Article

Temporal controls over inter-areal cortical projection neuron fate diversity

🔴 Check for updates

Esther Klingler[1], Ugo Tomasello[1,7], Julien Prados[2,7], Justus M. Kebschull[3,6], Alessandro Contestabile[1], Gregorio L. Galiñanes[1], Sabine Fièvre[1], Antonio Santinha[4], Randall Platt[4], Daniel Huber[1], Alexandre Dayer[1,2], Camilla Bellone[1] & Denis Jabaudon[1,5 ✉]

Interconnectivity between neocortical areas is critical for sensory integration and sensorimotor transformations[1–6]. These functions are mediated by heterogeneous inter-areal cortical projection neurons (ICPN), which send axon branches across cortical areas as well as to subcortical targets[7–9]. Although ICPN are anatomically diverse[10–14], they are molecularly homogeneous[15], and how the diversity of their anatomical and functional features emerge during development remains largely unknown. Here we address this question by linking the connectome and transcriptome in developing single ICPN of the mouse neocortex using a combination of multiplexed analysis of projections by sequencing[16,17] (MAPseq, to identify single-neuron axonal projections) and single-cell RNA sequencing (to identify corresponding gene expression). Focusing on neurons of the primary somatosensory cortex (S1), we reveal a protracted unfolding of the molecular and functional differentiation of motor cortex-projecting ($\vec{M}$) ICPN compared with secondary somatosensory cortex-projecting ($\vec{S2}$) ICPN. We identify SOX11 as a temporally differentially expressed transcription factor in $\vec{M}$ versus $\vec{S2}$ ICPN. Postnatal manipulation of SOX11 expression in S1 impaired sensorimotor connectivity and disrupted selective exploratory behaviours in mice. Together, our results reveal that within a single cortical area, different subtypes of ICPN have distinct postnatal paces of molecular differentiation, which are subsequently reflected in distinct circuit connectivities and functions. Dynamic differences in the expression levels of a largely generic set of genes, rather than fundamental differences in the identity of developmental genetic programs, may thus account for the emergence of intra-type diversity in cortical neurons.

Somatosensory input from the periphery reaches the primary somatosensory cortex (S1) via the thalamus, from which information is conveyed to other cortical areas and subcortical targets for further processing[1–6,16,17]. Within the cortex, S1 is strongly connected with motor areas (M) and with the secondary somatosensory area (S2), forming parallel pathways for sensorimotor coordination (S1–M connections) and sensory discrimination (S1–S2 connections)[2,17]. In adults, M- and S2-projecting ICPN (here termed $\vec{M}$ and $\vec{S2}$ ICPN, respectively) have largely mutually exclusive projections to these two areas[2,3,18,19], but how this specificity emerges during development is unknown. As a first foray into this question, we examined the laminar distribution of $\vec{M}$ and $\vec{S2}$ ICPN in S1 using retrograde labelling at postnatal day (P)5 (when thalamocortical synapses are forming in S1[20,21]), P7 and P14 (when S1 inter-areal projections are largely established)[7,22,23] (Fig. 1a, Extended Data Fig. 1a, b). This approach revealed that largely mutually exclusive projections to M and to S2 emerge from P5 onwards. At that age, only very few retrogradely labelled $\vec{M}$ and $\vec{S2}$ ICPN were visible in superficial

layers (SL) (that is, most axons had not reached their targets; Fig. 1a, Extended Data Fig. 1a, b), probably reflecting the sequential birth and development of deep layer (DL) followed by SL neurons[24]. Hence, during the first postnatal week, cortical sensorimotor function may be driven initially by DL ICPN, whereas SL ICPN have a role at later developmental stages.

We then used multiplexed analysis of projections by sequencing (MAPseq) to map the target-specific development of multiple axonal projections in single ICPN in S1. MAPseq enables high-throughput single-cell reconstruction of axonal projections to multiple remote targets using anterograde transport of a barcoded RNA from the soma[16,17] (Extended Data Fig. 1c, Supplementary Note 1). Six functionally relevant targets were examined: ipsilateral M, S2, auditory (A) and visual cortices (V), contralateral S1 (C) and subcortical targets (Sub) (that is, striatum and thalamus) at P5, P7 and P14 (Fig. 1b, Extended Data Fig. 1d). MAPseq mapping revealed that at all ages M, S2, and C were the main intracortical targets. $\vec{M}$ and $\vec{S2}$ ICPN represented around

[1]Department of Basic Neurosciences, University of Geneva, Geneva, Switzerland. [2]Department of Psychiatry, Geneva University Hospital, Geneva, Switzerland. [3]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, NY, USA. [4]Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland. [5]Clinic of Neurology, Geneva University Hospital, Geneva, Switzerland. [6]Present address: Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA. [7]These authors contributed equally: Ugo Tomasello, Julien Prados. ✉e-mail: denis.jabaudon@unige.ch
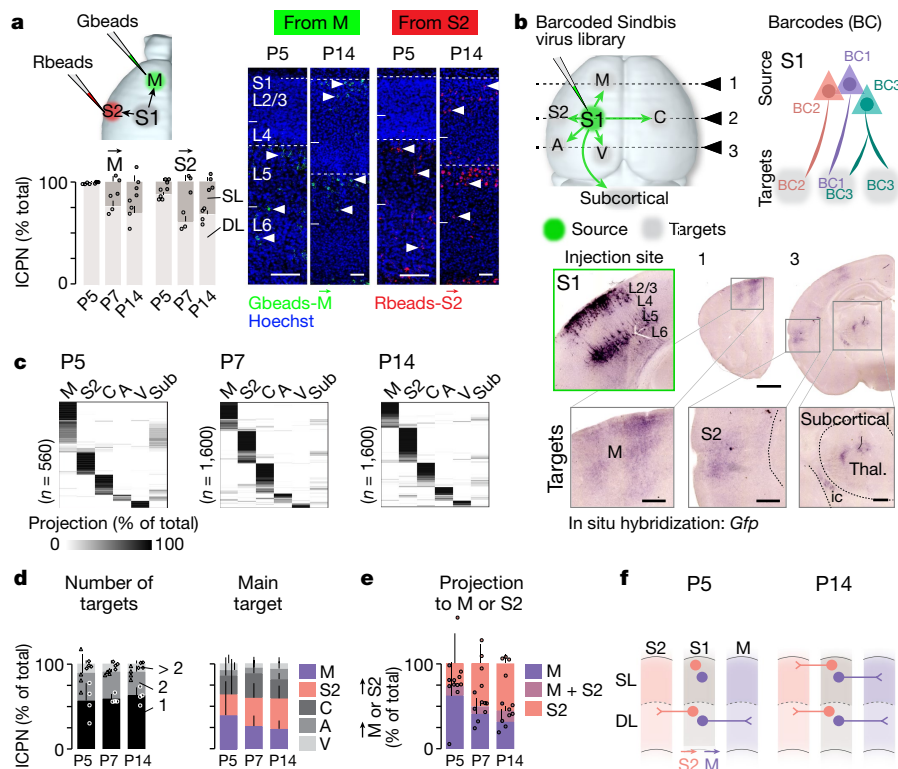
Fig. 1 | Postnatal emergence of intracortical connections from S1. a, Top left, retrograde labelling from M or S2 using green (Gbeads) or red (Rbeads) retrobeads. Right, Gbeads- and Rbeads-labelled $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN in S1 at P5 and P14 (arrowheads indicate retrogradely labelled ICPN). Bottom left, distribution of $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN in SL versus DL at P5, P7 and P14 (P5: $n = 4$, P7: $n = 3$, P14: $n = 4$ pups from two litters per target). **b**, Top, MAPseq mapping principle. Bottom, in situ hybridization at P14 shows transport of barcode-*Gfp* mRNA into axons at target sites. Green framed image: injection site. **c**, Multiplex axonal projections of single S1 ICPN at P5, P7 and P14 ($n = 4$ pups from 2 litters at each age). Numbers indicate number of ICPN. **d**, Number of targets (left) and main target (right) of ICPN. **e**, Projection to M or S2 or M + S2. **f**, Recapitulative diagram of S1 projections to M and S2 over postnatal development. Values are shown as mean ± s.e.m. (**a**, **d**, **e**). Scale bars, 100 μm (**a**), 1 mm (**b**, centre), 300 μm (**b**, bottom). A, auditory cortex; C, contralateral cortex; Thal, thalamus; ic, internal capsule; Sub, subcortical; V, visual cortex.

70% of S1 ICPN at all ages and around 30% of these neurons had multiple targets; $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN, however, shared essentially identical combinations of targets (Fig. 1c, d, Extended Data Figs. 1e–h, 2, Supplementary Table 1). As was the case with retrograde labelling, the small fraction of neurons projecting to both M and S2 remained stable across time, confirming that $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN are specified as such early on (Fig. 1e, Extended Data Fig. 1b). Together, these data indicate that S1 ICPN have a sparse and directed connectivity to M and S2 as soon as projections are detectable (Fig. 1f).

To investigate the developmental transcriptional dynamics of $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN, we combined MAPseq mapping with single-cell RNA sequencing (scRNA-seq) of S1 neurons, thereby linking the transcriptional identity of barcode-identified neurons with their corresponding projections, and called this approach ConnectID (Fig. 2a, Extended Data Figs. 3, 4, Supplementary Note 2). We performed ConnectID at P7, P9, and P14 (at P5 only few SL ICPN axons have reached M and S2; Fig. 1a), enabling us to identify diverse subtypes of ICPN, including $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN, and their corresponding transcriptional identities (Fig. 2b). Whereas ICPN molecular identity clearly related to neuronal age (Fig. 2b, left), it did not seem to correlate with connectivity (Fig. 2b, right). Thus, as is the case in adulthood[15], despite their distinct connectivities, $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN cannot be identified as distinct molecular cell types during postnatal development (Extended Data Fig. 5).

Given the molecular overlap between $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN, we hypothesized that ICPN diversity could emerge from largely generic transcriptional programs unfolding at different paces rather than from fundamental differences in the nature of these programs. To detect potential temporal differences in the expression levels of corresponding gene sets

in $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN, we mapped their developmental expression using pseudotime alignment of single cells[25] (Extended Data Fig. 6a). This approach identified three successive waves of gene expression in $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN: the first wave was early-onset (P7), genes in the second wave had a sustained expression across postnatal ages, and the third wave was late-onset (P14) (Fig. 2c, top, Extended Data Fig. 6b). Genes whose ontologies were associated with neuron differentiation showed prolonged expression in $\overrightarrow{M}$ ICPN compared with $\overrightarrow{S2}$ ICPN (Fig. 2c, bottom, Extended Data Fig. 6c, Supplementary Table 2). This transcriptional shift was no longer visible in wave 3 genes, suggesting a transient process occurring during the first two weeks of postnatal life. Temporally shifted gene sets included genes with ontologies related to axon development and dendrites (Fig. 2d, Extended Data Fig. 6c–f, Supplementary Table 3), suggesting potential functional relevance. Supporting such a possibility, $\overrightarrow{M}$ ICPN showed delayed axon extension and dendritic tree maturation compared to $\overrightarrow{S2}$ ICPN (Fig. 2e). Thus, compared with $\overrightarrow{S2}$ ICPN, $\overrightarrow{M}$ ICPN have delayed transcriptional programs, resulting in a comparatively more protracted morphological differentiation.

We next examined whether the temporal differences in gene expression and anatomical features of $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN identified above would affect the functional maturation of S1-to-M and S1-to-S2 connectivity. For this purpose, we used Cux2-*cre* × GCaMP6s transgenic pups (*Cux2* is expressed in SL ICPN)[26] and performed in vivo wide-field imaging of cortical activation upon whisker pad, forelimb and hindlimb stimulations (Fig. 2f, Extended Data Fig. 7a). We designed the imaging setup such that S2, S1 and M regions could be monitored in the same field of view, and performed recordings at P9—when pups start exploring their environment—and at P15—when exploration is well established[27,28]. We
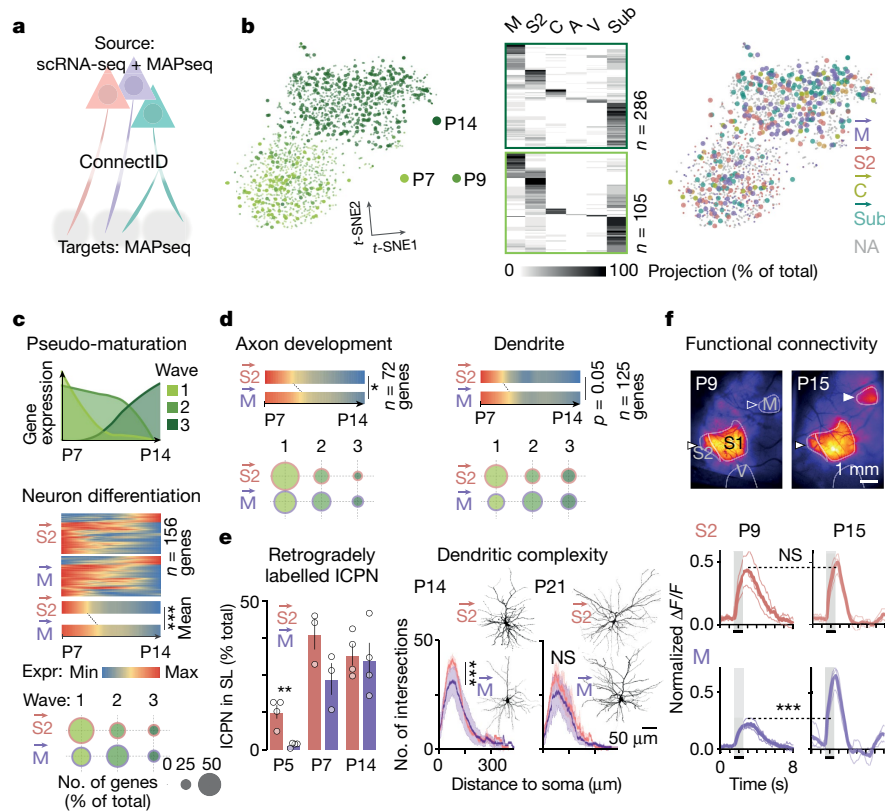
Fig. 2 | M and S2 ICPN have distinct maturation paces. a, ConnectID links transcriptome with projectome in single neurons. **b**, Left, *t*-distributed stochastic neighbor embedding (*t*-SNE) representation of the scRNA-seq ConnectID (P7, P9 and P14; *n* = 2,134 neurons) and retrogradely labelled (Rbeads) M and S2 ICPN (P9; *n* = 133 M, *n* = 183 S2 ICPN) datasets. Large dots, neurons with barcode identified in the target(s) (*n* = 391 ConnectID neurons). Centre, multiplex axonal projections of ConnectID neurons. Right, main target of ConnectID neurons (large dots) and Rbeads M and S2 ICPN (small dots). Grey dots, non-assigned (NA) neurons (that is, those for which no barcode was retrieved in any target). **c**, Top, gene-expression kinetics revealing three transcriptional waves. Bottom, individual gene-expression kinetics and number of genes related to neuron differentiation in each wave for S2 and M ICPN. Expr, expression; max, maximum; min, minimum. **d**, Mean expression

kinetics for axon development (left) and dendrite (right)-related genes and number of genes in each wave. **e**, Left, proportion of S2 and M ICPN in SL (P5, P14: *n* = 4 pups; P7: *n* = 3 pups). Right, dendrite tree complexity revealed by Sholl analysis of S2 and M SL ICPN patch-filled with biocytin at P14 (*n* = 13 M from 8 pups; *n* = 9 S2 ICPN from 6 pups) and P21 (*n* = 5 M from 3 pups; *n* = 5 S2 ICPN from 3 pups). Repeated-measures analysis of variance (ANOVA) with Geisser–Greenhouse correction. **f**, Calcium wide-field imaging upon sensory stimulation in anaesthetized Cux2-*cre* × GCaMP6s pups at P9 and P15 (*n* = 3 pups from 3 litters at each age). Representative images (top) and quantifications of responses upon whisker pad stimulation (bottom). Filled arrowhead indicates response. Empty arrowhead indicates absence of response. Grey, stimulation. ∆*F/F* traces are normalized to S1 responses. Two-way ANOVA with Sidak's multiple comparisons. NS, not significant.

found that at P9, whereas S2 activation had already reached the response strength they have at P15, this was not the case for M activation, which was much weaker. This reveals a sequential functional development of S1-to-S2 pathways followed by S1-to-M pathways during the first two postnatal weeks (Fig. 2f, Extended Data Fig. 7b, c). Together, these data indicate that M ICPN have a delayed molecular, anatomical and functional development compared with S2 ICPN. Thus, even within a single area and layer, closely related neuronal subtypes show distinct transcriptional and associated functional postnatal paces of maturation.

We next tested whether manipulation of temporally regulated transcriptional programs would affect the inter-areal connectivity of S1. For this purpose, we calculated the distance between M and S2 ICPN gene expression along pseudo-maturation for all wave 1 genes (that is, genes that are highly expressed when axon development occurs) (Fig. 2d, Supplementary Table 4). Focusing on transcription factors and axon development-related transcripts, we identified *Sox11* as an interesting candidate, as this transcription factor controls axon guidance in retinal ganglion cells[29] and is specifically expressed in SL neurons postnatally (Fig. 3a,b). Expression levels of Sox11 transcript and protein decrease with time between P7 and P14, but in S2 ICPN, *Sox11* is initially highly expressed and rapidly decreases, whereas in M ICPN,

initial expression is weaker (Fig. 3b); this was confirmed using immunohistochemistry following retrograde labelling from M and S2 (Extended Data Fig. 8a). To examine a potential role for SOX11 in postnatal ICPN development, we specifically overexpressed the *Sox11* gene in L2/3 of S1 by combining in utero electroporation of a Cre-dependent plasmid at embryonic day (E)15.5 (to target L2/3 ICPN) with stereotaxic injection of the adeno-associated virus AAV-pCMV-Cre in S1 at P0 (Fig. 3c, Extended Data Fig. 8b, c). Combined with retrograde labelling from M and S2, this approach revealed a near complete absence of S1-to-M projections in SOX11-overexpressing ICPN (Fig. 3d). This decrease in labelled M ICPN did not reflect cell death or mismigration, as demonstrated by similar numbers of SOX11-expressing L2/3 neurons at P17 compared with controls (Extended Data Fig. 8d). Instead, this result could reflect (1) the re-routing of M ICPN axons to other targets or (2) decreased axonal extension of these neurons. In relation to (1), quantification of anterograde axonal signal did not reveal increased projections to S2 (in fact S1-to-S2 projections were also decreased, albeit less markedly; Fig. 3e, Extended Data Fig. 8e). Similarly, projections to the contralateral cortex or subcortical projections were not increased (Extended Data Fig. 8e). Regarding (2), projections within the white matter were not detectably affected
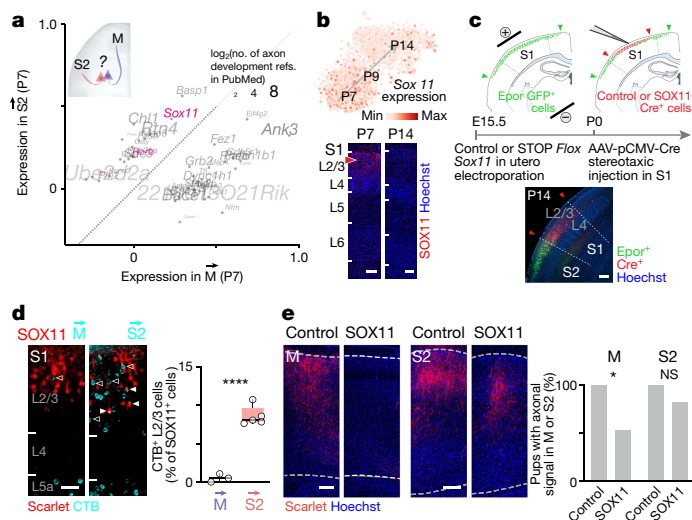
Fig. 3 | Postnatal SOX11 expression levels regulate S1 sensorimotor connectivity. a, Differential gene expression between $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN at P7, when axons are reaching their target. Total number for P7, P9 and P14 ConnectID neurons (used to calculate distance between $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN gene expression along pseudo-maturation for all wave 1 genes) $n = 114$ $\overrightarrow{M}$; $n = 76$ $\overrightarrow{S2}$ ICPN. The size of each gene symbol represents the number of PubMed abstracts with the term 'axon development' (Methods). Only genes with more than one PubMed abstract related to axon development are represented here. Transcription factors are shown in pink. **b**, *Sox11* is dynamically expressed during postnatal development (top), as confirmed at the protein level (bottom). **c**, Space- and time-restricted overexpression of *Sox11* using in utero electroporation at E15.5 combined with AAV-pCMV-*Cre* stereotaxic injection in putative S1 at P0. **d**, Retrograde labelling from M or S2 at P15 and quantification of S1 L2/3 SOX11-overexpressing (SOX11) $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN somas at P17 (M: $n = 3$; S2: $n = 5$ pups from 2 litters). Box plots indicate median ± s.d. and interquartile range (unpaired *t*-test: $P < 0.0001$). **e**, S1 Cre[+] axons expressing Scarlet in M and S2 at P17. Left, representative images of Scarlet[+] axons in each target. Right, fluorescence-based quantification of axonal projections ($n = 13$ control; $n = 11$ SOX11 pups from 4 litters per condition; Fisher's exact test, M: $P = 0.0149$; S2: $P = 0.1993$). Scale bars, 100 μm (**b**–**e**). CTB, Alexa Fluor 647-conjugated cholera toxin B; epor., electroporation; L, layer.

(Extended Data Fig. 8e), suggesting that axons were able to project beyond S1, but then did not invade their cortical target. Accordingly, axon growth was not perturbed in SOX11-overexpressing ICPN in vitro (Extended Data Fig. 8f), suggesting that SOX11 does not impair axon extension per se but instead probably regulates the response to target-derived molecular cues[29]. Consistent with an essential role of SOX11 in axon guidance[29], conditional ablation from P0 in S1 ICPN using an inducible CRISPR–Cas9 *Sox11* single guide RNA (sgRNA) plasmid construct led to an overall decrease in intracortical projections from S1 (Extended Data Fig. 8g). $\overrightarrow{M}$ ICPN projections were less affected by this manipulation, consistent with lower baseline levels of expression of SOX11 in these neurons. Thus, SOX11 is necessary for ICPN axons to invade their cortical targets. Overall, these data reveal that postnatal SOX11 expression levels are critical for proper inter-areal sensorimotor circuit assembly.

We next examined the behavioural consequences of disrupted sensorimotor connectivity following unilateral overexpression of SOX11 in S1 (Fig. 4, Extended Data Fig. 9). Spatial exploration was strongly impaired in S1 SOX11-overexpressing mice when tested in an open-field arena (Fig. 4a, b, Extended Data Fig. 9a, b). This probably reflects a genuine sensorimotor impediment rather than an anxiety-related behaviour, since the time spent in the centre of the open-field was similar to that of control mice (Extended Data Fig. 9b). Tactile spatial exploration was strongly asymmetric in SOX11-overexpressing mice, which spent most of their time exploring the wall contralateral to the targeted S1
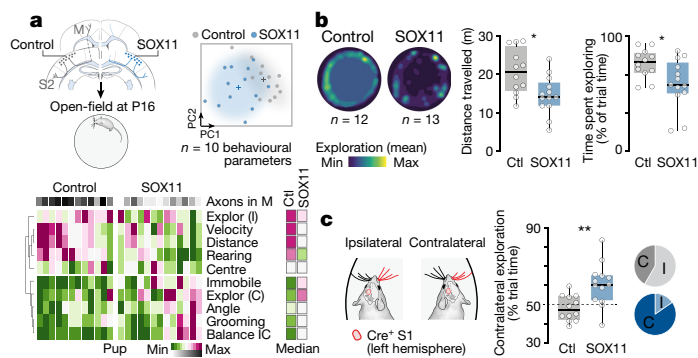
Fig. 4 | Subtype-specific ICPN connectivity underlies specific features of sensorimotor exploration. a, Top left, control (ctl) and SOX11-overexpressing (SOX11) P16 pups were monitored in an open-field arena for 10 min. Principal component analysis (top right) and heat map (bottom) of main exploratory parameters. C, contralateral; Explor, exploration; I, ipsilateral. **b**, Left, mean trajectories for all control and SOX11-overexpressing monitored pups. Right, quantification of distance travelled (unpaired *t*-test, $P = 0.0137$) and time spent exploring (unpaired *t*-test, $P = 0.0227$; $n = 12$ control pups, $n = 13$ SOX11 pups). **c**, Time spent exploring the open-field wall ipsilateral and contralateral to electroporated S1. Pie charts represent the proportion of pups spending time close to the ipsilateral or the contralateral wall ($n = 12$ control pups, $n = 13$ SOX11 pups; paired *t*-test, control: $P = 0.360$; SOX11: $P = 0.0073$; majority of time spent in ipsilateral versus contralateral wall by pup: Fisher's exact test, $P = 0.0414$). Box plots indicate median ± s.d. and interquartile range (**b**, **c**).

(Fig. 4c). Overall, behavioural modifications correlated with the extent of disrupted S1 connectivity (Extended data Fig. 9c, d), suggesting a direct causal relationship with sensorimotor exploration.

Our results reveal that a strategy for transcriptionally similar neurons to be wired differently is for them to mature at different paces. Specifically, dynamic controls over the levels of expression of shared genes may provide a parsimonious mechanism to generate a spectrum of projections across largely transcriptionally homogeneous ICPN subtypes during development. The repurposing of 'generic' transcriptional programs tweaked by distinct temporal regulators may have been selected for its metabolic fitness, enabling rapid and robust wiring of inter-areal networks required for exploratory behaviour. Fine-grained temporal dose-dependent controls may allow interaction with distinct extracellular molecular gradients, ultimately affecting axon guidance. Supporting such a process, differences in the timing of largely shared transcriptional programs underlie callosal versus anterior commissural crossing of interhemispheric axons in eutherian mouse versus marsupial dunnart[30]. In addition to these temporal controls, at earlier stages of development (including in the embryo) distinct and deterministic molecular programs may have roles within $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN to initiate axon target specificity and induce the temporal shift identified here.

In relation to behaviour, active environmental exploration emerges progressively during the first two weeks of life[27,28], which corresponds to the protracted maturation of S1-to-M connectivity identified here. Hence, the staggered development of $\overrightarrow{S2}$ ICPN and $\overrightarrow{M}$ ICPN may set the stage for a staggered development of sensory skills (for example, those required initially for breastfeeding) followed by motor skills (for example, those required to roam) and help establish sequential associations between sensory input and the coordinated motor output required for motor learning.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information,

1. Mao, T. et al. Long-range neuronal circuits underlying the interaction between sensory and motor cortex. *Neuron* **72**, 111–123 (2011).
2. Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L. & Helmchen, F. Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature* **499**, 336–340 (2013).
3. Yamashita, T. & Petersen, C. C. Target-specific membrane potential dynamics of neocortical projection neurons during goal-directed behavior. *eLife* **5**, e15798 (2016).
4. Glickfeld, L. L., Andermann, M. L., Bonin, V. & Reid, R. C. Cortico-cortical projections in mouse visual cortex are functionally target specific. *Nat. Neurosci.* **16**, 219–226 (2013).
5. Guo, Z. V. et al. Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**, 179–194 (2014).
6. Kwon, S. E., Yang, H., Minamisawa, G. & O'Connor, D. H. Sensory and decision-related activity propagate in a cortical feedback loop during touch perception. *Nat. Neurosci.* **19**, 1243–1249 (2016).
7. Fame, R. M., MacDonald, J. L. & Macklis, J. D. Development, specification, and diversity of callosal projection neurons. *Trends Neurosci.* **34**, 41–50 (2011).
8. Sorensen, S. A. et al. Correlated gene expression and target specificity demonstrate excitatory projection neuron diversity. *Cerebral Cortex* **25**, 433–449 (2015).
9. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
10. Aronoff, R. et al. Long-range connectivity of mouse primary somatosensory barrel cortex. *Eur. J. Neurosci.* **31**, 2221–2233 (2010).
11. Zingg, B. et al. Neural networks of the mouse neocortex. *Cell* **156**, 1096–1111 (2014).
12. Gong, H. et al. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. *Nat. Commun.* **7**, 12142 (2016).
13. Chen, X. et al. High-throughput mapping of long-range neuronal projection using in situ sequencing. *Cell* **179**, 772–786.e19 (2019).
14. Peng, H. et al. Morphological diversity of single neurons in molecularly defined cell types. *Nature* **598**, 174–181 (2021).
15. Yao, Z. et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).
16. Kebschull, J. M. et al. High-throughput mapping of single-neuron projections by sequencing of barcoded rna. *Neuron* **91**, 975–987 (2016).
17. Han, Y. et al. The logic of single-cell projections from visual cortex. *Nature* **556**, 51–56 (2018).
18. Yamashita, T. et al. Membrane potential dynamics of neocortical projection neurons driving target-specific signals. *Neuron* **80**, 1477–1490 (2013).
19. Yamashita, T. et al. Diverse long-range axonal projections of excitatory layer 2/3 neurons in mouse barrel cortex. *Front. Neuroanat.* **12**, 33 (2018).
20. Rebsam, A., Seif, I. & Gaspar, P. Refinement of thalamocortical arbors and emergence of barrel domains in the primary somatosensory cortex: a study of normal and monoamine oxidase a knock-out mice. *J. Neurosci.* **22**, 8541–8552 (2002).
21. Molnár, Z., Kurotani, T., Higashi, S., Yamamoto, N. & Toyama, K. Development of functional thalamocortical synapses studied with current source-density analysis in whole forebrain slices in the rat. *Brain Res. Bull.* **60**, 355–371 (2003).
22. Wang, C. L. et al. Activity-dependent development of callosal projections in the somatosensory cortex. *J. Neurosci.* **27**, 11334–11342 (2007).
23. Hand, R. A., Khalid, S., Tam, E. & Kolodkin, A. L. Axon dynamics during neocortical laminar innervation. *Cell Rep.* **12**, 172–182 (2015).
24. Klingler, E. et al. A Translaminar genetic logic for the circuit identity of intracortically projecting neurons. *Curr. Biol.* **29**, 332–339.e5 (2019).
25. Telley, L. et al. Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364**, eaav2522 (2019).
26. Madisen, L. et al. Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. *Neuron* **85**, 942–958 (2015).
27. Quairiaux, C., Megevand, P., Kiss, J. Z. & Michel, C. M. Functional development of large-scale sensorimotor cortical networks in the brain. *J. Neurosci.* **31**, 9574–9584 (2011).
28. van der Bourg, A. et al. Layer-specific refinement of sensory coding in developing mouse barrel cortex. *Cereb. Cortex* **84**, 401–416 (2016).
29. Kuwajima, T., Soares, C. A., Sitko, A. A., Lefebvre, V. & Mason, C. SoxC transcription factors promote contralateral retinal ganglion cell differentiation and axon guidance in the mouse visual system. *Neuron* **93**, 1110–1125.e5 (2017).
30. Paolino, A. et al. Differential timing of a conserved transcriptional network underlies divergent cortical projection routes across mammalian brain evolution. *Proc. Natl Acad. Sci. USA* **117**, 10554–10564 (2020).

Article

Methods

Mouse strains

C57Bl6/J (for MAPseq, ConnectID, $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN retrograde labelling) and CD1 (for in utero electroporation and *Sox11* overexpression phenotype analyses including axon anatomy and behaviour tests) male and female mice from Charles River Laboratory were used. For wide-field calcium imaging, Ai94(TITL-GCaMP6s) mice from Jackson Laboratories (B6;129S-Igs7tm94.1(tetO-GCaMP6s)Hze/J) were crossed with Cux2-IRES-Cre mice (B6(Cg)-Cux2tm1.1(cre)Mull/Mmmh). The experimental procedures described here were conducted in accordance with the Swiss laws and previously approved by the Geneva Cantonal Veterinary Authority. All mice were housed in the institutional animal facility under standard 12 h:12 h light:dark cycles with food and water ad libitum.

Plasmids

Sox11 gene was cloned into a double UP mNeon to Scarlet (dUP) backbone[31] (Addgene #125134) as follows. The multicloning site cassette (MCS) downstream Scarlet-P2A element was used to insert the mouse *Sox11* gene. First, the mouse *Sox11* was amplified by PCR from P0 mouse brain cDNA with primers carrying Cla1 and Nhe1 restriction sites respectively. Then, DNA digestion of MCS dUP and *Sox11* DNA was performed with Cla1 and Nhe1 and *Sox11* fragment inserted in the backbone by enzymatic ligation. dUP empty backbone was used as control plasmid. Forward *Sox11* Cla1 primer: TAAGCTatcgatATGGTGCAGCAGGCCGAG AGC; Reverse *Sox11* Nhe1 primer: TAAGCTgctagcTCTCAATACGTG AACACCAGGTCGG.

A plasmid vector was constructed to deliver sgRNA and Cre recombinase-dependent SpCas9 (*hU6-sgRNA-CAG-LSL-3xFLA G-NLS-SpCas9-NLS-P2A-eGFP*). A Cre-dependent SpCas9 transgene was amplified by PCR from Addgene plasmid #61408[32] and cloned under the control of the ubiquitous CAG promoter. A sgRNA expression cassette mediated by human U6 was PCR amplified from Addgene plasmid #52963[33] and cloned upstream of SpCas9 expression cassette. Two sgRNAs targeting *Sox11* were designed using GUIDES (http://guides.sanjanalab.org). Each sgRNA was cloned individually into the SpCas9 expression plasmid. In brief, complementary oligonucleotides (Integrated DNA Technologies) containing the sgRNA sequence and ligation overhangs were annealed and ligated by T7 ligase (NEB, M0318S) into BsmBI (ThermoFisher ER0451) digested SpCas9 plasmid. Sox11_sgRNA_001: AAAGCCCAAGACGGACCCAG; Sox11_sgRNA_002: GTTCCCCGACTACTGCACGC.

The two sgRNAs targeting *Sox11* were co-electroporated together to trigger a 720 bp- deletion of *Sox11* gene. The deletion was confirmed by PCR and sequencing of the deleted *Sox11* band after genomic DNA extraction from 10,000 sorted electroporated cortical neurons at P0 (Quick Extract DNA extraction buffer: 1 mM $CaCl_2$, 3 mM $MgCl_2$, 1 mM EDTA, 1% Triton X-100, 10 mM Tris pH 7.5; incubation at 65 °C for 15 min, 68 °C for 15 min, and 98 °C for 10 min). All samples ($n = 3$) co-electroporated with pCAG:*Cre* plasmid displayed the 200-bp deleted band, while Cre negative samples displayed the 920-bp band only. Forward_Sox_11_001: ACACTCTTTCCCTACACGACGCTCTTCC GATCTCAGCGAGAAGATCCCGTTCA; Reverse_Sox_11_002: GTGACTGG AGTTCAGACGTGTGCTCTTCCGATCGCGCCTCTCAATACGTGAAC.

In utero electroporation

In utero electroporations were performed as described[34]. Timed pregnant CD1 mice were anaesthetized with isoflurane (5% induction, 2.5% during the surgery) and treated with the analgaesic Temgesic (Reckitt Benckiser). Embryos were injected in the left lateral ventricle with ~1 μl of DNA plasmid solution (diluted in endotoxin-free TE buffer and 0.002% Fast Green FCF (Sigma)). Embryos were then electroporated by holding each head between circular tweezer-electrodes (5 mm diameter, Sonidel) across the uterine wall, while 5 electric pulses

(50 V, 50 ms at 1 Hz) were delivered with a square-wave electroporator (Nepa Gene, Sonidel).

Stereotaxic injections

Electroporated pups were anaesthetized by hypothermia at P0, and injected in putative electroporated S1 with 80 nl of AAV9-CMV::P1-Cre-rBG (Addgene, #CS1137, 3.92×10^{12} viral genomes per ml) using the following coordinates from lambda along antero–posterior (AP) and along medio–lateral (ML) axes: AP: 1.2; ML: 1.7.

Isoflurane anesthetized pups were placed in a stereotaxic apparatus on P5, P7 or P14 and were injected with barcoded Sindbis virus[16] (80 nl), or red Retrobeads IX and green Retrobeads from Lumafluor (100 nl), or Alexa 647-conjugated cholera toxin subunit B (CTB, Invitrogen, C-34775) (100 nl).

Coordinates of injection sites from lambda. M injection site: P5, AP: 3, ML:1; P7, AP: 3.3, ML: 1.3; P14, AP: 4.5, ML:1.5. S2 injection site: P5, AP: 2.5, ML: 3,8; P7, AP: 3, ML: 4.2; P14, AP: 3.2, ML: 4.5. S1 injection site: P5: AP: 2.5, ML: 2.5; P7, AP: 3, ML: 3; P14, AP: 3.2, ML: 3.2.

For the MAPseq mapping experiments (Fig. 1), Sindbis virus-injected pups were collected at either 14 h ($n = 2$ pups at P7; $n = 2$ pups at P14) or 24 h ($n = 5$ pups at P5; $n = 2$ pups at P7; $n = 2$ pups at P14) post-infection, when Sindbis *Gfp* RNA is present in axons (Fig. 1b; Supplementary Note 1). The 14 h time point allows preservation of the integrity of the infected cells (and thus access to transcriptomic identities) without affecting the efficiency of MAPseq mapping (Supplementary Note 1). For ConnectID experiments, pups were collected 14 h post-infection. Retrobeads- and CTB-injected pups were collected 36 h post-injection.

Biocytin patch filling and streptavidin staining

Three-hundred-micrometre-thick coronal slices from P14–15 and P21 mouse brain were cut in cooled artificial cerebrospinal fluid (ACSF) containing 125 mM NaCl, 2.5 mM KCl, 1 mM $MgCl_2$, 2.5 mM $CaCl_2$, 1.25 mM Na_2HPO_4, 26 mM $NaHCO_3$ and 11 mM glucose, 0.3% biocytin, oxygenated with 95% O_2 and 5% CO_2. Slices were kept at room temperature and allowed to recover for 1 h before cell filling. Under low magnification, the barrels in layer 4 could be readily identified, and high-power magnification was used to guide patch pipette onto Rbead+ identified neurons. Neurons were filled with 0.3% biocytin (Sigma CAS576-19-2) for 30 min. Sections were then fixed in 4% paraformaldehyde at 4 °C overnight and incubated with Alexa 647 coupled-streptavidin (Invitrogen S21374, 1:500 in PBS-10% tween) for 48 h at 4 °C. Sections were then rinsed with PBS.

Wide-field calcium imaging

Wide-field imaging was performed on P9–10 ($n = 3$) and P14–15 ($n = 3$) pups expressing GCaMP6s in layer 2/3 neurons obtained by crossing Ai94 with Cux2-*cre* mice (Fig. 2f, Extended Data Fig. 7). Thirty minutes before isoflurane anaesthesia, pups received a subcutaneous dose of buprenorphine (0.2 mg kg^{-1}). The temperature of the pups was maintained at 37 °C throughout all the procedures. After injecting lidocaine (3 μl subcutaneously), the scalp was excised exposing the parietal, frontal, interparietal and part of the squamosal bones of the left hemisphere in order to image the primary and secondary somatosensory along with the motor cortices. The periostium was removed and the edges of the incised skin were secured with a thin layer of cyanoacrilyc glue. A custom-made titanium head bar was attached to the interparietal bone using dental cement. During imaging experiments, the pitch and roll angles of the skull were adjusted to position motor and sensory cortices within the focal plane of the camera. Images (10 frames per second) were acquired with a 16-bit Retiga CCD camera controlled with Ephus[35] running on Matlab 2007. GCaMP fluorescence was obtained by illuminating the skull with a blue LED (460 nm, Roithner) using FICT excitation and emission filters. To improve bone translucence and avoid reflection artefacts, the skull was covered with ecographic gel and topped with a coverslip. Functional experiments started 30 min

after surgery and consisted of a series of vibratory stimuli delivered to the right whisker pad, forelimb and hindlimb. Each body part was stimulated for at least 60 trials (one trial every 10 s). Sensory stimuli consisted of 1 s of 100-Hz sinusoidal vibrations delivered with a modified galvanometric mirror controlled through Ephus. Images were analysed offline with custom Matlab scripts. $\Delta F/F$ was computed for every pixel of the raw images on a trial-by-trial basis. The obtained $\Delta F/F$ images were averaged across trials obtaining a $\Delta F/F$ average movie per experimental condition. $\Delta F/F$ movies were analysed using ImageJ. Calcium traces of S1, S2, motor and visual cortical areas were extracted using the Allen Brain Atlas landmarks. For every mouse, normalized $\Delta F/F$ traces were obtained using the average peak S1 fluorescence of their corresponding age group.

Primary culture of dissociated neurons after in utero electroporation

Embryos were electroporated in utero at E15.5 as described above. Half the litter was electroporated with dUP-Sox11 plasmid alone (allowing expression of GFP) and half with dUP-Sox11 and pCAG:*Cre* plasmids (inducing the recombination and the expression of Scarlet-SOX11). Parietal electroporated neocortices were microdissected at E16.5 in ice-cold HBSS (Gibco, 14175-053), and mechanically dissociated after trypsin (TrypLE Express, Gibco, 12604-013) incubation for 5 min at 37 °C. Dissociated cells were plated on 14-mm-diameter coverslips (50,000 cells per coverslip) coated with poly-L-lysine (Sigma, P4707) and natural mouse laminin (0.01 mg ml^{-1}, Invitrogen, 23017015), and cultured in Neurobasal medium (Gibco, 21103-049) supplemented with B27 (Gibco, 17504-044) serum-free supplement, L-glutamine (Gibco, 25030-024) and 1% penicillin-streptomycin at 37 °C in presence of 5% CO_2. GFP (control) and Scarlet-SOX11 neurons were cultured on the same coverslips to avoid any culture condition bias. Neurons were fixed for imaging after two days in vitro by gently adding in the medium 0.5× volume of 8% PFA for 20 min at room temperature, and then washed with PBS.

Open-field test

Unilaterally (left S1) control or SOX11-electroporated E15.5 and P0 Cre-injected female and male CD1 pups were tested at P16. All behaviour experiments and analyses were performed blindly by an independent experimenter to avoid any bias in observations. Pups were tested in the afternoon after a 1-h period of handling (10 min) and habituation to the behavioural room (50 min). Three litters of pups (minimum three pups per litter) were tested per group. Open-field tests were performed in a behavioural room with fixed low illumination (5–10 Lux) and with controlled humidity (40–50%) and temperature (22–24 °C). The open-field test was conducted in a cylindrical arena (diameter: 20 cm, height: 20 cm). The experimental pup was allowed to freely explore the apparatus for 10 min. The arena was cleaned with 1% acetic acid solution and dried between trials. Every session was video-tracked and recorded (24 frames per s) using Ethovision XT (Noldus), which provided an automated recording of the position of the nose, central and tail points. The distance moved, time in centre, time immobile and unilateral whisker exploring were assessed after the position points obtained by the automated recording. Vertical exploration and grooming were measured with manual scoring. The pups were then perfused at P17 and brains were sliced and stained with anti-GFP and anti-RFP antibodies to control for electroporation$^+$ or Cre$^+$ sites and Cre$^+$ neuron axonal projections.

Immunohistochemistry

Postnatal mice were perfused with 4% PFA and brains were fixed overnight in 4% PFA at 4 °C. Eighty-micrometre sections were cut using a vibrating microtome (Leica, VT1000S) and pre-incubated for 2 h at room temperature in a blocking–permeabilizing solution containing 5% bovine serum albumin and 0.3% triton X-100 in PBS, and incubated

for 2 days with primary antibodies at 4 °C. Sections were then rinsed 3 times in PBS and incubated with the corresponding Alexa-conjugated secondary antibodies (1:500; Invitrogen) for 2 h at room temperature. For SOX11 immunostaining, we performed antigen retrieval by incubating the sections in citrate buffer solution (Dako, S1699) for 20 min at 82 °C. Sections were rinsed 3 times in PBS before the blocking–permeabilizing step. Primary antibodies and their dilutions were: chicken anti-GFP (Invitrogen, A10262, 1:2,000), rabbit anti-CUX1 (Santa Cruz, SC-13024, 1:250), rat anti-RFP (Chromotek, 5F8, 1:500), rabbit anti-SOX11 (Millipore, ABN105, 1:500).

In situ hybridization

For antisense *Gfp* probe synthesis, DIG-labelled antisense RNA probe was obtained after in vitro transcription of cDNA obtained from a transgenic mouse line expressing GFP in brain, using specific primers (forward primer: CCATCCTGGTCGAGCTGG; T7Reverse primer: CGAT GTTAATACGACTCACTATAGGGCTTCTCGTTGGGGTCTTTGC). In situ hybridization on slides was performed according to methods described previously[36]. In brief, hybridization was carried out overnight at 60 °C with the digoxigenin (DIG)-labelled *Gfp* RNA probe. After hybridization, sections were washed and incubated with alkaline phosphatase-conjugated anti-DIG antibody (Roche, 11093274910, 1:2,000) for 2 days at 4 °C. After incubation, sections were washed and the colour reaction was carried out overnight at room temperature in a solution containing NBT (nitro-blue tetrazolium chloride) and BCIP (5-brom-4-chloro-3′-indoly phosphate *p*-toluidine salt) (Roche, 000000011681451001). After colour revelation, sections were washed, post-fixed for 30 min in 4% PFA and mounted with Fluoromount (Sigma).

Image acquisition and quantifications

All images from in situ hybridizations were acquired on a Zeiss Axioscan.Z1 slide scanner, equipped with a 10× NA 0.45 Plan Apochromat objective, and a Hitachi HV-F202FCL camera. For ConnectID single-cell quality control, Fluidigm HT800 chip wells imaged using ImageXpress (Molecular Devices) microscope with 10× S Fluor 0.50 objective and a CoolSnap HQ camera (Photometrics). All images from immunohistochemistry were acquired on a Nikon A1r confocal microscope.

For retrogradely labelled neuron laminar (Figs. 1, 2, Extended Data Fig. 1) and retrogradely labelled SOX11 neurons (Fig. 3) quantifications, images were acquired using a 20× 0.5 CFI Plan Fluor WD objective. The injection sites were controlled for both position and depth using an Eclipse 90i epifluorescence microscope (Nikon). Retrogradely labelled neurons in S1 were counted in 3 to 4 sections per pup ($n = 50$ to 200 labelled ICPN) using CellCounter ImageJ plugin. In Fig. 1a and Extended Data Fig. 1a, $n = 200$ $\vec{M}$ or $\vec{S2}$ ICPN were randomly selected for plotting (same number of neurons per pup).

Biocytin patch filled retrogradely labelled SL ICPN were acquired using a 40× 0.6 CFI ELWD S Plan Fluor WD objective. Stack images (z-step = 1 µm) were imported and neurons were reconstructed in 3D using Imaris image analysis software (semi-automatic reconstructions). Morphology analyses of dendrites were blindly performed to avoid any bias (retrogradely labelled neurons: P14–15: $n = 13$ $\vec{M}$; $n = 9$ $\vec{S2}$ ICPN; P21: $n = 5$ $\vec{M}$; $n = 5$ $\vec{S2}$ ICPN; 1 to 3 neurons were filled per pup, ensuring at least 3 pups per condition).

For quantifications of SOX11 immunolabeling fluorescence intensity (Fig. 3) and of in vitro axon length (Extended Data Fig. 8), images were acquired using a 40× 0.6 CFI ELWD S Plan Fluor WD objective. SOX11 immunofluorescence intensity was blindly measured in retrogradely labelled $\vec{M}$ and $\vec{S2}$ ICPN using ImageJ CellCounter plugin; positions of S1 retrogradely labelled ICPN were first recorded prior to measure SOX11 immunolabelling fluorescence, to avoid any selection bias. In order to compare the intensities from distinct sections and brains, for each brain section, fluorescence background intensity (measured in L2/3 inter-cellular space) was subtracted to raw fluorescence

intensities. These corrected values were next normalized to the maximal fluorescence intensity (measured as the most fluorescent cell of L2/3) ($n = 3$ pups per condition, $n = 2$ sections per pup; $n = 40$ to 130 retrogradely labelled neurons per pup; statistical analyses were performed on the average values for each pup). Quantifications of GFP (control) versus Scarlet (SOX11) neuron axon length in vitro were performed using NeuronJ plugin ($n = 53$ control; $n = 26$ SOX11 neurons). Data are from 3 different cortical cultures ($n = 2$ to 3 coverslips per culture), control and SOX11 neurons were pooled and cultured on the same coverslips to avoid any bias due to culture conditions.

For quantifications of axonal phenotype upon SOX11 overexpression (Fig. 3 & Extended Data Fig. 8), control versus SOX11 Scarlet axon images of S2, motor, and corpus callosum at the midline were acquired using a 10× 0.45 CFI Plan Apochromat WD objective. Semi-automated thresholding of Scarlet axons was performed using ImageJ to measure the area covered by the axons in the regions of interest ($n = 2$ to 3 sections per target per pup; $n = 13$ control, $n = 11$ SOX11 pups). All AAV9-CMV::P1-cre-rBG injection sites were controlled using an Eclipse 90i epifluorescence microscope (Nikon), and the number of Cre$^+$ Scarlet neurons in L2/3 was quantified using automated particle detection (minimum size = 10 µm) after semi-automated thresholding in ImageJ. For each pup, the average number of Scarlet cells in S1 was calculated on three sections (Extended Data Fig. 8) and used for normalizing the area covered by axons in the targets. In Fig. 3e and Extended Data Fig. 8e, the threshold bellow which we considered no axon in target was determined when we could not distinguish actual axon labelling from background fluorescence in each of the sections analysed for a given pup. All these quantifications were performed blindly to avoid any bias, with brain sections from the same pups as those used for open-field tests. The same approach was used to quantify the axonal phenotype upon SOX11 down-regulation ($n = 3$ control; $n = 6$ anti-Sox11 sgRNA) (Extended Data Fig. 8).

Tissue microdissection, cell sorting and scRNA-seq

MAPseq. Acute coronal brain sections were cut on a vibrating microtome (Leica, VT1000S) and brain regions were microdissected with micro-scalpel using a Leica Dissecting Microscope (Leica, M165FC) in ice-cold oxygenated ACSF under RNase-free conditions. Brains from individual pups were microdissected separately, on ice. P5: $n = 5$ pups; $n = 2$ litters; section thickness = 400 µm. P7: $n = 4$ pups; $n = 1$ litter; section thickness = 600 µm. P14: $n = 4$ pups; $n = 1$ litter; section thickness = 600 µm.

Microdissected brain tissues were collected in TRIzol reagent-containing tubes (ThermoFisher, 10296-010), mechanically dissociated and immediately stored at −80 °C. Throughout the procedure, sample cross-contamination was carefully avoided. Dissected samples were processed for sequencing as previously described[16,17]. Samples were mixed with spike-in RNA and processed for reverse transcription (using nested primers containing the 6-bp region per pup index and a 12-bp unique identifier (UMI)), production of double-stranded cDNA, treatment with ExonucleaseI (NEB), and two rounds of nested PCR using the following primers and Accuprime Pfx polymerase (ThermoFisher, 12344-040) as previously described[16,17].

PCR1: 5′-CTCGGCATGGACGAGCTGTA-3′, 5′-CAAGCAGAAGACGGCATA CGAGATCGTGATGTGACTGGAGTTCCTTGGCACCCGAGAATTCCA-3′. PCR2 (Solexa primers): 5′-AATGATACGGCGACCACCGA-3′, 5′-CAAGCAGAAGACGGCATACGA-3′.

The resulting PCR amplicons were gel-extracted using Qiagen Min-Elute Gel extraction kit according to the manufacturer's instructions and the cDNA purified with magnetic Agencourt AMPureXP beads (Beckman, A63881). Finally, we sequenced the library on an Illumina HiSeq4000 next generation sequencer using SBS kit (Illumina) pool of primers (HP10, HP12) for single-end 106 base-pair sequencing.

Example of MAPseq amplicon (barcode−common sequence−index− UMI):

GTACAACGATTCGACAAAAGCACCAGCCAA$_{30}$YY$_{32}$GTACTGCGGCC GCTANCTAATTGCCGNCGNGAGGTACGACCACCGCNAGCTGTACA$_{88}$CG TGAT$_{94}$GAGGCACCTCTA$_{106}$.

scRNA-seq. The primary somatosensory cortex was microdissected as aforementioned. Cells were further dissociated by incubating micro-dissected samples in 0.5 mg ml^{-1} pronase (Sigma, P5147) at 37 °C for 10 min, followed by incubation in 5% bovine serum albumin for 3 min, two washes in ACSF and manual trituration using pulled glass pipettes of decreasing diameters. Cells were then centrifuged for 10 min at 600 rpm and resuspended before filtration using a 70-µm cell strainer (ClearLine, 141379C). Cells were then incubated for 10 min at 37 °C with Hoechst (0.1 mg ml^{-1}) and isolated using a Beckman Coulter Moflo Astrios FAC-sorter. Singlet GFP$^+$Hoechst$^+$ cells were sorted according to their forward- and side-scattering properties, and their negativity for Draq7 (Viability dye, far red DNA intercalating agent, Beckman Coulter, B25595). five-thousand to ten-thousand cells were sorted for each experiment by fluorescence-activated cell sorting (FACS). Three microliters of C1 Suspension Reagent (Fluidigm) were added to 10 µl of FACS-sorted cells, which were captured into 800 well- AutoPrep integrated fluidic circuit (IFC) designed for 10- to 17−µm diameter-cells (Fluidigm HT800, 100-57-80) and imaged using the ImageXpress Micro Widefield High Content Screening System (Molecular Devices). cDNA synthesis and preamplification was processed following the manufacturer's instructions (C1 system, Fluidigm). cDNA libraries were prepared using Nextera XT DNA library prep kit (Illumina), quality controlled using 2001 Bioanalyzer from Agilent, and sequenced using HiSeq 2500 sequencer.

Analyses

All bioinformatics analyses were performed using the R programming language and Bioconductor packages as described below. Gene ontologies used QuickGO from EMBL-EBI Hinxton public database (https://www.ebi.ac.uk/QuickGO/) and GSEA (https://www.gsea-msigdb.org/gsea/index.jsp). Transcription factors are listed at TcoF-DB v2 database (https://tools.sschmeier.com/tcof/home/).

MAPseq mapping of projection patterns. Analyses were performed as described[16] (see also Supplementary Note 1). Reads of the FASTQ files were de-multiplexed by pup and region according to Sindbis index (read position 89–94 bp) with 0 mismatch tolerance against the expected target sequences. Reads with an N in the UMI (read position 95–106 bp) or in the barcode (read position 1–32 bp) sequence were filtered out, and barcodes with identical UMI sequence were kept only once. Then, barcodes were considered 'spike' if their tails (positions 25–32) matched with the sequence ATCAGTCA allowing for 1 mismatch; or considered 'viral' if the 2-bp tail (positions 31–32) matched YY.

Viral barcodes from the S1 injection site were used to build a reference library of barcodes for each pup. To correct for sequencing errors, S1 barcodes were mapped on themselves with bowtie v1.1.1 allowing for three mismatches. A graph of sequenced S1 barcodes was generated from the mapping result so that the node represented a barcode sequence, and edges linked two barcodes that differ by less than three mismatches. To identify barcodes with sequencing errors, the maximal weakly connected components of the graph were calculated with R package igraph. For each component, the most abundant sequence was kept as the error-corrected barcode sequence and the UMI counts were summed up. Error-corrected barcodes found in S1 injection site were checked against the known catalogue of barcode for barcoded Sindbis virus. Barcodes found in the target regions were mapped on this S1 reference library with bowtie v1.1.1, allowing for three mismatches. At the end of this procedure, barcodes sequenced in the target regions were associated to barcodes sequenced in S1 and thus establish a picture of S1 multiplex projections. All this procedure was made with a custom code contained in a docker container available at https://github.com/

pradosj/docker_sindbis. The contralateral thalamus (which does not receive input from the cortical injection site) was used as a negative control target. In this region, we found a mean of 4.3 ± 4.9 barcode counts, which we considered as noise value. We therefore excluded barcodes with less than 10 counts in at least one target, as well as those with less than 100 reads in the injection site (S1). Barcode counts were normalized by spike counts found in each target to avoid any experimental bias due to library preparation. ICPN were selected as the barcodes with maximum values in cortical target(s).

For Fig. 1c–e and Extended Data Fig. 1e–h, the same number of barcodes with maximum values in one of the cortical targets (that is, ICPN) were randomly selected per pup in order to normalize for potential variability in labelled cell populations owing to variability in the depth of injection across cortical layers. Projections of single neurons were normalized to 100% and then ordered into heat maps by their projection similarities. For Extended Data Fig. 1f–g, the projection heat maps were transformed into binary projection patterns (where more than 10% projection in a target was considered as above threshold) and the entropy was calculated, as a measure of the randomness of the projection patterns (reflecting their diversity). If the projections towards the 6 targets were totally random, 6 bits of information would be necessary to encode each pattern (0/1). If all cells had similar projection patterns, no bits would be required to display the data. The Kullback–Leibler divergence was then measured. It represents the number of bits per pattern necessary to add at a given age x to represent the patterns of another age y (when the model is optimal to represent age x).

For Extended Data Fig. 2, the same number of $\vec{M}$ and $\vec{S2}$ ICPN (that is, ICPN with more than 10% projection in either M or S2, but not in both) were randomly selected and the distances between projection patterns were calculated using Manhattan distance to define clusters of projection patterns (Ward.D2). Hierarchical clustering (complete linkage) was conducted so that patterns are grouped if they display less 30% difference between each other. Only patterns with representing more than 1% of total ICPN were kept for further analyses. The distribution of $\vec{M}$ and $\vec{S2}$ ICPN within these projection patterns was then analysed.

ConnectID: scRNA-seq combined with MAPseq. Reads were mapped on mouse genome GRCm38 following the same pipeline described in ref. [25] (see Supplementary Note 2). In brief, read1, which contains the UMI sequence, was appended at the end of read2 header. Read2 was further mapped to the mouse genome with Tophat v2.0.13. Resulting alignment files in BAM format were processed with umi_tools[37] to deduplicate reads with identical UMI. Gene-expression quantification was performed with R using summarizeOverlaps method of package GenomicAlignments. Only reads falling into exonic part of a gene are quantified (including 5′ and 3′ UTRs).

Each transcriptome was additionally associated to the result of a manual bright-field picture annotation, where the operator checked for the presence of a single cell in the wells of the fluidigm HT800 chips. Only wells where a single cell was observed were kept for further analyses (wells with no cell, cell(s) with convoluted shapes, multiple cells, or cell(s) with debris were excluded).

Reads that did not map on the mouse genome were aligned against Sindbis virus sequence with BWA[38] for quantification. In addition, unmapped reads were processed with a custom R script to identify the Sindbis barcode sequence of the cell. Using trimLRpatterns method of Biostrings package, this script looked in unmapped reads for the two 80 bp sequences surrounding the 32 bp random barcode of Sindbis virus. If both sequences matched a read with 5% mismatch tolerance and are separated by exactly 32 bp, the sequence in between was considered as the barcode of a Sindbis-infected cell. Single-cell barcodes identified in single cells of a given pup were further corrected for sequencing errors following a similar pipeline as described above (MAPseq). $n = 1,859$ out of 2,174 single cells with identified barcode(s) were obtained, with an average of 25 barcode counts per cell. Additionally, to ensure that the

same barcode was not sequenced in several single cells, it had to be at least three times more abundant in its associated cell than in any other cell. At the end of the procedure, a reference library of barcodes from S1 single cells was obtained. This was used to map barcodes found in the target regions and thus to infer connectivity from single-cell transcriptomes. For single cells with multiple barcodes ($n = 100$ out of 415 cells with identified barcode in target(s), see Extended Data Fig. 3 for distribution of barcodes within single cells), the distance between distinct barcode profiles was measured and cells high-distance profiles were removed ($n = 20$ out of 100 cells; see Supplementary Note 2, Extended Data Fig. 3).

All transcriptomic analyses were performed on reads per million (rpm) normalized ($\log_{10}$) gene expression.

Sindbis is an RNA virus that is likely to affect endogenous transcriptional processes in our experiments. To identify infection-related transcriptional processes and to reduce Sindbis-related transcriptional noise, the following quality controls and normalization procedure were applied: (1) The viral load in each single cell was first determined, that is, the number of sequenced reads which mapped on the sequence of Sindbis virus[16]. (2) An ordinal regression model was generated to identify genes with the strongest weight in ranking neurons based on their viral load. This identified genes whose expression was affected by the virus. (3) Ontologies of high-weight genes (false discovery rate (FDR) < 0.1, $n = 130$ genes) included ribosomal and translation-related processes (QuickGO from EMBL-EBI Hinxton database), consistent with the early stages of hijacking of the cell's translational machinery by the virus (see below). Significantly enriched ontologies assumed to represent reaction to viral infection are listed below; corresponding affected genes present in the dataset were excluded ($|(z)$ score$| > 0.5$ in the regression model). Ontology FDR: inter-species interaction between organisms, 6.11×10^{-7}; viral gene expression 9.29×10^{-6}; ribosome 2.58×10^{-5}; mitochondrion 5.55×10^{-5}. (4) Gene expression was corrected on the basis of viral load and on the number of expressed genes for each cell using a linear correction[39], thus mitigating the effect of viral infection. For this, a dataset of retrogradely labelled $\vec{M}$ and $\vec{S2}$ P9 S1 ICPN was used as control cells for the infection effect (Extended Data Fig. 4). These cells were collected and sequenced as described for ConnectID and were given viral load value = 0.

Before analysing the transcriptomes of single cells, cells with more than 30% of mitochondrial RNA or more than 50% of viral reads compared to total mapped reads were removed. All analyses were performed on $n = 2,450$ neurons ($n = 391$ ConnectID neurons, that is, with barcode identified in target(s) (ConnectID ICPN: P7: $n = 13 \vec{M}$, $n = 35 \vec{S2}$; P9: $n = 10 \vec{M}$, $n = 9 \vec{S2}$; P14: $n = 121 \vec{M}$, $n = 55 \vec{S2}$; retrogradely labelled ICPN: P9: $n = 133 \vec{M}$, $n = 183 \vec{S2}$ ICPN) and $n = 6,144$ gene-corrected expressions (>5 counts in at least 2.5% of the whole dataset).

For Extended Data Fig. 5, k-means clustering ($k = 4$) was used to cluster cells based on their t-SNE values. The same analysis was then performed independently for each age (P7, P9 and P14) and checked the distribution of neurons with identified projections to M, S2, C or Sub in the distinct clusters.

For Fig. 2, pseudo-maturation alignment and transcriptional waves were performed as described[25]. Regularized ordinal regression method was used to predict the age (maturation) of the cells (P7, P9 and P14), restricting the analysis to the top 2,000 variable genes identified with the FindVariableFeatures function of Seurat package (vst method). This method was implemented from bmrm R package (Bundle Methods for Regularized Risk Minimization Package, author: Julien Prados, year: 2018, R package version 3.10). The linear weight of the model is used to rank the genes according to their ability to predict each cell age and the prediction scores are defined by the linear combination of the core gene expression. All reported maturation predictions (pseudo-maturation score) were obtained by 10-fold cross-validations. $\vec{M}$ and $\vec{S2}$ ICPN from both ConnectID and Rbeads experiments were then aligned based on their pseudo-maturation score and their gene

expression along this axis was corrected using weighted expression in neighbour cells. This was performed separately for $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN. All transcriptional patterns were then normalized to the maximum value and further clustered in 3 groups (waves) based on their distances along pseudo-maturation using pam function of cluster package ($k = 3$). The average expression pattern was calculated for each cluster and the distances of all transcriptional patterns to this average were calculated. Only genes closely related to a given pattern in both $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN were kept for further analyses ($n = 1,293$ genes; Supplementary Table 2). Gene ontologies were performed using gene set enrichment analysis[40] (GSEA). Genes related to neuron differentiation, axon development, dendrite and synapse were identified using QuickGO gene ontology annotations from EMBL-EBI.

For Extended Data Fig. 5c, a machine learning approach was performed: a logistic regression model with regularization was used to build binary prediction models of superficial (SL) versus DL layers based on single cells from SL and DL microdissection. This implementation was provided by the bmrm R package and allowed reducing the possibility of overfitting as follow. We limited the linear model to the top and bottom 25 (for layer model) and re-trained a new model on these selected genes (2-stage learning process). Model performances were addressed by cross-validations on the subset of genes, which gave a prediction value to build receiver operating characteristic (ROC) curves, the reported performances generalizing accordingly. The genes selected by this approach were further used to predict the layer identity of all single cells. The use of logistic regression here was warranted by the fact that this dataset was obtained from biologically clearly distinct samples (that is, microdissection of SL and DL), providing a high level of confidence on the nature of the biological sample; previous work has shown these can be robustly distinguished using logistic regression together with regularization (see for example, refs. [34,41]). A cross-validation analysis to estimate the generalization error yielded an AUC of 0.97. In comparison, an SVM approach, which is well-accepted to handle high dimensionality, yielded a comparable AUC of 0.99, also consistent with the salient molecular and biological differences between SL and DL neurons.

For Fig. 3, the distances between $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN gene-expression profiles along pseudo-maturation were calculated for all genes from wave 1 (that is, genes dynamically regulated along postnatal development, with high expression at P7). Genes differentially expressed between $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN (that is, genes with distances >5.8 between $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN) were plotted and gene ontologies were performed on this gene set using GSEA[40] (Supplementary Table 4). We used the TcoF-DB v2 database (Massey University and CBRC@KAUST 2010–2017, v2.2.2) to identify transcription factors and easyPubMed package to automatically quantify the number of PubMed abstracts containing the term 'axon development' for each differentially expressed gene.

Statistical analyses

Statistical analyses were conducted with GraphPad Prism 8. Data are represented as the mean ± s.e.m. (bar plots) or median ± s.d. (box plots) and the significance threshold was set at 95% of confidence. All tests were two-sided unless otherwise specified. The following convention was used: NS, not significant; $*P < 0.05$; $**P < 0.001$, $***P < 0.0001$. Sample size was not predetermined statistically but corresponds to what is being used in the field.

Figure 1. MAPseq mapping. N target(s): two-way ANOVA (P5, P7, P14; 1, 2, >2 targets); age: $P > 0.9999$, $F_{(2,27)} = 2.278 \times 10^{-16}$; targets: $P < 0.0001$, $F_{(2,27)} = 61.62$. Main targets: two-way ANOVA (P5, P7, P14; M, S2, C, A, V targets); age: $P > 0.9999$, $F_{(2,45)} = 1.898 \times 10^{-18}$; targets: $P = 0.0029$; $F_{(4,45)} = 4.742$. M, S2 targets: two-way ANOVA (P5, P7, P14; M, S2, M + S2 targets); age: $P > 0.9999$; $F_{(2,27)} = 4.423 \times 10^{-16}$; targets: $P = 0.0005$; $F_{(2,27)} = 10.16$.

Figure 2. Neuron differentiation genes, wave in $\overrightarrow{M}$ versus $\overrightarrow{S2}$ ICPN: contingency test; $P = 0.0442$, $\text{Chi}^2 = 6.240$, d.f. = 2. Expression of genes

along pseudo-maturation: repeated-measures ANOVA based on general linear model with Geisser–Greenhouse correction; neuron differentiation ($n = 245$ genes): $P = 0.0005$, $F_{(1, 310)} = 12$; axon development ($n = 72$ genes): P value = 0.04, $F_{(1,142)} = 4$; dendrite ($n = 125$ genes): P value = 0.05, $F_{(1,248)} = 3$; synapse ($n = 245$ genes): $P = 0.1$, $F_{(1,488)} = 2$. Retrograde labelling: two-way ANOVA (P5, P7, P14; percent $\overrightarrow{M}$ versus $\overrightarrow{S2}$ ICPN in SL); age: $P < 0.0001$, $F_{(2,16)} = 24.07$; ICPN in SL: $P = 0.0133$, $F_{(1,16)} = 7.742$; multiple t-tests using false discovery rate approach with $Q = 1\%$: P5, $q = 0.0061$, d.f. = 6; P7, $q = 0.11$, d.f. = 4; P14, $q = 0.53$, d.f. = 6.

Sholl analysis of dendritic trees (retrogradely labelled SL ICPN patch filled with biocytin): two-way ANOVA; P14: $P < 0.0001$, $F_{(1,4260)} = 139.1$. Ca^{2+} imaging upon sensory stimulation: two-way ANOVA with Sidak's multiple comparisons test. P9 versus P14: M: $t = 0.000$, d.f. = 16, adjusted $P = 0.0004$; S2: $t = 0.6801$, d.f. = 16, adjusted $P = 0.9405$. Extended Data Fig. 7: M versus V: P9: $t = 0.7642$, d.f. = 12, adjusted $P = 0.8421$; P14: $t = 6.209$, d.f. = 12, adjusted $P = 0.0001$. S2 versus V: P9: $t = 3.343$, d.f. = 12, adjusted $P = 0.0174$; P14: $t = 3.535$, d.f. = 12, adjusted $P = 0.0123$.

Figure 3. SOX11 retrogradely labelled $\overrightarrow{M}$ versus $\overrightarrow{S2}$ ICPN: unpaired t-test: $P < 0.0001$; $t = 10.27$; d.f. = 6. Presence of Scarlet[+] axons in M/S2: Fisher's exact test; M: $P = 0.0149$; S2: $P = 0.1993$.

Extended Data Fig. 8. Expression of *Sox11* transcript in $\overrightarrow{M}$ versus $\overrightarrow{S2}$ ICPN along pseudo-maturation: Kolmogorov–Smirnov test; $P < 0.0001$; $D = 0.3544$. Enrichment of SOX11 protein in $\overrightarrow{M}$ versus $\overrightarrow{S2}$ ICPN at P7: unpaired t-test; $P = 0.0149$; $t = 2.936$; d.f. = 10. Cre[+] cells: unpaired t-test: $P = 0.0808$; $t = 1.823$; d.f. = 24. Scarlet[+] axons in M/S2 and corpus callosum at the midline: two-way ANOVA with Sidak's multiple comparisons test; Ctl versus SOX11: M: adjusted $P < 0.0001$; $t = 5.342$; d.f. = 46.00; S2: adjusted $P = 0.0008$; $t = 3.838$; d.f. = 46.00; M versus S2: Ctl: adjusted $P = 0.3186$; $t = 1.379$; d.f. = 46.00; SOX11: adjusted $P = 0.0251$; $t = 1.379$; d.f. = 46.00. Ctl versus *Sox11* sgRNA: M: adjusted $P = 0.0833$; $t = 2.435$; d.f. = 28.00; S2: adjusted $P = 0.0339$; $t = 2.826$; d.f. = 28.00. Axon length at DIV2: unpaired t-test: $P = 0.3101$; $t = 1.022$; d.f. = 77.

Figure 4 and Extended Data Fig. 9. Statistical outliers were identified with the ROUT method ($Q = 1$) and excluded from the analysis (1 animal was excluded from the analysis because the total distance travelled was less than 400 cm). The normality of sample distributions was assessed with the Shapiro–Wilk criterion and when violated non-parametric tests were used. When normally distributed, the data were analysed with unpaired two tailed t-tests. When normality was violated, the data were analysed with Mann–Whitney test. For the analysis of variance with two factors (two-way repeated-measures ANOVA), normality of sample distribution was assumed, and followed by Bonferroni post hoc test. Distance travelled: unpaired t-test; $P = 0.0137$; $t = 2.681$ and d.f. = 22; time spent exploring: unpaired t-test; $P = 0.0227$; $t = 2.443$ and d.f. = 23; velocity: unpaired t-test; $P = 0.0355$; $t = 2.234$, d.f. = 23; time immobile: Mann–Whitney test; P value = 0.0005; time in centre: unpaired t-test; $P = 0.9030$; $t = 0.1233$ and d.f. = 22; time spent grooming: unpaired t-test; $P = 0.1990$; $t = 1.324$ and d.f. = 22; time spent rearing: unpaired t-test; $P = 0.0007$; $t = 3.955$ and d.f. = 22; time spent exploring ipsilateral versus contralateral wall: Ctl (ipsilateral versus contralateral): paired t-test; $P = 0.3601$; $t = 0.955$ and d.f. = 11; SOX11 (ipsilateral versus contralateral): paired t-test; $P = 0.0073$; $t = 3.28$ and d.f. = 11. Majority of time spent in ipsilateral versus contralateral wall by pup: Fisher's exact test: $P = 0.0414$.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

scRNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under the following accessions: MAPseq, GSE118681; Retrobeads: GSE116944; ConnectID: GSE156080.

Code availability

https://github.com/pradosj/docker_sindbis

31. Taylor, R. J. et al. Double UP: a dual color, internally controlled platform for in utero knockdown or overexpression. *Front. Mol. Neurosci.* **13**, 82 (2020).
32. Platt, R. J., et al. CRISPR–Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).
33. Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
34. Vitali, I. et al. Progenitor hyperpolarization regulates the sequential generation of neuronal subtypes in the developing neocortex. *Cell* **174**, 1264–1276.e15 (2018).
35. Suter, B. A. et al. Ephus: multipurpose data acquisition software for neuroscience experiments. *Front. Neural Circuits* **4**, 100 (2010).
36. Golding, B. et al. Retinal input directs the recruitment of inhibitory interneurons into thalamic visual circuits. *Neuron* **81**, 1057–1069 (2014).
37. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
38. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
39. Nestorowa, S. et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**, e20–e31 (2016).
40. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
41. Oberst, P. et al. Temporal plasticity of apical progenitors in the developing mouse neocortex. *Nature* **573**, 370–374 (2019)

**Additional information**
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-04048-3.
Correspondence and requests for materials should be addressed to Denis Jabaudon.
Peer review information *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.
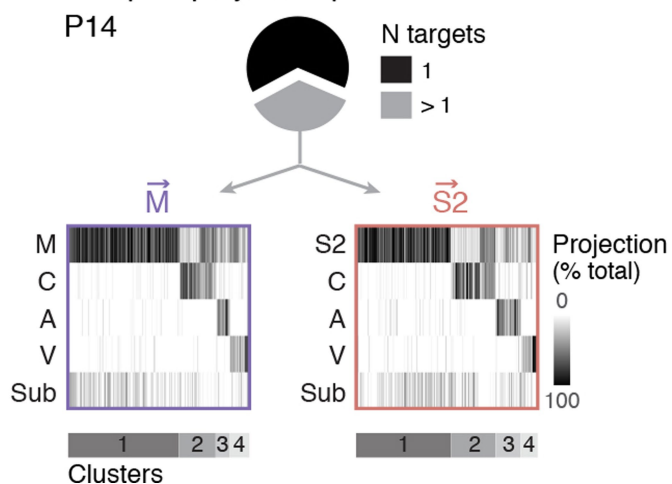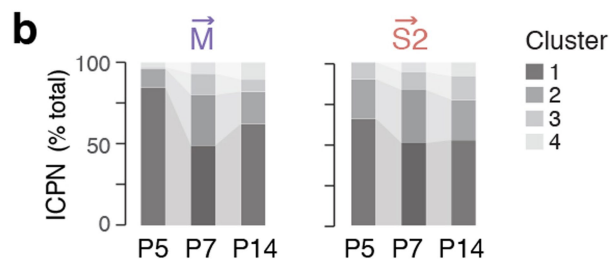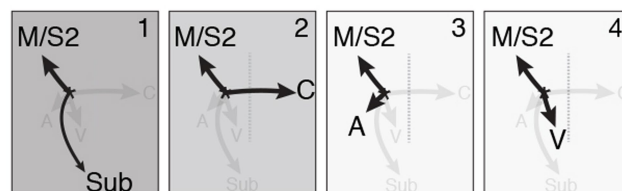**Reprints and permissions information** is available at http://www.nature.com/reprints.
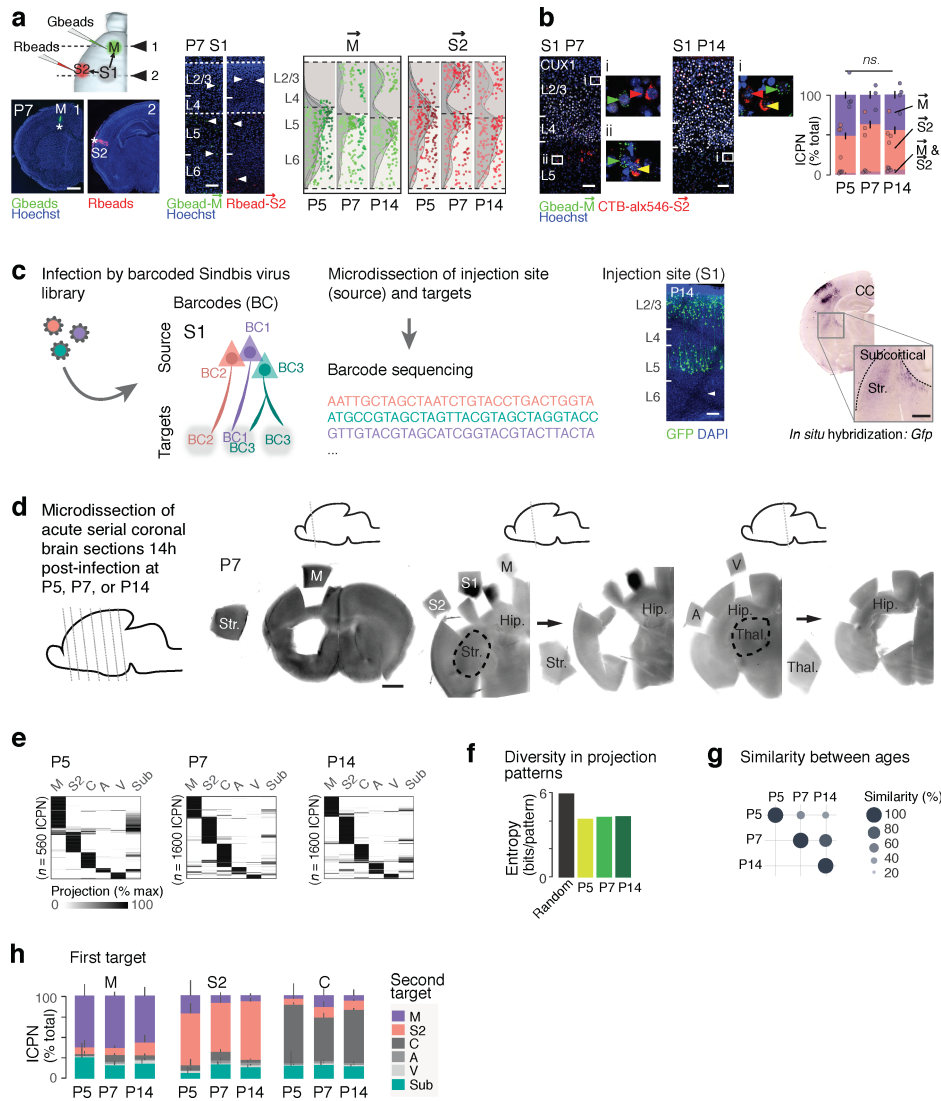
Extended Data Fig. 1 | MAPseq experimental procedure to study postnatal emergence of inter-areal cortical connectivity. a, Left, retrograde labeling from M or S2 using green (Gbeads) or red (Rbeads) retrobeads. *Injection site. Center, Gbeads M- and Rbeads S2- inter-areal cortical projection neurons ($\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN) in S1 at P7 (arrowheads, retrogradely-labeled cells). Right, Illustrations of $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN layer positions at P5, P7 and P14 (P5: $n = 4$, P7: $n = 3$, P14: $n = 2$ pups from 2 litters / target; 50 random quantified cells were plotted per condition; dot shades represent the different pups). **b**, Double retrograde labeling using Gbeads in M and alexa 546-conjugated CTB in S2 at P5, P7 and P14 (P5, P14: $n = 4$; P7: $n = 3$ pups from 2 litters; $n = 50$ to 200 $\overrightarrow{M}$ and $\overrightarrow{S2}$ ICPN per pup). Note that the proportion of co-labeled cells (*i.e.* cells projecting to both M and S2) is stable from P5 on (two-way ANOVA test: P value > 0.9999). **c**, Left, MAPseq principle. Center, infected neurons in S1 expressing Sindbis-GFP 14 hours after infection. Arrowhead shows an axon labeled with GFP. Right, *in situ* hybridization at P14 shows barcode-*Gfp* mRNA (level 2 from

Fig. 1b). **d**, Microdissections of injection and target sites. **e**, S1 ICPN multiplex projections at P5, P7 and P14 (showing data from Fig. 1c as percent max projection). **f**, Diversity of projection patterns measured by the entropy at each age (see Methods). In black is the value for random patterns (*i.e.* 6 bits / pattern). **g**, Similarity between projection pattern matrices, shown as 1 - relative Kullback-Liebler divergence value (see Methods). Note that P7 and P14 projection pattern matrices display the highest similarity. **h**, First and second targets of ICPN at P5, P7 and P14. ICPN with only one target are represented with same first and second targets. Values are shown as mean ± s.e.m. (**b**, **h**). Scale bars, 100 μm (**a**, **b**); 300 μm (**c**); 1 mm (**d**). A, auditory cortex; C, contralateral cortex; CTB, cholera toxin B; Gbeads, green retrobeads; Hip, hippocampus; M, motor cortex; P, postnatal day; Rbeads, red retrobeads; S1, primary somatosensory cortex; S2, secondary somatosensory cortex; Sub, subcortical; Str, striatum; Thal, thalamus; V, visual cortex.

a Multiplex projection patterns
P14

$\vec{M}$

$\vec{S2}$

N targets
■ 1
■ > 1

M
C
A
V
Sub

S2
C
A
V
Sub

Projection
(% total)

0

100

Clusters

M/S2 1
M/S2 2
C
A
M/S2 3
M/S2 4
V

b

$\vec{M}$

$\vec{S2}$

Cluster
■ 1
■ 2
■ 3
■ 4

ICPN (% total)

100

50

0

P5 P7 P14

P5 P7 P14

Extended Data Fig. 2 | $\vec{M}$ and $\vec{S2}$ ICPN have otherwise similar multiplex projection patterns. a, Cluster analysis of $\vec{M}$ and $\vec{S2}$ ICPN projecting in more than one target at P14 reveals 4 projection patterns ($n = 400$ $\vec{M}$ or $\vec{S2}$ ICPN). Note the similar distribution of patterns for both populations. **b**, Distribution of P5 and P7 $\vec{M}$ and $\vec{S2}$ ICPN projecting in more than one target in the P14 clusters using k-nearest neighbors (knn; see Methods). Note that $\vec{M}$ and $\vec{S2}$ ICPN show similar patterns at P7 and P14.
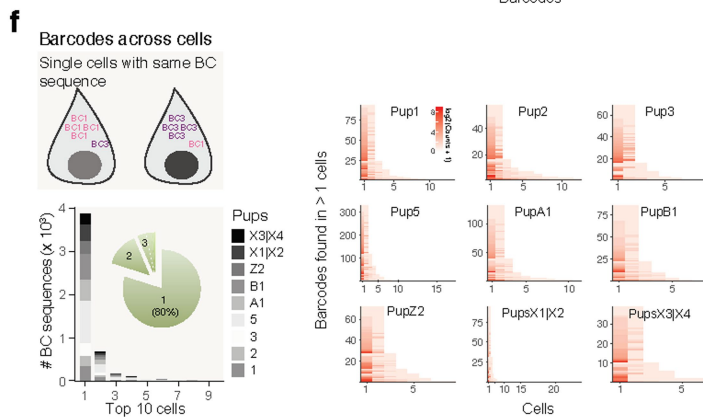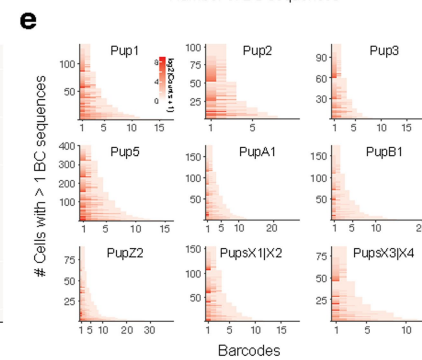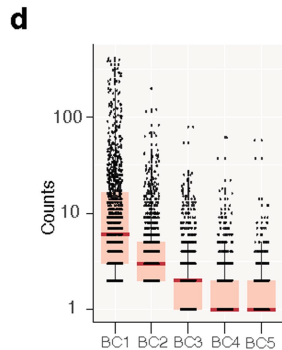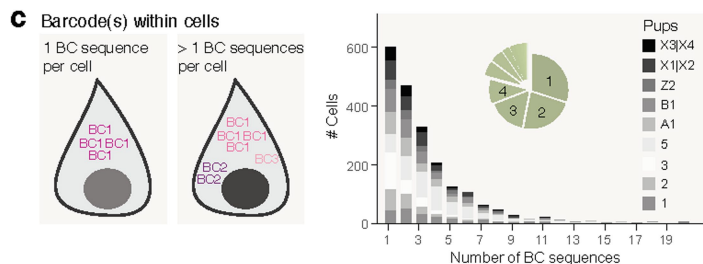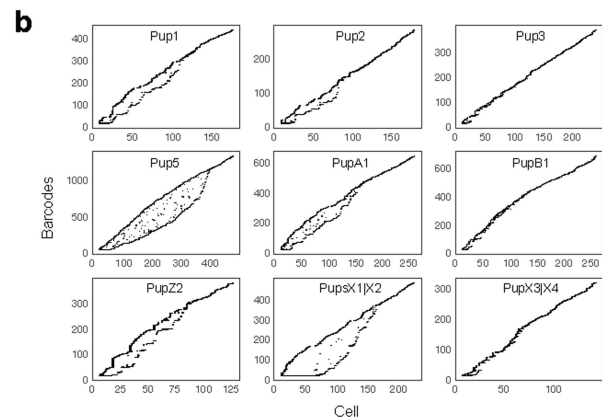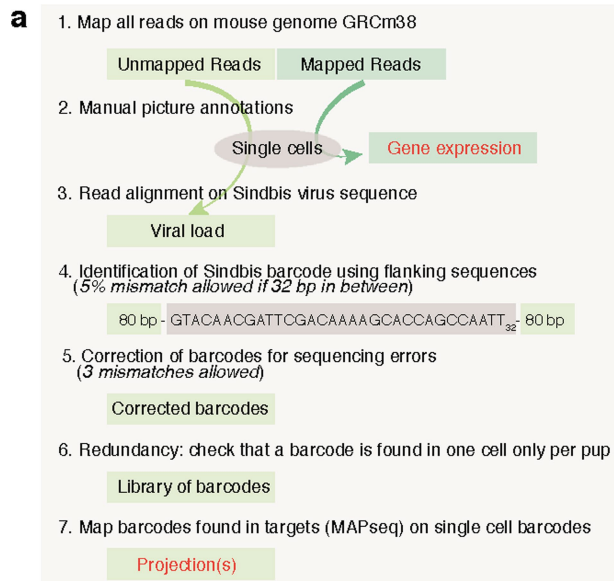
Extended Data Fig. 3 | ConnectID data preprocessing and quality controls. Data presented here are the raw data from the single-cell analysis, and include cells without detected projections in the targets. **a**, Data processing allowing retrieval of single-cell gene expression and projection(s). **b-f**, Non-filtered barcodes (BC) quality controls in single cells after sequencing error correction (step 5 in **a**) for each pup used in this study (*n* = 1859 cells). **b**, Single cells and their corresponding BC. Each column corresponds to a cell, and each line to a BC sequence. **c**, BC sequence(s) within cells. Number of BC with distinct sequences per cell. Pie chart represents summary of the data. ~80% of the cells express 1 to 4 BC sequences. **d**, BC counts for the top 5 most expressed BC. Box plots indicate median ± s.d. and interquartile range. **e**, Counts of the most expressed to the less expressed BC for cells with more than 1 BC sequence. **f**, Barcodes across cells. Number of BC sequences found in the top 9 cells (left). 80% of the BC sequences are found in only one cell. Counts for the BC found in several cells from the cell where it is the most abundant to the cell where it is the least abundant (right).
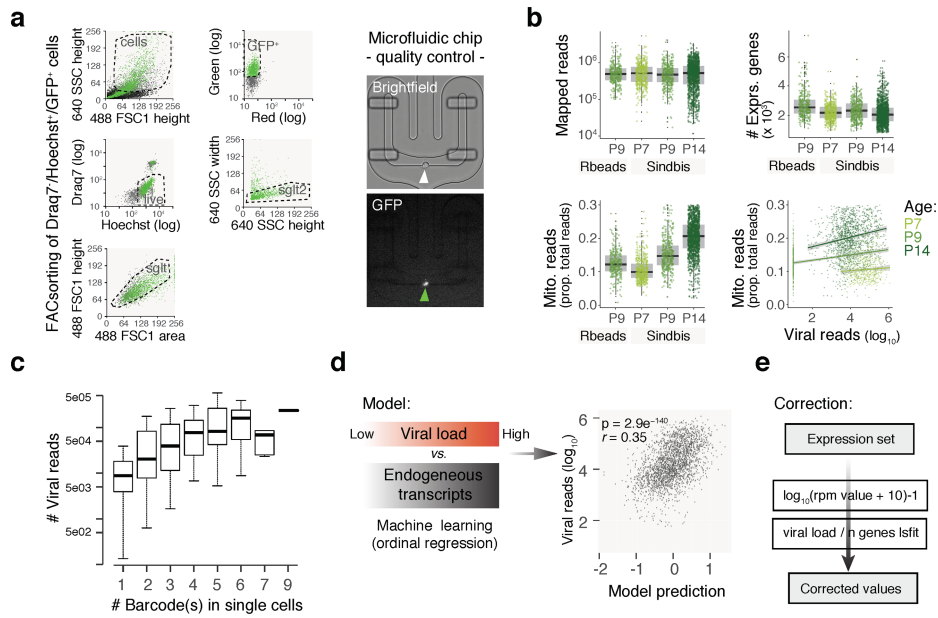
Extended Data Fig. 4 | Quality controls of single cells collected 14 hours after Sindbis infection. a, Fluorescence activated (FAC) sorting of Sindbis-GFP+/Hoechst+ cells (left), capture and quality controls of single cells in microfluidic wells using brightfield imaging (right). Wells with no cell, several cells, debris or with high mitochondrial reads (> 30% total mapped reads) were excluded. **b**, Number of mapped reads, of expressed (exprs.) genes, and proportion of mitochondrial (mito.) reads (on total mapped reads) per cell. Bottom right, note that the proportion of mitochondrial reads is not correlated with the viral load (*i.e.* with the number of Sindbis reads) in each cell.

Retrogradely-labeled ICPN (Rbeads) collected at P9 were used as control cells for the effect of the viral infection. **c**, Viral load increases with the number of barcode(s) per cell (*n* = 2450 cells). Box plots indicate median ± s.d. and interquartile range. **d**, Ordinal regression model identifies genes with the strongest weight in distinguishing neurons with a higher viral load from those with a lower viral load. **e**, "least square" fit regression was performed to the expression set using viral load and number of expressed genes as variables to regress. rpm, reads per million.
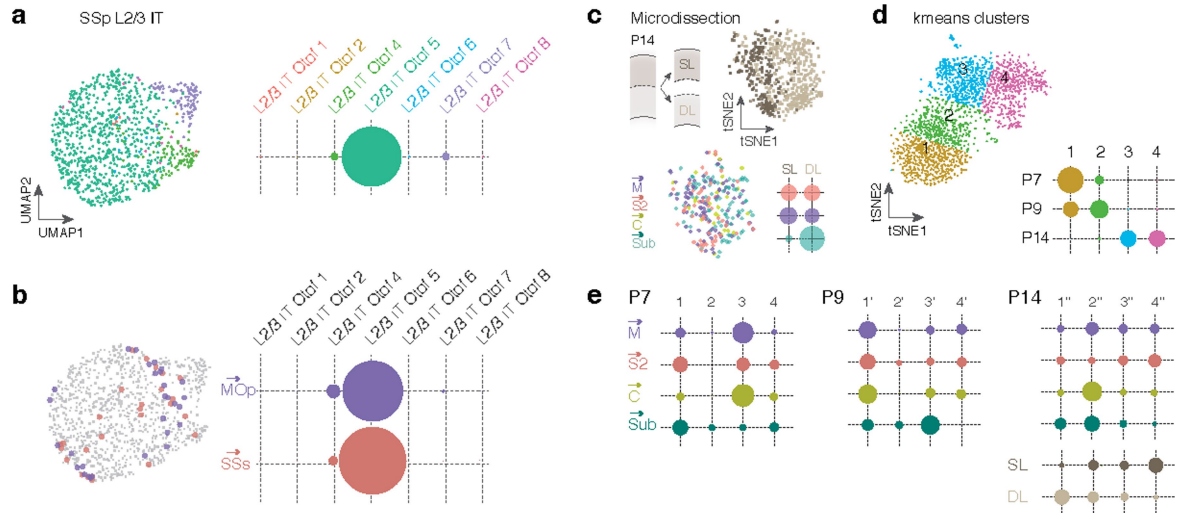
Extended Data Fig. 5 | S1 L2/3 ICPN have highly similar transcriptional identity. a, Left, UMAP of SSp L2/3 intratelencephalic (IT) neurons and clusters from Yao et al., 2021[15]. Right, distribution of cells across clusters. **b**, Distribution of neurons projecting to MOp (corresponding to "M") or SSs ("S2") retrogradely labeled (using retrograde AAV2 virus) in the UMAP (left) and within the clusters (right). Note that neurons projecting to MOp or SSs neither belonged to distinct transcriptional cluster nor clustered apart in the L2/3 IT Otof 5 cluster, suggesting that these subtypes have highly similar transcriptional identities. **c**, Top, tSNE representation of P14 neurons after layer microdissection. Bottom, main target and distribution of ConnectID neurons in superficial (SL) versus deep (DL) layers. **d**, Kmeans clustering of single cells based on their tSNE values (left) and distribution of cells in the 4 identified clusters (right). Note that cells cluster based on their age of collection. **e**, Kmeans clustering based on tSNE values calculated independently at each collection age (P7, P9, P14), and distribution of neurons with known projection within these clusters. P7 and P14 cells are from ConnectID experiments, while P9 cells are from both ConnectID and Rbeads experiments (see Methods). Note that developing ICPN are not detectably clustered by their axonal target, but rather by their layer position, as showed at P14 when SL and DL were microdissected. IT, intratelencephalic; L, layer; MOp; primary motor cortex; SSp, primary somatosensory cortex; SSs, secondary somatosensory cortex.
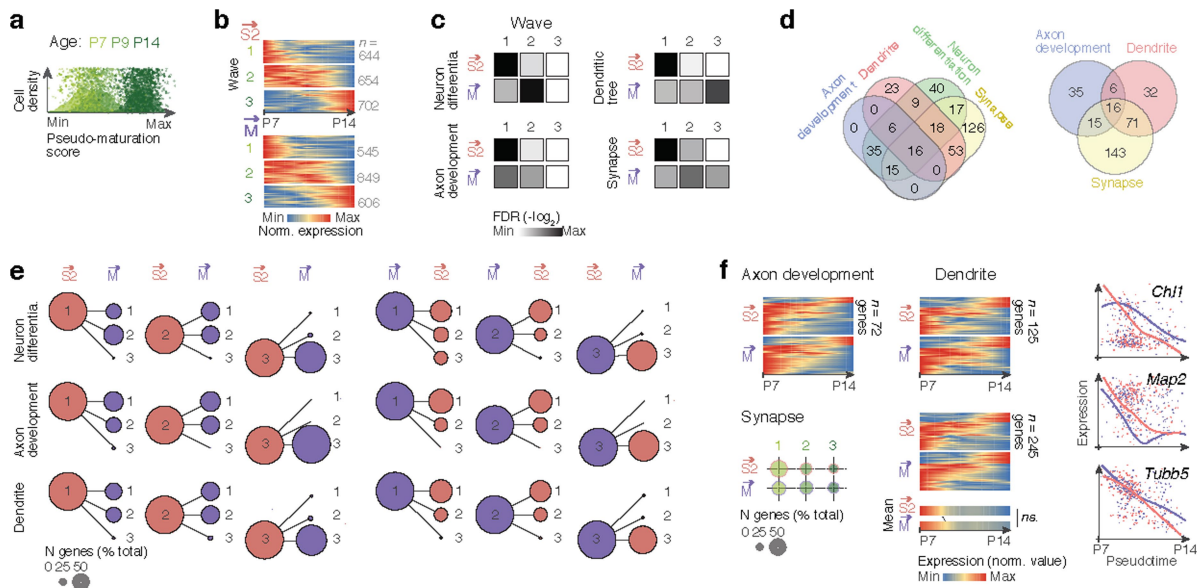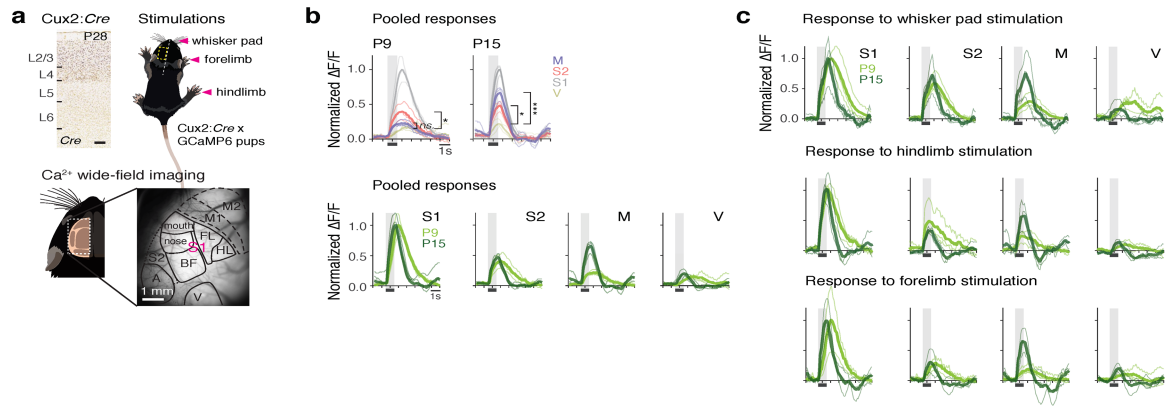
Extended Data Fig. 6 | Transcriptional maturation of $\overrightarrow{S2}$ and $\overrightarrow{M}$ ICPN.
a, Pseudo-maturation score calculated for each cell (see Methods). **b**, Expression and number of genes in each wave for $\overrightarrow{S2}$ and $\overrightarrow{M}$ ICPN. **c**, Gene ontology analysis of genes belonging to each wave in $\overrightarrow{S2}$ and $\overrightarrow{M}$ ICPN. **d**, Venn diagrams showing shared genes between ontologies. **e**, Distribution of $\overrightarrow{S2}$ ICPN wave-defined genes into $\overrightarrow{M}$ ICPN waves (left), and of $\overrightarrow{M}$ ICPN wave-defined genes into $\overrightarrow{S2}$ ICPN waves (right) by gene ontology. **f**, Top, expression of axon development- (left),

and dendrite- (center) related genes along pseudo-maturation in $\overrightarrow{S2}$ and $\overrightarrow{M}$ ICPN. Bottom, number of genes related to synapse in each wave for $\overrightarrow{S2}$ and $\overrightarrow{M}$ ICPN (left) and synapse-related gene expression along pseudo-maturation in $\overrightarrow{S2}$ and $\overrightarrow{M}$ ICPN (center) (repeated measures-ANOVA based on a general linear model with Geisser-Greenhouse correction). Right, examples of dynamics of gene expression for 3 classical axon (*Chl1* and *Tubb5*) and dendrite (*Map2*) development-related genes. Dots correspond to single cells.

Extended Data Fig. 7 | Functional maturation of M⃗ and S⃗2 inter-areal cortical projection neurons using calcium imaging upon somatosensory stimulations. a, Calcium wide-field imaging upon sensory stimulations in Cux2:*Cre* x GCaMP6s anesthetized pups at P9 and P15. Experimental paradigm and *in situ* hybridization for the *Cre* transcript in Cux2:*Cre* P28 pup (from the Allen Brain Institute database). **b**, Pooled responses to whisker pad, hindlimb and forelimb stimulations in S1, M, S2, and V at P9 and P15. At P9, response in M was not different than response in V (considered as background value) (*n* = 3 animals / age; two-way ANOVA with Sidak's multiple comparisons test: M versus V: P9: adjusted *P* value = 0.8421; P14: adjusted *P* value = 0.0001; S2 versus V: P9: adjusted *P* value = 0.0174; P14: adjusted *P* value = 0.0123).

Thinner traces represent the values for each individual pup. Grey, stimulation. Pooled responses are either represented by age (top) or by cortical area (bottom). Note the distinct response dynamics at P9 and P15, P15 responses decreasing faster after the stimulation than P9 responses. **c**, Response to whisker pad (top), hindlimb (middle), and forelimb (bottom) stimulations. The absence of response in M at P9 does not depend on the type of stimulation. Scale bar, 100 µm (**a**, top). Fluorescence (F) in targets was normalized to the mean fluorescence in S1. A, auditory cortex; BF, barrel field; FL, forelimb; HL, hindlimb; M, motor cortex; P, postnatal day; S1, primary somatosensory cortex; S2, secondary somatosensory cortex; V, visual cortex.
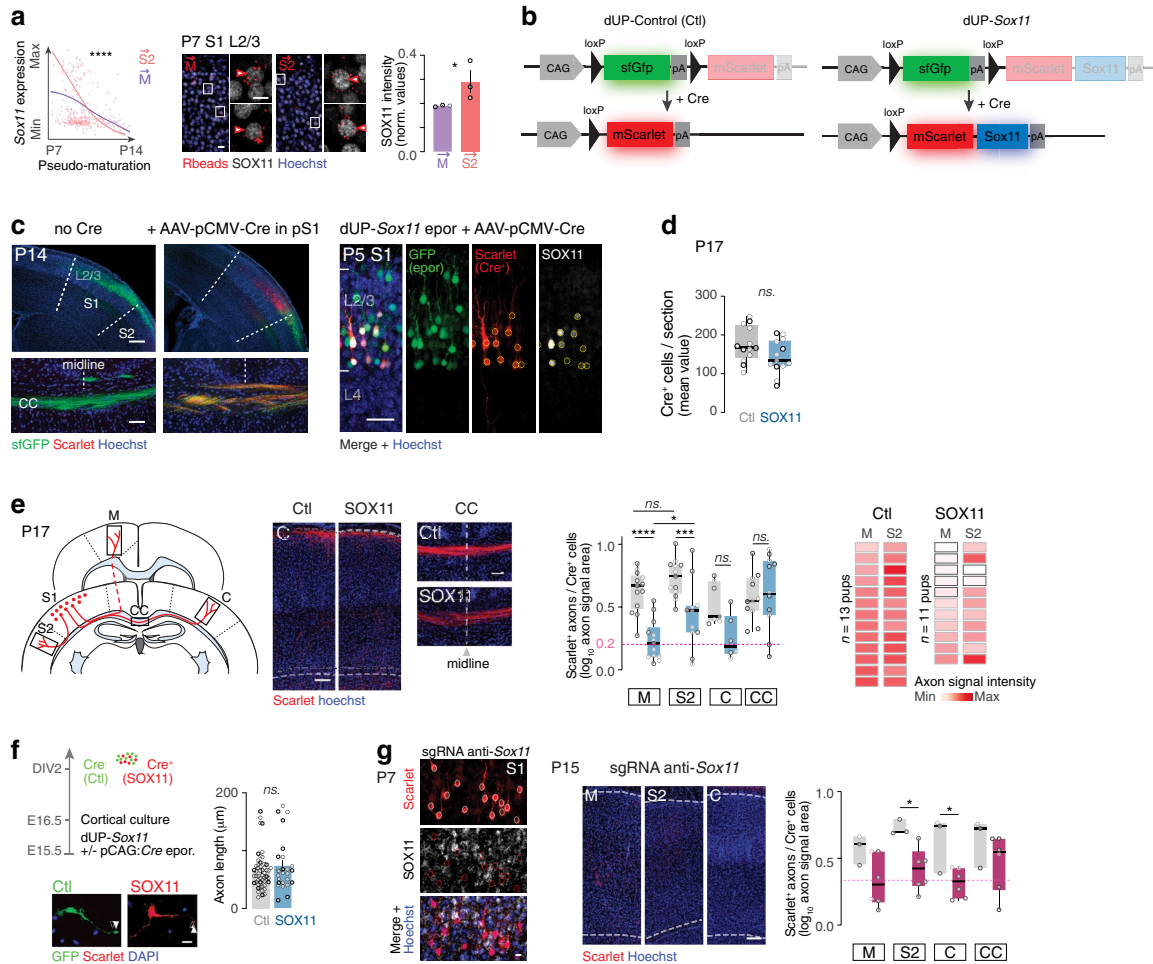
Extended Data Fig. 8 | Restricted up- and down-regulation of SOX11 in time and space alters intracortical connectivity. a, *Sox11* is transiently enriched in $\vec{S2}$ ICPN at P7 (Kolmogorov-Smirnov test: *P* value < 0.0001), as confirmed at the protein level in $\vec{S2}$ compared to $\vec{M}$ retrogradely-labeled ICPN (*n* = 3 pups from 2 litters per target; unpaired *t*-test: *P* value = 0.0149). Values are shown as mean ± s.e.m. **b**, Double-UP plasmid control (Ctl) / *Sox11* constructs and Cre-dependent recombination allowing the expression of Scarlet (Ctl) or Scarlet + SOX11 (SOX11). **c**, Left, GFP and Scarlet expression at P14 without (no Cre) or after AAV-pCMV-Cre stereotaxic injection at P0 in putative S1 (pS1). Note the restricted expression of Scarlet in S1 cells, while GFP electroporated cells are found both in S1 and S2. Right, overexpression of SOX11 protein in Scarlet+ cells at P5. **d**, Quantification of the number of Cre+ cells in S1 at P17 (data are the mean values over 3 sections per pup; unpaired *t*-test: *P* value = 0.0808). Note that SOX11 overexpression postnatally does not affect neuron migration or survival. **e**, Left, representative images of Scarlet axons in C and CC. Center, fluorescence-based quantification of the area covered by axonal signal (normalized by the number of Cre+ cells in S1) in M, S2, C and CC for each pup (log10 values, see Methods; M: *n* = 13 Ctl, *n* = 13 SOX11; S2: *n* = 13 Ctl, *n* = 11 SOX11; C: *n* = 7 Ctl, *n* = 6 SOX11; CC: *n* = 11 Ctl, *n* = 11 SOX11 pups from 4 litters per condition). Number of brains analyzed differ across targets because in some cases no quantification was possible. Pink dashed line, threshold below which no axon was detected. Right, heatmap of axon signal intensity in M and S2 for each pup. Dark grey, pups for which no axon was detected in the target. Only brains in which S2 and M projections could be assessed are shown here (in two SOX11 cases, only M projections could be assessed). **f**, Co-culture of SOX11-overexpressing and control L2/3 neurons. Left, experimental procedure. Half of the embryos were electroporated with dUP-*Sox11* alone (Ctl) and the other half with dUP-*Sox11* and pCAG:*Cre* plasmids (SOX11) at E15.5. Control and SOX11 cortical neurons were dissociated and cultured together at E16.5. Right, at DIV2, axon length was similar in both conditions. Values are shown as mean ± s.e.m. (unpaired *t*-test: *P* value = 0.3101). **g**, Left, inducible single-guide (sg) RNA against *Sox11* leads to absence of SOX11 expression after AAV-pCMV-Cre stereotaxic injection at P0 in putative S1. Right, representative images of Scarlet+ axons in each target and quantification of the axonal phenotype. Scale bars, 100 μm (**c, e, g** right); 10 μm (**a, f, g** left). **d, e, g**, Box plots indicate median ± s.d. and interquartile range. **e, g**, Two-way ANOVA with Sidak's multiple comparisons test: Ctl versus SOX11: M: adjusted *P* value < 0.0001; S2: adjusted *P* value = 0.0008; M versus 2: Ctl: adjusted *P* value = 0.3186; SOX11: adjusted *P* value = 0.0251; Ctl versus *Sox11* sgRNA: M: adjusted *P* value = 0.0833; S2: adjusted *P* value = 0.0339. CC, corpus callosum; DIV, days *in vitro*; dUP, double-UP; E, embryonic day; epor., *in utero* electroporation; M, motor cortex; L, layer; P, postnatal day; S1, primary somatosensory cortex; S2, secondary somatosensory cortex; sfGFP, super-folded GFP.
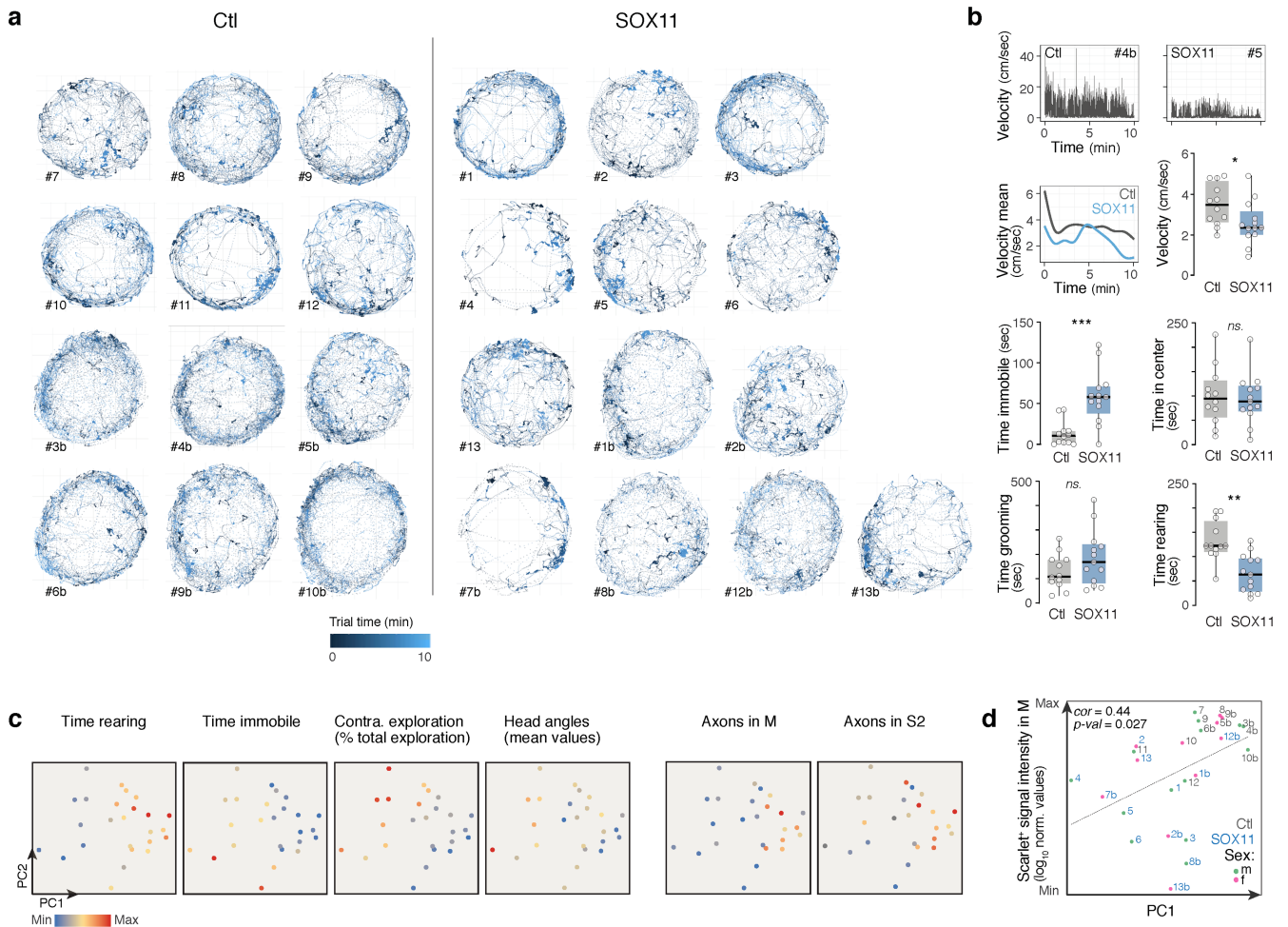
Extended Data Fig. 9 | Altered intracortical somatosensory connectivity impairs exploratory behavior. a, Trajectories in open-field arena for each pup. The name of each pup is indicated (#). Two different arenas were used accounting for slight differences in trajectory shapes. **b**, SOX11-overexpressing (SOX11) pups display decreased velocity (unpaired *t*-test: *P* value = 0.0355), decreased rearing (unpaired *t*-test: *P* value = 0.0007), increased time immobile (Mann Whitney test: *P* value = 0.0005) but similar time spent in the center of the open-field (unpaired *t*-test: *P* value = 0.9030) or time spent grooming (unpaired *t*-test: *P* value = 0.1990), compared to control pups (Ctl) (*n* = 12 Ctl pups; *n* = 13 SOX11 pups). Box plots indicate median ± s.d. and interquartile range. **c**, Feature plots of specific behavioral and anatomical parameters. **d**, Correlation between principal component (PC) 1 and the Scarlet axonal signal measured in M for each pup.

nature portfolio

Corresponding author(s): Denis Jabaudon

Last updated by author(s): Jul 9, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Single-cell RNA sequencing data were acquired using C1 system from Fluidigm and HiSeq 2500 sequencer. Gene ontology data were from QuickGO from EMBL-EBI Hinxton public database. TcoF-DB v2 database was used to identify transcription factors (© Massey University & CBRC@KAUST 2010-2017 // v2.2.2).
Data analysis	The custom code for data processing is provided here: https://github.com/pradosj/docker_sindbis. R studio (version 1.1.383) was used for all further analyses with the following packages: GenomicAlignments (summarizeOverlaps method, v1.14.1), Biostrings (trimLRpatterns, v2.46), Seurat (v3.1.1), bmrm (Bundle Methods for Regularized Risk Minimization Package, v3.1), and EasyPubMed (v2.13). GraphPad Prism (v8) was used for statistical analyses. Imaris (v8) was used for single cell morphological reconstructions. Fiji (ImageJ v2.1.0) was used for all image analyses.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Single-cell RNA sequencing data (GEO): MAPseq: GSE118681; Retrobeads: GSE116944; ConnectID: GSE156080.
Gene ontologies: QuickGO from EMBL-EBI Hinxton public database (https://www.ebi.ac.uk/QuickGO/); GSEA (https://www.gsea-msigdb.org/gsea/index.jsp)
Transcription factors: TcoF-DB v2 database (https://tools.sschmeier.com/tcof/home/)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not determined in advance but corresponds to what is being used in the field. For the majority of the experiments, based on consistency of results between pups, a sample size of 3-5 litters was considered sufficient. For behavioral experiments, a sample size of 12-13 pups was considered sufficient.
Data exclusions	The quality control for cells is described in detail in the Methods and in Extended Data Figures 3-4; all cells excluded from the collection point until final analysis are accounted for.
Replication	Experiments were performed on several animals and the analysis reveals a high degree of replicability across animals.
Randomization	Does not apply. Experiments of different conditions were performed on different days based on animal availability.
Blinding	All image acquisitions and analyses of SOX11 intensity in retrogradely-labeled neurons and axon signal upon SOX11 overexpression in cortical targets were performed blindly. All behavioral experiments and associated analyses were performed blindly by an independent experimenter. Blinding was not performed in other experiments due to the unambiguous nature of measurements and systematic analyses used in these experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	chicken anti-GFP (Invitrogen, #A10262, 1:2000, LOT: 1857662), rabbit anti-CUX1 (Santa Cruz, #SC-13024, 1:250, LOT: E2314), rat anti-RFP (Chromotek, #5F8, 1:500, LOT: 90228002AB),

rabbit anti-SOX11 (Millipore, #ABN105, 1:500; LOT: 3099692),
alkaline phosphatase-conjugated anti-DIG antibody (Roche, #11093274910, 1:2000, LOT: 32871921)

Validation

All antibodies used are commonly used in the field and have been validated in previous publications/by the manufacturer.
References and manufacturer validations can be found here:
chicken anti-GFP (Invitrogen, #A10262): https://www.thermofisher.com/antibody/product/GFP-Antibody-Polyclonal/A10262
rabbit anti-CUX1 (Santa Cruz, #SC-13024): https://www.scbt.com/scbt/product/cdp-antibody-m-222
rat anti-RFP (Chromotek, #5F8): https://www.chromotek.com/products/detail/product-detail/rfp-antibody-5f8/
rabbit anti-SOX11 (Millipore, #ABN105): https://www.merckmillipore.com/CH/fr/product/Anti-Sox11-Antibody,MM_NF-ABN105?
ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1
alkaline phosphatase-conjugated anti-DIG antibody (Roche, #11093274910): https://www.sigmaaldrich.com/CH/en/product/
roche/11093274910

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Mus musculus (CD1 and C57 Bl6J strains) P5, P7, P9, P14, P15, P16, P17 pups were used (both male and female). All mice were
housed in the institutional animal facility under standard 12 h/12 h light/dark cycles with food and water ad libitum.

Wild animals

The study did not involve wild animals.

Field-collected samples

The study did not involve field-collected samples.

Ethics oversight

The experimental procedures described here were conducted in accordance with the Swiss laws and previously approved by the
Geneva Cantonal Veterinary Authority.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Dissociated neurons from mouse primary somatosensory cortex at P7, P9 and P14.

Instrument

Beckman Coulter Moflo Astrios FAC-sorter

Software

Summite

Cell population abundance

We collected 5'000 to 10'000 GFP+ cells per collected brain.

Gating strategy

Singlet GFP+/Hoechst+ cells were sorted according to their Forward and Slide scattering properties, and their negativity for
Draq7TM (Viability dye, far red DNA intercalating agent, Beckman Coulter).

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.