Molecular characterization of 5' UTR of the *lycopene epsilon cyclase* (*lcyE*) gene among exotic and indigenous inbreds for its utilization in maize biofortification

Rajkumar Uttamrao Zunjare¹ · Rashmi Chhabra¹ · Firoz Hossain¹ · Aanchal Baveja¹ · Vignesh Muthusamy¹ · Hari Shanker Gupta¹

Received: 7 November 2017 / Accepted: 5 January 2018 / Published online: 13 January 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Maize grains are the important source of food and energy, but possess very low proA (< 2.5 µg/g) compared to target level of 15 µg/g set by HarvestPlus to alleviate VAD. Favorable allele having variation in 5' untranslated region (UTR) of *lycopene epsilon cyclase* (*lcyE*) gene enhances concentration of proA in maize. To identify the sequence variation in 5' UTR of *lcyE*, a set of diverse 13 inbreds of indigenous and exotic origin was characterized for allelic constitution of *lcyE*. Inbreds possessed wide variation in proA (1.62–23.12 µg/g) with a mean of 9.64 µg/g. The proA in CIMMYT-HarvestPlus genotypes having favorable allele of *lcyE* was very high (22.28 µg/g), whereas the Indian inbreds with the same allele possessed very low proA (2.48 µg/g). Eight genotypes viz., HKI161, HKI163, HKI161-PV, HKI163-PV, HKI193-1-PV, HKI193-2-PV, HP704-22 and HP704-23 revealed the presence of favorable allele, while VQL1, DMRIL47, MGU-PV-123/C6, HKI193-1 and HKI193-2 showed the presence of unfavorable allele of *lcyE* gene. Sequence comparison of favorable allele of Indian (HKI161 and HKI163) and exotic genotypes (HP704-22 and HP704-23) revealed seven *SNPs* having three transitions (*SNP1* and *SNP3*: G to A, *SNP2*: C to T) and four transversions (*SNP4*: C to G, *SNP5*: T to G, *SNP6*: G to C and *SNP7*: G to T). Four *SNPs* (*SNP1*: position 446, *SNP2*: position 458, *SNP3*: position 459 and *SNP4*: position 483) discriminated the low- and high- proA lines having favorable allele of *lcyE* 5'TE. These *SNPs* hold significance in enrichment of proA in maize for marker development and their use in marker-assisted selection.

Keywords Biofortification · Transposon · *LcyE* · *SNP* · Provitamin A · Maize

Introduction

Micronutrient deficiency is predominantly caused by consumption of staple food which lacks recommended level of micronutrients for normal body functions (Gupta et al. 2015; Neeraja et al. 2017). Two billion people suffer from micronutrient deficiency, while 795 million people are undernourished (IFPRI 2016). Among various micronutrients, vitamin A deficiency (VAD) has emerged as a serious health

problem that affects 4.4 million preschool-age children and nearly 20 million pregnant women worldwide (<http://www.harvestplus.org>). About 50 per cent of such cases occur in India with severe form of VAD. In extreme cases, VAD causes night blindness, keratomalacia, irreversible blindness, measles, diarrhea and respiratory diseases (Bouis and Saltzman 2017). These problems are mainly prevalent in the Africa, Meso-America and many Asian countries. Among cereals, maize serves as a major source of food and energy to millions of people in the developing and developed countries (Yadav et al. 2015). Traditional maize possesses < 2.5 µg/g of provitamin A (proA) compared to target level of 15 µg/g set by HarvestPlus (Pixley et al. 2013). Biofortification of maize for enrichment of proA therefore assumes great significance, as it is a viable approach for cost-effective and sustainable solution to alleviate VAD, compared to other approaches viz., food fortification, medical supplementation and dietary diversification (Tanumihardjo et al. 2007).

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-018-1100-y>) contains supplementary material, which is available to authorized users.

✉ Firoz Hossain
fh_gpb@yahoo.com

¹ ICAR-Indian Agricultural Research Institute (IARI),
New Delhi, India

In carotenoid biosynthesis pathway, *lycopene epsilon cyclase* (*lcyE*) and β -carotene hydroxylase (*crtRB1*) genes regulate the accumulation of proA carotenoids in maize kernel (Harjes et al. 2008; Bai et al. 2009; Yan et al. 2010; Babu et al. 2013). The gene *lcyE* located on chromosome 8 (bin 8.05; 3640 bp) consists of 10 exons, and is responsible for alteration of lycopene flux down to α -carotene versus β -carotene branches. Allele mining strategy employed by Harjes et al. (2008) revealed four natural *lcyE* polymorphisms explaining 58% of the variation and a threefold difference in proA carotenoids. Among the four identified polymorphisms viz., *lcyE* 5'TE [transposable element; in 5'-untranslated region (UTR)] alone causes significant increase in proA content over other variations viz., *lcyE* SNP216 (in exon 1), *lcyE* SNP2238 (in intron 4) and *lcyE* 3'InDel (in 3' UTR) (Harjes et al. 2008). The locus *crtRB1* identified by Yan et al. (2010) on chromosome 10 (bin: 10.05) with 5455 bp size had three polymorphisms viz., 5'TE (in 5' UTR), InDel4 (in coding region) and 3'TE (in exon 6 and 3' UTR), that were significantly associated with accumulation of high proA in maize kernel. The gene *crtRB1* has been utilized in the development of proA-rich maize cultivars worldwide (Muthusamy et al. 2014; Liu et al. 2015), while Dhlwayo et al. (2014) has developed maize with > 20 $\mu\text{g/g}$ proA through introgression of *lcyE*. Moreover, combination of *crtRB1* and *lcyE* accumulates more proA as compared to *crtRB1* alone, and therefore introgression of both the alleles is desirable to accumulate higher proA in maize kernels (Babu et al. 2013; Zunjare et al. 2017a).

DNA mutations like single nucleotide polymorphisms (SNPs) and insertion/deletions (*InDels*) are associated with the phenotypic diversity of crop plants (Vignesh et al. 2013; Saini et al. 2016; Udoh et al. 2017). The function of gene is altered when SNPs or *InDels* present in the coding region of gene causes change in the amino acid sequence and such nucleotide changes are called “functional nucleotide polymorphisms (FNPs)”. However, SNPs/*InDels* present in the 5'- and 3'-UTRs also affect gene regulation mainly at translational level through interaction of RNA-binding proteins (Makarevitch et al. 2015). Insertion and deletions of transposons in the genes create structural changes in DNA sequence and therefore affects gene function (Ito et al. 2013). In both *crtRB1* and *lcyE*, transposon had played a major role for creation of different alleles with diverse function (Harjes et al. 2008; Yan et al. 2010). Sequencing of alleles of candidate genes and subsequent discovery of FNPs would be useful in the development of functional markers for precise and accurate genotyping in marker-assisted breeding program (Saini et al. 2016; Udoh et al. 2017). Through sequence analysis of 3' UTR of *crtRB1* gene, Vignesh et al. (2013) discovered *SNP4*, *SNP13*, *InDel6* and *InDel7* as functional variations that clearly discriminated the high and the low β -carotene

genotypes belonging to Indian and exotic origin. These variations might be responsible for low β -carotene in Indian genotypes despite having favorable allele of *crtRB1*. However, no such sequence analyses have been undertaken for *lcyE*. The present study was, therefore, aimed to (i) characterize the sequence variation(s) in the 5' UTR of *lcyE* among Indian and exotic genotypes and (ii) identify the sequence variation(s) in the favorable allele of *lcyE* gene differentiating the high and low proA genotypes.

Materials and methods

Maize inbred panel

A set of 13 diverse maize inbreds originating from national institutes and CIMMYT-HarvestPlus program was selected for the present study (Table 1). Four inbreds from CCSHAU, Uchani (HKI161, HKI163, HKI193-1, HKI193-2); six from IARI, New Delhi (HKI161-PV, HKI163-PV, HKI193-1-PV, HKI193-2-PV, DMRIL47, MGU-PV-123/C6); one from VPKAS, Almora (VQL1) and two from CIMMYT-HarvestPlus program (HP704-22 and HP704-23) were selected for sequence analyses (Table 1). These inbred lines were characterized for the presence of favorable/unfavorable alleles of both *lcyE* and *crtRB1* through PCR amplification. Among the lines, HKI161-PV, HKI163-PV, HKI193-1-PV, and HKI193-2-PV were improved for proA using marker-assisted selection (MAS) (Muthusamy et al. 2014; Zunjare et al. 2017a). While, the other inbreds were identified through genotyping.

Estimation of provitamin A

Self-fertilized ears of each inbred were harvested at grain moisture ~ 12–14%. These ears were then dried for 3 days in the shade under ambient environment and then shelled using maize grain sheller. Separate packets with proper label were prepared, and placed in – 20 °C deep freezer. Proper care was taken at the time of harvesting, drying and seed bulking to avoid any loss of carotenoids due to light, heat and oxygen. The extraction of proA from inbred lines was carried out using procedure as per Zunjare et al. (2017a). Quantification of proA was carried out with a Dionex Ultimate 3000 UHPLC System (ultra high performance liquid chromatography; Thermo Scientific, Massachusetts, USA). Samples were eluted through YMC Carotenoid C30 column (5 μm , 4.6 $\times$ 250 mm; YMC) and detected with a diode array detector-3000 (RS). Significance for proA concentration among different genotypic classes was tested using 't' test.

Table 1 List of maize inbreds used in the characterization of 5' UTR of *lcyE*

S. No.	Genotypes	Pedigree	Institution
1	HKI161	Selection from CML161	CCSHAU, Karnal
2	HKI163	Selection from CML163	CCSHAU, Karnal
3	HKI193-1	Selection from CML193	CCSHAU, Karnal
4	HKI193-2	Selection from CML193	CCSHAU, Karnal
5	HKI161-PV	(HKI161 × HP467-13)-MAS-BC2P04-⊗-⊗-⊗-⊗	IARI, New Delhi
6	HKI163-PV	(HKI163 × HP704-22)-MAS-BC2P07-⊗-⊗-⊗-⊗	IARI, New Delhi
7	HKI193-1-PV	(HKI193-1 × HP704-23)-MAS-BC2P08-⊗-⊗-⊗-⊗	IARI, New Delhi
8	HKI193-2-PV	(HKI193-2 × HP704-22)-MAS-BC2P52-⊗-⊗-⊗-⊗	IARI, New Delhi
9	MGU-PV-123/C6	(MAP123 × HP465-20)-⊗-⊗-⊗-⊗-⊗	IARI, New Delhi
10	DMRIL47	(NAI116 × CM139)-⊗-⊗-⊗-⊗-⊗	IARI, New Delhi
11	VQL1	(V360 × CML189) BC2P1-bb##-b	VPKAS, Almora
12	HP704-22	(KUI carotenoid syn-FS11-1-1-B-B-B/(KU1409/DE3/KU1409)S2-18-2-B)-B-3(MAS:L4H1)-1-B-B-B	CIMMYT HarvestPlus
13	HP704-23	(KUI carotenoid syn-FS17-3-2-B-B-B/(KU1409/DE3/KU1409)S2-18-2-B)-B-1(MAS:L4H1)-1-B-B-B	CIMMYT HarvestPlus

DNA isolation and amplification of *lcyE* and *crtRB1*

Using standard CTAB procedure as per Murray and Thompson (1980), genomic DNA was isolated from 21 days old seedlings of each selected inbred. DNA was isolated from three different plants of each genotype and then pooled in equal quantity for analysis. Polymerase chain reaction (PCR) was performed using High-Fidelity Taq-DNA Polymerase (Thermo Scientific™) to amplify the unfavorable/favorable of *lcyE* using gene-based primers; forward: 5'AAG CAGGGAAGACATTCCAG3' and reverse: 5'GAGAGG GAGACGACGAGACAC3'. Favorable allele in the 5'TE region of the *lcyE* gene, i.e., allele 4 (650 bp) enhances the β-branch carotenoids by increasing the pathway flux towards β-branch, while unfavorable allele (300 bp) diverts lycopene to α-branch (Harjes et al. 2008). Inbreds were also characterized for the presence of unfavorable/favorable allele of *crtRB1* by primers; forward: 5'ACACCACATGGACAA GTTCG3', reverse1: 5'ACACTCTGGCCCATGAACAC3' and reverse2: 5'ACAGCAATACAGGGGACCAG3'. PCR conditions for both the genes were used as given in Zunjare et al. (2017b).

Sequencing and SNP discovery

PCR products of the *lcyE* among 13 inbreds were sequenced in both the directions using forward and reverse primers with two replications at M/s Sequencher Tech Pvt. Ltd., Ahmedabad. The contigs of each genotype were generated using both forward and reverse sequence with BioEdit program (Hall 1999). The sequences were then aligned using BioEdit and ClustalW (<http://www.ebi.ac.uk/tools/clustalw2/index.html>).

The aligned sequences were used to analyze the presence of *SNPs/InDels* specific to genotype.

Results

Genotypic characterization for *lcyE* and *crtRB1*

The selected 13 inbreds were characterized for the allelic constitution of *lcyE* and *crtRB1*. The eight genotypes viz., HKI161, HKI163, HKI161-PV, HKI163-PV, HKI193-1-PV, HKI193-2-PV, HP704-22 and HP704-23 revealed the presence of favorable allele (650 bp), while VQL1, DMRIL47, MGU-PV-123/C6, HKI193-1 and HKI193-2 possessed unfavorable allele (300 bp) of *lcyE* gene. For *crtRB1*, eight genotypes (HKI161-PV, HKI163-PV, HKI193-1-PV, HKI193-2-PV, DMRIL47, MGU-PV-123/C6, HP704-22 and HP704-23) amplified the favorable allele (543 bp) (Fig. S1), whereas five inbreds (VQL1, HKI161, HKI163, HKI193-1 and HKI193-2) had unfavorable allele (296 bp) (Table 2). Considering both loci, six inbreds (HKI161-PV, HKI163-PV, HKI193-1-PV, HKI193-2-PV, HP704-22 and HP704-23) possessed favorable alleles of both *lcyE* (650 bp) and *crtRB1* (543 bp). On the contrary, VQL1, HKI193-1 and HKI193-2 possessed unfavorable alleles of both the genes. The two inbreds (HKI161 and HKI163) possessed favorable allele of *lcyE*, but unfavorable allele of *crtRB1*. Conversely, DMRIL47 and MGU-PV-123/C6 carried unfavorable allele of *lcyE* and favorable allele of *crtRB1* gene (Table 2).

Table 2 Genotypic and biochemical evaluation of genotypes

S. No.	Genotypes	<i>crtRB1</i>	<i>lcyE</i>	Type with proA mean	proA ($\mu\text{g/g}$)
1	HKI193-1	296	300	Class-I (1.92 $\mu\text{g/g}$)	2.30
2	HKI193-2	296	300		1.62
3	VQL1	296	300		1.85
4	HKI161	296	650	Class-II (2.48 $\mu\text{g/g}$)	2.49
5	HKI163	296	650		2.46
6	DMRIL47	543	300	Class-III (6.30 $\mu\text{g/g}$)	2.31
7	MGU-PV-123/C6	543	300		10.30
8	HKI163-PV ^a	543	650	Class-IV (16.99 $\mu\text{g/g}$)	13.27
9	HKI193-1-PV ^a	543	650		15.43
10	HKI193-2-PV ^a	543	650		17.69
11	HKI161-PV ^b	543	650		11.00
12	HP704-22	543	650		21.44
13	HP704-23	543	650		23.12
Mean					9.64
Range					1.62–23.12
SE(d)					2.22

SE(d), standard error of difference

^aMAS derived lines with *crtRB1* and *lcyE* allele introgressed from exotic source (Zunjare et al. 2017a)^bMAS derived line with *crtRB1* introgressed from exotic source (Muthusamy et al. 2014)**Table 3** Mean comparison of the different genotypic classes

<i>t</i> -statistic	Class-I	Class-II	Class-III
Class-II	2.13	–	–
Class-III	1.47	0.95	–
Class-IV	5.37**	4.15**	2.41*

**, * Significant at $P = 0.01$ and $P = 0.05$, respectively

Kernel provitamin A in inbreds

The results revealed wide range of proA content (1.62–23.12 $\mu\text{g/g}$) among 13 genotypes (Table 2). The highest proA was possessed by HP704-23, while the lowest proA was recorded in HKI193-2, with a mean 9.64 $\mu\text{g/g}$. Based on their allelic status at *crtRB1* and *lcyE* genes, 13 genotypes were classified into four classes (Class I–IV). Class-I (*crtRB1*: 296 bp, *lcyE*: 300 bp) recorded mean proA of 1.92 $\mu\text{g/g}$. Class-II (*crtRB1*: 296 bp, *lcyE*: 650 bp), class-III (*crtRB1*: 543 bp, *lcyE*: 300 bp) and class-IV (*crtRB1*: 543 bp, *lcyE*: 650 bp) had mean proA of 2.47, 6.30 and 16.99 $\mu\text{g/g}$, respectively (Table 2). The *t* test performed among the means of four classes showed Class-IV was significantly different from Class-III, -II and -I suggesting the beneficial effects of favorable allele of *crtRB1* and *lcyE* (Table 3). In the genotype where favorable allele of *crtRB1* is present, mean proA (14.32 $\mu\text{g/g}$) is higher compared to genotypes where unfavorable allele of *crtRB1* is present (2.14 $\mu\text{g/g}$).

Sequence variation in 5' UTR of *lcyE*

The *lcyE*-unfavorable allele (VQL1, DMRIL47, MGU-PV-123/C6, HKI193-1 and HKI193-2) was aligned with *lcyE*-favorable allele to identify the transposon insertion in the 5' UTR of *lcyE*. Sequence alignment has revealed about ~ 380–400 bp insertion of transposon (Fig. S1). Among the inbreds of Indian origin, HKI161 and HKI163 only possessed favorable allele of *lcyE*, while two exotic inbreds (HP704-22 and HP704-23) had the *lcyE*-favorable allele. Other inbreds viz., HKI163-PV, HKI193-1-PV and HKI193-2-PV were developed through MAS in India (Zunjare et al. 2017a), and therefore possessed favorable alleles from exotic donor lines. On the contrary, HKI161-PV, HKI161 and HKI163 were naturally found inbreds with favorable allele of *lcyE* (Muthusamy et al. 2014, 2015). Importantly, HKI161 and HKI163 had very low proA viz., 2.49 and 2.46 $\mu\text{g/g}$, respectively, while HP704-22 and HP704-23 possessed very high proA measuring 21.44 and 23.12 $\mu\text{g/g}$, respectively. The sequence comparison of these lines revealed a total of 41 SNPs and five *InDels*. Of these variations, seven SNPs clearly differentiated the *lcyE*-favorable allele from Indian and exotic genotypes. The seven SNPs were named as *SNP1* (G to A) at position 447, *SNP2* (C to T) at position 459, *SNP3* (G to A) at position 460, *SNP4* (C to G) at position 484, *SNP5* (T to G) at position 557, *SNP6* (G to C) at position 566 and *SNP7* (G to T) at position 607 (Fig. S2). Among the seven point mutations, three SNPs (*SNP1*, *SNP2*, and *SNP3*) were of transition types, while four SNPs (*SNP4*, *SNP5*, *SNP6*, and *SNP7*) were of transversion type.

The biochemical evaluation of 13 inbreds clearly showed that favorable allele of *crtRB1* has major effect on proA enhancement over unfavorable allele. Therefore, to elucidate the effect of sequence variation in *lcyE* allele for proA, genotypes possessing favorable allele of *crtRB1* were used. Six genotypes (HKI161-PV, HKI163-PV, HKI193-1-PV, HKI193-2-PV, HP704-22 and HP704-23) possessing favorable alleles of both *crtRB1* and *lcyE* were selected for identifying possible phenotypic effect of sequence variations in 5' UTR of *lcyE*. These six genotypes could be classified into two groups depending upon their proA content. Group-A (mean proA: 12.13 µg/g)

had two inbreds (HKI161-PV, and HKI163-PV), while group-B comprised of four inbreds viz., HKI193-1-PV, HKI193-2-PV, HP704-22 and HP704-23 which had high proA (19.42 µg/g). The electropherograms of the *SNP1*, *SNP2*, *SNP3* and *SNP4* is presented as Fig. 1. A total of 26 *SNPs* and one *InDel* were observed among selected genotypes. *SNP1* (G to A) at position 446, *SNP2* (C to T) at position 458, *SNP3* (G to A) at position 459 and *SNP4* (C to G) at position 483 clearly discriminated group-A (low proA) and group-B (high proA). *SNP1*, *SNP2* and *SNP3* were of transition type whereas; *SNP4* was of transversion type (Table 4) (Figs. 1, S3).

Fig. 1 Electropherogram of *lcyE* 5' UTR among genotypes possessing favorable allele

Table 4 SNPs identified in the favorable allele of 5' UTR of *lcyE* among six genotypes

S. No.	SNP	Change	Position (bp)	Type of mutation
1	SNP1	G to A	446	Transition
2	SNP2	C to T	458	Transition
3	SNP3	G to A	459	Transition
4	SNP4	C to G	483	Transversion

Discussion

Insertion of transposon in *lcyE*

Transposons first called as 'controlling elements' by McClintock (1956), may result in mutant alleles by changing the reading frame and splicing pattern thereby affecting gene expression (Makarevitch et al. 2015). Firstly, Harjes et al. (2008) have reported the insertion of transposon in untranslated regions of *lcyE*. In the present study, we have further characterized the nature of insertion at 5' UTR of *lcyE*. Sequence comparison of *lcyE*-favorable and -unfavorable allele showed ~ 380–400 bp insertion of transposon in the 5' UTR of *lcyE*. Transposon insertions have been reported in many genes of maize and led to the creation of different haplotypes of *tb1* (*teosinte branched1*), and *vgt1* (*vegetative to generative transition1*) genes affecting the adaptation and stress response of plants (Studer et al. 2011; Castelletti et al. 2014). Expression profiling experiment revealed that lines with insertion of transposon near the start site of *lcyE* gene had much lower level of function (Harjes et al. 2008). The point of transposon insertion in the present study was doCT transcription factor binding site (CACGGCTC), which is a member of NAC (NAM, ATAF1/2 and CUC2) domain protein family involved in a variety of abiotic stresses mediated by synthesis of lutein (Shao et al. 2015). This disruption could be responsible for influencing ratio of α - to β -carotenoids, thereby suggesting that transposon insertion favors low lutein profile in maize endosperm.

Sequence variation in *lcyE* of indigenous and exotic inbreds

The importance of *lcyE* gene for maize biofortification was demonstrated by Harjes et al. (2008) through investigation on 288 diverse inbreds. Of the four polymorphisms, the favorable allele (650 bp) of *lcyE* 5'TE causes up to 30% reduction in the ratio of α - to β -branch carotenoids. Subsequently, Vallabhaneni and Wurtzel (2009) and Yan et al. (2010) suggested further enrichment of proA by combining *crtRB1* and *lcyE*-favorable alleles. However, the frequency of favorable allele of *lcyE* is very low in natural maize germplasm. Allele screening among 385 inbreds by Muthusamy

et al. (2015) revealed that only two indigenous and 11 CIM-MYT inbreds possessed *lcyE*-favorable allele. In present study, seven SNPs differentiated the favorable allele of *lcyE* from indigenous and exotic origin. Burt et al. (2011) showed one SNP per 36 nucleotides by characterizing *lcyE* sequence of 34 high carotenoid lines. We observed one SNP per ~ 15 nucleotides in 5' UTR of *lcyE* gene. The occurrence of SNP frequency of 150 bp was observed by Udoh et al. (2017) in *lcyE* gene sequence of 167 lines of cassava.

Effect of *crtRB1* versus *lcyE* on proA

The evaluation of the diverse inbreds recorded wide variation for proA, and functional markers revealed favorable and unfavorable alleles of *crtRB1* and *lcyE* genes (Table 2). It has been observed that the higher proA is due to favorable allele of *crtRB1* irrespective of *lcyE*. On the contrary, *crtRB1* unfavorable allele showed lower level of proA. This indicated *crtRB1* possessed pronounced effect on proA enrichment. Similarly, major effect of *crtRB1* for proA improvement was well recorded by several researchers (Vallabhaneni and Wurtzel 2009; Yan et al. 2010; Vignesh et al. 2012; Babu et al. 2013; Zunjare et al. 2017a). Subsequently, *crtRB1*-based introgression programs were carried out to develop biofortified maize varieties enriched with proA (Muthusamy et al. 2014; Liu et al. 2015). In the present study, variation for proA observed in genotypes possessing favorable allele of *crtRB1* could be due to other genes involved in carotenoid biosynthesis pathway (Zhou et al. 2012; Azmach et al. 2013; Kandianis et al. 2013; Muthusamy et al. 2014). Among various genes, *lcyE* has been implicated as an important gene that regulates proA accumulation in maize (Babu et al. 2013). In the present study, *lcyE*-favorable allelic class (650 bp) had high mean proA (13.36 $\mu\text{g/g}$) than *lcyE*-unfavorable allelic class (3.67 $\mu\text{g/g}$). Considering both loci, favorable genotypes had mean proA of 16.99 $\mu\text{g/g}$, which is much higher than unfavorable genotypes with a mean proA of 1.92 $\mu\text{g/g}$. Zunjare et al. (2017a) also suggested that the effect of *crtRB1* had major impact and *lcyE* could cause 24.6% increase in proA, suggesting that the pyramiding of *crtRB1* with *lcyE* would be promising for further proA enhancement. Moreover, recent estimates of 3.2:1 (Muzhingi et al. 2011), and 6.5:1 (Li et al. 2010) indicate that bioconversion of β -carotene from biofortified maize is considerably more efficient than the ratio of 12:1 proposed by the United States Institute of Medicine (2001).

Sequence variation at favorable allele of *lcyE* in high- and low- proA inbreds

In this study, we analyzed nucleotide differences at 5' UTR of *lcyE* gene in two groups that differed significantly for proA content. Genotypes comprising of group-A (~ 12 $\mu\text{g/g}$)

and -B (~ 20 µg/g) had favorable alleles of *crtRBI* and *lcyE* genes with varied levels of proA. Therefore, sequence difference could be correlated with variations at 5' UTR of *lcyE*. We have identified four single nucleotide differences (*SNP1*, *SNP2*, *SNP3* and *SNP4*) that discriminated the low and high proA inbreds. Similarly, *SNP4*, *SNP13*, *InDel6* and *InDel7* identified in the 3' UTR clearly differentiated the high and the low β-carotene genotypes in *crtRBI* gene (Vignesh et al. 2013). Translational efficiency largely depends on 5' UTR of genes due to the presence of control elements such as the m7G cap, secondary structure, length and nucleotide sequence nearby AUG start codon (Jackson et al. 2010). Also, specific sequence motifs located at 5' UTR are involved in controlling translational efficiency in either positive or negative manner (Kim et al. 2014). Here, the identified variations in the study could be affecting any of the controlling features present in the 5' UTR of the *lcyE* gene, thereby affecting its phenotypic expression. Another possibility could be the change in upstream open reading frames in the 5' UTR that can modulate the translational efficiency (Kim et al. 2014). Vignesh et al. (2013) also mentioned the role of *SNPs* and *InDels* at 3' UTR of *crtRBI* for proA accumulation through regulating the translational efficiency of the mRNA. However, the *SNPs* identified here require validation in a large set of diverse inbreds differing for proA.

Utilization in future breeding program

So far, selection of *lcyE*-favorable allele in the breeding program has been carried out using gene-based marker that produces 650 bp amplicon as the favorable allele (Harjes et al. 2008; Babu et al. 2013). However, the amplicon does not detect the presence of favorable or unfavorable *SNPs* PCR product that may lead to the selection of unfavorable *SNPs* thus *lcyE* allele. This *SNP*-based information, thus, can be used to develop co-dominant markers to differentiate the most effective allele of *lcyE* from other versions. *SNPs* were identified in cassava by sequencing of *lcyE* gene, and these *SNPs* were utilized to develop CAPS marker for utilization in MAS program focusing on proA improvement (Udoh et al. 2017). Similarly for quality improvement in *Brassica juncea*, two CAPS marker were developed from *SNPs* identified in *fatty acid elongase 1 (fae1)* gene that differentiated low and high erucic acid genotypes (Saini et al. 2016). Estimation of proA using HPLC in large segregating progenies involves high cost (Babu et al. 2013). On the other hand, PCR-based co-dominant marker system efficiently identifies the heterozygotes from homozygotes in the segregating populations, and desirable individual plants can be backcrossed without phenotyping (Collard et al. 2005). Considering this, MAS for *lcyE* especially in the genetic background of *crtRBI* has been advocated as the most viable

approach for proA enrichment in maize (Harjes et al. 2008; Babu et al. 2013). Using the co-dominant marker system, favorable allele of *crtRBI* has been effectively introgressed into elite inbreds to develop proA rich genotypes (Muthusamy et al. 2014; Liu et al. 2015). Thus, development of effective marker system for *SNPs* identified here can lead to the identification of the most favorable allele of *lcyE* for their effective utilization in the maize biofortification program. To best of our knowledge, this is the first ever report on identification of nucleotide variations in the favorable allele of *lcyE* in maize.

Conclusion

The nucleotide differences observed in the 5' UTR of *lcyE* among Indian and CIMMYT-HarvestPlus genotypes are linked with the proA accumulation through possible regulation of translational efficiency of mRNA. Despite the fact that HKI161-PV, HKI163-PV, had the favorable allele of *lcyE*, the accumulation of proA was low due to the presence of nucleotide polymorphisms. The four *SNPs* identified in the present study will be useful in designing allele-specific co-dominant primers for maize biofortification program.

Acknowledgements First author is thankful to University Grants Commission (UGC) for awarding the Rajiv Gandhi National Fellowship for his doctoral program. The financial help received from ICAR-Indian Agricultural Research Institute, New Delhi is gratefully acknowledged. The authors are thankful to Drs. Natalia Palacios-Rojas, Raman Babu, and Boddupalli M. Prasanna, CIMMYT for providing CIMMYT-HarvestPlus bred maize inbreds. We also thank the maize breeders of CCSHAU, Uchani and VPKAS, Almora for sharing their inbreds.

Authors' contribution Conduct of the experiment: RUZ; genotyping and sequence analysis: RUZ and RC; phenotyping for kernel carotenoids: AB and VM; drafting the manuscript: FH, VM and RUZ; design of experiment: FH and HSG

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

- Azmach G, Gedil M, Menkir A, Spillane C (2013) Marker-trait association analysis of functional gene markers for provitamin A levels across diverse tropical yellow maize inbred lines. *BMC Plant Biol* 13:227. <https://doi.org/10.1186/1471-2229-13-227>
- Babu R, Rojas NP, Gao S et al (2013) Validation of the effects of molecular marker polymorphisms in *LcyE* and *CrtRBI* on provitamin A concentrations for 26 tropical maize populations. *Theor Appl Genet* 126:389–399. <https://doi.org/10.1007/s00122-012-1987-3>
- Bai L, Kim EH, Dellapenna D, Brutnell TP (2009) Novel *lycopen epsilon cyclase* activities in maize revealed through perturbation

- of carotenoid biosynthesis. *Plant J* 59:588–599. <https://doi.org/10.1111/j.1365-3113X.2009.03899.x>
- Bouis HE, Saltzman A (2017) Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. *Glob Food Sec* 12:49–58. <https://doi.org/10.1016/j.gfs.2017.01.009>
- Burt AJ, Grainger CM, Smid MP et al (2011) Allele mining of exotic maize germplasm to enhance macular carotenoids. *Crop Sci* 51:991–1004. <https://doi.org/10.2135/cropsci2010.06.0335>
- Castelletti S, Tuberos R, Pindo M, Salvi S (2014) A MITE transposon insertion is associated with differential methylation at the maize flowering time QTL *Vgt1*. *G3 (Bethesda)* 4:805–812. <http://doi.org/10.1534/g3.114.010686>
- Collard BCY, Jahufer MZZ, Brouwer JB et al (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica* 142:169–196. <https://doi.org/10.1007/s10681-005-1681-5>
- Dhliwayo T, Palacios-Rojas N, Crossa J et al (2014) Effects of S1 recurrent selection for provitamin A carotenoid content for three open-pollinated maize cultivars. *Crop Sci* 54:2449–2460. <http://doi.org/10.2135/cropsci2013.11.0764>
- Gupta HS, Hossain F, Muthusamy V (2015) Biofortification of maize: an Indian perspective. *Indian J Genet Plant Breed* 75:1–22. <https://doi.org/10.5958/0975-6906.2015.00001.2>
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Harjes CE, Rocheford TR, Bai L et al (2008) Natural genetic variation in *lycopene epsilon cyclase* tapped for maize biofortification. *Science* 319:330–333. <https://doi.org/10.1126/science.1150255>
- Institute of Medicine (2001) Dietary reference intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. Panel on Micronutrients. National Academies Press, Washington, DC
- International Food Policy Research Institute (2016) Global hunger index. <http://www.ifpri.org/publication/2016-global-hunger-index-getting-zero-hunger>
- Ito H, Yoshida T, Tsukahara S, Kawabe A (2013) Evolution of the ONSEN retrotransposon family activated upon heat stress in Brassicaceae. *Gene* 518:256–261. <https://doi.org/10.1016/j.gene.2013.01.034>
- Jackson R, Christopher UT, Hellen TVP (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11:113–127. <https://doi.org/10.1038/nrm2838>
- Kandianis CB, Stevens R, Liu W et al (2013) Genetic architecture controlling variation in grain carotenoid composition and concentrations in two maize populations. *Theor Appl Genet* 126:2879–2895. <https://doi.org/10.1007/s00122-013-2179-5>
- Kim Y, Lee G, Jeon E et al (2014) The immediate upstream region of the 5' UTR from the AUG start codon has a pronounced effect on the translational efficiency in *Arabidopsis thaliana*. *Nucleic Acids Res* 42:485–498. <https://doi.org/10.1093/nar/gkt864>
- Li S, Nugroho A, Rocheford T et al (2010) Vitamin A equivalence of the β -carotene in β -carotene-biofortified maize porridge consumed by women. *Am J Clin Nutr* 92:1105–1112. <https://doi.org/10.3945/ajcn.2010.29802>
- Liu L, Jeffers D, Zhang Y et al (2015) Introgression of the *crtRB1* gene into quality protein maize inbred lines using molecular markers. *Mol Breed* 35:154. <https://doi.org/10.1007/s11032-015-0349-7>
- Makarevitch I, Waters AJ, West PT et al (2015) Transposable elements contribute to activation of maize genes in response to abiotic stress. *PLoS Genet* 11:e1004915. <https://doi.org/10.1371/journal.pgen.1004915>
- McClintock B (1956) Controlling elements and the gene. *Cold Spring Harb Symp Quant Biol* 21:197–216
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* 8:4321–4325
- Muthusamy V, Hossain F, Thirunavukkarasu N et al (2014) Development of β -carotene rich maize hybrids through marker-assisted introgression of *β -carotene hydroxylase* allele. *PLoS ONE* 9:1–22. <https://doi.org/10.1371/journal.pone.0113583>
- Muthusamy V, Hossain F, Thirunavukkarasu N et al (2015) Allelic variations for *lycopene- ϵ -cyclase* and *β -carotene hydroxylase* genes in maize inbreds and their utilization in β -carotene enrichment programme. *Cogent Food Agric* 1:1033141. <https://doi.org/10.1080/23311932.2015.1033141>
- Muzhingi T, Gadaga TH, Siwela AH et al (2011) Yellow maize with high β -carotene is an effective source of vitamin A in healthy Zimbabwean men. *Am J Clin Nutr* 94:510–519. <https://doi.org/10.3945/ajcn.110.006486>
- Neeraja CN, Ravindra BV, Ram S et al (2017) Biofortification in cereals-progress and prospects. *Curr Sci* 113(6):1050–1057
- Pixley K, Palacios-Rojas N, Babu R, et al (2013) Carotenoids and human health, chapter: biofortification of maize with provitamin A carotenoids. In: Tanumihardjo S (eds). Springer Science + Business Media, New York, pp 271–292. https://doi.org/10.1007/978-1-62703-203-2_17
- Saini N, Singh N, Kumar A et al (2016) Development and validation of functional CAPS markers for the FAE genes in *Brassica juncea* and their use in marker-assisted selection. *Breed Sci* 66:831–837. <https://doi.org/10.1270/jsbbs.16132>
- Shao H, Wang H, Tang X (2015) NAC transcription factors in plant multiple abiotic stress responses: progress and prospects. *Front Plant Sci* 6:1–8. <https://doi.org/10.3389/fpls.2015.00902>
- Studer A, Zhao Q, Ross-Ibarra J et al (2011) A transposon insertion was the causative mutation in the maize domestication gene *tb1*. *Nat Genet* 43:1160–1163
- Tanumihardjo SA, Anderson C, Kaufer-Horwitz M et al (2007) Poverty, obesity and malnutrition: an international perspective recognizing the paradox. *J Am Diet Assoc* 107:1966–1972
- Udoh L, Gedil M, Parkes EY et al (2017) Candidate gene sequencing and validation of SNP markers linked to carotenoid content in cassava (*Manihot esculenta* Crantz). *Mol Breed* 37:123. <https://doi.org/10.1007/s11032-017-0718-5>
- Vallabhaneni R, Wurtzel ET (2009) Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. *Plant Physiol* 150:562–572. <https://doi.org/10.1104/pp.109.137042>
- Vignesh M, Hossain F, Nepolean T et al (2012) Genetic variability for kernel β -carotene and utilization of *crtRB1* 3' TE gene for biofortification in maize (*Zea mays* L.). *Indian J Genet Plant Breed* 72:189–194
- Vignesh M, Nepolean T, Hossain F et al (2013) Sequence variation in 3' UTR region of *crtRB1* gene and its effect on β -carotene accumulation in maize kernel. *J Plant Biochem Biotechnol* 22:401–408. <https://doi.org/10.1007/s13562-012-0168-4>
- Yadav OP, Hossain F, Karjagi CG et al (2015) Genetic improvement of maize in India: retrospect and prospects. *Agric Res* 4(4):325–338. <https://doi.org/10.1007/s40003-015-0180-8>
- Yan J, Kandianis CB, Harjes CE et al (2010) Rare genetic variation at *Zea mays crtRB1* increases β -carotene in maize grain. *Nat Genet* 42:322–327. <https://doi.org/10.1038/ng.551>
- Zhou Y, Han Y, Li Z et al (2012) *ZmcrRB3* encodes a carotenoid hydroxylase that affects the accumulation of α -carotene in maize kernel. *J Integr Plant Biol* 54:260–269. <https://doi.org/10.1111/j.1744-7909.2012.01106.x>

Zunjare RU, Hossain F, Muthusamy V et al (2017a) Influence of rare alleles of *β -carotene hydroxylase* and *lycopene epsilon cyclase* genes on accumulation of provitamin A carotenoids in maize kernels. Plant Breed 136:872–880. <https://doi.org/10.1111/pbr.12548>

Zunjare RU, Chhabra R, Hossain F et al (2017b) Development and validation of multiplex-PCR assay for simultaneous detection of rare alleles of *crtRB1* and *lcyE* governing higher accumulation of provitamin A in maize. J Plant Biochem Biotechnol. <https://doi.org/10.1007/s13562-017-0432-8>