Comparative integromics on Eph family

MASUKO KATOH¹ and MASARU KATOH²

1M&M Medical BioInformatics, Hongo 113-0033; 2Genetics and Cell Biology Section, National Cancer Center Research Institute, Tokyo 104-0045, Japan

Received January 18, 2006; Accepted February 15, 2006

Abstract. EPHA1, EPHA2, EPHA3, EPHA4, EPHA5, EPHA6, EPHA7, EPHA8, EPHA10, EPHB1, EPHB2, EPHB3, EPHB4 and EPHB6 are Eph family receptors for Ephrin family ligands. Ephrin/EPH signaling pathway networks with the WNT signaling pathway during embryogenesis, tissue regeneration, and carcinogenesis. TCF/LEF-binding sites within the promoter region of human EPH family members were searched for by using bioinformatics and human intelligence. Because five TCF/LEF-binding sites were identified within the 5'-promoter region of the EPHA7 gene, comparative genomics analyses on EPHA7 orthologs were further performed. EPHA7-MANEA-FHL5 locus at human chromosome 6q16.1 and EPHA10-MANEAL-FHL3 locus at human chromosome 1p34.3 were paralogous regions within the human genome. Human EPHA7 mRNA was expressed in embryonic stem (ES) cells, neural tissues, duodenal cancer and parathyroid tumors, while mouse Epha7 mRNA was expressed in fertilized egg, Rathke's pouch, visual cortex, pituitary gland, other neural tissues, pancreas, lung tumors and mammary tumors. The chimpanzee EPHA7 gene and cow Epha7 gene were identified within NW_107969.1 and AC155055.2 genome sequences, respectively. Five TCF/LEF-binding sites within human EPHA7 promoter were conserved in the chimpanzee EPHA7 promoter, and three TCF/LEF-binding sites in the cow Epha7 promoter, but none in the mouse Epha7 promoter. Primates and cow EPHA7 orthologs were identified as evolutionarily conserved targets of the WNT/ß-catenin signaling pathway. D6S1056 microsatellite marker within EPHA7 gene is deleted in prostate cancer. Deletion and/or promoter CpG hypermethylation could explain the EPHA7 down-regulation in human tumors. EPHA7 is a target of systems medicine, especially in the fields of regenerative medicine and oncology.

Introduction

EPH family members are receptors for Ephrin family ligands (1-4). EPHA1, EPHA2, EPHA3, EPHA4, EPHA5, EPHA6, EPHA8 and EPHA10 are classified as EPHA subfamily members, while EPHB1, EPHB2, EPHB3, EPHB4 and EPHB6 are classified as the EPHB subfamily members. EPH family members consist of extracellular Ephrin-binding domain, cysteine-rich domain, two fibronectin type III repeats as well as cytoplasmic tyrosine kinase domain and C-terminal SAM motif.

EFNA1, EFNA2, EFNA3, EFNA4, EFNA5, EFNB1, EFNB2 and EFNB3 are Ephrin (EFN) family ligands for EPH family receptors (1-4). EFNA1, EFNA2, EFNA3, EFNA4 and EFNA5 are EFNA subfamily members characterized as GPI-anchored cell-surface proteins with EPH-binding domain, while EFNB1, EFNB2 and EFNB3 are EFNB subfamily members characterized as transmembrane proteins with an extracellular EPH-binding domain and a cytoplasmic PDZ-binding motif.

Ephrin/EPH signaling pathway networks with WNT signaling pathway in a variety of processes, such as axon guidance and gastrointestinal morphogenesis, during embryogenesis, tissue regeneration and carcinogenesis (5). Canonical WNT signaling pathway activates the transcription of target genes, such as DKK1, DKK4, FGF18 and FGF20, depending on the transcriptional complex consisting of TCF/LEF, ß-catenin, BCL9/BCL9L and PYGO1/PYGO2 (6-22).

Materials and methods

WNT target gene screening. Genome sequences corresponding to human EPHA1, EPHA2, EPHA3, EPHA4, EPHA5, EPHA6, EPHA8, EPHA10, EPHB1, EPHB2, EPHB3, EPHB4 and EPHB6 genes were searched for with BLAST programs (http://www.ncbi.nlm.nih.gov) as described previously (23-28). TCF/LEF-binding sites within the 5'-flanking promoter region of the above genes were searched for based on bioinformatics and manual inspection as described previously (29,30).
Identification of the chimpanzee and cow EPHA7 orthologs. Chimpanzee and cow genome sequences homologous to human EPHA7 were searched for with BLAST programs as described previously (31-36). TCF/LEF-binding sites within the 5'-flanking promoter region of EPHA7 orthologs were also searched for.

Comparative proteomics analysis. Phylogenetic analysis on EPH family proteins was performed by using the CLUSTALW program.

Comparative genomics analyses. Phylogenetic analysis on the promoter of EPHA7 orthologs was performed by using the CLUSTALW program. Promoter region of human, chimpanzee and cow EPHA7 orthologs were aligned by using the Genetyx program and manual curation.

In silico expression analyses. Expressed sequence tags (ESTs) derived from human EPHA7 gene and mouse Epha7 gene were searched for by using the BLAST programs. The sources of human EPHA7 ESTs and those of mouse Epha7 ESTs were listed up for in silico expression analyses.

Results

Screening of TCF/LEF-binding site within promoter region of EPH family genes. Human EPHA1 RefSeq (NM_005232.3), EPHA2 RefSeq (NM_004431.2), EPHA3 RefSeq (NM_005233.3), EPHA4 RefSeq (NM_004438.3), EPHA5 RefSeq (NM_004439.4), EPHA6 RefSeq (AK092565.1), EPHA7 RefSeq (NM_004440.2), EPHA8 RefSeq (NM_020526.3), EPHA10 coding sequence (A17821851.1), EPHB1 RefSeq (NM_004441.3), EPHB2 RefSeq (NM_017449.2), EPHB3 RefSeq (NM_004443.3), EPHB4 RefSeq (NM_004444.4) and EPHB6 RefSeq (NM_004445.2) were used as query sequences for the BLAST programs to identify genome clones corresponding to EPH family genes. The 5'-flanking promoter region of human EPHA1, EPHA2, EPHA3, EPHA4, EPHA5, EPHA6, EPHA8, EPHA10, EPHB1, EPHB2, EPHB3, EPHB4 and EPHB6 genes were identified within AC092214.3, AL451042.10, AC138973.2, AC079834.8, AC104137.5, AC109782.7, AL121966.10, AL035703.21, AC104336.2, AC016931.21, AL158086.32, AC112643.12, AF312032.1 and AC104597.3 genome sequences, respectively (Fig. 1A). TCF/LEF-binding sites within the 5'-promoter region of human EPH family genes were then searched for based on manual inspection. Five TCF/LEF-binding sites were identified within human EPHA7 promoter (Fig. 1A).

Comparative integromics analysis on EPHA7. Comparative proteomics analysis on EPHA7 was at first performed. Because human EPHA6 RefSeq encoded a C-terminally truncated isoform, mouse Epha6 was used for the phylogenetic analysis in this study. Phylogenetic analysis revealed that EPHA7 was more related to EPHA8 and EPHA10 than the other EPHA subfamily members (Fig. 1B).

Intra-species comparative genomics analysis was next performed. MDN1, CASP8AP2, CX62, BACH2 and MAP3K7 genes were located centromeric to EPHA7 gene, while MANEA, FUT9, KIAA0776 and FHL5 were located telomeric to EPHA7 gene. Paralogs corresponding to these genes around the EPHA7 locus were searched for by using the BLAST programs. MANEA-like (MANEAL) and FHL5 genes, located around the EPHA10 locus at human chromosome 1p34.3, were paralogs of MANEA and FHL3 genes, respectively. These facts indicate that EPHA7-MANEA-FHL5 locus at human chromosome 1p34.3 are paralogous regions within the human genome.
6q16.1 and EPHA10-MANEAL-FHL3 locus at human chromosome 1p34.3 were paralogous regions within the human genome (Fig. 1C).

Expression profile of human EPHA7 and mouse Epha7 mRNAs. In silico expression analyses were performed to compare the expression profile of human EPHA7 and mouse Epha7 mRNAs. Human EPHA7 mRNA was expressed in embryonic stem (ES) cells, neural tissues, duodenal cancer and parathyroid tumors, while mouse Epha7 mRNA was expressed in fertilized egg, Rathke’s pouch, visual cortex, pituitary gland, other neural tissues, pancreas, lung tumors and mammary tumors.

Identification of the chimpanzee and cow EPHA7 orthologs. BLAST programs using human EPHA7 RefSeq revealed that chimpanzee EPHA7 gene was located within NW_107969.1 genome sequence. Exon-intron boundaries of chimpanzee EPHA7 gene were determined based on the consensus sequence of exon-intron junctions. Although 3’-part of exon 8 was located within the sequence gap, chimpanzee EPHA7 gene was found consisting of 17 exons.

BLAST programs using human EPHA7 RefSeq revealed that exons 1-3 of cow Epha7 gene were located within the AC155055.2 genome sequence.

Comparative genomics analyses on EPHA7 promoters. Human EPHA7 promoter, chimpanzee EPHA7 promoter and cow Epha7 promoter were located within AL121966.10, NW_107969.1 and AC155055.2 genome sequences, respectively, as mentioned above. BLAST programs revealed that mouse and rat Epha7 promoters were located within BX000089.10 and AC106555.4 genome sequences, respectively (Fig. 2A). GC content of human, chimpanzee and cow EPHA7 promoters were 46.3%, that of mouse Epha7 promoter was 50.4%, and that of rat Epha7 promoter was 48.0%.

Five TCF/LEF-binding sites within human EPHA7 promoter were located about 1200 bp, 1150 bp, 1000 bp, 900 bp, and 550 bp upstream of the transcription start site (Fig. 2B). Five TCF/LEF-binding sites within human EPHA7 promoter have undergone nucleotide substitutions, the third to fifth TCF/LEF-binding sites within human EPHA7 promoter were completely conserved in the cow Epha7 promoter. On the other hand, the TCF/LEF-binding site was not identified within the mouse Epha7 promoter (Fig. 2B).

Phylogenetic analysis revealed that human, chimpanzee and cow EPHA7 orthologs were evolutionarily conserved targets of the WNT/β-catenin signaling pathway.

Discussion

TCF/LEF-binding sites within the promoter region of human EPHA1, EPHA2, EPHA3, EPHA4, EPHA5, EPHA6, EPHA8, EPHA10, EPHB1, EPHB2, EPHB3, EPHB4 and EPHB6 genes were searched for in this study. Because five TCF/LEF-binding sites were identified within the 5’-promoter region of EPHA7 gene at human chromosome 6q16.1 (Fig. 1A), comparative genomics analyses on EPHA7 orthologs were
Further performed, EPHA7-MANEAL-FHL5 locus at human chromosome 6q16.1 and EPHA10-MANEAL-FHL3 locus at human chromosome 1p34.3 were paralogous regions within the human genome (Fig. 1C).

The chimpanzee EPHA7 and cow Epha7 genes were identified within NW_107969.1 and AC155055.2 genome sequences, respectively (Fig. 2A). Five TCF/LEF-binding sites within human EPHA7 promoter were conserved in chimpanzee EPHA7 promoter, and three TCF/LEF-binding sites in the cow Epha7 promoter, but none in the mouse Epha7 promoter (Fig. 2B). Human EPHA7 mRNA and rodent Epha7 mRNA were expressed in early embryonic cells, neural tissues and several tumor types; however, primate EPHA7 promoters were significantly divergent from the rodent Epha7 promoters (Fig. 2C). More detailed comparison on the expression profile of human EPHA7 mRNA and rodent Epha7 mRNA should be performed in the future.

Primates and cow EPHA7 orthologs were identified as evolutionarily conserved targets of the WNT/β-catenin signaling pathway. Although WNT/β-catenin signaling pathway is activated in most cases of human colorectal cancer (37), expression of EPHA7 mRNA in human colorectal cancer was not detected by in silico expression analyses. Ephrin/EPH signaling, implicated in the maintenance of colorectal mucosal homeostasis, is down-regulated in mouse advanced colorectal cancer (5,38). Microsatellite markers D6S1942, D6S1293 and D6S1056 were identified within the human EPHA7 gene in this study. Among these microsatellite markers, D6S1056 is deleted in prostate cancer (39). EPHA7 might be down-regulated in colorectal cancer due to a deletion. Alternatively, EPHA7 mRNA might be down-regulated in extra-neural tissues due to promoter CpG hypermethylation.

Because EPHA7 is an evolutionarily conserved target of the WNT/β-catenin signaling pathway at least in primates and cow, EPHA7 is a target of systems medicine, especially in the fields of regenerative medicine and oncology.

References


