

1 **Intravenous gene transfer throughout the brain of infant Old World primates using AAV**

2

3 **Miguel R. Chuapoco<sup>a#</sup>, Nicholas C. Flytzanis<sup>a,b,e#</sup>, Nick Goeden<sup>a,b,e#</sup>, J. Christopher**  
4 **Octeau<sup>b</sup>, Kristina M. Roxas<sup>b</sup>, Ken Y. Chan<sup>a,f</sup>, Jon Scherrer<sup>b</sup>, Janet Winchester<sup>b</sup>, Roy J.**  
5 **Blackburn<sup>b</sup>, Lillian J. Campos<sup>c</sup>, Cynthia M. Arokiaraj<sup>a</sup>, Timothy F. Miles<sup>a</sup>, Min J. Jang<sup>a</sup>,**  
6 **Julia Vendemiatti<sup>a</sup>, Benjamin E. Deverman<sup>a,f</sup>, James Pickel<sup>d</sup>, Andrew S. Fox<sup>c</sup>, Viviana**  
7 **Gradinaru<sup>a\*</sup>**

8

9 <sup>a</sup> Division of Biology and Biological Engineering, California Institute of Technology, Pasadena,  
10 CA 91125, USA.

11 <sup>b</sup> Capsida Biotherapeutics, Thousand Oaks, CA 91320, USA

12 <sup>c</sup> Department of Psychology and the California National Primate Research Center, University of  
13 California-Davis, Davis, CA 95616, USA

14 <sup>d</sup> National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892 USA.

15 <sup>e</sup> Present address: Capsida Biotherapeutics, Thousand Oaks, CA 91320, USA

16 <sup>f</sup> Present address: Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard,  
17 Massachusetts Institute of Technology, Cambridge, MA 02142, USA

18

19

20

21 **# Authors contributed equally**

22 **\* Co-corresponding author: [viviana@caltech.edu](mailto:viviana@caltech.edu)**

23 **\* Co-corresponding author: [flytzanis.nicholas@capsida.com](mailto:flytzanis.nicholas@capsida.com)**

## 24 Abstract

25 Adeno-associated viruses (AAVs) can enable robust and safe gene delivery to the mammalian  
26 central nervous system (CNS). While the scientific community has developed numerous  
27 neurotropic AAV variants for systemic gene-transfer to the rodent brain, there are few AAVs that  
28 efficiently access the CNS of higher order primates. We describe here AAV.CAP-Mac, an  
29 engineered AAV variant that enables systemic, brain-wide gene delivery in infants of two Old  
30 World primate species—the rhesus macaque (*Macaca mulatta*) and the green monkey  
31 (*Chlorocebus sabaeus*). We identified CAP-Mac using a multi-species selection strategy, initially  
32 screening our library in the adult common marmoset (*Callithrix jacchus*) and narrowing our pool  
33 of test-variants for another round of selection in infant macaques. In individual characterization,  
34 CAP-Mac robustly transduces human neurons *in vitro* and Old World primate neurons *in vivo*,  
35 where it targets all lobes of cortex, the cerebellum, and multiple subcortical regions of disease  
36 relevance. We use CAP-Mac for Brainbow-like multicolor labeling of macaque neurons throughout  
37 the brain, enabling morphological reconstruction of both medium spiny neurons and cortical  
38 pyramidal cells. Because of its broad distribution throughout the brain and high neuronal efficiency  
39 in infant Old World primates compared to AAV9, CAP-Mac shows promise for researchers and  
40 clinicians alike to unlock novel, noninvasive access to the brain for efficient gene transfer.

## 41 Main Text

42 Adeno-associated viruses (AAVs) have emerged as dependable and ubiquitous tools for  
43 researchers and clinicians alike since their discovery as adenoviral contaminants in the 1960s<sup>1-3</sup>.  
44 In the nearly 4 decades following the earliest descriptions of recombinant AAV vectors<sup>4,5</sup>,  
45 hundreds of clinical trials serve as a testament that AAVs are nonpathogenic and have the  
46 potential to be used safely for long-term expression of genetic payloads<sup>6-9</sup>. There is, however,  
47 renewed concern about the safety of high dose systemic AAV following reports of adverse  
48 hepatotoxicity<sup>10,11</sup> and several patient deaths in the clinic<sup>12,13</sup>. The low therapeutic index of  
49 systemically administering natural AAV serotypes can require high doses to effectively penetrate  
50 tissue, highlighting the need for more efficient, and thus safer, AAVs. In recent years, the field has  
51 focused on engineering novel capsids to address this problem, while simultaneously aiming to  
52 expand the therapeutic opportunity landscape for gene therapy into disorders not previously  
53 approachable with natural AAV serotypes. In parallel, the neuroscience community has also  
54 utilized novel AAVs, as several engineered AAV variants can traverse the restrictive blood-brain-  
55 barrier (BBB) to systemically deliver genetically-encoded tools to the rodent brain<sup>14-17</sup>, such as  
56 GCaMP to detect intracellular calcium gradients<sup>18</sup>. However, the majority of AAV capsid  
57 engineering efforts targeting the brain have thus far focused on increasing gene-transfer to the  
58 rodent central nervous system (CNS), and direct efforts in non-human primates (NHPs) are  
59 sparse. Some recent capsids now enable systemic gene transfer to the brain of the common  
60 marmoset<sup>17</sup> (*Callithrix jacchus*), a New World primate species. But few comparable options exist  
61 for Old World primates, which are more evolutionarily related to humans compared to marmosets  
62 and are well-established animal models of human cognition, neurodevelopment, neuroanatomy,  
63 and physiology<sup>19-21</sup>. To enable research and for greater therapeutic translatability, it is imperative  
64 that we advance AAV development for systemic gene transfer to the brains of Old World primates  
65 such as the rhesus macaque (*Macaca mulatta*).

66 Advances in protein engineering, sequencing technologies, and our understanding of AAV  
67 structure and function have led to the new neurotropic capsid variants that target the rodent brain.  
68 The AAV9 variant AAV-PHP.B was the first variant to unlock efficient widespread gene transfer  
69 in adult mammals, traversing the BBB after systemic intravenous (IV) administration in mice<sup>14</sup>.  
70 Other variants have since followed with similar or enhanced properties, such as the ability to cross  
71 the BBB across different mouse strains, decreased transduction in non-CNS tissue, and biased  
72 tropism towards cell-types in the brain<sup>15-17,22,23</sup>. While two groups reported that the BBB-crossing  
73 tropism of AAV-PHP.B does not translate to the rhesus macaque<sup>24,25</sup>, studies reporting the  
74 systemic translatability of rodent neurotropic capsids in Old World primates are limited. Currently,  
75 no AAV enables efficient gene transfer broadly across the Old World primate brain.

76 In lieu of a capsid variant that can be used for systemic and brain-wide gene transfer in  
77 macaques, researchers and clinicians resort to direct injections to circumvent the BBB. However,  
78 due to limited spatial distribution, AAVs must typically be administered in multiple locations,  
79 invasively penetrating the brain parenchyma each time,<sup>26-33</sup> with each surgery requiring resource-  
80 intensive pre-planning and real-time monitoring of infusions<sup>30-39</sup>. More recently, several groups  
81 have utilized intrathecal routes of administration via lumbar puncture (LP)<sup>40</sup> or intra-cisterna  
82 magna (ICM)<sup>41</sup> injection to overcome the invasiveness and limited spatial distribution of direct  
83 injections. However, these intrathecal routes of administration have limited efficacy in the brain<sup>41-</sup>

84 <sup>45</sup>, and some groups report adverse transduction in non-brain tissue, especially in the dorsal root  
85 ganglia<sup>11,45–47</sup>.

86 Here, we describe AAV.CAP-Mac, an engineered AAV9 variant that efficiently transduces  
87 multiple neuronal subtypes throughout cortical and subcortical brain regions of infant Old World  
88 primates after IV administration. We initially identified CAP-Mac through 2 rounds of selection in  
89 the adult common marmoset (*Callithrix jacchus*). We performed a final round of selection in infant  
90 macaques, where CAP-Mac-delivered transgenes were 10- and 6-times more enriched than  
91 those delivered by AAV9 in viral DNA and whole RNA brain extracts, respectively. CAP-Mac  
92 efficiently transduces the brain in at least two Old World primate species, the rhesus macaque  
93 and the green monkey (*Chlorocebus sabaeus*), achieving broader CNS distribution via IV  
94 compared to intrathecal administration<sup>45,48</sup>. Furthermore, CAP-Mac targets neuronal cells in the  
95 CNS more effectively than its parent AAV9, highlighting the opportunity for broader and more  
96 diverse study of the Old World primate brain, as well as the potential for increased therapeutic  
97 benefit for disorders affecting neurons. As an example of CAP-Mac's immediate research utility,  
98 we capitalized on its neuronal bias to deliver a cocktail of three fluorescent proteins for Brainbow-  
99 like<sup>49,50</sup>, multicolor labeling and morphological tracing in Old World primate brain. By  
100 characterizing CAP-Mac in multiple NHP species, we aim to both expand the AAV toolbox  
101 available to researchers interested in studying the Old World primate CNS and to highlight the  
102 utility of engineering AAVs for increased translatability in higher order mammals.

## 103 Results

### 104 **AAV library selection in adult marmosets yields brain-enriched** 105 **variants**

106 We used a multi-species screening and characterization strategy to select for variants with  
107 enhanced BBB-crossing tropism in NHPs (Fig. 1a). To construct the starting capsid library, we  
108 inserted 21 degenerate bases ([NNK] x 7) after Q588 in the structural *cap* gene of AAV9, identical  
109 to our previously published engineering strategies<sup>14–16</sup> (Supplementary Fig. 1). We performed 2  
110 rounds of selection in 4 adult male marmosets (2 marmosets per round; 2 x 10<sup>12</sup> vector genomes  
111 [vg] of viral library in each marmoset via IV administration). During round 1 selection, we extracted  
112 the brain 4-weeks post-injection and isolated viral DNA from 4 coronal sections per marmoset (8  
113 samples total). Using next-generation sequencing (NGS), we recovered 33,314 unique variants  
114 from the 8 samples of brain. We included all recovered variants from round 1 selection plus a  
115 codon modified variant of each (66,628 total nucleotide sequences) for round 2 selection.

116 Notably, in the past we have used Cre-transgenic mouse lines as part of our Cre  
117 recombination-based AAV targeted evolution (CREATE) methodology to increase stringency  
118 during selections by only recovering variants that undergo cis-Cre-Lox mediated inversion<sup>14,16</sup>.  
119 However, since Cre-transgenic marmosets are currently unavailable, we were unable to confer  
120 this additional selective pressure during these selections and pursued other strategies to  
121 compensate for this loss. We previously demonstrated the utility of clustering capsid variants  
122 based on sequence similarity to generate network graphs as an aid in choosing variants for further  
123 characterization<sup>16</sup>. Briefly, we filter variants based on user-defined performance criteria and  
124 cluster high-performing variants into network graphs, wherein each node is a capsid variant and  
125 each edge represents shared sequence identity between related variants (i.e. the pairwise



126 reverse Hamming distance). We reasoned that through this clustering analysis, we could  
127 efficiently and productively sample variants from our selections to (1) limit the number of animals  
128 used for individual characterization and (2) partially overcome the absence of the selective  
129 pressure provided by Cre-transgenic mice in CREATE.

130 After round 2 selection, we used NGS to quantify the absolute read count of recovered  
131 sequences from the injected virus library and from 12 coronal sections (6 sections per marmoset)  
132 (Fig. 1b). From the variant read counts, we calculated library enrichment scores, filtered variants,  
133 and constructed 2 groups of network graphs (Fig. 1c). One group of network graphs consisted of  
134 113 variants assembled into 37 network graphs, with the largest network containing 22 variants  
135 (Fig. 1d). AAV.CAP-Mac (CAP-Mac) is the most interconnected node in this network, sharing an  
136 edge with 6 other nodes. A second group of network graphs contained 22 total variants assembled  
137 into 7 graphs, with the largest network containing 8 variants (Fig. 1e). AAV.CAP-C2 (CAP-C2),  
138 the most interconnected node in this network graph, connects to 4 other nodes. Because of their  
139 high connectivity to other variants, we selected CAP-Mac and CAP-C2 for further characterization.

### 140 **CAP-Mac is enriched in the infant rhesus macaque CNS compared to** 141 **other engineered variants**

142 Upon completing library selection in the adult marmoset, we initiated capsid-pool studies  
143 in infant rhesus macaques to assess the translatability of several of our engineered AAVs to Old  
144 World primates. By pooling several variants into equimolar vector genome doses and  
145 administering them to the same rhesus macaque, we could limit the number of animals used for  
146 characterization while also assessing variant performance head-to-head, as we have shown that  
147 inter-animal variability exceeds intra-animal variability<sup>51</sup>. We pooled a total of 8 capsid variants:  
148 CAP-Mac and CAP-C2 as well as the parent capsid, AAV9, plus five other previously engineered  
149 AAV controls<sup>15,17</sup>. Each variant packaged a single stranded human frataxin transgene fused to a  
150 hemagglutinin (HA) epitope tag under control of the ubiquitous CAG promoter (ssCAG-hFXN-HA)  
151 with a unique 12 bp molecular barcode in the 3' UTR. This construct design allowed us to assess  
152 protein expression and localization of the virus pool by staining for the HA tag. We then used NGS  
153 to quantify the relative enrichment of each individual barcode in bulk viral DNA and whole RNA  
154 extracts from tissue.

155 We administered  $1 \times 10^{14}$  vg/kg of the virus pool ( $1.25 \times 10^{13}$  vg/kg of each variant) to 2  
156 newborn rhesus macaques via the saphenous vein and extracted brains and livers 4 weeks post-  
157 injection. After staining, we observed widespread expression of the HA epitope throughout the  
158 brain (Fig. 2a), primarily in cells with neuronal morphology. In cortex and in hippocampus, we  
159 observed single cells with clear projections that resemble the apical dendrites of pyramidal cells.  
160 Furthermore, we saw increased HA epitope expression in the thalamus and the dorsal striatum  
161 (insets).

162 Given the robust expression of HA epitope, we extracted bulk viral DNA and whole RNA  
163 from the brain and liver to quantify the relative enrichment of each barcode. In the brain,  
164 AAV.CAP-Mac was the highest performing variant (Fig. 2b). The mean enrichment of the CAP-  
165 Mac delivered-barcodes were 10- and 6 -times higher than the AAV9-delivered barcodes in the  
166 viral DNA and whole RNA brain extracts, respectively. CAP-C2 barcode mean enrichment was  
167 approximately 4-fold higher than AAV9 barcodes in both viral DNA and whole RNA brain extracts.  
168 Interestingly in the viral DNA brain extracts, the barcodes of all other variants, which were selected  
169 in mice, were on par with AAV9. In the liver, variant barcodes were variably enriched relative to

170 AAV9. Notably, CAP-Mac and CAP-C2 were negatively enriched in the liver, along with some of  
171 the previously engineered controls shown to be de-targeted from the liver in rodents<sup>17</sup> (Fig. 2c).  
172 Due to its higher enrichment in the brain, we moved forward with characterizing the single variant  
173 CAP-Mac in two species of Old World primates.

## 174 **CAP-Mac efficiently transduces the CNS in rhesus macaque and green** 175 **monkey infants**

176 Due to the apparent neuronal bias of the capsid pool and CAP-Mac's overabundance in  
177 that pool, we explored using CAP-Mac as a research tool to define neuronal morphology with a  
178 mixture of fluorescent proteins<sup>15,49,50</sup>. To that end, we intravenously administered a cocktail of 3  
179 CAP-Mac vectors packaging ssCAG-mNeonGreen, ssCAG-mRuby2, and ssCAG-mTurquoise2  
180 in equimolar vector genome ratios into newborn rhesus macaques at a total dose of  $5 \times 10^{13}$  vg/kg,  
181 via the saphenous vein. We observed widespread expression of all 3 fluorescent proteins in  
182 cortex, thalamus (lateral geniculate nucleus), and cerebellum (Fig. 3a-c, respectively). Based on  
183 morphology, CAP-Mac primarily transduced neuronal cell types. In the cerebellum and thalamus  
184 specifically, we observed a high density of transduced cells, and the highest proportion of co-  
185 localization of multiple fluorescent proteins. However, co-localization of 2 or 3 fluorescent proteins  
186 appeared to be rare, suggesting that co-infection was uncommon after systemic administration.  
187 Fluorescent protein expression appeared to be similar across different coronal slices along the  
188 anterior-posterior axis (Fig. 4a). Fluorescent protein expression was robust in all four lobes of  
189 cortex and in subcortical areas like the dorsal striatum and hippocampus. We continued to see  
190 strong fluorescent protein expression in the thalamus, including in the lateral and medial nuclei,  
191 the lateral geniculate nucleus, and the pulvinar. Despite using the ubiquitous CAG promoter, we  
192 observed expression almost entirely in neuronal subtypes based on morphology in all regions.  
193 Given the broad and robust expression of fluorescent proteins throughout the brain, we were able  
194 to assemble morphological reconstructions of both medium spiny neurons (Fig. 4b) and cortical  
195 pyramidal cells (Fig. 4c), which are both known to be implicated in human disease<sup>52,53</sup>. In a  
196 separate study, we also attempted to administer CAP-Mac via LP administration in infant rhesus  
197 macaques, but we report here that efficiency throughout the brain was noticeably decreased  
198 compared to the IV administered animals (Supplementary Fig. 2) and especially low in subcortical  
199 structures, which has previously been reported<sup>41-45</sup>.

200 Other engineered AAV variants for BBB-crossing in mice are known to have strain-  
201 dependent behavior<sup>16,25,54,55</sup>. Therefore, in parallel to the rhesus macaque experiments, we  
202 characterized CAP-Mac in green monkeys, another Old World primate species. We administered  
203 AAV9 or CAP-Mac packaging ssCAG-eGFP to individual 8-month-old monkeys at a dose of  $7.5$   
204  $\times 10^{13}$  vg/kg via the saphenous vein. In the CAP-Mac-dosed green monkeys, we saw broad and  
205 strong neuronal expression in all four lobes of cortex and in various subcortical regions, including  
206 the putamen (Fig. 5a), which is consistent with the expression we observed in the pooled studies  
207 (Fig. 2a) and in the rhesus macaque (Fig. 3 and Fig. 4). Notably, we saw particularly strong  
208 enhanced green fluorescent protein (eGFP) expression throughout the cerebellum in the CAP-  
209 Mac-dosed green monkey. Consistent with other reports, AAV9 tropism appeared to be primarily  
210 biased towards cell types with astrocyte-like morphology<sup>41,45,56,57</sup>, and neuronal transduction was  
211 low throughout the cortex in AAV9-dosed green monkeys. Conversely, CAP-Mac is biased  
212 towards neurons throughout the brain, transducing a higher percentage of neurons than AAV9 in  
213 all cortical and subcortical regions that we sampled (Fig. 5b). Viral DNA extracts from various

214 brain regions corroborated this observed increased GFP expression, as recovered eGFP  
215 transgene in CAP-Mac green monkeys is consistently higher throughout the brain compared to  
216 AAV9-dosed monkeys, suggesting overall higher brain penetrance of CAP-Mac (Fig. 5c and  
217 Supplementary Fig. 3a). Interestingly, the cerebellum contained the fewest vector genomes per  
218 diploid genome in both CAP-Mac monkeys despite strong eGFP expression, most likely due to  
219 the high density of cells and processes within the cerebellum<sup>58,59</sup>. In most non-brain tissue, eGFP  
220 biodistribution and expression appeared to be comparable between CAP-Mac and AAV9 treated  
221 animals (Supplementary Fig. 3). However, because CAP-Mac transduces a higher percentage of  
222 neurons than AAV9, it is worth noting that cell-type tropism differences may continue to persist in  
223 non-brain tissue as well, and single-cell studies highlight that even in homogenous cell  
224 populations, there is significant viral infection variability<sup>51,60,61</sup>. As such, it is important to consider  
225 the role of differential infection dynamics in heterogenous cell types when interpreting viral DNA  
226 biodistribution, as measured viral DNA may not correlate to capsid penetrance in tissue in an  
227 identical manner across variants and cell types.

228 Given that we selected CAP-Mac in NHPs and due to its neuronal bias, we wanted to  
229 assess the capsid's utility in rodents. Therefore, we administered CAP-Mac packaging ssCAG-  
230 mNeonGreen to C57BL/6J, BALB/cJ, and DBA/2J adult mice via both IV and  
231 intracerebroventricular (ICV) administrations. Interestingly, the neuronal bias of CAP-Mac  
232 extended to mice when delivered to the adult brain through ICV administration but not IV, where  
233 it appeared to primarily transduce cells that make up the vasculature (Supplemental Fig. 4a and  
234 b). Furthermore, there is no apparent difference in CAP-Mac tropism across the three mouse  
235 strains. We also IV administered CAP-Mac to P0 C57BL/6J mice and observed ssCAG-  
236 mNeonGreen expression throughout the brain in various cell types, including neurons, astrocytes,  
237 and vasculature (Supplemental Fig. 4c). As rodent experiments are more accessible and have  
238 shorter timelines compared to NHPs, the strong neuronal tropism of CAP-Mac in adult mice after  
239 ICV administration offers a method to screen genetic cargo in mice prior to applying them in NHPs.

## 240 **CAP-Mac strongly transduces human neurons derived from** 241 **pluripotent stem cells**

242 Given the efficacy of CAP-Mac in penetrating the brain of infant Old World primates and  
243 motivated by our observations that CAP-Mac primarily transduces neurons, we wanted to verify  
244 whether CAP-Mac offered any improvement over its parent capsid, AAV9, in transducing human  
245 neurons. We differentiated cultured induced pluripotent stem cells (iPSCs) derived from humans  
246 into mature neurons (Fig. 6a) and incubated cultures with CAP-Mac or AAV9 packaging ssCAG-  
247 eGFP across a broad range of doses ranging from 1 vg/cell to 10<sup>6</sup> vg/cell. We found that GFP  
248 expression was noticeably increased in CAP-Mac-administered cultures compared to AAV9-  
249 administered cultures (Fig. 6b). Across the aforementioned dose range, we compared the number  
250 of transduced cells and eGFP expression per cell between the CAP-Mac and AAV9-treated  
251 cultures. AAV9 transduction efficiency achieved an EC<sub>50</sub>=10<sup>4.68</sup> vg/cell, while CAP-Mac achieved  
252 an EC<sub>50</sub>=10<sup>3.03</sup> vg/cell (Fig. 6c), a 45-fold increase in potency of CAP-Mac in transducing human  
253 neurons *in vitro*. Average per cell eGFP expression measured across the population of  
254 transduced cells fits a biphasic step function with CAP-Mac reaching the first plateau at a dose  
255 roughly two orders of magnitude lower than AAV9 (Fig. 6d). Overall, the increased potency of  
256 CAP-Mac in transducing mature human neurons *in vitro* is consistent with the neuronal tropism

257 observed in Old World primates, suggesting a similar mechanism of neuronal transduction across  
258 species.

## 259 Discussion

260 We describe here the selection and characterization of CAP-Mac, an engineered AAV  
261 capsid variant that enables brain-wide transgene expression in Old World primates. CAP-Mac  
262 outperforms AAV9 and broadly transduces neurons throughout both cortical and subcortical  
263 structures in two Old World primates, the rhesus macaque and the green monkey.

264 *In vivo* AAV capsid selections have been primarily conducted in mice, partially due to the  
265 utility of Cre-transgenic mouse lines to increase selective pressure during selections, as these  
266 Cre-based selections can yield neurotropic capsids in as few as two rounds of selection<sup>14,16,17</sup>.  
267 However, these engineered variants have thus far failed to translate to NHPs<sup>24,25</sup>. The notable  
268 exceptions are AAV.CAP-B10 and AAV.CAP-B22, which were identified using M-CREATE<sup>16</sup>  
269 selections in mice and retain their BBB-crossing and reduced liver tropism in the common  
270 marmoset<sup>17</sup>, a New World primate. However, our pool testing experiments here showed that these  
271 variants are on par with AAV9 within viral DNA extracts from the brains of infant macaques, an  
272 Old World primate. While mice last shared a common ancestor with humans approximately 80-  
273 90 million years ago (mya), marmosets and macaques are believed to have shared their last  
274 ancestor with humans 35-40 mya and 25-30 mya, respectively<sup>62</sup>. Given this distant shared  
275 ancestry, it is therefore not surprising that most variants selected in mice have failed to translate  
276 to Old World primates and human patients. Interestingly, we show here in pool studies that  
277 variants identified via Cre-independent selections in marmosets and chosen using network graphs  
278 (CAP-Mac and CAP-C2), generally outperformed variants identified via Cre-dependent selections  
279 in mice (Fig. 2b). This observation suggests that while technological advancements for enhancing  
280 selective pressures are important when evolving engineered AAVs *in vivo*, the evolutionary  
281 relatedness between the selection-species and target-species is vital to consider if the  
282 engineering goal is translatability in higher-order organisms. Notably, several transgenic  
283 marmoset lines are currently available<sup>63,64</sup>, and the generation of Cre-transgenic marmosets is  
284 currently underway<sup>65</sup>, opening the possibility of performing M-CREATE in NHPs. Interestingly,  
285 AAV.CAP-B10 and AAV.CAP-B22 were selected in mice but retain their BBB-crossing tropisms  
286 in marmosets. Given that the evolutionary distance between mouse and marmosets (40-55 mya)  
287 is slightly larger than the relatedness between marmosets and humans (35-40 mya), this offers  
288 hope that NHP selections can identify capsid variants that are efficacious in humans. Overall,  
289 CAP-Mac is a testament to the utility of using a selection species that is closely related  
290 evolutionarily to the target species.

291 In the Old World primate studies presented here, we used CAP-Mac exclusively in infants.  
292 While this presented several logistical benefits to partially de-risk our characterization efforts (e.g.  
293 infants are more likely to be seronegative for neutralizing AAV antibodies and it is more practical  
294 to produce sufficient virus to IV dose multiple primates if they weigh less), there is a perception  
295 that the developing BBB is “leaky” due to underdeveloped structural integrity, making it more  
296 permissive to molecules compared to the adult BBB. However, this interpretation of classical  
297 literature has been challenged more recently<sup>66-68</sup>, and contemporary studies report that tight  
298 junctions, the main structural unit that enables the BBB to act as a physical barrier between  
299 peripheral blood and the extracellular space, are formed embryonically and that the BBB is



300 structurally intact by birth<sup>69-71</sup>. This suggests that CAP-Mac's brain-penetrating tropism in infant  
301 primates and mice is not a passive process, and rather requires active transport across the BBB,  
302 most likely via receptor-mediated transcytosis. While the BBB's structural integrity may be static  
303 postnatally, other molecular and cellular developmental processes are dynamic during  
304 maturation<sup>69,70,72</sup>, which may dictate tropism differences across developmental states. Such  
305 developmental differences in BBB state may be apparent in our studies in mice, where IV  
306 administration of CAP-Mac to adult mice results in ssCAG-mNeonGreen expression in  
307 vasculature, but the same experiment in P0 mouse pups results in expression in neurons,  
308 astrocytes, and vasculature—a stark difference in tropism. The dynamic nature of the molecular  
309 and cellular properties of the BBB and the difference in CAP-Mac tropism between adult mice and  
310 P0 mouse pups shown here highlight the importance of considering the developmental state when  
311 using AAVs both in research and in the clinic.

312 Given this observation, we are actively working to fully characterize CAP-Mac in infants  
313 and adults of various primate species, as a major overarching goal of this study is to define and  
314 disseminate a suite of genetic tools to study the NHP brain, especially in Old World primates. This  
315 includes characterizing functional cargo that can be paired with CAP-Mac to study the macaque  
316 brain. The first such demonstration is included here, where we deliver a cocktail of three  
317 fluorescent proteins using CAP-Mac to achieve noninvasive, Brainbow-like<sup>49,50</sup> labeling in the  
318 macaque brain. Because CAP-Mac transduces neurons in cortical and subcortical areas, it  
319 provides researchers a more accessible and efficient method to elucidate the morphology of  
320 various neuronal cell types, including medium spiny neurons and pyramidal cells. These are both  
321 key cell types implicated in neurodegenerative disorders<sup>52,53,73,74</sup>, and the experiments shown here  
322 open the door for comparing these cell types in diseased and healthy macaques. Moving forward,  
323 the work here will enable major efforts under the NIH BRAIN Initiative<sup>75</sup>, as CAP-Mac-mediated  
324 labeling can be readily combined with tissue clearing and imaging techniques to map long-range  
325 projections within the Old World primate connectome<sup>76-78</sup>. Such studies present a novel  
326 opportunity to study the macaque brain, and more generally, an opportunity to understand the  
327 inner workings of the primate CNS.

328 In addition to CAP-Mac's utility as a tool to study the primate brain, it may also represent  
329 a compelling delivery vehicle for genetic medicine in humans. CAP-Mac, with its efficient neuronal  
330 transduction across the macaque brain when delivered IV, provides an unprecedented  
331 opportunity to deepen our understanding of the pharmacodynamics of genetic medicines in Old  
332 World primate models<sup>32,79,80</sup>. Furthermore, the broad and uniform distribution of CAP-Mac  
333 throughout the primate CNS has the potential to provide unprecedented therapeutic access to  
334 subcortical and midbrain regions, which has previously proven difficult in NHPs<sup>41-45</sup>. Of further  
335 importance for IV delivered gene therapies, the data suggests that CAP-Mac may have reduced  
336 tropism towards the liver across primate species. Additionally, CAP-Mac's enhanced transduction  
337 of cultured human neurons supports the potential of CAP-Mac as a clinically relevant gene-  
338 delivery vehicle in humans. Overall, the success of the capsid engineering approach we describe  
339 here to generate novel variants with BBB-crossing tropism and cell-type bias in Old World  
340 primates offers a roadmap for developing the next class of translational gene therapies with  
341 improved safety and efficacy profiles.

## 342 Methods

### 343 AAV DNA library generation

344 We initially generated our diversity at the DNA level, which is used to produce transfection  
345 material to produce the AAV capsid library. For the round 1 library, we introduced this genetic  
346 diversity using primers containing degenerate nucleotides that were inserted between CAP amino  
347 acids 588 and 589<sup>14-16</sup> (Supplementary Fig. 1a). We used a reverse primer containing 21  
348 degenerate nucleotides ([NNK] x 7) to randomly generate PCR fragments containing unique 7mer  
349 sequences inserted into the *cap* genome. For the round 2 DNA library, we used a synthetic oligo  
350 pool (Twist Bioscience) as a reverse primer, encoding only variants that we selected for further  
351 screening (66,628 DNA oligos total: 33,314 variants recovered after round 1 selections plus a  
352 codon-modified replicate of each). All reverse primers contained a 20 bp 5' overhang  
353 complementary with the CAP sequence near the AgeI restriction enzyme sequence and were  
354 paired with a forward primer containing a 20 bp 5' overhang near the XbaI restriction enzyme  
355 sequence. We then inserted the PCR fragments containing the diversified region into the rAAV-  
356 ΔCAP-in-cis-Lox plasmid via Gibson assembly to generate the resulting AAV DNA library, rAAV-  
357 CAP-in-cis-Lox, using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, E2621).

### 358 AAV capsid library production

359 We generated AAV capsid libraries according to previously published protocols<sup>16,81</sup>.  
360 Briefly, we transfected HEK293T cells (ATCC, CRL-3216) in 150 mm tissue culture plates using  
361 transfection grade, linear polyethylenimine (PEI; Polysciences, Inc). In each plate, we transfected  
362 4 plasmids: (1) the assembled rAAV-Cap-in-cis-Lox AAV DNA library, which is flanked by inverted  
363 terminal repeats (ITR) required for AAV encapsidation; (2) AAV2/9 REP-AAP-ΔCAP, which  
364 encodes the REP and AAP supplemental proteins required for AAV production with the C-  
365 terminus of the CAP gene excised to prevent recombination with the AAV DNA library and  
366 subsequent production of replication-competent AAV; (3) pHelper, which encodes the necessary  
367 adenoviral proteins required for AAV production; and (4) pUC-18, which contains no mammalian  
368 expression vector but is used as filler DNA to achieve the appropriate nitrogen-to-phosphate ratio  
369 for optimal PEI transfection. During preparation of the PEI-DNA mixture, we added 10 ng of our  
370 AAV DNA library (rAAV-Cap-in-cis-Lox) for every 150 mm dish and combined AAV2/9 REP-AAP-  
371 ΔCAP, pUC-18, and pHelper in a 1:1:2 ratio, respectively (40 μg of total DNA per 150 mm dish).  
372 At 60 hours post-transfection, we purified AAV capsid library from the both the cell pellet and  
373 media using polyethylene glycol precipitation and iodixanol gradient ultracentrifugation. Using  
374 quantitative PCR, we then determined the titer of the AAV capsid libraries by amplifying DNaseI  
375 resistant viral-genomes relative to a linearized genome standard according to established  
376 protocols<sup>81</sup>.

### 377 Marmoset experiments

378 All marmoset (*Callithrix jacchus*) procedures were performed at the National Institutes of  
379 Mental Health (NIMH) and approved by the local Institutional Animal Care and Use Committee  
380 (IACUC). Marmosets were born and raised in NIMH colonies and housed in family groups under  
381 standard conditions of 27°C and 50% humidity. They were fed ad libitum and received enrichment  
382 as part of the primate enrichment program for NHPs at the National Institutes of Health. For all  
383 marmosets used in this study, there were no detectible neutralizing antibodies at a 1:5 serum  
384 dilution prior to IV infusions (conducted by The Penn Vector Core, University of Pennsylvania).



385 They were then housed individually for several days and acclimated to a new room before  
386 injections. Four adult males were used for the library screening, 2 each for first- and second-round  
387 libraries. The day before infusion, the animals' food was removed. Animals were anesthetized  
388 with isoflurane in oxygen, the skin over the femoral vein was shaved and sanitized with an  
389 isopropanol scrub, and  $2 \times 10^{12}$  vg of the AAV capsid library was infused over several minutes.  
390 Anesthesia was withdrawn and the animals were monitored until they became active, upon which  
391 they were returned to their cages. Activity and behavior were closely monitored over the next 3  
392 days, with daily observations thereafter.

393 At 4 weeks post-injection, marmosets were euthanized (Euthanasia, VetOne) and  
394 perfused with 1X phosphate-buffered saline (PBS). After the round 1 library, the brain was cut  
395 into 4 coronal blocks, flash frozen in 2-methylbutane (Sigma Aldrich, M32631), chilled with dry  
396 ice, and stored at  $-80^{\circ}\text{C}$  for long term storage. After the round 2 library, the brain was cut into 6  
397 coronal blocks, and along with sections of the spinal cord and liver, was flash frozen and stored  
398 at  $-80^{\circ}\text{C}$  for long term storage.

### 399 **Viral library DNA extraction and NGS sample preparation**

400 To extract viral library DNA from marmoset tissue, we previously reported that viral library  
401 DNA and endogenous host RNA can be isolated using Trizol by precipitating nucleic acid from  
402 the aqueous phase<sup>14,16</sup>. As such, we homogenized 100 mg of spinal cord, liver, and each coronal  
403 block of brain in Trizol (Life Technologies, 15596) using a BeadBug (Benchmark Scientific,  
404 D1036) and isolated nucleic acids from the aqueous phase according to the manufacturer's  
405 recommended protocol. We treated the reconstituted precipitate with RNase (Invitrogen,  
406 AM2288) and digested with SmaI to improve downstream viral DNA recovery via PCR. After  
407 digestion, we purified with a Zymo DNA Clean and Concentrator kit (D4033) according to  
408 manufacturer's recommended protocol and stored the purified viral DNA at  $-20^{\circ}\text{C}$ .

409 To append Illumina adapters flanking the diversified region, we first PCR-amplified the  
410 region containing our 7mer insertion using 50% of the total extracted viral DNA as a template (25  
411 cycles). After Zymo DNA purification, we diluted samples 1:100 and further amplified around the  
412 library variable region with 10 cycles of PCR, appending binding regions for the next PCR  
413 reaction. Finally, we appended Illumina flow cell adapters and unique indices using NEBNext Dual  
414 Index Primers (New England Biolabs, E7600) via 10 more cycles of PCR. We then gel-purified  
415 the final PCR products using a 2% low-melting point agarose gel (ThermoFisher Scientific,  
416 16520050) and recovered the 210 bp band.

417 For the second-round library only, we also isolated the encapsidated AAV library ssDNA  
418 for NGS to calculate library enrichment scores, a quantitative metric that we use to normalize for  
419 differences in titer of the various variants in our library (see ref. 11 and the "NGS read alignment  
420 and analysis" method section below). To isolate the encapsidated viral genomes, we treated the  
421 AAV capsid library with DNaseI and digested capsids using proteinase K. We then purified the  
422 ssDNA using phenol: chloroform and amplified viral transgenes by 2 PCR amplification steps to  
423 add adapters and indices for Illumina NGS and purified after gel electrophoresis. This viral library  
424 DNA, along with the viral DNA extracted from tissue, was sent for deep sequencing using an  
425 Illumina HiSeq 2500 system (Millard and Muriel Jacobs Genetics and Genomics Laboratory,  
426 Caltech).

## 427 **NGS read alignment, analysis, and generating network graphs**

428 Raw fastq files from NGS runs were processed with custom-built scripts  
429 (<https://github.com/GradinaruLab/protfarm> and <https://github.com/GradinaruLab/mCREATE>)<sup>16</sup>.  
430 For the first-round library, we processed the pipeline to process these datasets involved filtering  
431 to remove low-quality reads, utilizing a quality score for each sequence, and eliminating bias from  
432 PCR-induced mutations or high GC-content. The filtered dataset was then aligned by a perfect  
433 string match algorithm and trimmed to improve the alignment quality. We then displayed absolute  
434 read counts for each variant during the sequencing run within each tissue, and all 33,314 variants  
435 that were found in the brain were chosen for round 2 selections.

436 After round two selections, we performed the same analysis to display variant absolute  
437 read count of the injected virus library and of each variant within each tissue. Additionally, we  
438 calculated the library enrichment<sup>16</sup> for each variant within each tissue:

$$439 \quad \widehat{RC}_{x, \text{injected library}} = \frac{RC_{x, \text{injected library}}}{\sum_{i=1}^{N_{\text{injected library}}} RC_{i, \text{injected library}}}$$

$$440 \quad \widehat{RC}_{x, \text{tissue}} = \frac{RC_{x, \text{virus}}}{\sum_{i=1}^{N_{\text{tissue}}} RC_{i, \text{tissue}}}$$

$$441 \quad \text{library enrichment} = \log_{10} \left( \frac{\widehat{RC}_{x, \text{injected library}}}{\widehat{RC}_{x, \text{tissue}}} \right)$$

442 such that for a given sample  $y$  (e.g. the injected virus library or a tissue sample),  $RC_{x,y}$  is the  
443 absolute read count of variant  $x$ ,  $N_y$  is the total number of variants recovered, and  $\widehat{RC}_{x,y}$  is the  
444 normalized read count.

445 To construct the CAP-Mac sequence clustering graph, we filtered the round 2 NGS data  
446 based on the following criteria: (1)  $\geq 100$  read count in the injected library sample (24,186/33,314  
447 variants), (2)  $\geq 0.7$  library enrichment score in more than 2 brain samples (415 variants), and (3)  
448 at least 2 more brain samples with  $\geq 0.7$  library enrichment than brain samples with  $< -0.7$  library  
449 enrichment (323 variants). To construct the CAP-C2 sequence graph, we filtered the round 2 NGS  
450 data based on the following criteria: (1)  $\geq 100$  read count in the injected library sample and (2)  
451 both codon replicates present in at least 2 brain samples with  $\geq 0.7$  library enrichment (95  
452 variants). These variants were then independently processed to determine pair-wise reverse  
453 Hamming distances (<https://github.com/GradinaruLab/mCREATE>) and clustered using  
454 Cytoscape (ver. 3.9.0) as described previously<sup>16</sup>. Networks presented show capsid variants  
455 (nodes) connected by edges if the pair-wise reverse Hamming distance is  $\geq 3$ .

## 456 **Cloning individual AAV capsid variants**

457 For single variant characterization, we cloned new variant plasmids by digesting a  
458 modified version of the pUCmini-iCAP-PHP.eB (Addgene ID: 103005) backbone using MscI and  
459 AgeI. We designed a 100 bp primer that contained the desired 21 bp insertion for each capsid  
460 variant and the regions complementary to the AAV9 template with  $\sim 20$  bp overlapping regions  
461 with the digested backbone. We then assembled the variant plasmid using NEBuilder HiFi DNA  
462 Assembly Master Mix, combining 5  $\mu\text{L}$  of 200 nM primer with 30 ng of digested backbone in the  
463 reaction mixture.

## 464 **Individual AAV production and purification**

465 To produce variants for pool testing, we followed our previously published protocol<sup>81</sup> using  
466 150 mm tissue culture dishes. For individual AAV.CAP-Mac and AAV9 characterization *in vivo*  
467 and *in vitro*, we adopted our published protocol to utilize ten-layer CellSTACKs (Corning, 3320)  
468 to efficiently produce viruses at high titer to dose rhesus macaques and green monkeys.  
469 Specifically, we passaged 20 150-mm dishes at approximately 70% confluency into a 10-layer  
470 CellSTACK 24 h before transfection. On the day of transfection, we prepared the DNA-PEI  
471 transfection mixture for 40 150-mm dishes and combined the transfection mixture with media and  
472 performed a complete media change for the CellSTACK. We collected and changed media at 72  
473 h post-transfection similarly to production in 150 mm dishes. At 120 h post-transfection, we added  
474 ethylenediaminetetraacetic acid (EDTA, Invitrogen, 15575020) to a final concentration of 10 mM  
475 and incubated at 37°C for 20 min, occasionally swirling and tapping the sides of the CellSTACK  
476 to detach the cells. We then removed the media and cell mixture and proceeded with the AAV  
477 purification protocol<sup>81</sup>. Of note, during the buffer exchange step after ultracentrifugation, we used  
478 centrifugal protein concentrators with polyethersulfone membranes (Thermo Scientific, 88533)  
479 instead of Amicon filtration devices and used Dulbecco's PBS supplemented with 0.001%  
480 Pluronic® F-68 (Gibco, 24040032).

## 481 **Rodent experiments**

482 All rodent procedures were performed at California Institute of Technology (Caltech) and  
483 were approved by the local IACUC. We purchased C57BL/6J (000664), BALB/cJ (000651), and  
484 DBA/2J (000671) mice (all males, 6–8 weeks old) from The Jackson Laboratory. For IV  
485 administration in mice, we delivered virus through the retro-orbital sinus<sup>81,82</sup> using a 31 G insulin  
486 syringe (BD, 328438). For intracerebroventricular administration in mice, we injected into the  
487 lateral ventricle. Briefly, we anesthetized mice using isoflurane (5% for induction, 1-3% for  
488 maintenance) with 95% O<sub>2</sub>/5% CO<sub>2</sub> (1 L/min) and mice were head-fixed in a stereotaxic frame.  
489 After shaving the head and sterilizing the area with chlorohexidine, we administered 0.05 mL of  
490 2.5 mg/mL bupivacaine subcutaneously, and a midline incision was made and the skull was  
491 cleaned of blood and connective tissue. After leveling the head, burr holes were drilled above the  
492 lateral ventricles bilaterally (0.6 mm posterior to bregma, 1.15 mm from the midline). Viral vectors  
493 were aspirated into 10 µL NanoFil syringes (World Precision Instruments) using a 33-gauge  
494 microinjection needle, and the needle was slowly lowered into the lateral ventricle (1.6 mm from  
495 the pial surface). The needle was allowed to sit in place for approximately 5 min and 3-5 µL of  
496 viral vector was injected using a microsyringe pump (World Precision Instruments, UMP3) and  
497 pump controller (World Precision Instruments, Micro3) at a rate of 300 nL/min. All mice received  
498 1 mg/kg of buprenorphine SR and 5 mg/kg of ketoprofen subcutaneously intraoperatively and 30  
499 mg/kg of ibuprofen and 60 mg/kg of Trimethoprim/ Sulfamethoxazole (TMPS) for 5 days post-  
500 surgery. After 3 weeks of expression, all mice were perfused with PBS and fixed in 4%  
501 paraformaldehyde (PFA). All organs were extracted, incubated in 4% PFA overnight, transferred  
502 into PBS supplemented with 0.01% sodium azide, and stored at 4°C for long-term storage. We  
503 sliced the brain into 100 µm sections by vibratome (Leica Biosystems, VT1200S), mounted in  
504 Prolong Diamond Antifade (Invitrogen, P36970), and imaged using a confocal microscope (Zeiss,  
505 LSM 880).

## 506 **Rhesus macaque experiments**

507 All rhesus macaque (*Macaca mulatta*) procedures were performed at the California  
508 National Primate Research Center (CNPRC) at UC Davis and were approved by the local IACUC.  
509 Infant macaques were weaned at birth. Within the first month, macaques were infused with AAV  
510 vectors either intravenously (IV) or intrathecally. For IV injections, animals were anesthetized with  
511 ketamine (0.1 mL) and the skin over the saphenous vein was shaved and sanitized. AAV (see  
512 Supplementary tables 1 and 2) was slowly infused into the saphenous vein over ~1 min in < 0.75  
513 mL of phosphate buffered saline. For IT injections, animals were administered a sedative  
514 intramuscularly and the area of skin at the neck was shaved and aseptically prepared. A needle  
515 was advanced into the cisterna magna to remove a small amount of CSF proportional to the  
516 amount of fluid injected. Then, a sterile syringe containing the sterile preparation of the AAV  
517 proportional to the amount of fluid collected was aseptically attached and slowly injected. All  
518 animals were monitored during recovery from sedation, throughout the day, and then daily for any  
519 adverse findings. All monkeys were individually housed within sight and sound of conspecifics.  
520 Tissue was collected 4-11 weeks after injection. Animals were deeply anesthetized and received  
521 sodium pentobarbital in accordance with guidelines for humane euthanasia of animals at the  
522 CNPRC. All material injected into rhesus macaques were free of endotoxins (<0.1 EU/mL), and  
523 protein purity was confirmed by sodium dodecyl sulphate–polyacrylamide gel electrophoresis  
524 (SDS-PAGE). See Supplementary tables 1 and 2 for route of administration, AAV variants, viral  
525 dose, genetic cargo, and duration of expression for each experiment.

### 526 *Pool testing in infants*

527 Macaques were perfused with ice cold RNase-free PBS. At the time of perfusion, one  
528 hemisphere of brain was flash-frozen and the other hemisphere was sectioned into 4 mm coronal  
529 blocks and post fixed in 4% PFA for 48 hours and transferred to Caltech for further processing.  
530 For HA staining, we incubated slices with rabbit anti-HA (1:200; Cell Signaling Technology, 3724),  
531 performed 3-5 washes with PBS, incubated with donkey anti-rabbit IgG (1:200; Jackson  
532 ImmunoResearch, 711-605-152), and washed 3-5 times before mounting. We diluted all  
533 antibodies and performed all incubations using PBS supplemented with 0.1% Triton X-100  
534 (Sigma-Aldrich, T8787) and 10% normal donkey serum (Jackson ImmunoResearch, 017-000-  
535 121) overnight at room temperature with shaking.

536 To isolate viral DNA and whole RNA, 100mg slices from brain and liver were homogenized  
537 in Trizol (Life Technologies, 15596) using a BeadBug (Benchmark Scientific, D1036) and total  
538 DNA and RNA were recovered according to the manufacturer's recommended protocol.  
539 Recovered DNA was treated with RNase, underwent restriction digestion with SmaI, and purified  
540 with a Zymo DNA Clean and Concentrator Kit (D4033). Recovered RNA was treated with DNase,  
541 and cDNA was generated from the mRNA using Superscript III (Thermo Fisher Scientific,  
542 18080093) and oligo(dT) primers according to the manufacturer's recommended protocol. We  
543 used PCR to amplify the barcoded using 50ng of viral DNA or cDNA as template. After Zymo DNA  
544 purification, we diluted samples 1:100 and further amplified around the barcode region using  
545 primers to append adapters for Illumina next-generation sequencing. After cleanup, these  
546 products were further amplified using NEBNext Dual Index Primers for Illumina sequencing (New  
547 England Biolabs, E7600) for ten cycles. We then gel-purified the final PCR products using a 2%  
548 low-melting point agarose gel (ThermoFisher Scientific, 16520050). Pool testing enrichment was  
549 identically to library enrichment, but is represented in Fig 2b and c on a linear scale.



## 550 *Individual characterization of CAP-Mac in infants*

551 Macaques were perfused with PBS and 4% PFA. The brain was sectioned into 4 mm  
552 coronal blocks and all tissue was post-fixed in 4% PFA for 3 days before storage in PBS. All tissue  
553 was transferred to Caltech for further processing. Brains and liver were sectioned into 100  $\mu$ m  
554 slices using vibratome. Sections of spinal cord were incubated in 30% sucrose overnight and  
555 embedded in Optimal Cutting Temperature Compound (Scigen, 4586) and sectioned into 50  $\mu$ m  
556 slices using cryostat (Leica Biosystems, CM1950). All slices were mounted using Prolong  
557 Diamond Antifade and imaged using a confocal microscope. For GFP staining of brain slices from  
558 the LP administered macaque, we incubated slices with chicken anti-GFP (1:500; Aves Bio, GFP-  
559 1020), performed 3-5 washes with PBS, incubated with donkey anti-chicken IgY (1:200; Jackson  
560 ImmunoResearch, 703-605-155), and washed 3-5 times before mounting. We diluted all  
561 antibodies and performed all incubations using PBS supplemented with 0.1% Triton X-100  
562 (Sigma-Aldrich, T8787) and 10% normal donkey serum (Jackson ImmunoResearch, 017-000-  
563 121) overnight at room temperature with shaking.

564 For morphological reconstruction, we sectioned brain into 300  $\mu$ m sections and incubated  
565 them in refractive index matching solution (RIMS)<sup>83</sup> for 72 hours before mounting on a slide  
566 immersed in RIMS. We imaged using a confocal microscope and 25x objective (LD LCI Plan-  
567 Achromat 25x/0.8 Imm Corr DIC) using 100% glycerol as the immersion fluid. We captured  
568 tiled, Z-stacks (1024x1024 each frame using suggested capture settings) around cells of interest,  
569 and cropped appropriate fields of view for tracing. Tracing was done Imaris using the semi-  
570 automated and automated methods.

## 571 **Green monkey experiments**

572 All green monkey (*Chlorocebus sabaeus*) procedures were performed at Virscio, Inc. and  
573 approved by their IACUC. All monkeys were screened for neutralizing antibodies and confirmed  
574 to have < 1:5 titer. At approximately 7-8 months of age, monkeys were dosed intravenously. Dose  
575 formulations were allowed to equilibrate to approximately room temperature for at least 10  
576 minutes, but no more than 60 minutes prior to dosing. IV dose volumes were based on Day 0  
577 body weights. Animals were sedated with ketamine (8 mg/kg) and xylazine (1.6 mg/kg). The  
578 injection area was shaved and prepped with chlorohexdrine and 70% isopropyl alcohol, surgically  
579 scrubbed prior to insertion of the intravenous catheter. Dosing occurred with a single intravenous  
580 infusion on Day 0 via a saphenous vein administered using a hand-held infusion device at a target  
581 rate of 1 mL/minute. General wellbeing was confirmed twice daily by cage side observation  
582 beginning one week prior to dosing. At the scheduled sacrifice time, monkeys were sedated with  
583 ketamine (8-10 mg/kg IM) and euthanized with sodium pentobarbital (100 mg/kg IV to effect).  
584 Upon loss of corneal reflex, a transcatheter perfusion (left ventricle) was performed with chilled  
585 phosphate buffered saline (PBS) using a peristaltic pump set at a rate of approximately 100  
586 mL/min until the escaping fluid ran clear prior to tissue collection. Cubes of tissue were collected  
587 from the left brain hemisphere and various other organs and frozen in the vapor phase of liquid  
588 nitrogen for further processing for biodistribution. The right brain hemisphere was removed and  
589 cut into ~4 mm coronal slices and post-fixed intact with approximately 20 volumes of 10% neutral-  
590 buffered formalin (NBF) for approximately 24 hours at room temperature.

591 Genomic DNA was extracted from CNS and peripheral tissues using the ThermoFisher  
592 MagMax DNA Ultra 2.0 extraction kit (Catalog number: A36570). DNA was assessed for yield by  
593 fluorometric quantification with the Qubit dsDNA assay. Approximately 20 ng of DNA was loaded

594 into each 20  $\mu$ L reaction and plates were run on the BioRad CFX Connect Real-Time PCR  
595 Detection System (Catalog number: 1855201). The viral copy number assay was validated for  
596 specificity by detection of a single amplified product, sensitivity by assessing the lower limit of  
597 detection to be greater than 10 copies per reaction and linearity by ensuring the standard curve  
598  $r^2$  was  $> 0.95$ . Reactions were assembled in FastStart Universal SYBR Green Master (Rox)  
599 (catalogue number: 4913850001). The following were the sequences of the primers: forward  
600 ACGACTTCTTCAAGTCCGCC, reverse TCTTGTAGTTGCCGTCGTCC. The following PCR  
601 protocol was used an initial denaturation step of 95 °C for 180 seconds, followed by 40 cycles of  
602 95 °C for 15 seconds, 60 °C for 60 seconds; with an imaging step following each 60 °C cycle.  
603 Standard curve generated with linearized plasmid containing the GFP template sequence present  
604 in the virus from 1e8-1e0 copies, diluted in naïve untreated macaque DNA samples prepared  
605 using an identical kit to the samples in this study to control for matrix effects. Copies of viral DNA  
606 were calculated from the standard curve using the equation for the line of the best fit. MOI values  
607 were calculated based on the measured total genomic weight of the host cell DNA per reaction.

608 Post fixation, tissues were placed into 10% > 20% > 30% sucrose for 24 hours each at 4  
609 °C then embedded in Optimal Cutting Temperature Compound and stored at -80 °C until  
610 cryosectioning. Tissue blocks were brought up to -20 °C in cryostat before sectioning into 30  $\mu$ m  
611 slices and dry-mounted onto slides after cryosectioning. After sectioning, the slides were left at  
612 room temperature overnight to dry. To assist in neuron quantification, we stained sections with  
613 the following antibodies and concentrations: rabbit anti-GFP (1:100; Millipore-Sigma, AB3080)  
614 and mouse anti-NeuN (1:500; Millipore-Sigma, MAB377). For secondary antibody staining, the  
615 following secondary antibodies and concentrations were used: donkey anti-rabbit Alexa Fluor 488  
616 (1:500; Invitrogen, A21206) and donkey anti-mouse Alexa Fluor 647 (1:500; Invitrogen, A31571).  
617 All antibodies were diluted with 1X PBS supplemented with 0.25% Triton X-100 (PBST) and 5%  
618 normal donkey serum. Primary antibody incubations were left overnight at room temperature.  
619 Sections were then washed with PBST. Secondary antibody incubations were 2 hours at room  
620 temperature. The sections were washed 3x in PBST. Sections were incubated in DAPI solution  
621 (1:10,000; Invitrogen, D1306) at room temperature for 5 minutes, then washed. Sections were  
622 coverslipped using Prolong Diamond Antifade.

623 3 sections per animal were stained and imaged. Each section was imaged in triplicate with  
624 each ROI having a total of 9 images. Tissue ROIs were imaged with a Keyence BZ-X800 with the  
625 following acquisition parameters: GFP (1/500 s), Cy5 (1 s), DAPI (1/12 s), High Resolution, Z-  
626 stack @ 1.2  $\mu$ m pitch. The following brain subregions were imaged frontal, parietal, temporal,  
627 occipital cortices, cerebellum, caudate, putamen, and thalamus (medial, ventral lateral, and  
628 ventral posterior nuclei). A semi-automated cell counting method was performed via ImageJ for  
629 quantification. Using thresholds and particle analysis, we were able to quantify NeuN positive and  
630 DAPI positive cells. Using ImageJ's cell counter, we manually counted GFP positive and GFP &  
631 NeuN double-positive cells.

### 632 **Induced pluripotent stem cell (iPSC) experiments**

633 Neuronal cultures were produced by differentiating and maturing iPSC-derived neural  
634 progenitor cells with Stemdiff™ Forebrain Differentiation and Maturation kits (StemCell #08600,  
635 #08605 respectively), according to their vendor protocols. Neural progenitor cells were produced  
636 by differentiation of the foreskin fibroblast-derived iPSC line: ACS™-1019 (ATCC# DYS-0100),  
637 with Stemdiff™ SMADi Neural Induction kits (StemCell I#08581), selection with Stemdiff™ Neural



638 Rosette Selection Reagent (StemCell l#05832), and expansion in Stemdiff™ Neural Progenitor  
639 Media (StemCell l#05833), according to their vendor protocols. Neurons were matured a minimum  
640 of 8 days prior to replating for transduction.

641 Mature neuronal cultures, seeded 15,000 cells/well in polyornithine and laminin coated,  
642 black-walled 96 well optical plates, were cultured an additional 4 days prior to transduction.  
643 Replicate wells were transduced with virus serially diluted across six orders of magnitude in 90%  
644 maturation media and 10% OptiproSFM. 4 days post-transduction, cultures were fixed with 4%  
645 paraformaldehyde and counterstained with 1 ug/ml Hoechst 33322. Identification of transduced  
646 cells was determined by imaging 60 fields/well, using two channel fluorescence detection  
647 (Hoechst at ex386/em440, eGFP ex485/em521) on a CellInsight CX5 HCS Platform. Individual  
648 cells were identified by Hoechst detection of their nuclei and applying size and contact constrained  
649 ring masks to each cell. Cell transduction was determined by measuring an eGFP fluorescence  
650 above a threshold level within an individual ring mask. For each population, the percentage of  
651 transduced cells was plotted vs the applied dose. Curve-fits and EC<sub>50</sub> values were determined  
652 with a Prism GraphPad [agonist] vs response (three parameter) regression method. To report per  
653 cell eGFP expression efficiencies, the eGFP spot fluorescence intensities were averaged from  
654 each ring mask across a minimum of 5000 cells/well. Curve fits were obtained using Prism  
655 GraphPad Biphasic, X as concentration regression method.  
656

## 657 References

- 658 1. Rose, J. A., Berns, K. I., Hoggan, M. D. & Koczot, F. J. EVIDENCE FOR A SINGLE-  
659 STRANDED ADENOVIRUS-ASSOCIATED VIRUS GENOME: FORMATION OF A DNA  
660 DENSITY HYBRID ON RELEASE OF VIRAL DNA. *Proc. Natl. Acad. Sci.* **64**, 863–869  
661 (1969).
- 662 2. Atchison, R. W., Casto, B. C. & Hammon, W. M. Adenovirus-Associated Defective Virus  
663 Particles. *Science (80-. )*. **149**, 754–756 (1965).
- 664 3. Hoggan, M. D., Blacklow, N. R. & Rowe, W. P. Studies of small DNA viruses found in  
665 various adenovirus preparations: physical, biological, and immunological characteristics.  
666 *Proc. Natl. Acad. Sci.* **55**, 1467–1474 (1966).
- 667 4. Tratschin, J. D., West, M. H., Sandbank, T. & Carter, B. J. A human parvovirus, adeno-  
668 associated virus, as a eucaryotic vector: transient expression and encapsidation of the  
669 procaryotic gene for chloramphenicol acetyltransferase. *Mol. Cell. Biol.* **4**, 2072–81 (1984).
- 670 5. Hermonat, P. L. & Muzyczka, N. Use of adeno-associated virus as a mammalian DNA  
671 cloning vector: transduction of neomycin resistance into mammalian tissue culture cells.  
672 *Proc. Natl. Acad. Sci. U. S. A.* **81**, 6466–6470 (1984).
- 673 6. Wang, D., Tai, P. W. L. & Gao, G. Adeno-associated virus vector as a platform for gene  
674 therapy delivery. *Nat. Rev. Drug Discov.* **18**, 358–378 (2019).
- 675 7. Zhu, D., Schieferecke, A. J., Lopez, P. A. & Schaffer, D. V. Adeno-Associated Virus Vector  
676 for Central Nervous System Gene Therapy. *Trends Mol. Med.* **27**, 524–537 (2021).
- 677 8. Samulski, R. J. & Muzyczka, N. AAV-mediated gene therapy for research and therapeutic  
678 purposes. *Annu. Rev. Virol.* **1**, 427–451 (2014).
- 679 9. Morris, J. A. *et al.* Next-generation strategies for gene-targeted therapies of central nervous  
680 system disorders: A workshop summary. *Mol. Ther.* **29**, 3332–3344 (2021).
- 681 10. Chand, D. *et al.* Hepatotoxicity following administration of onasemnogene abeparvovec  
682 (AVXS-101) for the treatment of spinal muscular atrophy. *J. Hepatol.* **74**, 560–566 (2021).
- 683 11. Hinderer, C. *et al.* Severe Toxicity in Nonhuman Primates and Piglets Following High-Dose  
684 Intravenous Administration of an Adeno-Associated Virus Vector Expressing Human SMN.  
685 *Hum. Gene Ther.* **29**, 285–298 (2018).
- 686 12. Harrison, C. High-dose AAV gene therapy deaths. *Nat. Biotechnol.* **38**, 910–910 (2020).
- 687 13. Morales, L., Gambhir, Y., Bennett, J. & Stedman, H. H. Broader Implications of Progressive  
688 Liver Dysfunction and Lethal Sepsis in Two Boys following Systemic High-Dose AAV. *Mol.*  
689 *Ther.* **28**, 1753–1755 (2020).
- 690 14. Deverman, B. E. *et al.* Cre-dependent selection yields AAV variants for widespread gene  
691 transfer to the adult brain. *Nat. Biotechnol.* **34**, 204–209 (2016).
- 692 15. Chan, K. Y. *et al.* Engineered AAVs for efficient noninvasive gene delivery to the central  
693 and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017).
- 694 16. Ravindra Kumar, S. *et al.* Multiplexed Cre-dependent selection yields systemic AAVs for  
695 targeting distinct brain cell types. *Nat. Methods* **17**, 541–550 (2020).
- 696 17. Goertsen, D. *et al.* AAV capsid variants with brain-wide transgene expression and  
697 decreased liver targeting after intravenous delivery in mouse and marmoset. *Nat. Neurosci.*  
698 **1–21** (2021). doi:10.1038/s41593-021-00969-4
- 699 18. Dana, H. *et al.* High-performance calcium sensors for imaging activity in neuronal  
700 populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).
- 701 19. Phillips, K. A. *et al.* Why primate models matter. *Am. J. Primatol.* **76**, 801–827 (2014).
- 702 20. Stonebarger, G. A., Bimonte-Nelson, H. A. & Urbanski, H. F. The Rhesus Macaque as a  
703 Translational Model for Neurodegeneration and Alzheimer’s Disease. *Front. Aging*  
704 *Neurosci.* **13**, 1–8 (2021).
- 705 21. Gray, D. T. & Barnes, C. A. Experiments in macaque monkeys provide critical insights into  
706 age-associated changes in cognitive and sensory function. *Proc. Natl. Acad. Sci. U. S. A.*

- 707 **116**, 26247–26254 (2019).
- 708 22. Choudhury, S. R. *et al.* Widespread central nervous system gene transfer and silencing  
709 after systemic delivery of novel AAV-AS vector. *Mol. Ther.* **24**, 726–735 (2016).
- 710 23. Nonnenmacher, M. *et al.* Rapid evolution of blood-brain-barrier-penetrating AAV capsids  
711 by RNA-driven biopanning. *Mol. Ther. - Methods Clin. Dev.* **20**, 366–378 (2021).
- 712 24. Liguore, W. A. *et al.* AAV-PHP.B Administration Results in a Differential Pattern of CNS  
713 Biodistribution in Non-human Primates Compared with Mice. *Mol. Ther.* **27**, 2018–2037  
714 (2019).
- 715 25. Hordeaux, J. *et al.* The Neurotropic Properties of AAV-PHP.B Are Limited to C57BL/6J  
716 Mice. *Mol. Ther.* **26**, 664–668 (2018).
- 717 26. De, A., El-Shamayleh, Y. & Horwitz, G. D. Fast and reversible neural inactivation in  
718 macaque cortex by optogenetic stimulation of GABAergic neurons. *Elife* **9**, 1–21 (2020).
- 719 27. Galvan, A., Hu, X., Smith, Y. & Wichmann, T. Effects of optogenetic activation of  
720 corticothalamic terminals in the motor thalamus of awake monkeys. *J. Neurosci.* **36**, 3519–  
721 3530 (2016).
- 722 28. El-Shamayleh, Y., Kojima, Y., Soetedjo, R. & Horwitz, G. D. Selective Optogenetic Control  
723 of Purkinje Cells in Monkey Cerebellum. *Neuron* **95**, 51-62.e4 (2017).
- 724 29. Watanabe, H. *et al.* Forelimb movements evoked by optogenetic stimulation of the  
725 macaque motor cortex. *Nat. Commun.* **11**, 1–9 (2020).
- 726 30. Oguchi, M. *et al.* Microendoscopic calcium imaging of the primary visual cortex of behaving  
727 macaques. *Sci. Rep.* **11**, 1–15 (2021).
- 728 31. Bollimunta, A. *et al.* Head-mounted microendoscopic calcium imaging in dorsal premotor  
729 cortex of behaving rhesus macaque. *Cell Rep.* **35**, 109239 (2021).
- 730 32. Weiss, A. R., Liguore, W. A., Domire, J. S., Button, D. & McBride, J. L. Intra-striatal  
731 AAV2.retro administration leads to extensive retrograde transport in the rhesus macaque  
732 brain: implications for disease modeling and therapeutic development. *Sci. Rep.* **10**, 1–14  
733 (2020).
- 734 33. Yazdan-Shahmorad, A. *et al.* A Large-Scale Interface for Optogenetic Stimulation and  
735 Recording in Nonhuman Primates. *Neuron* **89**, 927–939 (2016).
- 736 34. Bankiewicz, K. S. *et al.* Convection-enhanced delivery of AAV vector in Parkinsonian  
737 monkeys; in vivo detection of gene expression and restoration of dopaminergic function  
738 using pro-drug approach. *Exp. Neurol.* **164**, 2–14 (2000).
- 739 35. Kells, A. P. *et al.* Efficient gene therapy-based method for the delivery of therapeutics to  
740 primate cortex. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 2407–2411 (2009).
- 741 36. Salegio, E. A., Samaranch, L., Kells, A. P., Forsayeth, J. & Bankiewicz, K. Guided delivery  
742 of adeno-associated viral vectors into the primate brain. *Adv. Drug Deliv. Rev.* **64**, 598–  
743 604 (2012).
- 744 37. Fetsch, C. R. *et al.* Focal optogenetic suppression in macaque area MT biases direction  
745 discrimination and decision confidence, but only transiently. *Elife* **7**, 1–23 (2018).
- 746 38. Yazdan-Shahmorad, A. *et al.* Widespread optogenetic expression in macaque cortex  
747 obtained with MR-guided, convection enhanced delivery (CED) of AAV vector to the  
748 thalamus. *J. Neurosci. Methods* **293**, 347–358 (2018).
- 749 39. Stauffer, W. R. *et al.* Dopamine Neuron-Specific Optogenetic Stimulation in Rhesus  
750 Macaques. *Cell* **166**, 1564-1571.e6 (2016).
- 751 40. Gray, S. J., Nagabhashan Kalburgi, S., McCown, T. J. & Jude Samulski, R. Global CNS  
752 gene delivery and evasion of anti-AAV-neutralizing antibodies by intrathecal AAV  
753 administration in non-human primates. *Gene Ther.* **20**, 450–459 (2013).
- 754 41. Samaranch, L. *et al.* Adeno-associated virus serotype 9 transduction in the central nervous  
755 system of nonhuman primates. *Hum. Gene Ther.* **23**, 382–389 (2011).
- 756 42. Arotcarena, M.-L. *et al.* Pilot Study Assessing the Impact of Intrathecal Administration of  
757 Variants AAV-PHP.B and AAV-PHP.eB on Brain Transduction in Adult Rhesus Macaques.

- 758 *Front. Bioeng. Biotechnol.* **9**, 1–10 (2021).
- 759 43. Hinderer, C. *et al.* Widespread gene transfer in the central nervous system of cynomolgus  
760 macaques following delivery of AAV9 into the cisterna magna. *Mol. Ther. - Methods Clin.*  
761 *Dev.* **1**, 14051 (2014).
- 762 44. Bey, K. *et al.* Intra-CSF AAV9 and AAVrh10 Administration in Nonhuman Primates:  
763 Promising Routes and Vectors for Which Neurological Diseases? *Mol. Ther. - Methods*  
764 *Clin. Dev.* **17**, 771–784 (2020).
- 765 45. Meseck, E. K. *et al.* Intrathecal sc-AAV9-CB-GFP: Systemic Distribution Predominates  
766 Following Single-Dose Administration in Cynomolgus Macaques. *bioRxiv*  
767 2021.11.28.470258 (2021).
- 768 46. Hordeaux, J. *et al.* Toxicology Study of Intra-Cisterna Magna Adeno-Associated Virus 9  
769 Expressing Human Alpha-L-Iduronidase in Rhesus Macaques. *Mol. Ther. - Methods Clin.*  
770 *Dev.* **10**, 79–88 (2018).
- 771 47. Hordeaux, J. *et al.* Toxicology Study of Intra-Cisterna Magna Adeno-Associated Virus 9  
772 Expressing Iduronate-2-Sulfatase in Rhesus Macaques. *Mol. Ther. - Methods Clin. Dev.*  
773 **10**, 68–78 (2018).
- 774 48. Kondratov, O. *et al.* A comprehensive study of a 29-capsid AAV library in a non-human  
775 primate central nervous system. *Mol. Ther.* **29**, 2806–2820 (2021).
- 776 49. Cai, D., Cohen, K. B., Luo, T., Lichtman, J. W. & Sanes, J. R. Improved tools for the  
777 Brainbow toolbox. *Nat. Methods* **10**, 540–547 (2013).
- 778 50. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in  
779 the nervous system. *Nature* **450**, 56–62 (2007).
- 780 51. Brown, D. *et al.* Deep Parallel Characterization of AAV Tropism and AAV-Mediated  
781 Transcriptional Changes via Single-Cell RNA Sequencing. *Front. Immunol.* **12**, 1–24  
782 (2021).
- 783 52. McColgan, P., Joubert, J., Tabrizi, S. J. & Rees, G. The human motor cortex microcircuit:  
784 insights for neurodegenerative disease. *Nat. Rev. Neurosci.* **21**, 401–415 (2020).
- 785 53. Tabrizi, S. J., Flower, M. D., Ross, C. A. & Wild, E. J. Huntington disease: new insights into  
786 molecular pathogenesis and therapeutic opportunities. *Nat. Rev. Neurol.* **16**, 529–546  
787 (2020).
- 788 54. Hordeaux, J. *et al.* The GPI-Linked Protein LY6A Drives AAV-PHP.B Transport across the  
789 Blood-Brain Barrier. *Mol. Ther.* **27**, 912–921 (2019).
- 790 55. Huang, Q. *et al.* Delivering genes across the blood-brain barrier: LY6A, a novel cellular  
791 receptor for AAV-PHP.B capsids. *PLoS One* **14**, 1–17 (2019).
- 792 56. Bevan, A. K. *et al.* Systemic gene delivery in large species for targeting spinal cord, brain,  
793 and peripheral tissues for pediatric disorders. *Mol. Ther.* **19**, 1971–1980 (2011).
- 794 57. Gray, S. J. *et al.* Preclinical differences of intravascular aav9 delivery to neurons and glia:  
795 A comparative study of adult mice and nonhuman primates. *Mol. Ther.* **19**, 1058–1069  
796 (2011).
- 797 58. Herculano-Houzel, S. *et al.* Updated neuronal scaling rules for the brains of glires  
798 (rodents/lagomorphs). *Brain. Behav. Evol.* **78**, 302–314 (2011).
- 799 59. Keller, D., Erö, C. & Markram, H. Cell densities in the mouse brain: A systematic review.  
800 *Front. Neuroanat.* **12**, (2018).
- 801 60. Brandt, L., Cristinelli, S. & Ciuffi, A. Single-Cell Analysis Reveals Heterogeneity of Virus  
802 Infection, Pathogenicity, and Host Responses: HIV as a Pioneering Example. *Annu. Rev.*  
803 *Viol.* **7**, 333–350 (2020).
- 804 61. Suomalainen, M. & Greber, U. F. Virus infection variability by single-cell profiling. *Viruses*  
805 **13**, (2021).
- 806 62. Mitchell, J. F. & Leopold, D. A. The marmoset as a model for visual neuroscience. *Common*  
807 *Marmoset Captiv. Biomed. Res.* 377–413 (2018). doi:10.1016/B978-0-12-811829-  
808 0.00022-4



- 809 63. Sasaki, E. *et al.* Generation of transgenic non-human primates with germline transmission.  
810 *Nature* **459**, 523–527 (2009).
- 811 64. Park, J. E. *et al.* Generation of transgenic marmosets expressing genetically encoded  
812 calcium indicators. *Sci. Rep.* **6**, 1–12 (2016).
- 813 65. Okano, H. *et al.* Brain/MINDS: A Japanese National Brain Project for Marmoset  
814 Neuroscience. *Neuron* **92**, 582–590 (2016).
- 815 66. Saunders, N. R. *et al.* The rights and wrongs of blood-brain barrier permeability studies: A  
816 walk through 100 years of history. *Front. Neurosci.* **8**, 1–26 (2014).
- 817 67. Abbott, N. J., Patabendige, A. A. K., Dolman, D. E. M., Yusof, S. R. & Begley, D. J.  
818 Structure and function of the blood–brain barrier. *Neurobiol. Dis.* **37**, 13–25 (2010).
- 819 68. Blanchette, M. & Daneman, R. Formation and maintenance of the BBB. *Mech. Dev.* **138**,  
820 8–16 (2015).
- 821 69. Daneman, R. *et al.* The Mouse Blood-Brain Barrier Transcriptome: A New Resource for  
822 Understanding the Development and Function of Brain Endothelial Cells. *PLoS One* **5**,  
823 e13741 (2010).
- 824 70. Daneman, R., Zhou, L., Kebede, A. A. & Barres, B. A. Pericytes are required for blood–  
825 brain barrier integrity during embryogenesis. *Nature* **468**, 562–566 (2010).
- 826 71. Ek, C. J., Dziegielewska, K. M., Stolp, H. & Saunders, N. R. Functional effectiveness of the  
827 blood-brain barrier to small water-soluble molecules in developing and adult opossum  
828 (*Monodelphis domestica*). *J. Comp. Neurol.* **496**, 13–26 (2006).
- 829 72. Saunders, N. R., Liddelow, S. A. & Dziegielewska, K. M. Barrier Mechanisms in the  
830 Developing Brain. *Front. Pharmacol.* **3**, 1–18 (2012).
- 831 73. Genç, B. *et al.* Apical dendrite degeneration, a novel cellular pathology for Betz cells in  
832 ALS. *Sci. Rep.* **7**, 1–10 (2017).
- 833 74. Tsutsui-Kimura, I. *et al.* Dysfunction of ventrolateral striatal dopamine receptor type 2-  
834 expressing medium spiny neurons impairs instrumental motivation. *Nat. Commun.* **8**, 1–13  
835 (2017).
- 836 75. Ngai, J. BRAIN 2.0: Transforming neuroscience. *Cell* **185**, 4–8 (2022).
- 837 76. Yang, B. *et al.* Single-cell phenotyping within transparent intact tissue through whole-body  
838 clearing. *Cell* **158**, 945–958 (2014).
- 839 77. Xu, F. *et al.* High-throughput mapping of a whole rhesus monkey brain at micrometer  
840 resolution. *Nat. Biotechnol.* (2021). doi:10.1038/s41587-021-00986-5
- 841 78. Todorov, M. I. *et al.* Machine learning analysis of whole mouse brain vasculature. *Nat.*  
842 *Methods* **17**, 442–449 (2020).
- 843 79. McBride, J. L. *et al.* Discovery of a CLN7 model of Batten disease in non-human primates.  
844 *Neurobiol. Dis.* **119**, 65–78 (2018).
- 845 80. Lallani, S. B., Villalba, R. M., Chen, Y., Smith, Y. & Chan, A. W. S. Striatal Interneurons in  
846 Transgenic Nonhuman Primate Model of Huntington’s Disease. *Sci. Rep.* **9**, 1–9 (2019).
- 847 81. Challis, R. C. *et al.* Systemic AAV vectors for widespread and targeted gene delivery in  
848 rodents. *Nat. Protoc.* **14**, 379–414 (2019).
- 849 82. Yardeni, T., Eckhaus, M., Morris, H. D., Huizing, M. & Hoogstraten-Miller, S. Retro-orbital  
850 injections in mice. *Lab Anim. (NY)*. **40**, 155–160 (2011).
- 851 83. Treweek, J. B. *et al.* Whole-body tissue stabilization and selective extractions via tissue-  
852 hydrogel hybrids for high-resolution intact circuit mapping and phenotyping. *Nat. Protoc.*  
853 **10**, 1860–1896 (2015).
- 854

## 855 Acknowledgements

856 We wish to thank the entire Gradinaru laboratory and members of Capsida Biotherapeutics for  
857 helpful discussions. We thank Deborah Lidgate for her helpful input in writing and planning the

858 paper. We thank Catherine Oikonomou for help with manuscript editing. We thank Máté Borsos  
859 for his assistance in mouse pup injections. We thank Xinhong Chen, Anat Kahan, and Gerard M.  
860 Coughlin for their helpful discussion and input in planning rhesus macaque experiments and  
861 experimental cargo. Capsida would like to thank Michael Weed and the whole team at Virscio,  
862 Inc. for their help with the design and execution of green monkey experiments. We are grateful to  
863 Igor Antoshechkin and the Millar and Muriel Jacobs Genetics and Genomics Core at the California  
864 Institute of Technology for assistance with next-generation sequencing. We are grateful to the  
865 research and veterinarian staff at the California National Primate Research Center (CNPRC) for  
866 their aid with studies in rhesus macaques. This work was funded by grants from the National  
867 Institutes of Health (NIH) to V.G. (NIH Pioneer DP1OD025535), to the California National Primate  
868 Research Center (NIH: P51OD011107), and BRAIN Armamentarium U01 UMH128336A (to V.G.,  
869 T.F.M., and A.S.F.). Additional funding includes: (to V.G. and A.S.F.) The Michael J. Fox  
870 Foundation for Parkinson's Research as part of the Aligning Science Across Parkinson's Initiative  
871 (ASAP-020495). Figures were created using images from BioRender.com.

## 872 Author contributions

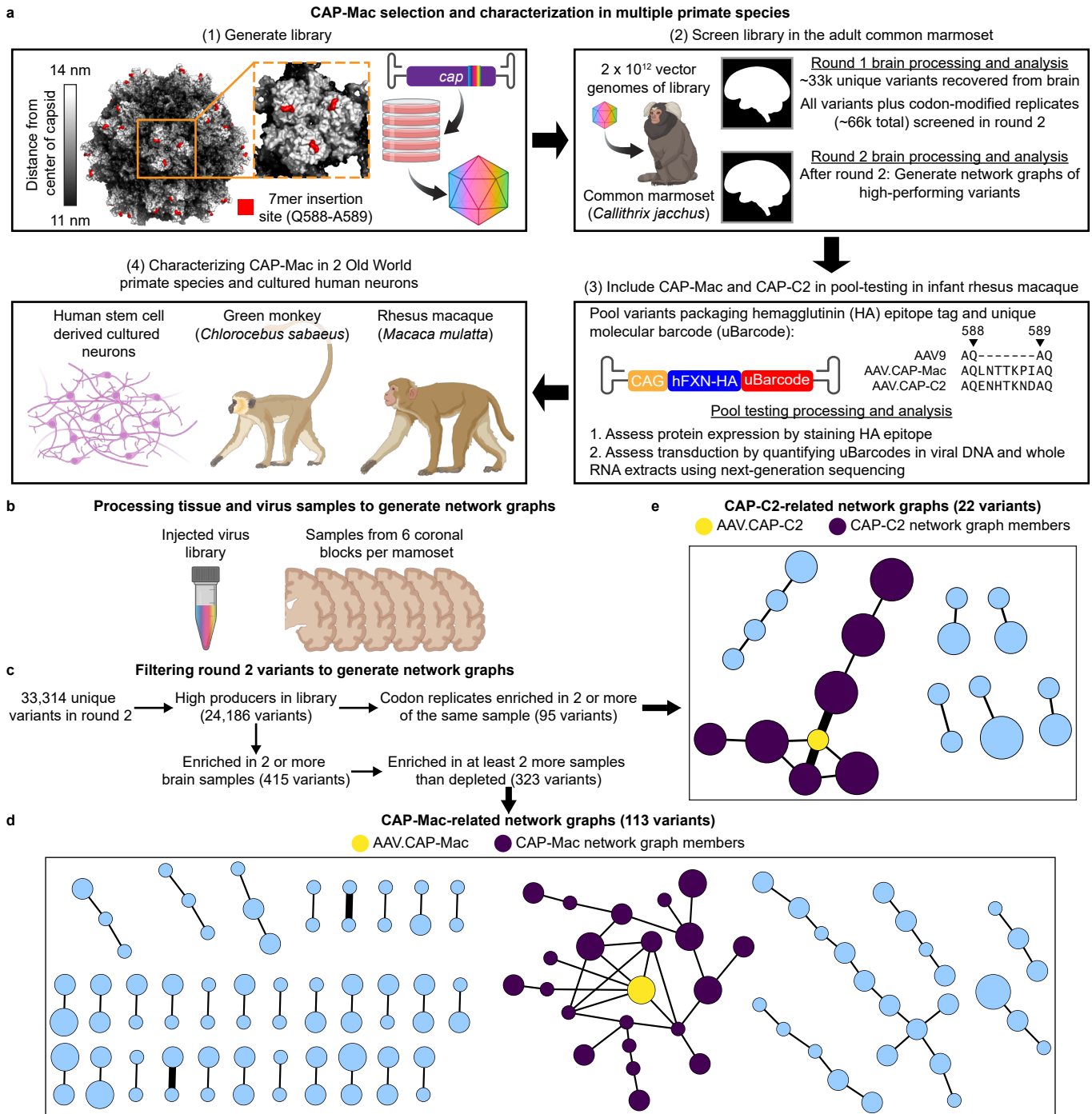
873 M.R.C. and N.C.F. wrote the manuscript, with input from all authors. M.R.C. designed, performed,  
874 and analyzed the data for the rhesus macaque and rodent experiments and prepared all figures.  
875 N.C.F., K.C. and B.E.D. designed, N.C.F. and K.C. performed, and N.C.F. analyzed the  
876 associated data for the viral library screening experiments in common marmosets. N.C.F. and  
877 N.G. designed, performed and analyzed the data of the pooled testing experiments in rhesus  
878 macaques. N.C.F. and N.G. designed, N.C.F. N.G., J.C.O. and K.M.R. performed the green  
879 monkey experiments, and J.C.O. and K.M.R. analyzed the associated data and helped prepare  
880 the associated figures. J.S., J.W. and R.J.B. designed and performed the human neuron  
881 experiments, analyzed the associated data and prepared the associated figures. L.J.C. designed  
882 and performed the rhesus macaque experiments. C.M.A. performed the rhesus macaque spinal  
883 cord and dorsal root ganglia analysis and imaging. T.F.M. analyzed the round 2 viral library  
884 screening experiment in the common marmoset and generated the sequence clustering graphs.  
885 M.J.J. helped with the imaging analysis. J.V. helped perform the rhesus macaque neuron tracing.  
886 J.P. supervised aspects of the common marmoset experiments. A.S.F. designed, performed, and  
887 supervised aspects of the rhesus macaque experiments. N.C.F. supervised all aspects of the  
888 green monkey and iPSC work. V.G. supervised all aspects of the library screening, pooled testing  
889 and rhesus macaque work and contributed to associated experimental design, data analysis, and  
890 manuscript writing.

## 891 Competing interests

892 The California Institute of Technology has filed and licensed patent applications for the work  
893 described in this manuscript, with N.C.F., N.G. and V.G. listed as inventors (US Patent application  
894 no. PCT/US21/46904). V.G. is a co-founder and board member and N.C.F. and N.G. are co-  
895 founders and officers of Capsida Biotherapeutics, a fully integrated AAV engineering and gene  
896 therapy company. The remaining authors declare no competing interests.

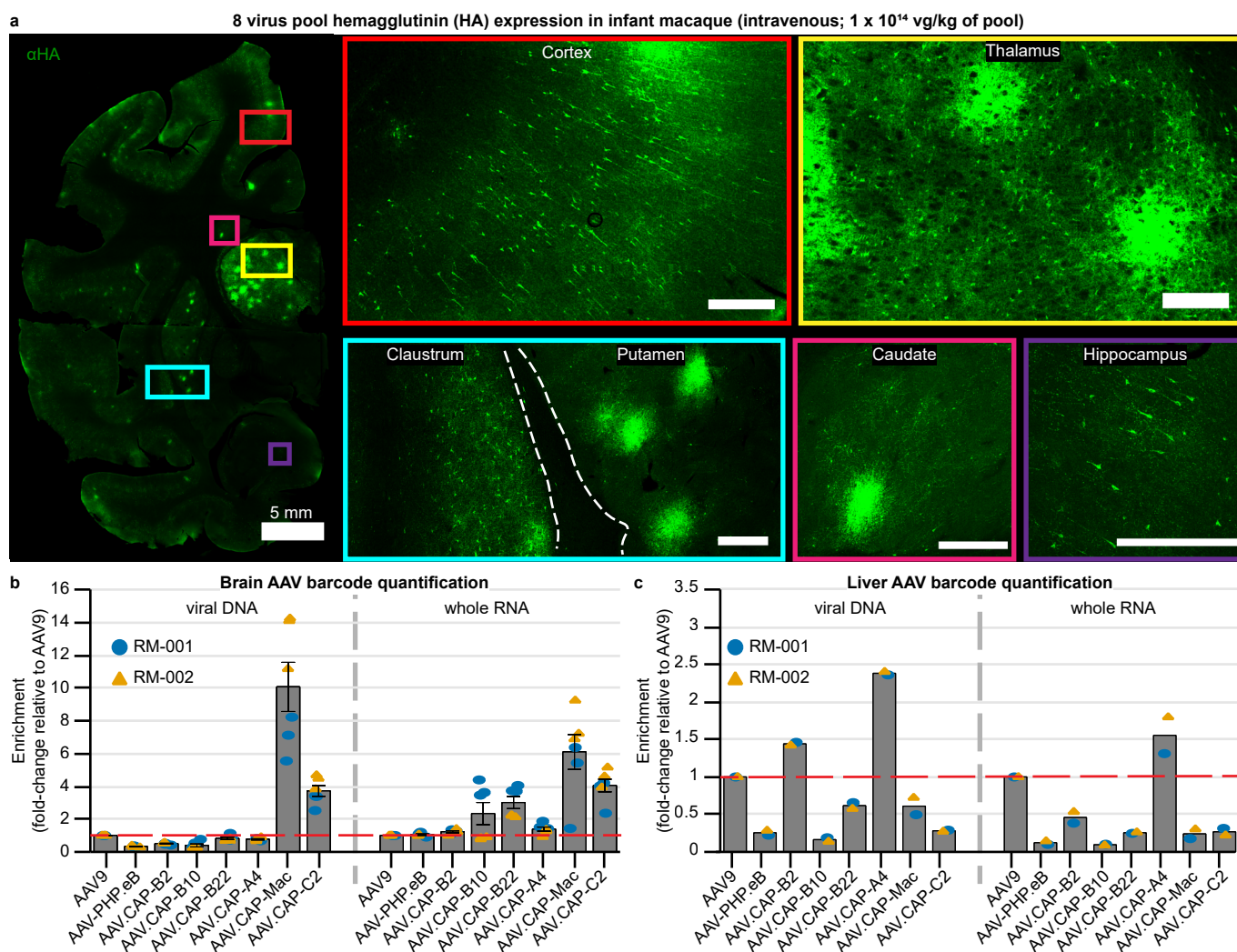


**Fig. 1: CAP-Mac selection and characterization strategy**



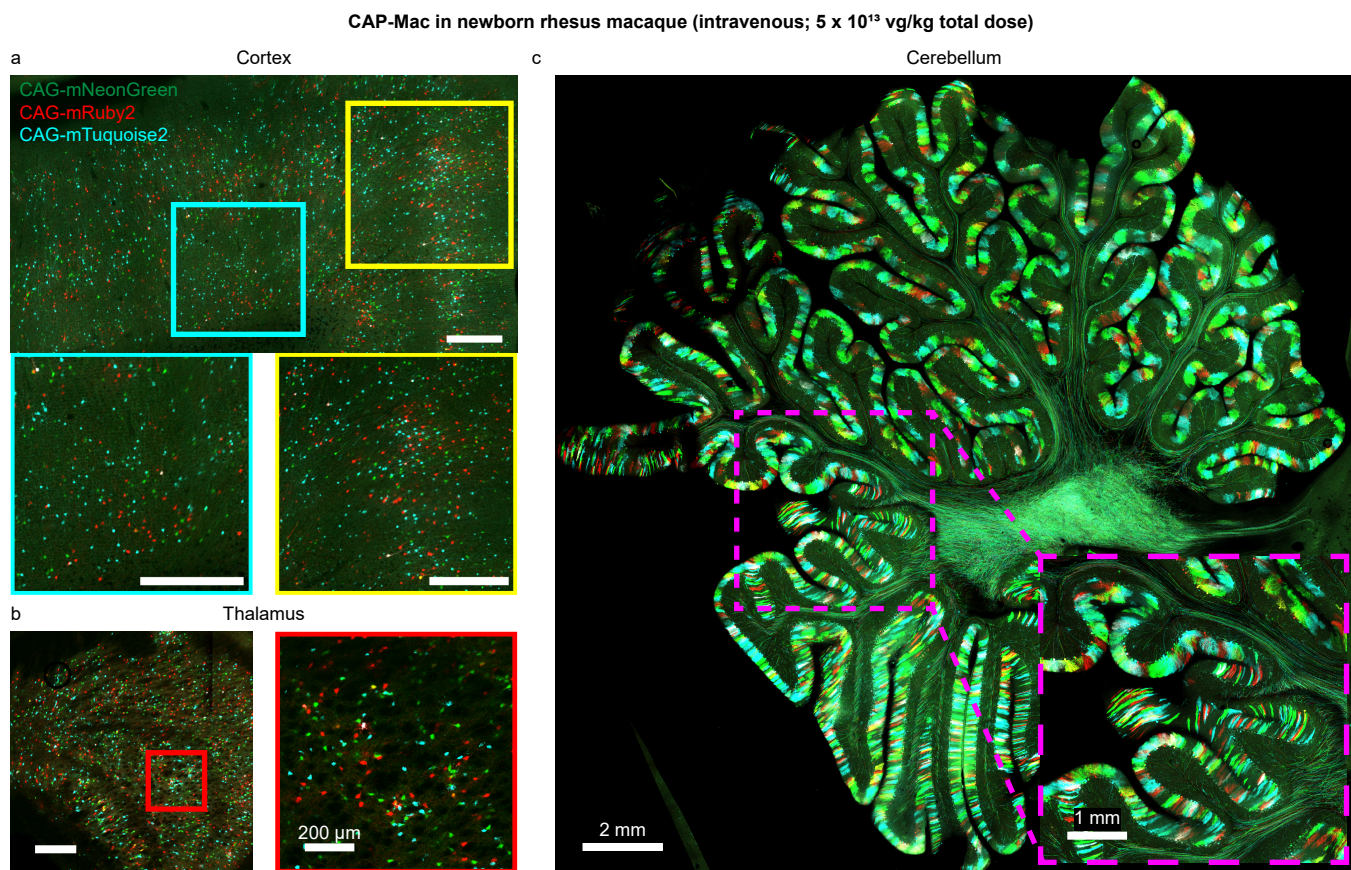
**Fig. 1: CAP-Mac selection and characterization strategy.** **a**, Schematic of the CAP-Mac selection strategy. (1) CAP-Mac is an AAV9 variant that we selected from a library screened in the adult common marmoset. We generated initial diversity by introducing 21 NNK degenerate codons after Q588 in the AAV9 cap gene. (2) We produced and screened the capsid library in 4 adult male marmosets (2 marmosets per round; 2 x 10<sup>12</sup> vector genomes of library administered intravenously in each marmoset). After the first round of selection, we recovered 33,314 unique amino acid sequences in the brain. For the second round of selection, we generated a synthetic oligo pool containing 66,628 sequences (the ~33k unique variants plus a codon modified replicate). After the second round of selection, we constructed network graphs of high-performing variants, and selected two capsids—AAV.CAP-Mac and AAV.CAP-C2—to be included in pool selections in infant rhesus macaques. (3) For pool selections, we packaged ssCAG-hFXN-HA with a unique molecular barcode (uBarcode) in the 3' UTR into 8 different capsids. The construct design enabled us to assess protein expression of the pool by staining for the hemagglutinin (HA) epitope and quantify barcodes in viral DNA and whole RNA extracts. (4) We then individually characterized AAV.CAP-Mac in 2 Old World primate species as well as human cultured neurons. **b**, To generate network graphs, we processed the injected virus library and sampled from each of the 6 brain sections from each animal. **c**, The filtering criteria used to identify high-performing variants. **d**, **e**, Network graphs for AAV.CAP-Mac (**d**) and AAV.CAP-C2 (**e**). Each node represents a unique variant recovered from the round 2 selection and each edge represents pairwise reverse Hamming distance  $\geq 3$ .

**Fig. 2: CAP-Mac outperforms other engineered variants in infant rhesus macaque in pool testing.**



**Fig. 2: CAP-Mac outperforms other engineered variants in infant rhesus macaque in pool testing. a,** Representative images of expression in cortex, thalamus, caudate nucleus, putamen, hippocampus and claustrum after intravenous administration of  $1 \times 10^{14}$  vg/kg of an 8-capsid pool ( $1.25 \times 10^{13}$  vg/kg of each variant) packaging hemagglutinin (HA) tagged human frataxin with unique barcode in each capsid. **b, c,** Unique barcode enrichments in viral DNA (left) and whole RNA (right) extracts from the brain (**b**) and the liver (**c**) of two newborn rhesus macaques. Each data point represents the fold-change relative to AAV9 within each sample of tissue. Mean  $\pm$  s.e.m. shown. The red dotted line denotes AAV9 performance in pool. All scalebars = 500  $\mu$ m unless otherwise stated.

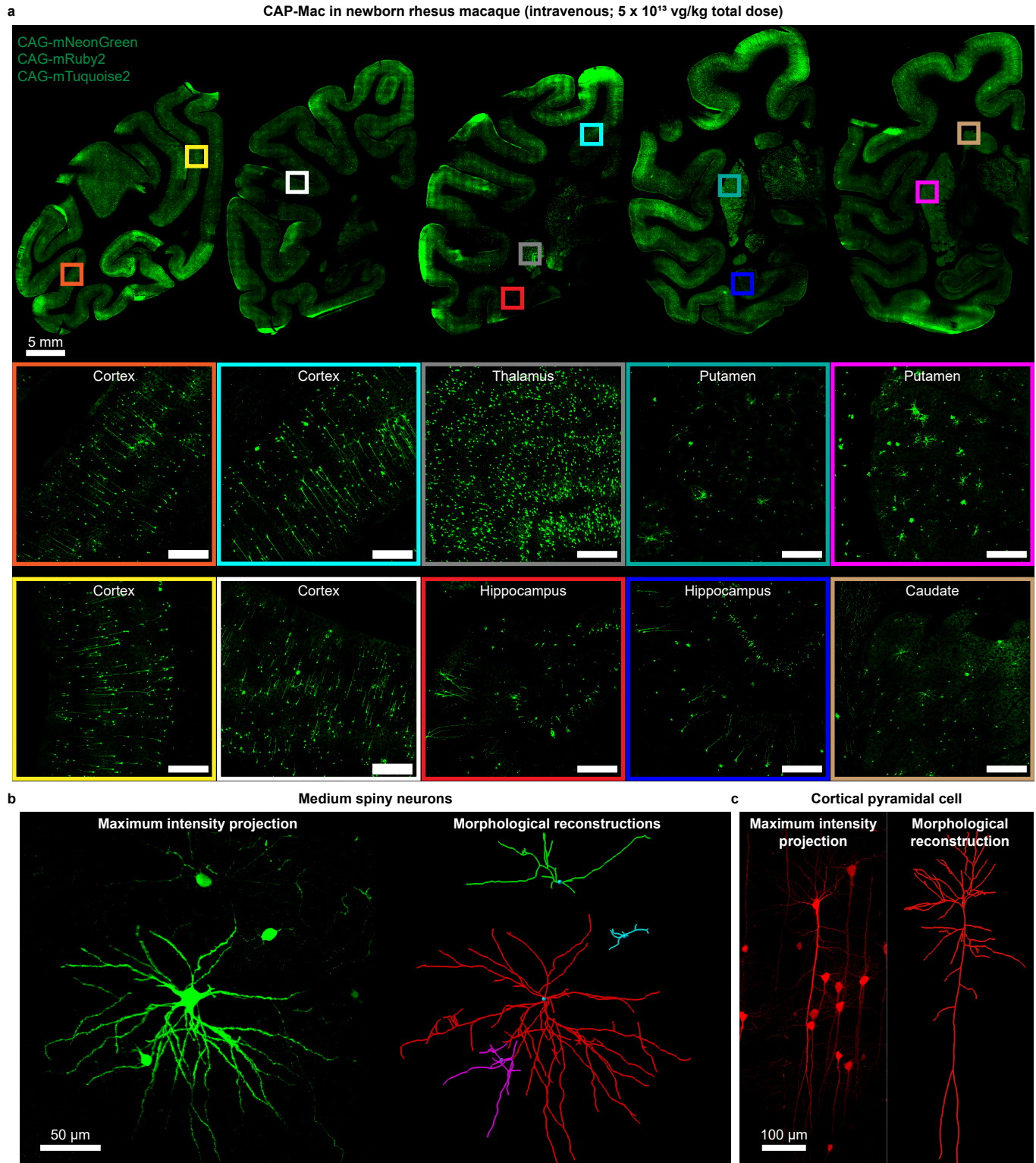
**Fig. 3: Brain-wide multicolor labeling of neurons in cortex, thalamus, and cerebellum after IV administering CAP-Mac packaging 3 fluorescent proteins in newborn rhesus macaque.**



**Fig. 3: Brain-wide multicolor labeling of neurons in cortex, thalamus, and cerebellum after IV administration of CAP-Mac packaging 3 fluorescent proteins. a, b, c, Representative images of cortex (a), thalamus (lateral geniculate nucleus) (b), and cerebellum (c) demonstrating widespread expression of a cocktail of 3 fluorescent proteins packaged in CAP-Mac ( $5 \times 10^{13}$  vg/kg total dose via intravenous administration). All scalebars = 500  $\mu$ m unless otherwise stated.**

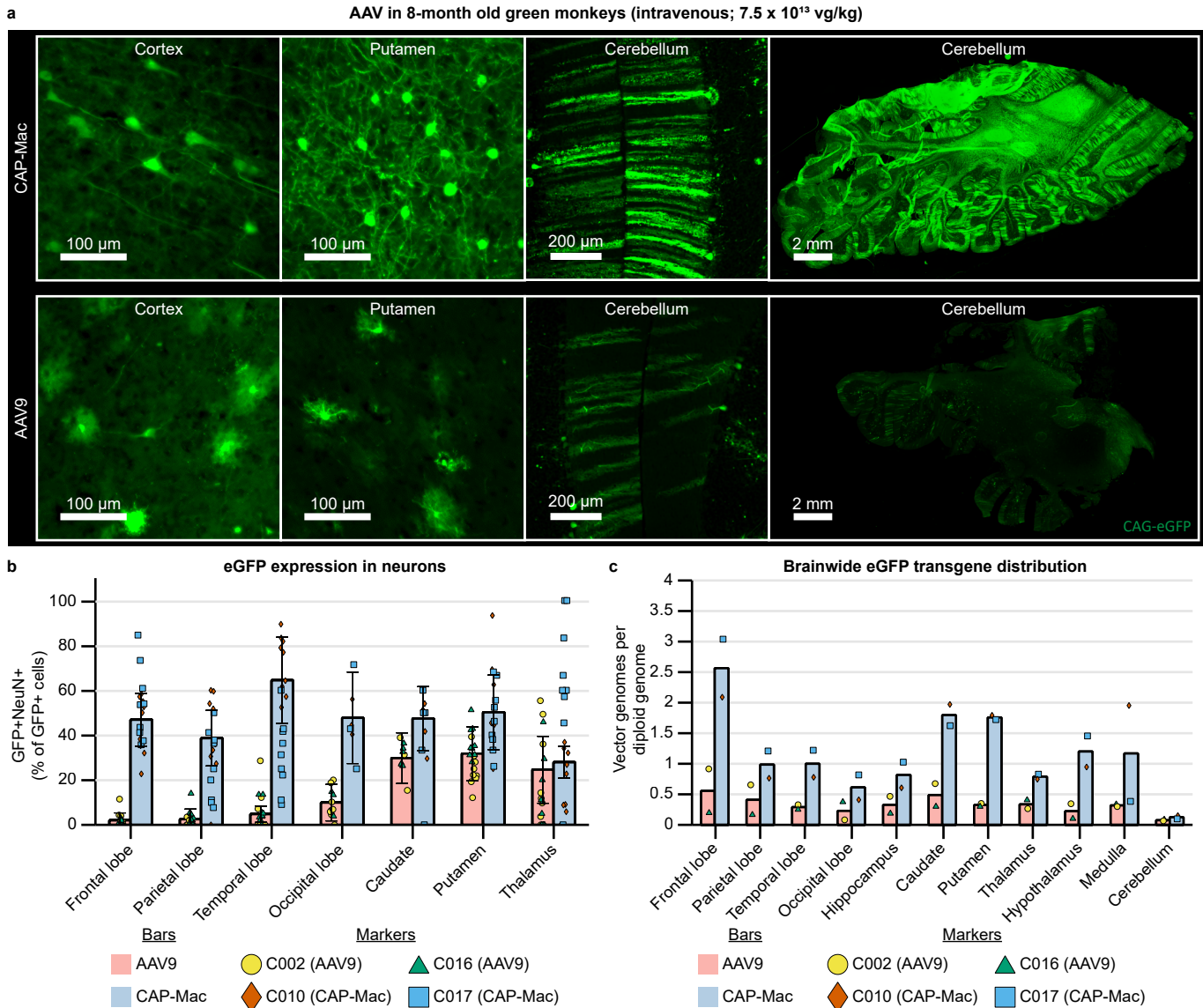


**Fig. 4: CAP-Mac transduces neurons throughout the newborn rhesus macaque brain, enabling morphological reconstruction of disease-relevant cell-types.**



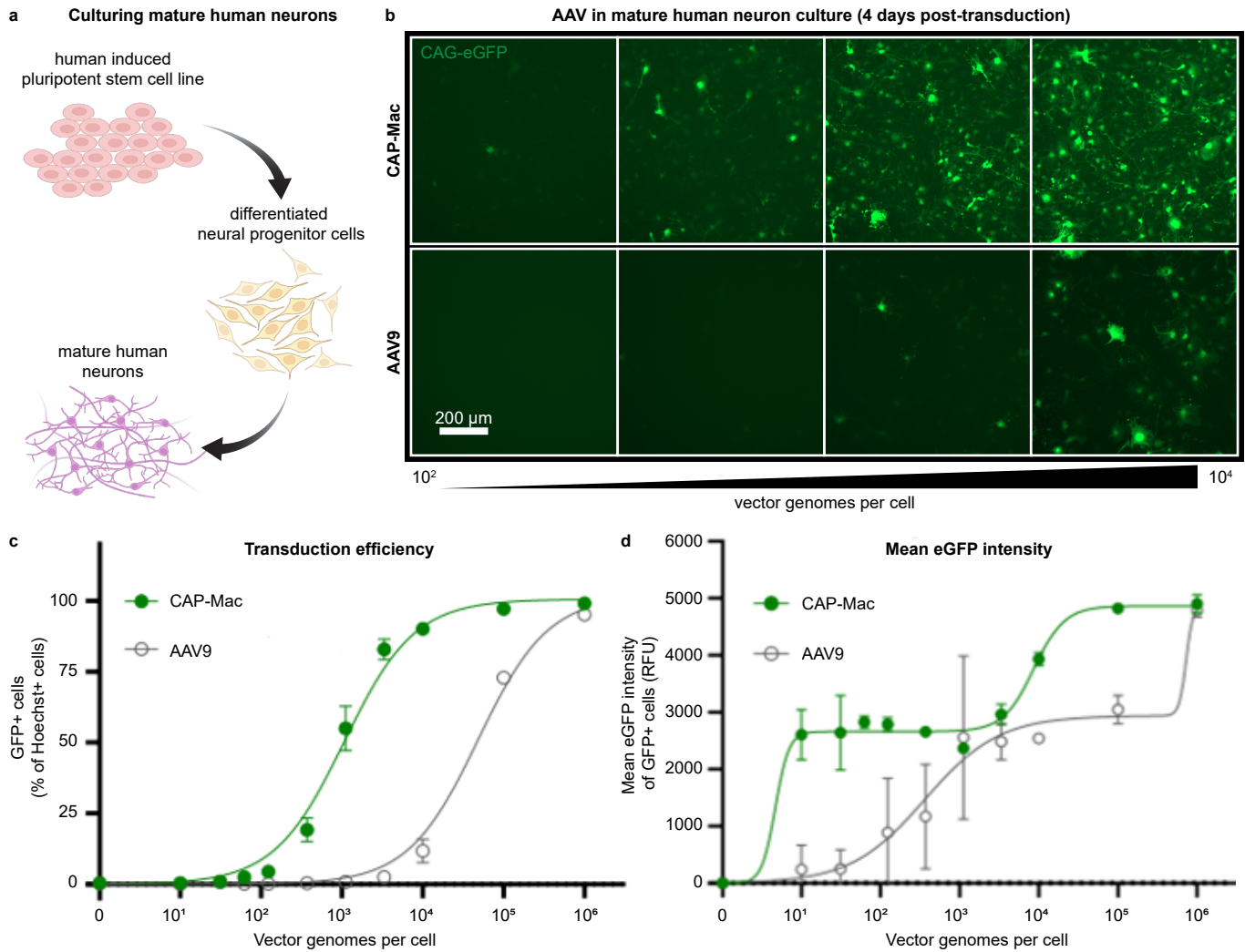
**Fig. 4: CAP-Mac transduces neurons throughout the newborn rhesus macaque brain, enabling morphological reconstruction of disease-relevant cell types.** **a**, Distribution of CAP-Mac expression across coronal slices showing fluorescent protein expression in cortical and subcortical brain regions (insets). Fluorescent proteins are identically pseudocolored. **b**, **c**, Morphological reconstruction of rhesus macaque medium spiny neurons (**b**) and a cortical pyramidal cell (**c**) in 300  $\mu$ m sections immersed in RIMS, enabled by intravenous administration of CAP-Mac packaging fluorescent protein. All scalebars = 500  $\mu$ m unless otherwise stated.

**Fig. 5: CAP-Mac transduces the green monkey brain more efficiently than AAV9 and is biased towards neurons.**



**Fig. 5: CAP-Mac transduces the green monkey brain more efficiently than AAV9 and is biased towards neurons.** **a**, Representative images of various brain regions from green monkeys dosed with CAP-Mac (top) or AAV9 (bottom) packaging ssCAG-eGFP ( $7.5 \times 10^{13}$  vg/kg via intravenous administration). **b**, NeuN quantification of GFP-transduced cells in images from 7 brain regions of green monkeys. Cells that are positive for both GFP and NeuN are expressed as a percentage of total GFP+ cells in that field of view. Each data point represents quantification of one field of view per condition. **c**, Distribution of CAP-Mac and AAV9-delivered eGFP transgene in 11 brain regions of green monkeys. Each data point represents measured vector genomes per diploid genome in a piece of tissue from each condition. Red bars: mean AAV9 values. Blue bars: mean CAP-Mac values. Yellow circles: measurements from AAV9-treated monkey, C002. Green triangles: measurements from AAV9-treated monkey, C016. Orange diamonds: measurements from CAP-Mac-treated monkey, C010. Blue squares: measurements from CAP-Mac-treated monkey, C017. Mean  $\pm$  s.e.m. shown (s.e.m. only calculated for samples with  $n > 2$ ).

**Fig. 6: CAP-Mac is more potent at transducing human cultured neurons compared to AAV9.**



**Fig. 6: CAP-Mac is more potent at transducing human cultured neurons compared to AAV9.** **a**, Differentiation process starting with a human induced pluripotent stem cell line that was differentiated into neural progenitor cells, which were further differentiated into mature neurons. **b**, Representative images of cultured human neurons after 4 days of incubation with either CAP-Mac (top) or AAV9 (bottom) packaging CAG-eGFP across 4 doses of AAV, ranging from  $10^2$ - $10^4$  vector genomes per cell. **c**, **d**, Dose response curves of AAV9 and CAP-Mac in mature human neuron culture measuring transduction efficiency (**c**) and mean eGFP intensity (**d**).