



Variation in anthocyanin pigmentation by *R1-navajo* gene, development and validation of breeder-friendly markers specific to *C1-Inhibitor* locus for in-vivo haploid production in maize

Nisrita Gain^{1,2} · Rashmi Chhabra¹ · Shivani Chandra² · Rajkumar U. Zunjare¹ · Suman Dutta¹ · Gulab Chand¹ · Konsam Sarika³ · Elangbam L. Devi⁴ · Aruna Kumar² · Jayanthi Madhavan¹ · Vignesh Muthusamy¹ · Firoz Hossain¹

Received: 20 September 2022 / Accepted: 13 December 2022 / Published online: 24 December 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Background In-vivo maternal haploids serve as the basis of doubled haploid (DH) breeding in maize. *R1-navajo* (*R1-nj*) gene governing anthocyanin colouration in the endosperm and embryo is widely used to identify haploid seeds. However, the expression of *R1-nj* depends on genetic-background of source-germplasm used for deriving DH-lines. Further, presence of *C1-Inhibitor* (*C1-I*) gene suppresses the expression of *R1-nj*, thus makes the selection of haploids difficult.

Methods In the present study, 178 subtropically-adapted maize inbreds were crossed with two *R1-nj* donors ‘that do not have haploid induction genes’. Of these, 76.4% inbreds developed purple colour in endosperm, while 23.6% did not show any colouration. In case of scutellum, 62.9% inbreds possessed colour and 37.1% were colourless. The anthocyanin intensity varied greatly, with 19.66% and 42.98% inbreds displayed the least intensity, while 16.85% and 0.84% inbreds showed the highest intensity in endosperm and scutellum, respectively. Two *C1-I* specific breeder-friendly markers (MGU-CI-InDel8 and MGU-C1-SNP1) covering (i) 8 bp InDel and (ii) A to G SNP, respectively, were developed. MGU-CI-InDel8 and MGU-C1-SNP1 markers predicted presence of *C1-I* allele with 92.9% and 84.7% effectiveness, respectively. However, when both markers were considered together, they provided 100% effectiveness.

Conclusions These markers of *C1-I* gene would help in saving valuable resources and time during haploid induction in maize. The information generated here assume great significance in DH breeding of maize.

Keywords Doubled haploid · Molecular markers · Endosperm · Color inhibitor

Introduction

Development of maize hybrids is a lengthy process as inbreds are developed through continuous selfing for ~8–9 generations [1, 2]. Doubled haploid (DH) technology has emerged as a powerful tool for accelerating the breeding

cycle to obtain inbreds in 2–3 generations [3, 4]. In maize, maternal haploids are developed by crossing the source population (as female) with a haploid-inducer (HI) line (as male) possessing *matrilineal* (*mtl*) [5] and *domain membrane protein* (*dmp*) genes [6]. Anthocyanin colour marker facilitates easy and quick identification of haploid seeds. Seeds with purple coloured endosperm and colourless scutellum are identified as haploid seeds [7]. *R1-navajo* (*R1-nj*) gene present on chromosome-10 governing anthocyanin pigmentation on the maize kernels, has been widely used as colour marker for haploid identification [8]. *R1-nj* expression requires the presence of functional *A1*, *A2*, *Bz1*, *Bz2*, *C1*, *C2* and *Pr1* genes involved in anthocyanin biosynthesis pathway [9]. The phenotypic expression of *R1-nj* causes purple colour in the aleurone layer on the crown region of endosperm and scutellum of the embryo [10]. However, purple colour conditioned by *R1-nj* is poorly expressed or completely suppressed in some germplasm, thereby making the visible

✉ Firoz Hossain
fh_gpb@yahoo.com

¹ ICAR-Indian Agricultural Research Institute, New Delhi, India
² Amity Institute of Biotechnology, AMITY University, Noida, India
³ ICAR Research Complex for NEH Region, Manipur Centre, Lamphelpat, India
⁴ ICAR Research Complex for NEH Region, Sikkim Centre, Gangtok, India

identification of haploid kernels difficult [11]. Presence of *CI-Inhibitor (CI-I)* - a dominant mutant of *CI* gene present on chromosome-9 in the genetic background has been identified as the key factor that interferes with the expression of kernel colour [12]. The *CI-I* produces transcriptional repressor that affects the function of *RI* gene [13].

Intensity of colour in the kernels upon crossing with HI lines is an important aspect as often very light colouration may lead to selection of false positives as haploids [14]. Thus, information on expression level of colour among the locally adapted germplasm would help in designing the right strategy in the DH programme [7]. Haploid induction process demands lots of resources and management of logistics. Thus, absence of colour in the endosperm and scutellum upon haploid induction due to presence of *CI-I* leads to loss of valuable resources and time. Thus, prior screening for presence of *CI-I* in source populations will be of great significance as such germplasm can be excluded before attempting haploid induction crosses. Through sequence analysis, Paz-Ares et al. [15] identified 8 bp InDel in the last exon of *CI* and *CI-I* allele, while Chaikam et al. [11] reported the presence of the same 8 bp InDel and a novel SNP (A–G transition) in exon-3 that differentiated the *CI* from *CI-I* allele. Chaikam et al. [11] also developed KASPar-based marker assay to screen maize germplasm for the presence of *CI-I* allele and also for A to G SNP. Though KASPar-based marker assay provides high throughput analysis, it involves high cost and sophisticated equipment besides skilled manpower, thereby limiting the use of the marker system in resource-poor developing countries. However, these high throughput marker technologies involve high cost and skilled personnel, and access of the services may not be feasible in many of the remote areas, hence limiting their wide-spread applications. On the other hand, PCR-based assay is simple, less costly and does not involve any sophistication. Thus, development of breeder-friendly PCR-based markers that can effectively differentiate the *CI-I* and *CI* in the maize germplasm holds great potential.

In India, maize is an important cereal crop after rice and wheat, and produces 30.5 million tonnes of grains annually [16]. So far, traditional selfing has been the mainstay of development of inbreds in the Indian maize breeding programme. Recently, various public sector organizations in India have initiated DH programme in maize [17, 18]. It is therefore important to analyze the expression pattern of *RI-nj* anthocyanin colour, and estimate the frequency of *CI-I* allele in the locally adapted subtropical germplasm for effective utilization of HI lines in the DH programme. The present study was therefore undertaken to (i) investigate the level of expression of colour marker conditioned by *RI-nj* gene, and (ii) develop and validate the effectiveness of breeder-friendly markers specific to *CI-I* gene using a set of diverse subtropically adapted inbreds.

Materials and methods

Plant materials

A set of 178 diverse locally adapted subtropical inbreds with no purple colouration (yellow/orange or white in colour) including exotic (91 CIMMYT inbreds) and indigenous origin (87 from Indian maize breeding programme) were crossed as female with two *RI-nj*-based inbreds (MGU-R1nj-101 and MGU-R1nj-102) having strong purple colouration in the crown region of the endosperm and scutellum as male during rainy season 2020. These *RI-nj* based inbreds were developed by DH breeding programme of ICAR-Indian Agricultural Research Institute, New Delhi.

These two *RI-nj*-based inbreds (MGU-R1nj-101 and MGU-R1nj-102) did not possess *mtl* and *dmp* genes as the goal was to check the true colouration pattern of the seeds specifically in scutellum. Two more inbreds viz., EC-994242 and EC-994243 with *CI-I* allele obtained from Maize Genetics Cooperation Centre, USA were used as female and crossed with the two *RI-nj* based inbreds (MGU-R1nj-101 and MGU-R1nj-102) as male. EC-994242 and EC-994243 were considered as control for the presence of *CI-I* allele in the experiment.

Characterization of anthocyanin colour expression

The crossed ears were harvested at physiological maturity stage and were dried until it reached 12–14% moisture [11]. Randomly selected 30 F₁ kernels from each of the crosses were scored on the basis of *RI-nj* marker expression. Kernels were scored on the scale of 0–5 based on increasing anthocyanin colour intensity on the endosperm as well as scutellum. Scale-0 indicated no colouration. Scale-1 possessed 10% expression, while scale-2, -3, -4 and -5 had the expression of 25%, 50%, 75% and 100% area in the crown region of the endosperm, respectively. While in scutellum, the scale-1 to 5 denoted the intensity of colour from very light to dark purple shades.

Sequence comparison and primer designing

The sequence of wild-type allele *CI* (Accession Number: X06333.1) and the inhibitor allele *CI-I* (Accession Number: X52201.1) were retrieved from EMBL database. Sequence of these two alleles were aligned and analyzed through BioEdit [19] software and MEGA tool using ClustalW. Two polymorphisms viz., an 8 bp insertion at position number 2051 and an SNP (A–G) at position 1795 differentiated the *CI* and *CI-I* allele (Fig. S1). Primers specific to the two

polymorphisms were designed using Primer3 online software (Table S1).

Genomic DNA isolation and PCR amplification

Genomic DNA of inbreds was isolated from leaf sample using modified CTAB extraction method [20], and quality was checked using 0.8% gel electrophoresis. PCR was carried out in 96 well plate Veriti 96-well thermal cycler (M/s. Applied Biosystems) with initial denaturation at 95 °C for 5 min, 35 cycles of amplification phase consisting of denaturation at 95 °C for 45 sec, primer annealing at 60 °C for 45 sec, primer extension at 75 °C for 45 sec and final extension at 72 °C for 8 min. The oligonucleotide primer was synthesized from M/s Macrogen. Final 20 µl reaction mixture consisted of 10 ng template DNA, 1 × Master mix (M/s G-Biosciences) and 0.25 µM primers. The amplified DNA fragments were separated using 2% Seakem LE agarose gel electrophoresis. AlphaImager (M/s Alpha Innotech, San Leandro, CA) gel documentation system was used for gel visualization.

Statistical analysis

Analysis of variance (ANOVA) for the expression of purple colouration in the endosperm and scutellum was undertaken using Windostat 8.0. Critical difference (CD) was used to establish difference between the two means was also calculated. The average and percentage calculations besides graphical representation were carried out using MS-Office-2019.

Results

Variation in expression of *R1-nj* phenotype among inbreds

ANOVA revealed significant differences among inbreds for colouration intensity in endosperm and scutellum (Table S2). Among the 178 inbreds under study, 136 genotypes possessed purple colour in the endosperm, while 112 inbreds developed colour in the scutellum (Fig. S2, Table S3). The expression of anthocyanin was also completely absent in two inbreds viz., EC-994242 and EC-994243 possessing *CI-I* allele. Expression of purple colour conditioned by *R1-nj* varied greatly in both endosperm and scutellum (Fig. 1). For instance, score-0 was observed in CM139, score-1 in ZL17333, while score-2 and 3 was shown by CML52 and CML384, respectively. While, score-4 was observed in CML373 and score-5 was observed in CML40. In scutellum, score-0 was observed in BAJIM5, while score-1, score-2, score-3, score-4 and score-5 were noticed in CAL1723,

CML543, CML556, CML582, and CML465, respectively. 25.28% inbreds possessed scale-5, while 20.22%, 8.99%, 5.06% and 16.85% inbreds had the scale of 4, 3, 2 and 1, respectively in the endosperm with respect to crosses with MGU-R1nj-101. When the inbreds were crossed with MGU-R1nj-102, 22.47% inbreds had the scale of 1, while 5.62%, 13.48%, 26.40% and 8.43% inbreds displayed the scale of 2, 3 4 and 5, respectively. However, in both the cases, 23.60% inbreds did not show any colour accumulation in the endosperm (Table 1). In the case of scutellum, 39.33% had the score of 1 when crossed with MGU-R1nj-101. Besides, 12.92% inbreds expressed score of 2, while score 3, 4 and 5 were observed in 6.18%, 3.37% and 1.12% of the inbreds, respectively (Fig. 2, Table 1). With respect to MGU-R1nj-102, 46.63% inbreds displayed scale-1, while 11.24%, 2.25%, 2.25% and 0.55% inbreds possessed scale-2, -3, -4 and -5, respectively. However, 37.08% of the inbreds scored 0 (no colour) for both MGU-R1nj-101 and MGU-R1nj-102. On an average, 19.66% inbreds displayed scale-1 (10% colour), while 5.34%, 11.24%, 23.31% and 16.85% inbreds showed scale -2 (25% colour), -3 (50% colour) -4 (75% colour) and -5 (100% colour) in the endosperm, respectively (Fig. 2). In case of scutellum, 42.98% inbreds possessed score-1 with 12.08%, 4.21%, 2.81% and 0.84% inbreds displayed scale-2, -3, -4 and -5, respectively. Of the total 178 inbreds, 112 inbreds showed *R1-nj* phenotype in both endosperm and scutellum, while 42 inbreds did not possess colour in both the tissues. Interestingly, 24 inbreds revealed purple phenotype in the endosperm without having any colour in the scutellum. However, no inbred with coloured scutellum and colourless endosperm was observed. The mean score among the CIMMYT inbreds was 2.56 in the endosperm, while the same was 2.26 among the inbreds developed in India. Similarly, the purple colouration score in scutellum was 1.05 and 0.84 among CIMMYT and Indian inbreds, respectively. Though, CIMMYT inbreds possessed higher intensity of purple colouration over the Indian lines, the difference was not significant (Table S3). Further, average colour score using MGU-R1nj-101 as male was 2.61 and 1.03, respectively, while the same was 2.22 and 0.88 for MGU-R1nj-102 in endosperm and scutellum, respectively. However, the difference between the score posed by MGU-R1nj-101 and MGU-R1nj-102 were non-significant.

Genotyping of inbreds using markers

PCR amplicon with 8 bp deletion was observed in the control inbred (PV-19A-BLSTBLKS-1858) having *CI* gene, while the insertion of 8 bp specific to *CI-I* in two control inbreds (EC-994242 and EC-994243) was detected (Figs. 3 and 4). Similarly, amplicon with SNP-G was observed in the controls with *CI-I*, while SNP-A was detected in *CI* allele. Genotyping using *MGU-CI-InDel8*

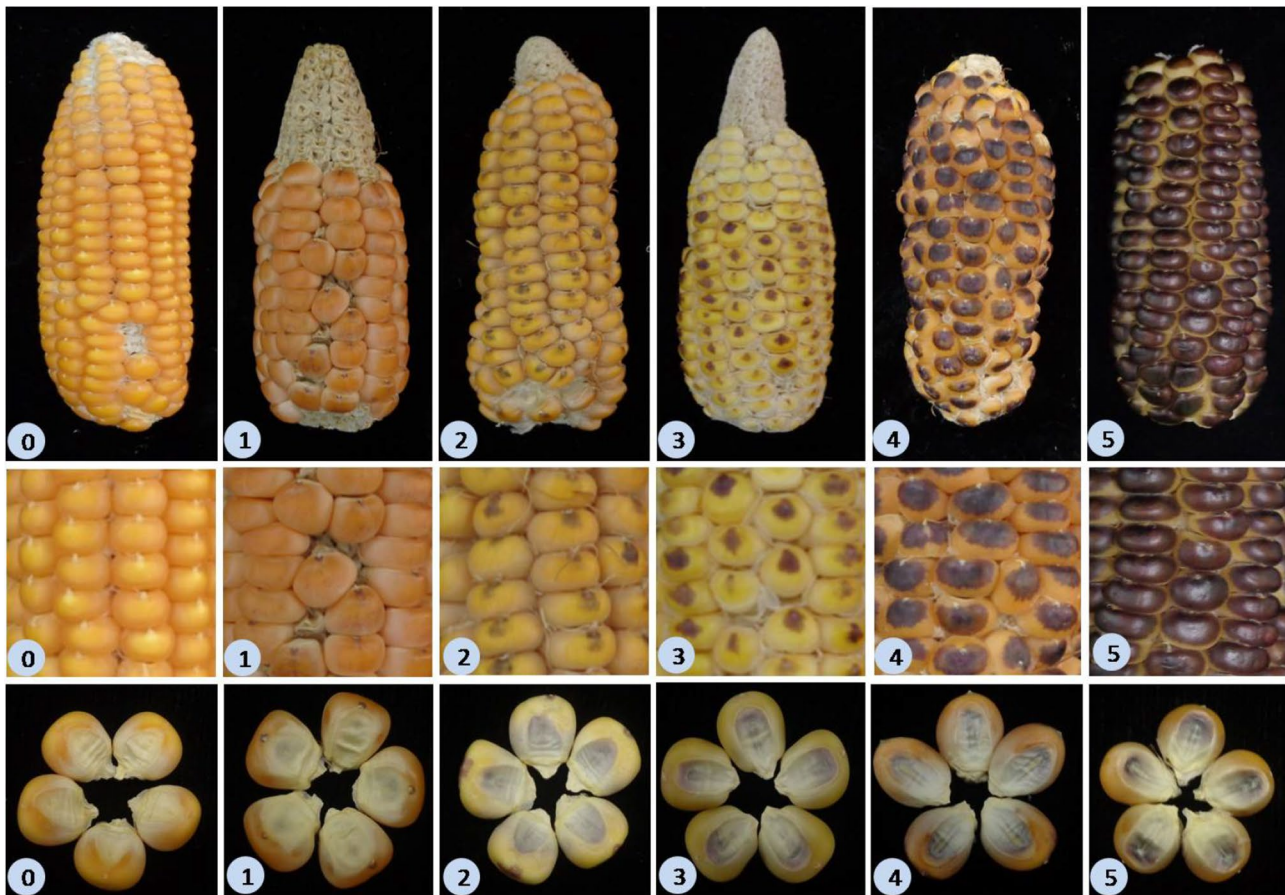


Fig. 1 Expression pattern of colour marker (0–5 scale) on endosperm and scutellum. Endosperm genotypes: score-0: CM139, score-1: ZL17333, score-2: CML52, score-3: CML384, score-4: CML373,

score-5: CML40. Scutellum genotypes: score-0: BAJIM5, score-1: CAL1723, score-2: CML543, score-3: CML556, score-4: CML582, score-5: CML465

Table 1 Expression of colour marker *R1-nj* on endosperm and scutellum among in breeds

Part of kernels (female)	<i>R1-nj</i> Donor (male)	Scale (level of expression in the crown region)						Total
		0	1	2	3	4	5	
Endosperm	A	42	30	9	16	36	45	178
	B	42	40	10	24	47	15	178
	Average	42	35	9.5	20	41.5	30	178
Scutellum	A	66	70	23	11	6	2	178
	B	66	83	20	4	4	1	178
	Average	66	76.5	21.5	7.5	5	1.5	178

A: MGU-R1nj-101, B: MGU-R1nj-102

revealed 14 inbreds possessed 8 bp insertion, while 164 inbreds had the 8 bp deletion. Similarly, assay with *MGU-CI-SNPI* revealed that 165 inbreds had SNP-A and 13 inbreds possessed SNP-G. When both the markers were considered together, nine inbreds had 8 bp insertion and SNP-A, and the same number of inbreds possessed 8 bp deletion with SNP-G. On the other hand, 8 bp insertion with SNP-G was observed in five inbreds, while 155 inbreds revealed 8 bp deletion with SNP-A.

Effectiveness of markers in detecting the presence of *C1* and *C1-I*

When the marker profiles were compared with the colour phenotypes, it was found that *MGU-CI-InDel8* correctly identified the presence of *C1-I* and *C1* with a success rate of 92.9 and 84.7%, respectively (Table 2). In case of *MGU-CI-SNPI*, the success rate in detecting the presence of *C1-I* was 39.5%, while the same was 83.6% for *C1*. However,

Fig. 2 Average expression (based on A and B donor of *RI-nj* inbreds) of anthocyanin colour marker *RI-nj* on endosperm and scutellum among maize inbreds. **A:** MGU-R1nj-101, **B:** MGU-R1nj-102. Scale-0: No colour, Scale-1 to 5: varying from light to dark

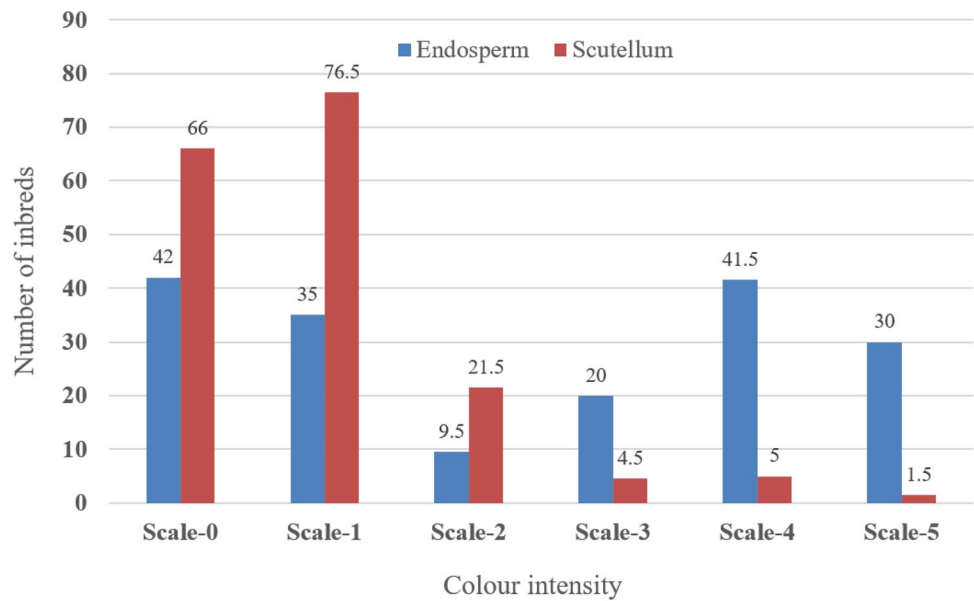


Fig. 3 Profile of *MGU-CI-InDel8* marker among selected set of inbreds; L: 50 bp ladder

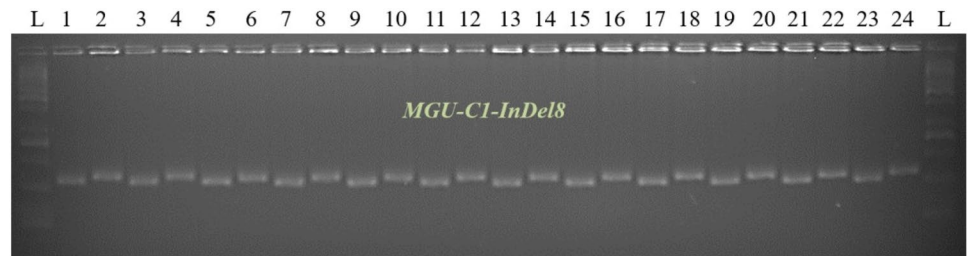
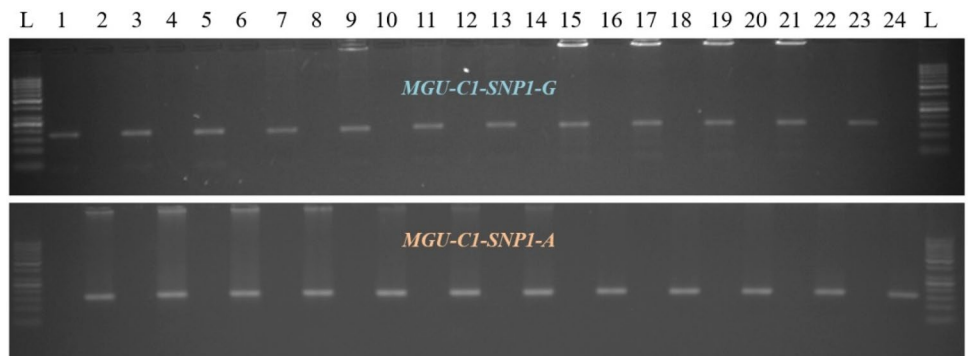


Fig. 4 Profile of *MGU-CI-SNPI* marker among a set of inbreds; L: 50 bp ladder



when both the markers were compared together, the success rate of detecting *CI-I* was 100%. Interestingly, the success rate of identifying *CI* was only 82.6% with a failure rate of 17.3% (Table 2).

Discussion

DH technology has emerged as an alternative to traditional methods of inbred development in maize breeding [4]. DH helps to accelerate the production of completely homozygous lines in much cost-effective way than the traditional selfing approach [21]. Selection of haploid seeds is an important step in DH production. Among various genetic systems, the *RI-nj* phenotypic marker system is widely used

Table 2 Relationship of marker polymorphisms with anthocyanin colouration

Marker	Polymorphism	Gene designation as per marker	Number of inbreds with anthocyanin colouration	Number of inbreds with no anthocyanin colouration	Total number of inbreds analyzed
MGU-CI-InDel8	8 bp insertion	<i>CI-I</i>	1	13	14
	8 bp deletion	<i>CI</i>	139	25	164
	Total seeds	–	140	38	178
MGU-C1-SNP1	SNP-G	<i>CI-I</i>	8	5	13
	SNP-A	<i>CI</i>	138	27	165
	Total seeds	–	146	32	178
MGU-CI-InDel8 + MGUC1-SNP1	8 bp insertion + SNP-G	<i>CI-I</i>	0	5	5
	8 bp deletion + SNP-A	<i>CI</i>	128	27	155
	8 bp insertion + SNP-A	–	1	8	9
	8 bp deletion + SNP-G	–	8	1	9
	Total seeds	–	137	41	178

Presence of anthocyanins either in endosperm and/or scutellum were scored as coloured, while absence in both were scored as colourless

for haploid identification [22]. *R1* codes for basic-Helix-Loop-Helix (bHLH) transcriptional factor that directly interacts with *CI* to activate transcription of anthocyanin biosynthesis genes in seeds [23]. The dominant allele *R1* confers colouration in the entire endosperm and scutellum, while its recessive *r1* allele does not impart colour in the seeds. *R1-nj*, an allele of *R1* also behaves dominantly, but causes purple colouration in the crown and its surrounding areas of the endosperm and scutellum. Majority of HI lines developed worldwide possess *R1-nj* allele [11]. However, the expression of *R1-nj* is not uniform across germplasm, and it may vary from light to dark colouration with even no formation of colour in specific germplasm [18]. In the present study, we analyzed the expression of purple colour on the seeds of the inbreds adapted to Indian subtropical conditions for estimating the effectiveness of *R1-nj* allele in the DH programme.

We observed that 62.9% of the inbreds had expression of *R1-nj* allele in both endosperm, and scutellum. Chaikam et al [11] reported that 27–49% of tropical inbreds showed inhibition of *R1-nj* colour marker, while it was 27% among the landraces. Khulbe et al. [17] analyzed 18 hybrids and 23 inbreds specifically adapted to the hilly regions of India, and reported that 97.6% of the maize genotypes expressed *R1-nj* colour marker. Among these, the expression level also varied drastically from extremely low to high intensity of purple colouration [17]. This could be due to varied expression levels of the key genes such as *A1*, *A2*, *Bz1*, *Bz2*, *CI*, *C2* and *Pr1* in the anthocyanin biosynthesis pathway [10]. Further, many modifier genes in the genetic-background might have influenced the expression of the *R1-nj* allele. It was also observed that purple colour intensity varied among the same inbred when crossed with the two *R1-nj* inbreds. MGU-R1nj-101 when used as male produced higher colour intensity in endosperm and scutellum over MGU-R1nj-102.

This suggested that genetic constitution of the *R1-nj* inbreds also influences the colour intensity. Earlier experiments by Chaikam et al. [11] and Khulbe et al. [17] reported anthocyanin pattern only with one *R1-nj* based HI line. Thus, the influence of the background of *R1-nj* lines on the expression of anthocyanin is a novel finding of the study. Since, various *R1-nj* based HI lines are used worldwide, the intensity of colour development posed by HI lines on the source germplasm for haploid identification would be of great significance. Further, colour intensity upon crossing with *R1-nj* lines was more among the CIMMYT-based inbreds compared to the inbreds developed in India. Similar observation of the differences in colour expression among various germplasm groups have been reported earlier by Prigge et al. [14] and Couto et al. [24]. This suggested that various modifier loci inhibiting the expression of purple colour may be more prevalent among Indian inbreds.

Further, 13.5% of the inbreds developed colour in the endosperm but not in the scutellum. This may be due to the presence of thicker and opaque pericarp (rather than thin and transparent pericarp) that blocked the viewing of coloured regions in the scutellum. Generally, such kind of seeds are treated as haploids when HI lines are used. To avoid the occurrence of such seeds, we deliberately selected *R1-nj* based inbreds without the haploid induction capacity as they did not possess *mtl* and *dmp* genes. Thus, seeds with coloured endosperm and colourless scutellum were not actually haploids but true diploids with uncharacteristic colour pattern. Such kind of seeds would cause identification of false positive haploids in the DH breeding programme [14]. While Khulbe et al. [17] did not report such kind of colourless scutellum, Chaikam et al. [11] mentioned the appearance of such kernels in their study. Since they used *R1-nj*-based HI lines, it was possible to have haploid seeds with coloured endosperm and colourless scutellum. However, in our study

we conclusively showed that the appearance of seeds with coloured endosperm and colourless scutellum in the *RI-nj* based crossing programme is a common phenomenon in certain genetic background. The *RI-nj* based selection of false positive haploids can be reduced by using alternative marker systems such as (i) purple coloured root and stem with the combination of *PII* and *BI* genes [25], (ii) purple coloured root marker [26], and (iii) embryo-based high oil marker-system [8].

The present analysis also suggested that 23.6% of the inbreds did not develop purple colour in both endosperm and scutellum. This is mainly due to the presence of *CI-I* allele among the inbreds [13]. The *CI-I* codes 252 amino acid protein, while *CI* coded protein is composed of 273 amino acids [13]. The *CI-I* coded suppressor protein out competes the functional *CI* protein for activator sites of the anthocyanin structural genes like *RI-nj* [27]. The presence of *CI-I* in the source population makes the visible selection of the haploid kernels difficult. In haploid induction crosses, generally two types of seeds are formed, viz., (i) F_1 seeds with purple colour in both endosperm and scutellum are treated as true diploids, and (ii) F_1 seeds with purple coloured endosperm and colourless scutellum are identified as haploids. Besides, F_1 seeds with no colour in endosperm and scutellum are also formed when *CI-I* allele is present in the source germplasm from which DH lines are derived. Thus, even though true haploid seeds are generated upon crossing with HI lines, they remain unidentified phenotypically. This leads to loss of valuable resources and time incurred in making unnecessary crosses during the haploid induction programme.

Therefore, prior identification of *CI-I* allele in the source germplasm assumes great significance in the DH programme. Here we developed two breeder-friendly PCR-based markers viz., *MGU-CI-InDel8* and *MGU-CI-SNP1* specific to 8 bp InDel and an SNP (A to G), respectively that differentiated the *CI* and *CI-I* allele. Although, the effectiveness of *MGU-CI-InDel8* and *MGU-CI-SNP1* in identifying *CI-I* was much higher (> 80%), none of the markers alone could detect the presence of *CI-I* with 100% success rate. However, when both the markers were considered together, they predicted with 100% accuracy for the presence of *CI-I* allele among the inbreds. Thus, haplotype having 8 bp insertion and SNP (G) provided the accurate prediction of presence of *CI-I* allele in the source populations. In contrast, Chaikam et al. [11] observed 84%, 79% and 81% success when scored with KASPar marker-assay having 8 bp InDel, SNP (A to G) and combination of both, respectively in a set of maize genotypes comprised of inbreds and landraces. The 8 bp insertion mutation in the last exon leads to premature termination of the open reading frame (ORF) in the *CI-I* protein, thereby shifting the activity from transcriptional activator to a transcriptional repressor [13]. On the other hand,

the SNP (A–G) conversion codes the same amino acid serine in both *CI* and *CI-I* allele, thereby suggesting no role in protein function. However, SNP-G might have evolutionary advantage and perpetuated in the population with 8 bp insertion due to selection pressure for not selecting *CI* allele in the traditional breeding programme. Earlier, Chaikam et al. [11] developed KASPar assay for the 8 bp and SNP polymorphisms to differentiate the *CI* and *CI-I* allele. Though KASPar assay is high throughput as it can detect 1536 SNPs, it involves very high cost with US\$246 per sample (26 samples' cost US\$6390) [28]. In a breeding programme where only two polymorphisms specific to *CI-I* are to be assayed, per sample cost for genotyping becomes too high. On the other hand, the breeder-friendly PCR assay developed here is cost effective as it needs only US\$1–1.5 per sample. The assay is simple and can be performed in any laboratory equipped with very few basic equipment like PCR machine, gel electrophoresis and gel documentation unit unlike costly high throughput system required for KASPar or any other next generation sequencing systems. The PCR-based marker system is quite affordable by the breeding programme especially in the developing countries where resources are limited.

However, the effectiveness of *MGU-CI-InDel8* marker in detecting *CI* allele was higher (> 80%) than the *MGU-CI-SNP1* (< 40%). When combined, both the markers however could not provide 100% success rate. This suggested that even though *CI* gene was present in the genetic background, there could be formation of no colour in the endosperm and scutellum. This is governed possibly due to the presence of other suppressor genes in the background. Various dominant inhibitory genes viz., *C2-Idf* (*inhibitor diffuse*), and *in-1D* affecting expression of anthocyanins in the maize endosperm and scutellum have been reported [29, 30]. *C2-Idf* prevents anthocyanin accumulation in homozygous genotypes and reduces the pigmentation in heterozygous genotypes through RNA silencing [31]. The semi-dominant mutation *in-D* inhibits anthocyanins production in the aleurone tissue by regulating the *Whp* gene, a homologue of *C2* gene [32]. Since, the markers were developed specific to *CI*, the presence of other suppressor loci such as *C2-Idf* and *in-1D* could not be verified. Therefore, identification of causal polymorphism and development of breeder-friendly marker system will pave the way for precise genotyping of the source population before attempting the induction crosses in the DH programme. This is the first comprehensive analysis on expression of *RI-nj* marker in subtropically adapted inbreds, and development of effective breeder-friendly markers to predict the presence of *CI-I* gene.

Conclusions

R1-nj marker showed wide range of variation in imparting purple colour in both endosperm and scutellum among the inbreds. Considerable number of inbreds did not develop anthocyanin colouration. Independently, two breeder-friendly markers specific to 8 bp InDel and SNP (A–G) showed high degree of effectiveness in identifying the *CI-I* allele. However, the two markers predicted the presence of *CI-I* allele with 100% accuracy when considered together. These two markers thus developed here can be used to detect the presence of *CI-I* in the source germplasm. This would save valuable resources and time in the DH breeding programme.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11033-022-08214-2>.

Acknowledgements Authors thank ICAR-IARI, New Delhi for providing the lab and field facilities. The support of AICRP centres and CIMMYT, Mexico for providing the inbred lines is also acknowledged.

Author contribution Conduct of the experiment: NG; Primer designing and genotyping: RC; Generation of crosses: RUZ; Maintenance of inbreds and raising of field experiment: VM; Recording of anthocyanin score: RC, SD and GC; Manuscript writing: NG and FH; Editing of manuscript: ELD, VM, JM and AK; Designing of experiment: FH, KS and SC.

Funding The financial support of IARI, New Delhi and DBT sponsored project on ‘Development of locally adapted haploid inducer lines in maize through marker-assisted introgression of pollen-specific *MAT-RILINEAL* phospholipase gene’ [BT/PR25134/NER/95/1035/2017] is duly acknowledged.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Seguí-Simarro JM, Jacquier N, Widiez T (2021) Overview of in vitro and in vivo doubled haploid technologies. Doubled haploid technol. https://doi.org/10.1007/978-1-0716-1315-3_1
- Trampe B, Batîru G, Pereira da Silva A, Frei UK, Lübberstedt T (2022) QTL mapping for haploid inducibility using genotyping by sequencing in maize. *Plants* 11(7):878. <https://doi.org/10.3390/plants11070878>
- Dermail A, Chankaew S, Lertrat K, Lübberstedt T, Suriharn K (2021) Selection gain of maize haploid inducers for the tropical Savanna environments. *Plants* 10(12):2812. <https://doi.org/10.3390/plants10122812>
- Uliana Trentin H, Frei UK, Lübberstedt T (2020) Breeding maize maternal haploid inducers. *Plants* 9(5):614. <https://doi.org/10.3390/plants9050614>
- Kelliher T, Starr D, Richbourg L, Chintamanani S, Delzer B, Nuccio ML, Green J, Chen Z, McCuiston J, Wang W, Liebler T (2017) Matrilineal, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* 542(7639):105–109. <https://doi.org/10.1038/nature20827>
- Zhong Y, Liu C, Qi X, Jiao Y, Wang D, Wang Y, Liu Z, Chen C, Chen B, Tian X, Li J (2019) Mutation of *ZmDMP* enhances haploid induction in maize. *Nat plants* 5(6):575–580. <https://doi.org/10.1038/s41477-019-0443-7>
- Chaikam V, Molenaar W, Melchinger AE, Boddupalli PM (2019) Doubled haploid technology for line development in maize: technical advances and prospects. *Theor Appl Genet* 132(12):3227–3243. <https://doi.org/10.1007/s00122-019-03433-x>
- Melchinger AE, Schipprack W, Mi X, Mirdita V (2015) Oil content is superior to oil mass for identification of haploid seeds in maize produced with high oil inducers. *Crop Sci* 55(1):188–195. <https://doi.org/10.2135/cropsci2014.06.0432>
- Geiger HH (2009) Doubled haploids handbook of maize. Springer, New York. https://doi.org/10.1007/978-0-387-77863-1_32
- Nanda DK, Chase SS (1966) An embryo marker for detecting monophloids of maize (*Zea Mays* L.) 1. *Crop Sci* 6(2):213–215. <https://doi.org/10.2135/cropsci1966.0011183X000600020036x>
- Chaikam V, Nair SK, Babu R, Martinez L, Tejomurtula J, Boddupalli PM (2015) Analysis of effectiveness of *R1-nj* anthocyanin marker for in vivo haploid identification in maize and molecular markers for predicting the inhibition of *R1-nj* expression. *Theor Appl Genet* 128(1):159–171. <https://doi.org/10.1007/s00122-014-2419-3>
- Coe EH Jr (1962) Spontaneous mutation of the aleurone color inhibitor in maize. *Genetics* 47(6):779. <https://doi.org/10.1093/genetics/47.6.779>
- Cone KC, Burr FA, Burr B (1986) Molecular analysis of the maize anthocyanin regulatory locus *C1*. *Proc Natl Acad Sci USA* 83(24):9631–9635. <https://doi.org/10.1073/pnas.83.24.9631>
- Prigge V, Sánchez C, Dhillon BS, Schipprack W, Araus JL, Bänziger M, Melchinger AE (2011) Doubled haploids in tropical maize: I. Effects of inducers and source germplasm on in vivo haploid induction rates. *Crop Sci* 51(4):1498–1506. <https://doi.org/10.2135/cropsci2010.10.0568>
- Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H (1987) The regulatory *c1* locus of *Zea mays* encodes a protein with homology to *myb* proto-oncogene products and with structural similarities to transcriptional activators. *EMBO J* 6(12):3553–3558. <https://doi.org/10.1002/j.1460-2075.1987.tb02684.x>
- FAOSTAT (2020) Trade/crops and livestock products. <https://doi.org/http://www.fao.org/faostat/en/#data/TP>
- Khulbe RK, Pattanayak A, Panday V (2019) *R1-nj* expression in parental inbreds as a predictor of amenability of maize hybrids to *R1-nj*-based doubled haploid development. *Indian J Genet Plant Breed* 79(04):678–684. <https://doi.org/10.31742/IJGPB.79.4.5>
- Prasanna BM, Chaikam V, Mahuku G (2012) Doubled haploid technology in maize breeding: theory and practice. CIMMYT, Batan
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp Series* 41:95–98
- Saghai-Marouf MA, Soliman KM, Jorgensen RA, Allard R (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci USA* 81(24):8014–8018. <https://doi.org/10.1073/pnas81248014>
- Battistelli GM, Von Pinho RG, Justus A, Couto EGO, Balestre M (2013) Production and identification of doubled haploids in

- tropical maize. *Genet Mol Res* 12(4):4230–4242. <https://doi.org/10.4238/2013>
22. Rádi F, Török K, Nagymihály M, Kereszt A, Dudits D (2020) Improved reliability in production of maize inbred lines by the combination of the *R1-navajo* marker with flow cytometry or microsatellite genotyping. *Cereal Res Commun* 48(4):423–430. <https://doi.org/10.1007/s42976-020-00054-9>
 23. Dellaporta SL, Greenblatt I, Kermicle JL, Hicks JB, Wessler SR (1988) Molecular cloning of the maize R-nj allele by transposon tagging with Ac. *Chromosome structure and function*. Springer, Boston, MA, pp 263–282
 24. Couto EGO, Cury MN, Bandeira E, Souza M, Granato ÍSC, Vidotti MS, Domingos Garbuglio D, Crossa J, Burgueño J, Fritsche-Neto R (2019) Effect of F₁ and F₂ generations on genetic variability and working steps of doubled haploid production in maize. *PLoS One* 14(11):e0224631. <https://doi.org/10.1371/journal.pone.0224631>
 25. Rotarenco V, Dicu G, Fuia S (2010) New inducers of maternal haploids in maize. *Maize Genet Coop Newsl* 84:21–22
 26. Chaikam V, Martinez L, Melchinger AE, Schipprack W, Boddupalli PM (2016) Development and validation of red root marker-based haploid inducers in maize. *Crop Sci* 56(4):1678–1688. <https://doi.org/10.2135/cropsci2015.10.0653>
 27. Goff SA, Cone KC, Fromm ME (1991) Identification of functional domains in the maize transcriptional activator C1: comparison of wild-type and dominant inhibitor proteins. *Genes Dev* 5(2):298–309. <https://doi.org/10.1101/gad.5.2.298>
 28. Semagn K, Babu R, Hearne S, Olsen M (2013) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Mol Breed* 33(1):1–4. <https://doi.org/10.1007/s11032-013-9917-x>
 29. Coe EH Jr, Neuffer MG, Hoisington DA (1988) The genetics of corn. *Corn Corn Improvement* 18:81–258. <https://doi.org/10.2134/agronmonogr18.3ed.c3>
 30. Stinard PS, Sachs MM (2002) The identification and characterization of two dominant *r1* haplotype-specific inhibitors of aleurone color in *Zea mays*. *J Hered* 93(6):421–428. <https://doi.org/10.1093/jhered/93.6.421>
 31. Della Vedova CB, Lorbiecke R, Kirsch H, Schulte MB, Scheets K, Borchert LM, Scheffler BE, Wienand U, Cone KC, Birchler JA (2005) The dominant inhibitory *chalcone synthase* allele *C2-Idf* (inhibitor diffuse) from *Zea mays* (L.) acts via an endogenous RNA silencing mechanism. *Genetics* 170(4):1989–2002. <https://doi.org/10.1534/genetics.105.043406>
 32. Franken P, Niesbach-Klößgen U, Weydemann U, Maréchal-Drouard L, Saedler H, Wienand U (1991) The duplicated chalcone synthase genes *C2* and *Whp* (white pollen) of *Zea mays* are independently regulated; evidence for translational control of *Whp* expression by the anthocyanin intensifying gene in. *EMBO J* 10(9):2605–2612. <https://doi.org/10.1002/j.1460-2075.1991.tb07802.x>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.