Comparative integromics on Ephrin family

YURIKO KATOH[1] and MASARU KATOH[2]

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

Received December 29, 2005; Accepted February 13, 2006

Abstract. EFNA1, EFNA2, EFNA3, EFNA4, EFNA5, EFNB1, EFNB2 and EFNB3 are EFN family ligands for EPH family receptors. EFN/EPH signaling pathway networks with the WNT signaling pathway during embryogenesis, tissue regeneration, and carcinogenesis. Comparative genomics analyses on *EFNB1, EFNB2* and *EFNB3* were performed by using bioinformatics and human intelligence (humint). *EFNB1* mRNA was expressed in human embryonic stem (ES) cells, neural tissues, diffuse type gastric cancer, pancreatic cancer, colon cancer, brain tumors and esophageal cancer, *EFNB2* mRNA in human ES cells, neural tissues and colon cancer, *EFNB3* mRNA in human ES cells, neural tissues, brain tumors, pancreatic cancer and colon cancer. Because triple TCF/LEF-binding sites were identified within the 5'-promoter region of human *EFNB3* gene, comparative genomics analyses on *EFNB3* orthologs were further performed. Chimpanzee *EFNB3* gene, consisting of five exons, was identified within AC164921.3 genome sequence. AY421228.1 was not a correct coding sequence for chimpanzee EFNB3. Chimpanzee *EFNB3* gene was found to encode a 340-amino-acid protein showing 99.4% and 96.6% total-amino-acid identity with human EFNB3 and mouse Efnb3, respectively. Three TCF/LEF-binding sites within human *EFNB3* promoter were conserved in chimpanzee *EFNB3* promoter, and the second TCF/LEF-binding site in rodent *Efnb3* promoters. CpG hypermethylation of *EFNB3* promoter with 63.2% GC content as well as deletion of *EFNB3* gene closely linked to *TP53* tumor suppressor gene at human chromosome 17p13.1 should be investigated to elucidate the mechanism of infrequent *EFNB3* upregulation in human colorectal cancer. *EFNB3*, identified as potential transcriptional target of WNT/ß-catenin signaling pathway, is a pharmacogenomics target in the fields of regenerative medicine and oncology.

Correspondence to: Dr Masaru Katoh, Genetics and Cell Biology Section, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
E-mail: mkatoh@ncc.go.jp

Key words: bioinformatics, comparative genomics, comparative proteomics, EFN, EPH, WNT, integrome network

Introduction

EFNA1, EFNA2, EFNA3, EFNA4, EFNA5, EFNB1, EFNB2 and EFNB3 are EPHRIN (EFN) family ligands for EPH family receptors (1-15). EFNA1, EFNA2, EFNA3, EFNA4 and EFNA5 are GPI-anchored cell-surface proteins with EPH-binding domain, while EFNB1, EFNB2 and EFNB3 are transmembrane proteins with extracellular EPH-binding domain and cytoplasmic PDZ-binding motif. EPH family members share common domain architecture, consisting of extracellular EFN-binding domain, cysteine-rich domain, two fibronectin type III repeats as well as cytoplasmic tyrosine kinase domain and C-terminal SAM motif. EPHA1, EPHA2, EPHA3, EPHA4, EPHA5, EPHA6, EPHA8 and EPHA10 are classified into EPHA subfamily, while EPHB1, EPHB2, EPHB3, EPHB4 and EPHB6 are classified into EPHB subfamily. EFN/EPH signaling pathway is implicated in a variety of processes, including axon guidance, angiogenesis, and gastrointestinal morphogenesis.

Canonical WNT signaling activation leads to transcriptional activation of *DKK1, DKK4, FGF18, FGF20*, etc. depending on the transcriptional complex consisting of TCF/LEF, ß-catenin, BCL9/BCL9L and PYGO1/PYGO2 (16-32). WNT/ß-catenin signaling pathway is implicated in the cell fate determination.

Mouse Efnb1 is expressed in intestinal differentiated cells, while Ephb2 and Ephb3 in intestinal proliferating cells depending on the WNT/ß-catenin signaling pathway (33,34). EFN/EPH and WNT signaling pathways network together during embryogenesis, tissue regeneration and carcinogenesis; however, direct transcriptional regulation of *EFN* family members by the WNT/ß-catenin signaling pathway remains unclear. Comparative genomics analyses on *EFNB1, EFNB2* and *EFNB3* were performed, and *EFNB3* was identified as potential target gene of the WNT/ß-catenin signaling pathway.

Materials and methods

WNT target gene screening. Genome sequences corresponding to human *EFNB1, EFNB2* and *EFNB3* genes were searched for with BLAST programs (http://www.ncbi.nlm.nih.gov) as described previously (35-39). 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 (28-32).

A

Human gene	Human chromosomal localization	Human RefSeq	Human genome clone including promoter	TCF/LEF binding site within human promoter	Mouse genome clone including promoter
EFNB1	Xq13.1	NM_004429.3	AL136092.8	0	AL671478.9
EFNB2	13q33.3	NM_004093.2	AL138689.22	0	AC090007.8
EFNB3	17p13.1	NM_001406.3	AC087388.9	3	AL731687.13

B

Exon No	Nucleotide sequences around exon – intron boundaries			Exon position within chimpanzee AC164921.3 genome sequence
1		GGTTTC-----TAAGAG	gtgagt	54062 - 54580
2	catag	GTTCCA-----TCATTG	gtactg	56810 - 57102
3	cccag	CCACAT-----GACAAA	gtgagt	57287 - 57379
4	tccag	GTCCCC-----TGCCAG	gtagga	57487 - 57591
5	gccag	GTGACC-----AACAAC		58078 - 60244

C

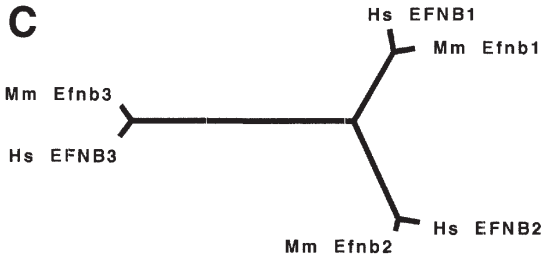

Figure 1. (A), *EFNB* gene family. Three TCF/LEF-binding sites exist within the *EFNB3* promoter. (B), Exon-intron structure of the chimpanzee *EFNB3* gene. Nucleotide sequences around exon-intron boundaries are shown by upper-case letters (exon) and by lower-case letters (intron). (C), Phylogenetic analyses on the EFNB family.

Identification of the chimpanzee EFNB3 orthologs. Chimpanzee genome sequences homologous to human *EFNB3* were searched for with BLAST programs as described previously (40-43). Exon-intron boundaries were determined based on the consensus sequence of exon-intron junctions ('gt ... ag' rule of intronic sequence) and codon usage within the coding region as described previously (44-47). Coding sequence of chimpanzee EFNB3 was determined by assembling exonic regions.

Comparative proteomics analysis. Phylogenetic analyses on mammalian EFNB family members were performed by using the CLUSTALW program.

Comparative genomics analyses. Promoter region of mammalian *EFNB3* orthologs were aligned by using the Genetyx program and manual curation. TCF/LEF-binding sites within the promoter region were determined as mentioned above.

In silico expression analysis. Expressed sequence tags (ESTs) derived from human *FENB1, EFNB2*, and *EFNB3* genes were searched for by using the BLAST programs. The sources of *FENB1, EFNB2*, and *EFNB3* ESTs were listed up for *in silico* expression analysis.

Results

Screening of TCF/LEF-binding site within promoter region of EFNB family genes. Human EFNB1 RefSeq (NM_004429.3), EFNB2 RefSeq (NM_004093.2) and EFNB3 RefSeq (NM_001406.3) were used as query sequences for the BLAST programs to identify genome clones corresponding to EFNB family genes. The 5'-flanking promoter region of human *EFNB1*, *EFNB2* and *EFNB3* genes were identified within AL136092.8, AL138689.22 and AC087388.9 genome sequences, respectively (Fig. 1A). TCF/LEF-binding sites within the 5'-promoter region of human *EFNB1, EFNB2* and *EFNB3* genes were then searched for based on manual inspection. Triple TCF/LEF-binding sites were identified within human *EFNB3* promoter (Fig. 1A).

Identification of the chimpanzee EFNB3 gene. BLAST programs using human EFNB3 RefSeq revealed that chimpanzee *EFNB3* gene was located within AC164921.3 genome sequence. Exon-intron boundaries of chimpanzee *EFNB3* gene were determined based on the consensus sequence of exon-intron junctions. Chimpanzee *EFNB3* gene was found consisting of five exons (Fig. 1B).

Because the chimpanzee AY421228.1 predicted sequence accompanied by sequence gaps within exons 2, 4 and 5 was not the correct chimpanzee EFNB3 sequence, complete coding sequence (CDS) of chimpanzee EFNB3 was determined in this study by assembling nucleotide sequences of five exons (Fig. 2).

Genetyx program revealed that nucleotide position 398-1420 was the coding region of chimpanzee EFNB3 complete CDS (Fig. 2). Chimpanzee *EFNB3* gene was found to encode a 340-amino-acid protein showing 99.4% and 96.6% total-amino-acid identity with human EFNB3 and mouse Efnb3, respectively.

Comparative proteomics analysis on mammalian EFNB3 family members. Phylogenetic analysis revealed that EFNB1 orthologs and EFNB2 orthologs were more related to each other than to EFNB3 orthologs (Fig. 1C). Extracellular EPH-binding domain as well as the C-terminal cytoplasmic region with five tyrosine residues and PZD-binding motif were well conserved among EFNB family members (Fig. 2B).

Expression of human EFNB1, EFNB2 and EFNB3 mRNAs. *In silico* expression analyses were performed to investigate the expression profile of *EFNB* family members. *EFNB1* mRNA was expressed in human embryonic stem (ES) cells, diffuse type gastric cancer, pancreatic cancer, colon cancer, brain tumors and esophageal cancer. *EFNB2* mRNA was expressed in human ES cells, neural tissues and colon cancer. *EFNB3* mRNA was expressed in human ES cells, neural tissues, brain tumors, pancreatic cancer and colon cancer.

Comparative genomics analyses on EFNB3 promoters. Human *EFNB3* promoter and chimpanzee *EFNB3* promoter were

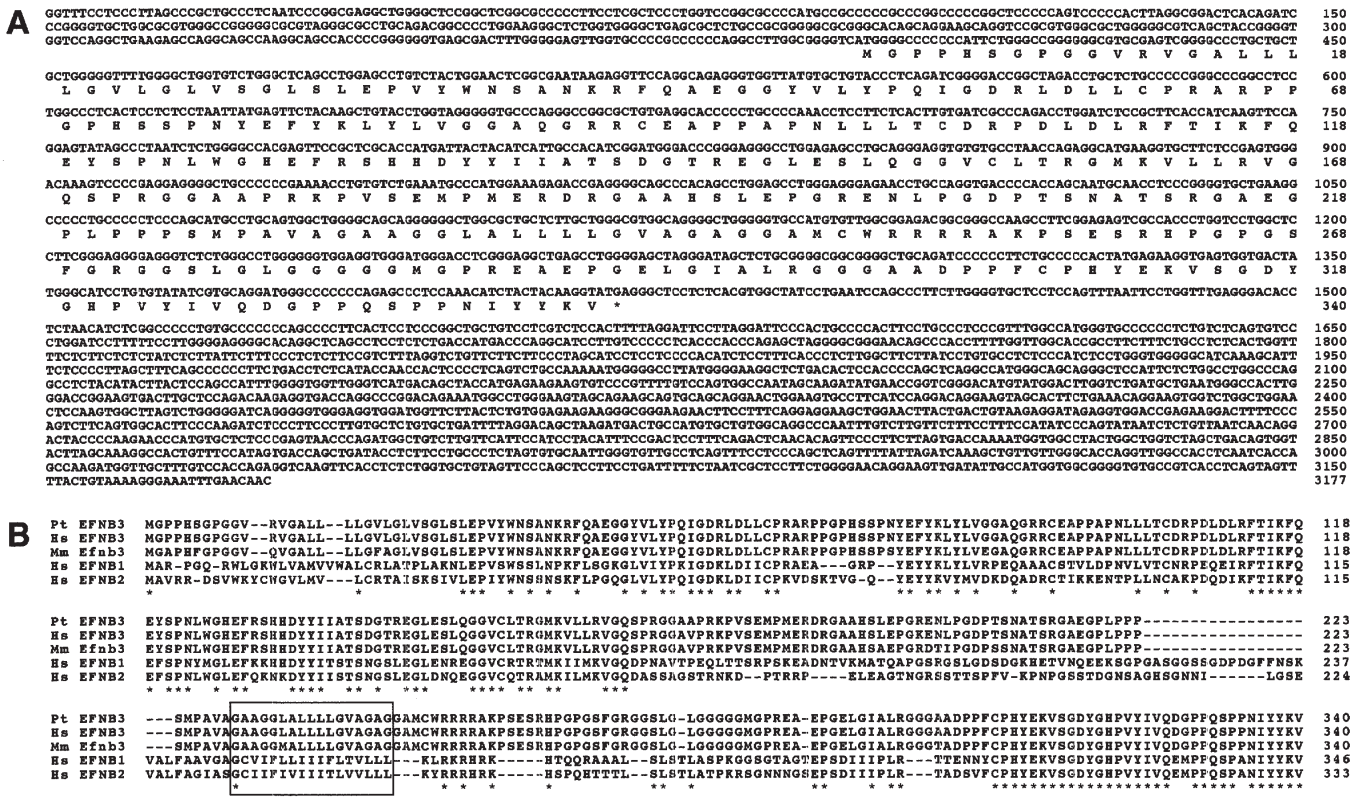

Figure 2. (A), Nucleotide and amino-acid sequences of chimpanzee EFNB3. Nucleotides and amino-acid residues are numbered on the right. (B), Alignment of EFNB family members. Pt, chimpanzee; Hs, human; Mm, mouse. Transmembrane domain is boxed. Amino-acid residues are numbered on the right. Conserved amino-acid residues are shown by asterisks.

Figure 3. Alignment of 5'-promoter region of mammalian *EFNB3* orthologs. Hs, human; Pt, chimpanzee; Mm, mouse; Rn, rat. Region corresponding to exon 1 of human *EFNB3* gene is boxed. Three TCF/LEF-binding sites conserved in primate *EFNB3* promoters are shown by double over-lines. The second TCF/LEF-binding site of primate *EFNB3* promoters is conserved in rodent *Efnb3* promoters. The first and the third TCF/LEF-binding sites of primate *EFNB3* promoters are not conserved in rodent *Efnb3* promoters due to nucleotide changes shown by a sharp.

located within AC087388.9 and AC164921.3 genome sequences, respectively, as mentioned above. BLAST programs revealed that mouse and rat *Efnb3* promoters were located within AL731687.13 and AC134317.3 genome sequences, respectively. Promoter alignment revealed that 5'-promoter region of human, chimpanzee, mouse and rat *EFNB3* orthologs were well conserved (Fig. 3).

GC content of human *EFNB3* promoter was 63.2%, that of chimpanzee *EFNB3* promoter was 63.1%, that of mouse *Efnb3* promoter was 55.5%, and that of rat *Efnb3* promoter was 56.3%. GC contents of primate *EFNB3* promoters were higher than the rodent *Efnb3* promoters.

Triple TCF/LEF-binding sites within human *EFNB3* promoters were located about 1100, 1000, and 400 bp upstream of the transcription start site (Fig. 3). Three TCF/LEF-binding sites within human *EFNB3* promoter were conserved in the chimpanzee *EFNB3* promoter. The second TCF/LEF-binding site within human *EFNB3* promoter was conserved in rodents *Efnb3* promoters.

Discussion

TCF/LEF-binding sites within *EFNB1, EFNB2, EFNB3* promoters were searched for to identify the WNT/ß-catenin target gene among the *EFNB* family in this study. The 5'-flanking promoter region of human *EFNB1, EFNB2* and *EFNB3* genes were identified within AL136092.8, AL138689.22 and AC087388.9 genome sequences, respectively. Because triple TCF/LEF-binding sites were identified within the 5'-promoter region of human *EFNB3* gene (Fig. 1A), comparative genomics analyses on *EFNB3* orthologs were further performed.

Chimpanzee *EFNB3* gene, consisting of five exons, was identified within the AC164921.3 genome sequence (Fig. 1B). AY421228.1 was not the correct coding sequence for chimpanzee EFNB3, and complete CDS of chimpanzee EFNB3 was determined in this study (Fig. 2). Chimpanzee *EFNB3* gene was found to encode a 340-amino-acid protein showing 99.4% and 96.6% total-amino-acid identity with human EFNB3 and mouse Efnb3, respectively.

Three TCF/LEF-binding sites within human *EFNB3* promoter were conserved in chimpanzee *EFNB3* promoter, while the only second TCF/LEF-binding site within human *EFNB3* promoter was conserved in rodent *Efnb3* promoters (Fig. 3). GC contents of primate *EFNB3* promoters were higher than the rodent *Efnb3* promoters. Although mammalian *EFNB3* promoters were relatively well conserved, TCF/LEF-binding sites were triplicated in primate *EFNB3* promoters compared with the rodent *Efnb3* promoters due to nucleotide changes during mammalian evolution.

Expression of human *EFNB1, EFNB2* and *EFNB3* mRNAs was investigated by using the *in silico* expression analyses. *EFNB1, EFNB2* and *EFNB3* mRNAs were expressed in human ES cells and neural tissues. In addition, *EFNB1* mRNA was expressed in a variety of tumors, such as gastric cancer, pancreatic cancer, colon cancer, brain tumors and esophageal cancer.

Mouse Efnb3, interacting with Epha4 on the axons, is implicated in axon repulsion during embryogenesis as well as inhibition of neurite outgrowth after traumatic spinal cord injury (48-50). Repression of *EFNB3* transcription could contribute to the acceleration of neurite outgrowth after traumatic spinal cord injury.

EFNB3 expression in colorectal cancer was relatively infrequent, although WNT/ß-catenin signaling pathway is frequently activated in colorectal cancer. GC content of *EFNB3* promoter was 63.2%, and *EFNB3* gene is closely linked to *TP53* tumor suppressor gene at human chromosome 17p13.1. CpG hypermethylation of *EFNB3* promoter as well as deletion of *EFNB3* gene might explain the relatively infrequent expression of *EFNB3* mRNA in colorectal cancer. Epigenetic changes and genetic alterations of *EFNB3* gene in colorectal cancer should be investigated in the future.

EFNB3 was identified as potential transcriptional target of WNT/ß-catenin signaling pathway in this study. *EFNB3* is a pharmacogenomics target in the fields of regenerative medicine and oncology.

References

1. Davis S, Gale NW, Aldrich TH, *et al*: Ligands for EPH-related receptor tyrosine kinases that require membrane attachment or clustering for activity. Science 266: 816-819, 1994.
2. Beckmann MP, Cerretti DP, Baum P, *et al*: Molecular characterization of a family of ligands for eph-related tyrosine kinase receptors. EMBO J 13: 3757-3762, 1994.
3. Fletcher FA, Huebner K, Shaffer LG, *et al*: Assignment of the gene (EPLG2) encoding a high-affinity binding protein for the receptor tyrosine kinase elk to a 200-kilobasepair region in human chromosome Xq12. Genomics 25: 334-335, 1995.
4. Bennett BD, Zeigler FC, Gu Q, *et al*: Molecular cloning of a ligand for the EPH-related receptor protein-tyrosine kinase Htk. Proc Natl Acad Sci USA 92: 1866-1870, 1995.
5. Pandey A, Lindberg RA and Dixit VM: Cell signalling. Receptor orphans find a family. Curr Biol 5: 986-989, 1995.
6. Cerretti DP, Lyman SD, Kozlosky CJ, *et al*: The genes encoding the eph-related receptor tyrosine kinase ligands LERK-1 (EPLG1, Epl1), LERK-3 (EPLG3, Epl3), and LERK-4 (EPLG4, Epl4) are clustered on human chromosome 1 and mouse chromosome 3. Genomics 33: 277-282, 1996.
7. Cerretti DP, Copeland NG, Gilbert DJ, *et al*: The gene encoding LERK-7 (EPLG7, Epl7), a ligand for the Eph-related receptor tyrosine kinases, maps to human chromosome 5 at band q21 and to mouse chromosome 17. Genomics 35: 376-379, 1996.
8. Gale NW, Flenniken A, Compton DC, *et al*: Elk-L3, a novel transmembrane ligand for the Eph family of receptor tyrosine kinases, expressed in embryonic floor plate, roof plate and hindbrain segments. Oncogene 13: 1343-1352, 1996.
9. Gale NW, Holland SJ, Valenzuela DM, *et al*: Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis. Neuron 17: 9-19, 1996.
10. Aasheim HC, Pedeutour F, Grosgeorge J and Logtenberg T: Cloning, chromosal mapping, and tissue expression of the gene encoding the human Eph-family kinase ligand ephrin-A2. Biochem Biophys Res Commun 252: 378-382, 1998.
11. Flanagan JG and Vanderhaeghen P: The ephrins and Eph receptors in neural development. Annu Rev Neurosci 21: 309-345, 1998.
12. Holder N and Klein R: Eph receptors and ephrins: effectors of morphogenesis. Development 126: 2033-2044, 1999.
13. Frisen J, Holmberg J and Barbacid M: Ephrins and their multitalented directors of embryonic development. EMBO J 18: 5159-5165, 1999.
14. Wilkinson DG: Multiple roles of Eph receptors and ephrinin neural development. Nat Rev Neurosci 2: 155-164, 2001.
15. Mancia F and Shapiro L: ADAM and Eph: how Ephrin-signaling cells become detached. Cell 123: 185-187, 2005.
16. Katoh M: *WNT* and *FGF* gene clusters. Int J Oncol 21: 1269-1273, 2002.
17. Katoh M: Regulation of WNT signaling molecules by retinoic acid during neuronal differentiation in NT2 cells: threshold model of WNT action. Int J Mol Med 10: 683-687, 2002.
18. Heller RS, Klein T, Ling Z, Heimberg H, Katoh M, Madsen OD and Serup P: Expression of *WNT*, *Frizzled, sFRP*, and *DKK* genes in adult human pancreas. Gene Expr 11: 141-147, 2003.
19. Katoh M: WNT2 and human gastrointestinal cancer. Int J Mol Med 12: 811-816, 2003.

20. Katoh M and Katoh M: Identification and characterization of human *BCL9L* gene and mouse *Bcl9l* gene *in silico*. Int J Mol Med 12: 643-649, 2003.
21. Garciadiego-Cazares D, Rosales C, Katoh M and Chimal-Monroy J: Coordination of chondrocyte differentiation and joint formation by α5ß1 integrin in the developing appendicular skeleton. Development 131: 4735-4742, 2004.
22. Katoh M and Katoh M: Comparative genomics on *WNT8A* and *WNT8B* genes. Int J Oncol 26: 1129-1133, 2005.
23. Katoh M: Molecular evolution of *WNT2B* orthologs. Int J Oncol 26: 1135-1139, 2005.
24. Katoh M: Comparative genomics on *WNT3-WNT9B* gene cluster. Int J Mol Med 15: 743-747, 2005.
25. Swain RK, Katoh M, Medina A and Steinbeisser H: *Xenopus* frizzled-4S, a splicing variant of Xfz4, is a context-dependent activator and inhibitor of Wnt/ß-catenin signaling. Cell Commun Signal 3: 12, 2005.
26. Katoh M: Epithelial-mesenchymal transition in gastric cancer. Int J Oncol 27: 1677-1683, 2005.
27. Katoh M: WNT2B: comparative integromics and clinical application. Int J Mol Med 16: 1103-1108, 2005.
28. Katoh Y and Katoh M: Comparative genomics on *DKK1* orthologs. Int J Oncol 27: 275-279, 2005.
29. Katoh Y and Katoh M: Comparative genomics on *DKK2* and *DKK4* orthologs. Int J Mol Med 16: 477-481, 2005.
30. Katoh M and Katoh M: Comparative genomics on *FGF20* orthologs. Oncol Rep 14: 287-290, 2005.
31. Katoh M and Katoh M: Comparative genomics on *FGF8, FGF17,* and *FGF18* orthologs. Int J Mol Med 16: 493-496, 2005.
32. Katoh Y and Katoh M: Comparative genomics on *FGF16* orthologs. Int J Mol Med 16: 959-963, 2005.
33. Battle E, Henderson JT, Beghtel H, *et al*: ß-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/EphrinB. Cell 111: 251-263, 2002.
34. Batlle E, Bacani J, Begthel H, *et al*: EphB receptor activity suppresses colorectal cancer progression. Nature 435: 1126-1130, 2005.
35. Katoh M: Paradigm shift in gene-finding method: from bench-top approach to desk-top approach. Int J Mol Med 10: 677-682, 2002.
36. Katoh M and Katoh M: Evolutionary conservation of *CCND1-ORAOV1-FGF19-FGF4* locus from zebrafish to human. Int J Mol Med 12: 45-50, 2003.
37. Katoh M and Katoh M: *CLDN23* gene, frequently down-regulated in intestinal-type gastric cancer, is a novel member of *CLAUDIN* gene family. Int J Mol Med 11: 683-689, 2003.
38. Katoh M and Katoh M: Identification and characterization of human *MPP7* gene and mouse *Mpp7* gene *in silico*. Int J Mol Med 13: 333-338, 2004.
39. Katoh M and Katoh M: Identification and characterization of *Crumbs* homolog 2 gene at human chromosome 9q33.3. Int J Oncol 24: 743-749, 2004.
40. Katoh M and Katoh M: Identification and characterization of human *HES2, HES3,* and *HES5* genes *in silico*. Int J Oncol 25: 529-534, 2004.
41. Katoh M and Katoh M: Identification and characterization of human *HESL*, rat *Hesl* and rainbow trout *hesl* genes *in silico*. Int J Mol Med 14: 747-751, 2004.
42. Katoh M and Katoh M: Comparative genomics on *WNT5A* and *WNT5B* genes. Int J Mol Med 15: 749-753, 2005.
43. Katoh Y and Katoh M: Comparative genomics on *WNT11* gene. Int J Mol Med 15: 879-883, 2005.
44. Katoh Y and Katoh M: Comparative genomics on *VANGL1* and *VANGL2* genes. Int J Oncol 26: 1435-1440, 2005.
45. Katoh Y and Katoh M: Comparative genomics on *SFRP1* orthologs. Int J Oncol 27: 861-865, 2005.
46. Katoh Y and Katoh M: WNT antagonist, SFRP1, is Hedgehog signaling target. Int J Mol Med 17: 171-175, 2006.
47. Katoh Y and Katoh M: Comparative genomics on HHIP family orthologs. Int J Mol Med 17: 391-395, 2006.
48. Kullander K, Croll SD, Zimmer M, *et al*: Ephrin-B3 is the midline barrier that prevents corticospinal tract axons from recrossing, allowing for unilateral motor control. Genes Dev 15: 877-888, 2001.
49. Kullander K, Butt SJ, Lebret JM, *et al*: Role of EphA4 and EphrinB3 in local neuronal circuits that control walking. Science 299: 1889-1892, 2003.
50. Benson MD, Romero MI, Lush ME, *et al*: Ephrin-B3 is a myelin-based inhibitor of neurite outgrowth. Proc Natl Acad Sci USA 102: 10694-10699, 2005.