Title: Inferring Ligand-Receptor Interactions between GABAergic and Glutamatergic neurons during somatosensory cortex 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 Univ, 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 and inhibitory neuron types, each characterized by distinct positional and synaptic connectivity patterns, yet the molecular mechanisms driving this stereotyped organization remain largely unknown. To identify ligand-receptor (LR) pairs that regulate interactions and connectivity between cortical neurons during embryonic and postnatal development, we tracked the transcriptional dynamics of all genes across all cortical neuron types at 17 time points spanning the entire course of cortical development. Using this data, we built a comprehensive bioinformatic atlas to infer significant LR-mediated interactions among all neuron types throughout cortical maturation. This atlas not only validated observations but also served as a powerful tool for hypothesis generation, revealing the critical role of two cadherin superfamily members in mediating perisomatic inhibition of deep and superficial layer excitatory neurons by neighboring parvalbumin-expressing basket cells.

Main Text:

The mammalian cerebral cortex is a six-layered tissue located at the brain's surface, responsible for 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 classes of neurons: excitatory glutamatergic neurons (GlutNs) and inhibitory GABAergic neurons (GABANs). Both GlutNs and GABANs are further divided into various subtypes, a diversity revealed through single-cell RNA sequencing (scRNA-seq). Based on biological significance, researchers have identified eleven main GlutN and sixteen main GABAN subtypes, totaling twenty-seven neuron subtypes [1–4]. These subtypes (1) exhibit specific morphological, electrophysiological, and transcriptomic properties, (2) preferentially localize to certain cortical layers, (3) form and receive synaptic connections from specific neuronal partners, often targeting particular subcellular compartments and (4) play defined roles in cortical computation, which are gradually being understood. The resulting cortical arrangement and associated synaptic connectivity, involving up to 729 potential neuron pairings, are crucial 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.

Therefore, characterizing how the final positions and synaptic specificities of the 729 potential neuronal *cell-type* pairs are established and regulated throughout development is crucial for both developmental and clinical neuroscience. Over the past decade, a tremendous effort has been made in characterizing the identity, laminar position and synaptic characteristics of cortical neuronal subtypes [3–5]. However, the molecular mechanisms that regulate these processes throughout development remain largely unknown, with only few exceptions [6–8].

It is likely that cell-cell communications via LR interactions play a key role, as these interactions are critical in the development of many tissues and organs. Several studies support this hypothesis, suggesting that excitatory glutamatergic neurons (GlutNs) may non-cell autonomously influence the recruitment of specific GABAergic neuron (GABAN) subtypes. For example, GlutN subtypes settle into their final positions earlier than the synchronically generated GABANs [9], and altering GlutN subtype identities in specific layers during development has been shown to change the allocation and synaptic connectivity of corresponding GABAN subtypes [10–12].

In this study, we utilize both new and previously published scRNA-seq datasets from mice to begin deciphering the molecular code through which GlutN and GABAN subtypes communicate to generate the finely functioning mammalian cortex. First, we track gene expression dynamics across all GlutN and GABAN subtypes throughout corticogenesis. Then, we build an atlas to infer LR interactions between these subtypes. Finally, we validate the utility of our atlas by demonstrating that it confirms known interactions, fills gaps in previous knowledge by identifying receptors in well-characterized cell pair interactions, and discovers new LR pairs within the cadherin family that are crucial for perisomatic inhibition in deep versus superficial cortical layers.

A single neuron transcriptomics atlas covering mouse somatosensory cortex development

First, we performed scRNA-seq and snRNA-seq to interrogate gene expression in neuronal subtypes of the mouse somatosensory cortex at 6 ages from postnatal day (P)0 to juvenile stages (P30), encompassing the laminar positioning of GABAN subtypes and the establishment of their

specific synaptic connectivity with GlutN subtypes. Specifically, we interrogated gene expression: at P0 and P2 (main period for radial migration and laminar allocation of GABANs [,13,14]), P5 (programmed cell death of deep layers GABANs [15,16]), P8 (end of radial migration of GABANs and programmed cell death of superficial layers GABANs [17]), P16 (GABAN circuit refinement), and P30 (when the cortical network is established). We collected 58 027 high quality cells across these 6 time points (Fig. 1A). We integrated to this dataset previously published scRNA-seq datasets corresponding to earlier (E11.5 to E18.5) and later (P53-P102, adult) developmental stages [3,18–23], with the goal of covering the entirety of somatosensory corticogenesis (Fig. 1A). Notably, we integrated datasets from embryonic ganglionic eminences (GEs) in order to include early transcriptional development of future cortical GABANs [20–23]. In total, our analysis covers 17 time points ranging between E11.5 to adulthood (Fig. 1A).

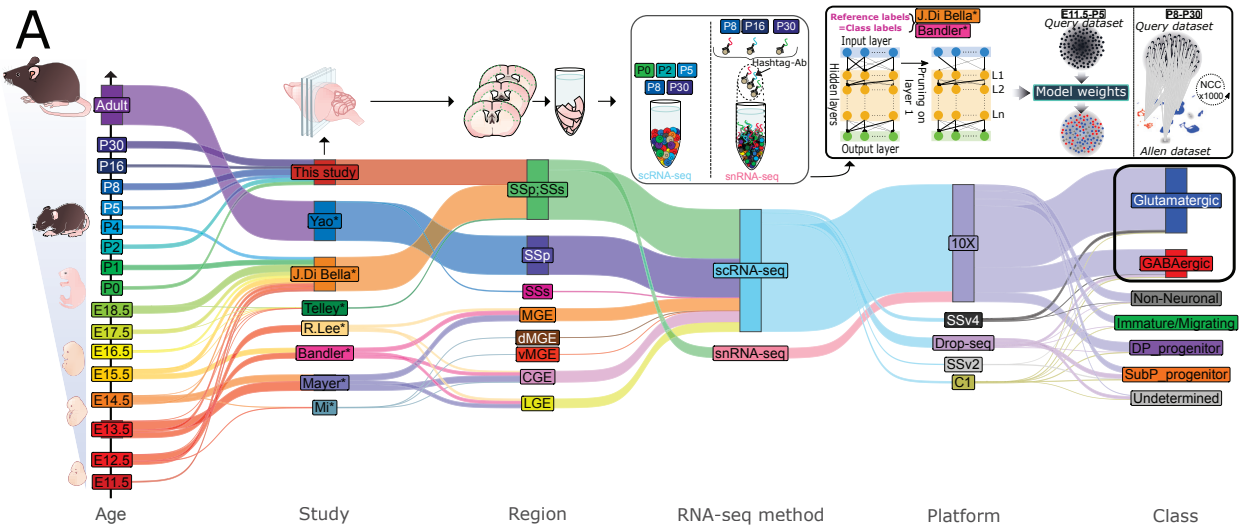
After stringent quality control and filtering (Fig. S1A-B and Methods), we first aimed to identify all postmitotic neuronal cell types and track their transcriptional dynamics over cortical development, from E11.5 to adulthood (Fig. S1A). For final cell type nomenclature, we chose as reference the current most comprehensive resource of transcriptomic *cell-types* in the somatosensory cortex made by the Allen Institute*, i.e.*, Yao et al. (2021b) [3] (from now on referred to as AllenRef21, see Methods), corresponding to the adult stage (P53-P102). Assigning cellular identities in early age datasets proved difficult with simple methods given that transcriptional signatures of specific *cell-types* evolve drastically during development [18,21]. To overcome this difficulty, we developed a hierarchical pipeline with different steps assigning identities at increasing resolution levels, similar to the method used by the Allen Institute [3].

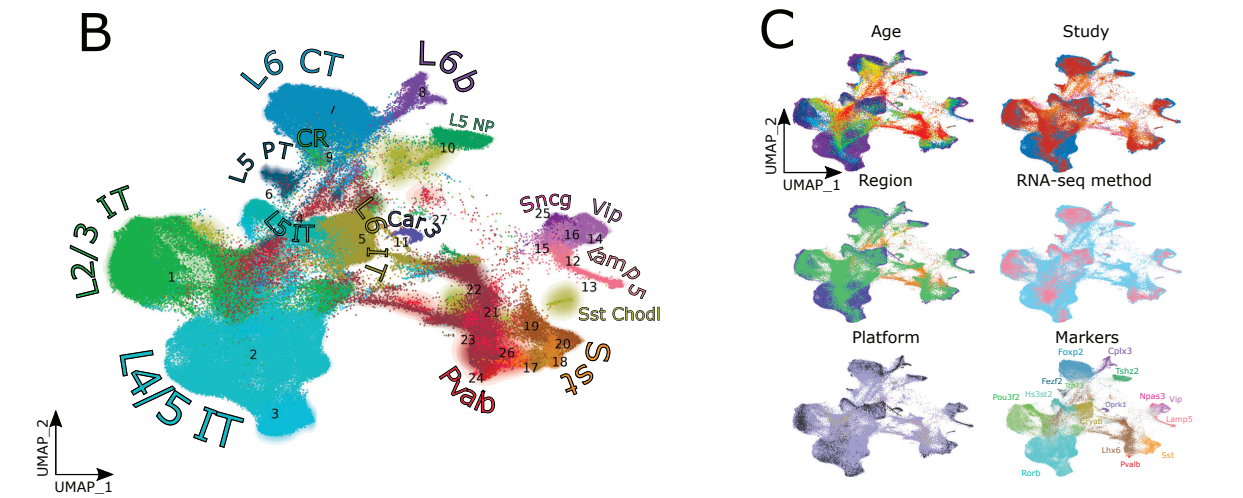
First, we assigned cell identities at a broad, low resolution level referred to as the *class* level (Glutamatergic, GABAergic, Non-Neuronal, Immature/Migrating, dorsal pallium progenitor and subpallium progenitor) [3]. For *classes* of embryonic and neonatal datasets (E11.5 to P5), we used an in-house Artificial Neural Network (ANN) approach and took as references the datasets of Di Bella et al. [19] (largest published perinatal dataset for somatosensory cortex, see Methods) and Bandler et al. [22] (corresponding to early GE GABANs destined to the cortex, see Methods). For P8 to P30, because cells were transcriptionally more similar to the adult reference, we used the nearest centroïd classifier of the R package scrattch.hicat [3,24], taking directly as reference the adult Allen dataset AllenRef21 [3] (Fig. S1A). To focus our hypothesis that a LR code between GlutNs and GABANs controls a large part of cortical microcircuits establishment, we kept only post-mitotic GlutNs and GABANs cells for downstream analyses (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 datasets coming from different labs using different dissociation methods and different library preparation methods (Fig. 1C, 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* and morpho-electro-connectomic types (*mec-types*) in patch-seq studies [1–3,25] (Fig. S2, S3; Methods). 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* - IT, extratelencephalic (ET), Other GlutNs - and 16 GABAN *cell-types* grouped in 5 *families* - Lamp5, Sst, Pvalb, Vip and Other GABANs - (Fig. 1B & 1D, Fig. S4).

In the UMAP embedding of all neurons from 17 ages, each GlutN *cell-type* formed a largely continuous gradient along time of corticogenesis, in which the gene expression of individual cells correlated well with their differentiation status (Fig. 1B, C). Likewise, each GABAN *cell-type* formed a temporal continuum (Fig. 1C, D). Cajal-Retzius (CR) cells were first detected at E11.5, defined by their expression of Trp73 (Fig. 1D).


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*
Tellev*
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
Lamp5
Sncg
Vip
Sst Chodl
Sst
Pvalb
L4/5 IT
UMAP_2
UMAP_1
C
Age
Study
Region
RNA-seq method
Platform
Markers

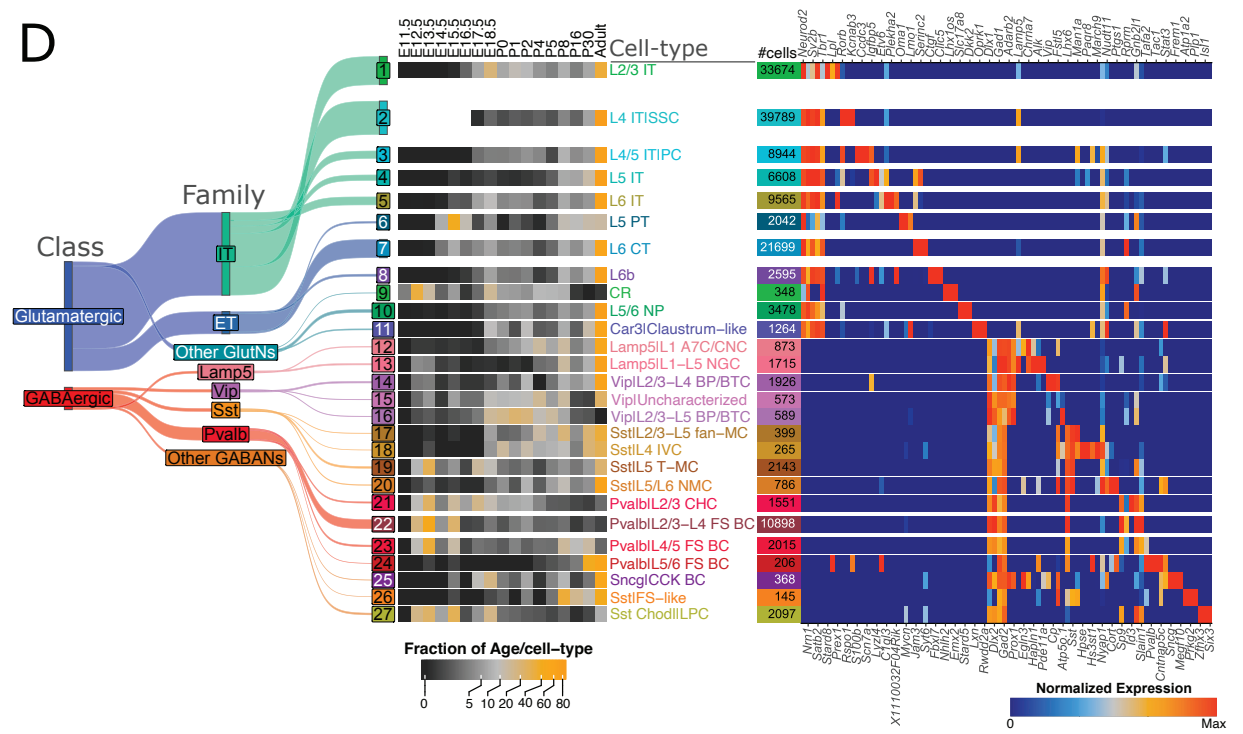

D
Class
Family
Glutamatergic
GABAergic
IT
ET
Other GlutNs
Lamp5
Vip
Sst
Pvalb
Other GABANs
Cell-type
#cells
L2/3 IT 33674
L4 ITISSC 39789
L4/5 ITIPC 8944
L5 IT 6608
L6 IT 9565
L5 PT 2042
L6 CT 21699
L6b 2595
CR 348
L5/6 NP 3478
Car3lClaustrum-like 1264
Lamp5lL1 A7C/CNC 873
Lamp5lL1-L5 NGC 1715
ViplL2/3-L4 BP/BTC 1926
ViplUncharacterized 573
ViplL2/3-L5 BP/BTC 589
SstlL2/3-L5 fan-MC 399
SstlL4 IVC 265
SstlL5 T-MC 2143
SstlL5/L6 NMC 786
PvalbIL2/3 CHC 1551
PvalbIL2/3-L4 FS BC 10898
PvalbIL4/5 FS BC 2015
PvalbIL5/6 FS BC 206
SncglCCK BC 368
SstlFS-like 145
Sst ChodlLPC 2097
Fraction of Age/cell-type
0 5 10 20 40 60 80
Normalized Expression
0 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 transcriptional *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 [26] (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) [27,28]. Our data uncover the molecular programs by which each neocortical *cell-type* shifts from intrinsic programs to programs allowing its interaction with other cells during development.

Spatial transcriptional dynamics over corticogenesis

Recent studies have indicated that, in adults, the GlutN IT *family* exhibits a spatial gradient of transcriptional variation along the cortical sheet [29]. We investigated the existence of such spatial transcriptomic gradients for each cell *family* at each age between E18.5 to P30. The correspondence between our datasets and published Patch-seq [1,2] and MERFISH [29] datasets allowed us to infer the average adult laminar position 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.

First, our analysis uncovered that the spatial transcriptional gradient described for the IT *family* in adults [29] 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 result is consistent with UL GlutNs being born later than DL GlutNs - due to cortical inside-out patterning - and therefore maturing later [18]. For GABANs, environment-sensing programs were already at play at E18.5 (Fig. S8, B and D), consistent with the hypothesis that these neurons need signals from previously settled GlutNs to reach their final position and connect the network appropriately [10–12,30].

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 for each gene within every cell *family* (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 areal patterning (Cux1, Fezf2, Reln, Fig. 2B-C) or in neuropeptidergic system maturation (Fig. 2D).

For validation, we verified whether the transcriptional landscapes of some specific LR pairs implicated in the processes of migration and/or synaptogenesis matched their known function. We took as a first example CXCL12-CXCR4 and CXCL12-ACKR3 that are essential for the tangential migration of GABANs [31,32]. The literature indicates that CXCL12 is expressed during the perinatal period by the meninges [33] and by immature subventricular zone/intermediate zone GlutNs [34]. A recent study also suggests expression in CR cells [35], but this requires confirmation. Concurrently, the CXCL12 co-receptors CXCR4 and ACKR3 are expressed by migrating GABANs [36,37], while CXCR4 is also found in CR cells [4]. The postnatal extinction of CXCL12 expression is thought to prompt the switch of GABANs from tangential migration to cortical plate invasion and lamination (Fig. 2E) [31,32,38]. Our expression atlas confirms these documented spatio-temporal expression patterns, as they are precisely mirrored in our transcriptional landscapes for Cxcl12, Cxcr4, and Ackr3 (Fig. 2E). Notably, the enrichment of Cxcl12 in CR cells was confirmed, and expressions of Cxcl12, Cxcr4, and Ackr3 were restricted to early time points (E18.5-P0), aligning with a role in migration but not 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) [39,40]. 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.

We next conducted a broader analysis of transcriptional landscapes using PCA and k-means clustering, which allowed us to identify 15 archetypal gene expression profiles spanning from early to late developmental stages and from deep to superficial cortical layers (Fig. 2F). This approach uncovered a wide range of spatiotemporal expression profiles, suggesting that certain genes are associated with early developmental processes, such as lamination, while others are linked to later stages, like synaptogenesis. Notably, the Cadherin gene family exhibited members across all 15 clusters (Fig. 2G), indicating a high degree of spatiotemporal diversity within this family.

Developmental gene expression functions have been made available using the gene expression function of a Shiny app that we developed and called *scLRSomatoDev* (weblink to be released), where researchers can visualize the spatiotemporal expression of any gene(s) in any neuronal *cell-type* over cortical development, using different visual representations (Fig. S11).

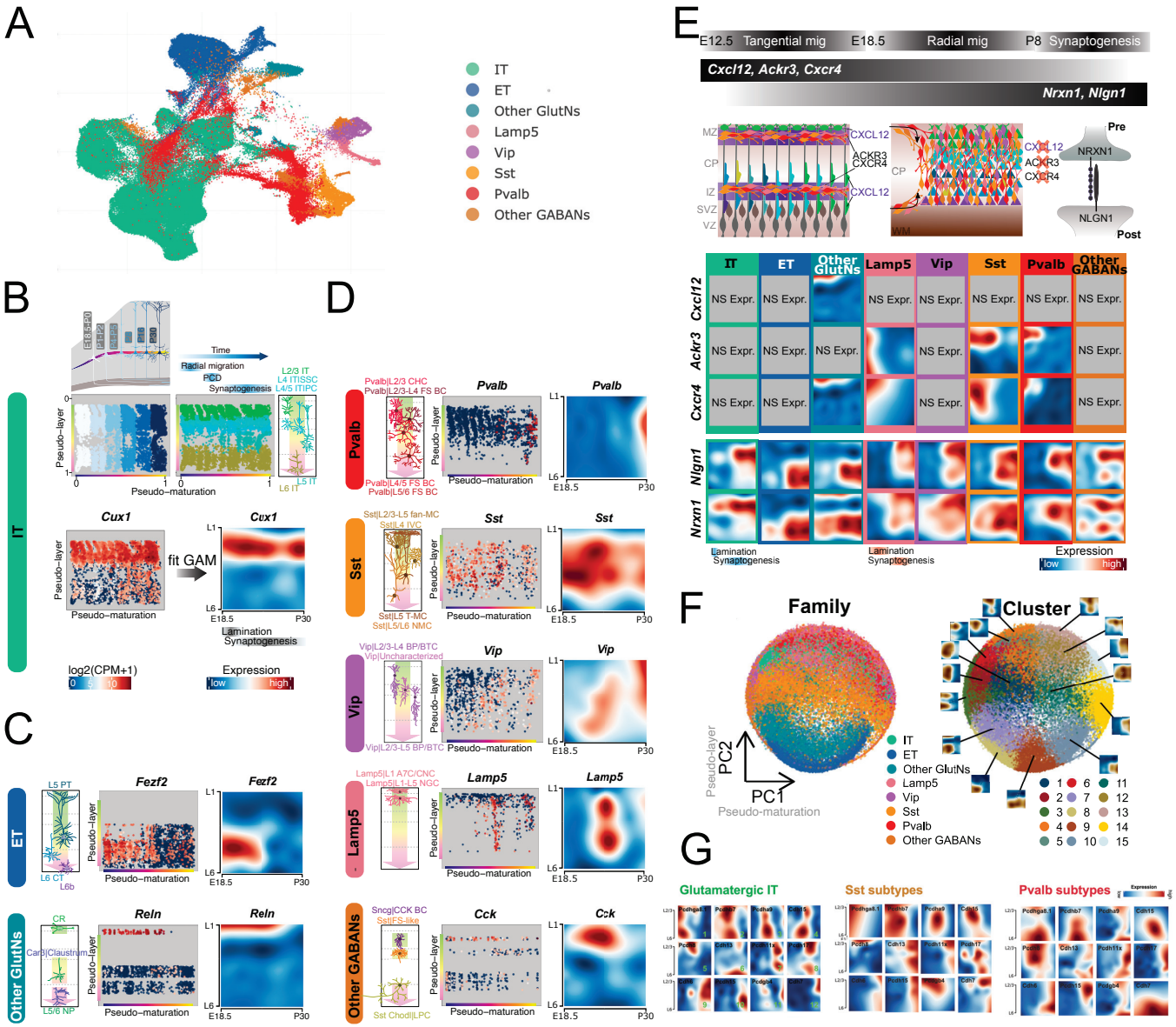


Figure 2. Spatiotemporal transcriptional dynamics for all neuronal cell-types during somatosensory cortex development. (A) UMAP visualization of all neurons color coded with cell family labels. (B) Top: 2D map in which cells are embedded according to their pseudo-maturation (x-axis) and pseudo-layer (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-S11.

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 Glutamatergic and GABAergic neurons play a pivotal role in orchestrating the stereotypical construction of the cortex. We thus anticipated observing highly correlated spatiotemporal expression patterns between GlutN ligands and their corresponding GABAN receptors, and vice versa, compared to randomly selected genes within the same *cell-types*. To explore this, we first created a curated ligand-receptor (LR) database by aggregating existing published LR databases and manually adding additional published LR pairs (Fig. 3A & S12, Table S1, see Methods). This LR database, which we called LR_DB, contains 7918 curated LR pairs with 1421 ligands and 1233 receptors. We categorized the ligands and receptors based on molecular functions or neurodevelopmental stages (see Methods).

Interestingly, for LR pairs associated with neurodevelopmental processes in general, such as the cadherin family, the transcriptional landscapes of ligands in IT neurons and of their corresponding receptors in GABAergic neurons were highly correlated (Fig. 3B, Fig. S13-S16). This finding suggests an essential role for these interactions in GlutN-GABAN subtype communication during development and supports recent hypotheses linking cadherins with areal specialization [30,41] and synaptic specificity [42]. LR pairs associated with specific developmental processes, i.e., migration, differentiation, cell death, cell adhesion and synaptogenesis (Table S2 and Methods) displayed varying levels of correlation in their spatiotemporal landscapes (Fig. 3B). Synaptogenesis was the process with the highest LR landscape correlations, suggesting that synaptogenesis relies more than earlier neurogenic processes on cell-cell communications while some processes like cell death or differentiation might depend more on intrinsic mechanisms (Fig. 3B).

To directly identify which ligand-receptor (LR) pairs drive cell-cell interactions and influence cortical development, we constructed a comprehensive LR atlas that infers significant LR interactions between neuronal cell types at various stages from E18.5 to adulthood. The inference of these interactions is based on an integrative approach that combines data from temporal cell-gene expression matrices with information from our LR database (Fig. 3C &S17, see Methods for further details).

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 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).
The quantification of intracellular signaling downstream of ligand-mediated receptor activation in the target cell type produced an intracellular score (Fig. 3C and S17). This score assesses the reliability of the inferred LR communication and helps identify the preferred receptor(s) utilized when a ligand interacted with multiple receptors.

Predictions of LR interactions over cortex development are accessible via a Shiny app called *scLRSomatoDev* (to-be-published, 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).

By which molecular code do specific LR pairs control position and connectivity of *cell-types* ? We performed a general analysis of our LR predictions in order to address this question. For all 729 *cell-type* combinations from E18.5 to adulthood, we analyzed the number of predicted LR pairs, examining the total number, unique pairs, widely shared pairs (shared by over 400 cell pairs), and very rare pairs (shared by fewer than 10 cell pairs) (Fig. 3D). The total number of LR pairs for specific *cell-type* combinations reached as high as ~600, while unique LR pairs were almost non-existent. Widely shared LR pairs were relatively uncommon, with fewer than ~30, whereas very rare LR pairs were more frequent, with up to ~60 identified. These findings suggest that a diverse set of LR pairs, blending both rare and widely shared interactions in specific combinations, underlies the unique cell-cell communication and connectivity patterns observed among different cortical neuron subtypes.

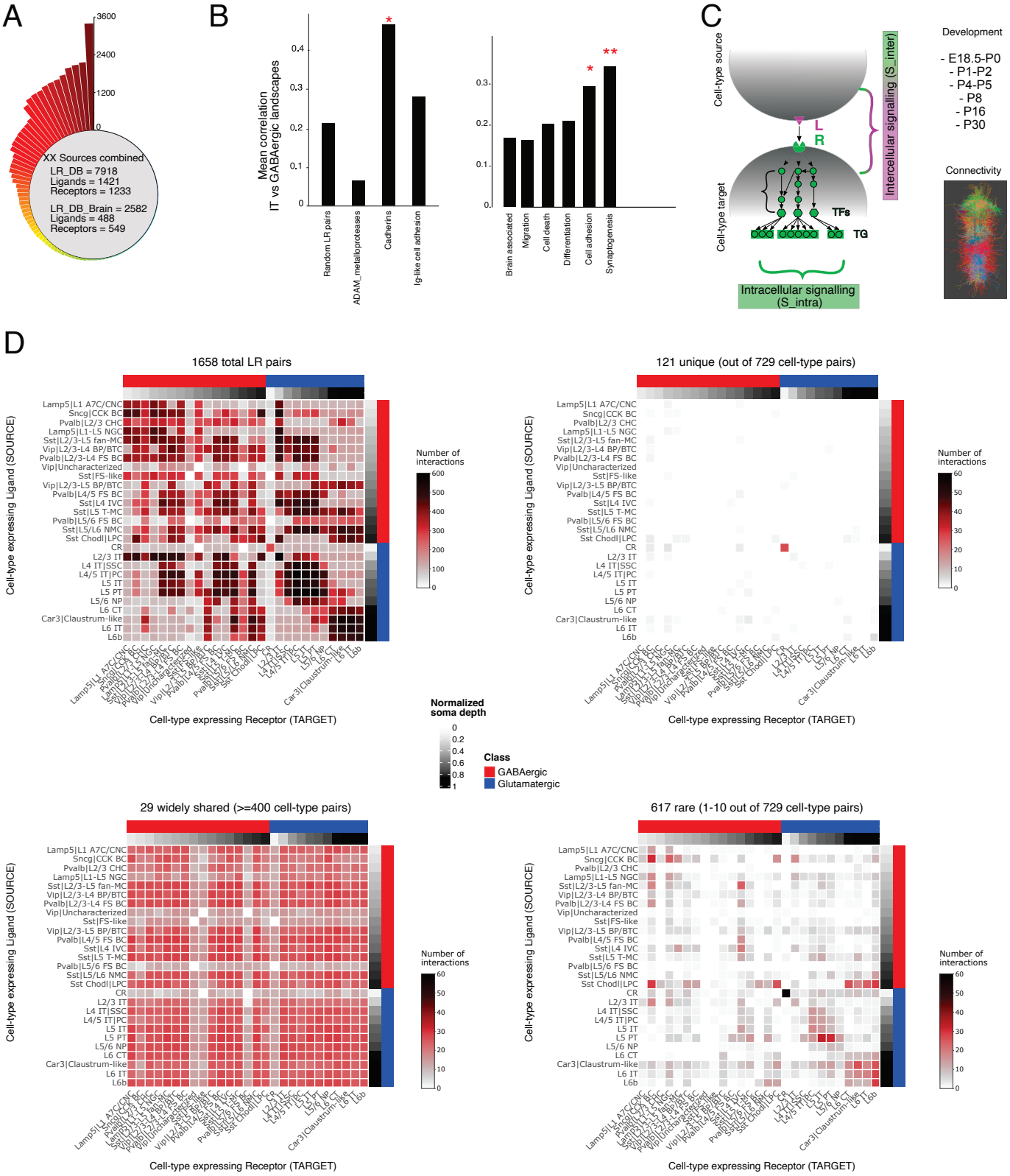

A
3600
2400
1200
0
XX Sources combined
LR_DB = 7918
Ligands = 1421
Receptors = 1233
LR_DB_Brain = 2582
Ligands = 488
Receptors = 549
B
Mean correlation IT vs GABAergic landscapes
Random LR pairs
ADAM_metalloproteases
Cadherins
Ig-like cell adhesion
Brain associated
Migration
Cell death
Differentiation
Cell adhesion
Synaptogenesis
C
Cell-type source
Cell-type target
L
R
TFs
TG
Intercellular signalling (S_inter)
Intracellular signalling (S_intra)
Development
- E18.5-P0
- P1-P2
- P4-P5
- P8
- P16
- P30
Connectivity
D
1658 total LR pairs
121 unique (out of 729 cell-type pairs)
29 widely shared (>=400 cell-type pairs)
617 rare (1-10 out of 729 cell-type pairs)
Cell-type expressing Ligand (SOURCE)
Cell-type expressing Receptor (TARGET)
Number of interactions
Normalized soma depth
Class
GABAergic
Glutamatergic

Figure 3: A LR interactome prediction atlas between GlutNs and GABANs over cortex development.

(A) Curated database of 7918 LR pairs. DB, database; L, ligand; R, receptor. (B) Left, intersections between LR pairs involved in different developmental processes. Right, high correlation of neurodevelopmental process-associated ligand/receptor transcriptional landscapes suggest a LR molecular code for neuronal adhesion and synaptogenesis. (C) Computational method for the inference of LR-mediated interactions. DB, database; L, ligand; TF, transcription factor; TG, target gene; R, receptor; S_inter, score for intercellular communication; S_intra, score for intracellular signaling. See Fig. S18 and Methods. (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 S12-19.

Temporal analysis suggests 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 LRs (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 [43]. 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.

Then, we calculated the percentage of LRs known to intervene in each of the 5 well-known successive neurodevelopmental processes (migration, differentiation, cell death, cell adhesion and synaptogenesis) that was effectively utilized for development. Strikingly, only synaptogenesis-associated LRs, and to a lesser extent cell adhesion-associated LRs, were predicted as strongly utilized across E18.5-adulthood (Fig. 4C). LRs associated with migration, differentiation and cell death were poorly utilized, suggesting that these processes rely less than synaptogenesis on LR-mediated cellular communications and more on cell-intrinsic processes. Notably, Cadherins were

quite strongly utilized from P4 onwards, indicating that Cadherins might regulate 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 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 all developmental stages (Fig. 4D). In contrast, LRs associated with autism and epilepsy were utilized at much higher rates, occasionally reaching up to 50%, and their prevalence increased progressively over developmental stages (Fig. 4D). This suggests that the etiology of autism and epilepsy may be more closely related to LR-mediated cell-cell communications 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 disease-associated LRs predicted to be utilized during cortical development were enriched in neurodevelopmental biological processes.

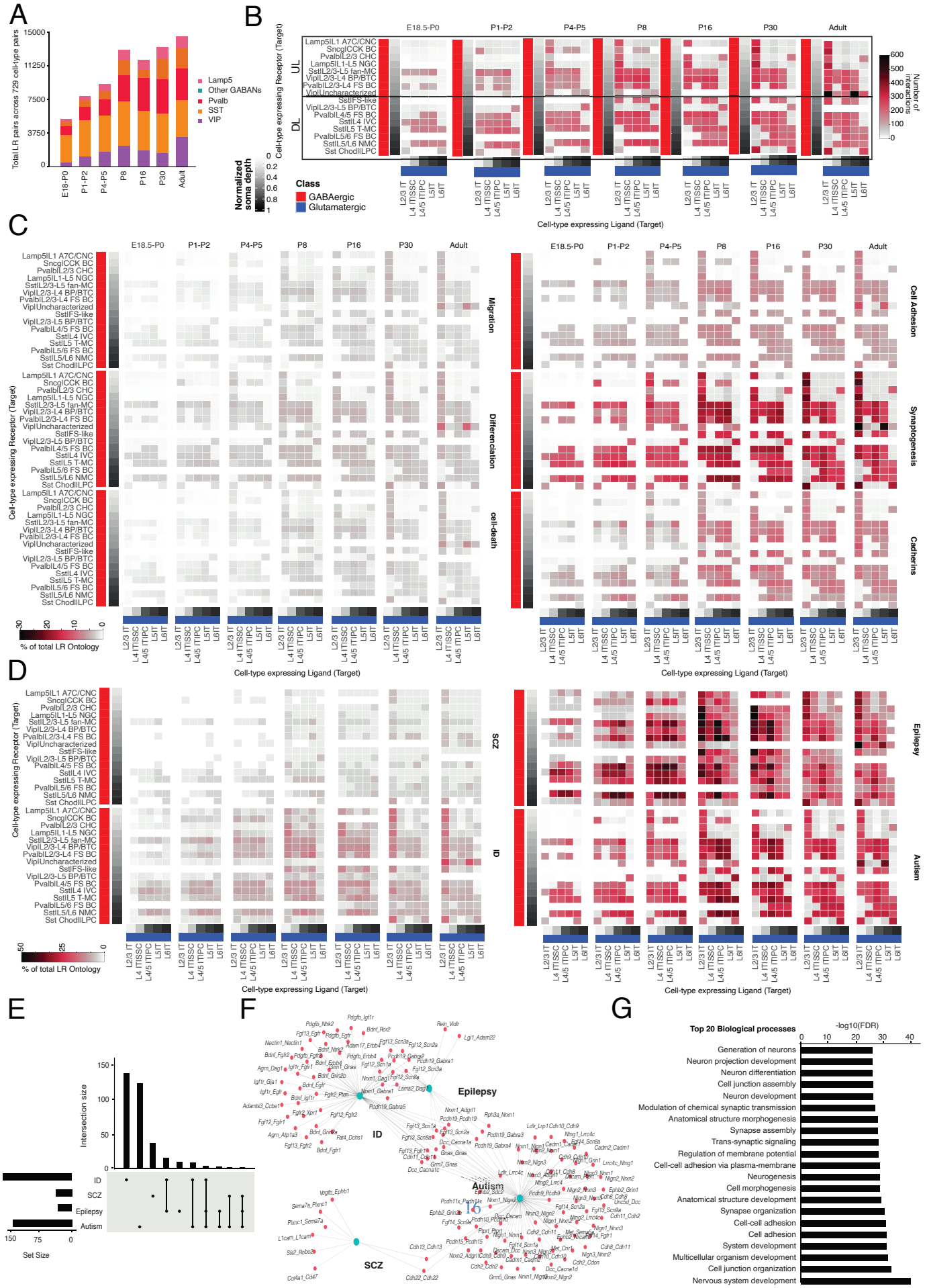

A
Total LR pairs across 729 cell-type pairs
Lamp5
Other GABANs
Pvalb
SST
VIP
E18-P0
P1-P2
P4-P5
P8
P16
P30
Adult
B
Cell-type expressing Receptor (Target)
Cell-type expressing Ligand (Target)
Number of interactions
Normalized soma depth
Class
GABAergic
Glutamatergic
UL
DL
C
Migration
Differenciation
cell-death
Cell Adhesion
Synaptogenesis
Cadherins
% of total LR Ontology
D
SCZ
ID
Epilepsy
Autism
E
Intersection size
Set Size
F
Epilepsy
ID
Autism
SCZ
G
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 disease associated LR pairs.

Validation of the LR 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 [44] (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 where when target cells were Pvalb BCs and Vip|L2/3−L5 BP/BTCs, two GABAN types known to require Erbb4 signaling for the formation of excitatory synapses [7,8,45]. Overall, our atlas supports the role of the Nrg3-Erbb4 interaction in excitatory synapse formation onto specific GABAN types during corticogenesis.

Utilizing the atlas to enhance understanding of known interactions: the Cbln4 example

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 [6,46] (Fig. 5D).

Then, we asked whether our gene expression and LR atlases can bring new information to known but incompletely understood cell-cell communications. We focused on CBLN4, a ligand that mediates the formation of synapses between axonal termini of Sst neurons and apical dendrites of GlutNs in the mouse cortex (Fig. 5D). 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 [46,47] (Fig. 5E) 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. 5F). 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. 5F) - more unexpectedly, most GABAN *cell-type*s were also predicted as "target" cells. Finally, the LR atlas revealed that among the 3 known Cbln4 receptors (Dcc, Glud1 and Neo1), only Glud1 and Neo1 were predicted to be involved based on intercellular signaling scores (Fig. 5G). These scores exhibited an increase from P4 to P8, which aligns with earlier findings underscoring the significance of Cbln4 during this developmental period [6] 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 a previous study suggested that GLUD1 might be the CBLN4 receptor in cortical GlutNs [46], the intracellular score predicts that NEO1 downstream signaling is much more active than GLUD1 signaling in "target" neurons (Fig. 5G). Neo1-associated genes in "target" GlutNs were notably linked with axon guidance and TGF-beta signaling pathways (Fig. 5G). Based on these predictions, we proceeded to investigate the presence of a direct interaction between CBLN4 and NEO1 or GLUD1 in situ. Utilizing proximity ligation assay (PLA), we identified robust CBLN4-NEO1 interactions in L1-2, whereas interactions between CBLN4 and GLUD1 were either weak or absent (Fig. 5H). 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.

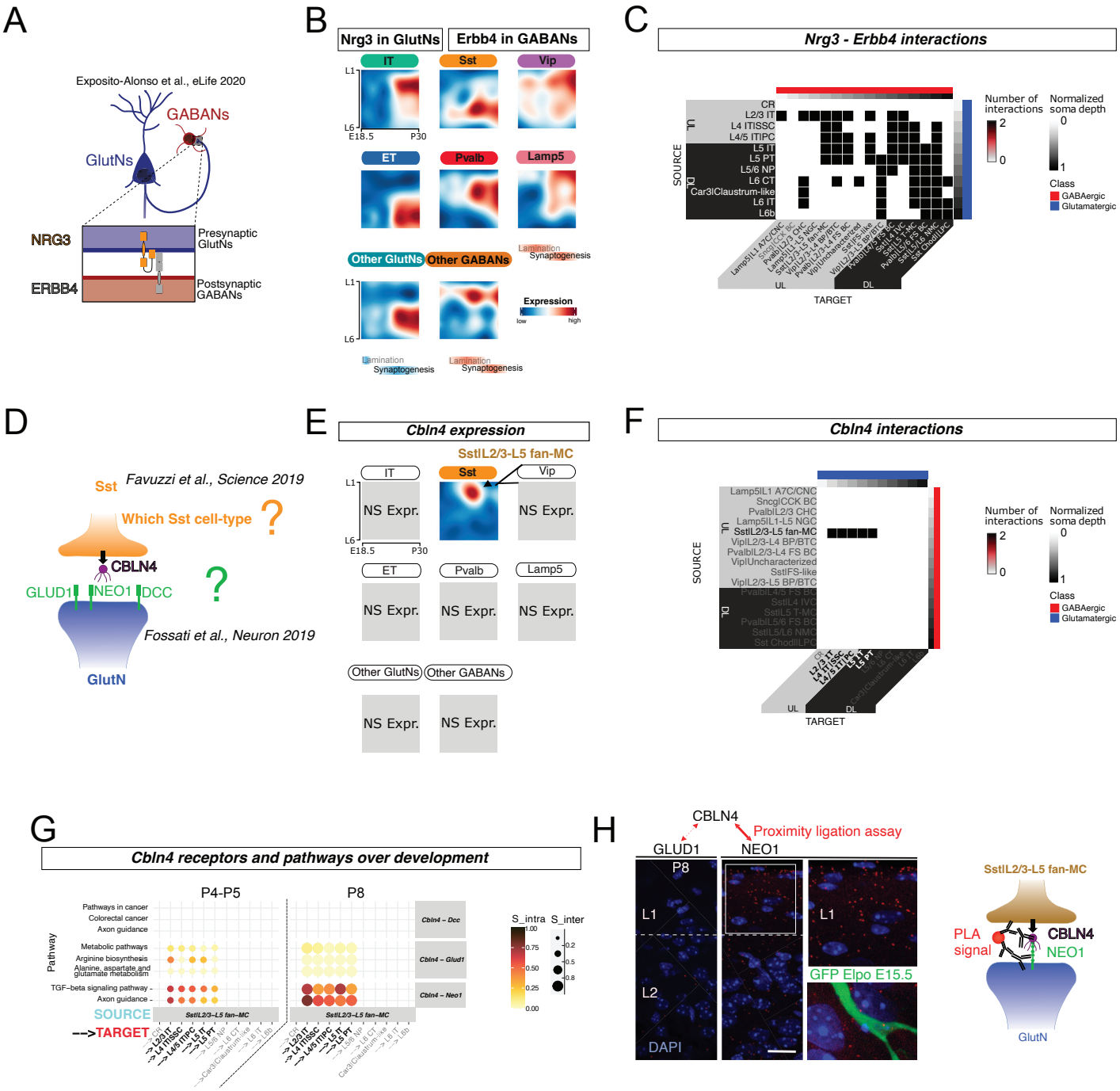


Figure 5. Validation and Application of the Ligand-Receptor Atlas for Knowledge Completion. (A) Diagram illustrating Nrg3-Erbb4 interaction between glutamatergic and GABAergic cells. (B) Nrg3 and Erbb4 expressions in transcriptional landscapes of glutamateric and GABAergic neuron *families*, respectively. (C) Heatmap depicting predictions of Nrg3-Erbb4 mediated cell-cell interactions between GlutN and GABAN *cell-types*. 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) Diagram illustrating Cbln4-

mediated inhibitory synapse formation from Sst cells to GlutN in the cortex. (E) Transcriptional landscapes showing Cbln4 expression across different neuronal families. (F) Heatmap of LR pairs involving Cbln4. (G) 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. (H) Left: Proximity ligation assay (PLA) results for interactions between CBLN4 and NEO1 (left) and CBLN4 and GLUD1 (middle) in cortical layers L1-2; Right: Summary of the experimental findings and predictions resulting from our analysis.

Using our LR atlas to uncover 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 [41,42,48]. 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 [49,50], was particularly interesting as it showed highest expression in DL GlutNs and in MGE-derived GABAN *families* (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 [6] (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 mediates interactions that are mainly between Sst / Pvalb GABANs and DL GlutNs. It was a strict rule at specific time points like P4-P5 (peak synaptogenesis) 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 predictions strongly suggest that Cdh13 is a candidate molecule crucial for establishing connectivity between MGE-derived Pvalb and Sst GABANs and DL GlutNs. We used the fact that synaptic connectivity between Pvalb and GlutNs is easily measurable with SYT2 staining [51] 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) while the mean volume of $SYT2^+$ puncta was unaltered (not shown). This result shows that post-synaptic CDH13 is necessary for proper Pvalb BC-perisomatic inhibition of DL GlutN somata.

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 GABANs of SL (Fig. 6E). Pcdh8 was also detected in L5 PT and Car3|Claustrum-like GlutNs, albeit to lower levels than L2/3 IT neurons. Our LR atlas predicted 15 GlutN-GABAN interactions with only 3/11

"source" *cell-types*, L2/3 IT, L5 PT and Car3|Claustrum-like neurons, and 9/16 "target" *cell-types*, including Pvalb and Sst neurons (Fig. 6F, G). 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 ofL2/3 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, in a negative way (Fig. 6H). Thus, our LR atlas allowed us to conclude that different cadherins modulate perisomatic inhibition in different ways depending on the layer examined.

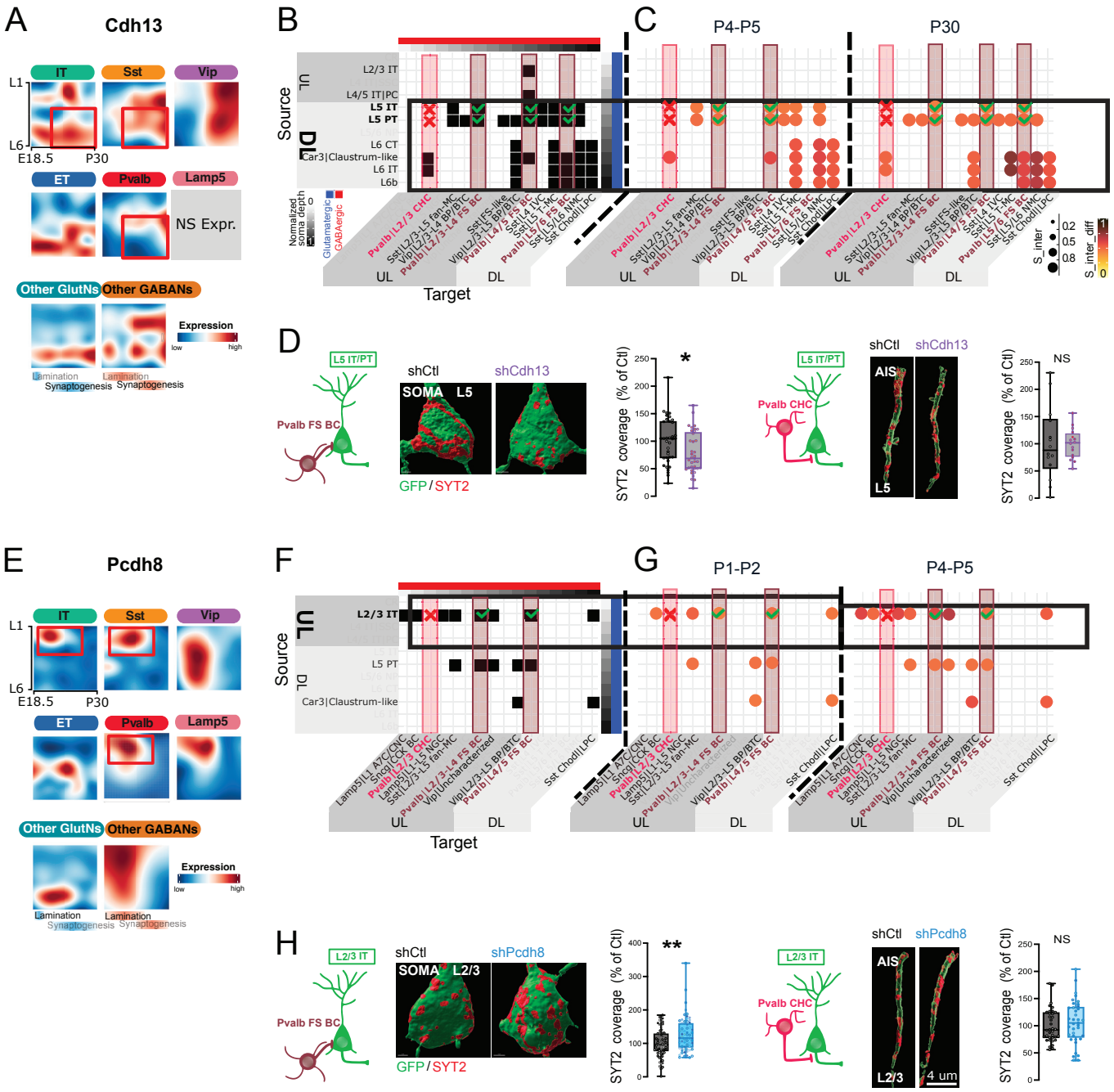


Figure 6. Atlas predictions identifies two cadherins regulating perisomatic inhibition in deep and superficial layers. (A) Cdh13 expression in the different families represented as transcriptional landscapes. (B) Heatmap representing Cdh13-Cdh13 predicted interactions between GlutN SOURCE cell-types and GABAN TARGET cell-types that are common between at least 2 contiguous ages. (C) Cdh13-Cdh13 utilization predictions across GlutNs and GABANs cell-types at P4-P5 and P30. (D) Perisomatic (left) and axon initial segment (AIS, right) inhibitory innervations of E13.5-electroporated L5 GlutNs by SYT2-expressing Pvalb BC and CHC boutons, respectively, when electroporarted GlutNs are control or knocked-down for Cdh13 via shCdh13 (Mann-Whitney test, *p-value < 0.05, n=3-4 mice/condition, > 36 and 14 neurons/condition for soma and AIS, respectively; each animal is represented by a color code). (E) Pcdh8 expression in the different families represented as transcriptional landscapes. (F) Heatmap representing the Pcdh8-Pcdh8 predicted interactions between GlutN SOURCE cell-types and GABAN TARGET cell-types that are common between 2 contiguous ages. (G) Analysis of the homophilic Pcdh8-Pcdh8 pair across P1-P2 and P4-P5. (H) Perisomatic (left) and axon initial segment (AIS, right) inhibitions of E15.5-electroporated L2/3 GlutNs by SYT2-expressing Pvalb BC and CHC boutons, respectively, when electroporarted GlutNs are control or knocked-down for Pcdh8 via shPcdh8 (Mann-Whitney test, **p-value < 0.01, n=3-4 mice/condition,> 61 and 35 neurons/condition, each animal is represented by a color code).

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 [10–12]. 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.

We characterized the transcriptome of all biologically-meaningful neuron types over the course of cortical development, and 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 is thought to shape most if not all 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* [52–54], 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 [55–58]. 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 [55]. 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 the LR atlas reveals that specific neuronal connections are defined by unique combinations of LR pairs for each of the 729 neuronal *cell-type* pairs. Most combinations consist of approximately 40 "pan" LR pairs, shared by over 400 *cell-type* pairs, along with several relatively rare pairs, for example some shared by fewer than 10 *cell-type* pairs. Ligand-receptor pairs unique to a single *cell-type* pair are very rare.

Recent studies have shown that members of the cadherin family of adhesion molecules are differentially expressed in subtypes of cortical excitatory and inhibitory neurons [30,41]. 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. S10-S13), 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 [59].

In contrast to the promotory effect of CDH13 in DL, we find that PCDH8 acts as a suppressor of perisomatic inhibition in SL. Thus, different cadherins or CAM family members might affect synaptogenic processes in opposite ways and in different cortical layers or areas. Nonetheless, given that variations in CDH13 and PCDH8 are associated with neurodevelopmental disorders [49,50,60], we propose that abnormal Pvalb BC-mediated perisomatic inhibition in specific cortical layers might be central to some of these disorders.

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 [61]. 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 [44]. 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 [61].

Currently, our knowledge of the spatial positions of cortical cell types relative to each other is limited to adult stages [62,4], 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 sharing his model Torch necessary to generate the connectivity score, 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

Supplementary Materials

Materials and Methods

Supplementary Text

Figs. S1 to S#

Tables S1 to S5

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. 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.
5. 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).
6. Favuzzi, E. *et al.* Distinct molecular programs regulate synapse specificity in cortical inhibitory circuits. *Science* **363**, 413–417 (2019).
7. Fazzari, P. *et al.* Control of cortical GABA circuitry development by Nrg1 and ErbB4 signalling. *Nature* **464**, 1376–1380 (2010).
8. Bernard, C. *et al.* Cortical wiring by synapse type–specific control of local protein synthesis. *Science* **378**, eabm7466 (2022).
9. 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).
10. Lodato, S. *et al.* Excitatory Projection Neuron Subtypes Control the Distribution of Local Inhibitory Interneurons in the Cerebral Cortex. *Neuron* **69**, 763–779 (2011).
11. Ye, Z. *et al.* Instructing Perisomatic Inhibition by Direct Lineage Reprogramming of Neocortical Projection Neurons. *Neuron* **88**, 475–483 (2015).
12. Wester, J. C. *et al.* Neocortical Projection Neurons Instruct Inhibitory Interneuron Circuit Development in a Lineage-Dependent Manner. *Neuron* **102**, 960-975.e6 (2019).
13. Miyoshi, G. & Fishell, G. GABAergic Interneuron Lineages Selectively Sort into Specific Cortical Layers during Early Postnatal Development. *Cereb. Cortex* **21**, 845–852 (2011).
14. 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).
15. Southwell, D. G. *et al.* Intrinsically determined cell death of developing cortical interneurons. *Nature* **491**, 109–113 (2012).
16. Wong, F. K. *et al.* Pyramidal cell regulation of interneuron survival sculpts cortical networks. *Nature* **557**, 668–673 (2018).
17. Priya, R. *et al.* Activity Regulates Cell Death within Cortical Interneurons through a Calcineurin-Dependent Mechanism. *Cell Rep.* **22**, 1695–1709 (2018).
18. Telley, L. *et al.* Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364**, eaav2522 (2019).
19. Di Bella, D. J. *et al.* Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature* **595**, 554–559 (2021).
20. Mi, D. *et al.* Early emergence of cortical interneuron diversity in the mouse embryo. *Science* **360**, 81–85 (2018).
21. Mayer, C. *et al.* Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).
22. Bandler, R. C. *et al.* Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).
23. Lee, D. R. *et al.* Transcriptional heterogeneity of ventricular zone cells in the ganglionic eminences of the mouse forebrain. *eLife* **11**, e71864 (2022).
24. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
25. 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).
26. Street, K. *et al.* Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
27. 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).
28. 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).
29. Zhang, M. *et al.* Spatially resolved cell atlas of the mouse primary motor cortex by MERFISH. *Nature* **598**, 137–143 (2021).

30. Libé-Philippot, B. *et al.* Auditory cortex interneuron development requires cadherins operating hair-cell mechanoelectrical transduction. *Proc. Natl. Acad. Sci.* **114**, 7765–7774 (2017).
31. 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).
32. López-Bendito, G. *et al.* Chemokine Signaling Controls Intracortical Migration and Final Distribution of GABAergic Interneurons. *J. Neurosci.* **28**, 1613–1624 (2008).
33. Borrell, V. & Marín, O. Meninges control tangential migration of hem-derived Cajal-Retzius cells via CXCL12/CXCR4 signaling. *Nat. Neurosci.* **9**, 1284–1293 (2006).
34. 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).
35. Causeret, F., Moreau, M. X., Pierani, A. & Blanquie, O. The multiple facets of Cajal-Retzius neurons. *Development* **148**, dev199409 (2021).
36. Stumm, R. K. *et al.* CXCR4 Regulates Interneuron Migration in the Developing Neocortex. *J. Neurosci.* **23**, 5123–5130 (2003).
37. Wang, Y. *et al.* CXCR4 and CXCR7 Have Distinct Functions in Regulating Interneuron Migration. *Neuron* **69**, 61–76 (2011).
38. Tanaka, D. H. *et al.* CXCR4 Is Required for Proper Regional and Laminar Distribution of Cortical Somatostatin-, Calretinin-, and Neuropeptide Y-Expressing GABAergic Interneurons. *Cereb. Cortex* **20**, 2810–2817 (2010).
39. 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).
40. 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. *Curr. Opin. Neurobiol.* **22**, 412–422 (2012).
41. 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).
42. Basu, R., Taylor, M. R. & Williams, M. E. The classic cadherins in synaptic specificity. *Cell Adhes. Migr.* **9**, 193–201 (2015).
43. Tuncdemir, S. N. *et al.* Early Somatostatin Interneuron Connectivity Mediates the Maturation of Deep Layer Cortical Circuits. *Neuron* **89**, 521–535 (2016).
44. Exposito-Alonso, D. *et al.* Subcellular sorting of neuregulins controls the assembly of excitatory-inhibitory cortical circuits. *eLife* **9**, e57000 (2020).
45. Batista-Brito, R. *et al.* Developmental Dysfunction of VIP Interneurons Impairs Cortical Circuits. *Neuron* **95**, 884-895.e9 (2017).
46. 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).
47. 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.* **119**, e2123421119 (2022).
48. Lv, X. *et al.* Patterned cPCDH expression regulates the fine organization of the neocortex. *Nature* **612**, 503–511 (2022).
49. 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).
50. 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).
51. 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).
52. 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).
53. Cabello-Aguilar, S., Fau, C., Lacroix, M. & Colinge, J. SingleCellSignalR: inference of intercellular networks from single-cell transcriptomics.
54. 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).
55. 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).
56. Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).

57. 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. *Brief. Bioinform.* **22**, 988–1005 (2021).
58. 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).
59. 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).
60. Butler, M., Rafi, S., Hossain, W., Stephan, D. & Manzardo, A. Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. *Int. J. Mol. Sci.* **16**, 1312–1335 (2015).
61. Unterauer, E. M. *et al.* Spatial proteomics in neurons at single-protein resolution. *Cell* **187**, 1785-1800.e16 (2024).
62. 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.