Translational Repression of a Splice Variant of Cynomolgus Macaque CXCL1L by Its C-terminal Sequence

Hisayuki Nomiyama,[1] Naoki Osada,[2] Ichiro Takahashi,[3] Keiji Terao,[3] Kazuya Yamagata,[4] and Osamu Yoshie[5]

[1] Department of Molecular Enzymology, Kumamoto University Graduate School of Medical Sciences, Honjo, Kumamoto 860-8556, Japan.
[2] Division of Bioengineering and Bioinformatics, Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido 060-0814, Japan
[3] Tsukuba Primate Research Center, National Institute of Biomedical Innovation, Health and Nutrition (NIBIOHN), Tsukuba, Ibaraki 305-0843, Japan
[4] Department of Medical Biochemistry, Faculty of Life Sciences, Kumamoto University, Honjo, Kumamoto 860-8556, Japan
[5] The Health and Kampo Institute, 1-11-10 Murasakiyama, Sendai, Miyagi 981-3205, Japan

Running Title: TRANSLATIONAL REPRESSION OF CXCL1Lβ

Corresponding author: Hisayuki Nomiyama
Department of Molecular Enzymology, Kumamoto University Graduate School of Medical Sciences, Honjo 1-1-1, Kumamoto 860-8556, Japan.
e-mail: Nomiyama@gpo.kumamoto-u.ac.jp

Keywords: chemokine, CXCL1L, translational regulation, gene duplication, evolution, cynomolgus macaque, Old World monkeys

Sequence data from this article have been submitted to the GenBank/EBI/DDBJ databases with the accession No. AB264543 (CXCL1Lβ)

ABSTRACT

We previously isolated a cDNA clone from cynomolgus macaque encoding a novel CXC chemokine that we termed CXCL1L from its close similarity to CXCL1. However, the cDNA consisted of three exons instead of four exons typically seen in other CXC chemokines. Here we isolated a cDNA encoding the full length variant of CXCL1L that we term CXCL1Lβ. CXCL1Lβ is 50 amino acids longer than the original CXCL1L, which we now term CXCL1Lα. The CXCL1Lβ mRNA is much more abundantly expressed in the cynomolgus macaque tissues than CXCL1Lα mRNA. However, CXCL1Lβ protein was poorly produced by transfected cells compared with that of CXCL1Lα. When the coding region of the 4th exon was fused to the C-terminus of CXCL1 or even to a non-secretory protein firefly luciferase, the fused proteins were also barely produced although the mRNAs were abundantly expressed. The polysome profiling analysis suggested that the inhibition was mainly at the translational level. Furthermore, we demonstrated the C-terminal five amino acids of CXCL1Lβ were critical for the translational repression. The present study thus reveals a unique translational regulation controlling the production of a splicing variant of CXCL1L. Since the CXCL1L gene is functional only in the Old World monkeys, we also discuss possible reasons for the conservation of the active CXCL1L gene in these monkeys during the primate evolution.

Introduction

Chemokines are a large family of chemotactic cytokines that plays important roles in inflammation and immunity (Gerber and others 2009; Mantovani and others 2010; Rollins 1997; Yoshie and others 2001; Zlotnik and Yoshie 2000; 2012). Based on the arrangement of the first two residues of conserved four cysteines, chemokines are classified into four groups, CXC, CC, C, and CX3C. Functionally, they can be further divided into two groups, inflammatory and homeostatic chemokines. In particular, the inflammatory CC and CXC chemokine genes have been generated by multiple tandem gene duplication events, forming the major CC and CXC gene clusters (Nomiyama and others 2010).

The gene duplication is a driving force for the generation of evolutionary novelty

(Eirin-Lopez and others 2012; Nei and Rooney 2005). After gene duplication, while one of the duplicated genes maintains the original function, the other could be released from the functional constraint to gain a new function or to be degenerated into a non-expressing pseudogene through nucleotide substitution or deletion. The chemokine superfamily thus provides an interesting model of gene duplication in evolution.

CXCL1 (also called Gro-α) is an inflammatory chemokine that mainly attracts neutrophils and is critically involved in inflammation, angiogenesis, and tumorigenesis (Amiri and Richmond 2003; Dhawan and Richmond 2002). Previously, we isolated a novel cynomolgus macaque CXC chemokine that we termed CXCL1L from its highest similarity and close genomic relationship to CXCL1 (Nomiyama and others 2007). While rodents have only one CXCL1 gene, there are CXCL1 and its duplicated CXCL1L gene in the genomes of cynomolgus macaque (*Macaca fascicularis*) and rhesus macaque (*Macaca mulatta*) (Nomiyama and others 2007), both of which are widely used in the medical research and belong to the same genus (*Macaca*) of Old World monkeys. These monkeys diverged from a common ancestor about 1.5 million years ago (Blancher and others 2008; Osada and others 2010). Our recent database search has further revealed that all seven Old World monkeys including *Macaca* monkeys have the active CXCL1L gene, while the gene has been inactivated in the genomes of higher primates including human (Nomiyama and others 2010; 2013). This may suggest a unique role of CXCL1L in the Old World monkeys.

Our original cynomolgus macaque cDNA clone of CXCL1L corresponded to an EST in the public database (Nomiyama and others 2007). CXCL1L shares 80% amino acid sequence similarity to cynomolgus macaque CXCL1, which we also cloned in the same study (Nomiyama and others 2007). Most CXC chemokine genes are known to have four exons. The comparison of the CXCL1L cDNA sequence and the corresponding gene, however, suggested that the cDNA was derived from an alternative splice product lacking the 4th exon sequence.

In this study, we therefore isolated the full-length CXCL1L cDNA clone of cynomolgus macaque consisting of the four exon sequences. We now term the products of the shorter cDNA and the full-length cDNA as CXCL1Lα and CXCL1Lβ, respectively. To elucidate the function of CXCL1Lα and CXCL1Lβ, we tried to produce their proteins. In contrast to the CXCL1Lα protein, however, we failed to obtain efficient production of the CXCL1Lβ protein by silk worms or transfected cells.

Our subsequent study revealed that Ex4 (we use this abbreviation in this report to indicate the coding region of the exon 4) represses the translation in spite of the abundant production of mRNA in transfected cells. Furthermore, the C-terminal five amino acids of Ex4 are critical for this suppression. Thus, we have revealed a unique translational regulation for the production of a splicing variant of CXCL1L. We also discuss the conservation of an active CXCL1L gene only in Old World monkeys in functional and evolutionary points of view.

Materials and Methods

Cloning and expression of CXCL1Lβ cDNA

Cynomolgus macaque CXCL1Lβ cDNA was cloned by RT-PCR using spleen cDNA library prepared as described (Nomiyama and others 2007). The primers used were based on the rhesus macaque CXCL1L gene and were 5'-AGTTCCCCTGCTCCTCTCAC and 5'-GAAAGAACCTTAATAGACAAGCTTTGA.

The expression plasmids for cynomolgus macaque CXCL1, CXCL1Lα and CXCL1Lβ were constructed using each cDNA as template. The primers used were: 5'-TTTGGTACCATGGCCCGCGCCGCGCTC and 5'-TTTTCTAGAGTTGGATTTGTCACAGTTCAG for CXCL1; 5'-TTTGGTACCATGGCCTGCGCCGCACCC and 5'-TTTTCTAGACCTGTTCAGCATCTTTTGCAT for CXCL1Lα; 5'-TTTGGTACCATGGCCTGCGCCGCACCC and 5'-TTTTCTAGATAAGAAGCAGTTAAACACATT for CXCL1Lβ. The synthesized cDNAs were digested with *Kpn*I and *Xba*I and cloned into pEF4/Myc-HisA vector containing polyhistidine tag (Thermo Fisher Scientific, MA). For construction of CXCL1-Ex4, the Ex4 sequence was amplified with primers 5'-TTTTCTAGAGGCAACGCTAACGGACCTAG and 5'-TTTTCTAGATAAGAAGCAGTTAAACACATT using CXCL1Lβ cDNA as template, and the *Xba*I digested fragment was inserted into the *Xba*I site, located downstream of the CXCL1 coding sequence, of the above CXCL1 expression plasmid. For construction of CXCL1-STOP-Ex4, the Ex4 fragment amplified with primers 5'-GGCAACGCTAACGGACCTAG and 5'-TAAGAAGCAGTTAAACACATT was

cloned into the *Pme*I site, located downstream the polyhistidine tag, of the CXCL1 expression vector. For construction of CXCL1-STOP-ARE, the AU-rich element in the 3'-noncoding region of the human granulocyte-macrophage colony-stimulating factor (GM-CSF) gene (Shaw and Kamen 1986) was amplified with primers 5'-CAGGCATGGCAGAAGAATGGGAATA and 5'-ATTACGGTAAAACATCTTGAATAAA using the human genome DNA as the template, and the obtained 135 bp fragment was inserted into the *Pme*I site of the CXCL1 expression vector.

The C-terminal serial deletion mutants were constructed as the construction of CXCL1-Ex4. The downstream primer of the primer pair was 5'-AAATCTAGACACATTAAACACAGTCCA (ΔC5), 5'-TTTTCTAGAAGTGTGCCCAGGAGCTAATTG (ΔC11), 5'-TTTTCTAGATTGTTTTTCTTCTTTCACTTT (ΔC17), 5'-AAATCTAGACACTTTGATTCCTCTAAGAATAGA (ΔC22), 5'-AAATCTAGACTTTAGGAATTGTAGCGG (ΔC31). Other C-terminal deletion mutants of short Ex-4 sequences were generated by inserting the annealed oligonucleotides, 5'-CTAGAGGCAACGCTAACGGACCTAGAGAGAAGGAGAAAT and 5'-CTAGATTTCTCCTTCTCTCTAGGTCCGTTAGCGTTGCCT (ΔC39), 5'-CTAGAGGCAACGCTAACGGAT and 5'-CTAGATCCGTTAGCGTTGCCT (ΔC45), into the *Xba*I site of the CXCL1 expression vector.

Luciferase-Ex4 (Luc-Ex4) and Luc-STOP-Ex4 plasmids were also constructed. Luciferase coding region was amplified with primers 5'-TTTAAGCTTATGGAAGACGCCAAAAACATAAA and 5'-AAAGAATTCCACGGCGATCTTTCCGCCCTTCT (Luc-Ex4) or 5'-AAAGAATTCTTACACGGCGATCTTTCCGCCCT (Luc-STOP-Ex4) using pGL3-Control vector (Promega, Madison, WI) as a template. Then, the Luc fragments were digested with *Eco*RI and *Hind*III and inserted into pcDNA3.1 (+) vector (Clontech, Palo Alto, CA).

RT-PCR

Relative expression levels of CXCL1Lα and CXCL1β mRNAs in cynomolgus macaque tissues were determined by RT-PCR. The primer pairs include one common

primer annealing to the first exon sequence and a downstream primer specific for either CXCL1Lα or CXCL1Lβ (Fig. 1). The common primer is 5'-CTGCTCGTGCTCCTGGTG, and the CXCL1Lα- and CXCL1Lβ-specific primers are 5'-GCCAGTATTTCTGACCAACG and 5'-TCTAGGTCCGTTAGCGTTGC, respectively. The amplification products are 299 bp for CXCL1Lα and 270 bp for CXCL1Lβ in length.

DNA transfection and Western and Northern blot analyses

Transfection of human embryonic kidney HEK293 cells and Chinese hamster ovary CHO cells with the constructed plasmids was performed using a transfection reagent LipoTrust EX Gene (Hokkaido System Science, Sapporo, Japan). The EGFP expression vector pEGFP (enhanced green fluorescent protein)-N1 (Clontech) was used as an internal control for Western blot analysis. Thirty-two hours after transfection, cell lysates were prepared with RIPA lysis buffer (Nakalai, Tokyo, Japan). Anti-EGFP, anti-His-tag, anti-luciferase polyclonal antibodies were obtained from Medical & Biological laboratories (Nagoya, Japan), and peroxidase-goat anti-rabbit IgG from ZYMED (San Francisco, CA). Total RNAs were prepared from the transfected cells by using Trisol reagent (Thermo Fisher Scientific). For Northern blot analyses, digoxigenin-labeled probes for CXCL1, luciferase, and β-actin were prepared by DIG RNA Labeling kit (Roche Diagnostics, Basel, Switzerland). The CXCL1 RNA probe was prepared from a PCR product of 210 bp in length that was amplified with 5'-GCTGCTCCTGCTCCTGGT and 5'-GATGCGGGGTTGAGACAAG primers. The CXCL1 probe that shares 92% identity with CXCL1L was used to detect both CXCL1 and CXCL1L mRNAs. The washing step after the hybridization was performed at low stringency (0.5X SSC, 60 ºC). ECL Plus Western blotting detection system (GE Healthcare, Pittsburgh, PA) or CDP-Star (Roche, Basel, Swiss) was used for the detection of Western and Northern signals, respectively. An image scanner LAS-3000 (Fujifilm, Tokyo, Japan) was used to determine relative signal intensity and an average was obtained using three independent experiments.

Polysome profile analysis

HEK293 cells transfected with plasmids in a 10 cm dish were lysed on the dish

with cold lysis buffer (5 mM Tris-HCl, pH 7.4, 1.5 mM KCl, 2.5 mM $MgCl_2$, 1 mM dithiothreitol, 100μg/ml cycloheximide, 1% Triton X-100, 1% deoxycholate). Polysome profile analysis was performed as described (He and Green 2013). Briefly, the lysate was homogenized with tissue grinder 10 times followed by centrifugation at 10,000 g for 10 minutes. The supernatant was then layered onto linear sucrose gradient (10-50% in 80 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-HCl, pH 7.4, and 1 mM dithiothreitol), and centrifuged at 27,000 rpm for 6 hrs. After centrifugation, fractions were collected while the polysome profiles were recorded with UV spectrometer. RNAs were extracted from the fractions and used for Northern blot analyses as described above.

Results

Isolation of cynomolgus macaque CXCL1Lβ cDNA

Previously we isolated a cDNA clone from cynomolgus macaque spleen encoding a novel CXC chemokine that we termed CXCL1L (Nomiyama and others 2007), which was apparently generated from CXCL1 by gene duplication at the early stage of primate evolution (Nomiyama and others 2013). Since the cynomolgus macaque genome sequence has not been determined at the beginning of the present study, we compared the cDNA sequence with that of the CXCL1L gene in the genome of rhesus macaque, the species closely related to the cynomolgus macaque (Nomiyama and others 2007). While the genes of most CXC chemokines including the original CXCL1 have four exons with the 4th exon coding the C-terminal region of a few amino acids, the cynomolgus macaque CXCL1L cDNA was found to be composed of only the first three exons. This suggested that the cDNA encoded a splice valiant of CXCL1L. We therefore tried to identify the CXCL1L exon 4 in the rhesus macaque genome sequence. We failed to identify the exon 4 by the similarity search using the exon 4 sequence of the rhesus macaque CXCL1 gene as a probe. However, by using a gene prediction software GENSCAN (Burge and Karlin 1997), we identified an exon sequence of 154 nucleotide long encoding a long peptide of 50 amino acids 812 bp downstream from the third exon of rhesus macaque CXCL1L. Furthermore, we have also identified a similar potential 4th exon sequence of CXCL1L in the cynomolgus macaque genome sequence recently reported (Higashino and others 2012; Yan and others 2011), which is located 532 bp downstream from the third exon in the cynomolgus macaque genome (Fig. 1A).

Based on the nucleotide sequence of the potential exon 4, we designed PCR primers and isolated the full-length cDNA clone from the spleen cDNA library prepared in our previous study (Nomiyama and others 2007). The exon 4 encodes a long peptide of 50 amino acids now termed Ex4, which is highly conserved between cynomolgus and rhesus macaques with only one amino acid substitution (Fig. 2A). We therefore renamed the spliced variant encoded by the first three exons as CXCL1Lα and the one from the four exons as CXCL1Lβ. The predicted CXCL1Lα and CXCL1Lβ proteins are 104 and 154 amino acids long, respectively (Fig. 2A).

Our previous RT-PCR data showed that the CXCL1L mRNA was expressed at a high level in the spleen and to a lesser level in the brain (Nomiyama and others 2007). In order to determine the expression of CXCL1Lα and CXCL1Lβ separately, we performed semi-quantitative RT-PCR using these two tissue cDNAs. The primer pairs contain one common primer that anneals to the first exon sequence and the downstream primers that are specific for either CXCL1Lα or CXCL1Lβ (Fig. 1A). The annealing site of the CXCL1Lα-specific primer is located downstream of the coding region of CXCL1Lα exon 3. This site is present in the third intron of the CXCL1Lβ. The results showed that CXCL1Lβ mRNA was dominantly expressed in both tissues (Fig. 1B).

Protein expression in cultured cells

In order to compare functions of CXCL1, CXCL1Lα and CXCL1Lβ, we tried to produce their proteins by silkworms transfected with the respective expression vectors. Unexpectedly, we could not detect CXCL1Lβ protein in the body fluid of the transfected silkworms (data not shown). We therefore constructed mammalian expression plasmids of CXCL1, CXCL1Lα, and CXCL1Lβ with His-tags at their C-terminal ends and compared their expression in transfected CHO cells. An EGFP expression plasmid was also transfected as an internal control. Thirty-two hours after transfection, Western blot analyses were performed by applying 1 μl from 600 μl of culture supernatant to each lane. While the protein bands of CXCL1 and CXCL1Lα were strongly visible, the CXCL1Lβ protein band was hardly detected (Fig. 3A). When the same experiments were repeated using HEK293 cells, CXCL1Lβ protein was again hardly detected in contrast to those of CXCL1 and CXCL1Lα (data not shown but see Fig. 3B). However, when 1 μl from 30 μl of the total cell lysates of transfected HEK293 cells was applied on the gel, a small amount of CXCL1Lβ protein was detected (Fig.

3B). These results suggested that the CXCL1Lβ protein was present at low levels in the endoplasmic reticulum but hardly detected in the culture supernatant.

In order to examine an effect of the Ex4 sequence on the protein production, a construct was generated to express CXCL1 with Ex4 fused to its C-terminal end (CXCL1-Ex4) (Fig. 3E). The results from the CXCL1-Ex4 construct were quite similar to those of CXCL1Lβ: namely, a poor production of CXCL1-Ex4 protein compared to that of CXCL1 (Fig. 3B). However, when we placed the Ex4 sequence at 6 nucleotides downstream of the stop codon of CXCL1 (CXCL1-STOP-Ex4) (Fig. 3E), the Ex4 sequence did not affect the protein abundance (Fig. 3B). These results suggested that the Ex4 sequence needed to be present within the coding region to suppress protein production.

We further examined whether Ex4 could inhibit the production of a non-secretory protein by using firefly luciferase. Again, the presence of the Ex4 sequence fused to the C-terminal end of luciferase (Luc-Ex4) (Fig. 3E) but not after the stope codon (Luc-STOP-Ex4) (Fig. 3E) strongly suppressed the protein production in cell lysates (Fig. 3C). This strongly suggested that the inhibitory activity of the Ex4 sequence was not at the secretion step but at the translation step.

Expression of mRNA in cultured cells

We also determined mRNA expression levels by Northern blot analysis. As shown in Fig. 3B, the level of CXCL1Lβ mRNA was also low compared with that of CXCL1Lα in transfected HEK293 cells. Furthermore, the presence of the Ex4 sequence fused to the C-terminal end or after the stop codon of CXCL1 also greatly reduced the mRNA levels. This might be due to the presence of AU-rich sequences that cause instability of mRNA (Hao and Baltimore 2009; Shaw and Kamen 1986). In fact, the CXCL1Lβ mRNA has at least one canonical AUUUA sequence in its 3' untranslated region, and the coding region of Ex4 is also AU-rich (60%) compared to those of the exons 1-3 (38%). Therefore, for comparison, we also constructed an expression vector of CXCL1 that contained an AU-rich element (ARE) of the human GM-CSF gene (Shaw and Kamen 1986) in the 3' non-coding region (CXCL1-STOP-ARE) (Fig. 3E). Transfection of the vector showed that its mRNA expression level was as low as those of the Ex4-containing CXCL1-Ex4 and CXCL1Lβ (Fig. 3D). However, the protein expression level of the CXCL1-STOP-ARE was comparable to that of

CXCL1-STOP-Ex4. It was also noted that the mRNA expression levels of Luc-Ex4 and Luc-STOP-Ex4 were not reduced compared to that of Luc (Fig. 3C). Similar phenomenon was observed with the constructs using fluorescent protein DsRed gene (data not shown). Collectively, these results suggest that, while the presence of the Ex4 sequence in the C-terminal end strongly reduces the protein production, its effect on the mRNA level is context-dependent.

Protein expression normalized by mRNA expression

Since the Ex4 reduced both the protein and mRNA levels, the effect of Ex4 on the protein expression levels of various constructs were compared after normalization with the levels of mRNA expression, assuming that mRNAs of the constructs were equally expressed and not degraded. Accordingly, averages of protein and mRNA signal intensities were determined using three independent experiments, two of which are shown in Figs. 3B and 3D. By setting the normalized protein level of CXCL1 as 100%, the normalized protein levels in the culture supernatants (mean of three experiments ± standard error) were 25.8±4.5% for CXCL1Lα, 1.3±0.2% for CXCL1Lβ, 8.8 ±0.7% for CXCL1-Ex4, and 32.3±2.4% for CXCL1-STOP-Ex4. The data clearly indicated that the protein expression of CXCL1Lβ and CXCL1-Ex4 was constantly low compared with CXCL1α, and CXCL1-STOP-Ex4, although CXCL1 exhibited an even higher protein expression level than CXCL1Lα and CXCL1-STOP-Ex4.

Luc-Ex4 mRNA binds less ribosomes than Luc mRNA

To see whether the decreased translation of proteins containing Ex4 was at the level of initiation or extension, the polysome profile analysis was performed. HEK293 cells transfected with Luc-Ex4 or Luc expression plasmids were homogenized and the cell lysates were applied to 15-50% sucrose gradient. This procedure separates mRNAs depending on the number of ribosomes binding to the mRNAs, which reflects the rate of translational initiation and elongation (He and Green 2013). After centrifugation, aliquots were collected. As shown in Fig. 4 (upper panel), the polysome profiles that represented the sum of those of various transcripts were quite similar between the cells transfected with Luc (A) or Luc-Ex4 (B). To determine the actual polysome numbers associated with Luc and Luc-Ex4 mRNAs, we next performed Northern blotting using

mRNAs extracted from the aliquots. As shown in Fig. 4 (lower panel), while the Luc mRNA was loaded with 1.5 to 6 ribosomes with the peak value of 3 ribosomes (A), the Luc-Ex4 mRNA was loaded with 1.5 to 2.5 ribosomes with the peak value of 2 ribosomes (B). The results indicated that the number of ribosomes bound to the Luc-Ex4 mRNA was much less than that of the Luc mRNA. The results also suggested that the translation was at least initiated. Although it could not be determined from this analysis whether the ribosomes were stalled on the mRNA, the low translation level of Ex4-containing mRNA might be mostly due to the repression of translational elongation.

The C-terminal five amino-acid sequence of Ex4 is important for the translational repression

To identify the amino acid sequence within the Ex4 that was responsible for the observed translational repression, successive C-terminal deletion mutants were made from the CXCL1-Ex4 construct (Fig. 5). We used CXCL1-Ex4 instead of the natural CXCL1Lβ because it was easier to replace the Ex4 sequence with truncated Ex4 sequences using the CXCL1-Ex4 construct. Furthermore, the expression of CXCL1 was much higher at both protein and mRNA levels than those of CXCL1Lα or CXCL1Lβ. The constructs were transfected into HEK293 cells, and Western blotting analyses were performed using cell culture supernatants and cell lysates. As shown in Fig. 5, the deletion of only five amino acids from the C-terminal end resulted in a tremendous increase in the CXCL1-Ex4 proteins in the cell supernatant, although it showed less effect on the protein abundance in the cell lysates. These results indicated that the C-terminal five amino acids (Phe-Asn-Cys-Phe-Leu) have the major impact on the level of protein secretion. Of note, the deletion of the C-terminal five amino acids also greatly affected the electrophoretic mobility of the protein band in the gel with an upward shift. This suggested altered tertiary structures of the C-terminally truncated proteins.

The mRNA expression levels were also gradually increased by the successive deletion of the Ex4 sequence. However, almost the half (~22 amino acids) of the Ex4 sequence is needed to be removed to make the expression levels comparable to that of CXCL1, suggesting that the region responsible for the reduced mRNA expression is much broader than that of the repression of protein translation.

Discussion

CXCL1L is a gene copy duplicated from CXCL1 in the primate lineage after the divergence of primates and rodents (Nomiyama and others 2013). Previously, we isolated a cDNA clone of CXCL1L from cynomolgus macaque (Nomiyama and others 2007). By comparing to the available rhesus macaque genome, however, we found that the cDNA was consisted of three exons instead of the standard four-exon structure of CXC chemokines. In the present study, therefore, we have isolated the full-length CXCL1L cDNA consisting of four exons. In contrast to a few amino acids encoded by the fourth exons of other CXC chemokines, however, the fourth exon of CXCL1L encodes a peptide of 50 amino acids long termed Ex4. We now call the short version of CXCL1L as CXCL1Lα and the long one as CXCL1Lβ. However, our subsequent study has revealed that the presence of Ex4 appears to have a negative impact on the stabilization and/or translation of CXCL1Lβ mRNA. Indeed, when the Ex4 sequence was fused to the genes of cynomolgus macaque CXCL1 or firefly luciferase, the translation of these genes was also greatly repressed (Fig. 3). However, the Ex4 sequence had no such repressive activity when present after the stop codon. Thus, the Ex4 sequence inhibits translation of a protein to which Ex4 is directly fused, suggesting that the presence of Ex4 at the C-terminus of a protein is required for the translational repression. Indeed, the presence of Ex in the C-terminus of firefly luciferase greatly affected the ribosome binding to its mRNA (Fig. 4). Furthermore, the C-terminal five amino-acid sequence of Ex4 is critical in suppression of protein production (Fig. 5). These results suggest that although the ribosome binds to mRNA and initiates translation of mRNA, the presence of Ex4 sequence at the C-terminus induces ribosome stalling, resulting in poor protein translation (Fig. 6).

Arresting of polysomes is a widely used mechanism in controlling the translational rate (Richter and Coller 2015), and the examples of specific amino acid sequences in the nascent peptides to arrest translational elongation of their own mRNAs are known in species from eubacteria to mammals (Chiba and Ito 2012; Cruz-Vera and others 2005; Fang and others 2000; Gong and Yanofsky 2001; 2002; Ito and others 2010; Nakatogawa and Ito 2002; Onouchi and others 2005; Tenson and Ehrenberg 2002; Vazquez-Laslop and others 2008; Wilson and Beckmann 2011). These sequences

generally contain basic amino acid residues and interact with the negatively-charged ribosome tunnel that exports proteins from the ribosome, and thus may cause ribosome stalling (Lu and Deutsch 2008). Some of such amino acid sequences have been found in proteins involved in physiological control systems. In the cases of *E. coli* secretion monitor SecM and *B. subtilis* membrane insertion/folding monitor MifM, the arrest can be released when the physiological situation is restored (Chiba and others 2009; Ito and others 2010). In the Ex4 sequence of the CXCL1Lβ, there are two Arg and six Lys residues, although they are not located in the crucial C-terminal five amino acid residues (Fig. 2A). It remains to be seen whether the ribosome pausing site on the Ex4 sequence corresponds to the C-terminal five amino-acid sequence. Ribosome profiling (Ingolia and others 2009; Ingolia and others 2011; Oh and others 2011) by next-generation sequencers may be helpful to detect the ribosome-protected mRNA fragments, thus allowing the identification of ribosome pausing sites.

Arrested nascent peptides are usually degraded by proteasome, although the mRNAs are not subjected to degradation (Dimitrova and others 2009; Inada 2013; Ito-Harashima and others 2007). Our preliminary data show, however, that proteasome inhibitor MG132 does not substantially affect the protein level of CXCL1Lβ, suggesting that the proteasome degradation is not the major mechanism for the reduction of protein production (data not shown).

The expression levels of Luc-Ex4 and Luc-STOP-Ex4 mRNAs were not appreciably decreased compared with that of Luc mRNA (Fig. 3C). Thus, the presence of the Ex4 sequence does not appear to affect the stability of these mRNAs. In contrast, CXCL1Lβ or CXCL1-Ex4 mRNAs were less abundant than those mRNAs without the Ex4 sequence and thus may be degraded (Figs. 3B, 3D, and 5). Given that the turnover of the Luc mRNA was not appreciably enhanced by the presence of Ex4 sequence (Fig. 3C), the effect of the Ex4 sequence on the mRNA stability may be context-dependent.

Translational elongation can be arrested by RNA secondary structures (Dimitrova and others 2009; Inada 2013). The Ex4 RNA sequence could form several stem-loop structures (unpublished results). Determination of ribosome pausing sites may elucidate whether the RNA secondary structure of Ex4 is also involved in the ribosome stalling.

So far, the active CXCL1L gene has been found only in seven Old World monkeys (family Cercopithecidae): cynomolgus macaque, rhesus macaque, Southern pig-tailed macaque (*Macaca nemestrina*), sooty mangabey (*Cercocebus atys*), drill

(*Mandrillus leucophaeus*), olive baboon (*Papio anubis*), and African green monkey (*Chlorocebus sabaeus*) (Figs. 2A and B). On the other hand, the gene has been inactivated by deletion of the third and fourth exons in other hominids such as human, chimpanzee, and gorilla. In Orangutan, the gene has been inactivated by nucleotide substitutions. Common marmoset (*Callithrix jacchus*) of New World monkeys, which diverged from Old World monkeys roughly 43.5 million years ago (Perelman and others 2011), also has a pseudogene of CXCL1L inactivated by nucleotide substitutions (our unpublished results). Thus, it appears that there might be a lineage-specific selective force for the preservation of active CXCL1L gene in Old World monkeys. Conversely, there might be a strong negative selection for the presence of active CXCL1L gene in the higher primates except for the Old World monkeys. Our genome analyses have also shown that Philippine tarsier (*Tarsius syrichta*), which is a separate relict lineage with an independent origin, and bushbaby (*Otolemur garnettii*) of Strepsirrhines seem to have no CXCL1L gene, though both of them have the CXCL1 gene.

Obviously, the CXCL1L gene was generated from CXCL1 by gene duplication during the evolution of primates. It seems likely that the gene had been active in the primates during the emergence of Old World monkeys but was inactivated in the lineages of New World monkeys and hominids. Notably, the gene also gained a new exon 4 in Old World monkeys. This might have conserved the CXCL1L gene in the Old World monkeys with a possible new function. The newly gained exon 4 encodes a long peptide of 50 amino acids, resulting in a quite new CXC chemokine with an exceptionally long C-terminal end compared to CXCL1 and other CXC chemokines. The newly acquired Ex4 sequence is relatively rich in polar amino acids (50% in Ex4 vs. 41% in exons 1-3) and also in charged residues (50% positively charged residues in Ex4 vs. 12% in the exons 1-3; 10% negatively charged residues in Ex4 vs. 2% in the exons 1-3). The pI of Ex4 is 9.36, while that of exons 1-3 is 10.11. The net pI of CXCL1Lβ is 9.90.

The function of Ex4 is at present unclear, since the Ex4 peptide has no sequence similarity to other proteins. Of note, however, the newly gained Ex4 sequence is highly conserved in Old World monkeys (Figs. 2A and 2B). This might suggest that CXCL1L with Ex4 has gained a new function important for the Old World monkeys and thus has been conserved in these monkey genomes even after they were diverged from hominids

20-38 million years ago (Perelman and others 2011).

Given the extent of evolutionary conservation of the active CXCL1L gene in the Old World monkeys (Fig. 2A), the CXCL1L proteins may have an essential role in these monkeys. For example, CXCL1Lα, which is the major protein product of the CXCL1L gene and lacks a few amino acids encoded by an original exon 4, might have acquired a new function in the Old World Monkeys by deleting the C-terminal amino acids. Alternatively, CXCL1Lβ, which has extra 50 amino acids encoded by the newly gained exon 4 and is the major product at the mRNA level, may have play a specific role in these monkeys in certain circumstances through coevolution of a mechanism that releases the ribosomal pausing of CXCL1Lβ mRNA in such conditions to promote its protein production. For example, since CXCL1 is one of the inflammatory chemokines, CXCL1Lβ may contribute to coping with particular infectious agents specific to these monkeys. The long C-terminal of CXCL1Lβ may promote the interaction with extracellular substrates such as heparin with affinities different from CXCL1, and thus may have acquired a new role in cell migration in those monkeys (Sierra and others 2004). Yet, another possibility, although may sound quite unusual, is that the CXCL1L gene in the Old World monkeys is also in the process of inactivation like its counterparts in New World monkeys and higher primates and in fact already behaves like a pseudogene with its self-inflicted suppression of mRNA translation.

In conclusion, it would be of quite interest to see whether CXCL1Lβ protein can be produced by cells or tissues of Old World monkeys in particular conditions. CXCL1Lα and/or CXCL1Lβ might also have some activities different from original CXCL1. The CXCL1L gene with its unique translational control not known in any other genes may also provide an interesting case for the study of evolution and function of duplicated genes.

Acknowledgements

We thank B. Okuno, T. Sata, and Y. Endo of Kumamoto Univ. and J. Kusuda of the NIBIOHN for their help.

Author Disclosure Statement

No competing financial interests exist.

References

Amiri KI, Richmond A. 2003. Fine tuning the transcriptional regulation of the CXCL1 chemokine. Progress in Nucleic Acid Research and Molecular Biology 74:1-36.

Blancher A, Bonhomme M, Crouau-Roy B, Terao K, Kitano T, Saitou N. 2008. Mitochondrial DNA sequence phylogeny of 4 populations of the widely distributed cynomolgus macaque (Macaca fascicularis fascicularis). Jou nal of Heredity 99(3):254-64.

Burge C, Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. Journal of Molecular Biology 268(1):78-94.

Chiba S, Ito K. 2012. Multisite ribosomal stalling: a unique mode of regulatory nascent chain action revealed for MifM. Molecular Cell 47(6):863-72.

Chiba S, Lamsa A, Pogliano K. 2009. A ribosome-nascent chain sensor of membrane protein biogenesis in Bacillus subtilis. EMBO Journal 28(22):3461-75.

Cruz-Vera LR, Rajagopal S, Squires C, Yanofsky C. 2005. Features of ribosome-peptidyl-tRNA interactions essential for tryptophan induction of tna operon expression. Molecular Cell 19(3):333-43.

Dhawan P, Richmond A. 2002. Role of CXCL1 in tumorigenesis of melanoma. J Leukoc Biol 72(1):9-18.

Dimitrova LN, Kuroha K, Tatematsu T, Inada T. 2009. Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. Journal of Biological Chemistry 284(16):10343-52.

Eirin-Lopez JM, Rebordinos L, Rooney AP, Rozas J. 2012. The birth-and-death evolution of multigene families revisited. Genome Dyn 7:170-96.

Fang P, Wang Z, Sachs MS. 2000. Evolutionarily conserved features of the arginine attenuator peptide provide the necessary requirements for its function in translational regulation. Journal of Biological Chemistry 275(35):26710-9.

Gerber PA, Hippe A, Buhren BA, Muller A, Homey B. 2009. Chemokines in tumor-associated angiogenesis. Biol Chem 390(12):1213-23.

Gong F, Yanofsky C. 2001. Reproducing tna operon regulation in vitro in an S-30 system. Tryptophan induction inhibits cleavage of TnaC peptidyl-tRNA. Journal

of Biological Chemistry 276(3):1974-83.
Gong F, Yanofsky C. 2002. Instruction of translating ribosome by nascent peptide. Science 297(5588):1864-7.
Hao S, Baltimore D. 2009. The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nature Immunology 10(3):281-8.
He SL, Green R. 2013. Polysome analysis of mammalian cells. Methods in Enzymology 530:183-92.
Higashino A, Sakate R, Kameoka Y, Takahashi I, Hirata M, Tanuma R, Masui T, Yasutomi Y, Osada N. 2012. Whole-genome sequencing and analysis of the Malaysian cynomolgus macaque (Macaca fascicularis) genome. Genome Biol 13(7):R58.
Inada T. 2013. Quality control systems for aberrant mRNAs induced by aberrant translation elongation and termination. Biochimica et Biophysica Acta 1829(6-7):634-42.
Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. 2009. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science 324(5924):218-23.
Ingolia NT, Lareau LF, Weissman JS. 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 147(4):789-802.
Ito K, Chiba S, Pogliano K. 2010. Divergent stalling sequences sense and control cellular physiology. Biochemical and Biophysical Research Communications 393(1):1-5.
Ito-Harashima S, Kuroha K, Tatematsu T, Inada T. 2007. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. Genes and Development 21(5):519-24.
Lu J, Deutsch C. 2008. Electrostatics in the ribosomal tunnel modulate chain elongation rates. Journal of Molecular Biology 384(1):73-86.
Mantovani A, Savino B, Locati M, Zammataro L, Allavena P, Bonecchi R. 2010. The chemokine system in cancer biology and therapy. Cytokine Growth Factor Rev 21(1):27-39.

Nakatogawa H, Ito K. 2002. The ribosomal exit tunnel functions as a discriminating gate. Cell 108(5):629-36.
Nei M, Rooney AP. 2005. Concerted and birth-and-death evolution of multigene families. Annu Rev Genet 39:121-52.
Nomiyama H, Osada N, Yoshie O. 2010. The evolution of mammalian chemokine genes. Cytokine Growth Factor Rev 21(4):253-62.
Nomiyama H, Osada N, Yoshie O. 2013. Systematic classification of vertebrate chemokines based on conserved synteny and evolutionary history. Genes to Cells 18(1):1-16.
Nomiyama H, Otsuka-Ono K, Miura R, Osada N, Terao K, Yoshie O, Kusuda J. 2007. Identification of a Novel CXCL1-Like Chemokine Gene in Macaques and Its Inactivation in Hominids. J Interferon Cytokine Res 27(1):32-7.
Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, Nichols RJ, Typas A, Gross CA, Kramer G, Weissman JS, Bukau B. 2011. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell 147(6):1295-308.
Onouchi H, Nagami Y, Haraguchi Y, Nakamoto M, Nishimura Y, Sakurai R, Nagao N, Kawasaki D, Kadokura Y, Naito S. 2005. Nascent peptide-mediated translation elongation arrest coupled with mRNA degradation in the CGS1 gene of Arabidopsis. Genes and Development 19(15):1799-810.
Osada N, Uno Y, Mineta K, Kameoka Y, Takahashi I, Terao K. 2010. Ancient genome-wide admixture extends beyond the current hybrid zone between Macaca fascicularis and M. mulatta. Molecular Ecology 19(14):2884-95.
Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y, Schneider MP, Silva A, O'Brien SJ, Pecon-Slattery J. 2011. A molecular phylogeny of living primates. PLoS Genetics 7(3):e1001342.
Richter JD, Coller J. 2015. Pausing on Polyribosomes: Make Way for Elongation in Translational Control. Cell 163(2):292-300.
Rollins BJ. 1997. Chemokines. Blood 90(3):909-28.
Shaw G, Kamen R. 1986. A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46(5):659-67.
Sierra MD, Yang F, Narazaki M, Salvucci O, Davis D, Yarchoan R, Zhang HH, Fales H,

Tosato G. 2004. Differential processing of stromal-derived factor-1 {alpha} and {beta} explains functional diversity. Blood 103:2452-2459.

Tenson T, Ehrenberg M. 2002. Regulatory nascent peptides in the ribosomal tunnel. Cell 108(5):591-4.

Vazquez-Laslop N, Thum C, Mankin AS. 2008. Molecular mechanism of drug-dependent ribosome stalling. Molecular Cell 30(2):190-202.

Wilson DN, Beckmann R. 2011. The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. Current Opinion in Structural Biology 21(2):274-82.

Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, van Gool AJ, Du H, Chen J, Chen R, Zhang P, Huang Z, Thompson JR, Meng Y, Bai Y, Wang J, Zhuo M, Wang T, Huang Y, Wei L, Li J, Wang Z, Hu H, Yang P, Le L, Stenson PD, Li B, Liu X, Ball EV, An N, Huang Q, Zhang Y, Fan W, Zhang X, Li Y, Wang W, Katze MG, Su B, Nielsen R, Yang H, Wang J, Wang X, Wang J. 2011. Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. Nature Biotechnology 29(11):1019-23.

Yoshie O, Imai T, Nomiyama H. 2001. Chemokines in immunity. Adv Immunol 78:57-110.

Zlotnik A, Yoshie O. 2000. Chemokines: a new classification system and their role in immunity. Immunity 12(2):121-7.

Zlotnik A, Yoshie O. 2012. The chemokine superfamily revisited. Immunity 36(5):705-16.

Address correspondence to:
Dr. Hisayuki Nomiyama
Department of Molecular Enzymology
Kumamoto University Graduate School of Medical Sciences
Honjo 1-1-1, Kumamoto 860-8556
Japan
E-mail: nomiyama@gpo.kumamoto-u.ac.jp

Figure Legends

FIG. 1. Identification and expression of splice variants of cynomolgus macaque CXCL1L. (**A**) The genomic structures of cynomolgus macaque CXCL1 and CXCL1L genes and splicing patterns of CXCL1Lα and CXCL1β mRNAs are shown with the sites of RT-PCR primers used in (B). Filled and white boxes are coding and noncoding regions, respectively. The exact 3' ends of CXCL1Lα and CXCL1β mRNAs are not known. The nucleotide sequences of the splice junctions between the third intron and its flanking exons are shown together with the translated amino acid sequences of CXCL1Lα and CXCL1Lβ. The genome sequences of cynomolgus macaque CXCL1 and CXCL1L were taken from AQIA010548 (GenBank accession number). (**B**) Expression of CXCL1Lα αnd CXCL1Lβ mRNAs in the spleen and brain of cynomolgus macaque. RT-PCR was performed using the primers shown in (A). The amplification products are 299 bp for CXCL1Lα and 270 bp for CXCL1Lβ in length. The β-actin was used as an internal control.

FIG. 2. The amino acid sequences of CXCL1Lα and CXCL1Lβ of seven Old World monkeys and the phylogenetic tree of primates. (**A**) The alignment of the CXCL1L amino acids. Cynomolgus macaque CXCL1 (GenBank, AB262778) is also shown for comparison. Signal sequences are shown as lowercase letters. The four conserved cysteine residues of chemokines are indicated by filled circles. The amino acid residues conserved among the CXCL1Lβ proteins are boxed. Ex4 indicates the amino acid sequence derived from the 4th exon. All the CXCL1Lα proteins form Old World monkeys are 104 amino acids long and end with R as indicated. The amino acid sequences were taken from GenBank; cynomolgus macaque (*M. fascicularis*, AB264543), rhesus macaque (*M. mulatta*, XM_001091688), Southern pig-tailed macaque (*M. nemestrina*, XM_011733753), sooty mangabey (*C. atys*, XM_012082996), drill (*M. leucophaeus*, XM_011968961), olive baboon (*P. anubis*, XM_003898769), and African green monkey (*C. sabaeus*, XM_007998877). (**B**) The phylogenetic tree of primates. New World monkeys diverged from a common ancestor with catarrhine species including Old World monkeys and humans roughly 43.5 million years ago (Perelman and others 2011). Divergence times (Mya, million years ago) (Perelman and others 2011) are not to scale.

FIG. 3. Western and Northern blot analyses of transfected cells. (**A**) The expression vectors depicted in (E) were transfected into CHO cells (A) or HEK293 cells (B, C, and D). An EGFP vector was co-transfected for transfection control. One μl from 600 μl of the cell culture supernatants and 1 μl from 30 μl of the cell lysates were loaded on a gel for western blot analyses. The vectors used were pEF4/Myc-HisA (A, B, and D) and pcDNA3.1 (C). STOP and AB indicate a stop codon and antibody, respectively. The probes used in the Northern blot analyses are shown to the right of each panel. The CXCL1 RNA probe hybridizes with both CXCL1 and CXCL1L mRNAs. (**A**) CHO cells were transfected with the expression constructs of CXCL1, CXCL1Lα, and CXCL1Lβ. (**B**) HEK293 cells were transfected with the expression constructs of CXCL1Lα, CXCL1Lβ, CXCL1, CXCL1-Ex4, and CXCL1-STOP-Ex4. (**C**) HEK293 cells were transfected with the expression constructs for Luc, Luc-Ex4, Luc-STOP-Ex4. (**D**) HEK293 cells were transfected with CXCL1-STOP-ARE together with the constructs used in (B). (**E**) The schematic depiction of used expression vectors. His-tags are indicated by small black boxes.

FIG. 4. Polysome profile analysis. HEK293 cells were transfected with Luc (**A**) or Luc-Ex4 (**B**), and extracts were prepared in the presence of cycloheximide. Extracts were subjected to velocity sedimentation in sucrose gradient (15-50%), and the fractions were monitored using UV absorbance at 260nm and collected. RNAs extracted from the fractions were analyzed by Northern blotting using luciferase probe.

FIG. 5. Expression of successive C-terminal deletion mutants of CXCL1-Ex4 in HEK293. The Ex4 sequence fused to the cynomolgus macaque CXCL1 was deleted successively from the C-terminal end as depicted in the right. The constructs were transfected into HEK293 cells. The cell culture supernatants, cell lysates, and total RNAs were prepared and analyzed as in Fig. 3. Protein bands shifted upward in the Δ5 and Δ11 lanes of supernatant and cell lysate samples are indicated by the arrows.

FIG. 6. Schematic illustration of translational repression of CXCL1Lβ mRNA by Ex4. More ribosomes bind cynomolgus macaque CXCL1Lα mRNA than CXCL1Lβ mRNA. Once the Ex4 sequence of CXCL1Lβ mRNA is translated, the nascent peptide may

inhibit translational initiation and/or elongation of the following ribosomes.

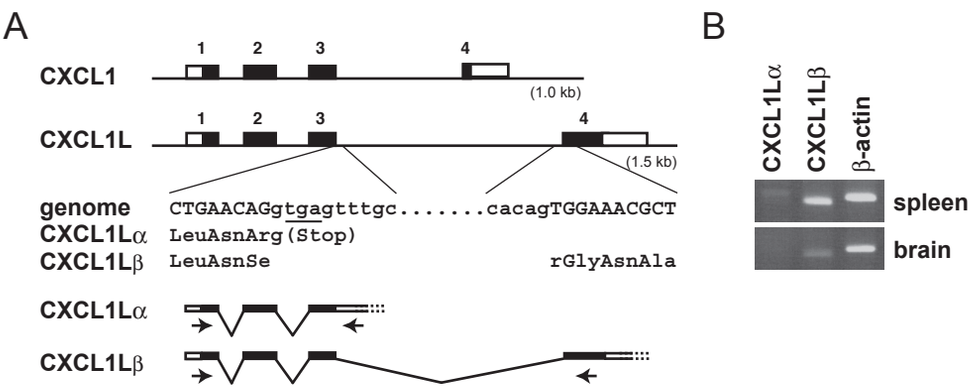
A
1
2
3
4
CXCL1
(1.0 kb)
CXCL1L
(1.5 kb)
genome
CTGAACAGgtgagtttgc.....cacagTGGAAACGCT
CXCL1Lα
LeuAsnArg(Stop)
CXCL1Lβ
LeuAsnSe
rGlyAsnAla
CXCL1Lα
CXCL1Lβ
B
CXCL1Lα
CXCL1Lβ
β-actin
spleen
brain

Fig. 1

A

```
  Old World monkeys CXCL1Lβ
cynomolgus macaque            macaaphavpgpprllrvallllvllvatcrraagAPVVTELRCQCLQTLQG 52
rhesus macaque                macsaphavpgpprllrvallllvllvatcrraagAPVVTELRCQCLQTLQG 52
Southern pig-tailed macaque   macaaphavpgpprllrvallllvllvatcrraagASVVTELRCQCLQTLQG 52
sooty mangabey                macaaphavpgpprllrvallllvllvatcrraagAPVVTELRCQCLQTLQG 52
drill                         macaaphavpgppwllrvallllvllvatcrraagAPVVTELHCQCLQTLQG 52
olive baboon                  macaaphaapgpprllrvallllvllvatcrraagAPVVTELRCQCLQTLQG 52
African green monkey           mawaaphaasgpprllrvallllvllvatcwraagAPVVTELRCQCLQTLQG 52
CXCL1 (cynomolgus macaque)    maraalsaapsnprflqvallllll-vatgrraagASVVTELRCQCLQTLQG 51

                                                                R (CXCL1Lα)
cynomolgus macaque            IHPKNIRSVNVKFPGPHCAQTEVIAALKNGQKVCLNPTAPMVQKIMQKMLNS 104
rhesus macaque                IHPKNIRSVNVKFPGPHCAQTEVIAALKNGQKVCLNPTAPMVQKIMQKMLNS 104
Southern pig-tailed macaque   IHPKNIRSVNVKFPGPHCAQTEVIAALKNGQKVCLNPTAPMVQKIMQKMLNS 104
sooty mangabey                IHPKNIRNVNVTFPGPHCAQTEVIAALKNGQKVCLNPTAPMVQKIIQKMLNS 104
drill                         IHPKNIRNVNVTFPGPHCAQTEVIAALKNGQKVCLNPTAPMVQKIIQQMLNS 104
olive baboon                  IHPKNIRSVNVTFPGPHCTQTEVLAALKNGQKVCLNPTAPMVQKIIQKTLNS 104
African green monkey           IHPKNIRSVNVTFPGPHCAQTEVIAALKNGQKVCLNPTAPMVQKIIQKMLNS 104
CXCL1 (cynomolgus macaque)    IHPKNIQSVNVKAPGPHCAETEVIATLKNGQKACLNPASPMVQKIIKKMLNC 103

                              Ex4
cynomolgus macaque            GNANGPREKEKAYPLQFLKESILRGIKVKEEKQLAPGHTWTVFNVFNCFL 154
rhesus macaque                GNANGPREKEKAYPLQFLKESILRGIKVKEEKQLAPRHTWTVFNVFNCFL 154
Southern pig-tailed macaque   GNANGPREKEKAHPLQFLKESILRGIKVKEEKQLAPGRTWTVFNVFNCFL 154
sooty mangabey                GNANGPREKEKAYPQQFLKESILRGIKVKEEKQLAPGRTWTVFNVFNCFL 154
drill                         GNANGPREKEKAYPQQFLKESILRGIKVKEEKQLAPGRTWTVFNVFNCFL 154
olive baboon                  GNANGPREKEKAYPLQFLKESILRGIKVKEEKQLAPGRTWTVFNVFNCFL 154
African green monkey           GNANGPREKEKAYPPQFLKESILGGIKVKEEKQLAPGGTWTVFNVFNCFL 154
CXCL1 (cynomolgus macaque)    DKSN                                                107
```

B

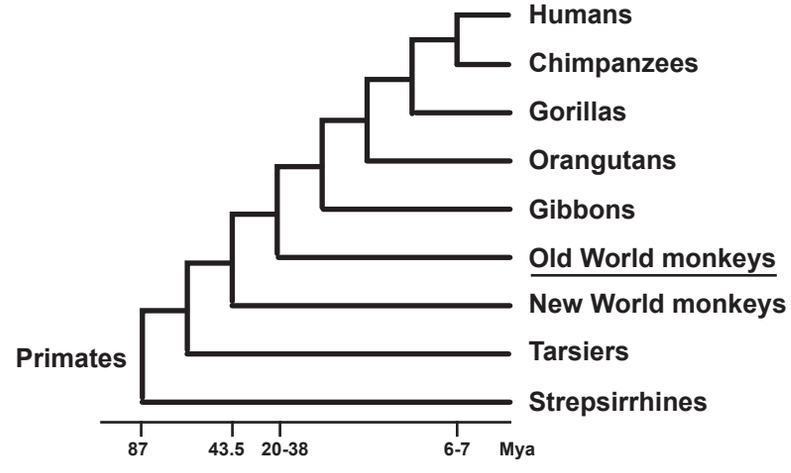

Fig. 2

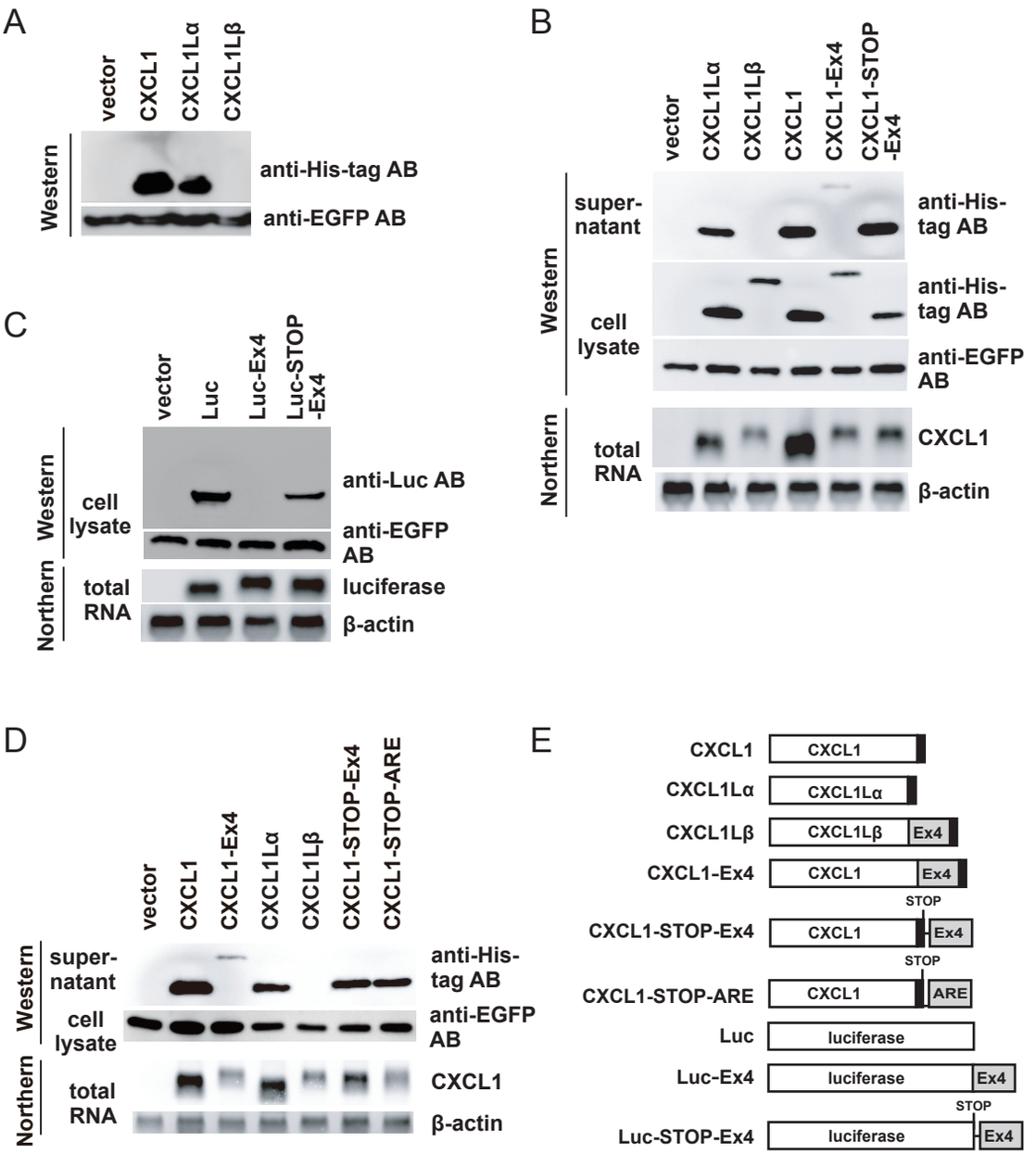
A
vector
CXCL1
CXCL1Lα
CXCL1Lβ
Western
anti-His-tag AB
anti-EGFP AB
B
vector
CXCL1Lα
CXCL1Lβ
CXCL1
CXCL1-Ex4
CXCL1-STOP -Ex4
Western
super-natant
anti-His-tag AB
cell lysate
anti-His-tag AB
anti-EGFP AB
Northern
total RNA
CXCL1
β-actin
C
vector
Luc
Luc-Ex4
Luc-STOP -Ex4
Western
cell lysate
anti-Luc AB
anti-EGFP AB
Northern
total RNA
luciferase
β-actin
D
vector
CXCL1
CXCL1-Ex4
CXCL1Lα
CXCL1Lβ
CXCL1-STOP-Ex4
CXCL1-STOP-ARE
Western
super-natant
anti-His-tag AB
cell lysate
anti-EGFP AB
Northern
total RNA
CXCL1
β-actin
E
CXCL1
CXCL1Lα
CXCL1Lβ
Ex4
CXCL1-Ex4
STOP
CXCL1-STOP-Ex4
STOP
CXCL1-STOP-ARE
ARE
Luc
luciferase
Luc-Ex4
STOP
Luc-STOP-Ex4

Fig. 3

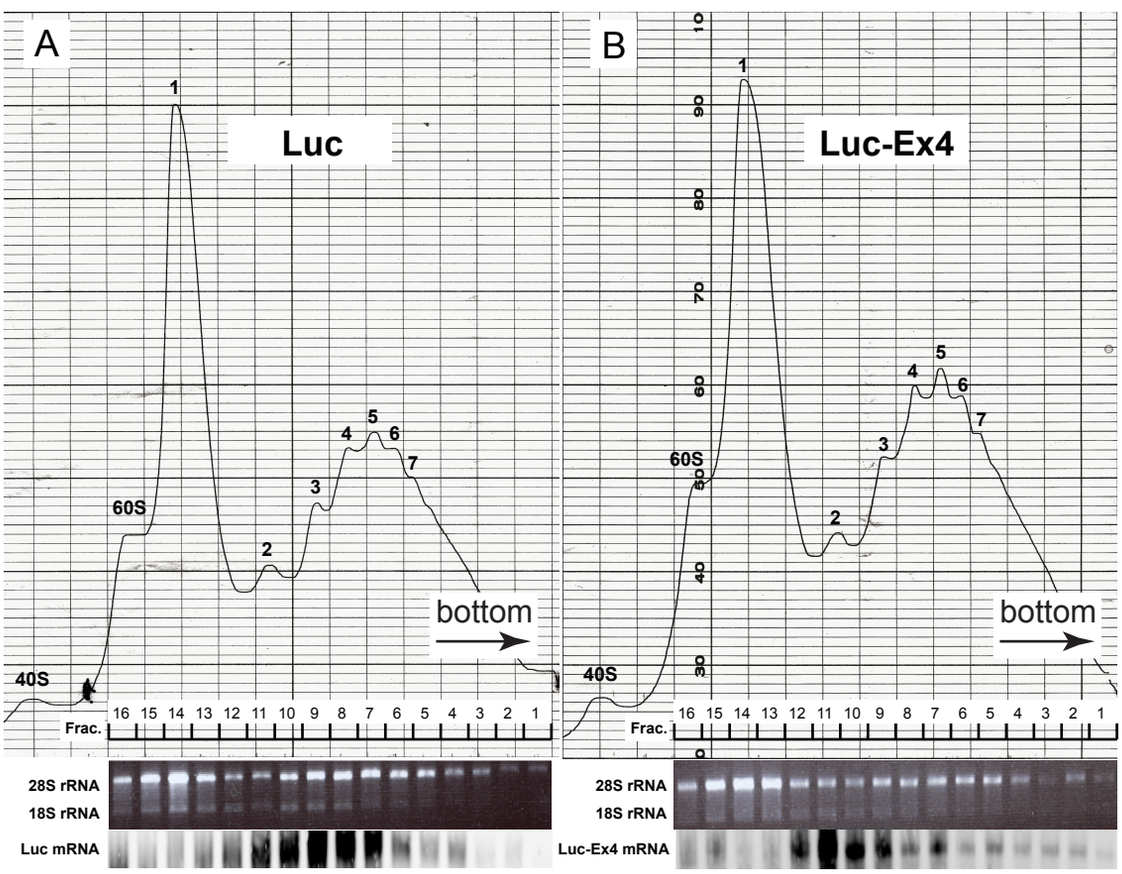
A
Luc
1
2
3
4
5
6
7
60S
40S
bottom
Frac.
16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
28S rRNA
18S rRNA
Luc mRNA
B
Luc-Ex4
1
2
3
4
5
6
7
60S
40S
bottom
Frac.
16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
28S rRNA
18S rRNA
Luc-Ex4 mRNA

Fig. 4

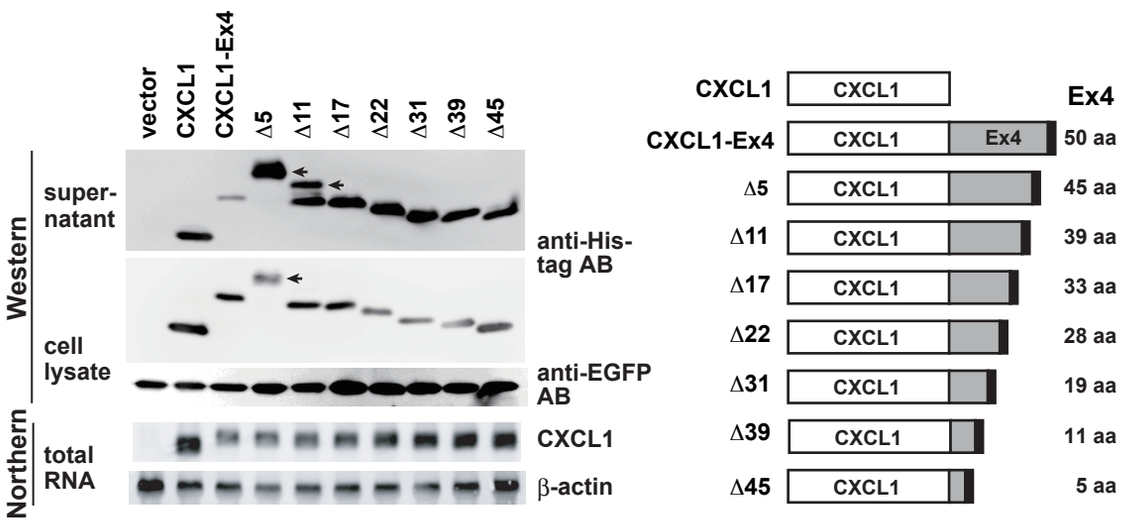
vector
CXCL1
CXCL1-Ex4
Δ5
Δ11
Δ17
Δ22
Δ31
Δ39
Δ45
Western
super-natant
cell lysate
anti-His-tag AB
anti-EGFP AB
Northern
total RNA
CXCL1
β-actin
CXCL1
CXCL1-Ex4
Δ5
Δ11
Δ17
Δ22
Δ31
Δ39
Δ45
Ex4
50 aa
45 aa
39 aa
33 aa
28 aa
19 aa
11 aa
5 aa

Fig. 5

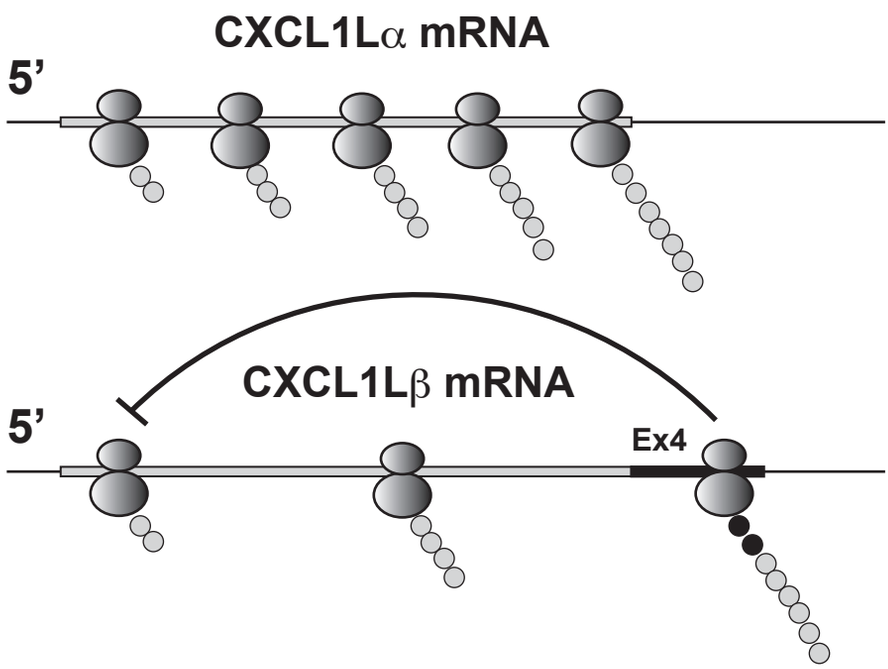

Fig. 6