

Comparative genomics on HHIP family orthologs

YURIKO 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 October 19, 2005; Accepted November 18, 2005

Abstract. Hedgehog, FGF, VEGF, and Notch signaling pathways network together for vascular remodeling during embryogenesis and carcinogenesis. HHIP1 (HHIP) is an endogenous antagonist for SHH, IHH, and DHH. Here, comparative integromics analyses on HHIP family members were performed by using bioinformatics and human intelligence. *HHIP1*, *HHIP2* (*HHIPL1* or *KIAA1822*) and *HHIP3* (*HHIPL2* or *KIAA1822L*) constitute human *HHIP* gene family. Rat *Hhip1*, *Hhip2*, and *Hhip3* genes were identified within AC107504.4, AC094820.6, and AC134264.2 genome sequences, respectively. HHIP-homologous (HIPH) domain with conserved 18 Cys residues was identified as the novel domain conserved among mammalian HHIP1, HHIP2, and HHIP3 orthologs. *HHIP1* mRNA was expressed in coronary artery endothelial cells, prostate, and rhabdomyosarcoma. *HHIP2* mRNA was expressed in trabecular bone cells. *HHIP3* mRNA was expressed in testis, thyroid gland, osteoarthritic cartilage, pancreatic cancer, and lung cancer. Promoters of *HHIP* family genes were not well conserved between human and rodents. Although GLI-, CSL-, and HES/HEY-binding sites were not identified, eleven bHLH-binding sites were identified within human *HHIP1* promoter. Expression of *HES/HEY* family members, including *HES1*, *HES2*, *HES3*, *HES4*, *HES5*, *HES6*, *HES7*, *HEY1*, *HEY2* and *HEYL*, in coronary artery endothelial cells was not detected *in silico*. Up-regulation of HHIP1 due to down-regulation of Notch-CSL-HES/HEY signaling cascade repressing bHLH transcription factors results in down-regulation of the Hedgehog-VEGF-Notch signaling cascade. On the other hand, down-regulation of HHIP1 due to up-regulation of Notch signaling in vascular endothelial cells during angiogenesis results in up-regulation of the Hedgehog-VEGF-Notch signaling cascade. Because HHIP1 is the key molecule for vascular remodeling, HHIP1 is the

pharmacogenomics target in the fields of oncology and vascular medicine.

Introduction

Hedgehog signaling pathway is implicated in a variety of processes during embryogenesis, chronic persistent inflammation, and carcinogenesis (1-3). Hedgehog family of secreted proteins consists of Sonic Hedgehog (SHH), Indian Hedgehog (IHH) and Desert Hedgehog (DHH) (4-6). PTCH1, PTCH2, DISP1, DISP2 and DISP3 are multi-transmembrane proteins with two PTCH/DISP homologous domains (7,8). PTCH1 and PTCH2 are Hedgehog receptors, regulating the Hedgehog signal transducer Smoothened (SMO) (9-13). KIF27, KIF7, STK36, SUFU and DZIP1 are Hedgehog signaling components (10-17). GLI family transcription factors are implicated in the transcriptional activation of Hedgehog target genes, such as *PTCH1*, *CCND2*, *IGFBP6*, and *FOXM1* (1,18-20).

HHIP1 (HHIP) is secreted-type Hedgehog-interacting protein, functioning as an endogenous antagonist for SHH, IHH, and DHH (21). *HHIP1* expression is down-regulated in a variety of tumors, such as gastric, pancreatic, colorectal, esophageal and lung cancer (22,23). *HHIP1* down-regulation in pancreatic cancer is due to epigenetic CpG hypermethylation of *HHIP1* promoter.

In contrast to the down-regulation of *HHIP1* in various types of human tumors, *HHIP1* is abundantly expressed in human aortic endothelial cells (22). However, mechanism of *HHIP1* expression in human aortic endothelial cells remains unclear.

Recently, we identified two other members of human *HHIP* gene family. Here, comparative genomics analyses on *HHIP* family members as well as expression analyses of *HHIP* family members were performed. Because *HHIP1* was expressed in coronary artery endothelial cells, transcriptional mechanism of *HHIP1* in coronary artery endothelial cells was further investigated.

Materials and methods

Identification of novel genes. Mouse cDNAs, ESTs, and rat genome sequences homologous to human HHIP1, HHIP2, and HHIP3 were searched for with the BLAST program (<http://www.ncbi.nlm.nih.gov>) as described previously (24,25). Exon-intron boundaries were determined by examining the consensus sequence of exon-intron junctions ('gt ag' rule of intronic

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, Hedgehog, VEGF, Notch, integrome network

A	Human				Mouse	Rat
	Gene	Alias	Chromosomal position	Complete CDS	Complete CDS	Complete CDS
	<i>HHIP1</i>	<i>HHIP</i>	4q31.21	NM_022475.1	NM_020259.3	This study
	<i>HHIP2</i>	<i>KIAA1822</i>	14q32.2	NM_032425.3	AK166269.1	This study
	<i>HHIP3</i>	<i>KIAA1822L</i>	1q41	NM_024746.2	This study	This study

B	Rat <i>Hhip1</i> gene			Rat <i>Hhip2</i> gene			Rat <i>Hhip3</i> gene		
	Exon								
	01	AGAAGC-----AACAG	gtaggc		GGCCTT-----TGCCAG	gtgagc		GGCCAA-----TGCCAG	gtgggt
	02	tttcag ATCTTT-----TTCCAG	gtaaaa	ctacag GAATGC-----AGAGAG	gtcgct		ttccag GAGTGC-----AGAGAG	gtgaga	
	03	ttaaag GTCTTC-----CAGCAG	gtccat	ctttag GATAAT-----AAACAA	gtacgc		ttccag GGTTCAT-----GAACAA	gtaagc	
	04	gtttag AAAGCA-----ATAAAG	gttgcc	ccgcag GTCAGC-----CCCTCG	gtgacc		tttcag AAGTTC-----CCTTGG	gtaagt	
	05	ctgcag GGAGGA-----ATCCAG	gtatca	cctcag ATGACG-----GAGCGG	gtaagt		tgacag ACGACA-----GAGTGG	gtaaga	
	06	atgcag GAAAAA-----GTTAAG	gtaaca	ttccag GCGACT-----AGGCTG	gtgagc		atctag TCGACT-----AAGCCG	gtaact	
	07	tctcag TGATTT-----AGGCAG	gtgagg	tcaaa GGGAGC-----CTCCAG	gtgagt		cottag GGGAGC-----CTCAAG	gtgaga	
	08	atgcag ATGTGC-----ATTACG	gtgagt	tggcag ACGGGC-----AGGAAA	gtaagt		ttacag GCGAGC-----GCGCAG	gtgagt	
	09	ttccag AAAGCG-----CAATGG	gtgtgt	gtgtag AGTTCA-----ATTTTC			ttccag CCACAG-----ATAAAT		
	10	tcatag GAATTT-----AATTAG	gtatcg						
	11	caacag GAGAGG-----AAAAAG	gtaaga						
	12	atgcag ACCTTT-----GAATTG	gttagt						
	13	ttgcag CCAAAAT-----TGAGAG							

Figure 1. (A) Mammalian *HHIP* gene family consisting of *HHIP1*, *HHIP2* and *HHIP3* orthologs. (B) Exon-intron structure of rat *Hhip1*, *Hhip2*, and *Hhip3* genes.

sequence) and the codon usage within the coding region as described previously (26,27).

Comparative proteomics analyses. Translation into amino-acid sequence and amino-acid alignment were performed with the Genetyx program as described previously (28,29). Signal peptide and transmembrane domain were searched for with the Kyte and Doolittle hydrophobicity analysis and PSORT II program. Domain architecture was at first searched for with the RPS-BLAST program (<http://www.ncbi.nlm.nih.gov>) as described previously (30,31). Novel domains were then searched for based on the conservation among related proteins as described previously (32,33).

Comparative genomics analyses. Genome sequences around human *HHIP1*, *HHIP2* and *HHIP3* genes were used as query sequences for the BLAST program to identify evolutionarily conserved regions. Transcription factor-binding sites within the promoter region were searched for with the Match program (<http://www.gene-regulation.com>) as well as by manual inspection.

Results

Human *HHIP* family genes. BLAST program using *HHIP1* amino-acid sequence NP_071920.1 as a query sequence revealed that NM_032425.3 and NM_024746.2 RefSeqs were derived from human *HHIP1*-related genes. Human genes corresponding to NM_032425.3 and NM_024746.2 were designated *HHIP2* (*HHIP1L* or *KIAA1822*) and *HHIP3* (*HHIP2L* or *KIAA1822L*), respectively (Fig. 1A).

Preliminary alignment of *HHIP* family members revealed that 529-aa NP_079022.1 was N-terminally truncated *HHIP3* protein. Although nucleotide position 644-2233 of NM_024746.2 RefSeq was translated for NP_079022.1, we found that nucleotide position 59-2233 was the real coding region. Instead of 529-aa N-terminally truncated *HHIP3* partial amino-

acid sequence, 724-aa full-length *HHIP3* amino-acid sequence translated from the real coding region of NM_024746.2 RefSeq was used for the following study.

Mouse *Hhip* family genes. Mouse cDNAs homologous to human *HHIP1*, *HHIP2*, and *HHIP3* were searched for with BLAST programs. Mouse NM_020259.3 RefSeq, AK166269 cDNA, and NM_030175.1 RefSeq were derived from mouse *Hhip1*, *Hhip2*, and *Hhip3* genes, respectively. NM_020259.3 RefSeq and AK166269 cDNA were representative full-length clones; however, NM_030175.1 RefSeq was a 5'-truncated partial clone (Fig. 1A).

BE305786 EST, corresponding to the 5'-UTR and N-terminal part of coding region of *Hhip3*, overlapped with NM_030175.1 5'-truncated partial RefSeq. Mouse *Hhip3* complete CDS was determined by assembling BE305786 EST and 5'-truncated NM_030175.1 RefSeq (Fig. 1A).

Rat *Hhip* family genes. Rat *Hhip1*, *Hhip2* and *Hhip3* genes were identified within AC107504.4, AC094820.6 and AC134264.2 genome sequences, respectively. Exon-intron boundaries of rat *Hhip1*, *Hhip2* and *Hhip3* genes were determined by examining the consensus sequence of exon-intron junctions and the codon usage. Rat *Hhip1*, *Hhip2* and *Hhip3* genes were found to consist of 13, 9, and 9 exons, respectively (Fig. 1B). Complete CDSs of rat *Hhip1*, *Hhip2* and *Hhip3* were determined by assembling exonic regions. Rat *Hhip1*, *Hhip2* and *Hhip3* genes were found to encode 700-, 791- and 712-amino-acid proteins, respectively (Fig. 2).

Comparative proteomics on *HHIP* family members. Membrane topology analyses were performed at first. *HHIP1* and *HHIP2* orthologs were secreted proteins with N-terminal signal peptide, while *HHIP3* orthologs were type II transmembrane proteins with short N-terminal cytoplasmic region.

Human, rat, and mouse *HHIP* family members were then aligned for comparative proteomics analyses. Although N- and



Figure 2. Alignment of vertebrate HHIP1, HHIP2 and HHIP3 orthologs. Hs, human. Rn, rat. Mm, mouse. HHIP homologous (HHIP) domain is boxed. Conserved Cys residues within the HHIP domain are shown by a cross.

C-terminal region of HHIP family members were divergent, core region corresponding to codon 68-595 of human HHIP1 was well conserved among mammalian HHIP family members (Fig. 2). The novel conserved region with 18 conserved Cys residues was designated the HHIP-homologous (HHIP) domain. These facts indicate that HHIP family members should be characterized as HHIP domain proteins.

Differential expression of HHIP1, HHIP2 and HHIP3 mRNAs. ESTs corresponding to HHIP1, HHIP2, and HHIP3 mRNAs were searched for by using the BLAST program. The sources of ESTs were then listed up for the *in silico* expression analyses. HHIP1 mRNA was expressed in coronary artery endothelial cells, prostate, and rhabdomyosarcoma. HHIP2 mRNA was expressed in trabecular bone cells. HHIP3 mRNA was expressed in testis, thyroid gland, osteoarthritic cartilage, pancreatic cancer, and lung cancer.

Comparative genomics on HHIP1, HHIP2 and HHIP3 orthologs. BLAST program as well as in house alignment of 5'-flanking regions revealed that HHIP1, HHIP2 and HHIP3 promoters were not well conserved between human and rodents.

Transcriptional regulation of human HHIP1 in coronary artery endothelial cells. We next analyzed the human HHIP1 promoter to elucidate the mechanism for HHIP1 expression

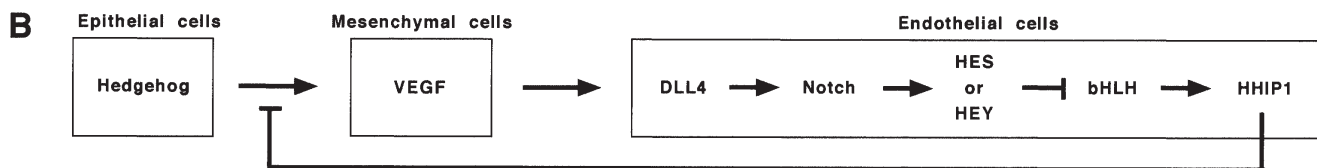
in coronary artery endothelial cells. GLI family transcription factors, TCF/LEF family transcription factors, and CSL transcription factor are implicated in the transcriptional regulation of Hedgehog, WNT, and Notch target genes, respectively (1,34-38). Because GLI-, TCF/LEF-, and CSL-binding sites were not identified within the human HHIP1 promoter (Fig. 3A), HHIP1 was not the direct transcriptional target gene of Hedgehog, WNT, and Notch signaling pathways.

Among HES/HEY family genes encoding transcriptional repressors with bHLH and orange domains, including HES1, HES2, HES3, HES4, HES5, HES6, HES7, HEY1, HEY2 and HEYL (39,40), at least HES1, HES5, HES7, HEY1, HEY2 and HEYL are best characterized Notch target genes. Expression of HES/HEY family members was not detected in coronary artery endothelial cells by using *in silico* expression analyses.

Eleven bHLH-binding E-boxes were identified within human HHIP1 promoter, while HES/HEY-binding N-box was not identified within human HHIP1 promoter (Fig. 3A). Because HES/HEY transcription factors repress bHLH factors, down-regulation of HES/HEY expression leads to up-regulation of HHIP1 mRNA depending on bHLH transcription factors (Fig. 3B).

Discussion

Mammalian HHIP family members were comprehensively identified and characterized in this study (Fig. 1). Complete



1. Katoh Y and Katoh M: Hedgehog signaling in gastric cancer. *Cancer Biol Ther* 4: 1050-1054, 2005.
2. Van den Brink GR, Bleuming SA, Hardwick JC, *et al*: Indian Hedgehog is an antagonist of Wnt signaling in colonic epithelial cell differentiation. *Nat Genet* 36: 277-282, 2004.
3. Garciadiego-Cazares D, Rosales C, Katoh M and Chimal-Monroy J: Coordination of chondrocyte differentiation and joint formation by $\alpha 5\beta 1$ integrin in the developing appendicular skeleton. *Development* 131: 4735-4742, 2004.
4. Marigo V, Roberts DJ, Lee SM, *et al*: Cloning, expression, and chromosomal location of *SHH* and *IHH*: two human homologues of the *Drosophila* segment polarity gene *hedgehog*. *Genomics* 28: 44-51, 1995.
5. Katoh Y and Katoh M: Identification and characterization of rat Desert hedgehog and Indian hedgehog genes *in silico*. *Int J Oncol* 26: 545-549, 2005.
6. Katoh Y and Katoh M: Comparative genomics on Sonic hedgehog orthologs. *Oncol Rep* 14: 1087-1090, 2005.

7. Johnson RL, Rothman AL, Xie J, *et al*: Human homolog of *patched*, a candidate gene for the basal cell nevus syndrome. *Science* 272: 1668-1671, 1996.
8. Katoh Y and Katoh M: Identification and characterization of *DISP3* gene *in silico*. *Int J Oncol* 26: 551-556, 2005.
9. Xie J, Murone M, Luoh SM, *et al*: Activating Smoothed mutations in sporadic basal-cell carcinoma. *Nature* 391: 90-92, 1998.
10. Hooper JF and Scott MP: Communicating with Hedgehogs. *Nat Rev Mol Cell Biol* 6: 306-317, 2005.
11. Briscoe J and Therond P: Hedgehog signaling: from the *Drosophila* cuticle to anti-cancer drugs. *Dev Cell* 8: 143-151, 2005.
12. Lum L and Beachy PA: The Hedgehog response network: sensors, switches, and routers. *Science* 304: 1755-1759, 2004.
13. Pasca di Magliano M and Hebrok M: Hedgehog signalling in cancer formation and maintenance. *Nat Rev Cancer* 3: 903-911, 2003.
14. Katoh Y and Katoh M: KIF27 is one of orthologs for *Drosophila* Costal-2. *Int J Oncol* 25: 1875-1880, 2004.
15. Katoh Y and Katoh M: Characterization of *KIF7* gene *in silico*. *Int J Oncol* 25: 1881-1886, 2004.
16. Taylor MD, Liu L, Raffel C, *et al*: Mutations in *SUFU* predispose to medulloblastoma. *Nat Genet* 31: 306-310, 2002.
17. Moore FL, Jaruzelska J, Dorfman DM and Reijo-Pera RA: Identification of a novel gene, DZIP (DAZ-interacting protein), that encodes a protein that interacts with DAZ (deleted in azoospermia) and is expressed in embryonic stem cells and germ cells. *Genomics* 83: 834-843, 2004.
18. Kinzler KW, Bigner SH, Bigner DD, *et al*: Identification of an amplified, highly expressed gene in a human glioma. *Science* 236: 70-73, 1987.
19. Teh MT, Wong ST, Neill GW, *et al*: *FOXMI* is a downstream target of Gli1 in basal cell carcinomas. *Cancer Res* 62: 4773-4780, 2002.
20. Yoon JW, Kita Y, Frank DJ, *et al*: Gene expression profiling leads to identification of GLI1-binding elements in target genes and a role for multiple downstream pathways in GLI1-induced cell transformation. *J Biol Chem* 277: 5548-5555, 2002.
21. Chuang PT and McMahon AP: Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature* 397: 617-621, 1999.
22. Olsen CL, Hsu PP, Glienke J, Rubanyi GM and Brooks AR: Hedgehog-interacting protein is highly expressed in endothelial cells but down-regulated during angiogenesis and in several human tumors. *BMC Cancer* 4: 43, 2004.
23. Martin ST, Sato N, Dhara S, *et al*: Aberrant methylation of the human *Hedgehog interacting protein (HHIP)* gene in pancreatic neoplasms. *Cancer Biol Ther* 4: 728-733, 2005.
24. Katoh M: Paradigm shift in gene-finding method: from bench-top approach to desk-top approach. *Int J Mol Med* 10: 677-682, 2002.
25. 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.
26. 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.
27. 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.
28. Katoh M and Katoh M: Identification and characterization of human *CKTSFIB2* and *CKTSFIB3* genes *in silico*. *Oncol Rep* 12: 423-427, 2004.
29. Katoh Y and Katoh M: Identification and characterization of rat *Wnt6* and *Wnt10a* genes *in silico*. *Int J Mol Med* 15: 527-531, 2005.
30. Katoh Y and Katoh M: Comparative genomics on *SFRP1* orthologs. *Int J Oncol* 27: 861-865, 2005.
31. Katoh M and Katoh M: Comparative genomics on *SFRP2* orthologs. *Oncol Rep* 14: 783-787, 2005.
32. Katoh Y and Katoh M: Comparative genomics on *DKK1* orthologs. *Int J Oncol* 27: 275-279, 2005.
33. Katoh Y and Katoh M: Comparative genomics on *DKK2* and *DKK4* orthologs. *Int J Mol Med* 16: 477-481, 2005.
34. Radtke F and Raj K: The role of Notch in tumorigenesis: oncogene or tumour suppressor? *Nat Rev Cancer* 3: 756-767, 2003.
35. Li JL and Harris AL: Notch signaling from tumor cells: a new mechanism of angiogenesis. *Cancer Cell* 8: 1-3, 2005.
36. Katoh M: Regulation of WNT signaling molecules by retinoic acid during neuronal differentiation in NT2 cells: threshold model of WNT action (Review). *Int J Mol Med* 10: 683-687, 2002.
37. Katoh M: *WNT* and *FGF* gene clusters (Review). *Int J Oncol* 21: 1269-1273, 2002.
38. Katoh M: *WNT2B*: comparative integromics and clinical applications (Review). *Int J Mol Med* 16: 1103-1108, 2005.
39. Katoh M and Katoh M: Identification and characterization of human *HES2*, *HES3*, and *HES5* genes *in silico*. *Int J Oncol* 25: 529-534, 2004.
40. 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.
41. 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.
42. 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.