ORIGINAL PAPER

Mapping QTLs for *opaque2* modifiers influencing the tryptophan content in quality protein maize using genomic and candidate gene-based SSRs of lysine and tryptophan metabolic pathway

B. Kalyana Babu · P. K. Agrawal · S. Saha · H. S. Gupta

Abstract

Key message **The mapping analysis resulted in identification of five significant QTLs for *opaque2* modifiers influencing the tryptophan content in quality protein maize using functional and genomic SSR markers.**

Abstract Quality protein maize (QPM) was developed by selecting genetic modifiers that convert *opaque2* mutant containing high lysine and tryptophan. There are several unlinked *opaque2* modifier loci (*Opm*) in QPM whose location, nature and mode of action are not clear. To identify these *Opm* QTLs, we developed a population of 218 $F_{2:3}$ individuals from a cross between VQL2 and VQL8, two isogenic QPM inbreds significantly differing in tryptophan content. Based on the data of the $F_{2:3}$ population, five significant QTLs on chromosomes 5, 7 and 9 with LOD values more than 2.5 were identified and together explained 38.6 % of the total phenotypic variance (R^2). The *Wx1* gene which has influence on the amino acid composition of the maize endosperm was mapped on chromosome 9 near the marker phi022 and also validated by bulk analysis. The QTL near the SSR marker ZmASK3, developed from the aspartate kinase 2 gene of the lysine pathway, mapped on chromosome 5 and had LOD of 2.7 with R^2 of 5.1 %. On chromosome 9, the QTL between the loci umc1430 and bnlg1401 had an LOD of 4.5 with R^2 of 9.1 %, whereas the QTL between the loci bnlg1401 and phi022 had an LOD of 4.2 with R^2 of 8.4 %. The third QTL was observed to be close to the marker umc2207 with an LOD of 4.8 and R^2 of 8.4 %. The identified QTLs will be very useful in the marker-assisted back-cross breeding and transgressive breeding for the development of QPM maize.

Introduction

Cereals supply more than half of the dietary proteins to human beings. However, these are nutritionally imbalanced as the most abundant storage proteins of cereals (prolamins) lack essential amino acids such as lysine, tryptophan and methionine. Maize is the most preferred staple food for more than one billion people in Sub-Saharan Africa, and Latin America and Asian continents. It is also primarily used as livestock feed, due to the strong economic growth and rapid urbanization experienced by many countries including India. The lysine and tryptophan contents in maize endosperm are considered to be one of the most important traits for determining the nutritional quality of food and feed. Maize endosperm contains 1.5–2.5 % lysine and 0.25–0.5 % tryptophan; however, 5 % lysine and 1.1 % tryptophan contents are required for optimal human nutrition (Young et al. 1998). The *Opaque2* (*o2*) mutation increases the lysine and tryptophan contents in maize endosperm by decreasing the synthesis of prolamin (*zein*)

B. K. Babu (✉) · P. K. Agrawal · S. Saha · H. S. Gupta
Vivekananda Parvateeya Krishi Anusanthan Sansthan (Indian Council of Agricultural Research), Almora 263601, Uttarakhand, India
e-mail: kalyan_biotek@yahoo.co.in

S. Saha · H. S. Gupta
Indian Agricultural Research Institute (Indian Council of Agricultural Research), New Delhi 110 012, India

proteins, thereby increasing the level of other types of endosperm proteins (Munck 1992). The *o2* maize was unpopular with breeders because of low yield and susceptibility to pests and diseases. The alleles controlling the soft and starchy texture of *o2* endosperm, are designated as *opaque2* modifiers (*Opm*) and these have been proved to be genetically complex in nature. The plant breeders at the CIMMYT, Mexico, and Pietermarizburg, South Africa (Geevers and Lake 1992), systematically introgressed *Opm* genes into *o2* germplasm to develop quality protein maize (QPM) (Vasal et al. 1980). It has the phenotype and yield of normal maize, but high lysine and tryptophan contents (Mertz 1992; Prasanna et al. 2001). The development and widespread use of QPM germplasm is limited because of the widespread distribution of *o2* modifiers among all the ten chromosomes and the technical complexity of the multiple loci (Vasal et al. 1980). This process could greatly be accelerated if the complexity and location of the *Opm* genes were known. However, relatively little is known about their number, chromosomal location and mechanism of action. These *Opm* influence the vitreousness of kernels, seed density and quality parameters such as lysine and tryptophan contents.

Candidate genes are sequenced genes of known biological function associated with the development or physiology of a trait. Candidate genes have been effectively used in several molecular breeding programs such as QTL mapping, association mapping, marker-assisted selection and development of transgenics for the improvement of agronomically important traits in crop plants (Barret et al. 2004). In maize, candidate genes associated with lysine metabolism were identified by Wang et al. (2007), who found lys-sensitive Asp kinase to be a better candidate gene for the QTL affecting free amino acid content than Thr-sensitive AK–HSDH2 (aspartate kinase–homoserine dehydrogenase 2). In another study, Wang and Larkins (2001) reported QTLs for aspartate kinase 2 and AK-homo-Ser dehydrogenase (HSDH). The exponential increase in publicly available databases made it possible to identify SSRs derived from these sequences which are desirable DNA markers, since they are abundant in plant genomes, highly polymorphic, rapid and less expensive, and can be used to identify specific chromosomal regions across populations (Chin et al. 1996). This is the first report on the utilization of candidate gene-based SSR markers of lysine and tryptophan metabolic pathways along with genomic SSR markers in mapping of *opaque2* modifiers in quality protein maize (QPM) using a QPM × QPM population to map the modifiers.

Holding et al. (2008) analyzed individual kernels for their hardness and seed density from a $F_{2:3}$ population and BSA studies, which confirmed the presence of QTLs for *Opm* on chromosomes 7 and 9. They also found additional loci associated with vitreousness on chromosomes 9 and 6. Yang et al. (2005) reported molecular genetic mapping of high lysine mutant gene (*opaque*16) by composite interval mapping. They identified a major QTL at either 3 or 2.2 cM distance from the marker umc1141 in the interval between the markers, umc1121 and umc1141, on the long arm of chromosome 8. The present study was undertaken to identify QTLs influencing the tryptophan content in quality protein maize. In our previous studies, the tryptophan content of the quality protein maize (QPM) inbreds was found to vary from 0.52 to 0.98 % among many newly developed isogenic lines. This variation is expected to be due to *o2* modifiers influencing the tryptophan content, since all of them were with the *o2o2* background. Among all the QPM inbreds, two genotypes (VQL2 and VQL8) derived from the same parent (CM145), but significantly differing in tryptophan content (0.52 and 0.94 %), were taken for generating an effective mapping population to identify highly influencing QTLs of *o2* modifiers for tryptophan content using genic and genomic SSR markers. This is the first kind of report on utilizing candidate gene-based EST-SSRs to identify QTLs influencing the tryptophan content of *o2* maize. Thus, the objective of the present study was to identify effective QTLs influencing *o2* modifiers for tryptophan content in QPM, using genomic SSRs and candidate gene-based SSRs, and their further utilization in marker-assisted selection for improvement of lysine and tryptophan contents in the locally adopted maize germplasm.

Materials and methods

Plant materials

The two inbred lines used in the present study, VQL2 and VQL8, were derived from a common parent, CM145, but differed in the tryptophan content to a great extent (Babu et al. 2009). These were considered as ideal parents for the mapping of the *o2* modifiers. VQL2 and VQL8 are the QPM inbred lines derived by introgression of *o2* gene from CML170 into CM145. The cross between VQL2 (low tryptophan content) and VQL8 (high tryptophan content) was raised at the Hawalbagh research farm, Vivekanand Parvatiya Krishi Anusanthan Sanstha (ICAR), Almora. The $F_{2:3}$ population was raised under mid hill condition at VPKAS experimental farm, Hawalbagh, Almora (29°36′ N and 79°30′ N situated at an elevation of 1,250 m above mean sea level) along with the parental lines. Recommended agronomic practices were followed for raising the crop. A total of 218 F_2 plants were generated from the F_1 plants. The seeds of the 218 F_3 ears were harvested from the F_2 plants. The F_2 plants of this cross were used for

genotyping using SSR markers, and the F_3 seeds (families) harvested from each of the F_2 plants were utilized for measuring the tryptophan content of the kernels. All the seeds were thoroughly sun dried and stored at room temperature for 6 months before the determination of tryptophan content.

Determination of tryptophan content using HPLC

Chromatographic separation of tryptophan was performed using a Waters Separations Module, connected with analytical columns (as described below), with Model 2475 fluorescence detector. The chromatographic separation was performed by the AccQ.Tag[TM] method. The tryptophan content was observed by the fluorescence detector using an excitation wavelength of 270 nm and emission wavelength of 340 nm (gain of 1) throughout the analysis. Before analysis, all eluents (eluent A: methanol:acetic acid:water, 10:2:88, v:v:v, eluent B: methanol) were filtered and degassed. Data acquisition and system suitability tests were performed with Waters Empower[TM] software. The concentration of tryptophan was estimated by comparing the area of standard solution using standard calibration curve and unknown sample. The tryptophan stock standard solutions containing 200 µg/mL were prepared in milliQ water and stored at 4 °C. Successive dilutions of stock standard solutions were made with milliQ water to obtain seven points of calibration curve (200, 100, 50, 25, 12.5, 6.25, 3.12 µg/mL). The calibration points were analyzed in triplicate with each point of line corresponding to the average of three measures.

DNA preparation and SSR analysis

DNA was extracted from the 218 F_2 individuals of the mapping population employing the procedure described in Shen et al. (1994). Approximately, 50 ng of DNA was used as template for PCR in a 15 µL reaction volume in a thermal cycler (MJ Research, USA), programmed for an initial denaturation of 2 min at 94 °C, followed by 35 cycles of amplification, at 94 °C for 30 s of annealing temperature (58–64 °C based on T_m values of different SSR primers), and extension for 2 min at 72 °C, with a final extension of 7 min at 72 °C and hold at 4 °C. The amplification products were separated on 3.5 % superfine resolution (SFR) agarose gel as per the procedure suggested by Senior et al. (1998). A total of 765 SSR loci uniformly spaced in the genome (http://www.maizegdb.org) were assessed for SSR polymorphism in the parental lines of the mapping population. Of the 765 SSR loci, 27 were found to be polymorphic, producing discriminatory banding pattern in the mapping population. These 27 polymorphic SSR loci were used for genotyping the 218 F_2 individuals.

Identification of candidate genes, their sequences and SSR primer designing

Candidate genes have been selected from the metabolic pathways of lysine and tryptophan and *opaque2* modifiers in maize as previously reported by Wang and Larkins (2001) and Wang et al. 2007. The details of the retrieval of candidate genes, their sequences and SSR primer designing have been reported earlier (Babu et al. 2012).

Linkage analysis and QTL detection

A linkage map for the F_2 population from the cross between VQL2 and VQL8 was determined using 27 informative SSR markers with the software Mapmanager (Manly et al. 2001) and QGene (Nelson 1997). By means of the Kosambi mapping function (Kosambi 1944), the values of recombinant fractions were converted into genetic map distance (cM). The QTL mapping was done with two softwares, viz., QTL cartographer and QGene. Three methods, forward regression, backward regression and forward and backward regression, of composite interval mapping (CIM) were used to examine the association between the tryptophan content and markers with Windows QTL Cartographer Version 2.0 (North Carolina State University, Raleigh, NC, USA) programmed by Wang et al. (2002). The parameters were set as follows: map function, Kosambi; distance units, centimorgan (cM); cross information, SF3 (self-cross3); walk speed, 2 cM; LOD = 2.5; CIM mode selection; model 6, i.e., standard model; background control, 5 of control marker numbers and 10.0 cM of window size.

QTL analysis was carried out on the set of 218 F_2 individuals with the phenotypic data for tryptophan content. The genotypic data were generated using 27 SSR loci and the phenotypic data (tryptophan content) were generated by HPLC. The method of simple interval mapping (SIM) and CIM was used to map the QTLs and estimate their genetic effects, as implemented in QTL Cartographer Version 2.5. Empirical threshold levels for declaring QTL significance at a type 1 error rate of 0.05 were obtained by performing 1,000 permutations of the data according to the method of Churchill and Doerge (1994). The best estimate of QTL location was assumed to correspond to the position having the peak significance level, and QTL support intervals were drawn as the entire region of the chromosome that exceeded the threshold value. The phenotypic variance (R^2) was calculated from the QGene software and multiplied with 100 for representing the R value in percentage.

Bulk segregant analysis

Bulk segregant analysis was done by selecting an equal amount of DNA from the ten highest and ten lowest tryptophan-containing natural genotypes with the above-mentioned 27 polymorphic SSR loci, including the candidate gene-based SSR loci.

Results

Analysis of tryptophan content in the mapping population

The tryptophan content of F_3 seeds from F_2 individual plants was determined using the HPLC method as described earlier. The lowest tryptophan content (percentage of protein) of the 218 F_2 genotypes was 0.162, while the highest value was 0.905 with a mean of 0.353. These values varied for replications 1 and 2 (Table 1). There were 90 F_2 individuals which had higher tryptophan level than

Table 1 The tryptophan content of the quality protein maize $F_{2:3}$ mapping population with two replications

Tryptophan content of $F_{2:3}$ individuals	Average (%)	Replication 1 (%)	Replication 2 (%)
Minimum	0.162	0.163	0.160
Maximum	0.905	0.890	0.998
Mean	0.353	0.352	0.354
Median	0.334	0.326	0.330
Variance	0.017	0.017	0.018

the average, while 128 had lower tryptophan content than the average value.

Candidate gene-based EST-SSR loci analysis

A total of 27 SSR loci were designed from 16 GenBank accessions representing the candidate genes of lysine and tryptophan metabolism and *Opm*. Out of the 27, two belonging to tryptophan synthase 1 could not be amplified. Hence, 25 SSR loci were used for the parental polymorphism analysis for the present study. Out of the 25 SSR loci, 4 developed from the GenBank accessions of lysine metabolic pathways were polymorphic and used for mapping purpose (Table 2).

SSR data analysis and linkage mapping

A total of 765 SSR loci including the candidate gene-based SSRs, spread across all the ten chromosomes, were screened for parental polymorphism and 27 were found (including 4 from candidate gene-based SSRs) to be polymorphic. These 27 polymorphic SSR loci were used for genotyping of 218 F_2 individuals. The gel pattern of the polymorphic markers between the parents, VQL2 and VQL8, is given in Fig. 1. Missing data in the genotyping of the mapping population across the 27 SSR loci were negligible (< 5 %). All the polymorphic SSR alleles were represented in five maize chromosomes in the linkage map spanning a total length of 733.1 cM at an average marker interval of 27.2 cM (Fig. 2). The details of the polymorphic primers with their bin location, repeat motif and annealing temperature are given in table S1. The

Table 2 The details of the polymorphic primers of candidate genes found between parents along with their sequences, contig numbers, position in the contig and expected product size

Candidate gene/putative function	Contig number	SSR loci	Primer sequence	Expected product size (bp)	Amplified product size (bp)	T_m (°C)
Lysine metabolism pathway candidate genes						
Aspartate kinase 2 (ask 2)	AJ400869	ZmASK2 F	5′ GCGATTGTTTGACCGGCT 3′	214	205–230	55
		ZmASK2 R	5′ TGTTGGCCCCATCCATTG 3′			
		ZmASK3 F	5′ TAGCACCCCCAGATGTTG 3′	105	110–150	53
		ZmASK3 R	5′ CA CAG GCC ATT CTT GAC 3′			
Homoserine dehydrogenase/ aspartokinase	BT085133	ZmHSDH1 F	5′ GCCGAAGATGTCCACTCCACCG 3′	450	490–700	58
		ZmHSDH1 R	5′ CCACAGCATCTTTCAGCAGG 3′			
	BT042726	ZmHSDH5 F	5′ CCGAAGATGTCCACTCCACCG 3′	170	260–310	58
		ZmHSDH5 R	5′ CTTCCGAGTGTGGTAGCG 3′			

Naming of the primers: ZmAKS, *Zea mays* aspartate kinase (initials of the candidate genes followed by serial number)

Fig. 1 The gel pattern of parental polymorphism using SSR markers (*1–24* correspond to primers umc1106, umc1071, umc1685, umc2229, umc2207, umc1430, umc1016, bnlg1832, bnlg1305, umc1456, phi022, umc2346, umc2160, bnlg2086, umc1153, umc1732, bnlg2160, umc1036, umc1567, phi028, umc2362, bnlg1401, umc2084 and ZmASK2, respectively)
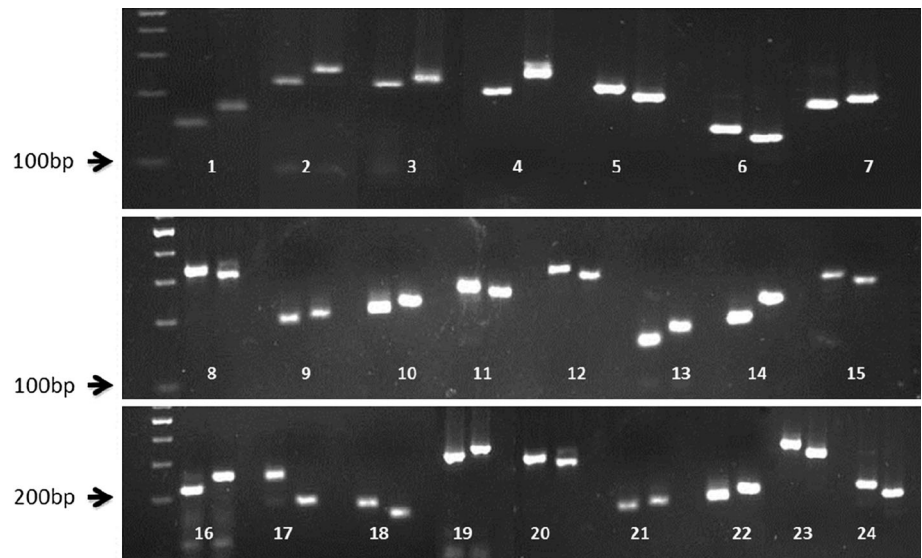
Fig. 2 The linkage map and location of the QTLs identified in the F$_{2:3}$ population of quality protein maize using SSR and EST-SSRs [the numerical values in the *left* side represent distance in centimorgan (cM)]
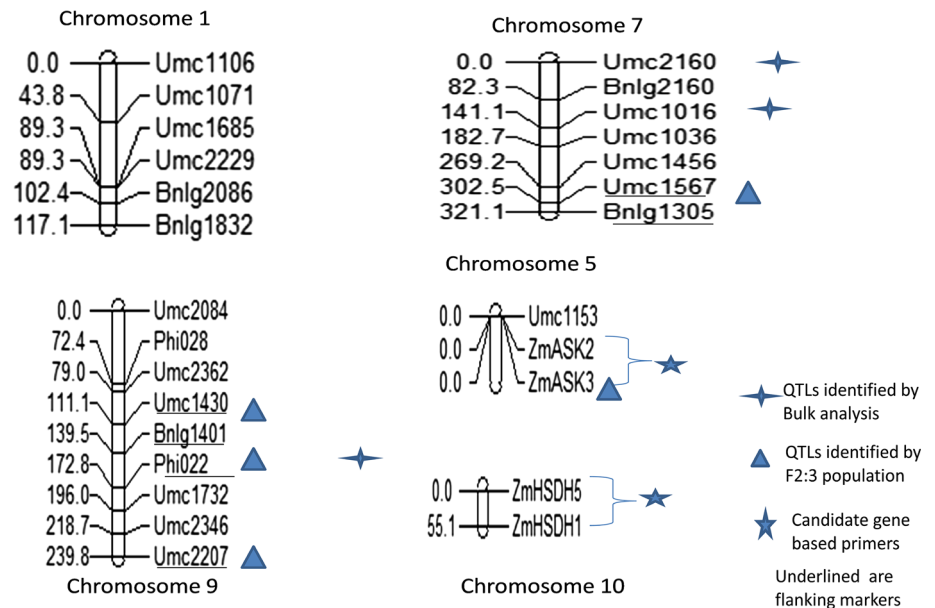
chromosome 1 covered a length of 117.1 cM with six polymorphic markers at an average distance of 19.5 cM. Chromosome 7 covered a length of 321.1 cM with seven polymorphic markers at an average distance of 45.9 cM, while the chromosome 9 covered 239.8 cM length with nine markers at an average distance of 26.7 cM. However, chromosome 10 covered a length of 55.1 cM with an average distance of 27.6 cM. The location of the newly identified candidate gene-based SSRs was done by the BLAST analysis. The sequences were BLAST against the genomic database found in the NCBI database (http://www.ncbi.nlm.nih.gov/). The two primers (ZmASK2 and ZmASK3) designed from the contig AJ400869 were found to have 100 % similarity with the part of the sequence of

chromosome 5; however, the primers ZmHSDH1 and ZmHSDH5 (designed from the contigs BT085133 and BT042726, respectively) were found to have 100 % similarity with the part of the sequence of chromosome 10 using the gramene website (http://www.gramene.org).

QTL mapping analysis for tryptophan content

The QTL analysis was done with two softwares, QTL cartographer and QGene, which showed similar results. However, the Qgene software found two additional QTLs on chromosomes 5 and 7. Single marker regression (SMR) analysis, simple interval mapping (SIM) and composite interval mapping (CIM) analysis based on the tryptophan

Table 3 The details of the identified QTLs with their LOD value, phenotypic variance (R^2) and additive effect

Chromosome	Marker	Average			Replication 1			Replication 2		
		LOD	R^2 (%)	Add effect	LOD	R^2 (%)	Add effect	LOD	R^2 (%)	Add effect
5	ZmASK3	2.7	5.1	−0.035	3.2	6.6	−0.037	–	–	–
7	Umc1567-bnlg1305	3.2	6.5	−0.644	3.0	6.3	−0.65	3	6.1	−0.629
9	Umc1430-bnlg1401	4.5	9.1	−0.031	3.8	7.8	−0.025	5	10.1	−0.041
	Phi022	4.2	8.4	−0.604	3.2	6.7	−0.541	5.1	10.2	−0.68
	Umc2207	4.8	9.5	−0.065	4.9	10		4.2	8.5	−0.08
	Total		38.6			37.4			34.9	

data of 218 individuals of the $F_{2:3}$ population and genotypic data of 27 markers were performed. The results from SMR and CIM were similar except for the replication 2 (R2) data, where only four QTLs were found on chromosome 5 for R2 data. However, broadly the results of SIM and CIM for both replicated data and average data were similar.

Based on CIM analysis, five QTLs were detected on chromosomes 5, 7 and 9 (each one on chromosomes 5 and 7; three on chromosome 9). Highly effective QTLs were detected on chromosome 9 which explained a phenotypic variance of 27 % by the three QTLs. The QTL between the loci umc1430 and bnlg1401 had an LOD value of 4.52 with an R^2 value of 9.1 %, which is the second highest among the QTLs reported here. The QTL near the loci phi022 spanning the marker interval between bnlg1401 and phi022 had LOD value of 4.2 with R^2 value of 8.4 %. The SSR loci phi022 was reported to be linked to the waxy1 (*Wx1*) gene, which influences the amino acid composition of maize endosperm. The third QTL on chromosome 9 is near the loci umc2207 and had LOD of 4.8 with R^2 value of 9.5 %, which is the highest among all the QTLs. All the three QTLs had an additive effect of 3.1, 60 and 6.5 %, respectively, by the parent VQL8. On chromosome 7, the QTL located near the loci umc1567 and bnlg1305 had LOD of 3.2 with R^2 of 6.5 % and an additive effect of 64 % by the parent VQL8. On chromosome 5, the QTL was located near the locus ZmASK3, which had LOD of 2.7 with R^2 of 5.1 % and an additive effect of 3.5 % by the parent VQL8. These loci were developed from the aspartate kinase 2 candidate gene involved in the conversion of aspartate to aspartate 4-semialdehyde, which further produced lysine. Similar results were obtained with replication 1 (R1) and replication 2 (R2) data, with the exception that in R2 none of the QTL was detected on chromosome 5 (Table 3). The QTLs on chromosome 9 with replication 2 showed more LOD value (5.1, 5.2 and 4.2, respectively) and phenotypic variance (28.8 %) than the replication 1 and average tryptophan results. Bulk analysis of ten genotypes having high tryptophan content and ten having low tryptophan content resulted in identification of 3 SSR

loci (UMC2160, UMC1016 and Phi022) out of the 27 polymorphic between the bulk samples. These three primers were then genotyped against the individual genotypes and also found to be polymorphic between high and low tryptophan-containing genotypes and were considered to be linked to the tryptophan content. The loci UMC2160 and UMC1016 were located on bin location 7.01 and 7.02, respectively, whereas Phi022 was on chromosome 9. The SSR marker Phi022 was found to be linked to QTL for tryptophan by both bulk segregant and linkage mapping analysis.

Discussion

The nature, location and relationship of *opaque2* modifier genes in most of the QPM germplasm are not clear. Many of these genes had a complex system of genetic control and spread throughout all the ten chromosomes with dosage effects and cytoplasmic effect (Belousov 1987). In our earlier study, we found that two QPM inbreds, derived from the same parent (CM145), significantly differed in tryptophan content to a large extent (VQL2, 0.54; VQL8, 0.94), which may be due to the influence of *o2* modifiers. It was thought that these two inbreds would be ideal parents for developing a mapping population for identifying QTLs for *o2* modifiers. The initial goal of this study was to investigate the QTLs influencing the *o2* modifiers for tryptophan content in QPM. Very low polymorphic markers between these two parents were observed, since they are isogenic. Significant variability in the tryptophan content among the $F_{2:3}$ population existed. Maize generally constitutes 0.25–0.50 % of tryptophan in normal maize endosperm and is increased to 0.50–1.1 % in the endosperm of the *o2* mutants (QPM). The tryptophan content among the $F_{2:3}$ population varied to a large extent between 0.16 and 0.91 %. The correlation coefficient between protein and tryptophan data showed negative correlation of 0.50. Till now, there has been no report on the identification of *Opm* influencing the tryptophan

content using a QPM × QPM population, and present study is the first report.

Based on simple interval mapping using QTL cartographer, six QTLs associated with tryptophan content were identified on chromosomes 5, 7 and 9. On chromosome 5, the QTL was detected near the ZmASK3 candidate gene-based locus, which was also found with the CIM analysis using Qgene software. On chromosome 7, there was only one QTL which was closely linked to the marker bnlg1305 (bin 7.03) at an LRS value of 7.35 as also confirmed by the QGene software. On chromosome 9, three QTLs were found to be significant, at 0.01 and 0.1 % level of significance. The three QTLs were found to be near the loci umc1430 (bin 9.02), phi022 (bin 9.03) and umc2207 (bin 9.06). The results obtained by SIM were well supported and correlated by the CIM method of regression analysis. However, Holding et al. (2008) found seven QTLs based on simple interval mapping. This showed that chromosomes 7 and 9 form the major hub of o2 modifiers.

Based on the CIM analysis, five QTLs were identified on chromosomes 5, 7 and 9, accounting for 38.6 % of the phenotypic variation. However, Holding et al. (2008) found 75 % phenotypic variance from seven QTLs. In our study, out of these five QTLs, two on chromosome 9 with high LOD values of 4.8 and 4.5 explained a greater amount of phenotypic variance (Table 3). The QTL with LOD value of 4.5 was in the marker interval of umc1430 (9.02)–bnlg1401 (9.03), nearest to bnlg1401 (Fig. 2), which explained the phenotypic variance of 9.1 %. The second QTL lies in the marker interval of bnlg1401–phi022, nearest to phi022, with an LOD value of 4.2 and R^2 of 8.4 %. These findings were well supported by earlier reports, where Holding et al. (2008) found QTLs at an LOD of around 3.0 and R^2 of 6 % from the F2 population. Similarly, Yang et al. (2005) reported QTLs linked to *opaque-16* (*o16*) gene influencing the lysine content by the markers umc1121 and umc1141 on chromosome 8, with LOD of 2.08 and R^2 of 6.7 %, in the $F_{2:3}$ population. These results showed that the QTLs found in the present study were significant and may be influencing the *o2* modifiers for tryptophan content. The marker Phi022 observed to be linked as deduced by the bulk analysis during the present study was also reported to be linked to a QTL from the bulk segregant analysis study for *opaque2* modifiers (vitreousness) by Holding et al. (2008). The SSR marker Phi022 might be playing an important role in influencing tryptophan and vitreousness as evidenced from our results (by both bulk segregant and linkage mapping analysis) and earlier studies (Gibbon et al. 2003). This showed that vitreousness and amino acid composition of the maize endosperm may be positively correlated. However by using $F_{2:3}$ populations, they found seven QTLs on chromosomes 1, 5, 7, 9 and 10. In our study, we also found five QTLs on

chromosomes 5, 7 and 9 and polymorphic primers on five chromosomes (1, 5, 7, 9 and 10). These results showed that the chromosomes 5, 7 and 9 might be a major hub of *opaque*2 modifiers. The SSR locus *phi*022 was observed to be linked to the Waxy1 (*Wx1*) gene encoding for granule - bound starch synthase 1 candidate gene. Gibbon et al. (2003) previously reported a link between starch structure and vitreousness in QPM. In mature kernels, alteration in starch structure was reported to be associated with interconnections between starch granules that resulted in a vitreous kernel phenotype. Hence, the loci mapped during the present study will be highly useful for further studies to improve the lysine and tryptophan contents and the vitreousness of maize endosperm. *Opaque2* modifiers generally influence several traits such as tryptophan and lysine contents of maize endosperm, free amino acids content, vitreousness, hardness and density of the maize kernel. The *waxy1* gene is positively correlated with the vitreousness of the maize endosperm and thus may be involved in more accumulation of lysine and tryptophan contents of maize endosperm. These results indicated that the region around the locus *phi*022 was expected to highly influence the *opaque*2 modifiers. Holding et al. (2011) characterized *o2* modifier QTLs and candidate genes in the recombinant inbred lines for density, vitreousness and breakage mean. They also identified a QTL near the loci *phi*022, which is also closely linked to the *o2* modifiers and waxy1 gene. In the present study, the third QTL on chromosome 9 lies in the marker interval of umc2346–umc2207, nearest to umc2207 at an LOD of 4.8 and R^2 9.5 %, which was the highest R^2 value among all the QTLs identified during the present study (Table 3). The marker umc2207 located on the bin 9.06 was very closely linked to this QTL and explained the high R^2, which may be very useful in the marker-assisted breeding programs. This QTL was expected to be close to the candidate gene EIN3-like (EIL) that maps on the bin location 9.05 and found to be up-regulated for its vitreousness. Holding et al. (2008) found a QTL near the marker umc1016; however, in the present study this marker was not linked to the tryptophan content. This may be because of the genotypes used in the studies.

Identification of SSR markers in candidate genes has become a novel approach for association mapping and QTL mapping and their effective utilization in future crop improvement programs through MAS. Among the candidate gene-based SSR markers, four loci designed from the candidate genes of lysine pathway were polymorphic and mapped to chromosome 5 (ZmASK2 and ZmASK3 for aspartate kinase 2) and chromosome 10 (ZmHSDH1 and ZmHSDH5 for AK–homoserine dehydrogenase). However, marker ZmASK3 designed from the aspartate kinase 2 (ask2) gene was observed to be linked to a QTL on chromosome 5 with LOD of 2.7 at a phenotypic variance of

5.1 %. The maize mono-functional Ask2 encodes a Lys-sensitive AK, which overproduces Thr, Lys and Met Hibberd and Green 1982; Dotson et al. 1990; Muehlbauer et al. 1994). In maize, candidate genes associated with lysine metabolism were identified by Wang et al. (2007). They found Lys-sensitive Asp kinase to be a better candidate gene for the quantitative trait locus affecting free amino acid content. Hence, the development of more candidate gene-based SSR loci will be quite useful in identifying QTLs for *opaque*2 modifiers influencing the lysine and tryptophan amino acid pathway. The activity of Asp kinase (AK) is up-regulated by *o2* (Brennecke et al. 1996); it is an important enzyme involved in the synthesis of several amino acids, including Thr, Lys, Met and Leu (Bryan 1990). The SSR markers, ZmHSDH1 and ZmHSDH5, developed from the AK–homoserine dehydrogenase gene were mapped on chromosome 10, but could not be linked to any QTLs. This may be due to the less number of markers found polymorphic between the parents. So, development of more candidate gene-based SSRs for the lysine and tryptophan pathway will be useful for the detection of effective QTLs.

On chromosome 7, only one QTL near the marker bnlg1305 (bin7.03) at LOD of 3.2 with phenotypic variance of 6.5 % was observed. The SSR marker bnlg1305 was reported to be near the 27-kDa δ-zein protein. Holding et al. (2011) also found QTL for *opaque*2 modifiers on chromosome 7 near the loci umc1305 (7.03), close to the zein protein. Saturation of the polymorphic loci in this region will help in locating more highly effective QTLs. The results of these experiments suggest that several different genes and metabolic pathways can contribute to the formation of more nutritious endosperm in QPM. Whether the candidate genes identified in the earlier studies are sufficient to find the *opaque*2 modifiers is unclear. At this point, it is difficult to assign importance to a particular gene or pathway, although a case can be made for the 27-kDa γ-zein protein (Woo et al. 2001). Bulk analysis of ten genotypes having high tryptophan content and ten having low tryptophan content resulted in identification of 3 linked markers (UMC2160, UMC1016 and Phi022) out of the 27 polymorphic markers. The loci UMC2160 on the bin location 7.01 is close to the *opaque*2 loci umc1066, which can be considered as *opaque*2 modifiers. The SSR loci UMC1016 was also identified as QTL by Holding et al. (2008) through bulk segregant analysis and by several other studies (Yang et al. 2005). This locus is reported to be closely linked to the kinase-associated protein phosphatase 1 (kpp1) gene on the bin location 7.02, which is also near the loci *opaque*2 and 27KDa gamma *zein* protein (http://www.maizegdb.org). These results indicate that the loci UMC2160, UMC1016 and Phi022 play an important role in the nutritional quality of maize endosperm.

The recombinant inbred lines (RIL) population using QPM × QPM inbreds will be very useful in high-resolution mapping for *o2* modifiers influencing the tryptophan content. In addition, identification of new candidate genes involved in amino acid pathways and their use in mapping will help obtain high-resolution map and mapping of amino acid modifiers. Though we used a large number of SSR loci for parental polymorphism, we could not get sufficient polymorphic loci, since both parents were isogenic lines differing in tryptophan content. So, SNP-based genotyping will be another future strategy for dissecting the *opaque*2 modifiers. The present study resulted in the identification of highly significant QTLs on chromosomes 5, 7 and 9 with high LOD values and phenotypic variance. The QTLs reported will be highly useful in studies such as marker-assisted selection and breeding of QPM maize. Based on the mapping results, one QTL near the marker umc2207 (LOD of 4.8 and R^2 9.5 %) and another near the marker Phi022 (LOD 4.2 and R^2 8.4 %) on chromosome 9 were found to be significant and can be used in marker-assisted breeding programs.

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

References

Babu BK, Agrawal PK, Mahajan V, Gupta HS (2009) Molecular and biochemical characterization of short duration quality protein maize. J Plant Biochem Biotech 18(1):93–96

Babu BK, Agrawal PK, Gupta HS, Bhatt JC, Anil Kumar (2012) Identification of candidate gene-based SSR markers for lysine and tryptophan metabolic pathways in maize (*Zea mays*). Plant Breed 131:20–27

Barret P, Brinkman M, Dufour P, Murigneux A, Beckert M (2004) Identification of candidate genes for in vitro androgenesis induction in maize. Theor Appl Genet 109:1660–1668

Belousov AA (1987) Genetic analysis of modified endosperm texture in opaque2 maize. Soviet Genetics 23:459–464

Brennecke K, Souza Neto AJ, Lugli J, Lea PJ, Azevedo RA (1996) Aspartate kinase in maize mutants *ask1-lt19* and *opaque-2*. Phytochemistry 41:707–712

Bryan JK (1990) Advances in the biochemistry of amino acid biosynthesis. In: Milfin BJ, Lea PJ (eds) The biochemistry of plants, vol 16. Academic Press, New York, pp 161–195

Chin ECL, Senior ML, Shu H, Smith JSC (1996) Maize simple sequence repetitive DNA sequences: abundance and allele variation. Genome 39:866–873

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138:964–971

Dotson SB, Frish DA, Somers DA, Gengenbach BG (1990) Lysine-insensitive aspartate kinase in two threonine-overproducing mutants in maize. Planta 182:546–552

Geevers HO, Lake JK (1992) Development of modified opaque-2 maize in south Africa. In: Mertz Et (ed) Quality protein maize. American Association of Cereal Chemists, St. Paul

Gibbon BC, Wang X, Larkins BA (2003) Altered starch structure is associated with endosperm modification in Quality Protein Maize. Proc Natl Acad Sci USA 100(26):15329–15334

Hibberd KA, Green CE (1982) Inheritance and expression of lysine plus threonine resistance selected in maize tissue culture. Proc Natl Acad Sci USA 79:559–563

Holding DR, Hunter BG, Chung T, Gibbon BC, Ford CF, Bharti AK, Messing J, Hamaker BR, Larkins BA (2008) Genetic analysis of opaque2 modifier loci in quality protein maize. Theor Appl Genet 117:157–170

Holding DR, Hunter BG, Klingler JP, Wu S, Guo X, Gibbon BC, Wu R, Schulze JM, Jung R, Larkins BA (2011) Characterization of opaque2 modifier QTLs and candidate genes in recombinant inbred lines derived from the K0326Y quality protein maize inbreds. Theor Appl Genet 122:783–794

Kosambi DD (1944) The estimation of the map from the recombination values. Ann Eugen 12:172–175

Manly KF, Cudmore RH Jr, Meer JM (2001) Map Manager QTX, cross-platform software for genetic mapping. Mamm Genome 12:930–932

Mertz (1992) Quality protein maize. American Association of Cereal Chemists, St. Paul

Muehlbauer GJ, Gengenbach BG, Somers DA, Donovan CM (1994) Genetic and amino acid analysis of two maize threonine-overproducing, lysine-insensitive aspartate kinase mutants. Theor Appl Genet 89:767–774

Munck L (1992) The case of high lysine barley breeding. In: Shewry P (ed) Barley, genetics, biochemistry, molecular biology and biotechnology. CAB International, Wallingford

Nelson JC (1997) QGENE: software for marker-based genomic analysis and breeding. Mol Breed 3:239–245

Prasanna BM, Vasal SK, Kassahun B, Singh NN (2001) Quality protein maize. Curr Sci 81:1308–1319

Senior ML, 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

Shen B, Carneiro N, Torresjerez I, Stevenson B, McCreey T, Helentjaris T, Baysdorfer C, Almira E, Ferl RJ, Habben JE, Larkins B (1994) Partial sequencing and mapping of clones from two maize cDNA libraries. Plant Mol Biol 26:1085–1101

Vasal SK, Villegas E, Bjarnason M, Gelaw B, Goertz P (1980) Genetic modifiers and breeding strategies in developing hard endosperm opaque2 materials. In: Pollmer WG, Phillips RH (eds) Quality traits for maize for grain and silage use. Martinus Nijhoff, London

Wang X, Larkins BA (2001) Genetic analysis of amino acid accumulation in opaque-2 maize endosperm. Plant Physiol 125:1766–1777

Wang S, Basten CJ, Zeng ZB (2002) Windows QTL Cartographer. WinQTLCart V2.0

Wang X, Jose A, Bryan CG, Lopez-Valenzuela JA, Gakiere B, Galili G, Larkins BA (2007) Characterization of monofunctional aspartate kinase genes in maize and their relationship with free amino acid content in the endosperm. J Exp Bot 58(10):2653–2660

Woo YM, Hu DWN, Larkins BA, Jung R (2001) Genomics analyses of genes expressed in maize endosperm identifies novel seed proteins and clarifies patterns of zein gene expression. Plant Cell 13:2297–2317

Yang W, Zheng Y, Zheng W, Feng R (2005) Molecular genetic mapping of a high-lysine mutant gene (opaque-16) and the double recessive effect with opaque-2 in maize. Mol Breed 15:257–269

Young VR, Scrimshaw NS, Pellet P (1998) Significance of dietary protein source in human nutrition: animal and/or plant proteins? In: Waterlow JC, Armstrong DG, Fowden L, Riley R (eds) Feeding a world population of more than eight billion people. Oxford University Press in associate with Rank Prize Funds, NewYork