

Evaluation of cassava (*Manihot esculenta* Crantz) genotypes reveals great genetic variability and potential selection gain

Matoso Francisco Avijala^{1,2}, Leonardo Lopes Bhering³, Leonardo de Azevedo Peixoto^{*3}, Cosme Damião Cruz³, Pedro Crescêncio Souza Carneiro³, Constantino Estevão Cuambe¹, Anabela Zacarias¹

¹Institute of Agricultural Research of Mozambique - IIAM, Avenida das FPLM nr. 1698, Caixa Postal nr. 3598, Maputo- Moçambique

²Genetics and Breeding, CNPq Scholarship, Federal University of Viçosa (UFV), Minas Gerais-36570-000, Brazil

³Department of General Biology, Federal University of Viçosa, Viçosa Campus, Minas Gerais-36570-900, Brazil

*Corresponding author: leonardo.bhering@ufv.br

Abstract

The aim of this study was to estimate genetic and phenotypic parameters of cassava in order to contribute in the orientation of selection strategies of breeding programs in cassava. In this sense, genetic and phenotypic variances, heritability coefficient, coefficients of genetic and phenotypic correlation, and expected gains from selection were estimated. Twenty-one genotypes of cassava collected from Mozambique were evaluated in a complete randomized block design (CRBD) with three replications for one year. The following traits were evaluated: plant height (PH); first branch height (FBH); shoot biomass weight (SBW); number of roots per plant (NRP); tuberous root yield (TRY); production of commercial roots (PCR); harvest index (HI) and dry matter content (DMC). Statistical analyses were carried out using the GENES program. Genetic and phenotypic parameters were estimated among the evaluated traits. The coefficient of genetic variation ranged from 8.86 to 54.74%, and the highest CV_g values were for shoot biomass weight (54.74%), and mean number of roots per plant (47.71%). The ratio of genetic and environmental variation coefficients was higher than 1 for 6 of the 8 evaluated traits. These same traits presented high values for heritability, indicating that most of the observed variation is of genetic nature. The genotypic correlations were higher than the phenotypic correlations in all cases. It was possible to identify positive and highly significant genetic correlations among SBW vs. TRY ($r_g = 0.85$), and NRP vs. TRY ($r_g = 0.94$). The graphical dispersion analysis, based on the first two canonical variables allows the formation of six clusters. Selection based on shoot biomass weight had indirect gain greater than direct gain for root yield. Therefore, based on these results, it was found that shoot biomass weight and number of roots per plant can be used as auxiliary criteria for selection of more productive cassava genotypes.

Keywords: *Manihot esculenta* Crantz, quantitative genetics, heritability, plant breeding, biometrics.

Abbreviation: PH_Plant Height; FHB_First Branch Height; SBW_Shoot Biomass Weight; NRP_Number of Roots per Plant; TRY_Tuberous Root Yield; PCR_Production of Commercial Roots; HI_Harvest Index; DMC_Dry Matter Content; CV_g _Genetic variation Coefficient; r_f _Phenotypic Correlation; r_g _Genotypic Correlation; PMG_{xy} and PMR_{xy} _mean products associated with the effects of genotypes and residue in relation to x and y characteristics; QMR_x and QMR_y _mean squares associated with the residue effects of x and y traits, respectively; QMG_x and QMG_y _mean squares associated with the effects of genotypes with respect to x and y characteristics, respectively; σ_{gxy} _Estimator of genotypic covariance between x and y; $\sigma^2_{g^x}$ and $\sigma^2_{g^y}$ _estimators of genotypic variances of x and y characteristics, respectively; GS_Gain Selection; h^2 _heritability; DS_Differential of selection; ANOVA_Analyses of Variance.

Introduction

Cassava is a native perennial crop of tropical America, and Brazil is its center of origin and diversity (Vavilov, 1951). It was widely disseminated by the Portuguese in the sixteenth and seventeenth centuries to tropical and subtropical areas of Africa, Asia and the Caribbean (Cock, 1985). It is a staple food for more than 700 million people in the tropics and sub tropics (FAO, 2009). Cassava's root is the third most important energetic food in the tropics, after rice and maize (FAO, 2009). It occupies about 17 million hectares in the world, which are located entirely in developing countries, with production of 185 million tons of tuberous roots (FAO, 2011). In Mozambique, cassava is the most important root crop, and it is grown throughout the country (Jorge, 1995). The amount of cassava produced per year exceeds corn in

terms of total calorie supply and market value. Cassava produced by small and medium producers represents 50% of national agricultural production, and contributes 55% to the potential relief of poverty, and to the improvement of familiar income. However, researches are limited, and the root yield achieved in the various regions of the country is low. Agronomic characterization and obtainment of clones resistant to disease and pests, and with high production capacity and superior agronomic traits, and which are able to replace the traditional cultivars are means used to increase crop yield in Mozambique. The objectives of a cassava breeding program are set according to the needs of production, processing and market, based on the resistance to pests and diseases, and especially on the increment of root

yield (Okogbenin et al., 2007). However, yield is a complex trait resulting from the expression and association of different components (Ceballos et al., 2004), which makes it necessary to understand and study the degree of association among these traits, as well as the estimates of correlation coefficients. Obtaining estimates of genetic and phenotypic parameters such as heritability, genetic and phenotypic correlations, and expected gains from selection is of great importance in breeding programs, as they allow the decision-making related to the choice of the most appropriate method, the traits that should be selected in early and advanced stages of a program, and also the weight that should be assigned to each character, separately or together (Cruz, 2005). Correlations measure the degree of association between two variables (Cruz, 2005); high correlation between two traits allows selection of a trait of interest through another correlated trait which is easier to be measured, especially when the former is of complex inheritance. This strategy allows faster progress on the use of direct selection in order to optimize the gains in breeding programs (Odjugo, 2008; Cruz et al, 2012). Correlation coefficients are appropriate to evaluate the association among traits since they are dimensionless and allow comparisons among different pairs of traits, unlike the covariance (Rubaihayo et al., 2001). The knowledge of the association among agronomic and morphological traits allows the breeder to predict the consequences of the traits' simultaneous change. This can result in greater efficiency in the selection of the traits to be improved, as well as in time, physical, financial and human resources savings, in relation to isolated selection of a particular trait. Phenotypic correlations have genetic and environmental causes; however, only the genetic associations are inheritable. Thus, the phenotypic correlation has little practical value, and it must be separated into causes of genetic and of environmental origin (Cruz, 2005). This study is important for the improvement of a population or cultivar that is directed to a set of traits simultaneously. The aim of this study was to estimate genetic and phenotypic parameters, such as: genetic and phenotypic variances, heritability coefficients, coefficients of genetic and phenotypic correlation, and expected gains from selection, in order to contribute in the orientation of selection strategies of breeding programs.

Results and Discussion

Analysis of variance

Prior to the analysis of variance (ANOVA), data were submitted to normality test using the Lilliefors test, which was reasonable to study the data through normal distribution ($p > 0.05$). The analysis of variance revealed significant results, at 1% probability by the test "F", among genotypes, for all the evaluated phenotypic traits. Thus, the different behavior of genotypes indicates the presence of variability for the agronomic traits studied. The presence of wide variability among genotypes was expected, since the evaluation was carried out with genotypes from different origins and levels of improvement, as previously reported by Farias Neto et al. (2013) and Nick et al. (2008). Genetic gain depends on the heritability of the trait under selection, on the intensity of the selection, and on the control of environmental conditions. The higher the expression level of genetic variability in the environment and if the proportion of this genetic variability is mostly due to additive effects, the greater the gains estimates for the next generation (Akinwale et al., 2010). Table 1

presents the estimates of genetic and phenotypic parameters of 8 quantitative traits evaluated in 21 cassava genotypes. Through F test of the analysis of variance, it was observed that there is genetic variability for the agronomic characteristics analyzed at 1% probability. There is great variability among the data, which is verified by the range of the results of the evaluated traits: plant height - 1.50 to 2.85; first branch height- 0.04 to 2.00; shoot biomass yield: 2.78 to 51.65; mean number of roots per plant - 0.39 to 14.14; mean tuberous roots yield - 2.78 to 58.89; production of commercial roots - 0.00 to 48.89; harvest index - 39.66 to 81.16; dry matter content - 15.15 to 43.25, with the respective means of 2.14 m; 1.18 m; 21.65 t ha⁻¹; 6.03; 29.61 t ha⁻¹; 20.28 t ha⁻¹; 58.78%, and 28.83%.

Estimate variance components

Table 1 shows that the phenotypic variance (σ_f^2) happens due to genetic (σ_g^2) rather than environmental causes (σ_e^2). This fact is proven by the highly significant heritability presented by the evaluated traits. Then, by means of variance components, it was possible to obtain the estimates of heritability (h^2). Heritability values were high for all the evaluated traits, except for HI, which was 60.24%, and which is considered on average. These heritability values are consistent with those obtained by Ntawuruhunga and Dixon (2010). Among the traits related to roots production, emphasis is given to the mean number of roots per plant and root yield, which presented the highest heritability, 91.5% and 87.4%, respectively (Table 1). Similar results were found by Brar and Sukhija (1980), who found high heritability value for roots yield (85%) in studies of estimates of genetic parameters in cassava. On the other hand, among the shoot traits (SBW, PH and FBH), shoot biomass yield was the one with the highest heritability (97.0%). Thus, the greater the heritability, the greater the genetic contribution to the total variability, which is desirable in a breeding program. Ceballos et al. (2004) argue that traits with low heritability tend to hinder the selection process due to the great environment influence. According to Allard (1971), heritability estimates together with the genetic variation coefficient (CVg) provide a better insight into the genetic progress expected with selection. Vencovsky and Barriga (1992) reported that in order to have a true idea of the situation of each trait that is intended to be improved, it is necessary to analyze the CVg together with the CVe, by CVg/CVe ratio of each trait, which must be greater than 1. Cruz et al. (2012) highlight that high estimates of heritability and CVg/CVe ratio close to or above 1 portray a very favorable situation for selection. The coefficients of genetic variation (CVg) in general indicated high genetic variability among the analyzed traits, since the genotypes evaluated have different genetic backgrounds and different improvement levels. These estimates ranged from 8.86 (PH) to 54.74% (SBW), and shoot biomass yield (54.74%) and number of roots per plant (47.71%) showed the highest coefficients of genetic variation. According to Valois and Son Miranda (1984), the knowledge of genetic variation coefficient is very important in the evaluation of genetic variability since it indicates the range of genetic variation of a character, with a view to assessing their use potential. The values found in this study show a possible environmental influence on the conditions of the experiment.

Table 1. Estimates of genetic parameters and general statistics for: plant height (PH), first branch height (FBH), shoot biomass (SBW), mean number of roots per plant (NRP); tuberous roots yield (TRY), production of commercial roots (PCR), harvest index (HI), dry matter in roots (DMC), evaluated in 21 cassava genotypes.

Parameters*	Agronomic Traits							
	PH	FBH	SBW	NRP	TRY	PCR	HI	DMC
Mean	2.14	1.18	21.65	6.03	29.61	20.28	58.76	28.83
Minimum	1.50	0.04	2.78	0.39	2.78	0.00	39.66	15.55
Maximum	2.85	2.00	51.67	14.44	58.89	48.89	81.16	43.25
σ^2_f	0.044	0.158	186.380	10.623	161.043	92.809	47.869	35.762
σ^2_e	0.007	0.014	5.568	0.903	20.341	12.361	19.033	9.073
σ^2_g	0.037	0.144	180.812	9.720	140.702	80.448	28.838	26.689
$h^2(\%)$	84.28	90.97	97.01	91.45	87.37	86.68	60.24	74.63
Ci	64.12	77.06	91.54	78.20	69.75	68.45	33.55	49.51
$CV_g(\%)$	8.86	31.69	54.74	47.71	37.08	40.77	9.35	18.04
CV_g/CV_e	1.34	1.83	3.29	1.89	1.52	1.47	0.71	0.99

* σ^2_f : mean phenotypic variance; σ^2_e : mean environmental variance; σ^2_g : mean genotypic variance; h^2 : heritability (mean of genotypes); ci: intraclass correlation coefficient; $CV_g\%$: genetic variation coefficient; CV_g/CV_e : ratio between genetic and environmental coefficient of variation, based on the means of genotypes. PH - plant height, FBH - first branch height, SBY - shoot biomass yield, NRP - mean number of roots per plant, PCR - production of commercial roots, HI- harvest index, DMC - dry matter content, RY - roots yield.

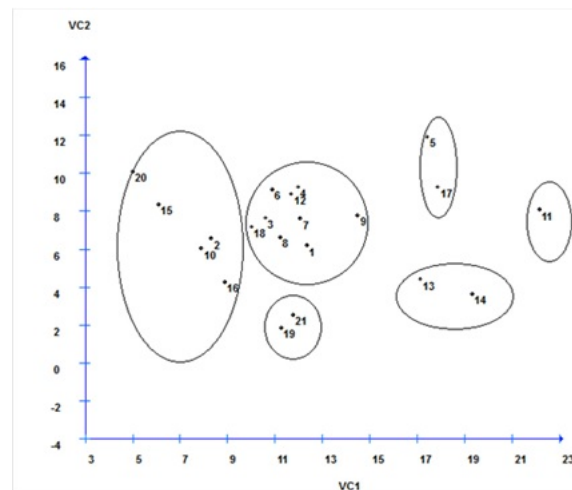


Fig 1. Graphical dispersion of scores for the two representative axes of the first two canonical variables (CV_1 e CV_2) for eight agronomic traits in 21 cassava genotypes: 1 = MzMg10/040; 2 = MzMg10/042; 3 = MzMg10/075; 4 = MzMg10/083; 5 = MzMg10/096; 6 = Mg10/107; 7 = MzMg10/111; 8 = MzMg10/118; 9 = MzMg10/119; 10 = MzMg10/131; 11 = MzMg10/162 12 = MzMg10/168; 13 = MzMg10/240; 14 = MzMg10/314; 15 = MzMg10/354; 16 = MzMg10/466; 17 = MzMg10/630; 18 = Mokhalana; 19 = Orera; 20 = Tomo; 21 = Variuaya.

The ratio of the coefficients of genetic and environmental variation (CV_g/CV_e) was greater than 1 for 6 of the 8 evaluated traits. These same traits presented high heritability values, indicating that most of the observed genetic variation is of genetic nature. This reflects a very favorable situation for the selection of important agronomic traits in cassava, such as shoot yield and number of roots per plant. Bhering et al. (2013) and Cruz et al. (2012) emphasize that CV_g/CV_e ratio can be used as indicative index of the easiness degree of progenies selection for each trait. According to Cruz et al., (2012), when heritability values (h^2) are greater than 80%, and the CV_g/CV_e ratio is greater than 1, satisfactory gains from selection can be obtained. Such situation was observed for plant height, first branch height, shoot biomass weight, mean number of roots per plant, tuberous roots yield and production of commercial roots (Table 1), indicating that simple improvement methods, such as mass selection, can be applied to provide considerable gains in the selection of these traits. Harvest index and dry matter in tuberous roots had

lower CV_g/CV_e ratio estimates, i.e., less than 1.0. According to Vencovsky and Barriga (1992), when the coefficients of CV_g/CV_e ratio are less than one, it indicates that the selection process should be carried out very carefully, using statistical genetic procedures with sufficient sensitivity.

Estimate of the phenotypic and genotypic correlations

Correlation measures the degree of association between two variables and is generally used to allow selection of a trait of complex inheritance through another correlated trait of easier measurement. This strategy allows faster progress in relation to the use of direct selection and the increase of gains from selection in breeding programs (Odjugo, 2008; Cruz, 2005; Cruz et al, 2012). Table 2 shows the results of phenotypic and genotypic correlations among the traits evaluated in cassava. Out of the 28 evaluated phenotypic correlations, nine significant correlations were observed, with values ranging from 0.44 to 0.93. This indicates a high degree of

Table 2. Estimates of genotypic (r_g - upper diagonal) and phenotypic (r_f - lower diagonal) correlations for plant height (PH), first branch height (FBH), shoot biomass yield (SBW), mean number of roots per plant (NRP); mean tuberous roots yield (TRY), production of commercial roots (PCR), harvest index (HI) and dry matter content (DMC), evaluated in 21 cassava genotypes.

Traits	PH	FBH	SBW	NRP	TRY	PCR	HI	DMC
PH	-	0.73 ⁺⁺	0.42	0.18	0.15	0.26	-0.59 ⁺	0.01
FBH	0.69 ^{**}	-	0.02	0.10	0.19	0.16	-0.03	0.25
SBW	0.31	0.02	-	0.85 ⁺⁺	0.85 ⁺⁺	0.74 ⁺⁺	-0.68 ⁺⁺	0.03
NRP	0.17	0.08	0.80 ^{**}	-	0.94 ⁺⁺	0.76 ⁺⁺	-0.24	0.13
TRY	0.15	0.07	0.80 ^{**}	0.93 ^{**}	-	0.92 ⁺⁺	0.39	0.17
PCR	0.26	0.14	0.70 ^{**}	0.75 ^{**}	0.91 ^{**}	-	-0.17	0.27
HI	-0.44 [*]	-0.02	-0.61 ^{**}	-0.11	0.33	-0.03	-	0.15
DMC	0.00	0.17	0.02	0.07	0.10	0.19	0.06	-

** And * significant at 1 and 5% probability by the t test, respectively.

++ And + significant at 1 and 5%, respectively, by the bootstrap method with 10,000 simulations.

PH - plant height, FBH - first branch height, SBY - shoot biomass yield, NRP - mean number of roots per plant, PCR - production of commercial roots, HI- harvest index, DMC - dry matter content, RY - roots yield.

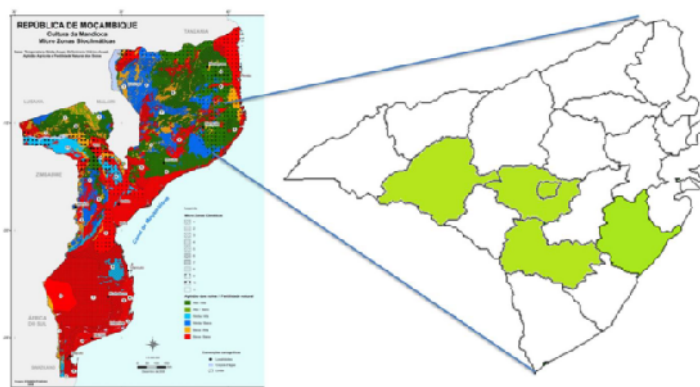


Fig 2. Map of Mozambique showing Nampula province, where the experiments were carried out and samples were collected.

Table 3. Accumulated eigenvalues corresponding to percentages of variation, explained by the canonical variables (CV), and weighting coefficients (eigenvectors) of eight agronomic traits evaluated in twenty one cassava genotypes evaluated in Agricultural Research Institute of Mozambique (IIAM) in the district of Mogincual, Nampula province, located in the North of Mozambique.

VC	Accumulated eigenvalues	Weighting coefficients associated to:							
		PH	ALTPR	SBY	NRP	RY	PCR	HI	DMC
VC ₁	58.27	-0.300	-0.244	1.596	0.956	-0.803	-0.703	1.239	0.342
VC ₂	78.02	0.505	0.703	0.059	-0.390	0.190	0.389	-0.445	-0.345
VC ₃	86.06	-0.621	0.660	0.443	0.525	-0.652	-0.622	0.943	-0.520
VC ₄	92.06	-0.592	0.407	-0.025	-0.359	-0.052	1.058	0.194	0.154
VC ₅	96.55	0.225	0.052	-0.317	0.899	-0.196	-0.307	0.071	0.681
VC ₆	98.21	0.294	-0.335	-0.812	0.188	1.268	-0.078	-1.013	-1.450
VC ₇	99.38	0.073	0.066	-0.186	-0.942	0.307	-1.194	-0.395	0.204
VC ₈	100.00	0.329	-0.132	0.803	-0.235	-0.937	-0.087	1.802	-0.028

PH - plant height, FBH - first branch height, SBY - shoot biomass yield, NRP - mean number of roots per plant, PCR - production of commercial roots, HI- harvest index, DMC - dry matter content, RY - roots yield.

Table 4. Estimates of the original mean (\bar{X}_o) heritability among genotypes means (h^2_m) and gain with selection (GS%), obtained by direct and indirect selection of 21 cassava genotypes.

Traits ¹	\bar{X}_o	\bar{X}_s	h^2_m	GS (%) ²								Total
				PH	FBH	SBW	NRP	TRY	PCR	HI	DMC	
PH	2.14	2.28	84.28	11.71	19.94	23.33	-3.82	-0.66	7.69	-6.51	-2.69	48.99
FBH	1.18	133	90.97	8.53	43.62	2.41	9.05	1.22	4.20	-1.33	0.74	68.44
SBW	21.65	37.07	97.01	1.28	-9.15	91.56	50.54	36.85	30.21	-11.09	2.06	192.26
NRP	6.03	9.03	91.45	5.02	6.56	74.30	61.40	39.40	31.47	-5.95	-1.25	210.95
TRY	29.61	45.04	87.37	3.48	8.02	69.15	45.45	44.03	50.98	-3.32	0.12	217.91
PCR	20.28	32.19	86.68	5.78	18.74	47.56	29.84	37.17	57.64	-0.57	0.62	196.78
HI	58.76	69.46	60.24	-7.58	-9.62	-35.29	10.84	8.95	3.65	12.94	8.54	-7.57
DMC	28.83	29.07	74.63	4.71	11.82	15.69	20.86	9.94	9.99	-0.89	17.97	90.09

Plant height (PH), first branch height (FBH), shoot biomass (SBW), mean number of roots per plant (NRP); tuberous roots yield (TRY), production of commercial roots (PCR), harvest index (HI), dry matter in roots (DMC). 2 selection criterion: increase in all traits. * Gains by direct selection in bold.

association. According to Shimakura and Ribeiro Junior (2006) (Table 3), in four of these significant correlations, estimate values were above 0.70, indicating high correlation. Regardless of the statistical significance of the *t* test, 7.1% of the phenotypic correlations were considered very high; 14.3% high; 10.7% moderate; 10.7% low; and 57.1% very low. Genotypic correlations were higher than phenotypic correlations in all the ratios evaluated (Table 2), and major contributions were of genetic factors in relation to environmental factors in correlations among traits. This is favorable to the selection process, since the genetic effects overlap the environmental ones in the phenotype manifestation. Odjugo (2008) points out that highly significant correlations are positive for indirect selection in breeding programs. In this way, in genotypic classification, 7.1% were considered very high; 17.9% high; 10.7% moderate; 17.9% low; and 46.4% very low. Due to the similarity and to its practical value in breeding studies, only the genetic correlations of greater significance and interest will be analyzed in more detail. Shoot biomass yield is an important trait in cassava improvement since it is closely related with the plant's ability to produce stem cuttings with the possibility of using the shoots as fodder in animal feed. It showed positive and high correlation with average number of roots per plant ($r_g = 0.85$), tuberous roots yield ($r_g = 0.85$) and production of commercial roots ($r_g = 0.74$), indicating that the improvement of a trait promotes the increment of the other. On the other hand, the direction and the significance of correlations indicate the possibility of selection of plants with higher number of roots per plant and higher tuberous roots yield through indirect selection of plants with higher shoot biomass yield. However, to know the shoot weight, it is necessary to cut and weight it, and therefore, it is a destructive method. Still, indirect selection through SBW may be carried out, because even if the shoot is removed, it has the ability to sprout and produce seeds due to the reserves stored in the root. These results are in agreement with those obtained by Ntawuruhunga and Dixon (2010), who found positive and high genetic correlations between tuberous roots yield and shoot biomass yield of 0.75. The presence of positive correlation between shoot biomass yield and tuberous root yield indicates that the increase of shoot material leads to higher tuberous roots yield, i.e., selection for higher shoot biomass yield provides indirect gains of tuberous roots yield and vice versa. On the other hand, the presence of a positive correlation may indicate a balance in the source-sink relationship. Fukuda et al. (2002) report the importance of shoot biomass yield when the goal is the indirect selection for tuberous root yield based on this trait. However, Peixoto et al. (2005) observed negative correlation among these traits and attributed this behavior to an imbalance in the source-sink relationship. The genotypes that presented higher plant height also showed greater shoot biomass yield, which was already expected. These traits present correlation significance of 0.42. These results indicate that the researcher can also select cassava plants using plant height. The fact that PH correlates with SBW, and presents great measurement easiness of data can be the basis for indirect selection of progenies with higher shoot weight. Similar results have been reported, and Ntawuruhunga and Dixon (2010) detected high correlation between plant height and shoot biomass yield. In turn, Fukuda et al. (2002), studying correlations among cassava traits, found that plant height positively correlated with tuberous roots yield, but at a lower rate compared to shoot biomass yield, similar to the present study. With respect to plant height and first branch

height, correlation ($r_g = 0.73$) was positive, indicating that the greater the plant height, the highest the insertion of the first branch. The first branch insertion height is an important trait, since producers prefer cultivars whose architecture is expressed in greater height of the first branch, and thus makes it easier when it comes to intercrops, weed control, and especially harvest. The mean number of roots per plant showed significant and positive correlation with tuberous roots yield ($r_g = 0.94$) and with the production of commercial roots ($r_g = 0.76$). Ntawuruhunga and Dixon (2010) found correlation coefficient of 0.75 and 0.68 between mean number of roots per plant and tuberous roots yield, respectively. The harvest index is the tuberous roots yield efficiency, and it negatively correlates with shoot biomass yield ($r_g = -0.68$), and positively correlates with tuberous roots yield ($r_g = 0.39$). It is expected that an increase in shoot biomass yield consequently reduces harvest index since it represents the ratio between tuberous roots weight and total plant weight. On the other hand, increasing tuberous roots yield induces higher harvest indexes (Fukuda and Guevara, 1998).

Canonical analysis

Table 3 lists the eigenvalues and the coefficients associated with the variables analyzed, expressing their relative importance in the genetic diversity study. Variables with little variability or which are correlated with other variability taken into account in the present study had coefficients of great magnitude in the last eigenvectors. Considering the last canonical variables, which represent less than 5% of the total variation, it is observed that, in future studies, harvest index (HI), production of commercial roots (PCR), dry matter content (DMC) and mean number of roots per plant (NRP) could be discarded. The use of the methodology of scores projections in 2D Cartesian plane in the genetic diversity study aims to identify more dissimilar genotypes in two-dimensional scatter plots, seeking to simplify results interpretation, in addition to the methods previously mentioned (Cruz et al. 2011). According to Cruz et al (2012), researchers have opted for graphical representation when the first two canonical variables have about 80% of the total variation. Only when this limit is not reached in the first two components, the analysis is complemented by the graphic dispersion in relation to the third component. However, in this study, the analysis revealed that the first two canonical variables accounted for about 78.02% of the total variation (58.27% for the first, and 19.75% for the second) (Table 3), which is satisfactory in this case. However, this analysis did not explain the genetic diversity by graphical dispersion of about 80% of the total variance, in relation to axes represented by these two canonical variables, with a negligible degree of distortion caused by the distance between genotypes. The graphical dispersion analysis, based on the first two canonical variables (Fig 1) allows the formation of six clusters. It was observed that Tomo and MzMg10/162 were the most distant genotypes in relation to the others, while MzMg10/083 and MzMg10/107 were the closest. Thus, they are reliable in the identification of parents with high diversity in order to point out promising crossings.

Estimate gain selection

Table 4 shows the estimated predicted gains obtained by the direct and indirect selection of 21 cassava genotypes, used in order to increase significance. This selection criterion was

Table 5. Identification list and pedigree of 21 cassava genotypes used for the study.

Varieties	Pedigree	Origin	Special attributes
Mokhalana	PAN 38 (OP)	IIAM Improved Variety	CBSD resistance, high yield
MzMg10/040	Mulaleia x IMM30025	IIAM Improved clone	CBSD and CMD resistance, high dry matter content
MzMg10/042	Likonde x Mulaleia	IIAM Improved clone	CBSD and CMD resistance, high dry matter content
MzMg10/075	Mulaleia (OP)	IIAM Improved clone	CBSD and CMD resistance, high dry matter content
MzMg10/083	Likonde x Chigoma máfia	IIAM Improved clone	High yield, CBSD and CMD resistance
MzMg10/096	Likonde x Mulaleia	IIAM Improved clone	High yield, good plant type, good leaves for consumption
MzMg10/107	Likonde x Mulaleia	IIAM Improved clone	High yield, good plant type, good harvest index
MzMg10/111	Mulaleia (OP)	IIAM Improved clone	High yield, good plant type
MzMg10/118	Mulaleia (OP)	IIAM Improved clone	CBSD tolerance, CMD resistance, high dry matter content
MzMg10/129	Mulaleia x Likonde	IIAM Improved clone	High yield, good plant type
MzMg10/131	Mulaleia (OP)	IIAM Improved clone	CBSD and CMD resistance, High yield, good type of plant
MzMg10/162	Mulaleia x Likonde	IIAM Improved clone	High yield, good plant type, good harvest index
MzMg10/168	Mulaleia x Chigoma máfia	IIAM Improved clone	High yield, good plant type, high dry matter content
MzMg10/240	Mulaleia x Chigoma máfia	IIAM Improved clone	High yield, CBSD resistance, high dry matter content
MzMg10/314	Mulaleia (OP)	IIAM Improved clone	High yield, CBSD and CMD resistance
MzMg10/354	Mulaleia x Likonde	IIAM Improved clone	CBSD tolerance, CMD resistance, high dry matter content
MzMg10/466	Mulaleia x Likonde	IIAM Improved clone	CBSD tolerance, CMD resistance, high dry matter content
MzMg10/630	Mulaleia x TMS30001	IIAM Improved clone	High yield, CBSD and CMD resistance
Orera	Likonde (OP)	IIAM Improved variety	CBSD and CMD resistance, good leaves for consumption
Tomo	Tomo	Landrace	Good type of plant, good leaves for consumption
Varuiaya	IMM30025 X Chigoma máfia	IIAM Improved Variety	CBSD Tolerance, high dry matter content

OP = Open pollination, CBSD = Cassava Brown Streak Disease, CMD = Cassava Mosaic Disease, IIAM= Agricultural Research Institute of Mozambique.

Table 6. Analysis of variance and mean square mathematical expectations for complete randomized block design.

VS	DF	SS	E(MS)	F
Blocks	b - 1	SQB	$\frac{2}{2} + \frac{g}{2} \frac{2}{b}$	
Genotypes	g - 1	SQG	$\frac{2}{2} + \frac{b}{2} \frac{2}{g}$	QMG/QMR
Residue	(b - 1)(g - 1)	SQR		
Total	gb - 1	SQT0		

Where b = number of replications and g = number of genotypes

adopted to get larger plants with high first branch; high number of roots per plant; harvest index aiming to obtain more productive genotypes with high dry matter content in tuberous roots and better adapted to the technological levels adopted by cassava producers in various regions of the country. Direct selection gains of all traits were always higher than the indirect selection gains (Table 4), and were close when the traits were highly correlated, as SBW vs. TRY and NRP vs. TRY. According to Falconer (1981), indirect selection can promote greater gains than direct selection, if the auxiliary trait presents higher heritability than the main one, and if the genetic correlation between them is positive and highly significant. Therefore, it was found that shoot biomass weight and number of roots per plant can be used as auxiliary criteria for selection of cassava genotypes. Shoot biomass yield (SBW) had the greatest single gain, and it presents higher coefficient of genetic variation (Table 1). Its selection may be more accurately, since the mean phenotypic value presented by genotypes in representing their genetic values is more reliable. Selection of TRY presented the highest estimate of the total gain. However, this strategy has resulted in negligible gains in DMC, and negative gains in HI, which are relevant agronomic traits, and usually positively correlated with yield.

Materials and Methods

Plant materials

The experiment was carried out in the experimental field of the Agricultural Research Institute of Mozambique (IIAM), in Mogincual district, province of Nampula, situated in the

North of Mozambique (Lat. 15°34'S, Long. 39°45' W, at 138m asl). The climate is Cwa type (humid temperate climate, with dry winter and hot summer), according to the Köppen classification. The district has annual 80% relative humidity, 26° C temperature (ranging from 17° C minimum average and 30 ° C maximum average), and the average annual rainfall ranges from 800 to 1,000 mm, with 65-70% of it concentrated in the period from December to March.

It was used twenty-one cassava genotypes from Mozambique. For experimental establishment, it was used a randomized block design with the twenty-one cassava genotypes, planted in three replications. The genotypes were selected for their contrasting agronomic performance, and also for being preferred by cassava farmers in northern region of Mozambique. The genotypes used in the experiment are shown in table 5. Fig 2 shows the location where each genotype was collected.

Experimental design

The experimental plots consisted of five rows with five plants each, spaced 1.0 m between rows and 1.0 m between plants. The useful plot area consisted of three rows and three central plants, covering an area of 9.0 m², consisting of 9 plants per plot. After 12 months plants emergence, data from eight quantitative traits were collected at harvest: plant height (PH), in meters; first branch height (FBH) in meters; shoot biomass yield (SBY) in t ha⁻¹; mean number of roots per plant (NRP); roots yield (RY) in t ha⁻¹; production of commercial roots (PCR) in t ha⁻¹; harvest index (HI), in%; and dry matter content (DM), in%.

Statistical analysis

Statistical analyses were carried out using the statistical program GENES (Cruz, 2013). Analysis of variance, F test, mean, and experimental variation coefficient were used in every evaluated trait. Complete randomized block design was used for the analysis of variance. The mathematical expectations of mean squares are shown in Table 6. The experimental data of the evaluated traits were submitted to analysis of variance and covariance using the GENES software (Cruz, 2013). Based on the analyses of variance, the components and the genetic parameters were estimated, according to Cruz et al. (2012): mean phenotypic variance, mean genotypic variance, mean environmental variance, heritability in the broad sense based on the genotypes mean, coefficient of genetic variation, and the ratio between the coefficients of genetic and environmental variations. Phenotypic (r_f) and genotypic (r_g) correlation coefficients among the evaluated traits were estimated by the following expressions (Cruz et al, 2012.):

$$r_f = \frac{PMG_{xy}}{\sqrt{QMG_x QMG_y}} \quad e \quad r_g = \frac{g_{xy}}{\sqrt{\sigma_{gx}^2 \sigma_{gy}^2}},$$

being:

$$g_{xy} = \frac{PMG_{xy} - PMR_{xy}}{b}$$

$$\sigma_{gx}^2 = \frac{QMG_x - QMR_x}{b}$$

$$\sigma_{gy}^2 = \frac{QMG_y - QMR_y}{b}$$

Where,

PMG_{xy} and PMR_{xy} - mean products associated with the effects of genotypes and residue in relation to x and y traits; QMR_x and QMR_y - mean squares associated with the residue effects of x and y traits, respectively; QMG_x and QMG_y - mean squares associated with the effects of genotypes with respect to x and y traits, respectively; σ_{gxy} - Estimator of genotypic covariance between x and y; σ_{gx}^2 and σ_{gy}^2 - estimators of genotypic variances of x and y traits, respectively.

The t statistic was used to evaluate the phenotypic correlation coefficients (r_f), and the bootstrap method with 10,000 simulations was used to evaluate genotypic correlation coefficients (r_g). The classification described by Shimakura and Ribeiro Junior (2006) was used for evaluating the correlation significance, as follows: 0.0 to 0.19 - very low; from 0.20 to 0.39 - low; 0.40 to 0.69 - moderate; 0.70 to 0.89 - high; and 0.90 to 1.00 - very high.

The analysis based on canonical variables technique was also used to study genetic diversity among cassava genotypes in order to better illustrate these differences, and to verify the accordance of the results obtained by these different analysis methods, which allows more reliable interpretation of the results.

The expected genetic gains by direct selection were estimated by the expression: $GS = h^2 \times DS$, where h^2 is the heritability coefficient; DS is the differential screening, which is given by $DS = \bar{X}_o - \bar{X}_s$ where \bar{X}_s is the mean of the selected genotypes, and \bar{X}_o is the original mean of genotypes.

Conclusion

We concluded that there is great genetic variability among genotypes for the evaluated agronomic traits, with great potential for selection of a group of genotypes to be used either as cultivars or for hybridization. Medium to high heritability values, as well as the coefficient of genetic variation and the ratio between coefficient of genetic variation and high coefficient of environmental variation indicate that the selection of genetic material may be carried out by simple selection methods on the basis of little environmental influence; genotype correlations were higher than the phenotypic correlations in all cases, indicating that the genetic factors contributed more than the environmental ones for correlations. Genotypic correlations SBW vs. TRY (0.85); NRP vs. TRY (0.94) and PCR vs. TRY (0.92) were positive and of highly significance, showing that the selection of a trait positively influences the others. Lastly, the number of roots per plant and shoot biomass yield can be used as auxiliary criteria for selection of cassava genotypes.

Acknowledgements

To the Ministry of Science and Technology (MCT), Mozambique, and to the National Council for Scientific and Technological Development (CNPq), Brazil, for the exchange program for graduate students, and for the scholarship granted for master's degree course. To FAPEMIG for providing resources for implementation of the Nucleo de Biometria (Biometrics Laboratory), where the analyses were carried out.

References

- Akinwale MG, Akinyele BO, Dixon AGO, Odiyi AC (2010) Genetic variability among forty-three cassava genotypes in three agro-ecological zones of Nigeria. J of Plant Breed Crop Sci. 2: 104-109.
- Allard RW (1971) Princípios do melhoramento genético das plantas. Edgard Blücher Ltda, São Paulo.
- Bhering LL, Barrera CF, Ortega D, Laviola BG, Alves AA, Rosado TB, Cruz CD (2013) Differential response of Jatropha genotypes to different selection methods indicates that combined selection is more suited than other methods for rapid improvement of the species. Ind Crop and Prod. 41:260-265.
- Brar JS, Sukhija BS (1980). Variability, heritability and genetic advance in carrot (*Daucus carota* L.). J of Res Pun Agri Uni. 17: 442-443.
- Ceballos H, Iglesias CA, Pérez JC, Dixon AGO (2004) Cassava breeding: opportunities and challenges. Plan Mol Bio. 56: 503-516.
- Cock JH (1985) Cassava: new potential for a neglected crop. Westview Press, Boulder, USA.
- Cruz CD (2013) GENES - a software package for analysis in experimental statistics and quantitative genetics. Acta Sci Agron. 35: 271-276.
- Cruz CD, Regazi AJ, Carneiro PCS (2012) Modelos biométricos aplicados ao melhoramento genético, fourth ed. UFV, Viçosa.
- Cruz CD, Ferreira MF, Pessoni AL (2011) Biometria aplicada ao estudo da diversidade genética. UFV, Viçosa, Minas Gerais.
- Cruz CD (2005) Princípio de genética quantitativa. UFV, Viçosa.

- Farias Neto JT, Moura EF, Resende MDV, Celestino Filho P, Augusto SG. Genetic parameters and simultaneous selection for root yield, adaptability and stability of cassava genotypes. *Pesqui Agropecu Bras.* 48: 1562-1568.
- Falconer DS (1981) *Introduction to quantitative genetics*, second ed. Longman, London.
- FAO (2011) Available in <<http://www.fao.org>>. Accessed in: August 30th, 2014.
- FAO (2009) Food and agriculture organization of the united nations. Available in: <<http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor>>. Accessed in: June 25th, 2014.
- Jorge MAB (1995) Germplasm collection and characterization of cassava and sweet potato in Mozambique. *Roots.* 2: 4-6.
- Fukuda WMG, Guevara CL (1998) *Descritores morfológicos e agronômicos para a caracterização de mandioca (Manihot esculenta Crantz)*: EMBRAPA-CNPMPF, Cruz das Almas, Brazil.
- Fukuda WMG, Silva SO, Iglesias C (2002) Cassava breeding. *Crop Breed Appl Biot.* 2: 617-638.
- Nick C, Carvalho M, Assis LHB, Carvalho SP (2008) Genetic dissimilarity in cassava clones determined by multivariate techniques. *Crop Breed Appl Biot.* 8: 104-110.
- Ntawuruhunga P, Dixon AGO (2010) Quantitative variation and interrelationship between factors influencing cassava yield. *J Appl Biosci.* 26: 1594 – 1602.
- Peixoto JR, Bernardes SR, Santos CM, Bonnas DS, Fialho JF, Oliveira JA (2005) Desempenho agronômico de variedades de mandioca mansa em Uberlândia. *Rev Bras Mand.* 8: 19-24.
- Okogbenin E, Porto MCM, Egesi C, Mba C, Espinosa E, Santos LG, Ospina C, Marín J, Barrera E, Gutiérrez J, Ekanayake I, Iglesias C, Fregen MA (2007) Marker-assisted introgression of resistance to cassava mosaic disease into latin american germplasm for the genetic improvement of cassava in Africa. *Crop Sci.* 47: 1895-1904.
- Odjugo PAO (2008) The impact of tillage systems on soil microclimate, growth and yield of cassava (*Manihot utilisima*) in Midwestern Nigeria. *Afri J of Agri res.* 3:225-233.
- Rubaihayo PR, Whyte JBA, Dixon AGO, Osiru DSO (2001) Inter-relationships among traits and path analysis for yield components of cassava: a search for storage root yield indicators. *Afri Crop Sci Jour.* 9: 599-606.
- Shimakura SE, Ribeiro Junior PJ (2006) *Estatística*. Available in: www.est.ufpr.br/~paulojus/CE003/ce003.html. Accessed in July 10th 2012.
- Vavilov NI (1951) *Phytogeographic basis of plant breeding. The origin, variation immunity and breeding of cultivated plants.* *Chron Bot.* 13: 1-366.
- Valois ACC, Miranda Filho JB (1984) *Estimação de componentes de variância na cultivar de milho central mex.* *Pesqui Agropecu Bras.* 19: 479-488.
- Vencovsky R, Barriga P (1992) *Genética biométrica no fitomelhoramento.* Sociedade Brasileira de Genética, Ribeirão Preto.