MARKER DEVELOPMENT FOR QUALITY PROTEIN MAIZE BREEDING AND AN INTERACTION STUDY BETWEEN Opaque-2 AND Ask2 GENES

Martha Hernández-Rodríguez1,*, Raman Babu2, Debra J. Skinner2, Natalia Palacios-Rojas2, Marina C. M. Martins3, J. Jesús García-Zavala1, Ricardo Lobato-Ortíz1, Amalio Santacruz-Varela1, Fernando Castillo-González1, Yunbi Xu2 and Boddupalli M. Prasanna2

1Colegio de Postgraduados, Campus Montecillo, PREGEP-Genética, Montecillo, Texcoco, Estado de México, México. 2International Maize and Wheat Improvement Center, Mexico City, México. 3Max-Planck-Institut für Molekulare Pflanzenphysiologie, Potsdam-Golm, Germany.

*Corresponding author (hernandez.martha@colpos.mx)

SUMMARY

Quality Protein Maize (QPM) kernels contain twice the amounts of lysine and tryptophan compared to normal corn kernels. Although the opaque-2 (o2) mutation is the underlying cause of this beneficial change, other genes such as Aspartate kinase-2 (Ask2) affect the amino acid content in the endosperm to a lesser degree. To date, reports on the interaction between both loci are scarce and there are no high-throughput assays for the identification of the alleles of these genes. The objectives of this research were: 1) to study the interaction between the o2 and Ask2 genes with respect to the accumulation of amino acids in the endosperm in an F2 population, 2) to identify conserved SNPs in the o2 gene that can be used as markers, 3) to estimate the frequency of an SNP of Ask2 associated with the accumulation of lysine in the endosperm in CIMMYT germplasm, and 4) to develop high-throughput marker assays for these SNPs. The interaction study showed a preponderant effect of o2 on the accumulation of 11 amino acids (P ≤ 0.01). Ask2 appeared only to act with o2 to enhance marginally lysine, histidine and methionine levels in the double recessive homozygotes. Sequencing of amplicons at the o2 locus led to the identification of an SNP in exon 1 that discriminated all QPM (C) from non-QPM (T) genotypes. Validation of this SNP through KASP™ indicated that was 92 % assertive in differentiating the double recessive homozygotes. The secuenciación de amplicones en el locus o2 condujo a la identificación de un SNP en el exón 1 que discriminó todos los genotipos QPM (C) de los genotipos no QPM (T). La validación de este SNP a través de los ensayos KASP™ indicó que fue 92 % asertivo en la diferenciación de los genotipos o2. In contrast, the frequency of the SNP of Ask2 in the germplasm QPM of the CIMMYT was low; however, an SSCP marker developed using this SNP detected five variants indicating that other unknown base changes may confer positive lysine-increasing responses. These markers could aid the marker-assisted selection of QPM cultivars.

Index words: Zea mays, Ask2, marker-assisted selection, Opaque-2, SNP.

RESUMEN

Los granos de maíz de alta calidad proteica (QPM) contienen el doble de lisina y triptófano en comparación con los granos de maíz normales. Aunque la mutación opaco-2 (o2) es la causa subyacente de este cambio beneficioso, otros genes como la aspartato quinasa-2 (Ask2) afectan en menor grado el contenido de aminoácidos en el endospermo. Hasta el momento, los informes sobre la interacción entre ambos loci son escasos y no existen ensayos de alto rendimiento para la identificación de los alelos de estos genes. Los objetivos de esta investigación fueron: 1) estudiar la interacción entre los genes o2 y Ask2 con respecto a la acumulación de aminoácidos en el endospermo de una población F2, 2) identificar SNPs conservados en el gen o2 que puedan ser utilizados como marcadores, 3) estimar la frecuencia de un SNP de Ask2, asociado con la acumulación de lisina en el endospermo, en germoplasma del CIMMYT, y 4) desarrollar ensayos de marcadores de alto rendimiento para estos SNPs. El estudio de interacción mostró un efecto preponderante del o2 sobre la acumulación de 11 aminoácidos (P ≤ 0.01). Al parecer, Ask2 solamente actúa con o2 para mejorar marginalmente los niveles de lisina, histidina y metionina en los homocigotos recesivos dobles. La secuenciación de amplicones en el locus o2 condujo a la identificación de un SNP en el exón 1 que discriminó todos los genotipos QPM (C) de los genotipos no QPM (T). La validación de este SNP a través de los ensayos KASP™ indicó que fue 92 % asertivo en la diferenciación de los genotipos o2. En contraste, la frecuencia del SNP de Ask2 en el germoplasma QPM del CIMMYT fue baja; sin embargo, un marcador SSCP desarrollado utilizando este SNP detectó cinco variantes, lo que indica que otros cambios de base desconocidos pueden conferir respuestas positivas de aumento de lisina. Estos marcadores podrían ayudar a la selección asistida en variedades QPM.

Palabras clave: Zea mays, Ask2, Opaco-2, selección asistida por marcadores, SNP.

INTRODUCTION

There are nine amino acids that human beings and mono-gastric animals cannot synthesize and need to be supplied through the diet (Galili et al., 2016). While most proteins of animal origin have suitable proportions of these amino acids, plant proteins are often suboptimal for human needs; for instance, most cereal proteins are poor in lysine and tryptophan, while those of legume origin are lower in methionine and other sulfur-containing amino acids (Mariotti and Gardner, 2019). In the developing world, prolonged consumption and over-reliance on cereal proteins can potentially lead to diseases such as pellagra, a niacin deficiency disease, and kwashiorkor, a form of protein-energy malnutrition that increases susceptibility to tuberculosis and gastroenteritis (Nuss and Tanumihardjo, 2011).

A mature corn kernel contains approximately 10 % protein and several mutants alter its amino acid profile (Zhang et al., 2011).
Molecular markers for QPM maize

The lysine value in QPM lines varies from 2.5 to 4.0 g per 100 g of protein (Vivek et al., 2008). Though o2 is the major determinant, a few minor modifier loci may likely work in unison with o2 to moderate the effect and alter the essential amino acid content in the homozygous (o2o2) recessive background (Vasal, 2001). In a QTL mapping study by Wang and Larkins (2001), free amino acid content in o2 maize lines was analyzed and shown to vary up to 12-fold in Oh545o2 and 3-10 times in OhS1Ao2 and W64Ao2 from non-o2 wild type counterpart. One of four loci identified was coincident with the Aspartate kinase 2 (Ask2) gene (Wang et al., 2001). This locus, located on chromosome 2, has a polymorphism that causes a single amino acid substitution in the C-terminal region of the ASK2 enzyme of Oh545o2 (Wang et al., 2007). Another study using a RIL population from the cross between B73o2 and CML161, identified QTLs for traits related to amino acid content on chromosomes 7 and 8, explaining up to 39% of the observed phenotypic variation (Gutiérrez-Rojas et al., 2010). These studies clearly established the existence of additional modifier loci that could favorably increase accumulation of amino acid levels in QPM germplasm. Of the several loci reported, the role of Ask2 is prominent and the responsible polymorphism has been identified and characterized (Wang et al., 2007).

The cloning and characterization of the O2 locus, located on chromosome 7 (Hartings et al., 1989; Schmidt et al., 1990), followed by detection of three SSR markers viz. phi057, phi112 and umc1066 (Chin et al., 1996) led to effective differentiation of the O2 and o2 alleles in elite QPM and non-QPM germplasm (Babu et al., 2005). These o2-specific SSR markers are the foundation for marker-assisted selection (MAS) in QPM breeding (Babu et al., 2005; Gupta et al., 2013; Pukalenthy et al., 2020); however, they are not enough to achieve the full effectiveness of molecular breeding for QPM genotypes. Marker umc1066 is easily visualized on agarose gels but it is commonly not polymorphic, at least in CIMMYT breeding populations; phi057 usually requires the use of polyacrylamide gels, while phi112, which is based on a deletion in the promoter region, is a dominant marker, and hence, cannot be used in the identification of heterozygotes in F1/BC generations (Babu and Prassana, 2014). These shortcomings, coupled with low throughput and time consuming of gel-based SSR analysis, result in a major hindrance to routine and large-scale adoption of marker technology in the breeding pipelines (Xu and Crouch, 2008). This requires the identification of SNP markers, which have propelled plant molecular breeding because of their ease of automation, high throughput, low cost and high read accuracy (Rasheed et al., 2017). Though the functional SNP conferring favorable response in Ask2 has been reported (Wang et al., 2007), high throughput assays are not yet available and the interaction between o2 and Ask2 loci, involving breeding germplasm, has not been documented. The objectives of this study were: 1) to evaluate the interaction between o2 and Ask2 in an F2 population derived from tropical maize germplasm and validate synergistic effect, if any, for enhanced amino acid accumulation in the endosperm, 2) to identify widely conserved SNPs within the o2 gene that can effectively discriminate the alleles in a range of elite genetic backgrounds, 3) to estimate the frequency of favorable Ask2 allele in CIMMYT germplasm and, 4) to develop medium to high throughput marker assays to facilitate the implementation of a comprehensive MAS scheme for rapid development of QPM germplasm.

Materials and Methods

Interaction analysis

Nine genotypic classes were determined after genotyping 400 seeds of an F2 population resulting from a cross between CML172 (QPM) and CML323 (non-QPM). Genotyping was carried out with molecular markers for o2 and Ask2 using DNA from endosperm. DNA extraction followed a protocol based on CTAB-laurolsarcosine described by Gao et al. (2008). To genotype o2, SSR marker phi057 was amplified by PCR reactions of 15 μL containing 1X Taq buffer, 2.5 mM MgCl2, 150 μM each dNTP, 0.3 μM of each primer (forward 5′-CTCATCAGGCGCTGGTCCAT-3′ and reverse 5′-CAGTCGCAAGAAACCGTTGCC-3′), 1 U homemade Taq polymerase and 20 ng of DNA. The PCR program consisted of an initial denaturation of 1 min at 94 °C followed by 30 cycles of 30 s at 94 °C, 1 min at 58 °C, 1 min at 72 °C and a final extension of 5 min at 72 °C using a MJ Tetrada thermalcycler (MJ Research, St. Bruno, Quebec.
Canada). Three microliters of each PCR product was migrated in 6 % acrylamide:bis-acrylamide (29:1) gels at 250 V for 2 h in 1X TBE (90 mM Tris-borate, 2 mM EDTA pH 8.0) in a MGV-116-33 system (CBS Scientific®, San Diego, California, USA). The reference of molecular weight was φX174/HaeIII. After electrophoresis, the gels were silver-stained according to CIMMYT (2005).

To genotype Ask2, primers flanking the functional SNP of Ask2 (Wang et al., 2007) were designed to amplify a product of 151 pb. PCR mix included 1X Taq buffer, 2.0 mM MgCl₂, 166 µM each dNTP, 0.1 µM of each primer (forward 5´-CTTGGTGGTCCATGACAGTG-3´ and reverse 5´-CTAATGGCTGTGGATTGTGC-3´), 1 U homemade Taq polymerase and 50 ng of DNA in a reaction volume of 15 µL. The PCR program consisted of one cycle of 94 °C for 2 min, 30 cycles at 94 °C for 1 min, 58°C for 40 s, 72°C for 1 min, and a final extension at 72 °C for 5 min performed in a MJ Tetrad thermalcycler (MJ Research, St. Bruno, Quebec, Canada). PCR products were resolved as single-stranded conformation polymorphisms (SSCP) in 16 % polyacrylamide gels applying a modified Laemmli’s protocol that consisted of preparing each layer of the gel (stacking and resolving) without SDS (Laemmli, 1970). The PCR products were diluted with 5 µL of loading buffer (10 mM NaOH, 95 % formamide, 0.05 % bromophenol blue, 0.05 % xylene cyanol), denatured during 10 min at 94 °C and quick quenching on ice. Five µL were run for 15-16 h at 150 V and 8 °C in a MGV-116-33 system (CBS Scientific®, San Diego, California, USA) using 1X TBE as running buffer. The gels were stained with 1X SYBR Gold (Life Technologies, Carlsbad City, California, USA) for 20 min and documented by scanning in a Typhoon 8600 imager (Amersham Biosciences Ltd., Chicago, Illinois, USA).

After genotyping, each marker was scored by hand and seeds with the same genotype were bulked to form a genotypic class. Only 20 seeds per class were used to quantify 11 amino acids by high-performance liquid chromatography (HPLC). Ethanolic extracts by HPLC (consisting of a pump P680, Dionex, autosampler 465 and fluorimetric detector SFM 25, Kontron, Eching, Germany) after derivation of the primary amino group with OPA reagent [25 mg OPA acid, mixed and kept for 2.5 min at 4 °C to allow complete derivation. The extract was injected into a pre-column (Hypersil 10 µm, Phenomenex, Torrance City, California, USA) before separation on a C18 column (Hypersil ODS, 150 x 4.6 mm, 3 µm particle size, Thermo Fisher Scientific, Bonn, Germany) maintained at constant 18 °C using a thermostatic column holder (560-CIL, ERC, Riemerling, Germany) at a flow rate of 0.8 mL min⁻¹. Eluent A was 8 mM Na₂HPO₄ (pH 6.8) containing 0.4 % (v/v) tetrahydrofuran. Eluent B was a mixture of 250 mL 40 mM Na₂HPO₄ (pH 6.8), 150 mL methanol and 105 mL acetonitrile. Two different gradients determined major and minor amino acids. For major amino acid columns were equilibrated for 10 min with 10 % A and separation conditions were as follows: 0-5 min, isocratic elution with 0 % B; 5-15 min, linear gradient from 0 % to 15 % B; 15-15.5 min, linear gradient from 15 to 100 % B; 15.5-21 min 100 % B. For minor amino acids, columns equilibrated for 10 min with 100 % A and separation conditions were as follows: 0-5 min, isocratic elution with 0 % B; 5-16 min, linear gradient from 0 to 15 % B; 16-26 min, linear gradient from 15 to 50 % B; 26-42 min, linear gradient from 50 to 60 % B; 42-46 min from 60 to 100 % B; 46-52 min 100 % B. Amino acids were identified and quantified by comparison with original standards (Sigma-Aldrich) using Chromleven software (version 6.8, Dionex). Individual analyses of variance per amino acid and principal component analysis were applied using PROC GLM and PRINCOMP of SAS, respectively (SAS Institute, 2002).

**Polymorphism identification**

Fifteen elite inbred lines were chosen to search for polymorphisms within the o2 gene. Seven inbred lines were QPM, of diverse agro-ecological zones, with excellent kernel modifying capacity and of extensive use in the QPM conversion programs. The other eight lines were non-QPM, with wide adaptability and superior performance in Asia, Latin America and East African highlands (Table 1). Leaves of four seedlings per line were used separately to extract genomic DNA following a protocol based on CTAB for small-scale isolation of high quality DNA (CIMMYT, 2005). Two sets of primers were designed to amplify two genomic fragments of o2 (477 and 530 pb) and to identify regions with high degree of conservation. The reference sequence X15544, which belongs to the zein regulatory gene o2 of Zea mays (http://www.ncbi.nlm.nih.gov/), plus six cDNA sequences of o2 gene (AJ41297, AJ491298, AJ491299, AJ491300, M29411 and X16618) were used as templates for the design of primers O2F1: 5´-CTAGTGTTTGCTTCTCCCT-3´ and O2R1: 5´-TCCTCAGTGTCATGACTCC-3´ and O2F3: 5´-TGTTGTGACCTCAGATCAACG-3´ and O2R3: 5´-TTCCAGTTCTTCTCCCTTCC-3´. The PCR reactions were optimized to produce single bands without smear by applying a lower number of cycles and a large volume of reaction (30 µL). In total, two PCR fragments per seedling were amplified and purified with Qiagen spin columns (Qiagen®, Hilden, Germany), and cloned into pGEM®-T Easy Vector (Promega, USA) according to manufacturer recommendations. Recombinant plasmids were selected...
MOLECULAR MARKERS FOR QPM MAIZE

Table 1. Maize inbred lines used to identify polymorphisms in the o2 locus.

<table>
<thead>
<tr>
<th>Grain color and adaptation zone</th>
<th>QPM line</th>
<th>Non-QPM line</th>
</tr>
</thead>
<tbody>
<tr>
<td>White, tropical lowland</td>
<td>CML159</td>
<td>CML343</td>
</tr>
<tr>
<td></td>
<td>CML491</td>
<td>CML495</td>
</tr>
<tr>
<td></td>
<td>CML502</td>
<td>CL RCW22</td>
</tr>
<tr>
<td></td>
<td>CML492</td>
<td></td>
</tr>
<tr>
<td>Yellow, tropical lowland</td>
<td>CML161</td>
<td>CML348</td>
</tr>
<tr>
<td></td>
<td>CML165</td>
<td></td>
</tr>
<tr>
<td>White, sub-tropical</td>
<td>CML176</td>
<td></td>
</tr>
<tr>
<td>White, highland</td>
<td>CML244</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CML349</td>
<td></td>
</tr>
<tr>
<td>White, Ethiopian highland</td>
<td>F7287</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A7018</td>
<td></td>
</tr>
</tbody>
</table>

after transforming DH5α competent cells of E. coli by heat-shock according to CIMMYT (2005). Plasmid DNA extraction was with Wizard® Plus SV Minipreps DNA Purification System (Promega, USA) and verification of inserts was by digestion with EcoR1. Two plasmids of each genotype were sequenced at CINVESTAV facilities (Guanajuato, México), aligned using Clustal X version 1.8 (Thompson et al., 1997) and refined by hand.

Markers assays validation

A panel of 88 inbred lines from the Maize Program or CIMMYT Genbank was used to validate polymorphisms using genomic DNA as template. DNA extraction protocol was the same as that used for polymorphism identification. Two SNPs were assayed at LGC Biosearch Technologies (https://www.biosearchtech.com/) using the KASP™ chemistry. One of them was the o2-specific SNP, which differentiated QPM and normal lines and the other was the Ask2-specific SNP amplified to determine genotypic classes. This last one was also screened as SSCP marker to explore the natural variation of Ask2 and to estimate the frequency of the favorable allele in CIMMYT germplasm.

RESULTS

Interaction analysis between o2 and Ask2

The interaction between o2 and Ask2 was studied in a tropical germplasm background by developing a segregating population from a cross between CML172, which is a QPM line possessing favorable alleles for both o2 and Ask2 loci (based on ‘2’ type SSCP conformer), and CML323, a non-QPM line. The interaction analysis was based on the concentration of 11 amino acids of nine genotypic classes (Table 2). Among these classes, the double recessive homozygous class (G1) contained significantly higher amounts of lysine, tryptophan and aspartic acid relative to the double dominant homozygous class (G5). In this regard, G1 had 2.3-fold more lysine, 1.3-fold more tryptophan and 5.3-fold more aspartic acid than G5. Lysine levels of G1 were also 60 % higher relative to G2, the recessive dominant class. A fast analysis using biplot analysis indicated that the first two components explained 96.24 % of the variance (Figure 1). PC1 explained 89.27 % of variation while PC2 explained 6.93 %. All vectors had a positive direction toward G1, G2 and G3 where o2 was recessive. No vector pointed toward genotypes with the dominant allele of o2 (G4 to G9). On the contrary, Ask2 locus on its own was not found to have any significant effect on any of the amino acids analyzed, but appeared to interact synergistically with o2 and it significantly enhanced lysine, histidine and methionine levels in the double recessive homozygous class (G1).

Polymorphism identification in the o2 locus

Sequencing the amplicons of o2 allowed to locate an important discriminatory SNP in the region of exon 1 (Figure 2). The sequencing included part of the 5' flanking region, four exons (1, 2, 3 and 4) and two introns (2 and 3). In total, a length of 1 kb of the genomic sequence was analyzed. 65 polymorphic sites were identified, and one polymorphism each 30 bp was averaged. In general, all QPM lines had the same sequence, indicating a possible common donor of o2 in CIMMYT lines.

Of the several polymorphisms that differentiated the QPM lines from non-QPM, the following four merit attention: 1) two SNP in intron 2 (positions 10,800,027 and 10,800,032 bp concerning to o2 reference sequence of Zea mays, ID 542375 of B73 RefSeq_v3) which defined three alleles with the haplotypes GT/AG/AT (donor/recipient/recipient), 2) two consecutive SNP in intron 2 (positions 10,800,038 and 10,800,039) which defined the alleles AT/TC/AC (donor/recipient/recipient), 3) two InDels in intron 2 (positions 10,799,978 to 10,799,982 and 10,799,988 to 10,799,990), which corresponded to 5-base and 3-base deletions in all QPM donor lines as well as for two non-QPM lines, and 4) an SNP of transition type in exon 1 (position 10,799,047 according to sequence ID 542375 of B73 RefSeq_v3, and position 1790 according to accession X15544 of the GenBank). The SNP corresponds to C in QPM lines and to T in non-QPM lines. This SNP was named 2-exon1-SNP, which causes an amino acid substitution (valine to alanine) and it significantly enhanced lysine, histidine and methionine levels in the double recessive homozygous class (G1).
Table 2. Amino acids concentration (nmol g⁻¹ FW) of kernels of the genotypic classes of o2 and Ask2 in CML172 × CML323.

<table>
<thead>
<tr>
<th>Genotypic class</th>
<th>o2</th>
<th>Ask2</th>
<th>Asp</th>
<th>Glu</th>
<th>Asn</th>
<th>Thr</th>
<th>His</th>
<th>Met</th>
<th>Trp</th>
<th>Phe</th>
<th>Ile</th>
<th>Leu</th>
<th>Lys</th>
</tr>
</thead>
<tbody>
<tr>
<td>CML172</td>
<td>1</td>
<td>1</td>
<td>4969.74</td>
<td>2759.53</td>
<td>4631.05</td>
<td>286.31</td>
<td>27.98</td>
<td>53.14</td>
<td>210.12</td>
<td>159.79</td>
<td>268.78</td>
<td>719.06</td>
<td></td>
</tr>
<tr>
<td>CML323</td>
<td>2</td>
<td>2</td>
<td>268.04</td>
<td>329.00</td>
<td>656.81</td>
<td>64.00</td>
<td>41.36</td>
<td>39.11</td>
<td>45.95</td>
<td>36.17</td>
<td>41.80</td>
<td>55.04</td>
<td>81.75</td>
</tr>
<tr>
<td>G1</td>
<td>1</td>
<td>1</td>
<td>3383.40</td>
<td>1203.85</td>
<td>1469.11</td>
<td>297.81</td>
<td>100.67</td>
<td>105.13</td>
<td>112.76</td>
<td>156.63</td>
<td>249.35</td>
<td>270.02</td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>1</td>
<td>2</td>
<td>3095.29</td>
<td>1294.23</td>
<td>2312.75</td>
<td>349.47</td>
<td>64.00</td>
<td>160.47</td>
<td>47.47</td>
<td>112.76</td>
<td>156.63</td>
<td>249.35</td>
<td>270.02</td>
</tr>
<tr>
<td>G3</td>
<td>1</td>
<td>H</td>
<td>1979.50</td>
<td>1108.96</td>
<td>1189.26</td>
<td>306.18</td>
<td>60.67</td>
<td>100.67</td>
<td>47.47</td>
<td>112.76</td>
<td>156.63</td>
<td>249.35</td>
<td>270.02</td>
</tr>
<tr>
<td>G4</td>
<td>2</td>
<td>1</td>
<td>409.54</td>
<td>393.07</td>
<td>351.05</td>
<td>179.56</td>
<td>68.63</td>
<td>31.10</td>
<td>37.67</td>
<td>48.73</td>
<td>62.50</td>
<td>129.35</td>
<td>124.24</td>
</tr>
<tr>
<td>G5</td>
<td>2</td>
<td>2</td>
<td>541.97</td>
<td>557.88</td>
<td>589.25</td>
<td>222.23</td>
<td>98.03</td>
<td>36.49</td>
<td>46.35</td>
<td>67.43</td>
<td>94.24</td>
<td>177.15</td>
<td>132.78</td>
</tr>
<tr>
<td>G6</td>
<td>2</td>
<td>H</td>
<td>460.78</td>
<td>366.26</td>
<td>374.36</td>
<td>180.02</td>
<td>104.14</td>
<td>38.35</td>
<td>49.94</td>
<td>48.73</td>
<td>62.50</td>
<td>129.35</td>
<td>124.24</td>
</tr>
<tr>
<td>G7</td>
<td>H</td>
<td>1</td>
<td>298.44</td>
<td>247.67</td>
<td>309.08</td>
<td>149.93</td>
<td>73.33</td>
<td>32.13</td>
<td>37.31</td>
<td>46.06</td>
<td>46.36</td>
<td>103.12</td>
<td>102.82</td>
</tr>
<tr>
<td>G8</td>
<td>H</td>
<td>2</td>
<td>471.34</td>
<td>342.96</td>
<td>385.82</td>
<td>170.18</td>
<td>95.32</td>
<td>39.26</td>
<td>38.24</td>
<td>50.65</td>
<td>72.13</td>
<td>128.60</td>
<td>126.01</td>
</tr>
<tr>
<td>G9</td>
<td>H</td>
<td>H</td>
<td>315.77</td>
<td>300.76</td>
<td>300.61</td>
<td>160.76</td>
<td>88.22</td>
<td>31.29</td>
<td>39.22</td>
<td>43.81</td>
<td>57.83</td>
<td>114.87</td>
<td>96.99</td>
</tr>
</tbody>
</table>

o2 classes

<p>| | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>G1,G2,G3</td>
<td>1</td>
<td>-</td>
<td>2819.40</td>
<td>1202.35</td>
<td>1657.04</td>
<td>317.82</td>
<td>170.72</td>
<td>69.61</td>
<td>94.82</td>
<td>141.74</td>
<td>187.52</td>
<td>241.55</td>
<td>355.55</td>
</tr>
<tr>
<td>G4,G5,G6</td>
<td>2</td>
<td>-</td>
<td>439.07</td>
<td>438.22</td>
<td>193.94</td>
<td>90.27</td>
<td>35.31</td>
<td>44.66</td>
<td>58.05</td>
<td>79.29</td>
<td>159.42</td>
<td>117.01</td>
<td>108.61</td>
</tr>
<tr>
<td>G7,G8,G9</td>
<td>H</td>
<td>-</td>
<td>361.85</td>
<td>297.13</td>
<td>331.83</td>
<td>159.99</td>
<td>85.62</td>
<td>34.23</td>
<td>38.25</td>
<td>46.84</td>
<td>58.77</td>
<td>115.53</td>
<td>110.28</td>
</tr>
</tbody>
</table>

Ask-2 classes

<p>| | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>G1,G4,G7</td>
<td></td>
<td>1</td>
<td>1363.79</td>
<td>614.86</td>
<td>709.74</td>
<td>209.10</td>
<td>124.25</td>
<td>54.54</td>
<td>60.04</td>
<td>86.59</td>
<td>109.13</td>
<td>166.45</td>
<td>209.59</td>
</tr>
<tr>
<td>G2,G5,G8</td>
<td></td>
<td>2</td>
<td>1369.53</td>
<td>731.69</td>
<td>1095.94</td>
<td>247.30</td>
<td>117.94</td>
<td>41.08</td>
<td>65.78</td>
<td>91.57</td>
<td>119.76</td>
<td>185.04</td>
<td>176.27</td>
</tr>
<tr>
<td>G3,G6,G9</td>
<td></td>
<td>H</td>
<td>918.69</td>
<td>591.99</td>
<td>621.41</td>
<td>215.36</td>
<td>104.42</td>
<td>43.44</td>
<td>51.91</td>
<td>68.48</td>
<td>96.68</td>
<td>165.02</td>
<td>195.31</td>
</tr>
</tbody>
</table>

Figure 1. Biplot of principal components of 11 free amino acids content in nine genotypic classes (G) from an F₂ population between CML172 (QPM) and CML323 (non-QPM). G1: o2/o2|ask2/ask2, G2: o2/o2|Ask2/Ask2, G3: o2/o2|Ask2/ask2, G4: O2/O2|ask2/ask2, G5: O2/O2|Ask2/Ask2, G6: O2/O2|ask2/ask2, G7: O2/ask2/ask2, G8: O2/ask2/Ask2, G9: O2/ask2/ask2.

Figure 2. The o2-exon1-SNP and its context sequence. The SNP is highlighted in yellow. Allele C discriminates to QPM lines (in italics). Allele T discriminates non-QPM lines (no italics). Name of each maize line plus its GenBank ID is to the left of the nucleotide sequence.
Validation of high throughput marker assays

Two high throughput KASP™ assays non gel-based were developed to facilitate large-scale MAS in pedigree/back-cross breeding programs involving QPM × non-QPM lines. One of these assays, designed to screen the \( o2\)-exon1-SNP, was firstly tested in 16 lines. The analysis of allele calls clearly grouped all genotypes into two clusters—one corresponding to allele C of QPM and another corresponding to allele T of non-QPM lines. Thus, this analysis was extended to a larger set of 88 lines in which 60 lines were QPM and the remaining non-QPM. Allele calls for \( o2\)-exon1-SNP differentiated two clusters again (Figure 3A), the cluster in red color corresponding to lines with the allele C and the cluster in blue color to lines with the allele T. The cluster with allele C included 53 QPM lines while the cluster with allele T comprised 35 lines, of which 28 were normal and seven were QPM. CML145, CML167, CML168, CML177, CML180, CML184 and CML188 were the QPM lines that possessed the ‘normal’ allele (T), contrary to expectation. Therefore, prediction value of \( o2\)-exon1-SNP was 92 % for the QPM trait.

The remaining assay was designed to test the functional polymorphism A (‘lysine enhancing’) of the C-terminal region of the ASK2 protein of the Oh545o2 maize line (Wang et al., 2007), which identified several lines possessing the favorable allele (Figure 3B). In this case, allele calls distinguished one cluster (in red) of seven inbred lines with the favorable allele A, of which five were QPM (CML148, CML161, CML162, CML179 and CML188) and two normal (CML328 and CML 370). The other cluster with unfavorable allele T (in blue) comprised 81 lines of which 55 were QPM and 26 non-QPM, indicating a low frequency of this allele in CIMMYT germplasm.

Natural variation of Ask2

Given that the favorable allele A at the Ask2 locus was only detected in five out of 60 QPM lines, to determine other possible functional variants of Ask2, a gel-based SSCP assay was developed and tested in 82 maize lines. Five distinct SSCP band patterns were unambiguously identified (Figure 4). SSCP variants were as follows: variant 1 in 60 lines (46 QPM and 14 normal), variant 2 in 11 lines (10 QPM and 1 normal), variant 3 in 5 lines (all QPM), variant 4 in 5 lines (all normal) and variant 5 in 1 line (QPM), suggesting that other polymorphisms could be involved in the high lysine level in the endosperm.

DISCUSSION

The interaction between \( o2\) and Ask2 in tropical germplasm background was studied in nine different genotypic classes resulting from an \( F_2\) segregating population. The single locus tests established the overriding effect of \( o2\) on most of the amino acids, irrespective of the genotypic constitution at Ask2 locus (Figure 1). On the contrary, Ask2 locus on its own was not found to have any significant effect on any of the amino acids analyzed but did interact synergistically with \( o2\) and significantly enhanced lysine, histidine and methionine levels in the

Figure 3. Genotyping cluster plot of KASP™ assays for \( o2\) (A, left) and Ask2 (B, right) SNP. Each data point represents an individual inbred line. Similar fluorescence values mean the same genotype. Red dots: homozygous for allele labeled with HEX, blue dots: homozygous for allele labeled with FAM, green dots: heterozygous, black dots: no template control.
double recessive homozygous class (G1).

The elevated levels of lysine in o2 endosperm are due to enhanced synthesis of lysine-containing proteins (Zhang et al., 2018) as well as high free lysine content (Wang and Larkins, 2001). Lysine is synthesized through a branch of aspartate pathway and is regulated by feedback inhibition loops, in which the accumulating lysine suppress the activity of two key enzymes, aspartate kinase (AK) and dihydrodipicolinate synthase (DHDPS) (Galili, 1995; Galili et al., 2016). The native mono-functional AK enzyme is a tetramer composed of two α- and two β-subunits, encoded by Ask1 and Ask2 genes, respectively (Dotson et al., 1989; 1990). Mutant alleles in either Ask1 or Ask2 loci result in AK enzyme that is less sensitive to feedback inhibition and lead to overproduction of lysine, threonine, methionine and isoleucine (Dotson et al., 1990). In maize, Ask2 explains a significant portion of variance for the free lysine levels in the o2 endosperm (Wang and Larkins, 2001), and thus, it represents an important target for enhancing free lysine levels in seeds. In this study, the mutation in the Ask2 locus that resulted in a favorable SSCP allele in CML172 is either insensitive or less sensitive to lysine inhibition and, thereby, contributed to increased accumulation of lysine and other aspartate pathway amino acids, similar to the mutant Ask2 allele of Oh545o2; however, it was apparent that the Ask2 locus on its own did not have any significant impact on the lysine or other amino acid levels in this investigation. Based on studies involving double mutants of Ask1 and o2, Brennecke et al. (1996) inferred that Ask1 is being directly or indirectly regulated by O2 because AK in these double mutants is less sensitive to lysine feedback inhibition than that of Ask1 mutation alone. Though there is no direct evidence for Ask2 being regulated by O2, such a possibility may partially explain the insignificant effect of Ask2 locus independent of o2.

Though the three o2-specific SSR markers offer an excellent opportunity for MAS, their use has a limited potential in plant breeding. In this investigation, an important discriminatory SNP in the exon 1 region of o2 was identified and an easy to use, high throughput assay, non-gel-based to facilitate large-scale MAS was developed. The lack of a perfect correlation between the underlying allelic state at this SNP locus and the QPM status indicates the non-causal nature of this variation, which however has a high predictive value for the QPM trait. Assays SSCP and KASP were designed for Ask2 which are able to screen a large collection of QPM and normal lines in the CIMMYT germplasm and to identify several lines possessing the favorable allele. Interestingly, the frequency of the unfavorable allele at the Ask2 locus was significantly higher in the QPM germplasm compared to normal endosperm, which could be due to the high selection pressure exerted by breeders during the process of QPM germplasm development through rigorous biochemical tests for high lysine and tryptophan. The higher free amino acid content in the o2 endosperm could either be due to increased lysine.
synthesis or loss of lysine ketoglutarate reductase (LKR) activity, which degrades lysine as endosperm matures (Arruda et al., 2000). Apart from Ask2, several genes have been postulated to play key roles in the enhanced lysine synthesis such as aspartate kinase-homoserine dehydrogenase-2 and dihydro dipicolinate synthase (Galili et al., 2016; Wang and Larkins, 2001). Some of such loci may likely be at play in those QPM lines that are rich in lysine but do not have the favorable Ask2 allele. Identifying and characterizing such loci in the high lysine QPM lines may provide further incremental gains on the enhanced amino acid levels conferred by o2 and Ask2.

The gel-based SNP detection assay developed in this investigation for the Ask2 locus combines reasonable throughput of modern gel electrophoresis platforms with robust discriminative power of the SSCP technique (Orita et al., 1989). The SSCP-SNP assay is efficient and sensitive enough for most of the PCR products in the size range of 100-750 bp, cost-competitive and requires no sophisticated instrumentation (Sunnucks et al., 2000). Notably, some of the lines (e.g. CML172) that were identified as possessing favorable allele at the Ask2 locus (‘2’ type), as per SSCP conformers, had indeed the unfavorable allele (T) according to KASP genotype calls. This discrepancy between SSCP and KASP genotypes in the Ask2 amplified region. Based on the results of these genetic analyses in the segregating F2 progenies, it was demonstrated that unknown base changes in the exon region of Ask2 may also confer positive responses (lysine enhancing) similar to the previously established functional SNP. It is indeed, possible to obtain higher genetic gains for enhanced lysine and tryptophan levels if the QPM lines identified in this investigation, that harbor favorable alleles (either A as per KASP assay or ‘2’ type SSCP conformer) at the Ask2 locus, are preferred as QPM donors in the new breeding programs and quickly becoming a compelling option for biplot analysis.

ACKNOWLEDGMENTS

The study was supported by the CGIAR Research Program Maize. First author acknowledges her PhD scholarship to Consejo Nacional de Ciencia y Tecnología (CONACYT) of Mexico. We also thank John Lunn from Max Planck Institute for Molecular Plant Physiology for his support for the free amino acid analysis and Dr. Gregorio Alvarado-Beltrán for biplot analysis.

BIBLIOGRAPHY


