QTL ANALYSIS FOR EARLY YIELD IN A PSEUDO F_2 POPULATION OF CASSAVA

E. OKOGBENIN[1, 2*], J. MARIN[1, 3] and M. FREGENE[1]

[1]International Center for Tropical Agriculture, CIAT, Km 17, Recta Cali-Palmira, AA6713, Cali, Colombia
[2]University of Ibadan, Ibadan, Nigeria
[3]Univerisdad del Tolima, Ibague, Tolima, Colombia

ABSTRACT

Genetic mapping of early bulking in a full-sib population of cassava was continued in a selfed family of 268 cassava plants derived from a single progeny of the full-sib population. The pseudo F_2 population was analysed with 122 segregating SSR markers. A previously constructed linkage map of cassava consisting of 22 linkage groups covering 1236.7 cM, with an average marker distance of 18 cM was used for this study. The F_2 population was evaluated for components of early yield, namely, dry root yield (DR) at 7 months, harvest index (HI), and weight of fresh foliage (FF). Interval mapping, with single- and two-QTL models, was used to identify QTLs. The single-QTL model identified three QTLs each for DR, FF, and HI. The two-QTL model approach identified groups of QTLs that together explained 33% for FF, 43.5% for DR and 36% for HI. The identification of QTLs involved in early yield is an important step toward understanding quantitative genetic variation of early yield and implies reconsideration of breeding strategies for improvement of this complex trait.

Key words: Linkage map, molecular marker-assisted selection, SSR markers

RÉSUMÉ

La cartographie génétique et l'epaississement précoce au sein d'une population 'full-sib' de manioc était rendue continuelle dans une famille isolée de 268 plantes de manioc dérivée à partir d'une lignée unique de la population full-sib. La pseudo-population F2 était analysée à l'aide de 122 marqueurs séparatifs SSR. Une carte de connexions de manioc précédemment établie, consistant en 22 groupes de connexions couvrant 1236,7 cM avec distance de marqueur de 18cM était utilisée pour cette étude. La population F_2 a été évalué pour les composantes de rendement précose notamment rendement de racine sèche (DR) à 7 mois, index de récoltes (HI) et poids de feuillage frais (FF). L'intervalle cartographique a été utilisé avec modèles QTL uniques et doubles en vue d'identifier les QTLs. Le modèle à QTL unique a identifié 3 QTLs chaque fois pour DR, FF et HI. L'approche du modèle à double QTL a identifié des groupes de QTLs qui, ensemble étaient responsables pour 33% pour FF, 43,5% pour DR et 36% pour HI. L'identification de QTLs impliqués dans le rendement précoce est une étape importante conduisant à la compréhension de la variation génétique quantitative de rendement précoce et implique la reconsidération des stratégies de croissance en vue d'une amélioration de ce complexe caractère.

Mots Clés: Carte de connections, séléction assistée par marquage moléculaire, marqueurs SSR

INTRODUCTION

Cassava (*Manihot esculenta* subsp. *esculenta* Crantz) is the principal or second most important source of calories for more than 500 million people (Cock, 1985; Best and Henry, 1992). As a staple food, it is the sixth most important crop worldwide (Mann, 1997). It is widely grown by rural subsistence farmers of Africa, Asia and Latin America where it can be an engine of rural development given the high perishability of its storage roots and the need to process the fresh roots close to where they are produced.

In spite of its global economic importance, it has traditionally received less researchers attention than temperate crops (Cock, 1985). As a consequence, many fundamental questions about the genetics of important traits have not been fully answered. With the development of the first molecular genetic map of cassava (Fregene *et al.,* 1997), the genetics of resistance of two of important cassava diseases, the African cassava mosaic disease (ACMD), and the cassava bacterial blight (CBB) have been studied (Akano *et al.,* 2002; Jorge *et al.,* 2000, 2001). The inheritance of early bulking and other important agronomic and morphological characters were also studied (Okogbenin and Fregene, 2002; Okogbenin and Fregene, 2003).

The first molecular genetic map and subsequent genetic analysis were carried out using a full-sib (an F_1 from non-inbred parents) intra-specific cross, based on the segregation of predominantly restriction fragment length polymorphism (RFLP) markers (Fregene *et al.,* 1997). Such a mapping population, however, cannot be used to test recessive or epistatic interactions and those may have limited utility for analysing important agronomic traits. Use of heterozygous parents alters QTL mapping by redefining "mating type" at a locus level rather than all loci in parents and also by allowing the detection of multiple QTL alleles using separate maps for each parent. The marker genotype in the F_1 progeny populations, result from the independent meioses and crossovers in the maternal and paternal parents, thus, individual maps are often constructed for each parent (Grattapaglia *et al.,* 1994; Groover *et al.,* 1994; van Eck *et al.,* 1994).

To overcome the problems associated with genetic analysis in an F_1 cross of non-inbred parents, and the use of RFLP markers which are expensive and labour intensive, genetic analysis with simple sequence repeat (SSR) markers in an F_2 mapping population was considered. SSR markers have proven to be highly polymorphic and useful as genetic markers in several plant species (Akkaya *et al.,* 1992, 1995; Senior and Heun, 1993; Jarret and Bowen, 1994; Plaschke *et al.,* 1995; Roder *et al.,* 1995; Rongwen *et al.,* 1995). SSR markers can also facilitate marker-assisted selection (MAS) in a modest cassava breeding programme once markers have been identified associated with traits of interest.

Early yield (bulking) is an important trait of cassava (Nweke *et al.,* 1994), critical to the crop's role as food security crop in sub-Saharan Africa. Genetic analysis of this trait, in the F_1 mapping population identified 3 traits, foliage, number of roots, and harvest index, and major QTLs for the traits that can be exploited in cassava breeding (Okogbenin and Fregene, 2002). We describe here genetic analysis of early bulking and component traits in an F_2 population, for better understanding of the genetics this trait and to facilitate marker-assisted improvement of yield of cassava. In addition, an attempt was made to verify QTLs detected earlier in the F_1 cross.

MATERIALS AND METHODS

Plant material. The F_1 cassava mapping population described by Fregene *et al.* (1997) was analysed for yield and related traits in 1998 and 1999. This population, which was developed from TMS 30572 (female parent and early bulking) and CM 2177-2 (male parent and late yielding) has been extensively used in mapping studies at CIAT. Based on results obtained in the F_1 population for root yield, foliage, and harvest index, and partly because of profuse flowering abilities, three F_1 individuals (K68, K145 and K150) were pre-selected for the development of an F_2 population. These individuals were selfed to generate F_2 seeds. The highest germination rate was recorded in K150 - also the most polymorphic of the 3 genotypes (with 372 seedlings), compared with 316 and 245 seedlings observed for K68 and

K145, respectively. The progeny of K150 was, therefore, selected for further analysis.

Molecular marker analysis. DNA was extracted from about 3g of fresh leaf tissue harvested from the F$_2$ plants according to Dellarporta *et al.* (1983). The population was genotyped and used for the construction of an SSR based genetic linkage map of cassava as described elsewhere (CIAT, 2002). Due to poor seedling development of certain genotypes (resulting in senescence in some few cases), 268 plants of the original 372 F$_2$ population progeny were used for molecular analysis.

Field experiment and trait evaluation. The F$_2$ seedlings were initially germinated in a screen house at CIAT headquarters in Palmira, under intensive management and care in February 2000. Seedlings were later transplanted to the field in July 2000 and harvested for woody stakes 11 months after planting (MAP). Of the 268 genotypes used for mapping analysis, only 207 with relatively sufficient stem cuttings (12 stakes) of about 25 cm long each, could be planted for the QTL mapping experiment at Santa Elena, a location 25 Km from CIAT headquarters in 2001. Severe inbreeding depression could be observed for a number of genotypes, leading to poor vigour and no woody stakes. Cassava, which is vegetatively propagated from cuttings (stakes) has low multiplication ratio normally requiring long period of time to multiply. Therefore, limited planting materials were generated from 10-month old plants for the newly developed F$_2$ population. Stakes from all 207 genotypes were planted in single row plots of 6 plants each, 0.8 m between plants and 1 m between rows, in a randomised complete block design of two replications. Average dry root yield, and yield components (fresh foliage and harvest index) were measured during harvest at 7 MAP for early yield evalauation. Growers routinely harvest at 11 to 24 months after planting.

Trait analysis. Distribution analyses for dry root yield (DR), fresh foliage (FF) and harvest index (HI) were performed using UNIVARIATE procedure of the SAS program (SAS Institute 1996). Normality of distribution was tested (P< 0.05) with the W test described by Shapiro and Wilk (1965). Phenotypic correlations coefficients between yield and components were estimated and tested for significance (P< 0.05). The mean of each genotype was used for the correlation analysis and QTL mapping.

QTL analysis. QTLs were detected by interval mapping analysis using MAPMAKER/QTL 1.1b (Lander *et al.*, 1987). A LOD score of 2.0 was chosen as the minimum to declare the presence of a QTL (Rector *et al.*, 1998; Lin *et al.*, 1998). The LOD score peak was used to estimate the most likely position of the QTL. The confidence interval for each QTL was set at one-LOD support interval, as described by Lander and Botstein (1989). Maximum likelihood estimates of both additive (a) and dominance (d) effects were calculated simultaneously during the genome scan for QTLs. The gene action of individual QTLs was determined to be largely additive, dominant or recessive by testing the hypothesis of a = 0 + d = 0 (the unconstrained "free" genetic model) by evaluating the relative likelihood of models. The additive model was tested by forcing the dominance term $d = 0$, a dominant model by forcing $d = a$, and a recessive model by forcing d = -a. We used a 1-LOD (10 fold) reduction in the likelihood to infer that the type of gene action was unlikely. The average degree of dominance for each QTL was calculated as the ratio d/a. To test for additional QTLs, the position and effect of one QTL was fixed, then the genome was re-scanned searching for other QTLs, using a two-QTL model (TQM). Two-QTL map estimates the fraction of variation explained by each locus while at the same time estimating the effect of others. The amount of unexplained noise in the model was reduced and the increased sensitivity helped in detecting new QTLs. We used the multi-locus model to explain how much of the phenotypic variance among the F$_2$ population for each trait was explained by the SQM or TQM model.

RESULTS

Phenotypic variation. Phenotypic data for DR, FF and HI are summarised in Table 1. The data range for each trait measured revealed wide variation in the F$_2$ population as was the case in the F$_1$ (Okogbenin and Fregene, 2002). All of the traits studied showed a continuous distribution as

TABLE 1. Performances of three traits evaluated in the F_2 population

Trait	Range	Mean	Standard deviation	Skewness	Kurtosis	W-statistic
Fresh foliage (g)	112.50-3900.00	1152.15	583.26	1.08	2.34	0.94*
Dry root (g)	9.5-562.6	213.24	120.62	0.27	-0.43	0.96*
Harvest Index	0.01-0.67	0.38	0.14	-0.4	-0.28	0.96*

expected for quantitative traits. None of the trait (FF, DR and HI) distributions fit a normal distribution; FF and DR were strongly skewed towards the high values while HI was slightly skewed to the less values.

Identification of QTLs affecting DR, FF and HI

Single QTL model (SQM). Significant peak values of LOD scores, the position of these peaks, the percentage phenotypic variance explained, and the estimated gene actions based on the analysis of MAPMAKER/QTL are shown in Table 2. Nine QTLs (LOD> 2.0) (three QTLs each) influencing FF, DR and HI were identified by interval mapping analysis on seven linkage groups. Three QTLs for DR were detected in this cross. The QTLs, *Dr1*, *Dr3* and *Dr13* were present on LGs 1, 3 and 13, respectively (Table 2). *Dr1*, *Dr3* and *Dr16a* accounted for 9, 7 and 6% of the phenotypic variance (PV) for DR, respectively. The QTLs controlling FF, *Ff3*, *Ff5* and *Ff9*, were located within three intervals on LGs 3, 5 and 9, respectively (Table 2). The single biggest QTL effect was observed for *Ff5* with Phenotypic variance of 31%, suggesting it to be a major QTL. The other two QTLs *Ff5* and *Ff9* individually explain 8 and 6% of the observed PV, respectively.

QTLs associated with HI, *Hi2*, *Hi9* and *Hi12* were located on LGs 2, 9 and 12, respectively. They accounted for between 5 and 15% of the total PV. When QTLs detected for all the traits were compared, the LOD score for *Hi2* (6.67) was higher than any other QTLs. The data showed that two QTLs (*Dr3* and *Ff3*) fell within a single interval (NS 928 – SSRY 153) separated by only 4 cM (Table 2). The direction of the genetic effects of these two QTLs were similar, suggesting that they were probably not different QTLs. Two putative QTLs (*Ff9* for FF and *Hi9* for HI;) affecting FF and HI were also found located in

two regions (38 cM apart) in the interval SSRY 12 – SSRY 91 on LG 9.

Two-QTL model (TQM). Additional QTLs were declared present, when significant interaction were observed between single QTLs (from SQM) and other loci elsewhere in the genome. Results from the analyses fitting two QTLs detected additional 15 QTLs for DR, 16 for FF and 13 for HI. However, only significant QTL interactions, which resulted in a minimum increase of 5% in phenotypic variance (PV) above values obtained in SQM, are reported here (Table 3). The improvement in both LOD scores and PVE obtained from fitting additional QTLs with single QTLs (from SQM) suggested that, such additional QTL detected in TQM (Table 3) contributed substantially in the phenotypic expression of the traits they controlled. The additional QTLs detected by TQM, shown in Table 3, were distributed over four LGs. Other significant QTLs detected but not reported here were considered as QTLs with small effects since they explained very little additional variance. Seven highly significant two-QTL interactions were identified for FF, and 4 each for DR and HI. PV explained for these interactions varied from 11 to 36% with LOD scores ranging between 2.74 and 8.97. Some of the QTLs identified in SQM significantly interacted with each other. Five of such interactions were observed amongst single QTLs (2 for HI and 1 each for FF and HI). In some instances, interactions led to a highly significant increase in LOD and PV explained. For example, *Dr1* significantly interacted with *Dr13* resulting in LOD of 5.14 and explained PV of 17.2%, which were higher than the sum of LOD scores and PVE for both QTLs under the SQM.

Some single QTLs identified in SQM for each trait interacted with similar genomic regions. For example, the three FF single QTLs (*Ff3*, *Ff5* and

TABLE 2. Biometrical parameters of individual QTLs affecting dry root yield (DR), fresh foliage (FF) and harvest index (HI) in the F$_2$ population

Trait	QTL	Flanking markers	Length (cM)	Linkage group	LOD	QTL position (cM)	PVE	a	d	d/a	Mode
FF	Ff3	NS 928 – SSRY 153	16.3	3	2.95	0.0	7.6	-87.84	274.90	-3.13	R
	Ff5	SSRY 35 – SSRY 284	28.1	5	2.26	10.0	31.1	-414.21	-556.36	1.34	D
	Ff9	SSRY 12 – SSRY 91	31.0	9	2.16	38.0	5.5	-211.98	-21.83	0.10	A
DR	Dr1	NS 911-NS 847	17.7	1	2.18	12.0	9.2	0.3	73.70	245.67	DR
	Dr3	SSRY 928 - SSRY153	16.3	3	2.14	4.0	7.3	-21.53	52.40	-2.43	R
	Dr16	NS 33 - SSRY 100	16.3	13	2.25	18.0	6.0	47.35	-64.61	-1.36	R
HI	Hi2	NS 149 - SSRY83	7.3	2	6.67	0.0	15.0	-0.08	0.00	0.05	A
	Hi9	SSRY52 - NS 340	3.1	9	2.25	0.0	54.3	0.05	0.00	0.00	A
	Hi12	NS 74 - NS 389	44.4	12	2.10	0.0	4.9	0.01	-0.06	-6.00	A

Individual QTL loci are named by trait (abbreviation indicated in titles) and linkage groups. The LOD score (LOD) and percent phenotypic variance explained (PVE) by the QTLs are presented from the single-QTL model with unconstrained gene action. The additive effect (a) dominance deviation (d), and ratio of dominance to additivity (d/a) for each QTL are presented in their original units. The possible pure modes of gene action (Mode) for each QTL are indicated based on testing of additive (A) and dominant (D, R) models as described in Materials and methods (if d = 0, then A, if d = a then D, if d = -a then R). If a model reduced likelihood by 10-fold or more, it was deemed unlikely. When two pure modes of gene action could not be deemed unlikely, the more likely mode was listed first (e.g. for Dr1, dominance (D) was most likely but recessivity (R) could not be deemed unlikely, thus the mode for this locus is denoted DR). QTL position is position of LOD peak given as distance from the first marker listed in the interval

Ff9) identified in SQM interacted with the same genomic region within the interval NS 717 – SSRY 3 (Table 4). LGs 4, 15 and 16, all give evidence for more than one QTL in one interval. TQM identified two QTLs each on three intervals, NS 717–SSRY3 (Ff4a, Ff4b), SSRY 50–SSRY 281 (Ff15a, Ff15b) and NS 33 – SSRY 100 (Dr16a, Dr16b) on LGs 4, 15 and 16 respectively. The two QTLs located in interval NS 717 – SSRY3 were spaced 6 cM apart whereas those in the interval SSRY 50–SSRY 281 were separated 12 cM apart. Ff4a, Ff4b Ff15a, and Ff15b were QTLs influencing FF. The other two QTLs (Dr16a and Dr16b) were associated with DR, spaced 6 cM apart in the interval NS 33- SSRY 100.

All additional QTLs identified in TQM (listed in Table 3) were fitted along with those identified in the SQM in a multi-locus model to determine total phenotypic variance explained for each trait. The total PV explained based on multiple QTL model are 33% for foliage, 44% for DR and 37% for HI. QTLs detected in SQM and TQM including those of minor effects not reported here revealed 13 intervals affecting more than one trait (Table 4), which is in agreement with the significant correlation observed between yield and the other two traits.

Gene action. The gene action of individual QTLs was evaluated by comparing the fits of individual QTL models (Lander and Botstein, 1989). Two of the QTLs identified for DR (Dr3 and Dr13) were consistent with recessive gene action. Our QTL map for Dr1 indicated that we can confidently rule out the possibility that this locus showed additive gene action suggesting that either recessive and dominance gene action were likely. Dominance effects of this QTL (Dr1) were positive, leading to increase in DR (Table 2). Two QTLs, Hi2 and Hi9 fit a pure additive model. The additive effect, which is the measurement of the change in a population mean when an allele of a QTL is substituted, showed that Hi9 increased harvest index while Hi2 decreased HI. The third QTL for HI (Hi12) was found consistent with recessive or dominance gene action (Table 2). The three QTLs detected for FF were different in their gene actions with Ff3 being consistent with a recessive gene. Ff5 showed dominant gene action while Ff9 exhibited additive gene action.

E. OKOGBENIN *et al.*

TABLE 3. Two-QTL interactions affecting Dry root yield, fresh foliage and harvest index

Trait	Interval 1	LG	QTL	QTL position (cM)	Interval 2	LG	QTL	QTL position (cM)	PVE	LOD
FF	SSRY 12 – SSRY 91	9	*Ff9*	38.0	NS 717 – SSRY 3	4	*Ff4a*	14.0	12.0	3.07
	SSRY 12 – SSRY 91	9	*Ff9*	38.0	NS 217 – NS74	12	*Ff12*	0.0	11.1	4.02
	SSRY 12 – SSRY 91	9	*Ff9*	38.0	SSRY50 – SSRY 281	15	*Ff15a*	40.0	11.9	2.74
	NS 928 – SSRY 153	3	*Ff3*	0.0	NS 717 –SSRY 3	4	*Ff4a*	14.0	16.0	3.97
	NS 928 – SSRY 153	3	*Ff3*	0.0	SSRY12 – SSRY91	9	*Ff9*	38.0	13.2	4.98
	NS 928 – SSRY 153	3	*Ff3*	0.0	SSRY 50 – SSRY281	15	*Ff15b*	28.0	15.4	3.56
	SSRY 35 – SSRY 284	5	*Ff5*	10.0	NS 717 – SSRY 3	4	*Ff4b*	8.0	36.0	3.48
DR	NS 928 – SSRY 153	3	*Dr3*	4.0	NS 717 – SSRY 3	4	*Dr4*	16.0	13.2	3.47
	NS 33 – SSRY 100	16	*Dr16a*	18.0	NS 74 – NS 319	12	*Dr12*	20.0	14.6	3.30
	NS 33 – SSRY 100	16	*Dr16a*	18.0	NS 33 – SSRY 100	16	*Dr16a*	12.0	15.4	3.23
	NS 911 – NS 847	1	*Dr1*	12.0	NS 33 – SSRY 100	16	*Dr16b*	18.0	17.2	5.14
HI	NS 149 – SSRY 83	2	*Hi2*	0.0	SSRY 182 – SSRY 148	17	*Hi8*	8.0	22.1	8.97
	SSRY 52 – NS 340	9	*Hi9*	0.0	NS 149 – SSRY 83	2	*Hi2*	0.0	17.0	7.64
	NS 74 – NS 389	12	*Hi12*	0.0	NS 267 – SSRY 1	18	*Hi18*	26.0	13.1	2.76
	NS 74 – NS 389	12	*Hi12*	0.0	NS 149 – SSRY 83	2	*Hi2*	0.0	18.3	8.25

See Table 2 legend. In cases where multiple QTLs affecting a trait were found along the same linkage group, the QTLs are distinguished by letters indicating the temporal order in which they were discovered (e.g. *Ff15a* and *Ff15b*). The LOD score (LOD) and percent phenotypic variance explained (PVE) by the QTLs are presented from the two-QTL model with unconstrained gene action

Additive effects of all three QTLs for FF resulted in decrease in FF.

QTLs common to the F_1 and F_2 mapping populations. QTL analysis identified similar regions controlling DR and HI in both F_1 and F_2 populations. Results indicated that six of the QTLs detected for DR in the F_1 either coincided with or were linked to intervals significantly associated with DR in the F_2 (Table 5). Similarly, 7 QTLs for HI were found to be common to both populations (Table 5).

DISCUSSION

SSR markers are advantageous to applied plant breeding because they are co-dominant, easily assayed and detect high levels of polymorphism (Morgante and Olivieri, 1993).

TABLE 4. Intervals in which one or more QTLs controlling more than one trait were detected

Interval	Linkage group	Traits
NS 267-SSRY 1	18	FF,DR, HI
NS 149-SSRY 83	2	FF,HI
SSRY 47-SSRY 62	22	FF,DR
NS 717-SSRY 3	4	FF,DR
SSRY 314-NS 82	20	FF,HI
SSRY 12-SSRY 91	9	FF,HI
NS 928-SSRY 153	3	FF,DR
NS 170-NS 207	6	FF,DR
NS 185-SSRY 97	13	FF,HI
SSRY 281-SSRY 82	15	FF,HI
SSRY 20-NS 308	18	DR,HI
SSRY 102-NS 170	6	DR,HI
NS 74-NS 319	12	DR,HI

Earlier studies of early yield in the F_1 indicate that foliage and harvest index are the two most important component traits influencing root bulking in cassava. Development of an F_2 map provides a different generation to study the QTLs and their genetic effects.

The statistical threshold is the most important in QTL analysis because the numbers of detected QTLs would be different when using different thresholds in QTL analysis (Lin *et al.*, 1998). Most of the thresholds employed in published QTL analyses have been between LOD 2.0 and 3.0 in MAPMAKER/QTL. Because our goal was to identify all possible putative QTLs, given that the F_2 linkage map constructed is not saturated, we used a threshold of LOD 2.0 for QTL mapping. In the F_2 population we were able to map three QTLs each for each trait under single-locus model. These QTLs have been described by location, magnitude effect on phenotype, additive effects and dominance deviations, by interaction with unlinked genetic factors and multiple effects on the different traits. The QTLs explain phenotypic variance ranging from 4.9 to 15% with one explaining 31%. These results along with previous studies in the F_1 mapping population support a model for quantitative inheritance for the traits studied.

More QTLs were identified with the two QTL model resulting in additional QTLs for FF, DR and HI. In an F_2 population, one can determine the effect of different gene action on phenotype because all three possible gene dosages at a locus are represented. This could not be done in an F_1 population. Thus, an F_2 population can be used to map recessive factors from either parent (Patterson

TABLE 5. Common genomic regions associated with dry root yield (DR) and harvest index (HI) between F_1 and F_2 populations

Trait	Significant Markers linked to trait in the F_1	Linkage group	Corresponding significant interval linked to trait in the F_2	Linkage group
DR	GY 181, GY42, Ai18b	D	NS 717 – SSRY 3	4
	GY 48	R	SSRY 47 – SSRY 62	22
	CBB1	L	SSRY 20 – NS 308	18
	CDY 131	L	SSRY 20 – NS 308	18
HI	rBEST	A	SSRY 314 - 319	10
	nGY 162	E	SSRY 20 – NS 308	18
	GY 34	J	NS 74 – NS 319	12
	GY 212, GY 142, GY153	L	NS 911 – NS 847	1

et al., 1991) unlike the F_1. For most of the QTLs studied herein, two modes of inheritance were found unlikely suggesting that in such cases the corresponding QTLs were clearly additive, dominant or recessive (Table 2). However, for two QTLs (*Drl and Ff9*), two or more gene actions types could not be rejected indicating that these QTLs may be partially dominant or recessive in gene action.

QTLs affecting different traits fell near one another more frequently than would be expected by chance. This is in agreement with observed correlations between yield and other two traits. Likelihood intervals for 13 QTLs affecting two or more traits were identified. This suggested that either some QTLs have pleoitropic effects (Gruneberg, 1938) or that different QTLs affecting these traits tend to be clustered together into closely linked groups. Close relationships between yield and FF as well as HI, has been suggested by numerous other studies, using either classical analyses or QTL mapping (El-Sharkawy and Cock, 1990; Kawano *et al.*, 1990; Okogbenin and Fregene, 2002). These results are in agreement with our observation in the F_1 mapping studies.

In total, the QTLs with major effects, which could be mapped in the F_2 accounted for 33.4% of the PV in FF, 43.5% in DR and 36.9% in HI. The remaining variation, which could not be explained by the QTL model, may be due to undetected QTLs with too small effects not resolvable by this experiment, interaction between QTLs with small effects, and interaction of individual F_2 genotypes with environment within the experiment (Patterson *et al.*, 1991). The PV explained by our model also probably underestimated total genetic variance because only additive, dominant and recessive genetic components are included in the model. Additional genetic variance may be due to epistasis. Past research in quantitative genetics suggest that interactions between QTL alleles at different loci have considerable influence on phenotype (Spickett and Thoday 1966, Allard, 1988).

Most of the QTLs identified were of small effects. Perhaps some QTLs appear to have small effects because they are dependent upon interaction with other loci. As variance explained by QTL decreases, the number of progeny required to detect QTLs, increases (Lander and Botstein, 1989). Thus, only QTLs with large effects will be detected in a specific cross while those with small effects will go unnoticed. Consequently, estimates of QTL numbers are considered as lower bounds in this study.

In a previous study using an F_1 mapping population, mostly RFLPs markers were employed. Because SSR markers were used in the F_2, integrating SSR markers into the F_1 map is essential for identifying corresponding QTLs in both maps. Current efforts have yielded great success leading to the mapping of over 200 SSR markers in the F_1 map (Zarate *et al.*, unpublished results, Libreros *et al.*, unpublished results). However, a majority of the SSR markers in the F_1 are yet to be mapped in the F_2, thus imposing some limitations on comparing both studies. Results in the F_1 showed that most of the QTLs were not consistently expressed across growth cycle, but were rather age specific, suggesting that different sets of regulatory factors may be involved in different periods of time. Due to the constraint associated with limited availability of planting materials for the newly developed F_2 population, this study could only be done for one location. This constraint also implied that sequential harvesting to determine age specific QTLs could also not be covered in this study since high quantity of planting materials are required for large plot sizes associated with multiple harvests. Consequently, only one single harvest was done at 7 MAP corresponding to the threshold harvest time for early bulking. However, since yield in this study was evaluated only at 7MAP, only a fraction of the QTLs detected in the F_1 expectedly, could be identified in the F_2. GXE interaction studies to determine QTLL stability is of important significance in breeding. Results from studies in the F_1 mapping population have revealed strong GxE effects in QTL expression in cassava (Okogbenin and Fregene, 2002, 2003). Efforts have been initiated to generate sufficient planting materials to validate and evaluate QTLs for early bulking in the F_2 population, over locations and seasons to address age-specific QTLs and GxE effects which were not covered in this study.

QTLs detected in the F_2 mapped to some of linkage groups identified for QTLs of early yield and trait components in the F_1. Similar the genomic regions were found involved in the genetic control

of early yield and harvest index between the F_1 and F_2 progeny, indicating that some QTLs were relatively stable across generations. Results indicated that six of the QTLs detected for DR in the F_1 either coincided with or were linked to intervals significantly associated with DR in the F_2 (Table 5). Similarly, 7 QTLs for HI were found common to both populations (Table 5). By adding more common markers to both the F_1 and F_2 maps, more QTLs for early yield that are useful from the point of view of breeding and stability in different genetic backgrounds, prerequisites for using molecular markers for marker-assisted selection, can be found. The identification of QTLs linked with early yield (bulking) in cassava may be useful for marker-aided selection (MAS) in breeding for early yield. We are testing this hypothesis by crossing these QTLs into different genetic backgrounds and also combining QTLs for the same traits. MAS for early bulking will allow a more accurate and efficient selection of superior genotypes and reduction in costs and time for improving this key trait.

CONCLUSION

Interval mapping analysis detected QTLs for early yield in cassava and its component traits (harvest index and foliage). The present study illustrates that QTL maps for early yield, harvest index and foliage are controlled by similar genomic regions, thus supporting earlier concept that yield in cassava is highly influenced by foliage and harvest index. Gene actions of the identified QTLs revealed that a good number of them displayed additive and dominance gene effects. Since cassava is vegetatively propagated, the utilization of both additive and non-additive variances through identified QTLs can be maximally exploited for rapid genetic gain for yield improvement in cassava. By cross-validating QTLs in different genetic backgrounds (F_1 and F_2 populations), stable QTLs for yield and component traits were identified. Through such stable QTLs, yield, which is a complex trait, can effectively be manipulated through marker assisted breeding in cassava improvement programmes.

REFERENCES

Allard, R.W. 1988. Genetic changes associated with the evolution of adaptedness in cultivated plants and their wild progenitors. *Journal of Heredity* 79:225-238.

Akano, A.O., Dixon, A.G.O., Mba, C., Barrera, E. and Fregene, M. 2002. Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theoretical and Applied Genetics* 105:521-525.

Akkaya, M.S., Bhagwat, A.A. and Cregan P.B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132:1131-1139.

Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A. and Cregan, P.B. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. *Crop Science* 35:1439-1445.

Best, R. and Henry, G. 1992. Cassava: towards the year 2000. In: *International network for cassava genetic resources. Report of the first meeting of the International Network for Cassava Genetics International Crop Network Series number 10. International Plant Genetic Resources Institute (IPGRI)* 10:3-11.

Cock, J.H. 1985. Cassava: new potential for a neglected crop. Westview press, Boulder Colorado USA.

CIAT. 2002. Annual report 2002. Cali, Colombia.

Dellarporta, S.L., Wood, J. and Hicks, J.R. 1983. A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter* 1:19-21.

Doerge, R.W. 1993. Statistical methods for locating quantitative trait loci with molecular markers. PhD dissertation. North Carolina State University.

El-Sharkawy, M.A. and Cock, J.H. 1990. Photosynthesis of cassava (*Manihot esculenta*). *Experimental Agriculture* 26: 325-340.

Fregene, M., Angel, F., Gomez, R., Rodriguez, F., Chavarriaga, P., Roca, W., Tohme, J. and Bonierbale, M. 1997 A molecular genetic map of cassava. *Theoretical and Applied Genetics* 95:431-441.

Grattapaglia, D. and Sederoff, R. 1994. Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Genetics* 137: 1121-1137.

Groover, A., Devey, M., Fiddler, T., Lee, J., Megraw, R., Mitchel-Olds, T., Sherman, B., Vujcic, S., Williams, C. and Neale, D. 1994. Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. *Genetics* 138: 1293-1300.

Gruneberg, J.B. 1938. An analysis of the pleiotropic effects of a new lethal mutation in the rat (*Mus norvegicus*). Proceedings of the Royal Society London Bulletin 125:123-144.

Jarret, R.L. and Bowen, N. 1994. Simple sequence repeats (SSRs) for sweet potato germplasm characterization. Plant Genetic Resources Newsletter 100:9-11.

Jorge, V., Fregene, M.A., Duque, M.C., Bonierbale, M.W., Tohme, J. and Verdier, V. 2000. Genetic mapping of resistance to bacterial blight disease in cassava (*Manihot esculenta* Crantz). *Theoretical Applied Genetics* 101:865-872.

Jorge, V., Fregene, M., Velez, C.M., Duque, M.C., Tohme, J. and Verdier, V. 2001. QTL analysis of field resistance to *Xanthomonas axonopodis* pv *manihotis* in cassava *Theoretical and Applied Genetics* 102:564-571.

Kawano, K. 1990. Harvest index and evolution of major food crop cultivars in the tropics. *Euphytica* 46:195-202.

Lander, E.S. and Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits by using RFLP linkage maps. *Genetics* 121:185-199.

Lander, E.S., Green, P., Abraham, J., Barlow, A., Daly, M.J., Lincoln, S.E. and Newburg, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.

Lin, S.Y., Sasa, T. and Yano, M. 1998. Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa*, using backcross inbred lines. *Theoretical and Applied Genetics* 96:997-1003.

Mann, C. 1997. Reseeding the green revolution. *Science* 277:209-220.

Mogante, M. and Olivieri, A.M. 1993. PCR-amplified microsatellite markers in plant genetics. *Plant Journal* 3:393-427.

Nweke, F.I., Dixon, A.G.O., Asiedu, R. and Folayan, S.A. 1994. Cassava varietal needs of farmers and potential for production growth in Africa. COSCA working paper 10.

Okogbenin, E. and Fregene, M. 2002. Genetic and QTL mapping of early root bulking in an F_1 mapping population of non-inbred parents in cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics* 106: 58-66.

Okogbenin, E. and Fregene, M. 2003. Genetic mapping of QTLs affecting productivity and plant architecture in a full-sib cross from non-inbred parents in cassava (*Manihot esculenta* Crantz). *Theoretical and Applied Genetics* 107:1452-1462.

Patterson, A.H., Damon, S., Hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., Lander, E.S. and Tanksley, S.D. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics* 127:181-197.

Plaschke, J., Ganal, M.W. and Roder, M.S. 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. *Theoretical and Applied Genetics* 91:1001-1007.

Rector, B.G., All, J.N., Parott, W.A., Boerma, H.R. 1998. Identification of molecular markers linked to quantitative trait loci for soybean resistance to corn earworm. *Theoretical and Applied Genetics* 96:786-790.

Roder, M.S., Plaschke, J., Konig, S.U., Borner, A., Sorrels, M.E., Tanksley, S.D. and Ganal, M.W. 1995. Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular and General Genetics* 246:327-333.

Rongwen, J., Akkaya, M.S., Bhagwat, A.A., Lavi, U. and Cregan, P.B. 1995. The use of microsatellite DNA markers for soybean genotype identification . *Theoretical Applied Genetics* 90:43-48.

SAS Institute Inc. 1996. SAS/STAT software: changes and enhancement for release 6.12, Cary, NC: SAS Institute Inc. 158pp.

Shapiro, S.S. and Wilk, M.B. 1965. An analysis of variance for normality (complete samples). *Biometrika* 52:591-611.

Senior, M.L. and Heun, M. 1993. Mapping maize microsatellites and polymerase chain reaction confirmation of the targeted repeats using a CT primer. *Genome* 36:884-889.

Spickett, S.G. and Thoday, J.M. 1966. Regular responses to selection 3. Interaction between located polygenes. *Genetical Research* 7:96-121.

van Eck, H.J., Jacobs, J.M.E., Stam, P., Ton, V.J., Stiekema, W.J. and Jacobsen, E. 1994. Multiple alleles for tuber shape in diploid potato detection by qualitative and quantitative genetic analysis using Ps. *Genetics* 137:303-309.