Evolutionarily Dynamic Alternative Splicing of GPR56 Regulates Regional Cerebral Cortical Patterning

Byoung-Il Baé,† Ian Tietjen,‡† Kutay D. Atabay,§ Gilad D. Evrony,‖ Matthew B. Johnson,‡ Ebenezer Asare,§ Peter P. Wang,‖ Ayako Y. Murayama, Kiko Im,∗ Steven N. Liss,‡ Lynne Overman,§ Renad Şestan,‡ Bernard S. Chang,‖ A. James Barkovich,‖ P. Ellen Grant,‡ Meral Topçu,§ Jeffrey Politsky,‡‡ Hideyuki Okano,† Xiānhuá Piao,* Christopher A. Walsh†§

The human neocortex has numerous specialized functional areas whose formation is poorly understood. Here, we describe a 15–base pair deletion mutation in a regulatory element of GPR56 that selectively disrupts human cortex surrounding the Sylvian fissure bilaterally including "Broca’s area," the primary language area, by disrupting regional GPR56 expression and blocking RFX transcription factor binding. GPR56 encodes a heterotrimeric guanine nucleotide–binding protein (G protein)–coupled receptor required for normal cortical development and is expressed in cortical progenitor cells. GPR56 expression levels regulate progenitor proliferation. GPR56 splice forms are highly variable between mice and humans, and the regulatory element of gyrencephalic mammals directs restricted lateral cortical expression. Our data reveal a mechanism by which control of GPR56 expression pattern by multiple alternative promoters can influence stem cell proliferation, gyral patterning, and, potentially, neocortex evolution.

Although most mammals have elaborate and species–specific patterns of folds ("gyri") in the neocortex, the genetic and evolutionary mechanisms of cortical gyrification are poorly understood (1–3). Abnormal gyrification, such as polymicrogyria (too many small gyri), invariably signals abnormal cortical development, so regional disorders of gyration are of particular interest, because they highlight mechanisms specific to cortical regions. The human cortex contains dozens of cortical regions specialized for distinct functions—including as language, hearing, and sensation—but it is unsolved how these cortical regions form and how human cortical regions evolved from those of primate ancestors.

Examination of >1000 individuals with gyral abnormalities identified five individuals from three

1Division of Genetics and Genomics, Montana Center for Orphan Disease, and Howard Hughes Medical Institute, Boston Children’s Hospital, Broad Institute of MIT and Harvard, and Departments of Pediatrics and Neurology, Harvard Medical School, Boston, MA 02115, USA. 2Department of Physiology, Keio University School of Medicine, Tokyo 160-8582, Japan. 3Division of Neuroradiology and Imaging, Department of Radiology, Boston Children’s Hospital, Harvard Medical School, Boston, MA 02115, USA. The MRC-Wellcome Trust Human Developmental Biology Resource (HDBR), Newcastle, Institute of Genetic Medicine, International Centre for Life, Central Parkway, Newcastle upon Tyne NE1 3BL, UK. 5Department of Neurobiology and Kavli Institute of Neuroscience, Yale University School of Medicine, New Haven, CT 06520, USA. 6Beth Israel Deaconess Medical Center, Comprehensive Epilepsy Center, Boston, MA 02215, USA. 7Departments of Radiology, Pediatrics, Neurology, and Neurosurgical Surgery, University of California San Francisco, San Francisco, CA 94143, USA. 8Department of Pediatrics, Hacettepe University Faculty of Medicine, Ankara, Turkey. 9Department of Neurology, Medical College of Georgia, Augusta, GA 30912, USA. 10Department of Newborn Medicine, Boston Children’s Hospital and Harvard Medical School, Boston, MA 02115, USA.

The authors contributed equally to this work. *Present address: Department of Anesthesiology, Pharmacology and Therapeutics, University of British Columbia, Vancouver, British Columbia V6T 3Z4, Canada. †Present address: Northeast Regional Epilepsy Group, Atlantic Neuroscience Institute Epilepsy Center, Summit, NJ 07901, USA. §Corresponding author. E-mail: christopher.walsh@childrens.harvard.edu
families (one Turkish and two Irish-American) with strikingly restricted polymicrogyria limited to the cortex surrounding the Sylvian fissure (Fig. 1, A and B; fig. S1; and movies S1 and S2), which suggests a rare, but genetically distinctive, condition. Affected individuals suffered intellectual and language difficulty, as well as refractory seizures (onset 7 months to 10 years), but had no motor disability (table S1). Magnetic resonance imaging (MRI) and quantitative gyral analysis showed abnormal inferior and middle gyri in prefrontal and motor cortex, with mildly affected temporal lobes. Broca’s area—the “motor center for speech” in the left hemisphere and the corresponding areas of the right hemisphere were most severely affected. Affected neocortical surface showed abnormally numerous, small gyral-like folds that fused in coarse, irregular patterns, with abnormal and highly irregular white matter protrusions, consistent with polymicrogyria (5, 6), along with widening of the Sylvian fissure (Fig. 1A and fig. S1B).

Genome-wide analysis identified a single linked locus on chromosome 16q12.2-21 (Fig. 1C) containing the GPR56 gene, which, when mutated in its coding region, leads to polymicrogyria of the entire neocortex, as well as cerebellar and white matter abnormalities (7–9). As we found no mutations in the exons of GPR56, we sequenced 38 conserved non-exonic elements (table S2), in one of which we identified a small deletion in all five individuals. The mutated element normally contains two copies of a 15-base pair (bp) tandem repeat, but all affected individuals have a homozygous deletion of one 15-bp repeat (Fig. 1, E and F). The deletion is heterozygous in parents of the affected individuals, who manifest no obvious clinical signs, and is absent from thousands of control chromosomes in the Single-Nucleotide Polymorphism Database and 1000 Genomes database. The two Irish-American families carry the mutation on the same chromosomal haplotype, which reflects a common founder. It is noteworthy that the Turkish family carries the same deletion on a distinct haplotype, which indicates that the mutation arose independently (Fig. 1D).

The element is located <150 bp upstream of the transcriptional start site of noncoding exon 1m (e1m) of GPR56, which suggests that it may regulate e1m expression as a cis-regulatory element. GPR56 has at least 17 alternative transcription start sites, each beginning from a different non-coding first exon; all of the start sites are predicted to drive transcription of mRNAs whose coding sequence starts from exon 3 (Fig. 2A and fig. S2A) and all of which encode the same GPR56 protein (10, 11). The diverse noncoding first exons have distinct expression profiles, with e1m being the most robustly transcribed first exon in fetal human brain but with several other alternative transcripts also expressed in fetal and adult brain (Fig. 2A and fig. S2, B to D).

To confirm that the 15-bp deletion disrupts perisylvian GPR56 expression, we generated transgenic mice with the 23-kb human GPR56 upstream region driving green fluorescent protein (GFP) expression. The 23-kb region encompasses 16 of the 17 transcription start sites containing e1m and ends before the translation start codon (Fig. 2A). This construct drives GFP expression in the entire central nervous system, including neocortex, and recapitulates the location and relative amount of expression of endogenous mouse GPR56 protein (Fig. 2B and fig. S3). In contrast, the 23-kb construct containing the 15-bp deletion drives expression in medial, but not lateral, cortex or lateral ganglionic eminence (Fig. 2B). These data suggest that the cis-regulatory element upstream of e1m drives GPR56 expression in the perisylvian and lateral cortex, whereas disruption of the element, with consequent impairment of e1m expression, causes the perisylvian malformation.

To elucidate how the 15-bp deletion in the cis-regulatory element disrupts e1m expression,
we performed yeast one-hybrid (Y1H) screening of a mouse forebrain cDNA library with the human cis-regulatory element as bait and obtained multiple yeast colonies encoding members of the regulatory factor X (Rfx) transcription factor family (Fig. 2C) (12). RFX1 and RFX3 bind the normal element in vitro, with binding decreased to 60 to 70% by the 15-bp deletion (Fig. 2D). Chromatin immunoprecipitation sequencing confirmed RFX3 binding to the element (fig. S4) (13). RFX1 and GPR56 colocalize in germlinal zones of fetal human brain (Fig. 2E). Dominant-negative RFX abrogates normal, but not mutant, e1m promoter activity on embryonic day 13.5 (E13.5) in mouse cortical cultures (Fig. 2F). Furthermore, genetic ablation of Rfx4 decreases Gpr56 expression in developing mouse brain (14). RFX and GPR56 expression patterns are correlated (fig. S5, A and B) (15), with RFX3 and RFX7 most prominent in human ventral lateral prefrontal cortex, the region affected in perisylvian polymicrogyria (Fig. 2G), which suggests that multiple RFX proteins regulate the element.

GPR56 encodes an adhesion heterotrimeric guanine nucleotide–binding protein (G protein)–coupled receptor that is highly expressed in cortical progenitors (7, 16) and binds extracellular matrix proteins (17). Loss of GPR56 disrupts radial glia and causes breaches in the pial basement membrane, through which some neurons over-migrate (9, 16). However, even where the pia is intact, we found that neocortical thickness and organization are irregular, with occasional thin regions in Gpr56 knockout mice (Fig. 3A). Postmortem analysis of a human with biallelic GPR56 coding mutations showed a very thin cortex, which suggested potential roles of GPR56 in neurogenesis as well (9). GPR56 protein is most highly expressed in progenitors in the ventricular and subventricular zones during neurogenesis in mice (16, 18). GPR56 expression in developing human and marmoset neocortex is highest in the ventricular zone, as well as in the outer subventricular zone, which is expanded in mammals with larger brains (2) (Fig. 3B and fig. S5, C and D).

Impairment and overexpression of GPR56 show that its expression regulates proliferation. Gpr56 knockout mice show fewer phosphohistone H3 (PH3)–positive mitotic progenitor cells and TBR2-positive intermediate progenitors than wild-type mice in the neocortex at E14.5. Conversely, mice carrying a transgene that directs overexpression of human GPR56 show increased

**Fig. 2. The noncoding mutation ablates lateral gene expression.** (A) Schematic of the human GPR56 locus showing 17 alternative transcription start sites. E1m is highly expressed in the human fetal brain [mRNA-seq (mRNA-Seq) track, arrow]. The 15-bp deletion is upstream of e1m, located within a cis-regulatory element as one of two tandem 15-bp repeats. (B) A 23-kb upstream region of human GPR56 drives GFP expression throughout the transgenic mouse neocortex (E14.5), which mirrors endogenous GPR56 protein expression. The 15-bp deletion eliminates GFP expression from lateral cortex but preserves medial cortex expression, consistent with lesions observed by brain MRI (fig. S1) (n = 4 to 6 embryos with identical patterns per construct). Scale bar, 200 μm. (C) Y1H screening reveals rfx transcription factor binding to the cis-regulatory element. See text for details. (D) The mutation decreases RFX binding to the cis-regulatory element in vitro. (E) RFX1 and GPR56 are colocalized in a human fetal brain 19 weeks after conception. Higher magnification of the outer subventricular zone is shown (right); v, ventricular zone; is, inner subventricular zone; os, outer subventricular zone; and i, intermediate zone. Scale bars, 100 μm (left) and 10 μm (right). (F) Dominant-negative RFX (white bars) abrogates normal e1m promoter activity. Black bars, GFP control. (G) Each RFX gene has distinct expression patterns in the fetal human brain. Each number means the corresponding RFX isoform. RFX3 and RFX7 are enriched in regions affected by perisylvian polymicrogyria (green boxes). pfc, prefrontal cortex; opfc, orbital pfc; dlpc, dorsolateral pfc; mpc, medial pfc; vlpfc, ventrolateral pfc; ms, motor-sensory cortex. *P < 0.001, t test.
mitotic progenitor cells and intermediate progenitor cells (Fig. 3, C and D). In utero electroporation (at E13.5 with analysis 48 hours later at E15.5) of a plasmid encoding GPR56 (as well as GFP, to mark the cells) caused cells to persist in proliferative zones compared with cells expressing GFP alone (Fig. 3E). Changes in the number of intermediate progenitors in transgenic and knock-out mice may be secondary to changes in the radial progenitor cells that generate them or might reflect a direct role of GPR56 in intermediate progenitors but is consistent with a report that loss of TBR2 (EOMES) also causes polymicrogyria in humans (19).

The cis-regulatory element upstream of GPR56 e1m is found in genomes of all placental mammals, but not monotremes, marsupials, or non-mammals, which suggests that it emerged after placental and nonplacental mammals diverged 85 to 100 million years ago (fig. S7B). The cis-regulatory element sequence is only found at the e1m locus in GPR56 but not elsewhere in the genome. E1m itself shows homology at its 3’ end to a long interspersed nuclear element (LINE)–4/RTE, a family of retrotransposons active in early mammals after divergence from marsupials (20). Another noncoding GPR56 exon (exon 2), present only in primates, derives from a primate-specific Alu insertion (fig. S7B). In contrast to the >17 alternative first exons in humans, we found only five noncoding first exons in the mouse Gpr56 gene (Fig. 4A and fig. S7A) (10, 11). Thus, GPR56 acquired many noncoding upstream exons and generated alternative splice forms with distinct expression patterns (fig. S2, B and D), in the lineage leading to humans. Transposable element insertion played a role in generating this diversity.

To test directly whether evolutionary changes in GPR56’s alternative splice forms have functional effects, we generated transgenic mice in which the β-galactosidase (β-gal) gene is driven by a minimal 300-bp e1m promoter from human, mouse, marmoset, dolphin, and cat (Fig. 2A and fig. S6A). The mouse e1m promoter drives β-gal expression broadly in the nervous system in diverse cell types, which suggests that this simple 300-bp e1m promoter is sufficient to recapitulate major features of the endogenous mouse GPR56 expression (16–18) (Fig. 4B and fig. S6B). In contrast, the corresponding human e1m promoter has a variety of deletions and single-nucleotide variants, relative to the mouse sequence (fig. S6A), and drives much more limited expression in rostral-lateral cortex (Fig. 4B and fig. S6B). Weak lateral cortical expression is visible in embryos carrying the mouse e1m promoter-β-gal transgene,

Fig. 3. GPR56 regulates neuroprogenitor proliferation. (A) In Gpr56 knockout mice, neurons overmigrate through breached pial basement membrane (arrowheads) or undermigrate (arrows) forming irregular cortical layers, as shown by immunostaining of Cox1, an upper layer (II to IV) marker (p9). Thin cortex is occasionally observed (asterisks). (B) GPR56 is highly expressed in human ventricular zone and outer subventricular zone at 12 to 17 weeks of gestation (GW), which suggests roles in neuroprogenitors. v, ventricular zone; s, subventricular zone; is, inner subventricular zone; os, outer subventricular zone; i, intermediate zone; c, cortical plate; and m, marginal zone. (C to D) Human GPR56 transgenic (Tg) mice have significantly more mitotic (PH3+) neuroprogenitor cells and intermediate progenitor (TBR2+) cells than wild-type (WT). In contrast, Gpr56 knockout (KO) mice have significantly fewer mitotic cells and intermediate progenitors than WT (E13.5 to E14.5). (n = 7 mice per genotype; *P < 0.005; **P < 0.001; paired t test). (E) The cells that are in utero electroporated (from E13.5 to E15.5) with human GPR56-IRE5-GFP [either side of the internal ribosome entry site (IRES), GFP expressing] persist in the germinal zones longer than the GFP control cells. Red, TBR2; blue, Hoechst. (n = 11 mouse embryos per construct; *P < 0.0001; chi-squared test). Scale bars, 500 μm (A) and 100 μm (B) to (E).
which suggests that, in humans, additional elements besides the 300-bp e1m promoter region are required to drive the full complement of GPR56 expression. E1m promoters from marmoset, dolphin, and cat drive expression patterns generally similar to human. The large number of noncoding first exons in gyrencephalic mammals and humans. Elaboration of additional alternative splice forms provides a mechanism for potentially independent evolution of these multiple forms.

Our studies show that levels of GPR56 control proliferation of progenitors in the neocortex. Loss of GPR56 expression impairs neurogenesis, and overexpression enhances proliferation and progenitor number. Selective GPR56 loss causes strikingly regional defects of cortical development. GPR56 likely influences progenitor proliferation by stabilizing the basal process of radial neuroepithelial progenitors, because (i) GPR56 protein localizes to the basal processes of radial neuroepithelial cells (16); (ii) GPR56 binds extracellular matrix proteins in the pial basement membrane, such as collagen type III (17) and tetraspanins, which bind integrins as well (21); (iii) GPR56 is required for normal attachment of basal processes to the pial basement membrane in mice (16); and (iv) basal processes regulate progenitor proliferation via integrin signaling (22, 23), and GPR56 interacts genetically with α3β1 integrin (24). The elaboration of the GPR56 locus in gyrencephalic mammals, and especially humans, to produce many alternative splice forms with diverse expression patterns presents GPR56 as a key target that could influence the dramatic changes in shape and folding that characterize the forebrain of many mammals. Elaboration and specialization of alternative transcripts with distinct transcription start sites is an evolutionary mechanism that has been difficult to study because of the lack of comprehensive catalogs of RNA splice forms, but continued RNA sequencing studies may soon provide the opportunity to assess its importance systematically.

References and Notes


Acknowledgments: Research performed on samples of human origin was conducted according to protocols approved by participating institutions, including Boston Children’s Hospital and Beth Israel Deaconess Medical Center. The human embryonic and fetal material was provided by the Joint Medical Research Council (grant no. G0700898—Wellcome Trust (grant no. G8082557) Human Developmental Biology Resource (www.hdbr.org) and the National Institute of Child Health and Human Development, NIH, Brain and Tissue Bank at the University of Maryland (contract no. HHSN27520090001C, reference no. NO1-HD-2-0011). Gpr56 knockout mice are from Genentech. This work was supported by the Strategic Research Program for Brain Sciences and from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Japan (H.O.); Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program) (H.O.); U51/MH081896 from National Institute of Mental Health, NIH (N.S.); 2R01NS035129 from National Institute of Neurological Disorders and Stroke, NIH (C.A.W.); and The Paul G. Allen Family Foundation (C.A.W.). Additional funding support listed in supplementary materials. C.A.W. is an investigator of the Howard Hughes Medical Institute. Gpr56 knockout mice are available from Genentech subject to a Material Transfer Agreement.

Supplementary Materials

www.sciencemag.org/content/343/6172/764/suppl/DC1 Materials and Methods

Supplementary Text Figs. S1 to S7 Tables S1 and S2 Movies S1 to S2 References (23–40)

7 August 2013; accepted 17 December 2013