Integrative genomic analyses on GLI1: Positive regulation of GLI1 by Hedgehog-GLI, TGFβ-Smads, and RTK-PI3K-AKT signals, and negative regulation of GLI1 by Notch-CSL-HES/HEY, and GPCR-Gs-PKA signals

YURIKO KATOH¹ and MASARU KATOH²

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

Received April 22, 2009; Accepted May 19, 2009

DOI: 10.3892/ijo_00000328

Abstract. GLI family members are zinc-finger transcription factors, which are involved in embryogenesis and carcinogenesis through transcription regulation of GLI1, CCND1, CCND2, FOXA2, FOXC2, RUNX2, SFRP1, and JAG2. GLI1 transcription is upregulated in a variety of human tumors, such as basal cell carcinoma, lung cancer, breast cancer, gastric cancer, pancreatic cancer, and esophageal cancer. Hedgehog signaling via Smoothened cascade and receptor tyrosine kinase (RTK) signaling via PI3K-AKT cascade induce stabilization of GLI1 protein, whereas G-protein coupled receptor (GPCR) signaling via Gs-PKA cascade induces degradation of GLI1 protein. Here we report integrative genomic analyses of the GLI1 gene. The GLI1 and ARHGAP9 genes are located in a tail-to-tail manner with overlapping 3'-ends. ARHGAP9 was expressed in bone marrow, spleen, thymus, monocytes, and macrophages, whereas GLI1 was almost undetectable in normal tissues or cells with predominant ARHGAP9 expression. Because overlapping sense and anti-sense transcripts are annealed to each other to give rise to double-stranded RNAs functioning as endogenous RNAi, GLI1 expression might be negatively regulated by GLI1 transcripts. GLI-binding element with one base substitution at the +1589-bp position from the transcriptional start site (TSS) of the human GLI1 gene was completely conserved in chimpanzee GLI1, mouse Gli1, and rat Gli1 genes. Ten Smad-binding elements, double E-boxes for EMT regulators, and double N-boxes for HES/HEY family members within intron 1 of the human GLI1 gene were also conserved in mammalian GLI1 orthologs. GLI1 transcription is upregulated due to Hedgehog, and TGFβ signaling activation, whereas GLI1 transcription is downregulated due to Snail/Slug, and Notch signaling activation. Together these facts indicate that Hedgehog, TGFβ, and RTK signals positively regulate GLI1, and that Notch, and GsPCR signals negatively regulate the GLI1.

Introduction

Sonic Hedgehog (SHH), Indian Hedgehog (IHH), and Desert Hedgehog (DHH) are secreted-type glycoproteins with lipid modifications involved in embryogenesis, adult tissue homeostasis, and carcinogenesis (1-4). Mature Hedgehog proteins bind to Patched family receptors to activate Smoothened signaling cascade, which results in GLI-mediated transcriptional upregulation of Hedgehog target genes (5-8).

GLI1, GLI2, and GLI3 are GLI family members (9-14). GLI1 consists of zinc finger domains, and C-terminal activator domain, whereas GLI2 and GLI3 consist of N-terminal repressor domain, zinc finger domains, and C-terminal activator domain. In the absence of Hedgehog signals, GLI1 is transcriptionally repressed, GLI2 is phosphorylated for the FBXW1/ßTRCP1-mediated processing and GLI3 is phosphorylated for the FBXW1/ßTRCP1-mediated processing into repressor (10-12). In the presence of Hedgehog signals, stabilized GLI family members induce transcriptional activation of target genes, such as GLI1, CCND1, CCND2, FOXA2, FOXC2, RUNX2, SFRP1, and JAG2 (15-21). Hedgehog-induced GLI1 upregulation augments the Hedgehog signaling cascades through a positive-feedback mechanism. GLI1 is upregulated in various types of human tumors, as summarized in Table I. GLI1 is amplified and overexpressed in glioma (9), rhabdomyosarcoma (22), and osteosarcoma (22). GLI1 is also amplified in B-cell lymphoma (23). Although GLI1 is almost undetectable in adult human tissue by using RNase A protection assay, GLI1 expression is detected in embryonal carcinoma (24). Because GLI1 is a representative target gene of the Hedgehog signaling cascades, GLI1 is upregulated in basal cell carcinoma (25), medulloblastoma (26), lung cancer (27), gastric cancer (28,29), pancreatic cancer (28,30), esophageal cancer (28,31), breast cancer (32), prostate...
GLI1 is also upregulated based on the EWS-FLI fusion protein generated by chromosomal translocation in Ewing sarcoma (39).

The GLI1 core promoter, defined as a region surrounding the transcription start site (TSS), is involved in the transcriptional regulation of human GLI1 gene (40). Hedgehog signals induce transcriptional upregulation of mouse Gli1 by using GLI-binding site within intron 1 (41), whereas TWIST1/TWIST induces transcriptional upregulation of human GLI1 by using E-boxes within intron 1 (42). Because proximal promoter region and intron 1 are involved in transcriptional regulation of GLI1, we carried out refined integrative genomic analyses of GLI1 to elucidate the mechanisms of GLI1 upregulation in human tumors with the emphasis on the stem-cell signaling network.

Materials and methods

Comparative genomic analyses. Human genome sequences corresponding to human GLI1 RefSeq (NM_005269.2) were searched for by using BLAST programs, as previously described (43,44). GLI1 expressed sequence tags (ESTs) were also searched for to identify GLI1 splicing variants (45,46). Conserved transcription factor-binding sites within GLI1 promoters were then searched for based on manual inspection, as previously described (47,48).

Regulatory network analyses. The literature on Hedgehog, TGFß, Notch and receptor tyrosine kinase (RTK) signaling cascades in PubMed and Medline databases was critically evaluated to extract knowledge. The mechanisms of GLI1 transcription were then investigated based on our data of conserved transcription factor-binding sites within GLI1 regulatory regions and in-house knowledgebase of transcription factors regulated by the stem-cell signaling network.

Results

Comparative genomic analyses of GLI1. BLAST programs using GLI1 RefSeq (NM_005269.2) as a query sequence revealed that GLI1 gene is located within human genome sequence AC022506.38. Human GLI1 gene, consisting of 12 exons, is about 12 kb in size, as previously reported by Liu et al (40). Human GLI1 gene is located between the INHBC-INHBE gene cluster and the ARHGAP9 gene at human chromosome 12q13.3. The GLI1 gene is the paralog of the GLI2 gene at human chromosome 2q14.12, whereas the INHBC and INHBE genes encoding Inhibin family members are paralogs of the INHBB gene neighboring the GLI2 gene (data not shown). The INHBC-INHBE-GLI1 locus at human chromosome 12q13.3 and the INHBB-GLI2 locus at human chromosome 2q14.12 are paralogous regions within the human genome.

The GLI1 and ARHGAP9 genes are located in a tail-to-tail manner with the overlapping 16 bases in their 3'-ends (Fig. 1). ARHGAP9 is expressed in peripheral blood leukocytes, spleen, thymus (49), and also in thyroid gland, testis, bone marrow, monocytes, and macrophages (RefEX database). GLI1 is almost undetectable in normal tissues with predominant ARHGAP9 expression (RefEX database). Because overlapping sense and anti-sense transcripts are annealed to
each other to give rise to double-stranded RNAs functioning as endogenous RNAi, GLI1 expression might be negatively regulated by the ARHGAP9 transcripts.

**Comparative genomic analyses of GLI1 orthologs.** BLAST programs using human GLI1 genome sequence as a query sequence revealed that chimpanzee GLI1 gene, mouse Gli1 gene, and rat Gli1 gene were located within NW_001223153.1, AC114678.20, and AC122965.5 genome sequences, respectively. GLI1 orthologs were well conserved not only within exonic regions, but also within the proximal promoter region and intron 1 (data not shown).

Consensus GLI-binding element within the regulatory regions of human GLI1 gene was searched for based on manual inspection, and a GLI-binding element with one base substitution was identified at the +1589-bp position from the TSS (Fig. 1). The GLI-binding element was completely conserved in intron 1 of chimpanzee GLI1, mouse Gli1, and rat Gli1 genes (Fig. 1). Vokes et al carried out chromatin-immunoprecipitation assay combined with microarray analyses (ChIP-chip assay) of GLI-binding sites in mouse embryonic stem (ES) cells, and showed that the major peak of GLI-binding site within the regulatory regions of mouse Gli1 gene was located within intron 1 (41). We confirmed that the conserved GLI-binding element identified in this study was located within the genome fragment immunoprecipitated by using the anti-Gli1 antibody (data not shown). Therefore, it was concluded that the GLI-binding site at about +1.5-kb position from the TSS was evolutionarily conserved among mammalian GLI1 orthologs (Fig. 1).

Smad-binding element (SBE) within the regulatory regions of human GLI1 gene was next searched for based on manual inspection, and 10 SBEs within intron 1 of human GLI1 gene were found to be conserved in mouse Gli1 gene (Fig. 1). These 10 SBEs were also conserved in chimpanzee GLI1 and rat Gli1 genes (Fig. 1). TGFβ signals induce Smad3-dependent upregulation of GLI1 and Gli2 in human NHDF fibroblasts, HaCaT keratinocytes, and MDA-MB-231 breast cancer cells (50); however precise mechanisms of Smad3-mediated GLI1 upregulation remained unclear. Conserved SBEs within intron 1 of mammalian GLI1 orthologs clearly elucidated the mechanism of TGFβ-induced GLI1 upregulation (Fig. 2).

Six bHLH-binding sites within intron 1 of human GLI1 gene were conserved in chimpanzee GLI1, mouse Gli1, and rat Gli1 genes (data not shown). Two bHLH-binding sites located within the 5'-region of intron 1 (Fig. 1) were the specific
E-boxes utilized as binding sites for TWIST1 (42). SNAI1/Snaill, SNAI2/Slug, ZEB1/EF1/ZFHX1A, ZEB2/SIP1/ZFHX1B, and TWIST2/DERMO1 also bind to the same E-boxes as TWIST1 to regulate the transcription of target genes (51-54). ZEB1, ZEB2, TWIST1, and TWIST2 function as transcriptional activators or repressors in a context-dependent manner, whereas SNAI1 and SNAI2 function as transcriptional repressors. Based on these facts, it was predicted that SNAI1 and SNAI2 repress GLI1 expression via the E-boxes (Fig. 2).

Double N-boxes within intron 1 of human GLI1 gene were also conserved in chimpanzee GLI1, mouse Gli1, and rat Gli1 genes (Fig. 1). Notch ligands induce Notch signaling activation through the NICD-CSL complex to upregulate Notch target genes, such as HES1, HES5, HEY1, HEY2 and HEYL (55-58). Consensus CSL-binding site was not conserved in mammalian GLI1 orthologs, whereas double N-boxes utilized as HES/HEY-binging sites within intron 1 of human GLI1 gene were conserved in chimpanzee GLI1, mouse Gli1, and rat Gli1 genes (Fig. 1). Together these facts indicate that Notch signals downregulate GLI1 expression via HES/HEY family members (Fig. 2).

Discussion

Integrative genomic analyses of GLI1 gene were carried out in this study. The GLI1 and ARHGAP9 genes are located in a tail-to-tail manner with overlapping 3'-ends (Fig. 1). ARHGAP9 was expressed in bone marrow, spleen, thymus, monocytes, and macrophages, whereas GLI1 was almost undetectable in normal hematopoietic tissues or cells with predominant ARHGAP9 expression. Because overlapping sense and anti-sense transcripts are annealed to each other to give rise to double-stranded RNAs functioning as endogenous RNAi, GLI1 expression might be negatively regulated by ARHGAP9 transcripts.

GLI-binding element with one base substitution at the +1589-bp position from the TSS of human GLI1 gene was completely conserved in chimpanzee GLI1, mouse Gli1, and rat Gli1 genes (Fig. 1). Ten Smad-binding elements, double E-boxes and double N-boxes within intron 1 of human GLI1 gene were also conserved in mammalian GLI1 orthologs (Fig. 1). Hedgehog and TGFß signals induce transcriptional upregulation of target genes via GLI and Smads, respectively (Fig. 2). On the other hands, Snail/Slug and Notch-HES/HEY signals induce transcriptional downregulation of target genes via E-box and N-box, respectively (Fig. 2). GLI1 transcription is upregulated due to Hedgehog, and TGFß signaling activation, whereas GLI1 transcription is downregulated due to Snail/Slug, and Notch signaling activation.

GLI1 is activated based on transcriptional activation as well as protein stabilization (Fig. 2). Hedgehog and TGFß signals induce transcriptional upregulation of GLI1, whereas Notch signals induce transcriptional downregulation of GLI1. RTK signals via the PI3K-AKT signaling cascade induce stabilization of GLI1 protein, whereas G-protein coupled
receptor (GPCR) signals via the Gs-PKA signaling cascade induce degradation of the GLI1 protein. Together these facts indicate that Hedgehog, TGFβ, and RTK signals positively regulate GLI1, and that Notch, and GsPCR signals negatively regulate the GLI1 (Fig. 2).

Hedgehog, TGFβ/BMP, RTK, Notch, and WNT signaling cascades constitute the stem-cell signaling network to orchestrate embryogenesis and carcinogenesis (59-63). Dysregulation of the stem-cell signaling network leads to GLI1 upregulation to promote carcinogenesis (Fig. 2). KAAD-cyclopamine and HhAntag are small-molecule compounds indirectly targeted to GLI1 and GLI2 via Smoothened (64,65), whereas GANT61 is a small-molecule compound directly targeted to GLI1 and GLI2 (66), GLI-targeted small-molecule compounds could show anti-tumor effects to a broader range of cancer than Smoothened-targeted small-molecule compounds. To realize personalized medicine in the sequencing era with a peta-scale supercomputer (67), development of GANT61-like compounds for clinical application is expected.

References