

## Short communication

Cloning of the  $\omega$ -secalin gene family in a wheat 1BL/1RS translocation line using BAC clone sequencing

Meng Jun Li, Ya Qing Li, Nan Zhang, Zhan Liang Shi \*

Shijiazhuang Academy of Agriculture and Forestry Sciences, Shijiazhuang, Hebei 050041, China

## ARTICLE INFO

## Article history:

Received 10 August 2015

Accepted 3 February 2016

Available online 19 February 2016

## Keywords:

Bacterial artificial chromosome (BAC)

Wheat

 $\omega$ -Secalin

1BL/1RS translocation line

## ABSTRACT

**Background:** Wheat 1BL/1RS translocation lines are planted around the world for their disease resistance and high yield. Most of them are poor in bread making, which is partially caused by  $\omega$ -secalins that are encoded by the  $\omega$ -secalin gene family, which is located on the short arm of rye chromosome 1R (1RS). However, information on the structure and evolution of the  $\omega$ -secalin gene family is still limited.

**Results:** We first generated a physical map of the  $\omega$ -secalin gene family covering 195 kb of the Sec-1 locus based on sequencing three bacterial artificial chromosome (BAC) clones of the 1BL/1RS translocation wheat cultivar Shimai 15. A BAC contig was constructed spanning 168 kb of the Sec-1 locus on 1RS. Twelve  $\omega$ -secalin genes were arranged in a head-to-tail fashion, separated by 8.2–21.6 kb spacers on the contig, whereas six other  $\omega$ -secalin genes were arranged head-to-tail, separated by 8.2–8.4 kb of spacers on clone BAC125. The 18  $\omega$ -secalin genes can be classified into six types among which eight  $\omega$ -secalin genes were expressed during seed development. The  $\omega$ -secalin genes with the 1074-bp open reading frame (ORF) represented the main population. Except for two pseudogenes, the N-terminal of the  $\omega$ -secalin gene was conserved, whereas variations in the C-terminal led to a change in ORF length. The spacers can be sorted into two classes. Class-1 spacers contained conserved and non-conservative sequences.

**Conclusion:** The  $\omega$ -secalin gene family consisted of at least 18 members in the 1BL/1RS translocation line cv. Shimai 15. Eight  $\omega$ -secalin genes were expressed during seed development. Eighteen members may originate from a progenitor with a 1,074-bp ORF. The spacers differed in length and sequence conservation.

© 2016 Pontificia Universidad Católica de Valparaíso. Production and hosting by Elsevier B.V. All rights reserved.

## 1. Introduction

1BL/1RS translocation lines are alien introgressions widely used in wheat breeding programs throughout the world. They were introduced in China during early 1970s and have accounted for 38% of the country's wheat cultivars. 1RS carries several important disease resistance genes that confer resistance to leaf rust (*Lr26*), stem rust (*Sr31*), stripe rust (*Yr9*), and powdery mildew (*Pm8*) [1]. Unfortunately, the benefits associated with 1RS are also accompanied by quality defects, in particular, sticky dough and a reduction in dough strength [2,3]. The poor quality of 1BL/1RS translocation lines is partially caused by  $\omega$ -secalin proteins, which are encoded by the Sec-1 locus on 1RS [4,5].

The  $\omega$ -secalin gene family contains approximately 15 members and is arranged head-to-tail at the Sec-1 locus [5]. Fiber FISH analysis

showed that 15  $\omega$ -secalin genes are arranged in a head-to-tail fashion and separated by 8 kb spacers, rendering a total length of 145 kb [6]. To date, some  $\omega$ -secalin genes, three from rye cv. Gazelle (GenBank No. X60294, X60295, and AF000227) [5,7], nine from DRA-1 containing a small interstitial segment from rye cv. Imperial (without accession numbers) [8], five from wheat 1BL/1RS translocation line cv. Lankao 906 (without accession numbers) [9], sixteen from Rye cv. Rogo [10], sixteen from hexaploid triticale cv. AC Ultima [10], fifteen from octoploid triticale line H93-5305 [10], fifteen from 1BL/1RS translocation line cv. Bobwhite [10], have been reported. However, information on the structure and evolution of the  $\omega$ -secalin gene family is currently limited and imperfect because of the research methods employed above.

Winter wheat cv. Shimai 15 is a high-yield 1BL/1RS translocation line deriving from Petkus rye. However, its bread making quality is poor. The genomic bacterial artificial chromosome (BAC) library of wheat cv. Shimai 15 was constructed and maintained in our laboratory. This made it possible to examine the structure and evolution of the  $\omega$ -secalin family at the genome level. We isolated 18 members of the  $\omega$ -secalin family and characterized any detected

\* Corresponding author.

E-mail addresses: lmj199612@aliyun.com (M.J. Li), liyaqing2002@126.com (Y.Q. Li), zhangnan-89@163.com (N. Zhang), shizhanl@tom.com (Z.L. Shi).

Peer review under responsibility of Pontificia Universidad Católica de Valparaíso.

structural variations and evolutionary features of the  $\omega$ -secalins and spacers. The information is essential and beneficial to manipulate the expression of  $\omega$ -secalin genes.

## 2. Materials and methods

### 2.1. Construction of a BAC library and four subclone libraries of wheat cv. Shimai 15

The BAC library of wheat cv. Shimai 15 was constructed using the *HindIII* site of pCC1BAC vector (Epicentre, Madison, WI, USA). The BAC library contained more than 1,000,000 clones and preserved in 1,020 pools. The average insert size of the BAC clones was about 85 kb.

Four subclone libraries were constructed using the *BamHI* site of the pCC1BAC vectors. The plasmids of four BAC clones, BAC112, BAC125, BAC356, and BAC736, were partially digested with *Sau3AI*. The 10-kb fragments of each BAC clone were extracted separately using a gel DNA isolation kit (GenStar, Beijing, China), and then the four fragments were ligated into the *BamHI* site of pCC1BAC vectors. Each subclone library was comprised of 384 clones with more than 20 $\times$  coverage of the corresponding BAC plasmid.

### 2.2. Sequencing of three BAC clones

Screening of the *HindIII* BAC library with primers of the  $\omega$ -secalin gene (RS1) (Table 1) was performed as described elsewhere [9]. Plasmid DNA was isolated using the standard alkaline-lysis protocol and used as templates. For the positive BAC clones, BAC112, BAC125, and BAC356, were selected for sequencing by Chinese National Human Genome Center (Shanghai, China). The DNA library preparation of BAC clones followed the Pacific Biosciences Protocol: “10 kb template preparation and sequencing”. Three libraries were sequenced using Pacific BioscienceSMRTT with DNA Sequencing Kit 2.0 (Pacific Biosciences, USA). The sequences produced were assembled by HGAP2, which is the module of SMRT Analysis Software.

For assembling and validation of the BACs, 10-kb subclone libraries were also constructed. For each library, primer RS2 designed based on the  $\omega$ -secalin gene (GenBank No. AF000227) (Table 1) was used to screen subclones containing  $\omega$ -secalin genes. The members of the  $\omega$ -secalin gene family in the positive subclones were sorted by sequencing using primer RS2F (Table 1), and then the sequences of the inserts in the positive subclones were gained by sequencing with BigDye<sup>R</sup> Terminator v3.1 Cycle Sequencing Kit (ABI) on an ABI 3730XL DNA Analyzer. The sequences of three BAC clones were determined by combining the sequencing data of BAC clones and subclones. Sequence assembly was performed using the Lasergene SeqMan II Module (DNASTar) (<http://www.DNASTar.com>).

### 2.3. Physical mapping of the 168-kb contig spanning *Sec-1* locus on 1RS

Plasmids of four BAC clones were purified using MIDI-PREP columns (Qiagen, Hamburg, Germany). Sequences generated through BAC end-sequencing were employed to develop BAC-end markers. Three

BAC clones, BAC112, BAC356, and BAC736, were assembled into a contig that was identified by polymerase chain reaction (PCR) method using the BAC end markers. Fifty clones of the 10-kb subclone libraries of BAC736 were randomly picked for sequencing. Based on the sequence data of BAC112, BAC356, and BAC736, physical mapping of the 168-kb contig spanning *Sec-1* locus on 1RS was constructed.

### 2.4. $\omega$ -Secalin gene expression in seed development

Winter wheat cv. Shimai 15 was grown in a farmland and the developing seeds from 7 to 28 d after pollination (DAP) were collected. Total RNA was isolated from developing seeds using a total RNA extraction kit (Bioteke, Beijing, China). Messenger RNA (mRNA) was isolated using the polyAtract mRNA isolation system III (Promega). One microgram of mRNA was reverse transcribed using PrimerScript<sup>TM</sup> 1st strand cDNA synthesis kit (TaKaRa, Dalian, China). One microliter of a 10-fold dilution of cDNA mixture was used as template for PCR amplification using PrimeSTAR HS DNA Polymerase (TaKaRa, Dalian, China) with three specific primers (RS1, RS3, and RS4). The 5' UTR and 3' UTR of the  $\omega$ -secalin genes were identified using the 5'/3' rapid amplification of cDNA ends (RACE) kit with primers 5' UTRP and 3' UTRP (Roche). The PCR products were cloned into the pMD18-T vectors (TaKaRa, Dalian, China). Sixty cDNA clones were randomly picked and sequenced with a BigDye<sup>R</sup> Terminator v3.1 cycle sequencing kit (ABI) on an ABI 3730XL DNA Analyzer. Primers used in  $\omega$ -secalin gene cloning are provided in Table 1.

### 2.5. cDNA, genomic DNA, and protein sequence analysis of the $\omega$ -secalin genes

Sequence analysis of the  $\omega$ -secalin genes and spacers was performed using Lasergene SeqMan II Module (DNASTar) (<http://www.DNASTar.com>), which included the coding sequence, amino acid sequence, molecular weight, and isoelectric point (pI). Multiple sequence alignments were analyzed using ClustalW 1.83 software (<http://www.ch.embnet.org/software/ClustalW.html>) and shaded using BoxShade server ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)).

## 3. Results

### 3.1. Isolating BAC clones containing the $\omega$ -secalin genes

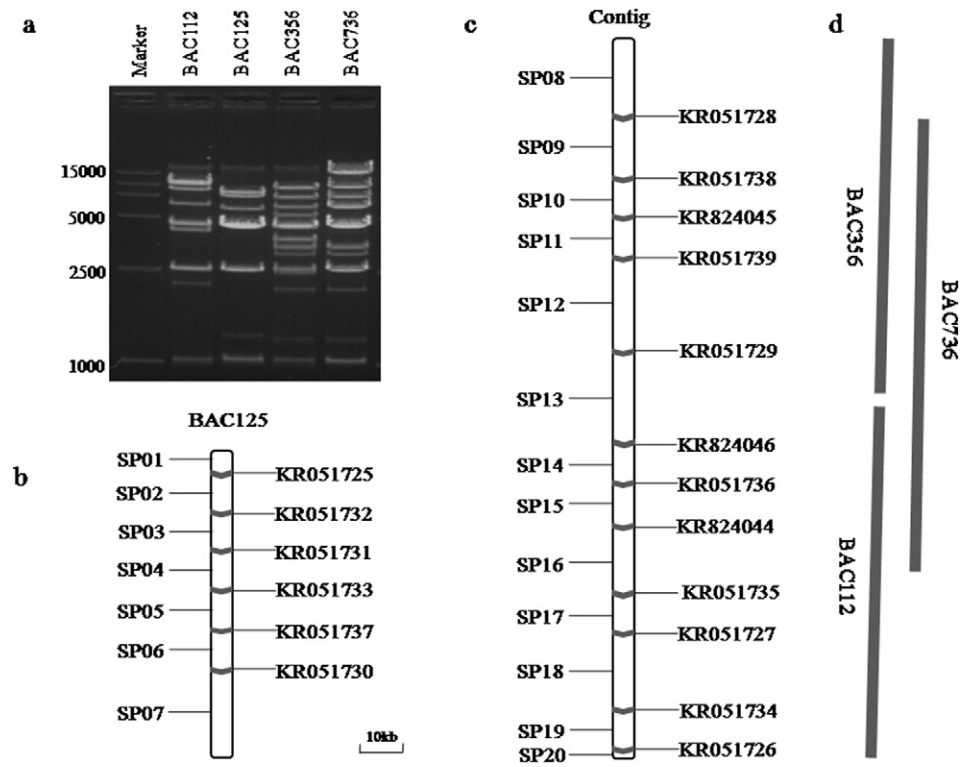
Based on published primers of the  $\omega$ -secalin gene (RS1; Table 1), six positive BAC pools, pool 112, pool 125, pool 356, pool 457, pool 484, and pool 736, were identified by three-round PCR screening using 1,020 BAC-pool plasmids of cv. Shimai 15 BAC library as templates. Then, only one type of positive BAC clone was isolated from each positive BAC pool as confirmed by restriction enzyme digestion (*BamHI*, *EcoRI*, and *SacI*) and BAC-end sequencing. Two pairs, BAC112/484 and BAC125/457, were overlapping BAC clones that were separately confirmed by restriction enzyme digestion and BAC end-sequencing. Restriction enzyme digestion fragment analysis (*SacI*) showed that four BAC clones, BAC112, BAC125, BAC356, and BAC736, displayed different restriction maps (Fig. 1a). PCR cloning and sequencing illustrated that the four BAC clones contained more than three members of the  $\omega$ -secalin gene family by using specific primers [9].

### 3.2. Sequencing of the three BAC clones

Pulsed-field gel electrophoresis (PFGE) showed that the inserts of three BAC clones, BAC112, BAC125, and BAC356, were about 80 kb long. Then, three BAC clones were sequenced using Pacific BioscienceSMRTT (Pacific Biosciences, USA). The sequencing data of BAC112, BAC125, and BAC356 were assembled into 72-kb, 52-kb,

**Table 1**  
Specific primers used in the study of the  $\omega$ -secalin gene family.

Primer name	Upstream primer (5' → 3')	Downstream primer (5' → 3')
RS1	ACCTTCCTCATCTTTGTCCT	CCGATGCCTATACCACTACT
RS2	CTTCCTCATCTTTGTCCTCGC	TGTGGTTGTTGGGAATGG
RS2F	CTTCCTCATCTTTGTCCTCGC	
RS3	ATCATCCAACCTCACGCACC	CGTTACATTGAACACTCCATTGAC
RS4	GAAACATCAATTCCAAGCAAGC	TGACTAAACAACGATGATCCACC
5'UTRP-1		GAGGACAAAGATGAGGAAGGTC
5'UTRP-2		GCGTGAAGTTTGGATGATGG
3'UTRP-1	CCCCAGCAACACAACAAC	
3'UTRP-1	CACAACAACCATTTTCCACG	



**Fig. 1.** Physical map of the  $\omega$ -secalin gene family of wheat cv. Shimai 15. a: Restriction enzyme digestion analysis of BAC clones; b: BAC125; c: Physical mapping of BAC112, BAC356, and BAC736; d: Schematic representation of the three BAC clones contigs.

and 82-kb sequences, respectively. Because of high similarities of the  $\omega$ -secalin genes and repeat sequences in the BAC clones, it was very difficult to accurately assemble the BAC sequences. To correct the assembled sequences of the three BAC clones, three 10-kb subclone libraries were screened by using PCR primers (RS2; Table 1). The types of  $\omega$ -secalin genes in positive subclones were identified using sequence primer RS2F (Table 1). Seven, six, and five  $\omega$ -secalin genes were identified in BAC112, BAC125, and BAC356, respectively (Table 2). The inserts of these positive subclones were sequenced

by primer walking. Then, the physical maps of three BAC clones were generated by performing sequence data analysis (Fig. 1b and c).

### 3.3. Physical mapping of the 168-kb contig spanning the *Sec-1* locus on 1RS

BAC end sequence analysis showed that the three BAC clones, BAC112, BAC356, and BAC736, assembled into a contig (Fig. 1d). The end sequences of 50 subclones in the BAC736 subclone library also confirmed the contiguous sequences of the three BAC clones. The

**Table 2**  
Features of the 18 members of the  $\omega$ -secalin gene family of wheat cv. Shimai 15.

Gene	Location (bp)	Accession	ORF (bp)	Deduced protein MWa (kDa)		Calculated pI <sup>b</sup>	
				Precursor	Mature	Precursor	Mature
Se01	BAC125/5090...6196	KR051725	1,107	42.2	40.1	8.00	7.20
Se02	BAC125/14,345...15,418	KR051732	1,074	41.6	39.5	8.57	7.16
Se03	BAC125/23,596...24,669	KR051731	1,074	41.5	39.4	8.53	7.16
Se04	BAC125/32,851...33,918	KR051733	1,068	41.5	39.4	9.31	8.63
Se05	BAC125/42,109...42,942	KR051737	834	32.3	30.2	10.04	9.88
Se06	BAC125/51,369...52,442	KR051730	1,074	41.4	39.3	10.04	9.93
Se07	Contig/18,592...19,665	KR051728	1,074	41.5	39.5	9.66	9.38
Se08	Contig/32,250...33,035	KR051738	786	30.5	29.5	9.76	10.06
Se09	Contig/41,492...42,564	KR824045 <sup>c</sup>	1,073	/	/	/	/
Se10	Contig/50,739...51,203	KR051739	465	17.9	15.8	9.29	8.97
Se11	Contig/72,806...73,879	KR051729	1,074	41.6	39.5	5.87	5.27
Se12	Contig/94,665...95,738	KR824046	1,074	41.5	39.4	9.74	9.57
Se13	Contig/103,929...104,831	KR051736	903	35.0	32.9	9.76	9.49
Se14	Contig/113,808...114,955	KR824044 <sup>c</sup>	1,148	/	/	/	/
Se15	Contig/129,400...130,401	KR051735	1,002	38.7	36.6	9.38	8.66
Se16	Contig/138,554...139,627	KR051727	1,074	41.6	39.4	8.61	7.16
Se17	Contig/156,189...157,238	KR051734	1,050	40.6	38.5	9.38	8.66
Se18	Contig/165,597...166,670	KR051726	1,074	41.6	39.5	7.16	5.86

<sup>a</sup> MW, molecular weight.

<sup>b</sup> pI, isoelectric point.

<sup>c</sup> Predicted pseudogene.



168-kb contig consisted of 12 members of the  $\omega$ -secalin family, including 2 pseudogenes. The  $\omega$ -secalin genes were arranged in a head-to-tail fashion with intergenic spacers ranging from 8.2 kb to 21.6 kb in the BAC contig (Fig. 1c). The  $\omega$ -secalin genes on BAC125 were also arranged head-to-tail with 8.2–8.4 kb intergenic spacers (Fig. 1b). The 18  $\omega$ -secalin genes covered a total of 195 kb of the *Sec-1* locus on 1RS.

### 3.4. Sequence analysis of 18 members of the $\omega$ -secalin gene family

Two pseudogenes and the open reading frame (ORF) of 16  $\omega$ -secalin genes were confirmed by PCR cloning and sequencing using subclone plasmids as template. The deduced ORF length of the  $\omega$ -secalin genes

ranged from 1,107 bp to 465 bp (Table 2). The ORF length changes were mainly the result of sequence variations within the ORF of the  $\omega$ -secalin genes (Fig. 2). The sequence variations can be classified into the following six types:

- (1)  $\omega$ -Secalin genes with a 1,107-bp ORF. Se01 with a frame shift was caused by 1-bp insertion at nucleotide 1,019, which resulted in 2 cysteine residues. The cysteine residues were not normally present in the  $\omega$ -secalin proteins [8]. Se01 was consistent with clone 123 from DRA-1 [8].
- (2)  $\omega$ -Secalin genes with a 1,074-bp ORF. Eight  $\omega$ -secalin genes with a 1,074-bp ORF were individually isolated from BAC112, BAC125,

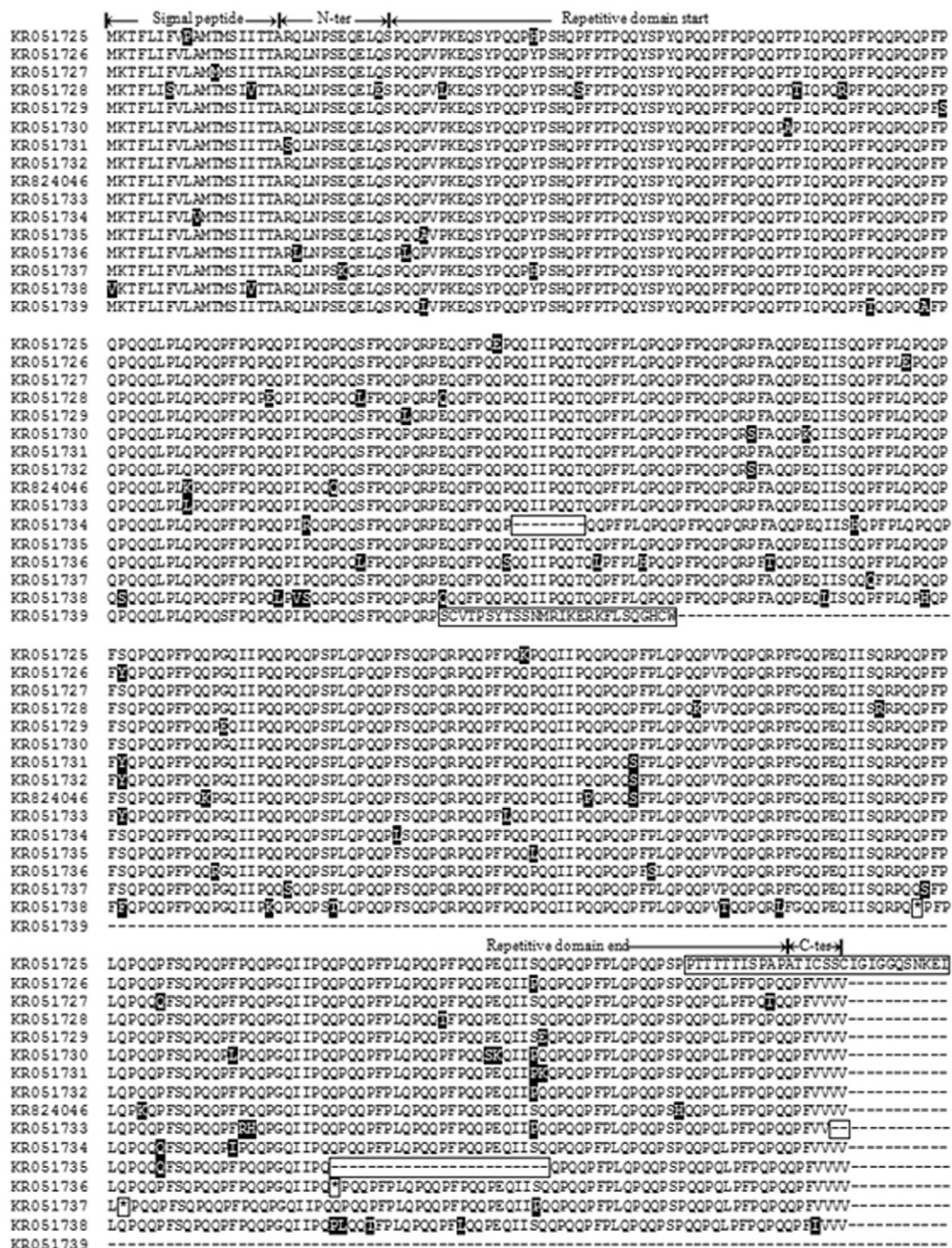


Fig. 2. Deduced protein sequence comparison of the  $\omega$ -secalin genes. Signal peptide, N-terminal, and C-terminal regions are labeled above the aligned sequences. An asterisk (\*) denotes a stop codon, whereas different amino acids are shaded.

and BAC356. The eight genes ranged from 97.6% to 99.6% nucleotide sequence identity and 94.4% to 99.2% deduced protein sequence identity. A total of 51 substitutions were detected within the 1,074-bp ORFs of the  $\omega$ -secalin genes, which resulted in a 38-amino-acid change. The 1,074-bp ORF of the  $\omega$ -secalin genes represented the main population in the  $\omega$ -secalin gene family of cv. Shimai 15.

- (3)  $\omega$ -Secalin genes with 1,068-bp, 1,050-bp, and 1,002-bp ORFs. The three  $\omega$ -secalin genes, Se04, Se17, and Se15, harbored 6-bp, 24-bp, and 72-bp deletions at nucleotides 1,066, 412, and 921, respectively. Se15 and Se17 were consistent with DNA1004-1 and DNA1052-1 from wheat cv. Lankao906 [9]. The 6-bp deletion of Se04 occurred in the conserved C-terminal region of the  $\omega$ -secalin gene.
- (4)  $\omega$ -Secalin genes with 903-bp, 834-bp, and 786-bp ORFs. The point mutations at nucleotides 901, 832, and 814 led to premature termination of the three  $\omega$ -secalin genes, Se13, Se05, and Se08. The point mutation (A  $\rightarrow$  G) of Se08 at nucleotide 1 led to a 10-amino acid deletion in the signal peptide of the  $\omega$ -secalin gene. Se08 was consistent with clone 8 from DRA-1 [8].
- (5)  $\omega$ -Secalin gene with a 465-bp ORF. Se10 was a special  $\omega$ -secalin gene. The C-terminal of Se10 from amino acid 129–154 was distinct from those of the other  $\omega$ -secalin genes. Two cysteine residues were detected within the region.

- (6) Two pseudogenes. The pseudogene Se14 resulted from a point mutation at nucleotide 86 (TTG to TAG), two deletions of 1-bp and 24-bp in size at nucleotides 709 and 909, and a 99-bp insertion at nucleotide 1,049 (Fig. 3). Se14 was similar to clone 182 from DRA-1 [8] and consistent with DNA1150 from wheat cv. Lankao906 [9]. Pseudogene Se09, which harbored a 1-bp deletion at nucleotide 37, resulted in a frame shift that was consistent with DNA1075-1 from wheat cv. Lankao906 [9].

### 3.5. Sequence analysis of the spacers

The 20 spacers were named as SP1–SP20 in BAC125 and contig, ranging in length from 1,579 bp to 21,602 bp including 16 intergenic spacers (Fig. 1b, c). The 16 intergenic spacers can be sorted into two classes, class-1 and class-2. Class-1 spacers consisted of conserved and non-conservative sequences while class-2 spacers only contained conserved sequences (Fig. 4). A sequence comparison showed that the conserved sequences of the spacers displayed more than 97% nucleotide sequence identity. The 5 intergenic spacers in BAC125 all belonged to class-2. Contrary to BAC125, 7 of 11 intergenic spacers in the contig were class-1 spacers. The 7 class-1 spacers in the contig can be further divided into 3 subclasses. The conserved sequences were replaced by non-conservative sequences in 4 subclass-1 spacers including SP09,

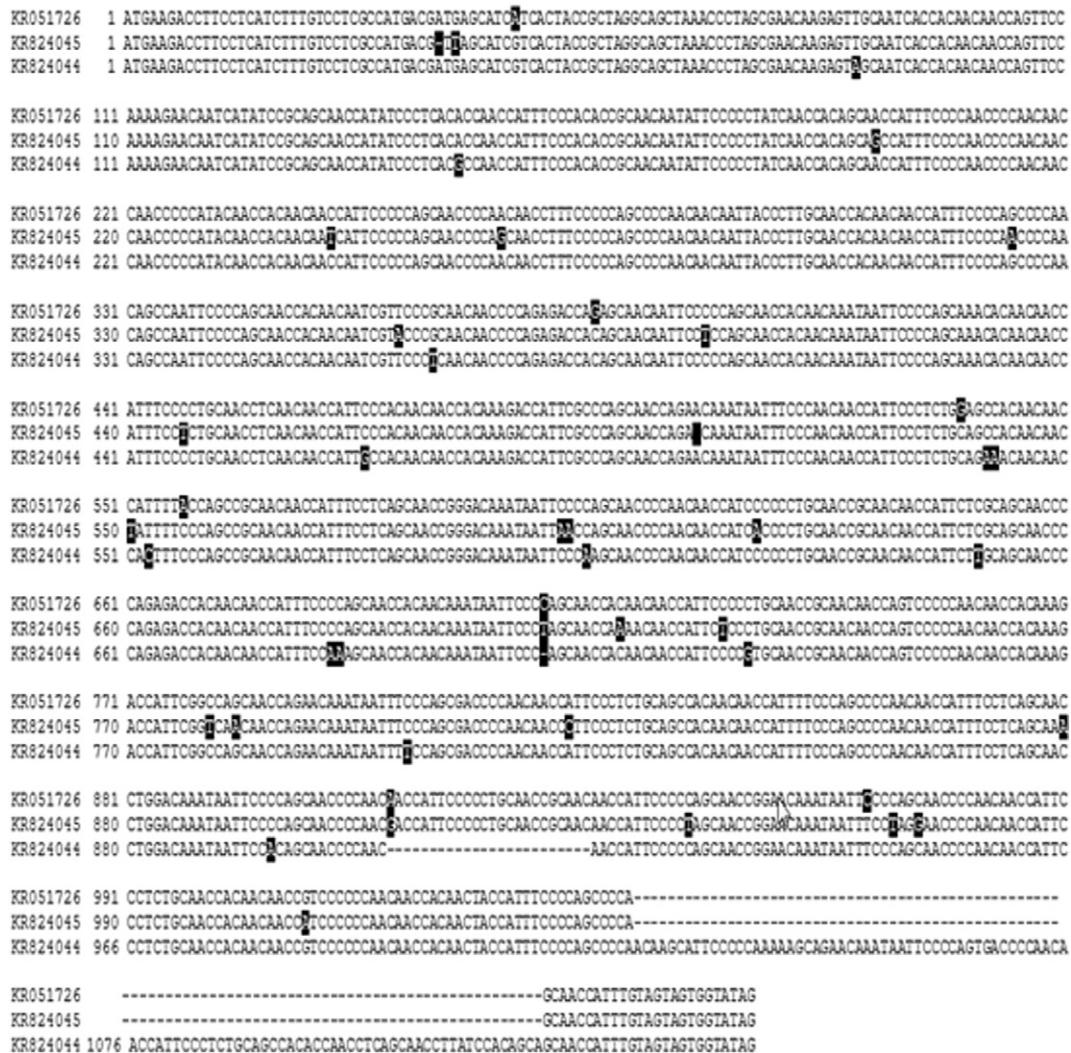


Fig. 3. Nucleotide sequence alignment of two pseudogenes with the  $\omega$ -secalin gene. The different nucleotides are shaded.



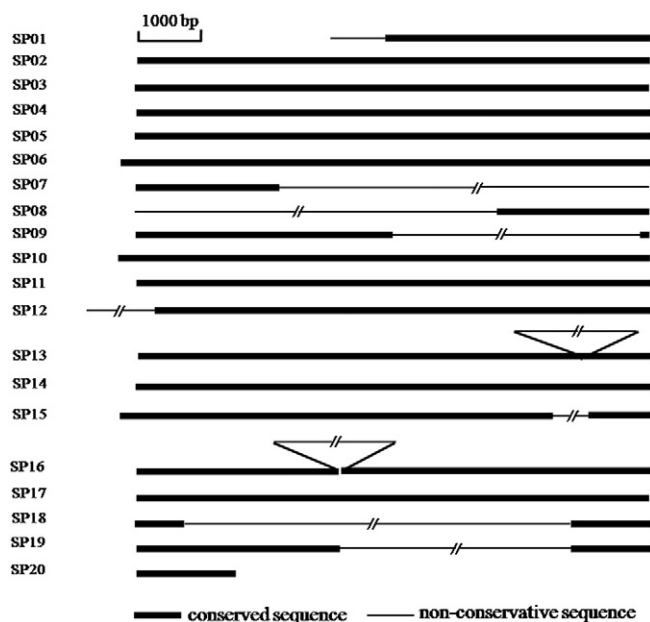


Fig. 4. Analysis on sequence conservation of the spacers.

SP15, SP18 and SP19 and separated by non-conservative sequences in 2 subclass-2 spacers including SP13 and SP16. The non-conservative sequence of subclass-3 spacer, SP12, located at the 5' end of the spacer. There were 4 putative genes, phosphatidylinositol 3-kinase, zinc finger MYM-type protein 1-like, GTP pyrophosphokinase, disease resistance RPP13-like protein 3, identified in the conserved sequences of the spacers by sequence annotation using blastx tool in NCBI database. Meanwhile, 5 genes were found in non-conservative sequences of 7 class-1 spacers in the contig including 3 retrotransposon genes, Ty3-gypsy retrotransposon protein (SP12), gag-polypeptide of copia-type (SP13), and Opie4 pol protein (SP18). Besides, a retrotransposon gene, Ty1-copia retrotransposon protein, existed in non-conservative sequences of SP07.

### 3.6. Genomic organization of the $\omega$ -secalin genes

The members of the  $\omega$ -secalin gene family expressing during seed development were investigated by RT-PCR by using three primers (RS1, RS3, and RS4) (Table 1). Fifty seven full-length cDNAs were gained by sequencing. Forty-eight cDNAs had corresponding genomic DNAs from three BAC clones, including eight members of the  $\omega$ -secalin gene family, Se03, Se06, Se07, Se11, Se12, Se15, Se16 and Se17. No genomic DNAs corresponding to the other nine  $\omega$ -secalin cDNAs were detected. This indicated that some other  $\omega$ -secalin genes may lay outside the three BACs. Comparison of genomic DNAs with their corresponding cDNAs showed no intron in the eight  $\omega$ -secalin genes.

To obtain the full-length cDNAs of the  $\omega$ -secalin genes, 5' RACE and 3'-RACE were performed. The 5' and 3' untranslated regions (UTRs) were aligned with the genomic DNA sequences of the  $\omega$ -secalin genes. The transcription start sites (CCTTCCATCATCC) were all located 71 bp upstream of the translation start codon. The length of the 3' UTR with two polyadenylation signal sequences (AATAAA) ranged from 171 bp to 188 bp. The poly(A) tails were added at 171 bp, 173 bp, 177 bp, and 188 bp downstream of the termination codon (CACAAGGACATTGCTGGT).

### 3.7. Promoter analysis of the $\omega$ -secalin genes

The cis-elements in the promoter sequences of the  $\omega$ -secalin genes were searched by the plantCARE online tool. Many cis-acting

regulatory elements involved in light responsiveness and enhancer elements were predicted (data not shown). The 1,125-bp sequences upstream from the transcription start site in the  $\omega$ -secalin genes were highly conserved among  $\omega$ -secalin genes except for Se08.

A comparison of the promoter sequences of 18  $\omega$ -secalin genes at 300 bp upstream from the transcription start site identified 24 substitutions and 1 insertion over the whole regions (8.3%), except for Se08 promoter sequence (Fig. 5). Se08 promoter sequence was relatively different from those of the other  $\omega$ -secalin genes. The pseudogenes Se09 and Se14, had two substitutions, T  $\rightarrow$  C and G  $\rightarrow$  T, within the conserved GCN 4 motif (TGAGTCA to TGAGCCA, TGAGTCA to TTAGTCA) of the  $-300$  elements. No mutation in the highly conserved promoter motif in the other  $\omega$ -secalin genes was detected, indicating that the two pseudogenes may be not actively transcribed.

## 4. Discussion

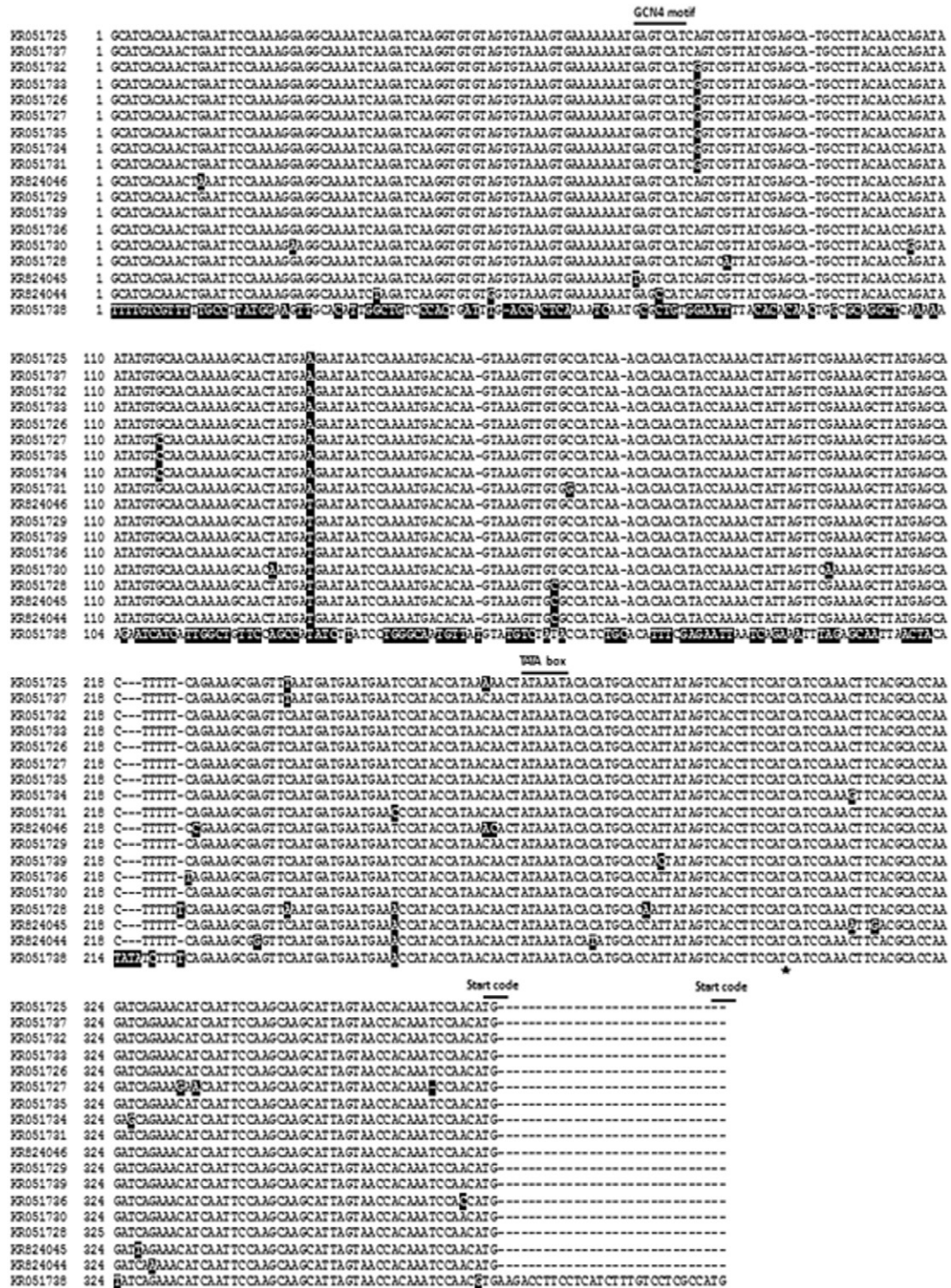
Omega-secalins have a negative impact on dough quality of the 1BL/1RS translocation lines. Less information on the  $\omega$ -secalin gene family is available, which was obtained by lambda clone enzyme cutting and sequencing [5,8], PCR analysis [9,10], or Fiber FISH [6]. However, it is difficult to accurately analyze the structure and evolution of the  $\omega$ -secalin gene family using these methods above.

Southern blotting showed that there were about 40–60  $\omega$ -secalin gene copies at Sec-1 in the wheat/rye addition lines [11] and 10–20 copies in wheat line DRA-1 which contains a small interstitial 1RS segment from rye cv. Imperial on 1DS [5]. Hybridization with DNA probe from pSec1 clone identified a 145 kb fragment which contained approximately 15  $\omega$ -secalin genes and also showed that some gene units lay outside the 145 kb fragment [5]. Fiber FISH analysis of rye support the model that the  $\omega$ -secalin genes are arranged in a head-to-tail fashion separated by 8 kb of spacer sequences [8]. In this study, we assembled BAC contig of 195 kb and identified totally 18  $\omega$ -secalin genes in wheat cv. Shimai 15. We also found another 9  $\omega$ -secalin genes by sequencing PCR product from cDNA mixture of developing seeds (data not shown). These results we achieved were not consistent with those of previous reports [5,6,9]. The first reason is that it's more efficient and accurate to isolate  $\omega$ -secalin genes with the BAC library compared to PCR cloning. The second reason is that the Sec-1 loci analyzed in different studies are from rye cv. Imperial [5,6] and rye cv. Petkus respectively.

For the  $\omega$ -secalin genes identified in wheat cv. Shimai 15, their 5' ends were much more conserved than the 3' end which suggested that the 5' end was under strong selection pressure. Among them, the 8 members with a 1,074-bp ORF represent the main population. The Se10 was very unique since its nucleotide sequence from 385 bp to 465 bp was not conserved to other members. The base insertions, deletions and point mutations in the 1,074-bp ORF resulted in frame shifts, stop codons, and amino acid deletions as shown in other 9  $\omega$ -secalin genes. Sequence comparison indicated that the members of the  $\omega$ -secalin gene family may originate from a 1,074-bp ORF progenitor.

Previous studies suggested that the  $\omega$ -secalin genes were separated by 8 kb spacers [5,6]. Similar results were obtained in the BAC125. However, only 4 spacers were around 8 kb in the contig. The intergenic class-1 spacers all located in the contig. Two pseudogenes, Se09 and Se14, were identified in the contig. The Se08 was also a special gene. These results above suggested that the members of the  $\omega$ -secalin gene family on the BAC125 and contig were produced by different mechanism. There were 4 retrotransposon genes found in non-conservative sequences of the spacers in the BAC125 and contig which indicated that retrotransposon genes involved in the evolution of the  $\omega$ -secalin gene family.

The cDNA sequences corresponding to 8 members were cloned by PCR using mRNA of different developing seeds as template. The number of expressed members showed no difference during seed



**Fig. 5.** Nucleotide sequence alignment of the promoter regions from the  $\omega$ -secalin genes. GCN motif, TATA box, and Start codon are labeled above the aligned sequences. An asterisk (\*) denotes the transcription start site, whereas different amino acids are shaded.

development (data not shown). Six of the eight expressed members belonged to the 1,074-bp ORF  $\omega$ -secalin gene, which may play a similar function. Cloning of two expressed members with 1,050-bp and 1,002-bp ORFs supports the results of Chai et al. [9], which was generated by PCR analysis. Compared to the 1,074-bp ORF member, the two members with a 1,050-bp ORF and 1,002-bp ORF were caused by the deletion of the repetitive domain in the  $\omega$ -secalin gene. The result indicated that the deletion had no effect on the function of the  $\omega$ -secalin gene.

In summary, 18 members of the  $\omega$ -secalin gene family were isolated by sequencing three BAC clones with spacers ranging from 8.2 kb to 21.6 kb. Except for two pseudogenes, the N-terminal of the  $\omega$ -secalin gene was conserved, whereas variations in the C-terminal led to a change in ORF length. Members with the 1,074-bp ORF predominated. Eight members were expressed during seed development. Whether  $\omega$ -secalins have other functions other than its negative impact on dough quality in 1BL/1RS translocation lines have yet to be determined.

## Accession codes

Sequence data can be found in the GenBank under accession numbers KU051382(BAC112), KU051383(BAC125) and KU051384(BAC356).

## Financial support

This research was supported by Natural Science Foundation of Hebei Province (No. C2014106075) and Shijiazhuang Science and Technology Pillar Program (No. 12149402A).

## Conflict of Interest

The authors have no conflicts of interest to report.

## References

- [1] Singh NK, Shepherd KW, McIntosh RA. Linkage mapping of genes for resistance to leaf, stem and stripe rusts and  $\omega$ -secalins on the short arm of rye chromosome 1R. *Theor Appl Genet* 1990;80:609–16. <http://dx.doi.org/10.1007/BF00224219>.
- [2] Martin DJ, Stewart BG. Dough stickiness in rye-derived wheat cultivars. *Euphytica* 1990;77–86. <http://dx.doi.org/10.1007/BF00022895>.
- [3] Graybosch RA, Peterson CJ, Hansen LE, Worrall D, Shelton DR, Lukaszewski A. Comparative flour quality and protein characteristics of 1BL/1RS and 1AL/1RS wheat-rye translocation lines. *J Cereal Sci* 1993;17:95–106. <http://dx.doi.org/10.1006/jcrs.1993.1010>.
- [4] Dhaliwal AS, Mares DJ, Marshall DR. Measurement of dough surface stickiness associated with the 1B/1R chromosome translocation in bread wheats. *J Cereal Sci* 1990;12:165–75. [http://dx.doi.org/10.1016/S0733-5210\(09\)80098-2](http://dx.doi.org/10.1016/S0733-5210(09)80098-2).
- [5] Clarke BC, Mukai Y, Appels R. The Sec-1 locus on the short arm of chromosome 1R of rye (*Secale cereale*). *Chromosoma* 1996;105:269–75. <http://dx.doi.org/10.1007/s004120050184>.
- [6] Yamamoto M, Mukai Y. High-resolution physical mapping of the secalin-1 locus of rye on extended DNA fibers. *Cytogenet Genome Res* 2005;109:79–82. <http://dx.doi.org/10.1159/000082385>.
- [7] Hull GA, Halford NG, Kreis M, Shewry PR. Isolation and characterisation of genes encoding rye prolamins containing a highly repetitive sequence motif. *Plant Mol Biol* 1991;17:1111–5. <http://dx.doi.org/10.1007/BF00037153>.
- [8] Clarke BC, Appels R. Sequence variation at the Sec-1 locus of rye, *Secale cereale* (Poaceae). *Plant Syst Evol* 1999;214:1–14.
- [9] Chai JF, Liu X, Jia JZ. Homoeologous cloning of  $\omega$ -secalin gene family in a wheat 1BL/1RS translocation. *Cell Res* 2005;15:658–64. <http://dx.doi.org/10.1038/sj.cr.7290335>.
- [10] Jiang QT, Wei YM, Andre L, Lu ZX, Pu ZE, Peng YY, et al. Characterization of  $\omega$ -secalin genes from rye, triticale, and a wheat 1BL/1RS translocation line. *J Appl Genet* 2010; 51:403–11. <http://dx.doi.org/10.1007/BF03208870>.
- [11] Hull G, Sabelli PA, Shewry PR. Restriction fragment analysis of the secalin loci of rye. *Biochem Genet* 1992;30:85–97. <http://dx.doi.org/10.1007/BF00554430>.