









ARTICLE

<https://doi.org/10.1038/s41467-019-12013-y>

OPEN

# In vivo epigenetic editing of *Sema6a* promoter reverses transcallosal dysconnectivity caused by *C11orf46/Arl14ep* risk gene

Cyril J. Peter <sup>1,9</sup>, Atsushi Saito<sup>2,9</sup>, Yuto Hasegawa<sup>2</sup>, Yuya Tanaka<sup>2</sup>, Mohika Nagpal<sup>2</sup>, Gabriel Perez<sup>2</sup>, Emily Alway <sup>2</sup>, Sergio Espeso-Gil <sup>1</sup>, Tariq Fayyad<sup>1</sup>, Chana Ratner<sup>1</sup>, Aslihan Dincer <sup>1</sup>, Achla Gupta<sup>1,3</sup>, Lakshmi Devi<sup>1,3</sup>, John G. Pappas <sup>4</sup>, François M. Lalonde <sup>5</sup>, John A. Butman <sup>6</sup>, Joan C. Han<sup>7,8</sup>, Schahram Akbarian <sup>1</sup> & Atsushi Kamiya<sup>2</sup>

Many neuropsychiatric risk genes contribute to epigenetic regulation but little is known about specific chromatin-associated mechanisms governing the formation of neuronal connectivity. Here we show that transcallosal connectivity is critically dependent on *C11orf46*, a nuclear protein encoded in the chromosome 11p13 WAGR risk locus. *C11orf46* haploinsufficiency was associated with hypoplasia of the corpus callosum. *C11orf46* knockdown disrupted transcallosal projections and was rescued by wild type *C11orf46* but not the *C11orf46*<sup>R236H</sup> mutant associated with intellectual disability. Multiple genes encoding key regulators of axonal development, including *Sema6a*, were hyperexpressed in *C11orf46*-knockdown neurons. RNA-guided epigenetic editing of *Sema6a* gene promoters via a dCas9-SunTag system with *C11orf46* binding normalized SEMA6A expression and rescued transcallosal dysconnectivity via repressive chromatin remodeling by the SETDB1 repressor complex. Our study demonstrates that interhemispheric communication is sensitive to locus-specific remodeling of neuronal chromatin, revealing the therapeutic potential for shaping the brain's connectome via gene-targeted designer activators and repressor proteins.

<sup>1</sup>Friedman Brain Institute and Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. <sup>2</sup>Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. <sup>3</sup>Department of Pharmacology and System Therapeutics, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. <sup>4</sup>Department of Pediatrics, New York University School of Medicine, New York, NY, USA. <sup>5</sup>Human Genetics Branch, National Institute of Mental Health, Bethesda, MD 20892, USA. <sup>6</sup>Diagnostic Radiology Department, The Clinical Center of the National Institutes of Health, Bethesda, MD 20892, USA. <sup>7</sup>Unit on Metabolism and Neuroendocrinology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institute of Health, Bethesda, MD 20892, USA. <sup>8</sup>Departments of Pediatrics and Physiology, University of Tennessee Health Science Center, and Children's Foundation Research Institute, Le Bonheur Children's Hospital, Memphis, TN 38103, USA. <sup>9</sup>These authors contributed equally: Cyril J. Peter, Atsushi Saito. Correspondence and requests for materials should be addressed to S.A. (email: [schahram.akbarian@mssm.edu](mailto:schahram.akbarian@mssm.edu)) or to A.K. (email: [akamiya1@jhmi.edu](mailto:akamiya1@jhmi.edu))

A wide range of neurodevelopmental disorders manifesting in infancy and early childhood (including intellectual disability and autism spectrum disorder) or young adulthood (including schizophrenia) are associated with disrupted interhemispheric communication, which in some but not all of the affected cases is accompanied by structural alterations of the corpus callosum, the brain's largest commissure<sup>1–3</sup>. The formation of interhemispheric connectivity involves a highly orchestrated multi-step process, including midline zipper glia promoting hemispheric fusion and midline crossing of pioneer fibers, followed by ingrowth of large numbers of transcallosal axons interconnecting the left and right cerebral cortex<sup>4</sup>. Early occurring disruptions of commissural development could manifest as agenesis of the corpus callosum<sup>4</sup>. However, alterations affecting later phases of development, including defective or misdirected axonal growth cone navigation with failure to cross the midline, or faulty axonal ingrowth into the contralateral cortex, could be responsible for the partial thinning, or developmental hypoplasia of the corpus callosum, extending either across its entire rostro-caudal axis or subportion thereof<sup>4,5</sup>. In addition, proper region- and lamina-specific axonal innervation and arborization within contralateral cortex are dependent on the functional activity of cortical projection neurons<sup>6</sup>. Furthermore, late occurring defects of transcallosal connectivity may not result in overt macroscopic defects of brain morphology including its commissures. Such a complex developmental program of commissural connectivity is, perhaps unsurprisingly, associated with a heavy footprint in the genetic risk architecture of neurodevelopmental disease. In fact, a number of chromosomal copy number variations (CNVs) have been linked to callosal agenesis<sup>7</sup>.

Of note, many cases with a chr. 22q11.2 micro-deletion or -duplication<sup>8</sup> as one of the most frequently diagnosed CNV in neuropsychiatric disease cohorts are affected by callosal hypoplasia and hyperplasia<sup>8</sup> and furthermore, interhemispheric dysconnectivity phenotypes have been associated with a rapidly increasing list of specific point mutations in cell surface signaling genes encoding key regulators of neurite outgrowth and connectivity, such as the *Roundabout guidance receptor 1 (ROBO1)*<sup>9</sup> or the *L1 cell adhesion molecule (L1CAM)*<sup>10</sup>. Much less is known about the role of nuclear signaling molecules for disease-relevant commissural phenotypes. For example, orderly formation of transcallosal connectivity critically depends on proper gene dosage and activity of a subset of chromatin regulators such as the transcriptional repressor *switch-insensitive 3 family member A (SIN3A)*<sup>11</sup> and the *Forkhead box G1 (FOXG1)* transcription factor<sup>12</sup>. However, molecular and cellular mechanisms linking the neuronal epigenome to interhemispheric connectivity remain poorly explored.

Here, we describe *C11orf46/ADP ribosylation factor like GTPase 14 effector protein (ARL14EP)*, a small (35 kDa) nuclear protein encoded in the chr. 11p13 Wilms Tumor, Aniridia, Genitourinary Abnormalities, intellectual disability (formerly referred to as Mental Retardation) (WAGR) risk locus, as a novel chromatin regulator of transcallosal connectivity. Importantly, we show that *C11orf46* is, as an RNA-guided Cas9 fusion protein, suitable for targeted epigenomic promoter editing to drive the expression of specific regulators of axonal development in immature transcallosal projection neurons, thereby offering a molecular tool to affect interhemispheric communication.

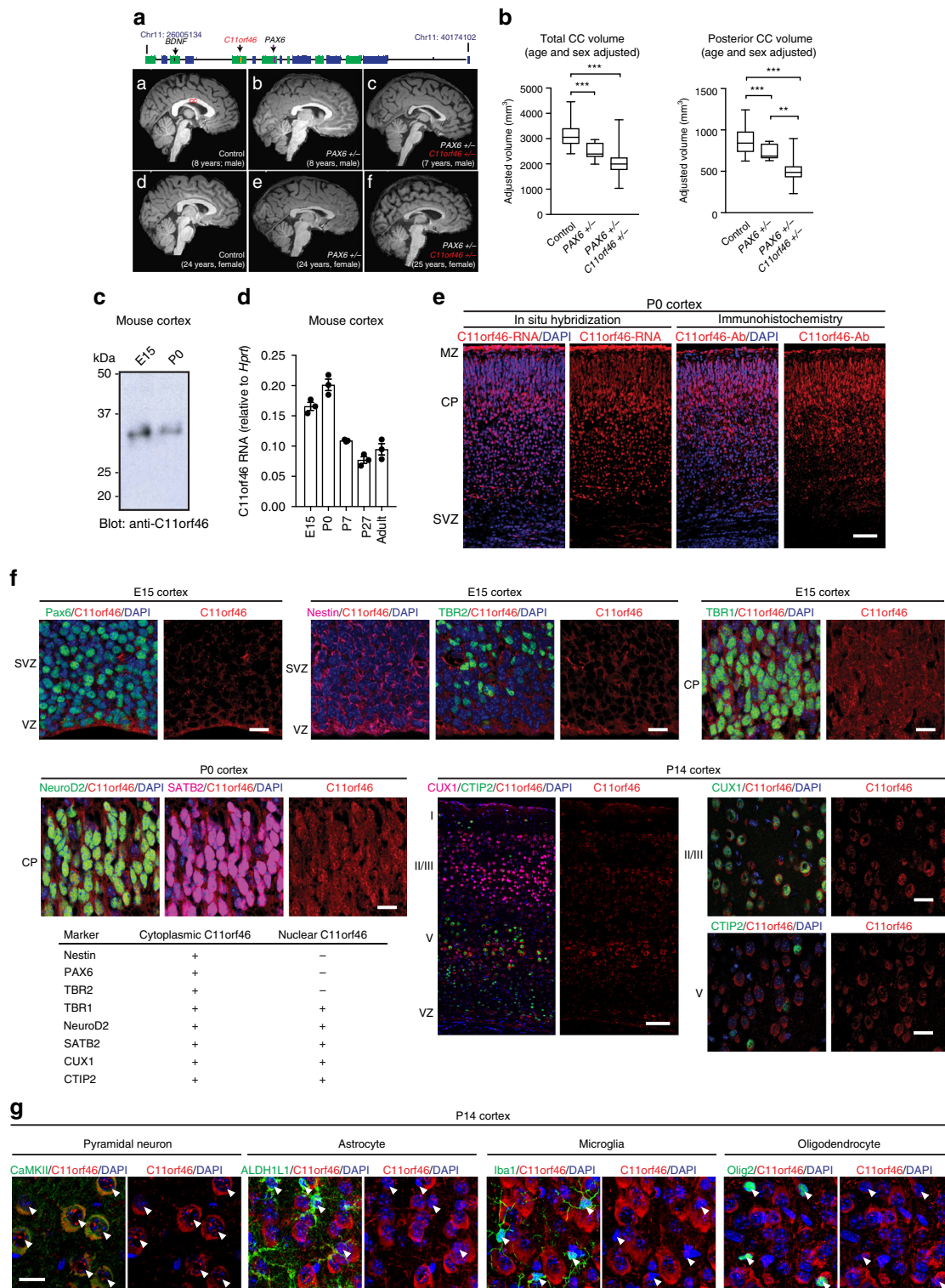
## Results

***C11orf46* is critical for transcallosal connectivity.** A recent study identified 50 novel genes associated with autosomal recessive forms of intellectual disability, including open reading frame *C11orf46*<sup>13</sup>. Importantly, moreover, this gene is located in the chr. 11p13-14 neurodevelopmental risk locus and deleted in some

cases of WAGR syndrome (Online Mendelian Inheritance of Man (OMIM) #194072), a rare copy number variant disorder with its core features caused by haploinsufficiency for the *Wilms tumor 1 (WT1)* and *Paired box 6 (PAX6)* homeobox gene<sup>14</sup>. Of note, *C11orf46* is positioned in between *PAX6* and the *Brain derived neurotrophic factor (BDNF)*, chr. 11p14.1 gene which is included in the deleted segment in approximately half of patients with WAGR syndrome and associated with obesity and more severe cognitive and behavioral deficits<sup>14</sup>. However, there is evidence that, independent of *BDNF*, additional coding genes in chr. 11p13-14 could contribute to neurodevelopmental disease. For example, exome sequencing studies in Middle Eastern pedigrees associated homozygous *C11orf46* (chr11:30,323,051-30,338,458, GRCh38/hg38) point mutations with cognitive disease and intellectual disability<sup>13</sup>. To begin the neurological exploration of *C11orf46*, we studied differences in brain morphology between WAGR microdeletion cases encompassing the *C11orf46* gene and those with less extensive chr. 11p13 segment haploinsufficiency.

To this end, we examined magnetic resonance T1-weighted images (MRI) in health control subjects ( $n = 23$ ), isolated *PAX6*+/- subjects ( $n = 12$ ), and individuals with WAGR-region deletions involving *PAX6*+/- and *C11orf46*+/- ( $n = 27$ ). Indeed, *C11orf46* haploinsufficiency was associated with significant differences observed in total corpus callosum (CC) and posterior CC volumes but not for other individual portions of the CC. After adjusting for age and sex, patients with isolated *PAX6*+/- compared to controls had significantly smaller total and posterior CC volumes [ANCOVA:  $F(4,57) = 29.43$ ,  $P < 10^{-8}$ ; post hoc two-sided LSD mean difference (95% CI):  $-705$  ( $-1046$  to  $-364$ )  $\text{mm}^3$ ,  $P = 0.0001$ ] and posterior CC volumes [ $F(4,57) = 36.38$ ,  $P < 10^{-10}$ ;  $-180$  ( $-282$  to  $-78$ )  $\text{mm}^3$ ,  $P = 0.0009$ ], a finding consistent with the role of *PAX6* in morphologic brain development<sup>15,16</sup>. However, patients with combined *PAX6/C11orf46*+/- had an even smaller posterior CC volume compared to both patients with isolated *PAX6*+/- [ $-173$  ( $-284$  to  $-62$ )  $\text{mm}^3$ ,  $P = 0.003$ ] and controls [ $-353$  ( $-436$  to  $-270$ )  $\text{mm}^3$ ,  $P < 10^{-10}$ ] (Unadjusted volumes are shown but  $P$ -values from ANCOVAs adjusting for covariates are depicted in the figure; estimated marginal means are provided in the figure legend) (Fig. 1a, b). Because of the established role of *BDNF* in brain development and intellectual functioning<sup>17</sup>, we also considered a potential confounding effect of *BDNF*+/- . We observed no significant differences in the *PAX6/C11orf46*+/- patients with ( $n = 17$ ) and without ( $n = 10$ ) *BDNF*+/- for total or posterior CC volumes [246 ( $-122$  to 614)  $\text{mm}^3$ ,  $P = 0.19$  and 41 ( $-73$  to 156)  $\text{mm}^3$ ,  $P = 0.47$ , respectively, for unadjusted comparisons; 281 ( $-84$  to 647)  $\text{mm}^3$ ,  $P = 0.13$  and 50 ( $-61$  to 162)  $\text{mm}^3$ ,  $P = 0.37$ , respectively, on ANCOVA adjusting for age and sex] (Supplementary Fig. 1a). Lymphoblastoid cell lines from *C11orf46* haploinsufficient WAGR cases have approximately 50% lower *C11orf46* mRNA levels, indicating that the expression of the remaining intact allele does not undergo compensatory upregulation (Supplementary Fig. 1b). Therefore, loss of one *C11orf46* copy in the context of chr. 11p13-14.1 microdeletions appears to be more highly detrimental for neurodevelopment and associated with more severe hypoplasia of the callosal commissure.

Having linked *C11orf46* haploinsufficiency to callosal hypoplasia, we next explored *C11orf46* expression patterns in mouse brain. In the cerebral cortex, *C11orf46* quantification at the level of mRNA, and protein with a custom-made anti-*C11orf46* antibody (Supplementary Fig. 1c) showed its expression across a wide age window, with higher levels in the pre- and post-natal developmental periods as compared to the adult brain (Fig. 1c–f). *C11orf46* was mainly expressed in the cortical plate (CP) with minimal labeling in the proliferative subventricular zone (SVZ) at



mRNA and protein levels (Fig. 1e). To examine cell type-specific expression pattern of C11orf46 during brain development, C11orf46 was co-stained with multiple cell type-specific markers. Notably, C11orf46 is expressed in a nuclear punctate pattern in NeuroD2-positive cells, but not in PAX6 and TBR2-positive cells, although we observed cytoplasmic expression of C11orf46 in cells expressing these markers (Fig. 1f and Supplementary Fig. 2a). Nuclear and cytoplasmic expression of C11orf46 was also observed in neuronal markers, such as TBR1, SATB2, CUX1, and CTIP2-positive cells (Fig. 1f and Supplementary Fig. 2a). In

Nestin-expressing radial glial cells, C11orf46 is expressed in the cytoplasm, but not in the nucleus (Fig. 1f and Supplementary Fig. 2a). Collectively, these results suggest that nuclear C11orf46 is primarily expressed in post-mitotic neurons. Additional double-labeling experiments confirmed that C11orf46 is predominantly expressed in Calmodulin-kinase II (CaMKII)-positive glutamatergic neurons but not in ALDH1L1-positive astrocytes or Olig2-positive oligodendrocytes and precursors, or Iba1-positive microglia (Fig. 1g and Supplementary Fig. 2b). C11orf46 immunoreactivity in GABAergic neurons was much weaker than

**Fig. 1** C11orf46 is a neuronal nuclear protein important for callosal development. **a** (top) Genomic map of chr. 11 WAGR deletion locus. (A–F) Representative sagittal T1 MRI images from midline were shown. A) 8-year-old male control; B) 8-year-old male isolated *PAX6*+/-; C) 7-year-old male heterozygous 11p13 deletion with *PAX6*+/- and *C11orf46*+/-; D) 24-year-old female control; E) 24-year-old female isolated *PAX6*+/-; F) 25-year-old female heterozygous 11p13 deletion with *PAX6*+/- and *C11orf46*+/- . **b** Hypoplastic corpus callosum (CC) in patients with 11p13 deletion encompassing *C11orf46*. ANCOVA including age and sex as covariates, compared CC volumes of control ( $n = 23$ ), isolated *PAX6*+/- ( $n = 12$ ), and heterozygous 11p13 deletion with *PAX6*+/- and *C11orf46*+/- ( $n = 27$ ). *P*-values are shown from two-sided post hoc LSD comparisons: \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . Box and whisker plots represent the distribution of unadjusted CC volume. The inner bar indicates median value. The upper and lower box ends represent first and third quantile. The upper and lower whisker ends represent the maximum and minimum values. **c** C11orf46 protein in mouse cortices at embryonic day 15 (E15) and postnatal day 0 (P0). **d** mRNA expression of C11orf46 in the mouse cerebral cortex during prenatal and postnatal development (mean  $\pm$  SEM).  $n = 3$ . **e** (Left) C11orf46 mRNA (red) is predominantly expressed in the cortical plate (CP) in the somatosensory cortex at P0. (Right) C11orf46 protein (red) is predominantly expressed in the CP at P0. Blue, nucleus. Scale bar, 50  $\mu$ m. MZ, marginal zone; SVZ, subventricular zone. **f** C11orf46 (red) is expressed in the cytoplasm, but not in the nucleus in Nestin-, PAX6-, and TBR2-positive cells, whereas C11orf46 is expressed in the cytoplasm and in the nucleus with a punctate pattern in TBR1-, NeuroD2-, SATB2-, CUX1-, and CTIP2-positive neuronal cells. (bottom left) Cell type specific C11orf46 expression are summarized in the table. Scale bar, 10 and 100  $\mu$ m (low magnification images). **g** Representative images of C11orf46 protein staining with cell type-specific markers (white arrowheads) at P14. C11orf46 is predominantly expressed in CaMKII-positive neurons, but not Olig2 (oligodendrocyte), ALDH1L1 (astrocyte), Iba-1 (microglia) -positive cells. Scale bar, 20  $\mu$ m

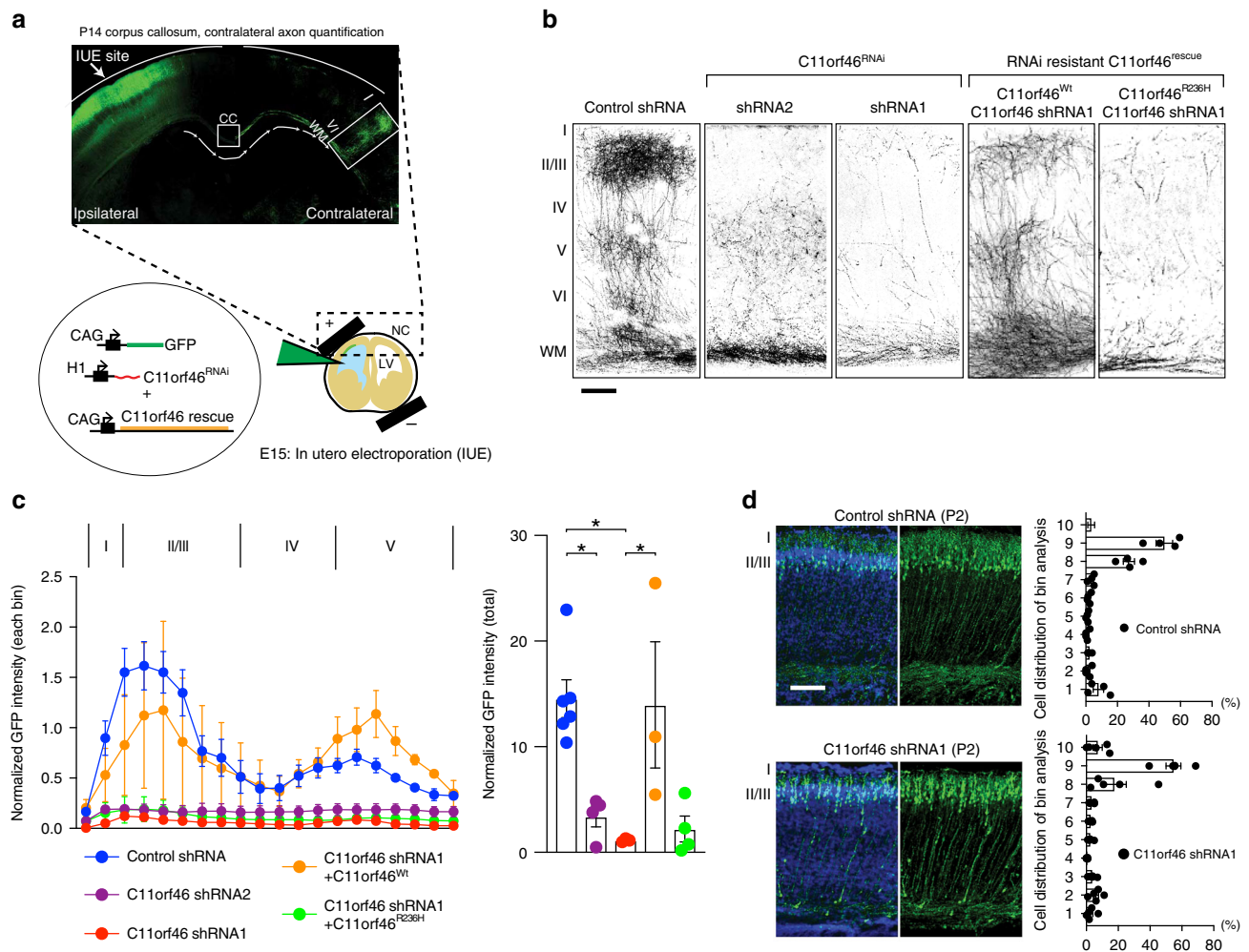
those in glutamatergic neurons (data not shown). We conclude that in the developing cerebral cortex, C11orf46 is primarily expressed in post-migratory glutamatergic cortical neurons.

To explore the role of C11orf46 in developing cortex, we knocked down its expression by in utero electroporation (IUE). We delivered C11orf46 (shRNA1, shRNA2) or control short hairpin RNAs (shRNAs) together with green fluorescent protein (GFP) expression plasmid using our published protocols<sup>18–21</sup>. We selected embryonic day 15 (E15) for IUE procedure with two reasons: first, C11orf46 is highly expressed at E15 (Fig. 1c, d). Second, at this stage IUE mostly targets progenitor cells in the ventricular zone (VZ) which later differentiate into cortical layer II/III pyramidal neurons<sup>22,23</sup>, many of which become transcallosal projection neurons<sup>24</sup> as our cell type-of-interest given the clinical phenotype reported above. Indeed, E15 IUE *C11orf46* knockdown resulted in severe arborization deficits in GFP-labeled callosal axons projecting into the contralateral cortex across all cortical layers at postnatal day 14 (P14) (Fig. 2a–c). This phenotype was highly dependent on the level of *C11orf46* knockdown. More severe axonal arborization deficits by C11orf46 shRNA1 (Fig. 2b, c) was associated with stronger knockdown effect on C11orf46 compared with shRNA2 (Supplementary Fig. 3a). Importantly, these axonal arborization deficits were partially rescued by co-expressing shRNA resistant wild-type C11orf46. Strikingly however, an shRNA resistant C11orf46 (R236H) mutant protein carrying a disease-associated non-synonymous point mutation that consists of arginine to histidine substitution at codon 236<sup>13</sup> did not rescue the transcallosal phenotype induced by *C11orf46* knockdown (Fig. 2b, c and Supplementary Fig. 3b). We also found reduction of thickness of corpus callosum that was elicited by knockdown of *C11orf46*, which is reversed by co-expression of RNAi-resistant wild-type C11orf46, but not by C11orf46 (R236H) mutant (Supplementary Fig. 3c). Furthermore, we observed no difference in percentage of SATB2 and CUX1 positive cells in GFP-labeled neurons in *C11orf46* knockdown brains, compared to those in controls, suggesting that *C11orf46* knockdown did not affect the number of cortico-cortical projecting neurons (Supplementary Fig. 3d). Importantly, the observed axonal phenotype resulting from C11orf46 knockdown was highly specific, because orderly radial neuronal migration into the cortex, when assessed at P2, was indistinguishable from control (Fig. 2d).

### C11orf46 regulates genes important for axonal connectivity.

Having shown that C11orf46 is widely expressed in cortical projection neurons and essential for transcallosal connectivity, we

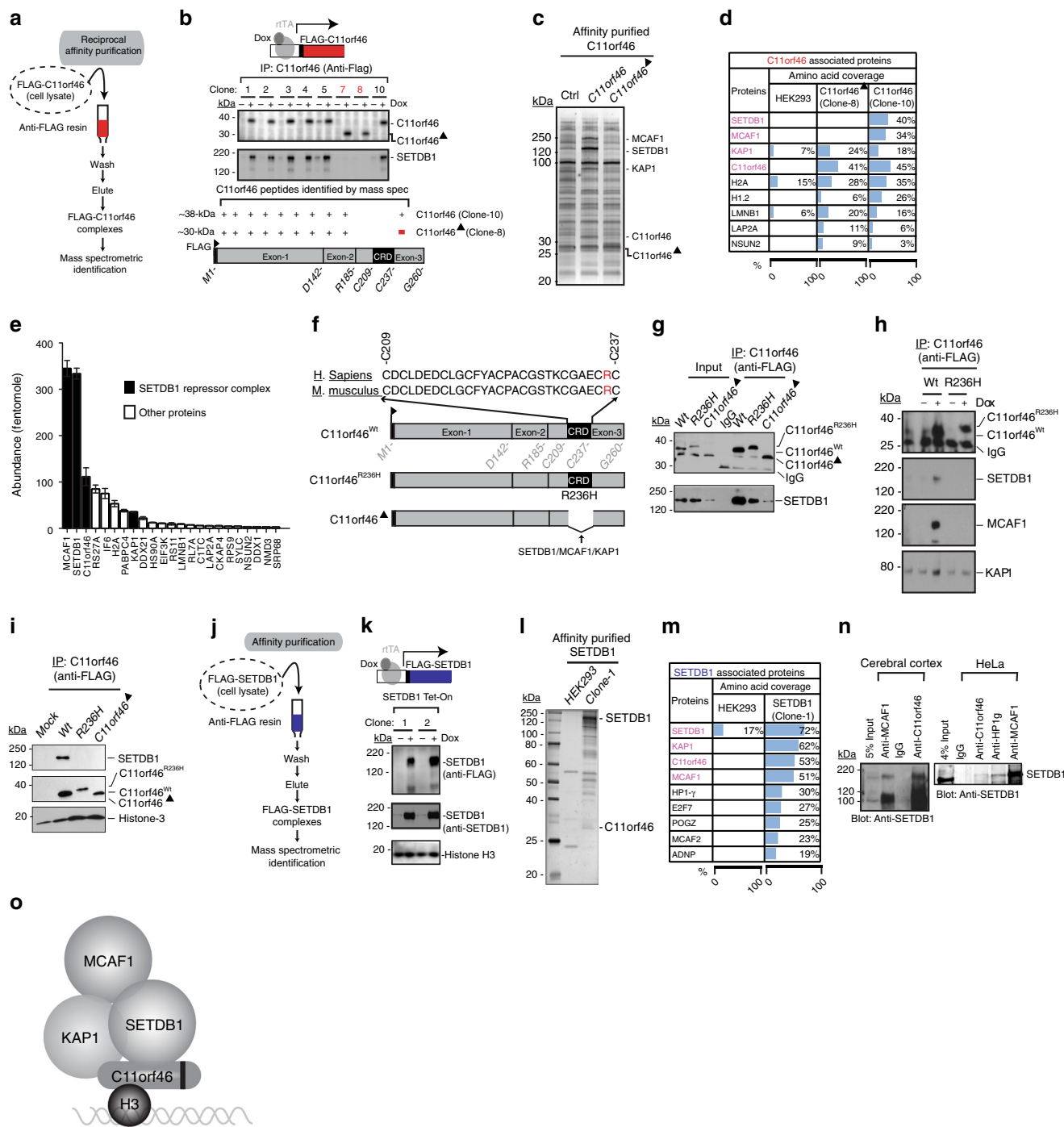
next wanted to gain initial insights into the molecular mechanisms underlying C11orf46 actions on the corpus callosum development. In the single report published to date on C11orf46 function, the protein was renamed ADP-Ribosylation Factor-Like 14 Effector Protein (ARL14EP), and ascribed with a role in major histocompatibility class (MHC) II antigen presentation in immune cells<sup>25</sup>. Therefore, we were very surprised about the predominant, if not exclusive, nuclear localization of C11orf46 in our immunohistochemical assays in the developing CP (Fig. 1e, f). However, a recent study listed C11orf46/ARL14EP among several hundred proteins serving as nuclear export cargo for the RanGTPase-driven exportin CRM1<sup>26</sup> and moreover, large scale total proteome protein–protein interaction mappings consistently identified chromatin-regulatory proteins as C11orf46 binding partners, including the SET domain bifurcated histone lysine methyltransferase 1 (SETDB1)/KMT1E (SETDB2/KMT1F)-MBD-containing-associated factor (MCAF1) histone lysine 9 (K9) methyltransferase repressor complex<sup>27–30</sup> (referred to as KMT-RC hereafter). To further examine C11orf46-associated proteins by affinity purification coupled to mass spectrometry (Fig. 3a), we generated inducible HEK293 cell lines expressing Flag-tagged C11orf46 (Fig. 3b). We produced two types of inducible clones, one for full length FLAG-tagged C11orf46 (clones 1–5 and 10) and one for a truncated form lacking amino acids 209–237 at C11orf46’s C-terminal region which corresponds to cysteine rich domain, CRD, (C11orf46 $\Delta$ ) (clone 7 and 8) (Fig. 3b and Supplementary Fig. 4a, b). Interestingly, nuclear proteins outranked all other C11orf46 binding partners, with the top scoring SETDB1-KRAB associated protein (KAP1)-MCAF1 chromatin repressor complex<sup>31,32</sup> (Fig. 3c–e and Supplementary Table 1), which is consistent with proteome interaction databases<sup>27–30</sup>. These findings were specific for full length C11orf46 because the truncated form C11orf46 $\Delta$  showed only very weak, or no interactions with nuclear proteins including the SETDB1-KAP1-MCAF1 repressor complex (Fig. 3d). Given the *C11orf46*<sup>R236H</sup> point mutation is located in the CRD domain which is required for protein binding of C11orf46 with SETDB1 (Fig. 3b–d), we wondered whether R236H may affect the interaction of C11orf46 and the SETDB1 complex. To test this hypothesis, in addition to inducible HEK293 cell lines expressing Flag-tagged C11orf46 and Flag-tagged C11orf46 $\Delta$ , we generated cell lines for inducible expression of FLAG-C11orf46<sup>R236H</sup> (Fig. 3f). We observed that deletion of CRD domain dramatically weakened the interaction of C11orf46 and SETDB1, whereas the R236H substitution abolished C11orf46-SETDB1 interaction (Fig. 3g, h). Binding of C11orf46 with MCAF1 and KAP1, other members of the SETDB1 protein complex, were also severely



**Fig. 2** Suppression of *C11orf46* impairs callosal development and interhemispheric connectivity. **a** (top) Schematic representation of in utero electroporation of GFP expressing plasmid into somatosensory cortex at E15. (bottom) P14 somatosensory cortex expressing GFP in pyramidal neurons at IUE injection site (Ipsilateral side) and axonal projections between layer I–VI (Contralateral side). **b** Knockdown of *C11orf46* in pyramidal neurons of layer II/III (Ipsilateral) show impaired axonal terminal arborization in the contralateral somatosensory cortex at P14. *C11orf46* shRNA1 had a stronger axonal arborization impairment than shRNA2 when compared with control shRNA (first 3 panels); this axonal deficit was partially rescued by co-expressing RNAi resistant wild-type *C11orf46* (*C11orf46*<sup>Wt</sup>), but not by R236H mutant (*C11orf46*<sup>R236H</sup>) (last 2 panels). Scale bar, 100  $\mu$ m. **c** Layer distributions (left) and total axonal density (right) of callosal axons in the contralateral site of the electroporated brains were shown as normalized immunofluorescence intensities of callosal axons.  $F(4, 15) = 7.9591, P = 0.0012$  was determined by one-way ANOVA with post hoc Bonferroni test. Three to six mice per condition. \* $P < 0.05$ . Impaired axon terminal arborization across all cortical layers elicited by *C11orf46* knockdown was restored by overexpression of *C11orf46*<sup>Wt</sup>, but not by overexpression of *C11orf46*<sup>R236H</sup>. **d** *C11orf46* knockdown did not have an effect on radial neuronal migration at P2. Scale bar, 50  $\mu$ m. Bar graphs indicate mean  $\pm$  S.E.M. (**c, d**)

decreased by R236H substitution (Fig. 3h). Notably, both *C11orf46* with R236H mutation and that with deletion of CRD domain still kept binding affinity with histone H3 (Fig. 3i). Also, purified, *E. coli* derived, human *C11orf46* specifically detected histone H3 tail modifications in a histone peptide array (Supplementary Fig. 4c–e). Similarly, reciprocal affinity purification with FLAG-tagged SETDB1 using inducible cell lines confirmed *C11orf46* as a top ranking binding partner together with other members of the complex, including KAP1 and MCAF1 (Fig. 3j–m and Supplementary Table 2). Additional co-immunoprecipitation experiments with mouse anti-*C11orf46* antibody in HeLa cells, and in homogenates of human cerebral cortex confirmed that *C11orf46* assembles with SETDB1, MCAF1 and other regulators of repressive chromatin, including heterochromatin-associated protein 1gamma (HP1 $\gamma$ ) (Fig. 3n and Supplementary Fig. 4f). These results suggest that *C11orf46* may function as a chromatin regulator in the SETDB1 complex (Fig. 3o).

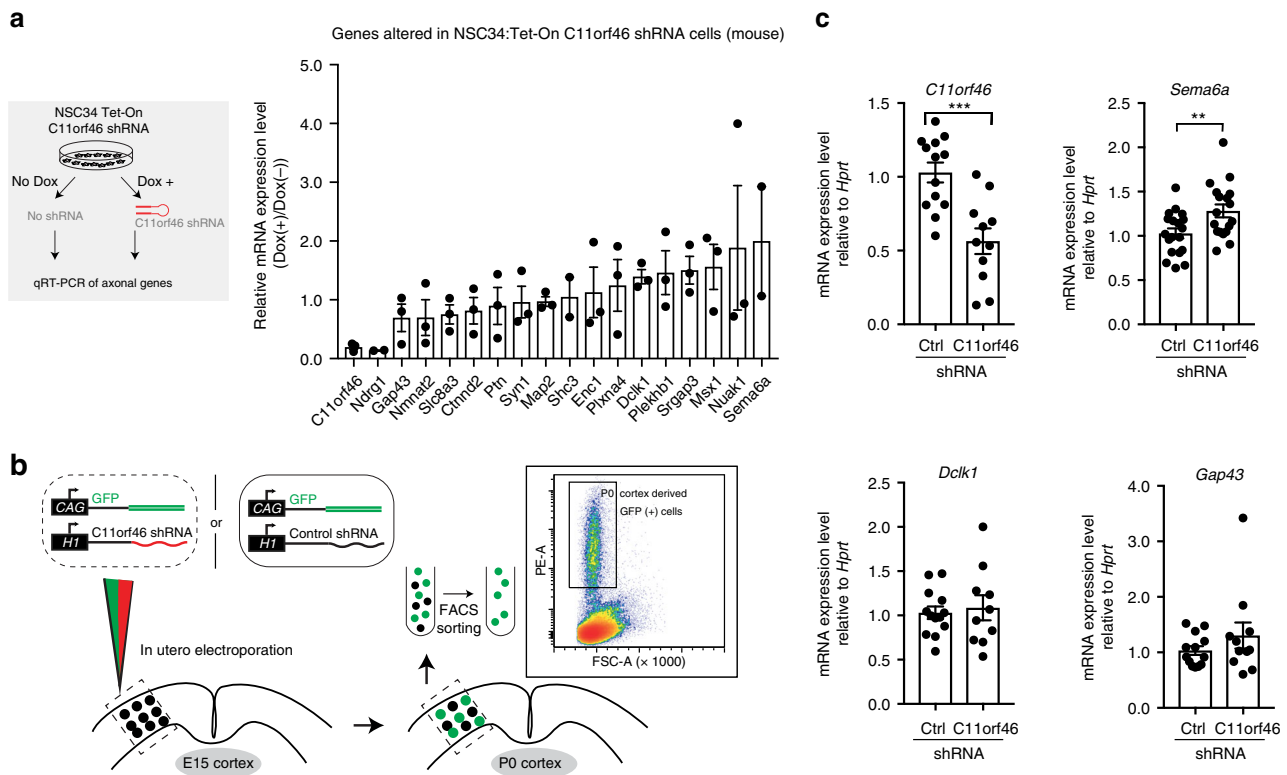
**Suppression of *C11orf46* altered expression of axonal genes.** Having shown that *C11orf46* is a neuronal nuclear protein bound to repressive chromatin regulators, we hypothesized that the transcallosal dysconnectivity resulting from *C11orf46* knockdown in fetal mouse cortex and of *C11orf46* deletions in WAGR patients could reflect dysregulated expression of genes critical for axonal growth and development. In order to obtain first insights into *C11orf46*-sensitive gene expression changes in WAGR deletion cases, we analyzed a RNA-seq dataset of lymphoblastoid cell lines obtained from the aforementioned 2 WAGR cases with *C11orf46* haploinsufficiency compared to 2 controls, including one WAGR case with a shorter microdeletion that retained intact *C11orf46* and one healthy control subject with no loss of genes for the entirety of chr. 11p13–14.1 (Supplementary Fig. 1b). We identified 227 differentially expressed genes (182 down-regulated genes and 45 upregulated genes) (FDR < 0.05, see Methods; Supplementary Fig. 5 and Supplementary Table 3) that include 41 genes which



have KAP1 binding sites<sup>33</sup> (Supplementary Table 4). Notably, among these 41 genes, we found that 17 genes are reportedly involved in axonal development (Fig. 4a, Supplementary Table 4 genes marked in bold), while the remaining 24 genes, such as CAP2 encoding an adenylyl-cyclase associated protein, are linked to various other cellular functions (Supplementary Table 4). However, given that the observed gene alterations were likely confounded by deletion of other proximal genes with *C11orf46* at WAGR locus, we switched in our next set of experiments to neuron-like (neuroblastoma) NSC34 Tet-On cells expressing *C11orf46* shRNA to further examine *C11orf46*-regulated genes. Nonetheless, we found that expression levels of the 17 axonal growth and development genes were not

consistently affected in response to *C11orf46* shRNA-mediated *C11orf46* knockdown in the neuroblastoma cell line. However, multiple putative *C11orf46* target genes, including *Semaphorin 6a* (*Sema6a*), *Doublecortin-like kinase 1* (*Dclk1*), *SLIT-ROBO Rho GTPase activating protein 3* (*Srgap3*), and others showed increased expression after *C11orf46* knockdown in the NSC-34 (Fig. 4a). Therefore, we wanted to confirm in neurons, in vivo, *C11orf46*'s role as a transcriptional regulator during brain development, focusing on the three candidate genes with KAP1 binding sites proximal to transcription start site (TSS) (+/-750b)<sup>33</sup>, including two genes (*Sema6a*, *Dclk1*) with expression changes in response to *C11orf46* knockdown in the NSC-34. *Dclk1* is essential for axon tract formation across the

**Fig. 3** C11orf46 binds to the SETDB1 complex. **a** Flow diagrams for immunoaffinity purification of C11orf46 followed by mass spectrometric analysis. **b** Eight clones of inducible HEK293 cell lines expressing FLAG-tagged C11orf46 (clone 1-5 and 10) and a truncated form C11orf46 lacking amino acids 209-237 at the C-terminal region which corresponds to cysteine rich domain (CRD) of C11orf46 (C11orf46Δ) (clone 7 and 8) in the presence of doxycycline (Dox+). SETDB1 was co-precipitated with C11orf46, but not with C11orf46Δ. **c** Silver staining of proteins co-precipitated with C11orf46 (clone 10) or C11orf46Δ (clone 8). SETDB1, MCAF1 and KAP1 were detected in co-precipitate from C11orf46, but not of C11orf46Δ. **d** Top scoring proteins co-precipitated with C11orf46 in cell lines overexpressed with C11orf46, C11orf46Δ, and control cells are shown with % of amino acid sequence coverage. **e** Protein abundance in affinity purified C11orf46 complexes. Notice prominence of SETDB1-MCAF1-KAP1 repressor proteins (black). Bar graphs indicate mean ± S.E.M.  $n = 3$  independent quantification. **f** Schematic diagrams indicating position of C11orf46 arginine to histidine substitution at codon 236 (R236H) in conserved cysteine rich domain (CRD; amino acids 209-237) between human and mouse. **g, h** Co-immunoprecipitation experiments using protein lysates from HEK293 cells overexpressing FLAG-tagged C11orf46, C11orf46Δ, or C11orf46<sup>R236H</sup> showed that absence of CRD domain and R236H mutation is associated with weakened or non-detectable C11orf46-SETDB1 binding and partial loss of binding to other components. R236H mutation does not fully disrupt C11orf46-KAP1 binding. **i** R236H mutation and CRD deletion of C11orf46 do not affect binding to histone H3. **j** Flow diagrams for immunoaffinity purification of SETDB1 followed by mass spectrometric analysis. **k** Inducible SETDB1 expression system in HEK293 cell lines (48 h after Dox treatment). **l, m** Silver staining of proteins co-precipitated with SETDB1. Both endogenous C11orf46 and SETDB1 binding partners were detected by mass spectrometry. Top scoring proteins co-precipitated with SETDB1 in cell lines overexpressed with SETDB1 and control cells are shown. **n** (left) C11orf46-SETDB1 co-immunoprecipitation in human cerebral cortex. (right) SETDB1 co-immunoprecipitates include C11orf46, HPIγ, and MCAF1 in HeLa nuclear extract. **o** Schematic summary of above data representing working model of C11orf46-SETDB1 complex in relation to histone H3



**Fig. 4** Selection of targeted genes by C11orf46-SETDB1 complex. **a** Mouse neuroblastoma cells NSC34 inducibly expressing C11orf46 shRNA upon doxycycline addition (schematic in left panel); selected subset of 17 differentially expressed genes implicated in axonal growth and development was tested in RNA expressed from NSC34:Tet-On C11orf46 shRNA cell (right panel). **b** Neuronal progenitors received control or C11orf46 shRNA and GFP marker plasmids at somatosensory cortex by IUE at E15; four days later at P0 GFP-positive cells were collected by fluorescence-activated cell sorting (FACS). **c** C11orf46 transcript levels were reduced by 50% in sorted GFP-positive cortical neurons electroporated with C11orf46 shRNA1 as compared to GFP-positive neurons with control shRNA ( $P = 0.0003$ ). Note increase of Sema6a transcript upon C11orf46 knockdown ( $P = 0.0085$ ).  $*P < 0.05$ ,  $**P < 0.01$ ,  $***P < 0.001$  determined by Student's  $t$ -test.  $n = 10-19$  independent mice per condition. All data were normalized by the Hprt transcript. Bar graphs indicate mean ± S.E.M. (**a, c**)

anterior commissure in the ventral forebrain and the corpus callosum<sup>34</sup>, Sema6a encodes a PLXNA4 protein ligand<sup>35,36</sup>, and our third candidate gene, Gap43, encodes a growth cone protein essential for commissural axon guidance<sup>37</sup> (Supplementary Table 4). To confirm C11orf46's role as a transcriptional regulator during brain development, we knocked down C11orf46 by IUE-mediated delivery of C11orf46 shRNA together with

GFP expression plasmid at E15 and GFP-labeled projection neurons were isolated by fluorescence-activated cell sorting (FACS) at P0 (Fig. 4b). Notably, Sema6a was significantly increased in C11orf46 knockdown projection neurons (Fig. 4c). These findings, taken together, strongly suggest that C11orf46 regulates neuronal gene expression, including an inhibitory effect on Sema6a.

**C11orf46- epigenome editing rescues a callosal phenotype.** Our results, reported above, showed that C11orf46 co-assembles with the KMT-RC. Apparently, this interaction is important because mutations such as C11orf46<sup>R236H</sup>, disrupting C11orf46's association with the complex, are highly detrimental to orderly brain development. We hypothesized that C11orf46's affinity to the KMT-RC could be exploited, in order to 'edit' neuronal gene expression via a chromatin-associated mechanism. Such type of epigenome editing could be accomplished, for example, by RNA-guided fusion proteins comprised of nuclease-deficient clustered, regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (dCas9) bound to a transcriptional activator or repressor<sup>38</sup>. However, because '1:1' systems (one transcriptional regulator per dCas9 copy) are often only minimally effective<sup>38,39</sup>, we instead used dCas9-SunTag protein scaffolds which assemble as a binary system with the protein of interest fused to a single-chain antibody variable fragment (scFv)-GFP cassette, with the scFv binding to the Cas9-fused protein scaffold carrying 10 or more copies of GCN4 (the scFv epitope)<sup>39</sup>. We built a dCas9-SunTag system to load ten copies of C11orf46, or the VP64 activator protein as a positive control, onto a single sgRNA-target sequence (Fig. 5a). We began with our epigenome editing experiments in HEK293 cells selecting *Sema6a*, *Dcl1*, and *GAP43*. For each gene, we obtained four sgRNAs positioned within 500 bp upstream of the target TSS. For each of the three target genes, three groups of HEK293 cells were compared, based on transfection with (1) dCas9-ST alone (sgRNA-dCas9-10xGCN4<sup>SunTag</sup>) or dCas9-ST together with one of the following plasmids, (2) scFv which recognizes the GCN4 epitope on dCas9-ST-superfold GFP-VP64, or (3) scFv-superfold GFP-wild-type C11orf46 (C11orf46<sup>wt</sup>) (Fig. 5a). After nuclear localization of the scFv-sfGFP-C11orf46 domain was confirmed (Fig. 5b), testing for each target gene experiments were conducted in parallel in HEK293 cells, with each of the four sgRNA tested individually (and in a pool of four for *SEMA6A* as well). Interestingly, introduction of two of the four sgRNAs to *SEMA6A* which is expressed at much higher levels as compared to the other test genes (Fig. 5c), resulting in significant decrease in *SEMA6A* expression when co-transfected with the dCas9-ST<sup>10x,C11orf46(wt)</sup> systems (Fig. 5d). Conversely, *DCLK1* and *GAP43*, which is expressed at extremely low levels at baseline (Fig. 5c), showed minimal changes when co-transfected with dCas9-ST<sup>10x,C11orf46(wt)</sup>, while responding to dCas9-ST<sup>10xVP64</sup> with a robust, at least 5–10 fold increase in their expression (Fig. 5d).

Our studies in HEK293 cells suggest that C11orf46, when tethered in multiple copies to specific promoter sequences via the CRISPR-Cas9-SunTag system, could actively repress transcribed genes. Next, we wanted to explore this approach in vivo in developing upper layer cortical projection neurons of somatosensory cortex. We focused on *Sema6a*, which was depressed after *C11orf46* knockdown in NSC34 cells and GFP-positive neurons (Fig. 4a, c). Both *Sema6a* and *C11orf46* are highly expressed in upper cortical layers II/III at P0, within the targeted time of brain development, via IUE (Supplementary Fig. 6a). Two sgRNAs, targeting different sequences within the first 500 bp upstream from the TSS of *Sema6a*, were selected for in vivo studies and delivered at E15 (Fig. 5e). Strikingly these *Sema6a* promoter-targeting sgRNAs rescued impaired midline crossing of callosal projections and axonal arborization deficit normally encountered upon *C11orf46* knockdown at E15 by IUE (Fig. 5f, g and Supplementary Fig. 6b). Consistent with C11orf46's association with KMT-RC, introduction of *Sema6a* sgRNA also resulted in increased H3K9me3 levels at the *Sema6a* promoter, in striking contrast with a decrease in H3K9me3 levels after *C11orf46* knockdown (Fig. 5h). These

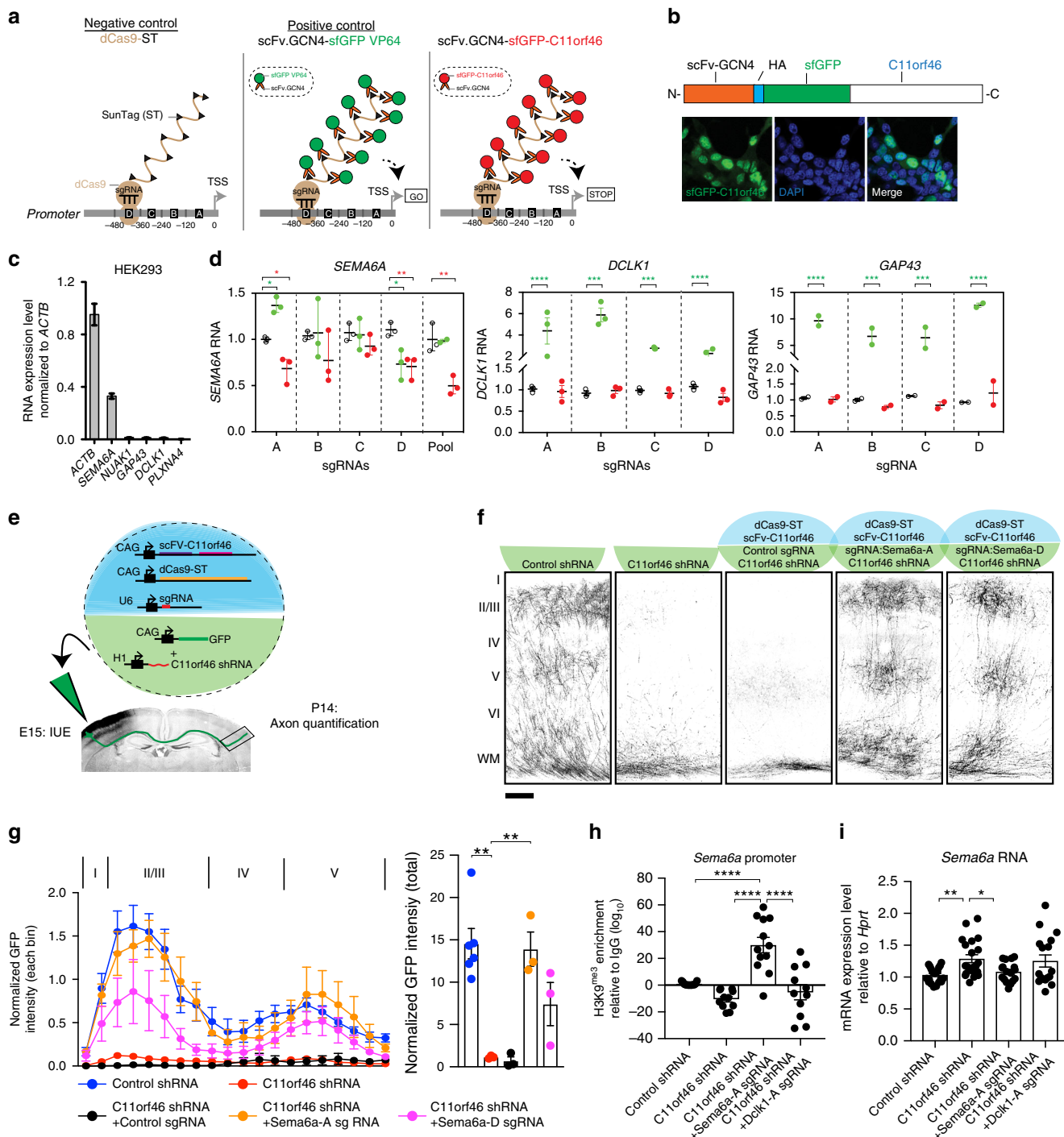
effects were highly specific for the target gene sequence, because sgRNAs directed against *Dcl1* promoter did not affect the *Sema6a* promoter (Fig. 5h). Likewise, increased *Sema6a* expression in *C11orf46* knockdown neurons was strongly suppressed by our SunTag-based C11orf46-mediated epigenome editing toolkit targeting the *Sema6a*, but not the *Dcl1* promoter (Fig. 5i). Consistently, overexpression of *Sema6a* elicited disturbance of midline crossing of callosal projections and axonal terminal arborization in the developing cerebral cortex (Supplementary Fig. 6c).

## Discussion

Axonal development is a pivotal process in establishing neural connectivity for brain functions, including cognition and memory<sup>40</sup>. This critical stepwise process coordinately regulates the establishment of synaptic connections to target sites with topographic specificity<sup>41</sup>. Disrupting these processes may lead to altered neural connectivity, which may underlie cognitive disabilities in neurodevelopmental psychiatric conditions such as intellectual disability and autism spectrum disorder<sup>42</sup>. While multiple molecular pathways including guidance cues and growth factors have been implicated in axonal development<sup>41</sup>, very little is known about chromatin-regulatory mechanisms<sup>43</sup>. Recently, the SET (histone methyltransferase) domain-containing transcriptional repressor, PR/SET domain 8 (PRDM8), was described as essential for the formation of trans-callosal and other inter-hemispheric fiber tracts<sup>44</sup>. Likewise, Heterochromatin-associated protein 1gamma (HP1γ) is required for the crossing of callosal axons into the contralateral hemisphere, while increased neuronal HP1γ expression promotes growth and differentiation of dendrites and axons<sup>45</sup>. Interestingly, PRDM8 promotes H3K9 methylation<sup>46</sup> and HP1γ is bound preferentially to compacted nucleosomal arrays bearing the repressive H3K9me3 mark<sup>47</sup>. Furthermore, the C11orf46-binding partner and repressive H3K9MTase, SETDB1, has been implicated in transcriptional regulation of S-type clustered Protocadherin genes, a group of cell adhesion molecules broadly important for neuronal connectivity<sup>48</sup>. These findings, when taken together with our results reported here, imply H3K9 methyl-regulators as important epigenomic determinant of the developing cortical connectome.

In this study, we showed that C11orf46 assembles with bona fide factors of histone-3 lysine-9 trimethyltransferase (H3K9me3) SETDB1 complex, it substrate histone H3 and a subset of nuclear proteins functionally associated with repressive and heavily H3K9-methylated (hetero) chromatin, including Lamin B1 (LMNB1), Lamina-associated protein 2, isoform alpha (LAP2A) and HP1γ. These multi-molecular assemblies of chromatin-bound proteins presumably require distinct domains of C11orf46 to associate with SETDB1 complex and histone H3. Importantly, non-synonymous genetic mutations G218E and R236H found in the CRD domain of C11orf46 are associated with neurodevelopmental disease conditions<sup>13,49</sup>. We show that specifically the R236H, disrupted the C11orf46 assembly with the SETDB1 complex but not C11orf46's association with histone H3. Therefore, mutations in critical domains of C11orf46 could disrupt the above assembly and potentially interfere with spatio-temporal gene silencing required for orderly development of transcallosal connectivity. Notably, purified C11orf46 showed strong affinity for histone H3 tail modifications in vitro; we were surprised to observe C11orf46's high affinity towards a histone mark antagonizing histone methylation, H3R2-citrulline (H3R2Cit)<sup>50,51</sup>. Future studies will explore in more detail the regulatory role for C11orf46. For example, C11orf46 could sense the levels of mono- or di-methylated H3K9 to promote





trimethylation, or block H3K9 methylation in response to high levels of antagonizing H3R2Cit rank.

In addition to protein components of the SETDB1 complex, such as KAP1 and MCAF1, our results of proteome screening identified LMNB1 and LAP2A, two nuclear lamina proteins that were previously discovered at the heterochromatin rich nuclear periphery<sup>52</sup>, as potential protein interactors of C11orf46. Interestingly, LMNB1 and LAP2A play critical roles for developmentally regulated dendritic and axonal connectivity<sup>53</sup>. Because C11orf46 resulted in impaired axonal development, but had no robust effect on dendritic structure, it will be important to further investigate epigenetic mechanisms to determine how C11orf46 specifically regulates axonal development.

To the best of our knowledge, this is the first study demonstrating the utility of the nuclease-deficient CRISPR/Cas9 system-mediated epigenome editing with IUE knockdown approach for addressing cell-autonomous epigenetic machinery for regulation of transcallosal connectivity. By targeting transcallosal projection neurons, we demonstrate that C11orf46 regulates axonal terminal arborization by control of expression of *Sema6a* via epigenomic promoter H3K9-methylation. The corpus callosum phenotypes in WAGR cases with *C11orf46* haploinsufficiency (Fig. 1) and the successful epigenomic rescue by sequence-specific C11orf46-mediated chromatin remodeling (Fig. 5), would imply that future chromatin-based therapies could be harnessed to correct developmental errors in the brain's connectome. Importantly, genetic

**Fig. 5** C11orf46- epigenomic editing of neurite-regulating genes rescues transcallosal dysconnectivity in C11orf46-deficient neurons. **a** Schematic representation showing epigenome editing using nuclease-deficient CRISPR/Cas9 SunTag (dCas9-ST) system. Single guide RNA (sgRNA) targeting neuronal gene promoters is introduced with dCas9-ST (10xGCN4) and scFv-sfGFP-C11orf46 (or -VP64 as a positive control) in HEK293 cells. **b** Overexpressed fGFP-C11orf46 is colocalized with DAPI signals in HEK293 cells. **c** Relative expression levels of neurite-regulating genes in HEK293 cells. **d** dCas9-ST mediated promoter loading with 10xVP64, and C11orf46<sup>R236H</sup> in comparison to 10xVP64 at *SEMA6A* in HEK293 cells. Four sgRNAs (**a-d**) for each gene were tested individually, and for *SEMA6A* also as a pool. Two-way ANOVA followed by Dunnett's multiple comparisons test was performed where C11orf46- or VP64-epigenetic editing effects were tested. **e** Epigenomic editing of transcallosal neurons by IUE. Five transgenes were delivered simultaneously at E15 by IUE into the developing somatosensory cortex, followed by quantification of axonal arborization at P14. **f, g** Axonal arborization was disrupted by *C11orf46* knockdown, and restored by dCas9-ST mediated recruitment of C11orf46 to *Sema6a* promoter (*Sema6a*-), using sgRNAs *Sema6a*-A, *Sema6a*-D, but not with non-targeting sgRNA. Scale bar, 100  $\mu$ m. Layer distributions and total axonal density in the contralateral site of the electroporated brains are shown (**g**).  $F(4, 13) = 13.1934$ ,  $P = 0.0002$  was determined by one-way ANOVA with post hoc Bonferroni test. Three to six mice per condition. **h** H3K9me3 enrichment at *Sema6a* promoter in GFP + neurons at P0, four days after IUE. Epigenome editing by *Sema6a*-A sgRNA increases H3K9me3 levels at *Sema6a* promoter. Dcl1-A sgRNA did not affect H3K9me3 levels at *Sema6a* promoter.  $F(3,44) = 8.718$ ,  $P = 0.0001$  was determined by one-way ANOVA with post hoc Bonferroni test. **i** *Sema6a* mRNA level in GFP + neurons at P0. Increased *Sema6a* expression caused by C11orf46 knockdown were suppressed by C11orf46-epigenome editing using *Sema6a*-A sgRNA, but not by Dcl1-A sgRNA.  $F(3,72) = 4.746$ ,  $P = 0.0045$ . \* $P < 0.05$ , \*\* $P < 0.01$  determined by one-way ANOVA with post hoc Bonferroni test. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , and \*\*\*\* $P < 0.0001$ . Bar graphs indicate mean  $\pm$  S.E.M.

deletion of *Sema6a* resulted in impaired neuronal connectivity<sup>54,55</sup>. SEMA6A is a transmembrane protein and is known as a ligand for Plexin A2 (PLXNA2) and Plexin A4 (PLXNA4) which also play critical roles for regulating axon guidance<sup>56</sup>. *Plxn2* and *Plxn4* knockout mice showed impaired axonal development similar to those observed in mouse models with genetic deletion of *Sema6a*<sup>56</sup>. Thus, dysregulation of C11orf46-mediated control of *Sema6a* expression may impair axonal development regulated by Plxns either in cell autonomous or non-autonomous manner, which warrants further investigation in the future.

Our results, however, should carefully be interpreted for several reasons. First, although IUE-knockdown approach is a powerful method to examine the cell-autonomous effect of C11orf46 in the specific time critical for axonal growth, off-target effect of shRNAs might be confounded in the phenotypes<sup>57</sup>. To ensure that observed results do not originate from their off-target effects, we confirmed dose-dependent effects on the axonal phenotypes induced by two shRNAs with independent target sequences. Rescue experiments by co-expression of RNAi-resistant wild-type C11orf46, containing silent mutations in the target sequence of their shRNA, also ensured that the observed phenotypes are specifically due to its knockdown effect. Nevertheless, C11orf46-mediated mechanisms for axonal development and resulting WAGR syndrome-related phenotypes should be investigated in mice with genetic deletion of C11orf46, which will be considered for future studies. Second, we utilized RNA-seq data obtained from lymphoblastoid cell lines to identify differentially expressed genes involved in axonal development in WAGR cases, compared to those in controls. Lymphoblastoid cells may not represent brain-specific epigenome. Nevertheless, lymphoblastoid cells remain to be a convenient and well-characterized tool to elucidate brain disease-related molecular phenotypes<sup>58</sup>. Importantly, our RNA-seq data sets with matched controls revealed not only lymphoid genes, but also genes critical for neuronal function, representing the disturbances of neuronal-like epigenome in these lymphoid cells. Third, the discrepancy of SEMA6A expression in lymphoblastoid cells and neuronal conditions likely arise from the confounding effect of other genes located in the WAGR risk locus, including PAX6 transcription factor. Epigenomic signatures of lymphoid cell line as well as the developing brain may also contribute to this discrepancy. Fourth, because human subjects in our study were recruited on the basis of having aniridia, all participants were, by definition, PAX6 haploinsufficient, and, therefore, we were unable to assess the effect of C11orf46<sup>+/-</sup> independent of PAX6<sup>+/-</sup> and the phenotype of isolated

C11orf46<sup>+/-</sup> is unknown (although homozygous mutations in C11orf46 have been reported in association with intellectual disability). PAX6 is a transcription factor regulating neurogenesis and rostrocaudal patterning during development<sup>59,60</sup>. PAX6 haploinsufficiency is associated with corpus callosum hypoplasia in both rodents and humans<sup>15,16,61,62</sup>. Thus, there remains the possibility that additional loss of PAX6<sup>+/-</sup> may be required for C11orf46<sup>+/-</sup> to cause significant defects in morphologic brain development. However, we were able to demonstrate that patients with combined loss of both genes had more severely reduced corpus callosum volumes compared to isolated PAX6<sup>+/-</sup>, confirming an additional role of C11orf46 in neurodevelopment. Whether the impact of mutations in the above molecules on axonal development is additive or synergistic remains to be determined. In addition, because the WAGR CNV encompasses a large genomic regions that encompass many genes, hampering identification of genetic drivers responsible for specific phenotypes shown in CNVs-associated disease conditions, our data do not exclude the possibility that other genes in 11p13 deletion region, besides *BDNF*, *PAX6*, and *C11orf46* may have independent effect on axonal and other anatomical phenotypes as well as behavioral outcomes in WAGR syndrome.

In summary, we identified a chromatin-associated mechanism underlying axonal development and show that C11orf46 plays an important role in transcallosal phenotypes in 11p13 deletion syndrome. Further investigation on epigenetic regulators in CNVs associated with neurodevelopmental disorders and altered neural connectivity is important for translational approaches that pave the way for novel treatment of neurodevelopmental psychiatric conditions in early life.

## Methods

**Human subjects.** Patients who had prior genetic testing confirming diagnosis of WAGR/11p13 deletion syndrome or isolated aniridia with known PAX6 mutation or deletion, as well healthy control subjects who had no chronic medical conditions were recruited through local advertisements and on-line postings. The study was approved by the institutional review board of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD and was registered at <https://clinicaltrials.gov/> as NCT00758108. Written informed consent was obtained from adult subjects who were competent to provide consent and from the parents or legal guardians of children and adults with cognitive impairment. Clinical studies were performed between December 2008 and May 2014 at the NIH Clinical Research Center.

**Genetic testing.** Deletion boundaries for each subject with WAGR syndrome were determined using oligonucleotide array comparative genomic hybridization using a custom-designed microarray platform (Agilent Technologies, Inc., Santa Clara, CA) containing 105,000 60-mer oligonucleotide probes using NCBI Build 36

(hg18) human reference sequence as previously described<sup>63</sup>. Probes for chromosome 11p were spaced at approximately 400 bp intervals (excluding repeat regions) using 57,925 probes; 121 probes were located within *BDNF*.

**Brain magnetic resonance imaging (MRI).** Brain MRI consisted of one cubic millimeter resolution, T1-weighted images collected on a 3.0 T Philips Achieva MRI scanner with an 8-channel phased array head coil. Corpus callosum volumes were calculated using FreeSurfer's (version 5.3) subcortical image processing pipeline and published methods<sup>64</sup>. Briefly, the pipeline uses prior probability of a given tissue class at a specific atlas location, the likelihood of the image intensity given the tissue class, and the probability of the local spatial configuration of labels given the tissue class. The measured corpus callosum volume extends 2.5 mm from the midline on both sides to mitigate against any residual misalignment after registration to the template. Analysis of MRI data was performed blinded to genetic diagnosis of each patient. Each subject had a single structural MRI scan. MRI brain scans were normalized to a T1-weighted template (MNI305) using a 12-parameter affine transformation. No artifact or structured noise removal was applied to the MRI scans.

**Lymphoblastoid cell lines.** Fasting venous blood samples were obtained from patients (Supplementary Table 3). Peripheral blood mononuclear cells were isolated using a Ficoll-Paque gradient (GE Healthcare Life Sciences, Pittsburgh, PA), then infected with Epstein Barr virus, and after sufficient cell line expansion, stored in freezing media at  $-80^{\circ}\text{C}$ .

**Antibodies.** A mouse monoclonal antibody against Flag-tag (M2) (Sigma, Cat #F1804) was used for affinity purifications. For immunoblots, the following antibodies were used: rabbit monoclonal antibody against the C-terminus of human SETDB1 (Cell Signaling technology, Cat #2196S), KAP1 (cell signaling technology, Cat#4124S); rabbit polyclonal antibody against SETDB1 (Santa Cruz, Cat #sc-66884X), MCAF1 (Novus biological, Cat #NB100-438), Histone H3 C-terminus (Millipore, Cat #07-690). Custom-made mouse monoclonal antibody and guinea pig polyclonal antibody were generated against *E. coli* derived, purified and mass spectrometrically verified full-length human His-tagged C11orf46 protein (Supplementary Fig. 4b). For details on the generation of mouse monoclonal antibody, see the references<sup>65,66</sup>. Altogether, total 45 individually picked anti-C11orf46 mouse monoclonal hybridoma clones were screened by Western blot using *E. coli* derived human C11orf46 as antigen. Five clones (clone numbers 9, 26, 28, 33, and 36) were positive in this assay (clone 26 shown as representative example in Supplementary Fig. 4b). Antibodies from these clones were pooled and used for the co-immunoprecipitation assay using tissue homogenates of human prefrontal cortex. The guinea pig antibodies were generated by Cocalico Biologicals (Reamstown, PA) using affinity purified human recombinant His-tagged full-length C11orf46. Rabbit polyclonal antibodies against histone H3 lysine-9 monomethyl (Millipore, Cat #07-450), H3 lysine-9 dimethyl (Upstate, Cat #07-441), and H3 lysine-9 trimethyl (Upstate, Cat #07-442) were used for co-immunoprecipitating endogenous SETDB1 from HeLa nuclear extract. The antibodies which were used for immunohistochemistry or chromatin immunoprecipitation (ChIP) are described in each section.

**Plasmids.** Plasmids expressing interfering shRNA were generated to suppress endogenous C11orf46 protein expression utilizing the pSUPERIOR.puro vector system (Oligoengine). Their target sequences are: C11orf46 shRNA-1 with strong suppression; 5'-CAAACCTGAATTGCTCCAGAA-3' and C11orf46 shRNA-2 with milder suppression; 5'-GAAGACAGCTTGTACTGGTT-3'. A scrambled sequence that shows no homology to any known messenger RNA was utilized to produce the Control shRNA (5'-ATCTCGCTTGGCGAGAGTAAG-3'). The HA and scFv-tagged RNAi resistant wild-type human C11orf46 expression constructs containing four silent mutations (underlined) in the target sequence of C11orf46 shRNA (5'-CAAACAGAGTTCGCACCAGAA-3') were produced and cloned into CAG promoter driven plasmid (pCAGGS1vector) for rescue experiments. dCas9-SunTag coding sequence was also transferred into pCAGGS1 vector. Two single guide RNA sequences each *Dcl1* or *Sema6a* promoter region were using online CRISPR design tool<sup>67</sup>. Those protospacer sequences were cloned into sgRNA cloning vector (Addgene #41824) with the protocol distributed on the addgene website (<https://media.addgene.org/data/93/40/adf444fe-5e77-11e2-9c30-003048dd6500.pdf>). The protospacer sequences are listed in Supplementary Table 5. CAG promoter driven mouse *Sema6a* expression plasmid (CAG-mSema6a) was cloned using *Sema6a* (NM\_018744) Mouse Tagged ORF clone (ORIGENE, Cat#MR211486).

**In vitro epigenomic editing using dCas9-SunTag-C11orf46.** For producing scFv-GCN4-sfGFP-C11orf46, full-length wild-type C11orf46 (C11orf46<sup>wt</sup>) or R236H mutant C11orf46 (C11orf46<sup>R236H</sup>) was cloned in-frame to the C-terminus of scFv GCN4 antibody using the BamHI and NotI sites (addgene plasmid #60904). dCas9-SunTag (2  $\mu\text{g}$ , addgene #60903) containing 10 copies GCN4 peptide epitope fused to catalytically inactive Cas9 was delivered into HEK293FT cells ( $1.5 \times 10^6$ , 60 mm dish) along with a 4 pooled sgRNAs (50 ng each sgRNA, U6-sgRNA-polyT amplicon) to direct GCN4 single-chain antibody conjugated fluorescent C11orf46

(2  $\mu\text{g}$ , scFv-GCN4-sfGFP-C11orf46) to neurite-regulating gene promoters. Three days after transfection using lipofectamine 2000, total RNA was extracted and qPCR was performed. scFv-GCN4-sfGFP-VP64 was used as a positive control for activating gene expression. dCas9 alone was tested as a negative control to account background gene activity. For producing U6-sgRNA-polyT, mouse [Dec.2011 (GRCh38/mm10)] and human promoter sequences (500 bp) were identified using PromoterWise (<http://www.ebi.ac.uk/Tools/psa/promoterwise/>). Four sgRNAs spaced approximately 500 bp away from each other in the promoter region of the target genes were identified based on their high gRNA score and low/no off-target site in the genome (Supplementary Table 5). sgRNAs were incorporated into U6-sgRNA-polyT cassettes by PCR amplification using U6-Fwd and U6sgRNATER-Rev primers which carried the reverse complement of part of the U6 promoter. These cassettes include the sgRNA (+85) scaffold with guide sequence and seven T nucleotides for transcriptional termination.

**In utero electroporation (IUE).** IUE targeting the somatosensory cortex was performed using our previously published methods with minor modifications<sup>20,21</sup>. Pregnant C57/BL6 mice were anesthetized at embryonic day 15 (E15) by intraperitoneal administration of a mixed solution of Ketamine HCl (100 mg/kg), Xylazine HCl (7.5 mg/kg), and Buprenorphine HCl (0.05 mg/kg). After the uterine horn was exposed by laparotomy, the shRNA plasmid (1  $\mu\text{g}/\mu\text{l}$ ) together with CAG promoter-driven GFP expression plasmids (1  $\mu\text{g}/\mu\text{l}$ ) (molar ratio, approximately 1:1) were injected into the lateral ventricles with a glass micropipette made from a micropipillary tube (Narishige, Cat #GD-1), and electroporated into the VZ of the CP at E15. C11orf46 expression constructs (1  $\mu\text{g}/\mu\text{l}$ ) were also co-introduced for rescue experiments. For in vivo epigenomic editing, CAG-dCas9-ST (1  $\mu\text{g}/\mu\text{l}$ ), scFv-C11orf46 (1  $\mu\text{g}/\mu\text{l}$ ) and sgRNAs (1  $\mu\text{g}/\mu\text{l}$ ) together with shRNAs and GFP expression plasmid will be introduced at E15. For overexpression of *Sema6a*, CAG-*Sema6a* (2  $\mu\text{g}/\mu\text{l}$ ) or tdTomato (1  $\mu\text{g}/\mu\text{l}$ ) with GFP expression plasmid will be introduced at E15. For electroporation, Electrode pulses (40 V; 50 ms) were charged four times at intervals of 950 ms with an electroporator (Nepagene, Cat #CUY21EDIT). All experiments were performed in accordance with the institutional guidelines for animal experiments of Johns Hopkins University.

**RNA-seq and bioinformatic analysis.** Total RNA was isolated from four lymphoblastoid cell lines including one derived from healthy control subject, one WAGR patient with *C11orf46* gene intact and two WAGR patients with *C11orf46* gene deletions using a total RNA extraction kit (Qiagen) according to manufacturer instructions. RNA integrity determined in Bioanalyzer (Agilent's Technologies) and a total RNA-seq was performed at Eurofins genomics. DNA libraries were prepared from total RNA samples, the adaptor-ligated libraries (125 bp) were enriched by PCR amplification and gel purified (average library size of ~270 bp). Libraries were sequenced with the Genome Analyzer II (Illumina) using pair-end 50-bp raw fastq reads. The sequencing approach resulted in 32,750,145 and 27,929,697 and 39,292,408, and 54,377,092 reads for lymphoblastoid cell line sample numbers one, two, three, and four, respectively. Sequencing raw reads were assessed for quality control using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). RNA-seq reads of 50 bp were aligned against the human genome (UCSC hg19) references with TopHat2 using default parameters with Bowtie as the internal aligner and a segment mapping algorithm to discover splice junctions<sup>68,69</sup>. In all, 77–82% of all unique sequences correctly aligned to the reference genome. Reads were counted in exon regions using featureCounts tool (subread v1.5.2) producing a table that was normalized to rpkm (Supplementary Table 3). Heatmaps were produced scaling rpkm values by scale function (base v3.5.1) and colors were adjusted by quantile breaking to allow colors represent an equal proportion of the data.

Two different techniques were used in order to assemble the transcriptome: (i) reference-guided assembly of RNA-seq reads using reference annotations from Refseq<sup>70</sup> and (ii) non-reference guided assembly (without the help of a reference gene model) to yield a 'transcript reference-free' characterization of RNA-seq reads. For the first approach, BioConductor packages DESeq (References) and EdgeR v2.4.6 in the R programming environment were used. HTSeq-count (<http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>) was used with UCSC annotation to generate a set of per-gene read counts for each sample, using default parameters including 'stranded' set to 'no'. The HTSeq-count-generated matrix count dataset was used as input. The quartile of genes with the lowest counts was excluded to increase the detection power at a given FDR. For the second approach, the Cufflinks-CuffDiff-CummeRbund pipeline was used<sup>69</sup>. Fragments per kilobase per millions of reads (FPKM) values were computed for each cell line sample. Genes that were significantly differentially expressed with a FDR of 0.05 in DESeq and EdgeR were selected for further analysis. Both methods yielded highly reproducible results in the expression measurements of the transcripts, and the majority of differentially expressed genes were also significant in an independent analysis using Cufflinks.

**Generation of tet-inducible cell lines.** Human full-length *SETDB1* complementary DNA (cDNA), identical to SETDB1 transcript variant 1 (NM\_001145415.1; 3876 bp, 1291 amino acids, predicted molecular weight 143 kDa), was reverse-transcribed using primer Cjk-SA1. Total RNA was isolated from

human brain derived U87-MG glioblastoma cells (ATCC, Cat #HTB-14) and was subsequently amplified by PCR (Cjk-SA2 & Cjk-SA4). Sall-NotI site was double digested and cloned into pENTR4 and transferred to the pFRT TO DESTFLAGHA vector by a LR reaction. The plasmid was completely sequenced. The FLAGHA epitope contributed 17 amino acids to the N-terminal of SETDB1. These amino acids fused at the N-terminus to the ~60 bp (1xFLAG)- (1xHA) epitope (FLAG-HA-SETDB1) downstream of a promoter containing the tetracycline-responsive element (pFRT-TO-FLAGHA-SETDB1) and stably transfected (hygromycin) into Flp-In™ T-Rex 293™ cells (ThermoFisher, Cat #R780-07) together with plasmid pOG44 (expressing FLP recombinase). Full-length *C11orf46* cDNA (NM\_152316.2; 270 amino acids; 783 bp) was amplified by PCR from an Incyte plasmid (Openbiosystem, Cat #IHS1380-97432875) using primers Cjk24 and Cjk25, cloned into Sall-NotI restriction sites of the pENTR4 vector. A similar procedure was used to generate HEK293 cells for tet-inducible expression of a FLAG-HA-tagged full length *C11orf46* cDNA (NM\_152316.1 784 bp, 260 amino acids). For cloning, high fidelity reverse transcriptase (AccuScript Hi-Fi; Agilent Technologies #600180) and DNA polymerase were used (Pfu Ultra DNA polymerase; Agilent Technologies #600380). All plasmids were sequence verified and the primers used in their generation are listed in Supplementary Table 6.

**Affinity purification of protein complexes.** Inducible (Tet-On) HEK293 cell lines expressing SETDB1 or C11orf46 or control cells ( $10^7$  cells/assay) were treated with 2 µg/ml of doxycycline for 48 h before harvest. Pelleted cells were suspended in CS100-0.02 buffer (Tris HCl, pH 7.4, 10 mM MgCl<sub>2</sub>, 2.5 mM NaCl, 100 mM 0.02% NP-40, 2 units/ml Benzamide, protease inhibitor, and phosphatase inhibitor) at four times, packed cell volume, sonicated, passed five times through a 26-gauge needle, centrifuged at 20,000 × g for 15 min, and filtered through a 0.2-µm filter (Millipore, Cat #SCGP00525). Benzamide was added to eliminate nucleic acid-dependent indirect interactions. Lysate was passed through an anti-FLAG agarose column (packed using 0.5 ml of 50% slurry per 100 mg of extracts) at 4 °C and washed four times with cold CS100-0.02, then four times with cold CS500-0.02 (Tris HCl, pH 7.4, 10 mM MgCl<sub>2</sub>, 2.5 mM NaCl, 500 mM 0.02% NP-40, protease and phosphatase inhibitors). Column-bound protein complexes were eluted using 10 bed volumes of the CS100-0.02-containing 3X-FLAG peptide (100 ng/µl) at 4 °C. Purity of protein complexes was verified with four to fifteen percent gradient gel/silver stain and analyzed using mass spectrometry at the proteomics core facility in University of Massachusetts.

**In situ hybridization.** The following Cyanine dyes-conjugated RNA probes were designed by assistance of Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>), C11orf46 antisense RNA probe: 5'-Cy3-AGCUUUAGCAGUCCU UCCAUUGA-3' and *Sema6a* antisense RNA probe: 5'-Cy5- GGAACUGGCG UGCCAUUCAGUGC-3' for each target gene, and NS (non-targeting) RNA probe-1: 5'-Cy3-AUCUCGCUUGGGCGAGAGUAAG-3' to determine background signal. All probes were synthesized by Integrated DNA Technologies, Inc. Mouse brains were extracted after perfusion with 4% paraformaldehyde (PFA). The fixed brains were embedded in cryocompound (Sakura) after replacement of PFA with 30% sucrose in PBS. Coronal sections were obtained at 20 µm with a cryostat (Leica, Cat #CM 3050S). The sections were washed with DEPC-treated PBS (DEPC-PBS), followed by permeabilization by ice-cold methanol for 30 min. Hybridization buffer (50% Formamide, 10% Dextran sulfate, 0.1 mg/ml Yeast tRNA, 2× SSC, 50 mM Sodium phosphate) including 15 ng/µl of RNA probes were heat-denatured at 95 °C for 5 min and 1/500 vol. RNase inhibitor (20U/µl) was added to the buffer. After washing by DEPC-PBS, the permeabilized brain slices were hybridized with RNA-probes in hybridization buffer overnight at 37 °C. The samples were then washed once with wash buffer (40% Formamide, 1× SSC) for 30 min at 37 °C. Final three-time washing steps by DEPC-PBS included nuclei labeling with DAPI (Roche, 10236276001). Fluorescent images were collected by a confocal microscope (Zeiss, LSM 700).

**Immunohistochemistry.** Immunohistochemistry was performed using our previously published methods with some modifications<sup>20,21</sup>. Mouse brains were extracted after perfusion with 4% PFA. The fixed brains were embedded in cryocompound (Sakura) after replacement of PFA with 30% sucrose in phosphate buffered saline (PBS). Coronal sections including somatosensory cortex were obtained at 40 µm with a cryostat (Leica, Cat #CM 3050S). The sections were washed with PBS containing 0.5% Triton X-100 and then blocked with 0.5% Triton X-100 and 1% bovine skin gelatin for one hour. After blocking, sections were incubated with various cell markers [mouse anti-Nestin antibody (BD Biosciences, Cat #556309), rabbit anti-Pax6 antibody (BioLegend, Cat #901301), rabbit anti-TBR2 antibody (abcam, Cat #ab23345), rabbit anti-NeuroD2 antibody (abcam Cat #ab109406), mouse anti-SATB2 antibody (Santa Cruz Cat #sc-81376), rabbit anti-CUX1 antibody (Santa Cruz Cat #sc-13024), or rat anti-CTIP2 antibody (abcam, Cat #ab18465)] and guinea-pig polyclonal anti-C11orf46 primary antibodies at 4 °C overnight, followed by incubation with secondary antibodies conjugated to Alexa 488, 568 or 647 (Invitrogen, Cat #A11001(mouse), #A11006(rat), #A11008(rabbit) for Alexa 488, A11075(guinea-pig) for Alexa568, A21235 (mouse), A21244 (rabbit) for Alexa 647, respectively) for 1 h. Nuclei were labeled with DAPI (Roche, 10236276001). For cell type marker staining, anti-CaMKII (Millipore, Cat #05-

532), anti-Olig-2 (Millipore, Cat #AB9610), anti-ALDH1L1 (abcam, ab177463), or anti-Iba-1 (Wako, Cat #019-19741) antibody were used. For analysis of neuronal migration and axonal outgrowth, coronal sections from fixed brains were obtained at 20 and 100 µm, respectively. The sections were washed with PBS containing 0.5% Triton X-100 and then blocked with 0.5% Triton X-100 and 5% normal goat serum for one hour. After blocking, sections were incubated with rat monoclonal anti-GFP (Nacalai, Cat #GF90R) antibodies at 4 °C overnight, followed by incubation with secondary antibodies conjugated to Alexa 488 for 1 h.

**Fluorescence-activated cell sorting (FACS).** The mice subjected to IUE at E15 underwent whole brain extraction at P0 or P14. The somatosensory cortex where GFP-positive neurons were localized was dissected using a stereomicroscope with a fluorescent flashlight (Night Sea, Cat #DFP-1). Cells were separated by utilizing papain dissociation kit (Worthington, Cat #LK003150) with minor modification<sup>71</sup>. For ChIP-qPCR experiments, approximate 0.5–1.0 × 10<sup>6</sup> GFP-positive neurons obtained by pooling 5–6 pups' GFP-positive cortices were collected by FACS into 1.5 ml Protein LoBind Tube (Eppendorf, Cat #Z666505).

**Quantitative real time PCR (qPCR).** FAC-sorted GFP positive somatosensory cortical neurons were suspended into RNAlater solution (ambion Cat#AM7021) and stored at –80 °C. Total RNA were prepared and treated by DNase™ RNA Isolation Kit, Applied Biosystems Cat#KIT0204). Total RNA was reverse transcribed to cDNA (SuperScript™ III First-Strand Synthesis System, Invitrogen Cat#18080051). qPCR was performed using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems Cat#A25742). *Hprt* was used as a reference gene to normalize the expression data by ddCt method. Target gene primers (*Sema6a*, *Dcl1*, *Gap43*, and *Hprt*) are listed in Supplementary Table 6.

**Chromatin immunoprecipitation (ChIP).** FAC-sorted GFP-positive cells were cross-linked with 1% formaldehyde for 15 min at room temperature, quenched with glycine to a final concentration of 0.125 M for another 10 min, and stored at –80 °C overnight before processing. After cell lyses with the lysis buffer (50 mM Tris-HCl, 10 mM EDTA (pH8.0), 1% SDS, and protease inhibitor cocktail (Roche)), chromatin was sonicated with a Q Sonica Sonicator (45% amplitude, 15 s 20 times with 1 min interval), cleared by centrifugation, and incubated overnight at 4 °C with 5–7 µg of the desired anti-H3K9me3 (abcam, Cat #ab8898). Five micrograms of chromatin was used for each immunoprecipitate. Immunocomplexes were immobilized with 30 µl of protein-G agarose beads (Active Motif) for 4 h at 4 °C, followed by stringent washes (wash buffers will be described) and elution. Eluates were reverse cross-linked overnight at 65 °C and deproteinized with proteinase K at 56 °C for 1 h. DNA was extracted with phenol chloroform, followed by ethanol precipitation. ChIP-qPCR analyses were performed in a QuantStudio (Applied Biosystems, 12K Flex) using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems, Cat#A25742). ChIP-qPCR signals were calculated as a percentage of input. Fold induction was calculated over a negative genomic region. All primers used in qPCR analyses are shown in Supplementary Table 6.

**Quantification of axonal phenotypes.** Axonal arborization was assessed by previously published methods with some modifications<sup>6,72</sup>. Tiled scanned images of the somatosensory cortex where GFP-positive transcallosal neurons projected from the IUE-targeted contralateral hemisphere were collected by a confocal microscope (Zeiss, LSM 700). Images were taken by the 10x objective lens (0.8× zoom-out), with 1024 × 1024 pixel resolution at a scan speed of seven per section. Single plain images at corpus callosum were taken by the 10× objective lens. Acquisition parameters were kept constant for all scans. Contralateral arborization signals were analyzed in images of whole cortex which were cropped with 9 inches width, and GFP signal from layer I to layer V was quantified by 20 equal spaces (20 bins) by ImageJ software. GFP signal intensity at corpus callosum was measured by ImageJ. Contralateral and callosal signal intensity were normalized by GFP-positive cell density in ipsilateral (electroporated) site. Three to six mice were used for quantification for each group.

**Quantitative bin analysis of brain slices.** The effect of knockdown of *C11orf46* on neuronal migration was assessed by performing quantitative bin analysis by our previously published methods<sup>19–21</sup>. The number of GFP-positive cells in the developing somatosensory cortex, including the CP, intermediate zone (IZ), and SVZ/VZ, were counted. The somatosensory cortex was divided into 10 equal spaces (10 bins) and the percentage of GFP-positive cells in each bin was determined. The numbers of neurons in each category from more than three independent experiments were counted in a blinded manner using ImageJ software (<http://rsb.info.nih.gov/ij/>).

**Statistical analyses.** For analysis of human MRI data, IBM SPSS Statistics (version 24) software was used. Kolmogorov-Smirnov testing was performed to confirm normal distribution of patient data. Univariate fixed effect ANCOVAs accounting for age and sex as covariates compared white matter structure volumes among patient groups, and post hoc pairwise two-sided Least Significant Difference (LSD) comparisons were performed. Bonferroni correction was applied due to

comparison of multiple corpus callosum regions among patient groups; nominal *p*-values are shown and threshold of  $P < 0.008$  ( $=0.05/6$ ) was considered significant. One-way ANOVA with the Bonferroni multiple comparison test was used for factorial analysis among more than three groups. Student's *t*-test was used for comparisons between two groups. All data are shown as mean  $\pm$  SEM, unless otherwise mentioned.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

The raw RNA-seq data that support the findings of this study have been deposited to GEO under the accession number GSE119360. All data are available from the corresponding authors, A.K. and S.A. upon reasonable request. Human subject data will be deidentified to protect confidentiality.

Received: 28 August 2018 Accepted: 12 August 2019

Published online: 11 September 2019

### References

- Peiker, I. et al. Perceptual integration deficits in autism spectrum disorders are associated with reduced interhemispheric gamma-band coherence. *J. Neurosci.* **35**, 16352–16361 (2015).
- Ribolsi, M., Daskalakis, Z. J., Siracusano, A. & Koch, G. Abnormal asymmetry of brain connectivity in schizophrenia. *Front. Hum. Neurosci.* **8**, 1010 (2014).
- Margari, L. et al. Clinical manifestations in children and adolescents with corpus callosum abnormalities. *J. Neurol.* **263**, 1939–1945 (2016).
- Lindwall, C., Fothergill, T. & Richards, L. J. Commissure formation in the mammalian forebrain. *Curr. Opin. Neurobiol.* **17**, 3–14 (2007).
- Mita, S. et al. Transcallosal projections require glycoprotein M6-dependent neurite growth and guidance. *Cereb. Cortex* **25**, 4111–4125 (2015).
- Wang, C. L. et al. Activity-dependent development of callosal projections in the somatosensory cortex. *J. Neurosci.* **27**, 11334–11342 (2007).
- Sajan, S. A. et al. Both rare and de novo copy number variants are prevalent in agenesis of the corpus callosum but not in cerebellar hypoplasia or polymicrogyria. *PLoS. Genet.* **9**, e1003823 (2013).
- Lin, A. et al. Mapping 22q11.2 gene dosage effects on brain morphometry. *J. Neurosci.* **37**, 6183–6199 (2017).
- Calloni, S. F. et al. Compound Heterozygous Variants in ROBO1 Cause a Neurodevelopmental Disorder with Absence of Transverse Pontine Fibers and Thinning of the Anterior Commissure and Corpus Callosum. *Pediatr. Neurol.* **70**, 70–74 (2017).
- Shieh, C., Moser, F., Graham, J. M. Jr., Watiker, V. & Pierson, T. M. Mutation in the sixth immunoglobulin domain of L1CAM is associated with migrational brain anomalies. *Neurol. Genet.* **1**, e34 (2015).
- Witteveen, J. S. et al. Haploinsufficiency of MeCP2-interacting transcriptional co-repressor SIN3A causes mild intellectual disability by affecting the development of cortical integrity. *Nat. Genet.* **48**, 877–887 (2016).
- Kortum, F. et al. The core FOXP1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. *J. Med. Genet.* **48**, 396–406 (2011).
- Najmabadi, H. et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* **478**, 57–63 (2011).
- Han, J. C. Rare syndromes and common variants of the brain-derived neurotrophic factor gene in human obesity. *Prog. Mol. Biol. Transl. Sci.* **140**, 75–95 (2016).
- Free, S. L. et al. Quantitative MR image analysis in subjects with defects in the PAX6 gene. *Neuroimage* **20**, 2281–2290 (2003).
- Abouzeid, H. et al. PAX6 aniridia and interhemispheric brain anomalies. *Mol. Vis.* **15**, 2074–2083 (2009).
- Han, J. C. et al. Association of brain-derived neurotrophic factor (BDNF) haploinsufficiency with lower adaptive behaviour and reduced cognitive functioning in WAGR/11p13 deletion syndrome. *Cortex* **49**, 2700–2710 (2013).
- Kamiya, A. et al. A schizophrenia-associated mutation of DISC1 perturbs cerebral cortex development. *Nat. Cell Biol.* **7**, 1167–1178 (2005).
- Kamiya, A. et al. Recruitment of PCM1 to the centrosome by the cooperative action of DISC1 and BBS4: a candidate for psychiatric illnesses. *Arch. Gen. Psychiatry* **65**, 996–1006 (2008).
- Saito, A. et al. Developmental alcohol exposure impairs activity-dependent S-nitrosylation of NDEL1 for neuronal maturation. *Cereb. Cortex* **27**, 3918–3929 (2017).
- Saito, A. et al. Early postnatal GABA receptor modulation reverses deficits in neuronal maturation in a conditional neurodevelopmental mouse model of DISC1. *Mol. Psychiatry* **21**, 1449–1459 (2016).
- Taniguchi, Y., Young-Pearse, T., Sawa, A. & Kamiya, A. In utero electroporation as a tool for genetic manipulation in vivo to study psychiatric disorders: from genes to circuits and behaviors. *Neuroscientist* **18**, 169–179 (2012).
- LoTurco, J., Manent, J. B. & Sidiqi, F. New and improved tools for in utero electroporation studies of developing cerebral cortex. *Cereb. Cortex* **19**, i120–i125 (2009).
- Fame, R. M., MacDonald, J. L. & Macklis, J. D. Development, specification, and diversity of callosal projection neurons. *Trends Neurosci.* **34**, 41–50 (2011).
- Paul, P. et al. A Genome-wide multidimensional RNAi screen reveals pathways controlling MHC class II antigen presentation. *Cell* **145**, 268–283 (2011).
- Kirli, K. et al. A deep proteomics perspective on CRM1-mediated nuclear export and nucleocytoplasmic partitioning. *Elife* **4**, e11466 (2015).
- Rolland, T. et al. A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226 (2014).
- Hein, M. Y. et al. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163**, 712–723 (2015).
- Stark, C. et al. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).
- Huttlin, E. L. et al. The BioPlex Network: a systematic exploration of the human interactome. *Cell* **162**, 425–440 (2015).
- Ayyanathan, K. et al. Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev.* **17**, 1855–1869 (2003).
- Ichimura, T. et al. Transcriptional repression and heterochromatin formation by MBD1 and MCAF/AM family proteins. *J. Biol. Chem.* **280**, 13928–13935 (2005).
- O'Geen, H. et al. Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs. *PLoS. Genet.* **3**, e89 (2007).
- Fu, X., Brown, K. J., Rayavarapu, S., Nagaraju, K. & Liu, J. S. The use of proteomic analysis to study trafficking defects in axons. *Methods Cell Biol.* **131**, 151–162 (2016).
- Haklai-Topper, L., Mlechkovich, G., Savariego, D., Gokhman, I. & Yaron, A. Cis interaction between Semaphorin6A and Plexin-A4 modulates the repulsive response to Sema6A. *EMBO J.* **29**, 2635–2645 (2010).
- Tamagnone, L. et al. Plexins are a large family of receptors for transmembrane, secreted, and GPI-anchored semaphorins in vertebrates. *Cell* **99**, 71–80 (1999).
- Shen, Y., Mani, S., Donovan, S. L., Schwob, J. E. & Meiri, K. F. Growth-associated protein-43 is required for commissural axon guidance in the developing vertebrate nervous system. *J. Neurosci.* **22**, 239–247 (2002).
- Powell, S. K., Gregory, J., Akbarian, S. & Brennand, K. J. Application of CRISPR/Cas9 to the study of brain development and neuropsychiatric disease. *Mol. Cell. Neurosci.* **82**, 157–166 (2017).
- Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).
- Stiles, J. & Jernigan, T. L. The basics of brain development. *Neuropsychol. Rev.* **20**, 327–348 (2010).
- Kalil, K. & Dent, E. W. Branch management: mechanisms of axon branching in the developing vertebrate CNS. *Nat. Rev. Neurosci.* **15**, 7–18 (2014).
- Lewis, T. L. Jr., Courchet, J. & Polleux, F. Cell biology in neuroscience: Cellular and molecular mechanisms underlying axon formation, growth, and branching. *J. Cell Biol.* **202**, 837–848 (2013).
- Akbarian, S. & Huang, H. S. Epigenetic regulation in human brain-focus on histone lysine methylation. *Biol. Psychiatry* **65**, 198–203 (2009).
- Ross, S. E. et al. Bhlhb5 and Prdm8 form a repressor complex involved in neuronal circuit assembly. *Neuron* **73**, 292–303 (2012).
- Oshiro, H., Hirabayashi, Y., Furuta, Y., Okabe, S. & Gotoh, Y. Up-regulation of HP1gamma expression during neuronal maturation promotes axonal and dendritic development in mouse embryonic neocortex. *Genes. Cells* **20**, 108–120 (2015).
- Eom, G. H. et al. Histone methyltransferase PRDM8 regulates mouse testis steroidogenesis. *Biochem. Biophys. Res. Commun.* **388**, 131–136 (2009).
- Mishima, Y. et al. Nucleosome compaction facilitates HP1gamma binding to methylated H3K9. *Nucleic Acids Res.* **43**, 10200–10212 (2015).
- Jiang, Y. et al. The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. *Nat. Genet.* **49**, 1239–1250 (2017).
- Alazami, A. M. et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep.* **10**, 148–161 (2015).
- Zhai, Q., Wang, L., Zhao, P. & Li, T. Role of citrullination modification catalyzed by peptidylarginine deiminase 4 in gene transcriptional regulation. *Acta Biochim. Biophys. Sin.* **49**, 567–572 (2017).

51. Sharma, P. et al. Citrullination of histone H3 interferes with HP1-mediated transcriptional repression. *PLoS Genet.* **8**, e1002934 (2012).
52. Dechat, T. et al. Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev.* **22**, 832–853 (2008).
53. Giacomini, C., Mahajani, S., Ruffilli, R., Marotta, R. & Gasparini, L. Lamin B1 protein is required for dendrite development in primary mouse cortical neurons. *Mol. Biol. Cell* **27**, 35–47 (2016).
54. Runker, A. E. et al. Mutation of Semaphorin-6A disrupts limbic and cortical connectivity and models neurodevelopmental psychopathology. *PLoS ONE* **6**, e26488 (2011).
55. Hakansson, K. et al. Semaphorin 6A knockout mice display abnormalities across ethologically-based topographies of exploration and in motor learning. *Neurosci. Lett.* **641**, 70–76 (2017).
56. Mitsogiannis, M. D., Little, G. E. & Mitchell, K. J. Semaphorin-Plexin signaling influences early ventral telencephalic development and thalamocortical axon guidance. *Neural Dev.* **12**, 6 (2017).
57. Baek, S. T. et al. Off-target effect of doublecortin family shRNA on neuronal migration associated with endogenous microRNA dysregulation. *Neuron* **82**, 1255–1262 (2014).
58. Gurwitz, D. Human iPSC-derived neurons and lymphoblastoid cells for personalized medicine research in neuropsychiatric disorders. *Dialog. Clin. Neurosci.* **18**, 267–276 (2016).
59. Bishop, K. M., Rubenstein, J. L. & O’Leary, D. D. Distinct actions of Emx1, Emx2, and Pax6 in regulating the specification of areas in the developing neocortex. *J. Neurosci.* **22**, 7627–7638 (2002).
60. Schuurmans, C. et al. Sequential phases of cortical specification involve Neurogenin-dependent and -independent pathways. *EMBO J.* **23**, 2892–2902 (2004).
61. Hiraoka, K. et al. Regional volume decreases in the brain of Pax6 heterozygous mutant rats: MRI deformation-based morphometry. *PLoS ONE* **11**, e0158153 (2016).
62. Boretius, S. et al. In vivo MRI of altered brain anatomy and fiber connectivity in adult pax6 deficient mice. *Cereb. Cortex* **19**, 2838–2847 (2009).
63. Han, J. C. et al. Brain-derived neurotrophic factor and obesity in the WAGR syndrome. *N. Engl. J. Med.* **359**, 918–927 (2008).
64. Fischl, B. et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
65. Gupta, A. & Devi, L. A. The use of receptor-specific antibodies to study G-protein-coupled receptors. *Mt. Sinai J. Med.* **73**, 673–681 (2006).
66. Gupta, A. et al. Post-activation-mediated changes in opioid receptors detected by N-terminal antibodies. *J. Biol. Chem.* **283**, 10735–10744 (2008).
67. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
68. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
69. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
70. Pruitt, K. D., Tatusova, T., Brown, G. R. & Maglott, D. R. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* **40**, D130–D135 (2012).
71. Saxena, A. et al. Trehalose-enhanced isolation of neuronal sub-types from adult mouse brain. *Biotechniques* **52**, 381–385 (2012).
72. Ageta-Ishihara, N. et al. Control of cortical axon elongation by a GABA-driven Ca<sup>2+</sup>/calmodulin-dependent protein kinase cascade. *J. Neurosci.* **29**, 13720–13729 (2009).

## Acknowledgements

We thank Dr. Kazuhiro Ishii, Ms. Nicole Rangos, Mr. Akarsh Sharma, and Mr. Tomonori Matsushita for technical and intellectual assistance of experiments; Dr. Xiaolei Zhu, Ms. Aisa Moreno-Megui and Ms. Nada Redradja for a critical reading of the manuscript; Lee Blosser and Raffaello Cimbri for sorting cells by FACS; Dr. Elliot Androphy for providing NSC34 cells. This work was supported by DA041208 (A.K.), MH091230 (A.K.), AT008547 (A.K.), MH094268 (A.K.), JHU Catalyst Award (A.K.), MH104341 (S.A.), MH117790 (S.A.), the Brain & Behavior Research Foundation (A.K., S.A., A.S.), ZIAHD008898 (J.C.H.), an NIH Bench-to-Bedside Award (J.C.H.), and S10OD016374 (Dr. Scot C. Kuo for use of confocal microscope LSM 700).

## Author contributions

C.J.P., A.S., S.A., and A.K. designed the study. C.J.P. performed biochemical and cellular experiments. A.S. performed biochemistry and in vivo experiments. T.F., C.R., A.D., A. G., and L.D. performed biochemical experiments. A.S. performed in utero electroporation surgery. Y.H., Y.T., M.N., and E.A. performed the data analysis of axonal phenotypes and qPCR. G.P. maintained and prepared animals for in utero electroporation. J.G.P. conducted gene search using Decipher database. A.D. and S.E. analyzed RNA-sequencing data. J.C.H. obtained human samples and generated lymphoblastoid cell lines. J.C.H., F. M.L., J.A.B. obtained and analyzed structural MRI data. C.J.P., A.S., S.A., and A.K. wrote the manuscript. All authors contributed to and have approved the final manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-12013-y>.

**Competing interests:** J.C.H. is the recipient of an unrestricted research grant from Rhythm Pharmaceuticals. The other authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019