4632–4638 Nucleic Acids Research, 1993, Vol. 21, No. 19

Snol, a novel alternatively spliced isoform of the *ski* proto-oncogene homolog, *sno*

Sonia Pearson-White
Departments of Pediatrics and Microbiology, University of Virginia Medical Center, MR-4 Building Box 1131, Charlottesville, VA 22908, USA

Received January 20, 1993; Revised and Accepted July 30, 1993 EMBL accession no. Z19588

ABSTRACT

We have cloned and sequenced a novel human isoform of *sno*, *snoI* for *I*nsertion. *SnoI* contains 1330 nucleotides inserted in place of 7 nucleotides of the *snoN* mRNA. *Sno* is a member of the *ski* proto-oncogene family, which has been implicated in muscle development. The two previously known *sno* alternatively spliced isoforms are *snoN* (684 amino acids), and *snoA* (415 amino acids); *snoI* encodes a truncated isoform of 399 amino acids (44,298 MW). Southern blot experiments show that *snoI* contains a third alternative exon from the *sno* gene; a single *sno* gene can express all three isoforms of *sno* by alternative splicing. All three isoforms contain the region that is most similar to the *ski* proto-oncogene. The relationship between *snoI* and *snoN* is analogous to that between ΔfosB and fosB, where a truncated form of the fosB transcription factor is produced by alternative splicing. We find conservation of human *snoI*-specific sequences in several mammalian species, in monkey, dog, cow, rabbit and pig, but not in rodents, whereas the common portion of the *sno* gene is conserved in all vertebrate species tested. *SnoN*, *snoA*, and *ski* mRNAs accumulate in many human tissues including skeletal muscle; the *snoI* alternative mRNA accumulates more specifically in skeletal muscle. *SnoI* is also expressed in rhabdomyosarcoma tumor, a tumor that contains differentiated skeletal muscle. The tissue-specific alternative splicing of human *snoI*, an mRNA in the *ski/sno* gene family, and the presence of *sno* mRNAs in muscle are consistent with a proposed role for the *sno* oncogene in muscle gene regulation.

INTRODUCTION

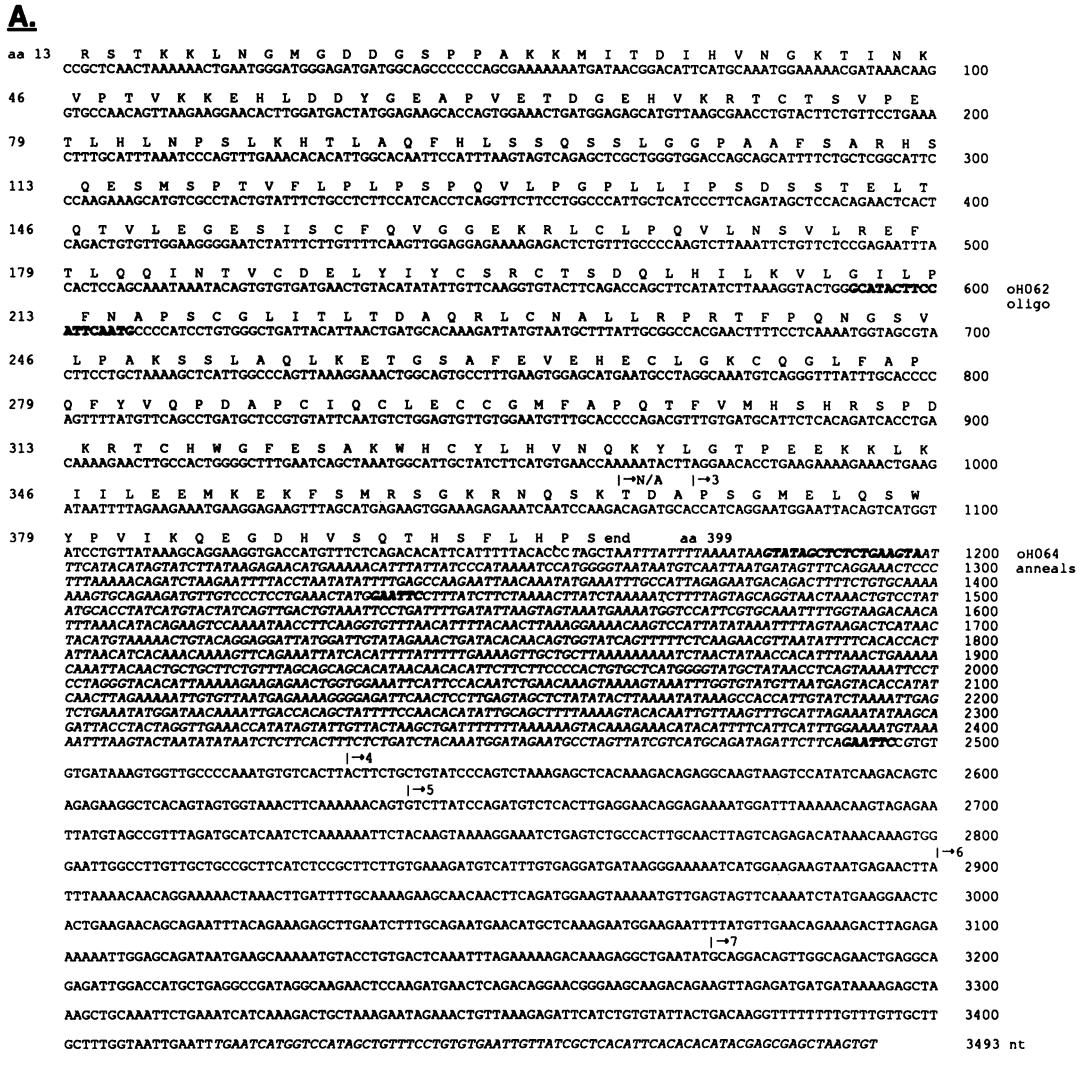
Two families of transcription factors have been implicated in muscle development, the MyoD family and the *ski/sno* proto-oncogene family. Members of the MyoD transcription factor family have been shown to play key roles in determination of the skeletal muscle lineage. The MyoD family is the muscle specific subfamily of the larger helix-loop-helix (bHLH) family and includes MyoD (1), myogenin (2, 3), myf5 (4), and MRF4/herculin/myf6 (5–7) in vertebrates. When any of these cDNAs are expressed in non-muscle cells of various types, the cells can be converted to determined myogenic cells that are capable of differentiating into skeletal muscle. These myogenic regulators are expressed in developing embryos in the myotomal compartment of the somites where cells become committed to the myogenic lineage (8–10).

Members of the *ski* family of putative transcription factors have also been implicated in muscle development; *ski* can convert non-muscle quail embryo cells to muscle (11). The v-*ski* oncogene (named for *S*loan-*K*ettering *I*nstitute) was isolated from Sloan-Kettering virus (SKV) transformed cells (12–14). The c-*ski* proto-oncoprotein is found in the nucleus, but is not a member of any of the major transcription factor gene families, such as bHLH, zinc finger, leucine zipper, homeobox or *rel*/NF-KB, although a weak similarity to *myc* has been suggested (15, 16). Since *ski* lacks the bHLH domain that is crucial for the myogenic determination activity of MyoD family proteins, it may represent a class of myogenic determination gene that acts via a different mechanism, the elucidation of which should offer insight into the myogenic determination process.

The involvement of *ski* in skeletal muscle development was first shown in avian cells, where non-myogenic cells are converted to the myogenic lineage by transfection with chicken v-*ski* or c-*ski* (11). After *ski* transfection endogenous quail MyoD (qmf1 (8)) and myogenin (qmf2 (9)) and muscle-specific gene expression, usually silent in these primary quail fibroblasts, is activated (17). A mutant *ski*, tdM5i, fails to convert the cells to muscle, but retains its ability to transactivate the endogenous MyoD and myogenin genes (17). This provides genetic evidence that the activation of MyoD and myogenin is insufficient to convert the cells to muscle in the presence of mutant *ski* protein. Thus *ski* plays a role in myogenic conversion and may function in myogenesis as part of the same pathway as MyoD, interacting directly or indirectly with the myogenic regulatory factors. Transgenic mice overexpressing a truncated cloned chicken c-*ski* cDNA in their skeletal muscles exhibit a marked enlargement of these muscles and diminished body fat (18). The enlarged skeletal muscle phenotype suggests that *ski* affects skeletal muscle development and/or maturation.

A related human gene *sno* (for *s*ki-related *no*vel) was cloned by low stringency hybridization of the chicken v-*ski* probe to human cDNA libraries by ourselves and others (19). Recently *sno*, like *ski*, has been shown to convert non-muscle quail embryo cells to skeletal muscle (20). Human *sno* and *ski* share a 106

amino acid region in the N-terminal half of each protein that is 82% identical; the two proteins are 37% identical throughout their length. By analogy with relationships within other such closely related gene families, *sno* and *ski* may share similar functions. Two isoforms of *sno*, *snoN* and *snoA*, have been previously cloned from several different human cDNA libraries (19). These isoforms differ by inclusion of different mutually exclusively spliced alternative exons A (Alu-containing) or N (non-Alu-containing). This alternative splicing pattern is common in muscle-specific genes, but is rarely found in other tissues. *SnoA* encodes a truncated, 415 amino acid protein, whereas *snoN* encodes a 684 amino acid protein. All but the last 49 amino acid residues of *snoA*, which are encoded by the alternative exon (19), are found in *snoN*. The tissue distribution of *snoN* and *snoA* (and



snoI) expression has not been described (19). We have identified a third *sno* isoform, *snoI*, and have examined the expression of all three alternatively spliced *sno* isoforms and *ski* in human tissues. We find that *snoN*, *snoA*, and *ski* are expressed in many tissues, including skeletal muscle, and that the *snoI* isoform is alternatively spliced more specifically in skeletal muscle. These results are consistent with a suggested role for *sno* in muscle development.

MATERIAL AND METHODS

Cell culture, RNA isolation

Primary human myoblasts were cultured from a normal muscle biopsy obtained during surgery on a 2 year old patient. Several clonal cell lines were isolated from these cultures. The myoblasts from the HMP2 clone used for library synthesis were proliferative and undifferentiated, but capable of differentiating into phenotypically normal striated muscle under appropriate culture conditions (i.e. growth factor depletion of the medium). RNA was isolated from subconfluent HMP2 myoblast cultures by a modification of the guanidinium thiocyanate procedure (21, 22). Poly(A^+) RNA was selected by chromatography on oligo(dT) cellulose using PL Biochemicals Type 7 (dT)-cellulose.

Library synthesis and screening

Poly (A^+) RNA isolated from HMP2 human myoblasts was used to make the cDNA library. HMP2 human myoblasts are a clone of primary human skeletal muscle myoblasts that can differentiate well in culture; HMP2 cells were obtained from Dr Mark Lovell at the University of Virginia. Double-stranded cDNA was synthesized by the method of Gubler and Hoffman (23), using RT-XL AMV reverse transcriptase (Life Sciences Inc., St. Petersburg, FL) for first strand cDNA synthesis. The cDNA was treated with EcoRI methylase (New England Biolabs) and EcoRI linkers were used for ligation to λgt10 phage arms (Stratagene); phage were packaged using Gigapack Gold extract (Stratagene). From 3.5 μg of poly(A^+) RNA 2.5×10^6 recombinants were isolated; the library was amplified once prior to screening. To isolate human *ski* homologs, 100,000 plaque forming units (pfu) of recombinant phage were plated at a density of 20,000 pfu per 150 mm plate. Duplicate plaque lifts on nitrocellulose filters (Schleicher and Schuell) were prehybridized in 6×SSC, 10×Denhardts (0.2% each of ficoll 400, polyvinylpyrrolidone, bovine serum albumin) for 5 hours, and hybridized in 1 M NaCl, 50 mM Tris pH 7.4, 2 mM EDTA, 10X Denhardts, 50 μg/ml sonicated, denatured salmon sperm DNA, 2.5 μg/ml each poly A,U, I, and C, and 0.5% SDS at 48°C for 60 hours (24). The filters were hybridized with random-hexamer-primed chicken v-*ski* oncogene pCCLski5 probe, generously provided by Ed Stavnezer (13, 14). Filters were washed at 48°C as described (24), plus a final 0.3×SSC wash at 48°C. Two duplicate positives were rescreened and plaque purified; one was the 3493 base pair (bp) *snoI* described here, the other was a 5300 bp human c-*ski* cDNA (ski81).

DNA sequencing

SnoI EcoRI fragments were subcloned in pEMBL 18^+ (a derivative of pEMBL 8 containing the pUC18 polylinker, provided by S.Lazarowitz) and single stranded DNA primed by M13 -20 primer or specific oligonucleotides was sequenced by the dideoxy chain termination method (25, 26) Sequences were read manually or with an IBI gel reader, and contigs were built using the Sequence Assembly Manager (SAM) program (27).

Filter hybridization

20 μg of genomic DNA, purified by standard methods (28, 29), was digested with restriction enzyme, purified by phenol and chloroform extraction, ethanol precipitated, then redigested and reextracted again. Samples were electrophoresed on horizontal 1% agarose gels. DNA fragments were capillary transferred (30) to Hybond-N (Amersham) in 20×SSC, then UV-irradiated at 500 μWatts/cm^2 for 1 minute. DNA fragments for hybridization probes were purified on agarose gels and labeled with random-hexamer-priming, using a Boehringer Mannheim kit, and ≥3000 Ci/mmol [α^{32}P]-dCTP (DuPont-New England Nuclear). The blots were hybridized and washed as described (31), with the substitution of 100 μg/ml *E.coli* DNA (Sigma) for salmon sperm DNA.

Reverse transcription-polymerase chain reaction (RT-PCR)

1 μg of total RNA from the indicated human tissue, tumor, or leukemic blood or bone marrow sample was reverse transcribed using random hexamer primers and AMV reverse transcriptase (Boehringer Mannheim). The first strand cDNA was then amplified using Vent DNA polymerase (New England Biolabs). The polymerase chain reaction (PCR) reaction conditions were

Figure 1. A. Nucleotide sequence of human *snoI*. The nucleotides are numbered at the right of each line and the amino acids at the left. The snoI cDNA begins at amino acid #13; the twelve additional amino acids that are encoded in *snoN* but missing from the 5′ end of the *snoI* cDNA are MENLQTNFSLVQ. The nucleotides (#1165–2497, #3417–3493) that differ from *snoN* are italicized. The EcoRI sites (*GAATTC*) that separate *snoI*(X), *snoI*(I), and *snoI*(Z) are highlighted. The synthetic EcoRI linkers at the beginning and end of the clone are not shown. The PCR primer oligonucleotide sequences are highlighted with bold lettering and identified in the right margin. The exon boundaries determined for the chicken *ski* gene are indicated with arrows and the exon number; we do not find the alternatively spliced exon 2 in mammalian *sno*, but have kept the numbering the same for consistency (32). The *snoI* nucleotide sequence is EMBL accession number Z19588; *snoN* is S06052; *snoA* is S06054. **B.** The *snoI*/*snoN* divergence region is shown, aligning the *snoI* and *snoN* nucleotide sequence identities flanking the *snoI* insertion. The *snoI* insertion of 1330 nucleotides is diagrammed schematically below the aligned region; it is contiguous with it at both ends. In the aligned comparison region, *snoI* nucleotide numbers are below the sequence line and *snoN* nucleotide numbers are above. They differ because *snoI* begins at nucleotide #746 of *snoN*. |→4 indicates the position of the exon 3/4 boundary in chicken c-*ski* (32). The EcoRI subfragments X, I and Z are indicated, and the EcoRI sites that separate them are drawn in the schematic diagram. Thus, the *snoI*(I) probe is specific for the insertion region of *snoI*, and *snoI*(X) and *snoI*(Z) hybridize to sequences common to all *sno* isoforms. Note that the *snoI*(X) probe (of 1436 bp) includes 276 bp of the *snoI*-specific insertion region. *SnoI*(Z) extends from the second EcoRI site at *snoI* nt # 2490, and is identical to *snoN*, except that it is 75 nucleotides longer (this region is shown in Figure 1A, italicized). **C.** The *snoA*, *snoN* and *snoI* amino acid sequences are aligned and presented from residue #360 (7 amino acid residues prior to the point of divergence) past the ends of *snoA* (at 415 residues) and *snoI* (at 399 residues). The *snoA* nucleotide sequence (not shown) differs from *snoN* starting at *snoN* nucleotide #1807. *SnoA* encodes a termination codon in this region, predicting the translation of a truncated protein of 415 amino acids with the C-terminus shown. *SnoI* also introduces a termination codon in this region, predicting the translation of a truncated protein of 399 amino acid residues, whereas *snoN* protein is 684 amino acid residues in length. |→°N/A indicates the first nucleotide where *snoA* differs from *snoN*. |→°4 indicates the position of the exon 3/4 boundary in chicken c-*ski* (32).

95°C for 2′, then 95°C 1′, 55°C 1′ and 72°C 1′ for 34 cycles, taking 12 μl samples (out of 100 μl total reaction volume) after 17 and 34 cycles. Samples were run on 4% Nusieve:agarose 3:1 gels (FMC), alkaline vacuum blotted to Hybond N+ (Amersham), and probed as described above. The following oligonucleotide primers (Operon) were used at 0.4 μM each:

oligo name	orientation	identity	length	sequence, 5′ to 3′
oH062	sense	snoI #590–607	18	GCATACTTCCATTCAATG
oH064	antisense	snoI-sp. #1180–1198	18	TACTTCAGAGAGCTATAC
oH065	antisense	snoN-sp. #1834–1854	20	ACATGTATAAGTAGTAGCTG
oH059	antisense	snoA-sp.	21	CATACAGGTTAAGTTACACAC
oH050	sense	ski #3138–3161	24	GACTGAGGCTGCAGCATTGGAACA
oH051	antisense	ski #3466–3491	26	CAGATTCAGTCGCTACTGAAGACAGA

RESULTS

Cloning and nucleotide sequence of *snoI*

We cloned a human *ski* homolog from a myoblast cDNA library using the chicken v-*ski* probe (16). We isolated a cDNA for human *ski* as well as a novel isoform of the related *sno* proto-oncogene (19). Figure 1A presents the nucleotide and translated amino acid sequence of the novel isoform of *sno*, which we call *snoI* (I for insertion). *SnoI* is identical to *snoN* except for a 1330 bp insertion region (italicized), ten out of seventeen *snoN* nucleotides differing from *snoI* at the site of the insertion (Figure 1B), and several apparent polymorphisms. The structure of the *snoI* insertion junction into *snoN* is shown in Figure 1B; the *snoI* insertion encodes an in-frame termination codon, which results in a 399 amino acid protein (44,298 MW), in contrast to the 684 amino acid product (76,998 MW) for *snoN*.

Sno is likely to share the intron/exon structure of the homologous chicken *ski* gene; the location of putative exons is indicated in Figure 1A. Exon 2 of *ski* is combinatorially spliced; it is absent from some chicken c-*ski* mRNAs; chicken v-*ski* contains exon 2 (32, 33). We have not found an exon 2 in mammalian *sno*, but we have numbered the exons in Figure 1A with reference to *ski*. *SnoN* and *snoA* diverge from eachother 9 nucleotides 5′ of the *ski* 1/3 exon junction, suggesting that this exon boundary is shifted slightly in *sno*. We do not find donor or acceptor splice boundaries near the sites where *snoN* and *snoI* diverge, and *snoI* is missing sequences contained in *snoN* (Figure 1B, at the site of divergence); thus, *snoI* is not an intron-containing version of *snoN*.

The *snoI* cDNA is 75 nucleotides longer than *snoN* at the 3′ end (Figure 1A). Neither cDNA has a polyadenylation site, and the mRNA size (9 kilobases (kb), Figure 6) is much longer than either cDNA. Two sequence differences noted in the *snoN*-identical regions of *snoI* are: *snoI* nt# 78 C is T in *snoN*, changing *snoN* valine to the similar alanine, and *snoI* nt#2733 A is T in *snoN*. This second change is after the *snoI* termination codon, but would predict lysine instead of asparagine in the *snoN* protein. These may be polymorphic differences between the two individuals whose DNA was sequenced, or sequencing errors or differences between the two laboratories. Near both ends of the 1330 nucleotide insertion, at *snoI* nucleotides #1159 and #2497, we also see different nucleotides (Figure 1B); these differences may have arisen after the *snoN* and *snoI* alternative exons diverged.

Figure 1C summarizes the three different predicted primary amino acid sequences in the *snoN/A/I* alternative region. Both *snoI* and *snoA* encode termination codons in this region. With the isolation of the new *snoI* isoform, we know that there are three alternative isoforms of *sno*, two of which predict proteins truncated in this region.

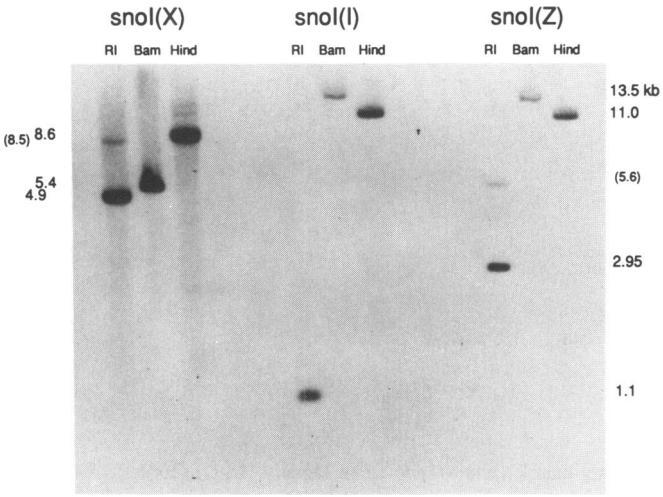


Figure 2. Human genomic southern blot analysis of the *snoI* molecule. 20 μg aliquots of human genomic DNA was digested to completion with RI (EcoRI), Bam (BamHI) or Hind (HindIII), electrophoresed on a 1% agarose gel, and capillary blotted to Hybond-N (Amersham). Identical blots were separately probed with subcloned *snoI* EcoRI subfragments X, I and Z under high stringency conditions (hybridized and washed at 65°C (31)), then exposed to XAR-5 film at −70°C with a Cronex intensifying screen for 3 days. *SnoI*(I) and *snoI*(Z) hybridizing band sizes are given to the right of the figure, and *snoI*(X) bands are listed to the left of the figure in kilobases (kb).

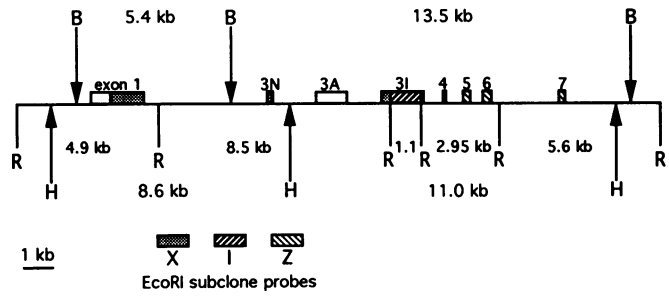


Figure 3. The human *sno* gene. This provisional map was derived from cDNA and genomic Southern blot data, plus published chicken c-*ski* exon boundaries (32). The fragment sizes (Figure 2) are written between each corresponding restriction enzyme designation (R for EcoRI, B for BamHI, or H for HindIII). No double digest sizes are presented, nor have they been determined, thus, the exact positions of the exons within each restriction fragment are not known. We have no evidence for chicken c-*ski* exon 2 in human *sno*, but have retained the chicken gene exon numbers, omitting exon 2. The regions of the exons that would hybridize with the X, I, and Z cDNA subclone probes are shaded to correspond to the key and the shading in Figure 1B.

Alternative splicing generates *snoI* from the same gene as *snoN*

To confirm that the *snoI* insertion is part of the same *sno* gene that encodes *snoN* and *snoA*, we carried out the genomic Southern blot experiment shown in Figure 2. The *snoI* cDNA insert can be separated into three fragments by digestion with EcoRI, *snoI*(X), *snoI*(I), and *snoI*(Z) (Figure 1A, 1B). *SnoI*(X) and *snoI*(Z) include sequences that can hybridize to all *sno* mRNAs, *snoN*, *snoA* and *snoI*. However, the last 276 nucleotides of the 1436 nucleotide *snoI*(X) probe are specific for the *snoI* isoform. The *snoI*(I) probe (1054 nucleotides) is specific for the *snoI* insertion. EcoRI, BamHI, and HindIII restriction digests of human genomic DNA were blotted in triplicate to membrane filters and probed with *snoI*(X), *snoI*(I), and *snoI*(Z) (Figure 2). *SnoI*(I) and *snoI*(Z) hybridize to identical 13.5 kb BamHI and 11.0 kb HindIII bands, showing that the *snoI* insertion is encoded in the same gene as *snoN*, and confirming that *snoI* is an alternatively spliced version of *sno*. The *snoI*(X) cDNA probe spans more than one exon, and hybridizes predominantly to a 4.9 kb EcoRI band, and to a lesser extent to an 8.5 kb band (Figure 2, *snoI*(X) EcoRI lane, and Figure 4, human lane). The 276 nucleotide portion of *snoI*(X) that is *snoI*-specific, from nt # 1158−1436, hybridizes faintly to the 11.0 kb HindIII band (Figure 2, *snoI*(X) HindIII panel), seen in the *snoI*(I) and *snoI*(Z) panels. (Figure 2, I and Z HindIII panels). Thus, the *snoI* insertion is part of the same *sno* gene as the previously identified *snoN*, and *sno* is encoded by a single gene in the human genome.

The genomic restriction digestion data, together with the exon structure from the chicken *ski* gene, are consistent with the provisional map of the gene shown in Figure 3 and have ruled out several other possible maps. The exact positions of the exons within each restriction fragment are not known, nor is the position of the 3A alternative exon relative to 3N and 3I. We have sought but not found evidence for the presence of an exon 2, as is found in chicken c-*ski* and v-*ski*. We cannot exclude the possibility that 'exon 3' is subdivided into two or more exons so that the *snoN/snoI*-identical region is a common exon, and the 17 nucleotide region of *snoN* that is *snoN*-specific and the difference region including the 1330 nucleotide insertion present only in *snoI* are alternative 4th exons; the remainder of 'exon 3' where *snoN* and *snoI* are again identical would then be a common exon 5. Our data are also consistent with the map shown (Figure 3), which diagrams a single human *sno* gene that preserves the *ski*

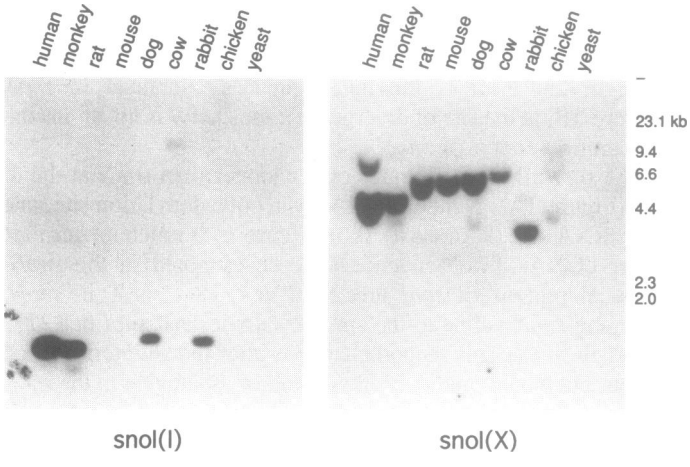


Figure 4. Conservation of the *sno* gene demonstrated by genomic Southern blot. The blot (Clontech Laboratories) contains 8 μg per lane of EcoRI-digested genomic DNA isolated from the indicated organisms separated on a 0.7% agarose gel, capillary blotted to a charge-modified nylon membrane (Nytran), and fixed by UV irradiation. This blot was probed with the *snoI*(I)-specific probe, then stripped and reprobed with the *snoI*(X) probe that hybridizes with every isoform of *sno*. The blot was hybridized at 60C, and washed at 55°C as described (31).

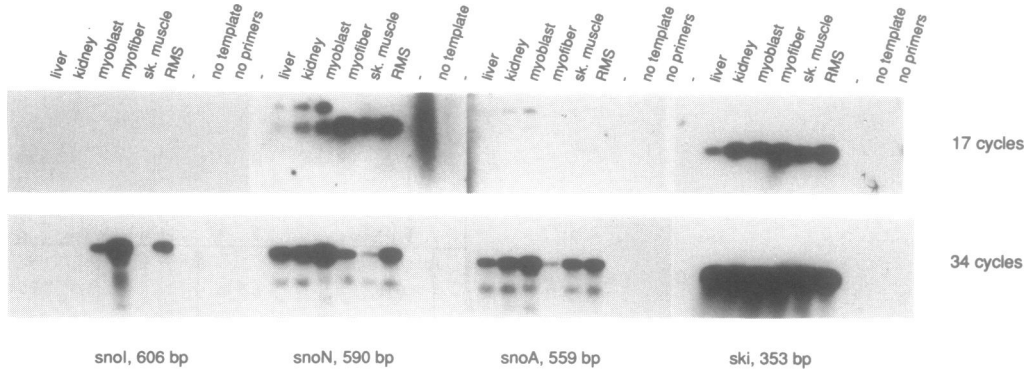


Figure 5. RT-PCR analysis of *snoN/A/I* and *ski* mRNA expression in human tissues, primary muscle cultures, and tumor samples. 1 μg of total RNA from the indicated human source was reverse transcribed, then subjected to the polymerase chain reaction, taking 12 μl samples out of 100 μl total reaction volume for gel electrophoresis after 17 and 34 cycles as indicated (18 and 34 cycles, respectively for *snoI*). Myoblasts and myofibers are HMP2 primary human myoblasts and differentiated myofibers in tissue culture. RMS is a rhabdomyosarcoma tumor sample. Sk. muscle is an adult skeletal muscle biopsy. The *snoA* 'no primers' control pertains to both *snoN* and *snoA* experiments, which were done together. The oligonucleotide sequences are highlighted in Figures 1A and 1B and presented in the Methods. The oH062 primer is identical to nucleotides #590−607 of *snoI*, which is in the region common to all *sno* transcripts, and was used for the *snoN, I,* and *A* experiments. The oH064 3′ primer is complementary to nucleotides #1180−1198 of *snoI*, and gives a 606 bp *snoI*-specific product. The oH065 3′ primer is from the *snoI* insertion junction (Figure 1B), and gives a 590 bp *snoN*-specific PCR product. The oH059 primer gives a 559 bp *snoA*-specific PCR product. 12 μl samples were run on a 4% Nusieve:agarose 3:1 gel (FMC), alkaline vacuum blotted to Hybond N^+ (Amersham), and probed with the *snoI*(X) cDNA probe that will hybridize to all *sno* PCR products, or the ski81 human *ski* cDNA (S. Pearson-White, unpublished) for the *ski* panel. Only the relevant parts of the autoradiograms are shown. Exposure times were 18 hours for the *snoI* 18 and *ski* 17 cycle panels, and 8 days for the *snoN* and *snoA* 17 cycle panels. Exposure times were 6 hours for the first 5 lanes (liver through sk. muscle) and 18 hours for the last 4 lanes (RMS=rhabdomyosarcoma through 'no primers') for the *snoI* 36 cycle panel. Exposures were 1.3 hours for *snoN* and *snoA* 34 cycle panels and 15 minutes for the *ski* 34 cycle panel. All exposures used one intensifying screen and XAR-5 film.

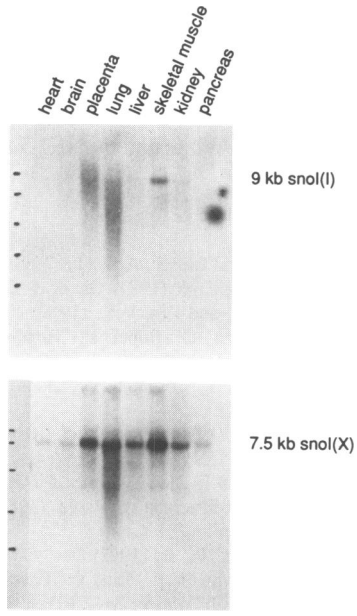


Figure 6. Northern blot analysis of *sno* and *snoI* accumulation in tissue samples. Duplicate Human Multiple Tissue Northern blots (Clontech), containing 2 μg of poly A^+ RNA per lane from the indicated tissues, were probed with random-prime-labeled *snoI*(I), a probe specific for the *snoI* insertion, or *snoI*(X), a probe that will hybridize with every variant of *sno*. The marks down the left side of each panel indicate the positions of RNA size markers, which are at 9.5, 7.5, 4.4, 2.4, and 1.35 kb. The blot was washed at high stringency (65°C in 40 mM sodium phosphate) as described (31). Exposures were overnight to XAR-5 film with one intensifying screen.

exon structure, and in which mutually exclusive alternative splicing of a *snoI* alternative exon # 3 (3I) in place of either 3N or 3A alternatives explains the differences between the *snoI* and *snoN* structures. In this model exons 3I and 3N contain regions of identity between the I and N versions, flanking an exon-specific region in which 3I has three different and 1330 inserted nucleotides and 3N has ten out of 17 *snoN*-specific nucleotides.

The *sno* gene can be found in a variety of mammalian species and in chicken (Figure 4, *snoI*(X) panel), as well as in quail and the amphibian *Xenopus laevis* (data not shown). . The multiple bands in Figure 4 reflect the structure of a single *sno* gene, rather than multiple genes; restriction digestion was done using EcoRI, which cleaves between genomic fragments hybridizing with the cDNA probe. The *snoI*-specific alternative exon is found in most mammalian species (monkey, dog, rabbit, cow, Figure 4) and pig (data not shown), but we have not been able to detect this exon in mouse or rat genomic DNA (Figure 4 and data not shown). We have probed genomic blots made from DNA from different species digested with several different restriction enzymes to rule out the possibility that *snoI*(I) hybridization is not detectable in rodents because the relevant EcoRI fragments are too short or too long to transfer to these blots. Thus, while the *sno* gene is present in all mammalian species examined (as well as in avian and amphibian), the *snoI*-specific insert is absent from rodents.

Sno expression

Because *sno* is a rare transcript and is alternatively spliced, we have used Reverse Transcription-Polymerase Chain Reaction (RT-PCR) to analyze accumulation of specific *sno* mRNA transcripts in various human tissues. We process samples after only 17 (or 18) cycles of PCR, as well as after 34 (or 36) cycles, to estimate qualitatively the relative abundance of RNAs in different tissues. The 5′ oligonucleotide oH062 (Figure 1A, and Methods) was used to measure expression of all three isoforms. Exon-specific 3′ antisense oligonucleotides were used to examine *sno*N/A/I; oH064 is specific for *snoI* (Figure 1A), oH065 is specific for *snoN* (Figure 1B), while oH059 is specific for *snoA* (not diagrammed, (19)). We have also used primers oH050 and oH051 to probe human *ski* gene expression. The oligonucleotide sequences are presented in the Methods.

Using RT-PCR, we detect *snoN*, *snoA*, and *ski* accumulation in every human tissue tested; however, *snoI* is more muscle-specific (Figure 5). RT-PCR reactions with human G3PDH primers confirmed that reverse transcribed RNA samples were potent and roughly equal in cDNA amount (data not shown). *SnoN*, *snoA* and *ski* 34 cycle bands were readily visible on ethidium staining of the gels, using 12 μl out of 100 μl total reaction volume, whereas the *snoI* band was not. We note that the exposure time for the *snoI* panel was 4.5 (myoblast and myofiber) and 13 (RMS) times longer than for the 34 cycle *snoN* and *snoA* panels. Exposure times for the *ski* panels were 5−8 times shorter than for *sno* panels, indicating that *ski* mRNA is expressed at higher levels than *sno* mRNAs. *SnoA* mRNA is less abundant than *snoN*, but can be readily detected after 34 cycles with RT-PCR. *SnoA* is difficult to detect after 17 cycles, and other workers were not able to detect *snoA* on Northern blots (19). *SnoN*, *snoA*, and *ski* are expressed in every human tissue or tumor sample tested (34 cycle panels, Figure 5). *SnoI* expression is detectable only in cultured primary myoblasts and myofibers and rhabdomyosarcoma (RMS) tumor (which contains skeletal muscle) after 36 cycles of amplification. (*SnoI* is also detectable in these tissues after 34 cycles, but the data shown used 36 cycles). *SnoI* expression has been detected in skeletal muscle biopsy tissue by RT-PCR (not in the data shown), but is lower than in cultured primary myoblasts and myofibers. Longer exposure obscures the skeletal muscle lane due to the robust *snoI* signal from differentiated myofibers.

The 17 cycle *snoN* PCR products migrate as doublet bands (Figure 5 *snoN* panel). The primers span the putative exon 2 location, so we considered the possibility that the doublets reflect the insertion of exon 2, adding 112 nt to some transcripts. However, the upper band, not the lower band, is the predicted size; the two band sizes differ by 100 nt. Therefore, our data, including many repetitions of this experiment not shown, provide no evidence for the presence of exon 2 in the transcripts. We have no explanation for the presence of the 100 nt shorter band in the 17 cycle *snoN* PCR reactions.

We have also examined *sno* expression with Northern blot analysis (Figure 6). As before, *snoI* expression is limited to skeletal muscle. *SnoI*(I) the *snoI*-specific probe, hybridizes with a 9 kb mRNA in skeletal muscle, but not with the other tissue RNAs. *SnoI*(X), a probe that hybridizes with all *sno* transcripts, hybridizes predominantly with a 7.5 kb mRNA in every tissue. The highest *sno* expression is found in skeletal muscle, followed by placenta and lung; the lowest expression is seen in heart, brain, and pancreas (Figure 6). The *snoI*(X) panel shows a faint hybridization signal to a 9 kb mRNA in skeletal muscle. The difference in size between the 9 kb and 7.5 kb transcripts can be approximately accounted for by the 1.33 kb insertion into *snoI*. In this size range of the blot, small differences in mobility greatly

affect the estimated length; this may explain why we reproducibly measure the major *sno* band size at 7.5 kb while other workers measured a 6.2 kb size (19). Thus, these data confirm that *sno* gene expression is widespread, accumulating *snoN* and/or *snoA* mRNA species in many tissues, but expression of the alternatively spliced *snoI* isoform is more specific to skeletal muscle.

DISCUSSION

We have cloned and sequenced a novel c-*sno* proto-oncogene cDNA, *snoI*, from human myoblasts. The *snoI* mRNA is transcribed from a single *sno* gene that also encodes the *snoN* alternatively spliced isoform. The *snoI* structure has revealed a third alternatively spliced isoform expressed from the *sno* gene, in addition to the previously described *snoN* and *snoA* isoforms (19). Mutually exclusive alternative splicing to include the *snoI* alternative exon would encode a truncated protein with respect to *snoN*. The invariant sequences shared by all three isoforms include the region most highly conserved between *ski* and *sno*. To date, very few genes have been described with three alternative exon choices. An example of a gene with this number of mutually exclusive alternatively spliced exons is the *Drosophila melanogaster* muscle myosin heavy-chain gene (34). Drosophila muscle myosin heavy-chain has two to five tandemly repeated and then diverged copies of five sets of exons, which are individually spliced in a mutually exclusive manner in the mature mRNAs (34).

Until we have determined the exon structure of the human *sno* gene, we cannot distinguish between the possibility that nucleotides 2494–2533 of *snoI*, which are identical to *snoN* and are located at the 3′ end of the *snoI* insertion before the position of the chicken *ski* exon 3/4 junction, come from an additional exon, and the possibility that they are part of alternatively spliced exon #3.

We have examined expression of all three alternative *sno* isoforms and *ski* in human tissues. *SnoN*, *snoA*, and *ski* are expressed in all tissues examined. *SnoI* is found more specifically in myoblasts and myofibers derived from a biopsy of skeletal muscle, and in skeletal muscle-containing rhabdomyosarcoma tumor. Regulated tissue-specific expression of this isoform suggests that it may have an important function in skeletal muscle.

It is intriguing that the *snoI* isoform is less well conserved than the *sno* gene itself, with *snoI* conserved in many mammalian species but lacking in rodents and birds. Rodents may produce a truncated form of *snoN* by another mechanism, such as a different alternative exon. A gain-of-function experiment to ectopically express human *snoI* in transgenic mouse muscle and evaluate the consequences on muscle development is in progress.

The relationship between *snoN* and *snoI* is analogous to that between FosB and ΔFosB, where production of a truncated form of a transcription factor occurs by alternative splicing (35). ΔFosB is co-expressed with FosB, although the ΔFosB increase after serum stimulation persists longer than FosB. ΔFosB exhibits several dominant negative effects, and *snoI* may function similarly to inhibit *snoN* activity in muscle. Unlike ΔFosB and FosB, however, *snoI* expression is regulated differently than *snoN*. *SnoI* accumulation is limited to skeletal muscle, induced in myoblast and myofiber explant cultures and the skeletal muscle-containing tumor rhabdomyosarcoma, while *snoN* and *ski* accumulation are more ubiquitous. The muscle-specificity of *snoI* points to a potential role in myogenesis; experiments are in progress to examine the roles of *snoI* and *snoN in vitro*.

ACKNOWLEDGEMENTS

I thank Bill Pearson for help with the GCG software package and reading the manuscript. This work was supported by American Cancer Society Institutional award # IN-149F and NIH RO1 HD27202 to S.Pearson-White and by NIH 5 PO1 DK40031.

REFERENCES

1. Davis, R.L., Weintraub, H. and Lassar, A.B. (1987) *Cell*, **51**, 987–1000.
2. Edmondson, D.G. and Olson, E.N. (1989) *Genes & Dev.*, **3**, 628–640.
3. Wright, W.E., Sassoon, D.A. and Lin, V.K. (1989) *Cell*, **56**, 607–617.
4. Braun, T., Buschhausen-Denker, G., Bober, E., Tannich, E. and Arnold, H.H. (1989) *EMBO J.*, **8**, 701–709.
5. Braun, T., Bober, E., Winter, B., Rosenthal, N. and Arnold, H.H. (1990) *EMBO J.*, **9**, 821–831.
6. Miner, J.H. and Wold, B. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 1089–1093.
7. Rhodes, S.J. and Konieczny, S.F. (1989) *Genes & Dev.*, **3**, 2050–2061.
8. Charles de la Brousse, F. and Emerson, C.P., Jr. (1989) *Genes & Dev.*, **4**, 567–581.
9. Pownall, M.E. and Emerson, C.P., Jr. (1992) *Dev. Biology*, **151**, 67–79.
10. Sassoon, D., Lyons, G., Wright, W.E., Lin, V., Lassar, A., Weintraub, H. and Buckingham, M. (1989) *Nature*, **341**, 303–307.
11. Colmenares, C. and Stavnezer, E. (1989) *Cell*, **59**, 293–303.
12. Stavnezer, E., Gerhard, D.S., Binari, R.C. and Balazs, I. (1981) *J. Virol.*, **39**, 920–934.
13. Stavnezer, E., Barkas, A.E., Brennan, L.A., Brodeur, D. and Li, Y. (1986) *J. Virol.*, **57**, 1073–1083.
14. Li, Y., Turck, M., Teumer, J.K. and Stavnezer, E. (1986) *J. Virol.*, **57**, 1965–1972.
15. Barkas, A.E., Brodeur, D. and Stavnezer, E. (1986) *Virology*, **151**, 131–138.
16. Stavnezer, E., Brodeur, D. and Brennan, L.A. (1989) *Mol. Cell. Biol.*, **9**, 4038–45.
17. Colmenares, C., Teumer, J.K. and Stavnezer, E. (1991) *Mol. Cell. Biol.*, **11**, 1167–1170.
18. Sutrave, P., Kelly, A.M. and Hughes, S.H. (1990) *Genes & Dev.*, **4**, 1462–72.
19. Nomura, N., Sasamoto, S., Ishii, S., Date, T., Matsui, M. and Ishizaki, R. (1989) *Nucleic Acids Res.*, **17**, 5489–5500.
20. Boyer, P.L., Colmenares, C., Stavnezer, E. and Hughes, S.H. (1993) *Oncogene*, **8**, 457–466.
21. Linzer, D.I.H. and Nathans, D. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4271–4275.
22. Chirgwin, J.M., Przybla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochem.*, **18**, 5294–5299.
23. Gubler, U. and Hoffman, B.J. (1983) *Gene*, **24**, 263–269.
24. Peden, K.W.C., Mounts, P. and Hayward, G.S. (1982) *Cell*, **31**, 71–80.
25. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
26. Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.
27. Lawrence, C.B., Shalom, T. and Honda, S. (1989) Molecular Biology Information Resource, Baylor College of Medicine
28. Blin, N. and Stafford, D.W. (1976) *Nucleic Acids Res.*, **3**, 2303–2307.
29. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor,
30. Southern, E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
31. Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
32. Sutrave, P. and Hughes, S.H. (1989) *Mol. Cell. Biol.*, **9**, 4046–51.
33. Grimes, H.L., Szente, B.E. and Goodenow, M.M. (1992) *Nuc. Acids Res.*, **20**, 1511–1516.
34. George, E.L., Ober, M.B. and C. P. Emerson, J. (1989) *Mol. Cell. Biol.*, **9**, 2957–2974.
35. Nakabeppu, Y. and Nathans, D. (1991) *Cell*, **64**, 751–759.