Significant progressive heterobeltiosis in banana crossbreeding

Michael Batte[1,2*], Moses Nyine[3], Brigitte Uwimana[1], Rony Swennen [4,5,6], Violet Akech[1], Allan Brown[4], Helena Persson Hovmalm[2], Mulatu Geleta [2], Rodomiro Ortiz.[2]

* Corresponding author: M.Batte@cgiar.org

Abstract

Background: Heterobeltiosis is the phenomenon when the hybrid's performance is superior to its best performing parent. Banana (*Musa* spp. AAA) breeding is a tedious, time-consuming process, taking up to two decades to develop a consumer acceptable hybrid. Exploiting heterobeltiosis in banana breeding will contribute to selecting breeding material with high compatibility, thus increasing banana breeding efficiency. The aim of this study was therefore to determine and document the level of heterobeltiosis of bunch weight and plant stature in the East African highland bananas, in order to identify potential parents that can be used to produce offspring with desired bunch weight and stature after a few crosses.

Results: This study found significant progressive heterobeltiosis in cross-bred 'Matooke' banana hybrids, also known as NARITAs, when grown together across years with their parents and grandparents in Uganda. All NARITAs exhibited positive heterobeltiosis for bunch weight, whereas half of them had negative heterobeltiosis for stature. NARITA 17 had the highest heterobeltiosis for bunch weight (248.7%), followed by 26666S-1 (229.3%), while the lowest was noted in NARITA 19 (1.2%). Heritability for yield potential and bunch weight were high (0.84 and 0.76 respectively), while that of plant stature was very low (0.0035). There was a positive significant correlation ($P \leq 0.05$) between grandparent heterobeltiosis for bunch weight and genetic distance between NARITAs' parents (r = 0.39, $P = 0.036$), bunch weight (r = 0.7, $P < 0.001$), plant

stature (r = 0.38, *P* = 0.033) and yield potential (r = 0.59, *P* = 0.0004). Grandparent heterobeltiosis for plant stature was significantly, but negatively, correlated to the genetic distance between NARITA parents (r = -0.6, *P* = 0.0004).

Conclusions: Such significant heterobeltiosis exhibited for bunch weight is to our knowledge the largest among main food crops. Since bananas are vegetatively propagated, the effect of heterobeltiosis is easily fixed in the hybrids and will not be lost over time after the release and further commercialization of these hybrids.

Key words: Bunch weight, East African highland banana, Genetic distance, Heterobeltiosis, *Musa* spp., NARITA

Background

Bananas and plantains (*Musa* spp. L) are important food and cash crops to millions of people in the tropical and subtropical regions of the world [1]. They are grown in more than 135 countries. In India, the largest banana producer, the crop occupies 20% of the area under fruit crops. Bananas and plantains rank among the most important food crops in the developing world [2]. In Uganda, bananas are grown by at least 75% of the farmers and cover an estimate of 38% of the total land under crops [3]. However, the production has declined over the past three decades due to mainly declining soil fertility and drought [4], plus pests and diseases. The most economically important pests for bananas in the Great Lakes region of Africa are the burrowing nematode (*Radopholus similis*) and banana weevil (*Cosmopolites sordidus*). The diseases are caused by pathogens which

thrive in tropical conditions, the most important of which are *Xanthomonas vasicola* pv. *musacearum* (formerly *Xanthomonas campestris* pv. *musacearum)* leading to banana bacterial wilt [5], *Pseudocercospora fijiensis*, causing black Sigatoka or black leaf streak disease [6, 7], and *Fusarium oxysporum* f. sp. *cubense* causing fusarium wilt or Panama disease [8]. Breeding of resistant/tolerant cultivars is the most sustainable intervention for banana health management [9, 10, 11]. However, plant breeding is a long process requiring efficient selection of suitable parents with desired traits to produce superior hybrids [12]. Utilization of heterosis or heterobeltiosis can speed up the process of generating superior hybrids. Heterosis, or hybrid vigour, is the superiority of the hybrid for a certain trait over the mean of the parents, whereas heterobeltiosis is a form of heterosis where the hybrid is superior to its best performing parent [13]. Jones [14] defined heterosis as the expression of dominance deviation, a variance from mid parent value, which may be explained by the additive effects of several desired dominant alleles, or as "overdominance," the combined effect of (two) different alleles at the same gene locus, or a combination of both. From the definitions, heterobeltiosis helps a breeder to make more stringent selections than heterosis, as also reported by Lamkey and Edwards [15]. Both positive and negative heterosis can be useful depending on the breeding objectives. Generally, positive heterosis is very useful when selecting for yield and its components, whereas negative heterosis is desired when selecting for short plant height and fast or early cycling [16,15]. Gowda et al. [12] reported that selection of promising parents to obtain superior hybrids primarily depends on the predominance of the genes for the additive effect due to heterosis and heterobeltiosis. The underlying genetic and molecular mechanisms of heterosis remain unknown [13]. Some of the theories for heterosis include dominance, over-dominance and epistasis [17,18]. Tao et al. [19] reported that it is possible to efficiently screen for superior parents and predict the heterosis of parental combinations. They

72 further pointed out that genetic differences between parents are the primary cause of heterosis.
73 Also, the correlation between the genetic distance and heterosis was said to depend on the type of
74 materials. Many studies have been done by plant breeders to investigate the genetic diversities
75 between parents and their relationship with heterosis. However, according to Hinze and Lamkey
76 [20], limitations in traditional methods based on geographic origins, genetic relationships,
77 morphological markers and isozymes make the prediction of heterosis difficult. The development
78 of molecular marker techniques is seen as a new and more effective way for heterosis prediction,
79 which will in turn improve the efficiency of hybrid breeding. van Ginkel and Ortiz [21] reported
80 that heterosis in selfing crops is often driven by additive and additive × additive gene action. They
81 further argued that this can be relatively easily fixed in homozygous lines, meaning that their seed
82 can simply be re-sown to express the heterosis, unlike nonadditive heterosis.

83 Goff [22] proposed a concept of heterosis which summarizes other theories that were earlier
84 proposed about the physiology of heterosis. It states that "heterosis is a result of allele-specific
85 expression, which favors the expression of the most energy-saving, stable alleles". In hybrids,
86 alleles at a locus are likely to be different, and there are multiple opportunities for allele-specific
87 expression of the more stable gene product. Hybrids are therefore more efficient in overall energy
88 use than their parents, with most loci in homozygous state and can use the saved energy for other
89 tasks. The saved energy can be invested in higher growth rates compared with the parental lines,
90 a phenomenon we perceive as heterosis. van Ginkel and Ortiz [21] reported that heterosis due to
91 dominance can be captured in homozygous individuals, as the favorable allele can be present twice
92 in homozygous lines or doubled haploids, unlike heterosis due to overdominance, which involves
93 different alleles of the same gene. More recent research is showing that, in selfing and some
94 outcrossing crops, dominance is more important than overdominance, implying that additive gene

95 expression exceeds non-additive gene action [23, 24]. However, Goldringer et al. [25] reported a
96 larger epistatic effect than additive genetic variance for grain yield in wheat. The more the additive
97 and additive × additive gene actions dominate in hybrids, the more effectively the F_1 performance
98 predicts the subsequent derived line performance.

99 Recent research gives an insight in gene actions driving heterosis in various crops. Heterosis for
100 grain yield components appears to be controlled by additive gene action [12], but also, as noted by
101 Beche et al. [13], by additive × additive gene effects. Early research in barley revealed that
102 heterosis in seed yield is due to additive and "homozygous–homozygous" gene effects [26, 27],
103 while heterosis for grain yield in rice seems to be determined by additive and additive × additive
104 gene action [28, 29, 30]. Scanty research results are available about heterosis in bananas [31],
105 though not in the East African highland bananas. The aim of this study was therefore to determine
106 and document the level of heterobeltiosis of bunch weight and plant stature in the East African
107 highland bananas, in order to identify potential parents that can be used to produce offspring with
108 desired bunch weight and plant stature after a few crosses, thereby improving the efficiency of the
109 banana breeding program.

110

111 **Results**

112 Broad sense heritability (H^2) for yield was 0.84, for bunch weight 0.76 and for plant stature 0.0035.
113 NARITA 23 had the highest bunch weight (29.3 kg), followed by NARITA 17 (29.0 kg) and
114 NARITA 18 (28.6 kg) while NARITA 19 had the smallest bunch weight (11.1 kg; Table 1).
115 However, NARITA 17 had the highest yield potential (35.6 t ha $^{-1}$ yr $^{-1}$), followed by NARITA 23
116 (35.0 t ha $^{-1}$ yr $^{-1}$) and NARITA 18 (34.4 t ha $^{-1}$ yr $^{-1}$) whereas NARITA 19 had the lowest yield

potential (14.7 t ha $^{-1}$ yr $^{-1}$). Similarly, NARITA 17 had the highest heterobeltiosis of 248.7% (Table 1), followed by 26666S-1 (229.3%), and NARITA 9 (201.2%) while NARITA 19 had the lowest heterobeltiosis of 1.2%. NARITA 7, the only released NARITA hybrid cultivar in Uganda so far, had a heterobeltiosis of 77.2%. All the 31 NARITAs available at the International Institute of Tropical Agriculture (IITA), Sendusu research station, with known pedigrees showed heterobeltiosis for bunch weight, compared to 'Matooke' (3*x* grandparents). Plant stature ranged from 0.16 to 0.21 (Table 2). NARITA 20 had the highest positive grandparent heterobeltiosis for stature (30.5%), followed by 29285S-20 (27.2%), and NARITA 17 (26.9%) while NARITA 1 had the highest negative grandparent heterobeltiosis (-18.4%). Half of the NARITAS had negative grandparent heterobeltiosis for plant stature (Table 2).

There was a positive significant correlation (at 95% confidence level) between grandparent heterobeltiosis for bunch weight and genetic distance between NARITA parents (r = 0.39, *P* = 0.036), bunch weight (r = 0.7, *P* = 1.028e-05), plant stature (r = 0.38, *P* = 0.033) and yield (r = 0.59, *P* = 0.0004) (Table 3). A significant and negative correlation between grandparent heterobeltiosis for plant stature and the genetic distance between NARITA parents (r = -0.6, *P* = 0.0004) was observed (Table 3). In a cladogram (Fig. 1), genotypes of the same known group clustered together such as NARITA cultivars, female parents of NARITAs, male parents of NARITAs and female grandparents of NARITAs, except 'cv. Rose' which clustered among the NARITAs between 29285S-20 (a progeny with 'cv. Rose' as the male parent) and NARITA 5. There was a significant ($P \leq 0.05$) progressive heterobeltiosis for bunch weight in bred 'Matooke' banana hybrids (NARITA), when grown together across years with their ancestors in Uganda (Fig. 2, Table 1, Table 3).

Discussion

Genetic factors explained a higher proportion of variance for yield and bunch weight than plant stature. The highest broad sense heritability (H^2) was recorded for yield (84%) followed by bunch weight (76%), while plant stature had the lowest H^2 (0.35%). These results differ from those reported by Tenkouano et al. [31] where the heritability estimates of yield components in plantains and derived hybrids were 42% for fruit circumference, 36% for bunch weight and fruit length and zero for number of hands and fruits. However, they argued that this medium heritability enabled yield improvement of individual plants through increased fruit size when recurrent selection was applied. Therefore, additional gains could be obtained through crossbreeding, despite the small recombinative heterosis. They further pointed out that diploid males contributed at least twice as much as tetraploid females to the yield of the progeny, implying that paternal phenotype was more predictive of progeny performance for this trait. This finding suggests that great yield gains are likely to be achieved when favorable alleles are accumulated in a diploid male parent. Incorporation of useful genes in the diploids is much easier than in polyploid parents. When these diploid males are crossed with higher ploidy level females, there is a higher probability of recovering hybrid offspring that show heterosis for the desired traits.

All the 31 NARITAs with known pedigrees showed positive heterobeltiosis for bunch weight, compared to their grandparents, the triploid ‘Matooke’ banana, despite the heterozygosity of the parents. This progressive heterosis, which does not ensue from crossing inbred lines [32], could be a result of favorable allele combinations that are kept in linkage disequilibrium through vegetative propagation [33]. Analysis of evolutionary history suggests that bananas underwent instant domestication followed by a few meiosis events. This may account for the very large

heterosis and grandparent heterobeltiosis noted in the most high-yielding NARITAs, which was above most of other food crops as per available knowledge [34]. It would be interesting to check if further crossbreeding of the NARITAs could maximize progressive heterosis responses resulting in higher-yielding third generation polyploid hybrids. Perrier et al. [35] postulated that the rise of cultivated triploid bananas from their direct wild ancestors, *M. acuminata* and *M. balbisiana* among others, was a three-step process. The first step was the anthropogenic circulation of pre-domesticated forms of diploid bananas extracted from the different wild genepools. The second step was the production of edible diploid hybrids, which occasionally produced unreduced gametes. Finally, sexual recombination among cultivated diploids followed by the fusion of reduced and unreduced gametes gave rise to the triploid varieties. The actual number of sexual events that gave rise to the diverse forms of bananas is unknown. However, Bakry and Horry [36] estimated it to be 7-14 events while Sardos et al. [37] estimated that the 208 cultivated diploids in their study may have arisen from 117 distinct sexual events, while 80 sexual events were estimated to be the origin of the 273 triploid accessions based on DArT markers. Yet, the East African highland bananas are believed to have arisen from a single ancestral clone that underwent population expansion by vegetative propagation [38].

Plant height is one of the agronomic traits that directly or indirectly influence yield. In cereals like wheat, the increase in yield during the 'Green Revolution' was attributed to mutant dwarfing alleles in the *RHt-1* gene which resulted in shorter plants that produced more tillers resulting in an increased number of grains and a reduced lodging by wind and rain [39]. Tall banana plants with slender pseudostems are more prone to wind damage especially after flowering due to the weight of the bunch. Half of the NARITAs expressed negative grandparent heterobeltiosis for plant stature

186 indicating that they were taller and slenderer than the grandparents, which is not desirable for
187 bananas due to high risk of breakage by wind. Lamkey and Edwards [15] and Alam et al. [16]
188 suggested that positive heterosis is desired in the selection for yield and its components, whereas
189 negative heterosis is desired for early cycling and short plant height. In our case however, a positive
190 heterobeltiosis for plant stature was desirable since it indicates that the hybrids are shorter or of
191 the same height as the grandparent but with more robust pseudostems. This is because plant stature
192 in the present study was calculated as a ratio of plant girth at 100 cm from the ground to the total
193 height of the plant at flowering (girth/height). A short plant with a large girth therefore would have
194 a higher value for stature than a tall, plant.

195

196 Although there is no unifying theory to explain the phenomenon of heterosis, several mechanisms
197 such as genetic diversity, overdominance, epistasis, and purging of deleterious alleles through
198 heterozygosity have been tested in different models and linked to observed heterosis in complex
199 traits [24]. In the present study, we observed a positive significant correlation between grandparent
200 heterobeltiosis for bunch weight and genetic distance between NARITA parents. These results
201 agree with those of Marcón et al. [40] who also reported a positive relationship between genetic
202 distances among parents and heterosis for forage yield in bahiagrass (*Paspalum notatum*).
203 However, these results contradict with what was observed by Tenkouano et al. [31] in plantains
204 and plantain-derived hybrids. They reported that hybrid performance was negatively but not
205 significantly correlated with the genetic relatedness between the parents. Sant et al. [41] and Joyce
206 et al. [42] also reported negative correlations between genetic distance between parents and hybrid
207 performance in Indian elite chickpea cultivars and white clover **(***Trifolium repens* L.), respectively.

208

The correlation between heterosis and genetic distance between the parents has been widely studied and, in many cases, a positive relationship has been established although not sufficient on its own to explain heterosis. Xu et al. [43] reported that the genetic distance between parents based on microsatellite data was significantly positively correlated with hybrid yield/yield heterosis in maize, but the coefficient of determination was low and therefore it was not possible to predict the yield heterosis. Genetic distance based on microsatellites was significantly positively correlated with yield heterosis in rice, but not significantly correlated with heterosis for other traits [19]. The correlation coefficient was however too low to be used to predict heterosis. Dias et al. [44] also observed a positive correlation between genetic distances based on random amplified polymorphic DNA markers and heterosis for wet seed weight per plant and wet seed weight per fruit in cacao. They suggested using this as a guide when choosing superior crosses.

Beche et al. [13] reported a positive and significant correlation between heterobeltiosis and grain yield per plant in spring wheat. They suggested using heterobeltiosis for indirect selection of a trait which positively and significantly correlates with the heterobeltiosis. In our study, bunch weight correlated positively and significantly with heterobeltiosis for bunch weight. Hence, this information assists in the indirect selection of parents that are likely to produce superior hybrids. For example, the parents of NARITA 17 (1438K-1 × 9719-7), 26666S-1(917K-2 × SH 3362), NARITA 9 (917K-2 × SH 3217), NARITA 22 (917K-2 × 9128-3) and 26874S-5 (917K-2 × 5610S-1), which had the highest heterobeltiosis for bunch weight are likely to produce superior hybrids and therefore might be selected for use in future crosses.

As indicated by Xu et al. [42], microsatellite markers showing high polymorphism can be used to assess genetic relationships and are widely used in assessing genetic diversity, identifying germplasm and characterizing population structures. The clustering of accessions in the cladogram based on microsatellite markers (Fig. 1) agreed with the known pedigree information as well as the defined *Musa* groups according to taxonomy. The high genetic variation among the NARITAs was attributed to diverse alleles from the diploid male parents because the 3*x* grandparents and the tetraploid parents clustered together indicating a low genetic diversity among these accessions. Boeven et al. [45] indicated that parents need to be genetically diverse to ensure heterosis in their hybrid offspring. However, genomic-led analysis revealed that diversity does not lead to heterosis [46, 47]. Indeed, there are various reports indicating positive or negative significant correlations between heterosis in hybrid offspring and the genetic distances among their parents. Hence, this association between parental divergence and heterosis does not have to be relevant when pursuing hybrid breeding. Correlations between parental genetic distances and phenotypic hybrid performance have been reported to be very low in most circumstances, which shows that genetic diversity alone is not enough to obtain heterosis. Although the genetic distance does not affect heterosis in a linear fashion, it is still important for obtaining heterosis in crosses. In many circumstances, the expression of heterosis is partly due to genetic diversity which is part of the genomic core for complex interactions of biological pathways that result into increased hybrid vigor.

Conclusion

Heterobeltiosis in high yielding banana hybrids was kept after two crossing generations, thus suggesting a progressive heterobeltiosis. Such a significant heterobeltiosis appears to be the largest among the main food crops as per available literature. Since bananas are vegetatively propagated, the effect of heterobeltiosis is easily fixed in the hybrids and will not be lost over time after release and further commercialization of the hybrids. The factors behind heterobeltiosis in banana are yet to be defined. Nonetheless, leveraging on this high heterobeltiosis there is a huge potential to improve banana production by developing high yielding banana hybrids in relatively few crossbreeding cycles.

Materials and methods

A field experiment was set up in 2015 at Namulonge- Sendusu in Uganda (00°31' 47″ N and 32°36′ 9″ E), comprising 34 NARITA cultivars (26 officially named and 8 not yet officially named), their parents, grandparents and local 'Matooke' banana cultivars as controls (Table 4). These cultivars were planted following a 7 × 8 rectangular lattice design using two replications, with a spacing of 3 m between rows and 2 m between plants within a row, thereby having a plant density of 1667 plants ha^{-1}. Data for bunch weight (kg) were collected at harvest for three crop cycles. Yield potential (t ha^{-1} yr^{-1}) was calculated as:

$$YLD = BW \times 365 \times 1667/ (DH \times 1000)$$

where YLD is yield potential (t ha^{-1} yr^{-1}), BW is bunch weight (kg) and DH is days to harvest.

The mean bunch weights and standard errors were calculated and used to determine heterobeltiosis using the formula:

Heterobeltiosis (%) = [("NARITA" mean bunch weight - "3*x* Grandparent" mean bunch weight)/ "3*x* Grandparent" mean bunch weight] × 100

Plant height and plant girth at 100 cm above the ground were measured at flowering. These data were used to estimate plant stature as the ratio of plant girth to height at flowering, which can be interpreted as a measure of the robustness of the pseudo-stem. The mean plant stature and standard errors were calculated and used to determine heterobeltiosis using the formula:

Heterobeltiosis (%) = [("NARITA" mean plant stature - "3*x* Grandparent" mean plant stature)/ "3*x* Grandmother" mean Plant stature] × 100.

Means of 3*x* grandparents were used to calculate heterobeltiosis of hybrids instead of their parents as the parents are not suitable for consumption and therefore not ideal for comparison. Hence, the type of heterobeltiosis calculated was grandparent heterobeltiosis.

Variance components were estimated using the mixed linear model with restricted maximum likelihood (REML) method as follows:

Model=lmer (Trait~Block+Rep+ (1|Cycle) +(1|Genotype) +(1|Genotype: Cycle), data = data)

Broad sense heritability (H^2) for yield, bunch weight, and plant stature was estimated using the formula:

Heritability = var (Genotype) / [var (Genotype) + var (cycle)/no. of cycles + var (Genotype: Cycle)/no. of years for exp't + var (Residual)/(no. of cycles × no. of years of experiment)]

where var (Genotype) is the variance component of the genotype, var (cycle) is the variance component of the cycle, var (Genotype: Cycle) is the variance component of the interaction

between genotype and cycle, and var (Residual) is the variance component of the residual. Since the data were recorded for 3 cycles during a period of 3 years, the formula used was:

$$\text{Heritability} = \text{var (Genotype)} / [\text{var (Genotype)} + \text{var(cycle)}/3 + \text{var (Genotype:Cycle)}/3 + \text{var(Residual)}/9]$$

Genotyping using SSR

To determine the effect of genetic distance on heterobeltiosis in banana, we genotyped the advanced hybrids (NARITAs), their parents and grandparents using simple sequence repeat (SSR) markers or microsatellites. Fresh young cigar leaf samples were collected from the field in Uganda and shipped under the cold chain to the Institute of Experimental Botany, Olomouc, Czech Republic. Leaf samples were lyophilized in Falcon tubes and stored at room temperature. Approximately 20 mg of lyophilized tissue was crushed into powder in 2 ml Eppendorf tubes using a tissuelyzer. DNA was extracted from tissue powder using NucleoSpin Plant II kit (Macherey-Nagel, Germany) following the manufacturer's instructions. The concentration and quality of DNA was assessed by a NanoDrop ND-1000 spectrophotometer. The working concentration of DNA was adjusted to ~10ng/µl. Genotyping was done using 19 informative *Musa* SSR primers following the protocol of Christelová et al. [48]. Two independent rounds of PCR were performed followed by fragment analysis. Alleles for each sample were inspected in a GeneMarker v1.75 (Softgenetics, State College, PA, USA) and manually scored for presence (1) or absence (0) only when concordance of alleles between PCR runs was observed. In case a sample showed inconsistency in allele sizes between two PCR runs, a third PCR run was performed to confirm the alleles. Squared Euclidean distances between genotypes were calculated using R software v3.4 [49] using the dist function provided in the package 'ape'. Hierarchical clustering was done with

the function hclust based on the ward. D2 method [50, 51]. Pearson's correlation coefficients between grandparent heterobeltiosis for bunch weight and the genetic distances between parents of NARITAs, genetic distance between NARITAs and their female parent's mother, yield, bunch weight and plant stature were calculated. Also, Pearson's correlation coefficients between grandparent heterobeltiosis for plant stature and the genetic distances between parents of NARITAs, genetic distance between NARITAs and their female parent's mother, yield, bunch weight, plant stature and grandparent heterobeltiosis for bunch weight were calculated using R software v3.4 [49].

List of abbreviations

BW: Bunch weight; DH: Days to harvest; GD: Genetic distance; GM: Mother of NARITA's female parent; HGP: Heterobeltiosis with female grandparent; PCR: Polymerase chain reaction; SE: Standard error; SSR: Simple sequence repeat; Var: Variance; YLD: Yield potential

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Not applicable

Competing interests

The authors declare that they have no competing interests.

Funding

This work was part of the Ph.D research of the first author/corresponding author (https://pub.epsilon.slu.se/16129/7/batte_m_190509.pdf), with funding from the International Institute of Tropical Agriculture coordinated Project "Improvement of Banana for Smallholder Farmers in the Great Lakes Region of Africa." The authors would furthermore like to thank all donors who supported this work through their contributions to the CGIAR Fund (https://www.cgiar.org/funders/) and in particular to the CGIAR Research Program for Roots, Tubers and Bananas (CRP-RTB). These funds were used for all the activities of this research involving field trial establishment and management, data collection and analysis.

Authors' contributions

MB, RS, BU, AB, HPH, MG and RO contributed to the conception and design of the study; MB, VA and MN conducted the experiment and collected data; MB and MN performed the statistical analysis; MB wrote the draft manuscript, while all co-authors contributed to manuscript revision

Acknowledgements:

Not applicable

Authors' information

[1] International Institute of Tropical Agriculture (IITA), P.O. Box 7878, Kampala, Uganda

[2] Department of Plant Breeding, Swedish University of Agricultural Sciences (SLU), P.O. Box 101, SE 23053 Alnarp, Sweden

[3] Department of Plant Pathology, Kansas State University, Manhattan, KS 66506, USA

[4] International Institute of Tropical Agriculture (IITA), C/o The Nelson Mandela African Institution of Science and Technology (NM-AIST) P.O. Box 447, Arusha, Tanzania

[5] Laboratory of Tropical Crop Improvement, Katholieke Universiteite Leuven (KUL), Willem De Croylaan 42, 3001 Leuven, Belgium

[6] Bioversity International, Willem De Croylaan 42, 3001 Heverlee, Belgium

Figure captions

Fig. 1. A cladogram showing clustering of NARITAs, their diploid ancestors (parents and 2*x* wild grandparent, 'Calcutta 4'), primary tetraploid parents, and triploid Matooke banana grandparents.

Fig. 2. Progressive heterobeltiosis for bunch weight in cross-bred 'Matooke' banana hybrids (NARITAs), when grown together with their parents and grandparents in Uganda; A: 'Entukura' (3*x* female grandparent), B: '1438K-1' (4*x* female parent) and C: 'NARITA 17' (3*x* hybrid)

References

1 Brown A, Tumuhimbise R, Amah D, Uwimana B, Nyine M, Mduma, H. et al. Bananas and Plantains (*Musa* spp.). In Campos H, Caligari PDS, editors. Genetic improvement of tropical crops. Springer, Cham. 2017. p. 219-40

2 Heslop-Harrison J S, Schwarzacher T. Domestication, genomics and the future for banana. Ann Bot. 2007; 100:1073–1084

3 Nowakunda K, Tushemereirwe W. Farmer acceptance of introduced banana genotypes in Uganda. Afr Crop Sci J. 2004; 12:1–6

4 Swennen R, Blomme G, van Asten P, Lepoint P, Karamura E, Njukwe E, et al. Mitigating the impact of biotic constraints to build resilient banana systems in Central and Eastern Africa. In: Vanlauwe B, Van Asten P, Blomme G, editors. Agro-Ecological Intensification of Agricultural Systems in the African Highlands. New York: Routledge; 2013. p. 85–104

5 Nakato G V, Christelova P, Were E, Nyine M, Coutinho T A, Dolezel J, et al. Sources of resistance in *Musa* to *Xanthomonas campestris* pv. *musacearum*, the causal agent of banana Xanthomonas wilt. Plant Pathol. 2018; 68:49–59. doi: 10.1111/ppa.12945

6 Arango Isaza R E, Diaz-Trujillo C, Dhillon B, Aerts A, Carlier J, Crane C F, et al. Combating a global threat to a clonal crop: banana black Sigatoka pathogen *Pseudocercospora fijiensis* (synonym *Mycosphaerella fijiensis*) genomes reveal clues for disease control. PLoS Genet. 2016; doi: 10.1371/journal.pgen.1005876

7 Alakonya AE, Kimunye J, Mahuku G, Amah D, Uwimana B, Brown A, et al. Progress in understanding *Pseudocercospora* banana pathogens and the development of resistant Musa germplasm. Plant Pathol. 2018; 67:759–770. doi: 10.1111/ppa.12824

8 Viljoen A, Mahuku G, Massawe C, Ssali R T, Kimunye J, Mostert G, et al. Banana diseases and pests: Field guide for diagnostics and data collection. International Institute of Tropical Agriculture, Ibadan, Nigeria;2017.

9 Simmonds N W. Bananas, MUSA cvs. In: Simmonds NW, editor. Breeding for durable resistance in perennial crops. FAO Technical Papers 70. Food and Agriculture Organization, Rome; 1986. p. 17-24

10 Rowe P R. Breeding bananas and plantains for resistance to fusarial wilt: the track record. In: Ploetz R C, editor. Fusarium wilt of bananas. APS, St. Paul, MN. 1990. p. 115-119.

11 Batte M, Swennen R, Uwimana B, Akech V, Brown A, Tumuhimbise R, Hovmalm H P, Mulatu Geleta M, Ortiz R. Crossbreeding East African highland bananas: lessons learnt relevant to the botany of the crop after 21 years of genetic enhancement. Front Plant Sci. 2019; 10:81. https://doi.org/10.3389/fpls.2019.00081

12 Gowda M, Longin C F H, Lein V, Reif J C. Relevance of specific versus general combining ability in winter wheat. Crop Sci. 2012; 52:2494–2500. doi:10.2135/cropsci2012.04.0245

13 Beche E, Lemes da Silva C, Pagliosa E S, Benin G. Hybrid performance and heterosis in early segregant populations of Brazilian spring wheat. Aust J Crop Sci. 2013; 7:51–57

14 Jones D F. Gene action in heterosis. Genetics. 1957; 42:93–103

15 Lamkey K R, Edwards JW. The quantitative genetics of heterosis. In: Coors JG, Pandey S, editors. The genetics and exploitation of heterosis in crops, edited by Crop Science Society of America, Madison, Wisconsin, 1999. p. 31–48

16 Alam MF, Khan MR, Nuruzzaman M. Genetic basis of heterosis and inbreeding depression in rice (*Oryza sativa* L.). J Zhejiang UNIV-SC. 2004; 4:406-411

17 Hochholdinger F, Hoecker N. Towards the molecular basis of heterosis. Trends Plant Sci. 2007; 12:427–432

18 Birchler JA, Yao H, Chudalayandi S, Vaiman D, Veitia RA. A review of heterosis: perspective. Plant Cell. 2010; 22:2105–2112

19 Tao Z, Xian-Lin N, Kai-Feng J, Hua-Feng D, Qing H, Qian-h-Hua Y, Li Y, Xian-Qi W, Ying-Jiang C, Jia-Kui Z. Relationship between heterosis and parental genetic distance based on molecular markers for functional genes related to yield traits in rice. Rice Science. 2010; 17:288–295. https://doi.org/10.1016/S1672-6308(09)60029-9

20 Hinze L L, Lamkey K R. Absence of epistasis for grain yield in elite maize hybrids. Crop Sci. 2003; 43:46–56

21 van Ginkel M, Ortiz R. Cross the best with the best, and select the best: HELP in breeding selfing crops. Crop Sci. 2018; 58:1–14. doi: 10.2135/cropsci2017.05.0270

22 Goff S A. A unifying theory for general multigenic heterosis: Energy efficiency, protein metabolism, and implications for molecular breeding. New Phytol. 2011; 189:923–937. doi:10.1111/j.1469-8137.2010.03574.x

23 Huang X, Yang S, Gong J, Zhao Y, Feng Q, Gong H, et al. Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. Nat Commun. 2015; 6:6258. doi:10.1038/ncomms7258

24 Kaeppler S. Heterosis: Many genes, many mechanisms—End the search for an undiscovered unifying theory. ISRN Bot. 2012; doi:10.5402/2012/682824

441 25 Goldringer I, Brabant P, Gallais A. Estimation of additive and epistatic genetic variances
442 for agronomic traits in a population of doubled-haploid lines of wheat. Heredity. 1997;
443 79:60–71. doi:10.1038/hdy.1997.123

444 26 Aastveit K. Heterosis and selection in barley. Genetics. 1964; 49:159–164

445 27 Smith E L, Lambert J W. Evaluation of early generation testing in spring barley. Crop
446 Sci. 1968; 8:490–493. doi:10.2135/cropsci1968.0011183X000800040029x

447 28 Xiao J, Li J, Yuan L, Tanksley SD. Dominance is the major genetic basis of heterosis in
448 rice as revealed by QTL analysis using molecular markers. Genetics. 1995; 140:745–754

449 29 Garcia AA F, Wang S, Melchinger A E, Zeng Z B. Quantitative trait loci mapping and
450 the genetic basis of heterosis in maize and rice. Genetics. 2008; 180:1707–1724.
451 doi:10.1534/genetics.107.082867

452 30 He Q, Zhang K, Xu C, Xing Y. Additive and additive x additive interaction make
453 important contributions to spikelets per panicle in rice near isogenic (*Oryza sativa* L.)
454 lines. J Genet Genomics. 2010; 37:795–803. doi:10.1016/S1673-8527(09)60097-7

455 31 Tenkouano A, Crouch J H, Crouch H K, Ortiz R. Genetic diversity, hybrid performance,
456 and combining ability for yield in *Musa* germplasm. Euphytica. 1998; 102:281–288

457 32 Groose R W, Talbert L E, Kojis W P, Bingham E T. Progressive heterosis in
458 autotetraploid alfalfa: studies using two types of inbreds. Crop Sci. 1989; 29:1173–1777

459 33 Ortiz R. Secondary polyploids, heterosis, and evolutionary crop breeding for further
460 improvement of the plantain and banana (*Musa* spp. L) genome. Theor Appl Genet. 1997;
461 94:1113–1120

34 Coors J G, Pandey S. The genetics and exploitation of heterosis in crops. Crop Science Society of America, Madison, Wisconsin; 1999.

35 Perrier X, De Langhe E, Donohue M et al. Multidisciplinary perspectives on banana (*Musa* spp.) domestication. Proceedings of the National Academy of Sciences of the USA. 2011; 108: 11311-11318

36 Bakry F, Horry JP. Advances in genomics: applications to banana breeding. Acta Hortic 2016; 1114: 171–180

37 Sardos J, Perrier X, J. Doležel J et al. DArT whole genome profiling provides insights on the evolution and taxonomy of edible Banana (Musa spp.). Ann Bot. 2016; 118: 1269-1278 doi:10.1093/aob/mcw170

38 Kitavi M, Downing T, Lorenzen J, Karamura D, Onyango M, Nyine M, et al., The triploid East African Highland Banana (EAHB) genepool is genetically uniform arising from a single ancestral clone that underwent population expansion by vegetative propagation. Theor Appl Genet. 2016; 129(3):547–561

39 Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, et al., 'Green revolution' genes encode mutant gibberellin response modulators. Nat. 1999; 400:256–261

40 Marcón F, Martínez EJ, Rodríguez GR. et al. Genetic distance and the relationship with heterosis and reproductive behavior in tetraploid bahiagrass hybrids. Mol Breeding. 2019; 39, 89. https://doi.org/10.1007/s11032-019-0994-3

41 Sant V J, Patankar A G, Sarode N D, Mhase L B, Sainani M N, Deshmukh R B, Ranjekar P K, Gupta V S. Potential of DNA markers in detecting divergence and in analysing heterosis in Indian elite chickpea cultivars. Theor Appl Genet. 1999; 98: 1217–1225

42 Joyce T A, Abberton M T, Michaelson-Yeates T P T, Forster J W. Relationships between genetic distance measured by RAPD-PCR and heterosis in inbred lines of white clover **(***Trifolium repens* L.). Euphytica. 1999; 107:159–165

43 Xu S, Liu J, Liu G. The use of SSRs for predicting the hybrid yield and yield heterosis in 15 key inbred lines of Chinese maize. Hereditas. 2004; 141:207–215

44 Dias L, Marita J, Cruz C, Barros E, Salomão T. Genetic distance and its association with heterosis in cacao. Braz Arch Biol Technol. 2003; 46:339–347

45 Boeven P H G, Longin C F H, Würschum T. A unified framework for hybrid breeding and the establishment of heterotic groups in wheat. Theor Appl Genet. 2016; 129:1231–1245. doi:10.1007/s00122-016-2699-x

46 Luo X, Ma C, Yi B, Tu J, Shen J, Fu T. Genetic distance revealed by genomic single nucleotide polymorphisms and their relationships with harvest index heterotic traits in rapeseed (*Brassica napus* L.). Euphytica. 2016; 209:41–47. doi:10.1007/s10681-015-1629-3

47 Tian H Y, Channa S A, Hu S W. Relationships between genetic distance, combining ability and heterosis in rapeseed (*Brassica napus* L.). Euphytica. 2016; 213:1. doi:10.1007/s10681-016-1788-x

48 Christelová P, Valárik M, Hřibová, Van den houwe I, Channelière S, Roux N, Doležel J. A platform for efficient genotyping in *Musa* using microsatellite markers. AoB Plants. 2011; doi:10.1093/aobpla/plr024

49 R Core Team. R: a language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. 2018. http://www.R-project.org/

50 Ward JH. Hierarchical grouping to optimize an objective function. J Am Stat Assoc. 1963; 58:236–244

51 Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? J Classif. 2014; 31:274–295