Theor Appl Genet (2015) 128:159–171
DOI 10.1007/s00122-014-2419-3

ORIGINAL PAPER

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

Vijay Chaikam · Sudha K. Nair · Raman Babu · Leocadio Martinez · Jyothsna Tejomurtula · Prasanna M. Boddupalli

Received: 9 June 2014 / Accepted: 14 October 2014 / Published online: 11 November 2014

Abstract

Key message ***R1-nj*** **anthocyanin marker inhibition is highly frequent in tropical maize germplasm considerably affecting efficiency of haploid identification. Molecular markers reliably differentiating germplasm with anthocyanin color inhibitor have been identified in this study.**

Abstract The *R1-Navajo* (*R1-nj*) color marker facilitates easy and quick identification of haploid kernels at the seed stage during in vivo haploid induction process in maize. However, the Navajo phenotype can be completely suppressed or poorly expressed in some germplasm, making it impossible or inefficient to identify haploids at the seed stage. In this study, we characterized the expression of *R1-nj* marker in a large array of tropical/subtropical inbred lines, breeding populations and landraces by crossing with the *R1-nj*-based tropicalized haploid inducer. There was a high frequency of inhibition of the Navajo phenotype in the maize inbred lines, which are used in tropical breeding programs. Genome-wide association mapping showed that the *C1* anthocyanin regulatory locus is the most significant genetic factor influencing inhibition of the Navajo phenotype. Molecular marker assays were designed based on polymorphism in the *C1* vs *C1-I* alleles. Analysis of a set of 714 inbred lines demonstrated that a combination of two gene-specific markers—8 bp *C1-I* InDel and *C1-I* SNP—could predict with high accuracy the presence of anthocyanin color inhibition in the germplasm analyzed. Information generated in this study aids in making informed decisions on the constitution of source populations for doubled haploid (DH) line development in tropical germplasm, particularly those derived from elite maize lines from CIMMYT. The *C1-I* gene-specific molecular markers identified and validated will facilitate high-throughput and cost-effective evaluation of a large pool of germplasm for the presence of the dominant color inhibitor in maize germplasm.

Communicated by Xianchun Xia.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2419-3) contains supplementary material, which is available to authorized users.

V. Chaikam · L. Martinez · J. Tejomurtula
International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico D.F, Mexico

S. K. Nair · R. Babu
International Maize and Wheat Improvement Center (CIMMYT), ICRISAT Campus, Patancheru, Greater Hyderabad 502324, India

P. M. Boddupalli (✉)
International Maize and Wheat Improvement Center (CIMMYT), ICRAF Campus, UN Avenue, Gigiri, Nairobi, Kenya
e-mail: b.m.prasanna@cgiar.org

Introduction

Accelerated development of homozygous maize parental lines, followed by selection for important traits and use in product development, is crucial for enhancing genetic gains and breeding efficiency. The doubled haploid (DH) technology, based on in vivo haploid induction (Prasanna et al. 2012) is now central to modern maize genetics and breeding. The technology has been effectively deployed by large-scale commercial breeding programs in Europe (Schmidt 2004), North America (Seitz 2005), and China (Chen and Li 2009). Several publications have highlighted the advantages of using DH technology in maize breeding programs, including simplified logistics, reduced time to generate parental lines and commercial hybrids, increased selection gain (Rober et al. 2005; Geiger and Gordillo 2009; Prasanna 2012; Battistelli et al. 2013), and enhanced

efficiency in trait introgression and transgene fixation (Forster et al. 2007; Lübberstedt and Frei 2012).

For large-scale production of DH lines using in vivo haploid induction, quick and accurate identification of a small proportion of haploid kernels among the progeny derived from a cross between a source germplasm (as female) and a haploid inducer (as male) is extremely important (Chaikam and Prasanna 2012). Chase and Nanda (1965), Nanda and Chase (1966), and Greenblatt and Bock (1967) describe an anthocyanin-based phenotypic marker for haploid identification induced by the *R1-nj* allele of the *R1* regulatory gene which regulates kernel anthocyanin biosynthesis. *R1-nj* expression requires other structural genes involved in the anthocyanin biosynthesis pathway, such as *A1*, *A2*, *C2*, *Bz1*, *Bz2,* and another regulatory gene *C1* (Chase 1969; Geiger 2009). The phenotypic expression of *R1-nj*, termed the "Navajo" phenotype, is characterized by purple coloration in the aluerone layer on the crown region of the endosperm and the scutellum of the embryo (Nanda and Chase 1966; Greenblatt and Bock 1967). Most of the haploid inducers described in the literature are based on the *R1-nj* marker system. However, the tropical maize germplasm used in breeding programs does not usually have the genetic constitution required for anthocyanin coloration. In a cross between source germplasm without anthocyanin coloration and a haploid inducer with the *R1-nj* marker system, the progeny display purple/red coloration on both the endosperm and diploid embryo, whereas the putative haploid kernels exhibit purple coloration on the endosperm, but not on the embryo, facilitating visual identification of haploid kernels (Chaikam and Prasanna 2012).

The *R1-nj* allele is dominant over the *r1* allele but its expression can vary significantly in different source germplasm when crossed with a haploid inducer, indicating the influence of the genetic background of the source germplasm, and possibly non-genetic or environmental factors (Chase 1952; Rober et al. 2005; Kebede et al. 2011; Prigge et al. 2011). In addition, some dominant inhibitory alleles of the genes involved in the anthocyanin pathway, such as *C1-I*, *C2-Idf* (inhibitor diffuse), and *in-1D*, were reported to cause inhibition of anthocyanin biosynthesis in the maize endosperm and embryo (Coe et al. 1988; Stinard and Sachs 2002). The *C1-I* allele is a dominant mutant of the wild-type *C1* allele (Coe 1962), whose gene product serves as a transcriptional repressor, outcompeting the functional *C1* protein for activator sites of the anthocyanin structural genes in heterozygous genotypes(*C1-I/C1*), preventing transcription of the structural genes (Paz-Ares et al. 1990). *C2-Idf* prevents anthocyanin accumulation in homozygous genotypes and reduces the pigmentation in heterozygous genotypes (Della Vedova et al. 2005). The inhibitory effect of *C2-Idf* occurs through RNA silencing (Della Vedova et al. 2005). The dominant wild-type allele of the Intensifier1 (*In1*) gene shares homology with the *R1/B1* gene family and suppresses the expression of anthocyanin coloration possibly by acting as a competitive inhibitor of *R1* while interacting with *C1* (Burr et al. 1996). *in-D* is a semi-dominant mutation of the *In (intensifier)* gene that inhibits overall production of anthocyanins in the aleurone tissue by regulating the *white pollen* (*Whp),* which is a homologue of *C2* gene (Coe et al. 1981; Franken et al. 1991). In a haploid induction cross, depending on the homozygosity or heterozygosity of the inhibitor alleles in the source population, *R1-nj* expression may be completely inhibited in all kernels or may segregate for color expression among the kernels. When color inhibition occurs due to the presence of dominant inhibitors like *C1-I*, it is not possible to visually identify haploid kernels. Also, when the expression is segregating among the kernels, it is not possible to identify all the haploids resulting from an induction cross. This reduces the efficiency of haploid identification based on crosses of source germplasm with the *R1-nj*-based haploid inducers.

To date, there are no published reports on the frequency of the dominant anthocyanin color inhibitors or the expression of the Navajo phenotype in tropical/subtropical maize germplasm when crossed with *R1-nj*-based genetic stock. This is of importance when considering the influence of such an occurrence on the amenability of the source germplasm for maternal haploid induction using the *R1-nj*-based haploid inducers. Consequently, one of the objectives of this study was to systematically document the prevalence of the kernel color inhibition trait in inbred lines, breeding populations, and landraces developed or conserved at International Maize and Wheat Improvement Center (CIMMYT). The phenotypic data and high-density genotypic information generated in this study were also used for a genome-wide association study (GWAS) to identify genetic factors that regulate kernel anthocyanain expression. High-density genotyping platforms, such as Illumina's infinium (MaizeSNP50 at http://res.illumina.com/) and genotyping by sequencing (GBS) (Elshire et al. 2011) are routinely used by maize researchers to discover useful marker-trait associations. The GBS platform in particular is high throughput in nature and often covers important regions of the genome that are inaccessible to sequence capture approaches.

Considering the time and resources used in haploid induction, it is important to effectively predict whether a given source germplasm has kernel color inhibitors in its genetic background. At present, source germplasm is crossed with *R1-nj*-based haploid inducer that has strong anthocyanin color expression in both the endosperm and the embryo, followed by visual analysis of the resulting progeny for expression of anthocyanin. The *C1* (wild type) allele and the *C1-I* allele of *C1*, which is

the most frequent kernel color inhibitor reported, have been sequenced (Cone et al. 1986; Paz-Ares et al. 1990). Sequence comparison of these alleles revealed that the most remarkable and potentially causative sequence change in *C1-I* (vis-à-vis *C1*) is an 8 base pair (bp) insertion mutation in the last exon, leading to premature termination of the open reading frame (ORF) in the *C1-I* protein. The *C1-I* protein has 252 amino acids, compared to 273 amino acids in *C1*, and the change in the 3′ terminus of the protein results in reduced acidity and conversion of this protein from a transcriptional activator (*C1*) to a transcriptional repressor (Cone et al. 1986). The second objective of the study was, therefore, to develop molecular marker assays for *C1-I* allele-based polymorphisms associated with anthocyanin color inhibition in the endosperm and embryo and to ascertain the utility of these assays as proxy screens for detecting genotypes with *C1-I* (color inhibition) before their use in the haploid induction process.

Materials and methods

Plant material and experimental crosses

Seed of the 82 tropical/subtropical advanced maize inbred lines used in the study was obtained from CIMMYT breeding programs based in Mexico, Kenya, and Zimbabwe, whereas 538 CMLs and 168 landraces were sourced from CIMMYT Gene Bank at El Batan, Mexico. A total of 380 inbred lines, comprising the Improved Maize for African Soils (IMAS) association mapping panel was also used. All inbreds and landraces were planted in the 2012 summer season at CIMMYT's Agua Fria Experimental Station, Mexico and were crossed with a *R1-nj*-based tropicalized haploid inducer, developed at CIMMYT in partnership with the University of Hohenheim (Prigge et al. 2012). A total of 157 breeding populations was crossed with the same inducer during the 2013 winter and summer seasons at Agua Fria station. Eighteen CMLs derived from the highland breeding program were planted at CIMMYT's El Batan experimental station located in Mexico State and crossed with the tropicalized inducer in the 2012 summer. Manual pollinations of each of the recipient lines/populations with the *R1-nj*-based haploid inducer stock were undertaken as described by Chaikam et al. (2012). Ears were harvested at physiological maturity and dried to optimum moisture content before being scored for the anthocyanin color phenotype. Inbreds that set none or only a few seed were excluded from further analysis. This reduced the final number of inbreds to 897, including 80 advanced maize inbred lines, 450 CMLs, and 367 IMAS lines. To validate the reproducibility of the experiments, we grew a set of 178 CMLs again in two replications in the 2013 summer and crossed them to the same haploid inducer as described above.

Characterization of anthocyanin color expression

The harvested and dried ears were arranged under day light, and the *R1-nj* phenotypic marker expression on the endosperm was recorded for inhibition of color marker expression, area marked, and color intensity by scoring on a scale of 1–5. From each ear, 10–15 kernels were randomly selected, shelled and scored on a scale of 1–5 for anthocyanin color inhibition and color intensity on the embryo. For tropical breeding populations and landraces, data were collected only for the endosperm traits. The score for color inhibition represents the proportion of kernels expressing the *R1-nj* marker; score of 1 for inhibition indicates that all kernels from the cross have color marker expression, a score of 2.5 indicates that nearly 50 per cent of the total kernels have color marker expression, and a score of 5 indicates complete color inhibition in all kernels. The score for the area marked trait represents the extent to which the endosperm expresses the Navajo phenotype; a score of 1 indicates anthocyanin color expression in almost the entire aleurone layer of the endosperm, a score of 2.5 indicates the expression of the Navajo phenotype in the crown region only, and a score of 5 represents a complete lack of Navajo expression. A score of 4.5 indicates that the Navajo phenotype was confined to anthocyanin expression at the silk attachment region. For intensity of the anthocyanin color expression on the kernel, a score of 1 indicates deep purple or red coloration that is easily recognized and a score of 5 is visually undetectable. To validate the accuracy of the phenotyping, we compared the scores for the 178 CML crosses in 2012 to those for the same CML crosses harvested in two replications in 2013. This comparison indicated high reproducibility and hence scores given in 2012 were used for further analysis. The *R1-nj* phenotypic expression data are summarized in supplementary Table 1. Variation in the Navajo phenotype, in terms of area marked by anthocyanin pigmentation, intensity of color, and anthocyanin color inhibition, is depicted in Fig. 1.

Phenotypic data analysis

To identify the extent of anthocyanin color expression or inhibition in different types of tropical germplasm, the genotypes were grouped into three categories: (a) all genotypes with a score of 1 and 1.5 were designated as having complete expression, (b) genotypes with a score of 2–4 were designated as displaying segregation for the trait, and (c) genotypes with a score above 4 were designated as having

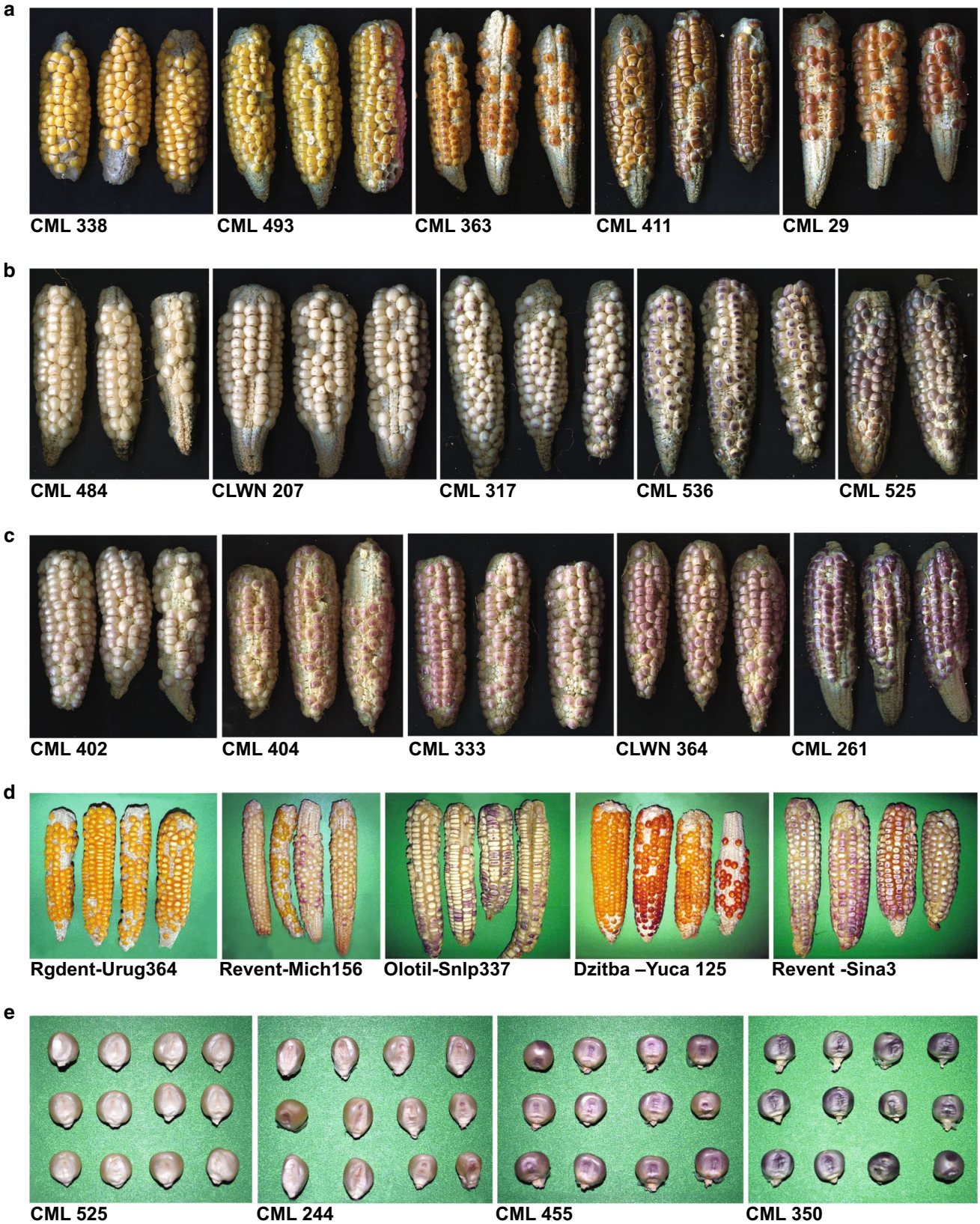

Fig. 1 Variation in *R1-nj* expression: **a** variation for endosperm area marked with Navajo expression in *yellow* kernel background; **b** variation for endosperm area marked with Navajo expression in *white* kernel background; **c** variation for intensity of *anthocyanin color*; **d** variation for *anthocyanin color* expression among landraces; **e** variation in the intensity of color expression on the embryo

Table 1 Three categories of *R1-nj* expression based on the proportion of kernels expressing the phenotype in different types of tropical germplasm

	Total entries	Full kernel color expression (%)	Segregating for color expression (%)	Complete color inhibition (%)	Likelihood ratio
Germplasm					
Breeding populations	157	56.05	40.13	3.82	<0.0001
Landraces	155	3.90	69.48	27.27	
Inbred lines	897	49.33	21.43	29.24	
Inbreds based on adaptation					
Lowland tropical	474	53.38	19.20	27.43	0.0082
Subtropical	358	44.41	25.42	30.17	
Highland	37	35.14	16.22	48.65	
Inbreds based on kernel color					
White	643	49.46	22.08	28.46	0.3586
Yellow	247	49.80	18.22	31.98	
Inbreds based on kernel texture					
Dent	115	51.30	21.74	26.96	0.9646
Flint	211	49.76	22.27	27.96	

complete inhibition of anthocyanin color (Table 1). Cumulative proportion graphs were used to compare the distribution of the proportion of entries with different scores for different traits. A likelihood ratio Chi square test was conducted to identify significant differences in the distribution of test scores, and *P* values were calculated. Significant differences ($P \leq 0.05$) are underlined in the cumulative graphs (Fig. 2).

Genotyping for genome-wide association mapping

Genotyping-by-sequencing (GBS) was undertaken at the Institute of Genomic Diversity (IGD), Cornell University. Genotypic data were generated following the standard GBS protocol (Elshire et al. 2011), using ApeKI as the restriction enzyme and multiplexing 96 samples on each Illumina flow cell lane. Raw reads from the machine were analyzed with approximately 18,000 additional maize samples, including Nested Association Mapping (NAM) and other linkage populations. The GBS discovery pipeline for species with a reference genome, available in TASSEL (version 3.0), was used. The data used for the GWAS were generated using a custom Java script that divided the entire SNP dataset into 1,024 SNP windows and looked for the most similar inbred line within each window to fill the missing data. The algorithm takes advantage of small identical by descent (IBD) regions shared between pairs of inbred lines in the collection; if the window from the closest neighbor had more than 5 % difference from the line being analyzed, the data point was recorded as missing. The GBS Zea database (approximately 22,000 samples) was used to search for the closest sample. All imputations were performed by the IGD, Cornell University, as part of GBS genotyping service. GBSv2.7 was used for GWAS analyses.

Genome-wide association analysis

Markers with a minor allele frequency (MAF) of less than 0.01 and call rate (CR) of less than 0.5 were discarded and the remaining (~410 K SNPs) were used in the GWAS analyses. A subset of high-quality GBS-SNPs with MAF $\geq$ 0.1 and CR $\geq$ 0.9 was identified for principal components-based structure analysis and to derive kinship matrices. Phenotypic data collected from 897 inbred lines for five different traits of the Navajo phenotype (inhibition of color on the endosperm, area marked on the endosperm, color intensity on the endosperm, color intensity on the embryo, and color inhibition on the embryo) were used for association mapping. Associations were determined using a General Linear Model (GLM) ("Q" correcting for population structure through the first 10 principal components) and Mixed Linear Model (MLM) ("Q+K" correcting for macro-population structure through 10 principal components and micro-population structure through a $n \times n$ kinship matrix) analyses for all traits (Flint-Garcia et al. 2005; Yu and Buckler 2006). All GLM and MLM computations used SVS software from Golden Helix (http://www.goldenhelix.com/SNPVariation/). The Mixed Model GWAS used a single locus (EMMAX) method (Kang et al. 2010) in SVS and a detailed description is available at http://doc.goldenhelix.com/SVS/latest/mixed_models.html. To determine the proportion of phenotypic variation explained by a combination of candidate SNPs, we performed a multiple linear regression for each trait with phenotype values as a response variable (y) and candidate SNPs as predictors (x), using the regression module in SVS.

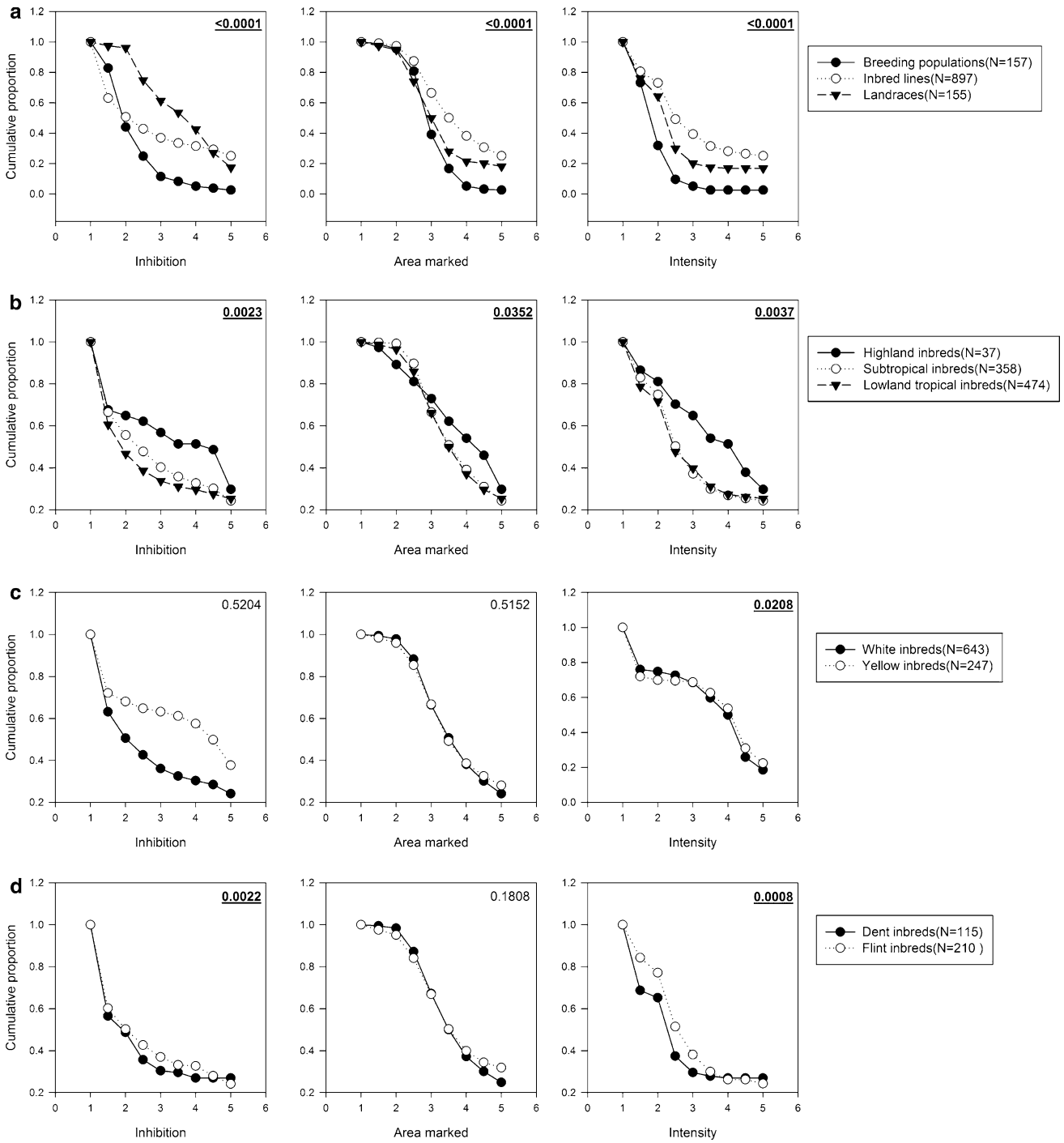

Fig. 2 *R1-nj* expression was characterized by giving scores for color marker inhibition, endosperm area marked, and color intensity on endosperm on a scale of 1–5. *Graphs* show score (*x-axis*) and cumulative proportion (*y-axis*). Likelihood Chi Square value is indicated on the *graph* and significant differences are *underlined*. Expression of *R1-nj* in **a** elite lines vs. breeding populations vs. land races; **b** tropical vs. subtropical vs. highland inbred lines; **c** *white* vs. *yellow*; **d** dent vs. flint

Marker assays and statistical analysis

The gene sequences of wild-type allele *C1* and the inhibitor allele *C1-I* were obtained from the EMBL database (accession numbers X06333.1 and X52201.1, respectively). Gene sequences of two additional *C1* alleles were obtained from the NCBI database (accession numbers AF320613.3 and AF320614.3). The 8-bp insertion in the terminal exon of the *C1-I* allele, which is the putative functional polymorphism, was identified using sequence alignment of

these four sequences with Clustal Omega. KASPar assays were developed for this polymorphism by LGC genomics (http://www.lgcgenomics.com).

Genomic DNA was extracted from young leaves (3–4 leaf stage) collected from a bulk of 15 plants per each inbred line from the plots planted for phenotyping studies, according to CIMMYT's laboratory protocols (CIMMYT 2005). DNA samples from 714 lines were assayed for the 8-bp insertion–deletion (InDel) and the KASPar assays did not identify any heterozygotes. The SNP marker, identified from the GWAS with the strongest association with the inhibition phenotype, was scored directly from GBS data developed for these 714 lines. The allele "G" was scored in the *C1-I* allele and allele "A" was scored in the non-*C1-I* alleles.

For the purpose of this analysis, all lines with an inhibition score of 4 or less were considered to be in the 'expression' category. The accuracy with which the *C1* gene-based markers can differentiate genotypes with and without the *C1-I* allele was determined by categorizing the tested lines into four categories as a 2 × 2 contingency table and testing the marker classification against the actual phenotype of the trait revealed by conventional crossing with a *R1-nj*-based haploid inducer. The terms "positive" and "negative" refer to the presence or absence of kernel color inhibition. The four classes in the 2 × 2 contingency table are (1) lines where the *C1-I* diagnostic marker and color inhibition are present (true positives; TP), (2) lines where the *C1-I* diagnostic marker is present but color inhibition is absent (false positives; FP), (3) lines where the *C1-I* diagnostic marker is absent and the color inhibition is present (false negatives; FN), and (4) lines where the *C1-I* diagnostic marker and color inhibition are absent (true negatives; TN). Fisher's Exact Test (Fisher 1922) was performed on the 2 × 2 contingency table to test the independence of the marker-based and the trait-based classifications. Statistics for the validity of screening tests were calculated (Altman and Bland 1994a, b) as follows:

True positive rate (TPR) = Proportion of lines having *C1-I* diagnostic marker among the total lines that showed color inhibition = TP/(TP + FN)
True negative rate (TNR) = Proportion of lines not having the *C1-I* diagnostic marker among the total lines that showed color expression = TN/(TN + FP)
Positive predictive value (PPV) = Proportion of samples with the *C1-I* diagnostic marker that are true positives = TPR × *P*/TPR × *P* + (1−TNR) × (1−*P*), where *P* = the prevalence of the color inhibition in the samples studied. Negative predictive value (NPV) = Proportion of samples without the *C1-I* diagnostic marker that are true negatives = TNR × (1−*P*)/(1−TPR) × *P* + TNR × (1−*P*), where *P* = the prevalence of the color inhibition in the sample studied. Accuracy (Number of true assessments)/Number of all assessments) = (TN + TP)/(TN + TP + FN + FP)

Results

Variability in expression of the Navajo phenotype

The Navajo phenotype expression could vary among progeny derived from crosses of landraces or inbred lines with the *R1-nj*-based haploid inducer, including presence/absence of anthocyanin color inhibition, segregation of kernels expressing the Navajo phenotype, area marked by anthocyanin pigmentation, and intensity of the color (Fig. 1). For majority of genotypes, the Navajo phenotype was present only on the crown region of the endosperm (which is typical of *R1-nj* expression), whereas in some inbreds (CML29 and CML525), the Navajo phenotype covered all of the endosperm. In a few inbreds the Navajo phenotype is limited to the silk attachment region or not expressed at all (Fig. 1a, b). The Navajo phenotype mostly occurs in purple or red with the intensity of the anthocyanin color expression varying from faint red to deep red or faint purple to deep purple in different genotypes (Fig. 1c). The progeny of crosses with some landrace accessions showed the Navajo phenotype in all kernels, while in some cases all kernels had complete anthocyanin color inhibition. In some landrace accessions, only a proportion of the total kernels expressed the Navajo phenotype (Fig. 1d). Navajo phenotype expression can vary in the scutellum (embryo), with some genotypes showing no expression, some showing faint expression, and some showing full expression, covering the entire embryo (Fig. 1e). Although the majority of the CMLs and elite inbred lines developed at the CIMMYT for lowland tropical, subtropical and highland agro-ecologies expressed the Navajo phenotype in the progeny when crossed with the *R1-nj*-based haploid inducer, broad range of variation were found among lines for inhibition, area marked, and intensity of the Navajo phenotype. A large number of inbreds showed inhibition and poor expression of the Navajo phenotype (Supplementary Tables 1, 2).

Data on different types of tropical germplasm with regard to the proportion of kernels expressing the Navajo phenotype in crosses with a *R1-nj*-based haploid inducer are presented in Table 1. Full expression and full inhibition indicates all kernels resulting from the cross are either expressing or inhibiting the Navajo phenotype, respectively, whereas segregation for color expression indicates that only a proportion of the total kernels are expressing the Navajo phenotype. Among the tropical breeding populations, the majority of the entries showed full expression (~56 %) or segregation (~40 %); only~ 4 % of the

populations analyzed showed complete color inhibition. In the landraces, complete color inhibition was observed in ~27 % of accessions and full expression in ~3 %. A large proportion of the landraces in this study (~70 %) exhibited segregation for the Navajo phenotype. Only 18 % of the inbred lines revealed segregation for color expression. More than 50 % showed full expression and a significant proportion of inbred lines (~30 %) showed complete color inhibition. Figure 2 presents the comparison of the Navajo marker expression in germplasm using cumulative proportion graphs for distribution of phenotypic scores for different aspects of *R1-nj* expression. Figure 2a shows that a higher proportion of landraces have inhibition scores between 2 and 4 and, therefore, a significantly higher level of segregation for color marker expression than inbred lines. Inbred lines had higher scores for endosperm area marked with the Navajo phenotype and color intensity than landraces and breeding populations.

Within the tropical germplasm, a higher proportion of inbreds (~49 %) developed from the highland breeding program of CIMMYT showed complete color inhibition than those developed from the subtropical (~30 %) or lowland (~28 %) breeding programs (Table 1). Moreover, the highland inbred lines had less color intensity and area marked with the Navajo phenotype than the subtropical or lowland lines (Fig. 2b). No significant differences were observed among topical and subtropical inbred lines with regard to these attributes. The yellow kernel inbreds had a higher proportion of color inhibition and lesser segregation for anthocyanin color expression than the white kernel inbreds (Table 1; Fig. 2c). There was no difference in area marked among yellow and white kernel inbreds, although anthocyanin color marker expression was less intense in the yellow kernel backgrounds (Fig. 2c). There were no significant differences between the CIMMYT dent and flint inbred lines with regard to color inhibition and the area marked (Table 1; Fig. 2d). In general, a higher proportion of flint inbreds had less color intensity than the dent inbreds (Fig. 2d).

Genome-wide association mapping

GWAS based on GLM and MLM for the different traits viz., endosperm color expression, intensity of color on the endosperm, area marked on the endosperm, embryo color expression, and intensity of color on the embryo, identified an important region on chromosome 9, which was conspicuous in the Manhattan plots (Fig. 3 and Supplementary Fig. 1) generated using ~410 K genome-wide markers. The SNP with the most significant *P* value (S9_9741377; 3.55E−42) was located within the *C1* gene, which has been well established to play a regulatory role in anthocyanin formation in the aleurone layer of the endosperm (Cone et al. 1986). This SNP was the most significant for all traits measured and explained approximately 30 % of the phenotypic variance for each. None of the other loci individually explained more than 3 % of the variance. A set of four genomic regions has been identified for endosperm color inhibition and is summarized in Table 2.

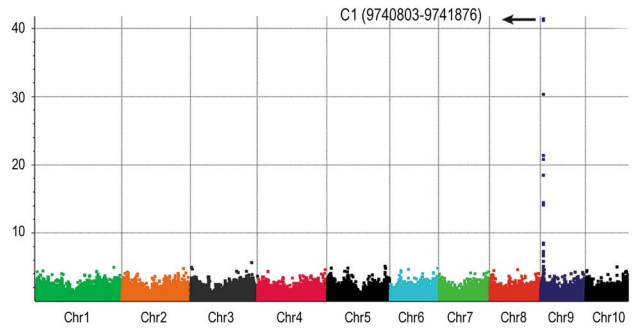

Fig. 3 Genome-wide association study (GWAS) based on ~410 K GBS-SNPs in 897 lines for endosperm inhibition established *C1* as the major gene influencing anthocyanin color expression. The most significant SNP (S9_9741377) on chromosome 9 was localized within the *C1* locus

Molecular markers for predicting kernel color inhibition

Of the 714 inbred lines analyzed, 209 lines showed complete inhibition of the Navajo phenotype. For the remaining 505 lines, 365 showed complete kernel color expression and 140 lines had scores ranging from 2 to 4, indicating incomplete expression of kernel color, which could be attributed to genetic and/or environmental factors that influence *R1-nj* color expression. The 505 inbred lines with scores ranging from 1 to 4 were determined to be in the 'expressing' group with regard to the trait under study, and the 209 lines with complete color inhibition in the 'non-expressing' group. The *C1* gene-based InDel marker developed in this study was scored in all 714 lines with the KASPar assay. The 8-bp insertion, diagnostic of the *C1-I* allele (Paz-Ares et al. 1990), was observed in 146 lines and the remaining 568 lines showed a non-*C1-I* allele (Table 3a). The most significant SNP identified from the GWAS was located in the intronic region of the *C1* gene that showed two alleles. Homozygotes and heterozygotes were scored for this marker. The SNP in homozygous state from the *C1-I* allele was observed in 246 lines (Table 3a), and the homozygotes for the non-*C1-I* allele was observed in 390 lines. Seventy-eight lines had this SNP in heterozygous state, and 55 showed color expression and were classified as homozygous non-*C1-I* allele in further analyses. The Fisher's Exact Tests for all marker-trait classifications were highly significant with P values of 2.79E−049 for *C1-I* insertion marker classification of the trait, 3.58E−047 for

Table 2 Significant genomic regions identified through genome-wide association study (GWAS) influencing the inhibition of the *R1-nj* marker system in the inbred lines analyzed

SNP	Minor allele frequency (MAF)	Mixed linear model (MLM)-*P*	Phenotypic average		R^2 (%)
			Minor allele	Major allele	
S9_9741377	0.40	3.55E-42	3.71	1.81	26.9
S5_200435045	0.31	6.29E-06	2.04	2.74	3.4
S10_108358852	0.25	8.08E-06	2.04	2.67	2.8
S1_272352634	0.35	9.75E-06	2.31	2.72	1.9

Table 3 Trait classification of inbred lines vis-à-vis (a) *C1-I*-specific InDel, (b) *C1-I* SNP and (c) *C1-I* InDel and SNP combination

	Color inhibition	Color expression	Total
(a)			
C1-I ins+	120 TP	26 FP	146
C1-I ins−	89 FN	479 TN	568
Total	209	505	714
(b)			
C1-I SNP+	154 TP	92 FP	246
C1-I SNP−	55 FN	413 TN	468
Total	209	505	714
(c)			
C1-I ins+/C1-I SNP+	172 TP	97 FP	269
C1-I ins-/C1-I SNP−	37 FN	408 TN	445
Total	209	505	714

TP true positives, *FP* false positives, *FN* false negatives, *TN* true negatives

C1-I SNP-based classification of the trait and 2.41E−064 for a combination of these two markers used for classifying the trait.

We also analyzed the efficiency of using *C1-I* allele-specific diagnostic markers by estimating their "sensitivity" and "specificity". The "sensitivity" of the test, if the diagnostic markers were used as a screening test for color inhibition, is indicated by the true positive rate (TPR), which describes the proportion of total lines with color inhibition that could be identified with the marker tested. TPR was 0.57 for the *C1-I*-specific 8 bp InDel, 0.74 for the *C1-I*-specific SNP, and 0.82 for a combination of these two markers (Table 4). Of the 209 color-inhibited lines, 120 could be detected with the *C1-I*-specific InDel, 154 with the *C1-I*-specific SNP, and 172 with a combination of these markers. The "specificity" of the test, if the diagnostic markers were used as a screening test for color expression, is measured by the true negative rate (TNR), which describes the proportion of all lines with color expression carrying the non-*C1-I* allele-specific marker. TNR was 0.95, 0.82, and 0.81 for the *C1-I* InDel, *C1-I*-specific SNP, and *C1-I* InDel-SNP combination, respectively. Although the *C1-I* InDel detects a large proportion of true negatives, which are lines that will express color, its ability to detect true positives is limited. Of the 209 lines that showed complete color inhibition, this marker picked up only 57 %, whereas the other two classifications detected approximately 74–82 % of this category. The positive and negative prediction values simulate practical situations where genotypes have to be predicted for the presence or absence of the dominant color inhibitor based solely on marker data. The positive predictive value (PPV) shows the efficacy of the prediction based on diagnostic markers to identify lines/genotypes with possible color inhibition, and the negative predictive value (NPV) shows the efficacy of the prediction to identify lines/genotypes with potential anthocyanin color expression when crossed with Navajo marker-based haploid inducer. In tropical lowland and subtropical maize inbred lines phenotyped in this study for endosperm color inhibition, approximately 30 % of the lines showed color inhibition (Table 1), and hence a trait prevalence of 0.30 was used in estimating PPV and NPV. PPV for the classification based on the *C1-I* InDel was 0.83, 0.64 for the *C1-I*-specific SNP and 0.65 for the combination of InDel and SNP. Of the 146 lines with *C1-I* InDel, 26

Table 4 Statistics associated with 2 × 2 contingency table of *C1-I* marker-based classification against the "gold standard"

	TPR	TNR	PPV	NPV	Accuracy
C1-I InDel	0.57 (0.50–0.64)	0.95 (0.93–0.97)	0.83 (0.75–0.88)	0.84 (0.81–0.87)	0.84 (0.80–0.87)
C1-I SNP	0.74 (0.67–0.80)	0.82 (0.78–0.85)	0.64 (0.56–0.69)	0.88 (0.85–0.91)	0.79 (0.73–0.85)
C1-I InDel/*C1-I* SNP	0.82 (0.76–0.87)	0.81 (0.77–0.84)	0.65 (0.58–0.70)	0.92 (0.89–0.94)	0.81 (0.79–0.85)

TPR true positive rate, *TNR* true negative rate, *PPV* positive predictive value, *NPV* negative predictive value

lines were false positives, whereas 92 of the 246 lines with *C1-I* SNP and 97 of 249 with both InDel and SNP were false positives. NPV for the classification based on the *C1-I* InDel was 0.84, 0.88 for the *C1-I* SNP, and 0.92 for the InDel-SNP combination. All three classifications have good predictive values for lines with color expression. The accuracy of the three classifications was 0.84, 0.79, and 0.81 for the *C1-I* InDel, *C1-I* SNP, and InDel-SNP combination, respectively.

Discussion

The *R1-nj* marker system is widely used for haploid identification as it enables visual differentiation of putative haploids from the diploids during the DH development process in maize (Greenblatt and Bock 1967). Inhibition or poor expression of the *R1-nj* marker can limit the use of DH technology in maize breeding programs with the presently available haploid inducers that do not have alternative marker systems. It is also possible to have considerable numbers of misclassified kernels due to incomplete expressivity of the Navajo phenotype when a source population is crossed with a *R1-nj*-based haploid inducer (Kebede et al. 2011). Our study showed that inhibition of the Navajo phenotype could be frequent (up to 25–30 %) in CIMMYT-derived maize inbred lines and landrace accessions from the CIMMYT Gene Bank, and there is considerable variation in the expression of the Navajo phenotype in different genetic backgrounds (Fig. 1, Supplementary Tables 1, 2).

As the DH technology is most commonly used for deriving homozygous lines from source populations developed by crossing elite inbred lines, it is encouraging to know that only 4 % of the populations exhibited complete color inhibition in the present study. However, more than 40 % of populations showed segregation for *R1-nj* expression indicating that it not possible to identify all the haploids induced in these populations. Segregation for the Navajo phenotype in a breeding population is possible when one or more inbred lines used in deriving the population possess color inhibitor gene/s. The *R1-nj* color marker system can, therefore, be inefficient for deriving DH lines from the tropical source populations, unless the source populations are carefully constituted using prior information about the absence of color inhibitors in the parental lines.

This study included a large number of tropical inbred lines developed at CIMMYT over the past five decades, including some of the advanced elite lines that are not released as CMLs. The majority of the inbred lines (49.33 %) showed full expression of the Navajo phenotype in a cross with a *R1-nj*-based inducer. However, 29.24 % of the inbred lines showed complete color inhibition which restricts their utility in DH-based breeding using the *R1-nj* marker system for haploid induction. The study also identified segregation of Navajo expression in 18 % of the tropical inbreds, which could be attributed to various factors including segregation for the color inhibitor (*C1-I*) and wild-type (*C1*) alleles or the early-generation nature of some inbred lines. Such inbred lines can be further purified or made completely homozygous using DH technology, with selection for the wild-type *C1* allele by discarding derivatives with the *C1-I* allele to enable anthocyanin expression in the lines. Among the inbred lines developed for diverse agro-ecologies, the highland inbred lines had a relatively higher frequency of color inhibition possibly due to a combination of factors including smaller sample size and a relatively narrow genetic base compared to tropical and subtropical lines. There was no significant difference between the tropical dent and flint inbreds in terms of color marker inhibition. The tropical yellow germplasm appeared to have a higher frequency of color inhibitor genes than the white maize germplasm (Table 1).

DH technology has been proposed as a valuable tool for purging deleterious alleles from maize landraces and deriving inbred lines that can be potentially used to broaden the genetic base of adapted germplasm (Strigens et al. 2013). In addition, DH technology is a cheaper and faster option for production of homozygous lines from landraces than the traditional selfing approach (Wilde et al. 2010). However, the study clearly demonstrated that the present *R1-nj*-based marker system may not be efficient for deriving DH lines from tropical maize landraces, as 28 % of landraces showed complete color inhibition and close to 70 % showed segregation for color marker expression. This is in addition to the difficulties associated with the presence of purple or red pericarp color in some landraces.

In addition to color inhibition, the area marked by the Navajo phenotype and intensity of anthocyanin color on the endosperm and embryo affect the efficiency of haploid identification using the *R1-nj* marker system. These traits are influenced by the genetic background of the source germplasm. For haploid identification, expression of the Navajo phenotype on the crown region of the endosperm and the scutellum area of the embryo with good intensity is important. In some genotypes, the whole endosperm area expresses the Navajo phenotype, whereas in other genotypes only the silk attachment region is marked. If such limited expression is combined with poor intensity of color on the endosperm or embryo (scutellum) this may result in high rates of misclassification (Prigge et al. 2011). Among the breeding populations, landraces, and inbred lines, the proportions of breeding populations exhibiting higher intensity of color and more area marked by the Navajo phenotype were higher than those of the landraces and inbred lines. Prigge et al. (2011) recorded higher rates of misclassification in induction crosses of landraces than

OPVs and elite single cross hybrids due to poor expression of the Navajo phenotype. There was no difference between yellow and white inbreds for area marked, although yellow kernel inbreds showed a relatively lower intensity of anthocyanin expression than the white inbreds. Combining the higher inhibition rates and lower intensity of color, haploid identification in germplasm with yellow kernels could be more challenging than for the white kernel inbreds. The dent inbreds were significantly different to the flint inbreds with flint inbreds tending to have poor intensity of color. A previous study using temperate flint and dent elite single crosses and temperate inducers did not find significant differences in the intensity of endosperm and embryo coloration but indicated higher misclassification rates in flint germplasm than in dent germplasm (Rober et al. 2005). In another study, misclassification rates were similar in tropical dent and flint genotypes (Prigge et al. 2011). Together these results indicate significant variation in the expression of the Navajo phenotype in different types of germplam when crossed with a *R1-nj*-based haploid inducer that could affect the efficiency of haploid identification. Due to the high frequency of complete color inhibition in tropical maize germplasm (~28–30 %) and poor expressivity of the Navajo phenotype in many tropical genetic backgrounds, it is necessary to develop and deploy alternative marker systems that can effectively replace or complement the *R1-nj* marker system to efficiently identify haploids. Rotarenco et al. (2010) developed inducer lines combining the *R1-nj* marker system with other anthocyanin genes *B1* and *Pl1,* which enable anthocyanin accumulation in seedling tissues like roots, to facilitate identification of haploids at the seedling stage when *R1-nj* expression is inhibited in the kernels. This system could aid in identification of false positives prior to chromosomal doubling treatments, thus saving costs. Another method of haploid identification based on kernel oil content and amenable to automation was proposed to overcome the problem of anthocyanin color inhibition (Rotarenco et al. 2007; Li et al. 2009; Melchinger et al. 2013, 2014). Also, the in vitro methods of haploid production based on anther or pollen culture do not require anthocyanin phenotypic markers; however, these methods were of limited usefulness in maize as they were noted to be inefficient, genotype dependent and technically challenging (Weber 2014). While better marker systems for in vivo haploid identification are being developed, it would be useful to understand the presence/absence of specific genetic factors responsible for inhibition of anthocyanin color expression in progenies derived by crossing the available *R1-nj*-based haploid inducer with source germplasm of interest.

GWAS to identify genome-wide genetic factors that regulate the expression of the Navajo phenotype revealed that the *C1* gene is the most significant genetic factor affecting *R1-nj* expression/inhibition (Fig. 3; Table 2), intensity, and area marked by the Navajo phenotype (Supplementary Fig. 1). This study did not identify genomic regions associated with *C2* (located on chromosome 4: 192.5 Mb) or the *In1* gene (located on chromosome 7: 19.35 Mb). *C2-IDF* was initially identified in two Peruvian strains and is very rare among the maize stock studied (Brink and Greenblatt 1954). It is possible that *C2-IDF* could be absent or occurs at very low frequency in the germplasm used in this study. Stinard and Sachs (2002) reported that the inhibitor allele *In-D* did not cause inhibition of the anthocyanin accumulation resulting from the *R1-nj* allele.

Considering the GWAS results and high frequency of *R1-nj* inhibition in tropical germplasm, the validity of the markers specific to the *C1-I* allele against the trait phenotype was tested. The observations from trait classifications of the inbred lines based on the allele-specific markers (*C1-I* InDel alone, *C1-I* SNP alone, and the combination of these two markers) and endosperm inhibition showed high accuracy ranging from 0.79 to 0.84. Inaccuracies using these marker assays can be attributed to the following factors: (1) although *C1-I* is a dominant inhibitor, in endosperm with genetic constitution *C1-I/C1-I/C*1 and *C1-I/C1/C1,* infrequent small patches of anthocyanin is expressed (Coe 1962; Coe et al. 1988). This may result in a lack of correlation between phenotype and genotype. When the sequence of *C1-I* allele with the wild type *C1* allele, eighteen positions were found altered in *C1-I*, including transitions, transversions, deletions and insertions (Paz-Ares et al. 1990). Although the 8-bp InDel polymorphism is putatively implicated in the functional change of *C1* from activator to dominant repressor (Paz-Ares et al. 1990), other changes in the sequence may also bring complete inhibition even in the absence of the putative functional InDel polymorphism or may alter *C1-I* inhibition capability. Frame shifts in the amino-terminal basic or carboxy terminal acidic region of *C1* were shown to generate dominant inhibitory protein in transient expression assays (Dooner et al. 1991; Goff et al. 1991). It may be possible that the InDel and SNP polymorphism used in our study may not be representative of all the functional changes from *C1* to *C1-I* and (3) there may be other inhibitory alleles affecting *R1-nj* expression other than *C1-I*, although present at low frequency in the germplasm, which could not be identified in this GWAS study. *C1-I* InDel and *C1-I* SNP marker combination had the highest sensitivity to correctly classify genotypes with kernel color inhibition, while *C1-I* InDel marker had high specificity that could by itself provide an effective means to correctly classify genotypes with potential for color expression. The practical value of the marker assay is to make a correct prediction of the trait without going through the laborious process of phenotypic data generation. The predictive power of these marker classifications,

determined using statistics PPV and NPV, was moderate for color inhibition and high for color expression. Hence, *C1-I* based InDel and SNP combination could potentially aid in accurate selection of genotypes that express color when crossed with *R1-nj*-based inducer. If breeders could effectively identify lines that do not have the *C1-I* allele, it would assist in developing/identifying selected pedigree crosses amenable for haploid induction using *R1-nj*-based haploid inducer. Other pedigree-based crosses which have the color inhibitor in the genetic background could be possibly advanced through a conventional selfing and selection scheme, until next-generation haploid inducers with better marker systems are developed and accessible in the public domain.

In conclusion, the *R1-nj* marker system is not expressed in a large proportion (~25–30 %) of tropical germplasm because of the presence of the dominant color inhibitor *C1-I*. KASPar assays developed in this study, using *C1-I* allele-specific diagnostic markers, identified germplasm that could express color with reasonably high accuracy. As demonstrated, these assays can assist in large-scale assaying of expression of *R1-nj* in a wider collection of germplasm across diverse genetic backgrounds, thereby optimizing the time and costs involved in DH line development using the *R1-nj*-based haploid inducers.

Author contributions PMB, VC, SN, and RB designed the experiments. VC, SN, RB, and LM performed the experiments. VC, SN, RB, and JT analyzed the data. PMB, VC, SN, and RB contributed materials/analysis tools. VC, SN, RB, and PMB wrote the manuscript.

Acknowledgments The authors gratefully acknowledge the financial support received from the CGIAR Research Program (CRP) MAIZE, Limagrain, and the MASAGRO project funded by the Secretariat of Agriculture, Livestock, Rural Development, Fisheries and Food (SAGARPA). The *R1-nj*-based tropical haploid inducers used in this study were the products of collaborative research efforts between CIMMYT and the University of Hohenheim, Germany. The assistance provided by Juan Burgueno for data analysis, and Miguel Mellado for the preparation of figures, is gratefully acknowledged.

Conflict of interest The authors declare that no competing interests exist.

Ethical standards The authors declare that the experiments comply with the laws of Mexico.

References

Altman DG, Bland JM (1994a) Diagnostic tests 2: predictive values. Br Med J 309:102

Altman DG, Bland JM (1994b) Diagnostic tests 3: receiver operating characteristic plots. Br Med J 309:188

Battistelli GM, Von Pinho RG, Justus A et al (2013) Production and identification of doubled haploids in tropical maize. Genet Mol Res 12:4230–4242

Brink RA, Greenblatt IM (1954) Diffuse, a pattern gene in Zea Mays epistatic to red pericarp. J Hered 45:47–50

Burr FA, Burr B, Scheffler BE et al (1996) The maize repressor-like gene *intensifier1* shares homology with the *r1/b1* multigene family of transcription factors and exhibits missplicing. Plant Cell 8:1249–1259

Chaikam V, Prasanna BM (2012) Maternal haploid detection using anthocyanin markers. In: Prasanna BM, Chaikam V, Mahuku G (eds) Doubled haploid technoogy in maize breeding: theory and practice. CIMMYT, Mexico, pp 20–23

Chaikam V, Mahuku G, Prasanna BM (2012) Design and implementation of maternal haploid induction. In: Prasanna BM, Chaikam V, Mahuku G (eds) Doubled haploid technoogy in maize breeding: theory and practice. CIMMYT, Mexico, pp 14–19

Chase SS (1952) Monoploids in maize. Iowa State College Press, Ames, pp 389–399

Chase SS (1969) Monoploids and monoploid-derivatives of maize (*Zea mays* L.). Bot Rev 35:117–168

Chase SS, Nanda DK (1965) Screening for monoploids of maize by use of a purple embryo marker. Maize Gen Coop Newslett 39:59–60

Chen S, Li LH (2009) Maize doubled haploid breeding (in Chinese). China Agricultural University Press, Beijing

CIMMYT (2005) Laboratory protocols: CIMMYT applied molecular genetics laboratory, 3rd edn. CIMMYT, Mexico

Coe EH Jr (1962) Spontaneous mutation of the aleurone color inhibitor in maize. Genetics 47:779

Coe EH, McCormick SM, Modena SA (1981) White pollen in maize. J Hered 72:318–320

Coe EH Jr, Nueffer MG, Hoisington DA (1988) The genetics of corn. In: Sprague GF, Dudley JW (eds) Corn andcorn improvement, 3rd edn. Madison, Wisconsin, pp 81–258

Cone KC, Burr FA, Burr B (1986) Molecular analysis of the maize anthocyanin regulatory locus *C1*. Proc Natl Acad Sci 83:9631–9635

Della Vedova CB, Lorbiecke R, Kirsch H et al (2005) The dominant inhibitory chalcone synthase allele *C2-Idf* (*Inhibitor diffuse*) from *Zea mays* L. acts via an endogenous RNA silencing mechanism. Genetics 170:1989–2002

Dooner HK, Robbins TP, Jorgensen RA (1991) Genetic and developmental control of anthocyanin biosynthesis. Annu Rev Genet 25:173–199

Elshire RJ, Glaubitz JC, Sun Q et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6:e19379

Fisher RA (1922) On the interpretation of χ^2 from contingency tables, and the calculation of P. J R Stat Soc 85:87–94

Flint-Garcia SA, Thuillet A-C, Yu J et al (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44:1054–1064

Forster BP, Heberle-Bors E, Kasha KJ, Touraev A (2007) The resurgence of haploids in higher plants. Trends Plant Sci 12:368–375

Franken P, Niesbach-Klösgen U, Weydemann U et al (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 10:2605

Geiger HH (2009) Doubled haploids. Handbook of maize. Springer, New York, pp 641–657

Geiger HH, Gordillo GA (2009) Doubled haploids in hybrid maize breeding. Maydica 54:485–489

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 5:298–309

Greenblatt IM, Bock M (1967) A commercially desirable procedure for detection of monoploids in maize. J Hered 58:9–13

Kang HM, Sul JH, Service SK et al (2010) Variance component model to account for sample structure in genome-wide association studies. Nat Genet 42:348–354

Kebede AZ, Dhillon BS, Schipprack W et al (2011) Effect of source germplasm and season on the in vivo haploid induction rate in tropical maize. Euphytica 180:219–226

Li L, Xu X, Jin W, Chen S (2009) Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. Planta 230:367–376

Lübberstedt T, Frei UK (2012) Application of doubled haploids for target gene fixation in backcross programmes of maize. Plant Breed 131:449–452

Melchinger AE, Schipprack W, Würschum T et al (2013) Rapid and accurate identification of in vivo-induced haploid seeds based on oil content in maize. Sci Rep 3:2129

Melchinger AE, Schipprack W, Friedrich Utz H, Mirdita V (2014) In vivo haploid induction in maize: identification of haploid seeds by their oil content. Crop Sci 54:1497–1504

Nanda DK, Chase SS (1966) An embryo marker for detecting monoploids of maize (*Zea mays* L.). Crop Sci 6:213–215

Paz-Ares J, Ghosal D, Saedler H (1990) Molecular analysis of the C1-I allele from *Zea mays*: a dominant mutant of the regulatory C1 locus. EMBO J 9:315

Prasanna BM (2012) Doubled haploid technology in maize breeding:an overview. In: Prasanna BM, Chaikam V, Mahuku G (eds) Doubled haploid technoogy in maize breeding: theory and practice. CIMMYT, Mexico, pp 1–8

Prasanna BM, Chaikam V, Mahuku G (2012) Doubled haploid technoogy in maize breeding: theory and practice. CIMMYT, Mexico

Prigge V, Sánchez C, Dhillon BS et al (2011) Doubled haploids in tropical maize: I. Effects of inducers and source germplasm on in vivo haploid induction rates. Crop Sci 51:1498–1506

Prigge V, Schipprack W, Mahuku G et al (2012) Development of in vivo haploid inducers for tropical maize breeding programs. Euphytica 185:481–490

Rober FK, Gordillo GA, Geiger HH (2005) *In vivo* haploid induction in maize-performance of new inducers and significance of doubled haploid lines in hybrid breeding. Maydica 50:275

Rotarenco VA, Kirtoca IH, Jacota AG (2007) The possibility of identifying kernels with haploid embryos using oil content. Maize Gen Coop NewsLett 81:11

Rotarenco VA, Dicu G, State D, Fuia S (2010) New inducers of maternal haploids in maize. Maize Genet Coop Newslett 84:15

Schmidt W (2004) Hybrid maize breeding at KWS Saat AG. Vereinigung der Pflanzenzüchter und Saatgutkaufleute Österreichs, St. Pölten, pp 1–6

Seitz G (2005) The use of doubled haploids in corn breeding. Proceedings 41st Annual Illinois Corn Breeders' School 2005. Urbana–Champaign, Illinois, pp 1–7

Stinard PS, Sachs MM (2002) The identification and characterization of two dominant *r1* haplotype-specific inhibitors of aleurone color in *Zea mays*. J Hered 93:421–428

Strigens A, Schipprack W, Reif JC, Melchinger AE (2013) Unlocking the genetic diversity of maize landraces with doubled haploids opens new avenues for breeding. PLoS One 8:e57234

Weber DF (2014) Today's use of haploids in corn plant breeding. In: Sparks DL (ed) Advances in agronomy. Academic press, Dublin, pp 123–144

Wilde K, Burger H, Prigge V et al (2010) Testcross performance of doubled-haploid lines developed from European flint maize landraces. Plant Breed 129:181–185

Yu J, Buckler ES (2006) Genetic association mapping and genome organization of maize. Curr Opin Biotechnol 17:155–160