Implications of an inverted duplication in the wheat 
*KN1*-type homeobox gene *Wknox1* for the 
origin of Persian wheat

Shigeo Takumi* and Ryoko Morimoto

Laboratory of Plant Genetics, Graduate School of Agricultural Science, Kobe University, 
Rokkodai 1-1, Nada, Kobe 657-8501, Japan

(Received 5 April 2015, accepted 28 April 2015)

Introgression between related species with different ploidy levels has played important roles in wheat subspecies differentiation. Persian wheat, a cultivated tetraploid wheat subspecies (*Triticum turgidum* subsp. *carthlicum*), is postulated to have evolved through interploidy hybridization between tetraploid and hexaploid wheats. Here, we report evidence for the origin of subsp. *carthlicum* based on the discovery of a new allele for the 5th-to-6th exon region of the *Wknox1b* *KNOTTED1*-type homeobox gene in a common wheat subspecies (*T. aestivum* subsp. *carthlicoides*). In this *Wknox1b* region, subsp. *carthlicoides* contains an inverted duplication mutation in the 3’ flanking region of a 157-bp MITE insertion site. This structural mutation resulted in the suppression of *Wknox1b* expression in subsp. *carthlicoides*, but no structural mutation was observed in the same region of subsp. *carthlicum*. In addition, the *carthlicum* allele for the *Wknox1b* 5th-to-6th exon region exhibited the same sequence as that in the wild emmer wheat subsp. *dicocoides*. These observations support an alternative hypothesis that subsp. *carthlicum* evolved by interploidy hybridization between subsp. *carthlicoides* and tetraploid wheat.

**Key words:** allopolyploidy, homeobox gene, homoeologous allele, hybridization, speciation

Polyploid wheat species evolved through natural hybridization and allopolyploidization, and the subsequent interploidy introgression between tetraploid and hexaploid wheat species was at least partly associated with subspecies diversification (Matsuoka, 2011). In cultivated tetraploid wheat, interploidy introgression is assumed to have contributed to the birth of two subspecies, Persian wheat, *Triticum turgidum* subsp. *carthlicum* (Neyski) Á. Löve & D. Löve (syn. *T. persicum* Vav.), and Georgian wheat, *T. turgidum* subsp. *paleocholchicum* (Menabde) Á. Löve & D. Löve. Subspecies *carthlicum* belongs to the cultivated tetraploid wheat group having the A and B genomes, and its cultivation area is distributed in the Transcaucasus region in countries such as Georgia and Armenia (van Slageren, 1994). Subspecies *carthlicum* is morphologically characterized by a long awn at an empty glume, controlled by the *tetraaristatus* (four-awned) gene on the long arm of chromosome 5A (Haque et al., 2011). The spike morphology of subsp. *carthlicum* resembles that of common wheat, *T. aestivum* L., rather than that of other subspecies of free-threshing tetraploid wheat. Moreover, the morphology of synthetic hexaploid wheat derived from crosses between subsp. *carthlicum* and *Aegilops tauschii* Coss., the D genome progenitor of common wheat, resembles that of common wheat (Kihara et al., 1950). Therefore, subsp. *carthlicum* is a candidate for the AB-genome donor of common wheat, whereas it is also considered to be a secondary species derived from an interspecific cross between emmer and common wheat (Vavilov, 1926, cited by Ohtsuka (1991)). Kuckuck (1979) found hexaploid wheat accessions showing the subsp. *carthlicum*-like morphology, and these accessions, called *T. aestivum* subsp. *carthlicoides* nom. nud., were distributed in the border region of Iran, Turkey and the Transcaucasus. Subspecies *carthlicum* was proposed to have originated from spontaneous hybridization between subsp. *carthlicoides* and cultivated emmer wheat, *T. turgidum* subsp. *dicoccon* (Schrank) Thell. (Kuckuck, 1979).

A wheat ortholog of maize *knotted1* (*kn1*) and rice *OSH1*, *Wknox1*, belongs to the class I *KN1*-type homeobox (*KNOX*) gene family on the homoeologous group 4 chromosomes (Takumi et al., 2000), which function in main-
tenance of shoot apical meristem (SAM) activity, determination of cell fate and vegetative organ development (Lincoln et al., 1994; Smith et al., 1995; Kerstetter et al., 1997). Our previous study showed that the three Wknox1 homeologous loci are functionally conserved in the three component genomes of common wheat, although many structural mutations containing MITE insertions in intron sequences have accumulated in each homoeoallele during wheat polyploid evolution (Morimoto et al., 2005).

Here, we report structural mutations newly found at the B-genome locus of Wknox1, designated Wknox1b, in subsp. carthlicoides. Based on a comparison of the carthlicoides Wknox1b allele with other alleles in related species, we discuss the evolution of subsp. carthlicum.

Cytological and molecular distinction of the two subspecies carthlicum and carthlicoides

Ten tetraploid accessions of T. turgidum (AABB) subsp. carthlicum, including KU-138, KU-139-1, KU-139-2, KU-187, KU-1800, KU-1801, KU-1807, KU-1808, CGN04221 and CGN08357, and two hexaploid accessions of T. aestivum L. (AABBDD) subsp. carthlicoides, KU-3724 and CGN08360, were used in this study. Plants were grown in a glasshouse of Kobe University for verification of species classification. Somatic chromosome numbers were determined from root-tip mitotic preparations of three seedlings from each accession of subsp. carthlicum and subsp. carthlicoides using the standard acetocarmine squash method. The morphology of the two carthlicoides accessions resembled that of subsp. carthlicum, as previously reported (Kuckuck, 1979; Haque et al., 2011), and the carthlicoides accessions had long awns at empty glumes, especially in late-emerging spikes. To validate ploidy levels of the subsp. carthlicum and subsp. carthlicoides accessions, we first checked chromosome numbers in root tips. Somatic chromosome numbers were 28 in all examined seeds from each accession of subsp. carthlicum, whereas the two carthlicoides accessions had 42 chromosomes.

The genomic sequences of the Wnnox1d 4th intron regions were amplified by PCR using the primer pair 5'-AAAAAAAAGGTTAATGGAC-3' and 5'-ACCTTATACATGATTGGGAA-3'. An insertion and deletion (indel) mutation site of the Wnnox1d locus was previously described (Morimoto et al., 2005), and the primer positions were set in the regions flanking this site. This primer pair precisely recognized and amplified the target region from the ancestral species of common wheat. A common wheat, T. aestivum cv. Chinese Spring (CS), and a durum wheat, T. turgidum subsp. durum cv. Langdon (Ldn), were also used for PCR. PCR was performed according to our previous study (Morimoto et al., 2005), and the amplified DNA fragments were visualized by ethidium bromide staining after electrophoresis on 1.5% agarose gels. In the 4th intron of Wnnox1d in common wheat, a 122-bp MITE insertion has been reported (Morimoto et al., 2005). The MITE-containing band was missing in tetraploid wheat accessions including all those examined of subsp. carthlicum (Fig. 1). The Wnnox1d-specific MITE insertion was also observed in the two carthlicoides accessions, indicating the presence of the D genome in subsp. carthlicoides as well as in other subspecies of common wheat.

**Fig. 1.** PCR-based haplotype analysis of the 4th intron of Wnnox1d. (A) Wnnox1d-specific primer positions in the 4th intron and 5th exon region. Arrows indicate the position of PCR primers. A 122-bp MITE insertion specifically observed at the D-genome locus of Wnnox1 generates the upper bands in hexaploid wheat accessions. (B) PCR amplification in 10 carthlicum accessions and two carthlicoides accessions. Arrows indicate PCR-amplified fragments. KU-7309 is an accession of T. turgidum subsp. dicoccon. Ldn, T. turgidum subsp. durum cv. Langdon; CS, T. aestivum cv. Chinese Spring.
Inverted duplication in a *Wknox1* homoeologous copy

Structural comparison of the *Wknox1b* locus in allopolyploid wheat  Similarly to the amplified region of *Wknox1d*, the *Wknox1b* 5th-to-6th exon region was amplified by PCR using the primer pair 5'-GCT-GAAGCACCATCTCCTGA-3' and 5'-CATGTAGAAGGGCG-GCGTTAG-3'. In addition to CS and Ldn, 10 accessions of wild tetraploid wheat, *T. turgidum* subsp. *dicoccoides*, and nine accessions of cultivated tetraploid wheat, *T. turgidum* subsp. *dicoccon*, were also used to analyze the 157-bp MITE indel polymorphism reported at the 5th intron of *Wknox1b* in tetraploid wheat (Fig. 2A; Morimoto et al., 2005). The accession numbers of the wild and cultivated tetraploid wheats are in our previous report (Morimoto et al., 2005). Common wheat and cultivated tetraploid wheat cultivars including the subsp. *dicoccon* accessions contained the MITE insertion in the *Wknox1b* 5th-to-6th exon region, whereas no fragments containing this insertion were observed in the subsp. *dicoccoides* accessions (Fig. 2B). The A- and D-genome progenitors *T. urartu* and *Ae. tauschii*, respectively, had missing *Wknox1b*-derived bands, and the subsp. *carthlicum* accessions showed the same-sized band, without the MITE insertion in the *Wknox1b* locus, as *Ae. speltoides* (Fig. 2C). The amplified fragments in the two *carthlicoides* accessions were clearly longer than those in either subsp. *carthlicum* or Ldn (Fig. 2, C and D). These observations indicated that the *carthlicoides* accessions contain a novel allele in the *Wknox1b* 5th-to-6th exon region.

The PCR fragments of the genomic regions near the *Wknox1b* 5th-to-6th exon in subsp. *carthlicum* and subsp. *carthlicoides* were cloned into the vector pGEM-T (Promega, Madison, WI, USA), and the nucleotide sequences were determined by an automated fluorescent DyeDeoxy terminator cycle sequencing system using an ABI PRISM 310 genetic analyzer (Applied Biosystems, Foster City, CA, USA). The nucleotide sequences of the inverted repeat region (see below) were determined using a CUGA sequencing system (Nippon Gene, Tokyo, Japan). The DNA sequences were analyzed by DNASIS software (Hitachi, Tokyo, Japan) and compared with those in other tetraploid and hexaploid wheats. The *carthlicoides* accessions contained the 157-bp MITE insertion in the 5th intron of *Wknox1b*, and the inserted MITE sequences were identical to those in CS, Ldn and the subsp. *dicoccon* accessions (Fig. 3). The genomic region from the 3’ end of the MITE sequence to the 6th exon was inverted in the *carthlicoides* accessions, and the inverted region contained a 1-bp insertion and a 13-bp deletion at the 5’ end. This inversion significantly disrupted the structure of the homeobox region encoding the DNA-binding motif, and made it impossible to form a precise homeodomain. Additionally, the central region of the 6th exon was duplicated in the *carthlicoides* accessions, and this duplication resulted in the appearance of longer PCR-amplified fragments for the *Wknox1b* 5th-to-6th exon region in the *carthlicoides* accessions compared with Ldn (Fig. 2D).

Subspecies *carthlicum* has been defined morphologically by its long awn at an empty glume, and the four-

---

**Fig. 2.** PCR-based haplotype analysis of the 5th-to-6th exon region of *Wknox1b*. A 157-bp MITE insertion generates the larger bands in *T. turgidum* subsp. *dicoccon*, Ldn and hexaploid wheat accessions. (A) *Wknox1b*-specific primer positions in the 5th-to-6th exon region. Arrows indicate the position of PCR primers. (B) PCR amplification in wild and cultivated emmer wheat accessions. Arrows indicate PCR-amplified fragments. (C) PCR amplification in the diploid progenitors, 10 *carthlicum* accessions and two *carthlicoides* accessions. Arrows indicate PCR-amplified fragments. (D) Comparison of three types of PCR-amplified fragments. Ldn, *T. turgidum* subsp. *durum* cv. Langdon; CS, *T. aestivum* cv. Chinese Spring; ura, *T. urartu* KU-199-1; spl, *Ae. speltoides* KU-1-2; tau, *Ae. tauschii* KU-20-1.
awned phenotype of each spikelet is controlled by the *tetraaristatus* gene on 5AL (Haque et al., 2011). Therefore, *Wknox1b* on chromosome 4B is unlikely to be directly associated with subsp. *carthlicum*. Nevertheless, our finding of the structural mutation at the *Wknox1b* locus of subsp. *carthlicoides* has implications for wheat speciation. Comparison of the genomic structure at the *Wknox1b* 5th-to-6th exon region among subspecies of tetraploid and hexaploid wheats strongly suggests that subsp. *carthlicoides* has not directly evolved from subsp. *carthlicum*, and is not a progenitor of other subspecies of hexaploid wheat (Fig. 3). It therefore appears likely that subsp. *carthlicoides* originated from other wheat hexaploids.

Subspecies *carthlicum* shares the same morphological characteristics as subsp. *carthlicoides* in spite of the differences in their genome constitution (Kuckuck, 1979; Haque et al., 2011). If a tight genetic relationship exists between subsp. *carthlicum* and subsp. *carthlicoides*, it is possible that subsp. *carthlicum* evolved from an interspecific cross between subsp. *carthlicoides* and another subspecies of *T. turgidum* based on comparison of their genomic structures in the *Wknox1b* 5th-to-6th exon region (Fig. 3). Interploidy hybridization between cultivated emmer wheat and subsp. *carthlicoides* was postulated to have resulted in the generation of subsp. *carthlicum* (Kuckuck, 1979). On the other hand, subsp. *paleocholchicum* is thought to have been generated by interploidy hybridization between wild emmer wheat and subsp. *carthlicoides*, and that subsp. *dicoccoides* is a parent of subsp. *carthlicum*, as well as of subsp. *paleocholchicum*. To clarify the parental subspecies of tetraploid wheat involved in interploidy hybridization, a wide survey of cultivated emmer wheat accessions lacking the 157-bp MITE in the *Wknox1b* 5th intron would be required. Structural analysis of the *Wknox1b* loci is useful for understanding the complicated evolutionary process of allopolyploid wheat speciation.

Comparative expression analysis of the *Wknox1* homoeologous copies To examine the effect of the inverted duplication in the 5th-to-6th exon region of *Wknox1b* on expression of homoeologous copies of *Wknox1* in subsp. *carthlicoides*, RT-PCR was conducted with primer pairs specific to each homoeologous copy using RNA from seedlings having a SAM. Total RNA was extracted using Sepasol-RNA I Super G solution (Nacalai Tesque, Kyoto, Japan) from shoots of 7-day-old seedlings. First-strand cDNA was synthesized from DNase I-treated RNA samples with oligo-dT primers and ReverTra Ace reverse transcriptase (Toyobo, Osaka, Japan), and used for PCR. In RT-PCR of the three homoeologous *Wknox1* transcripts, a common forward primer, 5'-GGAGGGTGAGACGCAACTCAACT-3', and the following three reverse primers, designed based on homoeolog-specific sequences in the 3' untranslated region, were used: 5'-CACCGACCAAGGTCACCAGT-3', 5'-CGCCGACCAAGGTGGACGGGC-3' and 5'-CCAAGGTCACCGGTAAACGT-3' for *Wknox1a*, *Wknox1b* and *Wknox1d*, respectively. For amplification of the ubiquitin gene as a control, the genespecific primer pair 5'-GCATGCAGATATTGTGAA-3' and 3'-GCATGCAGATATTGTGAA-5' were used.
Fig. 4. RT-PCR analysis of transcript accumulation from Wknox1 and its three homoeologous copies. Total RNA was extracted from SAM-containing shoots of each accession. The ubiquitin gene was used as an internal control. All of the amplified DNA fragments were separated by electrophoresis through a 2% agarose gel and stained with ethidium bromide. The number of PCR cycles is shown at the right side of the electropherograms.

and 5'-GGAGCTTACTGGCCAC-3' was used. For amplification of total Wknox1 transcripts, the following primer pair was used: 5'-TCTCCCACCCCCACTACTC-3' and 5'-GGAGCTTACTGGCCAC-3' was used. For amplification of DNA fragments were separated by electrophoresis through a 2% agarose gel and stained with ethidium bromide. The number of PCR cycles is shown at the right side of the electropherograms.

Thus, the inverted duplication in the Wknox1 5th-to-6th exon region could suppress the expression of Wknox1 in subsp. carthlicoides (Fig. 4), suggesting that carthlicoides Wknox1b is a null allele. The functions of the Wknox1 orthologs are essential for SAM maintenance and plant development, and loss-of-function mutants in Arabidopsis and maize show lethal or extreme morphological defects due to disruption of the SAM (Long et al., 1996; Kerstetter et al., 1997). Although many gain-of-function mutation alleles of Wknox1 orthologs have been well studied in maize, tomato and barley (Smith et al., 1996; Kuckuck, H. (1979) On the origin of Triticum persicum (=Triticum aestivum VAV.). Wheat Inform. Serv. 50, 1–5. Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K., and Hake, S. (1994) A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. Plant Cell 6, 1859–1876. Long, J. A., Moan, E. I., Medford, J. I., and Barton, M. K. (1996) polar paired (PP-1) homeobox gene family in Arabidopsis. Plant Cell 8, 1105–1117.}

REFERENCES

Chapman, E. J., and Carrington, J. C. (2007) Specialization and evolution of endogenous small RNA pathways. Nat. Rev. Genet. 8, 884–896.

Chen, J. J., Janssen, B. J., Williams, A., and Sinha, N. (1997) A gene fusion at a homeobox locus: alterations in leaf shape and implications for morphological evolution. Plant Cell 9, 1289–1304.

Dvorak, J., and Luo, M. C. (2001) Evolution of free-threshing and hulled forms of Triticum aestivum: old problems and new tools. In: Wheat Taxonomy: The Legacy of John Percival, (eds: Caligari, P. D. S., and Brandham, P. E.), pp. 127–136. Academic Press, London.

Haque, M. A., Takayama, A., Watanabe, N., and Kuboyama, T. (2011) Cytological and genetic mapping of the gene for four-awned phenotype in Triticum carthlicoides Nevski. Genet. Resour. Crop Evol. 58, 1087–1093.

Kerstetter, R. A., Laudencia-Chingcuano, D., Smith, L. G., and Hake, S. (1997) Loss-of-function mutations in the maize homeobox gene, knotted1, are defective in shoot meristem maintenance. Development 124, 3045–3054.

Kihara, H., Okamoto, M., Ikegami, M., Takushi, J., Suemoto, H., and Yamane, Y. (1950) Morphology and fertility of five new synthesized hexaploid wheats. Seiken Zoho 4, 127–140. (in Japanese with English summary)

Kuckuck, H. (1979) On the origin of Triticum carthlicoides NEYSKI (=Triticum persicum VAV.). Wheat Inform. Serv. 50, 1–5.

Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K., and Hake, S. (1994) A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. Plant Cell 6, 1859–1876.

Long, J. A., Moan, E. I., Medford, J. I., and Barton, M. K. (1996)
A member of the KNOTTED class of homeodomain proteins encoded by the STM gene of Arabidopsis. Nature 379, 66–69.

Matsuoka, Y. (2011) Evolution of polyploid Triticum wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. Plant Cell Physiol. 52, 750–764.

Morimoto, R., Kosugi, T., Nakamura, C., and Takumi, S. (2005) Intragenic diversity and functional conservation of the three homoeologous loci of the KN1-type homeobox gene Wknox1 in common wheat. Plant Mol. Biol. 57, 907–924.

Müller, K., Romano, N., Gerstner, O., Gracia-Marrotot, F., Pozzi, C., Salamini, F., and Rohde, W. (1995) The barley Hooded mutation caused by a duplication in a homeobox gene intron. Nature 374, 727–730.

Ohtsuka, I. (1991) Genetic differentiation in wheat nuclear genomes in relation to compatibility with Aegilops squarrosa cytoplasm and application to phylogeny of polyploid wheats. J. Fac. Agric., Hokkaido Univ. 65, Part 2, 127–198.

Shitsukawa, N., Tahira, C., Kassai, K., Hirabayashi, C., Shimizu, T., Takumi, S., Mochida, K., Kawaura, K., Ogihara, Y., and Murai, K. (2007) Genetic and epigenetic alteration among three homoeologous genes of a class E MADS-box gene in hexaploid wheat. Plant Cell 19, 1723–1737.

Smith, L. G., Greene, B., Veit, B., and Hake, S. (1992) A dominant mutation in the maize homeobox gene, Knotted-1, causes its ectopic expression in leaf cells with altered fates. Development 116, 21–30.

Smith, L. G., Jackson, D., and Hake, S. (1995) Expression of knotted1 marks shoot meristem formation during maize embryogenesis. Dev. Genet. 16, 344–348.

Takumi, S., Kosugi, T., Murai, K., Mori, N., and Nakamura, C. (2000) Molecular cloning of three homoeologous cDNAs encoding orthologs of the maize KNOTTED1 homeobox protein from young spikes of hexaploid wheat. Gene 249, 171–181.

van Slageren, M. W. (1994) Wild Wheats: A Monograph of Aegilops L. and Amblyopyrum (Jaub. & Spach) Eig (Poaceae). p. 512. Wageningen Agricultural University papers, The Netherlands.

Vavilov, N. I. (1926) Studies on the origin of cultivated plants. Trudi po Prikl. Bot., Gen. I Selek. (Leningrad) 16, 1–248. (in Russian)