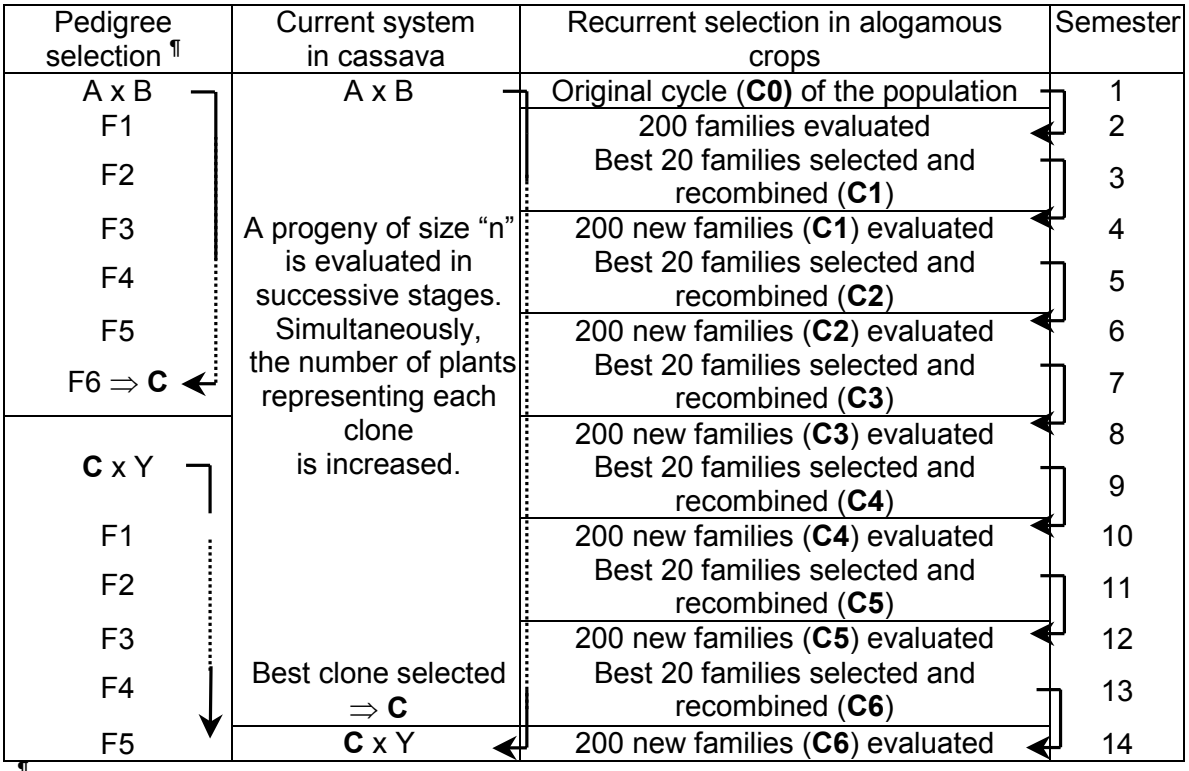
OUTPUT 4

Developing a new approach for cassava breeding integrating biotechnology tools.

Rationale

Cassava improvement has not been as consistent and efficient as in other crops due to many constraints. A typical scheme implies crossing elite clones to produce segregating families (Figure 4.1). Each individual produced is highly heterozygous. Once a superior genotype is identified (a process that requires about six years), it is vegetatively multiplied to take advantage of the reproductive habits of this crop. This system (except for the vegetative multiplication) has similarities with the ones used for autogamous crops (beans, wheat, rice, etc.) as well as for the hybrid maize industry. However, there is a major difference because cassava is never pushed to produce inbred (homozygous) lines from the segregating progenies of a given cross. The system also bears some similarities with recurrent selection used in allogamous crops (maize), but there is a significant difference because in cassava there is not a clearly defined population whose allelic frequencies are modified through evaluation and selection, as in true recurrent selection schemes.



¶ Used in autogamous crops and for the production of inbred lines in hybrid crops such as maize.

Figure 4.1. Illustration (highly simplified) of the differences in breeding systems employed for the genetic improvement of different type of crops.

From the simplified information provided in Figure 4.1, it is apparent that a major drawback of cassava breeding is also the length of each breeding cycle. For the reasons below, cassava breeding is slow and inefficient.

a. Because no inbreeding is carried out at any stage of cassava breeding, a sizable genetic load (undesirable or deleterious genes) is expected to prevent the crop fully achieving its actual yield potential (Hedrick, 1983).

b. There are no clearly defined populations (as defined by quantitative genetics), allelic frequencies cannot be efficiently modified. Cassava breeding in this regard resembles more the selection of segregating progenies from two parental lines in autogamous crops.

c. Because the highly heterozygous nature of the crop, dominance effects are likely to play a very important role in the performance of materials being selected. The current scheme can exploit dominance effects because, once an elite clone is identified, it can be propagated vegetatively, therefore carrying along the combination of genes responsible for dominance effects (Lamkey and Staub, 1998). However, selection of progenitors for the production of new segregating material is based on their *per se* performance. In that case, the current procedure has a bias because the breeding values of these clones are unlikely to be well correlated with their performance, precisely because of the distorting effects of dominance.

d. Production of recombinant seed is cumbersome in cassava. On average, only 0.6 viable seeds per pollination are produced. It takes about 16-18 months since a given cross is planned until an adequate amount of seed is produced.

e. When a desirable trait is identified, it is very difficult to transfer it from one genotype to another (even if a single gene controlled the trait). The back-cross scheme, one of the most common, successful and powerful breeding schemes for cultivated crops (Allard, 1960) is not feasible in cassava, because of the constant heterozygous state used throughout the breeding process.

f. Maintenance of genetic stocks is expensive and cumbersome. The only proven methods for long term storage of germplasm is through tissue culture procedures, which is expensive and requires several months to recover plants for planting in the field. The other alternative is maintaining representative plants in the field, which is also expensive when a large number of genotypes need to be maintained year after year, and also the stocks are vulnerable to gradual contamination by pathogens.

g. The occasional exchange of germplasm among cassava breeding programs in different countries is restricted to a few plants from few genotypes. Cassava breeding projects, effectively work in isolated conditions.

h. Lack of inbreeding in cassava implies that there are few opportunities for identifying useful recessive traits, which could have huge beneficial effects on the crop. For instance, acyanogenesis in the roots has been identified as a very desirable trait. It has been postulated that this trait may be recessive. Also worth mentioning are the several starch mutations that are generally recessive in most crops (Neuffer et al., 1997).

These may all be a valid explanation for the limited genetic gains for higher productivity observed in the crop, compared with that of other crops such as maize or rice. It should be emphasized that because the highly heterozygous condition of cassava in every stage of the breeding process, consolidation of genetic gains is difficult, due to the inherent genetic instability of heterozygosity.

From the practical point of view, implementing a traditional recurrent selection method in cassava offers some problems. Pollinations are slow and inefficient. It takes about 16-18 months since a given cross is planned until the recombinant seed is finally obtained (usually field operations have to adjust to the occurrence of the rainy season, and if that is the case, then planting can only be done 24 months after planning the cross). The third year would be used to grow the plant from the botanical seed. During the fourth year, a clonal evaluation could finally be carried out. Therefore a typical recurrent selection method would require no less than six to seven years (Figure 1). In this case, however, no selfing for reducing genetic load would have been included.

Sspecific Objectives

a) *To develop a methodology that will allow overcoming some of the most important drawback in the current cassava-breeding scheme.*
b) *To incorporate biotechnology tools into routine cassava breeding activities.*

Materials and methods

Advantages of inbreeding

When an elite clone is self-pollinated two important events occur: **a)** the unique, specific combination of genes present in the genotype is broken, therefore loosing the agronomic superiority that the clone might have. **b)** self pollination forces an average of half of the loci to become homozygous, thus facilitating the elimination of undesirable, deleterious alleles present in the original clone but hidden because of its predominant heterozygosity. In other words, selfed progenies allow for a reduction of the genetic load originally present in the clone, therefore becoming better progenitors themselves. In a way, selfing allows to "*concentrate*" the desirable genes originally present in the elite clone.

If inbreeding was pursued until near or complete homozygosity, then the transfer of desirable traits through the back-cross scheme becomes feasible. Also homozygosity "*captures*" genetic superiority because of its inherent genetic stability. Therefore each cassava improvement cycle would be a consolidated step that could help further progress in a more consistent and predictable way. On the other hand, each time a hybrid is used as parent, the process goes back to the initial step because of the genetic instability of the heterozygous material. In the latter case, progress cannot be easily consolidated or sustained through time, but in a rather inefficient way.

From the quantitative point of view the variability of a given population has traditionally been split into two major components: additive and dominant effects (Bos and Caligari, 1995; Hallauer and Miranda, 1988). Additive effects are very important because they define the breeding value of an individual, that is, its relative merit based on the quality of the progeny they produce. Dominance effects are also very important in plant breeding. They are the main contributors to the heterosis or hybrid vigor observed in hybrid cultivars, including

cassava. However, contrary to the additive effects, the dominance cannot be transmitted to the progeny. This means that dominance effects cannot be effectively exploited in a breeding program, unless a sophisticated breeding scheme (reciprocal recurrent selection) is employed (Hallauer and Miranda, 1988).

Table 4.1. Distribution of additive and dominance genetic variance among and within full-sib and inbred line families (Hallauer and Miranda 1988).

	Among families		Among plants within families	
Type of family	Additive variance	Dominance variance	Additive variance	Dominance variance
Half-sib family	1/4	0	3/4	1
Full-sib family	1/2	1/4	1/2	3/4
Inbred lines	2	0	0	0

The current breeding scheme is based on the selection of individual genotypes within half- or full-sib families. Table 1 illustrates how the total genetic variance and its components (additivity and dominance) are partitioned among and within different types of families. It is clear that all genetic effects will influence the selection of plants from half- or full-sib families. That is, 100% of additive variance and 100% of dominance variance are exploited during the selection process (this is so because the breeder selects the best families, and then, the best genotype within the best family). This is a convenient situation because, once a given clone is selected, both the additive and dominance effects determining its good performance can be perpetuated because of the vegetative reproduction of the crop. The specific combination of genes present in a clone can be maintained unaltered generation after generation, as long as there is only vegetative reproduction.

However, only the additive portion of the total genetic variance can effectively be passed on to a next generation when the same clone is used as parent in a breeding project. It is important to recognize that dominance strongly influences the selection of the best clones but has no effect on their breeding value. In other words, dominance effects can be beneficial for the *per se* performance of a clone, but it has a confounding effect of its actual value as progenitor.

Inbreeding is advantageous because it erases the dominance effects from the selection process. The resulting inbred lines do not possess any dominance effects, and therefore, there will be no heterosis or hybrid vigor expressing in their performance. That is precisely why inbred lines are inferior, agronomically speaking, compared with the non-inbred cassava materials. A striking feature of the data presented in Table 4.1 is that inbred lines show twice the additive variance originally present. In other words, when selecting among inbred lines the additive variance originally present (among F2 plants from an F1-hybrid) has been expanded, thus greatly facilitating the selection process of those additive effects that are precisely the only ones that define a superior progenitor.

Inbred lines are better material for selecting progenitors because, by definition, they carry lower levels of genetic load, no confounding dominance effect influence the selection process and because the additive genetic effects are expanded considerably, making the selection

more efficient. Also if a breeding process is based on the use of inbred lines the transfer of valuable traits is greatly facilitated because the back-cross scheme becomes feasible.

The availability of inbred lines in cassava would also benefit other areas in addition to breeding. Genetic and molecular marker analysis would be facilitated if homozygous lines were produced. The only way to maintain germplasm in cassava is either by growing the plants in the field or by tissue culture. Inbred lines could be maintained and shipped in the form of botanical seed. Phytosanitary problems could be reduced or eliminated if maintenance and/or multiplication of genetic stocks were partially based on botanical seed. Germplasm exchange among the few cassava-breeding projects in the world would be enhanced because it would be based on botanical seed, rather that vitro-plants. Finally, clones could be reproduced by sexual means. Although time – consuming, the first stage of evaluation could be based on many plants produced by the crossing of selected inbred lines. Currently the evaluation process takes three years to reach a stage for selection based in just 30 plants.

The problems of inbreeding

Cassava, being an out-crossed crop, abhors inbreeding and shows severe depression. As was the case of temperate maize in the early 1900s and tropical maize by the 1970s, cassava will need to be improved for its tolerance to inbreeding depression. A few recurrent selection cycles (selfing each elite clone down to the S2 level, and recombining the surviving progenies) should prepare elite cassava populations for the trauma of total homozygosity.

A recurrent selection involving the production of inbred lines would be difficult to implement because of the length of each cycle of selection. It is estimated that no less than **nine years** will be required from the time a group of elite clones are selected until recombinant seed from their inbred lines was obtained. Therefore, if a breeding scheme using inbred lines is to be implemented, a way to reduce the time required for each cycle of selection is urgently needed.

Doubled haploids have been produced and have benefited breeding efforts in many crops (Griffing, 1975). Upon producing an F1 plant, tissue culture techniques are applied to the reproductive tissue (typically anther culture). This process produces a haploid tissue that, quite frequently, doubles spontaneously to produce the doubled-haploid tissue, which by definition, is homozygous. There are other alternatives for producing similar materials (i.e. using inter-specific crosses). However these methods have seldom been incorporated as a routine tool in breeding projects.

If an efficient protocol for the production of doubled-haploids were available, it could be incorporated into the cassava breeding process with the advantages that inbred lines offer as explained above. From the practical point of view the protocol for the production of doubled-haploids would allow shortening the time required to produce hybrids from inbred lines down to **three years**.

Expected responses to selection with alternative breeding methods.

Breeding projects always search for maximizing the gains from selection. The genetic progress (**GP**) after one cycle of recurrent selection can be estimated as follows (Hallauer and Miranda, 1988; Simmonds and Smartt, 1999):

$$GP = i\,\sigma^2_A / \sigma_F$$

Where ***i*** is a factor related to the intensity of selection (proportion of the population selected to be parent for the next cycle); σ^2_A is the additive component of the genetic variance measured in the parental population and σ_F is the phenotypic standard deviation of the parental population. These σ^2_A and σ_F parameters will vary depending on the selection unit used (i.e. individual plant, mean family performance, mean of clones across replicated trials in different locations, etc.).

In turn, the σ_F can be defined as:

$$\sigma_F = \sqrt{\sigma^2_A + \sigma^2_D + \sigma^2_E}$$

Where σ^2_A and σ^2_D are the additive and dominance components of the genetic variance and σ^2_E is the environmental effect of the evaluation of the respective selection unit. It should be clear that to increase **GP** there are only few alternatives: **a)** Increase the value of ***i*** (that is increase the selection intensity); **b)** Increase the additive component of the genetic variance exploited by selection and/or **c)** Reduce one or all components of the phenotypic standard deviation. Changes in the selection intensity do not depend in the breeding scheme employed. On the other hand the proportion and magnitude of the additive component of the genetic variance exploited and the size of the phenotypic variance depend heavily on the breeding method implemented, as illustrated below.

The phenotypic standard deviation in phenotypic mass selection is very large because of the environmental component associated with single plant evaluations. In the clonal selection used in cassava, although single genotypes are selected (as in mass selection), the environmental variance is reduced because "***n***" plants representing the genotype are used in the evaluation and selection process. However, all the dominance effects remain as component of the denominator of the formula. Therefore, clonal selection would maximize **GP** when dominance effects are negligible, which is not the case of cassava.

When doubled-haploids are used, two important modifications are introduced into the formula for **GP**: **a)** The additive component in the numerator and denominator of the formula is now twice as large as before (but in the denominator the square root of the additive component is used); and **b)** The dominance component of the phenotypic variance disappears.

$GP_{(\text{Phenotypic mass selection})}$	=	$\dfrac{i\,\sigma^2_A}{\sqrt{\sigma^2_A + \sigma^2_D + \sigma^2_E}}$

$GP_{(\text{Clonal selection})}$	=	$\dfrac{i\,\sigma^2_A}{\sqrt{\sigma^2_A + \sigma^2_D + (\sigma^2_E / n)}}$

$GP_{(\text{Doubled-Haploids selection})}$	=	$\dfrac{i\,2\,\sigma^2_A}{\sqrt{2\,\sigma^2_A + (\sigma^2_E / n)}}$

The information provided illustrates the advantage of clonal evaluation over phenotypic mass selection. It also suggests that selection based on doubled-haploids would be better than the clonal evaluation (based on F1s) traditionally used for cassava breeding. However, because of the inbreeding depression observed in this crop, doubled-haploids would perform very poorly compared with the full vigor clones (F1s).

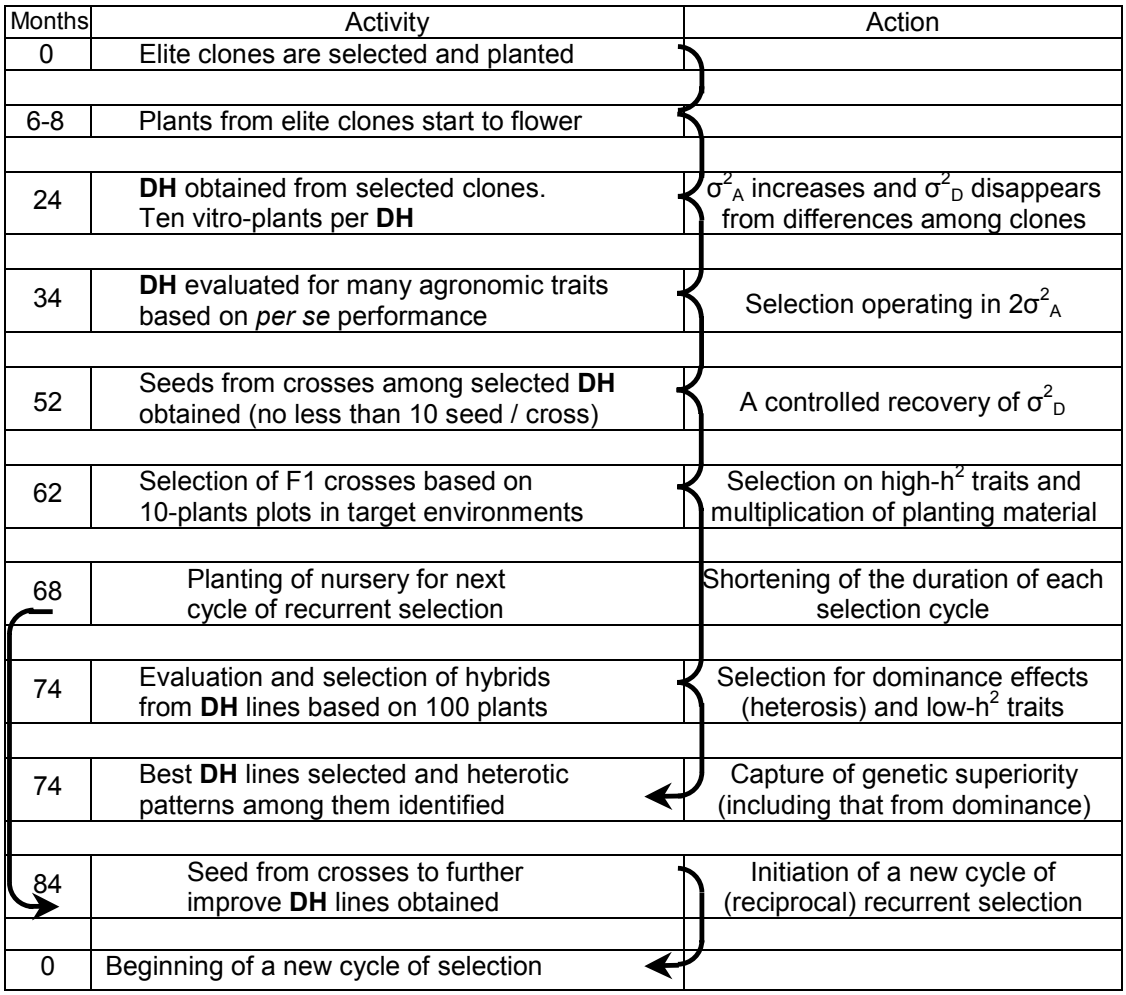


Figure 4.2. Illustration of a breeding scheme based on the production, evaluation and selection of doubled-haploid cassava lines to exploit additive and dominance effects in the production of superior hybrid clones.

It is difficult to make a fair comparison between the traditional and the new proposed scheme, because the latter introduces an intermediate stage when selection in the homozygous stage is conducted on the *per se* performance of the doubled haploids. This is the point where twice as much additive variance can be exploited. The second stage in the new scheme would be reconstituting the dominance effects in a controllable and predictable

way, by crossing specific pairs of doubled haploids lines. Hybrids developed from inbred cassava clones should perform better than current hybrids because: **a)** the elimination of deleterious genes through homozygosity; **b)** easier identification of "complementing parents" for the production of hybrids with maximized heterozygosity or hybrid vigor (dominance effects); and **c)** the possibility of building up, over several recurrent selection cycles, the dominance effects.

Results

A drastic change in the way cassava breeding is achieved should be introduced for taking advantage of the benefits of inbreeding. Figure 4.2 illustrates the general features of the proposed scheme.

*Production of doubled-haploid (**DH**) lines*

The process starts with the selection of elite clones themselves or after improvement for tolerance to inbreeding following the S2-recurrent selection described above. Once the planted material begins flowering, tissue will be taken for the induction of doubled haploidy through tissue culture protocols developed specifically for that purpose in cassava.

Upon the production of **DH** tissue or embryos, in vitro multiplication of each line will be carried out, to produce at least 10 hardened plants ready for transplantation to the field. This would take place at the end of the second year of activities.

Selection of doubled-haploid lines

Several **DH** lines will be produced and the ten plants representing each of them will be planted in a *Clonal Evaluation Trial* in the proper target environment. Hopefully these trials will involve at least 200 **DH** lines. Selection of these lines will be conducted for relevant characteristics with moderate to high heritability: resistances to diseases and/or insects, plant architecture, root dry matter content, root and parenchyma color, harvest index, etc. The selection at this stage operates with twice the additive genetic variance expected to be found in the original population under random mating conditions. Therefore, it is expected that large contrasts will be apparent at this stage. Lines surviving to this stage will have, by definition, reduced genetic load compared with the elite lines from which originated.

While the field evaluation is conducted lab analyses can be simultaneously carried out to obtain the molecular fingerprinting of each line. This will allow for further selection of characteristics difficult or impossible to determine from the field trials. For instance, marker assisted selection for CMD (Cassava Mosaic Disease) could be implemented in Colombia, although the disease is not present in this country. Also genetic distances among the lines could be determined to facilitate the following stage within the recurrent selection cycle.

*Production of hybrids from selected **DH** lines*

The following stage in the selection process involves the production of hybrids among the surviving **DH** lines. It is expected that from the 200 or more **DH** lines at least 30 will reach this stage. Although it is clear from the literature that genetic distances have failed to explain satisfactorily the heterosis among inbred lines in maize (Lamkey and Staub, 1988), genetic distances measured through molecular markers can be used at least to orient the crosses

that deserve some priority. This could be justified until an adequate definition of heterotic patterns is eventually reached. Since the parental materials (**DH** lines) are homozygous just a few seeds per cross are required at this stage. The only justification for obtaining more than one seed would be to accelerate the time required for evaluation with large number of plants representing each hybrid or clone.

With the production of hybrids from the selected **DH** lines, dominance effects (heterosis) are generated and, because of the breeding scheme proposed, will be fully exploitable by the cassava-breeding project.

*Evaluation of **DH**-derived hybrids.*
Depending on the number of hybrid seed produced the previous stage the evaluation and selection of hybrids can be conducted in two successive steps or just one growing cycle. In Figure 4.2, it is assumed that only ten plants from each cross can be obtained from botanical seed, and therefore that the evaluation and selection is conducted in two consecutive growing cycles.

The *first selection* is performed on all the hybrids produced and based on the 10 plants representing each hybrid clone. Because there is no replication, selection will be based only on high–heritability traits, and in the proper target environment, to allow for the pressure from biotic and abiotic limiting factors. The same evaluation plots are used as seed multiplication plots.

The *second stage* of selection and evaluation is conducted with about 100 plants (i.e. two replications at two locations with 25-plant plots). Only hybrids that survived the selection process the previous year will be included in this evaluation. Low-heritability traits are incorporated as selection criteria at this stage. Only a few clones will survive this selection and they will be included in *Regional Trials* for their eventual release as has been traditionally done up to now.

Preparation of nurseries for next cycle of recurrent selection
While the evaluation of hybrids is conducted, their parental **DH** lines will be planted in the field in such a way that they are about six-month old when the results of the hybrid trials become available. As soon as the hybrid trials yield results regarding the best **DH** progenitors and the identification of eventual heterotic patterns, they will be crossed to generate new genetic material for the following cycle of selection. The only purpose of these crosses will be to generate F1 plants from which to extract flower tissues for the production of a new generation of **DH** lines. Crosses will be made among DH lines from the same heterotic group. It is expected that the definition of heterotic patters will initially be weak (precisely because no breeding has been made to strengthen them), but as the process develops they will become stronger and clearer.

Hybrid trials will not only generate elite clones to be included in *Regional Trials* and eventually be released as new varieties, but also provide important information about the **DH** lines that generated the hybrid clones. This information will be used to determine lines with good general combining ability (i.e. that generate progenies with performances that are better than the mean of all the hybrids evaluated) as well as detecting heterotic patterns. This information is fundamental for deciding the kind of crosses that will be made for the next selection cycle.

Advantages of the proposed scheme

The capacity to produce inbred lines in cassava through the use of a dynamic process allows to drastically change the breeding process: **a)** the emphasis will shift from producing vast number of hybrids hoping that one (or few) will be genetically superior, towards the production of parental lines that will allow *'to design'* outstanding hybrids in a gradual, consistent and reliable fashion; **b)** genetic loads will be quickly reduced in elite cassava populations; **c)** hybrids produced from inbred lines will be better than hybrids produced from non-inbred progenitors because genetic load is reduced and because the system allows building up dominance effects; **d)** germplasm exchange will be greatly facilitated (botanical seed of outstanding parents) with obvious advantages for the cassava research community; **e)** gene exchange will also be greatly facilitated (currently it is very difficult to transfer one valuable gene from its source into an agronomically superior clone: the availability of inbred lines would make the back-cross scheme feasible for cassava); **f)** inbred materials are genetically stable, they allow the breeder to capture and efficiently exploit the genetic superiority contained in them, therefore, guaranteeing a sustainable and consistent genetic progress that cannot be observed nowadays; **g)** once a given combination of inbred lines is found (good performing hybrids) the same genotype could be produced at first using botanical seed, and from there by vegetative means. This implies not only a faster multiplication rate but also cleaner genetic stocks (from the phytosanitary point of view); **h)** the system allows for the identification of useful recessive traits.

References.

Allard, R.W. 1960. Principles of Plant Breeding. John Wiley & Sons. 485 pp.

Bos, I. and P. Caligari. 1995. Selection Methods in Plant Breeding. Chapman and Hall 347 pp.

Griffing, B. 1975. Efficiency changes due to use of doubled-haploids in recurrent selection methods. Theoretical and Applied Genetics 46:367-386.

Hallauer, A.R. and J.B. Miranda Fo. 1988. Quantitative Genetics in Maize Breeding. Second Edition. Iowa State University Press. Ames, IO, USA. 468 pp.

Hedrick, P.W. 1983. Genetics of populations. Science Books International. Van Nostrand Reinhold Company. 629 pp.

Lamkey, K.R. and J.E. Staub. 1998. Concepts and breeding of heterosis in crop plants. Crop Science Society of America, Madison, WI, USA. 127 pp.

Neuffer, M.G., E.H. Coe, and S.R. Wessler. 1997. Mutants of maize. Cold Spring Harbor Laboratory Press. USA. 468pp.

Simmonds, N.W. and J. Smartt. 1999. Principles of crop improvement. Blackwell Science.