nature neuroscience

TECHNICAL REPORT

https://doi.org/10.1038/s41593-020-0692-9

Check for updates

Viral manipulation of functionally distinct interneurons in mice, non-human primates and humans

Douglas Vormstein-Schneider[1,11], Jessica D. Lin[1,11], Kenneth A. Pelkey[2], Ramesh Chittajallu[2], Baolin Guo[1], Mario A. Arias-Garcia[1], Kathryn Allaway[1,3,4], Sofia Sakopoulos[1], Gates Schneider[1], Olivia Stevenson[1], Josselyn Vergara[1], Jitendra Sharma[5], Qiangge Zhang[5], Tom P. Franken[6], Jared Smith[6], Leena A. Ibrahim[1,3], Kevin J. M astro[1,3], Ehsan Sabri[7], Shuhan Huang[1,3], Emilia Favuzzi[1,3], Timothy Burbridge[1,3], Qing Xu[8], Lihua Guo[8], Ian Vogel[1], Vanessa Sanchez[1], Giuseppe A. Saldi[1,3], Bram L. Gorissen[1], Xiaoqing Yuan[2], Kareem A. Zaghloul[9], Orrin Devinsky[10], Bernardo L. Sabatini[1,3], Renata Batista-Brito[7], John Reynolds[6], Guoping Feng[1,5], Zhanyan Fu[1], Chris J. McBain[2], Gord Fishell[1,3] and Jordane Dimidschstein[1]✉

Recent success in identifying gene-regulatory elements in the context of recombinant adeno-associated virus vectors has enabled cell-type-restricted gene expression. However, within the cerebral cortex these tools are largely limited to broad classes of neurons. To overcome this limitation, we developed a strategy that led to the identification of multiple new enhancers to target functionally distinct neuronal subtypes. By investigating the regulatory landscape of the disease gene *Scn1a*, we discovered enhancers selective for parvalbumin (PV) and vasoactive intestinal peptide-expressing interneurons. Demonstrating the functional utility of these elements, we show that the PV-specific enhancer allowed for the selective targeting and manipulation of these neurons across vertebrate species, including humans. Finally, we demonstrate that our selection method is generalizable and characterizes additional PV-specific enhancers with exquisite specificity within distinct brain regions. Altogether, these viral tools can be used for cell-type-specific circuit manipulation and hold considerable promise for use in therapeutic interventions.

Large-scale transcriptomic studies are rapidly revealing where and when genes associated with neuropsychiatric disease are expressed within specific cell types[1–4]. Approaches for understanding and treating these disorders will require methods for targeting and manipulating specific neuronal subtypes; thus, gaining access to these populations in non-human primates and humans has become paramount. Recombinant adeno-associated viruses (rAAVs) are the method of choice for gene delivery in the nervous system, but they have a limited genomic payload and are not intrinsically selective for particular neuronal populations[5]. We and others have identified short regulatory elements capable of restricting viral expression to broad neuronal classes within the cerebral cortex (i.e., pan-pyramidal[5] and pan-interneuron[6]) or a select subpopulation of somatostatin (SST)-expressing interneurons[7]. In addition, systematic enhancer discovery has been accelerated by the recent development of technologies allowing for transcriptomic and epigenetic studies at single-cell resolution[6–12]. Despite these advances, the search space for enhancer selection remains enormous and, to date, success has been limited. To focus our enhancer selection, we chose to specifically examine the regulatory landscape of *Scn1a*, a gene expressed in distinct neuronal populations and the disruption of which is associated with severe epilepsy[13]. Combining the single-cell assay for transposable-accessible chromatin sequencing (scATAC-seq) data with sequence conservation across species, we nominated ten candidate regulatory sequences in the vicinity of this gene. By thoroughly investigating each of these elements for their ability to direct viral expression, we identified three enhancers that collectively target the breadth of neuronal populations expressing *Scn1a*. Among these, one particular short regulatory sequence was capable of restricting viral expression to PV-expressing cortical interneurons (PV cINs). To fully assess the utility of this element beyond reporter expression, we validated it in a variety of contexts, including synaptic tagging and calcium imaging, as well as opto- and chemogenetic approaches, both ex vivo and in vivo. Moreover, we show that this element allows for the selective targeting of PV cINs both during development and across

[1]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA, USA. [2]National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA. [3]Department of Neurobiology, Harvard Medical School, Boston, MA, USA. [4]NYU Langone Medical Center, New York University, New York, NY, USA. [5]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Systems Neurobiology Laboratories, The Salk Institute for Biological Studies, La Jolla, CA, USA. [7]Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY, USA. [8]Center for Genomics & Systems Biology, New York University, Abu Dhabi, UAE. [9]National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA. [10]Comprehensive Epilepsy Center, New York University School of Medicine, New York, NY, USA. [11]These authors contributed equally: Douglas Vormstein-Schneider, Jessica D. Lin. ✉e-mail: jordane@broadinstitute.org

species, including rodents, non-human primates and humans. Demonstrating that this approach provides a generalizable strategy for enhancer discovery, we further selected 25 regulatory elements in the vicinity of 7 additional genes with expression enriched in PV cINs. From these we identified an additional four PV-specific regulatory elements, each of which had remarkably selective expression for this population within specific brain regions. Together our findings demonstrate the generalizability of our enhancer selection method, as well as the utility of a variety of functionally tested tools that can be immediately utilized across animal models. These viral reagents can be employed to interrogate how functionally distinct neuronal cell types are affected in the context of neurological, neurodevelopmental and neurodegenerative diseases in non-human primates. Ultimately, these may provide the means to therapeutically normalize pathological neuronal activity or gene expression in specific neuronal populations.

Results

Identification of *Scn1a* enhancers. *Scn1a* encodes for Nav1.1, a sodium channel expressed in three nonoverlapping neuronal populations: fast-spiking cortical interneurons expressing PV cINs, disinhibitory cINs expressing the vasointestinal peptide (VIP cINs) and layer 5 pyramidal neurons[14–16]. Haploinsufficiency or pathogenic variants of *Scn1a* cause Dravet's syndrome, a common and intractable form of epileptic encephalopathy characterized by the early onset of seizures[17–19]. To devise a genetic strategy to target the distinct cortical populations expressing this gene, we developed an integrative method to systematically identify candidate enhancers within this locus. Regulatory sequences were selected based on the following three criteria: first, it has been posited that the proximity of the enhancer to the transcriptional start site (TSS) of a gene scales directly to the level of expression[20]. Therefore, to identify enhancers capable of driving functional levels of transgenes, we examined the intergenic and intronic regions of *Scn1a* closest to its TSS. Second, the location of active enhancers within a given cell type correlates with chromatin accessibility[21,22]. To assess the chromatin landscape of the cellular populations expressing *Scn1a*, we isolated and collected green fluorescent protein-positive (GFP$^+$) nuclei from interneurons in the primary visual cortex of Dlx6a-cre::Sun1-eGFP transgenic mice and performed scATAC-seq profiling[23,24] using the 10x Chromium platform. By applying stringent quality control criteria (Methods), we recovered 3,374 nuclei with on average 107,000 reads and 74,100 fragments detected per cell (62% of the total number of nuclei loaded on the 10x chip). The chromatin accessibility profile of these nuclei was then analyzed using the single nucleus analysis pipeline (Snap)ATAC analysis pipeline (Methods), allowing for the classification of these cells into the four major classes of cINs (Fig. 1a–c and Extended Data Fig. 1a). Third, as regulatory elements are subject to positive selection pressure, we identified sequences showing the highest conservation across mammalian species, including humans[25–29]. Thus, to nominate enhancers with therapeutic potential, we focused on ten selectively accessible intronic and intergenic regions near the TSS of *Scn1a* that are highly conserved across evolution (enhancer sequences E1–E10; Fig. 1d and Supplementary Table 1).

To examine the ability of the candidate enhancers to target the neuronal populations that express *Scn1a*, each sequence was inserted into an rAAV backbone containing a minimal promoter upstream of a red fluorescent reporter (rAAV-E[x]-dTomato). From these constructs, rAAVs were then produced with the PHPeB capsid[30] and systemically injected into adult mice. After 3 weeks, all viruses showed strong expression within the cortex, as well as across multiple brain regions. Except for E5, the vast majority of virally labeled cells expressed the pan-interneuron marker Gad1. However, the degree of co-localization for PV within cortical neurons varied, ranging from >90% for E2 to <5% for E6, with all remaining

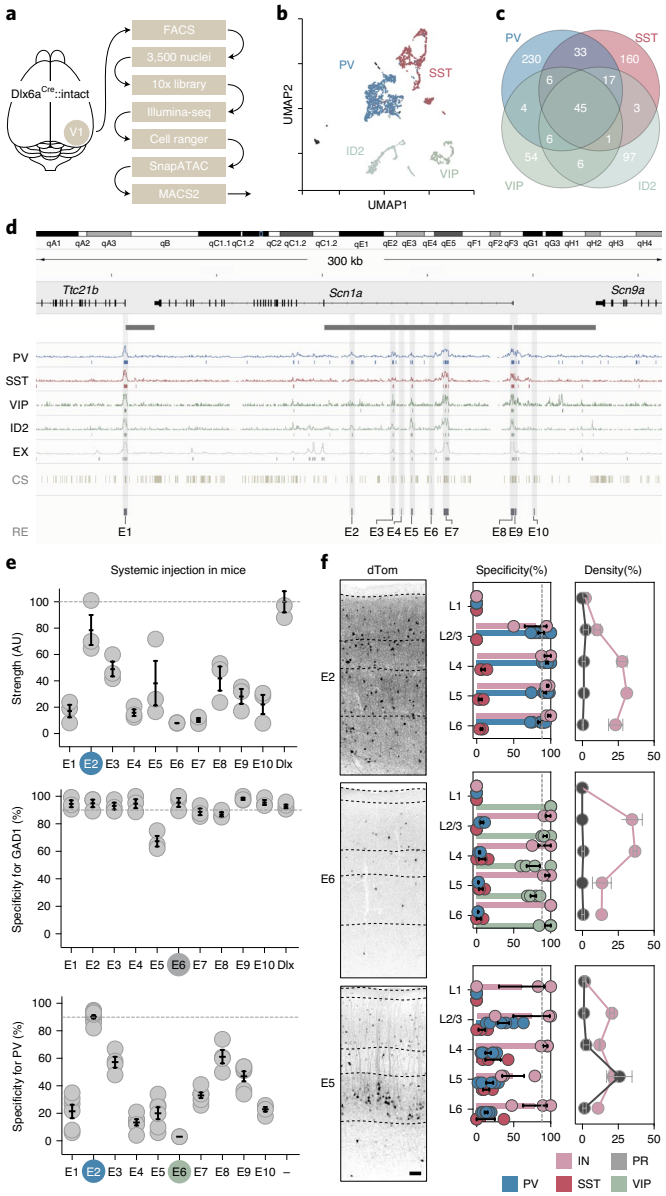


Fig. 1 | Identification of *Scn1a* enhancers. a, Schematic representation of the scATAC-seq pipeline. Interneurons were collected from the visual cortex of adult Dlx6a-cre::Sun1-eGFP mice. **b**, Plot of the 3,500 nuclei in UMAP space. The clusters obtained from the SnapATAC pipeline were lumped into the four cardinal classes of interneuron. **c**, Venn diagram showing the number of unique and shared peaks (in thousands) across the four interneuron populations. **d**, Schematic representation of the enhancer selection method at the *Scn1a* locus (see Methods for a complete description). **e**,**f**, Adult mice were injected systemically with the indicated rAAV-E[x]-dTomato and analyzed 3 weeks post-injection. IHC for the reporter and indicated markers in the S1 cortex was used to assess the strength of expression of the reporter (**e**, top) and the specificity of expression of the viral reporter for the indicated markers (all other panels). Representative fluorescent images of the indicated viral reporter are shown in the S1 cortex (**f**, left). Dashed lines represent the limits of anatomical structures. Scale bars, 50 µm. AU, arbitrary units. On the graphs, the dots represent individual measurements, and the lines represent the average ± s.e.m. Values for specificity, sensitivity and strength are listed in Supplementary Table 2.

enhancers displaying intermediate levels of PV specificity (that is, percentage of cells expressing the viral reporter that co-express PV; Fig. 1e). Next, we further examined the identity and layer

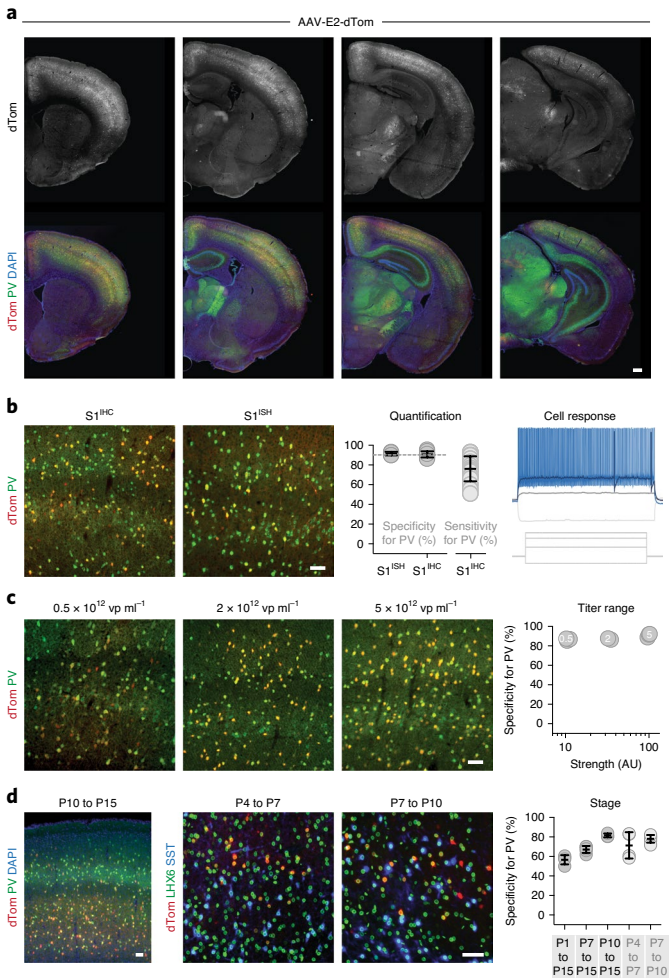


Fig. 2 | Viral targeting of PV cINs in mice. a–c, Adult mice were injected systemically (**a** and **b**) or locally (**c**) with rAAV-E2-dTomato and analyzed 3 weeks post-injection (systemic) or 1 week post-injection (local) by IHC or ISH for both the reporter and the PV. Slice recording of the intrinsic properties of virally labeled neurons is shown (**b**, right). See Methods for details on the reproducibility of the representative images presented in **a**. AU, arbitrary units; vp, viral particles. **d**, Mice were injected locally with rAAV-E2-dTomato and analyzed at the indicated developmental stages for the reporter and the indicated markers. Scale bars, 250 µm (**a**) and 50 µm (**b**–**d**). On the graphs, the dots represent individual measurements, and the lines represent average ± s.e.m. Values for specificity, sensitivity and strength are listed in Supplementary Table 2.

distribution of the neuronal populations captured by E2, E5 and E6. Consistent with their layer distribution, co-localization analysis with various markers revealed that the vast majority of cells expressing the viral reporter under the control of the E2 regulatory element co-localized with PV. We also found that E6 was highly selective for VIP interneurons. By contrast, the E5 regulatory element, although sparsely labeling interneurons across all layers, had a notable enrichment for pyramidal neurons in layer 5 (Fig. 1e,f and Extended Data Fig. 1b).

These data indicate that a large fraction of the neuronal populations expressing *Scn1a* in the cortex is mirrored by the collective expression of these three enhancers. Notably, these regulatory elements account for largely nonoverlapping expression in populations of neurons with distinct functions and developmental origins. The viral tools developed in the present report thus provide a means for dissecting neuronal subtypes and can be used to study their normal function, as well as abnormalities in diseased cortex.

Viral targeting of PV cINs in mice. The E2 regulatory element, with approximately 90% specificity for PV cINs, allows for the targeting of fast-spiking neurons that constitute 40% of all cINs. These neurons exert a strong level of inhibitory control over local networks and their dysfunction has been directly implicated in neurological and neuropsychiatric disorders, including Dravet's syndrome, focal epilepsy, autism spectrum disorder and schizophrenia[31–34]. As such, gaining control over their activity is of particular interest for both fundamental research and clinical applications. Thus, we focused our efforts on characterizing the E2 regulatory element to develop a viral tool with broad utility. Adult mice systemically injected with rAAV-E2-dTomato showed detectable expression of the viral reporter after 1 week and reached a high and stable level of expression after 3 weeks. Immunohistochemistry (IHC) and in situ hybridization (ISH) consistently showed that ~90% of virally labeled cells were PV INs in the cortex. Conversely, on average 75% of PV cINs expressed the viral reporter, reaching a maximum of 93% (Fig. 2a,b). To investigate whether the rAAV-E2-dTomato can target all PV subtypes, we further examined the intrinsic firing properties and the morphological features of virally labeled cells in both the somatosensory (S1) and the prefrontal cortex (PFC) (Extended Data Fig. 2a–c). In these regions, chandelier cells are preferentially found in the upper layers and can be distinguished from basket cells, based on their delay to first spike at rheobase[35]. In the superficial layers of the PFC, cells expressing the viral reporter showed either long or short delay to the first spike, consistent with the presence of both basket and chandelier cells (Extended Data Fig. 2c). Furthermore, we observed cells in the S1 cortex with the stereotypical features of chandelier cells. These cells had their soma located at the boundary of layers 1 and 2/3, and dendrites branching into layer 1 (Extended Data Fig. 2d). Together, these data suggest that the E2 is capable of targeting both basket and chandelier cells.

Although the viral reporter was predominantly confined to the cortex, some positive cells were observed in other brain regions closely corresponding to areas of *Scn1a* expression. E2 maintained high specificity for PV-expressing neurons within the primary visual (V1) cortex and cingulate cortex, subiculum, hippocampal CA1 and substantia nigra pars reticulata (Extended Data Fig. 2e). Importantly, virtually no viral reporter expression was observed outside the brain, with the exception of a few cells observed in the liver (which is expected on systemic delivery of any AAV) and in the lungs (where *Scn1a* is expressed at a low level; Extended Data Fig. 2f). This shows that, despite systemic delivery, the E2 vector can be used to selectively target PV-expressing neurons in various brain regions, with insignificant off-target expression outside the central nervous system.

Many experimental paradigms and clinical applications will require local rather than systemic injection. To be useful in these contexts, viral expression must retain a high level of specificity for PV cINs. Stereotactically guided injections typically lead to a higher number of viral particles per cell compared with systemic delivery, which may result in off-target expression. To test whether increasing the viral load altered the specificity, we locally injected the same volume of rAAV-E2-dTomato at various titers in the cortex of adult mice, and assessed the reporter expression within PV cINs after a week (Fig. 2c). The results show that, although higher titers had increased levels of reporter expression, no alteration of specificity was observed.

Targeting PV cINs at early postnatal stages has been hampered by the relatively late PV expression (around 15 d after birth—P15) and the lack of other early markers for this population. Their involvement in developmental disorders highlights the need to target and manipulate this population during cortical circuit

assembly. Although complex genetic strategies offer a partial solution to achieving this in mice (that is, Lhx6-Cre, Sst-Flp, and Cre- and Flp-dependent reporter), even this approach does not offer the means to easily manipulate these neurons before the second postnatal week. To test whether the E2 enhancer can target fast-spiking cINs before the onset of PV expression, we examined its activity at various postnatal stages. To this end, we tiled our analysis across the early postnatal period, through a series of stereotactically guided injections of rAAV-E2-dTomato (Fig. 2d). First, we assessed the selectivity of the reporter on the onset of PV expression at P15. This revealed that we could obtain >50% selectivity for PV cINs on injection at P1, increasing to 67% by P7 injection and >80% after P10 injection. We next wondered if we could use this approach to label PV cINs before P15. To identify fast-spiking cINs in this context, we relied on Lhx6-Cre/intact transgenic mice, in which GFP is expressed in MGE-derived interneurons (both PV and SST cINs). By co-staining for SST, the PV cINs can be distinguished as GFP^+/SST^-. We obtained 72% and 78% specificity for PV cINs with a P4–P7 or a P7–P10 time course, respectively. Our approach provides a means of studying these neurons during circuit maturation using a single viral injection.

Viral monitoring and manipulation of PV cINs in mice. Having demonstrated the fidelity of E2-directed expression for PV cINs with differing modes of injection and across developmental stages, we then explored the utility of this vector for studying connectivity (using a presynaptic reporter) and activity (using calcium imaging). When E2 was used to drive a synaptophysin (SYP)–tdTomato fusion gene[36], reporter expression was restricted presynaptically to PV cINs, with terminals perisomatically located on to pyramidal neurons (Fig. 3a). When this vector was used to drive GCaMP6f expression[37], we demonstrated that PV cINs were recruited on whisker stimulation (Fig. 3b and Extended Data Fig. 3a). Together, these results demonstrate that E2 provides an effective means of monitoring various aspects of PV cIN biology. We next examined whether E2 was sufficient to elicit functional changes in activity using chemo- or optogenetic approaches. E2 was used in adult animals to direct the expression of the chemogenetic receptor PSAM4-5HT3-LC[38]. PV cINs in brain sections collected from these animals, when exposed to the actuator varenicline, could be induced to fire when the current was clamped below threshold (Fig. 3c). Similar results were obtained using the chemogenetic receptor Gq-DREADD[39] (Extended Data Fig. 3b). Finally, both constant and high-frequency laser stimulation of PV cINs, expressing the red-shifted opsin C1V1 in brain slices[40], resulted in firing time-locked to the stimulus. Demonstrating that engagement of these neurons resulted in concomitant local inhibition, pyramidal neuron activity in the vicinity of virally labeled PV cINs was consistently interrupted by laser stimulation. Notably, this effect was abolished by treatment with picrotoxin (Fig. 3d and Extended Data Fig. 3c).

Having shown the efficacy of the method ex vivo, we next sought to explore our ability to alter excitatory networks in vivo by optogenetically stimulating PV cINs; 3 weeks after local injection of rAAV-E2-C1V1 into the V1 cortex of adult animals, single unit recordings within the infected region were performed both at baseline and on laser stimulation. The identity of recorded neurons was distinguished based on their spike width and maximal firing frequency. Reliably, inhibitory interneuron firing rates were increased by laser stimulation, whereas excitatory neuronal firing was silenced (Fig. 3e). Together, these data show that the E2 enhancer is able to functionally engage PV cINs and elicit network inhibition using chemo- or optogenetic approaches both ex vivo and in vivo.

Viral targeting and manipulation of PV cINs in primates. The high degree of sequence conservation of the E2 enhancer across mammalian species is suggestive of a conserved role in gene

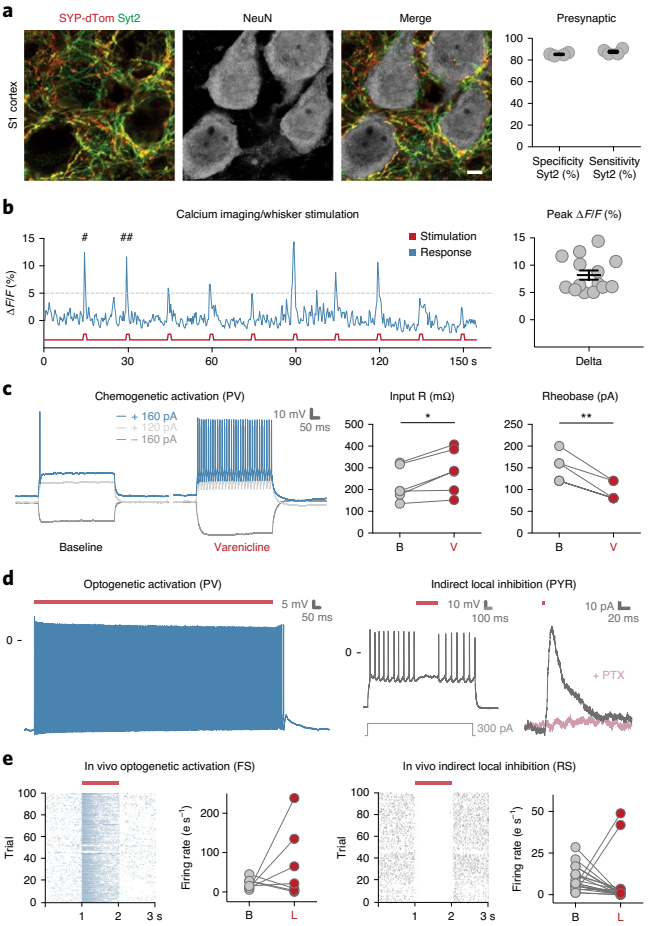


Fig. 3 | Viral monitoring and manipulation of PV cINs in mice. a–e, Mice were injected either locally in the S1 cortex (P10 local injection with rAAV-E2-SYP-dTomato (**a**); P14 local injection with rAAV-E2-GCaMP6f (**b**); and adult local injection with rAAV-E2-C1V1-eYFP (**d** and **e**)) or systemically (adult systemic injection with rAAV-E2-PSAM4-5HT3-LC-GFP (**c**)). **a**, Representative images of the co-localization between the SYP-dTomato reporter and the synaptic marker Syt2, 1 week post-injection, and corresponding quantification (see Methods for details on the reproducibility of the representative images presented). **b**, Ca^{2+} imaging on whisker stimulation was performed 2–3 weeks post-injection. The hash signs (#) indicate two peaks for which the widefield images are shown in Extended Data Fig. 3a. Right: the success rate was calculated as the proportions of $\Delta F/F$ peaks above threshold in response to whisker stimulation. **c**, Current-clamp recording was performed on brain sections 4 weeks after the injection. The traces show a representative cellular response at the indicated currents, both at baseline and after bath application of varenicline (left: $P = 0.0616$; $t = 3.564$; d.f. = 5; $^*P < 0.01$, $^{**}P < 0.001$, $^{***}P < 0.0001$; right: $P = 0.0015$; $t = 6.325$; d.f. = 5; $^*P < 0.01$, $^{**}P < 0.001$, $^{***}P < 0.0001$). **d**, Current-clamp recording was performed on brain sections 2 weeks after injection. Cells expressing the viral reporter were exposed to 2 s of constant laser stimulation (550 nm) while the voltage was recorded over 3 s. Neighboring pyramidal cells that did not express the viral reporter were also recorded during laser stimulation. PYR, pyramidal. **e**, In vivo single-unit analysis of neuronal activity. Raster plots of virally infected neurons on laser stimulation and corresponding population quantification data. Left: fast-spiking (FS) cells. Right: regular spiking excitatory (RS) cells. Notably, due to the mosaic nature of local viral injection, individual cell responses were bimodal. This presumably reflects whether or not particular cells were infected. Scale bars, 5 µm. The red bars represent laser stimulation. On the graphs, dots represent individual measurements and the lines represent average ± s.e.m. Values for specificity and sensitivity are listed in Supplementary Table 2.

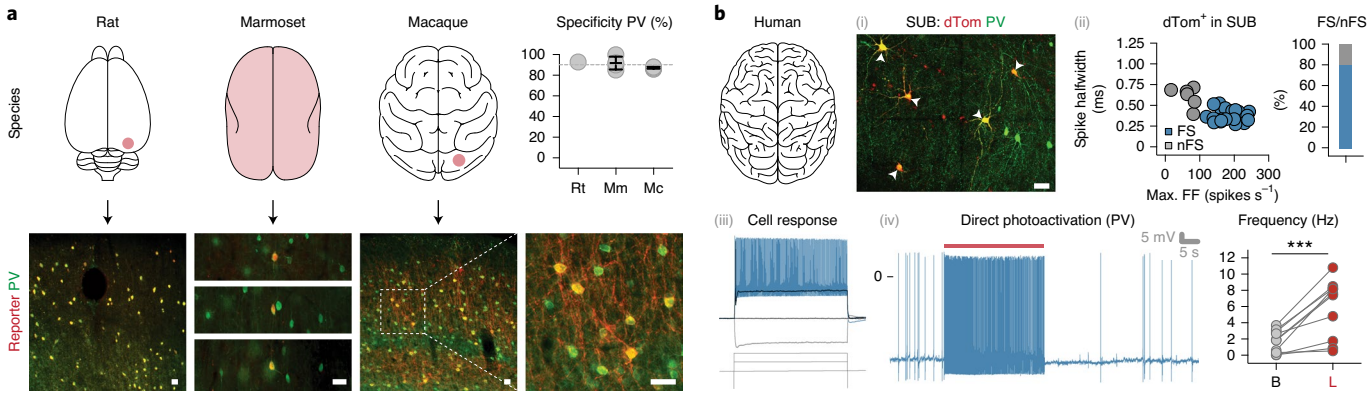


Fig. 4 | Viral targeting and manipulation of PV cINs in primates including humans. a, Animals from indicated species were locally (rat (Rt) and macaque (Mc)) or systemically (marmoset (Mm)) injected with rAAV-E2-C1V1-eYFP (macaque) or rAAV-E2-dTomato (rat and marmoset) and analyzed 2–8 weeks post-injection (see Methods for details on the reproducibility of the representative images presented). **b**, Human brain tissue obtained from surgical resection was exposed to either rAAV-E2-dTomato (i–iii) or rAAV-E2-C1V1-eYFP (see Methods for details on the reproducibility of the representative images presented in **b**(i)) and maintained in culture for 7–14 d (iv) ($P = 0.008$; $t = 4.899$; d.f. = 9; $^{*}P < 0.01$, $^{**}P < 0.001$, $^{***}P < 0.0001$). ii, The proportion of fast-spiking neurons among the virally labeled cells assessed by electrophysiological recordings of intrinsic properties. FF, firing frequency; FS, fast-spiker; nFS, non-fast-spiker; SUB, subiculum. iii, Representative trace from a neuron recorded in ii. iv, Electrophysiology current-clamp recording of virally labeled cells on laser stimulation. B, baseline; L, laser. Scale bars, 50 µm. The red bars represent laser stimulation, and the arrowheads point at neurons co-expressing PV and the viral reporter. On the graphs, dots represent individual measurements and the lines represent the average ± s.e.m. Values for specificity are listed in Supplementary Table 2.

regulation. We thus sought to establish whether this element could be used to target PV cINs across mammalian species. We administered the rAAV-E2 driving a fluorescent reporter systemically in marmosets (intravenous injection using capsid 9) or focally in rats and macaques (using the PHPeB capsid), and showed that we were able to target PV cINs with approximately 90% specificity across all three species (Fig. 4a).

Human brain tissue obtained during surgical resection can be cultured for prolonged periods[41]. Taking advantage of this exceptional ability of human brain tissue to remain healthy ex vivo, we exposed freshly resected subiculum or medial temporal cortex to the rAAV-E2 virus. Over the 2-week culture period, we observed the progressive appearance of fluorescently labeled cells. In regions where PV staining reflected the expected distribution of these cells, virally labeled cells were PV^{+} (Fig. 4b(i); see Methods for details). In addition, most cells within both the cortex and the subiculum showed the characteristic hallmarks of PV INs as indicated by multiple criteria. These include morphology, as well as maximum firing rate when evoked through direct depolarization or optogenetic light stimulation (Fig. 4b(ii–iv) and Extended Data Fig. 4a).

Importantly, the human version of the E2 enhancer showed a similar degree of specificity for PV cINs on injection in mice (Extended Data Fig. 4b), further demonstrating that noncoding regions of the genome characterized by a high degree of sequence conservation are likely to retain their functional properties across species. Finally, truncation of both the 5′- and the 3′-ends of this enhancer resulted in a drastic reduction of specificity, suggesting that the functional boundaries of the E2 enhancer have been optimally identified (Extended Data Fig. 4b). Altogether, these results indicate that E2 provides an effective tool for targeting and manipulating PV cINs across mammals, including humans.

Identification of viral enhancers with regional specificity. To demonstrate that our enhancer selection method is generalizable, we identified a further 25 candidates in the vicinity of 7 genes the expression of which was enriched in PV cINs across species (Methods). Systemic injection of AAVs containing these sequences revealed that four of them displayed >90% selectivity for PV cINs. Notably, among these enhancers, the relatively few virally labeled neurons that did not express PV were positive for the pan-interneuron marker Gad1 (Fig. 5a). Each of these four enhancers was specific for distinct but overlapping subsets of the PV-expressing neurons. Specifically, whereas E11 and E14 showed the strongest reporter expression in the upper cortical layers, E22 restricted expression almost exclusively to the cortex, with only a few cells showing low levels of expression elsewhere. By contrast, E29 directed the most global expression, targeting the entire population of PV-expressing neurons throughout the central nervous system. All enhancers selected for our screen have a high degree of sequence conservation, suggestive of conserved regulatory function across species. To directly test whether this correspondingly results in similar expression across species, we performed local injection of rAAV–E22–dTomato and rAAV–E29–dTomato viruses in the PFC and S1 cortex of macaques, respectively. This showed that, similar to mice, in both cases the expression of the reporter was restricted to PV cINs (Fig. 5b). The combination of regional selectivity and conservation of expression across species demonstrates the scalability of the enhancer selection method. This opens the possibility of using these tools for work in non-human primates and paves the way for the development of targeted therapies to correct abnormal brain function.

Sensitivity of viral enhancers for the PV cIN population. To estimate the percentage of the total targeted population that can be captured by a given enhancer, one needs to know accurately the percentage of the target population infected. As a result, the proportion of a given cellular population infected by a given rAAV depends on the mode of administration and the absolute number of viral particles injected. Estimating this is not trivial. In fact, the only practical solution to achieving this is by performing viral injections that reach saturation (that is, all cells are infected by at least one viral particle). This is currently impossible in primates. However, we attempted to estimate this for E2, E22 and E29 using local injections in macaques. As the diffusion of AAVs in parenchymal tissue on local injection is limited, the degree of capture of the target population varied with the distance to the center of the injection

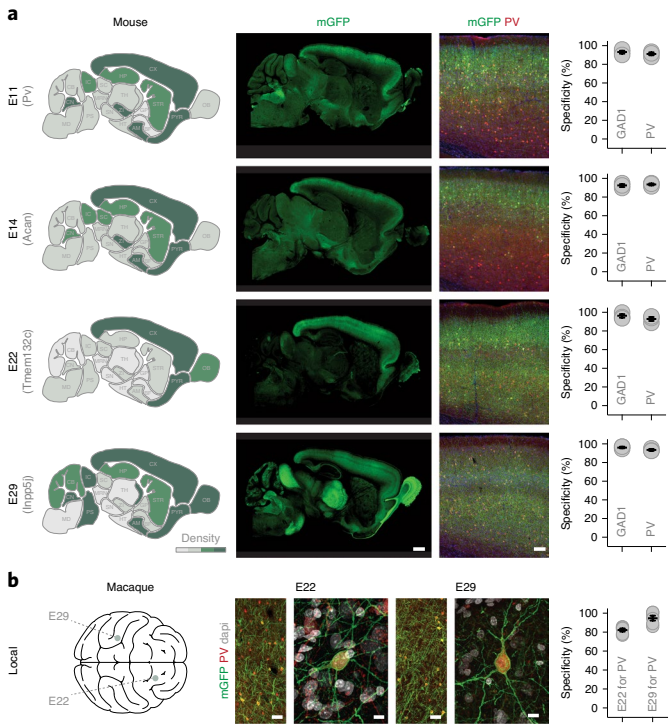


Fig. 5 | Identification of viral enhancers with regional specificity. a, Adult mice were injected systemically with the indicated rAAV-E[x]-mGFP and analyzed 3 weeks post-injection. IHC for the reporter and indicated markers in the S1 cortex were used to assess the density of neuronal cell bodies expressing the viral reporter (left) and the specificity of expression of the viral reporter for the indicated markers (right). Note that, for the E29 virus, there are no cell bodies in the thalamus with the exception of the thalamic reticular nucleus. CX, cortex; HP, hippocampus; STR, striatum; PYR, pyriform cortex; OB, olfactory bulbs; TH, thalamus; GP, globus pallidus; BF, basal forebrain; AM, amygdala; HT, hypothalamus; ZI, zona inserta; SN, substantia nigra; MRN, median raphe nucleus; SC, superior colliculus; IC, inferior colliculus; CB, cerebellum; CN, cerebellar nuclei; PS, pons; MD, medulla. **b**, One adult macaque was injected in the PFC with rAAV-E22-eGFP and another in the S1 cortex with AAV-E29-eGFP. Both animals were analyzed 8 weeks post-injection with IHC for the reporter and indicated markers. Scale bars, 100 µm (**a**), 50 µm (**b** left) and 10 µm (**b** right). On the graphs, dots represent individual measurements and the lines represent average ± s.e.m. Values for specificity are listed in Supplementary Table 2.

site. We thus quantified the percentage of PV cINs expressing the viral reporter using the coronal brain sections showing the highest density of reporter expression. By selecting subregions of the images where the density of expression of the viral reporter was the highest (that is, in the region nearest to the center of the injection), we could show that the sensitivity for PV (that is, the percentage of PV cINs that express the viral reporter) appears highest for both E2 and E29 but lower for E22 (capturing 87%, 85% and 40% of PV cINs, respectively; Extended Data Fig. 5). These numbers indicate that E2 and E29 are capable of targeting most PV cINs, whereas E22 might be restricted to a subset of the population in non-human primates. Importantly, these measures are almost certainly an underestimate of the sensitivity, because we strongly suspect that complete infectivity was never achieved.

Enhancer sequence composition. Although the degree of conservation between the mouse and the human enhancer sequences is high, we wondered whether the continuously conserved subregions of these sequences might be enriched with core transcription factor-binding sites (TFBSs). To assess this, we analyzed the TFBS enrichment profile for each cell-type-specific enhancer using the CiiiDER analysis pipeline with the Jaspar vertebrate database[42,43] (Methods). The analysis revealed that most TFBSs enriched in these sequences are found within subregions of the enhancer that are conserved between mouse and human (Extended Data Fig. 6). When conducted on the human homolog enhancer sequences, the outcome of this analysis was identical (data not shown). These results corroborate the initial assumption that the functional elements responsible for the specificity of the selected enhancers are nested within highly conserved subregions of the enhancer. Notably, a fraction of the mouse motifs enriched on the mouse enhancer sequence can be recognized by the equivalent human transcription factors (TFs). We speculate that these shared motifs constitute the core sequences that allow these sequences to maintain their functional properties across species.

Discussion

A challenge in understanding neurological disorders stems from the complexity of the neuronal types involved. In the present report, in an effort to deconvolve the cellular actions of a particular disease gene, we systematically dissected the *Scn1a* genomic locus. To this end, we identified ten different enhancers distributed across the intronic and intergenic regions of this gene. By creating AAVs with an expression dependent on each of these enhancers, we found three that recapitulate the global pattern of *Scn1a* gene expression. As loss of expression of *Scn1a* is especially associated with PV cIN dysfunction[15–19], we focused on E2, which was effective for selectively targeting these neurons, not only in rodents but also within various primates, including humans. This enhancer also proved useful for investigating various aspects of PV cIN function, including connectivity and excitability, and for manipulating their activity with chemo- and optogenetic effectors. The demonstration of the utility of E2 in a range of species highlights its usefulness for basic and clinical applications. Notably, this approach provides the means to evaluate and compare the circuit contributions and function of PV cINs across species, regardless of their genetic accessibility[44,45].

An obstacle for enhancer discovery has been the broad landscape over which regulatory elements can act. Indeed, it has been shown that enhancers can regulate gene expression up to 1 MB away from the TSS. As a consequence, the more distal the enhancer the harder it is to predict what gene it regulates. As the purpose of the present study was to identify cell-type-specific enhancers rather than to identify all the regulatory elements involved in *Scn1a* gene regulation, we limited the search to the intronic and immediate intergenic regions of this locus. This approach led to the identification of ten enhancers with various degrees of specificity, including three that collectively recapitulate the expression of *Scn1a* in the cortex. Our findings suggest that *Scn1a* expression is at least in part controlled by these elements, but does not exclude the possibility that additional sequences located further away from the TSS of *Scn1a* may also contribute. With the advent of CRISPR-based screening[20], this limitation will be addressed and will probably uncover additional regulatory elements that are more distal to the TSS of *Scn1a*.

We began the present study by examining enhancers at a specific disease locus, the *Scn1a* gene. By identifying key regulatory elements for each of the cell types that express this gene, we clarified its regulatory landscape. Many SNPs associated with the *SCN1A* locus map to the first intron[46–48]. The three enhancers we characterized were located within this region, perhaps indicating that these SNPs represent mutations affecting the expression of *Scn1a*. Corroborating this hypothesis, GTEx data show multiple expression quantitative trait loci within these enhancers that are associated with alterations in *SCN1A* expression in humans[49]. As *Scn1a* expression is largely restricted to the PV cINs[17,18], it is tempting to speculate that mutations of the E2 enhancer may be a direct cause of Dravet's syndrome.

A major impediment in examining early dynamics of circuit maturation in the context of normal or pathological development (including Dravet's syndrome) is the inaccessibility to specific cell types without the use of transgenic animals. Young PV cINs have been particularly problematic to target even with complex genetic strategies. Given their abundance (they represent 40% of all inhibitory cINs) and relatively late appearance, gaining access to them before the onset of PV expression has long been awaited by the field. The specificity of E2 at these developmental stages, and the simplicity of the viral injection over the use of transgenic animals, provide the means to study both their normal development and their role in disease.

More broadly, the enhancers identified in the present study provide access to neuronal populations with particular clinical relevance. Most obviously these enhancers can potentially be leveraged to alleviate debilitating aspects of Dravet's syndrome, through either gene therapy or modulation of neuronal activity[50]. We demonstrate that local and systemic injections can be used for effective viral delivery to the brain. With local injections, one could possibly ameliorate focal epilepsy, PFC dysfunction or hippocampal memory disorders. By contrast, systemic introduction of virus could be used in contexts where global interventions are necessary, for instance, to correct generalized seizures, or for psychiatric and neurodegenerative disorders. Our study shows that the rigorous identification of regulatory elements provides a roadmap for accessing specific cell types for therapeutic contexts.

Importantly, we demonstrated that our enhancer selection method is generalizable to other genes. Overall, we found a set of seven enhancers with unique specificity for both distinct neuronal populations and regions of the central nervous system. Even with the stringent criteria that we applied (>90% selectivity for the target population), our method has a remarkable (>20%) success rate. Moreover, as predicted by the high degree of sequence conservation, the subset of enhancers that we tested proved equally selective and effective across species including humans. Taken together, our findings support these methods providing a reliable means to systematically identify cell-type-specific enhancers that work across species.

Online content

Any methods, additional references, Nature Research reporting summaries, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-020-0692-9.

Received: 20 December 2019; Accepted: 10 July 2020;
Published online: 17 August 2020

References

1. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. *Nat. Genet.* **50**, 825–833 (2018).
2. Voineagu, I. et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).
3. Parikshak, N. N. et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013).
4. Camp, J. G., Platt, R. & Treutlein, B. Mapping human cell phenotypes to genotypes with single-cell genomics. *Science* **365**, 1401–1405 (2019).
5. Bedbrook, C. N., Deverman, B. E. & Gradinaru, V. Viral strategies for targeting the central and peripheral nervous systems. *Annu. Rev. Neurosci.* **41**, 323–348 (2018).
6. Dimidschstein, J. et al. A viral strategy for targeting and manipulating interneurons across vertebrate species. *Nat. Neurosci.* **12**, 1743–1749 (2016).
7. Hrvatin, S. et al. A scalable platform for the development of cell-type-specific viral drivers. *eLife* **8**, e48089 (2019).
8. Deverman, B. E., Ravina, B. M., Bankiewicz, K. S., Paul, S. M. & Sah, D. W. Y. Gene therapy for neurological disorders: progress and prospects. *Nat. Rev. Drug Discov.* **9**, 641–659 (2018).
9. de Leeuw, C. N. et al. rAAV-compatible mini-romoters for restricted expression in the brain and eye. *Mol. Brain* **9**, 52 (2016).
10. Jüttner, J. et al. Targeting neuronal and glial cell types with synthetic promoter AAVs in mice, non-human primates and humans. *Nat. Neurosci.* **22**, 1345–1356 (2019).
11. Blankvoort, S., Witter, M. P., Noonan, J., Cotney, J. & Kentros, C. Marked diversity of unique cortical enhancers enables neuron-specific tools by enhancer-driven gene expression. *Curr. Biol.* **13**, 2103–2114 (2018).
12. Mehta, P. et al. Functional access to neuron subclasses in rodent and primate forebrain. *Cell Rep.* **26**, 2818–2832 (2019).
13. Griffin, A. et al. Preclinical animal models for dravet syndrome: seizure phenotypes, comorbidities and drug screening. *Front. Pharmacol.* **9**, 573 (2018).
14. Ogiwara, I. et al. Nav1.1 localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an *Scn1a* gene mutation. *J. Neurosci.* **27**, 5903–5914 (2007).
15. Favero, M., Sotuyo, N. P., Lopez, E., Kearney, J. A. & Goldberg, E. M. A transient developmental window of fast-spiking interneuron dysfunction in a mouse model of Dravet syndrome. *J. Neurosci.* **38**, 7912–7927 (2018).
16. Goff, K. M. & Goldberg, E. M. Vasoactive intestinal peptide-expressing interneurons are impaired in a mouse model of Dravet syndrome. *eLife* **8**, e46846 (2019).
17. Cheah, C. S. et al. Specific deletion of Nav1.1 sodium channels in inhibitory interneurons causes seizures and premature death in a mouse model of Dravet syndrome. *Proc. Natl Acad. Sci. USA* **109**, 14646–14651 (2012).
18. Dutton, S. B. et al. Preferential inactivation of *Scn1a* in parvalbumin interneurons increases seizure susceptibility. *Neurobiol. Dis.* **49**, 211–220 (2013).
19. Yu, F. H. et al. Reduced sodium current in GABAergic interneurons in a mouse model of severe myoclonic epilepsy in infancy. *Nat. Neurosci.* **9**, 1142–1149 (2006).
20. Fulco, C. P. et al. Systematic mapping of functional enhancer-promoter connections with CRISPR interference. *Science* **354**, 769–773 (2016).
21. Mo, A. et al. Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* **86**, 1369–1384 (2015).
22. Luo, C. et al. Robust single-cell DNA methylome profiling with snmC-seq2. *Nat. Commun.* **9**, 3824 (2018).
23. Buenrostro, J. D. et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* **523**, 486–490 (2015).
24. Cusanovich, D. A. et al. Epigenetics. Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015).
25. Bejerano, G. et al. Ultraconserved elements in the human genome. *Science* **304**, 1321–1325 (2004).
26. Dimitrieva, S. & Bucher, P. UCNEbase—a database of ultraconserved non-coding elements and genomic regulatory blocks. *Nucleic Acids Res.* **41**(Database issue), D101–D109 (2013).
27. Andersson, R. et al. An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
28. Dousse, A., Junier, T. & Zdobnov, E. M. CEGA—a catalog of conserved elements from genomic alignments. *Nucleic Acids Res.* **44**, 96–100 (2016).
29. Dickel, D. E. et al. Ultraconserved enhancers are required for normal development. *Cell* **172**, 491–499 (2018).
30. Chan, K. Y. et al. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017).
31. Batista-Brito, R. et al. The cell-intrinsic requirement of Sox6 for cortical interneuron development. *Neuron* **63**, 466–481 (2009).
32. Rossignol, E., Kruglikov, I., van den Maagdenberg, A. M., Rudy, B. & Fishell, G. CaV2.1 ablation in cortical interneurons selectively impairs fast-spiking basket cells and causes generalized seizures. *Ann. Neurol.* **74**, 209–222 (2013).
33. Gandal, M. J., Nesbitt, A. M., McCurdy, R. M. & Alter, M. D. Measuring the maturity of the fast-spiking interneuron transcriptional program in autism, schizophrenia, and bipolar disorder. *PLoS ONE* **7**, e41215 (2012).
34. Barnes, S. A. et al. Disruption of mGluR5 in parvalbumin-positive interneurons induces core features of neurodevelopmental disorders. *Mol. Psychiatry* **20**, 1161–1172 (2015).
35. Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
36. Daigle, T. L. A suite of transgenic driver and reporter mouse lines with enhanced brain-cell-type targeting and functionality. *Cell* **174**, 465–480 (2018).
37. Chen, T. W. et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
38. Magnus, C. J. et al. Ultrapotent chemogenetics for research and potential clinical applications. *Science* **364**, eaav5282 (2019).
39. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl Acad. Sci. USA* **104**, 5163–5168 (2007).

40. Yizhar, O. et al. Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature* **477**, 171–178 (2011).
41. Eugène, E. et al. An organotypic brain slice preparation from adult patients with temporal lobe epilepsy. *J. Neurosci. Methods* **235**, 234–244 (2014).
42. Gearing, L. J. et al. CiiiDER: a tool for predicting and analyzing transcription factor binding sites. *PLoS ONE* **14**, e0215495 (2019).
43. Fornes, O. et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
44. Hodge, R. D. et al. Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
45. Boldog, E. et al. Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. *Nat. Neurosci.* **21**, 1185–1195 (2018).
46. Feenstra, B. et al. Common variants associated with general and MMR vaccine-related febrile seizures. *Nat. Genet.* **46**, 1274–1282 (2014).
47. International League Against Epilepsy Consortium on Complex Epilepsies. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol.* **13**, 893–903 (2014).
48. International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide megaanalysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat. Commun.* **9**, 5269 (2018).
49. GTEx Consortium. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
50. Walker, M. C. & Kullmann, D. M. Optogenetic and chemogenetic therapies for epilepsy. *Neuropharmacology* **168**, 107751 (2019).

Methods

Animals. Female C57BL/6J mice (*Mus musculus*; aged 10 weeks), were obtained from Jackson Labs (stock no. 000664). Male hemizygous Dlx6a-cre mice (*Mus musculus*; aged 10 weeks; Jax stock no. 008199) and female homozygous INTACT mice (*Mus musculus*; aged 10 weeks, flox-Sun1-eGFP, Jax stock no. 021039) were maintained at macroenvironmental temperature and humidity ranges of 17.8–26.1 °C and 30–70%, respectively. These parameters were monitored closely and controlled within rodent colony rooms. Sprague Dawley rats (*Rattus norvegicus*, aged 12 weeks, 150–250 g) were obtained from Charles River Labs. One female common marmoset (*Callithrix jacchus*, aged 6 y) was obtained from the colony at Massachusetts Institute of Technology (MIT). Adult (2-year-old) male macaques (*Macaca mulatta*) were obtained from the California National Primate Research Center at the University of California, Davis. All the animals were maintained in a 12 h:12 h light:dark cycle with a maximum of five animals per cage for mice and one animal per cage for rats. Marmosets and macaques were socially housed. All animal maintenance and experimental procedures were performed according to the guidelines established by the Institutional Animal Care and Use Committee at the Broad Institute of MIT and Harvard (mice), the McGovern Research Institute at MIT (rats and marmosets) and the Salk Institute for Biological studies (macaques).

ScATAC-seq library preparation and sequencing. Male hemizygous Dlx6a-cre mice (Jax stock no. 008199) were crossed with female homozygous INTACT mice (flox-Sun1-eGFP, Jax stock no. 021039) to yield Dlx6a-cre::INTACT offspring for scATAC-seq experiments. Brains from P28 Dlx6a-cre::Sun1-eGFP mice were harvested, sectioned coronally on a mouse brain slicer (Zivic Instruments), and the V1 cortex was dissected in ice-cold artificial cerebrospinal fluid (ACSF). Tissue was then transferred to a Dounce homogenizer containing lysis buffer (10 mM Tris-HCl, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Tween-20, and 0.01% IGEPAL CA-630, 0.001% digitonin). Tissue was homogenized with 10 strokes of pestle A and 10 strokes of pestle B, and incubated for 5 min on ice before being filtered through a 30-µm filter and centrifuged at 500*g* for 10 min at 4 °C. The pellet was resuspended in 1% bovine serum albumin for sorting GFP^+ nuclei on a Sony SH800S cell sorter. Nuclei were sorted into diluted nuclei buffer (10x Genomics). The scATAC library was prepared using the 10x Genomics platform with the Chromium Single Cell ATAC Library & Gel Bead Kit v.1.0 (PN-1000111), Chromium Chip E Single Cell kit (PN-1000156) and Chromium i7 Multiplex Kit N, Set A (PN-1000084) as instructed by the manufacturer. High-quality data were recovered from approximately 60% of the input nuclei. Libraries were sequenced using a Nova-Seq S2 100 cycle kit (Illumina).

ScATAC analysis. Raw sequencing data were processed with Cell Ranger ATAC (v.1.1.0) pipeline (10x Genomics) and sequencing reads were aligned to the mouse reference genome (GRCm38-mm10-Mus musculus). The fragments files from the output of this pipeline were then used to generate snap files for analysis using the SnapATAC package v.1 as described previously[51]. Cells were clustered using graph-based clustering ($k=15$, 24 principal components). Gene body accessibility was calculated as described in the SnapATAC package for interneuron subtype marker genes and used to determine clusters corresponding to interneuron cardinal classes. For each cardinal class, bigwig files were generated and peaks were called using MACS2 (v.2.2.7.1; https://github.com/taoliu/MACS) for input into the Integrated Genome Browser (v.2.3) and enhancer selection. Peaks across cardinal classes were compared using Bedtools (v.2.28.0).

Enhancer selection. Candidate regulatory elements were manually curated from a list of elements generated by intersecting the 'context' region (intergenic region limited to 100 kb + intron1 of *Scn1a* for the first part of the study or another seven genes for the larger screen described in Fig. 5), with both the scATAC-seq peak and the conservation data from the PhastCons Analysis. Three criteria as outlined below were used to make these selections (that is, gene selection, accessibility and conservation).

Gene selection. Scn1a was selected based on its known expression profile in PV cINs and its role in disease. The seven genes for the screen presented in Fig. 5 were selected based on a differential gene expression analysis performed independently on both the Tasic et al.[52] and Saunders et al.[53] scRNA-seq datasets. Specifically, each cluster from the original study was manually associated to one of seven meta-clusters (that is, the four cardinal classes in interneurons—PV, SST, VIP, ID2—and two classes of pyramidal neurons—intratelencephalic or subcortical projecting neurons and non-neuronal populations). A total of 92 genes showed significant enrichment in PV cINs in both datasets (that is, the average expression of the genes was at least two times higher in PV cINs compared with other interneuron meta-clusters and five times higher compared with pyramidal or non-neuronal cell types, with an adjusted $P<0.01$ based on pairwise comparison of each metacluster using Student's *t*-test and Bonferroni's correction for multiple comparisons). We then focused on the top 20 genes with the highest average expression in PV cINs, and manually selected a final subset of 7 genes that showed an expression pattern consistent with PV cIN enrichment using ISH data from the Allen Brain Atlas mouse brain (that is, sparse labeling in the S1 cortex with biased distribution toward the deep cortical layers). The seven genes included in the final selection were *Pvalb*, *Acan*, *Tmem132c*, *Lrrc38*, *Inpp5j*, *Mef2c* and *Pthlh*.

Accessibility. ScATAC-seq peaks obtained for each of the four classes of interneurons (see "Gene selection" section above) were examined. Bulk ATAC-seq data for cortical excitatory neurons generated in Mo et al.[21] were downloaded from the Gene Expression Omnibus (GEO) repository (accession no. GSE63137) and discretized as peaks using MACS2 (v.2.2.7.1; https://github.com/taoliu/MACS).

Conservation. The 'phastCons-Element60way' track was downloaded from the University of California, Santa Cruz portal (https://genome.ucsc.edu) using the table browser interface. We then used the bedtools suite to remove blocks shorter than 10 bp and subsequently merged any blocks separated by less than 50 bp. The resulting file was then imported to the IGV genome browser and used to identify genomic regions with a high degree of conservation across species.

Identification of human homolog sequence of each mouse enhancer. To find the human homolog of the mouse enhancer sequences, we used the EMBOSS Needle portal (https://www.ebi.ac.uk/Tools/psa/emboss_needle/). This tool is based on the Needleman–Wunsch alignment algorithm which finds an optimal global sequence alignment by penalizing starting and extending gaps. Using this approach, we compared the mouse sequence of a given enhancer with a region 100 kb up- and downstream around the human gene associated with the enhancer (that is, the mouse E2 enhancer sequence was compared with a ~200-kb region of the human genome centered on *SCN1A*). In all cases tested, we identified a human sequence of similar length showing a high degree of conservation with the mouse sequence.

Degree of conservation between mouse and human. We used the output of the analysis described above (see "Enhancer selection" and "Identification of human homolog sequence of each mouse enhancer" sections) and calculated the percentage of base-pairs from mice that are identical in humans. This criterion was used to prioritize which enhancers were tested in non-human primates.

TFBS enrichment analysis. TFBSs enriched in the mouse or human enhancer sequences were identified using the CiiiDER pipeline[42,43]. The analysis was performed using the mouse or human motifs from the Jaspar 2020 core redundant vertebrates TFBS database (http://jaspar.genereg.net/tools). TFBSs were classified based on whether the motif sequence predominantly overlaps with a conserved or nonconserved region, excluding TFBSs with exactly 50% overlap. The EMBOSS Needle tool was used to define regions of conservation by using the mouse sequence and human homolog as inputs. Next, putative human TFs that recognize motifs on the conserved regions of the mouse sequence were identified using the human Jaspar motifs and mouse sequence as inputs to CiiiDER. Finally, the putative human and mouse TFBSs on the mouse sequence were compared to define whether any equivalent TFs between the two species corresponded on the same location.

Recombinant AAV cloning and production. The enhancer sequences were amplified by PCR from mouse genomic DNA using the following primers: E1: caaagtggacagaggggagg and gtgctgttgggagtggtgga (1,280 bp); E2: aatctaacatggctgctata and caattgctcagagttatttt (618 bp); E3: ataaaatttttattttcctaa and gaggaaatcagctacggggc (832 bp); E4: tctgacagagcaagtcttga and tatcaaaattgtatattcag (261 bp); E5: aatgttttgatatttaggag and ttgactcttaaaatttaata (663 bp); E6: ttgtcactttgttactctac and ttaaatcttaaaatttttcct (606 bp); E7: gatactgtataattaattag and cttccttctggttccttttt (2430 bp); E8: attgatctccaacttttttaa and gttcatccaagtaataagag (1,644 bp); E9: atctcaagtgtatgtaacat and gtcttttttgttttttttttt (521 bp); E10: tattgcaaaaggaaggaatg and tcatggaaaaagaaaaaatc (547 bp); E11: actgtttgagcgggcagagc and agttcttaacattcagattt (504 bp); E14: tcggttctctctgtagtctg and gcggggaggggtgtcttcaa (496 bp); E22: tcgggtttaaccgttcaga and gggtttaaatatagagtccc (443 bp); and E29: atagtggcgctatgtgggag and agagtagaggctttgccttg (628 bp). The enhancers, reporters and effectors were cloned using the Gibson Cloning Assembly Kit (New England BioLabs, catalog no. NEB-E5510S) following standard procedures. Specifically, for AAV-E1:10-dTomato, we amplified the dTomato coding sequence from the plasmid Addgene no. 83897; for AAV-E2-SYP-dTomato, we amplified the SYP–tdTomato coding sequence from the plasmid Addgene no. 34881; for AAV-E2-GCaMP6f, we amplified the GCaMP6f coding sequence from the plasmid Addgene no. 83899; and for AAV-E2–C1V1–enhanced yellow fluorescent protein (eYFP), we amplified the C1V1–eYFP coding sequence from the plasmid Addgene no. 35499. For AAV-E11, -14, -22 and 29-ChR2–GFP, we amplified the ChR2-coding sequence from the plasmid Addgene no. 55639 and complemented it with a GFP–P2A–GFP cassette. The rAAVs were produced using standard production methods. Polyethylenimine was used for transfection and OptiPrep gradient (Sigma) was used for viral particle purification. Serotype 1 was used to produce the AAVs for local injections in mice and rats. Serotype 9 was used for systemic injection in marmosets and serotype PHPeB was used for both local injection in macaques and systemic injections in mice. Titer was estimated by quantitative PCR with primers for the WPRE sequence that is common to all constructs. All batches produced were in the range 10×10^{11} to 10×10^{13} viral genomes per milliliter.

Local and systemic viral injections. *Mouse local S1.* Local injection in adult mice was performed by stereotactically guided injections into the S1 cortex with the following coordinates: 1.0 mm posterior, 2.9 mm lateral and 0.7/0.45 mm ventral relative to Bregma with 75 nl of virus produced with the capsid PHPeB.

Mouse systemic. For systemic injection in adult mice, 10×10^{11} viral particles produced with the capsid PHPeB were injected into the retro-orbital sinus per animal. Postoperative monitoring was performed for 5 d post-injection. For rat local in the V1 cortex, local injection of the AAV produced with the capsid PHPeB in adult rats was performed by stereotactically guided injections into the P1 cortex with the following coordinates: 5.4 mm posterior, 4.2 mm lateral and 2.0 mm ventral relative to Bregma with 670 nl of virus.

Marmoset systemic injection. For systemic injection in adult marmosets, approximately 10×10^{12} viral particles produced with the serotype 9 in ~0.7 ml sterile phosphate-buffered saline (PBS) were injected into the saphenous vein, followed by another infusion with ~0.5 ml saline. After the final infusion, pressure was applied to the injection site to ensure hemostasis. The animal was returned to its home cage and monitored closely for normal behavior after anesthesia. The animal was euthanized 51 d after viral injection.

Macaque local. Local injection in adult macaques was performed by a stereotactically guided injection of virus produced with the capsid PHPeB at the following coordinates: 13 mm anterior, 19 mm lateral, 23 mm superior (V1 cortex), 27 mm anterior, 18 mm lateral, 34 mm superior (PFC), 5 mm posterior, 13 mm lateral and 40 mm superior (V1 cortex), relative to the center of the interaural line (based on the animal's magnetic resonance scan). A total of volume of 1,332 nl was injected, equally divided at 4 depths (that is, 1.8, 1.3, 0.8 and 0.3 mm from the cortical surface).

Electrophysiological recordings in mice. Virally injected mice were anesthetized with isoflurane. On the loss of reflexes, mice were transcardially perfused with ice-cold oxygenated ACSF containing the following (in mM): 87 NaCl, 75 sucrose, 2.5 KCl, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 10 glucose, 1 $CaCl_2$ and 2 $MgCl_2$. Mice were then decapitated and 300-µm-thick coronal slices were sectioned using a Leica VT-1200-S vibratome and incubated in a holding chamber at 32–35 °C for 15–30 min, followed by continued incubation at room temperature 20–23.5 °C for at least 45–60 min before physiological recordings. Slices containing the injection site were transferred to a recording chamber submerged with oxygenated ACSF containing the following (in mM): 125 NaCl, 2.5 KCl, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 10 glucose, 2 $CaCl_2$ and 1 $MgCl_2$ (pH 7.4, bubbled with 95% O_2 and 5% CO_2). Whole-cell current-clamp recordings were obtained from visually identified cells expressing the viral reporter using borosilicate pipettes (3–5 MΩ) containing (in mM): 130 potassium gluconate, 6.3 KCl, 0.5 (ethylenebis(oxonitrilo))tetra-acetate (EGTA), 10 4-(2-hydroxyethyl)-1-piperazine-ethanesulfonic acid (Hepes), 4 Mg-ATP, 0.3 Na-GTP and 0.3% biocytin (pH adjusted to 7.3 with KOH). On break-in, series resistance (typically 15–25 MΩ) was compensated for and only stable recordings (<20% change) were included. Data were acquired using a MultiClamp 700B amplifier (Molecular Devices), sampled at 20 kHz and filtered at 10 kHz. All cells were held at −60 mV with a DC current, and current-step protocols were applied to obtain firing patterns and to extract basic subthreshold and suprathreshold electrophysiological properties. Cells not expressing the viral reporter were selected according to their pyramidal cell-shaped soma under infrared–differential interference contrast visualization, and voltage-clamp recordings were made with pipettes containing (in mM): 130 cesium gluconate, 0.5 EGTA, 7 KCl, 10 Hepes, 4 Mg-ATP, 0.3 Na-GTP, 5 phosphocreatine, 5 QX-314 and 0.3% biocytin (pH adjusted to 7.3 with CsOH). Cells were held at 0 mV for baseline and optogenetic or chemogenetic stimulation. For both current- and voltage-clamp recording, a baseline of at least 2 min was recorded before stimulation. Small pulses (−20 pA or −5 mV, 100 ms at 0.2 Hz or 0.5 Hz) were applied throughout the baseline and clozapine-*N*-oxide (CNO) application to monitor series resistance changes. Data were analyzed offline using Clampfit software (v.10.2; Molecular Devises).

In vivo calcium imaging. Approximately 100 nl AAV-E2-GCaMP6 virus was injected into the barrel cortex at P10. At P27–P34, craniotomies were implanted over the injection site and widefield calcium imaging performed after recovery from a craniotomy procedure. Briefly, anesthetized (1.5% isoflurane) mice were imaged at 3–4 Hz with ×4 magnification (Thorlabs CCD camera; 1501M-USB, Thorlabs LED stimulation, DC4104), while air puffs (100–200 ms duration, Picospritzer III) at specific intervals (5–20 s) were directed at contralateral whiskers. Multiple recordings were performed and afterward the mouse was perfused for histological analysis. Recordings were analyzed in ImageJ by calculating the $\Delta F/F$ (change in fluorescence/average fluorescence) for each recording and synchronized whisker stimulation. A threshold of (5%) $\Delta F/F$ was set for both stimulated and spontaneous calcium signal response.

Electrophysiological recordings in humans. *Tissue preparation, culture protocol and inoculation of virus.* Four participants (two males/two females; age range 22–57 years) underwent a surgical procedure in which brain tissue (temporal lobe and hippocampus) was resected for the treatment of drug-resistant epilepsy. In all cases, each participant had previously undergone an initial surgery for placement of subdural and/or depth electrodes for intracranial monitoring to identify the location of seizure onset. The National institute of Neurological Disorders and Stroke (NINDS) Institutional Review Board (IRB) approved the research protocol (ClinicalTrials.gov ID NCT01273129), and we obtained informed consent from the participants for experimental use of the resected tissue. Slices of 300 µm from both hippocampus and temporal lobe were attained (Leica 1200S Vibratome; Leica Microsystems) in ice-cold, oxygenated, sucrose-based cutting solution (100 mM sucrose, 80 mM NaCl, 3.5 mM KCl, 24 mM $NaHCO_3$, 1.25 mM NaH_2PO_4, 4.5 mM $MgCl_2$, 0.5 mM $CaCl_2$ and 10 mM glucose, saturated with 95% O_2 and 5% CO_2) within 30 min of neurosurgical resection. Slices were then incubated in the sucrose cutting solution at 33 °C for 30 min and allowed to cool to room temperature for 15–30 min. The slices were transferred to culture medium[41] and placed in an incubator (5% CO_2) at 35 °C for 15 min of equilibration. Each individual slice was then transferred on to a 30-mm Millicell Cell Culture Insert (Millipore, catalog no. PICM0RG50) for interface culture. After 12 h, the culture medium was changed and 1–2 µl pAAV-S5E2–dTomato with or without pAAV-S5E2–C1V1–eYFP was directly pipetted on to each slice and placed back into the incubator. For hippocampal slices, the virus was targeted to the subiculum subfield. Electrophysiological recordings from cultured human slices were performed between 7 and 14 d after viral inoculation. Cultured human slices were transferred to a recording chamber perfused with extracellular solution (130 mM NaCl, 3.5 mM KCl, 24 mM $NaHCO_3$, 1.25 mM $NaH_2PO_4{\cdot}H_2O$, 10 mM glucose, 2.5 mM $CaCl_2$ and 1.5 mM $MgCl_2$ saturated with 95% O_2/5% CO_2 (pH 7.4; 300–310 mosmol) at a rate of 3–4 ml min^{-1} at 33 °C. Whole-cell patch-clamp recordings from pAAV_S5E2–dTomato- or pAAV-S5E2–C1V1–eYFP-infected neurons were performed with an intracellular solution of the following composition: 130 mM potassium gluconate, 10 mM Hepes, 0.6 mM EGTA, 2 mM $MgCl_2$, 2 mM Na_2ATP, 0.3 mM Na-GTP and 0.5% biocytin (pH adjusted to 7.4; osmolarity adjusted to 285–300 mosmol). Intrinsic membrane and firing properties were assayed as described previously (see "Electrophysiological recordings in mice" section above for details). Then, 550-nm light-stimulated optogenetic activation of C1V1 was delivered to the slices via the ×40 water immersion objective using a CoolLED pE-4000 Illumination system.

Immunohistochemistry. All antibodies used in the present study are commercially available and have been validated by the manufacturer. In addition to the validation statement that can be found by consulting the manufacturer's website using the references provided, the specificity of each primary antibody was examined and confirmed for the species for which it was used, based on examining the signal intensity, the density of staining and the consistency with morphological features of the cellular populations, and is presented in the relevant panels across the figures and Extended Data Figs. All animals injected with the virus were transcardially perfused with 4% paraformaldehyde (PFA). The brains were placed in 4% PFA overnight then sectioned at 40 µm using a Leica VTS1000 vibrosector. Floating sections were permeabilized with 0.3% Triton X-100, 5% normal donkey serum and PBS for 30 min. The sections were then incubated overnight in 0.1% Triton X-100 with the indicated combination of the following primary antibodies at 4 °C: chicken anti-GFP at 1:1,000 (Abcam, catalog no. ab13970); rabbit anti-DsRed at 1:1,000 (Clontech, catalog no. 632496); goat anti-PV at 1:1,000 (Swant, catalog no. PVG-213); guinea-pig anti-PV at 1:2,000 (Swant, catalog no. GP-72); rabbit anti-SST at 1:2,000 (Peninsula, catalog no. T-4103.0050); and mouse anti-Synaptotagmin-2 at 1:250 (ZFIN, catalog no. ZDB-ATB-081002-25). The sections were then washed three times with PBS, incubated with Alexa-Fluor-conjugated secondary antibodies at 1:1,000 (Invitrogen), counterstained with DAPI (Sigma) and mounted on glass slides using Fluoromount-G (Sigma).

In situ hybridization. The ISH probes (Gad1; product no. 400951, Pvalb; product no. 421931, VIP; product no. 415961) used were designed by Advanced Cell Diagnostics. The reagents in the RNAscope Multiplex Fluorescent Reagent Kit v.2 (product no. 323100), RNAscope Probe Diluent (product no. 300041), HybEZ oven (product no. 321710/321720), humidity control tray (product no. 310012) and HybEZ Humidifying Paper (product no. 310025) were also from Advanced Cell Diagnostics. TSA Plus Fluorescein, TSA Plus Cyanine 3 and TSA Plus Cyanine 5 were from PerkinElmer (nos. NEL741, NEL744 and NEL745). Brain tissue was processed as mentioned under Immunohistochemistry. Brain sections were washed once in PBS, followed by three washes in 0.1% Triton X-100 and PBS, mounted on Superfrost Plus glass slides (Thermo Fisher Scientific, 12-550-15) and baked at 60 °C in the HybEZ oven for 25 min. The slides were then submerged in 4% PFA for 30 min and washed three times in H_2O. RNAscope H_2O_2 was applied to each section for 5 min at room temperature. The slides were then washed three times in H_2O before being submerged in pre-warmed 90 °C H_2O for 15 s, followed by pre-warmed 90 °C RNAscope Target Retrieval for 15 min. Slides were washed three times in H_2O before RNAscope Protease III was applied to each section, and then incubated for 15 min at 40 °C in the HybEZ oven. Slides were washed three times in H_2O and then incubated with probe solution diluted to 1:50 with probe diluent for 2 h at 40 °C in the HybEZ oven. Next, the sections were washed three times in

RNAscope wash buffer followed by fluorescence amplification. Of note, probes against the RNA of the reporter revealed a non-specific staining that we speculate comes from the viral DNA. To reveal the viral reporter, we followed the RNAscope protocol with an IHC amplification of the dTomato. The sections were incubated in blocking solution (0.3% Triton X-100 + 5% normal horse serum in PBS) for 30 min. Then sections were incubated in antibody solution (0.1% Triton X-100 + 5% normal horse serum in PBS) with rabbit anti-DsRed at 1:250 (Clontech, 632496) at 4 °C overnight. The sections were then washed three times with PBS, incubated with Alexa-Fluor-conjugated secondary antibodies at 1:500 (Invitrogen), stained with DAPI (Sigma) and mounted on glass slides using Fluoromount-G (Sigma).

Quantifications, statistics and reproducibility. For strength of expression, the fluorescence images were taken at a standardized magnification and exposure time, and the average pixel intensity of the cell bodies of each cell expressing the viral reporter was recorded and reported as an average over all cells per enhancer. For quantification of co-localization, cells expressing the indicated reporter were counted using only the corresponding color channel and then, among these cells, the number of cells co-expressing the marker of interest was counted. A cell was considered to be positive for a given marker if the corresponding signal was above background fluorescence. The ratio of cells co-expressing both markers over the total number of cells expressing only the reporter was then calculated and reported as mean ± s.e.m. Quantifications were performed using a minimum of two independent biological replicates. Several sections from the same animal were used when indicated. Data collection and analysis were not performed blind to the conditions of the experiments, but experimenters from different research groups performed the quantification. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications[6]. Two-tailed, unpaired Student's *t*-tests were performed to estimate the statistical differences between the indicated test and control groups in Figs. 3c and 4b(iv) and Extended Data Fig. 3c. Equal variance between the two populations was not tested. The specific *P* values, *t* and degrees of freedom (d.f.) for each test are indicated in the legends of the figures. The images presented in Figs. 2a, 3a and 4a,b and Extended Data Figs. 1b, 2d,f and 3a show representative data that were obtained from at least three independent injections of the indicated virus in the indicated species. The image for the marmoset injection in Fig. 4a represents the specificity and cellular density observed across 12 sections from the same animal. The specific number of cells, animals and conditions, as well as the types of replicates, are reported in Supplementary Table 2. At least two biological replicates per data point were included for all quantifications used in the present study. In all cases where the number of biological replicates was not more than two, the replicates were highly consistent. For ethical and economic reasons, the data generated for systemic viral injection in a marmoset and for local injection in macaques came from a unique animal and were not repeated. The staining of PV IHC within human brain tissues was highly variable. As such, estimates of viral specificity were made within regions of the cortex and subiculum, where staining density reflected the known distribution and density of these cells. Sections where the PV IHC did not reflect the known distribution of these cells were excluded from the study.

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

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The scATAC datasets presented in the present study are available on GEO with the accession no. GSE152449. All AAV plasmids and their corresponding sequences are available from Addgene (pAAV-S5E2-dTom-nls-dTom (Addgene, no. 135630), pAAV-S5E2-GFP-fGFP (Addgene, no. 135631), pAAV-S5E2-GCaMP6f (Addgene, no. 135632), pAAV-S5E2-C1V1–eYFP (Addgene, no. 135633), pAAV-S5E2-ChR2-mCherry (Addgene, no. 135634), pAAV-S5E2-Gq-P2A-dTomato-short (Addgene, no. 135635), pAAV-S5E1-dTom-nlsdTom (Addgene, no. 135637), pAAV-S5E3-dTom-nlsdTom (Addgene, no. 135638), pAAV-S5E4-dTom-nls-dTom (Addgene, no. 135639), pAAV-S5E5-dTom-nlsdTom (Addgene, no. 135640), pAAV-S5E6-dTom-nlsdTom (Addgene, no. 135641), pAAV-S5E7-dTom-nls-dTom (Addgene, no. 135642), pAAV-S5E8-dTom-nlsdTom (Addgene, no. 135643), 018_pAAV-S5E9-dTom-nls-dTom (Addgene, no. 135644), pAAV-S5E10-dTom-nlsdTom (Addgene, no. 135645), pAAV-E11-ChR2GFP2x (Addgene, no. 153434), pAAV-E14-ChR2GFP2x (Addgene, no. 153435), pAAV-E22-ChR2GFP2x (Addgene, no. 153436) and pAAV-E29-ChR2GFP2x (Addgene, no. 153437)).

References

51. Fang, R. et al. Fast and accurate clustering of single cell epigenomes reveals *cis*-regulatory elements in rare cell types. Preprint at *bioRxiv* https://doi.org/10.1101/615179 (2019).
52. Tasic, B. et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19**, 335–346 (2016).
53. Saunders, A. et al. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030.e16 (2018).

Acknowledgements

We thank S. Gerard, P. Delvenne and C. Parotte for their useful comments on the manuscript. J.D. is supported by NIH grants (nos. R01-MH111529 and UG3MH120096) as well as the Simons Foundation Award (no. 566615) and a gift from the Friends-Of-FACES foundation. T.P.F. is supported by fellowships from the Belgian American Educational Foundation and the George E. Hewitt Foundation for Medical Research, an NARSAD Young Investigator grant from the Brain and Behavior Research Foundation, and a Kavli Institute for Brain and Mind Innovative research grant. C.J.M. was awarded an Intramural Research Program of NINDS and a Eunice Kennedy Shriver NICHD Intramural grant. J. Smith is supported by the NINDS (grant no. K99 NS106528). G. Fishell is supported by the NINDS (grant nos. NS081297, NS074972), NIMH (grant no. MH071679) and NIH (grant no. UG3MH120096), and the Harvard's Dean Initiative, as well as support from the Simons Foundation award (no. 566615). We thank the individuals who selflessly volunteered to participate in this study.

Author contributions

D.V. and J.D.L. designed and performed experiments, analyzed the data, prepared the figures and wrote the manuscript. K.A. performed and analyzed the scATAC-seq experiments. J.D.L. and G.A.S. provided computational support for the analysis. K.A.P., R.C., X.Y. and K.A.Z. performed experiments and analyzed the data related to human brain tissue. B.G., M.A.A., S.S., B.L.G., O.S., G.S., S.V., J.V., L.A.I., K.J.M., E.S., S.H., E.F, T.B., I.V. and V.S. performed experiments and analyzed the data related to mice. Q.X. and L.G. produced the AAVs in NYUAD using plasmids conceived and generated at the Broad Institute. J. Sharma and Q.Z. performed experiments and analyzed the data related to rats and marmosets. T.P.F. and J. Smith performed experiments and analyzed the data related to macaques. O.D., B.L.S., R.B., J.R., G. Feng, Z.F., C.J.M. and G. Fishell helped with the study design and provided collaborative support for the work. J.D. designed the study, conceived the viral construct, contributed to writing the manuscript and preparing the figures, and supervised the project. All the authors edited and approved the manuscript.

Competing interests

The Broad Institute of MIT and Harvard has filed patent applications related to this work with G. Fishell and J.D. listed as inventors.

Additional information

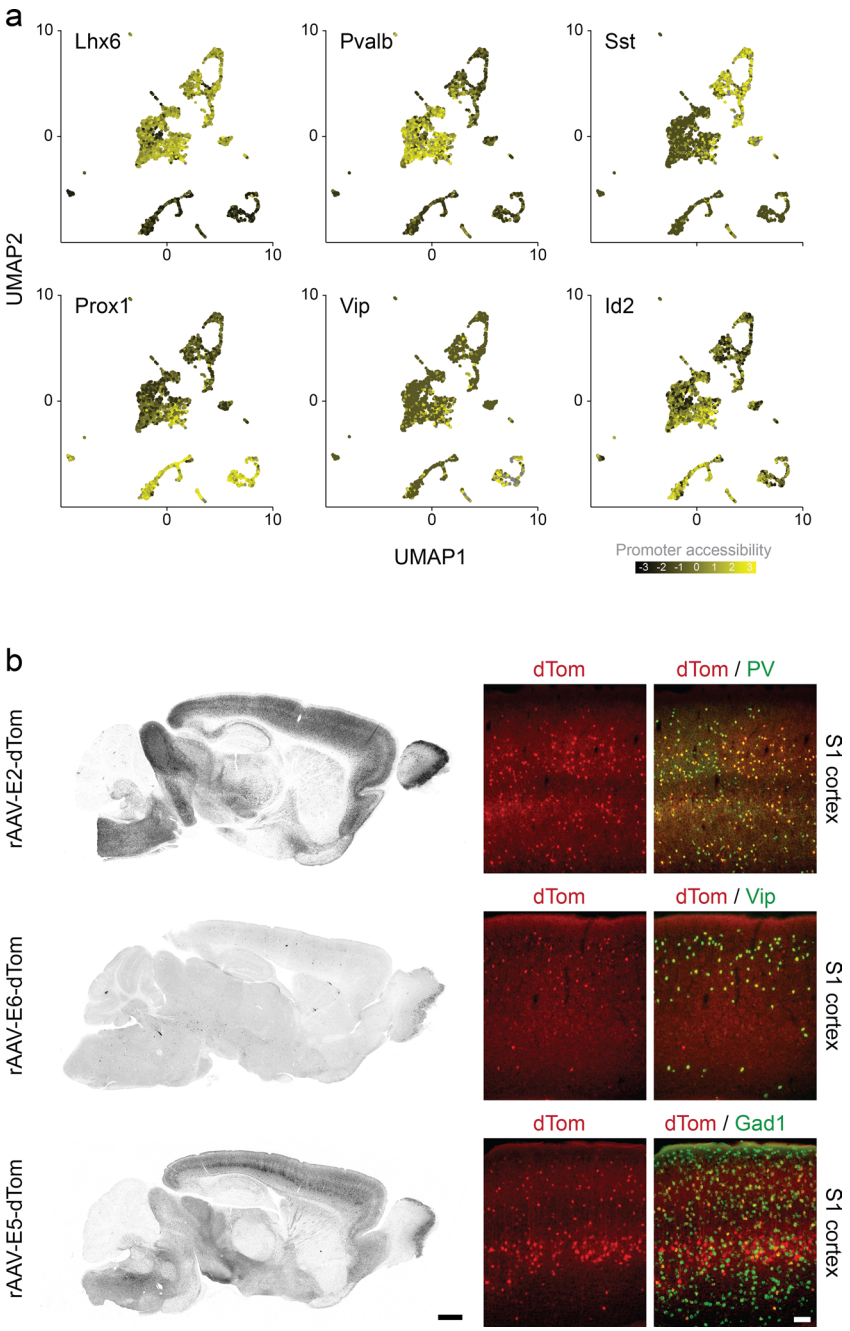
Extended data is available for this paper at https://doi.org/10.1038/s41593-020-0692-9.

Supplementary information is available for this paper at https://doi.org/10.1038/s41593-020-0692-9.

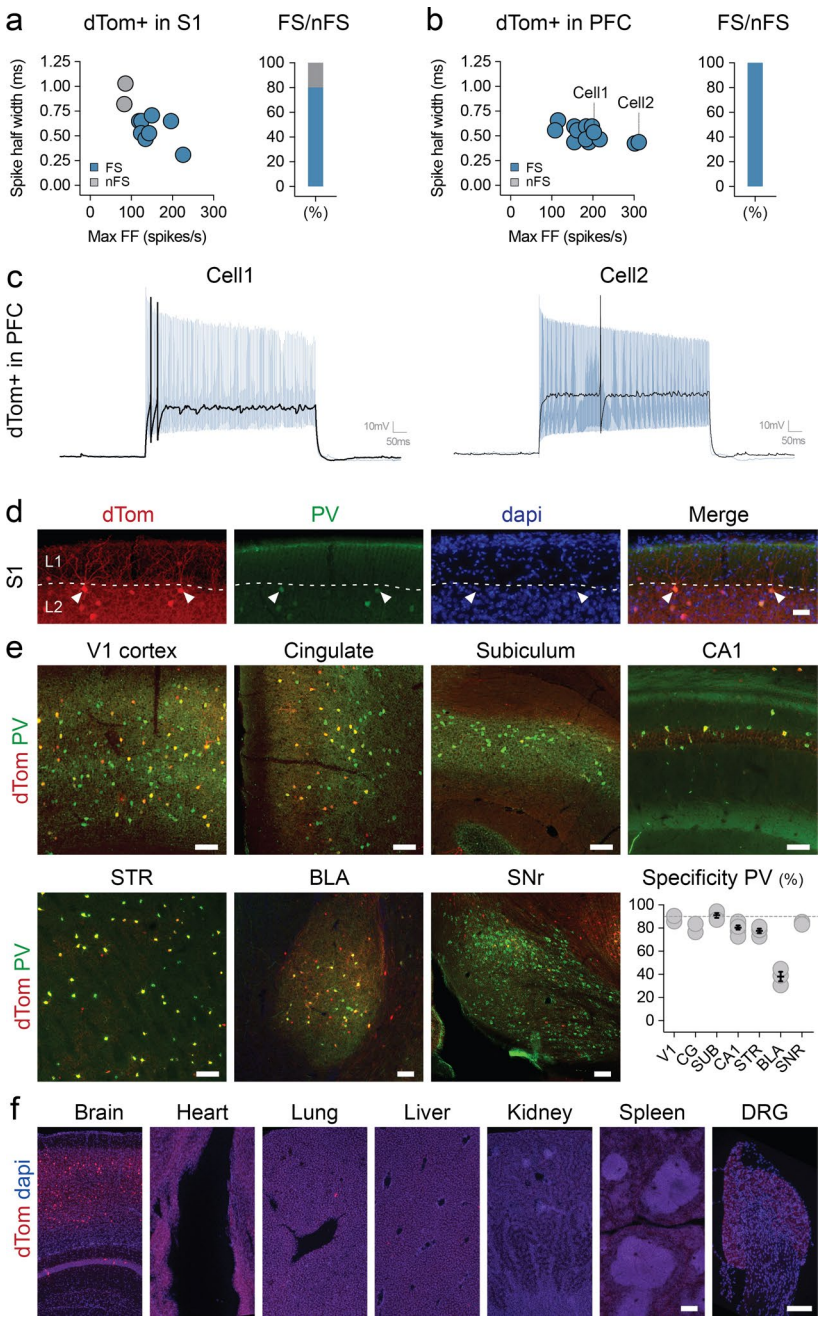
Correspondence and requests for materials should be addressed to J.D.

Peer review information *Nature Neuroscience* thanks Z. Josh Huang, Botond Roska and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

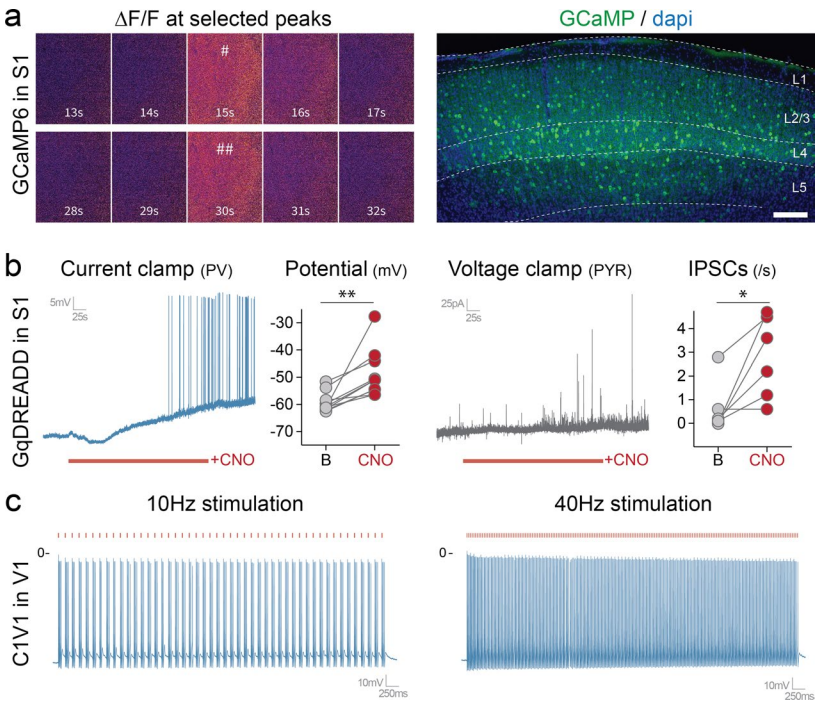
Reprints and permissions information is available at www.nature.com/reprints.



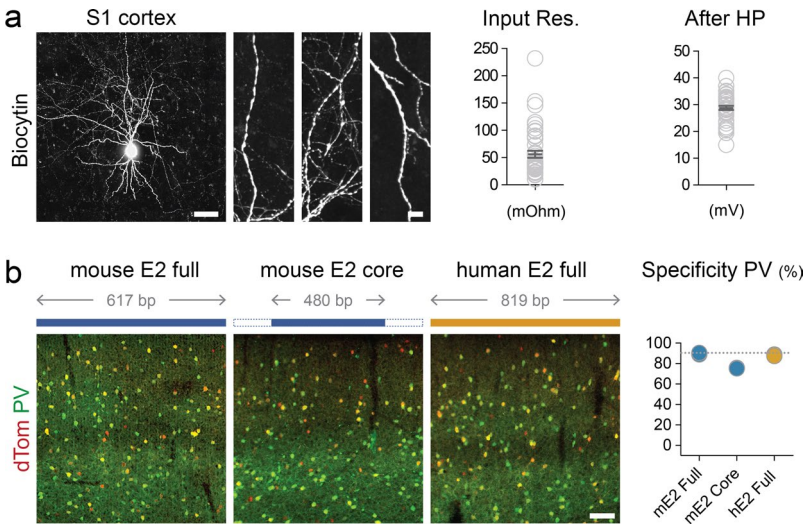
Extended Data Fig. 1 | Enhancer selection and top candidates. a, UMAP plot of 3500 neuronal nuclei collected from 4 Dlx6a-cre::Sun1-eGFP mice showing promoter accessibility of the indicated canonical interneuron markers. **b**, Fluorescent images of sagittal sections from adult mice that were injected systemically with the indicated rAAV-E[x]-dTom and analyzed 3 weeks post-injection with IHC for the viral reporter. Scale bar for left panels represents 500um; Scale bar for right panels represents 100um. See method section for details on the reproducibility of the representative images presented in panel **b**.



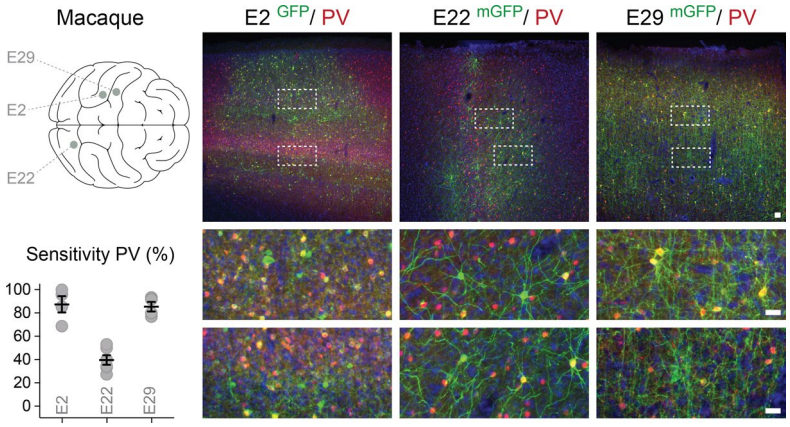
Extended Data Fig. 2 | E2 regulatory element to drive expression of reporters. Adult mice were injected systemically with rAAV-E2-dTomato. **a, b**, Slice recording of the intrinsic properties of virally labeled neurons in S1 cortex and PFC. The left panels show plots of recorded cells with the indicated intrinsic properties. The blue dots represent cells with stereotypical fast-spiking properties. The right panels indicate the proportion of fast spiking cells recorded. **c**, Representative slice recording traces of cells indicated in (**b**). (**d**) Representative image of virally labeled chandelier cells. (**e**) Coronal and sagittal sections were analyzed with IHC for the viral reporter and PV, and the specificity to PV was reported across brain regions. (**f**) The native viral expression was analyzed from the indicated organs. Scale bars represent 50um (**d**), 100um (**e**) and 250um (**f**). On the graphs, dots represent individual measurements and the lines represent average +/- s.e.m. Values for specificity are listed in the supplementary table 2. See method section for details on the reproducibility of the representative images presented in panel **d** and **f**.



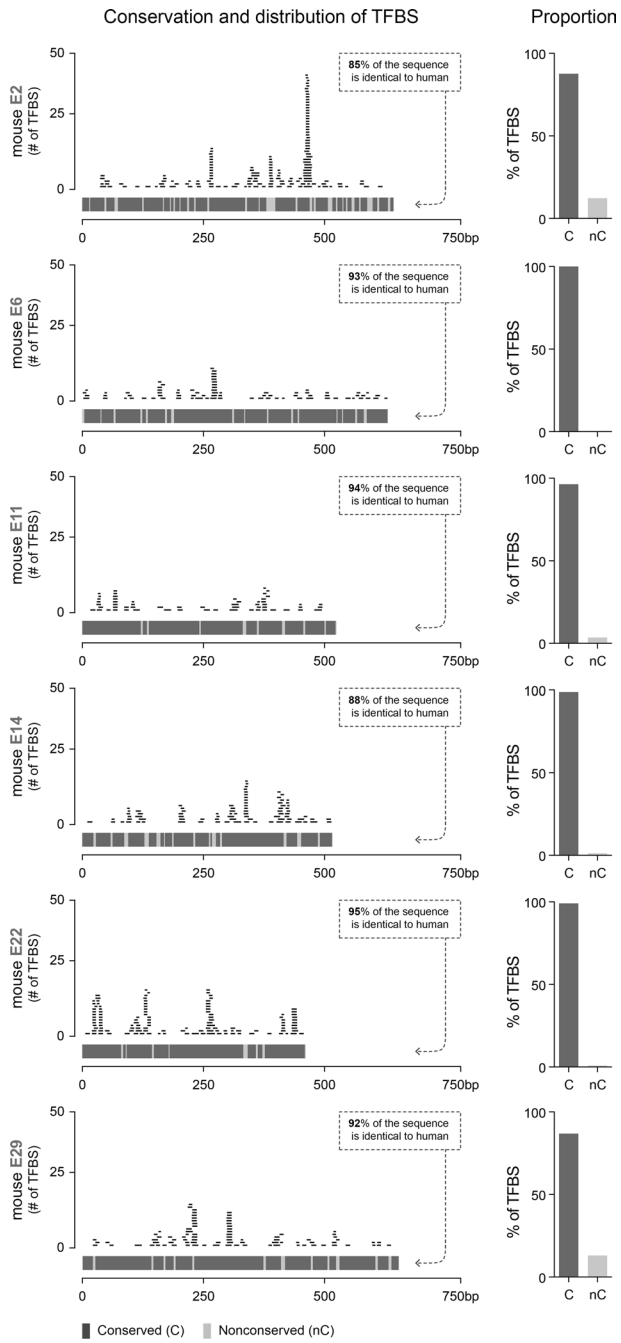
Extended Data Fig. 3 | E2 regulatory element to drive expression of effectors. Mice were injected locally with the following constructs (**a** - P14 injection with rAAV-E2-GCaMP6f; **b** – rAAV-E2-C1V1-eYFP; **c** – rAAV-E2-GqDREADD). (**a**) Mice were analyzed 1-week post-injection. The left panel shows widefield images of two representative peaks shown by the pound signs in Fig. 3. The right panel shows a fluorescent image taken after GCaMP recordings. (**b**) Slice electrophysiology current clamp recordings were performed 1-week post-injection. Cells expressing the viral reporter were targeted with either 10 Hz or 40 Hz laser stimulation (550 nm) while the voltage was recorded over 3 seconds. (**c**) Slice electrophysiology current clamp recordings were performed 1-week post-injection. The voltage was recorded before and after bath application of CNO. Scale bars represent 500um. The red bars represent laser stimulation. On the graphs, dots represent individual measurements. c right panel: p-value = 0.0039; t = 3.859; df = 9; c right panel: p-value = 0.0258; t = 3.135; df = 5); *, ** and *** correspond to a p-value < 0.01, <0,001 and < 0,0001 respectively. See method section for details on the reproducibility of the representative images presented in panel a.



Extended Data Fig. 4 | E2 regulatory element works across species. a, Human brain tissue obtained from surgical resection that was exposed to rAAV-E2-dTomato and maintained in culture for 7–14 days. Left - Representative image of the dendrites of virally labeled cells filled with Biocytin during the recording session. Right - Slice recording of the intrinsic properties of virally labeled neurons. The quantifications show the indicated parameters. Scale bars represent 100um for the left images and 2um for the right images. **b**, Adult mice were injected with the indicated modified rAAV-E2-dTomato construct and analyzed 3 weeks post-injection with IHC for the viral reporter and PV. The corresponding specificity is shown in the right panel. Scale bars represent 100um. On the graphs, dots represent individual measurements and the lines represent average +/- s.e.m. Values for specificity are listed in the supplementary table 2.



Extended Data Fig. 5 | Enhancer screen applied to additional genes. Adult macaques were injected locally in the prefrontal or S1 cortex with the indicated rAAV-E[x]-eGFP and analyzed 8 weeks post-injection with immunohistochemistry for the reporter and indicated markers. The right panels display the injection sites (above), and the boxed quantified regions (below). The corresponding sensitivity is shown in the bottom left. Scale bars represent 25um (lower panels) and 50um (upper panels). On the graphs, dots represent individual measurements and the lines represent average +/- s.e.m. Values for sensitivity are listed in the supplementary table 2.



Extended Data Fig. 6 | Transcription factor binding site enrichment. Each panel shows the indicated enhancer sequence displayed on a fixed region of 750 bp, where each block of the lower part of the left graph shows conserved (dark gray) and non-conserved (light gray) regions of the enhancer. On the upper part of the left graph, each trace shows an individual transcription factor binding site mapped using CiiiDER (see methods). The blue traces represent TFBS found only in mice and the orange traces represent the TFBS found both in mice and humans. The two bar charts show the proportion of the TFBS found either on conserved or non-conserved regions of the enhancer, for all TFBS and for the subset of conserved TFBS, respectively.

nature research

Corresponding author(s): Jordane Dimidschstein

Last updated by author(s): Jun 19, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect the data
Data analysis	- ImageJ (v1.8.0) was used to manipulate images and perform co-localization analysis. - 10X Cell-Ranger (v1.1.0) was used to process the scATACseq data. - snapATAC (v1) was used as the Analysis Pipeline for Single Cell ATAC-seq - macs2 (v2.2.7.1) was used to call peaks from the scATACseq data. - bedtools (v 2.28.0) was used to identify enhancer sequances - Integrated genome browser v2.3 was used to visualize the bed files generated for enhancer identification - EMBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) was used to align and compare enhancer sequences between mice and humans. - CiiiDER pipeline (http://ciiider.com) was used to analyze transcription binding site enrichment in the indicated enhancer sequences. - Clampfit software (v10.2) was used to analysis electrophysiological recordings.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The scATAC datasets presented in the study (figure 1 and supplementary figure 1) are available on GEO with the accession number GSE152449. All AAV plasmids and their corresponding sequences are available on Addgene (pAAV-S5E2-dTom-nlsdTom Addgene#135630; pAAV-S5E2-GFP-fGFP Addgene#135631; pAAV-S5E2-GCaMP6f Addgene#135632; pAAV-S5E2-C1V1-eYFP Addgene#135633; pAAV-S5E2-ChR2-mCherry Addgene#135634; pAAV-S5E2-Gq-P2A-dTomato-short Addgene#135635; pAAV-S5E1-dTom-nlsdTom Addgene#135637; pAAV-S5E3-dTom-nlsdTom Addgene#135638; pAAV-S5E4-dTom-nlsdTom Addgene#135639; pAAV-S5E5-dTom-nlsdTom Addgene#135640; pAAV-S5E6-dTom-nlsdTom Addgene#135641; pAAV-S5E7-dTom-nls-dTom Addgene#135642; pAAV-S5E8-dTom-nlsdTom Addgene#135643; 018_pAAV-S5E9-dTom-nls-dTom Addgene#135644; pAAV-S5E10-dTom-nlsdTom Addgene#135645; pAAV-E11-ChR2GFP2x Addgene#153434; pAAV-E14-ChR2GFP2x Addgene#153435; pAAV-E22-ChR2GFP2x Addgene#153436; pAAV-E29-ChR2GFP2x Addgene#153437).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical method was used to determine the sample size. We selected the highest number of samples we could perform under reasonable financial and logistical constraints to provide precise and accurate estimates of the data's central tendency and variance and allow for the computational of confidence intervals around estimates, with at least 2 biological replicates per conditions.
Data exclusions	The staining of PV IHC within human brain tissues was highly variable. As such, estimates of viral specificity were made within regions of cortex and subiculum where staining density was reflective of the known distribution and density of these cells. Sections where the PV-IHC was not reflective of the know distribution of these cells were excluded from the study.
Replication	At least two biological replicates per data point were included for all quantification presented in this study. In all cases where the number of biological replicates was not above 2, the replicates were highly consistent. For ethical and economic reasons, the data generated for systemic viral injection in a marmoset and for local injection in macaques come from a unique animal and was not repeated.
Randomization	No condition tested in this study required control versus test groups and thus no randomization was relevant
Blinding	No condition tested in this study required control versus test groups, as such blinding was not relevant to the study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	chicken anti-GFP at 1:1,000 (Abcam USA, ab13970); rabbit anti-DsRed at 1:1,000 (Clontech USA 632496); goat anti-PV at 1:1,000 (Swant USA, PVG-213); guinea-pig anti-PV at 1:2,000 (Swant USA, GP-72); rabbit anti-SST at 1:2,000 (Peninsula USA, T-4103.0050); mouse anti-Synaptotagmin-2 at 1:250 (ZFIN USA, #ZDB-ATB-081002-25)
Validation	All antibodies used in this study are commercially available and have been validated by the manufacturer. In addition to the validation statement that can be found by consulting the manufacturer's website using the references provided below, the specificity of each primary antibody used in these study was validated for the species for which it was used based on examining the signal intensity, the density of staining and the consistency with morphological features of the cellular populations and are presented in the relevant panels across the figures and supplementary figures of the manuscript. chicken anti-GFP at 1:1,000 (Abcam USA, ab13970); rabbit anti-DsRed at 1:1,000 (Clontech USA 632496); goat anti-PV at 1:1,000 (Swant USA, PVG-213); guinea-pig anti-PV at 1:2,000 (Swant USA, GP-72); rabbit anti-SST at 1:2,000 (Peninsula USA, T-4103.0050); mouse anti-Synaptotagmin-2 at 1:250 (ZFIN USA, #ZDB-ATB-081002-25)

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mice. Female C57BL/6J mice (Mus musculus; 10 weeks old) were obtained from Jackson Labs (Bar Harbor, ME - stock# 000664). Male hemizygous Dlx6a-cre mice (Mus musculus; 10 weeks old - Jax stock #008199) and female homozygous INTACT mice (Mus musculus; 10 weeks old - flox-Sun1-eGFP, Jax stock #021039). Mice were maintained at macroenvironmental temperature and humidity ranges of 64 to 79 °F (17.8 to 26.1 °C) and 30% to 70%, respectively. These parameters were monitored closely and controlled within rodent colony rooms. Rat. Sprague Dawley rats (Rattus norvegicus, 12 weeks old 150-250g) were obtained from Charles River labs, Kingston, NY. Marmosets. One female common marmoset (Callithrix jacchus, 6.0 years old) was obtained from the colony at Massachusetts Institute of Technology. Macaques. Adult (2 years old) male macaques (Macaca mulatta) were obtained from the California National Primate Research Center at the University of California, Davis. All the animals were maintained in a 12 light/12 dark cycle with a maximum of five animals per cage for mice and one animal per cage for rats at . Marmosets and macaques were socially housed. All animal maintenance and experimental procedures were performed according to the guidelines established by the Institutional Animal Care and Use Committee at the Broad Institute of MIT and Harvard (mice), McGovern research institute at MIT (rats and marmosets) and Salk Institute for Biological studies (macaques).
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve samples collected from the field.
Ethics oversight	All animal maintenance and experimental procedures were performed according to the guidelines established by the Institutional Animal Care and Use Committee at the Broad Institute of MIT and Harvard (mice), McGovern research institute at MIT (rats and marmosets) and Salk Institute for Biological studies (macaques).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Four participants (2 male / 2 female; age range 22-57 years) underwent a surgical procedure in which brain tissue (temporal lobe and hippocampus) was resected for the treatment of drug resistant epilepsy. In all cases, each participant had previously undergone an initial surgery for placement of subdural and/or depth electrodes for intracranial monitoring in order to identify the location of seizure onset.
Recruitment	Patients were selected based on their need for treatment of drug resistant epilepsy by resection surgery. Any self-selection bias that might have occurred is unlikely to have consequence on the results presented in this study: the brain tissue used for the study was collected at the margins of the epileptic focus and exposed to adeno-associated virus ex-vivo and no parameters directly relevant to seizures needed to be recorded in the context of this study. Notably, the results obtained were consistent across males and females participants and across the age range representative of adult mature cortical tissue.
Ethics oversight	The NINDS Institutional Review Board (IRB) approved the research protocol (ClinicalTrials.gov Identifier NCT01273129), and we obtained informed consent from the participants for experimental use of the ressected tissue.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NCT01273129
Study protocol	https://clinicaltrials.gov/ct2/show/NCT01273129
Data collection	Study Type : Observational Estimated Enrollment : 300 participants LLocation: Observational Model: Cohort Time Perspective: Prospective Official Title: Surgery as a Treatment for Medically Intractable Epilepsy Actual Study Start Date : December 7, 2010 United States, Maryland National Institutes of Health Clinical Center, 9000 Rockville Pike Recruiting Bethesda, Maryland, United States, 20892 Contact: For more information at the NIH Clinical Center contact Office of Patient Recruitment (OPR) 800-411-1222 ext TTY8664111010 prpl@cc.nih.gov Sponsors and Collaborators National Institute of Neurological Disorders and Stroke (NINDS)
Outcomes	Primary Outcome Measures : Change in seizure frequency [Time Frame: Baseline and 1 year] Change in seizure frequency, as measured by the Engel scale before and 1 year after treatment. Secondary Outcome Measures : 1. Neurophysiological correlates of human cognitive function and to provide invasive monitoring for patients with tumor related epilepsy [Time Frame: Baseline and 1 Year] The proportion of patients who are able to completely withdrawn from anti-epileptic medication (measured 2years after surgery; subjects will remain on antiepileptic medications for one year after surgery, and may be withdrawn from antiepileptic medications during the second year after surgery). 2. Neurophysiological correlates of human cognitive function and to provide invasive monitoring for patients with tumor related epilepsy [Time Frame: Baseline and 1 Year] The proportion of patients who are seizure-free (Engel Class I) one year after surgery. 3. Neurophysiological correlates of human cognitive function and to provide invasive monitoring for patients with tumor related epilepsy [Time Frame: Baseline and 1 Year] Mean Engel Class one year after surgery stratified by type of surgical procedure performed. 4. Neurophysiological correlates of human cognitive function and to provide invasive monitoring for patients with tumor related epilepsy [Time Frame: Baseline and 1 Year] Permanent neurological deficits, assessed one year after surgery. 5. Neurophysiological correlates of human cognitive function and to provide invasive monitoring for patients with tumor related epilepsy [Time Frame: Baseline and 1 Year] Neurophysiologic correlates of cognitive function and seizures 6. Neurophysiological correlates of human cognitive function and to provide invasive monitoring for patients with tumor related epilepsy [Time Frame: Baseline and 1 year] Outcomes for subjects with tumor related epilepsy will be assessed under a separate protocol, 16-N-0041, Tumor Related Epilepsy