

A. Tenkouano · J. H. Crouch · H. K. Crouch
D. Vuylsteke · R. Ortiz

Comparison of DNA marker and pedigree-based methods of genetic analysis in plantain and banana (*Musa* spp.) clones.

II. Predicting hybrid performance

Received: 29 October 1997 / Accepted: 14 July 1998

Abstract Pedigree and DNA marker-based methods were used to predict the performance of triploid progeny from tetraploid-diploid crosses, based on parental heterozygosity, genetic relatedness, and expected contribution to their progeny. There was no significant correlation between parental and progeny performance. Prediction of progeny bunch weight was best when based on genealogical distance and equal parental contribution. Predicted fruit size was most accurate when DNA marker data were used and the assumption of an unequal parental contribution was made. Consideration of parental heterozygosity produced larger residuals for all traits. No statistically significant differences were found between the mean residuals obtained under the assumption of an equal vs an unequal contribution of the 4x and 2x genotypes to their 3x progeny, regardless of the method used to estimate genetic relationships.

Key words Genealogical relationships · Genetic markers · Non-additive inheritance · Parental selection · Secondary triploid hybrids

IITA Journal Article no. 97/101/JA

Communicated by H. C. Becker

A. Tenkouano¹ (✉) · J. H. Crouch² · H. K. Crouch
D. Vuylsteke · R. Ortiz³
Plantain and Banana Improvement Program,
International Institute of Tropical Agriculture,
P.M.B. 5320, Oyo Road, Ibadan, Nigeria

Present addresses:

¹ IITA c/o L. W. Lambourn and Co., Carolyn House,
26 Dingwall Road, Croydon CR9 3EE, England
E-mail: iitaonne@satmail.bt.com

² Elsoms Seeds Ltd., Spalding, Lincolnshire,
PE11 1QG, England, U.K.

³ The Royal Veterinary and Agricultural University (KVL),
Department of Agricultural Sciences, 40 Thorvaldsensvej,
DK-1871 Frederiksberg C, Copenhagen, Denmark

Introduction

Progeny testing has been the predominant method for the identification of the combining ability of genotypes and the designation of heterotic groups in many crop species (Hallauer and Miranda 1988; Panter and Allen 1995). In this strategy, potential parental genotypes are selected on the assumption that subsequent offspring will be as good as the previous ones. Thus, progeny testing is inherently postdictive. Furthermore, this method of parental selection cannot be routinely used for *Musa* research due to the large land requirements (6 m² plant⁻¹) and long growth cycle (12–18 months) of this crop.

Predictive models for hybrid performance based on the past performance of parents or their relatives would be most useful for breeders who seek to make the most appropriate crosses on an *a priori* basis. Maize (*Zea mays* L.) breeders have traditionally used a semi-predictive model to select the parents of double-cross hybrids based on the mean of single crosses not involved in the double cross (Jenkins 1934). Other methods for predicting hybrid performance have been based on the genetic relationships among prospective parents and their mid-parental values (Bernardo 1992, 1994; Panter and Allen 1995). Genetic relationships among prospective parents may be estimated using Malécot's (1948) coefficient of co-ancestry or DNA marker polymorphisms (Staub and Serquen 1996; Saghai Maroof et al. 1997).

Prediction of the progeny mean from the mid-parent mean typically assumes three conditions: (1) the progeny of a biparental cross receives half of its genes from each parent, (2) the parents are inbred, homogeneous and unrelated, (3) the traits under consideration are strictly determined by additive inheritance (Panter and Allen 1995; Bernardo et al. 1996). However, current understanding of the meiotic behaviour of *Musa* spp. and of the inheritance of many traits suggests that these assumptions may not hold true in these species (Vuylsteke et al. 1997).

The objective of the present study was to examine the consequence of a departure from the above assumptions and to compare different methods of estimating the performance of secondary triploid *Musa* hybrids developed from interspecific tetraploid and diploid genotypes.

Materials and methods

Genetic materials and field experiments

The founding clones of the genotypes used in this study were three female-fertile triploid West African plantain landraces (*Musa* spp., AAB group) and two male-fertile diploid accessions (AA) from Southeastern Asia. The AAB accessions were 'Bobby Tannap' (BT), 'Obino l'Ewai' (OL) and a somaclonal French reversion mutant of 'Agbagba' (FR), while the AA accessions were *M. acuminata* subsp. *burmanicoides* 'Calcutta 4' (C4) and *M. acuminata* subsp. *malaccensis* 'Pisang liliin' (PL). Five tetraploid and five diploid AAB × AA progenies (Vuylsteke et al. 1993; Vuylsteke and Ortiz 1995) were crossed to produce putative secondary 3x hybrid seeds. The 4x clones were 1658-4 (OL × PL), 2796-5 (BT × PL), 4698-1 (OL × C4), 6930-1 (OL × C4) and 7002-1 (OL × C4), and the 2x clones were 1297-3 (FR × C4), 1448-1 (OL × C4), 2829-62 (BT × C4), 4281-2 (BT × C4) and 4400-8 (BT × C4).

Zygotic embryos were extracted from seeds, germinated in vitro and micropropagated using the methods of Vuylsteke et al. (1990). The parental genotypes were also clonally propagated using meristematic tissue from shoot tips (Vuylsteke et al. 1990). Two-month-old seedlings were transferred to the field in June 1995, at IITA's High Rainfall Station in Onne (4°43'N, 7°01'E, 10 masl), Southeastern Nigeria. The soil here is a deep and freely drained Typic Paleudult of the coarse-loamy, siliceous isohyperthermic family, with poor nutrient status and low pH (pH 4.3 in 1:1 H₂O in the upper 15 cm). Cultural practices were similar to those used by Swennen (1990) and Ortiz and Vuylsteke (1995). The genotypes were arranged in an unbalanced randomised complete block design with five replications for the progenies and two replications for the parents. For each genotype, data were recorded for two consecutive cycles, the plant crop and the first ratoon, on an individual plant bunch weight, average fruit length and average fruit circumference.

Estimation of progeny performance

The simplest approach to estimate bi-parental hybrid performance is to calculate mid-parental values, assuming parents are inbred, unrelated, and contribute equally to their progeny's genotype. Most *Musa* accessions, including the parental clones used in this study, are highly heterozygous, and there is evidence of an unequal genetic contribution of parents to their progeny (Ortiz 1997). On this basis, we modified the mid-parent approach to include terms describing the relative genetic contributions of the parents to their progeny, parental heterozygosity level and genetic relatedness. The following formula was developed for this purpose:

$$H_{ij} = \frac{c_i(1 + f_{ii})P_i + c_j(1 + f_{jj})P_j}{c_i(1 + f_{ii}) + c_j(1 + f_{jj})} \times \left[1 - \text{Ln} \frac{2 - f_{ij}}{\sqrt{(f_{ii})(f_{jj})}} \right], \quad (\text{Eq. 1})$$

where H_{ij} indicates the expected value of the hybrid produced from the i th and j th ($i \neq j$) parents. P_i and P_j are the observed values of the i th and j th parent, respectively. Equation 1 has two components: (1) an additive component which is simply the weighted average of parental phenotypes, and (2) a multiplicative component which reflects heterotic or inbreeding effects.

The terms c_i and c_j are the relative contributions of parents i and j , respectively, to their progeny. In a disomic situation, $c_i = c_j$; that is, the genomic contributions of parents to their offspring are assumed equal (Panter and Allen 1995; Bernardo et al. 1996). However, it is postulated that secondary triploid *Musa* hybrids receive two chromosomes from their tetraploid maternal parent for each chromosome donated by their paternal diploid parent ($c_i \neq c_j$), provided that 2n gametes are not produced (Ortiz 1997).

The terms f_{ii} and f_{jj} represent the probability of two alleles being identical at any locus in parents i and j , respectively, which is indicative of their homozygosity (inbreeding) level. The f_{ij} term is the coefficient of relationship (similarity) among parents i and j . The terms f_{ii} , f_{jj} , and f_{ij} were calculated using five combinations of pedigree and molecular data as described elsewhere (Tenkouano et al. 1998). The triplet (f_{ii} , f_{jj} , f_{ij}) was denoted as ($M_{S_{ii}}$, $M_{S_{jj}}$, $M_{S_{ij}}$) when estimated from DNA marker similarity, ($P_{S_{ii}}$, $P_{S_{jj}}$, $P_{S_{ij}}$) when based on genealogical similarity, ($PM_{S_{ii}}$, $PM_{S_{jj}}$, $PM_{S_{ij}}$) when estimated using a combination of genealogical and molecular data. Genetic relatedness was also expressed in terms of additive relationship coefficients, i.e., (f_{ii} , f_{jj} , f_{ij}) = ($P\phi_{ii}$, $P\phi_{jj}$, $P\phi_{ij}$) when coefficients were calculated from the expected parental contribution to progeny, and (f_{ii} , f_{jj} , f_{ij}) = ($PM\phi_{ii}$, $PM\phi_{jj}$, $PM\phi_{ij}$) when the parental contribution was estimated with DNA marker data (Tenkouano et al. 1998). Thus, with DNA markers and genealogical similarity methods, within-genotype relationships were equal to unity ($f_{ii} = f_{jj} = 1$). The coefficients of additive relationships based on the expected parental contribution to progeny were smaller ($f_{ii} = 0.219$ for tetraploid genotypes, $f_{jj} = 0.208$ for diploid genotypes). When the parental contribution was estimated with DNA markers, we obtained $0.202 \leq f_{ii} \leq 0.216$ for the tetraploid genotypes and $0.200 \leq f_{jj} \leq 0.208$ for the diploid genotypes (Tenkouano et al. 1998). Thus, the coefficients of additive relationships better reflected the heterozygous nature of the clones.

The logarithm term is analogous to Nei's (1972) formula for the calculation of genetic distance and reflects the postulate that the discrepancy between the expected progeny performance and the mid-parent value would vary as a function of the genetic distance among the parents. When parents are very similar genetically, the logarithm term approaches zero and the performance of the progeny is expected to approach the mid-parent value for the trait under consideration. For example, the progeny obtained from selfing an inbred line would be expected to express the same genotypic value as the inbred line. In contrast, crossing two unrelated parents should lead to an F_1 that expresses hybrid vigour, the magnitude of which would also depend on parental contribution and inbreeding status.

Statistical analysis of data

Parental and progeny data were separately subjected to an analysis of variance and the separation of means using the GLM procedure in SAS (SAS Institute 1989). For both groups, the statistical model was as follows:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma(\beta)_{jk} + \delta_l + \beta\delta_{jl} + \epsilon_{ijkl}, \quad (\text{Eq. 2})$$

where Y_{ijkl} is the observed performance of the k th clone of the j th genotype in the i th replication of the l th crop cycle, μ is the overall mean of the trait, α_i is the replication effect ($i = 1, 2$ for parents, $i = 1, \dots, 5$ for hybrids), β_j is the genotypic effect ($j = 1, \dots, 10$ for parents, $j = 1, 2, \dots, 20$ for hybrids), $\gamma(\beta)_{jk}$ is the clone-within-genotypes effect (k is different for each genotype), δ_l is the crop cycle effect ($l = 1, 2$), $\beta\delta_{jl}$ is the genotype × crop cycle interaction effect, and ϵ_{ijkl} is the random error associated with each observation.

Spearman's rank correlation coefficients were calculated among observed parental and progeny means to assess the predictive value of the parental phenotype on hybrid performance. Also, residuals were calculated as the difference between predicted values based on Eq. 1 and the observed average hybrid performance. Standard errors were derived for the residuals and two-tailed t -tests were used to

Table 1 Mean squares from the analysis of variance of bunch weight, fruit length and fruit circumference in parental and hybrid genotypes of *Musa* spp

Source of variation	df	Bunch weight (kg plant ⁻¹)	Fruit length (cm)	Fruit circumference (cm)
Parents				
Replications	1	70.67***	168.65***	30.04***
Genotypes (G)	9	199.91***	136.98***	83.66***
Clones (genotypes)	112	5.89	6.54	1.48
Crop cycle (C)	1	58.16	4.64	8.24
C × G	7	44.54***	9.83	5.97**
Residual	41	4.32	7.91	1.90
R ²		0.98	0.92	0.97
Progeny				
Replications	4	648.68***	1417.15***	554.71***
Genotypes (G)	19	57.18***	161.07***	70.09***
Clones (Genotypes)	805	10.90***	21.16***	18.29***
Crop cycle (C)	1	303.00**	304.64*	139.33*
C × G	18	26.67***	44.37***	19.72***
Residual	565	5.46	13.35	3.89
R ²		0.77	0.76	0.88

*, **, *** *F*-test significant at the 0.05, 0.01, and 0.001 probability levels, respectively.

Table 2 Average bunch weight, fruit length and circumference in diploid and tetraploid parents of secondary triploid *Musa* hybrids

Group	Genotypes ^a	Bunch weight (kg plant ⁻¹)		Fruit length (cm)		Fruit circumference (cm)	
		PC ^b	RT	PC	RT	PC	RT
Males	1297-3 (FR × C4)	5.9	5.0	19.3	17.6	9.4	8.8
	1448-1 (OL × C4)	5.0	6.1	14.7	13.8	8.7	8.6
	2829-62 (BT × C4)	1.5	2.4	11.9	13.1	4.9	5.6
	4281-2(BT × C4)	2.2	3.4	12.0	11.0	6.0	5.5
	4400-8 (BT × C4)	1.6	3.2	13.4	14.4	5.4	5.8
Females	1658-4 (OL × PL)	15.2	5.8	21.3	14.0	13.5	9.0
	2796-5 (BT × PL)	17.1	6.5	22.3	16.5	13.7	10.2
	4698-1 (OL × C4)	14.9	– ^c	22.0	–	11.8	–
	6930-1 (OL × C4)	14.4	–	19.6	–	12.6	–
	7002-1 (OL × C4)	13.1	7.4	22.0	19.5	13.5	10.0
LSD _{0.05}		4.1		5.5		2.7	

^a Parental genotypes are triploid (AAB) West African plantain landraces Obino l'Ewai (OL), Bobby Tannap (BT) and a somaclonal French reversion mutant of Agbagba (FR), and South East Asian diploid (AA) accessions Calcutta 4 (C4) and Pisang lilin (PL)

^b PC = plant (first) crop cycle, RT = ratoon (second) crop cycle

^c Data not available

examine the hypothesis that the residuals were equal to zero while paired *t*-tests were carried out to determine whether the models were significantly different from each other. The relative predictive accuracy of the models was also assessed by comparing the magnitude of the associated residuals.

Results and discussion

Parental phenotype and progeny performance

Significant ($P < 0.05$) differences were observed between and within the 4x-2x progenies for bunch weight, fruit length and fruit circumference, reflecting genetic differences between the parents (Tables 1–3). However, there was no significant difference between clonal repli-

cates of the parental genotypes, thereby indicating their somatic stability for these traits. There was no significant effect of crop cycle on the expression of all three traits in the parental genotypes, but there was a significant interaction between crop cycle and parental genotypes for bunch weight and fruit circumference (Table 1). In contrast, progeny yield and yield components were significantly influenced both by crop cycle and the interaction between crop cycle and hybrid genotypes (Table 1).

While significant intraploidy differences were observed among the parental genotypes, the tetraploid parents produced heavier bunches and bigger fruits than the diploid parents, particularly in the first crop cycle (Table 2). However, the bunch weight and the

Table 3 Average bunch weight, fruit length and circumference in 20 secondary triploid *Musa* hybrids derived from crosses between diploid and tetraploid accessions

Genotypes ^a	Bunch weight (kg plant ⁻¹)		Fruit length (cm)		Fruit circumference (cm)	
	PC ^b	RT	PC	RT	PC	RT
	1658-4 (OL × PL) × 1448-1 (OL × C4)	2.3	6.3	14.0	16.0	10.5
1658-4 (OL × PL) × 2829-62 (BT × C4)	5.2	3.9	16.8	14.3	8.9	8.1
1658-4 (OL × PL) × 4281-2 (BT × C4)	1.1	0.3	8.5	8.0	6.0	3.0
1658-4 (OL × PL) × 4400-8 (BT × C4)	5.0	3.5	14.4	14.2	9.8	10.0
2796-5 (BT × PL) × 1297-3 (FR × C4)	3.1	3.5	12.2	11.6	8.2	7.6
2796-5 (BT × PL) × 1448-1 (OL × C4)	1.4	3.6	9.1	11.7	6.4	7.6
2796-5 (BT × PL) × 2829-62 (BT × C4)	2.0	0.4	12.0	8.0	6.7	5.0
2796-5 (BT × PL) × 4400-8 (BT × C4)	2.8	3.7	11.4	11.2	7.4	7.2
4698-1 (OL × C4) × 1297-3 (FR × C4)	3.4	4.6	13.7	13.4	8.5	8.0
4698-1 (OL × C4) × 1448-1 (OL × C4)	3.1	3.5	12.1	11.2	7.0	7.0
4698-1 (OL × C4) × 4400-8 (BT × C4)	2.2	3.3	10.9	12.2	7.1	7.0
6930-1 (OL × C4) × 1297-3 (FR × C4)	3.7	4.0	13.0	13.1	8.3	8.1
6930-1 (OL × C4) × 1448-1 (OL × C4)	2.4	2.1	13.1	7.3	7.9	6.7
6930-1 (OL × C4) × 2829-62 (BT × C4)	3.2	5.8	15.0	14.0	8.0	7.3
6930-1 (OL × C4) × 4281-2 (BT × C4)	2.4	2.2	14.0	9.0	7.0	7.0
6930-1 (OL × C4) × 4400-8 (BT × C4)	3.2	3.5	12.1	11.7	7.2	6.8
7002-1 (OL × C4) × 1297-3 (FR × C4)	2.9	4.5	13.9	13.5	8.6	8.3
7002-1 (OL × C4) × 2829-62 (BT × C4)	3.8	3.1	14.8	10.9	9.7	7.3
7002-1 (OL × C4) × 4281-2 (BT × C4)	1.1	1.7	8.8	9.4	5.7	5.6
7002-1 (OL × C4) × 4400-8 (BT × C4)	2.4	3.8	13.6	13.5	7.3	7.2
LSD _{0.05}		3.2		5.1		2.7

^a Genotypes were derived from crosses between triploid (AAB) West African plantain landraces Obino l'Ewai (OL), Bobby Tannap (BT) and a somaclonal French reversion mutant of Agbagba (FR), and South East Asian diploid (AA) accessions Calcutta 4 (C4) and Pisang lilin (PL)

^b PC = plant (first) crop cycle, RT = ratoon (second) crop cycle

fruit size of the tetraploid parents were significantly reduced in the second cycle, while the diploid parents essentially maintained their performance for these traits. Tetraploid clones displayed a regulated suckering behaviour which resulted in delayed flowering in the second crop cycle, compared to diploid clones which had a non-regulated suckering behaviour. As a result, flowering and fruit filling of tetraploid clones coincided with the dry season, hence their reduced performance in the second cycle. The genotype 1297-3 (FR × C4) expressed the best combination of bunch weight, fruit length, and fruit circumference in both crop cycles, while the BT × C4 genotypes had the lowest values for these traits (Table 2). Although tetraploid genotypes had similar fruit length and circumference, they differed significantly in yield, with 2796-5 (BT × PL) being the best in overall yield (Table 2).

The average performance of hybrids from 4x-2x crosses was closer to that of their male parents than to the performance of their female parent (Table 3), although Spearman's rank (r_s) correlation coefficients between parental and progeny performance were not significant. However, the largest associations, i.e. the smallest P -values, were observed between diploid males and their half-sib progeny for fruit length ($r_s = -0.48$, $P = 0.16$) and circumference ($r_s = 0.41$, $P = 0.24$). The bunch weight of maternal half-sibs was negatively associated with that of their female parent ($r_s = -0.38$, $P = 0.35$), while the opposite ($r_s = 0.28$, $P = 0.42$) was

observed for the relationship between diploid males and their half-sibs for this trait. Correlation coefficients between mid-parent and full-sib performance were smaller than those between diploid males and paternal half-sibs, but greater than those between tetraploid females and maternal half-sibs. Thus, parental or mid-parental phenotype had little predictive value for progeny performance in this study. More specifically, the mid-parent value was higher than the progeny mean for all 20 populations. A similar result was reported by Cowen and Frey (1987) in segregating oat (*Avena sativa* L.) populations. This contrasts with the results of Busch et al. (1974) who reported a high positive correlation between hybrid yield and mid-parent values, using 25 segregating populations of wheat (*Triticum aestivum* L.). Souza and Sorrels (1991) suggested that the mid-parent value was a nearly perfect predictor of hybrid performance in oats since the hybrid mean was statistically equivalent to the mid-parent values in 20 populations.

Genealogical relationships and progeny performance

Average bunch weight in the first cycle was highest, 5.2 kg plant⁻¹, for the cross between 1658-4 (OL × PL) and 2829-62 (BT × C4) which also had no common parent. In the second crop cycle, the average bunch weight ranged from 0.3 kg plant⁻¹ for the cross

between 1658-4 (OL × PL) and 4281-2 (BT × C4) to 6.3 kg plant⁻¹ for the cross between the maternal half-sibs 1658-4 (OL × PL) and 1448-1 (OL × C4) (Table 3). The average bunch weight for the first crop cycle of the secondary triploid hybrids was smallest, 1.1 kg plant⁻¹, for the cross between the genotypes 1658-4 (OL × PL) and 4281-2 (BT × C4), which had no common parent, and for the cross between the paternal half-sibs 7002-1 (OL × C4) and 4281-2 (BT × C4) (Table 3).

Average fruit length of hybrids (Table 3) ranged from 8.5 cm in the cross 1658-4 (OL × PL) × 4281-2 (BT × C4) to 16.8 cm in the cross between 1658-4 (OL × PL) and 2829-62 (BT × C4) in the first crop cycle. However, in the second cycle, average fruit length was smallest, 7.3 cm, for the cross between the full-sibs 6930-1 (OL × C4) and 1448-1 (OL × C4) and highest, 16.0 cm, for the cross between the maternal half-sibs 1658-4 (OL × PL) and 1448-1 (OL × C4) (Table 3). The family 1658-4 × 1448-1 also displayed the largest average fruit circumference in both crop cycles, whereas the smallest values were recorded for progeny from the crosses 7002-1 (OL × C4) × 4281-2 (BT × C4) and 1658-4 (OL × PL) × 4281-2 (BT × C4) in the first and second crop cycles, respectively (Table 3). Thus, genealogical relationships alone were not predictive of progeny performance for yield and its components in this study. The data presented by Helms et al. (1997) for soybean (*Glycine max* L.) also indicate that the lowest yield was observed for a parental pair that had a zero coefficient of parentage and highest for the parental pairs with the highest coefficients of parentage.

Parental heterozygosity, genetic relatedness, and hybrid performance

Statistical models that take into account the performance of the parental genotypes, their actual or estimated relative genetic contributions to their progeny, their heterozygosity level, and their genetic relatedness (e.g. Eq. 1), could provide a better means for predicting the performance of the hybrids developed from such parents. Genetic relationships within and across parental genotypes were estimated from pedigree and molecular data (Tenkouano et al. 1998).

The average discrepancy between predicted and observed progeny bunch weight was not statistically different from zero when genetic relationships were estimated with the absolute distance method (Fig. 1). Yield prediction based on all other methods produced residuals which were significantly different from zero: methods based on DNA markers generally overestimated the bunch weight of the progenies, whereas the methods based on coefficients of additive relationships resulted in underestimation of this trait (Fig. 1).

There is no simple relationship between parental and progeny yield because bunch weight is a complex

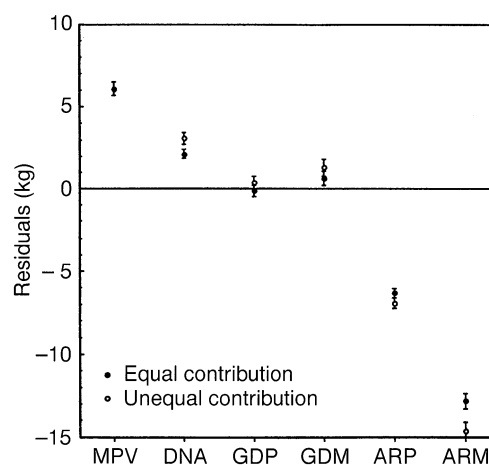


Fig. 1 Mean residuals associated with different methods of predicting the bunch weight of progeny from tetraploid × diploid crosses of *Musa* spp, based on parental performance, inbreeding status, genetic relatedness, and the contribution to the progeny genome. Inbreeding and genetic relatedness were estimated using DNA markers (*DNA*), genealogical distance solely (*GDP*) or in combination with molecular data (*GDM*), coefficients of additive relationships based solely on pedigree (*ARP*) or in conjunction with DNA marker data (*ARM*). *MPV* represents the mid-parental value (control). An equal contribution of 4x and 2x parents to the 3x genome assumes abnormal meiosis (unbalanced gametes) in both parents, whereas an unequal contribution assumes normal meiosis in both parents

character which is controlled by primarily non-additive genetic effects (Vuylsteke et al. 1997). Nevertheless, our results suggest that prediction methods based on the absolute distance calculated from pedigree data alone would provide conservative estimates of bunch weight in 4x-2x hybrids, and particularly when the parental contributions to progeny were assumed equal.

For fruit length, the prediction methods based on absolute distance and additive relationships produced large negative residuals, which were significantly different from zero (Fig. 2). Residuals obtained with DNA marker predictors were also negative but smaller than those associated with the pedigree-based methods (Fig. 2). Prediction of fruit length was best, i.e. the mean residual was not significantly different from zero, when genetic relationships were estimated with DNA markers and the parental contribution to the progeny was assumed to be unequal (Fig. 2).

Residuals associated with fruit circumference displayed the same pattern as for fruit length, but none were statistically equal to zero. However, the discrepancy between predicted and observed values was smallest when genetic relationships were estimated with DNA markers and the parental contribution was assumed to be unequal (Fig. 3).

Consideration of parental heterozygosity resulted in larger residuals (absolute values) for all traits and did not appear to significantly improve the prediction of any of these traits. Similarly, there was no statistical difference between the mean residuals obtained under

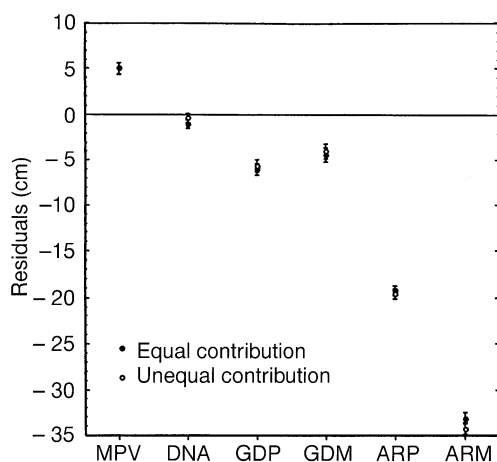


Fig. 2 Mean residuals associated with different methods of predicting the fruit length of progeny from tetraploid \times diploid crosses of *Musa* spp, based on parental performance, inbreeding status, genetic relatedness, and the contribution to the progeny genome. Inbreeding and genetic relatedness were estimated using DNA markers (*DNA*), genealogical distance solely (*GDP*) or in combination with molecular data (*GDM*), coefficients of additive relationships based solely on pedigree (*ARP*) or in conjunction with DNA marker data (*ARM*). *MPV* represents the mid-parental value (control). An equal contribution of 4x and 2x parents to the 3x genome assumes abnormal meiosis (unbalanced gametes) in both parents, whereas an unequal contribution assumes normal meiosis in both parents

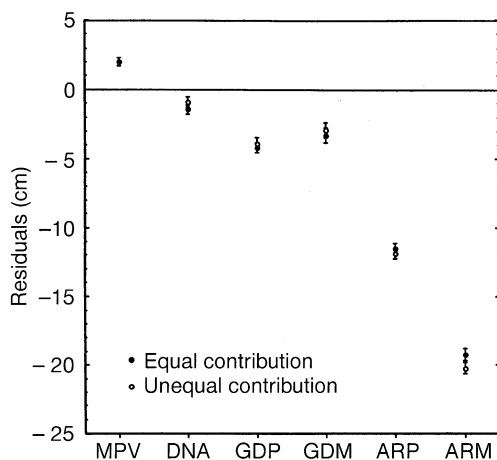


Fig. 3 Mean residuals associated with different methods of predicting the fruit circumference of progeny from tetraploid \times diploid crosses of *Musa* spp, based on parental performance, inbreeding status, genetic relatedness, and contribution to progeny genome. Inbreeding and genetic relatedness were estimated using DNA markers (*DNA*), genealogical distance solely (*GDP*) or in combination with molecular data (*GDM*), coefficients of additive relationships based solely on pedigree (*ARP*) or in conjunction with DNA marker data (*ARM*). *MPV* represents the mid-parental value (control). An equal contribution of 4x and 2x parents to the 3x genome assumes abnormal meiosis (unbalanced gametes) in both parents, whereas an unequal contribution assumes normal meiosis in both parents

the assumption of an equal vs an unequal contribution of the 4x and 2x genotypes to their 3x progeny. However, regardless of the method used to estimate genetic relationships, residuals associated with bunch weight

were smaller when the parental contribution to their progeny was assumed to be equal (Fig. 1). In contrast, residuals associated with fruit length and fruit circumference were smaller with the assumption of an unequal contribution (Figs. 2, 3), reflecting maternal effects in the determination of fruit size. Nevertheless, for the purpose of comparison, it would be acceptable to assume that parents are inbred and contribute equally to their progeny, especially when molecular data are not available.

The assumption that the parental genotypes were unrelated resulted in an overestimation of progeny performance by an average of 6 kg plant⁻¹ for bunch weight (Fig. 1), 5 cm for fruit length (Fig. 2) and 2 cm for fruit circumference (Fig. 3). Parental relatedness may have been underestimated with methods based on coefficients of additive relationships, resulting in an underestimation of progeny performance; hence the large negative residuals observed for all traits (Figs. 1–3). When parental relatedness was estimated with the absolute distance method, the discrepancy between predicted and observed bunch weight was less than 1 kg. Likewise, a consideration of parental relatedness estimated from the molecular data provided the best predictions for fruit size.

The relationship between parental divergence and hybrid performance has been investigated in several crop species. For example, Saghai-Marooif et al. (1997) estimated parental divergence in rice (*Oryza sativa* L.) from the level of RFLP heterozygosity of their F₁ progeny and detected high correlation coefficients between such heterozygosity and rough rice yield ($r = 0.79^{**}$) and head rice yield ($r = 0.82^{**}$). Similarly, the performance of F₁ hybrids was found to be correlated with the RFLP marker distance of their parents in the studies of Godshalk et al. (1990), Lee et al. (1989), Smith et al. (1990), and Dudley et al. (1991). However the extent that marker distance relates to hybrid performance was found to depend on the genetic background of the prospective parents (Bernardo 1992; Melchinger 1993). Whether such trends would apply to DNA polymorphisms based on the segregation of SSR loci is not known.

A significant correlation was also found between genealogical distance and heterosis in the F₁ generation (Smith et al. 1990), although in other cases genealogical distance was not a good predictor of hybrid performance (Cox and Murphy 1990).

Melchinger (1993) suggested that the use of genetic divergence could assist in assigning prospective parents to heterotic groups, particularly in crop species where heterotic patterns are not yet well defined. Our study suggests that, in so far as pedigree records are correct, genealogical distance could be used by *Musa* breeders to predict progeny yield and to reduce the production and field evaluation of potentially inferior experimental crosses. This would reduce costs and enhance the chances of identifying superior hybrids.

References

- Bernardo R (1992) Relationship between single-cross performance and molecular marker heterozygosity. *Theor Appl Genet* 83:628–634
- Bernardo R (1994) Prediction of maize single-cross performance using RFLPs and information from related hybrids. *Crop Sci* 34:20–25
- Bernardo R, Murigneux A, Karaman Z (1996) Marker-based estimates of identity by descent and likeness in state among maize inbreds. *Theor Appl Genet* 93:262–267
- Busch RL, Janke JC, Frohberg RC (1974) Evaluation of crosses among high and low yielding parents of spring wheat (*Triticum aestivum* L.) and bulk prediction of line performance. *Crop Sci* 14:47–50
- Cowen NM, Frey KJ (1987) Relationship between genealogical distance and breeding behavior in oats (*Avena sativa* L.). *Euphytica* 36:413–424
- Cox TS, Murphy JP (1990) The effect of parental divergence on F_2 heterosis in winter wheat crosses. *Theor Appl Genet* 79:241–250
- Dudley JW, Saghai-Maroo MA, Rufener GK (1991) Molecular markers and grouping of parents in maize breeding programs. *Crop Sci* 31:718–723
- Godshalk EB, Lee M, Lamkey KR (1990) Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. *Theor Appl Genet* 80:273–280
- Hallauer AR, Miranda JB (1988) Quantitative genetics in maize breeding. Iowa State Univ Press, Ames, Iowa
- Helms T, Orf J, Vallad G, McClean P (1997) Genetic variance, coefficient of parentage, and genetic distance of six soybean populations. *Theor Appl Genet* 94:20–26
- Jenkins MT (1934) Methods of estimating the performance of double crosses in corn. *J Am Soc Agron* 26:199–204
- Lee M, Godshalk EB, Lamkey KR, Woodman WW (1989) Association of restriction fragment length polymorphisms among maize inbreds with agronomic performance of their crosses. *Crop Sci* 29:1067–1071
- Malécot G (1948) *Les mathématiques de l'hérédité*. Masson et Cie, Paris
- Melchinger AE (1993) Use of RFLP markers for analysis of genetic relationships among breeding materials and prediction of hybrid performance. In: Buxton DR, Shibles R, Forsberg RA, Blad BL, Asay KH, Paulsen GM, Wilson RF (eds) *International Crop Science I*. CSSA, Madison, Wisconsin, USA, pp 621–628
- Nei M (1972) Genetic distance between populations. *Am Nat* 106:283–292
- Ortiz R (1997) Occurrence and inheritance of 2n pollen in *Musa*. *Ann Bot* 79:449–453
- Ortiz R, Vuylsteke D (1995) Recommended experimental designs for selection of plantain hybrids. *InfoMusa* 4:11–12
- Panter DM, Allen FL (1995) Using best linear unbiased predictions to enhance breeding for yield in soybean. II. Selection of superior crosses from a limited number of yield trials. *Crop Sci* 35:405–410
- Saghai Maroo MA, Yang GP, Zhang Q, Gravois KA (1997) Correlation between molecular marker distance and hybrid performance in U.S. southern long grain rice. *Crop Sci* 37:145–150
- SAS Institute Inc. (1989) *SAS/STAT users' guide*, version 6, 4th edn. SAS Institute Inc., Cary, North Carolina, U.S.A.
- Smith OS, Smith JSC, Bowen SL, Tenborg RA, Wall SJ (1990) Similarities among a group of elite maize inbreds as measured by pedigree, F_1 grain yield heterosis, and RFLPs. *Theor Appl Genet* 80:833–840
- Souza E, Sorrels ME (1991) Prediction of progeny variation in oat from parental genetic relationships. *Theor Appl Genet* 82:233–241
- Staub JE, Serquen FC (1996) Genetic markers, map construction, and their application in plant breeding. *HortScience* 31:729–741
- Swennen R (1990) Plantain cultivation under West African conditions: a reference manual. IITA, Ibadan, Nigeria
- Tenkouano A, Crouch JH, Crouch HK, Vuylsteke D, Ortiz R (1998) Comparison of DNA marker and pedigree-based methods of genetic analysis in plantain and banana (*Musa* spp.) clones. I. Estimation of genetic relationships. *Theor Appl Genet* 98:62–68
- Vuylsteke D, Ortiz R (1995) Plantain-derived diploid hybrids (TMP2x) with black sigatoka resistance. *HortScience* 30:147–149
- Vuylsteke D, Swennen R, De Langhe E (1990) Tissue culture technology for the improvement of African plantains. In: Fullerton RA, Stover RH (eds) *Sigatoka leaf spot diseases of bananas*. Proc Int Wksp, San Jose, Costa Rica, 28 March–1 April 1989. Int. Network for the Improvement of Banana and Plantain, Montpellier, France, pp 316–337
- Vuylsteke D, Swennen R, Ortiz R (1993) Registration of 14 improved tropical *Musa* plantain hybrids with black sigatoka resistance. *HortScience* 28:957–959
- Vuylsteke DR, Ortiz R, Ferris RBS, Crouch JH (1997) Plantain improvement. *Plant Breed Rev* 14:267–320