Volume 17 Number 14 1989 Nucleic Acids Research

# Isolation of human cDNA clones of *ski* and the *ski*-related gene, *sno*

Nobuo Nomura*, Shigemi Sasamoto, Shunsuke Ishii[1], Takayasu Date[2], Minami Matsui and Ryotaro Ishizaki

Molecular Oncology Laboratory, Nippon Veterinary and Zootechnical College, Sakuragi, 1-10-19 Uenosakuragi, Taito-ku, Tokyo 110, [1]Laboratory of Physical and Chemical Research, RIKEN, 3-1-1 Koya-dai, Tsukuba, Ibaraki 305 and [2]Division of Cancer Research, Medical Research Institute, Kanazawa Medical University, Uchinada, Ishikawa 920-02, Japan

Received May 8, 1989; Revised and Accepted June 26, 1989 EMBL accession nos X15217–X15219 (incl.)

## ABSTRACT

cDNA clones of *ski* and the *ski*-related gene, *sno*, were obtained by screening human cDNA libraries. The predicted open reading frame of h-*ski* could encode a protein of 728 amino acid residues. The h-*ski* protein is highly homologous with the v-*ski* protein. The overall homology between h-*ski* and v-*ski* is 91% at the amino acid level. DNA sequencing analysis revealed two types of cDNA clones from the *sno* (*s*ki-related *no*vel gene) gene, possibly due to alternative splicing. The first type, named *sno*N (*n*on Alu-containing), encoded a protein of 684 amino acid residues. The second type, named *sno*A (*A*lu-containing), encoded a protein of 415 amino acid residues. The first 366 amino acid residues of *sno*N and *sno*A are the same, but subsequent amino acids show divergence. Several transcripts of h-*ski* (6.0, 4.7, 3.8, 3.0, 2.1 and 1.8kb) were detected. The mRNAs of h-*sno* were 6.2, 4.4 and 3.2kb.

## INTRODUCTION

Sloan-Kettering viruses (SKVs) are acute transforming retroviruses that were isolated from cultured chicken embryo cells (CECs) (1–3). SKVs were shown to induce focus formation in monolayer cultures of CECs and colony formation in soft agar. They did not transform bone marrow cells *in vitro*. The SKVs isolated were shown to contain viral oncogene v-*ski* which is different from known oncogenes. Three novel *gag*-containing polyproteins $p125^{gag\text{-}ski}$, $p110^{gag\text{-}pol\text{-}ski}$ and $p55^{gag\text{-}ski}$ were synthesized in SKV-infected CECs and were shown to be located in the nucleus (4).

Gene families of nuclear oncogenes (5–13), of steroid and thyroid hormone receptor genes (14,15) and of protein-tyrosine kinases (16) have been reported. Each gene family harbors a conserved region which should encode proteins with common function. To determine the function of *ski* in transformation and cell growth, we used the strategy of isolating a human cDNA clone and looking for a gene(s) related to *ski*. Here we report the isolation and characterization of cDNA clones of the human *ski* gene and the *ski*-related gene, *sno*.

## MATERIALS AND METHODS

*Cells:*

Cell lines were obtained from a neuroblastoma (NB39-nu, NB1), Burkitt lymphoma (JBL-2, JBL-5), and carcinomas of the vulva (A431), the stomach (NMS92), the chorion (BeWo, GCH1), the thyroid (TC78), the lung (NMS83) and the prostate (PC3). NB39-nu and NB1 were from T.Suzuki (Fukushima Medical College), JBL-2 and JBL-5 were from I.Miyoshi (Kohchi Medical College), BeWo and GCH1 were from S.Sasaki (Nippon Medical School),

PC3 was from Y.Nakagami (Nippon Medical School), TC78 was from H.Ohami (Nippon Medical School), and A431 was from the Japanese Cancer Research Resources Bank.

*cDNA library:*

The human cDNA libraries used in this work were generously provided by J.E.Sadler [a λgt11 cDNA library from mRNA of endothelial cells from umbilical vein (17)], P.Chambon [λgt11, breast cancer cell line MCF-7 (18)], J.R.de Wet [λgt11, hepatoma cell line Li-7 (19)], S.L.C.Woo [λgt11, liver (20)], D.P.Dialynas [λgt10, T cell line HPB-MLT (21)], W.L.Miller [λgt10, adrenal (22)] and J.M.Puck [λgt11, peripheral blood lymphocytes (23)]. A human placenta cDNA library constructed in λgt11 phage vector was purchased from Clontech Lab., Inc.(Palo Alto, CA, U.S.A.).

*Screening of the cDNA library:*

A 1.15kb *Sst*I fragment of pCCL ski5 (3), 1.1kb *Eco*RI fragment of λ-ski1 (this work) and 1.4kb *Eco*RI fragment of λ-sno3 (this work) were random-primed with [ $\alpha$-$^{32}$P]dCTP (3000Ci/mmol) to a specific activity of $2\times10^6$cpm/ng. Hybridization was performed in a solution containing either 30% (relaxed condition) or 50% (stringent condition) formamide, 5×SSC, 0.5%SDS, 5×Denhardt's solution, 100μg/ml of sonicated salmon testis DNA and $^{32}$P-labeled probe ($2\times10^6$cpm/ml) at 37°C for 16 hours. After several washing in 1×SSC, 0.1%SDS at room temperature, filters were finally washed in 0.1×SSC, 0.5%SDS either at 35°C (relaxed condition) or 50°C (stringent condition) for 1 hour.

*DNA sequencing:*

Relevant DNA fragments were isolated from phage clones by digestion with restriction endonucleases and were cloned into M13mp11, M13mp18 and pUC18 (24). Some sequencing was performed by subcloning appropriate restriction fragments into M13mp11 and M13mp18. Sequence analysis was carried out by the dideoxynucleotide chain terminator method with modification (7-deaza dGTP instead of dGTP) (25,26). All sequences were confirmed by analyzing at least two overlapping λ phage clones. Both DNA strands were sequenced. When polymorphism was observed, another clone(s) was sequenced.

*Computer Analysis:*

Homology studies and other computer analyses were carried out with the UWGCG (27) and IDEAS (28) programs in a VAX/VMS computer (Institute of Medical Science, Tokyo University).

*Northern and Southern Blot Analyses:*

Cytoplasmic RNA from cell lines (29) was passed over oligo (dT)-cellulose. The glyoxylated poly(A)$^+$RNA (3μg) was fractionated on 0.7% agarose gel and transferred to a Biodyne A filter (Pall, New York, U.S.A.) (30). Genomic DNAs were digested with *Eco*RI (Takara Shuzo, Kyoto, JAPAN), electrophoresed in 0.7% agarose gel, treated, and blotted onto a Biodyne A filter essentially as described by Southern (31).

## RESULTS

*Isolation of cDNA clones of the ski and ski-related gene*

To obtain cDNA clones of the human-*ski* gene and the *ski*-related gene, we screened eight kinds of cDNA libraries with a v-*ski* probe, a 1.15kb *Sst*I fragment of pCCL ski5 (3), under conditions of reduced stringency. On screening $3\times10^5$ phages of each library, four positive clones were obtained. Physical mapping and partial sequence analysis revealed that one clone, λ-ski1, contains the human counterpart of v-*ski* and that the other three, λ-sno2, λ-sno3 and λ-sno4, harbour the *ski*-related gene, which was named *sno* (*s*ki-related

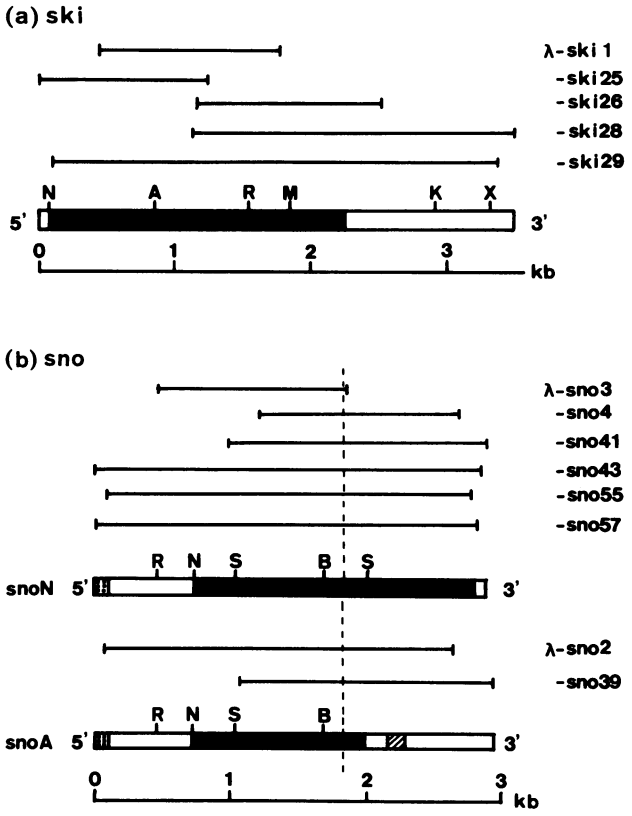



**Fig. 1** *cDNA clones of (a)human-ski and (b)human-sno*
(a) and (b) The solid and open boxes represent the coding and non-coding regions, respectively. Abbreviations: R,*Eco*RI; N,*Nco*I; A,*Apa*L1; M,*Mlu*I; K,*Kpn*I; X,*Xho*I; S,*Sst*I; B,*Bgl*I. (b) The stippled box represents the 5′ heterogeneous region. The hatched box in *sno*A represents the Alu-like sequence. The broken vertical line indicates the point where *sno*N and *sno*A show divergence.

*no*vel gene) (Fig.1). λ-ski1 and λ-sno2 were isolated from the umbilical vein cDNA library, while λ-sno3 and λ-sno4 were from the MCF-7 cDNA library.

*DNA sequence of the human-ski cDNA clones*

As the 1.3kb insert of λ-ski1 lacks both the N- and C-terminal regions [Fig.1 (a)], eight cDNA libraries were screened further with the 1.1kb *Eco*RI fragment of λ-ski1 as a probe under stringent conditions. Five clones, named λ-ski25, λ-ski26, λ-ski27, λ-ski28 and λ-ski29, were isolated from the umbilical vein cDNA library. No positive clones were obtained from the other seven libraries. λ-ski25 and λ-ski27 were found to be sister clones. The sequence of human-c-*ski* cDNA was determined by compiling the DNA sequences of these clones (Fig.2). An open reading frame of 2184bp starting with the first ATG codon at position 73 was identified. Although an in-frame termination codon is not present upstream of this ATG codon, the flanking nucleotides show a match with the consensus sequence of Kozak (32). The predicted open reading frame could encode a protein of 728

**Fig. 2** *h-ski cDNA nucleotide sequence and the deduced amino acid sequence*
The putative polyadenylation signal (AATAAG) is underlined. Polymorphism, which is boxed, is as follows: nucleotide position 3369, A in λ-ski29, G in λ-ski28; The end point (position 3373) of the insert of λ-ski29 with 7 additional A residues is shown by the symbol ( ⌋ ).

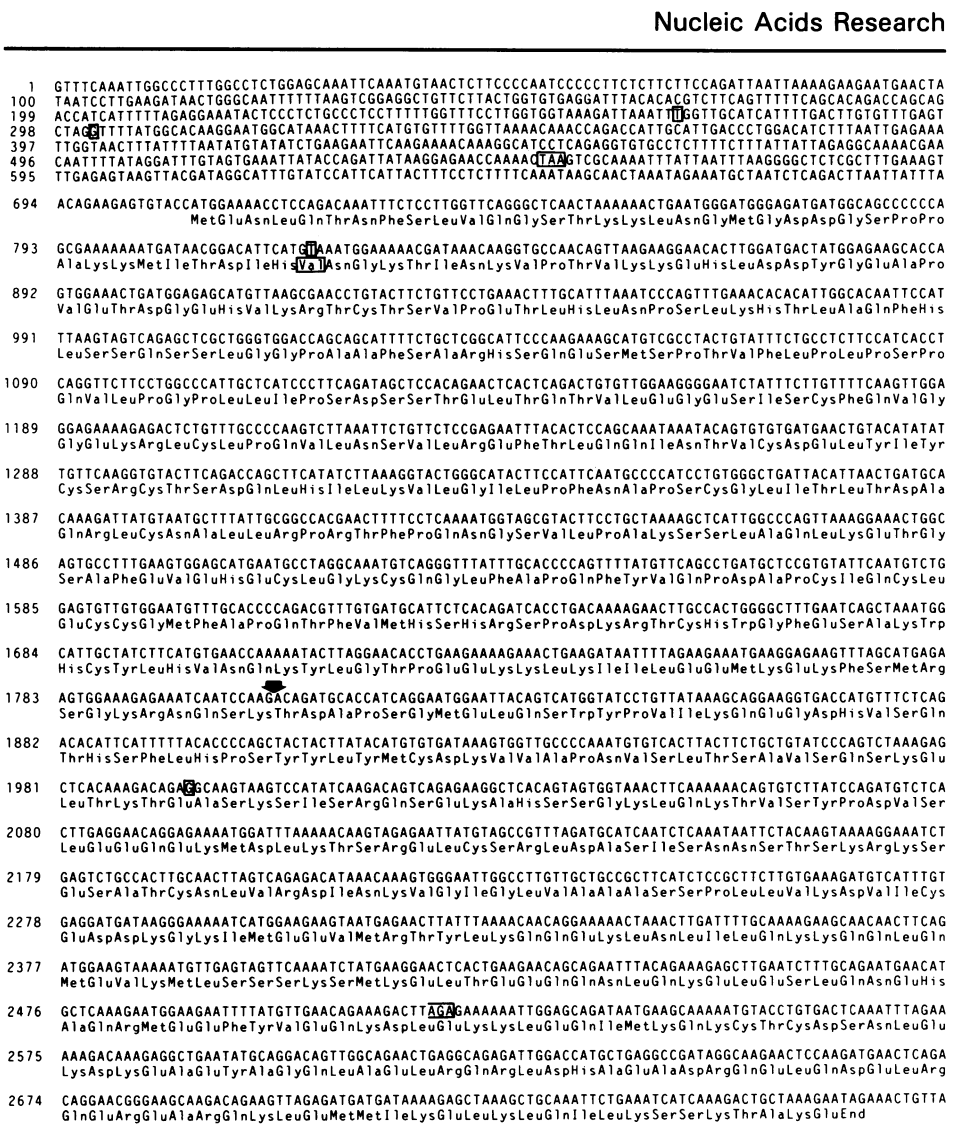


**Fig. 3** *h-snoN cDNA nucleotide sequence and the deduced amino acid sequence*
The solid arrow indicates the point where *sno*N and *sno*A show divergence. The stop codon, which is 156bp upstream of the putative initiation codon, is boxed. Polymorphism, which is also boxed, is as follows: nucleotide position 268, T in 7 clones including λ-sno55 and λ-sno57, C in λ-sno37; 302, G in 10 clones including λ-sno2, λ-sno55 and λ-sno57, A in λ-sno45 and λ-sno54; 821, T (Val)in λ-sno2 and λ-sno43, C (Ala) in λ-sno3; 1995, G in λ-sno42 and λ-sno43, A in λ-sno4. The corresponding amino acid is indicated in parenthesis. The symbol ( ·Ξ ) shows the end point (position 2519) of the insert of λ-sno56 with 8 additional A residues.

amino acid residues, whose calculated molecular weight is 80,004. Polymorphism was noted at nucleotide position 3369: A and G were detected in λ-ski29 and λ-ski28, respectively. The insert of λ-ski29 ended at position 3373 with 7 extra A residues.

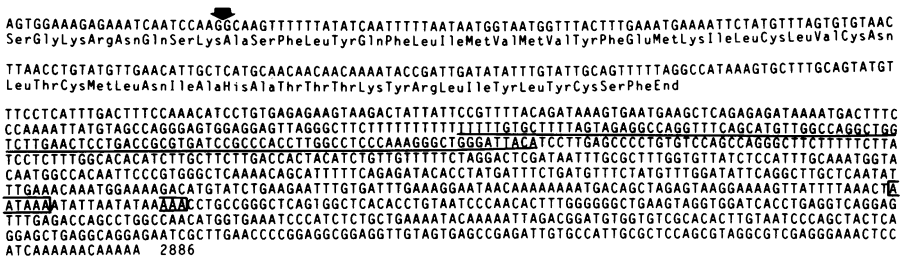



**Fig. 4** *h-snoA cDNA nucleotide sequence and the deduced amino acid sequence*
The solid arrow indicates the point where the *sno*A specific sequence starts. Only the *sno*A specific region is shown. The Alu-like sequence is underlined. The polyadenylation signal (AATAAA) is boxed. The symbol ( ⌋ ) shows the end point (position 2594) with 10 additional A residues in the insert of λ-sno2.

*DNA sequence of the human-sno cDNA clones*

The 3′ coding region of the insert of λ-sno2 has a different sequence from those of λ-sno3 and λ-sno4. This suggested the occurrence of two types of *sno* mRNAs with different coding frames in the C-terminal portion. To exclude the possibility that this was due to an artifact formed during construction of cDNA libraries, eight cDNA libraries were screened with the 1.4kb *Eco*RI fragment from λ-sno3 under stringent conditions. Another three clones were obtained from the MCF7 cDNA library, whereas 23 clones were isolated from the umbilical vein cDNA library. Single clones were obtained from the placenta and hepatoma cell Li-7 cDNA libraries, respectively. The other four libraries showed no positive signals. Southern blot hybridization of the insert of each clone (data not shown) and DNA seqencing analysis (Figs.3 and 4) confirmed the existence of two types of cDNA clones from the *sno* gene, possibly due to alternative splicing. Sixteen independent clones, including λ-sno3, λ-sno4, λ-sno41, λ-sno43, λ-sno55 and λ-sno57, were classified as the first type, named *sno*N (*n*on-Alu containing). Two independent clones, λ-sno2 and λ-sno39 from the umbilical vein cDNA library were classified as the second type, named *sno*A (*A*lu containing) (Fig.1). The other clones contained only the region common to *sno*N and *sno*A. *sno*N and *sno*A have divergent sequences downstream of position 1806. A possible translational start site of *sno* was identified at position 709, where Kozak's consensus sequence was observed. An in frame stop codon was found 156bp upstream of this ATG codon (Fig.3). The predicted open reading frames of *sno*N and *sno*A could encode proteins of 684 and 415 amino acid residues, respectively. The calculated molecular weights of the h-*sno*N and the h-*sno*A proteins are 77,003 and 46,362, respectively. An Alu-like sequence (33) was detected at positions 2130−2237 in the cDNA of *sno*A (Fig.4). In addition to 3′- heterogeneity, the 5′ non-coding region of the *sno* message seems to be multiple. Fig.5 shows two types of 5′ proximal region of the *sno* cDNA sequence. Among 10 independent clones sequenced, 4 clones, including λ-sno2 and λ-sno57, have the type (I) sequence, while 6 clones, including λ-sno43 and λ-sno55, contain the type (II) sequence. As *sno*N has both types at the 5′ end, *sno* mRNA molecules seem to be spliced alternatively in both the 5′ and 3′ region independently.

Polymorphisms of nucleotide sequences of *sno* were observed (Fig.3). λ-sno2 and λ-sno43 contain T at nucleotide position 821 and code for valine at the corresponding codon, whereas λ-sno3 retains C and codes for alanine. Polymorphism within the coding region of *sno*N was also noted at position 1995: G in λ-sno42 and λ-sno43, and A in λ-sno4.

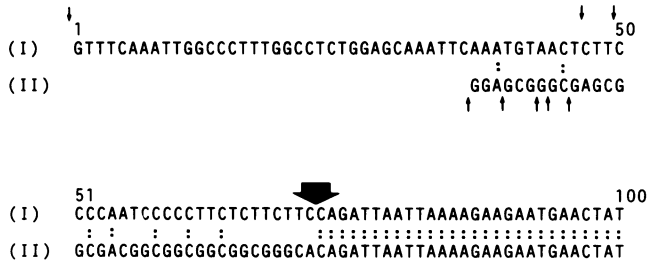



**Fig. 5** *5'-heterogeneity of the h-sno cDNA sequence*
The sequence of type (I) is the same as that in Fig.3. Nucleotides common to types (I) and (II) are indicated by (:). The solid arrow indicated the point of diversion. Thin arrows represent the 5′ end points of the inserts of chimeric phages.

In this case the same amino acid, glutamic acid, was encoded. Two polymorphic points were found in the 5′ nontranslated region. T was indentified in 7 clones including λ-sno55 and λ-sno57, whereas C was retained in λ-sno37 at nucleotide position 268. Ten clones including λ-sno2, λ-sno55 and λ-sno57 contained G at position 302, whereas λ-sno45 and λ-sno54 had A at this position. The insert of λ-sno56 ended at position 2519 in *sno*N with 8 additional A residues (Fig.3). The fragment of λ-sno2 ended at position 2594 in *sno*A and retained 10 extra A residues (Fig.4).

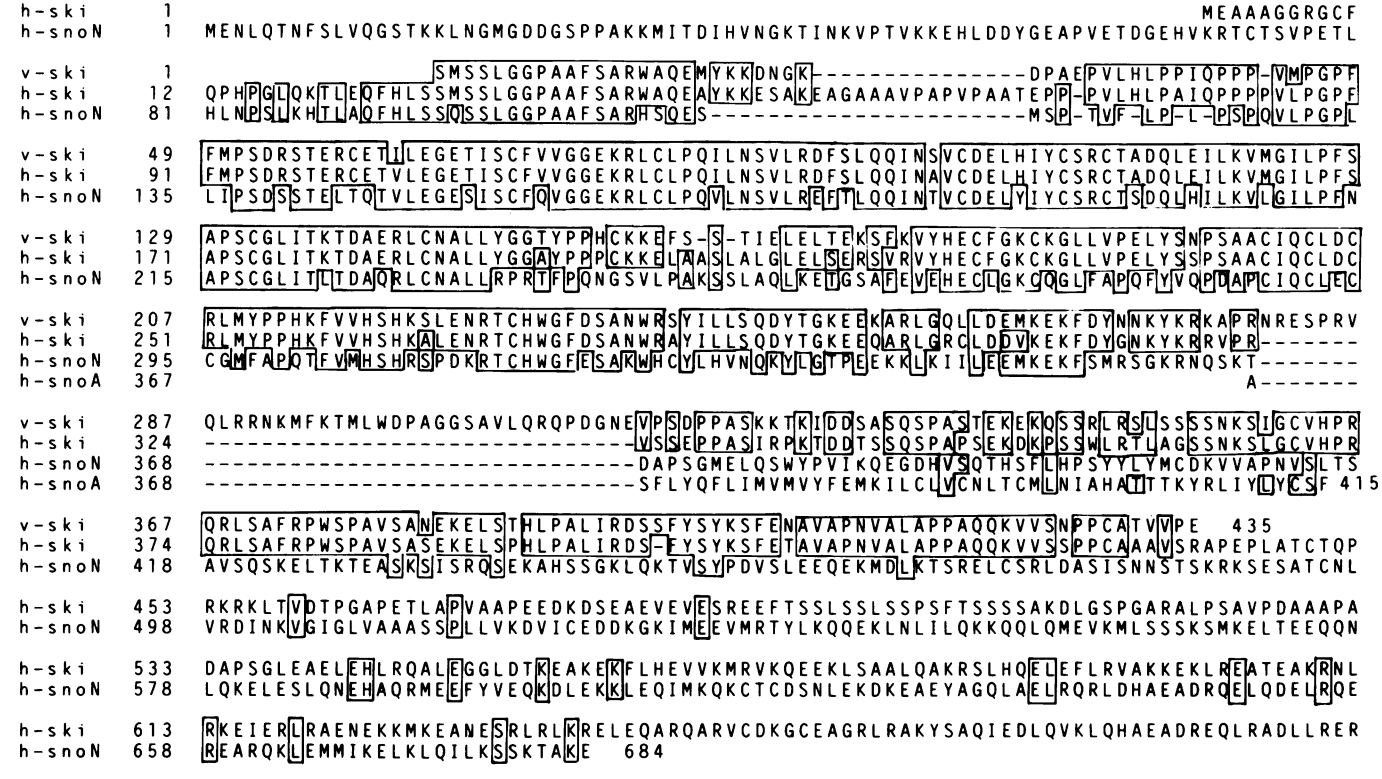
*Similarity between ski and sno*

Fig.6 shows the alignment of predicted amino acid residues of v-*ski* (34), h-*ski*, h-*sno*N and h-*sno*A. h-*ski* is highly homologous with v-*ski* except in two regions, a 15 amino acid insertion in h-*ski* (residues 55−69), and a 37 amino acid insertion in v-*ski* (residues 280−316). The latter corresponds to the exon II of the chicken c-*ski* gene (34). Therefore this suggests that the h-*ski* clones described here might lack the exon II of the human c-*ski* gene. Except in these two regions, the overall homology between h-*ski* and v-*ski* is 91% at the amino acid level. The h-*sno* protein shows high homology with both v-*ski* and h-*ski* in the N-terminal domain (residues 84−234), but has lower similarity with *ski* in the central domain (residues 263−355). The unique regions of h-*sno*N and h-*sno*A show no similarity with either v-*ski* or c-*ski*.

*mRNAs of ski and sno*

poly(A)$^+$ RNAs were prepared from various cell lines and Northern blotting was carried out as described by Thomas (30). After hybridization with the h-*ski* probe under stringent conditions, bands of 6.0, 4.7, 3.8, 3.0, 2.1 and 1.8kb were detected (Fig.7). All the cell lines, including neuroblastoma (NB39-nu, NB1), carcinoma from the vulva (A431), the stomach (NMS92), the chorion (BeWo, GCH1), the thyroid (TC78), the lung (NMS83) and the prostate (PC3) and Burkitt lymphoma (JBL-2, JBL-5) expressed c-*ski* mRNAs. As the major transcripts are 6.0kb and 4.7kb, the h-*ski* cDNA sequence determined, which is 3511bp in length, should not represent the entire c-*ski* mRNA. The G·C rich region in the 5′ noncoding region might block migration of reverse transcriptase. Therefore, the 5′ upstream domain might not be converted into the cDNA.

On hybridization with the 1.4kb *Eco*RI fragment from λ-sno3, *sno* mRNAs of 6.2, 4.4 and 3.2kb were detected in the A431, NMS92, TC78, NMS83, JBL-5 and JBL-2 cell lines. The *sno*N specific probe (nucleotide 2229−2642 in Fig.3) was also hybridized with the 6.2, 4.4 and 3.2 kb bands (data not shown). We were unable to detect the *sno*A specific


h-ski 1
h-snoN 1
MENLQTNFSLVQGSTKKLNGMGDDGSPPAKKMITDIHVNGKTINKVPTVKKEHLDDYGEAPVETDGEHVKRTCTSVPETL
v-ski 1
h-ski 12
h-snoN 81
SMSSLGGPAAFSARWAQEMYKKDNGK----------DPAEPVLHLPPIQPPP-VMPGPF
QPHPGLQKTLEQFHLSSMSSLGGPAAFSARWAQEAYKKESAKEAGAAAVPAPVPAATEPP-PVLHLPAIQPPPPVLPGPF
HLNPSLKHTLAQFHLSSQSSLGGPAAFSARHSQES----------MSP-TVF-LP-L-PSPQVLPGPL
v-ski 49
h-ski 91
h-snoN 135
FMPSDRSTERCETLLEGETISCFVVGGEKRLCLPQILNSVLRDFSLQQINSVCDELHIYCSRCTADQLEILKVMGILPFS
FMPSDRSTERCETVLEGETISCFVVGGEKRLCLPQILNSVLRDFSLQQINAVCDELHIYCSRCTADQLEILKVMGILPFS
LIPSDSSTELTQTVLEGESISCFQVGGEKRLCLPQVLNSVLREFTLQQINTVCDELYIYCSRCTSDQLHILKVLGILPFN
v-ski 129
h-ski 171
h-snoN 215
APSCGLITKTDAERLCNALLYGGTYPPHCKKEFS-S-TIELELTEKSFKVYHECFGKCKGLLVPELYSNPSAACIQCLDC
APSCGLITKTDAERLCNALLYGGAYPPPCKKELAASLALGLELSERSVRVYHECFGKCKGLLVPELYSSPSAACIQCLDC
APSCGLITLTDAQRLCNALLRPRTFPQNGSVLPAKSSLAQLKETGSAFEVEHECLGKCQGLFAPQFYVQPTAPCIQCLEC
v-ski 207
h-ski 251
h-snoN 295
h-snoA 367
RLMYPPHKFVVHSHKSLENRTCHWGFDSANWRSYILLSQDYTGKEEKARLGQLLDEMKEKFDYNNKYKRKAPRNRESPRV
RLMYPPHKFVVHSHKALENRTCHWGFDSANWRAYILLSQDYTGKEEQARLGRCLDDVKEKFDYGNKYKRRVPR-------
CGMFAPQTFVMHSHRSPDKRTCHWGFESAKWHCYLHVNQKYLGTPEEKKLKIILEEMKEKFSMRSGKRNQSKT-------
A-------
v-ski 287
h-ski 324
h-snoN 368
h-snoA 368
QLRRNKMFKTMLWDPAGGSAVLQRQPDGNEVPSDPPASKKTKIDDSASQSPASTEKEKQSSRLRSLSSSSNKSIGCVHPR
----------VSSEPPASIRPKTDDTSSQSPAPSEKDKPSSWLRTLAGSSNKSLGCVHPR
----------DAPSGMELQSWYPVIKQEGDHVSQTHSFLHPSYYLYMCDKVVAPNVSLTS
----------SFLYQFLIMVMVYFEMKILCLVCNLTCMLNIAHATTTKYRLIYLYCSF 415
v-ski 367
h-ski 374
h-snoN 418
QRLSAFRPWSPAVSANEKELSTHLPALIRDSSFYSYKSFENAVAPNVALAPPAQQKVVSNPPCATVVPE 435
QRLSAFRPWSPAVSASEKELSPHLPALIRDS-FYSYKSFETAVAPNVALAPPAQQKVVSSPPCAAAVSRAPEPLATCTQP
AVSQSKELTKTEASKSISRQSEKAHSSGKLQKTVSYPDVSLEEQEKMDLKTSRELCSRLDASISNNSTSKRKSESATCNL
h-ski 453
h-snoN 498
RKRKLTVDTPGAPETLAPVAAPEEDKDSEAEVEVESREEFTSSLSSLSSPSFTSSSSAKDLGSPGARALPSAVPDAAAPA
VRDINKVGIGLVAAASSPLLVKDVICEDDKGKIMEEVMRTYLKQQEKLNLILQKKQQLQMEVKMLSSSKSMKELTEEQQN
h-ski 533
h-snoN 578
DAPSGLEAELEHLRQALEGGLDTKEAKEKFLHEVVKMRVKQEEKLSAALQAKRSLHQELEFLRVAKKEKLREATEAKRNL
LQKELESLQNEHAQRMEEFYVEQKDLEKKLEQIMKQKCTCDSNLEKDKEAEYAGQLAELRQRLDHAEADRQELQDELRQE
h-ski 613
h-snoN 658
RKEIERLRAENEKKMKEANESRLRLKRELEQARQARVCDKGCEAGRLRAKYSAQIEDLQVKLQHAEADREQLRADLLRER
REARQKLEMMIKELKLQILKSSKTAKE 684
h-ski 693
EAREHLEKVVKELQEQLWPRARPEAAGSEGAAELEP 728

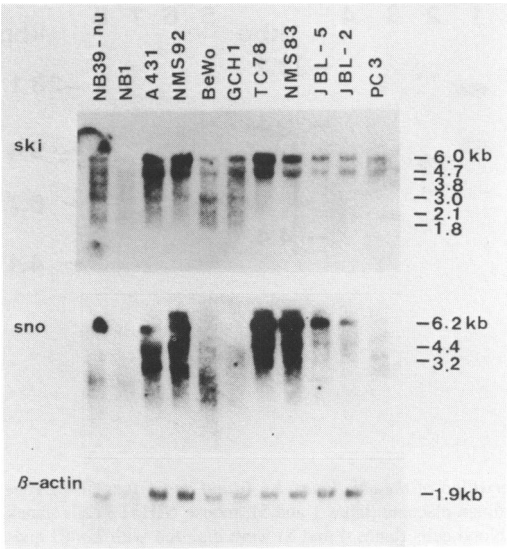



**Fig. 7** *mRNAs of h-ski and h-sno*
poly(A)$^+$RNA was analyzed as described in Materials and Methods. Hybridization was performed in a solution containing 50% formamide, 5×SSC, 5×Denhardt's solution, 100μg/ml of sonicated salmon testis DNA and $^{32}$P-labeled probe at 37°C for 16 hours. After several washing in 1×SSC, 0.1%SDS at room temperature, the filter was finally washed in 0.1×SSC, 0.5%SDS at 50°C for 1 hour. The *ski* probe used is the 1090bp fragment of λ-ski1 (nucleotides 456−1545 in Fig.2). The fragment used as probe for *sno* is the insert of λ-sno3 (nucleotides 429−1821 in Fig.3). Although this 1393bp fragment retains the 15bp *sno*N specific domain (nucleotides 1807−1821), almost the entire region is common to *sno*N and *sno*A. The same filter was rehybridized with the β-actin probe (45).

bands. This suggests that the level of *sno*A mRNA is lower than that of *sno*N mRNA. Actually sixteen *sno*N and two *sno*A clones were isolated in this work.

Although the expression patterns of *ski* and *sno* are similar, there are distinct differences. 6.2 and 4.4 kb *sno* mRNAs were not observed in NB39-nu, NB-1, BeWo, and GCH-1 cells which expressed 6.0 and 4.7kb *ski* mRNAs (Fig.7). Final washing with 0.1×SSC and 0.5% SDS at 60°C did not change the hybridization pattern compared to washing at 50°C (data not shown). Cross-reaction between *ski* and *sno* probes was not observed in Southern blot hybridization (Fig.8). These indicate that *ski* and *sno* probes detect specific mRNAs, respectively.

*Evolutionary conservation of ski and sno*

Southern blots of *Eco*RI-cleaved mammalian and avian genomic DNAs were hybridized with the h-*ski* or the h-*sno* probe. Fig.8 shows the evolutionary conservations of the *ski* and the *sno* gene. 140 human tumors, including 60 cell lines and 80 fresh tumors, were surveyed for possible amplification and/or rearrangement of c-*ski* and c-*sno*, but no structural aberrations were observed (data not shown).

**Fig. 6** *Comparison of the amino acid sequences of v-ski, human-ski, human-snoN and human-snoA*
Dashes indicate gaps that have been introduced for alignment. Only the specific region of h-*sno*A (367−415) is shown. Identical amino acids are boxed.

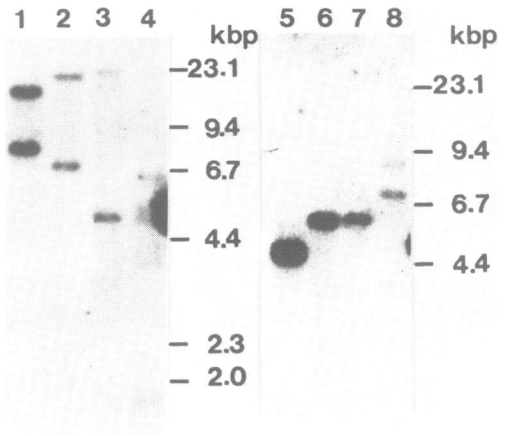



**Fig. 8** *Evolutionary conservation of the ski (lanes 1–4) and sno (lanes 5–8) genes*
DNAs (10μg each) from human placenta (lanes 1 and 5), mouse NIH3T3 cells (lanes 2 and 6), Fisher rat (lanes 3 and 7) and chicken red blood cells (lanes 4 and 8) were digested with *Eco*RI and analyzed by Southern blot hybridization. Hybridization was performed in a solution containing 40% formamide, 5×SSC, 0.5%SDS, 5×Denhardt's solution, 100μg/ml of sonicated salmon testis DNA and $^{32}$P-labeled probe at 37°C for 16hours. After several washing in 1×SSC, 0.1×SDS at room temperature, filters were finally washed in 0.1×SSC, 0.5%SDS at 40°C for 1 hour. The fragments used as probes are the same as Fig.7. *Hin*dIII digested λ DNA was used as a size marker.

## DISCUSSION

Eight kinds of cDNA libraries were screened with a v-*ski* probe, and cDNAs of the h-*ski* and the h-*sno* gene were isolated. As the v-*ski* product is localized in the nucleus (4), the gene products of h-*ski* and h-*sno* can also function in the nucleus. Highly conserved regions between the *ski* and *sno* genes are mapped within the N-terminal half, and these should be responsible for common functions of the *ski* gene family, such as binding to a specific region(s) of DNA and/or interaction with a specific protein(s). One of the nuclear oncogenes, c-*jun*, encodes a transcription factor AP-1 and cooperates with the *fos* protein (35–37). Therefore, it is tempting to speculate that the gene products of h-*ski* and h-*sno* could be transcriptional regulatory proteins. Experiments to test ability for DNA binding or transactivation are in progress. In preliminary work, c-*sno* was mapped in chromosome 3 of the human genome (M.Yoshida and N.Nomura, unpublished).

The *sno* messages are probably spliced alternatively within the coding region, producing two gene products. The functions of the proteins are not yet known, but the *sno*N and the *sno*A products might have different roles, as suggested for the c-*src* (38). A similar phenomenon was reported in the case of *erg*1 and *erg*2 (39). Alternative processing seems to occur in the 5′ nontranslated region of the *sno* gene in addition to the 3′ region. This may reflect a complex mechanism of gene expression, as shown in c-*abl* (40) and L-*myc* (41). Heterogeneous 5′ exons were also reported in N-*myc* (42) and *lck* (43).

To avoid artifacts during construction of cDNA library, we confirmed all sequences by analyzing at least two overlapping λ phage clones. When polymorphism was observed, another clone(s) was analyzed. The polymorphic point identified at position 302 in c-*sno* is most likely due to somatic mutation. The identification of each allele in multiple clones

excludes the possibility that the polymorphism was an artifact during construction of the cDNA library, such as an occasional reading error of reverse transcriptase (44). In the other positions, that is, 3369 in c-*ski*, 268 and 821 in c-*sno* and 1995 in *sno*N, only a single clone represents one of the two alleles. Therefore we could not determine whether the observed polymorphism was due to actual somatic mutation or an artifact during construction of the cDNA library. We chose the allele in the major group to represent the cDNA sequences.

Some of the clones have inserts ending around a polyA stretch with extra A residues. This finding suggests that reverse transcription primed with oligo (dT) might start at the polyA stretch in addition to the poly(A)$^+$ tail region.

## ACKNOWLEDGMENTS

We thank E.Stavnezer for plasmid pCCL ski5 and for letting us know about results before their publication, T.Kunieda for DNAs of mouse NIH3T3 cells and Fisher rat and M.Ide for excellent technical assistance. We also thank J.E.Sadler, P.Chambon, D.R.Helinski, S.L.C.Woo, T.Chandra, D.P.Dialynas, W.L.Miller and J.M.Puck for generously providing cDNA libraries. We are grateful to N.Miyajima, A.Ito, M.Yamamoto, H.Uchida, F.Yamao and H.Fujii for help in the computer analysis, T.Yamamoto, C.Colmenares, P.Sutrave and S.H.Hughes for helpful discussion and T.Suzuki, I.Miyoshi, S.Sasaki, Y.Nakagami and H.Ohami and the Japanese Cancer Research Resources Bank for cell lines. This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan.

*To whom correspondence should be addressed

## REFERENCES

1. Stavnezer, E., Gerhard, D.S., Binari, R.C. and Balazs, I. (1981) J.Virology *39*, 920–934.
2. Li, Y., Turck, C.M., Teumer, J.K. and Stavnezer, E. (1986) J.Virology *57*, 1065–1072.
3. Stavnezer, E., Barkas, A.E., Brennan, L.A., Brodeur, D., and Li, Y. (1986) J.Virology *57*, 1073–1083.
4. Barkas, A.E., Brodeur, D. and Stavnezer, E. (1986) Virology *151*, 131–138.
5. Kohl, N.E., Kanda, N., Schreck, R.R., Bruns, G., Latt, S.A. Gilbert, F. and Alt, F.W. (1983) Cell *35*, 359–367.
6. Schwab, M., Alitalo, K., Klempnauer, K.-H., Varmus, H.E., Bishop, J.M., Gilbert, F., Brodeur, G., Goldstein, M. and Trent, J. (1983) Nature (London) *305*, 245–248.
7. Kaye, F., Battey, J., Nau, M., Brooks, B., Seifter, E., De Greve, J., Birrer, M., Sausville, E. and Minna, J. (1988) Mol.Cell.Biol. *8*, 186–195.
8. Ryder, K., Lau, L.F. and Nathans, D., (1988) Proc.Natl.Acad.Sci.USA. *85*, 1487–1491.
9. Ryder, K., Lanahan, A., Perez-Albuerne, E. and Nathans, D. (1989) Proc.Natl.Acad.Sci.USA. *86*, 1500–1503.
10. Cohen, D.R. and Curran,T. (1988) Mol.Cell.Biol. *8*, 2063–2069.
11. Zerial, M., Toschi, L., Ryseck, R-P., Schuermann, M., Muller, R. and Bravo, R. (1989) EMBO J. *8*, 805–813.
12. Ingvarsson, S., Asker, C., Axelson, H., Klein, G. and Sumegi, J. (1988) Mol.Cell.Biol. *8*, 3168–3174.
13. Nomura, N., Takahashi, M., Matsui, M., Ishii, S., Date, T., Sasamoto, S. and Ishizaki, R. (1988) Nucleic Acids Res. *16*, 11075–11089.
14. Petkovich, M., Brand, N.J., Krust, A. and Chambon, P. (1987) Nature (London) *330*, 444–450.
15. Evans, R.M. (1988) Science *240*, 889–895.
16. Hunter, T. (1987) Cell *50*, 823–829.
17. Sadler, J.E., Shelton-Inloes, B.B., Sorace, J.M., Harlan, J.M., Titani, K. and Davie, E.W. (1985) Proc.Natl.Acad.Sci.USA. *82*, 6394–6398.
18. Walter, P., Green, S., Greene, G., Krust, A., Bornert, J.- M., Jeltsch, M., Staub, A., Jensen, E., Scrace, G., Waterfield, M. and Chambon, P. (1985) Proc.Natl.Acad.Sci. USA. *82*, 7889–7893.

19. de Wet, J.R., Fukushima, H., Dewji, N.N., Wilcox, E., O'Brien, J.S. and Helinski, D.R. (1984) DNA *3*, 437–447.
20. Kwok, S.C.M., Ledley, F.D., Dilella, A.G., Robson, K.J.H. and Woo, S.L.C. (1985) Biochemistry *24*, 556–561.
21. Dialynas, D.P., Murre, C., Quertermous, T., Boss, J.M., Leiden, J.M., Seidman, J.G. and Strominger, J.L. (1986) Proc.Natl.Acad.Sci.USA. *83*, 2619–2623.
22. Chung, B.-C., Matteson, K.J., Voutilainen, R., Mohandas, T.K. and Miller, W.L. (1986) Proc.Natl.Acad.Sci.USA. *83*, 8962–8966.
23. Gold, D.P., Puck, J.M., Pettey, C.L., Cho, M., Coligan, J., Woody, J.N. and Terhorst, C. (1986) Nature (London) *321*, 431–434.
24. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Gene *33*, 103–119.
25. Mizusawa, S., Nishimura, S. and Seela, F. (1986) Nucleic Acids Res. *14*, 1319–1324.
26. Sanger, F., Nicklen, S. and Coulson, A.T. (1977) Proc.Natl.Acad.Sci.USA. *74*, 5463–5467.
27. Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. *12*, 387–395.
28. Kanehisa, M. (1984) Nucleic Acids Res. *12*, 203–213.
29. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning (Cold Spring Harbor Laboratory) 191–193.
30. Thomas, P.S. (1983) In Wu, R., Grossman, L. and Moldave, K. (eds) Methods in Enzymology, Academic Press, New York. Vol. 100, pp255–266.
31. Southern, E. (1975) J.Mol.Biol. *98*, 503–517.
32. Kozak, M. (1984) Nucleic Acids Res. *12*, 857–872.
33. Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T. and Schmid, C.W. (1981) J.Mol.Biol. *151*, 17–33.
34. Stavnezer, E., Brodeur, D. and Brennan, L. (1989) Mol.Cell.Biol. (in press)
35. Bohmann, D., Bos, T.J., Admon, A., Nishimura, T, Vogt, P.K. and Tjian, R. (1987) Science *238*, 1386–1392.
36. Angel, P., Allehretto, E.A., Okino, S.T., Hattori, K., Boyle, W.J., Hunter, T. and Karin, M. (1988) Nature(London) *332*, 166–171.
37. Rauscher III, F.J., Cohen, D.R., Curran, T., Bos, T.J., Vogt, P.K., Bohmann, D., Tjian, R. and Franza, B.R.Jr. (1988) Science *240*, 1010–1016.
38. Martinez, R., Mathey-Prevot, B., Bernards, A. and Baltimore, D. (1987) Science *237*, 411–415.
39. Rao, V.N., Papas, T.S. and Reddy, E.S.P. (1987) Science *237*, 635–639.
40. Bernards, A., Rubin, C.M., Westbrook, C.A., Paskind, M. and Baltimore, D. (1987) Mol.Cell.Biol. *7*, 3231–3236.
41. Greve, J.De., Battey, J., Fedorko, J., Birrer, M., Evan, G., Kaye, F., Sausville, E. and Minna, J. (1988) Mol. Cell. Biol. *8*, 4381–4388.
42. Stanton, L.W. and Bishop, J.M. (1987) Mol.Cell.Biol. *7*, 4266–4272.
43. Voronova, A.F., Adler, H.T. and Sefton, B.M. (1987) Mol.Cell.Biol. *7*, 4407–4413.
44. Verma, I.M. (1977) Biochem.Biophys.Acta *473*, 1–38.
45. Cleveland, D.W., Lopata, M.A., MacDonald, R.J., Cowan, N.J., Rutter, W.J. and Kirschner, M.W. (1980) Cell *20*, 95–105.

**This article, submitted on disc, has been automatically converted into this typeset format by the publisher.**