Journal of Cereal Science 56 (2012) 282–288

Contents lists available at SciVerse ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Structure and evolutionary relationships among paralogous genes within the *Sec2* locus in rye

Yanqian Wang [a,b,c], Bo Zhang [a], Baolong Liu [a], Huaigang Zhang [a,**], Dengcai Liu [a,b,*]

[a] Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810001, PR China
[b] Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, PR China
[c] Graduate University, Chinese Academy of Sciences, Beijing 100049, PR China

ARTICLE INFO

Article history:
Received 12 November 2011
Received in revised form
6 May 2012
Accepted 11 May 2012

Keywords:
Rye
Sec2
Secalin
Seed storage protein

ABSTRACT

The 75K γ-secalins encoded by genes present at the locus *Sec2* on chromosome 2R are unique to rye and contribute about half of all rye storage proteins. However, there is a lack of sequence information for paralogous genes in this locus. For this study, 59 γ-secalin paralogous sequences in the *Sec2* locus were characterized from a cultivated rye and derived lines after crossing with bread wheat. They had similar structures with conserved sequences in their repetitive regions for the signal peptide, N-terminal, C-terminal and the repeat motif. Their high homology indicated that they originated from an ancestor sequence that existed before the speciation of the genus *Secale*. Duplication and divergence might have led to the formation of the paralogous genes at *Sec2*. Besides point mutations, these paralogs showed variations in DNA length due to insertion or deletion events in their repetitive regions. They encoded secalins with deduced molecular weight ranges between 22.2 and 54.5 kDa. These insertion or deletions may be caused by illegitimate recombination and this locus seemed to contribute to increased levels of protein content. However, the incorporation of locus *Sec2* may have a negative effect on flour processing quality since it reduced the SDS-sedimentation value.

1. Introduction

Secalins are an alcohol-soluble (prolamin) group of seed storage proteins found in rye (*Secale* L., 2N = 2X = 14) grain and account for around half the total nitrogen found in the mature grain (Shewry et al., 1983b). Secalins are a complex and highly polymorphic mixture of polypeptides that can be classified into four major groups by electrophoresis mobility on SDS-PAGE gel: high molecular weight (HMW) secalin, 75K γ-secalin, 40K γ-secalin and ω-secalin (Shewry et al., 1994). HMW secalins are encoded by genes present at the locus *Sec3* on chromosome arm 1RL, which consists of two paralogous alleles of duplication origin that encode the x and the y types of high-molecular weight secalin subunits, respectively. The subunits encoded by this locus reduce the dough and gluten strength (Kipp et al., 1996). The ω-secalins and 40K γ-secalins are encoded by locus *Sec1* on 1RS. The incorporation of this locus into wheat had a negative effect on wheat processing quality (Miller, 1984; Shewry et al., 1986). *Sec1* is a complex locus with families of closely linked genes (Chai et al., 2005; Jiang et al., 2010). A section encoding ω-secalins in this locus has been sequenced and shown to contain about 15 ω-secalin genes arranged in a head to tail fashion (Clarke et al., 1996). The 75K γ-secalins are encoded by locus *Sec2* found on 2R in cultivated rye (*Secale cereale*) or 6R in wild perennial rye (*Secale montanum*) (Lawrence and Shepherd, 1981; Shewry et al., 1984a,b; Hull et al., 1992). The *Sec2* locus may be derived from the successive translocation of genes from 1R to 6R in wild perennial rye and then to 2R in cultivated rye (Shewry et al., 1985, 1986). The 75K γ-secalins are present in endosperm with aggregates stabilized by intra/inter-molecular disulphide bonds (Field et al., 1983). An analysis of the 2RS.2BL translocation lines indicated that *Sec2* increased polymeric glutenin, which could potentially provide superior dough properties and increase protein levels in the endosperm (Gupta et al., 1989).

The *Sec2* locus encodes a unique family of 75K γ-secalins which does not have analogs in other cereals (Gellrich et al., 2005). This locus is thought to contain 3–6 genes and encodes 3–4 polypeptide

Abbreviations: GISH, Genomic in situ hybridization; HMW, High-molecular-weight; IPTG, Isopropyl β-Δ-thiogalactopyranoside; SDS-PAGE, Sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

* Corresponding author. Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, PR China. Tel.: +86 28 82650312; fax: +86 28 82650350.
** Corresponding author. Tel.: +86 971 6143630; fax: +86 971 6143282.
E-mail addresses: hgzhang@nwipb.cas.cn (H. Zhang), dcliu7@yahoo.com (D. Liu).

0733-5210/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.jcs.2012.05.012

bands (Hull et al., 1992; Murray et al., 2001). The 75K γ-secalins are the most abundant group of secalins, accounting for over 50% of the total protein in rye (Shewry et al., 1983a,b; Shewry and Field, 1982). To date, however, there is little sequence information for paralogous genes within the *Sec2* locus available, although a few alleles from different rye lines at this locus have been isolated by (Murray et al. (2001) and Chen et al. (2008).

In this study, 75K γ-secalin paralogous sequences in the *Sec2* locus were characterized from the cultivated rye var. Qinling, and lines derived from crosses with bread wheat. Based on these sequences, the structure and evolutionary relationship among paralogous genes in this locus were discussed.

2. Materials and methods

2.1. Plant materials

The cereals used in this study were: two common wheats, Chinese Spring and D-2-3-4 (*T. aestivum* L., 2n = 6x = 42, AABBDD); Chinese rye var. Qinling (*S. cereale* L., 2n = 2x = 14, RR); F_3 hybrid plant H-1, a cross between D-2-3-4 and Qinling, and two F_7 lines, YYW102 and YYW89. The two F_7 lines were derived from the successive selfing of the F_3 plant H-1. Rye var. Qinling had been selfed five times before being used.

2.2. Genomic in situ hybridization (GISH) analysis

The chromosome preparations were based on the procedure used by Zhang et al. (2007). GISH was based on the method according to Schwarzacher et al., 1992 with minor modifications. In this study, the proportion of rye DNA (probe) to wheat (Chinese spring) DNA (blocking DNA) was 1:160.

2.3. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and quality analysis

Protein was extracted and analyzed by SDS-PAGE according to Singh et al. (1991). The quality of dry seeds was analyzed with three repeats using a near-infrared spectroscopy analyzer (NIRSYS-TEMTR-3700C, Denmark FOSS). The three repeats were carried out using different seeds obtained from a mixed sample. Following the protocol provided by Denmark FOSS, data on protein content, gluten and sodium dodecyl sulfate sedimentation (SDS-sedimentation) were obtained from 7 g of whole grain per sample in a round sample cell. Student's *t*-test was undertaken using SPSS13.0 software (SPSS, Inc., Chicago, IL).

2.4. Cloning of the 75K γ-secalin genes and sequence analysis

Genomic DNA was extracted from the leaves of 10 days old plants as described by Yan et al. (2002). The complete coding region sequences for the genes encoding 75K γ-secalin at the *Sec2* locus were amplified by PCR using a pair of primers (F: 5′-ATGAAGACCTTACTCATGCTTGCAATA-3′ and R: 5′-TCAGTGGCCAACAATACCAGT-3′). The primers were designed using the previously reported sequences, AF201084 and EF432549 (Chen et al., 2008; Murray et al., 2001). PCR amplifications were carried out using a PTC-200 Thermocycler (MJ Research, Inc., Watertown, MA, USA) using Ex*Taq* DNA polymerase with a high fidelity (TaKaRa, Dalian, China) and followed the procedure outlined by (Chen et al., 2008). The PCR products were extracted from the agarose gel (Gel Extraction Kit, OMIGA) and the ligation products of the fragment were introduced into the pMD18-T plasmid vector (TaKaRa). This was then transformed into cells of *Escherichia coli* DH10B. The plasmid DNA was sent to a commercial company (Huada Gene, China) for sequencing.

The sequence analysis was carried out using DNAMAN version 5.2.2 (Lynnon Biosoft). Further adjustments to sequence alignments were performed manually. Based on the deduced amino acid sequences in the conserved regions (signal peptide, N-terminal and C-terminal), a neighbor-joining tree was constructed using MEGA4, the Kimura two-parameter distances model and a complete deletion of gaps (Kumar et al., 2004). The bootstrap values in the phylogenetic tree were estimated based on 1000 replications.

2.5. Bacterial expression of 75K γ-secalin genes

The intact open reading frames (ORF) of 75K γ-secalin genes were re-amplified using a pair of primers, P1 (5′-ACCCATATGAACATGCAAGTCAACCCTAGT-3′) and P2 (5′-TTCCTCGAGTCAGTGGCCAACAATA-3′), so that the signal peptide could be removed and restriction enzyme sites (*Nde* I and *Xho* I) could be added. PCR products were cloned into the bacterial expression vector, pET30a and transformed into *E. coli* strain BL21(DE3)pLsS. Polypeptide expression was induced by isopropyl β-Δ-thiogalactopyranoside (IPTG) (1 mM) over 4–6 h. The expressed proteins were extracted using two methods: the general extraction method, according to Singh et al. (1991), and the specific precipitation method, according to Murray et al. (2001).

3. Results

3.1. SDS-PAGE, cytology and quality analysis

SDS-PAGE analyses revealed four bands in the 75K γ-secalin region of rye var. Qinling (Fig. 1A). One of them was present in the F_3 hybrid H-1 and the F_7 lines, YYW89 and YYW102, while it was

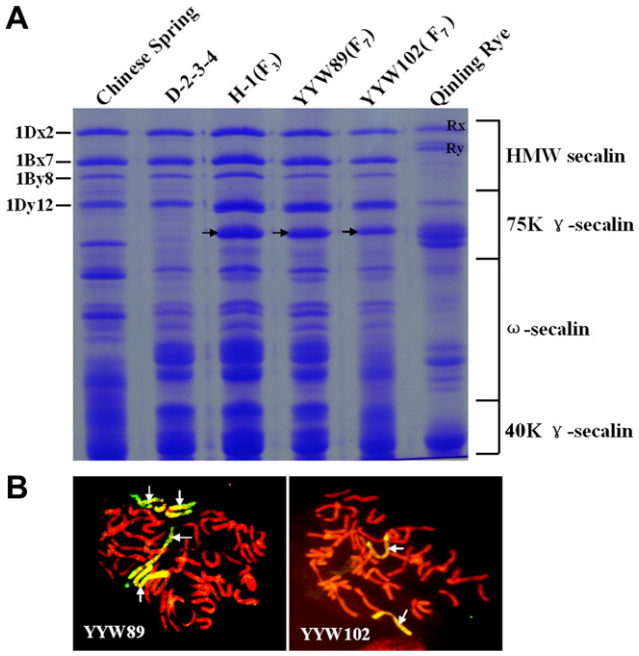

Fig. 1. A, SDS-PAGE analysis. The black arrows indicate the 75K γ-secalin protein from the hybrid derivatives. B, genomic in situ hybridization (GISH) analysis of YYW89 and YYW102. The white arrows indicate the rye chromosomes (yellow–brown). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Quality parameters comparison between the 2R addition line, YYW102, and its parents: rye var. Qinling and D-2-3-4 wheat.

Genotypes	Protein (d.m. %)[a]	Gluten (Pr. %)[b]	SDS (ml)[c]
Qinling rye	19.40 ± 0.25	40.25 ± 0.64	3.89 ± 0.40
D-2-3-4	16.19 ± 0.04*	33.66 ± 0.21	21.12 ± 2.40*
YYW102 (F_7)	17.21 ± 0.24*	34.05 ± 0.75	13.93 ± 0.72*

*significant level at $P<0.05$ by Student's *t*-test between D-2-3-4 and YYW102.
[a] In dry matter %.
[b] In protein %.
[c] Sodium dodecyl sulfate sedimentation (SDS-sedimentation) value.

absent in their common wheat parent, D-2-3-4. This indicated that the locus *Sec2*, of the parent rye, var. Qinling, was transferred into these derivatives. GISH analysis indicated that YYW102 (44 chromosomes) contained one pair of rye chromosomes and YYW89 (44 chromosomes), contained two pairs of rye chromosomes (Fig. 1B). The presence of 75K γ-secalins suggested that they carried chromosome 2R from rye. However, the absence of the rye HMW-secalin bands, Rx and Ry (Fig. 1A), indicated that they did not have chromosome 1R, the carrier of locus *Sec3* and *Sec1* (Hull et al., 1992).

Compared to its common wheat parent, D-2-3-4, the 2R addition line, YYW102, had a significantly higher protein content (Table 1). This suggested that incorporation of chromosome 2R increased the amount of protein in the wheat. However, YYW102 had a lower SDS-sedimentation value.

3.2. Cloning and sequence analysis

The PCR products with specific primers for 75K γ-secalin genes were cloned. Sixty clones, 12, 11, 15 and 22 from rye var. Qinling, the F_3 hybrid plant H-1 and the F_7 lines, YYW89 and YYW102, respectively, were successfully sequenced (Table 2). Two clones, one from H-1 and the other from YYW89, contained the same sequence, HQ266678. Of the 59 sequences (Table 2), six (HQ266723, HQ266724, HQ266725, HQ266726, HQ266727 and HQ266728) contained a stop codon, either TAG or TAA, due to a point mutation (C → T). Furthermore, a frameshift mutation had occurred in sequence HQ266689. The seven genes may not have been translated into secalin and were assumed to be pseudogenes (Table 2). Others had intact open reading frames (ORF) and were highly homologous to those previously reported for 75K γ-secalin (Fig. 2). They shared a similar primary structure, including four main structural regions: (a) a conservative signal peptide with 19 amino acids; (b) a steady short N-terminal region with 12 amino acids containing a cysteine; (c) a repetitive domain and (d) a conserved C-terminal domain with 143 amino acids, including seven or eight cysteines.

The variations among these sequences mainly appeared in the repetitive domain. The repetitive domain began with the conserved tetrapeptide PQ_3 and was rich in glutamine and proline. $PFPQ_{1-2}(PQQ)_{1-2}$ was the main conservative repeat motif in the repetitive region, although it did contain some variations and there were some amino acid substitutions in these sequences. An extra cysteine residue appeared in the repetitive region of HQ266709, HQ266679, HQ266692 and HQ266685 (Fig. 2, indicated by the box), which was the result of a point mutation of C → G in their genes.

An important feature of the repetitive region was that it was highly variable in length due to insertion or deletion (indels) events. These indels caused the sequence length to vary from 1440 bp to 609 bp. The molecular weights of their deduced amino acid residues ranged from 22.2 to 54.5 kDa (Table 2).

3.3. Bacterial expression analysis

Five sequences with different lengths: HQ266672 (1416 bp), HQ266693 (1239 bp), HQ266694 (1224 bp), HQ266715 (951 bp) and HQ266698 (945 bp), were expressed in *E. coli* and the expressed proteins were analyzed by SDS-PAGE (Fig. 3). Based on electrophoretic mobility, these expressed proteins showed differences in molecular weight, with a descending order of: HQ266672, HQ266694, HQ266693, HQ266715 and HQ266698. The bacterially expressed mature protein from HQ266672 showed an electrophoretic mobility identical to the 75K γ-secalins extracted from rye seeds (Fig. 3B). The other four sequences showed a much faster mobility than 75K γ-secalins and were located in the ω-secalin region (Fig. 3A).

3.4. Neighbor-joining analysis

In order to investigate the evolutionary relationships between 75K γ-secalin and other prolamins in wheat and rye, a neighbor-joining tree was constructed using their conserved regions (signal peptide, N-terminal and C-terminal regions). The results showed that the prolamins were divided into two clades (Fig. 4). The 75K γ-secalins, the 40K γ-secalin and the γ-gliadins were clustered at the top, while the ω-secalins and ω-gliadins were clustered at the bottom. In the top clade, the 75K γ-secalins were all found in one subgroup while the 40K γ-secalin and γ-gliadins were assigned to another subgroup. This suggested 40K γ-secalin was more closely related to γ-gliadins than to 75K γ-secalins.

4. Discussion

4.1. The Sec2 locus contained a number of genes with the typical structure of 75K γ-secalin

This study cannot exclude the existence of allelic heterozygosity at locus *Sec2* since rye is an out-crossing species, although

Table 2
Sequences isolated from the *Sec2* locus of rye var. Qinling and its derivatives.

Materials	GenBank accession number (Length[a] MW[b])
Qinling rye	HQ266670 (1440, 54.5), HQ266671 (1434, 54.2), HQ266672 (1416, 53.5), HQ266673 (1224, 45.9), HQ266674 (945, 35.1), HQ266675 (744, 27.4), HQ266676 (741, 27.2), HQ266677 (609, 22.2), HQ266723 (1395)[c], HQ266724 (1302)[c], HQ266725 (660)[c], HQ266726 (633)[c]
H-1 (F_3)	HQ266678 (1440, 54.4), HQ266679 (1212, 45.3), HQ266680 (1098, 40.9), HQ266681 (915, 33.9), HQ266682 (828, 30.6), HQ266683 (741, 27.2), HQ266684 (1344, 50.7), HQ266685 (1320, 49.7), HQ266686 (1296, 48.8), HQ266687 (1044, 38.9), HQ266688 (984, 36.8)
YYW89 (F_7)	HQ266678 (1440, 54.4), HQ266690 (1377, 52.0), HQ266691 (1314, 49.5), HQ266692 (1299, 49.0), HQ266693 (1239, 46.8), HQ266694 (1224, 45.9), HQ266695 (1152, 43,2), HQ266696 (1095, 40.8), HQ266697 (1068, 39.8), HQ266698 (945, 35.0), HQ266699 (936, 34.7), HQ266700 (840, 31.2), HQ266701 (741, 27.2), HQ266702 (660, 24.2), HQ266727 (876)[c]
YYW102 (F_7)	HQ266689 (851)[c], HQ266703 (1440, 54.4), HQ266704 (1356, 51.2), HQ266705 (1284, 48.4), HQ266706 (1272, 47.8), HQ266707 (1248, 46.9), HQ266708 (1224, 45.9), HQ266709 (1188, 44.5), HQ266710 (1146, 43.0), HQ266711 (1104, 41.2), HQ266712 (1083, 40.7), HQ266713 (1026, 38.2), HQ266714 (957, 35.8), HQ266715 (951, 35.5), HQ266716 (939, 34.9), HQ266717 (816, 30.3), HQ266718 (753, 27.8), HQ266719 (744, 27.4), HQ266720 (741, 27.3), HQ266721 (717, 26.3), HQ266722 (639, 23.4), HQ266728 (1140)[c]

Underlined sequences were used to express in *E. coli*.
[a] DNA sequence lengths.
[b] Molecular weight (kDa) of deduced amino acids.
[c] Represent the gene is a pseudogene.

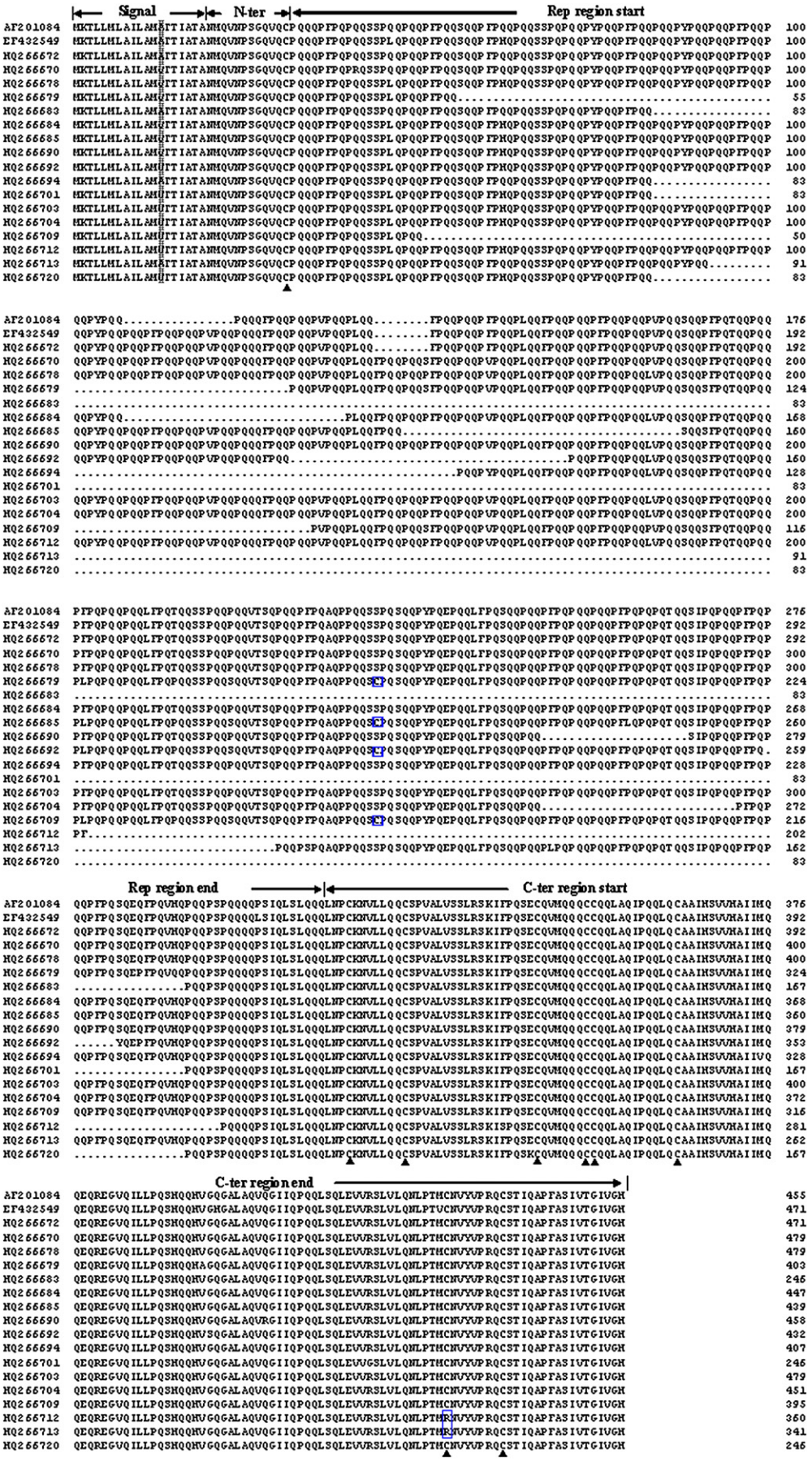

Fig. 2. Multiple sequence alignment of the deduced amino acid sequences of 75K γ-secalin genes from rye var. Qinling, H-1, YYW89 and YYW102 from this study; AF201084 from Murray et al. (2001) and EF432549 from Chen et al. (2008). The filled arrowhead indicates the cysteine residues. The cysteine residue mutants are boxed. The gray shading indicates differences in the signal peptide. Deletions are indicated by dots.

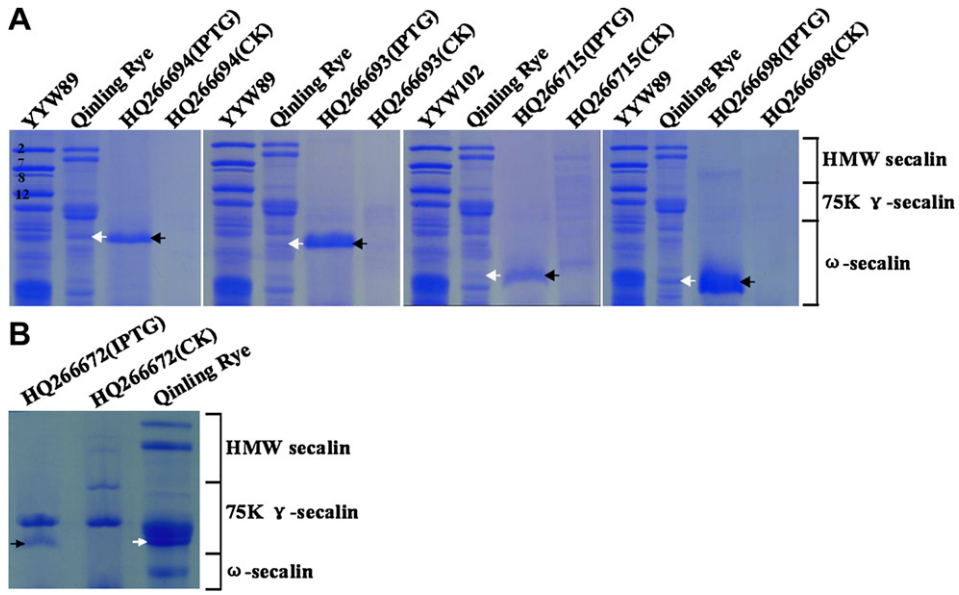

Fig. 3. SDS-PAGE analysis of the bacterially expressed genes isolated from *Sec2* Locus. IPTG: *E. coli* culture with isopropyl β-Δ-thiogalactopyranoside; CK: *E. coli* culture without IPTG; black arrows: the expressed mature protein from the *E. coli* culture; white arrows: the corresponding protein bands extracted from rye seed. A, the expressed proteins for HQ266693 (1239 bp), HQ266694 (1224 bp), HQ266715 (951 bp) and HQ266698 (945 bp) were extracted using the specific precipitation method according to Murray et al. (2001). B, the expressed proteins of HQ266672 (1416 bp) were extracted using the general extraction method according to Singh et al. (1991).

it had been selfed several times before being used. Therefore, the sequences isolated from the rye parent were not analyzed. The two F_7 lines and the F_3 plant H-1 line used in the present study were descended from the same haploid F_1 hybrid cross between rye and wheat (D-2-3-4) and thus contained the same *Sec2* locus. The 47 sequences isolated from the three derivatives can be assumed to represent paralogous genes from this locus. There could have been more paralogous genes present but only a limited number of clones were sequenced in this study. These results indicated that *Sec2* was a complex locus that contained a number of genes. It was unclear whether this was related to the high 75K γ-secalins content, which accounted for over 50% of the total seed storage protein in rye (Shewry et al., 1983a, b; Shewry and Field, 1982).

All the paralogous genes showed a similar structure with conserved sequences for the signal peptide, the N-terminal, the C-terminal and the repeat motif in the repetitive region. They were highly homologous with each other (Figs. 2 and 4). This indicated they originated from the same ancestor sequence. Based on the sequence similarity, the original 75K γ-secalin sequence may have been derived from a 40K γ-secalin gene (Fig. 4) (Gellrich et al., 2005; Murray et al., 2001; Shewry and Field, 1982). Since 75K γ-secalin genes appeared to be present in all species of *Secale* (Shewry et al., 1984a, b; Chen et al., 2008), the ancestor sequence must predate the speciation of this genus.

4.2. Duplication origin of paralogous genes at locus Sec2

It is probable that duplication and divergence led to the formation of multiple paralogous genes at *Sec2* (Shewry et al., 1984a). Duplication has been shown to play an important role in the gene evolution of prolamin storage proteins in common wheat and its related species (Kong et al., 2004; Shewry et al., 1994). The duplication origins of paralogous genes for rye HMW secalin at locus *Sec3* and ω-secalin at locus *Sec1* have been previously reported (Clarke et al., 1996). During duplication of these paralogous genes, there were insertions of spacer sequences or retrotransposons (Clarke et al., 1996; Kong et al., 2004; Wicker et al., 2003). Locus *Sec2* may also contain spacer sequences since the paralogous genes for 75K γ-secalins seem to be distributed inside a region with a genetic distance of about 0.5 cm (Malyshev et al., 1998).

As well as base sequence point mutations, the paralogous genes at locus *Sec2* were also highly variable in length due to insertion or deletion (indels) events in the repetitive region (Fig. 2). There were a number of conserved repeat motifs or direct repeat (DR) units in the repetitive region of these genes. The DR unit is a typical "signature", as Wicker et al. (2007) suggested, of an illegitimate recombinant event. Therefore, illegitimate recombination may be responsible for indels and thus for the length variation seen in the paralogous genes (Wicker et al., 2007; Yuan et al., 2011; Zhang et al., 2008).

4.3. The effect of Sec2 on wheat quality

The 75K γ-secalins accounted for over 50% of the total protein in rye (Shewry et al., 1983a,b; Shewry and Field, 1982). The incorporation of locus *Sec2* from rye into common wheat can increase the amount of protein in the mature grain from a 2RS.2BL chromosomal translocation line (Gupta et al., 1989). This study confirmed these findings by using a 2R addition line with a common wheat background. However, the incorporation of locus *Sec2* had a negative effect on flour processing quality since it reduced the SDS-sedimentation value (Table 1). This may be related to the distribution of cysteine residues. The eight cysteine residues in the C-terminal region can form four pairs of intra-chain disulphide bonds with each other, which may play an important role in stabilizing the glutenin polymers (Field et al., 1983; Murray et al., 2001; Shewry et al., 1994). However, the cysteine residue present in the N-terminal region was presumably able to form a disulphide bond with other unpaired cysteines in the gluten polymer, which would promote 'chain terminating' and stop lengthening of the gluten polymer. This would lead to a reduction in dough strength (Kasarda, 1989; Murray et al., 2001).

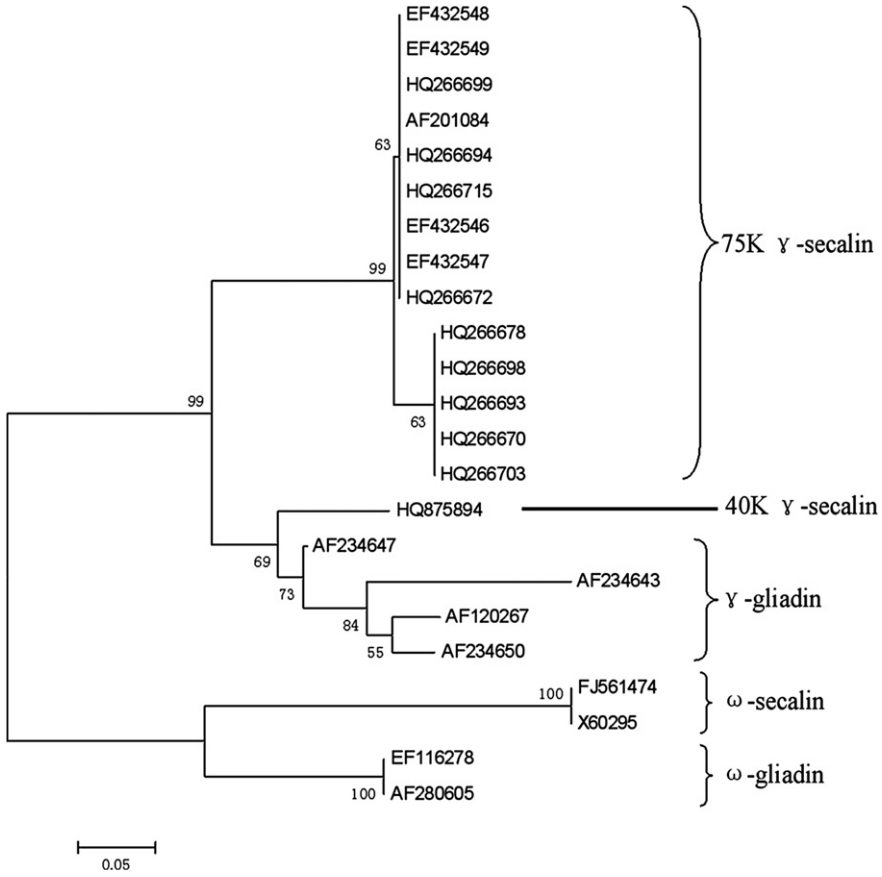

Fig. 4. Neighbor-joining tree of 75K γ-secalin and other prolamins, based on the alignment of the deduced amino acids in the conserved regions (signal peptide, N-terminal and C-terminal). Bootstrap values are indicated above or below the nodes.

In contrast, four of the proteins sequenced had an extra cysteine in the repetitive region (Fig. 2). The extra cysteine residue could make a secalin resemble a gliadin, which would lead to an improvement in dough quality (Chen et al., 2008). Two of the proteins studied had lost a cysteine residue in the C-terminal. The absence or presence of an additional cysteine in the C-terminal would significantly alter the way in which the protein binds to gluten polymers and may promote larger polymer formation and an increase in dough strength (Murray et al., 2001). Moreover, the change in protein length, due to insertion or deletion events in repetitive regions, may affect function. It is possible that the increased length could increase glutamine content and thus increase the viscosity and hydrogen bonding in the protein, in a similar way to C-hordeins and γ-gliadins (Chen et al., 2008).

4.4. Secalins with different molecular weights that are encoded by Sec2

With reference to their electrophoresis mobility on SDS-PAGE, secalins have been classified into HMW secalin, 75K γ-secalin, ω-secalin and 40K γ-secalin (Shewry et al., 1994). However, the electrophoresis mobility results from this study indicated that the molecular weight of deduced amino acid residues for some sequences with a typical 75K γ-secalin structure encoded by *Sec2*, were close to ω-secalin and 40K γ-secalin, which are encoded by *Sec1* (Table 1). This was also confirmed by the *E. coli* expression results (Fig. 3). From this study's results, it is difficult to differentiate the *Sec2* encoding secalins from *Sec1* based on molecular weight. The results from sequencing data on partial amino acids by Gellrich et al. (2005) also supported this finding. They also found amino acids specific for 75K γ-secalin by sequencing peptides in the 40K γ-secalin region.

Acknowledgments

This project was partially funded by the National Natural Science Foundation of China (31071417, 31071418 and 31101140), Project of Knowledge Innovation Engineering CAS (KSCX3-EW-N-02), West Action Plan (KZCX2-XB3-05) and the CAS 100 – Talent Program.

References

Chai, J.F., Liu, X., Jia, J.Z., 2005. Homologous cloning of ω-secalin gene family in a wheat 1BL/1RS translocation. Cell Research 15, 658–664.
Chen, Q.J., Yuan, Z.W., Zhang, L.Q., Yan, Z.H., Xiang, Z.G., Wan, Y.F., Zheng, Y.L., Liu, D.C., 2008. Molecular characterization and comparative analysis of four new genes from *Sec2* locus encoding 75K γ-secalin of rye species. Journal of Cereal Science 48, 111–116.
Clarke, B.C., Mukai, Y., Appels, R., 1996. The *Sec-1* locus on the short arm of chromosome 1R of rye (*Secale cereale*). Chromosoma 105, 269–275.
Field, J.M., Shewry, P.R., Miflin, B.J., 1983. Aggregation states of alcohol – stable storage proteins of barley, rye, wheat and maize. Journal of the Science of Food and Agriculture 34, 362–369.
Gellrich, C., Schieberle, P., Wieser, H., 2005. Studies of partial amino acid sequences of γ-40k secalins of rye. Cereal Chemistry 82, 541–545.
Gupta, R.B., Shepherd, K.W., MaCritchie, F., 1989. Effect of rye chromosome arm 2RS on flour properties and physical dough properties of wheat. Journal of Cereal Science 10, 169–173.
Hull, G., Sabelli, P.A., Shewry, P.R., 1992. Restriction fragment analysis of the secalin loci of rye. Biochemical Genetics 30, 85–97.

Jiang, Q.T., Wei, Y.M., Andre, L., Lu, Z.X., Pu, Z.E., Peng, Y.Y., Zheng, Y.L., 2010. Characterization of ω-secalin genes from rye, triticale, and a wheat 1BL/1RS translocation line. Journal of Applied Genetics 51, 403–411.

Kasarda, D.D., 1989. Glutenin structure in relation to wheat quality. In: Pomeranz, Y. (Ed.), Wheat is Unique. American Association of Cereal Chemists, St Paul, MN, pp. 277–302.

Kipp, B., Belitz, H.D., Seilmeier, W., Wieser, H., 1996. Comparative studies of high Mr subunits of rye and wheat. I. Isolation and biochemical characterization and effects on gluten extensibility. Journal of Cereal Science 23, 227–234.

Kong, X.Y., Gu, Y.Q., You, F.M., Dubcovsky, J., Anderson, O.D., 2004. Dynamics of the evolution of orthologous and paralogous portions of a complex locus region in two genomes of allopolyploid wheat. Plant Molecular Biology 54, 55–69.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA 3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5, 50–163.

Lawrence, G.J., Shepherd, K.W., 1981. Chromosomal location of genes controlling seed proteins in species related to wheat. Theoretical and Applied Genetics 59, 25–31.

Malyshev, S.V., Khmyl, T.O., Zabenkova, K.I., Voylokov, A.V., Korzun, V.N., Kartel, N.A., 1998. RFLP-based mapping of the *Sec-2* and *Sec-5* loci encoding 75K γ-secalins of rye. Plant Breeding 117, 329–333.

Miller, T.E., 1984. The homoeologous relationship between the chromosomes of rye and wheat. Current status. Canadian Journal of Genetics and Cytology 26, 578–589.

Murray, F.R., Skerritt, J.H., Appels, R., 2001. A gene from the *Sec2* (*Gli-R2*) locus of a wheat 2RS.2BL chromosomal translocation line. Theoretical and Applied Genetics 102, 431–439.

Schwarzacher, T., Anamthawat-Jónsson, K., Harrison, G.E., Islam, A.K.M.R., Jia, J.Z., King, I.P., Leitch, A.R., Miller, T.E., Reader, S.M., Rogers, W.J., Shi, M., Heslop-Harrison, J.S., 1992. Genomic in situ hybridization to identify alien chromosomes and chromosome segments in wheat. Theoretical and Applied Genetics 84, 778–786.

Shewry, P.R., Field, J.M., 1982. The purification and characterization of two groups of storage proteins (secalins) from rye (*Secale cereale* L.). Journal of Experimental Botany 33, 261–268.

Shewry, P.R., Kreis, M., Burgess, S.R., Parmer, S., Miflin, B.J., 1983a. The synthesis and deposition of the prolamin storage proteins (secalins) of rye. Planta 159, 439–445.

Shewry, P.R., Parmar, S., Miflin, B.J., 1983b. The extraction, separation and polymorphism of the prolamin storage proteins (secalins) of rye. Cereal Chemistry 60, 1–6.

Shewry, P.R., Bradberry, D., Franklin, J., White, R.P., 1984a. The chromosomal locations and linkage relationships of the structural genes for the prolamin storage proteins (secalins) of rye. Theoretical and Applied Genetics 69, 63–69.

Shewry, P.R., Miflin, B.J., Kasarda, D.D., 1984b. The structural and evolutionary relationships of the prolamin storage proteins of barley, rye and wheat. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 304, 297–308.

Shewry, P.R., Parmar, S., Miller, T.E., 1985. Chromosomal location of the structural genes for the Mr 75,000 γ-secalins in *Secale montanum* Guss: Evidence for a translocation involving chromosomes 2R and 6R in cultivated rye (*Secale cereale* L.). Heredity 54, 381–383.

Shewry, P.R., Parmar, S., Fulrath, N., Kasarda, D.D., Miller, T.E., 1986. Chromosomal locations of the structural genes for secalins in wild perennial rye (*Secale montanum* Guss.) and cultivated rye (*S. cereale* L.) determined by two-dimensional electrophoresis. Canadian Journal of Genetics and Cytology 28, 76–83.

Shewry, P.R., Miles, M.J., Tatham, A.S., 1994. The prolamin storage proteins of wheat and related cereals. Progress in Biophysics and Molecular Biology 61, 37–59.

Singh, N.K., Shepherd, K.W., Cornish, G.B., 1991. A simplfied SDS–PAGE procedure for separating LMW subunits of glutenin. Journal of Cereal Science 14, 203–208.

Wicker, T., Yahiaoui, N., Guyot, R., Schlagenhauf, E., Liu, Z.D., Dubcovsky, J., Keller, B., 2003. Rapid genome divergence at orthologous low molecular weight glutenin loci of A and A^m genomes of wheat. Plant Cell 15, 1186–1197.

Wicker, T., Yahiaoui, N., Keller, B., 2007. Illegitimate recombination is a major evolutionary mechanism for initiating size variation in plant resistance genes. The Plant Journal 51, 631–641.

Yan, Z.H., Wan, Y.F., Liu, K.F., Zheng, Y.L., Wang, D.W., 2002. Identification of a novel HMW glutenin subunit and comparison of its amino acid sequence with those of homologous subunits. Chinese Science Bulletin 47, 220–225.

Yuan, Z., Liu, D., Zhang, L., Chen, W., Yan, Z., Zheng, Y., Zhang, H., Yen, Y., 2011. Mitotic illegitimate recombination is a mechanism for novel changes in high-molecular-weight glutenin subunits in wheat-rye hybrids. PLoS ONE. http://dx.doi.org/10.1371/journal.pone.0023511.

Zhang, L.Q., Yen, Y., Zheng, Y.L., Liu, D.C., 2007. Meiotic restriction in emmer wheat is controlled by one or more nuclear genes that continue to function in derived lines. Sexual Plant Reproduction 20, 159–166.

Zhang, Y., Li, X., Wang, A., An, X., Zhang, Q., Pei, Y., Gao, L., Ma, W., Appels, R., Yan, Y., 2008. Novel x-type high-molecular-weight glutenin genes from *Aegilops tauschii* and their implication on the wheat origin and evolution mechanism of *Glu-D1-1* proteins. \ 178, 23–33.