Developmental landscape of shark telencephalon sheds light on the evolution of telencephalic cell types and the ancient origin of Cajal-Retzius cells

Idoia Quintana-Urzainqui[1]*, Phillip A. Oel[1], Tobias Gerber[1], Leslie Pan[1], Nikolaos Papadopoulos[2], Z. Gülce Serka[1], Ana Verbanac[3], Maite Börsig[1], Dorinda Torres-Sabino[4], Isabel Clara Rollán-Delgado[1], Luca Santangeli[1], Henrik Kaessmann[3], Detlev Arendt[1]*.

*Co-corresponding authors

[1]Developmental Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

[2]Department of Evolutionary Biology. Integrative Zoology, University of Vienna, Vienna, Austria.

[3]Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Heidelberg, Germany

[4]Department of Zoology Genetics and Physical Anthropology. Faculty of Veterinary. University of Santiago de Compostela, Lugo, Spain

Abstract

The emergence of predation and associated complex behaviors in jawed vertebrates (gnathostomes) have been major driving forces in brain evolution. To gain insight into the neuronal complexity of the last gnathostome common ancestor, we have studied the development of the shark telencephalon combining single-cell and spatial transcriptomics. Our findings suggest an ancient origin of basal progenitors for pallial indirect neurogenesis, which was so far only described for tetrapods. We also find evidence for migrating GABAergic neurons populating the dorsal pallial field, where we observe incipient layering. Finally, we provide evidence for the existence of Cajal-Retzius homologous cells in the developing shark telencephalon, with conserved molecular profile and tissue localisation. Our data antedates the evolutionary origin of major landmarks of the vertebrate pallium by several hundred million years, to the gnathostome stem.

Introduction

The emergence of jawed vertebrates, over 450 million years ago, was a major step in vertebrate evolution. The appearance of new ecological niches exerted profound changes in body and brain organization and complexity, likely enabling higher processing speed and improved associative learning. However, the exact path of how and when the telencephalon expanded and complexified during evolution remains unsolved, due to lack of comparative data from basal vertebrate branches. Actinopterygians (ray-finned fishes) and, in particular teleost fish, experienced additional rounds of genome duplication and exhibit unique patterns of brain development and morphology, which complicate direct comparisons and evolutionary inferences. On the contrary, cartilaginous fish are slow evolvers and show conserved features in brain development (*1–4*), making them excellent models to investigate conserved traits and pinpoint ancient innovations that have shaped early gnathostome brains.

Developmental comparisons are particularly well-suited for reconstructing evolutionary history because evolutionary change is often driven by changes in development. Furthermore, embryonic processes are more likely to retain ancestral features that are modified or even absent in adults. Therefore, the study of development is central to our understanding of how neuronal diversity and complexity evolved in vertebrate brains. In this study, we investigate the developmental landscape of the shark telencephalon to address fundamental questions about brain evolution. As a key site of sensory integration, association, and learning, the telencephalon is highly relevant to understanding vertebrate brain evolution; yet its step-wise elaboration and expansion is poorly understood. Our focus is on key aspects of telencephalic development that appear crucial for these changes.

While the neuron types derived from the ventral side of the telencephalon (subpallium), appear largely conserved, the structure and composition of the dorsal side (pallium) has proven highly variable, which has so far precluded insights into its evolutionary path. A longstanding question concerns the origin of specialized neural cell types such as Cajal-Retzius (CR) cells. CR cells are critical for cortical development in mammals, where they regulate laminar organization through reelin signaling. While CR cells have been extensively studied in mammals, their evolutionary history remains unclear.

In this study, we investigate the developmental landscape of the shark telencephalon to address fundamental questions about brain evolution. Using transcriptomic profiling, trajectory analysis and tissue mapping during embryonic development, we show that the shark embryonic telencephalon shows a degree of diversification of territories comparable to that of tetrapods, including an emerging dorsal pallium receiving gabaergic migrating neurons and pallial amplifying basal

progenitors. Most strikingly, we find bona fide Cajal-Retzius cells spatially correlating with some degree of cortical layering. Our results indicate that these features -some of which previously considered mammalian or amniote innovations- were already present in gnathostome ancestors and most likely contributed to the complexification of the telencephalon concomitant with the sophistication of their brains and behaviours.

Results

Single-nucleus transcriptomics captures telencephalic developmental trajectories

To construct a single-nucleus transcriptomic dataset of the developing shark telencephalon, we microdissected telencephalon (including the olfactory bulbs) from three *S. canicula* embryos at stages 29,30 and 31 (*5*) representing peak neurogenesis and early neuronal differentiation (*1*) (Fig. 1A) (equivalent to mouse E12-14). We isolated and dissociated the nuclei and ran three independent captures using 10X Chromium Next GEM Single Cell 3' v3.1, which yielded 21,686 high-quality nuclear transcriptomes after bioinformatic quality control filtering with Seurat (*6*) and integration using Harmony (*7*). Most cells expressed neuronal markers, with smaller non-neuronal populations identified as astrocytes (gfap, s100, sox9), choroid plexus cells (rbm47, glis3), and vascular-associated cells (flt1, piezo2) (Fig. 1B).

To investigate whether our data contained neurons in progressively different developmental stages, we calculated cell-state scores using gene sets diagnostic of telencephalic lineages in mice(*8*). Despite the large evolutionary distance between shark and mice, this revealed a clear differentiation axis from early progenitors to maturing neurons (Fig. 1C), supported by pseudotime analysis (Monocle3(*9*)), RNA velocity(*10*), and cell cycle scoring (Fig. 1D).

The vertebrate telencephalon comprises two major parts: the dorsal telencephalon or pallium (where excitatory glutamatergic neurons reside) and the ventral telencephalon or subpallium (where inhibitory GABAergic neurons originate). Our dataset captured two distinct developmental branches towards pallial and subpallial differentiation, emerging from their respective early progenitor clusters (Fig. 1B). In our UMAP, pallial markers (tbr1, lhx9) were restricted to the lower branch, while subpallial markers (gad1, dlx family) defined the upper one (Fig. 1E). Pax6 protein is highly expressed in early progenitors (ventricular zone, vz) and gradually fades away as neurons differentiate (Fig. 1F'); while Gad1 protein showed no expression in cycling subpallial progenitors (vz) and sharp upregulation in postmitotic neurons (outside of vz)(Fig. 1F''). This spatial expression pattern was consistent with pax6 and gad1 transcript expression in our single nucleus dataset (Fig. 1G).

Spatial transcriptomics maps the position of differentiated neuron types

Using Visium 10X spatial transcriptomics on stage-31 transverse sections, we were able to map the anatomical position of the main differentiating and differentiated neuron types (Fig. 1H). Pallial and subpallial markers showed complementary spatial distribution, delineating these parts (compare Fig. 1E and H). The olfactory bulb was enriched in peptide YY and ppp1r16b, linking it to a pallial cluster in our UMAPmap (Fig. 1E). Additionally, GABAergic signal was detected within pallial territory (arrowhead in Fig. 1H), likely representing migratory subpallial-derived cells, as observed at earlier stages (Fig. 1I), and previously reported by us (*1*).

Lhx6 and nkx2.1 are known to specify medial ganglionic eminence (MGE) in the subpallium (*11*). These markers were specifically expressed in a distinct trajectory of the subpallium in our single-nucleus dataset (Fig. 1E) and localized to a ventromedial bulge in the shark subpallium, defining this conserved structure across vertebrates (Fig. 1H). Interestingly, the shark MGE appears as a single medial bulge rather than bilateral structures seen in other jawed vertebrates. Overall, our single-nucleus and spatial transcriptomic datasets provide a comprehensive view of shark telencephalic development, recapitulating pallial and subpallial lineage progression and identifying major differentiated neuron types (Fig. 1B).

Progenitor diversity study reveals the presence of basal progenitors in the shark telencephalon

Cell-state score analysis predicted two broad progenitor classes in the developing shark telencephalon: apical progenitors (APs) and basal progenitors (BPs) (Fig. 1C). APs divide at the apical surface of the ventricular zone (vz) (Fig.2A), while BPs, which arise from asymmetric divisions of APs, migrate to the subventricular zone (svz) for mitosis (Fig. 2B). The potential presence of BPs in the pallium of sharks has significant evolutionary implications, as this cell type is central to indirect neurogenesis—a process classically linked to the evolutionary expansion of the mammalian neocortex.

Genes associated with early APs, such as gfap and notch family members, were enriched in clusters predicted to be APs (Fig. 2C). Within this population, we identified pallial APs expressing pax6 and lrp4, and subpallial APs enriched for meis2a and ntn1 (Fig. 2C), consistent with patterns observed in other vertebrates. Along the pallial lineage, we observed the sequential expression of pax6, eomes (previously known as tbr2), and tbr1 (Fig. 2D). In mice, these genes form a direct feedforward genetic cascade that regulates the sequential differentiation of pallial APs (Pax6 high) into BPs (Pax6 low and Eomes high) and finally into early excitatory projection neurons (Tbr1) (*12*). Immunohistochemistry confirmed high Pax6 expression in the VZ, lower levels in the SVZ, where Eomes transcript is specifically expressed (Fig. 2E). These findings align with previous observations (*2*) and reveal strong conservation of this genetic cascade in driving the early differentiation of telencephalic excitatory

neurons. This conservation accordingly dates back at least to the last gnathostome common ancestor and suggests the presence of a distinct BP state in the developing pallium of sharks.

One central question is whether this BP state in sharks represents a distinct progenitor type or just a transitional state towards differentiation. To investigate this, we examined cells with high BP scores in our single-nucleus dataset. These cells were distributed across both early differentiating branches and within progenitor clusters (Fig. 1C). We noticed that the mitotic gene expression signature grouped cells together into a “M-phase cluster” (Fig.2F), and that a proportion of BPs fell into this mitotic cluster (arrowheads in Fig. 1C). Likewise, top enriched pallial BP genes were co-expressed in the same mitotic cluster area than the one predicted by cell score analysis, confirming that a proportion of BPs must be mitotically active (arrowheads in Fig. 2G). In summary, BP identity in our dataset likely represents a state encompassing cells which just exited the cell cycle from AP state and are now entering into BP state, as well as actively dividing BPs and cells transitioning from BP state towards differentiation. The presence of mitotic cells (PH3-positive) cells in a subventricular position, noted in our previous work (*2*), further supports the existence of a distinct BP progenitor state in sharks. Among the pallial BP-enriched genes, we identified the transcription factor Ebf3, which we validated by HCR to be highly expressed in the pallial svz (Fig. 2E).

To unbiasedly assess potential homology between shark and murine BPs, we conducted transcriptional comparisons using published single-cell data (*8*, *12*) from the developing mouse telencephalon at equivalent embryonic stages (E12-14), where APs and BPs were easily identifiable. We annotated shark pallial mitotic BPs as a distinct cluster from mitotic APs, guided by BP marker expression profiles (Fig. 2G). Our cross-species comparison using SAMap revealed a clear correlation trend between shark and mouse APs and shark and mouse BPs (Fig. 2H).

Overall, these findings suggest that shark BPs closely resemble their mouse counterparts in transcriptomic profile, mitotic activity, and SVZ localization, supporting an evolutionary common origin of indirect neurogenesis dating back to early gnathostomes.

We next compared the transcriptomic profiles of APs and BPs across species. In order to capture robust transcriptomic signatures, we identified genes significantly enriched in each species and retained only those exhibiting a log2 fold change of ≥ 1.5. We then determined the set of shared and species-specific orthologous genes for each cell type. The shared gene signature revealed that APs in both species exhibit high expression of Notch-family genes, while BPs express the Delta gene Dll1. Only 11 orthologous genes were commonly enriched in shark and mouse BPs, most of which coincide with established BP markers in mammals. This conserved signature includes transcription factors such as Ngn1, Ngn2, Eomes, Insm1, and Lhx9 (Fig. 2I). Together, these findings suggest that the Delta-Notch signaling axis, combined with this core set of transcriptional regulators, constitutes an

evolutionary toolkit for indirect neurogenesis that has persisted for over 450 million years. Additionally, genes shared among APs appear to regulate stem cell maintenance, Hedgehog/FGF signaling pathways, and chromatin remodeling (Fig. 2I).

Regarding species-specific signatures, mouse APs display strong specific expression of Sfrp2, a Wnt pathway modulator, which could reflect adaptations to the increased proliferative complexity required for layered cortical development. Mouse BPs exhibit an enhanced proliferative profile compared to their shark counterparts, characterized by the acquisition of genes such as Hes6 and E2f1. Additionally, the retention of Argap10 (a Rho GFPase involved in cytoskeletal organization and cell migration) in mouse BPs could be associated with the radial expansion necessary for neocortical layering. In contrast, shark BPs show a stronger differentiation profile, with elevated expression of Ebf3 and Wnt7b, suggesting more limited BP amplification relative to mice. Together, our results indicate that core developmental mechanisms underlying pallial indirect neurogenesis were already established in the common ancestor of all jawed vertebrates and point to molecular evolutionary changes that may have promoted the retention of proliferative capacity in mammalian BPs, thereby contributing to mammalian pallial expansion.

To explore AP diversity, we isolated, reclustered and regressed the effect of the cell cycle of the AP clusters (including mitotic BPs in the analysis), revealing transcriptional variability along the pallial-subpallial axis. In the new embedding, the original M-phase clusters (red, blue and yellow in Fig. 2J), redistributed through this axis within their respective cell type identities. Based on known markers in other vertebrates and their relative position in the dorso-ventral axis, we propose the existence of at least five broad AP subtypes: medial ganglionic eminence (MGE)/preoptic area precursors (shh-high), lateral ganglionic eminence (LGE) precursors (meis2-high), dorsal LGE/pallial-subpallial border precursors (prdm16-etv1-high), ventrolateral pallium precursors (pax6-lrp4-high), and dorsomedial pallium precursors (lrp4-lhx9-foxp2-rspo1-high). Additionally, we detected a cluster of incipient BP-like cells expressing st18 and ebf3 (Fig. 1J). Interestingly, this cluster contains cells initially assigned to all the other clusters suggesting that all neuronal classes pass through a similar BP-like state during differentiation.

Trajectory study of the catshark subpallium shows three main differentiation pathways for gabaergic inhibitory neurons, including one migratory population directed to the pallium

To investigate the developmental trajectories of shark inhibitory cell types, we bioinformatically isolated and reclustered cells from the subpallial branch. Three major branches were apparent from the 3D UMAP structure, confirmed by trajectory inference algorithms (Monocle 3) (Fig.3A). We then identified differentially expressed genes along each sub-trajectory (Fig. 3 B,C). By integrating spatial

transcriptomic and HCR/ISH data, we assigned each trajectory to its corresponding terminal neuronal identity (Fig. 3C).

Our analysis revealed that early GABAergic neuron differentiation proceeds through an intermediate basal progenitor-like state, marked by specific expression of transcription factors Dlx1 and Isl1. From this state, the three distinct trajectories emerge. The first, the MGE trajectory, is characterized by upregulation of Lhx6 and Nkx2.1 and leads to the differentiation of MGE-derived interneurons and a second population rich in Nxph1, whose anatomical position is compatible with a location at the subpallial septum. Cross-species comparison (SAMap) confirmed a strong and exclusive correlation of shark MGE types with MGE and subpallial septum types of salamander (Fig. 3D).

The second main branch corresponds to an LGE-derived trajectory and is defined by sequential upregulation of Prox1, Zfhx4, and Six3, culminating in populations expressing Cap2 and Galanin. Anatomically, these cells correspond to part of the area superficialis basalis (ASB), a conspicuous mass of cells in the subpallium of all chondrichthyans, whose identity and homology to other vertebrates remain to be fully elucidated. It has been proposed that they contain striatal and amygdalar structures (*13*, *14*). The three clusters contained in this trajectory show high correlations with salamander cell types assigned to striatum and subpallial amygdala (Fig.3D).

Finally, we observed that the third main branch grouped together LGE interneurons with a migratory behaviour. This branch bifurcates into two subpopulations: one marked by Etv1 expression, followed by Pax6 and Kcns3a (cyan asterisk in Fig. 3C); and a second population co-expressing Foxp2 and Pax6 (magenta asterisk in Fig. 3C). According to their anatomical positions, the former likely corresponds to tangentially migrating interneurons destined for the pallium (cyan arrowheads in Fig. 3C), while the latter could correspond to dorsal-LGE (dLGE)-derived migratory neurons likely migrating to the olfactory bulb (magenta arrowheads in Fig. 3C). When confronted with salamander dataset, the migratory interneuron cluster showed strong correspondence with salamander olfactory bulb interneurons. The correlation with interneurons migrating to the pallium could not be assessed due to the lack of such a cluster in the salamander dataset.

These findings highlight conserved main developmental trajectories across vertebrates for inhibitory neurons such as: 1) the one giving rise to MGE and septum, 2) LGE-derived striatal and amygdalar neurons and 3) dLGE-derived olfactory bulb neurons. At the same time, they reveal potentially unique features in sharks, including pallial migratory interneurons that exhibit divergent gene expression profiles (e.g., Pax6) and appear to originate from the LGE rather than the MGE, as observed in mammals. Whether this represents an ancestral vertebrate condition or a lineage-specific adaptation remains an open question for future investigation.

Study of pallial developmental diversity uncovers four main excitatory neuron identities, and identifies a dorsal pallial homolog

The vertebrate pallium is one of the regions of the brain that have experienced the highest variability across animal groups and one of the most difficult to compare across species. Still, the existence of a "four-field pallium" (tetrapartite pallium) in all tetrapods has been proposed (*15*), with the dorsal pallial field giving rise to the layered mammalian neocortex, avian hyperpallium or reptile three-layered cortex. When this general pallial organization appeared in evolution is still a matter of debate. To shed new light on the matter, we performed a detailed analysis of the pallial components in our datasets, as well as cross-species comparisons with mouse and salamander developing pallial datasets.

In contrast to the subpallial trajectory, the pallial lineage did not exhibit distinct branching patterns but instead formed a more continuous landscape of cell states (Fig. 4A), likely reflecting gradual transitions in cell identity. Trajectory inference with the Monocle3 algorithm predicted two main paths diverging from early pallial neurons and converging at the end (Fig. 4B). By combining clustering, gene enrichment analysis, and the identification of spatially restricted genes in our single-nucleus and spatial transcriptomic datasets, we identified at least four molecularly and anatomically distinct cell types within the differentiated pallial neurons branch (Fig. 4C), excluding the olfactory bulb and separated pallial clusters, which will be considered later. First, and consistent with the trajectory prediction, the pallium appears to be divided into two broad regions: (1) a ventral-lateral pallial area, characterized by expression of Nr2f1 (Fig. 4D) and (2) a dorsal-medial pallial area marked by strong Sall1 expression (Fig.4E). These genes showed non-overlapping expression also in the spatial data. Lhx9 showed high and specific spatial expression at medial-most and ventral-most pallial regions (Fig. 4F,C), in agreement with the tetrapartite pallial model proposed for tetrapods (*15*, *16*), which considers this transcription factor as a specific medial and ventral pallial marker. The transcription factor Satb1 was highly expressed in the most medial pallial area, near the choroid plexus attachment (green arrows in Fig. 4G), and showed lower, yet specific, expression in the putative dorsal pallial area (black arrows in Fig. 4G). Alkal2 was detected at medial pallial levels, with some expression extending toward dorsal regions (Fig. 4H). Pcp4 was strongly expressed in the anatomical area identified as the DP (arrowheads in Fig. 4I), which also expressed specific markers such as gdf10 and slit2 (Fig. 4J,K). Grik1, a glutamate ionotropic receptor typically found in glutamatergic neurons, was highly expressed in the most mature neurons of the dorsal and lateral pallium (Fig. 4L). Collectively, these patterns support that the developing shark pallium comprises at least four main, partially overlapping neuronal populations, which we have labeled as the putative medial, dorsal, lateral and ventral pallial areas.

For cross-species comparisons, we used the pre-established annotation of the salamander dataset, which included the main developmental trajectories and four terminal populations, as well as a pallial amygdala cluster and two olfactory bulb clusters. We used the mouse dataset from (*17*), because it matched developmental timing better with our datasets. This mouse dataset was obtained from a purified pallial population lacking ventral pallium (Emx1-CRE FACS-sorted) and olfactory bulb clusters; we therefore annotated medial, dorsal and lateral pallium regions using established mammalian markers (*15*, *16*). Although we did not observe a strict one-to-one correspondence between specific pallial subdivisions across species—likely due to partial overlap of molecular identities, differences in annotation granularity, or slight discrepancies in developmental stages—a clear and noteworthy trend emerged from our analysis. The shark dorsal pallial cluster showed a very high correlation with the mouse dorsal pallial trajectory and with both the dorsal and medial pallial trajectories in salamander (asterisks in Fig. 4M). The shark lateral pallial area correlated most strongly with the mouse lateral pallium and with salamander olfactory bulb cell types, while displaying lower and more diffuse correlations with other salamander pallial types. Shark medial pallium found high correlation with mouse dorsal and medial pallium but also with developmental trajectories of various salamander types. Shark ventral pallium also showed a diffuse correlation across salamander and mouse types. Notably, the shark dorsal pallium was unique among all shark pallial clusters in that it exhibited no association with olfactory bulb-related cell types (arrowhead in Fig. 4M).

Outside of the main pallial branch, we identified a highly specific match between a previously uncharacterized shark pallial cluster and the salamander pallial amygdala, leading us to annotate this cluster as the shark putative pallial amygdala. As expected, the shark pallial olfactory bulb cluster showed a strong correlation with the salamander mature pallial olfactory bulb cluster.

Our findings suggest that the delineated dorsal pallial area of sharks possesses a distinct molecular signature that sets it apart from other pallial regions and its high correlation with developing dorsal pallium of mouse, and dorsomedial pallium of salamanders suggests an evolutionary common origin.

Presence of Cajal-Retzius-like cells correlating with pallial layering

Finally, we focused on one small cluster of pallial neurons (Fig. 5A) characterized by Tp73, Ebf3 and Lhx1 expression (Fig. 5B). These genes are the molecular signature of mammalian Cajal-Retzius cells (CRs) (*18*). CRs are a transient population of neurons found in the outermost layer of the developing cerebral cortex (dorsal pallium) and hippocampus (medial pallium) in mammals, best known for producing reelin, a signalling molecule crucial for guiding the radial migration and proper lamination of neurons of the neocortex. Their developmental functions are numerous and crucial for the correct architecture of the mammalian cortex and most CR cells undergo apoptosis after cortical

development is complete. Their presence in non-mammalian vertebrates is a matter of debate, with some evidence in reptiles, controversy in birds and no evidence to date in fish (*18–22*).

For an in-depth comparison of transcriptomic profiles, we used the mouse dataset with a previously annotated CR cluster (Fig. 5C). After processing the salamander dataset, we found a non-annotated cluster in specifically expressing Tp73, Ebf3 and Lhx1 (Fig. 5C), which we included in our three-way comparison. Integrating the three datasets based on shared orthogroups, CR cells from the three species co-distributed very closely in the new embedding (Fig. 5D). When integrating shark and mouse datasets, both CR types found each other in one of the most robust and individuated cluster of the whole dataset (Fig. 5E). Other integration methods yielded similar results. SAMap comparisons showed that the CR cluster was the only cell type that found a unique correlative across the three species (red lines in Fig. 5F). Taken together, these results provide strong evidence that CJ cells exist in sharks and suggest an ancient origin of CRs in vertebrates.

We next determined the localization of CR cells in the shark developing telencephalon. None of the three markers yielded conclusive results in our spatial transcriptomic dataset due to the limited resolution. We therefore performed HCR detection on tissue sections, where pallial cells expressing Ebf3 in the absence of Foxp2 demarcate CRs (orange arrows in Fig.5G); while Ebf3, Foxp2 co-expression labels pallial amygdala cell types (asterisk in Fig.5 G,H'). We found Foxp2-negative, Ebf3-positive putative CR cells lining the pial (external) surface throughout the pallium, particularly evident in close proximity to the choroid plexus attachment point -i.e. medial pallial regions- (Fig. 5H). At more caudal levels of the brain, we also found them extending through the pial surface, from medial pallial regions to areas close to the olfactory bulb (arrowheads in Fig.5H'''). We confirmed co-expression with the CR marker Tp73 in the same areas (Fig. 5I, I',I''). We observed only a few triple labelled neurons Ebf3-Tp73-Lhx1 at the caudal medial pallium location (Fig. 5I''). This could be indicative of subtypes of CR cells being distributed differently across the telencephalon. In summary, cells expressing CR markers were observed in close proximity to the choroid plexus attachment at medial and caudal levels of the medial pallium, as well as extending along the pial surface of the dorsal pallium, lateral pallium and the vicinity of the olfactory bulb (Fig. 5J). Both molecular profile and anatomical distribution of putative shark CRs thus parallel those observed in mammals.

We next investigated the expression pattern of reelin, known to be secreted by CR cells in mammals (*18*). In the single nucleus dataset, reelin transcript is expressed in the tip of the CR cluster, in the olfactory bulb and parts of the dorsomedial pallium (Fig. 5K). We confirmed the spatial distribution of reelin transcript in our spatial data as well as that of reelin protein with fluorescent immunohistochemistry (Fig. 5K,L). Pax6 and Gad1 distribution were used to delineate the different pallial regions and the tangentially migrating gabaergic neurons (arrowheads in Fig. 5M). The

distribution of Reelin indeed covered the areas where CRs are present, i.e., in the marginal zone of the dorsolateral pallium and olfactory bulb(Fig. 5K,L). As reported in our previous work (*1*), reelin distribution also coincides with that of the migratory interneurons into the pallium across dorsal subpial pallial areas (arrowheads in Fig. 5L). Importantly, Reelin is an extracellular matrix protein that regulates neuronal migration and layer formation in the developing mammalian brain (*23*).

To investigate possible layering of the developmental derivatives arising from the dorsal telencephalon in sharks, we electroporated a non-integrative expression plasmid (CAG-CMV-RFP) into the early telencephalon of five stage 29 embryos that we observed for 4 to 7 weeks (Fig. 5N) and sacrificed at different times, the most mature reaching late embryonic stage 32, when neuronal architecture starts to be similar to the mature brain. We achieved consistent labeling of the dorsal and lateral pallium, as well as of areas of the lateral subpallium (LGE). We observed apparent streams of migratory populations emerging around stage 31 (arrowhead in Fig. 5N). 1 week after electroporation we observed cells emerging from the ventricular zone with a radial migratory morphology, suggestive of potential lineages (arrowheads in Fig. 5O), as well as neurons in the mantle zone with a tangential orientation (Fig. 5O). 2 weeks after electroporation, when the embryo reaches stage 31, the same area showed a most sophisticated pattern, with numerous axons and cell bodies observed in an incipient layer (Fig. 5P). 4 to 7 weeks post electroporation, around stage 32, we found a more defined layered structure expanding through rostro-medial levels of the dorsal pallium, evident both in sagittal (Fig. 5Q') and transverse sections (Fig. 5R'). In line with this observation, we had previously described a layered distribution of Dlx2-positive GABAergic neurons situated beneath a layer of Tbr1-positive glutamatergic neurons in the rostral pallium of late embryos of *S.canicula* (see Fig. 6B,C in (1)). In summary, our new data indicates the emergence of a layered organization in the dorsal pallium of the shark telencephalon as early as embryonic stage 31.

Discussion

In this study, we present a single nucleus and spatial transcriptomics atlas of the developing shark telencephalon during key stages of neurogenesis and neural differentiation, capturing the developmental trajectories of pallial and subpallial cell types. This is an important resource for comparative developmental studies that will help to shed light on the origin and evolution of telencephalic cell types.

The existence of indirect neurogenesis and basal progenitors (BPs) in the pallium of fish has been debated, but our data provides compelling evidence that the subventricular mitoses previously identified in the embryonic shark pallium (*2*) correspond to a mitotic cell population with highest transcriptomic similarity to mammalian BPs. We further identify a core transcriptomic signature of

pallial BPs conserved from sharks to mammals, supporting the hypothesis that indirect neurogenesis emerged in the pallium of the first jawed vertebrates, promoting brain expansion.

Our reconstruction of subpallial trajectories supports conservation of MGE and LGE but not of the CGE, which may have evolved later as an extension of the MGE. Notably, the GABAergic neurons migrating to the shark pallium exhibit an LGE-like profile, in contrast to the MGE-derived interneurons in mammals, implying that MGE-derived pallial migration may be a later evolutionary innovation. This evidence is in line with previous research suggesting an evolutionary shift in cortical interneuron migration from the LGE to the MGE in mammals and possibly in birds (*24*). Whether the pallial-directed population we observe in sharks corresponds to a known basal vertebrate type or represents a lineage-specific feature remains to be determined.

The analysis of developing pallial cell types in sharks reveals at least four broad populations with overlapping identities. The dorsal pallium, in particular, displays a unique transcriptomic profile, distinct from other pallial regions and lacking olfactory bulb-associated signatures. Cross-species comparisons show a high correlation between shark dorsal pallial neurons and those of the mouse dorsal pallium and salamander dorsal and medial pallium, suggesting evolutionary correspondence. This evidence is significant in the context of longstanding debates on the evolutionary origins and homology of the dorsal pallium in vertebrates (*25*). The large, complex telencephalon of sharks has often been considered a derived feature, but our developmental and transcriptomic data argue for a shared origin of the dorsal pallium as a separate telencephalic subdivision. The appearance of a dorsal pallial field in the telencephalon of the first jawed vertebrates might have indeed been aided by the emergence of pallial indirect neurogenesis and basal progenitors.

A key finding of our study is the identification of a conserved Cajal-Retzius (CR) cell type in the developing shark brain, supported by transcriptomic identity and anatomical localization. This discovery antedates the evolutionary origin of CR cells by several hundred million years, to the gnathostome stem. CR cells have not been reported in fish, and their presence in birds and other amniotes (*21*, *26*, *27*) remains elusive. It now appears plausible that CR cells were lost or modified in these lineages. To refine the evolutionary trajectory of CR cells in vertebrates, it will be rewarding to investigate their presence and characteristics in agnathans, non-teleost fish, and basal sauropsids. Our additional observations that the distribution of CR across the periphery of the dorsal pallium correlates with reelin expression, and evidence for the existence of cortical layering yield strong support to the key hypothesis that the emergence of CR cells relates to a first expansion phase of the vertical pallium in the agnathan stem, which finds support by the apparent absence of CRs and cortical layering from agnathans (*28*). Concomitant with the emergence of pallial indirect neurogenesis, and separate medial, dorsal and lateral pallial subdivisions, CR cells could have

acquired enhanced migratory and signaling properties, facilitating the coordinated formation of new architectures and circuits. The appearance of layered structures in the pallium may have endowed early vertebrates with greater computational capacity, supporting the emergence of more sophisticated behaviors.

Acknowledgments

The authors thank Kresimir Crnokic for his support in animal husbandry, Matthias Janeschik for valuable discussion and exchanges, Ahmad Al Alwash for invaluable help with the spatial transcriptomics protocol, Ram Reshef (University of Haifa) for providing the reporter plasmid, GeneCore and ALMF imaging facilities from EMBL. The work was supported by the European Research Council Advanced grant 788921/NeuralCellTypeEvo (D.A.). I Q-U was supported by a fellowship from the EMBL Interdisciplinary Postdoc (EIPOD) program under Marie Skłodowska-Curie Actions COFUND.

Author contributions

Conceived the study: I.Q-U

Acquired funding: D.A., I.Q-U

Planned experiments: I.Q-U, P.A.O, L.P., G.S., A.V., M.B., I.C.R-D, D.T-S.

Executed experiments: I. Q-U, P.A.O, L.P., G.S, A.V, M.B, D.T-S, I.C.R-D, L.S

Designed and executed plasmid electroporation experiments: I.Q-U, I.C.R-D, M.B, D.T-S

Data processing: P.A.O., T.G, N.P

Analysed data and gave critical input: L.S., N.P., T.G.

Supervised the study and provided funding: D.A., H.K.

Interpreted the data and wrote the manuscript with input from all authors: I.Q-U, D.A.

References

1. I. Quintana-Urzainqui, I. Rodríguez-Moldes, S. Mazan, E. Candal, Tangential migratory pathways of subpallial origin in the embryonic telencephalon of sharks: evolutionary implications. *Brain Struct. Funct.* **220**, 2905–2926 (2015).

2. A. Docampo-Seara, R. Lagadec, S. Mazan, M. A. Rodríguez, I. Quintana-Urzainqui, E. Candal, Study of pallial neurogenesis in shark embryos and the evolutionary origin of the subventricular zone. *Brain Struct. Funct.* **223**, 3593–3612 (2018).

3. I. Rodríguez-Moldes, G. N. Santos-Durán, S. Pose-Méndez, I. Quintana-Urzainqui, E. Candal, "The Brains of Cartilaginous Fishes" in *Evolutionary Neuroscience* (Elsevier, 2020; https://linkinghub.elsevier.com/retrieve/pii/B9780128205846000064), pp. 101–123.

4. G. N. Santos-Durán, A. Menuet, R. Lagadec, H. Mayeur, S. Ferreiro-Galve, S. Mazan, I. Rodríguez-Moldes, E. Candal, Prosomeric organization of the hypothalamus in an elasmobranch, the catshark Scyliorhinus canicula. *Front. Neuroanat.* **09** (2015).

5. W. W. Ballard, J. Mellinger, H. Lechenault, A series of normal stages for development of *Scyliorhinus canicula* , the lesser spotted dogfish *(Chondrichthyes: Scyliorhinidae)*. *J. Exp. Zool.* **267**, 318–336 (1993).

6. Y. Hao, T. Stuart, M. H. Kowalski, S. Choudhary, P. Hoffman, A. Hartman, A. Srivastava, G. Molla, S. Madad, C. Fernandez-Granda, R. Satija, Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024).

7. I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

8. M. X. Moreau, Y. Saillour, A. W. Cwetsch, A. Pierani, F. Causeret, Single-cell transcriptomics of the early developing mouse cerebral cortex disentangle the spatial and temporal components of neuronal fate acquisition. *Development* **148**, dev197962 (2021).

9. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N. J. Lennon, K. J. Livak, T. S. Mikkelsen, J. L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).

10. G. La Manno, R. Soldatov, A. Zeisel, E. Braun, H. Hochgerner, V. Petukhov, K. Lidschreiber, M. E. Kastriti, P. Lönnerberg, A. Furlan, J. Fan, L. E. Borm, Z. Liu, D. Van Bruggen, J. Guo, X. He, R. Barker, E. Sundström, G. Castelo-Branco, P. Cramer, I. Adameyko, S. Linnarsson, P. V. Kharchenko, RNA velocity of single cells. *Nature* **560**, 494–498 (2018).

11. G. E. Elsen, F. Bedogni, R. D. Hodge, T. K. Bammler, J. W. MacDonald, S. Lindtner, J. L. R. Rubenstein, R. F. Hevner, The Epigenetic Factor Landscape of Developing Neocortex Is Regulated by Transcription Factors Pax6→ Tbr2→ Tbr1. *Front. Neurosci.* **12**, 571 (2018).

12. M. Manuel, K. B. Tan, Z. Kozic, M. Molinek, T. S. Marcos, M. F. A. Razak, D. Dobolyi, R. Dobie, B. E. P. Henderson, N. C. Henderson, W. K. Chan, M. I. Daw, J. O. Mason, D. J. Price, Pax6 limits the competence of developing cerebral cortical cells to respond to inductive intercellular signals. *PLOS Biol.* **20**, e3001563 (2022).

13. I. Rodríguez-Moldes, I. Quintana-Urzainqui, G. N. Santos-Durán, S. Ferreiro-Galve, S. Pereira-Guldrís, M. Candás, S. Mazan, E. Candal, Identifying Amygdala-Like Territories in ***Scyliorhinus canicula*** (Chondrichthyan): Evidence for a Pallial Amygdala. *Brain. Behav. Evol.* **96**, 283–304 (2022).

14. I. Quintana-Urzainqui, C. Sueiro, I. Carrera, S. Ferreiro-Galve, G. Santos-Durán, S. Pose-Méndez, S. Mazan, E. Candal, I. Rodríguez-Moldes, Contributions of Developmental Studies in the Dogfish ***Scyliorhinus canicula*** to the Brain Anatomy of Elasmobranchs: Insights on the Basal Ganglia. *Brain. Behav. Evol.* **80**, 127–141 (2012).

15. L. Medina, A. Abellán, Development and evolution of the pallium. *Semin. Cell Dev. Biol.* **20**, 698–711 (2009).

16. A. AbellÃ¡n, E. Desfilis, L. Medina, Combinatorial expression of Lef1, Lhx2, Lhx5, Lhx9, Lmo3, Lmo4, and Prox1 helps to identify comparable subdivisions in the developing hippocampal formation of mouse and chicken. *Front. Neuroanat.* **8** (2014).

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

18. F. García-Moreno, E. Anderton, M. Jankowska, J. Begbie, J. M. Encinas, M. Irimia, Z. Molnár, Absence of Tangentially Migrating Glutamatergic Neurons in the Developing Avian Brain. *Cell Rep.* **22**, 96–109 (2018).

19. A. Cabrera-Socorro, N. C. Hernandez-Acosta, M. Gonzalez-Gomez, G. Meyer, Comparative aspects of p73 and Reelin expression in Cajal-Retzius cells and the cortical hem in lizard, mouse and human. *Brain Res.* **1132**, 59–70 (2007).

20. A. M. Goffinet, The evolution of cortical development: the synapsid-diapsid divergence. *Development* **144**, 4061–4077 (2017).

21. T. Nomura, M. Hattori, N. Osumi, Reelin, radial fibers and cortical evolution: Insights from comparative analysis of the mammalian and avian telencephalon. *Dev. Growth Differ.* **51**, 287–297 (2009).

22. D. H. Tanaka, R. Oiwa, E. Sasaki, K. Nakajima, Changes in cortical interneuron migration contribute to the evolution of the neocortex. *Proc. Natl. Acad. Sci.* **108**, 8015–8020 (2011).

23. S. D. Briscoe, C. W. Ragsdale, Evolution of the Chordate Telencephalon. *Curr. Biol.* **29**, R647–R662 (2019).

24. B. Zaremba, A. Fallahshahroudi, C. Schneider, J. Schmidt, I. Sarropoulos, E. Leushkin, B. Berki, E. Van Poucke, P. Jensen, R. Senovilla-Ganzo, F. Hervas-Sotomayor, N. Trost, F. Lamanna, M. Sepp, F. García-Moreno, H. Kaessmann, Developmental origins and evolution of pallial cell types and structures in birds. *Science* **387**, eadp5182 (2025).

25. F. Lamanna, F. Hervas-Sotomayor, A. P. Oel, D. Jandzik, D. Sobrido-Cameán, G. N. Santos-Durán, M. L. Martik, J. Stundl, S. A. Green, T. Brüning, K. Mößinger, J. Schmidt, C. Schneider, M. Sepp, F. Murat, J. J. Smith, M. E. Bronner, M. C. Rodicio, A. Barreiro-Iglesias, D. M. Medeiros, D. Arendt, H. Kaessmann, A lamprey neural cell type atlas illuminates the origins of the vertebrate brain. *Nat. Ecol. Evol.* **7**, 1714–1728 (2023).

Figure 1

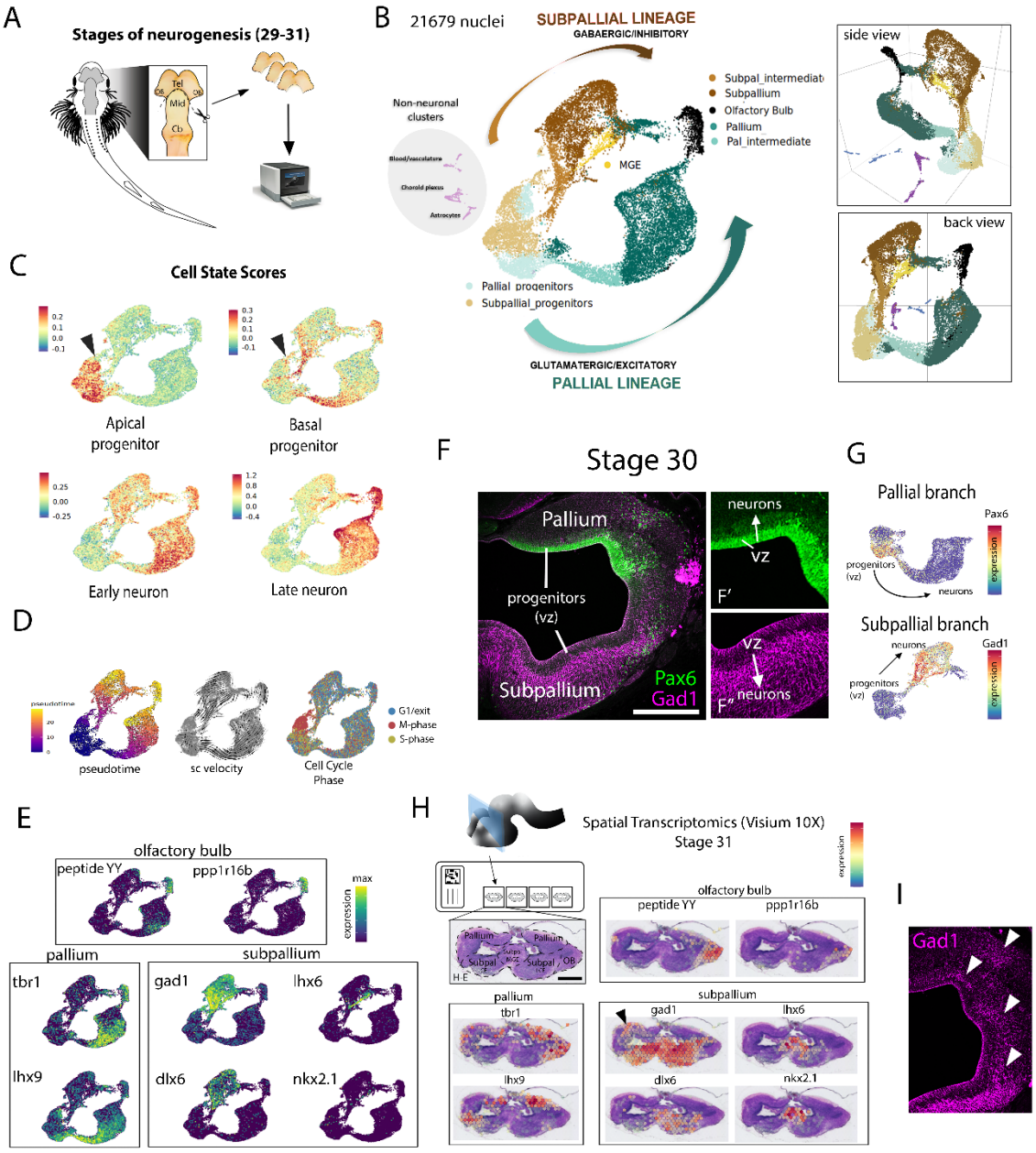
A
Stages of neurogenesis (29-31)
Tel
Mid
Cb
B
21679 nuclei
SUBPALLIAL LINEAGE
GABAERGIC/INHIBITORY
Non-neuronal clusters
Blood/vasculature
Choroid plexus
Astrocytes
Subpal_intermediate
Subpallium
Olfactory Bulb
Pallium
Pal_intermediate
MGE
Pallial progenitors
Subpallial_progenitors
GLUTAMATERGIC/EXCITATORY
PALLIAL LINEAGE
side view
back view
C
Cell State Scores
Apical progenitor
Basal progenitor
Early neuron
Late neuron
D
pseudotime
sc velocity
Cell Cycle Phase
G1/exit
M-phase
S-phase
E
olfactory bulb
peptide YY
ppp1r16b
max
expression
pallium
subpallium
tbr1
lhx9
gad1
lhx6
dlx6
nkx2.1
F
Stage 30
Pallium
progenitors (vz)
Subpallium
Pax6
Gad1
neurons
VZ
F′
F″
G
Pallial branch
Pax6
progenitors (vz)
neurons
expression
Subpallial branch
Gad1
H
Spatial Transcriptomics (Visium 10X)
Stage 31
H.E
I
Gad1

Figure 2

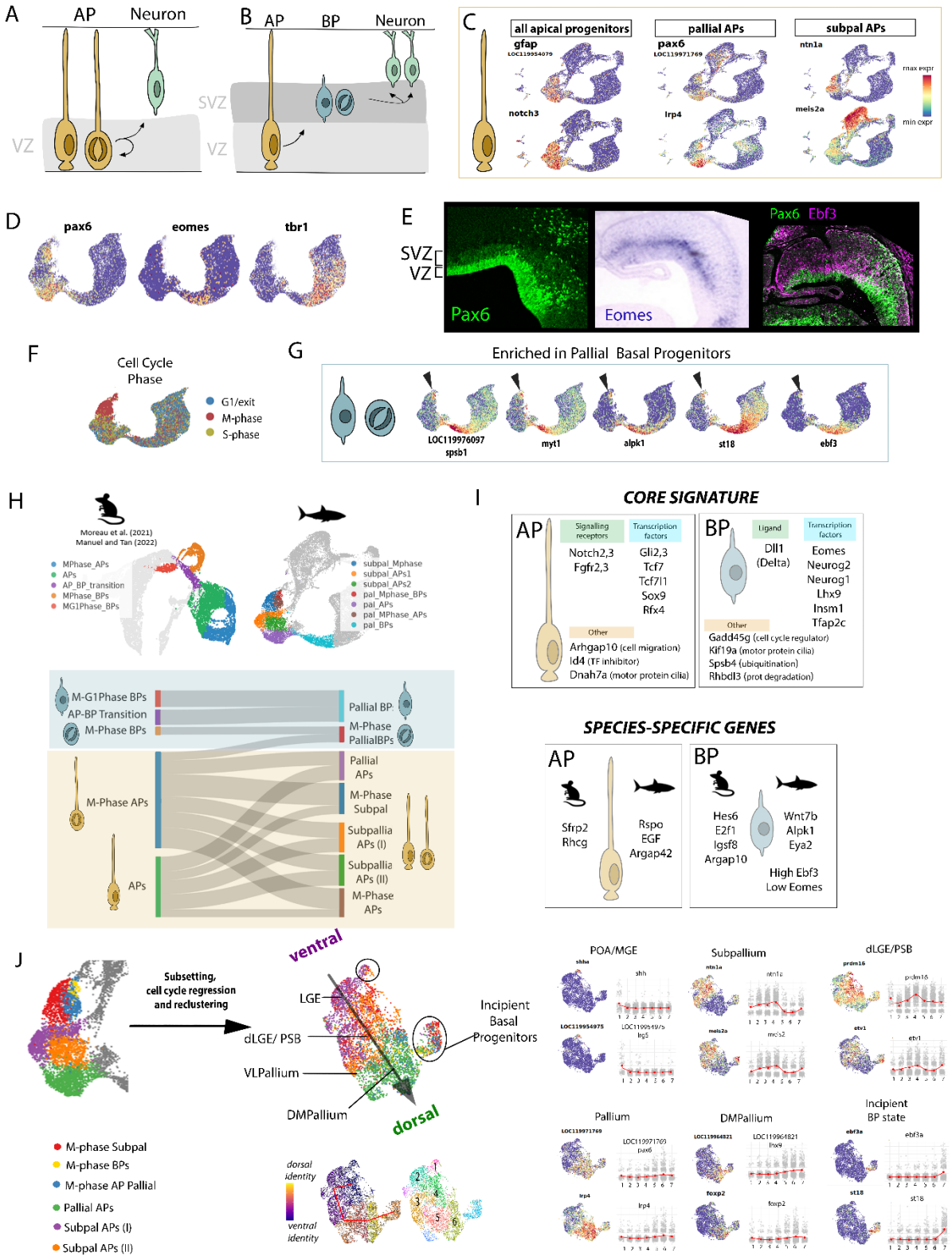
A
AP
Neuron
VZ
B
AP
BP
Neuron
SVZ
VZ
C
all apical progenitors
pallial APs
subpal APs
gfap
pax6
ntn1a
notch3
lrp4
meis2a
max expr
min expr
D
pax6
eomes
tbr1
E
SVZ
VZ
Pax6
Eomes
Pax6 Ebf3
F
Cell Cycle Phase
G1/exit
M-phase
S-phase
G
Enriched in Pallial Basal Progenitors
LOC119976097
spsb1
myt1
alpk1
st18
ebf3
H
Moreau et al. (2021)
Manuel and Tan (2022)
MPhase_APs
APs
AP_BP_transition
MPhase_BPs
MG1Phase_BPs
subpal_Mphase
subpal_APs1
subpal_APs2
pal_Mphase_BPs
pal_APs
pal_MPhase_APs
pal_BPs
M-G1Phase BPs
AP-BP Transition
M-Phase BPs
Pallial BP
M-Phase PallialBPs
M-Phase APs
APs
Pallial APs
M-Phase Subpal
Subpallia APs (I)
Subpallia APs (II)
M-Phase APs
I
CORE SIGNATURE
AP
Signalling receptors
Transcription factors
Notch2,3
Fgfr2,3
Gli2,3
Tcf7
Tcf7l1
Sox9
Rfx4
Other
Arhgap10 (cell migration)
Id4 (TF inhibitor)
Dnah7a (motor protein cilia)
BP
Ligand
Transcription factors
Dll1 (Delta)
Eomes
Neurog2
Neurog1
Lhx9
Insm1
Tfap2c
Other
Gadd45g (cell cycle regulator)
Kif19a (motor protein cilia)
Spsb4 (ubiquitination)
Rhbdl3 (prot degradation)
SPECIES-SPECIFIC GENES
AP
Sfrp2
Rhcg
Rspo
EGF
Argap42
BP
Hes6
E2f1
Igsf8
Argap10
Wnt7b
Alpk1
Eya2
High Ebf3
Low Eomes
J
Subsetting, cell cycle regression and reclustering
ventral
LGE
dLGE/ PSB
VLPallium
DMPallium
dorsal
Incipient Basal Progenitors
M-phase Subpal
M-phase BPs
M-phase AP Pallial
Pallial APs
Subpal APs (I)
Subpal APs (II)
dorsal identity
ventral identity
POA/MGE
Subpallium
dLGE/PSB
Pallium
DMPallium
Incipient BP state

Figure 3

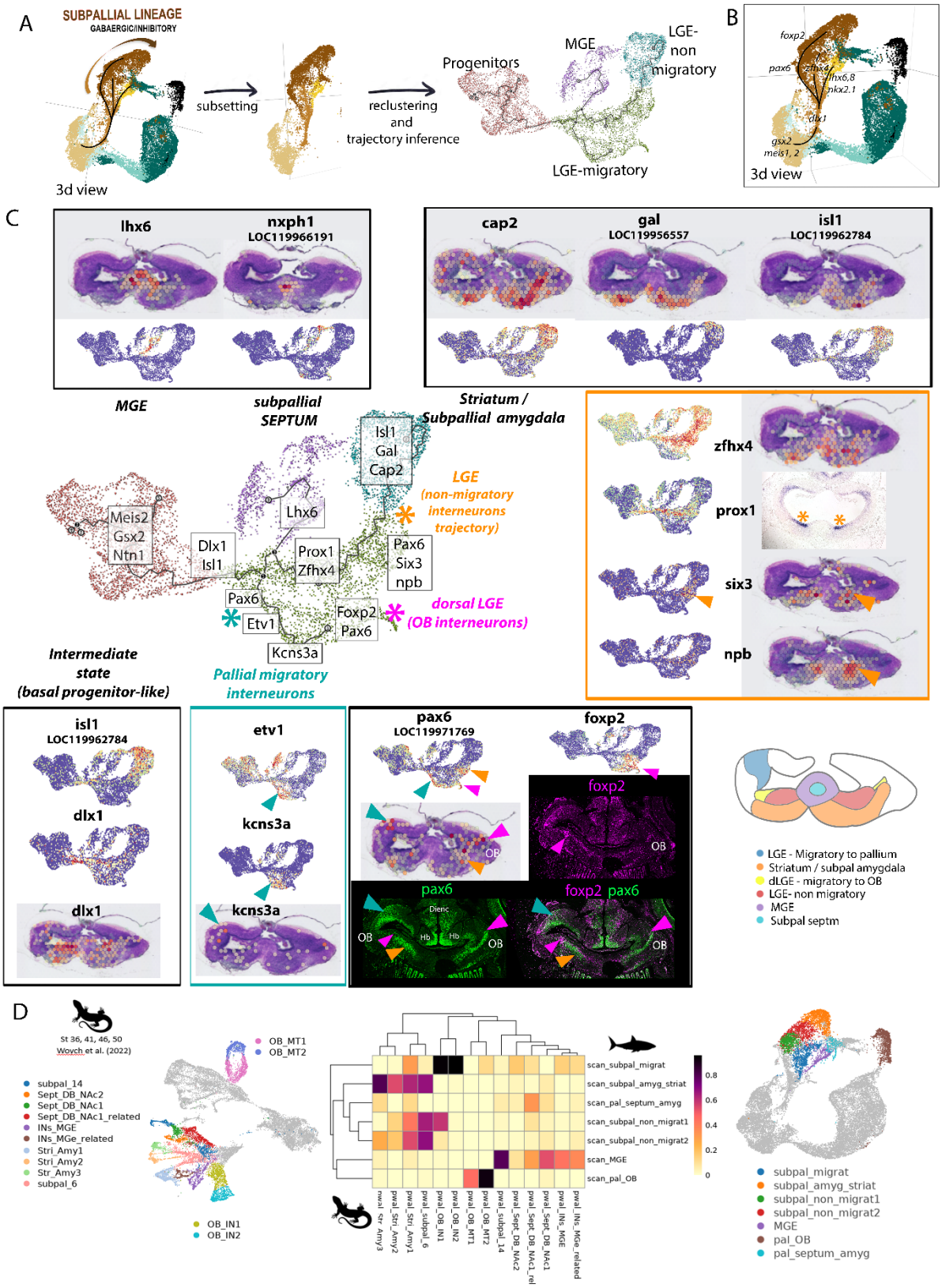
A
SUBPALLIAL LINEAGE
GABAERGIC/INHIBITORY
3d view
subsetting
reclustering
and
trajectory inference
Progenitors
MGE
LGE-
non
migratory
LGE-migratory
B
foxp2
pax6
zfhx4
lhx6,8
nkx2.1
dlx1
gsx2
meis1, 2
3d view
C
lhx6
nxph1
LOC119966191
cap2
gal
LOC119956557
isl1
LOC119962784
MGE
subpallial
SEPTUM
Striatum /
Subpallial amygdala
Isl1
Gal
Cap2
LGE
(non-migratory
interneurons
trajectory)
Meis2
Gsx2
Ntn1
Lhx6
Dlx1
Isl1
Prox1
Zfhx4
Pax6
Six3
npb
Pax6
Etv1
Foxp2
Pax6
dorsal LGE
(OB interneurons)
Kcns3a
Intermediate
state
(basal progenitor-like)
Pallial migratory
interneurons
zfhx4
prox1
six3
npb
isl1
LOC119962784
dlx1
dlx1
etv1
kcns3a
kcns3a
pax6
LOC119971769
foxp2
foxp2
OB
pax6
Dienc
OB
Hb
Hb
OB
foxp2 pax6
OB
LGE - Migratory to pallium
Striatum / subpal amygdala
dLGE - migratory to OB
LGE- non migratory
MGE
Subpal septm
D
St 36, 41, 46, 50
Wovch et al. (2022)
subpal_14
Sept_DB_NAc2
Sept_DB_NAc1
Sept_DB_NAc1_related
INs_MGE
INs_MGe_related
Stri_Amy1
Stri_Amy2
Str_Amy3
subpal_6
OB_MT1
OB_MT2
OB_IN1
OB_IN2
scan_subpal_migrat
scan_subpal_amyg_striat
scan_pal_septum_amyg
scan_subpal_non_migrat1
scan_subpal_non_migrat2
scan_MGE
scan_pal_OB
0.8
0.6
0.4
0.2
0
pwal_Str_Amy3
pwal_Str_Amy2
pwal_Stri_Amy1
pwal_subpal_6
pwal_OB_IN1
pwal_OB_IN2
pwal_OB_MT1
pwal_OB_MT2
pwal_subpal_14
pwal_Sept_DB_NAc2
pwal_Sept_DB_NAc1_rel
pwal_Sept_DB_NAc1
pwal_INs_MGE
pwal_INs_MGe_related
subpal_migrat
subpal_amyg_striat
subpal_non_migrat1
subpal_non_migrat2
MGE
pal_OB
pal_septum_amyg

Figure 4

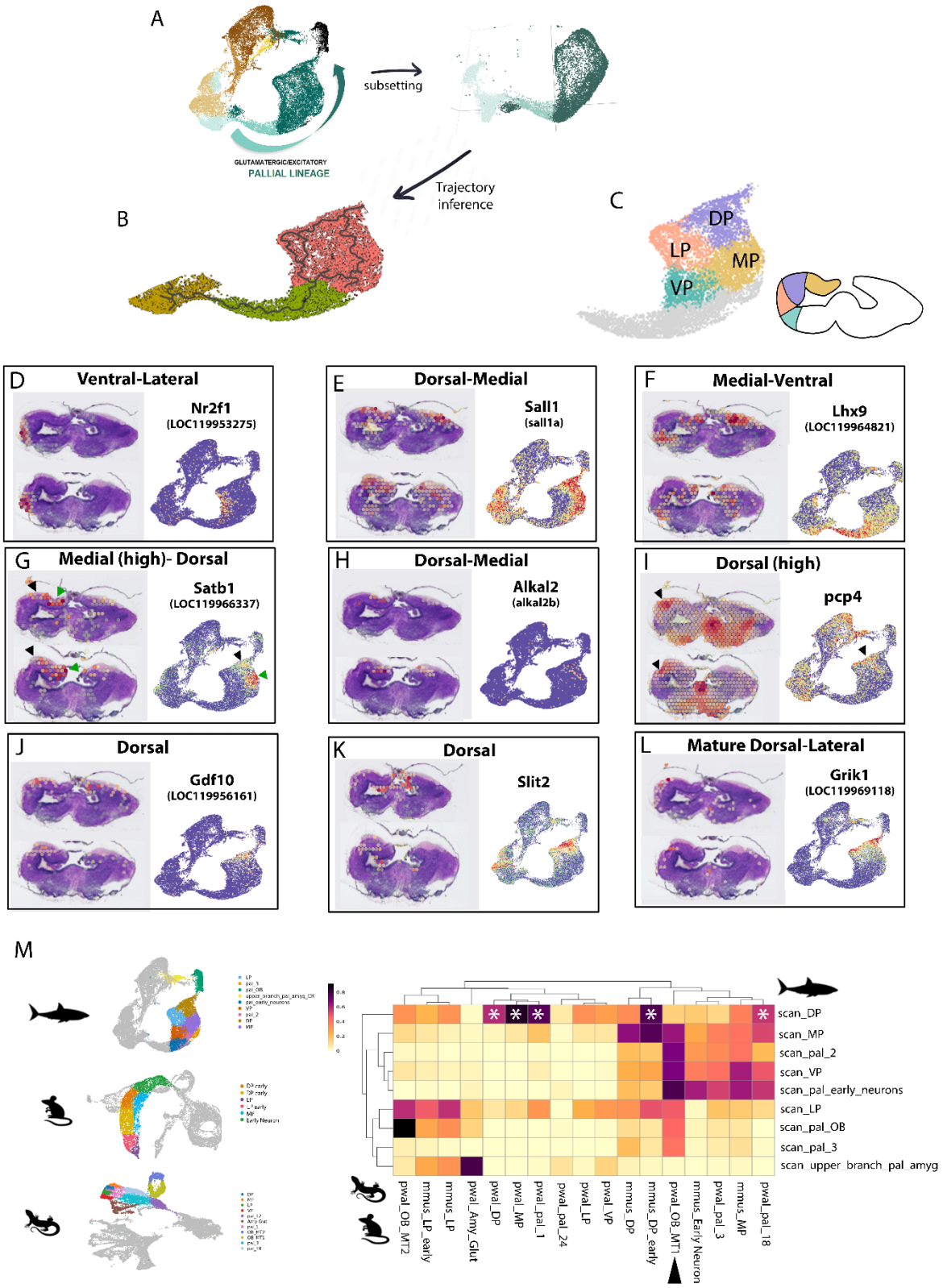
A
subsetting
GLUTAMATERGIC/EXCITATORY
PALLIAL LINEAGE
Trajectory inference
B
C
DP
LP
MP
VP
D Ventral-Lateral
Nr2f1
(LOC119953275)
E Dorsal-Medial
Sall1
(sall1a)
F Medial-Ventral
Lhx9
(LOC119964821)
G Medial (high)- Dorsal
Satb1
(LOC119966337)
H Dorsal-Medial
Alkal2
(alkal2b)
I Dorsal (high)
pcp4
J Dorsal
Gdf10
(LOC119956161)
K Dorsal
Slit2
L Mature Dorsal-Lateral
Grik1
(LOC119969118)
M
scan_DP
scan_MP
scan_pal_2
scan_VP
scan_pal_early_neurons
scan_LP
scan_pal_OB
scan_pal_3
scan_upper_branch_pal_amyg
pwal_OB_MT2
mmus_LP_early
mmus_LP
pwal_Amy_Glut
pwal_DP
pwal_MP
pwal_pal_1
pwal_pal_24
pwal_LP
pwal_VP
mmus_DP
mmus_DP_early
pwal_OB_MT1
mmus_Early Neuron
pwal_pal_3
mmus_MP
pwal_pal_18

Figure 5

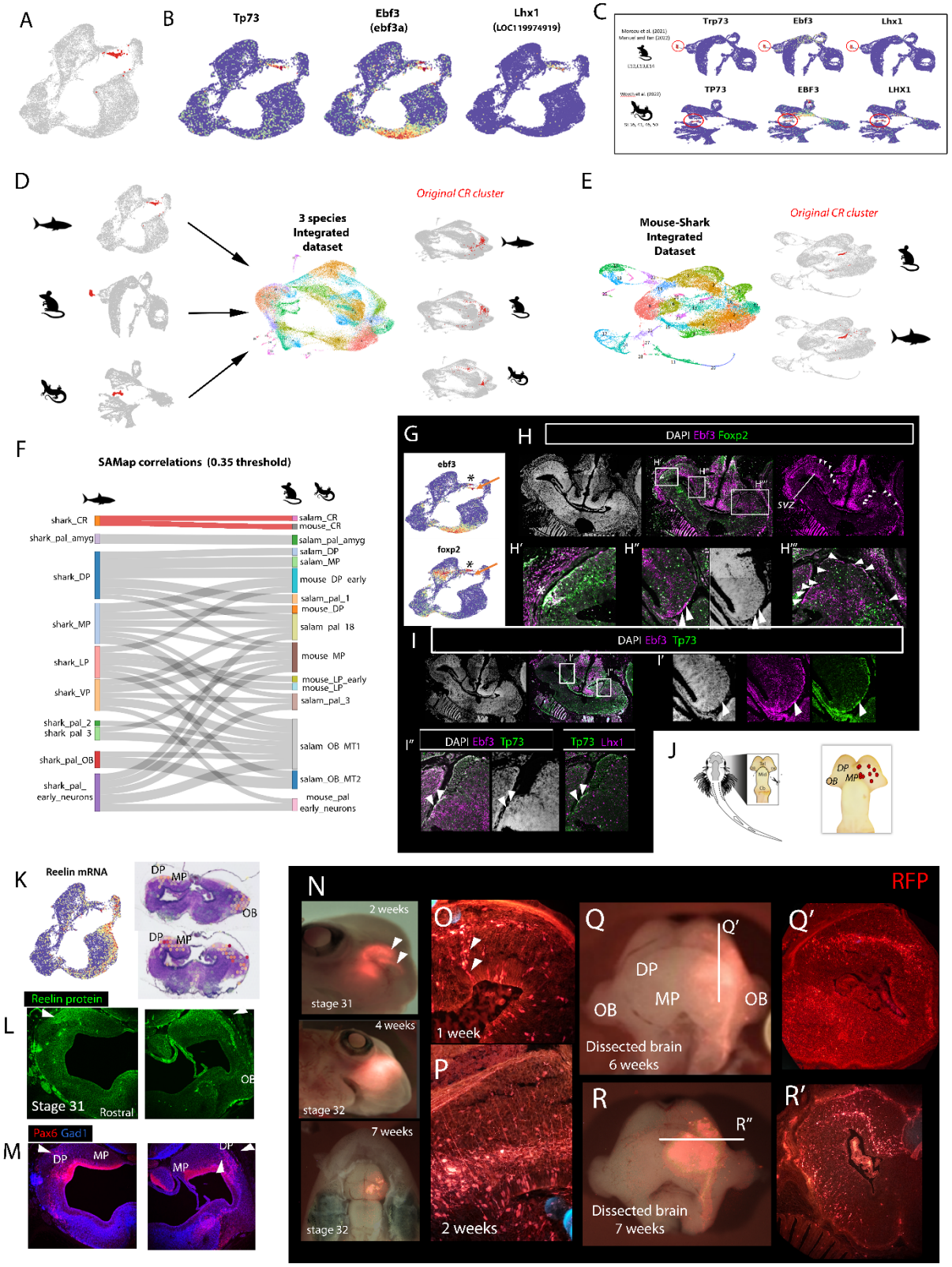
A
B
Tp73
Ebf3
(ebf3a)
Lhx1
(LOC119974919)
C
Trp73
Ebf3
Lhx1
TP73
EBF3
LHX1
D
3 species
Integrated
dataset
Original CR cluster
E
Mouse-Shark
Integrated
Dataset
Original CR cluster
F
SAMap correlations (0.35 threshold)
shark_CR
shark_pal_amyg
shark_DP
shark_MP
shark_LP
shark_VP
shark_pal_2
shark_pal_3
shark_pal_OB
shark_pal_
early_neurons
salam_CR
mouse_CR
salam_pal_amyg
salam_DP
salam_MP
mouse_DP_early
salam_pal_1
mouse_DP
salam_pal_18
mouse_MP
mouse_LP_early
mouse_LP
salam_pal_3
salam_OB_MT1
salam_OB_MT2
mouse_pal
early_neurons
G
ebf3
foxp2
H
DAPI Ebf3 Foxp2
SVZ
H′
H″
H‴
I
DAPI Ebf3 Tp73
I′
I″
DAPI Ebf3 Tp73
Tp73 Lhx1
J
DP
OB
MP
K
Reelin mRNA
DP
MP
OB
L
Reelin protein
Stage 31
Rostral
OB
M
Pax6 Gad1
DP
MP
N
RFP
2 weeks
stage 31
4 weeks
stage 32
7 weeks
stage 32
O
1 week
P
2 weeks
Q
Q′
DP
MP
OB
OB
Dissected brain
6 weeks
R
R″
Dissected brain
7 weeks
R′

Figure legends

Figure 1. Cell type atlas of the developing shark telencephalon. (A) Schema of the telencephalic dissections from three independent embryos at stages of neurogenesis 29, 30 and 31 (according to (*5*)). **(B)** (Left) Umap representation of the single nucleus dataset showing the main cell types and lineages identified (left). 3D Umap views (right). **(C)** Cell state score analysis reveals the presence of apical progenitors, basal progenitors, early neurons and late neurons depicting an axis of differentiation from left to right. **(D)** Pseudotime, Velocity and Cell Cycle score analysis confirm the presence of a developmental trajectory progression in our data. **(E)** Featureplots showing the expression of pallial, subpallial and olfactory bulb enriched genes. **(F)** (left) Fluorescent immunohistochemistry at stage 30 in a transverse section showing the expression of Pax6 and Gad1 antibodies. (right) Details of the pallium where Pax6 is expressed at high levels in the vz and gradually decreases its expression as neurons differentiate away from the ventricular surface; Gad1 is not expressed at the vz and strongly expressed in maturing subpallial neurons. **(G)** Featureplots showing the expression of pax6 in the pallial branch and gad1 in the subpallial branch matching that of the one showed in F. **(H)** Spatial expression of the same transcripts showed in E. **(I)** Gad1 protein expression in a transverse section through the catshark telencephalon. Arrowheads indicate gabaergic neurons migrating through pallial areas. Abbreviations: vz, ventricular zone. Scale bar: 200 µm

Figure 2. Profiling progenitor cell states in the catshark telencephalon. (A-B) Schematic drawing of direct neurogenesis (A) and indirect neurogenesis through basal progenitors (B). **(C)** Feature plots showing the expression of markers of the main types of apical progenitors. **(D)** Feature plots showing the sequential expression of pax6, eomes and tbr1 along the early pallial differentiation trajectory. **(E)** Tissue expression of Pax6 protein (left), eomes transcript (middle), pax6 and ebf3 transcripts (right). **(F)** Feature plot showing cell cycle phase scoring in the progenitors and pallial branch. Cells undergoing mitosis cluster together (red cluster). **(G)** Feature plots showing the expression of genes enriched in basal progenitor cluster. Arrowheads indicate expression of those markers within the mitotic cluster. **(H)** Cross-species comparison between mouse and shark developing telencephalic datasets using SAMap reveals a correlation trend across basal progenitors and apical progenitors of both species. **(I)** Lists of genes that show shared expression across species (up) and species specificity (down). **(J)** Study of apical progenitor variability after regressing the effect of the cell cycle. (Left) Cell cycle regression on progenitor clusters reveals variability along the dorso-ventral identity. (Right) Expression of genes showing variability along the dorso-ventral axis. Abbreviations: dLGE, dorsal-lateral ganglionic eminence; DMPal, dorsal-medial pallium; LGE, lateral ganglionic eminence; PSB,

pallial-subpallial border; SVZ, subventricular zone; VLPallium, ventral-lateral pallium; VZ, ventricular zone.

Figure 3. Single nucleus combined with spatial transcriptomics revealed subpallial trajectories, neuron subtypes and their anatomical positions in the shark developing telencephalon. (A) Subsetting of the subpallial lineage followed by reclustering and trajectory inference revealed three main trajectories. **(B)** 3D view of the umap showing the main subpallial trajectories and the genes upregulated along them. **(C)** (Centre) View of the three subpallial trajectories, main genes upregulated along them and how they correlate with anatomical positions. Expression of selected genes is shown in the umap alongside their spatial distribution (top, right and bottom). **(D)** Cross-species comparison with SAMap between shark and salamander developing telencephalic datasets. Abbreviations: dLGE, dorsal lateral ganglionic eminence; LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; OB, olfactory bulb.

Figure 4. Study of the pallial trajectory and pallial subtypes. (A) Schema showing the subsetting of the pallial branch from the main dataset. **(B)** Trajectory inference with Monocle3 in the pallial branch. **(C)** Schematic drawing showing clustering of the pallial mature neurons into four main cell identities and their anatomical localization in a transverse section. **(D-L)** Featureplots and spatiall distribution for selected gene that showed spatially-restricted expression into the different pallial areas. (M) Cross-species comparison across shark, salamander and mouse developing telencephalon datasets. (Left) Original datasets highlighting the pallial clusters included in the comparison; (Right) Heatmap showing correlation coefficients between the shark pallial clusters (rows) and mouse and salamander pallial clusters (columns). Abbreviations: DP, dorsal pallium; LP, lateral pallium; MP, medial pallium; VP, ventral pallium.

Figure 5. Characterization of a potential Cajal-Retzius (CR) homologous cell type in the shark embryonic telencephalon and identification of a layered strucrure in the pallium. (A) Umap of the whole shark dataset, highlighting in red the cluster of the putative CR cells. **(B)** Featureplot showing the expression of three genes enriched in the CR cluster, which are transcriptional signature of mammalian CRs. **(C)** Same three genes showed specific expression in the CRs cluster of mouse and salamander datasets. **(D-E)** Integration of three (D) and two (E) species dataset into one common shared space where genes from each species were converted to their respective orthogroups at the verterbate taxonomic level. **(F)** Sankey plot showing correlation across pallial mature clusters from shark (left column) versus mouse and salamander (right column). Highlighted in red is the correlation between the CR clusters across species. **(G-J)** Localization of CR cells in transverse sections of the embryonic brain at stage 31. (G) Feature plots showing the expression of ebf3 and foxp2. Asterisk indicate the upper cluster where both markers colocalize and corresponds to the pallial amygdala.

Orange arrow indicates the CR cluster, which expresses ebf3 but not foxp2. (H-H'') HCR in a transverse section of the shark telencephalon showing expression of ebf3 and foxp2. Insets in H represent the area showed at higher magnification in H'-H''. Arrowheads indicate the position of CR cells throughout the pial surface of the pallium. (I-I'') HCR in a transverse section of the shark telencephalon showing expression of ebf3 and tp73. Insets indicate the position if the area showed at higher magnification in I' and I''. Arrowheads point to cells showing colocalization of the markers indicated. (J) (Left) Schematic drawing of a dorsal view of a shark embryo at around stage 30-31 and the position of the brain. (Right) Representation of a dorsal view of the anterior embryonic shark brain. Red circles represent the position of the identified CR-like cells. **(K)** Reelin transcript expression in the single nucleus (left) and spatial dataset (right). **(L-M)** (L) Transverse section through the telencephalon of a stage 31 embryo showing the expression of Reelin (L) and Pax6-Gad1 (M) proteins. **(N-O)** Injection and electroporation of a CMV-RFP reporter plasmid in the telencephalon of stage 29 embryos labels a neuronal population in the dorsomedial telencephalon organized in a layer. (N) External view of the head of catshark embryos at different times after electroporation. Arrowheads show apparent streams of migrating neurons in the dorsal pallial area. (O-P) Sections through the pallium one week (O) and two weeks (P) after electroporation shows RFP-positive neurons emerging from the ventricular zone (arrowheads) and extending tangentially into an incipient layer. (Q-R') 6-7 weeks after electroporation, strong RFP labelling appears in the dorsal telencephalon forming a noticeable neuronal layer visible in both saggital (Q') and transverse (R') sections. Lines in Q and R indicate the approximated section level of Q' and R', respectively. Abbreviations: Cb, cerebellum; DP, dorsal pallium; Mid, midbrain; MP, medial pallium; OB, olfactory bulb; Tel, telencephalon.

Materials and Methods

Animal use

Eggs of *S. canicula* were supplied by the Marine Biological Model Supply Service of the CNRS UPMC Roscoff Biological Station (France). Upon arrival, eggs were kept in sea water tanks under controlled conditions of 16°C, pH 7.5-8.5 and salinity of 35 g/L. Suitable measures were taken to minimize animal pain or discomfort.

Embryos were deeply anesthetized with 0.5% tricaine methane sulfonate (MS-222, Sigma, St. Louis, MO) in filtered sea water before proceeding with dissection or fixation. Embryos were staged based on external features following **(*5*)**. For fixation, embryos were decapitated and immersed in 4% paraformaldehyde (PFA) in PBS, overnight at 4°C.

All animal experiments were carried out according to the guidelines of the Committee for Animal Welfare and Institutional Animal Care and Use (IACUC) under EMBL's Policy on the Protection and Welfare of Animals Used for Scientific purposes (IACUC code 21-010_HD_NP).

Sample collection and nuclei preparation

Deeply anaesthetized embryos were staged and immersed in cold sea water containing 0.5% of MS222. After telencephalic dissections, the tissue was flash-frozen in an Eppendorf tube immersed in liquid nitrogen and stored at -80°C. Single nuclei were isolated from fresh frozen tissue using a modified protocol from **(*29*)**. Dissected tissue was placed in 1.5 ml Eppendorf LoBind tubes, and flash-frozen in liquid nitrogen. Frozen tissues were then placed on ice and homogenized in 300 µl of HB buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl pH 8.0, 0.1% IGEPAL, 1 µM DTT, 0.4 U/µl Murine RNase Inhibitor [NEB], 0.2 U/µl Superase-In [ThermoFisher]) using micro pestles (Axygen). Homogenization was performed on ice and completed within 5 minutes Unlysed tissue fragments were pelleted by centrifugation at 100 × g for 1 minute at 4 °C. The supernatant containing nuclei was transferred to a new LoBind tube and centrifuged at 400 × g for 4 minutes at 4 °C. After discarding the supernatant, the pellet was washed and resuspended in 300 µl of fresh HB buffer and nuclei were pelleted again at 400 × g for 4 minutes at 4 °C. Finally, the nuclei pellet was resuspended in 50 to 100 µl of ice-cold Dulbecco's PBS supplemented with 1 µM DTT, 0.4 U/µl Murine RNase Inhibitor (NEB), and 0.2 U/µl Superase-In (ThermoFisher). The suspension was passed through a 40 µm Flowmi filter to remove aggregates. Nuclei quality and concentration were assessed by staining with SYBR Safe DNA stain and imaging on a Zeiss Axio Imager using a C-Chip hemocytometer (Neubauer improved).

Single Nuclei Capture and Library Preparation

Single nuclei capture and library preparation were performed using the Chromium Single Cell 3' Gene Expression Kit (v3.1 chemistry) and the Chromium Controller (10x Genomics), following the manufacturer's protocol with slight modifications. Approximately 15,000 nuclei were loaded per capture channel. cDNA amplification was performed with 15 cycles for samples at stage 29–30 and 12 cycles for samples at stage 30–31, using a 3-minute extension time. Final libraries were quantified using a Qubit fluorometer, and fragment sizes were assessed using an Agilent Bioanalyzer. Libraries were sequenced with 250 million PE reads on an Illumina NextSeq 2000 using the following read configuration: 28 cycles for Read 1, 10 cycles each for the i5 and i7 indices, and 90 cycles for Read 2.

Visium 10X Spatial Transcriptomics and Library Preparation and data processing

One fresh dissected telencephalon at stage 31, was embedded in OCT mounting medium, cryosectioned into 14µm transverse sections and placed on a Visium Spatial Gene Expression slide from 10X Genomics. Samples were processed according to manufacturer's instructions. Briefly, slides were fixed in methanol and stained with Hematoxylin and Eosin. Images were taken with a Nikon Ti2 at 20x magnification. The slide was processed using the Visium Spatial Gene Expression Reagent kit. Permeabilization time was tested and determined the optimal at 8 minutes.

cDNA libraries were amplified using 12 cycles for samples at stage 30–31 with a 1-minute extension time. Final libraries were quantified using a Qubit fluorometer, and fragment sizes were assessed using an Agilent Bioanalyzer. Sequencing was performed on an Illumina NextSeq 2000 with the following read configuration: 28 cycles for Read 1, 10 cycles each for the i5 and i7 indices, and 90 cycles for Read 2 with 200 million PE reads per library.

Spatial transcriptomics data was processed using the Space Ranger (v2.0.1) software (10x Genomics) with default parameters. Raw sequencing data were aligned to the *S.canicula* reference genome (see above), and spatially resolved gene expression matrices were generated. Quality control, normalization, and downstream analyses, including clustering and visualization of spatial gene expression patterns, were performed using Seurat (v5) in R. Default workflows and parameter settings were used for data filtering, dimensionality reduction, clustering, and spatial mapping.

Spatial transcriptomics data generated using the 10x Genomics Visium platform were processed and analyzed in R (v4.3.1) using the Seurat package (v5.1.0). Following standard quality control and normalization procedures, spatially enriched genes were identified using the FindSpatiallyVariableFeatures() function with default parameters. This method detects genes

exhibiting significant spatial variability across tissue sections by leveraging spatial autocorrelation statistics, thereby highlighting genes with non-random spatial expression patterns. The resulting list of spatially enriched genes was used for downstream analyses and interpretation of spatial gene expression architecture within the tissue.

Genome alignment, scRNA-seq preprocessing and quality control:

A mapping index was created with CellRanger v6.1.1 using the S. canicula v1.1 genome (GCF_902713615) with default parameters. Demultiplexed FASTQ files from snRNA-seq were mapped against the index with "cellranger count" using default parameters.

Single nucleus RNA sequencing data processing

Data processing was performed using Seurat (v5.1.0) in R (v4.3.1). Raw count matrices from three independent experiments were initially processed separately. Following quality control, individual Seurat objects were created and log-normalized. These objects were subsequently merged while preserving experiment-specific metadata identifiers. To address batch effects, we implemented a dual integration strategy: first applying SCTransform normalization with glmGAMPOI variance stabilization while regressing out mitochondrial content, followed by Harmony batch correction using experiment origin as the batch covariate across the first 50 principal components. The integrated dataset was clustered using the Louvain algorithm via the FindClusters() function, with subsequent dimensionality reduction and visualization using UMAP. Differential gene expression analysis between clusters was conducted through Wilcoxon rank-sum tests (minimum 25% cells expressing genes in either cluster, log-fold change threshold of 0.25, unless otherwise specified), followed by cell type annotation.

RNA velocity

Raw reads of the 10x sequencing libraries were mapped individually using STARsolo (v2.7.9)(*30*) with the soloFeature Velocyto following the recommendations for 10x libraries (--soloMultiMappers Uniform --soloCBmatchWLtype 1MM_multi_Nbase_pseudocounts --soloUMIfiltering MultiGeneUMI_CR --soloUMIdedup 1MM_CR --soloCellFilter EmptyDrops_C). Spliced and Unspliced data matrices obtained were combined across all samples and loaded into python to run scVelo (*31*) using 30 PCs and 30 neighbours for the moments' calculation. The resulting velocity fields were projected onto the existing UMAP embedding to visualize the predicted future states and lineage relationships of individual nuclei. This analysis provided additional insight into differentiation trajectories and complemented pseudotime inference from Monocle3.

Monocle trajectory inference

Single nucleus RNA-seq data were analyzed using Seurat (v5) and Monocle3 to reconstruct developmental trajectories. After preprocessing and clustering in Seurat, the dataset was converted to a Monocle3 cell_data_set object. Cluster identities, UMAP embeddings, and cell metadata from Seurat were transferred to Monocle3. Cells were clustered and a trajectory graph was learned using UMAP-based dimensionality reduction, with parameters adjusted to optimize graph structure. Cells were ordered in pseudotime, and differential gene expression along trajectories was identified using Monocle3's graph_test function. Visualization of gene expression dynamics and pseudotime was performed with plot_cells and FeaturePlot.

AP diversity study

Apical progenitor populations were subsetted based on cluster identities, and data normalization, scaling, and dimensionality reduction were performed using SCTransform and principal component analysis (PCA). Re-integration was performed using Harmony, followed by UMAP embedding, neighborhood graph construction, and clustering. Cell cycle effects were regressed out using ScaleData() with the vars.to.regress argument set to S and G2M scores. Marker genes for each cluster were identified using differential expression analysis. For trajectory inference, the processed Seurat object was converted to a Monocle3 cell_data_set using as.cell_data_set(). Cluster identities, UMAP embeddings, and metadata were transferred to the Monocle3 object, and cells were reclustered with cluster_cells() using UMAP reduction. Principal graph learning was performed with learn_graph(), with parameters such as ncenter, minimal_branch_len, prune_graph, and geodesic_distance_ratio adjusted for optimal trajectory structure. Cells were ordered in pseudotime using order_cells(), and pseudotime values were extracted with pseudotime(). Differentially expressed genes along pseudotime and at branch points were identified using graph_test(), and gene modules were detected with find_gene_modules(). Visualization of gene expression dynamics and trajectories was performed using plot_cells(), plot_genes_in_pseudotime(), and ggplot2-based plotting functions.

Trajectory analysis using Monocle3

Clusters of interest were subsetted and re-clustered using the Louvain algorithm in Seurat (v5.1.0) to refine population heterogeneity. Trajectory inference was performed with Monocle 3 (v1.3.4) by converting the subsetted Seurat object into a CellDataSet using default parameters. A principal graph was learned over the precomputed UMAP space with learn_graph(), and pseudotime ordering was calculated by specifying a root node at the progenitor-enriched. Branch-specific differentially

expressed genes were identified using graph_test() with Moran's I statistic to prioritize spatially coherent transcriptional changes along inferred differentiation trajectories.

Cell state and cell cycle scoring analysis

Cell cycle phase assignment was performed using the CellCycleScoring() function in Seurat (v5.1.0), which calculates S and G2/M phase scores for each cell based on the expression of canonical marker genes. For this analysis, lists of S and G2/M phase genes originally curated in mouse were translated to their shark orthologs using OrthoFinder (v2.5.4) and used as input. Cells were classified into G1, S, or G2/M phases according to their highest phase score, and these assignments were retained as metadata for downstream analyses. Similarly, cell state scores were computed using the AddModuleScore() function in Seurat, with gene sets representing four cell states (apical progenitor, basal progenitor, early neuron, late neuron) curated from a published mouse study **(*8*)** and translated to shark orthologs using OrthoFinder. The resulting module scores quantified the relative enrichment of each cell for the selected gene sets, facilitating the interpretation of cellular heterogeneity in terms of both cell cycle phase and developmental cell state.

Cell cycle regression and AP dorso-ventral variability

To further dissect AP diversity independent of cell cycle effects, we isolated AP clusters-including mitotic basal progenitors (BPs)-and performed reclustering using Seurat (v5.1.0). To minimize confounding transcriptional variability attributable to cell cycle phase, we regressed out S and G2/M/S scores -calculated in the previous stage- using the ScaleData() function, following standard Seurat cell cycle regression protocols. Dimensionality reduction and visualization were then repeated on the corrected expression matrix. This approach revealed transcriptional heterogeneity along the pallial-subpallial axis, with cells from the original M-phase clusters redistributing according to their cell type identities rather than proliferative state.

Cross-species comparisons

Mouse object was created by integrating two published datasets from **(*8*, *17*)**. Briefly, we aligned the reads against the mouse GRCm38 genome, created separate Seurat objects, merged them and integrated them following the same pipeline indicated above. Raw data salamander objects from different stages were obtained from **(*32*)** and aligned to the published transcriptome from the same publication. We merged and integrated the objects following the same pipeline and parameters indicated above.

Cross-species comparisons using common orthogroups

We estimated the transcriptome of the last vertebrate common ancestor by assigning each modern gene to an orthogroup at the vertebrate taxonomic level with the emapper tool (v.2.1.3) (*33*) and convert each modern animal's count matrix from a matrix of counts of each modern gene per cell, to a matrix of counts of each orthogroup per cell. In some cases, true lineage-specific gene duplications will therefore be assigned the same vertebrate orthogroup, and in such cases the counts for the true genes are summed to estimate the expression of the orthogroup. While this can obscure some of the differences between closely-related, recently-derived modern cell types that rely on novel gene duplications in the modern animal, it also enables us to focus on the similarities between species, whose common ancestor likely had a simpler set of cell types relative to either descendant lineage.

Proteomes for each animal were collected by and using each animal's genome resources (mouse GRCm38 and salamander transcriptome from (*32*)) we kept only the longest protein encoded per gene, such that we had a single protein per gene. In the resulting simplified proteome, we then renamed each protein identifier to match the gene identifier that was used in the single cell RNAseq count matrix. We ran this simplified proteome through the emapper software (v.2.1.3) (*33*) with the command *emapper.py -m mmseqs -i* . For each unique peptide in the emapper output matrix, we obtained a gene identifier that matched the genes reported in the single cell count matrix of the relevant animal, and also an orthogroup assignment at a range of taxonomic levels, including for the vertebrate orthogroups relevant for this study. We used this information to relate genes to orthogroups for both species by creating a simple lookup table, with one column for genes and a second column with their orthogroup assignment.

Cross-species comparison with SAMap

The single-cell objects were converted to the H5AD format using the "sceasy" R library (v0.0.7) and their metadata (gene and cell IDs) lightly edited to ensure cross-compatibility with SAMap. Pairwise all-against-all sequence alignments of the peptide sequences of the three species were conducted using MMseqs2 (commit 67949d7, cite). Pairwise SAMap comparisons between shark and mouse as well as shark and salamander were performed (SAMap v1.0.15, cite) using the previously established consensus annotation. The mapping tables of shark to mouse/salamander clusters were extracted from the corresponding SAMap object and concatenated with each other. The resulting matrix was visualized as simple and clustered heatmaps using seaborn v0.13.2 (cite) and plotly v6.0.1 (cite), with euclidean, cosine, and correlation distance measures.

Immunofluorescence

Three primary antibodies were used in this study. GAD (Glutamic acid decarboxylase) Polyclonal sheep anti-GAD65/67 provided by Dr. E. Mugnaini; Pax6 Cat. no. PRB-278P Polyclonal rabbit anti-Pax6 Covance, Emeryville, CA; RELN (Reelin; clone 142) Cat.no. MAB5366 Monoclonal mouse anti-RELN Millipore, Billerica, MA. Briefly, cocktail of primary antibodies at appropriate concentrations (GAD, 1:30000; Pax6 (1:200); Reelin (1:150) were incubated in PBS containing 0.5% Tween (PBST), overnight at room temperature. Secondary antibodies were incubated at a 1:100 concentration in a humid chamber for 2 hours at room temperature in the dark. We used Alexa 488-conjugated donkey anti-mouse; Alexa 546-conjugated donkey anti-rabbit. All of them from Molecular probes Eugene, OR. Sections were rinsed in distilled water, allowed to dry and mounted in MOWIOL 4-88 Reagent (Calbiochem, MerkKGaA, Darmstadt, Germany).

HCR

Cryosectioned tissue slides were thawed for 15 minutes at room temperature. Pre-hybridization was performed at 37°C for 30 minutes in a humidified chamber using pre-hybridization buffer supplemented with 10 mM VRC RNase inhibitor (5 µl per 100 µl buffer). Hybridization was carried out at 37°C with 8 nM probe concentration and 10 mM VRC RNase inhibitor in a humid chamber. Post-hybridization washes were performed at 37°C with decreasing ratios of wash buffer to 5xSSCT, followed by equilibration in 5xSSCT at room temperature. For signal amplification, hairpins were snap-cooled (2 µl of each 3 µM stock hairpin per 100 µl amplification buffer for a final 60 nM concentration), heated at 95°C for 90 seconds, and cooled to room temperature in the dark for 30 minutes. Amplification was performed by incubating slides overnight at room temperature in a humidified chamber with 100 µl of amplification buffer containing the hairpins. Slides were then washed in 5xSSCT, counterstained with DAPI (5 µg/ml final concentration, 10 minutes), washed, and mounted in Prolong Gold antifade (ThermoFisher scientific).

Plasmid injection an electroporation

Embryos at stages 29 and 30 were removed from their egg cases, deeply anaesthetized as indicated above and placed in cold filtered sea water containing anaesthetics. We used a plasmid harbouring a CMV enhancer and Beta-acting promoter driving the expression of RFP. We injected 1 µl of plasmid dilution in the telencephalic ventricle (plasmid concentration of 200ng/µl) and electroporated into the telencephalic ventricular wall using flat electrodes flaking the head area. Electroporation settings (previously tested for maximizing efficiency and reducing damage) were as follows: 25V, pulse length 5 msec, interval 500 msec, 4 pulses, decay rate 10%, positive polarity (poring pulse); 10V, pulse length 50 msec, pulse interval 500 msec, 5 pulses, decay rate 40%, positive polarity (transfer pulse). We used a NepaGene electroporator (Nepa21 Type II).

After electroporation, embryos were immersed in oxygenated sea water and monitor their recovery. They were regularly photographed using a Leica M205 FCA (optics carrier – 12803510). Embryos were eventually deeply anaesthesized, and processed for cryosectioning.
Fluorescent images were taken in a Stellaris 8 confocal microscope.

Data/code availability

Code used for RNA mapping and SAMap comparison is available at http://git.embl.de/npapadop/scan-scrnaseq and archived at https://doi.org/10.5281/zenodo.15233987. The cross-species comparison data is available at https://doi.org/10.5281/zenodo.15234059.