

# Imaging and multi-omics datasets converge to define different neural progenitor origins for ATRT-SHH subgroups

---

Received: 10 June 2022

---

Accepted: 9 October 2023

---

Published online: 20 October 2023

---

 Check for updates

---

María-Jesús Lobón-Iglesias<sup>1,18</sup>, Mamy Andrianteranagna<sup>1,2,18</sup>, Zhi-Yan Han<sup>1,18</sup>, Céline Chauvin<sup>1</sup>, Julien Masliah-Planchon<sup>3</sup>, Valeria Manriquez<sup>4</sup>, Arnault Tauziede-Espariat<sup>5,6</sup>, Sandrina Turczynski<sup>1</sup>, Rachida Bouarich-Bourimi<sup>1</sup>, Magali Frah<sup>1</sup>, Christelle Dufour<sup>7</sup>, Thomas Blauwblomme<sup>8</sup>, Liesbeth Cardoen<sup>9</sup>, Gaelle Pierron<sup>3</sup>, Laetitia Maillot<sup>3</sup>, Delphine Guillemot<sup>3</sup>, Stéphanie Reynaud<sup>3</sup>, Christine Bourneix<sup>3</sup>, Célio Pouponnot<sup>10</sup>, Didier Surdez<sup>11,12</sup>, Mylene Bohec<sup>13</sup>, Sylvain Baulande<sup>13</sup>, Olivier Delattre<sup>3,11</sup>, Eliane Piaggio<sup>4</sup>, Olivier Ayrault<sup>10</sup>, Joshua J. Waterfall<sup>14,15</sup>, Nicolas Servant<sup>2,19</sup>, Kevin Beccaria<sup>8,19</sup>, Volodia Dangouloff-Ros<sup>16,19</sup> & Franck Bourdeaut<sup>1,17,19</sup> ✉

Atypical teratoid rhabdoid tumors (ATRT) are divided into MYC, TYR and SHH subgroups, suggesting diverse lineages of origin. Here, we investigate the imaging of human ATRT at diagnosis and the precise anatomic origin of brain tumors in the Rosa26-Cre<sup>ERT2</sup>::Smarb1<sup>flox/flox</sup> model. This cross-species analysis points to an extra-cerebral origin for MYC tumors. Additionally, we clearly distinguish SHH ATRT emerging from the cerebellar anterior lobe (CAL) from those emerging from the basal ganglia (BG) and intra-ventricular (IV) regions. Molecular characteristics point to the midbrain-hindbrain boundary as the origin of CAL SHH ATRT, and to the ganglionic eminence as the origin of BG/IV SHH ATRT. Single-cell RNA sequencing on SHH ATRT supports these hypotheses. Trajectory analyses suggest that *SMARCB1* loss induces a de-differentiation process mediated by repressors of the neuronal program such as *REST*, *ID* and the NOTCH pathway.

Atypical teratoid rhabdoid tumors (ATRT) are rare and aggressive malignancies of the central nervous system (CNS) affecting infants and young children. They are characterized by a biallelic inactivation of *SMARCB1* tumor suppressor gene in an otherwise very stable genome<sup>1</sup>. Based on both methylation and expression profiling, recent studies have highlighted the molecular diversity of these tumors<sup>2,3</sup>, which are now divided into at least three subgroups, i.e., the so-called MYC-, TYR-, and SHH-ATRTs<sup>4</sup>. This molecular diversity suggests multiple lineages of origin for each sub-type, though none of these origins is definitively identified at the present time. In this respect, the analysis of expression profiling has been weakly informative, giving at most some insights on the lineage (neurogenic or melanogenic features for the SHH and TYR subtypes respectively) or on some recurrently expressed

developmental genes and pathways (SHH and NOTCH pathway for the SHH subtype; HOX clusters for the MYC subtype). Methylation profiling analyses including ATRT as well as extra-cranial rhabdoid tumors (ECRT) have revealed that MYC ATRT and ECRT tend to cluster together, apart from TYR and SHH ATRT, suggesting that MYC ATRT may share some commonalities with tumors that emerge outside of the brain<sup>5</sup>.

These results are consistent with single-cell RNA sequencing (scRNAseq) results failing to find any convincing correlation between tumor cells from MYC ATRT and any known cell types from normal brain development<sup>6</sup>. These data were consistent with our own findings in an inducible mouse model, where gene expression profiling of mouse Myc ATRT suggested a neural-crest-derived lineage<sup>7</sup>. Likewise,

---

A full list of affiliations appears at the end of the paper. ✉ e-mail: [franck.bourdeaut@curie.fr](mailto:franck.bourdeaut@curie.fr)

a genotyping approach, comparing tumor cells with adjacent normal tissues, suggested that ECRT derive from ancestors shared with Schwann cells, thus pointing to a neural-crest-derived origin<sup>8</sup>. These results were also partly supported by the development of typical ATRT-like tumors in *PO-Cre::Smad1<sup>fllox/fllox</sup>* mouse, again pointing to neural-crest cells as putative cells of origin<sup>9</sup>. Finally, applying scRNAseq to murine tumors, Graf et al. have suggested that primordial germ cells could be the origin for at least a subset of murine *Myc* ATRT<sup>10</sup>. Altogether, while a neuronal origin is considered for SHH ATRT and a non-neuronal origin for *MYC* ATRT, the question of the lineage of origin of each ATRT subtype remains mostly speculative.

In the present study, we thoroughly investigated the imaging at diagnosis of a series of primary ATRT, aiming to describe the epicenter and thereby, the site of origin for each case. We next correlated these findings with the molecular profiles, including the new imaging results in the definition of the subgroups together with methylation and expression profiling. Finally, we performed a similar approach on tumors from the *Rosa26-Cre<sup>ERT2</sup>::Smad1<sup>fllox/fllox</sup>* model, and added a single-cell RNAseq analysis on both human and mouse tumors to investigate the putative cells of origin for the various ATRT subtypes.

## Results

### Radiological description of ATRTs' epicenter suggests clearly distinct origins for each molecular subtype

We first reviewed a series of 54 human brain ATRT (Supplementary Fig. 1A; Supplementary Data 1), aiming to categorize them following the classically used locations, i.e. infratentorial or supratentorial. Using the online DKFZ brain tumor DNA methylation-based classifier tool<sup>11</sup> (<http://www.moleculareuropathology.org/>), we also assigned a molecular subgroup to this series and found that, while the infratentorial location is not per se suggestive of any tumor subgroup, the supratentorial location could suggest both *MYC* or SHH subtypes (Fig. 1A, B; Supplementary Data 1). For a higher-resolution analysis of anatomic location, we then categorized the tumors from their presumptive epicenter, aiming to give a more precise anatomic origin. This led us to define eight anatomical categories of ATRT: (i) cranial nerves, tumors from extra-axial structures such as interpeduncular cistern (IIIrd cranial nerve), internal acoustic canal (VII/VIIIth cranial nerves), cavernous sinus (IIIrd, IVth and Vth nerves) and jugular foramen, (IX/X/XIth nerves); (ii) cerebellar anterior lobe (CAL), tumors spreading from the quadrigeminal cistern to the anterior vermis; (iii) tumors from the middle cerebellar peduncles and inferior cerebellar vermis (MCP/ICV); (iv) peripheral tumors located in the cerebral cortex, pressing the normal parenchyma towards the ventricles; (v) intraventricular (IV) tumors, which were often large tumors in close relationship with the basal ganglia region; (vi) basal ganglia (BG) tumors, centered on the basal ganglia, pushing the brain parenchyma to the periphery; (vii) septal tumors, located in the interventricular septum; and finally, (viii) spinal cord tumors (Supplementary Fig. 1B). Overall, we ended up with four locations belonging to the infratentorial region and four locations belonging to the supratentorial region (Fig. 1C).

Combining these detailed anatomic locations with the DNA methylation subgroups, we found different localization patterns for each molecular subtype. Specifically, (i) almost all *TYR* ATRT emerged from the middle cerebellar peduncle and the inferior cerebellar vermis; (ii) the SHH subgroup was composed of tumors located in the CAL, the BG and the intraventricular region; and (iii) the *MYC* subgroup was mainly composed by tumors located in the cranial nerves, in the cerebral cortex and the spinal cord (Supplementary Fig. 1C). Altogether, these findings showed a correlation between anatomical location and molecular subgroups that suggested different lineages of origin. To further investigate this, we performed unsupervised analyses (UMAP and hierarchical clustering) on this DNA methylation dataset. We found that ATRT

samples clustered according to their anatomical location (Figs. 1D, E). Furthermore, the SHH subgroup seemed to be composed of two distinct sub-clusters: one composed almost exclusively of tumors from CAL and the other mainly composed of BG and IV tumors (Fig. 1E).

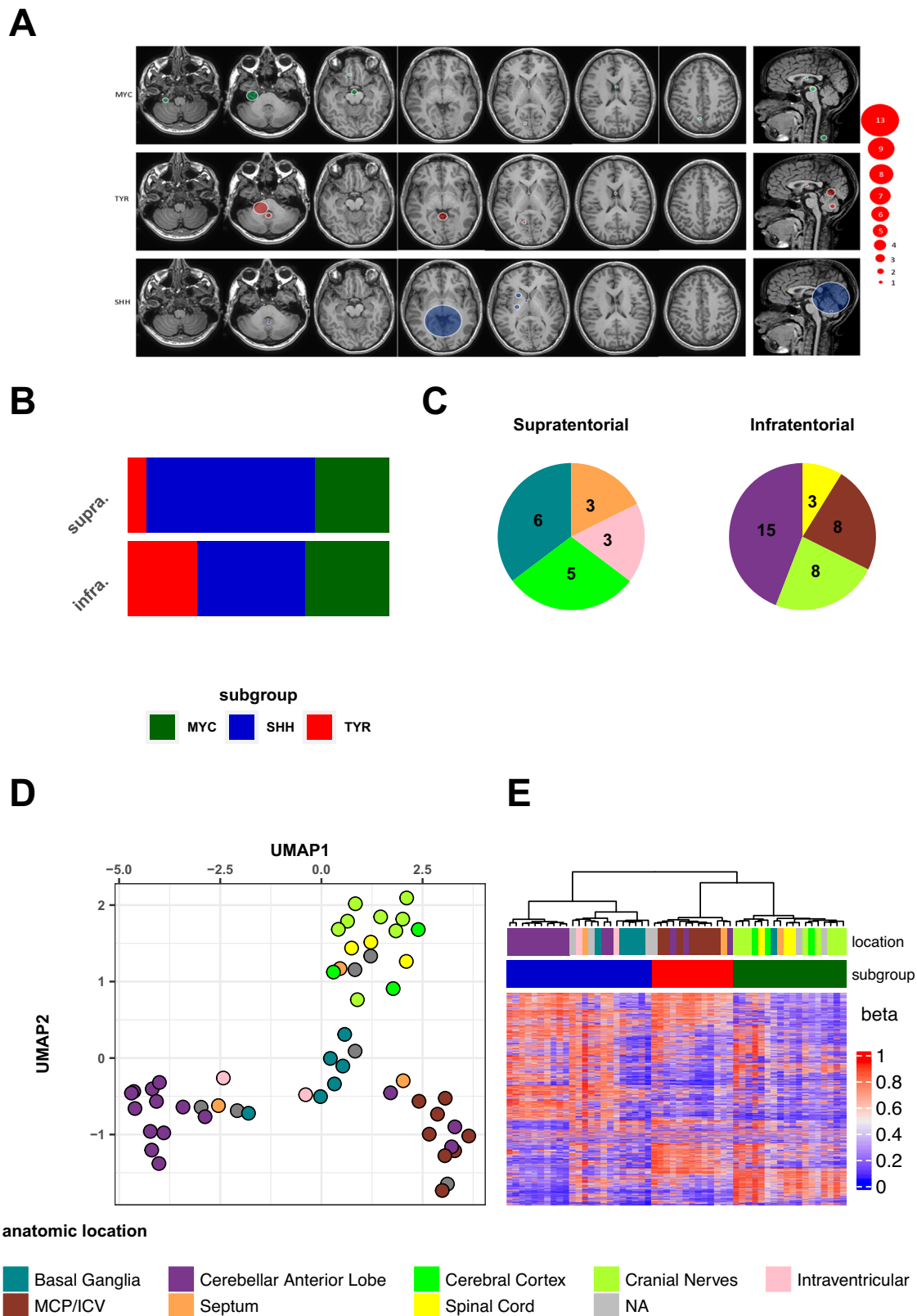
### Integrative analysis identified four anatomical-molecular subgroups and splits SHH ATRT in two subgroups with distinct anatomical locations and transcriptional profiles

Next, we sought to explore how the anatomically defined ATRT classification behaved at the transcriptomic level. We first performed an unsupervised analysis on the RNAseq data from 49 samples based on the 2000 most variable genes. Hierarchical clustering showed three molecular subgroups that corresponded to *TYR*, SHH, and *MYC* (Fig. 2A).

To achieve a more comprehensive subgrouping, we next performed a data integration approach using the three layers of information: DNA methylation, gene expression, and anatomical location. We then performed consensus clustering based on the transcriptomic dataset (Fig. 2B, C, Supplementary Fig. 2A). Assuming the three subtypes, we found some discrepancy between methylation and RNAseq subgroupings (Fig. 2B); however, the consensus clustering solved this discrepancy with  $k = 4$  by separating BG/IV (SHH methylation signature but *MYC* gene expression profile) from CAL SHH ATRT (SHH by methylation and transcriptomic signatures) and *MYC* (*MYC* by both methylation and transcriptomic signatures) (Fig. 2C). Sparse Partial Least Square Determinant Analysis (sPLS-DA) applied to RNAseq also supported the four subgroups (Fig. 2D, Supplementary Fig. 2B) and allowed us to identify genes that were specific to each subgroup (Supplementary Data 2). Consistently, a kernel-based data integration method<sup>12</sup>, aiming to combine gene expression and DNA methylation datasets before unsupervised analysis, confirmed the existence of four molecular subgroups that fitted with our anatomical classification (Fig. 2E). These results established that the SHH ATRT, canonically defined by DNA methylation, could be divided in two distinct subgroups, with distinct gene expression profiles and specific anatomic locations.

### Genetically engineered mouse models reveal specific anatomical origins for murine *Myc* and *Shh* ATRT

We next aimed to investigate how these four anatomical-molecular subgroups, based on anatomic location and molecular profiling, were also relevant in our previously published mouse model of ATRT. We developed a mouse model (*Rosa26-Cre<sup>ERT2</sup>::Smad1<sup>fllox/fllox</sup>*) which, by an inducible inactivation of *Smad1* in unrestricted cell types, gives rise to *Myc* and *Shh* subtypes of ATRT<sup>7</sup> (Supplementary Data 3). To obtain more robust correlations between tumor molecular subtypes and anatomical sites of origin, we generated an increased number of tumors and profiled them by RNAseq and gene expression microarray. Using unsupervised hierarchical clustering on these datasets, we first confirmed the two *Shh* and *Myc* subgroups previously described. We thereby found that intracranial *Myc* clustered with the extracranial tumors (Fig. 3A and Supplementary Fig. 3A, B), in line with aforementioned observations in humans relating ECRT to *MYC* ATRT<sup>5</sup>. This corroborated the hypothesis that, in mice as in humans, *Myc* ATRT and extra-cranial tumors share certain molecular similarities, potentially revealing similar non-neuronal lineages of origin. Furthermore, a careful description of tumor locations clearly demonstrated that *Myc* intracranial tumors were in fact of extra-CNS origin, arising from the periphery of the brain, in meningeal spaces (Figs. 3A, B, C c, d; Supplementary Fig. 3A and 3B). These results fit well with the extra-axial origin and peripheral location of some human *MYC* ATRTs, as deduced from diagnostic magnetic resonance imaging (MRI) (Supplementary Fig. 1B), strengthening the hypothesis of a non-neuronal origin for this subgroup.



In contrast, all murine Shh ATRTs developed within the brain parenchyma (Fig. 3A); more precisely, 15/22 tumors were localized in the ventral sub-ventricular region, and invaded the basal ganglia (Fig. 3B, C a, b, Supplementary Fig. 3C: a, b), a location highly consistent with the description of human BG/IV SHH ATRT. In addition, we occasionally found tumor-like Smarcb1-deficient areas at the

junction of the midbrain and the posterior fossa (Fig. 3C b, Supplementary Fig. 3C: c, d), which never reached the bulky volume observed in the basal ganglia region. This indicated that, although the CAL SHH ATRT was not properly recapitulated in this model, the corresponding region in mice allowed some abnormal proliferation of Smarcb1-deficient neurons. Of note, Shh/intra-brain parenchyma

**Fig. 1 | Radiological description of ATRTs' epicenter suggests clearly distinct origins for each molecular subtype.** **A** MRI showing the most frequent tumor locations according to molecular subgroups. The round size indicates the number of tumors. The last column represents the tumors located on the midline, which are not additional cases, except the spinal MYC tumors. The different colors correspond to the molecular subgroups based on DNA methylation data (DKFZ brain tumor classifier v11b4): MYC (green), TYR (red) and SHH (blue). **B** Bar plot showing the distribution of the different molecular groups assigned according to the DNA

methylation profile at the supra- and infratentorial level. **C** Pie charts showing the distribution of the different ATRT anatomical locations at infra and supratentorial levels. NA: not available anatomical location. The color code is referred to at the bottom of Fig. 1. MCP/ICV: middle cerebellar peduncle and inferior cerebellar vermis. **D** UMAP analysis performed on human ATRT DNA methylation array data. **E** Unsupervised hierarchical clustering of ATRT samples based on DNA methylation data. Top annotation indicates ATRT anatomical location and molecular subgroup. Source data are provided as a Source Data file.

(intra-CNS) tumors were exclusively obtained after the earliest inactivation of *Smarch1*, i.e. E6 to E7; conversely, Myc/extra-brain parenchyma (extra-CNS) tumors were obtained within a broader developmental window (E6 to E10) (Fig. 3A, B). This suggested earlier and more time-restricted progenitors for Shh ATRTs than for Myc ATRTs. In addition, the single sample Gene set Enrichment Analysis (GSEA) definitively pointed to a neuronal origin for Shh, whereas the Myc tumors were mainly characterized by an immune signature (Fig. 3D), as previously described<sup>13</sup>. These results supported the hypothesis that mice Shh ATRT and supratentorial human SHH ATRTs (BG/IV SHH ATRT) develop from similar anatomical structures.

### BG/IV SHH ATRT and murine Shh ATRT show a unique expression pattern suggesting a ganglionic eminence origin

Since we were confident that BG/IV SHH ATRT formed a distinct group from CAL SHH ATRT and shared similar location with Shh mice tumors, we addressed the specific gene expression signatures in both species. In humans, we compared the four anatomically and molecularly defined subgroups. When compared with all other groups, BG/IV SHH ATRT were characterized by the overexpression of genes involved in forebrain development (*FOXG1*, *EMX2*, *ARX*, and *NRG1*), neurogenesis, synapse and neuronal plasticity (*ARC*, *LRRTM3* and *BDNF*), glial markers (*FABP7*, *OLIG2*, *GFAP*, *MLC1*) and pluripotency genes (*DPPA4*) (Fig. 4A, Supplementary Fig. 4G, Supplementary Data 4). When compared with MYC ATRT only, BG/IV SHH ATRT again showed a neuronal signature (Fig. 4A, B, D; Supplementary Fig. 4A, B, D; Supplementary Data 5). In contrast, when compared with CAL SHH ATRT only (Supplementary Fig. 4B, C), BG/IV SHH ATRTs were characterized by genes related to immune response (Fig. 4C; Supplementary Fig. 4F; Supplementary Data 6). Using deconvolution and immune infiltration estimation methods on both transcriptomic and DNA methylation datasets, we found that BG/IV SHH ATRT shared with MYC ATRT a more prominent immune infiltrate than TYR and CAL SHH ATRT, which likely accounted for their «MYC-like» expression signature (Supplementary Fig. 4H–O). However, the immune response observed in MYC ATRT and BG/IV SHH ATRT showed some differences such as a higher regulatory T cell infiltrate in MYC ATRT and a higher level of NK cells in BG/IV SHH ATRT (Supplementary Fig. 4N, O). These results were confirmed by immunostaining using CD3, CD8, and CD163, showing a more prominent immune infiltrate in BG/IV SHH ATRT than in CAL SHH ATRT (Supplementary Fig. 4P, Q).

Besides genes related to immune response, the top-expressed genes in BG/IV SHH ATRT contained the typical core set of genes expressed in the ganglionic eminence (Fig. 4D). The ganglionic eminence is a transitional structure with a key role during forebrain development; it is a progenitor domain that gives rise to cortical interneurons and, eventually, basal ganglia<sup>14,15</sup>. The expression of the ganglionic eminence core set of genes is consistent with the anatomic location of BG/IV SHH ATRT and suggested that cells from the ganglionic eminence could be at the developmental origin of this subgroup.

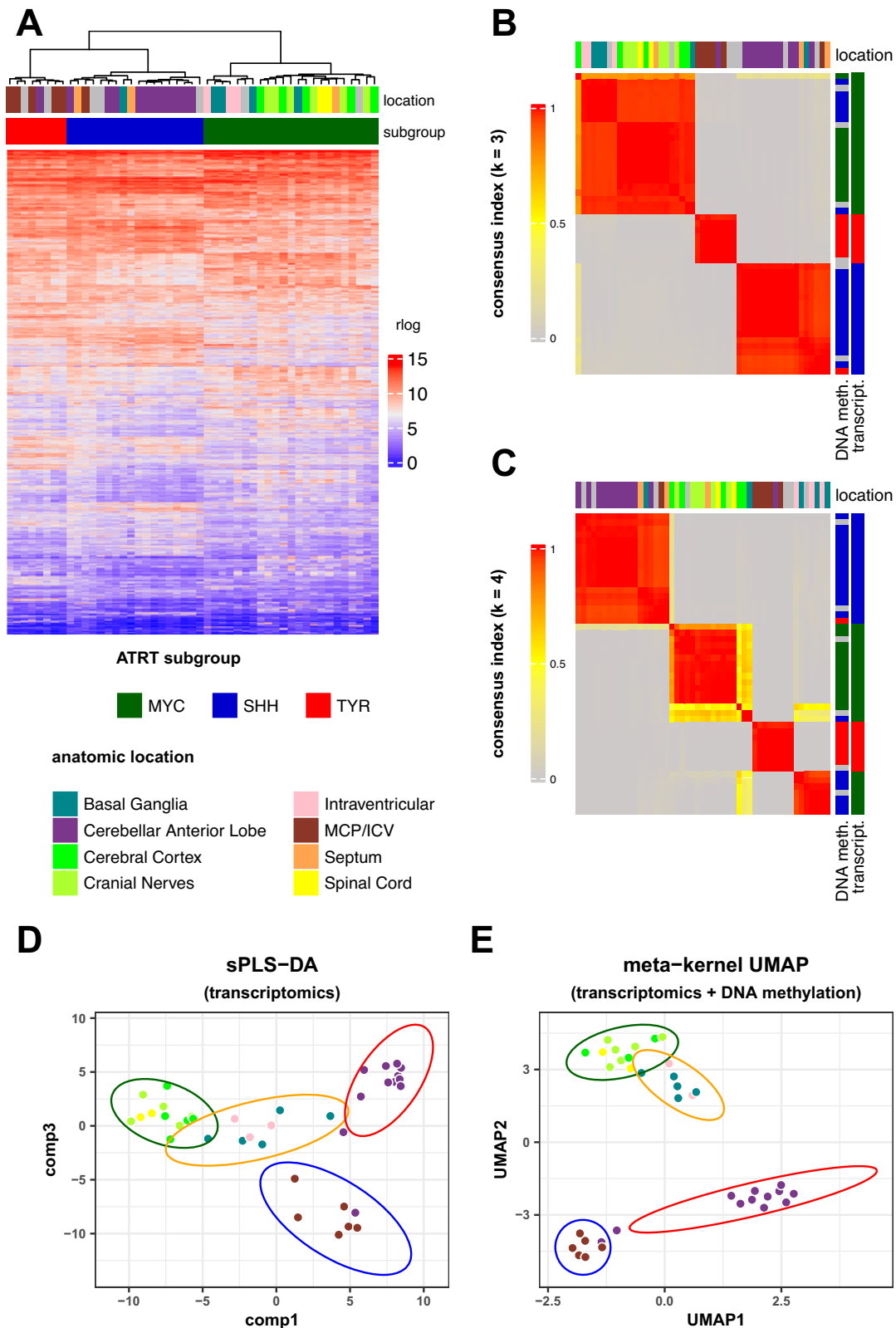
To investigate how mouse Shh ATRT actually recapitulates human BG/IV SHH ATRT, we first performed an unsupervised hierarchical clustering pooling mouse and human dataset. This showed that Shh mouse tumors clustered with CAL SHH ATRT rather than with BG/IV

SHH ATRT, despite their clearly distinct locations (Supplementary Fig. 5A). However, considering that the immune infiltrate noticeably influenced the transcriptome-based clustering of BG/IV with MYC ATRT, despite their clearly neuronal features, we hypothesized that the difference between mouse Shh ATRT and BG/IV ATRT could be related to the influence of micro-environment rather than pointing to different origins. We therefore considered that the cell identity could be also assessed by the analysis of specific transcription factors (TF), known to play a critical role in neuronal lineages' differentiation. In that line, ganglionic eminence marker genes that characterized BG/IV ATRT were also clearly overexpressed in the mouse Shh tumors (Fig. 4E), while TF found in CAL were much less clearly overexpressed (Supplementary Fig. 5B). Altogether, this shared TF expression in tumors localized in similar brain regions (Fig. 3C a, b) still suggested relevant homologies between Shh mouse and BG/IV ATRT. Together with the single sample GSEA pointing to embryonic brain neuron signatures (Fig. 3D), these results suggested neural progenitors from the ganglionic eminence as putative candidate cells of origin for basal ganglia Shh ATRT.

### Single-cell RNAseq suggests ganglionic eminence neural progenitors as putative cells of origin of BG/IV-ATRT

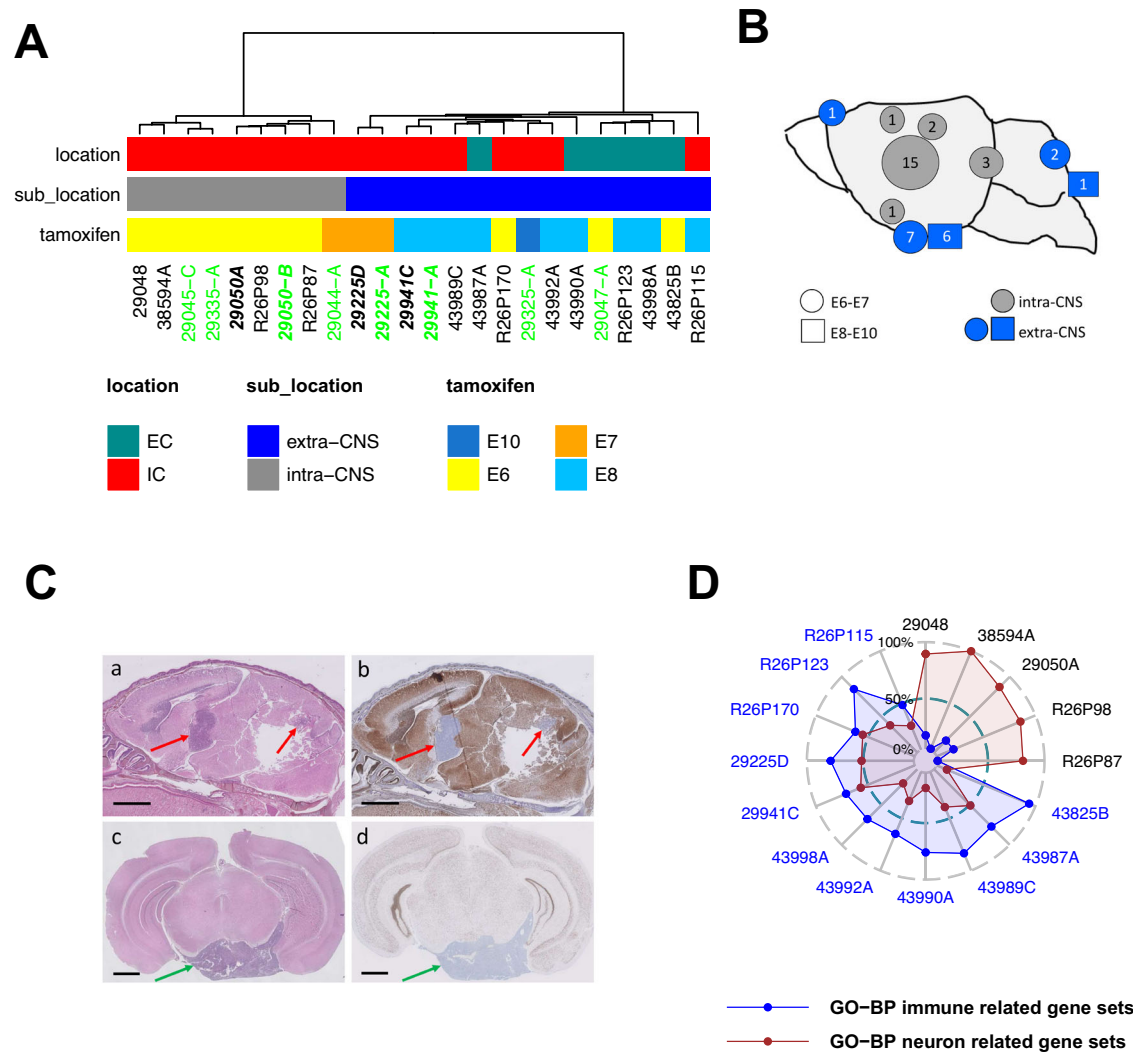
Taking advantage of the similarity between anatomical location and transcription factor modules of BG/IV and murine Shh ATRT, we assumed that single-cell transcriptomic analysis on mice tumors could further inform on the putative cells of origin in both species. We therefore performed scRNAseq on three murine Shh basal ganglia ATRT. We first checked the expression of genes that were known to be specific to SHH ATRT molecular subgroup identified in previous studies<sup>2,3</sup>; as expected, most of the genes related to the SHH ATRT signature were not homogeneously expressed, demonstrating the transcriptional intratumoral heterogeneity (Supplementary Fig. 6A). Then, our clustering approach suggested 13 different cell populations (Fig. 5A), for which marker genes were identified by differential expression analyses (Supplementary Data 7). Based on these marker genes, gene expression atlas databases (see Methods), as well as literature curation, we were able to infer the biological identity of each cell population (Fig. 5A, B; Supplementary Fig. 6B). In addition, in order to identify potential master regulator candidates for each cell population, we applied gene regulatory network (GRN) analysis using the SCENIC framework<sup>16</sup>, leading to the identification of transcription factors (TFs) characteristic for each cluster.

We first identified five clusters marked by non-neuronal marked genes: clusters 5 and 8 were characterized by the expression of genes of G1 to S and G2 to M cell cycle transitions, cluster 10 by immune response genes (*C1qa*, *Fcer1g*), cluster 11 by astro-glial gene markers (*Gfap*, *Aqp4*) and finally cluster 12 by endothelial gene markers (*Rassf9*, *Tm4sf1*) (Supplementary Fig. 6B). Interestingly, all the other clusters shared a subset of markers from the neuronal lineage (Fig. 5B, C). These clusters could be gathered in three main subgroups: (i) clusters 1 and 6 («neural progenitor 1» and «neural progenitor 2») characterized by the expression of neuronal progenitor markers (*Sox2*, *Dlx1*, *Dlx2*) and the activation of typical ganglionic eminence TFs (Cluster 1: *Gsx1*; Cluster 6: *Dlx1*); (ii) clusters 2, 7 and 9 («neuron restricted progenitors 1», «neuron restricted progenitors 2» and «committed neuronal progenitors») characterized by the expression of neuronal differentiation genes (*Dcx*,



**Fig. 2 | Integrative analysis identified four anatomical-molecular subgroups and splits SHH ATRT in two subgroups with distinct anatomical locations and transcriptional profiles.** **A** Unsupervised hierarchical clustering of human ATRT samples based on RNAseq dataset. Anatomic location as well as the assigned RNAseq subgroup for each sample is indicated in the top annotations. **B, C** Consensus clustering of human ATRT samples based on RNAseq data with  $k = 3$  (**B**) and  $k = 4$  (**C**). **D** UMAP of human ATRT based on the integrated DNA methylation

and transcriptomics datasets (kernel-based approach). Points indicate tumor samples, colors indicate anatomic location and the ellipses indicate tumor anatomical-molecular subgroups. **E** sPLS-DA individual plot using the two major components (comp 1, and comp3). Points indicate tumor samples, colors indicate anatomic location and the ellipses indicate tumor anatomical-molecular subgroups. sPLS-DA was applied on RNAseq data. Source data are provided as a Source Data file.



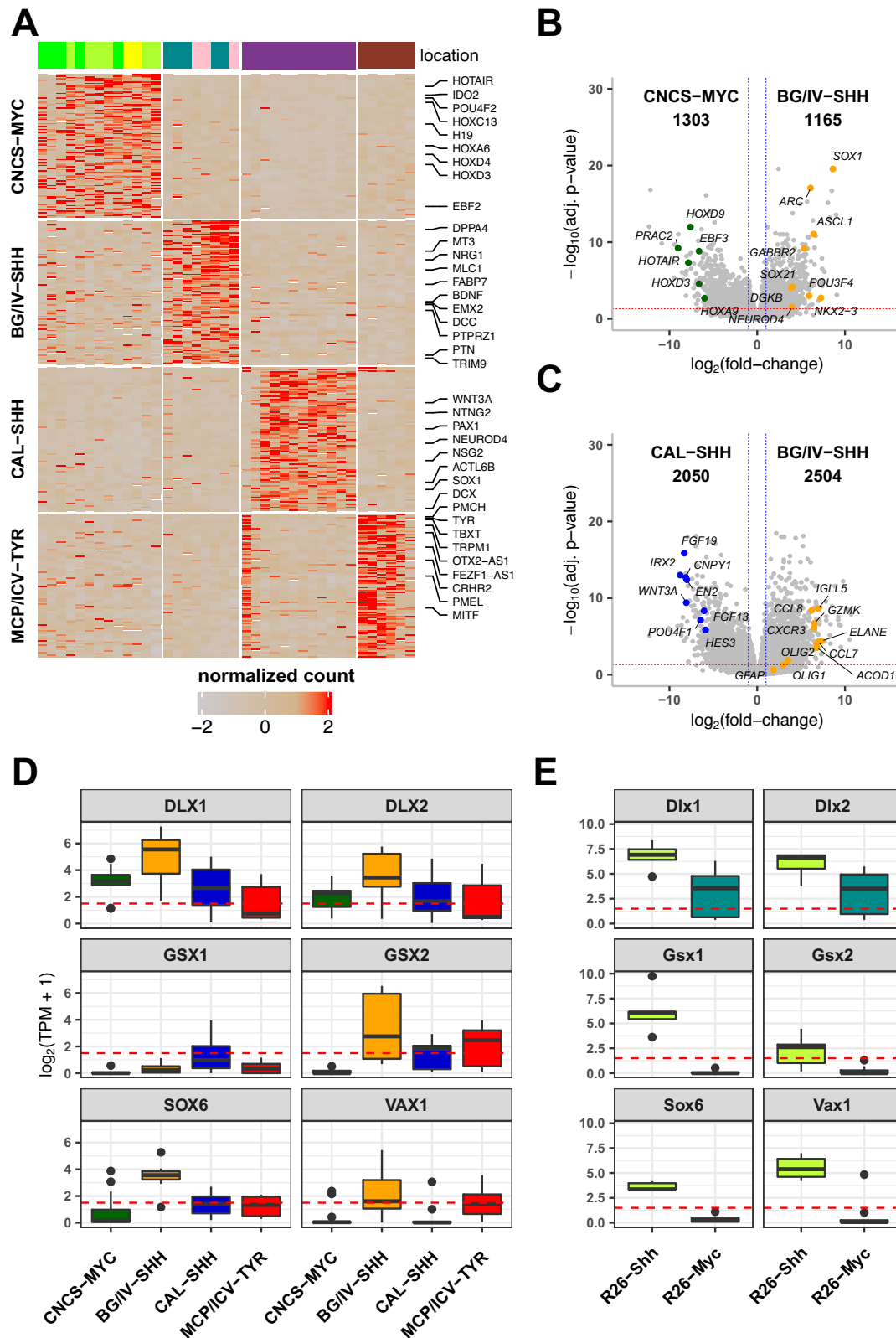
**Fig. 3 | Genetically engineered mouse models reveal specific anatomical origins for murine Myc and Shh ATRT.** **A** Unsupervised hierarchical clustering of mouse RT ( $n = 21$  biologically independent mice) based on transcriptomic dataset of combined RNAseq dataset ( $n = 16$  biologically independent samples, sample name colored in black) and gene expression microarray dataset ( $n = 8$  biologically independent samples, sample name colored in green). 3 samples (bold sample names) were processed using the two platforms and used to control the cross-platform data integration. **B** Anatomical distribution of Rosa26::CreERT2/Smadcb1flox/flox (R26) RT; circles correspond to Tamoxifen injection at E6-E7 while squares correspond to E8-E10; gray indicates intra-CNS tumors, blue, indicates extra-CNS

tumors. **C** Hematoxylin Eosin Safran (HES, a, c) and BAF47 staining (b, d) of sagittal (a, b) and coronal (c, d) sections of brains from the R26 mice. Red arrows point to intra-CNS tumors (a, b); green arrows point to extra-CNS tumors (c, d). Scale bars: 1 mm **(D)** Single sample gene set enrichment analysis of murine RT samples based on RNAseq data ( $n = 16$ ). Each dot corresponds to one sample; sample names are colored according to molecular subgroups (black: Shh; blue: Myc). Average enrichment scores (in percentage) of the top 5 differentially enriched neuronal development-related (red) and immune-related (blue) gene sets are considered and indicated for each sample. Source data are provided as a Source Data file.

*Stmn2* and *Tubb3*) and regulated by the transcription factors *Neurod1* and *Pou4f1*; (iii) clusters 3 and 4 (“neural stem cell like 1” and “neural stem cell like 2”), both characterized by neural stem cells markers (cluster 4: *Sox2*, *Sox9*, *Ttyh1*; cluster 3: *Sox2*, *Nes*) and a large variety of “non-neuronal” markers (Fig. 5B). Finally, cluster 0 fell apart from all others since it was massively characterized by non-coding RNAs, and the absence of marker genes of any specific cell type; it was therefore further referred to as “unspecified” cluster (Supplementary Data 7).

Aiming to identify the source of this cell diversity, we next performed trajectory inference analysis on the “neuronal” clusters (clusters 1, 2, 3, 4, 6, 7, and 9), using two different algorithms, the PAGA algorithm implemented in Monocle<sup>3</sup> and the elastic principal graph algorithm implemented in EIPIGraph<sup>18</sup>. In addition, RNA velocity analysis based on dynamical modeling<sup>19</sup> was used to identify the root of the trajectory. These analyses suggested that the tumor cell populations studied here originated from the “neural progenitors 1 and 2”

(Fig. 5D, E). These clusters were characterized by the expression of repressors of the terminal neuronal differentiation (*Rest* and *Id* genes) and the ganglionic eminence markers (*Ascl1*, *Dlx1*, *Dlx2*) (Fig. 5B, F). These cells then followed two antagonistic paths: (i) one towards the more committed neurons (clusters 7, 2, 9), which lost the neuronal repressors *Rest* and *Id* and switched to the post-mitotic *Actl6b* Swi/Snf member (Fig. 5D–F) and (ii) the other to the less differentiated cells, still characterized by neuronal lineage markers such as *Sox2* and *Sox9*, but also expressing more pleiotropic markers of mesodermal lineages (*Pmp22*, *Twist*) (Fig. 5B, F). Altogether, this trajectory analysis, in agreement with our hypothesis based on anatomic location and bulk RNAseq in both species, suggested that the murine Shh ATRT, and therefore possibly human BG/IV SHH ATRT, arise from neural progenitors from the ganglionic eminence, a subset of which will still be capable of neuronal commitment while another subset will dedifferentiate in more pleiotropic cells.



**CAL SHH ATRT anatomical location and molecular profile suggest a neuronal progenitor from the midbrain-hindbrain boundary as cell of origin**

Since we felt confident that CAL SHH ATRT formed a distinct group, we sought to identify its specific transcriptomic profile as we did for BG/IV SHH ATRT. Differential gene expression analysis between CAL SHH ATRT and all other subgroups identified the over-expression of genes

related to neurogenesis and neuronal migration (*SOX1*, *NTNG2*, *NEUROD4*, *NSG2*, *DCX*), WNT and FGF signaling pathways (Fig. 4A, Supplementary Data 4) and genes involved in midbrain-hindbrain boundary (MHB) patterning and cerebellum development (Fig. 6A)<sup>20-24</sup>. The MHB is an embryonic region delimiting the midbrain and the hindbrain and organizing the fate of neuronal progenitors from both sides of the edge, an embryonic structure of particular

**Fig. 4 | BG/IV SHH ATRT and murine Shh ATRT show a unique expression pattern suggesting a ganglionic eminence origin.** **A** Heatmap of gene expression using the 100 most differentially expressed genes between anatomical-molecular subgroups in a “one versus all others” manner. Top annotation indicates sample anatomical location. Genes of interest are listed at the left of the heatmap; expression levels are ranked from the lowest (gray, -2) to the highest (red). **B, C** Volcano plots showing differential gene expression analysis results of BG/IV SHH versus MYC (b) and BG/IV SHH versus CAL SHH (c). The x axis indicates the log<sub>2</sub> transformed fold-change and the y axis indicates the reverse of the log<sub>10</sub> transformed adjusted *p*-value. Horizontal red line corresponds to adjusted *p*-value equals to 0.05 and two vertical blue lines indicate log<sub>2</sub>(fold-change) respectively equal to = -1 (left) and 1 (right). Differentially expressed genes of interest are

labeled. Negative binomial GLM and Wald test were applied for gene expression comparison and generated *p*-values were corrected using the Benjamini and Hochberg method. **D, E** Boxplots of ganglionic eminence gene expression in (A) human ATRT anatomical molecular subgroups (*n* = 39 total of independent samples: *n*<sub>CNC5-MYC</sub> = 13, *n*<sub>BG/IV-SHH</sub> = 8, *n*<sub>CAL-SHH</sub> = 12, *n*<sub>MCP/ICV-TYR</sub> = 6) and in (B) mouse RT subgroups (*n* = 16 total of independent samples: *n*<sub>R26-SHH</sub> = 5, *n*<sub>R26-MYC</sub> = 11). x axis indicates subgroups and y axis indicates the level of expression in log<sub>2</sub>(TPM + 1). The box part of the boxplots represents the interquartile range while the whisker bonds of the boxplots indicate the highest and smallest values within 1.5 times interquartile range above and below the 75th and 25th quantiles respectively. Source data are provided as a Source Data file.

interest given the location of CAL SHH ATRT at the edge of supra- and infratentorial regions. This region is specified by the expression of typical genes and signaling, such as *FGF8* and *WNT3A* at the edge, markers of the rostral and caudal structures (Iroquois homeobox genes (*IRX1*, *IRX2*) and Engrailed family (*EN1*, *EN2*)), *HES3* and *PAX3*<sup>25–31</sup>. All these MHB core genes were characteristic for CAL SHH ATRTs (Fig. 6A), while GSEA pointed to embryonic neuronal development and midbrain or hindbrain patterning gene sets (Fig. 6B).

Focusing on differential expression analysis between CAL and BG/IV SHH ATRT only, we confirmed that CAL SHH ATRT were characterized by WNT and FGF signaling and MHB signature (Figs. 4C and 6C). Finally, considering that CAL SHH ATRT and infant SHH medulloblastomas develop in the cerebellum of young children and activate the SHH pathway, we assumed that a comparison of gene expression of these two types of tumors could also be relevant on the actual identity of CAL SHH ATRT (Supplementary Fig. 7; Supplementary Data 8). The MHB signature characterized again the CAL SHH ATRTs, strengthening our previous results. In addition, compared to SHH medulloblastomas, CAL SHH ATRT also showed the overexpression of genes involved in *NOTCH*, *WNT*, and *FGF* signaling pathways and the overexpression of stem cell and neuronal progenitor genes markers (Fig. 6D).

### Single-cell RNAseq analysis reveals transcriptional intra-tumoral heterogeneity of cerebellar anterior lobe ATRTs and neuronal progenitors as putative cells of origin

As aforementioned, we failed to obtain bulky tumors from the midbrain/hindbrain boundary that would recapitulate the CAL SHH ATRT subgroup in the *Rosa26-Cre<sup>ERT2</sup>::Smadcb1<sup>fllox/fllox</sup>* model. Thus, to get further insights on putative cells of origin of the CAL SHH ATRT, we performed scRNAseq on four fresh human tumor specimens. As expected, the expression of typical SHH ATRT genes identified in Johann et al.<sup>2</sup> and Torchia et al.<sup>3</sup> was found heterogeneously among cells, suggesting the transcriptional intra tumoral heterogeneity of CAL SHH ATRT (Supplementary Fig. 8A). Different clustering approaches were applied (see Methods) leading to the identification of 10 clusters (Fig. 7A). Aiming to determine the biological identity of each cluster, we performed differential expression analyses (Supplementary Data 9) as well as GRN and identified marker genes and activated TFs characteristic for each cluster, as described above for mice Shh ATRT. Using atlas databases (see Methods) and literature review, we assigned a biological identity to each of the identified clusters (Fig. 7A, B). We first isolated two cycling clusters corresponding to G1 to S and G2 to M transitions (clusters 2 and 3, Supplementary Fig. 8B). We next found two clusters (clusters 4 and 6) characterized by the expression of different neuronal markers such as *GADD45G*, *NHLHI*, *DCX*, *NRN1*, *CBLN2*, and *MAPT*<sup>6,32–37</sup>. These two clusters also showed the activation of the same neurogenic TFs such as *ISL1*, *NHLHI*, *NEUROD2*, *DLX2* and *SHOX2*, suggesting that they were closely related (GRN analysis, Fig. 7C). In addition, cluster 4 presented specific expression of *LHX9*, *GAP43* and *ELAV2* genes, related to neuronal differentiation, whereas cluster 6 showed a specific expression of *ASCL1*, *JAG1*, *DLL3*, *DLX5* and *NEUROG1*, also

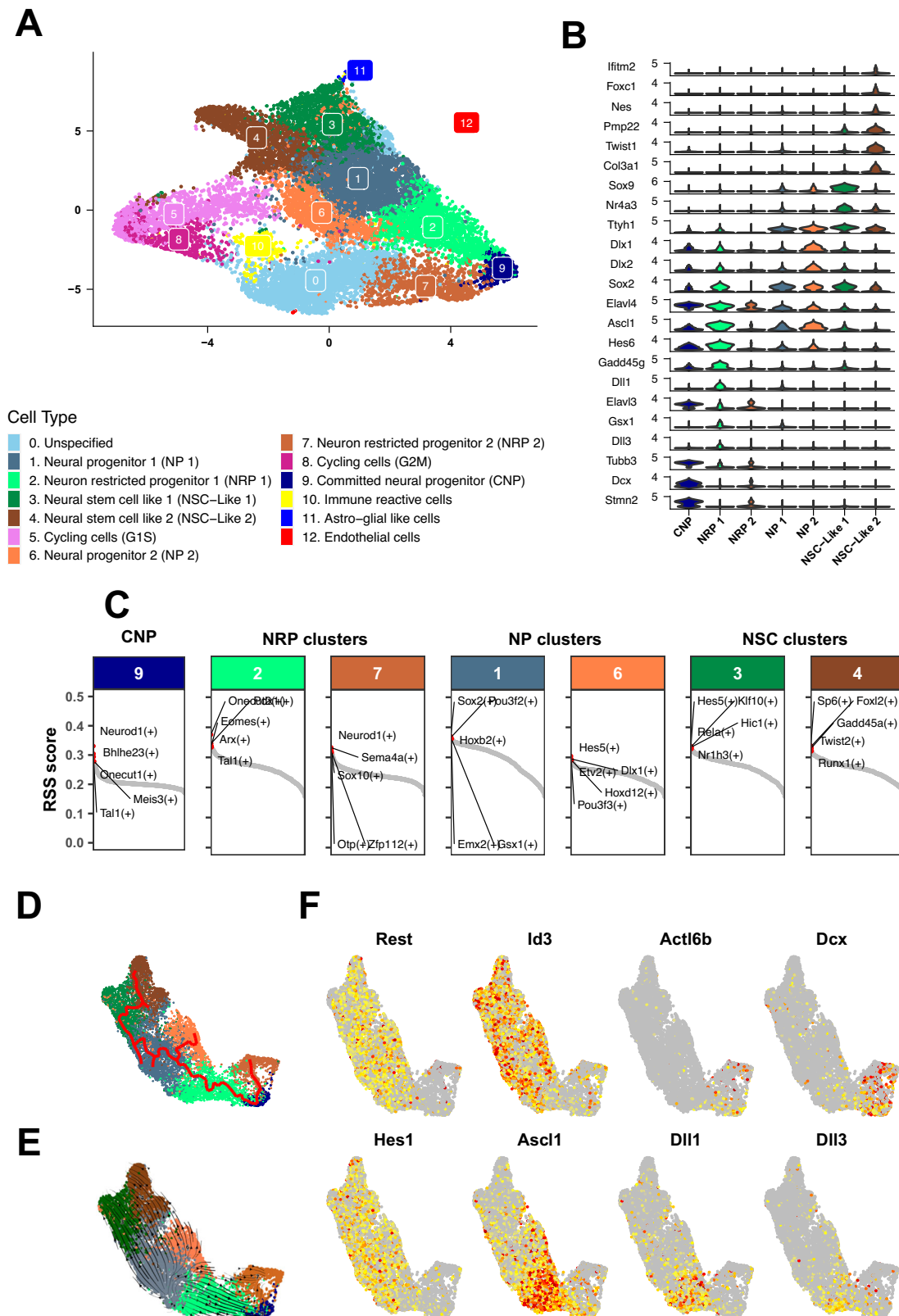
suggesting a neuronal commitment but with a prominent NOTCH pathway activation and a less differentiated state than cluster 4 (Fig. 7B). Of note, the MHB marker genes identified from bulk RNA-seq analyses (*IRX1*, *EN1*) were found expressed in these two clusters, although not restrictively (Fig. 7D). These two clusters were then referred to as “Neuronal Progenitor-like 1” (cluster 4) and “Neuronal Progenitor like 2” (cluster 6). In addition, we identified three clusters characterized by the expression of endothelial and glial genes markers (clusters 7, 8 and 9) that were referred to as “non-neuronal” clusters (Fig. 7A, Supplementary Fig. 8B, C) and one cluster characterized by the expression of genes involved in the hypoxic inflammatory response (cluster 5, Supplementary Fig. 8B). Finally, clusters 0 and 1 didn’t show any marker genes specific of known normal cell types, (Fig. 7B, Supplementary Data 9), which lead us to assign them as “undifferentiated”. However, they showed activation of TFs related to stem cell and pluripotency maintenance (*NANOG*<sup>38–40</sup>), epithelial-mesenchymal transition, and mesodermal commitment pathway (*PRRX2*, *MEOX1*, *FOXA1*, and *TBX2*)<sup>41</sup>, and neuroglial fate determination (*SOX9*)<sup>42,43</sup>, indicating some stemness features (Fig. 7C). The exclusive expression of stemness-related (*OTX2*) and neuronal committed-related markers (*DCX*) were confirmed by immunostaining (Supplementary Fig. 8D).

To investigate the relationship between the different clusters showing neuronal or stem cell features, we performed trajectory inference analyses on the “Neuronal progenitor like” and the “Undifferentiated” clusters, using similar approaches as described for murine Shh ATRT. These approaches gave consistent results and identified a path linking “Neuronal progenitor” clusters to the “undifferentiated” clusters (Fig. 7E, Supplementary Fig. 8E). Trajectory inference analysis using CytoTRACE<sup>44</sup> shows gradual change on cell differentiation state throughout the trajectory, suggesting differentiation or dedifferentiation processes involved in the tumor development (Fig. 7F). RNA velocity analysis showed that the “Neuronal progenitors” from cluster 4 was the origin of the “undifferentiated” cells (Fig. 7G), while pseudotime analyses highlighted a consistent temporal progression, starting from “Neuronal progenitors” from cluster 4 and leading to the “undifferentiated” cells (Supplementary Fig. 8F).

### Notch pathway plays a role in the differentiation blockade

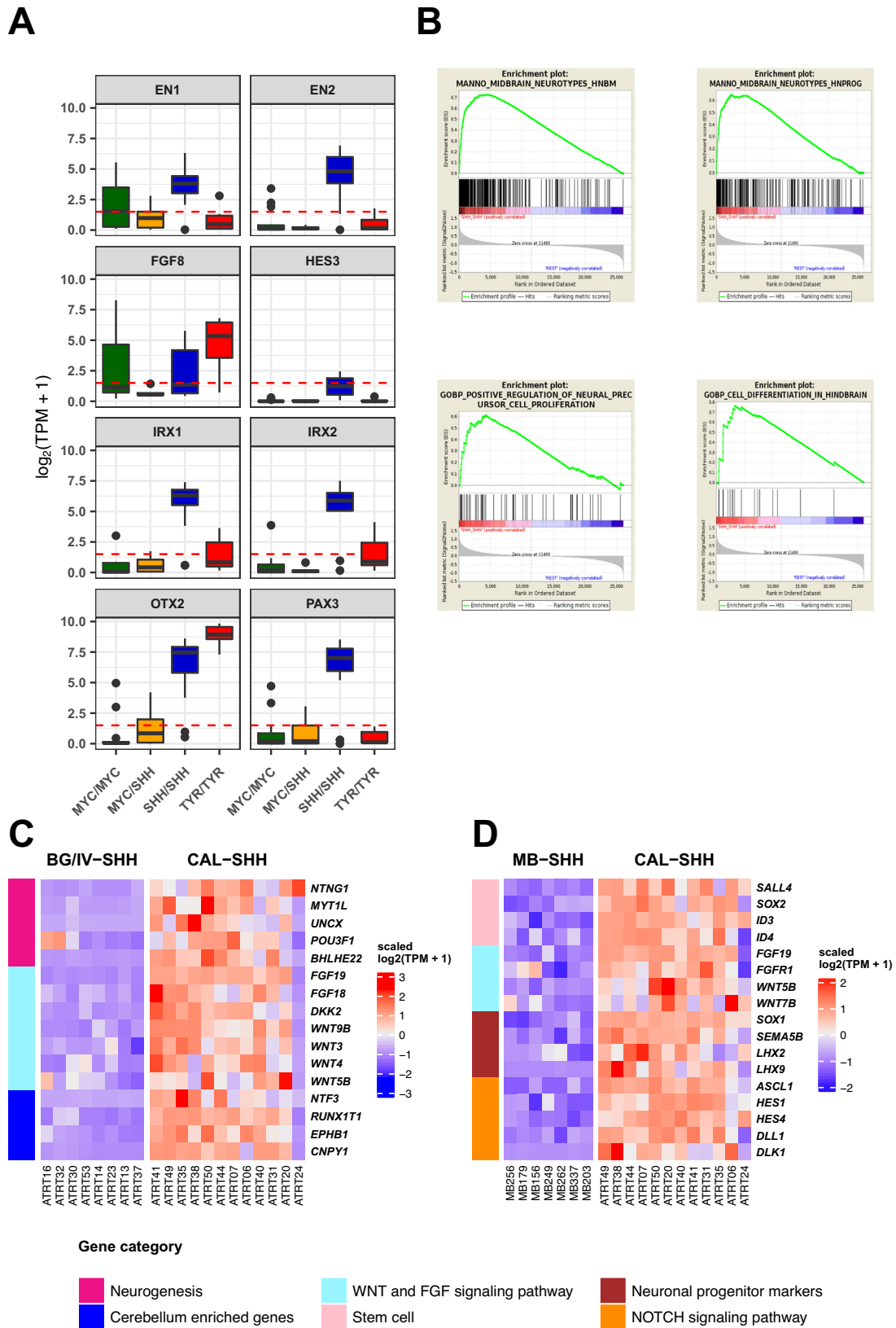
To understand the molecular mechanisms underlying this transition from “Neuronal progenitors” to the “Undifferentiated” cell clusters, we explored the expression of genes throughout the trajectory. Interestingly, we found a clear split between clusters expressing repressors of neuronal differentiation (*REST* and *ID* genes)<sup>45–47</sup>, neural stem cell markers (*SOX2*)<sup>31,48</sup> and *ACTL6A*, and clusters expressing markers of neuronal differentiation (*DCX*, *ACTL6B*) (Fig. 7H). Similarly, we observed a clear separation between clusters expressing NOTCH pathway ligands, and clusters expressing NOTCH receptors and intracellular signaling actors (Supplementary Fig. 8G), suggesting a cross-talk between clusters via the NOTCH pathway. To confirm this hypothesis, we performed ligand-receptor interaction analysis using the CellPhoneDB framework<sup>49</sup> and found that the ligands of





**Fig. 5 | Single-cell RNAseq suggests ganglionic eminence neural progenitors as putative cells of origin of BG/IV-ATRTR. A** UMAP visualization of the mouse *Shh* single cell clusters obtained after the integration of 3 independent samples. Colors distinguish the different clusters; assigned names are reported at the right of the UMAP. **B** Violin plots showing the specific marker genes for neuronal clusters. **C** Regulon specificity score (RSS) for each transcription factor (TF) in Neuronal like clusters. In each cluster, regulons (TF along with their direct targets) are ordered according to their RSS. Interesting regulons are labeled. **D** Trajectory inference

analysis using the PAGA algorithm (Monocle3). Colors distinguish the different clusters as reported in the legend of (A). **E** Embedding streams (RNA velocity - scVelo) showing the transcriptional dynamics throughout the trajectory. Colors distinguish the different clusters as reported in the legend of (A). **F** Heatmap showing the gradual expression along the trajectory of: the Notch signaling (*Hes*, *Dlk*, *Ascl1*), the SWI/SNF subunit (*Actl6b*) and neuronal differentiation gene (*Rest*, *Id*, *Dcx*) expression. Source data are provided as a Source Data file.



**Fig. 6 | CAL SHH ATRT anatomical location and molecular profile suggest a neuronal progenitor from the midbrain-hindbrain boundary as cell of origin.** **A** Boxplot showing the expression levels of genes involved in midbrain-hindbrain boundary (MHB) patterning in the human ATRT anatomical molecular subgroups ( $n = 39$  total of independent samples:  $n_{\text{NCNS-MYC}} = 13$ ,  $n_{\text{BG/IV-SHH}} = 8$ ,  $n_{\text{CAL-SHH}} = 12$ ,  $n_{\text{MCP/ICV-TYR}} = 6$ ). *x* axis indicates subgroups and *y* axis indicates the expression level in  $\log_2(\text{TPM} + 1)$ . The box part of the boxplots represents the interquartile range while the whisker bonds of the boxplots indicate the highest and smallest values

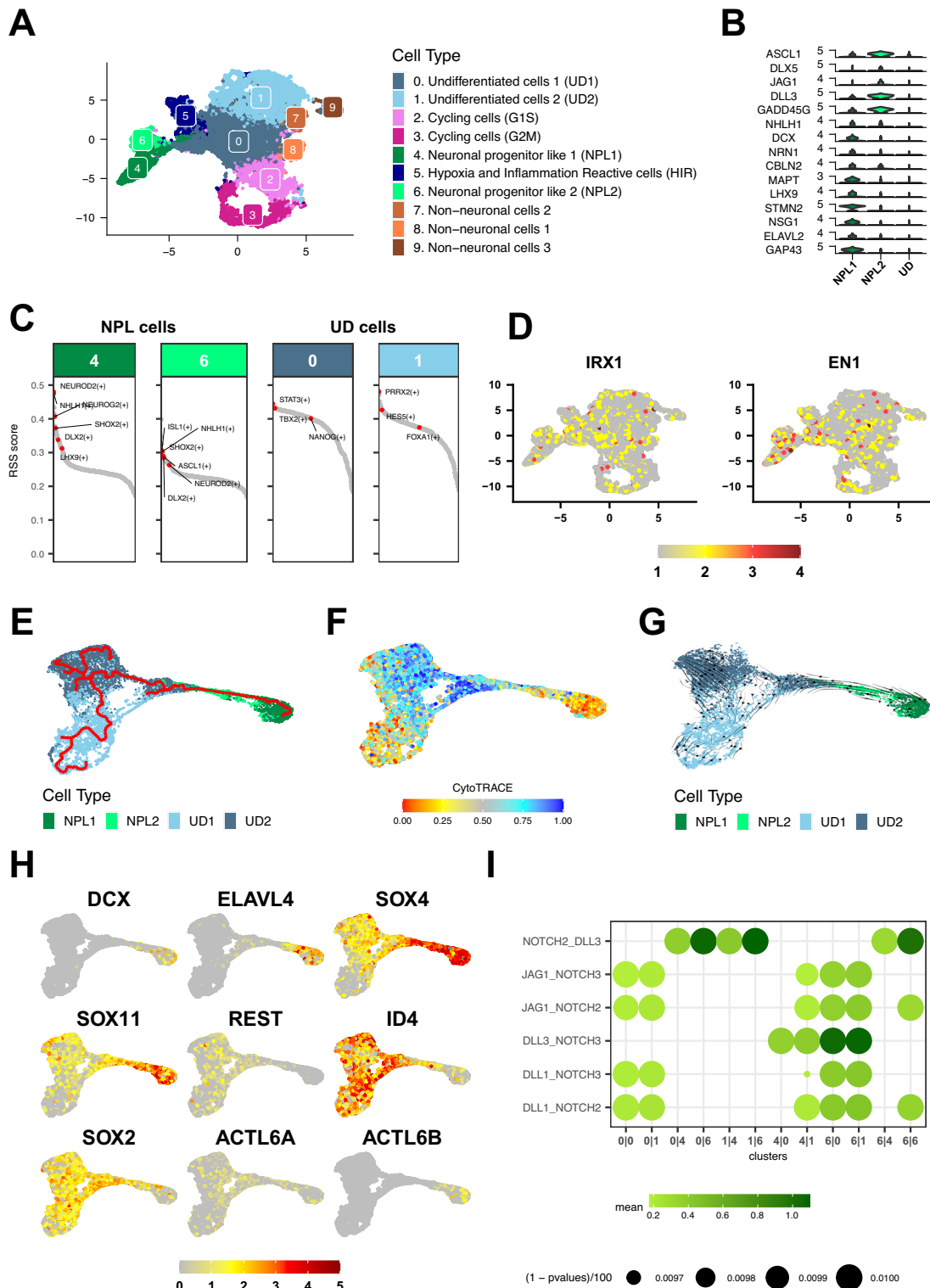
within 1.5 times interquartile range above and below the 75th and 25th quantiles respectively. **B** Enrichment plot of midbrain/hindbrain gene sets in CAL-SHH resulting from GSEA between CAL SHH versus all other subgroups. **C, D** Heatmap showing the level of expression of a selection of genes differentially expressed in CAL SHH ( $n = 12$ ) compared to BG/IV SHH ( $n = 8$ ) (c) and to medulloblastoma SHH subgroup ( $n = 7$ ) (d). Color codes at the right of the heatmaps refer to biological functions that are depicted below figures **C** and **D** Source data are provided as a Source Data file.

NOTCH signaling were specifically expressed in the neuronal progenitor clusters, while the NOTCH receptors were specifically expressed in undifferentiated clusters (Fig. 7I). Altogether, these results suggested that CAL SHH ATRT emerge from neuronal progenitors at the MHB, that, following *SMARCB1* inactivation, mostly dedifferentiate by promoting neuronal program repressors.

Interestingly, re-analysing the only single cell sequencing of CAL SHH ATRT sample available in the literature (Jessa et al.<sup>6</sup>, sample ATRT5), we also clearly found that genes related to undifferentiated

states (*SOX2*, *FABP7*, *OTX2*, *ID4*, *TTYH1*) distinguished the most prominent clusters from a minor population expressing more committed markers (*STMN4*, *DCX*), findings that corroborate our observations on an independent dataset (Supplementary Fig. 9A). Again, genes involved in the NOTCH pathway (*HES5*, *DLL3*) distinguished these clusters.

Our results suggested that the activation of the NOTCH pathway could mimic a lateral inhibition process that is also known to play a role in repressing the normal neuronal differentiation<sup>50–53</sup>. To get an insight



**Fig. 7 | Single-cell RNAseq analysis reveals transcriptional intra-tumoral heterogeneity of cerebellar anterior lobe ATRTs and neuronal progenitors as putative cells of origin.** **A** UMAP visualization of the human CAL SHH cell clusters obtained after integration 4 independent samples. Colors distinguish the different clusters; assigned names are reported at the right of the UMAP. **B** Violin plots showing the specific marker genes for Neuronal progenitor like cells and undifferentiated cells. **C** Heatmap of Midbrain/Hindbrain boundary gene signatures (*IRX1*, *EN1*) expression on the UMAP of integrated human CAL-SHH ATRT samples. **D** Regulon specificity score (RSS) for each transcription factors (TF) in Neuronal progenitor and undifferentiated cell populations. In each cluster, regulons (TF along with their direct targets) are ordered according to their RSS. Regulons of interest are labeled. **E** Trajectory inference analysis using the PAGA algorithm (Monocle3) showing a path from NPL1 cluster to UD clusters via NPL2 cells. Color code indicates cell clusters as referred in (A). Dark green: NPL1 cells, light green: NPL2 cells, light blue: undifferentiated cells. **F** Heatmap of CytoTRACE score at

single cell level. The color gradient indicates a differentiated state (red) to an undifferentiated state (blue). **G** Embedding streams (RNA velocity – *scVelo*) showing the transcriptional dynamics throughout the trajectory. Genes specifically expressed in NPL1, NPL2, UD1, and UD2 cell clusters were used. Dark green: NPL1 cells, light green: NPL2 cells, light blue: undifferentiated cells. **H** Heatmap of gene expression showing the gradual expression along the trajectory of neurogenic TFs (*SOX4*, *SOX11*), neuronal differentiation genes (*DCX*, *ELAVL4*), neuronal repressor (*REST*), stem cell and pluripotency markers (*ID4*, *SOX2*) and the SWI/SNF subunits genes (*ACTL6A*, *ACTL6B*). The color gradient indicates the expression levels, from the lowest (gray) to the highest (red). **I** Dot plot of NOTCH signaling ligand-receptor interaction between two cell clusters. Color gradient indicates the average expression of the ligand-receptor partner in the two clusters while the diameter of the dot indicates the corresponding *p*-value. Source data are provided as a Source Data file.

on how NOTCH pathway could participate to cell reprogramming, we treated two SHH ATRT cell lines (CHLA-02, and IC-032) with DAPT, a gamma secretase inhibitor (Fig. 8A). Although the treatment did not significantly affect cell viability (Supplementary Fig. 9B), it clearly impacted gene expression (Fig. 8B, C, Supplementary Data 10–11). As assessed by gene ontology analyses, differentially expressed genes were involved in developmental and differentiation programs (Fig. 8D–G). NOTCH signaling actor genes (*HES5*, *HES1*, *HES4*...) were clearly repressed as well as stem markers (*TTYHI*, *GFAP*, *FABP7*, *ID4*) typical of SHH ATRT (Fig. 8H, I); in CHLA-02 at least, the treatment also induced neuronal progenitor markers (*ISL1*, *NEUROG2*, *NEUROD1*), findings consistent with an overall dysregulation of development-related gene ontologies.

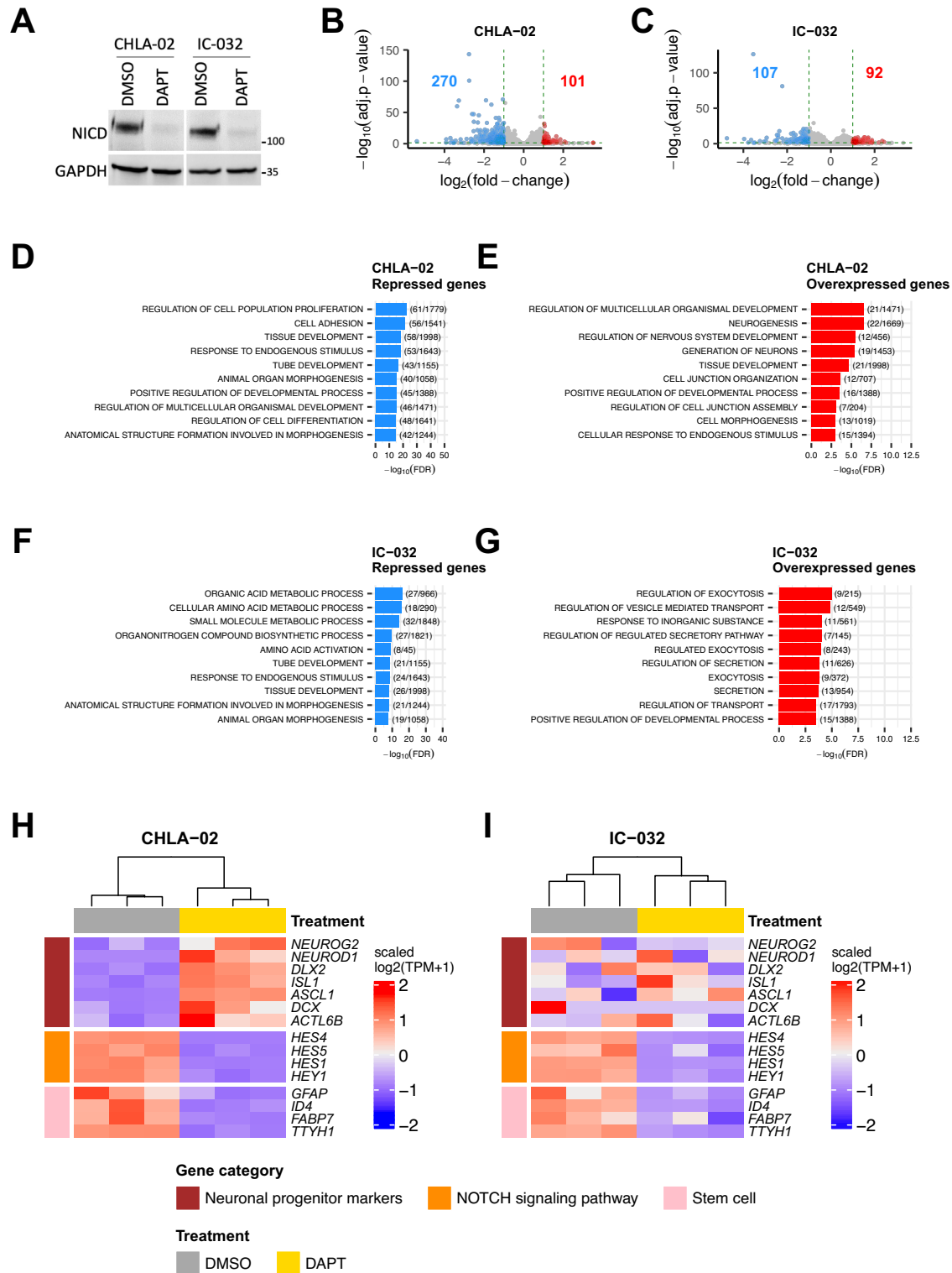
## Discussion

To date, multiple studies in humans have attempted to correlate ATRT molecular subgroups with predominant supra or infratentorial anatomical location, failing to accurately correlate the anatomical site with the molecular profiles and to provide information about the putative lineage of origin of each ATRT subgroup<sup>2–4,54,55</sup>. Here, we have investigated by an integrative analysis whether the precise anatomical description could improve the understanding of ATRT development. This approach allowed us to identify four distinct anatomical-molecular subgroups, which in turn allowed us to investigate the putative lineage of origin. We first found that a significant subset of MYC ATRT emerge from extra-axial locations such as cranial nerves; this feature has been repeatedly described in previous case reports<sup>56–62</sup> that we believe, from our own findings, could be related to the lineage of origin of MYC tumors. The location of Myc murine ATRT, arising outside of the brain in the *Rosa26-Cre<sup>ERT2</sup>::Smarcb1<sup>fllox/fllox</sup>* model, corroborated the hypothesis of an extra-axial origin for this subtype of tumors. These findings are consistent with (i) the *SMARCB1*-related schwannomatosis, which tightly links *SMARCB1* to oncogenesis in mature peripheral nerves<sup>63–66</sup> (ii) the previously published phylogenetic links between Schwann cells and ECRT, again indicating a possible neural-crest origin for rhabdoid tumors<sup>8</sup>, (iii) the unique clustering to which MYC ATRT and ECRT belong, already suggesting a common non brain origin for these types of rhabdoid tumors<sup>5</sup>, and (iv) the typically extra-axial origin of ATRTs in the *PO-Cre::Smarcb1<sup>fllox/fllox</sup>* model, originating from the cranial nerves and the periphery of the brain, locations very similar to the ones we described for human MYC ATRT, and also pointing to neural-crest cell precursors as lineage of origin. In a recently published work, Graf et al. performed scRNAseq on a series of Myc murine ATRT, which was not done in the present manuscript<sup>10</sup>. Of note, those authors related this tumor subset to primordial germ cells based on correlation with atlases, a finding that does not fully fit with the aforementioned studies. Whether these results reveal some specificity for murine Myc tumors, some diversity

among the potential origins of Myc tumors, or simply indicate that *SMARCB1* abrogation can, in some contexts, erase all specific pre-existing gene expression in such a deep manner that the transcriptome matches with primordial germ cells, remains to be further demonstrated. Further work is therefore needed on human and mouse MYC ATRT to fully elucidate this enigma.

Our study also established that SHH ATRT, canonically defined using DNA methylation profiling, are in fact composed by at least two anatomical-molecular subgroups. While we were preparing this manuscript, Federico et al. published similar findings on an independent series of SHH ATRT<sup>67</sup>, depicting three subgroups, with clearly distinct locations and gene expression profiles. “ATRT SHH-2” in the Federico study overlaps with our CAL SHH ATRT regarding the location and the overexpression of genes such as *EN2*, supporting our results about the origin of these tumors in the midbrain hindbrain boundary. Moreover, “ATRT SHH-1A” largely corresponds to our BG/IV SHH ATRT, and, interestingly, overexpress *OLIG2*, a glial marker that we also found overexpressed in the bulk RNAseq analyses, together with several neuronal markers, again underscoring the multiple lineage markers characteristic for this group. Our series comprises less SHH supratentorial cases with more restricted age distribution (between 0.7 and 6.2); this fact may explain why we don’t identify the third subgroup recently described by Federico et al., which could be part of our BG/IV SHH ATRT. By including a comprehensive radiological review, our study helped orienting the interpretation of the bulk expression profiling and extends the results of Federico et al. by adding scRNAseq for two of these subgroups. This combined approach taking into account the precise anatomical location to interpret gene expression profiling allowed us to raise hypotheses for both CAL and BG/IV SHH ATRT, i.e., progenitors from the MHB and the ganglionic eminence, respectively. These hypotheses now need to be definitely confirmed using mouse models expressing Cre in specifically restricted progenitors.

One unexpected finding of our study was the immune signature distinguishing BG/IV from CAL SHH ATRT, that accounted for the discrepancy between methylation-based and gene expression-based classifications for this particular subgroup. Previous studies depicting the immune infiltrate showed partly discrepant results regarding SHH ATRT immune infiltrate<sup>13,68</sup>, which may be explained by the use of different subgrouping techniques. It also suggests that the abundance of the immune infiltration may be due to some anatomical constraints as much as to intrinsic tumor cell properties. Interestingly, the low immune infiltrate seen in CAL ATRT is to be compared to the similarly low immune infiltrate in mouse *Shh* ATRT. Altogether, this suggests that not only the SHH or MYC methylation profiling, linked to the neuronal versus non-neuronal features of ATRT, should be considered to predict potential efficacy of immunotherapies, but rather immunohistochemistry and/or precise subtype definition.



Since these neural progenitor regions play a crucial role in brain segmentation processes<sup>69-72</sup>, SHH ATRT could be considered as a disease of embryonic brain segmentation. Moreover, single cell transcriptomic approaches suggested that *SMARCB1* loss in the appropriate neural progenitors drives a dedifferentiation process through the induction of neuronal program repressors such as *REST* and *ID* genes, leading to the emergence of neural stem cell markers including *ACTL6A* and *SOX2*. In none of these SHH ATRT did the SHH signaling seem to play a major role in cluster distinction. In

contrast, we clearly noted a prominent role of NOTCH signaling which has already been described to be characteristic for SHH ATRT<sup>2,67</sup>. Our results suggest that lateral inhibition involving the NOTCH pathway may play an important role in the inter-cluster cross-talks, and possibly in the dedifferentiation process, as suggested by the effects of NOTCH inhibition on SHH ATRT cell lines. Targeting NOTCH signaling would therefore deserve further explorations in an interceptive strategy for the treatment of both types of SHH ATRT.

**Fig. 8 | Inhibition of NOTCH signaling in rhabdoid cell lines.** **A** Western blot analyses for Notch intracellular domain NICD in CHLA-02 and IC-032 cell lines treated with DAPT NOTCH inhibitor or DMSO as a control. The experiment has been repeated 2 times for CHLA-02. **B, C** Volcano plot showing the differential gene expression analysis results of DAPT treatment versus DMSO in CHLA-02 (**B**) and in IC-032 (**C**) cell lines. The X axis indicates the  $\log_2$  transformed fold-change and the Y axis indicates the reverse of the  $\log_{10}$  transformed adjusted *p*-value. The number of significantly repressed and overexpressed genes are labeled in blue and in red, respectively. Genes having an absolute fold-change higher than 2 and an adjusted *p*-value lower than 0.05 are colored either in blue (for genes repressed in DAPT-treated cells) or in red (for genes overexpressed in DAPT-treated cells). The dotted green horizontal and vertical lines correspond to a *p*-value = 0.05 and to absolute

fold-change = 2, respectively. Negative binomial GLM and Wald test were applied for gene expression comparison and generated *p*-values were corrected using the Benjamini and Hochberg method. **D, E** Top ten significantly enriched Gene Ontology (Biological Processes only) gene sets in genes that are differentially repressed (**D**) or overexpressed (**E**) in DAPT-treated CHLA-02 cells compared with DMSO. **F, G** Top ten significantly enriched Gene Ontology (Biological Processes only) gene sets in genes that are differentially repressed (**F**) or overexpressed (**G**) in DAPT-treated IC-032 cells compared with DMSO. **H, I** Heatmap showing the level of expression of a selection of interested genes in DAPT-treated cells compared with DMSO in CHLA-02 (**H**) and in IC-032 (**I**) cell lines. Color codes at the left and at the top of the heatmaps are shown below the figure. Source data are provided as a Source Data file.

## Methods

### Animals

Mouse strain Rosa26-Cre<sup>ERT2</sup>::Smadcb1<sup>fllox/fllox</sup> (Han et al.<sup>7</sup>) was genotyped according to this reference. All experiments were performed on mixed background (129/SV×C57BL/6). The sex ratio within groups was in equilibrium. Protocol and animal housing were in accordance with national regulation and international guidelines<sup>73</sup>.

Approval for this study was received from the Institutional CEST review board (Comité d'Evaluation et de Suivi de Recherche Translacionnelle) from Curie Institute, and from the Direction Generale de la Recherche et de l'Innovation, Ministère de l'Enseignement Supérieur et de la Recherche (authorization number 6,150).

Tamoxifen administration to recombine Smadcb1 locus was performed according to the same reference. Mice were sacrificed according to the ethical approval and when neurological behaviors reached the ethical limit endpoints.

### Mouse tumor monitoring

Mice were monitored for tumor formation at least 3 times per week. The observation period encompassed at least 18 months. Intra-CNS tumors could not be measured. Mice weight was checked daily from any neurological symptom (weakness, circling, tremor, paralysis) and mice were sacrificed when weight loss reached 20%. For the rare extracranial tumor, mice were sacrificed when tumors reached 1000 mm<sup>3</sup>. Tumors and all organs were taken and fixed in AFA for histology or were frozen in -80 °C until RNA extraction. Three fresh tumors were processed for scRNA-seq.

### Mouse sample histological examination

Organs were collected, frozen on dry ice and processed for cryo-sectioning and macro-dissection or fixed in AFA (Carlo Erba, ref: 526263001) for histological examination. BAF47 immunohistochemistry was performed on fixed paraffin-embedded tissue using BD, code 612111, clone 25/BAF47, dilution 1/50<sup>74</sup>. Immunohistochemistry Ki67: primary antibody (sc-7846) incubation 1 h (dilution 1:300), secondary antibody anti-goat BIOT (705-006-147 Jackson) incubation 25 min (dilution 1:250), staining with Vectastin Kit.

### Mouse tumors macrodissection and RNA extraction

Frozen brains were serially sectioned using a cryostat at 4 mm; quick Hematoxylin stainings were performed on each section until a tumor could be identified; macrodissection was then performed with a sterile scalpel. Small pieces of tissue containing the tumor cells were frozen at -80 °C until RNA preparation. The tumor RNAs were extracted using a miRNeasy mini kit (Qiagen ref: 217004).

### ATRT Tumor samples

Freshly resected and snap-frozen human ATRT samples were collected following written informed consent of parents regarding tumor banking and use for research; approval of these consents was obtained by the internal review board from Curie Institute and Necker Hospital

for Sick Children (Paris, France, IRB approved protocol number DC-2009-955).

### Immunohistochemical analyses

A representative section was selected for each case. Unstained 3- $\mu$ m-thick slides of formalin-fixed, paraffin-embedded tissues were obtained and submitted for immunostaining. The primary antibodies employed included programmed death-ligand 1 (PD-L1) (1:100, clone EIL3N, Cell Signaling Technology, Beverly, USA), PD-1 (1:250, clone EPR4877(2), Epitomics, Cambridge, USA), CD3 (1:50, clone F7.2.38, Dako, Carpinteria, USA), CD4 (1:80, clone 4B12, Leica Biosystems, Wetzlar, Germany), CD8 (1:25, clone C8/144B, Thermo Scientific, Waltham, USA), CD45 (1:500, clone PD7/26 and 2B11, Dako, Carpinteria, USA), CD57 (1:40, clone HNK-1, BD Biosciences, Franklin Lakes, USA), CD68 (1:400, clone KP1, Dako, Carpinteria, USA), CD163 (1:50, clone IHC163, Diagnostics, Blagnac, France), FOXP3 (1:50, clone 206D, BioLegend, USA), Granzyme B (ready to use, clone 11F1, Leica Biosystems, Wetzlar, Germany), OX40 (1:100, clone ACT-35, Thermo Scientific, Waltham, USA), OTX2 (1:600, clone IH12C4B5, Thermo Fisher, Rockford, USA) and DCX (1:200, clone EPRI9997, Abcam, Cambridge, United Kingdom). All slides were stained using previously optimized conditions including positive and negative controls (human placenta for PD-L1 and human tonsil for other markers). PD-L1 expression was evaluated in the tumor cells using H-score, which includes the percentage of positive cells showing membrane staining pattern (0–100) and intensity of the staining (0–3+), with a total score ranging from 0 to 300. All other immunemarkers were evaluated as density of cells (0: absent; 1: scarce; 2: moderate, and 3: diffuse), defined as the number of positive cells per area (1 mm<sup>2</sup>) regardless of the intensity. The final score for each marker was expressed as the average score of the five areas analyzed within the tumor region. The final scores for each marker from each patient were then transferred to a database for statistical analysis.

### Cell culture and viability assays

ATRT cell line CHLA-02-ATRT (#CRL-3020, ATCC) was cultured according to the manufacturer's protocol. IC-032 cell line, established in Curie Institute from a supra-tentorial ATRT, was cultured in DMEM/F12 supplemented with 20 ng/mL FGF, 20 ng/mL EGF, and 1x B27 supplement (#15360284, Fisher Scientific). To assess the effect of NOTCH inhibition on cell viability,  $8 \times 10^3$  cells per well were plated in triplicate in 96-well plates and treated either with DMSO vehicle control or varying doses of DAPT gamma-secretase inhibitor (#HY-13027, MedChem Express). After 7 days of treatment, resazurin (Sigma-Aldrich) was added (20  $\mu$ g/ml) to assess cell viability and cells were incubated for another 2–6 h, depending on the cell line. Fluorescence signals proportional to the number of cells were recorded in a FLUOstar Omega plate reader (BMG labtech SARL).

### Western blot analysis

For NOTCH inhibition validation, CHLA-02-ATRT and IC-032 cell lines were treated with DMSO or DAPT at 10  $\mu$ M for 7 days. Cells were then

washed once with cold PBS and scraped on ice with lysis buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.5% NP40, 120 mM NaCl) supplemented with protease inhibitors (Roche). After centrifugation and quantification, protein extracts were resolved by sodium dodecyl sulfate–polyacrylamide gel electrophoresis before transfer onto nitrocellulose membrane. Immunoblots were done with monoclonal rabbit cleaved Notch1 (#4147, Cell signaling) and HRP-conjugated GAPDH (#HRP-60004, proteintech) antibodies. Cleaved Notch1-blotted membrane was then incubated with an anti-rabbit immunoglobulin G horseradish peroxidase-coupled secondary antibody (1:3000, NA934; Amersham Biosciences). Proteins were detected by enhanced chemiluminescence (Biorad).

### Gene expression profiling

**RNA-seq library preparation.** Total RNAs were obtained from ATRT ( $n=10$ ) and mouse RT ( $n=16$ ) frozen samples using Qiagen QIAamp RNAeasy kit, according to the manufacturer's procedures (Qiagen, Venlo, Netherlands). For cell lines, CHLA-02-ATRT and IC-032 cells were treated with DMSO or DAPT at  $10\ \mu\text{M}$  for 7 days. RNAs were extracted using the Nucleospin II kit (Macherey-Nagel). Experiments were performed in triplicate. The tumor cell content was visually estimated before RNA extractions. Barcoded Illumina compatible libraries were generated from 750 ng of total RNA for each sample using TruSeq Stranded mRNA Library Preparation Kit (Illumina, San Diego, California, U.S.). Libraries were sequenced using the Illumina HiSeq 2000/2500 or Illumina NovaSeq platforms in the 100 bp paired-end mode. FASTQ samples were generated after demultiplexing the resulting BCL files.

**RNA-seq data processing.** Raw data were processed using an in-house pipeline developed at the Institut Curie Bioinformatics Core Facility, following standard analysis in the field and available at <https://github.com/bioinfo-pf-curie/RNA-seq>. Briefly, read mapping and counting were performed using STAR version 2.5.3a aligner<sup>75</sup>. The human reference genome hg19 and the mouse reference genome mm10 were used. The 10 ATRT RNA-seq data were combined with 9 from Andrianteranagna et al.<sup>76</sup> and 29 from Leruste et al.<sup>13</sup> to form a cohort of 49 RNA-seq samples that were re-analyzed from FASTQ files using the same pipeline.

**RNA-seq statistical analysis.** For each of the human and mouse datasets, all genes having 0 counts in all samples were filtered out before subsequent analyses. Variance stabilization process was then applied using the *rlog()* function of DESeq2 version 1.30.1 (Love et al.<sup>77</sup>) Bioconductor package. Genes having rlog counts lower than 7.5 in 99% or more of the samples were filtered out. Only the 5000 top variable genes (based on IQR) were kept for the unsupervised analyses.

Principal component analyses (PCA) were performed using the *prcomp()* function of the R base package *stats* on centered and scaled data. Hierarchical clustering analyses were conducted using the *Heatmap()* function of the *ComplexHeatmap* version 2.6.2 (Gu et al.<sup>78</sup>) Bioconductor package. Pearson correlation and Ward's method were used respectively as similarity metric and linkage method. Consensus clustering analyses were conducted to estimate the stability of the number of clusters. They were performed with the *ConsensusClusterPlus* version 1.54.0 (Wilkerson and Hayes<sup>79</sup>) (Bioconductor package using the same metrics of similarity and linkage as set for the hierarchical clustering. All other settings were set by default except the *pFeature* that we set to 0.8. The *ComplexHeatmap* package was used to visualize the consensus clustering result.

The Uniform Manifold Approximation and Projection (UMAP)<sup>80</sup> non-linear dimensionality reduction algorithm was applied for visualization purpose. UMAP analyses were performed using the *umap* (version 0.2.7.0) CRAN packages.

Sparse Partial Least Squares Discriminant Analysis (sPLS-DA)<sup>81</sup> was conducted using the mixOmics framework version 6.14.1 (Rohart et al.<sup>82</sup>) Bioconductor package. The optimal number of component as

well as the number of gene per component were determined by running the *perf()* and *tune.splsda()* functions using 3-fold cross-validation repeated 50 times. Finally, sPLS-DA analysis was run using the *splsda()* function using 3 components with respectively the 90, 100 and 50 previously selected genes (Supplementary Fig. 2B).

Differential gene expression analyses were performed using *DESeq()* function of the DESeq2 package using the filtered raw counts. Resulting *p*-values were corrected using the Benjamini and Hochberg method (*a.k.a.* FDR). Immune cells and stromal cells infiltration scores were computed using the ESTIMATE version 1.0.11<sup>83</sup> R-Forge package, a marker-based single sample gene set enrichment method. Immune cells (T cell CD8+, T cell CD4+, T cell regulatory, NK cell) relative fraction were computed using deconvolution-based quanTIseq algorithm (Finotello et al., 2019; Sturm et al., 2019) implemented in the immunedeconv version 2.0.4R Bioconductor package. Analyses were performed inside R environment (version 4.0.2).

Functional enrichment analyses on human dataset were performed using the web application available in <https://www.gseamsigdb.org/gsea/msigdb/annotate.jsp> and the GSEA tool (version 2.2.3) on the GO:BP gene set collection (version 7.4). Functional enrichment analyses on mouse dataset were performed using the GSVA (version 1.38.2) bioconductor package and the msigdb (version 7.4.1) CAN package.

**Gene expression array data processing and analysis.** Eight Smarcb1flox/flox;Rosa26-CreERT2 mouse primary RT affymetrix (MOE430 2.0 array type) samples from Han et al.<sup>7</sup> were re-analyzed. Data were normalized using the RMA method implemented in the affy version 1.70.0 Bioconductor package with the custom Brain array CDF annotation packages version 23.0.0 (<http://brainarray.mbnj.med.umich.edu/Brainarray/Database/CustomCDF/23.0.0/entrezg.asp>). Gene filtering was performed using the PCA based approach described in Lu et al. (2011) and implemented in the pvac version 1.40.0 Bioconductor package. Hierarchical clustering was performed using the ComplexHeatmap version 2.9.4 Bioconductor package. GSEA are applied using the GSEA software version 2.2.3 and the MSigDB database version 5.2 downloaded from <https://www.gsea-msigdb.org/gsea/msigdb/>. Data analyses were performed using R environment (version 4.1.1)

**Mouse RNA-seq and gene expression array data integration.** 8 mouse RT affymetrix data and 16 mouse RT RNA-seq data were combined based at gene level. The two datasets were previously normalized, filtered and scaled separately before merging. Platform technical effect was assessed using PCA and corrected using linear model implemented in *comBat()* function of SVA version 3.40.0 Bioconductor R package. Hierarchical clustering was applied to the combined dataset using the ComplexHeatmap version 2.9.4 package and the 5000 most variable genes (based on IQR). Euclidean distance and Ward linkage method were used.

**Human and Mouse gene expression data integration.** Human and mouse RNA-seq data were merged based on the orthologous genes (16864) between the two species identified using biomaRt package version 2.48.3. Genes having an expression lower than a threshold equals to 6 regularized counts in more than  $\frac{3}{4}$  of the samples were filtered out. Sample-wise correlation between human and mouse datasets using genes of interest was performed based on the Pearson's method. Organism effect correction was applied using the *ComBat()* function of the SVA package version 3.40.0 before unsupervised clustering on the merged dataset. The 1000 most variable genes (based on IQR value) were kept for the hierarchical clustering performed on the merged dataset. The analyses were conducted within R environment (version 4.1.1).

## DNA methylation array

**DNA methylation array data processing.** Infinium methylationEPIC array data from 54 human ATRT samples were collected (30 were retrieved from et al.<sup>13</sup>) and processed using *RnBeads* package version 1.6.1 (Assenov et al., 2014) Bioconductor package. All samples and all probes were kept after quality control. Probe intensities were normalized using the *rnb.execute.normalization()* function with the “illumina” method. Probes outside CpG context (2991 probes), targeting single nucleotide polymorphisms (17369 probes) or targeting X and Y chromosomes (19457 probes) were all filtered out before subsequent analysis.

**DNA methylation array statistical analysis.** Beta-value at probe level were extracted using *meth()* function. For a given CpG site, beta-value is the ratio of signal from methylated probes relative to signal from both methylated and unmethylated probes. The 5,000 highest variable probes (based on beta-value IQR) among the retained probes (828,109) were selected for the unsupervised analyses. Hierarchical clustering samples was performed with the *Heatmap()* function of *ComplexHeatmap* package. Euclidean and 1-Pearson were used as distance metrics respectively for probes and samples clustering. Ward’s method was used as linkage criterion for both sample and probes clustering. The UMAP non-linear dimensionality reduction algorithm was applied for visualization purpose using the *umap* R package. Leucocytes infiltration scores were computed based on the leucocytes unmethylated probes (LUMP) identified by Aran et al.<sup>84</sup>. For a given sample, this score is calculated as 1 subtracted by the mean LUMP beta-values divided by 0.85. Lymphocytes score was computed, for each sample, as the mean of the lymphoid-specific hypermethylated probes identified by Killian et al.<sup>85</sup>. Data processing and analyses were performed inside R environment (version 4.0.2).

## Human gene expression and DNA methylation data integration

Gene expression (RNA-seq) and DNA methylation (EPIC array) data (“early”) integration were performed using the kernel-based method implemented in *mixKernel* version 0.7 CRAN package<sup>12</sup>. The analysis was conducted on the 43 samples in which both RNA-seq and DNA methylation data are available. For RNA-seq dataset, the rlog count matrix including only the 2,000 most variable genes (based on IQR value) was used while for DNA methylation dataset, the beta-value matrix including only the 5000 most variable genes (based on IQR value) was considered. The kernel matrix of each dataset was computed using the *compute.kernel()* function with the “linear” kernel method. The two kernel matrices were combined using the *combine.kernels()* function with the “full-UMKL” option. PCA as well as UMAP were applied on the meta-kernel using respectively the *prcomp()* function of the R base package *stats* and the *umap* version 0.2.7.0 CRAN package. Alluvial diagrams were generated using the *alluvial* version 0.1-2 CRAN package. Data analysis was performed inside R environment (version 4.0.2).

## Single cell RNA sequencing (scRNA-seq)

**Tissue processing and cell population enrichment for human samples.** Fresh tumor samples were cut in small pieces then dissociated 30 min at 37 °C in CO<sub>2</sub>-independent medium (Gibco) + 0,4 g/l of human albumin (Vialebex) with Liberase TL (Roche) 150 mg/ml and DNase 1 (Sigma) 150 mcg/ml. Dissociated cells were then filtered with a 40 mm cell strainer, then washed and resuspended with CO<sub>2</sub>-independent medium + 0,4 g/l of human albumin. Cells were then continuously maintained on ice or at 4 °C. In case of lot of blood cells, the Debris removal kit (Miltenyi Biotec) was used according to the manufacturer’s protocol. To enrich in tumoral cells (human samples) the Tumor Cell Isolation Kit (Miltenyi Biotec) was used according to the manufacturer’s protocol. Cells were then resuspended in PBS + BSA 0.04%. Samples were prepared for concentration of 800 cell/ml.

Tissues were processed within 2 h after tumor resection and loaded in 10x Chromium instrument within 4 h.

**Preparation of single cell suspensions for mouse samples.** Fresh tumor samples were cut in small pieces then dissociated 30 min at 37 °C in CO<sub>2</sub>-independent medium (Gibco) + 0,4 g/l of human albumin (Vialebex) with Liberase TL (Roche) 150 mg/ml and DNase 1 (Sigma) 150 mcg/ml. Dissociated cells were then filtered with a 100 mm cell strainer, then washed and resuspended with CO<sub>2</sub>-independent medium + 0,4 g/l of human albumin. Cells were then continuously maintained on ice or at 4 °C. In case of lot of blood cells. Cells were then resuspended in PBS + BSA 0.04%. Samples were prepared for concentration of 800 cell/ml. Tissues were processed within 2 h after tumor resection and loaded in 10x Chromium instrument within 4 h.

**Single cell RNA sequencing.** Sample preparations were loaded on a 10x Chromium instrument (10x Genomics) and libraries were prepared using a Single Cell 3’ Reagent Kit (V2 chemistry, 10X Genomics) according to the manufacturer’s protocol, targeting 1000 recovered cells per sample. Single cells were included and barcoded into droplets together with gel beads coated with unique barcodes, unique molecular identifiers (UMI), and poly(dT) sequences, followed by in droplet reverse transcription to generate barcoded full-length cDNA. cDNA was subsequently recovered from droplets, then cleaned up with Dynabead MyOne Silane Beads (Thermo Fisher Scientific), then amplified with the following protocol: 98 °C-3 min; 12x (98 °C-15s, 67 °C-20s, 72 °C-1 min); held at 4 °C. Amplified cDNA product was cleaned up using the SPRI select Reagent Kit (Beckman Coulter). Indexed libraries were constructed following these steps: 1. Fragmentation, end repair and A-tailing; 2. Size selection with SPRI select beads; 3. Adaptor ligation; 4. Post-ligation cleanup with SPRI select beads; 5. Sample index PCR and final cleanup with SPRI selects beads. Library and quality assessment were achieved using dsDNA High Sensitivity Assay Kit and Bioanalyzer Agilent System. Indexed libraries were tested for quality, equimolarly pooled and sequenced on an Illumina HiSeq2500 using paired-end 26 × 98 bp as sequencing mode, targeting at least 50,000 reads per cell.

**scRNA-seq data pre-processing.** Raw data generated by the sequencing Illumina HiSeq2500 platform were pre-processed using the cellranger (version 3.1.0, <https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger>) including demultiplexing, mapping with the hg19 or the mm10 reference genome, gene counting and aggregation steps.

**scRNA-seq cell filtering.** Single sample adaptive filtering strategy was applied to remove “bad cells” in both human and mouse samples. The number of detected genes and the percentage of mitochondrial RNA were considered as the filtering criteria. Cells with both a low number of genes and a high proportion of mitochondrial RNA were discarded. The threshold of the minimum number of detected genes was set as the 5th percentile of the distribution of the number of detected genes in all cells (Supplementary Figs. 10A and 12A). The thresholds of the maximum proportion of mitochondrial genes were set individually for each sample based on the visual inspection of the plot of the number of detected genes versus the percentage of mitochondrial gene (Supplementary Figs. 10B and 12B). In addition, cell clustering was performed in individual sample to check if some covariates such as the number UMIs per gene ration (UMIs/gene) drive the clustering of cells (Supplementary Fig. 10D). For IRT003 samples, two clusters (3 and 4, Supplementary Fig. 10D) are characterized by a relatively low UMIs/gene and, therefore, cells belonging to these two clusters of IRT003 were not included in the integrated dataset.

Since mouse samples were not FACS-sorted, additional filtering aiming to remove non-tumoral cells was performed based on the expression of *Ptprc*, *Epcam*, *Smarcb1* genes (Supplementary Fig. 13).



The three samples were first integrated using the CCA-based Seurat version 3.2.2 methods before UMAP embedding.

**scRNA-seq data integration.** scRNA-seq data integration was performed using the CCA-based approach described in Stuart et al.<sup>86</sup> and implemented in Seurat version 3. Data integration quality was assessed by plotting the cell cycle phases on the integrated data (Supplementary Figs. 11D and 11E).

**scRNA-seq cell clustering.** PCA was applied to reduce the dimensionality of the integrated data using the *RunPCA()* function. The integrated data matrix was previously scaled using the *ScaleData()* function before PCA. The clustering was conducted using the graph-based modularity optimization Louvain algorithm implemented in Seurat v3. KNN graph is first built using the *FindNeighbors()* function on a user defined number of PCs. Then, the clustering was performed using the *FindClusters()* function with a specific resolution. Since the clustering result depends on the chosen number of PCs and the resolution value and we do not have an a priori about the expected number of clusters in our dataset, we choose to explore our data by running the clustering algorithm using different combination of number of PCs (from 12 to 50 increment by 1) and resolutions (from 0 to 1.5 increment by 0.1). The IKAP approach published in Chen et al.<sup>87</sup> was used to perform this recursive clustering and to assess their stability. In addition, iterative clustering was also applied by removing cells belonging to well-defined clusters and re-running the clustering with the remaining cells. This approach leads to the identification of 10 clusters which can be obtained using a 12 PCs with a clustering resolution equals to 0.2 and recapitulating all the possible clusters identified throughout the analyses performed above. UMAP was used to visualize the clustering result in reduced two dimensions.

**scRNA-seq cluster marker genes and cell type annotation.** To biologically annotate the identified clusters, differential expression analyses in one versus others manner using the *FindAllMarkers()* function were carried out to identify marker genes for each cluster. The default Wilcoxon Rank Sum test was applied. Genes with a  $\log_2(\text{fold-change})$  higher than 0.5, an adjusted  $p$ -value lower than 0.01 and detected in more than 25% of the cells of the given cluster were defined as gene markers for this cluster. These markers were investigated by knowledge-based using literature curations to identify the biological closest cell type of the cluster. In addition, on-line data bases such as <http://mousebrain.org/>, <http://dropviz.org/>, <https://portal.brain-map.org/> and <https://www.proteinatlas.org/> were used to explore marker genes cell type.

**scRNA-seq gene regulatory network analysis.** Gene regulatory network (GRN) analysis was conducted using the SCENIC framework<sup>16</sup>. For more details about the analysis steps, please refer to Sande et al.<sup>88</sup>.

The *pyscenic* program version 0.10.3 was used to conduct the analysis. For the analysis of human samples, the TFs list (containing 1,839 human TFs) downloaded from <https://github.com/aertslab/pySCENIC/blob/master/resources/> on August 7th, 2020, the cis-Target database (<https://resources.aertslab.org/cistarget/>) and the human motif annotation table downloaded from <https://resources.aertslab.org/cistarget/motif2tf/> on August 12th 2020 were for the analysis. For the analysis of mouse samples, the TFs list (containing 1721 mouse TF) downloaded from <https://github.com/aertslab/pySCENIC/blob/master/resources/> on January 20th 2022, the cis-Target database and the mouse motif annotation table downloaded from <https://resources.aertslab.org/cistarget/motif2tf/> on January 20th 2022 were used. The regulon specificity score (RSS) was computed for each cluster and for each regulon in order to identify the most specific regulon compared to other clusters. Python version 3.6.11 was used for the analysis.

**scRNA-seq trajectory inference (TI) and pseudotime analysis.** TI analyses were conducted using reverse graph embedding approaches implemented in (1) EIPiGraph version 1.0.0 (Albergante et al., 2020) R package and based on the elastic principal graph<sup>89</sup> and (2) the Monocle3 version 0.2.3.0 (Cao et al.<sup>17</sup>) R package and based on the Partition-based graph abstraction (PAGA)<sup>90</sup>. For the EIPiGraph tool, both the *computeElasticPrincipalTree()* and the *computeElasticPrincipalCurve()* function were run using 30 nodes. Lambda and Mu parameters were set to 0.01 and 0.1 respectively. All other parameters are set to their default values. For the Monocle3 tool, the dimensionality of the data was reduced using UMAP (based on 12 PCs) before cell clustering and graph inference. TI analyses were conducted within R version 4.0.2 environment.

**scRNA-seq RNA velocity analysis.** RNA velocity analysis was performed using the scVelo version 0.2.3 (Bergen et al.<sup>19</sup>) tool using the dynamical modeling. Spliced and unspliced matrices were generated separately for each sample in loom file using the *velocito* tool version 0.17.17<sup>91</sup>. RNA velocity was estimated using the marker genes for considered clusters and visualized on the embedding UMAP coordinates generated by Monocle3 (see TI analysis part). RNA velocity analysis was performed inside python version 3.7.8 environment.

**scRNA-seq ligand-receptor interaction analysis.** Ligand-receptor interaction analysis was performed using the cellPhoneDB framework<sup>49</sup>. To conduct this ligand-receptor interaction analysis, the *cellphonedb* binary (version 2.1.7) was run on the integrated expression matrix including all genes. Only ligand-receptor interaction with mean higher than 0.075 and a  $p$ -value lower than 0.05 were considered as significant interaction. The analysis was conducted inside Python 3.7.10 environment.

**scRNA-seq differentiation state prediction.** Cell differentiation state was estimated using the CytoTRACE<sup>44</sup> tool (version 0.3.3, Gulati et al.<sup>44</sup>) with default options. CytoTRACE assesses the relative differentiation state of each cell using a predictive model based on the gene counts (number of expressed genes) and the average expression of genes that are highly correlated with gene counts (see Gulati et al.<sup>44</sup> for more details). For each cell, it generates a score between 0 and 1 indicating a relatively more differentiated and less differentiated state. The analyses were conducted inside R 4.0.2 environment.

### Single nucleus data analysis

Single nucleus data of an ATRT sample (ATRT5) localized in the pineal area published in Jessa et al.<sup>6</sup> was re-analyzed. Gene read counts per cell (generated by cellranger) were provided by the author. Very basic data filtering was applied. Nuclei having number of detected genes lower than the 5th percentile were filtered out. Sequencing depth normalisation was performed using the *logNormalize* function of Seurat (Version 3.2.2). Dimensionality reduction using PCA was applied to the dataset using the 2000 highly variable genes (Seurat flavor). Gene marker expressions were assessed using the heatmap density produced by the *nebulosa*<sup>92</sup> R package (Version 1.0.1) on a UMAP embedding visualization based on the 12 first PCs.

### Reporting summary

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

### Data availability

Raw and processed data included in this study have been deposited in NCBI's Gene Expression Omnibus and accessible through GEO Super-Series accession number [GSE242090](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE242090). Cell lines bulk RNA-seq data are available through GEO accession number [GSE241733](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE241733). Bulk RNA-seq data of mouse rhabdoid tumor model are accessible through GEO

accession number [GSE241734](#). Mouse gene expression array (Affymetrix®) from Han et al., 2016 are available in GEO under accession number [GSE64019](#). Single cell RNA-seq data from mouse model are deposited in GEO under accession code [GSE241736](#). Single cell RNA-seq data from human primary ATRT are deposited in GEO under accession code [GSE241737](#). Bulk RNA-seq data of human primary ATRT are accessible through [GSE241831](#) accession code. Bulk RNA-seq data from Andrianteranagna et al., 2021 are stored in GEO under accession code [GSE175891](#). Bulk RNA-seq data from Leruste et al., 2019 are deposited in dbGaP database under accession code [phs001915.v1.p1](#). DNA methylation array (Illumina Infinium MethylationEPIC) data are accessible through [GSE242089](#) code. DNA methylation array data from Andrianteranagna et al., 2021 are available in GEO under accession number [GSE175892](#). Source data are provided with the manuscript, without restriction. Source data are provided with this paper.

## Code availability

The codes are available in GitHub following the link: <https://github.com/bourdeautlab/LobonJglesias2023>.

## References

- Lee, R. S. et al. A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J. Clin. Invest.* **122**, 2983–2988 (2012).
- Johann, P. D. et al. Atypical Teratoid/Rhabdoid tumors are comprised of three epigenetic subgroups with distinct enhancer landscapes. *Cancer Cell* **29**, 379–393 (2016).
- Torchia, J. et al. Integrated (epi)-genomic analyses identify subgroup-specific therapeutic targets in CNS Rhabdoid tumors. *Cancer Cell* **30**, 891–908 (2016).
- Ho, B. et al. Molecular subgrouping of atypical teratoid/rhabdoid tumors - A reinvestigation and current consensus. *Neuro. Oncol.* **22**, 613–624 (2020).
- Chun, H.-J. E. et al. Identification and analyses of extra-cranial and cranial rhabdoid tumor molecular subgroups reveal tumors with cytotoxic T cell infiltration. *Cell Rep.* <https://doi.org/10.1016/j.celrep.2019.10.013> (2019).
- Jessa, S. et al. Stalled developmental programs at the root of pediatric brain tumors. *Nat. Genet.* **51**, 1702–1713 (2019).
- Han, Z. Y. et al. The occurrence of intracranial rhabdoid tumors in mice depends on temporal control of Smarcb1 inactivation. *Nat. Commun.* **7**, 10421 (2016).
- Custers, L. et al. Somatic mutations and single-cell transcriptomes reveal the root of malignant rhabdoid tumors. *Nat. Commun.* <https://doi.org/10.1038/s41467-021-21675-6> (2021).
- Vitte, J., Gao, F., Coppola, G., Judkins, A. R. & Giovannini, M. Timing of Smarcb1 and Nf2 inactivation determines schwannoma versus rhabdoid tumor development. *Nat. Commun.* **8**, 1–12 (2017).
- Graf, M. et al. Single-cell transcriptomics identifies potential cells of origin of MYC rhabdoid tumors. *Nat. Commun.* **13**, 1544 (2022).
- Capper, D. et al. DNA methylation-based classification of central nervous system tumors. *Nature* **555**, 469–474 (2018).
- Mariette, J. & Villa-Vialaneix, N. Unsupervised multiple kernel learning for heterogeneous data integration. *Bioinformatics* **34**, 1009–1015 (2018).
- Leruste, A. et al. Clonally expanded T cells reveal immunogenicity of Rhabdoid tumors. *Cancer Cell* **36**, 597–612 (2019).
- Donkelaar. Development of the basal ganglia and the basal forebrain. *Brain Mapp. Encycl. Ref.* **2**, 357–365 (2015).
- Encha-Razavi, F. & Sonigo, P. Features of the developing brain. *Child's Nerv. Syst.* **19**, 426–428 (2003).
- Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
- Cao, J. et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
- Albergante, L. et al. Robust and scalable learning of complex intrinsic dataset geometry via ELPIGraph. *Entropy* **22**, 296 (2020).
- Bergen, V., Lange, M., Peidl, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
- Buckles, G. R., Thorpe, C. J., Ramel, M. C. & Lekven, A. C. Combinatorial Wnt control of zebrafish midbrain-hindbrain boundary formation. *Mech. Dev.* **121**, 437–447 (2004).
- Di Giovannantonio, L. G. et al. Otx2 cell-autonomously determines dorsal mesencephalon versus cerebellum fate independently of isthmic organizing activity. *Development* **141**, 377–388 (2014).
- Gasser, E., Johannssen, H. C., Rülcke, T., Zeilhofer, H. U. & Stoffel, M. Foxa1 is essential for development and functional integrity of the subthalamic nucleus. *Sci. Rep.* **6**, 1–15 (2016).
- Hirata, H. & Tomita, K. Hes1 and Hes3 regulate maintenance of the isthmic organizer and development of the mid/hindbrain. *EMBO J.* **20**, 4454–4466 (2001).
- Lobe, C. G. Expression of the helix-loop-helix factor, Hes3, during embryo development suggests a role in early midbrain-hindbrain patterning. *Mech. Dev.* **62**, 227–237 (1997).
- Hanks, M., Wurst, W., Anson-Cartwright, L., Auerbach, A. B. & Joyner, A. L. Rescue of the En-1 mutant phenotype by replacement of En-1 with En-2. *Science* **269**, 679–682 (1995).
- Joyner, A. L., Herrup, K., Auerbach, B. A., Davis, C. A. & Rossant, J. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the En-2 homeobox. *Science* **251**, 1239–1243 (1991).
- Martinez, S., Andreu, A., Mecklenburg, N. & Echevarria, D. Cellular and molecular basis of cerebellar development. *Front. Neuroanat.* **7**, 1–12 (2013).
- Matsumoto, K. et al. The prepattern transcription factor Irx2, a target of the FGF8/MAP kinase cascade, is involved in cerebellum formation. *Nat. Neurosci.* **7**, 605–612 (2004).
- Hibi, M. & Shimizu, T. Development of the cerebellum and cerebellar neural circuits. *Dev. Neurobiol.* **72**, 282–301 (2012).
- Nakamura, H., Katahira, T., Matsunaga, E. & Sato, T. Isthmus organizer for midbrain and hindbrain development. *Brain Res. Rev.* **49**, 120–126 (2005).
- Peretz, Y. et al. A new role of hindbrain boundaries as pools of neural stem/progenitor cells regulated by Sox2. *BMC Biol.* **14**, 1–20 (2016).
- Dennis, D. J., Han, S. & Schuurmans, C. bHLH transcription factors in neural development, disease, and reprogramming. *Brain Res.* **1705**, 48–65 (2019).
- Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. & Kriegstein, A. R. Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat. Neurosci.* **24**, 584–594 (2021).
- La Manno, G. et al. Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell* **167**, 566–580.e19 (2016).
- Parisian, A. D. et al. SMARCB1 loss interacts with neuronal differentiation state to block maturation and impact cell stability. *Genes Dev.* **34**, 1316–1329 (2020).
- Peukert, D., Weber, S., Lumsden, A. & Scholpp, S. Lhx2 and Lhx9 determine neuronal differentiation and compartment in the caudal forebrain by regulating Wnt signaling. *PLoS Biol.* **9**, e1001218 (2011).
- Yuzwa, S. A. et al. Developmental emergence of adult neural stem cells as revealed by single-cell transcriptional profiling. *Cell Rep.* **21**, 3970–3986 (2017).
- Carvalho, J. Cell reversal from a differentiated to a stem-like state at cancer initiation. *Front. Oncol.* **10**, 1–6 (2020).
- Zhao, W., Ji, X., Zhang, F., Li, L. & Ma, L. Embryonic stem cell markers. *Molecules* **17**, 6196–6236 (2012).

40. Deisch, J., Raisanen, J. & Rakheja, D. Immunohistochemical expression of embryonic stem cell markers in malignant rhabdoid tumors. *Pediatr. Dev. Pathol.* **14**, 353–359 (2011).
41. Soldatov, R. et al. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364**, eaas9536 (2019).
42. Vong, K. I., Leung, C. K. Y., Behringer, R. R. & Kwan, K. M. Sox9 is critical for suppression of neurogenesis but not initiation of gliogenesis in the cerebellum. *Mol. Brain* **8**, 0–17 (2015).
43. Leto, K. et al. Consensus paper: cerebellar development. *Cerebellum* **15**, 789–828 (2016).
44. Gulati, G. S. et al. Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405–411 (2020).
45. Otto, S. J. et al. A new binding motif for the transcriptional repressor REST uncovers large gene networks devoted to neuronal functions. *J. Neurosci.* **27**, 6729–6739 (2007).
46. Ballas, N., Grunseich, C., Lu, D. D., Speh, J. C. & Mandel, G. REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis. *Cell* **121**, 645–657 (2005).
47. Nechiporuk, T. et al. The REST remodeling complex protects genomic integrity during embryonic neurogenesis. *Elife* **5**, 1–28 (2016).
48. Ellis, P. et al. SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev. Neurosci.* **26**, 148–165 (2004).
49. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
50. Roesse-Koerner, B., Stappert, L. & Brüstle, O. Notch/Hes signaling and miR-9 engage in complex feedback interactions controlling neural progenitor cell proliferation and differentiation. *Neurogenesis* **4**, e1313647 (2017).
51. Machold, R. P., Kittell, D. J. & Fishell, G. J. Antagonism between Notch and bone morphogenetic protein receptor signaling regulates neurogenesis in the cerebellar rhombic lip. *Neural Dev.* **2**, 5 (2007).
52. Ables, J. L., Breunig, J. J., Eisch, A. J. & Rakic, P. Not(ch) just development: Notch signalling in the adult brain. *Nat. Rev. Neurosci.* **12**, 269–283 (2011).
53. Boareto, M., Iber, D. & Taylor, V. Differential interactions between Notch and ID factors control neurogenesis by modulating Hes factor autoregulation. *Dev* **144**, 3465–3474 (2017).
54. Dho, Y. S. et al. Investigation of the location of atypical teratoid/rhabdoid tumor. *Child's Nerv. Syst.* **31**, 1305–1311 (2015).
55. Nowak, J. et al. Magnetic resonance imaging surrogates of molecular subgroups in atypical teratoid/rhabdoid tumor. *Neuro. Oncol.* **20**, 1672–1679 (2018).
56. Beschorner, R. et al. Atypical teratoid-rhabdoid tumor spreading along the trigeminal nerve. *Pediatr. Neurosurg.* **42**, 258–263 (2006).
57. Oh, C. C. et al. Atypical teratoid/rhabdoid tumor (ATRT) arising from the 3rd cranial nerve in infants: a clinical-radiological entity? *J. Neurooncol.* **124**, 175–183 (2015).
58. Rizzo, D. et al. SMARCB1 deficiency in tumors from the peripheral nervous system: a link between schwannomas and rhabdoid tumors? *Am. J. Surg. Pathol.* **36**, 964–972 (2012).
59. Verma, A. & Morriss, C. Atypical teratoid/rhabdoid tumor of the optic nerve. *Pediatr. Radiol.* **38**, 1117–1121 (2008).
60. Wang, X. et al. Atypical teratoid/rhabdoid tumor (AT/RT) arising from the acoustic nerve in a young adult: a case report and a review of literature. *Medicine* **94**, 1–4 (2015).
61. Wu, P.-S. Atypical teratoid rhabdoid tumor (AT/RT) at the cerebellopontine angle arising from acoustic nerve with facial nerve palsy: report of a case. *Int. J. Clin. Exp. Pathol.* **13**, 2377–2380 (2020).
62. Wykoff, C. C. et al. Atypical teratoid/rhabdoid tumor arising from the third cranial nerve. *J. Neuro-Ophthalmol.* **28**, 207–211 (2008).
63. Smith, M. J., Wallace, A. J., Bowers, N. L., Eaton, H. & Evans, D. G. R. SMARCB1 mutations in schwannomatosis and genotype correlations with rhabdoid tumors. *Cancer Genet.* **207**, 373–378 (2014).
64. Kehrer-Sawatzki, H., Farschtschi, S., Mautner, V. F. & Cooper, D. N. The molecular pathogenesis of schwannomatosis, a paradigm for the co-involvement of multiple tumor suppressor genes in tumorigenesis. *Hum. Genet.* **136**, 129–148 (2017).
65. Kresak, J. & Walsh, M. Neurofibromatosis: a review of NF1, NF2, and Schwannomatosis. *J. Pediatr. Genet.* **05**, 098–104 (2016).
66. Evans, D. G. et al. Schwannomatosis: A genetic and epidemiological study. *J. Neurol. Neurosurg. Psychiatry* <https://doi.org/10.1136/jnnp-2018-318538> (2018).
67. Federico, A. et al. ATRT-SHH comprises three molecular subgroups with characteristic clinical and histopathological features and prognostic significance. *Acta Neuropathol.* **143**, 697–711 (2022).
68. Melcher, V. et al. Macrophage-tumor cell interaction promotes ATRT progression and chemoresistance. *Acta Neuropathol.* **139**, 913–936 (2020).
69. Langenberg, T. & Brand, M. Lineage restriction maintains a stable organizer cell population at the zebrafish midbrain-hindbrain boundary. *Development* **132**, 3209–3216 (2005).
70. Belmonte-Mateos, C. & Pujades, C. From cell states to cell fates: how cell proliferation and neuronal differentiation are coordinated during embryonic development. *Front. Neurosci.* **15**, 1–15 (2022).
71. Delgado, R. N. & Lim, D. A. Maintenance of positional identity of neural progenitors in the embryonic and postnatal telencephalon. *Front. Mol. Neurosci.* **10**, 1–9 (2017).
72. Puelles, L. Evolution of the nervous system brain segmentation and forebrain development in amniotes. *Brain Res. Bull.* **55**, 695–710 (2001).
73. Workman, P. et al. Guidelines for the welfare and use of animals in cancer research. *Br. J. Cancer* **102**, 1555–1577 (2010).
74. Bourdeaut, F. et al. hSNF5/INI1-deficient tumours and rhabdoid tumours are convergent but not fully overlapping entities. *J. Pathol.* **211**, 323–330 (2007).
75. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
76. Andrianteranagna, M. et al. SMARCA4-deficient rhabdoid tumours show intermediate molecular features between SMARCB1-deficient rhabdoid tumours and small cell carcinomas of the ovary, hypercalcaemic type. *J. Pathol.* **255**, 1–15 (2021).
77. Love, M. I., Huber, W., & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome. Biol.* **15**, 550 (2014).
78. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–9 (2016).
79. Wilkerson, M. D., Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
80. McInnes, L., Healy, J. & Melville, J. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.* (2018).
81. Lê Cao, K. A., Boitard, S. & Besse, P. Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinforma.* **12**, 0–16 (2011).
82. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K. A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS. Comput. Biol.* **13**, e1005752 (2017).
83. Yoshihara, K. et al. Inferring tumor purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
84. Aran, D., Sirota, M. & Butte, A. J. Systematic pan-cancer analysis of tumor purity. *Nat. Commun.* **6**, 1–12 (2015).

85. Killian, J. K. et al. Imprints and DPPA3 are bypassed during pluripotency -and differentiation-coupled methylation reprogramming in testicular germ cell tumors. *Genome Res.* **26**, 1490–1504 (2016).
86. Stuart, T. et al. Comprehensive Integration of Single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
87. Chen, Y. C. et al. IKAP-identifying K mAjor cell population groups in single-cell RNA-sequencing analysis. *Gigascience* **8**, 1–8 (2019).
88. Van de Sande, B. et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276 (2020).
89. Gorban, A. & Zinovyev, A. Elastic principal graphs and manifolds and their practical applications. *Comput. (Vienna/N. Y.)* **75**, 359–379 (2005).
90. Wolf, F. A. et al. PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 1–9 (2019).
91. La Manno, G. et al. RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
92. Alquicira-Hernandez, J. & Powell, J. E. Nebulosa recovers single-cell gene expression signals by kernel density estimation. *Bioinformatics* **37**, 2485–2487 (2021).

## Acknowledgements

This project has been financed by the INCA PRK 19-052 (FB), SFCE Enfants-Cancer et Santé, Etoile de Martin, the SIRIC of Institut Curie, the Saint Baldrick Foundation, ANR-10-IDEX-0001-02 PSL, INCa-DGOS-Inser12554, CIC IGR-Curie 1428, and the association Marabout de Ficelle. María-Jesús Lobón Iglesias is a recipient of the Fondation pour la Recherche Medicale PhD grant. High-throughput sequencing was performed by the ICGex NGS platform of the Institut Curie supported by the grants ANR-10-EQPX-03 (Equipex) and ANR-10-INBS-09-08 (France Génomique Consortium) from the Agence Nationale de la Recherche (“Investissements d’Avenir” program), by the ITMO-Cancer Aviesan (Plan Cancer III) and by the SiRIC-Curie program (SiRIC Grant INCa-DGOS-465 and INCa-DGOSInser12554). Data management, quality control and primary analysis were performed by the Bioinformatics platform of the Institut Curie. We thank Virginie Dangles-Marie, David Andreau, and all members of animal facility, Andre Nicolas and all members of PathEx for kind help. We thank Florence Cavalli and Pierre Gestraud for helpful discussions and the Curie Bioinformatics Core Facility data managers for data management. We thank François Doz, Gudrun Schleiermacher, Sergio Roman-Roman, Dominique Gallier, Nicolas André, Franck Bielle, Ulrich Schuller, for their kind support and help for the realization of the study and fruitful criticism of the work.

## Author contributions

M.-J.L.-I. prepared the samples for multi-omic analyses and the single-cell RNAseq experiments; she interpreted the results and did bioinformatics analyses with M.A. Z.-Y.H. performed all the experiments with mouse models, with the help of M.F. All co-first authors wrote the manuscript, under the supervision of co-last authors. F.B. conceived and designed the study. V.D.-R. supervised all the imaging review. N.S.

supervised the bioinformatics analyses. K.B. set up the clinical procedures to have access to human samples. J.M.-P., G.P., L.M., D.G., S.R., and C.B. ensured the access to nucleic acids for all human samples. C.C. performed the ATRT cell line experiment with NOTCH inhibitor and corresponding bioinformatics analysis. A.T.-E. reviewed all histological slides. V.D.-R. and L.C. participated to the review of MRI. V.M. and R.B.-B. helped for all scRNAseq experiments. C.C. and S.T. helped in the realization of several experiments. C.D., T.B. and O.A. gave access to medulloblastoma samples and RNAseq results for comparison of SHH medulloblastoma with CAL SHH ATRT. D.G. and S.R. prepared nucleic acid extractions for human and mouse samples. M.B. and S.B. helped in the realization of the single-cell RNAseq experiments. O.D., J.W., O.A., E.P., D.S. and C.P. helped in writing the manuscript and having access to the biological samples.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-42371-7>.

**Correspondence** and requests for materials should be addressed to Franck Bourdeaut.

**Peer review information** *Nature Communications* thanks Ankur Sharma, and the other, anonymous, reviewers for their contribution to the peer review of this work.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

<sup>1</sup>INSERM U830, Laboratory of Translational Research In Pediatric Oncology, PSL Research University, SIREDO Oncology center, Institut Curie Research Center, Paris, France. <sup>2</sup>INSERM U900, Bioinformatics, Biostatistics, Epidemiology and Computational Systems Unit, Institut Curie, Mines Paris Tech, PSL Research University, Institut Curie Research Center, Paris, France. <sup>3</sup>Somatic Genetic Unit, Department of Pathology and Diagnostic and Theranostic Medecine, Institut Curie Hospital, Paris, France. <sup>4</sup>INSERM U932, Immunity and Cancer, PSL Research University, Institut Curie Research Center, Paris, France. <sup>5</sup>Department of Neuropathology, GHU Paris-Psychiatry and Neurosciences, Sainte-Anne Hospital, Paris, France. <sup>6</sup>Paris Psychiatry and Neurosciences Institute (IPNP), UMR S1266, INSERM, IMA-BRAIN, Paris, France. <sup>7</sup>Department of Children and Adolescents Oncology, Gustave Roussy, Paris Saclay University, Villejuif, France. <sup>8</sup>Department of Pediatric Neurosurgery-AP-HP, Necker Sick Kids Hospital, Université de Paris, Paris, France. <sup>9</sup>Imaging Department, Institut Curie Hospital, Paris, France. <sup>10</sup>CNRS UMR 3347, INSERM U1021, Institut Curie, PSL Research University, Université Paris-Saclay, Orsay, France. <sup>11</sup>INSERM U830, Diversity and Plasticity of Childhood Tumors Lab, PSL Research University, SIREDO Oncology Center, Institut Curie Research Center, Paris, France. <sup>12</sup>Balgrist University Hospital, Faculty of Medicine, University of Zurich (UZH), Zurich, Switzerland. <sup>13</sup>Institut Curie, PSL University, Single Cell Initiative, ICGex Next-Generation

Sequencing Platform, PSL University, 75005 Paris, France. <sup>14</sup>INSERM U830, Integrative Functional Genomics of Cancer Lab, PSL Research University, Institut Curie Research Center, Paris, France. <sup>15</sup>Department of Translational Research, PSL Research University, Institut Curie Research Center, Paris, France. <sup>16</sup>Pediatric Radiology Department, AP-HP, Necker Sick Kids Hospital and Paris Cite University INSERM 1299 and UMR 1163, Institut Imagine, Paris, France. <sup>17</sup>Department of Pediatric Oncology, SIREDO Oncology Center, Institut Curie Hospital, Paris, and Université de Paris, Paris, France. <sup>18</sup>These authors contributed equally: María-Jesús Lobón-Iglesias, Mamy Andrianteranagna, Zhi-Yan Han. <sup>19</sup>These authors jointly supervised this work: Nicolas Servant, Kevin Beccaria, Volodia Dangoulouff-Ros, Franck Bourdeaut. ✉ e-mail: [franck.bourdeaut@curie.fr](mailto:franck.bourdeaut@curie.fr)