Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 313 (2004) 135–141
BBRC
www.elsevier.com/locate/ybbrc

Last intron of the chemokine-like factor gene contains a putative promoter for the downstream *CKLF* super family member 1 gene☆

Mingxu Xu,[a,*] Wenling Han,[a] Mingping Qian,[b] Xiaotu Ma,[b] Peiguo Ding,[a] Ying Wang,[a] Donglan Xia,[a] Min Rui,[a] Lu Wang,[a] Yinmei Zhang,[a] and Dalong Ma[a]

[a] *Laboratory of Medical Immunology, School of Basic Medical Science, Peking University, 38 Xueyuan Road, Beijing 100083, PR China*
[b] *School of Mathematics, Peking University, Beijing 100871, PR China*

Received 27 October 2003

Abstract

The genes for chemokine-like factor (*CKLF*) and four chemokine-like factor super family members (*CKLFSF1–4*) are tightly linked on chromosome 16, with only 325 bp separating *CKLF* and *CKLFSF1*. We used Northern blotting and RT-PCR to show that these two genes are expressed independently of one another. We then used a novel computational promoter prediction method based on the interaction among transcription factor binding sites (TFBSs) to identify a putative promoter region for the *CKLFSF1* gene. Our method predicted a promoter region in the last intron of the upstream gene, *CKLF*. We PCR amplified the predicted promoter region and used a luciferase assay to show that the region was able to drive the luciferase gene. DNA decoy experiments indicated that 214 bp fragment neighboring the TATA box markedly inhibited *CKLFSF1* gene expression. Sequence analysis of the region revealed a typical TATA box (TATATAA) and multiple potential transcription factor binding sites, providing further evidence for this being a functional promoter for *CKLFSF1*. This work provides the first evidence of a promoter from one gene located in an intron of another.

Keywords: Chemokine-like factor; Chemokine-like factor super family; Luciferase assay; Intron

Beyond their involvement in splicing reactions, some introns contain enhancer-like or silencer-like elements that play an important role in transcriptional regulation. Intronic regulatory sequences have been identified in genes related to cell growth [1], granulocyte development [2], and tissue-specific gene expression [3]. Within introns, AP-1 sequences [4,5], glucocorticoid response elements [6,7], and thyroid responsive elements [8] play important regulatory roles in gene transcription. Furthermore, some gene promoters are located in downstream intronic regions. For example, a functional eukaryotic promoter is contained within the first intron of the hGH-N coding region [9]; the mouse mdm2 gene contains two promoters: one located upstream of the gene that is active in the absence of p53 and one within the first intron that requires p53 for transcriptional activity [10]; and of the two promoters of the human BMP-4 gene, one is located upstream of exon 1 and the other is located upstream of exon 2, within the first intron [11]. However, to our knowledge, our discovery that the *CKLFSF1* promoter overlaps the last intron of the *CKLF* gene is the first report of a promoter region for one gene being found in an intron of an upstream gene.

Chemokine-like factor (CKLF) is novel cytokine that our laboratory cloned from U937 cells using Suppression Subtractive Hybridization [12]. Further, we used in silico gene identification from human genome sequence and EST databases to identify eight additional novel genes sharing sequence identity with *CKLF* [13]. These genes were designated chemokine-like factor super family 1–8 (*CKLFSF1–8*) by the International Human Gene Nomenclature Committee. One of the interesting aspects of these genes is that *CKLF*, *CKLFSF1*, *2*, *3*, and *4* exist as a gene cluster on chromosome 16, and the *CKLF* and *CKLFSF1* genes are very closely linked

☆ This work was supported by Grants from the National Natural Sciences Foundation of China (30170871).
* Corresponding author. Fax: +86-10-8280-1149.
E-mail address: xumingxu@bjmu.edu.cn (M. Xu).

0006-291X/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2003.11.100

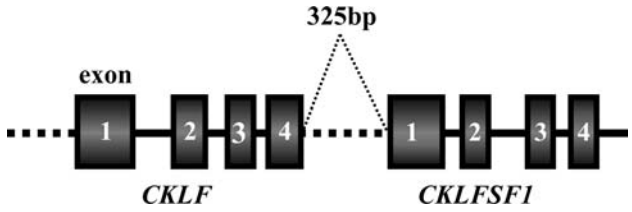

Fig. 1. Schematic representation of the *CKLF* and *CKLFSF1* gene structure. The boxes show the exons with their relative size and the positions in the genes. Numerals in the boxes indicate exon numbers.

(within 325 bp of each other) (Fig. 1). Recently, we have discovered that *CKLFSF1* is a complicated gene with at least 23 splice variants and it is expressed at particularly high levels in testes (submitted data). In this study, Northern blot and RT-PCR analyses of *CKLF* and *CKLFSF1* demonstrated that the two genes are independent, leading us to question where the regulatory elements of *CKLFSF1* are located. We used a recently developed method for recognizing compositional features of promoter regions (manuscript in preparation) to predict the *CKLFSF1* gene promoter region. The results indicated that a potential promoter region for the *CKLFSF1* gene was located within the last intron of the upstream gene, *CKLF*. We PCR amplified the predicted promoter fragment and inserted it into a promoterless luciferase reporter gene vector, which we then transfected into several kinds of cell lines. The results of the luciferase assay clearly demonstrated that the *CKLFSF1* upstream region displayed promoter activity in some cell lines. Decoy assay indicated that 214 bp fragment neighboring the TATA box markedly inhibited *CKLFSF1* gene expression. These suggest that the predicted promoter is functional and represents the first report of a gene promoter being located within an intron of an upstream gene.

Materials and methods

Northern blot analysis. The full-length cDNAs for human *CKLF* and *CKLFSF1* were labeled with [α-^{32}P]dCTP using a random hexamer labeling kit (Amersham). The probes were hybridized to a multiple tissue Northern blot following the manufacturer's instructions (Clontech). The blot was washed with 0.1× SSC at 68 °C and exposed to Cyclone (Packard, Japan) for 3 h at room temperature.

Computational identification of the promoter region of the human CKLFSF1 gene. We used the following novel method to recognize compositional features of the *CKLFSF1* promoter region. First, we identified all transcription factor binding sites (TFBSs) between 5 and 12 bp in length as predicted by the TRANSFAC35 program [14]. We then counted the formed in the learning promoter dataset (565 sequences of experimentally verified promoters) [15] and in the control dataset (890 coding sequences) [Compiled and provided by M. Reese (LBNL, mgreese@lbl.gov), D. Kulp (UCSC, dkulp@cse.ucsc.edu), A. Gentles (Stanford), and U. Ohler (UCB, ohler@fruitfly.berkeley.edu)]. Next, the TFBS pairs that were >4 times more frequent in the learning dataset than in the control dataset were used for our 'dictionary' of possible promoters. The *potential functioning score* for each pair was calculated from the ratio of learning dataset frequencies to control dataset frequencies. Next, each learning and control sequence was compared to our dictionary, and the potential functioning scores of the pairs found in the sequence were summed together to create a *promoter-potential score*. Finally, a threshold was set to equalize the misclassification rates for both promoter and control sequences.

Cloning of the promoter region of CKLFSF1. The putative *CKLFSF1* promoter region was PCR amplified using primers designed from genomic sequences flanking the identified region. The forward primer contained an *Nhe*I restriction site (underlined, see below). The reverse primer was designed to include an arbitrary sequence (in lowercase) that added an *Nhe*I restriction site (underlined) to the amplified product. Human genomic DNA was isolated from HeLa cell lines according to standard protocol. PCRs were performed in 50 μl reaction buffer (50 mM KCl, 10 mM Tris, pH 8.5, and 1.5 mM $MgCl_2$) with 250 ng genomic DNA, 10 mM dNTPs, 50 pmol of each primer, and 2.5 U LA *Taq* DNA polymerase (Takara, Japan). The amplification conditions were as follows: 4 min at 94 °C, followed by 35 cycles of 94 °C for 30 s; 60 °C for 30 s; and 72 °C for 40 s, followed by 72 °C for 7 min. The amplification products were separated on a 1% agarose gel and visualized by ethidium bromide staining.

Forward primer: 5′-GTATAGTAACCATAGTTGCTAGCC-3′.

Reverse primer: 5′-ctagctAGCCACTGCCTCCTAGGTTCTG-3′.

Construction of luciferase reporter vector. To prepare the *CKLFSF1* promoter region reporter gene construction, the PCR fragment from above was *Nhe*I digested and cloned into the *Nhe*I site of the pGL3-Basic vector (Promega) to obtain plasmids pGL3-CKLFSF1. The vector lacks a promoter but contains the entire luciferase gene, SV40 T intron, and polyadenylation signal. DNA sequencing was performed by the dideoxynucleotide chain-termination method [16] using an ABI automated DNA sequencer (3100 Genetic Analyzer). The nucleotide sequences were performed on both DNA strands of the reporter plasmids using sense and antisense vector-specific primers to confirm the sequence and orientation. Sequencing data were BLAST compared with the appropriate human genomic sequence.

Cell culture, transfections, and luciferase assay. HeLa and 293T (SV40 large T antigen) cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco) supplemented with 10% fetal calf serum (FCS, Hyclone), 100 U/ml penicillin G sodium, and 100 mg/ml streptomycin sulfate. A549, MCF7, and HT29 cells were maintained in RPMI 1640 medium supplemented with 10% FCS and 2 mM glutamine, 100 U/ml penicillin, and 100 mg/ml streptomycin sulfate. All cell lines were grown in 100-mm dishes containing 15 ml of the medium and incubated at 37 °C in 5% CO2. For transient transfections, the cells were seeded into 24-well plates at a density of 4×10^4 cells/well in the medium and incubated overnight. Two hours prior to transfection, the medium was renewed with the medium. Cells were transfected by the calcium phosphate precipitation method [17]. Transfection reactions contained a total of 1.2 μg of plasmid DNA/well, consisting of 1 μg of reporter plasmid and 0.2 μg pEGFP, which was used to correct for transfection efficiency. After transfection, cells were incubated in fresh medium for an additional 24 h. Cells were then washed twice with calcium/magnesium-free phosphate-buffered saline and incubated in 50 μl of cell lysis buffer at room temperature for 15 min. Ten microliters of cell lysate was mixed with 50 μl of Luciferase Assay Reagent (Promega) at room temperature and the light intensity of the samples was counted immediately in a POLARstar galaxy spectrometer (BMG Labtechnologies) for 10 s at 25 °C. Each transfection was performed thrice and the measurements were normalized with equivalent quantities of GFP.

Production of decoy fragment and stability analysis. DNA decoy fragment was produced by PCR amplification using the *CKLFSF1* sequence-specific sense primer (5′ CTAGCTAGCTGGGATTACAG GCGTGAG 3′) and antisense primer (5′ CTAGCTAGCTTATATA CAGCGGTTTC 3′) and pGL3-CKLFSF1 plasmid DNA as template. As control DNA an unrelated 249 bp plasmid DNA fragment generated by PCR on the pGEM-plasmid was used.

The 'in solution' stability of PCR-generated DNA fragments was analyzed incubating 300 ng DNA at 37 °C for 48 h in water, DMEM, DMEM supplemented with 10% FBS, and HeLa cell conditioned medium. The cell conditioned medium was obtained after the harvesting of HeLa cells grown to confluence. After incubation, DNA samples were analyzed by electrophoresis in 1% agarose gel.

Electroporation. HeLa cells which were maintained confluent for 2 days were used for these experiments. The cells were trypsinized and washed with 10 ml RPMI 1640 medium. After centrifugation the pellets were resuspended in RPMI 1640 medium in a volume to give 10^7 cells/ml. Cells (200 μl) were placed in a 4-mm electroporation cuvette (Eurogentec) and 10 μg (20 μl) of DNA solution was added. The cells were electroporated with a single pulse (120 V, 20 mS) with a Division of Genetronics electroporator system. The cells were kept in the electroporation cuvette for 5 min. The cells were then seeded into 6 cm plates with DMEM containing 10% FCS, incubated for 48 h, and then harvested.

RT-PCR analysis. Cells were collected by scraping. Cell pellets were washed with PBS and total RNA was isolated as described above. cDNA was synthesized from 3 μg of total RNA using ThemoScript RT-PCR system (Invitrogen) and random primers. The resulting cDNA products were used to amplify fragments of *CKLFSF1* and G3PDH cDNA in 2.5 U LA *Taq* DNA polymerase (Takara), on GenAmp PCR System 9700 (PE Applied Bio Systems) with 30 cycles for *CKLFSF1* and 25 cycles for G3PDH cDNA. Following primers and conditions were used:

CKLFSF1: F = 5′ AATGAGTCATTTATAACAATCACAAG 3′
R = 5′ CAATGTAAATAGGTCAGCAAGTGGTG 3′

The cycles were 30 s at 94 °C, 30 s at 58 °C, and 30 s at 72 °C;

G3PDH: F = 5′ CTAGCTAGCTGGGATTACAGGCGTGAG 3′
R = 5′ CTAGCTAGCTTATATACAGCGGTTTC 3′

The cycles were 30 s at 94 °C, 30 s at 58 °C, and 60 s at 72 °C.

The amplification products were separated on 1% agarose gel and visualized by ethidium bromide staining. The PCR product was subcloned into pGEM-T Easy vector (Promega) and subjected to DNA sequencing by the dideoxynucleotide chain-termination method.

Results

Northern blot analysis of CKLF and CKLFSF1

To identify whether *CKLF* and *CKLFSF1* are two independent genes, the complete ORF of each gene was used as template for probe labeling (Fig. 2). The *CKLF* (459 bp) probe identified a band at 0.6 bp and the *CKLFSF1* (503 bp) identified a transcript of 1.4 bp. There was no cross-reactivity, suggesting that the two transcripts are independent of each other. In addition, RT-PCR with an upstream primer within the *CKLF* sequence and a downstream primer within the *CKLFSF1* sequence produced no specific PCR product (data not shown). Taken together, these results suggest that *CKLF* and *CKLFSF1* are two independent genes.

Computational identification of the promoter region of the human CKLFSF1 gene

We used a novel method to predict the *CKLFSF1* gene promoter region. The promote potential score of a 300-bp sliding window is shown in Fig. 3. The largest peak prior to the first exon of *CKLFSF1* is predicted to be the promoter region.

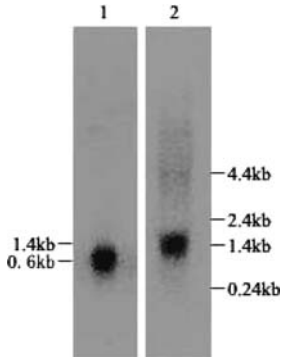

Fig. 2. Northern blot analysis of expression of *CKLF* and *CKLFSF1* mRNA. Commercially purchased multiple-tissue Northern blot filters were hybridized with the probe corresponding to the complete ORF encoding human *CKLF* (lane 1) or *CKLFSF1* (lane 2) successively.

Sequence analysis of an 874 bp PCR fragment including the putative promoter (Fig. 4) revealed that the region contains a typical TATA box (TATATAA) [18] from −197 to −191. In addition, this region contains multiple bHLH binding site sequences (consensus CANNTG) [19]. Several other potential binding sites for transcription factors were identified by sequence comparison, including two AP1 sites [20,21] at nt −292 to −299 and −829 to −836; a GATA1 binding site found at nt −719 to −724 [20]; and an AP2 binding site [22,23] located between nt −622 and −631. The presence of these multiple transcription factor binding sites in the 5′-upstream sequence further suggested that the fragment may contain the functional promoter.

Cloning of the promoter region of the human CKLFSF1 gene

We PCR amplified the potential *CKLFSF1* promoter region from human genomic DNA. The resulting 874 bp fragment was cloned into the promoterless pGL3 reporter plasmid (Promega) in the forward orientation to give pGL3-CKLFSF1. Sequencing with vector-specific primers confirmed the *CKLFSF1* genomic DNA sequence (GenBank Accession No. AC010542).

Functional characterization of the putative promoter

We used a luciferase assay to determine the functionality of the putative *CKLFSF1* promoter. The test construct (pGL3-CKLFSF1), the positive control construct (pGL3-SV40), and the negative control construct (pGL3-Basic) were individually transfected into human cervical carcinoma (HeLa, ATCC), human lung carcinoma (A549, ATCC), human breast carcinoma (MCF7, ATCC), human colon carcinoma (HT29, ATCC), and human embryonic kidney [293T (expressing SV40 large

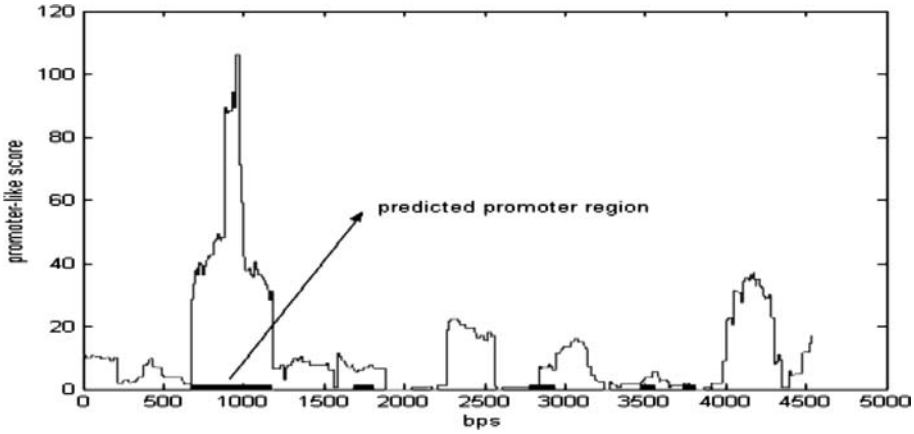

Fig. 3. Moving promoter-potential score of the *CKLFSF1* gene sequence. The highest peak was predicted to be promoter due to the fact that there are four coding exons (indicated by bold black lines) following it.

```
-874                                    ccac atggggctac tgagcactgg aaatgtggct -841
                                        bHLH
-840 agtctgaatc aagatcagct gtaagtataa aaatacacat ttaatttcaa agatttatta -781
       AP1       bHLH                                    TCF1
-780 cccccaacaa ggaatgtaga atattttgtt aataactttg ttgattacat gttaaatgat -721
                                                                GATA-1
-720 aatatttttag acatactgag ctaaatatat taaaatctat gacttatttt tttctttttt -661

-660 ttttttttct gagacagagt ctgtctctgt cgcccaggct ggtggagtgc agtggtgcta -601
                                        AP2
-600 tctcagctca ctgcaagctc tgcctcccgg gttcacgcca ttctcctgcc tcagcctccc -541
       CAP                                 CAP                CAP
-540 aagtagctgg gactacaggc atccgccacc atgcctggct aatttttgt attttagta -481

-480 gagacggagt ttcaccgtgt taacaaggat ggtctcgatc ttccgacgtc gtgatccgcc -421

-420 tgcctcggtc tcccaaagtg ctgggattac aggcgtgagc caccgcgccc ggccgacttt -361
                TCF1                        ZESTE    GCF
-360 tttctttaca tttataatgt ggctgttaaa aaatctttaa ttatgtggct gtgttacact -301
       gamma_IRE                           gamma_IRE
-300 ggtgactcat attatgctgc tttcagacag cactgactag agtgtgagcc ccgtcctcaa -241
       AP1                                 LPB
-240 gttgcttaga gcttaacaca ggagacacaa gtggaaaccg ctgTATATAA gctgttgaca -181
                TCF1                    bHLH
-180 gggagctgca tggtgcagaa aagaggaaac tacctctctg gcagtggaat tgctgtttgc -121
                        TCF1
-120 tgagaataaa ggaaaaaaaa tggaaaaggt tccaggtaga ggagccaata gtgtacagag  -61
                                            NFI
-60  tgccagtcat gaaaagattg gtaggaggta ttcagtgtgg tcagaaccta ggaggcagtg   -1
```

Fig. 4. Nucleotide sequence of the upstream region of *CKLFSF1* gene. Potential *cis*-acting sequences are *underlined*. A TATA-like sequence is in *boldface*. The region includes several kinds of potential *cis*-acting regulatory elements including multiple bHLH binding sites and AP1 binding sites, AP2 binding site and several other elements (underline). Nt −1 is in front of the last nucleotide of the cloned genomic DNA.

T antigen), ATCC] cell lines, along with the pEGFE plasmid for internal control. The cells were harvested 24 h after transfection and luciferase activities were determined by standard procedures. GFP activity was used to normalize for variations in transfection efficiency. All experiments were performed three times and normalized for transfection efficiency. The luciferase assay results (Fig. 5) demonstrated that the *CKLFSF1* upstream region fragment displayed promoter activity in HeLa, A549, and MCF7 cells but not in 293T and HT29 cells. In contrast, no promoter activity was observed with transfection of the negative control construct pGL3-Basic. These data suggest that the cloned upstream region of the *CKLFSF1* gene may contain a functional promoter to drive *CKLFSF1* gene expression in some cell lines.

Decoy assay

To confirm if the cloned upstream region of the *CKLFSF1* could drive *CKLFSF1* gene expression in vivo, the decoy technique was employed. As much as 214 bp fragment neighboring the TATA box (test DNA sample) and 249 bp unrelated plasmid DNA fragment

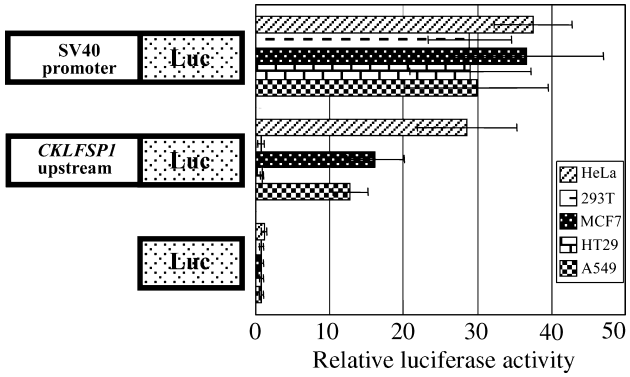

Fig. 5. Promoter activity of *CKLFSF1* gene 5′-upstream region. The left and right panels show the structures of the constructs and the results of luciferase assay, respectively.

(control DNA sample) were amplified by PCR (Fig. 6A). First, to assess stability of PCR-generated DNA fragments in different experimental condition, test DNA sample and control DNA sample were incubated in water, DMEM, DMEM plus 10% FBS, and cell conditioned medium, subsequent analysis by electrophoresis. As shown in Fig. 6B, both DNA fragments were stable for at least 48 h as shown by the presence of electrophoretic bands corresponding to the original molecular weight.

In addition, morphology of HeLa cells treated with the PCR products did not differ from untreated cells when examined under a phase-contract microscope (data not shown).

DNA decoy experiments were performed on HeLa cells treated with PCR production using electroporation. Effects on *CKLFSF1* gene expression were studied by RT-PCR analysis. The same amount of RNA isolated from test and control cells was reverse transcribed with random primer. The cDNA obtained was amplified by PCR using *CKLFSF1* sequence-specific primers and G3PDH sequence-specific primers. G3PDH PCR amplification was used as an internal control. The results was indicated in Fig. 6C. The 104 bp of DNA fragment was amplified by PCR. It is confirmed that this fragment is a part of *CKLFSF1* cDNA by DNA sequencing. The treatment of HeLa cells with 214 bp decoy resulted in a significant inhibition of *CKLFSF1* mRNA compared with unrelated 249 bp fragment, while G3PDH gene expression is essentially unaffected, suggesting that the effect of this DNA decoy was considered specific. These results clearly indicate that the cloned upstream region of the *CKLFSF1* gene can drive *CKLFSF1* gene expression.

Discussion

Here, we report the first identification of the promoter of one gene being located within a neighboring gene. This identification was made through use of a new method for promoter identification that is based on the fact that transcription factor binding sites (TFBSs) appear in higher density in promoter sequences than in other regions [24]. There are several existing programs that use the high density of TFBSs in promoters as a method for finding promoter regions [25], but our method improves upon these by assuming that promoter function involves the interaction of a group of transcription factors corresponding to different binding sites. Using the method, we identified a putative promoter for

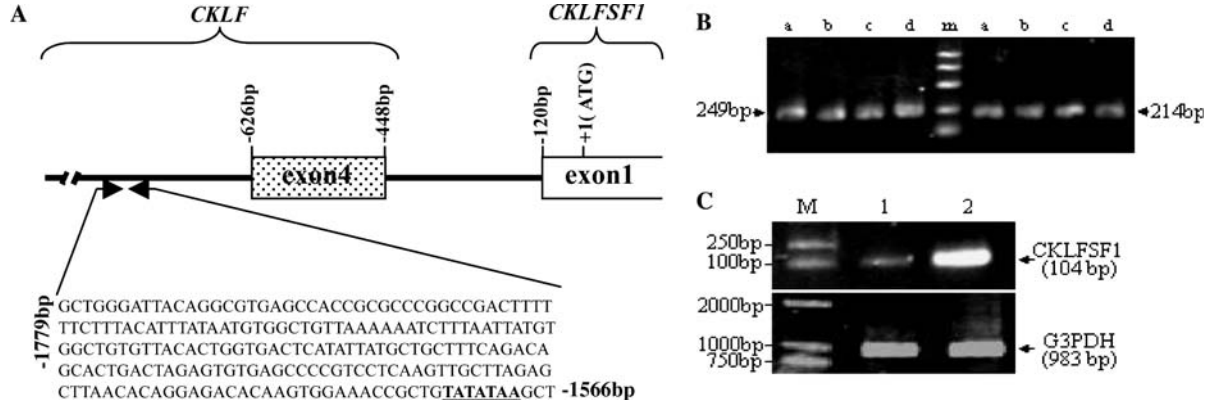

Fig. 6. (A) Scheme of the 5′-upstream region of *CKLFSF1* gene. The position of exon 1 and the upstream region of *CKLFSF1* are shown. The location and sequence of decoy fragment are indicated. A TATA-box sequence is in *boldface*. (B) Analysis of the stability of PCR-generated fragments. The 214 and 249 bp PCR products were incubated at 37 °C for 48 h in (a) water, (b) DMEM, (c) DMEM plus 10% FBS, and (d) HeLa cell conditioned medium, subsequent analysis by 1% electrophoresis. m, Molecular weight marker (DL 2000). (C) Effect of PCR-generated DNA decoy on modulation of gene expression. Representative electrophoretic gel of RT-PCR products obtained from total RNA of HeLa cells was shown. Cells were treated with 214 bp *CKLFSF1*-specific decoy (lane 1) or 249 bp unrelated decoy (lane 2). Expression of *CKLFSF1* mRNA was visualized by ethidium bromide standing. M, molecular weight marker (DL 2000). Specific RT-PCR products were arrowed. G3PDH PCR amplification was used as an internal control.

the *CKLFSF1* located in the last intron of upstream gene, *CKLF*.

Sequence analysis of the potential promoter region revealed that the region contains a typical TATA box (TATATAA), as well as multiple potential transcription factor binding sites including bHLH, AP1, AP2, and GATA1 binding sites. These results suggest that the region could be a functional promoter region and prompted us to test the promoter region in a functional assay. According to the result of computational predicted promoter region of the *CKLFSF1*, we have cloned and sequenced the 0.8 kb 5′-upstream region of *CKLFSF1*. Reporter assays using this region attached to a luciferase gene showed cell-specific promoter activity in HeLa, A549, and MCF7 cells, which expressed *CKLFSF1* mRNA, but not in 293T and HT29 cells, which had no detectable *CKLFSF1* mRNA by RT-PCR (data not shown), further suggesting that a functional promoter for *CKLFSF1* might be located in the last intron of upstream gene, *CKLF*. To obtain more evidence about the cloned upstream region belonging to *CKLFSF1* gene promoter, we performed decoy experiments in HeLa cells, which expressed *CKLFSF1* mRNA by RT-PCR and Northern blot (data not shown). The results have shown that the 214 bp upstream fragment neighboring the TATA box markedly inhibited *CKLFSF1* gene expression in HeLa cells. These suggest that the predicted promoter is functional and can drive *CKLFSF1* gene expression.

We believe that this is the first report of a promoter being located within the intron of another gene, though there are several examples of same-gene promoters being found within introns. This can occur in genes that have more than one promoter. Usually, one promoter is located upstream of the gene and another is found within an intron. For example, the genes for rat acetyl-CoA carboxylase gene [26], rat cAMP phosphodiesterase gene [27], human growth hormone gene [9], and human mdm2 [10], each have two promoters, one upstream of exon 1 and the second in the first intron, upstream of exon 2. Also, a promoter within intron 35 of the human C4A gene initiates abundant adrenal-specific transcription of that gene [28]. However, these examples are all within the same gene. Here, we present evidence that a functional promoter for *CKLFSF1* is located within an intron of the upstream gene, *CKLF*.

CKLF and *CKLFSF1–4* are closely linked in a high-density gene cluster on chromosome 16, with 325 and 312 bp separating the genes for *CKLF* and *CKLFSF1*, and *CKLFSF1* and *CKLFSF2*, respectively. In addition, these genes show high degrees of similarity to each other and to their mouse homologues (submitted data). Because of this, we would suggest that *CKLF*, *CKLFSF1*, and *CKLFSF2* may share an ancestral gene, which formed several independent but overlapping genes during evolution. It is interesting to speculate that promoter regions of other closely linked genes may overlap into neighboring genes and what the functional consequences of this might be.

References

[1] M. Jung, A. Dritschilo, G. Mark, U. Kasid, Identification of multiple repeat sequences and transcription-related elements within introns 4, 8 and 9 of human Raf-1, Biochem. Biophys. Res. Commun. 190 (1993) 462–469.

[2] M. Yamada, M. Yoshida, K. Hashinaka, Identification of transcriptional cis-elements in introns 7 and 9 of the myeloperoxidase gene, J. Biol. Chem. 268 (1993) 13479–13485.

[3] A.R. Brooks, B.D. Blackhart, K. Haubold, B. Levy-Wilson, Characterization of tissue-specific enhancer elements in the second intron of the human apolipoprotein B gene, J. Biol. Chem. 266 (1991) 7848–7859.

[4] H. Katai, J.D. Stephenson, C.P. Simkevich, J.P. Thompson, R. Raghow, An AP-1-like motif in the first intron of human Pro alpha 1(I) collagen gene is a critical determinant of its transcriptional activity, Mol. Cell. Biochem. 118 (1992) 119–129.

[5] R.G. Oshima, L. Abrams, D. Kulesh, Activation of an intron enhancer within the keratin 18 gene by expression of c-fos and c-jun in undifferentiated F9 embryonal carcinoma cells, Genes Dev. 4 (1990) 835–848.

[6] J.M. Mathis, W.H. Houser, E. Bresnick, J.A. Cidlowski, R.N. Hines, R.A. Prough, E.R. Simpson, Glucocorticoid regulation of the rat cytochrome P450c (P450IA1) gene: receptor binding within intron I, Arch. Biochem. Biophys. 269 (1989) 93–105.

[7] A.J. Lange, C. Espinet, R. Hall, M.R. el-Maghrabi, A.M. Vargas, R.J. Miksicek, D.K. Granner, S.J. Pilkis, Regulation of gene expression of rat skeletal muscle/liver 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Isolation and characterization of a glucocorticoid response element in the first intron of the gene, J. Biol. Chem. 267 (1992) 15673–15680.

[8] J. Sap, L. de Magistris, H. Stunnenberg, B. Vennstrom, A major thyroid hormone response element in the third intron of the rat growth hormone gene, EMBO J. 9 (1990) 887–896.

[9] A.F. Kolb, W.H. Gunzburg,, G. Brem, V. Erfle, B. Salmons, A functional eukaryotic promoter is contained within the first intron of the hGH-N coding region, Biochem. Biophys. Res. Commun. 247 (1998) 332–337.

[10] A. Zauberman, D. Flusberg, Y. Haupt, Y. Barak, M. Oren, A functional p53-responsive intronic promoter is contained within the human mdm2 gene, Nucleic Acids Res. 23 (1995) 2584–2592.

[11] A. Van den Wijngaard, M.A. Pijpers, P.H. Joosten, J.M. Roelofs, E.J. Van zoelen, W.J. Olijve, Functional characterization of two promoters in the human bone morphogenetic protein-4 gene, J. Bone Miner. Res. 14 (1999) 1432–1441.

[12] W.L. Han, Y.X. Lou, J.M. Tang, Y.M. Zhang, Y.Y. Chen, Y. Li, W.F. Gu, J.Q. Huang, L.M. Gui, Y. Tang, F. Li, Q.S. Song, C.H. Di, L. Wang, Q. Shi, R.H. Sun, D.L. Xia, M. Rui, J. Tang, D.L. Ma, Molecular cloning and characterization of chemokine-like factor 1 (CKLF1), a novel human cytokine with unique structure and potential chemotactic activity, Biochem. J. 357 (2001) 127–135.

[13] W.L. Han, P.G. Ding, M.X. Xu, L. Wang, M. Rui, S. Shi, Y.N. Liu, Y. Zheng, Y.Y. Chen, T. Yang, D.L. Ma, Identification of eight genes encoding chemokine-like factor superfamily members 1–8 (*CKLFSF1–8*) by in silico cloning and experimental validation, Genomics 81 (2003) 609–617.

[14] E. Wingender, X. Chen, E. Fricke, R. Geffers, R. Hehl, I. Liebich, M. Krull, V. Matys, H. Michael, R. Ohnhauser, M. Pruss, F. Schacherer, S. Thiele, S. Urbach, The TRANSFAC system on gene expression regulation, Nucleic Acids Res. 29 (2001) 281–283.

[15] R. Cavin Perier, T. Junier, P. Bucher, The eukaryotic promoter database EPD, Nucleic Acids Res. 26 (1998) 353–357.
[16] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, Proc. Natl. Acad. Sci. USA 24 (1977) 104–108.
[17] F. Graham, A. van der Eb, A new technique for the assay of infectivity of human adenovirus 5 DNA, Virology 52 (1973) 456–457.
[18] F.C. Wefald, B.H. Devlin, R.S. Williams, Functional heterogeneity of mammalian TATA-box sequences revealed by interaction with a cell-specific enhancer, Nature 344 (1990) 260–262.
[19] X.H. Sun, N.G. Copeland, N.A. Jenkins, D. Baltimore, Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins, Mol. Cell. Biol. 11 (1991) 5603–5611.
[20] S. Faisst, S. Meyer, Compilation of vertebrate-encoded transcription factors, Nucleic Acids Res. 20 (1992) 3–26.
[21] W. Lee, A. Haslinger, M. Karin, R. Tjian, Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40, Nature 325 (1987) 368–372.
[22] V. Lemarchandel, J. Ghysdael, V. Mignotte, C. Rahuel, P.H. Romeo, GATA and Ets cis-acting sequences mediate megakaryocyte-specific expression, Mol. Cell. Biol. 13 (1993) 668–676.
[23] S.J. Courtois, D.A. Lafontaine, F.P. Lemaigre, S.M. Durviaux, G.G. Rousseau, Nuclear factor-I and activator protein-2 bind in a mutually exclusive way to overlapping promoter sequences and trans-activate the human growth hormone gene, Nucleic Acids Res. 18 (1990) 57–64.
[24] A. Wagner, Genes regulated cooperatively by one or more transcription factors and their identification in whole eukaryotic genomes, Bioinformatics 15 (1999) 776–784.
[25] J.W. Fickett, A.G. Hatzigeorgiou, Eukaryotic promoter recognition, Genome Res. 7 (1997) 861–878.
[26] F. Lopez-Casillas, K.H. Kim, Heterogeneity at the 5′ end of rat acetyl-coenzyme A carboxylase mRNA. Lipogenic conditions enhance synthesis of a unique mRNA in liver, J. Biol. Chem. 264 (1989) 7176–7184.
[27] L. Monaco, E. Vicini, M. Conti, Structure of two rat genes coding for closely related rolipram-sensitive cAMP phosphodiesterases. Multiple mRNA variants originate from alternative splicing and multiple start sites, J. Biol. Chem. 269 (1994) 347–357.
[28] M.K. Tee, G.O. Babalola, P. Aza-Blanc, M. Speek, S.E. Gitelmanm, W.L. Miller, A promoter within intron 35 of the human C4A gene initiates abundant adrenal-specific transcription of a 1 kb RNA: location of a cryptic CYP21 promoter element? Hum. Mol. Genet. 4 (1995) 2109–2116.