The marked diversity of unique cortical enhancers enables neuron-specific tools by Enhancer-Driven Gene Expression

Authors: Stefan Blankvoort[1], Menno P. Witter[1], James Noonan[2,3], Justin Cotney[2,4,*], Cliff Kentros[1,5,6,*]

Author affiliations

1. Kavli Institute for Systems Neuroscience and Centre for Neural Computation, NTNU, Norway;
2. Department of Genetics, Yale School of Medicine, New Haven, CT, USA;
3. Kavli Institute for Neuroscience, Yale University, New Haven, CT, USA;
4. Department of Genetics and Genome Sciences, University of Connecticut Health Centre, Farmington, CT, USA;
5. Institute of Neuroscience, University of Oregon, Eugene, OR, USA;
6. Lead Contact.

* Corresponding authors: clifford.kentros@ntnu.no and cotney@uchc.edu

Highlights

- Enhancer ChIP-seq of cortical subregions reveals 59372 putative enhancers.
- 3740 of these are specific to particular cortical subregions.
- This reflects the remarkable anatomical diversity of the adult cortex.
- Unique enhancers provide a means to make targeted cell-type specific genetic tools.

SUMMARY

Understanding neural circuit function requires individually addressing their component parts: specific neuronal cell types. However, not only do the precise genetic mechanisms specifying neuronal cell types remain obscure, access to these neuronal cell types by transgenic techniques also remains elusive. While most genes are expressed in the brain, the vast majority are expressed in many different kinds of neurons, suggesting that promoters alone are not sufficiently specific to distinguish cell types. However, there are orders of magnitude more distal genetic cis-regulatory elements controlling transcription (i.e. enhancers), so we screened for enhancer activity in microdissected samples of mouse cortical subregions. This identified thousands of novel putative enhancers, many unique to particular cortical subregions. Pronuclear injection of expression constructs containing such region-specific enhancers resulted in transgenic lines driving expression in distinct sets of cells specifically in the targeted cortical subregions, even though the parent gene's promoter was relatively nonspecific. These data showcase the promise of utilizing the genetic mechanisms underlying the specification of diverse neuronal cell types for the development of genetic tools potentially capable of targeting any neuronal circuit of interest, an approach we call Enhancer-Driven Gene Expression (EDGE).

Keywords

Neural circuits, transgenic animals, transgene expression, transgenic methods, transgenics, enhancers, epigenetics, transcriptional control, entorhinal cortex.

41 # INTRODUCTION

42 The mammalian brain is arguably the most complex biological structure known, composed of around 10^{11} neurons in
43 humans[1]. While this number is generally accepted, the same is not true for how many different *kinds* of neurons exist.
44 Indeed, there is not even a clear consensus as to how to define a neuronal cell type: by morphology, connectivity, gene
45 expression, receptive field type, or some combination of the above? If one takes the expansive view (i.e. all of the
46 above), the numbers quickly become astronomical. For example, current estimates of retinal cell types range between
47 100 and 150[2], and dozens of cell types have been proposed for a single hypothalamic area based solely on which
48 genes are expressed[3]. However, gene expression alone is a poor basis for defining cell types, because although most
49 genes are expressed in the adult brain, the vast majority of them are expressed in many different cell types[4].
50 Identification of neuronal cell types is much more than an issue of taxonomy, it is crucial to understanding brain
51 function. The past two decades have seen the development of revolutionary molecular tools which allow one to
52 determine the precise connectivity of neurons[5, 6] as well as manipulate[7-9] and observe[10] their activity. Yet, the
53 utility of these powerful tools is often limited by the inability to deliver them at the level of particular neuronal cell
54 types. Almost all existing neuron-specific lines are either made by non-homologous recombination of minimal promoter
55 constructs[5, 11, 12] or knocking the transgene into the native transcript[13, 14], made much easier by the advent of
56 CRISPR-Cas. Both of these techniques depend upon the specificity of a native promoter, which can recapitulate the
57 expression of the native gene, but that gene will almost always be expressed in multiple neuronal cell types[15-17].

58 Still, there must be *some* genetic basis for the remarkable diversity of neuronal cell types. Investigations of eukaryotic
59 transcriptional regulation have revealed that spatiotemporally precise gene expression is achieved by the modular and
60 combinatorial action of a variety of trans–acting factors (i.e. DNA-binding proteins) interacting with cis-regulatory
61 elements (i.e. regions of noncoding DNA termed enhancers)[18]. While the exact number of enhancers remains
62 unknown, estimates run into the millions[19, 20]. This is many times the number of genes or promoters, suggesting that
63 the same gene is expressed in distinct cell types via the activation of different sets of enhancers. Enhancers may
64 therefore enable the generation of molecular genetic tools that are more specific than what is possible using promoter-
65 based methods. Indeed, many of the most specific neuronal driver lines are likely the result of random integration next
66 to a highly specific enhancer[12, 21]. Fortunately, investigators studying the mechanisms of transcription have
67 developed a variety of techniques enabling the identification of the enhancers active in any tissue sample[22-26]. We
68 reasoned that because different cell types are found in different brain regions, enhancers active only in particular brain

69 regions could enable the generation of region- and/or cell type-specific molecular genetic tools, an approach that we

70 call Enhancer Driven Gene Expression or EDGE (Figure 1).

71 # RESULTS

72 ## Enhancer ChIP-seq of cortical subregions reveals a striking diversity of unique enhancers

73 Because promoter based techniques generally lead to gene expression throughout the telencephalon, we specifically
74 targeted closely related subregions of cortex in the hopes of obtaining regionally specific tools. The following brain
75 regions from two adult (P56) male C57BL6J mice were microdissected (for details see methods and Figure S1): the
76 medial entorhinal cortex (MEC), the lateral entorhinal cortex (LEC), the retrosplenial cortex (RSC), and the anterior
77 cingulate cortex (ACC). Each mouse was processed separately and the samples were used as biological replicates for
78 further analysis. We performed ChIP-seq on homogenized tissue against the active-enhancer-associated histone
79 modifications H3K27ac and H3K4me2 for samples of each of the four brain regions. The regions enriched for H3K27ac
80 reproducibly identified similar numbers of active promoters and distal cis-regulatory sequences between two replicates
81 of each brain subregion (Figure 2A). Nearly 90% of all active promoters were identified in at least two subregions with
82 the remainder being active in only one subregion (17032 total, 2045 unique).

83 When we analyzed more distal sites (>5kb from a transcriptional start site) we identified a total of 59372 reproducibly
84 active enhancers in at least one subregion. Of these 31% were only identified in a single cortical subregion (18185
85 unique relative to other subregions). Surprisingly the number of subregion specific enhancers in the cortex was similar
86 to the number of total enhancers active in any single tissue in the body thus far investigated[20, 27]. Furthermore, 81%
87 (48077) of enhancers identified in these subregions were not identified in bulk cortex tissue, presumably due to signal-
88 to-noise ratios. The fact that so many novel and unique enhancers were isolated from a tiny minority of cortical regions
89 demonstrates the potentially vast repertoire of enhancers active in the brain.

90 Interestingly, when comparing the total number of reproducible peak calls in these 4 cortical subregions (59372) to the
91 number identified in bulk cortex treated in the same way (13472), the number of putative active enhancers one obtains
92 from the four cortical subregions is far greater than what one obtains from the entire cortex, even though these four
93 cortical regions compose only a small minority of the entire cortex. Of course, this is comparing 4 pooled samples to a
94 single sample, but each of the individual samples gives numbers similar to bulk cortex (Figure 2). In our view, the most
95 likely explanation for this superficially puzzling result is a reduction in signal to noise ratio when pooling heterogeneous
96 sets of tissues for ChIP-seq. This would tend to favor those enhancers that are expressed throughout many cortical
97 subregions at the expense of more specific ones. In support of this, 89% of cortical enhancers were found in one or

98 more cortical subregions, and 78% were found in at least 2 cortical subregions. Compare this to the fact that fully 31%
99 of the enhancers we found in our subregions were specific to that single subregion.

100 While many of these enhancers identified by peak calls alone are specific to this small number of cortical subregions
101 the goal of this study was to identify very specific regulatory sequences with limited activity within other regions of the
102 brain as well as the rest of the body. To ensure the identification of such sequences and exclude regions with weak
103 activity elsewhere we expanded our comparisons to include a variety of published mouse adult tissues and cultured cell
104 types[27]. We first identified active putative enhancers in these additional mouse samples and merged them to create
105 a unified set of enhancers for consistent comparisons across all samples. We then extracted normalized H3K27ac counts
106 at 108299 discrete regions from the subregions profiled in this study as well as those from 17 mouse ENCODE
107 samples[27]. Hierarchical clustering of samples revealed two main groups of mouse tissues: neuronal and non-neuronal
108 (Figure 2B). Amongst non-neuronal tissues, the strongest correlations were observed in developmental stages of heart
109 and tissues that make up the immune system: bone marrow, thymus, and spleen. In neuronal tissues the four cortical
110 subregions profiled here were well correlated across all enhancers assayed but clustered distinctly from cerebellum,
111 olfactory bulb, and embryonic brain.

112 We then utilized k-means clustering to identify enhancers that were significantly more active in each cortical region
113 versus each other (Figure 2C) and the other 17 mouse tissues. Those enhancers that were identified as most specifically
114 active in a given cortical subregion were then further filtered to ensure that they were never identified by peak calling
115 in any other mouse tissue. Even though this does not exclude identification of enhancers as unique while they are
116 actually highly enriched, it does increase the specificity and thus the chances of identifying unique enhancers. This
117 stringent analysis yielded 165 to 1824 novel and unique putative distal enhancers for each cortical subregion (Figure
118 2C, Data S1). We then assigned these novel enhancers to putative target genes based upon the GREAT algorithm[28].
119 Gene ontology analysis suggests these novel enhancers are enriched near genes associated with a variety of neuronal
120 functions (Figure S2). We prioritized these novel putative enhancers based on specificity of the H3K27ac signal relative
121 to other regions and conservation across 30 species. We then cloned a subset of them specific to the entorhinal cortices
122 (EC) upstream of a heterologous minimal promoter driving the tetracycline transactivator (tTA[29]) for transgenesis
123 (Figure S3).

Region-specific enhancers drive transgene expression in targeted cortical subregions

Of course, just because a sequence is identified by ChIP-seq does not mean that it is a valid enhancer, let alone that it can drive region- or cell type-specific transgene expression. Even a single case of expression in a particular tissue type is not sufficient because one can obtain specific transgene expression by randomly inserting a minimal promoter/reporter construct into the genome. This technique is known as an "enhancer trap" because it relies upon random insertion near a native enhancer to drive the transgene expression[12, 30]. To ensure that the expression pattern comes from the enhancer construct and not from the insertion site, the standard way to validate a putative enhancer is to show that at least three distinct transgenic embryos (with three distinct random insertion sites) have similar expression patterns[26]. We therefore injected enough oocytes to get at least three genotypically-positive founders for each putative enhancer construct. But, since our aim was to generate modular genetic tools rather than simply to validate the enhancers, we could not sacrifice the founders to validate the enhancer as is typically done. Instead, the founders were crossed to tTA dependent reporter mice for visualization of expression patterns.

We selected 8 (notionally) MEC-specific and 2 LEC-specific enhancers for transgenesis. Transgenesis via pronuclear injection is not an efficient process because it involves random integration into the genome. While one typically only publishes the ones that work, it is worth specifying what issues have to do with transgenesis in general versus using EDGE. When making any transgenic line, some founders do not successfully transmit the transgene to offspring, while others fail to express presumably due to negative insertional (a.k.a. positional) effects. For these reasons, only 45 lines derived from 105 genotypically-positive founders expressed in the brain when mated to a tetO reporter line, a number that is typical regardless of the injection construct. Notably, nearly all of them (41) expressed the reporter in the EC, including at least one from each of the 10 enhancer constructs (Figure S4 and Table S1). Since an enhancer trap would lead to random expression patterns, this alone suggests that the specificity of expression comes from the transgenic enhancer. At least as compelling is the fact that when we obtained multiple distinct founders with a given enhancer construct, almost all of them had similar expression patterns (see Figure S5 for examples).

Figure 3A shows an example of the results of our bioinformatic analysis for one of the eight MEC enhancers (MEC-13-81, see methods for nomenclature) which GREAT associated with the gene *Kitl*. Note that the promoter region (vertical yellow band) is a strong peak in all brain regions, consistent with expression of the *Kitl* mRNA throughout the brain (Figure 3B). The same is true for other putative enhancers (horizontal black bars). In contrast, the downstream enhancer

151 peak used for transgenesis (MEC-13-81, Figure 3A blow-up), while not as strong as some of the other peak calls, is
152 greatly enriched in MEC. Figure 3C shows the result of crossing the transgenic line MEC-13-81B to an tetO-ArChT
153 payload line[31]. Remarkably, even though the *Kitl* promoter expresses throughout the brain (including multiple layers
154 of the EC, Figure 3B), the tetO-ArChT payload is confined to layer II of MEC (Figure 3C). In other words, one can obtain
155 highly specific targeted gene expression from regionally specific cis-elements of non-specific genes.

156 The same basic result of highly specific expression from single enhancers of non-specific genes was also true for 4/8
157 MEC- and 2/2 LEC-specific enhancer constructs we injected, although the correspondence between the ChIP-seq signal
158 and the expression was not always as tight. Figure 4 compares the expression patterns of representative transgenic
159 driver lines made with other injection constructs containing either MEC-specific enhancers (Figure 4A to 4C, right
160 column) or LEC-specific enhancers (Figure 4D and E, right panel. Extended medial-lateral of sections range in Figure S4A-
161 F) compared to the expression pattern of the presumed associated native gene (Figure 4 left column). Note that while
162 each associated gene is broadly expressed in the brain, the transgenic lines all express more or less specifically in the
163 brain region the enhancers were isolated from. When crossed with broadly expressing tTA lines (such as CaMKIIa [11]),
164 these tetO payload lines express broadly (see references for published lines and Figure S6 for our as-of-yet unpublished
165 tetO-GCaMP6 line). These data show that one can obtain targeted region-specific (and possibly even cell type-specific)
166 expression from elements of a non-specific promoter by using one of its region-specific enhancer to drive a
167 heterologous core promoter. Even those enhancers that were less specific still gave rise to lines that were enriched in
168 the EC relative to the expression of the native gene (Figure S4G-J). This in effect solves the problem that most genes
169 are expressed in multiple cell types in the brain: using EDGE one can dissect out the individual genetic components
170 which underlie the expression of a "nonspecific" gene in multiple cell types.

171 Region versus cell type-specific expression?

172 The above results show that subregion specific expression can result from subregion specific enhancers. Whether such
173 enhancers drive expression in specific cell types in the targeted brain region is a more difficult question to answer, in
174 large part because there is no consensus as to the number of cell types in the brain or how to classify them. However,
175 there are indications that some these enhancers can specify particular cell types, at least to the level of granularity
176 current knowledge permits. First, the different EC enhancers tend to drive expression in different layers of the EC (Figure
177 3 and 4 and S7), and neurons in different cortical layers are almost by definition different cell types. By the same logic,

178 some of these enhancers are clearly not cell type-specific (Figure S4G-J). Since four of the enhancers drive expression
179 in layer II, this raises the question of whether they specify the same cell type, or distinct biological subpopulations. We
180 therefore investigated the expression of immunohistochemical markers used to characterize cell types of EC in two layer
181 II expressing lines derived from MEC-specific enhancers (Figures 5A, B and 6A, B). The underlying logic is that if the two
182 distinct enhancers drive transgene expression in subsets of the exact same cell type(s), they should both express the
183 same proportions of neurochemical markers. Neither of the two enhancers appear to drive expression in inhibitory
184 neurons (Figure 5I-L and 6I-L), so the question becomes whether they express in different types of excitatory neurons.
185 Excitatory neurons in EC layer II are typically further subdivided into reelin positive cells and calbindin positive cells[32].
186 Line MEC-13-53A expressed exclusively in reelin+ neurons (Figure 5C-H, L), while line MEC-13-104B roughly corresponds
187 to the relative densities of the two celltypes (Figure 6C-H, L). Thus it appears that MEC-13-53A is a stellate cell specific
188 enhancer, whereas MEC-13-104B is found in both neurochemical kinds of excitatory cells of layer II described to date.
189 This means that some enhancers specify different subsets of cells even within a single cortical layer, showing the
190 potential of enhancers to distinguish between cell types with a finer granularity than possible with native promoters. Of
191 course, the functional significance, if any, of these subsets of cells remains to be demonstrated.

192 # DISCUSSION

193 We demonstrate the existence of thousands of previously undescribed putative enhancers uniquely active in targeted
194 cortical subregions of the adult mouse brain. We took a small subset (10/3740) of the enhancers that were specific to
195 the EC and combined them with a heterologous minimal promoter to make transgenic mice expressing the tTA
196 transactivator. When crossed to tetO payload lines, we obtained transgene expression specific to the EC, and possibly
197 even particular cell types in the targeted region. The genes that these enhancers (presumably) act upon are nowhere
198 near that specific. Most genes express in multiple cell types in the brain. Since there are only around 24.000 genes (and
199 around 46.000 promoters[19]), but estimated millions of putative enhancers[19, 20], this implies that the same gene is
200 expressed in different cell types by using different sets of enhancers acting upon the same promoter. In turn, this
201 suggests that there may be a genetic diversity in the brain beyond most estimates of the number of distinct neuronal
202 cell types in the cortex[33-36]. Moreover, this provides a strategy to make genetic tools with far greater cell type and
203 regional specificity of expression than promoter-based methods, by far the dominant means to generate neuron specific
204 transgenic animals to date.

205 ## EDGE is a method to create neuron-specific tools for targeted brain regions

206 While the above discussion illustrates the power of this technique, it is important to be clear about what is and is not
207 novel about what has been presented here. A variety of forms of enhancer ChIP-seq have existed for roughly a
208 decade[22, 26], and the general concept that the same gene is expressed in different tissues by the use of different
209 enhancers is even older[30]. Hundreds of thousands of putative enhancers have already been identified in the mouse
210 genome by dissection of distinct tissues (including cortex) followed by ChIP-seq[20, 27]. Indeed, a molecular geneticist
211 in the transcription field may find the results presented here unsurprising, as generation of a transgenic animal is how
212 putative enhancers are biologically verified, although the transgenic founders are typically killed in the process[22, 25,
213 26, 37]. In short, we have not invented any novel techniques, but we demonstrate how the application of these existing
214 technologies to the adult brain could potentially provide systems neuroscientists with a means to make cell type-specific
215 tools for any brain region of interest, greatly facilitating the study of the functional circuitry of the adult brain.

216 Putting our results into context requires discussing the rich literature that inspired our approach. A variety of recent
217 papers used various techniques to suggest a highly diverse chromatin landscape in the adult brain, indicative of a

218 diversity of enhancers. One group has performed ChIP-seq on 136 different dissected human brain regions, obtaining
219 over 80.000 putative enhancers[38]. Another group has used ATAC-seq to profile open chromatin in transgenically-
220 defined excitatory cells from different layers of the mouse visual cortex[39]. They found a diversity of putative cis-acting
221 sequences even within single layers of a single type of cortex, implying distinct classes of cells. Finally, using single cell
222 methylomes, Luo et al. have shown that neuron type classification is supported by the epigenomic state of regulatory
223 sequences[40]. Nonetheless, in none of these cases were these putative enhancers biologically verified, nor used to
224 make molecular genetic tools, which is the point of this paper.

225 Conversely, many enhancers derived from the developing brain have in fact been biologically verified, and even used to
226 make transgenic lines and viruses[41]. Evolutionarily conserved single enhancers demonstrably label specific subsets of
227 cells during development[25, 26, 37, 42], with different subsets active in different developmental epochs[43]. Of
228 particular interest is a pair of papers from the Rubenstein lab examining the activity of enhancers derived from the
229 developing (E11.5) telencephalon. They made CreER lines from the pallium (14 lines[44]) and subpallium (10 lines[45])
230 to illustrate the fatemaps of the telencephalic subdivisions by comparing expression patterns at several timepoints
231 during development and young adulthood. By examining in vivo transcription factor occupancy they showed that
232 broadly expressed transcription factors interact with far more specific enhancer elements[44].

233 Taken together, all these studies provide part of the basis for what is presented here. However, their focus was on the
234 transcriptional and developmental mechanisms of neural cell fate relatively early in development. As these and other
235 studies demonstrated, every neuroepithelial cell present at this time will have many daughter cells which will further
236 differentiate during development into many more neuronal and non-neuronal (e.g. glia) cell types[46, 47]. Presumably,
237 for this reason, these enhancers show relatively broad expression in the adult brain [45]. Subpallial enhancers as
238 expected tended to drive expression in GABAergic cells, but do not distinguish between the various known subtypes of
239 GABAergic interneurons[41, 48]. Therefore, although these tools are valuable to the elucidation of cell lineages, they
240 are not necessarily more specific than promoter-based transgenic lines[15], which as noted earlier are not always
241 specific enough for the analysis of native neural circuits.

242 Thus, a seemingly trivial difference in technique results in a large increase in utility for systems neuroscience. Applying
243 the same methods discussed above to microdissected adult cortical subregions allows one to make molecular genetic
244 tools apparently specific to particular cell types of the targeted brain regions. The microdissection is not in fact a trivial

245 feature: by examining four subregions of the cortex separately, we found around four times as many reproducible peak
246 calls as was obtained from the entire cortex[27], even though these four subregions together comprise a small minority
247 of the cortex. This implies that individual cortical subregions contain their own epigenetically distinct cell types, which
248 are washed out when pooled. Similarly, there is relatively little overlap between the enhancers active in the embryonic
249 brain and those we have obtained from adult brain (Figure 2). Hence, it would be interesting to work backwards and
250 study the developmental expression of EDGE lines made from subdivisions of the adult brain to investigate the genetic
251 signatures of the pre- and postnatal processes that specify the enormous variety of neuronal cell types present in the
252 fully differentiated adult brain. In sum, we do not claim to have discovered anything novel about transcription in the
253 brain, although the sheer number of novel putative enhancers unique to particular cortical subregions was surprising,
254 nor the methods described herein. What we claim is both novel and significant is the application of these methods to
255 the generation of anatomically-specific tools enabling the study of the circuit dynamics of the adult brain[49].

256 It is worth mentioning that EDGE is not the same as enhancer traps. In enhancer traps[12, 30], one randomly inserts a
257 minimal promoter construct into the genome in the hopes of integrating near a specific enhancer while EDGE involves
258 the identification and use of enhancers specific to particular brain regions. The key advantage of EDGE over enhancer
259 traps is anatomical targeting. To illustrate, we can compare our results to those of a recently published enhancer trap
260 study[12] using a lentiviral vector containing the exact same minimal promoter we used. Since we are interested in the
261 EC, we consider the creation of EC specific lines the goal, as in the current study. The total number of genotypically
262 positive founders that express in the brain are similar (45/105: 43% herein vs. 42/151: 28%), and both techniques can
263 yield very specific expression patterns. The key difference is the numbers of lines expressing in the EC at all (41/45: 91%
264 herein vs. 6/42: 14%) and especially those more or less specifically expressing in the EC (16/45: 36% vs. 0/42: 0%). This
265 illustrates the difference between the two approaches: enhancer traps result in expression in random cell types
266 throughout the brain (and indeed the entire body), while EDGE targets those cell types found in particular brain regions
267 of interest.

268 Of course, not everyone is interested in the EC. While we subtracted out any enhancers which expressed anywhere but
269 the MEC (or LEC), other investigators interested in other brain regions can use the same strategy to develop tools
270 specifically targeting *their* brain regions of interest. This process can occur for any brain regions, potentially providing
271 cell type-specific tools to interrogate any neural circuit. Moreover, the more subdivisions of the brain one collects, the

272 more one can subtract, so therefore the more specific the resulting putative enhancers will be. With this in mind we
273 have initiated a second round of enhancer ChIP-seq with over 20 brain subregions which will provide a much more
274 generally useful resource. Finally, the relatively small size of the enhancers means they can fit easily in viral vectors. If
275 EDGE viruses recapitulate the anatomical specificity seen in transgenic mice, this will potentially bring EDGE to bear on
276 any species[41]. This could revolutionize not only systems neuroscience, but ultimately provide a novel therapeutic
277 avenue to rectify the circuit imbalances that underlie disorders of the central nervous system.

278 Do enhancers specify neuronal cell types in the brain?

279 One of the most interesting questions in neuroscience is how we should think about the 100 or so billion neurons in our
280 brains- as unique actors, or as repeated elements in a printed circuit? The answer is likely in between. Several
281 investigators have proposed a canonical circuit for the neocortex[50, 51] , with regional variations, and there are clearly
282 commonalities in neocortical circuits, particularly with regards to layer-specific connectivity. Yet, within this general
283 canonical theme there are uniquely specialized cell types in individual cortical subregions. Our results demonstrate that
284 there are thousands of putative enhancers unique to cortical subregions, a number far larger than the number of genes
285 that are specific to these subregions (indeed to our knowledge there are no EC specific genes). Why do the same genes
286 use different enhancers to express in different cortical subregions? There is not yet enough data for a satisfactory
287 answer, but the developmental literature discussed above would suggest a combinatorial code of transcription factors
288 and active enhancers for each unique cell fate. If so, enhancer usage could provide a finer grained differentiation of cell
289 type than gene expression alone. The fact that there are hundreds to thousands of unique enhancers in individual
290 cortical subregions means that the genetic machinery exists to have a similar number of differentiable cell types. In
291 support of this, a recent study of the transcriptome of thousands of individually sequenced neurons from two different
292 cortical regions finds a large number of distinct transcriptional profiles between excitatory, but not inhibitory
293 neurons[52]. This (as well as the fact that inhibitory neurons are a small minority of cortical neurons) may explain why
294 we only obtained expression in excitatory neurons when we selected region-specific enhancers.

295 EDGE allows the generation of tools that provide a means to investigate the nature of neuronal cell types. For example,
296 three of the enhancer constructs presented here drive expression in layer II of MEC, two of which (MEC-13-53 and MEC-
297 13-81) exclusively in reelin-positive neurons (Figure 5 and data not shown for MEC-13-81). MEC LII reelin-positive
298 neurons are stellate cells, which is arguably a cell type, but neither line expresses in 100% of reelin-positive neurons.

There are two possible explanations for this is the first one being that these distinct enhancers drive expression in functionally distinct subsets of stellate cells[52, 53]. The other, possibility is that each enhancer drives expression in stellate cells as part of a co-regulated network of enhancers specifying this cell type[38]. If so, the difference in percentage of expression in stellate cells is largely artefactual, resulting from differential penetrance of transgene expression of otherwise identical cells due to mosaicism arising from insertional effects. The exhaustive biochemical, anatomical and electrophysiological characterization of each line necessary to provide a definitive answer to the relationship between these enhancers and cell types is beyond the scope of this paper. However, the fact that there are so many enhancers unique to specific cortical subregions implies the potential for a more direct connection between the molecular identity of a cell and cell type in other terms. Moreover, it is entirely possible that further subdivisions of the cells specified by these transgenic lines could provide even more specific expression. This could be achieved in a variety of ways, for example by finer manual microdissection, laser capture microscopy or even nested ChIP-seq of transgenically-labeled cells isolated by a cell sorter from microdissected tissue.

Regardless, our results certainly do not suggest that every enhancer defines a distinct cell type, in fact several of our lines express in more than one layer. There is not necessarily a one-to-one correspondence between cell types and enhancers: a single cell type could be specified by multiple unique enhancers, i.e. a co-regulated enhancer network[38]. Conversely, different cell types may arise from distinct combinatorial codes of active enhancers, meaning the number of different cell types may conceivably be even larger than the number of unique enhancers. Finally, there are other reasons for differential sets of active enhancers beyond definition of cell type: neural activity changes the chromatin landscape[54]. This means that activity of differential enhancers does not automatically imply different cell types but changes in function of a given cell. Nevertheless, differential enhancer utilization does signify distinct epigenetic signatures, even if their functional significance is currently unclear. We therefore maintain that a powerful way to investigate the relationship of diverse cis-acting elements of the genome to the functional circuitry of the brain is to create and study enhancer-specific tools like those presented here.

Acknowledgements

This work was supported by the FRIPRO ToppForsk grant Enhanced Transgenics (90096000) of the Research Council of Norway, the Kavli Foundation, the Centre of Excellence scheme of the Research Council of Norway – Centre for Biology of Memory and Centre for Neural Computation, The Egil and Pauline Braathen and Fred Kavli Centre for Cortical

Microcircuits and the National Infrastructure scheme of the Research Council of Norway – NORBRAIN. We would like to thank Haiyan Wu and Qiangwei Zhang for their help with *In Situ hybridization*, Ute Hostick of the transgenic mouse facility in Eugene, Oregon for her help with pronuclear injection and Hanne Mali Møllergård for her help with genotyping.

Author contribution

Conceptualization: J.C. and C.K. ; Software: J.C. ; Validation: S.B. ; Methodology, Formal analysis, Data curation, and Visualization: J.C. and S.B. ; Investigation and Resources: S.B., M.P.W. and J.C. ; Writing – original draft: S.B., J.C. and C.K. ; Writing – review and editing: S.B., M.P.W., J.C. and C.K. ; Supervision: J.C., J.N. and C.K. ; Project administration: J.N. and C.K. ; Funding acquisition: C.K.

Declaration of Interests

C.K., J.C. and S.B. are inventors on US Patent Application No. 62/584,282, Appl. Norwegian University of Science and Technology (NTNU), which is related to this work. The authors have no other competing interests to declare.

338 1. von Bartheld, C.S., Bahney, J., and Herculano-Houzel, S. (2016). The search for true numbers of neurons and
339 glial cells in the human brain: A review of 150 years of cell counting. The Journal of comparative neurology
340 *524*, 3865-3895.
341 2. Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path
342 forward. Nat Rev Neurosci *18*, 530-546.
343 3. Bonnavion, P., Mickelsen, L.E., Fujita, A., de Lecea, L., and Jackson, A.C. (2016). Hubs and spokes of the lateral
344 hypothalamus: cell types, circuits and behaviour. J Physiol *594*, 6443-6462.
345 4. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway,
346 K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. Nature *445*,
347 168-176.
348 5. Feng, G., Mellor, R.H., Bernstein, M., Keller-Peck, C., Nguyen, Q.T., Wallace, M., Nerbonne, J.M., Lichtman,
349 J.W., and Sanes, J.R. (2000). Imaging neuronal subsets in transgenic mice expressing multiple spectral variants
350 of GFP. Neuron *28*, 41-51.
351 6. Wickersham, I.R., Lyon, D.C., Barnard, R.J., Mori, T., Finke, S., Conzelmann, K.K., Young, J.A., and Callaway, E.M.
352 (2007). Monosynaptic restriction of transsynaptic tracing from single, genetically targeted neurons. Neuron
353 *53*, 639-647.
354 7. Boyden, E.S., Zhang, F., Bamberg, E., Nagel, G., and Deisseroth, K. (2005). Millisecond-timescale, genetically
355 targeted optical control of neural activity. Nature neuroscience *8*, 1263-1268.
356 8. Alexander, G.M., Rogan, S.C., Abbas, A.I., Armbruster, B.N., Pei, Y., Allen, J.A., Nonneman, R.J., Hartmann, J.,
357 Moy, S.S., Nicolelis, M.A., et al. (2009). Remote control of neuronal activity in transgenic mice expressing
358 evolved G protein-coupled receptors. Neuron *63*, 27-39.
359 9. Magnus, C.J., Lee, P.H., Atasoy, D., Su, H.H., Looger, L.L., and Sternson, S.M. (2011). Chemical and genetic
360 engineering of selective ion channel-ligand interactions. Science (New York, N.Y.) *333*, 1292-1296.
361 10. Mank, M., Santos, A.F., Direnberger, S., Mrsic-Flogel, T.D., Hofer, S.B., Stein, V., Hendel, T., Reiff, D.F., Levelt,
362 C., Borst, A., et al. (2008). A genetically encoded calcium indicator for chronic in vivo two-photon imaging.
363 Nature methods *5*, 805-811.
364 11. Mayford, M., Bach, M.E., Huang, Y.Y., Wang, L., Hawkins, R.D., and Kandel, E.R. (1996). Control of memory
365 formation through regulated expression of a CaMKII transgene. Science (New York, N.Y.) *274*, 1678-1683.
366 12. Shima, Y., Sugino, K., Hempel, C.M., Shima, M., Taneja, P., Bullis, J.B., Mehta, S., Lois, C., and Nelson, S.B.
367 (2016). A Mammalian enhancer trap resource for discovering and manipulating neuronal cell types. Elife *5*,
368 e13503.
369 13. Capecchi, M.R. (1989). Altering the genome by homologous recombination. Science (New York, N.Y.) *244*,
370 1288-1292.
371 14. Heintz, N. (2001). Bac to the future: The use of bac transgenic mice for neuroscience research. Nat Rev
372 Neurosci *2*, 861-870.
373 15. Huang, Z.J. (2014). Toward a genetic dissection of cortical circuits in the mouse. Neuron *83*, 1284-1302.
374 16. Madisen, L., Garner, A.R., Shimaoka, D., Chuong, A.S., Klapoetke, N.C., Li, L., van der Bourg, A., Niino, Y., Egolf,
375 L., Monetti, C., et al. (2015). Transgenic mice for intersectional targeting of neural sensors and effectors with
376 high specificity and performance. Neuron *85*, 942-958.
377 17. Luo, L., Callaway, E.M., and Svoboda, K. (2008). Genetic dissection of neural circuits. Neuron *57*, 634-660.
378 18. Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40
379 DNA sequences. Cell *27*, 299-308.
380 19. ENCODE, P.C. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
381 20. Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A.,
382 Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes.
383 Nature *518*, 317-330.
384 21. Yasuda, M., and Mayford, M.R. (2006). CaMKII activation in the entorhinal cortex disrupts previously encoded
385 spatial memory. Neuron *50*, 309-318.
386 22. Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C.,
387 Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and
388 enhancers in the human genome. Nature genetics *39*, 311-318.
389 23. Visel, A., Rubin, E.M., and Pennacchio, L.A. (2009). Genomic views of distant-acting enhancers. Nature *461*,
390 199-205.

24. Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R.V., McKinsey, G.L., Pattabiraman, K., Silberberg, S.N., Blow, M.J., et al. (2013). A high-resolution enhancer atlas of the developing telencephalon. Cell *152*, 895-908.
25. Cotney, J., Leng, J., Yin, J., Reilly, S.K., DeMare, L.E., Emera, D., Ayoub, A.E., Rakic, P., and Noonan, J.P. (2013). The evolution of lineage-specific regulatory activities in the human embryonic limb. Cell *154*, 185-196.
26. Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.
27. Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.
28. McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology *28*, 495-501.
29. Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proceedings of the National Academy of Sciences of the United States of America *89*, 5547-5551.
30. Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development *118*, 401-415.
31. Weible, A.P., Moore, A.K., Liu, C., DeBlander, L., Wu, H., Kentros, C., and Wehr, M. (2014). Perceptual gap detection is mediated by gap termination responses in auditory cortex. Current biology : CB *24*, 1447-1455.
32. Witter, M.P., Doan, T.P., Jacobsen, B., Nilssen, E.S., and Ohara, S. (2017). Architecture of the Entorhinal Cortex A Review of Entorhinal Anatomy in Rodents with Some Comparative Notes. Front Syst Neurosci *11*, 46.
33. DeFelipe, J., Lopez-Cruz, P.L., Benavides-Piccione, R., Bielza, C., Larranaga, P., Anderson, S., Burkhalter, A., Cauli, B., Fairen, A., Feldmeyer, D., et al. (2013). New insights into the classification and nomenclature of cortical GABAergic interneurons. Nat Rev Neurosci *14*, 202-216.
34. Taniguchi, H. (2014). Genetic dissection of GABAergic neural circuits in mouse neocortex. Frontiers in cellular neuroscience *8*, 8.
35. Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature neuroscience *19*, 335-346.
36. Briggs, F., and Callaway, E.M. (2001). Layer-specific input to distinct cell types in layer 6 of monkey primary visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *21*, 3600-3608.
37. Reilly, S.K., Yin, J., Ayoub, A.E., Emera, D., Leng, J., Cotney, J., Sarro, R., Rakic, P., and Noonan, J.P. (2015). Evolutionary genomics. Evolutionary changes in promoter and enhancer activity during human corticogenesis. Science (New York, N.Y.) *347*, 1155-1159.
38. Vermunt, M.W., Reinink, P., Korving, J., de Bruijn, E., Creyghton, P.M., Basak, O., Geeven, G., Toonen, P.W., Lansu, N., Meunier, C., et al. (2014). Large-scale identification of coregulated enhancer networks in the adult human brain. Cell Rep *9*, 767-779.
39. Gray, L.T., Yao, Z., Nguyen, T.N., Kim, T.K., Zeng, H., and Tasic, B. (2017). Layer-specific chromatin accessibility landscapes reveal regulatory networks in adult mouse visual cortex. Elife *6*.
40. Luo, C., Keown, C.L., Kurihara, L., Zhou, J., He, Y., Li, J., Castanon, R., Lucero, J., Nery, J.R., Sandoval, J.P., et al. (2017). Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. Science (New York, N.Y.) *357*, 600-604.
41. Dimidschstein, J., Chen, Q., Tremblay, R., Rogers, S.L., Saldi, G.A., Guo, L., Xu, Q., Liu, R., Lu, C., Chu, J., et al. (2016). A viral strategy for targeting and manipulating interneurons across vertebrate species. Nature neuroscience *19*, 1743-1749.
42. Prabhakar, S., Visel, A., Akiyama, J.A., Shoukry, M., Lewis, K.D., Holt, A., Plajzer-Frick, I., Morrison, H., Fitzpatrick, D.R., Afzal, V., et al. (2008). Human-specific gain of function in a developmental enhancer. Science (New York, N.Y.) *321*, 1346-1350.
43. Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. Cell *155*, 1521-1531.
44. Pattabiraman, K., Golonzhka, O., Lindtner, S., Nord, A.S., Taher, L., Hoch, R., Silberberg, S.N., Zhang, D., Chen, B., Zeng, H., et al. (2014). Transcriptional regulation of enhancers active in protodomains of the developing cerebral cortex. Neuron *82*, 989-1003.

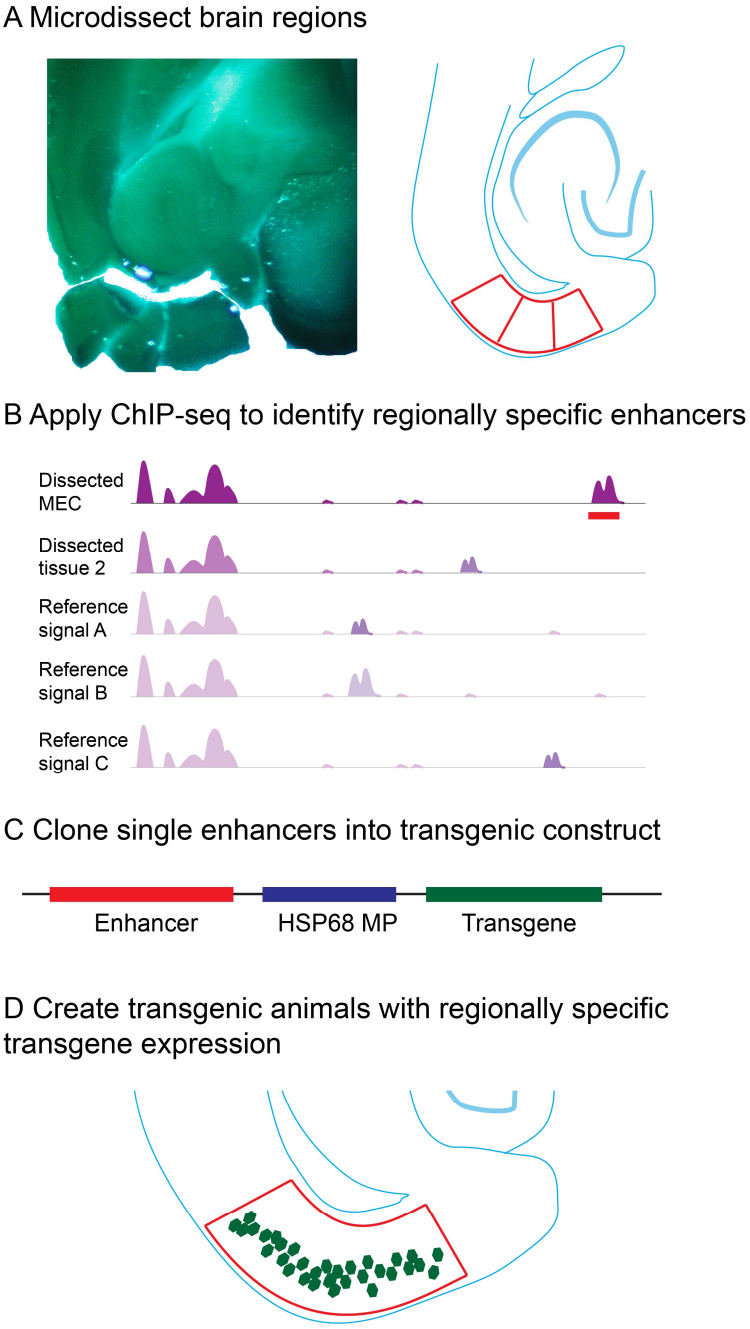
45. Silberberg, S.N., Taher, L., Lindtner, S., Sandberg, M., Nord, A.S., Vogt, D., McKinsey, G.L., Hoch, R., Pattabiraman, K., Zhang, D., et al. (2016). Subpallial Enhancer Transgenic Lines: a Data and Tool Resource to Study Transcriptional Regulation of GABAergic Cell Fate. Neuron *92*, 59-74.
46. Merkle, F.T., and Alvarez-Buylla, A. (2006). Neural stem cells in mammalian development. Curr Opin Cell Biol *18*, 704-709.
47. Fogarty, M., Grist, M., Gelman, D., Marin, O., Pachnis, V., and Kessaris, N. (2007). Spatial genetic patterning of the embryonic neuroepithelium generates GABAergic interneuron diversity in the adult cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *27*, 10935-10946.
48. Zerucha, T., Stuhmer, T., Hatch, G., Park, B.K., Long, Q., Yu, G., Gambarotta, A., Schultz, J.R., Rubenstein, J.L., and Ekker, M. (2000). A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. The Journal of neuroscience : the official journal of the Society for Neuroscience *20*, 709-721.
49. Kanter, B.R., Lykken, C.M., Avesar, D., Weible, A., Dickinson, J., Dunn, B., Borgesius, N.Z., Roudi, Y., and Kentros, C.G. (2017). A Novel Mechanism for the Grid-to-Place Cell Transformation Revealed by Transgenic Depolarization of Medial Entorhinal Cortex Layer II. Neuron *93*, 1480-1492 e1486.
50. Harris, K.D., and Shepherd, G.M. (2015). The neocortical circuit: themes and variations. Nature neuroscience *18*, 170-181.
51. Douglas, R.J., and Martin, K.A. (2007). Recurrent neuronal circuits in the neocortex. Current biology : CB *17*, R496-500.
52. Tasic, B., Yao, Z., Smith, K.A., Graybuck, L.T., T.N., N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2017). Shared and distinct transcriptomic cell types across neocortical areas. bioRxiv.
53. Fuchs, E.C., Neitz, A., Pinna, R., Melzer, S., Caputi, A., and Monyer, H. (2016). Local and Distant Input Controlling Excitation in Layer II of the Medial Entorhinal Cortex. Neuron *89*, 194-208.
54. Malik, A.N., Vierbuchen, T., Hemberg, M., Rubin, A.A., Ling, E., Couch, C.H., Stroud, H., Spiegel, I., Farh, K.K.-H., Harmin, D.A., et al. (2014). Genome-wide identification and characterization of functional neuronal activity-dependent enhancers. Nature neuroscience *17*, 1330-1339.
55. Boccara, C.N., Kjonigsen, L.J., Hammer, I.M., Bjaalie, J.G., Leergaard, T.B., and Witter, M.P. (2015). A three-plane architectonic atlas of the rat hippocampal region. Hippocampus *25*, 838-857.
56. Sugar, J., and Witter, M.P. (2016). Postnatal development of retrosplenial projections to the parahippocampal region of the rat. Elife *5*.
57. O'Reilly, K.C., Flatberg, A., Islam, S., Olsen, L.C., Kruge, I.U., and Witter, M.P. (2015). Identification of dorsal-ventral hippocampal differentiation in neonatal rats. Brain Struct Funct *220*, 2873-2893.
58. Jones, B.F., and Witter, M.P. (2007). Cingulate cortex projections to the parahippocampal region and hippocampal formation in the rat. Hippocampus *17*, 957-976.
59. Witter, M.P. (2011). The hippocampus. In The Mouse Nervous System, 1st Edition, G. Paxinos, L. Puilles and C. Watson, eds. (Academic Press), pp. pp 112-139.
60. Cotney, J.L., and Noonan, J.P. (2015). Chromatin immunoprecipitation with fixed animal tissues and preparation for high-throughput sequencing. Cold Spring Harb Protoc *2015*, 419.
61. Weible, A.P., Schwarcz, L., Wickersham, I.R., Deblander, L., Wu, H., Callaway, E.M., Seung, H.S., and Kentros, C.G. (2010). Transgenic targeting of recombinant rabies virus reveals monosynaptic connectivity of specific neurons. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 16509-16513.
62. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.
63. Mikkelsen, T.S., Xu, Z., Zhang, X., Wang, L., Gimble, J.M., Lander, E.S., and Rosen, E.D. (2010). Comparative epigenomic analysis of murine and human adipogenesis. Cell *143*, 156-169.
64. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.
65. Habegger, L., Sboner, A., Gianoulis, T.A., Rozowsky, J., Agarwal, A., Snyder, M., and Gerstein, M. (2011). RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. Bioinformatics *27*, 281-283.
66. de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.
67. Saldanha, A.J. (2004). Java Treeview--extensible visualization of microarray data. Bioinformatics *20*, 3246-3248.

498 68. Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier,
499 L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast
500 genomes. Genome research *15*, 1034-1050.

501

502

503 FIGURES



504

505 **Figure 1. Experimental summary of Enhancer Driven Gene Expression (EDGE).**

506 (A) Samples of brain regions of interest are microdissected by hand. (B) ChIP-seq is performed on these samples and
507 genome-wide H3K27ac and H3K4me2 signals for each sample are compared to reference signals and signals from the
508 other samples. Bioinformatic analysis algorithms output unique peaks as potential region-specific enhancers (red bar).
509 (C) Single putative enhancers are cloned into constructs containing a heterologous minimal promoter to drive
510 transgene expression. (D) Following pronuclear injection of these constructs, the resulting founder mice are crossed
511 to reporter lines and evaluated for desired expression patterns. See also Figure S1.

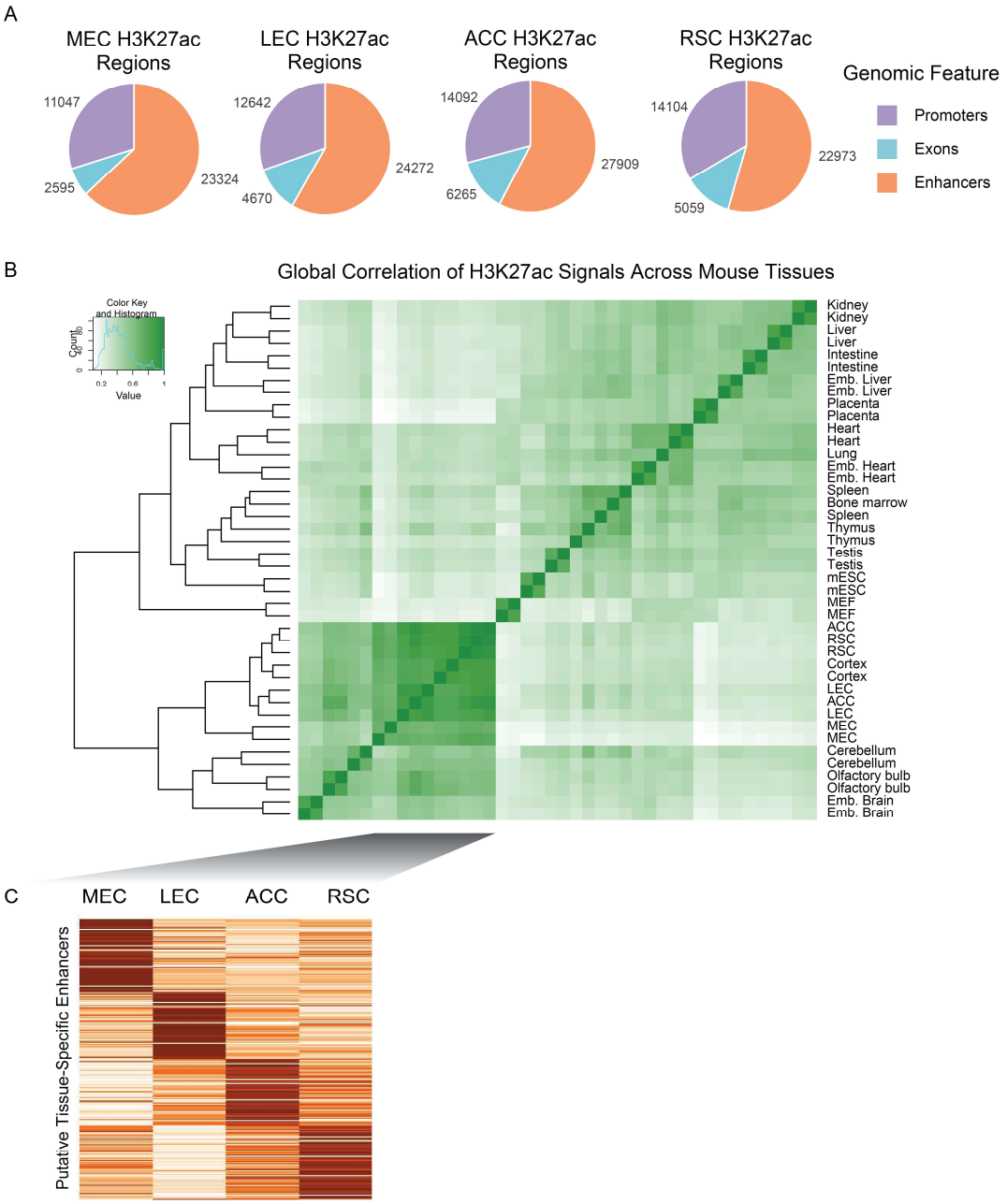
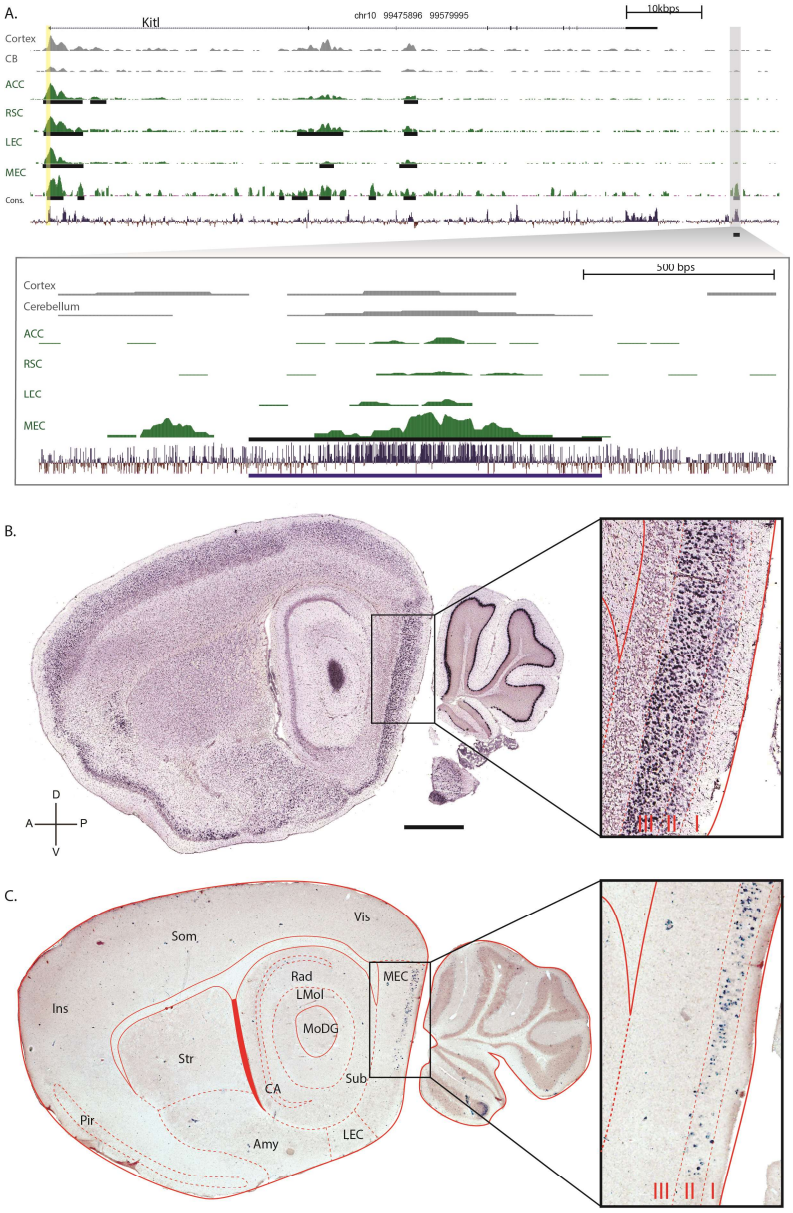


Figure 2. ChIP-seq reveals a striking diversity of unique and novel enhancers in different cortical subregions.

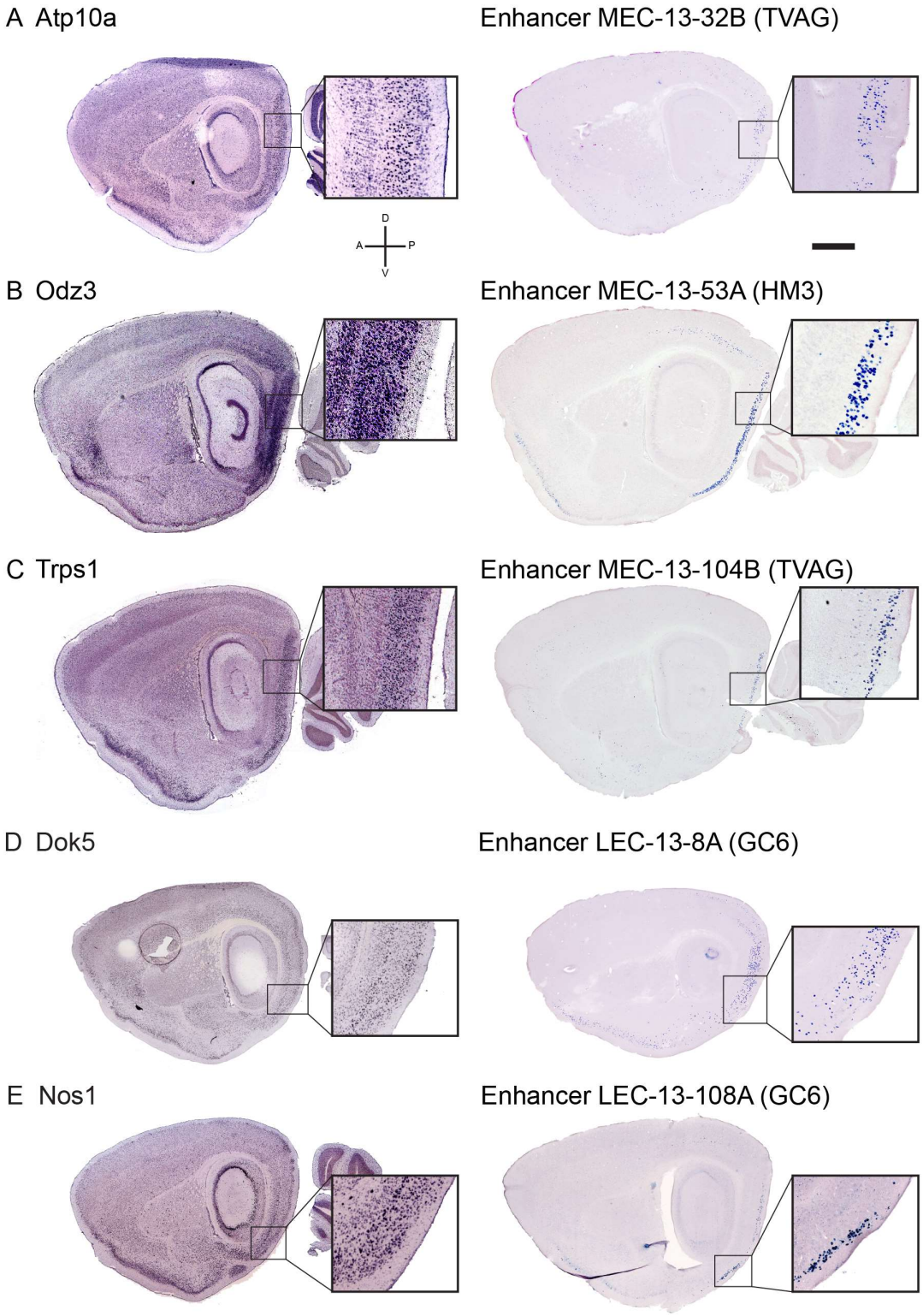
(A) Pie charts showing the proportions (and numbers) of distinct active genomic elements identified by H2K27ac ChIP-seq of the 4 cortical subregions. These numbers are roughly similar to those found by ChIP-seq of other organs. (B) Dendrogram (left) and correlation matrix of the H3K27ac signals (right) from replicates of the cortical subregions dissected in this experiment versus those from ENCODE were used for subtraction. Note the relatively high correlation of replicates (except ACC) and clustering of signal from cortical tissues. (C) Heatmaps showing some of the tissue-specific putative enhancers identified in the microdissected cortical subregions. See also Figure S2 and Data S1.



521

522 **Figure 3. The enhancers of non-specific genes drive region-specific transgene expression.**

523 (A) A genomic view of one of the 165 MEC-specific enhancers yielded by ChIP-seq analysis. The top panel indicates the
524 location and coding regions of *Kitl* as well as H3K27Ac signal for two regions from Roadmap epigenome (Cortex and
525 Cerebellum), the four regions we analyzed (ACC, RSC, LEC and MEC), and conservation over 30 species. The vertical
526 yellow column indicates the promoter region upstream of the transcriptional start site. Peak calls, are denoted by the
527 black horizontal lines. The specific genomic region containing the enhancer (MEC-13-81) is blown up in the bottom
528 panel. (B) ISH (brain-map.org) of *Kitl*, the gene associated with enhancer MEC-13-81 shows expression throughout
529 cortex, hippocampus and cerebellum. (C) tTA dependent transgene Arch driven by the enhancer (ranked number 81)
530 is expressed in MEC LII. Scalebar is 1000µm. Sagittal plane, Dorsal-Ventral and Anterior-Posterior axis are indicated.
531 Abbreviations are: Ins: insular cortex, Som: somatosensory cortex, Vis: visual cortex, Pir: Piriform cortex, Str: striatum,
532 Amy: amygdala and associated regions, Rad: stratum radiatum of the hippocamlus, LMol: molecular layer of the
533 hippocampus, CA: both cornu ammonis fields of the hippocampus, sub: subiculum, MoDG: moclecular layer of the
534 Dentate Gyrus, LEC: lateral entorhinal cortex, MEC: medial entorhinal cortex, Layers I, II and III of the MEC are indicated
535 in the blow-up to the right. See also Figure S3, S4, S5, S6 and Table S1.



536

537 **Figure 4. Distinct MEC-specific enhancers drive transgene expression in distinct sets of cells in MEC.**

538 (A through E, left column) ISH showing expression patterns of native genes associated with EC-specific enhancers. (A
539 through E, right column) ISH showing EC-specific expression of transgenes driven by the corresponding EC-specific
540 enhancers, tTA driven transgenes in parentheses. ISH for the native genes from brain-map.org. Scalebar in A is
541 1000µm. See also Figures S4, S5 and S6 and Table S1.

542

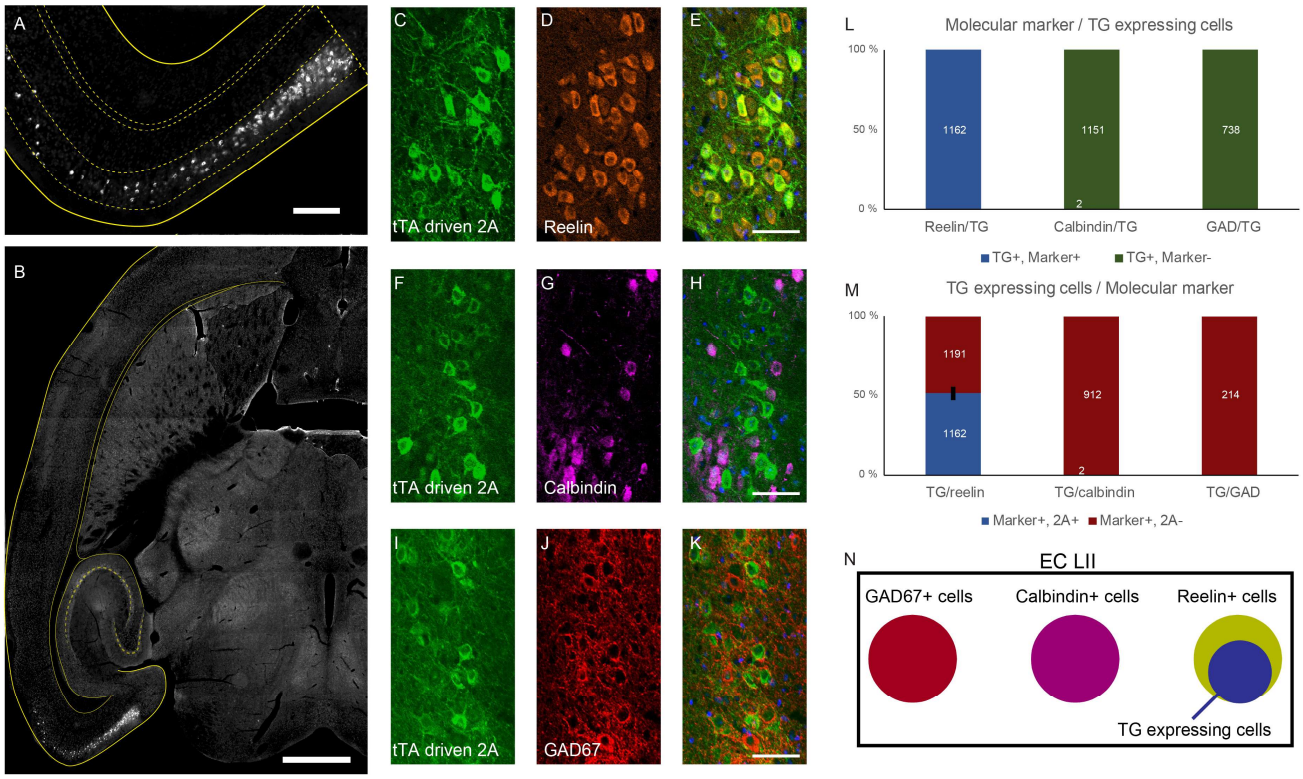


Figure 5. Single enhancers can drive expression in histochemically-defined subsets of MEC LII cells.

(A,B) Horizontal section of a mouse cross between MEC-13-53A and TVAG. Immunohistochemical transgene detection with anti-2A Ab shows layer II EC-specific expression. (C,F,I) Anti-2A histochemistry; (D) Anti-Reelin; (G) Anti-Calbindin; (J) Anti-GAD67; (E,H,K) Overlays of the two signals, each row is the same section. (L) 100% (1162/1162 counted cells) of transgenic cells co-localize with Reelin but there is essentially 0% co-localization with calbindin (2/1151) and GAD67 (0/738). (M) 49.4% (1162/2353) of all Reelin positive cells were positive for the transgene, essentially none of the other cell populations had any transgene expressing cells. Total numbers of cells counted in white. (N) Schematic summary of the data in C to M. Scale bars are 1000µm in B, 200µm in A and 50µm in C-K. In all graphs bars show the mean ±SEM. See also Figure S7.

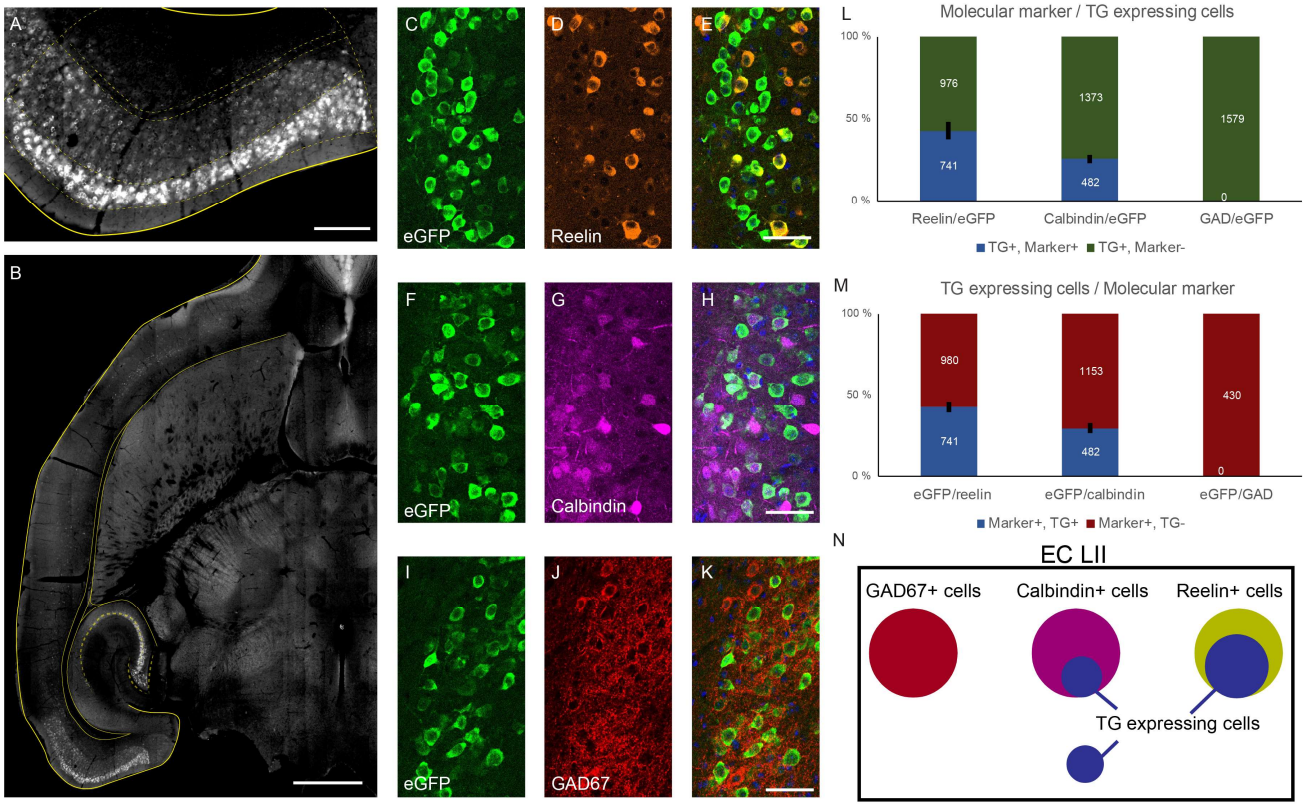


Figure 6. Different single enhancers can drive expression in histochemically-distinct subsets of MEC LII cells.

(A,B) Horizontal section of a mouse cross between MEC-13-104B and tetO-eGFP. Immunohistochemical transgene detection with anti-GFP Ab shows expression in layer II of the EC. (C,F,I) Anti-GFP histochemistry; (D) Anti-Reelin; (G) Anti-Calbindin; (J) Anti-GAD67; (E,H,K) Overlays of the two signals, each row is the same section. (L) 43.1% (741/1717 counted cells) of transgenic cells in layer II of the EC co-localize with Reelin while 26% (482/1855) of them co-localize with calbindin. 0% (0/1579) co-localize with GAD67. (M) 43.1% (741/1721) of all Reelin positive cells in layer II of the EC were positive for the transgene and 28.5% (482/1635) of all Calbindin positive cells in layer II of the EC were positive for the transgene, while 0% (0/430) of the GAD67 positive population had any transgene expressing cells. Total numbers of cells counted in white. (N) Schematic summary of the data in C to M. Scale bars are 1000µm in B, 200µm in A and 50µm in C-K. In all graphs bars show the mean ±SEM. See also Figure S7.

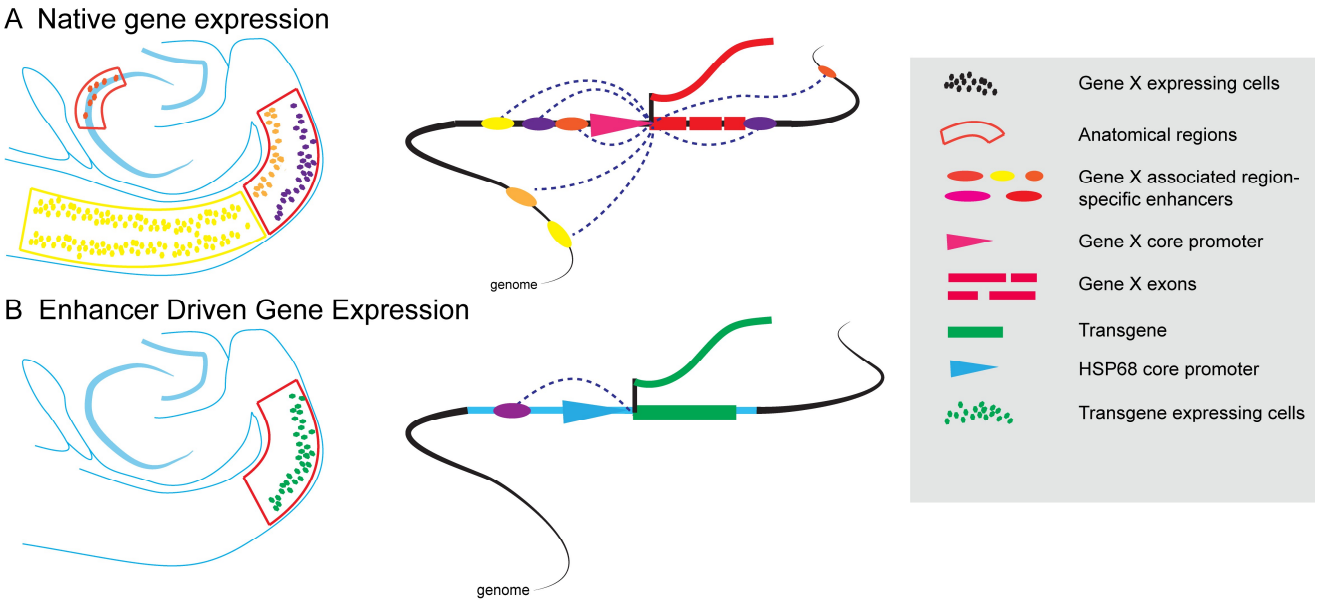


Figure 7. Schematic of putative genetic basis for EDGE technology.

(A) Native gene expression: A gene "X" is expressed in multiple cell types in distinct brain areas. Expression in each cell type is driven by distinct sets of color-coded active enhancers acting upon the native core promoter (pink triangle). Promoter-based methods of transgene expression such as BAC transgenesis and Knock-ins respectively include several or all of the native enhancers, thereby recapitulating some or all of the expression pattern of the native gene. (B) Enhancer-Driven Gene Expression: a single active enhancer isolated from a particular brain region drives transgene expression from a heterologous minimal promoter (blue). This leads to transgene expression that is restricted to a particular region-specific subset of the cell types that the native promoter expresses in, greatly increasing the anatomical specificity relative to promoter-based methods or the native gene.

METHODS

Contact for reagent and resource sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Cliff Kentros (clifford.kentros@ntnu.no).

Experimental model and subject details

The experimental model used in this study is the rodent *M. musculus*. For the microdissection two C57BL/6J mice (Jackson Laboratory, stock no. 000664, P56, one male, one female) were used. For pro-nuclear injection B6D2F1 mice were used. The resulting offspring were backcrossed with C57BL/6J mice and various reporter lines. The mice used for histology were a mix of male and female, taken from several litters when multiple individual mice were investigated, and were at least 6 weeks old.

All mice were kept on a 12-h light/12-h dark schedule in a humidity- and temperature-controlled environment. All experiments in Norway were performed in accordance with the Norwegian Animal Welfare Act and the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes. All experiments involving animals in Oregon (pronuclear injection and husbandry of the resulting animals) were performed in accordance with guidelines approved by University of Oregon's Animal Care and Use Committee and the National Institutes of Health Guide for the Care and Use of Laboratory Animals (National Institutes of Health Publications No. 80-23).

Method details

Microdissection

Two C57BL/6J mice (P56) were deeply anesthetized by injection with pentobarbital (100mg/ml in 96% ethanol, Ås produksjonslab AS). The brains were removed and horizontal or coronal 500 µm sections were cut on a Leica VT 1000 S microtome and kept at 4 °C until dissection. Bilateral dissection was performed, while watching the tissue through a dissection microscope with transmitted and reflected white light (Zeiss Discovery V8 stereomicroscope) applying architectonic criteria [55-59] to unstained tissue. The tissue samples were snap-frozen in liquid nitrogen, kept at -80^{0}C and shipped on dry ice.

All dissections avoided border regions (i.e., centered in the identified cortical area). In horizontal sections, MEC is easily recognized by the marked shape of the cortex, the prominent white, opaque lamina dissecans and the radial organization of the layers deep to the lamina dissecans. Layer II neurons are large spherical neurons, which differ markedly in level of opacity from those in layer III. The medial border between MEC and parasubiculum is characterized by the loss of the differentiation between layers II and III, and the border with the laterally adjacent postrhinal cortex is characterized by the loss of the large spherical neurons in layer II [55-59]. We only sampled the more dorsal and central portions of MEC. LEC shares the large layer II neurons with MEC, but the radial organization in layer V is absent. The anterior and dorsal border of LEC with the perirhinal cortex is characterized by the abrupt disappearance of the large layer II neurons. We only sampled the most lateral portions of LEC, as to avoid contamination with ventromedially adjacent components of the amygdaloid complex. ACC and RSC were sampled from the medial wall of the lateral hemisphere above the corpus callosum, avoiding the most anterior part of ACC and the posteroventral part of RCS. Since the border between the two areas coincides with the dorsal-anterior tip of the hippocampal formation, all samples avoided that border region.

In coronal sections, ACC and RSC samples were taken dorsal to the corpus callosum, just below the shoulder of the medial wall of the hemisphere down to, but not touching the corpus callosum, as to avoid inclusion of the indusium griseum. Samples were taken from sections anterior to the most anterodorsal tip of the hippocampal formation in case of ACC and posterior to the tip in case of RSC. Samples of LEC were collected one section after the disappearance of the piriform cortex characterized by a densely packed thick layer II, a polymorph lightly packed deeper cell layer and the presence of the endopiriform nucleus. LEC shows cytoarchitectonic features similar to those described above. We sampled only from the vertical part of LEC, directly below the rhinal fissure. For MEC, samples were collected from more posterior coronal sections, using shape of the section, the presence of the ventral hippocampus and cytoarchitectonic features as described above as our selection criteria.

624 ## ChIP-seq

625 All dissected brain tissues were briefly homogenized and cross-linked with 1% formaldehyde at room temperature
626 with rotation for 15 min. Cross-linking was quenched with glycine (150mM in PBS), then tissue was washed and flash
627 frozen. Chromatin was extracted as previously described [25, 60]. Briefly, nuclei were extracted, lysed and sonicated
628 (30 min, 10-sec pulses) to produce sheared chromatin with an average length of ~250 bp. 1-10µg of final soluble
629 chromatin was used for each ChIP and combined with Protein G Dynabeads (Invitrogen, cat# 10004D) prebound with
630 5µg of antibodies to H3K4me2 (Abcam ab7766) or H3K27ac (Abcam ab4729). Immunoprecipitated chromatin was
631 washed five times with 1mL of wash buffer and once with TE. Immunoprecipiated chromatin was eluted, cross-links
632 were reversed, and DNA was purified. Libraries were prepared for sequencing using NEBNext ChIP-Seq Library Prep
633 reagents and sequenced on the Illumina HiSeq 2000 platform at the Yale Center for Genome Analysis.

634 ## Cloning of transgenic constructs

635 The putative enhancers sequences were cloned from BACs (chori.org) and transferred to pENTR^{tm}/D-TOPO® vectors
636 by TOPO® cloning (Invitrogen, K2400-20). The putative enhancers were transferred to injection plasmids by gateway
637 cloning® (Invitrogen, 11791-019). The resulting plasmids consist of a putative enhancer followed by a mutated
638 heatshock promoter 68 (HSP68), a tTA gene, a synthetic intron and a WPRE element (Figure S3).

639 ## Pronuclear injection

640 The 10 injection plasmids were linearized by enzyme digestion to keep the relevant elements but remove the bacterial
641 elements of the plasmids. Linearized vectors were run on a 1% agarose gel and isolated using a Zymoclean Gel DNA
642 Recovery Kit (Zymo research, D4001). Fertilized eggcells were injected with 1 µl of DNA at concentrations of 0.5 to 1
643 ng/µl, leading to surviving pups of which 96 were genotypically positive for MEC and 9 were genotypically positive for
644 LEC (Table S1). Pronuclear injections were done at the transgenic mouse facility of the University of Oregon.

645 ## Mouse husbandry

646 Mouse lines were named after the ranked enhancers identified in this study, as specified in Data S1. The nomenclature
647 consists of firstly the targeted region, secondly the year of microdissection, thirdly the rank of the enhancer that
648 corresponds with the row in Data S1 and finally a letter for the founder. To illustrate, line MEC-13-53A is based on
649 MEC tissue isolated in 2013, where the particular enhancer was ranked 53 and the founder is specified by the "A".

650 All genotypically positive founders based on MEC enhancers were initially mated with histone GFP mice (Jackson
651 laboratory, Tg(tetO-HIST1H2BJ/GFP)47Efu, stocknr. 005104), while those based on LEC enhancers were mated with
652 GCaMP6 mice (in house made). Double positive pups were used for further analysis. Subsequent crosses were done
653 with GCaMP6 mice (in house made), TVAG mice (Line TVAG5 from [61]), ArChT mice [31], tetO-eGFP (Jackson
654 laboratory, C57BL/6J-Tg(tetO-EGFP/Rpl10a)5aReij/J_JAX) and hM3 mice [8].

655 ## Genotyping

656 Genotyping was done on ear tissue using a Kapa mouse genotyping kit (Kapa Biosystems, Cat# KK7302). Primer pairs
657 for the appropriate gene and internal controls (Table S2) are added to the PCR mixture at a final concentration of
658 10µM. The PCR reaction was done by an initial step of 4 minutes at 95°C, then 20 cycles of 1 minute at 95°C, 30 seconds
659 at 70°C reduced by 0.5°C each cycle, and 30 seconds at 72°C. This is followed by 20 cycles of 30 seconds at 95°C, 30
660 seconds at 60°C, and 30 seconds at 72°C and a final 7 minute step at 72°C. The products are run on a 1% agarose gel
661 along with positive and negative controls.

662 ## In situ hybridization

663 Double positive mice (tTA+/-, reporter gene+/-) were deeply anesthetized with pentobarbital and transcardially
664 perfused with 0.9% saline first and freshly made 4% formaldehyde (in 1x DPBS, thermofisher, Cat# 14200075) second.
665 At least 2 mice from different litters were investigated on consistent expression patterns. Brains were removed and
666 postfixated overnight in 4% paraformaldehyde. Subsequently the brains were dehydrated for at least 24h with 30%
667 sucrose in 1x PBS. The brains were sectioned sagittally at 30µm on a Cryostat, mounted directly (on Fisherbrand
668 Superfrost Plus microscope slides (Fisher Scientific Cat #12-550-15)) and dried overnight at room temperature. Slides
669 were stored at -80^{0}C.

670 Slides were thawed in closed containers. Sections were outlined with a PAP pen (Sigma, cat# Z377821-1EA). The probe
671 was diluted (usually 0.1-1µgm/ml) in hybridisation buffer (1:10 10x salt solution, 50% deionized formamide (sigma,

672 cat# D-4551), 10% dextran sulfate (sigma, cat# D-8906), 1mg/ml rRNA (sigma, Cat#R5636), 1x Denhardt's (Sigma cat#
673 D-2532). Salt solution (10x) was made with 114g NaCl, 14.04g TrisHCl, 1.3g TrisBase, 7.8g NaH2PO4.2H2O, 7.1g
674 Na2HPO4 in H_2O to 1000ml with a final concentration of 0.5M EDTA). The probe was denatured for 10 min at 62°C,
675 added to the section and coverslipped (Fisher, cat# 12-548-5P). The slides were incubated overnight at 62°C in a closed
676 box with filter paper wetted in 1x SSC with 50% formamide.

677 The slides were transferred to polypropylene Coplin jars containing 1x SSC with 50% formamide and 0.1% Tween-20
678 warmed to 62°C for 10 minutes to allow the coverslips to fall off. The slides were washed 3x30 minutes at 62°C. Then
679 the slides were washed 3x30 minutes in MABT (11.6g Maleic acid (sigma, cat#M0375-1kg), 8.76g NaCl, 5ml 20% tween,
680 pH 7.5, ddH_2O to 1000ml) at room temperature.

681 The slides were drained (not dried) and re-circled with a PAP pen. Then blocking solution was added (600µl MABT,
682 200µl sheep serum, 200µl 10% blocking reagent (Roche cat#11 096 176 001)) and slides were incubated in a Perspex
683 box with wetted filter paper at room temperature for 2-3 hours. The slides were drained and 1:5,000 sheep anti-dig
684 AP in blocking solution was added followed by overnight incubation.

685 4g of polyvinyl alcohol was dissolved into 40ml AP staining buffer (100mM NaCl, 50mM $MgCl_2$, 100mM Tris pH9.5,
686 0.1% Tween-20) by heat and cooled to 37°C. The slides were washed in MABT 5 times for 4 minutes. And subsequently
687 washed 2x10 minutes in AP staining buffer. Nitroblue tetrazolium chloride (Roche, cat# 11 383 213 001. At 3.5 µl/ml),
688 5-Bromo-4-chloro-3-indolyl-phosphate,4-toluidene salt (Roche, cat# 11 383 221 001. At 2.6 µl/ml) and Levamisole
689 (Vector, cat# SP-5000. At 80µl/ml) was added to the cool polyvinyl alcohol solution. This was shaken well and
690 transferred to a Coplin jar. The slides were added to the jar and incubated at 37°C for 3 to 5 hours. The reaction was
691 stopped by washing in 2xPBS with 0.1% Tween-20. The slides were subsequently wash 2X in ddH2O, and dehydrated
692 quickly through graded ethanols from 50%, 70%, 95% to 100% ethanol. Finally the slides were cleared in xylene and
693 coverslipped.

694 Immunohistochemistry

695 Double positive mice (tTA+/-, TVAG+/- or eGFP +/-) were deeply anesthetized with pentobarbital and transcardially
696 perfused with approximately 30ml 0.9% saline first and approximately 30ml freshly made 4% paraformaldehyde (in 1x
697 DPBS, thermofisher, Cat# 14200075) second. Brains were removed and postfixated for 24 hours in 4%
698 paraformaldehyde. Subsequently the brains were dehydrated with 30% sucrose in 1x PBS. The brains were sectioned
699 horizontally at 50µm and kept in TCS (tissue collection solution, 25% glycerol, 35% ethyl glycol, 50% 1xDPBS) at -20°C.

700 Immunohistochemistry was done by two initial 10 minute washes in 1xDPBS and subsequent permeabilized by a 60
701 minute wash in 1% Triton X-100 (Sigma, Cat#T9284) in 1xDPBS. Then the tissue is incubated in primary antibody in
702 1xDPBS with 1% trition X-100 and 5% donkey serum (Sigma, Cat# D9663) for 48 hours at 4°C. Primary antibodies and
703 dilutions were: Rabbit-anti-2A (1:2000, Millipore, cat#ABS31), Mouse-anti-reelin (1:1000, Millipore, cat# Mab5364),
704 Mouse-anti-GAD67 (1:1000, Millipore, cat# Mab5406), Mouse-anti-calbindin (1:10000, Swant, cat# CB300).

705 After incubation with primary antibodies, sections were washed 4x in 1xDPBS (10 minutes per wash) and 2x in 1xDPBS
706 with 1% Triton X-100. Then sections were incubated for 6h at room temperature in secondary antibody (all secondary
707 antibodies were raised in Donkey and diluted 1:250). The secondary antibodies were: anti-Rabbit-AF488 (Jackson
708 ImmunoResearch, Cat# 711-545-152) and anti-Mouse-Cy^{tm}3 (Jackson ImmunoResearch, Cat# 715-165-151)

709 The sections were DAPI stained by a single 10 minute wash in 1xDPBS with 0.2µg/ml DAPI (thermofisher, D1306) and
710 finally washed 5x (10 minutes per wash) in 1x DPBS. Sections were mounted on superfrost® plus glass slides (VWR,
711 Cat# 631-9483) and coverslipped with polyvinyl alcohol with 2.5% DABCO (Sigma, Cat# D27802).

712 Imaging

713 From mice in the lines MEC-13-53A x TVAG and MEC-13-104B x tetO-eGFP MEC was imaged in sections from three
714 different dorsal-ventral levels with a Zeiss Meta 880 confocal microscope. Two channels were imaged: one for AF488
715 with maximum excitation wavelength at 488nm and maximum emission wavelength at 528nm and one for Cy3 with
716 maximum excitation wavelength at 561nm and maximum emission wavelength at 595nm.

717 For display images, sections were imaged on Zeiss Axio.scan Z1 scanners in three preset channels: DAPI, dl488 and
718 dl549.

Image processing

From the Zeiss proprietary file format .lsm, .tiff files were exported. These were processed in Adobe Photoshop and all alterations in levels were made on the entire images. In some cases, images were processed to remove visual artifacts and background.

Quantification and statistical analysis

ChIP-seq data analysis

ChIP-Seq data was initially processed as previously described (Reilly et al 2015). Briefly, reads were aligned to the mm9 version of the mouse genome using bowtie (v1.1.1) [62]. Enriched regions were identified in individual replicates using a sliding window method as previously described [63]. Enriched regions were divided into functional categories based on overlaps with genomic features as annotated by Ensembl v67 using Bedtools (2.19.0) [64]. Reproducibly enriched regions were determined as the union of overlapping regions identified in both biological replicates. Putative enhancer regions from intergenic and intronic portions of the genome were then assigned target genes using GREAT. H3K27ac ChIP-Seq reads were retrieved from Encodeproject.org for 17 mouse tissues [27] and uniformly processed as above. Enhancers for all cell types were combined and merged to generate a uniform annotation of all possible enhancers. H3K27ac counts at each enhancer from each tissue were calculated using mrfQuantifier [65]. Pearson correlations for all enhancer signals were calculated and plotted using R (https://www.r-project.org/). K-means clustering of H3K27ac count matrix was performed using Cluster (v3.0) [66]. Rows were centered on the mean value of the row and normalized, the k parameter was the total number of tissues, and 100 runs were performed. The clustering result was then visualized using Java TreeView [67]. Subregion specific clusters of enhancers were intersected with peak calls from all other tissues to identify enhancers with likely tissue specific function. Subregion specific enhancers were assigned two target genes using GREAT, ranked by H3K27ac signal, and overlapped with vertebrate conserved sequences [68].

Counting

Mice in the lines MEC-13-53A x TVAG and MEC-13-104B x tetO-eGFP MEC was imaged on sections from three different dorsal-ventral levels. For each strain 4 mice from at least 2 litters each were used. For each section, three to seven slices in the Z direction with 1.5µm spacing were taken, with a 20x objective and tiling to cover the entire MEC. Counts were made on the confocal images for single positive cells expressing transgenes, cells expressing native genes (GAD67, Reelin, Calbindin) and cells expressing both. Graphs were made in Microsoft Excel and statistical analysis was done in SPSS (v22, IBM).

Data and software availability

No new software was generated in this study.

Raw data can be accessed on GEO under accession number GSE112897.

Processed data can be viewed here.

753 **Supplemental Information**: Figures S1-S7 and Tables S1 and S2

754 **Data S1:** Ranked regionally specific enhancers for brain regions MEC, LEC, ACC and RSC, related to Figure 2.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
H3K4me2	Abcam	Cat# ab7766, RRID:AB_2560996
H3K27ac	Abcam	Cat# ab4729, RRID:AB_2118291
Rabbit-anti-2A	Millipore	Cat# ABS31, RRID:AB_11214282
Mouse-anti-reelin	Millipore	Cat# Mab5364, RRID:AB_2179313
Mouse-anti-GAD67	Millipore	Cat# Mab5406, RRID:AB_2278725
Mouse-anti-calbindin	Swant	Cat# CB300, RRID:AB_10000347
anti-Rabbit-AF488	Jackson ImmunoResearch	Cat# 711-545-152, RRID:AB_2313584
anti-Mouse-Cy^{tm}3	Jackson ImmunoResearch	Cat# 715-165-151, RRID:AB_2315777
Biological Samples		
Healthy, adult microdissected brain tissue (mouse)	This paper	N/A
Critical Commercial Assays		
NEBNext ChIP-Seq Library Prep	NEB	E6240L
Deposited Data		
Raw data	GEO	GSE112897
Processed data	NCBI session	http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=Jcotney&hgS_otherUserSessionName=Kavli
Experimental Models: Organisms/Strains		
C57BL/6J mice	Jackson Laboratory	Stock no. 000664, RRID: IMSR_JAX:000664
B6D2F1 mice	Jackson Laboratory	Stock no. 100006, RRID:IMSR_JAX:100006
TetO-GCaMP6 mice	This paper	P8966
TetO-histoneGFP mice Tg(tetO-HIST1H2BJ/GFP)47Efu	Jackson laboratory	Stock no. 005104, RRID:IMSR_JAX:005104
TetO-TVAG mice	From [61]	Line TVAG5
TetO-ArChT mice	From [31]	Line tetO-ArchT2
tetO-eGFP mice C57BL/6J-Tg(tetO-EGFP/Rpl10a)5aReij/J_JAX	Jackson laboratory	Stock no. 024898, RRID:IMSR_JAX:024898
TetO-HM3 mice	From [8]	Line TRE-hM3Dq, RRID:IMSR_JAX:014093
MEC-13-32B	This paper	N/A
MEC-13-53A	This paper	N/A
MEC-13-53B	This paper	N/A
MEC-13-53D	This paper	N/A
MEC-13-53G	This paper	N/A
MEC-13-81A	This paper	N/A
MEC-13-104B	This paper	N/A
MEC-13-123A	This paper	N/A

MEC-13-123B	This paper	N/A
MEC-13-123C	This paper	N/A
LEC-13-8B	This paper	N/A
LEC-13-108A	This paper	N/A
Oligonucleotides		
Genotyping primers, Table S2	This paper	N/A
Recombinant DNA		
Injection construct (Gateway destination vector)	This paper	pDest-HSP68-tTA-WPRE
Software and Algorithms		
Bowtie (v1.1.1)	[62]	http://bowtie-bio.sourceforge.net/index.shtml
Bedtools (2.19.0)	[64]	http://bedtools.readthedocs.io/en/latest/
mrfQuantifier	[65]	http://archive.gersteinlab.org/proj/rnaseq/rseqtools/
GREAT	[28]	http://great.stanford.edu/public/html/
R	R-project	https://www.r-project.org/
Cluster (v3.0)	[66]	http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm
Java TreeView	[67]	http://jtreeview.sourceforge.net/
SPSS statistics v. 22	IBM	https://www.ibm.com/products/spss-statistics
Creative Cloud (photoshop, illustrator) 2017 edition	Adobe	https://www.adobe.com/no/creativecloud.html
Excel 2016	Microsoft	www.microsoft.com
Other		
Illumina HiSeq 2000	Illumina	N/A

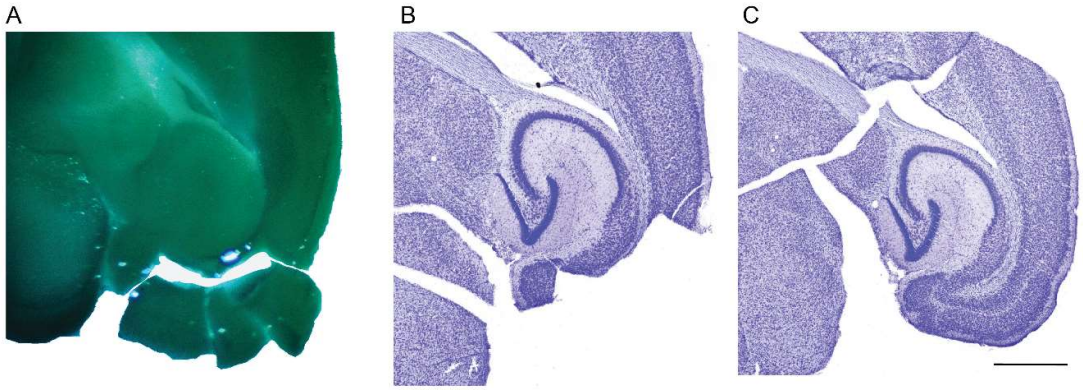


Figure S1. Example of microdissection, related to STAR methods and Figure 1.

(A) 500µm thick section during microdissection. (B) Contralateral side of the same section as in (A), re-sectioned to 50µm and Nissl stained. (C) Re-sectioned (50µm), Nissl stained tissue from (A). Scalebar is 1000µm.

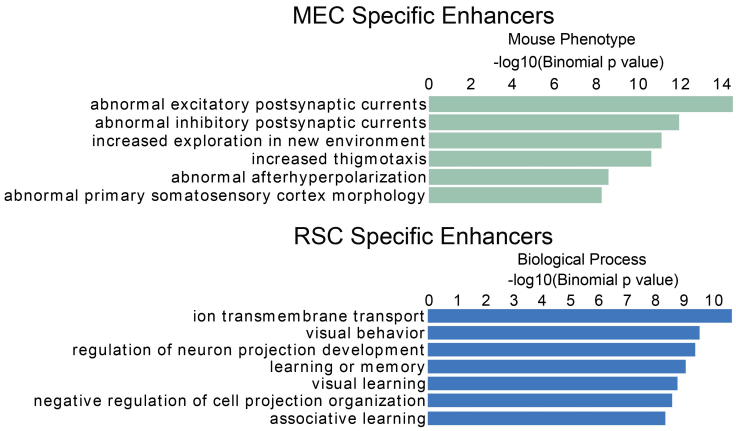


Figure S2. Gene Ontology of selected cortical subregions, related to Figure 2.

Peak calls for each tissue type were used in the GREAT algorithm to assign genes. Ontology terms for these genes were ranked on Binominal P-value.

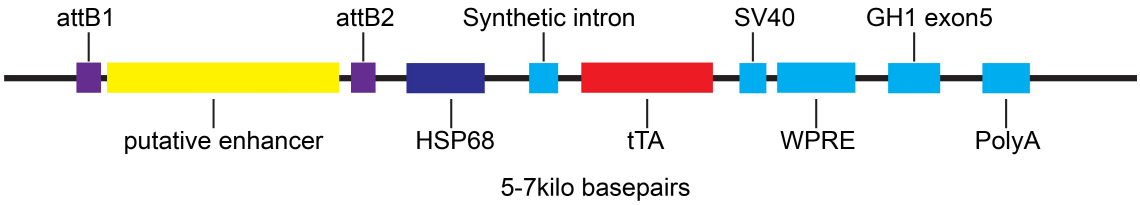


Figure S3. Injection construct, related to Figure 3.

The putative enhancer of 0.7 to 3kbp was cloned to the injection construct by gateway® cloning. The synthetic intron, SV40, WPRE and growth hormone 1 exon 5 are present for optimal mRNA stability and expression of the tetracycline TransActivator (tTA). The construct is linearized with appropriate restriction enzymes depending on exact sequence of the putative enhancer.

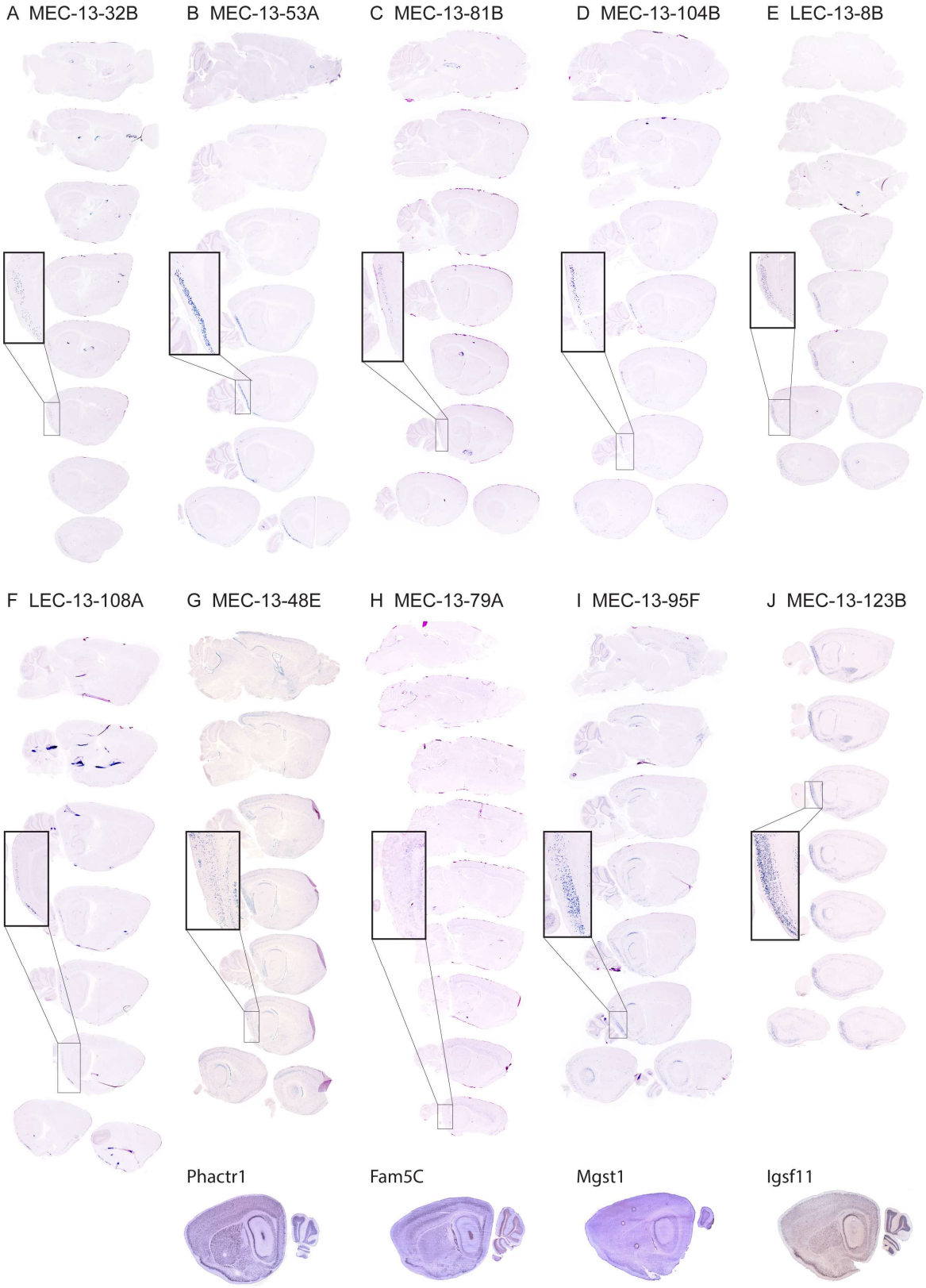


Figure S4. Extended medial lateral coverage of sagittal sections, related to Figure 3 and 4.

(A)-(F) Enhancer lines based 6 different enhancers used (MEC-13-32B, MEC-13-53A, MEC-13-81B, MEC-13-104B, LEC-13-8B, LEC-13-108A) show specific transgene expression in the EC. (G)-(J) Enhancer lines based on 4 different enhancers (MEC-13-48E, MEC-13-79A, MEC-13-95F, MEC-13-123B) show enriched, but not specific transgene expression in the EC. The bottom row shows *in situ hybridization* (taken from brain-map.org) of associated genes.

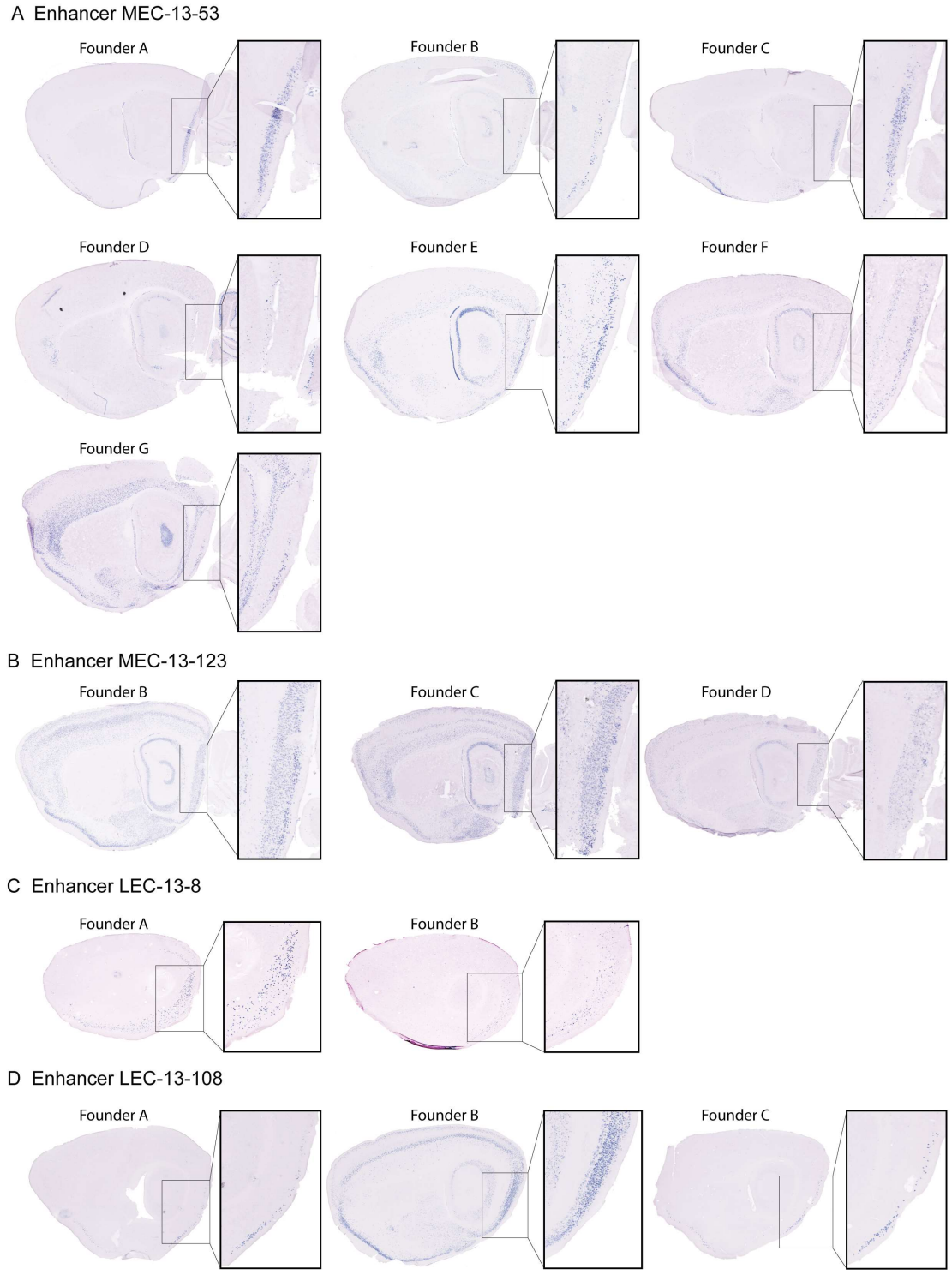


Figure S5. Enhancer driven transgene expression of various genomic insertions, related to Figure 4.

Sagittal sections of approximately similar levels. Different founders based on the same enhancers show roughly similar expression patterns. All mice based on MEC specific enhancers were crosses with hGFP reporter mice, while all mice based on LEC specific enhancers were crossed with GC6 payload mice. (A) Enhancer MEC-13-53 reproducibly shows expression in LII of the EC in 6 of the 7 analyzed mouse lines. We do find expression in other regions, such as visual cortex in founder B, the CA fields of the hippocampus in founder E, and deep layers of cortex in founder G. But since all of these patterns of expression occur only once within the 7 analyzed lines, we consider them to be positional effects. (B) Enhancer MEC-13-123 reproducibly shows expression in LIII of the EC, CA3 and select cortical layers. (C) Enhancer LEC-13-8 reproducibly shows expression in LIII of the EC. (D) Enhancer LEC-13-108 reproducibly shows expression in LII of the EC. We consider the additional expression in the “founder B” line to be another positional effect.

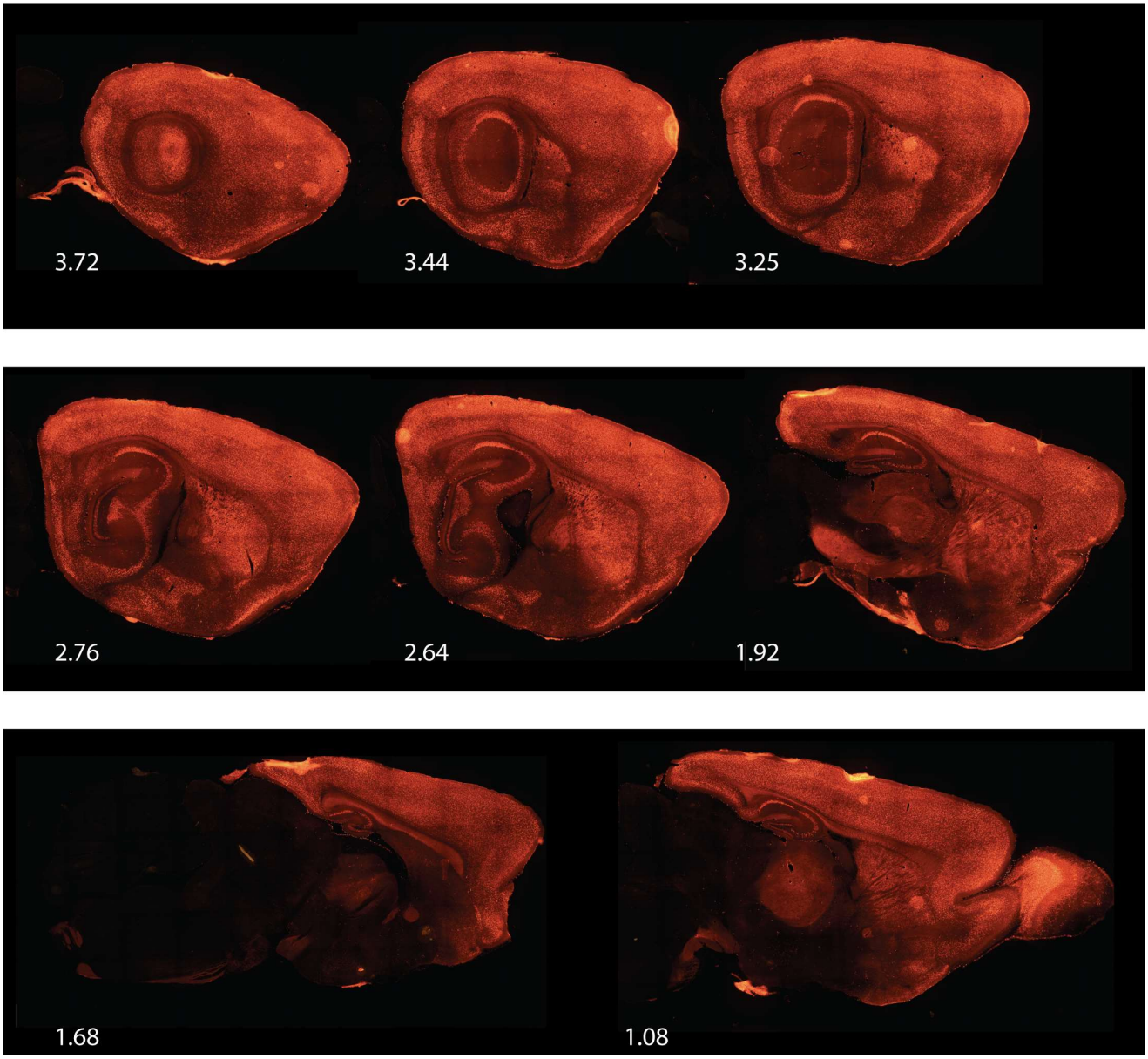


Figure S6. Expression of tTA dependent transgene gCaMP6 in a CaMKIIa driver line, related to main text describing Figures 3 and 4.

Fluorescent signal of the mCherry conjugated to the tTA dependent gCaMP6 in sagittal sections (approximate level in mm lateral to midline indicated in white) show expression throughout the brain.

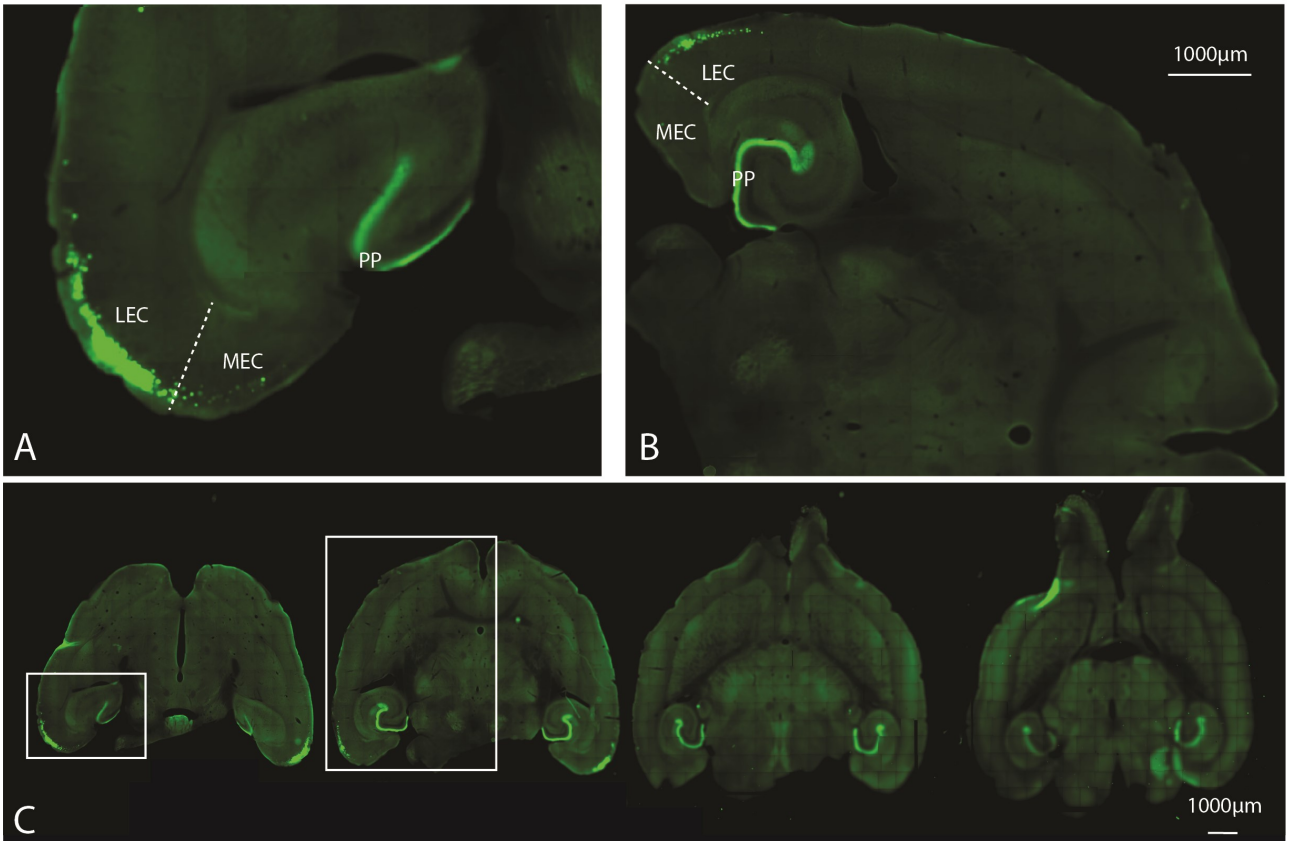


Figure S7. Expression of tTA dependent transgene gCaMP6 in enhancer line LEC-13-108A, related to main text describing Figures 5 and 6.

The labeling of axons in the molecular layer in the dentate gyrus illustrates the transgenic labeling of layer II reelin+ cells.

Genomic coordinates enhancer	chr1:148,339,129-148,340,300	chr16:39,750,789-39,753,454	chr10:99,573,051-99,574,981	chr15:50,913,896-50,916,356	chr7:65,916,198-65,918,580	chr6:138,334,728-138,335,952	chr8:49,906,388-49,908,569	chr13:42,782,503-42,784,035	MEC	chr2:171,158,079-171,159,156	chr5:118,194,653-118,195,333	LEC	Shima et al.
Enhancer rank	79	123	81	104	32	95	53	48		8	108		
Associated genes	Fam5c	Igsf11	Kitl Gm4301	Trps1 Eif3h	Ube3a Atp10a	Lmo3 Mgst1	Odz3	Phactr1 Tbc1d7	total	Dok5 Cbln4	Nos1 Ksr2	total	variable integration
tTA positive founders	6	6	10	12	8	16	18	20	96	5	4	9	151
Lines analyzed	5	4	7	8	4	13	16	16	73	4	3	7	151
GFP signal in EC	2	4	3	3	3	7	7	7	36	2	3	5	6
No GFP signal in brain	2	0	4	4	1	5	9	8	33	2	0	2	109
GFP in brain but no GFP in EC	1	0	0	1	0	1	0	1	4	0	0	0	36

Table S1. Overview of generated transgenic lines, related to Figures 3 and 4.

Genomic coordinates are in mm9 and indicate the exact region used for the enhancer in the transgenic construct, rather than the putative enhancer as indicated by ChIP-seq. Enhancer rank indicates the rank as a result of the ChIP-seq analysis. Associated genes are a result of analysis of GREAT. All indications of expression are based on the initial round of assessment where mice based on MEC specific enhancers were crossed with histone GFP reporter mice and mice based on LEC specific enhancers were crossed with gCaMP6. The numbers in the last column are taken from Shima et al. 2016, table 1 and supplemental table 1 to identify the lines expressing in the EC (53L, 56L, TCAO, TCAR, TCIF, TCLC).

Gene	Primer 1 (5'-3')	Primer 2 (5'-3')	Product size (bp)	Interal control primer 1 (5'-3')	Interal control primer 2 (5'-3')	Product size (bp)
tTA	GGACAAGTCCAAGGTGATCAAC	CCTGGTGGTCGAACAGCTCG	591	CTA GGC CAC AGA ATT GAA AGA TCT	GTA GGT GGA AAT TCT AGC ATC ATC C	324
Histone GFP	TGGGGACGGTGATGCGGTCT	ACGTGGCGAAGCTCTGCTGC	~300	CAA ATG TTG CTT GTC TGG TG	GTC AGT CGA GTG CAC AGT TT	200
TVAG	GTCCGGTAACGGTTCTTTTG	GCTCTTGTCAGGCACCAG	391	CGT CTT TAA TTG GAT TAC AAT GCT	CTA GCA AGT GGT TGT GGT CA	181
Arch	CTTCTCGCTAAGGTGGATCG	CACCAAGACCAGAGCTGTCA	246	CTA GGC CAC AGA ATT GAA AGA TCT	GTA GGT GGA AAT TCT AGC ATC ATC C	324
GCamp6	TGGGGACGGTGATGCGGTCT	ACGTGGCGAAGCTCTGCTGC	~300	CAA ATG TTG CTT GTC TGG TG	GTC AGT CGA GTG CAC AGT TT	200
HM3	ACC GTC AGA TCG CCT GGA GA	TCA TCG GTG GTA CCG TCT GGA G	200	TCC TCA AAG ATG CTC ATT AG	GTA ACT CAC TCA TGC AAA GT	340

Table S2. Primers used for genotyping, related to STAR methods.

All primers were used in a concentration of 10µM. All genes are genotyped individually, ie. tTA x TVAG crosses were genotyped using two separate reactions, one for tTA and one for TVAG.