Multi-omics reveal critical differentiation target for Parkinson's Disease-vulnerable midbrain dopaminergic neurons

Author Information

Roberto Garcia-Swinburn[1*#], Guochang Lyu[1,3#], Judith C. Kreutzmann[1,2,] Anqi Xiong[1], Rika Kojima[2,5], Carmen Abaurre[1,3], Chiara Tremolanti[1], Sandra Gellhaar[1,3], Per Uhlén[1], Per Svenningsson[2,3,4], Gonçalo Castelo-Branco[1], Carmen Salto[1], Onur Dagliyan[1*], Ernest Arenas[1,3†]

Affiliations

[1]Laboratory of Molecular Neurobiology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden

[2]Department of Clinical Neuroscience, Karolinska Institutet, 171 76 Stockholm, Sweden

[3]Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, MD, 20815.

[4]Department of Basic and Clinical Neuroscience, King's College London, WC2R 2LS London, UK

[5]Present address: Sheffield Institute for Translational Neuroscience, University of Sheffield, Western Bank, Sheffield S10 2TN, United Kingdom

*Corresponding author: roberto.garcia.swinburn@ki.se or onur.dagliyan@ki.se

[#]These authors contributed equally

† Deceased 15/09/2024

Abstract

Parkinson's disease (PD) is characterized by the progressive loss of midbrain dopaminergic (mDA) neurons, leading to severe debilitating motor impairment. Recent single-cell analyses have revealed that specific mDA subpopulations, such as SOX6$^+$ AGTR1$^+$ neurons, are most vulnerable to degeneration. Current human pluripotent stem cell differentiation protocols fail to selectively generate these subtypes. Here, we describe a multiomic-guided strategy that enriches SOX6$^+$ mDA

neurons by combining enhancing Sonic Hedgehog agonism with prolonged Wnt activation. This approach accelerates floor plate specification, increases expression of mDA developmental markers, and substantially increases the yield of SOX6$^+$ mature mDA neurons compared to prior attempts. Following intrastriatal transplantation into hemiparkinsonian mice, these cells restored motor function in approximately 4 months and generated grafts containing SOX6$^+$ A9-like neurons. Our work thus establishes a reproducible differentiation platform for generating the PD-susceptible mDA subtype, providing a foundation for precision disease modelling and subtype-targeted cell replacement therapies.

Introduction

Parkinson's disease (PD), the second most prevalent neurodegenerative disorder, lacks curative treatments, severely impacting patients' quality of life. This condition is characterized by the progressive degeneration of midbrain dopaminergic (mDA) neurons in the substantia nigra pars compacta (SNpc). The loss of these neurons, which project primarily to the dorsal striatum, depletes dopaminergic input and leads to the characteristic motor symptoms of PD[1]. Although advances in pharmacological therapies have been made, these treatments are often insufficient in halting disease progression or restoring lost neuronal function in later stages[2]. Cell therapy offers a promising alternative, with the potential to replace the damaged or lost DA neurons critical to PD pathology[3].

Advances in stem cell and midbrain development research have enabled the generation of transplantable mDA progenitors from human pluripotent stem cells (hPSCs)[4–7]. Success in preclinical models has led to the development of clinical-grade mDA progenitors for PD clinical

49 trials[8–11] with recent phase I/II clinical trials reporting safe outcomes, thereby paving the way for
50 future medical use of hPSCs[12,13].

51 However, the translational promise of hPSC-derived cell therapies hinges on the precise
52 composition and quality of the grafted cells. Single-cell transcriptomic studies have begun to
53 molecularly define the diversity of human mDA neurons[14,15], both in development and in hPSC-
54 derived midbrain cultures[7,15–17]. These studies have revealed a critical challenge: not all mDA
55 neurons are equally susceptible to PD. Recent work has reported complications like dyskinesias,
56 which is linked to serotonergic neuron contamination in fetal transplants[18]. Thus, the inability of
57 existing differentiation protocols, which generate heterogeneous mixes of mDA subtypes, to
58 specifically target this critical population represents a significant limitation for both disease
59 modelling and the development of precision cell therapies.

60 Single-cell transcriptomics has revealed significant diversity within mDA neurons in the adult
61 human SNpc, categorizing them into three major subtypes based on the expression of *SOX6*,
62 *CALB1*, and *GAD2*, establishing that the SOX6 subpopulation is the most susceptible to
63 degeneration[19,20]. These subtypes emerge during human mDA neuron development[14], and recent
64 single-cell epigenomic profiling (e.g., ATAC-seq) further elucidates midbrain differentiation[21]. To
65 generate PD-vulnerable SOX6^{+} mDA neurons, we first developed a simplified differentiation
66 protocol based on recently identified key developmental regulators[7], but molecular phenotyping
67 revealed unstable phenotypes at later stages, indicating insufficient patterning. This prompted
68 reviewing multiomic data leading to implement two key protocol modifications: enhanced
69 ventralization via stronger Sonic Hedgehog (SHH) pathway agonization and prolonged Wnt
70 activation improved midbrain patterning, with elevated *OTX2*, *EN1* and *SOX6* expression in
71 engineered cells. Transplanting these engineered cells into hemiparkinsonian (hemiPD) rodent

72 models demonstrated their therapeutic efficacy, yielding grafts enriched with SOX6^{+} mDA
73 neurons and restoring motor function. Thus, we here report a multiomic-guided differentiation
74 strategy that specifically enriches for the critical PD-vulnerable mDA subtype and validates its
75 therapeutic potential *in vivo*.

76 **Results**

77 Longitudinal multiomic single-nuclei sequencing delineates the cell-type
78 composition throughout the process of midbrain dopaminergic neuron
79 differentiation *in vitro*

80 Building upon our recently published method[22], we designed a new protocol focused on precise
81 midbrain floor plate (mFP) patterning and improving cell viability to generate sufficient progenitor
82 cells, while expanding compatibility to a broader range of human induced pluripotent stem cell
83 lines. Our protocol enabled sequential differentiation of human induced pluripotent stem cells
84 (hiPSCs) into midbrain progenitors, positive in mFP progenitor markers such as FOXA2, OTX2
85 and LMX1A, and functional neurons expressing canonical mDA markers like CORIN, TH or
86 Nurr1 (Fig. 1a, Extended Fig. 1).
87 To delineate developmental trajectories, we performed single-cell multiomic sequencing
88 (scATAC-seq + scRNAseq) in benchmarked wild-type iPSCs (KOLF2.1J) at five developmental
89 stages (Days *in vitro* (DIV): DIV11, 16, 28, 42 and 56; Fig. 1b). Following quality control and
90 filtering, we annotated 9,233 high-quality nuclei using integrated manual-automatic classification
91 based on cluster-specific markers and developing human brain references[14] categorized into major
92 types, with sequential differentiation to a mature DA identity based on their enriched GO terms
93 (Fig. 1c,d). Chronologically, early stages (DIV11-16) showed high expression of regional

specification markers (LMX1A, OTX2). By DIV28, dopaminergic commitment marker NR4A2 upregulated, while late stages (DIV42-56) exhibited elevated expression of maturation markers (such as TH, NR4A2, KCNJ6), indicating functional specification (Fig. 1e, Extended Fig. 2). Further chromatin accessibility analysis showed precursor cell types such as floor plate progenitors (ProgFP), medial neuroblasts (NbM) and DA clusters to recapitulate expected chromatin access (*SOX6* and *LMX1A* for mDA progenitors, and *DDC* for mature functional mDA neurons) with ProgFP having a higher potential to differentiate to SOX6 populations (Fig. 1g). We then validated our annotations by comparing them to publicly available scRNA-seq datasets from human fetal midbrain samples[14], by quantifying cell type identity across individual cells. As a result, our annotated cell types closely resembled their endogenous counterparts, as demonstrated by hierarchical clustering (Fig. 1f). High similarity scores (>0.8) were predominantly observed during early stages of the midbrain differentiation and were primarily enriched in ProgFP and NbM, underscoring the effectiveness of our approach in recapitulating endogenous early mDA neuron development. Radial glia cell subtypes, Rgl2 and Rgl3, also displayed transcriptomic profiles akin to their counterparts *in vivo*, hinting at the establishment of an mDA signaling niche to deliver cell type-specific developmental factors for neuronal maturation *in vitro*[23]. At advanced stages, mDA neurons, closely matched their endogenous cells from reference datasets (reference DA and DA0 vs experimental DA similarity score >0.8). These findings show that while mDA neurons matched reference profiles, later stages may benefit from extended maturation cues to enhance postmitotic neuronal survival.

Reconstructed differentiation trajectory resolves transcriptional dynamics during mDA lineage specification and maturation

116 To move beyond chronological staging and define the molecular logic of *in vitro* differentiation,
117 we reconstructed the developmental trajectory of the dopaminergic lineage. We isolated progenitor
118 and neuronal populations (ProgFP, NbM and mDA neurons) from our multiomic dataset and
119 inferred their relationships using pseudotime ordering and RNA velocity. Both analyses confirmed
120 a direct progression from ProgFP and NbM progenitors to mature mDA neurons, validating the
121 protocol's ability to recapitulate this fundamental lineage path (Fig. 1h).
122 Uniform manifold approximation and projection (UMAP) analysis revealed a continuous
123 maturation trajectory for mDA neurons. Gene Ontology (GO) analysis of genes ordered along this
124 path showed a developmentally-relevant progression of biological processes: initial neurite
125 outgrowth was followed by neurotransmitter release, synaptic assembly, and finally synaptic
126 modulation (Extended Fig. 2b), matching the pattern observed in pseudotime analysis.
127 Intriguingly, we observed heterogeneity in the maturation routes. Intriguingly, we observed
128 heterogeneity in the maturation routes. While most cells followed the canonical path, a subset of
129 late-stage cells (DIV42–56) appeared to bypass the axonal development phase characteristic of
130 DIV28, suggesting the existence of alternative maturation mechanisms, potentially influenced by
131 laminar identity cues established earlier in development[24].
132 We next applied single-cell regulatory network inference and clustering (SCENIC) to identify
133 stage-specific regulons. Clustering of AUCell scores for activities of each regulon revealed distinct
134 segregation between early neural progenitors, ProgFP, and NbM, and their descendant mDA
135 neurons along the development trajectory (Fig. 1i). We identified key transcription factors (TFs)
136 orchestrating each developmental transition. In ProgFP cells, the NR3C1 (glucocorticoid receptor)
137 regulon was upregulated, promoting progenitor pool expansion by skewing cells toward
138 proliferation over differentiation[25]. Concurrently, SREBF1, a pro-neural transcription factor

139 controlling mDA neurogenesis[26], activity drove the subsequent transition to neurogenic
140 progenitors (NProg). In NbM cells, pro-neural TFs NEUROD1 and NHLH1 emerged as top
141 regulons, initiating neurogenesis and differentiation in postmitotic cells[27,28]. Subtypes of mDA
142 neurons shared a core set of transcription factors, including PBX1, a canonical regulator of mDA
143 specification and survival[29]. In contrast, transcription factors as POU2F2 and BCL11A displayed
144 more specialized roles at distinct maturation stages[17,30]. In conclusion, trajectory reconstruction
145 confirms that our differentiation protocol recapitulates key embryological processes of mDA
146 development. More importantly, this approach provides a powerful *de novo* framework for
147 identifying the developmental regulatory modules, offering deeper insights into dopaminergic
148 neuron differentiation and maturation.

149 An iterative enhancement in ventralization improves floorplate patterning and 150 accelerates neuronal conversion

151 Given that robust regional patterning underpins high-quality mDA neuron generation, and our
152 single-cell data suggested suboptimal ventralization, we hypothesized that enhanced ventralization
153 would improve patterning outcomes. We induced Engrailed 1 (EN1), a well-established marker
154 for mDA neuron specification[36], while streamlining the protocol for scalability. EN1 plays a
155 pivotal role from the earliest stages of regional specification and maintenance of mDA neurons in
156 adulthood[31]. It supports the midbrain-hindbrain boundary formation during development and
157 sustains mDA function by regulating key genes (*TH, VMAT2, DAT,* and *ALDH1A1*) in postmitotic
158 neurons[31,32]. Thus, EN1 serves as both a surrogate for enhanced mDA neuron specification quality
159 and a predictor of transplantation outcome[9].

160 To enhance EN1 expression, we applied a higher concentration of the SHH agonist
161 purmorphamine[6], and optimized culture conditions to maintain cell viability (Fig. 2a).
162 Subsequently, elevated purmorphamine significantly boosted EN1 expression at DIV11
163 (Optimized: 88.53% ± 6.8%; Original: 20.00% ± 3.65%) without altering core floor plate markers
164 FOXA2 or LMX1A (Fig. 2b,e; Extended Fig. 3a). The effects of this enhanced ventralization were
165 more pronounced at later stages. By DIV16, during the fate transition from NbM towards early
166 mDA neurons, new protocol cultures exhibited an accelerated TH^+ neuron generation (Optimized:
167 26.12% ± 3.93%; Original: 3.57% ± 1.68%) and increased co-expression of TH^+, $EN1^+$ and
168 $SOX6^+$ (Optimized: 75.92% ± 1.26%; Original: 7.56% ± 5.11%; Fig. 2c,f). This improved
169 molecular signature persisted through maturation (DIV28), with sustained co-expression of EN1
170 and SOX6 (Optimized: 65.44% ± 7.06%; Original: 5.35% ± 3.75%; Fig. 2d,g, Extended Fig. 3b).
171 These neurons exhibited functional maturity, displaying calcium dynamics characteristic of mDA
172 neurons by DIV42 (Fig. 2h).
173 To molecularly characterize the improvements driven by enhanced ventralization, we performed
174 single-nuclei RNA sequencing (snRNA-seq) on cells from the new protocol at DIV14,16 and 28.
175 The new protocol accelerated differentiation, with a higher proportion of NProg and NbM
176 emerging at earlier stages, accompanied by a reduction in ProgFP cells (Fig. 3b). Using the same
177 annotation methods as in the previous dataset, the resulting mDA lineage cell types were enriched
178 with classic mDA neuron marker genes (Fig. 3c), including *FOXA2, LMX1A, OTX2, TH,* and
179 *NR4A2*, and showed an expected developmental trajectory across different cell populations (Fig.
180 3d). Critically, postmitotic cells exhibited higher EN1, sustained OTX2 and SOX6, and elevated
181 maturation markers TH, SLC18A2 (VMAT2), and CALB1[33] (Fig. 3f).

182 SCENIC analysis identified the key transcription factors activated by the new protocol (Fig. 3e).
183 The paired-box transcription factor PAX7 transiently became the dominant regulon in NProg,
184 suggesting precise rostro-caudal patterning as it contributes to determine the midbrain-hindbrain
185 boundary during regionalization[34,35]. Other leading regulons in NProg and NbM cells are also
186 actively involved in guiding the neuronal identity acquisition. For instance, POU3F1 and ZEB1
187 are essential regulators of neural lineage commitment, often working synergistically to coordinate
188 correct orders of neuronal development, as POU3F1 promotes early neural fate decisions, while
189 ZEB1 modulates the developmental timing and prevents premature differentiation[36,37]. Notably,
190 PITX3 emerged as a leading regulon in mDA neurons, which reflects the progressive acquisition
191 of mature mDA phenotypes[38] and highlights their molecular similarity to endogenous SNpc mDA
192 neurons (Fig. 3e,f). Regulon activity profiles from the integrated dataset revealed a distinct and
193 significant regulatory impacts of TCF4 and TCF12 under the new protocol (Extended Fig. 4c).
194 Interestingly, the new protocol also appeared to enhance SREBF1 expression throughout
195 differentiation, linking it to its known role in regulating midbrain dopaminergic neurogenesis[26],
196 and its involvement in controlling neuronal dendrite expansion through lipid homeostasis[39].
197 Overall, a targeted intervention to enhance ventralization, guided by multiomic insight, profoundly
198 accelerated and improved the specification of mDA neurons, driving them toward a transcriptomic
199 profile that closely resembles that of endogenous, vulnerable SNpc neurons.

200 Transplanted hiPSC-derived neurons mature into striatum-innervating dopaminergic 201 neurons and restore motor function in hemiPD mice

202 Having established that our optimized protocol enriches ventralized mDA subpopulation *in vitro*,
203 we next evaluated the therapeutic potential of our engineered cells *in vivo* using hemiparkinsonian

204 rodent models. We unilaterally lesioned immunocompromised B6RGS mice by injecting 6-
205 OHDA into the medial forebrain bundle (MFB). Upon confirming hemiparkinsonism, we
206 intrastriatally transplanted 200,000 cells into randomly selected mice. Motor recovery was
207 assessed longitudinally by amphetamine-induced ipsilateral rotations, showing that grafted
208 hemiPD animals returned to pre-lesion state 4 months post-implant (m.p.i.), which was maintained
209 1 month later, ending our grafting experiment (Sham: -0.24 ± 0.73; 6-OHDA: 13.41 ± 1.43;
210 Grafted: -0.70 ±1.18 net ipsilateral rotations/min). To ensure overall motor recovery, we also
211 analysed complex motor coordination by tape removal test (Sham: -0.40 ± 0.34; 6-OHDA: 19.51
212 ± 2.50; Grafted: -2.82 ±0.96 net ipsilateral removal time) and pole test (Sham: 5.00 ± 0.59; 6-
213 OHDA: 21.25 ± 3.53; Grafted: 8.56 ± 1.13 seconds), and spontaneous motor activity by cylinder
214 wall test (Sham: 51.82 ± 3.73%; 6-OHDA: 8.85 ± 4.75%; Grafted: 40.89 ± 3.73% ipsilateral
215 contacts). Human cells were detected in these 5-month grafts by STEM121^{+} staining, confirming
216 that grafted cells did not migrate to other regions (Fig. 4c). Furthermore, human DANs were
217 detected by double positive staining of TH and HuNu, with these cells showing extended
218 processes, characteristic of mDA neurons (Fig. 4d,e). Using stereological quantification of
219 TH^{+}/HuNu^{+} cells (Fig. 4F), grafts were estimated to have 160000 human nuclei (155700 ± 35893),
220 of which about 10% were also stained positively for TH (10.36 ± 0.73%). Additionally, no actively
221 proliferative human cells were found within the grafts (Extended Fig. 5)
222 After establishing the presence of DANs within our grafts, we aimed to analyse if these DANs
223 recapitulated the phenotypical aspects of the most PD-vulnerable portion of nigral DANs which
224 have been reported to be A9-subtype DANs from the SNpc, more specifically the
225 SOX6^{+}/AGTR1a^{+} subpopulation located in the ventral SNpc[19,40]. To this aim, we performed a
226 series of IHC analysis of the grafted DANs. A majority of grafted TH^{+} cells were found to be

227 positive for several markers distinctive to the PD-vulnerable A9-subtype (Fig. 4g,h), such as SOX6
228 (84.67 ± 3.03%), ALDH1A1 (50.67 ± 2.37%) and AGTR1 (64.16 ± 4.97%). In contrast, a minority
229 of the DANS stained positively for the A10-subtype marker Calbindin (29.92 ± 3.45%). Another
230 aspect that we wanted to evaluate is the expression of canonical markers for mDA neurons in the
231 grafted TH^+ cells (Fig. 4i,j). We performed IHC and we identified approximately all of the grafted
232 DANs as $LMX1A^+$ (92.09 ± 3.37%), $FOXA2^+$ (96.66 ± 1.88%), $NR4A2^+$ (96.19 ± 2.16%) and
233 $PITX3^+$ (85.14 ± 1.80%). Finally, we performed another series of IHC to evaluate expression of
234 proteins involved in DA regulation and secretion (Fig. 4k,l). We were able to determine that the
235 grafted DANs expressed DAT (90.03 ± 1.23%), GIRK2 (96.78 ± 2.12%), VMAT (94.24 ± 1.47%)
236 and DDC (100.00 ± 0.00%). Overall, these results show that grafted cells mature into a large
237 percentage of DANs expressing canonical markers for A9-subtype mDA neurons that are able to
238 ameliorate motor degeneration in rodent PD models.

239 snRNA-seq reveals A9-subtype mDA neuron enrichment in grafts

240 To unbiasedly characterize graft compositionand transcriptional fidelity to the transplanted
241 neurons to native human neuron subtypes, we performed snRNA-seq on graft tissue from three
242 transplanted mice. After quality control, we retained 9196 high-quality human nuclei.
243 Unsupervised clustering revealed that the graft composition consisted of four main cell lineages:
244 neurons, radial glia, gliobasts and oligodendroglial lineage cells, with no presence of vascular
245 leptomeningeal meningeal cells (VLMC) contamination, an off-target population in mDA neuron
246 differentiation[41] (Fig. 5a, Extended Fig. 6). Further analysis of the neuronal clusters (defined by
247 RBFOX3) resolved into four distinct populations: dopaminergic (DAergic), glutamatergic,
248 GABAergic ,and a small cluster of neural progenitors (defined by high expression of INA)
249 (Extended Fig. 7a, Extended Fig. 8a).

250 Critically, the DAergic cluster was comprised of 3 distinct subclusters: DA_SOX6, DA_CALB1
251 and DA_ONECUT2 (Fig. 5b). DA clusters exhibited high expression of canonical dopaminergic
252 markers (*TH, DDC, NR4A2, LMX1A, PBX1, SLC18A2* [VMAT2], *SLC6A3* [DAT], *PITX3, KCNJ6*
253 [GIRK2]), whereas GABA and glutamatergic clusters lacked these but expressed *GAD1* and *GAD2*
254 and *SLC17A6* [VGLUT2], respectively. INA clusters were distinguished from glutamatergic
255 clusters by elevated *SLC17A6*, *STMN2* and *INA* expression (Extended Fig. 7b,c, Extended Fig.
256 8b,c) . Within the DA lineage, subclusters were resolved by differential expression of SOX6,
257 CALB1 and ONECUT2 (Fig. 5c,d). Roughly half of DA cells were classified as DA_SOX6 (44.73
258 ± 4.79%), with the remainder predominantly DA_CALB1 and a minority DA_ONECUT2 (Fig.
259 5e). These results confirm our protocol's capacity to produce PD-susceptible $SOX6^+$ mDA subtype
260 *in vivo*.

261 Transcriptomic comparison validates fidelity of grafted neurons to native human 262 midbrain subtypes

263 To rigorously benchmark the identity of our grafted dopaminergic neurons, we compared their
264 transcriptional profiles directly to those of native adult human mDA cells. We performed a
265 MetaNeighbor analysis, a quantitative method for assessing cell type similarity, comparing our
266 graft-derived DA subclusters to the well-annotated dopaminergic subtypes from a reference single
267 cell atlas of the human substantia nigra[19] (Fig. 5f). This analysis showed high distinctive
268 resemblance with a score >0.8 between our DA_SOX6 subcluster and the SOX6_AGTR1
269 subcluster from Kamath dataset, which have been reported to be the most susceptible
270 subpopulation to PD[19,40]. In contrast, the DA_ONECUT2 subcluster showed resemblance to the
271 human CALB1_TRHR reported to be the most resilient. However, the DA_CALB1 subcluster

272 lacked exclusive alignment with any Kamath cluster, instead showing broad similarity to multiple
273 CALB1$^+$ subtypes, indicating incomplete subtype commitment or ongoing differentiation in this
274 population. These results confirm successful *in vivo* differentiation into SOX6$^+$ mDA neurons
275 mirroring PD-vulnerable human counterparts reinforces that we were successfully able to generate
276 cells that differentiate *in vivo* to cells resembling human SOX6 mDA neurons.

277 Gene regulatory network analysis confirms stable commitment of SOX6$^+$ 278 dopaminergic neurons

279 Given the transcriptional similarity of grafted DA neurons to native human subtypes, we assessed
280 whether SOX6$^+$ subtype represented a stably committed fate or a transitional state. GO analysis
281 revealed distinct functional specializations among the dopaminergic subclusters, consistent with
282 advanced maturation (Fig. 5g). The DA_SOX6 population was enriched for terms related to
283 synaptic transmission and organization, hallmarks of functional, integrated neurons. In contrast,
284 the DA_CALB1 and DA_ONECUT2 clusters were associated with terms like "learning" and
285 "regulation of cell projection," suggesting nuanced, subtype-specific functional roles. (Fig. 5g).
286 To directly infer developmental potential, we performed RNA velocity analysis on our DA and
287 immature neurons subclusters. We found that immature neurons have the potential to become
288 DA_CALB1 cells. Interestingly, the DA_CALB1 subclusters have the potential to become either
289 DA_ONECUT2 or DA_SOX6, which in contrast their velocity show that these identities were a
290 stable endpoint, not a transient phase (Fig. 5h, Extended Fig. 8d,e).
291 Further evaluation of the gene regulatory network by SCENIC established that the DA clusters
292 were enriched in exclusive TFs expression profile, not found in the immature neuron, GABA or
293 Glut clusters (Fig. 5i, Extended Fig. 8f). For example, PBX1 regulon networks have been
294 associated with ventral midbrain DA organization and is highly enriched (score >0.8) in our DA

295 cluster. Moreover, the DA_SOX6 is differentially highly enriched in SOX6 (score >0.8) and
296 CUX2 (score >0.6) within the DA clusters. Interestingly, the CUX2 regulon network has been
297 reported to be highly associated with the $SOX6^{+}AGTR1^{+}$ clusters of neurons in human datasets,
298 further validated in PD macaque models[19,42]. In conclusion, this multi-faceted computational
299 analysis demonstrates that our grafts contain mature, synaptically active neurons whose subtype
300 identities are locked in by stable gene regulatory networks. The DA_SOX6 population, in
301 particular, represents a terminal, committed fate, confirming that our protocol generates the
302 precise, stable, and vulnerable neuronal subtype required for accurate disease modeling and
303 effective cell therapy for PD.
304

305

306 ## Discussion

307 This study establishes a multiomic-guided framework for differentiating human iPSCs into mDA
308 neurons with unprecedented molecular and functional fidelity to the PD-vulnerable
309 $SOX6^{+}AGTR1^{+}$ subtype of the substantia nigra. By leveraging single-cell transcriptomic and
310 epigenomic data to identify and overcome a key bottleneck (insufficient ventral patterning), we
311 developed an optimized protocol that achieves accelerated, reproducible production of these
312 therapeutically critical neurons across multiple cell lines.

313 Our longitudinal multi-omic atlas of *in vitro* mDA differentiation revealed stage-specific
314 chromatin accessibility and transcriptional shifts mirroring human midbrain development. Late
315 maturation (DIV42-56) coincided with elevated TH, DDC and KCNJ6 expression, and PBX1
316 regulon activity, a known regulator of mDA survival and functional maturation. Crucially,
317 transcriptome-informed ventralization enhancement (via strategic agonism of SHH signaling)
318 boosted EN1 expression, a key regulator of midbrain identity and a known predictor of transplant
319 success, yielding ventral SNpc neurons with enhanced molecular fidelity to human native
320 counterparts. Furthermore, we were able to identify regulatory networks involving mDA neuron
321 maturation enriched in the new protocol, such as TCF12 and TCF4. Specifically, TCF12 plays a
322 dual role in the early commitment of neural progenitors and the late specification of neuronal
323 subsets, with its loss leading to altered expression of TH and AHD2 (ALDH1A1), the latter being
324 crucial for the specification and maintenance of dopaminergic neurons, particularly in the A9
325 dopaminergic neuronal group[43]. TCF4, on the other hand, is likely involved in the fine-tuning of
326 neuron identity during the onset of mDA neuron subtype specification. In a mouse study, deletion
327 of TCF4 led to delayed maturation of dopaminergic neurons, primarily through transient

328 suppression of *Ahd2* and *Nxph3*[44]. A key advancement in the optimized protocol was the precise
329 modulation of midbrain floor plate (mFP) patterning by the biphasic WNT activation (CHIR boost)
330 and SSH agonism, combined with optimized coating conditions using LN511, which generated a
331 substantial population of mFP progenitors by day 11 and enhanced the expression of critical mDA
332 markers such as FOXA2, LMX1A, and OTX2. This improved patterning was accompanied by a
333 reduced contamination by non-neural cells (a common issue in region-specific neuron generation).
334 Furthermore, the specification to committed progenitors allowed us to omit the MEK/ERK
335 inhibitor PD0325901 and the FGF receptor inhibitor SU5402, previously used to suppress
336 expansion of $SOX2^+$ progenitors[45], therefore harnessing the intrinsic neuronal differentiation
337 mechanisms without any further external influence for producing mature mDA neurons
338 downstream. The addition of synthetic LXR ligands (GW3965) further accelerated neurogenesis,
339 leading to the earlier appearance of TH^+ neurons at day 16 compared to day 21 in prior protocols.
340 This earlier establishment enables functional assays (such as calcium imaging and neurotransmitter
341 release analysis) to be performed sooner, thereby expediting detection of phenotypic differences.
342 As an example, live-cell calcium recordings confirmed that the derived mDA neurons displayed
343 spontaneous electrical activity and robust depolarization responses, consistent with functionally
344 active neurons[46]. Finally, as a proof-of concept validation for therapeutic approaches, we
345 determined that the new protocol could produce mFP progenitors and NbM able to differentiate *in*
346 *vivo* to mDA neurons resembling the A9-subtype nigral neurons in majority, based on expression
347 of proteins such as SOX6, AGTR1, GIRK2 or ALDH1A, reported to be more susceptible to PD in
348 human and animal models[19,40,47,48]. Furthermore, evaluation of the transcriptomic profile of the
349 DA neurons in the graft showed high resemblance to the human counterpart, specially the SOX6-
350 AGTR1A mDA neurons. Moreover, the lack of proliferative progenitor cells at our endpoint is

351 promising for putative long-term transplants. This proposes that these cells could have great
352 potential for directed replacement therapies. While these advancements could be of great interest
353 to the PD research community, we have to address several caveats. First, we only have established
354 the replicability of this protocol on benchmarked hIPSC. But further in-depth characterization with
355 more well known mutations found in familial PD, such as GBA^{E326K}, $LRRK2^{G2019S}$ or the triplicate
356 SNCA, could elucidate the validity of this enhanced protocol. Furthermore, the use of human
357 embryonic stem cells has not been proposed for this study. More *in vitro* studies would be of
358 interest to validate further this new protocol. Another caveat to address is the pre-clinical aspect,
359 as the *in vivo* experiments were only used as proof-of-concept for therapeutical approaches.
360 Further grafting experiments with PD-mutated hiPSC cell lines and a more in-depth preclinical
361 analysis could detail if this protocol would be a much better advantage that the current available
362 protocols, which so far have shown to be safe in clinical trials[12,13]. Moreover, the longevity of the
363 grafting experiment was decided upon stable motor recovery, so we do not have an overview of
364 the long-term effects of these grafts, whether other contaminants can appear or tumorigenesis in
365 aged individuals. In conclusion, we have developed a multiomic-guided differentiation protocol
366 which addresses the major challenge of subtype specificity in generating human mDA neurons
367 relevant to PD. By leveraging single-cell genomics and epigenomics to identify and correct
368 insufficient ventral patterning, we achieved robust enrichment of the PD-vulnerable $SOX6^{+}$ A9
369 subtype, which yielded neurons with enhanced molecular fidelity, functional maturity, and the
370 capacity to restore motor function upon transplantation. Our work provides a robust platform for
371 generating more accurate human cellular models of PD and establishes a foundational strategy for
372 developing next-generation, subtype-enriched cell replacement therapies.
373

374 # Methods

375 ## Cell lines

376 KOLF 2.1J cell line was obtained through Jackson Laboratories[49]. All experiments performed with
377 these cells followed the Ethical Permit 2021-06727-01 issued by the Swedish Ethical Review
378 Authority. Karyotipe by G-banding, genotypes by polymerase chain reaction (PCR), pluripotency
379 by pluripotency marker microscopy analysis, and mycoplasma infection by PCR were routinely
380 assessed before starting new experiments. Applied iPSC lines were adapted to grow on Laminin
381 521 (Biolamina, LN521)-coated plates and maintained in Essential 8 Flex media (E8 Flex, Thermo
382 Fisher Scientific) at 37ºC, 5% CO_2, until starting differentiation. The medium was changed daily,
383 and cells were passaged at 70% confluency using TrypLE™ Select Enzyme (Thermo Fisher
384 Scientific). When thawing and replating, 10 µM Rho-associated protein kinase (ROCK) inhibitor
385 (Y-27632) (Tocris) was added into the culture media overnight.

386 ## Midbrain dopaminergic neuron differentiation

387 Two days prior to the mDA differentiation, adapted iPSC cells were seeded at a density of 200,000
388 cells/cm^2 onto plates coated with Geltrex (Life Technologies) and Laminin-511 (Biolamina,
389 LN511), and maintained in E8 Flex media (Thermo Fisher Scientific) for expansion. The
390 differentiation protocol was initiated when the cells reach approximately 80% confluency. For the
391 first iteration, cells were cultured at day 0-4 on plates coated with Geltrex and LN511, replated at
392 day 11 on LN511, and replated again at day 16 on LN511 + poly-L-ornithine. Cultured cells were
393 kept in Neurobasal/N2/B27 (Life Technologies) medium and 2 mM L-glutamine (Invitrogen),
394 along with various supplements for the first stage of the protocol. From day 0 to day 4, the culture
395 medium was conditioned with 250 nM LDN193189 (Stemgent), 10 µM SB431542, 1 µM

396 Purmorphamine, and 0.7 μM CHIR99021 from day 0 to day 3, with the inclusion of 10 μM Y27632
397 (Tocris Bioscience) during the first 24 hours. Boosted CHIR99021 (7.5μM) was administered from
398 day 4 to day 11, while LDN193189, SB431542 and Purmorphamine were withdrawn after day 7
399 and N2 supplement after day 10. From day 10 to day 28, cells were switched to the differentiation
400 media: Neurobasal/B27/L-Glu media supplemented with 20 ng/mL brain-derived neurotrophic
401 factor (BDNF, R&D Systems), 20 ng/ml glial cell line-derived 455 neurotrophic factor (GDNF,
402 R&D Systems), 1 ng/ml transforming growth factor type b3 (TGFb3, R&D Systems), 0.2 mM
403 ascorbic acid (Sigma), 0.2 mM Dibutyryl cAMP (Sigma). On day 11 and 16, the resulting cells
404 were dissociated into single cells, replated at a density of 800,000 cells/cm2 onto plates coated
405 with LN511 and polyornithine/LN511 respectively, and treated with 10 μM Y27632 for 48 hours.
406 To promote neurogenesis, the mDA differentiation media was supplemented with 10 μM DAPT
407 (Sigma) from day 16. The small molecule GW3965 (10 μM) was administered from day 12 to day
408 15, followed by a cocktail of small molecules consisting of 1 μM PD0325901 and 5 μM SU5402
409 (Sigma), from day 16 to day 21. For the second iteration, conditions were similar except for the
410 following: Culture plates were coated with Geltrex + LN511 until day 14, and poly-L-ornithine +
411 LN511 from day 14 onwards. Cells were only replated once at day 14 and boosted CHIR99021
412 (7.5μM) was administered from day 4 to day 14. Day 0-7 Purmorphamine concentration was
413 increased to 2 μM, and both PD0325901 and SU5402 were removed from the protocol.

414 Immunocytochemistry

415 Progenitor cells at different stages were seeded in a 96-well plate at a density of about 300,000
416 cells per well for the intended cell culture purposes. Harvested cells were fixed by 4%
417 paraformaldehyde (PFA) for 20 min at 4ºC, and then pre-incubated by 5% donkey serum in
418 phosphate buffered saline containing 0.3% Triton X-100 (PBST) for 1 hour. The samples were

419 incubated with primary antibodies at 4ºC overnight. The samples were incubated with either
420 Alexa488 or Alexa555 or Alexa647-conjugated secondary antibodies (Thermo Fisher Scientific)
421 for 30 min and then were incubated with 4',6-diamidino-2-phenylindole (DAPI, Sigma, D9542)
422 for 15 min. The primary antibodies were used as follows: CORIN (rat, 1:1,000, R&D Systems,
423 MAB2209), DCX (goat, 1:500, SantaCruz, sc-8066), EN1 (mouse 1:50, DSHB, 4G11), FOXA2
424 (goat, 1:500, R&D Systems, AF2400), GIRK2 (rabbit, 1:400, Alomone, APC006), LMX1A
425 (rabbit, 1:4,000, Millipore, AB10533), LMO3 (goat, 1:200, SantaCruz, sc-82647), MAP2 (mouse
426 1:1,000, Sigma, M4403), NURR1 (rabbit, 1:500, SantaCruz, sc-990), OTX2 (goat, 1:1,000, R&D
427 Systems, AF1979), TH (rabbit, 1:1,000, Millipore, AB152), TH (mouse, 1:500, ImmunoStar,
428 22941), and TH (sheep, 1:500, Novus Biological, NB300). Mounted cells were further visualized
429 using confocal microscopy (Zeiss LSM880 or 980) with a 10X or 20X lens. Quantification of
430 images was performed using QuPath[58] by generating a "single measurement classifier" or
431 "composite classifier" for positive or double-positive cells respectively.

432 Calcium imaging

433 The 35mm culture dishes were loaded with the $Ca2^{+}$-sensitive fluorescence indicator Fluo-4/AM
434 (5 µM; Invitrogen, USA) and Pluronic F-127 (0.625%, ThermoFisher Scientific) and incubated
435 for 30 min at 37ºC in Krebs-Ringer buffer, containing: NaCl (119 mM), KCl (2.5 mM), NaH2PO4
436 (1 mM), CaCl2 (2.5 mM), MgCl2 (1.3 mM), HEPES (20 mM) and D-Glucose (11 mM), with pH
437 adjusted to 7.4. Dishes were washed twice and $Ca2^{+}$-measurements were performed in Krebs-
438 Ringer buffer at 37ºC on a confocal spinning disk microscope (Nikon CrEST X-Light V3)
439 equipped with a 20x Air lens. Excitation was assessed at 480 nm at a sampling frequency of 1 Hz.
440 The equipment was controlled, and data was acquired using the Nikon NIS-Elements software.

441 FIJI, MATLAB (R2020b, MathWorks, USA) and FluoroSNNAP (https://github.com/tapan-
442 patel/FluoroSNNAP) were used to process and analyze the data.

443 Single cell sequencing libraries preparation and data processing

444 A total of 5 samples (DIV 11, 16, 28, 42, and 56) from KOLF2.1J mDA neuron cultures generated
445 with the optimized protocol were collected for time-course multiomic analysis (ATAC + Gene
446 Expression); in addition, 3 samples (DIV 14, 16, and 28) from the new protocol trial of KOLF2.1J
447 mDA differentiation were collected for single-nuclei transcriptomic profiling. All 8 samples were
448 prepared with commercially available protocols (10x Genomics) and the resulting Fastqc files
449 were processed through the 10X Genomics Cellranger pipeline[50], while 5 multiomic libraries were
450 aligned to the hg38 reference genome using Cellranger-arc default settings[51]. After quality-control
451 filtering, we obtained high-quality 23,534 cells (Fig. 3a, Extended Fig. 5a) and then we integrated
452 transcriptomic profiles from both conditions using Harmony[52], enabling a comparative analysis by
453 isolating the mDA neuronal lineage from different datasets (Extended Fig. 5b). Resulting
454 transcriptomic and epigenomic data from individual samples were combined by projects using the
455 'aggr' function in Cellranger or Cellranger-arc, without normalization applied. Objects produced
456 with Seurat V4.4.0 were employed to accommodate aggregated count matrix[53], with the fragment
457 files specifically integrated into Signac v1.9.0 chromatin assays[54]. Quality metrics for gene
458 expression and peak accessibility, such as mitochondrial ratio, cell cycle score, nucleosome signal,
459 and Transcription Start Site (TSS) enrichment, were thereafter calculated to determine quality
460 control cutoffs. Depending on sample quality, following thresholds were applied for downstream
461 analysis: a maximum of 300,000 and a minimum of 1,000 ATAC counts per cell, a maximum of
462 25,000 and a minimum of 600 RNA counts, at least 250 genes detected, a maximum nucleosomal
463 signal of 1.5, TSS enrichment of at least 2, and mitochondrial content no higher than 15%. These

464 cutoffs reduced the cell count by less than 20% without significantly affecting clustering
465 downstream. Peak calling was performed with MACS2 v2.2.9 using the aggregated fragments
466 file[55], and those mapped to non-standard chromosomes or overlapping hg38 genome blacklist
467 regions were excluded. DoubletFinder v2.0 was applied to identify doublets in gene expression
468 data by flagging instances with a high resemblance to artificial neighbors in gene expression
469 space[56]. AMULET (ATAC-seq MULtiplet Estimation Tool) was additionally utilized to detect
470 potential multiplets in snATAC-seq datasets by enumerating regions with greater than two
471 uniquely aligned reads across the genome[57]. All predicted doublets were removed from the
472 downstream analysis. After merging high-quality cells from the relevant projects, count data was
473 normalized using SCTransform V2 function with regression out of mitochondrial or cell-cycle
474 genes. The top 2,000 most variable genes were initially identified through variance-stabilizing
475 transformation, along with 95% of the most frequent peaks in multiomic datasets, as a part of the
476 significant feature selection process. Subsequent dimensionality reduction was performed using
477 the principal components analysis (PCA), in which the top 30/40 principal components were
478 chosen from the elbow plot to represent most of the diversity of datasets. For single-nuclei ATAC-
479 seq datasets, singular value decomposition (SVD) was performed with the top 25 of the Latent
480 Semantic Indexing (LSI) layers to reduce the dimensions of the most common peaks. Notably, the
481 first LSI component was excluded from downstream analysis as it often captures sequencing depth
482 (technical variation) rather than biological variation. Putative clusters were identified by a spiking
483 neural network (SNN) modularity optimization-based clustering algorithm, and the optimal
484 resolution of clustering was determined by iterating over different resolution for the most stable
485 segregation. Finally, the calculated nearest-neighbor graph was projected to two dimensions to
486 visualize the distribution of instances by representing with uniform manifold approximation and

projection (UMAP). To accurately identify and interpret cell states *in vitro*, we leveraged combined manual cell type annotation using cluster-specific marker genes[58,59], with automatic annotation methods using reference datasets that characterize first-trimester developing human brain[14,60]. Initially, a label transfer was performed with Celltypist[61] and 'TransferData' function in Seurat to automatically classify selected dataset with cell type labels from the reference dataset, based on the similarity of query gene expression profiles to selected anchors in the reference dataset. The identity of each cell will be determined by the highest prediction score, and for stringent comparison, cells with a prediction score below 0.51 will be excluded from downstream analysis. After data trimming and reclustering, a manual cluster-based annotation was performed, using both expression levels of canonical midbrain development marker genes and top 50 upregulated genes identified by the 'FindAllMarkers' function. These upregulated genes were used for pathway and biological process enrichment analysis via Metascape[59], and the top-ranked terms were used to refine annotations and determine cell subtypes or states. To reconstruct the putative trajectory of mDA differentiation *in vitro*, we employed Monocle3[62] and scVelo[63,64] to infer pseudo- and latent-temporal ordering of cells using default parameters[65]. The root of trajectory for pseudotime analysis was set as ProgFP or NbM at the earliest timepoint. at the earliest timepoint. To represent quantitatively the cell identity similarities with other published human datasets, we performed Metaneighbor[66]. SCENIC[67] was used to visualize regulatory networks and regulon in our developing cells using AUCell scores.

Animals

Healthy female adult (8-12 weeks old) C57BL/6N-Rag2Tm1-IL2rgTm1-SirpaNOD/Rj (B6RGS) mice were purchased from Janvier Labs. Groups of 5 animals were housed in ventilated cages under a 12 h light-dark cycle with food (standard chow) and water ad libitum at 20-22ºC and 50-

510 55% humidity. Tunnels, cardboard nesting material and chewing wooden bars were used for
511 environmental enrichment. All surgical procedures followed aseptic techniques, and for post-
512 surgical recovery, additional Dietgel Recovery (Scanbur) and food pellets were supplied on the
513 cage floor. Animals were randomly allocated to sham control group (n=6), hemiPD group (n=8)
514 and a grafted group (n=8), ensuring all cages had individuals of each group. Briefly, the exclusion
515 criteria were animals which lost over 10% of their pre-surgical weight after 4 weeks were discarded
516 from the experiment. No animals were discarded based on this criterion. The experimental unit
517 were the individual mouse. All procedures were approved by the local ethical committee
518 (Stockholms djurforsoksetiska namnd) and followed the EU directive 2010/63/EU under the
519 ethical permit of Ernesto Arenas Cases and Sandra Gellhaar (12265-2021; 2-3143/23; 2-3595/24;
520 2-3147/24).

521 ## 6-OHDA mouse model

522 B6RGS mice were acclimatized 2-3 weeks before surgery to the animal facility and the researcher
523 and to a recovery diet 2-3 days before surgery. For 6-hydroxidopamine (6-OHDA, Sigma) lesions
524 (or sham), animals were injected analgesia (Carprofen 5mg/kg + Buprenorphine 0.1 mg/kg) 15
525 minutes before isoflurane induction (4%). Animals were anesthetized at 1-2% isoflurane in air and
526 injected intracranially (ICR) using a Neurostar Robotic Stereotaxic Frame equipped with a 10 µl
527 syringe with 33G blunt needle (Hamilton) 1 µl of 6-OHDA 3.6 µg/ µl in sterile ascorbic acid/saline
528 0.02% (or sham vehicle) in the right medial forebrain bundle [respect to Bregma: AP -1.2 mm,
529 ML +1.2 mm, DV -5.0 mm] at a flow of 0.2 µl/min, raising the needle 500 µm for 3 minutes. The
530 needle was left inside for 2 additional minutes. Animals were placed in warm cages (28 °C)
531 between 3 days and 1 week for better recovery and rehydrated with Ringer Acetate solution if
532 weight decreased by 5%. 6-OHDA hemiPD was confirmed by amphetamine-induced rotation.

533 Briefly, 4 weeks after the lesion, mice were injected intraperitoneally with d-amphetamine in saline
534 (5 mg/kg). 10 minutes after injection, the mice were recorded for 30 minutes inside a 15 cm wide
535 cylinder in dark conditions, and the number of net ipsilateral rotations (ipsilateral-contralateral)
536 per minute in active movement was calculated using Ethovision XT 17 (Noldus). Mice are
537 considered hemiPD if the values are >6 rotations/minutes.

538 hiPSC transplant

539 HemiPD-confirmed mice of healthy weight (22-30 g) were randomly selected for ICR hPSC
540 injection in the right striatum [respect to Bregma: AP +0.5 mm, ML +2 mm, DV -3.5 mm].
541 Analgesia and anesthesia were followed as above. For ICR transplants, 2 µl of 100.000 ± 20.000
542 hPSCs/µl in Neurobasal medium was injected to achieve 200.000 hPSC transplants. For cell
543 injection, the Neurostar robot was coupled with a 10 µl syringe with a 27G sharp syringe. A flow
544 of 0.5 µl/min was used, and the needle remained for 5 additional minutes. Cells were injected
545 between 1-6 hours after thawing.

546 Amphetamines-induced rotations

547 Mice were analysed a week before surgery, 4 weeks after surgery and hereon once a month until
548 endpoint was reached. Mice were injected intraperitoneally with 5 mg/kg d-amphetamine in saline
549 (0.5 mg/ml). 10 minutes after injection, the mice were recorded for 30 minutes inside a 15 cm wide
550 cylinder in dark conditions, and the number of net ipsilateral rotations (ipsilateral-contralateral)
551 per minute in active movement (body center traveling at >0.5 cm/s) was calculated using
552 Ethovision XT 17.

553 Pole test

554 Mice were analysed a week before surgery, 4 weeks after surgery and at the endpoint. Animals
555 were placed facing upwards on the top of a 20 cm vertical smooth wooden pole with 2 mm grooves
556 notched every 0.5 cm down. The time to climb down was counted starting when the mouse rotates
557 facing downwards. Animals were trained for a minimum of 2 days before testing day, by
558 performing 3 mock tests each day. If the animals took more than 1 minute to descend, the trial was
559 interrupted and repeated at least 10 minutes later. A total of 3 trials were used to calculate a mean
560 value per animal.

561 Cylinder wall

562 Mice were analysed a week before surgery, 4 weeks after surgery and at the endpoint. Animals
563 were placed inside a 12 cm diameter glass cylinder in red light and quiet conditions. For 15
564 minutes, animals were recorded and videos were analysed at x8 speed. Unilateral full paw contacts
565 were counted. Paw dragging or paws not fully placed on glass were disregarded.

566 Tape removal test

567 Mice were analysed a week before surgery, 4 weeks after surgery and at the endpoint. A green
568 adhesive paper tape 2x3 mm was placed on each paw individually. The time it took to remove the
569 tape was counted, beginning when the mouse recognized it and ending when it was removed from
570 the paw. Net contralateral time (ipsilateral-contralateral) was calculated to better visualize and
571 eliminate external factors. Before testing day, animals were trained for a minimum of 2 days,
572 performing 3 mock trials for each paw. If the animal took more than a minute, the trial was
573 interrupted and repeated at least 15 minutes later. A total of 3 trials were used to calculate a mean
574 value.

575 Immunohistochemistry

576 5 animals were terminally anesthethized with IP pentobarbital 100 mg/kg dose and then perfused
577 intracardially with 50 ml of pH-corrected PBS1X followed by 50 ml of formaldehyde 4 % (FA)
578 both at a flow of 9 ml/min. Brains were extracted and placed in FA overnight, after which they
579 were placed in sucrose 30%/PBS solution at 4ºC for 2-3 days. After, brains were kept at -80ºC.

580 Brains were equilibrated at -20ºC for 4 hours before slicing. Samples were sliced in series at -20ºC
581 at 20 µm using a Cryostar NX70 (Epredia) and stored for further use at -20ºC in cryoprotective
582 solution (Ethylenglycol 30%, Glycerol 30%, PBS1X). A total of 5 brains were processed.

583 Samples were stained using free-floating protocols. Briefly, 2 series of 5 minutes of PBS1X,
584 followed by 15 minutes of PBS-TritonX 0.3% (PBTx0.3%), then 2 series of 5 minutes of
585 PBTx0.1%. Samples were blocked for 1 hour with blocking solution (BSA 1 mg/ml, Normal
586 Donkey Serum 5%, PBTx0.1%) and incubated o/n at 4ºC with primary antibody diluted in
587 blocking solution. Primary antibody was washed with 2 series of 5 minutes PBTx0.1% and
588 samples were incubated for 4 hours with secondary antibody diluted in blocking solution and
589 washed twice for 5 minutes with PBS1X. Tissue was counterstained with DAPI 1:1000 in PBS1X
590 for 10 minutes and then washed twice for 5 minutes with PBS1X. Stained samples were mounted
591 on Superfrost Plus slides (Epredia) and Fluoromount-G mounting medium (00-4958-02, Thermo
592 Scientific). Primary antibodies used were STEM121 (mouse, 1:1000, Tanaka Bio, Y40410), TH
593 (rabbit, 1:1000, Pelfreeze, P4010; sheep, 1:250 Pelfreeze, P60101), HuNu (mouse, 1:1000,
594 Millipore, MAB1281), SOX6 (rabbit, 1:200, Abcam, ab30455), ALDH1A (rabbit, 1:500, Atlas
595 Antibodies, HPA002123), AGTR1 (rabbit, 1:100, Invitrogen, MA5-56587), Calbindin-1 (rabbit,
596 1:200, Invitrogen, PA5-85669), LMX1 (rabbit, 1:100, Millipore, AB10533), FOXA2 (rabbit,
597 1:200, Cell Signaling, 3143S), NR4A2 (rabbit, 1:500, Invitrogen, PA5-78097), PITX3 (rabbit,
598 1:100, Invitrogen, 35-2850), DAT (rat, 1:250, Abcam, ab5990), GIRK2 (rabbit, 1:500, Alomone

599 Labs, APC006), VMAT2 (rabbit, 1:500, Sigma-Aldrich, AB1598P), DDC (rabbit, 1:500, Novus
600 Biological, NB300-252), Ki67 (rabbit, 1:200, Abcam, ab15580), SOX2 (goat, 1:200, R&D
601 Systems, AF2018). All fluorescent secondary antibodies were from Invitrogen, donkey as host and
602 conjugated with Alexa Fluor 488, 555 or 647. The HRP-conjugated antibody was a goat anti-
603 mouse from Invitrogen. Images were captured either with laser confocal microscope Zeiss
604 LSM800 equipped with x20 lens or high-throughput epifluorescence microscope Zeiss Axioscan
605 equipped with x10 lens. Quantification of images was performed using QuPath[58] by generating a
606 'single measurement classifier' or 'composite classifier' for positive or double-positive cells
607 respectively.

608 Single-nuclei RNA sequencing of grafts and data analysis

609 3 animals were euthanized by cervical dislocation and brains were promptly extracted and flash-
610 frozen in dry-ice chilled isopentane for 10-15 seconds. Flash-frozen brains were stored at -80 °C
611 until graft extraction. A total of 3 brains with grafts were processed independently.

612 On the day of nuclei extraction, brains were equilibrated to -20ºC and sliced in -20ºC cooled
613 chambers. Brain matrices were used to stabilize the brain, and 1 mm-thick slices for a total of 4 (2
614 before the center of graft implant location and 2 after) were cut using chilled razors. 1 mm punches
615 of each slice of the dorsal striatum were extracted using 1 mm biopsy-punchers (49101, PFM
616 Medical).

617 Nuclei were dissociated using the 10X Genomics Chromium nuclei extraction kit (1000494) using
618 manufacturer's instructions. For better extraction, eppendorfs used for tissue dissociation were
619 coated with BSA 1% o/n at 4ºC, briefly washed with filtered sterile PBS1X before procedure. 10X
620 Genomics Chromium GEM-X Single Cell 3' Kit v4 (1000691), Chromium GEM-X Single Cell 3'

621 Chip Kit v4 (1000690) and Dual Index Kit TT Set A (1000215) were used for GEM generation
622 using manufacturer's instructions with the 10X Controller iX. GEMs were sent for RNA-
623 sequencing to National Genomics Infrastructure of SciLifeLab (Karolinska Institute, Sweden)
624 using the NovaSeq XPlus 25B. Data analysis was performed as the *in vitro* RNAseq dataset,
625 described above. To ensure analysis of human cells, we aligned the libraries to the h38 human
626 genome dataset and the m39 mouse genome dataset, and cells with a ratio of human:mouse gene
627 count >1 were processed.

628 Statistics

629 The N for the animals was calculated using G-Power 3.1[68] based on prior data from several
630 publications of hemiPD rodent model behavior[69]. Briefly, using the "ANOVA: repeated measures,
631 between factors" a priori statistical test, the effect size was calculated to be 0.82, α =0.05, β=0.95,
632 groups =3, measurements=7 and correlation among repeated measurements=0.8.

633 GraphPad Prism v10.00 was used for generating all graphs and performing statistical analyses. For
634 pairwise comparisons, Two-tailed Student's t-test was utilized. One-way ANOVA or two-way
635 ANOVA with Tukey's or Bonferroni's multiple comparison tests were used to compare multiple
636 variables. Error bars indicate the standard error of the mean (SEM). F are reported with degrees of
637 freedom for numerator and denominator ($F_{DFn,DFd}$). Significance levels were set as follows: ns =
638 not significant; * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$; **** = $p<0.0001$.

639 **Data Availability**

640 Single-cell RNA-seq and ATAC-seq raw data have been deposited at EGA (EGA accession
641 number: EGAD50000002054). Key Resource Table (including Material RRID, protocol.io links
642 for protocols used, accession numbers for sequencing datasets), microscopy data and images have

been deposited at Zenodo at https://doi.org/10.5281/zenodo.17099440 and are publicly available as of the date of publication. All reused code used for this paper has been deposited at Github and linked to Zenodo at https://doi.org/10.5281/zenodo.17099440 and is publicly available as of the date of publication. This paper analyzes existing, publicly available data, accessible at https://doi.org/10.1038/s41593-022-01061-1; https://doi.org/10.1126/science.adf1226; https://doi.org/10.1126/science.add7046. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Code Availability

No new code was generated for this paper. All data cleaning, preprocessing, analysis, and visualization was performed using Graphpad Prism v10.00.

References

1. Poewe, W. *et al.* Parkinson disease. *Nat Rev Dis Primers* **3**, 17013 (2017).
2. Brichta, L., Greengard, P. & Flajolet, M. Advances in the pharmacological treatment of Parkinson's disease: targeting neurotransmitter systems. *Trends Neurosci* **36**, 543–54 (2013).
3. Björklund, A. & Parmar, M. Dopamine Cell Therapy: From Cell Replacement to Circuitry Repair. *J Parkinsons Dis* **11**, S159 (2021).
4. Doi, D. *et al.* Isolation of human induced pluripotent stem cell-derived dopaminergic progenitors by cell sorting for successful transplantation. *Stem Cell Reports* **2**, 337–350 (2014).
5. Kim, T. W. *et al.* Biphasic Activation of WNT Signaling Facilitates the Derivation of Midbrain Dopamine Neurons from hESCs for Translational Use. *Cell Stem Cell* **28**, 343-355.e5 (2021).
6. Kriks, S. *et al.* Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* **480**, 547–551 (2011).
7. Nishimura, K. *et al.* Single-cell transcriptomics reveals correct developmental dynamics and high-quality midbrain cell types by improved hESC differentiation. *Stem Cell Reports* **18**, 337–353 (2023).

672 8. Kirkeby, A. *et al.* Preclinical quality, safety, and efficacy of a human embryonic stem cell-
673 derived product for the treatment of Parkinson's disease, STEM-PD. *Cell Stem Cell* **30**,
674 1299-1314.e9 (2023).

675 9. Piao, J. *et al.* Preclinical Efficacy and Safety of a Human Embryonic Stem Cell-Derived
676 Midbrain Dopamine Progenitor Product, MSK-DA01. *Cell Stem Cell* **28**, 217-229.e7
677 (2021).

678 10. Schweitzer, J. S. *et al.* Personalized iPSC-Derived Dopamine Progenitor Cells for
679 Parkinson's Disease. *N Engl J Med* **382**, 1926–1932 (2020).

680 11. Takahashi, J. iPS cell-based therapy for Parkinson's disease: A Kyoto trial. *Regen Ther*
681 **13**, 18–22 (2020).

682 12. Tabar, V. *et al.* Phase I trial of hES cell-derived dopaminergic neurons for Parkinson's
683 disease. *Nature 2025 641:8064* **641**, 978–983 (2025).

684 13. Sawamoto, N. *et al.* Phase I/II trial of iPS-cell-derived dopaminergic cells for Parkinson's
685 disease. *Nature 2025 641:8064* **641**, 971–977 (2025).

686 14. Braun, E. *et al.* Comprehensive cell atlas of the first-trimester developing human brain.
687 *Science* **382**, eadf1226 (2023).

688 15. La Manno, G. *et al.* Molecular Diversity of Midbrain Development in Mouse, Human, and
689 Stem Cells. *Cell* **167**, 566-580.e19 (2016).

690 16. Xu, P. *et al.* Human midbrain dopaminergic neuronal differentiation markers predict cell
691 therapy outcomes in a Parkinson's disease model. *J Clin Invest* **132**, (2022).

692 17. Tiklová, K. *et al.* Single-cell RNA sequencing reveals midbrain dopamine neuron
693 diversity emerging during mouse brain development. *Nature Communications 2019 10:1*
694 **10**, 1–12 (2019).

695 18. Greene, P. E. *et al.* Persistent dyskinesias in patients with fetal tissue transplantation for
696 Parkinson disease. *NPJ Parkinsons Dis* **7**, (2021).

697 19. Kamath, T. *et al.* Single-cell genomic profiling of human dopamine neurons identifies a
698 population that selectively degenerates in Parkinson's disease. *Nature Neuroscience 2022*
699 *25:5* **25**, 588–595 (2022).

700 20. Siletti, K. *et al.* Transcriptomic diversity of cell types across the adult human brain.
701 *Science* **382**, eadd7046 (2023).

702 21. Mannens, C. C. A. *et al.* Chromatin accessibility during human first-trimester
703 neurodevelopment. *Nature* 1–8 (2024) doi:10.1038/s41586-024-07234-1.

704 22. Nishimura, K., Ásgrímsdóttir, E. S., Yang, S. & Arenas, E. A protocol for the
705 differentiation of human embryonic stem cells into midbrain dopaminergic neurons. *STAR*
706 *Protoc* **4**, 102355 (2023).

707 23. Sif Ásgrímsdóttir, E. *et al.* Midbrain Dopaminergic Neuron Development is Regulated by
708 Two Molecularly Distinct Subtypes of Radial Glia Cells. *bioRxiv* 2024.03.08.584031
709 (2024) doi:10.1101/2024.03.08.584031.

710 24. Desai, A. R. & McConnell, S. K. Progressive restriction in fate potential by neural
711 progenitors during cerebral cortical development. *Development* **127**, 2863–2872 (2000).

712 25. Nürnberg, E., Horschitz, S., Schloss, P. & Meyer-Lindenberg, A. Basal glucocorticoid
713 receptor activation induces proliferation and inhibits neuronal differentiation of human
714 induced pluripotent stem cell-derived neuronal precursor cells. *J Steroid Biochem Mol*
715 *Biol* **182**, 119–126 (2018).

716 26. Toledo, E. M. *et al.* Srebf1 Controls Midbrain Dopaminergic Neurogenesis. *Cell Rep* **31**,
717 107601 (2020).

718 27. Kim, W. Y. NeuroD1 is an upstream regulator of NSCL1. *Biochem Biophys Res Commun*
719 **419**, 27–31 (2012).

720 28. Murdoch, J. N., Eddleston, J., Leblond-Bourget, N., Stanier, P. & Copp, A. J. Sequence
721 and expression analysis of Nhlh1: a basic helix-loop-helix gene implicated in
722 neurogenesis. *Dev Genet* **24**, 165–177 (1999).

723 29. Dennis, D. J., Han, S. & Schuurmans, C. bHLH transcription factors in neural
724 development, disease, and reprogramming. *Brain Res* **1705**, 48–65 (2019).

725 30. Tolve, M. *et al.* The transcription factor BCL11A defines distinct subsets of midbrain
726 dopaminergic neurons. *Cell Rep* **36**, 109697 (2021).

727 31. Simon, H. H., Saueressig, H., Wurst, W., Goulding, M. D. & O'Leary, D. D. M. Fate of
728 Midbrain Dopaminergic Neurons Controlled by the Engrailed Genes. *Journal of*
729 *Neuroscience* **21**, 3126–3134 (2001).

730 32. Albéri, L., Sgadò, P. & Simon, H. H. Engrailed genes are cell-autonomously required to
731 prevent apoptosis in mesencephalic dopaminergic neurons. *Development* **131**, 3229–3236
732 (2004).

733 33. Poulin, J. F. *et al.* Mapping projections of molecularly defined dopamine neuron subtypes
734 using intersectional genetic approaches. *Nature Neuroscience 2018 21:9* **21**, 1260–1271
735 (2018).

736 34. Agoston, Z., Li, N., Haslinger, A., Wizenmann, A. & Schulte, D. Genetic and physical
737 interaction of Meis2, Pax3 and Pax7 during dorsal midbrain development. *BMC Dev Biol*
738 **12**, 10 (2012).

739 35. Nakamura, H., Katahira, T., Matsunaga, E. & Sato, T. Isthmus organizer for midbrain and
740 hindbrain development. *Brain Res Rev* **49**, 120–126 (2005).

741 36. Di Bella, D. J. *et al.* Molecular logic of cellular diversification in the mouse cerebral
742 cortex. *Nature 2021 595:7868* **595**, 554–559 (2021).

743 37. Gheldof, A., Hulpiau, P., van Roy, F., De Craene, B. & Berx, G. Evolutionary functional
744 analysis and molecular regulation of the ZEB transcription factors. *Cell Mol Life Sci* **69**,
745 2527–41 (2012).

746 38. Smidt, M. P. *et al.* Early developmental failure of substantia nigra dopamine neurons in
747 mice lacking the homeodomain gene Pitx3. *Development* **131**, 1145–1155 (2004).

748 39. Ziegler, A. B. *et al.* Cell-Autonomous Control of Neuronal Dendrite Expansion via the
749 Fatty Acid Synthesis Regulator SREBP. *Cell Rep* **21**, 3346–3353 (2017).

750 40. Yaghmaeian Salmani, B. *et al.* Transcriptomic atlas of midbrain dopamine neurons
751 uncovers differential vulnerability in a Parkinsonism lesion model. *Elife* **12**, (2024).

752 41. Storm, P. *et al.* Lineage tracing of stem cell-derived dopamine grafts in a Parkinson's
753 model reveals shared origin of all graft-derived cells. *Science Advances* **10**, 3057 (2024).

754 42. Tang, L. *et al.* A primate nigrostriatal atlas of neuronal vulnerability and resilience in a
755 model of Parkinson's disease. *Nature Communications 2023 14:1* **14**, 1–13 (2023).

756 43. Chung, S. *et al.* The homeodomain transcription factor Pitx3 facilitates differentiation of
757 mouse embryonic stem cells into AHD2-expressing dopaminergic neurons. *Molecular and*
758 *Cellular Neuroscience* **28**, 241–252 (2005).

759 44. Mesman, S., Wever, I. & Smidt, M. P. Tcf4 Is Involved in Subset Specification of
760 Mesodiencephalic Dopaminergic Neurons. *Biomedicines 2021, Vol. 9, Page 317* **9**, 317
761 (2021).

762 45. Koch, P., Opitz, T., Steinbeck, J. A., Ladewig, J. & Brüstle, O. A rosette-type, self-
763 renewing human ES cell-derived neural stem cell with potential for in vitro instruction and
764 synaptic integration. *Proc Natl Acad Sci U S A* **106**, 3225–30 (2009).

765 46. Pissadaki, E. K. & Bolam, J. P. The energy cost of action potential propagation in
766 dopamine neurons: clues to susceptibility in Parkinson's disease. *Front Comput Neurosci*
767 **7**, (2013).

768 47. del Rey, N. L. G. *et al.* Calbindin and Girk2/Aldh1a1 define resilient vs vulnerable
769 dopaminergic neurons in a primate Parkinson's disease model. *npj Parkinson's Disease*
770 *2024 10:1* **10**, 1–15 (2024).

771 48. Nie, S. *et al.* Sox6 and ALDH1A1 Truncation by Asparagine Endopeptidase Defines
772 Selective Neuronal Vulnerability in Parkinson's Disease. *Advanced Science* **12**, 2409477
773 (2025).

774 49. Pantazis, C. B. *et al.* A reference human induced pluripotent stem cell line for large-scale
775 collaborative studies. *Cell Stem Cell* **29**, 1685-1702.e22 (2022).

776 50. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells.
777 *Nature Communications 2017 8:1* **8**, 1–12 (2017).

51. Satpathy, A. T. *et al.* Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nature Biotechnology 2019 37:8* **37**, 925–936 (2019).

52. Korsunsky, I. *et al.* Fast, sensitive, and accurate integration of single cell data with Harmony. *Nat Methods* **16**, 1289 (2019).

53. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

54. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nature Methods 2021 18:11* **18**, 1333–1341 (2021).

55. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

56. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337.e4 (2019).

57. Thibodeau, A. *et al.* AMULET: a novel read count-based method for effective multiplet detection from single nucleus ATAC-seq data. *Genome Biol* **22**, 252 (2021).

58. Clarke, Z. A. *et al.* Tutorial: guidelines for annotating single-cell transcriptomic maps using automated and manual methods. *Nature Protocols 2021 16:6* **16**, 2749–2764 (2021).

59. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).

60. Xu, C. *et al.* Automatic cell-type harmonization and integration across Human Cell Atlas datasets. *Cell* **186**, 5876-5891.e20 (2023).

61. Domínguez Conde, C. *et al.* Cross-tissue immune cell analysis reveals tissue-specific features in humans. *Science* **376**, eabl5197 (2022).

62. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology 2014 32:4* **32**, 381–386 (2014).

63. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology 2020 38:12* **38**, 1408–1414 (2020).

64. Gayoso, A. *et al.* Deep generative modeling of transcriptional dynamics for RNA velocity analysis in single cells. *Nature Methods 2023 21:1* **21**, 50–59 (2023).

65. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature 2019 566:7745* **566**, 496–502 (2019).

812 66. Crow, M., Paul, A., Ballouz, S., Huang, Z. J. & Gillis, J. Characterizing the replicability
813 of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nature*
814 *Communications 2018 9:1* **9**, 1–12 (2018).

815 67. Aibar, S. *et al.* SCENIC: Single-cell regulatory network inference and clustering. *Nat*
816 *Methods* **14**, 1083–1086 (2017).

817 68. Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. G*Power 3: a flexible statistical power
818 analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*
819 **39**, 175–191 (2007).

820 69. Björklund, A. & Dunnett, S. B. The Amphetamine Induced Rotation Test: A Re-
821 Assessment of Its Use as a Tool to Monitor Motor Impairment and Functional Recovery in
822 Rodent Models of Parkinson's Disease. *J Parkinsons Dis* **9**, 17–29 (2019).

823

824 ## Acknowledgements

825 This work was mainly supported by the European Union's Horizon 2020 (H2020) Research and
826 Innovation Programme (Neurostemcell-Reconstruct 874758/EA) and the European Research
827 Council (ERC) PreciseCellPD grant (Grant No. 884608/EA). Other contributors: Knut and Alice
828 Wallenberg Foundation Wallenberg Scholar 2018.0232/EA, Vetenskaprådet VR2020-01426/EA,
829 Chan Zuckerberg Initiative and the Silicon Valley Community Foundation (2018-191929)/EA,
830 H2020 MSCA-ITN ASCTN/EA, H2020 Grant No. 899687 - HS-SEQ, Aligning Science Across
831 Parkinson's (ASAP-20505)/EA, Hjärnfonden project FO2024-0195/EA. JCK and PU were
832 supported by the German Science Foundation (516641042) and the Wenner-Gren foundation
833 (UPD2022-0159) respectively.

834 The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded
835 by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish
836 Research Council, and NAISS/Uppsala Multidisciplinary Center for Advanced Computational
837 Science for assistance with massively parallel sequencing and access to the UPPMAX
838 computational infrastructure. We are grateful for the assistance of the Molecular Neurobiology

839 unit within the Department of Medical Biochemistry and Biophysics at the Karolinska Institute.
840 We would like to posthumously thank Prof. Ernest Arenas, without whom this project would not
841 have been possible.

842

843 Ethics declarations

844 Competing interests

845 None of the authors declare competing interests.

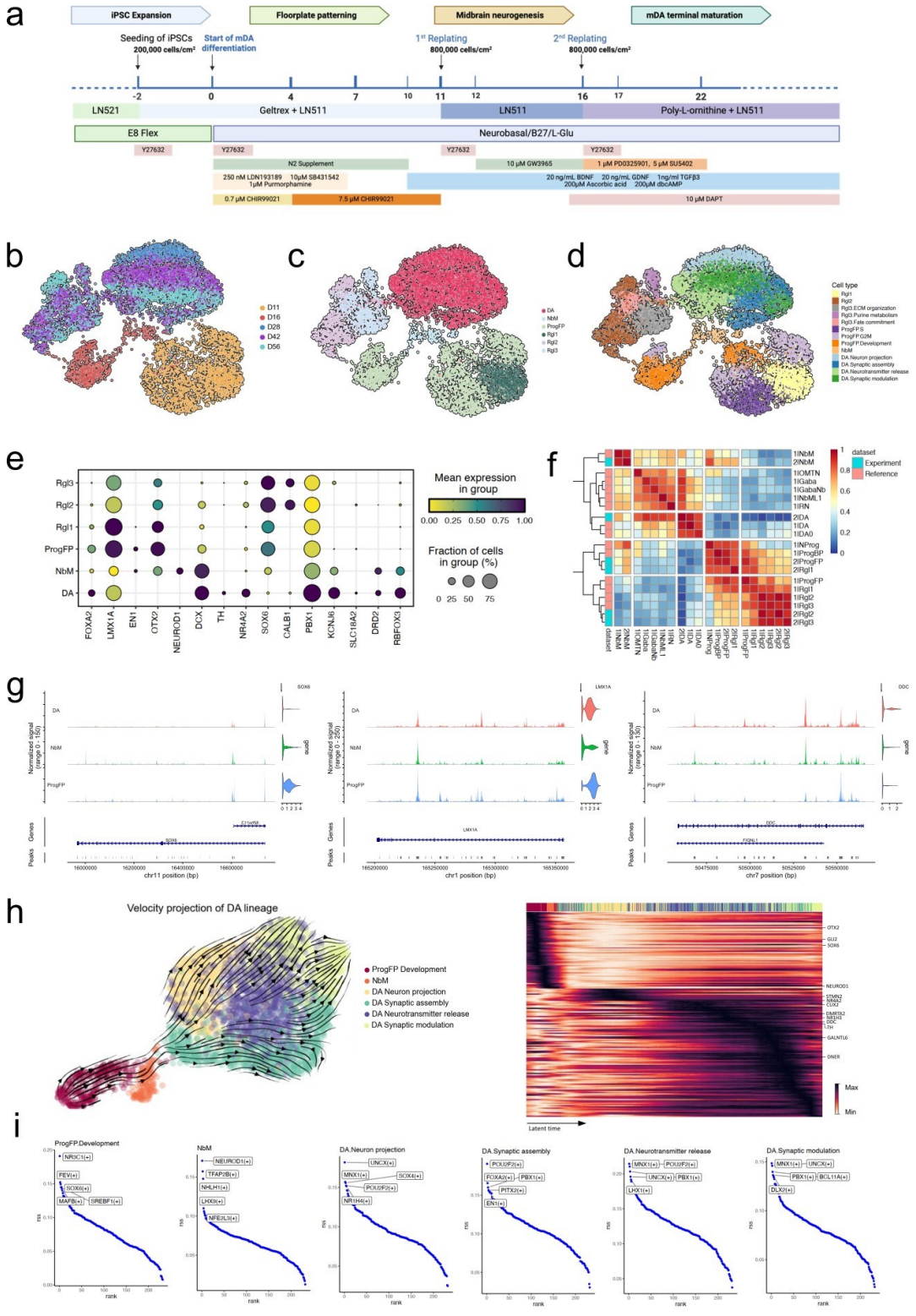
846 Contributions

847 R.G.S., G.L. and E.A. conceptualized the experiments. R.G.S. and G.L. redacted the manuscript.
848 R.G.S. performed *in vivo* experiments; G.L., R.K., C.S., C.T. and A.X. contributed to the *in vitro*
849 protocol development. J.C.K. contributed to the calcium imaging experiment. G.L. and A.X.
850 contributed to the *in vitro* transcriptional analysis; R.G.S. and C.A. performed the graft nuclei
851 isolation, and R.G.S. and G.L. the transcriptional analysis. C.S., G.C.B. and O.D. contributed to
852 supervision and reviewed the manuscript.

853

854 **Fig. 1: Single nuclei multi-omic sequencing revealed the diversity of cellular composition**

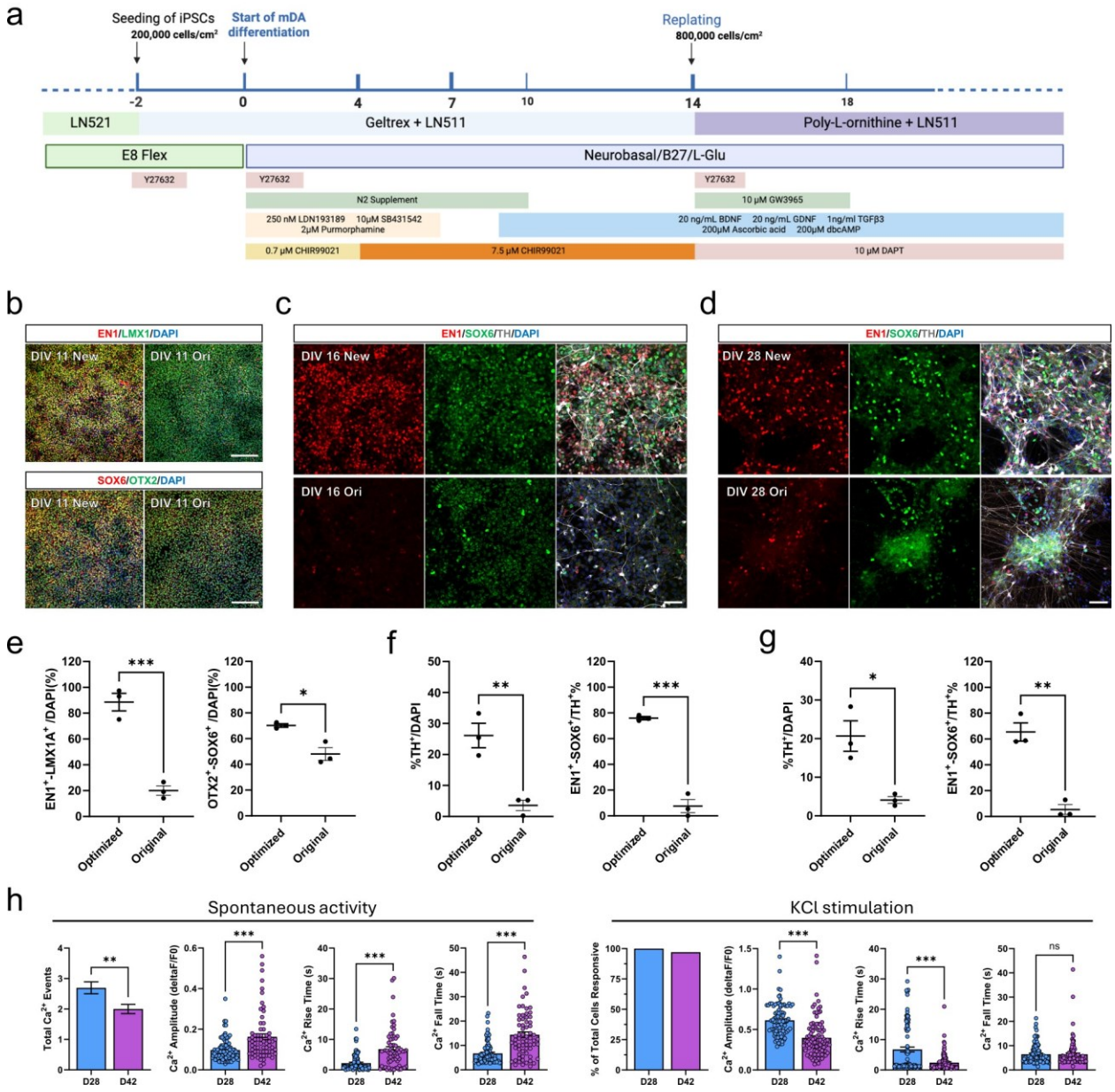
855 **during mDA neuron differentiation *in vitro*.**



856

857 **a,** Schematic diagram of the initial differentiation protocol used on the iPSC. **b,** UMAP of time-course
858 samples profiled with simultaneous snRNA-seq and snATAC-seq showing clustering of cells based on both
859 transcriptomic and chromatin accessibility profiles. **c,** Annotated bimodal UMAP clusterings based on
860 major cell types. **d,** Subsets of cell types were manually annotated with top-ranked marker genes and their
861 enriched GO terms. **e,** Dot plot visualizing the expression levels and the percentage of cells expressing
862 curated key genes across cell types during mDA neuron differentiation. **f,** Heatmap of Area Under the
863 Receiver Operating Characteristic Curve (AUROC) scores comparing annotated cell types derived from in
864 vitro differentiation in this study with those from a publicly available first-trimester human midbrain dataset.
865 **g,** Normalized chromatin accessibility tracks at genomic loci of genes of interest (SOX6, LMX1A and
866 DDC), along with enriched accession peaks and links. **h,** UMAP clustering showing RNA velocity
867 projection in mDA-related cell types (left), and a heatmap displaying gene expression dynamics of cluster
868 driver genes ordered by their inferred latent time (right). **i,** Regulon specificity score (RSS) plots
869 highlighting the top five cell type-specific transcription factors involved in mDA differentiation.
870

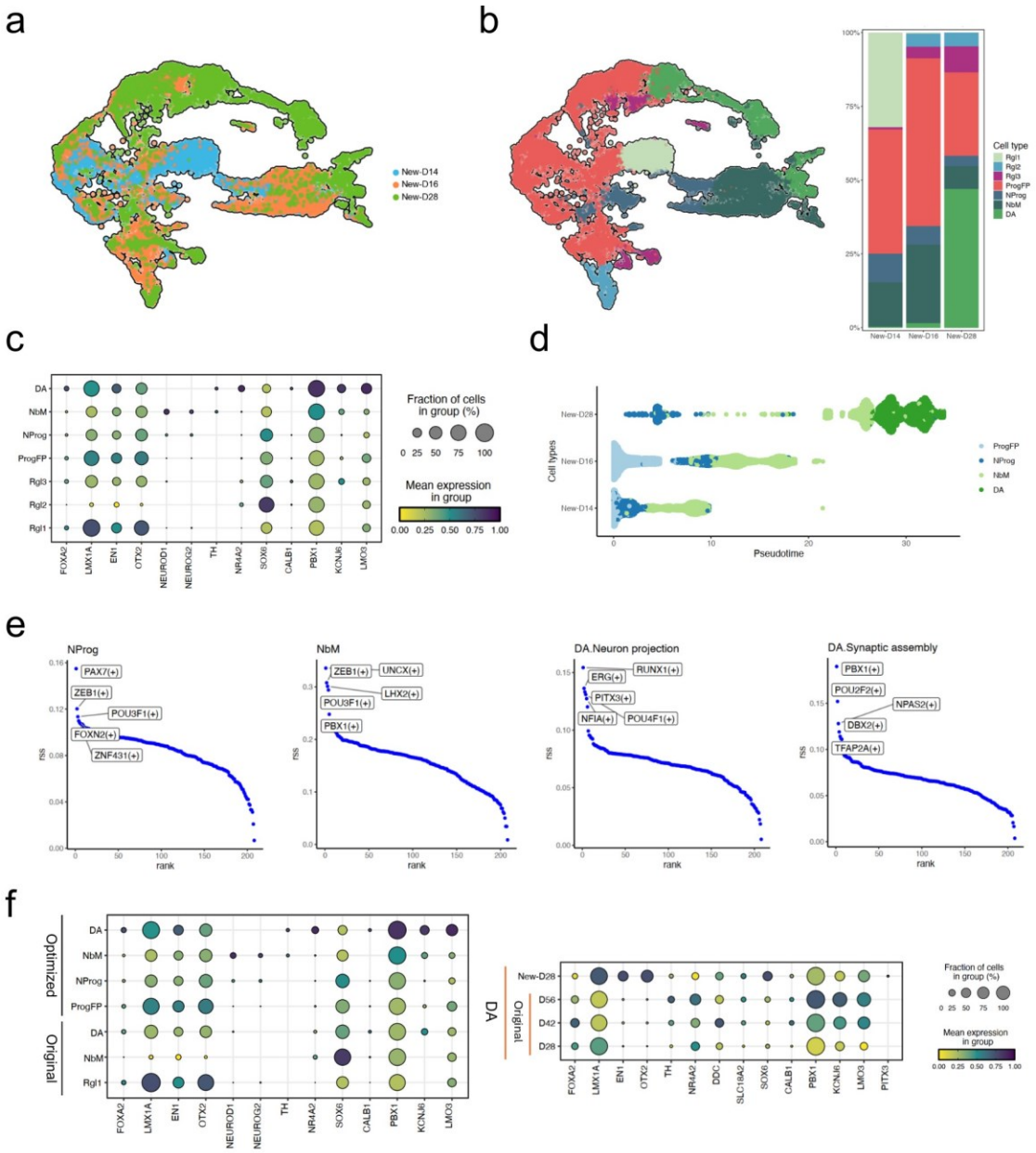
Fig. 2: Prolonged ventralization promotes accelerated differentiation of mDA neurons.



a, Schematic of the optimized differentiation protocol, illustrating the application of pharmacological compounds and growth factors. Representative immunocytochemistry images and quantification from both original protocol and new protocol showing **b,e**, efficacies of midbrain floorplate patterning at DIV11, showing co-localizations of $LMX1^{+}/EN1^{+}$ and $SOX6^{+}/OTX2^{+}$; and $EN1^{+}/SOX6^{+}/TH^{+}$ neurons at **c,f,** DIV 16 and **d,g,** 28 . White scale bars = 1 mm in B, 50 µm in **c**,**d** (n = 3 biological replicates). **h**, Calcium imaging recordings of spontaneous and KCl-induced activities on DIV 28 and 42, presenting the total

number of calcium events per neuron recorded, and the percentages of cells responding to KCl-induced depolarization. Metrics of activities, including amplitude, rise time, and fall time of calcium signals, are depicted in the subsequent bar plots. Each dot represents one cell. Two-tailed T-tests in **e-h** were performed. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$. Error bars = SEM.

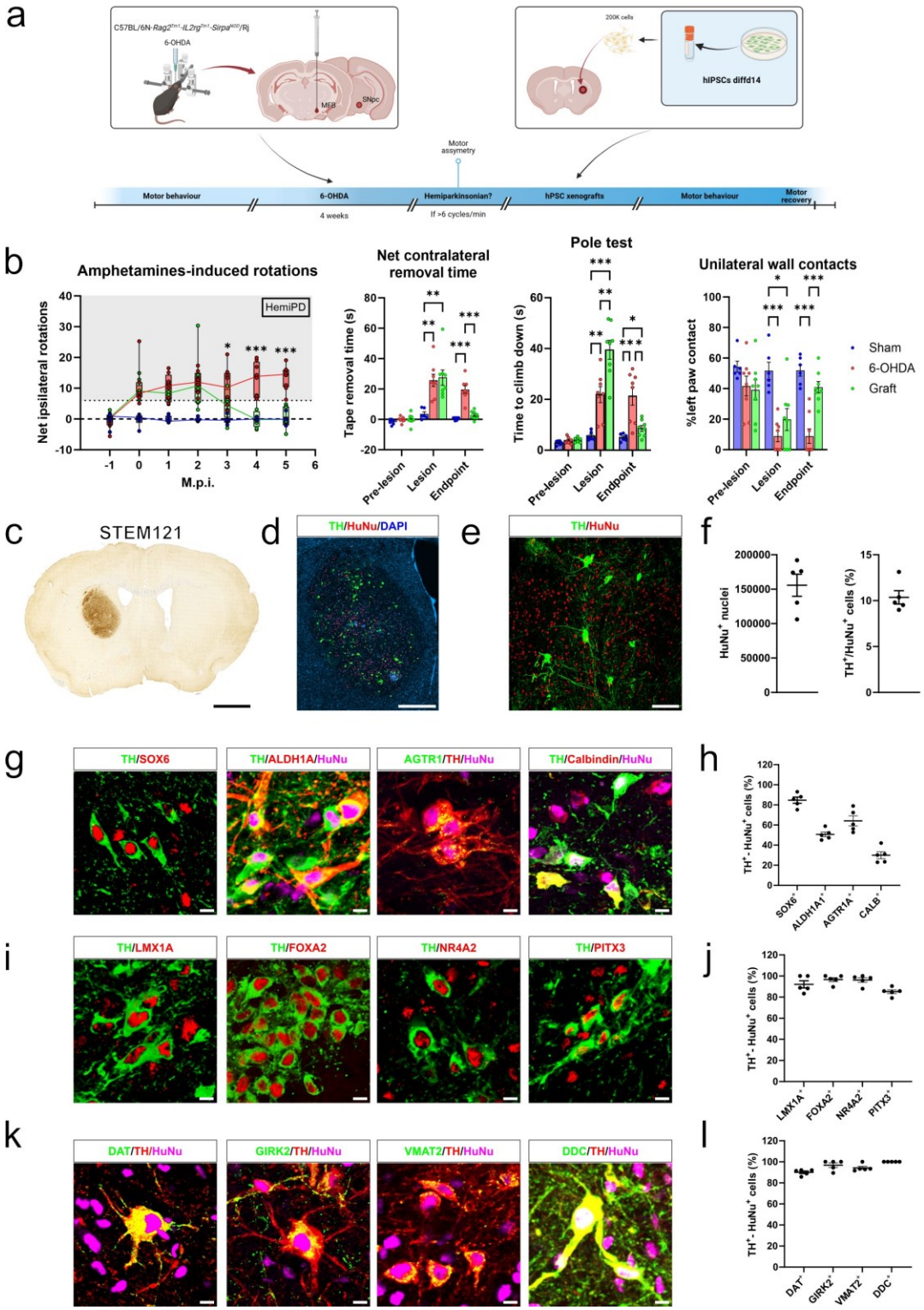
Fig. 3: snRNA-seq confirms enhanced molecular patterning under the optimized protocol.



a, UMAP embeddings of mDA differentiation in vitro using KOLF2.1J iPSCs under the new protocol, color coded by different time points at DIV 14, 16 and 28. **b,** UMAP embeddings of detailed cell types were manually annotated using top-ranked marker genes and their enriched GO terms, and proportion of cells classified into each cell type per differentiation time point. **c,** Dot plot visualizing the expression levels and the percentage of cells expressing curated key genes across distinct cell types generated under the new protocol. **d**, Pseudotime trajectory analysis showing progression from progenitor-like populations (ProgFP,

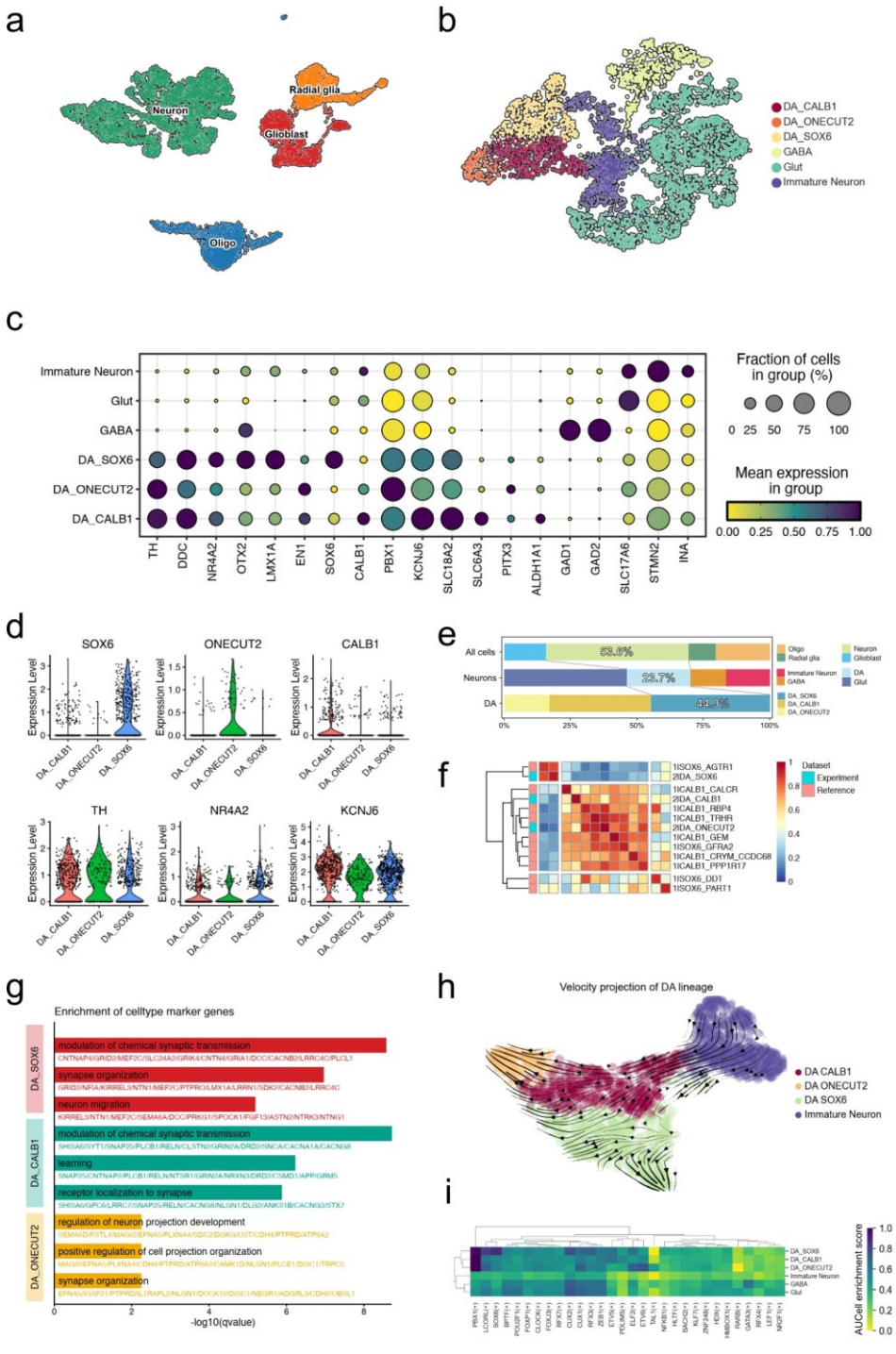
892 NProg, and NbM) to differentiated DA neurons. Cells are ordered by inferred pseudotime along the x-axis.
893 **e**, Regulon specificity score (RSS) plots highlighting the top 5 cell type-specific transcription factors driving
894 mDA differentiation under the new protocol. (f) Dot plots showing the expression levels and the percentage
895 of cells expressing midbrain dopaminergic lineage marker genes after integration in (left) midbrain cells
896 categorized as dopaminergic neurons (DA) (right), dopaminergic lineage cell types across all timepoints
897 and both conditions.
898

Fig. 4: Transplanted neurons restore motor function and mature into A9-subtype mDA neurons *in vivo*.



902 **a**, Diagram of the experimental design for the implantation of day 14-differentiated hiPSC in B6RGS mice.
903 **b**, Motor recovery is observed by continuous analysis of amphetamine-induced ipsilateral rotation, endpoint
904 tape removal test, endpoint pole test and endpoint cylinder wall analysis (right). Of note, grafted animals
905 recover pre-lesion state while the 6-OHDA control animals maintain the lesioned state. **c**, Example of
906 grafted tissue, stained against for STEM121, showing the human cells within the graft. **d,** Micrograph of
907 graft stained against for TH and HuNu, showing hDANs all over the graft. **e,** Closeup of grafted hDANs
908 showing elongated and ramified extensions. **f,** Stereological estimate of HuNu+ nuclei (left) and percentage
909 of hDANs (TH+-HuNu+) in grafts (right). **g,** Representative examples of hDANs stained positive for A9-
910 subtype marker, SOX6, ALDH1A1, AGTR1A and A10 marker, Calbindin. **h,** Quantification of percentage
911 of hDANs (TH+-HuNu+) expressing SOX6, ALDH1A1, AGTR1A and Calbindin. **i,** Representative
912 examples of hDANs stained positive for mDA neuron makers, LMX1A, FOXA2, NR4A2 and PITX3. **j,**
913 Quantification of percentage of hDANs (TH+-HuNu+) expressing LMX1A, FOXA2, NR4A2 and PITX3.
914 **k,** Representative examples of hDANs stained positive for proteins involved in DA excretion, DAT, GIRK2,
915 VMAT2 and DDC. **l,** Quantification of percentage of hDANs (TH+-HuNu+) expressing DAT, GIRK2,
916 VMAT2 and DDC. m.p.i = months post-implant. n for animal behavioral experiment **b** = sham: 6; 6-OHDA:
917 8; hIPSC-grafted: 8. N for IHC **f,h,j,l**, = 5. Scale bar = **c**: 1mm; d: 500 µm; **e**: 100 µm; **g,i,k**: 10 µm. Means
918 are represented as bars (**b**) or lines (**f,h,j,l**). One-way ANOVA with Tukey correction between treatments
919 was performed separately on each timepoint in the continuous amphetamine test ($F_{12,114} = 7.876$), as well
920 as for each separate timepoint in the endpoint experiments (Pole test $F_{4,38,} = 27.32$; Tape Removal $F_{4.38} =$
921 8.40; Cylinder Wall $F_{4,38} = 3.97$). * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$. Error bars = SEM.
922

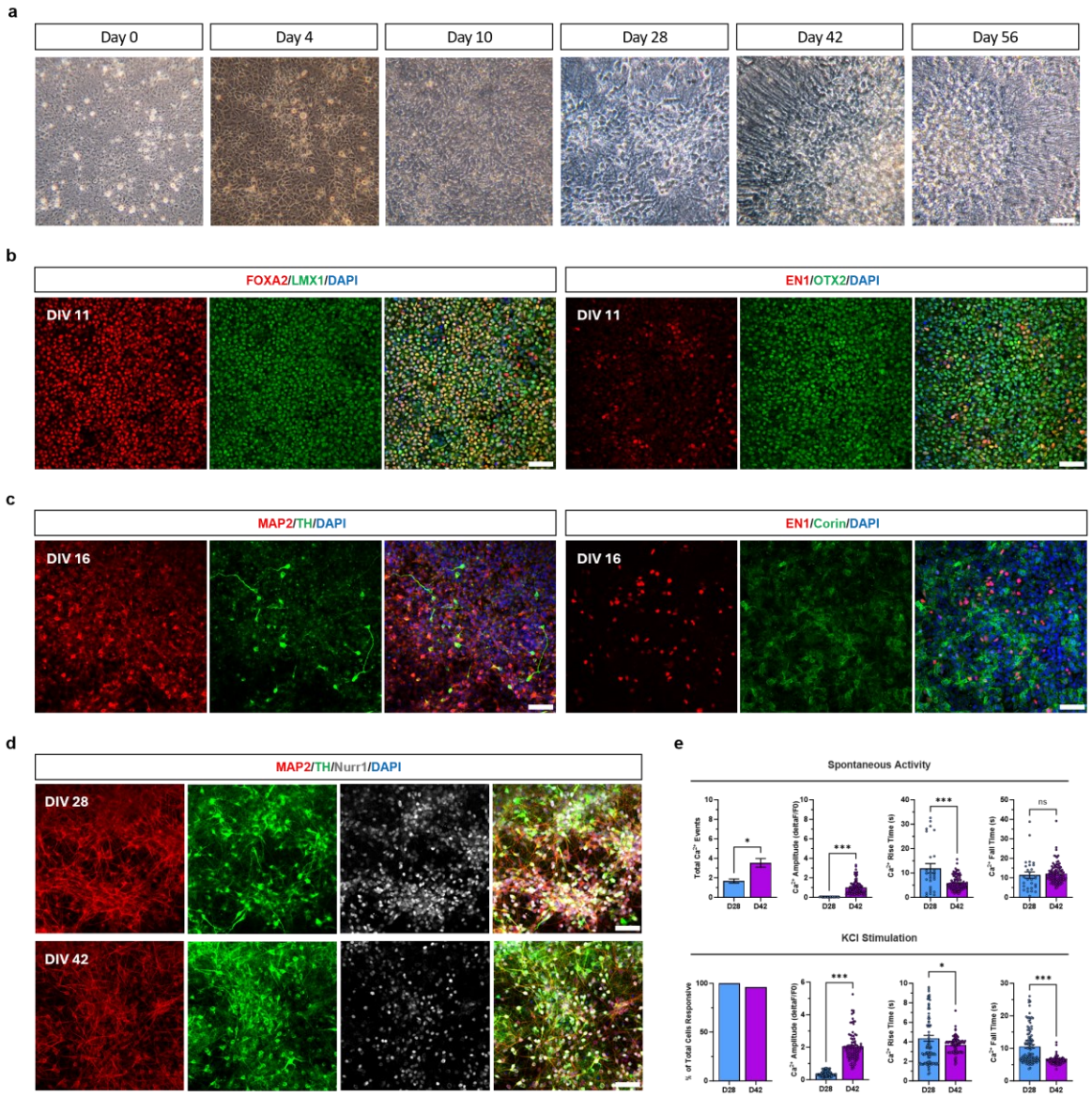
Fig. 5: snRNA-seq of grafts reveals presence of transcriptionally-defined A9-subtype neurons



a, Annotated UMAP embeddings of 9,196 high-quality human nuclei recovered from three grafted mice at five months post-transplantation, with cells classified into neurons, radial glia, gliobasts and

928 oligodendrocyte precursors; and **b**, neuronal population was consisted of immature neurons, GABAergic
929 neurons, Glutamatergic neurons and 3 subtypes of mDA neurons, each characterized by specific
930 transcription factors, as DA_CALB1, DA_ONECUT2, and DA_SOX6. **c,** Dot plot showing average gene
931 expression and percentage of expressing cells for selected marker genes across annotated neuronal cell
932 types. **d**, Violin plots of representative marker genes for dopaminergic subtypes (SOX6, ONECUT2,
933 CALB1), pan-DA markers (TH, NR4A2, KCNJ6), illustrating subtype-specific and shared expression
934 patterns. **e**, Average proportion of yielded human nuclei in each cellular hierarchy across three individual
935 datasets, demonstrating consistent representation of DA subtypes (N=3). **f**, MetaNeighbor analysis between
936 transplanted mDA neurons in this study and a reference dataset of DA neurons from the SNpc region of PD
937 patients and matched controls, suggesting transcriptional similarities between DA_SOX6 subtype and
938 SOX6_AGTR1 subtype from the reference dataset. **g,** GO enrichment analysis of DA subtype-specific
939 markers, showing distinct biological processes enriched in DA_SOX6, DA_CALB1, and DA_ONECUT2
940 populations. **h**, RNA velocity projection of mDA neuron trajectories in UMAP embeddings, suggesting
941 differentiation pathways among different DA subtypes and immature neurons. **i**, Heatmap of average
942 regulon activities of the top 30 transcription factors across mDA and other neuronal subtypes.
943

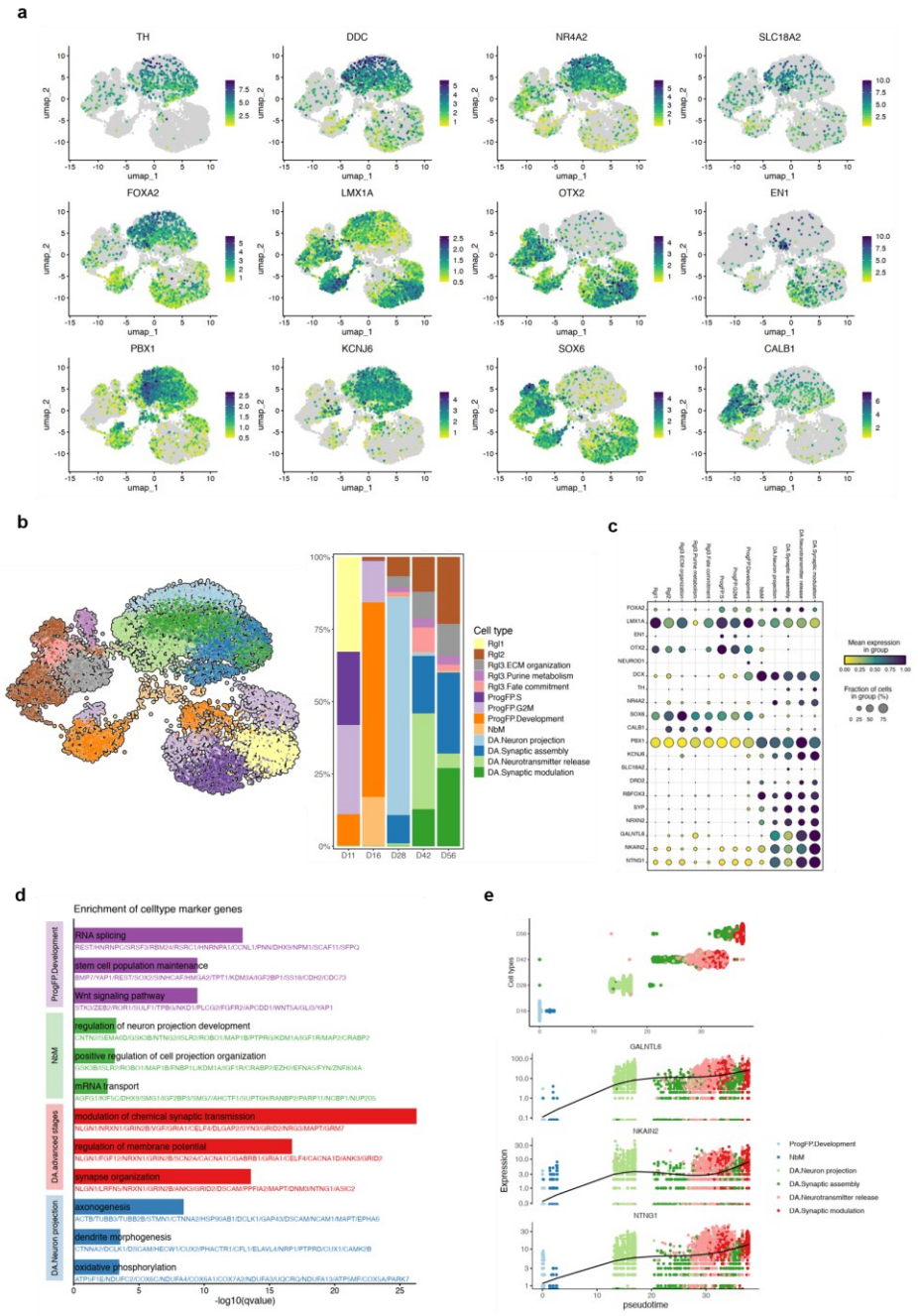
Extended Fig1. Chronological characterization of midbrain dopaminergic neurons derived from human induced pluripotent stem cell lines.



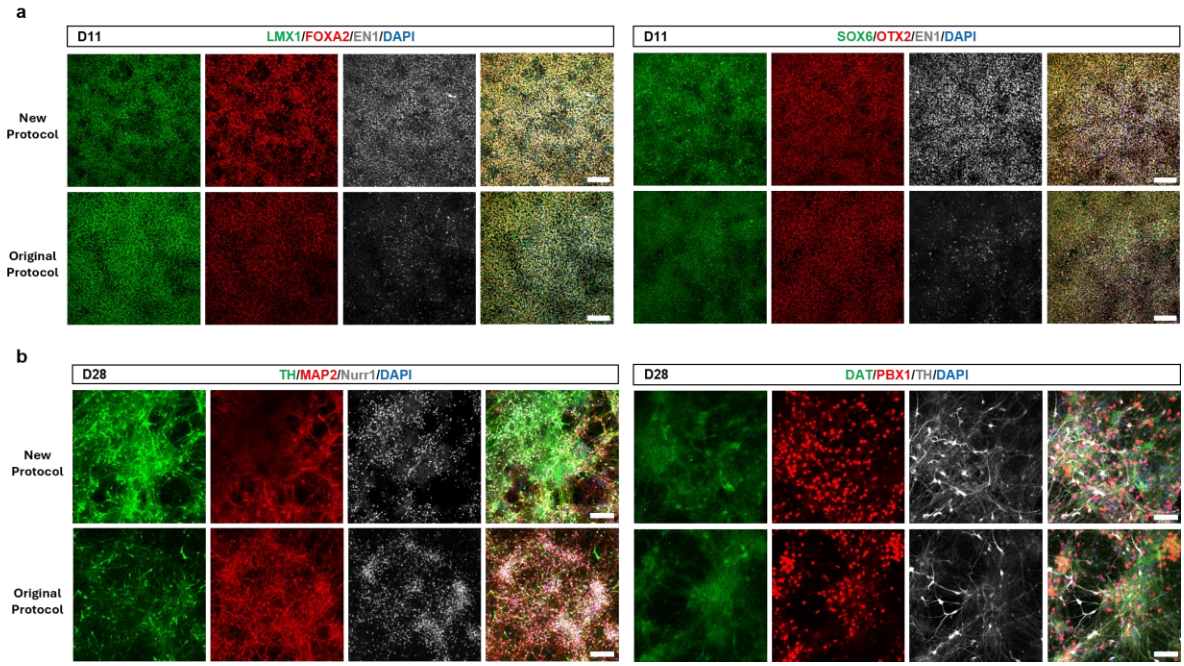
a, Brightfield images showcase the sequential progression of neuronal morphology in mDA neurons. Representative immunocytochemical images showing **b,** staining of canonical transcription factors for midbrain floor plate patterning in the process of *in vitro* differentiation at DIV 11 and **c,** 16. **d**, Maturation of midbrain dopaminergic neurons at DIV 28 and 42. **e,** Calcium imaging recordings of spontaneous (top) and KCl-induced (bottom) activities on DIV 28 and 42, presenting the total number of calcium events per

neuron recorded, and the percentages of cells responding to KCl-induced depolarization. Metrics of activities, including amplitude, rise time, and fall time of calcium signals, are depicted in the subsequent bar plots. White scale bars = 50 µm

Extended Fig2. Multiomic single-cell analysis of human iPSCs at different stages of differentiation.



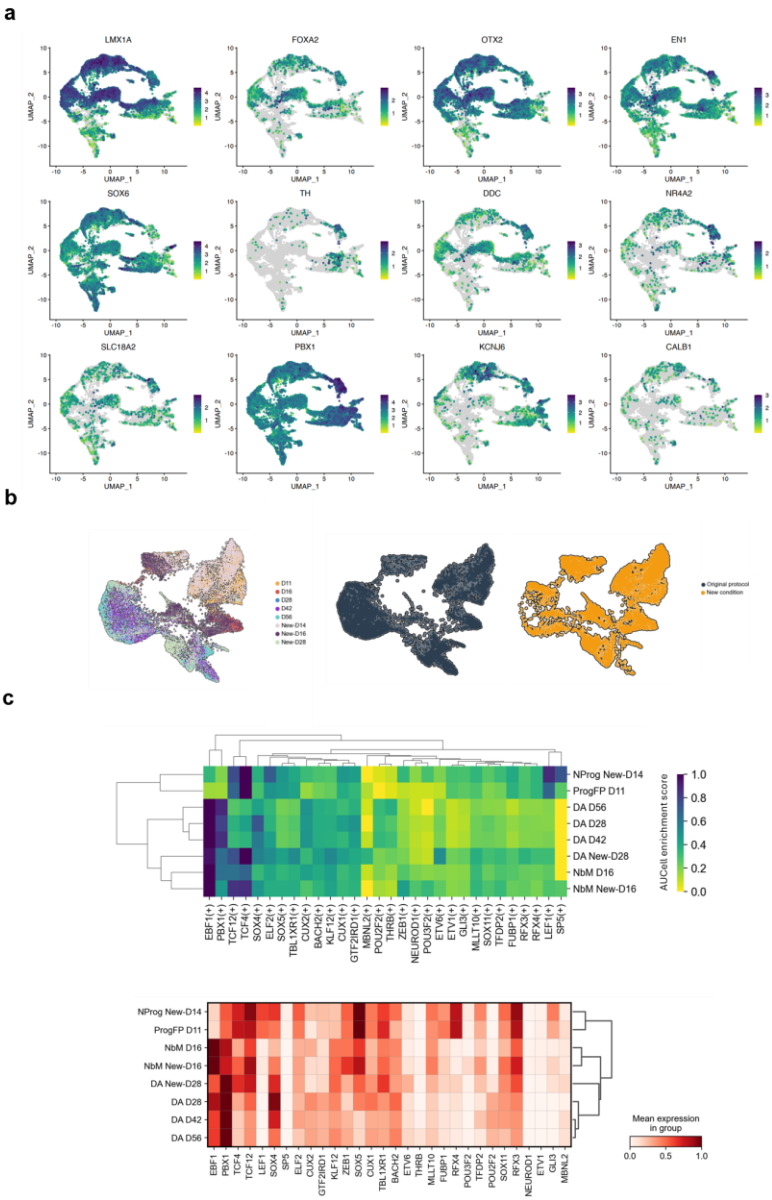
958 **a,** Single-cell gene expression of classic dopaminergic marker genes (TH, DDC, NR4A2, and SLC18A2),
959 midbrain regionalization marker genes (FOXA2, LMX1A, OTX2, and EN1), and subtyping marker genes
960 (PBX1, KCNJ6, SOX6, and CALB1) across cells during dopaminergic neuron differentiation on the
961 UMAP projection. **b**, Proportion of cells belonging to each cell type at each differentiation time point
962 (left), and proportion of cells sub-setted by functionality (right) . Detailed cell types were manually
963 annotated with top-ranked marker genes and their enriched GO terms. **c**, Dot plot visualizing the
964 expression levels and the percentage of cells expressing curated key genes across cell types during mDA
965 neuron differentiation in detailed annotations. **d,** Gene Ontology (GO) enrichment analysis of mDA-
966 lineage cells, showing distinct and related biological processes enriched in different cell types. **e,**
967 Expression trends of selected genes along the pseudotime axis, these genes participate in specific
968 biological processes pertinent to neuron maintenance or function.
969
970 **Extended Fig3. Further characterization of midbrain dopaminergic patterning of hiPSC comparing**
971 **different protocols.**



972

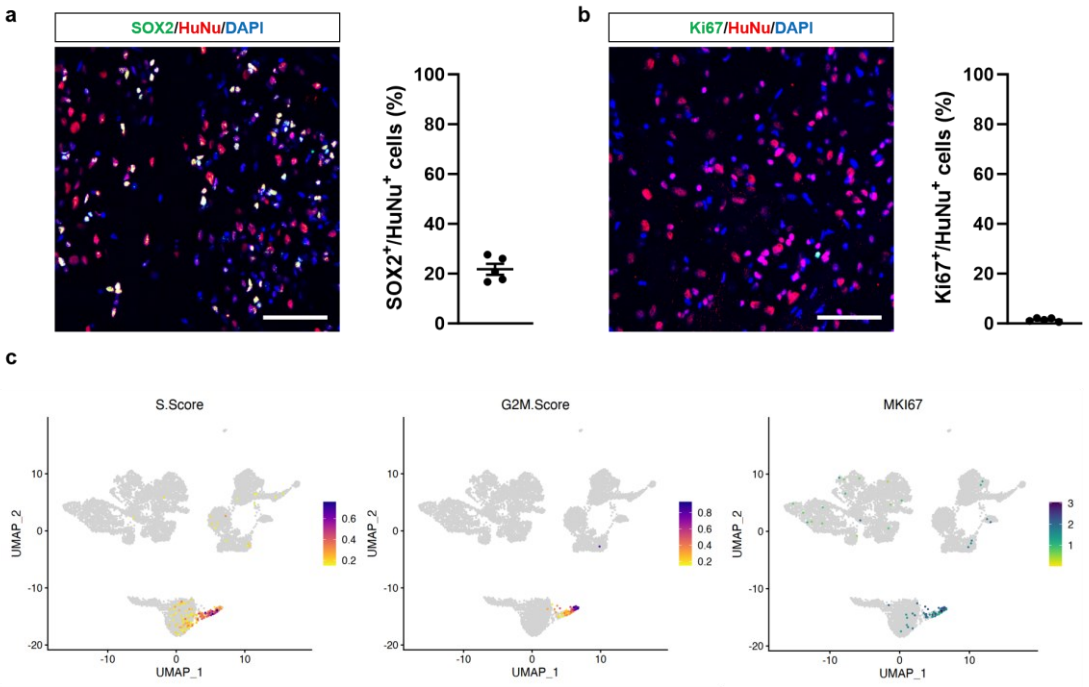
a, Representative immunocytochemical images showing staining of canonical transcription factors for midbrain floor plate patterning of *in vitro* differentiation at DIV 11. Of note, SOX6 and EN1 presented higher intensity in the new protocol. **b**, Representative immunocytochemical images of markers for the maturation of midbrain dopaminergic neurons at DIV 28, where we can observe a higher amount of TH intensity in the new protocol. White scale bars = 100 µm

Extended Fig4. Single-nuclei analysis of midbrain cells generated from the new protocol.



a, Single-nuclei gene expression of midbrain regionalization marker genes (LMX1A, FOXA2, OTX2, and EN1), classic dopaminergic marker genes (TH, DDC, NR4A2, and SLC18A2), and subtyping marker genes (PBX1, KCNJ6, SOX6, and CALB1) across cells during dopaminergic neuron differentiation on the UMAP projection. **b**, Integrated UMAP embedding of mDA differentiation in vitro using KOLF2.1J iPSCs, with nuclei from both the original (Figure 1) and new protocols (Figure 3). **c**, Heatmap of regulon activities (top) and expression levels of top 30 transcription factors (bottom) across selected mDA lineage cell types after the integration.

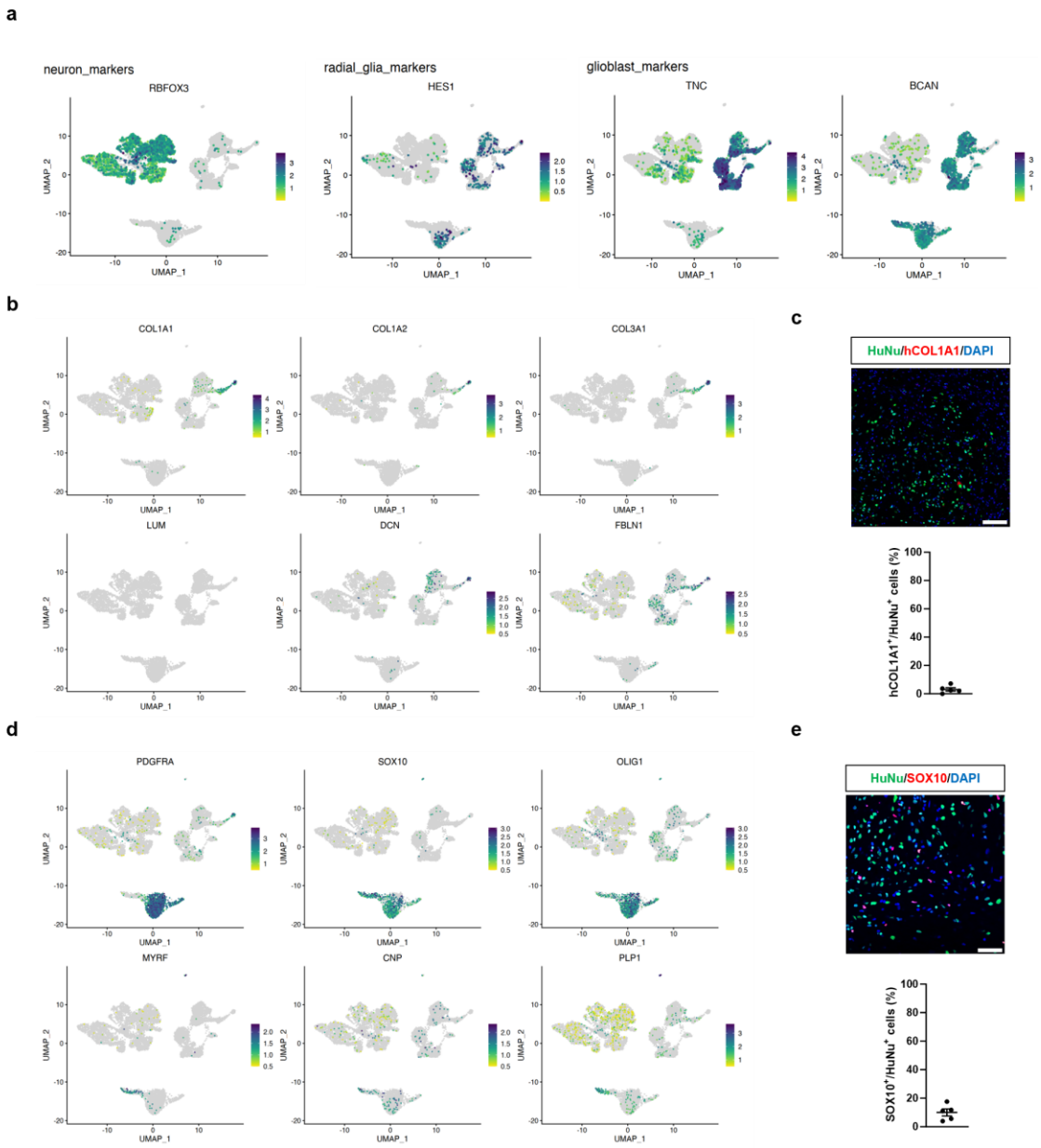
Extended Fig5. Lack of cell proliferation within grafted human DA progenitor cells.



Immunohistochemical analysis of progenitor cells in human xenografts by **a**, SOX2 staining and **b**, active proliferation by Ki67 staining in all human nuclei (HuNu). Of note, no Ki67+ were found to be human. **c**, Single-nuclei analysis of active proliferative cells by S phase scoring, G2M phase scoring and expression of Ki67. Interestingly, proliferative cells were found in the oligodendrocyte lineage cluster. Bar scale in A,B = 50 µm. Each dot represents each graft, error bar = SEM.

997 **Extended Fig6. Further identity of cells by single-nuclei marker expression in *in vivo* xenografted**

998 **cells.**



999

1000 **a**, UMAP representation of top neuronal markers (RBFOX3), radial glia markers (HES1), and glioblasts

1001 (TNC, BCAN1). **b**, UMAP representation of expression of VLMC markers (COL1A1, COL1A2,

1002 COL3A1, LUM, DCN, FLBN1). **c**, Representative micrograph of immunohistochemistry of human

1003 grafted cells against human VLMC marker hCOL1A1 (top) with quantification of human cells positive

1004 for hCOL1A1 (bottom). **d**, UMAP representation of oligodendroglial markers (PDGFRA, SOX10,

1005 OLIG1, MYRF, CNP, PLP1). **e**, Representative micrograph of immunohistochemistry of human grafted

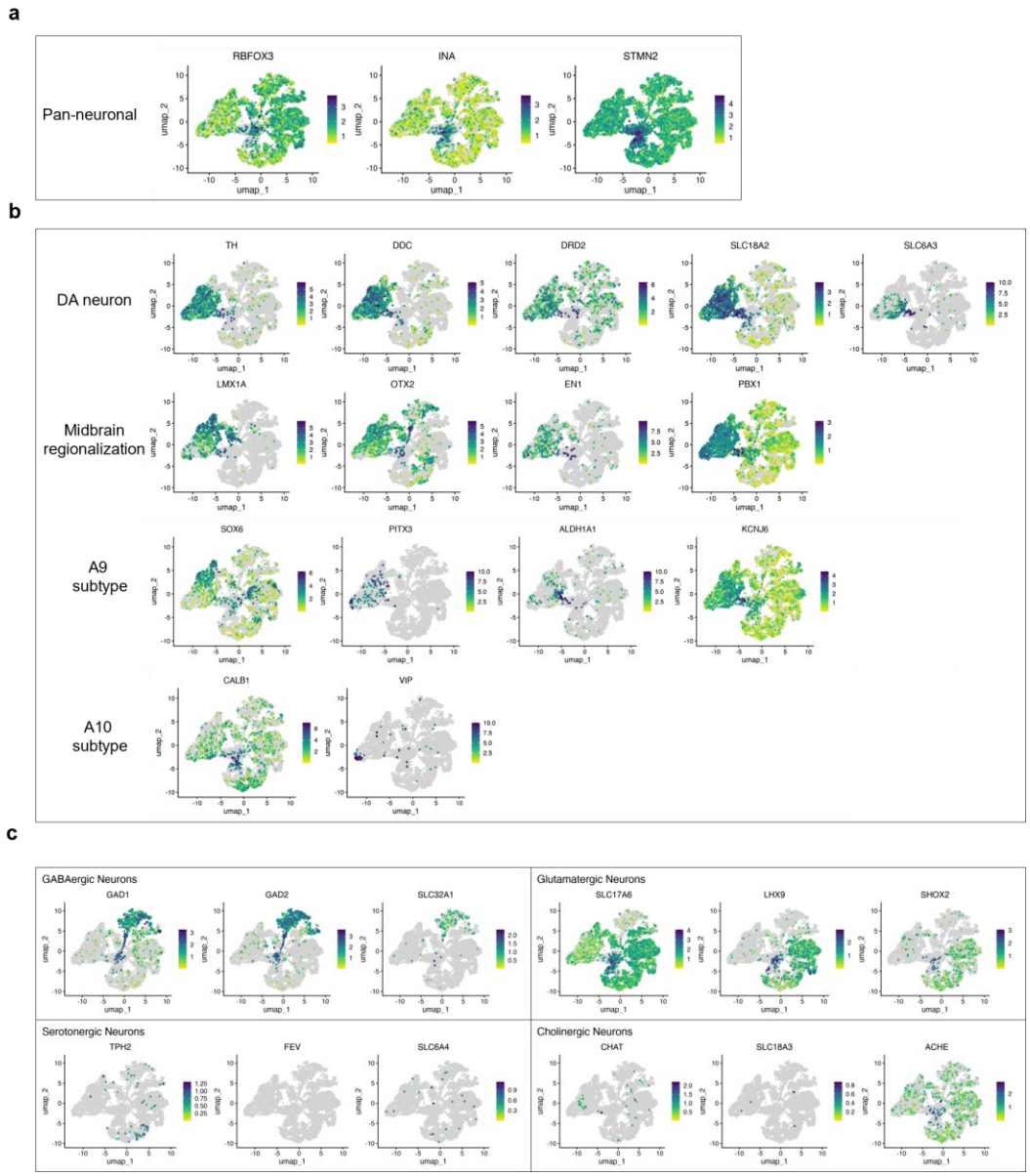
1006 cells against OPC marker SOX10 (top) with quantification of human cells positive for SOX10 (bottom).

1007 Scale bar = 100 μm.

1008

1009 **Extended Fig7. Identification of specific markers within the neuronal cluster of the human grafted**

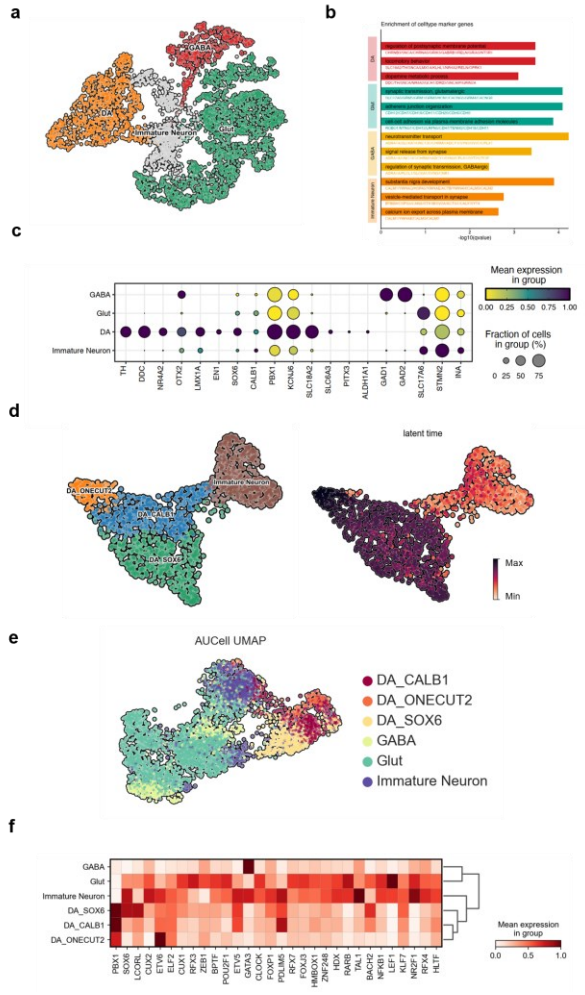
1010 **cells.**



1011

a, UMAP representation of pan-neuronal markers (RBFOX3, INA, STMN2). **b,** UMAP representation of DA neurons specificity based on DA neuronal general markers (TH, DDC, DRD2, SLC18A2, SLC6A3), for midbrain regionalized DA neurons (LMX1A, OTX2, EN1, PBX1), for A9 subtype DA neurons (SOX6, PITX3, ALDH1A1 and KCNJ6), and A10 subtype (CALB1, VIP). **c,** UMAP representation of non-DAergic neurons such as GABAergic neurons (GAD1, GAD2, SLC32A1), glutamatergic neurons (SCL17A6, LMX3, SHOX2), serotonergic neurons (TPH2, FEV, SLC6A4) and cholinergic neurons (CHAT, SLC18A3, ACHE). Of note, the grafted neuronal cluster lacks cells identifying as serotonergic neurons.

Extended Fig8. **Single-nuclei analysis of neuronal clusters of *in vivo* xenograft.**



1023 **a**, Annotated UMAP embeddings of neuronal population, including immature neurons, GABAergic
1024 neurons, glutamatergic neurons and dopaminergic neurons. **b**, Gene Ontology (GO) enrichment analysis
1025 of different types of neurons, showing distinct biological processes enriched according to their specific
1026 marker genes. **c,** Dot plot showing average gene expression and percentage of expressing cells for
1027 selected marker genes across annotated neuronal cell types. **d**, Visualization with UMAP embeddings
1028 showing mDA-related cell types, denotated in their subtypes and predicted latent time. **e**, UMAP
1029 embeddings of AUCell scores across different subtypes of neurons. **f**, Heatmap of expression levels of
1030 top 30 specific transcription factors across selected cell types.