dominant SSR and RFLP markers combined with the dominant DArT markers, these were later separated due to difficulties in placing the DArT markers, separate maps were therefore developed.

A preliminary evaluation of CNP in leaf tissue and roots was conducted in the AM320 S_1 family last year based upon 3 plants and a single replication. Evaluations were conducted on a single root and leaf tissue harvested from each of the plants and CNP determined according to the enzymatic protocol developed by Cooke (1978) and modified by O'Brien (1991). This year 200 S_1 progenies of the AM320 family were re-established in single plant plots replicated eight times in CIAT, Palmira, for evaluation of the CNP phenotype. The trial will be harvested piece-meal at 4, 6 and 10 months after planting to measure the accumulation of CNP over a range of growth period.
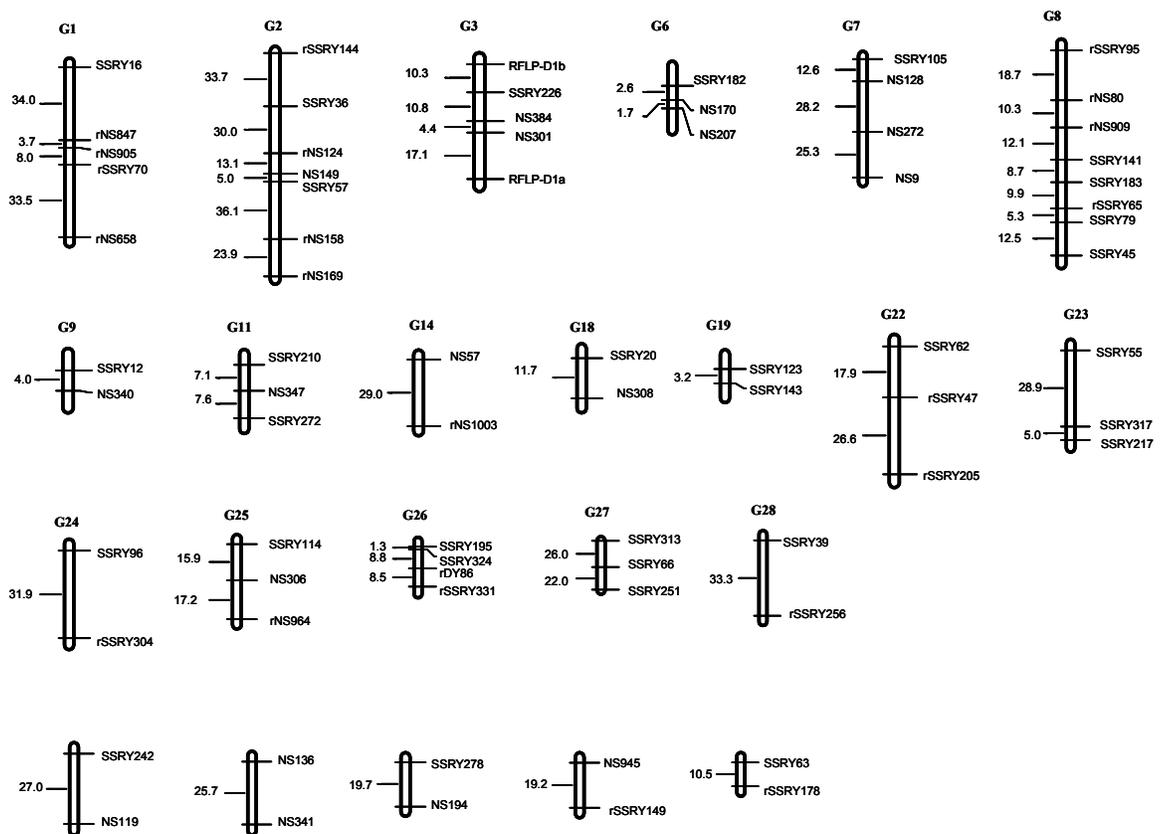
Figure 12.8. A SSR and RFLP genetic map of the AM320S1 family

Results

So far, a total of 100 SSR markers have been found to segregate in the AM320 S_1 family, while 208 polymorphic DArT markers were found. Less than 20 of the SSR markers and more than 90% of the DArT showed segregation distortion at a 0.05% confidence level in the Chi square test. Seventy-four of these SSR markers were organized into 23 linkage groups that covered 819.5cM of the cassava genome by linkage analysis while 26 markers remained unlinked. The 2 cytochrome P-450 gene CYPD1 was polymorphic with the restriction enzyme *Eco*RI used in the RFLP parental survey but CYPD1 was monomorphic with the 5 enzymes used (Figure 12.8). RFLP segregation data for CYPD1 revealed possible duplicated loci, one segregating at the expected ration of 1:2:1, at a 0.05% confidence level in the Chi square test, and the second at a ratio of 3:1. Linkage analysis permitted the mapping of both loci to linkage group 3. Efforts are ongoing to use many more restriction enzymes to look for RFLPs with CYPD2 to enable genetic mapping of this gene.

Preliminary results of cyanogenic glucoside content in the AM320 family revealed a wide segregation for the trait and the appropriateness of this family for mapping the trait. In leaves, 5% of the progenies had below 1075ppm, 85% had a range of 1075 -3048, while 10% had between 3049 -5071ppm. In the roots, 11% of the family had below 258ppm, 76% had between 259 and 878ppm, while 13% had higher than 1294ppm. The distribution of the trait in both leaves and roots was normal suggesting a quantitative trait. However, the above data is based upon a single replication and cannot be said to be an accurate representation of CNP in these genotypes.
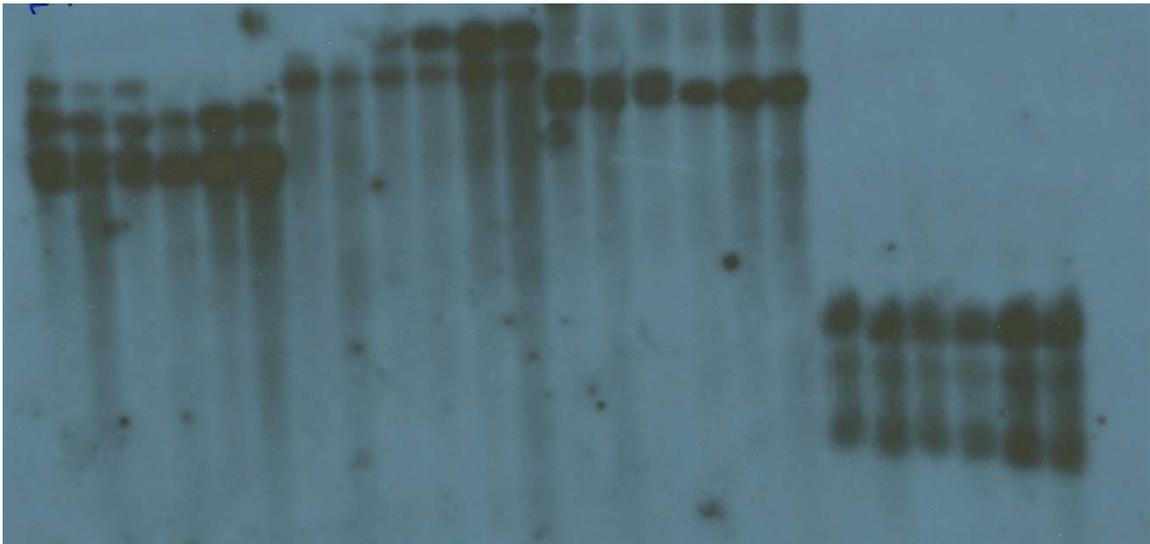
Figure 12.9. Southern hybridization of the cyanogenic glucoside biosynthetic gene CYP79D1 with MTAI 8 and 4 S_1 progenies with the restriction enzymes, from left to right, *EcoRI, EcoRV, HindIII*, and *HaeIII*.

Conclusions and Perspectives

A partial molecular genetic map of cassava has been constructed using SSR and RFLP markers and the cytochrome P-450 gene CYPD1, in the S_1 family AM320 derived from MTAI8. Preliminary evaluation of cyanogenic glucoside content in this family revealed wide segregation of the trait. A proper evaluation of the trait is being carried out this year in preparation for association of cyanogenic glucosides content with the biosynthetic genes and QTL mapping. Work is also ongoing to identify RFLPs for the second gene, CYPD2 so that it can also be placed on the genetic map.

References

ANDERSEN, M.D., P. K. BUSK, I. SVENDSEN, B. L. MØLLER. 2000. Cytochrome P-450 from Cassava (*Manihot esculenta* Crantz) Catalyzing the First Steps in the Biosynthesis of the Cyanogenic Glucosides Linamarin and Lotaustralin. The Journal of Biological Chemistry 275 (3): 1966-1975.

COOKE, R. D. 1978. An enzymatic assay for the total cyanide content of cassava (*Manihot esculenta* Crantz). J. Sci. Food. Agric. 29: 345-352.

DELLAPORTA, S. L., J. WOOD, J. B. HICKS. 1983. A plant DNA minipreparation: Version II. Plant Molecular Biology Reporter 1 (14): 19-21.

O'BRIEN, G. M., A. J. TAYLOR, N. H. POULTER. 1991. Improved enzymatic assay for cyanogens in fresh and processed cassava. J. Sci. Food Agric. 56: 277-289.

Activity 12.6. Development of Mapping Populations for Gene Tagging of Post Harvest Physiological Deterioration (PPD), Resistance to Hornworms and Whiteflies Found in Wild Relatives of Cassava

Collaborators:

A. Lopez, N.Morante, O.Akinbo, H.Ceballos, A. Bellotti, M. Fregene.

Funding:

CIAT

Important Outputs

1) Development of 8 populations for QTL mapping of post-harvest deterioration (PHD), resistance to whiteflies and hornworm.

Rationale

Post harvest physiological deterioration (PPD) and arthropod pests are severe marketing and production constraints respectively in cassava. It has been estimated that cassava farmers, typically resource-poor farmers, lose 48 million tons of fresh root valued at US$1.4billion every year to pests, diseases, and PPD; some 30% of total world production. Wild relatives of cassava are important sources of genes for resistance to pests and diseases and longer shelf life. The only source of dramatically delayed PPD has been identified in an inter-specific hybrid between cassava and *Manihot walkerae* (Sanchez et al. 2003, unpublished data). The delayed PPD trait, originally from the wild *Manihot* parent, was successfully transferred to an F_1 inter-

specific hybrid suggesting a dominant or additive gene action of gene(s) involved. The only source of resistance to the cassava hornworm and a widely deployed source of resistance to the cassava mosaic disease (CMD) were identified in 4th backcross derivatives of *M. glaziovii* (Chavariagga et al. 2004). Moderate to high levels of resistance to white flies have been found in inter-specific hybrids of *M. esculenta* sub spp *flabellifolia* (CIAT, unpublished data). Again, resistance was recovered easily in F_1 inter-specific hybrids, suggesting a simple inheritance of the trait. For several years now molecular marker tools and a modified Advanced Back Cross QTL (ABC-QTL) scheme have been tested for cost-effective pyramiding of useful genes from cultivated and wild gene pool through the elimination of phenotypic evaluations in each breeding cycle. We describe here the development of populations for QTL mapping of post-harvest deterioration (PHD), resistance to whiteflies and hornworm.

Methodology

Segregating populations for the identification of molecular markers for the introgression of delayed PPD, resistance to the cassava hornworm and white flies (presently as sexual seeds) include BC_1 as well as S_1 families to enable identification of recessive genes. The inter-specific hybrid from *Manihot walkerae* with delayed PHD, CW429-1, was crossed extensively to the elite cassava genotypes MTAI8, CM523-7, and SM909-25 to create 3 BC_1 families (BC_1 only in the sense of crosses to cassava). This genotype, CW429-1, was also selfed to generate an S_1 family. The variety MNG11, a BC_4 derivative of *M.glaziovii* with cassava as recurrent parent, having resistance to the hornworm was also crossed to MTAI 8 and selfed to produce BC_1 and S_1 families respectively. The inter-specific hybrid CW251-3, a progeny of *M.esculenta* sub spp flabellifolia (OW189-1) and a high beta-carotene cassava land race CM1734, showing a high level of resistance to white flies was crossed to MTAI 8 and selfed to produce BC_1 and S_1 families respectively. All the above-mentioned crosses were done in the 2003-2004 season.

Table 12.8. List of crosses made to date for development of populations for QTL mapping pf PDH, resistance to whiteflies and hornworm

Female	Male	Seeds available or expected
PHD		
MTAI 8	429 - 1	75
429 -1	MTAI- 8	127
CW 429 -1	SM 909-25	157
CW 429 -1	CM 523-7	143
429 -1	429 - 1	270
White flies		
CW 251-3	MTAI- 8	120
MTAI- 8	CW 251-3	43
CW 251-3	CW 251-3	165
Horn worm		
MNIG19	TAI - 8	87
TAI - 8	MNIG19	136
MNIG19	MNIG19	124

Results
Between 50 and 150 crosses per cross combinations have been made for the development of BC_1 and S_1 gene tagging populations for PHD, resistance to whiteflies and hornworm (Table 12.8). Sexual seeds will be harvested later in the season in preparation for *in vitro* establishment next year. At least 200 progenies, including reciprocals, of each BC_1 and S_1 populations will be established *in vitro* from embryo axes and multiplied to obtain 10 plants per genotype. The tissue culture plantlets will be transferred to the screen house and then to the field as a single row trial (SRT) of ten plants.

The following year progenies will be re-established in a QTL mapping trial of single row plots of 8 plant with 5 replications in one location. Great efforts will be made to ensure that the trials are kept free of weeds, pests, diseases, and nutritional deficiencies to minimize environmental variation.

Conclusions and Future Perspectives
Eight populations have been developed for QTL mapping of post-harvest deterioration (PHD), resistance to whiteflies and hornworm. The segregating populations will be established early next year from embryo axes and multiplied in preparation for field evaluations of PHD and green house evaluations of hornworm and whiteflies resistance. Based on the results of the phenotypic evaluations, bulks of extreme phenotypes will be made for bulk segregant analysis (BSA) with 600 SSR and RAPD markers as described earlier. Polymorphic markers will be evaluated in individuals of the segregating populations and strength of association measured by simple regression. Should BSA fail to identify markers, then a standard QTL procedure, including development of a genetic map with SSR markers, will be followed.

References
Chavarriaga P.; S. Prieto; C.J. Herrera; D. López; A. Bellotti[1]; J. Tohme (2004). Screening transgenics unveils apparent resistance to hornworm (*E. ello*) in the non-transgenic, African cassava clone 60444. In: Alves and Tohme. Adding Value to a Small-Farmer Crop: Proceedings of the Sixth International Scientific Meeting of the Cassava Biotechnology Network. 8-14 March 2004, CIAT, Cali Colombia. Book of Abstracts Pp4.

Activity 12.**7 Generation Challenge Program: Comparison of Simple Sequence Repeats (SSR) and Diversity Array Technology (DArT) Markers for Structural Characterization of Diversity in Cassava**

Collaborators:
P. Hurtado, C. Buitrago, C. Ospina, J. Marin, C. De Vicente, M. Fregene (CIAT) (IPGRI), Prapit Wongtiem (Field Crop Research Station, Rayong, Thailand), Andrzej Killian, P. Wenzel (CAMBIA, Canberra, Australia)

Important Outputs
1) Structural characterization of genetic diversity in cassava with 251 polymorphic dominant DArT markers compared to that with 36 SSR co-dominant markers
2) A clear trade-off between number of loci and amount of information provided by each locus

Rationale
At the heart of the Generation Challenge Programme (GCP) is a vision to harness advances in molecular biology and the rich natural variation found in crop genetic resources to create a new generation of hardy crops for small farmers. Characterizing structural and functional diversity of 11 mandate crops: Barley, Maize, Rice, Sorghum, Wheat, Chickpea, Cowpea, Common Bean, Cassava, Potato and *Musa*, is the entry point of the GCP. SP1 is the subprogram in charge of ensuring a scientifically sound scheme to put germplasm collections to work for the discovery of new genes and alleles that will contribute to solve the important challenges of modern agriculture. By examining the genetic structure of a large and representative sample of a collection revealed with molecular markers, SP1 proposes to re-sample the original germplasm and select a subset that will be subject to fine phenotyping and association studies.

For many reasons up to date, the markers of choice for germplasm characterization have been microsatellites (SSR). SSR are abundant in most genomes, highly polymorphic and easily assayed. SSR marker mutations are formed by slipped strand mis-pairings. A newer marker tool, Diversity Array Technology (DArT) is a DNA hybridization-based system based on single nucleotide polymorphisms, insertion-deletions and DNA methylation changes. DArT offers the highest throughput available up to date at a fraction of the cost of SSR markers and allows for whole genome scanning in a speedy manner.

A pilot experiment was designed to test the usefulness of DArT markers as an alternative to SSR for detecting structural variation in a more cost-effective way. A randomly selected set of 436 accessions of cassava (*Manihot esculenta* Crantz) were analyzed with DArT and SSR markers and results compared.   The hypothesis is whether DArT markers are more adept at uncovering genetic diversity structure.

Methodology
Plant material for the pilot study included accessions from the International Institute of Tropical Agriculture (IITA) in Nigeria, principally local varieties from West Africa and elite IITA varieties, and from CIAT, selected at random from South and Central American varieties held at the world germplasm collection. DArT analysis was conducted at the Center for the Application of Molecular Biology to Agriculture (CAMBIA), Canberra, Australia in collaboration with IPGRI and CIAT. A National Program Scientist, Ms Prapit Wongtiem, of the Field Crop Research Station, Rayong, Thailand, participated in the DArT analysis. A cassava DArT array of approximately 1000 polymorphic clones constructed earlier (Ling et al. 2004) was the source of markers. The cassava DArT chip was developed using DNA from accessions originating from 17 different countries, improved varieties from CIAT and IITA, and wild accessions of *Manihot esculenta* sub spp flabellifolia, *M. carthaginensis* and *M. walkerae*. At

CIAT, the same accessions were analyzed with 36 microsatellites selected from 18 linkage groups of the cassava genetic map. Data analysis was conducted at CIAT. Cluster analysis of the DArT and SSR data were performed using principal coordinate analysis of a similarity matrix derived by the Jaccard method using NTSYS-PC (Rohlf 1993).

Results

A total of 251 loci were sampled with DArT markers compared to 36 loci for SSR markers. Results of cluster analysis by PCoA gave similar outcomes of distinct clusters, but there was a clear separation of the Latin American and African accessions with SSR markers (Figure 12.10). Three clusters common in both markers were a group of genotypes from Guatemala, a sub-set of accessions from Nigeria, and a third large conglomeration of genotypes from the rest of the world, a total of 20 countries. SSR markers separated this third clusters into two according to geographic origin of the germplasm. These results agree with a previous attempt to elucidate the organization of genetic diversity in cassava using 67 SSR markers, that study revealed a high level of genetic differentiation between a group of genotypes from Guatemala and a separation between African and Latin American accessions.

A   **B**
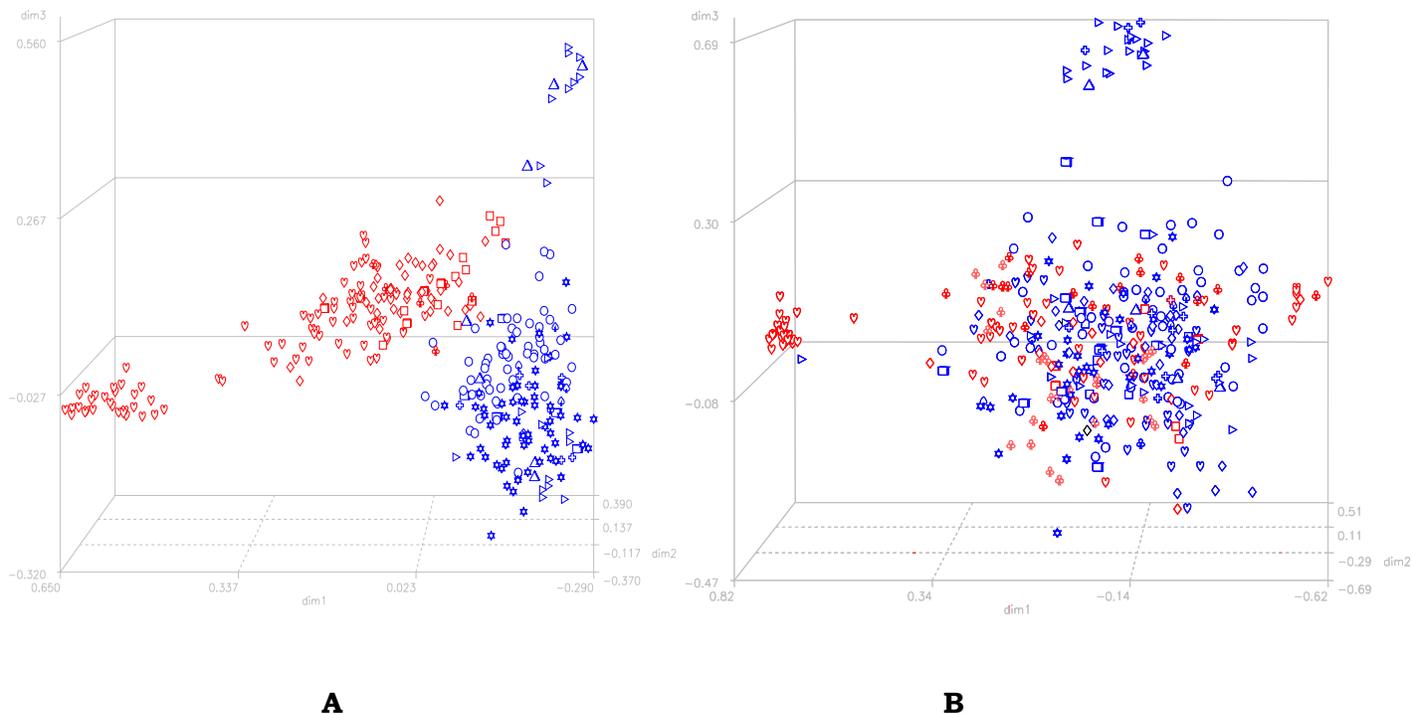
Figure 12.10. Principal coordinates analysis (PcoA) of SSR markers (A) and DArT markers (B). The African accessions are represented in red while the Latin American accessions are in blue color. Three distinct groups can be seen with both markers but there a separation of the Latin American and African accessions is evident with SSR markers.

Possible sources of the observed structure could be founder effects (geographic dispersal to the old world), selection (especially for diseases prevalent in Africa), small effective sample sizes (as in the case of spread of cassava to Africa from Latin America), migration (introgression from wild relatives), independent domestication events especially for the Guatemalan accessions, and mutations. The possibilities of introgression from wild relatives into accessions from Guatemala is quite high, the geographical origin of these Guatemalan accessions overlaps with that of 2 *Manihot* species unique to Central America. It is also remotely possible that the accessions from Guatemala represent a second center of domestication, similar to other crops like common beans (*Phaseolus vulgaris*) and pepper (*Capsicum* spp.) that were independently domesticated in Central and South America. The separation between the African and Latin American accessions could be due to selection and/or small effective sample sizes, suggesting that Africa could yet benefit from introgression of germplasm from Latin America. The cluster made up of some Nigerian accessions was also observed in an earlier study (Raji 2002, unpublished data), these accessions are from the Northern part of the country and it is not clear if this is due to small effective sample sizes and selection for tolerance to drought prevalent in the Northern part of the country, this again suggests a need to broaden the germplasm base in this part of the country.

A large number of alleles were detected by each SSR loci (an average of 10 alleles per locus) compared to DArT markers (2 per loci) although DArT markers sampled many more loci (251) of the genome compared to SSR (36) in this study. This suggests a trade-off between information and number of loci. Because DArT loci can be significantly increased up to 1000 with minimal additional costs for molecular characterization it can be speculated that the level of resolution obtained with DArT could be increased to a similar level as SSR markers. Furthermore, given that the investment (labor and consumables) for development of both marker systems is approximately similar, even if both technologies provided a similar level of resolution, DArT would appear as an attractive marker alternative. This is so because of the cost per assay (lots of data points in one single assay vs two data points per assay with SSR), which would make DArT especially interesting for orphan crop species that do not count on existing marker systems.

Conclusion and perspectives
In conclusion, diversity estimated with 36 co-dominant SSR markers is more efficient than 251 DArT dominant markers. These results reveal a trade-off between amount of information and number of loci provided by each locus, in this study DArT sampled a larger number of loci of the genome, but they are dominant markers and consequently have less information compared to co-dominant SSR markers. But the hypothesis that DArT markers are more useful than SSR markers in detecting structural variation cannot be accepted, more conclusive evidence will have to await analysis using a denser DArT array and a larger data set of accessions from each country and region.

References
Ling Xia, Kaiman Peng, Yang Shiying, Peter Wenzl, M Carmen de Vicente, Martin Fregene, Andrzej Kilian (2004) DArT for High-Throughput Genotyping of Cassava (*Manihot esculenta*) and its Wild Relatives. Theor Appl Genet (in review)

Activity 12.8 Generation Challenge Program: Assembling Germplasm and Molecular Markers Sets for Analysis of Structural Diversity in Cassava

Collaborators:
Paula Hurtado, C. Buitrago ,Cesar Ospina, Jaime Marín, Paola Alfonso, Graciela Mafla, Alfredo Alves, Daniel Debouck, Hernán Ceballos, Joe Tohme, Martin Fregene (CIAT)

Funding:
Generation Challenge Program

Important Outputs
1) Selection of a set of 3000 cassava accessions for structural diversity characterization using SSR markers
2) Selection of 36 SSR markers for molecular analysis of the germplasm set.

Rationale
The objective of sub-programme 1 of the GCP is the selection of a representative sub-set of germplasm and the molecular analysis of structural diversity to identify population structures as a guide for future association mapping studies. At a meeting to select marker systems for target GCP crops held at the Plant and Animal (PAG) genome 2004 it was decided that 3000 cassava accessions, represented by 1500 accessions from CIAT's world germplasm collection, 1000 accessions from Africa (IITA) and 500 accessions from EMBRAPA will be selected for the study. DNA from these accessions will be extracted at each institution and sent to CIAT for re-distribution to all three participating institutions. Molecular markers for analysis of structural diversity will be 36 SSR markers, 2 each from the 18 linkage groups of the cassava map, that gave clear and reproducible allele patterns and high PIC will be used. IITA and CIAT will analyze 14 and 16 SSR markers respectively while CNPMF will analyze 6 SSR markers, in the 3000 accessions. CIAT will sub contract CNPMF.

Following the PAG meeting and further discussions with SP1 colleagues, a pilot study was proposed to analyze, a sub set, 500 genotypes, of the larger selection with SSR and DArT markers to fine-tune final selection criteria for the larger set of germplasm. The pilot study was also to compare the power of DArT and SSR markers to detect underlying genetic diversity structure. The DArT analysis was to be carried out by a young female national program scientist from Thailand at CAMBIA, Canberra, Australia in collaboration with IPGRI, CIAT. Selections for the pilot study were in the same proportions as the full set of 3000 genotypes, in other words170, 80, and 250 accessions from IITA, CNPMF and CIAT respectively. The comparison of SSR and DArT analysis of the 436 accessions was presented as one of the 'success story' at the GCP annual meeting in Brisbane. Result of the pilot study is also reported in Activity 12.17 of this report. We describe here selection of the set of 3000 accessions, selection of 36 SSR markers and work carried out so far on SSR analysis of genetic diversity in the selected germplasm.

Methodology
The selection of a set of 3000 cassava accession was based on a selection criteria that emphasizes a very broad genetic diversity and key agronomic traits such as Drought tolerance, resistance to major pests and diseases, adaptation to different ecologies, etc. The complete set of criteria used to select the germplasm set is listed in Table 12.9.

Table 12.9 Description of selection criteria for assembly of 3000 cassava accessions for SSR analysis

Selection Criteria	
1	CIAT and IITA Core Collections
2	CIAT and IITA Elite Clones
3	Officially released varieties from CIAT, IITA, and CNPMF
4	Drought tolerance
5	Tolerance to all major cassava pests and diseases worldwide
6	Good culinary quality
7	Tolerance to acid soils and low phosphorus soils
8	High beta-carotene varieties (yellow varieties)
9	Low and high level of amylose and amylopectin
10	Low and high level of carotene, Iron, Zinc and HCN in the root
11	High level of protein in the root
12	Good adaptation to different tropical and sub-tropical agro-ecologies of South and Central, America, Southeast Asia and sub-Saharan Africa
13	Dwarf accessions
14	Wild accessions (no more than 5% of total number)

For selection of molecular marker sets, criteria was a marker system with:
1. High level of information or polymorphism information content (PIC) per locus
2. Easily assayed in most cassava research labs around the world, for example PCR-based
3. Have been used previously in analyzing cassava diversity and can resolve close relationships in cassava germplasm
4. Amenable to automation

The marker system that best fits the above criteria in cassava is by far simple sequence repeat (SSR) markers. More than 600 of these markers exist for cassava of which about 200 are mapped, 67 of these markers have also been used to assess diversity in a sub-set of 300 genotypes from all over the world, in other words PIC values exist for them. The cassava team also agreed to do a pilot study to compare another marker system, DArTs with SSRs in assessing structural diversity using a random sub-set of 426 accessions from the larger collection of 3000 accessions. The result of this study is reported in Activity 12.7 of this report.

Leaf tissue from the CIAT selection was obtained from tissue culture plantlets from the genetic resources unit (GRU) or plants maintained in the field or screen house. DNA isolation from the selected germplasm was by a Dellaporta (1983) mini prep extraction method. DNA extracted from selections at CIAT, IITA and CNPMF was shipped to CIAT for re-distribution to all three participating institutions and also to CAMBIA for DArT analysis. CIAT as lead institute will collate and analyze the molecular data from all markers and all accessions, as well as compile passport data, including the local names, source (Country/State/Province/Region/Village), geographical position (Longitude, Latitude, Altitude) and the main agronomic traits, from all accessions into a data base that is accessible to the entire cassava research community.

Results

A total of 3000 accessions were selected for SSR analysis: 1500 at CIAT, 1000 at IITA and 500 at CNPMF, an excel file of the selection can be viewed at www.ciat.cgiar.org/molcas. Due to a delay in obtaining a permit from the Brazilian GR council to access materials from CNPMF, in the best of circumstances it could take 4-5 months to obtain a permit, another list of 500 accessions from CNPMF held in the CIAT world cassava collection was made for immediate access and molecular analysis. This new Brazilian germplasm set has 200 in common with the one selected at CNPMF. An agreement has been reached with CNPMF to analyze the CIAT selection with the participation of a CNPMF scientist and later analyze the 300 outstanding CNPMF accessions in Brazil.

DNA samples from 170 accessions were received from IITA, Ibadan of which 155 had sufficient quantity and quality for molecular analysis, and DNA from 281 genotypes were prepared at CIAT for the pilot study (Figure 12.11). An aliquot of all 426 DNA samples were sent to IITA Nairobi (ILRI-Bioscience facility) where molecular analysis will be conducted. No DNA samples were sent to CNPMF because an agreement on access to CNPMF's germplasm, a pre-requisite for release of funds from CIAT for CNPMF's sub-contract, could not be reached due to delays in obtaining a permit from the Brazilian genetic resources (GR) council. SSR analysis of the 436 accessions was carried out only in CIAT due to technical problems with the genotyping facility at IITA-Nairobi.

All DNA samples from the complete set of 1000 accessions selected at IITA, Ibadan have been received at CIAT. At CIAT, DNA isolation is been conducted as materials are received from the genetic resource unit (GRU), to date DNA isolation has been completed for roughly half of the 2000 selection: 1500 from CIAT and 500 from CNPMF. DNA extraction is expected to be complete by the second week in October.
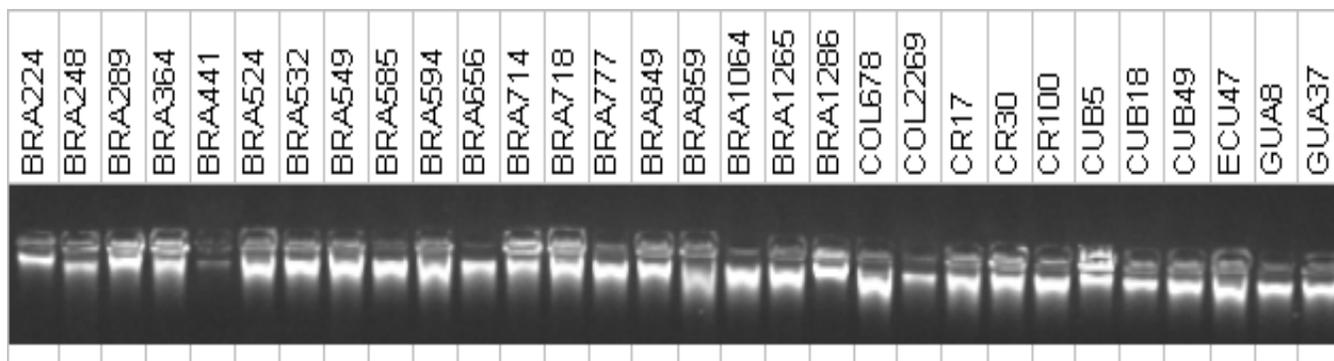
Figure 12.11. Agarose gel showing quality of some DNA samples from CIAT germplasm.

In previous studies of genetic diversity of cassava accessions from 14 countries in Africa and the Neotropics, 36 markers, 2 from every one of the 18 linkage groups that represent the 18 haploid chromosomes of cassava, with high PIC and that give very reproducible patterns were used. We proposed to use these 36 SSR markers (see Table 12.10.) with broad coverage of the cassava genome for structural diversity analysis of 3000 cassava accessions. These markers were used in genotyping 426 accessions of the pilot study and they revealed a structure in the accessions based upon region of origin and other unknown factors.

Conclusion and Perspectives
Selection of a set of 3000 cassava accessions and 36 SSR markers for structural diversity characterization of cassava has been completed. Also completed was a pilot study to characterize 426 accessions with the 36 SSR and DArT markers. Ongoing activities include completion of DNA isolation and SSR analysis of the rest of the germplasm data set.

References
Dellaporta SL. Wood J., Hicks JR, 1983. A plant DNA minipreparation: versión II. Plant Mol Biol. Rep 1: 19-21.
Mba R. E.C., P. Stephenson, K. Edwards, S. Melzer, J. Nkumbira, U. Gullberg, K. Apel, M. Gale, J. Tohme, and M. Fregene (2001) Simple Sequence Repeat (SSR) Markers Survey of the Cassava (*Manihot esculenta* Crantz) Genome: Towards a SSR-Based Molecular Genetic Map of Cassava. Theor and Appl Genet 102:21-31

Table 12.10. SSR markers selected to study structural diversity in Cassava.

Place of Evaluat.	SSR locus	Type of repeat	Left primer (Reverse)	Right primer (Forward)	Product Size (bp)	T * (°C)	MgCl$_2$ (mM)[A]	Thermocycler Program[a]	Genotyping place (first subset: 500 acc.)
IITA	SSRY4	(GA)$_{16}$	ATAGAGCAGAAGTGC AGGCG	CTAACGCACACGACT ACGGA	287	45	1.5	M	CIAT
IITA	SSRY9	(GT)$_{15}$	ACAATTCATCATGAGT CATCAACT	CCGTTATTGTTCCTG GTCCT	278	55	1.5	M	CIAT
CIAT	SSRY12	(CA)$_{19}$	AACTGTCAAACCATT CTACTTGC	GCCAGCAAGGTTTGC TACAT	266	55	1.5	M	CIAT
CIAT	SSRY19	(CT)$_8$(CA)$_{18}$	TGTAAGGCATTCCAA GAATTATCA	TCTCCTGTGAAAAGT GCATGA	214	55	1.5	M	CIAT
CIAT	SSRY20	(GT)$_{14}$	CATTGGACTTCCTAC AAATATGAAT	TGATGGAAAGTGGTT ATGTCCTT	143	55	1.5	M	CIAT
CIAT	SSRY21	(GA)$_{26}$	CCTGCCACAATATTG AAATGG	CAACAATTGGACTAA GCAGCA	192	55	1.5	M	CIAT
Brazil	SSRY34	(GGC)$_5$GGTGGC (GGT)$_2$	TTCCAGACCTGTTCC ACCAT	ATTGCAGGGATTATT GCTCG	279	55	1.5	M	CIAT
Brazil	SSRY38	(CA)$_{17}$	GGCTGTTCGTGATCC TTATTAAC	GTAGTTGAGAAAACT TTGCATGAG	122	55	1.5	M or N	CIAT
CIAT	SSRY51	(CT)11CG(CT)$_{11}$ (CA)$_{18}$	AGGTTGGATGCTTGA AGGAA	GGATGCAGGAGTGCT CAACT	298	55	1.5	M or Y	CIAT
Brazil	SSRY52	(GT)$_{19}$	GCCAGCAAGGTTTGC TACAT	AACTGTCAAACCATT CTACTTGC	266	55	1.5	M or Y	CIAT
CIAT	SSRY59	(CA)$_{20}$	GCAATGCAGTGAACC ATCTTT	CGTTTGTCCTTCTGA TGTTC	158	55	1.5	M1	CIAT
Brazil	SSRY63	(GA)$_{16}$	TCAGAATCATCTACCT TGGCA	AAGACAATCATTTTGT GCTCCA	290	55	1.5	M or Y	CIAT
Brazil	SSRY64	(CT)$_{13}$CG(CT)$_6$	CGACAAGTCGTATAT GTAGTATTCACG	GCAGAGGTGGCTAAC GAGAC	194	55	1.5	M or Y	CIAT
Brazil	SSRY69	(CT)$_{18}$ATT(AT)$_2$ (N)$_7$ (CTTT)2	CGATCTCAGTCGATA CCCAAG	CACTCCGTTGCAGGC ATTA	239	55	1.5	N	CIAT
CIAT	SSRY82	(GA)$_{24}$	TGTGACAATTTTCAGA TAGCTTCA	CACCATCGGCATTAA ACTTTG	211	55	1.5	M or Y	CIAT
CIAT	SSRY100	(CT)$_{17}$TT(CT)$_7$	ATCCTTGCCTGACAT TTTGC	TTCGCAGAGTCCAAT TGTTG	210	55	1.5	N	CIAT
IITA	SSRY102	(GT)$_{11}$	TTGGCTGCTTTCACT AATGC	TTGAACACGTTGAAC AACCA	179	55	1.5	N	
IITA	SSRY103	(GA)$_{22}$	TGAGAAGGAAACTGC TTGCAC	CAGCAAGACCATCAC CAGTTT	272	55	1.5	N	
IITA	SSRY105	(GT)$_6$GC(GT)$_2$(GA)$_{16}$	CAAACATCTGCACTTT TGGC	TCGAGTGGCTTCTGG TCTTC	225	55	1	N	CIAT

cont. Table 12.10.

Place of Evaluat.	SSR locus	Type of repeat	Left primer (Reverse)	Right primer (Forward)	Product Size (bp)	T * (°C)	MgCl$_2$ (mM)^A	Thermocycler Program^a	Genotyping place (first subset: 500 acc.)
CIAT	SSRY106	(CT)$_{24}$	GGAAACTGCTTGCACAAAGA	CAGCAAGACCATCACCAGTTT	270	55	1.5	N	CIAT
CIAT	SSRY108	(CT)$_{24}$CCT	ACGCTATGATGTCCAAAGGC	CATGCCACATAGTTCGTGCT	203	55	1.5	M or Y	CIAT
CIAT	SSRY110	(GT)$_{12}$	TTGAGTGGTGAATGCGAAAG	AGTGCCACCTTGAAAGAGCA	247	55	1.5	N	CIAT
IITA	SSRY135	(CT)$_{16}$	CCAGAAACTGAAATGCATCG	AACATGTGCGACAGTGATTG	253	45	1.5	Y	CIAT
IITA	SSRY147		GTACATCACCACCAACGGGC	AGAGCGGTGGGGCGAAGAGC	113	45	1.5	Y	
IITA	SSRY148		GGCTTCATCATGGAAAAACC	CAATGCTTTACGGAAGAGCC	114	45	1.5	Y	
CIAT	SSRY151		AGTGGAAATAAGCCATGTGATG	CCCATAATTGATGCCAGGTT	182	45	1.5	N	CIAT
CIAT	SSRY155		CGTTGATAAAGTGGAAAGAGCA	ACTCCACTCCCGATGCTCGC	158	55	1	Y	CIAT
IITA	SSRY161	(CT)$_{11}$TT(CT)$_{21}$(CA)$_{19}$	AAGGAACACCTCTCCTAGAATCA	CCAGCTGTATGTTGAGTGAGC	220	55	1.5	Y	
CIAT	SSRY164	(GA)$_{29}$	TCAAACAAGAATTAGCAGAACTGG	TGAGATTTCGTAATATTCATTTCACTT	187	45	1.5	N	CIAT
CIAT	SSRY169	(GA)$_{19}$(A)$_3$(GAA)$_2$	ACAGCTCTAAAAACTGCAGCC	AACGTAGGCCCTAACTAACCC	100	55	1	Y	CIAT
CIAT	SSRY171	(TA)$_5$CATA(GATA)$_8$GC(GA)$_{23}$	ACTGTGCCAAAATAGCCAAATAGT	TCATGAGTGTGGGATGTTTTTATG	291	55	1.5	N	CIAT
IITA	SSRY177	(CCT)$_6$CT(N)$_{65}$(CT)$_4$AT(CT)$_{18}$	ACCACAAACATAGGCACGAG	CACCCAATTCACCAATTACCA	268	45	1.5	Y	
IITA	SSRY179	(GA)$_{28}$	CAGGCTCAGGTGAAGTAAAGG	GCGAAAGTAAGTCTACAACTTTTCTAA	226	55	1.5	M	
IITA	SSRY180	(GA)$_{16}$(G)$_4$(GA)$_5$	CCTTGGCAGAGATGAATTAGAG	GGGGCATTCTACATGATCAATAA	163	55	1.5	M	
IITA	SSRY181	(GA)$_{22}$(G)$_3$C(GA)$_3$GGAA(GA)$_4$	GGTAGATCTGGATCGAGGAGG	CAATCGAAACCGACGATACA	199	55	1.5	Y	IITA
IITA	SSRY182	(CA)$_{17}$(N)$_{31}$GAGG (GA)$_8$	GGAATTCTTTGCTTATGATGCC	TTCCTTTACAATTCTGGACGC	253	55	1.5	Y	IITA

T*: Annealing Temperature. ^{A}: **MgCl$_2$** final concentration per PCR reaction (final volume in the PCR reaction: 25μ
^{a} M= MICROBC1; N= NEWBC1; Y= YUCADIV
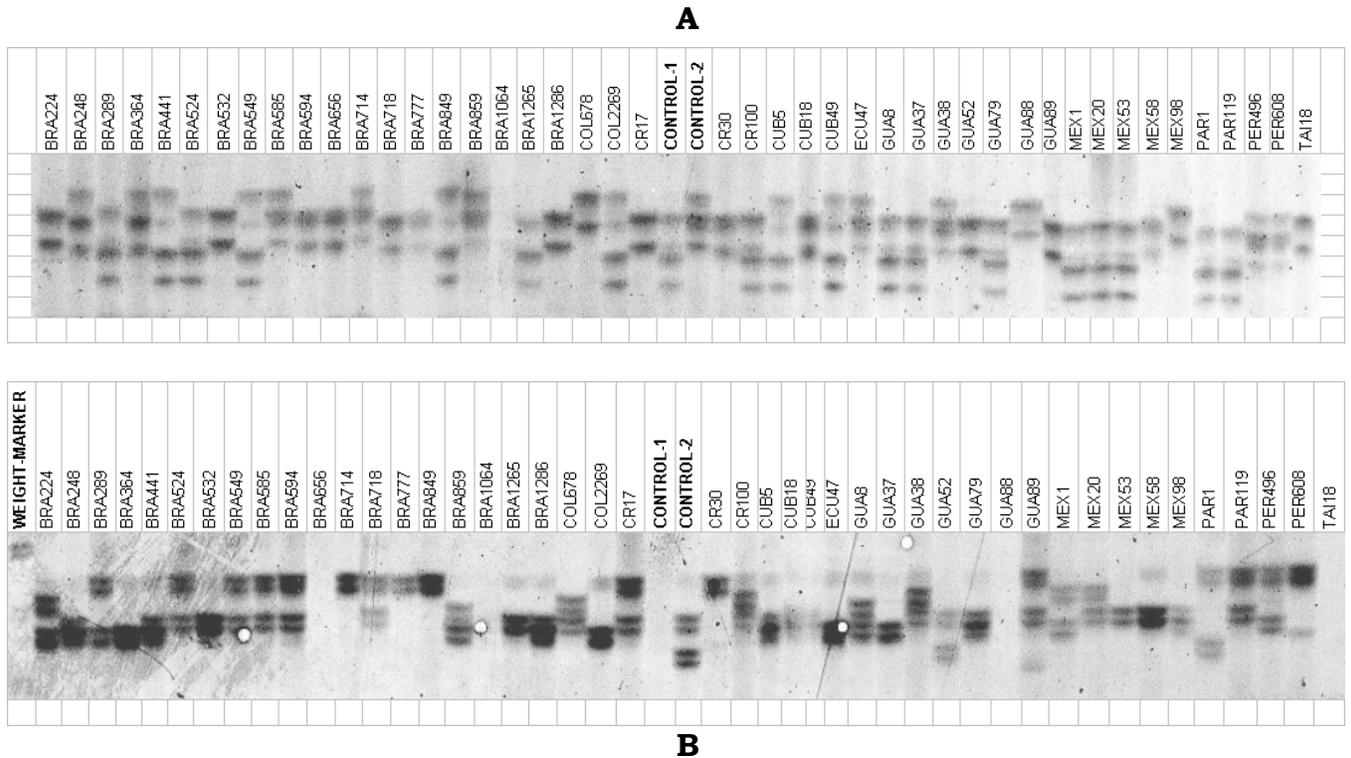
Figure 12.12. PAGE gel of PCR amplification of cassava accessions with SSR markers: SSRY135(A) and SSRY4 (B).

Activity 12.9 Analysis of genetic diversity in a cassava germplasm collection from Cuba using SSR markers

Collaborators:
Yoel Beovides, Marilys D. Milián (INIVIT Cuba), Janneth P. Gutiérrez, Edgar Barrera, Charles Buitrago, Jaime A. Marin, Alfredo Alves, Martin Fregene (CIAT)

Funding:
The Cassava Biotechnology Network (CBN)

Important Outputs
Assessment of genetic diversity of a collection of Cuban cassava land races and detection of a structure in this collection

Rationale

The assessment of genetic diversity of cassava germplasm from Cuba using 36 SSR markers was started last year. The study, concluded this year, seeks to understand the organization of diversity and genetic differentiation, with respect to germplasm from the rest of the world, of local cassava varieties from Caribbean island in light of evidence of a possible second center of diversity of cassava in Central America. A second objective was to provide cassava breeders in Cuba information to better exploit genetic diversity in their cassava collection. The study was carried out as collaboration with INIVIT, Cuba, with funding from the cassava biotechnology network (CBN)

Methodology

Plant material was 94 accessions selected from a collection of cassava held at INIVIT in Cuba, selection criteria were the economic importance and origin in Cuba. A set of 54 clones from Africa and the Neotropics: 12 from Nigeria, 10 from Tanzania, 12 from Guatemala, 20 from South America, and 13 improved genotypes from CIAT, were included. These genotypes a representative of a larger set of germplasm from these countries based upon previous SSR studies (Fregene et al 2003), were included for estimation of genetic differentiation. DNA from all accessions was obtained using the Dellaporta et al. method (1983). Concentration and quality of the DNA was checked by flourometer and agarose gel electrophoresis respectively. The DNA samples were diluted to a working concentration of 10ng/ul for subsequent PCR amplification.

PCR amplification, gel analysis and date collection of the DNA samples with 36 SSR markers were as described earlier (CIAT 2003; Mba et al. 2003). The raw SSR data was used to calculate estimates of genetic diversity and differentiation using the computer package GENSURVEY (Vekeman et al 1997). Genetic differentiation was estimated using the statistic F_{ST} (theta) and G_{ST} (Nei 1978) using the FSTAT computer program (Goudet 1990). Confidence intervals of F_{ST} and G_{ST} were calculated by jackknifing (200 replications) or by bootstrapping (1000 bootstraps). Pair-wise values of F_{ST} between countries were used in drawing a dendogram by the UPGMA method and the program NTSYS-PC (Rohlf 1993). Other analyses conducted with the SSR data include calculation of pair-wise genetic distance based upon the proportion of shared alleles (PSA), using the computer microsat (Minch 1993, **http://www.lotka.stanford.edu/microsat.html**) and cluster analysis of the genetic distance matrix using principal coordinate analysis (PCoA) and multiple correspondence analysis (MCA), using the computer package NTSYS-PC (Rohlf 1993).

Results

The evaluation of 36 SSR markers in 142 accessions yielded a high level of polymorphism, with the exception of 2 (SSRY127 and SSRY132) that were monomorphic. Number of alleles for the polymorphic markers ranged from 2 to 10 for each SSR loci. Seventeen alleles unique to accessions from Cuba were identified in the following SSR markers: SSRY 4 (0.04), SSRY 20 (0.006), SSRY 38 (0.005), SSRY 59 (0.006 y 0.079), SSRY 63 (0.033), SSRY 69 (0.023), SSRY 100 (0.011), SSRY 103 (0.052, 0.012, 0.012 and 0.006), SSRY135 (0.005), SSRY 151 (0.05), SSRY 171 (0.012 and 0.036) and SSRY 177 (0.014). Unique alleles were also found in some genotypes from Colombia (6), Nigeria (3), Tanzania (3) and Guatemala (1). Average gene diversity was high, with an average of 0.6292 ± 0.0120, for all the samples analyzed but highest for those from Cuba and Tanzania (Table 12.11)

Estimates of genetic differentiation between the country samples ranged from 0.04 to 0.06, with the highest being between Cuban and Guatemalan/Tanzanian accessions (Figure 12.13). Cluster analysis by principal coordinate analysis and PCA gave basically the same results. A pattern of clustering was observed between the African and Latin American accessions and within the Cuban accessions (Figure 12.14).

Table 12.11. Genetic diversity within samples of cassava accessions from 5 countries and standard deviation for *jackknifing* over loci (200 replications). H_t, H_s, D_{st}, y G_{st} are given over loci and samples (country).

Country sample	n	#loc	#loc_P	PLP	K	K_P	HO_p	HE_p	HEc_p
CUBA	86	34	34	100.0	5.8	5.8	0.6016	0.6314	0.6351
GUATEMALA	10	34	34	100.0	4.2	4.2	0.5556	0.6063	0.6385
COLOMBIA	11	34	33	97.1	4.5	4.6	0.5675	0.6087	0.6396
NIGERIA	16	34	33	97.1	4.5	4.6	0.5885	0.5949	0.6136
TANZANIA	10	34	31	91.2	4.2	4.5	0.6459	0.5869	0.6190
Average				97.06	4.64	4.72	0.5918	0.6057	0.6292
Stand. Dev.				3.60	0.65	0.61	0.0351	0.0169	0.0120

	Ht	Hs	Dst	Gst
Average	0.6538	0.6057	0.0482	0.0740
Stand. Dev.	0.1770	0.1682	0.0253	0.0377
95% CI	0.5780	0.5341	0.0383	0.0618
95% CI	0.7137	0.6663	0.0585	0.0878

n: Number of individuals
#loc: Number of loci
#loc_P: Number of polymorphic loci
PLP: Percentage of polymorphic loci
K: Number of alleles per locus
K_P: Number of alleles per polymorphic loci
HO_p: Observed Heterozygosity
HE_p: Expected Heterozygosity
Hec_p: Expected Heterozygosity corrected for small sample sizes (Nei, 1978)
Ht: Total Genetic diversity
Hs: Average genetic diversity within populations
Dst: Average genetic diversity between populations
Gst: Coefficient of genetic differentiation.
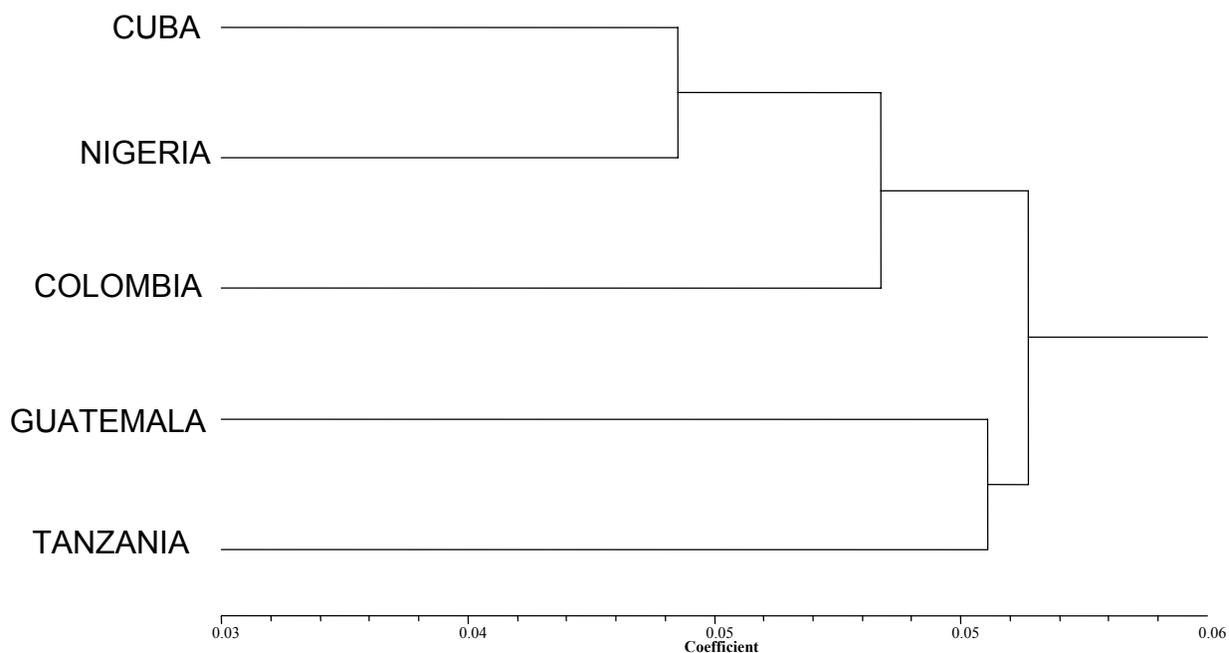
Figure 12.13. UPGMA dendogram of pair-wise comparison of the fixation index (F_{ST}) between samples of cassava from Cuba, Nigeria, Colombia, Guatemala and Tanzania.
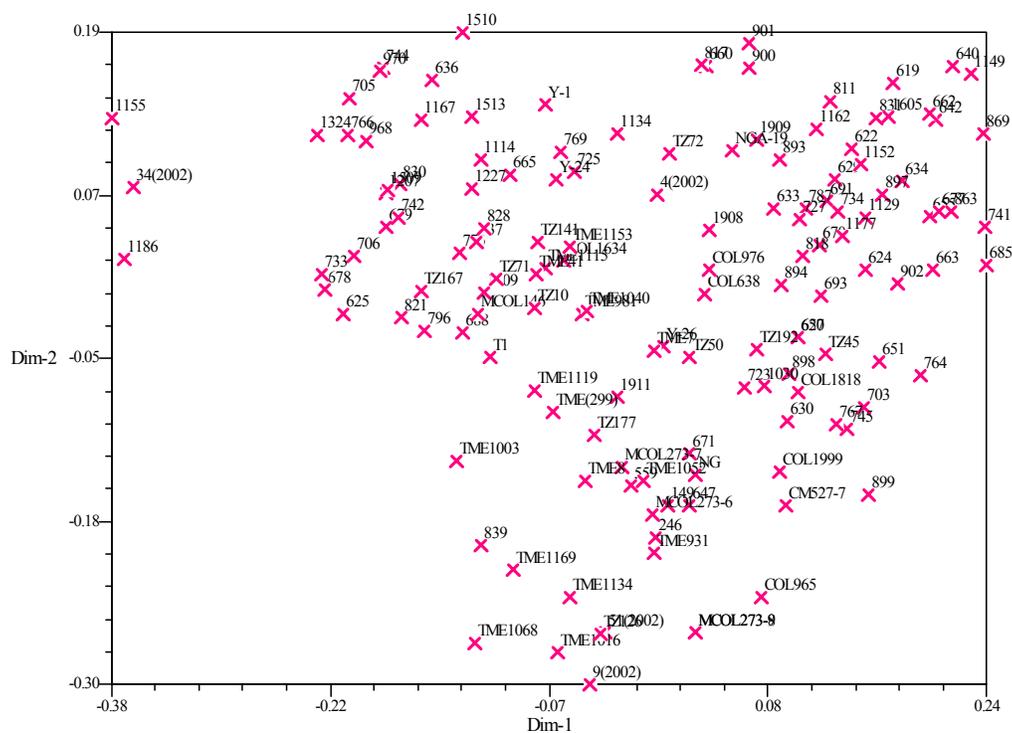
Figure 12.14. Principal coordinate analysis of cassava accessions from Cuba and reference accessions from Nigeria, Colombia, Guatemala and Tanzania.

Conclusion and Perspectives

Genetic Diversity of cassava in the Caribbean island of Cuba follows diversity found for cassava in the rest of the world, a high genetic diversity and low genetic differentiation. However a structure was observed within the Cuban collection using cluster analysis. This structure is the basis of future work on linkage disequilibrium mapping of dry matter content using candidate starch biosynthesis genes.

References

Fregene M, Suárez M, Mkumbira J, Kulembeka H, Ndedya E, Kulaya A, Mitchel S, Gullberg U, Dixon AGO, Dean R, Kresovich S (2003) Simple sequence repeats marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. Theoretical and Applied Genetics, accepted.

Mba REC, Stephenson P, Edwards K, Melzer S, Mkumbira J, Gullberg U, Apel K, Gale M, Tohme J, Fregene MA (2001) Simple Sequence Repeat (SSR) Markers Survey of the cassava (*Manihot esculenta* Crantz) Genome: Towards an SSSR-Based Molecular Genetic Map of Cassava. Theoretical and Applied Genetics. 102: 21-31.

Minch (1993), http://www.lotka.stanford.edu/microsat.html

Dellaporta SL, Wood J, Hicks JR (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1:19-21.

Goudert J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485-486.

Kimura M, Crow JT (1964) The number of alleles that can be maintained in a finite population. Genetics 49:725738.

Nei M (1978) Estimation of average heterozygosity and genetics distances from a small number of individuals. Genetics 89:583-590.

Rohlf DJ (1993) NTSYS-PC numerical taxonomy and multivariate analysis system. Version 1.8 Exeter Publ., Setauket, New York, USA.

Vekemans X, Lefebvre C (1997) On the evolution of heavy metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. J Evol Biol 10:175-199.

Wright S (1951) The genetic structure of populations. Ann Eugen 15:323-354.

Activity 12.10 Studies in Market Preferences of Cassava Cultivars in Malawi using SSR Markers

Collaborators:

Linley Chiwona Karltun (SLU, Uppsala, Sweden), Prof Janice Jiggins (WAU, Wageningen, Netherlands) A. Akerbolm (IPICs, Sweden), C. Buitrago, F. Rojas, J. Marin, C. Ospina, M. Fregene (CIAT)

Important Outputs:
1) Identification of 54 unique genotypes that form 3 broad clusters that might represent market classes of cassava in Northern and Southern Malawi

Rationale
Cassava is the second most important staple crop after maize in Malawi. Two other factors have emerged in Africa that will further increase the role of cassava as a staple crop: 1) The increasingly unpredictable rain pattern causing large fluctuations in maize harvests and 2) a large increase in prices of chemical fertilizers and hybrid maize seeds make maize growing a less viable, or even impossible, alternative, for the smallholder farmer. As in many parts of Africa, the growing of cassava has increased considerably over the past years in Malawi. From 1992 to 1996 cassava production in Malawi is officially reported to have increased from about 100 000 tons/yr to more than 500 000 tons/yr. Since the cassava mosaic and mealy bug outbreak in the 1980s there has been a rapid shift in cassava cultivars and there is an obvious need for new cassava varieties with resistance to the major pests and diseases. Major disease outbreaks, like the one caused by the mealy bug in the eighties in Malawi or the cassava mosaic virus outbreak in Uganda, could be less devastating if resistant cassava material was available and accepted by the farmers within their farming systems. We conceive that the methods for provision of new cultivars to farmers can be considerably improved. A project to identify, evaluate and discover traits that make certain local varieties popular amongst consumers was developed and funded by SAREC. We describe here Molecular marker technology for tracking gene flow in trading of cassava cultivars and to explore how molecular markers can be used to identify cultivar preferences in urban cassava markets

Methodology
The study was conducted in two geographic areas, in the north and the south of Malawi, respectively. In the north, Nkhata-Bay close to Lake Malawi and has a fairly long dry period, mean daily temperature (DMT) during the growing season, above 20°C, and relatively low population density. The area in the south, Mulanje, is less dry and more densely populated. In each area, 5 farmers who are recognized by other farmers as cassava farming enthusiasts were recruited into the study. Using a combination of interviews, cassava cultivars were collected from markets in the above areas and planted at the Namiganzi farm center, Malawi. A total of 54 cultivars were collected. The cultivars were subjected to molecular analysis using 36 SSR markers for cultivar identification. SSR marker and data analysis were as described earlier Fregene et al. (2003).

Results
The 54 cultivars collected from Malawi revealed a fairly high amount of SSR allele diversity (Figure 12.15). The total number of alleles per locus ranged between 3 and 7. The fairly large number of alleles per locus enabled the unique identification of every single cultivar collected. Analysis of genetic relationships between the clones revealed that they formed 3 loose clusters that appear to represent market classes of cassava in Malawi (Fig8.16). Detailed and conclusive interpretation of the above results will have to wait for agronomic data of the cultivars from the replicated trial at the Namiganzi farm center to be collected later in the year and additional information collected during the interviews.

Figure 12.15. Silver stained PAGE gel electrophoresis of PCR amplification of the cassava genotypes collected in Malawi using the primers SSRY51 and SSRY63.

Figure 12.16. Principal component analysis (PCA) of pairwise genetic distances based on proportion of shared alleles (PSA) between the cassava accessions from Malawi

Conclusions and Future Perspective

An attempt has been made to use molecular markers in the identification of cultivar and therefore their preferences in urban cassava markets in Malawi. Using 36 SSR markers, 54

unique genotypes could be identified from 54 cassava cultivars collected from markets in Northern and Southern Malawi. The cultivars also formed 3 clusters that might represent market classes. Data from agronomic trial is being awaited to obtain conclusive evidence of this and also to identify what makes these cultivars preferred as a first step in making cassava breeding projects more relevant to end users.

References

Fregene M., Suarez M., Mkumbira J. , Kulembeka H. , Ndedya E. , Kulaya A. , Mitchel S. Gullberg U. , Rosling H., Dixon A., Kresovich S. (2003) Simple Sequence Repeat (SSR) Diversity of Cassava (*Manihot esculenta* Crantz) Landraces: Genetic Structure in a Predominantly Asexually Propagated Crop Theor Appl Genetics 107:1083-1093.

Activity 12.11 Data base of the Molecular Diversity Network of Cassava (MOLCAS)

Collaborators:
Charles. Buitrago, F. Rojas, J. Marin, C. Ospina, M. Fregene (CIAT)

Funding:
IPICs, Uppsala, Sweden

Results
This year, the MOLCAS database of country studies of SSR diversity of cassava was updated with data from country studies from Cuba, Sierra Leona, and Malawi. The new data can be viewed at: **http://www.ciat.cgiar.org/molcas**

Table 12.12. Record of visits to the MOLCAS website between May1 and September 5 2004

Request	Hits	Sep-04	Aug 2004	Jul 2004	Jun 2004	May 2004
/molcas/imagen.jsp	22,676	35	1,807	993	1,644	1,066
/molcas/locus.jsp	21,328	12	2,293	821	1,992	833
/molcas/alelosp.jsp	15,108	8	2,114	746	1,458	1,358
/molcas/imagenbioquim.jsp	7,836	5	293	82	76	7
/molcas/	7,800	57	253	214	232	1,229
/molcas/estudios.jsp	6,010	10	290	186	216	290
/molcas/markers-det.jsp	3,390	6	183	138	128	139
/molcas/intrap_data2.jsp	2,177	6	118	70	90	83
/molcas/appendix1.jsp	2,126	2	118	71	58	94
/molcas/appendix2.jsp	1,335	0	101	47	57	43
/molcas/pcr_cond.jsp	975	4	64	47	58	42
/molcas/studies.jsp	784	3	36	33	34	20
/molcas/images.jsp	150	0	0	14	11	92
/webapps/molcas/	104	0	0	0	0	0
	91,799	148	7,670	3,462	6,054	5,296

The MOLCAS web-site is increasingly becoming a useful asset for the research community as demonstrated by the over 260% increment in the visits to the MOLCAS this year, a total of 91,799 visits between May and September this year, compared to a total of 34,477 visits the same period of time in 2003 (Table 12.12).

Activity 12.12   Mining the Primary Gene Pool of Cassava: Introgression of High Root Protein from Accessions of Manihot esculenta sub spp Fabellifolia and Manihot Tristis into Cassava

Collaborators:
J. Gutierrez, O. Akinbo, N. Morante, T. Sanchez, J. Marin, C. Ospina, H. Ceballos, L. Santos, A. Alzate, S. Moreno, M. Fregene (CIAT)

Funding:
CIAT core funds

Important outputs
1) Generation of back cross families for QTL mapping and development of improved varieties with high protein content.
2) Evaluation of protein content in Ghana of varieties from Central America confirms the high protein content observed earlier
3) Standardization of a SDS-PAGE analysis method towards a proper characterization of the proteins contained in high protein content cassava and inter-specific genotypes

Rationale
As a major staple throughout the tropics, cassava can serve as a cheap means of deploying adequate protein requirement amongst poor people as well as for animal feeds. An advanced back cross QTL (ABC-QTL) to introgress high protein content genes from wild relatives into cassava is in its third year at CIAT. Similarly high protein content cassava varieties mostly from Central America were re-evaluated in another environment, Wenchi, Ghana, this year. The cassava varieties have also been re-established in the field at CIAT for a second year of evaluation. Genetic crosses of these high protein varieties are being made to elite parents of the CIAT cassava gene pools for breeding for high protein content and for QTL mapping studies. Other activities continued this year include standardization of the SDS-PAGE methodology for the determination of kind and size of proteins found in high protein accessions.