Alternative splicing results in a lack of starch synthase IIa-D in Chinese wheat landrace

Journal:	<i>Genome</i>
Manuscript ID	gen-2017-0246.R1
Manuscript Type:	Article
Date Submitted by the Author:	04-Jan-2018
Complete List of Authors:	<p>Zhou, Yan-Jie ; Sichuan Agricultural University, Triticeae Research Institute Yang, Qiang; Sichuan Agricultural University, Triticeae Research Institute Zhong, Xiao-Juan; Sichuan Agricultural University, Triticeae Research Institute Deng, Mei; Sichuan Agricultural University, Triticeae Research Institute Ma, Jian; Triticeae Research Institute, Sichuan Agricultural University Qi, Peng-Fei; Triticeae Research Institute, Sichuan Agricultural University Wang, Ji-Rui; Sichuan Agricultural University, Triticeae Research Institute Chen, Guo-Yue; Triticeae Research Institute, Sichuan Agricultural University Liu, Yaxi; Triticeae Research Institute, Sichuan Agricultural University Lu, Zhen-Xiang; Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada Li, Wei; Triticeae Research Institute, Sichuan Agricultural University Lan, Xiu-Jin; Sichuan Agricultural University, Triticeae Research Institute Wei, Yu-Ming; Sichuan Agricultural University, Triticeae Research Institute Zheng, Youliang; Sichuan Agricultural University, Triticeae Research Institute Jiang, Qian-Tao; Sichuan Agricultural University, Triticeae Research Institute</p>
Is the invited manuscript for consideration in a Special Issue? :	N/A
Keyword:	Chinese wheat landrace, starch, amylose, TaSSIIa, alternative splicing

Alternative splicing results in a lack of starch synthase IIa-D in Chinese wheat landrace

Yan-Jie Zhou¹, Qiang Yang¹, Xiao-Juan Zhong¹, Hua-Ping Tang¹, Mei Deng¹, Jian Ma¹, Peng-Fei Qi¹, Ji-Rui Wang¹, Guo-Yue Chen¹, Ya-Xi Liu¹, Zhen-Xiang Lu², Wei Li³, Xiu-Jin Lan¹, Yu-Ming Wei¹, You-Liang Zheng¹, Qian-Tao Jiang^{1*}

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

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

³*College of Agronomy, Sichuan Agricultural University, Chengdu, Sichuan 611130, China*

*Corresponding author

Dr. Qian-Tao Jiang

E-mail: qiantaojiang@sicau.edu.cn

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

Abstract: We evaluated the SGP-1 protein composition of 368 Chinese wheat landraces using SDS-PAGE. The SGP-D1 null type was identified in three accessions (Xiaoqingmang, Pushanbama, and P119). An 18-bp deletion and 9-bp variation were found at the junction region of the 7th intron and 8th exon, leading to deletion of the intron-exon junction recognition site AG when aligned the 8261-bp DNA sequence of *TaSSIIa-D* in Pushanbama with that of Chinese Spring. Four cDNA types with mis-spliced isoforms were subsequently detected through amplification of *TaSSIIa-D* cDNAs. Among these, 9 type II cDNAs with a 16-bp deletion in the 8th exon were detected, indicating that the major transcriptional pattern of *TaSSIIa* in Pushanbama is type II. In the type IV cDNA, a 97-bp sequence remains undeleted in the end of the 5th exon. The amylose content in Pushanbama was significantly higher than that in all control lines under field conditions, which suggested that deletion of SGP-D1 has an efficient impact on amylose content. As the *TaSSIIa* gene plays an important role in regulating the content of amylose, it is anticipated that these natural variants of *TaSSIIa-D* will provide useful resources for quality improvement in wheat.

Key Words: Chinese wheat landrace, starch, amylose, *TaSSIIa*, alternative splicing

Introduction

Starch is a major component (65%–75%) of bread wheat grains. Its composition directly determines wheat yield, and together with its physicochemical properties can affect flour processing quality. Furthermore, starch also serves as the primary carbohydrate component in the human diet (Rahman et al. 2000). Starch is a complex carbohydrate consisting of two types of glucose polymers (amylose and amylopectin), which differ primarily in terms of the degree of polymerization and the number of side branches. Amylose is a linear chain consisting of α -(1,4)-linked glucose units with a low degree of polymerization, whereas amylopectin is a very large, highly branched glucose polymer in which α -(1,4)-linked polymers are connected by α -(1,6)-linkages with a high degree of polymerization (Zobel et al. 1988). In normal wheat grains, the amylose to amylopectin ratio of total starch is approximately 1:3 (Hung et al. 2006). However, marked changes in this ratio can occur in certain mutant genotypes of wheat, barley, and rice, with amylose increasing by up to 80% or amylopectin by up 99%–100% (Nakamura et al. 1995; Regina et al. 2015). The different ratios of starch components lead to differences in the structure and physicochemical properties of wheat, and thus end-product quality.

High-amylose starch has attracted particular interest because a positive relationship between resistant starch (RS) and amylose content has been established through several studies (Lee et al. 2013; Morell et al. 2003; Rahman et al. 2007; Yamamori et al. 2006; Yang et al. 2006). However, in addition to the actual amylose content, modified amylopectin also contributes to RS via the influence of longer amylopectin chain lengths and low 6' OH phosphate content (Regina et al. 2012). The products of RS degradation are not absorbed in the small intestine but are instead

fermented in the colon, producing short-chain fatty acids. This enhances colonic function and reduces the risk of intestinal ailments by increasing the amount of dietary fiber (Nugent et al. 2005; Robertson et al. 2012). In addition, RS has also been shown to reduce the risk of dietary protein-induced DNA damage (Conlon et al. 2012) and to improve cholesterol metabolism. Since RS is not digested in the small intestine, it also reduces the glycemic index of food (Shamai et al. 2003).

To date, a series of enzymes have been shown to participate in the synthesis of starch, which include granule-bound starch synthase (GBSS, involved mainly in amylose synthesis), and the soluble isoforms: sucrose synthase I (SSI, involved in the synthesis of small chains of amylopectin), SSII and SSIII (which play a major role in amylopectin synthesis), SSIV (recently found to be involved in the control of starch granule number and starch granule initiation), starch branching enzyme (SBE), and debranching enzymes (DBE, which induces an efficient clustering of branch points and the formation of crystalline lamellae) (Delvalle et al. 2005; Maddelein et al. 1994; Zhang et al. 2005; Delrue et al. 1992; Wattebled et al. 2008; Coppin et al. 2005). SSII has two forms, SSIIa and SSIIb, the former of SSIIa is expressed predominantly in the endosperm, whereas transcripts of the SSIIb are detected mainly in leaves at low abundance (Harn et al. 1998). SSIIa, also known as SGP-1 protein, plays an important role in the synthesis of amylopectin in all cereals, and its reduction or complete loss can alter the amount of amylose and total starch in the endosperms of barley, wheat, rice, and maize (Morell et al. 2003; Yamamori et al. 2000; Umemoto et al. 2002). Moreover, the three SGP-1 proteins, SGP-A1, B1, and D1, are produced by three homoeologous *SSII* genes, *SSII-A*, *B*, and *D*, which are located on the short arms of group-7 chromosomes (7AS, 7BS, and 7DS) (Yamamori et al. 1996).

Several studies have focused on the creation of genetic variation in *SSIIa* and have discussed its role in wheat quality breeding (Botticella et al. 2016; Li et al. 2016; Hogg et al. 2013; Luo et al. 2015; Yamamori et al. 2000). Two *SSIIa* null mutations lines selected for durum wheat showed novel rheological behavior and an increased content of all major dietary fiber components and RS (Botticella et al. 2016). Furthermore, some frame shift mutations of *SSIIa* induced by EMS mutagenesis have also been shown to affect gene function (Li et al. 2016).

Wheat landraces are known to possess abundant genetic diversity and variation compared to cultivars. In our recent work, we demonstrated that a transposon insertion in one novel null *Wx-B1* allele (*Wx-B1n*) resulted in the loss of its function in Chinese wheat landrace (Zhang et al. 2017). However, to the best of our knowledge the SGP-1 composition of Chinese wheat landraces has not been previously explored. In this study, we accordingly evaluated the SGP-1 protein composition of 368 accessions of Chinese wheat landrace, which enabled us to detect some null *TaSSII* genes. Furthermore, we characterized the molecular mechanisms of silencing in these genes and the effects on starch composition ratio.

Materials and methods

Plant materials

In total 368 wheat accessions, including 325 wheat landraces and 43 wheat cultivars (lines), were used for evaluation of variation in SGP-1 proteins (Supplementary Table 1). All the accessions of Chinese landrace and cultivar accessions were kindly provided by the National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences and the Triticeae Research Institute, Sichuan Agricultural

University, respectively. As controls, we used diploid *Triticum monococcum* and *Aegilops tauschii*, tetraploid *Triticum turgidum*, Chinese Spring nullisomic/tetrasomic (NT) lines (Sears et al. 1985) , and Chinese Spring.

Starch extraction and gel electrophoresis of SGP-1 proteins

SGP-1 protein extraction was performed according to Yamamori and Endo (Yamamori et al. 1996). The whole flour from single mature grains was mixed with 700 μ L of buffer I [55 mM Tris-HCl, pH 6.8, 2.3% (w/v) sodium dodecyl sulfate, 5% (v/v) β -mercaptoethanol, 10%(v/v) glycerol]. The homogenate was then centrifuged at $10,000 \times g$ for 5 min. The pellets were washed three times with buffer I, twice with 1 mL distilled water, and then twice with 800 μ L acetone. The residue was mixed with 0.1 mg/ μ L of buffer I, containing 0.005% (w/v) bromophenol blue, heated in a boiling bath for 5 min, cooled in ice, and then centrifuged as described above. Aliquots of the supernatant (10 μ L) were separated using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in a discontinuous Tris-HCl-SDS buffer system (pH: 6.8/8.8) with a polyacrylamide concentration of 12% (w/v) (Ayala et al. 2015). Electrophoresis was performed at a constant current of 100 V/gel and was continued for 2 h after the tracking dye had migrated off the gel. Protein bands were detected by silver staining. For each cultivar (line), we evaluated at least three independent seeds as three replications. Identification of the SGP-1 protein was performed as described by Yamamori and Endo (Yamamori et al. 1996) with reference to control material (*Triticum monococcum*, *Aegilops tauschii*, tetraploid *Triticum turgidum*, Chinese Spring NT lines (N7AT7B and N7DTB), and Chinese Spring).

Isolation of genomic sequence of *TaSSIIa-D*

Genomic DNA was extracted from the leaves of two-week-old seedlings using a

Plant Genomic DNA Kit (TIANGEN). The complete DNA of *TaSSIIa-D* was amplified using five overlapping primers pairs (SSIIa-D-F1/SSIIa-D-R1, SSIIa-D-F2/SSIIa-D-R2, SSIIa-D-F3/SSIIa-D-R3, SSIIa-D-F4/SSIIa-D-R4, and SSIIa-D-F5/SSIIa-D-R5) that were designed based on the conserved 5'- or 3'-end nucleotide regions of the available *TaSSIIa* sequences (Table 1). All these primers pairs were used to amplify the five sections of *TaSSIIa-D* DNA, including the predicted 1367 bp of the 5'upstream sequence, and 138 bp of the 3' downstream sequence (Fig. 1). Errors that could be introduced in the sequences by PCR were minimized by using high-fidelity LA Taq polymerase (TaKaRa, Dalian, China) in GC buffer for GC-rich templates. PCR was performed in 30 μ L reaction volumes, containing 15 μ L 2 $\times$ GC buffer I, 10 μ M of dNTP, 20 μ mol of each primer, 1 U of LA Taq, and $\sim$ 100 ng of template DNA. All reagents were from Takara Biotechnology Co., Ltd. (<http://www.takara.com.cn>). Reaction conditions were 94°C for 5 min, followed by 35 cycles of 94°C for 45s, annealing at 55~59°C for 45s, 72°C for 2 min, and a final extension of 72°C for 10 min. The PCR products were separated on 2% agarose gels, and all DNA fragments were recovered, purified, and ligated into pEASY-T5 (TransGen Biotech, Beijing, China). The ligated mixtures were transformed into *Escherichia coli* DH5 α competent cells. DNA sequencing was performed by TsingKe Biotech (Beijing, China). The final nucleotide sequence for each fragment was determined from the sequencing results of at least three independent clones. The resultant complete genomic sequences of *TaSSIIa-D* were assembled based on overlapping sequences.

Sequencing mis-spliced *TaSSIIa-D* transcripts

Developing seeds at approximately 10 days post-anthesis were selected for RNA extraction. Total RNA was extracted using a Plant Genomic RNA Kit (BIOFIT) and RNAs

were reverse transcribed using a PrimeScript™RT reagent Kit with gDNA Eraser (TAKARA). The paired primers SSIIa-D-F6 and SSIIa-D-R6 were designed to amplify the complete cDNA of *TaSSIIa-D*, including 152 bp of the 5'UTR and 149 bp of the 3'UTR. PCR was also performed using high-fidelity LA Taq polymerase, the cycling parameters and PCR program were the same as described for the isolated genomic DNA. Plasmid DNAs from single clones were completely sequenced.

Sequences analysis

All sequences of each clone were aligned to IWGSC wheat scaffold database to conform its original chromosome. The nucleotide sequences were then assembled and aligned using DNAMAN V8.0 and DNASTAR V7.1. In addition, the *TaSSIIa-D* gene of Chinese Spring and Fielder *TaSSIIa-A* (NCBI accession: AB201445, AJ269502), *TaSSIIa-B* (NCBI accession: AB201446, AJ269503), and *TaSSIIa-D* (NCBI accession: AB201447, AJ269504) were used for reference in multiple sequence alignments.

Starch test

The contents of total starch and amylose in the sampled grains were determined using a "Total starch assay kit" (Pcode: K-TSTA) and an "Amylose assay kit" (Pcode: K-AMYL), respectively, which were from Megazyme International Ireland Ltd. Each sample was tested five times. Mutant lines and control materials were grown both in a greenhouse and in the field.

Results

Identification of SGP-1 protein varieties

SGP-1 bands were identified in all 368 wheat landraces and wheat cultivars (lines) by SDS-PAGE. Two distinct bands were detected in 365 lines, which were determined to be SGP-B1 and SGP-D1 by comparison with the bands derived from

control diploid (*Triticum monococcum*, *Aegilops tauschii*), tetraploid (*Triticum turgidum*), and Chinese Spring NT lines (N7AT7B, N7DTB). Of these, SGP-B1, which has a lower molecular weight, showed high mobility (also detected in *Triticum turgidum*, N7AT7B, and N7DTB controls), whereas SGP-D1 with a higher molecular weight showed lower mobility (also detected in *Aegilops tauschii* and N7AT7B controls). However, in the three accessions Pushanbamai, Xiaoqingmang, and P119 we detected only one of these bands (SGP-B1), whereas SGP-D1 was deleted (Fig. 2, Supplementary Fig. 1). Among these three accessions, we selected Pushanbamai for further cloning experiments.

Sequence analysis of *TaSSIIa-D* DNA in Pushanbamai

On the basis of an NCBI database search, *TaSSIIa-D* contains a total of eight exons and seven introns (6932 bp) as compared to the homologous sequence in Chinese Spring (NCBI accession: AB201447). The primer sets SSIIa-D-F1/SSIIa-D-R1, SSIIa-D-F2/SSIIa-D-R2, SSIIa-D-F3/SSIIa-D-R3, SSIIa-D-F4/SSIIa-D-R4, and SSIIa-D-F5/SSIIa-D-R5 were designed to clone the genomic sequences in Pushanbamai between the first exon and the third exon, the sequence between the third intron and the seventh exon, the sequence of the seventh intron, the eighth exon, and the gDNA sequence upstream of the start codon of *TaSSIIa-D*, respectively, which yielded sequences of 1513 bp, 2210 bp, 2208 bp, 1652 bp, and 1574 bp, respectively. In order to assemble the full length *TaSSIIa-D* genomic sequence, each two sequences amplified by the designed primers were overlapped by at least 200 bp. All sequences of single clones were blasted against the Chinese Spring shotgun assembly database in order to determine the *TaSSIIa-D* product (https://urgi.versailles.inra.fr/blast/?dbgroup=wheat_all&program=blastn). On the

basis of the sequence results for the five PCR products, we assembled an 8261bp full-length genomic sequence of *TaSSIIa-D* in Pushanbamai.

The assembled *TaSSIIa-D* gene contains 1367 bp upstream of the start codon, 6756 bp from ATG to TGA, and 138 bp downstream of the stop codon. Alignment of the *TaSSIIa-D* genomic sequence showed that the upstream sequence in Pushanbamai was completely consistent with that in Chinese Spring, whereas an 18-bp deletion and 9-bp variation were detected at the junction region of the 7th intron and 8th exon. The 18-bp deletion was located at the end of the 7th intron and the 9-bp variation was located at the start of the 8th exon region (Fig. 3).

Characterization of *TaSSIIa-D* cDNA and amino acid sequences in Pushanbamai

The sequences of *TaSSIIa-D* cDNA were amplified using the primer set SSIIa-D-F6/SSIIa-D-R6, and a total of 12 independent clones were sequenced from the cloning vector using the M13 primer. Four types of *TaSSIIa-D* cDNA were identified in Pushanbamai and all types were found to be incorrectly spliced. These four types of mis-spliced *TaSSIIa-D* cDNAs were named type I, type II, type III, and type IV. The sequences of all four types included 152 bp at the 5'UTR and 149 bp at the 3'UTR, whereas the length from the start codon to the stop codon was 2335 bp, 2384 bp, 2439 bp, and 2680 bp, for type I, type II, type III, and type IV, respectively. Among the 12 mis-spliced *TaSSIIa-D* clones detected, nine type II sequences were identified, whereas each of the remaining three types was detected only once (Fig. 4).

When compared with the *TaSSIIa-D* mRNA sequence of Chinese Spring, it was found that type I is missing 65 bp at the beginning of the 8th exon, leading to a shorter cDNA length, which results in a frame-shift that generates a premature

termination codon within exon 8. This correspondingly gives rise to a shorter protein of only 488 amino acids (Fig. 5a). Similarly, type II is missing 16 bp at the beginning of the 8th exon, resulting in a frame-shift that generates a premature termination codon within exon 8, and a shorter protein of 514 amino acids (Fig. 5b). Type III has a 39-bp less deletion in the 7th intron and a 9-bp variation (GGCCGTGGC/ACCGAGTTG) within the 8th exon, which result in a longer, but in-frame, cDNA encoding a SGP-1 protein of 812 amino acids containing a 13-amino acid insert (LFHSSVLRWANSF), with a three-amino acid variation (GRG/TEL) at the start of the 8th exon (Fig. 5c). Type IV has a longer sequence with a 183-bp less deletion in the 7th intron and the same variation at the start of the 8th exon as type III. Furthermore, a 97-bp sequence of the whole 5th intron of the DNA is not been excised from the type IV cDNA during the process of transcription, which has resulted in a frame-shift that generates a premature termination codon at the 5th exon, thereby generating a shorter protein of only 416 amino acids (Fig. 5d).

Starch test

The contents of total starch and amylose were determined under both greenhouse and field conditions. In the greenhouse, there was a highly significant difference between the total starch contents of Pushanbamai and Chuannong16. However, no significant difference was observed between Pushanbamai and Shumai969 and Chinese Spring. Moreover, the amylose content of Pushanbamai was significantly lower than that in three control wheat lines (Chuangnong16, Shumai969, and Chinese Spring). Interestingly, the content of amylose in Pushanbamai was significantly higher than that in all control lines under field conditions, whereas the total starch content was comparable with that of the control lines.

Discussion

Starch is an important reserve in wheat endosperm, which contains both amylopectin and amylose. Wheat grain quality is determined by the amylose/amylopectin ratio and the structure of amylopectin. Generally, natural variation or mutagenesis leading to novel genetic variants has played a key role in crop genetic improvement. Of these, lines showing natural variation have high environmental adaptability, which can be directly exploited in breeding.

SSIa is a candidate gene for controlling differences in alkali disintegration of rice grains and the chain-length distribution of amylopectin in rice and wheat (Umemoto et al. 2002; Konik-Rose et al. 2007). Natural variation in rice has been detected and the potential importance of two of the amino acid residues for starch association of *SSIa* in changing of amylopectin quantities and shape has been identified (Umemoto et al. 2004). Lines carrying null alleles for each of these SGP-1 proteins indicated that the loss of each SGP was due to MITE insertion and alternative splicing respectively (Yamamori et al. 2000; Shimbata et al. 2005). However, the different effects of *SSIa* null mutation on endosperm amylose content of barley, wheat, and rice indicated that regulation of starch synthesis in *SSIa* mutants and the inactive *SSIa* variant may occur at the post-translational level or that the altered amylopectin structure reduces the affinity of *SSI* and *SBEIIb* for amylopectin stroma in cereals (Luo et al. 2015). Landrace collections in general show a considerably higher level of genetic diversity than cultivars (Zhang et al. 2017; Cavanagh et al. 2013; Hao et al. 2010; Hao et al. 2008). There is a widely recognized imperative for breeding programs to use this genetic diversity by introducing landraces as donors, which will provide basic information on the unique characteristics of adaptation, performance, stress

tolerance, and end-use quality exhibited by landrace cultivars (Moore 2015). Thus, these lines have been widely used for detection of allelic variation and distribution of functional genes (Yang et al. 2009; Chen et al. 2006; Zhang et al. 2008). The investigation of Chinese landraces is providing new insights into the genetic diversity among starch-related genes in non-domesticated lines.

Pre-mRNA splicing produces extensive proteomic and regulatory diversity during development. Accurate modulation of pre-mRNA splicing is essential for normal cellular functions in wheat, rice, and Arabidopsis (Fang et al. 2017; Xia et al. 2017; Marquez et al. 2012; Su et al. 2017). In plants, approximately 20% of the genes have multiple isoforms of transcripts that play various role in photosynthesis, disease resistance, flowering, quality, and other physiology processes (Wang et al. 2006; Barbazuk et al. 2008; Reddy 2007; Sestili et al. 2015).

Alternative splicing in *TaSSIIa* has previously been detected at the region surrounding the junction of the fifth exon and intron of the null *TaSSII-D* allele (Shimbata et al. 2005). Another study has shown that only a portion (35%) of transcripts was correctly spliced and the majority (65%) of transcripts were incorrectly spliced due to multiple mis-splicing events, and all of the mis-spliced cDNA transcripts observed resulted in a premature stop codons due to a shift in the open reading frame. Thus, the mis-spliced cDNAs have been deduced to encode incomplete proteins (Fang et al. 2017). In the present study, a complete exon 5 was retained in the type IV cDNA of Pushanbamai in the same region and showed a different type of splicing. However, we found that genomic DNA sequences of Pushanbamai are completely consistent with those of Chinese Spring, indicating that multiple splicing events probably occur in this region of Pushanbamai. However, a

novel 9-bp variation of the 3' splicing site in Pushanbamai leads to loss of the donor of the splicing acceptor site and expectation of abnormal excision of intron 7 from the pre-mRNA. Further cDNA sequencing confirmed this expectation, and as the result, four types of cDNA were detected. Among these, sequencing determined that type II-spliced cDNA was the main form and contained a premature stop codon following exon 8, resulting in SGP-1 deletion. Therefore, this is a novel variant among SGP-1 deletion line.

With regards to the contents of starch component, we obtained differing results under greenhouse and field conditions. We found that the amylose content of Pushanbamai was higher than that in control lines when grown under field condition. Generally, most researchers have essentially all shown that a significantly increase of triple null lines on in amylose content (Konik-Rose et al. 2007, Shimbata et al. 2012). Nevertheless, SGP-1 triple null line was unsearched in natural wheat accessions since now, mainly because SGP-1 triple or double null alleles would reduce plant fitness. Therefore, 368 Chinese wheat or landraces original growing from various growing environments have been screened, and we only detected 3 SGP-D1 null lines, consistent with that SGP-D1 null have slightly increase amylose content (Yamamori et al. 2000). SGP-D1 null line Pushanbamai was a strong winter wheat accession, its amylose content inconsistent with field environment, probably due to unsuitable growth characterize in green house. Furthermore, this result highlight that appropriate amylose probably reflected in fitting environment for plant. Notably, another important protein Wx, showing completely bands in Pushanbamai and the control lines. Thus, the increasing of amylose mainly at least partially, due to lack of SGP-1 in this study, which was consistent to previously study (Yamamori et al. 2000).

Importantly, this difference was attributed to a novel mechanism whereby mis-splicing of the *TaSSIIa* gene produced abnormal transcripts lacking SGP-1, eventually leading to an increase in amylose content. In contrast, no significant differences in total starch content was observed between Pushanbamai and all control lines, which was probably because the function of SGP-1 can be implemented by other starch synthesis-related enzymes. Thus, we anticipate identifying *TaSSIIa-A*, *TaSSIIa-B*, and mutants of other related genes for further pyramiding in breeding.

In summary, in the current study we identified a mechanism (alternative splicing) that produced three SGP-1 deletion lines, which reveal the natural diversity of *TaSSIIa-D* deletion mechanisms and can also be used to produce novel lines with different amylopectin content in wheat improvement.

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

- Ayala, M., Alvarez, J. B., Yamamori, M., & Guzmán, C. 2015. Molecular characterization of waxy alleles in three subspecies of hexaploid wheat and identification of two novel wx- b1, alleles. *Theor Appl Genet*, 128(12), 2427.
- Barbazuk, W. B., Fu, Y., & McGinnis, K. M. 2008. Genome-wide analyses of alternative splicing in plants: opportunities and challenges. *Genome Res*, 18(9), 1381-1392.
- Botticella, E., Sestili, F., Ferrazzano, G., Mantovani, P., Cammerata, A., & D'Egidio, M. G., et al. 2016. The impact of the *ssii* null mutations on grain traits and composition in durum wheat. *Breeding Sci*, 66(4), 572.
- Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., & Kiani, S., et al. 2013. Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc Natl Acad Sci*, 110(20), 8057-62.
- Chen F, He ZH, Xia XC, Xia LQ, Zhang XY, Lillemo M, Morris CF. 2006. Molecular and biochemical characterization of puroindoline a and b alleles in Chinese landraces and historical cultivars. *Theor Appl Genet* 112:400-409
- Conlon, M. A., Kerr, C. A., Mcsweeney, C. S., Dunne, R. A., Shaw, J. M., & Kang, S., et al. 2012. Resistant starches protect against colonic dna damage and alter microbiota and gene expression in rats fed a western diet. *J Nutr*, 142(5), 832-40.
- Coppin, A., Varré, J. S., Lienard, L., Dauvillée, D., Guérardel, Y., & Soyergobillard, M. O., et al. 2005. Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. *J Mol Evol*, 60(2), 257.
- Delrue, B., Fontaine, T., Routier, F., Decq, A., Wieruszeski, J. M., & Van, D. K. N., et al. 1992. Waxy *Chlamydomonas reinhardtii*: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. *J Bacteriol*, 174(11), 3612.
- Delvallé, D., Dumez, S., Wattebled, F., Roldán, I., Planchot, V., & Berbezy, P., et al. 2005. Soluble starch synthase i: a major determinant for the synthesis of amylopectin in *Arabidopsis thaliana* leaves. *Plant J*, 43(3), 398.
- Fang T, Carver BF, Hunger RM, Yan L. 2017. Mis-Spliced Lr34 Transcript Events in Winter Wheat. *PLOS One* 12(1): e0171149. doi:10.1371/journal.pone.0171149
- Hao, C. Y., Dong, Y. C., Wang, L. F., You, G. X., Zhang, H. N., & Ge, H. M., et al. 2008. Genetic diversity and construction of core collection in Chinese wheat genetic resources. *Sci Bull*, 53(10), 1518-1526.
- Hao, C., Wang, L., Ge, H., Dong, Y., & Zhang, X. 2010. Genetic diversity and linkage disequilibrium in Chinese bread wheat (*Triticum aestivum* L.) revealed by SSR markers. *Plos One*, 6(2), e17279.

- Harn, C., Knight, M., Ramakrishnan, A., Guan, H., Keeling, P. L., & Wasserman, B. P. 1998. Isolation and characterization of the zSSIa and zSSIb starch synthase cDNA clones from maize endosperm. *Plant Mol Biol*, 37(4), 639-649.
- Hogg, A. C., Gause, K., Hofer, P., Martin, J. M., Graybosch, R. A., & Hansen, L. E., et al. 2013. Creation of a high-amylose durum wheat through mutagenesis of starch synthase ii (ssii). *J Cereal Sci*, 57(3), 377-383.
- Hung, P. V., Maeda, T., & Morita, N. 2006. Waxy and high-amylose wheat starches and flours-characteristics, functionality and application. *Trends Food Sci Tech*, 17(8), 448-456.
- Konik-Rose, C., Thistleton, J., Chanvrier, H., Tan, I., Halley, P., & Gidley, M., et al. 2007. Effects of starch synthase iia gene dosage on grain, protein and starch in endosperm of wheat. *Theor Appl Genet*, 115(8), 1053.
- Lee, J. S., Ee, M. L., Chung, K. H., & Othman, Z. 2013. Formation of resistant corn starches induced by gamma-irradiation. *Carbohydr Polym*, 97(2), 614-617.
- Li, W., Guo, H., Wang, Y., Xie, Y., Zhao, L., & Gu, J., et al. 2016. Identification of novel alleles induced by ems-mutagenesis in key genes of kernel hardness and starch biosynthesis in wheat by tilling. *Genes Genom*, 1-9.
- Luo, J., Ahmed, R., Kosar-Hashemi, B., Larroque, O., Butardo, V. M., Tanner, G. J., & Millar, A. 2015. The different effects of starch synthase IIa mutations or variation on endosperm amylose content of barley, wheat and rice are determined by the distribution of starch synthase I and starch branching enzyme IIb between the starch granule and amyloplast stroma. *Theor Appl Genet*, 128(7), 1407-1419.
- Maddelein, M. L., Libessart, N., Bellanger, F., Delrue, B., D'Hulst, C., & Koornhuyse, N. V. D., et al. 1994. Toward an understanding of the biogenesis of the starch granule determination of granule-bound and soluble starch synthase functions in amylopectin synthesis. *J Biol Chem*, 269(40), 25150-7.
- Marquez, Y., Brown, J. W., Simpson, C., Barta, A., & Kalyna, M. 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in arabidopsis. *Genome Res*, 22(6), 1184.
- Moore, G. 2015. Strategic pre-breeding for wheat improvement. *Nat Plants*, 1(3), nplants201518.
- Morell, M. K., Behjat, K., Mark, C., Samuel, M. S., Peter, C., & Sadequr, R., et al. 2003. Barley sex6 mutants lack starch synthase iia activity and contain a starch with novel properties. *Plant J*, 34(2), 173-85.
- Nakamura, T., Yamamori, M., Hirano, H., Hidaka, S., & Nagamine, T. 1995. Production of waxy (amylose-free) wheats. *Mol Genet Genomics*, 248(3), 253-259.
- Nugent, A. P. 2005. Health properties of resistant starch. *Nutrition Bulletin*, 30(1), 27-54.
- Rahman, S., Bird, A., Regina, A., Li, Z., Ral, J.P., McMaugh, S., Topping, D. and Morell, M. 2007. Resistant starch in cereals: exploiting genetic engineering and genetic variation. *J. Cereal Sci*. 46, 251–

260.

Rahman, S., Li, Z., Batey, I., Cochrane, M. P., Appels, R., & Morell, M. 2000. Genetic alteration of starch functionality in wheat. *J Cereal Sci*, 31(1), 91-110.

Reddy, A. S. 2007. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annu. Rev. Plant Biol.*, 58, 267-294.

Regina, A., Berbezy, P., Kosar-Hashemi, B., Li, S., Cmiel, M., & Larroque, O., et al. 2015. A genetic strategy generating wheat with very high amylose content. *Plant Biotechnol J*, 13(9), 1276.

Regina, A., Blazek, J., Gilbert, E., Flanagan, B. M., Gidley, M. J., & Cavanagh, C., et al. 2012. Differential effects of genetically distinct mechanisms of elevating amylose on barley starch characteristics. *Carbohydr Polym*, 89(3), 979.

Robertson, M. D. 2012. Dietary-resistant starch and glucose metabolism. *Curr Opin Clin Nutr*, 15(4), 362.

Sears ER, Miller TE. 1985. The history of Chinese spring wheat. *Cereal Res Commun* 13:261–263

Sestili, F., Palombieri, S., Botticella, E., Mantovani, P., Bovina, R., & Lafiandra, D. 2015. TILLING mutants of durum wheat result in a high amylose phenotype and provide information on alternative splicing mechanisms. *Plant Sci*, 233, 127-133.

Shamai, K., Bianco-Peled, H., & Shimoni, E. 2003. Polymorphism of resistant starch type iii. *Carbohydr Polym*, 54(3), 363-369.

Shimbata, T., Nakamura, T., Vrinten, P., Saito, M., Yonemaru, J., & Seto, Y., et al. 2005. Mutations in wheat starch synthase ii, genes and pcr-based selection of a sgp-1 null line. *Theor Appl Genet*, 111(6), 1072-1079.

Shimbata, T., Ai, Y., Fujita, M., Inokuma, T., Vrinten, P., & Sunohara, A., et al. 2012. Effects of homoeologous wheat starch synthase iia genes on starch properties. *J Agr Food*, 60(48), 12004.

Su, C., Zhao, H., Zhao, Y., Ji, H., Wang, Y., Zhi, L., & Li, X. 2017. RUG3 and ATM synergistically regulate the alternative splicing of mitochondrial nad2 and the DNA damage response in *Arabidopsis thaliana*. *Sci Rep*, 7.

Umemoto, T., Aoki, N., Lin, H., Nakamura, Y., Inouchi, N., Sato, Y., & Maruyama, S. 2004. Natural variation in rice starch synthase IIa affects enzyme and starch properties. *Funct Plant Biol*, 31(7), 671-684.

Umemoto, T., Yano, M., Satoh, H., Shomura, A., & Nakamura, Y. 2002. Mapping of a gene responsible for the difference in amylopectin structure between japonica -type and indica -type rice varieties. *Theor Appl Genet*, 104(1), 1.

Wang, B. B., & Brendel, V. 2006. Genomewide comparative analysis of alternative splicing in plants. *Proc Natl Acad Sci*, 103(18), 7175-7180.

- Wattebled, F., Planchot, V., Dong, Y., Szydlowski, N., Pontoire, B., & Devin, A., et al. 2008. Further evidence for the mandatory nature of polysaccharide debranching for the aggregation of semicrystalline starch and for overlapping functions of debranching enzymes in arabidopsis leaves. *Plant Physiol*, 148(3), 1309-1323.
- Xia, T., Zhang, L., Xu, J., Wang, L., Liu, B., Hao, M. & Liu, D. 2017. The alternative splicing of EAM8 contributes to early flowering and short-season adaptation in a landrace barley from the Qinghai-Tibetan Plateau. *Theor Appl Genet*, 130(4), 757-766.
- Yamamori, M., & Endo, T. R. 1996. Variation of starch granule proteins and chromosome mapping of their coding genes in common wheat. *Theor Appl Genet*, 93(1), 275-281.
- Yamamori, M., Fujita, S., Hayakawa, K., Matsuki, J., & Yasui, T. 2000. Genetic elimination of a starch granule protein, sgp-1, of wheat generates an altered starch with apparent high amylose. *Theor Appl Genet*, 101(1), 21-29.
- Yamamori, M., Kato, M., Yui, M., & Kawasaki, M. 2006. Resistant starch and starch pasting properties of a starch synthase iia-deficient wheat with apparent high amylose. *Aust J Agr Res*, 57(5), 531-535.
- Yang, C. Z., Shu, X. L., Zhang, L. L., Wang, X. Y., Zhao, H., Ma, C. X., & Wu, D. X. 2006. Starch properties of mutant rice high in resistant starch. *J Agr Food Chem*, 54(2), 523-528.
- Yang, Y., Chen, X. M., He, Z. H., Röder, M., & Xia, L. Q. 2009. Distribution of vp-1 alleles in chinese white-grained landraces, historical and current wheat cultivars. *Cereal Res Commun*, 37(2), 169-177.
- Zhang, L. L., Chen, H., Luo, M., Zhang, X. W., Deng, M., Ma, J & Pu, Z. E. 2017. Transposon insertion resulted in the silencing of Wx-B1n in Chinese wheat landraces. *Theor Appl Genet*, 130(6), 1321-1330.
- Zhang, X. K., Xiao, Y. G., Zhang, Y., Xia, X. C., Dubcovsky, J., & He, Z. H. 2008. Allelic variation at the vernalization genes *vrn-a1*, *vrn-b1*, *vrn-d1*, and *vrn-b3* in chinese wheat cultivars and their association with growth habit. *Crop Sci*, 42(2), 1690-1694.
- Zhang, X., Myers, A. M., & James, M. G. 2005. Mutations affecting starch synthase iii in arabidopsis alter leaf starch structure and increase the rate of starch synthesis. *Plant Physiol*, 138(2), 663-674.
- Zobel, H. F. 2010. Starch crystal transformations and their industrial importance. *Starch - Starke*, 40(1), 1-7.

Figure legends

Fig. 1. The primer-walking strategy used for isolation of the *TaSSIIa-D* allele. P1, P2, P3, P4, and P5 are PCR products of primer pairs SSIIa-D-F1/SSIIa-D-R1, SSIIa-D-F2/SSIIa-D-R2, SSIIa-D-F3/SSIIa-D-R3, SSIIa-D-F4/SSIIa-D-R4, and SSIIa-D-F5/SSIIa-D-R5, respectively.

Fig. 2. SDS-PAGE results of SGAPs in hexaploid wheat. 1. Xiaoqingmang, 2. Pushanbaimai, 3. P119, 4. Laoqimai, 5. Mazhamai, 6. Xianmai, 7. Chaoanxiaomai, 8. Tumangmai, 9. Chinese Spring. The arrows indicated the area where the absent SGP1-D band might be.

Fig. 3. Amplified results for the junction region between the 7th intron and 8th exon of *TaSSIIa-D* in Pushanbaimai and Chinese Spring. Lowercase letters represent the intron and uppercase letters represent the exon. Arrows shows the splice site.

Fig. 4. Amplified results for the junction region between the 7th intron and 8th exon of *TaSSIIa-D* cDNA in Pushanbaimai. Lowercase letters represent the intron and uppercase letters represent the exon.

Fig. 5. Alignment of sequences at the splice region in Pushanbaimai and Chinese Spring.

Supplementary Table 1. The materials used in this study.

Supplementary Fig. 1. SDS-PAGE results of SGP-1 in diploid and tetraploid wheat (1. Xiaoqingmang, 2. Pushanbaimai, 3. P119, 4. Y46, 5. AS2404, 6. AS313, 7. N7AT7B, 8. N7DTB, 9. CS, and M. Marker).

Fig. 1. The primer-walking strategy used for isolation of the wSSIIa-D allele. P1, P2, P3, P4, and P5 are PCR products of primer pairs SSIIa-D-F1/SSIIa-D-R1, SSIIa-D-F2/SSIIa-D-R2, SSIIa-D-F3/SSIIa-D-R3, SSIIa-D-F4/SSIIa-D-R4, and SSIIa-D-F5/ SSIIa-D-R5, respectively.

397x91mm (96 x 96 DPI)

Draft

Fig. 2. SDS-PAGE results of SGAPs in hexaploid wheat. 1. Xiaoqingmang, 2. Pushanbaimai, 3. P119, 4. Laoqimai, 5. Mazhamai, 6. Xianmai, 7. Chaoanxiaomai, 8. Tumangmai, 9. Chinese Spring. The arrows indicated the area where the absent SGP1-D band might be.

238x109mm (150 x 150 DPI)

Fig. 3. Amplified results for the junction region between the 7th intron and 8th exon of wSSIIa-D in Pushanbamai and Chinese Spring. Lowercase letters represent the intron and uppercase letters represent the exon. Arrows shows the splice site.

397x130mm (96 x 96 DPI)

Draft

Fig. 4. Amplified results for the junction region between the 7th intron and 8th exon of wSSIIa-D cDNA in Pushanbamai. Lowercase letters represent the intron and uppercase letters represent the exon.

292x183mm (150 x 150 DPI)

254x190mm (150 x 150 DPI)

Table 1. Primers used to isolate the *SSIIa-D* gene

Name	Sequence (5'-3')	bp	Tm	Amplified target
SSIIa-D-F1	GACCAACCCGCGAACCGTA	1513	55	P1
SSIIa-D-R1	ATTACGGAGGGAGTATGAC			
SSIIa-D-F2	TCTTATAGTAGTTTACAGACGGAG	2210	55	P2
SSIIa-D-R2	CACAGACACACATACACAAGACA			
SSIIa-D-F3	TGTATGTGTGTCTGTGTGGGGAG	2208	59	P3
SSIIa-D-R3	AGCCCAACGGAGTACCGAAGAAT			
SSIIa-D-F4	TCGCCCCAGCCATAGATGCCTTGTT	1652	58	P4
SSIIa-D-R4	AATCAGCGTCTCAACCACA			
SSIIa-D-F5	CGTCACTTGTTCCCTTATTATCT	1574	56.5	P5
SSIIa-D-R5	GTCGATCCCCGCGTCCTTCTT			
SSIIa-D-F6	GACCCCGTTTCGTTTACCCCA	2400	55	cDNA
SSIIa-D-R6	GGCCGGATTGGAATCAGCG			

Draft

Table 2. Relative amylase and total starch contents of mutant SSIIa-D and control

Line	Amylose	Total Starch
	%	—g 100 g ⁻¹ —
Greenhouse		
Pushanbamai	19.91 ± 1.20 ^b	55.35 ± 0.71 ^b
Chinese Spring	23.28 ± 0.41 ^a	55.13 ± 1.16 ^b
Shumai969	21.53 ± 0.79 ^a	55.52 ± 0.71 ^b
Chuannong16	22.88 ± 0.46 ^a	60.18 ± 0.66 ^a
Field		
Pushanbamai	31.72 ± 1.76 ^a	61.22 ± 2.24 ^a
Chinese Spring	25.24 ± 0.89 ^{cd}	61.28 ± 1.51 ^a
Hongmangmai	28.08 ± 1.07 ^b	62.04 ± 3.54 ^a
Dabaimai	23.34 ± 1.96 ^d	64.05 ± 2.78 ^a
Shumai969	24.27 ± 1.97 ^d	60.60 ± 1.37 ^a
Chuanmai104	27.69 ± 0.75 ^{bc}	65.14 ± 2.91 ^a
Liangmai4	25.24 ± 0.83 ^{cd}	62.22 ± 3.15 ^a

The different superscript letter represent P<0.05.