nature genetics

Regulation of a remote *Shh* forebrain enhancer by the Six3 homeoprotein

Yongsu Jeong[1,7], Federico Coluccio Leskow[1,7], Kenia El-Jaick[2], Erich Roessler[2], Maximilian Muenke[2], Anastasia Yocum[3], Christele Dubourg[4], Xue Li[5], Xin Geng[6], Guillermo Oliver[6] & Douglas J Epstein[1]

In humans, *SHH* haploinsufficiency results in holoprosencephaly (HPE), a defect in anterior midline formation[1,2]. Despite the importance of maintaining *SHH* transcript levels above a critical threshold, we know little about the upstream regulators of *SHH* expression in the forebrain. Here we describe a rare nucleotide variant located 460 kb upstream of *SHH* in an individual with HPE that resulted in the loss of Shh brain enhancer-2 (SBE2) activity in the hypothalamus of transgenic mouse embryos. Using a DNA affinity-capture assay, we screened the SBE2 sequence for DNA-binding proteins and identified members of the Six3 and Six6 homeodomain family as candidate regulators of *Shh* transcription. Six3 showed reduced binding affinity for the mutant compared to the wild-type SBE2 sequence. Moreover, Six3 with HPE-causing alterations failed to bind and activate SBE2. These data suggest a direct link between Six3 and *Shh* regulation during normal forebrain development and in the pathogenesis of HPE.

Shh expression must be regulated in a temporally and spatially restricted manner in order for it to fulfill its multiple functions during forebrain and craniofacial development (reviewed in refs. 3,4). Three tissues, including the prechordal plate, ventral forebrain and facial ectoderm, have been identified as critical sources of *Shh* that promote distinct aspects of ventral forebrain and craniofacial morphogenesis[5–8]. Interference with Shh signaling from any of these sites results in HPE, a spectrum of brain and craniofacial malformations, the severity of which correlates with the timing of Shh perturbation[6,9,10]. In humans, *SHH* haploinsufficiency is the predominant cause of HPE, indicating that the level of *SHH* expression is important for proper forebrain and craniofacial development[1]. Several downstream effectors of the SHH and NODAL signaling pathways have also been identified as targets of mutation in HPE, whereas mutations in *SIX3* cause HPE through poorly defined mechanisms[2]. Although much is known about the signal transduction pathway functioning downstream of Shh, relatively little is known of the genes acting upstream in the pathway regulating *Shh* transcription in key signaling centers mediating forebrain and craniofacial development.

Previous efforts to address this issue have focused on determining the genomic location of functional *Shh* regulatory elements[11]. These experiments have identified six enhancers distributed over a 500-kb interval surrounding the *Shh* gene that directed reporter activity to most areas of *Shh* expression in the mouse central nervous system, including the ventral forebrain (**Fig. 1a**). In particular, the highly conserved Shh brain enhancer-2 (SBE2), located 460 kb upstream of the *SHH* coding sequence, has been identified as unique in its ability to regulate *Shh*-like expression throughout the hypothalamus.

To identify functionally relevant nucleotides in SBE2, we screened the 1.1-kb sequence for mutations in humans with HPE. We reasoned that HPE-associated variants in SBE2 could aid in identifying critical *cis* and *trans* determinants of *SHH* expression in the forebrain. By analyzing 474 individuals with HPE, we identified one individual who was heterozygous for a cytosine to thymine base change at nucleotide position 444 (g.444C>T) of the enhancer sequence (human chr7:155,754,267; NCBI build 36.3). The C/T variant is situated within a block of 10 nucleotides that have been maintained in human, mouse, chicken and frog for over 350 million years (**Fig. 1a**). We did not observe this C/T nucleotide variant in DNA samples from 450 unrelated control individuals. The affected female showed features of semilobar HPE including microcephaly, midfacial hypoplasia, cleft-lip and palate, diabetes insipidus and moderate fusion of the hypothalamus and basal ganglia. The parents' genotype revealed that the father is an unaffected carrier, and the mother is homozygous for the wild-type SBE2(C) allele. It is known that approximately 30% of individuals heterozygous for loss-of-function mutations in *SHH* show no evidence of HPE[12]; that is, these mutations are often nonpenetrant.

[1]Department of Genetics, University of Pennsylvania School of Medicine, 415 Curie Boulevard, Philadelphia, Pennsylvania 19104, USA. [2]Medical Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA. [3]Department of Pharmacology, University of Pennsylvania School of Medicine, 421 Curie Boulevard, Pennsylvania 19104, USA. [4]Génétique et Développement, Centre National de la Recherche Scientifique UMR 6061, Université de Rennes1, 2 avenue du Pr Léon Bernard, 35043 Rennes Cedex, France. [5]Department of Surgery/Urology, Children's Hospital of Boston, Harvard Medical School, 300 Longwood Ave., Boston, Massachusetts 02115, USA. [6]Department of Genetics, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [7]These authors contributed equally to this work. Correspondence should be addressed to D.J.E. (epsteind@mail.med.upenn.edu).

Received 11 June; accepted 22 July; published online 5 October 2008; doi:10.1038/ng.230

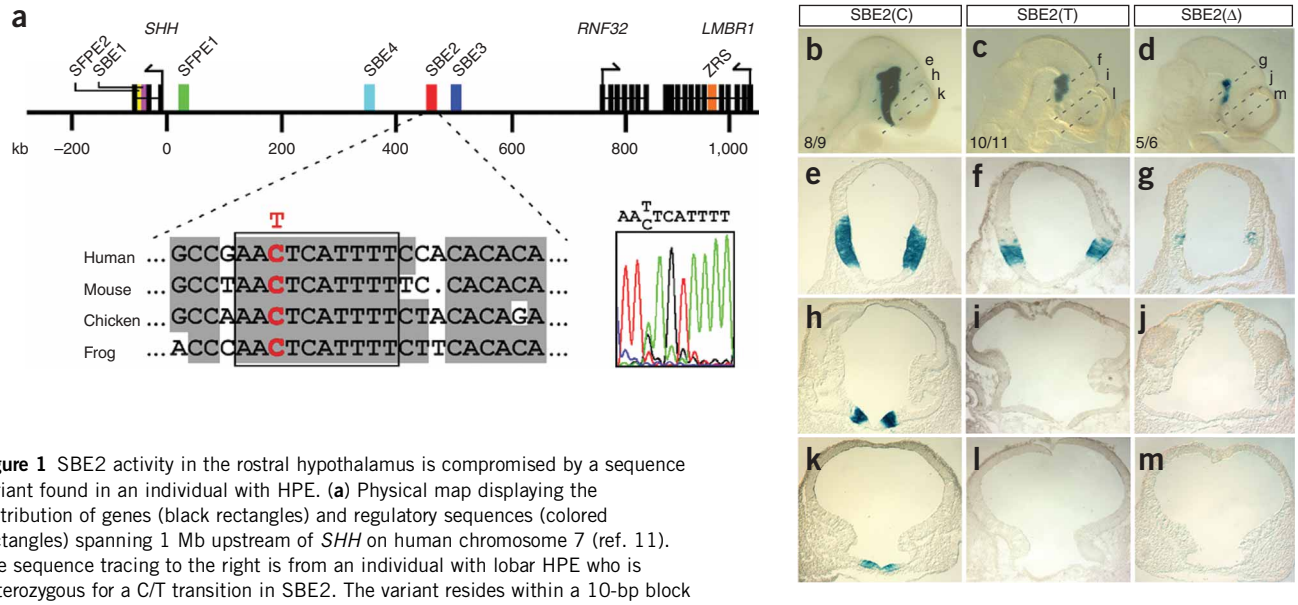

Figure 1 SBE2 activity in the rostral hypothalamus is compromised by a sequence variant found in an individual with HPE. (**a**) Physical map displaying the distribution of genes (black rectangles) and regulatory sequences (colored rectangles) spanning 1 Mb upstream of *SHH* on human chromosome 7 (ref. 11). The sequence tracing to the right is from an individual with lobar HPE who is heterozygous for a C/T transition in SBE2. The variant resides within a 10-bp block of SBE2 sequence that was 100% conserved in human, mouse, chicken and frog (red base in boxed sequence alignment). (**b**–**m**) X-gal staining of representative embryos carrying wild-type SBE2(C) (**b**,**e**,**h**,**k**), mutant SBE2(T) (**c**,**f**,**i**,**l**) or a 10-bp deletion SBE2(Δ) (**d**,**g**,**j**,**m**) at E10.5. Dashed lines in **b**–**d** indicate the planes of section shown in **e**–**m**. The number of embryos showing representative reporter activity over the total number of transgenic embryos is indicated for each construct in **b**–**d**. SBE, Shh brain enhancer; SFPE, Shh floor plate enhancer; ZRS, zone of polarizing activity regulatory sequence.

Thus, the finding that the carrier father is unaffected does not discount the possibility that SBE2(T) confers an increased risk of HPE. As mutations in known HPE-associated genes were not detected in the affected female, we sought to determine whether the single-nucleotide change could alter SBE2 activity and thus provide a molecular basis for her phenotype.

We tested human SBE2 sequences containing either the wild-type SBE2(C) or variant SBE2(T) residue for their ability to drive *lacZ* expression in transgenic mouse embryos. Embryos carrying the wild-type SBE2(C) reporter construct showed little variability in the spatial distribution of X-gal staining, recapitulating *Shh* expression in the hypothalamus from the mammillary region caudally to the preoptic area rostrally ($n = 8$ of 9; **Fig. 1b**,**e**,**h**,**k**). In contrast, embryos carrying SBE2(T) consistently showed a loss of reporter activity from the level of the optic vesicles to the rostral extent of the diencephalon ($n = 10$ of 11; **Fig. 1c**,**f**,**i**,**l**). The reduction in SBE2 reporter activity was similar in embryos carrying SBE2(Δ), a construct in which the highly conserved 10-bp sequence overlapping the C/T substitution was deleted ($n = 5$ of 6; **Fig. 1d**,**g**,**j**,**m**). Notably, the area of the ventral diencephalon that showed decreased X-gal staining in embryos carrying SBE2(T) correlated with the sites of malformation shown by the individual with HPE. The anterior region of the ventral diencephalon is an important source of Shh for the development of the face and pituitary gland[10,13,14].

On the basis of these findings, we hypothesized that the conserved 10-bp SBE2 sequence serves as a binding site for a transcriptional regulator whose function is required for the activation of *Shh* expression in the anteroventral portion of the hypothalamus. In the presence of the SBE2(T) variant, assembly of this transcriptional activation complex is compromised, likely resulting in the reduction of *Shh* expression below a critical threshold. The analysis of the 10-bp sequence in question did not reveal informative transcription factor binding sites in the TRANSFAC database (see URLs section in Methods). Therefore, we sought to identify the putative SBE2-binding protein using a DNA affinity-capture assay[15].

We first incubated a biotinylated 18-bp double-stranded SBE2 probe with nuclear extracts prepared from adult mouse brain. Next,

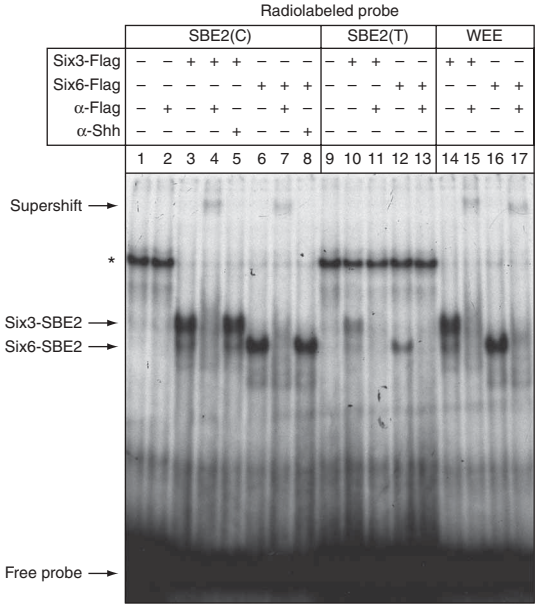

Figure 2 Six3 and Six6 proteins bind directly to SBE2. EMSAs done with COS-1 cell extracts transfected with Flag-tagged Six3 (lanes 3–5,10,11) or Six6 (lanes 6–8,12,13) expression vectors and incubated with SBE2(C) (lanes 1–8), SBE2(T) (9–13) or WEE (14–17) radiolabeled probes. Specific protein-DNA complexes were supershifted in the presence of an antibody to Flag (α-Flag; lanes 4,7,15,17) but not a nonspecific antibody (lanes 5,8). Note that in addition to the supershift, incubation with the antibody to Flag also disrupted Six3/Six6-SBE2 and Six3/Six6-WEE complex formation (lanes 4,7,11,13,15,17). The asterisk indicates the formation of a nonspecific complex that is more effectively competed away in the presence of Six3 and Six6 and high-affinity probes.

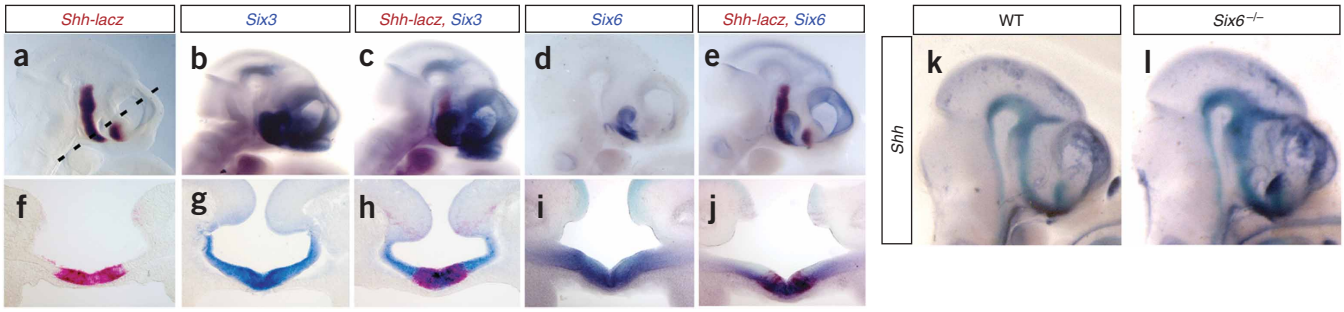

Figure 3 Overlap of *Shh* and *Six3* and *Six6* expression in the ventral diencephalon. (**a**–**j**) Whole mount (**a**–**e**) and transverse sections (**f**–**j**) showing the colocalization of *Shh-lacZ* (salmon-gal), *Six3* and *Six6* (alkaline phosphatase) expression in the mouse embryonic forebrain at E10.5. The expression of *Shh* was monitored using a lacZ reporter line (447L17βlacZ) that recapitulates endogenous *Shh* expression in the hypothalamus in an SBE2-dependent manner[11]. (**k**–**l**) *Shh* expression (alkaline phosphatase) in wild-type and *Six6*$^{-/-}$ embryos at E10.5.

we pulled down DNA-protein complexes with streptavidin-coated agarose beads and analyzed their protein content by mass spectrometry (see Methods). To control for nonspecific DNA-binding proteins, we compared SBE2-extracted proteins to those pulled down with an SBE2 probe containing multiple nucleotide mismatches in highly conserved residues. Only DNA-binding proteins specific for SBE2 were considered further. Of the six transcription factors identified (**Supplementary Table 1** online), the one of greatest interest was Six6, a homeodomain-containing protein belonging to the optix family of transcriptional regulators that includes Six3, a protein implicated in HPE[16,17].

To validate the binding of Six3 and Six6 proteins to SBE2, we carried out electrophoretic mobility shift assays (EMSAs). COS-1 cell lysates transfected with Flag-tagged versions of full-length Six3 and Six6 formed specific complexes when incubated with a radiolabeled SBE2(C) probe (**Fig. 2**, lanes 1,3,6). These protein-DNA complexes were supershifted when exposed to an antibody to Flag (α-Flag) (**Fig. 2**, lanes 2,4,7), but not a nonspecific antibody (**Fig. 2**, lanes 5,8), indicating that the binding of Six3 and Six6 to SBE2 was direct. Specific complexes formed with similar mobility when we used an SBE2(T) probe; however, the intensity of the bands was noticeably weaker (**Fig. 2**, lanes 9–13). We also observed similar protein-DNA

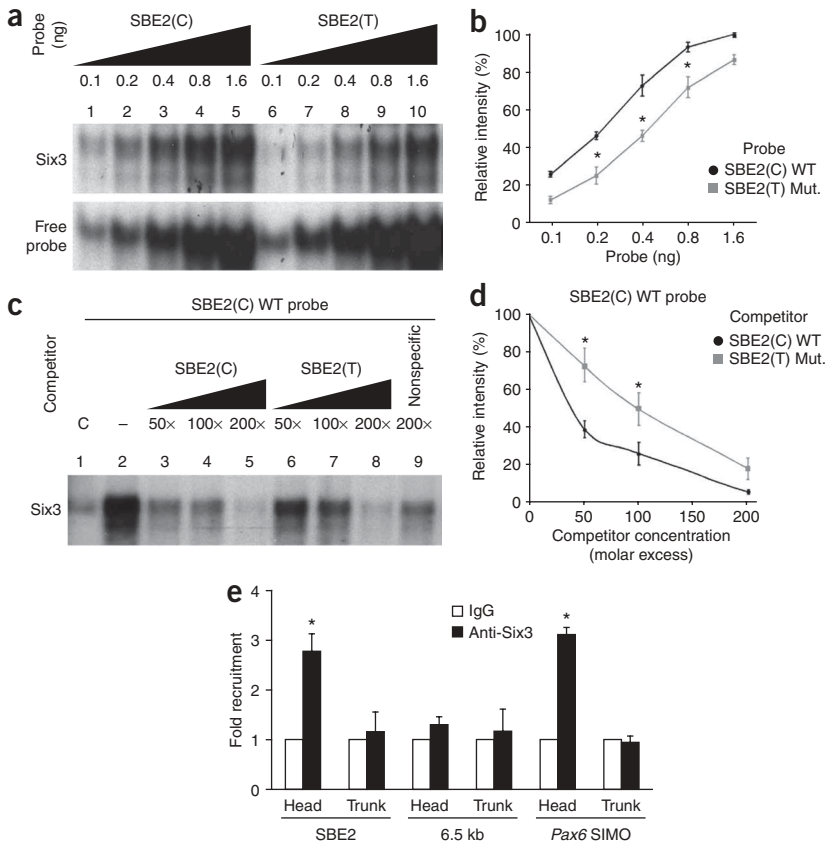

Figure 4 Six3 binds SBE2(C) with higher affinity than SBE2(T). (**a**) EMSAs done with increasing amounts of radiolabeled SBE2(C) (lanes 1–5) and SBE2(T) (lanes 6–10) probes incubated with COS-1 cell lysates transfected with pCDNA3-Six3-Flag. (**b**) Dose response curves for data shown in **a**. Each point along the curve is the average band intensity from three independent experiments (*$P < 0.05$, Student's *t*-test). (**c**) Competitive EMSAs showing the binding of Six3 to a radiolabeled SBE2(C) probe. COS-1 cell lysates transfected with pCDNA3-Flag (lane 1) or pCDNA3-Six3-Flag (lanes 2–8) were analyzed for binding to a 33-bp probe overlapping wild-type SBE2(C). Increasing concentrations of cold wild-type SBE2(C) competitor (lanes 3–5) were more efficient at displacing radiolabeled SBE2(C) probe from Six3, compared to increasing concentrations of cold SBE2(T) competitor (lanes 6–8). A nonspecific probe (lane 9) did not significantly alter the shifted complex. C, control. (**d**) Graphical representation of the data in **c**. The relative intensities of the retarded bands were quantified and plotted against competitor concentration. Each data point on the curve is an average of five independent experiments. At subsaturating concentrations of competitor (50×, 100×), SBE2(C) (black line) was significantly more effective at interfering with complex formation than SBE2(T) (gray line) (*$P < 0.05$, Student's *t*-test). (**e**) ChIP from embryos using antibodies to Six3 or IgG. QPCR results from three independent experiments reveal a significant enrichment of SBE2 DNA in Six3- versus IgG-bound chromatin isolated from forebrain but not posterior trunk regions of E8.75 mouse embryos (*$P < 0.01$, Student's *t*-test). A negative control sequence, 6.5 kb downstream of SBE2, was not enriched in Six3-bound chromatin, whereas a positive control sequence, *Pax6* SIMO, was enriched to a similar degree as SBE2. Error bars, s.d.

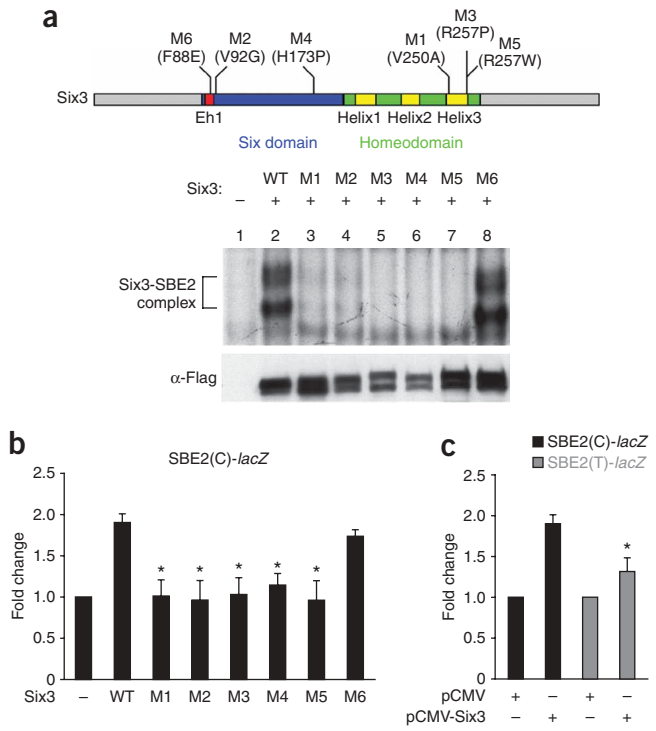

Figure 5 HPE-causing mutations in Six3 alter binding and activation of SBE2. (**a**) Top: Schematic of Six3 protein showing the location of amino acid substitutions resulting from five different HPE-causing point mutations (M1–M5) affecting either the Six domain or homeodomain. An additional mutation (M6) in the Groucho interaction domain interferes with Six3 repressor activity[19] but was not identified in individuals with HPE. Bottom: Cell lysates transfected with pCDNA3-Flag (lane 1), pCDNA3-Six3-Flag (WT, lane 2) or pCDNA3-mutant Six3-Flag (M1–M6, lanes 3–8) were analyzed for binding to a 33-bp probe overlapping wild-type SBE2(C). The Six3-SBE2 complex is indicated with a bracket. Weak or no complex formation was observed for M1 and M2 (lanes 3,4) and M3–M5 (lanes 5–7), respectively, whereas DNA binding activity was retained by M6 (lanes 8). α-Flag immunoblot demonstrates that wild-type and mutant Six3 proteins were expressed at equivalent levels. (**b**) SBE2(C)-*lacZ* activation is compromised by mutations in Six3. Wild-type and F88E (M6) forms of Six3 activated reporter expression, whereas the other Six3 mutants (M1–M5) showed reduced capacity to stimulate SBE2 (black bars). Each bar represents an average of three replicates. Asterisk indicates significant difference from wild-type ($P < 0.001$). (**c**) Six3 regulates SBE2 activity in Cos-1 cells. pCMV-Six3 stimulated wild-type SBE2(C)-*lacZ* expression (black bars), compared to the empty expression plasmid. This transcriptional activation by Six3 was attenuated in cells expressing SBE2(T)-*lacZ* (gray bars). Each bar represents an average of six replicates. Error bars, s.d.

complexes using a probe overlapping a *Wnt1* enhancer element (WEE) that was previously shown to contain a consensus Six3 binding motif[18,19] (**Fig. 2**, lanes 14–17). These results confirm the existence of a unique Six3 and Six6 binding site in SBE2.

Six3 and *Six6* show dynamic patterns of expression in the developing forebrain[16,20]. For Six3 or Six6 to be considered a direct regulator of *Shh*, their spatial and temporal expression profiles should overlap. In comparing *Six3* and *Six6* expression with SBE2-dependent *Shh-lacZ* reporter activity, we noted that in the ventral forebrain, from the level of the optic vesicles to the rostral extent of the diencephalon, *Shh-lacZ* was embedded within the *Six3* and *Six6* expression domains (**Fig. 3**). Both *Six3* and *Six6* showed a broad distribution throughout the ventral portion of the anterior hypothalamus, whereas *Shh-lacZ* expression was restricted medially within the *Six3* and *Six6* domains (**Fig. 3f–j**). Of note, the region of overlap between *Shh-lacZ* and *Six3* and *Six6* is precisely where mutant SBE2(T) reporter activity was diminished (**Fig. 1c**).

We next determined whether Six6 is required to regulate *Shh* expression. *Six6*$^{-/-}$ mouse embryos have reduced proliferation of retinal and pituitary progenitors but do not show overt signs of HPE[21]. Consistent with this milder phenotype, *Shh* expression was unaffected in the ventral forebrain of *Six6*$^{-/-}$ embryos (**Fig. 3k,l**). In contrast, *Six3*$^{-/-}$ mouse embryos show severe forebrain truncations, including in rostral regions of the diencephalon[18]. Moreover, the combination of findings that mutations in *SIX3* cause HPE in humans[17] and that mouse embryos carrying a knock-in allele of an HPE-causing point mutation in *Six3* show reduced *Shh* expression in the forebrain[22] is consistent with an essential role for Six3 in the direct regulation of *Shh* transcription.

Our finding that the DNA sequence containing SBE2(C) functions as a Six3 and Six6 binding site raised the possibility that the SBE2(T) variant interferes with the recruitment of Six3 to this site. To test this hypothesis, we evaluated the affinity of Six3 for SBE2(C) compared to SBE2(T). We generated a dose-response curve by varying the amount of radiolabeled probe exposed to a constant amount of Six3 protein and quantifying band intensity as a measure of Six3-SBE2 complex formation. SBE2(C) consistently showed a stronger association with Six3 compared to SBE2(T) at all doses of probe tested (**Fig. 4a,b**). Similar results were observed with Six6 (data not shown).

We also quantified the amount of radiolabeled probe displaced from Six3 in the presence of increasing amounts of unlabeled (cold) competitor. In the presence of 50- and 100-fold excess wild-type SBE2(C) unlabeled competitor, the majority (60% and 75%, respectively) of the radiolabeled SBE2(C) probe was displaced from Six3 (**Fig. 4c**, lanes 2–4, and **Fig. 4d**). In comparison, when 50- and 100-fold excess SBE2(T) unlabeled competitor was introduced, significantly less (27% and 50%, respectively) of the radiolabled SBE2(C) probe was displaced from Six3 (**Fig. 4c** lanes 2,6,7 and **Fig. 4d**). We obtained similar results when we used SBE2(T) as the radiolabeled probe (data not shown). These results indicate that the SBE2(T) variant weakens the affinity of Six3 binding by approximately twofold in relation to SBE2(C) (**Fig. 4b,d**).

We also confirmed Six3 binding to SBE2(C) *in vivo* by chromatin immunoprecipitation (ChIP). SBE2 was significantly enriched in Six3-bound chromatin isolated from forebrain compared to posterior trunk regions of E8.75 mouse embryos (**Fig. 4e**). A control sequence, 6.5 kb downstream of SBE2, was not enriched in Six3-bound chromatin. The extent of Six3 binding to SBE2 is comparable to that of another Six3 target sequence identified in the *Pax6* SIMO lens enhancer[23] (**Fig. 4e**). Taken together, these data are consistent with our hypothesis that Six3 is a direct regulator of SBE2 activity, and that the SBE2(T) variant compromises the recruitment of Six3 and subsequent activation of *Shh* transcription. Although our data suggest that the SBE2(T) variant may cause HPE, formal proof of this claim must await the evaluation of mice carrying a targeted knock-in of SBE2(T) into the mouse genome.

Six3 functions as a context-dependent activator or repressor of target gene expression in the developing eye and forebrain[18,19,23–25]. In addition, Six3 promotes the proliferation of forebrain progenitors by antagonizing geminin, a DNA replication inhibitor[26]. This aspect of Six3 function is independent of its DNA-binding properties. Hence, the mechanism by which individuals carrying mutations in *SIX3* develop HPE may be due to heightened geminin function, improper

regulation of target gene expression, or both. To determine whether HPE-causing mutations in Six3 impair the ability of the mutant proteins to bind SBE2, we carried out EMSAs. COS-1 cell extracts expressing equivalent amounts of wild-type or mutant forms of mouse Six3 were incubated with radiolabeled SBE2(C) probe (**Fig. 5a**). Three independent substitutions in the homeodomain either reduced (V250A) or prevented (R257P, R257W) Six3-SBE2 complex formation (**Fig. 5a**, lanes 2,3,5,7). Unexpectedly, two of three substitutions in the Six domain also showed greatly reduced (V92G) or absent (H173P) Six3-SBE2 complex formation (**Fig. 5a**, lanes 2,4,6). Only the F88E substitution in the Groucho interaction domain retained the ability to bind SBE2 (**Fig. 5a**, lanes 2,8). The Six3 mutants that failed to bind SBE2(C) were also impaired in their ability to stimulate SBE2(C)-*lacZ* expression in Cos-1 cells (**Fig. 5b**). Similarly, Six3-dependent activation of SBE2(T)-*lacZ* expression in COS-1 cells was significantly reduced compared to that of SBE2(C)-*lacZ* (**Fig. 5c**). How mutations affecting the Six domain interfere with the DNA-binding properties of Six3 is unclear, but may indicate a previously unappreciated interaction between the Six domain and the homeodomain. On the basis of these results, we deem it likely that the mechanism by which HPE manifests in individuals carrying point mutations in *SIX3* is due, in part, to a failure in the binding of *SHH* regulatory sequences and subsequent activation of *SHH* transcription in the ventral forebrain.

Results from this study provide a better understanding of the transcriptional control mechanisms regulating *SHH* expression during normal forebrain development and in the pathogenesis of HPE. Our data suggest that Six3 is a direct regulator of *Shh* expression in the anterior diencephalon. Moreover, the approaches taken in assigning function to a putative HPE-causing variant in a remote *SHH* regulatory element should be generally applicable for studying the growing number of rare, as well as common, regulatory SNPs that modulate gene expression in normal and disease states[27–29].

METHODS

Sequence analysis of SBE2. We screened genomic DNA from 474 individuals with sporadic or familial HPE registered at the National Institutes of Health (NIH) and 450 normal controls for mutations in SBE2. Genomic DNA was extracted from either lymphocytes or established lymphoblastoid cell lines by routine methods. All samples were obtained by informed consent according to the guidelines of the National Human Genome Research Institute Institutional Review Board.

For PCR amplification and direct sequencing, we designed one pair of primers (FBE) to amplify the 1.1-kb SBE2 region (**Supplementary Table 2** online). Sequencing was done using five internal primers (available upon request), allowing sequence reads of both strands. Amplification of genomic DNA was done in a 35 µl reaction volume, using 60–100 ng DNA template, 3.5 µl of 10× PCR Amplification Buffer (Invitrogen), 1.75 µl of PCR Enhancer solution (Invitrogen), 1 µl of 50 mM $MgSO_4$ (Invitrogen), 0.3 µl of 25 mM dNTP stock mixture (Amersham Biosciences), 1 µl of each 20 pmol primer (Invitrogen) and 0.5µl of AmpliTaq 5U/µl (Applied Biosystems). The PCR cycling parameters used for amplification were 95 °C for 4 min followed by 30 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min, and a final extension of 72 °C for 7 min. We used the Big Dye terminator cycle sequencing kit 3.1 (Applied Biosystems) for direct DNA sequencing and analyzed the reactions on an ABI 3100 Genetic Analyzer.

DNA affinity-capture assay. We preincubated 500 µl of brain nuclear extracts (2 µg/µl) prepared from adult mice (Sigma NuCLEAR Extraction kit) for 20 min at 4 °C in 300 µl of binding buffer (10 mM Tris-Cl pH 7.4, 50 mM NaCl, 1 mM EDTA, 1 mM DTT and 5% glycerol), 25 µl of dIdC (200 ng/µl), 25 ml of BSA (1µg/µl) and 550 µl of H_2O. The extracts were precleaned with 100 µl of a 50% slurry of streptavidin agarose beads (Invitrogen) at 4 °C for 30 min with rotation, centrifuged at high speed for 30 s and transferred (supernatant) into a new tube. A double-stranded DNA oligonucleotide (18-mer) overlapping a conserved 10-bp segment of SBE2 (**Supplementary Table 2**) was synthesized with a terminal 5′ biotin modification (Invitrogen) and incubated with the brain nuclear extracts for 3 h at 4 °C with rotation. We also used an SBE2 mismatch oligo (**Supplementary Table 2**) in a parallel experiment as a negative control for nonspecific binding proteins. We added 50 µl of a 50% slurry of streptavidin agarose beads to the mixture for 30 min at 4 °C with rotation. The beads were centrifuged, washed twice in binding buffer, twice in wash buffer (10 mM Tris-Cl pH 7.4, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1% Nonidet P-40) and twice in PBS. The beads were stored at −20 °C until trypsin digestion and subjected to reversed-phase liquid chromatography/tandem mass spectrometry analysis at the University of Pennsylvania proteomics core facility. The raw mass spectrometry data were submitted to Bioworks Browser (Thermo Electron) and batch searched through Turbo-SEQUEST against an indexed mouse RefSeq database (version updated December 2004; details available upon request).

Electromobility shift assays (EMSA). We transfected pCDNA3-Flag, pCDNA3-Six3-Flag, pCDNA3-(M1–6)-Six3-Flag and pCDNA3-Six6-Flag plasmids into COS-1 cells using FuGENE 6 transfection reagent (Roche). After 48 h, whole-cell lysates were prepared in a buffer containing 50 mM Tris-HCl pH 7.4, 1 mM EDTA, 1 mM DTT, protease inhibitor cocktail and 25% glycerol. For EMSA, we incubated 10 µg of protein from the cell lysates for 10 min at room temperature (25 °C) in a DNA-binding buffer containing 10 mM Tris-HCl pH 7.4, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, 5% glycerol, 200 ng poly(dI-dC) and 1 µg BSA in the presence or absence of competitor double-stranded oligonucleotides. After 0.1 ng (5×10^4 to 10×10^4 c.p.m.) of probe was added to the mixture, incubation was continued for an additional 20 min. Supershifts were done by incubating protein-DNA mixtures with 0.5 µl of mouse monoclonal antibody to Flag M2 (F3165; Sigma-Aldrich), or rabbit polyclonal antibody to Six3 (G. Oliver, St. Jude Children's Hospital) for 5 min before gel loading. We separated protein-bound DNA complex from free probe on a 6.5% acrylamide gel run in 1× TBE (Tris-borate-EDTA) buffer. The nucleotide sequences of the sense strand of probes and competitors are listed in **Supplementary Table 2**. Competitive EMSAs were done by incubating protein-DNA complexes with 50, 100 or 200 molar excess of cold probe. After overnight exposure, autoradiographs were scanned and the intensity of individual bands corresponding to the protein-DNA complex of interest was quantified using NIH ImageJ. Values were plotted as the ratio (percentage) of band intensities in the presence and absence of specific competitor and compared using the Student's *t*-test.

Transient transfection and dual reporter assay. We seeded COS-1 cells at 50–70% confluence and transfected them using Fugene 6 (Roche Applied Science) according to the manufacturer's instructions. We mixed 20 ng of pRL-TK vector (Promega), which contains the *Renilla* luciferase gene as a transfection efficiency control, and 500 ng of SBE2(C) or SBE2(T) LacZ reporter plasmids with 125 ng of empty pCMV, pCMV-Six3, or pCMV-mutant Six3 (M1–M6). Lysates were prepared 36 h after transfection by adding 100 µl of lysis solution (Dual-Light System; Applied Biosystems). We determined β-galactosidase activity by the accumulated product of Galacton-Plus substrate reaction (Applied Biosystems), and normalized it to that produced by *Renilla* luciferase. Enhancer activity was expressed as fold induction relative to that of cells transfected with the empty pCMV vector. We did at least three independent experiments for each construct.

Chromatin immunoprecipitation. The chromatin immunoprecipitation (ChIP) assay followed a modified version of a previously described protocol[30]. Briefly, pooled embryos at the 13- to 17-somite stage were fixed with 1% formaldehyde for 15 min with shaking. After a 5 min incubation with 100 mM glycine, heads and trunks were separated and disrupted in lysis buffer (20 mM Tris-HCl pH 7.5, 1% Triton X-100, 1 mM EDTA) and protease inhibitor cocktail (Sigma) by passing through 30G needles. Chromatin was sonicated and diluted with 20 mM Tris-HCl pH 7.5, 140 mM NaCl, 0.5% Triton X-100, and protease inhibitor cocktail. After preclearing with protein A agarose beads (Upstate), the chromatin was incubated overnight with 3 µl of antibody to Six3

(Rockland) or antibody to IgG (Santacruz), followed by incubation with protein A agarose beads, and washed with 20 mM Tris-HCl pH 7.5, 140 mM NaCl and 0.5% Triton X-100. After elution with 20 mM Tris-HCl pH 7.5, 10 mM EDTA and 1% SDS and decrosslinking, DNA was purified with QIAquick kit (Qiagen) and subject to quantitative PCR (QPCR) using the primers listed in **Supplementary Table 2**. QPCR was done using Brilliant SYBR Green QPCR Master Mix (Stratagene) on an Mx4000 instrument (Stratagene). Each reaction contained 0.6 µl of a 1 in 1,600 diluted reference dye, 2 µl of chromatin DNA, 1 µl of 2 uM primer A, 1 µl of 2 µM primer B and 10 µl of 2× Master Mix in a final reaction volume of 20 µl. PCRs were amplified for 1 cycle at 95 °C for 10 min and 40 cycles at 95 °C for 30 s, 57 °C for 1 min and 72 °C for 30 s. PCRs of three independent replicates were each done in triplicate. We used differences in threshold cycle (Ct) number to quantify relative amounts of target DNA template. We normalized Ct number for each chromatin sample to Ct number for input PCR. Relative enrichment of target chromatin DNA was determined by $2^{-(Ct1-Ct2)}$, assuming that one Ct number difference represents a twofold difference in the amount of starting template.

Plasmid construction. We cloned the SBE2 reporter constructs into a vector containing the *Shh* minimal promoter, *lacZ* gene and SV40 poly(A) signal. A construct harboring a deletion of the 10-bp sequence (AACTCATTTT) from human SBE2 was generated by ligating two PCR products flanking the region of interest that were amplified with primer pairs listed in **Supplementary Table 2**. The cloning of full-length human Six3 and Six6 cDNAs into the pCDNA3-Flag expression vector was described previously[20].

Production and genotyping of transgenic mice. Transgenes were prepared for microinjection as described[11]. Transient transgenic embryos were generated by pronuclear injection into fertilized eggs derived from the (BL6 × SJL) F_1 mouse strain (Jackson Laboratories). The generation of *Six6*$^{-/-}$ embryos has been described previously[21].

Whole-mount β-galactosidase staining and *in situ* hybridization. We detected β-galactosidase activity in whole-mount embryos by using X-gal (Sigma) or salmon-gal (Biosynth) as substrates. Whole-mount RNA *in situ* hybridization was done using digoxygenin-UTP-labeled riboprobes against *Shh*, *Six3* and *Six6* according to a previously described protocol[11]. For double-labeling experiments, embryos were initially fixed for 60 min in 4% paraformaldehyde, stained in salmon-gal substrate for 2 h then postfixed overnight in 4% paraformaldehyde. Whole-mount *in situ* hybridization was then done essentially as described[11]. We photographed stained embryos after dehydration in methanol and clearing in a 1:1 solution of benzyl alcohol and benzyl benzoate. Representative embryos were rehydrated, sunk in 30% sucrose overnight, embedded and frozen in O.C.T. compound (Tissue Tek) and sectioned at 20 µm on a cryostat.

URLs. TRANSFAC, http://helixweb.nih.gov/biobase/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank the families for their participation in these studies. We also thank J. Richa and his staff at the University of Pennsylvania Transgenic and Mouse Chimeric Facility for their assistance in transgenic mouse production. We are grateful to V. Cheung, D. Kessler and T. Kadesch for their helpful comments on the manuscript. We are also grateful to K. Ewens and W. Ankener (R. Spielman laboratory) for the control human genotyping data and P. Bovolenta (Instituto Cajal, CSIC, Madrid, Spain) for kindly providing the human Six3 and Six6 expression constructs. This work was supported by NIH grants R01 NS39421 from NINDS (D.J.E.), R01 NS052386 (G.O.), March of Dimes grant #1-FY05-112 (D.J.E.), a Pew Scholar Award in the Biomedical Sciences (D.J.E.), Cancer Center Support CA-21765 (G.O.), the American Lebanese Syrian Associated Charities (ALSAC) (G.O.) and the Division of Intramural Research, National Human Genome Research Institute, National Institutes of Health (M.M.).

AUTHOR CONTRIBUTIONS

Y.J. performed the transgenic, EMSA, transfection and ChIP assays. F.C.L. designed and performed the DNA affinity capture assay, competitive EMSA and *in situ* hybridization. K.E.-J., E.R., C.D. and M.M. sequenced SBE2 from individuals with HPE. A.Y. generated the mass spectrometry data. X.L. provided the *Six6*$^{-/-}$ embryos. X.G. and G.O. generated the Six3 expression constructs. D.J.E. conceived and supervised the project and wrote the paper.

Published online at http://www.nature.com/naturegenetics/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Roessler, E. *et al.* Mutations in the human *Sonic Hedgehog* gene cause holoprosencephaly. *Nat. Genet.* **14**, 357–360 (1996).
2. Dubourg, C. *et al.* Holoprosencephaly. *Orphanet J. Rare Dis.* **2**, 8 (2007).
3. Helms, J.A., Cordero, D. & Tapadia, M.D. New insights into craniofacial morphogenesis. *Development* **132**, 851–861 (2005).
4. Fuccillo, M., Joyner, A.L. & Fishell, G. Morphogen to mitogen: the multiple roles of hedgehog signalling in vertebrate neural development. *Nat. Rev. Neurosci.* **7**, 772–783 (2006).
5. Shimamura, K. & Rubenstein, J.L. Inductive interactions direct early regionalization of the mouse forebrain. *Development* **124**, 2709–2718 (1997).
6. Fuccillo, M., Rallu, M., McMahon, A.P. & Fishell, G. Temporal requirement for hedgehog signaling in ventral telencephalic patterning. *Development* **131**, 5031–5040 (2004).
7. Marcucio, R.S., Cordero, D.R., Hu, D. & Helms, J.A. Molecular interactions coordinating the development of the forebrain and face. *Dev. Biol.* **284**, 48–61 (2005).
8. Hu, D., Marcucio, R.S. & Helms, J.A. A zone of frontonasal ectoderm regulates patterning and growth in the face. *Development* **130**, 1749–1758 (2003).
9. Chiang, C. *et al.* Cyclopia and defective axial patterning in mice lacking *Sonic hedgehog* gene function. *Nature* **383**, 407–413 (1996).
10. Cordero, D. *et al.* Temporal perturbations in sonic hedgehog signaling elicit the spectrum of holoprosencephaly phenotypes. *J. Clin. Invest.* **114**, 485–494 (2004).
11. Jeong, Y., El-Jaick, K., Roessler, E., Muenke, M. & Epstein, D.J. A functional screen for *Sonic hedgehog* regulatory elements across a 1 Mb interval identifies long range ventral forebrain enhancers. *Development* **133**, 761–772 (2006).
12. Ming, J.E. & Muenke, M. Multiple hits during early embryonic development: digenic diseases and holoprosencephaly. *Am. J. Hum. Genet.* **71**, 1017–1032 (2002).
13. Treier, M. *et al.* Hedgehog signaling is required for pituitary gland development. *Development* **128**, 377–386 (2001).
14. Roessler, E. *et al.* Loss-of-function mutations in the human *GLI2* gene are associated with pituitary anomalies and holoprosencephaly-like features. *Proc. Natl. Acad. Sci. USA* **100**, 13424–13429 (2003).
15. Park, S.S., Ko, B.J.a & Kim, B.G. Mass spectrometric screening of transcriptional regulators using DNA affinity capture assay. *Anal. Biochem.* **344**, 152–154 (2005).
16. Jean, D., Bernier, G. & Gruss, P. Six6 (Optx2) is a novel murine Six3-related homeobox gene that demarcates the presumptive pituitary/hypothalamic axis and the ventral optic stalk. *Mech. Dev.* **84**, 31–40 (1999).
17. Wallis, D.E. *et al.* Mutations in the homeodomain of the human *SIX3* gene cause holoprosencephaly. *Nat. Genet.* **22**, 196–198 (1999).
18. Lagutin, O.V. *et al.* Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. *Genes Dev.* **17**, 368–379 (2003).
19. Zhu, C. *et al.* Six3-mediated auto repression and eye development requires its interaction with members of the Groucho-related family of co-repressors. *Development* **129**, 2835–2849 (2002).
20. Conte, I., Morcillo, J. & Bovolenta, P. Comparative analysis of Six3 and Six6 distribution in the developing and adult mouse brain. *Dev. Dyn.* **234**, 718–725 (2005).
21. Li, X., Perissi, V., Liu, F., Rose, D.W. & Rosenfeld, M.G. Tissue-specific regulation of retinal and pituitary precursor cell proliferation. *Science* **297**, 1180–1183 (2002).
22. Geng, X. *et al.* Haploinsufficiency of Six3 fails to activate *Sonic hedgehog* expression in the ventral forebrain and causes holoprosencephaly. *Dev. Cell* **15**, 236–247 (2008).
23. Liu, W., Lagutin, O.V., Mende, M., Streit, A. & Oliver, G. Six3 activation of Pax6 expression is essential for mammalian lens induction and specification. *EMBO J.* **25**, 5383–5395 (2006).
24. Carl, M., Loosli, F. & Wittbrodt, J. Six3 inactivation reveals its essential role for the formation and patterning of the vertebrate eye. *Development* **129**, 4057–4063 (2002).
25. Lopez-Rios, J., Tessmar, K., Loosli, F., Wittbrodt, J. & Bovolenta, P. Six3 and Six6 activity is modulated by members of the groucho family. *Development* **130**, 185–195 (2003).
26. Del Bene, F., Tessmar-Raible, K. & Wittbrodt, J. Direct interaction of geminin and Six3 in eye development. *Nature* **427**, 745–749 (2004).
27. Cheung, V.G. *et al.* Mapping determinants of human gene expression by regional and genome-wide association. *Nature* **437**, 1365–1369 (2005).
28. Emison, E.S. *et al.* A common sex-dependent mutation in a *RET* enhancer underlies Hirschsprung disease risk. *Nature* **434**, 857–863 (2005).
29. Haiman, C.A. *et al.* A common genetic risk factor for colorectal and prostate cancer. *Nat. Genet.* **39**, 954–956 (2007).
30. Cawley, S. *et al.* Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**, 499–509 (2004).