Comparative integromics on VEGF family members

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 February 6, 2006; Accepted March 7, 2006

Abstract. VEGF, Hedgehog, FGF, Notch, and WNT signaling pathways network together for vascular remodeling during embryogenesis, tissue regeneration, and carcinogenesis. VEGFA (VEGF), VEGFB, VEGFC, VEGFD (FIGF) and PGF (PlGF) are VEGF family ligands for receptor tyrosine kinases, including VEGFR1 (FLT1), VEGFR2 (KDR) and VEGFR3 (FLT4). Bevacizumab (Avastin), Sunitinib (Sutent) and Sorafenib (Nexavar) are anti-cancer drugs targeted to VEGF signaling pathway. TCF/LEF binding sites within the promoter region of human *VEGF* family members were searched for by using bioinformatics and human intelligence (Humint). Because four TCF/LEF-binding sites were identified within the 5'-promoter region of human *VEGFD* gene within AC095351.5 genome sequence, comparative genomics analyses on *VEGFD* orthologs were further performed. *ASB9-ASB11-VEGFD* locus at human chromosome Xp22.2 and *ASB5-VEGFC* locus at human chromosome 4q34 were paralogous regions within the human genome. Human *VEGFD* mRNA was expressed in lung, small intestine, uterus, breast, neural tissues, and neuroblastoma. Mouse *Vegfd* mRNA was expressed in kidney, pregnant oviduct, and neural tissues. Chimpanzee *VEGFD* promoter, cow *Vegfd* promoter, mouse *Vegfd* promoter and rat *Vegfd* promoter were identified within NW_121675.1, AC161065.2, AL732475.6 and AC130036.3 genome sequences, respectively. Three out of four TCF/LEF-binding sites within human *VEGFD* promoter were conserved in chimpanzee *VEGFD* promoter, and one in cow *Vegfd* promoter. TCF/LEF-binding site, not conserved in human *VEGFD* promoter, occurred in cow, mouse and rat *Vegfd* promoters. At least five out of six bHLH-binding sites within human VEGFD proximal promoter region were conserved in chimpanzee *VEGFD* proximal promoter region, while only one in cow *Vegfd* proximal promoter region. Together these facts indicate that relatively significant promoter evolution occurred among mammalian *VEGFD* orthologs. Human *VEGFD* was characterized as a potent target gene of WNT/ß-catenin signaling pathway. VEGFD, implicated in angiogenesis and lymphatic metastasis, is a pharmacogenomics target in the field of 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, VEGF, WNT, integrome network, systems medicine

Introduction

VEGFA (VEGF), VEGFB, VEGFC, VEGFD (FIGF) and PGF (PlGF) are VEGF family proteins implicated in vascular remodeling during embryogenesis, tissue regeneration, and carcinogenesis (1-4). VEGF family members are ligands for receptor tyrosine kinases, including VEGFR1 (FLT1), VEGFR2 (KDR) and VEGFR3 (FLT4). Bevacizumab (Avastin) is a humanized anti-VEGFA monoclonal antibody, while Sunitinib (Sutent) and Sorafenib (Nexavar) are multi-kinase inhibitors for VEGFR1 and other protein kinases. Bevacizumab, Sunitinib and Sorafenib are anti-cancer drugs targeted to VEGF signaling pathway (5-8).

VEGF-induced expression of DLL4 in vascular endothelial cells leads to the activation of Notch signaling, and the following down-regulation of HHIP1 (9-11). HHIP1 down-regulation then leads to SHH activation, which results in the activation of VEGF signaling. Therefore, VEGF-induced down-regulation of HHIP1 during angiogenesis leads to the positive feedback to the Hedgehog-VEGF-Notch signaling cascade (11). VEGF, Hedgehog, FGF, Notch, and WNT signaling pathways network together for vascular remodeling during embryogenesis, tissue regeneration, and carcinogenesis (9-17).

Canonical WNT signals are transduced to the transcriptional complex consisting of TCF/LEF, ß-catenin, BCL9/BCL9L and PYGO1/PYGO2 to activate transcription of target genes, such as *DKK1, DKK4, FGF18* and *FGF20* (18-28).

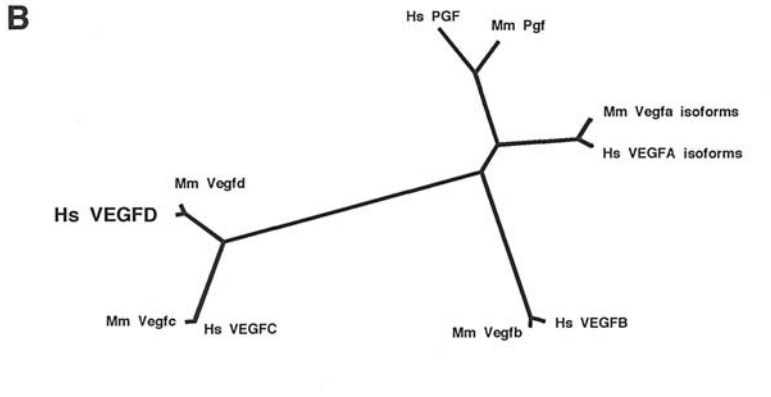
TCF/LEF binding sites within the promoter region of human *VEGF* family members were searched for by using bioinformatics and human intelligence (Humint). Because four TCF/ LEF-binding sites were identified within human *VEGFD* promoter, comparative genomics analyses on *VEGFD* orthologs were further performed.

Materials and methods

Screening of WNT target gene. Genome sequences corresponding to human *VEGFA, VEGFB, VEGFC, VEGFD* and *PGF* genes were searched for with BLAST programs

A

Human gene	Human complete CDS	Human genome clone including promoter	TCF/LEF binding site within human promoter
VEGFA	NM_001025366.1	AL136131.15	1
VEGFB	NM_003377.3	AP001453.6	0
VEGFC	NM_005429.2	AC092673.3	0
VEGFD	NM_004469.2	AC095351.5	4
PGF	NM_002632.4	AC006530.4	0



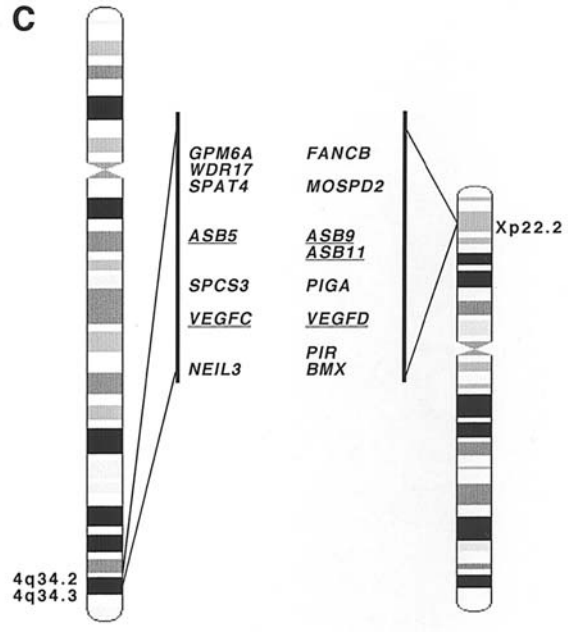


Figure 1. (A), Human *VEGF* gene family. Gene symbol, complete coding sequence, genome sequence and the number of TCF/LEF-binding sites within promoter region of five *VEGF* family genes are listed. Four TCF/LEF-binding sites exist within the human *VEGFD* promoter. (B), Phylogenetic analysis on VEGF family. (C), Intra-species comparative genomics on *VEGFD* and *VEGFC* loci. *ASB9-ASB11-VEGFD* locus at human chromosome Xp22.2 and *ASB5-VEGFC* locus at human chromosome 4q34 are paralogous regions within the human genome.

(http://www.ncbi.nlm.nih.gov) as described previously (29-32). 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 (23-27,33).

Identification of chimpanzee and cow VEGFD orthologs. Chimpanzee and cow genome sequences homologous to human *VEGFD* were searched for with BLAST programs as described previously (34-37). TCF/LEF-binding sites within the 5'-flanking promoter region of *VEGFD* orthologs were also searched for.

Comparative integromics analyses. Phylogenetic analyses on VEGF family proteins as well as on promoters of mammalian *VEGFD* orthologs were performed by using the CLUSTALW program. Human and chimpanzee *VEGFD* promoters were then aligned by using Genetyx program and manual curation as described previously (38-41).

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

Results

Screening of the TCF/LEF-binding site within promoter region of VEGF family genes. Human VEGFA RefSeq (NM_001025366.1), VEGFB RefSeq (NM_003377.3), VEGFC RefSeq (NM_005429.2), VEGFD RefSeq (NM_004469.2) and PGF RefSeq (NM_002632.4) were used as query sequences for the BLAST programs to identify genome clones corresponding to *VEGF* family genes. The 5'-flanking promoter region of human *VEGFA, VEGFB, VEGFC, VEGFD* and *PGF* genes were identified within AL136131.15, AP001453.6, AC092673.3, AC095351.5 and AC006530.4 genome sequences, respectively (Fig. 1A). TCF/LEF-binding sites within the 5'-promoter region of human *VEGF* family genes were then searched for based on manual inspection. Four TCF/LEF-binding sites were identified within human *VEGFD* promoter (Fig. 1A).

Comparative integromics analyses on VEGFD. Comparative proteomics analysis was performed at first. Phylogenetic analysis on human and mouse VEGF family members revealed that VEGFD orthologs were relatively well conserved (Fig. 1B).

Intra-species comparative genomics analysis was next performed. *FANCB, MOSPD2, ASB9, ASB11, PIGA, PIR* and *BMX* genes were located around the *VEGFD* gene, while *GPM6A, WDR17, SPAT4, ASB5, SPCS3* and *NEIL3* genes were located around the *VEGFC* gene (Fig. 1C). Paralogs corresponding to genes around the *VEGFD* locus were searched for by using the BLAST programs. *ASB9* and *ASB11* genes, encoding ankyrin repeat and SOCS box-containing proteins, were the paralog of *ASB5* gene. These facts indicate that *ASB9-ASB11-VEGFD* locus at human chromosome Xp22.2 and *ASB5-VEGFC* locus at human chromosome 4q34 were paralogous regions within the human genome (Fig. 1C).

Expression profile of human VEGFD and mouse Vegfd mRNAs. *In silico* expression analyses were performed to compare the expression profile of human *VEGFD* and mouse *Vegfd* mRNAs. Human *VEGFD* mRNA was expressed in lung, small intestine, uterus, breast, neural tissues, and neuroblastoma. Mouse *Vegfd* mRNA was expressed in kidney, pregnant oviduct, and neural tissues.

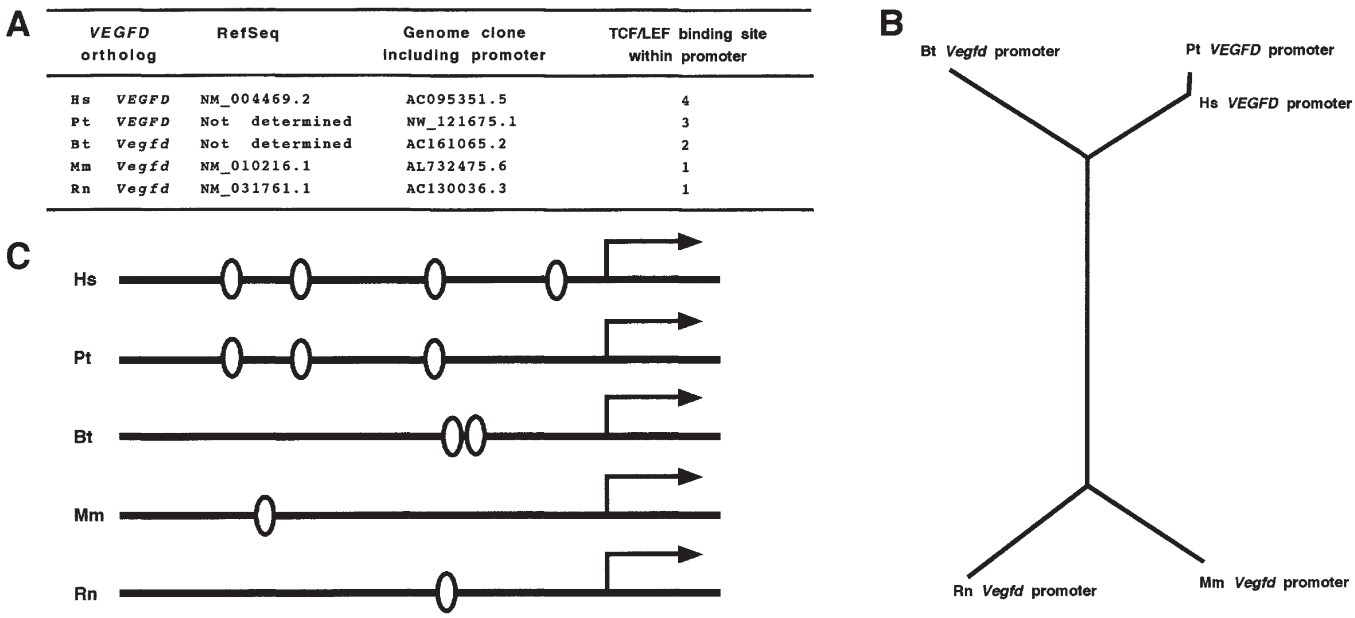


Figure 2. Inter-species comparative genomics analyses on *VEGFD* promoters. Hs, human; Pt, chimpanzee; Bt, cow; Mm, mouse; Rn, rat. (A), Promoter of mammalian *VEGFD* orthologs. (B), Phylogenetic analysis on mammalian *VEGFD* promoters. (C), Schematic representation of mammalian *VEGFD* promoters. TCF/LEF-binding sites are shown by an oval symbol.

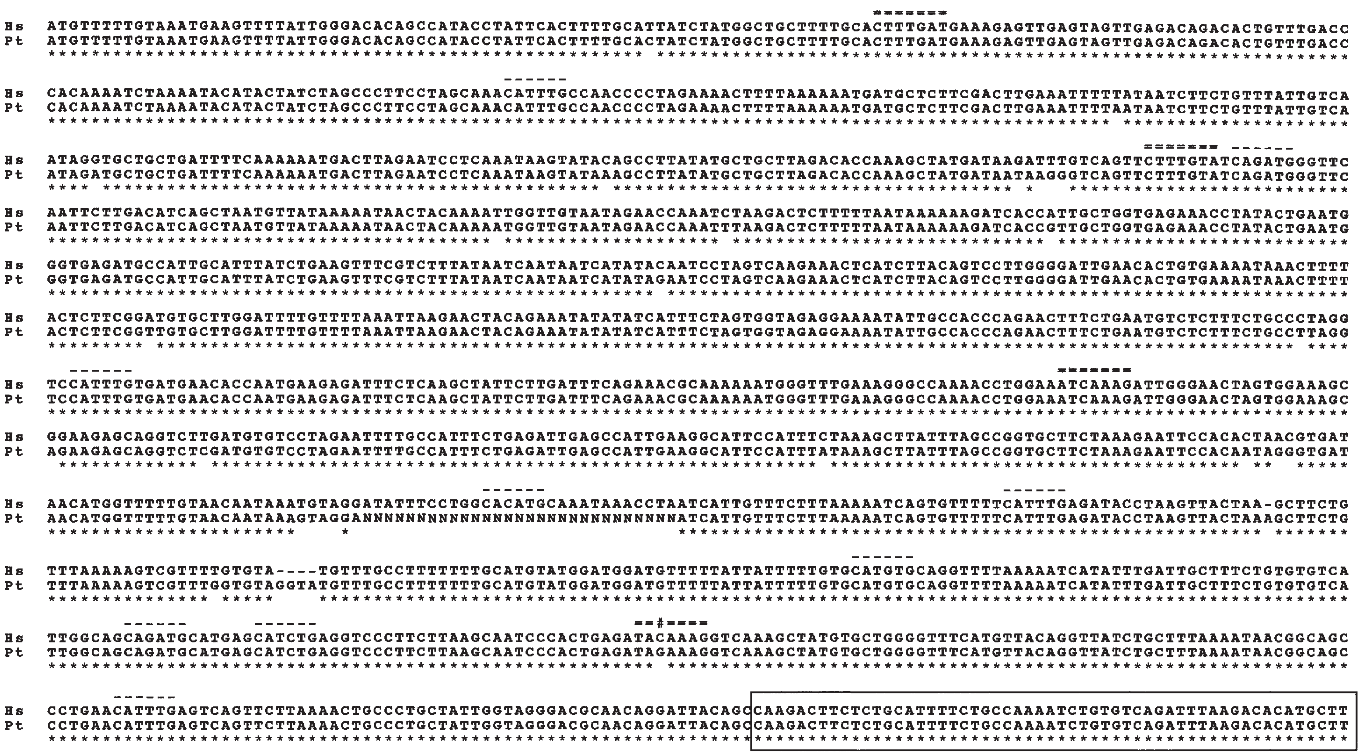


Figure 3. Alignment of the human and chimpanzee *VEGFD* promoters. Hs, human; Pt, chimpanzee. Region corresponding to exon 1 of human *VEGFD* gene is boxed. TCF/LEF-binding sites (double over-lines) as well as bHLH-binding sites are shown. Nucleotide changes to disrupt the most proximal TCF/LEF-binding site in chimpanzee *Vegfd* promoter are shown by a sharp.

Identification of the chimpanzee VEGFD ortholog. BLAST programs using human VEGFD RefSeq revealed that chimpanzee *VEGFD* gene was located within NW_121675.1 genome sequence (Fig. 2A). Exon-intron boundaries of chimpanzee *VEGFD* gene were determined based on the consensus sequence of exon-intron junctions. Chimpanzee *VEGFD* gene was found consisting of seven exons. Compared with human VEGFD RefSeq, one-base insertion occurred at two positions within the N-terminal coding region of chimpanzee *VEGFD* gene. Re-sequencing of the genome

sequence around exon 1 of chimpanzee *VEGFD* gene should be done in the future to evaluate whether one-base insertions are real insertions or sequencing errors.

Comparative genomics analyses on VEGFD promoters. Human and chimpanzee *VEGFD* promoters were located within AC095351.5 and NW_121675.1 genome sequences, respectively, as mentioned above. BLAST programs next revealed that cow, mouse and rat *Vegfd* promoters were located within AC161065.2, AL732475.6 and AC130036.3 genome sequences, respectively (Fig. 2A). GC content of human, chimpanzee, cow, mouse and rat *VEGFD* promoters were 36.3%, 35.5%, 37.4%, 37.3% and 38.2%, respectively. Phylogenetic analysis revealed that human, chimpanzee and cow *VEGFD* promoters were significantly divergent from mouse and rat *Vegfd* promoters (Fig. 2B).

Four TCF/LEF-binding sites within human *VEGFD* promoter were located about 1300, 1000, 600, and 100 bp upstream of the transcription start site (Fig. 2C). Three out of four TCF/LEF-binding sites within human *VEGFD* promoter were conserved in chimpanzee *VEGFD* promoter, and one in cow *Vegfd* promoter. TCF/LEF-binding site, not conserved in the human *VEGFD* promoter, occurred in cow, mouse and rat *Vegfd* promoters (Fig. 2C). Although human and chimpanzee *VEGFD* promoters were well conserved, most proximal TCF/LEF-binding site within human *VEGFD* promoter was disrupted in chimpanzee *VEGFD* promoter due to a single nucleotide substitution (Fig. 3).

Six bHLH-binding sites were clustered within the proximal region of human *VEGFD* promoter (Fig. 3). At least five out of six bHLH-binding sites were conserved in chimpanzee *VEGFD* proximal promoter region, while only one in cow *Vegfd* proximal promoter region.

Discussion

TCF/LEF binding sites within the promoter region of human *VEGFA, VEGFB, VEGFC, VEGFD* and *PGF* genes were searched for in this study. Because four TCF/LEF-binding sites were identified within the 5'-promoter region of human *VEGFD* gene (Fig. 1A), comparative genomics analyses on *VEGFD* orthologs were further performed. *ASB9-ASB11-PIGA-VEGFD* locus at human chromosome Xp22.2 and *ASB5-SPCS3-VEGFC* locus at human chromosome 4q34 were paralogous regions within the human genome (Fig. 1C).

Chimpanzee *VEGFD* promoter, cow *Vegfd* promoter, mouse *Vegfd* promoter and rat *Vegfd* promoter were identified within NW_121675.1, AC161065.2, AL732475.6 and AC130036.3 genome sequences, respectively (Fig. 2A). Three out of four TCF/LEF-binding sites within human *VEGFD* promoter were conserved in chimpanzee *VEGFD* promoter, and one in cow *Vegfd* promoter. TCF/LEF-binding site, not conserved in the human *VEGFD* promoter, occurred in cow, mouse and rat *Vegfd* promoters (Fig. 2C). At least five out of six bHLH-binding sites within human VEGFD proximal promoter region were conserved in chimpanzee *VEGFD* proximal promoter region, while only one in cow *Vegfd* proximal promoter region. Together these facts indicate that relatively significant promoter evolution occurred among mammalian *VEGFD* orthologs.

Human *VEGFD* was identified as the target gene of WNT/ß-catenin signaling pathway in this study. Human *VEGFD* mRNA was expressed in lung, small intestine, uterus, breast, neural tissues, and neuroblastoma. VEGFD, implicated in angiogenesis and lymphatic metastasis, is a pharmacogenomics target in the field of oncology.

References

1. Shibuya M: Structure and function of VEGF/VEGF-receptor system involved in angiogenesis. Cell Struct Funct 26: 25-35, 2001.
2. Tammela T, Enholm B, Alitalo K and Paavonen K: The biology of vascular endothelial growth factors. Cardiovasc Res 65: 550-563, 2005.
3. Coultas L, Chawengsaksophak K and Rossant J: Endothelial cells and VEGF in vascular development. Nature 438: 937-945, 2005.
4. Alitalo K, Tammela T and Petrova TV: Lymphangiogenesis in development and human disease. Nature 438: 946-953, 2005.
5. Chen HX, Gore-Langton RE and Cheson BD: Clinical trials referral resource: current clinical trials of the anti-VEGF monoclonal antibody bevacizumab. Oncology 15: 1023-1026, 2001.
6. O'Farrell AM, Abrams TJ, Yuen HA, *et al*: SU11248 is a novel FLT3 tyrosine kinase inhibitor with potent activity *in vitro* and *in vivo*. Blood 101: 3597-3605, 2003.
7. Hotte SJ and Hirte HW: BAY 43-9006: early clinical data in patients with advanced solid malignancies. Curr Pharm Des 8: 2249-2253, 2002.
8. Ferrara N and Kerbel RS: Angiogenesis as a therapeutic target. Nature 438: 967-974, 2005.
9. Patel NS, Li JL, Generali D, *et al*: Up-regulation of Delta-like 4 ligand in human tumor vasculature and the role of basal expression in endothelial cell function. Cancer Res 65: 8690-8697, 2005.
10. Li JL and Harris AL: Notch signaling from tumor cells: a new mechanism of angiogenesis. Cancer Cell 8: 1-3, 2005.
11. Katoh Y and Katoh M: Comparative genomics on HHIP family orthologs. Int J Mol Med 17: 391-395, 2006.
12. Lawson ND, Vogel AM and Weinstein BM: Sonic hedgehog and vascular endothelium growth factor act upstream of Notch pathway during arterial endothelial differentiation. Dev Cell 3: 127-136, 2002.
13. Katoh Y and Katoh M: Hedgehog signaling in gastric cancer. Cancer Biol Ther 4: 1050-1054, 2005.
14. 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.
15. Katoh Y and Katoh M: WNT antagonist, SFRP1, is Hedgehog signaling target. Int J Mol Med 17: 171-175, 2006.
16. Katoh M: Epithelial-mesenchymal transition in gastric cancer. Int J Oncol 27: 1677-1683, 2005.
17. Katoh Y and Katoh M: FGF signaling inhibitor, SPRY4, is evolutionarily conserved target of WNT signaling pathway in progenitor cells. Int J Mol Med 17: 529-532, 2006.
18. Katoh M: *WNT* and *FGF* gene clusters. Int J Oncol 21: 1269-1273, 2002.
19. 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.
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. 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.
22. 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.
23. Katoh Y and Katoh M: Comparative genomics on *DKK1* orthologs. Int J Oncol 27: 275-279, 2005.
24. Katoh Y and Katoh M: Comparative genomics on *DKK2* and *DKK4* orthologs. Int J Mol Med 16: 477-481, 2005.
25. Katoh Y and Katoh M: Comparative genomics on *FGF16* orthologs. Int J Mol Med 16: 959-963, 2005.

26. Katoh M and Katoh M: Comparative genomics on *FGF8, FGF17,* and *FGF18* orthologs. Int J Mol Med 16: 493-496, 2005.
27. Katoh M and Katoh M: Comparative genomics on *FGF20* orthologs. Oncol Rep 14: 287-290, 2005.
28. Katoh M: WNT2B: comparative integromics and clinical application. Int J Mol Med 16: 1103-1108, 2005.
29. Katoh M: Paradigm shift in gene-finding method: from bench-top approach to desk-top approach. Int J Mol Med 10: 677-682, 2002.
30. Katoh M and Katoh M: Identification and characterization of human *HES2, HES3,* and *HES5* genes *in silico*. Int J Oncol 25: 529-534, 2004.
31. 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, 2005.
32. Katoh Y and Katoh M: Identification and characterization of rat *Wnt6* and *Wnt10a* genes *in silico*. Int J Mol Med 15: 527-531, 2005.
33. Katoh Y and Katoh M: Comparative genomics on *SLIT1, SLIT2,* and *SLIT3* orthologs. Oncol Rep 14: 1351-1355, 2005.
34. Katoh Y and Katoh M: Identification and characterization of rat *Wnt1* and *Wnt10b* genes *in silico*. Int J Oncol 26: 841-845, 2005.
35. Katoh M and Katoh M: Comparative genomics on *WNT8A* and *WNT8B* genes. Int J Oncol 26: 1129-1133, 2005.
36. Katoh M: Molecular evolution of *WNT2B* orthologs. Int J Oncol 26: 1135-1139, 2005.
37. Katoh M: Comparative genomics on *WNT3-WNT9B* gene cluster. Int J Mol Med 15: 743-747, 2005.
38. Katoh M and Katoh M: Comparative genomics on *WNT5A* and *WNT5B* genes. Int J Mol Med 15: 749-753, 2005.
39. Katoh Y and Katoh M: Comparative genomics on *WNT11* gene. Int J Mol Med 15: 879-883, 2005.
40. Katoh Y and Katoh M: Comparative genomics on *VANGL1* and *VANGL2* genes. Int J Oncol 26: 1435-1440, 2005.
41. Katoh Y and Katoh M: Comparative genomics on *SFRP1* orthologs. Int J Oncol 27: 861-865, 2005.