Structure and expression analysis of genes encoding ADP-glucose pyrophosphorylase large subunit in wheat and its relatives

Journal:	<i>Genome</i>
Manuscript ID	gen-2016-0007.R2
Manuscript Type:	Article
Date Submitted by the Author:	06-Apr-2016
Complete List of Authors:	<p>Zhang, Xiao-Wei; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Li, Si-Yu; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Zhang, Ling-Ling; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Yang, Qiang; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Jiang, Qian-Tao; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Ma, Jian; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Qi, Peng-Fei; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Li, Wei; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Chen, Guo-Yue; Triticeae Research Institute, Sichuan Agricultural University</p> <p>Lan, Xiu-Jin; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Deng, Mei; Sichuan Agricultural University, Triticeae Research Institute</p> <p>Lu, Zhen-Xiang; Agriculture and Agri-Food Canada Lethbridge Research Centre</p> <p>Liu, Chunji; CSIRO Plant Industry</p> <p>Wei, Yu-Ming; Triticeae Research Institute</p> <p>Zheng, You-Liang; Sichuan Agricultural University,</p>
Keyword:	Wheat, AGP-L, Gene structure, Starch accumulation, Expression

Structure and expression analysis of genes encoding ADP-glucose pyrophosphorylase large subunit in wheat and its relatives

Xiao-Wei Zhang¹⁺, Si-Yu Li¹⁺, Ling-Ling Zhang¹, Qiang Yang¹, Qian-Tao Jiang^{1*}, Jian Ma¹, Peng-Fei Qi¹, Wei Li¹, Guo-Yue Chen¹, Xiu-Jin Lan¹, Mei Deng¹, Zhen-Xiang Lu², Chunji Liu³, Yu-Ming Wei^{1*}, You-Liang Zheng⁴

¹*Triticeae Research Institute, Sichuan Agricultural University, Chengdu, Sichuan, 611130, China*

²*Lethbridge Research Centre, Agriculture and Agri-Food Canada, Lethbridge T1J 4B1, Canada*

³*CSIRO Agriculture Flagship, 306 Carmody Road, St Lucia, Brisbane, QLD 4067, Australia*

⁴*Key Laboratory of Southwestern Crop Germplasm Utilization, Ministry of Agriculture, Sichuan Agricultural University, Ya'an, Sichuan, 625014, China*

+The first and second authors contributed equally to this paper.

*Authors for correspondence:

Dr. Qian-Tao Jiang

E-mail: qiantaojiang@sicau.edu.cn

Tel: +86 28 86290958; Fax: +86 28 82650350

Dr. Yu-Ming Wei

E-mail: ymwei@sicau.edu.cn

Tel: +86-28-86290909; Fax: +86-28-82650350

Abstract

ADP-glucose pyrophosphorylase (AGP), which consists of two large subunits (AGP-L) and two small subunits (AGP-S), controls the rate-limiting step in the starch biosynthetic pathway. In this study, a full-length open reading frame (ORF) of AGP-L gene (named as *Agp2*) in wheat and a series of *Agp2* gene sequences in wheat relatives were isolated. The coding region of *Agp2* contained 15 exons and 14 introns including a full-length ORF of 1566 nucleotides, and the deduced protein contained 522 amino acids (57.8 kDa). Generally, the phylogenetic tree of *Agp2* indicated that sequences from A and D genome donor species were most similar to each other and sequences from B genome donor species contained more variation. Starch accumulation and *Agp2* expression in wheat grains reached their peak at 21 and 15 days post anthesis (DPA), respectively.

Keywords AGP-L · Expression · Gene structure · Starch accumulation · Wheat

Draft

Introduction

Common wheat (*Triticum aestivum* L.) is one of the most important cereal crops in the world. Starch is the most abundant storage reserve carbohydrate in wheat seed and comprises approximately 70% of wheat grain dry weight (Morell et al. 1995). The starch content influences wheat yield; starch also influences the processing quality of wheat flour (Blazek and Copeland 2008; Lee et al. 2001). Synthesis of adenosine diphosphate glucose (ADP-glucose), the substrate for starch synthesis, is catalyzed by ADP-glucose pyrophosphorylase (AGP), which controls the rate-limiting step in the starch biosynthetic pathway (Hannah 1997). In plants, AGP is an allosteric heterotetramer, consisting of two large subunits (AGP-L) and two small subunits (AGP-S). In most cases, the small subunit is catalytic, while the large subunit affects interactions between the small subunit and 3-phosphoglyceric acid (3-PGA) (Ballicora et al. 2004). Recently, catalytic AGP-Ls have been discovered (Ventriglia et al. 2008). 3-PGA and orthophosphate (Pi) are positive and negative allosteric effectors for AGP, respectively (Hannah 1997; Preiss 1997).

In *Arabidopsis thaliana* (L.) mutants, both AGP activity and starch synthesis decreased owing to the lack of one of the subunits (Li and Preiss 1992; Lin et al. 1988). Low levels of AGP activity were correlated with low levels of starch in maize mutants (Dickinson and Preiss 1969). Mutations created by site-specific mutagenesis in the maize AGP-L gene (*shrunk2*) increased maize seed weight by 18% without changing starch content (Giroux et al. 1996). An altered maize AGP-L gene (*Sh2r6hs*) that was transformed into and expressed in wheat increased total plant biomass, seed yield including seed number and AGP activity in developing seed endosperm (Meyer et al. 2004; Smidansky et al. 2007).

Subcellular localization has demonstrated that AGP may be plastidial or cytosolic in different species. In maize endosperm, AGP-Ls and AGP-Ss are encoded by two separate genes (Giroux and Hannah 1994). In barley, plastidial and cytosolic AGP-Ss are produced from a single gene (Thorbjørnsen et al. 1996). In wheat, genes encoding AGP-L were designated as *Agp2* (Ainsworth et al. 1995). Up to now, our knowledge on plastidial and cytosolic AGP-L is insufficient. In previous reports, only a cDNA sequence and a gene sequence of *Agp2* from the 1B chromosome of the Chinese Spring cultivar were identified (Ainsworth et al. 1995; Thorncroft et al. 2003). In addition, histochemical analysis of GUS expression initiated by the *Agp2B* promoter indicated that GUS expression was detected in the endosperm and aleurone but not in leaves (Thorncroft et al. 2003). Wheat relatives like *Triticum urartu* Thumanjan ex Gandilyan, *Aegilops speltoides* Tausch, and *A. tauschii* Coss. are

highly similar to wheat because they or their close relatives were the donors of the wheat A, B, and D genomes, so they could contribute to research on functional genes. To date, no *Agp2* sequences have been reported from relatives of common wheat.

Materials and methods

Plant material

Triticum aestivum cv. 'Bobwhite 7' (AABBDD, $2n = 42$) and 18 accessions of *Triticum urartu* (A^uA^u genome, $2n = 14$), *Triticum monococcum* L. (A^mA^m , $2n = 14$), *Aegilops speltoides* (SS, $2n = 14$), *Aegilops bicornis* (Forssk.) Jaub. & Spach (S^bS^b , $2n = 14$), *Aegilops longissima* Schweinf. & Muschl. (S^lS^l , $2n = 14$), *Aegilops sharonensis* Eig ($S^{sh}S^{sh}$, $2n = 14$), and *Aegilops tauschii* (DD, $2n = 14$) were kindly provided by USDA-ARS (<http://www.arsgrin.gov>) (Table 1). The accessions were grown in a glasshouse with a daytime temperature of 23°C and a nighttime temperature of 18°C, with a photoperiod of 16 h. Wheat seeds were collected at different developmental stages and stored at -80°C until DNA and RNA were extracted. Three replicates were used for all samples.

Isolation of the *Agp2* open reading frame (ORF) and gene sequences

Genomic DNA and total RNA were extracted from endosperm at 20 days post anthesis (DPA), using the modified SDS method described by Aljanabi and Martinez (1997) and Trizol (Tiangen, Beijing, China), respectively. cDNA was obtained from reverse transcription using a PrimeScript™ 1st Strand cDNA Synthesis Kit (Takara, Dalian, China). A set of primers (Table 2) was designed from the cDNA sequence of *Agp2* in wheat (Ainsworth et al. 1995) to isolate the ORF or gene sequences of *Agp2* from wheat and 18 diploid *Triticeae* accessions. PCRs were performed in 50-μL reactions containing 5 U PrimerStar HS DNA polymerase, 25 μL 2× PrimeSTAR GC Buffer (Mg^{2+} Plus), 200 μM of each dNTP, 200 nM of each primer, and 100 ng of DNA. The PCR program included an initial denaturation at 94°C for 5 min, followed by 35 cycles of 94°C for 1 min, 60°C for 1 min, 72°C for 2 min, and a final extension at 72°C for 10 min. Target bands were purified and cloned into the pEasy-T1 (Transgen, Beijing, China) vector and then transformed into *E. coli* (DH5α) (Transgen, Beijing, China). For each target band, three positive clones were selected and sequenced by Invitrogen (Shanghai, China).

Structure analysis of gene and protein

Intron-exon structures of *Agp2* were analyzed from the cDNA and gene sequences with CLC Genomics Workbench 8.0.2. Predictions of conserved domains in proteins were performed with Conserved Domain Database Tools (CDD) (<http://www.ncbi.nlm.nih.gov/cdd>). Gene and protein sequences were aligned with DNAMAN 7.0 (Lynnon Biosoft, San Ramon, CA, USA). AGP-L protein sequences from *A. thaliana* (NP_197423.1), *Hordeum vulgare* L. (CAX51355.1), *Oryza sativa* L. (ACJ71342.1), *Phaseolus vulgaris* L. (BAC66692.1), *Sorghum bicolor* (L.) Moench (XP_002456012.1), and *Zea mays* L. (AFP90368.1) were obtained from GenBank (<http://www.ncbi.nlm.nih.gov/protein/>). Amino acid secondary structures were predicted with PredictProtein (<http://www.predictprotein.org>) (Rost et al. 2004).

Phylogenetic analysis of *Agp2*

The phylogenetic tree was constructed from 18 accessions of *Triticum urartu*, *Triticum monococcum* L., *Aegilops speltoides*, *Aegilops bicornis*, *Aegilops longissima*, *Aegilops sharonensis* and *Aegilops tauschii* (Table 1). A dendrogram showing the evolutionary relationships of *Agp2* genes was generated by CLC Main Workbench 7.6.4 (CLC bio, Aarhus, Denmark) with the maximum likelihood method (Felsenstein 1981).

Starch accumulation pattern in wheat grains

The total starch content in grains of Bobwhite 7 was measured at 0, 3, 6, 9, 12, 15, 18, 21, 24, 27, and 30 DPA. The grains were ground in liquid nitrogen and assessed with the Total Starch Assay Kit (Megazyme, Bray, Wicklow, Ireland). Three replicates were used for all samples.

Expression analysis by real-time quantitative PCR (RT-qPCR)

The expression levels of *Agp2* in Bobwhite 7 grains were detected at 0, 3, 6, 9, 12, 15, 18, 21, 24, 27, and 30 DPA by RT-qPCR. The glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene was used as the internal reference. Primers used for RT-qPCR are listed in Table 1; the E values of primers for *Agp2* and *GAPDH* were 1.031 and 1.041, respectively. Total RNA extraction and cDNA synthesis were performed with the same methods outlined previously. The RT-qPCR program included an initial denaturation at 95°C for 2 min, followed by 40 cycles of 95°C for 5 s, 60°C for 15 s, 72°C for 15 s, and a melt curve stage. Relative expression levels of *Agp2* were calculated using the Pfaffl method (Pfaffl 2001).

Results

Sequence analysis of *Agp2* in wheat and related species

A full-length clone of the *Agp2* ORF from Bobwhite 7 and sequences of *Agp2* from its relatives were successfully isolated (Table 1). Gene structures and ORFs of *Agp2* were identified from these sequences with CLC Genomics Workbench 8.0.2. (CLC bio). The coding region of *Agp2* contained 15 exons and 14 introns (Fig. 1a). All species had an ORF of 1566 bp and the full gene ranged from 3264 to 3350 bp. The ORFs were 98.7% similar and the full gene was 93% similar across the species (Fig. S1). Most variation in the gene was located in intron 2 and exon 15. Intron 2 ranged from 425 to 508 bp.

Analysis of AGP-L protein structure

Agp2 from hexaploid wheat and its relatives encoded AGP-L consisting of 522 amino acid residues, including a putative transit peptide of 62 amino acids. The predicted secondary structure composition of AGP-L was 13.4% helix, 21.7% strand, and 65.0% loop (Fig. 1b). An alignment indicated that similarity of amino acid sequences between hexaploid wheat and its relatives was 99.4% and the similarity between *Triticeae* species and others (i.e., rice, maize, and *Arabidopsis*) was 63.5%. BlastP results against CDD databases indicated that conserved domains (i.e., oligomer interface, dimer interface, N-terminal domain interface, sulfate 1 binding site, ligand binding site) (Jin et al. 2005) were present in AGP-L from members of *Triticeae* and rice, maize, and *Arabidopsis* (Fig. 1b, Fig. 2).

Phylogenetic analysis of *Agp2*

The phylogenetic analysis of *Agp2* from 18 *Triticeae* accessions showed that the evolutionary tree was divergent and contained two clades: one was for all of A, B, D genome donor species and the other was only for B genome donor species (Fig. 3). In the upper clade, PI 428170 (*Triticum monococcum*) was categorized as an individual branch; A and D genome donor species (except PI 428170) were clustered in a branch and all D genome donor species composed a subgroup; B genome donor species (*A. bicornis*, *A. sharonensis* and *A. speltoidea*) composed the rest of branches. In the nether clade, 5 accessions of *A. longissima*, *A. sharonensis* and *A. speltoidea* were clustered in different subgroups.

Starch accumulation and *Agp2* expression analysis in wheat grain

Total starch content and total expression levels of *Agp2* were measured in wheat grains at different developmental stages to analyze the relationship between starch biosynthesis and *Agp2* expression. Starch content rose steadily after 6 DPA and reached a peak at 21 DPA before plateauing (Fig. 4a). Total expression levels of *Agp2* rose slowly from 3 DPA to 9 DPA before they decreased slightly. The

expression levels rose rapidly from 12 DPA and reached a peak at 15 DPA before slowly declining (Fig. 4a). The starch accumulation curve was similar to the expression curve (Fig. 4b), and there was a high correlation ($R^2=0.89$) between starch accumulation and *Agp2* expression.

Discussion

AGP controls the rate-limiting step in the starch biosynthetic pathway (Hannah 1997), and AGP-L increases the activity of AGP-S (Ballicora et al. 2004). In this study, a full-length ORF of *Agp2* encoding AGP-L from common wheat Bobwhite 7 and sequences of *Agp2* from its relatives were successfully isolated. As previously reported (Thorneycroft et al. 2003), the coding region of *Agp2* isolated from these species contained 15 exons and 14 introns (an additional intron has been reported from the 3' untranslated region), including a full-length ORF of 1566 nucleotides. The deduced protein includes 522 amino acids ($\approx$ 58 kDa) and includes a putative transit peptide of 62 amino acids (6.5 kDa) (Ainsworth et al. 1995).

The large subunit of AGP in higher plant is less conserved (50%–60% identity) than the small subunit (85%–95% identity) (Smith-White and Preiss 1992), which probably reflects differences in regulation of small subunit sensitivity to allosteric activation and inhibition in different tissues and species (Ballicora et al. 2004). In this study, there was a low similarity (63.5%) between AGP-L protein sequences from *Triticeae* and other species (rice, maize, and *Arabidopsis*), but a high similarity (99.4%) within members of *Triticeae*. The sequence conservation suggests that these *Triticeae* species use AGP-L in a similar manner. On the whole, the phylogenetic analysis of *Agp2* indicated that sequences from A and D genome donor species (except PI 428170) are most similar to each other and sequences from B genome donor species contain more variation. These sequences could be useful references for future research on AGP in wheat and its relatives.

Starch steadily accumulated from 6 DPA, but the peak starch content in Bobwhite 7 was recorded at 21 DPA (Fig. 4a), two weeks earlier than reported by Laudencia-Chingcuanco et al. (2007). Although the peak expression was observed at a similar time (15 DPA), the relative expression levels of *Agp2* in Bobwhite 7 changed 8-fold and just 2.6-fold in earlier report (Stamova et al. 2010). These differences might be due to differences between wheat varieties or growth conditions. In addition, both our study and earlier studies (Laudencia-Chingcuanco et al. 2007; Stamova et al. 2010) showed that *Agp2* expression levels reached their peak earlier than starch content peaked. In this study, *Agp2* expression reached its peak one week earlier than starch content peaked, indicating that the

accumulation of starch grains is a protracted and complicated process. Despite the time lag, there was a high correlation ($R^2=0.89$) between starch accumulation and *Agp2* expression.

Wheat relatives not only contribute to research on wheat functional genes, but also play important roles in wheat breeding. Sequences and expression of *Agp2* in *Triticeae* species were insufficiently known. Our study provides a better understanding of AGP-L, and should be a useful reference for further studies of AGP in wheat and its relatives.

Acknowledgments

This work was supported by the International Science & Technology Cooperation Program of China (No. 2015DFA30600), and the International Science & Technology Cooperation project of Sichuan province, China (2016HH0057).

References

- Ainsworth, C., Hosein, F., Tarvis, M., Weir, F., Burrell, M., Devos, K.M., Gale, M.D. 1995. Adenosine diphosphate glucose pyrophosphorylase genes in wheat: differential expression and gene mapping. *Planta*, **197**(1): 1-10. doi: 10.1007/BF00239933. PMID: 7580858.
- Aljanabi, S.M., Martinez, I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* **25**(22): 4692-4693. doi: 10.1093/nar/25.22.4692. PMID: 9358185.
- Ballicora, M.A., Iglesias, A.A., Preiss, J. 2004. ADP-glucose pyrophosphorylase: a regulatory enzyme for plant starch synthesis. *Photosynth. Res.* **79**(1): 1-24. PMID: 16228397.
- Blazek, J., Copeland, L. 2008. Pasting and swelling properties of wheat flour and starch in relation to amylose content. *Carbohydr. Polym.* **71**(3): 380-387. doi:10.1016/j.carbpol.2007.06.010.
- Dickinson, D.B., Preiss, J. 1969. Presence of ADP-glucose pyrophosphorylase in *shrunk-2* and *brittle-2* mutants of maize endosperm. *Plant Physiol.* **44**(7): 1058-1062. Available from <http://dx.doi.org/10.1104/pp.44.7.1058>. PMID: 16657157.
- Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**(6): 368-376. PMID: 7288891.
- Giroux, M., Hannah, L. 1994. ADP-glucose pyrophosphorylase in *shrunk-2* and *brittle-2* mutants of maize. *Mol. Gen. Genet.* **243**(4): 400-408. doi: 10.1007/BF00280470.

- Giroux, M.J., Shaw, J., Barry, G., Cobb, B.G., Greene, T., Okita, T., Hannah, L.C. 1996. A single mutation that increases maize seed weight. *Proc. Natl. Acad. Sci. U. S. A.* **93**: 5824-5829. PMID: 8650177.
- Hannah, L.C. 1997. Starch synthesis in the maize seed. *In* Cellular and Molecular Biology of Plant Seed Development, Volume 4. Edited by B.A. Larkins and I.K. Vasil. Springer Netherlands, pp. 375-405.
- Jin, X., Ballicora, M.A., Preiss, J., Geiger, J.H. 2005. Crystal structure of potato tuber ADP-glucose pyrophosphorylase. *EMBO J.* **24**(4): 694-704. doi: 10.1038/sj.emboj.7600551.
- Laudencia-Chinguanco, D.L., Stamova, B.S., You, F.M., Lazo, G.R., Beckles, D.M., Anderson, O.D. 2007. Transcriptional profiling of wheat caryopsis development using cDNA microarrays. *Plant Mol. Biol.* **63**(5): 651-668. doi: 10.1007/s11103-006-9114-y. PMID: 17211515.
- Lee, M-R., Swanson, B.G., Baik, B-K. 2001. Influence of amylose content on properties of wheat starch and breadmaking quality of starch and gluten blends. *Cereal Chem.* **78**(6): 701-706. Available from <http://dx.doi.org/10.1094/CCHEM.2001.78.6.701>.
- Li, L., Preiss, J. 1992. Characterization of ADPglucose pyrophosphorylase from a starch-deficient mutant of *Arabidopsis thaliana* (L.). *Carbohydr. Res.* **227**: 227-239. doi:10.1016/0008-6215(92)85074-A.
- Lin, T-P., Caspar, T., Somerville, C., Preiss, J. 1988. Isolation and characterization of a starchless mutant of *Arabidopsis thaliana* (L.) Heynh. lacking ADPglucose pyrophosphorylase activity. *Plant Physiol.* **86**(4): 1131-1135. doi: <http://dx.doi.org/10.1104/pp.86.4.1131>. PMID: 16666044.
- Meyer, F.D., Smidansky, E.D., Beecher, B., Greene, T.W., Giroux, M.J. 2004. The maize *Sh2r6hs* ADP-glucose pyrophosphorylase (AGP) large subunit confers enhanced AGP properties in transgenic wheat (*Triticum aestivum*). *Plant sci.* **167**(4): 899-911. doi:10.1016/j.plantsci.2004.05.031.
- Morell, M., Rahman, S., Abrahams, S., Appels, R. 1995. The biochemistry and molecular biology of starch synthesis in cereals. *Aust. J. Plant Physiol.* **22**(4): 647-660. doi:10.1071/PP9950647.
- Pfaffl, M.W. 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**(9): e45-e45. doi: 10.1093/nar/29.9.e45.
- Preiss, J. 1997. Modulation of starch synthesis. *In* A molecular approach to primary metabolism in higher plants. Edited by C.H. Foyer and W. Quick. Taylor and Francis, London, pp. 81-104.

- Rost, B., Yachdav, G., Liu, J. 2004. The predictprotein server. *Nucleic Acids Res.* **32**: W321-W326. doi: 10.1093/nar/gkh377. PMID: 15215403.
- Smidansky, E.D., Meyer, F.D., Blakeslee, B., Weglarz, T.E., Greene, T.W., Giroux, M.J. 2007. Expression of a modified ADP-glucose pyrophosphorylase large subunit in wheat seeds stimulates photosynthesis and carbon metabolism. *Planta*, **225**(4): 965-976. doi: 10.1007/s00425-006-0400-3. PMID: 17021802.
- Smith-White, B.J., Preiss, J. 1992. Comparison of proteins of ADP-glucose pyrophosphorylase from diverse sources. *J. Mol. Evol.* **34**(5): 449-464. doi: 10.1007/BF00162999. PMID: 1318389.
- Stamova, B.S., Laudencia-Chingcuanco, D., Beckles, D.M. 2010. Transcriptomic analysis of starch biosynthesis in the developing grain of hexaploid wheat. *International journal of plant genomics*, 2009. Available from <http://dx.doi.org/10.1155/2009/407426>.
- Thorbjørnsen, T., Volland, P., Kleczkowski, L.A., Olsen, O.A. 1996. A single gene encodes two different transcripts for the ADP-glucose pyrophosphorylase small subunit from barley (*Hordeum vulgare*). *Biochem. J.* **313**: 149-154. PMID: 8546676.
- Thorneycroft, D., Hosein, F., Thangavelu, M., Clark, J., Vizir, I., Burrell, M.M., Ainsworth, C. 2003. Characterization of a gene from chromosome 1B encoding the large subunit of ADPglucose pyrophosphorylase from wheat: evolutionary divergence and differential expression of *Agp2* genes between leaves and developing endosperm. *Plant Biotech. J.* **1**(4): 259-270. doi: 10.1046/j.1467-7652.2003.00025.x.
- Ventriglia, T., Kuhn, M.L., Ruiz, M.T., Ribeiro-Pedro, M., Valverde, F., Ballicora, M.A., Preiss, J., Romero, J.M. 2008. Two *Arabidopsis* ADP-glucose pyrophosphorylase large subunits (APL1 and APL2) are catalytic. *Plant Physiol.* **148**(1): 65-76. Available from <http://dx.doi.org/10.1104/pp.108.122846>. PMID: 18614708.

Table 1 Plant materials used in this study.

GRIN accession no.	Species	Genome	GenBank accession no.
PI 428320	<i>Triticum urartu</i>	A ^u	KU170096
PI 427328	<i>Triticum urartu</i>	A ^u	KU170097
PI 428170	<i>Triticum monococcum</i> subsp. <i>monococcum</i>	A ^m	KU170098
PI 428174	<i>Triticum monococcum</i> subsp. <i>monococcum</i>	A ^m	KU170099
PI 170204	<i>Aegilops speltoides</i>	S	KU170100
PI 487232	<i>Aegilops speltoides</i>	S	KU170101
PI 542249	<i>Aegilops speltoides</i> var. <i>ligustica</i>	S	KU170102
Clae 70	<i>Aegilops bicornis</i>	S ^b	KU170103
Clae 47	<i>Aegilops bicornis</i>	S ^b	KU170104
PI 604123	<i>Aegilops longissima</i>	S ^l	KU170105
PI 604122	<i>Aegilops longissima</i>	S ^l	KU170106
PI 584388	<i>Aegilops sharonensis</i>	S ^{sh}	KU170107
PI 584358	<i>Aegilops sharonensis</i>	S ^{sh}	KU170108
PI 584369	<i>Aegilops sharonensis</i>	S ^{Sh}	KU170109
AS 2386	<i>Aegilops tauschii</i>	D	KU170110
AS 2404	<i>Aegilops tauschii</i>	D	KU170111
YX 1212	<i>Aegilops tauschii</i>	D	KU170112
YX 1216	<i>Aegilops tauschii</i>	D	KU170113
PI 614034	<i>Triticum aestivum</i>	ABD	KU170114

Table 2 Primer sequences used in this study

Primer	Sequence(5'→3')	Description
APL-KHF1	GGTACCATGTCATCGATGCAGTTCAGCA	Primer for cDNA isolation
APL-KHR1	AAGCTTCTACACGACGGTGCCGTCC	Primer for cDNA isolation
APL-F1	ATGTCATCGATGCAGTTCAGCA	Primer for DNA isolation
APL-R1	GCGGTGAATGTGACGATTAAGA	Primer for DNA isolation
APL-F2	TCGATATTCCCATGAGCAACTG	Primer for DNA isolation
APL-R2	GGTTGAGCAGAACATCTCTTTTGA	Primer for DNA isolation
APL-F3	GTGGACACCAGTTTTCTCAATTC	Primer for DNA isolation
APL-R3	CTACACGACGGTGCCGTCC	Primer for DNA isolation
qAPL-F	TCAATTTGCGCCATCGACGAC	Primer for real-time quantification PCR
qAPL-R	ACATTGTGGTCATGCAGAGC	Primer for real-time quantification PCR
GAPDH-F	GAGTCTGCCCACCCATTCGTAA	Primer for real-time quantification PCR
GAPDH-R	GACATGCCATAGGTTTCAGCGAC	Primer for real-time quantification PCR

Figure legends

Fig. 1 Gene and protein structures of AGP-L. **a** Gene structure of *Agp2* and location of primers. Black boxes and thick lines represent exons and introns, respectively. Regions that vary within *Agp2* from the species examined are indicated by red boxes. The positions and directions of primers are indicated by arrows. **b** Protein structure of AGP-L. Symbols in gray boxes indicate conserved features in AGP-L. TP = transit peptides, LBS = ligand binding site, DI = dimer interface, and NDI = N-terminal domain interface. The blue boxes represent strands and the red boxes represent helices.

Fig. 2 Sequence alignment of AGP-L amino acid sequences from 10 species. Amino acids that constitute conserved domains have been labeled. The violet regions indicate a similarity of 100%; the pink regions indicate a similarity $\geq 75\%$, the blue regions indicate a similarity was $\geq 50\%$, and the white regions indicate a similarity was $< 50\%$.

Fig. 3 Phylogenetic relationships from 18 *Triticeae* accessions based on *Agp2* gene sequences. The bootstrap values were calculated based on 1000 replications and were showed as percentages.

Fig. 4 Analysis of starch accumulation and *Agp2* expression pattern. **a** Total starch contents and expression levels of *Agp2* in wheat grains at different developmental stages. **b** Correlations between starch accumulation and *Agp2* expression in grains of wheat.

Gene and protein structures of AGP-L. a Gene structure of AGP-L and location of primers. Black boxes and thick lines represent exons and introns, respectively. The varied regions in AGP-L from a range of species are indicated by red box. The position and direction of primers are indicated by arrows. b Protein structure of AGP-L. Symbols in gray boxed indicate conserved features in AGP-L. TP represent transit peptides, LBS represent ligand binding site, DI represent dimer interface and NDI represent N-terminal domain interface. The blue boxes represent strand structure and the red boxes represent helix structure.

1833x1077mm (72 x 72 DPI)

Sequence alignment of AGP-L amino acid sequences from 10 species. Amino acids that constitute conserved domains have been labeled. The violet regions indicate a similarity of 100%; the pink regions indicate a similarity $\geq 75\%$, the blue regions indicate a similarity was $\geq 50\%$, and the white regions indicate a similarity was $< 50\%$.
1500x2016mm (72 x 72 DPI)

Phylogenetic relationships from 18 Triticeae accessions based on Agp2 gene sequences. The bootstrap values were calculated based on 1000 replications and were showed as percentages.

1000x558mm (72 x 72 DPI)

Analysis of starch accumulation and Agp2 expression pattern. a Total starch contents and expression levels of Agp2 in wheat grains at different developmental stages. b Correlations between starch accumulation and Agp2 expression in grains of wheat.
1389x1968mm (72 x 72 DPI)

Fig. S1 DNA sequence alignment of Agp2 from 18 Triticeae accessions. The violet regions indicate a similarity of 100%; the pink regions indicate a similarity $\geq 75\%$, the blue regions indicate a similarity was $\geq 50\%$, and the white regions indicate a similarity was $< 50\%$.

Draft

P1428320	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1090
P1428170	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1095
P1428174	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1098
P1487232	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1043
P1542249	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1029
C1AE70	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1043
P1604123	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1043
P1604122	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1043
P1584388	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1043
P1584369	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1043
A23386	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1054
A24044	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1054
YK1216	GATGTTACAGCGTCACGGA	ATTTCCCATGAGCAACGTTCTTCA	AGTGGCATCAACAGAGATTGCTATG	CAGGTC	CAACTCGGG	TTAAATGTCACATTCA	CGCACTACTCTGGGGGGGGATCAATTT	1054

[illegible][illegible][illegible]

P1428320	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1560
P1427328	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1561
P1428170	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1562
P1428174	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1563
P1170209	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1564
P1487232	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1565
AY522449	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1566
C1A670	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1567
CA1427	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1568
P141203	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1569
P1604122	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1570
U854388	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1571
U854358	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1572
P1484369	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1573
A52386	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1574
A52404	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1575
YX1212	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1576
YX1216	CTTTCAGAGTCTGGTGGGCTCTTTACACGCTCAGTGATGATGTTGTTGAATCGCGCGCAGCGCTGTCGACCTCATCTTATCATGCGCCCTGT	1577

[illegible][illegible]