Article

Genetic and Epigenetic Regulation of the Smoothened Gene (SMO) in Cancer Cells

Hong Lou [1],[†], Hongchuan Li [2],[†], Andrew Huehn [3],[4], Nadya Tarasova [3], Bahara Saleh [3], Stephen Anderson [2],[3],[*] and Michael Dean [5],[*]

[1] Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, Leidos Biomedical Research, Inc., National Laboratory for Cancer Research, Gaithersburg, MD 20892, USA; louho@mail.nih.gov

[2] Basic Science Program, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA; lihongchu@mail.nih.gov

[3] Laboratory of Cancer Immunometabolism, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702, USA; tarasovn@mail.nih.gov;   baharasaleh93@gmail.com; andersonst@mail.nih.gov

[4] Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06510, USA; andrew.huehn@yale.edu

[5] Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Gaithersburg, MD 20892, USA; deanm@mail.nih.gov

[†]These authors equally contributed to this paper.

[*] Authors to whom correspondence should be addressed.

Correspondence: deanm@mail.nih.gov (M.D.); andersonst@mail.nih.gov (S.A.); Tel.: +1-240-760-6484 (M.D.); +1-301-846-1330 (S.A.); Fax: +1-xxx-xxx-xxxx (M.D.); +1- 301-846-1673 (S.A.)

Abstract: 1) Background: The hedgehog (HH) signaling pathway is a key regulator of embryonic patterning, tissue regeneration, stem cell renewal, and cancer growth. The smoothened (SMO) protein regulates the HH signaling pathway and has demonstrated oncogenic activity. **2) Methods:** To clarify the role of the HH signaling pathway in tumorigenesis, the expression profile of key HH signaling molecules, including *SMO*, *PTCH1*, *GLI1*, *GLI2*, and *GLI3*, were determined in thirty-three cancer cell lines. We performed a computational analysis of the upstream region of the *SMO* gene to identify the regulatory elements. **3) Results:** Three potential CpG islands and several putative *SMO* promoter elements were identified. Luciferase reporter assays mapped key *SMO* promoter elements, and functional binding sites for SP1, AP1, CREB, and AP-2α transcription factors in the core SMO promoter region were confirmed. A hypermethylated *SMO* promoter was identified in several cancer cell lines suggesting an important role for epigenetic silencing of *SMO* expression in certain cancer cells. **4) Discussio**n: These results have important implications for our understanding of regulatory mechanisms controlling HH pathway activity and the molecular basis of *SMO* gene function. Moreover, this study may prove valuable for future research aimed at producing therapeutic downregulation of *SMO* expression in cancer cells.

Keywords: Hedgehog pathway; Smoothened; Promoter; Methylation; Transcription factors

1. Introduction

The hedgehog (HH) pathway is one of the key signaling pathways regulating embryonic patterning, tissue regeneration, stem cell renewal, and cancer growth [1-4]. Canonical HH signaling is triggered by the binding of HH ligand to its receptor PTCH1, resulting in the release of PTCH1 mediated repression of the seven-transmembrane protein smoothened (SMO). Activation of SMO ultimately triggers GLI dependent expression of downstream target genes through a complex network of post-translational processes and translocations [5]. In the absence of HH ligands, PTCH inhibits SMO, and GLI2, GLI3 are phosphorylated and undergo partial proteasome degradation, resulting in repressive forms of GLI2 and GLI3 (GLI2/3 R), that translocate into the nucleus where they inhibit the transcription of HH target genes [6,7]. HH pathway activation amplifies the signal

by increasing GLI1 levels, and in contrast, potentiates negative regulators such as PTCH1 and HH interacting protein (HHIP) [8,9]. The positive and negative feedback loops ensure that the activity of HH signaling is kept within an optimal range. Constitutive activation of the HH pathway has been observed in various types of malignancies caused either by mutations in the pathway, such as PTCH1 loss-of function or SMO activation in basal cell carcinoma, or through HH overexpression, as observed in small-cell lung cancer, glioma, endometrial carcinoma, digestive tract tumors, pancreas, and prostate [10-16].

The switch between active and inactive states of the HH pathway involves rapid translocation of SMO. The SMO protein is the key positive regulator of the HH pathway, and GLI family proteins play a critical role in the regulation of HH signaling pathway activity. Despite a strong link between SMO expression, HH pathway activity, and cancer development, the basis for SMO gene regulation has not been well characterized. Therefore, an investigation of the mechanisms controlling the expression of SMO and additional HH pathway genes may provide valuable insight into HH signaling alterations associated with cancer development. SMO also is the major target for pharmaceutical agents that modulate HH pathway activity [17-19], such as vismodegib [20] and sonidegib [21]. We previously studied SMO peptides, and found that specific lipopeptides can serve as effective inhibitors [22,23].

DNA methylation of HH pathway genes is a potential regulatory mechanism in the progression of cancers. Several epigenetic factors that act on the HH signaling pathways have been associated with cancer initiation and progression [10,24]. It was reported that distinct subgroups of cancers have an exceptionally high frequency of cancer specific CpG island hypermethylation [25,26]. Methylation has been studied as a clinical biomarker for the diagnosis and prognostic evaluation of various cancers, especially in breast cancer [27]. Recently, SMO methylation was used as a biomarker for the occurrence and development of breast cancer [28].

In the current study, we have developed a qRT-PCR method to accurately determine the expression levels of SMO, PTCH1, GLI1, GLI2 and GLI3 in a panel of cancer cell lines. Different SMO expression patterns in the cancer cell lines led us to characterize SMO gene regulatory elements. The SMO 5'-flanking region and Exon 1 was analyzed in silico, revealing that the region surrounding the SMO transcriptional start site (TSS) has an extremely high GC content (70%+) that prevents its PCR amplification by traditional methods. We used a touchdown PCR method to amplify SMO promoter fragments and determined their promoter activity using a dual luciferase assay. EMSA analysis identified binding sites for the transcription factors, SP1, AP1, CREB, and AP-2, which likely play an important role in SMO transcriptional activity in cancer cells. In order to gain insight into the epigenetic regulation of SMO, bisulfite sequencing PCR (BSP) and methylation specific PCR (MSP) were carried out to determine methylation status of the potential SMO promoter region. The relationship between the methylation status and SMO mRNA expression was analyzed.

2. Results

2.1 Expression of HH signaling molecules in cancer cell lines

In order to gain a greater understanding of the transcriptional regulation of the HH pathway components, a careful analysis of mRNA levels for key HH signaling genes was conducted. Quantification of mRNA expression levels using a real-time PCR method is increasingly used to determine the activity of HH signaling genes. However, most studies use a relative RT-PCR method, which is less precise, and does not provide meaningful comparisons of gene expression between different cell lines. Therefore, an accurate quantitation of HH pathway mRNA expression is necessary. In this study, we developed and validated a standard curve based on a Taqman qRT-PCR method to measure key HH signaling genes, including SMO, PTCH1, GLI1, GLI2 and GLI3.

The results produced a broad linear dynamic range of detection of at least 6 logs, and a small quantitative variation produced by triplicate analysis. The slope of the curve was used to determine the reaction efficiency. Efficiency = [10(-1/slope)] – 1 [29]. The efficiency of standard curve for all genes in is greater than 92% and R2 is greater than 0.99.

The expression levels of HH signaling components were determined in 33 cancer cell lines using the quantitative Taqman RT-PCR method. Expression levels of the HH pathway genes, the HH signaling receptors PTCH and SMO, and the target transcription factors GLI2, and GLI3, varied significantly among the cancer cell lines (Figure 1). The SMO gene exhibited the highest level of mRNA expression and the greatest variation between cell lines (22.98±31.80, 95% CI 11.70~31.25), compared with PTCH1 (8.72±10.44, 95% CI 5.02~12.42), GLI2 (1.25±2.08, 95% CI 0.56~1.99) and GLI3 (18.04±20.26, 95% CI 10.85~25.22), whereas GLI1 (0.17±0.25, 95% CI 0.08~0.26) that functions as an amplifier of HH signal, consistently showed low expression in all cancer cell lines. In addition, the absence of SMO expression was confirmed in seven cell lines, including 5 breast cancer cell lines and the stomach cancer AGS cell line, whereas the lack of SMO was accompanied by undetectable GLI3 in colon cancer HT29 cells (Table 1).
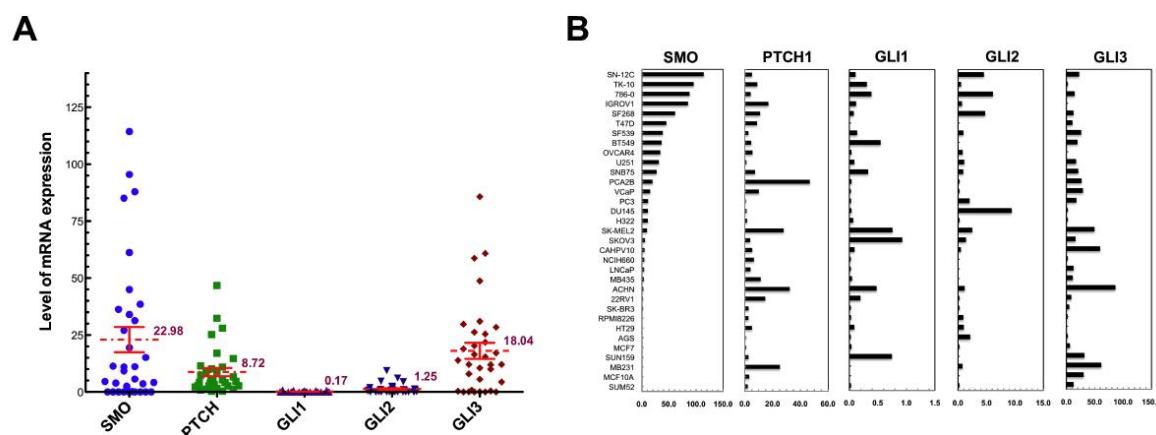

Figure 1. (A) Distribution of mRNA levels of SMO, PTCH, GLI1, GLI2 and GLI3 in 33 cancer cell lines. Mean values ± SCE of each gene are indicated by horizontal bars. (B) Comparison of mRNA levels. The mRNA levels of the five genes were quantitated from the Taqman RT-PCR as described in Materials and Methods. Data are presented by cell line in decreasing order of SMO mRNA level.

Table 1.   Expression level of hedgehog pathway genes in 33 tumor cell lines

Tissue	Cell lines	Level of mRNA expression*				
		SMO	PTCH	GLI1	GLI2	GLI3
Prostate	PCA2B	19.46	46.74	0.03	0.003	26.22
	VCaP	15.15	10.15	0	0.1	28.43
	PC3	11.33	0.39	0.02	1.96	17.13

|						
	DU145	11.14	1.27	0.03	9.36	0.47
	CAHPV10	4.59	5.31	0.08	0.43	58.71
	NCIH660	4.18	6.61	0.01	ND	0.02
	LNCaP	3.9	4.1	0.01	ND	12.09
	22RV1	0.18	14.68	0.19	0.01	7.91
Breast	T47D	44.96	8.71	ND	ND	9.99
	BT549	36.26	4.64	0.54	ND	18.91
	MB435	3.56	11.43	0.04	ND	10.56
	SK-BR3	0.02	2.72	0	0	4.39
	SUM52	ND	2.16	0.02	0	12.09
	MCF10A	ND	3.24	ND	ND	29.75
	MB231	ND	25.2	ND	0.71	60.85
	SUN159	ND	2.68	0.74	ND	31.01
	MCF7	ND	0.87	0.03	0.01	5.54
Kidney	SN-12C	114.33	5.11	0.1	4.47	21.85
	TK-10	95.49	8.95	0.3	0.46	0.08
	786-0	87.9	4.22	0.38	6.07	13.87
	ACHN	2.54	32.37	0.47	1.09	85.73
Glioblastoma	U251	31.36	1.19	0.08	1.01	16.5
	SF539	38.5	2.61	0.13	0.85	25.44
	SNB75	27.07	7.33	0.32	0.83	20.28
	SF268	61.21	10.97	0.07	4.7	11.96
Ovary	IGROV1	85.11	17	0.11	0.62	ND
	OVCAR4	33.99	5.46	0.03	0.74	ND
	SKOV3	5.65	3.85	0.92	1.32	15.48
Others						
Stomach	AGS	ND	0.37	0.01	2.07	0.08
Skin	SK-MEL2	9.28	27.98	0.75	2.43	48.77
Colon	HT29	ND	5.09	0.08	0.93	ND
Lung	H322	11	1.64	0.06	0.19	0.75
Myeloma	RPMI8226	0.02	2.64	0.02	0.85	0.37

Abbreviations: ND, not detectable

*Relative mRNA expression = (Target gene / 18s rRNA) x 1,000,000

Significant positive correlations were identified between SMO and GLI2 transcript levels (Pearson's correlation=0.359, p=0.040), and between PTCH1 and GLI3 (Pearson's correlation=0.532,

p=0.001). No correlation was found between the expression levels of SMO/PTCH or SMO/GLI1 (Figure S1).

2.2. Interspecies comparison of genomic SMO sequences

For a comparison of mammalian SMO genes, we analyzed 20,500 bp of SMO genome sequence including the upstream, 5'-UTR, exon 1 and part of intron 1 regions for human, mouse, and rhesus monkey. Multiple sequence alignment of the 20,500 bp of SMO was performed by the mVista web-tool (Figure 2). The macaque sequence is highly homologous to the human sequence, as 87.5% of the 20-kb region showed at least 88% sequence identity over a 100 bp window. In contrast, the mouse sequence shares 71.2 % identity with human, with several conserved noncoding sequences (CNS) showing at least 70% identity over 100 bps.

The global genomic sequence comparison showed significant highly conserved regions among the three genes immediately upstream of the transcriptional start site (P = 3.9x10-25). Five additional CNS were identified, but all had a lower p value that was greater than 0.005. A 924 bp region consisting of the full length 5'-UTR and exon 1 regions, along with 293 bp of upstream and 20 bp of intron 1 sequence is shown in Figure 2b. This region possesses 83% identity between the mouse and human SMO genes, and 95% between human and rhesus.
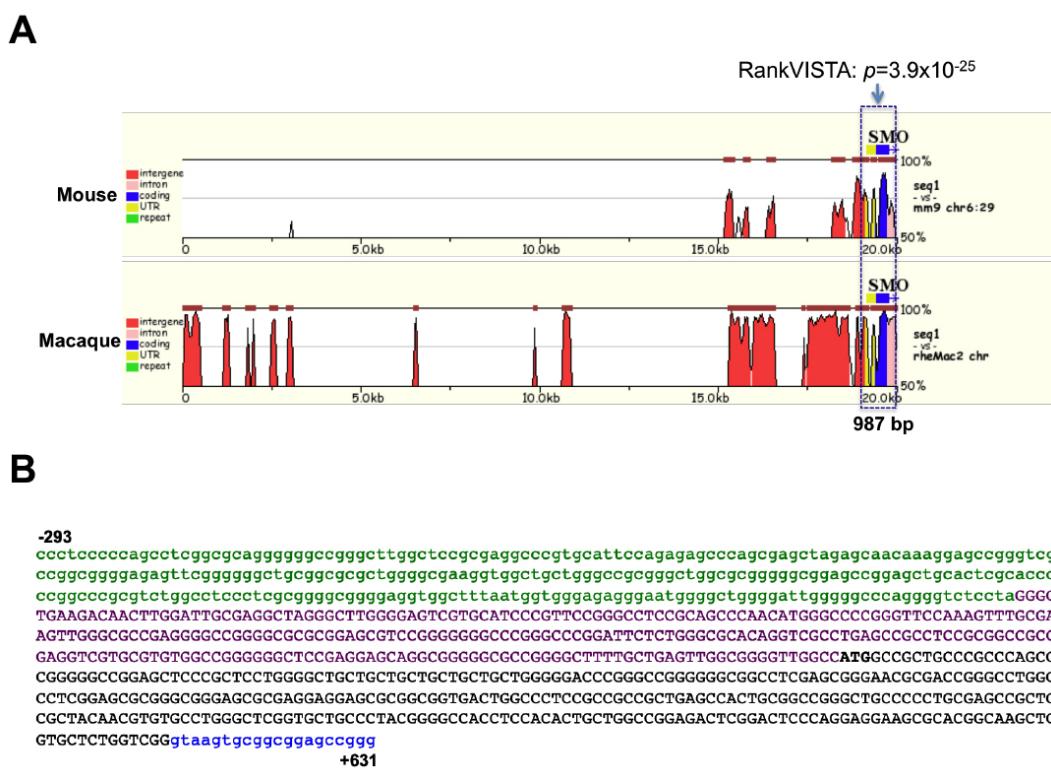
Figure 2. (A) Alignment of the 5'-flanking region of three mammalian SMO genes. 20 kb of upstream sequence including exon-1 from the mouse, macaque and human SMO genes was aligned by the MLAGAN algorithm of the mVista program. The sequences of mouse and macaque are aligned to the human SMO sequence (x-axis); numbering is relative to the transcriptional start site. Conserved regions (> 70% homology over 100 bp window) are shaded. The box indicates conserved regions among the three sequences, as determined by RankVista (P ≤ 10-5), with the P-values given above. (B) The structure of 924 bp conserved sequence. The 5'-flanking region is defined as sequence upstream (from the 5' end) of the transcript start site and shown in green lower-case letters. Intron is shown in blue lower-case letters. The exon is shown in uppercase letters, with UTR as purple and coding sequence as black color, respectively.

2.3. In silico analysis of the SMO upstream regulatory region

Submission of a 1,611 bp sequence, including 1,000 bp 5'-upstream region and the full Exon 1 of the human SMO gene, to the MatInspector software program (core similarity > 0.85; matrix similarity optimized) returned 336 potential TFBS, distributed over the entire sequence. Furthermore, we performed analysis with the PromoterInspector program and found a 1028 bp potential promoter region within this sequence, located from position -508 to +520 bp. By using the ModelInspector program, twenty-five models were identified, including SMAD-MIT, SMAD-AP1, YY1-SMAD, ETF-AP1, SP1-ETS. SP1F-NF1, IKRS-AP2, EGR-SP1, SP1-KLFS, GATA-SP1, CAAT-CAAT, NFKB-SP1, SP1-CAAT.

2.4. Functional analysis of the core SMO promoter

Sequence analysis revealed that the 5'-flanking region of the human SMO gene exhibits a high GC content, and lacks a consensus TATA element. Three potential CpG islands were identified surrounding the SMO gene TSS using the MethPrimer program (Figure 3a). We evaluated different experimental PCR conditions and programs for amplification of the SMO promoter sequence. The GC-rich sequences contained within the SMO gene promoter region were effectively amplified by a touchdown program in presence of 3% DMSO [30,31]. Comparison of the touchdown PCR results with PCR under standard condition (fixed annealing temperature, 35 cycles) program for five primer-pairs corresponding to upstream region of SMO was shown in Figure S2 and Table S1.
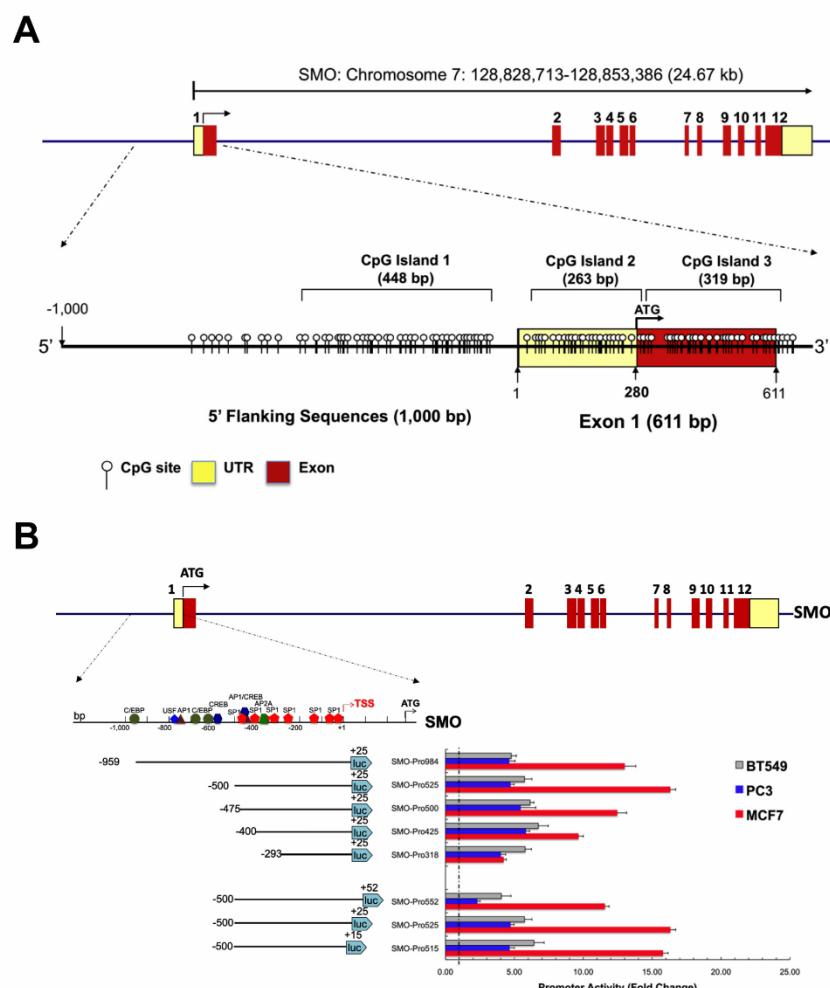
Figure 3. (A) Structure of the SMO gene. A schematic representation of the exon-intron organization and UTR region. 12 exons are indicated by the numbered rectangles. Distribution of

CpG dinucleotides in a 1,611 bp fragment of the SMO gene harboring 1,000 bp 5′ upstream region and full exon 1 is shown. Each vertical line represents a single CpG site. Numbering is relative to the transcriptional start site at exon 1. Transcription orientations are indicated by arrows. (B) Functional localization of the SMO promoter. A schematic of the SMO gene structure is shown above. Twelve SMO exons are indicated by the numbered rectangles. A schematic diagram of the 1,500 bp 5′-flanking region of SMO and serial truncation constructs of the SMO promoter and their corresponding luciferase activities in different cell types are shown. Serial deletions at the 5′ and the 3′ ends of the promoter fragment of SMO are shown on the left. The promoter activities measured after transfection into PC3, BT549 and MCF7 cells are shown on the right. The relative size and position of fragments cloned into the pGL3 vector are indicated by the lines below the schematic, and the numbers in parentheses on either side of each fragment indicate the distance in nucleotides upstream from the SMO start codon of the 5′ and 3′ ends of each fragment. The luciferase activity of the pGL3 constructs is shown as fold-increase of corrected light units relative to an empty pGL3 vector control. Values represent the mean, and error bars indicate the SEM of at least three independent experiments.

The potential promoter region upstream of the SMO gene was analyzed using interspecies comparison with the Genomatix package. Multiple regulatory elements are located surrounding the TSS of SMO and may play a role in the regulation of SMO expression. In order to determine the minimal sequences required for promoter function and identify cis-acting elements controlling SMO promoter activity, a series of truncated luciferase constructs were generated by progressive deletion from the 5′ end of a 984 bp fragment (region from -959 to +25 relative to the TSS), to produce 5 constructs (Figure 3b and Table S1), based on our 5′ truncation analysis result (Figure S3). Plasmids containing SMO gene fragments were transiently transfected into three cancer cell lines (prostate cancer line PC3; breast cancer lines BT549 and MCF7), and the luciferase activities of these constructs were measured.

The highest promoter activity was observed in MCF7 cells, and moderate activity was found in PC3 and BT549 cells. In MCF7 cells, increased promoter activity was detected upon removal of 459 bp of 5′ sequence up to position -500 bp (relative to the transcription initiation site), indicating the presence of negative regulatory element(s) in the region from -959 bp to -500 bp in MCF7 cells. When truncated to -470 bp, the promoter activity returned to the full-length promoter activity, and deletion of additional sequence to either -400 or -293 further reduced promoter activity, suggesting the presence of positive regulatory element (s) in the region -500 bp to -293 bp. In PC3 and BT549 cells, the 5′ truncations had little effect and maximal activity was observed with the PGL3-SMO-400/+25 construct. The PGL3-SMO-500/+25 construct exhibited the highest promoter activity in MCF7 cells therefore, this reporter vector was used for subsequent 3′ deletion analysis (Figure 3b).

To further identify the 3′ boundary of the core promoter, three plasmids were generated sharing the same 5′ boundary at position -500, and variable 3′ ends from +50 to -15. In contrast to the results from the 5′ deletion analysis, luciferase activity with the 3′ deletions showed similar effects in all three cell lines. The promoter activities were comparable between the PGL3-SMO-500/+25 and PGL3-SMO-500 /+15 in the three cell lines, while addition of 25 bp of 3′ sequence to +50 resulted a decreased activity (Figure 3b), indicating an absence of downstream promoter element activity in the SMO promoter.

Taken together, these results demonstrated that the region between the -500 and +15 bp is important for transcriptional activity of the SMO promoter, and both negative and positive regulatory regions can affect the promoter activity of SMO gene depending on cell context.

2.5. Identification of transcription factor binding sites in the SMO gene promoter

EMSA experiments were conducted to investigate the binding of nuclear proteins to the core proximal SMO promoter sequences in nuclear extracts of PC3, BT549 and MCF7 cells. (Figure 4). Six overlapping oligonucleotide probes covering the region between -500 and -357 bp (Fig. 4a) that significantly enhanced promoter activity in MCF7 cells were prepared to investigate their DNA-protein binding activity. Double-stranded DNA probes spanning ~30 bp, covering the regions -500 to -471 (SMO-500P); -472 to -444 (SMO-472P); -454 to -421 (SMO-454P); -422 to -398 (SMO-422P); -400 to -373 (SMO-400P); and -383 to -357 (SMO-383P) were prepared. The SMO-472P, SMO-454P, SMO-422P and SMO-383P probes showed clear binding with nuclear proteins extracted from PC3, BT549 and MCF7 cells, while the SMO-500P and SMO-400P probes did not produce strong complexes with nuclear proteins from any of the cell lines tested (Figure 4b). The SMO-472P, SMO-422P and SMO-383P showed strong binding with nuclear protein from MCF7, but strong binding in the SMO-454P region was observed with nuclear proteins from PC3 and BT549 (Figure 4b).
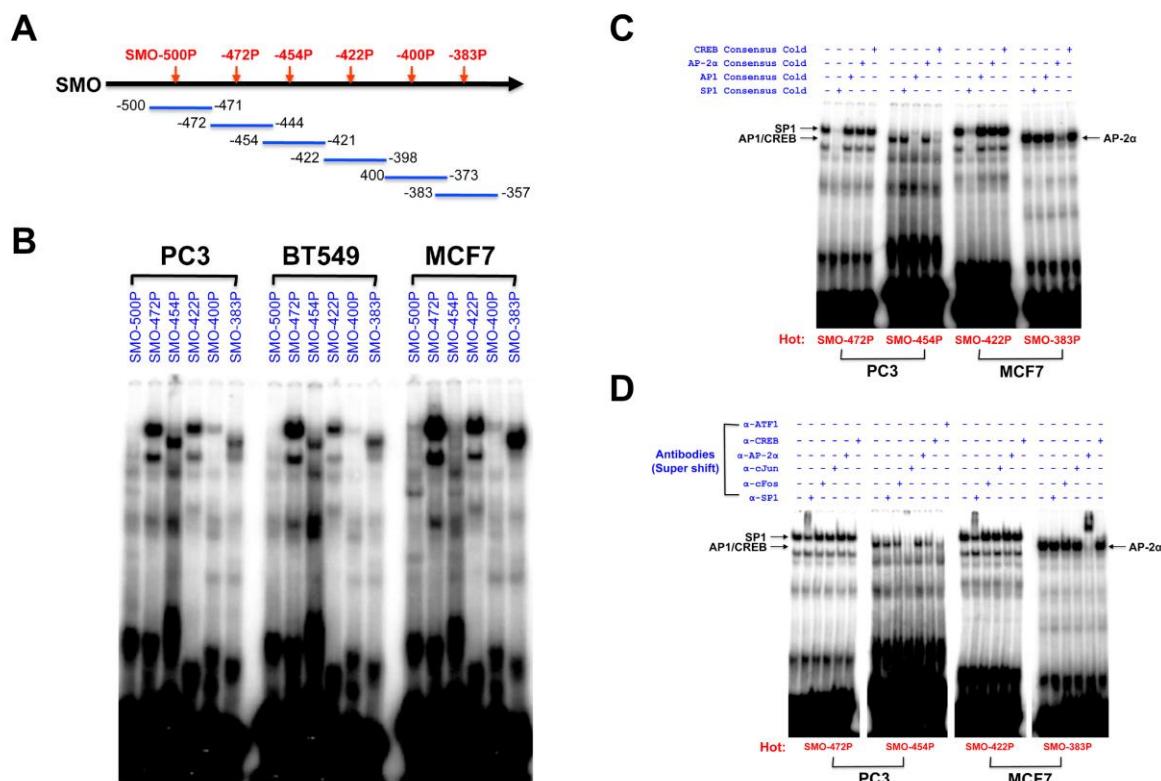
Figure 4. EMSA analysis of the core region of SMO promoter. (A) Binding assay using overlapping probes in the SMO promoter. A schematic illustration of six probes used for EMSA is shown in the upper panel. (B) The 32P-labeled probes were incubated with nuclear extracts from PC3, BT549 and MCF7 cells, respectively. (C) Competition analysis using a 50-fold excess of unlabeled oligonucleotides (cold). The 32P-labeled SMO-472P and SMO-422P were incubated with nuclear extracts from PC3 cells in the presence of a 50-fold excess unlabeled consensus SP1, AP1, AP-2 , and CREB oligonucleotides, respectively. Similarly, the competition analysis of 32P-labeled SMO-454P and SMO-383P were performed with nuclear extracts from MCF7 cells. D. Supershift analysis using specific antibodies. Antibodies (2 g), including anti-SP1, anti-c-Fos, anti-c-Jun, anti-AP-2 , anti-CREB and anti-ATF1, were preincubated with 10 g of nuclear extracts before the

addition of the 32P-labeled probes. The bands of SP1, AP1/CREB, and AP-2 binding are indicated by arrows.

Specificity of binding was tested in competition experiments using excess 9nlabeled oligonucleotide carrying consensus sequences, and specific versus nonspecific antibodies. A competition assay was conducted with four probes, SMO-472P, SMO-454P, SMO-422P and SMO-383P, to confirm DNA-protein complexes. The DNA-protein complexes formed by the SMO-472P and SMO-422P probes were reduced or disappeared completely in the presence of a 50-fold excess of the consensus SP1-binding oligonucleotide, but not in the presence of excess unlabeled consensus oligonucleotides for other TFs (Figure 4c). The major complex formed by the SMO-454P probe disappeared completely in the presence of a 50-fold excess of unlabeled consensus AP1 and was greatly reduced by CREB oligonucleotides, but not by the addition of excess unlabeled consensus SP1 and AP-2 oligonucleotides. The major band observed with the SMO-383P probe was reduced only in the presence of a 50-fold excess of unlabeled consensus AP-2 oligonucleotide (Figure 4c).

A supershift assay with specific antibodies was performed using nuclear extracts from PC3 and MCF7 cells to confirm the identity of the TFs generating the complexes observed in Figure 4c. As shown in Figure 4d, the DNA-protein complexes in the SMO-472P and SMO-422P regions were supershifted by anti-SP1 antibody. The complex in the SMO-383P region was supershifted by anti-AP-2 antibody. The complex in the SMO-454P region was inhibited by anti-c-Jun, anti-ATF1 and anti-CREB antibodies. Therefore SP1, AP1 (c-Jun/ATF1) / CREB, and AP-2 all appear to play a role in SMO gene regulation.

2.6. Hypermethylation of the SMO gene in non-expressing cancer cell lines

We have identified three CpG islands in the 5'-flanking region of the SMO gene, and CpG island 1 is located in the proximal promoter region (Figure 3a). We therefore analyzed the methylation status of CpG island 1 of the SMO gene in 33 cancer cell lines using the MSP and BSP methods (Figure 5 and Table 2). The sequences and locations of the primer pairs used in BSP and MSP are shown in Table S1 and Figure 5a. The methylation frequency was determined using MSP real time PCR. The seven cell lines that did not express SMO (SUM52, MCF10A, MB231, SUN159, MCF7, AGS and HT29) were all hypermethylated in the amplification region (Table 1 and Figure 5b). In contrast, less than 10% methylation was found in the other cell lines which expressed SMO (Table 2). To determine the methylation status in all three CpG islands, we performed BSP for AGS, MCF7, SKBR3 and PC3 cells. Ten clones of the amplified region of the putative SMO promoter for each cancer cell line were sequenced, and methylation status was established for three CpG islands in this region using bisulfite sequencing. The full methylation of all three CpG islands was confirmed in AGS and MCF7 cells, whereas, no methylation was found in the cell lines SKBR3 and

PC3 that express SMO (Figure 5c).
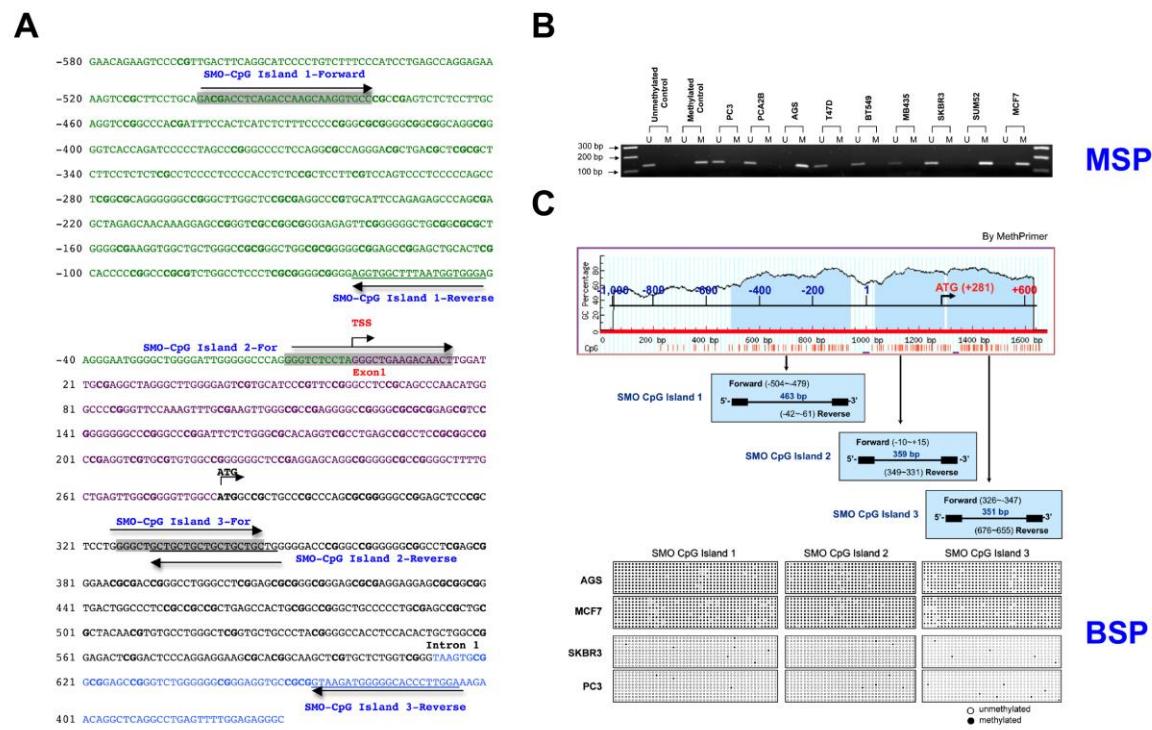
Figure 5. Hypomethylation of the 5'-flanking region of the SMO gene. (A) Bases are numbered relative to the transcription start site at position + 1. CpG sites are shown in bold. The primers used for amplification and sequencing of bisulfate modified DNA were indicated by grey for forward and underline for reverse primer. The long arrows indicate the orientation. (B) Methylation-specific PCR analysis of the SMO upstream regulatory region in methylated/unmethylated controls and nine cancer cell lines. M indicates hypermethylated SMO; U indicates unmethylated SMO. (C) SMO promoter methylation analysis by MethPrimer. Three CpG-rich regions surrounding SMO TSS in a span of the 1,611 bases, and results of bisulfite DNA sequencing were shown.

Table 2.   Epigenetic regulation of SMO mRNA expression in tumor cell lines

Tissue	Cell lines	SMO mRNA Level*	SMO Methylation Frequency (%)
Prostate	PCA2B	19.46	0.47
	VCaP	15.15	1.65
	PC3	11.33	0.61
	DU145	11.14	1.99
	CAHPV10	4.59	2.67
	NCIH660	4.18	1.41
	LNCaP	3.9	1.93
	22RV1	0.18	4.9

Breast	T47D	44.96	4.05
	BT549	36.26	1.47
	MB435	3.56	1.41
	SK-BR3	0.02	2.32
	SUM52	ND	83.57
	MCF10A	ND	98.28
	MB231	ND	91.52
	SUN159	ND	99.6
	MCF7	ND	99.66
Kidney	SN-12C	114.33	0.71
	TK-10	95.49	3.7
	786-0	87.9	0.33
	ACHN	2.54	6.88
Glioblastoma	SF268	61.21	0.36
	SF539	38.5	9.06
	U251	31.36	0.31
	SNB75	27.07	0.77
Ovary	IGROV1	85.11	2.37
	OVCAR4	33.99	0.84
	SKOV3	5.65	3.1
Stomach	AGS	ND	99.36
Skin	SK-MEL2	9.28	3.39
Colon	HT29	ND	96.22
Lung	H322	11	1.98
Myeloma	RPMI8226	0.02	0.57

Abbreviations: ND, not detectable

*Relative mRNA expression = (Target gene / 18s rRNA) x 1,000,000

To confirm the role of methylation in silencing SMO gene expression, the breast cancer cell line MCF7 that lacked SMO gene expression, and the prostate cancer cell line PC3 with moderate SMO expression, were treated with 5-aza-dC for 72 hr. Treatment with 5-aza-dC resulted in expression of SMO in MCF7 cells, however, treatment with 5-aza-dC decreased SMO expression in PC3 cells (Table 3).

Table 3. Demethylation treatment restored SMO mRNA expression in MCF7 cell

Cell line	Control		5-Aza treatment (1 µM)	
	Methylation frequency (%)	mRNA level	Methylation frequency (%)	mRNA level
MCF7	99.6	0	57.9	16.1
PC3	0.6	113.3	1.5	86.5

3. Discussion

The HH pathway drives oncogenesis in many cancers, and strategies targeting this pathway have been developed, most notably through inhibition of SMO which is a key step involved in the regulation of the seven-transmembrane oncoprotein. SMO is able to activate the glioma-associated oncogene (GLI) family of transcription factors, leading to hyperproliferation of epithelial cells [32].

CpG islands represent a common epigenetic element that regulates transcription at many promoters through methylation-induced silencing. In this study, the most striking feature revealed by in silico analysis of the SMO promoter is the abundance of CpG dinucleotides and multiple SP1 binding sites (7xSP1 sites) close to the TSS. SP1 has been widely described as a general transcription factor involved in the transcription of gene promoters that lack a TATA box. CpG-rich promoters bound by DNA sequence-specific transcription factors including SP1 have the highest expression level, and deletion of SP1 binding sites results in significantly decreased promoter activity [33,34].

We have isolated and cloned DNA fragments containing the predicted SMO promoter region, and the strongest promoter activity was identified in the 5'-UTR region around -500/+25 that contains a high GC content, is CpG rich and lacks a canonical TATA box. Evaluation of the predicted SP1 consensus sites revealed that the SMO-472P (-472 to -444 bp) and SMO-422P (-422 to -398 bp) regions are functional in binding to the SP1 transcriptional factor as shown in EMSA/supershift experiments. In addition, inducible TF factors binding to AP1 and AP2 sites were also identified in the promoter region and may play a role in modulating SMO expression.

The epigenetic regulation of SMO transcription was characterized in 33 cancer cell lines. We determined the mRNA expression of the major HH pathway genes, SMO, PTCH, GLI1, GLI2 and GLI3 and found the highest level of mRNA expression was observed in the SMO gene. DNA methylation is a major epigenetic regulatory mechanism of gene expression and is involved in the progression of cancer [35]. The absence of SMO expression in 7 cell lines derived from breast (5 of 9), stomach (1 of 1), and colon (1 of 1) cancer tissues was correlated with a high level of gene methylation. The level of SMO mRNA was negatively correlated with the methylation status of the SMO promoter (Figure 6). In this study, SMO and GLI3 were undetectable in the HT29 colon cancer cell line. This result is consistent with a previous study that showed SMO methylation leads to silencing of GLI3 expression [36]. Moreover, full methylation was confirmed in all three CpG islands in the MCF7 breast cancer cell line, and stomach cancer cell line AGS, and SMO expression in MCF7 cells was restored after 1 μM 5-Aza treatment. The results indicate that DNA methylation of the SMO gene may play an important role in the development of cancer. Cell lines from breast cancer tissue showed the highest methylation frequency, 56% (5/9), whereas 8 prostate cancer cell lines had no detectable methylation in the SMO gene. Whether the degree of SMO methylation

correlates with the tissue specificity remains to be explored.
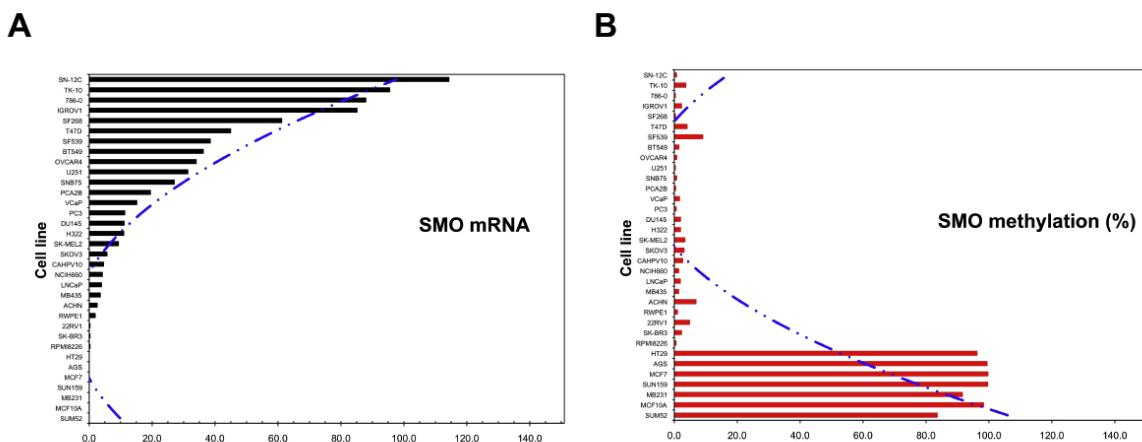
Figure 6. Correlation of mRNA level and methylation frequency of SMO gene in 33 cancer cell lines. (A) mRNA level of SMO in decreasing order. (B) Distribution of methylation frequency (%) in the cell lines by a quantitative MSP method.

The full mechanistic details of HH signal transduction are still under investigation. Abnormal HH activation has been implicated in tumorigenesis in a wide variety of tumors, and SMO and GLI play a critical role in this pathway. GLI2 is suggested to function primarily as a transcriptional activator, and GLI3 as a repressor [4]. A significant positive correlation was identified between the expression of SMO and GLI2 in our study. This is consistent with previous reports that overexpression of either GLI1 or GLI2 leads to tumor development in transgenic mice, suggesting that GLI1 or GLI2 contribute to tumorigenesis by increasing SMO expression that is associated with the development of cancer [17,37]. PTCH1 has emerged as a tumor suppressor gene and developmental regulator. Although the role of GLI3 as a negative regulator of HH signaling is well established in the context of normal development, its role in cancer has largely been ignored [38-40]. In this study, we have demonstrated that GLI3expression is significantly positive correlated with PTCH1 levels (Pearson's correlation=0.532, p=0.001). This result is indirectly supported by the study of Ohba et al. They reported that PTCH1 deficient (PTCH1 +/-) cells showed increase adult bone mass in patients with nevoid basal cell carcinoma syndrome through reduced GLI3 [41]. Moreover, different mutations in GLI3 [42] and PTCH1 haploinsufficiency are associated with distinct autosomal dominant syndromes [43]. These data suggest that the repressive effect of PTCH1 and GLI3 on HH signaling has a crucial role in cancer development.

4. Materials and Methods

4.1. Computational analysis of the SMO gene

Identification and sequence analysis of evolutionary conserved regions (ECRs) of the SMO was performed with the ECR Browser, and the publicly available web-based tool mVista [44] using the MLAGAN algorithm. A search for potential TFBS in the upstream regulatory region of the SMO

gene was performed online at Genomatix using the MatInspector program (http://www.cbrc.jb/research/db/TFSEARCH.html).

4.2. Cell lines and 5-Aza treatment

All cancer cell lines were purchased from American Type Culture Collection (ATCC, Rockville, MD, USA) and grown according to the ATCC instructions. 5-Azacytidine (5-Aza, Sigma-Aldrich) was freshly prepared in PBS before use. A vehicle control consisting of culture medium alone was included in the analysis. MCF7 and PC3 cells were pre-cultured for 24 h, then treated with 1 M 5-Aza for 72 h. Cells were collected by centrifugation, then genomic DNA and RNA were extracted and analyzed.

4.3. Real Time quantitative RT-PCR

Total cellular RNA was isolated and purified by RNeasy columns (QIAGEN Valencia, CA, USA) according to the manufacturer's instructions with on-column and in-solution DNaseI digestion. RNA quality and quantity were determined using Agilent RNA 6000 Nano Chip (Agilent Technologies Inc., CA, USA). cDNA synthesis was carried out using Random Hexamer primer, Taqman Reverse Transcription Reagents kit (Applied Biosystems Foster City, CA, USA).Taqman real time RT-PCR primers and probes for target genes were designed by using the Primer Express software. SMO Fwd: 5'-GAGACTCGGACTCCCAG-3'; Rev: 5'-GTATACGGCACACAGCAG-3' and probe: 5'(FAM)-TCGGGCCTCCGGAAT-(MGB)3'. PTCH1 Fwd: 5'-GCATAGGAGTGGAGTTCA-3'; 5'-CCCTGCGGTTCTTGTC-3' and probe 5'(FAM)-TTGGCCTTTCT-(MGB)3'. GLI1 Fwd: 5'-GTCTCAAACTGCCCAGC-3'; Rev: 5'-CGTTCAAGAGAGACTGGG-3' and Probe: 5'(FAM)-TCCCACACCGGTACCA-(MGB)3'. PTCH2, (Assay ID=Hs01085642_ml), GLI2 (Assay ID=Hs00257977_ml) and GLI3 (Assay ID=Hs00609233_ml). Taqman real time RT-PCR was used to determine the expression profile, with a 18S rRNA plasmid as the standard reference gene using primers Fwd: 5'-CCGAAGCGTTTACTTTGAAAAAA-3'; Rev: 5'-TTCCATTATTCCTAGCTGCGGTAT-3' and probe 5'(VIC)-AGTGTTCAAAGCAGGCC-(MGB) 3' [45]. The PCR reactions were performed in 20μl final volume containing 5 ng of cDNA, 1 x Master Mix (TaqMan Universal PCR Master Mix, ABI, CA), 900nM of each primer and 200nM of each probe, respectively. The thermal cycling condition are 40 cycles of PCR amplification (UNG incubation: 50ºC, 2min; Ampli TaqGold activation: 95ºC, 10 min; denaturation: 95ºC, 15 s; annealing/extension: 60ºC, 1 min) (ABI PRISM 7900HT Sequence Detection System, CA). All assays were performed in triplicate, and each plate contained the same standard and positive quality control sample. For each unknown sample, the copy number of each gene is calculated using linear regression analysis from their respective standard curves (Figure S1). The relative mRNA expression level of target genes was normalized by the following formula: (copy number of target gene) / (copy number of 18S rRNA) x 10,000. The standard curves were generated using a dilution series of plasmids containing SMO, PTCH1, PTCH2, GLI1, GLI2 and GLI3 from full length of cDNA (ATCC, Rockville, MD). The copy number of plasmid cDNA was calculated by optimal density according to the exact molar mass derived from the sequences. Serial dilutions were made to obtain 101 to 107 copies. The slope and intercept were calculated for each run using linear regression analysis of log copy number versus threshold cycle (Ct) value for both target genes and 18S rRNA standard curves [46].

4.4, Bisulfite modification and bisulfite sequencing PCR (BSP)

DNA was extracted using the QIAamp DNA Mini Kit (Qiagen, Valencia, CA). Bisulfite modification of 1 g of genomic DNA was performed with the EpiTect Bisulfite Kit (Qiagen) as described by the manufacturer. Primers for BSP and identification of predicted CpG islands in SMO promoter region were carried out with the assistance of Methyl Primer Express Software v1.0 (Applied Biosystems) and MethPrimer (http://www.urogene.org/methprimer/). The primer

sequences used for methylation analysis are summarized in Table S1. PCR reactions were performed in a volume of 25 µl containing 10 ng of bisulfite-converted DNA and 20 pmol of each primer using Platinum® PCR SuperMix (Invitrogen, Carlsbad, CA). Thermal cycling conditions were 95°C for 2 min, followed by 35 cycles of 95°C for 15 s, 56°C for 30 s, 72°C for 10 s, and final extension at 72°C for 5 min. For sequence analysis, the PCR products were subcloned into a pCR2.1 vector using a TOPO TA Cloning kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. At least twelve clones were sequenced in an ABI 3730 DNA Sequencer (Applied Biosystems) for each cell line tested. Methylation analysis was performed using BiQ Analyzer software [47].

4.5. Methylation-specific PCR (MSP)

The MSP products were 132 bp long. Unmethylated and methylated human DNA was used as a negative and a positive control, respectively (Qiagen). Water blanks and PCR mixtures without template were also used as experimental controls in each assay. The primers were listed in Table S1. The amplification cycles performed were 38 cycles. After PCR, products were separated on a 2% agarose gel, and stained with ethidium bromide. Bisulfite treatment and MS-PCR assays were performed in duplicate for all samples. Each experiment was performed at least three times.

4.6. Real-time quantitative MSP

The bisulfite–converted genomic DNA was amplified using fluorescence-based real-time MSP using FastStart SYBR Green Master Kit (Roche). Methylation of the SMO gene was examined using actin as the internal control for DNA quantification. The beta actin gene contains no CpG dinucleotides and is not affected by DNA methylation status or bisulfite treatment. The primers for quantitative MSP are same as the normal MSP shown in Table 1. Real-time PCR conditions were 95 °C for 10 min followed by 40 cycles of 94 °C for 15 s, 59 °C for 60 s with data acquisition after each cycle. At the end, properties of real-time PCR conditions and amplification products were checked by melting curve analysis. PCRs were done in two replicates of each sample with the 7900HT Fast Real-Time PCR System (Applied Biosystems).

4.7. Touchdown PCR of the SMO promoter region

PCR was carried out in a volume of 50 µl containing 100 ng of genomic DNA, 20 pmol of each primer and 3% Dimethyl Sulfoxide (DMSO, Sigma-Aldrich, St. Louis, MO, USA) using Platinum® PCR SuperMix Kit (Invitrogen, Carlsbad, CA, USA). A modified touchdown PCR was performed with the following cycling conditions: the templates were denatured at 94°C for 3 min, and then 20 cycles composed of 20 s at 95°C, 30 s annealing with a stepwise reduction of annealing temperature from 68°C to 58°C decreasing by 0.5°C every cycle, and an elongation step of 4 min at 72°C. Twenty additional cycles were then performed at 94°C 20 s, 58°C for 40 s, and 72°C for 40 s. The standard PCR program was 35 cycles at 94°C 20 s, 58°C for 40 s, and 72°C for 40 s. All PCR products were analyzed by electrophoresis on a 1.0% agarose gel stained with ethidium bromide.

4.8. Generation of luciferase reporter plasmids

A series of truncated SMO promoter constructs, including 5 deletions from the 5' side and 3 deletions on the 3' side, were created by PCR using the primers shown in Table S1. PCR products were cloned into the TOPO-TA vector, and inserts were excised with SacI and XhoI and cloned into pGL3 (Promega, Madison, WI, USA) to generate constructs in the forward orientation. All subclones were verified by sequencing. Sequence analysis was performed with the Molecular Evolutionary Genetics Analysis (MEGA) software version 7.

4.9. Cell transfection and luciferase assays

Two breast cancer cell lines, MCF7 and BT549, and prostate cancer cell line PC3, were used for the analysis of promoter constructs. The cells were plated at 1x105 cells per well in a 24-well plate the day before transfection and incubated overnight at 37°C in 5% CO2. For each well, 5 µl of HilyMax transfection reagent (Dojindo, Rockville, MD, USA) was diluted in 30 µl of growth medium without serum and incubated at room temperature for 5 min. The DNA mixture containing 2   g of the specific reporter construct plus 2 ng of Renilla luciferase pRL-SV40 control DNA was then added to each well, and incubated at room temperature for 20 min. Luciferase activity was assayed at 48 hr using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, USA) according to the manufacturer's instructions. Measurement of the firefly luciferase activity of the SMO promoter constructs was normalized relative to the activity of the Renilla luciferase produced by the pRLSV40 control vector and each construct was tested in triplicate in at least three independent experiments.

4.10. Electrophoretic mobility shift assays (EMSA)

Nuclear extracts were prepared from PC3, BT549 and MCF7 cells using the CellLytic NuCLEAR extraction kit (Sigma-Aldrich, St. Louis, MO). Six double-stranded DNA oligonucleotide probes corresponding to -500 to -357 bp of SMO promoter were synthesized (Figure 5a, upper panel). Labeling, DNA-protein binding reactions and antibody supershift experiments were performed as previous described [46,48]. For supershift experiments, antibody was added after addition of labeled DNA-probe, and the binding reaction was incubated for additional 20 min at room temperature. For competition analyses, a 50-fold excess of consensus unlabeled-competitor oligonucleotides, SP1, AP1, Ap-2 , and CREB (Santa Cruz Biotechnology, Santa Cruz, CA), were included in the binding reaction.

4.11. Statistical analysis

The correlation of mRNA expression levels of SMO, PTCH1, GLI1, GLI2 and GLI3 were assessed by Pearson's correlation coefficient using GraphPad Prism 7 software. All p values reported were two-tailed, with significance defined as $p < 0.05$.

5. Conclusions

In conclusion, our data provide strong experimental and computational evidence for genetic and epigenetic regulatory mechanisms of the SMO gene. The SMO promoter has been characterized and its major regulatory elements, including multiple CpG islands and SP1 binding sites, were identified. A correlation between SMO/GLI2 and PTCH1/GLI3 expression was observed. Moreover, SMO expression is correlated with the degree of CpG island methylation. Our results reveal a central role for epigenetic regulation of SMO gene transcription that may be exploited for the development of new therapeutic strategies to treat hedgehog-driven tumors.

Supplementary Materials:

Figure S1. Pearson's correlation analysis of gene expression levels of SMO, PTCH, GLI1, GLI2 and GLI3 in cancer cells. Pearson's correlation coefficients were performed to estimate the correlations of mRNA expression levels using the JMP program.

Figure S2. Comparison of PCR conditions to amplify SMO promoter sequence in BT549 DNA. The figure showed PCR products from the five different primer sets used in the 5'-deletion assay to amplify 984bp, 525bp, 500bp, 425bp and 318bp fragments of the SMO promoter (M&M section), respectively. Equal amounts (15 µl) were loaded on a 1.0% agarose gel. Left panel showed PCR results from the touchdown PCR, and right panel showed PCR results from the standard PCR program. Lane M, molecular weight marker (1kb plus DNA ladder).

Figure S3. Identification of SMO promoters by 5' truncation analysis. A schematic representation of the SMO bidirectional element is shown with the potential transcription factor binding sites. Serial deletions at the 5' end

of the promoter fragment of SMO are shown. The luciferase activity of pGL3 constructs containing fragments cloned in either the reverse (left panel) or forward (right panel) orientation is shown as fold-increase of corrected light units relative to an empty pGL3 vector control [48]. Values represent the mean±SEM of at least three independent experiments.

Table S1. Sequences of primers used in this study.

Author Contributions: Conceptualization, H.L., HC.L., N.T., S.A. and M.D.; methodology, H.L., HC.L. A.H. and B.S.; software, HC.L.; validation, H.L. and HC.L.; formal analysis, H.L., and HC.L..; investigation, H.L., HC.L., S.A. and M.D.; resources, H.L. and HC.L.; data curation, H.L., HC.L. and S.A.; writing—original draft preparation, H.L and HC.L.; writing—review and editing, S.A. and M.D.; visualization, H.L. and HC.L.; supervision, S.A. and M.D.; project administration, S.A. and M.D.; funding acquisition, S.A. and M.D. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix

References

1. Raleigh, D.R.; Reiter, J.F. Misactivation of Hedgehog signaling causes inherited and sporadic cancers. J Clin Invest 2019, 129, 465-475, doi:10.1172/JCI120850.

2. Beachy, P.A.; Karhadkar, S.S.; Berman, D.M. Tissue repair and stem cell renewal in carcinogenesis. Nature 2004, 432, 324-331, doi:10.1038/nature03100.

3. Berman, D.M.; Karhadkar, S.S.; Maitra, A.; Montes De Oca, R.; Gerstenblith, M.R.; Briggs, K.; Parker, A.R.; Shimada, Y.; Eshleman, J.R.; Watkins, D.N., et al. Widespread requirement for Hedgehog ligand stimulation in growth of digestive tract tumours. Nature 2003, 425, 846-851, doi:10.1038/nature01972.

4. Ingham, P.W.; McMahon, A.P. Hedgehog signaling in animal development: paradigms and principles. Genes Dev 2001, 15, 3059-3087, doi:10.1101/gad.938601.

5. Zhang, B.; Zhuang, T.; Lin, Q.; Yang, B.; Xu, X.; Xin, G.; Zhu, S.; Wang, G.; Yu, B.; Zhang, T., et al. Patched1-ArhGAP36-PKA-Inversin axis determines the ciliary translocation of Smoothened for Sonic Hedgehog pathway activation. Proc Natl Acad Sci U S A 2019, 116, 874-879, doi:10.1073/pnas.1804042116.

6. Pietrobono, S.; Gagliardi, S.; Stecca, B. Non-canonical Hedgehog Signaling Pathway in Cancer: Activation of GLI Transcription Factors Beyond Smoothened. Front Genet 2019, 10, 556, doi:10.3389/fgene.2019.00556.

7. Price, M.A.; Kalderon, D. Proteolysis of the Hedgehog signaling effector Cubitus interruptus requires phosphorylation by Glycogen Synthase Kinase 3 and Casein Kinase 1. Cell 2002, 108, 823-835, doi:10.1016/s0092-8674(02)00664-5.

8. Vokes, S.A.; Ji, H.; McCuine, S.; Tenzen, T.; Giles, S.; Zhong, S.; Longabaugh, W.J.; Davidson, E.H.; Wong, W.H.; McMahon, A.P. Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. Development 2007, 134, 1977-1989, doi:10.1242/dev.001966.

9. Dean, M. Towards a unified model of tumor suppression: lessons learned from the human patched gene. Biochim Biophys Acta 1997, 1332, M43-52, doi:10.1016/s0304-419x(96)00043-1.

10. Wolf, I.; Bose, S.; Desmond, J.C.; Lin, B.T.; Williamson, E.A.; Karlan, B.Y.; Koeffler, H.P. Unmasking of epigenetically silenced genes reveals DNA promoter methylation and reduced expression of PTCH in breast cancer. Breast Cancer Res Treat 2007, 105, 139-155, doi:10.1007/s10549-006-9440-4.

11. Evangelista, M.; Tian, H.; de Sauvage, F.J. The hedgehog signaling pathway in cancer. Clin Cancer Res 2006, 12, 5924-5928, doi:10.1158/1078-0432.CCR-06-1736.

12. Ohta, M.; Tateishi, K.; Kanai, F.; Watabe, H.; Kondo, S.; Guleng, B.; Tanaka, Y.; Asaoka, Y.; Jazag, A.; Imamura, J., et al. p53-Independent negative regulation of p21/cyclin-dependent kinase-interacting protein 1 by the sonic hedgehog-glioma-associated oncogene 1 pathway in gastric carcinoma cells. Cancer Res 2005, 65, 10822-10829, doi:10.1158/0008-5472.CAN-05-0777.

13. Karhadkar, S.S.; Bova, G.S.; Abdallah, N.; Dhara, S.; Gardner, D.; Maitra, A.; Isaacs, J.T.; Berman, D.M.; Beachy, P.A. Hedgehog signalling in prostate regeneration, neoplasia and metastasis. Nature 2004, 431, 707-712, doi:10.1038/nature02962.

14. Watkins, D.N.; Berman, D.M.; Burkholder, S.G.; Wang, B.; Beachy, P.A.; Baylin, S.B. Hedgehog signalling within airway epithelial progenitors and in small-cell lung cancer. Nature 2003, 422, 313-317, doi:10.1038/nature01493.

15. Thayer, S.P.; di Magliano, M.P.; Heiser, P.W.; Nielsen, C.M.; Roberts, D.J.; Lauwers, G.Y.; Qi, Y.P.; Gysin, S.; Fernandez-del Castillo, C.; Yajnik, V., et al. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. Nature 2003, 425, 851-856, doi:10.1038/nature02009.

16. Berman, D.M.; Karhadkar, S.S.; Hallahan, A.R.; Pritchard, J.I.; Eberhart, C.G.; Watkins, D.N.; Chen, J.K.; Cooper, M.K.; Taipale, J.; Olson, J.M., et al. Medulloblastoma growth inhibition by hedgehog pathway blockade. Science 2002, 297, 1559-1561, doi:10.1126/science.1073733.

17. Espinosa-Bustos, C.; Mella, J.; Soto-Delgado, J.; Salas, C.O. State of the art of Smo antagonists for cancer therapy: advances in the target receptor and new ligand structures. Future Med Chem 2019, 11, 617-638, doi:10.4155/fmc-2018-0497.

18. Lou, H.; Dean, M. Targeted therapy for cancer stem cells: the patched pathway and ABC transporters. Oncogene 2007, 26, 1357-1360, doi:10.1038/sj.onc.1210200.

19. Dean, M. Cancer stem cells: redefining the paradigm of cancer treatment strategies. Mol Interv 2006, 6, 140-148, doi:10.1124/mi.6.3.5.

20. De Smaele, E.; Ferretti, E.; Gulino, A. Vismodegib, a small-molecule inhibitor of the hedgehog pathway for the treatment of advanced cancers. Curr Opin Investig Drugs 2010, 11, 707-718.

21. Burness, C.B. Sonidegib: First Global Approval. Drugs 2015, 75, 1559-1566, doi:10.1007/s40265-015-0458-y.

22. Johannessen, L.; Remsberg, J.; Gaponenko, V.; Adams, K.M.; Barchi, J.J., Jr.; Tarasov, S.G.; Jiang, S.; Tarasova, N.I. Peptide structure stabilization by membrane anchoring and its general applicability to the development of potent cell-permeable inhibitors. Chembiochem 2011, 12, 914-921, doi:10.1002/cbic.201000563.

23. Remsberg, J.R.; Lou, H.; Tarasov, S.G.; Dean, M.; Tarasova, N.I. Structural analogues of smoothened intracellular loops as potent inhibitors of Hedgehog pathway and cancer cell growth. J Med Chem 2007, 50, 4534-4538, doi:10.1021/jm0705657.

24. Du, P.; Ye, H.R.; Gao, J.; Chen, W.; Wang, Z.C.; Jiang, H.H.; Xu, J.; Zhang, J.W.; Zhang, J.C.; Cui, L. Methylation of PTCH1a gene in a subset of gastric cancers. World J Gastroenterol 2009, 15, 3799-3806, doi:10.3748/wjg.15.3799.

25.   Issa, J.P. CpG island methylator phenotype in cancer. Nat Rev Cancer 2004, 4, 988-993, doi:10.1038/nrc1507.

26.   Toyota, M.; Ahuja, N.; Ohe-Toyota, M.; Herman, J.G.; Baylin, S.B.; Issa, J.P. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999, 96, 8681-8686, doi:10.1073/pnas.96.15.8681.

27.   An, N.; Shi, Y.; Ye, P.; Pan, Z.; Long, X. Association Between MGMT Promoter Methylation and Breast Cancer: a Meta-Analysis. Cell Physiol Biochem 2017, 42, 2430-2440, doi:10.1159/000480196.

28.   Feng, L.; Jin, F. Screening of differentially methylated genes in breast cancer and risk model construction based on TCGA database. Oncol Lett 2018, 16, 6407-6416, doi:10.3892/ol.2018.9457.

29.   Peters, I.R.; Helps, C.R.; Hall, E.J.; Day, M.J. Real-time RT-PCR: considerations for efficient and sensitive assay design. J Immunol Methods 2004, 286, 203-217, doi:10.1016/j.jim.2004.01.003.

30.   Lorenz, T.C. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. J Vis Exp 2012, 10.3791/3998, e3998, doi:10.3791/3998.

31.   Jensen, M.A.; Fukushima, M.; Davis, R.W. DMSO and betaine greatly improve amplification of GC-rich constructs in de novo synthesis. PLoS One 2010, 5, e11024, doi:10.1371/journal.pone.0011024.

32.   Yang, C.; Chen, W.; Chen, Y.; Jiang, J. Smoothened transduces Hedgehog signal by forming a complex with Evc/Evc2. Cell Res 2012, 22, 1593-1604, doi:10.1038/cr.2012.134.

33.   Gebhard, C.; Benner, C.; Ehrich, M.; Schwarzfischer, L.; Schilling, E.; Klug, M.; Dietmaier, W.; Thiede, C.; Holler, E.; Andreesen, R., et al. General transcription factor binding at CpG islands in normal cells correlates with resistance to de novo DNA methylation in cancer cells. Cancer Res 2010, 70, 1398-1407, doi:10.1158/0008-5472.CAN-09-3406.

34.   Brandeis, M.; Frank, D.; Keshet, I.; Siegfried, Z.; Mendelsohn, M.; Nemes, A.; Temper, V.; Razin, A.; Cedar, H. Sp1 elements protect a CpG island from de novo methylation. Nature 1994, 371, 435-438, doi:10.1038/371435a0.

35.   Jaenisch, R.; Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 2003, 33 Suppl, 245-254, doi:10.1038/ng1089.

36.   Zhu, Y.; James, R.M.; Peter, A.; Lomas, C.; Cheung, F.; Harrison, D.J.; Bader, S.A. Functional Smoothened is required for expression of GLI3 in colorectal carcinoma cells. Cancer Lett 2004, 207, 205-214, doi:10.1016/j.canlet.2003.10.025.

37.   Rui, Z.; Li-Ying, P.; Jia-Fei, Q.; Ying-Ying, H.; Feng, C.; Tie-Jun, L. Smoothened gene alterations in keratocystic odontogenic tumors. Head Face Med 2014, 10, 36, doi:10.1186/1746-160X-10-36.

38.   Hilton, M.J.; Tu, X.; Cook, J.; Hu, H.; Long, F. Ihh controls cartilage development by antagonizing Gli3, but requires additional effectors to regulate osteoblast and vascular development. Development 2005, 132, 4339-4351, doi:10.1242/dev.02025.

39.   Litingtung, Y.; Dahn, R.D.; Li, Y.; Fallon, J.F.; Chiang, C. Shh and Gli3 are dispensable for limb skeleton formation but regulate digit number and identity. Nature 2002, 418, 979-983, doi:10.1038/nature01033.

40.   Wang, B.; Fallon, J.F.; Beachy, P.A. Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. Cell 2000, 100, 423-434, doi:10.1016/s0092-8674(00)80678-9.

41.   Ohba, S.; Kawaguchi, H.; Kugimiya, F.; Ogasawara, T.; Kawamura, N.; Saito, T.; Ikeda, T.; Fujii, K.; Miyajima, T.; Kuramochi, A., et al. Patched1 haploinsufficiency increases adult bone mass and modulates Gli3 repressor activity. Dev Cell 2008, 14, 689-699, doi:10.1016/j.devcel.2008.03.007.

42. Johnston, J.J.; Olivos-Glander, I.; Killoran, C.; Elson, E.; Turner, J.T.; Peters, K.F.; Abbott, M.H.; Aughton, D.J.; Aylsworth, A.S.; Bamshad, M.J., et al. Molecular and clinical analyses of Greig cephalopolysyndactyly and Pallister-Hall syndromes: robust phenotype prediction from the type and position of GLI3 mutations. Am J Hum Genet 2005, 76, 609-622, doi:10.1086/429346.

43. Gorlin, R.J. Nevoid basal-cell carcinoma syndrome. Medicine (Baltimore) 1987, 66, 98-113, doi:10.1097/00005792-198703000-00002.

44. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: computational tools for comparative genomics. Nucleic Acids Res 2004, 32, W273-279, doi:10.1093/nar/gkh458.

45. Muthukumar, T.; Dadhania, D.; Ding, R.; Snopkowski, C.; Naqvi, R.; Lee, J.B.; Hartono, C.; Li, B.; Sharma, V.K.; Seshan, S.V., et al. Messenger RNA for FOXP3 in the urine of renal-allograft recipients. N Engl J Med 2005, 353, 2342-2351, doi:10.1056/NEJMoa051907.

46. Lou, H.; Yeager, M.; Li, H.; Bosquet, J.G.; Hayes, R.B.; Orr, N.; Yu, K.; Hutchinson, A.; Jacobs, K.B.; Kraft, P., et al. Fine mapping and functional analysis of a common variant in MSMB on chromosome 10q11.2 associated with prostate cancer susceptibility. Proc Natl Acad Sci U S A 2009, 106, 7933-7938, doi:10.1073/pnas.0902104106.

47. Bock, C.; Reither, S.; Mikeska, T.; Paulsen, M.; Walter, J.; Lengauer, T. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics 2005, 21, 4067-4068, doi:10.1093/bioinformatics/bti652.

48. Li, H.; Pascal, V.; Martin, M.P.; Carrington, M.; Anderson, S.K. Genetic control of variegated KIR gene expression: polymorphisms of the bi-directional KIR3DL1 promoter are associated with distinct frequencies of gene expression. PLoS Genet 2008, 4, e1000254, doi:10.1371/journal.pgen.1000254.