Theor Appl Genet (2008) 116:213–221
DOI 10.1007/s00122-007-0660-8

ORIGINAL PAPER

Characterization of phytoene synthase 1 gene (*Psy1*) located on common wheat chromosome 7A and development of a functional marker

X. Y. He · Y. L. Zhang · Z. H. He · Y. P. Wu · Y. G. Xiao · C. X. Ma · X. C. Xia

Received: 11 May 2007 / Accepted: 27 September 2007 / Published online: 18 October 2007

Abstract Phytoene synthase (Psy), a critical enzyme in the carotenoid biosynthetic pathway, demonstrated high association with the yellow pigment (YP) content in wheat grain. Characterization of Psy genes and the development of functional markers for them are of importance for marker-assisted selection in wheat breeding. In this study, the full-length genomic DNA sequence of a Psy gene (*Psy-A1*) located on chromosome 7A, was characterized by in silico cloning and experimental validation. The cloned *Psy-A1* comprises six exons and five introns, 4,175 bp in total, and an ORF of 1,284 bp. A co-dominant marker, *YP7A*, was developed based on polymorphisms of two haplotypes of *Psy-A1*, yielding 194 and 231-bp fragments in cultivars with high and low YP content, respectively. The marker *YP7A* was mapped on chromosome 7AL using an RIL population from cross PH82-2/Neixing 188, and a set of Chinese Spring nullisomic–tetrasomic lines and ditelosomic line 7AS. *Psy-A1*, co-segregating with the STS marker *YP7A*, was linked to SSR marker *Xwmc809* on chromosome 7AL with a genetic distance of 5.8 cM, and explained 20–28% of the phenotypic variance for YP content across three environments. A total of 217 Chinese wheat cultivars and advanced lines were used to validate the association between the polymorphic band pattern and grain YP content. The results showed that the functional marker *YP7A* was closely related to grain YP content and, therefore, could be used in wheat breeding programs targeting of YP content for various wheat-based products.

Communicated by P. Langridge.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0660-8) contains supplementary material, which is available to authorized users.

X. Y. He · Z. H. He (✉) · Y. P. Wu · Y. G. Xiao · X. C. Xia (✉)
Institute of Crop Science, National Wheat Improvement Center/ The National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences (CAAS), Zhongguancun South Street 12, Beijing 100081, China
e-mail: xiaxianchun@caas.net.cn; zhhe@public3.bta.net.cn

Y. L. Zhang · C. X. Ma
College of Agronomy, Anhui Agricultural University, Hefei, Anhui 230036, China

Z. H. He
International Maize and Wheat Improvement Center (CIMMYT) China Office, c/o CAAS, Zhongguancun South Street 12, Beijing 100081, China

Introduction

Yellow pigment (YP) content is an important trait in wheat quality evaluation in both common and durum wheats. A bright yellow colour is desirable for yellow alkaline noodles, which are consumed in Japan and southeastern Asia, and thus increasing the YP content is beneficial in wheat cultivars for use in these regions (Kruger et al. 1992). In China, however, a bright white to creamy colour is preferred for white salted noodles, Chinese dry white noodles and steamed bread, requiring a low YP content in wheat grains (He et al. 2004). Therefore, developing wheat cultivars with either high or low YP content, depending on end-use products, is an important objective in wheat breeding programs.

Common wheat is an allohexaploid crop, with three homoeologous genomes of A, B and D, and each was constituted of seven chromosomes. A number of studies on YP content have been undertaken over recent decades. Although YP content is affected by environment (Miskelly 1984), it is mainly determined by genotype. Parker et al.

(1998) analyzed the inheritance of YP content using a single seed descent (SSD) population derived from the common wheat cross Schomburgk/Yarralinka and estimated its heritability at 0.67. A higher heritability was found in durum wheat, ranging from 0.88 to 0.95 (Clarke et al. 2006). YP content is mainly conferred by the genes located on homoeologous group 7 chromosomes. Using 150 SSD lines from Schomburgk/Yarralinka, Parker et al. (1998) found two major QTLs on chromosomes 7A and 3A, explaining 60 and 13% of the phenotypic variance, respectively. An AFLP marker, *Xwua26-7A.4*, closely linked to the QTL on chromosome 7A, was then converted into an STS marker (Parker and Langridge 2000). Mares and Campbell (2001) detected two QTLs associated with YP content on chromosomes 7A and 3B in a Sunco/Tasman mapping population. These accounted for 27 and 20% of the phenotypic variance, respectively. Kuchel et al. (2006) mapped a major QTL for flour yellowness b* on chromosome 7B, explaining 48 and 61% of phenotypic variance in different cropping seasons. Zhang et al. (2006) detected a major QTL controlling kernel YP content and flour yellowness b* on chromosome 7A, accounting for 12.9–37.6% of phenotypic variance across five environments. In durum wheat, likewise, major QTLs for YP content were found on chromosomes 7A and 7B (Elouafi et al. 2001; Pozniak et al. 2007). QTLs for YP content were also detected on homoeologous group 1 chromosomes (Ma et al. 1999), chromosomes 4A and 5A (Hessler et al. 2002), 2D and 4D (Zhang et al. 2006), and 2A, 4B and 6B (Pozniak et al. 2007), indicating multigenic control of YP content in wheat grain in addition to the major genes on homoeologous group 7 chromosomes.

Carotenoids are the main components of flour yellow pigment (Miskelly 1984). Adom et al. (2003) analyzed carotenoid contents in 11 wheat cultivars and found that lutein was the most abundant type, followed by zeaxanthin and β-cryptoxanthin, with significantly different contents in different cultivars. The biosynthetic pathway of lutein, zeaxanthin and β-cryptoxanthin involves more than ten enzymatic steps (Hirschberg 2001), among which the step catalyzed by phytoene synthase (Psy), dimerizing two geranylgeranyl pyrophosphate molecules, is assumed to be rate-limiting in the carotenoid biosynthesis (Lindgren et al. 2003). The maize *Y1* gene (*Psy1*) is associated with the accumulation of carotenoid in endosperm (Buckner et al. 1996). Palaisa et al. (2003) investigated the associations of *Psy1* and *Psy2* with endosperm yellowness, and found that *Psy1*, but not *Psy2*, exhibited a strong association with YP content of endosperm in maize. By means of database mining and ortholog-specific universal PCR, Gallagher et al. (2004) identified duplicated Psy genes (*Psy1* and *Psy2*) in 12 species (including common wheat) from eight subfamilies of the grass family. The accumulation of *Psy1* transcript was associated with carotenoid content in maize endosperm (Gallagher et al. 2004), thus demonstrating the significance of the gene in the grain carotenoid biosynthetic pathway.

The *Psy1* was mapped on chromosome 6 both in maize (Palaisa et al. 2003) and rice (Gallagher et al. 2004). Both chromosomes share extensive synteny with wheat group 7 chromosomes (Gale and Devos 1998). Considering the QTL mapping results mentioned above, the genes for kernel YP content residing on wheat homoeologous group 7 are most likely *Psy1*. More recently, Pozniak et al. (2007) mapped the *Psy1* on linkage group 7 chromosomes in durum wheat, providing further evidence for the hypothesis. To date, however, no full-length DNA sequence of *Psy1* has been cloned in either common or durum wheat.

Development of functional markers is very important for accurate discrimination of contrasting alleles in marker-assisted selection (Bagge et al. 2007; Tommasini et al. 2006; Yan et al. 2004). Previously, several molecular markers were identified on chromosome 7A closely linked to the QTLs for kernel YP content (Kuchel et al. 2006; Mares and Campbell 2001; Parker et al. 1998; Zhang et al. 2006). However, all were SSR, AFLP or SCAR markers located away from the genes, and no more useful gene-specific functional markers co-segregating with the genes of interest were developed. The objectives of this study were to characterize allelic variations of *Psy1* on chromosome 7A, evaluate its association with YP content in Chinese wheat cultivars, and develop a functional STS marker for the locus.

Materials and methods

Plant materials

Two hundred and seventeen Chinese wheat cultivars and advanced lines were used for validation of a newly developed STS marker. A recombinant inbred line (RIL) population with 240 lines derived from the cross PH82-2/Neixiang 188 was used to map both the *Psy* gene and the STS marker. PH82-2 has a low YP content, whereas Neixiang 188 has high YP content. The RIL population was previously genotyped with 188 SSR and four protein markers (Wu et al. 2007). A set of Chinese Spring nullisomic–tetrasomic lines (except nullisomic 2A and nullisomic 4B) and ditelosomic line 7AS, kindly provided by Prof. R. A. McIntosh, University of Sydney, were employed to verify the chromosomal location of the STS marker. Nullisomic 2A–tetrasomic 2B plants were selected from the self-pollinated progeny of a monosomic 2A–tetrasomic 2B (M2A–T2B) individual, kindly provided by Prof. Peidu Chen, Nanjing Agricultural University.

Field trials and grain yellow pigment assay

During the 2001–2002 and 2002–2003 cropping seasons, 217 Chinese wheat cultivars and advanced lines were sown in a randomized complete block with three replicates in the Anyang experimental station of the Chinese Academy of Agricultural Sciences, located in Henan Province. Each plot consisted of two 2-m rows spaced 25 cm apart, with 100 plants in each row. A total of 240 $F_{5:6}$ lines from the cross PH82-2/Neixiang 188 were planted in a latinized alpha-lattice design in Jiaozuo and Anyang (Henan) and Taian (Shandong) in the 2004–2005 cropping season (Wu et al. 2007). Test plots were managed according to local practices. All field trials were kept free of weeds and diseases, with two applications of broad-range herbicides and fungicides, respectively. The procedure for estimating YP content in wheat grains followed the recommended AACC method (AACC 1995).

Strategies for cloning and sequence analysis of the *Psy* gene

The corresponding cDNA sequence of maize *Psy1* gene (GenBank accession U32636) was used for a BLAST search against the wheat EST database in GenBank. All wheat ESTs sharing high similarity with the reference gene were subjected to contig assembly (E value < e−50, score > 300 bp and identity > 70%, http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/glossary2.html). The constructed contigs were then aligned with the genomic DNA sequence of U32636 to determine the putative insertion sites of introns. Primers were designed based on the constructed contigs using the software Premier Primer 5 (http://www.premierbiosoft.com) and synthesized by Beijing Augct Biological Technology Co., Ltd (http://www.augct.com). Each primer combination was tested with the Chinese Spring nullisomic–tetrasomic lines N7A–T7B, N7B–T7D and N7D–T7A to confirm a chromosome 7A-specific location.

Genomic DNA was extracted from single seeds using a method modified from Lagudah et al. (1991). PCR reactions were performed in an MJ Research PTC-200 thermal cycler in a total volume of 20 µl, including 20 mM of Tris–HCl (pH 8.4), 20 mM of KCl, 100 µM of each of dNTP, 1.5 mM of $MgCl_2$, 5 pmol of each primer, 1 unit of *Taq* DNA polymerase (TIANGEN Biotech Co., Ltd., Beijing, http://www.tiangen.com) and 50 ng of template DNA. Reaction conditions were 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 62°C for 30 s, and 72°C for 1.5 min, with a final extension of 72°C for 5 min. The PCR products were separated by electrophoresis either on a 2% agarose gel stained with ethidium bromide and visualized using UV light, or on a 6% polyacrylamide gel and resolved by silver staining (Bassam et al. 1991). Targeted fragments with expected sizes based on the primer-binding sites in the DNA sequence of U32636 were recovered and cloned into the pMD18-T vector, and sequenced by Shanghai Sangon Biological Engineering Technology & Service Co. Ltd. (http://www.sangon.com).

Six wheat cultivars with a diverse range of YP content were used for cloning the full-length sequence of *Psy1* on chromosome 7A (designated *Psy-A1*) and for identifying allelic variations at the locus. The PCR reaction and DNA sequencing were repeated 2–4 times for each primer set to ensure the accuracy of nucleotide sequences. The genomic DNA sequence and deduced amino acids for different allelic variants of *Psy-A1* gene were aligned using the software DNAMAN (http://www.lynnon.com). Intron positions were determined by alignment of the amplified genomic DNA sequence and constructed contig sequences using DNAMAN as well.

Development of a functional marker for *Psy-A1*

Seven wheat cultivars with high and seven with low YP content were used for screening the newly developed STS markers. Based on the sequence divergence of the two haplotypes at the *Psy-A1* locus characterized in this study, two primer sets were designed to identify allelic variants of *Psy-A1* gene in the 14 cultivars. The divergence of phenotypic values was assumed to be associated with *Psy-A1*, which was then validated using 217 Chinese wheat cultivars and advanced lines, and confirmed with the RIL population. PCR reactions were 95°C for 5 min, followed by 35 cycles of 95°C for 30 s, 65°C for 30 s, and 72°C for 30 s, with a final extension of 72°C for 5 min..

Statistical analyses

For the 217 cultivars and advanced lines, the YP content of each genotype was measured in each of two cropping seasons and averaged to verify the association between YP content and the newly developed functional marker. Analysis of variance was conducted by PROC MIXED in the Statistical Analysis System (SAS Institute 2000) with genotype classes as a categorical variable to derive the mean YP content for each class and to test the significant level for the two classes. The genotype class, indicated by two types of fragments (194 and 231 bp) amplified with the STS marker *YP7A*, was treated as fixed effects, while genotypes nested in class as random. Linkage analysis was performed with the software Map Manager QTX (Manly et al. 2001) and the Kosambi function was used for calculating genetic distance (Kosambi 1944). QTL analysis was conducted using the software QTL Cartographer V2.5 (Wang et al. 2005), and a LOD score of 3.0 was

set as the threshold for declaring significant linkage and QTL detection.

Results

Cloning and characterization of *Psy-A1*

Fourteen wheat ESTs were detected through a BLAST search using the cDNA sequence of U32636 as a probe. These were assembled into five contigs based on their sequence similarities (data not shown). Of them, contigs 1 and 2 were mapped on chromosome 7A and contig 3 on chromosome 7B, by means of chromosome-specific primer localization, using Chinese Spring nullisomic–tetrasomic lines N7A-T7B, N7B-T7D and N7D-T7A as template DNA (Fig. 1). Contigs 4 and 5, however, cannot be localized using this method. Of the contigs mapped on chromosome 7A, contig 1 comprised ESTs AL822090 and BM137136, and contig 2 included CJ685595, CJ579028, CJ617730 and CJ721718. These two contigs covered the full coding sequence (CDS) of U32636 (Fig. 2).

Four primer sets were generated to amplify the full-length genomic DNA sequence of *Psy-A1* based on the sequences of the two contigs (Table 1). Each primer set was tested with Chinese Spring nullisomic–tetrasomic lines N7A–T7B, N7B–T7D and N7D–T7A, and three of the four primer sets used for cloning *Psy-A1*, namely, *Y2*, *Y3* and *Y4* (Table 1), were mapped to chromosome 7A using this method (Fig. 1). The primer set *Y1*, however, amplified an approximate 1,000-bp fragment in each of the three nullisomic–tetrasomic lines, showing its cross-amplifying nature. All three bands amplified from N7A–T7B, N7B–T7D and N7D–T7A were recovered and cloned, and eight clones from each band were sequenced. Two divergent sequences were obtained in the clones from N7B–T7D, and one of them was the same as that in the clones from N7A–T7B and the other was the same as that from N7D–T7A, implying that the primer set *Y1* co-amplified *Psy1* loci from chromosomes 7A and 7D. Thus the sequence obtained in the clones from N7D–T7A was the upstream sequence of the *Psy1* gene on chromosome 7A, and it was then assembled with sequences from other three primer sets to construct the full-length DNA sequence of *Psy-A1*.

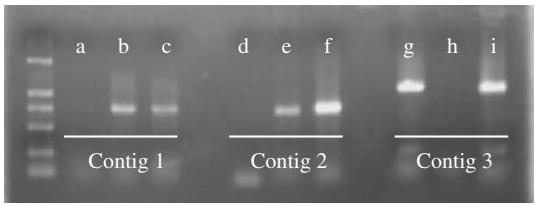

Fig. 1 Chromosomal localization of assembled contigs by amplifying nullisomic–tetrasomic lines of Chinese Spring N7A-T7B (lanes *a*, *d*, *g*), N7B-T7D (lanes *b*, *e*, *h*) and N7D-T7A (lanes *c*, *f*, *i*) with primer sets *Y5*, *Y6* and *Y7*, respectively

The complete genomic DNA sequence of *Psy-A1* comprised 4,175 bp, with six exons and five introns, as well as the 5′ and 3′ flanking sequences (Fig. 3 and Supplementary Fig. 1). The exon–intron structure of *Psy-A1* was very similar to that of maize *Psy1* (U32636), indicating a close phylogenetic relationship. The exons of the maize and wheat *Psy1* genes showed high sequence identities, ranging from 72.2 to 86.3%, whereas the introns exhibited low sequence similarities of 14.3–39.3%. Moreover, the sizes of exons were more conserved between the two species than those of introns (Fig. 3).

The cDNA sequence of the cloned *Psy-A1* was 1,811 bp, containing an open reading frame (ORF) of 1,284 bp, a 221-bp 5′ untranslated region (UTR) and a 306-bp 3′ UTR (Sup-Fig. 1). The deduced pre-protein sequence of *Psy-A1* consisted of 428 amino acid residues, with a predicted molecular mass of ~47.7 kDa. Based on the alignment of the deduced amino acids of *Psy-A1* with those from other plants, the putative cleavage site was located in the 70th amino acid. Cleavage at this site would generate a mature protein of 358 amino acid residues with a molecular mass of ~40.6 kDa and a transit peptide of 70 amino acids with a molecular mass of ~7.1 kDa. Using the deduced pre-protein sequence of *Psy-A1* for a BLASTP search, a putative conserved domain for phytoene synthase was found in the region 144th to 418th amino acid, with an *E* value of

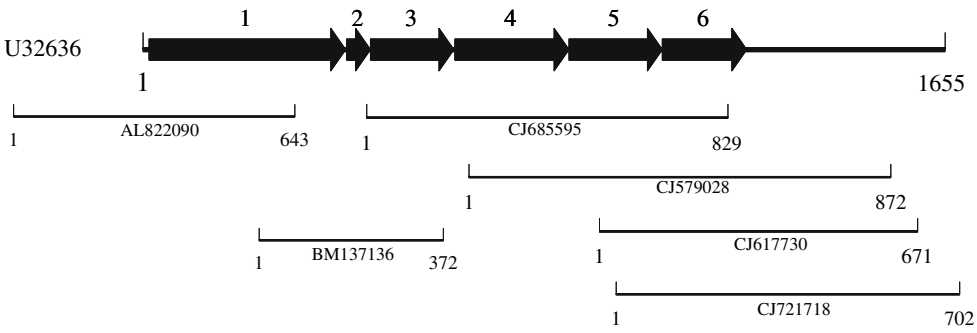

Fig. 2 The six ESTs involved in two contigs localized to chromosome 7A and their relative positions to the cDNA sequence of maize *Psy1* (GenBank accession number U32636). The *numbered solid arrows* on U32636 denote the exons; the *lines* under U32636 represent the aligned wheat ESTs with their sizes and accession numbers indicated

Table 1 Four primer sets used for cloning *Psy-A1*, three for mapping contigs 1, 2 and 3 (primer sets *Y5*, *Y6* and *Y7*, respectively) and the functional STS marker *YP7A* developed for the gene

Primer set or marker	Primer sequence (5′–3′)	Amplified region[a]	Size of PCR fragments (bp)
Y1	Forward: GGCAGGCTAGTGGTCGGTA	1–993	993
	Reverse: TGACGGTCTGAAGTGAGAATGA		
Y2	Forward: GGACCTTGCTGATGACCGAG	800–1,602	803
	Reverse: GGGGAACTTGGTGATGGTGTC		
Y3	Forward: TATGGTGCAGGAGGACAGAC	1442–2,679	1,238
	Reverse: CAAGATGGTGGATTCAGGCTC		
Y4	Forward: AGCTGAGAGCGTCTATGGC	2,422–4,175	1,754
	Reverse: CGGGACCAACAACGAGTATA		
Y5	Forward: CGCCGTTATATGTCACTCAC	−65–822	887
	Reverse: CTCCTCGGTCATCAGCAAG		
Y6	Forward: GGACCTTGCTGATGACCGAG	800–1,602	803
	Reverse: GGGGAACTTGGTGATGGTGTC		
Y7	Forward: GGACCTCAAGAAGGCAAGAT	80–1,282[b]	1,203
	Reverse: CGGGACCGACAACGAGTATA		
YP7A	Forward: GGACCTTGCTGATGACCGAG	800–993 or 800–1,030	194 or 231
	Reverse: TGACGGTCTGAAGTGAGAATGA		

[a] Amplified region was counted from the genomic DNA sequence of *Psy-A1* cloned in this study

[b] The amplified region indicated here are counted from the first nucleotide of contig 3 that mapped on chromosome 7B, including ESTs CA598664, CD862515, CD862556, CD875290, CD898848

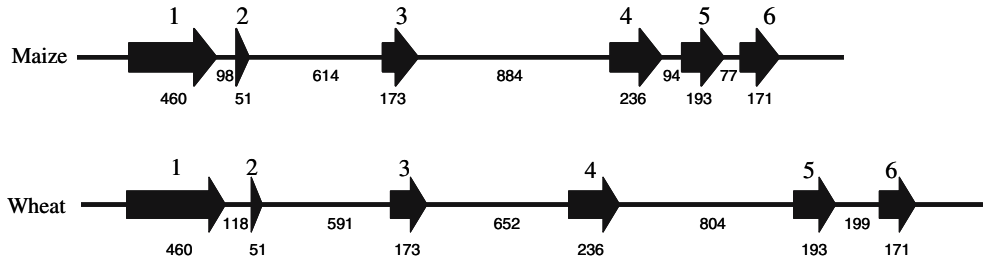

Fig. 3 The exon–intron structure of *Psy1* from maize (U32636) and wheat (EF600063). The *numbered solid arrows* denote exons, and the *lines* between exons represent introns. The *numbers* under exons and introns indicate their size (bp)

5e−84. The putative mature protein of wheat *Psy1* also shared higher sequence identities of 74–85% with those of other plants than the presumed transit peptides with 9–46% of sequence identities.

Identification and validation of the functional marker for *Psy-A1*

The amplified genomic DNA sequences of *Psy-A1* from wheat cultivars with diverse grain YP contents revealed the presence of two haplotypes, *Psy-A1a* and *Psy-A1b* (Sup-Fig. 1, GenBank accessions EF600063 and EF600064, respectively). An insertion of a 37-bp fragment was observed in the 5′-end of the second intron of *Psy-A1b* (Fig. 4), and two SNPs were detected with one in the first exon and the other in the fourth intron (Sup-Fig. 1). The SNP in exon was a synonymous mutation.

One of the two primer sets designed from the insertion of the 37-bp polymorphic sequence, designated *YP7A* (Table 1), demonstrated polymorphic band profile between the accessions with high and low YP content (Fig. 5), yielding a 194-bp fragment in cultivars with allele *Psy-A1a* and a 231-bp fragment in those with allele *Psy-A1b*. In the test of 217 Chinese wheat cultivars and advanced lines using *YP7A*, the 194-bp fragment was detected in 135 genotypes, and the 231-bp fragment was detected in 82 genotypes. The group with the 194-bp fragment had significantly higher YP content than that with 231-bp fragment (Table 2), confirming the association of PCR band profile with YP phenotype. Thirty-four of the 36 accessions with YP contents higher than 2.2 mg kg^{-1} associated with the 194-bp fragment, whereas 20 of the 29 accessions with YP contents lower than 1.0 mg kg^{-1} possessed the 231-bp fragment (data not shown, 1.0 and 2.2

Fig. 4 Alignment of partial DNA sequences of *Psy-A1a* (EF600063) and *Psy-A1b* (EF600064). The first and second introns are *underlined*; the sequences of wheat EST BM137136 are *shadowed*; the sequence of the forward primer and complementary sequence of the reverse primer for *YP7A* are *boxed*

```
                                                                    YP7A Forward primer
Psy-A1a  GGAGCAGAAGGTGTACGACGTGGTGGTGAAGCAGGCGGCATTGCTCAAGCGCCAGCTGCGCCCCTCGCAG  560
Psy-A1b  GGAGCAGAAGGTGTACGACGTGGTGGTGAAGCAGGCGGCATTGCTCAAGCGCCAGCTGCGCCCCTCGCAG  560

Psy-A1a  CAACAGCAGCAGGCGCCGCCAGCCGTCGCCAGGGAGCTGGACGCGCCGCGCGGCGGGCTCGGGGAGGCCT  630
Psy-A1b  CAACAGCAGCAGGCGCCGCCAGCCGTCGCCAGGGAGCTGGACGCGCCGCGCGGCGGGCTCGGGGAGGCCT  630

Psy-A1a  ACGCCCGCTGCGGCGAGATCTGCGAGGAGTACGCCAAGACCTTCTACCTCGGTACGCCACTCCTTCGTGG  700
Psy-A1b  ACGCCCGCTGCGGCGAGATCTGCGAGGAGTACGCCAAGACCTTCTACCTCGGTACGCCACTCCTTCGTGG  700

Psy-A1a  ATACTCTGTTTTTCTTGAGCCATGGTGGCAGGCTGCGTGCCAAGCCGGTGTTCCGGTGATCATGGAGCTC  770
Psy-A1b  ATACTCTGTTTTTCTTGAGCCATGGTGGCAGGCTGCGTGCCAAGCCGGTGTTCCGGTGATCATGGAGCTC  770

                                                                    YP7A Forward primer
Psy-A1a  ACTCGTTCATGTCTGGTCGTGCATGGCAGGGACCTTGCTGATGACCGAGGAGCGGCGGCGCGCCATATGG  840
Psy-A1b  ACTCGTTCATGTCTGGTCGTGCATGGCAGGGACCTTGCTGATGACCGAGGAGCGGCGGCGCGCCATATGG  840

Psy-A1a  GCCATCTACG.....................................GTAATCTGAAAATTCACCATGCC  873
Psy-A1b  GCCATCTACGGTAATCTGAAAATTCGCCGCCATATGGGCCATCTACGGTAATCTGAAAATTCACCATGCC  910

Psy-A1a  TGGTTTGGACCCTCCATTGTTGCTCCCCTGTTGTGGTATCAGTATGTGTCACACAGTGTTAGTTAGTGTC  943
Psy-A1b  TGGTTTGGACCCTCCATTGTTGCTCCCCTGTTGTGGTATCAGTATGTGTCACACAGTGTTAGTTAGTGTC  980

                                                                    YP7A Reverse primer
Psy-A1a  AGTAATGTGACTGAAAATTCAGCTAGTTTCATTCTCACTTCAGACCGTCAGAAAGGGCATGCCCACATTT 1013
Psy-A1b  AGTAATGTGACTGAAAATTCAGCTAGTTTCATTCTCACTTCAGACCGTCAGAAAGGGCATGCCCACATTT 1050

Psy-A1a  TGCATCAGTTAAATTGCTACATATTGTATTTAACAGCAACTTGCAAGAATCTTCAACACTCCCCAAGAAA 1083
Psy-A1b  TGCATCAGTTAAATTGCTACATATTGTATTTAACAGCAACTTGCAAGAATCTTCAACACTCCCCAAGAAA 1120

Psy-A1a  ATTGGCCACTTTAAAGTTAATGGTGTGAACTAGTTCTGGATGCGAATAATGGCAAATAGAAACATTGCTG 1153
Psy-A1b  ATTGGCCACTTTAAAGTTAATGGTGTGAACTAGTTCTGGATGCGAATAATGGCAAATAGAAACATTGCTG 1190

Psy-A1a  AACTTGCATGCTATGTGTTTACAGATACTCCTATATACGTAGTATAGTCAGTGAAGAATAAAGGGTTCGT 1223
Psy-A1b  AACTTGCATGCTATGTGTTTACAGATACTCCTATATACGTAGTATAGTCAGTGAAGAATAAAGGGTTCGT 1260

Psy-A1a  ATAACACTTTTTTATATGCCATTATGTGTGGAAGCATCAAATTAGGCTTTTTGTTGGCTAAATGGCTTCA 1293
Psy-A1b  ATAACACTTTTTTATATGCCATTATGTGTGGAAGCATCAAATTAGGCTTTTTGTTGGCTAAATGGCTTCA 1330

Psy-A1a  ATAGGATCAAAGTACACGAGAAAAGGTTGCAAGAACATATTCCTCAAATTGCCTGGGGACATGAATCTGA 1363
Psy-A1b  ATAGGATCAAAGTACACGAGAAAAGGTTGCAAGAACATATTCCTCAAATTGCCTGGGGACATGAATCTGA 1400
```

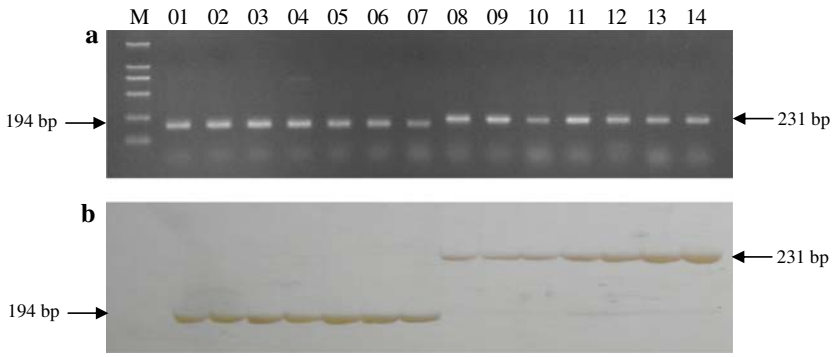

Fig. 5 Test of polymorphism for PCR fragments amplified with *YP7A* in seven cultivars with high and seven with low YP content on a 2% agarose gel (**a**) and a 6% denaturing polyacrylamide gel (**b**). M DNA ladder DL2000; 01 CA9648 (3.100 mg kg^{-1}); 02 Nongda 3291 (3.176); 03 99G66 (2.905); 04 99G80 (3.100); 05 Shannong 1355 (3.356); 06 Yun 97169 (3.416); 07 Xuzhou 826 (2.905); 08 Xinong 336 (0.482); 09 Xinong 6426 (0.753); 10 Ning 97–18 (0.346); 11 Huaimai 16 (0.572); 12 Wanmai 33 (0.662); 13 Mianyang 940112 (0.677); 14 Mianyang 960107 (0.692). The *number in parentheses* indicates the yellow pigment content in wheat grain (mg kg^{-1})

Table 2 Statistical analysis of the association between PCR band profile and YP content in 217 wheat accessions, including 104 with and 113 without the 1B•1R translocation, respectively, tested with *YP7A*

Line	Allele[a]	Accession number tested	Mean YP content[b]	Range
All lines	*Psy-A1a*	135	1.82a	0.62–3.42
	Psy-A1b	82	1.30b	0.35–2.88
Lines with 1B•1R translocation	*Psy-A1a*	79	2.07a	0.93–3.42
	Psy-A1b	25	1.58b	0.86–2.88
Lines with Non- 1B•1R translocation	*Psy-A1a*	56	1.48a	0.62–2.91
	Psy-A1b	57	1.18b	0.35–2.15

[a] *Psy-A1a* and *Psy-A1b* amplified by the 194 and 231-bp fragments, respectively, using *YP7A*

[b] Different letters following the mean YP content indicate highly significant differences between the two groups ($P < 0.01$)

are calculated from one standard deviation from the mean value of YP content).

Chromosomal location and QTL analysis of *YP7A*

Linkage analysis showed that *YP7A* was linked to *Xwmc809*, an SSR marker located on chromosome 7AL, with a genetic distance of 5.8 cM (Fig. 6). A major QTL for YP content on chromosome 7A co-segregated with *YP7A* and explained from 20 to 28% of the phenotypic variance in the RIL population across three environments. The location of *YP7A* on chromosome 7AL was confirmed by amplifying genomic DNA from a set of Chinese Spring nullisomic–tetrasomic lines and ditelosomic line 7AS (Fig. 7). No PCR product was detected in N7A–T7B (lane 20) and DT7AS (lane 23).

Discussion

Cloning and characterization of *Psy-A1*

Cloning genes from common wheat is complicated due to its enormous genome size and allohexaploid nature. It is very necessary to make sure that each primer set used for cloning genes is chromosome-specific. In this study, three primer sets *Y2*, *Y3* and *Y4* were confirmed to be chromosome 7A specific by amplifying nullisomic–tetrasomic lines N7A–T7B, N7B–T7D and N7D–T7A. However, the primer set *Y1* did not show any genome specific in the PCR amplification, and thus the PCR products from the nullisomic–tetrasomic lines for *Y1* were sequenced, and the sequence from N7D–T7A was determined to be amplified from chromosome 7A. Through this strategy, the complete genomic DNA sequence of Psy gene on chromosome 7A was constructed.

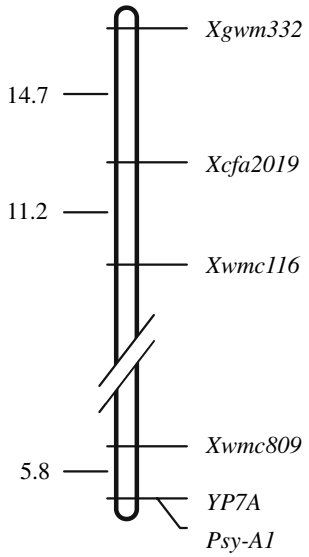

Fig. 6 Linkage map constructed with *YP7A* and four SSR markers on wheat chromosome 7AL in the RIL population derived from PH82-2/Neixiang 188

The genomic DNA sequences of the *Psy1* gene from wheat, tomato (Ray et al. 1992), maize (Buckner et al. 1996), and rice (Gallagher et al. 2004) showed similar exon–intron characteristics, i.e. six exons interspaced by five introns. Except the first exon, the second to fifth exons of *Psy 1* exhibited the same size in wheat, maize, and potato, as well as rice with the same sizes for the second to the fourth exons. The four species reflected diverse sequences of the *Psy1* transit peptides, which direct the pre-protein to the target plastids, chloroplasts or chromoplasts (Karvouni et al. 1995).

The amino acid sequence of wheat *Psy-A1* showed a high degree of identity and similar molecular weight with those from other plants, such as rice (Gallagher et al. 2004), maize (Buckner et al. 1996), melon (Karvouni et al. 1995), and sunflower (Fambrini et al. 1993). The pre-protein of wheat *Psy1* gene shared similarities with those from tomato (Bartley et al. 1992), daffodil (Schledz et al. 1996) and sunflower (Salvini et al. 2005), i.e. a transit peptide of 50–70 amino acids and a mature protein of ~350 residues.

A molecular mechanism for the association between allelic variation of *Psy-A1* and grain YP content

Introns can influence gene transcription through alternative splicing (Fedorova and Fedorov 2003). In previous studies on wheat grain polyphenol oxidase (PPO) genes, divergence in the introns of PPO genes showed high associations with grain PPO activity (Sun et al. 2005; He et al. 2007). Likewise, a 37-bp insertion was found in the 5′-end of the second intron of *Psy-A1b*, followed by a GT dinucleotide, suggesting a similar molecular mechanism for intron splicing.

In contig assembly, the wheat EST BM137136 aligned well with the first and second exons, but not the third exon, of the maize *Psy1* gene (U32636). Alignment of the EST with *Psy-A1b* showed that its downstream sequence comprised two segments of the second intron of *Psy-A1b* (Fig. 4). Interestingly, in *Psy-A1b*, the intron fragment between the two segments present in BM137136 was bordered with GT-AG, a typical structure for intron splicing. This strongly implied that alternative splicing occurred in post-transcriptional modification of *Psy-A1b*, resulting in a nonsense mutation generating a premature translation termination before the third exon; as a consequence this influences the expression of *Psy-A1b* leading to reduced grain YP content.

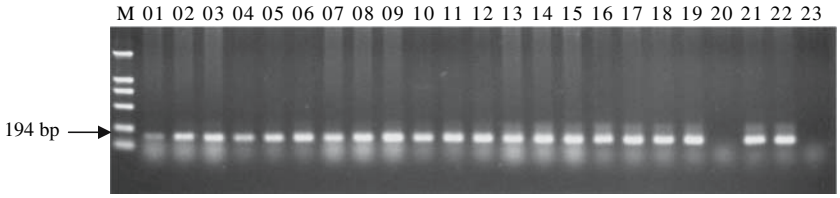

Fig. 7 PCR amplification of Chinese Spring, its nullisomic–tetrasomic lines, and ditelosomic line 7AS with STS marker *YP7A*. M DNA ladder DL2000; 01 Chinese Spring; 02 Chinese Spring nullisomic 1A–tetrasomic 1D (N1A-T1D); 03 N1B-T1D; 04 N1D-T1B; 05 N2A-T2B; 06 N2B-T2A; 07 N2D-T2B; 08 N3A-T3D; 09 N3B-T3D; 10 N3D-T3A; 11 N4A-T4B; 12 DT4BS; 13 N4D-T4B; 14 N5A-T5B; 15 N5B-T5A; 16 N5D-T5B; 17 N6A-T6B; 18 N6B-T6D; 19 N6D-T6B; 20 N7A-T7B; 21 N7B-T7D; 22 N7D-T7A; 23 DT7AS

Evaluation of the functional marker *YP7A*

Functional markers, developed from polymorphic sites within genes that causally affect phenotypic variation, are ideal tools for marker-assisted selection in wheat breeding (Bagge et al. 2007; Tommasini et al. 2006; Yan et al. 2004). In the present study, a co-dominant functional marker *YP7A* for wheat grain YP content was developed and validated. The marker was shown to be highly relevant to YP content and thus can be used in wheat breeding programs aimed at improving the colour of flour for various wheat-based end products.

Nevertheless, nine cultivars with the 194-bp fragment still showed a YP content of less than 1.0 mg kg^{-1}, and two cultivar with the 231-bp fragment possessed a YP content higher than 2.2 mg kg^{-1}. This may be attributed in part to other genes for grain YP content, especially the one on chromosome 7B (Elouafi et al. 2001; Kuchel et al. 2006; Pozniak et al. 2007). Developing functional markers for the Psy gene on chromosome 7B, and implementing them for wheat YP evaluation in combination with *YP7A*, might be necessary for breeding wheat cultivars with low or high YP content. As for the Psy gene on chromosome 7D, it may be silenced or hardly expressed since no QTLs associated with grain YP content have been reported at this locus. The results of BLAST search performed in this study are also in favor of this hypothesis, in which 12 out of 14 ESTs were mapped to either chromosomes 7A or 7B, and two unmapped, strongly suggesting that the Psy genes on these two chromosomes are actively expressed, whereas the one on chromosome 7D is not. The phenomenon of different expression of homoeologous genes (homoeoalleles) has been extensively studied in common wheat (Bottley et al. 2006; Nomura et al. 2005).

Furthermore, the 1B•1R translocation also exerts a significant influence on grain YP. In the RIL population derived from the cross of PH82-2 and Neixiang 188, apart from *YP7A*, we also detected a major QTL on chromosome 1B, explaining 30.6–54.0% of phenotypic variance (data not shown). Since 42.6% of the cultivars in the Northern China Winter Wheat Region, the most important wheat growing area in China, have the 1B•1R translocation (Zhou et al. 2007), its influence on YP content cannot be neglected. Nevertheless, *YP7A* still showed a significant effect upon grain YP content either in the lines with or without 1B•1R translocation (Table 2), exhibiting its nature as a reliable and competent marker for molecular-assisted selection of grain YP content in wheat breeding programs.

Acknowledgments The authors are very grateful to Prof. R. A. McIntosh, Plant Breeding Institute, University of Sydney for his critical review of this manuscript. This study was supported by the National Science Foundation of China (30771335), National Basic Research Program (2002CB11300), and National 863 Program (2006AA10Z1A7 and 2006AA100102).

References

AACC (1995) Approved methods of the American association of cereal chemists, 9th edn. St. Paul, MN, USA

Adom KK, Sorrells ME, Liu RH (2003) Phytochemical profiles and antioxidant activity of wheat varieties. J Agric Food Chem 51:7825–7834

Bagge M, Xia XC, Lübberstedt T (2007) Functional markers in wheat. Curr Opin Plant Biol 10:211–216

Bartley GE, Viitanen PV, Bacot KO, Scolnik PA (1992) A tomato gene expressed during fruit ripening encodes an enzyme of the carotenoid biosynthesis pathway. J Biol Chem 267:5036–5039

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal Biochem 196:80–83

Bottley A, Xia GM, Koebner RMD (2006) Homoeologous gene silencing in hexaploid wheat. Plant J 47:897–906

Buckner B, Miguel PS, Janick-Buckner D, Bennetzen JL (1996) The y1 gene of maize codes for phytoene synthase. Genetics 143:479–488

Clarke FR, Clarke JM, McCaig TN, Knox RE, Depauw RM (2006) Inheritance of yellow pigment concentration in seven durum wheat crosses. Can J Plant Sci 86:133–141

Elouafi I, Nachit MM, Martin LM (2001) Identification of a microsatellite on chromosome 7B showing a strong linkage with yellow pigment in durum wheat (*Triticum turgidum* L. var. *durum*). Hereditas 135:255–261

Fambrini M, Pugliesi C, Vernieri P, Giuliano G, Baroncelli S (1993) Characterization of a sunflower (*Helianthus annuus* L.) mutant, deficient in carotenoid synthesis and abscisic-acid content, induced by in-vitro tissue culture. Theor Appl Genet 87:65–69

Fedorova L, Fedorov A (2003) Introns in gene evolution. Genetica 118:123–131

Gale MD, Devos KM (1998) Comparative genetics in the grasses. Proc Natl Acad Sci USA 95:1971–1974

Gallagher CE, Matthews PD, Li F, Wurtzel ET (2004) Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. Plant Physiol 135:1776–1783

He XY, He ZH, Zhang LP, Sun DJ, Morris CF, Fuerst EP, Xia XC (2007) Allelic variation of polyphenol oxidase (PPO) genes located on chromosomes 2A and 2D and development of functional markers for the PPO genes in common wheat. Theor Appl Genet 115:47–58

He ZH, Yang J, Zhang Y, Quail K, Pena RJ (2004) Pan bread and dry white Chinese noodle quality in Chinese winter wheats. Euphytica 139:257–267

Hessler TG, Thomson MJ, Benscher D, Nachit MM, Sorrells ME (2002) Association of a lipoxygenase locus, *Lpx-B1*, with variation in lipoxygenase activity in durum wheat seeds. Crop Sci 42:1695–1700

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol 4:210–218

Karvouni Z, John I, Taylor JE, Watson CF, Turner AJ, Grierson D (1995) Isolation and characterisation of a melon cDNA clone encoding phytoene synthase. Plant Mol Biol 27:1153–1162

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

Kruger JE, Matsuo RR, Preston K (1992) A comparison of methods for the prediction of Cantonese noodle colour. Can J Plant Sci 72:1021–1029

Kuchel H, Langridge P, Mosionek L, Williams K, Jefferies SP (2006) The genetic control of milling yield, dough rheology and baking quality of wheat. Theor Appl Genet 112:1487–1495

Lagudah ES, Appels R, McNeil D (1991) The *Nor-D3* locus of *Triticum tauschii*: natural variation and genetic linkage to markers in chromosome 5. Genome 34:387–395

Lindgren LO, Stalberg KG, Hoglund AS (2003) Seed-specific overexpression of an endogenous *Arabidopsis* phytoene synthase gene results in delayed germination and increased levels of carotenoids, chlorophyll, and abscisic acid. Plant Physiol 132:779–785

Ma W, Daggard G, Sutherland M, Brennan P (1999) Molecular markers for quality attributes in wheat. In: Williamson P, Banks P, Haak I, Thompson J, Campbell A (eds) Proceedings of the ninth assembly of the Wheat Breeding Society of Australia, Toowoomba, vol 1, pp 115–117

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

Mares DJ, Campbell AW (2001) Mapping components of flour and noodle colour in Australian wheat. Aust J Agric Res 52:1297–1309

Miskelly DM (1984) Flour components affecting paste and noodle colour. J Sci Food Agric 35:463–471

Nomura T, Ishihara A, Yanagita RC, Endo TR, Iwamura H (2005) Three genomes differentially contribute to the biosynthesis of benzoxazinones in hexaploid wheat. Proc Natl Acad Sci USA 102:16490–16495

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15:1795–1806

Parker GD, Chalmers KJ, Rathjen AJ, Langridge P (1998) Mapping loci associated with flour color in wheat. Theor Appl Genet 97:238–245

Parker GD, Langridge P (2000) Development of a STS marker linked to a major locus controlling flour colour in wheat (*Triticum aestivum* L.). Mol Breed 6:169–174

Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. Theor Appl Genet 114:525–537

Ray J, Moureau P, Bird C, Bird A, Grierson D, Maunders M, Truesdale M, Bramley P, Schuch W (1992) Cloning and characterization of a gene involved in phytoene synthesis from tomato. Plant Mol Biol 19:401–404

Salvini M, Bernini A, Fambrini M, Pugliesi C (2005) cDNA cloning and expression of the phytoene synthase gene in sunflower. J Plant Physiol 162:479–484

Schledz M, Al-Babili S, von Lintig J, Haubruck H, Rabbani S, Kleinig H, Beyer P (1996) Phytoene synthase from *Narcissus pseudonarcissus*: functional expression, galactolipid requirement, topological distribution in chromoplasts and induction during flowering. Plant J 10:781–792

Sun DJ, He ZH, Xia XC, Zhang LP, Morris CF, Appels R, Ma WJ, Wang H (2005) A novel STS marker for polyphenol oxidase activity in bread wheat. Mol Breed 16:209–218

Tommasini L, Yahiaoui N, Srichumpa P, Keller B (2006) Development of functional markers specific for seven *Pm3* resistance alleles and their validation in the bread wheat gene pool. Theor Appl Genet 114:165–175

Wang S, Basten CJ, Zeng ZB (2005) Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

Wu YP, Zhang YL, Xiao YG, Yan J, Zhang Y, Zhang XK, Zhang LM, Xia XC, He ZH (2007) QTL mapping for important quality traits in common wheat. Sci Agric Sin (in press)

Yan L, Helguera M, Kato K, Fukuyama S, Sherman J, Dubcovsky J (2004) Allelic variation at the *VRN-1* promoter region in polyploid wheat. Theor Appl Genet 109:1677–1686

Zhang LP, Yan J, Xia XC, He ZH, Sutherland MW (2006) QTL mapping for kernel yellow pigment content in common wheat. Acta Agron Sin 32:41–45

Zhou Y, He ZH, Sui XX, Xia XC, Zhang XK, Zhang GS (2007) Genetic improvement of grain yield and associated traits in the Northern China winter wheat region from 1960 to 2000. Crop Sci 47:245–253