Title : Inferring Ligand-Receptor Interactions between neuronal subtypes during mouse cortical development

Authors: Rémi Mathieu[1], Léa Corbières[1], Tangra Draia-Nicolau[1], Annousha Govindan[1], Vianney Bensa[1], Emilie Pallesi-Pocachard[1], Lucas Silvagnoli[1], Alfonso Represa[1], Carlos Cardoso[1], Ludovic Telley[2]*, Antoine de Chevigny[1]*

Affiliations:

[1] INMED, INSERM, Aix Marseille University, Marseille, France

[2] Department of Fundamental Neurosciences, University of Lausanne, Lausanne, Switzerland

*Correspondence:

Antoine de Chevigny

antoine.de-chevigny@inserm.fr

Ludovic Telley

ludovic.telley@unil.ch

`

Abstract: The cerebral cortex hosts a diverse array of excitatory glutamatergic and inhibitory GABAergic neuron types, each characterized by distinct positional and synaptic connectivity patterns. However, the molecular mechanisms orchestrating this precise organization remain largely unknown. To identify ligand-receptor (LR) pairs regulating interactions and connectivity among cortical neurons during embryonic and postnatal development, we analyzed the transcriptional dynamics of all genes across major cortical neuron subtypes at 17 developmental time points using single-cell transcriptomics. From these data, we constructed a comprehensive bioinformatic atlas that inferred significant LR-mediated interactions between glutamatergic and GABAergic neurons throughout cortical maturation. This atlas not only corroborated known interactions but also enabled the discovery of novel regulators, identifying two cadherin superfamily members as key mediators of perisomatic inhibition in deep and superficial layer excitatory neurons by parvalbumin-expressing basket cells. These findings underscore the power of large-scale transcriptional profiling to unravel fundamental molecular mechanisms driving cortical circuit assembly.

`

Main Text:

The mammalian cerebral cortex, a six-layered structure located at the brain surface, is responsible for executing the most complex cognitive functions. To achieve these functions, it must undergo highly orchestrated and conserved developmental processes ultimately leading to the finely organized communication between two primary neuron classes: excitatory glutamatergic neurons (GlutNs) and inhibitory GABAergic neurons (GABANs). Both GlutNs and GABANs are further subdivided into numerous subtypes, a diversity primarily uncovered through single-cell RNA sequencing (scRNA-seq). Tens of cortical GlutN and GABAN subtypes have been identified [1–5], which (1) exhibit specific morphological, electrophysiological and transcriptomic properties, (2) preferentially localize to certain cortical layers, (3) form synaptic connections with specific pre and post-synaptic partners, preferentially targeting specific subcellular compartments and (4) play specialized roles in cortical computation. The resulting cortical arrangement and associated synaptic connectivity, involving hundreds to thousands of potential neuron pairs, is responsible for cortical function. Even a disruption in the communication between just one pair of neuron subtypes at a particular developmental stage might lead to cortical dysfunction and subsequent neurodevelopmental disorders [6]. Therefore, understanding how relative physical positions and synaptic connectivity rules between the hundreds of potential neuron pairs are established and regulated is crucial for advancing developmental and clinical neuroscience.

Over the past decade, substantial efforts have been directed toward elucidating the spatial relationships and synaptic connectivity rules among cortical neuronal subtypes. In particular, a wealth of groundbreaking studies has provided high-resolution insights into the laminar organization and connectivity patterns of cortical GlutNs and GABANs in the adult mouse brain. These studies have demonstrated the remarkable specificity, stereotypy, and evolutionary conservation of this intricate cellular and subcellular architecture across mammals [7–10]. Despite this progress, the molecular mechanisms underlying corticogenesis that culminate in such a finely orchestrated final organization remain largely elusive. Only a handful of pioneering studies have begun to uncover specific molecules driving the cellular interactions and processes shaping cortical structure during development [11–13].

It is likely that neuron-neuron communications via ligand-receptor (LR) interactions play a key role in general [11–13], as these interactions are critical for the development of many tissues. This hypothesis is supported by several studies suggesting that GlutNs may non-cell autonomously influence the recruitment of specific GABAN subtypes. Indeed, GlutN subtypes settle into their final positions earlier than the synchronically generated GABANs [14], and altering specific GlutN subtype identities during development modifies the allocation and synaptic connectivity of corresponding GABAN subtypes [15–17].

In this study, we leverage both newly generated and previously published scRNA-seq datasets from mice to elucidate the molecular mechanisms underlying subtype-specific neuron communication during cortical development. First, we tracked gene expression dynamics across all GlutN and GABAN subtypes throughout corticogenesis. Then, we constructed a comprehensive

atlas to infer LR interactions between these neuronal subtypes. Finally, we validated the utility of our atlas by (1) confirming known LR interactions, (2) bridging gaps in existing knowledge by identifying receptors involved in well-characterized neuron-neuron interactions, and (3) uncovering novel LRs within the cadherin family that are crucial for perisomatic inhibition in deep versus superficial cortical layers. The data can be explored through an interactive Shiny App, currently accessible via the instructions provided in the README file, which will be available as a web-based interface upon publication of the paper.

Comprehensive single neuron transcriptomics atlas covering mouse somatosensory cortex development

To infer LR-mediated cell-cell interactions during mouse neurodevelopment, our strategy was to cross-reference single-cell RNA sequencing (scRNA-seq) transcriptomic data from all cortical neurons across all multiple stages with publicly-known LR information. The purpose of scRNA-seq was to both identify neuronal subtypes and predict LR expression over time.

First, we generated a comprehensive scRNA-seq transcriptomic dataset of all cortical neurons throughout somatosensory corticogenesis. While prenatal and adult stages are well-documented in the mouse, the P0-P30 period remains underexplored, despite being crucial for synaptogenesis. We conducted scRNA-seq and snRNA-seq at six key stages within this timeframe: P0-P2 (radial migration and laminar allocation of GABANs [18,19]), P5-P8 (programmed cell death of GlutNs and GABANs [20–22]) , and P16-P30 (circuit refinement and synaptogenesis completion). We integrated to this dataset previously published scRNA-seq data covering earlier embryonic (E11.5-E18.5) and adult stages (P53-P102) [3,23–28], including ganglionic eminences (GEs) to capture early transcriptional development of future cortical GABANs [25–28] (see Methods). In total, our analysis spans 17 time points from E11.5 to adulthood (Fig. 1A).

After stringent quality control and filtering (Fig. S1A-B and Methods), we first aimed to identify all postmitotic neuronal cell types and track their transcriptional dynamics over cortical development, from E11.5 to adulthood. For final cell type nomenclature, we chose as reference the most comprehensive resource of transcriptomic cell-types in the somatosensory cortex that existed before December 2023*, i.e.*, Yao et al. (2021) [3] (from now on referred to as AllenRef21), corresponding to the adult stage (P53-P102) (Fig. S1A). Assigning cellular identities in early age datasets proved difficult with simple methods given that transcriptional signatures of specific cell-types evolve drastically during development [23,26]. To overcome this difficulty, we developed a hierarchical pipeline assigning identities at increasing resolution levels, similar to the method used by the Allen Institute [3] (Fig. S1A).

First, we assigned cell identities at a broad, low resolution level referred to as the *class* level (Glutamatergic, GABAergic, Non-Neuronal, Immature/Migrating, dorsal pallium progenitor and subpallium progenitor) [3]. For *classes* of embryonic and neonatal datasets (E11.5 to P5), we took as references the datasets of Di Bella et al. [24] (largest published perinatal dataset for somatosensory

cortex) and Bandler et al. [27] (corresponding to early GE GABANs destined to the cortex). For P8 to P30, because cells were transcriptionally more similar to the adult reference, we took directly as the reference the adult dataset AllenRef21 [3] (Fig. S1A). To test our hypothesis that a ligand-receptor (LR) code between GlutNs and GABANs plays a key role in the establishment of cortical microcircuits, we focused our downstream analyses on post-mitotic GlutN and GABAN cells (Fig. 1A, B). In total, 182,084 post-mitotic neurons were further analyzed. To analyze biological variation in gene expression and get rid of technical variation due to different studies, RNA-seq methods (scRNA-seq and snRNA-seq) and sequencing platforms (Fig. 1C), we integrated all data with the Seurat SCT workflow (Fig. S1C). The resulting integrated dataset was split into *classes*, i.e., GlutNs and GABANs. For each *class*, cell identities were further delineated at increasing resolutions; they were determined first at the *subclass* level, and then at the *supertype* level in each subclass, as defined in AllenRef21 [3]. A limitation of the *supertype* level is that it does not correspond to known morpho-functionally identified cortical neuronal subtypes. To identify biologically meaningful cell types in our analysis we leveraged the fact that Yao et al. provided the correspondence between *supertypes* from scRNA-seq studies and morpho-electro-connectomic types (*mec-types*) from patch-seq studies [1–3,29] (Fig. S2, S3). Pooling *supertypes* belonging to the same *mec-type* allowed us to reach a biologically meaningful *cell-type* annotation (Fig. S4). Ultimately, our analysis led to the identification of 159 290 cells, representing 27 *cell-types* (Fig. 1D) including 11 GlutN *cell-types* grouped in 3 *families* – intratelencephalic (IT), extratelencephalic (ET), Other GlutNs - and 16 GABAN *cell-types* grouped in 5 *families* - Lamp5, Sst, Pvalb, Vip and Other GABANs - (Fig. 1B-D, Fig. S4).

In the UMAP embedding of all neurons from 17 ages, each GlutN and GABAN *cell-type* exhibited a largely continuous temporal gradient of transcriptional variation across corticogenesis, with gene expression profiles of individual cells correlating closely with their differentiation states / mouse developmental age (Fig. 1B, C). the integration and annotation efforts presented here establish a valuable new reference for the developmental neurobiology community.

`

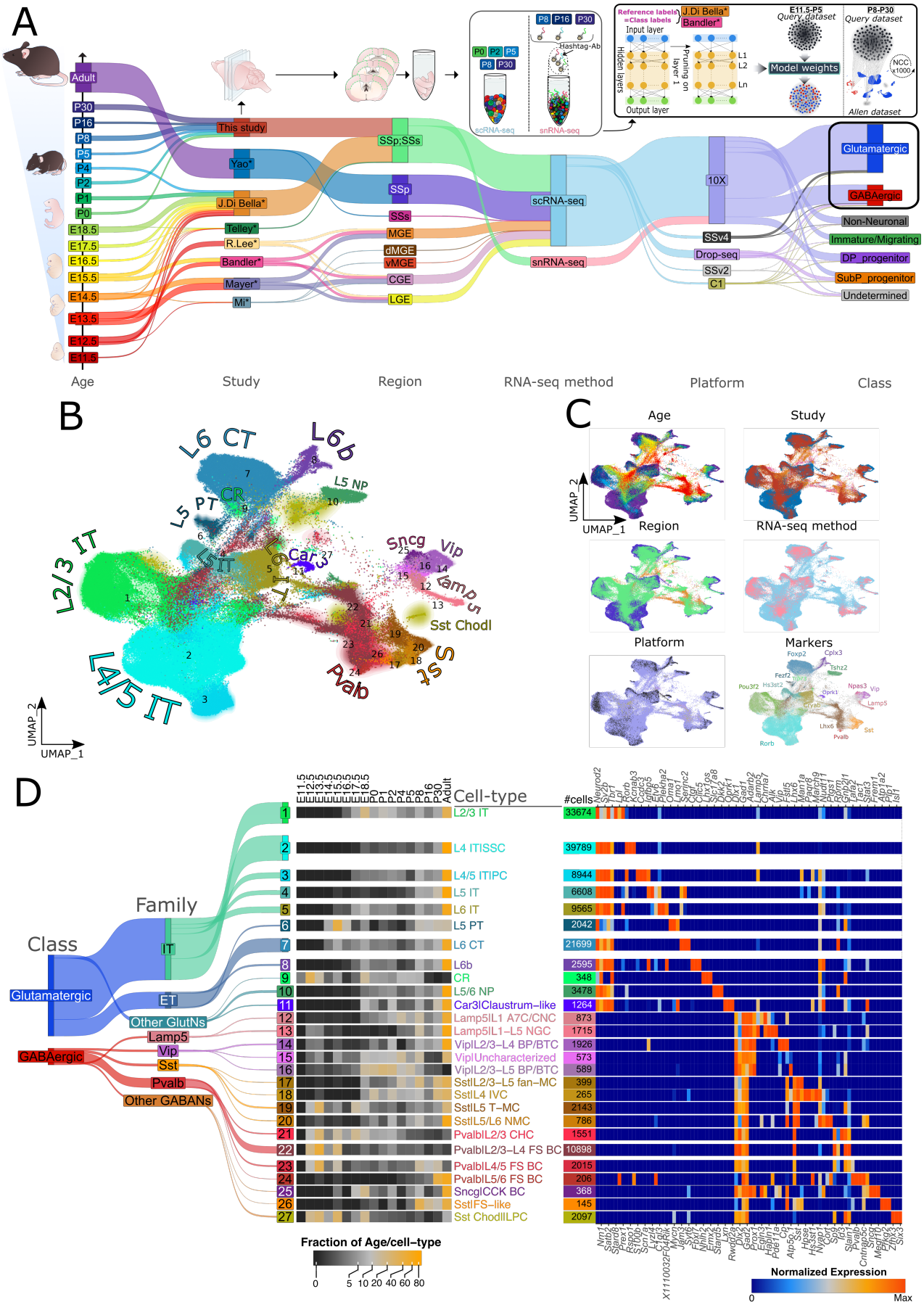

A
Adult
P30
P16
P8
P5
P4
P2
P1
P0
E18.5
E17.5
E16.5
E15.5
E14.5
E13.5
E12.5
E11.5
This study
Yao*
J.Di Bella*
Telley*
R.Lee*
Bandler*
Mayer*
Mi*
SSp;SSs
SSp
SSs
MGE
dMGE
vMGE
CGE
LGE
scRNA-seq
snRNA-seq
10X
SSv4
Drop-seq
SSv2
C1
Glutamatergic
GABAergic
Non-Neuronal
Immature/Migrating
DP_progenitor
SubP_progenitor
Undetermined
Age
Study
Region
RNA-seq method
Platform
Class
B
L6 CT
L6b
L5 NP
L5 PT
CR
L2/3 IT
L5 IT
Car3
Sncg
Vip
Lamp5
Sst Chodl
Sst
Pvalb
L4/5 IT
UMAP_2
UMAP_1
C
Age
Study
Region
RNA-seq method
Platform
Markers
D
Cell-type
#cells
Class
Family
Glutamatergic
GABAergic
IT
ET
Other GlutNs
Lamp5
Vip
Sst
Pvalb
Other GABANs
L2/3 IT
L4 ITISSC
L4/5 ITIPC
L5 IT
L6 IT
L5 PT
L6 CT
L6b
CR
L5/6 NP
Car3|Claustrum-like
Lamp5|L1 A7C/CNC
Lamp5|L1-L5 NGC
Vip|L2/3-L4 BP/BTC
Vip|Uncharacterized
Vip|L2/3-L5 BP/BTC
Sst|L2/3-L5 fan-MC
Sst|L4 IVC
Sst|L5 T-MC
Sst|L5/L6 NMC
Pvalb|L2/3 CHC
Pvalb|L2/3-L4 FS BC
Pvalb|L4/5 FS BC
Pvalb|L5/6 FS BC
Sncg|CCK BC
Sst|FS-like
Sst Chod|LPC
Fraction of Age/cell-type
Normalized Expression
Max

Figure 1. Comprehensive neuronal transcriptomic atlas of mouse somatosensory cortex development. (A) Sankey diagram depicting the experimental paradigm for data collections and integration of published datasets. Yao*: Yao et al., 2021; J. Di Bella*: J. Di Bella et al., 2021, Telley*: Telley et al., 2019; R.Lee*: R.Lee et al., 2022; Bandler: Bandler et al., 2021; Mayer*: Mayer et al., 2018; Mi*: Mi et al., 2018; SSp: primary somatosensory cortex; SSs: supplemental somatosensory cortex; MGE: medial ganglionic eminence; dMGE: dorsal MGE; vMGE: ventral MGE; CGE: caudal ganglionic eminence; LGE: lateral ganglionic eminence; scRNA-seq: single-cell RNA sequencing; snRNA-seq; single-nuclei RNA sequencing; 10X: 10X Genomics; SSv4: Smart-seq version 4; SSv2: Smart-seq version 2; C1: Fluidigm C1. (B) UMAP visualization of post-mitotic neurons from 17 time points after integration. Cells are colored by *cell-type* assignment. (C) UMAP visualization is colored by age, study, region, RNA-seq method, sequencing platform, and gene markers. (D) Left: Sankey plot representing the different hierarchical levels of *cell-type* assignment; middle: heatmap representing the fraction of each cell type per age; right: heatmap representing some cell type-specific genes. "Normalized expression" represents the 25% trimmed mean log2(CPM + 1) expression normalized by row. See also Figs. S1-S4.

Temporal transcriptional dynamics of cortical neuron subtypes

To track the temporal maturation of each neuronal *cell-type* over migration, programmed cell death, dendritic/axonal development and synaptogenesis in a continuous rather than discrete way, we used an approach in which single cells were ordered on a linear path on the basis of their transcriptional profile [30] (see Methods) (Fig. S5A, Fig. S6A-B). This "pseudo-maturation" axis correlated well with the actual age of the cells (Fig. S5A, Fig. S6A-B), indicating gradual variation in gene expression over neuronal differentiation.

For each *cell-type*, we found genes showing significant variation along the "pseudo-maturation" axis and used them to display temporal gene dynamics in 6 consecutive waves (Fig. S5A, Fig. S6A-B). During early development (E11.5-E18.5), most *cell-types* showed transcriptional signatures related with cell-intrinsic properties (Fig. S5B, Fig. S6C), while later on (neonatal to P30), genetic programs related with cell-cell and cell-environment interactions predominated (Fig. S5B, Fig. S6C) [31,32]. These findings reveal the molecular transitions through which each neocortical cell type shifts from intrinsic programs to extrinsic, interaction-driven programs essential for their integration into the developing cortical network.

Spatial transcriptional gradients at given developmental ages

In adults, GlutN ITs exhibit a spatial gradient of transcriptional variation along the cortical sheet [3,33]. We investigated the existence of such spatial transcriptomic gradients for each cell *family* at

each age analyzed between E18.5 to P30. The correspondence between our datasets and published Patch-seq [1,2] and MERFISH [33] datasets allowed us to determine the median adult laminar distribution for each *cell-type* (Fig. S2 & S3). Using a similar approach as for pseudo-maturation, for each *family* that displayed spatial organization along the cortical sheet and for each age, we fitted a curve along UMAP continua, and called this axis the "pseudo-layer" score.

The spatial transcriptional gradient described for the IT *family* in adults [33] was also found at all earlier ages examined, from E18.5 to P30 (Fig. S7A). Strikingly, we also identified clear spatial transcriptional continua for ET (Fig. S7C), Sst (Fig. S8A), Pvalb (Fig. S8C), Vip (Fig. S9A) and Lamp5 (Fig. S9C) *families* at all ages examined. For each *family* and at every age, we identified genes with significant variation along pseudo-layers and used them to discretize gene dynamics in 6 spatial transcriptional waves. In GlutNs of neonatal mice (E18.5 to P5), the first waves, which corresponded to upper layer GlutNs (UL), displayed genetic signatures of immaturity such as intrinsic properties (Fig. S7, B and D); this was not the case for late waves, which correspond to deep layers (DL). This is consistent with UL GlutNs being born later than DL GlutNs - due to cortical inside-out patterning - and therefore maturing later [23]. In contrast, for GABANs, environment-sensing programs were already active by E18.5 (Fig. S8, B and D), supporting the hypothesis that these neurons mature later than the earlier-settled GlutNs and rely on signals from them to reach their final positions and establish proper connections [15–17,34].

Spatiotemporal transcriptional landscapes

On the basis of the pseudo-maturation and pseudo-layer scores, cells of each *family* were embedded within 2D graphs, so as to generate spatiotemporal transcriptional landscapes of gene expression (Fig. 2A-D and S10). For any gene queried in a given *family*, the 2D transcriptional map shows temporal dynamics in the x axis and spatial dynamics in the y axis (Fig. 2B-D). This allowed us to confirm the spatiotemporal expression of genes known to be involved in cortical layer patterning (Cux1, Fezf2, Reln, Fig. 2B-C) or in neuropeptidergic system maturation in specific neuronal *families* (Fig. 2D).

For validation, we examined whether the transcriptional landscapes of specific LRs aligned with their known functions in migration and/or synaptogenesis. We focused first on CXCL12 and its receptors CXCR4 and ACKR3, critical for the tangential migration of GABANs [35,36]. The literature clearly indicates that CXCL12 is expressed during the perinatal period by the meninges [37] and by immature subventricular zone/intermediate zone GlutNs [38]. Additionally, CXCL12 might also be expressed in CR cells [39], although this requires further confirmation. Meanwhile, CXCR4 and ACKR3 are consistently reported to be co-expressed on migrating GABANs [40,41], with CXCR4 also expressed in CR cells [40]. The postnatal extinction of CXCL12 expression has been proposed to drive the transition of GABANs from tangential to radial migration and cortical plate invasion (Fig. 2E) [35,36,42]. Our transcriptional landscapes corroborated these documented spatio-temporal expression patterns fully (Fig. 2E). Specifically, the enrichment of Cxcl12 in CR cells was confirmed, and the expression of Cxcl12, Cxcr4, and Ackr3 was restricted to early developmental

`

stages (E18.5-P0). This timing aligns with their established role in migration, while the absence of expression at later stages (P0 onward) supports their lack of involvement in synaptogenesis (Fig. 2E).

We then investigated a LR pair associated with synaptogenesis rather than migration. It is well established that the formation of synapses requires specific adhesion molecules, including the NRXN1-NLGN1 LR pair (Fig. 2E) [43,44]. Our landscapes showed Nrxn1 and Nlgn1 expression in all cell types during time windows consistent with synaptogenesis (Fig. 2E), supporting the role of these proteins in this process.

Analysis of transcriptional landscapes using PCA and k-means clustering allowed us to identify 15 archetypal gene expression profiles. These profiles span early to late developmental stages and map to neurons in deep versus superficial cortical layers (Fig. 2F). This analysis uncovered diverse spatiotemporal patterns: some genes were associated with early developmental processes like lamination, while others were linked to later processes such as synaptogenesis. Additionally, certain genes were specific to neuron types in superficial layers, others to deep layers, while some exhibited a combination of temporal and spatial specificity. Intriguingly, members of the Cadherin gene family were distributed across all 15 clusters (Fig. 2G), highlighting their extensive spatiotemporal diversity and roles throughout cortical development.

We have developed a Shiny App called *scLRSomatoDev* (weblink to be released) to facilitate the exploration of developmental gene expression functions. This app allows researchers to visualize the spatiotemporal expression patterns of any gene(s) across neuronal cell types during cortical development. Users can choose from various visual representations to explore the data, providing an intuitive and comprehensive view of gene expression dynamics (Fig. S11).

`

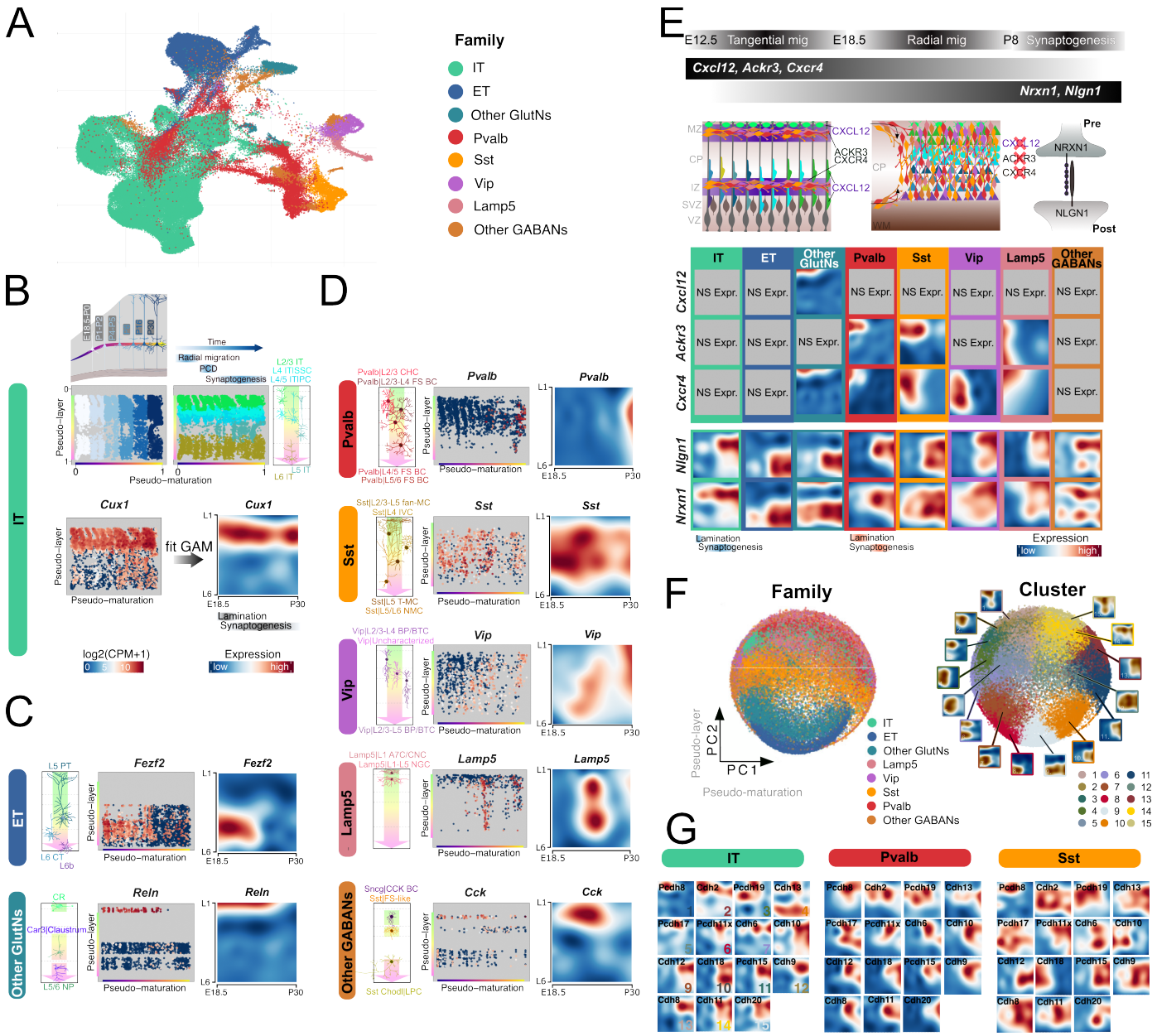

A
Family
IT
ET
Other GlutNs
Pvalb
Sst
Vip
Lamp5
Other GABANs
B
IT
Time
Radial migration
PCD
Synaptogenesis
Pseudo-layer
Pseudo-maturation
Cux1
fit GAM
Lamination
Synaptogenesis
log2(CPM+1)
Expression
low
high
C
ET
Fezf2
Other GlutNs
Reln
D
Pvalb
Sst
Vip
Lamp5
Other GABANs
Cck
E
E12.5 Tangential mig
E18.5
Radial mig
P8 Synaptogenesis
Cxcl12, Ackr3, Cxcr4
Nrxn1, Nlgn1
CXCL12
ACKR3
CXCR4
Pre
NRXN1
NLGN1
Post
NS Expr.
Lamination
Synaptogenesis
Expression
low
high
F
Family
Cluster
PC1
PC2
Pseudo-maturation
G
IT
Pvalb
Sst

Figure 2. Spatiotemporal transcriptional dynamics of all neuronal cell-types during somatosensory cortex development. (A) UMAP visualization of all neurons color coded with cell family labels. (B) Top: 2D map in which cells are embedded according to their pseudo-maturation (x-axis) and pseudo-layer (y-axis) scores for the IT *family*. Bottom: A generalized additive model (GAM) was applied to generate 2D maps for the spatiotemporal expression of genes (”transcriptional landscape”) throughout somatosensory cortex development; Cux1 expression is represented as an example. (C) Transcriptional landscapes for *Fezf2* in ET (top) and for Reln in other GlutNs (bottom). (D) Transcriptional landscapes for representative genes in each GABAN *family*. (E) Example of transcriptional landscape profiles for genes implicated in GABAN migration (*Cxcl12*, *Ackr3* and *Cxcr4*), and in synaptogenesis (*Nlgn1* and *Nrxn1*). (F) Gene map obtained by performing a PCA on all the significantly expressed genes colored by cell family (left) and by spatiotemporal cluster (right). Average cluster’s transcriptional landscapes are displayed around the gene map. (G) Varying spatiotemporal clusters for the Cadherin gene family among 3 neuronal subtypes. CP, cortical plate ; IZ, intermediate zone ; MZ, marginal zone ; SVZ, subventricular zone ; VZ, ventricular zone ; WM, white matter. See also Figs S5-S16.

An atlas for inference of LR mediated neuronal interactions during cortical development

The underlying hypothesis driving our study posits that interactions mediated by ligand-receptor (LR) pairs between neuronal subtypes, and in particular between GlutNs and GABANs, play a pivotal role in orchestrating the stereotypical construction of the cortex. We thus anticipated that these interactions would result in highly correlated spatiotemporal expression patterns between GlutN ligands and their corresponding GABAN receptors, and vice versa, compared to random gene pairs. To investigate this, we first assembled LR_DB_2025, the largest curated ligand-receptor database to date. This database was created by aggregating existing LR datasets and manual add-ons from published studies (Fig. 3A & S12, Table S1, see Methods). LR_DB_2025 contains 8,789 curated LR pairs, encompassing 1,421 ligands and 1,233 receptors. Each ligand and receptor were further categorized based on their known molecular functions or their association with specific neurodevelopmental stages (see Methods).

For LRs associated with developmental processes, such as cadherins, the transcriptional landscapes of ligands in IT neurons and their receptors in GABANs showed high correlations (Fig. 3B, Fig. S13-S16). This highlights the critical role of cadherin-mediated interactions in GlutN-GABAN communication during development, supporting hypotheses linking cadherins to areal specialization [34,45] and synaptic specificity [46].

LRs tied to migration, differentiation, cell death, cell adhesion, and synaptogenesis (Table S2, Methods) exhibited varying degrees of spatiotemporal correlation (Fig. 3B). Notably, synaptogenesis displayed the strongest LR correlations, suggesting it heavily depends on cell-cell communication, whereas processes like cell death or differentiation may rely more on intrinsic mechanisms (Fig. 3B).

`

To identify which LRs drive cell-cell interactions influencing cortical development, we constructed a comprehensive LR atlas that infers likely strong LR interactions between neuronal cell types from E18.5 to adulthood. This approach integrates temporal cell-gene expression matrices with data from our LR database (Fig. 3C and S17, see Methods).

First, we implemented and adapted to our dataset the *scSeqComm* method, which computes intercellular LR scores for each cell pair, based on the relative expression levels of ligands in source cells and receptors in target cells [47].

To improve accuracy, we created a cell pair score that prioritized LRs based on known developmental timing and connectivity patterns. The cell pair score is the modified sum of a development score (Fig. 3C & S17, Table S3) accounting for when cell types populate the cortex and a connectivity score that estimates synaptic connections based on physical proximity. The connectivity score was calculated from the morphological reconstruction of 1000 neurons in a cortical column from patch-seq studies (Fig. 3C, S17-18; Table S4). LRs with secreted ligands, we gave a lower weight to the cell pair score and a higher weight to the LR intercellular score to account for their ability to act over greater distances. To minimize false positives further, we only included LR pairs that were significantly present across contiguous developmental stages (Fig. S17).

To enhance robustness and minimize false positives in our LR predictions, we incorporated weights into each selected *cell-type* pair that reflect known patterns of developmental synchronicity and connectivity. Specifically, a *development* score (Fig. 3C & S17, Table S3) accounted for the varying times at which different cell types populate the cortex, affecting the likelihood of certain interactions depending on the mouse's age (Fig. S18). A *connectivity* score estimated the degree of synaptic connectivity between cell types, based on their physical proximity (Fig. 3C, S17-18; Table S4). These scores were integrated to generate a *cell-type* score, which was applied to each cell pair to incorporate prior knowledge into the threshold calculation for predicting significant LR-mediated intercellular communications (Fig. 3C and S17, see Methods). For LR pairs involving secreted ligands, we adjusted the *cell-type* score by assigning a lower weight, acknowledging that these interactions can occur over greater physical distances compared to non-secreted ligands (Fig. S17). To further reduce false positives and ensure robust identification of LR pairs, we retained only those pairs that were significantly present between two contiguous developmental stages (Fig. S17).

For LRs with available information, our modified *scSeqComm* pipeline quantified intracellular signaling pathways downstream of receptor activation in target cell types, by integrating data from public regulatory gene databases [47] (see Methods). This approach generated an intracellular signaling score (Fig. 3C and S17) that allowed us to evaluate the reliability of inferred LR communications. It also facilitated the identification of the preferred receptor(s) engaged when a ligand interacted with multiple receptors, enhancing the understanding of specific LR dynamics within cellular contexts.

`

Inferences of LR interactions over cortex development are accessible via the Shiny App *scLRSomatoDev* (to-be-released, Fig. S19), allowing researchers to visualize both number and identities of predicted LR interactions between any of the 729 possible neuronal *cell-type* pairs at each of 7 developmental stages covering E18.5-adulthood. Numbers of LR interactions are depicted by heatmaps (Fig. 3D, S19A), while the identities of LR pairs and associated gene ontologies (GO) pathways are represented by dot plots (Fig. S19B).

Which molecular codes allow specific LR pairs to control the position and connectivity of cell types? To explore this, we conducted a comprehensive analysis of our LR predictions. Across 729 cell-type combinations from E18.5 to adulthood, we quantified the number of predicted LR interactions maintained between at least two consecutive stages in the E18-Adult temporal window. We categorized them into unique, "widely shared" (shared by over 400 cell pairs), and "rarely shared" (used by ≤ 10 cell pairs) (Fig. 3D). The total number of predicted LR interactions for specific cell combinations reached up to ~600, whereas interactions exclusive to unique cell pairs were exceedingly rare. Widely shared LRs, i.e., LRs shared by ≥ 400 cell pairs, were relatively infrequent, accounting for less than 10% of total inferred LRs (Fig. 3D). In contrast, rarely shared LR pairs (by ≤ 10 cell pairs) showed variable frequency, with up to 60 detected in CR-CR interactions. These results suggest that cortical neuron subtypes rely on a diverse repertoire of LR pairs, combining both rare and widely shared interactions in specific patterns to drive their distinct communication and connectivity. Examining the number of inferred LR interactions at different developmental ages further reinforced the conclusion from the global analysis: LRs used exclusively by a single cell pair remained rare regardless of developmental age. This held true even when fixing one of the two cell types in the pair while varying the other (Fig. S20–S22).

By analyzing the number of significant interactions between GlutN and GABAN cell types, we observed that when GlutNs served as the ligand sender, DL GlutNs, as well as UL GlutNs, exhibited stronger interactions with DL GABANs during early developmental stages (Fig. S23). The predominant interactions between UL GlutNs and DL GABANs (compared to UL GABANs) decreased with age (Fig. S23). Consistent with previous findings, UL GlutNs at early developmental stages relied predominantly on intrinsic genetic programs rather than extrinsic inputs (Fig. S7), which aligns with the low number of interactions observed between UL GlutNs and UL GABANs at young ages, followed by an increase in these interactions with maturation. In the reverse direction, when GABANs acted as the ligand source and GlutNs as the target, both UL and DL GABANs interacted significantly more with GlutNs within the same layer from E18.5-P0 to P8 (Fig. S23). By P16 and P30, however, a notable exception emerged, with DL GABANs displaying increased interactions with UL GlutNs (Fig. S23).

`

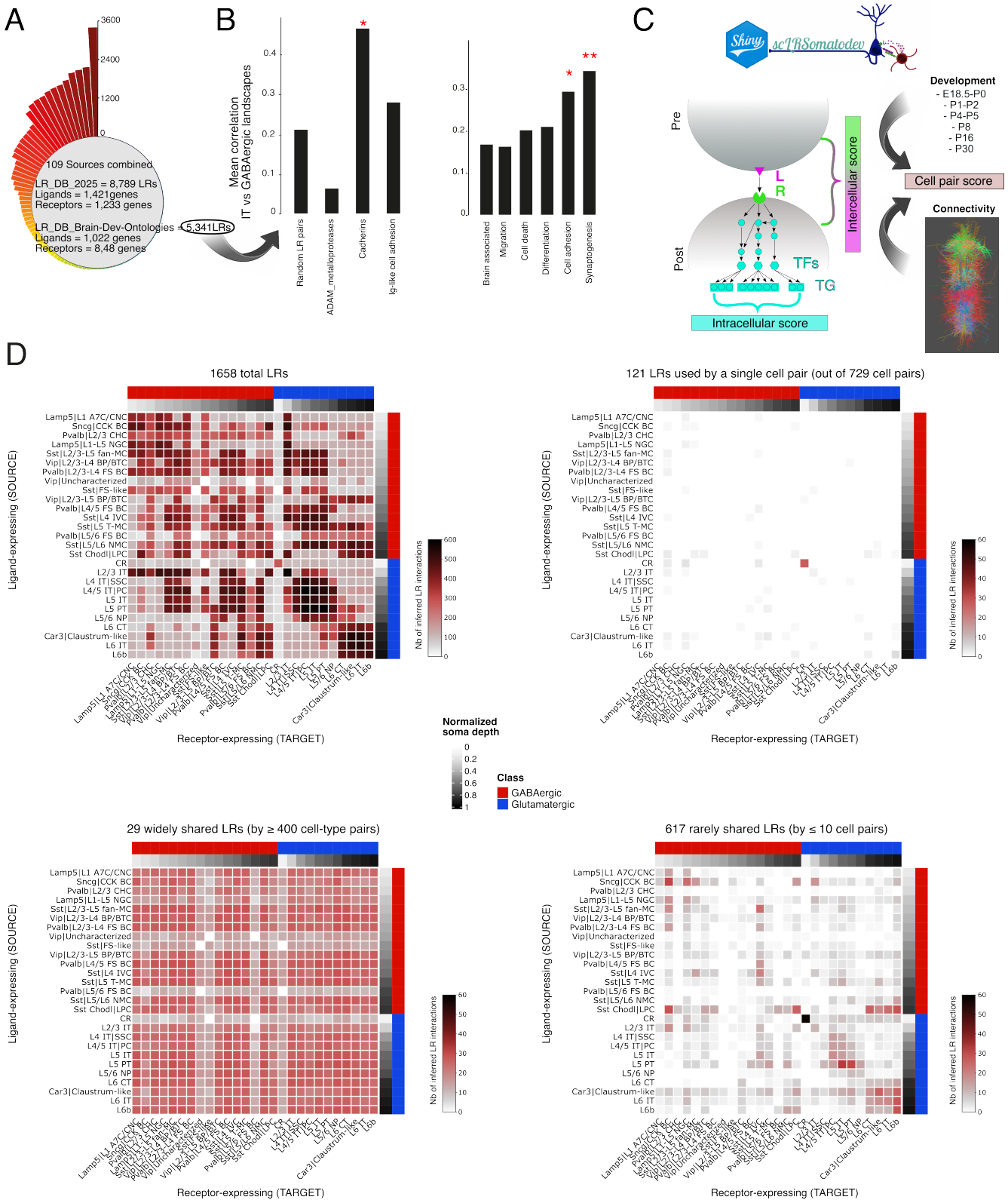

A
3600
2400
1200
0
109 Sources combined
LR_DB_2025 = 8,789 LRs
Ligands = 1,421genes
Receptors = 1,233 genes
LR_DB_Brain-Dev-Ontologies = 5,341LRs
Ligands = 1,022 genes
Receptors = 8,48 genes
B
Mean correlation
IT vs GABAergic landscapes
Random LR pairs
ADAM_metalloproteases
Cadherins
Ig-like cell adhesion
Brain associated
Migration
Cell death
Differentiation
Cell adhesion
Synaptogenesis
C
scRNSomatodev
Development
- E18.5-P0
- P1-P2
- P4-P5
- P8
- P16
- P30
Pre
Post
L
R
TFs
TG
Intercellular score
Intracellular score
Cell pair score
Connectivity
D
1658 total LRs
121 LRs used by a single cell pair (out of 729 cell pairs)
29 widely shared LRs (by ≥ 400 cell-type pairs)
617 rarely shared LRs (by ≤ 10 cell pairs)
Ligand-expressing (SOURCE)
Receptor-expressing (TARGET)
Nb of inferred LR interactions
Normalized soma depth
Class
GABAergic
Glutamatergic

Figure 3: *scLRSomatoDev*, a LR interactome prediction atlas between GlutNs and GABANs over cortex development. (A) Curated database of 8,789 LR pairs. DB, database; L, ligand; R, receptor. LR_DB_2025 contains all LRs, LR_DB_2025_Brain contains brain-expressed LRs with significant expression landscapes. See Tables S1 and S2. (B) High correlation of neurodevelopmental process-associated LR transcriptional landscapes suggest a LR molecular code for neuronal adhesion and synaptogenesis, as well as for the Cadherin family of adhesion molecules. (C) Simplified computational method for the inference of LR-mediated interactions. TF, transcription factor; TG, target gene; S_inter, score for intercellular communication; S_intra, score for intracellular signaling. See Fig. S17-18. (D) Heatmaps illustrating the number of inferred LR interactions that persist across at least two consecutive ages for all 729 *cell-type* pairs. From left to right and top to bottom: all possible interactions; interactions unique to a single *cell-type* pair, widely shared interactions (by >400 pairs) and very rare interactions (shared by <10 pairs). See also Figs S12-23.

Key roles for LR interactions during synaptogenesis and in neurodevelopmental diseases

We analyzed temporal LR predictions between GlutNs and GABANs. The total number of LRs predicted increased over time (Fig. 4A), which aligns with the gradual establishment of synaptic connectivity. Among GABAN subtypes, Sst and Pvalb neurons were the first to receive a high number of LR interactions from GlutNs (Fig. 4A), while Lamp5, Vip, and other GABAN populations exhibited a later increase. Notably, Pvalb cells initially had fewer LRs than Sst neurons but gradually caught up, consistent with studies showing that Sst neurons precede Pvalb neurons in the formation of cortical networks [48]. Furthermore, the fact that Sst and Pvalb cells received the most LRs (Fig. 4A) suggests that these subtypes may rely more heavily on non-cell-autonomous signals for their integration into cortical networks than other GABAN subtypes.

We then examined the temporal progression of LR predictions for each *cell-type*, using IT and GABAN subtypes as examples of source and target cells, respectively. Overall, the number of predicted LRs increased over time (Fig. 4B), aligning with the trends observed in the global analysis (Fig. 4A). However, different patterns emerged for UL and DL interactions, i.e., interactions involving UL vs DL GABAN targets (Fig. 4B). In UL interactions, LRs were minimal at early stages but increased progressively, reaching high levels in adulthood. Conversely, DL interactions exhibited relatively high LR numbers early on, with only modest increases as the organism aged. This suggests that the earlier developmental onset of DL neurons enables them to establish LR connections at a younger age compared to UL neurons.

We then calculated the proportion of LRs associated with five key neurodevelopmental processes—migration, differentiation, cell death, cell adhesion, and synaptogenesis—that are actively utilized during corticogenesis.

LRs associated with migration, differentiation, and cell death showed minimal inferred utilization, suggesting that these neurodevelopmental processes rely less on LR-mediated cell-cell communication and more on cell-intrinsic mechanisms. In contrast, synaptogenesis-associated

`

LRs, and to a lesser extent cell adhesion-associated LRs, were predicted to be strongly utilized from E18.5 to adulthood (Fig. 4C). Notably, Cadherins exhibited strong utilization from P4 onwards, highlighting their potential role in regulating cell-cell interactions during adhesion and synaptogenesis in the cortex.

We next investigated the timing and involvement of LRs associated with neurodevelopmental diseases during development. Using documented gene-disease associations, we categorized the LRs from the LR_DB_2025 into four major neurodevelopmental diseases: autism, epilepsy, schizophrenia (SCZ), and intellectual disability (ID) (Table S5). Notably, the percentage of utilized LRs linked to SCZ and ID remained consistently low across developmental stages (Fig. 4D). In contrast, LRs associated with autism and epilepsy were utilized at high rates, occasionally reaching up to 50%, and their prevalence increased progressively over time (Fig. 4D). This suggests that the etiology of autism and epilepsy may be more closely related to LR-mediated cell-cell communications in cortical areas compared to SCZ and ID. Additionally, some LRs were associated with multiple diseases (Fig. 4E), and notably, the only two significantly utilized LRs common to both autism and SCZ were the Cadherins, *Cdh13* and *Cdh22* (Fig. 4F). Consistent with expectations, gene ontology analysis revealed that predicted disease-associated LRs were enriched in neurodevelopmental biological processes (Fig. 4G).

`

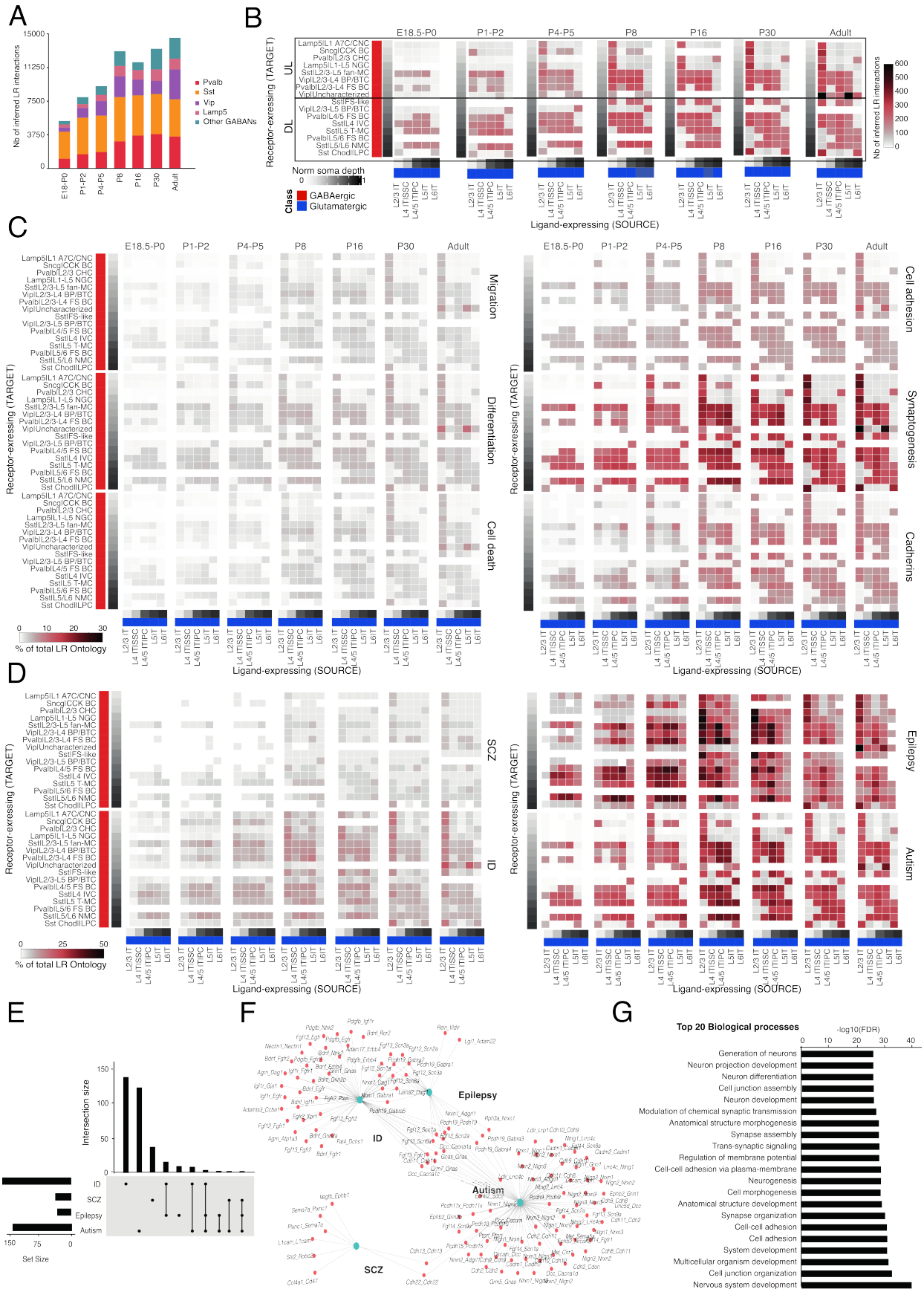

A
B
C
D
E
F
G
Nb of inferred LR interactions
Pvalb
Sst
Vip
Lamp5
Other GABANs
E18.5-P0
P1-P2
P4-P5
P8
P16
P30
Adult
Receptor-exressing (TARGET)
UL
DL
Norm soma depth
Class
GABAergic
Glutamatergic
Ligand-expressing (SOURCE)
Migration
Differentiation
Cell death
Cell adhesion
Synaptogenesis
Cadherins
% of total LR Ontology
SCZ
ID
Epilepsy
Autism
Intersection size
Set Size
Top 20 Biological processes
-log10(FDR)
Generation of neurons
Neuron projection development
Neuron differentiation
Cell junction assembly
Neuron development
Modulation of chemical synaptic transmission
Anatomical structure morphogenesis
Synapse assembly
Trans-synaptic signaling
Regulation of membrane potential
Cell-cell adhesion via plasma-membrane
Neurogenesis
Cell morphogenesis
Anatomical structure development
Synapse organization
Cell-cell adhesion
Cell adhesion
System development
Multicellular organism development
Cell junction organization
Nervous system development

Figure 4: Temporal analysis of LR pairs associated with neurodevelopmental processes and diseases. (A) Number of LR pairs predicted to be utilized per GABAN family as target cells, from E18 to adulthood. (B) Number of predicted LR interactions between IT and GABAN *cell-type*s as source and target cells, respectively, from E18 to adulthood. (C) Percentage of LR pairs of 5 main neurodevelopmental ontologies and of the cadherin LR family that are predicted as utilized between IT and GABAN *cell-type*s as source and target cells, respectively, from E18 to adulthood. (D) Percentage of LR pairs with both L and R associated with neurodevelopmental diseases that are predicted as utilized between IT and GABAN *cell-type*s as source and target cells, respectively, from E18 to adulthood. (E) Upset plot showing intersections between LR pairs of the 4 diseases studied. (F) Interaction network plot showing the LR pairs associated with each disease. (G) Gene ontologies associated with inferred, disease-associated LR pairs.

Validation of the LR prediction atlas: Nrg3-Erbb4

We queried our LR atlas to validate known interactions between GlutNs and GABANs, focusing on the Nrg3-Erbb4 pair, which is recognized for mediating the development of excitatory synapses onto GABANs and is linked to neurodevelopmental disorders [49] (Fig. 5A). First, consistent with the literature, transcriptional landscapes confirmed Nrg3 expression in GlutNs and the restriction of Erbb4 expression to GABANs (Fig. 5B). Nrg3 and Erbb4 expression levels peaked in the second half of the developmental timeline, aligning with synaptogenesis (Fig. 5B). Second, our atlas predicts significant Nrg3-Erbb4 interactions from E18.5-P0 to P30 between GlutNs (Nrg3 source) and most GABANs (Erbb4-expressing target cells) (Fig. 5C). Regardless of mouse age, the intracellular signaling score for Nrg3-Erbb4 was strongest in pairs for which target cells were Pvalb BCs and Vip|L2/3−L5 BP/BTCs, two GABAN types best known to require Erbb4 signaling for the formation of excitatory synapses (Fig. 5D) [12,13,50]. Overall, our atlas supports the role of the Nrg3-Erbb4 interaction in excitatory synapse formation onto specific GABAN types during corticogenesis.

NEO1 is the main CBLN4 receptor at Martinotti-glutamatergic developing synapses

We then investigated whether our gene expression and LR atlases could shed new light on known but incompletely understood cell-cell communications. We focused on CBLN4, a ligand that facilitates synapse formation between the axonal termini of Sst neurons and the apical dendrites of GlutNs in the mouse cortex [11,51] (Fig. 5E).

Despite this knowledge, which precise Sst *cell-type(s)* express(es) Cbln4 and what its receptor(s) is/are in GlutN apical dendrites are two pending questions [51,52] (Fig. 5F) that we addressed using our gene expression and LR atlases. Our analysis of transcriptional profiles revealed that among the 4 identified Sst *cell-types* (Sst|L2/3-L5 fan-MC, Sst|L4 IVC, Sst|L5 T−MC, and Sst|L5/L6

`

NMCs), only the Sst|L2/3-L5 fan-MCs exhibited strong expression of *Cbln4* (Fig. 5G). Then, our LR interactome atlas predicted that CBLN4 was involved in only very few interactions during peak synaptogenesis (P5 to P30), *i.e.*, between Sst|L2/3-L5 fan-MCs as "source" cells and some GlutN types (L2/3 IT, L4 IT|SSC, L4/5 IT|PC, L5 IT and L5 PT) as "target" cells (Fig. 5G) - more unexpectedly, most GABAN *cell-type*s were also predicted as "target" cells. Finally, intercellular signaling scores revealed that among the 3 known CBLN4 receptors (DCC, GLUD1 and NEO1), only GLUD1 and NEO1 were predicted to be involved (Fig. 5H). These intercellular scores exhibited an increase from P4 to P8, whichaligns with earlier findings underscoring the significance of CBLN4 during this developmental period [11] and supports its role in synaptic formation.

We then took advantage of the intracellular score to determine whether GLUD1 or NEO1 might predominate as a CBLN4 receptor in the cortex. Interestingly, while earlier studies showed that GLUD1 is a CBLN4 receptor in cortical GlutNs around P21-P30 [11,51], our intracellular score predicts that NEO1 downstream signaling is more active than GLUD1 downstream signaling in CBLN4 "receiving" neurons at peak synaptogenesis around P4-P8 (Fig. 5H). NEO1-associated genes in "target" GlutNs were notably linked with axon guidance and TGF-beta signaling pathways (Fig. 5H). Based on these predictions, we investigated the presence of a direct interaction between CBLN4 and NEO1 or GLUD1 at P8 in situ. Using proximity ligation assay (PLA), we detected strong interactions between CBLN4 and NEO1 in cortical layers L1-2, while interactions between CBLN4 and GLUD1 were minimal or undetectable, as indicated by the sparse or absent PLA signals (Fig. 5I). Overall, these findings underscore the utility of our gene expression and LR atlases in pinpointing the precise cellular sources of ligands and their primary receptors in target cells.

`

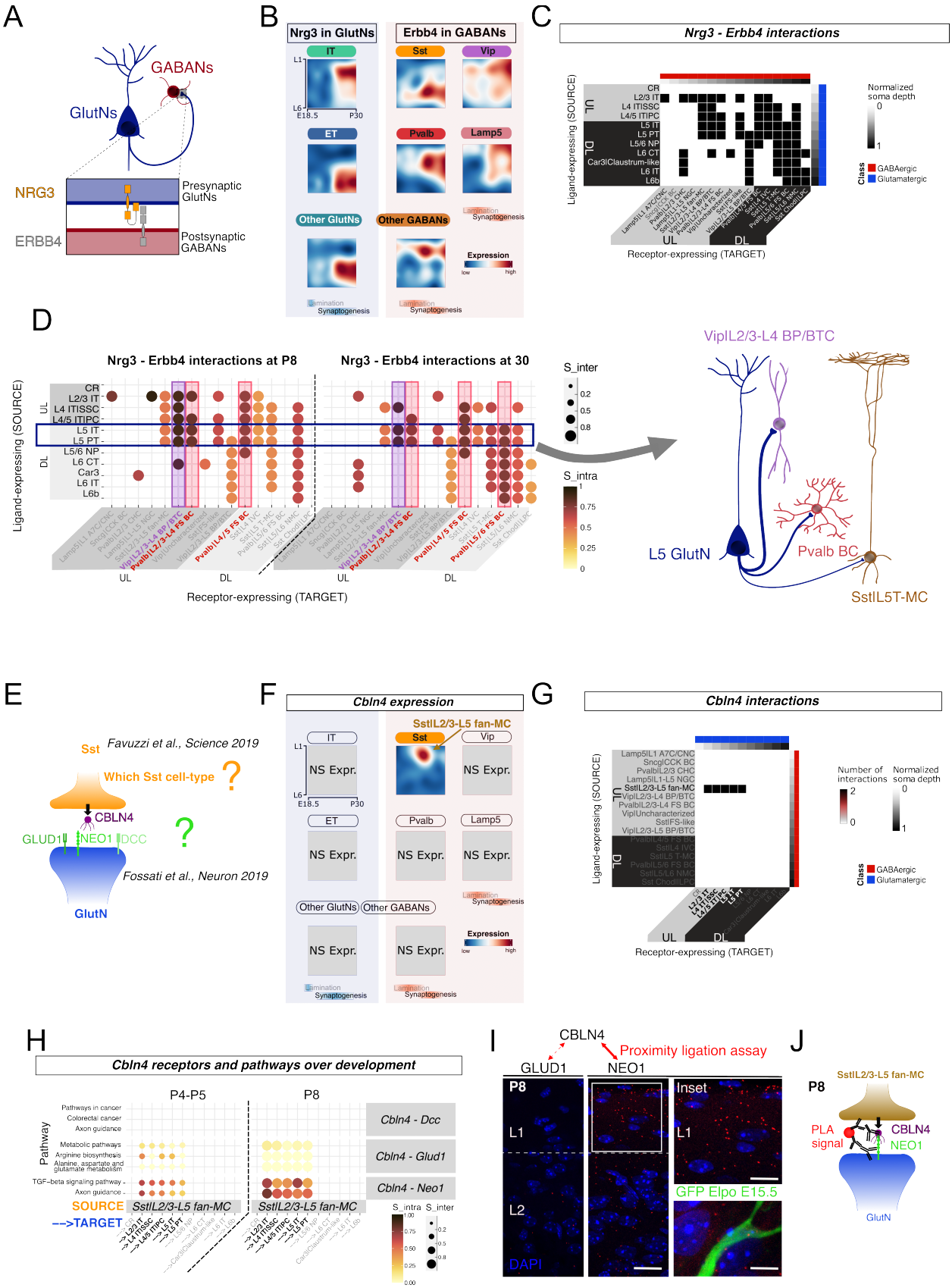

A
GABANs
GlutNs
NRG3
ERBB4
Presynaptic GlutNs
Postsynaptic GABANs
B
Nrg3 in GlutNs
Erbb4 in GABANs
IT
Sst
Vip
ET
Pvalb
Lamp5
Other GlutNs
Other GABANs
Lamination
Synaptogenesis
Expression
low
high
C
Nrg3 - Erbb4 interactions
Ligand-expressing (SOURCE)
Receptor-expressing (TARGET)
UL
DL
Normalized soma depth
Class
GABAergic
Glutamatergic
D
Nrg3 - Erbb4 interactions at P8
Nrg3 - Erbb4 interactions at 30
S_inter
S_intra
VipIL2/3-L4 BP/BTC
L5 GlutN
Pvalb BC
SstIL5T-MC
E
Sst
Favuzzi et al., Science 2019
Which Sst cell-type
?
CBLN4
GLUD1
NEO1
DCC
Fossati et al., Neuron 2019
GlutN
F
Cbln4 expression
SstIL2/3-L5 fan-MC
NS Expr.
G
Cbln4 interactions
Number of interactions
H
Cbln4 receptors and pathways over development
P4-P5
P8
Pathway
Pathways in cancer
Colorectal cancer
Axon guidance
Metabolic pathways
Arginine biosynthesis
Alanine, aspartate and glutamate metabolism
TGF-beta signaling pathway
Cbln4 - Dcc
Cbln4 - Glud1
Cbln4 - Neo1
SOURCE
-->TARGET
I
CBLN4
Proximity ligation assay
GLUD1
NEO1
P8
Inset
L1
L2
GFP Elpo E15.5
DAPI
J
SstIL2/3-L5 fan-MC
PLA signal
CBLN4
NEO1
GlutN

Figure 5. Validation and Application of the Ligand-Receptor Atlas for Knowledge Completion. (A) Diagram illustrating NRG3-ERBB4 interaction between glutamatergic and GABAergic cells. (B) *Nrg3* and *Erbb4* expressions in transcriptional landscapes of glutamateric and GABAergic neuron *families*, respectively. (C) Heatmap depicting predictions of *Nrg3-Erbb4* mediated cell-cell interactions between GlutN and GABAN *cell-type*s. Blue represents GlutN types, red represents GABAN types. The white to black gradient indicates the median normalized soma depth, with 0 at the pial surface and 1 at the white matter. (D) Left: Dot plot of predicted *Nrg3-Erbb4* mediated cell-cell interactions at P8 and P30. GlutNs (blue) are sources of *Nrg3* and GABANs are targets expressing *Erbb4*. Right: Schematic representation illustrating the predicted interaction strength among three example cell pairs, highlighting preferred connectivity patterns based on computational predictions. (E) Diagram illustrating Cbln4-mediated inhibitory synapse formation from Sst cells to GlutN in the cortex. (F) Transcriptional landscapes showing *Cbln4* expression across different neuronal families. (G) Heatmap of LR pairs involving *Cbln4*. (H) Dot plot of predicted Clbn4 mediated cell-cell interactions at P4-P5 and P8. Pathways are shown on the left. CBLN4-receptor pairs are represented on the right. (I) Proximity ligation assay (PLA) results for interactions between CBLN4 and GLUD1 and CBLN4 and NEO1 in cortical layers L1-2. A magnification of a green process from a GFP-electroporated neuron (E15.5) with red puncta on it is shown on the lower right. (J) Summary of the experimental findings and predictions resulting from our analysis. S_inter, intercellular score; S_intra, intracellular score.

Uncovering novel interactions: CDH13 and PCDH8 as mediators of perisomatic inhibition in deep and superficial layers, respectively

We asked whether our LR atlas can identify novel LR pairs regulating the establishment of specific connectivity patterns. We focused on the cadherin *family* of adhesion molecules that are thought to play such a role in the cortex [45,46,53]. Analysis of transcriptional landscapes revealed notable diversity in the spatio-temporal expression profiles of cadherin family members across different cell types (Fig. 2G).

The atypical *Cdh13*, associated with neurodevelopmental disorders [54,55], was particularly interesting as it showed highest expression in DL GlutNs and MGE-derived GABANs (Pvalb & Sst) (Fig. 6A). *Cdh13* expression was notably high in DL GlutNs neurons through cortical development, and in Pvalb and Sst neurons around the P4-P30 synaptogenesis period [11] (Fig. 5A). This suggests a role for *Cdh13-Cdh13* mediated intercellular communication in the establishment of connectivity between MGE-derived GABANs and DL GlutNs. Accordingly, our LR atlas indeed inferred that *Cdh13-Cdh13* might mediate interactions that are mainly between Sst / Pvalb GABANs and DL GlutNs. It was a strict rule at specific time points like P4-P5 and P30 (Fig. 6C), and there were only exceptions when considering all time points collectively (33/41 involved Pvalb/Sst->DL GlutN pairs, Fig. 6B). These findings identify Cdh13 as a potential candidate

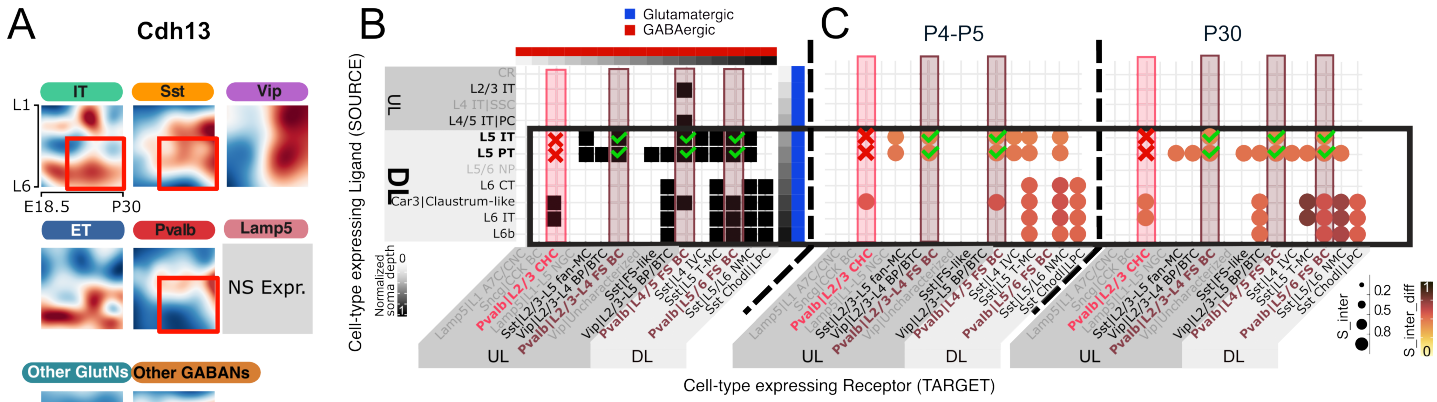
`

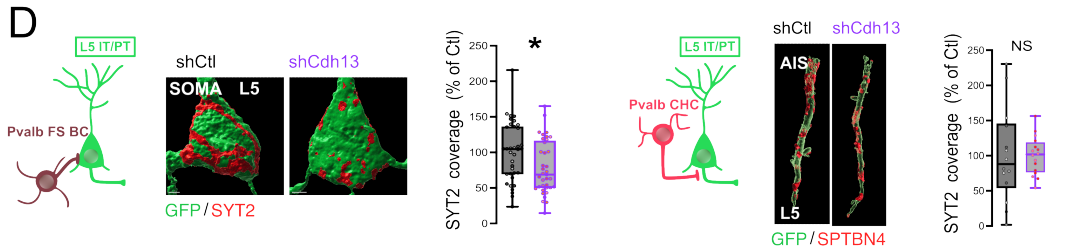
molecule for establishing connectivity between Pvalb/Sst GABANs and DL GlutNs, consistent with recent evidence suggesting that Cdh13 mediates the perisomatic inhibition of Pvalb BCs onto layer 5 extratelencephalic (L5 ET) neurons [56]. We used the fact that synaptic connectivity between Pvalb and GlutNs is measurable with SYT2 staining [56,57] to investigate the implication of CDH13 in the development of perisomatic inhibition mediated by Pvalb BCs on DL GlutN somata. After *Cdh13* knock-down in L5 GlutNs by in utero electroporation, we assessed SYT2 perisomatic inhibition at P28, after migration and establishment of connectivity have occurred. As expected, the surface area of DL GlutN soma covered by $SYT2^+$ boutons was significantly reduced by *Cdh13* knock-down (Fig. 6C). This result shows that post-synaptic CDH13 is necessary for proper Pvalb BC-perisomatic inhibition of DL GlutN somata. The specificity of this regulation was further supported by the observation that Pvalb chandelier cell (CHC) innervation of the axon initial segment was unaffected, as predicted by our LR atlas (Fig. 6D).

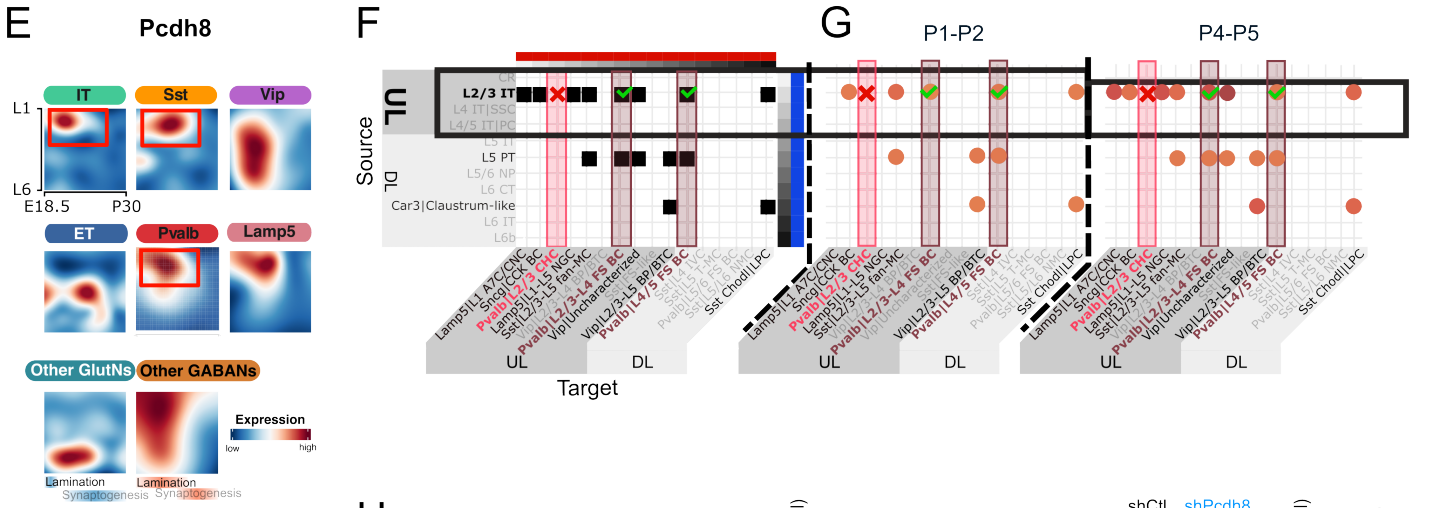
We then turned our focus to PCDH8 because its laminar expression is relatively complementary to the one of DL-enriched CDH13, i.e., *Pcdh8* is enriched in both GlutNs and Pvalb/Sst subtypes of UL (Fig. 6E). Out of the only 8.5% of all possible GlutN-GABAN interactions predicted to involve *Pcdh8* (15/176), L2/3 ITs displayed the highest fraction, i.e., 50% (8/16 possible interactions). The fraction of their possible interaction with Pvalb BC types reached 66.7% (2/3 possible interactions) (Fig. 6F). Also, inferred *Pcdh8*-mediated interactions were more prominent in cell pairs involving L2/3 IT neurons compared to those involving L5 PT or Car3|Claustrum-like neurons (Fig. 6G).

Therefore, we used in utero electroporation of a *Pcdh8* shRNA at E15.5 to investigate the role of PCDH8-PCDH8 interactions in Pvalb-mediated perisomatic inhibition of L2/3 IT neurons. PCDH8 knock-down significantly increased SYT2 coverage of GFP-electroporated cells compared to a control situation, which indicates that PCDH18 regulates perisomatic inhibition of L2/3 IT neurons negatively (Fig. 6H). Thus, our LR atlas allowed us to conclude that different cadherins modulate perisomatic inhibition in different ways depending on the layer examined.

`


A
Cdh13
IT
Sst
Vip
L1
L6
E18.5
P30
ET
Pvalb
Lamp5
NS Expr.
Other GlutNs
Other GABANs
Expression
low
high
Lamination
Synaptogenesis
B
Glutamatergic
GABAergic
Cell-type expressing Ligand (SOURCE)
UL
DL
Normalized soma depth
C
P4-P5
P30
S_inter
S_inter_diff
Cell-type expressing Receptor (TARGET)


D
L5 IT/PT
Pvalb FS BC
shCtl
shCdh13
SOMA
L5
GFP / SYT2
SYT2 coverage (% of Ctl)
*
Pvalb CHC
AIS
GFP / SPTBN4
NS


E
Pcdh8
IT
Sst
Vip
L1
L6
E18.5
P30
ET
Pvalb
Lamp5
Other GlutNs
Other GABANs
Expression
low
high
Lamination
Synaptogenesis
F
Source
UL
DL
Target
G
P1-P2
P4-P5

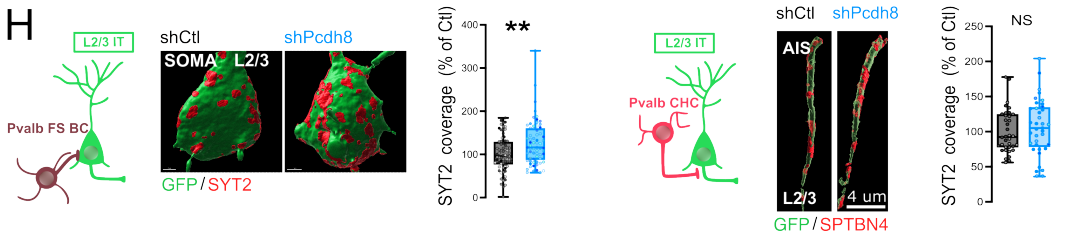

H
L2/3 IT
Pvalb FS BC
shCtl
shPcdh8
SOMA
L2/3
GFP / SYT2
SYT2 coverage (% of Ctl)
**
Pvalb CHC
AIS
4 um
GFP / SPTBN4
NS

Figure 6. Cadherins Cdh13 and Pcdh8 regulate perisomatic inhibition in deep and superficial cortical layers. (A) Transcriptional landscapes of Cdh13 expression across different cell families. (B) Heatmap of predicted Cdh13-Cdh13 interactions between GlutN and GABAN cell types shared by at least two consecutive ages. (C) Utilization of Cdh13-Cdh13 interactions across GlutNs and GABANs at P4-P5 and P30, with interaction strength quantified by the inter_diff score. (D) Perisomatic (left) and axon initial segment (AIS, right) inhibitory innervation of E13.5-electroporated L5 GlutNs by SYT2-expressing Pvalb BC and CHC boutons, respectively, in control or Cdh13-knockdown (shCdh13) GlutNs (Mann-Whitney test, *p-value < 0.05, n=3-4 mice/condition; ≥ 36 and 14 cells/condition for soma and AIS, respectively). (E) Transcriptional landscapes of Pcdh8 expression across different cell families. (F) Heatmap of predicted Pcdh8-Pcdh8 interactions between GlutN and GABAN cell types shared by at least two consecutive ages. (G) Predicted Pcdh8-Pcdh8 interactions at P1-P2 and P4-P5, with stronger expression in L2/3 IT-involving cell pairs. (H) Perisomatic (left) and axon initial segment (AIS, right) inhibitory innervation of E15.5-electroporated L2/3 GlutNs by SYT2-expressing Pvalb BC and CHC boutons, respectively, in control or Pcdh8-knockdown (shPcdh8) GlutNs (Mann-Whitney test, **p-value < 0.01, n=3-4 mice/condition; > 61 and 35 cells/condition).

Discussion

Seminal studies in the last decades have indicated that the stereotyped laminar positioning and connectivity of cortical GABAN subtypes are both in large part extrinsically regulated [15–17]. These data suggest that a molecular code based on LR interactions between firstly settled GlutNs and secondarily arriving GABANs might orchestrate cortical circuit assembly. In this work, we utilized high-throughput scRNA-seq to address this question, by attempting to identify LR pairs that are essential for GlutN-GABAN interaction and cortical circuit assembly. All results from the LR interaction inference, along with gene expression visualizations, are available in a Shiny App *scLRSomatoDev* (to-be-released).

We first characterized the transcriptome of all the main known neuron types over the course of cortical development. We provide this dataset, which results from an extensive integration and annotation effort, as a new reference to capture the dynamic transcriptional landscape of corticogenesis (Shiny App). We then focused, for LR interaction assessment, on the E18-P30 critical period during which GABANs reach their preferred laminar position and establish synaptic connections with GlutNs and other GABAN partners. We developed computational tools to: (1) assess spatiotemporal expression of all genes in all cortical neuron types and (2) infer the number and identities of significant LR interactions between all neuron types that might control stereotypical cortical assembly. We provide evidence that these tools can be utilized to: (1) validate

`

true ground observations (e.g., Nrg3 - Erbb4), (2) complete knowledge about known LR interactions (e.g., Cbln4 – Neo1 for Sst MC -> GlutN interaction) and (3) identify novel LR pairs with a role in the formation of specific connections (*e.g.*, the *Cdh13* and *Pcdh8* cases).

Ligand-receptor mediated cell-cell interactions constitute a fundamental process that participates in shaping most biological tissues. Before the development of scRNA-seq, studying cell-cell interactions was low-throughput, limited to a short selection of genes/proteins and one to few *cell-type* pairs. Since the advent of scRNA-seq, a growing number of computational tools has been developed to infer cell-cell communication. Most tools are based solely on the gene expression level of the LR pair between 2 *cell-types* [58–60], but few tools provide a more accurate reflection of ground truth by considering the response triggered by LR binding in the cell expressing the receptor [47,61–63]. Here, we benchmarked most of these methods and ended up choosing one of the most recent ones that has the advantage of inferring both intercellular and intracellular signaling for a given LR pair between *cell-type* pairs: scSeqComm [47]. The fact that scSeqComm returns these scores between 0 and 1 makes it easier to increment a threshold to identify relevant LR pairs. For the intercellular score, we determined a threshold using prior knowledge taking into account the final connectivity between *cell-types* (connectivity score) as well as the age (development score) at which the analysis was conducted. This prior knowledge allowed us to increase stringency and robustness of our inferences. Furthermore, by computing an intracellular score, scSeqComm quantifies the evidence of ongoing intracellular signaling triggered by the LR binding in the target cell. This intracellular score allowed us to increase the reliability of inferred communication. Custom thresholds can be set for both intercellular and intracellular scores according to the desired reliability of the inferred interactions. To further increase the reliability of the inferred interactions, we created a meticulously curated LR database by amalgamating existing published databases. Additionally, we incorporated LR pairs from the literature to enrich its content.

Our analysis of LR predictions reveals that neuronal connections are often determined by specific combinations of LRs with distinct specificities. Notably, a core group of approximately 40 broadly expressed LRs mediate interactions in over 50% of cell-type pairs, complemented by a larger subset of LRs with varying degrees of specificity. Importantly, LRs unique to individual neuronal connections are exceedingly rare, suggesting that neuronal connectivity is predominantly shaped by shared and context-dependent LRs rather than by unique or isolated interactions.

Recent studies have shown that members of the cadherin family of adhesion molecules are differentially expressed in subtypes of cortical excitatory and inhibitory neurons [34,45]. This suggested that a cadherin code may orchestrate the stereotypical organization of cortical circuits. Our findings, which demonstrate that cadherin superfamily members exhibit *cell-type*-pair specific spatiotemporal expression profiles (Fig. S13-S16), and that CDH13 and PCDH8 regulate critical aspects of deep versus superficial circuit formation (Fig. 5), support this hypothesis. Future functional studies based on our LR atlas will be crucial in confirming or refuting this hypothesis.

Based on hypotheses generated by our LR atlas, we experimentally identified two cadherin superfamily members, CDH13 and PCDH8, as regulators of Pvalb BC mediated perisomatic

`

inhibition in DL and SL, respectively. Our observation that *Cdh13-Cdh13* interactions are necessary for perisomatic inhibition by Pvalb, but not Sncg, BCs in DL, aligns with a recent study [56].

In contrast to the promotive role of CDH13 in DLs, we find that PCDH8 acts as a suppressor of perisomatic inhibition in ULs. These findings highlight the predictive power of our LR atlas in identifying both facilitatory and inhibitory interactions, a crucial feature given the role of repulsive interactions in forebrain neurodevelopment. Furthermore, the data suggest that different cadherins or CAM family members may modulate synaptogenesis in region-specific and divergent ways across the cortex.

The suppressive effect of PCDH8 on inhibitory synapse assembly in ULs parallels a similar role observed for another protocadherin, PCDH18, in regulating Sst neuron-to-GlutN synapse assembly [11]. Collectively, given the established associations of CDH13 and PCDH8 with neurodevelopmental disorders [54,55,64], we propose that disrupted Pvalb basket cell-mediated inhibition in specific cortical layers could play a key role in the pathophysiology of these conditions.

In addition to neurons, non-neuronal cells also play a role in the early stages of cortical circuit assembly. For example, blood vessels and ventral oligodendrocyte precursor cells regulate the tangential migration of GABANs from the subpallium to the cortical plate through specific *Cxcl12*-SEMA6A/B-PLXNA3 unipolar contact repulsion [65]. Future studies should include single-cell transcriptomics-based analyses of neuron-non-neuronal cell communication to further explore these interactions.

Limitations of the approach – Considerations for Future Improvements in Ligand-Receptor Atlases

Although our approach is robust and comprehensive in terms of the genes investigated for LR interactions, it does not provide information on the subcellular expression of ligands and receptors across different cell types, which is likely important for explaining specific connectivity patterns [49]. Furthermore, our approach focuses on mRNA expression as a proxy for protein expression, although the two may differ in reality. Recent advances in *in vitro* single cell high-throughput proteomics suggest it will soon be possible to include the required subcellular protein expression information in future atlases [66].

Currently, our knowledge of the spatial positions of cortical cell types relative to each other is limited to adult stages [67,68], which impedes precise inference of timely LR mediated cell-cell interactions at specific developmental stages, whether transient or stable. The advent of spatial transcriptomic approaches with exceptional sensitivity, throughput, and resolution holds promise to bridge this gap in the near future.

Acknowledgments:

We thank Julien Prados (UNIGE) for providing his Torch model for artificial neural network (ANN)-based cell type identification, and members of the Cardoso laboratory for their input, and

`

the Molecular and Cellular Biology Facility (PBMC), the Animal Core Facility and the Imaging Facility (inMagic) INMED platforms.

Funding: This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Agence Nationale de la Recherche with ANR-13-JSV4-0006 SynD2 and ANR-23-CE16-0021 CALIN (A.d.C.), NeuroMarseille ICR+ Grant 2021 (A.d.C.), Fondation pour la Recherche sur le Cerveau 'Développement et vieillissement' (A.d.C.), Fondation Lejeune (A.d.C.), European Community 7th Framework programs (Development and Epilepsy—Strategies for Innovative Research to improve diagnosis, prevention and treatment in children with difficult to treat Epilepsy [DESIRE], Health-F2-602531-2013 (A. R., C.C.), and by an Excellence Initiative of Aix–Marseille University/A*MIDEX grant (CALIN-R24002AA) of the French 'Investissements d'Avenir' programme (C.C., A.d.C.). Research in the Telley laboratory was supported by ERC starting grant CERDEV_759112 and a SNSF grant 31003A_182676/1.

Author contributions: A.d.C. and R.M. initiated the study. A.d.C. and L.T. conceptualized and supervised the study. A.d.C., L.T. and R.M. designed and conceptualized the experiments. A.d.C. and R.M. performed most experiments. R.M. analyzed most experiments, supervised by L.T. and A.d.C.. T.D.N. performed some LR analyses for figures 3 and 4. R.M. generated the Shiny App and L.S. added some improvements. C.C., A.G., L.C. and V.B. performed and analyzed shRNA design/production and in utero electroporations. E.P. validated shRNAs and performed proximity ligation assays. A.d.C. and R.M. wrote the manuscript with inputs from all authors.

Competing interests: The authors declare no competing interests.

Data and materials availability: All annotated data will be available online once published. All other data are available in the main text or the supplementary materials. License information: Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

`

References:

1. Gouwens, N. W. *et al.* Integrated Morphoelectric and Transcriptomic Classification of Cortical GABAergic Cells. *Cell* **183**, 935-953.e19 (2020).

2. Scala, F. *et al.* Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature* **598**, 144–150 (2021).

3. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).

4. Yao, Z. *et al.* A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).

5. Zhang, M. *et al.* Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature* **624**, 343–354 (2023).

6. del Pino, I. *et al.* Erbb4 Deletion from Fast-Spiking Interneurons Causes Schizophrenia-like Phenotypes. *Neuron* **79**, 1152–1168 (2013).

7. Jiang, X. *et al.* Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, aac9462 (2015).

8. Motta, A. *et al.* Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science* **366**, eaay3134 (2019).

9. Campagnola, L. *et al.* Local connectivity and synaptic dynamics in mouse and human neocortex. *Science* **375**, eabj5861 (2022).

10. Schneider-Mizell, C. M. *et al.* Cell-type-specific inhibitory circuitry from a connectomic census of mouse visual cortex. Preprint at https://doi.org/10.1101/2023.01.23.525290 (2023).

11. Favuzzi, E. *et al.* Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. *Science* **363**, 413–417 (2019).

12. Fazzari, P. *et al.* Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. *Nature* **464**, 1376–1380 (2010).

13. Bernard, C. *et al.* Cortical wiring by synapse type–specific control of local protein synthesis. *Science* **378**, eabm7466 (2022).

14. Pla, R., Borrell, V., Flames, N. & Marín, O. Layer Acquisition by Cortical GABAergic Interneurons Is Independent of Reelin Signaling. *J. Neurosci.* **26**, 6924–6934 (2006).

15. Lodato, S. *et al.* Excitatory Projection Neuron Subtypes Control the Distribution of Local Inhibitory Interneurons in the Cerebral Cortex. *Neuron* **69**, 763–779 (2011).

16. Ye, Z. *et al.* Instructing Perisomatic Inhibition by Direct Lineage Reprogramming of Neocortical Projection Neurons. *Neuron* **88**, 475–483 (2015).

17. Wester, J. C. *et al.* Neocortical Projection Neurons Instruct Inhibitory Interneuron Circuit Development in a Lineage-Dependent Manner. *Neuron* **102**, 960-975.e6 (2019).

18. Miyoshi, G. & Fishell, G. GABAergic Interneuron Lineages Selectively Sort into Specific Cortical Layers during Early Postnatal Development. *Cerebral Cortex* **21**, 845–852 (2011).

19. Baudoin, J.-P. *et al.* Tangentially Migrating Neurons Assemble a Primary

`

Cilium that Promotes Their Reorientation to the Cortical Plate. *Neuron* **76**, 1108–1122 (2012).

20. Southwell, D. G. *et al.* Intrinsically determined cell death of developing cortical interneurons. *Nature* **491**, 109–113 (2012).

21. Wong, F. K. *et al.* Pyramidal cell regulation of interneuron survival sculpts cortical networks. *Nature* **557**, 668–673 (2018).

22. Priya, R. *et al.* Activity Regulates Cell Death within Cortical Interneurons through a Calcineurin-Dependent Mechanism. *Cell Reports* **22**, 1695–1709 (2018).

23. Telley, L. *et al.* Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364**, eaav2522 (2019).

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

25. Mi, D. *et al.* Early emergence of cortical interneuron diversity in the mouse embryo. *Science* **360**, 81–85 (2018).

26. Mayer, C. *et al.* Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).

27. Bandler, R. C. *et al.* Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).

28. Lee, D. R. *et al.* Transcriptional heterogeneity of ventricular zone cells in the ganglionic eminences of the mouse forebrain. *eLife* **11**, e71864 (2022).

29. BRAIN Initiative Cell Census Network (BICCN) *et al.* A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature* **598**, 86–102 (2021).

30. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

31. Mandai, K., Rikitake, Y., Mori, M. & Takai, Y. Nectins and Nectin-Like Molecules in Development and Disease. in *Current Topics in Developmental Biology* vol. 112 197–231 (Elsevier, 2015).

32. Friedman, L. G., Benson, D. L. & Huntley, G. W. Cadherin-Based Transsynaptic Networks in Establishing and Modifying Neural Connectivity. in *Current Topics in Developmental Biology* vol. 112 415–465 (Elsevier, 2015).

33. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).

34. Libé-Philippot, B. *et al.* Auditory cortex interneuron development requires cadherins operating hair-cell mechanoelectrical transduction. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7765–7774 (2017).

35. Li, G. *et al.* Regional Distribution of Cortical Interneurons and Development of Inhibitory Tone Are Regulated by Cxcl12/Cxcr4 Signaling. *J. Neurosci.* **28**, 1085–1098 (2008).

36. López-Bendito, G. *et al.* Chemokine Signaling Controls Intracortical Migration and Final Distribution of GABAergic Interneurons. *J. Neurosci.* **28**, 1613–1624 (2008).

37. Borrell, V. & Marín, O. Meninges control tangential migration of hem-derived Cajal-Retzius cells via CXCL12/CXCR4 signaling. *Nat Neurosci* **9**, 1284–1293 (2006).

`

38. Tiveron, M.-C. *et al.* Molecular Interaction between Projection Neuron Precursors and Invading Interneurons via Stromal-Derived Factor 1 (CXCL12)/CXCR4 Signaling in the Cortical Subventricular Zone/Intermediate Zone. *J. Neurosci.* **26**, 13273–13278 (2006).

39. Causeret, F., Moreau, M. X., Pierani, A. & Blanquie, O. The multiple facets of Cajal-Retzius neurons. *Development* **148**, dev199409 (2021).

40. Stumm, R. K. *et al.* CXCR4 Regulates Interneuron Migration in the Developing Neocortex. *J. Neurosci.* **23**, 5123–5130 (2003).

41. Wang, Y. *et al.* CXCR4 and CXCR7 Have Distinct Functions in Regulating Interneuron Migration. *Neuron* **69**, 61–76 (2011).

42. Tanaka, D. H. *et al.* CXCR4 Is Required for Proper Regional and Laminar Distribution of Cortical Somatostatin-, Calretinin-, and Neuropeptide Y-Expressing GABAergic Interneurons. *Cerebral Cortex* **20**, 2810–2817 (2010).

43. Graf, E. R., Zhang, X., Jin, S.-X., Linhoff, M. W. & Craig, A. M. Neurexins Induce Differentiation of GABA and Glutamate Postsynaptic Specializations via Neuroligins. *Cell* **119**, 1013–1026 (2004).

44. Krueger, D. D., Tuffy, L. P., Papadopoulos, T. & Brose, N. The role of neurexins and neuroligins in the formation, maturation, and function of vertebrate synapses. *Current Opinion in Neurobiology* **22**, 412–422 (2012).

45. Sun, Y.-C. *et al.* Integrating barcoded neuroanatomy with spatial transcriptional profiling enables identification of gene correlates of projections. *Nat Neurosci* **24**, 873–885 (2021).

46. Basu, R., Taylor, M. R. & Williams, M. E. The classic cadherins in synaptic specificity. *Cell Adhesion & Migration* **9**, 193–201 (2015).

47. Baruzzo, G., Cesaro, G. & Di Camillo, B. Identify, quantify and characterize cellular communication from single-cell RNA sequencing data with *scSeqComm*. *Bioinformatics* **38**, 1920–1929 (2022).

48. Tuncdemir, S. N. *et al.* Early Somatostatin Interneuron Connectivity Mediates the Maturation of Deep Layer Cortical Circuits. *Neuron* **89**, 521–535 (2016).

49. Exposito-Alonso, D. *et al.* Subcellular sorting of neuregulins controls the assembly of excitatory-inhibitory cortical circuits. *eLife* **9**, e57000 (2020).

50. Batista-Brito, R. *et al.* Developmental Dysfunction of VIP Interneurons Impairs Cortical Circuits. *Neuron* **95**, 884-895.e9 (2017).

51. Fossati, M. *et al.* Trans-Synaptic Signaling through the Glutamate Receptor Delta-1 Mediates Inhibitory Synapse Formation in Cortical Pyramidal Neurons. *Neuron* **104**, 1081-1094.e7 (2019).

52. Liakath-Ali, K., Polepalli, J. S., Lee, S.-J., Cloutier, J.-F. & Südhof, T. C. Transsynaptic cerebellin 4–neogenin 1 signaling mediates LTP in the mouse dentate gyrus. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2123421119 (2022).

53. Lv, X. *et al.* Patterned cPCDH expression regulates the fine organization of the neocortex. *Nature* **612**, 503–511 (2022).

54. Lesch, K.-P. *et al.* Molecular genetics of adult ADHD: converging evidence from genome-wide association and extended pedigree linkage studies. *J Neural Transm*

`

115, 1573–1585 (2008).

55. Sanders, S. J. *et al.* Multiple Recurrent De Novo CNVs, Including Duplications of the 7q11.23 Williams Syndrome Region, Are Strongly Associated with Autism. *Neuron* **70**, 863–885 (2011).

56. Jézéquel, J. *et al.* Cadherins orchestrate specific patterns of perisomatic inhibition onto distinct pyramidal cell populations. Preprint at https://doi.org/10.1101/2023.09.28.559922 (2023).

57. Sommeijer, J.-P. & Levelt, C. N. Synaptotagmin-2 Is a Reliable Marker for Parvalbumin Positive Inhibitory Boutons in the Mouse Visual Cortex. *PLoS ONE* **7**, e35323 (2012).

58. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat Protoc* **15**, 1484–1506 (2020).

59. Cabello-Aguilar, S., Fau, C., Lacroix, M. & Colinge, J. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics.

60. Hou, R., Denisenko, E., Ong, H. T., Ramilowski, J. A. & Forrest, A. R. R. Predicting cell-to-cell communication networks using NATMI. *Nat Commun* **11**, 5011 (2020).

61. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* **17**, 159–162 (2020).

62. Cheng, J., Zhang, J., Wu, Z. & Sun, X. Inferring microenvironmental regulation of gene expression from single-cell RNA sequencing data using scMLnet with an application to COVID-19. *Briefings in Bioinformatics* **22**, 988–1005 (2021).

63. Hu, Y., Peng, T., Gao, L. & Tan, K. CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data. *Sci. Adv.* **7**, eabf1356 (2021).

64. Butler, M., Rafi, S., Hossain, W., Stephan, D. & Manzardo, A. Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. *IJMS* **16**, 1312–1335 (2015).

65. Lepiemme, F. *et al.* Oligodendrocyte precursors guide interneuron migration by unidirectional contact repulsion. *Science* **376**, eabn6204 (2022).

66. Unterauer, E. M. *et al.* Spatial proteomics in neurons at single-protein resolution. *Cell* **187**, 1785-1800.e16 (2024).

67. Zhang, M. *et al. A Molecularly Defined and Spatially Resolved Cell Atlas of the Whole Mouse Brain*. http://biorxiv.org/lookup/doi/10.1101/2023.03.06.531348 (2023) doi:10.1101/2023.03.06.531348.

68. Yao, Z. *et al. A High-Resolution Transcriptomic and Spatial Atlas of Cell Types in the Whole Mouse Brain*. http://biorxiv.org/lookup/doi/10.1101/2023.03.06.531121 (2023) doi:10.1101/2023.03.06.531121.

`