Natural antisense transcripts of *Trifolium repens* dehydrins

Irina I Vaseva* and Urs Feller

1Plant Stress Molecular Biology Department; Institute of Plant Physiology and Genetics; Bulgarian Academy of Sciences; Sofia, Bulgaria; 2Institute of Plant Sciences and Oeschger Centre for Climate Change Research (OCCR); University of Bern; Bern, Switzerland

**Keywords:** Dehydrins, natural antisense transcripts, sense-antisense pairs, splice variants, *Trifolium repens*

The recently described complex nature of some dehydrin-coding sequences in *Trifolium repens* could explain the considerable variability among transcripts originating from a single gene. For some of the sequences the existence of natural antisense transcripts (NATs), which could form sense-antisense (SAS) pairs, was predicted. The present study demonstrates that *cis*-natural antisense transcripts of 2 dehydrin types (*Y*<sub>nK</sub> and *Y*<sub>nSK</sub>) accumulate in white clover plants subjected to treatments with polyethylene glycol (PEG), abscisic acid (ABA), and high salt concentration. The isolated *Y*<sub>nK</sub> *cis*-NATs mapped to sequence site enriched in alternative start codons. Some of the sense-antisense pairs exhibited inverse expression with differing profiles which depended on the applied stress. A natural antisense transcript coding for an ABC family protein (a *trans*-NAT) which shares short sequence homology with *Y*<sub>nSK</sub> dehydrin was identified in plants subjected to salt stress. Forthcoming experiments will evaluate the impact of NATs on transcript abundances, elucidating the role of transcriptional and post-transcriptional interferences in the regulation of dehydrin levels under various abiotic stresses.

Dehydrins (group 2 LEA proteins) tend to accumulate late in embryogenesis and in response to stress leading to cell dehydration (e.g., drought, low temperature, and salinity). The expression pattern of group 2 LEA genes is frequently associated with higher tolerance of crop plants to abiotic stresses such as cold and drought. All dehydrins have at least one conserved, lysine-rich 15-amino acid domain, EKK-GIMDKIKEKLP, named the K-segment near the C-terminus, and some of them may have a track of Ser residues (the S-segment), and/or a consensus motif (in one or more copies), T/VDEYGNP (Y-segment) located near the N-terminus. The number and order of the Y-, S-, and K-segments define different dehydrin sub-classes: *Y*<sub>nSK</sub>, *Y*<sub>nK</sub>, *SK*<sub>n</sub>, *K*<sub>n</sub>, and *K*<sub>nS</sub>. Detailed analyses of promoter regions of some DHN genes provided evidence for a close relationship between dehydrin expression patterns and the upstream and downstream cis-regulatory elements present in the sequences. Published data outlined that the regulation of expression of some dehydrin genes is elaborate and could be a result of several interacting factors.

Results from a recent study aimed at the identification of different dehydrin types in white clover (*Trifolium repens*) demonstrated the complex nature of dehydrin-coding sequences, which may lead to a high variability among the transcripts originating from a single gene. The aim of the present study was to confirm experimentally the existence of previously predicted dehydrin NAT transcripts and to compare their abundance in white clover grown under different abiotic stress treatments.

Natural antisense RNAs could potentially regulate the expression of their sense partner(s) at either transcriptional or post-transcriptional level. A recent work described that a rice *cis*-natural antisense RNA acts as a translational enhancer for its cognate mRNA. SAS pairs in plants have the potential to become substrates for the ribonuclease III-like enzyme Dicer to produce short interfering RNAs (siRNAs) and natural antisense microRNAs (nat-miRNAs) with regulatory potential. Many studies have confirmed that abiotic or biotic stresses induce production of the so-called nat-siRNA (natural-small interfering RNA) from *cis*-NATs.

Plants used in the present study (*Trifolium repens*, cv Apis) were grown hydroponically on a standard nutrient solution for 2 weeks (23 °C night / 26 °C day; 80% relative humidity; 14 h photoperiod; photosynthetic active irradiation of 200 μmol m<sup>-2</sup> s<sup>-1</sup>). Some of them were subjected to 72 h treatment with PEG (100 g PEG-6000 added to 1 L nutrient solution), and others were transferred on standard nutrient medium supplemented with 1 μM ABA or 75 mM NaCl and were grown for additional 14 d. Total RNA extracted from differently treated plants was reversely transcribed in sense and in antisense direction, with oligo(dT<sub>23</sub>) (Sigma-Aldrich) and forward gene specific primers (*Y*<sub>nK</sub> F: ATGAATATGG AAACCCAGTG).
and \( Y_{\text{SK}} \) F: ACTGTCAACACTCCTAATCCAACTTC, respectively. The expression of \( T. repens \) \( \alpha \)-tubulin (GenBank ID: AY192359.1) in sense and antisense direction (Fig. 1) was used as a reference. 21 It is known that approximately 8–10% of Arabidopsis mRNAs have naturally occurring antisense transcripts. 22,23 When the sense and antisense transcripts originate from the same genomic locus but in opposite directions, the natural antisense transcripts are called \( \text{cis} \)-NATs. Usually they are characterized by perfect RNA–RNA sequence complement. As reported earlier, a \( Y_{\text{nKn}} \) dehydrin gene could be expressed in several splice variants. 1 Previously, Jen et al. 22 have noticed that this was a specific characteristic of \( \text{cis} \)-NAT encoding genes where the RNA splicing could be induced by antisense transcripts.

The primer pair (Forward CTTATCAACC ATTCCTTTTT CACC/ Reverse CAGCAACAATATGGAGACAGGG) used in the \( Y_{\text{nKn}} \) RT-PCR of differently treated samples was designed to amplify sense products with expected sizes 308 (transcripts with GenBank IDs: KC756187 and KC756188), 170 (GenBank ID: KC247805), and 140 b.p. (GenBank ID: KC756186, which was not detected in the tested samples). \( Y_{\text{nKn}} \) splice variants formed sense-antisense (SAS) transcript pairs (Fig. 1), which was confirmed by direct sequencing of the amplification products. The isolated \( \text{cis} \)-NATs manifested a full overlap with transcript variants KC756187 and KC756188, both coding for \( Y_{\text{nKn}} \) dehydrin types (Fig. 2A). Vector NTI analysis of \( \text{cis} \)-NATs position on \( Y_{\text{nKn}} \) dehydrin-coding DNA revealed that they both map to a site enriched in alternative start codons (Fig. 2B).

Direct sequencing confirmed that the antisense \( Y_{\text{SKn}} \) 144 b.p. (obtained with primers Forward GGTGCTTATTG GTCGCGGTGCA/ Reverse CTTGAACCTGGAGGACGAGCAT) accumulating in PEG- and NaCl-treated leaves (Fig. 1), was a full overlap \( \text{cis} \)-NAT of the target \( Y_{\text{SKn}} \) dehydrin gene showing high homology to transcript variant KC247804, as well as to JF748411 and KF234077 (Fig. 3A). The last 2 were previously identified as drought-inducible. 1,24 The 144 b.p. \( \text{cis} \)-NAT exhibited comparable expression to its sense partner in leaf samples from the experiment with younger clovers treated with PEG for 72 h. It has been proposed that the similar sense and antisense transcript abundance is controlled by shared regulatory elements or processes related to chromatin structure. 25 Although a certain level of transcriptional noise resulting from bidirectional promoter activity cannot be excluded, the observed differing expression patterns of SAS under various stresses imply a possible physiological function.

The \( \text{trans} \)-NATs are pairs of overlapping transcripts from different genomic loci and often code for a protein or microRNA. 12,23 A genome-wide screening of \( \text{trans} \)-NATs in Arabidopsis thaliana has identified 1320 putative \( \text{trans} \)-NAT pairs, and it has been concluded that 430 transcripts had both putative \( \text{cis} \) and \( \text{trans} \)-NATs. 26 \( \text{trans} \)-NATs may have imperfect sequence complementarities. A natural antisense transcript (561 b.p.), sharing short nucleotide homology (24 base pairs) with \( Y_{\text{SKn}} \) dehydrin gene, was isolated as well (Fig. 1).
sequence codes for ATPase component of ABC transporters (ABC F family, Gen Bank ID: KF373076) which lack a transmembrane domain and function in processes other than transport, most probably in ribosome recycling and translational control.27,28 The transcript was specifically induced by salt stress both in leaves and in roots (Fig. 1). Direct sequencing and BLAST analysis of the antisense RT-PCR band migrating at position around 200 b.p. (GenBank ID: KF552071) visible in controls, ABA- and PEG-treated leaf samples, confirmed that it was a splice variant of ABC F coding sequence. The alignment of the isolated trans-NATs and YnSKn-coding sequence showed that the short homology between them lays upstream in the intron and downstream at the beginning of ORF3 (Fig. 3B).

The present results will initiate extended studies on the duplex formation between sense and antisense dehydrin transcripts and their potential as functional regulators of diverse gene expression and transcript stability.

Disclosure of Potential Conflicts of Interest
There are no potential conflicts of interest.

Acknowledgments
This study was supported by a grant from the Scientific Exchange Programme between Switzerland and the New Member States of the European Union – Sciex-NMSch (Project No. 11.113: “Identification of Dehydrin Types involved in Abiotic Stress Responses in Trifolium repens” – IDAST) and the NCCR Climate (Project PLANT/SOIL).
Figure 3. (A) Multiple sequence alignment (MegAlign, DNASTAR, Lasergene) of Y_SK transcripts with the isolated cis-NAT (144 b.p.). The alignment showed that the isolated cis-NAT shares 93.8% identity (5.8% divergence) with transcript variants JF748411 and KF234077. (B) Vector NTI analyses of the genomic T. repens Y_SK sequence (GenBank ID: GU443965.1) and position of the isolated cis-NAT (red block) and trans-NATs (GenBank ID: KF373076) coding for ABC F protein (blue blocks). The first exon of Y_SK genomic sequence, which codes for 3 Y-segments and for the stretch of 9 Ser-residues, contains 2 open reading frames – ORF 1 (blue arrow1–399 b.p.; alternative starts at positions 10 and 223 b.p., indicated with green arrows) and the nested ORF 2 (173–343 b.p., alternative starts at 182, 188, and 191 b.p., indicated with green arrows). The second exon comprises ORF 3 (blue arrow 607–831 b.p., alternative starts at 655 and 667 b.p., indicated with green arrows), and it codes for 2 K-segments.

References


22. Jen C-H, Michalopoulos I, Westhead DR, Meyer P. Natural antisense transcripts with coding capacity in *Arabidopsis* may have a regulatory role that is not linked to double-stranded RNA degradation. Genome Biol 2005; 6:R51; PMID:15960803; http://dx.doi.org/10.1186/gb-2005-6-6-r51


