Post-translational Tuning of Human Cortical Progenitor Neuronal Output

Julien Pigeon[1], Tamina Dietl[2], Myriame Abou Mrad[1], Ludovico Rizzuti[2], Miguel V. Silva[3], Natasha Danda[1], Corentine Marie[1], Clarisse Brunet Avalos[4], Hayat Mokrani[5], Laila El Khattabi[1], Alexandre D. Baffet[4], Diogo S. Castro[3], Carlos Parras[1], Boyan Bonev[2] and Bassem A. Hassan[1,6,*].

[1] Institut du Cerveau-Paris Brain Institute (ICM), Sorbonne Université, Inserm, CNRS, Hôpital Pitié-Salpêtrière, 75013 Paris, France.

[2] Institute of Epigenetics and Stem Cells, Helmholtz Center Munich, Munich, Germany.

[3] I3S Institute for Research and Innovation in Health, University of Porto, Rua Alfredo Allen 208, 4200-135 Porto, Portugal.

[4] Institut Curie, PSL University, CNRS UMR144, Paris, France.

[5] Sorbonne Université, INSERM, Centre de Recherche Saint-Antoine, CRSA, AP-HP, Paris, France.

[6] Lead contact

*Correspondence: Bassem.hassan@icm-institute.org

ABSTRACT

The striking expansion of the human neocortex is driven by the ability of radial glial cells (RGCs) to prolong neurogenesis without exhausting their progenitor pool, thereby producing more neurons, particularly in upper layers. The mechanisms underlying enhanced human neurogenic output remain unresolved. We hypothesized that post-translational modification of conserved proteins could provide a flexible mechanism supporting species-specific patterns of neurogenesis. Here, we identify phosphorylation of the proneural factor NEUROGENIN2 as a regulator of human RGC neuronal output. Inhibiting NEUROGENIN2 phosphorylation remodels RGC chromatin landscape and gene regulatory networks, leading to increased neuronal production biased towards upper-layer neurons. Our work proposes a conceptual framework for how subtle modifications of conserved processes encode tunable mechanisms for species-specific trajectories of brain development.

KEYWORDS

Human cortical neurogenesis, Cortical organoids, Cell fate, Deep-learning, Single-Cell Multiomics, Chromatin, AP-1, Post-translational modification, Proneural Proteins.

INTRODUCTION

The evolutionary expansion of the human neocortex underlies higher-order cognitive abilities such as reasoning, language, and abstract thought [1,2]. Architecturally defined by six layers of neurons formed in an inside-out sequence, with early generation of deep layer neurons, followed by later generation of upper layer neurons[3–5], the human neocortex contains more neurons than any other species, with approximately 16 billion in humans compared to 6 billion in chimpanzees, 1.7 billion in rhesus macaques, 31 million in rats, and 14 million in mice [6–8]. This expansion of the primate neocortex is most pronounced in the upper layers, which show increases in both neuron number and molecular subtype diversity [9–12]. Increased neurogenesis is facilitated by a prolonged neurogenic period[13–16] during which cortical progenitors called Radial Glial Cells (RGCs) engage fin a temporally controlled balance between proliferative divisions that expand their pool and neurogenic divisions that generate neurons either directly from RGCs (direct neurogenesis), or indirectly through the generation of intermediate progenitors (IPs; indirect neurogenesis). It remains unclear how human RGCs evolved their species-specific extension of the neurogenic period to increase neuronal output. Recent work has focused on human-specific evolutionary driving forces[17] such as gene duplications of ARHGAP11B [18–20], NOTCH2NL [21,22], CROCCP2 [23] , and SRGAP2C [24,25], and regulatory innovations with human-accelerated regions (HARs) like HAR5 [26,27]. However, these genomic level human-specific innovations do not generate new mechanisms. Instead, they converge on classic developmental pathways such as Notch, WNT, or PI3K-Akt [28], to modify the activity, timing, or expression of neurogenic programs that are otherwise shared across vertebrates. This raises a fundamental question: can human-specific features of cortical development arise from regulatory tuning of deeply conserved mechanisms, without changes to gene sequence? In particular, can we fine-tune the neuronal production of human cortical progenitors through post-translational modifications of conserved transcription factors that gate neurogenic programs?

In the developing mouse neocortex, Neurogenin2 (Neurog2), a highly conserved basic-helix-loop-helix (bHLH) transcription factor, is both necessary and sufficient for deep-layer cortical neurogenesis, but is dispensable for the generation of upper layer neurons [29–34] – the main target of human cortical expansion. By integrating upstream signals such as Notch, FGF, and WNT into transcriptional programs, Neurog2 acts as a hub to promotes neuronal differentiation and fate commitment during mouse cortical development [29,35–38]. Surprisingly, whether NEUROG2 is required for human cortical neurogenesis has not been investigated.

Proneural proteins, like NEUROG2, do not function as binary switches. Instead, their activity is finely tunable through post-translational modifications (PTMs) such as phosphorylation to modulate their stability, activity, and transcriptional output [37,39–41]. A conserved phosphorylation site in NEUROG2, threonine 149 (T149), located at the junction between the loop and the second helix of its bHLH domain,

directly modulates its neurogenic activity. Overexpression of a phospho-mutant form (T149A) of Neurog2 in developing mouse cortical RGCs enhances its neurogenic activity, promoting early neurogenesis and favoring deep-layer neuron production at the expense of upper-layer neurons [39]. Given the species-specific expansion of human RGCs and their increased neuronal output, particularly in upper layers, we asked whether there may be a species-specific alteration to NEUROG2 function such that phosphorylation enables context-specific functional plasticity to modulate the neurogenic capacities of human RGCs. This would provide a tunable cell biological mechanism to extend neurogenic trajectories throughout neocortical development without requiring human-specific genetic changes.

To test this hypothesis, we used CRISPR/Cas9 genome editing in two different human iPSC lines to generate two different mutants; NEUROG2 KO to ask whether NEUROG2 is required for human cortical neurogenesis, and NEUROG2 T149A to ask whether it's function can be tuned by PTM. We then combined live single progenitor lineage tracing of ~400 RGCs, and developed a machine learning based pipeline to quantify ~100 million cells across ~200 hCOs to characterize the division patterns and neurogenic output of control and NEUROG2 mutant human cortical RGCs. Next, we leveraged paired single-nucleus RNA-seq and ATAC-seq, Gene Regulatory Network (GRN) analysis and experimental validation to discover the molecular mechanisms underpinning NEUROG2 T149 modulation in human RGCs. At the cellular level, we find that, in contrast to mice, NEUROG2 expression is largely restricted to RGCs and is required for both deep and upper layer neurogenesis. The T149 phosphorylation site acts as a rheostat of RGC division patterns to fine tune the balance between proliferative and neurogenic divisions. At the molecular level, inhibiting T149 phosphorylation uncouples NEUROG22's chromatin modulating activity from its transcriptional activity, promoting the opening of JUN and FOS binding sites specifically in RGCs. This alters RGC gene regulatory networks and drives enhanced neurogenic commitment, leading to increased production of upper-layer neurons. Together, our results support the hypothesis that conserved and tunable cell biological mechanisms can underlie species-specific expansion in brain size.

RESULTS

Human NEUROG2 is largely restricted to RGCs

To investigate neurogenesis in hCOs, we used a 3D cortical differentiation protocol (Fig. S1a) adapted from Sloan et al.,[42] and we confirmed the presence of all key cell subtypes found in the developing human neocortex (Fig. S1c-g). At day 70, we found all progenitor subtypes, such as radial glia cells (RGCs) expressing SOX2 and PAX6 (Fig.S1c & S1d), and Intermediate progenitors (IPs), expressing TBR2 (Fig.S1c & S1d). At this stage of cortical development, these progenitors mostly produce TBR1+ deep-layer neurons and are just starting to produce SATB2+ upper-layer neurons (Fig.S1e). However, at later

97 stages of neocortical development, around day 140, these SOX2+ progenitors produce both deep and upper-
98 layer neurons (Fig.S1f & S1g). The high number and high density of cells in 3D hCOs poses a significant
99 challenge for quantification; thus, we developed a semi-automated image analysis pipeline integrating Deep
100 Learning (DL) and Machine Learning (ML) algorithms to measure cell proportions with precision (Fig.1a,
101 see Methods for details). We fine-tuned the StarDist segmentation model [43,44] to better detect individual
102 nuclei in very dense regions of the hCOs where we found that the F1 score improved from 0.64 to 0.83,
103 indicating a substantial increase in segmentation accuracy (Fig.S2b-e). Next, we used an ML classifier
104 based on random forests in the open-source software *ilastik*[45] to categorize each nucleus into distinct
105 subtypes defined by standard cell-type markers. This image analysis pipeline allowed us to quantify
106 thousands of sections containing hundreds of thousands of cells. In all the experiments described below, all
107 nuclei within each section of each organoid used in this study were quantified to obtain the data presented.

108 We next examined the expression pattern of NEUROG2 in the progenitor populations of hCOs at day 70
109 through immunostaining for NEUROG2, SOX2 (pan-progenitor marker, RGCs and IPs), and TBR2
110 (intermediate progenitor marker, IPs only) (Fig.1b). In two independent clones, we observed that 4-5% of
111 all progenitors expressed NEUROG2 (Fig.1c). To determine the identity of the cells expressing NEUROG2,
112 we quantified the proportion of RGCs (SOX2+/TBR2-) and IPs (TBR2+) within the NEUROG2+
113 population. In both sets of control organoids, the majority of NEUROG2+ cells were RGCs (86.5% and
114 82.2%), while a smaller fraction corresponded to IPs (13.5% and 17.8%) (Fig.1d). To assess the relevance
115 of these findings to human brain development, we next examined NEUROG2 expression in the Human
116 Fetal Neocortex (HFC). We found similar proportions in the developing HFC at gestational week 18,
117 developmentally comparable to D70-90 in hCOs, where 6% of all progenitors expressed NEUROG2
118 (Fig.1c). Among these NEUROG2+ cells, 76.3% were RGCs and 23.7% were IPs (Fig.1d).

119 Together, these results indicate that in the human developing neocortex, NEUROG2 is predominantly
120 expressed in RGCs, a notable divergence from the murine developing neocortex, where NEUROG2
121 expression predominates in IPs at mid-stage of neocortical development[34,46].

122 ***NEUROG2 is required for efficient and timely human cortical neurogenesis***

123 Given this species-specific difference, we sought to examine the role of ENUROG2 in human cortical
124 neurogenesis. In the developing mouse neocortex, NEUROG2 is considered the main driver of deep-layer
125 cortical neurogenesis, but it is dispensable for the generation of upper-layer neurons [34,46]. To this end, we
126 used CRISPR/Cas9 genome editing to generate from our main iPSC line, WTSli008-A, two independent
127 NEUROG2 knockout iPSC clones referred to as KO1 and KO2, each carrying early stop codons (Fig.S3a).
128 A third clone (Ctrl), exposed to the editing protocol but unmodified, served as the isogenic control. We
129 validated the NEUROG2 KO in a different genetic background with the GM25256*E iPSCs line (Fig.S3d).

130 We first validated in the main iPSC line (WTSli008-A) the loss of NEUROG2 protein expression and
131 assessed early progenitor dynamics at D70 of differentiation by immunostaining COs for SOX2 (pan-
132 progenitor marker), TBR2 (IP marker), and NEUROG2 (Fig.2a). In control organoids, 4.9% ± 0.5% of
133 progenitors expressed NEUROG2, whereas NEUROG2 expression was reduced to background levels in
134 KO1 (0.61% ± 0.5%) and KO2 (0.26% ± 0.5%) organoids (Fig.S3b & S3c), confirming the loss of
135 NEUROG2 in these clones.

136 To assess the consequences of NEUROG2 depletion on neurogenesis, we quantified the proportions of
137 progenitors and neurons in hCOs at mid (D70) and late (D140) stages using immunostaining for SOX2
138 (progenitors) and NEUROD2 (excitatory neurons) (Fig.2a & 2b). At mid stages, in control organoids we
139 found that 75.8% ± 9.2% of the cells were progenitors and 24.2% ± 9.2% were neurons (Fig.2c, d). In
140 contrast, KO clones showed a significant increase in progenitors (KO1 90.4% ± 9.2%, KO2 85.4% ± 9.2%)
141 and an approximately 2-fold reduction in neurons (KO1 9.6% ± 9.2%, KO2 14.6% ± 9.2%) (Fig.2c, d).
142 These results were also validated in our second iPSC cell line (Fig.S3e-g). At later stages, control organoids
143 contained 46.6% ± 7.2% progenitors and 53.4% ± 7.2% neurons, whereas KO organoids displayed a
144 persistent neurogenic deficit, with significantly elevated proportions of progenitors (KO1: 65.0% ± 7.2%;
145 KO2: 62.1% ± 7.2%) and a ~1.5 fold reduction in neurons (KO1 35.0% ± 7.2%, KO2 37.9% ± 7.2%)
146 (Fig.2e, f). These findings indicate a sustained impairment in neurogenic output and prolonged maintenance
147 of progenitor identity upon NEUROG2 loss in hCOs.

148 To examine whether the loss of NEUROG2 specifically affects the generation of deep-layer neurons, as
149 previously found in mice[32], we quantified neuronal subtype identity among all NEUROD2+ cells using
150 TBR1 (deep layer) and SATB2 (upper layer) at D140 (Fig.2g). In control organoids, 60.3% ± 10.6%
151 neurons were $SATB2^+$ upper-layer neurons, while the remaining 38.8% ± 9.8% were $TBR1^+$ deep-layer
152 neurons (Fig.2h). In contrast, KO1 and KO2 organoids exhibited a shift toward deep-layer identity, with an
153 increased proportion of $TBR1^+$ neurons (KO1: 55.7% ± 9.8%; KO2: 54.2% ± 9.8%) and a corresponding
154 reduction in upper-layer neurons (KO1: 43.0% ± 10.5%; KO2: 44.7% ± 10.5%) (Fig.2h).

155 Thus, in contrast to the mouse, human NEUROG2 is required for the generation upper-layer neurons.

156

157 ***A conserved phosphorylation site regulates NEUROG2 proneural potency***

158 Given the critical role of NEUROG2 in human cortical neurogenesis, we investigated mechanisms
159 controlling its activity. Previous work in mice showed that phosphorylation of a conserved serine/threonine
160 residue within the bHLH domain, the threonine 149 (T149) in NEUROG2, acts as a regulatory switch for
161 the transcriptional and proneural potency of proneural proteins [39]. This residue is conserved from insects to

162 primates, including humans (Fig.3a), suggesting a strong evolutionary pressure to maintain its function.
163 The gnomAD database comprising approximately 1.6 million sequenced alleles, identifying that
164 NEUROGENIN2 T149 is rarely mutated. Only two heterozygous cases of a T149A substitution were
165 detected where NEUROG2 cannot be phosphorylated anymore (Fig.3b). Moreover, pathogenicity
166 predictions using AlphaMissense indicate that substitutions at this position are highly deleterious (score =
167 0.99, Fig.3b), among the most highly pathogenic in the bHLH domain highlighting its potential importance
168 in human brain development.

169 To test its importance, we used CRISPR/Cas9 genome editing to prevent T149 phosphorylation by
170 replacing the T149 with an Alanine on both alleles (T149A, henceforth TA/TA) of our two iPSC lines, the
171 WTSli008-A (Fig.S4a) and GM25256*E (Fig.S5a). For the main WTSli008-A iPSCs line, we confirmed
172 with optical sequencing the absence of major DNA remodification post editing (Fig.S4b & S4c). In mid
173 stages hCOs, we found that the TA/TA mutation did not alter the proportion of NEUROG2+ cell (4.15%
174 for the WT/WT and 4.45 % for the TA/TA) (Fig.3c & 3d), nor their distribution: with the majority of
175 NEUROG2+ cells being RGCs in both WT/WT (83.6% ± 4.8%) and TA/TA (83.6% ± 4.8%) and the rest
176 being IPs (WT/WT 16.4% ± 4.8%) and (TA/TA 16.4% ± 4.8%) (Fig.3e).

177 We next examined early (D70, Fig.3f) and late (D140, Fig.3g) neurogenesis. In D70 COs, we observed a
178 significant 10% reduction in progenitors (SOX2+) with a corresponding increase in neurons (NEUROD2+)
179 (Fig.3h & 3i), which was confirmed with our second iPSC line GM25256*E (Fig.S5b-d). This significant
180 change was persistent but milder at D140, with a 5% decrease in progenitors and a corresponding increase
181 in neurons (Fig.3j & 3k). Finally, we assessed cortical laminar identities at D140 by staining for deep-layer
182 ($TBR1^+$) and upper-layer ($SATB2^+$) markers (Fig.3l). Quantification revealed a significant shift in neuronal
183 subtype specification: with WT/WT organoids exhibiting approximately equal proportions of deep-layer
184 (49.8%) and upper-layer (42.4%) neurons, whereas TA/TA organoids showed a significant decrease in
185 deep-layer neurons (44.8%) and a corresponding increase in upper-layer neurons (49.8%) (Fig.3m).

186 In summary, NEUROG2 T149A shows the reciprocal phenotypes to the NEUROG2 KO, increasing
187 neuronal production throughout cortical neurogenesis, notably of upper-layer neurons.

188

189 ***Inhibiting NEUROG2 Phosphorylation Drives RGCs into a Neurogenic State***

190 So far, we have seen that NEUROG2 is mostly expressed by RGCs and that when the T149 is not
191 phosphorylated, it drives increased neurogenesis in both mid and late stages of human cortical development.
192 However, it remains unclear whether the increased neuronal output observed in the phospho-mutant arises
193 from RGCs with direct neurogenesis or IPs with indirect neurogenesis. To address this, we performed

194 single-cell lineage tracing using live imaging combined with correlative immunostaining [47]. We infected
195 D50–70 WT/WT and TA/TA cortical organoids with a GFP-expressing retrovirus and tracked individual
196 progenitors (n = 386) over 48–96 hours, allowing us to capture 1-2 divisions per progenitor (Fig.4a). We
197 then used immunostaining for cell fate markers (SOX2 for RGCs, TBR2 for IPs, and NEUROD2 for
198 neurons) to determine the fate of GFP+ progeny, enabling us to classify each division as either proliferative
199 or neurogenic (Fig.4a). RGCs were classified as undergoing proliferative divisions if both daughter cells
200 remained RGCs or neurogenic divisions if at least one daughter became an IP or neuron (Fig.4b). Similarly,
201 IP divisions were classified as proliferative (producing two IPs) or neurogenic (producing one or two
202 neurons) (Fig.4c).

203 In WT/WT RGCs, the vast majority of divisions (90.7%) were proliferative, with only 9.3% producing IPs
204 and no direct neuronal production. In contrast, TA/TA phospho-mutant RGCs exhibited a marked decrease
205 in proliferative divisions (75.1%) and a significant increase in neurogenic divisions, generating more IPs
206 (21.3%) and neurons directly (3.5%) (Fig.4d), indicating enhanced direct and indirect neurogenesis. Tracing
207 IP lineages revealed a similar trend. While most WT/WT IPs underwent proliferative divisions 93.9%, only
208 6.1% produced neurons. In TA/TA organoids, however, only 56.4% of IP divisions were proliferative, with
209 43.6% yielding neurons (Fig.4e). These results suggest that phospho-mutant RGCs not only produce more
210 IPs but also bias these IPs toward neurogenic divisions, thereby amplifying neuronal output across
211 successive generations.

212 This shift in division dynamics led to substantial changes in progenitor pools: live imaging revealed a
213 reduction in RGCs from 88% to 60.5%, alongside increased proportions of IPs, from 11.6% to 24.1%, and
214 neurons, from 0.4% to 15.4%, among TA/TA progenitor progeny over 2–4 days (Fig.4f). To validate these
215 findings at a larger scale, we analyzed fixed organoid slices—including those not subjected to live
216 imaging—by quantifying GFP+ cells across random fields of view. This analysis confirmed the same trend,
217 with TA/TA organoids exhibiting decreased RGCs (from 81.2% to 64.5%) and increased IPs (from 9.7%
218 to 14.6%) and neurons (from 9.1% to 21%) (Fig.4g). Together, these results demonstrate that blocking
219 NEUROG2 phosphorylation shifts progenitors toward neurogenic trajectories, enhancing both direct and
220 indirect neurogenesis.

221

222 ***T149 attenuates NEUROG2's proneural function***

223 This raised the question of how the T149 residue regulates NEUROG2 activity, specifically, whether it
224 alters NEUROG2's intrinsic transcriptional capacity? To validate this, we first probed NEUROG2's direct
225 reprogramming potential with the rapid neuronal reprogramming protocol[48] by using an all-in-one, rtTA-
226 inducible lentiviral construct encoding the human NEUROG2 fused to eGFP that will reprogram iPSCs

227 into neurons in less than a week. To prevent any positive feedback loop of NEUROG2, we used one of our
228 NEUROG2 KO iPSCs clones and transduced it with two different lentiviruses, one expressing wild-type
229 NEUROG2 and the other NEUROG2 T149A. We then selected the iPSCs that had been infected with
230 lentiviruses through puromycin treatment to obtain stable clones with inducible NEUROG2 WT or
231 NEUROG2 T149A. After NEUROG2 induction through doxycycline we quantified the fraction of $MAP2^+$
232 neurons among GFP^+ cells after six days in culture (Fig.5a). While WT NEUROG2 leads to a high
233 conversion efficiency of 99.6%, the NEUROG2 T149A produced only 62.3% $MAP2^+$ neurons (Fig.5b),
234 suggesting that, surprisingly, the T149A mutation decreases NEUROG2's transactivation capacity. To test
235 this directly, we compared the transactivation capacity of NEUROG2 WT and T149A in P19 cells, which
236 lack endogenous NEUROG2 expression. We used luciferase reporters driven by either the Neurod1
237 promoter (long 2.2 kb variants; Fig.5c) or a multimerized E-box element from the Dll1 promoter $(E2Z)_6$-
238 luc (Fig.5d) as targets [49,50]. Although NEUROG2 T149A retained the ability to activate both Neurod1 and
239 Dll1 reporters, its activity was substantially reduced by ~2-fold relative to wild-type NEUROG2,
240 confirming the direct reprogramming results (Fig.5c-d & Fig.S6a-f). Thus, paradoxically, preventing the
241 phosphorylation of NEUROG2 T149 enhances its proneural function in cortical RGCs but it also
242 substantially diminishes its transactivation potency.

243 This paradox suggests that the increased neurogenesis observed in phosphomutant hCOs must result from
244 another type of activity. More and more evidences suggest that NEUROG2 can induce a reshaping of the
245 chromatin landscape by promoting DNA demethylation, increasing chromatin looping, and gene
246 accessibility [51–55]. Therefore, we examined gene accessibility and gene expression with paired single-
247 nucleus multiome profiling (RNA-seq and ATAC-seq on the same nuclei, Fig.5e and Fig.S7a-d) on eight
248 control organoids and eight phospho-mutant organoids collected from two independent iPSCs 3D
249 differentiation batches at day 70. We next integrated single-nucleus ATAC-seq and RNA-seq data into a
250 unified UMAP projection, which resolved seven major cell populations (Fig.5f). Cortical clusters (FOXG1)
251 contained the expected cell types, namely RGCs, mitotic RGCs (RGC_M: CDK1, TPX2, TOP2a), IPs, and
252 two clusters of excitatory neurons at different stages of maturation (N1: SSTR2, NEUROD2, NEUROD6,
253 and N2: MYT1L, NELL2, RUNX1T1, BCL11B) (Fig.5g).

254 To determine the effects of NEUROG2 T149A on the molecular program of cortical differentiation, we
255 used Monocle3[56] to reconstruct a pseudotemporal trajectory from the transcriptomes of all nuclei by
256 designating RGCs as the origin of all the other cell subtypes (Fig.5h). The inferred path recapitulated the
257 key stages of human cortical development starting with RGC, then IP, N1, and N2. Interestingly, TA/TA
258 RGCs and IPs occupy significantly more advanced pseudotime positions than their WT/WT counterparts
259 (Fig.5i, upper panel), indicating premature progression through progenitor states. Although N2 neurons

exhibit higher pseudotime values than N1, consistent with their greater maturity, neither neuronal population showed a significant genotype difference (Fig.5i, lower panel).

These results suggest a priming of phospho-mutant RGCs towards more neurogenic states at the transcriptional level, but are they also biased towards faster maturation? To address this question, we applied FateID's[57] random-forest classifier to compute, for each RGC and IP, the probability of adopting an N1 versus an N2 identity. We observed that TA/TA RGCs and IPs are significantly more biased towards generating more mature neurons (N2) than the WT/WT cells (Fig.5j), confirming the more advanced neurogenic potential observed in cellular analyses.

Phospho-mutant NEUROG2 remodels the RGC chromatin landscape

Taking advantage of the combination of single-nucleus RNAseq and ATAC-seq, we linked distal ATAC peaks, which are candidate cis-regulatory regions (cCREs), to their putative target genes using co-accessibility and correlated expression profiles (Fig.6a). Plotting the normalized accessibility of cCREs alongside the normalized expression of their linked genes for each cell subtype (RGC, IP, N1, N2) revealed both genes and cCREs with variable expression and accessibility patterns in TA/TA compared to WT/WT hCOs. To understand which transcription factors (TFs) might be underlying this variability, we performed TF-motif enrichment on linked cCREs (Fig.6b). We found an enrichment of motifs bound by different members of the JUN/FOS superfamily of transcription factors, collectively known as the AP-1 complex, in TA/TA RGCs compared to WT/WT RGCs (Fig.6b). Interestingly, the same AP-1 motifs were less enriched in TA/TA IP than in WT/WT IP cCRE, highlighting a transient and cell-specific chromatin remodeling. Consistent with this, JUN shows higher expression in RGCs than IPs (Fig.S6e), and enhancers containing JUN motifs are associated with higher gene expression in RGCs while decreasing as RGCs enter neuronal differentiation (Fig.S6f). Finally, regression-based enrichment analysis showed that while NEUROG2 is the most enriched motif in WT/WT RGCs, JUN, ATF4, and CTCF are among the most enriched TF motifs in TA/TA RGCs. This suggests a potential shift in the transcriptional program of RGCs as WT/WT RGCs rely more on NEUROG2 while TA/TA RGCs increasingly engage in AP-1-centered programs (Fig.6c & Fig.S6g).

Next, we assessed whether JUN family proteins were differentially bound to their motifs in TA/TA versus WT/WT progenitors using footprinting analysis, a method to detect physical binding of TF to their motifs. We found that the JUN footprint is stronger in TA/TA RGCs compared to WT/WT RGCs (Fig.6d), and while this footprint remains stronger in TA/TA IPs, it is reduced overall, and the JUN motif no longer appears among the enriched motifs in IP peaks (Fig.6e, Fig.S6h). Finally, we quantified the effect of TF motif presence in cCREs on the expression of their linked target genes. For that purpose, RGC-specific

genes were split into two sets. The first set comprised genes linked to distal enhancers having the AP-1 TFBS or one of its variants present in their sequence and the second set comprised genes linked to enhancers without those motifs. Genes linked to RGC-specific peaks with FOS or JUN family motifs show significantly higher expression in TA/TA compared to those without these motifs (Fig.6f). This effect is specific to RGCs and is not observed in IPs or neurons (Fig.S6f). Furthermore, this effect was observed only in RGCs but not in IPs or neurons (Fig.S6f).

NEUROG2 modulates the RGC gene regulatory network via JUN and FOS

Collectively, these observations establish that the phospho-mutant NEUROG2 leads to a remodeling of the chromatin landscape in RGCs, increasing accessibility and functional engagement of AP-1–associated regulatory elements. The enriched AP-1 motifs, JUN expression in RGCs, a stronger JUN footprint, and increased expression of AP-1 linked target genes suggest a potential rewiring of the gene regulatory network (GRN) of phosphomutant RGCs.

To explore the architecture of the GRN in sequenced RGCs, we used CellOracle[58]. The overall network topology remained similar between TA/TA and WT/WT RGCs, with a number of nodes, representing either a transcription factor or their target genes of 2025 nodes in TA/TA RGCs compared to 2005 nodes in WT/WT RGCs (Fig.7a, b). We also observed a slight increase in the number of edges, or regulatory connections, in the TA/TA RGCs with 9499 edges compared to 8993 edges in WT/WT RGCs (Fig.7a, c), suggesting a mild GRN expansion. To assess the influence of JUN and FOS on these networks, we computed three metrics: the number of direct connections a TF has (degree centrality), how often a TF acts as a regulatory bottleneck (betweenness centrality), and the number of a TF's connections and the importance of its neighboring nodes (eigenvalue centrality). We found that both JUN and FOS showed increased centrality in TA/TA RGCs across the three parameters, indicating that these AP-1 factors regulate more target genes and are more central to key regulatory modules in the phospho-mutant context.

However, it remains unclear whether AP1-centric transcriptional control drives the increased neurogenesis in phosphomutant RGCs in hCOs. To address this, we performed in silico knockdowns (KDs) of JUN and FOS to measure their potential impact on neuronal trajectories. We isolated cortical cells from our dataset and visualized their organization in UMAP. We found, that N1 and N2 neuronal populations form distinct clusters (Fig.7f) and WT/WT and TA/TA cells occupy different spaces (Fig.7g). To perform the perturbations, we built the developmental landscape using a grid point 2D vector field approach based on earlier pseudotime analysis (Fig.7h). Next, we performed in silico KDs of JUN and FOS (Fig.7i & 7k). By computing the inner product of the perturbation vectors with the developmental vectors of TA/TA cells, we found that JUN and FOS KDs reduce the differentiation potential of TA/TA RGCs and IPs (indicated by

326 the purple color and back-flowing arrows) and revert them toward a WT/WT-like transcriptional state
327 (indicated by the green color and upward flowing arrows) (Fig.7j & 7l).

328 To validate these analyses experimentally, we used 2D cultures of human RGCs kept under proliferative
329 conditions. TA/TA RGCs displayed a ~4-fold increase in the expression of the immature neuron marker
330 DOUBLECORTIN (DCX) compared to WT/WT RGCs (Fig.7m & n), validating the increased
331 neurogenesis of phosphomutant RGCs in 2D cultures. Inhibiting AP1 activity with the well-established
332 molecule SR11302 [59,60] for 48 hours, corresponding to the length of one cell division in human RGCs,
333 reduced the number of immature neurons (DCX+) by 1.8-fold (Fig.7n). Therefore, we demonstrated that
334 NEUROG2 can promote neurogenic cell state transitions in human cortical cells through the AP-1 complex.

335

336 **DISCUSSION**

337 One of the central questions in neurobiology is how species-specific features of the brain arise during
338 development. While genetic innovations such as duplications of ARHGAP11B[18] and NOTCH2NL[21,22], or
339 human-accelerated regions[27], have been proposed to drive cortical expansion, our findings suggest that post-
340 translational modifications of conserved proteins like NEUROG2 may offer a complementary, reversible,
341 mechanism for fine-tuning neurogenesis and thus potentially contribute to species-specific cortical
342 evolution.

343 ***A robust and efficient Deep-Learning based image analysis workflow***

344 To detect quantitative alterations with high confidence and at single cell resolution, a central technical
345 advance of our study was the development of a scalable image analysis pipeline, powered by deep learning,
346 to quantify progenitor and neuronal dynamics across entire cortical organoid datasets. By combining
347 automated nuclear segmentation with supervised machine learning–based classification, we analyzed over
348 100 million cells across thousands of sections. This unprecedented scale allowed us to detect subtle but
349 robust shifts in progenitor fate and lineage progression that would be difficult to capture with traditional
350 manual or low-throughput approaches. Beyond enabling our discoveries, this pipeline offers a powerful,
351 adaptable resource for the general scientific community, allowing researchers to uncover hidden phenotypes
352 in any model subjected to genetic perturbations, environmental challenges, or drug treatments.

353 ***Species-Specific Roles and Post-Translational Modulation of NEUROG2 in Human Cortical***
354 ***Neurogenesis***

355 At the cellular level, our findings reveal that NEUROG2 plays a broader role in human cortical neurogenesis
356 than previously appreciated. In contrast to mice, where NEUROG2 is largely restricted to intermediate
357 progenitors (IPs) and primarily promotes deep-layer neuron production, we show that in human models,

358 NEUROG2 protein is predominantly expressed in radial glial cells (RGCs) and is required for the
359 generation of both deep and upper-layer neurons. Importantly, this cannot be solely explained by a
360 difference in the expression pattern, since electroporation of mouse Neurog2 into developing mouse RGCs
361 accentuates deep layer neurogenesis at the expense of upper layer neurogenesis. This establishes
362 NEUROG2 as a central regulator of human cortical neurogenesis and suggests a species divergence in its
363 function. In addition, our findings support the hypothesis that decoupling protein activity from its
364 expression could enable evolutionary changes in developmental timing, progenitor cell state transitions and
365 neuronal output. NEUROG2 phospho-mutant RGCs shift toward neurogenic divisions, both directly
366 producing neurons and generating more IPs, which themselves are biased toward neurogenesis during mid-
367 stages of cortical development. This effect persists into later developmental stages, specifically enhancing
368 production of late born upper layer neurons, a defining hallmark of human neocortical expansion.
369 Importantly, this occurs without depleting the progenitor pool, suggesting that NEUROG2 phosphorylation
370 status contributes to balancing self-renewal and differentiation in a way that maximizes neuronal output
371 over extended periods.

372 Together, these findings support a general principle: cellular behaviors need not arise solely from new
373 genes, but can also reflect context-dependent modulation of conserved regulators via PTMs.

374 ***T149 phosphorylation disentangles transcriptional and pioneer-like activities of NEUROG2***

375 At the molecular level, NEUROG2 has been known to function both as a transcriptional activator and as a
376 pioneer-like factor capable of inducing reshaping of the chromatin landscape by promoting DNA
377 demethylation, increasing chromatin looping, and accessibility [51–55]. However, whether and how these dual
378 activities, transcriptional activation and chromatin remodeling, are mechanistically separated and
379 differentially regulated was unknown. We discovered that phosphor-regulation of T149 uncouples
380 NEUROG2's transcriptional activity from its pioneer-like function to remodel the chromatin. Surprisingly,
381 phospho-mutant NEUROG2 exhibited reduced transactivation capacity in direct reprogramming and
382 reporter assays, despite promoting increased neuronal differentiation in organoids. This apparent paradox
383 is resolved by chromatin accessibility analyses, which revealed that inhibiting T149 phosphorylation leads
384 to widespread premature opening of chromatin regions in RGCs, effectively priming progenitors for
385 differentiation before they normally would. Importantly, this shift does not reflect accelerated maturation
386 of neurons. Rather, it represents an earlier transition of progenitors into more differentiated trajectories, as
387 demonstrated by pseudotime analysis. Furthermore, these chromatin changes show remarkable specificity
388 toward particular transcription factor networks, notably the JUN and FOS family. This was accompanied
389 by increased accessibility and elevated expression of AP-1 target genes, as well as enhanced JUN
390 transcription factor footprints in NEUROG2 phospho-mutant RGCs. Interestingly, these AP-1–associated

391 chromatin changes were largely restricted to RGCs and faded in downstream IPs and neurons, suggesting
392 that phospho-mutant NEUROG2 primes RGCs to enter neurogenic trajectories via selective enhancer
393 activation. Pharmacological inhibition of AP-1 activity partially rescued the premature neurogenic
394 phenotype after one cell division, confirming its direct functional role in driving the observed differentiation
395 bias. Thus, it appears that NEUROG2 can drive neurogenesis through two different mechanisms. The first
396 through its classical role by inducing the expression of genes like NEUROD1 and a previously unkniwn
397 role of increasing RGCs responsiveness to the AP-1 complex.

398 ***The AP-1 complex: a hub for human cortical evolution?***

399 This mechanistic link position the AP-1 transcription factor networks as a central mediator of NEUROG2
400 phosphorylation effects. This aligns with accumulating evidence implicating AP-1 in species-specific
401 aspects of cortical development. For instance, JUNB and FOSL2 have been shown to alter progenitor
402 dynamics and extend neurogenesis in human models, while ectopic JUNB expression in mice confers
403 human-like progenitor properties [61]. Similarly, FOSL2 has been identified as a driver of human-specific
404 gene regulatory networks in cortical progenitors [62]. Thus, our results suggest NEUROG2 phosphorylation
405 as an upstream modulator of a potential evolutionary hub: the AP-1 transcription factor network. By
406 demonstrating that NEUROG2's phosphorylation state selectively tunes AP-1–driven chromatin
407 remodeling in human RGCs, our study suggests that NEUROG2 itself has been evolutionarily co-opted to
408 engage AP-1 pathways in a human-specific context. Moreover, the evolutionary significance of AP-1 motifs
409 extends beyond progenitors. Recent work has shown that species-specific changes in AP-1 binding sites
410 underlie activity-dependent gene regulation in neurons [63], suggesting that AP-1 represents a generalizable
411 axis of evolutionary plasticity throughout cortical development, from progenitor fate decisions to synaptic
412 maturation. In this light, NEUROG2 phosphorylation emerges not merely as a developmental modulator
413 but as a potential entry point into species-specific regulatory programs centered around the AP-1 complex.

414 ***Cellular context regulation of NEUROG2 activity?***

415 Our study shows that phosphorylation of a single conserved residue, T149, can uncouple distinct activities
416 of NEUROG2, thereby rewiring progenitor behavior and neuronal output in humans. This suggests that the
417 interplay between kinases and phosphatases creates context-dependent windows of NEUROG2 activity that
418 modulate neuronal production. Because phosphorylation depends on cellular energy and signaling states,
419 this mechanism provides a direct link between metabolism, developmental cues, and progenitor fate. In this
420 way, neurogenesis can be regulated at multiple levels by diverse key players, offering a tunable layer of
421 control that may differ not only across progenitor subtypes but also between species. This raises a broader
422 question: might evolution have harnessed reversible post-translational modifications such as T149
423 phosphorylation to diversify cortical trajectories, increasing or limiting neuronal output in a species-specific

manner? If so, the evolutionary expansion of the human cortex may not solely reflect the appearance of new genes, but also the flexible re-use of ancient proteins under novel regulatory dynamics—an elegant strategy for generating complexity without rewriting the genome. This conceptual framework offers a new lens on how evolutionary innovations in brain development can arise: through the convergence of genetic novelty with regulatory plasticity of conserved proteins.

LEAD CONTACT

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bassem Hassan (bassem.hassan@icm-institute.org).

ACKNOWLEDGMENTS

We thank Dr. Nicolas Renier for comments and discussions. We thank the Paris Brain Institute ICV-iPSCs core facility for CRISPR/Cas 9 genome editing, the ICM-Quant facility for help with microscopy and especially David Akbar for his precious recommendations, and the Data Analysis Core for help with statistical analysis. This work was supported by Paris Brain Institute core funding, ANR grant NEURONPP (ANR-21-CE13-0041), the Roger De Spoelberch Prize (to BAH), the New Frontiers in Research Fund – Transformation, supported by the Canadian Tri-agencies - CIHR, NSERC, and SSHRC (to DSC), the Bettencourt Schueller Foundation Impulscience grant (to ADB), ERA-NET Neuron (MOSAIC), and European Research Council Consolidator Grant (EpiCortex, 101044469) (to BB). C.B.A. was funded by a Fondation pour la Recherche Médicale (FRM) post-doctoral fellowship.

AUTHOR CONTRIBUTIONS

Conceptualization, J.P. and B.A.H.; methodology, J.P., B.A.H., ; Investigation, J.P.; T.D.; M.A.M.; L.R; M.V.S; N.D.; C.M.; C.B.A; H.M.; writing—original draft, J.P., T.D, B.B, B.A.H.; writing—review & editing, J.P. and D.C.; C.B.A; A.D.B; C.P.; B.B.; B.A.H.; funding acquisition, L.E.K; B.B; B.A.H.; resources, L.E.K; B.B. and B.A.H.; supervision, B.B. and B.A.H.

DECLARATION OF INTERESTS

The authors declare no conflicting interests.

METHOD DETAILS

Experimental models

Two human induced pluripotent stem cell (iPSC) lines were used in this study: WTSIi008-A, obtained from the European Bank for Induced Pluripotent Stem Cells (EBISC), and GM25256*E, obtained from the Coriell Institute. Both lines underwent genomic characterization using Bionano optical mapping and were used for CRISPR/Cas9-mediated genome editing. WTSIi008-A served as the primary line for generating edited clones used in the main figures. Bionano analysis confirmed that no additional structural variants or recombination events occurred after CRISPR/Cas9 genome editing (Fig.S3). The GM25256*E served as the secondary line to validate the results obtained from the WTSIi008-A in another genetic background.

Human fetal neocortical tissue at gestational week 18 (GW18) was obtained with prior informed consent and in full compliance with institutional and legal ethical guidelines. The protocol was approved by the

461 French Biomedical Agency (Agence de la Biomédecine, approval number: PFS17-003). Fresh cortical
462 tissue was collected from spontaneous miscarriages or medically indicated pregnancy terminations during
463 autopsies performed at Robert Debré Hospital and Necker-Enfants Malades Hospital (Paris, France). A
464 fragment of prefrontal cortex was dissected from one hemisphere and transported on ice to the laboratory.
465 Samples were fixed in 4% paraformaldehyde (PFA) for 2 hours prior to further processing.
466

467 ***iPSCs culturing and maintenance***

468 iPSCs were cultured on Geltrex LDEV-Free hESC-qualified Reduced Growth Factor Basement Membrane
469 Matrix (1%, Thermo Fisher Scientific, A1413302) coated dishes (B6 or B10 dishes) in mTeSR™ Plus
470 (STEMCELL Technologies, #100-0276) supplemented with Antibiotic-Antimycotic (0,1%, Thermo Fisher
471 Scientific, 15240062). The medium was changed every other day, and the iPSCs were passaged when iPSCs
472 reached 80% confluency with 300µL of cGMP ReLeSR™ (STEMCELL Technologies, #100-0483) at
473 various dilutions depending on the needs for the different experiments.

474 ***CRISPR/Cas9 genome editing in human iPSCs***

475 Guide RNAs (gRNAs) were designed using Benchling[64]. A total of 1×10^6 human iPSCs were nucleofected
476 with a ribonucleoprotein (RNP) complex composed of 225 pmol each of crRNA and tracrRNA-ATTO+,
477 and 120 pmol of Cas9 protein. Twenty-four hours after nucleofection, ATTO+ iPSCs were isolated by
478 fluorescence-activated cell sorting (FACS) and seeded at low density (10–20 cells/cm²) on Synthemax II-
479 coated surfaces in medium supplemented with CloneR2 to support clonal expansion. After one week, 192
480 individual iPSC clones were manually picked under a stereomicroscope and transferred to Geltrex-coated
481 96-well plates for expansion. Once confluent, clones were split for parallel cryopreservation and genomic
482 DNA extraction. Genotyping was performed by PCR amplification followed by Sanger sequencing. List of
483 primers and gRNA sequences used for the generation of the NEUROG2 KO clones of both iPSCs lines:

484 crRNA = GTCTGGTACACGATTGCAAA CGG

485 primer FWD and sequencing = CTGACCTTGGTTAGCACTGCC

486 primer REV = GGAATTGGAGGACACGGAGG

487 List of primers and gRNA sequences used for the generation of the NEUROG2 T149A clones of both iPSCs
488 lines:

489 crRNA = AGCTCACCAAGATCGAGACC

490 primer FWD and sequencing = CTGACCTTGGTTAGCACTGCC

491 primer REV = CAGAAAGGCTACACCTGCCC

492 ***Optical genome mapping***

493 We used optical genome mapping (OGM) to evaluate genome stability in iPSCs following CRISPR-Cas9–
494 mediated genome editing. This technique enables the detection of a wide range of chromosomal aberrations,
495 both numerical and structural, with a resolution over 10,000 times greater than traditional karyotyping[65].
496 OGM has been shown to be a powerful tool for iPSC quality control[66].

497 Briefly, ultra-high molecular weight DNA was extracted from frozen cell pellets using the Bionano Prep
498 Blood & Cell Culture DNA Isolation Kit (Bionano Genomics, San Diego, CA, USA). DNA was then
499 labeled with the Direct Label and Stain (DLS) enzyme using the Bionano Prep DNA Labeling Kit,
500 following the manufacturer's instructions. The labeled DNA molecules were loaded onto a Saphyr
501 nanofluidic chip (Bionano Genomics) for linearization and imaged using the Saphyr instrument. The
502 resulting images were converted into digital data and analyzed using Bionano Solve software.

503 *De novo* assembly was performed, followed by chromosomal analysis using two pipelines: a copy number
504 variation (CNV) pipeline for detecting large, unbalanced events based on molecule coverage, and a
505 structural variation (SV) pipeline for identifying both balanced and unbalanced rearrangements by
506 comparing labeling patterns to a reference genome. Results were analyzed and visualized using Bionano
507 Access software.

508 Variants detected in the edited lines (heterozygous and homozygous) but absent in the parental line were
509 manually reviewed for validation.

510 ***Generation of human cortical organoids from iPSCs***

511 Cortical organoids were generated from human iPSCs using a previously reported protocol[42] with
512 modifications. hiPS cells were washed with PBS and incubated with 1mL of StemPro™ Accutase™ Cell
513 Dissociation Reagent (Thermo Fisher Scientific, A1110501) for 5 min at 37 °C and dissociated into single
514 cells. To obtain uniformly sized spheroids, approximately 3×10^6 single cells were added per well in the
515 AggreWell 800 plate (STEMCELL Technologies, 34815) with mTeSR™ Plus medium supplemented with
516 Stemgent hES Cell Cloning & Recovery Supplement (1X, Ozyme, STE01-0014-500) and incubated at 37°C
517 with 5% CO_2. After 24 hours, spheroids from each microwell were collected by pipetting medium in the
518 well up and down and transferred into Corning® non-treated culture dishes (Merck, CLS430591-500EA)
519 in TeSR™-E6 (StemCell Technologies, #05946) supplemented with two inhibitors of the SMAD signaling
520 pathway, dorsomorphin (5 μM, STEMCELL Technologies, #72102) and SB-431542 (10 μM, STEMCELL
521 Technologies, #72234). From day 2 to day 5, TeSR™-E6 supplemented with dorsomorphin and SB-431542
522 was changed daily.

523 On day 6, the medium was replaced by Neurobasal™-A (Thermo Fisher Scientific, 10888022), B-27™
524 Supplement minus vitamin A (50X) (1X, Thermo Fisher Scientific 12587010), supplemented with
525 GlutaMAX™ (1%, Thermo Fisher Scientific, 35050038), 2-mercaptoethanol (0.1mM, Thermo Fisher
526 Scientific, 31350010) and Antibiotic-Antimycotic (0.1%, Thermo Fisher Scientific, 15240062).

527 This medium was supplemented with 20 ng/mL Human Recombinant EGF, ACF (STEMCELL
528 Technologies, #78136) and 20 ng/mL Human Recombinant bFGF, ACF (STEMCELL Technologies,
529 #78134.1) and changed daily until day 12. From day 12 to day 24, the medium was changed every other
530 day.

531 On day 25, the medium was replaced by Neurobasal™ Plus (Thermo Fisher Scientific, A3582901),
532 supplemented with B-27™ Plus Supplement (50X) (1X, Thermo Fisher Scientific, A3582801),
533 GlutaMAX™ (1%, Thermo Fisher Scientific, 35050038), 2-mercaptoethanol (0,1mM, Thermo Fisher
534 Scientific, 31350010), L-Ascorbic acid (200µM, Sigma-Aldrich, A4403), and Antibiotic-Antimycotic
535 (0,1%, Thermo Fisher Scientific, 15240062). From day 25 to day 43, this medium was supplemented with
536 Human Recombinant BDNF, ACF (STEMCELL Technologies, #78133), and Human/Mouse Recombinant
537 NT-3 (STEMCELL Technologies, # 78074) and changed twice a week.

538 From day 43 until the last time points of interest, the same medium was changed twice a week without
539 BDNF nor NT3.

540 ***Cryopreservation of organoids***

541 Organoids were collected in 2mL Eppendorf tubes, washed with PBS, and fixed in 4% PFA (Electron
542 Microscopy Sciences, 15714) at RT for 6 hours in the dark. Organoids were washed several times in PBS
543 before the incubation in sucrose 30% for at least 24h. When organoids are at the bottom of the tubes, they
544 were collected and placed in the cryomolds for cryopreservation with Epredia™ Neg-50™ Frozen Section
545 Medium (Fisher Scientific, 12688086) and frozen on dry ice. Organoids were then transferred to -80°C
546 upon slicing.

547 ***Slicing of organoids***

548 For immunostaining, COs were sectioned at a thickness of 14 µm using a Leica cryostat (Leica CM3050S)
549 and collected on Superfrost Plus glass slides 25x75 (Fischer scientific, 11950657). The slides were stored
550 at -80°C upon immunostaining.

551 **Stainings, image acquisition and analysis**

To reduce the potential bias included by the experimenter and external conditions that vary throughout the years, all the cortical organoids of the same differentiation batch were stained, imaged and analyzed together. Here is the general pipeline from staining to image acquisition and analysis.

Immunostainings

Slides were incubated at room temperature (RT) for 2 hours until completely dry. Slides were then washed with phosphate-buffered saline (PBS) to remove all cryoprotectant and blocked for 1 hour at RT with a solution containing 5% heat-inactivated horse serum (Thermo Fisher Scientific, 26050070), 3% bovine serum albumin (BSA) (Sigma-Aldrich, A9647-100G), and 0.3% Triton X-100 (Sigma Aldrich, X100-500mL). The sections were then incubated overnight at 4°C with primary antibodies diluted in the same blocking solution. After three washes with PBS for 5 minutes each, the sections were incubated with secondary antibodies and 4',6-diamidino-2-phenylindole (DAPI) (1:5000, Sigma-Aldrich, D9564) diluted in the same blocking solution for 1 hour at RT. Following three additional washes with PBS, slides were mounted with VECTASHIELD Vibrance Antifade Mounting Medium (Eurobio Scientific, H170010) and stored at 4°C until imaging. Nail polish was added after polymerization of the mounting medium to ensure airtight sealing for long-term storage.

Imaging of whole sliced organoids

Slides were positioned in the 4-slide plate (inverted-face down) and left at RT to warm before imaging with an inverted Nikon confocal AXR microscope. An initial low-resolution overview of the entire plate was acquired using a 4X objective (Resonant mode, 512x512 pixels, pinhole: 6 AU, high gain) in the 405 nm channel to capture DAPI signal. This step allows spatial registration of individual organoid slices and alignment with the identifiers assigned by the NIS software. High-resolution imaging was performed using a 20X air objective with Perfect Focus System enabled to maintain consistent z-plane positioning throughout acquisition. Each organoid slice was imaged as a tiled mosaic with a 15% overlap for better stitching using Galvano scanning (dwell time: 0.8 µs, pinhole: 1 AU, resolution: 2048x2048 pixels). Channels were acquired sequentially from longest to shortest wavelength (647, 555, 488, 405 nm) to prevent fluorophore cross-excitation. The mosaic dimensions were standardized across all slices based on the largest sample to ensure full coverage. Then image acquisition was launched after setting up all the parameters.

Renaming of the images

After imaging, all slices were renamed following the same nomenclature throughout the entire project to store metadata in the title of the files. Here is the nomenclature:

Clone_Diff_Day_Staining_Objective_Parameters_OGnumber_SliceNumber

584 For example, a title would be: H9_dG_J70_SNd2_20Xa_g2dt08_OG7_s4 which corresponds to clone H9,
585 a NEUROG2 T149A phospho-mutant from differentiation batch G fixed and stained at day 70 for the
586 immunostaining of SOX2-NEUROD2 imaged with the 20X air objective using the mode galvano with an
587 averaging of 2 and a dwell time of 0.8. This image corresponds to the slice 4 (s4) of organoid number 7
588 (OG7. The number of the organoid was set once for each batch and each time point and the number of slices
589 corresponds to the number of slices obtained at the cryostat (usually 4 to 6 slices).

590 ***Image processing***

591 First, images were all opened manually in ImageJ for visual inspection. Any image with either an
592 immunostaining too faint or with organoid slices not imaged entirely or the focus lost during acquisition
593 were discarded from further analysis. Then the remaining images were manually cropped to save storage
594 space and saved as tiff images for segmentation and object classification.

595 **Fine-tuning of StarDist, a deep learning algorithm for nuclear segmentation**

596 ***Training***

597 StarDist is a state-of-the-art machine learning-based method for cell segmentation that employs a star-
598 convex polygon shape model to accurately segment nuclei[43,44]. While the pre-trained version of StarDist
599 for fluorescent images yields very good results on images with sparse nuclei, it performs poorly on images
600 with dense nuclei, as shown in Fig.S2a. Thus, 134 images were taken using the Nikon AXR confocal
601 microscope at different resolutions (Resonnant and Glavano mode) with either 1024x1024 pixels resolution
602 or 2048x2048 pixels with the 20X air or 40X oil objectives to ensure the model to segment nuclei at
603 different scale so it can be used on images taken from drosophila, mice and human samples. As StarDist
604 internally uses convolutional neural network (CNN), the images don't need to have the same x and y axis.
605 Thus, images were cropped randomly to contain at least 100 to 700 cells each. All these images used for
606 the training were manually annotated using Labkit, an ImageJ plugin. Furthermore, the original images and
607 their corresponding masks used for training the model, which the Stardist authors published, were also
608 downloaded. This resulted in a total of 447 images being downloaded from Stardist's GitHub,
609 corresponding to the stage1_train images from the Kaggle Data Science Bowl, which are available in full
610 from the Broad Bioimage Benchmark Collection. The folder with the 447 images can be downloaded in a
611 link provided in additional material. To enhance the training further, all the images were augmented before
612 the training with techniques such as flipping the axis followed by rotations of 90°, 180° and 270° to multiply
613 the number of images by 16. Then, in the training notebook, further augmentation was done by applying
614 random intensity changes and Gaussian blur.

615

616 ***Testing the model***

617 In parallel, 9 images with various nuclei densities were taken as described above and manually labeled to
618 evaluate the Stardist haug2 model fine-tuned. These images were not used for the training or the validation
619 of our StarDist haug2 model. The model performances were assessed and are shown in Fig.S2b. In short,
620 the F1 score, which measures the model's ability to avoid false alarms, i.e., predicting positive when the
621 actual label is negative. The pretrained version of StarDist Versatile (fluorescent nuclei) scored an F1 score
622 of 0.64, while the fine-tuned StarDist: stardit_haug2 has an F1 score of 0.83 (Fig.S2c-e) on the test images.
623 The model was then exported for use in Python but also in ImageJ or FIJI with version of TensorFlow 1.12
624 and 1.15, and in QuPath. The fine-tuned model, StarDist haug2 can be downloaded from the GitHub page
625 of the lab and will be published on Bioimage.io for downloads. When used in Python, the Probability/Score
626 Threshold and the overlap threshold are set automatically. In FIJI and QuPath, they have to be set manually
627 to 0.51 and 0.3 respectively.

628 ***Image segmentation***

629 In Python, all images were then segmented using the fine-tuned model, StarDist haug2 on the DAPI channel
630 (corresponding to channel 0 in the metadata of our images) with a tiling of 4x4 or 6x6 for faster processing
631 of large images. The corresponding masks were saved either as uint16 when the image contains less than
632 65,535 nuclei or as uint32 bits when it contains more than 65,535 nuclei to optimize storage.

633 ***Image selection for training the object classifier and assessing its performance***

634 For better classification, the mean intensity of the nuclear markers of the different channels per image was
635 calculated to plot the distribution of image intensities for a channel of interest across all the images. In
636 general, 3 images were selected per genotypes for the training of the object classifier, with overall high,
637 mean, and low intensities for the markers of interest.

638 ***Object classification***

639 We used a free and user-friendly software, ilastik, to train object classifiers [45]. Training images and their
640 corresponding masks were loaded in a new project for object classification. For training the random forest,
641 we selected specific features among all the ones proposed by ilastik such as the mean intensity, the total
642 intensity, the skewness of intensities, and the object's area. Categories were created based on what type of
643 cell population were stained. The training was done by assigning nuclei, and thus segmented objects, to
644 each subcategory created in all the training images. The classifier measures all the previously selected
645 features to establish a general profile of the cell. We then developed a simple python code to aggregate all
646 the single files generated by ilastik for each image into a single file that corresponds to the RAW_DATA
647 of the quantification done for the specific staining of the batch analyzed. With this file, we can then count

648 the number of cells of each subcategory based on ilastik predictions per organoid slices and then the sum
649 of all counted slice per organoid to get all the cells quantified per organoid. Organoids with less that 4 slices
650 were discarded as the quantification might not reflect the full diversity of cell population across the entire
651 organoid.

652 **Correlative immunostaining**

653 The following technique is based on Coquand et al [47].

654 ***Viral infection with retroviruses***

655 Cortical organoids at both week 7 and week 9, corresponding to 55 and 63 days in culture, respectively,
656 were collected and embedded in UltraPure low-gelling agarose (INVITROGEN, 16520-050) for subsequent
657 slicing. A Leica VT1200S vibratome was used to slice the organoids into 250 µm – 300 µm thick sections
658 at a 200µm/sec speed, while maintaining the samples in cold DMEM/F-12 (Thermo Fisher Scientific,
659 10565018). Slices were transferred in a 24 wells plate for retroviral transduction (MSCV-IRES-GFP,
660 Addgene plasmid #20672) diluted at 1/30 in DMEM/F-12. After 2 hours of incubation, slices were washed
661 using PBS and incubated in DMEM/F-12 supplemented with B-27 (Thermo Fisher Scientific, 10565018),
662 N2 (1X, Thermo Fisher Scientific, 17502048), 10ng/mL FGF2 (STEMCELL Technologies, #78133.1),
663 10ng/mL EGF (STEMCELL Technologies, #78136), 5% fetal bovine serum (1X, Thermo Fisher Scientific,
664 A5670801) and 5% horse serum (Thermo Fisher Scientific, 26050070) for 2 days before live imaging.

665 ***Live Imaging***

666 In order to follow RGCs and IP divisions over several days, slices were transferred onto a permeable
667 membrane (Millicell Cell Culture Insert, 30 mm, hydrophilic PTFE, 0.4 µm, PICM0RG50, Merck) to create
668 an air-liquid interface. The membrane was theninserted onto a 35mm glass bottom dishes (Fluorodish WPI,
669 FD35-100) with 1mL of culture medium supplemented with EGF and FGF at a final concentration of 10
670 ng/mL under the membrane. Based on the GFP signal, positions were chosen and imaged every 15 minutes
671 in 100µm stacks to capture migrating progenitors. However, the z step for each stack depends on the number
672 of positions you have so that all the imaging fits within a 15-minute time window, where you have a good
673 temporal resolution of cell divisions and migration to correlate the types of divisions with the daughter cell
674 fate. At the end of the live imaging, low magnification images were taken with a 4X objective for alignment
675 of the movies with the fixed, stained and mounted slices of hCOs.

676 ***Immunostaining and movie processing***

677 Subsequently, the slices were fixed using 4% PFA, with 1 mL applied under the membrane and 1 mL on
678 top, to ensure maximal penetration of PFA. The slices were then washed with PBS for three 10-minute

679 intervals and distributed into the wells of a 24-well plate. They were incubated for 1 hour in a blocking
680 buffer consisting of PBS supplemented with 0.3% Triton-X and 2% donkey serum. Primary antibodies were
681 incubated in the same solution overnight at 4°C on a shaking plate. Following this, the slices were washed
682 using PBS + 0.05% Tween-20 before the incubation with secondary antibodies (diluted 1:1000 from our
683 stock) in the same blocking buffer. The slices were then washed and mounted on Superfrost slides using
684 Aquapolymount (18606-20, PolySciences). The slides were left in the dark at room temperature for
685 polymerization and subsequently stored at 4°C until use.

686 ***iPSCs infection with the NEUROG2 lentiviruses***

687 The NEUROG2 KO clone 1 from the iPSCs line WTS8, lacking NEUROGENIN2 was infected with the
688 two different lentiviruses containing either the WT version of the human Neurog2 or the NEUROG2 T149A
689 phospho-mutant. On day 0, 60,000 iPSCs were incubated in a coated plate with 12 wells with the classical
690 conditions as described earlier. On day 1, the cells were transduced with the lentiviruses at a MOI of 4 in
691 mTeSR™ Plus (STEMCELL Technologies, #100-0276) supplemented with Stemgent hES Cell Cloning &
692 Recovery Supplement (1X, Ozyme, STE01-0014-500) and 8µg/mL polybrene (Hexadimethrine bromide,
693 Sigma-Aldrich, H9268-5G) for 3 days. After, we started the selection phase by changing the medium every
694 day with increasing doses of puromycin (Sigma-Aldrich, P9620-10ML). For two days at 2µg/mL followed
695 by three days at 1µg/mL. Then the selected iPSCs were then amplified and frozen in CryoStor CS10
696 (STEMCELL Technologies, 07930) until their use.

697 ***iPSCs conversion into neurons***

698 On day -1, 60,000 iPSCs were incubated in a well of a 12-well plate coated with Geltrex LDEV-Free hESC-
699 qualified Reduced Growth Factor Basement Membrane Matrix (1%, Thermo Fisher Scientific, A1413302)
700 in mTeSR™ Plus supplemented with hES. On day 0, the induction of NEUROG2 and NEUROG2 phospho-
701 mutant expression was done by changing the medium with mTeSR™ Plus supplemented with 0.4µg/mL
702 doxycycline (Sigma-Aldrich, D9891-5G). On day 1, the medium was replaced by Neurobasal Plus
703 supplemented with BDNF and NT3 at a final concentration of 20 ng/mL with 0.4 µg/mL doxycycline. On
704 day 3, the medium was replaced with the same medium as day 1, supplemented with cytosine β-D-
705 arabinofuranoside (Sigma-Aldrich, C6645), to remove any potential dividing iPSCs that had escaped the
706 selection process. After two days, on day 5, the medium was replaced with the exact same medium prepared
707 on day 1. On day 6, neurons were fixed using PFA 4% for 10 minutes and washed three times very gently
708 to prevent neuron detachment from the coverslips. Immunostaining using antibodies anti-GFP and anti-
709 MAP2 was done to assess the conversion efficiency of each *NEUROG2* construct.

710

711 ***Luciferase assay***

712 P19 embryonic carcinoma cells were maintained in Dulbecco´s modified Eagle's high glucose medium
713 (Sigma, D0422) supplemented with 10% Fetal Bovine Serum (Biowest, S008Y30304), 100 U/ml
714 Penicillin-Streptomycin (Gibco, 15140122) and 2 mM L-Glutamine (Gibco, 25030149). Transfections
715 were performed in quadruplicate in 48-well plates by using Lipofectamine 2000 (Thermo Fisher, 11668019)
716 according to manufacturer's protocol. Each well was seeded 1 day earlier with 6x104 cells and transfected
717 with 200ng appropriate expression plasmid, 100ng luciferase reporter plasmid, and 200ng CMV-b-gal as
718 internal control. Cells were lysed 24h after transfection with Glo Lysis Buffer (Promega, E2661), and
719 extracts were assayed for luciferase and b-gal activities. Data are represented as means of quadruplicates.

720 **Quantification and statistical analysis**

721 All statistical analyses of the data generated by our image analysis pipeline were quantified using RStudio
722 with a linear mixed-effects model that accounts for the biological variability across independent iPSCs 3D
723 differentiation batches. For each staining and timepoint, the percentage of neuronal or progenitor cells per
724 organoid section was modeled using the lmer() function from the lme4 R package, specifying Group (e.g.,
725 genotype) as a fixed effect and Batch as a random intercept to correct for inter-experimental variation.
726 Group-level comparisons were performed on the fitted models using the emmeans package, which
727 computes estimated marginal means and pairwise contrasts via Students t-tests with Satterthwaite's
728 approximation for degrees of freedom. Model assumptions were assessed by visual inspection of residual
729 plots. Outliers were defined as observations with standardized residuals exceeding ±2 standard deviations
730 and were excluded before re-fitting the models.

731 All other statistical analyses were done with GraphPad Prism9, and normality of samples was verified by
732 the Shapiro-Wilk test, followed by Student t tests when comparing only two conditions. Otherwise, we used
733 a one-way ANOVA followed by a Tukey post hoc analysis for multiple comparisons.

734 For multiomic analyses, all statistical details, test and p-values, are given in the results sections and the
735 figure legends.

736 **Multiomics analyses**

737 ***QC and cell filtering***

738 First, the raw sequencing data were demultiplexed with Cellranger (cellranger-arc-2.0.2) and mapped to the
739 human hg38 reference genome (refdata-cellranger-arc-GRCh38-2020-A-2.0.0). On the feature-barcode
740 matrix, QC metrics were computed in R using Seurat and Signac. Cells with minimally 2,000 and
741 maximally 18,000 transcriptomic reads, minimally 2,000 and maximally 37,000 fragment counts, a minimal

742 nucleosome signal of 0.2 and a maximum of 2, minimal TSS enrichment of 1 and maximal of 20, and
743 minimally 1000 genes detected were retained (Fig S6 A-D). After QC filtering, there were 6,773 cells
744 remaining. After peak calling with macs2[67], non-standard chromosomes and blacklisted regions were
745 removed.

746 ***Dimensionality reduction and joint embedding***

747 Next, the gene expression data were analyzed with Seurat. First, the transcriptomic counts were log-
748 normalized, and the most variable features were selected. On the scaled and centered data, PCA was applied.
749 The first 15 dimensions were used as input for a UMAP. Analogously, the chromatin accessibility data
750 were analyzed with Signac and ArchR[68]. Normalization was done via term frequency-inverse document
751 frequency (TF-IDF) normalization. The most variable peaks were identified, including features with >10
752 total counts. Those top peaks were scaled as well. Dimensionality reduction was done via singular value
753 decomposition (SVD).

754 The first 15 and second to 15th dimensions of the RNA and ATAC modality respectively were used to
755 construct a weighted nearest neighbor (WNN) graph which was visualized via a UMAP representation.
756 Doublet scores were computed with scDblFinder[69]. Harmony[70] was used to correct for batch leading to an
757 updated WNN and UMAP. Cells were clustered using Louvain. One of the clusters had a median doublet
758 score of 0.99 and was therefore removed as putative doublet cluster. The final clustering comprises of 7
759 clusters, which were annotated using marker gene expression per cluster.

760 ***Cell fate bias***

761 The gene expression data of RGCs and IPs was tested per condition for a cell fate bias towards neurons (N1
762 and N2) with FateID[57] using 5 cells per target cluster for classification per iteration and 10 cells per training
763 set for classification

764 ***Monocle 3 pseudotime***

765 To obtain pseudotime values per cell type and condition, the monocle3[56] algorithm was applied to the
766 transcriptomic data using all cells in the RGC cluster as starting point.

767 ***ATAC-seq analysis***

768 Using ArchR, a reproducible peak set was established, and Transcription factors (TF) were annotated using
769 JASPAR2020. Accessible peaks were linked to putative target genes (P2G) via Pearson correlation of peak
770 accessibility and gene expression in a 500kb window. Only peaks that had a significant correlation (FDR <
771 0.1) and a minimal distance of 5kb to the TSS were retained as putative candidates of distal enhancers.
772 Correlations were filtered for a value above 0.35 or below -0.35.

773 To decipher which TFBS in the peaks could be driving the gene expression in the target genes, a binned TF
774 motif enrichment was done for the peaks linked to genes aggregated per cell type and condition with
775 MonaLisa[71].

776 ***TF motif enrichment in RGC marker peaks***

777 Marker peaks per cell type were identified using a Wilcoxon test using the TA condition as group and the
778 WT condition as background while correcting for TSS enrichment and number of unique fragments per cell
779 via ArchR. To identify TF binding sites, a TF motif enrichment in the RGC-specific marker peaks was done
780 with MonaLisa using 5 equal sized bins based on the log2 fold change (LFC) of the marker peaks in the
781 TA condition of RGC. Human PWMs were extracted from JASPAR2020 and filtered for expressed TFs.
782 Additionally, a regression-based enrichment using randomized lasso stability selection was done for both
783 RGC and IP, predicting the LFC of the Wilcox test of the marker peak selection with the TF motif hits in
784 the peak sequences using a selection cutoff of 0.6. Motif footprints were determined in the marker peaks
785 for JUN and FOS motifs using a modified version of the ArchR implementation. The footprints were
786 normalized by subtracting Tn5 sequence bias.

787 ***Effect of TF motif presence on gene expression***

788 RGC specific genes and their linked peaks were extracted from the P2G matrix. TF motifs of the expressed
789 TFs were added with Signac. Then, the peak set was split into a positive peak set containing at least one or
790 more motifs of the TF of interest and a negative peak set containing none of these motifs. The RGC-specific
791 genes were split into a positive and negative set depending on whether their linked peak contained these
792 motif hits or not. These gene sets were compared across conditions in their average LFC across cells and
793 in their expression per cell over the cell's pseudotime values. To obtain LFC values, RGC TA vs WT was
794 tested for differentially expressed genes in pseudobulk gene expression per cell type and condition with
795 Seurat FindMarkers using TA as group and WT as background using DESeq2.

796 ***Gene regulatory network construction***

797 Gene regulatory networks (GRNs) were inferred using CellOracle. A base GRN was constructed from
798 ATAC-seq fragments by performing motif scanning (JASPAR 2020) and associating transcription factors
799 (TFs) with putative target genes based on promoter and enhancer accessibility. The multiomic dataset was
800 first filtered to retain only cortical cells. Each cortical cluster identified earlier during the multiome analysis,
801 was then subdivided by condition (WT or TA). GRNs were independently computed for each cell subtype
802 and condition using CellOracle's default thresholding parameters. To enrich the regulatory model, we
803 manually extended the selected gene set to include the genes of the AP-1 complex and their putative target

genes, resulting in a total of approximately 4500 genes used for GRN inference. No other specific filtering was applied.

In silico gene perturbation

In silico perturbation analyses were performed using the GRNs derived from TA cells and plotted on the general UMAP computed after filtering for cortical cells. We simulated knockdowns of **JUN** and **FOS** by reducing their expression to 30% of baseline levels (0.3). CellOracle was used to compute the resulting vector fields representing predicted changes in gene expression and cell state across the existing low-dimensional space. These simulations were projected onto the precomputed pseudotime embedding to visualize how perturbation of AP-1 components would influence cell fate trajectories in the TA condition.

ADDITIONAL RESOURCES

Downloading the images used for the pretrained versions of StarDist: https://github.com/stardist/stardist/releases/download/0.1.0/dsb2018.zip

Multiomic data can be found with GEO accession number: GSE305338

All the code for the image analysis pipeline can be found on GitHub at: https://github.com/Julien-Pgn/nuclei_quant

831

832 **REFERENCES**

833 1. Molnár, Z. *et al.* New insights into the development of the human cerebral cortex. *Journal of Anatomy*
834 **235**, 432–451 (2019).
835 2. Kaas, J. H. Chapter 3 - The origin and evolution of neocortex: From early mammals to modern humans.
836 in *Progress in Brain Research* (ed. Hofman, M. A.) vol. 250 61–81 (Elsevier, 2019).
837 3. Angevine, J. B. & Sidman, R. L. Autoradiographic Study of Cell Migration during Histogenesis of
838 Cerebral Cortex in the Mouse. *Nature* **192**, 766–768 (1961).
839 4. Caviness, V. S. Neocortical histogenesis in normal and reeler mice: a developmental study based upon
840 [3H]thymidine autoradiography. *Brain Res* **256**, 293–302 (1982).
841 5. Rakic, P. & Lombroso, P. J. Development of the Cerebral Cortex: I. Forming the Cortical Structure.
842 *Journal of the American Academy of Child & Adolescent Psychiatry* **37**, 116–117 (1998).
843 6. Herculano-Houzel, S. The human brain in numbers: a linearly scaled-up primate brain. *Front. Hum.*
844 *Neurosci.* **3**, (2009).
845 7. Herculano-Houzel, S. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain
846 and its associated cost. *Proc Natl Acad Sci U S A* **109 Suppl 1**, 10661–10668 (2012).
847 8. Herculano-Houzel, S., Catania, K., Manger, P. R. & Kaas, J. H. Mammalian Brains Are Made of These:
848 A Dataset of the Numbers and Densities of Neuronal and Nonneuronal Cells in the Brain of Glires,
849 Primates, Scandentia, Eulipotyphlans, Afrotherians and Artiodactyls, and Their Relationship with
850 Body Mass. *Brain Behav Evol* **86**, 145–163 (2015).
851 9. Bakken, T. E. *et al.* Comparative cellular analysis of motor cortex in human, marmoset and mouse.
852 *Nature* **598**, 111–119 (2021).
853 10. Yu, Y. *et al.* Interneuron origin and molecular diversity in the human fetal brain. *Nat Neurosci* **24**,
854 1745–1756 (2021).
855 11. Vanderhaeghen, P. & Polleux, F. Developmental mechanisms underlying the evolution of human
856 cortical circuits. *Nat Rev Neurosci* **24**, 213–232 (2023).

857 12. Stepien, B. K., Vaid, S. & Huttner, W. B. Length of the Neurogenic Period—A Key Determinant for
858 the Generation of Upper-Layer Neurons During Neocortex Development and Evolution. *Front Cell*
859 *Dev Biol* **9**, 676911 (2021).
860 13. Geschwind, D. H. & Rakic, P. Cortical Evolution: Judge the Brain by Its Cover. *Neuron* **80**, 633–647
861 (2013).
862 14. Rakic, P. Evolution of the neocortex: Perspective from developmental biology. *Nat Rev Neurosci* **10**,
863 724–735 (2009).
864 15. Libé-Philippot, B. & Vanderhaeghen, P. Cellular and Molecular Mechanisms Linking Human Cortical
865 Development and Evolution. *Annu Rev Genet* **55**, 555–581 (2021).
866 16. Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M. & Sestan, N. The Cellular and Molecular Landscapes
867 of the Developing Human Central Nervous System. *Neuron* **89**, 248–268 (2016).
868 17. Ohno, S. Gene duplication and the uniqueness of vertebrate genomes circa 1970-1999. *Semin Cell Dev*
869 *Biol* **10**, 517–522 (1999).
870 18. Florio, M. *et al.* Human-specific gene ARHGAP11B promotes basal progenitor amplification and
871 neocortex expansion. *Science* **347**, 1465–1470 (2015).
872 19. Namba, T., Nardelli, J., Gressens, P. & Huttner, W. B. Metabolic Regulation of Neocortical Expansion
873 in Development and Evolution. *Neuron* **109**, 408–419 (2021).
874 20. Namba, T. *et al.* Human-Specific ARHGAP11B Acts in Mitochondria to Expand Neocortical
875 Progenitors by Glutaminolysis. *Neuron* **105**, 867-881.e9 (2020).
876 21. Fiddes, I. T. *et al.* Human-Specific NOTCH2NL Genes Affect Notch Signaling and Cortical
877 Neurogenesis. *Cell* **173**, 1356-1369.e22 (2018).
878 22. Suzuki, I. K. *et al.* Human-Specific NOTCH2NL Genes Expand Cortical Neurogenesis through
879 Delta/Notch Regulation. *Cell* **173**, 1370-1384.e16 (2018).
880 23. Van Heurck, R. *et al.* CROCCP2 acts as a human-specific modifier of cilia dynamics and mTOR
881 signaling to promote expansion of cortical progenitors. *Neuron* **111**, 65-80.e6 (2023).

882 24. Fossati, M. *et al.* SRGAP2 and Its Human-Specific Paralog Co-Regulate the Development of
883 Excitatory and Inhibitory Synapses. *Neuron* **91**, 356–369 (2016).
884 25. Schmidt, E. R. E. *et al.* A human-specific modifier of cortical connectivity and circuit function. *Nature*
885 **599**, 640–644 (2021).
886 26. Boyd, J. L. *et al.* Human-Chimpanzee Differences in a FZD8 Enhancer Alter Cell-Cycle Dynamics in
887 the Developing Neocortex. *Current Biology* **25**, 772–779 (2015).
888 27. Liu, J. *et al.* A human-specific enhancer fine-tunes radial glia potency and corticogenesis. *Nature*
889 (2025) doi:10.1038/s41586-025-09002-1.
890 28. Pollen, A. A. *et al.* Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution.
891 *Cell* **176**, 743-756.e17 (2019).
892 29. Bertrand, N., Castro, D. S. & Guillemot, F. Proneural genes and the specification of neural cell types.
893 *Nature Reviews Neuroscience* **3**, 517–530 (2002).
894 30. Schuurmans, C. *et al.* Sequential phases of cortical specification involve Neurogenin-dependent and -
895 independent pathways. *EMBO J* **23**, 2892–2902 (2004).
896 31. Fode, C. *et al.* A role for neural determination genes in specifying the dorsoventral identity of
897 telencephalic neurons. *Genes Dev* **14**, 67–80 (2000).
898 32. Mattar, P. *et al.* Basic helix-loop-helix transcription factors cooperate to specify a cortical projection
899 neuron identity. *Mol Cell Biol* **28**, 1456–1469 (2008).
900 33. Oproescu, A.-M., Han, S. & Schuurmans, C. New Insights Into the Intricacies of Proneural Gene
901 Regulation in the Embryonic and Adult Cerebral Cortex. *Front Mol Neurosci* **14**, 642016 (2021).
902 34. Kovach, C. *et al.* Neurog2 Simultaneously Activates and Represses Alternative Gene Expression
903 Programs in the Developing Neocortex. *Cerebral Cortex* **23**, 1884–1900 (2013).
904 35. Guillemot, F. Cell fate specification in the mammalian telencephalon. *Progress in Neurobiology* **83**,
905 37–52 (2007).
906 36. Ota, M. & Ito, K. BMP and FGF-2 regulate neurogenin-2 expression and the differentiation of sensory
907 neurons and glia. *Dev Dyn* **235**, 646–655 (2006).

908 37. Hindley, C. *et al.* Post-translational modification of Ngn2 differentially affects transcription of distinct
909 targets to regulate the balance between progenitor maintenance and differentiation. *Development* **139**,
910 1718–1723 (2012).
911 38. Li, S. *et al.* GSK3 temporally regulates neurogenin 2 proneural activity in the neocortex. *J Neurosci*
912 **32**, 7791–7805 (2012).
913 39. Quan, X.-J. *et al.* Post-translational Control of the Temporal Dynamics of Transcription Factor Activity
914 Regulates Neurogenesis. *Cell* **164**, 460–475 (2016).
915 40. Hand, R. *et al.* Phosphorylation of Neurogenin2 specifies the migration properties and the dendritic
916 morphology of pyramidal neurons in the neocortex. *Neuron* **48**, 45–62 (2005).
917 41. Ali, F. R. *et al.* The phosphorylation status of Ascl1 is a key determinant of neuronal differentiation
918 and maturation in vivo and in vitro. *Development* **141**, 2216–2224 (2014).
919 42. Sloan, S. A., Andersen, J., Pașca, A. M., Birey, F. & Pașca, S. P. Generation and assembly of human
920 brain region–specific three-dimensional cultures. *Nat Protoc* **13**, 2062–2085 (2018).
921 43. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. Cell Detection with Star-Convex Polygons. in
922 *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (eds. Frangi, A. F.,
923 Schnabel, J. A., Davatzikos, C., Alberola-López, C. & Fichtinger, G.) 265–273 (Springer International
924 Publishing, Cham, 2018). doi:10.1007/978-3-030-00934-2_30.
925 44. Weigert, M., Schmidt, U., Haase, R., Sugawara, K. & Myers, G. Star-convex Polyhedra for 3D Object
926 Detection and Segmentation in Microscopy. in *2020 IEEE Winter Conference on Applications of*
927 *Computer Vision (WACV)* 3655–3662 (2020). doi:10.1109/WACV45572.2020.9093435.
928 45. Berg, S. *et al.* ilastik: interactive machine learning for (bio)image analysis. *Nat Methods* **16**, 1226–
929 1232 (2019).
930 46. Britz, O. *et al.* A role for proneural genes in the maturation of cortical progenitor cells. *Cereb Cortex*
931 **16 Suppl 1**, i138-151 (2006).

932 47. Coquand, L. *et al.* A cell fate decision map reveals abundant direct neurogenesis bypassing intermediate
933 progenitors in the human developing neocortex. *Nat Cell Biol* (2024) doi:10.1038/s41556-024-01393-
934 z.
935 48. Zhang, Y. *et al.* Rapid Single-Step Induction of Functional Neurons from Human Pluripotent Stem
936 Cells. *Neuron* **78**, 785–798 (2013).
937 49. Huang, H. P. *et al.* Regulation of the pancreatic islet-specific gene BETA2 (neuroD) by neurogenin 3.
938 *Mol Cell Biol* **20**, 3292–3307 (2000).
939 50. Castro, D. S. *et al.* Proneural bHLH and Brn Proteins Coregulate a Neurogenic Program through
940 Cooperative Binding to a Conserved DNA Motif. *Developmental Cell* **11**, 831–844 (2006).
941 51. Noack, F. *et al.* Multimodal profiling of the transcriptional regulatory landscape of the developing
942 mouse cortex identifies Neurog2 as a key epigenome remodeler. *Nat Neurosci* **25**, 154–167 (2022).
943 52. Pereira, A. *et al.* Direct neuronal reprogramming of mouse astrocytes is associated with multiscale
944 epigenome remodeling and requires Yy1. *Nat Neurosci* **27**, 1260–1273 (2024).
945 53. Zhu, Y. *et al.* Spatiotemporal transcriptomic divergence across human and macaque brain development.
946 *Science* **362**, eaat8077 (2018).
947 54. Aydin, B. *et al.* Proneural factors Ascl1 and Neurog2 contribute to neuronal subtype identities by
948 establishing distinct chromatin landscapes. *Nat Neurosci* **22**, 897–908 (2019).
949 55. Smith, D. K., Yang, J., Liu, M.-L. & Zhang, C.-L. Small Molecules Modulate Chromatin Accessibility
950 to Promote NEUROG2-Mediated Fibroblast-to-Neuron Reprogramming. *Stem Cell Reports* **7**, 955–
951 969 (2016).
952 56. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–
953 502 (2019).
954 57. Herman, J. S., Sagar, null & Grün, D. FateID infers cell fate bias in multipotent progenitors from
955 single-cell RNA-seq data. *Nat Methods* **15**, 379–386 (2018).
956 58. Kamimoto, K. *et al.* Dissecting cell identity via network inference and in silico gene perturbation.
957 *Nature* **614**, 742–751 (2023).

958 59. Huang, C. *et al.* Blocking activator protein-1 activity, but not activating retinoic acid response element,
959 is required for the antitumor promotion effect of retinoic acid. *Proc Natl Acad Sci U S A* **94**, 5826–
960 5830 (1997).
961 60. Mishra, D. K. & Kim, M. P. SR 11302, an AP-1 Inhibitor, Reduces Metastatic Lesion Formation in Ex
962 Vivo 4D Lung Cancer Model. *Cancer Microenviron* **10**, 95–103 (2017).
963 61. Javed, A. *et al.* Developmental gene expression patterns driving species-specific cortical features.
964 2025.02.18.638637 Preprint at https://doi.org/10.1101/2025.02.18.638637 (2025).
965 62. Song, J. H. T. *et al.* Human-chimpanzee tetraploid system defines mechanisms of species-specific
966 neural gene regulation. 2025.03.31.646367 Preprint at https://doi.org/10.1101/2025.03.31.646367
967 (2025).
968 63. Carter, A. C. *et al.* FOS binding sites are a hub for the evolution of activity-dependent gene regulatory
969 programs in human neurons. 2025.03.31.646366 Preprint at
970 https://doi.org/10.1101/2025.03.31.646366 (2025).
971 64. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the
972 guide RNA selection tool CRISPOR. *Genome Biology* **17**, 148 (2016).
973 65. Mantere, T. *et al.* Optical genome mapping enables constitutional chromosomal aberration detection.
974 *Am J Hum Genet* **108**, 1409–1422 (2021).
975 66. Kitano, Y. *et al.* Generation of hypoimmunogenic induced pluripotent stem cells by CRISPR-Cas9
976 system and detailed evaluation for clinical application. *Mol Ther Methods Clin Dev* **26**, 15–25 (2022).
977 67. Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat*
978 *Protoc* **7**, 1728–1740 (2012).
979 68. Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin
980 accessibility analysis. *Nat Genet* **53**, 403–411 (2021).
981 69. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M. D. Doublet identification
982 in single-cell sequencing data using scDblFinder. *F1000Res* **10**, 979 (2021).

983 70. Korsunsky, I. *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat*
984 *Methods* **16**, 1289–1296 (2019).
985 71. Machlab, D. *et al.* monaLisa: an R/Bioconductor package for identifying regulatory motifs.
986 *Bioinformatics* **38**, 2624–2625 (2022).
987

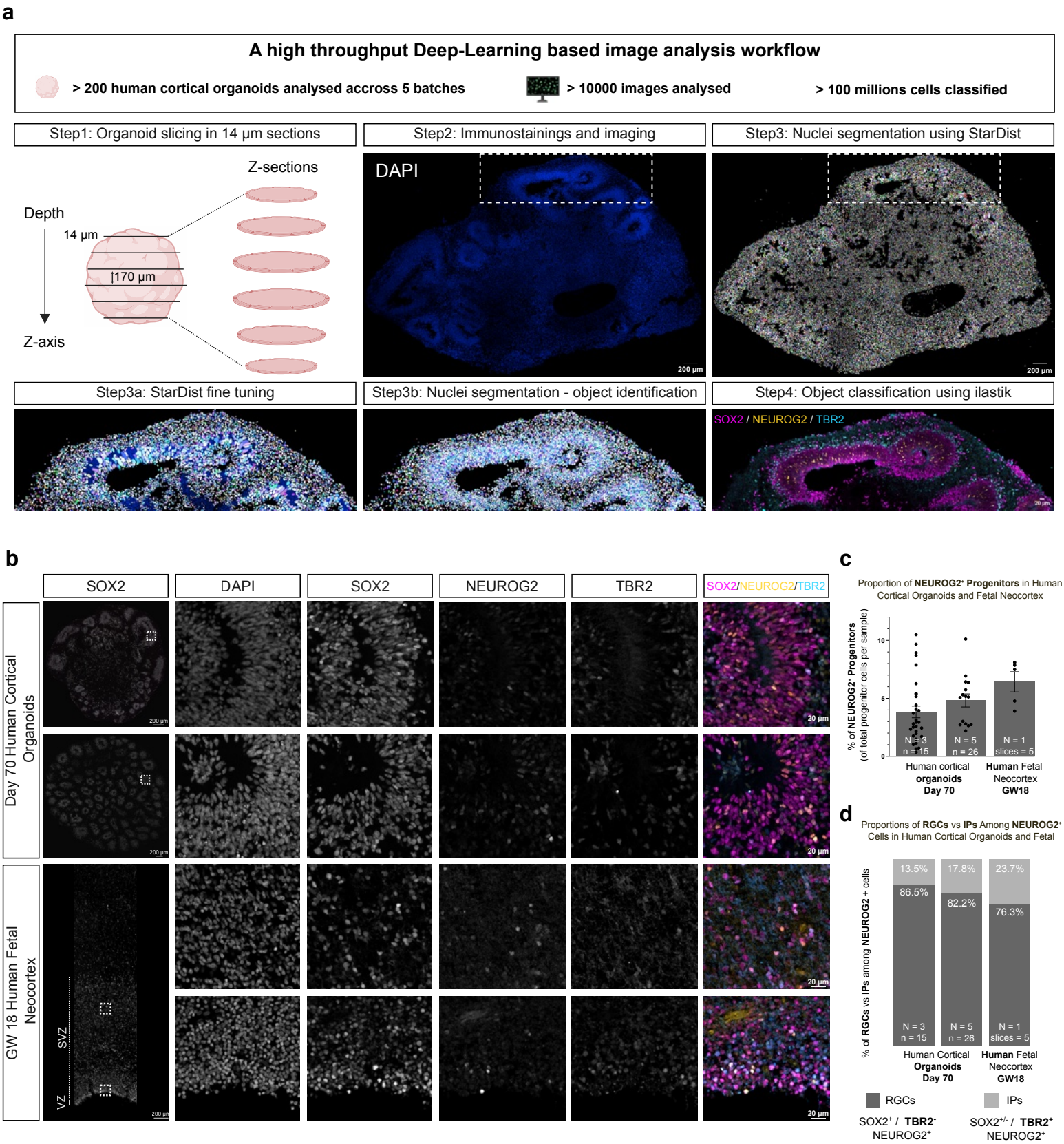


Fig.1: High-throughput deep learning-based image analysis of cell subtypes in human cortical organoids and fetal neocortex.
(a) Overview of the high-throughput image analysis workflow. Step 1: Organoids were sectioned at 14 µm thickness to generate multiple slices across the tissue depth. Step 2: Sections were immunostained and imaged for DAPI and relevant markers. Step 3: Nuclei were segmented using StarDist, a deep learning-based algorithm. Step 3a: initial segmentation using the pretrained StarDist model. Step 3b: segmentation refined using a custom fine-tuned StarDist model (haug2). Step 4: Cells were classified based on marker expression using ilastik. The pipeline was applied to 200 organoids across 5 independent differentiations, analyzing >10,000 images and classifying over 100 million cells. **(b)** Representative confocal images of two wild-type clones of iPSC-derived cortical organoids at day 70 (top rows) and GW18 human fetal neocortex (bottom rows, corresponding to the subventricular and ventricular zones), stained for DAPI, SOX2, NEUROG2, and TBR2. Merged images show SOX2 (magenta), NEUROG2 (yellow), and TBR2 (cyan). Dotted boxes indicate regions shown at higher magnification on the right. Scale bars, 200 µm (overview), 20 µm (zoom). **(c)** Quantification of NEUROG2+ progenitors, performed using the pipeline described in (A), expressed as a percentage of total SOX2+ and/or TBR2+ progenitor cells in day 70 human cortical organoids and GW18 fetal neocortex. **(d)** Proportion of radial glia cells (RGCs, SOX2+TBR2−) and intermediate progenitors (IPs, TBR2+ and SOX2+/-) among all NEUROG2+ cells in human cortical organoids and fetal neocortex.
N = number of biological replicates (organoid differentiations or embryos); n/slices = number of organoids or tissue sections analyzed.

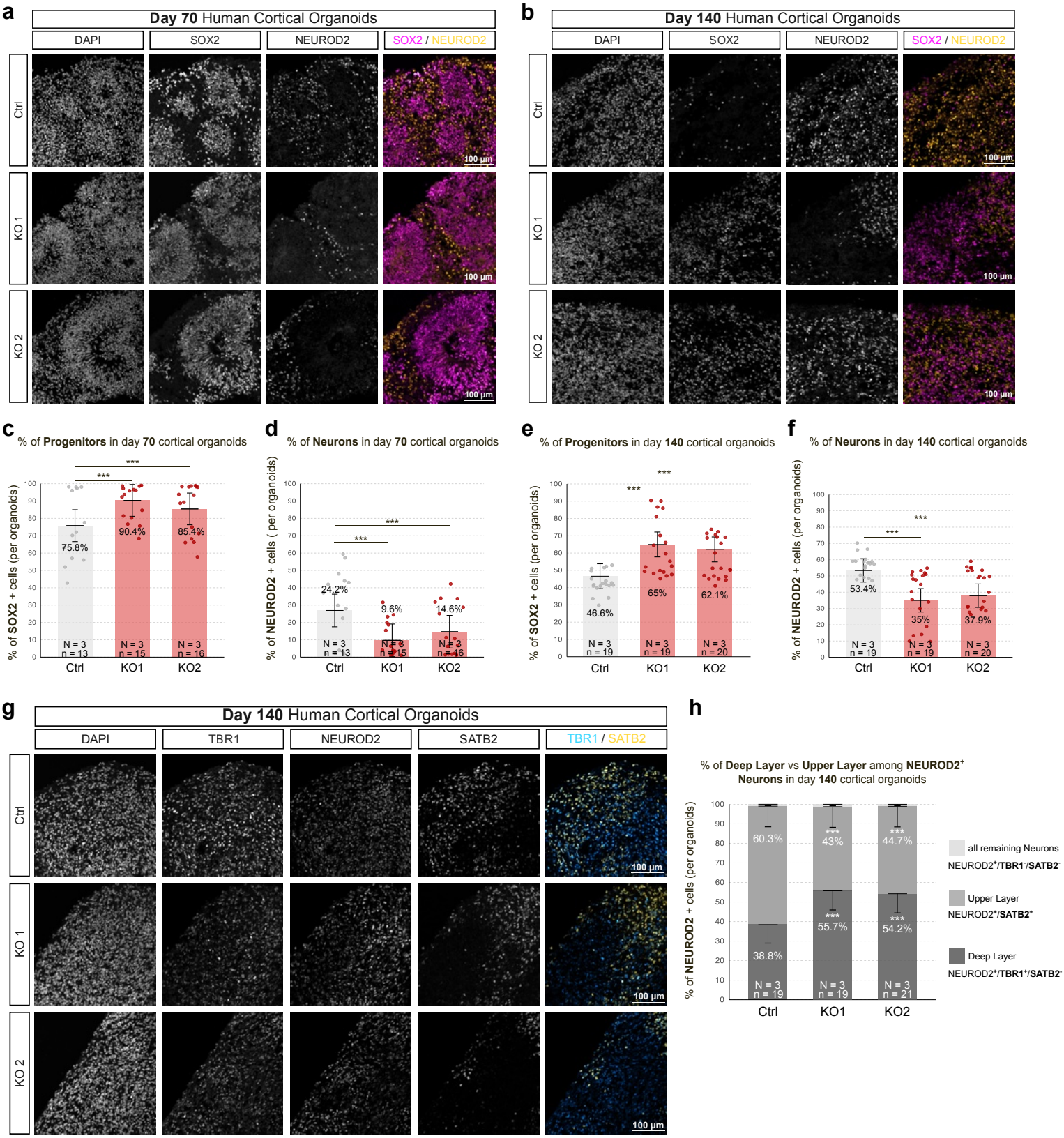


Fig.2: NEUROG2 regulates human cortical neurogenesis.
(a–b) Representative confocal images of human cortical organoids at day 70 **(a)** and day 140 **(b)**, stained for SOX2 (magenta), NEUROD2 (yellow), and DAPI (gray). Ctrl: wild-type clone; KO1 and KO2: NEUROG2 knockout clones. Scale bars, 100 µm. **(c)** Quantification of the proportion of SOX2⁺ progenitors at day 70, expressed as a percentage of all antibody-stained cells. **(d)** Quantification of the proportion of NEUROD2⁺ neurons at day 70, expressed as a percentage of all antibody-stained cells. **(e)** Quantification of the proportion of SOX2⁺ progenitors at day 140, expressed as a percentage of all antibody-stained cells. **(f)** Quantification of the proportion of NEUROD2⁺ neurons at day 140, expressed as a percentage of all antibody-stained cells. **(g)** Representative confocal images of day 140 organoids stained for TBR1 (cyan), SATB2 (yellow), and NEUROD2 (gray). Scale bars, 100 µm. **(h)** Quantification of the proportion of deep-layer (TBR1⁺) and upper-layer (SATB2⁺) neurons among all NEUROD2⁺ cells, expressed per organoid.Bar graphs show adjusted means ± SEM derived from a linear mixed-effects model (group as fixed effect; differentiation batch as random effect). Pairwise comparisons were performed using estimated marginal means (emmeans); outliers (residuals > 2 SD) were excluded. N = number of independent differentiations; n = number of organoids. $p < 0.05$; $p < 0.01$; $p < 0.001$.

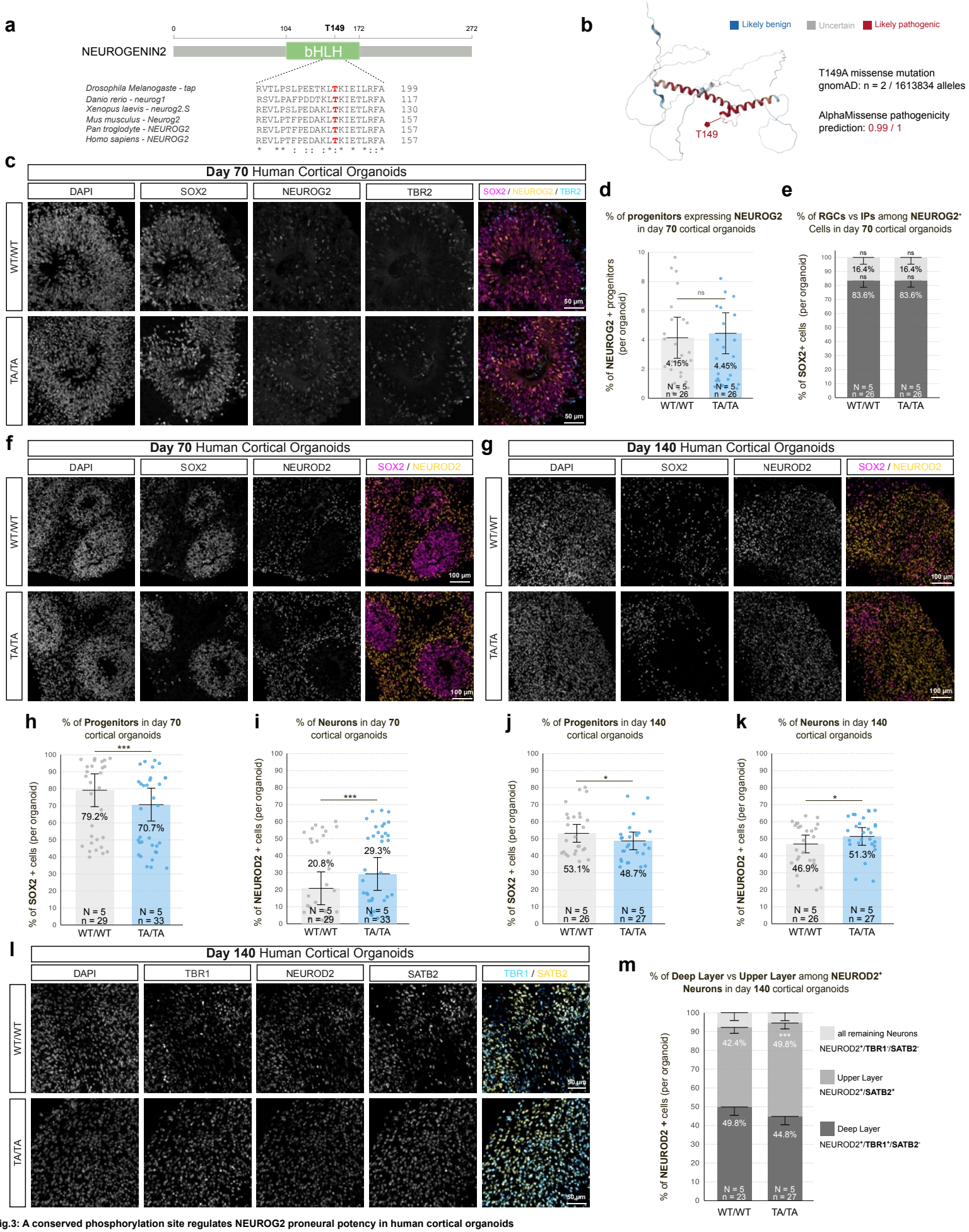


Fig.3: A conserved phosphorylation site regulates NEUROG2 proneural potency in human cortical organoids

(a) Multiple sequence alignment of the bHLH domain from NEUROG2 orthologs across species, highlighting the conserved threonine residue (T149 and equivalents) in red. Asterisks (*) denote fully conserved residues. Sequences shown correspond to: *tap* (Drosophila melanogaster), *neurog1* (Danio rerio), *neurog2.S* (Xenopus laevis), *Neurog2* (Mus musculus), *NEUROG2* (Pan troglodytes), and *NEUROG2* (Homo sapiens). **(b)** Lollipop plot of the human NEUROG2 protein (RefSeq NP_076924) showing the T149A missense variant (allele frequency 1.24×10^{-6}), observed in two heterozygous individuals (gnomAD v4.0). The bHLH domain (residues 104–172) is shaded. AlphaFold2 predicts high pathogenicity for missense variants in this domain (score: 0.99/1). **(c)** Representative confocal images of day 70 human cortical organoids stained for DAPI (gray), SOX2 (magenta), NEUROG2 (yellow), and TBR2 (cyan). WT/WT: wild-type; TA/TA: NEUROG2 T149A phosphomutant clone. Scale bars, 100 µm. **(d)** Quantification of the proportion of progenitor cells co-expressing NEUROG2 at day 70, expressed as a percentage of all antibody-stained cells. **(e)** Proportion of radial glial cells (SOX2⁺/TBR2⁻) and intermediate progenitors (SOX2⁺ or SOX2⁻ but TBR2⁺) among all NEUROG2⁺ cells in day 70 organoids. **(f–g)** Representative confocal images of day 70 **(f)** and day 140 **(g)** organoids stained for SOX2 (magenta), NEUROD2 (yellow), and DAPI (gray). Scale bars, 100 µm. **(h)** Quantification of the proportion of SOX2⁺ progenitors at day 70, expressed as a percentage of all antibody-stained cells. **(I)** Quantification of the proportion of NEUROD2⁺ neurons at day 70, expressed as a percentage of all antibody-stained cells. **(j)** Quantification of the proportion of SOX2⁺ progenitors at day 140, expressed as a percentage of all antibody-stained cells. **(k)** Quantification of the proportion of NEUROD2⁺ neurons at day 140, expressed as a percentage of all antibody-stained cells. **(i)** Representative confocal images of day 140 organoids stained for TBR1 (blue), SATB2 (orange), and NEUROD2 (gray). Scale bars, 100 µm. **(m)** Proportion of deep-layer (TBR1⁺) and upper-layer (SATB2⁺) neurons among all NEUROD2⁺ cells in day 140 organoids.Bar graphs show adjusted means ± SEM derived from a linear mixed-effects model (group as fixed effect; differentiation batch as random effect). Pairwise comparisons were performed using estimated marginal means (emmeans); outliers (residuals > 2 SD) were excluded. N = number of independent differentiations; n = number of organoids. $p < 0.05$; $p < 0.01$; $p < 0.001$.

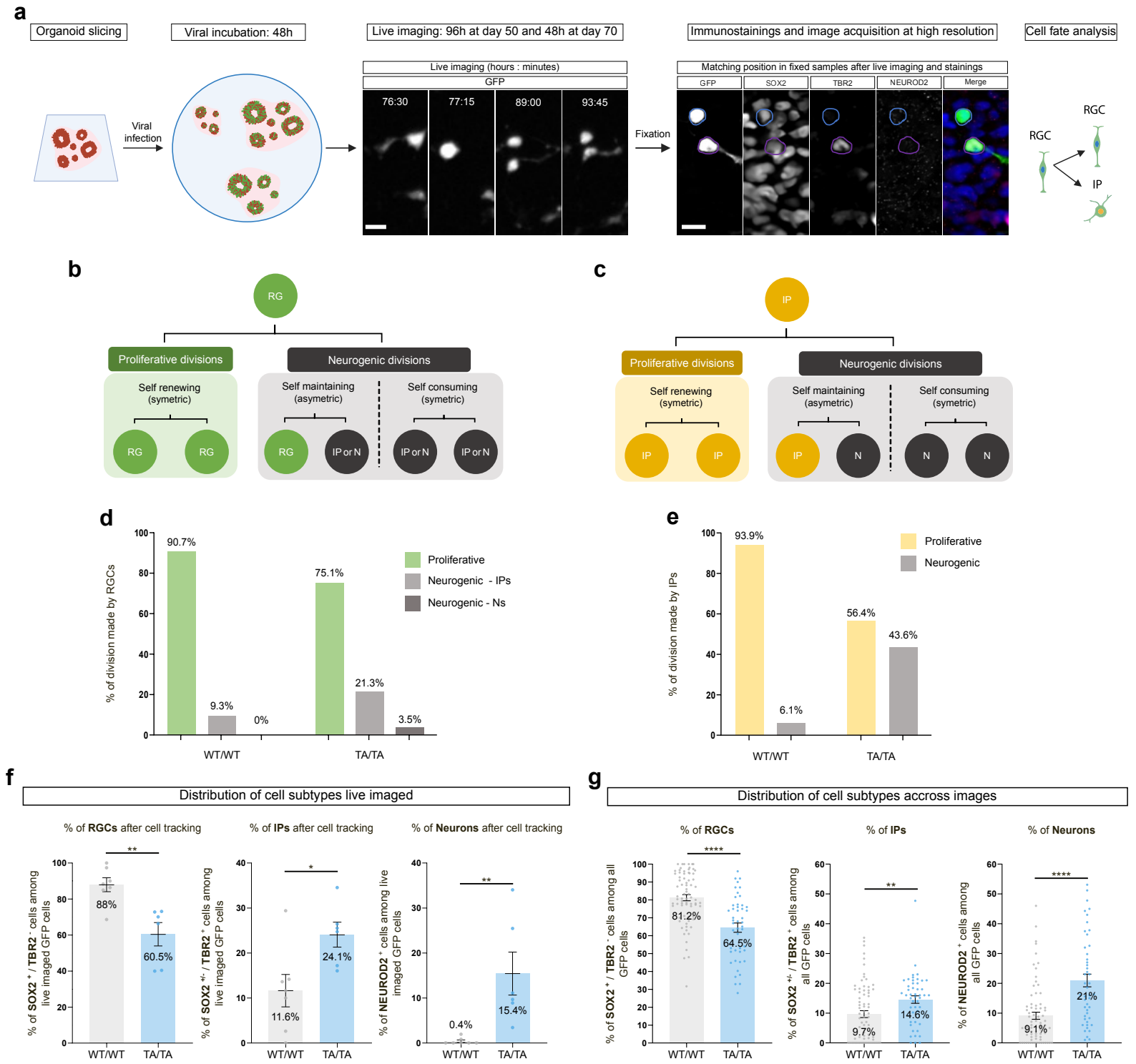


Fig.4: Phosphorylation of NEUROG2 at T149 regulates the balance between proliferative and neurogenic divisions in human cortical organoid progenitors.
(a) Schematic overview of the correlative live imaging and immunostaining workflow. Human cortical organoids were sliced and infected with a GFP-expressing retrovirus for 48h at day 48 followed by live-imaging for 96 h. A second batch has been infected at day 50 or for 48 h followed by live imaging for 48 hours. After imaging, slices were fixed, stained, and re-imaged at high resolution. The division shown in the figure correspond to an asymetric neurogenic division. Data from day 50 and day 70 imaging were pooled. **(b)** Diagram summarizing possible division modes of RGCs, including proliferative (symmetric or asymmetric) and neurogenic divisions. **(c)** Diagram summarizing division modes of intermediate progenitors (IPs), which can self-renew or differentiate into neurons. **(d)** Quantification of division types for tracked WT/WT and TA/TA RGCs. **(e)** Quantification of division types for tracked WT/WT and TA/TA IPs. **(f)** Distribution of final cell subtypes (RGCs, IPs, Neurons) among all tracked progeny after live imaging. **(g)** Distribution of progenitor and neuronal subtypes among all GFP+ cells identified in fixed tissue sections, quantified across multiple fields of view from all infected organoid slices (N = 5 organoids per genotype).
Bar plots represent means ± SEM. Pairwise comparisons performed using Student's t-test. *p < 0.05; **p < 0.01; ***p < 0.001; ****p<0.0001.

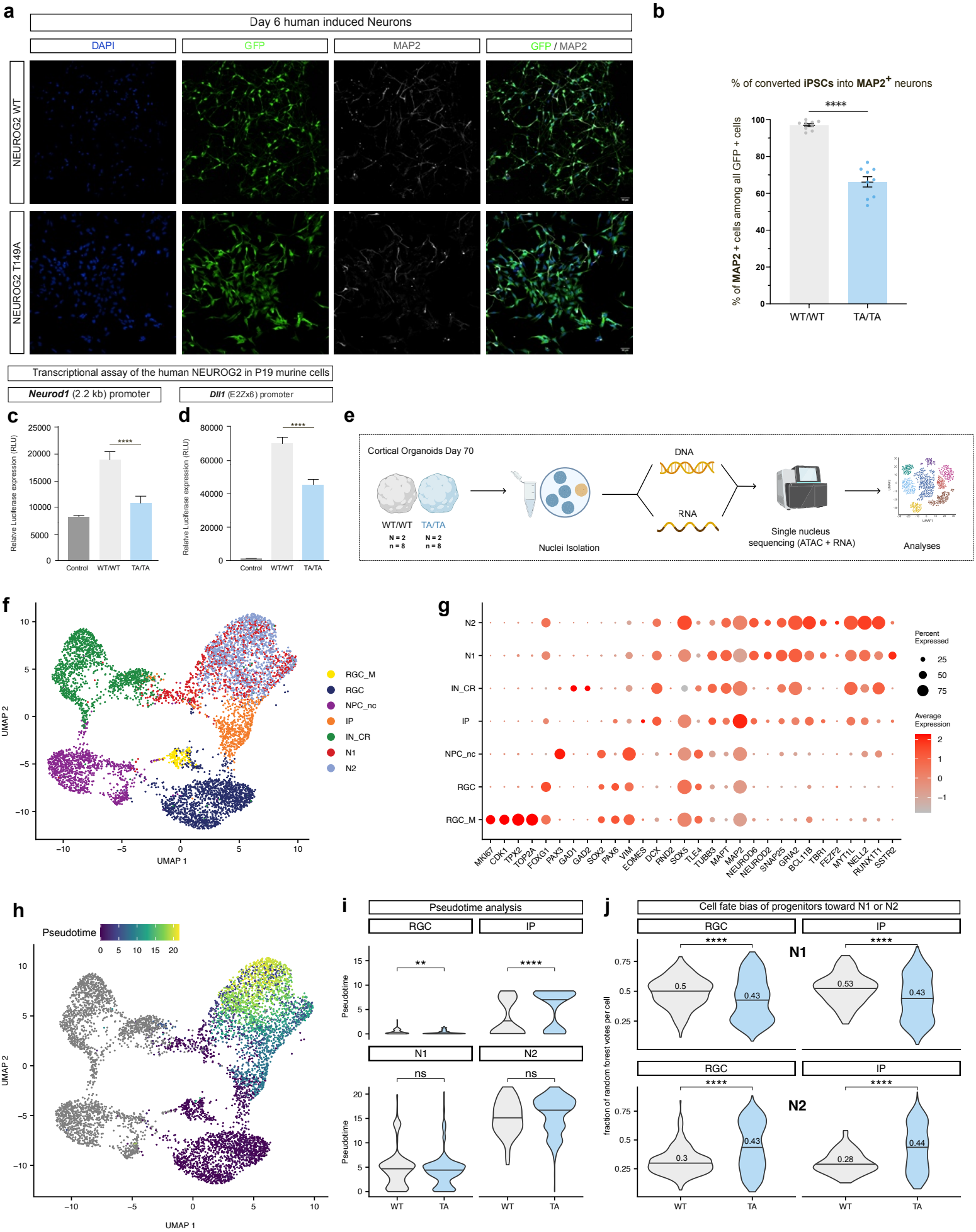


Fig.5: T149 regulate NEUROG2 activty

(a) Immunostaing for GFP (in green) and MAP2 (in gray) in converted iPSCs after 6 days in culture. **(b)** Quantification of MAP2 + neurons among all transfected GFP + cells. **(c)** Quantification of the relative expression of Luciferase in transcriptional assays accross 3 independant experiment for the **(c)** Neurod1 2.2kb promoter or the **(d)** Dll1 promoter, in control: empty plasmid, WT: plasmid containing NEUROG2 WT, TA/TA: plasmid containing NEUROG2 T149A. **(e)** Schematic of the single-nucleus multiomic experimental design. Day 70 WT/WT and TA/TA human cortical organoids (N = 2 differentiations per genotype, n = 8 organoids per differentiation) were processed for parallel snRNA-seq and snATAC-seq. Nuclei were isolated and analyzed using a joint modality embedding strategy. **(f)** UMAP embedding of WT/WT and TA/TA samples, showing integrated clustering of major cortical populations based on a weighted combination of RNA and ATAC profiles. Cell types include radial glia (RGC, RGC_M), intermediate progenitors (IP), early neurons (N1), late neurons (N2), and interneuron-like cells (IN_CR). **(g)** Scaled gene expression of canonical marker genes across identified clusters, shown as dot plots. Dot size represents the percentage of cells expressing the gene; color represents the average scaled expression. **(h)** UMAP embedding colored by pseudotime, with radial glia set as the root. Colors reflect the inferred developmental progression across cortical lineage states. **(i)** Pseudotime values across WT and TA conditions in RGCs, IPs, and neurons (N1 and N2). **(j)** Cell fate bias scores in RGCs and IPs toward either early (N1, top) or late (N2, bottom) neuronal states in WT and TA conditions. Values represent the median fraction of random forest classification votes supporting a given fate. ****p ≤ 0.0001, Wilcoxon test. Violin plots represent the median compared using Wilcoxon test and bar plots represent means ± SEM. *p < 0.05; **p < 0.01; ***p < 0.001; ****p<0.0001.

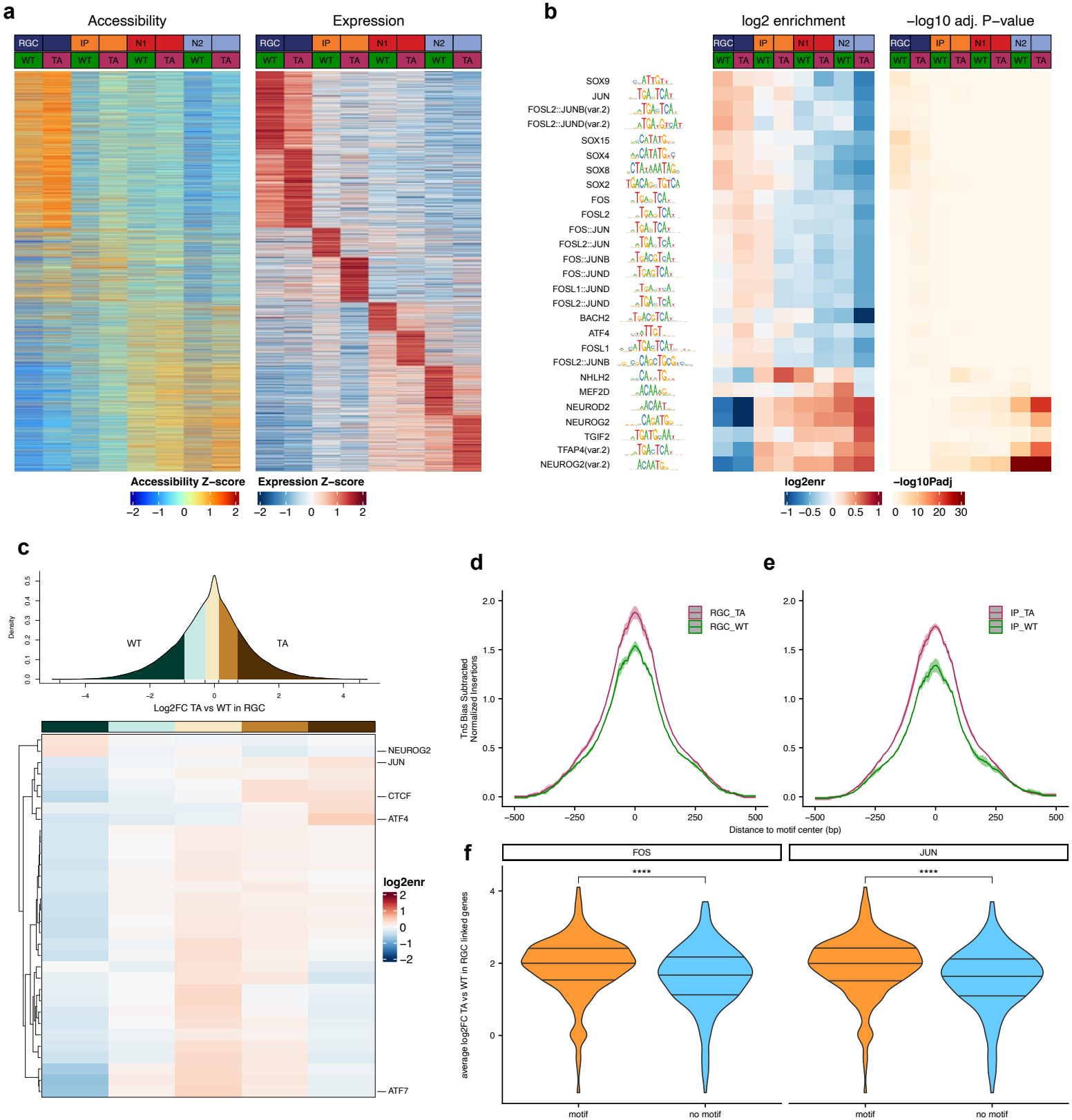


Fig.6: NEUROG2 T149 induces remodifications of the RGCs chromatin landscape in human cortical organoids
(a) Z-score heatmaps of chromatin accessibility (left) and gene expression (right) for enhancer–target gene pairs in radial glia (RGC), intermediate progenitors (IP), and neurons (N1, N2) in WT/WT and TA/TA conditions. Enhancers were linked to genes using co-accessibility and proximity criteria. **(b)** Transcription factor (TF) motif enrichment in distal regulatory elements linked to target genes from (A), per cell type and condition. Heatmaps show log2 enrichment (left) and –log10 adjusted p-values (right) for each motif. **(c)** Enrichment of motifs for expressed TFs in RGC marker peaks, binned by log2 fold change (TA/TA vs. WT/WT). Density plot (top) indicates the distribution of motif-containing peaks per bin; heatmap (bottom) shows motif enrichment (log2 enrichment) per bin, with key TFs labeled. **(d–e)** Aggregate Tn5 transposase insertion profiles (footprints) around JUN motif centers in RGC **(d)** and IP **(e)** marker peaks, comparing TA/TA and WT/WT conditions. **(f)** Violin plots showing pseudobulk average log2 fold change (TA/TA vs. WT/WT) in RGCs for genes linked to peaks containing (or lacking) FOS or JUN motifs. ****p ≤ 0.0001, Wilcoxon test.

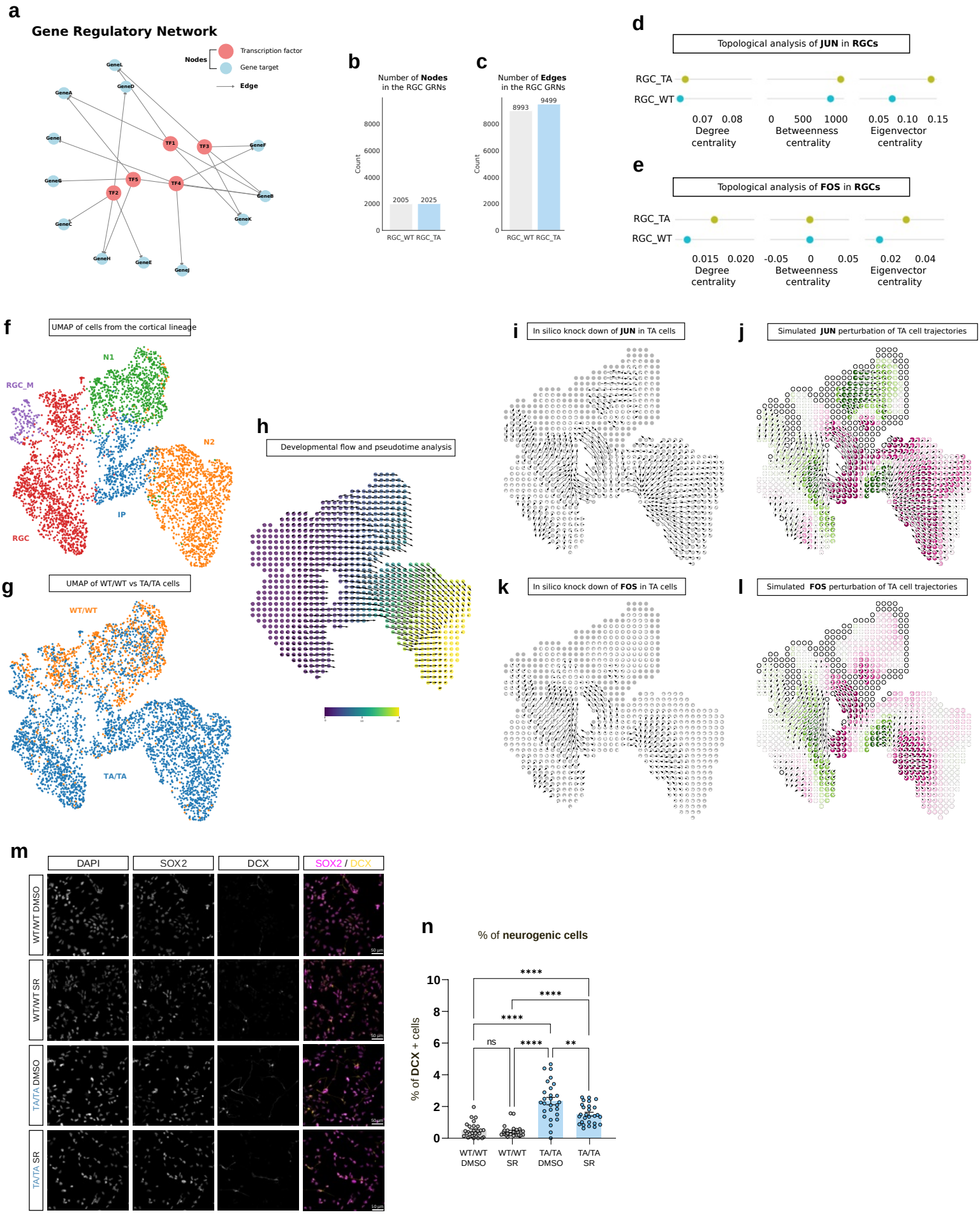


Fig.7: Gene regulatory network analysis identifies the AP-1 complex as a key regulator of neurogenesis in RGCs.
(a) Schematic representation of a gene regulatory network (GRN), with nodes representing genes (TFs or targets) and edges representing regulatory interactions. **(b)** Bar plot showing the number of nodes in GRNs constructed from WT/WT versus TA/TA radial glial cells (RGCs). **(c)** Bar plot showing the number of edges in GRNs from WT/WT versus TA/TA RGCs. **(d)** Topological properties of JUN in GRNs from WT/WT and TA/TA RGCs: degree, betweenness, and eigenvector centrality. **(e)** Topological properties of FOS in GRNs from WT/WT and TA/TA RGCs, as in (d). **(f)** UMAP of all cortical lineage cells included in the multiomic analysis and used for GRN inference. **(g)** Same UMAP as in (f), colored by genotype (WT/WT or TA/TA). **(h)** Grid-based transformation of the UMAP embedding showing pseudotime overlay and CellOracle-derived 2D developmental vector field representing directionality of differentiation. **(i)** Grid plot showing the perturbation vector field following JUN knock-down in TA/TA cells. **(j)** Inner product of the JUN perturbation vectors with the developmental vectors across the grid space. Purple indicates inhibition of differentiation, green indicates promotion. **(k)** Grid plot showing the perturbation vector field following FOS knock-down in TA/TA cells. **(l)** Inner product of the FOS perturbation vectors with the developmental vectors, as in (j). **(m)** Representative confocal images of neural progenitor cells from cortical organoids treated with DMSO or the AP-1 inhibitor SR-11302, stained for SOX2 (magenta), DCX (yellow), and DAPI (gray). Scale bar, 25 µm. **(n)** Quantification of DCX⁺ cells shown in (m), expressed as a percentage of total cells. One Way ANOVA followed by Tukey-corrected pairwise tests: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

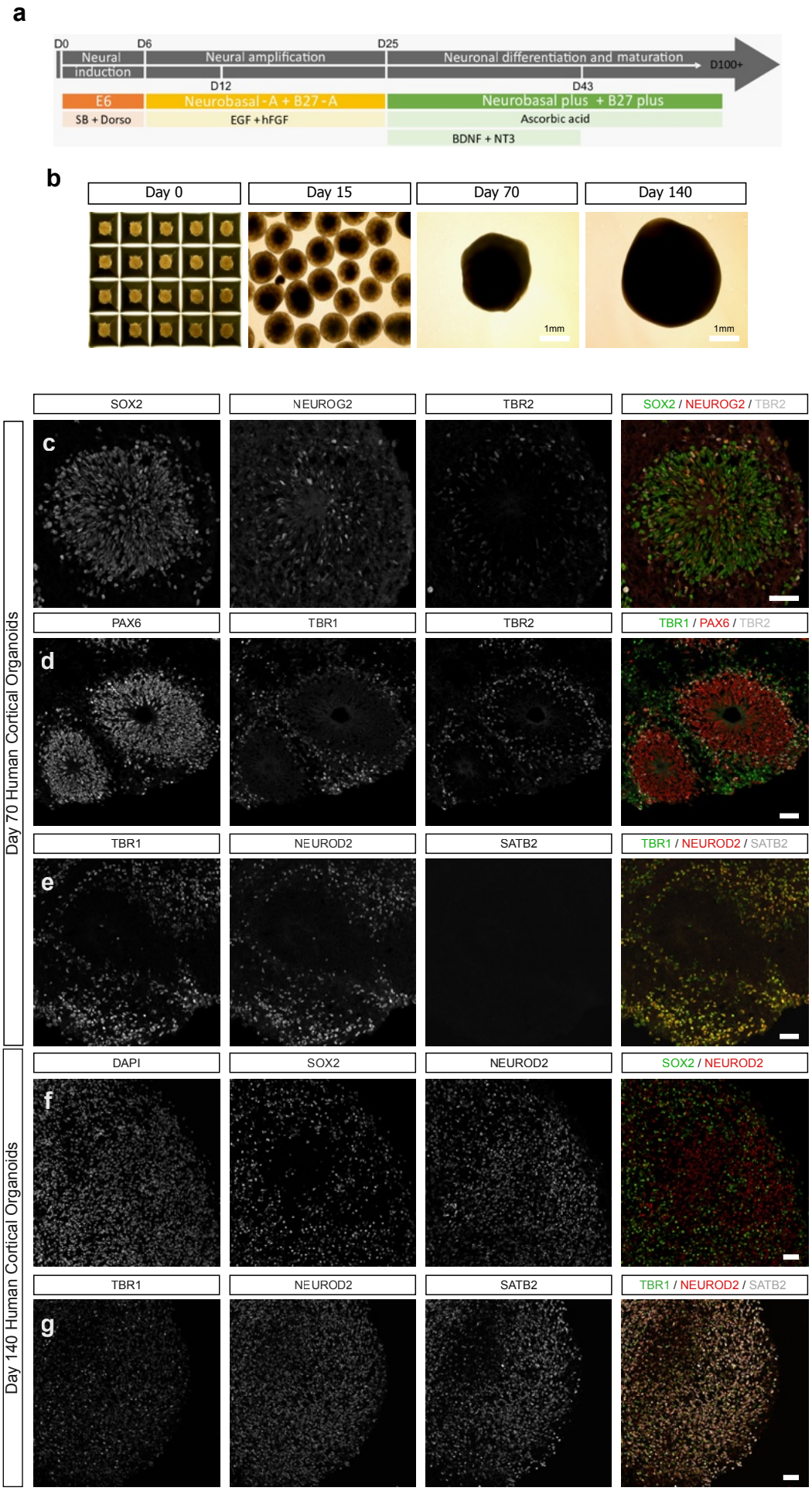
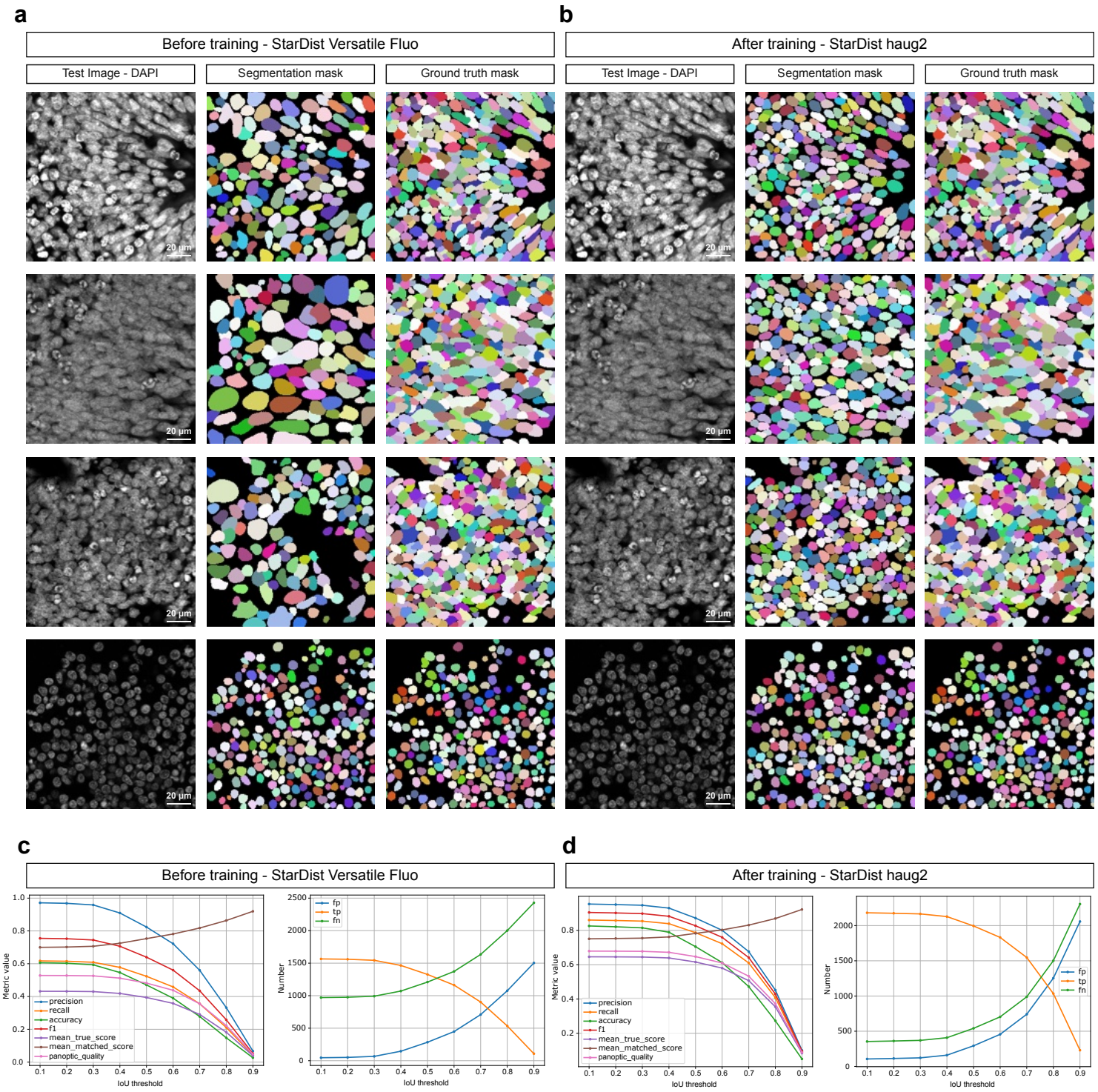


Fig.S1: Validation of the 3D differentiation protocol of human iPSCs into cortical organoids.
(a) Schematic representation of the protocol used, adapted from Sloan et al., 2018. **(b)** Brightfield images of human cortical organoids at different developmental stages: from day 0 in AggreWell plates to day 15 (rosette formation), day 70 (early/mid-stage), and day 140 (late-stage) of cortical organoid development. **(c)** Confocal images of hCOs at day 70 stained for SOX2, NEUROG2, and TBR2. **(D)** Confocal images of hCOs at day 70 stained for PAX6, TBR1, and TBR2. **(e)** Confocal images of hCOs at day 70 stained for TBR1, NEUROG2, and SATB2. **(f)** Confocal images of hCOs at day 140 stained for DAPI, SOX2, and NEUROD2. **(g)** Confocal images of hCOs at day 140 stained for TBR1, NEUROD2, and SATB2. In all images, scale bars represent 50 µm.



e

Metric comparison before and after fine tuning of StarDist at IoU = 0.5

Metric	StarDist 2D fluo (pretrained)	StarDist haug2 (fine tuned)
Precision	0.82	0.87
Recall	0.52	0.79
F1 score	0.64	0.83
Accuracy	0.47	0.71
Panoptic Quality (PQ)	0.48	0.65
Mean Matched Score	0.75	0.78

Fig.S2: Fine-tuning of the StarDist model using manually annotated nuclei of human cortical organoids.
(a) Representative test images (DAPI) used to evaluate the pretrained StarDist 2D Versatile Fluorescence model segmentation performance shown in the following panel with the annotated ground truth masks. Scale bars: 20µm. **(b)** Representative test images (DAPI) used to evaluate the fine tuned StarDist haug2 model segmentation performance shown in the following panel with the annotated ground truth masks. Scale bars: 20µm. **(c)** Performance metrics of the pretrained StarDist 2D model plotted across different intersection over union (IoU) thresholds. **(d)** Performance metrics of the fine-tuned StarDist haug2 model across IoU thresholds, showing improved prediction accuracy and precision. **(e)** Comparison of segmentation metrics at IoU = 0.5 for the pretrained and fine-tuned models, including Precision, Recall, F1 score, Accuracy, Panoptic Quality (PQ), and Mean Matched Score.

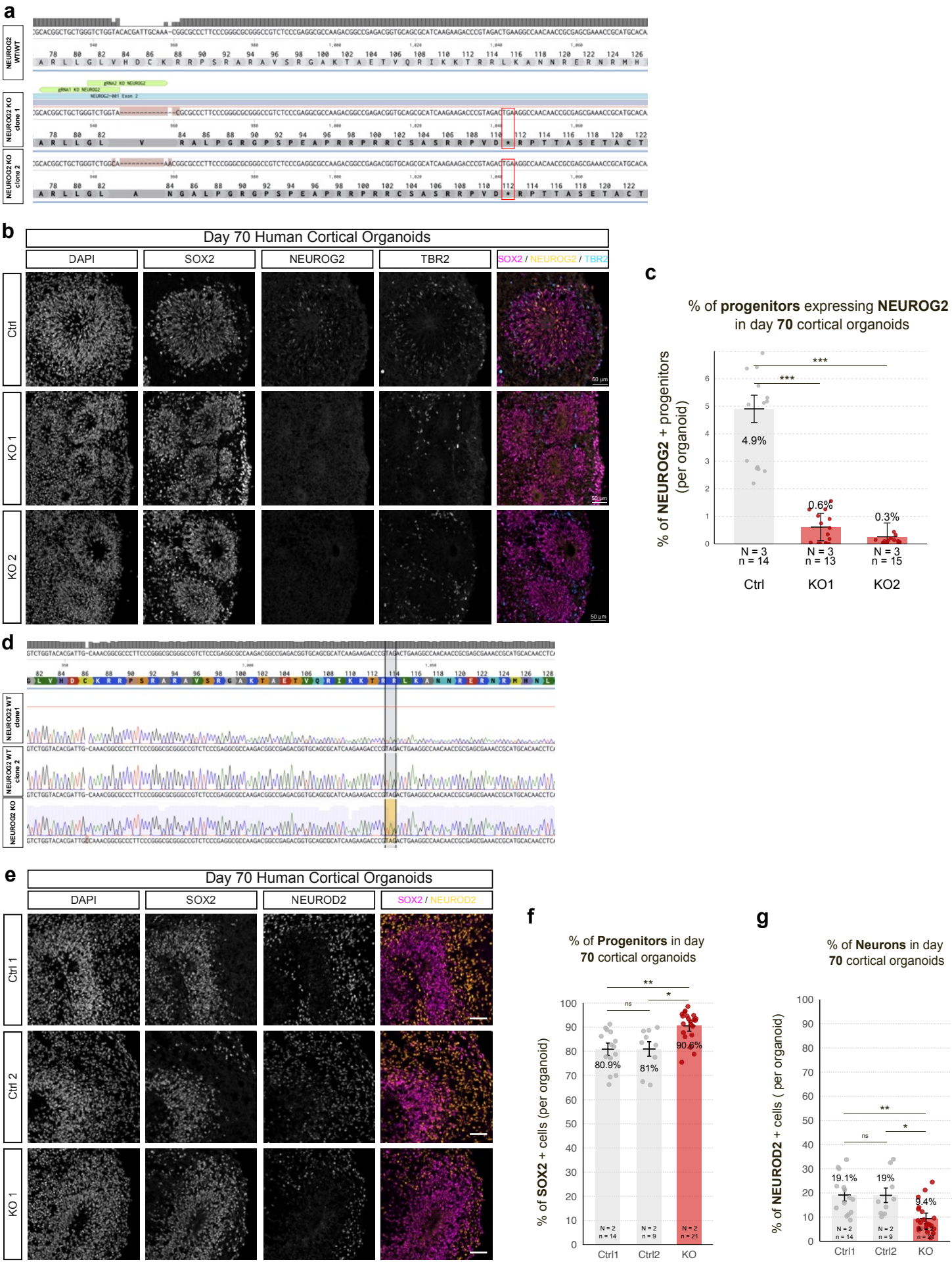
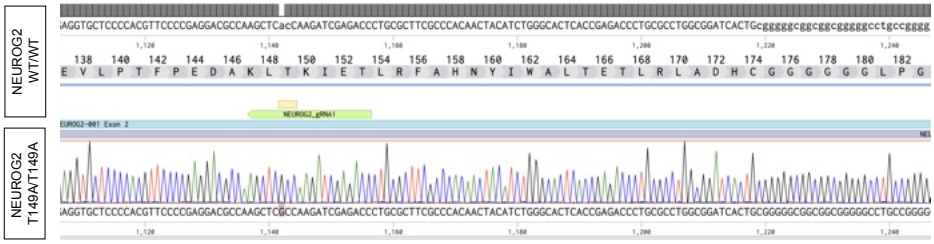


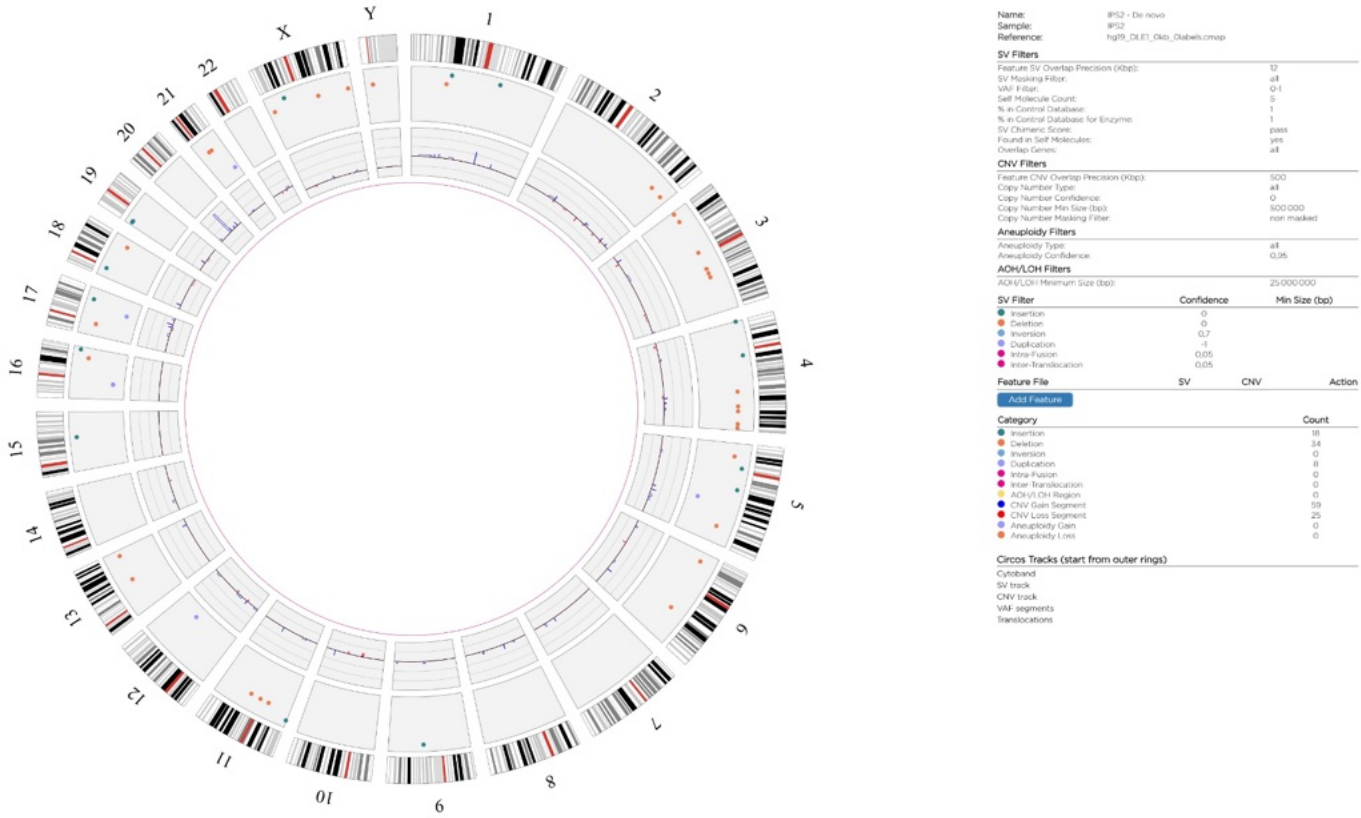
Fig.S3: Validation of NEUROG2 CRISPR/Cas9 knockout and associated phenotype in two human iPSC lines.
(a) Comparison of sequencing results from the control clone and two NEUROG2 KO clones of the WTSIi008-A iPSC line, showing deletions of 14 and 11 nucleotides, respectively, leading to premature stop codons (indicated by *). **(b)** Representative confocal images of human cortical organoids at day 70 stained for SOX2 (magenta), NEUROG2 (yellow), and TBR2 (cyan), plus DAPI (gray), for clones derived from the WTSIi008-A iPSC line: Ctrl: wild-type clone 1; KO1: NEUROG2 KO clone 1; KO2: NEUROG2 KO clone 2. Scale bars, 50 µm. **(c)** Quantification of the proportion of NEUROG2⁺ progenitors in day 70 cortical organoids, expressed as a percentage of all stained cells. **(d)** Comparison of sequencing results from two control clones and one NEUROG2 KO clone of the GM25256E iPSC line, showing a 1-nucleotide insertion leading to a premature stop codon (indicated by brackets). **(e)** Representative confocal images of day 70 human cortical organoids stained for SOX2 (magenta), NEUROD2 (yellow), and DAPI (gray), for clones derived from the GM25256*E iPSC line: Ctrl1: wild-type clone 1; Ctrl2: wild-type clone 2; KO1: NEUROG2 KO clone 1. Scale bars, 50 µm. **(f)** Quantification of the proportion of SOX2⁺ progenitors in day 70 cortical organoids, expressed as a percentage of all antibody stained cells. **(g)** Quantification of the proportion of NEUROD2⁺ neurons in day 70 cortical organoids, expressed as a percentage of all antibody stained cells.Bar graphs show adjusted means ± SEM derived from a linear mixed-effects model (group as fixed effect; differentiation batch as random effect). Pairwise comparisons were performed using estimated marginal means (emmeans); outliers (residuals > 2 SD) were excluded. N = number of independent differentiations; n = number of organoids. $p < 0.05$; $p < 0.01$; $p < 0.001$.

a



b

Parental clone of the WTSli008-A iPSCs line (before CRISPR/Cas9 genome editing)



c

NEUROG2 T149A phosphomutant clone of the WTSli008-A iPSCs line (after CRISPR/Cas9 genome editing)

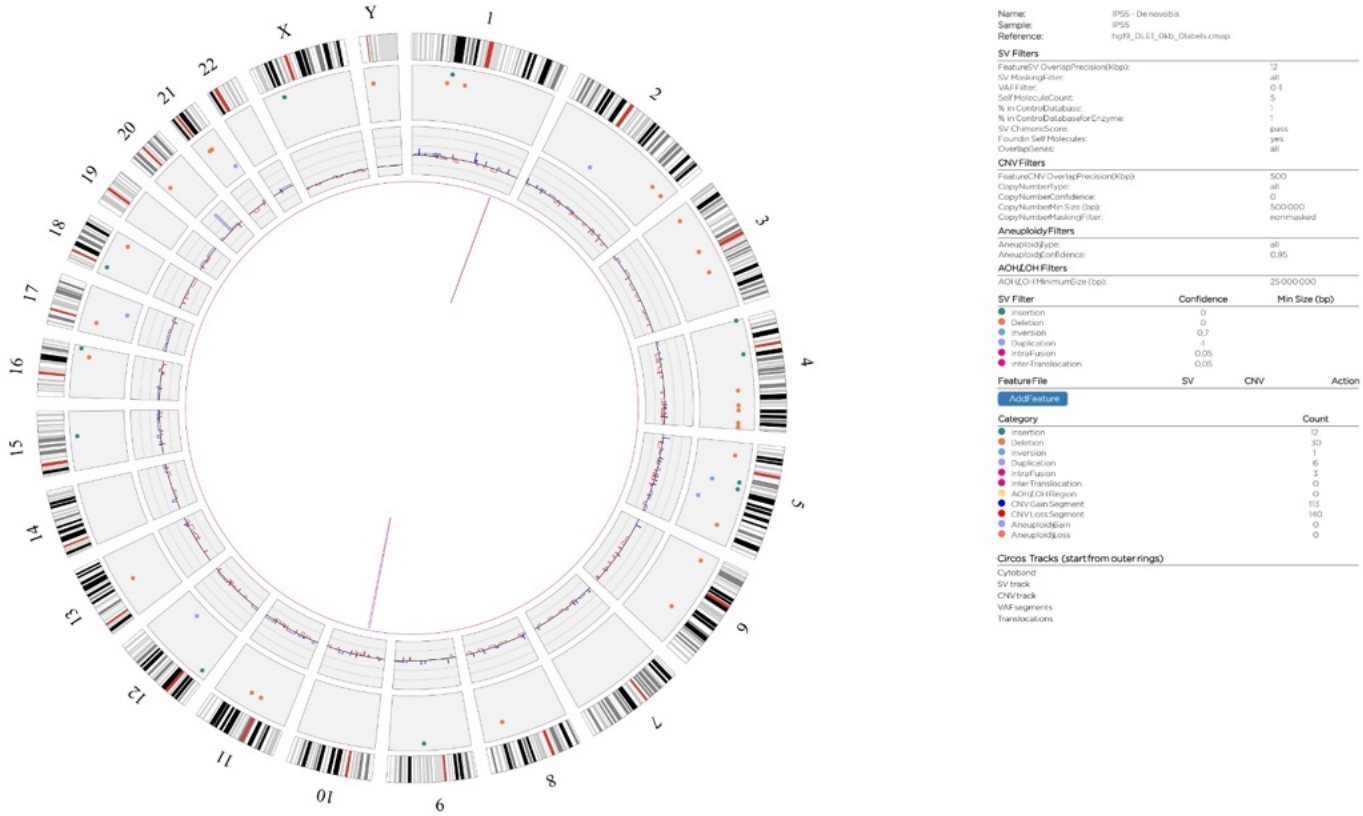


Fig.S4: Validation of NEUROG2 T149A genome editing by CRISPR/Cas9 in the WTSli008-A iPSC line.
(a) Comparison of sequencing results from the NEUROG2 T149A phosphomutant and the wild-type NEUROG2 sequence after CRISPR/Cas9 genome editing. **(b)** Circos plot and associated filter table from optical genome mapping of the parental WTSli008-A iPSC line using Bionano. Tracks are shown from outermost to innermost: chromosome ideogram, interstitial structural variants, copy number variants, and proximity lines for translocations, insertions, and inversions. In short a large duplication of the chromosome 20q has been found in the parental line. **(c)** Circos plot and corresponding filter table from genome analysis of the NEUROG2 T149A/T149A phosphomutant after CRISPR/Cas9 editing, validating the absence of off-target genomic alterations induced by the editing process. All minor detected structural variants compared to pannel (b) have been classified as false positive after careful manual verification.

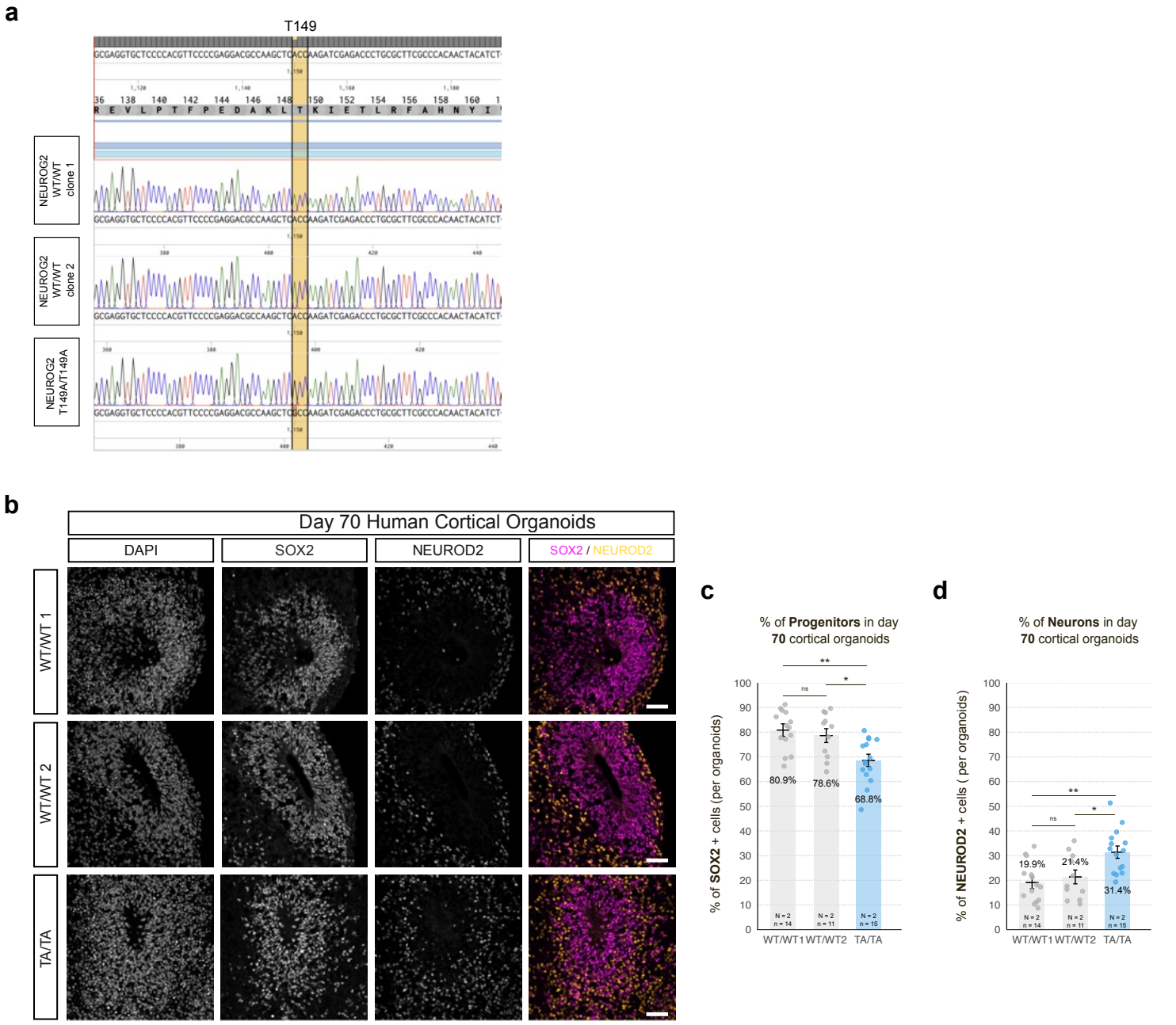


Fig.S5: Validation of NEUROG2 T149A genome editing by CRISPR/Cas9 and associated phenotype in GM25256*E iPSC-derived cortical organoids.
(a) Comparison of sequencing results from the NEUROG2 T149A phosphomutant (TA/TA) and two wild-type controls (WT/WT1 and WT/WT2), aligned to the reference NEUROG2 sequence. **(b)** Representative confocal images of day 70 human cortical organoids stained for SOX2 (magenta), NEUROD2 (yellow), and DAPI (gray), for clones derived from the GM25256*E iPSC line: WT/WT1: wild-type clone 1; WT/WT2: wild-type clone 2; TA/TA: NEUROG2 T149A phosphomutant. Scale bars, 50 µm. **(c)** Quantification of the proportion of SOX2⁺ progenitors at day 70, expressed as a percentage of all antibody-stained cells. **(d)** Quantification of the proportion of NEUROD2⁺ neurons at day 70, expressed as a percentage of all antibody-stained cells.Bar graphs show adjusted means ± SEM derived from a linear mixed-effects model (group as fixed effect; differentiation batch as random effect). Pairwise comparisons were performed using estimated marginal means (emmeans); outliers (residuals > 2 SD) were excluded. N = number of independent differentiations; n = number of organoids. $p < 0.05$; $p < 0.01$; $p < 0.001$.

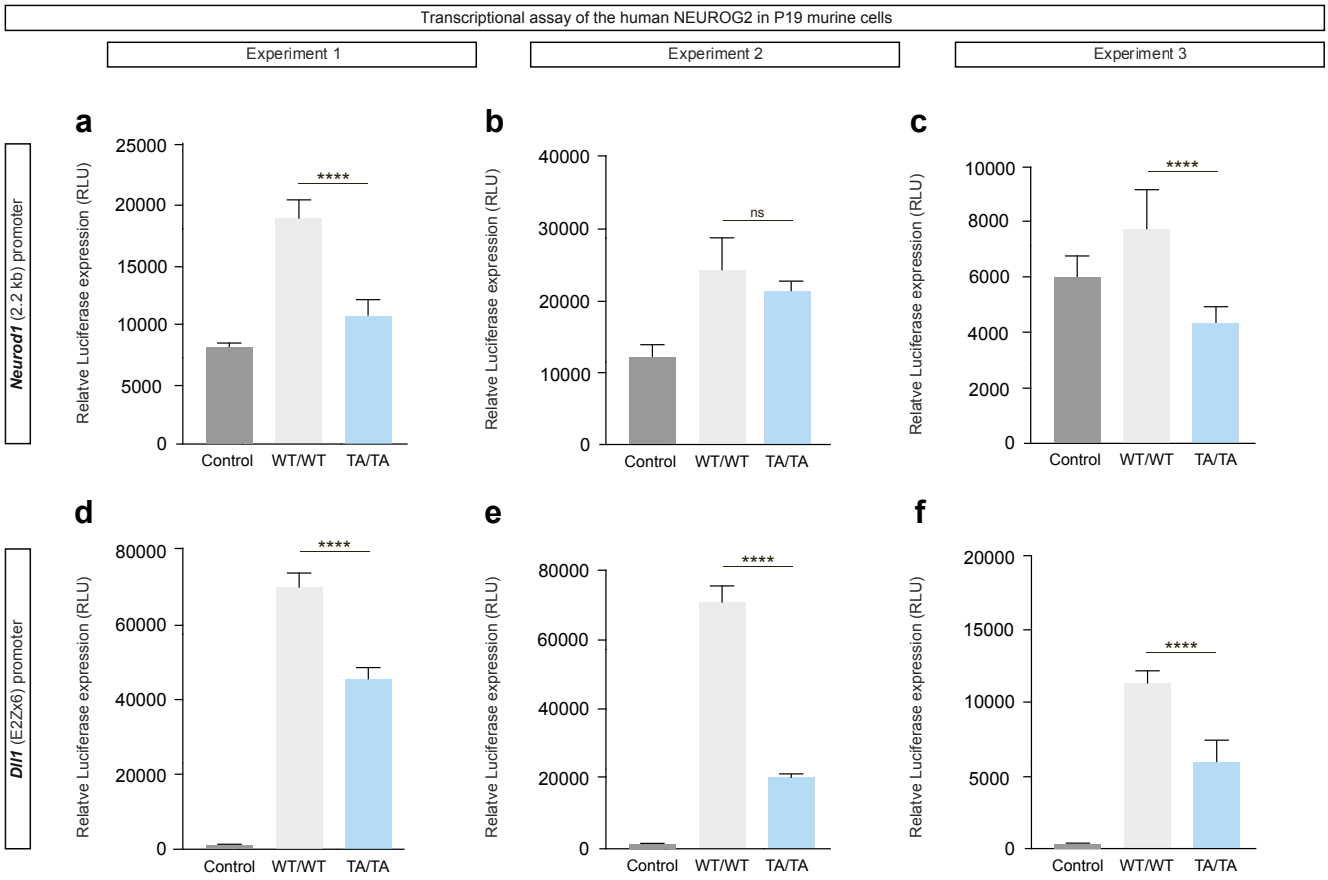


Fig.S6: Unphosphorylated T149 induces a partial loss of transactivational properties of NEUROG2
(a-c) Quantification of the relative expression of Luciferase in transcriptional assays accross 3 independant experiment for the *Neurod1* 2.2kb promoter in control: empty plasmid, WT: plasmid containing *NEUROG2* WT, TA/TA: plasmid containing *NEUROG2* T149A. **(d-f)** Quantification of the relative expression of Luciferase in transcriptional assays accross 3 independant experiments for the *Dll1* promoter in control: empty plasmid, WT: plasmid containing *NEUROG2* WT, TA/TA: plasmid containing *NEUROG2* T149A.
Bar plots represent means ± SEM. Pairwise comparisons performed using Student's t-test. *p < 0.05; **p < 0.01; ***p < 0.001; ****p<0.0001.

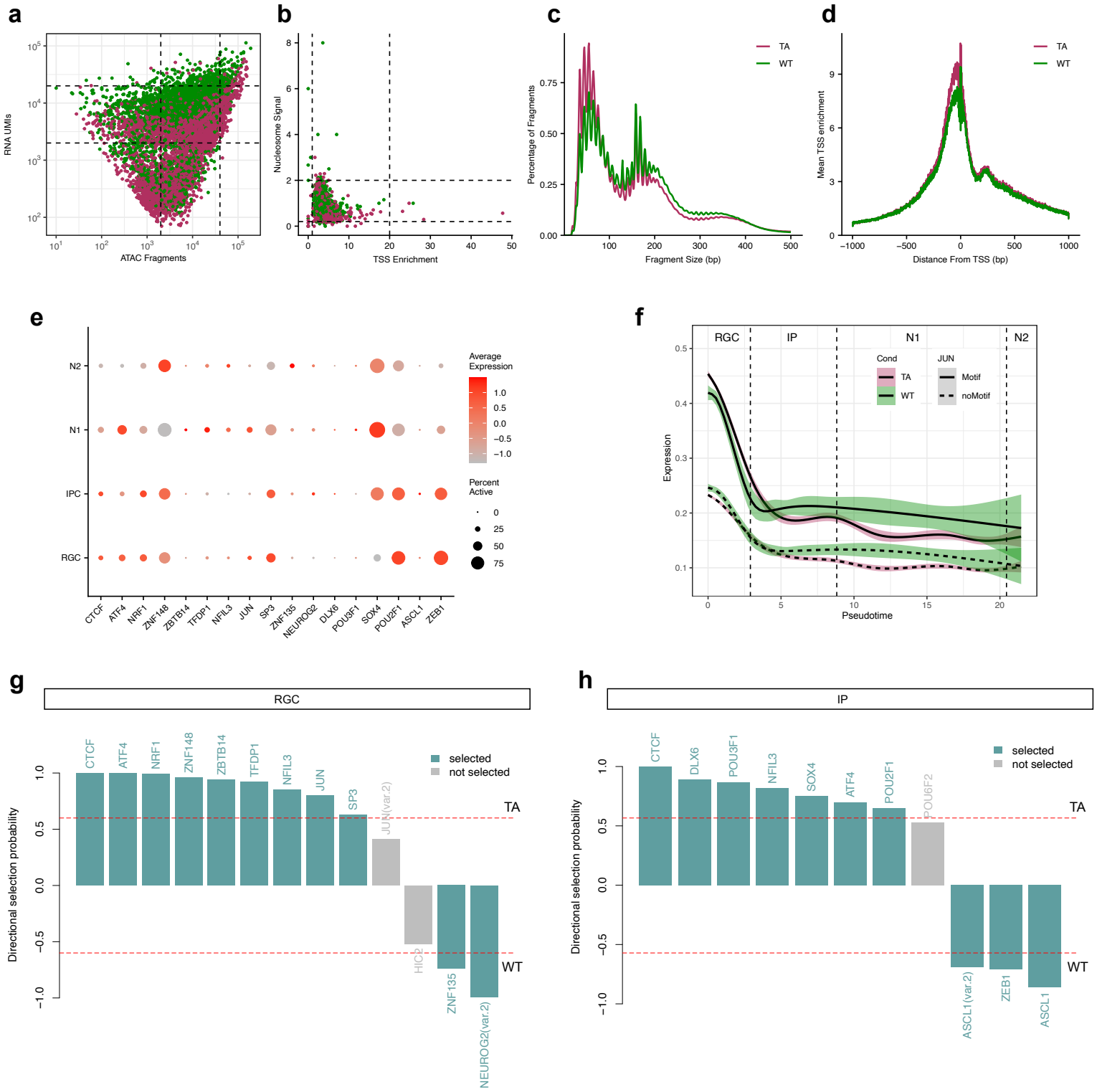


Fig.S7: Quality control and transcription factor activity in single-nucleus multiomic data from human cortical organoids.
(a) Scatter plot of RNA UMIs versus ATAC fragments per nucleus prior to filtering. Dashed lines indicate quality control thresholds used to exclude low-quality cells. **(b)** Scatter plot of nucleosome signal versus transcription start site (TSS) enrichment score before filtering. Dashed lines represent quality control cutoffs. **(c)** Fragment size distribution after filtering of low-quality cells, shown separately for WT/WT and TA/TA conditions. **(d)** Aggregate TSS enrichment profiles post-filtering, aligned to TSSs for WT/WT and TA/TA samples. **(e)** Dot plot showing scaled expression and percentage of cells expressing transcription factors with the highest selection probabilities from regression analysis in RGC and IP populations. **(f)** Gene expression dynamics of RGC-specific genes over pseudotime in IPs, stratified by presence (solid lines) or absence (dashed lines) of a JUN motif in the linked peak. TA/TA and WT/WT conditions are shown in pink and green, respectively. **(g–h)** Regression-based transcription factor motif enrichment analysis for expressed TFs, based on $\log_2$ fold change (TA/TA vs. WT/WT) of marker peaks in RGCs **(g)** and IPs **(h)**. Bars represent directional selection probabilities: positive values indicate enrichment in TA/TA, negative values in WT/WT. TFs with significant selection are shown in color, with selected status indicated.