



OPEN

# Marker based enrichment of provitamin A content in two tropical maize synthetics

Dejene Kebede<sup>2,3</sup>, Wende Mengesha<sup>1✉</sup>, Abebe Menkir<sup>1</sup>, Ayodeji Abe<sup>3</sup>, Ana Luisa Garcia-Oliveira<sup>4</sup> & Melaku Gedil<sup>1</sup>

Most of the maize (*Zea mays* L.) varieties in developing countries have low content of micronutrients including vitamin A. As a result, people who are largely dependent on cereal-based diets suffer from health challenges due to micronutrient deficiencies. Marker assisted recurrent selection (MARS), which increases the frequency of favorable alleles with advances in selection cycle, could be used to enhance the provitamin A (PVA) content of maize. This study was carried out to determine changes in levels of PVA carotenoids and genetic diversity in two maize synthetics that were subjected to two cycles of MARS. The two populations, known as HGA and HGB, and their advanced selection cycles (C1 and C2) were evaluated at Ibadan in Nigeria. Selection increased the concentrations of  $\beta$ -carotene, PVA and total carotenoids across cycles in HGA, while in HGB only  $\alpha$ -carotene increased with advances in selection cycle.  $\beta$ -cryptoxanthine increased at C1 but decreased at C2 in HGB. The levels of  $\beta$ -carotene, PVA, and total carotenoids increased by 40%, 30% and 36% respectively, in HGA after two cycles of selection.  $\alpha$ -carotene and  $\beta$ -cryptoxanthine content improved by 20% and 5%, respectively after two cycles of selection in HGB. MARS caused changes in genetic diversity over selection cycles. Number of effective alleles and observed heterozygosity decreased with selection cycles, while expected heterozygosity increased at C1 and decreased at C2 in HGA. In HGB, number of effective alleles, observed and expected heterozygosity increased at C1 and decreased at C2. In both populations, fixation index increased after two cycle of selections. The greatest part of the genetic variability resides within the population accounting for 86% of the total genetic variance. In general, MARS effectively improved PVA carotenoid content. However, genetic diversity in the two synthetics declined after two cycles of selection.

Maize (*Zea mays* L.) is a major crop for millions of people which depend on it as a prime source of food in sub-Saharan Africa<sup>1,2</sup>. The projected annual maize demand in SSA had been estimated at 500 million tons by the year 2020, which will surpass the demand for both wheat and rice<sup>3</sup>. It is also regarded as a vital crop for global nutrition<sup>1,2</sup>. However, most of the maize varieties grown in these parts of the region are white and they are all devoid of nutritionally important micronutrients such as pro-vitamin A (PVA), zinc, iodine and iron<sup>4</sup>. As a result, the people who depend on maize as a major food source suffer from diseases due to micronutrient deficiencies<sup>5</sup>. Micronutrient deficiencies, mainly caused by lack of vitamin A, are major health problems globally, with high severity in the developing world such as sub-Saharan Africa and South America<sup>6</sup>. Even though the required amount of micronutrients are small, their deficiencies have a wide range of negative health impacts which could eventually result in death<sup>7</sup>. According to<sup>8</sup> 33 to 50% of the world population is estimated to be affected by malnutrition mainly caused by the deficiency of provitamin A.

Vitamin A deficiency (VAD) results from low level of PVA content of major staple food crops such as maize. As reported by<sup>9</sup> more than 250 million children are globally affected by VAD. VAD is responsible for several disorders that include impaired iron mobilization, growth retardation, blindness and depressed immune response, increased susceptibility to infectious disease and childhood mortality and morbidity affecting 190 million preschool-age children and 19 million pregnant women in Africa and South Asia<sup>10,11</sup>. Moreover, the

<sup>1</sup>International Institute of Tropical Agriculture, PMB 5320 Oyo Road, Ibadan, Nigeria. <sup>2</sup>Department of Plant Breeding and Biotechnology, Collage of Agricultural and Environmental Sciences, Makerere University, Kampala, Uganda. <sup>3</sup>Department of Agronomy, University of Ibadan, Ibadan, Nigeria. <sup>4</sup>International Maize and Wheat Improvement Center (CIMMYT), ICRF House, UN Avenue, PO Box, Nairobi 1041-00621, Kenya. ✉email: W.mengesha@cgiar.org

increased susceptibility to several major diseases, such as anemia, diarrhea, measles, malaria and respiratory infections, which accounts for about 70% of childhood deaths throughout the world, is the most disastrous effects of VAD<sup>12,13</sup>.

Different strategies have been employed to alleviate VAD including food supplementation, food fortification, dietary diversification and staple crop bio-fortification<sup>6</sup>. Supplementation and food fortification are expensive, unsustainable and not amenable to implementation in most rural communities. However, bio-fortification through breeding, to improve the nutritional content of staple crops is an effective and sustainable method whereby farmers can easily access PVA foods by growing improved varieties, as nutritionally improved varieties are the long-term solutions to the problem through sustainable consumption of PVA carotenoids from the crops they grow<sup>6,29</sup>. The dietary habit of many people in the world, particularly in sub-Saharan Africa (SSA) where it is consumed daily makes maize a major staple crop of choice for bio-fortification<sup>15,23</sup>.

According to<sup>37</sup>, maize has a substantial genetic variation for kernel carotenoids, with total carotenoid content ranging from virtually zero to 15  $\mu\text{g g}^{-1}$ . The dominant carotenoids in maize kernels, in order of increasing concentration are  $\alpha$ -carotene,  $\beta$ -cryptoxanthine,  $\beta$ -carotene, zeaxanthine, and lutein<sup>15,18</sup>. However, most maize varieties grown and consumed throughout the world have only 0.5–1.5  $\mu\text{g/g}$  of  $\beta$ -carotene<sup>19</sup>. This genetic variability makes it possible to develop maize varieties with increased level of PVA carotenoid content.

Maize improvement for PVA carotenoid through conventional breeding approach involves the use of seed color and quantification of the concentration of carotenoids using high performance or ultra-high-performance liquid chromatography (HPLC or UPLC) laboratory assays. However, selection of genotypes for PVA based on seed color may not be reliable as seed color and PVA carotenoid contents are not always positively correlated with PVA concentration<sup>20</sup>. Furthermore, the cost of carotenoids quantification by HPLC or UPLC assays is too expensive to integrate in routine inbred line development or population improvement<sup>4,14</sup>. Marker-assisted selection (MAS) using DNA markers to select traits of interest could be used to complement conventional approaches. Markers that are tightly linked to target locus or loci are used for the identification of the causal loci which is an important pre-requisite to enable MAS in breeding programs. Genes or loci associated with PVA carotenoids have been identified and markers that aid in the selection of this trait have been developed and validated<sup>4,20,21</sup>. Therefore the use of these DNA markers in MAS has led to accelerated improvement of maize genotypes for their PVA contents<sup>19</sup>. The use of MAS can also shorten the time required for evaluation of breeding materials, improve the efficiency of selection and reduce the cost of screening and development of synthetics.

The use of MARS to develop synthetics with better yield potential and good nutritional quality is an important step in maize breeding. Synthetics can be released as variety per se besides serving as a source of better performing inbred parents<sup>22</sup>. These synthetics were synthesized from provitamin A rich elite inbred lines adapted to low-land tropical environments and are tolerant to major constraints in their target environments. Development of advanced inbred lines for the formation of provitamin A rich maize hybrids with better agronomic performance through recurrent selection has a paramount importance in improving the food and nutritional security of the target population.

Recurrent selection is a breeding procedure involving cycles of hybridization, selection and recombination and used to improve breeding populations and inbred lines for combining ability in the hybrid breeding program. MARS is one of the commonly used methods to increase the frequency of favorable alleles in subsequent selection cycles while maintaining genetic diversity<sup>23</sup>. Reported increased  $\beta$ -cryptoxanthine,  $\beta$ -carotene and zeaxanthine contents after three cycles of S1 recurrent selection. The study also indicated that this selection method was effective in improving the PVA content of different maize populations. Rapid cycling recurrent selection has also been employed to improve the total carotenoid content (TTC) as well as  $\beta$ -carotene in cassava<sup>24</sup>. Two cycles of marker assisted recurrent selection have been carried out to improve pro-vitamin A content in two maize synthetics. The present study was, therefore, conducted to determine the genetic gain in pro-vitamin A (PVA) content and examine changes in genetic diversity in two synthetics after two cycles of selection.

## Materials and methods

**Marker assisted enrichment of PVA.** Two maize synthetics (HGA and HGB) were subjected to recurrent selection to improve their provitamin A content. The synthetics were developed from a cross of eight provitamin A rich inbred lines belonging to each heterotic group (HGA or HGB). There was no allelic information for the eight inbred lines constituting the synthetics. However, the inbred lines were subjected to phenotypic selection based on HPLC information regarding their carotenoid content, and inbred lines with  $> 15 \mu\text{g g}^{-1}$  were used to constitute the synthetics. Moreover, visual selection based on kernel color (deep orange kernel) were also used to select the inbred lines. The recurrent selection was carried out in IITA for two cycles. The plant selection was carried out using PVA molecular markers of *lycE* and *crtRB1* genes and complemented with the phenotypic performance of plants. Sixty rows were planted from each population in 5 m long rows spaced 0.75 m apart and 0.25 m within a row. Two hundred and eighty-eight plants were tagged and plants that had at least two favorable alleles of the PVA markers and desirable ears were selected from each of HGA and HGB. The selected 50 and 52 plants were recombined to form the first cycle (C1) of HGA and HGB. To form the C2, 60 rows of plants from C1 of each HGA and HGB were planted. Two hundred and eighty-eight (288) plants were tagged and plants that had at least two favorable alleles of the PVA markers with good agronomic performances were selected from HGA and HGB, respectively. The selected 151 and 126 plants were recombined to form the C2 of HGA and HGB. Pollination was done by using bulk pollen from male rows to pollinate the female rows in both cases. Different numbers of plants and ears with favorable alleles were selected and recombined in each cycle for both synthetics (Table 1). The genetic materials used in the present study involved neither wild species nor landrace collections received from gene banks. Consequently, permission to collect samples for carotenoid analysis was not required

	Total number of plants selfed and genotyped in each cycle	Selected plants carrying favorable alleles		Number of ears selected and recombined	
		C1	C2	C1	C2
HGA	288	60	193	50	151
HGB	288	197	165	52	126

**Table 1.** Number of plants (with favorable alleles) and ears selected for each synthetic per cycle.

from gene banks. All experiments were conducted in line with relevant international, national and institutional guidelines and legislation.

**Experimental design and field layout.** An experiment consisting six rows of each MARS cycle of the two synthetic of different selection cycles were planted at IITA, Ibadan. The experiment was arranged in randomized complete block design (RCBD) with three replications and was planted in 5 m long rows with 0.75 m space between rows and 0.25 m within a row. NPK15:15:15 fertilizer was applied at the rate of 60 kg N ha<sup>-1</sup>, 60 kg P ha<sup>-1</sup>, and 60 kg K ha<sup>-1</sup>, at planting and an additional 60 kg N ha<sup>-1</sup> was applied four weeks later. Manual weeding was done to keep the experiments weed free. The middle four rows were self-pollinated, and seed samples were collected for carotenoids analysis from self-pollinated plants.

**Analysis of carotenoids.** The extraction protocol and carotenoid analysis used was the method described in<sup>25</sup>. Briefly, 0.5 g finely ground sample of each entry was transferred into a 50 ml glass centrifuge tube to which 6 ml of Ethanol plus 0.1% butylated hydroxyl toluene were added, vortexed for 15 s, and incubated in 85 °C water bath for 5 min. 500 µl of 80% potassium hydroxide (w/v) was added to each sample, vortexed for 15 s, and incubated in 85 °C water bath for 10 min with vortexing at about 5 min interval. Samples were then immediately placed on ice and 3 ml ice cold deionized water added to each of them, vortexed for 15 s, and 200 µl internal standard β-Apo-8'-carotenal and 4 ml hexane added. After vortexing and centrifugation, the top hexane layer formed was transferred into a new test tube. The hexane extraction was repeated twice, adding 3 ml hexane each time. Samples were allowed to dry down completely under nitrogen gas using a Turbovap LV concentrator (Caliper Life Sciences) and reconstituted in 500 µl of 50:50 Methanol: Dichloroethane.

Fifty micro-liter aliquots of each extract were injected into an HPLC system (Water Corporation, Milford, MA). The Water's HPLC components was operated with Empower 1 software and included a 717 Plus auto sampler with temperature control set at 5 °C, Waters 1525 binary HPLC pump, and a 2996 photodiode array detector for carotenoid quantification. Carotenoids were separated by C30 Carotenoid Column (4.6 × 250 mm; 3 µm) eluted by a mobile phase gradient from 100% methanol/water (92:8 v/v) with 10 mM ammonium acetate to 50% methyl tertiary butyl ether. The flow rate was 1.0 mL/min and the solvents were HPLC grade. To maximize detection of carotenoids, absorbance was measured at 450 nm. Alpha-carotene, β-carotene (cis and trans isomers), β-cryptoxanthin, lutein, and zeaxanthin were quantified. Total carotenoid was calculated as the sum of concentrations of α-carotene, lutein, β-carotene, β-cryptoxanthin and zeaxanthin. Provitamin A was calculated as the sum of β-carotene and half of each of β-cryptoxanthin and α-carotene concentrations, since the latter two contribute 50% of the value of β-carotene as provitamin A<sup>20</sup>. Other derived carotenoid traits were also calculated as indicated in<sup>20,21</sup>, namely the ratio of carotenoids in β to a branch of the carotenoid pathway, the ratio of β-carotene to β-cryptoxanthin and the ratio of β-carotene to total carotenoids. The natural logarithms of the ratios were calculated to allow statistical analysis of the data, as the ratios followed a highly non-normal distribution. All concentrations were described in µg g<sup>-1</sup> dry weight (DW).

**DNA extraction and genotyping.** In cycle one, 288 individual plants were genotyped from HGA with markers crtRB1-5' TE, LycE-3' Indel, LycE-SNP(216), whereas 197 individual plants from HGB were genotyped with markers crtRB1-3' TE, LycE-3' Indel, LycE-SNP(216) as described in similar studies<sup>26</sup>. Previous studies have confirmed that crtRB1-3' TE and crtRB1-5' TE are in LD<sup>26</sup>. In cycle 2, the same genotyping method was used but with only two markers from each gene—crtRB1-3' TE and LycE-SNP(216). For KASP assay, DNA was extracted, and the quality of DNA checked by running 2 µl of the diluted DNA sample on 1% agarose gel and quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., Denver, CO, USA). Genomic DNA samples were lyophilized to dry powder and sent to LGC genomics (UK) for single nucleotide polymorphism (SNP) genotyping using KASP assay platform.

**Deployment of high throughput SNP genotyping assay.** CrtrB1-KASP specific PVA SNP markers developed by CIMMYT were used to evaluate the improvement of PVA across the different cycles for HGA and HGB populations. These markers fall on chromosome 10 of the crtRB1 gene, which regulates the hydroxylation of β-carotene to β-cryptoxanthin. Studies indicated that, this region is responsible for higher β-carotene accumulation in maize. Selection for the favorable SNPs of these markers are thought to enhance β-carotene and PVA content. Hence, the markers were employed for genotyping of the two synthetic populations. KASP assay genotyping platform was used for the genotyping. Alignment blast of all 7 SNPs to the crtRB1 gene nucleotide

Source of variation	DF	Mean squares of carotenoids						
		Lutein ( $\mu\text{g g}^{-1}$ )	Zeaxanthine ( $\mu\text{g g}^{-1}$ )	$\beta$ -cryptoxanthine ( $\mu\text{g g}^{-1}$ )	$\alpha$ -carotene ( $\mu\text{g g}^{-1}$ )	$\beta$ -carotene ( $\mu\text{g g}^{-1}$ )	provitamin A ( $\mu\text{g g}^{-1}$ )	Total carotenoid ( $\mu\text{g g}^{-1}$ )
Run	1	0.021	0.001	0.0001	0.005	0.01	0.02	0.002
Rep (Run)	4	5.9	4.2	1.6**	0.2*	9.1	9.7	92.1**
Cycle	5	196.5***	4.7***	0.6***	0.05*	56.5***	51.6***	350.7***
Run*Cycle	5	0.003	0.002	0.0004	0.009	0.06	0.083	0.12
Error	20	3.4	2.5	0.3	0.05	7.1	6.1	21.7

**Table 2.** Mean squares from analysis of variance for pro-vitamin A and other carotenoids of HGA and HGB populations improved through marker assisted recurrent selection. \*, \*\*, \*\*\* = Significant at  $P < 0.05$ ,  $P < 0.01$ , and  $P < 0.001$ , respectively.

sequence showed that only SNP ZM0015 is located on the gene. The rest were either upstream or downstream of the crtRB1 gene.

**Assessment of genetic diversity after two cycles of selection.** In order to examine the changes in allele frequencies caused by MARS after two cycles of selection, single nucleotide polymorphic (SNP) markers distributed across the genome were selected based on information from previous studies and SNPs routinely used as quality control (QC) set in IITA Bioscience Center. In total, 14–23 SNP markers from each chromosome were selected. KASP assay genotyping platform was employed for genotyping.

**Statistical analysis.** A combined analysis of variance for PVA and carotenoids were computed with PROC GLM in SAS 9.3 using a mixed model<sup>26</sup>, considering cycles as fixed effects and blocks and runs as random effects. The significance of mean squares for the main and interaction effects were tested using the appropriate mean squares, obtained from the RANDOM option in SAS 9.3<sup>26</sup>. Repeatability values for PVA and each carotenoid were estimated using PROC MIXED procedure in SAS 9.3<sup>26</sup>. Mean separation was done using Least Significant Difference (LSD) at 0.05 level of probability<sup>26</sup>.

For each trait, cycle means were regressed as dependent variables on cycle numbers (0, 1, and 2) as independent variables using PROC Reg procedure in SAS 9.3<sup>26</sup>. The coefficient of linear regression ( $b$  value) provided an estimate of the gain per cycle, which was divided by the intercept and multiplied by 100 to obtain the percent response per cycle<sup>27</sup>.

GenAlEx 6.5 software<sup>28</sup> was used to analyze the frequency of favorable SNP alleles in the course of recurrent selection in both populations. Out of the six crtRB1-KASP specific PVA markers used for genotyping one marker was excluded from the analysis due to high missing data. SNPs missing data greater than 20% in a population were excluded from the analysis.

Out of the 185 SNP markers used for the genetic diversity assessment, fifteen were missing. An additional three markers exhibited greater than 20% missing data and were excluded from further analysis. Genetic diversity indices such as number of effective alleles ( $A_e$ ), average observed heterozygosity ( $H_o$ ), expected heterozygosity, fixation index and percent of polymorphic loci were calculated using GenAlEx 6.5<sup>28</sup>.

Analysis of molecular variance (AMOVA) was used to partition genetic variation among cycles and within cycles and Pairwise population  $F_{st}$  were analyzed using GenAlEx 6.5. Genetic distance among cycles was estimated according to<sup>29</sup> method in GenAlEx 6.5. Percent of missing data for each cycle was also computed by taking the proportion of missing markers out of the total number of markers for each cycle.

**Human and animal rights.** This research does not involve human/animal.

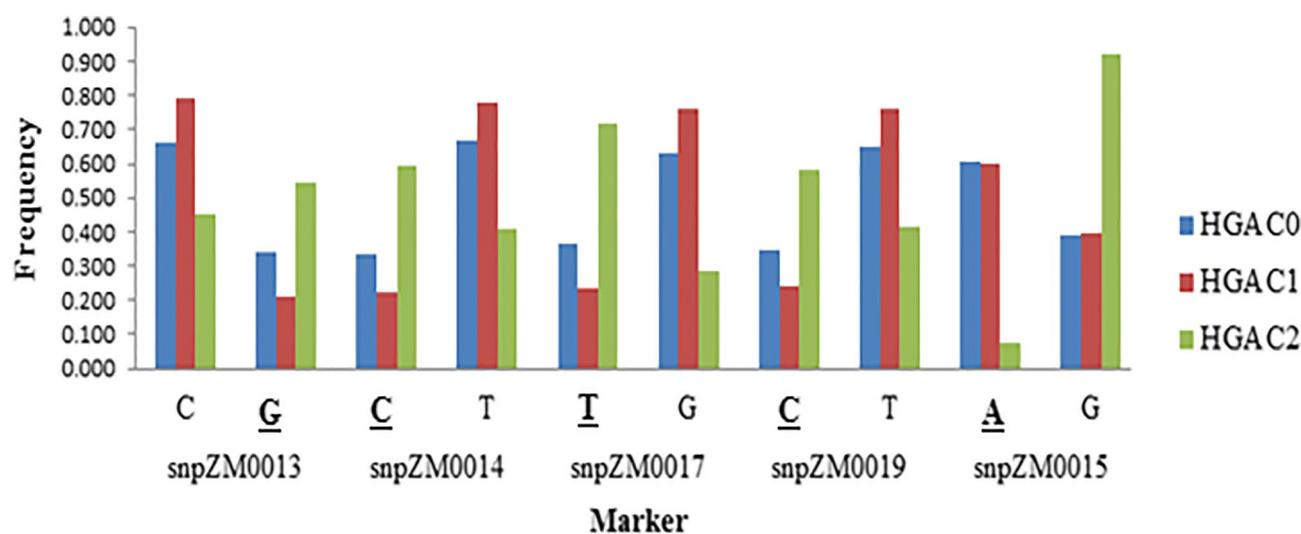
## Results

**Genetic gain for pro-vitamin A and other carotenoids.** In the combined analysis of variance selection cycle was a significant source of variation for the five carotenoids, including pro-vitamin A (PVA) and total carotenoids (Table 2). Run  $\times$  block interaction was significant for  $\beta$ -cryptoxanthine,  $\alpha$ -carotene and total carotenoids (Table 2). Run and run  $\times$  cycle interaction was not significant for all carotenoids (Table 2). Higher repeatability estimates were observed for all carotenoids including PVA and total carotenoids except for  $\alpha$ -carotene (Table 3). The number of plants with favorable allele and ears selected in HGA was higher in C2 as compared to the previous cycle (C1). In HGB there was no consistent increase in numbers of plants selected from C1 to C2, however the number of ears selected in C2 were higher than the ones selected in C1.

The content of  $\beta$ -carotene, PVA and total carotenoids increased with selection cycles in HGA (Table 3). After two cycles of selection the levels of  $\beta$ -carotene, PVA and total carotenoids were increased by 40%, 30% and 36%, respectively in HGA. In HGB, only the content of  $\alpha$ -carotene increased with selection cycle, while  $\beta$ -cryptoxanthine increased at C1 and decreased at C2 (Table 3). In HGA, the genetic gains per cycle for  $\beta$ -carotene, PVA and total carotenoid were  $2.6 \mu\text{g g}^{-1}$ ,  $2.5 \mu\text{g g}^{-1}$  and  $7.8 \mu\text{g g}^{-1}$ , respectively. The content of  $\alpha$ -carotene and  $\beta$ -cryptoxanthine were improved by 20% and 5%, respectively, after two cycles of selection in HGB. The genetic gain per cycle was small ( $0.1 \mu\text{g g}^{-1}$ ) for both  $\alpha$ -carotene and  $\beta$ -cryptoxanthine in HGB (Table 3). Even if the PVA and total carotenoids were not consistently increased in HGB, other important carotenoids ( $\alpha$ -carotene

Cycle	Lutein ( $\mu\text{g g}^{-1}$ )	Zeaxanthine ( $\mu\text{g g}^{-1}$ )	B-cryptoxanthine ( $\mu\text{g g}^{-1}$ )	a-carotene ( $\mu\text{g g}^{-1}$ )	$\beta$ -carotene ( $\mu\text{g g}^{-1}$ )	PVA ( $\mu\text{g g}^{-1}$ )	Total carotenoid ( $\mu\text{g g}^{-1}$ )
HGAC0	11.3	12.7	4.6	1.1	13.7	16.5	43.3
HGAC1	10.5	11.0	4.7	1.0	14.0	16.8	41.2
HGAC2	23.6	11.2	3.9	1.2	18.9	21.5	58.8
HGBC0	22.3	11.7	4.2	1.0	19.3	21.9	58.4
HGBC1	13.1	13.3	4.7	1.1	12.0	14.9	44.2
HGBC2	18.8	11.4	4.4	1.2	17.6	20.4	53.4
Mean	16.6	11.9	4.4	1.1	15.9	18.7	49.9
Gain cycle <sup>-1</sup> HGA	6.2	- 0.7	- 0.4	0.04	2.6	2.5	7.8
% Resp cycle <sup>-1</sup> HGA	69.7	- 5.8	- 7.5	3.9	20.3	15.6	19.4
Gain cycle <sup>-1</sup> HGB	- 1.7	- 0.1	0.1	0.1	- 0.9	- 0.8	-2.5
% Resp cycle <sup>-1</sup> HGB	- 8.8	- 1.1	2.2	11.1	- 5	- 3.8	-4.6
Repeatability	0.99	0.58	0.64	0.10	0.90	0.91	0.96
CV	11	13	12	21	17	13	7.1
LSD(0.05)	1.9	ns	ns	ns	2.8	2.6	6.4

**Table 3.** Means and genetic gains for PVA and other carotenoids of HGA and HGB and their derived cycles improved through marker assisted recurrent selection. *Resp* response, *ns* non-significant.

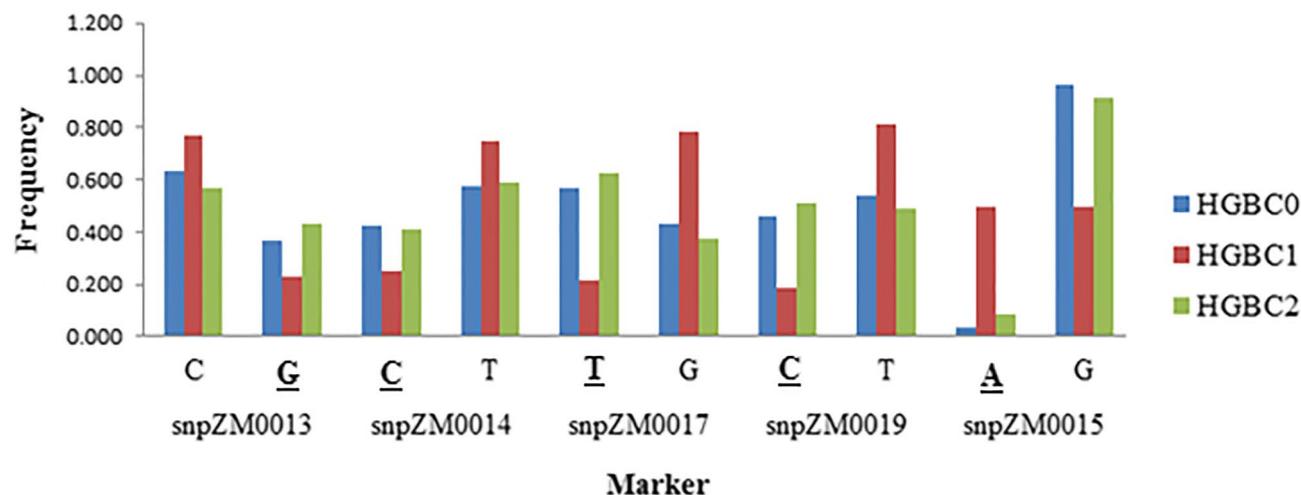


**Figure 1.** Allele frequency changes of crtRB1-KASP specific PVA marker alleles in HGA and its derived cycles. \*The favorable alleles are bolded and underlined.

and  $\beta$ -cryptoxanthine) showed an increase after two cycles of selection, which confirms the possibilities of enriching the two synthetics through recurrent selection.

**Changes in frequency of favorable SNP marker alleles.** The mean frequency of the favorable SNPs of four crtRB1-KASP specific markers namely; snpZM0013, snpZM0014, snpZM0017 and snpZM0019, reduced at C1 and then increased at C2 in HGA (Fig. 1). In HGB, the mean frequency of the favorable alleles of three markers (snpZM0013, snpZM0017 and snpZM0019) decreased at C1 and increased at C2 (Fig. 2). However, the frequency of the favorable SNP of snpZM0015 decreased in the advanced cycle of selection (C2) in both populations. The frequency of the favorable alleles of snpZM0014 did not improve after two cycles of selection in HGB (Fig. 2). None of the marker loci got fixed in the course of selection.

**Genetic diversity study with SNP markers.** In order to evaluate genetic diversity changes that occurred in the course of selection, randomly selected 60 plants from each cycle of the two synthetics were genotyped using SNP markers. Number of effective alleles and observed heterozygosity decreased with selection cycles, while expected heterozygosity increased at C1 and decreased at C2 in HGA (Table 4). On the other hand, in HGB the observed heterozygosity, number of effective alleles and expected heterozygosity slightly increased at C1 but decreased at C2 (Table 5). The percentage of polymorphic loci was slightly higher in selection cycles of



**Figure 2.** Allele frequency changes of crtRB1-KASP specific PVA marker alleles in HGB and its derived cycles. \*The favorable alleles are bolded and underlined.

Cycles	# Samples <sup>1</sup>	Missing data (%) <sup>2</sup>	% Polymorphic loci <sup>3</sup>	# Effective alleles <sup>4</sup>	Observed heterozygosity <sup>5</sup>	Expected heterozygosity <sup>6</sup>	Fixation Index <sup>7</sup>
HGAC0	60	4.6	93	1.6 ± 0.026	0.33 ± 0.015	0.32 ± 0.014	0.01 ± 0.018
HGAC1	60	3.9	94	1.6 ± 0.027	0.29 ± 0.013	0.33 ± 0.013	0.11 ± 0.019
HGAC2	60	5.5	93	1.5 ± 0.027	0.24 ± 0.013	0.27 ± 0.013	0.13 ± 0.022
Mean	60	4.7	93	1.5 ± 0.016	0.29 ± 0.008	0.31 ± 0.008	0.08 ± 0.012

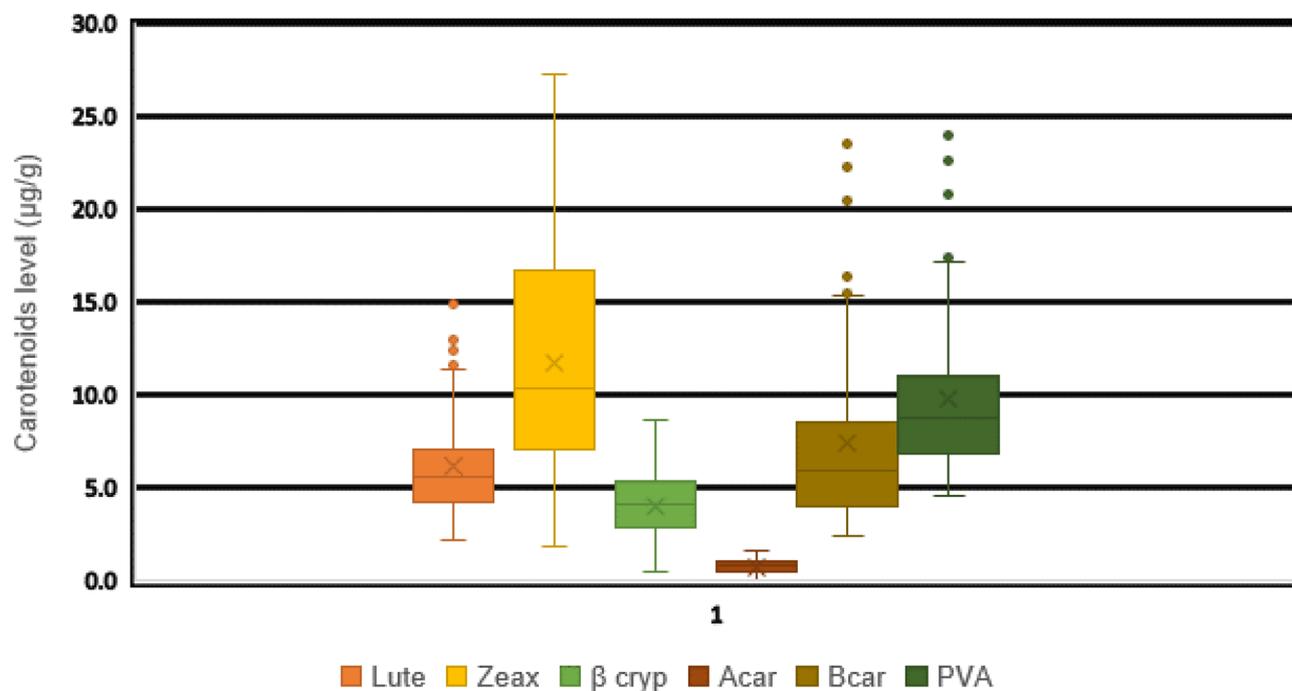
**Table 4.** Genetic variability based on SNP markers computed for HGA and its selection cycles. <sup>1</sup># Number of samples correspond to number of plants from which leaf tissue was assayed. <sup>2</sup>Percent (%) of missing data is the percentage of markers that did not return data computed per cycle. <sup>3</sup>Percent of polymorphic loci is the percentage of polymorphic loci to the total number of loci (polymorphic and monomorphic) per cycle. <sup>4</sup>Number of effective alleles are the number of alleles that can be present in the population. <sup>5</sup>Observed heterozygosity is the number of markers within a population that are heterozygous calculated per cycle based. <sup>6</sup>Expected heterozygosity is the probability that at a single locus, any two alleles chosen at random from the population are different to each other. <sup>7</sup>Fixation index is the measure of reduction in heterozygosity.

Cycles	# Samples <sup>1</sup>	Missing data (%) <sup>2</sup>	% Polymorphic loci <sup>3</sup>	# Effective alleles <sup>4</sup>	Observed heterozygosity <sup>5</sup>	Expected heterozygosity <sup>6</sup>	Fixation Index <sup>7</sup>
HGBC0	60	3.7	88.6	1.5 ± 0.027	0.25 ± 0.015	0.27 ± 0.014	0.081 ± 0.02
HGBC1	60	6.7	93.4	1.6 ± 0.027	0.3 ± 0.014	0.32 ± 0.013	0.054 ± 0.01
HGBC2	60	5.8	94.6	1.5 ± 0.026	0.26 ± 0.014	0.29 ± 0.013	0.083 ± 0.021
Mean	60	5.4	92.2	1.5 ± 0.016	0.27 ± 0.008	0.29 ± 0.008	0.07 ± 0.011

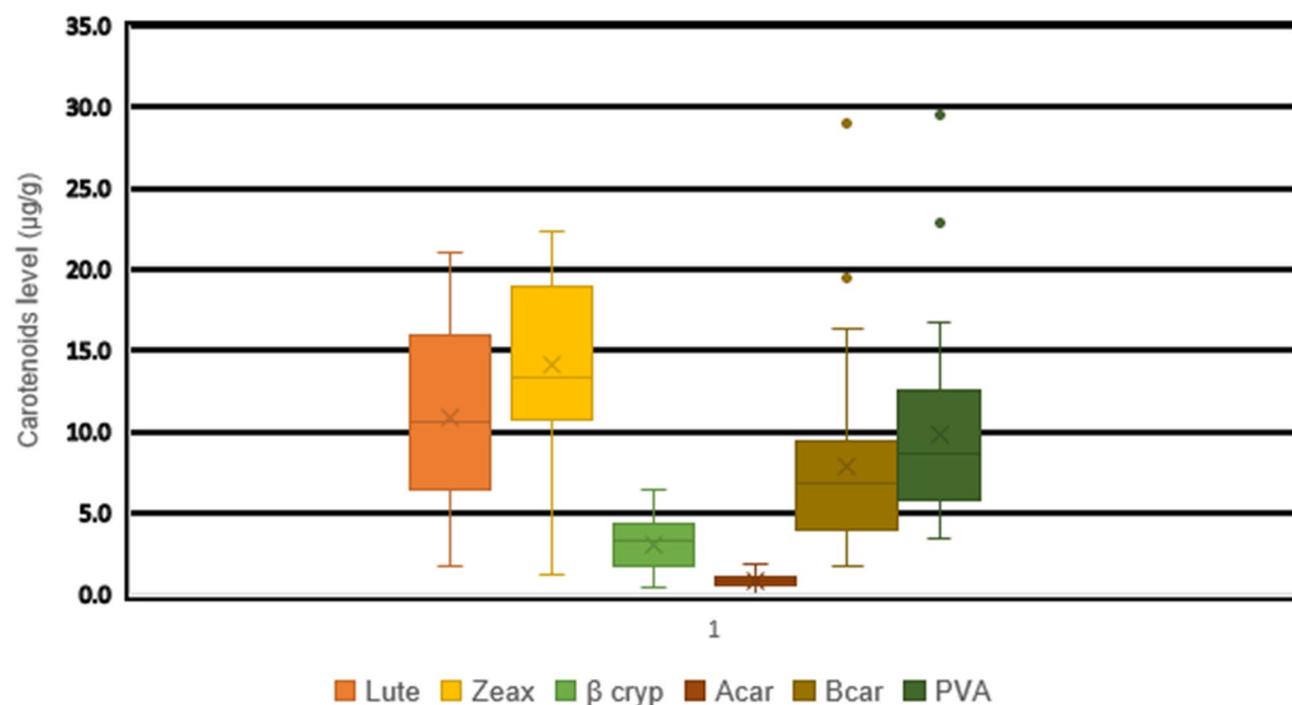
**Table 5.** Genetic variability based on SNP markers computed for HGB and its selection cycles. <sup>1</sup># Number of samples correspond to number of plants from which leaf tissue was assayed. <sup>2</sup>Percent (%) of missing data is the percentage of markers that did not return data computed per cycle. <sup>3</sup>Percent of polymorphic loci is the percentage of polymorphic loci to the total number of loci (polymorphic and monomorphic) per cycle. <sup>4</sup>Number of effective alleles are the number of alleles that can be present in the population. <sup>5</sup>Observed heterozygosity is the number of markers within a population that are heterozygous calculated per cycle based. <sup>6</sup>Expected heterozygosity is the probability that at a single locus, any two alleles chosen at random from the population are different to each other. <sup>7</sup>Fixation index is the measure of reduction in heterozygosity.

HGA than those of HGB. The fixation index varied from 0.01 to 0.13 in HGA and 0.081 to 0.083 in HGB selection cycles.

**Carotenoid concentrations of early generation lines derived from HGA and HGB.** The result of analyses showed the presence of considerable variability among the inbred lines for important carotenoids considered in this study (Figs. 3 and 4). The genetic variability among the S4 lines derived from HGA was broader for Zeaxanthin and B-carotene. In HGB the genetic differences among inbred lines was higher for lutein and



**Figure 3.** Distribution of the different carotenoids across the S4 inbred lines derived from HGA.



**Figure 4.** Distribution of the different carotenoids across the S4 inbred lines derived from HGB.

zeaxanthin (Figs. 3 and 4). The differences among S4 lines was small for  $\alpha$ -carotene content in the two synthetics. Inbred lines derived from the two synthetics (HGA and HGB) displayed considerable variability in provitamin A content. Thirty-one S4 lines in HGA and 13 S4 lines in HGB had provitamin A contents exceeding 10 µg/g. Fourteen inbred lines in HGA and 4 in HGB had provitamin A contents of more than 15 µg/g, which is above the target level set under HarvestPlus Global Challenge Program (Figs. 3 and 4).

## Discussion

The present study was conducted to determine the genetic gain for pro-vitamin A (PVA) content and examine the extent of change in genetic diversity in two synthetics. The two synthetics were formed from pro-vitamin A rich maize inbred lines belonging to different heterotic groups developed at IITA and they were improved for pro-vitamin A content through two cycles of MARS. The observed significant difference between selection cycles for all carotenoids as well as PVA and total carotenoids suggest that selection induced significant changes in carotenoid accumulations across selection cycles, consistent with the results reported earlier<sup>23</sup>. The higher repeatability estimates for content of all carotenoids as well as PVA and total carotenoids indicate that these nutrients were accurately measured. This result was consistent with that reported by<sup>30</sup> who observed higher repeatability estimates of carotenoids concentrations for major carotenoids including PVA and total carotenoids in maize.

Linear increase was observed for  $\beta$ -carotene, PVA and total carotenoids with selection in HGA, consistent with the results reported by<sup>23,24</sup> also demonstrated significant improvement for  $\beta$ -carotene and total carotenoids with selection cycles in cassava in rapid cycling recurrent selection program. The content of PVA and other carotenoids were higher in the base population but decreased with advances in selection cycle in HGB. The possible explanation for this could be loss of plants with favorable alleles during the process of selection for desirable ears. Since a combination of molecular and phenotypic selection of ears for bright yellow to orange color and semi flint to flint kernel texture were employed for selection of plants for recombination, some plants with favorable PVA alleles could have been lost.<sup>20</sup> observed that selection of plants for PVA content based on seed color may not be reliable as seed color and PVA carotenoid contents are not always positively correlated.

<sup>31</sup> reported the existence of duplicated beta-carotene hydroxylase genes elsewhere in the maize genome, suggests a complex regulatory system for maize carotenoid biosynthesis, with each hydroxylase gene playing functionally different roles. Such scenario is the most probable case in our study, where there was steady increase in provitamin A content in HGA, but showed fluctuation in case of HGB. Likewise, the observed reduction in PVA content at an advanced cycle (C2) of HGB may also be due to the linkage drag from the provitamin A donor parents or loss of diversity associated to intensive selection on the synthetics<sup>31</sup>. The other potential causes of inconsistency could also arise from random changes in allele frequencies or genetic drift, which can cause changes in population means, particularly for a relatively small sized synthetics typical of our recurrent selection experiments<sup>32</sup>. Larger population size would be the correct strategy to overcome the problem of segregation distortion effects on selection in the breeding program for high provitamin A level using markers. However, our result still confirms the possibilities of developing products with higher carotenoid contents through the deployment of recurrent selection using relatively narrow based synthetics as a base population.

The current study is in line with<sup>31</sup> who reported the results of 26 tropical maize populations, demonstrating the effectiveness of marker-assisted selection (MAS) for alleles of *cr1RB1*, which is increasingly being implemented by breeding programs involved in the development of varieties with enhanced levels of provitamin A and other essential carotenoids. In this study,  $\beta$ -carotene was more abundant than other PVA carotenoids and increased consistently with selection cycles in HGA. After two cycles of MARS, the  $\beta$ -cryptoxanthine content did not increase in HGA and its improvement was very small in HGB. Similarly, the content of zeaxanthin did not increase after two cycles of selection in both populations. The result of this study suggests that hydroxylation reducing alleles of *cr1RB1* were selected in the process of the MARS population development.

Recently, CIMMYT has developed *cr1RB1* KASP specific SNP markers which are being used to screen and evaluate maize breeding populations for PVA due to their cost effectiveness and shortening the time required for genotyping. Selection for the favorable SNPs of these markers enables the selections of favorable alleles of *cr1RB1* (5'TE, InDel4 and 3'TE) that could double the  $\beta$ -carotene content in maize endosperm (Yan et al., 2010). In this study, the frequency of the *cr1RB1* KASP SNP markers associated with favourable alleles for PVA increased after two cycles of MARS. Out of the six *cr1RB1* KASP markers used for genotyping, the frequency of favorable alleles of four and three markers increased after two cycle of selection in HGA and HGB, respectively. The gradual increase in the frequency of favorable alleles was in line with the findings of<sup>27,33</sup> who observed that MARS increased the frequencies of the favorable allele.

After two cycles of MARS, the PVA content of HGA was  $21.5 \mu\text{g g}^{-1}$ , which is higher than the  $17.3 \mu\text{g g}^{-1}$  previously reported by<sup>26</sup> for PVA inbred lines. This result was also greater than the PVA target level of  $15 \mu\text{g g}^{-1}$  set by HarvestPlus Global Challenge Program to overcome vitamin A deficiency<sup>19</sup>. The higher response per cycle of selection and genetic gain observed for PVA and other carotenoids suggest that marker assisted recurrent selection was effective in improving the PVA content and other carotenoids when there is adequate diversity in a population under selection<sup>23</sup>.

The genetic variability in a population could possibly be maintained through random mating of individuals which was the case in the present study where the observed heterozygosity and number of effective alleles slightly increased from C0 to C1 and decreased at C2 in HGB. However, there was reduction in the number of effective alleles, observed heterozygosity and expected heterozygosity after two cycle of marker assisted recurrent selection for HGA, which is consistent with other earlier works (Vogel 2010). The reason for the reduction of these genetic diversity indices could be the effect of changes in frequency of favorable and unfavorable alleles<sup>27</sup>. It could also be because of selection and genetic drift<sup>4,35</sup>. A similar result was reported by<sup>33,36</sup>. The average percent polymorphic loci were higher in HGA (94%) than HGB (92.2%) suggesting that HGA had higher genetic variability than HGB. Expected heterozygosity was generally greater than observed heterozygosity within selection cycles for both HGA and HGB. This might suggest that the two populations were different with different alleles being fixed at advanced selection cycles. This result was consistent with that of<sup>35</sup> who reported larger expected heterozygosity than observed heterozygosity for two maize RRS selection populations.

The increase in fixation index in the course of selection suggests that homozygosity increased at the advanced selection cycle<sup>37</sup>.<sup>38</sup> who studied the genetic diversity of two composite maize recurrent selection populations

reported an increase in the proportion of homozygotes with selection cycles. The decrease in the average effective number of alleles from 1.6 at C0 to 1.5 at C2 in HGA as well as the increase in fixation index after two cycles of selection in both populations suggests that the populations are gradually moving towards allele fixation<sup>34</sup>. The reduction in the level of heterozygosity and increased fixation index from C0 to C2 imply loss of heterozygosity possibly because of positive assortative mating and selection of homozygotes<sup>28</sup>. The reduction in expected heterozygosity and the increase in the level of inbreeding resulting from response to selection were in line with the findings of<sup>33</sup> who observed a decrease in level of gene diversity, increase level of inbreeding and changes in allele frequency in a bi-parental maize population improved through MARS.

Allele frequency changes allows for the identification of favorable or unfavorable alleles. <sup>39,40</sup> found a high mean frequency at C1 while a third population showed no difference in mean frequency of favorable alleles between C1 and C2. Our results suggest that breeders may conduct two cycles of MARS to develop superior inbred lines with enhanced levels of essential carotenoids including provitamin A. The combination of favorable alleles in the present study reveals that majority of the S4 inbred lines derived from both synthetics, had higher frequencies of favorable alleles for provitamin A and other important carotenoids. This identification of S4 lines with provitamin A contents of up to 29 µg/g suggested that breeding scheme was effective in enhancing genetic gain for carotenoid contents in the synthetics through recombination of alleles.

In summary, improvement in PVA and changes in genetic diversity of the two maize synthetics subjected to MARS was investigated. The PVA content increased with advances in selection cycles in HGA. MARS effected desirable changes in the frequency of favorable alleles for PVA in both populations. It was therefore, concluded that it can be used for improving PVA content of tropically adapted maize germplasm.

Received: 5 February 2021; Accepted: 7 July 2021

Published online: 22 July 2021

## References

1. International Institute of Tropical Agriculture (IITA). New varieties to boost maize output in West and Central Africa [http://www.iita.org/maize/-/asset\\_publisher/jeR0/content/new-varieties-to-boost-maize](http://www.iita.org/maize/-/asset_publisher/jeR0/content/new-varieties-to-boost-maize) (2010).
2. Nuss, E. T. & Tanumihardjo, S. A. Maize: a paramount staple crop in the context of global nutrition. *Compr. Rev. Food Sci. Food Saf.* **9**, 417–436 (2010).
3. FAOSTAT. Statistical database of the food and agriculture of the United Nations. <http://www.fao.org> (2018).
4. Babu, R., Rojas, N. P., Shibin, G., Yan, J. & Pixley, K. Validation of the effects of molecular marker polymorphisms in LcyE and crtRB1 on provitamin A concentrations for 26 tropical maize populations. *Theor. Appl. Genet.* **126**(2), 389–399 (2013).
5. Mayer, J. E., Pfeiffer, W. H. & Beyer, P. Biofortified crops to alleviate micronutrient malnutrition. *Curr. Opin. Plant Biol.* **11**, 166–170 (2008).
6. Bouis, H. E., Hotz, C., McClafferty, B., Meenakshi, J. V. & Pfeiffer, W. H. Biofortification: A new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* **32**(1), 31–40 (2011).
7. Bailey, L. R., West, K. P. Jr. & Robert, B. E. The epidemiology of global micronutrient deficiencies. *Ann. Nutr. Metab.* **66**(2), 22–33 (2015).
8. Miller, D. D. & Welch, R. M. Food system strategies for preventing micronutrient malnutrition. *Food Policy* **42**, 115–128 (2013).
9. West, K. P. Extent of vitamin A deficiency among preschool children and women of reproductive age. *J. Nutr.* **132**, 2857S–2866S (2002).
10. Rice, A.L., West, K.P Jr.& Black, R.E. Vitamin A deficiency. In: M. Ezzati, A.D. Lopez, A. Rodgers, and C.J.L. Murray, editors, Comparative quantification of health risks—Global and regional burden of disease attributed to selected major risk factors. Vol. 1. The World Health Organization, Geneva, Switzerland (2004).
11. World Health Organization. Global prevalence of vitamin A deficiency in populations at risk 1995–2005: WHO global database on vitamin A deficiency. *WHO Iris* **55**. (2009).
12. Black, R. E. *et al.* Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* **371**, 243–260 (2008).
13. Vignesh, M. *et al.* Genetic variability for kernel beta -carotene and utilization of crt1b1 3' te gene for biofortification in maize *Zea mays L.* *Indian Journal of Genetics and Plant Breeding* **72**(2), 189–194 (2012).
14. Pixley, K., Rojas, N.P., Babu, R., Mutale, R., Surles, R. & Simpugwe, E. *Preventive Nutrition*. S. A. Tanumihardjo. Ed. 5th Ed. Springer International Publishing Switzerland. <http://www.springer.com/series/7659> (2015).
15. Safawo, T. *et al.* Exploitation of natural variability in maize for β -carotene content using HPLC and gene specific markers. *Electr. J. Plant Breed.* **1**(4), 548–555 (2010).
16. Menkir, A., Maziya-Dixon, B., Mengesha, W., Rocheford, T. & Alamu, E. O. Accruing genetic gain in pro-vitamin A enrichment from harnessing diverse maize germplasm. *Euphytica* **213**(5), 1–12 (2017).
17. Warburton, M. L. Identification and use of rare natural genetic sequence variation to improve levels of β-carotene in maize. *Agric. Environ. Biotechnol.* **23**, 38 (2010).
18. Zhou, Y. *et al.* ZmcrtRB3 encodes a carotenoid hydroxylase that affects the accumulation of -carotene in maize kernel. *J. Integr. Plant Biol.* **54**(4), 260–269 (2012).
19. HarvestPlus Biofortification Progress Briefs. Retrieved from <http://www.harvestplus.org> (2019).
20. Harjes, C. E. *et al.* Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **319**(5861), 330–333 (2008).
21. Yan, J. *et al.* Rare genetic variation at *Zea mays L.* crtRB1 increases β -carotene in maize grain. *Nat. Genet.* **42**(4), 322–327 (2010).
22. Pixley, K.V. Hybrid and open-pollinated varieties in modern agriculture. In: K.R. Lamkey and M. Lee, editors, Plant breeding: The Arnel R. Hallauer International Symposium. Blackwell Publ. Professional, Ames, Iowa. p. 234–250 (2006).
23. Dhlwayo, T., Palacios-rojas, N., Crossa, J. & Pixley, K. V. Effects of S1 recurrent selection for provitamin A carotenoid content for three open-pollinated maize cultivars. *Crop Sci.* **54**, 2449–2460 (2015).
24. Ceballos, H. *et al.* Rapid cycling recurrent selection for increased carotenoids content in cassava roots. *Crop Sci.* **53**(6), 2342–2351 (2013).
25. Howe, J. A. & Tanumihardjo, S. A. Carotenoid-biofortified maize maintains adequate vitamin a status in mongolian gerbils. *J. Nutr.* **136**, 2562–2567 (2006).
26. Azmach, G., Gedil, M., Menkir, A. & Spillane, C. Marker-trait association analysis of functional gene markers for provitamin A levels across diverse tropical yellow maize inbred lines. *BMC Plant Biol.* **13**, 227 (2013).
27. SAS Institute *Statistical Analysis Software .SAS User's Guide* 9.3. Cary, NC: SAS Institute Inc. (2009).

28. Abdulmalik, R. O. *et al.* Genetic gains in grain yield of a maize population improved through marker assisted recurrent selection under stress and non-stress conditions in West Africa. *Front. Plant Sci.* **8**(841), 1–11 (2017).
29. Peakall, R. & Smouse, P. E. GenALEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537–2539 (2012).
30. Nei, M. *Molecular Evolutionary Genetics* (Columbia University Press, 1987).
31. Menkir, A., Gedil, M., Tanumihardjo, S., Adepoju, A. & Bossey, B. Carotenoid accumulation and agronomic performance of maize hybrids involving parental combinations from different marker-based groups. *Food Chem.* **148**, 131–137 (2014).
32. Babu, R., Natalia, P. & Prasanna, B.M. *Biofortified maize—a genetic avenue for nutritional security, in translational genomics for crop breeding: abiotic stress, yield and quality.* <https://doi.org/10.1002/9781118728482.ch10> (2013a).
33. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* 4th edn. (Pearson Prentice Hall, 1996).
34. Bankole, F. *et al.* Genetic gains in yield and yield related traits under drought stress and favorable environments in a maize population improved using marker assisted recurrent selection. *Front. Plant Sci.* **8**(808), 1–10 (2017).
35. Vogel, K.E. Genetic characterization of maize populations selected for grain methionine content, MSc. Thesis. *Iowa State University, Ames, Iowa, USA* (2010).
36. Solomon, K. F., Martin, I. & Zeppa, A. Temporal genetic structure patterns in tropical maize populations under reciprocal recurrent selection. *Euphytica* **176**(2), 239–249 (2010).
37. Huang, S. H., Teng, W. T., Wang, Y. J. & Dai, J. R. Genetic diversity analysis of maize recurrent selection populations by SSR marker. *Acta Genet. Sin.* **31**, 73–80 (2004).
38. Ishii, T. & Yonezawa, K. Optimization of the marker-based procedures for pyramiding genes from multiple donorlines: II. Strategies for selecting the objective homozygous plant. *Crop Sci* **47**, 1878–1886 (2007).
39. Kolawole, A. O. *et al.* Genetic divergence in two tropical maize composites after four cycles of reciprocal recurrent selection. *Plant Breeding* **136**, 41–49 (2017).
40. Johnson, R. Marker-assisted selection. *Plant Breed. Rev.* **24**, 293–310 (2004).

## Acknowledgements

This research was conducted at the International Institute of Tropical Agriculture. The authors express their appreciation to all staff that participated during planting, data recording, harvesting and management of the trial at all the locations and the staff who conducted the carotenoid analyses.

## Author contributions

Conceptualization, D.K and A.M.; Writing—original draft, D.K, W.M, A.M, A.A; Writing—review & editing, D.K, A.M, W.M, A.A, M.G and A.O.

## Funding

The research was supported by Bill and Melinda Gates foundation through HarvestPlus Challenge Program Grant Number: OPP1019962 (HarvestPlus 3). The authors have no relevant financial or non-financial interests to disclose.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to W.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021