Development 134, 1977-1989 (2007) doi:10.1242/dev.001966

Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning

Steven A. Vokes[1], Hongkai Ji[2,3], Scott McCuine[4], Toyoaki Tenzen[1], Shane Giles[4], Sheng Zhong[3,*], William J. R. Longabaugh[5], Eric H. Davidson[6], Wing H. Wong[3] and Andrew P. McMahon[1,7,†]

Sonic hedgehog (Shh) acts as a morphogen to mediate the specification of distinct cell identities in the ventral neural tube through a Gli-mediated (Gli1-3) transcriptional network. Identifying Gli targets in a systematic fashion is central to the understanding of the action of Shh. We examined this issue in differentiating neural progenitors in mouse. An epitope-tagged Gli-activator protein was used to directly isolate cis-regulatory sequences by chromatin immunoprecipitation (ChIP). ChIP products were then used to screen custom genomic tiling arrays of putative Hedgehog (Hh) targets predicted from transcriptional profiling studies, surveying 50-150 kb of non-transcribed sequence for each candidate. In addition to identifying expected Gli-target sites, the data predicted a number of unreported direct targets of Shh action. Transgenic analysis of binding regions in Nkx2.2, Nkx2.1 (Titf1) and Rab34 established these as direct Hh targets. These data also facilitated the generation of an algorithm that improved in silico predictions of Hh target genes. Together, these approaches provide significant new insights into both tissue-specific and general transcriptional targets in a crucial Shh-mediated patterning process.

KEY WORDS: Chromatin immunoprecipitation, Gli, Neural patterning, Sonic hedgehog, Mouse

INTRODUCTION

The Hedgehog (Hh) signaling pathway represents one of approximately six core signaling pathways that dictate development in most bilateria (Davidson and Erwin, 2006; Hooper and Scott, 2005). Hh ligands regulate diverse processes, including morphogen-mediated patterning, cell cycle regulation and cell polarity, operating in a variety of cellular contexts from the forming segments of the early *Drosophila* embryo to the digit-forming mesenchyme of the vertebrate limb (McMahon et al., 2003). Genetic and biochemical studies indicate that all Hh signaling is likely to be mediated by Ci/Gli zinc finger domain-containing transcription regulators, corresponding in vertebrates to Gli1-3, through a Gli-consensus binding sequence, TGGGTGGTC (reviewed by Hooper and Scott, 2005; Jacob and Briscoe, 2003; Ruiz i Altaba et al., 2003). In the absence of Hh signaling in Hh-responsive tissues in mice, Gli3 and, to a far lesser extent, Gli2 (Pan et al., 2006) undergo proteosome-mediated carboxyl cleavage to repressor forms (Gli^{REP}) that silences one set of Hh targets. Hh signaling at low levels counters cleavage, leading to de-repression of targets. High levels of Hh signaling are essential for direct activation of a second group of targets. All three Gli members can function as activators (Gli^{ACT}), although only Gli1, itself a direct target of Hh signaling, is thought to act exclusively as an activator (Bai et al., 2004; Dai et al., 1999; Wang et al., 2000). The primary transcriptional activator appears to be Gli2; however, substitution of Gli1 into the Gli2 locus rescues the Gli2 null phenotype. Thus, Gli1 and Gli2 activator forms have similar properties (Bai and Joyner, 2001). While not comprehensively explored, evidence from *Drosophila* suggests that Gli3 repressor and activator forms appear to bind a common set of target sites (Muller and Basler, 2000).

One crucial role of sonic hedgehog (Shh) lies in patterning of the vertebrate neural tube. Here, Shh secreted from the midline notochord acts as a morphogen to specify five classes of ventral neuronal progenitors in a concentration-dependent fashion; from dorsal to ventral these are progenitors for V0, V1, V2 interneuron pools, motoneurons (MN) and V3 interneurons. In addition, the highest levels of Shh signaling at the ventral midline induce a second ventral midline domain of Shh-expressing cells, the floor plate (FP), immediately ventromedial to V3 progenitors (reviewed by Briscoe and Ericson, 2001; Jessell, 2000; McMahon et al., 2003). Genetic studies have demonstrated that whereas V0, V1, V2 and MN identities can be specified in the absence of Hh signaling if Gli^{REP} is removed, the most ventral identities, V3 and FP, require Hh signaling and Gli^{ACT} forms (Bai et al., 2004). Recent work indicates that relatively small differences in the ratio of these forms dictate distinct transcriptional outputs (Stamataki et al., 2005), suggesting a fine-tuned response to Gli regulators in the cis-regulatory regions of target genes expressed at distinct Hh thresholds. Clearly, testing this or any alternative model requires a thorough understanding of the cis-regulatory mechanisms in direct target genes that dictate Gli-mediated regulation.

Here we have initiated a systematic effort to define sets of genes that are directly regulated by Gli activator. Gli1-directed chromatin immunoprecipitation (ChIP) products isolated during the course of Hh-mediated patterning of neural progenitors were used to interrogate custom genomic arrays, successfully identifying and mapping a large number of novel in vivo target sites. Bioinformatic analysis and expression studies in cell culture and transgenic mouse embryos validated these data and provide novel insights into the regulation of both tissue-specific and more general components of a Hh transcriptional response. Further, the analysis of in vivo targets facilitated the development of a novel algorithm for Gli-target site prediction. Based on our data, we present a model for ventral

[1]Department of Molecular and Cellular Biology and [2]Department of Statistics, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA. [3]Department of Statistics, Stanford University, Sequoia Hall, 390 Serra Mall, Stanford, CA, 94305, USA. [4]Agilent Technologies, 245 First Street, Suite 105, Cambridge, MA 02142, USA. [5]Institute for Systems Biology, Seattle, WA, 98103, USA. [6]Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA. [7]Harvard Stem Cell Institute, 16 Divinity Avenue, Cambridge, MA 02138, USA.

*Present address: Bioengineering Department, University of Illinois at Urbana-Champaign, 1304 West Springfield Avenue, Urbana, IL 61801, USA
[†]Author for correspondence (e-mail: mcmahon@mcb.harvard.edu)

Accepted 5 March 2007

neuronal specification in which Gli-activator and Gli-repressor forms differ substantially in their selection of binding sites. This uneven competition drives the transcriptional interpretation of the morphogen gradient.

MATERIALS AND METHODS

Generation of ES cells and transgenics

A full-length mouse Gli1 construct was cloned with an in-frame C-terminal 3XFLAG tag (Sigma). The tagged Gli induced luciferase expression 111±6-fold compared to 254±27-fold with the untagged form. This was cloned upstream of an IRES2, allowing bicistronic expression of the Venus YFP protein (Nagai et al., 2002), and the module was cloned into the pBigT shuttle vector, then into Rosa26PAS (Srinivas et al., 2001). The linearized construct was electroporated into YFP3-1 (Rosa26YFP/β-gal) embryonic stem (ES) cells (Mao et al., 2005) and neomycin-resistant colonies that passed initial visual screens (loss of β-gal or YFP expression) were assayed by Southern blot. We used one ES cell line, $Gli1^{FLAG}$, for all further experiments. To obtain a constitutively active Rosa Gli1 construct (constitutive $Gli1^{FLAG}$), the conditional ES cells were grown for several passages in 1 μm 4-OH Tamoxifen, and then cloned by serial dilution (three independent lines were isolated and we utilized one for all subsequent experiments). For leptomycin B treatment, ES cells were incubated for 5 hours with 5 nM leptomycin B.

The pHSP68lacZ2XINS vector was constructed by PCR amplifying the 2X Chick β-globin insulator motif from XB3+Insulator with primers containing *Asc*1 sites and inserting this into the *Asc*1 site of pHSP68lacZ (Sasaki and Hogan, 1996). With the exception of the *Nkx2.1* (also known as *Titf1* – Mouse Genome Informatics) peak, where only one copy of the enhancer was used, transgenic constructs were generated by cloning four copies of test DNA upstream of the minimal promoter in pHSP68lacZ2XINS. The sequence coordinates of the *Ptch1* peak2 enhancer and the *Nkx2.2* (*Nkx2-2*) Gli enhancers are described in the text. In addition, a 575 bp sequence (chr12:53,245,800-53,246,374) encompassing the Nkx2.1 peak and a 291 bp sequence (chr11:77915085-77915375) spanning the Rab34 peak were used to generate the respective constructs. To construct mutated Gli sites, we changed the site 5′-TGGGTGGTC-3′ to 5′-TCCCACGTC-3′ (Nkx^{M}-HSP68lacZ), a mutant form that removes an invariant G essential for Gli binding and several other residues that contact the zinc finger motif of Gli1 (Pavletich and Pabo, 1993).

Embryoid bodies (EBs) were generated using previously described procedures (Wichterle et al., 2002). After 2 days, the media was changed to DFNB containing 500 nM retinoic acid and, when applicable, 1 μM Hh agonist (HH-Ag) (Frank-Kamenetsky et al., 2002). EBs were grown for an additional 3 days to induce neural progenitor stages, at which point cultures were harvested for ChIP.

Immunostaining with Pax6, Nkx6.1 (Nkx6-1), Nkx2.2 and FoxA2 was performed using previously published methods (Wijgerde et al., 2002). For western blots, samples were run out on 4-12% BisTris gradient gels and protein levels were normalized using levels of β-gal. The M2 FLAG antibody (Sigma) or anti-β-galactosidase (Promega Z3781) were each used at a 1:2000 dilution. For immunostains, YFP was detected using an anti-GFP at 1:2000 (Abcam ab290).

Chromatin immunoprecipitation and hybridization

We modified a published protocol (Odom et al., 2004) to reduce the number of input cells to approximately 2×10^6 cells per precipitation using a FLAG M2 mouse monoclonal antibody (Sigma). After isolating chromatin, we incubated it with previously prepared M2-coated magnetic beads. To prepare beads, magnetic sheep anti-mouse IgG beads (Dynal) were incubated overnight with 0.8 μg M2 Mouse monoclonal anti-Flag antibody (Sigma) and 10 μl of beads per precipitation. The following day, the beads were rinsed and added to chromatin and incubated overnight at 4°C. Samples were then rinsed five times with RIPA buffer, and the antibody was stripped from the beads by incubating in 1% SDS at 65°C for 15 minutes and cross-linking was reversed by incubating overnight at 65°C. The next day, samples were sequentially treated with RNAse A and Proteinase K, phenol:chloroform extracted, ethanol precipitated with 20 μg glycogen and resuspended in 60 μl 10 mM Tris pH 8.0. ChIP-enriched DNA samples were amplified before hybridization to the tiling array. For this, DNA samples were blunted using T4 DNA polymerase (50 μl of each ChIPed sample with 0.6 units of T4 DNA polymerase in a 110 μl volume at 12°C for 15 minutes), phenol:chloroform extracted with 10 μg glycogen, ethanol precipitated and resuspended in 25 μl water. This material was then blunt-ligated to unidirectional linkers (Ren et al., 2000) using 15 μM annealed linkers and a Quick Ligation Kit (New England Biolabs) in a 21 μl volume at 12°C overnight. The reaction was then ethanol precipitated, resuspended in 25 μl water and PCR amplified using 1 μM primer oJW102 with an annealing temperature of 60°C and a 1 minute extension at 72°C for 26 or 29 cycles. To ensure that samples were uniformly amplified, we re-assayed the samples by qPCR.

For each ChIP, 2 μg amplified ChIP sample or Input control was labeled with Cy3 or Cy5-dUTP (Perkin Elmer) using the Bioprime Array CGH kit (Invitrogen) and competitively hybridized using 5 μg labeled material per channel with the Oligo CGH Hybridization Kit (Agilent Technologies) on a custom tiled 44K array from Agilent Technologies (see Table S1 in the supplementary material). Complete details for array construction and all raw array data may be obtained from GEO (GEO #GSE5683). Three independent biological samples were employed in the analysis. Within each sample, dye-swap experiments were performed in duplicate for a total of four technical replicates per biological sample.

The hybridizations were done in Agilent SureHyb hybridization chambers, and in an Agilent rotating hyb oven at a speed of 10 rpm for 72 hours at 65°C. Each hybridization was performed in duplicate and by exchanging labeled dyes (dye-swapping) for a total of four technical replicates per biological sample. Samples were washed sequentially for 5 minutes at room temperature using Oligo aCHG Wash Buffer 1 (Agilent Technologies), 1 minute at 37°C in Oligo aCGH wash buffer 2 (Agilent Technologies), 1 minute in Acetonitrile at room temperature and 30 seconds in Stabilization and Drying Solution (Agilent Technologies). Finished arrays were scanned using an Agilent scanner at a 10 μm resolution with both channels at a PMT setting of 100. Array images were extracted using Agilent's Feature Extraction Software (version 8.5.1.1) using the standard CGH protocol included with the software. The settings included spatial detrending of the extracted array data and a linear median normalization. Log fold changes were obtained for each probe, and the four technical replicates within each biological replicate were averaged. The correlation coefficient between any two technical replicates using the same dye labeling was r=0.948 (s.d.=0.020), and between two technical replicates that use different dye labeling was r=0.773 (s.d.=0.095). In contrast, the correlation between two biological replicates was r=0.068 (s.d.=0.075). T-scores were obtained using TileMap in the moving average mode (Ji and Wong, 2005). All coordinates in this manuscript refer to the Build 34 assembly of the mouse genome.

Mapping of Gli sites, determination of conserved regions and motif identification

Potential binding regions (see Table S2 in the supplementary material) were extended 200 bp from both ends and repeats were masked (A. F. A. Smit, R. Hubley and P. Green, RepeatMasker at http://repeatmasker.org). We then mapped the Gli consensus-binding pattern TGGGTGGTC (Kinzler and Vogelstein, 1990), allowing only one bp mismatch. Candidate binding regions were binned into groups of size ten according to their initial rank, and the number of Gli sites (n_1) and the total number of surveyed non-repeat base pairs (n_2) were counted for each tier and reported in Table S3 in the supplementary material. To determine the background level of Gli occurrence, the Gli consensus sites were also mapped to all tiled regions present on the array, yielding 1264 sites in a total of 5,608,267 bp non-repeat sequences. Gli-binding-site enrichment was then computed for each tier by comparing its site density to the site density of control regions, i.e. $r_1=(n_1/n_2)_{target}/(n_1/n_2)_{control}$.

In a separate approach, we examined the degree of cross-species conservation among binding sites. The multiple species alignment among mouse, rat, human, dog and zebrafish were used to compute a conservation score for each position of mouse genome. Sites whose mean conservation score is among the top 10% of the genome were called conserved sites. The

Table 1. Summary statistics of ChIP peaks

Rank	Gene	Peak	Chr.	Start	End	qPCR	Orientation	Gli concensus	Gli matrix
1	*Ptch1*	2	13	60949397	60950371	34	TSS –1.1k	–137, 219	–154, –137
2	*Nkx2-2*	1	2	146644626	146645570	29.2	TSS –1.9k	–51	–98, –51
3	*Ptch1*	3	13	60951377	60952953	45.3	TSS –3.7k	78, 437	–760, 78
4	*Nkx2-9*	2	12	53347310	53348354	16.2	TSS –8.7k	–25	–222, –25, 45, 289
5	*Ptch1*	5	13	60956284	60956777	65	TSS –7.9k	–78	–78
6	*Nkx2-9*	1	12	53340519	53341015	37.7	TSS –1.9k	–119	–119
7	*Rab34*	1	11	77915220	77915560	28.1	Intron 1,2	–71, 264	–106, –71
8	*Ptch2*	1	4	116057320	116057946	5.6	Intron 2	–106	–106
9	*Ptch1*	1	13	60945578	60945972	30	Intron 2	101	
10	*Nkx 2-1*	1	12	53245748	53246351	9.1	TES +11.5k	152	–355, 152
11	*Gli1*	1	10	127076571	127078504	7.2/4.7/12.5	Promoter, Intron 1,2	–115, 543, 1422, 1580	–115, 1580
12	*Hhip*	1	8	79270092	79270492	5.6	Intron 1	–51	–184, –51
13	*FoxA2*	1	2	147493701	147494245	6.4	TES +5.8k	–148, 331	–148, 331
14	*Ptch2*	2	4	116054385	116054954	48.3	TSS –0.3k	–136, 61	–136, 65
15	*Cart1*	1	10	102925949	102926177	–	Intron 3	NA	NA
16	*Ptch1*	4	13	60953950	60955072	9	TSS –6.3k	–36	380
17	*Prdx2*	1	8	84215218	84215559	1.4	TSS –22.5k	NA	–78
18	*Cart1*	2	10	102938108	102938308	–	Intron 2	–75	–75
19	*Ptch1*	6	13	61010241	61010714	2.8	TSS –61.9k	NA	NA
20	*Hhip*	2	8	79268419	79268867	8.9	Intron 1	NA	NA
21	*Flrt3*	1	2	140191468	140191692	–	TSS –6.3k	NA	NA
22	*Pax9*	1	12	53443055	53443311	–	TES +5.3k6	–35	NA
23	*Ncor2*	1	5	124340743	124341009	–	TSS –0.4k	60, 335	54, 60,199
24	*Zic3*	1	X	52789693	52789967	–	TES –0.8k (3′ UTR)	NA	NA
25	*Hand2*	1	8	56373391	56373692	2.7	TES +10.3k	–154	–181, –154

First-tier peaks (ranked 1-13) include all known Gli enhancer sites (gene name is underlined in 'Gene' column). Negative values indicate regions upstream (5′) of the specified reference site in the orientation column, and positive values are downstream (3′) of the site. Gli consensus and matrix sites were mapped to the defined peak regions and are displayed as values relative to the peak center (see Table S2 in the supplementary material). Our Gli-site mapping tolerates only a 1 bp mismatch from the TGGGTGGTC consensus. Sites that are identified by both consensus and matrix mapping are underlined. Chr., chromosome; NA, not applicable (no Gli sites present); TSS, transcriptional start site; TES, transcriptional end site; UTR, untranslated region.

number of conserved Gli sites (n_3) and the total number of surveyed base pairs that are conserved (n_4) were counted for each tier and control regions. We computed the enrichment of conserved Gli sites in conserved regions as:

$$(1)\ r_2=(n_3/n_4)_{target}/(n_3/n_4)_{control};$$

and

$$(2)\ r_3=(n_3/n_2)_{target}/(n_3/n_2)_{control}.$$

r_2 measures the extent to which Gli sites are enriched in conserved regions, and r_3 measures the combined gain we can obtain using both binding-sequence specificity and cross-species conservation information.

To perform de novo motif discovery, Gibbs motif sampler was applied to the top 13 tier 1 peaks after extending each defined peak 100 bp from both ends. The Gli matrix identified by motif analysis were visualized using WebLogo (Crooks et al., 2004). To map Gli consensus sites, the core-binding pattern TGGGTGGTC was used to scan genomic sequences; up to one mismatch was allowed. We computed cross-species conservation scores using mouse, rat, human, dog and zebrafish alignments. Motif sites that reside within the top 10% most conserved regions in the mouse genome are defined as phylogenetically conserved sites.

In silico predictions of Gli target genes and enhancers

To generate an experimental dataset for predictions, retinoic acid-treated EBs were grown in the presence or absence of HH-Ag. We did not observe induction of ventral Hh markers in RA-treated constitutive $Gli1^{FLAG}$ EBs and used these for the control, baseline set. To pick genes that were upregulated by Hh signaling, a modified t-statistic (Ji and Wong, 2005) was computed for each gene to test if its expression level is higher in Hh-induced samples compared with non-induced samples and the top 365 probe sets corresponding to 249 non-redundant genes were used in further analyses (see Table S5 in the supplementary material). We then identified their orthologs from human, dog, cow, chicken and zebrafish. Each gene was extended 50 kb upstream from transcription start and 50 kb downstream from transcription end, and predictions were made based on these sequences.

In order to make EEL predictions, EEL was run using the same parameters as Hallikas et al. (Hallikas et al., 2006). In addition to the previous 107 motif-position-specific weight matrices, we also included the Gli matrix recovered in this current study (see Fig. S2 in the supplementary material). Pairwise alignments between mouse-human, mouse-dog, mouse-cow, mouse-chicken and mouse-zebrafish were generated. As EEL did not provide a way to combine these pairwise alignments into multiple alignments and to rank predicted enhancers accordingly, we decided to use mouse-human alignments as the basis for making our predictions. Each predicted enhancer was then annotated to reflect whether it was also found in other pairwise alignments. Predictions shorter than 2000 bp, with an EEL score ⩾100 and a combined Gli score ⩾25 were picked, retained and ranked by their combined Gli scores.

To perform MCA, we first collected conserved, non-coding genomic segments from the mouse genome. Cross-species conservation scores were computed based on percent identities measured in a 50 bp sliding window using multiz mouse, rat, human, dog and zebrafish alignments (downloaded from http://genome.ucsc.edu). The scores ranged from 0 to 255, with 255 corresponding to the most conserved status. We collected all continuous segments with a score ⩾40 (top 10%), discarding segments shorter than 50 bp. These first-stage segments were joined to form second-stage segments if their end-to-end distance (gap) was ⩽50 bp. The second-stage segments were filtered further to remove segments <200 bp or if they did not contain at least one first-stage segment >100 bp. The remaining segments were termed third-stage segments, comprising the final collection of conserved genomic regions.

To measure Gli enrichment, we mapped the Gli matrix recovered from the ChIP to all third-stage conserved genomic regions. The Gli matrix was compared to a third-order background Markov model, and a likelihood ratio ⩾500 was used as a cutoff to define Gli sites. The occurrence of Gli sites in conserved genomic segments was then modeled as Poisson processes. If a segment is a Gli-binding region (the alternative hypothesis H_1), the occurrence rate of Gli sites was assumed to be λ_1 (site/bp) and each Gli site was assumed to be generated from the Gli-binding matrix. In contrast, if a segment is not a Gli-binding region (the null hypothesis H_0), the occurrence rate of Gli sites was assumed to be λ_0 and all such sites were assumed to be

Table 2. Global versus tissue-restricted enhancers

Region	Coordinates	Size (bp)	Conserved Gli sites	*Gli1* induction	St. dev.
Ptch1 peak 1	chr13:60,945,834-60,946,202	369	chr13:60946069-60946077(–)	13.5	2.3
Ptch1 peak 2	chr13:60949285-60949464	180	chr13:60949424-60949435(–)	11.2	3.7
Ptch1 peak 3	chr13:60952107-60952356	250	chr13:60952273-60952284(+)	3.3	0.8
Ptch1 peak 4	chr13:60,954,478-60,955,071	594	chr13:60954710-60954718(–) /// chr13:60954791-60954799(–)	3.8	0.4
Ptch1 peak 5	chr13:60956197-60956431	235	chr13:60956316-60956327(+)	2.6	0.5
Ptch1 peak 6	chr13:61,010,209-61,010,707	499	–	3.1	1.3
Ptch2 peak 1	chr4:116,057,081-116,057,672	592	chr4:116057540-116057551(+)	7	2.1
Ptch2 peak 2	chr4:116,054,618-116,055,117	500	chr4:116054800-116054808(+) /// chr4:116054812-116054823(+) /// chr4:116054906-116054914(+)	129	24
Hhip1 peak 1	chr8:79,269,855-79,270,597	743	chr8:79270163-79270171(+) /// chr8:79270296-79270304(–)	7.7	2.2
Hhip1 peak 2	chr8:79,268,170-79,268,855	686	–	1.8	0.7
Nkx2.2	chr2:146645052-146645290	239	chr2:146645185-146645196(+)	3.2	1
Negative control	chr12:53255352-53255521	170	–	2.6	0.6
Rab34	chr11:77915199-77915431	233	–	6.2	1.1
Empty vector	–	–	–	2.1	0.5

Conserved Gli motifs in ChIPed regions near predicted global responders to Hh signaling were used to obtain putative Gli enhancer elements that were cloned into a minimal promoter luciferase vector and transfected into murine NIH3T3 fibroblasts in the presence or absence of Gli1. We detect no specific induction for the neural-specific Nkx2.2 peak, suggesting that this assay only identifies global responders of Hh signaling. chr, chromosome; St. dev., standard deviation.

false sites and generated from the background Markov model. The rate λ_0 was estimated based on the mapping results in general conserved regions (λ_0=0.0002/bp), and λ_1 was estimated using conserved regions covered by the top 25 peaks (λ_1=0.0015/bp). For each segment, the log likelihood ratio between H_1 and H_0 was computed as:

$$S = n \times \log_{10}(\lambda_1/\lambda_0) - (\lambda_1-\lambda_0) \times l \times \log_{10}(e) + \log_{10}(L_1) - \log_{10}(L_0).$$

Here, l is the length of a conserved segment, n is the number of Gli sites in the segment obtained by matrix mapping, L_1 is the probability to generate the sites according to Gli matrix, and L_0 is the probability to generate the sites by background Markov model. Finally, S was used to measure the Gli enrichment and rank conserved segments.

We also mapped Gli consensus TGGGTGGTC (allowing ⩽1 mismatch) to the third-stage conserved genomic regions. All regions that contain at least one Gli consensus site were selected as a potential Gli cis-regulatory module (CRM). The Gli CRMs are then ranked by Gli enrichment score S. For each gene, all Gli CRMs located within introns, UTRs, 50 kb upstream regions (from transcription start) and 50 kb downstream regions (from transcription end) were collected and the enrichment scores were added together to derive the combined Gli-binding strength for that gene.

Cell culture

Gli luciferase assays in murine NIH3T3 cells used methods described previously (Nybakken et al., 2005). We used 25 ng Gli1 cloned into pCIG (Megason and McMahon, 2002) or empty pCIG vector for controls. Candidate enhancers (Table 2) were PCR-amplified from genomic DNA and cloned into the pGL3-Promoter vector (Promega).

Transcriptional profiling

To generate the list of Gli target candidate genes, a variety of mouse stages and Hh-pathway mutants (*Ptch1*, *Smo* and *Shh*) conditions were analyzed by transcriptional profiling (T.T. and A.P.M., unpublished). Briefly, for each combination, three different samples were profiled using a combination of Affymetrix mouse U74Av2 and MOE430 A&B arrays. The screen included 6-8-somite embryo (E8.5), 10-13-somite embryo (E8.75), head (E10.5) and limb (E10.5). Data were analyzed by dChip (Li and Wong, 2001) and PowerExpress (H.J. and W.H.W., unpublished) to identify and rank putative Hh targets. The EB transcriptional profiling screen was conducted by screening quadruplicate samples of neuralized EBs treated with Hh agonist (Frank-Kamenetsky et al., 2002) and 500 nM retinoic acid (under identical conditions to the ChIPs) versus EBs treated with 500 nM retinoic acid alone. Samples were profiled using the Affymetrix MOE4302.0 arrays and results were analyzed using dChip (GEO #GSE4936).

RESULTS

Gli1 ES cell line and differentiation into neuralized embryoid bodies

Our analysis focused on Gli1, the only Gli member that appears to function exclusively as an activator in Hh signaling. In order to circumvent the lack of useful Gli antibodies, we generated a biologically active C-terminal FLAG-tagged Gli1 construct. Gene targeting introduced this into the constitutive *ROSA26* locus such that Cre-mediated excision of a selection cassette was required to activate $Gli1^{FLAG}$ production (see Fig. S1A in the supplementary material). The Cre-dependent inducibility of the $Gli1^{FLAG}$ construct in ES cells was confirmed by immunostaining and western blot analysis (see Fig. S1B,D in the supplementary material) and a clonal line was established in which $Gli1^{FLAG}$ was constitutively produced (see Fig. S1C in the supplementary material). As previously reported, $Gli1^{FLAG}$ was predominately restricted to the cytoplasm but accumulated rapidly in the nucleus when nuclear export was inhibited by the addition of leptomycin B (Kogerman et al., 1999) (see Fig. S1C in the supplementary material).

As a strategy for isolating Gli1 targets in the ventral neural tube, we focused on the demonstrated ability of Shh to recapitulate its ventral neuralizing activity in the specification of ES cells differentiated to form neuralized EBs (Wichterle et al., 2002). In this assay, addition of Shh, or small molecule agonists that activate the Shh pathway, led to a quantitative ventralization of neural progenitors closely mirroring the neural patterning activity of Shh in the ventral neural tube. In the absence of Hh agonist (Hh-Ag), the $Gli1^{FLAG}$ line was not sufficient to specify $Nkx6.1^{+}$- (V2, MN, V3, FP progenitors), $Nkx2.2^{+}$- (V3 progenitors) or $FoxA2^{+}$- (FP) producing cell types (Fig. 1B-D,J-L). The only differential response detected between the induced $Gli1^{FLAG}$ line and control, parental EBs was a large decrease in the number of $Pax6^{high}$ cells, although the total number of Pax6-producing cells was not altered (70% σ=10% of cells are $Pax6^{high}$ in control EBs compared with 2% σ=4% for $Gli1^{FLAG}$ EBs; Fig. 1A,I,Q); Pax6 repression has been shown to be Shh-dependent (Ericson et al., 1997). In contrast to this weak response, $Gli1^{FLAG}$ resulted in a strong synergy with Hh-Ag (Frank-Kamenetsky et al., 2002). At a dose of 1 μM, EBs were only weakly ventralized ($Nkx6.1^{+}$, $Nkx2.2^{-}$, $FoxA2^{-}$). In contrast, when $Gli1^{FLAG}$ was present, an identical dose of Hh-Ag resulted in the

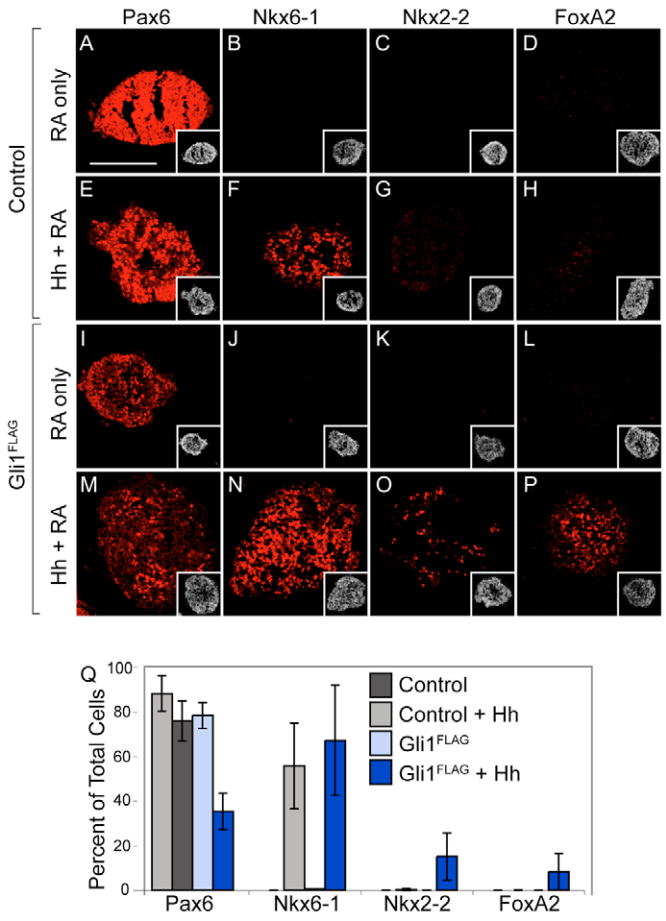

Fig. 1. EBs containing Gli1^{FLAG} exhibit an amplified response to Hh stimulation. (**A-P**) Immunostaining of EBs to detect neural progenitor types in mouse. While both control (A-H) and Gli1^{FLAG} samples (I-P) respond to Hh signaling by activating Nkx6.1 expression (F,N), only the latter samples contain Nkx2.2^{+} and FoxA2^{+} cells on Hh-Ag treatment (O,P). (**Q**) Distinct neural progenitors demarcated by expression of progenitor-type-specific markers were counted and represented as a percentage of all DAPI-positive cells in three independent EBs. The error bars represent the standard deviation. Scale bar: 100 μm.

induction of ventralmost neural (Nkx2.2^{+}, V3 progenitor) and FP (FoxA2^{+}) identities (compare Fig. 1F-H with N-P), as well as a substantial reduction in Pax6^{+} cells consistent with Nkx2.2 repression of Pax6 (Fig. 1E,M) (Ericson et al., 1997). Nkx2.2^{+} and FoxA2^{+} cells comprised approximately 10% each of the EB-derived population. The requirement of Shh for Gli1 activation – even when under control of an ectopic promoter – was previously noted by Bai and Joyner (Bai and Joyner, 2001).

Chromatin immunoprecipitation and statistical validation of putative Gli-binding sites

In an independent study, the developing neural tube and several additional Hh target populations in the early mouse embryo were transcriptionally profiled in an attempt to identify possible targets (positive and negative) of Hh signaling (Tenzen et al., 2006) (T.T. and A.P.M., unpublished). We ranked these genes using PowerExpress (http://biogibbs.stanford.edu/~jihk/CisGenome/microarray.htm) (H.J. and W.H.W., unpublished) and chose the top 122 genes to place on the custom tiling array. For each gene, this method uses the TileMap probe level hierarchical model to compute a posterior probability that the gene has a positive or negative Hh-responsive expression pattern. Among the top targets on this list were several general Hh pathway components, including *Ptch1*, *Gli1*, *Hhip1* and *Gli3*, and previously reported direct neural (*FoxA2*) (Sasaki et al., 1997) and mesodermal (*Myf5*) (Teboul et al., 2003) targets of Hh signaling. A total of 122 genes were selected, together with approximately ten additional genes drawn from the literature to generate arrays for ChIP-mediated identification of Gli1 targets (see Table S1 in the supplementary material). The inclusion of non-neural targets provides an important control for context-dependent regulation in these studies. Custom arrays were generated after repeat masking in which genomic sequence encompassing 25 to 75 kb 5′ and 3′ of the transcriptional start site (TSS) was represented by a 60-mer DNA probe at a spacing of one 60-mer every 125 bp for a total of approximately 5.6 mb of surveyed genomic sequence. Because of the large variability in the size of a given transcriptional unit, probes covered an extensive region 5′ to the TSS, but a variable extent of intronic and/or other 3′ regions.

To verify Gli1^{FLAG}-specific enrichment of expected cis-regulatory regions in Hh-Ag-induced EBs, we adapted a standard ChIP protocol to perform ChIP on 2×10^{6} cells (Odom et al., 2004). Chromatin extracts were incubated with anti-FLAG antibody and following IP, the enrichment of known Gli targets in IP fractions was examined by quantitative RT-PCR (qPCR). qPCR confirmed specific enrichment of Gli-binding sites in several known target regions, including the *FoxA2* FP enhancer (Sasaki et al., 1997), a conserved cis-regulatory region upstream of *Ptch1* previously associated with Shh-dependent regulation in human *Ptch1* (Agren et al., 2004), and *Gli1* itself (Dai et al., 1999). *Gli1* expression is absolutely dependent on prior Hh signaling (Bai et al., 2004).

Having verified the ChIP procedure, we screened the custom tiling array. The regions significantly enriched in the IP fraction were ranked by a T-score, identifying 47 peaks having a false discovery rate of ≤25% (see Table S2 in the supplementary material). As the top 13 ranked sites contained all previously known Gli targets known to be expressed in the neural tube (Table 2, underlined), we used these regions to determine if we could uncover any enriched motif that might represent a transcription-factor-binding site without any prior knowledge of that site. Employing a Gibbs motif sampler (Lawrence et al., 1993; Liu, 1994), we obtained two distinct motifs. The first motif (Fig. 2I) was very similar to the Gli-binding motif (TGGGTGGTC; Fig. 2J; see Fig. S2 in the supplementary material) determined by oligonucleotide binding to recombinant proteins (Hallikas et al., 2006; Kinzler and Vogelstein, 1990; Vortkamp et al., 1995), a confirmation that our strategy specifically isolates sequences directly bound by Gli1^{FLAG}. Further, the analysis predicted nucleotide preferences extending beyond the core Gli consensus region (Fig. 2I). A second, lower scoring motif with a GC-rich pattern was the only other significantly enriched sequence; this is not enriched in conserved regions and is also present in multiple, unrelated ChIP datasets, suggesting that it is non-specific (data not shown).

In an initial effort to determine the biological significance of these regions, we used the enrichment and phylogenetic conservation of Gli motif sites as a measurement of the likelihood that a given peak represents specific Gli1 binding. The ranked sites were grouped into bins of ten, the Gli consensus motif TGGGTGGTC (≤1 mismatch) was enriched in peak regions with high T-scores, when compared

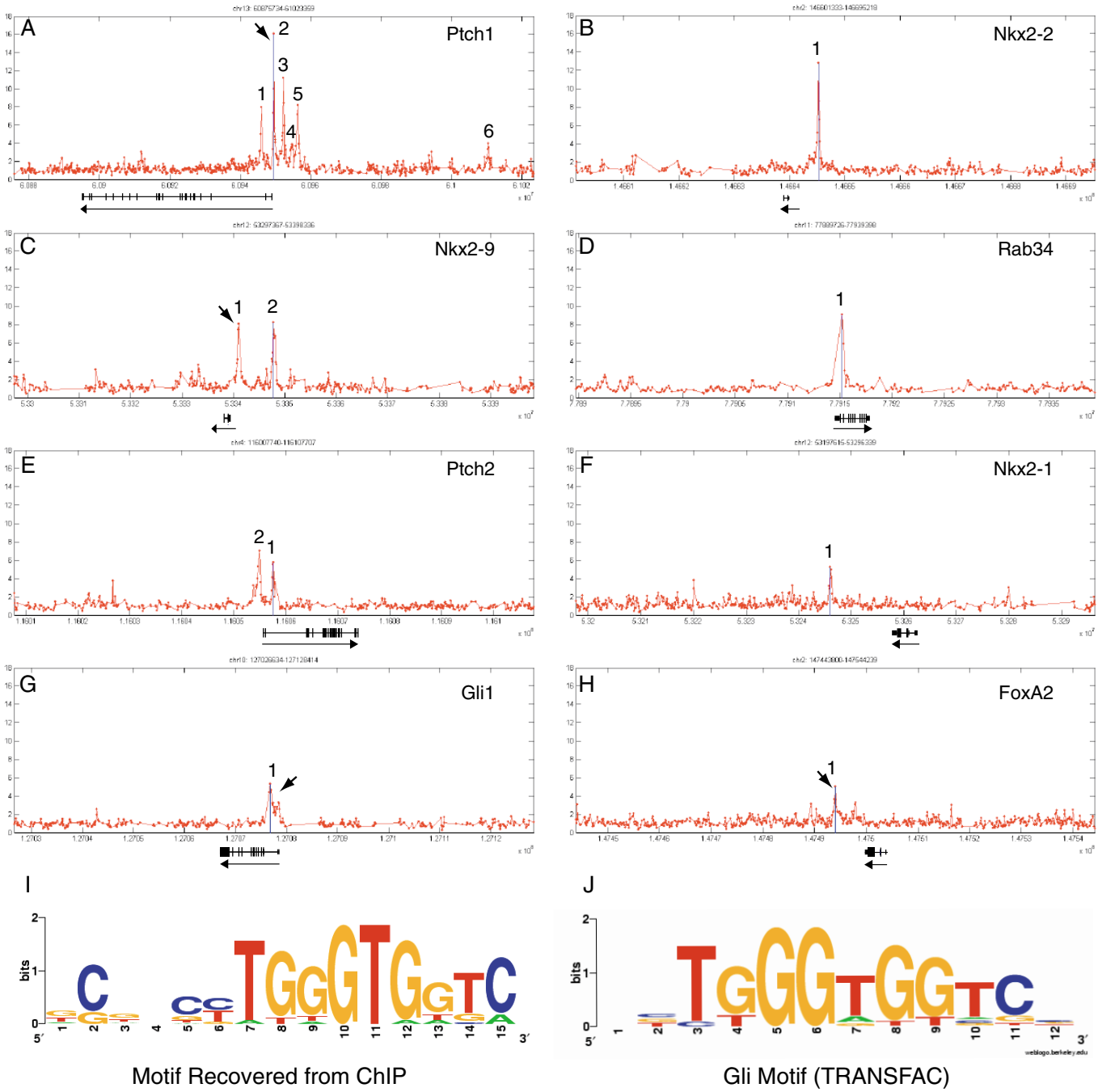

Fig. 2. Chromatin immunoprecipitation of neuralized murine EBs. (**A-H**) Selected positive targets showing the mean fold enrichment of $Gli1^{FLAG}$ ChIPs representing three biological replicates, each with four technical replicates. The plots show the mean fold enrichment of ChIPed sequence versus input. Approximately half of the peaks lie outside the proximal promoter region. Multiple peaks are numbered to correspond to the peaks listed in Table 1 and in Table S2 in the supplementary material; arrows indicate previously described Gli regulatory regions. Below each graph, the position of exons (rectangles) and the direction transcription (arrows) are shown relative to the ChIPed region. (**I**,**J**) De novo motif analysis recovers a consensus site (I) similar to a previously described Gli1 consensus sequence (J) (Transfac # M01037) (I).

with control regions; however, we observed a distinct drop in enrichment from about 5-fold in the first two bins to 3.5-fold in the third bin (see Table S3 in the supplementary material). Of the five motif sites detected in the third bin, four of them occurred in regions ranked 21-25, and only one in regions ranked 26-30. Consequently, we focused on the top group of 25 predictions in subsequent analyses. Fig. 2 shows examples of the data for a subset of genes studied here. These 25 regions contained 28 Gli consensus motif sites in a total of 23,656 bp of non-repeat sequences. The width of the peak itself is predicted to depend both on the number of Gli sites within a particular region and the length of chromatin fragments generated on DNA shearing. The first, second (median) and third quantiles of region length were 295, 474 and 707 bp, respectively. Many, but not all, peaks are centered about a strong Gli consensus

sequence. In total, we observed a 5.25-fold enrichment of predicted Gli target sites within peaks when compared with all regions tiled on the array.

If we were to assume that the top 25 ChIP peaks represented all the Gli-binding events on the array, only 28/1264 Gli consensus sites (2.22%) were bound by Gli (some ChIP peaks had multiple consensus sites, while others did not contain any consensus sites). Among the 1264 consensus sites tiled in the array, 299 were phylogenetically conserved, and 20 of these (6.69%) were associated with Gli binding. These regions were associated with 16 genes, indicating that a significant fraction of the targets contained multiple Gli-binding modules. Interestingly, only seven of the regions identified (28%) were within the 5 kb proximal promoter region. Remaining sites were located in the first or second intron at varying distances from the TSS (24%), at a range greater than 5 kb 5′ of the TSS (24%), after the end of transcription (TES) (16%) or elsewhere within the transcript (8%) (Table 1). These results highlight the problem of local promoter analyses for large-scale mining of transcriptional regulatory mechanisms in the mammalian genome.

To verify that ChIP-chip screening reflects enrichment of a given region of DNA before amplification, we assayed most of the top 25 regions by qPCR using unamplified ChIP material and input chromatin (pre-ChIP baseline) that was a technical replicate of one of the amplified samples. When compared with baseline negative control values (see Table S4 in the supplementary material), all of the peaks present in the top half showed significantly higher qPCR signal in the ChIPed samples. In contrast, this was true for only half of the assayed peaks in the second half of the ranking (3/6; Table 1). Some of these peaks (i.e. *Hhip1* peak2) lacked a Gli matrix site but still exhibited a very significant enrichment by qPCR (Table 1), suggesting that other mechanisms, not simply direct DNA binding, might bring Gli factors to their target sites (see Discussion). As an additional specificity control, analysis of ChIP products in the parental ES cell line before $Gli1^{FLAG}$ activation failed to show any enrichment, indicating that the ChIP results were indeed $Gli1^{FLAG}$-dependent (data not shown).

Identification of neural Gli enhancers

We next examined this list for the presence of known Gli-binding sites that mapped within reported cis-regulatory regions in *Ptch1* (Agren et al., 2004; Hallikas et al., 2006), *Gli1* (Dai et al., 1999), *Nkx2-9* (Santagati et al., 2003) and *FoxA2* (Sasaki et al., 1997); each of these predicted target sites was identified (arrows in Fig. 2A,G,C,H, respectively). To the best of our knowledge, these represent all published, mapped Gli-binding sites in targets expected to be expressed within the EB-generated ventral neural target population. Based on the presence of these previously validated sites, we further divided the ranked list into a first tier of 13 regions that included all the previously characterized Gli enhancer sites and a second tier of 12 regions that represented additional peaks of high statistical quality (Table 1, Table S2 in the supplementary material, and data below). As expected from our previous transcriptional profiling data, the transcripts adjacent to these peaks were upregulated in $Ptch1^{lacZ/lacZ}$ embryos (Goodrich et al., 1997) and downregulated in $Smo^{-/-}$ embryos (Zhang et al., 2001), where Hh signaling was enhanced or absent, respectively (see Fig. S3 in the supplementary material). In addition to the one previously reported Gli enhancer in *Nkx2.9* (Santagati et al., 2003) (Fig. 2C, peak 1), we detected an additional site (Fig. 2C, peak 2) approximately 8.7 kb 5′ to the TSS. In human *Ptch1*, a single 2981 bp enhancer/promoter fragment has been reported from studies in NIH3T3 and HEK293T cells; two conserved, predicted Gli-binding sites are predicted in this region, one of which has been shown to be functional (Agren et al., 2004). The homologous region was detected in our analysis (peak 2, arrow in Fig. 2A). We also identified five additional peaks, including three that were highly enriched (Fig. 2A). One of these, peak 5, lies immediately upstream of an uncharacterized alternatively spliced full-length form of *Ptch1*. While human *Ptch1* contains three alternatively spliced full-length transcripts (Agren et al., 2004), there are only two full-length forms represented among mouse expressed sequence tags (ESTs). A peak immediately upstream of *Gli1* is unusually broad, possibly reflecting an extended region of *Gli1* that also binds Gli3 (Fig. 2G) (Dai et al., 1999). We also detected strong peaks in *Ptch2* (Fig. 2E, two peaks), *Hhip1* (data not shown, two peaks), *Nkx2.2* (Fig. 2B, one peak), *Nkx2.1* (Fig. 2F, one peak) and *Rab34* (Fig. 2D, one peak); with the exception of *Nkx2.2* (Lei et al., 2006), none of these putative Gli-regulatory regions have been previously reported (Fig. 2). *Ptch2* and *Hhip1* are members of the Hh pathway that have been reported to show Hh-dependent regulation (Chuang and McMahon, 1999; Motoyama et al., 1998); *Nkx2.2* encodes a transcriptional determinant of the Shh-dependent V3 interneurons (Briscoe et al., 1999); *Nkx2.1* encodes a Shh-dependent transcriptional regulator within ventral neural populations of the forebrain (Pabst et al., 2000); and *Rab34* encodes a small GTP-binding protein not previously associated with Hh signaling that is predicted from our transcriptional array data to be a rather general target of Hh regulation in several tissues.

Differences between generic and tissue-restricted Gli enhancers

The genes identified as targets in our analysis exhibited either a global (multiple Hh-responsive tissues) or tissue-restricted pattern of expression (see Fig. S3 in the supplementary material). To test Gli-mediated regulation in a putative cis-regulatory enhancer region, we identified 12 candidate enhancers (Table 2) and asked if these could respond to Gli1 stimulation. These assays were performed in the presence and absence of Gli1 using murine NIH3T3 cells, a fibroblast line that has been used to define several cis-regulatory regions, including a *Ptch1* enhancer region (*Ptch1*, peak 2) identified in this screen (Agren et al., 2004). We focused our analysis of candidate sites among global Hh responders (*Ptch1*, *Ptch2*, *Hhip1* and *Rab34*) and compared it with the putative regulatory region in *Nkx2.2*, the expression of which has only been shown to require Hh signaling in the context of neural progenitor specification (see Fig. S3B-B″ in the supplementary material). As expected, we did not detect any specific Gli-mediated induction for the neural-specific *Nkx2.2* candidate enhancer in the reporter assay (Table 2). Nor did we observe reporter activation for either of the peaks in *Ptch1* (Fig. 2A, peak 6) or *Hhip1* (peak 2, data not shown) that failed to show a consensus Gli motif. However, we observed a robust induction with the Gli site in *Ptch1* peak 2 (Fig. 2C), as well as in previously unknown sites in *Ptch1* Intron2 (Fig. 2C, peak 1), and *Ptch2* (Fig. 2E, peak 2). A more modest response was detected in *Ptch2* peak 1 (Fig. 2E), *Hhip1* peak 1 (data not shown) and *Rab34* (Fig. 2D). Taken together, these data (Table 2) indicate that many Gli-responsive elements are confirmed in this assay; those that are not may reflect the absence of context-dependent factors (see Discussion).

To further explore the properties of a subset of these putative enhancers in more detail, we selected four of the regions for more extensive transgenic analysis. We chose a 180 bp Ptch1 peak 2

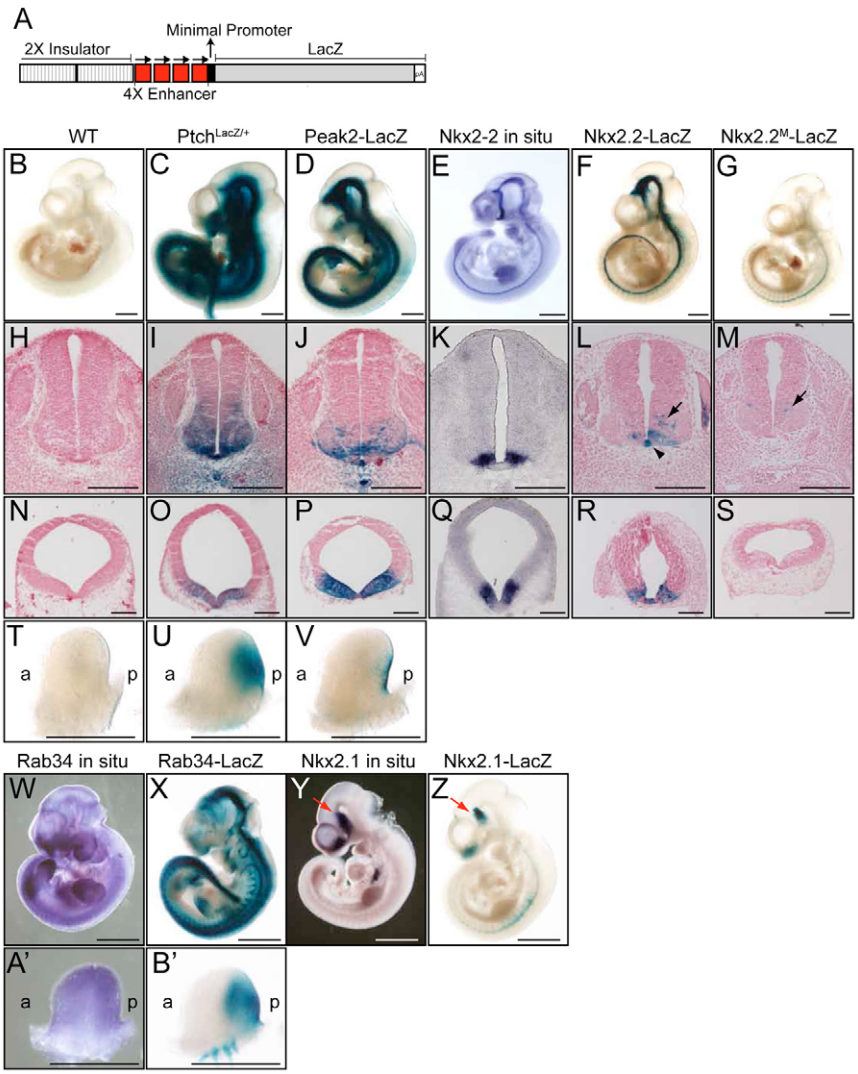

Fig. 3. Transgenic validation of Gli-binding regions demonstrates Gli-dependent expression in Hh target tissue. Selected candidate enhancers (*Ptch1* peak2 and *Nkx2.2* peak1, *Rab34* peak1 and *Nkx2.1* peak1) were cloned into minimal promoter *lacZ* reporter construct (**A**) to generate transgenic embryos and assayed at embryonic day 10.5. (**B-G**,**W-Z**) The embryo in panel E is cleared in benzyl alcohol and benzyl benzoate; all other whole mounts are in 80% glycerol. Histological sections through the spinal cord (forelimb level) (**H-M**) and brain (**N-S**) of transgenic and control embryos. Forelimbs (**T-V**,**A′**,**B′**) show are dissected from the corresponding whole mount. Negative control embryos and sections are shown in B,H,N,T. When compared with *Ptch*$^{lacZ/+}$ embryos, *Ptch1* peak2-*LacZ* transgenics express a subset of the normal domain β-gal activity (compare C,I,O,U with D,J,P,V), lacking expression in the limb bud mesenchyme (U,V). In situ hybridizations of *Nkx2.2* (E,K,Q), transgenic Nkx2.2-*lacZ* embryos (F,L,R) or transgenics with the Gli-binding site mutated (G,M,S). In situ hybridizations of *Rab34* (W,A′) and transgenic Rab34-*lacZ* embryos (X,B′). In situ hybridizations of *Nkx2.1* (Y) and transgenic Nkx2.1-*lacZ* embryo (Z). The arrows in X and Z indicate the domain of expression within the ventral diencephalon. Unlike the other transgenics, Nkx2.1-*lacZ* is driven by only one copy of the enhancer. All limb specimens are oriented with anterior to the left and distal up. Scale bars: 200 μm in H-S; 1 mm in A-G,T-Z,A′,B′.

enhancer region (Fig. 2A; Table 2) representing a global target gene. As tissue-specific candidates, we chose a 239 bp region (a region non-responsive in the NIH3T3 assay) representing the single, Gli target prediction within a peak located approximately 1.9 kb 5′ of the TSS in Nkx2.2, a transcriptional regulator with neurally restricted activity associated with V3 interneuron progenitor specification (Fig. 2B; Table 2). Very recently, this site was shown to be Gli responsive (Lei et al., 2006). In addition, we selected a 575 bp region encompassing the single peak in Nkx2.1 11.5 kb downstream of the TES (Fig. 2F) of this forebrain restricted gene (see Fig. S3F in the supplementary material). Finally, we selected a 291 bp region around the single, novel peak in Rab34 intron 1 (see above, Fig. 2D).

Three of the four independent *Ptch1* Peak2-*lacZ* transgenic founder lines exhibited similar β-gal activity. We focused on one of these lines for more extensive analysis, comparing the activity with that of *Ptch1*$^{lacZ/+}$ reporter embryos in which a *lacZ* cassette had been knocked into the *Ptch1* locus (Goodrich et al., 1997). *Ptch1* Peak2-*lacZ* transgenics expressed β-gal in the ventral forebrain, midbrain, hindbrain, spinal cord and in the notochord in a similar, although more ventrally restricted and mosaic, pattern within the developing CNS compared with *Ptch1*$^{lacZ/+}$ embryos (Fig. 3C,I,O with D,J,P). However, the absence of reporter expression in the branchial arches and most notably in the limb mesenchyme indicates that other cis-regulatory regions are essential for these components of the 'generic' response to Shh signaling (Fig. 3U,V). Several expected aspects of Hh regulation were observed at later stages in Shh and Ihh target fields (e.g. whisker and chondrocytes respectively, data not shown). In summary, *Ptch1* peak 2 is sufficient for the correct spatial and temporal readout of Hh signaling in some but not all Hh target tissues. Further, its activity in the neural tube is consistent with its identification in the neuralized EB patterning assay.

Nkx2.2-lacZ transgenic embryos were analyzed directly at E10.5; nine of 15 showed detectable β-gal activity; of these, seven had ventral-specific staining in the presumptive brain and spinal cord that clearly encompassed the normal *Nkx2.2* domain (Fig. 3E,F,K,L,Q,R). Examination of sections indicated two ectopic regions of transgenic activity when compared to *Nkx2.2* expression in the neural tube at this stage (Fig. 3E,K,Q). One was in the floor plate (arrowhead in Fig. 3L), the other dorsal to the normal *Nkx2.2* domain (arrow in Fig. 3L). To determine whether the observed expression was Gli-dependent, the single predicted Gli site was mutated, and expression of the resulting Nkx2.2^{M} *lacZ* transgene was analyzed in G0 embryos at E10.5. In contrast to the previous results, only four out of 11 transgenic embryos showed any β-gal activity, and activity was generally weaker and more restricted. Expression in the normal Nkx2.2 domain (in both the brain and

spinal cord) and ectopic expression in the floor plate were entirely dependent upon the Gli-binding region (Fig. 3G,M,S); the weak ectopic, ventrolateral spinal cord domain, however, was Gli-site-independent (arrowhead in Fig. 3M). In summary, this peak region encodes a CRM that regulates Gli-dependent expression of *Nkx2.2* in a domain that encompasses its native expression domain. However, Gli-site dependent expression in the FP region indicates that additional cis-regulatory elements are required to repress *Nkx2.2* in the floor plate; *Nkx2.2* is first activated within ventral midline cells and is co-expressed together with the FP determinant FoxA2 at the outset of FP formation (Jeong and McMahon, 2005). In addition, ectopic ventrolateral Gli-independent activity suggests that additional regulatory regions normally silence non-Hh dependent activation of Nkx2.2 outside its normal domain. Previous data indicate that *Pax6*, and potentially, *Tcf4* play an important role in the ventral restriction of Nkx2.2 (Ericson et al., 1997; Lei et al., 2006).

Nkx2.1 exhibits neural, Hh-dependent expression in a domain in the ventral diencephalon and an additional domain in the ventral telencephalon. Of the two transgenics we obtained, one clearly showed β-gal activity specifically within the expected ventral diencephalic domain (arrows in Fig. 2Y,Z). The absence of expression within the ventral telencephalon suggests that additional sequences are required for expression in this region. Given that the in vitro assay system reflects more posterior neural progenitors, the more anterior progenitors representing the telencephalon were not expected.

Rab34 has a broad expression pattern that would not normally be suggestive of a Hh target gene (Fig. 3W). Nonetheless, transgenic embryos expressed a restricted subset of this expression that correlates well with Hh-responsive regions – 11/17 transgenics expressed β-gal and 10/11 were in a similar, restricted pattern (Fig. 3X). For example, expression of *Rab34* in the limb bud is diffuse, but β-gal driven by the predicted Gli target region was sharply restricted to the posterior, Hh-responsive portion of the limb bud in 7/10 transgenics (Fig. 3A′,B′). This expression seemed to commence after the onset of Hh expression, as the three transgenics without limb bud expression were slightly younger and only the oldest transgenics exhibited expression within the later arising hindlimb.

Gli target specificity is determined by the nuclear availability of Gli1

After characterizing these enhancers, we asked if Gli1 could bind to targets in the absence of gene expression. As shown in Fig. 1, Nkx2.2, FoxA2 and other markers of Hh-dependent ventral neuronal cells were not expressed in RA-treated EBs in the absence of Hh agonist. We examined the degree of Gli1 binding to the enhancer sites of several of these tissue-restricted genes as well as the degree of Gli1 binding to several enhancers that have more global response using the standard ChIP assay. As an additional experiment, we also asked whether we could facilitate Gli1 binding to targets by sequestering it in the nucleus by treatment with leptomycin B immediately prior to harvesting.

Our experiments demonstrate that tissue-specific Gli enhancers generally contain little or no Gli1 binding in RA-only treated EBs. For example, the Nkx2.2 enhancer exhibited 6.9-fold enrichment in ChIPs from Hh-stimulated EBs where gene expression occurs, but was enriched only 1.3-fold in RA-only samples; the Nkx2.9 enhancer was enriched 10-fold in Hh-stimulated versus 1.7-fold in the RA-only samples. Similar trends were seen for Nkx2.1 and FoxA2 (data not shown). In contrast, global enhancers were enriched by Gli1 in both the presence and absence of Hh stimulation. Here, Ptch1 peak 2 was enriched in both RA-only samples (14.5-fold) and in Hh-stimulated EBs (25.2-fold). Enrichment of Gli binding in the absence of Hh stimulation was also seen for the other global enhancers Ptch1 peak1, Ptch2peak2, Rab34 and Gli1 (data not shown). Thus, Gli1 is able to bind to global enhancers in the absence of Hh stimulation, but is unable to bind efficiently to tissue-specific enhancers in the absence of corresponding gene expression. Surprisingly, when RA-only EBs were treated with 5 nM leptomycin B (Lepto) for 4 hours before ChIP to cause nuclear accumulation of Gli1 (see Fig. S1C in the supplementary material), they exhibited marked increases in the levels of binding to tissue-specific enhancers. Gli1 enrichment of the Nkx2.2 enhancers was 3.3-fold in RA+Lepto samples compared with 1.3-fold in RA-only samples; similarly, the Nkx2.9 enhancer was enriched 9.4-fold in RA+Lepto-treated samples compared with 1.7 fold in RA-only samples. Leptomycin treatment caused similar increases in enrichment for Nkx2.1 and FoxA2, while maintaining the levels of enrichment seen previously for global enhancers in RA-only samples (data not shown). These data argue that Gli1 nuclear concentration is a major determinant of enhancer-binding specificity.

Table 3. Predicted Gli target genes in the ventral neural tube

	EEL predictions		MCA predictions	
Rank	EEL	qPCR	MCA	qPCR
1	*Ptch1*	+	*Ncam1*	–
2	*BC017647*	–	*Ptch1*	+
3	*Kif5a*	–	*Ctnna2*	–
4	*Ntn1*	–	*Gpc3*	+
5	*Ngfr*	–	*Robo2*	+
6	*4831426I19Rik*	+	*Nkx2-9*	+
7	*Abcb8*	–	*D230005D02Rik*	–
8	*Olfm1*	–	*Ebf3*	–
9	*Nkx2-9*	+	*Ebf1*	+
10	*2600003E23Rik*	–	*Ascl1*	–
11	*Phox2a*	–	*FoxA2*	+
12	*Rad541*	–	*3732412D22Rik*	NA
13	*Hoxb5*	–	*Lrrtm1*	NA
14	*Snx26*	–	*Ssh3*	NA
15	*1810059G22Rik*	–	*Hoxb8*	–
16	*FoxA2*	+	*Gpm6b*	–
17	*Mtap1b*	–	*Trpm3*	–
18	*Cmkor1*	NA	*Neurog2*	–
19	*Tgfbr3*	NA	*Ret*	+
20	*5730411O18Rik*	NA	*Pappa*	NA
	Total positive	4/17	Total positive	7/16

Predictions, derived using a pool of Hh-upregulated candidates obtained from transcriptional profiling of embryoid bodies, were ranked according to either EEL or MCA and candidate enhancers were tested by qPCR. Underlined genes were also detected in the preceding ChIP-chip studies. Columns marked NA indicate that the cis-regulatory modules were not tested for that gene. EEL, enhancer element locater; MCA, module cluster analysis; qPCR, quantitative real-time polymerase chain reaction.

In silico prediction of Gli targets in the ventral neural tube

Next we sought to incorporate information from the Gli target identification to develop a general, testable method for predicting Gli targets in a tissue-specific fashion. We generated a pool of candidate targets by transcriptional profiling of EBs in the presence and absence of Hh signaling and selected 249 genes upregulated in response to Hh-Ag treatment, a group that includes relevant Gli targets in the ventral neural tube (see Table S5 in the supplementary material).

We then applied two distinct methods: one a recently published algorithm, Enhancer Element Locater (EEL) (Hallikas et al., 2006); and the other a novel method, Module Cluster Analysis (MCA). Our ChIP-chip data indicated that several Gli targets contain multiple peaks that presumably correspond to multiple CRMs (e.g. *Ptch1*, *Nkx2.9*). In MCA, Gli-binding affinities from multiple CRMs were combined, and the combined signal was used as a predictor of Gli target genes (see Materials and methods).

When EEL was applied to the 249 genes upregulated in the Hh-treated EBs, the top 20 predicted enhancers, ranked by their combined Gli-binding strength, included five regions identified in our ChIP screens. We performed qPCR on the remaining 15 predictions and verified one additional peak corresponding to an EST (see Table S6 in the supplementary material), resulting in an effective prediction rate of 1/15 enhancer sites.

When MCA was applied on a genome-wide basis, *Ptch1* and *Hhip1* were both present in the top 10 genes (see Table S7 in the supplementary material), suggesting that the combined Gli-binding strength serves as a good predictor of Gli target genes. We intersected this genome-wide dataset with the Hh-EB upregulated gene list to predict Gli target genes specifically within the ventral neural tube. We then ranked all predicted CRMs that were associated with the top 20 neural Gli target genes and tested the top 20 individual CRMs by qPCR on ChIPed material. Within these top-ranking CRMs, 45% (9/20) were indeed Gli-activator targets by this assay (see Table S8 in the supplementary material). This set included six peaks that had been previously identified by our ChIP-chip studies, giving a new discovery rate of 3/14, compared to 1/15 with EEL.

To summarize, 5/22 of the MCA predicted enhancers (or 11/28 including predictions covered by ChIP peaks) were bound by Gli1. At the gene level, at least 7/20 MCA predicted genes were verified to be direct Gli targets. Thus, pooling information from multiple potential CRMs in the MCA approach improved the ability to predict Gli target genes computationally.

DISCUSSION

In this study, we have developed a general approach for detecting targets of Gli transcriptional regulation by ChIP and applied this to a model of Shh-mediated patterning of the developing mammalian neural tube. The approach was validated by the identification of all previously reported ventral neural tube and general Gli-responsive enhancers that were represented on the custom oligonucleotide arrays, and confirmed by multiple statistical and biological methods. These analyses provide considerable confidence that the nine previously uncharacterized peaks present in the top tier of Table 1 represent genuine Gli targets. In total, the six new Gli target genes represented by the peaks in the top tier of Table 1 and the five new target genes predicted by the MCA in silico analysis and validated by qPCR of ChIP products almost double the total number of known mammalian Gli targets (summarized in Table S9 in the supplementary material). Because Hh-mediated neuronal specification is a relatively well-characterized system, this model provides a particularly good opportunity for validating candidate targets. The core of the strategy, a genetically inducible system for epitope-tagging transcription factors, should be broadly applicable to the study of other transcription factors. Further, the method can be adapted to enable Gli target identification in embryonic and adult tissues using either pre-selected or whole-genome tiling arrays. Modifying the endogenous Gli1 transcript to encode an epitope-tagged Gli1 protein is an alternative strategy; however, this approach restricts analysis to the normal Gli1 expression domain, often representing only a subset of the potential target field.

Identification of enhancer elements during neural tube development

Genome-scale transcription factor binding approaches have been used with great success in yeast, and more recently in ES cells (Lee et al., 2006) and *Drosophila* embryos (Alekseyenko et al., 2006; Sandmann et al., 2006). Embryonic analysis in mammalian systems presents several challenges, including the limited number of cells present in a given target population at a crucial regulatory stage and the dynamic expression of transcription factors during development. By modifying existing protocols we were able to generate robust results with an approximately 50-fold reduction in input ChIP product (approximately 2×10^6 cells per ChIP). The degree of enrichment we obtained with the $Gli1^{FLAG}$ construct is impressive considering that: (1) Gli1 is predominately cytoplasmic, and Hh signal-directed nuclear accumulation is transient (Kogerman et al., 1999); (2) the epitope-tagged construct must compete with endogenously transcribed, untagged Gli-activator; and (3) the target population is a mosaic in which relatively few cells represent a given ventral progenitor type (Fig. 1Q). Indeed, we were able to identify bona-fide Gli-activator target sites within enhancers in *Nkx2.2* and *FoxA2*, where $Nkx2.2^+$ and $FoxA2^+$ cells represented only 10% of the input population. Interestingly, a subset of predicted peaks did not contain any Gli consensus sites. These regions might be directly Gli-responsive despite the absence of a consensus Gli DNA-binding site. Alternatively, they might represent binding sites for other transcriptional regulators that could then make secondary protein-protein interactions with Gli1. While no specific interactions have been demonstrated for Gli-activator forms, Gli3 proteins have previously been shown to interact with members of the histone deacetylase complex (Cheng and Bishop, 2002; Dai et al., 2002) as well as with Smad proteins (Liu et al., 1998).

In addition to previously characterized *FoxA2* and *Nkx2.9* enhancers, we identified *Nkx2.2*, *Nkx2.1* and *Rab34* as direct targets through transgenic analysis and several new additional targets through validated bioinformatics predictions. However, direct transgenic analysis of *Nkx2.2* illustrates one major challenge the researcher faces in constructing regulatory circuitry from this data: that is, the identification of other cis-regulatory regions that act in conjunction with Gli factors to give appropriate temporal and spatial regulation. In the example of the *Nkx2.2* element we assayed, sequences that repress Gli-dependent FP expression and silence non-Gli dependent ventrolateral neural expression were not present and remain to be identified.

Surprisingly, our data indicate that a global Hh response represented by *Ptch1*, the best characterized transcriptional response to Hh signaling, may in fact be controlled by cell-type-specific response elements. Our analysis of one of four strong, independent Gli-binding regions demonstrates that the region selected represents a subset of the full range and activity of the *Ptch1* transcriptional response. This leaves open the possibility that other regions may encode limb response elements, or that a specific limb response element that normally acts in conjunction with the Gli site in *Ptch1* peak 2 lies outside the regulatory module assayed here. Our assay for putative enhancers using NIH3T3 cells allows us to predict several additional enhancers that are strong candidates for global regulators of a Hh signaling response (Table 2). Consistent with this interpretation, several genes associated with these enhancers are expressed in multiple Hh-responsive tissues (see Fig. S3 in the supplementary material), and all these targets seem to participate in negative feedback loops to regulate expression levels (see model in Fig. 4A).

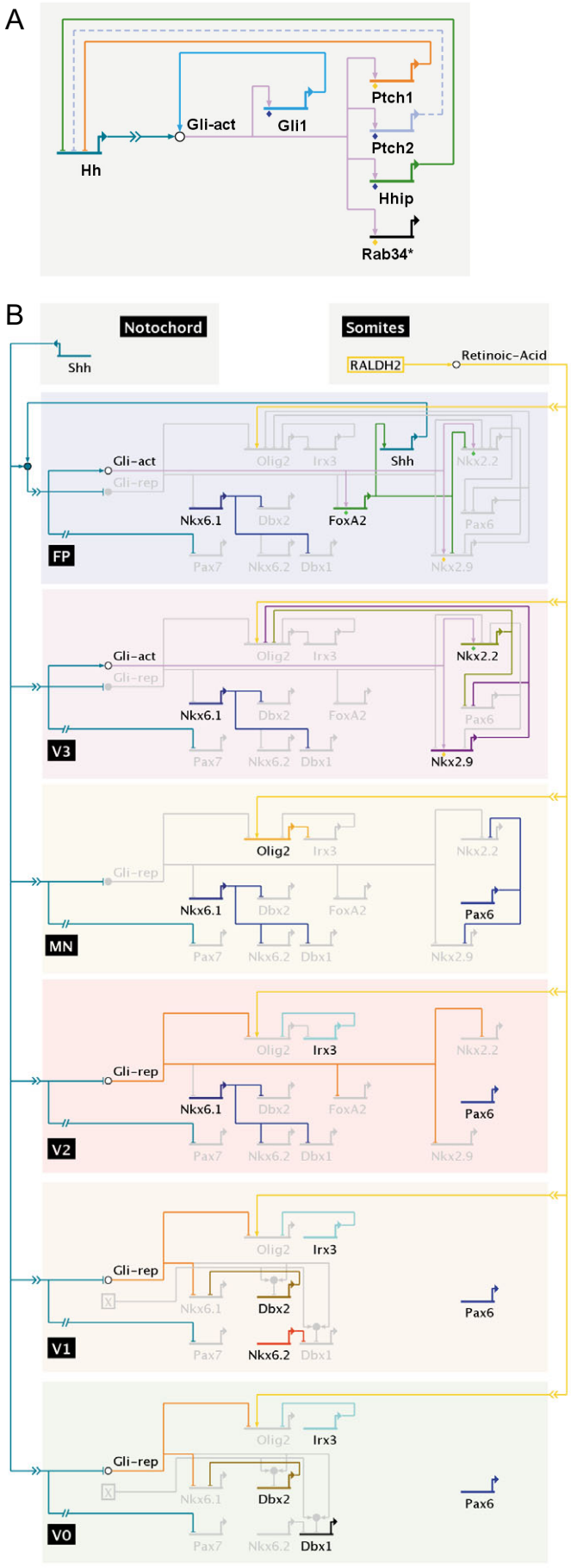

Fig. 4. BioTapestry models of Hh-driven cis-regulatory networks in mouse. Transcriptional targets of the Hh pathway can be defined by a globally responsive signaling cassette (**A**), containing components that are either known (unbroken lines) or likely (broken lines) to be part of a negative feedback loop. (**B**) A model for a Shh-driven transcriptional network underlying ventral neuronal specification. In this diagram, depicted in standard BioTapestry nomenclature (Longabaugh et al., 2005), neuronal specification is depicted as a sequential series of cell states. All genes not expressed are in gray, whereas currently expressed genes are in black or other colors. Similarly, inactive links (active in previous stages of specification) are depicted in gray, whereas active activation or repression is depicted using lines of other colors. An animation of these events can be viewed in Movie 1 in the supplementary material. This diagram focuses explicitly on ventral cell specification; thus previous events in general neuronal specification are not shown. Validated Gli targets (all identified or confirmed in our study) are indicated by blue diamonds (ChIP peak), orange diamonds (transgenic validation) or green diamonds (mutation of binding site in transgenic embryos).

Among the novel targets, *Rab34* is particularly intriguing given the role of another small GTP-binding protein, Rab23, as a crucial regulator of Hh responsiveness (Eggenschwiler et al., 2006; Eggenschwiler et al., 2001) and an independent link between Smo and Ptch1 activity in endosomal trafficking (Zhu et al., 2003). While transcriptional profiling identified *Rab34* as a general, positive target of Hh action in many tissues, *Rab34* is itself rather broadly expressed and therefore, based on expression alone, would not be viewed as a strong candidate for a specific role in the Hh pathway. However, the ChIP data, coupled with studies of Rab34 expression in *Ptch1* and Smo mutants (see Fig. S3D-D″ in the supplementary material), indicate that a component of the Rab34 expression pattern is Hh/Gli-dependent, an observation supported by our transgenic studies (Fig. 2W,X,A′,B′).

Gli sites and target gene regulation

Our data have allowed us to develop, and importantly, to then experimentally test, improved algorithms for the detection of Gli-binding sites, a significant advance for an experimental mammalian dataset. By combining ChIP analysis with transcriptional profiling data and MCA analysis, we were able to correctly predict Gli target genes with an optimal predictive rate of greater than 40% accuracy. Surprisingly, MCA predicts multiple Gli sites within a given locus, whereas only a subset were actually bound by Gli1. This raises the question of whether these are bound by other Gli factors, and if so, whether there is a mechanism that could impart specificity given that all Gli factors bind a similar consensus sequence (Hallikas et al., 2006; Vortkamp et al., 1995). One possibility is that in vivo target sites may have differential affinities for Gli factors.

As noted, we have failed to detect Gli-binding sites in 150 kb of tiled sequence flanking the *Nkx6.1*, one of the very earliest responders to the Hh pathway (see Table S1 in the supplementary material) (Jeong and McMahon, 2005). Whereas the possibility of a more distant Gli1-binding enhancer region remains open, an alternative explanation is that Gli activator binds only to the subset of Gli targets that are expressed in the ventralmost FP and pV3 domains. Because Nkx6.1 is still expressed in mice lacking the activity of all Gli proteins (both activator and repressor) in the neural tube, its expression depends principally upon loss of Gli repression rather than Gli activation (Bai et al., 2004; Lei et al., 2004; Wijgerde et al., 2002).

Our findings, together with published reports that have documented the relationship between Gli factors and Shh signaling in specification and the interactions among specific transcriptional regulators downstream of Shh input, lead to the model depicted in Fig. 4B (Briscoe et al., 2000; Ericson et al., 1997; Fogarty et al., 2005; Novitch et al., 2003; Santagati et al., 2003; Sasaki and Hogan, 1996; Vallstedt et al., 2001; Wijgerde et al., 2002). Shh signaling within the neural tube is effectively interpreted by region-specific combinations of Gli activator and Gli repressor forms. Initially only two cell types are present – one is a V0/V1 bipotential progenitor that depends only indirectly on a loss of Gli repression, which would in turn result in the loss of an inhibitory dorsal signal (Fogarty et al., 2005; Wijgerde et al., 2002). In the second cell state, there is an initial V2 progenitor, which would then sequentially give rise to V2-FP progenitor domains. Most dorsally, where Shh is never received, Gli repressors silence all Hh target genes. At the dorsal limits of Hh signaling (at the dorsal-ventral intersect), Gli repressor would be reduced, thereby relieving the inhibitory action of an unknown factor X on Dbx1 and Dbx2. Dorsal BMP signals have previously been hypothesized to play such a role (Wijgerde et al., 2002); this interaction with Gli repressor is especially attractive, as Gli3 has previously been shown to bind Smad proteins, transcriptional mediators of BMP signaling (Liu et al., 1998). Importantly, the regulation of these domains would not depend on any differential activity of Gli repressor – only on the attenuation of the dorsal inhibitory input. In domains that receive a higher Hh input, Gli repression would be progressively attenuated to the pMN domain, and loss of this repressor is sufficient for activation of targets. Within the pV3 and FP domains, however, specification requires varying levels of Gli-activator activity. Recent experiments in chick argue persuasively for a ratiometric assessment of Gli repressor and activator in dictating the position-specific Hh response both in the limb and the neural tube (Davey et al., 2006) (E. Dessaud and J. Briscoe, personal communication). Finally, the available data indicate an important dynamic component to ventral patterning, in which specific neural progenitor populations may change their identity on integrating Hh signaling over time (see Movie 1 in the supplementary material).

In summary, by applying biochemical and bioinformatics approaches, we have a new view of Hh targets, their regulation and the general parameters that characterize a Gli-activator response in patterning the vertebrate neural tube. These data provide a solid grounding for explaining the possible roles of newly defined targets (e.g. Rab34) and for determining the relationships between Gli proteins and other transcriptional regulators in coordinating precise transcriptional outcomes through local cis-acting regulatory modules in central Hh target genes.

We thank Curis for providing Hh-Ag, Alex Joyner and David Rowitch for providing constructs, Zhenjuan Wang (Harvard Genome Manipulation Facility) for pronuclear injections, Jennifer Couget (GGR) for expert guidance on performing microarrays, Julia Zeitlinger for helpful advice on optimizing ChIP, and Renate Hellmiss-Peralte for advice and assistance on figures. We are indebted to Leo Brizuela, Brett Chevalier and Mel Kronick at Agilent Technologies for supporting this project. We thank Dr James Briscoe for discussion of unpublished material. The Pax6 antibody (developed by the Kawakami Laboratory), Nkx2.2 and FoxA2 antibodies (developed by the Jessell laboratory) were obtained from the Developmental Studies Hybridoma Bank. We acknowledge support from the NIH-NINDS #R37 NS033642 (A.P.M.), NIH #R01 GM67250 and HG3903-01 (W.W. and H.J.),Helen Hay Whitney Foundation (S.V.) and the Charles King Trust Fellowship (S.V.).

Supplementary material

Supplementary material for this article is available at http://dev.biologists.org/cgi/content/full/134/10/1977/DC1

References

Agren, M., Kogerman, P., Kleman, M. I., Wessling, M. and Toftgard, R. (2004). Expression of the PTCH1 tumor suppressor gene is regulated by alternative promoters and a single functional Gli-binding site. *Gene* **330**, 101-114.

Alekseyenko, A. A., Larschan, E., Lai, W. R., Park, P. J. and Kuroda, M. I. (2006). High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. *Genes Dev.* **20**, 848-857.

Bai, C. B. and Joyner, A. L. (2001). Gli1 can rescue the in vivo function of Gli2. *Development* **128**, 5161-5172.

Bai, C. B., Stephen, D. and Joyner, A. L. (2004). All mouse ventral spinal cord patterning by hedgehog is Gli dependent and involves an activator function of Gli3. *Dev. Cell* **6**, 103-115.

Briscoe, J. and Ericson, J. (2001). Specification of neuronal fates in the ventral neural tube. *Curr. Opin. Neurobiol.* **11**, 43-49.

Briscoe, J., Sussel, L., Serup, P., Hartigan-O'Connor, D., Jessell, T. M., Rubenstein, J. L. and Ericson, J. (1999). Homeobox gene Nkx2.2 and specification of neuronal identity by graded Sonic hedgehog signalling. *Nature* **398**, 622-627.

Briscoe, J., Pierani, A., Jessell, T. M. and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell* **101**, 435-445.

Cheng, S. Y. and Bishop, J. M. (2002). Suppressor of Fused represses Gli-mediated transcription by recruiting the SAP18-mSin3 corepressor complex. *Proc. Natl. Acad. Sci. USA* **99**, 5442-5447.

Chuang, P. T. and McMahon, A. P. (1999). Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature* **397**, 617-621.

Crooks, G. E., Hon, G., Chandonia, J. M. and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188-1190.

Dai, P., Akimaru, H., Tanaka, Y., Maekawa, T., Nakafuku, M. and Ishii, S. (1999). Sonic Hedgehog-induced activation of the Gli1 promoter is mediated by GLI3. *J. Biol. Chem.* **274**, 8143-8152.

Dai, P., Shinagawa, T., Nomura, T., Harada, J., Kaul, S. C., Wadhwa, R., Khan, M. M., Akimaru, H., Sasaki, H., Colmenares, C. et al. (2002). Ski is involved in transcriptional regulation by the repressor and full-length forms of Gli3. *Genes Dev.* **16**, 2843-2848.

Davey, M. G., Paton, I. R., Yin, Y., Schmidt, M., Bangs, F. K., Morrice, D. R., Smith, T. G., Buxton, P., Stamataki, D., Tanaka, M. et al. (2006). The chicken talpid3 gene encodes a novel protein essential for Hedgehog signaling. *Genes Dev.* **20**, 1365-1377.

Davidson, E. H. and Erwin, D. H. (2006). Gene regulatory networks and the evolution of animal body plans. *Science* **311**, 796-800.

Eggenschwiler, J. T., Espinoza, E. and Anderson, K. V. (2001). Rab23 is an essential negative regulator of the mouse Sonic hedgehog signalling pathway. *Nature* **412**, 194-198.

Eggenschwiler, J. T., Bulgakov, O. V., Qin, J., Li, T. and Anderson, K. V. (2006). Mouse Rab23 regulates hedgehog signaling from smoothened to Gli proteins. *Dev. Biol.* **290**, 1-12.

Ericson, J., Rashbass, P., Schedl, A., Brenner-Morton, S., Kawakami, A., van Heyningen, V., Jessell, T. M. and Briscoe, J. (1997). Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. *Cell* **90**, 169-180.

Fogarty, M., Richardson, W. D. and Kessaris, N. (2005). A subset of oligodendrocytes generated from radial glia in the dorsal spinal cord. *Development* **132**, 1951-1959.

Frank-Kamenetsky, M., Zhang, X. M., Bottega, S., Guicherit, O., Wichterle, H., Dudek, H., Bumcrot, D., Wang, F. Y., Jones, S., Shulok, J. et al. (2002). Small-molecule modulators of Hedgehog signaling: identification and characterization of Smoothened agonists and antagonists. *J. Biol.* **1**, 10.

Goodrich, L. V., Milenkovic, L., Higgins, K. M. and Scott, M. P. (1997). Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* **277**, 1109-1113.

Hallikas, O., Palin, K., Sinjushina, N., Rautiainen, R., Partanen, J., Ukkonen, E. and Taipale, J. (2006). Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* **124**, 47-59.

Hooper, J. E. and Scott, M. P. (2005). Communicating with Hedgehogs. *Nat. Rev. Mol. Cell Biol.* **6**, 306-317.

Jacob, J. and Briscoe, J. (2003). Gli proteins and the control of spinal-cord patterning. *EMBO Rep.* **4**, 761-765.

Jeong, J. and McMahon, A. P. (2005). Growth and pattern of the mammalian neural tube are governed by partially overlapping feedback activities of the hedgehog antagonists patched 1 and Hhip1. *Development* **132**, 143-154.

Jessell, T. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat. Rev. Genet.* **1**, 20-29.

Ji, H. and Wong, W. H. (2005). TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics* **21**, 3629-3636.

Kinzler, K. W. and Vogelstein, B. (1990). The GLI gene encodes a nuclear protein which binds specific sequences in the human genome. *Mol. Cell. Biol.* **10**, 634-642.

Kogerman, P., Grimm, T., Kogerman, L., Krause, D., Unden, A. B., Sandstedt, B., Toftgard, R. and Zaphiropoulos, P. G. (1999). Mammalian suppressor-of-fused modulates nuclear-cytoplasmic shuttling of Gli-1. *Nat. Cell Biol.* **1**, 312-319.
Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F. and Wootton, J. C. (1993). Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* **262**, 208-214.
Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevalier, B., Johnstone, S. E., Cole, M. F., Isono, K. et al. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* **125**, 301-313.
Lei, Q., Zelman, A. K., Kuang, E., Li, S. and Matise, M. P. (2004). Transduction of graded Hedgehog signaling by a combination of Gli2 and Gli3 activator functions in the developing spinal cord. *Development* **131**, 3593-3604.
Lei, Q., Jeong, Y., Misra, K., Li, S., Zelman, A. K., Epstein, D. J. and Matise, M. P. (2006). Wnt signaling inhibitors regulate the transcriptional response to morphogenetic Shh-Gli signaling in the neural tube. *Dev. Cell* **11**, 325-337.
Li, C. and Wong, W. H. (2001). Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc. Natl. Acad. Sci. USA* **98**, 31-36.
Liu, F., Massague, J. and Ruiz i Altaba, A. (1998). Carboxy-terminally truncated Gli3 proteins associate with Smads. *Nat. Genet.* **20**, 325-326.
Liu, J. S. (1994). The collapsed Gibbs sampler with applications to a gene regulation problem. *J. Am. Stat. Assoc.* **89**, 958-966.
Longabaugh, W. J. R., Davidson, E. H. and Bolouri, H. (2005). Computational representation of developmental genetic regulatory networks. *Dev. Biol.* **283**, 1-16.
Mao, J., Barrow, J., McMahon, J., Vaughan, J. and McMahon, A. P. (2005). An ES cell system for rapid, spatial and temporal analysis of gene function in vitro and in vivo. *Nucleic Acids Res.* **33**, e155.
McMahon, A. P., Ingham, P. W. and Tabin, C. J. (2003). Developmental roles and clinical significance of hedgehog signaling. *Curr. Top. Dev. Biol.* **53**, 1-114.
Megason, S. G. and McMahon, A. P. (2002). A mitogen gradient of dorsal midline Wnts organizes growth in the CNS. *Development* **129**, 2087-2098.
Motoyama, J., Takabatake, T., Takeshima, K. and Hui, C. (1998). Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog. *Nat. Genet.* **18**, 104-106.
Muller, B. and Basler, K. (2000). The repressor and activator forms of Cubitus interruptus control Hedgehog target genes through common generic gli-binding sites. *Development* **127**, 2999-3007.
Nagai, T., Ibata, K., Park, E. S., Kubota, M., Mikoshiba, K. and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87-90.
Novitch, B. G., Wichterle, H., Jessell, T. M. and Sockanathan, S. (2003). A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification. *Neuron* **40**, 81-95.
Nybakken, K., Vokes, S. A., Lin, T. Y., McMahon, A. P. and Perrimon, N. (2005). A genome-wide RNA interference screen in Drosophila melanogaster cells for new components of the Hh signaling pathway. *Nat. Genet.* **37**, 1323-1332.
Odom, D. T., Zizlsperger, N., Gordon, D. B., Bell, G. W., Rinaldi, N. J., Murray, H. L., Volkert, T. L., Schreiber, J., Rolfe, P. A., Gifford, D. K. et al. (2004). Control of pancreas and liver gene expression by HNF transcription factors. *Science* **303**, 1378-1381.
Pabst, O., Herbrand, H., Takuma, N. and Arnold, H. H. (2000). NKX2 gene expression in neuroectoderm but not in mesendodermally derived structures depends on sonic hedgehog in mouse embryos. *Dev. Genes Evol.* **210**, 47-50.
Pan, Y., Bai, C. B., Joyner, A. L. and Wang, B. (2006). Sonic hedgehog signaling regulates Gli2 transcriptional activity by suppressing its processing and degradation. *Mol. Cell. Biol.* **26**, 3365-3377.
Pavletich, N. P. and Pabo, C. O. (1993). Crystal structure of a five-finger GLI-DNA complex: new perspectives on zinc fingers. *Science* **261**, 1701-1707.
Ren, B., Robert, F., Wyrick, J. J., Aparicio, O., Jennings, E. G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E. et al. (2000). Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306-2309.
Ruiz i Altaba, A., Nguyen, V. and Palma, V. (2003). The emergent design of the neural tube: prepattern, SHH morphogen and GLI code. *Curr. Opin. Genet. Dev.* **13**, 513-521.
Sandmann, T., Jensen, L. J., Jakobsen, J. S., Karzynski, M. M., Eichenlaub, M. P., Bork, P. and Furlong, E. E. (2006). A temporal map of transcription factor activity: mef2 directly regulates target genes at all stages of muscle development. *Dev. Cell* **10**, 797-807.
Santagati, F., Abe, K., Schmidt, V., Schmitt-John, T., Suzuki, M., Yamamura, K. and Imai, K. (2003). Identification of Cis-regulatory elements in the mouse Pax9/Nkx2-9 genomic region: implication for evolutionary conserved synteny. *Genetics* **165**, 235-242.
Sasaki, H. and Hogan, B. L. (1996). Enhancer analysis of the mouse HNF-3 beta gene: regulatory elements for node/notochord and floor plate are independent and consist of multiple sub-elements. *Genes Cells* **1**, 59-72.
Sasaki, H., Hui, C., Nakafuku, M. and Kondoh, H. (1997). A binding site for Gli proteins is essential for HNF-3beta floor plate enhancer activity in transgenics and can respond to Shh in vitro. *Development* **124**, 1313-1322.
Srinivas, S., Watanabe, T., Lin, C. S., William, C. M., Tanabe, Y., Jessell, T. M. and Costantini, F. (2001). Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1**, 4.
Stamataki, D., Ulloa, F., Tsoni, S. V., Mynett, A. and Briscoe, J. (2005). A gradient of Gli activity mediates graded Sonic Hedgehog signaling in the neural tube. *Genes Dev.* **19**, 626-641.
Teboul, L., Summerbell, D. and Rigby, P. W. (2003). The initial somitic phase of Myf5 expression requires neither Shh signaling nor Gli regulation. *Genes Dev.* **17**, 2870-2874.
Tenzen, T., Allen, B. L., Cole, F., Kang, J. S., Krauss, R. S. and McMahon, A. P. (2006). The cell surface membrane proteins Cdo and Boc are components and targets of the Hedgehog signaling pathway and feedback network in mice. *Dev. Cell* **10**, 647-656.
Vallstedt, A., Muhr, J., Pattyn, A., Pierani, A., Mendelsohn, M., Sander, M., Jessell, T. M. and Ericson, J. (2001). Different levels of repressor activity assign redundant and specific roles to Nkx6 genes in motor neuron and interneuron specification. *Neuron* **31**, 743-755.
Vortkamp, A., Gessler, M. and Grzeschik, K. H. (1995). Identification of optimized target sequences for the GLI3 zinc finger protein. *DNA Cell Biol.* **14**, 629-634.
Wang, B., Fallon, J. F. and Beachy, P. A. (2000). Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. *Cell* **100**, 423-434.
Wichterle, H., Lieberam, I., Porter, J. A. and Jessell, T. M. (2002). Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110**, 385-397.
Wijgerde, M., McMahon, J. A., Rule, M. and McMahon, A. P. (2002). A direct requirement for Hedgehog signaling for normal specification of all ventral progenitor domains in the presumptive mammalian spinal cord. *Genes Dev.* **16**, 2849-2864.
Zhang, X. M., Ramalho-Santos, M. and McMahon, A. P. (2001). Smoothened mutants reveal redundant roles for Shh and Ihh signaling including regulation of L/R symmetry by the mouse node. *Cell* **106**, 781-792.
Zhu, A. J., Zheng, L., Suyama, K. and Scott, M. P. (2003). Altered localization of Drosophila Smoothened protein activates Hedgehog signal transduction. *Genes Dev.* **17**, 1240-1252.

Table S1. Genomic loci present on the 44K genomic tiling array

Gene	ID	Chromosome	Strand	TSS	TES	Coverage	Category
Gli	NM_010296	10	-	127066524	127078216	150001	Positive
Ptch1	NM_008957	13	-	60894523	60948510	150001	Positive
Ptch2	NM_008958	4	+	116055261	116073736	148978	Positive
Hhip	NM_020259	8	-	79185727	79272113	150001	Positive
Foxa2	NM_010446	2	-	147499859	147503905	150001	Positive
Nkx6-1	NM_144955	5	-	100679979	100685482	149959	Positive
Nkx2-2	NM_010919	2	-	146640088	146643295	150001	Positive
Hox D cluster		2	n/a	74133959	74195987	62029	Positive
Nkx2-1_Nkx2-9_Pax9		12	n/a	53197523	53497991	300469	Positive
Neurog2	NM_009718	3	+	126516140	126518624	50001	Positive
Ascl1	NM_008553	10	-	87463687	87466041	50001	Positive
Nkx6-2	NM_183248	7	-	133975237	133978657	49826	Positive
Ntn1	NM_008744	11	-	67938664	68111778	50001	Positive
Hsd11b2	NM_008289	8	+	104814482	104819145	50001	Positive
Ccnd1	NM_007631	7	-	139343751	139352397	50001	Positive
Cthrc1	AK003674	15	+	38976515	38986748	50001	Positive
3110038L01Rik	NM_026524	X	+	8956434	8958633	50001	Positive
Rab34	NM_033475	11	+	77914493	77917846	50001	Positive
Foxd1	NM_008242	13	+	94542754	94545191	50001	Positive
Foxf2	NM_010225	13	+	31105401	31111024	50001	Positive
Foxc2	NM_013519	8	+	120472104	120474411	50001	Positive
–	NM_177863	4	-	81884119	82007294	50001	Positive
AI504062	AI504062	5	+	3536607	3537009	56001	Positive
Sall1	NM_021390	8	-	88313728	88327923	50001	Positive
Shh	NM_009170	5	-	26906865	26917112	50001	Positive
Foxf1a	NM_010426	8	+	120439995	120442399	50001	Positive
Cspg2	AK034871	13	-	85826083	85857173	50001	Positive
Leprel1	NM_173379	16	-	24747134	24892594	50001	Positive
9430010M06Rik	AK020409	9	-	24616095	24617172	50001	Positive
Rhoe	NM_028810	2	-	51062603	51081273	50001	Positive
2310037P21Rik	NM_001002267	9	-	123277459	123278318	50001	Positive
Colec11	BC000078	12	-	25160227	25182803	50001	Positive
3110070M22Rik	NM_026084	Un_random	+	19413072	19414122	50001	Positive
Tpd52l1	NM_009413	10	-	31372578	31486119	50001	Positive
Ncor2	NM_011424	5	-	124178459	124340350	50001	Positive
Galt	NM_016658	4	+	41894196	41897373	50001	Positive
3230401I01Rik	NM_026140	7	-	115916679	115944921	50001	Positive
Homer2	NM_011983	7	-	75417865	75514933	50001	Positive
Ccl27	NM_011336	4	-	41908172	41908905	50001	Positive
Nkx2-5	NM_008700	17	-	24633727	24636672	50001	Positive
Nkx3-1	NM_010921	14	+	63724262	63728228	50001	Positive
Gdf1	NM_008107	8	+	69476431	69485066	50001	Positive
–	BB344065	X	+	118149099	118149801	0	Positive
Gpsm1	NM_153410	2	+	26247702	26280397	50001	Positive
Ccnd2	NM_009829	6	-	127802646	127824339	50001	Positive
Cdk2ap1	NM_013812	5	-	123516784	123525971	47196	Positive
Cdk6	NM_009873	5	+	3350317	3528230	50001	Positive
Arl4	NM_007487	12	-	36632029	36636075	50001	Positive
Fbn2	NM_010181	18	-	58225408	58424955	50001	Positive
Slit3	AF144629	11	+	34915546	35502170	50001	Positive
Foxc1	NM_008592	13	+	31286268	31290256	50001	Positive
Nsg2	NM_008741	11	+	31895253	31953996	50001	Positive
Hoxa11	NM_010450	6	-	52386167	52389779	50001	Positive
Hoxd11	NM_008273	2	+	74380142	74381886	50001	Positive
Hoxd13	NM_008275	2	+	74366059	74367884	50001	Positive
Hoxa13	NM_008264	6	-	52403288	52404984	50001	Positive

Foxa1	NM_008259	12	-	54272108	54277598	50001	Positive
Myf5	NM_008656	10	-	107369978	107373204	50001	Positive
Vtn	NM_011707	11	+	78224945	78227982	50001	Positive
Pitx1	NM_011097	13	-	54441613	54447983	50001	Positive
Hlx	NM_008250	1	-	184225321	184230785	50001	Positive
Hand2	NM_010402	8	+	56361191	56363149	50001	Positive
Cdh5	NM_009868	8	+	103396382	103438830	50001	Positive
Csnk2a1	NM_007788	2	+	151683883	151738781	50001	Positive
Jag1	NM_013822	2	-	136595317	136630150	50001	Positive
Hey1	NM_010423	3	-	8650291	8653917	50001	Positive
Meox2	NM_008584	12	+	33709011	33783546	50001	Positive
Pou3f1	NM_011141	4	+	123684765	123686282	50001	Positive
Angptl6	NM_145154	9	-	20751107	20756927	49924	Positive
Ror1	NM_013845	4	+	99054696	99401450	50001	Negative
–	BM211445	6	-	49390692	49391335	50001	Negative
–	BB759978	1	-	78388625	78389095	50001	Negative
Igfbp2	NM_008342	1	+	73126184	73154145	50001	Negative
Irx3	NM_008393	8	-	91085099	91087946	50001	Negative
Msx1	NM_010835	5	-	36327005	36330969	49989	Negative
Msx2	NM_013601	13	-	52026240	52031427	50001	Negative
Snai2	NM_011415	16	+	13416441	13419962	50001	Negative
Crabp1	NM_013496	9	+	54882658	54891162	50001	Negative
Crabp2	NM_007759	3	+	87692354	87697104	50001	Negative
Fap	NM_007986	2	-	62356345	62429421	50001	Negative
Irx2	NM_010574	13	+	68682462	68687568	50001	Negative
Ptn	NM_008973	6	-	36850202	36945899	50001	Negative
Irx5	NM_018826	8	+	91644914	91648572	50001	Negative
5031439A09Rik	NM_026582	3	+	158855550	158939401	50001	Negative
Flrt3	NM_178382	2	-	140172058	140185326	50001	Negative
6330404E20Rik	NM_028707	18	+	36188770	36238339	50001	Negative
6720477E09Rik	NM_178660	9	-	116564518	117239065	50001	Negative
Paip1	NM_145457	13_random	+	16918866	16946076	50001	Negative
Gsn	NM_146120	2	+	35214602	35240054	49986	Negative
MGC57096	NM_183191	3	-	63364164	63519393	50001	Negative
–	BE685845	7	+	122858915	122859376	50001	Negative
Fabp7	NM_021272	10	+	58050984	58054546	49901	Negative
Cbx1	NM_007622	11	+	96610246	96629727	50001	Negative
Dcc	NM_007831	18	-	71491611	72583420	50001	Negative
Dlgh1	NM_007862	16	+	30473410	30682713	50001	Negative
Efna5	NM_207654	17	-	60297718	60574961	50001	Negative
Lhx2	NM_010710	2	+	38283470	38301890	50001	Negative
Ndrg1	NM_010884	15	-	66953527	66993844	50001	Negative
Nfil3	NM_017373	13	-	51530723	51544552	50001	Negative
Prrx1	NM_011127	1	-	163163407	163227993	50001	Negative
Prdx2	NM_011563	8	+	84237800	84242986	50001	Negative
Zic2	NM_009574	14	+	117045015	117049406	50001	Negative
Taf7	NM_011901	18	-	37865744	37867869	50001	Negative
Axl	NM_009465	7	-	20922057	20953324	50001	Negative
Zic5	NM_022987	14	-	117028709	117035238	50001	Negative
2410012A13Rik	NM_023396	2	-	54009632	54011091	50001	Negative
BC030477	NM_177618	11	+	71476361	71515302	50001	Negative
–	BB534113	16	-	51367678	51368359	50001	Negative
Egr2	NM_010118	10	+	67591672	67595978	50001	Negative
Sfrp2	NM_009144	3	+	83502189	83510183	49997	Negative
Spry2	NM_011897	14	-	100435328	100439630	49984	Negative
Zfh4	NM_030708	3	+	5216397	5389233	50001	Negative
Boc	NM_172506	16	-	43368375	43441912	50001	Negative
Pde2a	NM_001008548	7	+	95527997	95619102	50001	Negative

Cart1	NM_172553	10	-	102921768	102942802	50001	Negative
Efemp1	NM_146015	11	+	28762563	28821523	50001	Negative
Cpeb2	NM_175937	5	+	41990144	42042375	50001	Negative
Pbx3	NM_016768	2	-	34104071	34304198	50001	Negative
Tcfap2a	NM_011547	13	-	40277825	40296013	50001	Negative
Twist1	NM_011658	12	+	30570330	30572488	50001	Negative
Cntnap2	NM_025771	6	+	47445151	47500249	50001	Negative
Sox10	BC023356	15	-	79206196	79215611	50001	Negative
Gas1	NM_008086	13	-	58831933	58834893	150001	Negative
Zic1	NM_009573	9	-	91251442	91256620	150001	Negative
Zic3	NM_009575	X	+	52785357	52790617	150001	Negative
Gli3	NM_008130	13	+	14911831	15175564	150001	Negative
Pax6	NM_013627	2	+	105381207	105401822	150001	Negative
Cdon	NM_021339	9	+	35364530	35416802	150001	Negative
Pax7	NM_011039	4	-	138619342	138714229	150001	Negative

Genes were tiled over the range defined by the above coordinates. Coverage refers to the distance tiled (not including regions that were masked for repeats). Positive genes were upregulated by Hh signaling in transcriptional profiling studies, whereas negative genes were downregulated (combinations of genes in the HoxD cluster and Nkx2.1_Nkx2.9_Pax9 refer to additional loci that contained extensive tiling throughout the region).

Table S2. Summary statistics of Gli1 ChIP peaks

Rank	Gene	Peak	Chr	Start	End	Size of region	Center	T-Stat	FDR	Fold change
1	Ptch1	Peak 2	13	60949397	60950371	974	60949566	5.507	0	16.0503
2	Nkx2-2	Peak 1	2	146644626	146645570	944	146645242	5.144	0	12.8085
3	Ptch1	Peak 3	13	60951377	60952953	1576	60952201	4.8769	0	11.2259
4	Nkx2-9	Peak 1	12	53347310	53348354	1044	53347899	4.7249	0	6.9036
5	Ptch1	Peak 5	13	60956284	60956777	493	60956284	4.2212	0	7.1708
6	Nkx2-9	Peak 2	12	53340519	53341015	496	53340862	4.1573	0	4.1739
7	Rab34	Peak 1	11	77915220	77915560	340	77915220	4.0735	0	9.0651
8	Ptch2	Peak 1	4	116057320	116057946	626	116057561	3.9081	0	5.5392
9	Ptch1	Peak 1	13	60945578	60945972	394	60945972	3.8484	0	7.9795
10	Nkx2.1	Peak 1	12	53245748	53246351	603	53246120	3.7201	0	4.0266
11	Gli1	Peak 1	10	127076571	127078099	1528	127076975	3.4976	0	4.6198
12	Hhip	Peak 1	8	79270092	79270492	400	79270240	3.4535	0	4.9558
13	Foxa2	Peak 1	2	147493701	147494245	544	147493953	3.4405	0	3.1634
14	Ptch2	Peak 2	4	116054385	116054954	569	116054597	3.3145	0	3.9994
15	Cart1	Peak 1	10	102925949	102926177	228	102926042	3.116	0	3.559
16	Ptch1	Peak 4	13	60953950	60955072	1122	60954831	3.0549	0	3.8567
17	Prdx2	Peak 1	8	84215218	84215559	341	84215273	3.0089	0	4.0679
18	Cart1	Peak 2	10	102938208	102938208	0	102938208	2.9954	0	3.2221
19	Ptch1	Peak 6	13	61010241	61010714	473	61010369	2.9851	0	3.3462
20	Hhip	Peak 2	8	79268419	79268867	448	79268732	2.9702	0	4.4496
21	Flrt3	Peak 1	2	140191468	140191692	224	140191623	2.8017	0	3.5177
22	Pax9	Peak 1	12	53443055	53443311	256	53443186	2.7947	0	2.9971
23	Ncor2	Peak 1	5	124340743	124341009	266	124340909	2.7832	0.0435	3.5798
24	Zic3	Peak 1	20	52789693	52789967	274	52789693	2.6971	0.0769	2.6119
25	Hand2	Peak 1	8	56373391	56373692	301	56373552	2.6907	0.0769	2.7515
26	Cart1		10	102963671	102964187	516	102963911	2.6876	0.0769	4.6047
27	Twist1		12	30577542	30577932	390	30577763	2.6562	0.1034	4.3133
28	Nkx2-9		12	53333143	53333408	265	53333313	2.6414	0.1034	3.5581
29	Myf5		10	107361999	107362096	97	107362096	2.6329	0.1034	3.9894
30	Ptn		6	36849581	36849581	0	36849581	2.5351	0.1667	3.5856
31	Ncor2		5	124356413	124356633	220	124356540	2.4556	0.1818	3.1064
32	Nkx2.1		12	53220111	53220111	0	53220111	2.3832	0.1818	2.9039
33	Foxa2		2	147497880	147498001	121	147497880	2.3664	0.1818	3.2767
34	Pax9		12	53399381	53399596	215	53399475	2.3388	0.1842	3.3657
35	Ncor2		5	124346282	124346436	154	124346375	2.3308	0.1842	2.9774
36	Ptch1		13	60955660	60955823	163	60955660	2.3207	0.1842	3.6777
37	Foxf2		13	31129412	31129501	89	31129501	2.3084	0.1842	2.8392
38	Mid1ip1		20	8973477	8973606	129	8973606	2.3043	0.1842	3.3893
39	Nkx6-1		5	100666378	100666493	115	100666493	2.2766	0.2051	3.4918
40	Ptch1		13	60946669	60946815	146	60946815	2.237	0.2128	2.8874
41	Angpt16		9	20734531	20734687	156	20734531	2.222	0.2128	3.4968
42	Foxc1		13	31302115	31302268	153	31302268	2.2081	0.2128	2.8619
43	5031439A09Rik		3	158868339	158868516	177	158868516	2.2043	0.2128	2.9474
44	Fabp7		10	58049170	58049170	0	58049170	2.2036	0.2128	2.8
45	Dlgh1		16	30486849	30487046	197	30486849	2.1944	0.2128	2.7407
46	Hhip		8	79267697	79267697	0	79267697	2.1859	0.2128	2.5583
47	Pbx3		2	34297540	34297625	85	34297625	2.1846	0.2128	2.1798

Peaks were defined and ranked using TileMap. Using an FDR cut-off of ⩽25%, there are a total of 47 peaks. Chromosomal coordinates (Build 34) showing the boundary coordinates for peaks. The size of the region is measured by subtracting these values. In the instances where the region has a size of zero, there is only one probe that is significantly enriched. The T-statistic gives the value for the highest peak probe in the region. FDR, false discovery rate.

Table S3. Enrichment of Gli-binding motif in candidate regions

Peak Rank	n1/n2	r1	FDR(1/r1)	n3/n4	r2/r3
1-10	13/10820	5.33	19%	10/5328	7.73/17.34
11-20	11/9579	5.09	20%	8/4740	6.95/15.66
21-30	5/ 6280	3.53	28%	3/2720	4.54/8.96
31-40	4/ 4959	3.58	28%	1/1594	2.58/3.78
41-50	1/ 4919	0.9	100%	0/1826	0.00/0.00
51-60	2/ 4159	2.13	47%	2/1241	6.64/9.02
Control	1264/5608267	1		299/1231633	1.00/1.00

The Gli consensus motif TGGGTGGTC, with a tolerance for 1 bp mismatch was mapped to statistically ranked peak regions. Peaks were grouped into bins of size 10. The number of Gli sites (n_1) and the total number of surveyed non-repeat base pairs (n_2) were counted for each bin as well as control regions (i.e. all regions tiled in the array). $r_1 = (n_1/n_2)_{target}/(n_1/n_2)_{control}$ measures the Gli enrichment level for each bin. FDR=$1/r_1$ measures the false discovery rate of Gli-binding sites as predicted by consensus. Sites that reside within the top 10% of most-conserved regions of the genome are defined as conserved binding sites. n_3 and n_4 are the number of conserved Gli consensus sites and the total number of conserved base pairs in each bin, respectively. $r_2 = (n_3/n_4)_{target}/(n_3/n_4)_{control}$ measures the extent to which Gli sites are enriched in conserved regions, and $r_3 = (n_3/n_2)_{target}/(n_3/n_2)_{control}$ measures the combined gain we can obtain using both binding-sequence specificity and cross-species conservation information. The level of Gli enrichment (r_1, r_2 and r_3) is reduced in lower-scoring regions

Table S4. Validation with qPCR on unamplified ChIPs showing enrichment values, primers and coordinates

Primer	Locus	Peak	ΔΔC(T)	St. dev.	Fold enrichment	Amplicon coordinates	Forward primer	Reverse primer
807808	Ptch1	Peak1	4.91	0.8	30.0	13:60945711-60945972	TTATAAAGCAGGTCCCCAAC	ACGCATGTTTGCAAAGTAGA
727728	Nkx2.2	Peak1	4.87	0.6	29.2	2:146645052-146645290	CGGGTTTCTAGAGCCTTGTG	GACAAAGGCTTCCAACAGGA
801802	Rab34	Peak1	4.81	0.6	28.1	11:77915199-77915431	CAAAAGCTCCATTCTCACAA	AGGAAACTGGGAGAACATCA
793794	Nkx2.9	Peak1	4.01	0.7	16.2	12:53347646-53347851	GAGTCTTAAATCGGGGATGA	GAGACAATTTTCTGCCCATC
731732	Gli1	Peak1	3.65	0.6	12.5	10:127077945-127078145	GGACAAAGAGACCTGGGACA	AGGAGATGCTCTGACGCCTA
803804	Nkx2.1	Peak1	3.18	0.6	9.1	12:53245985-53246192	AATTGCTTAGGGTTCAGGTG	TTTGTTGAAGGCCTGTAGAA
879880	Gli1	Peak1	2.85	0.7	7.2	10:127076806-127076959	ATCCTGAAAGCAGGCAGTAG	AGGAATCCAAGGTGTCCTCT
713714	FoxA2	Peak1	2.68	0.6	6.4	2:147493520-147493736	GGGGCCCAGTCTAGATCTTC	CACACCACAAACACACAGCA
777778	Hhip1	Peak1	2.48	0.8	5.6	8:79,270,322-79,270,491	CGCAGGAGCTGTTTACATGA	AAGCCTGCAGTTGTCAAGGT
799800	Ptch2	Peak1	2.47	0.7	5.6	4:116057694-116057895	GGCTAGTCCCTGCAGTTATG	TCAGCTCTCAACTCCATCAA
729730	Gli1	Peak1	2.24	0.6	4.7	10:127077694-127077964	CCACAGGTCAGGAGTCCAAT	TGTCCCAGGTCTCTTTGTCC
855856	Hand2	Peak1	1.45	0.8	2.7	8:56373405-56373599	TTGGCTCTGTACTTTGCTCA	TAGAGGCTCCCAGTTCTTTG
791B792	Prdx2	Peak1	0.44	0.6	1.4	8:84215391-84215629	TGAACACCCTGAATGAAATG	TAGTAATGCCTTCCCCTCTG
833834	Ptch2	Peak2	5.59	0.6	48.3	4:116054701-116054867	ACTTACACTTGGCCAACGAC	CGGATCCGTTGCTAGGTA
805806	Nkx2.9	Peak2	5.24	0.6	37.7	12:53340854-53341077	CTACCAAGCGTGCCTAAAGT	TTTATCCCAGGGAGCTAAAG
711712	Ptch1	Peak2	5.09	0.6	34.0	13:60949317-60949518	GAAGTATTGCATGCGAGAGG	CTGTCAGATGGCTTGGGTTT
841842	Hhip1	Peak2	3.15	0.7	8.9	8:79268572-79268810	TAAAAGGGCACACTTGAAAA	AATTGCTGCAGACCCTAAAT
795796	Ptch1	Peak3	5.50	0.6	45.3	13:60952107-60952356	CCAGGCTCTCTCTAACAAGC	ACAACATTATCGCTGGGAAT
831832	Ptch1	Peak4	3.17	0.8	9.0	13:60954689-60954893	GAAGGACTCTGGGTGACAAC	TTGGCTCCAGTACTTCAACA
797798	Ptch1	Peak5	6.02	0.7	65.0	13:60956197-60956431	ACAAAAGGAACGGAAAGTGT	AGCACGTTCCTCCAGTTTAC
829830	Ptch1	Peak6	1.48	0.6	2.8	13:61010364-61010604	GCAGGGTAATGGGATTTATG	GTCTTTCCTGAAAGCCAAAA
825826	Negative control	Neg Ctrl	1.32	0.7	2.5	7:133989909-133990158	CAAGCTCCAGCCATTAAGTT	GGTACTCAGTGTGCTTGCTG
861862	Ptch1 locus Gli b.s. (–)	Neg Ctrl	2.28	0.6	4.8	13:60946932-60947110	GGCTAGTCCCTGCAGTTATG	TCAGCTCTCAACTCCATCAA
863864	Ptch1 locus Gli b.s. (–)	Neg Ctrl	1.09	0.7	2.1	13:60953479-60953701	GGGCTTGACTCTCTTCACAT	ACTGTAAGCTCCTGGGTCTG
867868	Nkx2.2 locus Gli b.s. (–)	Neg Ctrl	0.75	0.6	1.7	2:146647208-146647417	ATCCAAGCTCTGAGTTGAGG	AAGTGCCAAACAGGTCTCTC
859860	Ptch1 locus Gli b.s. (–)	Neg Ctrl	0.67	0.7	1.6	13:60942920-60943069	TCACCCTATCTCTGGAAAGC	GTTGGGGAAATCAGTGAAAG
689690	B-actin	Neg Ctrl	0.44	2.8	1.4	5:141969468-141969691	AGAAGGACTCCTATGTGGGTGA	ACTGACCTGGGTCATCTTTTC
865866	Ptch1 locus Gli b.s. (–)	Neg Ctrl	0.16	0.7	1.1	13:60958637-60958864	TTCTTTGGAGCTCAATTTCC	GTTTTCCCCGATTTTAAGGT
811812	Negative control region 2	Neg Ctrl	0.00	0.7	1.0	13:51535150-51535387	CTGGCCTCCATACACACATA	AGTCAGCAGGATCCACACTT
827828	Negative control region 10	Neg Ctrl	-0.14	0.7	0.9	15:66979588-66979802	AGAGGTATGGCAGAGGTGAG	CTTGGTGAGTCTTTGGGAAC
869870	Nkx2-1 lcous Gli b.s (–)	Neg Ctrl	-0.36	0.6	0.8	12:53255352-53255521	ATTTCTTTTCCTGGGAGGTC	TAGGAGCTGGGTTTAAGGTG
875876	Hhip1 locus Gli b.s. (–)	Neg Ctrl	-0.51	0.8	0.7	8:79272744-79272928	TGAGGTGGTTTGGAGACTTT	AGTCTGGGTGTACAGGTGGT
691692	Rosa26	Neg Ctrl	-0.54	0.7	0.7	6:113638429-113638553	CACCACAAATCGAGGCTGTA	CCTGACGGGAGAGGTGATAG
809810	Negative control region 1	Neg Ctrl	-1.04	1.2	0.5	10:127016990-127017237	AAGCCAGATGTGATGTAGCA	TTATGGAAGGACAGACAGCA

All enrichment values are determined using primer pair 811-812 for baseline values. Negative controls (see 'Peak' column) fall into several categories: (I) loci referred to by name only are random loci that are not affected by Hh signalling; (ii) loci referred to as negative control refer to random tiled regions of the genome that were not enriched in the chip; (iii) loci containing 'Gli b.s.' contain a Gli site with no more than two deviations from consensus. This final class consists of Gli-like motifs that were adjacent to positive peak regions but which showed no ChIP enrichment in the microarray data set.

Table S5. The 365 probe-sets exhibiting maximal differences in transcriptional profiling studies of EBs treated with retinoic acid and Hh agonist versus retinoic acid only

Rank	Probeset_id	Score	FDR	Gene	Gene ID	Accession
1	1422284_at	-1.80E+01	0.00E+00	Nkx2-9	18094	NM_008701
2	1451342_at	-1.42E+01	0.00E+00	Spon1	233744	NM_145584
3	1455907_x_at	-1.37E+01	0.00E+00	Phox2b	18935	NM_008888
4	1416286_at	-1.20E+01	0.00E+00	Rgs4	19736	NM_009062
5	1418054_at	-1.14E+01	0.00E+00	Neurod4	11923	NM_007501
6	1422833_at	-1.07E+01	1.25E-02	Foxa2	15376	NM_010446
7	1436694_s_at	-1.06E+01	1.25E-02	Neurod4	11923	NM_007501
8	1422720_at	-1.04E+01	1.25E-02	Isl1	16392	NM_021459
9	1450723_at	-9.60E+00	2.00E-02	Isl1	16392	NM_021459
10	1458536_at	-9.07E+00	2.00E-02	Ccni	12453	NM_017367
11	1428741_at	-8.99E+00	2.08E-02	Elavl4	15572	NM_010488
12	1425828_at	-8.82E+00	2.08E-02	Nkx6-1	18096	NM_144955
13	1416287_at	-8.38E+00	4.23E-02	Rgs4	19736	NM_009062
14	1452894_at	-8.24E+00	4.29E-02	Elavl4	15572	NM_010488
15	1450258_a_at	-7.77E+00	5.33E-02	Elavl4	15572	NM_010488
16	1419549_at	-7.47E+00	5.58E-02	Arg1	11846	NM_007482
17	1445047_at	-7.39E+00	5.58E-02	C79246	96891	---
18	1448475_at	-7.21E+00	5.58E-02	Olfml3	99543	NM_133859
19	1416232_at	-7.17E+00	5.58E-02	Olig2	50913	NM_016967
20	1426255_at	-7.16E+00	5.58E-02	Nefl	18039	NM_010910
21	1415885_at	-7.11E+00	5.58E-02	Chgb	12653	NM_007694
22	1433906_at	-7.07E+00	5.58E-02	4933402J24Rik	74438	NM_028940
23	1436319_at	-6.94E+00	5.58E-02	Sulf1	240725	NM_172294
24	1422520_at	-6.85E+00	5.58E-02	Nef3	18040	NM_008691
25	1440056_at	-6.79E+00	5.58E-02	---	---	---
26	1423977_at	-6.79E+00	5.58E-02	4930453N24Rik	67609	XM_148353
27	1458140_at	-6.75E+00	5.93E-02	Slit2	20563	NM_178804
28	1460000_at	-6.68E+00	6.25E-02	D830007B15Rik	330096	XM_485619
29	1459852_x_at	-6.54E+00	6.71E-02	Trib2	217410	NM_144551
30	1447709_at	-6.44E+00	6.71E-02	Epb4.1l3	13823	NM_013813
31	1425412_at	-6.34E+00	6.71E-02	Cias1	216799	NM_145827
32	1416149_at	-6.31E+00	6.71E-02	Olig1	50914	NM_016968
33	1423281_at	-6.25E+00	6.71E-02	Stmn2	20257	NM_025285
34	1460181_at	-6.24E+00	6.71E-02	Stmn3	20262	NM_009133
35	1454672_at	-6.19E+00	6.71E-02	---	---	---
36	1452842_at	-6.07E+00	7.22E-02	Hspb9	75482	---
37	1459524_at	-5.99E+00	7.57E-02	Mcc	328949	XM_140309
38	1436359_at	-5.94E+00	7.89E-02	Ret	19713	NM_009050
39	1451990_at	-5.89E+00	8.26E-02	Mapre2	212307	NM_153058
40	1455865_at	-5.86E+00	8.26E-02	Insm1	53626	NM_016889
41	1438989_s_at	-5.86E+00	8.26E-02	B130021B11Rik	320860	---
42	1418620_at	-5.83E+00	8.26E-02	Phox2a	11859	NM_008887
43	1438200_at	-5.81E+00	8.26E-02	Sulf1	240725	NM_172294
44	1415999_at	-5.60E+00	1.05E-01	Hey1	15213	NM_010423
45	1457403_at	-5.48E+00	1.16E-01	---	---	---
46	1436287_at	-5.46E+00	1.18E-01	---	---	---
47	1416313_at	-5.42E+00	1.18E-01	Mllt11	56772	NM_019914
48	1425078_x_at	-5.42E+00	1.18E-01	5830484A20Rik /// LOC546062	109032 /// 546062	NM_175397 /// XM_620658
49	1423537_at	-5.39E+00	1.22E-01	Gap43	14432	NM_008083
50	1421399_at	-5.36E+00	1.23E-01	Insm1	53626	NM_016889
51	1458229_at	-5.23E+00	1.40E-01	Robo2	268902	NM_175549
52	1449468_at	-5.21E+00	1.40E-01	St6galnac5	26938	NM_012028
53	1442710_at	-5.15E+00	1.53E-01	Pdlim5	56376	NM_019808 /// NM_019809 ///

						NM_022554
54	1450492_at	-5.11E+00	1.57E-01	Cngb3	30952	NM_013927
55	1459941_at	-5.09E+00	1.57E-01	4933402J24Rik	74438	NM_028940
56	1438258_at	-5.04E+00	1.57E-01	Vldlr	22359	NM_013703
57	1454660_at	-5.02E+00	1.57E-01	1100001E04Rik	75404	XM_135854
58	1453916_at	-5.01E+00	1.57E-01	4933431K23Rik	74475	---
59	1426864_a_at	-5.00E+00	1.57E-01	Ncam1	17967	NM_010875
60	1425928_at	-4.99E+00	1.57E-01	Xkr6	219149	NM_173393
61	1442180_at	-4.93E+00	1.57E-01	Dleu7	239133	NM_173419
62	1437558_at	-4.92E+00	1.57E-01	B130021B11Rik	320860	---
63	1456002_at	-4.92E+00	1.57E-01	Xpa	22590	NM_011728
64	1458275_at	-4.91E+00	1.57E-01	2310045A20Rik	231238	NM_172710
65	1450683_at	-4.89E+00	1.57E-01	Tagln3	56370	NM_019754
66	1454245_at	-4.88E+00	1.57E-01	Csmd3	239420	XM_139502
67	1445279_at	-4.88E+00	1.57E-01	BC034902	228642	NM_177654
68	1440839_x_at	-4.87E+00	1.57E-01	1700065I16Rik	78462	---
69	1460083_at	-4.84E+00	1.58E-01	Adam10	11487	NM_007399
70	1437191_at	-4.84E+00	1.58E-01	BC050789	213234	NM_172515
71	1421958_at	-4.83E+00	1.58E-01	L1cam	16728	NM_008478
72	1422232_at	-4.78E+00	1.63E-01	Phox2b	18935	NM_008888
73	1437467_at	-4.77E+00	1.63E-01	Alcam	11658	NM_009655
74	1448628_at	-4.71E+00	1.73E-01	Scg3	20255	NM_009130
75	1435396_at	-4.68E+00	1.74E-01	C85317	97823	---
76	1455123_at	-4.68E+00	1.74E-01	St18	240690	NM_173868
77	1421359_at	-4.67E+00	1.74E-01	Ret	19713	NM_009050
78	1459249_at	-4.64E+00	1.80E-01	Tdrd3	219249	NM_172605
79	1452123_s_at	-4.62E+00	1.82E-01	Frmd4b	232288	NM_145148
80	1436931_at	-4.60E+00	1.82E-01	Rfx4	71137	NM_001024918 /// NM_027689
81	1418415_at	-4.59E+00	1.82E-01	Hoxb5	15413	NM_008268
82	1446144_at	-4.58E+00	1.82E-01	Pex2	58869	NM_021483
83	1426865_a_at	-4.56E+00	1.83E-01	Ncam1	17967	NM_010875
84	1420484_a_at	-4.54E+00	1.86E-01	Vtn	22370	NM_011707
85	1448289_at	-4.53E+00	1.89E-01	Crmp1	12933	NM_007765
86	1418452_at	-4.51E+00	1.92E-01	Gng2	14702	NM_010315
87	1452010_at	-4.49E+00	1.94E-01	Chrna3	110834	NM_145129
88	1427867_at	-4.48E+00	1.97E-01	Myh1	17879	XM_354615
89	1440005_at	-4.46E+00	1.98E-01	C730009D12	328974	---
90	1442893_at	-4.45E+00	2.00E-01	Lrrtm1	74342	NM_028880
91	1457146_at	-4.39E+00	2.17E-01	Dock4	238130	NM_172803
92	1459613_at	-4.37E+00	2.17E-01	---	---	---
93	1435165_at	-4.37E+00	2.17E-01	Cntn2	21367	NM_011531 /// NM_177129
94	1441163_at	-4.36E+00	2.17E-01	Med12l	329650	---
95	1457399_at	-4.34E+00	2.20E-01	---	---	---
96	1452357_at	-4.33E+00	2.20E-01	Gp1bb /// Sept5	14724 /// 18951	NM_001001999 /// NM_010327 /// NM_213614
97	1448759_at	-4.31E+00	2.27E-01	Il2rb	16185	NM_008368
98	1450661_x_at	-4.29E+00	2.27E-01	Nfic	18029	NM_008688
99	1431045_at	-4.29E+00	2.27E-01	D12Ertd553e	76820	NM_029758
100	1447586_at	-4.28E+00	2.27E-01	Cspg2	13003	XM_488510
101	1418105_at	-4.28E+00	2.27E-01	Stmn4	56471	NM_019675
102	1418451_at	-4.28E+00	2.27E-01	Gng2	14702	NM_010315
103	1436557_at	-4.27E+00	2.27E-01	Kb36	223915	NM_212485
104	1418877_at	-4.26E+00	2.27E-01	Foxd1	15229	NM_008242
105	1423506_a_at	-4.25E+00	2.28E-01	Nnat	18111	NM_010923 /// NM_180960
106	1422839_at	-4.25E+00	2.28E-01	Neurog2	11924	NM_009718

107	1435195_at	-4.24E+00	2.28E-01	D930046M13Rik	104880	---
108	1452399_at	-4.22E+00	2.28E-01	Rgs6	50779	NM_015812
109	1439789_at	-4.22E+00	2.28E-01	---	---	---
110	1459929_at	-4.22E+00	2.28E-01	C80731	243905	XM_145503
111	1442553_at	-4.20E+00	2.29E-01	Mapre2	212307	NM_153058
112	1422401_at	-4.20E+00	2.29E-01	Sprr3	20766	NM_011478
113	1442675_at	-4.19E+00	2.29E-01	2300006M17Rik	69466	---
114	1433551_at	-4.18E+00	2.29E-01	AI427515	270097	NM_173016
115	1453300_at	-4.17E+00	2.29E-01	Slc35d2	70484	NM_001001321
116	1436148_at	-4.17E+00	2.29E-01	---	---	---
117	1416646_at	-4.16E+00	2.30E-01	Afp	11576	NM_007423
118	1432516_at	-4.15E+00	2.30E-01	4930447F04Rik	74862	XM_135902
119	1458306_at	-4.15E+00	2.30E-01	F630022B06Rik	239827	NM_172822
120	1447929_at	-4.09E+00	2.53E-01	Ssh3	245857	NM_198113
121	1442257_at	-4.08E+00	2.53E-01	---	---	---
122	1426215_at	-4.06E+00	2.59E-01	Ddc	13195	NM_016672
123	1425111_at	-4.05E+00	2.59E-01	Sorcs3	66673	NM_025696
124	1448944_at	-4.04E+00	2.59E-01	Nrp1	18186	NM_008737
125	1428157_at	-4.03E+00	2.59E-01	Gng2	14702	NM_010315
126	1438036_x_at	-4.02E+00	2.59E-01	AW061290	381110	NM_201361
127	1435941_at	-4.02E+00	2.59E-01	Rhbdl4	246104	NM_139228
128	1454110_at	-4.01E+00	2.59E-01	4930408K08Rik	76865	---
129	1440256_at	-4.01E+00	2.59E-01	Rgs9bp	243923	NM_145840
130	1455027_at	-4.00E+00	2.62E-01	D5Bwg0860e	52822	NM_027530
131	1456121_at	-4.00E+00	2.63E-01	Lrriq2	74201	NM_028815
132	1455447_at	-3.99E+00	2.63E-01	D430019H16Rik	268595	---
133	1455033_at	-3.97E+00	2.69E-01	B430201A12Rik	329739	XM_283903
134	1442230_at	-3.96E+00	2.69E-01	---	---	---
135	1453351_at	-3.95E+00	2.70E-01	Tbx20	57246	NM_020496 /// NM_194263
136	1448233_at	-3.93E+00	2.75E-01	Prnp	19122	NM_011170
137	1417280_at	-3.92E+00	2.76E-01	Slc17a1	20504	NM_009198
138	1445737_at	-3.91E+00	2.76E-01	1700057H21Rik	73375	---
139	1418055_at	-3.90E+00	2.78E-01	Neurod4	11923	NM_007501
140	1431556_at	-3.89E+00	2.78E-01	4930435M08Rik	74651	---
141	1444080_at	-3.89E+00	2.78E-01	E430014L09Rik	320937	---
142	1420662_at	-3.87E+00	2.82E-01	4933439F18Rik	66771	NM_025757
143	1417988_at	-3.86E+00	2.82E-01	Resp18	19711	NM_009049
144	1454449_at	-3.86E+00	2.82E-01	Rabepk	227746	NM_145522
145	1415978_at	-3.85E+00	2.82E-01	Tubb3	22152	NM_023279
146	1454974_at	-3.84E+00	2.82E-01	Ntn1	18208	NM_008744
147	1440422_at	-3.84E+00	2.82E-01	Htr1f	15557	NM_008310
148	1424403_a_at	-3.84E+00	2.82E-01	D5Bwg0860e	52822	NM_027530
149	1458793_at	-3.82E+00	2.90E-01	---	---	---
150	1424719_a_at	-3.82E+00	2.90E-01	Mapt	17762	NM_010838
151	1428156_at	-3.81E+00	2.93E-01	Gng2	14702	NM_010315
152	1448991_a_at	-3.81E+00	2.93E-01	Ina	226180	NM_146100
153	1455266_at	-3.80E+00	2.94E-01	Kif5c	16574	NM_008449
154	1434670_at	-3.79E+00	2.94E-01	Kif5a	16572	NM_008447
155	1430222_at	-3.78E+00	2.94E-01	9130007G19Rik	74550	---
156	1442613_at	-3.78E+00	2.94E-01	Spon1	233744	NM_145584
157	1416301_a_at	-3.76E+00	2.94E-01	Ebf1	13591	NM_007897
158	1443086_at	-3.76E+00	2.94E-01	Alcam	11658	NM_009655
159	1443855_at	-3.76E+00	2.94E-01	Kcnc1	16502	NM_008421
160	1435323_a_at	-3.76E+00	2.94E-01	Oact1	218121	NM_153546
161	1431484_at	-3.75E+00	2.96E-01	2010007L08Rik	70114	---
162	1450779_at	-3.73E+00	3.03E-01	Fabp7	12140	NM_021272
163	1459760_at	-3.71E+00	3.13E-01	Ndufs4	17993	NM_010887

164	1440264_at	-3.70E+00	3.16E-01	---	---	---
165	1443723_at	-3.70E+00	3.16E-01	Trpm3	226025	NM_177341
166	1448250_at	-3.70E+00	3.16E-01	9030425E11Rik	71566	NM_133733
167	1450397_at	-3.69E+00	3.16E-01	Mtap1b	17755	NM_008634
168	1459840_s_at	-3.69E+00	3.16E-01	Ccdc28b	66264	NM_025455
169	1438607_at	-3.68E+00	3.16E-01	Zdhhc19	245308	NM_199309
170	1446300_at	-3.68E+00	3.16E-01	---	---	---
171	1429816_at	-3.68E+00	3.16E-01	Armc3	70882	XM_130012 /// XM_622876
172	1441118_at	-3.67E+00	3.17E-01	BC017647	216971	NM_145430
173	1432591_at	-3.66E+00	3.25E-01	Pappa	18491	NM_021362
174	1422586_at	-3.64E+00	3.31E-01	Ecel1	13599	NM_021306
175	1440991_at	-3.63E+00	3.36E-01	Plekhg1	213783	XM_136911
176	1417625_s_at	-3.63E+00	3.37E-01	Cmkor1	12778	NM_007722
177	1448099_at	-3.62E+00	3.37E-01	---	432995	XM_488763
178	1433776_at	-3.62E+00	3.37E-01	Lhfp	108927	NM_175386
179	1459186_at	-3.61E+00	3.39E-01	C80165	97726	---
180	1419524_at	-3.61E+00	3.39E-01	Tph1	21990	NM_009414
181	1435321_at	-3.60E+00	3.40E-01	3732412D22Rik	77569	NM_001001980
182	1435404_at	-3.60E+00	3.40E-01	Disp2	214240	NM_170593
183	1446237_at	-3.59E+00	3.45E-01	Akap9	100986	NM_194462
184	1442347_at	-3.59E+00	3.45E-01	Lrp8	16975	NM_053073
185	1415844_at	-3.57E+00	3.52E-01	Syt4	20983	NM_009308
186	1438042_at	-3.56E+00	3.52E-01	Shox2	20429	NM_013665
187	1435142_at	-3.56E+00	3.52E-01	Sft2d2	108735	NM_145512
188	1459702_at	-3.54E+00	3.55E-01	---	---	---
189	1457762_at	-3.54E+00	3.55E-01	Ttc15	217449	NM_178811
190	1423713_at	-3.54E+00	3.55E-01	Abcb8	74610	NM_029020
191	1456198_at	-3.54E+00	3.55E-01	1810007D17Rik	69055	---
192	1439795_at	-3.54E+00	3.57E-01	Gpr64	237175	NM_178712
193	1455037_at	-3.53E+00	3.57E-01	Plxna2	18845	NM_008882
194	1435408_at	-3.53E+00	3.57E-01	---	---	---
195	1419672_at	-3.52E+00	3.59E-01	Spock1	20745	NM_009262
196	1423280_at	-3.52E+00	3.59E-01	Stmn2	20257	NM_025285
197	1456206_at	-3.52E+00	3.59E-01	1810020C19Rik	69113	XM_130317
198	1439286_at	-3.51E+00	3.59E-01	Grik2	14806	NM_010349
199	1438169_a_at	-3.51E+00	3.60E-01	Frmd4b	232288	NM_145148
200	1445215_at	-3.50E+00	3.60E-01	2600003E23Rik	70292	NM_027373
201	1416713_at	-3.50E+00	3.60E-01	2700055K07Rik	67971	NM_026481
202	1442428_at	-3.46E+00	3.78E-01	---	---	---
203	1446324_at	-3.45E+00	3.79E-01	---	---	---
204	1439663_at	-3.45E+00	3.79E-01	Ptch1	19206	NM_008957
205	1457260_at	-3.45E+00	3.79E-01	---	---	---
206	1423274_at	-3.43E+00	3.84E-01	Ddx26	18130	NM_008715
207	1418178_at	-3.43E+00	3.84E-01	Ina	226180	NM_146100
208	1445227_at	-3.43E+00	3.84E-01	Tnfsf13	69583	NM_023517
209	1419420_at	-3.42E+00	3.86E-01	St6galnac5	26938	NM_012028
210	1428136_at	-3.41E+00	3.88E-01	Sfrp1	20377	NM_013834
211	1441764_at	-3.40E+00	3.88E-01	Prdm10	382066	XM_356146
212	1429474_at	-3.40E+00	3.88E-01	Zadh1	77219	NM_029880
213	1421851_at	-3.40E+00	3.88E-01	Mtap1b	17755	NM_008634
214	1429019_s_at	-3.40E+00	3.88E-01	Pon2	330260	NM_183308
215	1421739_a_at	-3.40E+00	3.88E-01	Matk	17179	NM_010768
216	1434502_x_at	-3.40E+00	3.88E-01	Slc4a1	20533	NM_011403
217	1450686_at	-3.39E+00	3.88E-01	Pon2	330260	NM_183308
218	1444212_at	-3.38E+00	3.90E-01	Hip2	53323	NM_016786
219	1442638_at	-3.38E+00	3.90E-01	---	---	---
220	1437086_at	-3.37E+00	3.93E-01	Ascl1	17172	NM_008553

221	1443809_at	-3.36E+00	3.93E-01	Ptn	19242	NM_008973
222	1443372_at	-3.36E+00	3.93E-01	Chat	12647	NM_009891
223	1416359_at	-3.36E+00	3.97E-01	Snag1	170625	NM_130796
224	1454507_at	-3.35E+00	4.00E-01	8430432A02Rik	71524	---
225	1417871_at	-3.35E+00	4.01E-01	Hsd17b7	15490	NM_010476
226	1441440_at	-3.34E+00	4.01E-01	Apg4c	242557	NM_175029
227	1426442_at	-3.34E+00	4.01E-01	Gpm6a	234267	NM_153581
228	1455090_at	-3.34E+00	4.03E-01	Angptl2	26360	NM_011923
229	1450164_at	-3.33E+00	4.03E-01	Ascl1	17172	NM_008553
230	1457600_x_at	-3.33E+00	4.03E-01	---	---	---
231	1423091_a_at	-3.33E+00	4.03E-01	Gpm6b	14758	NM_023122
232	1434828_at	-3.32E+00	4.03E-01	B430201A12Rik	329739	XM_283903
233	1442269_at	-3.32E+00	4.03E-01	Seh1l	72124	NM_028112
234	1455874_at	-3.32E+00	4.03E-01	1810059G22Rik	67706	NM_026325
235	1429978_at	-3.32E+00	4.03E-01	5830467E07Rik	76107	---
236	1426551_at	-3.32E+00	4.03E-01	Sidt1	320007	NM_198034
237	1459355_at	-3.31E+00	4.03E-01	Tgfbr3	21814	NM_011578
238	1448961_at	-3.31E+00	4.03E-01	Plscr2	18828	NM_008880
239	1426045_at	-3.31E+00	4.03E-01	Kng1	16644	NM_023125
240	1424402_at	-3.31E+00	4.03E-01	D5Bwg0860e	52822	NM_027530
241	1450438_at	-3.30E+00	4.04E-01	Ncam1	17967	NM_010875
242	1419009_at	-3.30E+00	4.08E-01	Actl7a	11470	NM_009611
243	1459028_at	-3.29E+00	4.11E-01	3830408P06Rik	70683	XM_125867
244	1421712_at	-3.28E+00	4.11E-01	Sele	20339	NM_011345
245	1456551_at	-3.27E+00	4.17E-01	---	---	---
246	1447389_at	-3.26E+00	4.20E-01	AI662168	102573	---
247	1441017_at	-3.26E+00	4.20E-01	Zcchc14	142682	NM_080855
248	1446355_at	-3.26E+00	4.20E-01	9530085L11Rik	320362	---
249	1445672_at	-3.26E+00	4.20E-01	---	---	---
250	1448895_a_at	-3.25E+00	4.26E-01	Ctnna2	12386	NM_009819 /// NM_145732
251	1425784_a_at	-3.24E+00	4.27E-01	Olfm1	56177	NM_019498
252	1417381_at	-3.24E+00	4.29E-01	C1qa	12259	NM_007572
253	1432873_at	-3.24E+00	4.29E-01	4932702M13Rik	74418	---
254	1442619_at	-3.23E+00	4.30E-01	Klf15	66277	NM_023184
255	1453926_at	-3.23E+00	4.30E-01	Rad54l	19366	NM_009015
256	1440525_at	-3.22E+00	4.30E-01	Acvr2b	11481	NM_007397
257	1429651_at	-3.22E+00	4.30E-01	Phactr3	74189	NM_001007154 /// NM_028806
258	1423287_at	-3.22E+00	4.30E-01	Cbln1	12404	NM_019626
259	1449757_x_at	-3.22E+00	4.30E-01	Dntt	21673	NM_009345
260	1457979_at	-3.22E+00	4.30E-01	---	---	---
261	1447497_at	-3.21E+00	4.30E-01	Rab21	216344	NM_024454
262	1440402_at	-3.21E+00	4.30E-01	4831426I19Rik	212073	NM_172500
263	1436815_x_at	-3.20E+00	4.31E-01	---	---	---
264	1445168_at	-3.20E+00	4.31E-01	---	---	---
265	1436470_at	-3.20E+00	4.31E-01	Rims2	116838	NM_053271
266	1421946_at	-3.19E+00	4.33E-01	Crp	12944	NM_007768
267	1421898_a_at	-3.18E+00	4.33E-01	Mr1	15064	NM_008209
268	1420432_at	-3.18E+00	4.33E-01	Tas1r1	110326	NM_031867
269	1460383_at	-3.18E+00	4.33E-01	Gnao1	14681	NM_010308
270	1421605_a_at	-3.18E+00	4.33E-01	Aqp9	64008	NM_022026
271	1440304_at	-3.18E+00	4.33E-01	BB214985	100485	---
272	1452537_at	-3.17E+00	4.33E-01	Igk-V1	16081	---
273	1449925_at	-3.16E+00	4.33E-01	Cxcr3	12766	NM_009910
274	1419746_at	-3.16E+00	4.33E-01	Arhgap23	58996	NM_021493
275	1458519_at	-3.16E+00	4.33E-01	---	---	---
276	1440820_x_at	-3.16E+00	4.33E-01	Tmco2	69469	---
277	1457008_at	-3.16E+00	4.33E-01	Chrnb4	108015	NM_148944

278	1427866_x_at	-3.16E+00	4.33E-01	Hbb	15127	XM_489729
279	1432006_at	-3.16E+00	4.33E-01	Ap2a2	11772	NM_007459
280	1428048_at	-3.15E+00	4.33E-01	BC027582	269472	---
281	1420190_at	-3.15E+00	4.33E-01	---	---	---
282	1440973_at	-3.15E+00	4.33E-01	LOC552874	552874	---
283	1450288_at	-3.15E+00	4.33E-01	Cdh6	12563	NM_007666
284	1443044_at	-3.15E+00	4.33E-01	A830091E24	414271	---
285	1445687_at	-3.14E+00	4.33E-01	---	380732	XM_354645
286	1421104_at	-3.14E+00	4.33E-01	Mpa2 /// LOC547126	17472 /// 547126	NM_008620 /// XM_621976
287	1431732_at	-3.14E+00	4.33E-01	Spag16	66722	NM_025728 /// NM_029160
288	1447619_at	-3.14E+00	4.33E-01	Rtn3	20168	NM_001003930 /// NM_001003933 /// NM_001003934 /// NM_053076
289	1436188_a_at	-3.14E+00	4.33E-01	Ndrg4	234593	NM_145602
290	1427157_at	-3.13E+00	4.33E-01	E030025D05Rik	216613	NM_181577
291	1431698_at	-3.13E+00	4.33E-01	Ubash3a	328795	NM_177823
292	1424029_at	-3.13E+00	4.33E-01	Tspyl4	72480	NM_030203
293	1425395_at	-3.13E+00	4.33E-01	Adam26a	13525	NM_010085
294	1459357_at	-3.13E+00	4.33E-01	Nfasc	269116	NM_182716
295	1445096_at	-3.12E+00	4.39E-01	---	---	---
296	1458061_at	-3.11E+00	4.43E-01	C130033P17Rik	402769	---
297	1440148_at	-3.10E+00	4.44E-01	Gpr6	140741	NM_199058
298	1425833_a_at	-3.10E+00	4.45E-01	Hpca	15444	NM_010471
299	1420543_at	-3.10E+00	4.45E-01	ORF28	246738	NM_138664
300	1433965_at	-3.09E+00	4.45E-01	Atp8a1	11980	NM_009727
301	1427667_s_at	-3.09E+00	4.49E-01	Tcrb-J /// LOC330294 /// LOC333791	21580 /// 330294 /// 333791	XM_284182 /// XM_289900
302	1420618_at	-3.08E+00	4.50E-01	Cpeb4	67579	NM_026252
303	1444428_at	-3.08E+00	4.50E-01	---	---	---
304	1452072_at	-3.08E+00	4.50E-01	Myct1	68632	NM_026793
305	1456536_at	-3.08E+00	4.50E-01	Thap6	381650	---
306	1431684_at	-3.07E+00	4.50E-01	4933402J24Rik	74438	NM_028940
307	1456562_x_at	-3.07E+00	4.50E-01	KIAA0415	231855	NM_172725
308	1436178_at	-3.07E+00	4.50E-01	Leprel1	210530	NM_173379
309	1457289_at	-3.07E+00	4.50E-01	Nr2e1	21907	NM_152229
310	1436025_at	-3.07E+00	4.50E-01	A430106J12Rik	108686	NM_176841
311	1443005_at	-3.07E+00	4.50E-01	Zfhx1a	21417	NM_011546
312	1445580_at	-3.06E+00	4.50E-01	Matn2	17181	NM_016762
313	1431091_at	-3.06E+00	4.50E-01	Pygo1	72135	XM_134865
314	1421483_at	-3.05E+00	4.55E-01	Lhx4	16872	NM_010712
315	1444145_at	-3.05E+00	4.55E-01	---	---	---
316	1455701_at	-3.05E+00	4.55E-01	Snx26	233071	NM_178252
317	1446525_at	-3.05E+00	4.55E-01	Gpc3	14734	NM_016697
318	1435113_x_at	-3.05E+00	4.55E-01	Stmn3	20262	NM_009133
319	1439018_at	-3.04E+00	4.58E-01	---	---	---
320	1422606_at	-3.04E+00	4.59E-01	C1qtnf3	81799	NM_030888
321	1454903_at	-3.04E+00	4.59E-01	Ngfr	18053	NM_033217
322	1445331_at	-3.03E+00	4.59E-01	---	---	---
323	1448748_at	-3.03E+00	4.59E-01	Plek	56193	NM_019549
324	1430511_at	-3.03E+00	4.59E-01	1500037O19Rik	70254	---
325	1417803_at	-3.03E+00	4.59E-01	1110032A04Rik	66183	NM_133675
326	1446327_at	-3.03E+00	4.59E-01	---	---	---
327	1448785_at	-3.02E+00	4.59E-01	Cbfa2t1h	12395	NM_009822
328	1458673_at	-3.02E+00	4.59E-01	1700093J21Rik	74296	XM_484572 /// XM_622717
329	1444784_at	-3.02E+00	4.59E-01	4930564K09Rik	78212	---

330	1428902_at	-3.01E+00	4.59E-01	Chst11	58250	NM_021439
331	1453969_at	-3.01E+00	4.59E-01	MGI:1920494	73244	NM_028477
332	1458643_at	-3.01E+00	4.59E-01	Dtx1	14357	NM_008052
333	1440363_at	-3.00E+00	4.65E-01	---	---	---
334	1454719_at	-3.00E+00	4.65E-01	5730411O18Rik	279766	NM_177370
335	1443347_at	-3.00E+00	4.65E-01	---	---	---
336	1427845_at	-2.99E+00	4.68E-01	Fgfr4	14186	NM_008011
337	1444608_at	-2.99E+00	4.70E-01	Ankrd17	81702	NM_030886 /// NM_198010
338	1450344_a_at	-2.98E+00	4.72E-01	Ptger3	19218	NM_011196
339	1439912_at	-2.97E+00	4.76E-01	9430098F02Rik	103793	---
340	1425905_at	-2.97E+00	4.76E-01	---	---	---
341	1452644_at	-2.97E+00	4.76E-01	Mbd3l2	234988	NM_144934
342	1420582_at	-2.97E+00	4.76E-01	Cd209e	170780	NM_130905
343	1433770_at	-2.97E+00	4.76E-01	Dpysl2	12934	NM_009955
344	1444778_at	-2.96E+00	4.76E-01	Sept3	24050	NM_011889
345	1432169_at	-2.96E+00	4.76E-01	4930523O13Rik	74726	---
346	1459250_at	-2.96E+00	4.76E-01	MGI:2153084	228911	NM_080455
347	1444423_at	-2.96E+00	4.76E-01	---	---	---
348	1416992_at	-2.96E+00	4.76E-01	Mfng	17305	NM_008595
349	1435280_at	-2.95E+00	4.80E-01	AI452195	105178	---
350	1440396_at	-2.95E+00	4.81E-01	Mpp6	56524	NM_019939
351	1442615_at	-2.95E+00	4.82E-01	Adk	11534	NM_134079
352	1447857_at	-2.94E+00	4.87E-01	D230005D02Rik	239188	NM_172813
353	1448812_at	-2.94E+00	4.87E-01	Hpcal1	53602	NM_016677
354	1439557_s_at	-2.93E+00	4.91E-01	Ldb2	16826	NM_010698
355	1449637_at	-2.93E+00	4.91E-01	Cdh4	12561	NM_009867
356	1445816_at	-2.93E+00	4.91E-01	---	---	---
357	1452467_at	-2.93E+00	4.91E-01	Mmab	77697	NM_029956
358	1432518_at	-2.92E+00	4.91E-01	4930430J20Rik	78109	---
359	1441628_at	-2.92E+00	4.91E-01	Diap3	56419	NM_019670
360	1438772_at	-2.92E+00	4.91E-01	Zfp367	238673	NM_175494
361	1440302_at	-2.92E+00	4.93E-01	Slc25a3	18674	NM_133668
362	1441831_x_at	-2.91E+00	4.96E-01	Clcn5	12728	NM_016691
363	1449874_at	-2.91E+00	4.96E-01	Ly96	17087	NM_016923
364	1446155_at	-2.91E+00	4.97E-01	---	---	---
365	1452751_at	-2.90E+00	5.00E-01	Ebf3	13593	NM_010096

Title refers to the gene title, Probeset_id refers to the Affymetrix probe ID. The score refers to the PowerExpress value that incorporates both fold-change and *P*-value. FDR indicates the false discovery rate, and the gene ID is given in the final column.

Table S6. The top-20 predicted enhancer elements using EEL

Rank	Gene	Coordinates	qPCR
1	Ptch1	chr13:60949059-60949576	+
2	BC017647	chr11:77879370-77881080	
3	Kif5a	chr10:127004002-127004301	
4	Ptch1	chr13:60956203-60957612	+
5	Ntn1	chr11:68031783-68031998	
6	Ngfr	chr11:95396268-95396731	
7	4831426I19Rik	chr12:100377810-100377958	+
8	Abcb8	chr5:22884482-22884701	
9	Olfm1	chr2:28152536-28152588	
10	Nkx2-9	chr12:53347501-53347755	+
11	2600003E23Rik	chr5:34368835-34369174	
12	Phox2a	chr7:95939471-95939679	
13	Rad54l	chr4:115033946-115035434	
14	Hoxb2	chr11:96176306-96176394	
15	Nkx2-9	chr12:53332902-53334901	
16	Snx26	chr7:25981542-25982431	
17	1810059G22Rik	chr19:7999574-8001182	
18	Nkx2-9	chr12:53340887-53341102	+
19	Foxa2	chr2:147493732-147495645	+
20	Mtap1b	chr13:95603717-95603819	

Enhancers also identified by location analysis are underlined. Conditions were performed as described (Hallikas et al., 2006) (λ=2, μ=0.12, ν=200, ξ=200, module length $\leqslant$ 2000, Gli score $\geqslant$25) except that we required EEL score $\geqslant$100 instead of $\geqslant$500, and the predictions were ranked by combined Gli score instead of by total EEL score. We chose our current criteria since they resulted in the inclusion of validated enhancers in Ptch1 and Nkx2-9 into the top 20 predictions. Enhancers were tested by qPCR on LMPCR amplified ChIP tissue.

Table S7. Genome-wide screen of Gli target genes in mouse using MCA

Rank	Gene	Rank	Gene	Rank	Gene
1	Zswim2	18	1700011F03Rik	35	Ntn4
2	Ppp3ca	19	**Gpc3**	36	Cntnap2
3	Lrfn5	20	Utrn	37	Nol4
4	Ncam1	21	Nfia	38	Camk2g
5	**Hhip**	22	Gpc5	39	**Nkx2-9**
6	**Ptch1**	23	Eda	40	Ntrk2
7	A2bp1	24	1700108E19Rik	41	Ahdc1
8	Stk32a	25	2810426N06Rik	42	1700022L09Rik
9	Odz2	26	Arid5b	43	Foxf1a
10	Bcas3	27	Adamts2	44	BC052066
11	Srisnf2l	28	Tex264	45	Pitpnb
12	Mapk4	29	Xlr3a	46	Ndrg3
13	Ntng2	30	Prkag1	47	Sh2d1b
14	Opcm1	31	Dcbld1	48	Ank2
15	Reln	32	Gopc	49	Mar1
16	Grm3	33	Nrf1	50	Prkce
17	Ctnna2	34	**Robo2**	51	Atp5h

Genes are ranked by combined Gli-binding affinity. The neural targets that were subsequently validated by qPCR are denoted by bold, underlined text.

Table S8. Top-20 Gli enhancers identified by MCA present in the pool of EB genes upregulated by Hh treatment

Rank	MCA	Coordinates	qPCR
1	Ncam1	chr9:49835328-49835543	
2	Ebf3	chr7:131683012-131684612	
3	Nkx2-9	chr12:53347461-53347776	+
4	Ptch1	chr13:60948913-60949595	+
5	Gpm6b	chrX:159911226-159911457	
6	Gpc3	chrX:47041532-47041785	
7	Robo2	chr16:73388698-73389494	+
8	Gpc3	chrX:47086023-47086378	+
9	Ret	chr6:118619345-118619637	+
10	Neurog2	chr3:126511586-126512041	
11	Ebf1	chr11:44728066-44728630	
12	Ctnna2	chr6:77576792-77577343	
13	Ptch1	chr13:60951919-60952536	+
14	Ctnna2	chr6:77525938-77526505	
15	Ptch1	chr13:60955950-60956548	+
16	Nkx2-9	chr12:53340885-53341147	+
17	Ascl1	chr10:87473182-87474379	
18	Hoxb	chr11:96102714-96103772	
19	Trpm3	chr19:21679064-21679850	
20	FoxA2	chr2:147493773-147494127	+

Note that the Gpc3 enhancers predicted here differ from a validated enhancer present in the mesenchyme ventral to the neural tube. This target, identified in a previous study using EEL (Hallikas et al., 2006), is not present as a top-20 target in Table S6 because it is sorted based on Gli score instead of total EEL score.

Table S9. Summary of qPCR-validated Gli-binding regions and procedures that identified sites

Gene	Coordinates	Identification
FoxA2 – peak1	chr2:147493701-147494245	ChIP-on-chip, MCA, EEL
Gli1 – peak1	chr 10:127076571-127078504	ChIP-on-chip
Gpc3	chr X:47086023-47086378	MCA
Hhip – peak1	chr 8:79270092-79270492	ChIP-on-chip
Hhip – peak2	chr 8:79268419-79268867	ChIP-on-chip
Nkx 2-1 – peak1	chr 12:53245748-53246351	ChIP-on-chip
Nkx2-2 – peak1	chr 2:146644626-146645570	ChIP-on-chip
Nkx2-9 – peak1	chr 12:53340519-53341015	ChIP-on-chip, MCA, EEL
Nkx2-9 – peak2	chr 12:53347310-53348354	ChIP-on-chip, MCA, EEL
Ptch1 – peak1	chr 13:60945578-60945972	ChIP-on-chip
Ptch1 – peak2	chr 13:60949397-60950371	ChIP-on-chip, MCA, EEL
Ptch1 – peak3	chr 13:60951377-60952953	ChIP-on-chip, MCA
Ptch1 – peak4	chr 13:60953950-60955072	ChIP-on-chip
Ptch1 – peak5	chr 1360956284-60956777	ChIP-on-chip, MCA, EEL
Ptch2 – peak1	chr 4:116057320-116057946	ChIP-on-chip
Ptch2 – peak2	chr 4:116054385-116054954	ChIP-on-chip
Rab34 – peak1	chr 11:77915220-77915560	ChIP-on-chip
Ret	chr 6:118619345-118619637	MCA
Robo2	chr 16:73388698-73389494	MCA
4831426I19Rik	chr 12:100377810-100377958	EEL

All Gli1-binding sites identified in this study that showed enrichment in additional qPCR-based ChIP assays are listed along with the identification methods that identified each site (see results). When more than one method identified overlapping coordinates, the values refer to the ChIP-on-chip coordinates defined in Table 1 (other coordinates refer to Tables S6 and S8). Note that several sites were identified only by the EEL or MCA algorithms; in all cases, these regions were not present on the custom tiling array.