Article

Transcriptomic and spatial organization of telencephalic GABAergic neurons

https://doi.org/10.1038/s41586-025-09296-1

Received: 10 June 2024

Accepted: 12 June 2025

Published online: 5 November 2025

Open access

Check for updates

Cindy T. J. van Velthoven[1]✉, Yuan Gao[1], Michael Kunst[1], Changkyu Lee[1], Delissa McMillen[1], Anish Bhaswanth Chakka[1], Tamara Casper[1], Michael Clark[1], Rushil Chakrabarty[1], Scott Daniel[1], Tim Dolbeare[1], Rebecca Ferrer[1], Jessica Gloe[1], Jeff Goldy[1], Junitta Guzman[1], Carliana Halterman[1], Windy Ho[1], Mike J. Huang[1], Katelyn James[1], Rachel McCue[1], Beagan Nguy[1], Trangthanh Cardenas[1], Kara Ronellenfitch[1], Emma D. Thomas[1], Amy Torkelson[1], Chelsea M. Pagan[1], Lauren Kruse[1], Nick Dee[1], Lydia Ng[1], Jack Waters[1], Kimberly A. Smith[1], Bosiljka Tasic[1], Zizhen Yao[1] & Hongkui Zeng[1]✉

The telencephalon of the mammalian brain contains multiple regions and circuits that have adaptive and integrative roles in a variety of brain functions. GABAergic neurons in the telencephalon are diverse; they have many circuit functions, and dysfunction of these neurons has been implicated in various brain disorders[1–3]. Here we conducted a systematic and in-depth analysis of the transcriptomic and spatial organization of GABAergic neuronal types in all regions of the mouse telencephalon and their developmental origins. This was accomplished using 611,423 young adult single-cell transcriptomes and 614,569 single-cell transcriptomes collected from multiple prenatal and postnatal developmental timepoints. We present a hierarchically organized adult telencephalic GABAergic neuronal cell-type taxonomy of 7 classes, 52 subclasses, 284 supertypes and 1,051 clusters, as well as a corresponding developmental taxonomy of 1,688 clusters across ages from embryonic day 7 to postnatal day 14. Detailed charting efforts reveal extraordinary complexity whereby relationships among cell types reflect both spatial locations and developmental origins. Transcriptomically and developmentally related cell types are often found in distant and diverse brain regions, indicating that long-distance migration and dispersion is a common characteristic of nearly all classes of telencephalic GABAergic neurons. Moreover, we find various spatial dimensions of both discrete and continuous variation among related cell types that are correlated with gene expression gradients. Lastly, we find that cortical, striatal and some pallidal GABAergic neurons undergo extensive postnatal diversification, whereas septal, preoptic and most pallidal GABAergic neuronal types emerge in a burst during the embryonic stage with limited postnatal diversification. Overall, the telencephalic GABAergic cell-type taxonomy will serve as a foundational reference for molecular, structural and functional studies of cell types and circuits by the entire community.

The telencephalon, the anterior part of the mammalian brain, encompasses key structures that function as command centres for hierarchically organized brain networks driving behaviour and cognition. It consists of the cerebral cortex (CTX; shell) and cerebral nuclei (CNU; core), derived from the developmental pallium and subpallium, respectively. The CTX consists of isocortex, hippocampal formation (HPF), olfactory areas (OLF) and cortical subplate (CTXsp), and the CNU consists of striatum (STR) and pallidum (PAL). Each structure contains functionally distinct subregions (together, over 100 regions; see Supplementary Table 1 for anatomical ontology from the Allen Mouse Brain Common Coordinate Framework version 3 (CCFv3; RRID: SCR_020999))[4], each containing many cell types.

A general organizing principle of the telencephalon involves several parallel corticostriatopallidal circuit pathways[5]. In a highly simplified view, the dorsal pathway, involving the isocortex, caudate putamen (CP) and globus pallidus external and internal segments (GPe and GPi), mediates sensory and motor functions; the ventral pathways, from the prefrontal cortex to the nucleus accumbens (ACB) and ventral PAL (PALv), or from the amygdala to the bed nucleus of the stria terminalis (BST), regulate affective functions; and the hippocampal–septal pathway supports learning and cognition. These circuits rely on highly diverse glutamatergic and GABAergic neurons, of which the heterogeneity supports regional specialization. Understanding their molecular identity, spatial distribution and inter-relationships

[1]Allen Institute for Brain Science, Seattle, WA, USA. ✉e-mail: cindy.vanvelthoven@alleninstitute.org; hongkuiz@alleninstitute.org

is key to understanding how brain functions emerge from this complexity.

Glutamatergic excitatory neurons are the dominant neuronal class of the CTX and arise from the ventricular and subventricular zones (SVZs) of the developing pallium. By contrast, most telencephalic GABAergic neurons originate from five subpallial progenitor domains: medial (MGE), caudal (CGE) and lateral (LGE) ganglionic eminences, septum and preoptic area (POA). These neurons are born embryonically and migrate along defined routes to disperse throughout the forebrain[3,6,7]. Many GABAergic neurons migrate to the pallium, where they intermingle with glutamatergic neurons. Cell fate specification in the progenitor domains is orchestrated by a combination of transcription factors (TFs) and morphogens[6–15].

We conducted a systematic and in-depth analysis of the transcriptomic and spatial organization of GABAergic neuronal types in all regions of the mouse telencephalon and their developmental origins, using our high-resolution whole-brain atlas[16] and an additional single-cell RNA-sequencing (scRNA-seq) dataset from prenatal and postnatal development. Our study identified a diverse set of distinct cell types and molecular gradients that collectively shape the cellular landscape that underlies the functions of various telencephalic regions and circuits. We defined a comprehensive set of TFs that delineate all major subclasses and supertypes of the adult-stage telencephalic GABAergic neurons, linking their expression to specific developmental regions. This enabled us to infer the developmental origins of these diverse neuronal types. The results reveal two prominent features of the telencephalic GABAergic neurons: (1) long-distance migration and dispersion as a common characteristic of nearly all classes of telencephalic GABAergic neurons; (2) differential extents of postnatal diversification between cortical and striatal, and septal, preoptic and most pallidal GABAergic neuronal types.

GABAergic neuronal types in the mouse telencephalon

In a previous study, we created a high-resolution transcriptomic and spatial cell-type atlas of the entire adult mouse brain, combining scRNA-seq and multiplexed error-robust fluorescence in situ hybridization (MERFISH) spatial transcriptomics[16] (Supplementary Table 2). This atlas included four hierarchical levels of classification: 34 classes, 338 subclasses, 1,201 supertypes and 5,322 clusters. Neuronal cell types represented a large portion of the atlas, with 29 classes (85%), 315 subclasses (93%), 1,156 supertypes (96%) and 5,205 clusters (98%). However, that study provided a coarse, class/subclass level overview. Here we conduct a more in-depth analysis and introduce the most complete taxonomy to date of all of the GABAergic neuronal types in the entire telencephalon, at all levels of the hierarchy.

The telencephalic GABAergic taxonomy, defined as the subpallium GABA neighbourhood, includes subclasses 39–90 across seven classes: OB (olfactory bulb)–IMN (immature neurons) GABA, CTX–CGE GABA, CTX–MGE GABA, CNU–MGE GABA, CNU–LGE GABA, LSX (lateral septal complex) GABA and CNU–HYa (anterior hypothalamus) GABA (Fig. 1). These 52 subclasses encompass 284 supertypes and 1,051 clusters, containing 611,423 high-quality single-cell transcriptomes (Supplementary Table 3). We provide several visualizations, including dendrograms (Fig. 1a), uniform manifold approximation and projections (UMAPs; Fig. 1b–d) and constellation diagrams (Extended Data Fig. 1c,d) to depict the multidimensional relationships among different subclasses or supertypes.

Using MERFISH data registered to the Allen Mouse Brain CCFv3 (Fig. 1e), we determined the distribution of GABAergic supertypes across telencephalic regions (Extended Data Fig. 1a). Most supertypes span neighbouring regions, except those in the LSX and the main (MOB) and accessory (AOB) olfactory bulb, which are region specific. Gini and Shannon indices revealed strong spatial localization, with the exception of more widely distributed types in the isocortex (Extended Data Fig. 1a).

As reported previously[16], a vast majority of the clusters are purely GABAergic (Fig. 1a and Supplementary Table 3). However, we also identified glutamate–GABA co-releasing types expressing *Slc17a8* in 11 clusters across several subclasses. We identified 11 cholinergic neuronal clusters in subclass 58 (PAL–STR Gaba–Chol), with complex GABA and/or glutamate co-release patterns (see below), and 2 GABA–cholinergic clusters in subclass 69 (LSX Nkx2-1 Gaba) (Supplementary Table 3). We also identified four GABA–dopamine co-releasing clusters in subclass 44 (OB Dopa–Gaba).

We identified 26 neuropeptide genes that are differentially expressed across telencephalic GABAergic types (Supplementary Table 3), many of which have been previously used as cell-type markers. Most exhibit restricted expression, such as *Edn1*, which is largely confined to subclass 52 (Pvalb Gaba). Others, like *Penk* and *Pnoc*, are more broadly expressed across subclasses within each GABAergic class. These expression patterns further support the transcriptomic identity and potential functional roles of specific cell types.

GABAergic neuronal types in the olfactory bulb

Olfactory sensory neurons transmit odour information to the OB, where neurons with the same odorant receptor converge on specific glomeruli. These neurons synapse with interneurons and excitatory mitral and tufted cells, which then project to various brain regions for further processing. The OB has an exceptionally high ratio of inhibitory to excitatory neurons compared to, for example, the isocortex.

GABAergic neuronal types in the MOB and AOB mostly originate from the LGE and migrate to the OB during development[11]. These neurons continue to be replaced in the adult through neurogenesis in the SVZ and migration through the rostral migratory stream[17]. In the OB–IMN GABA class, we defined six OB GABAergic subclasses (39 to 45; Extended Data Fig. 2).

Subclass 45 (OB–STR–CTX Inh IMN), which includes seven supertypes, contains immature neurons from the SVZ that migrate to the OB (Extended Data Figs. 2a–c,k and 3a,b). Supertypes 166, 168, 171 and 172 are the most immature based on gene expression, spatial location and pseudotemporal ordering[16,18]. The remaining three supertypes appear transitional, resembling mature subclasses (Extended Data Fig. 2a–c). Supertype 170 forms a transition from immature neurons to subclass 39 (OB Meis2 Thsd7b Gaba), which is found mostly in the internal plexiform and mitral (Ipl/Mi), external plexiform (EPl) and glomerular (Gl) layers (Extended Data Fig. 2a,b,d,e). Within subclass 39, supertype 143 represents *Calb1*$^+$ Blanes cells[19] populating the Gl layer (Extended Data Fig. 2e). Supertype 167 forms a transition to subclass 41 (OB–in Frmd7 Gaba), representing granule cells in the granule layer (GrO), Ipl/Mi and partly EPl. Supertype 150 populates the EPl and includes Gl cells modulated by olfactory activity[20] (Extended Data Figs. 2a,b,d,j and 3a). Lastly, immature supertype 169 forms a transition to subclass 42 (OB–out Frmd7 Gaba), which consists of *Calr*$^+$ periglomerular cells (PGCs) populating the Gl layer (Extended Data Fig. 2a,b,d,g).

Subclass 40 (OB Trdn Gaba) contains the population of *Rprm*$^+$ granule cells[20] (Extended Data Fig. 2c,f) and subclass 43 (OB–mi Frmd7 Gaba) represents *Pvalb*$^+$ OB neurons that are postnatally generated neurons populating the EPl[19] (Extended Data Fig. 2c,h). The Gl layer, where sensory input is first processed, contains several interneuron populations that can be broadly divided into three categories, *Calb1*$^+$ PGCs from subclass 39, *Calr*$^+$ PGCs from subclass 42, and the *Slc6a3*$^+$ and *Th*$^+$ dopaminergic PGCs from subclass 44 (Extended Data Figs. 2c,e,g,i and 3a).

GABAergic neuronal types in the cerebral cortex

The CTX (pallium) comprises four major structures: the isocortex, HPF, OLF and CTXsp. GABAergic neurons across these regions fall into two main classes, CTX–CGE GABA (Extended Data Fig. 4) and CTX–MGE GABA (Fig. 2), named after their spatial location and predominant

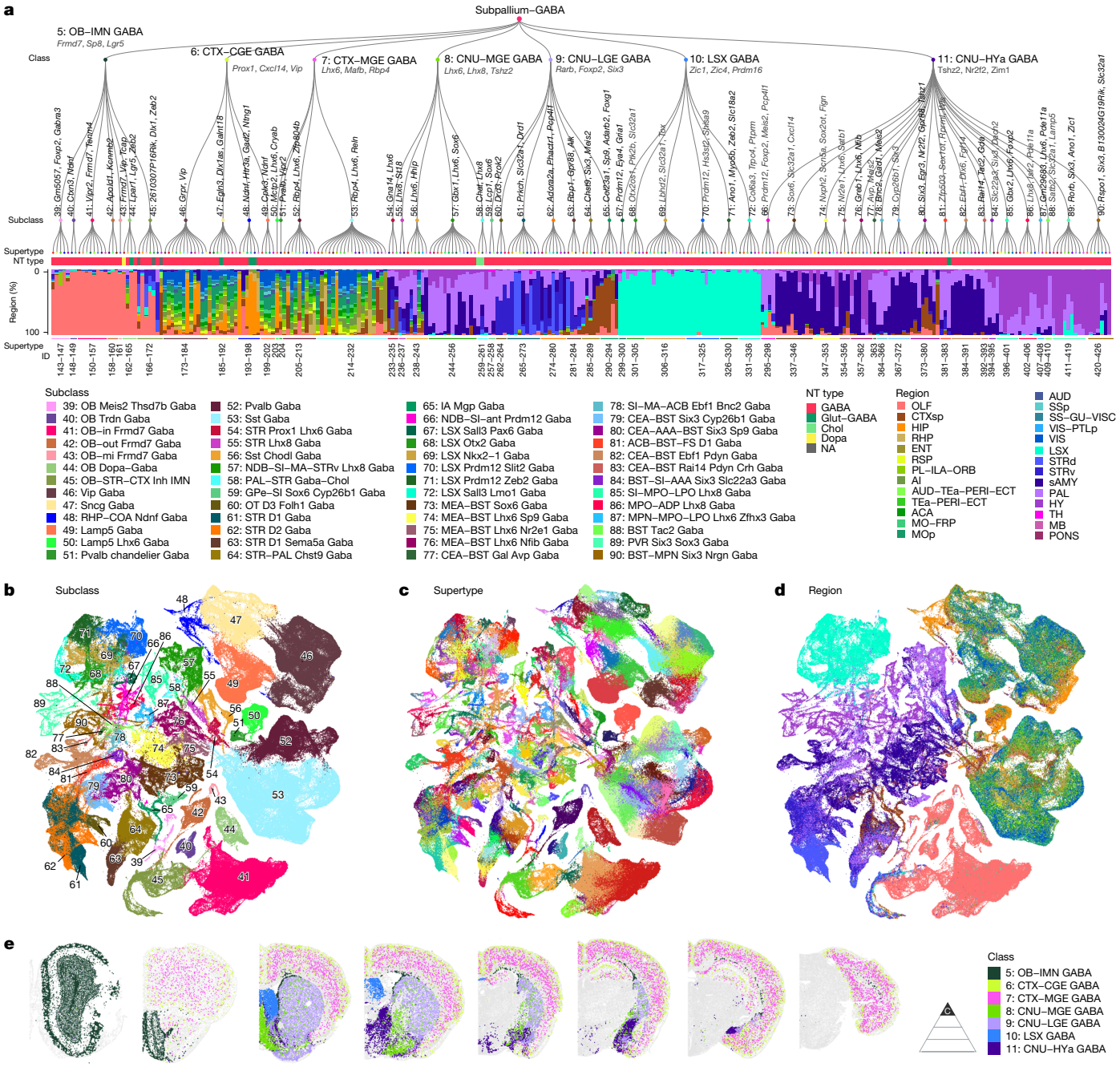


Fig. 1 | Transcriptomic taxonomy of telencephalic GABAergic neuronal types in mice. a, The transcriptomic taxonomy of 284 supertypes organized in a dendrogram (10xv2, $n = 269,307$ cells; 10xv3, $n = 342,116$ cells). At each class or subclass branch, the top marker genes are shown. From the top down, the bar plots represent major neurotransmitter (NT) type and region distribution of profiled cells. NA, not applicable. **b**–**d**, UMAP representation of all cell types coloured by subclass (**b**), supertype (**c**) and the dissection region (**d**). **e**, Representative MERFISH sections of the adult mouse brain across forebrain structures coloured by cell class. Each class is labelled by its ID and shown in the same colour as in the dendrogram in **a**. The triangular schematic denotes the hierarchical level shown in the panel. C, class. All class, subclass and region abbreviations are listed in Supplementary Table 1.

developmental origins. Notably, around 10% of neurons in these classes arise from the embryonic POA[7,14,21]. The current classification aligns with our previous transcriptomic taxonomy of the mouse isocortex and HPF[22], as well as MET-types in Patch-seq studies of the mouse visual cortex[23] (Extended Data Fig. 5), and extends the framework into the OLF and CTXsp areas.

In the CTX–CGE GABA class, we defined four subclasses: three known (46, Vip Gaba; 47, Sncg Gaba; and 49, Lamp5 Gaba)[22], and one newly defined (48, retrohippocampus (RHP)–COA Ndnf Gaba) (Extended Data Fig. 4a–g). The *Vip*$^+$, *Sncg*$^+$ and *Lamp5*$^+$ subclasses correspond to the well-known bipolar, multipolar, neurogliaform and other types of GABAergic interneurons found throughout isocortex and HPF, and also in OLF and CTXsp (Extended Data Fig. 4c–g). Vip clusters in supertypes 179, 181 and 182 are largely hippocampus (HIP) specific (Extended Data Fig. 4b–d). Most GABAergic neurons in the CTX are local interneurons, except for a few long-range projecting (LRP) types[24,25]. The LRP types in the HIP are among the best described cortical LRP neurons. On the basis of the expression of genes, including *Chrna4*, *Pcp4*, *Nos1* and *Htr3a*, supertype 179, which is located in the HIP, was identified as a potential LRP population[26].

The newly defined subclass 48 (RHP–COA Ndnf Gaba) expresses *Ndnf* and *Ntng1*, is enriched in the HPF, OLF and CTXsp, and shows strong

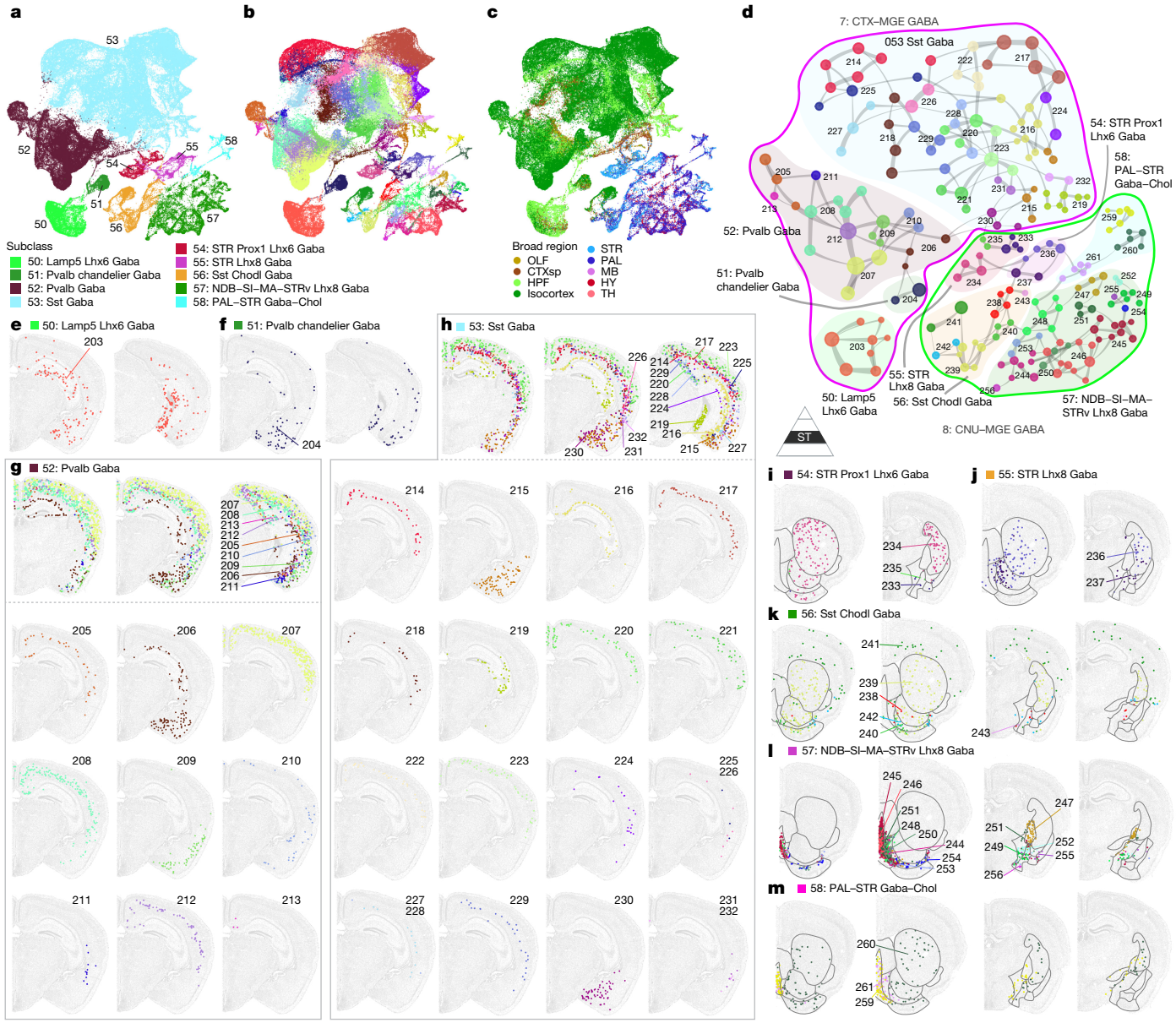


Fig. 2 | MGE-derived GABAergic neuronal types in the CTX and CNU. **a**–**c**, UMAP representation of all MGE cells coloured by subclass (**a**), supertype (**b**) or the broad brain region (**c**). **d**, Constellation plot of MGE clusters using UMAP coordinates shown in **b**. The nodes are clusters coloured by supertype (ST) and grouped in bubbles by subclass. The lines around the subclass bubbles denote the two classes—CTX–MGE GABA and CNU–MGE GABA. **e**–**m**, Representative MERFISH sections showing the location of supertypes in MGE subclasses 50 (Lamp5 Lhx6 Gaba) (**e**), 51 (Pvalb chandelier Gaba) (**f**), 52 (Pvalb Gaba) (**g**), 53 (Sst Gaba) (**b**), 54 (STR Prox1 Lhx6 Gaba) (**i**), 55 (STRv Lhx8 Gaba) (**j**), 56 (Sst Chodl Gaba) (**k**), 57 (NDB–SI–MA–STRv Lhx8 Gaba) (**l**) and 58 PAL–STR Gaba–Chol (**m**). Cells are coloured and labelled by supertype. The triangular schematic denotes the most granular hierarchical level shown in the panels. All class, subclass and region abbreviations are listed in Supplementary Table 1.

regional specificity. Supertypes 195 and 198 localize to the entorhinal cortex (ENT) and RHP, respectively; supertype 194 localizes to the cortical amygdalar area (COA); and the remaining supertypes localize to the HIP (Extended Data Fig. 4c,f). This subclass includes the *Meis2*$^{+}$ GABAergic population in supertype 198, previously described as originating from the embryonic pallial–subpallial boundary and residing in white matter[22,27] (Extended Data Fig. 4a–c). Supertype 198 is the only cortical GABAergic cell type expressing *Meis2* (Extended Data Fig. 4c). Subclass 48 (RHP–COA Ndnf Gaba) lacks *Prox1*, in contrast to other CGE-derived GABAergic neurons, but expresses the CGE markers *Nr2f2* and *Htr3a* and the LGE marker *Sp9*, suggesting an origin at the posterior LGE–anterior CGE boundary.

In the CTX–MGE GABA class, we defined four subclasses: three previously defined ones[22], subclass 53 (Sst Gaba), subclass 52 (Pvalb Gaba) and subclass 51 (Pvalb chandelier Gaba), and a newly reclassified type, subclass 50 (Lamp5 Lhx6 Gaba). The latter was previously part of the CGE-derived *Lamp5*$^{+}$ subclass but is now assigned to the MGE class on the basis of expression of the MGE-specific TF *Lhx6* (Fig. 2a–d). The *Sst*$^{+}$, *Pvalb*$^{+}$ and chandelier subclasses correspond to known interneuron types distributed across the isocortex and HPF, and we now show that they are also present in the OLF and CTXsp regions (Fig. 2f–h and Extended Data Fig. 6a).

Two *Pvalb*$^{+}$ chandelier clusters were identified: one that is broadly distributed, and another that is mostly HIP specific. Among nine *Pvalb*$^{+}$ supertypes, four are isocortex predominant, and two are enriched in the OLF, CTXsp and HPF (Extended Data Fig. 6a). The *Sst*$^{+}$ subclass is highly diverse, with 19 supertypes and 71 clusters showing variable regional preferences. Seven supertypes are enriched in the OLF, CTXsp

and HPF (Extended Data Fig. 6a (red dots)), while others are broadly present in the isocortex. Several *Pvalb*$^+$ and *Sst*$^+$ supertypes (such as, 209, 216, 219, 228 and 232) are HIP specific. (Fig. 2g,h and Extended Data Fig. 6a). The *Lamp5*$^+$*Lhx6*$^+$ subclass is similarly enriched in the HIP with a minor presence in the isocortex (Fig. 2e and Extended Data Fig. 6a).

Sst$^+$ supertypes 215 and 216 probably correspond to LRP neurons expressing *Sst*, *Npy* and *Nos1* (ref. 25), localized to the OLF, CTXsp and HPF, respectively (Fig. 2i). These match LRP3 and LRP4 types described previously[28], which align with CTX–HPF type 87 Sst Etv1 (ref. 22) and show a mixed transcriptional profile characteristic of both interneurons and LRP neurons[28] (Extended Data Fig. 5b).

Overall, CTX–CGE and CTX–MGE GABAergic neuron types in the CTX display clear spatial segregation between the isocortex and the other structures that are considered evolutionarily more ancient, including the HPF, OLF and CTXsp. Some types are situated in transition zones, for example, *Vip*$^+$ supertype 177 is mostly located in the isocortex, is less dense in the retrosplenial cortex (RSP) and is sparse in the HIP (Extended Data Fig. 4d).

MGE-derived GABAergic types in the cerebral nuclei

MGE-derived GABAergic neurons are widely distributed and fall into two main classes: CTX–MGE, located in pallial areas, and CNU–MGE, mainly in the CNU. Both classes express *Lhx6*, a developmental pan-MGE marker. CNU–MGE GABAergic neurons form smaller, more heterogeneous clusters than those in the CTX–MGE class, highlighting a marked difference in complexity between the two (Fig. 2a–d).

STR and PAL in the CNU are key components of the basal ganglia, a group of interconnected subcortical nuclei involved in motor behaviour, learning and cognition[29,30]. These regions are predominantly composed of GABAergic neurons, which fall into four transcriptomic classes (Fig. 1). The CNU–MGE GABA and CNU–LGE GABA classes include striatal and pallidal neurons derived from the MGE and LGE, respectively. The LSX GABA class contains lateral septum neurons derived from the embryonic septum[31]. The CNU–HYa GABA class includes GABAergic neurons located in the STR-like amygdalar nuclei (sAMY), caudal PAL (PALc) and POA of the HYa (Extended Data Fig. 1a); these probably arise from multiple developmental origins, including the LGE, MGE and embryonic POA.

In the CNU–MGE class, we defined five subclasses: 54 (STR Prox1 Lhx6 Gaba), 55 (STR Lhx8 Gaba), 56 (Sst Chodl Gaba), 57 (NDB–SI–MA–STRv Lhx8 Gaba) and 58 (PAL–STR Gaba–Chol) (Fig. 2a,d and Extended Data Fig. 6a). These subclasses span multiple regions, including the dorsal and ventral STR (STRd and STRv) and dorsal, ventral and medial PAL (PALd, PALv, PALm), with no single subclass confined to a specific region (Fig. 2i–m). Although local interneurons make up less than 10% of the striatal neurons, they are molecularly diverse[32]. Subclasses 54 (*Pvalb*$^+$) and 55 (*Pvalb*$^-$), located in the STRd and STRv, probably represent striatal GABAergic interneurons (Fig. 2i,j and Extended Data Figs. 3d and 6a). Subclass 56 (Sst Chodl Gaba) is notable for spanning both pallial and subpallial structures. It includes isocortical *Sst*$^+$*Chodl*$^+$ cells (supertype 241, the cortical *Sst*$^+$*Chodl*$^+$ cells[22,33]), CTXsp (supertype 242) and *Sst*$^+$ interneurons in the STR and PAL (supertypes 238–240 and 243) (Fig. 2d,k and Extended Data Figs. 1a and 6a). These findings indicate that the previously identified cortical *Sst*$^+$*Chodl*$^+$ LRP neurons[23,28,33] are transcriptionally related to striatal *Sst*$^+$ interneurons, all classified within subclass 56.

Subclass 58 (PAL–STR Gaba–Chol) includes basal forebrain cholinergic neurons, divided into three supertypes (Fig. 2a,d,m and Extended Data Figs. 6a and 7). Supertypes 259 and 260 contain cholinergic neurons, whereas supertype 261 includes closely related GABAergic neurons primarily from the medial septum (MS) and diagonal band nucleus (NDB). Supertype 259 represents cholinergic projection neurons that co-release GABA and, in some cases, glutamate (Supplementary Table 3), and show spatially specific clustering: MS and NDB (clusters 923–925), substantia innominata (SI) and NDB (clusters 926 and 928) or GPe and SI (cluster 927) (Extended Data Fig. 7b). Supertype 260 includes striatal cholinergic interneurons located in the CP, ACB and olfactory tubercle (OT), which may co-release glutamate through *Slc17a8* but do not co-release GABA (Extended Data Fig. 7c,e and Supplementary Table 3). Moreover, supertype 307 of subclass 69 (LSX Nkx2-1 Gaba) contains clusters 1081 and 1084, which are both GABAergic and cholinergic and are located in the lateral septum (Extended Data Fig. 7a,d,e).

LGE-derived GABAergic types in the cerebral nuclei

The CNU–LGE GABA class consists of seven LGE-derived subclasses (59–65) (Fig. 3 and Extended Data Fig. 6b). These subclasses resemble the well-known D1- and D2-type medium spiny neurons (MSNs; also known as spiny projection neurons) in the STR, with subclasses 61 (STR D1 Gaba) and 62 (STR D2 Gaba) being the prototypic striatal D1 and D2 MSNs and the other subclasses being newly defined homologous cell types. Among them, subclasses 60, 63 and 64 also express the dopamine receptor gene *Drd1*, and subclass 60 additionally expresses *Drd3* strongly.

Subclass 59 (GPe–SI Sox6 Cyp26b1 Gaba) is predominantly localized to the GPe, with some cells observed in the SI (Fig. 3c). The GPe—a part of the basal ganglia involved in action control, decision-making and reward—contains two main types of GABAergic neurons: prototypical and arkypallidal[34]. The prototypical neurons mostly express the TF gene *Nkx2-1*, exhibit a fast firing rate and project to the subthalamic nucleus. Supertype 247 of the MGE-derived subclass 57 (NDB–SI–MA–STRv Lhx8 Gaba) is located in the GPe and corresponds to these *Nkx2-1*$^+$ and *Lhx6*$^+$ neurons (Fig. 2l and Extended Data Fig. 6a). By contrast, arkypallidal neurons, characterized by the expression of *Penk* and *Foxp2*, have a low firing rate and project back to the STR[34,35]. Supertype 257 of the LGE-derived subclass 59 (GPe–SI Sox6 Cyp26b1 Gaba) probably represents these *Penk*$^+$ and *Foxp2*$^+$ GPe arkypallidal neurons (Fig. 3c and Extended Data Fig. 6b).

Subclass 60 (OT D3 Folh1 Gaba), which strongly expresses *Drd3* alongside *Drd1*, is specifically localized in the Islands of Calleja in the OT (Fig. 3f and Extended Data Fig. 3c,d). Subclasses 63 (STR D1 Sema5a Gaba), 64 (STR–PAL Chst9 Gaba) and 65 (IA Mgp Gaba) collectively form a new group of GABAergic neuronal types that resemble MSNs based on their gene expression profiles but are distinct based on their spatial locations (Fig. 3d,e,g). The spatial distribution of many supertypes and clusters in these subclasses is unique, scattered along the borders of different striatal and pallidal regions, forming distinct patches or streaks.

Supertype 281 within subclass 63 (STR D1 Sema5a Gaba) is found in the STRd and contains hybrid MSNs that show robust expression of *Drd1* and modest expression of *Drd2* (Fig. 3d and Extended Data Fig. 6b). Within cluster 991 of supertype 281, 32% of cells express both *Drd1* and *Drd2* above a threshold of $\log_2$[counts per million (CPM)] > 3. However, the gene expression signature of these *Drd1* and *Drd2* co-expressing cells is insufficient to define a separate cluster. These neurons may be those described before as co-expressing *Drd1a* and a shortened variant of the D2 receptor[36]. These neurons resemble the exo-patch MSNs, which are located in the striatal matrix but physiologically resemble the MSNs from patches[32,36,37] (Extended Data Fig. 3c,d). We also identified a small proportion of *Drd1*/*Drd2*-co-expressing neurons in subclasses 61 (8.3%) and 62 (8.8%).

Spatial patterns suggest that subclasses 64 (STR–PAL Chst9 Gaba) and 65 (IA Mgp Gaba) reside in the interstitial nucleus of the posterior limb of the anterior commissure (described in the Paxinos's atlas[38]) or intercalated amygdalar nucleus (IA) (Fig. 3e,g). Subclass 65 includes five supertypes, all expressing *Foxp2* and *Tshz1*, markers of *Sp8*$^+$ dLGE-derived intercalated cells (ITCs)[39,40] (Extended Data Fig. 6b). Subclass 65 (IA Mgp Gaba) shows similarity to subclass 39 (OB Meis2 Thsd7b Gaba) (Fig. 1b,c,e,f), hinting at shared developmental origins. Within subclass 65, supertype 293 stands out. On the basis of

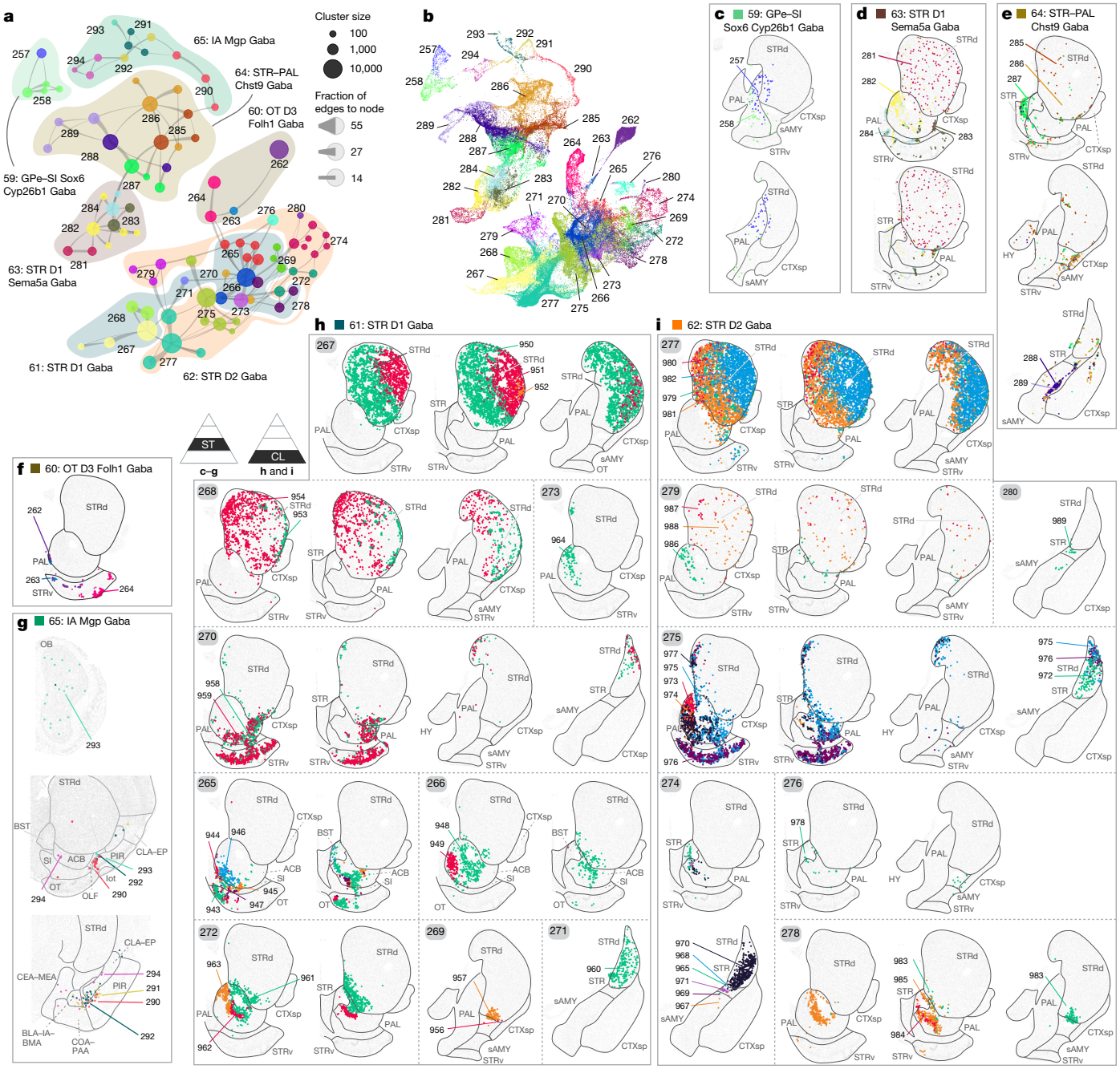


Fig. 3 | LGE-derived GABAergic neuronal types of the CNU. a, Constellation plot of LGE clusters using UMAP coordinates shown in **b**. Nodes are labelled by cluster ID, coloured by supertype and grouped in bubbles by subclass. **b**, UMAP representation of all LGE cells coloured by supertype. **c**–**g**, Representative MERFISH sections showing the location of LGE subclasses 59 (GPe Sox6 Cyp26b1 Gaba) (**c**), 63 (STR D1 Sema5a Gaba) (**d**), 64 (STR–PAL Chst9 Gaba) (**e**), 60 (OT D3 Folh1 Gaba) (**f**) and 65 (IA Mgp Gaba) (**g**). Cells are coloured and labelled by supertype. **h**,**i**, Representative MERFISH sections showing the location of CNU–LGE subclasses 61 (STR D1 Gaba) (**h**) and 62 (STR D1 Gaba) (**i**). Each box contains one supertype and cells are labelled and coloured by cluster to highlight the diversity. The triangular schematics denote the most granular hierarchical level shown in **c**–**g** (supertype) and **h** and **i** (cluster (CL)). BLA, basolateral amygdalar nucleus; BMA, basomedial amygdalar nucleus; CLA, claustrum. All class, subclass and region abbreviations are listed in Supplementary Table 1.

the dissection information, these cells were isolated from the OLF (48%) and CTXsp (36%) (Extended Data Fig. 6b). MERFISH data show that cells in supertype 293 are located in the MOB, anterior olfactory nucleus and, sparsely, in the IA (Fig. 3g). A previous STICR lineage-tracing study showed that the ITCs of the amygdala are clonally related to OB interneurons and MSNs in the STR[41].

We defined 9 and 7 supertypes in subclasses 61 (STR D1 Gaba) and 62 (STR D2 Gaba), respectively. These supertypes and their clusters show diverse spatial patterns across the STRd and STRv (Fig. 3h,i). Some span broad areas while others are confined to specific medial–lateral, anterior–posterior or dorsal–ventral subdomains. Paired D1 and D2 clusters often co-localize. Clusters D1 951 and D2 982 are located in the lateral CP, while clusters D1 950 and D2 981 are in the medial CP. Clusters D1 960 and D2 970 are located in the posterior tail of the CP, while clusters D1 957 and D2 983 are in the posterior ventral tip of the CP. Within subclass 61 (STR D1 Gaba), supertype 268 is enriched in the

striatal patches and striosomes and supertype 267 is enriched in the matrix (Fig. 3h). Similarly, in subclass 62, supertype 279 is striosome enriched and supertype 277 is matrix enriched (Fig. 3i).

We observed complex subregional enrichment of cell types in the ACB. D1 supertypes 265, 266, 270, 272 and 273, and D2 supertypes 275 and 278 are mostly restricted to the ACB and OT (Fig. 3h,i). The ACB can be divided into core and shell subregions. D1 supertypes 272 and 273 and D2 supertype 278 are predominantly located in the core. D1 supertype 266 has two clusters: 948 in the core and 949 in the shell (Fig. 3h). D2 supertype 275 includes clusters 973 and 974 in the core and the others in the shell or OT (Fig. 3i). The ACB core and shell can be further subdivided along mediolateral and anteroposterior axes. In the D1 272 and D2 278 supertypes, clusters D1 963 and D2 985 are located in the medial–anterior subdomain of the ACB core, while clusters D1 961 and D2 983 are in a more lateral–posterior position (Fig. 3h,i). In the ACB shell, D1 supertypes 265 and 270 occupy medial–anterior and lateral–posterior subdomains, respectively, and, within the D2 275 supertype, clusters 977 and 975 show a similar pattern. Despite these enrichments, many cell types overlap across ACB subregions, as previously reported[32].

GABAergic neuronal types in the lateral septum

The LSX GABA class is uniquely localized to the LSX and is distinct from other CNU GABAergic types. It contains six complex and intertwined subclasses (67–72) with partially overlapping spatial patterns (Extended Data Fig. 8a–j). For example, subclasses 67 and 68 are found in both the rostral and ventral LSX (LSr and LSv), with subclass 68 also extending dorsally (Extended Data Fig. 8d–f). Similarly, subclasses 69 and 70 overlap in the LSr and LSv, with subclass 70 extending further into the anterior and posterior LSX (Extended Data Fig. 8d, g, h).

GABAergic neuronal types in the sAMY and POA

The CNU–HYa GABA class is highly complex, with 19 subclasses that are predominantly localized in the CNU but also extending into the POA of the HYa (subclasses 66 and 73–90; Extended Data Fig. 9). Neurons in this class are located in specific sAMY and PAL regions, including the central amygdalar nucleus (CEA), medial amygdalar nucleus (MEA), anterior amygdalar area (AAA), BST, SI and magnocellular nucleus (MA) (Fig. 4a and Extended Data Fig. 9d). Each subclass within the CNU–HYa GABA class is not specific to a single region but contains neurons from multiple regions, for example, subclass 80 contains neurons from the CEA, AAA and BST (Fig. 4a,c and Extended Data Fig. 9d). Conversely, each region contains multiple subclasses, for example, subclasses 74, 79, 80 and 82 are co-localized in the CEA, AAA and BST, while subclasses 78 and 84 are co-localized in the SI, NDB and MA.

BST functions as a hub for processing limbic information and monitoring emotional valence and is at the centre of a vast connectivity network. It regulates mood and arousal through sAMY, dorsal raphe and VTA connections, and monitors feeding, drinking and reproductive behaviours through brainstem and hypothalamic inputs, as well as LSX and MEA projections. These diverse functions correspond to distinct cell types. As part of the extended amygdala, BST is a key output region for CEA and MEA neurons[42,43]. Transcriptomic analysis reveals strong similarity between the BST, CEA and MEA, with each region marked by distinct gene signatures (Extended Data Fig. 9d).

Types within subclasses 73, 74, 75, 76 and 88 are mostly located in the MEA and/or BST (Figs. 1a,e and 4b and Extended Data Fig. 9d), probably sharing developmental origins based on conserved TF expression (see below). These MEA–BST types are transcriptomically similar to subclasses 85–87, 89 and 90 in the POA, which are linked by shared circuits and functions[44–46] (Fig. 1e and Extended Data Fig. 9d). The MEA integrates olfactory and vomeronasal inputs and has a central role in regulating social and reproductive behaviours[47]. For example, subclasses 74, 75 and 76 express the posterodorsal MEA marker *Lhx6* (ref. 48,49) and neurons in this region are involved in reproductive behaviours[47,49]. Supertypes 349 and 351 (subclass 74) express *Crhr2* and *Ucn3* (Extended Data Fig. 9d), markers of stress-responsive neurons[50]. Supertypes 357–359 (subclass 76) include the BNSTpr$^{Tac1/Esr1}$ type, which is known to regulate male social behaviour[46] (Extended Data Fig. 3e).

Subclasses 77–84 are primarily located in the CEA, BST, AAA and SI (Extended Data Fig. 9d), show similarity to striatal CNU–LGE subclasses (Figs. 1e and 4a,c) and may share developmental origins based on conserved TF expression. The CEA is a STR-like GABAergic structure containing both interneurons and long-range projection neurons and drives fear responses through projections to the hypothalamus and brainstem[51]. Most supertypes in subclass 79 (CEA–AAA–BST Six3 Cyp26b1 Gaba) are located in the lateral and capsular parts of CEA (CEAl and CEAc), express the known MSN markers *Penk*, *Pax6*, *Gpr88* and *Ppp1r1b*[36,37,48], and are related to subclass 62 (STR D2 Gaba) (Figs. 1e,f and 4c and Extended Data Fig. 9d). Supertype 371 in subclass 79, located mostly in the CEAc, is transcriptomically and spatially similar to D2 supertypes 274 and 280 (Figs. 1e,f, 3i and 4c). Supertype 368 corresponds to the previously described CeA Prkcd-Ezr type that is highly responsive to cued fear conditioning[48] (Extended Data Fig. 3f). Subclasses 77 (CEA–BST Gal Avp Gaba), 82 (CEA–BST Ebf1 Pdyn Gaba) and 83 (CEA–BST Rai14 Pdyn Crh Gaba) are mainly located in the medial part of CEA (CEAm; Fig. 4c).

We extracted all BST neurons from the MERFISH data and performed spatial clustering based on gene expression (500-gene panel) and spatial proximity (Methods). This revealed distinct domains, several of which align with known BST subdivisions and with MEA–BST and CEA–BST subclasses (Fig. 4d). BST subdivisions have been proposed along mediolateral, anteroposterior and dorsoventral axes based on cytoarchitecture, developmental origin and monoaminergic innervation[42,52]. MEA primarily projects to posteromedial BST[42], where MEA–BST subclasses (73–76) are enriched, while CEA projects to the anterolateral BST[53], where CEA–BST subclasses (77 and 79–83) are located (Fig. 4e,f). Thus, neurons of the same supertype may migrate to distant yet interconnected regions such as the CEA and anterolateral BST or MEA and posteromedial BST.

Gene expression gradients in telencephalic regions

MGE cortical gradient

During adulthood, CGE-derived GABAergic neurons in subclasses 46 (Vip Gaba) and 47 (Sncg Gaba) are evenly distributed in deep and superficial cortical layers and show minimal spatial gene expression gradients (Extended Data Figs. 1b and 4d,e). By contrast, MGE-derived *Pvalb*$^+$ and *Sst*$^+$ neurons exhibit distinct laminar distributions. The birthdate of neurons has been linked to laminar patterning whereby MGE-derived GABAergic neurons are generated in an inside-out pattern similar to their excitatory counterparts[54,55]. However, this is only part of a more complex inside-outside-in pattern formation resulting in the distinct laminar distribution seen in *Pvalb*$^+$ and *Sst*$^+$ neurons[56]. *Sst*$^+$ GABAergic neurons are among the first types to diversify and mature and are more abundant in infragranular than in supragranular layers of the isocortex, while *Pvalb*$^+$ GABAergic neurons can be found throughout all layers except for layer 1 (Fig. 2g,h and Extended Data Fig. 10a–c).

We performed independent component analysis (ICA) separately on subclasses 52 (Pvalb Gaba) and 53 (Sst Gaba), projecting scRNA-seq data onto MERFISH data and identifying top loading genes driving the strongest spatial gradients (Methods). We found 20 genes driving a spatial gradient along the cortical depth in the *Pvalb*$^+$ subclass and 45 such genes in the *Sst*$^+$ subclass (Extended Data Fig. 10d–g). Among these genes, only six genes are shared between the subclasses (*Gm32647*, *Il1rapl2*, *Col25a1*, *Cnr1*, *Nkain3* and *Parm1*), indicating that the observed gradient is unique for each subclass. And only three genes (*Rbp4*, *Pdyn* and *Ndst3*) are subclass or supertype markers, suggesting that the

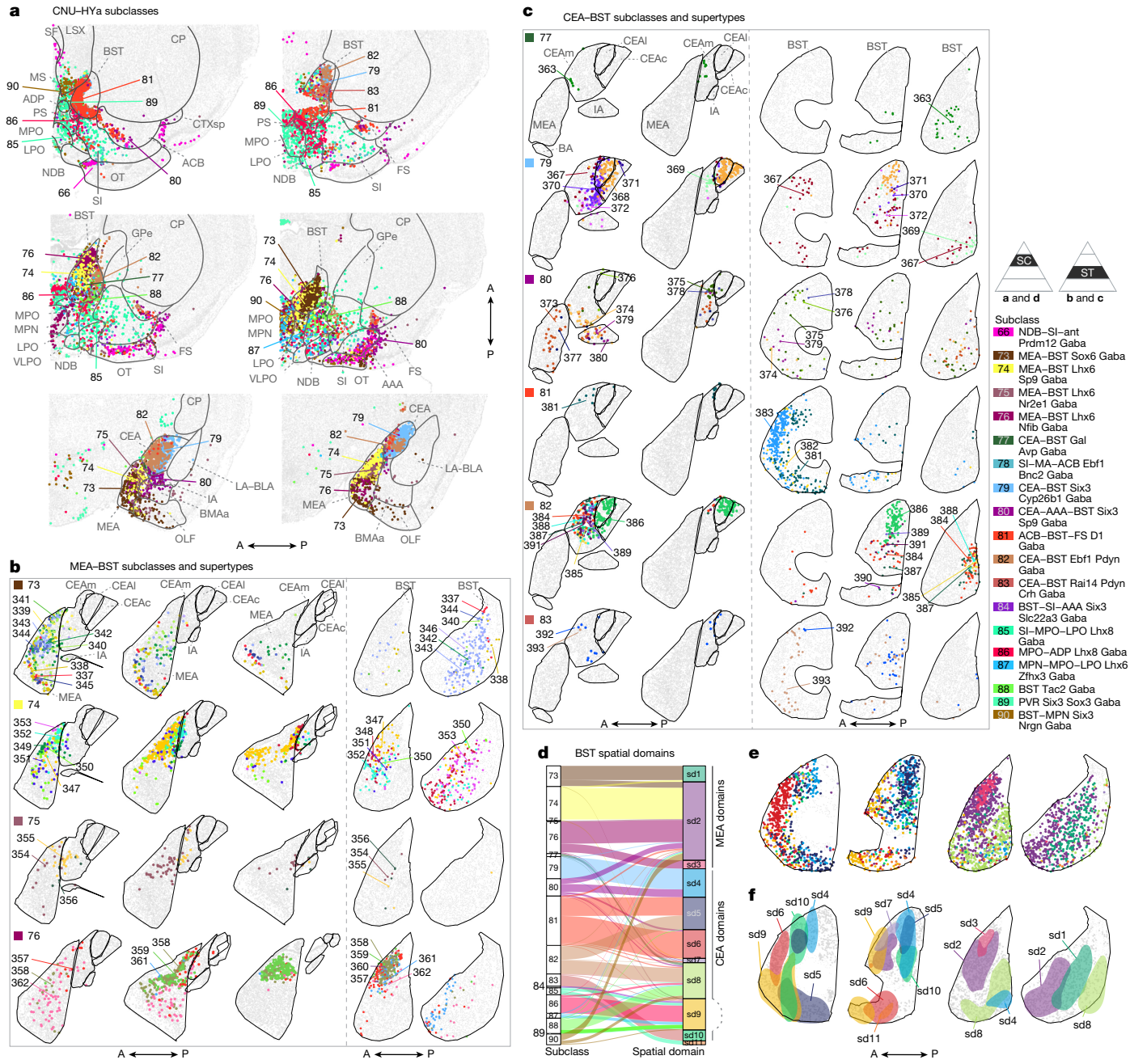


Fig. 4 | Organization of GABAergic neuronal types across the sAMY and BST. a, Representative MERFISH sections showing the locations of subclasses belonging to the CNU–HYa GABA class. Cells are coloured and labelled by subclass (SC). A, anterior; P, posterior. **b**,**c**, Representative MERFISH sections showing the locations of cells belonging to the MEA–BST subclasses (**b**) and CEA–BST subclasses (**c**). Each row shows one subclass, and cells are coloured and labelled by supertype. Within each subclass, each supertype's locations in both the MEA and BST (**b**) or both the CEA and BST (**c**) are shown, indicating the existence of the same cell type in these locations. **d**–**f**, Spatial domain clustering within BST neurons using BANKSY. The alluvial plot (**d**) shows the relationship between subclasses present in the BST and the spatial domains. Spatial domains corresponding to subclasses that are shared with MEA or CEA are labelled as MEA domains or CEA domains, respectively. Representative MERFISH images show the location of the spatial domains, with cells coloured by spatial domain identity (**e**) and the area covered by each domain (**f**). Triangular schematics denote the most granular hierarchical level shown in the panels. All class, subclass and region abbreviations are listed in Supplementary Table 1.

gradients are largely independent of supertype diversity within each subclass.

Homology and gradients in MSN populations

GABAergic neuron subclasses 61 (STR D1) and 62 (STR D2) exhibit highly similar spatial patterns (Fig. 3h,i). The constellation plot shows pairs of D1 and D2 clusters with the highest level of transcriptome similarity (Methods and Extended Data Fig. 11a). MERFISH analysis validated co-localized spatial distributions for five closely related D1–D2 cluster pairs (Extended Data Fig. 11b–f). ICA on D1 and D2 subclasses, combined with scRNA-seq projection onto MERFISH data, identified two gene modules forming a dorsolateral-to-ventromedial gradient (Methods and Extended Data Fig. 11g,h), with gene module scores shown on the scRNA-seq UMAP and representative MERFISH sections (Extended Data Fig. 11i–p). Dorsolateral D1/D2 types are marked by expression of genes encoding various neuroactive receptors, such as *Grm8*, *Htr2a*

and *Cnr*, while ventromedial types express genes encoding proteins involved in cGMP–PKG signalling, including *Csgalnact1*, *Prkg1* and *Slc8a1* (Extended Data Fig. 11g,h).

Gradients in lateral septum

The LSX—a basal forebrain structure—integrates cortical and subcortical inputs and relays signals to hypothalamic and midbrain nuclei, playing a key part in social behaviours such as anxiety and aggression. Most neurons in the LSX are GABAergic and contain receptors for a variety of neuromodulators and neuropeptides (Supplementary Table 3). LSX subclasses exhibit overlapping spatial distributions, and transcriptomic subtypes are not confined to single LSX segments. Using scRNA-seq imputed into MERFISH space, we identified five major spatial gradients (Methods and Extended Data Fig. 12). Gene modules driving these gradients often cross subclass and supertype boundaries, with only two subclass markers, *Zeb2* and *Six3*, and just one supertype marker, *Foxp2* (Extended Data Fig. 12a). The strongest spatial gradients in the LSX represent dorsoventral or mediolateral gradients but no strong anteroposterior gradient exists (Extended Data Fig. 12c). Although LSX input domains poorly match its molecular structure, molecular patterns align with projection targets[57,58]. For example, module 3 includes *Foxp2* and *Ndst4*, genes that are enriched in a subregion projecting to the medial and lateral POA (MPO and LPO)[57,58]. These findings suggest that LSX cell types are arranged along multidimensional gradients that may reflect both input and output pathways.

Persistent developmental signatures

The telencephalic GABAergic neurons arise mostly from the five principal progenitor domains of the subpallium and from there migrate to populate various regions of the telencephalon. We identified a comprehensive set of TF marker genes defining the adult neuronal types at the class and subclass levels (Fig. 5a).

To trace the developmental origins of telencephalic GABAergic cell types, we generated scRNA-seq data across nine stages (embryonic day 11.5 (E11.5) to postnatal day 14 (P14)) and incorporated eight external datasets spanning E7–P2 (refs. 31,59–63) (Supplementary Tables 4 and 5). After stringent quality control and selection of subpallium GABA neurons, we compiled a developmental dataset of 614,569 cells, 387,148 in house (Allen Institute for Brain Science, AIBS) and 227,421 external sources (Supplementary Tables 4 and 5 and Extended Data Fig. 13). This dataset was integrated with adult P56 telencephalic GABAergic neurons (10xv3 only) to visualize gene expression dynamics across time (Fig. 5b). We inferred progenitor origins for adult classes and subclasses (Fig. 5c,d) by label transfer to developmental cells and subclass-specific de novo clustering (Methods). The label accuracy was validated against the original dataset annotations (Extended Data Fig. 14).

We established a developmental telencephalic GABAergic neuronal type taxonomy (E7–P14) with 1,688 clusters within 268 supertypes, 62 subclasses and 13 classes (Supplementary Table 6). Of these, 1,525 clusters, containing all postnatal developmental cells, were assigned with the adult cell-type identities through label transfer, recovering 242 out of the 284 adult supertypes and all adult subclasses and classes. The remaining 163 clusters were assigned developmental identities: 6 classes (51, radial glial cell (RGC); 52, neuroblast; 53, CGE immature; 54, MGE immature; 55, LGE immature; 56, CNU–HYa immature), 10 subclasses and 50 supertypes. The RGC class, probably representing SVZ radial glial cells, is closely related to OB–IMN GABA and comprises mainly early postnatal cells. The neuroblast class includes early embryonic progenitors across domains and may give rise to multiple neuronal lineages (Fig. 5b,c).

GABAergic neurons from the five progenitor domains are identifiable early by distinct gene expression patterns (Fig. 5e and Extended Data Figs. 15a and 16), with spatial TF specificity verified using the developing mouse ISH atlas[64] and the P56 MERFISH dataset (Extended Data Fig. 15b).

There is considerable gene expression heterogeneity in the newly defined immature classes. Both classes 54 (MGE immature) and 55 (LGE immature) contain clusters that are actively going through cell cycle as shown by expression of *Top2a* and *Cdk1* (Extended Data Fig. 16a,b,d). Within each of these classes, genes marking specific progenitors are expressed as well as genes that are expressed in RGCs and/or intermediate progenitor cells (IPCs). *Nkx2-1*, *Dlx1* and *Ascl1* are expressed in MGE RGCs and IPCs regardless of the cell cycle phase (Extended Data Fig. 16a,d). Within the MGE immature class, several clusters express *Zic1*, *Zic4* and low levels of *Pax6*, potentially marking progenitors from rostral MGE or septum[7,65]. Immature LGE cells express *Dbi*, *Fabp7*, *Gsx2* and *Ascl1*, which are known to be expressed in LGE RGCs and IPCs[66] (Extended Data Fig. 16b,d). As the cells start differentiating into MSN D1 or D2, *Isl1* and *Ebf1* are upregulated in D1, while *Six3* and *Tshz1* mark D2. In the 56 CNU–HYa immature class, we could not resolve developmental stages; instead, this class, like its adult counterpart, includes clusters from diverse developmental domains, marked by *Nkx2-1* and *Zic1* (MGE/septum), *Meis2* and *Pax6* (LGE), and *Otp* (hypothalamic anlage)[67,68] (Extended Data Fig. 16c).

Developmental CGE and MGE

The cortical CGE- and MGE-derived GABAergic classes are predominantly located in all regions of the isocortex, OLF, HPF and CTXsp (Extended Data Fig. 1a), and are marked by expression of the developmental TF *Maf* (Fig. 5a and Extended Data Fig. 15a). The CTX–CGE GABA class specifically expresses the developmental TFs *Nr2f2* and *Prox1*. MGE gives rise to GABAergic neurons that populate both CTX (CTX–MGE GABA class) and STR and PAL (CNU–MGE GABA class). Both classes are marked by expression of *Lhx6* and the CNU–MGE GABA class also specifically expresses *Lhx8* (Fig. 5a,e and Extended Data Fig. 15b).

For both CTX–CGE and CTX–MGE neurons, expansion of distinct types occurs between P0 and P14, whereas CNU–MGE types begin diversifying earlier, during late embryonic stages (Fig. 6a–c and Extended Data Fig. 17a–d). Before P0, we identified 6 out of 8 subclasses and 8 out of 60 supertypes in the CTX–CGE GABA and CTX–MGE GABA classes, increasing to all 8 subclasses and 53 supertypes by P14. In the CNU–MGE class, all 5 subclasses and 14 out of 27 supertypes were detected before P0, rising to 23 by P14. Among cortical types, we could verify that *Sst*$^+$ GABAergic neurons emerge and diversify earlier than *Pvalb*$^+$ neurons[13]. *Pvalb*$^+$ chandelier cell signatures[69] appear around E14 but do not form distinct clusters until P0 (Extended Data Fig. 17e). Early postnatal circuit activity probably influences subtype specification[70,71].

In corroboration with previous results, subclass 50 (Lamp5 Lhx6 Gaba) expresses both *Lhx6* and *Adarb2*, markers of the MGE and CGE (Extended Data Fig. 17f). As stated above, in the current whole-mouse brain (WMB) taxonomy, this subclass is assigned to the CTX–MGE class due to expression of the MGE marker *Lhx6*. *Lhx6* and *Adarb2*, marking MGE- and CGE-originated inhibitory neurons, respectively, are co-enriched in the POA-derived neurogliaform cells, indicating that the *Lamp5*$^+$*Lhx6*$^+$ subclass might be POA derived.

In the adult data, we described the existence of a laminar gradient among the *Pvalb*$^+$ and *Sst*$^+$ types (Extended Data Fig. 10). This laminar gradient can also be seen in the *Pvalb*$^+$ and *Sst*$^+$ types during postnatal development starting as early as P0 (Extended Data Fig. 18a,b). This might indicate that the gradient is established very early on as the new neurons enter the isocortex or it is preprogrammed into newborn neurons.

Developmental LGE

The OB–IMN GABA and CNU–LGE GABA classes both arise from the LGE domain (Figs. 5b–e and 6d,e and Extended Data Fig. 19a,b). OB–IMN GABA types are found in the MOB, AOB and the SVZ lining the lateral ventricle, and express *Sp8*, *Sp9* and *Meis2* (Fig. 5a,e and Extended Data Fig. 15). The types within the CNU–LGE GABA class are predominantly located in the STRd and STRv and are marked by expression of *Rarb* and

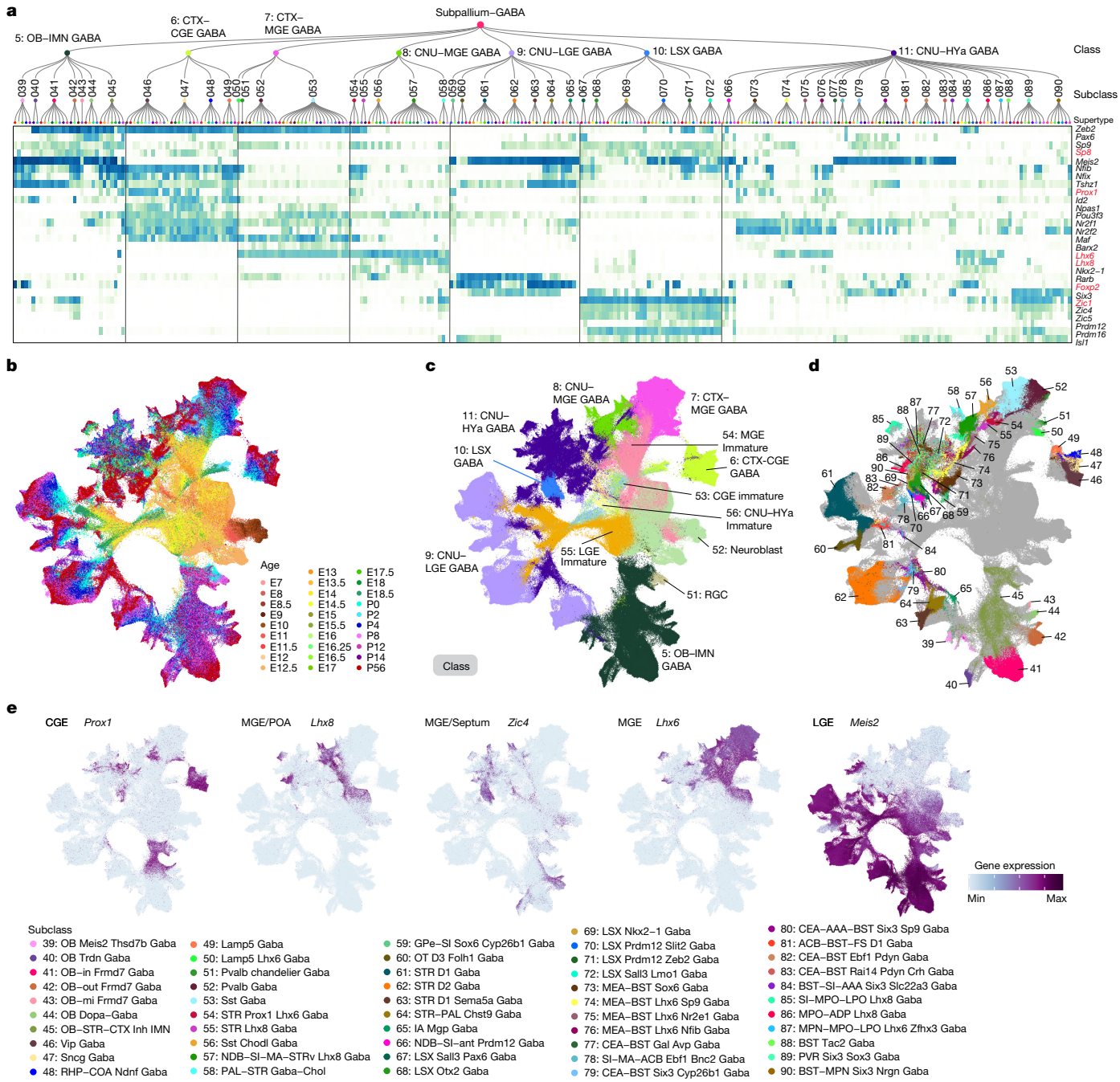


Fig. 5 | TFs and developmental trajectories of telencephalic GABAergic neurons. a, Expression of key TFs in each supertype in the adult subpallium GABAergic neuron taxonomy tree. Genes labelled in red have exemplary developmental in situ hybridization (ISH) images shown in Extended Data Fig. 15b. **b**–**e**, UMAP representation of all cell types across the developmental time course from E7 to P56, coloured by age (**b**), class (**c**), adult subclass (**d**) and major developmental lineage gene markers (**e**). In this global UMAP, only the 10xv3 cells are included for the P56 timepoint. All class, subclass and region abbreviations are listed in Supplementary Table 1.

Foxp2 (Fig. 5a and Extended Data Fig. 15). Before P0, we identified 6 out of 14 subclasses and 12 out of 65 supertypes in the OB–IMN GABA and CNU–LGE GABA classes. At P14, we identified all 14 subclasses and 64 out of 65 supertypes within these classes.

In the adult GABAergic cell-type taxonomy, we observed a relatedness between subclass 39 (OB Meis2 Thsd7b Gaba) and subclass 65 (IA Mgp Gaba) from CNU that would suggest that these subclasses share their developmental origin. Subclass 39, containing the neurons that populate the Ipl/Mi, EPl and Gl layers of the OB, is sequestered away from most other OB cell types and more closely related to subclass 65 containing immature cells populating IA (Fig. 6d,e).

In the developmental taxonomy, we observed that supertypes 265 and 270 of subclass 61 (STR D1 Gaba) are closely related to subclass 60 (OT D3 Folh1 Gaba) (Fig. 6d,e). Notably, these supertypes share a physical location with the OT neurons in subclass 60 in the adult (Fig. 3f,h). This might indicate that the neurons in these locations share a specific developmental origin.

Although subclasses 61 (STR D1 Gaba) and 62 (STR D2 Gaba) are transcriptomically similar in adulthood, their distinction is established early. A shared immature subclass (507 STR immature) emerges at E12.5, followed by distinct D1 (508) and D2 (509) subclasses at E13.5 and E14.5, respectively (Fig. 6d–g and Extended Data Fig. 19a–e). D1 and D2

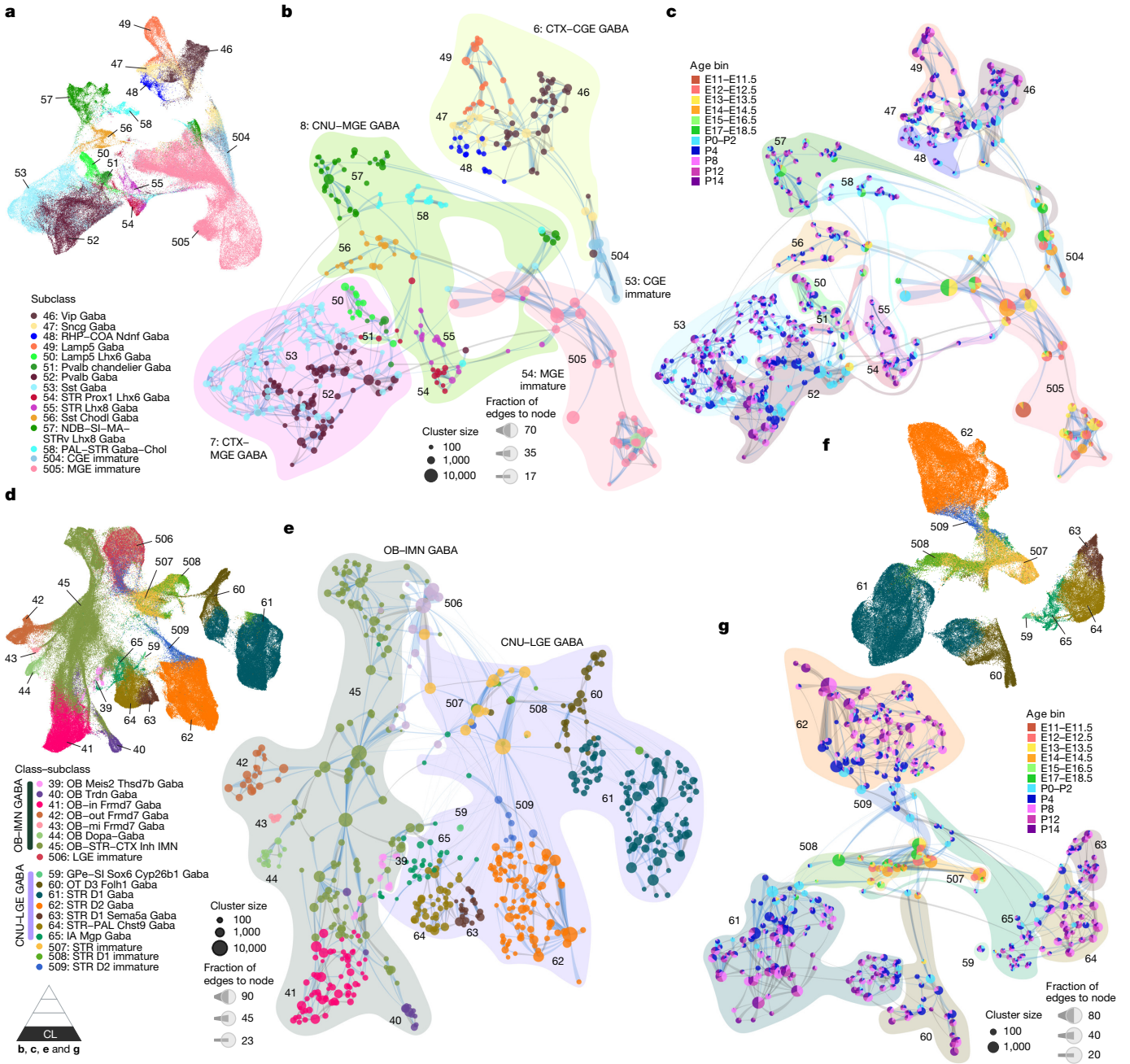


Fig. 6 | Developmental trajectories of CGE-, MGE- and LGE-derived neurons. **a**, UMAP representation of all neurons from embryonic ages to P14 that will form the CTX–CGE GABA, CTX–MGE GABA and CNU–MGE GABA classes. Cells are coloured and labelled by subclass. **b**,**c**, Constellation plots showing all clusters using UMAP coordinates from **a**. Nodes are clusters that are coloured and labelled by subclass (**b**) or the proportion of age bins (**c**). The bubbles represent classes (**b**) and subclasses (**c**). **d**, UMAP representation of all neurons from embryonic ages to P14 that will form the OB–IMN GABA and CNU–LGE GABA classes. Cells are coloured and labelled by subclass. **e**, Constellation plot showing all clusters using UMAP coordinates from **d**. Nodes are coloured and labelled by subclass, and bubbles behind constellation are coloured by class. **f**, UMAP representation of all neurons from embryonic ages to P14 that will form the CNU–LGE GABA class. Cells are coloured and labelled by subclass. **g**, Constellation plot showing all clusters using UMAP coordinates from **f**. Nodes are clusters that are coloured by the proportion of age bins. The bubbles represent subclasses. The triangular schematic denotes the most granular hierarchical level shown in the panels (supertype). All class, subclass and region abbreviations are listed in Supplementary Table 1.

populations originate from different progenitor populations based on expression of distinct gene sets, for example, *Isl1* and *Ebf1* for D1, and *Sp9* and *Sox2* for D2 (Extended Data Fig. 19e–k). Initially distinct, their gene expression profiles converge postnatally, with formerly specific markers shared by P14 and P56 (Extended Data Fig. 19e,l–q). Transcriptomic similarity increases over time, as shown by reduced distance and emergence of a shared gene expression gradient in adulthood (Extended Data Figs. 11g,h, 18c,d and 19r). This suggests that D1 and D2 types follow separate developmental trajectories and later converge as they integrate into shared circuits.

Developmental CNU–HYa and LSX

The LSX GABA class, restricted to LSX, expresses TFs *Zic1*, *Zic4*, *Zic5* and *Prdm16* (Fig. 5a,e and Extended Data Figs. 15 and 20f). The CNU–HYa GABA class includes types found in both the CNU (notably sAMY and

PALc) and POA. Both classes are highly diverse with multiple embryonic origins. LSX contains a population of *Nkx2-1*$^+$ neurons from the septal eminence[31]. CNU–HYa types in the CEA, AAA and BST express *Meis2* and *Six3*, types in the MEA and certain BST cell types express *Lhx6* and *Nr2f2*, and adult POA types (in the medial preoptic nucleus (MPN), MPO and LPO) express *Lhx6*, *Lhx8*, *Nkx2-1* and *Zic1*, suggesting origins in the ventral MGE, embryonic POA or septal eminence[7] (Fig. 5a and Extended Data Fig. 20f). Developmental mapping could not precisely resolve origins for all adult subclasses (Extended Data Fig. 20a–e). Both LSX and CNU–HYa classes share similar progenitor populations from E11.5 to E14.5 that shift substantially by E16.5. In contrast to cortical and striatal GABAergic types, LSX and CNU–HYa cell-type diversification largely occurs before birth: by E18, 23 out of 25 subclasses and 87 out of 132 supertypes were identified, with near-complete adult profiles (all 25 subclasses and 107 out of 132 supertypes) by P14. This early maturation parallels POA-derived hypothalamic neurons[63]. The 56 CNU–HYa immature class features heterogeneous signatures from multiple regional domains, in contrast to 54 MGE and 55 LGE immature classes, which reflect temporal stages of differentiation (Extended Data Fig. 16).

Discussion

Here we present a comprehensive cell-type taxonomy and spatial atlas of GABAergic neurons in the mouse telencephalon, integrating high-resolution scRNA-seq and adult whole-brain MERFISH data. This enables a unified view of the molecular and anatomical organization of GABAergic neurons, revealing a hierarchical structure in which cell-type relationships reflect both spatial location and developmental origin. Our taxonomy aligns with previous studies across telencephalic regions and identifies many novel types. This resource provides a foundation for future molecular, anatomical and functional investigations of GABAergic circuits. Moreover, we identify key organizational principles, including spatial gene expression gradients, long-range migration and dispersion of related cell types, and marked differences in the timing of diversification between cortical and striatal neurons and between septal, pallidal and preoptic neurons.

GABAergic neurons of the telencephalon originate from five embryonic subpallial domains: MGE, CGE, LGE, POA and septum[3,7,72]. From here, the immature neurons migrate to distant locations and differentiate to form mature brain structures. These processes are regulated by morphogen-regulated TF modules. Many TFs regulating neuronal differentiation during development are also expressed in mature neurons, enabling us to infer developmental origins based on TF expression patterns. Recent studies show that some of these TFs not only specify but also maintain neuronal identity. In invertebrates, the term 'terminal selector genes' describes such TFs, while, in mice, the concept of master regulators has been used[73,74]. Many of these TFs have been identified by the fact that genetic removal during development results in the failure of specific neuronal classes to develop properly. Our study reveals a large set of persistent TFs in telencephalic GABAergic neurons (Fig. 5), with several candidates potentially acting as terminal selectors, to be explored in future genetic perturbation experiments.

A near-universal feature of telencephalic GABAergic neurons is that neurons from distant brain regions can belong to the same cell type. This suggests that neurons sharing developmental origins migrate extensively, populating disparate areas. Building on previous studies, we show that most CTX–CGE and CTX–MGE supertypes and clusters are broadly distributed across the isocortex, HPF, OLF and CTXsp, while some are selectively restricted to the HPF, OLF or CTXsp (Fig. 2 and Extended Data Fig. 4). The CNU–MGE class, containing all striatal interneurons and many pallidal neurons, shares a common MGE origin[75,76] with cortical *Pvalb*$^+$ and *Sst*$^+$ interneurons (Fig. 2). The OB–IMN class, containing olfactory bulb GABAergic neurons, developmentally originates in the LGE and is transcriptomically related to the CNU–LGE class containing the striatal D1 and D2 MSNs and related cell types (Fig. 3 and Extended Data Fig. 2). The OB–IMN class also includes immature neurons generated by adult neurogenesis in the SVZ. Finally, the CNU–HYa class encompasses a highly heterogeneous group of cell types that are widely distributed across the amygdala, PALm/PALv/PALc and hypothalamic POA. This class derives from multiple origins, including MGE, LGE and POA, and may even contain some diencephalic-origin neurons.

At finer resolution, subclass 56 (Sst–Chodl Gaba) in the CNU–MGE class contains both cortical *Sst*$^+$*Chodl*$^+$ long-range projection neurons[33,77,78] and *Sst*$^+$ striatal interneurons, which share closer transcriptomic similarity to each other than to cortical *Sst*$^+$ interneurons (Fig. 2). Similarly, the cholinergic neurons in subclass 58 (PAL–STR Gaba–Chol) include basal forebrain cholinergic projection neurons and striatal cholinergic interneurons, with spatial specificity at the supertype or cluster level (Fig. 2 and Extended Data Fig. 7). Lastly, within the CNU–HYa class, certain supertypes span functionally connected regions, such as the CEA and BST, or the MEA and BST, indicating that a single-cell type can occupy separate but connected nodes of the same neural circuit (Fig. 4 and Extended Data Fig. 9).

As a consequence of widespread migration and dispersion, most telencephalic regions (except for the LSX and OB) contain a heterogeneous mixture of GABAergic neuronal types (Extended Data Fig. 1) with distinct developmental origins and probably distinct connectivity and circuit functions. Within each major progenitor domain, several functional subdomains have been identified that generate distinct GABAergic neurons. For example, the LGE's SVZ is divided dorsoventrally into four progenitor zones (pLGE1–4). The dorsal LGE (dLGE; pLGE1,2) gives rise to OB interneurons and amygdala ITCs, while the ventral LGE (vLGE; pLGE3,4) predominantly produces MSNs[39,79]. Recent data suggest that pLGE3 and pLGE4 preferentially generate D2 and D1 MSNs, respectively[80]. Likewise, MGE is divided into five subdomains (pMGE1–5) generating diverse GABAergic neurons[7,81]. Ventral MGE (pMGE4,5) produces many striatal and pallidal neurons, while dorsal MGE (pMGE1–3) generates cortical interneurons[7,81]. Future studies will be crucial for systematically mapping the spatiotemporal emergence of GABAergic diversity within each progenitor domain.

After reaching their final destinations, GABAergic neurons further refine their identities under the influence of the local environment, evidenced by gene expression gradients and spatial variation. MGE-derived cortical *Sst*$^+$ and *Pvalb*$^+$ interneurons show gradual transcriptomic changes from deep to superficial layers (Extended Data Fig. 10). Similarly, LSX subclasses exhibit multidimensional spatial gradients (Extended Data Fig. 12) that might align with their input/output connections[57,82]. The most pronounced spatial gradients were observed among D1 and D2 MSNs in the STR, correlating with the dorsolateral-to-ventromedial axis (Extended Data Fig. 11). This transcriptional gradient is partly driven by gene modules related to cGMP–PKG signalling, which is crucial for regulating long-term changes in striatal synaptic efficacy[83,84]. Meanwhile, the ventromedial striatal MSNs express neuroactive receptor genes such as *Cnr1* (ref. 36), and are aligned with topographically organized excitatory afferent projections[85].

We identified distinct cholinergic neuronal types in STR, PAL and lateral septum. Most telencephalic cholinergic neurons originate from the MGE, embryonic POA and embryonic septum[86]. The cholinergic precursors originate from the *Nkx2-1*$^+$ domain and are further specified by combinatorial expression of additional TFs[7]. *Lhx6* is essential for specification and migration of MGE-derived GABAergic interneurons in both the CTX and STR[8,87], and *Lhx8* has been associated with the specification of a cholinergic phenotype by actively inducing cholinergic properties[8]. These TFs remain expressed in adulthood. Cholinergic neurons born between E12 and E16 acquire distinct identities, with early- and late-born neurons populating different regions. The late-born population, absent in *Gbx2* knockouts[88], probably corresponds to striatal supertype 260, which still expresses *Gbx2* (Extended Data Fig. 7). A subset of the striatal cholinergic interneurons in this supertype expresses the type-3

vesicular glutamate transporter (*Slc17a8*) and can mediate glutamatergic transmission, which is required for cholinergic signalling onto fast spiking interneurons (subclass 54, STR Prox1 Lhx6 Gaba) as well as acetylcholine-dependent inhibition of MSNs[89]. While striatal cholinergic neurons act primarily as interneurons, pallidal cholinergic neurons mostly project to the isocortex, HIP and amygdala. Their function is linked to topographic organization: dorsal prefrontal cortex receives input from medial SI and NDB cholinergic neurons, whereas ventral regions are innervated by more lateral basal forebrain nuclei. The HIP and ENT mainly receive cholinergic input from MS and NDB neurons[90].

Developmental scRNA-seq data reveal a key difference between GABAergic neurons derived from the MGE, CGE and LGE (CTX–CGE, CTX–MGE, CNU–MGE, CNU–LGE and OB–IMN classes) and those from the embryonic septum and POA (LSX and CNU–HYa classes). The former exhibit extensive postnatal diversification, with clear transcriptomic shifts from late embryonic to adult stages (Fig. 6 and Extended Data Figs. 17, 19 and 20). New types, particularly in CTX–CGE and CTX–MGE, emerge as late as P21 (ref. 91), potentially including postnatally migrating *Htr3a*$^{+}$ CGE-derived neurons[92]. By contrast, septal, preoptic and most pallidal GABAergic neurons transition primarily between E12 and E18, with minimal postnatal diversification, indicating that their adult identities are largely established prenatally[63]. This distinction reflects a broader dorsal–ventral dichotomy. Dorsal cell types (for example, cortical and hippocampal) tend to be larger and more transcriptomically distinct, while ventral types (for example, pallidal and hypothalamic) form numerous small, similar clusters[16]. Evolutionarily, ventral structures are more conserved and support homeostatic functions such as feeding, sleep and reproduction, whereas dorsal structures have evolved to support adaptive behaviours. Dorsal region maturation depends more on sensory inputs, in contrast to ventral regions, which develop largely independent of experience[6,10,63,71,93].

In conclusion, our study provides a detailed transcriptomic characterization of GABAergic neurons in the telencephalon, their spatial locations and their potential developmental origins. Although we have captured several rare cell types, it is possible that, with more cells profiled, a more comprehensive catalogue of cell types could be obtained. However, even this level of diversity poses challenges for linking cell types to morphology, connectivity, physiology and function. Although the spatial organization of transcriptomic types aligns with known circuit organization, further research is needed to link the transcriptomic types to specific projection and connectivity patterns. Also, whereas the current developmental data link adult cell types to their origins, further molecular studies are needed to fully understand the diversification process leading to the large repertoire of telencephalic GABAergic neuronal types.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-09296-1.

1. Tang, X., Jaenisch, R. & Sur, M. The role of GABAergic signalling in neurodevelopmental disorders. *Nat. Rev. Neurosci.* **22**, 290–307 (2021).
2. Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).
3. Hu, J. S., Vogt, D., Sandberg, M. & Rubenstein, J. L. Cortical interneuron development: a tale of time and space. *Development* **144**, 3867–3878 (2017).
4. Wang, Q. et al. The Allen Mouse Brain Common Coordinate Framework: a 3D reference atlas. *Cell* **181**, 936–953 (2020).
5. Swanson, L. W. *Brain Architecture: Understanding the Basic Plan* (Oxford Univ. Press, 2012).
6. Lim, L., Mi, D., Llorca, A. & Marín, O. Development and functional diversification of cortical interneurons. *Neuron* **100**, 294–313 (2018).
7. Flames, N. et al. Delineation of multiple subpallial progenitor domains by the combinatorial expression of transcriptional codes. *J. Neurosci.* **27**, 9682–9695 (2007).
8. Fragkouli, A., van Wijk, N. V., Lopes, R., Kessaris, N. & Pachnis, V. LIM homeodomain transcription factor-dependent specification of bipotential MGE progenitors into cholinergic and GABAergic striatal interneurons. *Development* **136**, 3841–3851 (2009).
9. Anderson, S. A., Eisenstat, D. D., Shi, L. & Rubenstein, J. L. Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. *Science* **278**, 474–476 (1997).
10. Nóbrega-Pereira, S. et al. Origin and molecular specification of globus pallidus neurons. *J. Neurosci.* **30**, 2824–2834 (2010).
11. Tufo, C. et al. Development of the mammalian main olfactory bulb. *Development* **149**, dev200210 (2022).
12. Li, J. et al. Transcription factors Sp8 and Sp9 coordinately regulate olfactory bulb interneuron development. *Cereb. Cortex* **28**, 3278–3294 (2018).
13. Bandler, R. C., Mayer, C. & Fishell, G. Cortical interneuron specification: the juncture of genes, time and geometry. *Curr. Opin. Neurobiol.* **42**, 17–24 (2017).
14. Turrero García, M. & Harwell, C. C. Radial glia in the ventral telencephalon. *FEBS Lett.* **591**, 3942–3959 (2017).
15. Schmitz, M. T. et al. The development and evolution of inhibitory neurons in primate cerebrum. *Nature* **603**, 871–877 (2022).
16. Yao, Z. et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).
17. Lim, D. A. & Alvarez-Buylla, A. The adult ventricular-subventricular zone (V-SVZ) and olfactory bulb (OB) neurogenesis. *Cold Spring Harb. Perspect. Biol.* **8**, a018820 (2016).
18. Cebrian-Silla, A. et al. Single-cell analysis of the ventricular-subventricular zone reveals signatures of dorsal and ventral adult neurogenesis. *eLife* **10**, e67436 (2021).
19. Batista-Brito, R., Close, J., Machold, R. & Fishell, G. The distinct temporal origins of olfactory bulb interneuron subtypes. *J. Neurosci.* **28**, 3966–3975 (2008).
20. Tepe, B. et al. Single-cell RNA-seq of mouse olfactory bulb reveals cellular heterogeneity and activity-dependent molecular census of adult-born neurons. *Cell Rep.* **25**, 2689–2703 (2018).
21. Gelman, D. et al. A wide diversity of cortical GABAergic interneurons derives from the embryonic preoptic area. *J. Neurosci.* **31**, 16570–16580 (2011).
22. Yao, Z. et al. A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222–3241 (2021).
23. Gouwens, N. W. et al. Integrated morphoelectric and transcriptomic classification of cortical GABAergic cells. *Cell* **183**, 935–953 (2020).
24. Pelkey, K. A. et al. Hippocampal GABAergic inhibitory interneurons. *Physiol. Rev.* **97**, 1619–1747 (2017).
25. Urrutia-Piñones, J., Morales-Moraga, C., Sanguinetti-González, N., Escobar, A. P. & Chiu, C. Q. Long-range GABAergic projections of cortical origin in brain function. *Front. Syst. Neurosci.* **16**, 841869 (2022).
26. Luo, X. et al. Transcriptomic profile of the subiculum-projecting VIP GABAergic neurons in the mouse CA1 hippocampus. *Brain Struct. Funct.* **224**, 2269–2280 (2019).
27. Frazer, S. et al. Transcriptomic and anatomic parcellation of 5-HT3AR expressing cortical interneuron subtypes revealed by single-cell RNA sequencing. *Nat. Commun.* **8**, 14219 (2017).
28. Fisher, J. et al. Cortical somatostatin long-range projection neurons and interneurons exhibit divergent developmental trajectories. *Neuron* https://doi.org/10.1016/j.neuron.2023.11.013 (2023).
29. Fang, L. Z. & Creed, M. C. Updating the striatal-pallidal wiring diagram. *Nat. Neurosci.* **27**, 15–27 (2024).
30. Courtney, C. D., Pamukcu, A. & Chan, C. S. Cell and circuit complexity of the external globus pallidus. *Nat. Neurosci.* **26**, 1147–1159 (2023).
31. Turrero García, M. et al. Transcriptional profiling of sequentially generated septal neuron fates. *eLife* **10**, e71545 (2021).
32. Chen, R. et al. Decoding molecular and cellular heterogeneity of mouse nucleus accumbens. *Nat. Neurosci.* **24**, 1757–1771 (2021).
33. Tasic, B. et al. Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).
34. Hegeman, D. J., Hong, E. S., Hernández, V. M. & Chan, C. S. The external globus pallidus: progress and perspectives. *Eur. J. Neurosci.* **43**, 1239–1265 (2016).
35. Abdi, A. et al. Prototypic and arkypallidal neurons in the dopamine-intact external globus pallidus. *J. Neurosci.* **35**, 6667–6688 (2015).
36. Stanley, G., Gokce, O., Malenka, R. C., Südhof, T. C. & Quake, S. R. Continuous and discrete neuron types of the adult murine striatum. *Neuron* **105**, 688–699 (2020).
37. Gokce, O. et al. Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. *Cell Rep.* **16**, 1126–1137 (2016).
38. Paxinos, G. & Franklin, K. B. J. *Paxinos and Franklin's the Mouse Brain in Stereotaxic Coordinates*. (Academic, 2019).
39. Kuerbitz, J. et al. Loss of intercalated cells (ITCs) in the mouse amygdala of Tshz1 mutants correlates with fear, depression, and social interaction phenotypes. *J. Neurosci.* **38**, 1160–1177 (2018).
40. Waclaw, R. R., Ehrman, L. A., Pierani, A. & Campbell, K. Developmental origin of the neuronal subtypes that comprise the amygdalar fear circuit in the mouse. *J. Neurosci.* **30**, 6944–6953 (2010).
41. Bandler, R. C. et al. Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).
42. Pardo-Bellver, C., Cádiz-Moretti, B., Novejarque, A., Martínez-García, F. & Lanuza, E. Differential efferent projections of the anterior, posteroventral, and posterodorsal subdivisions of the medial amygdala in mice. *Front. Neuroanat.* **6**, 33 (2012).
43. Wang, Y. et al. Multimodal mapping of cell types and projections in the central nucleus of the amygdala. *eLife* **12**, e84262 (2023).
44. Nguyen, Q. A. T. et al. Hypothalamic representation of the imminence of predator threat detected by the vomeronasal organ in mice. *eLife* **12**, RP92982 (2024).
45. Miller, S. M., Marcotulli, D., Shen, A. & Zweifel, L. S. Divergent medial amygdala projections regulate approach-avoidance conflict behavior. *Nat. Neurosci.* **22**, 565–575 (2019).
46. Knoedler, J. R. et al. A functional cellular framework for sex and estrous cycle-dependent gene expression and behavior. *Cell* **185**, 654–671 (2022).

47. Raam, T. & Hong, W. Organization of neural circuits underlying social behavior: a consideration of the medial amygdala. *Curr. Opin. Neurobiol.* **68**, 124–136 (2021).
48. Hochgerner, H. et al. Neuronal types in the mouse amygdala and their transcriptional response to fear conditioning. *Nat. Neurosci.* **26**, 2237–2249 (2023).
49. Choi, G. B. et al. Lhx6 delineates a pathway mediating innate reproductive behaviors from the amygdala to the hypothalamus. *Neuron* **46**, 647–660 (2005).
50. Shemesh, Y. et al. Ucn3 and CRF-R2 in the medial amygdala regulate complex social dynamics. *Nat. Neurosci.* **19**, 1489–1496 (2016).
51. Pare, D. & Duvarci, S. Amygdala microcircuits mediating fear expression and extinction. *Curr. Opin. Neurobiol.* **22**, 717–723 (2012).
52. Hammack, S. E., Braas, K. M. & May, V. Chemoarchitecture of the bed nucleus of the stria terminalis: neurophenotypic diversity and function. *Handb. Clin. Neurol.* **179**, 385–402 (2021).
53. Forray, M. I. & Gysling, K. Role of noradrenergic projections to the bed nucleus of the stria terminalis in the regulation of the hypothalamic-pituitary-adrenal axis. *Brain Res. Rev.* **47**, 145–160 (2004).
54. Rymar, V. V. & Sadikot, A. F. Laminar fate of cortical GABAergic interneurons is dependent on both birthdate and phenotype. *J. Comp. Neurol.* **501**, 369–380 (2007).
55. Valcanis, H. & Tan, S.-S. Layer specification of transplanted interneurons in developing mouse neocortex. *J. Neurosci.* **23**, 5113–5122 (2003).
56. Sultan, K. T. et al. Progressive divisions of multipotent neural progenitors generate late-born chandelier cells in the neocortex. *Nat. Commun.* **9**, 4595 (2018).
57. Besnard, A. & Leroy, F. Top-down regulation of motivated behaviors via lateral septum sub-circuits. *Mol. Psychiatry* **27**, 3119–3128 (2022).
58. Reid, C. M. et al. Multimodal classification of neurons in the lateral septum. Preprint at *bioRxiv* https://doi.org/10.1101/2024.02.15.580381 (2024).
59. Allaway, K. C. et al. Genetic and epigenetic coordination of cortical interneuron development. *Nature* **597**, 693–697 (2021).
60. La Manno, G. et al. Molecular architecture of the developing mouse brain. *Nature* **596**, 92–96 (2021).
61. Lee, D. R. et al. Transcriptional heterogeneity of ventricular zone cells in the ganglionic eminences of the mouse forebrain. *eLife* **11**, e71864 (2022).
62. Mayer, C. et al. Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).
63. Kaplan, H. S. et al. Sensory input, sex and function shape hypothalamic cell type development. *Nature* https://doi.org/10.1038/s41586-025-08603-0 (2025).
64. Thompson, C. L. et al. A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain. *Neuron* **83**, 309–323 (2014).
65. Su-Feher, L. et al. Single cell enhancer activity distinguishes GABAergic and cholinergic lineages in embryonic mouse basal ganglia. *Proc. Natl Acad. Sci. USA* **119**, e2108760119 (2022).
66. Li, Z. et al. Transcription factor Sp9 is a negative regulator of D1-type MSN development. *Cell Death Discov.* **8**, 301 (2022).
67. García-Moreno, F. et al. A neuronal migratory pathway crossing from diencephalon to telencephalon populates amygdala nuclei. *Nat. Neurosci.* **13**, 680–689 (2010).
68. Chen, Y.-J. J. et al. Use of "MGE enhancers" for labeling and selection of embryonic stem cell-derived medial ganglionic eminence (MGE) progenitors and neurons. *PLoS ONE* **8**, e61956 (2013).
69. Taniguchi, H., Lu, J. & Huang, Z. J. The spatial and temporal origin of chandelier cells in mouse neocortex. *Science* **339**, 70–74 (2013).
70. Dehorter, N. et al. Tuning of fast-spiking interneuron properties by an activity-dependent transcriptional switch. *Science* **349**, 1216–1220 (2015).
71. De Marco García, N. V., Karayannis, T. & Fishell, G. Neuronal activity is required for the development of specific cortical interneuron subtypes. *Nature* **472**, 351–355 (2011).
72. Huilgol, D. & Tole, S. Cell migration in the developing rodent olfactory system. *Cell. Mol. Life Sci.* **73**, 2467–2490 (2016).
73. Hobert, O. Regulatory logic of neuronal diversity: terminal selector genes and selector motifs. *Proc. Natl Acad. Sci. USA* **105**, 20067–20071 (2008).
74. Deneris, E. S. & Hobert, O. Maintenance of postmitotic neuronal cell identity. *Nat. Neurosci.* **17**, 899–907 (2014).
75. Harwell, C. C. et al. Wide dispersion and diversity of clonally related inhibitory interneurons. *Neuron* **87**, 999–1007 (2015).
76. Mayer, C. et al. Clonally related forebrain interneurons disperse broadly across both functional areas and structural boundaries. *Neuron* **87**, 989–998 (2015).
77. Tomioka, R. et al. Demonstration of long-range GABAergic connections distributed throughout the mouse neocortex. *Eur. J. Neurosci.* **21**, 1587–1600 (2005).
78. Paul, A. et al. Transcriptional architecture of synaptic communication delineates GABAergic neuron identity. *Cell* **171**, 522–539 (2017).
79. Stenman, J., Toresson, H. & Campbell, K. Identification of two distinct progenitor populations in the lateral ganglionic eminence: implications for striatal and olfactory bulb neurogenesis. *J. Neurosci.* **23**, 167–174 (2003).
80. Xu, Z. et al. SP8 and SP9 coordinately promote D2-type medium spiny neuron production by activating Six3 expression. *Development* **145**, dev165456 (2018).
81. Wonders, C. P. et al. A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence. *Dev. Biol.* **314**, 127–136 (2008).
82. Risold, P. Y. & Swanson, L. W. Connections of the rat lateral septal complex. *Brain Res. Rev.* **24**, 115–195 (1997).
83. Calabresi, P., Picconi, B., Tozzi, A. & Di Filippo, M. Dopamine-mediated regulation of corticostriatal synaptic plasticity. *Trends Neurosci.* **30**, 211–219 (2007).
84. Jouvert, P. et al. Activation of the cGMP pathway in dopaminergic structures reduces cocaine-induced EGR-1 expression and locomotor activity. *J. Neurosci.* **24**, 10716–10725 (2004).
85. Hunnicutt, B. J. et al. A comprehensive excitatory input map of the striatum reveals novel functional organization. *eLife* **5**, e19103 (2016).
86. Magno, L. et al. NKX2-1 is required in the embryonic septum for cholinergic system development, learning, and memory. *Cell Rep.* **20**, 1572–1584 (2017).
87. Flandin, P. et al. Lhx6 and Lhx8 coordinately induce neuronal expression of Shh that controls the generation of interneuron progenitors. *Neuron* **70**, 939–950 (2011).
88. Chen, L., Chatterjee, M. & Li, J. Y. H. The mouse homeobox gene Gbx2 is required for the development of cholinergic interneurons in the striatum. *J. Neurosci.* **30**, 14824–14834 (2010).
89. Higley, M. J. et al. Cholinergic interneurons mediate fast VGluT3-dependent glutamatergic transmission in the striatum. *PLoS ONE* **6**, e19155 (2011).
90. Gielow, M. R. & Zaborszky, L. The input-output relationship of the cholinergic basal forebrain. *Cell Rep.* **18**, 1817–1830 (2017).
91. Gao, Y. et al. Continuous cell-type diversification in mouse visual cortex development. *Nature* https://doi.org/10.1038/s41586-025-09644-1 (2025).
92. Inta, D. et al. Neurogenesis and widespread forebrain migration of distinct GABAergic neurons from the postnatal subventricular zone. *Proc. Natl Acad. Sci. USA* **105**, 20994–20999 (2008).
93. Fishell, G. & Kepecs, A. Interneuron types as attractors and controllers. *Annu. Rev. Neurosci.* **43**, 1–30 (2020).

Methods

Sample collection and data analysis for the P56 dataset

Most of the methods that apply to the adult P56 10x scRNA and MERFISH datasets used for this paper were described previously[16] and the following methods are either newly introduced or a modified version was used for this Article. The P56 dataset used in this study is from the subpallium GABA neighbourhood in the WMB cell atlas[16], with a total of 611,423 high-quality single-cell transcriptomes (10xv2: 269,307 cells, 3,567 ± 1,264 (mean ± s.d.) genes per cell, 9,328 ± 5,502 unique molecular identifiers (UMIs) per cell; 10xv3: 342,116 cells, 5,949 ± 1,625 genes per cell, 26,476 ± 14,943 UMIs per cell) (Supplementary Table 3).

UMAP projection

We performed principal component analysis (R package stats, v.4.4.1, RRID: SCR_025968) based on the imputed gene expression matrix of 4,895 marker genes using the 10xv3 reference. We selected the top 100 principal components, then removed one principal component with more than 0.7 correlation with the technical bias vector, defined as $\log_2$[gene count] for each cell. We used the remaining principal components as an input to create 2D and 3D UMAPs (v.0.5.6, RRID: SCR_018217)[94] using the parameters nn.neighbors = 25 and md = 0.4.

Constellation plot

To generate the constellation plot, each transcriptomic supertype was represented by a node (circle), of which the surface area reflected the number of cells within the supertype in log scale. The position of nodes was based on the centroid positions of the corresponding supertypes in UMAP coordinates. The relationships between nodes were indicated by edges that were calculated as follows. For each cell, 15 nearest neighbours in reduced-dimension space were determined and summarized by supertypes. For each supertype, we then calculated the fraction of nearest neighbours that were assigned to other supertypes. The edges connected two nodes in which at least one of the nodes had >5% of nearest neighbours in the connecting node. The width of the edge at the node reflected the fraction of nearest neighbours that were assigned to the connecting node and was scaled to node size. For all nodes in the plot, we then determined the maximum fraction of outside neighbours and set this as edge width = 100% of node width. The function for creating these plots, plot_constellation, is included in scrattch.bigcat (v.0.0.5; https://github.com/AllenInstitute/scrattch.bigcat)[16].

Imputation of scRNA-seq data into the MERFISH space

The MERFISH dataset was collected using only 500 genes. To obtain the spatial distribution of all of the genes, we projected gene expression of the MERFISH dataset to the 10xv3 scRNA-seq dataset using the impute_knn_global function in the scrattch.bigcat package[16]. We used the self-imputed 10xv3 dataset as a reference, meaning that the expression of each 10xv3 cell was first imputed based on its nearest 15 neighbours in the reduced principal component space. This decision was made to ensure that, in the following hierarchical imputation step, the transitions between major cell types were preserved. The imputation was conducted in the order specified by the hierarchy defined by class and subclass. At the root, we imputed the expression for all of the genes for each MERFISH cell based on the average expression of their nearest neighbours from the reference 10xv3 dataset, defined by the cosine similarity using all 500 MERFISH genes. In each of the following iterations, we selected the node to which each MERFISH cell was assigned and imputed only the expression of the differentially expressed genes (DEGs) based on pairwise comparison for all of the clusters under this node. The nearest neighbours for imputation were selected from the clusters under this node in the reference dataset, using only the subset of DEGs that were present on the MERFISH gene panel. We repeated this procedure until reaching the leaf node. This strategy enabled us to preserve the cell-type resolution during imputation, making it less susceptible to the global platform differences between MERFISH and scRNA-seq.

Analysis of spatial gene expression gradients

We performed ICA using fastICA (v.1.2-5.1; RRID: SCR_013110)[95] to decompose the gene expression matrix into independent components. These components were then projected onto the imputed MERFISH data to determine whether the component represents a spatial gene expression program. For components that represent a spatial gene expression program, the top loading genes were selected and visualized on both the UMAP in scRNA-seq space and on sections in the imputed MERFISH space. We evaluated the gene modules in the identified individual components and applied UCell (v.2.8.0; https://doi.org/10.18129/B9.bioc.UCell)[96] to assign a gene module score based on both positive and negative genes to each cell.

Spatial domain clustering

We used BANKSY (v.0.99.12; https://github.com/prabhakarlab/Banksy)[97] to perform spatial domain clustering within BST neurons. This algorithm implements a feature-augmentation approach to map domains by integrating the transcriptional profiles of individual cells with their physical distances and tissue neighbourhood context. As the MERFISH data were registered to the CCFv3, it enabled us to subset BST neurons from the MERFISH data. We used the 8,988 BST neurons, their spatial location and their 500-gene expression profile as input for BANKSY.

Assessing the concordance of cell-type taxonomy between the adult subpallium GABAergic cell-type atlas and adult external datasets

We performed mapping of cells from each adult external dataset[18,20,32,36,46,48] to the 10xv3 whole-brain dataset using treeMap function from scrattch.mapping package (v.0.55; https://github.com/AllenInstitute/scrattch-mapping)[98]. The reference cell-type taxonomy was organized in a hierarchy defined by class and subclass. At each node, the top markers were selected that best discriminate the clusters belonging to different child nodes. Starting at the root, cells were assigned to the closest cluster centroid from all of the clusters under the given node based on the selected node markers using the cosine similarity metric. This mapping procedure was repeated until reaching the leaf nodes. To assess the mapping confidence, we subsampled 80% of the markers at each node and repeated the mapping process 100 times. In each bootstrapping step, we also computed the cosine similarity of the cell to the mapped cluster based on the markers for all of the nodes along the mapping path and calculated the average similarity across all 100 bootstrapping iterations. This score was used to assess the quality of the mapping. Cells with a score above 0.5 were used to generate a confusion matrix showing the proportion of cells jointly found between two types and their Jaccard similarity score.

Measuring the similarity between MSN D1 and D2 clusters

We computed the nearest neighbours from STR D1 MSNs for STR D2 MSNs and vice versa using the cosine similarity metric based on the same marker list used to define the edge weights for the constellation plots. To select the most similar pairs between D1 and D2 MSN types, we selected all of the pairs with a Pearson correlation of greater than 0.93 and with at least 30% of the *k*-nearest neighbours from one cluster belonging to the other cluster in the pair.

Developmental scRNA-seq data collection

Mouse breeding and husbandry. All experimental procedures related to the use of mice were approved by the Institutional Animal Care and Use Committee of the AIBS, in accordance with NIH guidelines. Mice were housed in a room with temperature (21–22 °C) and humidity (40–51%) control within the vivarium of the AIBS. Mice were

provided food and water ad libitum and were maintained on a regular 14 h–10 h light–dark cycle. Mice were maintained on the C57BL/6J (JAX, 000664) background. We excluded any mice with anophthalmia or microphthalmia.

The mothers of all experimental mice were placed into a fresh cage when the embryos were aged about E8. We used 6 embryos to collect 74,550 cells from ages E11.5 ($n = 1$), E12.5 ($n = 1$), E13.5 ($n = 2$) and E14.5 ($n = 2$). From ages E11.5 and E12.5, we collected whole-brain tissue and, from ages E13.5 and E14.5, we collected cerebrum and brain stem (CH–BS). From P0 pups ($n = 7$), we collected 171,041 cells; from P4 pups ($n = 8$) we collected 259,172 cells; from P8 pups ($n = 7$) we collected 326,245 cells; from P12 pups ($n = 7$) we collected 260,336 cells; and from P14 pups ($n = 8$) we collected 331,137 cells. Cells from postnatal mice were collected from both male and female mice across six dissection regions of interest (ROIs): OLF, CTXsp, isocortex, HPF, CNU and HY. No statistical methods were used to predetermine sample size. All donor animals used for the developmental scRNA-seq data generation are listed in Supplementary Table 4. Brain dissections for all groups took place in the morning. Randomization and blinding were not performed in this study. scRNA-seq data were generated and analysed using standardized, automated computational pipelines. As such, data collection and analysis were not influenced by operator bias. Potential batch effects were accounted for using established normalization and integration methods during downstream analysis.

Single-cell isolation. Single cells were isolated according to a cell-isolation protocol developed at AIBS[99]. The brain was dissected, submerged in artificial cerebrospinal fluid (ACSF), embedded in 2% agarose and sliced into 350 µm coronal sections on a compresstome (Precisionary Instruments). Block-face images were captured during slicing. ROIs were then microdissected from the slices and dissociated into single cells.

Dissected tissue pieces were digested with 30 U ml^{-1} papain (Worthington PAP2) in ACSF for 30 min at 30 °C. Owing to the short incubation period in a dry oven, we set the oven temperature to 35 °C to compensate for the indirect heat exchange, with a target solution temperature of 30 °C. Enzymatic digestion was quenched by exchanging the papain solution three times with quenching buffer (ACSF with 1% FBS and 0.2% BSA). The samples were incubated on ice for 5 min before trituration. The tissue pieces in the quenching buffer were triturated through a fire-polished pipette with 600 µm diameter opening approximately 20 times. The tissue pieces were allowed to settle and the supernatant, which now contained suspended single cells, was transferred to a new tube. Fresh quenching buffer was added to the settled tissue pieces, and trituration and supernatant transfer were repeated using 300 µm and 150 µm fire-polished pipettes. The single-cell suspension was passed through a 70 µm filter into a 15 ml conical tube with 500 µl of high-BSA buffer (ACSF with 1% FBS and 1% BSA) at the bottom to help cushion the cells during centrifugation at 100*g* in a swinging-bucket centrifuge for 10 min. The supernatant was discarded, and the cell pellet was resuspended in the quenching buffer. The concentration of the resuspended cells was quantified, and cells were immediately loaded onto the 10x Genomics Chromium controller.

cDNA amplification and library construction. The E11.5 to E14.5 cell suspensions were processed using the Chromium Single Cell 3′ Reagent Kit v3 (1000075, 10x Genomics). We followed the manufacturer's instructions for cell capture, barcoding, reverse transcription, cDNA amplification and library construction[100]. We loaded 8,283 ± 703 (mean ± s.d.) cells per port. We targeted a sequencing depth of 120,000 reads per cell; the actual average achieved was 70,324 ± 62,149 (mean ± s.d.) reads per cell across 9 libraries.

The P0 cell suspensions were processed using the Chromium Single Cell 3′ Reagent Kit v3.1 (1000268, 10x Genomics). We followed the manufacturer's instructions for cell capture, barcoding, reverse transcription, cDNA amplification and library construction[101]. We loaded 11,468 ± 1,735 cells per port. We targeted sequencing depth of 120,000 reads per cell; the actual average achieved was 114,853 ± 30,439 reads per cell across 13 libraries.

The P4 to P14 cell suspensions were processed using the Chromium Next GEM-X Single Cell 3′ Reagent Kit v4 (1000691, 10x Genomics). We followed the manufacturer's instructions for cell capture, barcoding, reverse transcription, cDNA amplification and library construction. We loaded 38,740 ± 8,097 cells per port. We targeted sequencing depth of 120,000 reads per cell; the actual average achieved was 67,175 ± 24,394 reads per cell across 58 libraries.

Sequencing data processing and QC. Processing of 10x Genomics scRNA-seq libraries was performed as described previously[22]. In brief, libraries were sequenced on the Illumina NovaSeq 6000 system, and sequencing reads were aligned to the mouse reference transcriptome (M21, GRCm38.p6) using the 10x Genomics CellRanger pipeline (v.6.1.1) with the default parameters. To remove low-quality cells, we applied similar quality-control analysis and thresholding based on gene detection, quality control score and doublet score as described previously[16]. Doublets were identified using a modified version of the DoubletFinder algorithm, which is available in scrattch.hicat (v.0.1.0, RRID: SCR_01809)[102]. We removed cells using a gene count cut-off of 2,000, quality-control score cut-off of 200 and doublet score cut-off of 0.3. Cells were collected from six distinct dissection ROIs: OLF, CTXsp, isocortex, HPF, CNU and HY. The subpallium GABAergic cells were selected based on marker gene expression. In brief, Seurat (v.5.1.0, RRID: SCR_016341) was used for initial clustering (default parameters). A gene module score was calculated for the following genes: *Dlx1*, *Dlx2*, *Dlx5*, *Dlx6* and *Arx*. Clusters with a score of <0 were removed from further analysis. In total, 387,148 newly generated cells were included in the dataset (Supplementary Table 4).

Published datasets. We combined our developmental dataset with eight published datasets from other laboratories covering the embryonic to early postnatal period (E7 to P2)[31,59–63] (Supplementary Table 5). Raw counts and, if available, cell-type and metadata annotations for each of the datasets were downloaded from their respective sources. We applied minimal quality-control thresholds to the external datasets (Supplementary Table 5). Cells with fewer than 1,000 genes per cell were removed from the dataset, resulting in a total of 227,421 remaining cells for integration with our own datasets. The combined datasets vary greatly in number of genes and UMIs detected (Extended Data Fig. 13a,b).

Integration of scRNA-seq data

We used all developmental datasets and the 10xv3 portion of the adult P56 dataset, totalling 956,213 cells for integration. Raw counts were log-normalized and subsequently used to select 4,000 highly variable genes. For dimensionality reduction and batch correction, the raw counts of the highly variable genes were used to train the scvi-tools (v.0.17.1, RRID: SCR_026673) scVI variational autoencoder (parameters: batch_key = 'dataset', n_latent = 32, gene_likelihood = 'nb', n_layers = 3, continuous_covariate_keys = log2ngene)[103]. The latent dimensions were used to generate a *k*-nearest neighbours graph for UMAP generation and subclustering.

Clustering scRNA-seq data

To assign cell-type identity to cells at P14, each cell was mapped onto the WMB taxonomy using hierarchical approximate nearest neighbour (HANN) mapping available in the scrattch-mapping package[98]. For P12 cells, we assigned each cell's identity by mapping it to the nearest supertype centroid in the adjacent older age group, P14, using scrattch.mapping. P8 cells were mapped to P12 and P4 cells were mapped to

P8. For cells from E7 to P2, ages were binned together based on the similarity between temporal age and predicted age (Extended Data Fig. 13d). Predicted age was calculated as follows; within the integrated scVI latent space, the ten nearest neighbours for each single cell were determined. Each cell was assigned a predicted age based on the common age among the ten neighbours. This process was repeated ten times, using the predicted age from the previous iteration as starting point. Cells from P0 and P2 were binned and mapped to P4, cells from E15 to E18.5 were binned and mapped to P0–P2, cells from E11 to E14.5 were binned and mapped to E15–E18.5, and cells from E7 to E10 were binned and mapped to E11–E14.5.

After assigning the broad cell types, iterative clustering was performed within assigned subclasses in the scVI latent space using the scrattch.bigcat package as described previously[16]. Clustering was performed using the following DEG criteria: q1.th = 0.4, q.diff.th = 0.7, de.score.th = 120, min.cells = 10. Cell-type annotation of the developmental scRNA-seq dataset is shown in Supplementary Table 6.

Reporting summary

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

Data availability

The scRNA-seq and MERFISH datasets for this study are part of the Allen WMB cell type atlas and are accessible through the Neuroscience Multi-omic Data Archive (NeMO, RRID: SCR_016152) and the Brain Image Library (BIL, RRID: SCR_017272). The adult 10x scRNA-seq datasets and the developmental scRNA-seq datasets are being made available through the BRAIN Initiative Cell Atlas Network (BICAN, RRID: SCR_022794) (FASTQ files) and are available at NeMO (https://assets.nemoarchive.org/dat-zqurqvh). The MERFISH dataset is available at BIL (https://doi.org/10.35077/g.610). The subpallium–GABA cell type taxonomy is available from the Allen Brain Cell Atlas (RRID: SCR_024440) to visualize both sc/snRNA-seq and MERFISH datasets. Instructions for access of the processed adult 10x scRNA-seq data are available at GitHub (https://github.com/AllenInstitute/abc_atlas_access/blob/main/descriptions/WMB-10X.md), and instructions for access to the processed MERFISH data are available at GitHub (https://github.com/AllenInstitute/abc_atlas_access/blob/main/descriptions/MERFISH-C57BL6J-638850.md). The following published developmental datasets were used: ref. 59 (GSE164570); ref. 63 (GSE280964); ref. 61 (GSE167013 and GSE190593); ref. 62 (GSE103983 and GSE104156); ref. 60 (PRJNA637987); and ref. 31 (GSE184879).

Code availability

The R package scrattch.bigcat is available at GitHub (https://github.com/AllenInstitute/scrattch.bigcat). Notebooks with examples of data analysis code used in this manuscript are available at GitHub (https://alleninstitute.github.io/scrattch.example/).

94. McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
95. Marchini, J. L., Heaton, C. & Ripley, B. D. fastICA: FastICA algorithms to perform ica and projection pursuit (2021); https://CRAN.R-project.org/package=fastICA
96. Andreatta, M. & Carmona, S. J. UCell: Robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798 (2021).
97. Singhal, V. et al. BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat. Genet.* **56**, 431–441 (2024).
98. Johansen, N., Miller, J., Lee, C. & Kapen, I. AllenInstitute/Scrattch.Mapping: V0.55. *Zenodo* https://doi.org/10.5281/zenodo.10939013 (2024).
99. Allen Institute for Brain Science. Mouse whole cell tissue processing for 10x Genomics platform V.9. *protocols.io* https://doi.org/10.17504/protocols.io.q26g7b52klwz/v9 (2022).
100. Allen Institute for Brain Science. 10xV3 Genomics sample processing protocol. *protocols.io* https://doi.org/10.17504/protocols.io.bq7cmziw (2021).
101. Allen Institute for Brain Science. 10Xv3.1 Genomics sample processing. *protocols.io* https://doi.org/10.17504/protocols.io.dm6gpwd8jlzp/v3 (2024).
102. Yao, Z. et al. AllenInstitute/Scrattch.Hicat: Doi_release. *Zenodo* https://doi.org/10.5281/zenodo.11405898 (2024).
103. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).

Acknowledgements We thank the members of the Transgenic Colony Management, Lab Animal Services, Molecular Biology and Histology teams at the Allen Institute for technical support; and the Allen Institute founder, Paul G. Allen, for his vision, encouragement and support. The research was funded by the U19MH114830 and U01MH130962 grants from the National Institute of Mental Health to H.Z., under the BRAIN Initiative of the National Institutes of Health (NIH). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH and its subsidiary institutes. This work was also supported by the Allen Institute for Brain Science.

Author contributions Conceptualization: H.Z. Data analysis lead and coordination: C.T.J.v.V. Data generation scRNA-seq: C.T.J.v.V., T. Casper, M.C., R.F., J. Gloe, J. Guzman, C.H., W.H., K.J., R.MC., T. Cardenas, K.R., E.D.T., A.T., N.D., K.A.S., Z.Y. and H.Z. Data processing and analysis (scRNA-seq): C.T.J.v.V., Y.G., C.L., A.B.C., R.C., T.D., J. Goldy, B.N., K.A.S., Z.Y. and H.Z. Data generation (MERFISH): M.K., D.M., J.W. and H.Z. Data processing and analysis (MERFISH): C.T.J.v.V., M.K., D.M., S.D., M.J.H., L.N., J.W., Z.Y. and H.Z. Project management: C.M.P., L.K. and K.A.S. Management and supervision: C.T.J.v.V., D.M., N.D., L.N., J.W., K.A.S., B.T., Z.Y. and H.Z. Manuscript writing and figure generation: C.T.J.v.V., Z.Y. and H.Z. Manuscript review and editing: C.T.J.v.V. and H.Z.

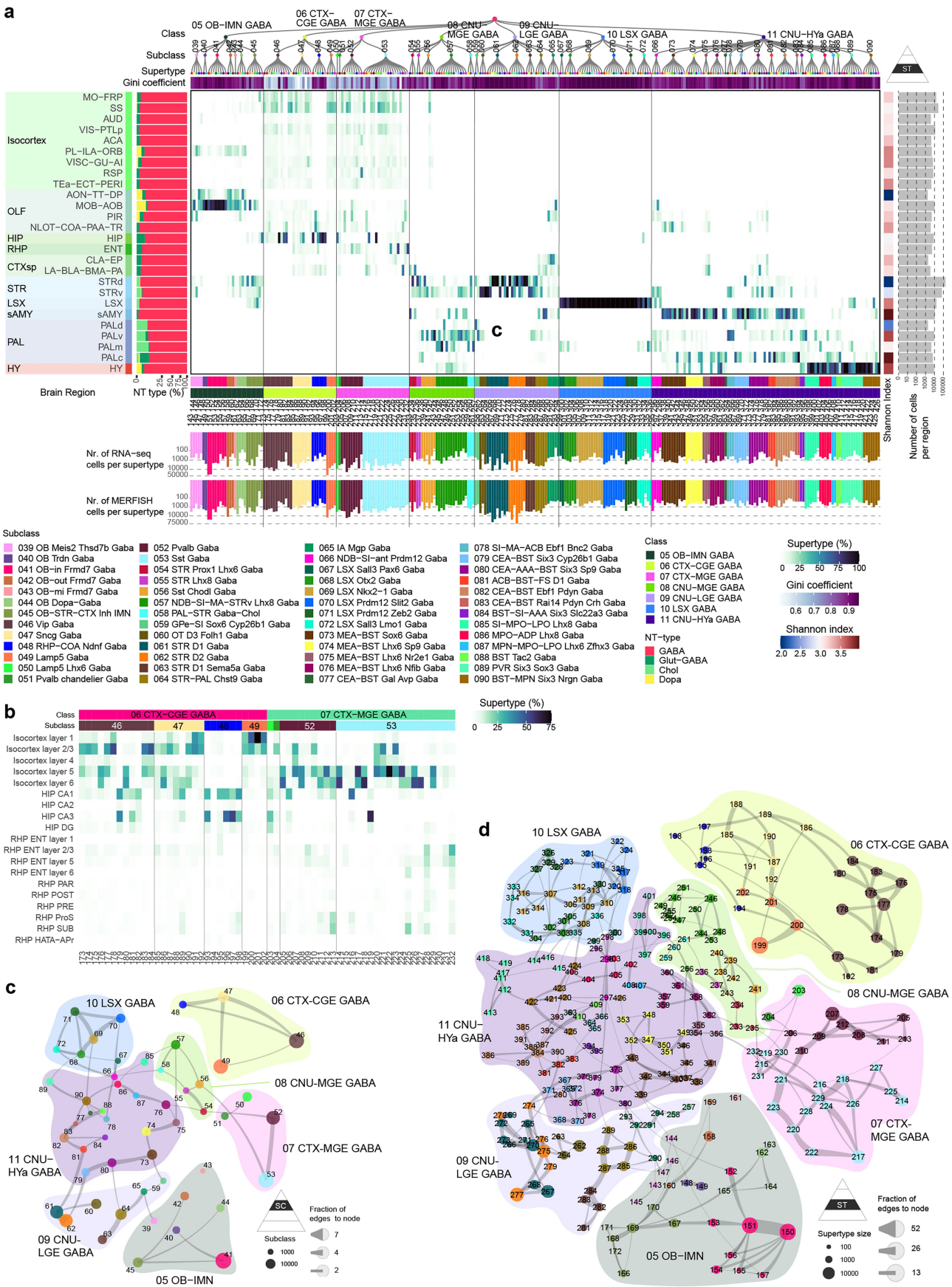
Competing interests H.Z. is on the scientific advisory board of MapLight Therapeutics. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-09296-1.

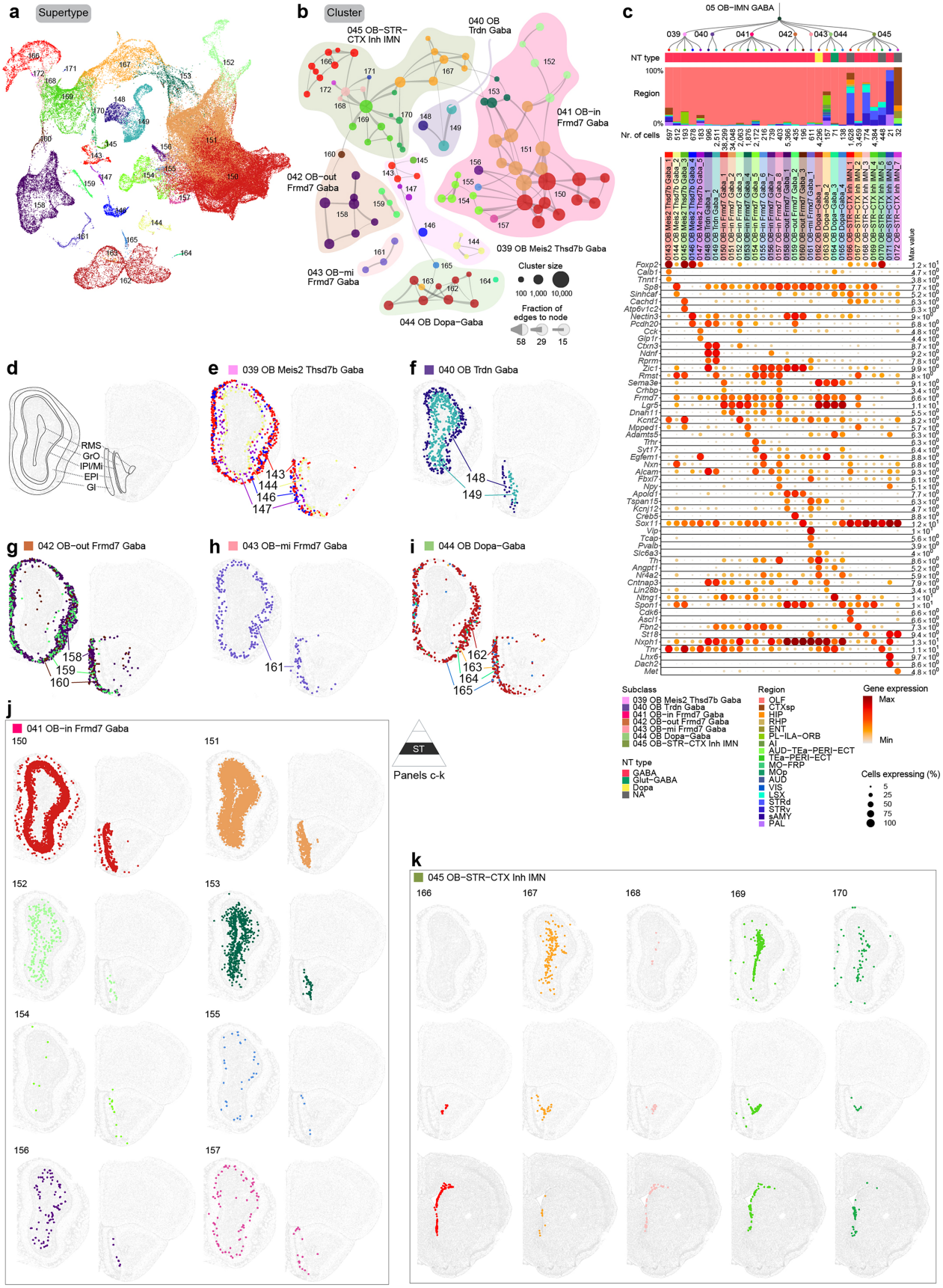
Correspondence and requests for materials should be addressed to Cindy T. J. van Velthoven or Hongkui Zeng.

Peer review information *Nature* thanks Lynette Lim and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



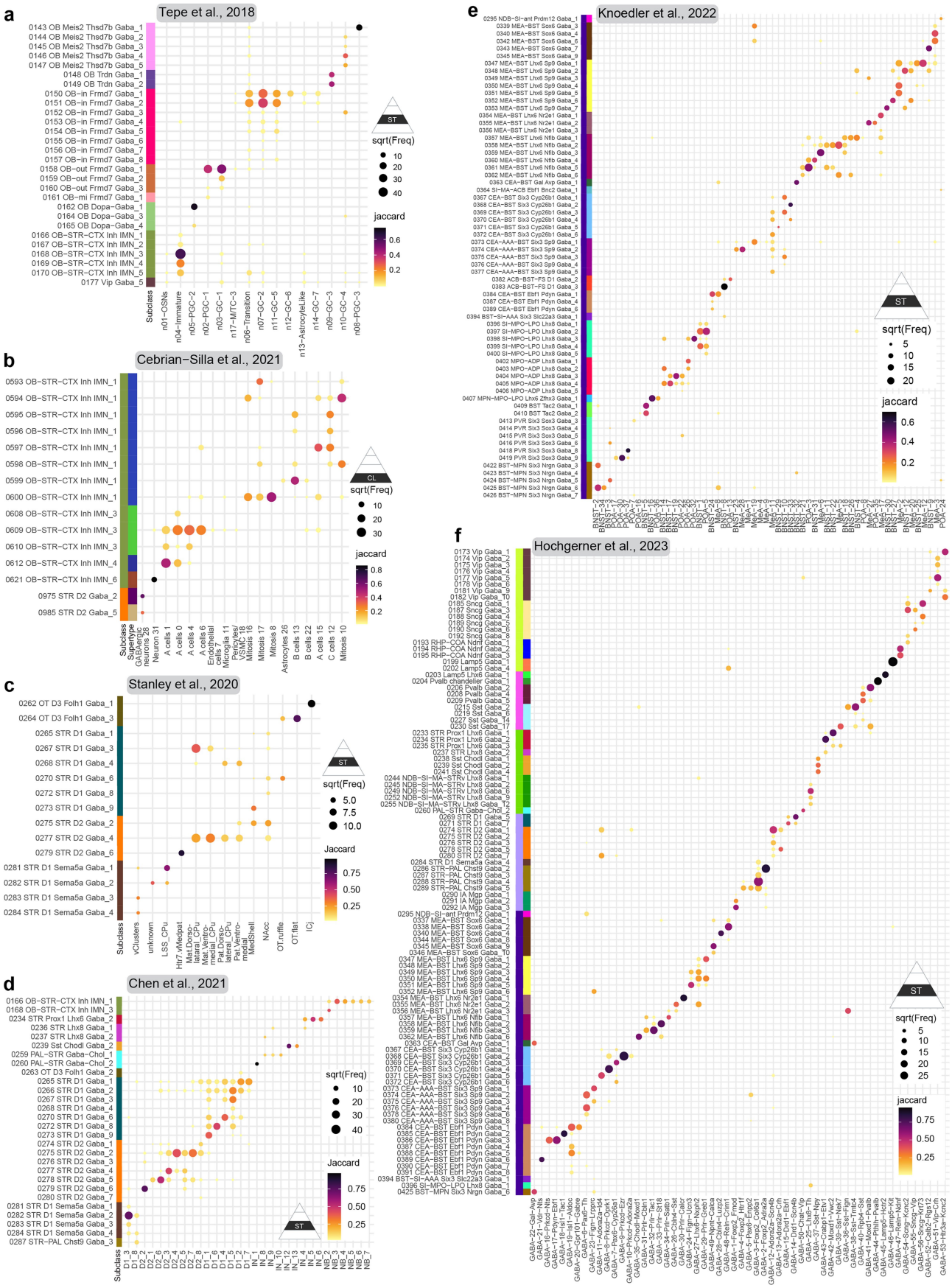
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | GABAergic neuronal type composition in different regions of the telencephalon. (a) Heatmap showing the proportion of cells in each broad region of the telencephalon per GABAergic supertype. Triangular schematic denotes the most granular hierarchical level shown in the panel (ST: supertype). (**b**) Heatmap showing the proportion of cells in each supertype from 06 CTX-CGE GABA and 07 CTX-MGE GABA classes in each cortical layer and substructure of the hippocampal formation. **(c-d)** Constellation plots of the global relatedness between subclasses (c) or supertypes (d). Each subclass or supertype is represented by a disk, labelled by the subclass or supertype ID, and positioned at the subclass or supertype centroid in UMAP coordinates shown in Fig. 1b or c. The size of the disk corresponds to the number of cells within each subclass or supertype, and the edge weights correspond to the fraction of shared neighbours (see Methods) between subclasses or supertypes. In d, each supertype disk is coloured by the subclass it belongs to. Bubbles drawn around subclasses or supertypes outline the major classes. Triangular schematics denote the specific hierarchical level depicted in each panel (SC: subclass; ST: supertype).



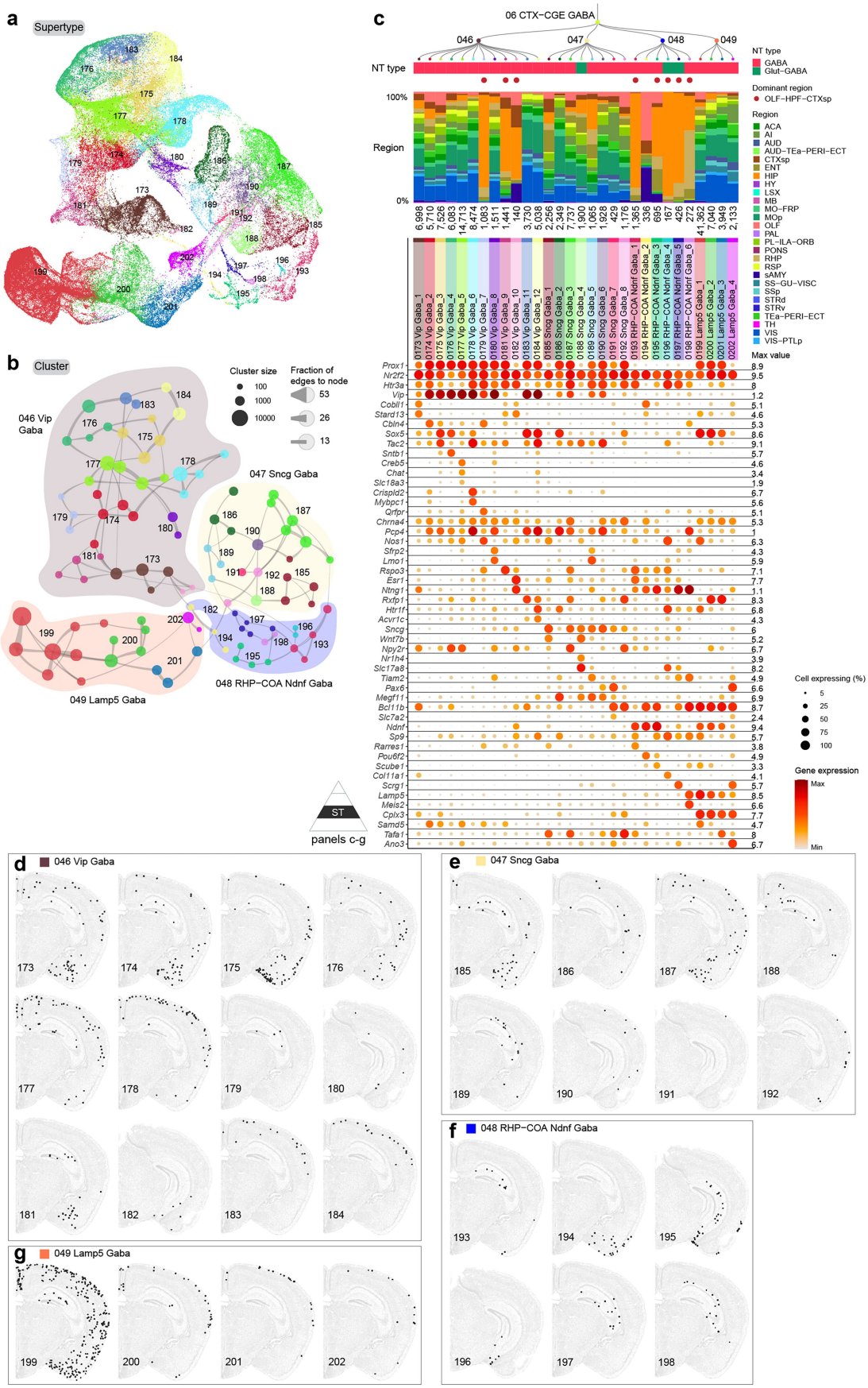
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Olfactory bulb GABAergic and immature neuronal cell types. **(a)** UMAP representation of OB GABAergic and immature neuronal types, coloured and labelled by supertype. **(b)** Constellation plot of OB-IMN GABA clusters using UMAP coordinates shown in a. Nodes are clusters coloured by supertype and grouped in bubbles by subclass. **(c)** Dendrogram of OB-IMN GABA supertypes followed by bar graphs showing major neurotransmitter type and region distribution of profiled cells, followed by dot plot showing marker gene expression in each supertype from the 10xv3 dataset. Dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively. **(d)** Schematic drawing of anatomical structure in MOB (left) and AOB (right). Abbreviations: RMS, rostral migratory stream; GrO, granular layer; IPl, internal plexiform layer; Mi, mitral layer; EPl, external plexiform layer; Gl, glomerular layer. **(e-k)** Representative MERFISH sections showing the location of OB subclasses 39 OB Meis2 Thsd7bGaba (e), 40 OB Trdn Gaba (f), 42 OB-out Frmd7 Gaba (g), 43 OB-mi Frmd7 Gaba (h), 44 OB Dopa-Gaba (i), 41OB-in Frmd7 Gaba (j), and 45 OB-STR-CTX Inh IMN (k). Cells are coloured and labelled by supertype. Triangular schematic denotes the most granular hierarchical level shown in the panels (ST: supertype).



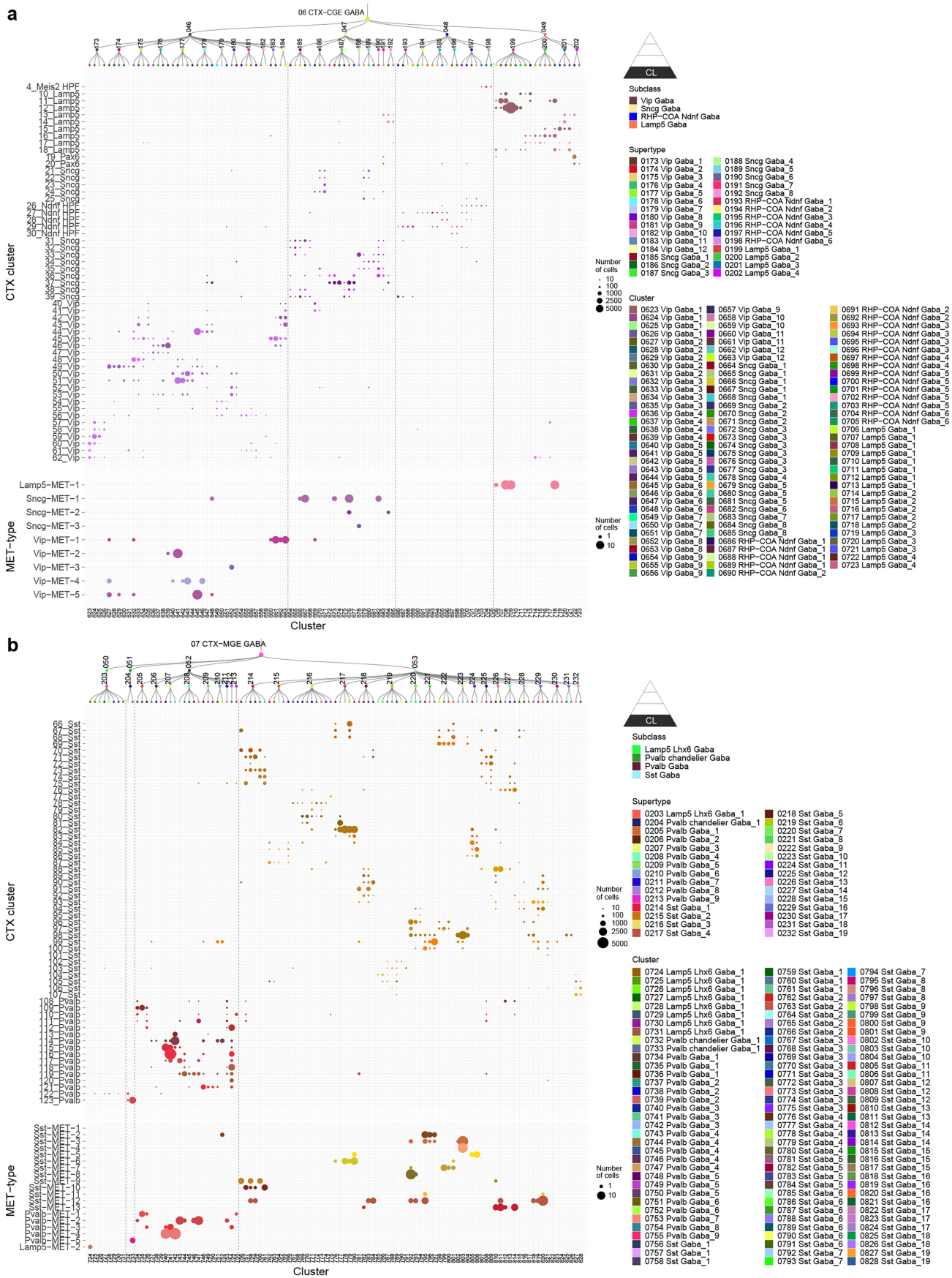
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Correspondence between the current transcriptomic taxonomy of OB-IMN GABA,CNU-LGE GABA, and CNU-HYa GABA classes and previously published ones. Correspondence was determined by mapping cells from previously published datasets to the current taxonomy as described before[22]. **(a,b)** Mapping of Tepe et al. [20] (a) and Cebrian-Silla et al. [18] (b) to the OB-IMN GABA class. **(c,d)** Mapping of Stanley et al. [47] (c) and Chen et al. [38] (d) to the CNU-LGE GABA class. **(e,f)** Mapping of Knoedler et al. [46] (e) and Hochgerner et al., 2023 (f) to supertypes in the CNU-HYa class or the entire telencephalic GABAergic taxonomy. Triangular schematics denote the specific hierarchical level shown in each panel (ST: supertype; CL: cluster).



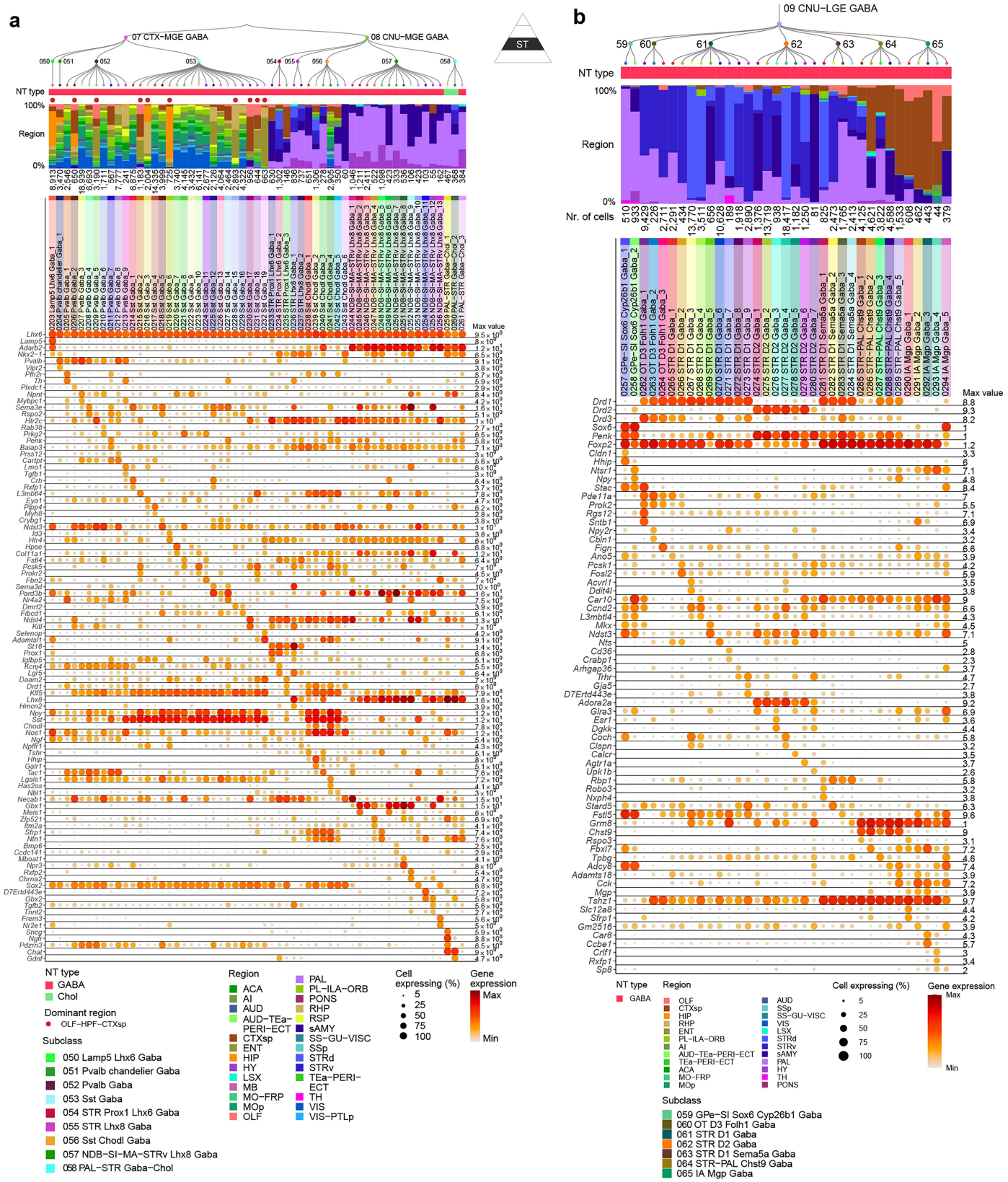
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | CGE-derived GABAergic neuronal types in the cerebral cortex. (a) UMAP representation of all CGE clusters coloured and labelled by supertype. **(b)** Constellation plot of CGE clusters using UMAP coordinates shown in a. Cluster nodes are coloured by supertype and grouped in bubbles by subclass. **(c)** Dendrogram of CGE supertypes followed by bar graphs showing major neurotransmitter type, region distribution of profiled cells, dominant region (red dot), and number of cells within supertype, followed by dot plot showing marker gene expression in each supertype from the 10xv3 dataset. The dominant region was assigned if more than 70% of cells are from the assigned region. For the gene expression dot plot, dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively. **(d-g)** Representative MERFISH sections showing the location of supertypes in CGE subclasses 46 Vip Gaba (d), 47 Sncg Gaba (e), 48 RHP-COA Ndnf Gaba (f), and 49 Lamp5 Gaba (g). Cells are coloured and labelled by supertype. Triangular schematic denotes the most granular hierarchical level shown in the panels (ST: supertype).

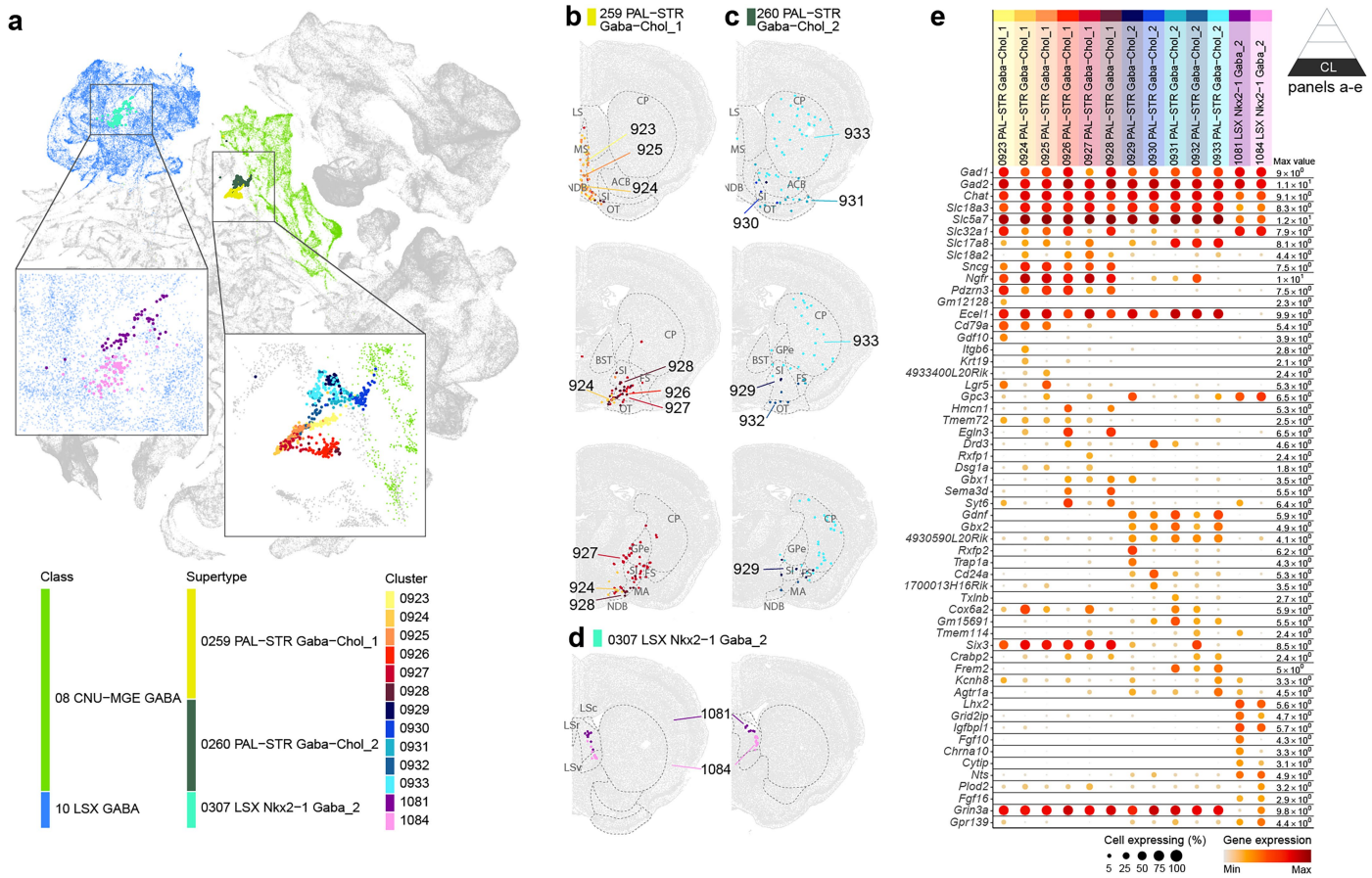


Extended Data Fig. 5 | See next page for caption.

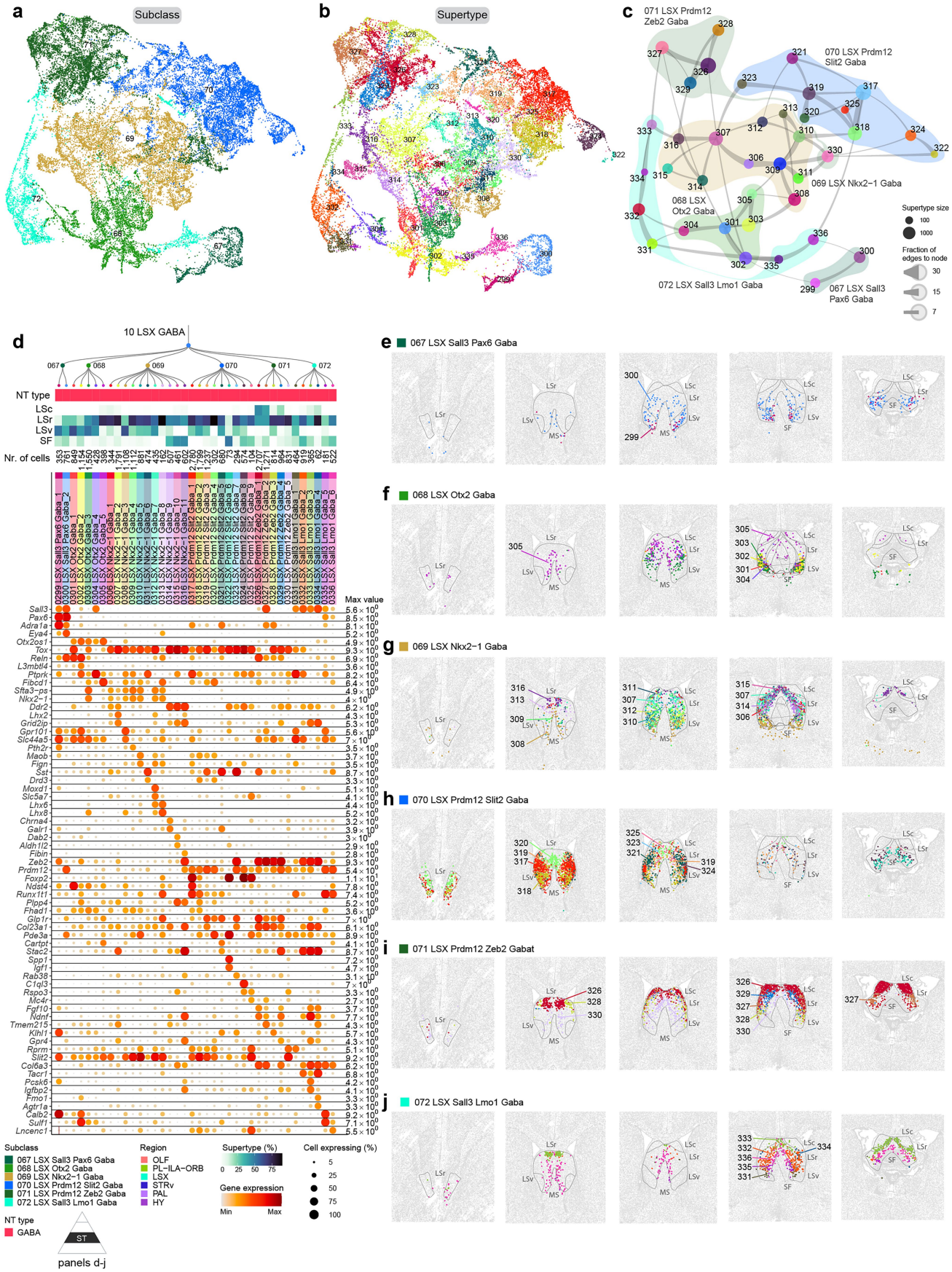
Extended Data Fig. 5 | Correspondence of CGE and MGE GABAergic neuronal types with previously published cell-type taxonomies. (**a-b**) CGE (a) and MGE (b) GABAergic cell types identified in this study are compared to cell types in CTX-HPF study[22] and VISp Patch-seq study[23]. The size of the dots corresponds to the number of overlapping cells in corresponding taxonomies. Columns are separated by supertypes, and rows are separated manually based on subclass in the corresponding dataset. Triangular schematics denote the most granular hierarchical level shown in the panels (CL: cluster).



Extended Data Fig. 6 | Subclass and supertype marker genes within CTX-MGE, CNU-MGE, and CNU-LGE GABAergic neuronal types. (a-b) Dendrogram of MGE (a) and CNU-LGE (b) supertypes followed by bar graphs showing major neurotransmitter type, region distribution of profiled cells, dominant region, and number of cells within supertype, followed by dot plot showing marker gene expression in each supertype from the 10xv3 dataset. The dominant region was assigned if more than 70% of cells are from the OLF-HPF-CTXsp regions. For the gene expression dot plot, dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively. Triangular schematic denotes the most granular hierarchical level shown in the panel (ST: supertype).

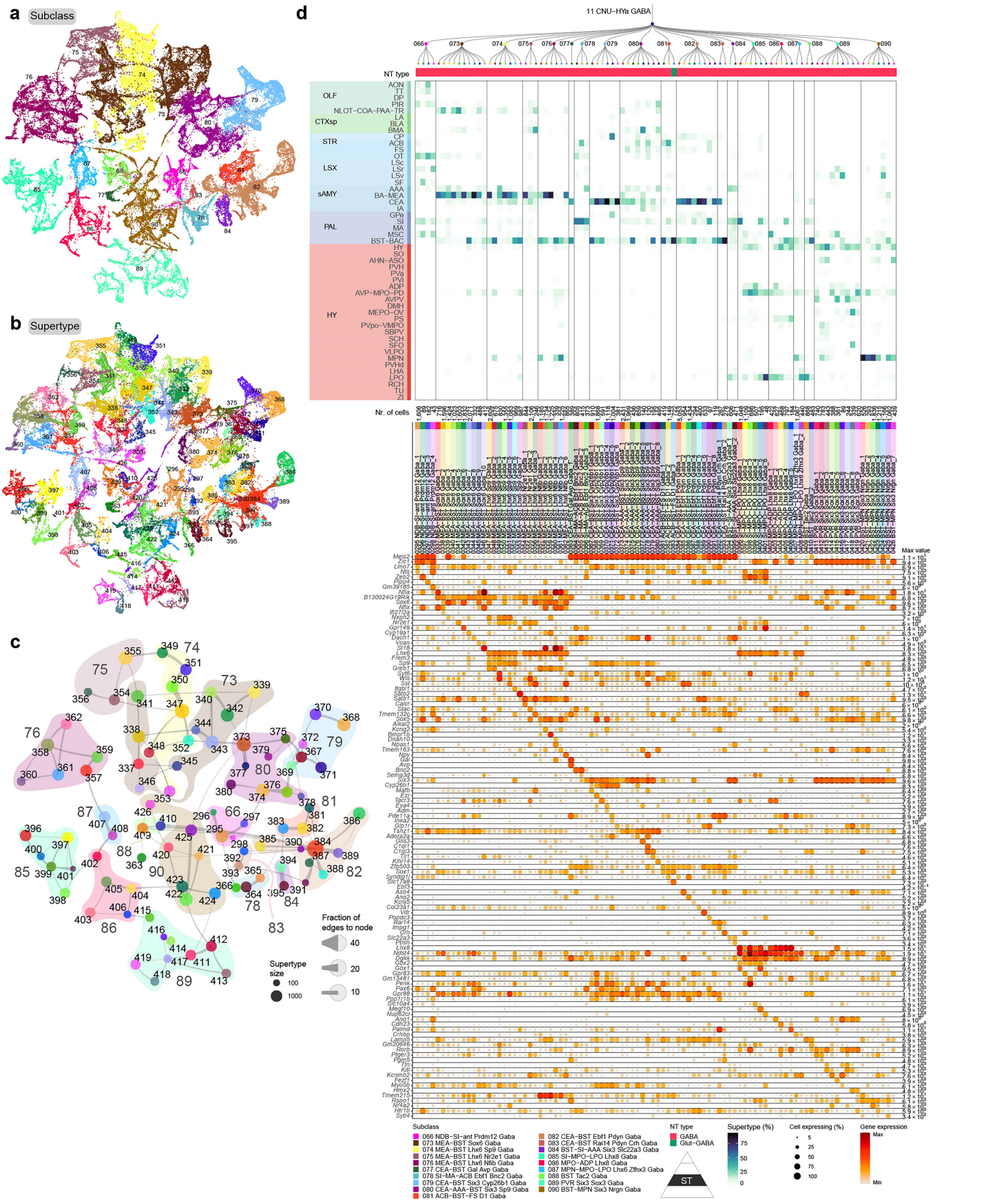


Extended Data Fig. 7 | Distribution of telencephalic cholinergic neurons. **(a)** UMAP representation of all telencephalic GABAergic neurons as in Fig. 1b–d, cells in background are coloured classes that contain cholinergic neurons and foreground cells are coloured by supertype. Inserts show the cholinergic neurons belonging to the CNU-MGE GABA and LSX GABA classes, respectively, coloured by cluster. **(b-d)** Representative MERFISH sections showing cholinergic neurons coloured by cluster for supertypes 259 PAL-STR Gaba-Chol_1 (b), 260 PAL-STR Gaba-Chol_2 (c), and 307 LSX Nkx2-1 Gaba_2 (d). **(e)** Dot plot showing marker gene expression in each cholinergic cluster from the 10xv3 dataset. Dot size and colour indicate proportion of expressing cells and average expression level in each cluster, respectively. Triangular schematic denotes the most granular hierarchical level shown in the panels (CL: cluster).



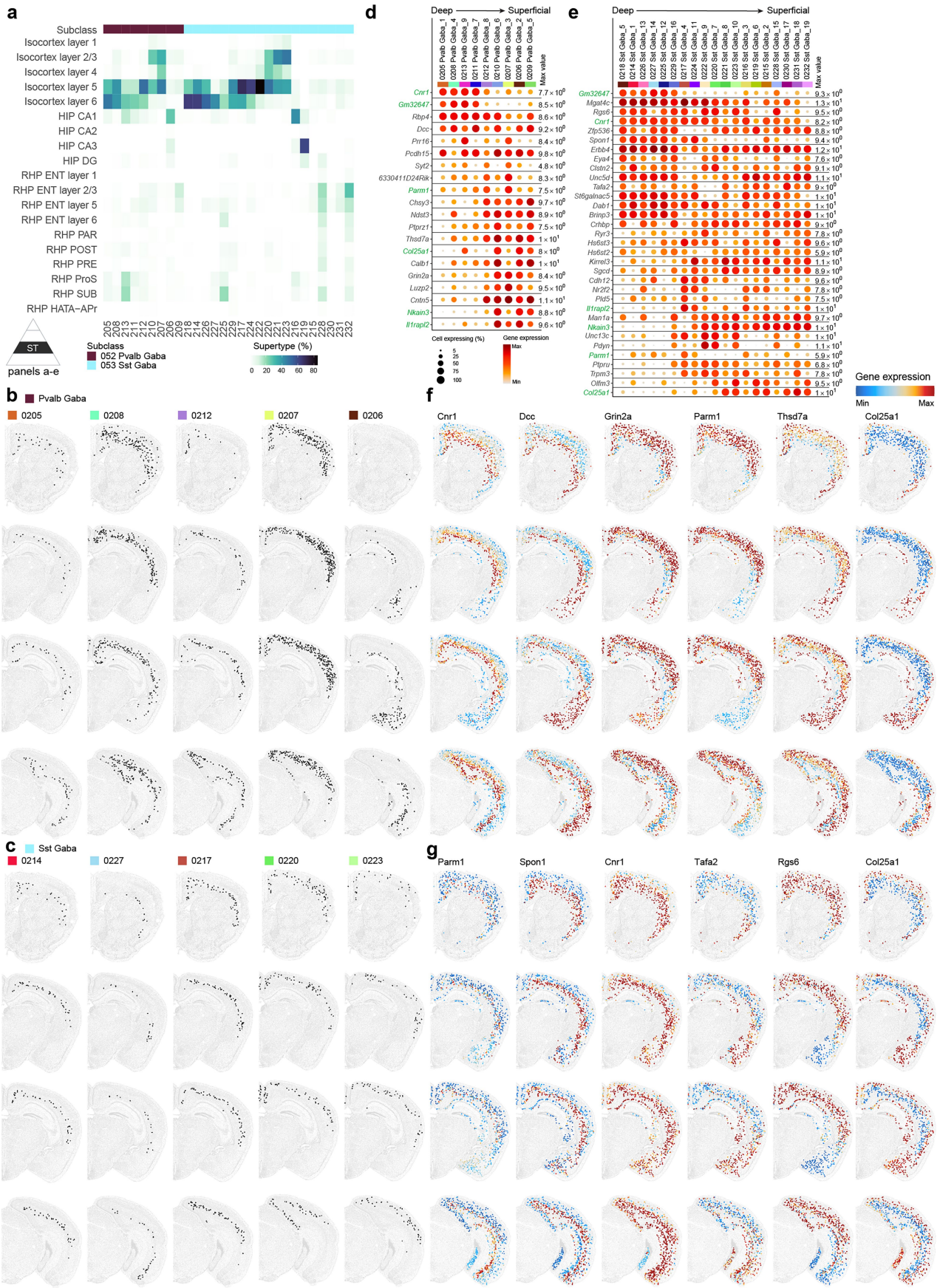
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | GABAergic cell types of the lateral septum. (a-b) UMAP representation of all LSX cells coloured by subclass (a) or supertype (b). **(c)** Constellation plot of LSX supertypes using UMAP coordinates shown in b. Nodes are coloured by supertype and grouped in bubbles by subclass. **(d)** Dendrogram of LSX supertypes followed by tiles showing major neurotransmitter type, followed by a heatmap showing the region distribution of profiled cells, then followed by dot plot showing marker gene expression in each supertype from the 10xv3 dataset. Dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively. **(e-j)** Representative MERFISH sections showing the location of supertypes in LSX subclasses 67 LSX Sall3 Pax6 Gaba (e), 68 LSX Otx2 Gaba (f), 69 LSX Nkx2-1 Gaba (g), 70 LSX Prdm12 ve Gaba (h), 71 LSX Prdm12 do Gaba (i), and 72 LSX Sall3 Lmo1 Gaba (j). Triangular schematic denotes the most granular hierarchical level shown in the panels (ST: supertype).



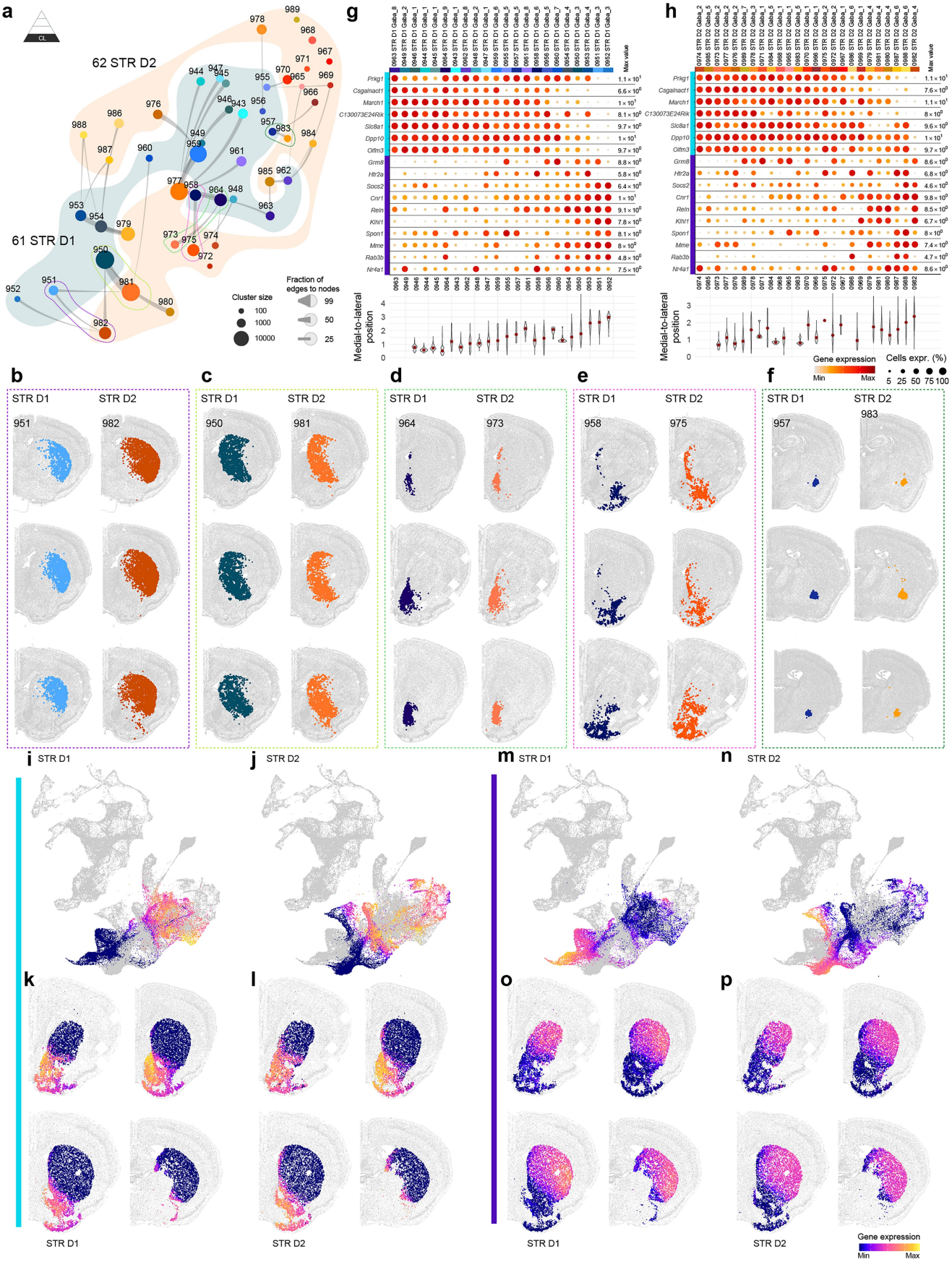
Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | GABAergic cell types of the CNU and anterior hypothalamus (HYa). (a-b) UMAP representation of all cells in the CNU-HYa GABA class coloured and labelled by subclass (a) or supertype (b). **(c)** Constellation plot of CNU-HYa supertypes using UMAP coordinates shown in b. Nodes are coloured by supertype and grouped in bubbles by subclass. **(d)** Dendrogram of CNU-HYa supertypes followed by tiles showing major neurotransmitter type, followed by a heatmap showing the region distribution of profiled cells, then followed by dot plot showing marker gene expression in each supertype from the 10xv3 dataset. Dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively. Triangular schematic denotes the most granular hierarchical level shown in the panels (ST: supertype).



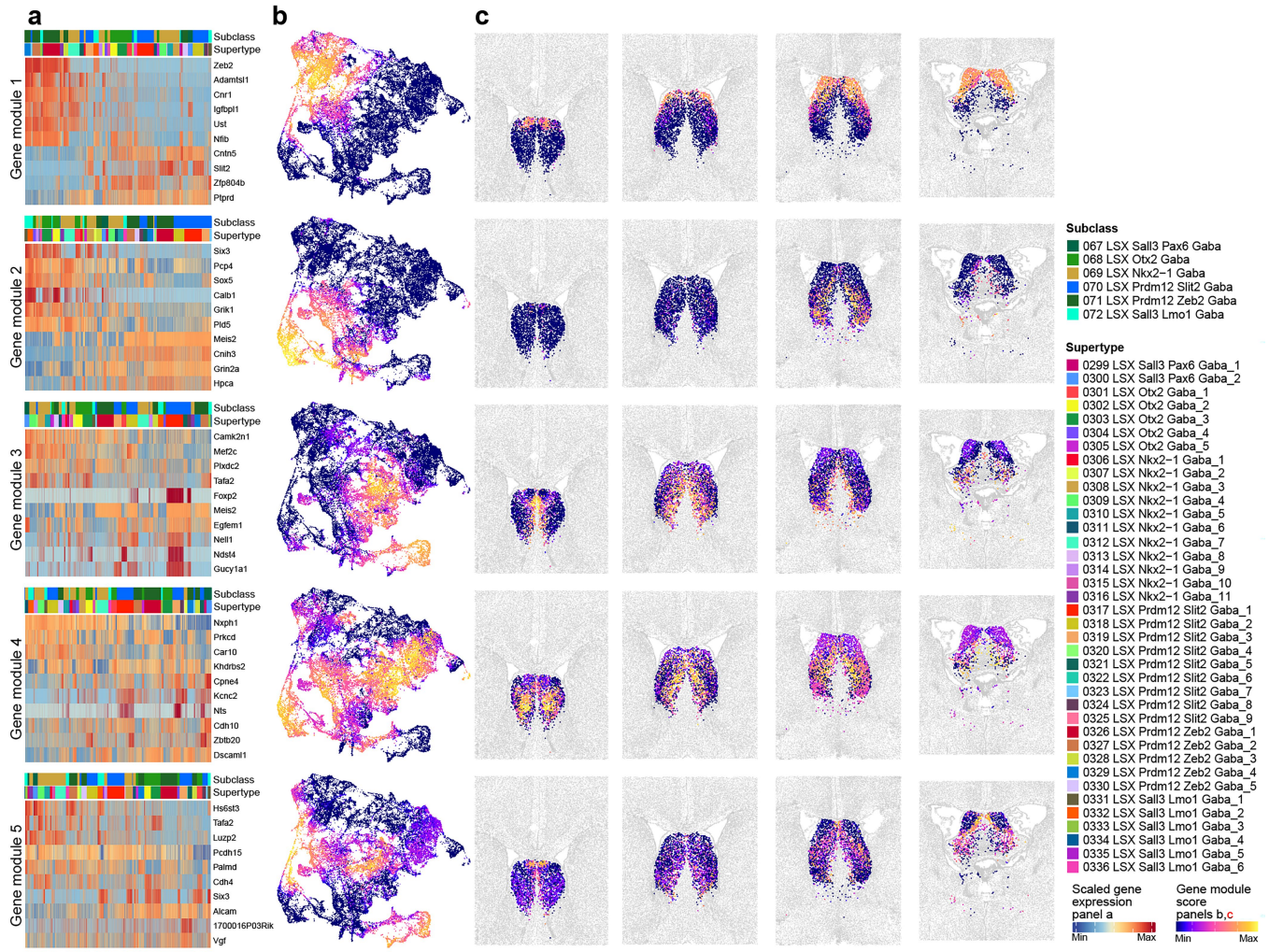
Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Laminar distribution and gene expression gradients of MGE GABAergic neurons in cortex and hippocampal formation. (**a**) Heatmap showing the proportion of cells in each layer or region of the isocortex and HPF for supertypes in the 52 Pvalb Gaba and 53 Sst Gaba subclasses. (**b,c**) Representative MERFISH sections showing the distribution of neurons across isocortex and HPF in select supertypes from subclasses 52 Pvalb Gaba (b) and 53 Sst Gaba (c). (**d,e**) Dot plots showing expression level of genes driving the gene expression gradient along the cortical depth for supertypes (ordered from deep to superficial) within the 52 Pvalb Gaba (d) and 53 Sst Gaba (e) subclasses. Dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively. Genes labelled in green are shared between the two subclasses. (**f, g**) Representative examples of genes that drive the laminar gene expression gradient, plotted on the same MERFISH section as in panels b and c, shown for the 52 Pvalb Gaba (f) and 53 Sst Gaba (g) subclasses. Each cell is coloured by relative gene expression level. Triangular schematic denotes the most granular hierarchical level shown in the panels (ST: supertype).

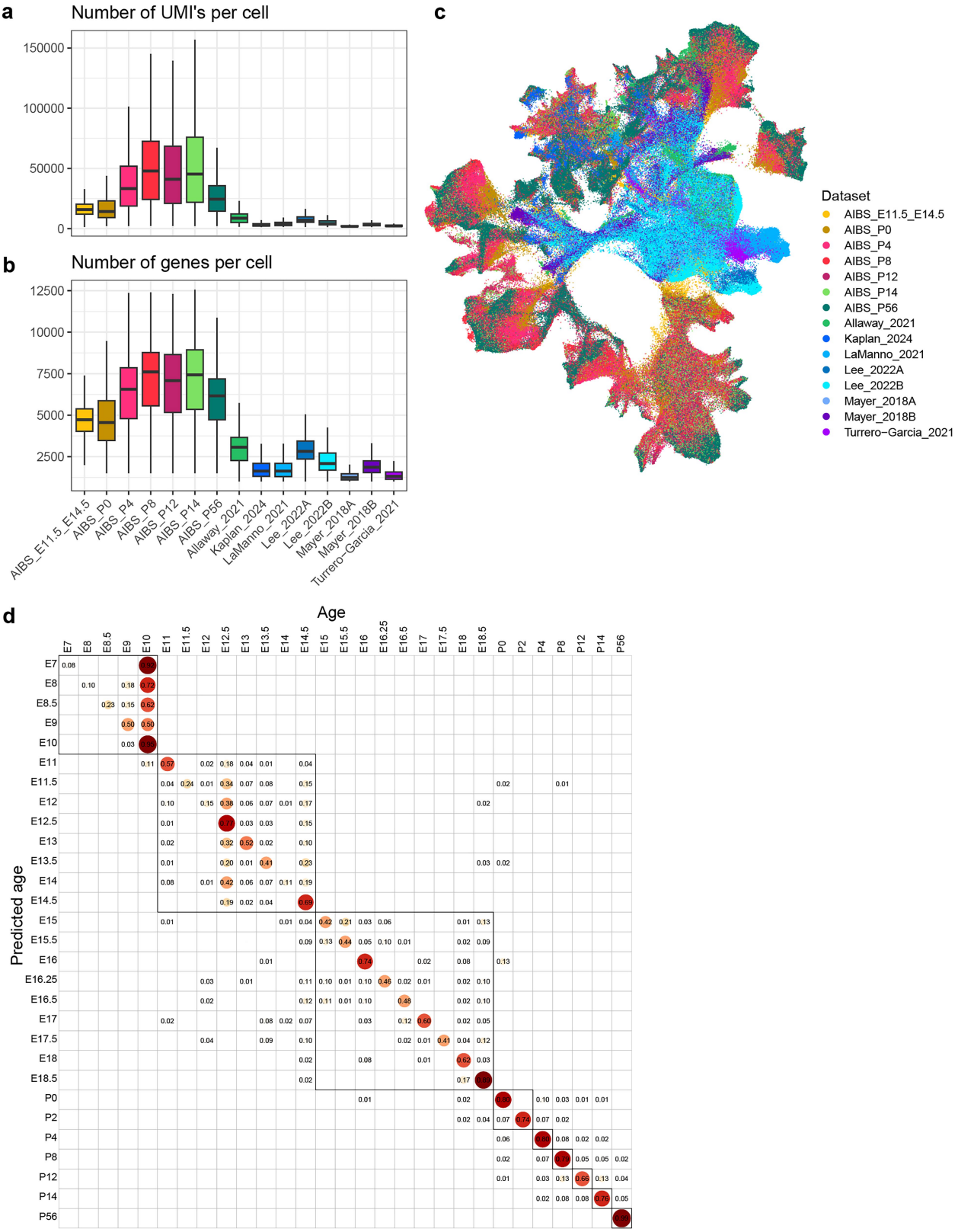


Extended Data Fig. 11 | See next page for caption.

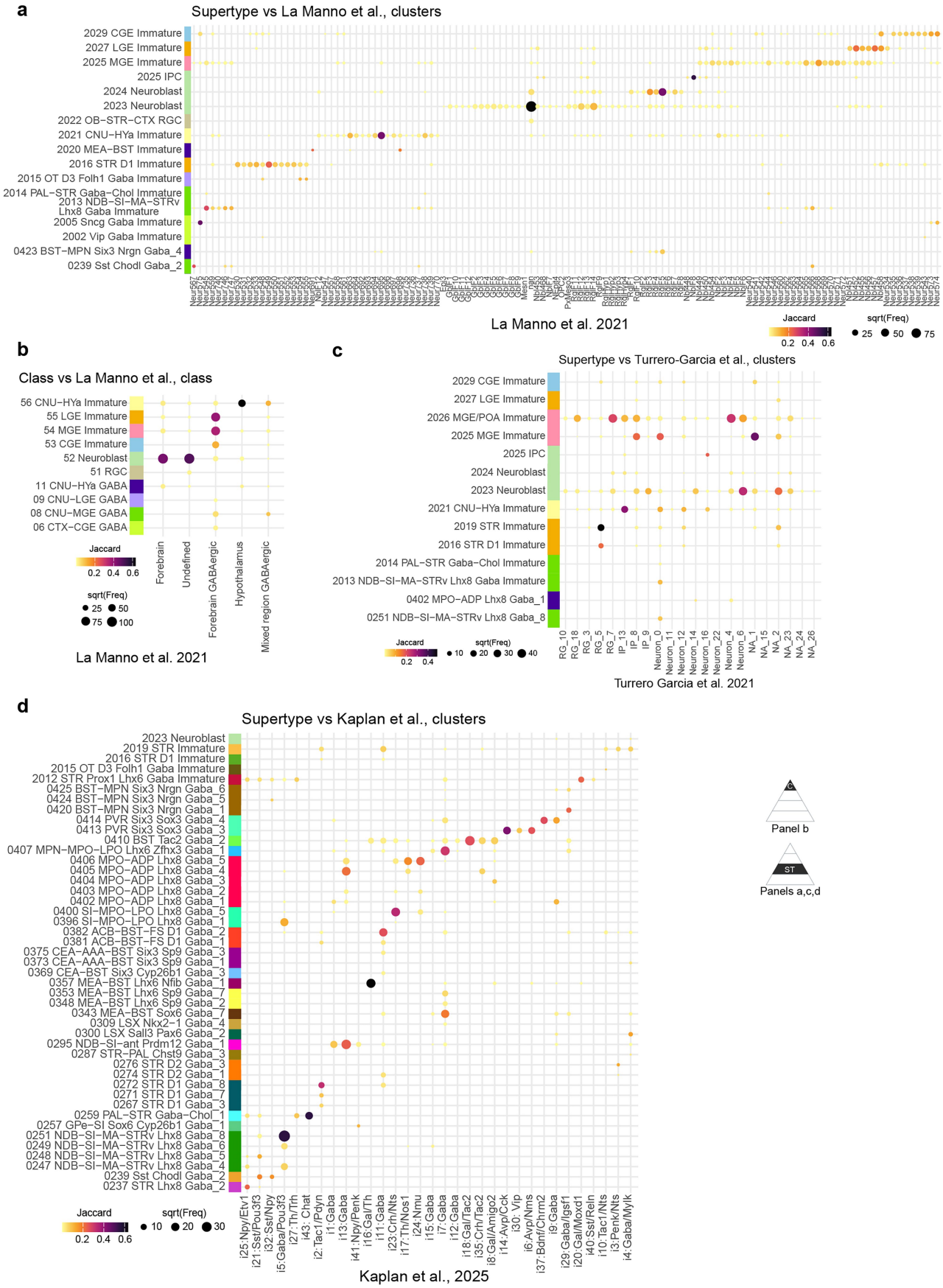
Extended Data Fig. 11 | Gene signatures defining shared gradients in D1 and D2 medium spiny neurons. (a) Constellation plot of clusters showing the pairs of similar clusters between 61 STR D1 Gaba and 62 STR D2 Gaba subclasses. Clusters are represented by a disk coloured by cluster, labelled by cluster ID, and disks with a coloured border are highlighted exemplars in panels b to f. (**b-f**) Representative MERFISH sections showing five examples of STR D1 and STR D2 cluster pairs from panel a, and their spatial distribution patterns. Sections are coloured by cluster identity. (**g-h**) Gene expression dot plots of two major gene modules driving the spatial gradient among STR D1 (g) and STR D2 (h) clusters. Dot size and colour indicate proportion of expressing cells and average expression level in each cluster, respectively. Underneath the dot plot, a violin plot of the medial-lateral coordinate for each MERFISH cell per cluster is shown. **(i-p)** Based on the gene modules (light blue and dark blue) identified in panels g-h, a cumulative gene score was calculated using UCell. UMAPs showing CNU-LGE GABAergic neurons (i,j,m,n), and representative MERFISH sections (k,l,o,p), coloured by the light blue gene module score for STR D1 (i,k) and STR D2 (j,l) types, or coloured by the dark blue gene module score for STR D1 (m,o) and STR D2 (n,p) types. Triangular schematic denotes the most granular hierarchical level shown in the panels (CL: cluster).



Extended Data Fig. 12 | Spatial gradients in LSX. (a) Heatmaps showing expression of genes that drive the spatial gradients among subclasses in LSX in the top five gene modules. Each module contains genes that are both up- and down-regulated along the spatial gradient. **(b-c)** Gene signature score for these five modules is calculated for every cell and shown in the scRNA-seq UMAP representations (b) and representative MERFISH sections (c).

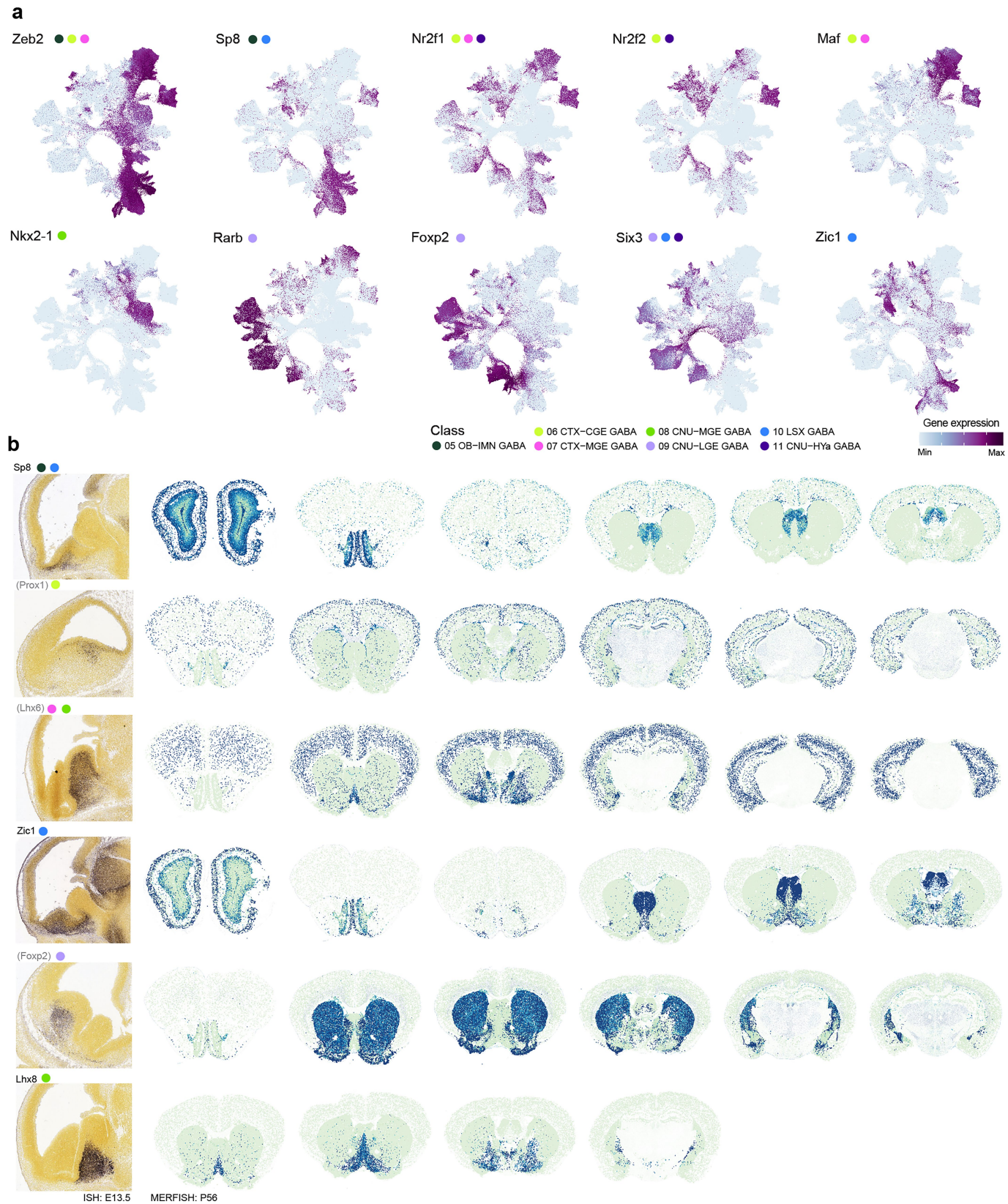


Extended Data Fig. 13 | Developmental datasets integrated with the P56 Subpallium GABAergic neuron taxonomy. (a,b) Comparison of UMI counts (a) and gene counts (b) per cell for the post-QC developmental and adult (10xv3) datasets. Boxplots show median, interquartile range and 1.5× interquartile range. Number of cells in each dataset: AIBS_E11.5_E14.5, n = 9,296; AIBS_P0, n = 54,247; AIBS_P4, n = 83,650; AIBS_P8, n = 94,535; AIBS_P12, n = 68,636; AIBS_P14, n = 76,784; AIBS_P56, n = 341,644; Allaway_2021, n = 18,131; Kaplan_2024, n = 26,344; LaManno_2021, n = 39,772; Lee_2022A, n = 18,200; Lee_2022B, n = 72,589; Mayer_2018A, n = 7,777; Mayer_2018B, n = 27,472; Turrero-Garcia_2021, n = 17,136. **(c)** UMAP representation of all cells across embryonic to P56 timepoints following scVI integration, coloured by dataset. **(d)** Confusion matrix between chronological age and predicted age (see Methods).

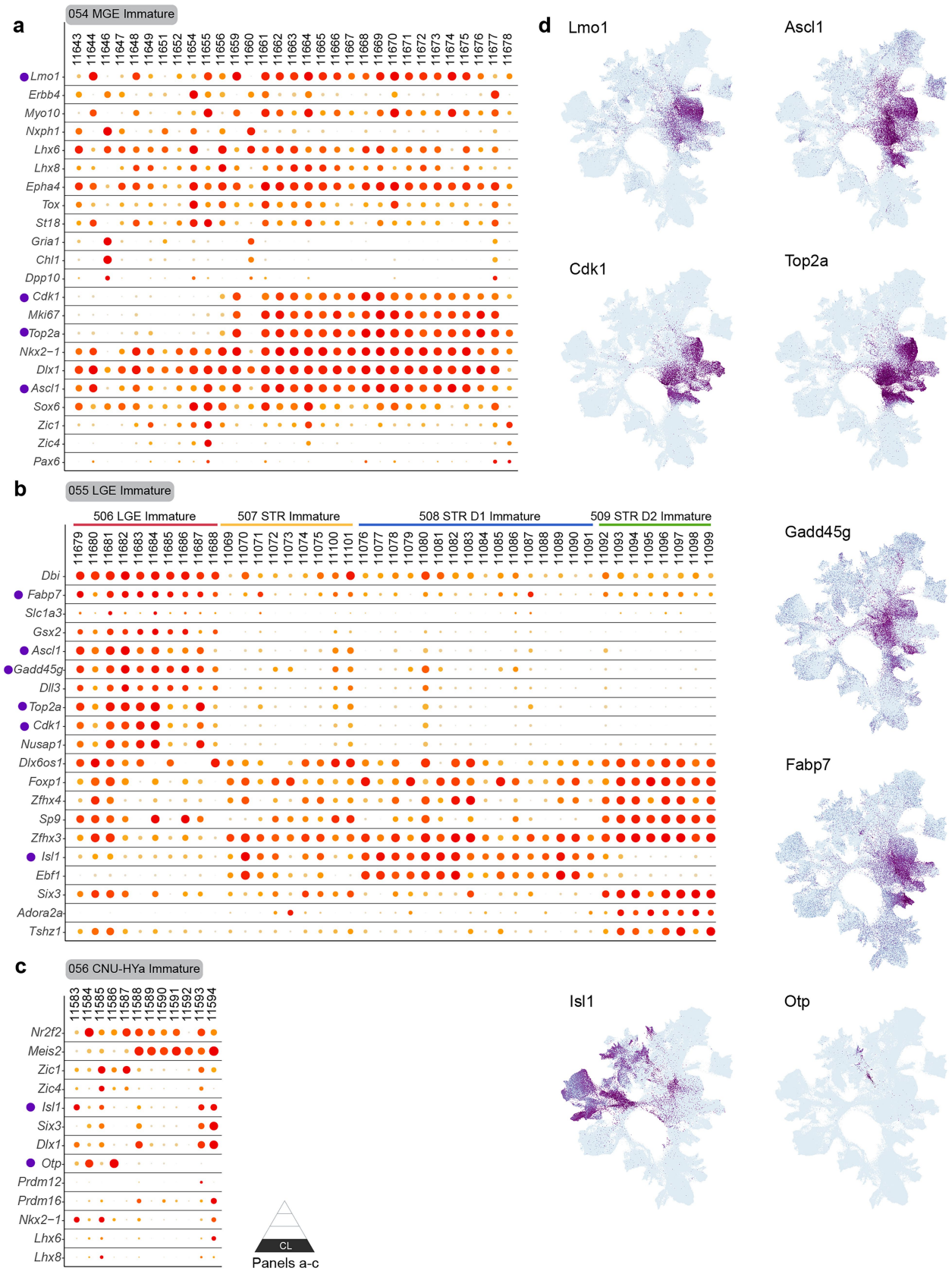


Extended Data Fig. 14 | See next page for caption.

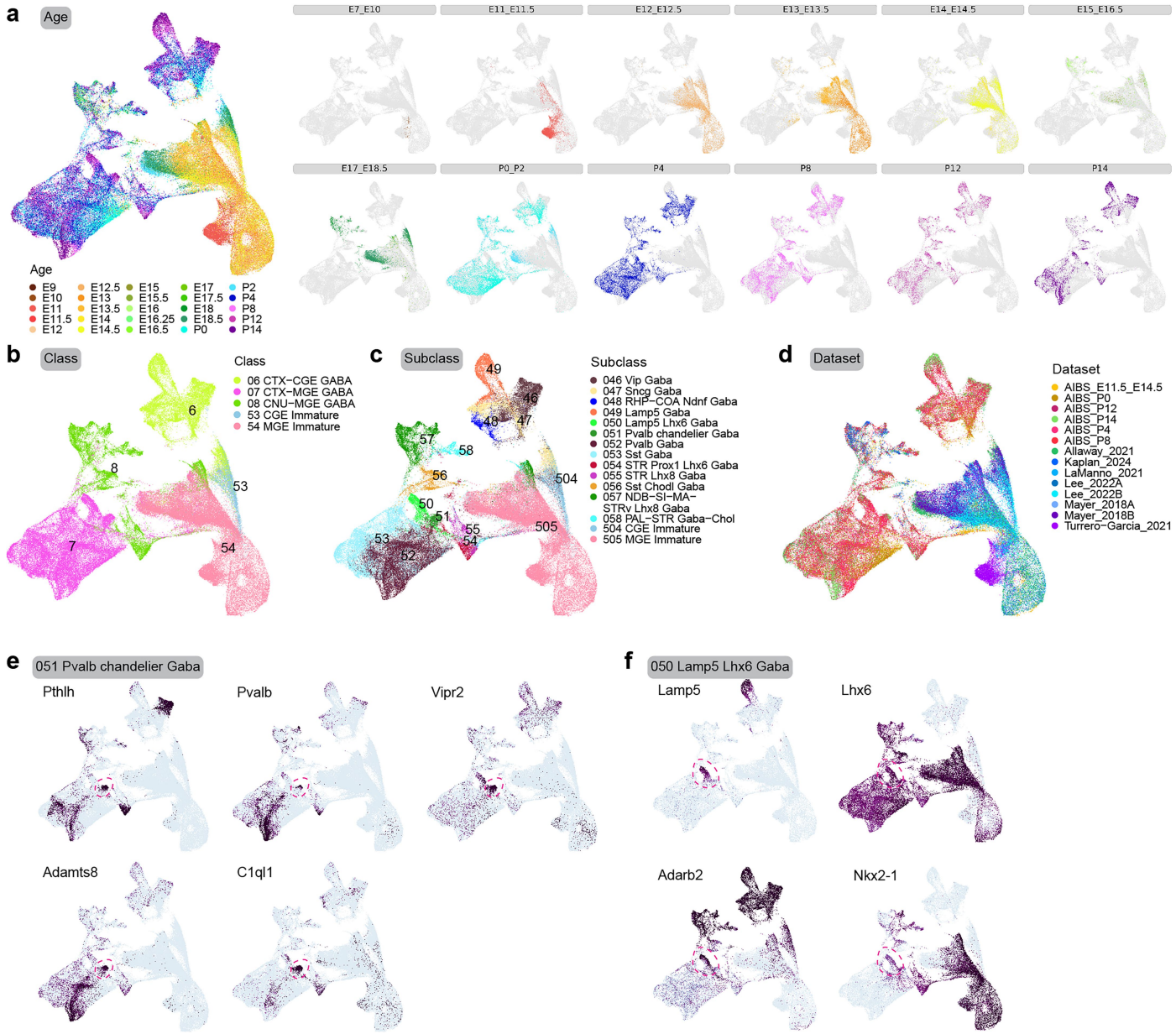
Extended Data Fig. 14 | Comparison of cell type annotations between the subpallium GABAergic neuronal type taxonomy and published developmental studies. (a-d) Correspondence of cell type annotations from published studies included in the developmental dataset and the current subpallium GABAergic neuronal type taxonomy. Cell type annotations from the current taxonomy are compared to the La Manno et al. dataset at supertype-to-cluster level (a) and class-to-class level (b). The current taxonomy is compared to the Turrero-Garcia et al. dataset at supertype-to-cluster level (c) and compared to the Kaplan et al. dataset at supertype-to-cluster level (d). Triangular schematics denote the specific hierarchical level shown in the panels (C: class, ST: supertype).



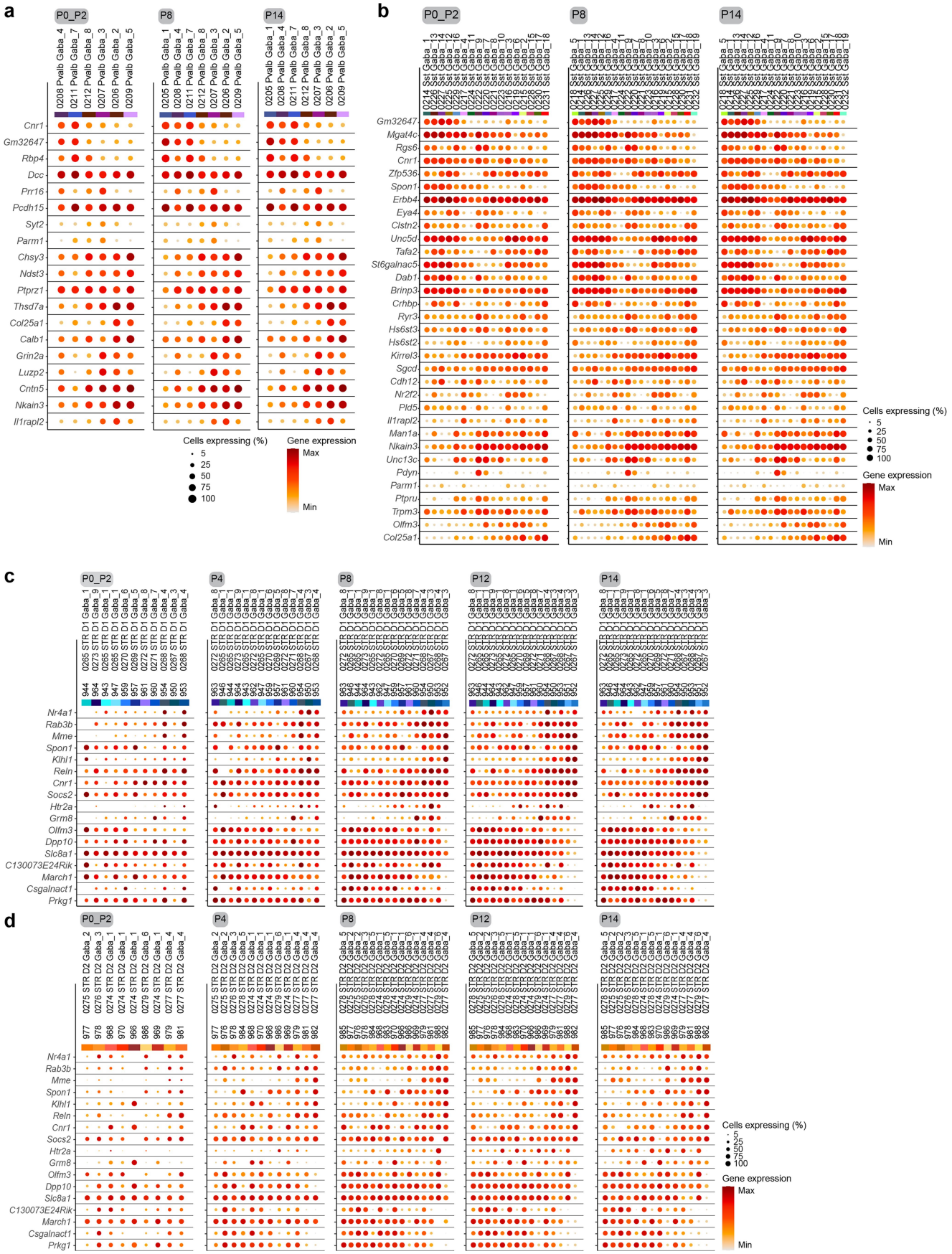
Extended Data Fig. 15 | Transcription factor expression marking developmental lineages. (a) UMAP representation of all cells across embryonic to P56 timepoints coloured by expression level of major lineage marker genes. The coloured dots next to each gene name show the adult cell classes that express that gene. **(b)** Representative developmental ISH (E13.5) and adult MERFISH sections showing expression of key TFs in GABAergic neurons. Genes marked with parentheses indicate imputed spatial gene expression pattern on MERFISH sections (see Methods).



Extended Data Fig. 16 | Gene signatures of immature neuronal populations. **(a-c)** Dot plots showing gene expression signatures for clusters in the 54 MGE Immature class (a), 55 LGE Immature class (b), and 56 CNU-HYa Immature class (c). Genes marked with purple dots are shown on a UMAP representation of all cells across embryonic to P56 timepoints coloured by expression level **(d)**. Triangular schematic denotes the specific hierarchical level shown in the panels (CL: cluster).

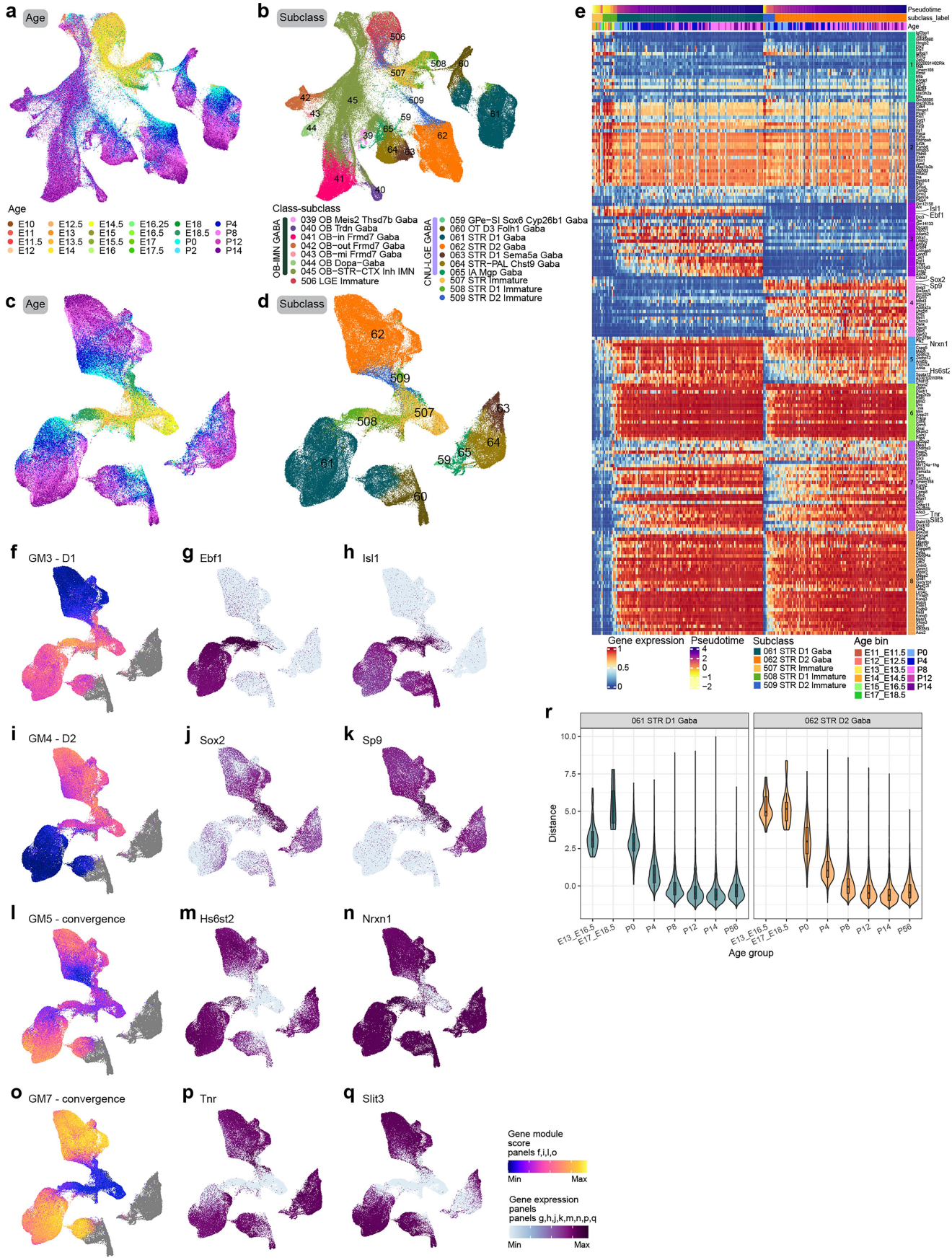


Extended Data Fig. 17 | Developmental trajectory of CGE- and MGE-derived neurons. (a-d) UMAP representation of all neurons from embryonic ages to P14 that will form the 06 CTX-CGE GABA, 07 CTX-MGE GABA, and 08 CNU-MGE GABA classes, coloured by age (a), class (b), subclass (c), or dataset (d). UMAPs on the right side of panel a show cells in each age bin individually. **(e)** UMAPs showing expression of genes that mark the 51 Pvalb chandelier Gaba subclass (dashed oval). **(f)** UMAPs showing expression of genes that link the 50 Lamp5 Lhx6 Gaba subclass (dashed oval) to both MGE and CGE origins.



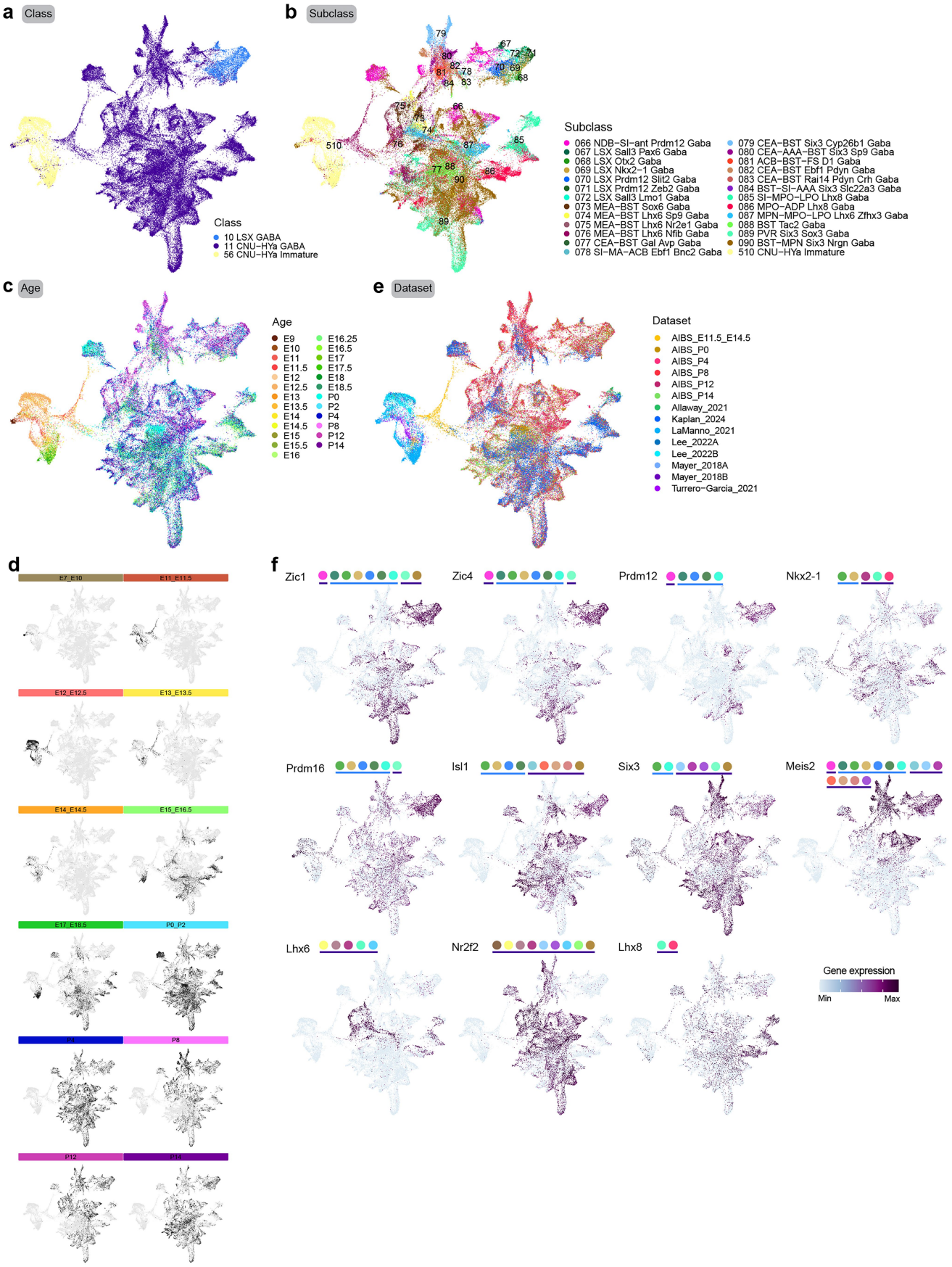
Extended Data Fig. 18 | See next page for caption.

Extended Data Fig. 18 | Gene expression gradient formation during development. (**a,b**) Dot plots showing expression level of genes driving the gene expression gradient along the cortical depth identified in the P56 data within the Pvalb Gaba (a) and Sst Gaba (b) subclasses for ages P0-P2, P8, and P14. The genes and supertypes are ordered exactly as shown in Extended Data Fig. 10d, e. (**c,d**) Gene expression dot plots of two major gene modules driving the spatial gradient among STR D1 (c) and STR D2 (d) clusters identified in the P56 data for ages P0-P2, P4, P8, P12 and P14. The genes and supertypes are ordered exactly as shown in Extended Data Fig. 11g, h. Dot size and colour indicate proportion of expressing cells and average expression level in each supertype, respectively.



Extended Data Fig. 19 | See next page for caption.

Extended Data Fig. 19 | Developmental trajectory of LGE-derived neurons. **(a-b)** UMAP representation of all neurons from embryonic ages to P14 that are derived from LGE and will populate the OB-IMN GABA and CNU-LGE GABA classes. UMAPs are coloured by age (a) or subclass (b). **(c-d)** UMAP representation of all neurons from embryonic ages to P14 that will form the CNU-LGE GABA class. UMAPs are coloured by age (c) or subclass (d). **(e)** Heatmap showing differentially expressed genes in the STR D1 and STR D2 trajectories across time. Eight gene modules were identified that show various modes of expression along and between subclasses over time. **(f,i)** Gene module scores marking different stages along the maturation path of D1 and D2 neurons. Gene modules (GM) 3 and 4 highlight STR D1 (f) and STR D2 (i) neurons respectively. **(g,h,j,k)** UMAP representation like in panels c-d coloured by major trajectory markers. **(l-n)** Gene module 5 contains genes highlighting the delayed maturation of STR D2 vs STR D1 neurons (l), such as two exemplar genes *Hs6st2* (m) and *Nrxn1* (n). **(o-q)** Gene module 7 contains genes whose expressions converge along the maturation trajectory of STR D1 and STR D2 neurons (o), such as two exemplar genes *Tnr* (p) and *Slit3* (q). **(r)** Violin plot showing the transcriptomic distance between STR D1 and STR D2 subclass transcriptomes across the time course. Left panel: *n* for increasing age in the 061 STR D1 Gaba subclass: 120, 76, 4242, 7873, 8559, 12117, 35578, 14536 edges between 061 STR D1 Gaba and closest 062 STR D2 Gaba clusters. Right panel: *n* for increasing age in the 062 STR D2 Gaba subclass: 45, 52, 2604, 7224, 7700, 11195, 36193, 13549 edges between 062 STR D2 Gaba and closest 061 STR D1 Gaba clusters. Boxplots within violins show median, interquartile range and 1.5× interquartile range.



Extended Data Fig. 20 | See next page for caption.

Extended Data Fig. 20 | Developmental trajectory of LSX and CNU-HYa GABAergic neurons. (a-e) UMAP representation of all neurons from embryonic ages to P14 that will form the 10 LSX GABA and 11 CNU-HYa GABA classes, coloured by class (a), subclass (b), age (c,d), and dataset (e). UMAPs in panel d show cells in each age bin individually. **(f)** UMAP representation as in panels a-e coloured by expression level of subclass marker genes. The coloured dots next to each gene name show the subclasses expressing that gene.

nature portfolio

Corresponding author(s): Cindy van Velthoven & Hongkui Zeng

Last updated by author(s): Jun 5, 2025

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	scRNA-seq data was collected using the 10x Genomics Chromium Single Cell 3' v3, v3.1 and v4 kits. Sequencing was performed on the Illumina HiSeq2500 and NovaSeq6000. MERFISH data were acquired using the Vizgen MERSCOPE platform.

Data analysis	Most of the methods used to analyze the data have been described before (Yao et al., 2023) and the following methods are either newly introduced or modified version for this manuscript. We performed PCA (R package stats, v4.4.1, RRID:SCR_025968, https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html). Followed by creation of UMAPs (v0.5.6, RRID:SCR_018217, https://github.com/lmcinnes/umap) Constellation plots were generated using scrattch.bigcat (v0.0.5, https://github.com/AllenInstitute/scrattch.bigcat). We performed independent component analysis using fastICA (v1.2-5.1, RRID:SCR_013110, https://cran.r-project.org/web/packages/fastICA/). We applied UCell (v2.8.0, DOI: 10.18129/B9.bioc.UCell) to assign a "gene module score" to each cell. BANKSY was used to perform spatial domain clustering on MERFISH data (v0.99.12, https://github.com/prabhakarlab/Banksy) We performed mapping of cells from each adult external dataset to the 10xv3 whole-brain dataset using treeMap function from scrattch.mapping package (v0.55, https://github.com/AllenInstitute/scrattch-mapping). Raw scRNA-seq fastq files were processed with the CellRanger v6.1.1 pipeline using default parameters. Doublets were identified using a modified version of the DoubletFinder algorithm which is available in scrattch.hicat (v0.1.0, RRID:SCR_01809, https://github.com/AllenInstitute/scrattch.hicat). Seurat was used for initial clustering of developmental data (v5.1.0, RRID:SCR_016341) For dimensionality reduction and batch correction, the raw counts of the highly variable genes were used to train the scvi-tools (v0.17.1, RRID:SCR_026673) scVI variational autoencoder.

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

Data

Policy information about availability of data

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

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

Sequencing reads were aligned to the mouse reference transcriptome (M21, GRCm38.p6)

The scRNA-seq and MERFISH datasets for this study are part of the Allen whole mouse brain (WMB) cell type atlas and are accessible through Neuroscience Multi-omic Data Archive (NeMO, RRID:SCR_016152) and Brain Image Library (BIL, RRID:SCR_017272). The adult 10x scRNA-seq datasets and the developmental scRNA-seq datasets are being made available through BRAIN Initiative Cell Atlas Network (BICAN, RRID:SCR_022794) (FASTQ files) and are available at NeMO under identifier https://assets.nemoarchive.org/dat-zqurqvh. The MERFISH dataset is available at BIL under DOI https://doi.org/10.35077/g.610.

The Subpallium-GABA cell type taxonomy is available from the Allen Brain Cell (ABC) Atlas (RRID:SCR_024440) to visualize both sc/snRNA-seq and MERFISH datasets. Instruction for access of the processed 10x scRNA-seq data is available at https://github.com/AllenInstitute/abc_atlas_access/blob/main/descriptions/WMB-10X.md, and instruction for access of the processed MERFISH data is available at https://github.com/AllenInstitute/abc_atlas_access/blob/main/descriptions/MERFISH-C57BL6J-638850.md.

The following published developmental datasets were used; Allaway et al. 2021, GSE164570; Kaplan et al. 2025, GSE280964; Lee et al. 2022, GSE167013 and GSE190593; Mayer et al. 2018, GSE103983 and GSE104156; La Manno et al. 2021, PRJNA637987; Turrero García et al. 2021, GSE184879.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	This study does not involve human research participants.
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

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

Field-specific reporting

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

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

Life sciences study design

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

Sample size	No statistical methods were used to predetermine sample sizes For scRNA-seq data, sample size was determined based on extensive prior experience using these technologies in brain (Zeisel et al. 2018, Yao et al 2021). Post-hoc analysis of clustering shows that it is robust to downsampling, which indicates that the sample size is sufficient to distinguish cell types as described in the manuscript. For MERFISH data we evenly sampled one mouse brain to capture most brain regions and get an even distribution of cells types throughout the brain.
Data exclusions	Cells from scRNA-seq data underwent stringent QC and cell quality was assessed based on gene detection, qc score, and doublet score. The qc score was calculated by summing the log transformed expression of a set of genes whose expression level is decreased significantly in poor quality cells. We used this qc score to quantify the integrity of cytoplasmic mRNA content, which tended to show bimodal distribution. Cells at the low-end were very similar to single nuclei, which we removed for downstream analysis. Doublets were identified using a modified version of the DoubletFinder algorithm (available in scrattch.hicat, https://github.com/AllenInstitute/scrattch.hicat, v1.0.9) and removed when doublet score > 0.3. Post-clustering we excluded noise clusters which are clusters with significantly lower gene detection due to extensive drop out, and clusters due to doublets or contamination. We first identified doublet clusters based on the co-expression of any pair of broad class marker genes using find_doublet_by_marker function in scrattch.bigcat package. To identify other doublet clusters, we searched for triplets of clusters A, B and C, wherein A was the putative doublet cluster, such that up-regulated genes of A relative to B largely overlapped with up-regulated genes in C relative to B, and up-regulated genes in A relative to C largely overlapped with up-regulated genes of B relative to C. After removing all doublet clusters, we then identified clusters with lower gene detection. To do that, we identified pairs of clusters such that one cluster with at least 50% fewer UMIs or >100 lower QC score, smaller size, and no more than one up-regulated gene relative to another cluster was identified as the low-quality cluster. For MERFISH data the cell-by-gene table containing segmented cells was filtered to keep cells with a volume > 100 µm3 and < 3,000 µm3, that have at least 15 genes detected and contain a minimum of 40 but no more than 3,000 mRNA molecules to remove low quality cells and doublets that are outside of these ranges.
Replication	Most of the scRNA-seq experiments were carried out at least twice independently and at least 2 mice and multiple brain dissections per mouse were used. All attempts at replication were successful. For MERFISH, 59 sections from one mouse brain were generated.
Randomization	Randomization of animals to different groups is not relevant to our study design. There were no experimental vs. control groups.
Blinding	Prior to clustering, single cell transcriptomes were analyzed for previously known marker genes and were segregated into large groups: non-neuronal, glutamatergic and GABAergic. Clustering was then performed blind to the cell source or any other metadata that could reveal sample identity.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	All animals used in this study were house mice (Mus musculus) maintained on the C57BL/6J background. Animals were euthanized at E11.5-E14.5 (n=6), P0 (n=14), P4 (n=12), P8 (n=16), P12 (n=12), P14 (n=14). For the developmental dataset, each animal's unique ID, sex, age, and genotype are listed in Supplementary Table 4. The donor information for the adult mice can be found in Yao et al., 2023 (https://doi.org/10.1038/s41586-023-06812-z) Supplemental Table 2. Mice had ad libitum access to food and water and were group-

	housed within a temperature- (21-22°C), humidity- (40-51%), and light- (14/10hr light/dark cycle) controlled room within the vivariums of the Allen Institute for Brain Science.
Wild animals	This study did not involve wild animals.
Reporting on sex	For ages P0 to P14, both male and female mice were used to collect scRNA-seq data. For the embryonic ages E11.5 to E14.5 we collected brains and performed post-hoc sex identification. Therefore, sexes are not fully balanced for the embryonic samples.
Field-collected samples	This study did not involve field-collected samples.
Ethics oversight	All experimental procedures related to the use of mice were approved by the Institutional Animal Care and Use Committee of the Allen Institute for Brain Science, in accordance with NIH guidelines.

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

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*