Embryonic origins of forebrain oligodendrocytes revisited by combinatorial genetic fate mapping

Yuqi Cai[1#], Zhirong Zhao[1#], Mingyue Shi[1], Mingfang Zheng[1], Ling Gong[1], Miao He[1*]

[1] Institutes of Brain Science, State Key Laboratory of Medical Neurobiology and MOE Frontiers Center for Brain Science, Department of Neurobiology, Zhongshan Hospital, Fudan University, Shanghai 200032, China.

[#]These authors contributed equally to this work.

[*]Corresponding author hem@fudan.edu.cn

Abstract

Multiple embryonic origins give rise to forebrain oligodendrocytes (OLs), yet controversies and uncertainty exist regarding their differential contributions. We established intersectional and subtractional strategies to genetically fate map OLs produced by medial ganglionic eminence/preoptic area (MGE/POA), lateral/caudal ganglionic eminences (LGE/CGE) and dorsal pallium. We found that, contrary to the canonical view, LGE/CGE-derived OLs make minimum contributions to the neocortex and corpus callosum, but dominate piriform cortex and anterior commissure. Additionally, MGE/POA-derived OLs, instead of being entirely eliminated, make small but sustained contribution to cortex with a distribution pattern distinctive from those derived from the dorsal origin. Our study provides a revised and more comprehensive view of cortical and white matter OL origins, and established valuable new tools and strategies for future OL studies.

Introduction

Oligodendrocytes (OLs) are an important class of macroglia responsible for producing the myelin sheaths that insulate and protect neuronal axons. Forebrain OLs arise from multiple embryonic origins. Previous fate mapping study using *Nkx2.1*^{Cre}[1], *Gsh2*^{Cre}[2] and *Emx1*^{Cre}[3] reported consecutive and competing waves of OLs derived from medial ganglionic eminence/preoptic area (MGE/POA), lateral/caudal ganglionic eminences (LGE/CGE) and dorsal pallium[2]. The first-wave of OLs generated by MGE/POA ($_{MP}$OLs) were believed to be eliminated postnatally, while those from the second and third waves ($_{LC}$OLs and dOLs) survive and populate the cortex and corpus callosum at comparable proportions. Several other studies provided both supporting and contradicting evidence to this model[4-9]. Moreover, *Gsh2* was recently found to be expressed in dorsal progenitors[10], casting doubt on the interpretation of lineage tracing data from *Gsh2*^{Cre}.

In this study, we generated new genetic tools and combinatorial fate mapping strategies which allow direct visualization and comparison among OLs derived from different origins. We found that neocortical OLs are primarily composed of dOLs, rather than similar proportions of $_{LC}$OLs and dOLs. In contrast, $_{LC}$OLs and dOLs made comparable contributions to piriform cortex. We also found that although $_{MP}$OLs only make a small contribution, they do persist in the cortex beyond adulthood with a unique spatial pattern distinct from that of the dOLs. In the two major white matter commissure tracts, dOLs are the vast majority in corpus callosum but make little contribution to anterior commissure, while $_{LC}$OLs behaved the opposite. These findings significantly revised the classical view and provided a new and more comprehensive picture of cortical and white matter OL origins.

45 **Results and Discussion**

46 To unambiguously track OLs from different embryonic origins, we first generated a knock-in
47 driver, *Opalin*$^{P2A\text{-}Flpo\text{-}T2A\text{-}tTA2}$ (Figure 1), orthogonal to Cre drivers that label dorsal or ventral
48 progenitors (*Progenitor*^{Cre}). *Opalin* (also known as *Tmem10*) encodes oligodendrocytic myelin
49 paranodal and inner loop protein that are specifically expressed in differentiated
50 oligodendrocytes[11-14]. In *Opalin*$^{P2A\text{-}Flpo\text{-}T2A\text{-}tTA2}$, *Flpo* and *tTA2* were inserted before the STOP
51 codon and linked by self-cleavage peptide P2A and T2A (Figure 1A-C), allowing co-
52 transcription and translation with *Opalin*. Flp-mediated recombination by this driver
53 (hereinafter referred to as *Opalin*^{Flp} for simplicity) enables highly specific, efficient and
54 irreversible OL labeling, while the tTA2 component offers the flexibility for labeling in tunable
55 densities (Figure 1D-F).

56 Next, we established two types of genetic combinatorial fate mapping strategies to directly
57 visualize OLs from different embryonic origins (Figure 2): (1) combining *Opalin*^{Flp} and
58 *Progenitor*^{Cre} with intersectional reporters *Ai65* to label OLs derived from Cre+ progenitor
59 domain by RFP (Figure 2A); (2) combining *Opalin*^{Flp} and *Progenitor*^{Cre} with *RC::FLTG*[15] to
60 simultaneously label OLs derived from Cre+ progenitors by GFP and OLs derived from the
61 complementing Cre- progenitors by RFP (Figure 2B). The first approach allowed us to track
62 dOLs and $_{MP}$OLs (Figure 2C, E). The second approach empowered us to observe and compare
63 OLs generated from dorsal and ventral origins, or those from Gsh2+ and Gsh2- progenitors, in
64 the same brain (Supplementary Figure 1). Importantly, the subtraction power enabled us to
65 target OLs derived from LGE/CGE progenitors that express neither *Emx1* nor *Nkx2.1* (Figure
66 2D). In addition, these strategies greatly facilitated the identification of OLs derived from

67 specific origin which exist at relatively low density in certain regions.

68 Deploying these strategies, we assessed the differential contributions of dOLs, $_{LC}$OLs and

69 $_{MP}$OLs by analyzing RFP+ cells in the following mice: *Opalin*^{Flp}::*Emx1*^{Cre}::*Ai65* (Figure 2C),

70 *Opalin*^{Flp}::*Emx1*^{Cre}::*Nkx2.1*^{Cre}::*RC::FLTG* (Figure 2D) and *Opalin*^{Flp}::*Nkx2.1*^{Cre}::*Ai65* (Figure

71 2E). We observed two significant differences from the traditional model in the neocortex. The

72 first major deviation is that, instead of comparable contributions by dOLs and $_{LC}$OLs, the vast

73 majority of neocortical OLs were dOLs but not $_{LC}$OLs. The density of dOLs is significantly

74 higher than that of $_{LC}$OLs in the motor cortex (Mo) (504.5±25.1 vs. 5.3±0.7; P<0.001; n=3)

75 and somatosensory cortex (SS) (518.0±22.6 vs.21.6±2.3; P<0.001; n=3) (Figure 2C, D, F). Co-

76 staining with aspartoacylase (ASPA)[16], a marker for mature OL, also revealed a very high

77 neocortical OL labeling rate in *Opalin*^{Flp}::*Emx1*^{Cre}::*Ai65* (RPF+ASPA+/ASPA+: 94.6±2.0% in

78 Mo and 90.5±2.2% in SS; n=3).Considering the possibility of incomplete recombination in

79 combinatorial reporters, and the relatively low Cre activity in the dorsal MGE of *Nkx2.1*^{Cre}[1],

80 the genuine contribution of $_{LC}$OLs to the neocortex could be even lesser than our current

81 observation. Therefore, the large quantity of neocortical OLs labeled by *Gsh2*^{Cre} in previous

82 study[2] or by GFP in *Opalin*^{Flp}::*Gsh2*^{Cre}:: *RC::FLTG* (Supplementary Figure 1B) most likely

83 were predominantly dOLs generated by Gsh2+ dorsal progenitors[10], rather than bona fide

84 $_{LC}$OLs.

85 The second major deviation is that cortical $_{MP}$OLs are not completely depleted postnatally,

86 instead, they make a small but continued contribution with a unique spatial distribution pattern

87 (Figure 2D and Supplementary Figure 2). $_{MP}$OLs display a clear rostrocaudal density decline

88 (Supplementary Figure 2A-B) and a laminar preference towards layer 4 (L4) in SS

(Supplementary Figure 2C-E). In contrast, the distribution of dOLs does not vary significantly across the rostrocaudal axis but exhibits increased density towards deeper layers (Supplementary Figure 2A-E). Importantly, we have observed cortical $_{MP}$OLs in mice as old as one year (Supplementary Figure 2F), well beyond the age analyzed in previous reports[2, 7, 8], suggesting a persisted contribution.

We then turned our attention to the lateral three-layer archicortex, piriform cortex (Pir). Different from the neocortex, Pir contains comparable density of dOLs and $_{LC}$OLs (114.1±12.5 vs. 171.7±23.6; P=0.058; n=3) (Figure 2F). $_{MP}$OLs make the lowest contribution at a density similar to SS(16.4±2.4 vs.15.8±1.4; P=0.966; n=4) (Figure 2E).

These combinatorial models also grant us the opportunity to revisit the differential contributions of dOLs, $_{LC}$OLs and $_{MP}$OLs to the two commissural white matter tracts, corpus callosum (cc) and anterior commissure (ac), which contain higher density of OLs than the grey matter (Figure 3). We found that, similar to the neocortex, cc is mainly populated by dOLs and supplemented by very low density of $_{LC}$OLs and $_{MP}$OLs(Figure 3A-D). Interestingly, $_{LC}$OLs and $_{MP}$OLs seem to show preferential distribution in the lateral and medial regions of cc (cc-l and cc-m), respectively (Figure 3B-D, boxed regions in right panels). Different from cc, ac is mainly populated by $_{LC}$OLs (2454.6±254.1, n=3) and $_{MP}$OLs (1146.3±101.6; n=4) and supplemented by very low density of dOLs (85.2±34.5; n=3) (Figure 3D).

To substantiate the above results, we further breed *Opalin*^{Flp}::*Emx1*^{Cre}::*Nkx2.1*^{Cre}::*Ai65* to label dOLs together with $_{MP}$OLs by RFP and co-stained them with ASPA (Supplementary Figure 3). RFP-ASPA+ cells were difficult to find in Mo, SS and cc, but were more easily observed in Pir and ac, consistent with the respective low and high $_{LC}$OL densities in these regions.

In summary, our findings significantly revised the canonical model of forebrain OL origins (Figure 4A), and provided a new and more comprehensive view (Figure 4B). We demonstrated that neocortical OLs are mainly derived from dorsal origin with small but lasting contribution from the ventral origin (Figure 2, Supplementary Figure 1A). Our data showed that LGE/CGE makes little contribution to neocortex and cc, but makes major contribution to piriform cortex and ac. This finding is supported by another report[17] in which in utero electroporation failed to label LGE-derived cortical OLs. Taken together, these results suggest that the lack of $_{LC}$OL in neocortex is less likely caused by competitive postnatal elimination, but more likely due to limited production and/or allocation. We further discovered that MGE/POA makes a small but persistent contribution to the neocortex with a distinct distribution pattern featured by a rostral-high to caudal-low gradient and a preference towards L4 in SS. Whether their enduring existence and highly biased localization has functional implications awaits future exploration. In addition, we found that the cc showed a similar OL composition as the neocortex, but the Pir and the ac each exhibited distinct OL compositions in term of their embryonic origins. $_{LC}$OLs are the major contributor to both regions, while dOLs and $_{MP}$OLs mainly contribute to Pir and ac, respectively.

In addition to the new framework of forebrain OL origins (Figure 4), we also generated a new driver (Figure 1) and established multiple combinatorial genetic models (Figure 2) for efficient tracking and direct visualization of OLs from different embryonic origins without interference from other cells types. These tools set up a firm foundation and will provide reliable experimental access for future inquiries on the development and function of diverse OLs in healthy and disease brains[18], especially to uncover the relationship between developmental

origins and the functional and molecular heterogeneity of OLs.

Material and Methods

Mice

All mouse studies were carried out in strict accordance with the guidelines of the Institutional Animal Care and Use Committee of Fudan University. All husbandry and experimental procedures were reviewed and approved by the same committee. All applicable institutional and/or national guidelines for the care and use of animals were followed. The following transgenic mouse lines were used in this study: *Nkx2.1*^{Cre} (Jax 008661)[1],*Gsh2*^{Cre}(Jax 025806)[2], *Emx1*^{Cre} (Jax 005628)[3], *Ai65* (Jax 021875)[19], *RC::FLTG* (Jax 026932)[15]. The tTA2-dependent tdTomato reporter (*TRE-RFP*) was derived from *Ai62* (Jax 022731) [19], by removing LoxP-STOP-LoxP with *E2a-Cre* (Jax 003724). The Flp dependent H2B-GFP reporter (*HG-FRT*) was derived from *HG-dual* (Jax 028581) via removal of loxP flanking STOP cassette by *CMV-Cre*[20]. The *Opalin*$^{P2A\text{-}Flpo\text{-}T2A\text{-}tTA2}$ allele was generated by targeted insertion of the T2A-Flpo-P2A-tTA2 sequence immediately before the STOP codon of the endogenous *Opalin* gene using homologous recombination. Gene targeting vector was generated using PCR-based cloning approach as described before[20]. More specifically, a 4.7 kb 5' homology arm, a loxP flanking Neo positive selection cassette, a T2A-FlpO-P2A-tTA2 cassette and a 2.7 kb 3' homology arm were cloned into a building vector containing the DTA negative selection cassette to generate the targeting vector. Targeting vector was linearized and transfected into a C57/black6 ES cell line. ES clones that survived through negative and positive selections were first screened by genomic PCR, then confirmed by Southern blotting using appropriate DIG-dUTP labelled probes. One positive ES cell clone was used for blastocyst injection to obtain

155 male chimera mice carrying the modified allele following standard procedures. Chimera males
156 were bred with C57BL/6J females to confirm germline transmission by genomic PCR. The Neo
157 selection cassette was self-excised during spermatogenesis of F0 chimeras. Heterozygous F1
158 siblings were bred with one another to establish the colony. Targeting vector construction, ES
159 cell transfections and screening, blastocyst injections and chimera breeding were performed by
160 Biocytogen Co, China.

161 **Genomic PCR**

162 Genomic DNA was prepared from mouse tails. Tissue was lysed by incubation in tail lysis
163 buffer (Viagen, 102-T) with 0.1 mg/ml proteinase K (Diamond, A100706) overnight at 55°C
164 followed by 45 min at 90°C in an air bath to inactivate proteinase K. The lysate was cleared by
165 centrifugation at maximum speed (21,130 G) for 15 min in a table-top centrifuge. Supernatant
166 containing genomic DNA was used as the PCR template for amplifying DNA products. The
167 following primers were used:

168 *Opalin-F*: 5'- GGCCTATGTTTGATTTCCAGCACTG -3'
169 *Opalin-R*: 5'- AGCACTTATGACTGCTGAGCCGTTC -3'

170 **Immunohistochemistry and microscopy**

171 Mice were anaesthetized by intraperitoneal injection of 1.5% sodium pentobarbital (0.09 mg/g
172 body weight) and then intracardially perfused with saline followed by 4% paraformaldehyde in
173 0.1 M PB. Following post fixation at 4°C for 24 hours, brain samples were sectioned at 30 μm
174 using a vibratome (Leica VT1000S), or transferred into 30% sucrose in 0.1M PB for
175 cryoprotection, OCT embedded, and sectioned using a cryostat (Leica CM1950). For CC1

176 immunostaining, antigen retrieval was performed prior to blocking by boiling for 3 min in 10
177 mM citrate buffer (pH6.0). Sections were blocked in PBS containing 0.05% Triton and 5%
178 normal donkey serum and then incubated with the following primary antibodies in the blocking
179 solution at 4 ℃ overnight: RFP (goat polyclonal antibody, 1:2000, SICGEN AB0081-200;
180 rabbit polyclonal antibody, 1:2000, Rockland 600-401-379), GFP (chicken polyclonal antibody,
181 1:1000, Aves Labs, GFP-1020), MBP (rat polyclonal antibody, 1:500, AbD Serotec, MCA409S),
182 CC1 (rabbit polyclonal antibody, 1:500, Oasis Biofarm, OB-PRB070, mouse polyclonal
183 antibody, 1:300, Millipore, OP80), ASPA (rat polyclonal antibody, 1:200, Oasis Biofarm, OB-
184 PRT005), Sox9(rabbit polyclonal antibody, 1:2000, Chemicon, AB5535), NeuN (mouse
185 monoclonal antibody, 1:500, Millipore, MAB377). Sections were then incubated with
186 appropriate Alexa fluor dye-conjugated IgG secondary antibodies (1:500, Thermofisher
187 Scientific) or CF dye-conjugated IgG secondary antibodies (1:250, Sigma) in blocking solution
188 and mounted in Aqua-mount (Southern Biotech, 0100-01). Sections were counterstained with
189 DAPI. Sections were imaged with confocal microscopy (Olympus FV3000), fluorescence
190 microscopy (Nikon Eclipse Ni; Olympus VS120; Olympus VS200) and fluorescent stereoscope
191 (Nikon SMZ25). Anatomical regions in the postnatal brain were identified according to the
192 *Paxinos "The Mouse Brain" Atlas* and the *Allen Reference Atlas*, and their areas were measured
193 in ImageJ, whenever applicable. One out of every four sections within the range of selection
194 from each brain and at least three brains per genotype were analyzed. To quantify density and
195 co-localization, cells were identified and counted in Image J or QuPath.

196 **Statistical analysis**

GraphPad Prism version 8.0.1 was used for statistical calculations. No statistical methods were used to predetermine sample sizes, but our sample sizes are similar to those reported in previous publications. Data collection and analysis were performed blind to the conditions of the experiments whenever possible. No animals or data points were excluded from the analysis. Normalcy was assessed using Shapiro-Wilk test. Equal variances were assessed using F test or Bartlett's test. Statistical significance was tested using two-tailed unpaired t-test, Welch's t-test, one-way ANOVA and two-way ANOVA followed by Tukey's or Bonferroni post hoc test, wherever appropriate. Data are presented as mean ± standard error of the mean (S.E.M.). $P < 0.05$ was considered significant. Significance is marked as $^{*}P < 0.05$; $^{**}P < 0.01$ and $^{***}P < 0.001$.

Acknowledgements. We thank Dr. Yilin Tai for helpful discussion and Dr. Min Jiang from core facility of IOBS, Fudan University, for technical support on imaging experiments. This study was supported by funds from the National Science and Technology Innovation 2030 Major Projects of China (STI2030-Major Projects-2022ZD0206500), National Natural Science Foundation of China (32171087, 32200918, 32371145).

Author contributions.

Conceptualization: MH

Methodology: YC, MH

Investigation: YC, ZZ, MS

Visualization: YC, MH, LG

Supervision: MH, LG

Writing—original draft: MH, YC, LG

Writing—Review and editing: MH, YC, ZZ, MZ

Declaration of interests. The authors declare no competing interests.

Reference

1. Xu, Q., M. Tam, and S.A. Anderson, *Fate mapping Nkx2.1-lineage cells in the mouse telencephalon.* J Comp Neurol, 2008. **506**(1): p. 16-29.
2. Kessaris, N., et al., *Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage.* Nat Neurosci, 2006. **9**(2): p. 173-9.
3. Gorski, J.A., et al., *Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage.* The Journal of neuroscience : the official journal of the Society for Neuroscience, 2002. **22**(15): p. 6309-14.
4. Nakahira, E., et al., *Direct evidence that ventral forebrain cells migrate to the cortex and contribute to the generation of cortical myelinating oligodendrocytes.* Dev Biol, 2006. **291**(1): p. 123-31.
5. Tsoa, R.W., et al., *Spatiotemporally different origins of NG2 progenitors produce cortical interneurons versus glia in the mammalian forebrain.* Proc Natl Acad Sci U S A, 2014. **111**(20): p. 7444-9.
6. Naruse, M., et al., *The Dorsoventral Boundary of the Germinal Zone is a Specialized Niche for the Generation of Cortical Oligodendrocytes during a Restricted Temporal Window.* Cereb Cortex, 2016. **26**(6): p. 2800-2810.
7. Orduz, D., et al., *Developmental cell death regulates lineage-related interneuron-oligodendroglia functional clusters and oligodendrocyte homeostasis.* Nat Commun, 2019. **10**(1): p. 4249.
8. Liu, R., et al., *In Vivo Clonal Analysis Reveals Development Heterogeneity of Oligodendrocyte Precursor Cells Derived from Distinct Germinal Zones.* Adv Sci (Weinh), 2021. **8**(20): p. e2102274.
9. Shen, Z., et al., *Distinct progenitor behavior underlying neocortical gliogenesis related to tumorigenesis.* Cell Rep, 2021. **34**(11): p. 108853.
10. Zhang, Y., et al., *Cortical Neural Stem Cell Lineage Progression Is Regulated by Extrinsic Signaling Molecule Sonic Hedgehog.* Cell Rep, 2020. **30**(13): p. 4490-4504 e4.
11. Kippert, A., et al., *Identification of Tmem10/Opalin as a novel marker for oligodendrocytes using gene expression profiling.* BMC Neurosci, 2008. **9**: p. 40.
12. Yoshikawa, F., et al., *Opalin, a transmembrane sialylglycoprotein located in the central nervous system myelin paranodal loop membrane.* J Biol Chem, 2008. **283**(30): p. 20830-40.
13. Jiang, W., et al., *Identification of Tmem10 as a novel late-stage oligodendrocytes marker for detecting hypomyelination.* Int J Biol Sci, 2013. **10**(1): p. 33-42.
14. Marques, S., et al., *Oligodendrocyte heterogeneity in the mouse juvenile and adult central nervous system.* Science, 2016. **352**(6291): p. 1326-1329.
15. Plummer, N.W., et al., *Expanding the power of recombinase-based labeling to uncover cellular diversity.* Development, 2015. **142**(24): p. 4385-93.
16. Huang, H., et al., *Immunological Markers for Central Nervous System Glia.* Neurosci Bull, 2023. **39**(3): p. 379-392.

17. Li, J., et al., *The Lateral Ganglionic Eminence Does Not Generate Cortical Oligodendrocytes.* bioRxiv, 2023: p. 2023.12.01.569674.
18. Gong, L., et al., *Emerging strategies for the genetic dissection of gene functions, cell types, and neural circuits in the mammalian brain.* Mol Psychiatry, 2022. **27**(1): p. 422-435.
19. Madisen, L., et al., *Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance.* Neuron, 2015. **85**(5): p. 942-58.
20. He, M., et al., *Strategies and Tools for Combinatorial Targeting of GABAergic Neurons in Mouse Cerebral Cortex.* Neuron, 2016. **91**(6): p. 1228-1243.

Figures

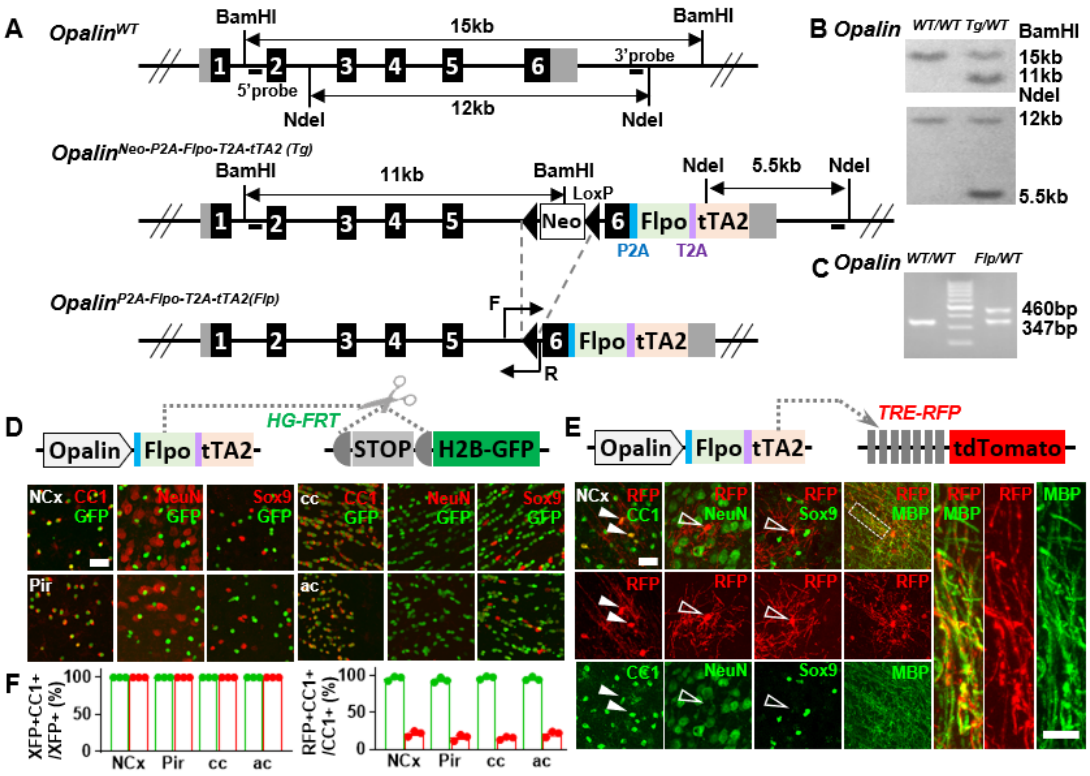


Figure 1. A new driver mouse for efficient and specific OL labeling

(A) Scheme for generating the *Opalin*$^{P2A\text{-}Flp\text{-}T2A\text{-}tTA2}$allele. (B) Southern blot confirmation of correctly targeted ES clone. (C) Genomic PCR to genotype F1 offspring. (D) OL labeling by Flp. (E) OL labeling by tTA2. High magnification images of the boxed region showing co-localization of RFP with MBP staining, which further demonstrated the myelination ability of labeled OLs. (F) Quantification of marker staining. Left panel quantified the labeling specificity. Both systems are highly specific, as shown by the complete co-localization with OL marker (CC1) and lack of co-staining by the neuronal marker (NeuN) or astrocyte marker (Sox9). Right panel quantified the labeling efficiency. Close to complete OL labeling was achieved by Flp-dependent H2B-GFP reporter in all analyzed regions, while sparser labeling with variable regional density was achieved by tTA2-dependent tdTomato reporter driven by TRE promoter. Scale bar: 50 μm in low magnification images, 5μm in high magnification images. Quantification: n=3. Dots represent data from individual mice.

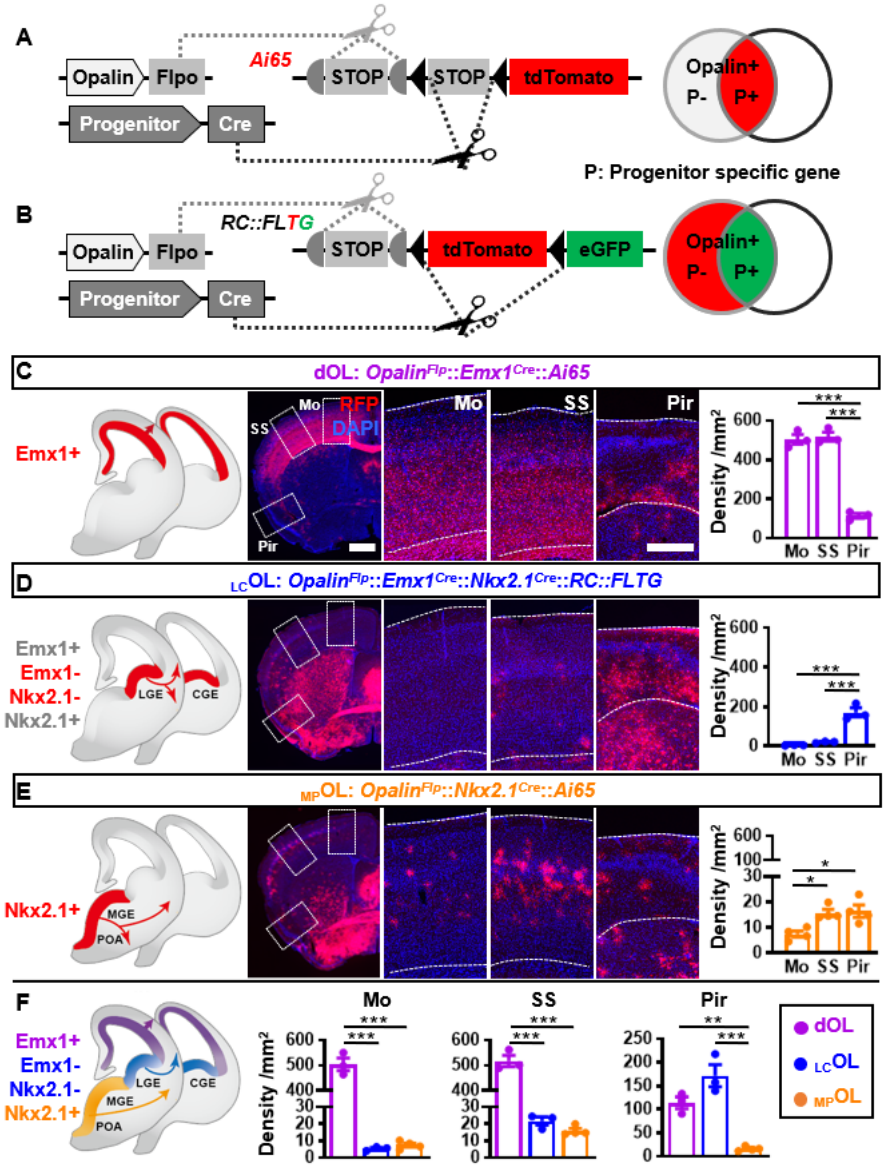


Figure 2. Combinatorial fate mapping of dOLs, $_{MP}$OLs and $_{LC}$OLs

(A) Strategy for intersectional labeling. Flp-AND-Cre labels OLs from Cre-expressing progenitors with RFP. (B) Strategy for simultaneous intersectional and subtractional labeling of OLs derived from complementary origins. Flp-NOT-Cre labels OLs from non-Cre-expressing progenitors with RFP, while Flp-AND-Cre labels OLs from Cre-expressing progenitors with GFP. (C-E) Intersectional labeling of dOLs in *Opalin*^{Flp}::*Emx1*^{Cre}::*Ai65,* subtractional labeling of $_{LC}$OLs in *Opalin*^{Flp}::*Emx1*^{Cre}::*Nkx2.1*^{Cre}::*RC::FLTG*, and intersectional labeling of $_{MP}$OLs in *Opalin*^{Flp}::*Nkx2.1*^{Cre}::*Ai65* and their cortical distribution. Schematics, representative images and quantification of RFP+ cell density in motor cortex (Mo), somatosensory cortex (SS) and piriform cortex (Pir) were displayed. (F) Differential contribution of OLs from the three embryonic origins to Mo, SS and Pir.

Scale bar: 1mm in low magnification images, 500 μm in high magnification images of the boxed area. Quantification: n=3 for dOLs and $_{LC}$OLs; n=4 for $_{MP}$OLs. Dots represent data from individual mice. **P < 0.01, ***P < 0.001. Error bar: S.E.M.

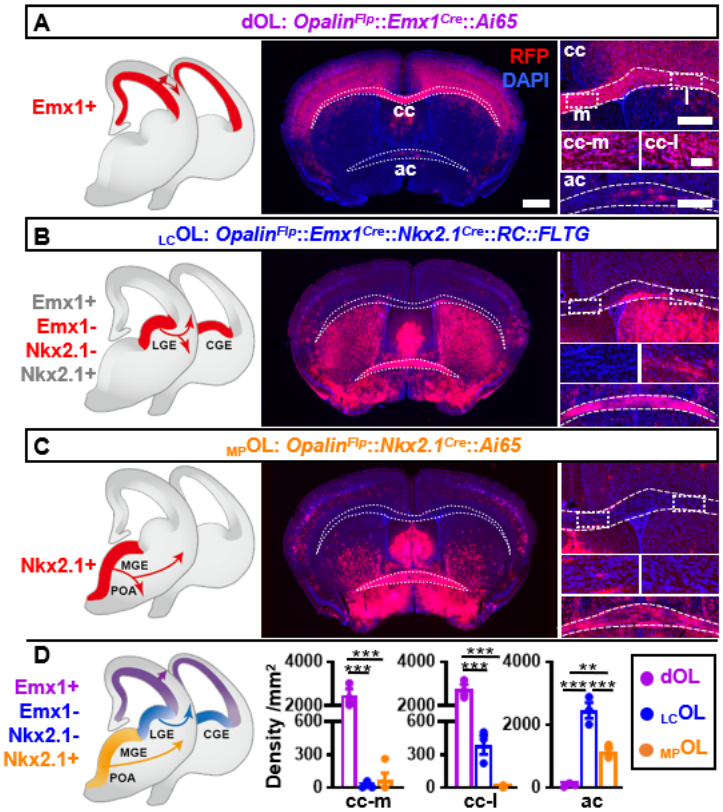


Figure 3. dOLs, $_{MP}$OLs and $_{LC}$OLs differentially contribute to the two major white matter tracts

(A-C) Schematics and representative images of dOLs, $_{MP}$OLs and $_{LC}$OLs in the two major commissure white matter tracts: corpus callosum(cc) and anterior commissure (ac). $_{MP}$OLs and $_{LC}$OLs preferentially reside in the medial and lateral cc (cc-m and cc-l), respectively. (D) Quantification of dOL, $_{MP}$OL and $_{LC}$OL densities in cc-m, cc-l and ac. Scale bar: 1mm in low magnification images, 500μm in cc and ac, 100 μm in high magnification images of the boxed area (cc-m and cc-l) in cc. Quantification: n=3 for dOLs and $_{LC}$OLs; n=4 for $_{MP}$OLs. Dots represent data from individual mice. **P < 0.01, ***P < 0.001. Error bar: S.E.M.

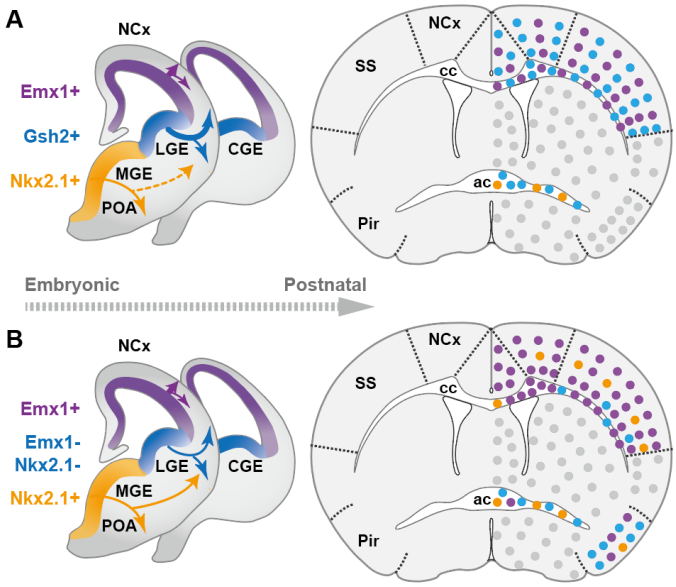
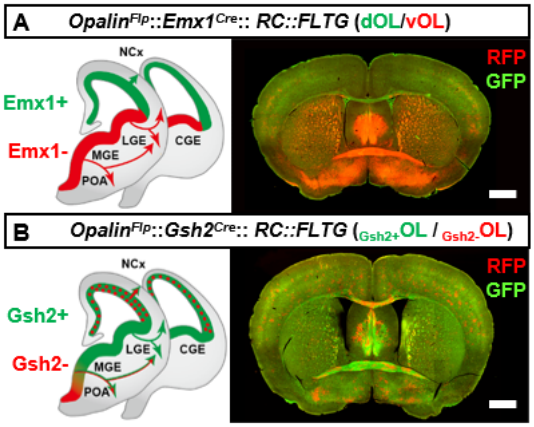


Figure 4. The classical and revised model of forebrain OL origins

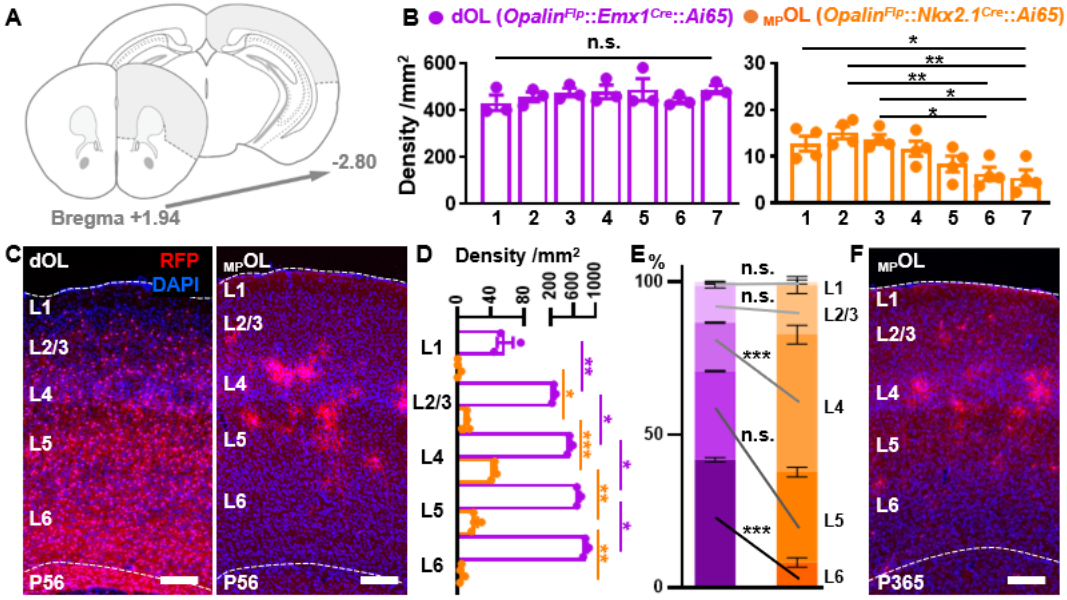
(A) In the classical model[2], OLs derived from MGE/POA (orange) were largely eliminated postnatally (thin dashed line), while those from LGE/CGE (blue) and dorsal origin (purple) survive at similar proportions (thick solid line). Therefore, neocortex (NCx) and corpus callosum (cc) contain comparable density of $_{LC}$OLs (blue dots) and dOLs (purple dots) and are devoid of $_{MP}$OLs (orange dots). (B) In the new model, NCx and cc mainly contains dOLs with very low contribution from the ventral origins. $_{LC}$OLs mainly contribute to Pir and ac. $_{MP}$OLs makes a small but sustained contribution to NCx, with a strong laminar preference towards layer 4 in somatosensory cortex (SS). In addition, dOLs and $_{MP}$OLs also make substantial contributions to Pir and ac, respectively. Grey dots indicate OLs in unanalyzed regions.

Supplementary information



Supplementary Figure 1. Simultaneous differential labeling of OLs derived from complementary embryonic origins

(A) Coronal sections showing GFP+ OLs from dorsal origin (dOLs) and RFP+ OLs from ventral origin (vOLs) in *Opalin*^{Flp}::*Emx1*^{Cre}::*RC::FLTG*. (B) Coronal sections showing GFP+ OLs derived from Gsh2+ progenitors ($_{Gsh2+}$OL) and RFP+ OLs derived from Gsh2- progenitors ($_{Gsh2-}$OL) in *Opalin*^{Flp}::*Gsh2*^{Cre}::*RC::FLTG*. Scale bar: 1mm.

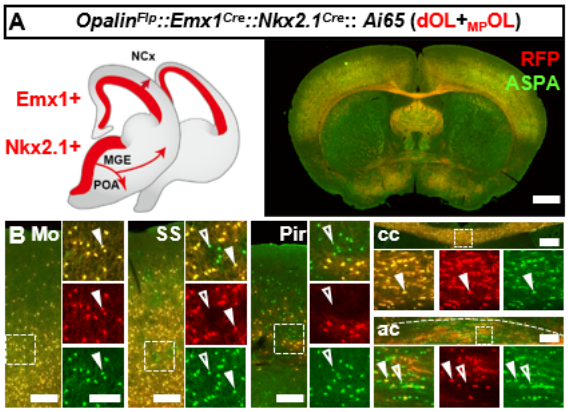


332

333 **Supplementary Figure 2. The distribution pattern of neocortical dOLs and $_{MP}$OLs**

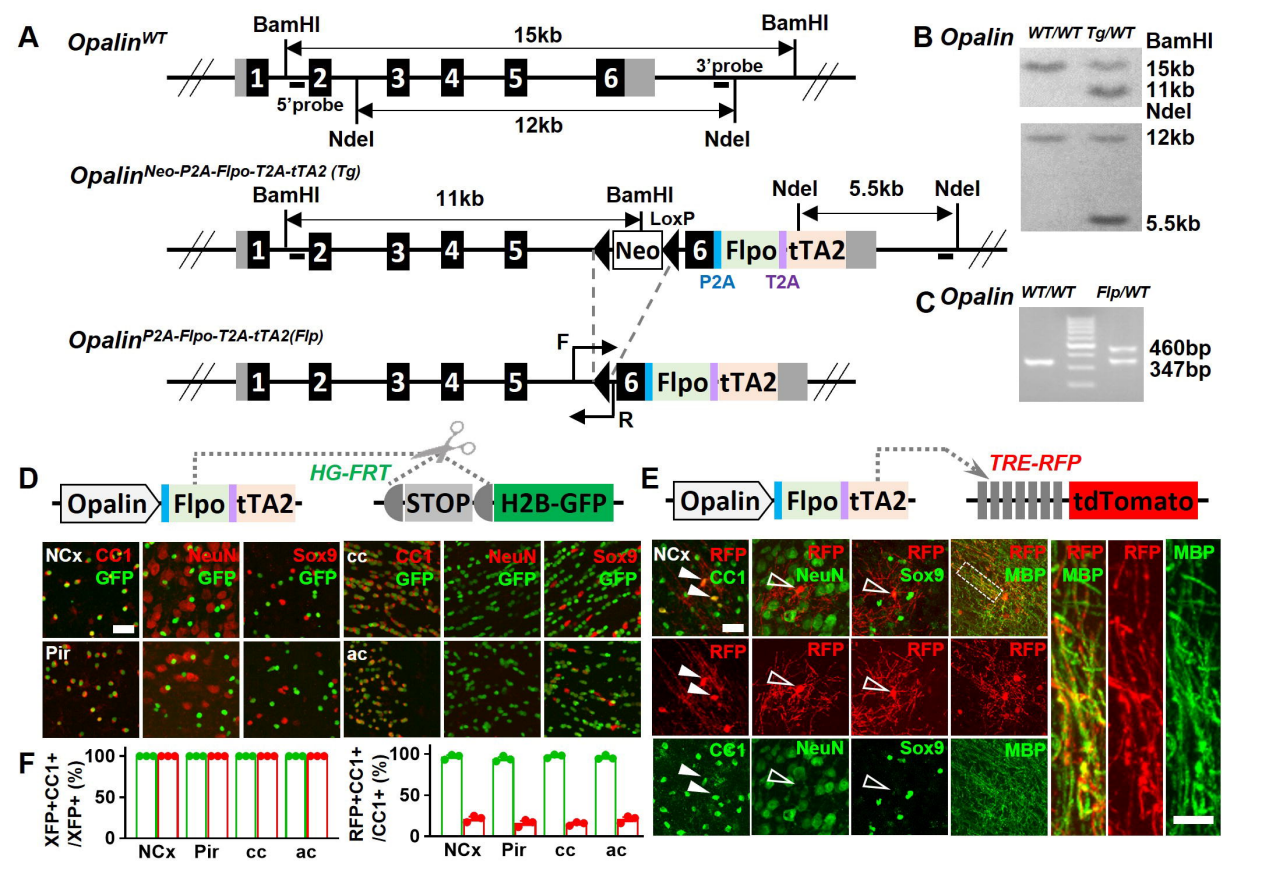
334 (A-B) Rostrocaudal distribution of neocortical dOLs and $_{MP}$OLs. (A) Within the range of
335 analysis, RFP+ OLs in neocortex of every fourth section were quantified. (B) Slices were
336 grouped into seven bins numbered rostrocaudally. Density of dOLs showed no significant
337 change across bins, while $_{MP}$OLs exhibited lower density in more caudal regions. (C-E) Laminar
338 distribution of dOLs and $_{MP}$OLs in somatosensory cortex (SS). (C) Representative images of
339 SS from P56 mice. (D) Distribution of dOLs and $_{MP}$OLs across 6 layers. More dOLs reside in
340 deeper layers, while $_{MP}$OLs are highly enriched in L4. (E) dOLs and $_{MP}$OLs exhibited significant
341 differences in their relative laminar distributions. (F) Representative image of SS from 1 year
342 old *Opalin*^{Flp}::*Nkx2.1*^{Cre}::*Ai65* mouse. Scale bar: 200 μm. Dots represent data from individual
343 mice. Error bar: S.E.M. *P < 0.05, **P < 0.01, ***P < 0.001.

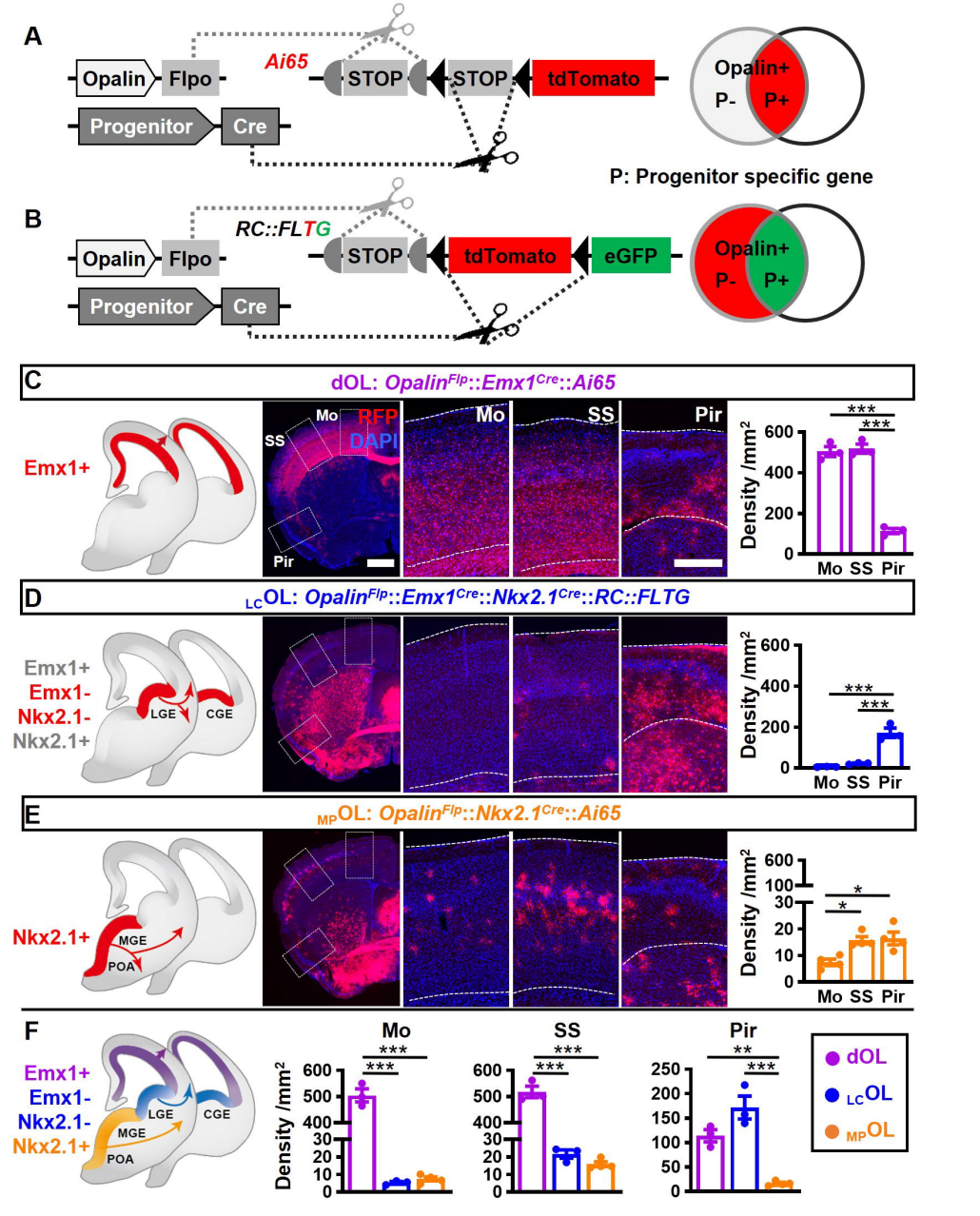
344

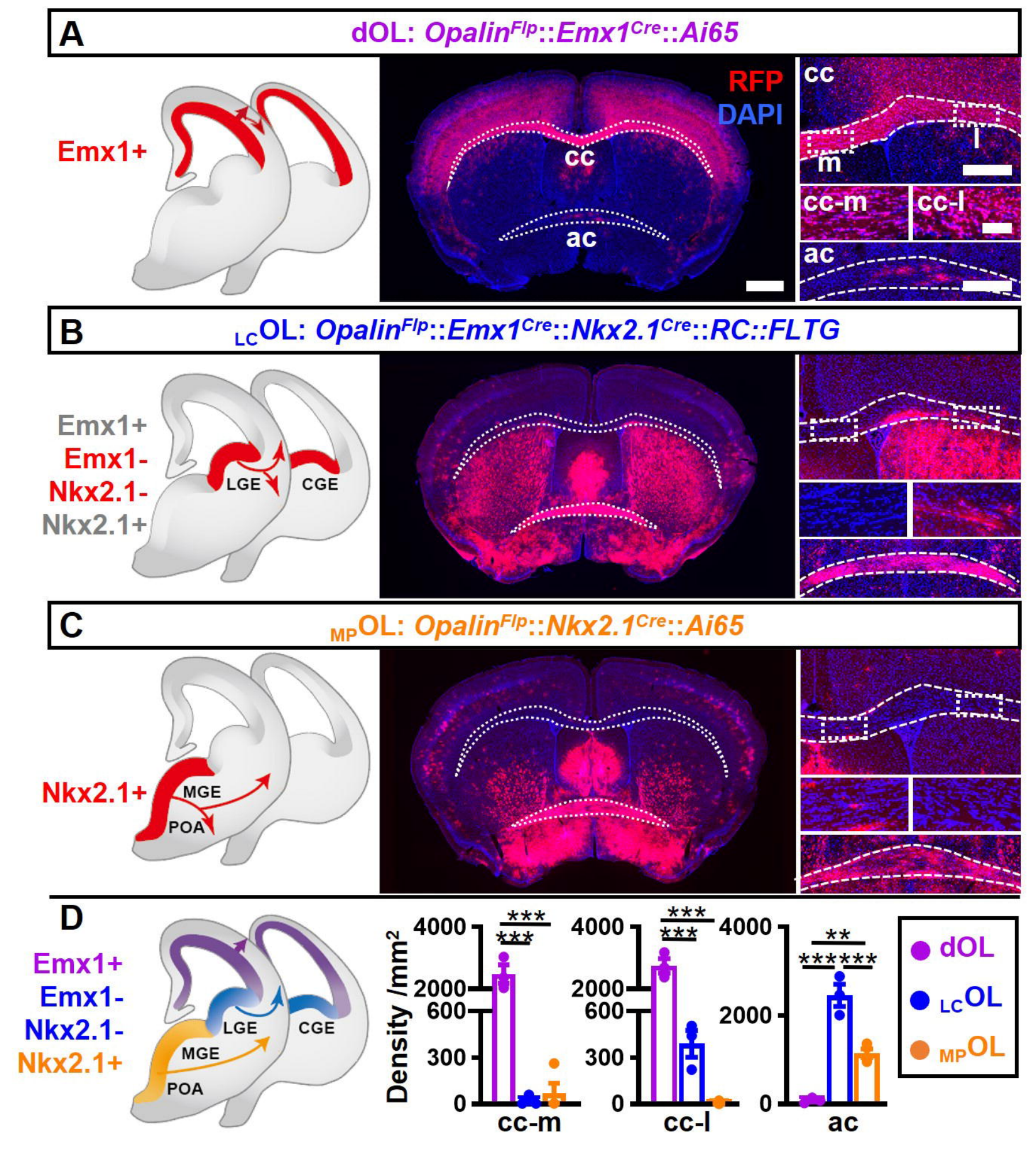


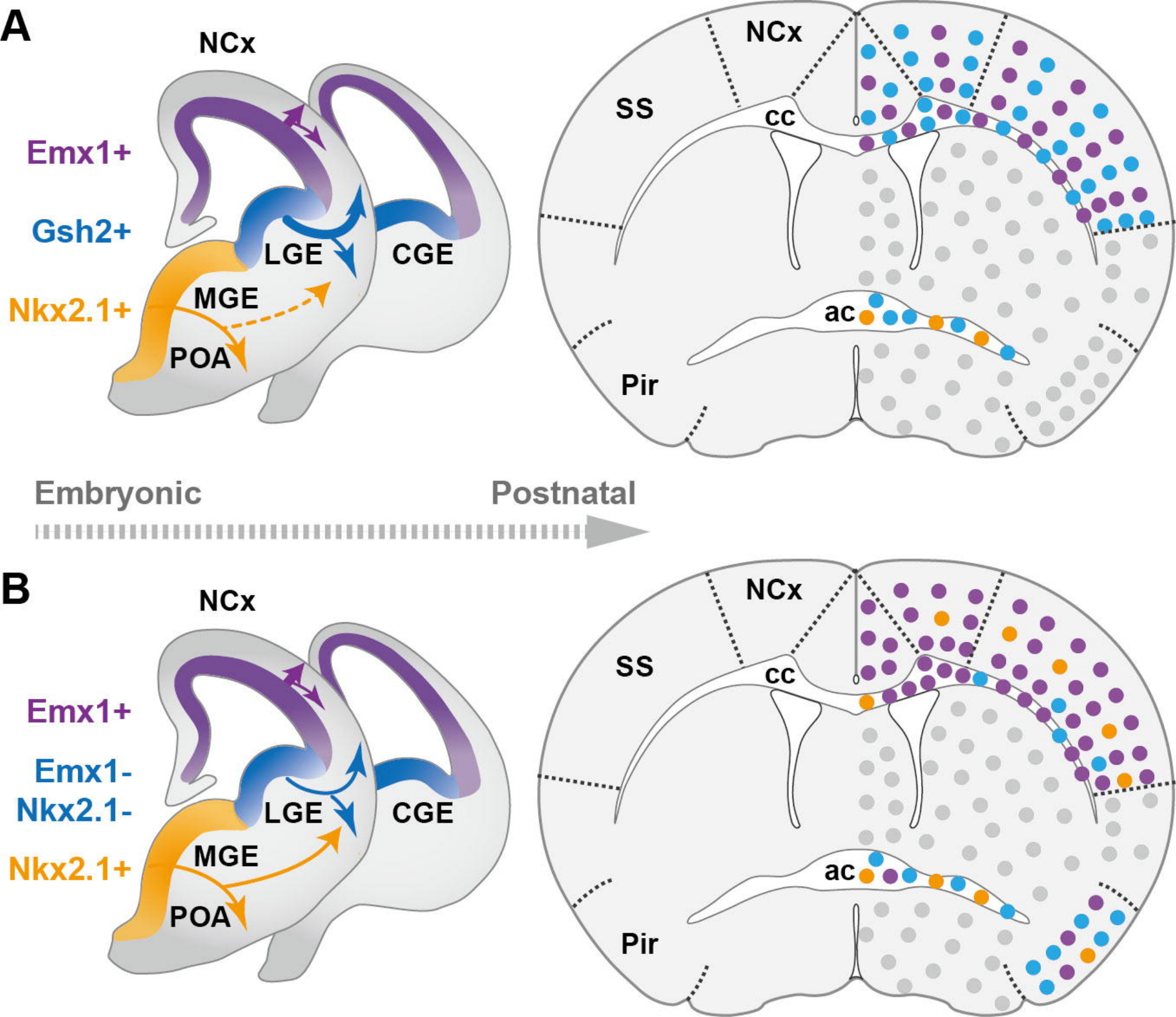
Supplementary Figure 3. Intersectional labeling of OLs derived from both dorsal origin and MGE/POA

(A) Coronal sections showing both dOLs and $_{MP}$OLs labeled by RFP in *Opalin*^{Flp}::*Emx1*^{Cre}::*Nkx2.1*^{Cre}::*Ai65*. (B) Higher magnification images showing ASPA+ RFP+ dOLs/$_{MP}$OLs (closed arrow heads) and ASPA+RFP- putative $_{LC}$OLs (open arrow heads).The latter cells were difficult to find in neocortical regions such as motor cortex (Mo) and somatosensory cortex (SS), and corpus callosum (cc), but were frequently encountered in piriform cortex (Pir) and anterior commissure (ac). Scale bar: 1mm in low magnification images, 200μm in Mo, SS, cc and ac, 10 μm in high magnification images of boxed area (all shared the same scale).

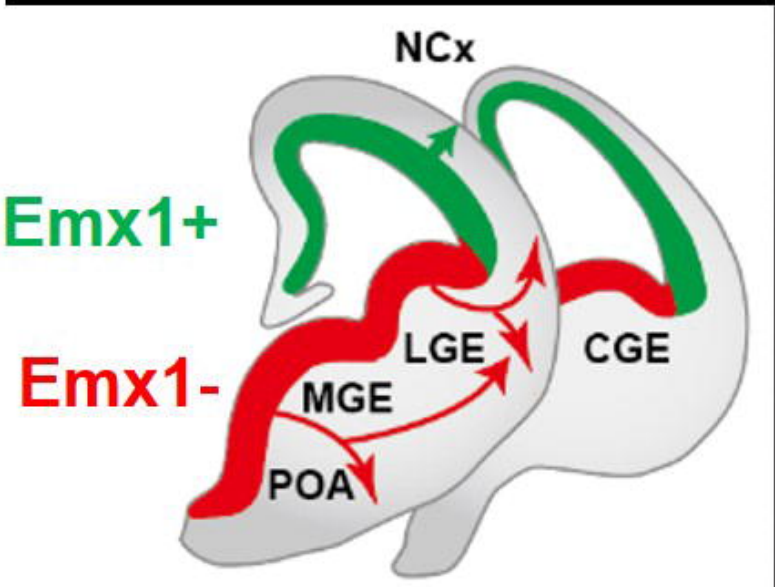

A
Opalin^WT
BamHI
15kb
BamHI
3'probe
5'probe
NdeI
12kb
NdeI
Opalin^Neo-P2A-Flpo-T2A-tTA2 (Tg)
BamHI
11kb
BamHI
LoxP
NdeI
5.5kb
NdeI
Neo
Flpo
tTA2
P2A
T2A
Opalin^P2A-Flpo-T2A-tTA2(Flp)
F
R
Flpo
tTA2
B Opalin
WT/WT Tg/WT
BamHI
15kb
11kb
NdeI
12kb
5.5kb
C Opalin
WT/WT
Flp/WT
460bp
347bp
D
Opalin
Flpo
tTA2
HG-FRT
STOP
H2B-GFP
E
Opalin
Flpo
tTA2
TRE-RFP
tdTomato
NCx
Pir
cc
ac
CC1
NeuN
Sox9
GFP
RFP
MBP
F
XFP+CC1+ /XFP+ (%)
RFP+CC1+ /CC1+ (%)
100
50
0
NCx
Pir
cc
ac


A
Opalin
Flpo
Progenitor
Cre
Ai65
STOP
STOP
tdTomato
Opalin+
P-
P+
P: Progenitor specific gene
B
RC::FLTG
Opalin
Flpo
Progenitor
Cre
STOP
tdTomato
eGFP
Opalin+
P-
P+
C
dOL: Opalin^Flp::Emx1^Cre::Ai65
Emx1+
Mo
SS
Pir
RFP
DAPI
Density /mm²
600
400
200
0
***
***
Mo SS Pir
D
LcOL: Opalin^Flp::Emx1^Cre::Nkx2.1^Cre::RC::FLTG
Emx1+
Emx1-
Nkx2.1-
Nkx2.1+
LGE
CGE
Density /mm²
600
400
200
0
***
***
Mo SS Pir
E
MPOL: Opalin^Flp::Nkx2.1^Cre::Ai65
Nkx2.1+
MGE
POA
Density /mm²
600
100
30
20
10
0
*
*
Mo SS Pir
F
Emx1+
Emx1-
Nkx2.1-
Nkx2.1+
LGE
CGE
MGE
POA
Mo
SS
Pir
Density /mm²
***
***
**
dOL
LcOL
MPOL

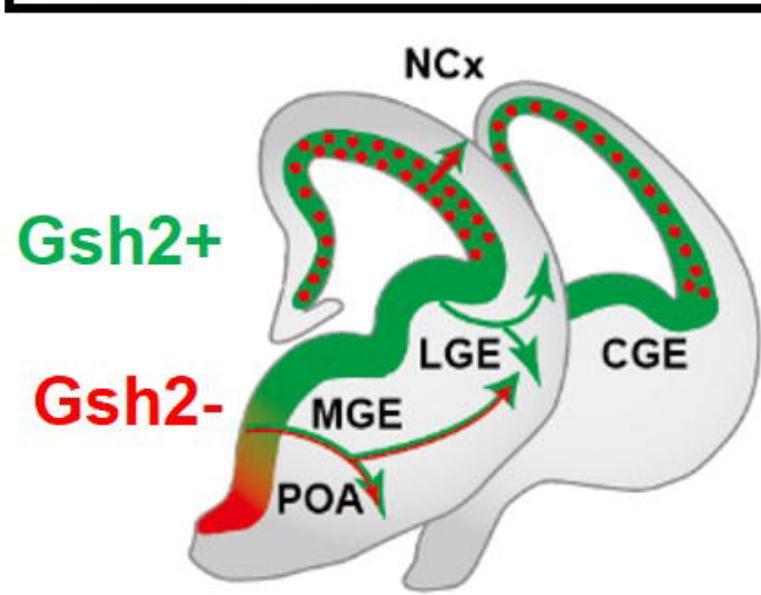

A
dOL: Opalin^Flp::Emx1^Cre::Ai65
Emx1+
RFP
DAPI
cc
ac
cc
m
l
cc-m
cc-l
ac
B
LCOL: Opalin^Flp::Emx1^Cre::Nkx2.1^Cre::RC::FLTG
Emx1+
Emx1-
Nkx2.1-
Nkx2.1+
LGE
CGE
C
MPOL: Opalin^Flp::Nkx2.1^Cre::Ai65
Nkx2.1+
MGE
POA
D
Emx1+
Emx1-
Nkx2.1-
Nkx2.1+
LGE
CGE
MGE
POA
Density /mm²
4000
2000
600
300
0
***
***
***
***
**
******
cc-m
cc-l
ac
dOL
LCOL
MPOL

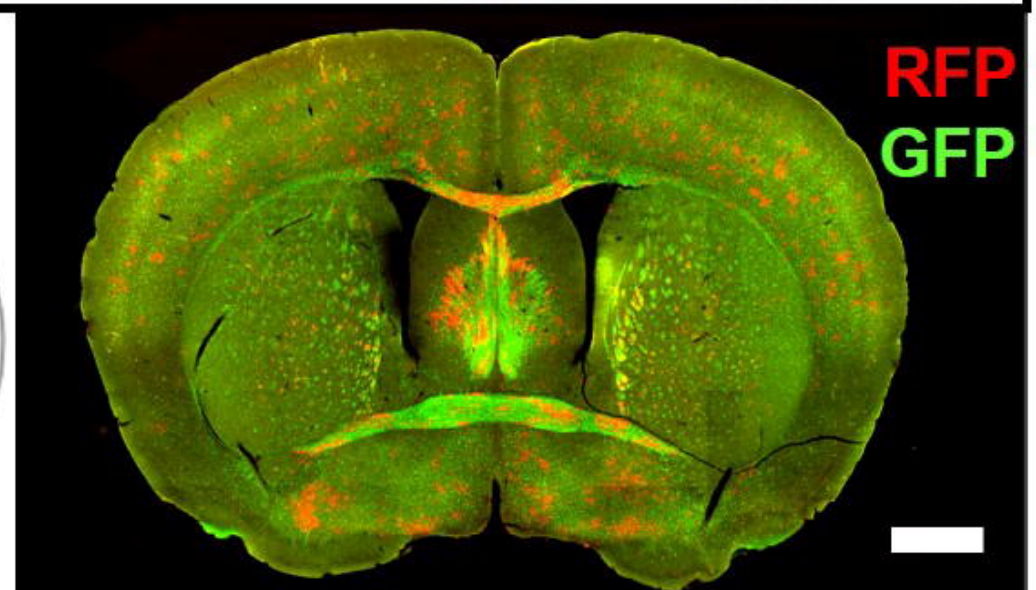

A
NCx
Emx1+
Gsh2+
Nkx2.1+
LGE
CGE
MGE
POA
NCx
SS
cc
ac
Pir
Embryonic
Postnatal
B
NCx
Emx1+
Emx1-
Nkx2.1-
Nkx2.1+
LGE
CGE
MGE
POA
NCx
SS
cc
ac
Pir

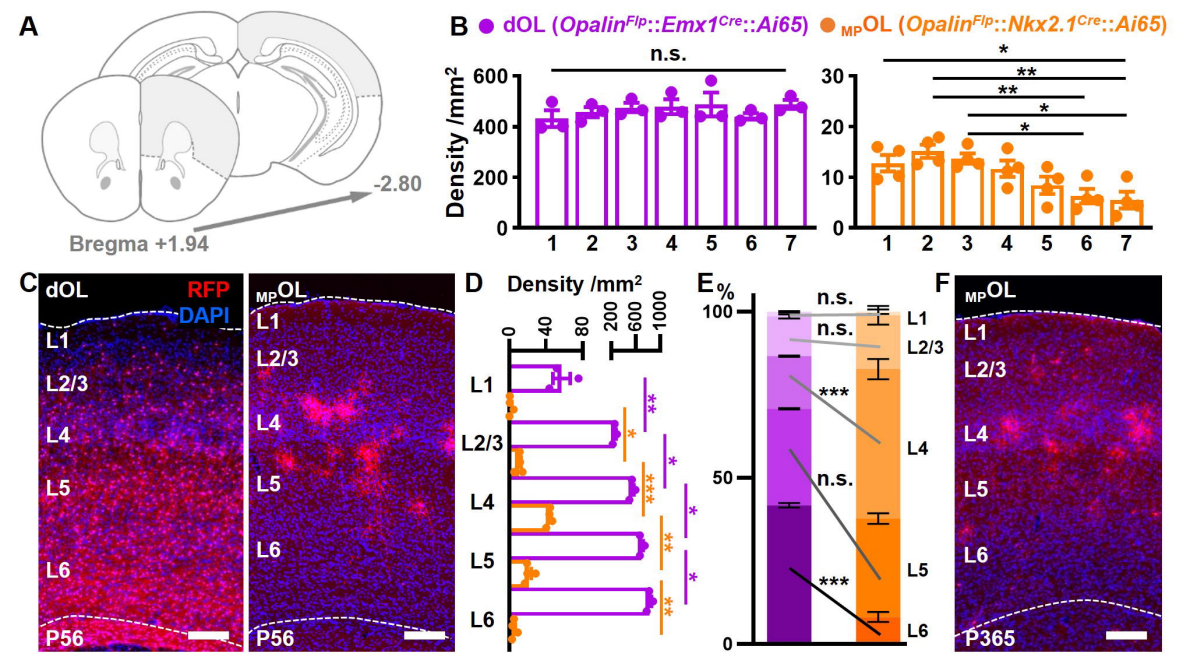
A *Opalin^{Flp}::Emx1^{Cre}:: RC::FLTG* (dOL/vOL)

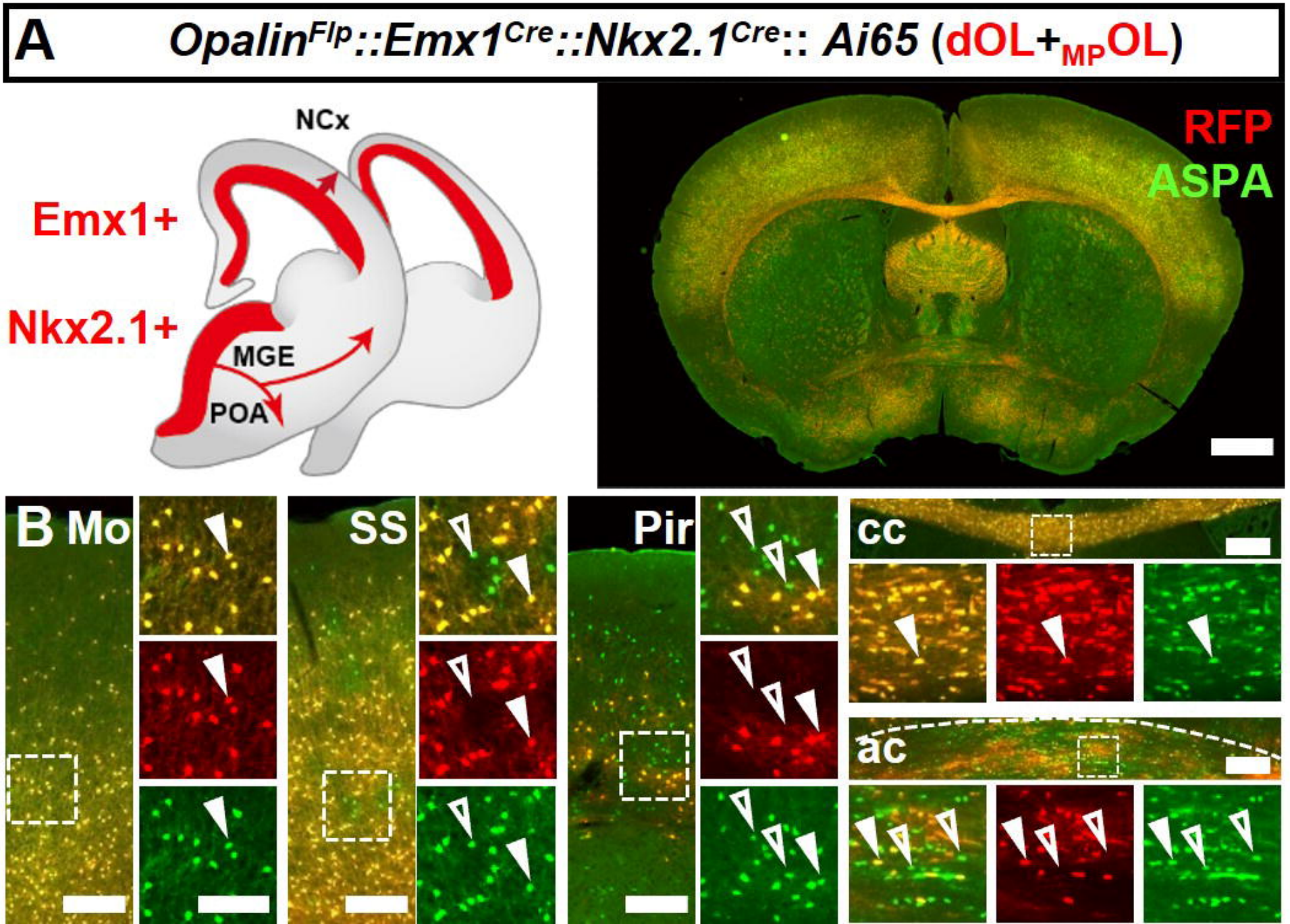


B *Opalin^{Flp}::Gsh2^{Cre}:: RC::FLTG* ($_{Gsh2+}$OL / $_{Gsh2-}$OL)






A
-2.80
Bregma +1.94
B
dOL (Opalin^Flp::Emx1^Cre::Ai65)
MPOL (Opalin^Flp::Nkx2.1^Cre::Ai65)
n.s.
Density /mm²
600
400
200
0
30
20
10
1 2 3 4 5 6 7
C
dOL
RFP
DAPI
MPOL
L1
L2/3
L4
L5
L6
P56
D
Density /mm²
0 40 80
200 600 1000
E
%
100
50
0
n.s.
***
F
P365


A
Opalin^Flp::Emx1^Cre::Nkx2.1^Cre:: Ai65 (dOL+MPOL)
NCx
Emx1+
Nkx2.1+
MGE
POA
RFP
ASPA
B Mo
SS
Pir
cc
ac