DOI: 10.1534/genetics.107.074765

Note

The Role of Double-Stranded Break Repair in the Creation of Phenotypic Diversity at Cereal *VRN1* Loci

James Cockram, Ian J. Mackay and Donal M. O'Sullivan[1]

National Institute of Agricultural Botany, Cambridge CB3 0LE, United Kingdom

Manuscript received July 27, 2007
Accepted for publication October 19, 2007

ABSTRACT

Nonhomologous repair of double-stranded breaks, although fundamental to the maintenance of genomic integrity in all eukaryotes, has received little attention as to its evolutionary consequences in the generation and selection of phenotypic diversity. Here we document the role of illegitimate recombination in the creation of novel alleles in *VRN1* orthologs selected to confer adaptation to annual cropping systems in barley and wheat.

DURING their lifecycle, plants are exposed to a variety of environmental and endogenous factors that may damage DNA integrity, such as ionizing radiation, DNA replication failure, and retrotransposon activity. Two pathways are involved in repair of potentially lethal double-stranded breaks (DSBs) in eukaryotes (reviewed by Puchta 2005). Homologous recombination (HR) utilizes DNA sequence homology from an intact copy of the damaged region (for example from the sister chromatid) as a template for break repair. Alternatively, DSBs can be repaired by nonhomologous end-joining (NHEJ, also known as illegitimate recombination) where breakpoints are rejoined end to end, requiring little or no sequence homology. Much effort has focused on understanding the process of HR for its potential applications in breeding and biotechnology. NHEJ (which is less accurate, but more efficient and widespread in nature) has been less studied, chiefly in plants by analysis of inducible DSBs in transgenic model systems (*e.g.*, Salomon and Puchta 1998; Kirik *et al.* 2000). Although comparison of truncated transposable elements (TEs) shows NHEJ to have profound implications for reduction of plant genome size (reviewed by Bennetzen 2007), the action of this mechanism on sequences other than TEs has not been extensively documented.

Here, we document the impact of DSB repair on phenotypic diversity using the orthologous *VRN1* genes of wheat and barley (reviewed by Trevaskis *et al.* 2007) as a model system in which a major adaptive change in phenotype—conversion of the ancestral vernalization-responsive winter growth habit to a nonresponsive spring growth habit—is frequently conferred by deletions within regulatory noncoding regions of orthologous MADS-box transcription factor genes *VRN-H1* (barley), *VRN-A1*, *-B1*, and *-D1* (hexaploid wheat) (Fu *et al.* 2005). To investigate the molecular mechanisms underlying these deletions, we sequenced seven intron I rearrangements associated with spring *Vrn-H1* alleles identified in a screen of 429 European barley cultivars (further details presented in Cockram *et al.* 2007). In addition, intron I and promoter deletions associated with previously sequenced spring *Vrn1* alleles from hexaploid wheat and its diploid wheat relatives *Triticum monococcum* and *Aegilops tauschii* were also analyzed.

Sequence alignment of spring *Vrn1* alleles with winter alleles, presumed to be ancestral to spring forms (Takahashi and Yasuda 1971), permitted precise localization of deletion breakpoints. Twenty spring barley and wheat *Vrn1* alleles were analyzed, including nine intron I deletions and five promoter deletions, as well as six instances of TE insertion into the promoter or intron I (Table 1). Of the 14 spring *Vrn1* alleles carrying putative functional deletions, 10 display short repeated sequences of 3–7 bp immediately flanking the deleted regions (Figure 1), which range in size from 20 bp in the *T. monococcum VRN-A^{m}1* promoter to 8.9 kb within *VRN-H1* intron 1. Probability tests on base composition show these motifs are significantly associated with deletion breakpoints at $p < 0.03$ (Figure 1A). The presence of small patches of nucleotide homology

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. EF591635–EF591650.

[1]*Corresponding author:* National Institute of Agricultural Botany, Huntingdon Rd., Cambridge CB3 0LE, United Kingdom.
E-mail: donal.osullivan@niab.com

TABLE 1

Full length *VRN1* promoter and intron I sequences from barley and wheat, detailing location and size of putative functional deletions, the short sequence repeats that flank them, and details of TEs

Species (genome)	Accession	Strain/cultivar (allele designation)	Allele	Insertion/deletion type (size)	Flanking nucleotide repeat	Notes
Tm (A^m)	AY188331[a]	DV92	Winter	NA	NA	Intron I (9.8 kb): 2 MITEs, *Sukkula* retroelement[a]
Tm (A^m)	DQ146421, DQ146423	PI503874 (*Vrn1f*), PI306540 (*Vrn1h*)	Spring	Intron I TE insertion	NA	TE insertion into intron I
Tm (A^m)	[b]	G2528	Spring	Promoter, CArG-box (20 bp)	AACCC	
Tm (A^m)	DQ146422	PI326317 (*Vrn1g*); PI1349049	Spring	Promoter, CArG-box (34 bp)	TTT	
Tm (A^m)	[b]	PI355515	Spring	Promoter, CArG-box (48 bp)	CCTCCCC	
Tt (A)	AY616462	(*Vrn-A1d*)	Spring	Promoter, CArG-box (32 bp)	ATCC	
Tt (A)	AY616463	(*Vrn-A1e*)	Spring	Promoter, CArG-box (54 bp)	GGCCA	
Ta (A)	AY747600[a]	Tripple Dirk C	Winter	N/A		Intron I (8.5 kb): MITE, *Sukkula* retroelement[a]
Ta (A)	AY747599	IL369	Spring	Intron I deletion (6.8 kb)	CCAC	3′ breakpoint flanked by *Sukkula* retroelement
Ta (A)	AY747601	Tripple Dirk D	Spring	Promoter TE insertion		Intron I deletion as above, TE insertion into promoter
Ta (A)	AY616458	Tripple Dirk TDD, isolate 1, (*Vrn-A1a*)	Spring	Promoter TE insertion		TE insertion into promoter
Ta (A)	AY616459	Tripple Dirk TDD, isolate 2	Spring	Promoter TE insertion, truncated (91 bp)	TGAATGACA	TE insertion into promoter, partial TE deletion not associated with a change in growth habit
Ta (A)	AY616460	Anza	Spring	Promoter TE insertion, truncated (179 bp)	GGTCTCATA	TE insertion into promoter, partial TE deletion not associated with a change in growth habit
Ta (B); *Tt* (B)	AY747604[a], AY747602	Tripple Dirk C; Langdon	Winter	NA	NA	Intron I (9.7 kb): 2 AU SINE elements[a]
Ta (B)	AY747603	Tripple Dirk B	Spring	Intron I deletion (6.8 kb)	Repeated CCGG, CCA filler DNA at 5′ breakpoint	AU SINE flanking 3′ breakpoint
Ta (D); *At* (D)	AY747606[a], AY747605	Tripple Dirk C; unknown	Winter	NA	NA	29-bp sequence repeat flanking 109 bp del outside "vernalization critical" region
Ta (D)	AY747597	Tripple Dirk E	Spring	Intron I deletion (4.2 kb)	GCCT	
Hv (H)	AY750994, AY866487, EF591645, AY750993, AY866490[a]	Dicktoo; Kompolti Korai; Pearl (haplotype 1A); Strider	Winter	NA	NA	Intron I (11 kb): MITE, Loalog solo LTR[a]
Hv (H)	DQ924860, EF591649	Etu (haplotype 1B), Maskin	Spring	Intron I deletion (3.9 kb)		5′ breakpoint 28 bp upstream of haplotype 2 TE insertion. 23 bp filler DNA at 3′ breakpoint identical to sequence <45 bp downstream.

(*continued*)

TABLE 1

(Continued)

Species (genome)	Accession	Strain/cultivar (allele designation)	Allele	Insertion/deletion type (size)	Flanking nucleotide repeat	Notes
Hv (H)	EF611825, EF591650	Tremois; Varunda (haplotype 2)	Spring	Intron I TE insertion	NA	TE insertion into intron I
Hv (H)	AY750996, EF591636	OWB-D; Dandy (haplotype 3)	Spring	Intron I deletion (6.3 kb)		5′ breakpoint 27 bp downstream of MITE
Hv (H)	AY871789, EF591639	Triumph; Optic (haplotype 4A)	Spring	Intron I deletion (8.9 kb)	GCCT	
Hv (H)	EF591641	Prisma (haplotype 4B)	Spring	Intron I deletion (5.2 kb)	AAC	Same deletion as haplotype 5A
Hv (H)	AY750995, EF591640	Morex; Pohto (haplotype 5A)	Spring	Intron I deletion (5.2 kb)	AAC	Same deletion as haplotype 4B
Hv (H)	AY866494, EF591646, AY866495	Albacete; Calicuchima-sib; Oriol (haplotype 5B)	Spring	Intron I deletion (4.1 kb)	ACCCCGA	Repeat recessed by 2 bp at 3′ breakpoint (Figure 1A)
Hv (H)	DQ924859, EF591647	Ager; Express (haplotype 5C)	Winter	Intron I deletion (486 bp)		Deletion within intron I Loalog solo LTR

Species abbreviations: *At*, *Aegilops tauschii*; *Hv*, *Hordeum vulgare* ssp. *vulgare*; *Ta*, *Triticum aestivum*; *Tm*, *T. monococcum*; *Tt*, *T. turgidum* ssp. *dicoccoides*.

[a] Spring cereal *Vrn-A*^{m}*1*, *-A1*, *-B1*, *-D1*, and *-H1* alleles are aligned in Figure 1 to their respective reference winter alleles (underlined).

[b] No GenBank accession given; DNA sequence from Yan *et al.* (2004).

flanking deletions repaired by NHEJ are a prerequisite for the operation of a single-strand annealing (SSA)-like mechanism (reviewed by Puchta 2005). Such flanking motifs have previously been observed in plants surrounding deletions within TEs (Devos *et al.* 2002; Wicker *et al.* 2003; Ma *et al.* 2004; Bennetzen *et al.* 2005). Their presence flanking putative functional deletions within *VRN1* genes suggests this mechanism is responsible for the creation of the majority of spring *Vrn1* alleles. The error prone nature of NHEJ is further supported by the observation of filler DNA segments of variable length at other breakpoint junctions (Figure 1B). In the case of *VRN-B1* and *VRN-H1* haplotype 3 intron I deletions, these 6–7 bp filler segments flanking breakpoint boundaries are identical to motifs <30 bp away. Interestingly, where barley and wheat intron I breakpoints lie within conserved regions, colocation of breakpoints can be observed (Figure 1), possibly reflecting the independent utilization of conserved repeat motifs in these regions for DSB repair.

TEs have played a major role in the expansion of genome size in plant species (Bennetzen 2000) and in the creation of novel alleles selected during crop domestication and varietal differentiation (Doebley *et al.* 2006). In addition, transposon activity/retroelement integration has been suggested to play a role in genome reduction due to induction of DSBs, followed by NHEJ via a SSA-like mechanism (Puchta 2005). Although a limited number of TEs are found within cereal *VRN1* genes (Table 1), their location immediately adjacent to deletion breakpoints in three of the spring alleles studied (Figure 1) suggests that insertion/excision events may have played a role in the formation of DSBs that led to the deletions within *VRN1* intron 1. Short sequence repeats flanking the truncation of duplicated foldback elements within the *VRN-A1* gene have previously been observed, although these reductions were not associated with a change in phenotype (Yan *et al.* 2004). However, it is interesting to note that the small *VRN1* promoter deletions (20–54 bp) associated with spring alleles in *T. monococcum* all flank the insertion position of the same foldback element referred to above, suggesting that the spring alleles observed may be associated with the activity of this element, rather than due to random replication slippage as previously suggested (Yan *et al.* 2004).

We conclude that similar mechanisms operate in the repair of DSBs that have resulted in independent selection of spring alleles at orthologous cereal *VRN1* loci and suggest NHEJ via a SSA-like mechanism is the predominant pathway utilized. Although deletion mutation due to NHEJ is assuming rapidly increasing prominence in human disease genetics (*e.g.*, Kozak *et al.* 2006; Le Guédard *et al.* 2007), this report is the first example in which this mechanism is implicated in the creation of naturally occurring adaptive variation in plants. This does not necessarily mean that such variation is not both abundant and potentially of great significance. It may simply be that the right type of study—comparison of long tracts of coding and noncoding genomic sequence from divergent specimens of closely related taxa—has not yet been carried out. A glimpse of the

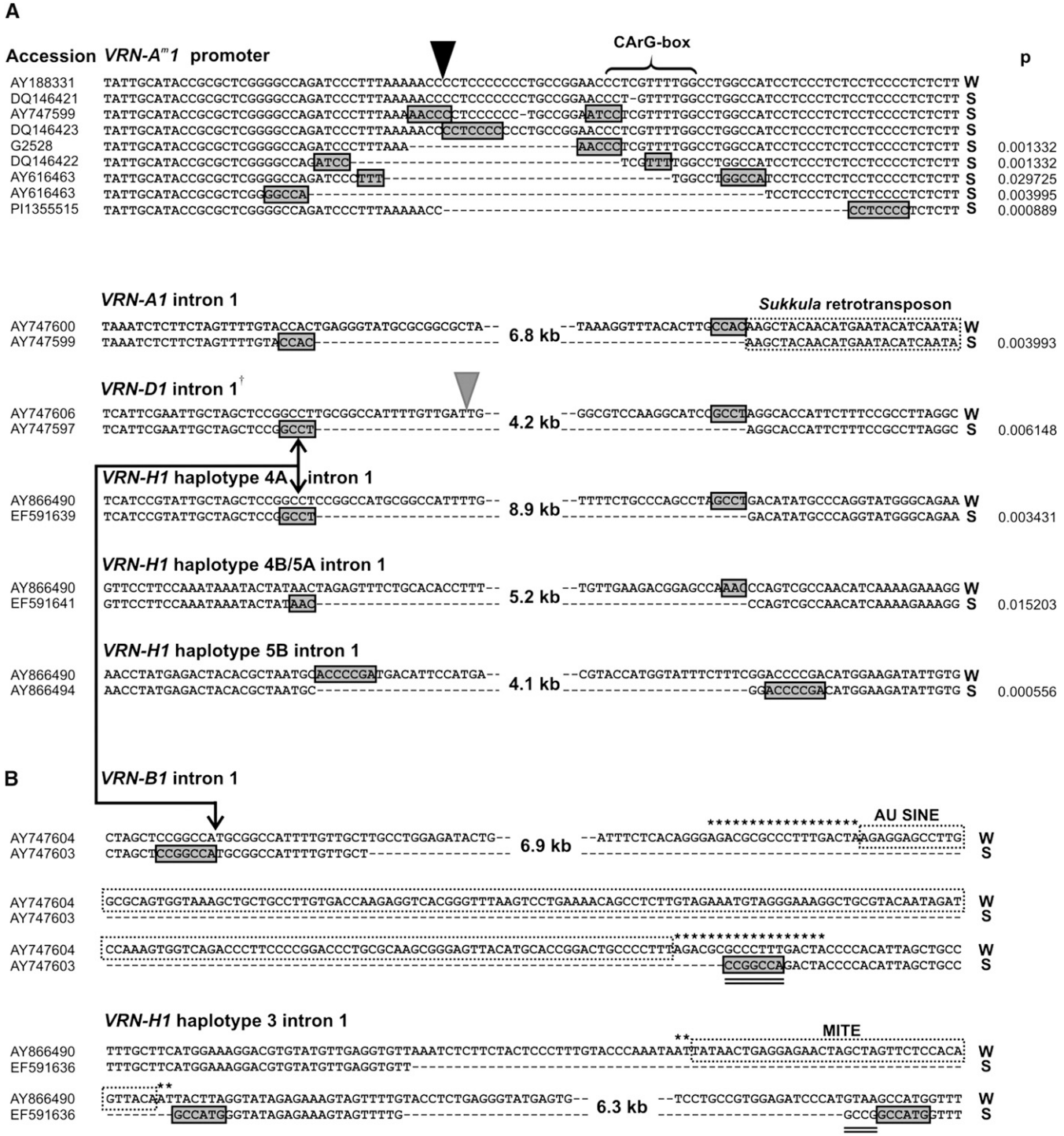

FIGURE 1.—Nucleotide sequence alignments flanking deletions in spring and winter alleles of orthologous cereal *VRN1* genes. (A) Promoter and intron deletions exhibiting short flanking repeat motifs precisely flanking the breakpoint with no filler sequence. (B) Intron I deletions where flanking repeats are interspersed with short flanking filler sequences (double underscored). The probability (p) that short sequence repeat motifs flank deletion breakpoints is adapted from the methods described by DEVOS *et al.* (2002): the number of times the short sequence repeat occurs within intron I of the reference winter allele is divided by the number of possible sequences of identical length within the intron. Where a repeat motif starts at the nth nucleotide following the breakpoint rather than the first, the probability is multiplied by n to correct. Deletions within the promoter are indicated by dashed lines; intron I deletions are indicated by dashed lines, separated by a gap of indicated size. Short sequence repeats flanking deletion boundaries are boxed in gray. Transposable elements (TEs) are boxed by dashed lines. Conserved 5′ breakpoints in spring alleles from barley and wheat are indicated by arrowed lines. W, winter allele; S, spring allele; TE insertion footprint is indicated by an asterisk. The position of TE insertion events in wheat and barley are indicated by solid and shaded triangles, respectively. The single dagger indicates flanking repeats in D-genome spring allele noted by FU *et al.* (2005).

potential is afforded from the example of Indica and Japonica rice, where whole genome draft sequences of the subspecies were compared. In this analysis, 4 of 78 inferred deletions with short flanking repeats mined from alignments were associated with introns or exons, and a majority of the remaining deletions were judged to be in single copy regions (Ma and Bennetzen 2005). These findings suggest the current focus in plant genetics on SNP variation within coding sequences should be complemented with approaches such as multiplex ligation-dependent probe amplification (Schouten *et al.* 2002) and comparative genomic hybridization using whole genome tiling arrays (*e.g.* Urban *et al.* 2006), which can robustly call the "null" alleles associated with deletions for detection and location of functionally important variation in plant species.

LITERATURE CITED

Bennetzen, J. L., 2000 Transposable element contributes to plant gene and genome evolution. Plant Mol. Biol. **42:** 251–269.

Bennetzen, J. L., 2007 Patterns in grass genome evolution. Curr. Opin. Plant Biol. **10:** 176–181.

Bennetzen, J. L., J. Ma and K. M. Devos, 2005 Mechanisms of recent genome size variation in flowering plants. Annals of Bot. **95:** 127–132.

Cockram, J., E. Chiapparino, K. Stamati, S. A. Taylor, P. Donini *et al.*, 2007 Haplotype analysis of vernalization loci in European barley germplasm reveals novel *VRN-H1* alleles and a predominant winter *VRN-H1/VRN-H2* multilocus haplotype. Theor. Appl. Genet. **115:** 993–1001.

Devos, K. M., J. K. M. Brown and J. L. Bennetzen, 2002 Genome size reduction through illegitimate recombination counteracts genome expansion in *Arabidopsis*. Genome Res. **12:** 1075–1079.

Doebley, J. F., B. S. Gaut and B. D. Smith, 2006 The molecular genetics of crop domestication. Cell **127:** 1309–1321.

Fu, D. L., P. Szűcs, L. L. Yan, M. Helguera, J. S. Skinner *et al.*, 2005 Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. Mol. Genet. Genomics **274:** 442–443.

Kirik, A., S. Salomon and H. Putcha, 2000 Species-specific double-stand break repair and genome evolution in plants. EMBO J. **19:** 5562–5566.

Kozak, L., E. Hrabincova, J. Kintr, O. Horky, P. Zapletalova *et al.*, 2006 Identification and characterization of large deletions in the phenylalanine hydroxylase (PAH) gene by MLPA: evidence for both homologous and non-homologous mechanisms of rearrangement. Mol. Genet. Metab. **89:** 300–309.

Le Guédard, S., V. Faugère, S. Malcolm, M. Claustres and A.-F. Roux, 2007 Large genomic rearrangements within the *PCDH15* gene are a significant cause of USH1F syndrome. Mol. Vision **13:** 102–107.

Ma, J., K. M. Devos and J. L. Bennetzen, 2004 Analysis of LTR-retrotransposon structure reveal ancient and rapid genomic DNA loss in rice. Genome Res. **14:** 860–869.

Ma, J., and J. L. Bennetzen, 2005 Rapid recent growth and divergence of rice nuclear genomes. Proc. Natl. Acad. Sci. USA **101:** 12404–12410.

Puchta, H., 2005 The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J. Expt. Bot. **56:** 1–14.

Salomon, S., and H. Puchta, 1998 Capture of T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J. **17:** 6086–6095.

Schouten, J. P., C. J. McElgunn, R. Waaijer, D. Zwijnenburg, F. Diepvens *et al.*, 2002 Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res. **30:** e57.

Takahashi, R., and S. Yasuda, 1971 Genetics of earliness and growth habit in barley, pp. 388–408 in *Barley Genetics II, Proceedings of the Second International Barley Genetics Symposium*, edited by R. A. Nilan. Washington State University Press, Pullman, WA.

Trevaskis, B., M. N. Hemming, E. S. Dennis and W. J. Peacock, 2007 The molecular basis of vernalization-induced flowering in cereals. Trends Plant Sci. **12:** 352–357.

Urban A. E., J. O. Korbel, R. Selzer, T. Richmond, A. Hacker *et al.*, 2006 High-resolution mapping of DNA copy alterations in human chromosome 22 using high-density tiling oligonucleotide arrays. Proc. Natl. Acad. Sci. USA **103:** 4534–4539.

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

Yan, L., A. Loukolanov, G. Tranquilli, M. Helguera, T. Fahima *et al.*, 2003 Positional cloning of the wheat vernalization gene *VRN1*. Proc. Natl. Acad. Sci. USA **100:** 6263–6268.

Yan, L., M. Helguera, K. Kato, S. Fukuyama, J. Sherman *et al.*, 2004 Allelic variation at the *VRN-1* promoter region in polyploidy wheat. Theor. Appl. Genet. **109:** 1677–1686.

Communicating editor: E. S. Buckler