Conserved POU-binding site linked to SP1-binding site within FZD5 promoter: Transcriptional mechanisms of FZD5 in undifferentiated human ES cells, fetal liver/spleen, adult colon, pancreatic islet, and diffuse-type gastric cancer

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 November 8, 2006; Accepted December 15, 2006

Abstract. Canonical WNT signals are transduced through Frizzled (FZD) family receptor and LRPS/LRP6 co-receptor to upregulate FGF20, JAG1, DKK1, WISP1, CCND1 and MYC genes for cell-fate determination, while non-canonical WNT signals are transduced through FZD family receptor and ROR2/PTK7/RYK co-receptor to activate RHOA/RHOU/RAC/CDC42, JNK, PKC, NLK and NFAT signaling cascades for the regulation of tissue polarity, cell movement, and adhesion. We previously reported molecular cloning and characterization of human FZD5, which showed six amino-acid substitutions with human Hfz5. FZD5, functioning as WNT5A receptor, is the key molecule in the fields of oncology, regenerative medicine, cardiology, rheumatology, diabetology, and gastroenterology. Here, comparative integromics analyses on FZD5 orthologs were performed by using bioinformatics (Techint) and human intelligence (Humint). Chimpanzee FZD5 and cow Fzd5 genes were identified within NW_104292.1 and AC166656.2 genome sequences, respectively. FZD5 orthologs were seven-transmembrane proteins with extracellular Frizzled domain, leucine zipper motif around the 5th transmembrane domain, and cytoplasmic DVL- and PDZ-binding motifs. Ser523 and Ser529 around the DVL-binding motif of FZD5 orthologs were putative aPKC phosphorylation sites. POU5F1 (OCT4)-binding site linked to SP1-binding site within the 5'-promoter region of human FZD5 gene was evolutionarily conserved among mammalian FZD5 orthologs. POU5F1 was more related to POU2F and POU3F subfamily members. POU5F1 was preferentially expressed in undifferentiated human embryonic stem (ES) cells, pancreatic islet, and diffuse-type gastric cancer. POU2F1 (OCT1) was expressed in ES cells, fetal liver/spleen, adult colon, POU2F2 in ES cells, fetal liver/spleen, and POU2F3 in diffuse-type gastric cancer. Multiple SP1/KLF family members, other than KLF2 or KLF4, were expressed in undifferentiated human ES cells. Together, these facts indicate that POU5F1 and POU2F subfamily members play a pivotal role for the FZD5 expression in undifferentiated human ES cells, fetal liver/spleen, adult colon, pancreatic islet, and diffuse-type gastric cancer.

Introduction

Canonical WNT signals are transduced through Frizzled (FZD) family receptor and LRPS/LRP6 co-receptor to upregulate FGF20, JAG1, DKK1, WISP1, CCND1 and MYC genes for cell-fate determination (1-13), while non-canonical WNT signals are transduced through FZD family receptor and ROR2/PTK7/RYK co-receptor to activate RHOA/RHOU/RAC/CDC42, JNK, PKC, NLK and NFAT signaling cascades for the regulation of tissue polarity, cell movement, and adhesion (14-20). WNT signals are context-dependently transduced to canonical and non-canonical signaling cascades. Cross-talk of WNT signaling pathway and FGF, Notch, Hedgehog and BMP/Nodal/TGFß signaling pathways constitute the stem-cell signaling network, which is implicated in embryogenesis and adult tissues homeostasis (21-31). Dysregulation of the stem cell signaling networks results in the pathological conditions, such as cancer, rheumatoid arthritis, and metabolic syndrome.

We previously reported molecular cloning and characterization of human FZD5 (32), which showed six amino-acid substitutions with human Hfz5 (33). FZD5, functioning as WNT5A receptor, is the key molecule in the fields of oncology, regenerative medicine, cardiology, rheumatology, diabetology, and gastroenterology. Here, comparative integromics analyses on FZD5 orthologs are described. Chimpanzee FZD5 and cow Fzd5 genes were identified and characterized in this study. In addition, comparative genomics analyses revealed conserved POU- and SP1/KLF-binding sites within the 5'-promoter region of FZD5 orthologs. POU5F1 and POU2F subfamily members were found to play a pivotal role for the expression of FZD5 in undifferentiated human embryonic stem (ES) cells, and other cells or tissues.
Materials and methods

Identification and characterization of chimpanzee and cow FZD5 orthologs. Chimpanzee and cow genome sequences homologous to human FZD5 were searched for with BLAST programs as described previously (34-36). 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 (37-39). Complete coding sequence (CDS) of chimpanzee FZD5 or cow Fzd5 was determined by assembling exonic regions.

Comparative genomics analyses. Human genome sequences around the FZD5 gene was compared with chimpanzee, cow, mouse and rat genome sequences to identify evolutionarily conserved regions. Binding sites for transcription factors, such as POU5F1, SOX2, NANOG and KLF, were then searched for.

Comparative proteomics analyses. Domain architecture of FZD5 orthologs were analyzed by using RPS-BLAST and PSORT II programs. Phylogenetic analyses on FZD and POU family proteins were performed by using the CLUSTALW program.

In silico expression analyses. Expressed sequence tags (ESTs) derived from human FZD5, POU family and SPI/ KLF family members were searched for by using the BLAST programs as described previously (40-42). The sources of human ESTs were listed up for in silico expression analyses.

Results

Chimpanzee FZD5 and cow Fzd5 genes. BLAST programs using human FZD5 RefSeq (NM_003468.3) revealed that chimpanzee FZD5 gene was located within NW_104292.1 genome sequence. Exon-intron boundaries of chimpanzee FZD5 gene were determined based on the consensus sequence of exon-intron junctions. Exon 1 corresponded to the nucleotide position 758775-758476 of NW_104292.1, while exons 2 to nucleotide position 758207-751775. Chimpanzee FZD5 gene was found consisting of two exons. Complete CDS of chimpanzee FZD5 was determined by assembling exonic regions. Genetyx program revealed that nucleotide position 555-2312 was the coding region. Chimpanzee FZD5 gene was found to encode a 585-amino-acid protein (Fig. 1A).

BLAST programs revealed that cow Fzd5 gene was located within AC166656.2 genome sequence. Complete CDS of cow Fzd5 was determined by assembling two exonic regions. Genetyx program revealed that nucleotide position 538-2295 was the coding region. Cow Fzd5 gene was found to encode a 585-amino-acid protein (Fig. 1B).

Comparative proteomics analyses on FZD5 orthologs. Chimpanzee FZD5 and cow Fzd5 showed 99.7% and 96.8% total-amino-acid identity with human FZD5, respectively.
amino-acid substitutions occur between human FZD5 and human Hfz5. Because all of the six amino-acids responding to human FZD5 were conserved among mammalian FZD5 orthologs (Fig. 2), it was concluded that six amino-acid substitutions in Hfz5 are due to sequencing errors or cloning artifacts.

FZD5 orthologs were seven-transmembrane proteins with extracellular Frizzled domain, leucine zipper motif around the 5th transmembrane domain, and cytoplasmic DVL- and PDZ-binding motifs (Fig. 2). Asn47 and Asn151 within the N-terminal extracellular region FZD5 orthologs were Asn-linked glycosylation sites. Ser523 and Ser529 around the DVL-binding motif of FZD5 orthologs were very similar to Ser554 and Ser560 of *Drosophila* Frizzled, which are phosphorylated by human aPKC.

In silico expression analysis on human FZD5. We have previously reported FZD5 mRNA expression in placenta, fetal liver, lung, kidney, adult pancreas, liver, colon, lung, kidney, K562 cells, SW480 cells, A549 cells, G361 cells, and HeLa S3 cells by using Northern blot analyses (32). We have also reported FZD5 mRNA expression in diffuse-type gastric cancer cells (44), and pancreatic islet (45). In this study, FZD5 mRNA was detected in pregnant uterus, placenta, amnion, undifferentiated human ES cells, fetal liver/spleen, adult colon, pancreatic islet, bone marrow stem cells, germ cell tumors, diffuse-type gastric cancer, colon cancer, and prostate cancer by using *in silico* expression analysis.

Transcriptional mechanism of FZD5 within undifferentiated human ES cells. Because FZD5 was preferentially expressed in undifferentiated human ES cells rather than differentiated human ES cells, POU5F1 (OCT4)-, SOX2-, and NANOG-binding sites around human FZD5 gene were searched for by using manual inspection. POU5F1-, SOX2- and NANOG-binding sites were identified within 5'-promoter region and 3'-UTR of human FZD5 gene. Among these sites, POU5F1-binding site at about 600-bp upstream of transcriptional start site for human FZD5 gene was conserved in chimpanzee, cow, mouse, and rat Fzd5 orthologs (Fig. 3A). NANOG-binding site at about 1500-bp upstream of transcriptional start site for human FZD5 gene was conserved in chimpanzee, cow and mouse FZD5 orthologs, but not in rat Fzd5 (Fig. 3A). Although SOX2-binding site was not linked to POU5F1-binding site within FZD5 5'-promoter region, SP1-binding site was linked to POU5F1-binding site with an interval of 3 bp (Fig. 3A). POU5F1-binding site linked to SP1-binding site within the 5'-promoter region of human FZD5 gene was evolutionarily conserved among mammalian FZD5 orthologs (Fig. 3A).

**Figure 2. Mammalian FZD5 orthologs.** Hs, human; Pt, chimpanzee; Bt, cow; Mm, mouse; Rn, rat. Signal peptide and seven-transmembrane domains (TM1-TM7) are boxed. Conserved Cys residues (cross) and Asn-linked glycosylation sites (sharp) within the N-terminal extracellular Frizzled region, leucine zipper motif around the TM5 domain (over line), DVL-binding and PDZ-binding motifs within the C-terminal cytoplasmic region (double over line), Ser523 and Ser529 around the DVL-binding motif (open arrowhead) are conserved among mammalian FZD5 orthologs. Six amino-acid substitutions between human FZD5 and human Hzd5 ([]) are also shown.

---

**In silico expression analyses revealed that SP1, SP2, SP3, and SP4 promoters are bound by RNA polymerase II in undifferentiated human ES cells (46).** In *in silico* expression analysis, we further investigated the transcriptional mechanism of FZD5 within undifferentiated human ES cells by using *in silico* expression analysis. Together, these facts indicate that multiple SP1/KLF family members and POU5F1 are implicated in FZD5 expression in undifferentiated human ES cells, and also that POU5F1 plays a key role in the FZD5 expression in undifferentiated human ES cells.

**Figure 2. Mammalian FZD5 orthologs.** Hs, human; Pt, chimpanzee; Bt, cow; Mm, mouse; Rn, rat. Signal peptide and seven-transmembrane domains (TM1-TM7) are boxed. Conserved Cys residues (cross) and Asn-linked glycosylation sites (sharp) within the N-terminal extracellular Frizzled region, leucine zipper motif around the TM5 domain (over line), DVL-binding and PDZ-binding motifs within the C-terminal cytoplasmic region (double over line), Ser523 and Ser529 around the DVL-binding motif (open arrowhead) are conserved among mammalian FZD5 orthologs. Six amino-acid substitutions between human FZD5 and human Hzd5 ([]) are also shown.
POU2F3, POU3F1, POU3F2, POU3F3 and POU3F4 (Fig. 3B). POU5F1-binding sequence within the 5'-promoter of FZD5 gene was identical to that within the 3'-UTR of mouse Fgf4 gene, to which POU5F1, POU2F1 (OCT1) and POU3F1 (OCT6) can bind (47). Based on these facts, expression of POU5F1, POU2F1, POU2F2, POU2F3, POU3F1, POU3F2, POU3F3 and POU3F4 were next searched for to investigate the mechanism of FZD5 expression in fetal liver/spleen, adult colon, pancreatic islet, and diffuse-type gastric cancer. POU5F1 was preferentially expressed in pancreatic islet, and diffuse-type gastric cancer. POU2F1 (OCT1) was expressed in fetal liver/spleen, and adult colon. POU2F2 was expressed in fetal liver/spleen. POU2F3 was expressed in diffuse-type gastric cancer. POU5F1 and POU2F subfamily members were implicated in FZD5 expression in fetal tissues, adult tissues, and cancer (Fig. 3C).

**Discussion**

Comparative integomics analyses on FZD5 orthologs were performed in this study, because FZD5 is the key molecule in a variety of fields, such as oncology, regenerative medicine, and gastroenterology. Chimpanzee FZD5 and cow Fzd5 genes were identified within NW_104292.1 and AC166656.2 genome sequences, respectively. Chimpanzee FZD5 and cow Fzd5 genes, consisting of two exons, encoded a 585-amino-acid protein (Fig. 1). Human FZD5 showed 99.7% and 96.8% total-amino-acid identity with chimpanzee FZD5 and cow Fzd5, respectively.

FZD5 orthologs were seven-transmembrane proteins with extracellular Frizzled domain, leucine zipper motif around the 5th transmembrane domain, and cytoplasmic DVL- and PDZ-binding motifs (Fig. 2). Leucine zipper motif around the 5th transmembrane domain might be implicated in homo- or hetero-dimerization of Frizzled family members.

Ser523 and Ser529 around the DVL-binding motif within the C-terminal cytoplasmic region of FZD5 orthologs were putative aPKC phosphorylation sites. Because Ser554/Ser560 phosphorylation of Drosophila Frizzled by aPKC results in the inhibition of planar polarity signaling, Ser523 and Ser529 are critical amino-acid residues to determine the FZD5 signaling to canonical or non-canonical WNT signaling cascades.

POU5F1-binding site linked to SP1-binding site within the 5'-promoter region of human FZD5 gene was evolutionarily conserved among mammalian FZD5 orthologs (Fig. 3A). Klf4 and Klf2 are implicated in the maintenance of undifferentiated mouse ES cells (48), and Klf4 is also implicated in the ES-like transformation of mouse embryonic fibroblast (49). KLF1, KLF2 and KLF4, constituting a subfamily among KLF/SP1 family members, bind to CACCC motif rather than CGCCC motif (50). KLF1, KLF2 and KLF4 were not expressed in undifferentiated human ES cells, while SP1, SP2, SP3, KLF10 and KLF13 were expressed in undifferentiated human ES cells as well as in differentiated human ES cells. Therefore, multiple SP1/KLF family members, other than KLF2 or KLF4, were implicated in FZD5 expression in undifferentiated human ES cells.

POU5F1 was preferentially expressed in undifferentiated human ES cells, pancreatic islet, and diffuse-type gastric cancer. POU2F1 was expressed in ES cells, fetal liver/spleen, adult colon, pancreatic islet, and diffuse-type gastric cancer. Together, these facts indicate that POU5F1 and POU2F subfamily members play a pivotal role in the FZD5 expression in undifferentiated human ES cells, fetal liver/spleen, adult colon, pancreatic islet, and diffuse-type gastric cancer.
References