

Molecular characterization and expression analysis of mouse epidermal growth factor-like domain 8

IK-JIN SONG^{1*}, MUHAMMAD IKRAM^{2,3*}, FAZLI SUBHAN^{2,3}, DA-JEONG CHOI^{2,3},
JA-RANG LEE⁴, HEUI-SOO KIM⁴, YOUNG-TAK LIM¹ and SIK YOON^{2,3}

Departments of ¹Pediatrics and ²Anatomy and Medical Research Institute,
³Immune-Network Pioneer Research Center, Pusan National University School of Medicine, Yangsan,
Gyeongsangnam-do 626-870; ⁴Department of Biological Sciences, College of Natural Sciences,
Pusan National University, Busan 609-735, Republic of Korea

Received November 4, 2014; Accepted May 20, 2015

DOI: 10.3892/ijmm.2015.2252

Abstract. Epidermal growth factor (EGF)-like (EGFL) domain, a common structural module in numerous secreted or transmembrane proteins, is generally involved in protein-protein interactions. To date, several EGFL proteins have been identified and characterized, but little is known about EGFL domain 8 (EGFL8). The present study reported the molecular characterization and expression analysis of EGFL8 in mice. Mouse EGFL8 amplified using a reverse transcription-polymerase chain reaction approach was sequenced and characterized. Mouse EGFL8 encodes a protein of 293 amino acids with two EGFL domains, an Emilin-like domain and a Ca²⁺-binding EGFL domain, which has a molecular mass of 32 kDa. The coding sequence has a high degree of amino acid sequence identity across species, and the EGFL domain has been highly conserved in various species during evolutionary radiation. A phylogenetic tree calculated using the neighbor-joining method revealed that EGFL8 and EGFL7 are more closely associated with each other than either is to EGFL3, and they cluster with EGFL6. It was found that mouse EGFL8 protein was highly expressed in diverse mouse tissue types, including the thymus, lymph nodes, testis, ovaries, epididymis, ductus deferens, ileum, colon, stomach, esophagus, lung, uterus, urinary bladder, skin, spleen, adrenal glands and penis. These results are of great use in understanding the biological roles of mouse EGFL8 for further study.

Introduction

Epidermal growth factor (EGF)-like domain 8 (EGFL8) is a member of the EGFL domain family and was identified in mice as a paralog of EGFL7 using the Basic Local Alignment Search Tool (1). EGFL8 displays the same overall domain structure as that of EGFL7 (1). EGFL domain, which consists of 30-40 amino acids and has a significant identity with EGF, is a common structural module in numerous secreted or transmembrane proteins and is generally involved in protein-protein interactions (2). The EGFL domain of each protein has diverse physiological functions. For instance, the EGFL domain of CD93 has a central role in eliciting angiogenesis by stimulating proliferation, migration and *in vitro* tube formation of human umbilical vein endothelial cells (3). The EGFLs in thrombospondins induce synaptogenesis by interacting with neuronal cell-surface receptors such as $\alpha 2\delta$ -1 (4). The EGFL structures in a large number of the coagulation factors as well as in thrombomodulin mediate complex formation, which is a critical process in the blood coagulation cascade and in the thrombomodulin-protein C anti-coagulant pathway (5). In addition, the EGFLs of the Notch family proteins have essential roles in controlling the Notch signaling pathway by binding their ligands on adjacent cells, and their mutations and polymorphisms are implicated in the development of severe diseases (6). For example, polymorphisms in the EGFL domain of Notch-3 result in symptomatic ischemic cerebrovascular disease (7).

To date, several EGFL proteins have been identified using a complementary DNA (cDNA) library from human brain tissue, but less is known about their physiological features. EGFL7 is secreted by endothelial cells and is implicated in the regulation of blood vessel formation and cell migration through interaction with receptors of the Notch family (1,8-11). EGFL7 binds to the extracellular domains of all four Notch receptor isoforms and inhibits Jagged-induced Notch signaling (12). EGFL7 has recently been suggested as a novel target for the design of therapies aimed at preventing cancer progression by interfering with the processes of tumor immune evasion, as the regulatory role of EGFL7 in tumor endothelial cell activities has been implicated in tumor escape from immunity through downregulation of immune cell extravasation (13). EGFL7

Correspondence to: Professor Sik Yoon, Department of Anatomy, Pusan National University School of Medicine, 49 Busandaehak-ro, Yangsan, Gyeongsangnam-do 626-870, Republic of Korea
E-mail: sikyoon@pusan.ac.kr

*Contributed equally

Key words: molecular characterization, epidermal growth factor-like domain 8, expression analysis

expression is significantly higher in low-grade invasive lesions and is correlated with favorable prognosis in human breast cancer. Furthermore, EGFL7 has been identified as an inhibitor of neural stem cell maintenance (14). EGFL6, an extracellular matrix protein, has been found in human subcutaneous adipose tissue and is a paracrine/autocrine growth factor of adipose tissue in obesity (15). EGFL6 is overexpressed in benign meningioma tissues and serum (16).

Downregulation of EGFL8 expression has been observed in gastric cancer, and is significantly correlated with high tumor-node-metastasis stage and poor prognosis in gastric and colorectal cancer (17,18). Recent studies by our group have demonstrated that EGFL8 has inhibitory effects on mouse thymic epithelial cells and thymocytes, indicating that it is a negative regulatory molecule in T-cell development in the mouse thymus (19,20). However, the physiological characteristics and biological significance of EGFL8 have remained to be elucidated. In addition, little information is available on the molecular characterization and expression profile of EGFL8 (1). Thus, the present study was performed to investigate the molecular and expressional characterization of EGFL8 in various mouse tissue types.

Materials and methods

Cell lines and cell culture. The generation, maintenance and functional characterization of mouse thymic sub-capsular cortex or thymic nurse epithelial cells (SNECs) have been described previously (21). SNECs constitutively express the SV40 T antigen transgene and class I antigens of the major histocompatibility complex, and they can be induced to express major histocompatibility complex class II antigens via stimulation with recombinant interferon- γ (IFN- γ) as well as produce granulocyte-macrophage colony-stimulating factor (GM-CSF) (21). SNECs were provided by Dr Barbara B. Knowles (Jackson Laboratory, Bar Harbor, ME, USA) and were cultured in Dulbecco's modified Eagle's medium containing 10% (v/v) fetal bovine serum (both from Gibco-BRL, Invitrogen Life Technologies, Carlsbad, CA, USA), 2 mM glutamine (Sigma-Aldrich, St. Louis, MO, USA), 100 U/ml penicillin, and 100 μ g/ml streptomycin (both from Invitrogen Life Technologies) at 37°C in a 5% CO₂-enriched atmosphere.

RNA extraction. Total RNA was extracted from each tissue or cell sample. In case of tissues, they were transferred to a mortar containing liquid nitrogen and ground to fine powder using a pestle. Total RNA from tissue or cell samples was isolated using TRIzol reagent (Invitrogen Life Technologies) following the manufacturer's instructions. After the addition of chloroform, the mixture was vigorously shaken by hand for 15-30 sec and incubated at room temperature for 5 min. The mixture was then centrifuged at 15,300 \times g for 15 min at 4°C. The upper aqueous phase containing the total RNA was carefully removed and placed in a fresh 50 ml tube, where it was precipitated with isopropanol, washed with ethanol and re-suspended in diethylpyrocyanate-treated water. RNA concentration and purity were determined using a spectrophotometer (NanoDrop 2000c; Thermo Scientific, Wilmington, DE, USA) at absorbances of 260 and 280 nm. Samples exhibiting a (260/280) absorbance ratio of ≥ 1.7 and strong 28S- and 18S-ribosomal RNA bands on 1% agarose gels were used for further analysis. RNA

samples were stored at -80°C until required. Polyadenylated [poly(A)] RNA was purified from total RNA using a messenger RNA (mRNA) isolation kit (200347; Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions. Briefly, 5 mg total RNA was mixed with 5 ml elution buffer and the mixture was hybridized to 0.2 g oligo(dT) cellulose at room temperature with gentle agitation followed by high-salt (low-stringency) and low-salt (high-stringency) washes. These washes removed unwanted components of the crude lysate, including proteins, carbohydrates, lipids, DNA, transfer RNA and a significant amount of ribosomal RNA from poly(A) mRNA. The oligo(dT) cellulose was loaded into a push column and mRNA was eluted with 65°C elution buffer. The mRNA was concentrated via ethanol precipitation, and the mRNA concentration was determined by measuring the absorbance at 260 nm.

Cloning cDNA of the open reading frame (ORF) region of mouse EGFL8. The first strand of cDNA was synthesized by the reverse transcription using 1-2 μ g of poly(A) mRNA. The ORF region of mouse EGFL8 was amplified with the following primers: EGFL8-sense (5'-AGC CCG CTC CCG CAC CAT-3') and EGFL8-anti-sense (5'-gcc ctt ggc tgc acg etc a-3'), derived from mouse EGFL8 (GenBank accession no. NM_152922). PCR amplification of the cDNA was performed for 25 cycles at 94°C for 30 sec, 60°C for 30 sec and 72°C for 60 sec in an automated thermal cycler (TC-312; Techne, Teddington, UK) in a final volume of 25 μ l containing 4 μ l cDNA solution, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.1% Triton X-100, 0.2 mM deoxynucleotide triphosphate mixture (Invitrogen Life Technologies), 0.5 pmol of each primer and 5 units TaqDNA polymerase (Promega). After PCR, the amplified products were separated on a 1.5% agarose gel, purified with a QIAquick gel extraction kit (Qiagen, Hilden, Germany) and cloned into a pGEM-T-easy vector (Promega, Madison, WI, USA). The cloned DNA was isolated using a High Pure plasmid isolation kit (Roche, Basel, Switzerland).

Sequence analysis. Sequence alignments of EGFL8 and other homologous EGFL family genes were performed using the BioEdit sequence alignment editor program. To analyze functional domains, hydrophobicity, nucleotide and amino acid sequence composition, and synonymous/non-synonymous substitution of EGFL8, online tools were used, including Pfam (22), Expasy (23), MEGA5 (24) and the National Center for Biotechnology Information (NCBI) Conserved Domain Database (CDD; <https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Neighbor-joining phylogenetic analysis (25) was performed using CLUSTAL W (26) and MEGA (24). EGFL sequences were retrieved from the GenBank database (<http://www.ncbi.nlm.nih.gov/genbank/>) with the aid of the Basic Local Alignment Search Tool network server (27).

RT-PCR amplification. Mouse EGFL8 transcripts were analyzed through RT-PCR amplification. First-strand cDNA was obtained via reverse transcription using 2 μ g total RNA. The reaction was performed in 20 μ l buffer containing 0.5 μ g oligo(dT)₁₂₋₁₈ primer, 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl₂, 40 mM dithiothreitol, 0.5 mM deoxynucleotide triphosphate mixture, 10 units RNase inhibitor and 200 units Moloney murine leukemia virus reverse transcriptase (all from

Invitrogen Life Technologies). After incubation at 37°C for 60 min, the reaction was stopped via heating at 70°C for 15 min. To remove the remaining RNA, 1 μ l *Escherichia coli* RNase H (4 mg/ml; Promega) was added to the reaction mixture followed by incubation at 37°C for 30 min. The cDNA was used as a template for PCR amplification with gene-specific primers. As a standard control, GAPDH was amplified using the primer pair GAPDH-sense (5'-GAA ATC CCA TCA CCA TCT TCC AGG-3') and GAPDH-anti-sense (5'-GAG CCC CAG CCT TCT CCA TG-3'), derived from mouse GAPDH (GenBank accession no. NM_002046). The mouse full-length EGFL8 transcript was amplified using the primer pair EGFL8-sense (5'-TTT CAA AGA GAG TTT GGG AGT G-3') and EGFL8-anti-sense (5'-CAC CAC GTG TGT CTG TGG TA-3'), derived from mouse EGFL8 (GenBank accession no. NM_152922). PCR amplification of the cDNA was performed in an automated thermal cycler (TC-312; Techne) in a final volume of 25 μ l containing 4 μ l cDNA solution, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, 1.5 mM MgCl₂, 0.1% Triton X-100, 0.2 mM deoxynucleotide triphosphate mixture (Invitrogen Life Technologies), 0.5 pmol of each primer and 5 units TaqDNA polymerase (Promega). All PCR reactions were performed in 30-35 cycles at 94°C for 30 sec, 57°C for 30 sec and 72°C for 30 sec. After PCR, the amplified products were separated by 1.5% agarose gel and visualized using ethidium bromide staining under ultraviolet light using a Gel Doc XR system (Bio-Rad, Hercules, CA, USA).

Molecular cloning, sequencing and nucleotide sequence accession numbers. Since an additional band to the specific EGFL8 band was identified in certain mouse organs when RT-PCR analysis of the mouse EGFL8 was performed in various mouse tissue types, it was hypothesized that this band may represent an isoform of mouse EGFL8. To test the hypothesis, the RT-PCR products were separated using a 1.5% agarose gel and purified using the QIAquick® Gel Extraction kit (Qiagen) and cloned into a pGEM-T-easy vector (Promega). The cloned DNA was isolated using a High Pure plasmid isolation kit (Roche), amplified using the primer pair, EGFL8-2-sense (5'-TTG GGG TCC CTT TGA GAC CT-3') and EGFL8-2-anti-sense (5'-TGC TGG TTC CTC TCC GTT AC-3'), and sequenced by Macrogen (Seoul, Korea) to confirm their identities. The nucleotide sequences of the EGFL8 isoform (EGFL8-2) from mouse thymic epithelial cells reported herein have been deposited in the DNA Data Bank of Japan/European Molecular Biology Laboratory/GenBank nucleotide sequence databases with the accession number AB613266.

Experimental animals. Adult male specific pathogen-free C57BL/6 mice (weighing 22-25 g) were purchased from Dae Han Bio Link (Seoul, Korea). Six animals were sacrificed by cervical dislocation. Mice were used at 8-10 weeks of age. Animal care and all experimental procedures were performed in accordance with the Guide for Animal Experiments edited by the Korean Academy of Medical Sciences.

Thymocyte and thymic stromal cell isolation. For the isolation of mouse thymocytes and thymic stromal cells, 2-3 thymi were dissected from mice immediately after sacrifice and trimmed of fat and connective tissues. Small incisions (2-3 mm) were made into the capsules with a pair of razors, and the thymi

were gently agitated in 30 ml RPMI-1640 (Gibco-BRL) using a magnetic stirrer at 4°C for 40 min. The resulting thymic fragments and supernatant were transferred into separate tubes. For the isolation of thymocytes, the supernatant was passed three times through 70 μ m mesh and centrifuged at 3,800 x g for 2 min at 4°C. The cell pellet was re-suspended in 20 ml ammonium chloride-potassium lysis solution (0.15 M NH₄Cl, 1 mM KHCO₃ and 0.1 mM disodium EDTA) to remove red blood cells. The thymocyte suspension was washed three times with Hank's buffered salt solution (HBSS) buffer. The thymocytes were then re-suspended in HBSS buffer, and viable cells were counted using a hemocytometer after trypan blue staining. For the isolation of thymic stromal cells, the thymic fragments were transferred into 5 ml RPMI-1640 containing 0.125% (w/v) collagenase D and 0.1% (w/v) DNase I (both from Roche) and incubated for 15 min with gentle shaking in a water bath at 37°C. The thymic fragments in the enzyme mixtures were carefully dispersed several times with a Pasteur pipette, and the supernatant was removed after fragments had settled and replaced with fresh enzyme mixture. Gentle mechanical agitation was provided using a 5-ml syringe and 18G, 21G and 23G needles. Tissue fragments were allowed to settle, and the supernatant was discarded. This digestion process was repeated four times until the tissue was fully digested. Cells liberated by the fourth, fifth and sixth digest were saved, filtered through 100- μ m mesh to remove undigested particles and washed three times with HBSS buffer. They were then re-suspended in HBSS buffer and viable cells were counted using a hemocytometer after trypan blue staining.

Western blot analysis. Tissue proteins were isolated using a protein extraction solution (PRO-PREP Protein Extraction Solution; Intron Biotechnology, Seoul, Korea) from 27 adult mouse tissues, including the thymus, spleen, lymph nodes, stomach, ileum, colon, esophagus, tongue, liver, kidney, testis, epididymis, penis, ductus deferens, ovary, uterus, urinary bladder, adrenal glands, trachea, heart, submandibular gland, skin, fat, lung, pancreas, aorta and skeletal muscle. Total proteins from the cultured cells were then extracted using a protein extraction solution (Intron Biotechnology) supplemented with a protease inhibitor mixture (Sigma-Aldrich). The lysates were centrifuged at 17,900 x g for 15 min at 4°C. Protein concentrations were determined using the bicinchoninic acid protein assay method (B9643; Sigma-Aldrich). Equal amounts of protein samples were heated for 10 min at 95°C in sample buffer and separated by 10% SDS-PAGE, using a Mini-Protean III system (Bio-Rad). The proteins were transferred onto a polyvinylidene fluoride membrane via semi-dry transfer (both from Bio-Rad), and the membrane was incubated overnight at 4°C with rabbit polyclonal anti-EGFL8 (AV42656; Sigma-Aldrich) at a dilution of 1:500 and mouse monoclonal anti- β -actin-horseradish peroxidase (HRP) (Ab8226; Abcam, Cambridge, UK) antibody at a dilution of 1:1,000 in Tris-buffered saline (TBS; 20 mM Tris-HCl, 150 mM NaCl, pH 7.4) containing 2% skimmed milk. After three washes with TBS containing 0.1% Tween-20 (TBS-T) and 1% skimmed milk, the membrane was incubated for 2 h at room temperature with secondary antibodies, donkey anti-rabbit immunoglobulin G-HRP (sc-2313; Santa Cruz Biotechnology, Inc., Dallas, TX, USA), diluted 1:10,000 and washed three times with TBS-T. Immunoreactivity was detected with enhanced chemilumines-

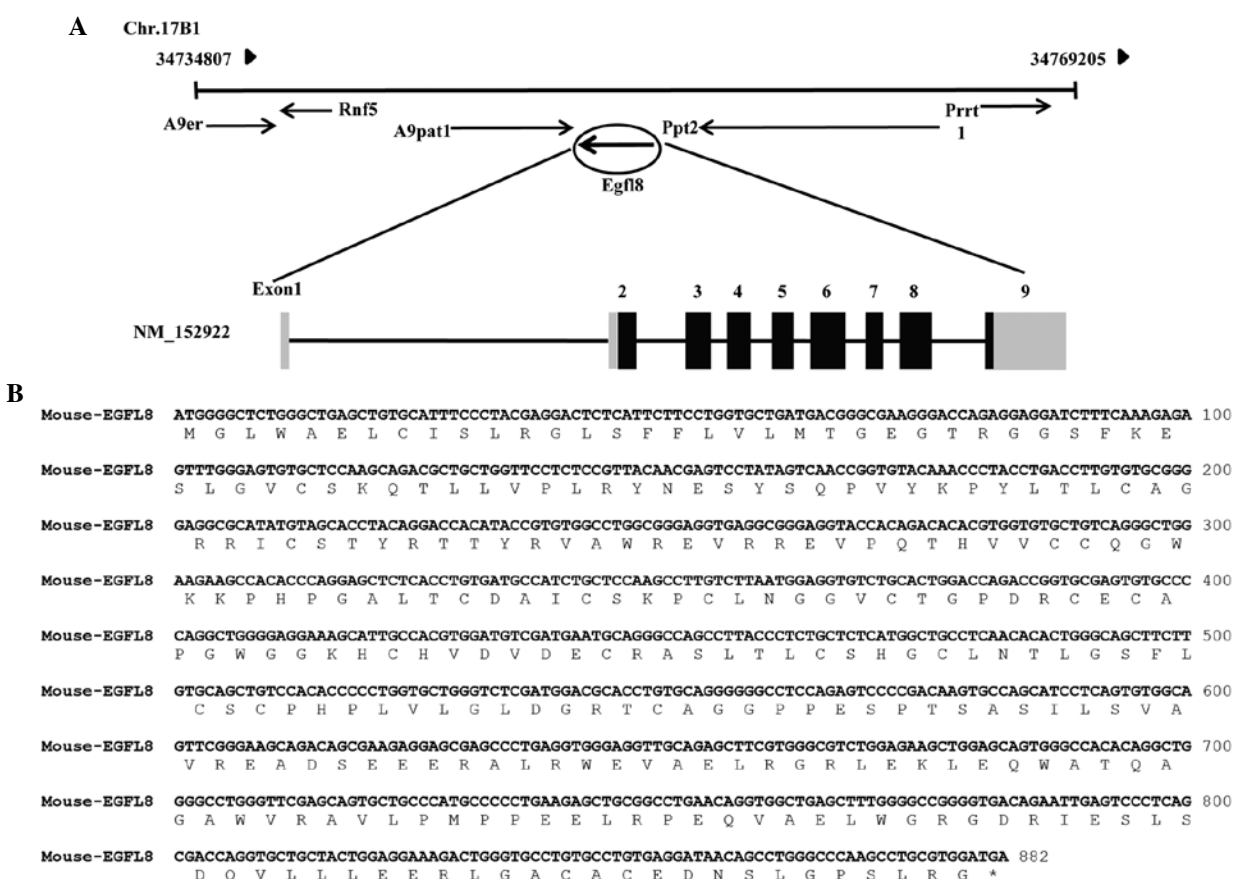


Figure 1. (A) Analysis of the genomic structure of the mouse EGFL8 gene using the NCBI server (<http://www.ncbi.nlm.nih.gov/>) and (B) open reading frame region nucleotide sequence identification. The nucleotide sequence of mouse EGFL8 was identified from mouse thymus tissue. Amino acids are aligned with the first nucleotide of each codon using the ORF finder from the NCBI server (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>).

cence (SuperSignal West Pico Chemiluminescent Substrate kit; Pierce, Rockford, IL, USA) according to manufacturer's instructions. Images were captured and quantified using the LAS-3000 imaging system (Fujifilm, Tokyo, Japan).

Results and Discussion

Molecular cloning and characterization of mouse EGFL8. To isolate the full-length EGFL8 cDNA, total RNA from thymic epithelial cells from C57BL/6 mice was amplified using RT-PCR, and the amplicon was cloned and sequenced. Searching the full-length cDNA of EGFL8 in the mouse genome using NCBI server (<http://blast.ncbi.nlm.nih.gov/>), mouse EGFL8 was located on chromosome 17B1 from 34,734,807 to 34,769,205 bp and was identified to consist of nine coding exons (Fig. 1A). The cDNA contains an ORF of 879 bp, which encodes a putative 32-kDa protein of 293 amino acids (Fig. 1B). Based on the deduced amino acid sequence from the mature peptide region of this EGFL8 cDNA, a 3-dimensional molecular model (Fig. 2) was generated using the automated protein-modeling server Swiss-Model (<http://swissmodel.expasy.org>). The gene product was a slightly acidic protein with an isoelectric point of pH 6.13 and had a molecular mass of 32 kDa (Fig. 3).

Base composition and putative amino acids were calculated from the cDNA sequence data. The average G+C content of EGFL8 was 61.4% in the species examined (Fig. 4A). Mouse EGFL8 sequences had a G+C content of 60.8%, whereas

human EGFL8 sequences had a G+C content of 63.3%. Amino acids with high prevalence in mouse EGFL8 are leucine (Leu) (12.84%), glycine (Gly) (9.8%) and glutamic acid (Glu) (8.78%) (Fig. 4B).

Multiple alignment of the amino acid sequences of EGFL8 proteins from various vertebrates revealed that the EGF-like domain is highly conserved in various species (human, monkey, cattle, pig, rat and frog) (Fig. 5). This coding sequence shared 78.4% nucleotide sequence similarity and 79.7% amino acid sequence identity with the homologous sequences of human EGFL8 cDNA. Sequence identity of amino acids was 87.5% between mouse EGFL8 and rat EGFL8 (Table I). The domain structure of EGFL8 was determined using the SignalP server (<http://www.cbs.dtu.dk/services/SignalP/>) (28) to predict signal sequences and the Pfam database search (<http://www.sanger.ac.uk/resources/databases/pfam.html>) was used to identify protein domains. EGFL8 contained a predicted N-terminal signal sequence, suggesting that EGFL8 is a secreted protein composed mainly of an Emilin-like (EMI) domain, followed by two EGF domains and a sub-class of EGF domains that bind Ca^{2+} (Figs. 5 and 2B). Of note, EGFL8 contained an amino acid sequence similar to that of the Delta:Serrate:LAG2 domain conserved in ligands of Notch receptors (29).

The level of synonymous substitutions per site (Ks) and non-synonymous substitutions (Ka) in the EGFL8 genes was also calculated (Table II). As Ks values are influenced by selection to a greater extent than Ka values, their ratio is a good indicator of

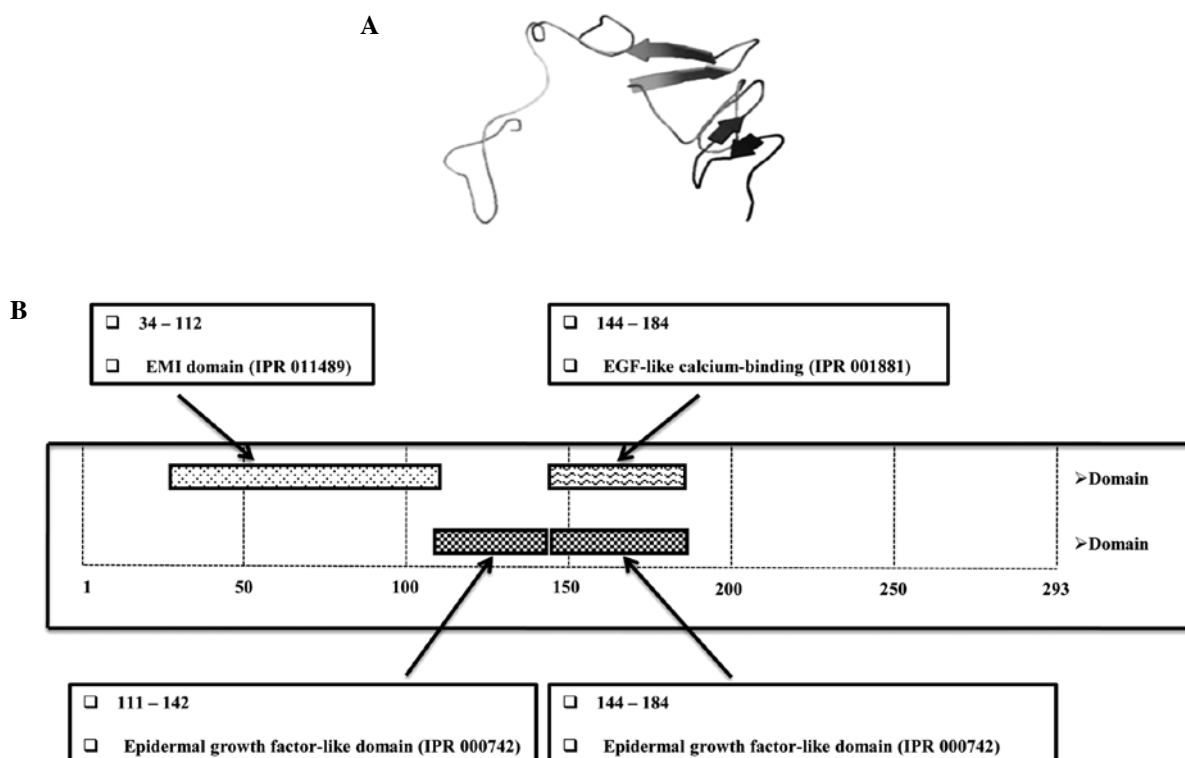


Figure 2. (A) 3-Dimensional structure predicted by the automated protein-modeling server Swiss-Model (<http://swissmodel.expasy.org>). (B) Four domains and their positions in mouse EGF-like domain 8 protein determined by the SignalP server (<http://www.cbs.dtu.dk/services/SignalP/>) predicted by Swiss-Model. EMI, Emilin-like; EGF, epidermal growth factor.

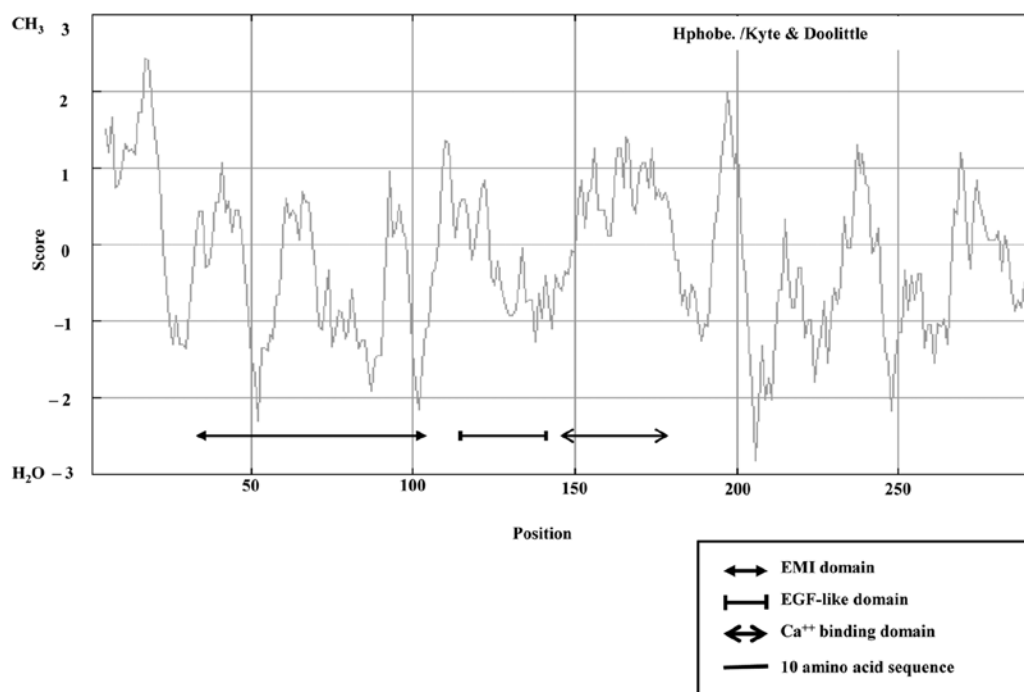


Figure 3. Hydrophobicity plot of the deduced protein sequence of mouse epidermal growth factor-like domain 8 using the ProtScale tool from the ExPASy server (<http://web.expasy.org/protscale/>). Hydrophobic regions lie above the horizontal line at Y-axis = 0 and hydrophilic regions lie below. The scale at the bottom is in amino acids, with the N terminus on the left. Designated regions include an EMI domain, an EGFL domain and a Ca²⁺-binding domain, a sub-class of EGFL domains that bind Ca²⁺. EMI, Emilin-like; EGFL, epidermal growth factor-like.

selection. For the comparison of mouse to human and mouse to rat, Ka/Ks values of 0.32 and 0.46 were determined, respectively. The number of synonymous substitutions was higher than that

of non-synonymous substitutions in EGFL8 (Table II). Positive selection has not occurred in EGFL8. Therefore, EGFL8 genes were stable and amino acid changes were deleterious.

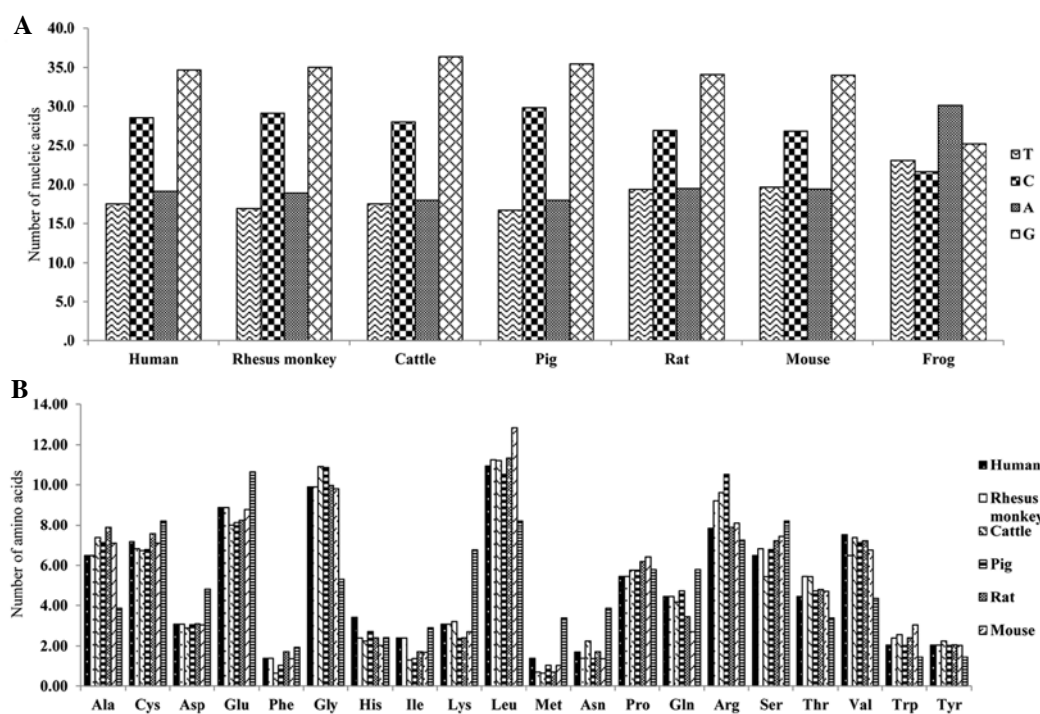


Figure 4. Composition of (A) nucleic acids and (B) amino acids in the open reading frame (coding DNA sequence) region of epidermal growth factor-like domain 8.

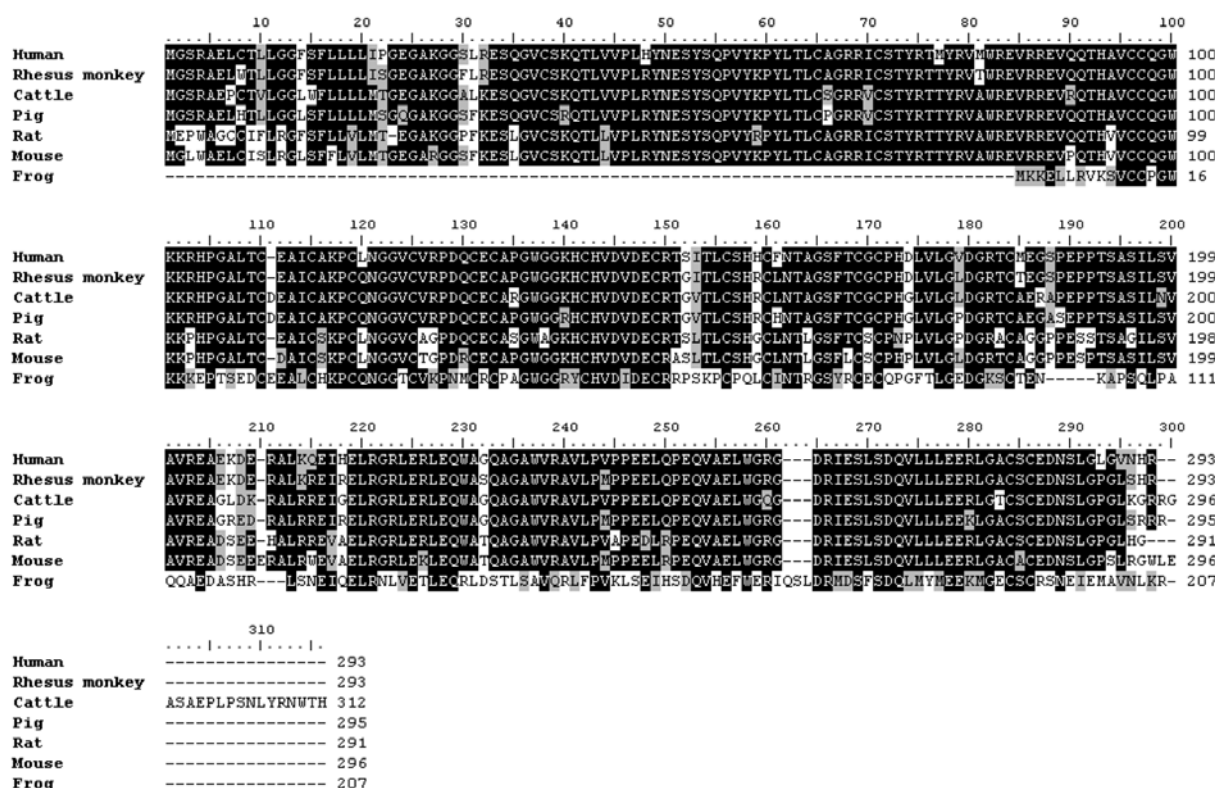


Figure 5. Multiple alignment analysis of amino acid sequences of EGFL8 in various species using the BioEdit sequence alignment editor program. Open and closed sequences indicate the variant and conserved sequences, respectively. Designated regions include an Emilin-like domain, EGFL domain and a Ca^{2+} -binding domain, a sub-class of EGFL domains that bind Ca^{2+} . EGFL, epidermal growth factor-like.

In order to illustrate the evolutionary associations among mouse EGFL family members, a phylogenetic tree was constructed by the neighbor-joining method using cDNA

sequences. Mouse EGFL8 and EGFL7 are more closely related to each other than either is to *EGFL3*, and they cluster with *EGFL6* (Fig. 6).

Table I. Similarity and identity of the open reading frame region for epidermal growth factor-like domain 8 genes (%).

Species	Human	Rhesus monkey	Cattle	Pig	Rat	Mouse	Frog
Human	-	93.5	81.7	86.7	79.5	79.7	29.1
Rhesus monkey	96.3	-	83.9	90.1	80.2	80.7	28.8
Cattle	85.5	86.5	-	85.5	75.0	75.7	27.6
Pig	85.8	86.8	90.4	-	78.9	80.1	29.8
Rat	79.3	80.0	80.1	78.6	-	87.5	26.9
Mouse	78.4	79.5	79.0	78.2	90.2	-	26.6
Frog	36.1	35.8	37.1	36.3	34.4	34.8	-

Table II. Synonymous substitutions per site (Ks) and non-synonymous substitutions per site (Ka) in epidermal growth factor-like domain 8 genes (expressed as Ka/Ks).

Species	Human	Rhesus monkey	Cattle	Pig	Rat	Mouse	Frog
Human	-	0.66	0.30	0.28	0.32	0.32	0.56
Rhesus monkey	0.03/0.05	-	0.26	0.18	0.29	0.29	0.50
Cattle	0.09/0.31	0.08/0.29	-	0.28	0.33	0.26	0.59
Pig	0.09/0.32	0.06/0.34	0.06/0.20	-	0.28	0.27	0.55
Rat	0.16/0.49	0.14/0.48	0.14/0.43	0.14/0.51	-	0.46	0.60
Mouse	0.16/0.50	0.14/0.49	0.13/0.50	0.14/0.51	0.07/0.15	-	0.65
Frog	0.61/1.08	0.60/1.20	0.60/1.01	0.61/1.10	0.68/1.12	0.68/1.04	-

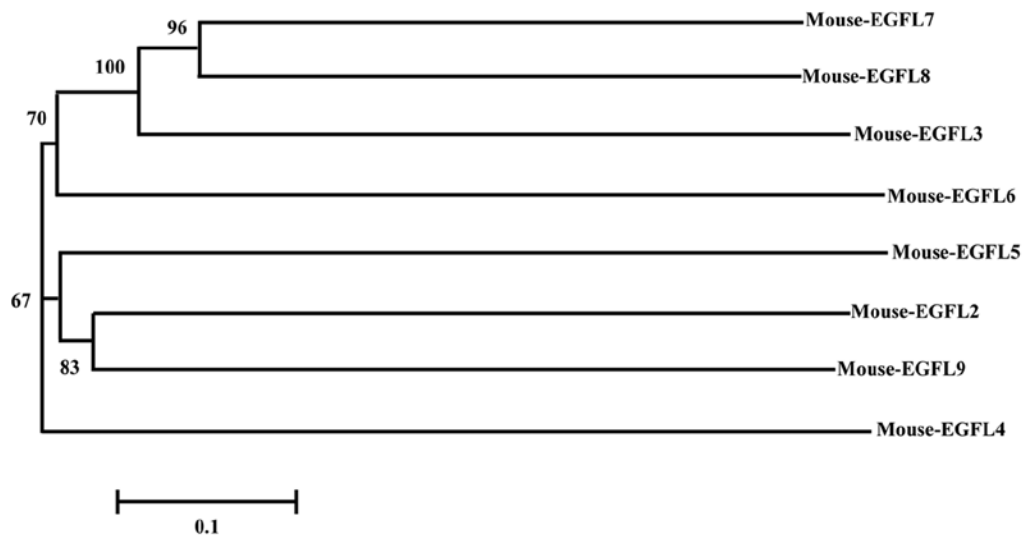


Figure 6. Phylogenetic analysis with the neighbor-joining method by CLUSTAL W (26) and MEGA5 (24) using the complementary DNA sequences of the mouse EGFL family. The branch lengths are proportional to the distances between taxa. Values at the branch points indicate the percentage support for a particular node after 1,000 bootstrap replicates were performed. EGFL, epidermal growth factor-like.

The EGFL domain, an evolutionarily conserved protein domain found in large numbers in most animal proteins, constitutes an expanding family of proteins involved in cellular activities, including blood coagulation, fibrinolysis, cell adhesion, and neural and vertebrate development (30,31). EGFL8, also known as C6 or f8, NG3 and VE-statin-2, is a secreted protein of 293 amino acids that contains two EGFL domains, an EMI domain and a Ca^{2+} -binding EGFL domain. Via these domains, EGFL8 may participate in protein-protein interactions that

correlate with cellular proliferation and developmental signaling events. Although the characterization of *EGFL8* and the presence of *EGFL8* mRNA in certain mouse tissues, including the thymus, lungs and kidney, have been mentioned briefly in the literature (1), the complete characterization of *EGFL8* and the expression pattern of *EGFL8* have remained to be elucidated.

Tissue distribution of mouse *EGFL8*. *EGFL8* was first identified in 2004 and shares a high degree of structural similarity

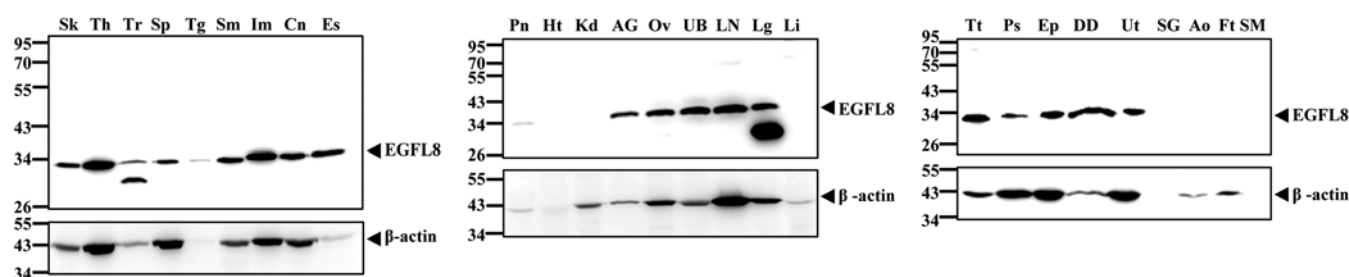


Figure 7. Representative western blot analysis of mouse EGFL8 expression in various mouse tissues: Sk, skin; Th, thymus; Tr, trachea; Sp, spleen; Tg, tongue; Sm, stomach; Im, ileum; Cn, colon; Es, esophagus; Pn, pancreas; Ht, heart; Kd, kidney; AG, adrenal gland; Ov, ovary; UB, urinary bladder; LN, lymph node; Lg, lung; Li, liver; Tt, testis; Ps, penis; Ep, epididymis; DD, ductus deferens; Ut, uterus; SG, submandibular gland; Ao, aorta; Ft, fat; and SM, skeletal muscle. β -actin was used as a loading control. EGFL, epidermal growth factor-like.

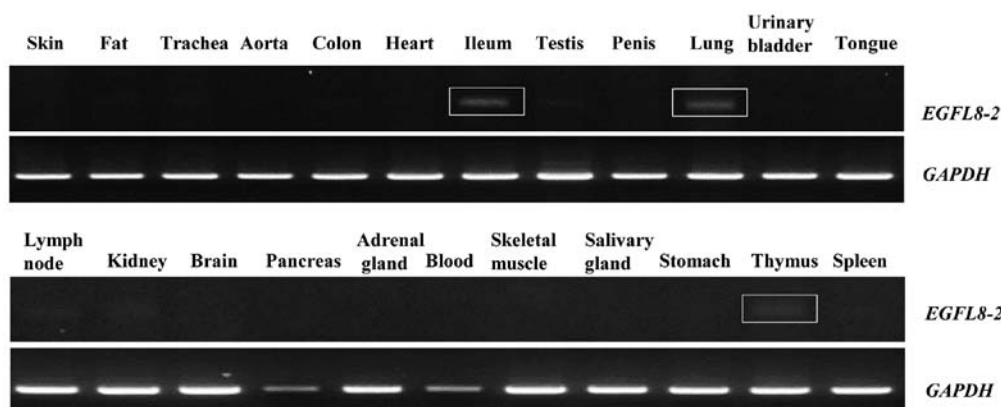


Figure 8. Representative reverse transcription polymerase chain reaction analysis of *EGFL8-2* gene expression in various mouse tissues. *EGFL8-2* amplicon was detected in the ileum, lung and thymus (box). GAPDH transcript was used as a loading control. EGFL, epidermal growth factor-like.

with EGFL7 (1). A previous study has described the tissue distribution of EGFL8 in mice (19) but tested a limited panel of mouse tissues. As little is known regarding the expression of EGFL8, the present study investigated the expression pattern of EGFL8 protein in normal C57BL/6 mouse tissues. To determine the distribution more broadly, EGFL8 protein was examined in various tissues by western blot analysis. The results showed that EGFL8 was expressed in a wide variety of tissue types in C57BL/6 mice. Western blot analysis of 27 mouse tissues showed a specific band of ~32 kDa, which represented EGFL8. EGFL8 proteins were highly expressed in diverse mouse tissues, including the thymus, lymph nodes, testis, ovary, epididymis, ductus deferens, ileum, colon, stomach, esophagus, lung, uterus, urinary bladder, skin, spleen, adrenal glands and penis (Fig. 7). However, almost no EGFL8 protein was expressed in the heart, kidney, submandibular gland, aorta, fat and skeletal muscle, while EGFL8 protein was expressed at low levels in the trachea, tongue and pancreas (Fig. 7). At the same time, a band that was smaller in size than the typical EGFL8-specific band was detected only in the lung and trachea among all organs examined in the present study. Although it is difficult to determine the precise nature of the smaller-sized band, the anti-EGFL8 antibody used in the present study appeared to be able to bind either non-specifically or specifically to a certain protein present particularly in the organs of the respiratory system including the lung and trachea. Further studies are required to clarify whether this band represents an EGFL8 isoform or not. Of note, high levels of EGFL8 mRNA were detected in diverse organs. The expression pattern of EGFL8 mRNA as determined by RT-PCR

analysis was also virtually in accord with that of EGFL8 protein as determined by western blot analysis (data not shown). These findings supported the validity of the results regarding EGFL8 expression in the present study.

Alternative splicing variant of EGFL8 in specific mouse organs. Of note, the 296-bp EGFL8 amplicon found in the present study was present in the ileum, lung and thymus (Fig. 8). To determine whether this abnormal phenomenon resulted from an alternative splicing process of the EGFL8 gene, the amplicon was cloned and sequenced. The 296-bp EGFL8 isoform, EGFL8-2 (GenBank accession no. AB613266) was shown to contain an ORF resulting from alternative splicing of the mouse full-length EGFL8 gene transcript. Analysis with the NCBI CDD indicated that mouse EGFL8-2 contains an EMI domain, which is also present in the full-length EGFL8 (Fig. 9).

EGFL8 expression in the mouse thymus, thymocytes and thymic stromal cells. The expression of EGFL8 was assessed in the mouse thymus using RT-PCR and western blotting, which demonstrated the presence of high levels of EGFL8 mRNA (Fig. 10A) and protein expression (Fig. 10B). In addition, western blot analysis of EGFL8 expression in the mouse thymus showed that EGFL8 equally expressed in freshly isolated thymocytes (Fig. 10C) and thymic stromal cells in the mouse (Fig. 10D).

The presence of EGFL8 expression in the thymus implied that the protein may have a physiological role in T-cell development, as evidenced by the findings of a recent study by our group, which indicated that EGFL8 inhibits the expression

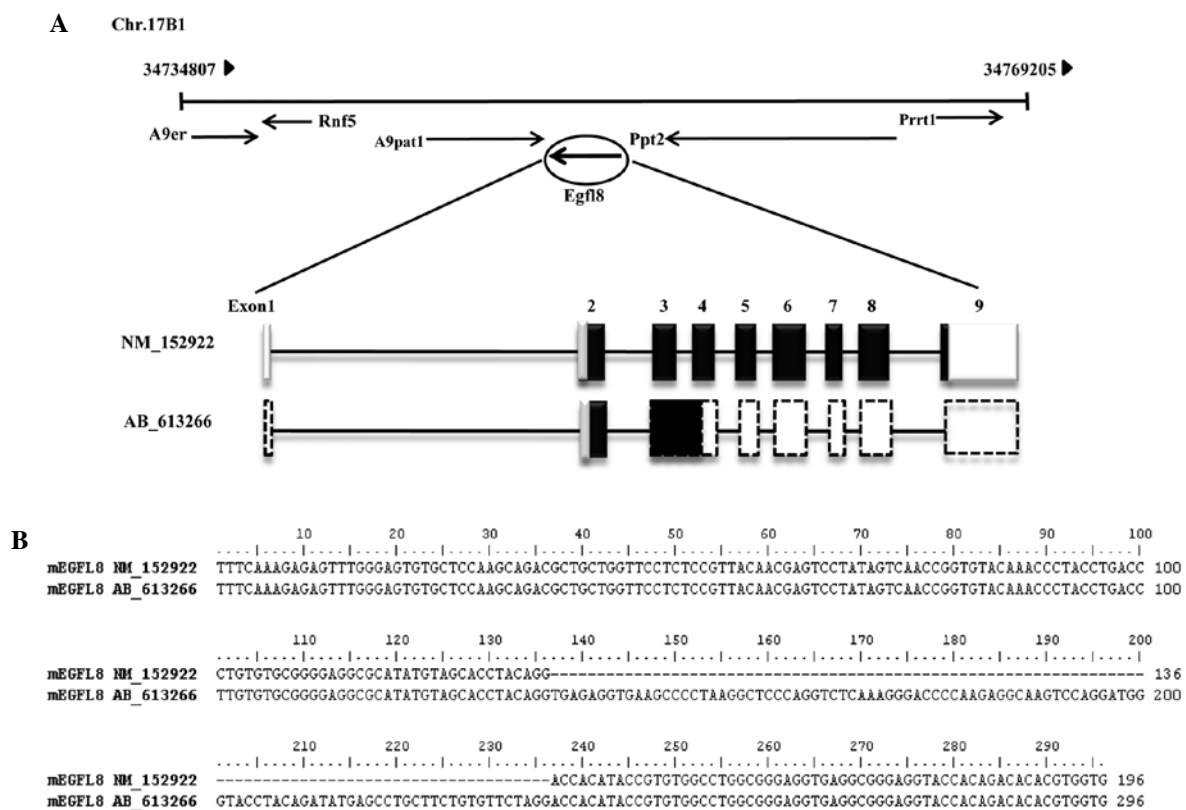


Figure 9. Comparative analysis of (A) the structure of mouse full-length EGFL8 and a newly identified 296-bp alternative transcript using the BLAT tool from UCSC genome browser (<https://genome.ucsc.edu/>) and (B) their sequences using the BioEdit sequence alignment editor program. mEGFL, mouse epidermal growth factor-like.

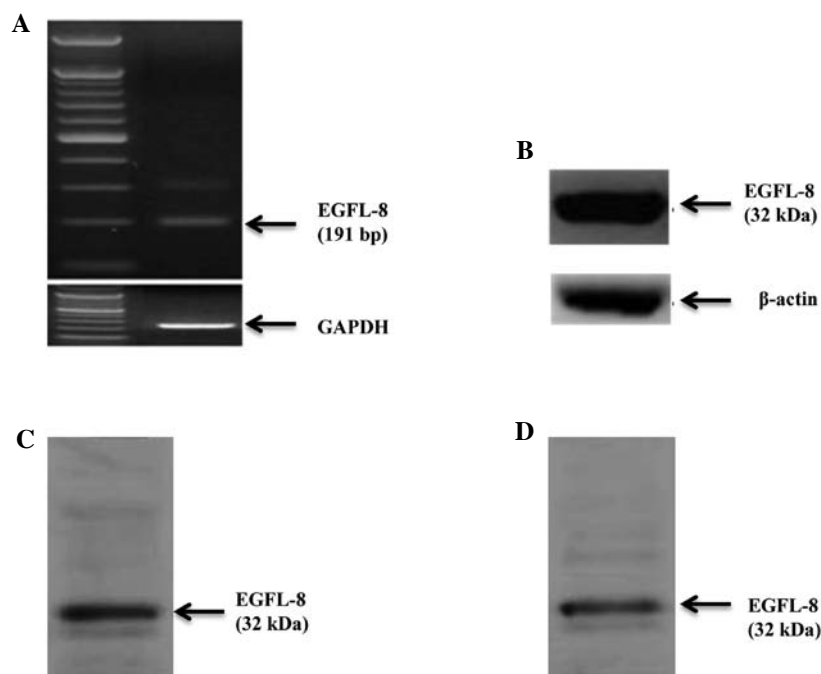


Figure 10. (A) Representative reverse transcription polymerase chain reaction analysis shows EGFL8 mRNA in mouse thymus. (B) Western blot analysis of mouse EGFL8 transcripts and protein in mouse thymus. Western blot analysis of mouse EGFL8 protein in (C) mouse thymocytes and (D) thymic stromal cells. EGFL, epidermal growth factor-like.

of the Notch downstream effectors Hes-1 and Hey-1 in the thymocytes and thymic epithelial cells of mice, indicating that EGFL8 has a role in the inhibition of T-cell development in the

mouse thymus (19,20). EGFL8 may have important implications in the physiology of the thymus, since the Notch signaling pathway is crucial in T-cell development (32); this is further

supported by the finding of the present study that the EGFL domain of EGFL8 contains a domain structure similar to the Delta:Serrate:LAG2 domain conserved in ligands of Notch receptors. These results also suggested that EGFL8 may be a physiologically important inhibitory ligand of Notch family proteins. Furthermore, it is of critical importance to develop therapeutic strategies to modulate Notch signaling, since it has been well recognized that alterations in Notch signaling lead to the development of various diseases, including T-cell leukemia, tumors, Alagille syndrome, Hajdu-Cheney syndrome and cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), a syndrome associated with progressive dementia, mood disorders, migraine and recurrent sub-cortical cerebral infarctions (33-37). In conclusion, the results of the present study suggested that EGFL8 may have great potential as a therapeutic target for various diseases associated with de-regulated Notch signaling.

Acknowledgements

This study was supported by the Pioneer Research Center Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (NRF-2012-000-9667).

References

- Fitch MJ, Campagnolo L, Kuhnert F and Stuhlmann H: Egf17, a novel epidermal growth factor-domain gene expressed in endothelial cells. *Dev Dyn* 230: 316-324, 2004.
- Koeppel JR, Beach MA, Baerga-Ortiz A, Kerns SJ and Komives EA: Mutations in the fourth EGF-like domain affect thrombomodulin-induced changes in the active site of thrombin. *Biochemistry* 47: 10933-10939, 2008.
- Kao YC, Jiang SJ, Pan WA, Wang KC, Chen PK, Wei HJ, Chen WS, Chang BI, Shi GY and Wu HL: The epidermal growth factor-like domain of CD93 is a potent angiogenic factor. *PLoS One* 7: e51647, 2012.
- Eroglu C, Allen NJ, Susman MW, *et al.*: Gabapentin receptor $\alpha 2\delta$ -1 is a neuronal thrombospondin receptor responsible for excitatory CNS synaptogenesis. *Cell* 139: 380-392, 2009.
- Koutsis A, Papapanagiotou A and Papavassiliou AG: Thrombomodulin: From haemostasis to inflammation and tumorigenesis. *Int J Biochem Cell Biol* 40: 1669-1673, 2008.
- Kojika S and Griffin JD: Notch receptors and hematopoiesis. *Exp Hematol* 29: 1041-1052, 2001.
- Ito D, Tanahashi N, Murata M, Sato H, Saito I, Watanabe K and Fukuuchi Y: Notch3 gene polymorphism and ischaemic cerebrovascular disease. *J Neurol Neurosurg Psychiatry* 72: 382-384, 2002.
- Campagnolo L, Leahy A, Chitnis S, Koschnick S, Fitch MJ, Fallon JT, Loskutoff D, Taubman MB and Stuhlmann H: EGFL7 is a chemoattractant for endothelial cells and is up-regulated in angiogenesis and arterial injury. *Am J Pathol* 167: 275-284, 2005.
- Nichol D, Shawber C, Fitch MJ, Bambino K, Sharma A, Kitajewski J and Stuhlmann H: Impaired angiogenesis and altered Notch signaling in mice overexpressing endothelial Egf17. *Blood* 116: 6133-6143, 2010.
- Nichol D and Stuhlmann H: EGFL7: A unique angiogenic signaling factor in vascular development and disease. *Blood* 119: 1345-1352, 2012.
- Parker LH, Schmidt M, Jin SW, *et al.*: The endothelial-cell-derived secreted factor Egf17 regulates vascular tube formation. *Nature* 428: 754-758, 2004.
- Dikic I and Schmidt MH: Notch: Implications of endogenous inhibitors for therapy. *BioEssays* 32: 481-487, 2010.
- Pinte S and Soncin F: Egf17 promotes tumor escape from immunity. *Oncimmunology* 1: 375-376, 2012.
- Schmidt MH, Bicker F, Nikolic I, Meister J, Babuke T, Picuric S, Müller-Esterl W, Plate KH and Dikic I: Epidermal growth factor-like domain 7 (EGFL7) modulates Notch signalling and affects neural stem cell renewal. *Nat Cell Biol* 11: 873-880, 2009.
- Oberauer R, Rist W, Lenter MC, Hamilton BS and Neubauer H: EGFL6 is increasingly expressed in human obesity and promotes proliferation of adipose tissue-derived stromal vascular cells. *Mol Cell Biochem* 343: 257-269, 2010.
- Wang X, Gong Y, Wang D, *et al.*: Analysis of gene expression profiling in meningioma: Deregulated signaling pathways associated with meningioma and EGFL6 overexpression in benign meningioma tissue and serum. *PLoS One* 7: e52707, 2012.
- Wu F, Shirahata A, Sakuraba K, *et al.*: Down-regulation of EGFL8: A novel biomarker for advanced gastric cancer. *Anticancer Res* 31: 3377-3380, 2011.
- Wu F, Shirahata A, Sakuraba K, *et al.*: Down-regulation of EGFL8: A novel prognostic biomarker for patients with colorectal cancer. *Anticancer Res* 31: 2249-2254, 2011.
- Choi H-J, Yoon TD, Muhammad I, Jeong MH, Lee J, Baek SY, Kim BS and Yoon S: Regulatory role of mouse epidermal growth factor-like protein 8 in thymic epithelial cells. *Biochem Biophys Res Commun* 425: 250-255, 2012.
- Subhan F, Yoon TD, Choi HJ, Muhammad I, Lee J, Hong C, Oh SO, Baek SY, Kim BS and Yoon S: Epidermal growth factor-like domain 8 inhibits the survival and proliferation of mouse thymocytes. *Int J Mol Med* 32: 952-958, 2013.
- Faas SJ, Rothstein JL, Kreider BL, Rovera G and Knowles BB: Phenotypically diverse mouse thymic stromal cell lines which induce proliferation and differentiation of hematopoietic cells. *Eur J Immunol* 23: 1201-1214, 1993.
- Sonnhammer EL, Eddy SR and Durbin R: Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins* 28: 405-420, 1997.
- Appel RD, Bairoch A and Hochstrasser DF: A new generation of information retrieval tools for biologists: The example of the ExPASy WWW server. *Trends Biochem Sci* 19: 258-260, 1994.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M and Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739, 2011.
- Saitou N and Nei M: The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425, 1987.
- Thompson JD, Higgins DG and Gibson TJ: CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680, 1994.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ: Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402, 1997.
- Emanuelsson O, Brunak S, von Heijne G and Nielsen H: Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc* 2: 953-971, 2007.
- Lindsell CE, Shawber CJ, Boulter J and Weinmaster G: Jagged: A mammalian ligand that activates Notch1. *Cell* 80: 909-917, 1995.
- Downing AK, Knott V, Werner JM, Cardy CM, Campbell ID and Handford PA: Solution structure of a pair of calcium-binding epidermal growth factor-like domains: Implications for the Marfan syndrome and other genetic disorders. *Cell* 85: 597-605, 1996.
- Nagata K, Kohda D, Hatanaka H, Ichikawa S, Matsuda S, Yamamoto T, Suzuki A and Inagaki F: Solution structure of the epidermal growth factor-like domain of heregulin- α , a ligand for p180erbB-4. *EMBO J* 13: 3517-3523, 1994.
- Shah DK and Zúñiga-Pflücker JC: An overview of the intrathymic intricacies of T cell development. *J Immunol* 192: 4017-4023, 2014.
- Andersson ER and Lendahl U: Therapeutic modulation of Notch signalling - are we there yet? *Nat Rev Drug Discov* 13: 357-378, 2014.
- Lobry C, Oh P, Mansour MR, Look AT and Aifantis I: Notch signaling: Switching an oncogene to a tumor suppressor. *Blood* 123: 2451-2459, 2014.
- Hernandez Tejada FN, Galvez Silva JR and Zweidler-McKay PA: The challenge of targeting notch in hematologic malignancies. *Front Pediatr* 2: 54, 2014.
- Li Y, Ma J, Qian X, Wu Q, Xia J, Miele L, Sarkar FH and Wang Z: Regulation of EMT by Notch signaling pathway in tumor progression. *Curr Cancer Drug Targets* 13: 957-962, 2013.
- Aster JC: In brief: Notch signalling in health and disease. *J Pathol* 232: 1-3, 2014.