Early emergence of cortical interneuron diversity in the mouse embryo

Da Mi,1,2* Zhen Li,2* Lynette Lim,1,2 Mingfeng Li,3 Monika Moissidis,1,2 Yifei Yang,4 Tianliuyun Gao,3 Tim Xiaoming Hu,5,6 Thomas Pratt,4 David J. Price,5 Nenad Sestan,3,5†; Oscar Marin1,2,†

GABAergic interneurons (GABA, γ-aminobutyric acid) regulate neural-circuit activity in the mammalian cerebral cortex. These cortical interneurons are structurally and functionally diverse. Here, we use single-cell transcriptomics to study the origins of this diversity in the mouse. We identify distinct types of progenitor cells and newborn neurons in the ganglionic eminences, the embryonic proliferative regions that give rise to cortical interneurons. These embryonic precursors show temporally and spatially restricted transcriptional patterns that lead to different classes of interneurons in the adult cerebral cortex. Our findings suggest that shortly after the interneurons become postmitotic, their diversity is already determined in their diverse transcriptional programs, which subsequently guide further differentiation in the developing cortex.

The mammalian cerebral cortex contains more than two dozen GABAergic cell types (GABA, γ-aminobutyric acid) with particular morphological, electrophysiological, and molecular characteristics (1–3). Interneuron diversity has evolved to increase the repertoire of cortical computational motifs through a division of labor that allows individual classes of interneurons to control information flow in cortical circuits (4–6). Although a picture about cortical interneuron cell types is emerging (7, 8), the mechanisms that generate interneuron diversity remain controversial. One model proposes that interneurons acquire the potential to differentiate into a distinct subtype at the level of progenitors or shortly after becoming postmitotic, before they migrate; the competing model postulates that interneuron identity is established relatively late in development, after they have migrated to their final location, through interactions with the cortical environment (9).

To study cell diversity in the germinal zones of cortical interneurons (10), we dissected tissue from three regions in the mouse subpallium, the dorsal and ventral medial ganglionic eminence (dMGE and vMGE, respectively) and the caudal ganglionic eminence (CGE), across two stages that coincide with the peak of neurogenesis for cortical interneurons [embryonic (E) days 12.5 and 14.5] (11) (Fig. 1A). We prepared single-cell suspensions and sequenced the transcriptome of individual cells, which, following quality control (fig. S1, A to E), led to a final data set of 2003 cells (fig. S1F), covering on average of about 3200 genes per cell. We performed regression analysis on these cells to remove the influence of cell cycle-dependent genes in cell type identification (fig. SIG).

We used principal components analysis (PCA) to identify the most prominent sources of variation. We found that developmental stage and anatomical source contribute to cell segregation (Fig. 1, B and C, and figs. S2 and S3). To distinguish between dividing and postmitotic cells, we conducted random forest (RF) feature selection and classification, starting with a list of established genes to sort cells into these categories (Fig. 1D). Subsequently, we reduced the dimensionality of our data using t-SNE (t-distributed stochastic neighbor embedding) to visualize the segregation of progenitor cells from neurons (figs. S4A and S5A). These analyses revealed gene expression patterns that distinguish progenitor cells and neurons at each developmental stage (figs. S4B and S5B).

We took a semisupervised clustering approach to explore variation across all progenitor cells and identified progenitor clusters with characteristic regional and developmental patterns (fig. S6, A to D). This analysis revealed a prominent temporal segregation of progenitor clusters (fig. S6, B and D), which suggests that progenitor cells in the ganglionic eminences (GE) may have a rapid turnover during embryonic development. To identify distinctive features of E12.5 and E14.5 progenitor cells, we further investigated progenitor cell diversity at each stage individually. We first used RF feature selection and classification, starting with a list of established genes to distinguish between ventricular zone (VZ) radial glial cells and subventricular zone (SVZ) intermediate progenitors (12, 13) (fig. S7A). We then carried out semisupervised clustering and distinguished VZ and SVZ progenitor clusters at both developmental stages (Fig. 2A and fig. S7B), independent of their cell cycle state (fig. S8), and found characteristic patterns of gene expression (fig. S9). Cross-validation using MetaNeighbor (14) confirmed cluster robustness and identity (fig. S10A). Many progenitor clusters found at E12.5 did not seem to have a direct transcriptional equivalent at E14.5 (fig. S10B), which reinforces the notion that the GE contains highly dynamic pools of progenitor cells during development.

Analysis of progenitor cell clusters confirmed that radial glial cells and intermediate progenitors have distinct identities across different regions of the subpallium, with characteristic and often complementary expression of transcription factors (e.g., Nkx2-1 and Pax6 in VZ, Lhx6 and Foxp2 in SVZ) (Fig. 2B and figs. S11 to S14). Thus, based on transcriptomic signatures, the diversification of progenitor cells in the GE seems to emerge primarily within the highly neurogenic SVZ.

We next turned our attention to the neurons that are being generated in the GE during this temporal window of high progenitor cell diversity. The MGE and CGE generate different groups of cortical interneurons (17–19). Most parvalbumin (PV)–expressing and somatostatin (SST)–expressing interneurons are born in the MGE, whereas the CGE is the origin of vasoactive intestinal peptide (VIP)–expressing interneurons and neurogliofascicular (NDNF) cells (20). We took a completely unsupervised approach to explore the emergence of neuronal diversity in the GE. Unbiased clustering of all neurons identified 13 groups of newborn neurons with distinctive genetic expression profiles, as well as specific temporal and regional identities (figs. S15, A to D, and S16). This analysis revealed that regional identity segregates more clearly among E14.5 neuronal clusters (fig. S15C), which suggests that neurons become more transcriptionally heterogeneous over time. Similar results were obtained when neuronal clusters were identified for both stages separately (fig. S17).

GO (Gene Ontology) enrichment analysis revealed different states of maturation across neuronal clusters (fig. S15F), which reiterates that some aspects of this diversity might be linked to the differentiation of newborn neurons and not cell identity. Analysis of the expression of region- and cell type–specific genes revealed the emerging signature of the main groups of cortical interneurons (fig. S15F). For example, clusters primarily populated by MGE-derived cells can be further segregated into those with features of SST+ interneurons (N3 and N4) and those without (N1, N2, and N9), which presumably include neurons that will differentiate into PV+ interneurons. The profile of emerging CGE-specific interneuron classes, such as those characterized by the expression of Meis2 (21), is also

1Centre for Developmental Neurobiology, Institute of Psychiatry, Psychology, and Neuroscience, King’s College London, London SE1 9UL, UK. 2Medical Research Council Centre for Neurodevelopmental Disorders, King’s College London, London SE1 9UL, UK. 3Department of Neuroscience and Kavli Institute for Neuroscience, Yale School of Medicine, New Haven, CT 06510, USA. 4Biomedical Sciences, University of Edinburgh, Edinburgh EH8 9XD, UK. 5Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02446, USA. 6Department of Physiology, Anatomy, and Genetics, University of Oxford, Oxford, UK.

*These authors contributed equally to this work. †Corresponding author. Email: oscar.marin@kcl.ac.uk (O.M.); nenad.sestan@yale.edu (N.S.)

Downloaded from http://science.sciencemag.org/ on July 8, 2021
delineated at this stage. This analysis also revealed that the CGE gives rise to neurons with molecular profile of SST+ interneurons (N13), which reinforces the view that this anatomical region contains a molecularly heterogeneous pool of progenitor cells (15).

The adult mouse cerebral cortex contains more than 20 distinct classes of interneurons with characteristic transcriptional profiles (7, 8). We asked whether any of these classes of interneurons would be identifiable shortly after becoming postmitotic in the GE. To this end, we used a publicly available single-cell RNA sequencing (RNA-seq) data set of 761 adult GABAergic interneurons from the adult mouse visual cortex (8) and identified highly variable genes in both adult and embryonic data sets. We employed the resulting data set to identify the features that best represent each of the 23 interneuron cell types found in the adult mouse cortex (8). We then carried out RF feature selection and classification based on those features to assign the identity of adult interneurons into distinct cell types. We were unable to identify all cell types originally described in the adult data set (8), which suggests a difference in transcriptomic and cell-type diversity between embryo and adult. We then used the identified adult interneuron cell types to annotate the embryonic data set using the RF classification workflow and found six prospective interneuron subtypes among embryonic neurons (Fig. 3A and fig. S18). Cross-data set validation between the emerging embryonic subtypes and adult interneurons confirmed the robustness of these annotations (fig. S19).

We used a second, independent approach to assign embryonic neurons to adult interneuron subtypes. In brief, we conducted canonical correlation analysis (CCA) to identify the sources of variation that are shared between embryonic and adult interneurons. We then assigned prospective identities to embryonic neurons based on transcriptional similarity with adult interneurons. This analysis provided evidence for early cell type differentiation: All 11 classes of cortical interneurons were identified among embryonic neurons (Fig. 3C), which exhibit characteristic patterns of gene expression (Fig. 3D and fig. S20) and robustness in cross-validation analyses (fig. S21). Comparison between the two independent approaches identified eight conserved interneuron subtypes among the assigned
embryonic neurons (Fig. 3E). Analysis of the contribution of E12.5 and E14.5 neurons to these identities revealed timing biases for the generation or maturation of some interneuron subtypes (Fig. 3F and fig. S22). Altogether, these results strongly suggested that interneurons exhibit a great diversity of transcriptional signatures shortly after becoming postmitotic in the GE.

We hypothesized that the patterns of gene expression identified in progenitor cells and newborn neurons delineate specific lineages of cortical interneurons. To investigate this idea, we limited our analysis to embryonic neurons that were assigned to the same subtype identity by both RF and CCA methods, which we named “consensus” neurons and which belong to three interneuron subtypes: PV1, SST1, and SST2. We conducted MetaNeighbor analysis to identify possible links between progenitor cell clusters and consensus neurons. This analysis revealed putative SST+ and PV+ progenitor cell clusters at E12.5 and E14.5 (Fig. 4A and fig. S23).

We then carried out differential gene expression between E12.5 progenitor clusters P5 and P7 (Fig. 4B), which exhibited the highest association with PV1 and SST1, respectively (Fig. 4A). We found early PV (Ccnd2 and St18) and SST (Epha5, Cdk14, and Map2) markers in these progenitor pools (Fig. 4C and D), which are subsequently maintained in specific subtypes of
newborn interneurons (Fig. 3, A and D). To validate these observations, we investigated the function of \textit{Maf} in the delineation of MGE interneuron lineages. We infected progenitor cells in \textit{Nkx2-1-Cre} embryos with conditional retroviruses expressing Cre-dependent control or \textit{Maf} vectors during the period of SST$^+$ interneuron production (E12.5) and explored the identity of labeled interneurons in the cortex of young adult mice (Fig. 4E and fig. S24, A and B). We found that widespread expression of \textit{Maf} in MGE progenitors increases the relative proportion of SST$^+$ interneurons at the expense of PV$^+$ cells (Fig. 4F and fig. S24C). Conversely, conditional loss of \textit{Maf} from MGE progenitor cells decreases the density of cortical SST$^+$ interneurons (Fig. 4, G and H). We also observed that overexpression of \textit{Maf} at the peak of PV neurogenesis (E14.5) represses PV$^+$ interneuron fates (fig. S24, D and E). Altogether, these results indicated that \textit{Maf} regulates the potential of interneurons to acquire SST$^+$ interneuron identity.

Our study reveals that GABAergic interneurons have a propensity toward a defined fate long before they occupy their final position in the cerebral cortex during early postnatal development. This suggests that interneuron diversity does not emerge in response to activity-dependent mechanisms in the cortex (1, 9) but rather is established early, before these cells reach the cortex, by specific transcriptional programs that then unfold over the course of several weeks. Activity-dependent mechanisms undoubtedly influence development, maturation, and plasticity of cortical interneurons (24–26), but most aspects that are directly linked to the functional diversity of

Fig. 3. Emergence of cortical interneuron diversity in the ganglionic eminences. (A) Heat map showing average expression of differentially expressed (DE) genes among six classes of interneurons identified by RF classification of embryonic neurons. (B) Integration of embryonic neurons and adult cortical interneurons in t-SNE space following canonical correlation analysis (CCA). (C) Embryonic neurons assigned to specific interneuron lineages by k-nearest neighbor (KNN) analysis are depicted in the same t-SNE space. Unassigned embryonic neurons are omitted. (D) Heat map illustrating the expression of DE genes among the 11 classes of interneurons identified by CCA. (E) Heat map of mean AUROC (area under the receiver operating characteristic curve) scores for assigned interneuron cell types using the two independent approaches (RF and CCA). AUROC scores of eight conserved interneuron subtypes: PV1$_{RF}$-PV1$_{CCA}$ = 0.95; PV1$_{RF}$-PV3$_{CCA}$ = 0.90; PV1$_{RF}$-PV4$_{CCA}$ = 0.90; SST1$_{RF}$-SST1$_{CCA}$ = 1; SST2$_{RF}$-SST2$_{CCA}$ = 0.95; VIP2$_{RF}$-VIP2$_{CCA}$ = 0.90; VIP2$_{RF}$-VIP3$_{CCA}$ = 0.85; NDNF1$_{RF}$-NDNF1$_{CCA}$ = 0.7. (F) Histogram illustrating the relative contribution of E12.5 and E14.5 neurons to conserved interneuron subtypes.
cortical interneurons seem to be intrinsically determined (14, 27).

Our analysis identifies early markers for many different classes of cortical interneurons, whose functional validation may eventually illuminate the mechanisms regulating the differentiation of GABAergic interneurons into specific subtypes and, through comparative analyses, inform the use of stem cell biology for the generation of distinct classes of human cortical interneurons (28, 29). Thus, core aspects of interneuron identity are drafted early in development, forming the foundation on which later interactions with other neurons must function.

REFERENCES AND NOTES

ACKNOWLEDGMENTS
We thank S. E. Bae for technical assistance and M. Fernández for general lab support. We are grateful to M. Zhong at Yale Stem Cell Center Genomics and Bioinformatics Core for conducting RNA sequencing. N. Flames and B. Rico for critical reading of the manuscript; S. Anderson, C. Birchmeier, F. Cage, and G. Fishell for reagents and mouse strains; and members of the Marín, Rico, and Sestan laboratories for stimulating discussions and ideas. Funding: This work was supported by grants from the Wellcome Trust Investigator. Author contributions: D.M. and O.M. designed the study; single-cell experiments were performed by D.M., Z.L., and O.M., with feedback from all authors. Competing interests: None declared. Data and materials availability: Sequencing data have been deposited at the National Center for Biotechnology Information BioProjects Gene Expression Omnibus and are accessible through GEO Series accession number GSE109196. c-Maf flx/− mice are available from C. Birchmeier under a material agreement with the Max-Delbrück-Centrum für Molekulare Medizin (Berlin). The supplementary materials contain additional data. All data needed to evaluate the conclusions in the paper are present in the paper or the supplemental materials.

SUPPLEMENTARY MATERIALS
www.sciencemag.org/content/360/6384/81/suppl/DC1
Materials and Methods.
Figs. S1 to S4.
Tables S1 to S4.
References (30–44).
6 December 2017; accepted 14 February 2018.
Published online 22 February 2018.
10.1126/science.aar6821

Fig. 4. Maf regulates SST+ interneuron fate. (A) AUROC values for putative lineages linking E12.5 progenitor clusters (P) with specific interneuron subtypes (AUROC scores above 0.9). (B) t-SNE plot illustrating progenitor cell clusters at E12.5. Two SVZ clusters, P5 and P7, are highlighted by color. (C) Violin plots for selected DE genes between P5 and P7 clusters. (D) RNAscope labeling of MGE SVZ progenitor cells with a Maf probe. (E) Coronal sections through the somatosensory cortex of P21 mice after viral infection with Gfp or Maf-P2A-Gfp retroviruses in the MGE at E12.5. (F) Quantification of the proportion of PV−/SST+, PV−, and SST+ interneurons; n = 5; t2 test, ***P < 0.001. Post-hoc analysis was performed with binomial pairwise comparison with adjusted P value by Bonferroni correction; PV−/SST+ versus PV−, ***P < 0.001; PV− versus SST+, **P < 0.01. (G) Coronal sections through the somatosensory cortex of P21 control and conditional Maf mutants. (H) Quantification of the density of GFP+/SST+ and GFP+/PV− interneurons; n = 4, one-way analysis of variance with Tukey correction, *P < 0.05. Scale bars, 15 μm (D) and 100 μm (E) and (G).
Early emergence of cortical interneuron diversity in the mouse embryo

Da Mi, Zhen Li, Lynette Lim, Mingfeng Li, Monika Moissidis, Yifei Yang, Tianliuyun Gao, Tim Xiaoming Hu, Thomas Pratt, David J. Price, Nenad Sestan and Oscar Marín

Science 360 (6384), 81-85.
DOI: 10.1126/science.aar6821 originally published online February 22, 2018

Embryonic hints of adult diversity

The adult brain contains dozens of different types of interneurons that control and refine neuronal circuits. Mi et al. used single-cell transcriptomics to investigate when these subtypes emerge during interneuron development in the mouse. Transcriptomes of embryonic interneurons showed similarities to adult classes of differentiated interneurons, thus dividing the immature embryonic interneurons themselves into classes. Nearly a dozen classes of embryonic neurons could be identified soon after their last mitosis by transcriptomic similarity with known classes of adult cortical interneurons. Thus, the fate of embryonic interneurons can be read in their transcriptomes well before the neurons migrate and reach their final sites of differentiation and circuit integration.

Science, this issue p. 81