Nucleic Acids Research, Vol. 20, No. 7 1511–1516

C-*ski* cDNAs are encoded by eight exons, six of which are closely linked within the chicken genome

H.Leighton Grimes[1,2], Brian E.Szente[1,2] and Maureen M.Goodenow[1,2,3]*
[1]Graduate Program in Immunology and Molecular Pathology, [2]Department of Pathology and [3]Department of Pediatrics, Box 100275, J.Hillis Miller Health Center, University of Florida, Gainesville, FL 32610, USA

Received January 27, 1992; Revised and Accepted March 6, 1992 EMBL accession no. X62159

ABSTRACT

The c-*ski* locus extends a minimum of 65 kb in the chicken genome and is expressed as multiple mRNAs resulting from alternative exon usage. Four exons comprising approximately 1.5 kb of cDNA sequence have been mapped within the chicken c-*ski* locus. However, c-*ski* cDNAs include almost 3 kb of sequence for which the exon structure was not defined. From our studies using the polymerase chain reaction and templates of RNA and genomic DNA, it is clear that c-*ski* cDNAs are encoded by a minimum of eight exons. A long 3′ untranslated region is contiguous in the genome with the distal portion of the ski open reading frame such that exon 8 is composed of both coding and noncoding sequences. Exons 2 and 3 are separated by more than 25 kb of genomic sequence. In contrast, exons 3 through 8, representing more than half the length of c-*ski* cDNA sequences, are closely linked within 10 kb in the chicken genome.

INTRODUCTION

A variety of mechanisms have evolved that expand the coding repertoire of the eukaryotic genome. Whereas somatic rearrangement of DNA occurs in relatively restricted cell lineages such as lymphocytes, alternative splicing of RNA has evolved as a more universal mechanism whereby genes can be regulated and protein diversity can be generated. Regulated alternative splicing plays a pivotal role in initiation or maintenance of developmental programs, for example by acting as an on/off switch in sexual differentiation or by spatial/temporal modulation of homeotic genes. In a variety of species, alternative splicing is prevalent in genes encoding contractile proteins or cell adhesion molecules, and in genes expressed in muscle cells (1–4).

The proto-oncogene c-*ski* is expressed as a number of RNA transcripts in chicken, human, and mouse cells. Different c-*ski* cDNAs have been isolated from chick embryo cells suggesting that multiple c-*ski* mRNAs can be produced by alternative splicing (5). Isoforms of Ski proteins expressed by alternatively spliced cDNAs in retrovirus vectors display subtle variations in their half-life, in their pattern of nuclear localization, and in their biological properties (6,7).

Similar to many proto-oncogenes, c-*ski* was identified by virtue of its homology with v-*ski*, the oncogene transduced by the avian retroviruses SKVs (8–10). The 1.3-kb v-*ski* sequences represent a truncated version of the longest c-*ski* cDNA, which is 4.2 kb (5,10). Infection with avian retroviruses containing either v-*ski* or c-*ski* sequences has a potent effect on proliferation and morphological transformation of chicken embryo cells, and induces myogenic differentiation in quail embryo cells in culture (7,11–14). The dominant effect of ski in myogenesis in culture is associated with molecular and biochemical programs involved in normal myoblast differentiation (13,14). The effects of ski in myogenesis are not confined to cells in culture, but can be demonstrated in animals as well. Transgenic mice carrying a portion of the chicken c-*ski* proto-oncogene develop a distinctive muscular phenotype due to selective hypertrophy of fast skeletal muscle fibers and a concomitant decrease in body fat (15).

The unusual dual functions in proliferation and in differentiation of ski sequences expressed in avian retroviruses or in transgenic animals led us to examine expression of the endogenous c-*ski* locus in chickens. To assess the relationship between c-*ski* transcripts and c-*ski* cDNAs containing different exons, it was necessary to develop probes specific for each of the c-*ski* exons. It has been proposed that the chicken c-*ski* locus is organized into at least seven coding exons and a long 3′ noncoding region (5). However, only four ski exons (exons 1, 2, 4, and 5), representing about one-third of the known c-*ski* locus in chickens, were defined and mapped directly from genomic clones (10). Exon 1 appears to be the largest of the c-*ski* exons, extending almost 1 kb in the genome. In contrast, exons 2, 4, and 5 are small exons of 111, 115, and 265 bp, respectively. The four exons are dispersed over more than 65 kb of the chicken genome (10). Exon 1 and exon 2 are separated by at least 20 kb, whereas exons 4 and 5 are only 841 bp apart. Genomic clones containing exons 1 and 2 have not been linked to clones containing exons 3 and 4, in part because exon 3, defined as the 126 bp found between exons 2 and 4 in v-*ski* as well as in c-*ski* cDNAs (5,10), has not been located in the genome. Likewise, the genomic

* To whom correspondence should be addressed

organization of 2.5 kb of c-*ski* sequences downstream of exon 5 in the longest c-*ski* cDNAs has not been examined. Organization of exons 6 and 7 was inferred from the structure of cDNAs, although the exon structure of the 3′ noncoding region and its relationship with the coding exons in the c-*ski* locus was not determined (5).

Our primary goals were to assess if sequences proposed as exons in c-*ski* cDNAs were indeed organized as such in the chicken genome and to determine the exon structure of the long 3′ noncoding region. A second aim was to map these exons within the c-*ski* locus. Because the unmapped c-*ski* exons were generally small and might well be dispersed over vast expanses of the chicken genome, we were reluctant to screen conventional genomic libraries to locate the unmapped sequences. Rather, we adopted a strategy using enzymatic amplification of genomic DNA with primers specific for known or proposed c-*ski* exons to examine the structure of the sequences and their organization within the c-*ski* locus. Our studies show that sequences encoding the longest c-*ski* cDNA are organized into eight exons. Exon 8 includes the distal portion of the open reading frame, as well as the entire 3′ noncoding region. Exons 3 through 8, comprising almost 3.0 kb of c-*ski* cDNA, are closely linked in the chicken genome.

MATERIAL AND METHODS

Southern analysis

Genomic DNA was extracted by standard methods (16) and digested with restriction enzymes according to the manufacturer's protocols (BRL Gibco). Restricted DNA was electrophoresed in 0.7% agarose gels and transferred to NYTRAN membrane (Schleicher and Schuell) according to the method of Southern (17). Probes were prepared from double-stranded fragments, which were excised with appropriate restriction enzymes from plasmid vectors, radiolabelled with ^{32}P-dATP to a specific activity of at least 3×10^9 cpm/μg by the random primed method (18), and separated from unincorporated nucleotides by Sephadex G-50 (Pharmacia) spun column (16). Filters were prehybridized for 1 h at 60°C in phosphate hybridization buffer (500 mM $NaPO_4$ pH 7.2, 1 mM EDTA pH 8.0, 7% SDS, 1% BSA) (19) followed by 16 h at 60°C with 1×10^6cpm of radiolabelled probe per ml of buffer. After washing for one hour at 60°C in 40 mM $NaPO_4$ with 0.1% SDS, filters were exposed to Kodak XAR 5 film with intensifying screens at −70°C for two days.

RNA extraction and cDNA synthesis

RNA was extracted from 10-day Spafas chick embryo torsos according to standard methods (16). Tissues were processed in a Virtis homogenizer in guanidine thiocyanate buffer (4 M guanidine thiocyanate, 25 mM sodium citrate pH 7.0, 0.5% Sarkosyl, 0.1 M betamercaptoethanol). The homogenate was layered onto CsCl (5.7 M CsCl, 0.1 M NaOAc pH 5.0, 100 mM EDTA), and the RNA was pelleted by ultracentrifugation at 200,000×g for 16 hours in a Beckman SW40 rotor. The RNA pellet was resuspended in ETS (5mM EDTA, 10 mM Tris, 0.05% SDS), and passed through an oligo(dT)-cellulose column to select poly(A) RNA. The poly(A) RNA was ethanol precipitated, pelleted for 30 minutes at 26,500×g in a Beckman SW13.1 rotor, resuspended in H2O, and quantitated by absorbance at 260 nm.

First strand cDNA synthesis was performed using Moloney murine leukemia virus reverse transcriptase without RNase H activity (Superscript, BRL Gibco), according to the manufacturer's protocol except that 20 ng random hexamers (Pharmacia) were used to prime the cDNA synthesis from 1 μg of poly(A) RNA at 45°C.

Polymerase chain reaction and DNA sequencing

Oligonucleotide primers specific for proposed c-*ski* exons were prepared on an Applied Biosystems DNA synthesizer in the DNA synthesis core facility of the Interdisciplinary Center for Biotechnology Research at the University of Florida. The location and sequence of the primers are summarized in Fig. 1.

C-*ski* sequences were amplified from 1 μg of chicken genomic DNA or one-fifth of a first strand cDNA synthesis reaction in 100 μl reactions with polymerase chain reaction (PCR) buffer (1.5 mM MgCl2, 500 mM KCl, 200 mM TRIS pH 8.4, 0.1% BSA), 100 pmol of each primer, 200 μmol of each dNTP, and 2.5 units Amplitaq (Perkin Elmer) overlayed with 20 μl mineral oil (20). Amplification reactions were carried out in an automated thermal cycler (Perkin Elmer Cetus) programmed for one cycle of denaturing (94°C for 10 min), 35 cycles of amplification (each cycle included denaturation for 1 min at 94°C, 1 min primer annealing at 60°C, and 2 min polymerization at 72°C), followed by one cycle of polymerization (72°C for 10 min). Products of the amplification reactions were electrophoresed in agarose gels, then transferred to NYTRAN and probed as described for genomic DNA samples. Amplified DNA fragments were cloned into pGEM-3Z for double stranded sequencing with Sequenase Version 2.0 (United States Biochemical) according to the manufacturer's protocols.

RESULTS

Proposed exons 3 and 6 are exons

Primers MG48 and MG52, complementary to cDNA sequences between exon 2 and exon 4, were used with chick embryo cDNA or genomic DNA as templates. To assure specificity of the

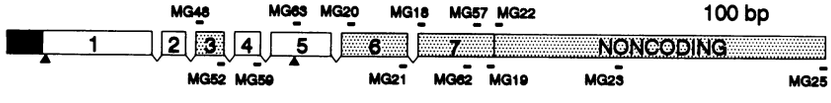

Figure 1. Position of primers used to analyze proposed exon organization of c-*ski*. Exons 1, 2, 4, and 5 (open boxes) were defined from genomic clones (10). The 5′ noncoding region of exon 1 is indicated in black. Proposed exons 3, 6, and 7, including the 3′ noncoding region, are indicated by shaded boxes (5,10). Arrowheads beneath exon 1 and exon 5 indicate the boundaries of v-*ski* (10). Forward amplification primers are indicated above the exons. Reverse amplification primers are designated below. The nucleotide sequence of the primers (5′ to 3′) and their location from the cDNA sequence (nucleotide numbers in parenthesis from ref.5) are as follows: MG18, CTGGAGTTCCTCAGAGTGGC (2004 to 2023); MG19, TTCTCCATGTTGCTCGTTGTG (2416 to 2396); MG20, CCATCC-TTTACTTCATCCAGC (1728 to 1748); MG21, CTGATGTAGGCTGCGTTTGGC (2000 to 1980); MG22, CGTGTGTAAGTGCATGTGTGAG (2482 to 2503); MG23, CTCATACAACCACAGGACAGG (3033 to 3013); MG25, GTCAGCATTGCCTATAGAAGATGC (4144 to 4121); MG48, GTCCCTTCAGATCCTCC-TGCT (1191 to 1211); MG52, ATGAACTGGATAAGGACCGT (1308 to 1289); MG57, GAGGACCTACAGGTTAAGCT (2235 to 2254); MG59, CTGTCT-CGGATCAATGCA (1432 to 1415); MG62, TTCACAACCCTTGTCGCA (2198 to2181); MG63, CAGAAAGTTGTGAGCAACC (1503 to 1521).

templates, cDNA was synthesized from random-primed, poly(A)-selected RNA isolated from the torso of 10-day chick embryos, while genomic DNA was treated with RNase before amplification. Products of identical size were amplified from both cDNA and genomic DNA templates, but not in the absence of template (Fig. 2, lanes 1–3). The amplified product from genomic DNA was cloned, sequenced, and found to be identical to the 126 bp located between exons 2 and 4 in c-*ski* cDNAs. The genomic sequence we determined for exon 3 included a T residue corresponding to the T in c-*ski* cDNAs at nucleotide 1284 (5). This T residue represents the single nucleotide substitution in the region of homology between c-*ski* and v-*ski* that would change the predicted amino acid sequence from a tryptophan in the c-*ski* encoded protein to an arginine in the v-*ski* oncoprotein (10).

A similar amplification strategy was used to evaluate the genomic organization of the 301 bp proposed to be exon 6 in c-*ski* cDNA. Amplification of either cDNA or genomic DNA with primers MG20 and MG21, complementary to sequences near the 5′ and 3′ termini of designated exon 6, produced a 301-bp fragment that was not detected in the absence of template (Fig. 2, lanes 4–6). When the product amplified from genomic DNA was cloned and sequenced, it was found to be identical with that of the cDNA sequence (5). Results of our analyses provide evidence that two of the proposed exons in cDNAs are organized as exons in the genome.

Proposed exon 7 is two exons

Downstream of exon 6 in c-*ski* cDNAs are more than 2 kb of coding and noncoding sequences that have been designated exon 7 (5). To analyze the exon structure of this region, we first used primers MG18 and MG19 to amplify the 420 bp of coding sequences comprising proposed exon 7 (Fig. 3A). When cDNA was used as template in the amplification reaction, a fragment of the expected 420-bp size resulted (Fig. 3B, lane 2). In contrast, when genomic DNA was used as a template with primers MG18 and MG19, a product of approximately 2.2 kb, rather than the predicted 420 bp, was obtained (Fig. 3B, lane 1). Because the 2.2-kb product hybridized with a probe, MT211, derived from the coding sequences of proposed exon 7 (Fig. 3A), we considered the possibility that designated exon 7 in c-*ski* cDNAs was organized as more than a single exon in the genome. The possibility was supported by results of Southern analysis of genomic DNA with the MT211 probe, which hybridized to two HindIII fragments, as well as two EcoRI or PstI fragments (Fig. 3C). The presence of a HindIII site in this region of the genome was expected from the cDNA sequence (5) (Fig. 3A). However, neither the EcoRI nor the PstI sites identified by Southern analysis of this region in the genome were predicted from the cDNA sequence of proposed exon 7 (5) (Fig. 3A).

The 2.2-kb fragment amplified from the genome was cloned, analyzed with restriction enzymes, and partially sequenced. The results revealed that the first 231 bp of the coding region of proposed exon 7 were separated from the distal 189 bp by an intervening sequence of 1.8 kb (Fig. 3D). Consensus splice donor and acceptor signals were identified at the 5′ and 3′ ends of the intervening sequence, respectively (21) (Table 1). The intervening sequence contained EcoRI and PstI restriction sites consistent with the results of Southern analysis of genomic DNA. Although the sequence of the entire intron was not determined, a probe derived from the intron by digestion with HindIII and XbaI (Fig. 3D) did not hybridize to poly(A) RNA from thirteen different chicken tissues, indicating that additional exons are not lurking within

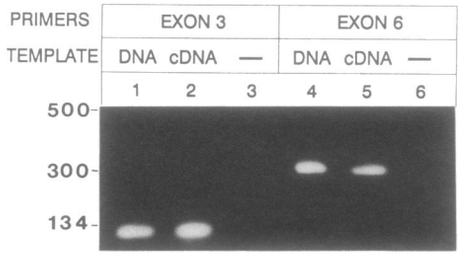

Figure 2. Amplification of proposed c-*ski* exons 3 and 6 from chicken genomic DNA and cDNA. Primers for exon 3 were MG48 and MG52; primers for exon 6 were MG20 and MG21 (see Materials and Methods). One-tenth of amplified products were analyzed by electrophoresis through 1.2% agarose gels stained with ethidium bromide. Size markers in base pairs are indicated at the left. Primers: lanes 1–3, exon 3; lanes 4–6, exon 6. Templates: lanes 1 and 4, genomic DNA; lanes 2 and 5, cDNA; lanes 3 and 6, no template.

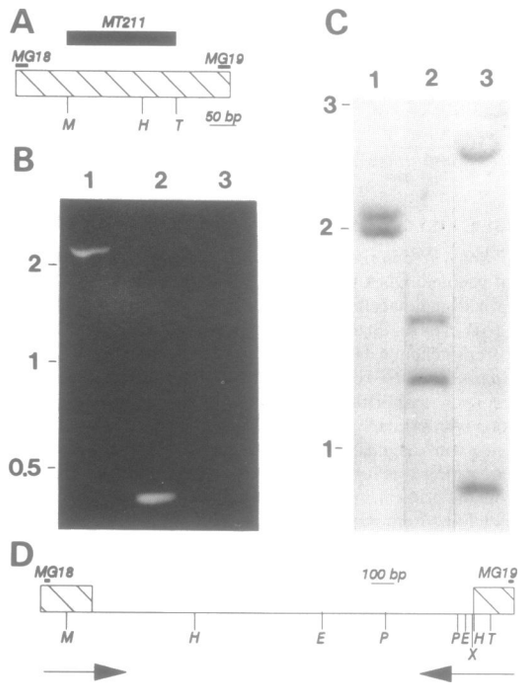

Figure 3. Proposed exon 7 is more than one exon. A. Diagram of the 420-bp coding region of proposed exon 7 with selected restriction sites mapped from the cDNA sequence (5). Symbols: M, MboI; H, HindIII; and T, TaqI. There are no EcoRI or PstI sites in the cDNA sequence of proposed exon 7. MG18 and MG19 indicate position of primers used for amplification. MT211, a 211-bp fragment derived by MboII and TaqI digestion, is a double-stranded probe which does not include either MG18 or MG19 primer sequences. B. Amplification of genomic DNA (lane 1), cDNA (lane 2), or no template (lane 3) with MG18 and MG19 primers. One-tenth of the amplified products were electrophoresed through ethidium bromide stained 1.2% agarose gel. Size markers in kb are indicated at the left. C. Southern analysis of genomic DNA digested with HindIII (lane 1), EcoRI (lane 2), and PstI (lane 3) hybridized with the radiolabelled MT211 probe. D. Diagram of exon and intron organization with restriction sites detected by sequencing or by restriction analysis. Hatched boxes indicate exon sequences while the intron sequences are represented by a line. Arrows represent direction and extent of sequencing. Symbols: E, EcoRI; X, XbaI; P, PstI; and, as above.

Table 1. Splice Junction Sequences

	Splice Donor	Splice Acceptor
Consensus	$\frac{C}{A}AG/GT\frac{A}{G}AGT$	$(\frac{T}{C})_n N\frac{C}{T}AG/G$
Exon 3	AAG/GTGATT	CTTATTCCAG/G
Exon 6	CAG/GTAATG	TATTTTTCAG/T
Exon 7	CAG/GTAAGA	CGTGTTGTAG/G
Exon 8		CTTTTCTCAG/A

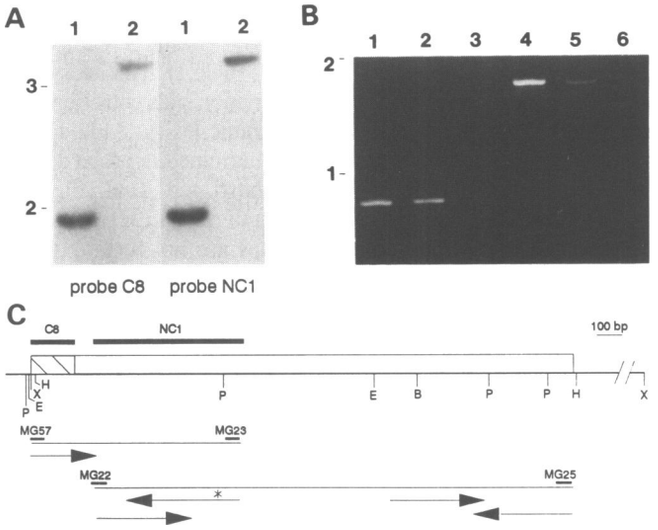

Figure 4. Exon 8 is comprised of both coding and noncoding regions. A. Southern analysis of genomic DNA digested with HindIII (lane 1) or XbaI (lane 2) and hybridized with radiolabelled probes for exon 8 (C8) or the non-coding region (NC1) of c-*ski* cDNAs. Size markers are indicated in kb at the left. B. Products amplified from genomic DNA (lanes 1 and 4), cDNA (lanes 2 and 5), or no template (lanes 3 and 6) were analyzed in an ethidium-bromide stained agarose gel. Lanes 1–3, amplification with primers MG57 and MG23; lanes 4–6, amplification with primers MG22 and MG25. C. Diagram of exon 8 indicating coding region with diagonal lines and noncoding region in open box. Probes C8 and NC1 are indicated in heavy bars above. Regions amplified are indicated below with appropriate primers. Arrows indicate direction and extent of sequencing of products amplified from genomic DNA. Asterisk designates position of the 25-bp insertion found in genomic DNA. Symbols as above.

this region of the genome (data not shown). Thus, we conclude that the coding region of proposed exon 7 is actually comprised of two exons, which we have designated exon 7 and exon 8.

Exon 8 contains coding and noncoding sequences

The c-*ski* open reading frame in cDNA clones is followed by a noncoding region extending more than 1.7 kb (5). We next wanted to determine the genomic organization of the noncoding region and to analyze the relationship between the coding and the noncoding exons in the c-*ski* locus. Therefore, genomic DNA was digested with HindIII or XbaI and hybridized with probes specific for exon 8 (C8) and the proximal noncoding region (NC1) (Fig. 4). Results indicated that the noncoding region was closely linked, or perhaps even contiguous, to the newly designated exon 8 on a 2.0-kb HindIII fragment and a 3.8-kb XbaI fragment (Fig. 4A, lanes 1 and 2).

The organization of this region of the c-*ski* locus was examined in greater detail using amplification primers for two overlapping

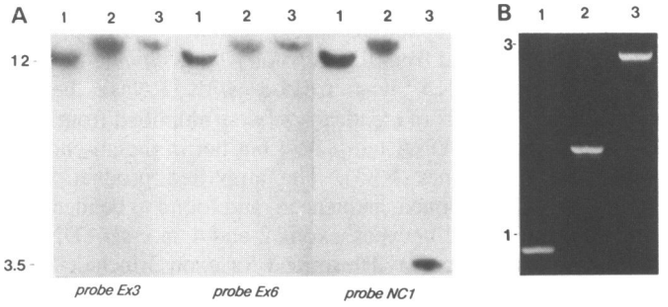

Figure 5. Exons 3, 6, and 8 are closely linked. A. Genomic DNA digested with BamHI (lane 1), KpnI (lane 2), or XbaI (lane 3) was hybridized with probes for exon 3 (Ex3), exon 6 (Ex6), or the noncoding portion of exon 8 (NC1) (see Figs. 2 and 4). B. Ethidium-bromide stained products resulting from amplifications between exons in genomic DNA: lane 1, exon 3 to 4 with primers MG48 and MG59; lane 2, exon 5 to 6 with primers MG63 and MG21; lane 3, exon 6 to exon 7 with primers MG20 and MG62.

segments. Primers MG57 and MG23 flanked cDNA sequences extending from the 5′ end of exon 8 through the proximal portion of the noncoding region. Primers MG22 and MG25 were complementary to sequences at the proximal and distal termini of the noncoding region, respectively. Amplification from both cDNA and genomic DNA templates resulted in identical products of approximately 800 bp or 1.7 kb, according to which primer pairs were used (Fig. 4B). All amplified products hybridized with the NC1 probe (data not shown).

The 800-bp fragment amplified from the genome with primers MG57 and MG23 was cloned and sequenced to show that the c-*ski* coding region in exon 8 was juxtaposed to the noncoding region precisely as found in embryonic cDNAs. However, the proximal portion of the noncoding region amplified from the genome, although very similar, was not identical to sequences determined from cDNA clones (5). We identified an additional 25 bp (CCAGCCGTCTCGGAAGCCATCACAG) inserted between nucleotides 3005 and 3006 within a T-rich region in the c-*ski* cDNA sequence (5) (Fig. 4C). In addition, we found a transversion of G to C at nucleotide 2594, a deletion of T at position 2653, and a transition from T to C at nucleotide 2658. None of the sequence differences altered the noncoding region by introducing an open reading frame. It is unlikely that the sequence variations in the amplified NC1 region resulted from anomalies of Taq polymerase activity because we found no sequence discrepancies in the distal 664 nucleotides of the noncoding region between amplified and cDNA sequences. Moreover, the same sequence variation, including the 25-bp insertion, was found in c-*ski* cDNAs that we isolated from conventional cDNA libraries (Grimes and Goodenow, unpublished data).

From these studies it is clear that exon 8 is comprised of both coding and noncoding sequences and that the entire 3′ noncoding region is contiguous in the genome. Consequently, exon 8 extends a minimum of 2 kb in the c-*ski* locus, making it the largest c-*ski* exon identified.

Exons 3 through 8 are closely linked in the genome

Once the exon organization of c-*ski* cDNA sequences was established, we examined the linkage of exons 3, 6, and 8 to exons previously defined from genomic clones (10). Southern analysis using exon-specific probes indicated that exons 3, 6, and 8 mapped to a single KpnI fragment that is larger than 20 kb,

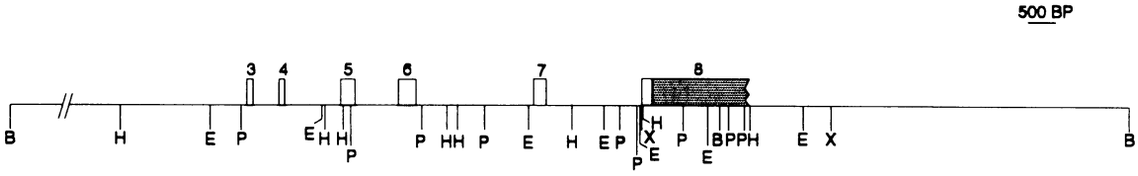

Figure 6. Organization of exons 3 through 8 in the chicken c-*ski* locus. Coding regions are indicated in open boxes while the 3′ noncoding portion of exon 8 is stippled. Broken line at distal boundary of exon 8 indicates that the 3′ terminus has not been defined precisely (5). Symbols for restriction endonuclease sites as in previous figures.

and a BamHI fragment of approximately 15 kb (Fig. 5A, lanes 1 and 2). Exons 3 and 6 both mapped to an XbaI fragment distinct from the 3.5-kb XbaI fragment containing exon 8 (Fig. 5A, lanes 3). This was expected from the location of an XbaI site in the intron just 5′ of exon 8 (see Fig. 3C). Results of additional Southern analysis indicated that exons 3 through 8 were not widely dispersed in the genome and suggested that organization of the exons might be assessed by amplification of genomic DNA.

Linkage between exons 3 and 4 was determined by amplification from genomic DNA using primers MG48 and MG59, complementary to the 5′ boundary of exon 3 and the 3′ boundary of exon 4, respectively. The resulting product was approximately 650 bp (Fig. 5B, lane 1). The fragment hybridized with an exon 3-specific probe, thereby positioning exon 3 within 500 bp of exon 4. When we determined the sequence upstream of exon 4 in a plasmid subclone of the region kindly provided by Dr. Ed Stavnezer, exon 3 was located 449 bp 5′ of exon 4, flanked by consensus splice acceptor and splice donor sequences (Table 1).

To map exon 6 precisely, chicken DNA was amplified from exon 5 to exon 6 with primers MG63 and MG21, or from exon 6 to exon 7 with primers MG20 and MG62. The products resulting from each reaction were 1.3 or 2.6 kb, respectively (Fig. 5B, lanes 2 and 3). Both products were within size ranges predicted by our Southern analysis and hybridized with an exon 6-specific probe devoid of primer sequences (data not shown). The amplified fragments were cloned and partially sequenced to identify concensus splice donor and acceptor sequences at the intron/exon boundaries flanking exon 6 and preceding exon 7 (Table 1). Additional exons do not appear to be located within this region of the genome because probes derived from the introns failed to hybridize to transcripts in Northern analysis (data not shown).

Results of both Southern analysis and enzymatic amplification mapped c-*ski* exons 3 through 8 to a relatively small region in the chicken genome (Fig. 6). It is apparent that most of the genomic sequences comprising chicken c-*ski* cDNAs, including approximately half of the open reading frame, are closely linked within 10 kb. Although we located exon 3 and linked it to exon 4, we have been unable to establish the linkage between exon 3 and exon 2 by Southern analysis of genomic DNA with exon-specific probes (data not shown). However, our results suggest that there is a minimum of 25 kb of genomic sequence between exons 2 and 3.

DISCUSSION

Our studies have established that sequences contained within c-*ski* cDNAs are organized as eight exons in the c-*ski* locus. We have defined the genomic structure of four of these exons, exons 6, 7, 8, and the elusive exon 3, and identified their boundaries with intron sequences. The exons defined by our analysis comprise almost two-thirds of c-*ski* cDNA sequences.

It is essential to establish the exon structure of the chicken c-*ski* locus because different c-*ski* cDNAs have been isolated, indicating that alternative splicing of c-*ski* transcripts is likely to occur. Although there is evidence that the entire ski gene has yet to be identified, protein isoforms encoded by c-*ski* cDNAs containing exons 1 through 8 display striking biological activities in both proliferation and differentiation (7,15). In addition to cDNAs containing eight exons, cDNAs without sequences encoded by exon 2 or exon 6 have been isolated (5). There was unequivocal evidence that exon 2, one of the four exons originally defined from genomic clones, is alternatively spliced in chicken c-*ski* transcripts (5,10). In contrast, the boundaries and alternative splicing of exon 6 were inferred from the structure of a single cDNA clone, even though c-*ski* transcripts without exon 6 were undetectable in RNAs from chick embryos (5). It was not clear whether a c-*ski* cDNA without exon 6 was an artefact of cloning or representative of an extremely rare c-*ski* transcript. This is significant because splicing of exon 6 introduces a translation termination codon at the junction between exon 5 and exon 7. Transcripts without exon 6 would encode a truncated Ski protein similar to the protein encoded by v-*ski* (Fig. 1). Our studies provide direct evidence that proposed exon 6 is indeed an exon by identifying splice junction sequences flanking the exon and by mapping the exon within the c-*ski* locus. Thus, it is likely that c-*ski* cDNAs without exon 6 reflect authentic, but rare c-*ski* mRNAs. Identification of temporal or tissue specific regulation of differential exon 6 splicing could provide insights into the function(s) of endogenous c-*ski*.

Likewise, there is significance in determining that proposed exon 7 in chickens is organized as two exons. There is more than 90% identity in the deduced carboxy-terminal sequence of Ski proteins encoded by chicken or human c-*ski* genes (22). The carboxyl terminus of the human Ski protein has been implicated in DNA binding by deletion analysis (23). Although exon organization within the human c-*ski* locus has yet to be determined, alternative splicing of exon 7 from chicken c-*ski* transcripts would delete more than half of the sequences encoding the putative DNA binding domain.

Sequence motifs implicated in reduced half-life of RNAs are found in the long 3′ untranslated region of c-*ski* cDNAs (5), raising the possibility that ski mRNAs could be regulated, at least in part, by differential splicing of 3′ noncoding regions. Our results establish that the long 3′ noncoding region in c-*ski* cDNAs is part of exon 8 in the genome. Such a genomic organization diminishes the likelihood that c-*ski* transcripts encoding different isoforms of Ski proteins will be distinguishable by alternatively spliced noncoding regions.

In addition to defining the structure of four exons, we were able to map their location within the ski locus. In view of the expanse of the c-*ski* locus in chickens, it was unexpected, but fortuitous, that 3 kb of ski cDNAs encoded by exons 3 through 8 were localized within 10 kb of the genome. The close proximity of the six exons provided an opportunity to define their genomic organization using PCR amplification. Using the probes specific for exons 6, 7, and 8 there was no indication for the presence of genes related to c-*ski*, at least at the stringency of the Southern analysis. Moreover, from our analyses of c-*ski* mRNA, it appears unlikely that additional exons, as part of the major c-*ski* transcripts, map within this region of the chicken ski locus. Studies of c-*ski* transcripts are in progress to assess directly differential exon usage and to examine their relationship to c-*ski* cDNAs containing different exons.

ACKNOWLEDGEMENTS

The authors are grateful to M. Giovannella Moscovici, Carlo Moscovici, Michael Ambrose, Maurice Swanson, and William Farmerie for helpful discussion and comments. This work was supported by funding from the American Cancer Society Institutional Grant; from the American Cancer Society Florida Division (Sigurd M. Johnson Research Grant and R.G. Thompson Memorial Summer Research Program), the Division of Sponsored Research at the University of Florida, and by NIH T32 CA09126.

REFERENCES

1. Bingham,P., Chou,T.-B., Mims,I., and Zachar,Z. (1988) Trends Genet. 4, 134–140.
2. Smith,C.W., Patton,J.G., and Nadal-Ginard,B. (1989) Annu.Rev.Genet. 23, 527–577.
3. McKeown,M. (1990) Genet.Eng. 12, 139–181.
4. Maniatis,T. (1991) Science 251, 33–34.
5. Sutrave,P. and Hughes,S.H. (1989) Mol.Cell.Biol. 9, 4046–4051.
6. Sutrave,P., Copeland,T.D., Showalter,S.D., and Hughes, S.H. (1990) Mol.Cell.Biol. 10, 3137–3144.
7. Colmenares,C., Sutrave,P., Hughes,S.H., and Stavnezer,E. (1991) J.Virol. 65, 4929–4935.
8. Stavnezer,E., Gerhard,D.S., Binari,R.C., and Balazs,I. (1981) J.Virol. 39, 920–934.
9. Stavnezer,E., Barkas,A.E., Brennan,L.A., Brodeur,D., and Li,Y. (1986) J.Virol. 57, 1073–1089.
10. Stavnezer,E., Brodeur,D., and Brennan,L. (1989) Mol.Cell.Biol. 9, 4038–4045.
11. Li,Y., Turk,C.M., Teumer,J.K., and Stavnezer,E. (1986) J.Virol. 57, 1065–1072.
12. Barkas,A.E., Brodeur,D., and Stavnezer,E. (1986) Virology 151, 131–138.
13. Colmenares,C. and Stavnezer,E. (1989) Cell 59, 293–303.
14. Colmenares,C., Teumer,J.K., and Stavnezer,E. (1991) Mol.Cell.Biol. 11, 1167–1170.
15. Sutrave,P., Kelly,A.M., and Hughes,S.H. (1990) Genes Dev. 4, 1462–1472.
16. Sambrook,J., Fritsch,E.F., and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2nd ed.
17. Southern,E.M. (1975) J.Mol.Biol. 98, 503–517.
18. Feinberg,A.P. and Vogelstein,B. (1984) Anal.Biochem. 137, 266–267.
19. Church,G.M. and Gilbert,W. (1984) Proc.Natl.Acad.Sci. USA 81, 1991–1995.
20. Goodenow,M.M., Huet,T., Saurin, W., Kwok,S., Sninsky,J., and Wain-Hobson,S. (1989) J.Acquired Immune Defic.Syndr. 2, 344–352.
21. Shapiro,M.B. and Senapathy,P. (1987) Nucleic Acids Res. 15, 7155–7174.
22. Nomura,N., Sasamoto,S., Ishii,S., Date,T., Matsui,M., and Ishizaki,R. (1989) Nucleic Acids Res. 17, 5489–5500.
23. Nagase,T., Mizuguchi,G., Nomura,N., Ishizaki,R., Ueno,Y., and Ishii,S. (1990) Nucleic Acid Res. 18, 337–343.