UC Davis
UC Davis Previously Published Works

Title
Cereal genes similar to Snf2 define a new subfamily that includes human and mouse genes

Permalink
https://escholarship.org/uc/item/9g07t162

Journal
Molecular Genetics and Genomics, 268(4)

ISSN
1617-4615

Authors
Yan, L
Echenique, V
Busso, C
et al.

Publication Date
2002-12-01

DOI
10.1007/s00438-002-0765-3

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Mol Genet Genomics (2002) 268: 488–499
DOI 10.1007/s00438-002-0765-3

ORIGINAL PAPER

L. Yan · V. Echenique · C. Busso · P. SanMiguel
W. Ramakrishna · J.L. Bennetzen · S. Harrington
J. Dubcovsky

Cereal genes similar to *Snf2* define a new subfamily that includes human and mouse genes

Received: 21 March 2002 / Accepted: 23 September 2002 / Published online: 24 October 2002
© Springer-Verlag 2002

Abstract Genes from the SNF2 family play important roles in transcriptional regulation, maintenance of chromosome integrity and DNA repair. This study describes the molecular cloning and characterization of cereal genes from this family. The predicted proteins exhibit a novel C-terminal domain that defines a new subfamily designated *SNF2P* that includes human and mouse proteins. Comparison between genomic and cDNA sequences showed that cereal *Snf2P* genes consisted of 17 exons, including one only 8 bp long. Two barley alleles differed by the presence of a 7.7-kb non-LTR retrotransposon in intron 6. An alternative annotation of the orthologous Arabidopsis gene would improve its similarity with the other members of the subfamily. Intron 2 was not spliced out in approximately half of the rice *Snf2P* mRNAs present in leaves, resulting in a premature stop codon. Transcripts from the barley and wheat *Snf2P* genes were found in apexes, leaves, sheaths, roots and spikes. The *Snf2P* genes exist as single copies on wheat chromosome arm $5A^{m}L$ and in the colinear regions on barley chromosome arm 4HL and rice chromosome 3. High-density genetic mapping and RT-PCR suggest that *Snf2P* is not a candidate gene for the tightly linked vernalization gene *Vrn2*.

Keywords SNF2 family · Alternative splicing · Comparative genomics · Wheat · Vernalization gene *Vrn2*

The first two authors contributed equally to this work

Communicated by A. Kondorosi

L. Yan · V. Echenique · C. Busso · J. Dubcovsky (✉)
Dept. of Agronomy and Range Science,
One Shields Avenue, University of California,
Davis, CA 95616, USA
E-mail: jdubcovsky@ucdavis.edu
Tel.: +1-530-7525159
Fax: +1-530-7524361

P. SanMiguel
Purdue University Genomics Core,
Purdue University, West Lafayette,
IN 47907, USA

W. Ramakrishna · J.L. Bennetzen
Dept. of Biological Sciences,
Purdue University, West Lafayette,
IN 47907, USA

S. Harrington
Dept. of Plant Breeding,
Cornell University, Ithaca,
NY 14853, USA

Present address: V. Echenique
Dept. de Agronomía, Univ. Nacional del Sur,
8000 Bahía Blanca, Argentina

Introduction

Members of the SNF2 family of proteins share the SNF2 domain, defined by the presence of seven motifs with sequence similarity to DNA helicases (Coleman et al. 2000). Despite the presence of these helicase-related motifs, none of the proteins from this family has yet been demonstrated to have an helicase function. In general, members of the SNF2 family show the capacity to use the energy released by their DNA-dependent ATPase activity to stabilize or perturb protein-DNA interactions (Pazin and Kadonaga 1997; Muthuswami et al. 2000).

Among the best-characterized members of this family are the genes included in the *SNF2* subfamily. Recent whole-genome expression studies have shown that genes from the *SNF2* subfamily control the transcription of about 6% of all genes in *Saccharomyces cerevisiae* (Kingston et al. 1996; Sudarsanam et al. 2000), and their products are required in vivo for the establishment of a transcriptionally active chromatin structure (Gavin and Simpson 1997; Wu and Winston 1997). These genome-wide expression analyses also suggested that control was exerted at the level of individual promoters rather than over chromosomal domains (Sudarsanam et al. 2000). Genetic and biochemical studies of SNF2 and related proteins

indicated that they can destabilize nucleosome structure and thereby facilitate the binding of transcription factors to chromatin in promoter regions (Kingston et al. 1996).

Whole-genome mRNA expression studies suggested that genes from the SNF2 family could also act as transcriptional repressors. Almost half (25–40%) of the genes affected in defective *snf2* mutants in yeast express increased levels of mRNA. Another example is the *Drosophila domino* gene, which contributes to the silencing of homeotic genes (Ruhf et al. 2001).

Proteins from the SNF2 family have been classified into 13 subfamilies with a range of biological functions, including transcriptional regulation, maintenance of chromosome integrity, DNA repair, and maintenance of genomic methylation (Eisen et al. 1995; Jeddeloh et al. 1999; Coleman et al. 2000). Most of the proteins included in the studies by Coleman et al. (2000) were from bacteria, yeast, *Drosophila*, mouse, and human; they looked at no examples from plant species. Recently, numerous proteins related to members of the different SNF2 subfamilies have been predicted from the genome sequence of *Arabidopsis thaliana* (Lin et al. 1999; Kaneko et al. 2000; Salanoubat et al. 2000), but no formal classification was attempted.

We found sequences homologous to one of the *Snf2*-related Arabidopsis genes (GenBank AC007659_19) in diploid wheat (*Triticum monococcum* L., hereafter referred to as wheat), barley and rice during a chromosome walk toward the vernalization gene *Vrn2* of wheat (Dubcovsky et al. 1998; Dubcovsky 2001; L. Yan and J. Dubcovsky, unpublished results). *Vrn2* is involved in the repression of flowering in barley and diploid wheat (Tranquilli and Dubcovsky 1999), and is located on the distal part of the long arm of homoeologous group 4 (in a segment translocated to chromosome 5A in both diploid and polyploid wheats; Dubcovsky et al. 1998). This gene is different from *Vrn-B1,* which was known as *Vrn2* in the old *Vrn* classification (McIntosh et al. 1998).

The involvement of different members of the SNF2 family in transcriptional repression (Ruhf et al. 2001) and in the maintenance of genomic methylation (Jeddeloh et al. 1999), and the absence of recombination in a mapping population of 150 gametes, suggested that the *Snf2*-related gene might be a good candidate for the *Vrn2* gene. Sequences from complete cDNAs from barley and rice were used to determine the gene structure in cereals and to compare genes from winter and spring barley varieties. Expression profiles and high-density genetic maps are presented to show that this *Snf2*-related gene in Triticeae is not a strong candidate gene for the vernalization gene *Vrn2*. In spite of this negative result, valuable information was obtained on the structure and expression of the *Snf2*-related genes in these experiments. We also show that the proteins coded by these plant genes, together with proteins from human and mouse, define a new subfamily within the SNF2 family.

Materials and methods

Sequence alignment and nomenclature

The 350 amino acids extending from the C-terminus of the Arabidopsis protein (GenBank AC007659_19) and the related proteins identified in this study were used to search the non-redundant database using BLASTP (Altschul et al. 1997). Human protein AAH01171 and mouse protein NP_080815 were the only ones to show a significantly similar C-terminal domain, and were aligned with the proteins encoded by the plant genes using the multiple sequence alignment program ClustalW1.8 (Thompson et al. 1994). The amino acid sequence alignment figure was produced with Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html).

BLASTP searches using the conserved SNF2 domain showed that the most closely related proteins belong to the *SNF2*, *SNF2L*, *CHD1* and *YFK8* subfamilies. A phylogenetic study was performed using the six SNF2-related proteins listed above plus three members of each of the four most related SNF2 subfamilies and one member of the other subfamilies as defined by Coleman et al (2000). The Accession Nos. for the members of the *SNF2* subfamily included in this study were S66910 (human), M61703 (yeast), and NP_187252 (Arabidopsis); for the *SNF2L* subfamily M89907 (human), S46122 (yeast), and NP_187291 (Arabidopsis); for the *CHD1* subfamily XP_004000 (human), NP_011091 (yeast), and AAD28668 (Arabidopsis); and for the *YFK8* subfamily NP_032260 (mouse), P43610 (yeast), and NP_201476 (Arabidopsis). The Arabidopsis member of the *YFK8* subfamily is the *DDM1* gene, which is involved in the maintenance of genomic DNA methylation (Jeddeloh et al. 1999). Only one member each was included from the less related subfamilies *CSB* (NP012569, yeast), *RAD54* (M63232, yeast), *YSCD* (Z48618, yeast), and *MOT1* (M83224, yeast). Finally, only one member each was included from the two pairs of distantly related groups including subfamilies *HEPA1-HARP* (XP046726, human) and *HEPA2-RAD16* (M86929, yeast) (Coleman et al. 2000).

Phylogenetic trees were generated from the ClustalW sequence alignments using multiple distance- and parsimony-based methods available in the MEGA2.1 computer software package (http://www.megasoftware.net/; see Kumar et al. 1994). All sites containing alignment gaps were removed before the calculations (complete-deletion option). Distances between each pair of proteins were calculated and trees were constructed using four different methods to determine the robustness of the groups to variations in clustering techniques. The four techniques compared in this study included Neighbor-Joining (NJ), the Unweighted Pair Group Method using Averages (UPGMA), Maximum Parsimony (MP) and Minimum Evolution (ME) with the default parameters implemented in MEGA2.1. Consensus trees and confidence values for the nodes were calculated for each of the four analyses using 500 bootstraps (MEGA2.1).

BLASTP searches of the C-terminal domain and consistent clusters from the phylogenetic studies were used to define the subfamilies. Based on these criteria it will be shown that the group of proteins including the Arabidopsis AC007659_19, the cereal proteins described in this study, the human AAH01171 and mouse NP_080815 define a new subfamily that will be designated *SNF2P* throughout the text. Genes coding for the proteins included in the *SNF2P* subfamily will be indicated in lower case italics (*Snf2P*) and preceded by two letters indicating the species to facilitate gene comparisons (*At*, *Arabidopsis thaliana*; *Tm*, *Triticum monococcum*; *Hv*, *Hordeum vulgare*; *Os*, *Oryza sativaç Hs*, *Homo sapiens*; *Mm*, *Mus musculus*). Subfamily names will be indicated in italics to differentiate the SNF2 family from the *SNF2* subfamily.

Selection and sequencing of BAC clones

BAC clones from wheat, barley and rice were selected with the genomic RFLP probes UCW1 (Nucellin; Chen and Foolad 1997) and UCW2, which were previously found to be tightly linked to the vernalization gene *Vrn2* (Dubcovsky et al. 1998; L. Yan and

J. Dubcovsky, unpublished results). Shotgun sequencing libraries were constructed for *T. monococcum* BAC 455C17 (Lijavetzky et al. 1999), barley BAC 615K1 (Yu et al. 2000) and rice *Hin*dIII BAC 49F5 [M.A. Budiman, M.-Z. Luo, J.P. Thomkins, H.-R. Kim and R.A. Wing (2001) "Construction and characterization of two deep-coverage BAC libraries for the rice (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare) genome sequencing project"; Clemson University Genomics Institute] following a previously described procedure (Dubcovsky et al. 2001). The BAC clones were completely sequenced with an error rate lower than 1/10,000 bp. Assembled sequences were confirmed experimentally by restriction digestion and hybridization with two fragments of the barley *Snf2P* gene.

Genomic and cDNA sequencing of the *Snf2P-2* gene from the winter barley variety Dairokkaku

The genomic sequence of *Snf2P-1* from the spring barley variety Morex (BAC 615K1) was compared with the genomic sequence of *Snf2P-2* from the winter barley variety Dairokkaku. Seven overlapping PCR fragments (for primer sequences, see Table 1) including the 5′ untranslated leader, the complete coding region, and the 3′ untranslated region, were cloned and sequenced (Table 1)

A cDNA from barley variety Dairokkaku was also cloned in five overlapping fragments including the stop codon and sequenced (Table 1). Total RNA was extracted using the Trizol method (Gibco/BRL, Catalog No. 15596-026), mRNA was isolated with the PolyATtract kit from Promega, and first-strand cDNA was generated using Superscript II (Gibco/BRL). The PCR products generated by RT-PCR were purified on PCR Wizard columns (Promega) and cloned with the pCR 4-TOPO TA cloning kit (Invitrogen). Ligation reactions were transformed into competent Electro MAX DH10B cells.

Mapping of the *Snf2P* genes in cereal plants

In wheat, the *Snf2* gene was first mapped in a population of 74 F_2 plants from the cross between *T. monococcum* ssp. *monococcum* DV92 (spring growth habit) and *T. monococcum* ssp. *aegilopoides* G3116 (winter growth habit) previously used to map the vernalization gene *Vrn2* (Dubcovsky et al. 1998). In addition, 2775 new F_2 plants from the same cross were analyzed for recombination between *Vrn2* and *Snf2P*. Plants showing a recombinant chromosome in this region were selected using the proximal RFLP marker WG199 (later replaced by UCW26) and the distal marker UCW1(Nucellin). Clones used in this study were described before (Dubcovsky et al. 1996), with the exception of UCW1(Nucellin) (a genomic clone including exons 4–6 from the barley Nucellin gene (Chen and Foolad 1997), and UCW2 and UCW26, which are low-copy-number genomic clones from *T. monococcum* with no significant similarity to known genes.

In *H. vulgare* L., a population of 146 F_2 plants from a cross between *H. vulgare* ssp. *vulgare* cv. Morex × *H. vulgare* ssp. *spontaneoum* (C. Koch) Thell., kindly provided by Dr. A. Kleinhofs, was used to map the barley *Snf2P* gene. The rice *Snf2P* gene was mapped using 83 BC_1 plants from the cross *O. sativa* L./*O. longistaminata* A.Chev et Roher//*O. sativa* (Causse et al. 1994). DNA extraction, and gel-blot and hybridization protocols were performed as previously described (Dubcovsky et al. 1994).

Expression profiles of *Snf2P* genes

Expression of the *Snf2P* genes was studied in winter and spring accessions of *T. monococcum* (DV92: *vrn2* allele for spring growth habit; G3116 and G2528: *Vrn2* allele for winter growth habit; Tranquilli and Dubcovsky 1999), and barley (Dairokkaku and Hayakiso: *Vrn2* allele for winter growth habit; and Morex and Iwate Mensury C: *vrn2* allele for spring growth habit). Total RNA was extracted from vegetative apexes, leaves, sheaths and roots from the different genotypes 3–4 weeks after sowing in the greenhouse. Total RNA from leaves was also extracted from wheat and barley genotypes with a winter growth habit before and after 6 weeks of vernalization at 8°C (after 1 day at room temperature). Young spikes at the boot stage were harvested from vernalized plants of *T. monococcum* G3116 and barley Dairokkaku.

RNA extraction, first-strand cDNA synthesis, and PCR conditions for the expression profile studies were as described above for the cDNA cloning experiment. Although detailed quantitative analyses of the expression of *Snf2P* genes were not included in this study, all first-strand cDNAs were generated from the same amount of mRNA (500 ng) to standardize the initial conditions of

Table 1 Sequences and locations of primers used

Primer No.[a]	Sequence (5′→3′)	Location	Orientation
1	CAAACGCTGACTCTCG TAGCC	5′-UTR	Forward
2	ATGGCGGCCACCGCCTA CGA	Exon 1	Forward
3	GGCAATTTCTCTGCTGAGTAC	Exon 3	Forward
4	CACCGACATACTGGAT GACCT	Exon 4	Reverse
5	AACGCTTTATGATGCCAAGG	Exon 6	Forward
6	GAAAACTGAAGGCATGCAGAA	Exon 6	Reverse
7	TGAAGCTAACCAAGCAAACAG	Exon 7	Forward
8	GGGCAAAAATCCAAGAAGGTT	Intron 7	Reverse
9	TCATCGTGTTGTGCTATTTGC	Exon 10	Forward
10	TGTTTGTCAGCCTGAGGATTC	Exon 12	Reverse
11	CTTCAATTGTGCGCTGTGATA	Exon 12	Reverse
12	TCATGCGAAGAGCAGAGAG	Exon 13	Forward
13	CGGATCCCGAGTCAGAAATA	Exon 13	Reverse
14	GGGACATGGAGGAATGTTTG	Exon 14	Forward
15	TAGCTCAATCAAATGAAGAT	Exon 14	Reverse
16	TGCACCTCTATGGGTTGCAC	Exon 15	Forward
17	GCTGAGAAGGCCGCCTTTGA	Exon 15	Reverse
18	CCCGGATTGGTCAGAGAAGT	Exon 16	Forward
19	ATGAAGCGATGCGTATT TCC	Exon 16	Reverse
20	CAACAGCCGGATAAAAC TGC	3′-UTR	Reverse
21	GCGGTTTTAGTTTTGATTAT	3′-UTR	Reverse
22	AACGTCCTGCTCGGTGAC	Exons 1 and 2	Forward
23	TGCCAAGGAATTTGTGA AAG	Exon 5	Reverse
24	ATGTGGATATCAGGAA GGA	Actin-U	Forward
25	CTCATACGGTCAGCAATAC	Actin-L	Reverse

[a]Primers 1–21 were used to amplify the genomic sequences and the complete cDNA from *HvSnf2P-2* (winter barley variety Dairokkaku). Primers 22–23 from rice *OsSnf2P* and primers 24–25 from the wheat actin gene were used for the expression studies

the RT-PCR, and the experiments were repeated twice. All samples were split in two and amplified with specific *Snf2P* primers and actin primers as an internal control (Table 1).

Two sets of conserved primers were selected from the barley and wheat *Snf2P* genomic sequence for the expression studies. Primers 14 and 17 (Table 1) were expected to amplify a 1217-bp fragment from barley genomic DNA and a 1228-bp product from wheat genomic DNA, but only a 266-bp fragment from cDNAs of both these species. Therefore, the size of the amplification products could be used to test for contamination of the mRNA sample with genomic DNA. Primers were designed from the non-conserved region upstream of the helicase motifs to increase their specificity. To confirm the results obtained with primers 14 and 17, a second set of primers was designed to amplify a 581-bp segment from exon 13 (primers 12 and 13, Table 1).

Results

The *Hordeum vulgare HvSnf2P* gene

An 18,750-bp contig from barley BAC 615K1 (spring variety Morex) included a complete *Snf2P* gene designated *HvSnf2P-1* (GenBank Accession No. AF459084). This contig included a 3-kb region upstream of the 5′ end of the gene, a 1-kb region downstream of the 3′ end, and extended for 14,750 bp from start to stop codon. A large 7731-bp non-LTR retrotransposon was detected in intron 6. This non-LTR retrotransposon was flanked by two 14-bp direct repeats of host DNA (GCAGTCC GCTCTTT) and included an ORF predicted to code for a reverse transcriptase that is 72% similar to a putative non-LTR retroelement present in Arabidopsis (GenBank Accession No. AP002071).

A second genomic sequence was obtained from the winter barley variety Dairokkaku for comparative purposes. A genomic contig of 7962 bp was sequenced from seven overlapping PCR products including 114 bp of the 5′ untranslated leader, 7090 bp of coding region and 758 bp of the 3′ untranslated region. The *Snf2P* gene in Dairokkaku barley was designed *HvSnf2P-2* (GenBank Accession No. AF459085). *HvSnf2P-1* and *HvSnf2P-2* genomic sequences were very similar in both the exon regions (99% identity) and the intron regions (97% identity). The large non-LTR retrotransposon found in *HvSnf2P-1* intron 6 was absent in *HvSnf2P-2* (Fig. 1).

No barley or wheat EST clones for the *Snf2P* gene were found in GenBank databases. Therefore, a complete *HvSnf2P-2* cDNA was sequenced to determine the *Snf2P* gene structure. Sequence from five overlapping RT-PCR products (2843 bp) included 2649 bp of coding region and 194 bp of the 3′ untranslated region. Comparison of the cDNA and genomic DNA sequences of Dairokkaku (100% identity) showed that the *HvSnf2P* gene had 17 exons and 16 introns (Fig. 1). One surprising aspect of the gene structure of both alleles was the presence of a very small exon (Exon 2: GCGACGAG) flanked by perfect AG-GT splicing sites.

The *T. monococcum TmSnf2P* gene

A 10,943-bp contig from *T. monococcum* (spring accession DV92) BAC 455C17 including a complete *TmSnf2P* gene was sequenced (GenBank Accession No. AF459088). This contig comprised 6,943 bp from start to stop codon, 3 kb of 5′ untranslated region and 1 kb of 3′ untranslated region.

Alignment of the wheat genomic sequence with the exons previously identified in barley revealed the existence of 17 exons of identical sizes to those found in barley. The predicted cDNA sequence from wheat was 96.1% identical to that of the barley cDNA. The small 8-bp exon 2 was also found in *TmSnf2P*, coding for amino acids identical to the ones predicted in barley.

Comparison of the 3-kb regions upstream of the ATG start codon in wheat and barley (BLASTN) identified an 85-bp region of high homology (88%). This homologous region was located immediately upstream

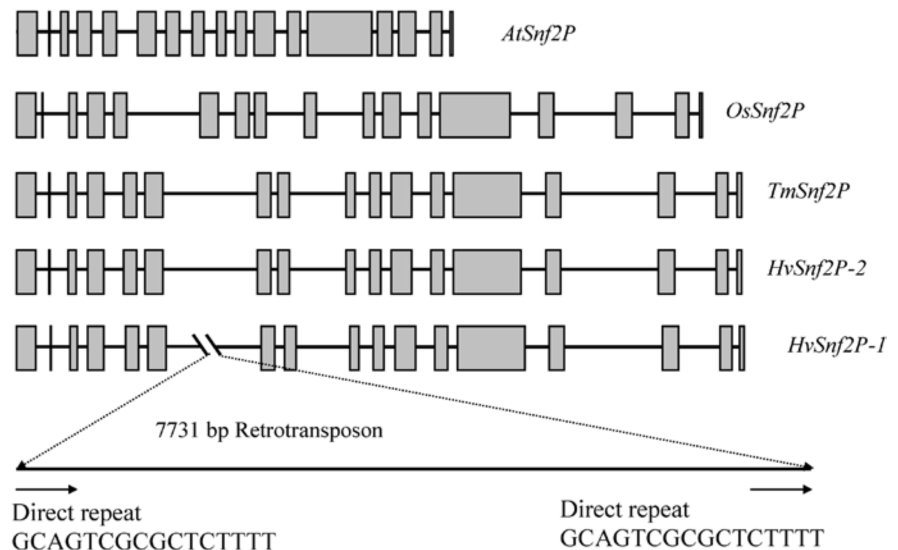


Fig. 1 Structure of plant *Snf2P* genes. The *filled boxes* represent exons. Sizes of exons and introns are proportional to their lengths in bp

from the start codon and might correspond partially or completely to the untranslated leader sequence. No additional conserved elements were found in this region. In the 3′ UTR region, identity between the *TmSnf2P* and *HvSnf2P* genomic sequences extended throughout the rest of the available genomic sequence (1 kb, average 93.9% identity).

The *Oryza sativa OsSnf2P* gene

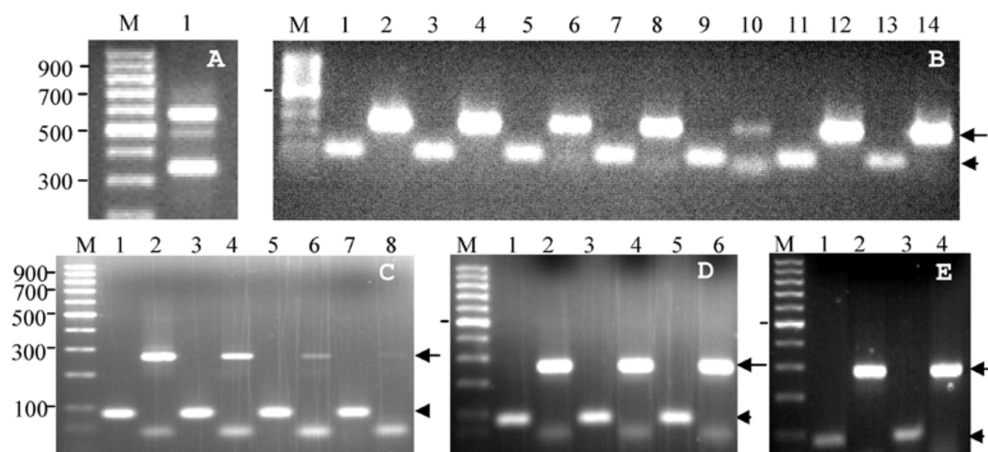
Four positive BACs were detected by screening Nippombare rice BAC libraries with the barley *Snf2P* clone. Based on contigs produced at the Clemson University Genomics Institute (http://www.genome.clemson.edu/projects/rice/fpc/), *Hin*dIII BAC 49F5 which included *Snf2P* in a central location, was sequenced (the complete BAC sequence without annotation was deposited in GenBank under Accession No. AF485811). An 8353-bp region from this BAC with an annotated *Snf2P* gene (*OsSnf2P*) was deposited in GenBank as Accession No. AF459086.

A BLASTN search of the EST database with the *OsSnf2P* genomic sequence detected the presence of a cDNA clone, C12105, with partial 5′ (457 bp, GenBank AU100698) and 3′ sequences [713 bp including poly(A)$^+$, GenBank C26323] that matched the first and last predicted exons of the genomic sequence perfectly (E values were e^{-109} and $2e^{-82}$, respectively). Dr. T. Sasaki kindly provided this cDNA, which we sequenced completely; the sequence was deposited in GenBank (Accession No. AF459087). The 3182-bp cDNA clone C12105 [with the exception of the 24-bp poly(A)$^+$ region] showed 100% identity to the sequence from the *OsSnf2P* gene present in rice *Hin*dIII BAC 49F5.

One region of the rice cDNA did not agree with the exon-intron structure of *OsSnf2P* predicted from the barley cDNA. Rice cDNA clone C12105 had a 232-bp segment (positions 205–436 from start codon) that showed no homology with Triticeae *Snf2P* genes. Translation of the complete 232-bp region would result in a premature stop codon. This 232-bp segment was flanked at the 5′ end by sequences similar to barley exons 1 and 2, and at the 3′ end by sequences similar to barley exon 3. Splicing sites GT-AG were observed at the ends of the 232-bp region, suggesting that it could be an intron that was not spliced out during mRNA processing.

To investigate this hypothesis, an RT-PCR experiment was carried out using one PCR primer that overlapped the last 13 bp of exon 1 and the first 5 bp of exon 2 (primer 22, Table 1) and a second primer from exon 5 (primer 23, Table 1). These two primers were 741 bp apart in the genomic sequence but, as expected from the design of the first primer, did not produce any amplification product from genomic DNA. Two bands of similar intensity were amplified in the RT-PCR, which was repeated twice (Fig. 2A). Rice cDNA clone C12105 hybridized with both bands. The smaller band had the expected size for a cDNA fragment without introns (347 bp), whereas the larger band had the expected size of a cDNA fragment including only intron 2 (579 bp) (Fig. 2A). Cloning and sequencing of the 347-bp band confirmed that the 232-bp intron was absent in these cDNAs and that two forms of *Snf2P* mRNAs were present in the mRNA population in rice leaves, one including intron 2 and the other lacking this intron.

Fig. 2A–E RT-PCR of mRNA from tissues of rice, wheat and barley. **A** RT-PCR of mRNA from rice leaves using primers 22 and 23. A 347-bp band is expected from a cDNA fragment without any intron and a 579-bp band is expected if intron 2 is present. **B–D** RT-PCR products amplified with *Snf2P* primers 14 and 17 (Table 1; *arrows*) and actin primers (*arrowheads*). **B** Leaves from *T. monococcum* accessions [lanes 1 and 2, DV92 (*vrn2*); 3 and 4, G2528 (*Vrn2*); 5 and 6, G3116 (*Vrn2*)] and *H. vulgare* varieties [lanes 7 and 8, Morex; 9 and 10, Iwate Mensury C; 11 and 12, Dairokkaku; 13 and 14, Hayakisu]. **C** Different tissues from barley variety Dairokkaku [lanes 1 and 2, vegetative apexes; 3 and 4, leaves; 5 and 6, sheaths; and 7 and 8, roots]. **D** RT-PCR products from different tissues of *T. monococcum* G3116 obtained using twice the amount of cDNA and five more PCR cycles than in Fig. 2C. Lanes 1 and 2, sheaths; 3 and 4, roots; and 5 and 6, spikes. **E** RT-PCR of mRNA from *T. monococcum* G3116 leaves obtained from non-vernalized (lanes 1 and 2) and vernalized (lanes 3 and 4) plants. Lanes M contained a 100-bp ladder (Promega). The 500-bp fragment is indicated by a *tick mark*



The *A. thaliana AtSnf2P* gene

Conceptual translation of the *AtSnf2P* gene (GenBank AC007659_19; see Lin et al. 2000) suggested the presence of 15 exons and 14 introns. However, information from the cereal *Snf2P* genes reported in this study can be integrated into the annotation of *AtSnf2P* to maximize the similarity between the monocotyledonous and dicotyledonous *Snf2* genes, as done before for other genes (Dubcovsky et al. 2001).

First, a sequence homeologous to the 8-bp exon 2 from cereal *Snf2* genes is present in the Arabidopsis genomic DNA (NC003071) 351–358 bp from the start codon. This predicted exon 2 had perfect GT-AG splicing sites and would encode the same amino acid sequence (GDE) as the exons described for the other members of the *SNF2P* subfamily (Figs. 1 and 3). Introduction of the small exon 2 in Arabidopsis, however, forces changes in the beginning of exon 3 to maintain the reading frame (the new AG site was 11 bp downstream of the annotated AG). Translation of the modified exon 3 showed that the first 11 amino acids were now identical to those observed in the cereal *Snf2P* proteins.

Another modification was the selection of an alternative GT splicing site for *AtSnf2P* intron 6 (previous intron five), 9 bp after the one currently annotated. This alternative annotation would result in an *AtSnf2P* exon 6 of identical size to those described in the three cereal species.

Finally, an alternative annotation at the end of *AtSnf2P* results in an additional exon (17). Translation of this additional exon predicted five amino acids identical to the ones described in rice (YYYRR, Figs. 1 and 3). A new GT splicing site was predicted 1 bp upstream of the annotated TAA stop codon followed by the addition of a new 25-bp exon that extended the sequence of *AtSnf2P* to 4327 bp from start to stop codon.

Comparative analysis of SNF2P proteins

The locations of the conserved helicase domains (Gorbalenya and Koonin 1993) are indicated in the alignment of the protein sequences in Fig. 3. The predicted barley HvSNF2P protein was 97% identical to the predicted wheat protein, 79% identical to the predicted rice protein OsSNF2P and 61% identical to the

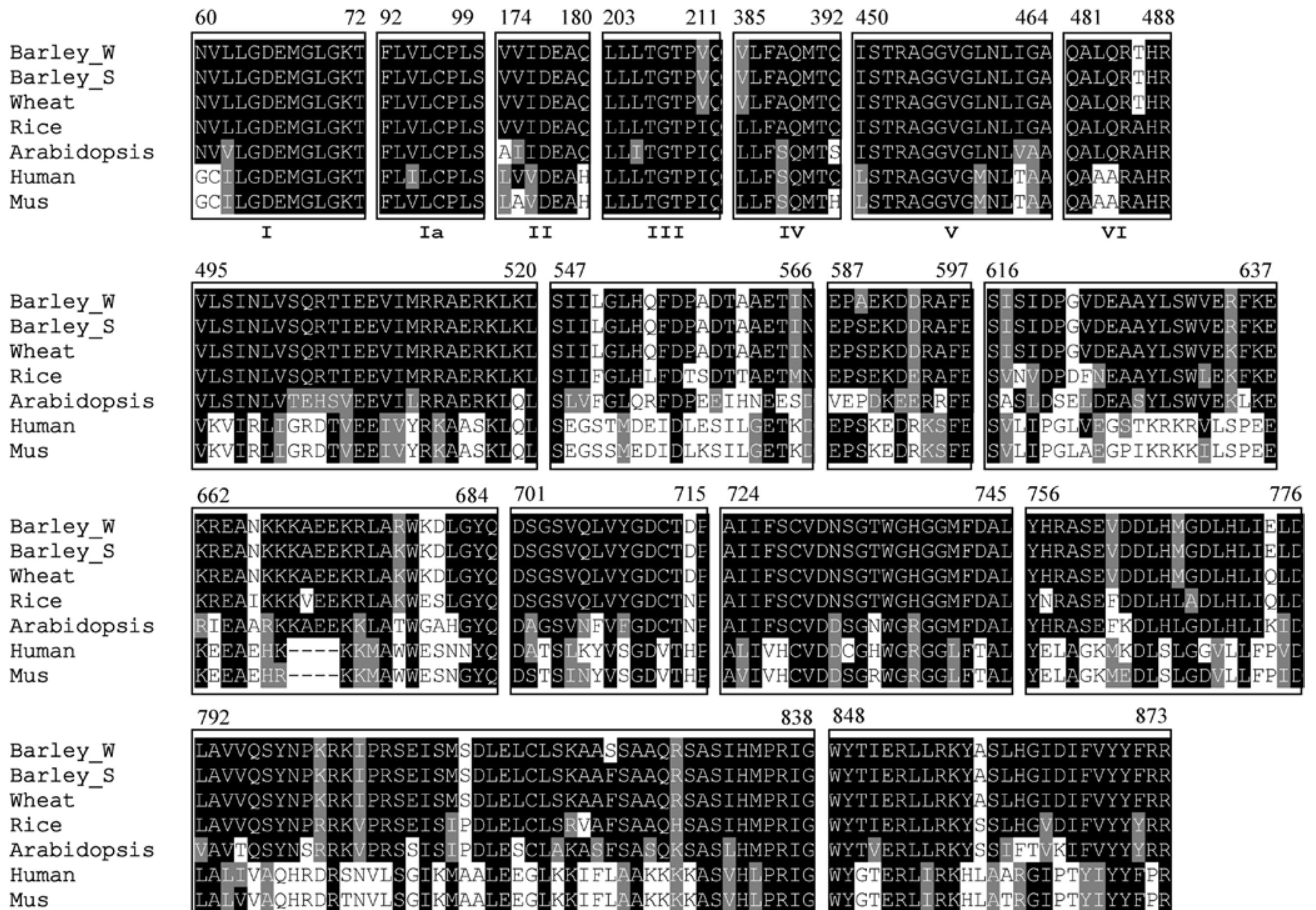


Fig. 3 Sequence alignment of SNF2P proteins from Arabidopsis (AC007659_19), rice (AF459086), barley (Morex: barley_S AF459084; Dairokkaku: barley_W AF459085), wheat (GenBank AF459088), human (AAH01171), and mouse (NP_080815). The *first row* includes the seven helicase motifs and the *last three rows* indicate the conserved regions in the C-terminal domain outside the helicase region

predicted AtSNF2P protein. Eleven of the 17 exons from barley were of identical length to the corresponding exons in rice and 10 to the corresponding exons in Arabidopsis. The eight splicing sites located within codons (at the end of exons 1, 3, 6, 7, 13, 14, 15, and 16) were conserved in all the *Snf2* genes analyzed from plants, indicating additional structural similarities (Fig. 1).

Sequence analysis of cereal SNF2P proteins using BLASTP (NCBI) showed significant similarities over the entire length of the conserved motifs SNF2_N (pfam 00176, SNF2 N-terminal domain, Expect = $2e^{-70}$) and helicase_C (pfam00271, helicase conserved C-terminal domain, Expect = $9e^{-15}$). These two conserved motifs spanned the first twelve exons of the plant SNF2P proteins and included the seven conserved helicase motifs (I, Ia, and II–VI) that are characteristic of helicase superfamily SF2 (Gorbalenya and Koonin 1993) (Fig. 3, numbered boxes).

Database searches using the last five exons outside the SNF2 domain showed significant similarities only with recently released human (AAH01171, *E* value = $1e^{-17}$) and mouse (NP_080815, *E* value = $3e^{-19}$) proteins. These proteins were not included in previous evolutionary studies of the SNF2 family (Coleman et al. 2000). No similarity was found in this region with members of the other 13 subfamilies, demonstrating that the SNF2P proteins constitute an independent subfamily. This was further confirmed by the construction of phylogenetic trees including the SNF2P proteins and members of the other subfamilies (see Materials and methods for a complete list of GenBank Accession Nos.).

The bootstrap consensus tree based on the Minimum Evolution method showed that the human HsSNF2P, the mouse MmSNF2P, and the plant SNF2P proteins formed a single cluster with a high bootstrap value (97) separated from the other subfamilies (Fig. 4). These proteins were also clustered in a separate group in the Neighbor-Joining (bootstrap = 98), UPGMA (bootstrap = 83), and Maximum Parsimony (bootstrap = 62) analyses, confirming that they belong to a distinctive subfamily.

The Minimum Evolution and UPGMA trees suggested that the *SNFP* subfamily was more closely related to subfamilies *SNF2*, *SNF2L*, *CHD1* and *YFK8* than to the other members of the SNF2 family (bootstrap values: ME = 80, UPGMA = 99). The Maximum Parsimony method also included subfamily *YSCD* in this group but with a lower bootstrap value (56). Finally, the

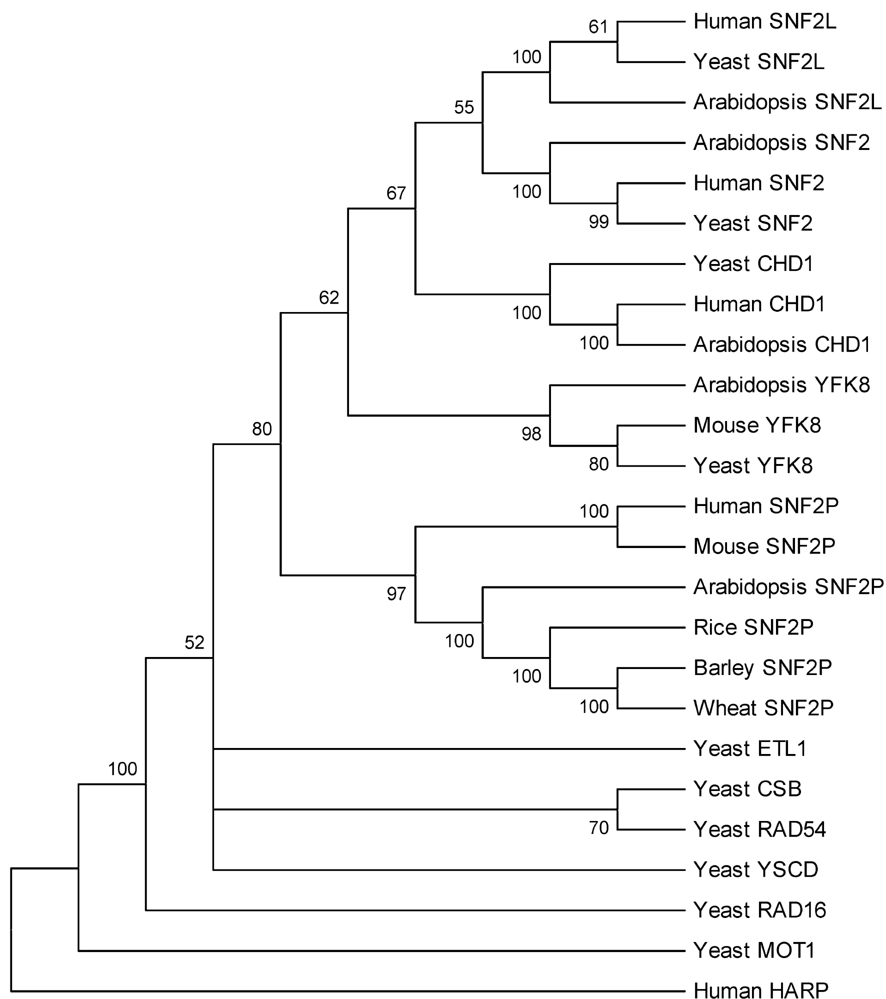


Fig. 4 Minimum Evolution consensus tree for members of the different subfamilies of the SNF2 family (see Materials and methods for GenBank Accession Nos.). Confidence values for the nodes are based on 500 bootstraps

Neighbor-Joining method included all subfamilies with the exception of *MOT1* and *HARP* in an unresolved basal node.

The new subfamily was designated *SNF2P* because of its similarity with the *SNF2* subfamily and "*P*" because it was first identified based on comparison of sequences from plants.

Genetic mapping of cereal *Snf2P* genes

The *Snf2P* genes were mapped on barley, wheat and rice genomes (Fig. 5) to determine their copy number and to explore putative linkage relationships with known traits or QTLs previously mapped in these species. Hybridization of Southern blots bearing genomic DNAs from wheat, barley and rice with probes amplified from two different segments of the *HvSnf2P* gene (primer pairs 14–17 and 12–13, Table 1) showed single RFLP fragments with most of the restriction enzymes analyzed. These results indicated that there is a single copy of the *Snf2P* gene in the genomes of each of these three species. This parallels the presence of a single copy in Arabidopsis and humans as determined by BLASTP searches of the Arabidopsis and Human databases. This gene was not found in the yeast genome.

Initially, the *TmSnf2P* gene was mapped in the 74-member F_2 population used to map the *Vrn2* gene in *T. monococcum* (Dubcovsky et al. 1998) (Fig. 5). In this population, *TmSnf2P* was completely linked to RFLP loci *Xucw2* and *Vrn2*, and was flanked by the distal locus *Xucw1(Nucellin)* (0.7 cM) and the proximal locus *Xwg199* (1.4 cM), in the distal region of the long arm of chromosome $5A^m$ (Dubcovsky et al. 1998).

The first 493 F_2 plants used for the high-density map were screened for recombination between *Xwg199* and *Xucw1(Nucellin)*. In the screen of the other 2282 F_2 plants, RFLP marker WG199 was replaced by UCW26, which was found to map closer to the *TmSnf2P*/*Vrn2* genes after the analysis of the first set of F_2 plants (Fig. 5). Plants were grown in a greenhouse without vernalization treatment (temperatures always above 15°C) and under long-day conditions. All plants homozygous at both flanking loci for the alleles from the *T. monococcum* accession with spring growth habit (DV92) flowered early, whereas plants heterozygous or homozygous for the alleles from the other genotype (G3116) at the same loci, did not flower under these conditions (winter growth habit). This result confirmed that the *Vrn2* gene was responsible for the observed segregation in growth habit in this mapping population, as observed before (Dubcovsky et al. 1998).

Nineteen recombinant plants were identified between the flanking markers and were further characterized for *Snf2P* and UCW2. Recombination between *Vrn2* (scored as a dominant trait) and the molecular markers was determined in the F_2 generation in those recombinant plants where one of the flanking markers was homozygous for the *T. monococcum* DV92 allele. Progeny tests including 15–20 F_3 plants were performed to determine the *Vrn2* genotype of the corresponding F_2 plant when one of the flanking markers was heterozygous and the other one was homozygous for the G3116 alleles.

One recombinant plant, F_2-#498, showed recombination between *XSnf2P* (homozygous DV92) and *Xucw26* (heterozygous) loci. This plant showed a clear winter growth habit indicative of a recombination event between the *TmSnf2P* and *Vrn2* genes. Locus *Xucw2,*

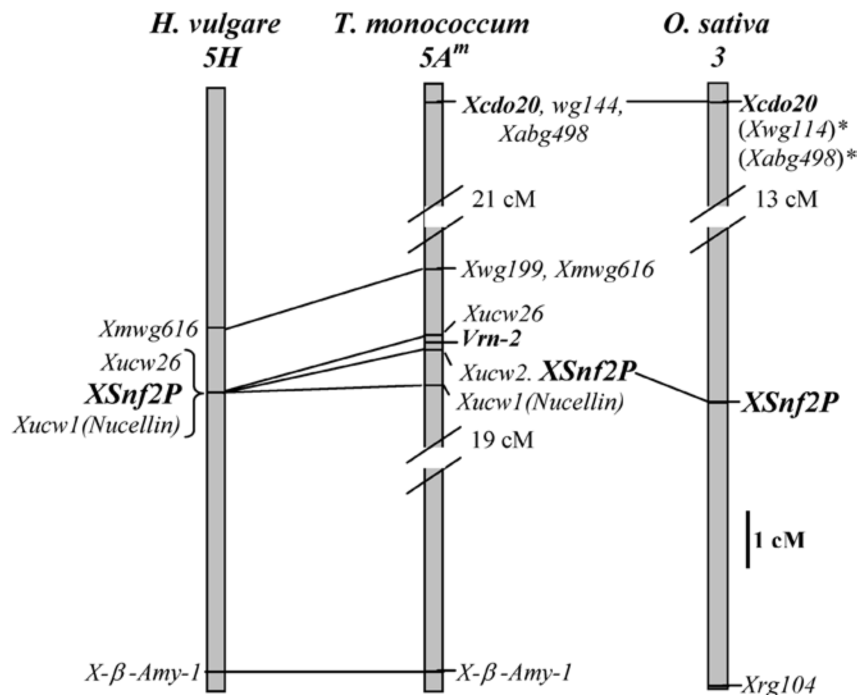


Fig. 5 Locations of *Snf2P* genes on the genetic maps of various cereal species. *Xwg114** and *Xabg498** were mapped in a different population (Saghai Maroof et al. 1996)

which was located 7 kb upstream from the 5′ end of *TmSnf2P* (L. Yan and J. Dubcovsky, unpublished results), was mapped in plant F_2-#498 and showed that the crossover occurred between *Xucw2* and *Vrn2*, at least 7 kb from the start codon of *TmSnf2P*. This result indicated that differences in the *TmSnf2P* gene were not the reason for the differences in growth habit observed between the spring and winter *T. monococcum* parental lines that were used in this cross.

In the smaller mapping population of barley plants, the *HvSnf2P* gene was found to be completely linked to *Xucw1(Nucellin)*, and *Xucw26* loci and lie between *Xmwg616* and *XβAmy* (Fig. 5). This region of barley chromosome 4H was previously known to be colinear with the distal region of *T. monococcum* chromosome $5A^{m}L$ (Devos et al. 1995; Dubcovsky et al. 1998).

In rice, the *OsSnf2P* gene was added to the map developed by Causse et al. (1994) and mapped on the distal region of chromosome 3 between the distal loci *Xcdo20* and *Xrg104* (Fig. 5). Rice *Hin*d III BAC 49F5 containing *OsSnf2P* gene was located by the Rice Genome Project at Clemson University within Contig 65 of the physical map, close to the distal end of chromosome 3 of rice (updated 10/20/01, http://www.genome.clemson.edu).

In previous studies, the most distal colinear marker between rice chromosome 3 and Triticeae homeologous group 4 was *Xcdo20* (Dubcovsky et al. 1998). The presence of the *Snf2P* gene distal to *Xcdo20* in wheat, barley and rice extends by 13 cM the colinearity between the distal region of rice chromosome 3 and the distal region of the long arm of Triticeae homeologous group 4 (Fig. 5).

Expression profiles of the genes *TmSnf2P* and *HvSnf2P*

The expression patterns of *Snf2P* genes in different tissues, genotypes and conditions were explored by RT-PCR in wheat and barley (Fig. 2B–E). In wheat, no differences in the intensity of the 266-bp band amplified with primers 14 and 17 were observed between the wheat genotypes carrying the *vrn2* (DV92; Fig. 2B, lane 2) and *Vrn2* (G3116, G2528) alleles (Fig. 2B, lanes 4 and 6). A similar result was obtained with primers 12 and 13 (Table 1, data not shown).

A lower amount of RT-PCR product was amplified with primers 14 and 17 from barley variety Iwate Mensury C compared with the other barley genotypes (Fig. 2B, lane 10). A similar result was obtained with primers 12 and 13 (Table 1), which yielded almost no product with Iwate Mensury C, and fragments of lower intensity from Morex compared with barley varieties Hayakiso and Dairokkaku. The actin control showed similar levels of expression in the different genotypes (Fig. 2B).

RT-PCR bands of the expected size for both pairs of primers were observed in mRNA extracted from apexes, leaves, sheaths, roots, and spikes in both barley (Fig. 2C) and wheat (Fig. 2D). Even though no quantitative estimates were made, the intensity of the observed RT-PCR fragments was consistently higher in apexes and leaves than in roots and sheaths, with both pairs of primers in both barley (Fig. 2C) and wheat (data not shown). The actin control showed similar levels of expression in all tissues. A second RT-PCR with double the amount of initial cDNA and incorporating five additional PCR cycles (40 cycles) confirmed the presence of *Snf2P* mRNA in roots, sheaths and spikes in wheat (Fig. 2D) and barley (data not shown).

No difference in the intensity of the RT-PCR products was observed between the samples extracted from leaves of vernalized (6 weeks, 8°C) and non-vernalized *T. monococcum* G3116 plants (Fig. 2E) or barley variety Dairokkaku. This result suggested that the expression of *Snf2P* was not significantly affected by the vernalization treatment.

Discussion

Characteristics of plant SNF2P proteins

The SNF2 family has been included in the large helicase SF2 superfamily based on the sequence of the seven conserved helicase-like motifs (Gorbalenya and Koonin 1993). It has recently been suggested that these helicase-related domains are actually motifs that permit the recognition of transition states between double-stranded and single-stranded DNA that are similar to the structures generated by helicases (Muthuswami et al. 2000). Such regions have been found in the processes of DNA metabolism that involve *SNF2* members as participants.

Motifs I, Ia, and II from the SNF2 family make up the nucleotide-binding site and are characterized by a phosphate-binding loop, often referred as the Walker A and B boxes (Koonin 1993). These two boxes are highly conserved among enzymes that hydrolyze ATP and GTP. Particularly well conserved is Motif II, which includes the signature amino acid sequence DExD/H, which is referred as the DEAD box.

All SNF2P proteins from plants analyzed so far have shown a distinctive glutamine (Q) amino acid at the end of the DEAD/H box (DEAQ). This change in the signature of Motif II differentiates plant proteins from the human and mouse members of the *SNF2P* subfamily and from most of the members of the SNF2 family (Bork and Koonin 1993). However, functional helicases from other groups, such as UvrD from *Escherichia coli* have a glutamine amino acid in this position (DEFQ) (Lohman 1993). This suggests that this glutamine in the DEAD box is compatible with an helicase function.

The conserved helicase motif I is encoded by a single exon in humans but by three different exons in plants. One of these three exons is the small 8-bp exon that codes for three amino acids that are conserved in all the members of the *SNF2P* subfamily. Small internal exons

have previously been described in muscle genes (N-CAM has a 3-bp exon). Usually, these small exons require special enhancing sequences in addition to strong splice sites for normal splicing (for a review, see Berget 1995).

It is interesting to point out that the intron that was not spliced out in some mRNAs from rice leaves was adjacent to the small exon 2. It would be interesting to investigate if there is a relationship between the unusually short length of exon 2 and the regulation of plant *Snf2P* genes via a regulated splicing mechanism. Another example of incomplete splicing in rice is the *Waxy* pre-mRNA, which shows normal splicing in varieties with high amylose content, but retains intron 1 in varieties that contain no amylose (Cai et al. 1998).

The gene prediction programs used to determine the structure of the *Snf2* gene in Arabidopsis failed to detect the small exon 2 (and also missed exon 17). Results from the comparison of cereal and Arabidopsis *Snf2* genes from this study provide an additional example of the use of comparative genomic studies as a tool to improve gene annotation in plants (Dubcovsky et al. 2001). These examples suggest that the Rice Genome Sequencing Project will provide valuable information to improve the annotation of the Arabidopsis genome.

Relationship between *Snf2P* and *Vrn2*

Plants carrying the dominant *Vrn2* allele require a long period at low temperature (vernalization) to induce flowering, whereas plants carrying the recessive *vrn2* mutation have a spring growth habit and do not require vernalization for flowering. The vernalization gene *Vrn2* has not yet been cloned, but genetic studies suggested that it is involved in the repression of genes in the flowering pathway (Tranquilli and Dubcovsky 1999). Based on the initial linkage between the genes *Vrn2* and *Snf2P*, and their related putative functions as transcriptional repressors, it was initially hypothesized that the *Snf2P* gene could be a potential candidate gene for *Vrn2*.

The discovery of one F_2 plant with a crossover between these two genes indicated that the differences in growth habit between *T. monococcum* accessions DV92 (spring) and G3116 (winter) were not causally related to putative differences in the sequences of *TmSnf2P* in these two accessions. However, this result does not rule out the possibility that a regulatory element located upstream from the *TmSnf2P* gene on the other side of the crossover event detected in plant #498 could be related to the observed differences in growth habit.

No mutations that could result in a premature stop codon or that could change substantially the conserved regions of the protein were found in the spring barley variety Morex or the spring *T. monococcum* accession DV92, both of which carry recessive *vrn2* alleles (Fig. 3). In addition, no differences were detected in the expression of *Snf2P* between *T. monococcum* genotypes carrying *Vrn2* (G3116) and *vrn2* (DV92) alleles in the RT-PCR experiments (Fig. 2). Finally, no differences were observed in *Snf2P* expression in leaves from winter wheat and barley plants before and after 6 weeks of vernalization. If *Snf2P* was involved in the repression of flowering genes and was regulated by vernalization, a lower (or null) level of expression was expected after the vernalization process.

All RT-PCR experiments were repeated with two different combinations of primers, included an internal actin control, and were performed with similar starting amounts of mRNA to standardize the RT-PCR conditions. However, since no detailed quantitative measures were performed in this study, it is not possible to rule out the possibility that there were small differences in *Snf2P* expression that were associated with variation at the *Vrn2* locus or the vernalization treatment, or that *Snf2P* is regulated at a posttranscriptional level.

RT-PCR products in the *vrn2* spring barley varieties Iwate Mensury C and Morex were of lower intensity than products amplified from similar samples from *Vrn2* winter barley varieties in three out of the four RT-PCR experiments. These differences were not observed in the RT-PCR experiments in which wheat genotypes with contrasting *Vrn2*/*vrn2* alleles were compared. Based on these two results, it is tempting to speculate that the lower levels of transcription of *HvSnf2* inferred in barley varieties Iwate Mensury C and Morex might be associated with the insertion of the large non-LTR retrotransposon element in intron 6 in these two varieties. This large intron is flanked by small exons (130–180 bp) and, theoretically, it would be possible to obtain normal splicing according to models that suggest an exonic recognition of splice sites ("exon definition"; Berget 1995). Testing the hypothesis of an effect of the non-LTR retrotransposon on the level of expression of *Snf2P* will require more sensitive quantitative measures to confirm the preliminary differences observed in this study, and the creation of isogenic lines to rule out epistatic effects from other genes present in these different varieties.

Based on the different results from these experiments the hypothesis that the *Snf2P* gene might be a candidate for the *Vrn2* gene was abandoned, and other candidate genes from the region are currently being evaluated.

Origin and function of *Snf2* proteins

The fact that the *Snf2P* genes from humans and plants are more similar to each other than these genes are to other members of the SNF2 family in each species (Fig. 4) indicates that the *SNF2P* subfamily arose prior to the divergence between animals and plants. The absence of proteins from the *SNF2P* subfamily in the proteomes of yeast and bacteria suggests that this subfamily originated in higher eukaryotic organisms.

SNF2P proteins from plants and animals showed high levels of similarity both within the SNF2 domain and in the C-terminal amino acids coded by the last five exons, suggesting the possibility that they serve similar functions. The conservation of blocks of amino acids between the plant and animal proteins (Fig. 3) should be useful in helping to delimit the functionally important domains of these proteins. Although no evidence is currently available for the function of these genes, the fact that the products of this subfamily are more similar to those of the *SNF2*, *SNF2L*, *CHD1* and *YFK8/DDM1* subfamilies may indicate some related function. However, this needs to be demonstrated experimentally since the observed similarities between these subfamilies and the members of the *SNF2P* subfamily were restricted to the conserved SNF2 domain.

The *Snf2P* transcripts were found in all the tissues tested, suggesting that the proteins they encode are involved in normal cellular functions or housekeeping activities such as transcriptional regulation. Many of the other genes from the related *SNF2* subfamily that regulate transcription also show ubiquitous patterns of gene expression (Coleman et al. 2000).

The experimental determination of the *Snf2P* gene structure in plants, the identification of conserved regions between SNF2P proteins from plants and animals, and the preliminary characterization of the expression profile of the plant *Snf2P* genes described in this study will provide the basis for future analysis of the function of this new subfamily of transcription factors.

Acknowledgments This work was supported by the United States Department of Agriculture, National Research Initiative, Grant 2000-1678 (PI Dr. J. Dubcovsky), and by the National Science Foundation Grant 9975793 (PI Dr. J. Bennetzen). The authors express their gratitude to A. Kleinhofs, M. Sorrells, and A. Graner for supplying clones, to A. Kleinhofs for the seeds of the Morex × *H. sopontaneoum* cross, and to S. McCouch for her help with the rice mapping population.

References

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389–3402

Berget SM (1995) Exon recognition in vertebrate splicing. J Biol Chem 270:2411–2414

Bork P, Koonin EV (1993) An expanding family of helicases within the DEAD/H superfamily. Nucleic Acids Res 21:751–752

Cai XL, Wang ZY, Xing YY, Zhang JL, Hong MM (1998) Aberrant splicing of intron 1 leads to the heterogeneous 5′ UTR and decreased expression of *waxy* gene in rice cultivars of intermediate amylose content. Plant J 14:459–465

Causse MA, Fulton TM, Cho YG, Ahn SN, Chunwongse J, Wu K, Xiao J, Yu Z, Ronald PC, Harrington SE, Second G, McCouch S, Tanksley SD (1994) Saturated molecular map of the rice genome based on an interspecific backcross population. Genetics 138:1251–1274

Chen FQ, Foolad MR (1997) Molecular organization of a gene in barley which encodes a protein similar to aspartic protease and its specific expression in nucellar cells during degeneration. Plant Mol Biol 35:821–831

Coleman MA, Eisen JA, Mohrenweiser HW (2000) Cloning and characterization of HARP/SMARCAL1: a prokaryotic HepA-related SNF2 helicase protein from human and mouse. Genomics 65:274–282

Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor Appl Genet 91:282–288

Dubcovsky J (2001) Plant gene cloning may lead to better timing of flowering. NRI Research Highlights. USDA-CSREES

Dubcovsky J, Galvez AF, Dvorak J (1994) Comparison of the genetic organization of the early salt stress response gene system in salt-tolerant *Lophopyrum elongatum* and salt-sensitive wheat. Theor Appl Genetics 87:957–964

Dubcovsky J, Luo M-C, Zhong G-Y, Bransteiter R, Desai A, Kilian A, Kleinhofs A, Dvorak J (1996) Genetic map of diploid wheat, *Triticum monococcum* L., and its comparison with maps of *Hordeum vulgare* L. Genetics 143:983–999

Dubcovsky J, Lijavetzky D, Appendino L, Tranquilli G (1998) Comparative RFLP mapping of *Triticum monococcum* genes controlling vernalization requirement. Theor Appl Genet 97:968–975

Dubcovsky J, Ramakrishna W, SanMiguel P, Busso C, Yan L, Shiloff B, Bennetzen J (2001) Comparative sequence analysis of colinear barley and rice BACs. Plant Physiol 125:1342–1353

Eisen JA, Sweder KS, Hanawalt PC (1995) Evolution of the Snf2 family of proteins: subfamilies with distinct sequences and functions. Nucleic Acids Res 23:2715–2723

Gavin IM, Simpson RT (1997) Interplay of yeast global transcriptional regulators Ssn6p-Tup1p and Swi-Snf and their effect on chromatin structure. EMBO J 16:6263–6271

Gorbalenya AE, Koonin EV (1993) Helicases – amino acid sequence comparisons and structure function relationships. Curr Opin Struct Biol 3:419–429

Jeddeloh JA, Stokes TL, Richards EJ (1999) Maintenance of genomic methylation requires a SW12/SNF2-like protein. Nat Genet 22:94–97

Kaneko T, Katoh T, Sato S, Nakamura Y, Asamizu E, Tabata S (2000) Structural analysis of *Arabidopsis thaliana* chromosome 3. II. Sequence features of the 4,251,695-bp regions covered by 90 P1, TAC and BAC clones. DNA Res 7:217–221

Kingston RE, Bunker CA, Imbalzano AN (1996) Repression and activation by multiprotein complexes that alter chromatin structure. Genes Dev 10:905–920

Koonin EV (1993) A common set of conserved motifs in a vast variety of putative nucleic acid-dependent ATPases including MCM proteins involved in the initiation of eukaryotic DNA replication. Nucleic Acids Res 21:2541–2547

Kumar S, Tamura K, Nei M (1994) MEGA: molecular evolutionary Genetics Analysis Software for Microcomputers. Comput Appl Biosci 10:189–191

Lijavetzky D, Muzzi G, Wicker T, Keller B, Wing R, Dubcovsky J (1999) Construction and characterization of a bacterial artificial chromosome (BAC) library for the A genome of wheat. Genome 42:1176–1182

Lin XY, et al (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–768

Lohman TM (1993) Helicase-catalyzed DNA unwinding. J Biol Chem 268:2269–2272

McIntosh RA, Hart GE, Devos KM, Gale MD, Rogers WJ (1998) Catalogue of gene symbols for wheat. In: Proc 9th Int Wheat Genetics Symp Saskatchewan Canada 2–7 August 1998, pp 119–120

Muthuswami R, Truman PA, Mesner LD, Hockensmith JW (2000) A eukaryotic SWI2/SNF2 domain, an exquisite detector of double-stranded to single-stranded DNA transition elements. J Biol Chem 275:7648–7655

Pazin MJ, Kadonaga JT (1997) SW12/SNF2 and related proteins: ATP-driven motors that disrupt protein-DNA interactions? Cell 88:737–740

Ruhf ML, Braun A, Papoulas O, Tamkun JW, Randsholt N, Meister M (2001) The *domino* gene of Drosophila encodes novel

members of the SW12/SNF2 family of DNA-dependent ATPases, which contribute to the silencing of homeotic genes. Development 128:1429–1441

Saghai Maroof MA, Yang GP, Biyashev RM, Maughan PJ, Zhang Q (1996) Analysis of the barley and rice genomes by comparative RFLP linkage mapping. Theor Appl Genet 92:541–551

Salanoubat M, et al (2000) Sequence and analysis of chromosome 3 of the plant *Arabidopsis thaliana*. Nature 408:820–822

Sudarsanam P, Iyer VR, Brown PO, Winston F (2000) Whole-genome expression analysis of *snf/swi* mutants of *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 97:3364–3369

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Tranquilli GE, Dubcovsky J (1999) Epistatic interactions between vernalization genes *Vrn-A*^{m}*1* and *Vrn-A*^{m}*2* in diploid wheat. J Hered 91.:304–306

Wu L, Winston F (1997) Evidence that Snf-Swi controls chromatin structure over both the TATA and UAS regions of the *SUC2* promoter in *Saccharomyces cerevisiae*. Nucleic Acids Research 25:4230–4234

Yu Y, Tomkins JP, Waugh R, Frisch DA, Kudrna D, Kleinhofs A, Brueggeman RS, Muehlbauer GJ, Wing RA (2000) A bacterial artificial chromosome library for barley (*Hordeum vulgare*). Theor Appl Genet 101:1093–1099