Theor Appl Genet (2005) 111: 888–897
DOI 10.1007/s00122-005-0011-6

ORIGINAL PAPER

R. Babu · S. K. Nair · A. Kumar · S. Venkatesh
J. C. Sekhar · N. N. Singh · G. Srinivasan · H. S. Gupta

Two-generation marker-aided backcrossing for rapid conversion of normal maize lines to quality protein maize (QPM)

Received: 31 January 2005 / Accepted: 31 May 2005 / Published online: 21 July 2005

Abstract The low nutritive value of maize endosperm protein is genetically corrected in quality protein maize (QPM), which contains the *opaque2* gene along with numerous modifiers for kernel hardness. We report here a two generation marker-based backcross breeding program for incorporation of the *opaque2* gene along with phenotypic selection for kernel modification in the background of an early maturing normal maize inbred line, V25. Using the flanking marker distances from *opaque2* gene in the cross V25×CML176, optimum population size for the BC_2 generation was computed in such a way that at least one double recombinant could be obtained. Whole genome background selection in the BC_2 generation identified three plants with 93 to 96% recurrent parent genome content. The three BC_2F_2 families derived from marker identified BC_2 individuals were subjected to foreground selection and phenotypic selection for kernel modification. The tryptophan concentration in endosperm protein was significantly enhanced in all the three classes of kernel modification viz., less than 25%, 25–50% and more than 50% opaqueness. BC_2F_3 lines developed from the hard endosperm kernels were evaluated for desirable agronomic and biochemical traits in replicated trials and the best line was chosen to represent the QPM version of V25, with tryptophan concentration of 0.85% in protein. The integrated breeding strategy reported here can be applied to reduce genetic drag as well as the time involved in a conventional line conversion program, and would prove valuable in rapid development of specialty corn germplasm.

Communicated by R. Bernardo

R. Babu (✉) · S. K. Nair · A. Kumar · H. S. Gupta
Vivekananda Institute of Hill Agriculture (ICAR), Almora, Uttaranchal, 263601, India
E-mail: rbabu_icar@rediffmail.com
Tel.: +91-5962-230060
Fax: +91-5962-231539

S. Venkatesh · J. C. Sekhar
Directorate of Maize Research, Amberpet, Hyderabad, India

N. N. Singh
Directorate of Maize Research, Pusa, New Delhi, 110012, India

G. Srinivasan
International Center for Wheat and Maize Research (CIMMYT), Apdo Postal 6-641, 06600, Mexico DF, Mexico

Introduction

Maize (*Zea mays* L.) plays a very important role in human and animal nutrition. However, the normal maize protein is of poor nutritional quality due to a deficiency in two essential amino acids (lysine and tryptophan) and high leucine–isoleucine ratio. A breakthrough came in the 1960s, with the discovery of the enhanced nutritional quality of the maize mutant *opaque2* (Mertz et al. 1964). Isolation and characterization of the *opaque2* gene revealed that it encodes a transcriptional factor that regulates the expression of zein genes and a gene encoding a ribosomal inactivating protein (Schmidt et al. 1990; Lohmer et al. 1991; Bass et al. 1992). The *o2* gene significantly reduces the level of 22-kD alpha-zeins while increasing the content of non-zein proteins particularly, EF-1 alpha, which is positively correlated with lysine content in the endosperm (Habben et al. 1995). The protein quality of *opaque2* maize is 43% higher than that of common maize and 95% of the value of casein (Mertz 1992). Plant breeders throughout the world made vigorous efforts to incorporate *opaque2* into high yielding commercial cultivars but the numerous agronomic and processing problems associated with *opaque2* prevented its acceptance (Glover and Mertz 1987). The *opaque2* maize did not become popular with farmers as well as consumers mainly because of reduced grain yield, soft endosperm, chalky and dull kernel appearance and susceptibility to ear rots and stored grain pests.

The low nutritive value of normal endosperm and poor agronomic and keeping quality of *opaque2* maize is corrected in a genetically improved, hard endosperm quality protein maize (QPM). QPM is a genotype in which *opaque2* gene has been incorporated along with associated modifiers. QPM contains twice the amount of lysine and tryptophan as compared to normal maize endosperm. Also, it is essentially interchangeable with normal maize in both cultivation and agronomic characteristics. A wide array of QPM germplasm combining hard endosperm characters and high lysine has been developed by International Center for Maize and Wheat Research (CIMMYT), Mexico through conventional breeding (Vasal et al. 1993a). Genetic analysis of QPM, opaque2 and normal maize has revealed the existence of at least two additive modifier genes that significantly influence the endosperm modification in two populations viz., W64Ao2×pool 33 and pool 33×W22o2 (Lopes et al. 1995). One modifier locus was tightly linked with the gamma-zein coding sequences near the centromere of chromosome 7, while the other near the telomere of the 7L (Lopes et al. 1995).

Despite the nutritional superiority and improved agronomic performance of QPM, large-scale QPM cultivation and adoption is yet to gain significant momentum in many developing countries. QPM is likely to gain wider acceptance if hybrids are produced that have agronomic performance similar to normal hybrids and retain an enhanced nutritional quality. The QPM hybrids have several advantages over the open pollinated QPM varieties viz., higher yield potential comparable to the best normal hybrids, assured seed purity, more uniform and stable endosperm modification and less monitoring for ensuring protein quality in seed production. A location-specific QPM hybrid development strategy necessitates an efficient line conversion program for enhancing the protein quality of normal inbreds that are heterotic to each other and well adapted to different region-specific intensified cropping systems. Maize holds a unique position as food, feed and industrial crop in India, where it is currently cultivated in 7.4 million hectares with the annual production of 14.7 million tonnes (Annual Report 2004–05). Development of early maturing QPM hybrids will meet the demands of high cropping intensity as well as contribute significantly towards household nutritional security in many parts of India.

The *opaque2* gene is recessive and the modifiers are polygenic. Their introgression into elite inbreds is not a straight forward procedure because of three major reasons viz., (1) each conventional backcross generation needs to be selfed to identify the *opaque2* recessive gene and a minimum of six backcross generations are required to recover satisfactory levels of recurrent parent genome (2) in addition to maintaining the homozygous *opaque2* gene, multiple modifiers must be selected and (3) rigorous biochemical tests to ensure enhanced lysine and tryptophan levels in the selected materials in each breeding generation require enormous labor, time and material resources. Although conventional breeding procedures have been used to convert commercial lines to QPM forms, the procedure is tedious and time consuming. Rapid advances in genome research and molecular technology have led to the use of DNA marker-assisted selection which holds promise in enhancing selection efficiency and expediting the development of new cultivars with higher yield potential. (Ribaut and Hoisington 1998). While marker-assisted foreground selection (Tanksley 1983; Melchinger 1990) helps in identifying the gene of interest without extensive phenotypic assays, marker-assisted background selection (Young and Tanksley 1989; Hospital et al. 1992; Visscher et al. 1996; Frisch et al. 1999a, b) expedites significantly the rate of genetic gain/recovery of recurrent parent genome in a backcross breeding program. With the development and access to reliable PCR-based allele specific markers such as SSRs, and SNPs, marker assisted selection is becoming an attractive option for simply inherited traits (Babu et al. 2004)

The objective of this study was to examine the feasibility of combining high protein quality and hard endosperm characters in early maturing normal maize inbreds through a combination of marker aided and phenotypic selection techniques. We report here a rapid line conversion strategy with a two-generation backcross program that employs foreground selection for the *opaque2* gene in both the backcross generations, background selection at non-target loci only in the BC_2 generation, and phenotypic selection for kernel modification and other desirable agronomic traits in two subsequent selfed generations.

Materials and methods

Plant materials

A set of nine normal maize and five QPM inbred lines were analyzed for polymorphism with *opaque2* specific SSR markers. The normal inbred lines (CM212, CM145, V25, V335, V338, V340, V345, V346 and V348) targeted for conversion to QPM version are parental lines of extra-early, high yielding single cross hybrids of maize in India. Based on the parental polymorphism analysis three normal inbred lines viz., V25, CM212 and CM145 and two QPM donors viz., CML173 and CML176 were chosen for line conversion. In this study, we report conversion V25 using CML176 as QPM donor. QPM inbred lines obtained from CIMMYT, Mexico, viz., CML170, CML173, CML176, CML180 and CML184 were tested as donors for the *opaque2* gene and associated modifiers. These late maturing lines had tryptophan content ranging from 0.80 to 1.05% of total protein, where as normal inbred lines possessed tryptophan content from 0.38 to 0.49% of endosperm protein.

Target gene and marker assays

Three SSR markers, viz., *phi057*, *phi112* and *umc1066* located as internal repetitive elements within the *opaque2* gene, were used in initial polymorphism analysis with all the nine normal and five QPM inbred lines, for checking their feasibility to be used as foreground selection marker for the *opaque2* gene. A total of 200 SSR markers spanning all the bin locations in a maize SSR consensus map (http://www.maizegdb.org) were selected for background selection. Of the 200 markers, 77 were found to be polymorphic between V25 and CML176. About 2 g of young and healthy leaf tissue was ground into a fine powder in liquid nitrogen using a pre-chilled mortar-pestle. 50–100 mg of ground material was transferred into a 2 ml microcentrifuge tube containing 800 µl extraction buffer (Tris -HCL 75 mM, pH 8.0, EDTA 20 mM, NaCl 500 mM, SDS 1.2% and PVP 1% to be added after autoclaving in case of mature leaf tissue) and the contents were mixed gently by swirling action. After suspending the tissue in the buffer for 5–10 min, 500 µl of ice-cold chloroform: iso-amyl alcohol (24:1) mixture was added, mixed well with the suspended tissue and spun at 10,000 rpm for 1 min preferably at 4°C. Without disturbing the interface, the supernatant was transferred to another 2 ml microcentrifuge tube and 1 ml absolute alcohol was added to the supernatant and mixed gently. The tube was spun for 3 min in a microcentrifuge with maximum speed and the supernatant was decanted. DNA was suspended in 200 µl of 1XTE (Tris-HCl 10 mM, pH 8.0, EDTA 1 mM) buffer containing 10 mg/ml of RNase A and then stored at −20°C for long term storage or used directly for PCR amplification reactions. PCR cycling consisted of initial denaturation at 94°C for 2 min, followed by 30–35 cycles of amplification at 94°C for 1 min, 55–65°C for 1 min, and 72°C for 2 min. A final extension step at 72°C for 7 min was followed by termination of the cycle at 4°C. The amplified products (15 µl) were resolved on a 3.5% SFR (Amresco) agarose gel at 125 V according to Senior et al. (1998).

Linkage map construction

Of the 23 SSR markers (spanning the chromosome 7) analyzed for polymorphism between V25 and CML176, 13 were found to be polymorphic. Genotyping was carried out on 148 BC_1 individuals of the cross V25× (V25×CML176). Standard χ^2 test was used to test the segregation pattern at each marker locus from the expected Mendelian segregation ratio of 1:1 for a backcross population. The map distances and order of the markers on the linkage group was calculated using the BC_1 algorithm of MAPMAKER/EXP 3.0 (Lincoln et al. 1992). The maximum likely map order for the markers was determined with a LOD score threshold of 3.0, and map distances were calculated using Haldane's map function. The order of the markers on the linkage group was compared with the SSR consensus map in MaizeGDB (http://www.maizegdb.org).

Selection procedure

The F_1 was made using the recurrent parent, V25 as female and the donor line CML176 as the pollen parent (Fig. 1). The first backcross was made with the F_1 hybrid as the female and 178 BC_1 seeds were planted.

BC_1 selection

Selection in BC_1 individuals was carried out on the basis of two criteria, (1) select heterozygotes for the *opaque2* gene specific SSR marker and (2) out of the selected individuals, identify individuals that were single or double recombinants for either or both of the flanking markers. Seven individuals were selected as per these criteria, of which three were single recombinants with the closest flanking marker, *bnlg1200*.

BC_2 selection

The minimum population size required in BC_2 generation to identify with 99% probability at least one plant which is a double recombinant with heterozygosity at target locus and homozygosity for recurrent parent alleles at flanking marker loci was calculated using the program Popmin (Hospital and Decoux 2002). Accordingly, 106 seeds from each of the three selected single recombinants were sown and later thinned to 53 individuals from each selected BC_1F_1 plant to maintain a population of 159 plants. Selection in these individuals was carried out based on the criteria that they were heterozygous at the marker locus linked to the target gene and homozygous for recurrent parent alleles at the two flanking loci. At this stage, whole genome selection employing 72 markers was done on 74 individuals heterozygous at the *opaque2* locus to identify individuals with the highest amount of the recurrent parent genome. This analysis was carried out using the software program GGT (Van Berloo 1999). Three plants were identified based on these criteria and were self-pollinated at flowering.

BC_2F_2 selection

Seeds from the self-pollinated ears of three selected BC_2F_1 individuals with highest proportion of recurrent parent genome were planted for the next round of selection. In this case, the DNA extracted from the leaf samples was assayed for the *opaque2* specific marker and only those plants that were recessive homozygous for *opaque2* were selected and tagged. These plants were evaluated for days to 50% silking, plant vigor, resistance

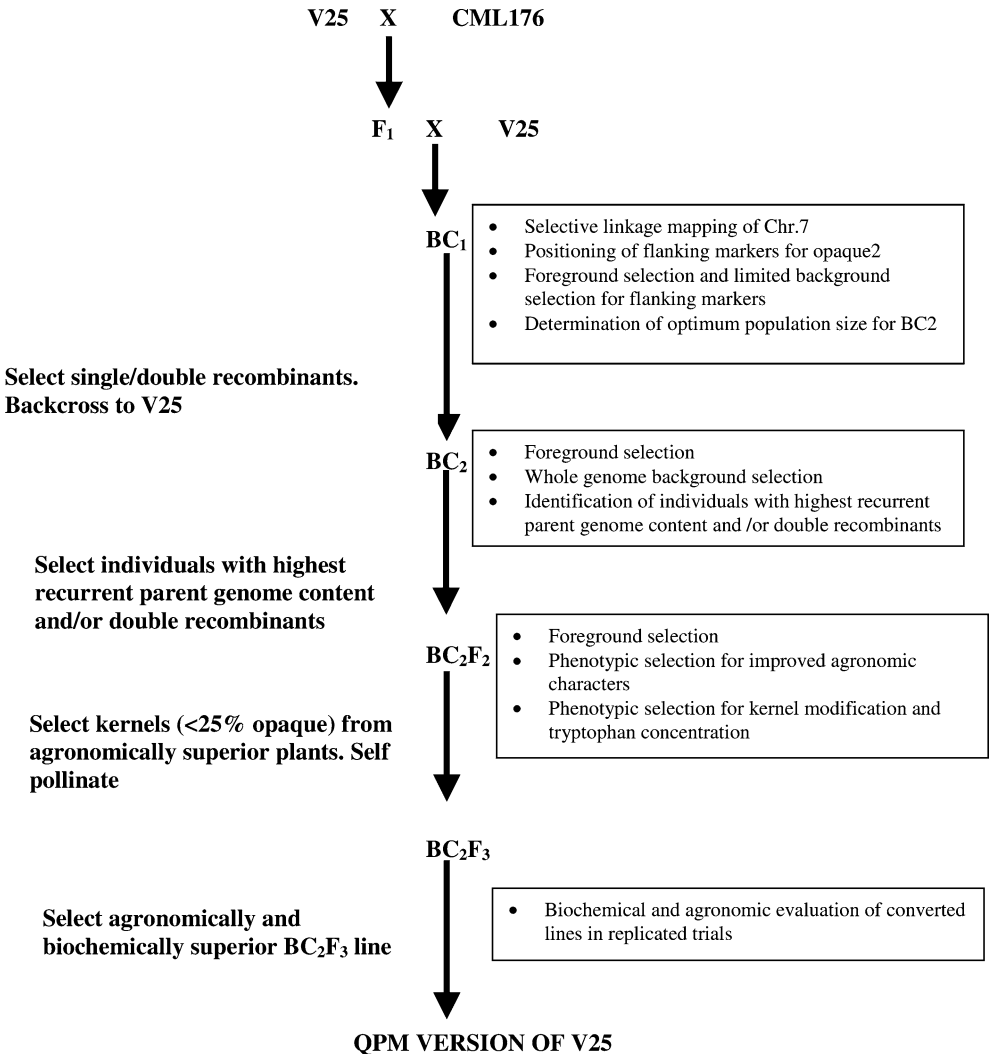

Fig. 1 Schematic representation of conversion of normal maize inbred line, V25 to its QPM version employing the QPM donor, CML176

to *Exserohilum turcicum* and *Bipolaris maydis* for elimination of undesirable individuals at Hawalbagh Farm (India) in 2003. All the selected plants were self-pollinated and upon harvest the kernels were subjected to selection for kernel modification under transmitted light using a light box. The kernels of each selected self-pollinated BC_2F_2 ear were grouped under three classes of modification viz., less than 25%, 25–50% and more than 50% opaque. The biochemical analyses for total protein content in the endosperm and tryptophan concentration in endosperm protein in each class of kernel modification were carried out according to standard procedures developed by Villegas (1975) and Villegas et al. (1984). In short, the grain samples were de-germed after removing the pericarp and finely ground. The resulting flour was defatted and total nitrogen content was determined by Microjeldahl procedure (AOAC 1965) and percentage of protein was calculated by multiplying the N content with a factor of 6.25. The tryptophan concentration in endosperm protein was estimated by the colorimetric method of Hernandez and Bates (1969). All the biochemical analyses were performed in duplicate samples. Kernel modification was assessed subjectively and grains were grouped as hard, semi-soft and soft. Hard endosperm kernels with less than 25% opaqueness were selected for raising further BC_2F_3 generation.

BC_2F_3 selection

The 16 selected BC_2F_2 entries [six from P-7 (P-7-A to P-7-F), five from P-56 (P-56-A to P-56-E) and five from P-73 (P-73-A to P-73-E)] were evaluated in an alpha lattice design with two replications along with recurrent and donor parents as checks in Hawalbagh Farm, VPKAS (ICAR), Almora, India in summer 2003 Two-row plots, each 5 m long and spaced 75 cm apart were used. One row plots were used wherever sufficient seeds were not available for specific entries. All the plants in BC_2F_3 were self-pollinated. The trial was hand harvested and grain yield was estimated at 80% of ear weight and adjusted to 15% moisture. The following agronomic and biochemical traits were recorded in the $BC2F_3$ trial: plant height (cm), ear height (cm), days to 50% silking,

days to 50% pollen shed, resistance to *E. turcicum*, and *B. maydis*, grain yield (t/ha), total protein content (%) and tryptophan concentration in endosperm protein (%).

Results

Parental polymorphism analysis between normal and QPM inbred lines

The SSR marker, *phi112* exhibited dominant polymorphism between all nine normal and five QPM inbreds. Absence of approximately 150 bp DNA fragment in the PCR amplified product of QPM inbreds clearly distinguished them from the normal inbreds. The presence of this band in one of the QPM inbred viz., CML 184 indicated possible pollen contamination from the normal maize during seed maintenance. The markers viz., *phi057* and *umc1066* exhibited co-dominant polymorphism between normal and QPM inbreds. *phi057* amplified around 160 bp fragment in all the nine normal inbreds and 170 bp fragment in three QPM inbred lines viz., CML 170, CML173 and CML189. *umc1066* amplified 150 bp fragment in all the five QPM inbreds and 160–170 bp fragments in six normal inbred lines (Fig. 2a)

Linkage mapping of chromosome 7

The map of chromosome 7 spanned a distance of 133.2 cM with an average density of 10.3 cM. The marker order was consistent with maize SSR consensus map (http://www.maizegdb.org). The two flanking markers viz., *bnlg2160* and *bnlg1200* on chromosome 7 were 4.2 and 3.8 cM from the *opaque2* locus in this particular mapping/MAS population

Foreground selection in backcross and selfed generations

The SSR marker *umc1066* identified 71 heterozygous individuals in a population of 148 plants in BC_1 generation, 74 in 159 plants of BC_2 generation and 53 recessive homozygotes in 225 plants of BC_2F_2 generation (Fig. 2b). The chi-square test for goodness of fit showed that the marker segregated according to the expected Mendelian ratio of 1:1 for a BC1 generation and 1:2:1 for an F_2 generation. Seven out of 71 plants heterozygous for the *opaque2* locus were single recombinants on either side of *opaque2*, while no double recombinants were identified. Out of the seven single recombinants, three were recombinants between *opaque2* gene and *bnlg1200* and four were between *opaque2* gene and *bnlg2160*.

Determination of optimum population size for MAS

Given the various input parameters in Popmin viz., d1 (distance of the target gene from the closest flanking marker, *bnlg1200*) as 3.8 cM, d2 (distance of the target gene from the distant flanking marker, *bnlg2160*) as 4.2 cM, maximum number of BC generations (t_{max}) as 2, risk level as 0.01, probability of transmission of the target gene to a BC progeny as 0.5, and indicating the initial marker genotype (single recombinant for the closest marker, *bnlg1200*) selected in the BC_1 generation, the suggested optimum population size for BC_2 generation was 157 with the cumulated probability of 0.99 for success over two generations (Table 1).

Background selection in BC_2 generation

The average recurrent parent genome content of the BC_2 generation was 78.4% while, the donor genome content

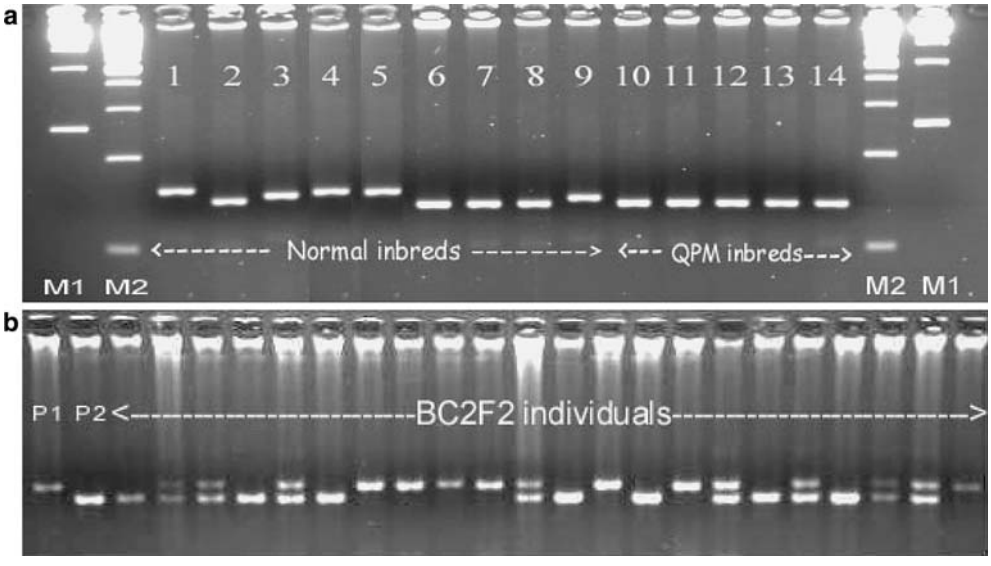

Fig. 2 **a** Parental polymorphism analysis using *opaque2* specific SSR marker, *umc1066* between normal and QPM inbreds. Lane M1: 1 kb marker, Lane M2: 100 bp marker, Lanes1: CM212, 2: CM145, 3: V25, 4: V335, 5: V338, 6: V340, 7: V345, 8: V346, 9: V348, 10: CML173, 11: CML176, 12: CML180, 13: CML184 and 14: CML189. **b** Foreground selection for identification of recessive homozygotes for *opaque2* using *umc1066* in the BC_2F_2 population of V25×CML176. P1: V25, P2: CML176, Lanes 1–14: BC_2F_2 individuals

Table 1 Optimum population size computation based on flanking marker distances using POPMIN program

Generation	Population size	Cumulative population size	Probability of success	Cumulative probability of success	Actual population size used
BC1	70	70	0.759	0.759	148
BC2	157	227	0.231	0.990	159

Input parameters for POPMIN program: Flanking marker, d1 = 3.8 cM, Flanking marker, d2 = 4.2 cM, Number of BC generations, $t_{max} = 2$, risk factor, $r = .01$, Probability of transmission of the target gene in BC generation, $P = 0.5$ and Initial marker genotype = 1 (single recombinant for the closest flanking marker)

was 9.9%. The proportion of unexplained variation (either due to unknown allele type or missing data) was 11.7%. A total of 14 BC_2 progenies had more than average recurrent parent genome content (87.5%) (Table 2). Plant No.7 (P-7) had the highest proportion of recurrent parent genome of 95.75%, while plant No. 56 (P-56) and Plant No. 73 (P-73) had 93.25 and 92.75%, making them suitable candidates for developing further selfed generations.

Out of the 74 plants in BC_2 generation, 27% of them had 75–80% of recurrent parent genome while 24% of the population had 80–85%. An extremely low frequency of individuals (1.35%) was found on the upper and lower extremes of recurrent parent genome content. There were five double recombinants in the BC_2 population that contained a heterozygous locus for *opaque2* and flanking markers of recipient allele type on either side of the gene. Two of the three plants viz., P-7 and P-73 that had the highest proportion of recurrent parent genome were also found to be among the five double recombinants.

Phenotypic selection in BC_2F_2 and BC_2F_3

Employing the *opaque2* specific SSR marker, *umc1066*, 53 recessive homozygous individuals out of 225 plants were identified from P-7, which possessed the highest recurrent parent genome content. Similarly, 43 out of 200 and 49 out of 213 BC_2F_2 individuals were recessive homozygous for *opaque2* from P-56 and P-73, respectively. Based on the agronomic evaluation for plant vigour, days to 50% silking and resistance to turcicum and maydis blight, 25 BC_2F_2 progenies from P-7, 23 from P-56 and 23 from P-73 were selected. Several classes of kernel modification were observed in each population of BC_2F_2. Among the harvested ears of 53 BC_2F_2 plants (from P-7) that were homozygous for *opaque2* gene, varying proportions of 4 classes of kernel opaqueness *viz*., less than 25%, 25–50%, more than 50% and 100% opaque were identified. Two out of 53 BC_2F_2 plants produced ears with entirely 100% opaque kernels while no plant produced completely modified kernels. In the other BC_2F_2 populations, the frequency of plants producing ears with 100% opaque or completely modified kernels was likewise very low or zero. The rest of the plants in the BC_2F_2 populations from P-56 and P-73 produced ears with varying proportions of less than 25%, 25–50% and more than 50% opaque kernels. Of the 49 marker selected BC_2F_2 plants from P-73, one plant produced kernels at 10.3% frequency that were completely modified as that of the recurrent parent, V25 (Fig. 3).

The donor QPM inbred line, CML 176 had 7.7% total protein in the endosperm and a tryptophan concentration of 1.05% (Table 3). The recipient line, V 25 had 9.6% total protein in the endosperm and a tryptophan concentration of 0.41%. Among the three

Table 2 Recurrent and donor parent genome content of 'above average' BC_2F_1 individuals (more than 87.5% of recurrent genome) as revealed by background marker analysis

Individual plant no.	Recurrent parent genome content	Donor parent genome content	Unexplained variation
P-7	95.75	04.25	0.0
P-8	87.70	12.30	0.0
P-12	89.25	10.75	0.0
P-20	87.90	03.60	8.5
P-28	90.25	09.75	0.0
P-30	87.80	09.20	3.0
P-34	88.10	11.90	0.0
P-43	87.90	03.60	8.5
P-51	87.80	12.20	0.0
P-54	88.00	12.00	0.0
P-56	93.25	06.75	0.0
P-64	89.90	03.60	6.5
P-73	92.75	07.25	0.0
P-74	90.80	09.25	0.0

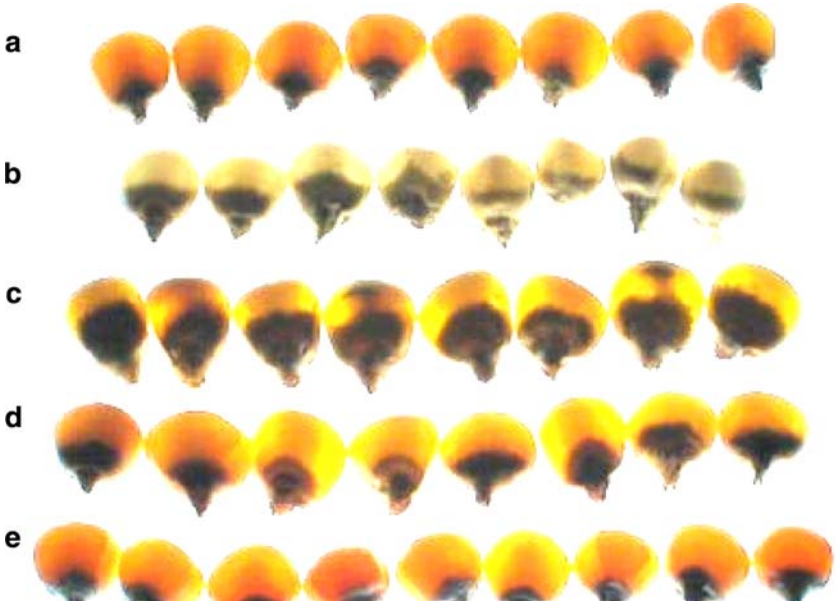

Fig. 3 Screening kernels of BC_2F_2 individuals under transmitted light for kernel modification. **a** V25, **b** CML176, **c** 50% or more opaque kernels, **d** less than 25% opaque kernels, **e** completely modified kernels

Table 3 Total protein content and protein quality parameters in the parental lines and three selected BC_2F_2 families of the target cross V25×CML176

Parent/progeny	Total protein content (%)	Tryptophan in protein (%)	Moisture at harvest (%)	Endosperm modification
CML176 (QPM donor parent)	7.7 ± 0.51	1.05 ± 0.10	22.5 ± 3.3	Hard
V25 (Recurrent parent)	9.6 ± 0.42	0.41 ± 0.03	20.8 ± 3.5	Hard
Selected BC_2F_2 family mean from P-7				
< 25% opaque kernels	9.1 ± 0.71	0.85 ± 0.08	23.9 ± 4.9	Hard
25–50% opaque kernels	8.8 ± 0.53	0.88 ± 0.07	29.5 ± 3.4	Semi-soft
> 50% opaque kernels	8.6 ± 0.49	1.01 ± 0.08	35.6 ± 4.3	Soft
Selected BC_2F_2 family mean from P-56				
< 25% opaque kernels	8.7 ± 0.91	0.83 ± 0.12	22.2 ± 4.1	Hard
25–50% opaque kernels	8.6 ± 0.98	0.92 ± 0.09	26.5 ± 5.2	Semi-soft
> 50% opaque kernels	8.8 ± 0.60	0.90 ± 0.11	29.8 ± 3.1	Soft
Selected BC_2F_2 family mean from P-73				
< 25% opaque kernels	9.0 ± 1.20	0.78 ± 0.10	20.6 ± 3.6	Hard
25–50% opaque kernels	8.5 ± 0.93	0.87 ± 0.08	27.8 ± 6.2	Hard
> 50% opaque kernels	8.3 ± 0.75	0.94 ± 0.10	31.8 ± 7.5	Soft

Mean ± standard error

BC_2F_2 families, the mean total protein content ranged from 8.3 to 9.1%, while mean tryptophan in protein ranged from 0.78 to 0.94%. The moisture content at harvest ranged from 20.6 to 35.6% with the donor parent at 22.5% and recurrent parent at 20.8%. In general, less than 25% opaque kernels had hard endosperm texture, while the 25–50% and more than 50% opaque kernels had semi-soft and soft texture respectively.

Days to 50% silking ranged from 50.2 to 58.1, days to pollen shed from 49.3 to 58.0, grain yield from 3.5 to 4.9 t/ha, total protein in endosperm from 8.1 to 9.2% and tryptophan content from 0.69 to 0.82% (Table 4). Based on the selection index constructed giving equal weight to agronomic and biochemical traits, three BC_2F_3 lines viz., P-7-D, P-7-B and P-56-A were identified representing partially converted QPM versions of recurrent parent, V25.

Discussion

Marker assisted selection for *opaque2* gene

In this study, distinct polymorphism could be observed between the normal and QPM inbred lines with all the three SSR markers. However, the nature of polymorphism was different with respect to *phi112*, which exhibited dominant (presence-absence) polymorphism, restricting its potential utility in identifying the three forms of genotypes for the *opaque2* gene. Nevertheless, this marker could be of use in checking the seed purity during routine field maintenance of QPM inbred lines. Co-dominant nature of the polymorphism exhibited by *phi057* and *umc1066* enables their potential utility in MAS programs to successfully discriminate between homozygotes and heterozygotes. Identification of het-

Table 4 Agronomic and biochemical features of the recurrent parent (V25), donor parent (CML176) and converted QPM version of V25 (BC_2F_3 line)

Traits/descriptors	V25 (recurrent parent)	CML176 (QPM donor parent)	Converted QPM version of V25 (BC_2F_3 line)
Plant height (cm)	180.1 ± 16.0	205 ± 18	192 ± 17
Ear height (cm)	85.3 ± 7.1	94.5 ± 9.0	92.3 ± 9.3
Reaction to Turcicum blight (1–5)[a]	2.75 ± 0.18	1.0 ± 0.08	1.5 ± 0.12
Reaction to Maydis blight (1–5)[a]	3.0 ± 0.21	1.5 ± 0.09	1.5 ± 0.19
Days to 50% silking	53.1 ± 3.6	69.0 ± 5.4	55.1 ± 4.2
Days to 50% pollen shed	52.4 ± 5.1	67.5 ± 4.8	53.5 ± 4.2
Total protein content (%)	9.4 ± 0.48	8.1 ± 0.62	8.9 ± 0.76
Tryptophan in protein (%)	0.39 ± 0.04	0.99 ± 0.10	0.82 ± 0.07
Kernel hardness	Hard	Hard	Hard
Grain yield[b] (Tons/ha)	3.8 ± 0.25	5.10 ± 0.19	4.2 ± 0.35

Mean ± standard error
[a] 1—Resistant and 5—Susceptible
[b] Upon selfing

erozygotes in the seedling stage prior to pollination aided in the rejection of non-target BC progenies resulting in substantial saving of labor and material resources. However, as demonstrated by our results, polymorphism may not be obtained for all the normal inbred lines against QPM inbred lines with these SSR markers, warranting cautious selection of QPM donors in a marker aided QPM breeding program. With respect to the gene specific markers such as *phi057* and *umc1066*, which are located within the *opaque2* gene itself, the individual plants in any segregating population could be scored directly for the gene, eliminating the probability of occurrence of false positives and false negatives. While this kind of relationship between the markers and the target gene is the most preferred, it is also difficult to find such markers for most of the traits/genes, unless they are already cloned and sequenced.

Determination of optimum population size in backcross generation

There are two major issues in a backcross breeding program that are critical in deciding upon the optimum population size of BC generations either a priori or posteriori (Frisch et al. 1999a, b). (1) reduction of the proportion of the donor genome on the carrier chromosome around the target allele (linkage drag) and (2) reduction of the donor genome on the non-carrier chromosomes to the maximum extent. In this study, the linkage drag around the *opaque2* locus was minimized by deliberately mapping more polymorphic markers near to the locus (bin 7.01). The flanking markers, *bnlg1200* and *bnlg2160* were positioned at 3.8 and 4.2 cM respectively from the *opaque2* locus. However, if the flanking marker distances from the target locus are already known from any of the mapping populations involving the target recipient and donor line, then it would obviate the need for selective linkage mapping in BC_1 generation. Based on the earlier studies by Lee (1995) and Ribaut et al. (2002), it could be generalized that whenever a target gene is introduced for the first time presumably from either wild or unadapted germplasm, flanking markers as close as 2 cM is considered the ideal option, while in the transfer of the same target gene in subsequent phases from elite into elite lines, positioning the flanking markers even at 12 cM might be more effective in reducing the required size of the backcross population. In this study, although the donor QPM inbred line, CML176 is an elite line with proven combining ability, the flanking markers were positioned as close as 3.8 and 4.2 cM because it is from a sub-tropical germplasm, with white kernels and belongs to late maturity group.

In this study, we define the optimum population size as the minimal number of individuals that need to be genotyped at each generation so that there is a 99% probability of obtaining at least one double recombinant latest by the second BC generation along with maximum proportion of recurrent parent genome. We combined the salient features of a priori (Hospital and Decoux 2002) and posteriori (Frisch and Melchinger 2001) approaches and selected three individuals in the BC_1 generation that were single recombinants at the *opaque2* locus with the closest flanking marker, *bnlg1200* and calculated the population size of the BC_2 generation using the Popmin program, specifying the initial genotypes selected. In a practical QPM breeding program, selecting for a single individual in the BC_1 generation as suggested in a priori approach could be a risky proposition considering the possible unforeseen problems, such as poor germination, ear rot susceptibility and poor pollen quality (Ribaut et al. 2002). Besides, selection of more than one individual in BC_1 generation aids in establishing wider genetic variability and provides ample scope for phenotypic selection of desirable agronomic traits in later generations.

Background selection for recovery of maximum recurrent parent genome content

The objective of the background/whole genome selection is to recover rapidly maximum proportion of

recurrent parent genome at non target loci through markers that are distributed evenly throughout the genome (Young and Tanksley 1989; Hospital et al. 1992; Visscher et al. 1996; Frisch et al. 1999a, b). Availability of robust anchored marker maps in maize renders application of marker aided background selection an easy and attractive proposition. However, the cost of employing background selection in each BC generation could be prohibitive to many public sector breeding programs. The earlier simulation studies of Frisch et al. (1999a, b) and Ribaut et al. (2002) have indicated that application of background selection in one later generation along with foreground selection in each BC generations could be efficient and cost-effective. In the present investigation, we followed a two generation marker based breeding program in which whole genome background selection at non-target loci was applied only in the BC_2 generation. As predicted by Ribaut et al. (2002), the selection response for background analysis depends on the recombination frequency between the target gene and the flanking markers and on the densities of markers on the carrier and non-carrier chromosomes. The critical issue of number and distribution of unlinked target loci to be screened depends upon the required levels of line conversion, the extent of saturation of the base molecular marker map and availability of technical resources at a given point of time. In our study, we selected markers representing each established bin location on all the chromosomes of maize genome as per the SSR consensus map available in public domain (http://www.maizegdb.org). Although it has been demonstrated through simulation studies that increasing the number of markers to more than three per non-carrier chromosome was not efficient at early generations (Hospital et al. 1992), care needs to be taken in a practical MAS program to avoid sampling error and exaggerated estimates of recurrent parent genome associated with smaller number of marker data points. The distribution of individuals with low, moderate and high recurrent parent genome content in the BC_2 population in more or less symmetrical fashion in the present study indicated unbiased sampling and sufficient number of marker data points.

Based on the marker aided background analysis, ideally one BC_2 individual with highest proportion of recurrent parent genome needs to be chosen for the next selfing generation. Nevertheless, considering the unpredictable field problems especially with QPM germplasm (which is prone to poor germination and insufficient kernel modification), we selected the first three BC_2 individuals with highest proportion of recurrent parent genome for developing further BC_2F_2 families. There could be several modifications to the procedure we have followed in this study depending upon the requirements and objectives of each breeding scheme. For instance, if reducing linkage drag around the target locus is of paramount importance in a given breeding program, screening the BC_2 population initially for the double recombinants and then exercising whole genome selection on selected individuals could prove to be cost-effective.

Phenotypic selection: a key to success in QPM breeding

The *opaque2* allele is recessive and the endosperm modifiers are polygenic with, no reliable molecular markers identified for kernel modification. Phenotypic screening of the individual kernels under transmitted light and selection of kernels that have less than 25% opaqueness is by far the most convenient and efficient strategy employed in all the QPM breeding programs (Vasal et al. 1993a, b).

The very low frequency of 100% opaque (soft) and completely modified (hard) kernels in the three BC_2F_2 families corroborated the earlier findings of Lopes and Larkins (1995) that kernel modification in QPM is governed by more than one minor gene. We preferred kernels with less than 25% opaqueness over 25–50% and more than 50% opaqueness due to the semi-soft/soft nature of the endosperm and susceptibility to ear rot of latter categories (Vasal et al. 1993a, b). The biochemical analysis of the three classes of kernel modification showed that tryptophan concentration in protein, the chief indicator of protein quality was enhanced more than twice in all the three classes as compared to original recipient line. Lysine proportion in protein was not determined due to the strong positive correlation between the two essential amino acids in endosperm protein (Pixley and Bjarnasan 1993).

In addition, phenotypic selection was exercised in the marker-selected *opaque2* homozygous individuals in the three BC_2F_2 families for other desirable agronomic traits such as days to 50% silking, resistance to turcicum and maydis blight, pollen shedding ability (data not shown) and seed yield. Thus the phenotypic selection in the marker selected *opaque2* homozygous BC_2F_2 individuals provided an excellent opportunity to combine the desirable agronomic traits with superior protein quality and hard endosperm characters. The stable expression of these traits could be confirmed through BC_2F_3 families, which represented the final converted QPM versions of the original recipient lines. The single cross QPM hybrids developed using converted QPM versions of V25, CM 212 and CM145 showed 9–12% superiority in grain yield over their normal versions with tryptophan concentration of above 0.72% in endosperm protein (Unpublished).

The rapid line conversion strategy outlined in this investigation brings together the salient features of both marker aided and phenotypic selection approaches such as fixing the large segregating generation for the target locus, reduction of the linkage drag by selecting for flanking markers of recipient allele type, recovery of maximum amount of recurrent parent genome within two BC generations and provides ample scope for exercising phenotypic selection for as many desirable agronomic and biochemical traits as possible. With the

advent of third generation marker technology such as SNPs and non gel based detection of PCR amplified products viz., Taqman and molecular beacons (Salvi et al. 2001; Tyagi et al. 2001), the power and efficiency of genotyping is expected to improve in the coming decades.

Acknowledgements This study was, in part, supported by a competitive research grant awarded to RB and HSG by Dept. of Biotechnology, Govt. of India. The authors thank the breeder, V. P. Mani, who developed the original normal inbreds viz., V25, CM212 and CM145. Thanks are also due to the International Maize Testing Unit of CIMMYT, Mexico for providing QPM donor lines. The technical help rendered by K. S. Koranga, G. S. Bisht, Ram Murti, Mohan Bhatt and Ohm Pratap is gratefully acknowledged.

References

Annual Report (2004–05) Ministry of Agriculture, Department of Agriculture & Cooperation, Govt of India

AOAC (1965) Official methods of analysis of the association of official agricultural chemists. 10th ed, pp 744–745

Babu R, Nair SK, Prasanna BM, Gupta HS (2004) Integrating marker assisted selection in crop breeding-prospects and challenges. Curr Sci 87:607–619

Bass HW, Webster C, O'Brien GR, Roberts JKM, Boston RS (1992) A maize ribosome inactivating protein is controlled by the transcriptional activator *opaque2*. Plant Cell 4:225–234

Frisch M, Bohn M, Melchinger AE (1999a) Comparison of selection strategies for marker assisted backcrossing of a gene. Crop Sci 39:1295–1301

Frisch M, Bohn M, Melchinger AE (1999b) Comparison of selection strategies for marker assisted backcrossing of a gene. Crop Sci 39:967–975

Frisch M, Melchinger AE (2001) Marker assisted backcrossing for introgression of a recessive gene. Crop Sci 41:1485–1494

Glover DV, Mertz ET (1987) Corn. In: Olson RA, Frey KJ (eds) Nutritional security of cereal grains: genetic and agronomic improvement. Agron Monogr 28 ASA CSSA and SSSA Madison, WL

Habben JE, Moro GL, Hunter BG, Hamaker R, Larkins BA (1995) Characterization of the proteins that define the nutritional quality of maize endosperm. In: Larkins BA, Mertz ET (eds) Quality protein maize: 1964–1994. Proc. International Symposium on Quality Protein Maize, EMBRAPA/CNPMS, Sete Lagaos, MG, Brazil, Dec 1–3, 1994

Hernandez HH, Bates LS (1969) A modified method for rapid tryptophan analysis of maize. CIMMYT Research Bull 13 CIMMYT, Mexico

Hospital F, Chevalet C, Mulsant P (1992) Using markers in gene introgression programs. Genetics 132:1199–1210

Hospital F, Decoux G (2002) Popmin: a program for the numerical optimization of population sizes in marker assisted backcross programs. J Hered 93:383–384

Lee M (1995) DNA markers and plant breeding programs. Adv Agron 55:265–344

Lincoln S, Daly M, Lander E (1992) Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Tech Rep, 3rd ed

Lohmer S, Maddaloni M, Motto M, Difonzo N, Hartings H, Salamini F, Thomson RD (1991) The maize regulatory locus *opaque2* encodes a DNA binding protein which activates the transcription of the B-32 gene. Embo J 10:617–624

Lopes MA, Larkins BA (1995) Genetic analysis of *opaque2* modifier gene activity in maize endosperm. Theor Appl Genet 119:274–281

Lopes MA, Takasaki K, Bostwick DE, Helentjaris T, Larkins B (1995) Identification of two *opaque2* modifier loci in QPM. Mol Gen Genet 247:603–613

Melchinger AE (1990) Use of molecular markers in plant breeding. Plant Breed 104:1–19

Mertz ET (1992) Discovery of high lysine, high Tryptophan cereals. In: Mertz ET (ed) Quality protein maize. American Society of Cereal Chemistry, St Paul, MN, pp 1–8

Mertz ET, Bates LS, Nelson OE (1964) Mutant that changes protein composition and increases lysine content of maize endosperm. Science 145:279–280

Pixley KV, Bjarnasan MS (1993) Combining ability for yield and protein quality among modified endosperm *opaque2* tropical maize inbreds. Crop Sci 33:1229–1234

Ribaut JM, Hoisington DA (1998) Marker assisted selection: new tools and strategies. Trends in Pl Sci 3:236–239

Ribaut JM, Jiang C, Hoisington DA (2002) Simulation experiments on efficiencies of gene introgression by backcrossing. Crop Sci 42:557–565

Salvi S, Tuberosa R, Phillips RL (2001) Development of PCR based assays for allelic discrimination in maize by using the 5′ nuclease procedure. Mol Breed 8:169–176

Schmidt RJ, Burr FA, Aukerman MJ, Burr B (1990) Maize regulatory gene *opaque2* encodes protein with a "leucine zipper" motif that binds to zein DNA. Proc Natl Acad Sci 87:46–50

Senior MC, Murphy JP, Goodman MM, Stuber CW (1998) Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. Crop Sci 38:1088–1098

Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Bio Rep 1:3–8

Tyagi s, Bratu DP, Kramer FR (1998) Multicolor molecular beacons for allele discrimination . Nature Biotech 16:49–53

VanBerloo R (1999) GGT: software for display of graphical genotypes. J Hered 90:328–329

Vasal SK, Srinivasan G, Pandy S, Gonzalez CF, Crossa J, Beck DL (1993a) Heterosis and combining ability of CIMMYT's quality protein maize germplasm I: lowland tropical. Crop Sci 33:46–51

Vasal SK, Srinivasan G, Pandy S, Gonzalez CF, Crossa J, Beck DL (1993b) Heterosis and combining ability of CIMMYT's quality protein maize germplasm II: sub tropical. Crop Sci 33:51–57

Villegas E (1975) An integral system for chemical screening of quality protein maize In: High qulity protein maize. Proc CIMMYT-Purdue Int Symp on protein quality in maize, El Batan, Mexico

Villegas E, Ortega E, Bauer R (1984) Chemical methods used at CIMMYT for determining protein quality in corn. CIMMYT, Mexico

Visscher PM, Haley CS, Thompson R (1996) Marker assisted introgression in backcross breeding programs. Genetics 144:1923–1932

Young ND, Tanksley SD (1989) RFLP analysis of the size of chromosomal segments retained around Tm-2 locus of tomato during backcross breeding. Theor Appl Genet 77:353–359