



**HAL**  
open science

# Integrative Clinico-Molecular Analysis and Single-Cell RNA Sequencing To Unravel the Origin of Atypical Teratoid Rhabdoid Tumors

María-Jesús Lobón Iglesias

► **To cite this version:**

María-Jesús Lobón Iglesias. Integrative Clinico-Molecular Analysis and Single-Cell RNA Sequencing To Unravel the Origin of Atypical Teratoid Rhabdoid Tumors. Cancer. Université Paris-Saclay, 2022. English. NNT : 2022UPASL010 . tel-03961470

**HAL Id: tel-03961470**

**<https://theses.hal.science/tel-03961470>**

Submitted on 29 Jan 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Integrative Clinico-Molecular Analysis and Single-cell RNA Sequencing To Unravel The Origin of Atypical Teratoid Rhabdoid Tumors

*Analyse intégrative clinico-moléculaire et séquençage ARNseq de  
cellule-unique afin d'identifier l'origine des tumeurs rhabdoïdes  
tétratoïdes atypiques*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 582 Cancérologie : biologie-médecine-santé (CBMS)  
Spécialité de doctorat : Sciences de la vie et de la santé  
Graduate School : Sciences de la vie et santé. Référent : Faculté de médecine

Thèse préparée dans l'unité de recherche **Cancer, Hétérogénéité, Instabilité et  
Plasticité U830 (Institut Curie, Université PSL)**  
sous la direction de **Franck BOURDEAUT**, MD PhD, Institut Curie

Thèse soutenue à l'Institut Curie, le 28 Janvier 2022, par

**María-Jesús LOBÓN IGLESIAS**

## Composition du Jury

<b>Veronique MINARD</b> PU-PH, Université Paris Saclay	Présidente
<b>Nicolas ANDRE</b> PU-PH, Aix-Marseille Université	Rapporteur & Examineur
<b>Ulrich SCHÜLLER</b> Professeur, UKE Hamburg-Eppendorf	Rapporteur & Examineur
<b>Franck BIELLE</b> PU-PH, Sorbonne Université	Examineur
<b>Cédric MAURANGE</b> PhD DR, Aix-Marseille Université	Examineur
<b>Franck BOURDEAUT</b> MD PhD, Institut Curie	Directeur de thèse

**Integrative Clinico-Molecular Analysis**  
**and**  
**Single-cell RNA Sequencing**  
**To Unravel The Origin of Atypical Teratoid**  
**Rhabdoid Tumors**

*A Sebastian C.*

*A Georgios K.*

*A Raquel Muñoz Navarro.*

*Tu Alianza Rebelde.*

*"Observar sin pensar es tan peligroso como pensar sin observar."*

Santiago Ramón y Cajal

*"... it is easier to smash an atom than a prejudice."*

Albert Einstein

## Acknowledgements

*I want to express my deepest gratitude to all the jury members.*

*Franck, merci de m'avoir donné l'opportunité de travailler dans votre équipe. Te estaré siempre agradecida.*

*Gracias a todos los que me han inspirado antes de llegar, porque, por muchos años que pasen, sigo pensando en vosotros, Lola Lluch, Olivier Dulac, Jacques Grill.*

*Merci à tout l'équipe RTOP, chacun et chacune d'entre vous, pour votre façon d'être, drôle et magique, merci pour votre gentillesse et votre bienveillance pendant ces 3 difficiles années.*

*Gracias Irene por escucharme siempre. Zhiyan, pour être si méthodique et si efficace, pour ta patience et ton soutien. Céline et Sandrina, pour les discussions, pour votre esprit critique, Yasmine, pour râler et sourire au même temps, Rachida, pour ta sincérité, Charles et Angela, pour être le meilleure duo dynamique ... Alex pour les rappels...Mandy et Solène..Jaydutt... pour m'accueillir dans le petit bureau bio-info...*

*Merci à Mam, pour avoir été sans doute le meilleur collègue de bureau que je pouvais avoir, merci d'avoir été si pédagogue et d'avoir rendu la difficulté facile, grâce à toi la bio-informatique fait maintenant partie de ma vie.*

*Merci a Dominique G. y a Sergio R-R., por todos y cada uno de vuestros buenos días, por darme la sensación de llegar a casa.*

*Gracias a Amaury por tu recibimiento, por tu cercanía. Gracias a todo el equipo de Eliane, a Leticia, Jimena, Valeria...sin duda no podía haber tenido mejores maestr@s en la animaleria y en el single cell.*

*Merci à l'équipe Baffet, je te fais confiance Magali pour arriver loin avec nos petits cerveaux Nestin!*

*Mil merci à la Plateforme NGS, pour votre magnifique organisation et votre disponibilité.*

*Merci à la Plateforme Imagerie, a Chloe Guedj et Anne-Sophie Mace, pour tant que vous m'avez appris.*

*Merci les amis d'ici et d'ailleurs, Lucie, Elsa, Anne-So, Clara, Aphaia, Arielle ... Marijose, Candela, Peter, Inma, Andrea, Macarena, Rocio...por estar ahí.*

*A Raquel, por tu preciosa ayuda, por tu amistad, por tu lucha. Aunque tu ausencia sea irrespirable, tu amor y tu fuerza nos impulsa cada día.*

*A mi familia, a mi padre, por su ejemplo, el valor del trabajo, y de la vida.*

*Y por su supuesto, a Jacob y Olivier, que han dado forma a esta tesis.*

## TABLE OF CONTENTS

<i>Abstract</i> .....	6
<i>Résumé</i> .....	7
<i>List of abbreviations</i> .....	9
<i>Introduction</i> .....	11
<b>General aspects</b> .....	12
Historical context of Atypical Teratoid Rhabdoid Tumors.....	12
Clinical presentation and diagnosis .....	12
Genetics of ATRT.....	13
Standard of care and outcome .....	14
<b>The SWI/SNF chromatin remodeling complex</b> .....	15
Structure and function.....	15
The role of SWI/SNF in normal brain development.....	16
The particularities of the BAF47 ( <i>SMARCB1</i> ) subunit .....	18
The SWI/SNF complex and Human Genetic Disorders .....	19
<b>The Cell of Origin</b> .....	25
Molecular Heterogeneity of ATRTs.....	25
Mouse Models of ATRTs .....	29
Phylogenetic analyses and single-cell approaches.....	31
Location, location, location .....	33
<i>Project Rationale and Objectives</i> .....	35
<i>Results</i> .....	37
<i>Material et Methods</i> .....	86
<i>Discussion</i> .....	100
<i>Perspectives</i> .....	106
<b>The Nestin Mouse Model</b> .....	107
Ex-vivo characterization of <i>Smarb1</i> <sup>F/F</sup> : <i>Nestin</i> <sup>Cre+</sup> E18 brains.....	107
Single Cell RNAseq analysis of Nestin model E18 brains.....	108
<i>Bibliography</i> .....	115
<i>Appendix Figures</i> .....	133
<i>Appendix Tables</i> .....	139

## Abstract

Atypical teratoid rhabdoid tumors (ATRT) are rare and aggressive malignancies of the central nervous system (CNS) affecting infants and young children and characterized by a biallelic inactivation of SMARCB1 tumor suppressor gene in an otherwise very simple genome. Based on methylation and expression profiling, recent studies have pointed out the molecular diversity of these tumors, that are now divided in at least three subgroups, i.e. the so-called MYC-, TYR, and SHH-ATRTs. This molecular diversity is suggestive of various cells of origin for each subtype, though none of these putative cells of origin is undoubtedly identified at 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).

In this work, we have provided a detailed description of different anatomical ATRT locations to obtain precise information about the putative tissue of origin. Then, we have performed integrative analysis of anatomical location (n= 51), transcriptomic (n= 49) and epigenetic data (n= 54). This analysis has allowed us to describe and characterize 4 distinct anatomical-molecular ATRT subgroups: cranial nerves, cerebral cortex and spinal cord tumors, CNCS-ATRT; Basal ganglia and Intraventricular tumors, BG/IV-ATRT; cerebellar anterior lobe tumors, CAL-ATRT; middle cerebellar peduncle and inferior cerebellar lobe tumors, MCP/ICV-ATRT. Next, we have sought to investigate the relevance of our mouse models for human disease. We have found that the Rosa26: MYC mouse model perfectly recapitulates molecular features and clinical phenotypes of a subset of human ATRT (CNCS-ATRT) and support the hypothesis of an origin in neural crest cells. The Rosa26: SHH tumors are phenotypically similar to the BG/IV-ATRT and are characterized by the expression of gene markers from the ganglionic eminences. Finally, the CAL-ATRT are characterized by the expression of midbrain-hindbrain boundary (MHB) genes. Although we don't have a faithful mouse model for this location, the integrative analysis has allowed us to identify the CAL-ATRT subgroup and lead us to perform single cell analysis of three human samples exclusively from this location. Consecutively, we have demonstrated at the single-cell level, by the first time so far, the dedifferentiation of neuronal progenitor cells and the involvement of Notch pathway in the malignant transformation of CAL-ATRT.

Key words: ATRT subgroups; Anatomical Correlation; Integrative Analysis; Single Cell RNAseq; Notch signalling pathway

## Résumé

Les tumeurs rhabdoïdes tératoïdes atypiques (ATRT) sont des tumeurs malignes rares et agressives du système nerveux central (SNC) affectant les nourrissons et les jeunes enfants et caractérisées par une inactivation biallélique du gène suppresseur de tumeur SMARCB1 dans un génome par ailleurs très simple. Sur la base du profil de méthylation et d'expression, des études récentes ont mis en évidence la diversité moléculaire de ces tumeurs, qui sont maintenant divisées en au moins trois sous-groupes, à savoir les tumeurs dites MYC-, TYR- et SHH-ATRT. Cette diversité moléculaire suggère l'existence de diverses cellules d'origine pour chaque sous-type, bien qu'aucune de ces cellules d'origine n'ait été identifiée à l'heure actuelle. L'analyse du profil d'expression a été faiblement informative sur le tissu d'origine mais a donné des indications sur le lignage (neurogénique et mélanogénique pour les sous-types SHH et TYR respectivement) et sur certains gènes et voies de développement exprimés de manière récurrente (SHH et voie NOTCH pour le sous-type SHH ; groupes HOX pour le sous-type MYC).

Dans ce travail, nous avons fait une description détaillée des différentes localisations anatomiques des ATRT afin d'obtenir des informations précises sur le tissu d'origine. Ensuite, nous avons réalisé une analyse intégrative de la localisation anatomique (n=51), avec les données transcriptomiques (n=49) et épigénétiques (n=54). Cette analyse nous a permis de décrire et de caractériser 4 sous-groupes anatomico-moléculaires d'ATRT : les tumeurs des nerfs crâniens, du cortex cérébral et de la moelle épinière, CNCS-ATRT ; les tumeurs des ganglions de la base et intraventriculaires, BG/IV-ATRT ; les tumeurs du lobe antérieur du cervelet, CAL-ATRT ; et les tumeurs du pédoncule cérébelleux moyen et du vermis cérébelleux inférieur, MCP/ICV-ATRT. Ensuite, nous avons cherché à étudier la pertinence de nos modèles murins chez l'homme. Nous avons constaté que le modèle de souris Rosa26 : MYC récapitule parfaitement les caractéristiques moléculaires et les phénotypes cliniques d'un sous-ensemble d'ATRT humains (CNCS-ATRT) et soutient l'hypothèse d'une origine dans les cellules de la crête neurale. Les tumeurs Rosa26 : SHH sont phénotypiquement similaires aux BG/IV-ATRT et sont caractérisées par l'expression de marqueurs génétiques des éminences ganglionnaires. Enfin, les tumeurs CAL-ATRT humaines sont caractérisées par l'expression de gènes typiques de la région entre le cerveau moyen et le cerveau postérieur pendant le développement embryonnaire. Bien que nous ne disposions pas d'un modèle de souris récapitulant cette localisation, l'analyse intégrative nous a permis d'identifier le sous-groupe CAL-ATRT et nous a conduit à effectuer une analyse RNAseq cellule-unique de trois

échantillons humains provenant exclusivement de cette localisation. Nous avons consécutivement démontré au niveau cellule-unique, pour la première fois à ce jour, la dédifférenciation des cellules progénitrices neuronales et l'implication de la voie Notch dans la transformation maligne des CAL-ATRT.

Mots clés : ATRT sous-groupes ; corrélation anatomique ; analyse intégrative ; RNAseq cellule-unique; voie de signalisation Notch

## **List of abbreviations**

ATRT: Atypical Teratoid Rhabdoid Tumor

BG/IV-ATRT: Basal Ganglia and Intraventricular- ATRT

BG: Basal Ganglia

CAL-ATRT: Cerebellar Anterior Lobe- ATRT

CN: Cranial Nerves

CNCS-ATRT: Cranial Nerves, Cerebral Cortex and Spinal Cord- ATRT

CNS: Central Nervous System

Cx: Cerebral Cortex

CC: Cell Cluster

CSS: Coffin-Siris Syndrome

DEA: Differential Expression Analysis

ECRT: Extra-cranial Rhabdoid Tumor

ELPiGraph: Elastic Principal Graph Algorithm

EMT: Epithelia Mesenchymal Transition

ETMR: Embryonal Tumor with Multiple Rosettes

GE: Ganglionic Eminences

GCP: Granule Cell Precursor

HES: Haematoxylin Eosin Staining

IHC: Immunohistochemistry

GSEA: Gene Set Enrichment Analysis

GBM: Glioblastoma

MCP/ICV-ATRT: Middle Cerebellar Peduncle and Inferior Cerebellar Vermis- ATRT

MHB: Midbrain-Hindbrain Boundary

NDD: Neuro Developmental Disorder

NCBRS : Nicolaidis-Baraitser Syndrome

NCC : Neural Crest Cell

NF2 : Neurofibromatosis type 2

NN : Non Neuronal

NP: Neuronal Progenitor

NPL1: Neuronal Progenitor Like 1

NPL2: Neuronal Progenitor Like 2

PNS: Peripheral Nervous System

RT: Rhabdoid Tumors

RTPS: Rhabdoid Tumor Predisposition Syndrome

SC: Schwann cell

SC: Spinal Cord

sPLS-DA: sparse Partial Least Square Discriminant Analysis

TF: Transcription Factor

UMAP: Uniform Manifold Approximation and Projection

UD: Undifferentiated

# **Introduction**

## General aspects

### Historical context of Atypical Teratoid Rhabdoid Tumors

Embryonal brain tumors (EBTs) are highly aggressive and the most common CNS tumor type in the first 4 years of life (Li et al., 2020). They are undifferentiated or poorly differentiated tumors and have significant overlap in the histologic and radiologic appearance. Although EBT is a category defined since 1979, it has greatly evolved thanks to the improvement of imaging and molecular techniques. Thus, EBT encompasses today different types of tumors such as medulloblastoma, atypical teratoid rhabdoid tumors (ATRT), pinealoblastoma, embryonal tumor with multiple rosettes (ETMR), and embryonal tumors with FOXR2 activation or BCOR alteration, that includes molecular parameters for their diagnosis (Louis et al., 2021).

ATRT are EBT firstly described by Lefkowitz et al in 1987 as a tumor containing an unusual mixture of primitive neuroepithelial, surface epithelial and mesenchymal elements (Rorke et al., 1996). These tumors also contained varying amounts of rhabdoid cells (Figure 1 a, b) similar to the already described infant malignant rhabdoid tumors located outside the CNS i.e. in kidney, liver or soft tissues (extracranial rhabdoid tumors, ECRT). The recognition of bi-allelic inactivation of the tumor suppressor gene *SMARCB1/BAF47* (95%) (Versteeg, 1998) and rarely *SMARC4A/BRG1* (5%) (Schneppenheim et al., 2010) in rhabdoid tumors (RT) allowed a better distinction as separate entity from other EBT (Figure 1c). Until its official inclusion in the WHO classification in 2000, ATRT were probably often misdiagnosed as primitive neuroectodermal tumors (PNET) or medulloblastomas (Rorke et al., 1996). ATRT constitute 65% of the cases of RT and are the focus of my PhD work.

### Clinical presentation and diagnosis

ATRT are specially suspected in children less than 3 years and their clinical presentations are heterogeneous and related to the location and age at the time of diagnosis. They may be revealed by different symptoms such as lethargy, vomiting, failure to thrive, hemiplegia, headaches or cranial nerve problems (Rorke et al., 1996). After clinical suspicion and MRI confirming the presence of a CNS tumor in a very young children, surgery or biopsy of the tumor are mandatory depending on its location, extent and the patient's condition.

The diagnosis of ATRT is validated based on neuropathologist observations of haematoxylin and eosin (HE) and immunohistochemistry (IHC) staining of tumor resection or biopsy. ATRT may show typical rhabdoid cells with an eccentric round nucleus with a prominent nucleolus

and a plump cell body adjacent to or surrounded by epithelial or mesenchymal components. Immunohistochemical studies may show positivity for epithelial membrane antigen (EMA), vimentin, smooth-muscle actin and typical diffuse loss of BAF47 nuclear expression in tumor cells(+ref). Finally, molecular and cytogenetic analysis will confirm the bi-allelic loss of function of *SMARCB1*.

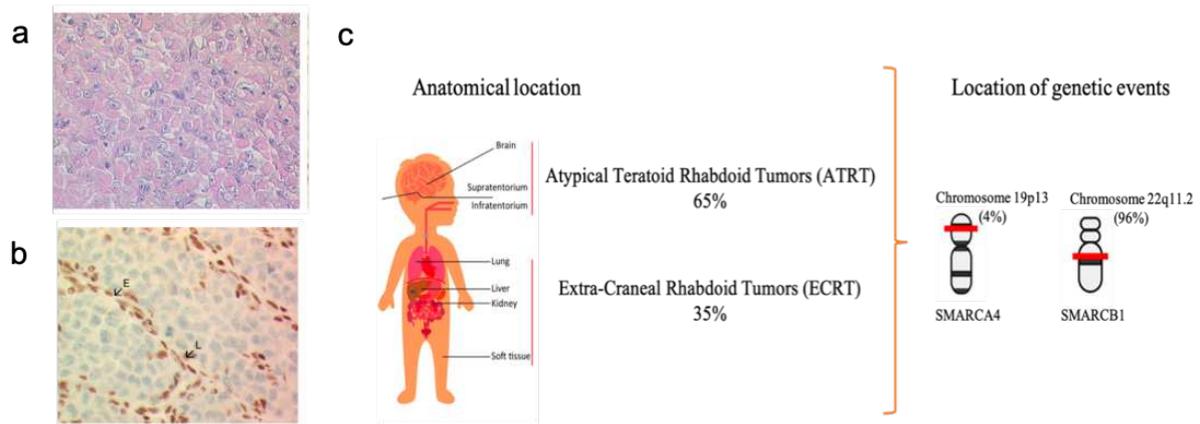


Figure 1. a. High-power photomicrograph shows hyper cellularity and pleomorphic differentiation (teratoid appearance), with a few rhabdoid cells demonstrating a plump eosinophilic cytoplasm and an eccentrically placed nucleus. (H-E stain.) b. High-power photomicrograph from BAF47 immunohistochemistry shows diffuse loss of nuclear expression/positivity in the tumor cells. There is retained staining (brown nuclei) in normal endothelial cells, which serve as a positive control. This finding is diagnostic of AT/RT (adapted from Bourdeaut et al., 2007). c. Schematic representation of different of malignant rhabdoid tumors and molecular and location and frequency of most common genetic events (adapted from Tang et al. Trends in Cancer, 2016 and Del Baldo et al. Frontiers in Oncology, 2021).

### Genetics of ATRT

The majority of RTs arise as a consequence of biallelic inactivation of the *SMARCB1* (aliases BAF47/INI1/hSNF5) gene (Versteeg, 1998) and rarely *SMARCA4* (encoding the BRG1 protein) (<5%). They are both members of the human SWI/SNF chromatin-remodeling complex (Wilson and Roberts, 2011) which will be described in detail in the following section. Although other tumors such as chordoma and renal medullary carcinoma share with ATRT a similar complete loss of *SMARCB1* expression, the great particularity of ATRTs is that they don't show any other recurrent genetic event driving to the malignant transformation in an overall simple genome (Kieran, 2012; Lee et al., 2012). The spectrum of events leading to this complete loss of function includes whole-gene deletions, large intragenic deletions/duplications, small out-of-frame intragenic deletion/insertions, splice-site mutations

and nonsense mutations. Importantly, the *SMARCB1* biallelic-inactivation can occur in a context of predisposition cancer syndrome (called Rhabdoid Tumor Predisposition Syndrome, RTPS). This establishes *SMARCB1* as a bona fide tumor suppressor gene, as defined by Knudson in the “two-hit model”. RTPS is an autosomal dominant cancer predisposition syndrome involving *SMARCB1* gene (RTPS1) or rarely *SMARCA4* (RTPS2) (Sredni and Tomita, 2015). RTPS1 constitutes up to 25% of newly diagnosed ATRTs and are characterized by the occurrence before 12 months of age and the tendency to develop synchronous or multifocal tumors with aggressive clinical features (Biegel et al., 1999; Frühwald, 2020). RTPS2 is much less frequent and related to the occurrence of SMARCA4-deficient rhabdoid tumors in infants and small cell carcinomas of the ovary, hypercalcemic type (SCCOHT), in slightly older females (Del Baldo et al., 2021). Of note, *SMARCB1* pathogenic variants observed in RTPS1 are almost exclusively truncating mutations, while more hypomorphic mutations (i.e. missense or splice-site variants) rather predispose to schwannomas or meningiomas, more indolent tumors that usually show a mosaic pattern of expression of BAF47 (Del Baldo et al., 2021; Smith et al., 2014). Beside their universal stable genome, three ATRTs molecular subgroups have been described (Han et al., 2016; Johann et al., 2016; Torchia et al., 2016), now consensually referred to as ATRT-SHH, ATRT-TYR and ATRT-MYC (Ho et al., 2020) with still uncertain prognostic implications (Frühwald, 2020; Reddy et al., 2020; Upadhyaya et al., 2021) (more details below). More recently, other authors reported SMARCA4 deficient ATRT as a distinct ATRT subgroup (Holdhof et al., 2021a). In my thesis I have focused in ATRT Smarcb1 deficient since they are far more frequent.

### **Standard of care and outcome**

RT are classically treated by multimodal treatment regimens, including surgery, conventional chemotherapy, intrathecal chemotherapy, high dose chemotherapy with stem cell rescue, radiotherapy and maintenance treatment (Bartelheim et al., 2016; Chi et al., 2009; Fossey et al., 2017; Lafay-Cousin et al., 2012; Reddy et al., 2020; Schrey et al., 2016). Despite this intensive therapeutic approach, ATRTs remain an aggressive deadly disease with recent studies showing 40% of 4 year event free and overall survival (Reddy et al., 2020). Furthermore, the management of these patients is frequently limited by the complexity of delivering these intensive treatments and radiotherapy to very young children. In addition, significant lifelong disability and neurocognitive impairment are commonly observed in survivors of pediatric

CNS tumors (Brinkman et al., 2018). Thus, therapies for ATRT remain medically and ethically challenging.

New therapeutic insights focus on the key signal transduction pathways identified in the pathogenesis of rhabdoid tumors (EZH2, DNMT, HDAC, CDK4/6/ Cyclin D1/Rb, AURKA, SHH/GLI1, Wnt/ $\beta$ -Catenin pathway) (Jagani et al., 2010; Mora-Blanco et al., 2014) as well as on the growing evidence of the potential role of targeted immunotherapy (Chun et al., 2019; Leruste et al., 2019). Advances in the understanding of ATRT biology could lead to the discovery of newer specific diagnostic markers, subgrouping with clinicopathologic significance and effective disability-spared targeted therapies (Nemes and Frühwald, 2018). Although it is true that this is a common objective of all laboratory groups working on pediatric brain tumors, it is especially so in our group specialized in translational research on ATRT. These necessary advances are closely related to the discovery of cells of origin and the disruption of the developmental pathways leading to the tumor initiation that are the concern of my PhD work.

## **The SWI/SNF chromatin remodeling complex**

### **Structure and function**

The SWI/SNF chromatin remodeling complexes constitute a highly related family of multisubunit complexes. They consist of 12–15 subunits and uses energy obtained from adenosine triphosphate (ATP) hydrolysis to remodel nucleosomes and modulate gene transcription (activation or repression). To accommodate changes in strategies of gene regulation, the SWI/SNF complex has had to evolved extensively in the past from yeast to mammals (Fig. 1)(Kadoch and Crabtree, 2015). In this sense, the BAF complex subunits are widely expressed, and the specific combinatorial assemblies and subunit switching contribute to the formation of cell lineage specific BAF complexes capable of instructing specific cell fates such as heart development, muscle development, embryonic stem cell pluripotency and neural development and function (more details in the next section).

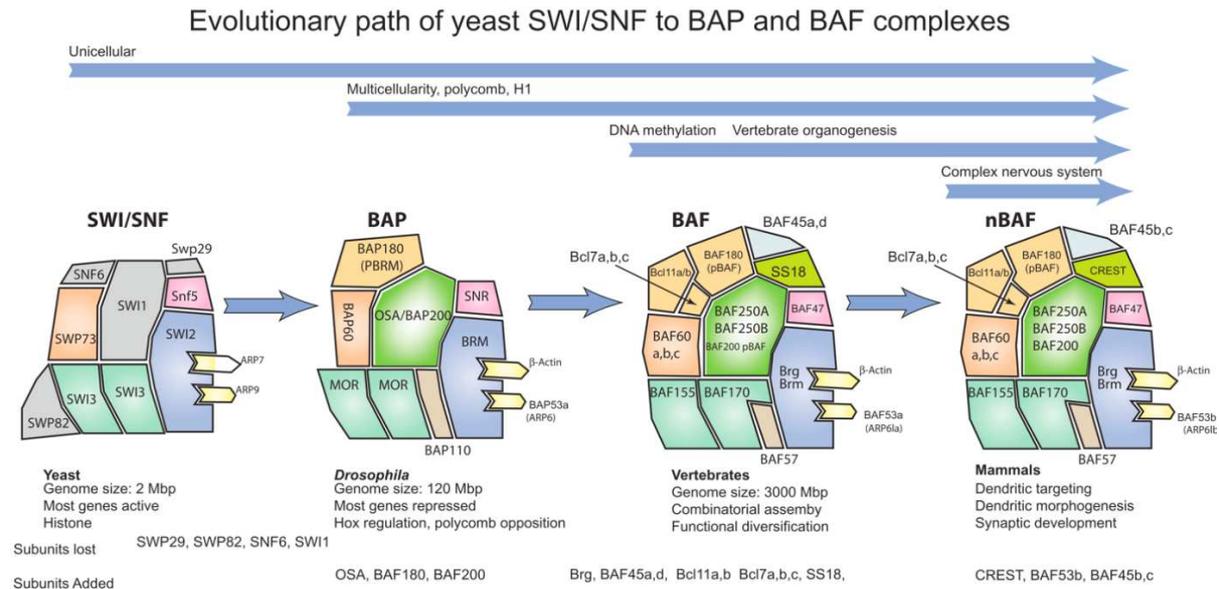


Figure 2. Evolution of the yeast SWI/SNF complexes to the fly BAP and vertebrate BAF complexes (colors indicate homology). The development of multicellularity and the need to repress most genes is coupled with the appearance of polycomb-mediated repression, histoneH1, and major changes in the subunit structure of SWI/SNF complex. The emergence of vertebrates, appearance of a much larger genome, DNA methylation, and vertebrate complexity is accompanied by another transition in subunit structure and combinatorial assembly. Finally, with the emergence of a complex nervous system, four new neuron-specific subunits enter the complex and are essential for dendritic morphogenesis, synaptogenesis, and connectivity within the nervous system. Adapted from Kadoch and Crabtree, 2015 (Kadoch and Crabtree, 2015). To note, BAF47 is also known as SMARCB1 and BRG1 as SMARCA4.

### The role of SWI/SNF in normal brain development

The SWI/SNF complex has been shown to be required in neural development and the establishment of fully functional nervous system (Bachmann et al., 2016; Narayanan et al., 2015; Nguyen et al., 2016). An important switch in BAF complex composition occurs from embryonic stem cell (ESC) BAF (esBAF) to self-renewing neural stem and progenitor cell BAF (npBAF) and post-mitotic neuron BAF (nBAF) (Figure 3). The shift from npBAF to nBAF coincides with the mitotic exit of neural precursors and is distinguishable by the replacement of Baf53a by 53b, SS18 by CREST, Baf45a/d by Baf45b/c and changed expression levels of Baf155 and Baf170 (Lessard et al., 2007; Olave et al., 2002; Wu et al., 2007). Accordingly, Baf45A and Baf53A are essential for the proliferation of neural stem and progenitor cells and Baf53B has essential roles in developing neurons (Cenik and Shilatfard, 2021).

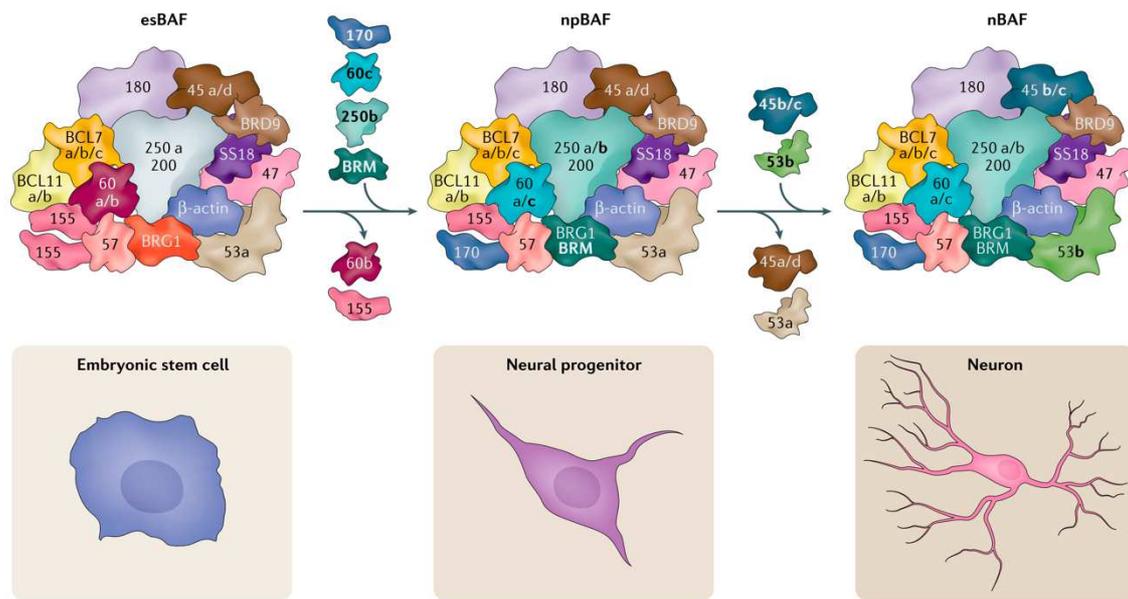


Figure 3. Subunit switching of BAF complexes during neurogenesis. This figure shows stage- specific subunit compositions at different developmental time points. Embryonic stem cell-specific BAF (esBAF) uses BRG as its central ATPase, whereas neural progenitor BAF (npBAF) and neuronal BAF (nBAF) complexes can incorporate either BRG or BRM. Other subunits are also evicted and replaced by different proteins, as indicated in the figure. Adapted from Cenik et al. 2021

The npBAF complexes modulate Notch and Sonic hedgehog (Shh) signaling to support proliferation and keep cells in a state poised for differentiation to post-mitotic neurons (Lessard et al., 2007). Moreover, the switch from BAF53A to BAF53B is mediated by the microRNAs (miRNAs) miR-9\* and miR-124, that are repressed by REST in neural progenitors and highly expressed in post mitotic neurons. This BAF53B subunit is essential for dendritic growth and branching (Wu et al., 2007) and for long-term memory and synaptic plasticity (Vogel-Ciernia et al., 2013).

The subunits of the SWI/SNF complex that are not specific for neural BAF complexes (npBAF or nBAF) have also important roles in neural development (Figure 4). For instance, Brg1 and Baf155 have an important and dosage-sensitive role in neural differentiation and development since heterozygous mutants exhibit neural tube closure defects and exencephaly (Bultman et al., 2000; Kim et al., 2001). Furthermore, BRG1 subunit is essential for the development of major essential brain structures (Holdhof et al., 2020; Holdhof et al., 2021b) and required during neuronal and glial differentiation (Marathe et al., 2013; Matsumoto et al., 2006; Yu et al., 2013). The BAF170-containing BAF complexes modulate neurogenesis and control cortex

size (Tuoc et al., 2013) and along with BAF155 are essential for forebrain development (Narayanan et al., 2015). Finally, in the peripheral nervous system, BAF60A is essential for the differentiation of Schwann cells (Weider et al., 2012).

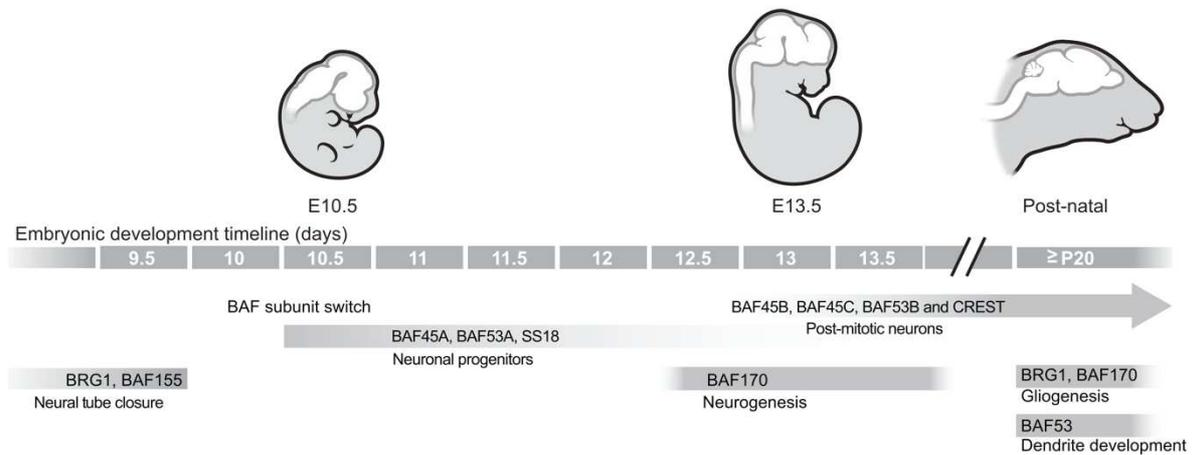


Figure 4. Chromatin remodeler functions and transitions in neural development. Stages of mammalian neural development are pictured, with mouse embryonic day (E) indicated below each pictured stage. Key demonstrated roles for specific chromatin remodeling factors are positioned at relevant stages below the developmental timeline. Adapted from Hota et Bruneau, 2016.

### The particularities of the BAF47 (*SMARCB1*) subunit

*SMARCB1* is essential for embryonic development since homozygous knockout of *Smarb1* leads to early embryonic lethality (Guidi et al., 2001; Han et al., 2016; Klochendler-Yeivin et al., 2000). In addition, it has been shown that *SMARCB1* is a key regulator of stem cell-associated programs and the equilibrium of the stem versus differentiated cell balance depends on its epigenetic antagonism with the Polycomb gene *EZH2* (Wilson et al., 2010).

In order to better understand the occurrence of ATRT, the specific role of *SMARCB1* in neural development should be further investigated. Parisian et al. have studied the potential interactions between *SMARCB1* loss and the process of neural development by an inducible *SMARCB1* loss-of-function system into human induced pluripotent stem cells (iPSCs) by directed neuronal differentiation or differentiation into cerebral organoids. They showed that *SMARCB1* loss during neuronal differentiation leads to impairment of neuronal commitment and lack of stability among neural progenitors (NPCs). Most importantly, they showed that the transcriptional effects of *SMARCB1* loss can vary dramatically in different cell types leading to opposite phenotypic and transcriptional effects (Parisian et al., 2020). Moreover and in agreement with Han et al., they showed that cells are especially sensitive to the effects of *SMARCB1* loss in an early developmental window (Han et al., 2016).

In a more restricted way, Moreno et al. studied specifically the impact of *SMARCB1* loss in cerebellar granule cell precursors and found that it resulted in severe decrease proliferation of granule neuron precursors leading to an hypoplastic cerebellum. They also showed that *SMARCB1* loss in hGFAP-positive neural precursors led to lamination disturbances and a thinning of the cerebral cortex. These results suggest that *SMARCB1* is important for cerebellar and forebrain development (Moreno et al., 2014). Finally, other authors have recently shown that *SMARCB1* is essential for hESC super-enhancer silencing in neural differentiation conditions (Langer et al., 2019).

Taken together, these studies show that *SMARCB1* is important for normal brain development and this fact is reinforced by the existence of neurodevelopmental disorders caused by *SMARCB1* mutations (more details below).

### **The SWI/SNF complex and Human Genetic Disorders**

Mutations in genes encoding components of the SWI/SNF complex are found in neurodevelopmental disorders (NDD) and in cancer. They belong to the most frequently mutated genes in human NDD, and close to 20% of all human cancers harbor mutations in these genes (Kadoch and Crabtree, 2015; Masliah-Planchon et al., 2015; Santen et al., 2012; Sokpor et al., 2017). Interestingly, heterozygous germline mutations predispose to both conditions but up to now it is unclear how such distinct disease entities can be caused by mutations in identical genes. Hereafter we will first describe the most known syndromes and their relationship with the different components of SWI/SNF complex. In the second part, we will detail the role of SWI/SNF complexes in tumorigenesis. In both parts, we will pay a special attention to BAF47 encoded by *SMARCB1* gene.

### **SWI/SNF complex and NeuroDevelopmental Disorders (NDD)**

As previously detailed, proper functioning of the BAF complexes plays critical roles in neural development, including the establishment and maintenance of neural fates and functionality. Indeed, BAF complex subunits are linked to neurodevelopmental disorders such as Coffin-Siris syndrome (CSS), Nicolaides-Baraitser syndrome (NCBRS), Kleefstra's syndrome spectrum, autism spectrum disorder and schizophrenia (Figure 5).

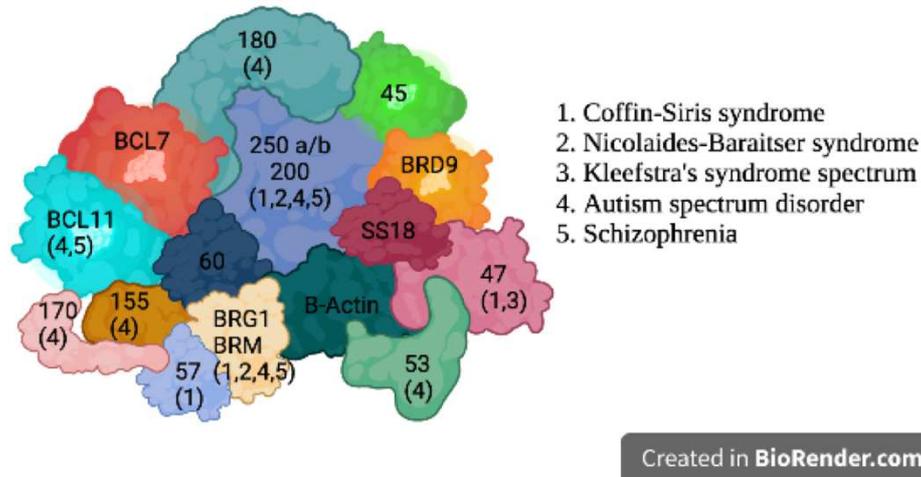


Figure 5. Chromatin remodeling BAF (mSWI/SNF) complex in neural development and disorders. The figure shows BAF subunits implicated in the different neurological disorders. For example, BAF250a, BAF250b and BAF200 have been involved in CSS, NCBRS, autism spectrum disorder and schizophrenia, while BAF57 has been only involved in Coffin-Siris syndrome.

*ARID1B* (BAF250b subunit) is the most frequently mutated gene in SWI/SNF-related NDD in general (Wright et al., 2018) and also the most frequently mutated gene in Coffin–Siris syndrome. Besides this, *SMARCA2* is predominantly associated with intellectual disabilities and missense mutations of *SMARCA2* is the primarily cause of NCBRS. Interestingly, their paralogs (*ARID1A* and *SMARCA4*) are more closely related to cancer.

From a clinical point of view, NCBRS and CSS share remarkable features including intellectual disability and typical somatic characteristics, especially sparse hair, low frontal hairline, large mouth with thick and everted lips, and hands and feet anomalies. The hallmark differences between NCBRS and CSS come from specific features of the hands and feet (Kosho et al., 2014a). Although the reason why the different pathogenic variants of BAF subunits lead to specific neurological phenotypes in patients remains largely unknown, certain correlations between genotype and phenotype have been described (Kosho et al., 2014b). For instance, CSS patients with *SMARCB1* mutations display the most striking physical features and deep problems associated with developmental delay (Kosho et al., 2014b), but with least distal limb anomalies (Mari et al., 2015). To note, *SMARCB1* mutations in CSS are *de novo* germline heterozygous non-truncating missense mutations or small in-frame deletions within exons 8 and 9. Interestingly, Filatova et al. generated mice (*Smarchb1<sup>+/-</sup>inv NesCre<sup>+/-</sup>*) with a heterozygous reversible *Smarchb1* disruption in neural stem/progenitor cells. They obtain a mouse model recapitulating brain abnormalities encountered in CSS individuals: reduction or

absence of forebrain commissures, absence or hypoplasia of the septum pellucidum and, curiously, cerebellar vermis hypoplasia. They noticed that in this model the *Smarca1* transcript levels are reduced by about 30% in embryonic brain tissue, while there is no reduction in *Smarca1* transcript levels in heterozygous germline *Smarca1* knockout animals (*Smarca1*<sup>+/-</sup>) suggesting that the neurodevelopmental disorders or malignant tumors could depend on a transcriptional compensation by the second intact *Smarca1* allele. These findings reinforce the idea that a better understanding of normal and pathological neurodevelopment can help us to unravel the mechanism of tumor development.

### **SWI/SNF complex and Cancer**

The SWI/SNF chromatin remodeling complex is the most frequently involved in human malignancies. Mutations in SWI/SNF subunit genes, structural abnormalities, or epigenetic modifications that lead to reduced or aberrant expression of members of the SWI/SNF complex have been reported in 20% of human cancers (Kadoch et al., 2013; Shain and Pollack, 2013). The most commonly affected BAF subunit in cancer is BAF250A (*ARID1A*), mainly by frameshift and nonsense mutations resulting in a truncated protein product. Interestingly, some tumors are related to mutations in only a specific subunit (i.e. ovarian endometrial cancer is only related with mutations in *ARID1A* but in no other subunits), while others tumors are related to mutations in different subunits (i.e. urothelial cancer) (Figure 6)(Mittal and Roberts, 2020). In addition, some tumors show mutations only in a specific subunit and in nearly all the cases (i.e. small-cell carcinoma of the ovary, hypercalcaemic type and rhabdoid tumors). These findings suggest differential roles for individual SWI/SNF components and tissue-specific vulnerabilities (Shain and Pollack, 2013).

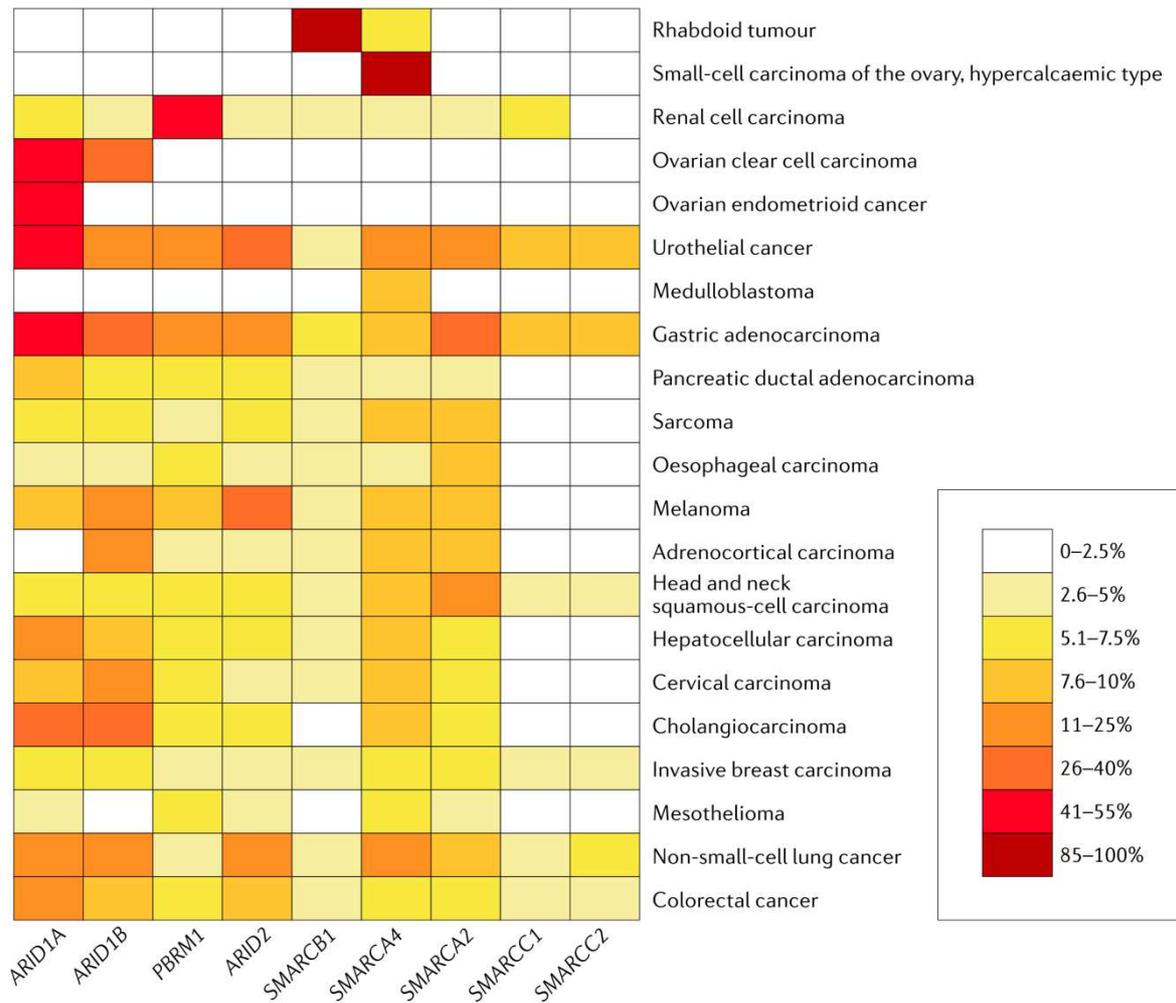


Figure 6. Frequency and pattern of SWI/SNF subunit mutations across human cancers. The heatmap depicts the frequency of non-synonymous mutations and deletions in select genes encoding components of SWI/SNF complexes across cancer types. Overall, the figure depicts the high prevalence of mutations affecting nine SWI/SNF subunits and the context-specificity of these mutations. Adapted from Mittal et Roberts, 2020.

Importantly, there are malignancies in which SWI/SNF complex proteins act as oncogenes and those in which they act as tumor suppressors. Hereafter I will describe two of examples of cancer driven respectively by these two mechanisms: synovial sarcomas and rhabdoid tumors.

### SWI/SNF complex proteins can act as oncogenes

The synovial sarcoma are tumors caused by t(X;18) chromosomal translocation, fusing 78 amino acids of the protein SSX to the dedicated BAF complex subunit SS18 in nearly all patients. Unlike rhabdoid tumors previously detailed, caused by a biallelic loss of BAF47, tumorigenesis in synovial sarcoma occurs in spite of a remaining wild-type allele. In fact, the transcription of the wild-type allele is decreased in sarcoma cells in general and the SS18-SSX

fusion protein is preferably incorporated into the complex, leading to the degradation of the monomeric wild-type protein. In these conditions, the fusion protein integrates into the BAF complex, but evicts the core subunit BAF47. This pathological complex can be retargeted to several developmental loci, including SOX2, PAX6, PAX3 and PAX7, leading to loss of repressive H3K27me3, misexpression of these genes and characteristic stem-cell expression pattern.

### **SWI/SNF complex proteins can act as tumor suppressors**

SWI/SNF proteins most commonly act as tumor suppressors, although the molecular mechanism can vary among cancers. The tumor suppressor role of SMARCB1 has been highlighted in different studies with genetically engineered mouse models. *SMARCB1*-heterozygous mice develop tumors consistent with rhabdoid tumors (Guidi et al., 2001; Klochendler-Yeivin et al., 2000; Roberts and Galusha, 2000) and conditional biallelic inactivation results in fully penetrant phenotype with the development of lymphomas and rhabdoid tumors at a median onset of only 11 weeks (Roberts et al., 2002). Although SWI/SNF complexes have been implicated in several mechanisms of DNA-damage repair (Brownlee et al., 2015; Chen et al., 2019a; Qi et al., 2015), these results along with early onset of rhabdoid tumor in children abrogates for transcriptional dysregulations of lineage-specific transcriptional programs as the central mechanism of tumorigenesis (Mathur and Roberts, 2018; Shain and Pollack, 2013). Furthermore, it has been suggested that the ability of SMARCB1 deletion to cause tumorigenesis may be dependent on the epigenetic environment of a particular stage in cellular differentiation (Mathur and Roberts, 2018; Wilson and Roberts, 2011).

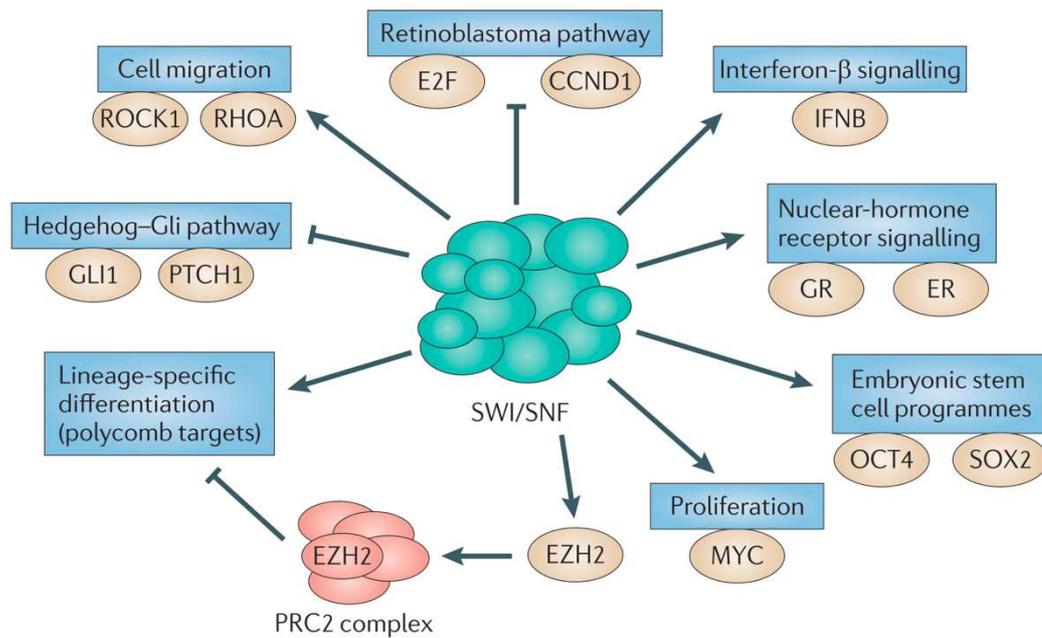


Figure 7. Target pathways implicated in the tumor suppressor activity of SWI/SNF complexes. The SWI/SNF complexes regulate a broad range of genetic programs, and several target pathways (shown in blue boxes) have been implicated in their tumor suppressor activity. Selected proteins the expression or activity of which is regulated by SWI/SNF complexes are shown in ovals. CCND1, cyclin D1; ER, oestrogen receptor; GR, glucocorticoid receptor; IFNB, interferon- $\beta$ ; ROCK1, RHO-associated protein kinase 1. Adapted from Wilson et Roberts, 2011.

In addition, the tumor suppressor activity of SWI/SNF complex involves different signaling pathways (Figure 6). During differentiation, SWI/SNF complexes cooperate with tissue-specific transcription factors to coordinate the balance between self-renewal and differentiation and the disruption of this balance has been implicated in the transformation. It has been shown that an imbalance in the antagonistic relationship between EZH2 and SNF5 has essential roles in the genesis of tumors. In fact, PcG proteins contribute to gene silencing during lineage-specific development and increased polycomb-repressive complex 2 (PRC2) binding (Wilson et al., 2010) and skewed SMARCB1-deficient BAF complex binding at super-enhancers (Nakayama et al., 2017; Wang et al., 2017) have been suggested mechanisms of tumorigenesis due to SMARCB1 loss. Indeed, Roberts et al. showed that BAF47-deficient rhabdoid tumors displayed marked increases in the H3K27me3 repressive mark, a mark known to be placed only by PRC2 complexes and known to drive rhabdoid tumor formation (Wilson et al., 2010). Besides this, p16INK4A expression is downregulated following SNF5 inactivation and has been also implicated in tumorigenesis (Isakoff et al., 2005; Oruetxebarria et al., 2004). In fact, p16INK4A is a cyclin-dependent kinase inhibitor that regulates the RB tumor suppressor

pathway and thus has essential roles in cell cycle regulation and differentiation. Cyclin D1 may also have essential roles during tumorigenesis *in vivo*, since *Snf5*-heterozygous mice fail to develop spontaneous tumors on a background of cyclin D1-deficiency (Tsikitis et al., 2005). SWI/SNF complexes also interact with MYC whose expression is also high in SNF5-deficient rhabdoid tumors and may have an important role in tumorigenesis (Cheng et al., 1999; Gadd et al., 2010; Nagl et al., 2006). Finally, nuclear hormone receptors (NHRs) and alteration of Hedgehog pathway signaling has been also implicated in oncogenic transformation (Jagani et al., 2010; Lemon et al., 2001; Lessard et al., 2007).

## **The Cell of Origin**

As we previously described, RT have a heterogeneous clinical presentation and can occur in many different locations throughout the body. Among ATRT, clinical presentation is also heterogeneous. For instance, ATRT occur at all levels of the central nervous system (i.e. supratentorial, infratentorial and spinal) and at different ages (i.e. mostly infants and toddlers, but also children and rarely adults). Moreover, the tumor location and age at diagnosis are key decision-making factor in treatment and may therefore impact survival. Although ATRT are most usually an aggressive disease, some ATRT will present a more indolent behavior (Upadhyaya et al., 2021). These observations bring to light that despite their homogeneous genetic features, ATRTs are probably a heterogeneous group of diseases which may lead to more tailored treatments. Below, we will describe the cumulative studies suggesting that different cells of origin may contribute to ATRT heterogeneity and the contribution of the transgenic mouse to the better understanding of ATRT formation. Then, we will detail recent works based on phylogenetic analysis and single-cell transcriptomics. Finally, we will rationalize the relevance of tumor location in the brain to better understand the biology of the tumor.

## **Molecular Heterogeneity of ATRTs**

The first study came from Torchia et al in 2015. They showed that ATRT were a biologically heterogeneous disease comprising at least two molecular subtypes with distinct clinico-pathological associations: group 1 tumors were predominantly supratentorial and showed overexpression of genes involved in brain or neural development, axonal guidance and NOTCH signaling pathway; and group 2 tumors, were predominantly infratentorial ATRT showing enrichment of genes involved in mesenchymal differentiation and the bone morpho

genetic protein (BMP) signaling pathway. In addition, they identified *ASCL1* as a positive prognostic factor and related their expression to the supratentorial location. Next year, in 2016, Johann et al. demonstrated that ATRTs are composed of three epigenetic subgroups with distinct enhancer landscapes that have different clinical characteristics, and identified subgroup-specific regulatory networks that suggest potential therapeutic targets (Johann et al., 2016). They coined for the first time the term of ATRT-SHH, ATRT-TYR and ATRT-MYC for the three different subgroups. In this work, ATRT-SHH tumors located both infra- and supratentorially were characterized by overexpression of genes involved in SHH signaling (*MYCN* and *GLI2*) and notch signaling. ATRT-TYR tumors were more common in infratentorial regions of very young children and were characterized by overexpression of several melanosomal markers (*MITF*, *TYR*, *DCT*) and genes involved in ciliogenesis. ATRT-MYC tumors mostly occurred in the supratentorial compartment of older children and were characterized by the overexpression of the *MYC* oncogene, *HOTAIR* and many other *HOX* cluster genes. Notably, the type of inactivating *SMARCB1* mutation clearly differed between the ATRT subgroups: broad *SMARCB1* deletions for ATRT-TYR and focal *SMARCB1* aberrations and deletions for ATRT-SHH and ATRT-MYC, respectively. Furthermore, they found remarkable epigenetic differences between subgroups with ATRT-TYR subgroup and, to a lesser extent also the ATRT-SHH subgroup, showing a hypermethylated genome while ATRT-MYC subgroup were characterized by a hypomethylated genome. Ultimately, they identified distinct molecular ATRT subgroups with different preferred locations in the brain and distinct transcriptional and epigenetic landscape and suggested they may originate from different precursors cells.

In the same year, Torchia et al. published a new study where they showed that ATRTs comprised three epigenetic subtypes (1, 2A and 2B) that correlate with distinct tumor locations, patient age, lineage-enriched methylation and transcriptional signatures and differential cellular responses to a panel of signaling and epigenetic inhibitors (Torchia et al., 2016). In detail, group 1 arose predominantly in the supratentorial location and in the oldest children and showed overexpression of neurogenic genes (*ASCL1*, *FABP7*), genes involved in Notch, glutamate receptor and axonal guidance signaling; group 2A tumors arose infratentorial locations and in youngest children and showed overexpression of genes involved in visual cortex/hindbrain development, retinol pathway and tyrosine (*TYR*) metabolism; group 2B ATRTs included infra/supratentorial locations and all spinal tumors and the majority of patients older than 3 years of age and showed overexpression of *MYC* and *HOXB/C* clusters, interferon signaling, cell adhesion, and cytoskeletal genes. Moreover, they found that BMP signaling and

mesenchymal differentiation genes were most differentially in group 2A/B and that many group 2A enriched genes had functions in pluripotency and EMT. Most importantly, the distinct methylation and enrichment gene expression profile and clinical presentation in the three subgroups led them to hypothesize radial glial neural progenitors and mid/hindbrain neural progenitors as potential cell of origin for group 1 and 2 respectively (Torchia et al., 2016). Finally, they also showed that the type of genetic event leading to SMARCB1 loss also differed between ATRT subgroups, indicating SMARCB1 genotype: phenotype correlations in ATRTs. In summary, they state, in line with Johann et al., that diverse mechanisms of SMARCB1 loss in different cellular contexts together with additional epigenetic and genetic events underlie the clinical heterogeneity of human ATRTs, and go a step further pointing to possible cell types as cells of origin.

More recently, other studies (Chun et al., 2019) including rhabdoid tumors from multiple locations (ATRTs and ECRT) have demonstrated similarities between the ATRT-MYC and ECRT, including global DNA hypomethylation and overexpression of HOX genes and genes involved in mesenchymal development, distinguishing them from other ATRT subgroups. They also tried to correlate the transcriptome profiles of the different molecular subgroups to various progenitor cell types and showed that ATRT-MYC were correlated with CD56+ mesodermal progenitor cells and ATRT-SHH were correlated with cranial neural crest cells, neuronal progenitors and brain tissues. More interestingly, they also found that MYC-ATRT and ECRT exhibit cytotoxic T cell infiltration and expression of immune checkpoint regulators, consistent with a potential role for immunotherapy in a subset of patients with rhabdoid tumors. These findings have been more extensively corroborated and addressed recently by our group (Leruste et al., 2019).

Considering that all the aforementioned works included different numbers of ATRT, various subgrouping techniques and naming, Ho et al. performed a meta-analysis of unpublished and previously published DNA methylation, gene expression and clinicopathological data to generate a consensus on nomenclature and clinicopathological significance of ATRT subgroups (Ho et al., 2020). They identified 3 main molecular subgroups that closely matched with the previously published and named them as ATRT-SHH, ATRT-TYR, and ATRT-MYC, based on the nomenclature proposed by Johann et al. The figure 8 shows the consensus overview of ATRTs subgroups and the correspondence between the different subgroups already described. Interestingly, they found that the ATRT-SHH subgroup segregated further into 2 subtypes both expressing marker genes from the NOTCH and SHH pathways and associated with a predominant location (supratentorial, ATRT-SHH-1 or infratentorial, ATRT-

SHH-2). The existence of subclusters within the SHH and also the MYC subgroup had been previously suggested, but not strongly supported (Chun et al., 2019; Johann et al., 2016). In addition, they pointed out the resemblance of ATRT-TYR with cribriform neuroectodermal tumors (CRINETs) and suggested that they may represent 2 histological variants with a common neuroectodermal cell of origin. The effort to establish a correct classification is essential to investigate subgroup-specific features of ATRT and to design clinical trials accordingly since this classification may impact the prognosis (Frühwald, 2020; Reddy et al., 2020; Upadhyaya et al., 2021). Furthermore, it allows performing subgroup-specific experiments using properly classified preclinical *in vitro* and *in vivo* models. In the next part, different works attempting to establish rhabdoid tumor mouse models and their contribution to a better understanding of the cell of origin will be described.

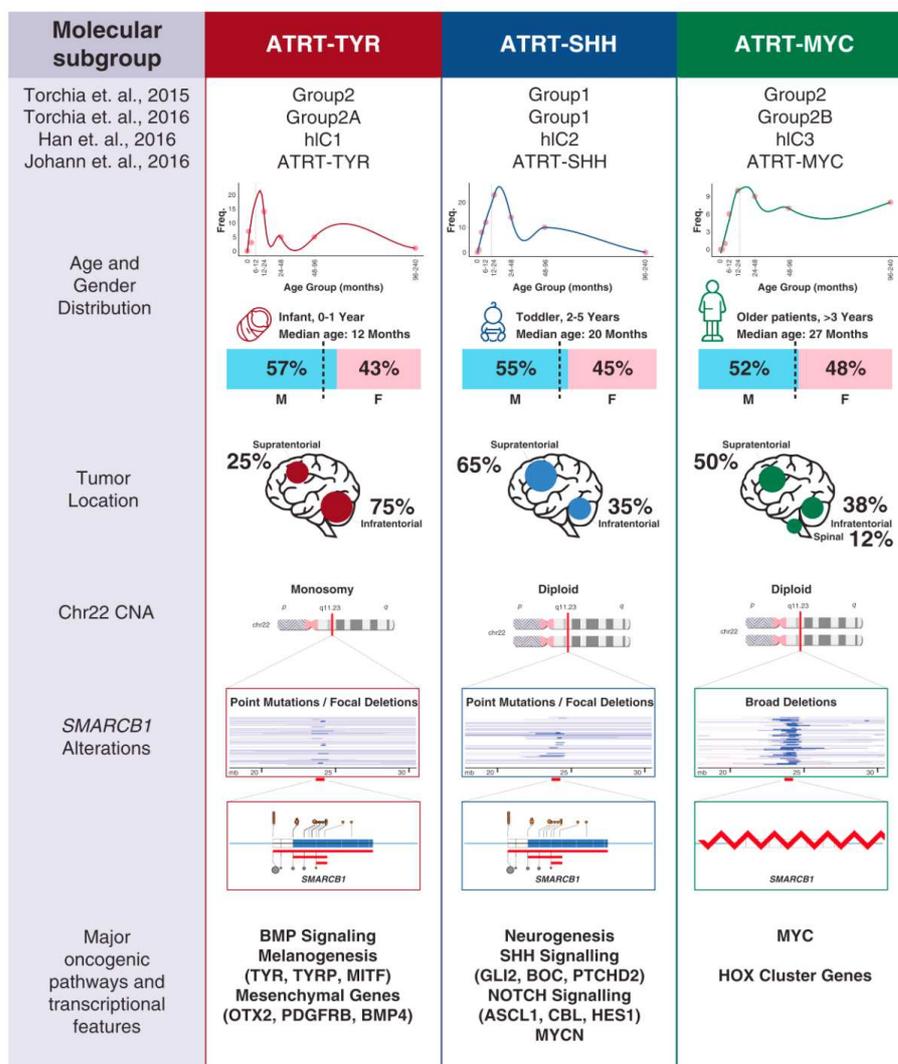


Figure 8. Consensus overview of ATRTs subgroups. *Adapted from Ho et al.*

## Mouse Models of ATRTs

Several groups have attempted to recapitulate rhabdoid tumors but until recently authors failed to obtain a reliable model for ATRTs. In fact, prior studies showed that the homozygous knockout of *Smarcb1* led to early embryonic lethality and that *Smarcb1* heterozygous mice were prone to develop tumors resembling human rhabdoid tumors. However, these tumors develop with long latency (Klochender-Yeivin et al., 2000) and/or almost exclusively from soft parts (Guidi et al., 2001; Roberts and Galusha, 2000). In detail, Klochender-Yeivin et al. generated *Snf5* heterozygous mice and observed that 32% of them developed tumors located in over half of the cases at intra-cranial and para-vertebral sites. However, since these tumors occur when mice were adults and no molecular comparisons were done with human tumors, it was uncertain to which extent this provided a faithful model for rhabdoid tumor of young children. The immunoreactivity for PS100, NGF-R and GFAP in some tumors induced the authors to suggest the neural crest cell as cell of origin. On the other hand, other authors failed to obtain brain tumors (Guidi et al., 2001; Roberts and Galusha, 2000) but soft tissue tumors from structures derived from the branchial arches; of note, they also observed a high expression of *Ini1* in branchial arches that are composed of neural crest cells during normal development. In 2002, Roberts et al. performed conditional inactivation using a *Smarcb1<sup>inv</sup>;Mx-Cre* model and obtained highly aggressive and fully penetrant tumors in short delay (median: 11 weeks) but failed to obtain CNS tumors (Roberts et al., 2002). These tumors were demonstrated to be exclusively CD8 mature peripheral T-cell lymphomas (20/23) and rarely rhabdoid tumors (3/23).

The most faithful murine model for ATRTs was published in 2016 by our team (Han et al., 2016). In this study, by intercrossing *Rosa26-CreERT2* and *Smarcb1<sup>flox/flox</sup>* mice, we generated genetically modified mice that allow conditional inactivation of *Smarcb1* at different developmental stages. The injection of tamoxifen at various time points allowed us to sequentially inactivate *Smarcb1* from early embryonic stages to adulthood and explore the associated phenotypes. As a result, the injection before E6, at birth or at 2 months of age recapitulated previously described phenotypes including embryonic lethality, hepatic toxicity or development of T-cell lymphomas, respectively (Guidi et al., 2001; Klochender-Yeivin et al., 2000; Roberts et al., 2002). Notably, the inactivation of *Smarcb1* at E6-E7 resulted in almost constantly intra-cranial or spinal tumors with short delays (median: 3 months). The tumor penetrance decreased when tamoxifen was administered from E8 to E10 and was null from E11 to E18. Furthermore, the tumors obtained showed anatomical, morphological and gene expression profiles consistent with those of human ATRTs. In addition, intra- and inter-

species comparisons of tumors revealed that human and mouse rhabdoid tumors split into different entities that may underline the variety of rhabdoid tumors cells of origin. In particular, we pointed out cephalic neural crest-derived cells as cells of origin for mE/IC and neural stem cells and neural progenitors for mIC/hIC2, corresponding respectively to MYC-ATRT and SHH-ATRT (Figure 8 and Figure 9).

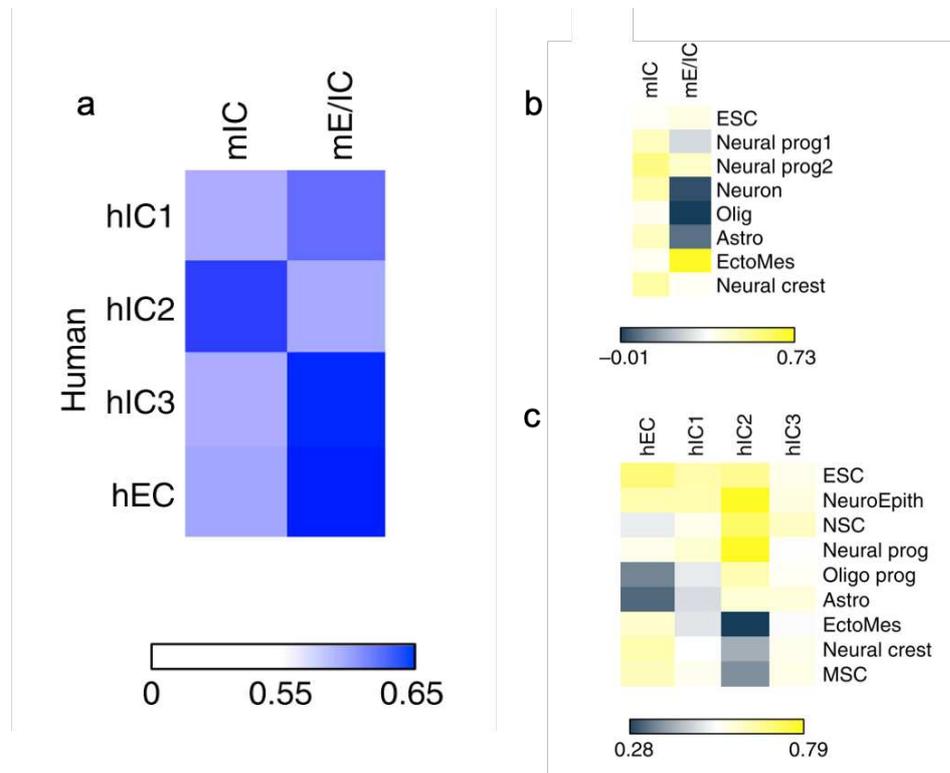


Figure 9. a. The Mouse CNS tumors reflect the diversity of human RTs. mIC, murine intra-cranial tumors corresponds to SHH-ATRT ; mE/IC, murine extra-cranial/Intracranial corresponds to MYC-ATRT. hIC, human intracranial; hEC, human extra-cranial; hIC1 corresponds to TYR-ATRT, hIC2 corresponds to SHH-ATRT and hIC3 and hEC corresponds to MYC-ATRT. hEC group. The figure shows the AGDEX score performed on each mouse and human subgroups. b and c. Transcriptome profiles suggest various cells of origin for human and mice SMARCB1-deficient tumors. The figure b shows the pearson correlation between transcriptome profiles of mIC and mE/IC from one part, and various mouse embryonal tissues from the other part. ESC: embryonic stem cell. Neural Prog 1, neural progenitor; neural Prog 2, neural progenitor; neuron; olig, oligodendrocytes derived from neural stem cells; astro, astrocytes derived from neural stem cells; EctoMes, ectomesodermal tissue from the palate; neural crest. c. The figure c shows the Pearson correlation between transcriptome profiling of hICs and hEC from one part, and various embryonal tissues from the other part. NeuroEpith, NeuroEpithelium; NSC, neural stem cell obtained from foetal cortex; neural Prog; glial Prog, glial progenitors; oligo Prog, adult oligodendrocytes progenitors; astro, astrocytes; MSC, mesenchymal stem cells; neural crest; EctoMes, ecto-mesenchymal stem cells. Adapted from Han et al. 2016.

One year later, another group developed a mouse model in which *Smarb1* is inactivated, precisely, in the Neural Crest Cells (NCC) and Schwann Cell (SC) lineage (Vitte et al., 2017).

Since *Smadcb1* germline mutations predisposes to rhabdoid tumors and familial schwannomatosis, these authors sought to investigate the mechanisms by which the same gene lead to early aggressive rhabdoid tumors versus late-onset benign peripheral nervous system tumors (Vitte et al., 2017). They showed that *Smadcb1* loss in early neural crest lead to rhabdoid tumors while *Smadcb1* and *Nf2* loss at later developmental stage in the Schwann cell lineage promote schwannoma development. More in detail, they developed the P0-CreC; *Smadcb1*<sup>F/F</sup> mouse model which specifically targets NCC and schwannoma precursor cells. Interestingly, they obtained that one third of P0-CreC;*Smadcb1*<sup>F/F</sup> mice developed rhabdoid tumors; even more interesting, these rhabdoid tumors were exclusively located in the cranial nerves (olfactory, trigeminal, oculomotor, optic and vestibulocochlear), meninges and the spinal nerve roots. To note, they only obtained, in line with our work (Han et al., 2016), RT when *Smadcb1* was inactivated at early developmental stages.

Taken together these works highlight the fact that not all cells are able to survive to *Smadcb1* inactivation and give subsequently rise to ATRTs. Or in other words, some cells would be exquisitely sensitive to *Smadcb1* deletion but only in a restricted developmental window.

### **Phylogenetic analyses and single-cell approaches**

In addition to molecular analysis and mouse models other approaches are used to investigate the origin of rhabdoid tumors. For instance, Jessa et al. mapped single-cell transcriptomes from ATRTs to a dataset containing developing mouse pons and forebrain (E12.5–P6) cell types and prenatal human brain-stem (17–19 postconception weeks) cell types. The aim of this approach is to find a match between tumor cell types and developing brain cell types consistent with a model of stalled differentiation where tumor cells retain features of the lineage of origin. Despite they did not find specific match of malignant cells to any specific cell type of the dataset they suggest that SHH-ATRT may arise from an earlier progenitor (prior to E12.5) and MYC/TYR-ATRT likely from cells outside the neuroectoderm (Jessa et al., 2019).

Recently, combined phylogenetic analyses and single-cell mRNA studies in patient-derived organoids have been used to study the cell of origin of extra-cranial rhabdoid tumors (ECRT) (Custers et al., 2021). This phylogenetic approach using whole-genome sequencing (WGS) consist on the comparison of somatic mutations shared between cancer and surrounding normal tissues. In addition, they performed single-cell mRNA readouts of ECRT differentiation by reverting the SMARCB1 loss in patient-derived organoids. As a result, they found that ECRT are phylogenetically related to neural crest-derived tissues likely blocked en route to

differentiating into mesenchyme. Interestingly, they found in one of the cases that some normal Schwann cells harbored partly the somatic genome of ECRT, including homozygous loss of SMARCB1. This finding suggests that SMARCB1 loss on its own may not suffice to generate tumors, or to impair normal cellular differentiation. Consecutively, factors others than embryological timing of disruption of SMARCB1 would influence tumor formation in humans. This approach could be used to study ATRT with some limitations since it requires samples of the tumor, blood but also healthy adjacent tissues that in the case of ATRT could be obtained only post-mortem.

## Location, location, location

Major advances have been done since Gilbertson et al. published the milestone paper supporting the notion that histologically similar tumors from different brain regions are molecularly distinct because they arise from distinct populations of site-restricted progenitor cells (Gilbertson and Gutmann, 2007). Today, growing evidence supports that brain tumors arise from deregulated developmental processes (Azzarelli et al., 2018; Baker et al., 2016; Bruschi et al., 2021; Curry and Glasgow, 2021; Jessa et al., 2019). Therefore, understanding the processes that regulate normal development is extremely important to understand brain tumor pathogenesis. More interesting, childhood brain tumors and their driver mutations show a specific spatiotemporal distribution (Marino and Gilbertson, 2021) (Figure 10).

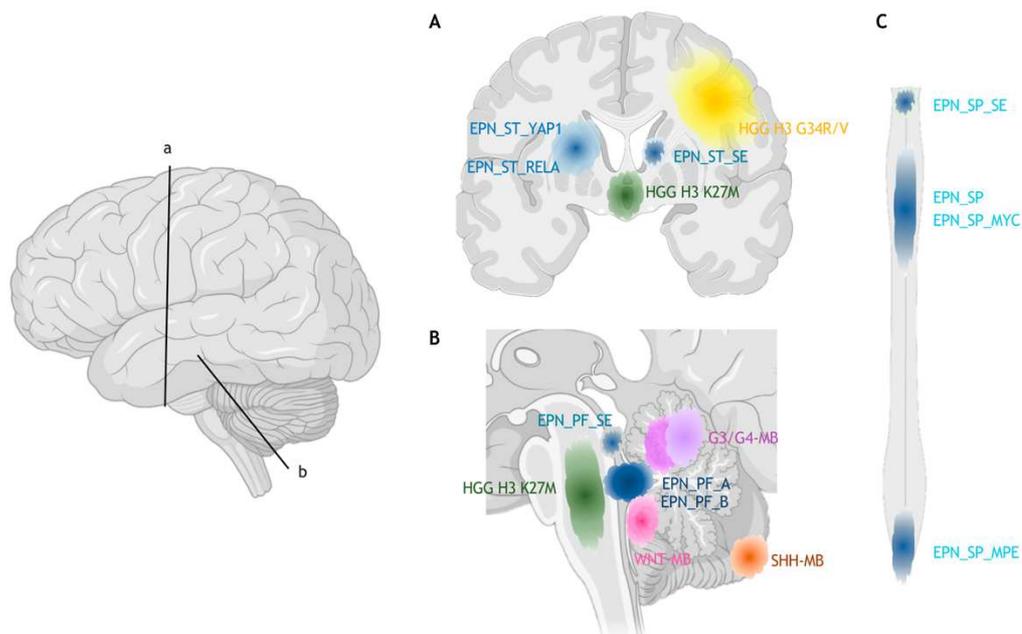


Figure 10. Genetically defined paediatric brain tumors occur at specific CNS locations. In the forebrain (A), YAP1 fusion (EPN\_ST\_YAP1) and C11orf95 fusion, (EPN\_ST\_RELA), ependymomas and subependymoma (EPN\_ST\_SE) arise within the radial glia lineage in the lateral ventricles, whereas histone mutant gliomas are found at the midline (HGG H3 K27M) and in the hemispheres (HGG H3 G34R/V). In the hindbrain (B), ependymoma A (EPN\_PF\_A), B (EPN\_PF\_B) and subependymoma (EPN\_PF\_SE) as well as certain forms of medulloblastoma (WNT-MB, SHH-MB and G3/G4-MB) arise from various progenitor populations in the cerebellum or in the wall of the 4th ventricle. In contrast, histone mutant gliomas arise in the pons (HGG H3 K27M) and WNT medulloblastoma arise in the dorsal brainstem from mossy fibre precursors. The spinal cord (C) is the site of subependymoma (EPN\_SP\_SE), as well as spinal (EPN\_SP), MYC amplified (EPN\_SP\_MYC) and myxopapillary ependymomas (EPN\_SP\_MPE). Adapted from Marino and Gilbertson, 2021.

In the case of ATRT, the question of anatomical localization has been scarcely addressed and mostly limited to supra- or infratentorial description (Dho et al., 2015; Ho et al., 2020; Johann et al., 2016; Nowak et al., 2018; Torchia et al., 2016). This may have led some authors to state that "... ATRT are a rare exception regarding spatiotemporal patterns..." (Jessa et al., 2019). However, if we consider that ATRTs could come from different cells of origin it would make perfect sense they do not present a single spatiotemporal pattern. In this sense, consider ATRT according to their specific location could bring new insight into their origin. To this effect, in our work we sought to integrate this information layer to the multi-omics analysis.

# **Project Rationale and Objectives**

Considering that:

1. ATRT are embryonal brain tumors; they are clinically heterogeneous and arise in very young children; they have extremely poor prognosis and need urgently new therapies.
2. Pediatric brain tumors arise from deregulated developmental processes; the roots are related with specific cell lineages within the developing brain; for other pediatric brain tumors, such as medulloblastomas, it has been demonstrated that specific brain locations are related to different cells of origin present during brain development.
3. Three different molecular subgroups of ATRTs suggesting different cells of origin have been described. These cells could correspond to different spatiotemporal pattern and explain ATRT heterogeneity.
4. Few studies have addressed the issue of the anatomical location of ATRT; or have merely described them as supra or infratentorial tumors. These studies do not provide precise information about the putative tissue of origin.
5. We assumed that precise anatomical description could improve the understanding of ATRT development.

The specific aims were:

1. To provide a detailed description of different anatomical ATRT locations.
2. To perform an integrative analysis of anatomical location, epigenetic and transcriptomic dataset.
3. To investigate the putative cells of origin taking into account this new layer of information.
4. To guide the development of murine models that recapitulates the occurrence of human tumors.

# Results

Authors' contributions;

Radiological analysis was carried out by María-Jesús Lobón-Iglesias and Volodia Dangouloff-Ross. Animal experiments were carried out by Zhiyan Han and María-Jesús Lobón-Iglesias. Bioinformatics analysis was carried out by María-Jesús Lobón-Iglesias and Mamy Andrianteranagna. These results are exclusively presented in the María-Jesús Lobón-Iglesias's doctoral thesis.

# **Anatomical-molecular integration study defines a new ATRT subgrouping and reveal potential lineages of origin**

María-Jesús Lobón-Iglesias<sup>1\*</sup>, Mamy Andrianteranagna<sup>1,10,11\*</sup>, ZhiYan Han<sup>1\*</sup>,  
Julien Masliah-Planchon<sup>4</sup>, Arnault Tauziède-Espariat<sup>5</sup>, Christelle Dufour<sup>6</sup>, Celine Chauvin<sup>1</sup>,  
Rachida Bouarich<sup>1</sup>, Sandrina Turczynski<sup>1</sup>, Didier Surdez<sup>3</sup>, Olivier Delattre<sup>3</sup>, Kevin Beccaria<sup>7</sup>,  
Olivier Ayrault<sup>8</sup>, Joshua Waterfall<sup>3</sup>, Volodia Dangouloff-Ross<sup>9</sup>, Nicolas Servant<sup>2,10,11</sup>,  
Franck Bourdeaut<sup>1,3</sup>

<sup>1</sup> Institut Curie, Translational Research In Pediatric Oncology Department, Paris, France

<sup>2</sup> Institut Curie, PSL Research University, F-75005 Paris, France

<sup>3</sup> SIREDO: Care, Innovation and Research for Children, Adolescents and Young Adults with Cancer, Institut Curie, Paris, France

<sup>4</sup> PSL Research University, Institut Curie Hospital, Laboratory of Somatic Genetics, Paris, France

<sup>5</sup> Sainte-Anne Hospital, Department of Neuropathology, Paris, France

<sup>6</sup> Department of Pediatric and Adolescent Oncology, Gustave Roussy, Rue Edouard Vaillant, 94805 Villejuif, France

<sup>7</sup> AP-HP, Necker Hospital, Department of Neurosurgery, Paris, France

<sup>8</sup> PSL, CNRS UMR 3347, INSERM U1021, Institut Curie, Orsay, France

<sup>9</sup> AP-HP, Necker Hospital, Department of Radiology, Paris, France

<sup>10</sup> INSERM, U900, F-75005 Paris, France

<sup>11</sup> IMINES ParisTech, PSL Research University, CBIO-Centre for Computational Biology, F-75006 Paris, France

The results of my PhD work are presented below in a manuscript format which is currently in preparation.

## Introduction

Atypical teratoid rhabdoid tumors (ATRT) are rare and aggressive malignancies of the central nervous system (CNS) affecting infants and young children and characterized by a biallelic inactivation of *SMARCB1* tumor suppressor gene in an otherwise very simple genome (Lee et al., 2012). Based on methylation and expression profiling, recent studies have pointed out the molecular diversity of these tumors (Han et al., 2016; Johann et al., 2016; Torchia et al., 2016), that are now divided in at least three subgroups, i.e. the so-called MYC-, TYR-, and SHH-ATRTs (Ho et al., 2020). This molecular diversity is suggestive of various cells of origin for each sub-type, though none of these putative cells of origin is undoubtedly identified at 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).

In the present manuscript we aimed to provide a detailed description of different anatomical ATRT locations in order to obtain precise information about their putative tissue of origin. Then, we have performed integrative analysis of anatomical location (n=51), transcriptomic (n=49) and epigenetic data (n=54). This analysis allowed us to describe and characterize 4 distinct anatomico-molecular ATRT subgroups: cranial nerves, cerebral cortex and spinal cord tumors, CNCS-ATRT; basal ganglia and intraventricular tumors, BG/IV-ATRT; cerebellar anterior lobe tumors, CAL-ATRT; and middle cerebellar peduncle and inferior cerebellar vermis tumors, MCP/ICV-ATRT. Next, we sought to investigate the relevance of our mouse models for human disease. We found that the Rosa26: MYC mouse model perfectly recapitulates molecular features and clinical phenotypes of a subset of human ATRT (CNCS-ATRT) and supports the hypothesis of an origin in neural crest cells. The Rosa26: SHH tumors are phenotypically similar to BG/IV-ATRT and characterized by the expression of genes markers from the ganglionic eminence. Finally, the CAL-ATRT are characterized by the expression of midbrain-hindbrain boundary (MHB) genes. Although we don't have a faithful mouse model for this subgroup, the integrative analysis allowed us to identify the CAL-ATRT subgroup and lead us to perform single cell analysis of three human samples coming exclusively from this location. Consecutively, we demonstrated at the single-cell level, by the first time so far, the involvement of neuronal progenitor cells, neuronal program inhibition and Notch pathway in the malignant transformation of CAL-ATRT.

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

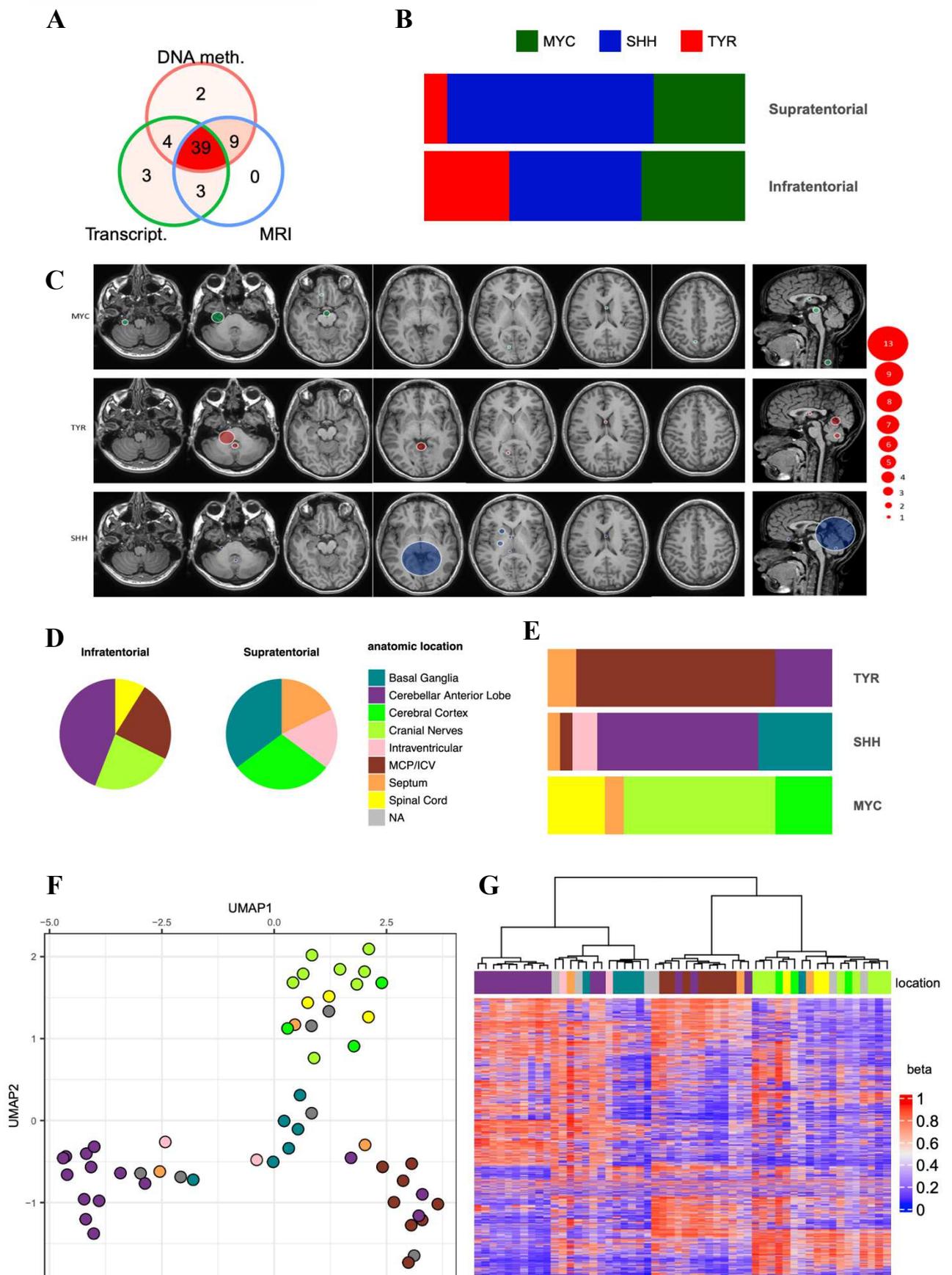
To investigate the site of origin of ATRTs, we first reviewed a series of 51 human brain ATRTs aiming to categorize them following the classically used locations, i.e; infratentorial, or supratentorial (Figure 1A, B, C).

We next assigned a molecular subgroup for those having DNA methylation data using the online DKFZ brain tumor classifier tool (<https://www.moleculareuropathology.org; v11b4>) (Capper et al., 2018). This classification based on DNA methylation data uses random forest algorithm to assign a sample among the existing 64 subtypes of tumors 2800 references samples (Capper et al., 2018). Raw data of DNA methylation were uploaded in the web application and the predicted ATRT subtypes with the prediction score are shown Figure S1B. We found that, while the infra-tentorial location is not *per se* suggestive of any tumor group, the supra-tentorial location can suggest both MYC or SHH subtypes (Figure 1B). Considering that this infra/supratentorial classification was not helpful enough determining an origin for ATRT subtypes, we then endeavored to categorize the tumors from their presumptive epicenter, aiming to give a more precise anatomic origin. This led us to define 8 anatomical categories of ATRTs: i) cranial nerves (CN), 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 (SC) tumors (Figure S1A). Overall, we end up with 4 locations belonging to infratentorial region and 4 locations belonging to supratentorial region (Figure 1D).

Combining these detailed anatomic locations with the DNA methylation subgroups we found that i) almost all TYR-ATRTs emerged from the middle cerebellar peduncle and the inferior cerebellar vermis; ii) SHH subgroup was composed by tumors located in the CAL, the BG and the intraventricular region; and iii) MYC subgroup was mainly composed by tumors located in the cranial nerves or the cerebral cortex and by spinal cord tumors (Figure 1E).

Altogether, these findings spotlighted a new correlation between anatomical location and molecular subgroups suggesting different lineages of origin. To further investigate this, we performed unsupervised analysis based on methylation array profiling, using UMAP and hierarchical clustering approaches. Strikingly, we found that ATRT samples tended to cluster according to their anatomical location (Figure 1F, 1G, S1B).

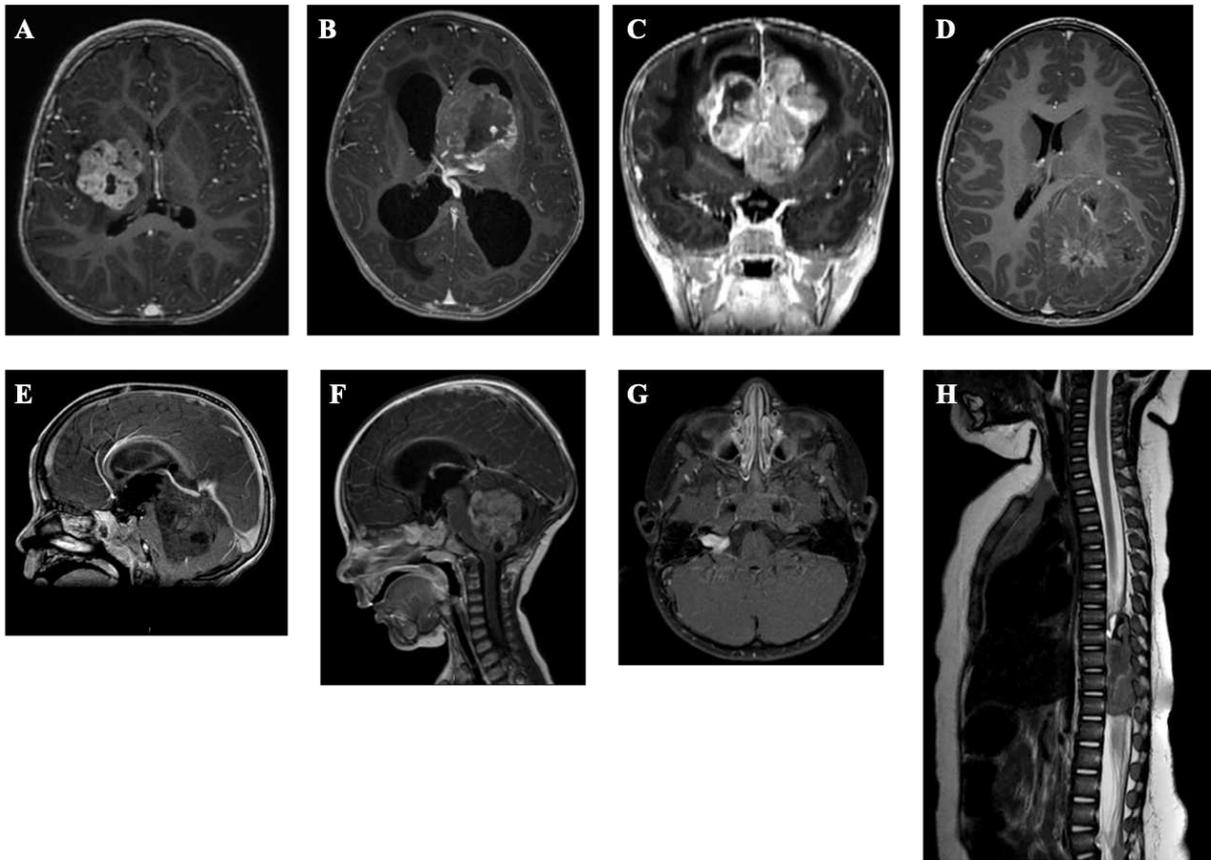
**Figure 1. Radiological description of ATRTs' epicenter suggests clearly distinct origins for each molecular subtype**



**Figure 1. Radiological description of ATRTs' epicenter suggests clearly distinct origins for each molecular subtype.**

- A. Venn diagram recapitulating the number of samples in each dataset: DNA methylation (DNA meth., red), magnetic resonance imaging (MRI, blue) and transcriptomic dataset (Transcrip.,green). DNA methylation data are obtained using the Illumina Human Infinium EPIC array and the transcriptomic data are from RNA sequencing technology.
- B. Bar plot showing the distribution of the different molecular groups assigned according to the DNA methylation profile (DKFZ brain tumor classifier v11b4) at the supra- and infratentorial level.
- C. MRI images showing the most frequent tumor locations according to molecular subgroups. The round size indicates the number of 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).
- D. Pie charts showing the distribution of the different ATRT anatomical locations at infra and supratentorial level. NA: unknown anatomical location.
- E. Bar plot showing the fraction of anatomical ATRT locations according to the molecular subgroup based on DNA methylation data (DKFZ brain tumor classifier v11b4).
- F. UMAP analysis performed on DNA methylation data of tumor samples.
- G. Hierarchical clustering of ATRT samples based on DNA methylation data. Top annotation indicates ATRT anatomical location.

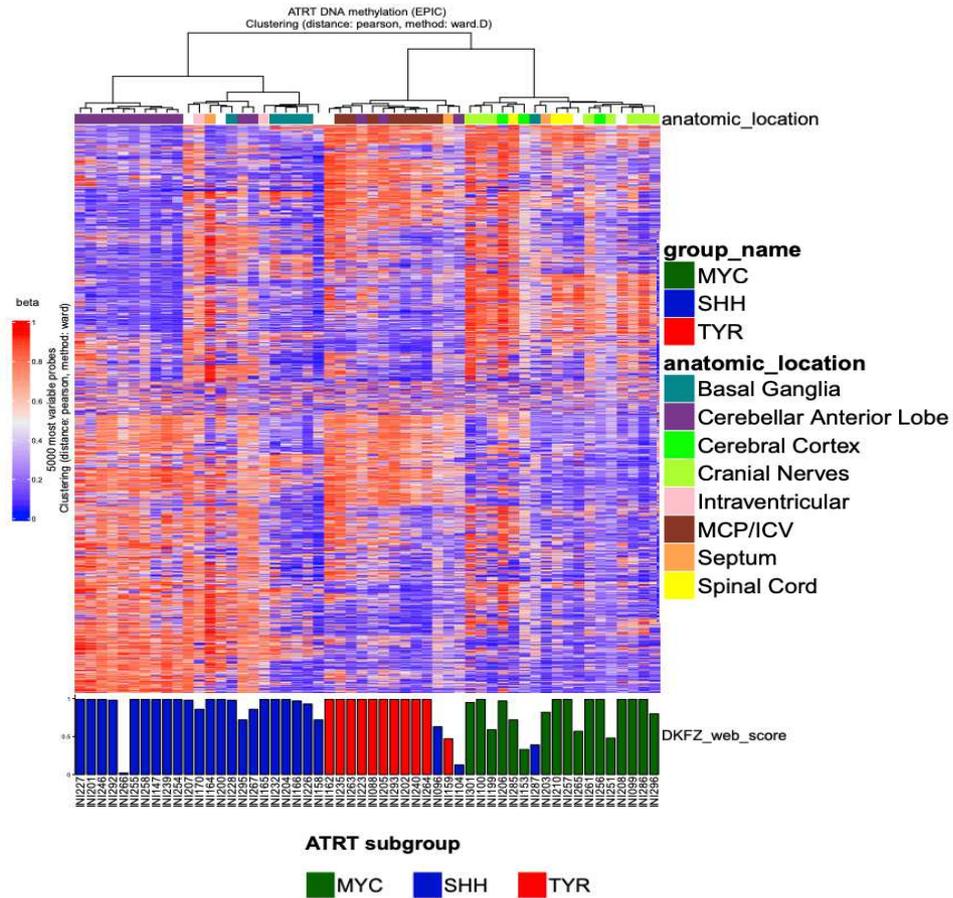
**Figure S1A. Anatomical locations categories of ATRT.**



MRI of tumors located at the supratentorial (A-D) and infratentorial level (E-H), respectively.

- A. Brain Axial-MRI showing an ATRT located in the basal ganglia region.
- B. Brain Axial-MRI showing an ATRT located in the intraventricular region.
- C. Brain Coronal-MRI showing an ATRT located in the septum ventricular.
- D. Brain Axial-MRI showing an ATRT located in the cerebral cortex.
- E. Brain Sagittal-MRI showing an ATRT located in the upper part of the cerebellum.
- F. Brain Sagittal-MRI showing an ATRT from the inferior cerebellar vermis.
- G. Brain Axial-MRI showing an ATRT located in the internal acoustic canal.
- H. Brain Sagittal-MRI showing an ATRT located in the spinal cord.

**Figure S1B. Unsupervised hierarchical clustering of ATRT samples based on DNA methylation data.**



The ATRT subgroups according the DKFZ classifier are indicated. White color in the anatomic\_location row indicates unknown anatomical location.

### **Anatomical-molecular integrative analysis defines 4 ATRT subgroups**

Next, we sought to explore the relevance of our anatomical clustering at the transcriptional level. At this aim, we performed an unsupervised analysis on the RNA-seq data of 49 samples based on the 5000 most variable genes. At first glance, the unsupervised hierarchical clustering showed as expected 3 molecular subgroups that corresponded to TYR, SHH and MYC (Figure S2A); however, the consensus clustering (set from  $k=2$  to  $k=7$ ) showed that the most stable cluster numbers were when  $k$  was set to 3 and 4 (Figure S2B, S2C). Interestingly, clustering with  $k=3$ , corresponding to the classically recognized number of subgroups, revealed some discrepancy between methylation and RNAseq profiling for a category of SHH ATRTs (Figure 2A). In contrary, clustering with  $k=4$  resolved this discrepancy between methylation and gene-expression based classifications by isolating a consistent group having a SHH methylation signature but a MYC expression profile (Figure 2B). Remarkably, all those tumors fell in the basal ganglia and intraventricular anatomical category.

To confirm this finding, we applied a kernel-based data integration approach (see material et methods) aiming to combine gene expression and DNA methylation datasets before unsupervised analysis. The combined kernel was visualized using UMAP representation and suggest the existence of 4 anatomico-molecular subgroups (Figure 2D).

In summary, the 4 subgroups showed the following correlation of anatomico-molecular profile (methylation/transcriptomics): i) Cerebellar anterior lobe tumors with a SHH/SHH profile (CAL-ATRT); ii) Cranial nerves, cerebral cortex and spinal tumors with MYC/MYC profile (CNCS-ATRT); iii) Middle cerebellar peduncle and inferior cerebellar vermis tumors with TYR/TYR profile (MCP/ICV-ATRT); and finally, iv) Basal Ganglia and Intraventricular tumors with SHH/MYC profile (BG/IV-ATRT).

To better understand the specificities of these subgroups, we performed one versus all others differential expression analysis (DEA). Among the top 100 differentially overexpressed genes for each subgroup (Figure 2C), previously identified ATRT subgroup signature genes as well as new signature genes are identified (detailed in the following sections).

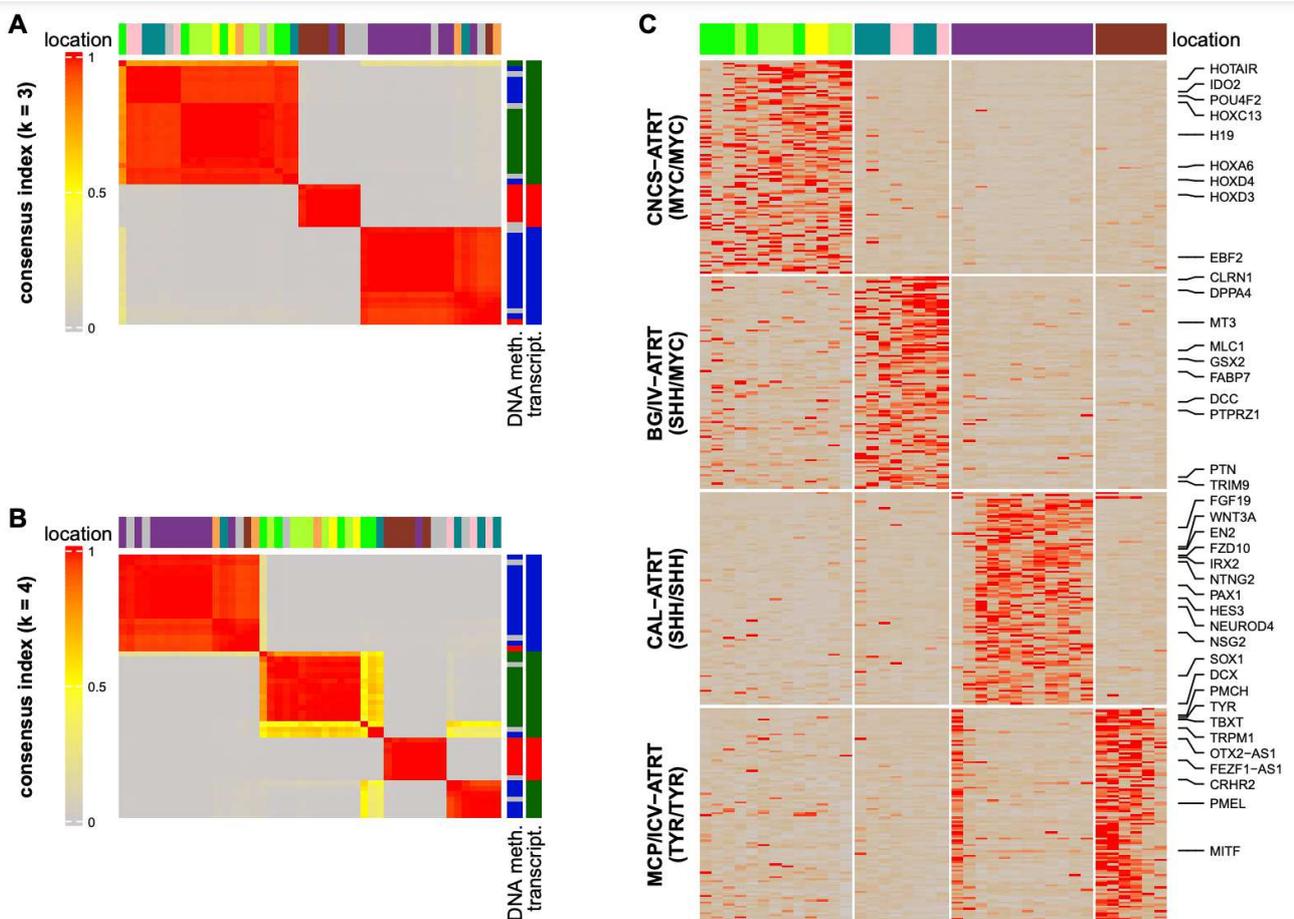
In order to confirm the consistency of this new subgrouping as well as to identify subgroup signature genes, we performed sparse Partial Least Square Determinant Analysis (sPLS-DA) (Figure 2E). In addition, this analysis allowed us to select subgroup signature genes. Our analysis showed that 3 components with respectively 90, 100 and 50 genes can separate our 4 subgroups with high performance (data not shown). Genes that are contributing to separate each subgroup, shown in Appendix Table 1, corroborate with the DEA results and supports

again the difference between the transcriptomics profiles of the 4 anatomico-molecular subgroups.

In order to reveal the biological function underpinning each subgroup gene set enrichment analysis (GSEA) were conducted using either the entire collections of MSigDB or specific gene set collections. The results of DEA, GSEA and the sPLS-DA analyses allowed us to identify subgroup specific genes and to characterize the transcriptomic profiles that we will develop in the followings sections.

Taken together, these results demonstrated that the anatomical location of ATRT allows to refine the difference between their molecular profiles. Furthermore, these results definitely established that SHH ATRT can be divided in 2 distinct groups, defined by their anatomic location, their gene expression signatures and their methylation profiles.

**Figure 2. Anatomical-molecular integrative analysis defines 4 ATRT subgroups.**



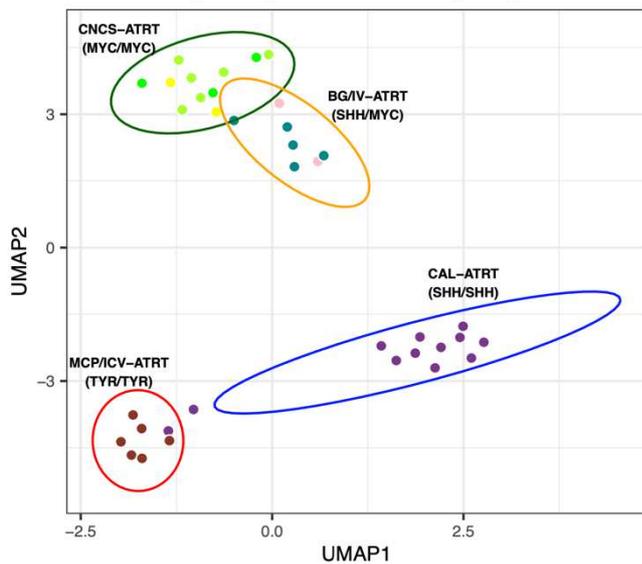
ATRT subgroup

MYC SHH TYR

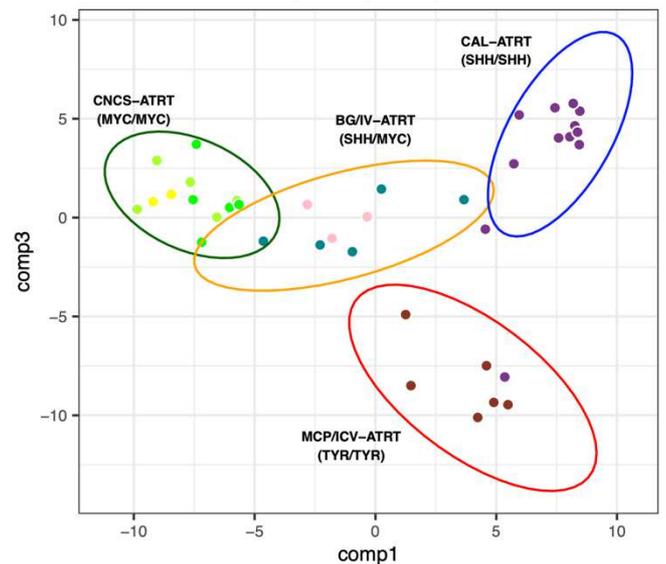
anatomic location

Basal Ganglia Cerebral Cortex Intraventricular Septum  
Cerebellar Anterior Lobe Cranial Nerves MCP/ICV Spinal Cord

**D** meta-kernel UMAP  
(transcriptomics + DNA methylation)



**E** sPLS-DA  
(transcriptomics)



**Figure 2. Anatomical-molecular integrative analysis defines 4 ATRT subgroups.**

- A. Consensus clustering analysis performed on transcriptomics data (RNA-seq) using the top 5000 most variable genes and a number of cluster  $k=3$ . Top annotation indicates anatomical location by different colors (grey color indicates unknown location). Right annotations indicate molecular subgroups according to transcriptomics and DNA methylation profile (DKFZ brain tumor classifier v11b4).
- B. Consensus clustering analysis performed on transcriptomics data (RNA-seq) using the top 5000 most variable genes and a number of cluster  $k=4$ . Top annotation indicates anatomical location by different colors (grey color indicates unknown location). Right annotations indicate molecular subgroups according to transcriptomics and DNA methylation profile (DKFZ brain tumor classifier v11b4).
- C. Heatmap of gene expression using the 100 most differentially expressed genes between different anatomical molecular subgroups. Top annotation indicates anatomical location by different colors (see legend). Genes of interest are indicated.
- D. UMAP performed on the meta-kernel matrix generated after combining the kernel of DNA methylation data and the kernel of the RNA-seq data (see method for details).
- E. Individual plot of sPLS-DA on grouped location. The figure shows the composante 1 in the x axis explaining 19% of the variability and in y axis the composante 3 accounting for 10% of variability. Colors indicate the sample anatomical location and the ellipse colors indicate the anatomical molecular subgroups.



**Figure S2.**

- A. Hierarchical clustering of ATRT sample based on RNA-seq data and using the 5000 most variable genes (based on IQR value). Top annotations indicate respectively from the top: the transcriptomics subgroup, the DNA methylation subgroup and the sample anatomic location.
- B. Average silhouette showing a stable clustering at 3 and 4 classes. Silhouette analysis was performed on RNA-seq data using the 5000 most variable genes.
- C. Consensus clustering of ATRT samples based on RNA-seq data showing results from 2 to 7 classes. Top annotation indicates anatomical location. NA: unknown anatomical location. Right annotation indicates molecular subgroups according respectively to transcriptomics and DNA methylation profile (DKFZ brain tumor classifier v11b4).

### **Integrated anatomico-molecular analyses suggest an extra-CNS lineage for MYC ATRT**

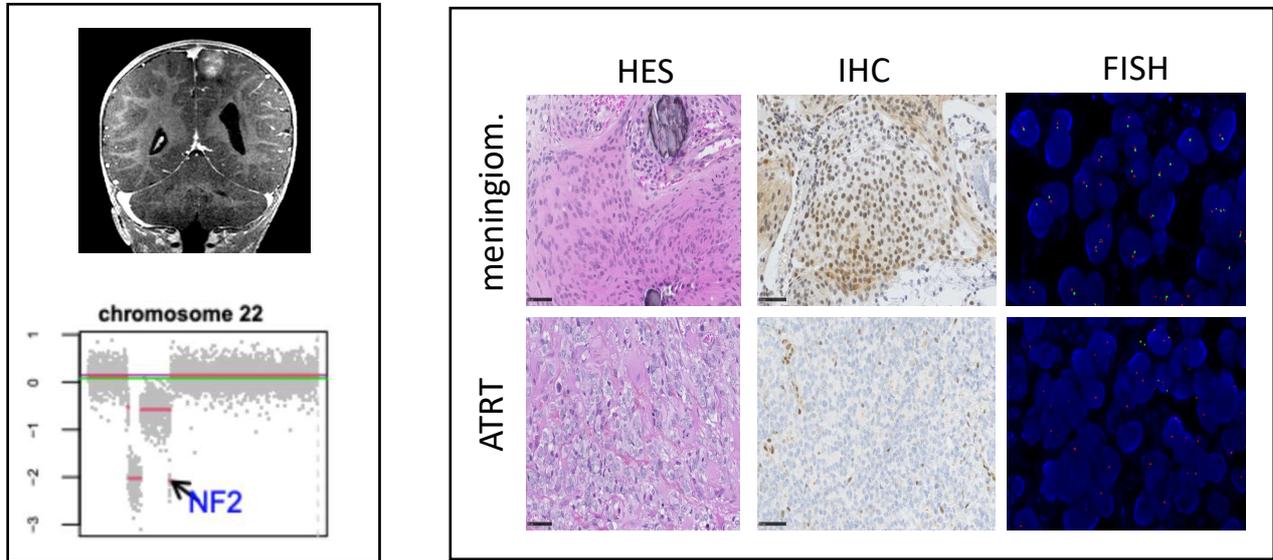
The cranial nerve and meningeal location of some ATRTs was somewhat reminiscent of the location of tumors observed in SMARCB1-related schwannomatosis and Neurofibromatosis type 2 (NF2), a genetic disorder which increases the risk of developing acoustic nerve schwannomas and meningiomas. In line with this similarity of tumor location, two CNCS-ATRT showed a biallelic inactivation of *NF2*. The first tumor was a typical cortical ATRT; a careful histopathological review of this tumor showed that it had two components, one with meningiomatosis and the other with rhabdoid features, the former retaining and the latter losing SMARCB1 expression (Figure 3A). This suggested that this MYC tumor emerged from a preexisting SMARCB1-retaining, meningeal lesion. The second case emerged from the acoustic nerve, with a long-lasting evolution in a teenager; it showed a homozygous deleterious mutation in *NF2* (c.340del/ p.(Thr114Hisfs\*9) within a broader 22q11.2 isodisomy; this case again suggested that a NF2/SMARCB1 inactivated acoustic schwannoma could be a path leading to CNCS-ATRT. Since those two cases were otherwise molecularly undistinguishable from other CNCS-ATRT, we sought to investigate if CNCS-ATRT could derive from a similar process in the meninges or cranial nerves. However, we found no other tumor with loss of NF2 expression, nor specific genetic alteration or splicing abnormality (data not shown). At the transcriptomic level, we wondered whether cranial nerve MYC tumors displayed other signatures relating them with NF2-related tumors such as meningiomas or acoustic nerves schwannomas. CNCS-ATRT showed a trend toward lower NF2 expression; among cephalic neural crest lineage markers, CNCS-ATRTs expressed genes of the “neural crest specification module” such as *SNAI2*, *TWIST* and *MYC* (Parker et al., 2018) (Figure 3B). As previously reported, they also expressed *HOXA* and *HOXB* cluster of genes, which are sequentially expressed in neural crest cells for hindbrain segmentation and generation of cranial nerves (Krumlauf, 2016; Méndez-Maldonado et al., 2020; Parker and Krumlauf, 2020; Parker et al., 2018).

Altogether, these data suggested that MYC tumors may occasionally derive from NF2-inactivated tumors such as meningiomas or acoustic nerve schwannomas, thus emerging from cephalic neural crest structures; the kinship with other MYC tumors suggested either similar biological origins for all cortical and cranial nerves ATRT, or similar effects of SMARCB1 loss on various lineages.

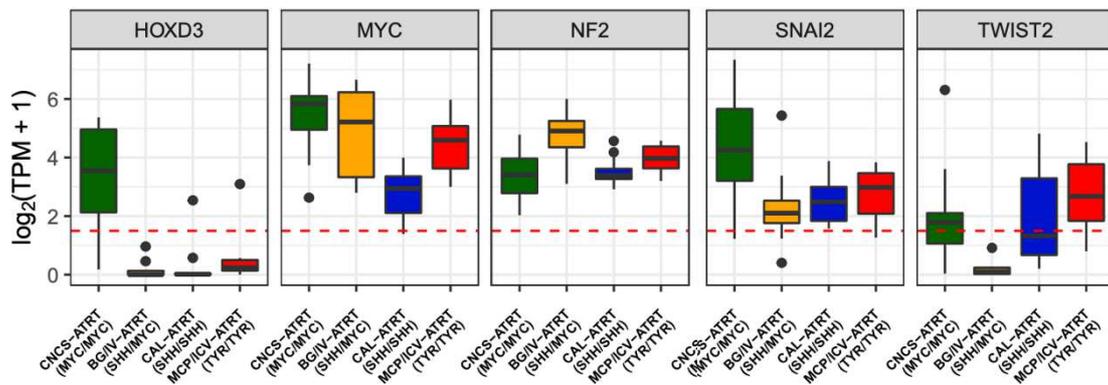
Finally, the GSEA comparing this subgroup to all others using the GO Biological Process gene sets showed, an enrichment of immune system related gene sets (Figure 3C).

**Figure 3. Integrated anatomico-molecular analyses suggest an extra-CNS lineage for MYC ATRT.**

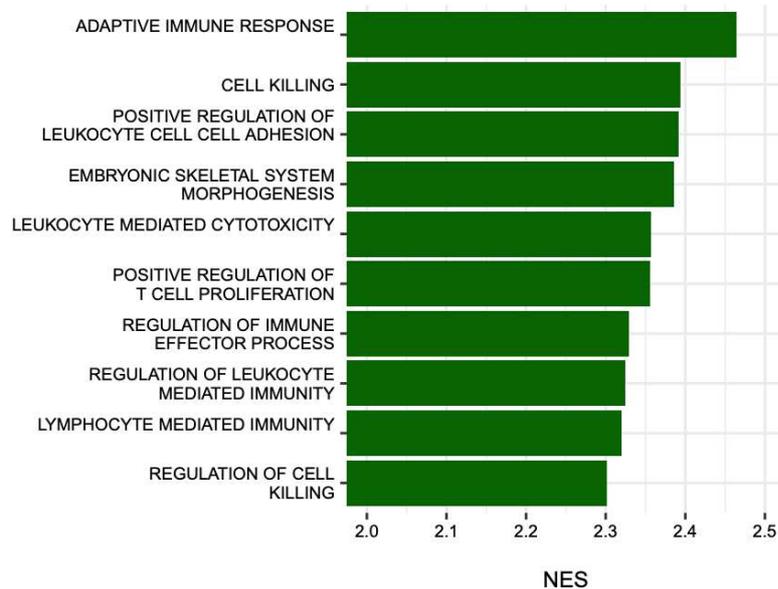
**A**



**B**



**C**



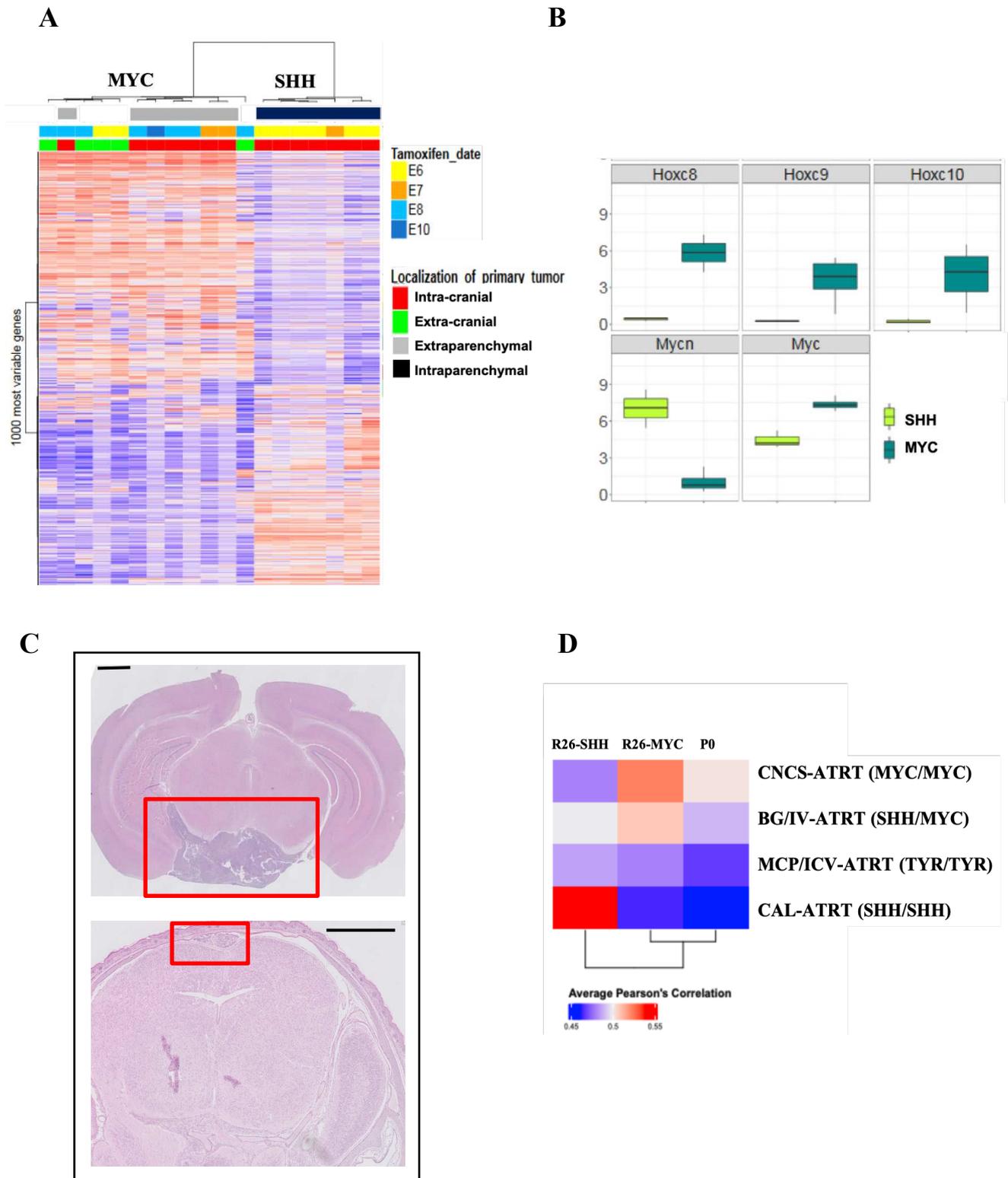
**Figure 3. Integrated anatomico-molecular analyses suggest an extra-CNS lineage for MYC ATRT.**

- A. CNCS-ATRT (MYC/MYC) case arising in continuity with a meningioma tumor. Left panel: coronal-MRI showing cortical ATRT (top) and copy number alteration plot zoomed on chromosome 22 showing NF2 deletion (bottom). Right panel: histopathological review showing meningiomatosis and ATRT components corresponding to the cortical tumor of the left panel. The IHC of the meningiomatosis component shows positive SMARCB1 staining while the IHC of the ATRT component shows negative SMARCB1 staining. FISH: Green probe= SMARCB1 locus (22q11.23); Orange probe = control locus, KREMEN1 locus (22q12.1-q12.2). Scale bar = 50  $\mu$ m.
- B. Boxplot of gene expression showing levels of HOXD3 (cranial neural crest gene), MYC, NF2 and SNAI and TWIST2 (genes related to neural crest) in each ATRT anatomico-molecular subgroup.
- C. GSEA: top 10 enriched GO Biological Process gene sets in CNCS-ATRT (MYC/MYC) compared to all others anatomical-molecular subgroups.

### **Genetically engineered mouse models suggest a neural crest cell origin for MYC ATRT**

We then investigated whether mouse models could bring further arguments in favor of this hypothesis. We have previously developed a faithful mouse model for ATRT by crossing Rosa26-cre<sup>ERT2</sup> and Smarcb1<sup>flox/flox</sup> strains (Han et al., 2016). We interestingly observed that tumors restrictedly emerged when *Smarcb1* inactivation occurred at early embryonic development steps, from E6 to E10. The transcriptome profiling of tumors revealed that the same mouse model generates tumors recapitulating both the MYC- and SHH-ATRT subgroups (Han et al., 2016). By generating an increased number of tumors profiled by RNAseq, we reinvestigated the link between developmental window, precise anatomic location and molecular subgroup (Figure 4A, B). Remarkably, Myc mouse tumors, characterized by similar expression profiling as MYC human ATRTs (Figure 4B), were obtained after Smarcb1 recombination from E6 to E10, thus within the widest developmental window; more interestingly, intracranial Myc clustered with the few extra cranial tumors we obtained (Figure 4A), in line with previous works in humans, suggesting that Myc ATRT and extra-cranial tumors share some similarities (Chun et al., 2019). Finally, a careful description of tumor locations clearly demonstrated an extraparenchymal origin of mouse Myc ATRT, which developed around the brain, from meningeal areas (Figure 4C). These results fitted well with our previous results suggesting a neuro-ectodermic origin for those tumors (Han et al., 2016), but the lack of specificity of Rosa26 promoter precluded from definitive conclusions. These results prompted us to reinvestigate the characteristics of tumors developed by the P0-CreC::Smarcb1<sup>flox/flox</sup> (P0:SB1) published by Vitte et al, consecutive to the *Smarcb1* loss in early cranial neural crest. As highlighted by the authors, the anatomic location of these tumors were typically in cranial nerves and meninges, locations strongly reminding the anatomic sites we observed in human (CNCS-ATRT; MYC/MYC). We therefore hypothesized that the P0-CreC::Smarcb1<sup>flox/flox</sup> (P0:SB1) model would the best recapitulate CNCS-ATRTs. Re-analysing the transcriptome profiling of these mouse tumors compared to our human ATRT cohort, we clearly found that P0:SB1 P0 correlated most strongly with the CNCS-ATRT subgroup and this correlation was less significant for the other subgroups (Figure 4D). Interestingly, pooled with RNAseq data from the R26:SB1 Myc tumors, we found a higher correlation of this model with CNCS-ATRT and to a lesser extent to the others subgroups (Figure 4D). Altogether, these results supported the hypothesis of a cranial neural-crest origin for CNCS-ATRT.

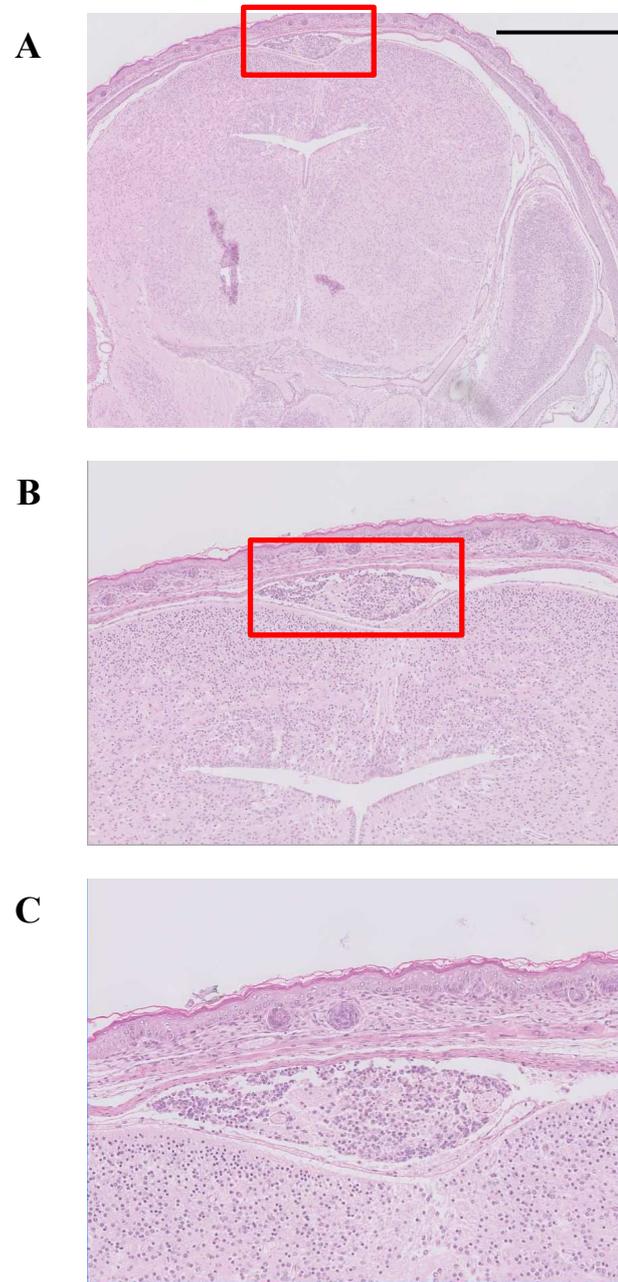
**Figure 4. Genetically engineered mouse models suggest a neural crest cell origin for MYC ATRT.**



**Figure 4. Genetically engineered mouse models suggest a neural crest cell origin for MYC ATRT.**

- A. Hierarchical clustering of tumors from R26 mouse model based on RNA-seq data (n=18). Top annotations indicate tamoxifen data corresponding to the time of inactivation of *Smarcb1*: tumors obtained after *Smarcb1* deletion at E6 (yellow), E7 (orange), E8 (light blue) and E10 (dark blue) and primary tumor location: extracranial tumors in green, intracranial tumors in red.
- B. Boxplots expression of genes typical for SHH (light green) and for MYC (dark green) ATRT subgroups.
- C. Typical examples of tumors emerging from extra-parenchymal meningeal tissues (axial section, top panel; coronal section, lower panel). Scale bar: 1mm
- D. Average Pearson correlation between human ATRT anatomico-molecular subgroups and mouse models (Rosa26: SHH, Rosa26: MYC and P0).

**Figure S4. Tumor emerging from extra-parenchymal/meningeal tissues.** High power magnification of the lower panel of Figure 4C (coronal section). Scale bar: 1mm.



### **BG\_IV-ATRT subgroup molecular pattern suggests neuronal origin and point out the ganglionic eminence region as tissue of origin**

We then focused on BG/IV-ATRT subgroup composed by tumors localized either in the basal ganglia or intraventricular region, both in the direct vicinity of subventricular zones. Remarkably, BG/IV-ATRT belong to the SHH subgroup according to methylation subgrouping but showed MYC gene expression profile (Figures 2B, S2A). When performing GSEA of this subgroup compared to all other subgroups using the GO-BP gene set collections, the top enriched gene sets are related to immune response (Figure 5A).

We next investigated whether, beside the immune signature, RNAseq data could also shed a light on the neuronal origin of these tumors, and in particular on neuron progenitors likely to be linked to the basal ganglia location. Noteworthy, when we looked at the top 100 most overexpressed genes in this subgroup compared to all others, we noted the overexpression of genes involved in forebrain development (FOXP1, EMX2, ARX and NRG1), neurogenesis, synapse or neuronal plasticity (ARC, BDNF), glial markers (FABP7, MLC1) and pluripotency genes (DPPA4) (Figure 2C, Figure 5B).

Even more strikingly, we noticed a clear over expression of a set of transcription factors such as ASCL1, SOX6, GSX2, DLX1, DLX2 and VAX1, which constitute a typical core set of genes expressed in the ganglionic eminences (Figure 5B).

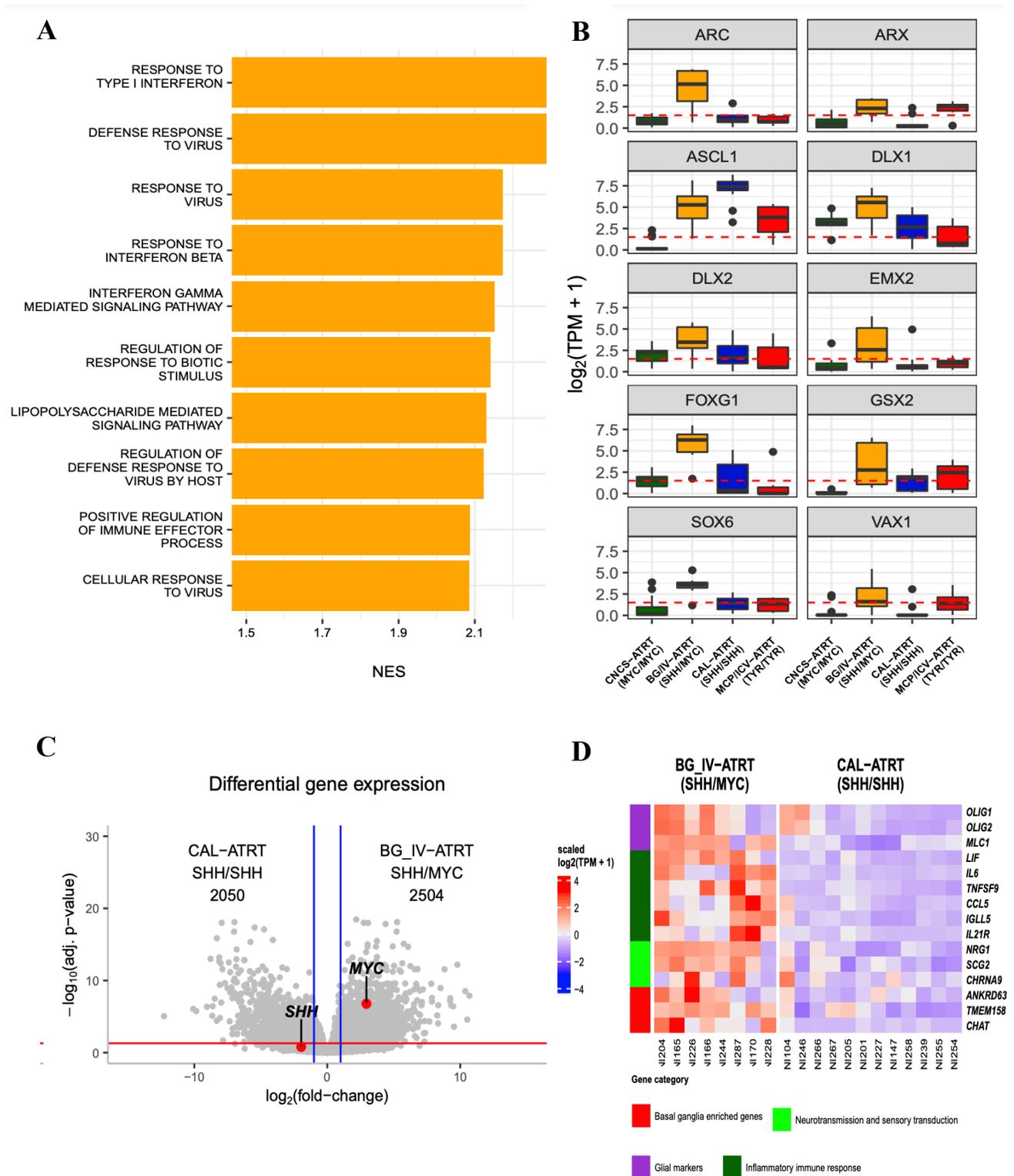
Ganglionic eminences are a progenitor domain during forebrain development and laid from the ventricular to the sub ventricular zones, that will eventually give rise to cortical neurons (Miyoshi et al., 2010; Wonders and Anderson, 2006) and the basal ganglia (Donkelaar, 2015; Kelly et al., 2018; Turrero García and Harwell, 2017). Notably, *SOX6* (SRdY-box 6) is involved in the specification of telencephalic progenitors and interneuron heterogeneity and is usually used as a marker for medial GE progenitors (Azim et al., 2009; Chen et al., 2013). The GE pan-marker *GSX2* regulates progenitor expansion and neurogenesis and oligodendroglial fate specification in GE progenitors (Chapman et al., 2012; De Mori et al., 2019; Roychoudhury et al., 2020); and *ARX* plays a role in neuronal progenitor proliferation and development of cerebral cortex, thalamus, hippocampus, striatum and olfactory bulbs (Fulp et al., 2008). In summary, all these genes are extremely important in the fate specification of GE progenitor and the development of the basal ganglia.

Next, we sought to investigate the specificities of the two subtypes of SHH ATRT. The DEA of BG/IV-ATRT versus CAL-ATRT showed the overexpression of genes involved in neurotransmission and sensory transduction, inflammatory immune response, glial markers and finally, genes normally enriched in basal ganglia region (Figure 5C, D). Interestingly,

besides the GE marker we found the overexpression of glial genes known to be also expressed in neural progenitors of the VZ and SVZ of the ganglionic eminence (OLIG2) (Jakovcevski and Zecevic, 2005; Szu et al., 2021) and in multipotent neural precursors of the SVZ of the lateral ventricle (MLC1) (Schmitt et al., 2003).

Taking together, these results suggested that BG/IV-ATRT are tumors from neuronal origin and pointed out the ganglionic eminence neuronal progenitors present during forebrain development as candidate cells of origin.

**Figure 5. BG\_IV-ATRT subgroup molecular pattern suggests neuronal origin and point out the ganglionic eminence region as tissue of origin.**



**Figure 5. BG/IV-ATRT subgroup molecular pattern suggests neuronal origin and point out the ganglionic eminence region as tissue of origin.**

- A. GSEA result showing the top 10 enriched GO Biological Process gene sets in BG\_IV-ATRT compared to all other anatomico-molecular subgroups.
- B. Boxplot showing the expression levels of typical ganglionic eminence gene markers (ASCL1, DLX1, DLX2, GSX2, VAX1, SOX6) and forebrain development (FOXP1, EMX2, ARX) in the ATRT anatomico-molecular subgroups.
- C. Volcano plot showing differential gene expression analysis result of BG/IV-ATRT versus CAL-ATRT. The X axis indicates the log<sub>2</sub> transformed fold-change (log<sub>2</sub>(fold-change)) and the Y axis indicates the reverse of the log<sub>10</sub> transformed adjusted *p*-value (-log<sub>10</sub>(adj. *p*-value)). MYC and SHH genes are labelled. 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).
- D. Heatmap showing the expression level of interested genes differentially expressed in BG/IV-ATRT compared to CAL-ATRT.

### **BG/IV-ATRT and CNCS-ATRT inflammatory immune response explain their close transcriptomics profile**

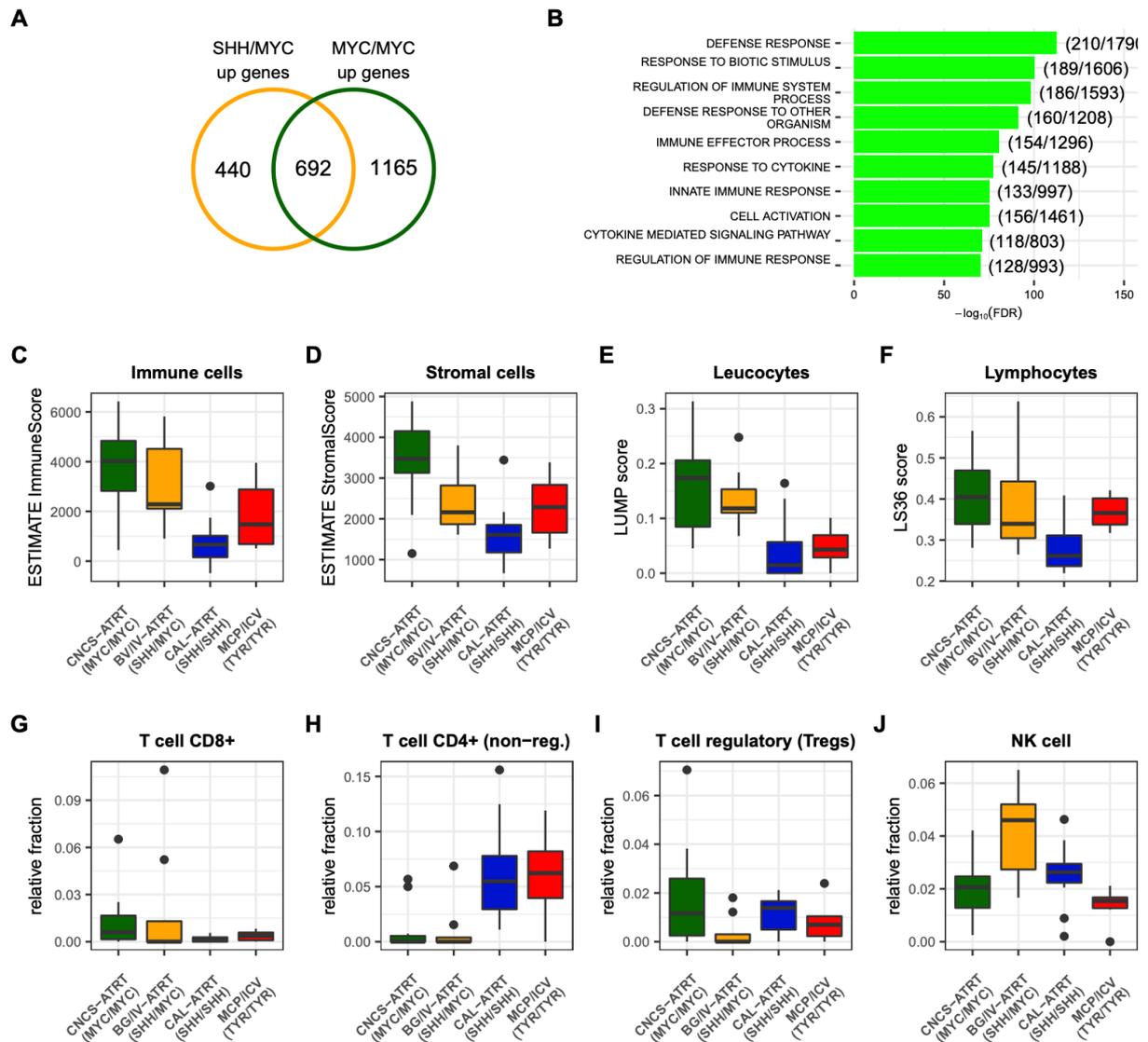
Despite their SHH methylation profiling and their clear neuronal origin, BG/IV-ATRTs showed a remarkably inflammatory immune response (Figure 5A) as we had previously seen in CNCS-ATRTs belonging to the MYC subgroup (Figure 3C). This made us to hypothesize that this feature brings these two subgroups close together at transcriptomics level (Figure S2A). To confirm this hypothesis, we performed pairwise DEA between BV/IV-ATRT, CNCS-ATRT and CAL-ATRT and look at the overexpressed genes in BV/IV-ATRT and CNCS-ATRT compared to CAL-ATRT (Figure 6A). The Gene Ontology (GO) of these common overexpressed genes were clearly associated with the inflammatory immune response terms (Figure 6B).

We thus attempted to infer the immune cell content of each subgroup based on their transcriptomics data and using the ESTIMATE algorithm (as previously described in Leruste et al). We found that the BG/IV-ATRT and the CNCS-ATRT subgroups showed the highest immune score, while the CAL-ATRT subgroup showed the lowest stromal score (Figure 6 C, D). Consistently, estimating immune cells infiltration rate, based on DNA methylation data using LUMP and LS36 scoring (see material et methods), showed again the highest lymphocytes and leucocytes scores for BG/IV ATRTs and CNCS-ATRTs (Figure 6 E, F).

Finally, to better characterize this inflammatory immune response, we performed deconvolution-based analysis approach using Quantiseq algorithm (Finotello et al., 2019) and infer the relative content of different immune cells for each sample. We found that, in average, CNCS-ATRTs and BG/IV-ATRTs had similar great CD8 T cells and similar low CD4 T cells content as compared to CAL-ATRT and MCP/ICV-ATRT (Figure 6 G, H), but CNCS-ATRT showed highest T cell regulatory and BG/IV-ATRT highest NK cell suggesting some specificities related to their different location (Figure 6 I, J).

Altogether, these results showed that BG/IV-ATRTs share with CNCS-ATRTs a more prominent immune infiltrate than MCP/ICV-ATRT and CAL-ATRT, that accounted for their « MYC » expression signature in spite of their typically neuronal location and SHH methylation profile.

**Figure 6. BG/IV-ATRT and CNCS-ATRT inflammatory immune response explain their close transcriptomics profile.**



**Figure 6. BG/IV-ATRT and CNCS-ATRT inflammatory immune response explain their close transcriptomics profile.**

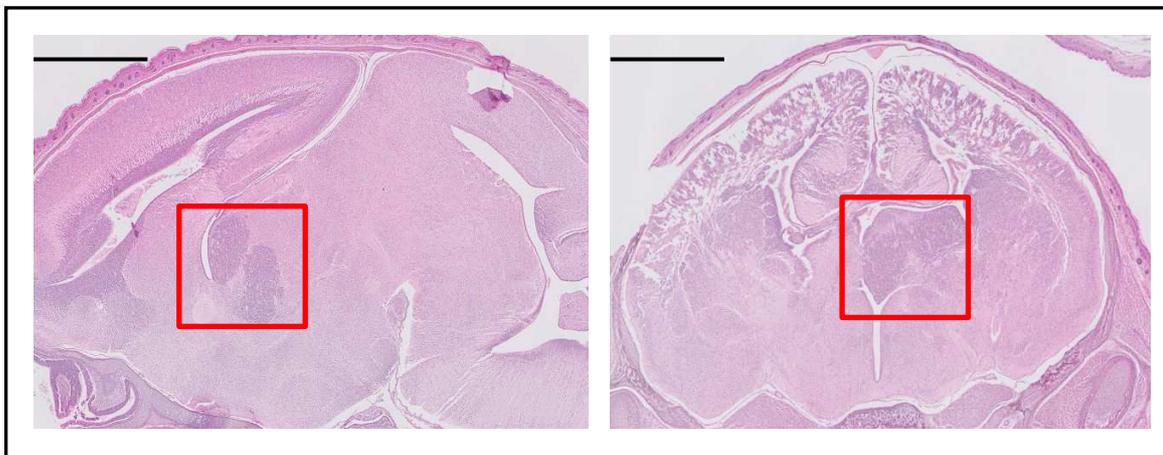
- A. Venn diagram showing the number of genes overexpressed from differential expression analysis between BG/IV-ATRT (SHH/MYC) versus CAL-ATRT (SHH/SHH) (orange circle); CNCS-ATRT (MYC/MYC) versus CAL-ATRT (SHH/SHH) (green circle). The intersection shows the number of overexpressed genes shared by BG\_IV-ATRT and CNCS-ATRT anatomical-molecular subgroups.
- B. Gene Ontology of common overexpressed genes shared by BG/IV-ATRT and CNCS-ATRT.
- C. In silico estimation of immune cell infiltration rate based on gene expression data (RNA-seq) stratified by ATRT anatomical-molecular subgroup using ESTIMATE algorithm.
- D. In silico estimation of stromal cell infiltration rate stratified by ATRT anatomical-molecular subgroups using the ESTIMATE algorithm.
- E. In silico estimation of leucocytes infiltration rate based on DNA methylation data, stratified by ATRT anatomical-molecular subgroups using the LUMP scoring.
- F. In silico estimation of lymphocytes infiltration rate based on DNA methylation data, stratified by ATRT anatomical-molecular subgroups using the LS36 scoring.
- G. - J. Relative fraction of T Cell CD8+, T Cell CD4+, T cell regulatory and NK cell using Quantiseq algorithms and stratified by anatomical-molecular ATRT subgroup.

### **Mouse Shh tumors recapitulate the human BG/IV-ATRT**

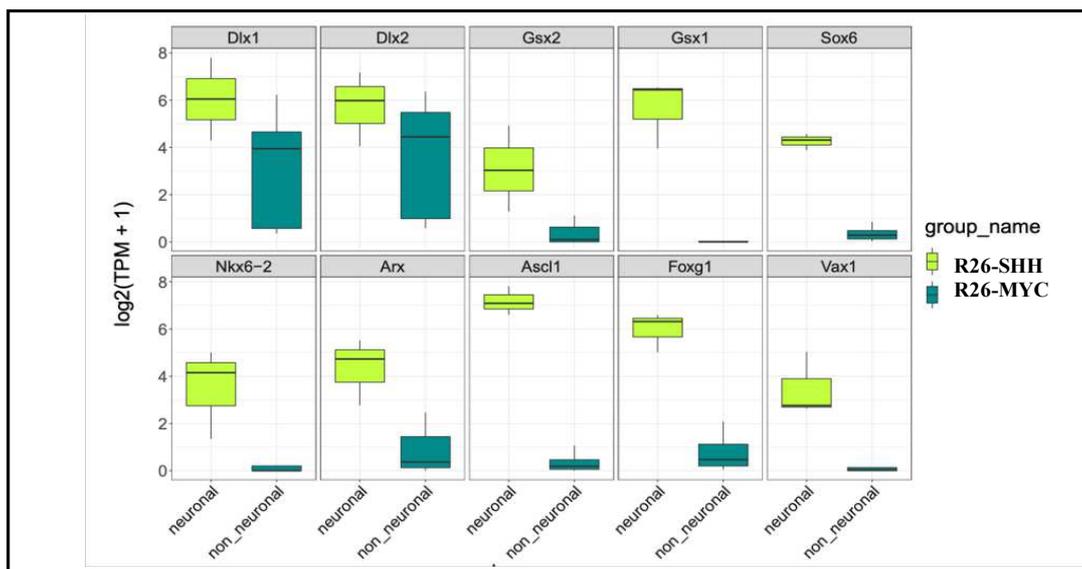
We then investigated whether mouse tumors could corroborate our hypothesis. As aforementioned, we generated an extended series of tumors from our previously published Rosa26-Cre<sup>ERT2</sup>::*Smarcb1*<sup>flox/flox</sup> model, and characterized them by RNAseq profiling. While MYC tumors constantly showed an extra-CNS location, all SHH ATRTs developed from intraparenchymal regions (Figure 4A). More precisely, almost all these intraparenchymal tumors were located in the sub ventricular region, and invaded the basal ganglia, a location fitting perfectly with the description of human BG/IV-ATRT (Figure 7A, Figure S7). Remarkably, Shh tumors were exclusively obtained at the earliest inactivation of *Smarcb1*, i.e. E6 to E7, suggesting earlier progenitors for BG/IV-ATRTs (Figure 4A). Finally, as their human counterparts, the murine tumors were characterized by the expression of typical GE and forebrain development gene markers (Figure 7B).

**Figure 7. Mouse *Shh* tumors recapitulate the human BG SHH ATRT**

**A**



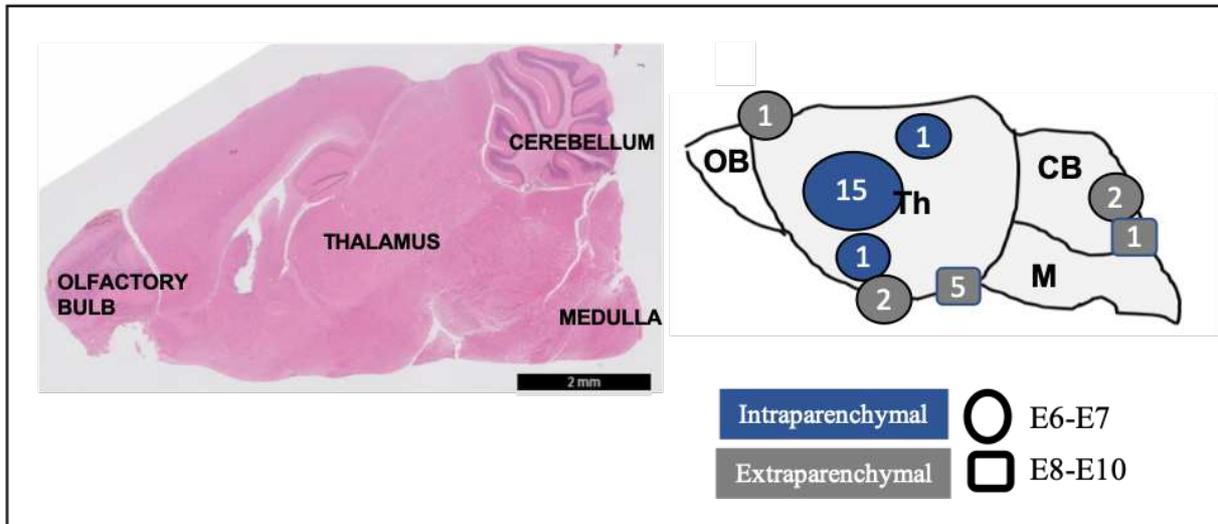
**B**



A. Sagittal (left) and coronal (right) brain slides showing intraparenchymal tumors located in the subventricular region. The red box indicates precise tumor location in the hypothalamic (left) and basal ganglia (right) area. Scale bar: 1 mm.

B. Boxplot of gene expression showing the level of expression of typical ganglionic eminence gene markers (DLX1, DLX2, GSX1, GSX2, SOX6, NKX6.2, ASCL1, VAX1, SOX6, NKX6.2) and forebrain development (FOXG1, ARX) in the Rosa26-SHH compared to Rosa26-MYC tumors.

**Figure S7**



**Intracranial tumor (Intraparenchymal and extraparenchymal) localization according to the time of *Smarcb1* inactivation (E6-E7 versus E8-E10).** Sagittal IHC mouse brain slice showing the different anatomical regions (left) and schematic view of the same slide showing the intracranial tumor location (intraparenchymal, dark blue; extraparenchymal, grey) and the number of tumors. Most E6-E7 intracranial tumors are intraparenchymal (17/22, 77%); all E8-E10 intracranial tumors are extraparenchymal (6/6) and are located in the meninges. OB: Olfactory Bulb; Th: Thalamus; CB: Cerebellum; M: Medulla.

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

Our data clearly demonstrated that CAL-ATRTs subgroup was distinct from BG/IV-ATRTs (Figure 1F and Figure 5C) and were composed exclusively by tumors spreading quadrigeminal cistern appearing to originate from the cerebellar anterior lobe (Figure 1F, Figure S1). To further characterize this subgroup, we performed differential gene expression analysis and GSEA between this subgroup and all other groups. The top 100 most overexpressed genes were composed by genes that were related to neurogenesis and neuronal migration (such as *SOX1*, *NTNG2*, *NEUROD4*, *NSG2*, *DCX*), WNT and FGF signaling pathway and finally, genes involved in midbrain-hindbrain boundary (MHB) patterning and cerebellum development (Buckles et al., 2004; Di Giovannantonio et al., 2014; Gasser et al., 2016; Hirata and Tomita, 2001; Lobe, 1997) (Figure 2C, Figure 8A).

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. This region is specified by the expression of *FGF8* and *WNT3A* at the edge, markers of the rostral (midbrain) and caudal (hindbrain) structures such as Iroquois homeobox genes (*IRX1*, *IRX2*), Engrailed family (*EN1*, *EN2*) (Hanks et al., 1995; Joyner et al., 1991; Martinez et al., 2013) and *HES3* and *PAX3*. In particular, Iroquois homeobox genes (*IRX1*, *IRX2*) are pre-patterning factors expressed in the presumptive cerebellum that positively regulate proneural genes (Matsumoto et al., 2004). Engrailed family (*EN1*, *EN2*) are expressed early in cerebellar and mesencephalic primordial neuroepithelium and is involved in the formation of the cerebellum (Hanks et al., 1995; Joyner et al., 1991; Martinez et al., 2013). Finally, *HES3* and *PAX3* are involved in midbrain-hindbrain patterning and tectum development (Hibi and Shimizu, 2012; Nakamura et al., 2005; Peretz et al., 2016). We strikingly found that all these MHB core set of genes were characteristic for CAL SHH ATRTs. Consistently, GSEA clearly pointed to enrichment of embryonic neuronal development gene sets (Positive Regulation of Neural Precursor Cell Proliferation) and to midbrain/hindbrain patterning gene sets (Midbrain Neurotypes Human Neuroblasts and Human Neuronal Progenitor, Cell Differentiation in Hindbrain) in CAL-ATRT (Figure 8B).

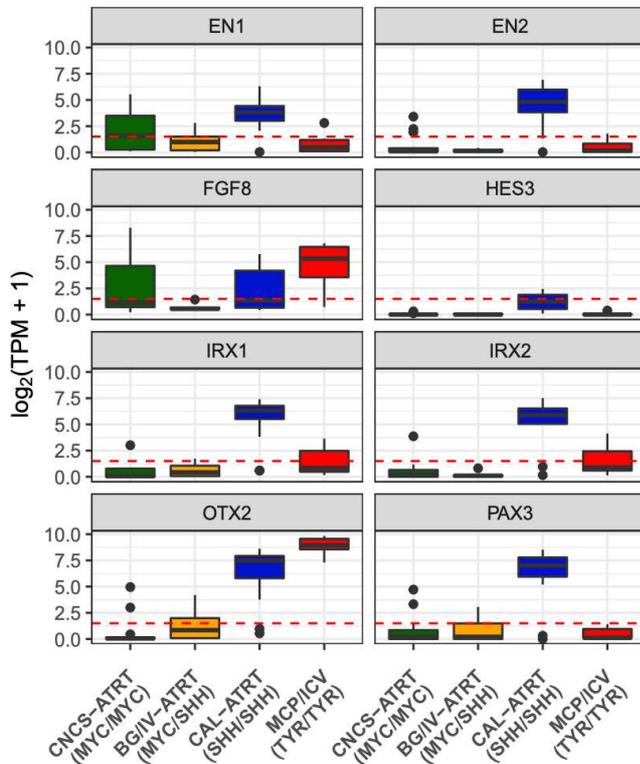
In order to better delineate the molecular specificity of the CAL-ATRT subgroup, we performed two rows of differential analyses. First, we compared it to the BG/IV-ATRTs and found in addition to the already described signature of MHB, the overexpression of genes involved in neurogenesis, WNT and FGF signaling pathway and genes that are normally enriched in cerebellum such as *RUNX1T1*, *EPHB1* and *CNPY1* (also involved in the

establishment of the MHB) (Partanen, 2007) (Figure 8C). CAL-ATRT were then compared to infants SHH medulloblastomas, considering that both types of tumors develop in the cerebellum of young children and activate the SHH pathway, but also that their anatomic location could result from different lineages of origin. We again found that CAL-ATRTs were characterized by the MHB signature, strengthening our previous results; in addition, CAL-ATRT showed overexpression of genes involved in NOTCH, WNT and FGF signaling pathways and the overexpression of stem cell and neuronal progenitor genes markers (Figure 8D). SHH MB showed an overexpression of genes related with granule cell precursors (GCP), HIST1 family of genes and, as previously described, non-coding miRNAs and lncRNAs (Zhao et al., 2021) (Appendix Figure 1).

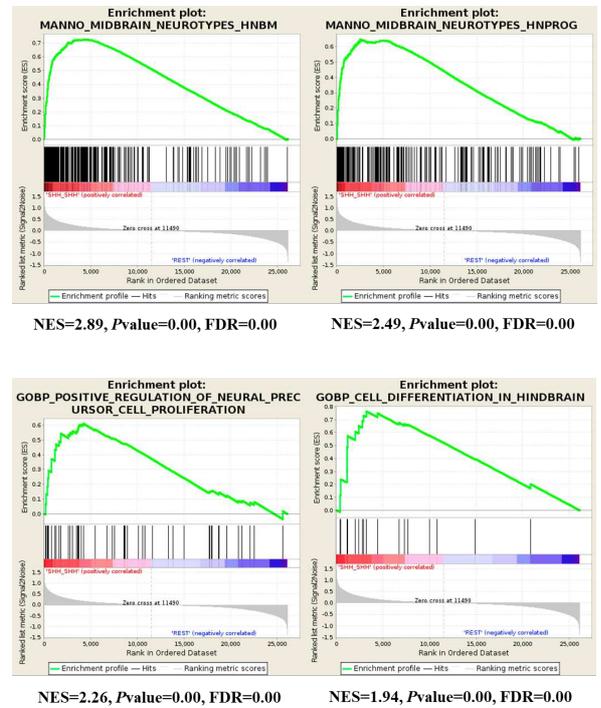
All these findings suggested that the CAL-ATRT subgroup are related to neuronal cell progenitors involved in the MHB which is highly consistent with their anatomical location.

**Figure 8. CAL-ATRT anatomical location and molecular profile suggest neuronal progenitor from the midbrain-hindbrain boundary as cell of origin.**

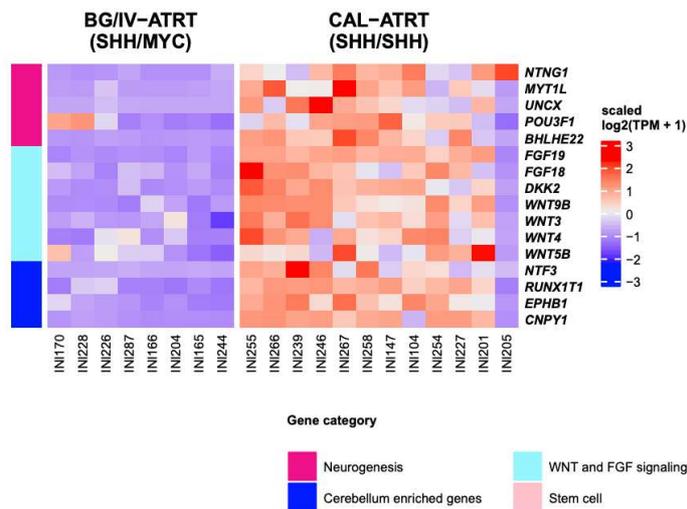
**A**



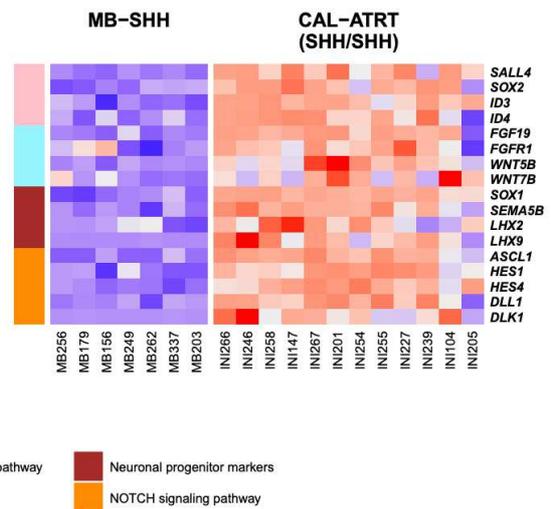
**B**



**C**



**D**



**Figure 8. CAL-ATRT anatomical location and molecular profile suggest 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 ATRT anatomical-molecular subgroups.
- B. GSEA enrichment plots of gene sets enriched in CAL-ATRT compared to all other anatomical-molecular subgroups.
- C. Heatmap showing the level of expression of interested genes differentially expressed in CAL-ATRT (SHH/SHH) compared to BG/IV-ATRT.
- D. Heatmap showing the level of expression of interested genes differentially expressed in CAL-ATRT (SHH/SHH) compared to medulloblastoma-SHH (MB-SHH).

## Single-cell RNAseq analysis reveals transcriptional intra-tumoral heterogeneity of cerebellar anterior lobe ATRTs.

To get further insights on putative cells of origin of the CAL-ATRTs we performed single-cell RNAseq analyses (scRNAseq) on three fresh tumor specimens using 10x Genomics Chromium technology.

We applied different clustering approaches (see methods) that were highly congruent corroborating the existence of 10 clusters (Figure 9A). From this clustering, we first checked the expression of genes that were known to be specific to ATRT-SHH molecular subgroup identified in previous studies (Johann et al., 2016; Torchia et al., 2016). We found that most of these ATRT-SHH signatures were differentially expressed among clusters. For instance, *ASCL1* and *DLL3* were specific to cluster 5 and *STMN4* to cluster 3, while *TTYH1* was expressed in all clusters but cluster 3 and 5. This finding demonstrated the transcriptional intratumoral heterogeneity of ATRT-SHH (Figure S9A).

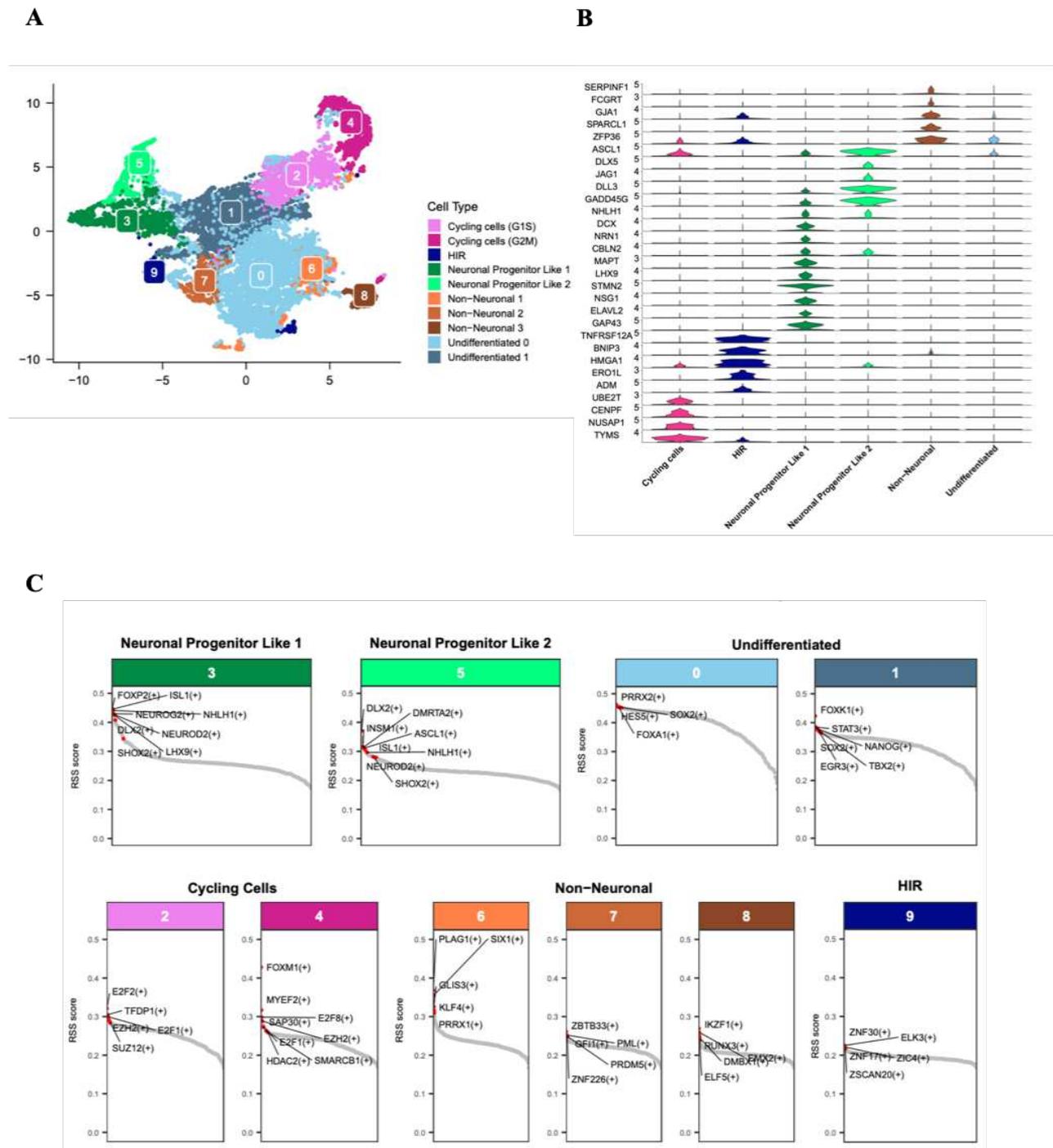
Aiming to unravel the biological identity of each identified cell cluster, we performed differential expression analyses in a one versus all others manner and identified marker genes characteristic for each cluster. Gene expression atlas databases (see material et methods section) and literature curations were then used to define the biological identity of each cluster based on their respective gene markers. We first identified two clusters (cluster 2 and cluster 4) that were associated with cell cycle features. Cluster 2 was composed by cells at G1 to S transition (*MCM3/4/6/7*, *GINS2*, *PCNA* in the so-called G1S CC) while cluster 4 contained cells at G2 to M transition (*TOP2A*, *BIRC5*, *NUSAP*, *CDC20*, *CCNB1*, *PLK1*, *DLGAP5* in the so-called G2M CC) (Figure S9B, C). We next found two clusters (clusters 3 and 5) characterized by the expression of different neuronal markers such as *GADD45G*, *NHLH1*, *DCX*, *NEUROD1*, *NRN1*, *CBLN2*, *MAPT* (Dennis et al., 2019; Eze et al., 2021; Jessa et al., 2019; La Manno et al., 2016; Parisian et al., 2020; Peukert et al., 2011; Yuzwa et al., 2017). In particular, *GADD45G* and *NHLH1* and *DCX* are described to be expressed in progenitor neurons and early newborn and migrating neurons. In addition to these common neuronal genes markers, cluster 3 presented specific expression of *LHX9*, *NSG1*, *GAP43*, *STMN2* and *ELAV2* genes and cluster 5 showed specific expression of *ASCL1*, *DLX5*, *JAG1*, *DLL3*, *NEUROG1* genes (Figure 9B).

Additionally, we found 4 clusters with marker genes of non-neuronal cell types (Figure S9C) in both human and mouse reference atlases; three of them showed endothelial and glial genes markers such as *IFITM3* and *SERPINF1*, cluster 6, referred as NN1; *CD82*, *ZFP36* and *ATPIA2*, cluster 7, referred as NN2; *GJA1*, *SPARCL1* and *NKAIN4*, cluster 8, referred as

*NN3*. The cluster 9 expressing *TNFRSF12A*, *ERO1L* and *BNIP3* was referred to as “hypoxic inflammatory response (HIR)”. We finally defined two clusters as “undifferentiated” (UD1 and UD2) since they did not show specific well-defined marker genes (Figure 9B, S9E). However, as these were the two largest clusters in terms of number of cells (Figure S9D), we considered them as the core of the tumoral cells. Thus, these 10 identified clusters could be broadly grouped into 6 biological cell types (Figure 9A, B): cycling cells (cluster 2 and 4), Neuronal Progenitor-like 1 (NPL1) (cluster 3), Neuronal Progenitor-like 2 (NPL2) (cluster 5), Non-Neuronal cells (NN) (cluster 6, 7, 8), hypoxic inflammatory response (HIR) (cluster 9) and Undifferentiated cells (UD) (cluster 0 and 1). Genes defining the identity of each cell type are shown in Figure 9B.

To further characterize the biological meaning of each cluster, we performed gene regulatory network analysis based on the SCENIC framework (Aibar et al., 2017). We found that each cluster could be characterized by the activity of specific Transcription Factors (TF) (Figure 9C). Interestingly, we found that neuronal clusters (3 and 5) shared specific neurogenic TFs such as *ISL1*, *NHLH1*, *NEUROG2*, *DLX2* and *SHOX2* suggesting they were closely related and likely shared a common origin. In addition, UD0 and UD1 showed TFs related to epithelia-mesenchymal transition and mesodermal commitment pathway (*PRRX2*, *MEOX1*, *FOXA1* and *TBX2*), stem cell and pluripotency maintenance (*SOX2*, *NANOG*, *STAT3*), neuroglial fate determination (*SOX9*) and cell growth and differentiation (Forkhead box). As expected, cycling clusters (2 and 4) showed the activity of TFs involved in cell cycle progression and cell proliferation and also those of SWI/SNF (*SMARCB1*), PRC2 complex (*EZH2*, *SUZ12*) subunits and histone deacetylases (*SAP30* and *HDAC2*). Finally, NN clusters showed an over-representation of zinc finger related proteins (i.e. *GLIS3*, *KLF4*, *ZBTB33*, *ZNF226*) (Figure 8C). These results were consistent with signatures found by differential gene expression analyses, especially for the two neuronal clusters.

**Figure 9. Single-cell RNAseq analysis reveals transcriptional intra-tumoral heterogeneity of cerebellar anterior lobe ATRTs.**

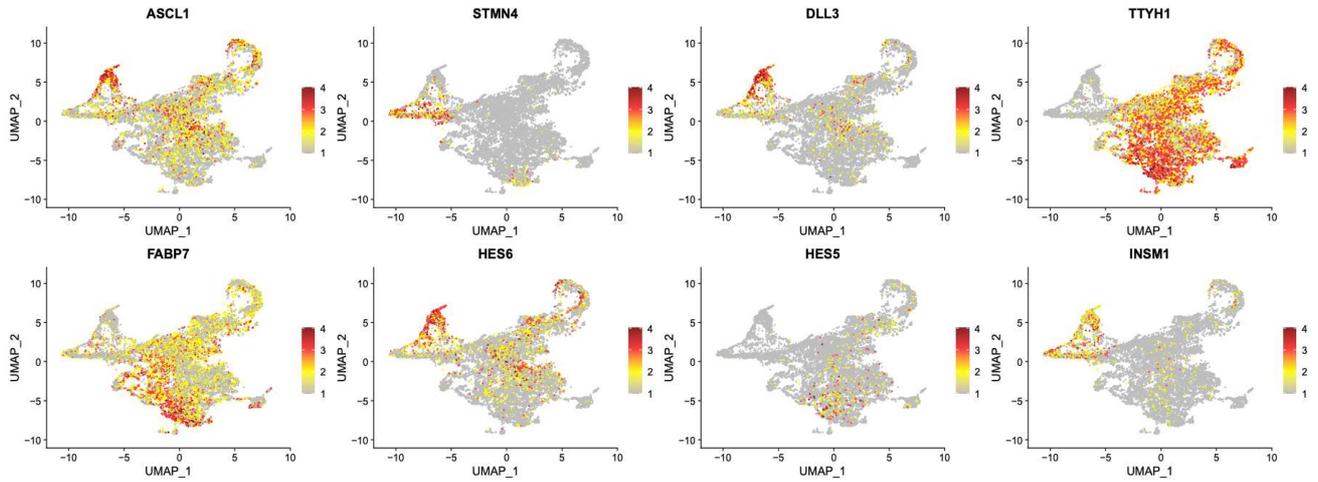


**Figure 9. Single-cell RNAseq analysis reveals transcriptional intra-tumoral heterogeneity of cerebellar anterior lobe ATRT**

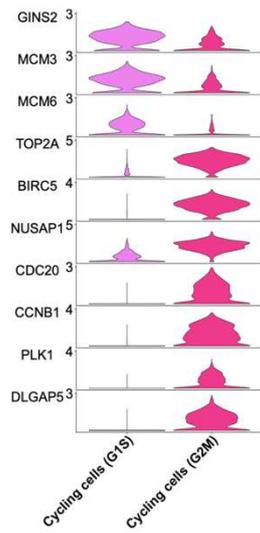
- A. Visualization of 10 cell groups using uniform manifold approximation projection (UMAP).
- B. Violin plots showing specific gene markers for each grouped cell cluster.
- C. Regulon specificity score (RSS) for each cell cluster. In each cluster, regulons (TF and their direct targets) are ordered according to their RSS score. Interesting regulons are labelled.

**Figure S9.**

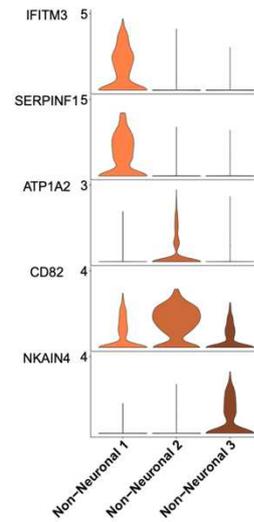
**A**



**B**

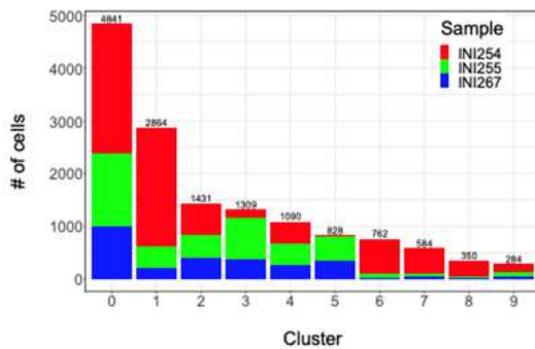


**C**



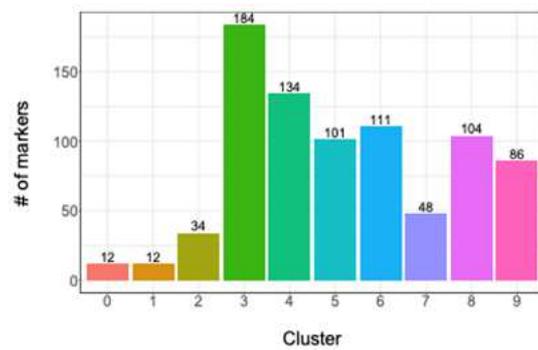
**D**

**Figure: Number of cells**



**E**

**Figure: Number of markers**



**Figure S9.**

- A. UMAP showing intra-tumoral heterogeneity by the expression of typical known markers of SHH-ATRT subgroup described by Johann et al., 2016 and Torchia et al., 2016.
- B. Violin plots showing cell-specific gene markers for cycling clusters (G1S, G2M).
- C. Violin plots showing cell-specific gene markers for non-neuronal clusters (NN-1, NN-2 and NN-3).
- D. Bar plot showing the number of cells for each cell cluster.
- E. Bar plot showing the number of gene markers for each cell cluster.

## **Neuronal Differentiation Impairment involving Notch signaling pathway leads to malignant transformation in CAL-ATRT**

Since we identified neuronal clusters in the previous analysis, we hypothesized that these clusters could be the cells of origin of the undifferentiated clusters, the core of the tumor cells.

To test this hypothesis, we first performed trajectory inference analysis on neuronal and undifferentiated clusters. Two different algorithms were applied: the PAGA algorithm implemented in Monocle3 and the elastic principal graph algorithm implemented in EIPiGraph (see material et methods). Both approaches give consistent results and identify a path from NPL1 to UD cells via NPL2 (Figure 10A) suggesting that NPL1 are a possible cell of origin of the tumor cells.

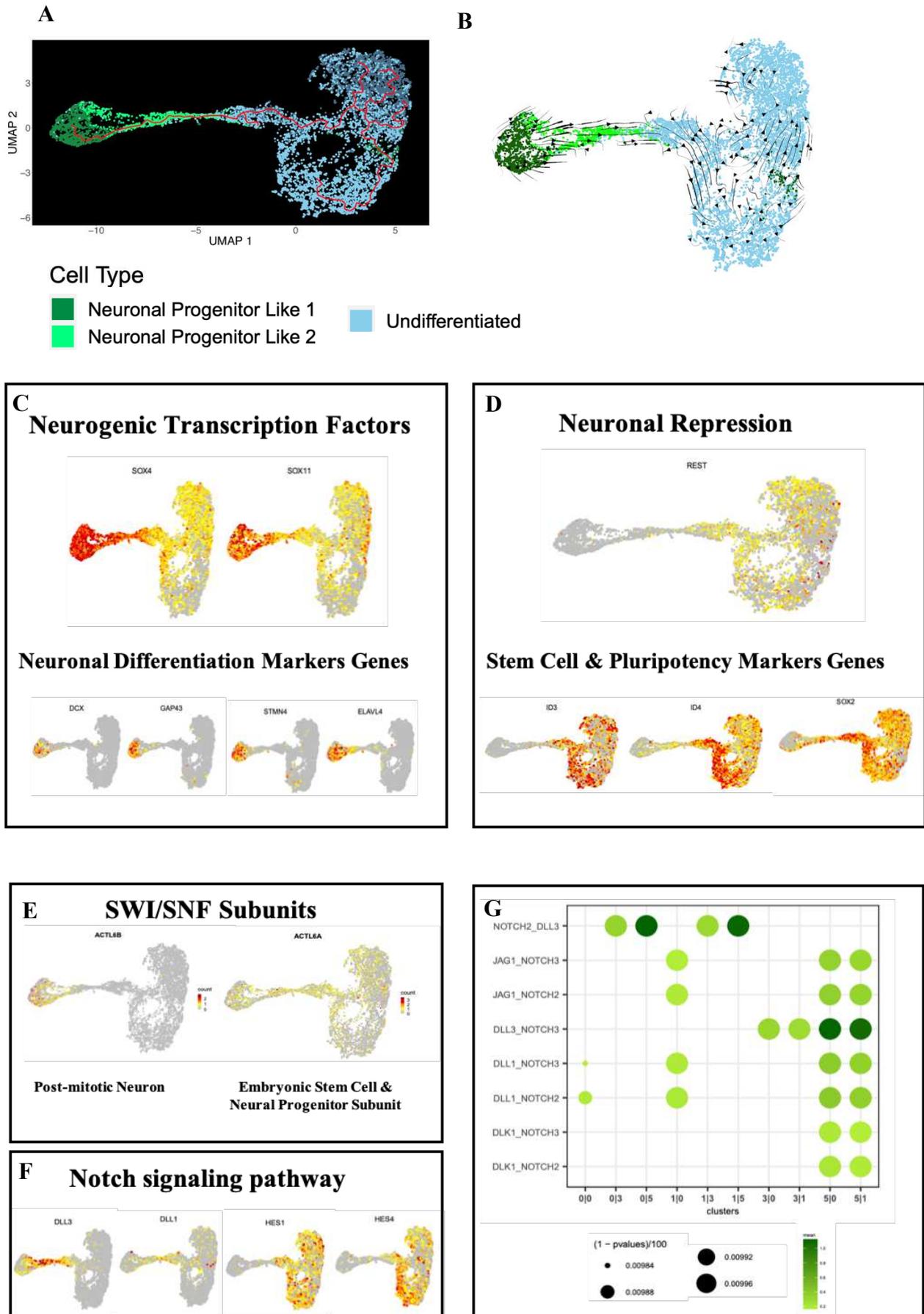
To confirm that the NPL1 are the root of the trajectory, we performed RNA velocity analysis on neuronal and undifferentiated clusters using scVelo tool (La Manno et al., 2018). The idea of this approach is to infer the future state of each cell based on the transcriptional dynamics of each gene deduced from the spliced/unspliced transcript ratios. We use the marker genes of neuronal clusters (3 and 5, NPL1 and NPL2) and the undifferentiated tumoral clusters (0 and 1, UD1 and UD2) to perform this analysis and found streams of cells starting from NPL1 to the undifferentiated tumoral cells (Figure 10B) confirming our hypothesis. Pseudotime and latent time analyses highlight the temporal progression from NPL1 to the undifferentiated cells (Figure S10A, B).

Interestingly, in line with the trajectories, we found the gradual expression of the neurogenic transcription factor SOX4 and SOX11 in the NPL clusters; this expression is attenuated progressively in the undifferentiated clusters (Figure 10C). Consistently with the function of these TF to enable the expression of pan-neuronal genes, we found the expression of genes involved in neuronal differentiation (Figure 10C). Besides this, we observed the decreasing expression of ACTL6B, the SWI/SNF subunit restricted to post-mitotic neurons, from NPL to UD clusters (Figure 10E), together with the progressive increase of neuronal repressors, stem cell and pluripotency genes markers (Figure 10D).

In addition, we observed the expression of genes involved in NOTCH signaling pathway throughout this trajectory suggesting a role of NOTCH signaling in the malignant transformation (Figure 10F). To confirm this hypothesis, we performed ligand-receptor interaction analysis using CellPhoneDB database and tool (Efremova et al., 2020). The results shown in Figure 10G demonstrated that the ligand of NOTCH signaling are specifically expressed in the NPL2 cluster while the NOTCH receptors are specifically expressed in the

UD cluster (Figure S10C). Other signaling pathways such as BMP, FGFR and EPHB pathways were also found to be activated in undifferentiated cells suggesting their possible role in the tumor cell development (Figure S10D, S10E, S10F).

**Figure 10. Neuronal Differentiation Impairment involving Notch signaling pathway leads to malignant transformation in CAL-ATRT .**

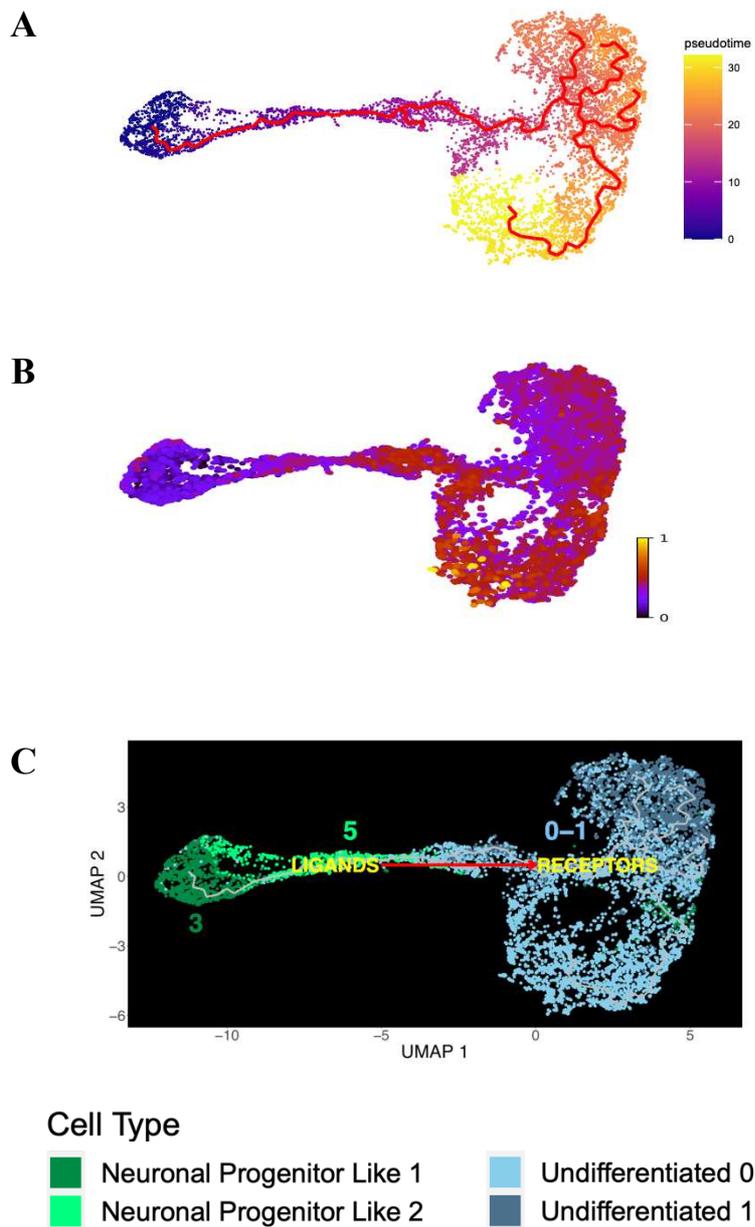


**Figure 10. Neuronal Differentiation Impairment involving Notch signaling pathway leads to malignant transformation in CAL-ATRTR.**

- A. Trajectory inference analysis using the PAGA algorithm (Monocle3) showing a path from NPL1 cluster to UD clusters via NPL2 cluster.
- B. RNA velocity analysis using scVelo tool showing streams of cells starting from NPL1 through NPL2 to the undifferentiated tumoral cells.
- C. Heatmap showing the gradual expression of the neurogenic transcription factors (SOX4 and SOX11) and the expression of genes involved in neuronal differentiation (STMN4, ELAVL4, GAP43, DCX).
- D. Heatmap showing the expression of neuronal repression, stem cell and pluripotency markers genes.
- E. Heatmap showing the expression of SWI/SNF chromatin remodeling complex subunits ACTL6A (embryonic stem cell and neural progenitor marker) and ACTL6B (postmitotic neuron marker).
- F. Heatmap showing the expression of NOTCH pathway genes.
- G. Notch pathway ligand-receptor interaction between NP clusters (3 and 5) and UD clusters (0 and 1) using Cellphone DB.

## Figure S10

- A. Pseudotime showing the temporal progression from NPL1 to undifferentiated clusters.
- B. Latent time analyses showing the temporal progression from NPL1 to undifferentiated clusters.
- C. Schematic representation showing the relationship between ligands and receptors of the NOTCH pathway in Neuronal Progenitor and Undifferentiated clusters.

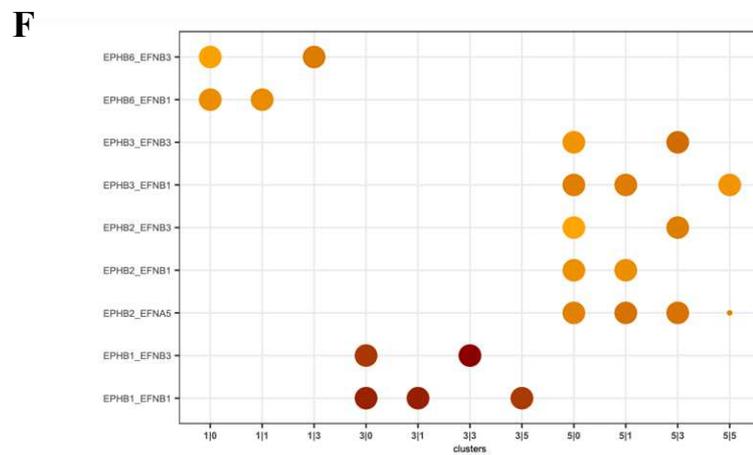
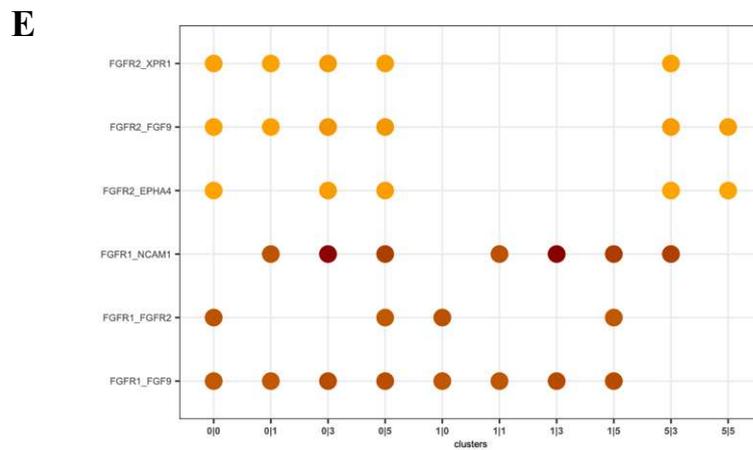
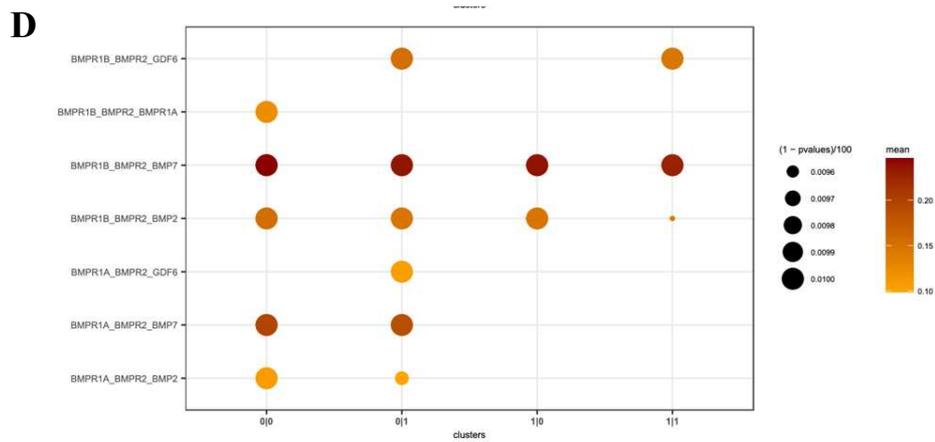


**Figure S10 (continuation)**

D. BMP pathway ligand-receptor interaction between NPL clusters (3 and 5) and UD clusters (0 and 1).

E. FGFR pathway ligand-receptor interaction between NPL clusters (3 and 5) and UD clusters (0 and 1).

F. EPHB pathway ligand-receptor interaction between NPL clusters (3 and 5) and UD clusters (0 and 1).



## **Material et Methods**

## **Animals**

The following mouse strains were used in this study: Smarcb1<sup>f/f</sup>;R26-Cre<sup>ERT2</sup> (Han et al;2016), NestinCre (Tronche et al., 1999), genotyped according to these references. 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 (Workman et al., 2010).

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

## **Tamoxifen administration (Smarcb1<sup>f/f</sup>; R26-Cre<sup>ERT2</sup>)**

Tamoxifen administration according to the reference (Han et al., 2016).

## **Mouse Tumor monitoring (Smarcb1<sup>f/f</sup>; R26-Cre<sup>ERT2</sup>)**

Mice were monitored for tumor formation at least 3 times per week. The observation period encompassed at least 18 months. The animal who shows clinical behaviour was euthanized, tumors and all organs were taken and fixed in AFA for histology or were frozen in -80°C until RNA extraction.

## **Immunofluorescence staining for mouse sample**

Mouse embryonic brains were dissected out of the skull, fixed in 4% PFA for 4 hours, and store it in at 4° until use (no more than 1 month). Inclusion of brain in Agarose gel 4% and store at 4° C to allow it to solidify faster. Although it is usually solidified within 20 minutes, for Smarcb1F/F; NestinCre+, due to their thin cortex, it is recommended to leave it for at least 1 hour. Then, 80-120 µm-thick slices were prepared with a Leica VT1200S vibratome in PBS. Slices were then blocked in PBS-Triton X100 0.3%-Donkey serum 2% at room temperature for 2 hours, incubated with primary antibody overnight at 4°C in blocking solution, washed in PBS-Tween 0.05%, and incubated with secondary antibody overnight at 4°C in blocking solution before final wash and mounting in aquapolymount. Image analysis, modifications of brightness and contrast were carried out with Fiji.

### **Immunofluorescence staining antibodies**

Primary antibodies: Goat anti Sox2 (AF2018-100UG R and D System, 1/500), p3 Rabbit anti Phospho Histone3 ( ab47297 Abcam, 1/2000), rabbit anti Pax6 (Biolegend 901301, 1/500).

Secondary antibodies: donkey Alexa Fluor 555 anti-rabbit (Jackson laboratories 711-165-152, 1/250), donkey Alexa Fluor 488 anti-goat (Jackson laboratories 705-545-147, 1/250)

### **Mouse sample histological examination**

Organs were collected, frozen on dry ice and processed for cryosectioning and macrodissection) 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, as described in Bourdeaut et al., 2007. Ki67 (sc-7846, dilution 1/300).

### **Mouse tumours 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 tumour could be identified; macrodissection was then performed with a sterile scalpel. Small pieces of tissue containing the tumour cells were frozen at 80 °C until RNA preparation. The tumour 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). All cases of ATRT were SMARCB1-deficient. These data have not been previously published.

## **RNA sequencing (RNA-seq)**

### **RNA-seq library preparation**

Total RNAs were obtained from ATRT (n = 49) frozen samples using Qiagen QIAamp RNAeasy kit, according to the manufacturer's procedures (Qiagen, Venlo, Netherlands). 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 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 (Dobin et al., 2013). The human reference genome hg19 and the ref-Flat table gene annotation (downloaded in GTF format on October 2nd 2018 from UCSC website and including 27,979 transcripts) were respectively used for mapping and read counting.

### **RNA-seq statistical analysis**

All genes having 0 counts in all samples (1865 genes) were filtered out before subsequent analyses. Variance stabilization process was then applied to the filtered raw counts using the `rlog()` function of DESeq2 version 1.30.1 (Love et al., 2014) Bioconductor package. An additional filtering step were applied consisting on all genes with at least 1 sample with an expression level higher than 7.5 in `rlog`. This threshold value was set by visual inspection of the `rlog` count distribution. Finally, the 5000 highest variable genes (based on IQR range) among the remaining genes (16,904) were kept for the unsupervised analyses.

Principal component analyses 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., 2016) Bioconductor package. Pearson correlation and Ward's method were used respectively as similarity metric and agglomeration criterion. 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, 2010) 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) (McInnes et al., 2018) 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), (Lê Cao et al., 2011) was conducted using the mixOmics framework version 6.14.1 (Rohart et al., 2017), 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.

Differential gene expression analysis were performed using *DESeq()* function of the DESeq2 package. Resulting *p*-values were corrected using the Benjamini and Hochberg method (a.k.a. FDR). Genes having an FDR lower than 0.05 and a  $\log_2(\text{fold-change})$  higher than 1 are considered to be significantly differentially expressed.

Analyses were performed inside R version 4.0.2 environment.

Immune cells and stromal cells infiltration scores were computed using the ESTIMATE version 1.0.11 (Yoshihara et al., 2013) R 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 2.0.4 R package.

Analyses were performed inside R version 4.0.2 environment.

## **DNA methylation array**

### **DNA methylation array data and processing**

DNA methylation data (n = 54) from ATRT samples were obtained by the Illumina HumanInfiniumEPIC array. *RnBeads* Bioconductor package version 1.6.1 (Assenov et al., 2014) was used to preprocess the data. Raw data, in IDAT format, were imported using the *rnb.execute.import()* function. Bisulfite conversion efficiency and sample quality control were assessed using visual inspection of different plots (of control probes) generated by the *rnb.run.qc()* function (data not shown). All samples and all probes were kept after quality control. Probes intensity were normalized using the *rnb.execute.normalization()* function with the “illumina” method. Probes outside CpG context (2,991 probes), targeting single nucleotide polymorphisms (17,369 probes) or targeting X and Y chromosomes (19,457 probes) were all filtered out using, respectively, *rnb.execute.context.removal()*, *rnb.execute.snp.removal()* and *rnb.execute.sex.removal()* functions.

### **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* version 2.6.2 (Gu et al., 2016) Bioconductor package. Euclidean and 1-Pearson were used as distance metrics respectively for probes and samples clustering. Ward's method were used as linkage criterion for both sample and probes clustering.

The UMAP non-linear dimensionality reduction algorithm was applied for visualization purpose. UMAP analysis were performed using the *umap* version 0.2.7.0 CRAN package.

Leucocytes infiltration scores were computed based on the leucocytes unmethylated probes (LUMP) identified by Aran et al. 2015. 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 a each sample, as the mean of the lymphoid-specific hypermethylated probes identified by Killian et al., 2016.

Data processing and analyses were performed inside R environment (version 4.0.2).

## **Whole exome sequencing (WES)**

### **WES library preparation**

WES was conducted for ATRT INI256 case. Genetic tests were performed with the appropriate written informed consent. Tumour DNA was extracted from fresh-frozen tissue. WES libraries were generated using the SureSelect Agilent-XT2 protocol and sequencing was performed on an Illumina HiSeq2500 platform in paired-end 100 base-pair (bp) mode by the NGS Core Facility of Curie Institute.

### **WES data processing**

Reads sequencing quality were assessed using the *fastqc* tool version 0.11.5 (Babraham Institute). Reads alignment to the reference genome, somatic variant detection and copy number alteration analyses were performed using an internal pipeline. Basically, reads alignment were conducted using the human hg19 reference genome and the *BWA-MEM* tool version 0.7.15 (Li, 2013). Reads with alignment quality (MAPQ) lower than 20 as well as reads that are marked as duplicated by *Picard (markduplicate)* tool (Broad Institute, version 2.6.0) were excluded. In addition, only alignment intersecting the Agilent exome regions file “SureSelect\_Clinical\_Research\_Exome\_Regions.bed” were kept for further analysis. Local realignment around indels and base score recalibration were conducted with *GATK* framework version 4.0.21 (Broad Institute) using dbSNP database release 137 (Sherry et al., 2001) .

### **WES somatic variant calling**

Single nucleotide variants (SNVs) and small insertion-deletions (InDels) calling and annotation were respectively performed using *Mutect2 (GATK)* and *AnnoVar* release 2017-07-16 (Wang et al., 2010). Population-based annotations were performed using the *1000 genome* (Auton et al., 2015) release (all populations), Exome Sequencing Project (*ESP6500*) version 2 (all populations) and the Exome Aggregation Consortium (*ExAC*) all populations data base (Lek et al., 2016) while gene-based annotation was performed using the *RefSeq* gene annotation release oct. 2015. Variant filtering was applied as follow: SNVs and InDels with a depth of coverage lower than 10 or covered by less than 5 reads were filtered out as well as variants reported in more than 1% of the population based on the 1000 genomes, ESP6500 and ExAC annotations. Synonymous variants were also filtered out. Finally, we decided to keep only variants that are localized within exons and splice sites and with allelic frequency higher or equal to 10. Normal blood sample were analysed jointly with tumor sample for germline mutation filtering.

## **WES copy number alteration analysis**

Copy number alteration analysis (CNA) were conducted using *Facets* tool version 0.5.11 (Shen and Seshan, 2016) based on the hg19 single nucleotide polymorphisms (SNPs) VCF file downloaded from NCBI (november 2016). All SNPs covered by less than 25 reads and by higher than 3000 reads in the normal sample were filtered out. The segmentation value was set to 150. The *MAFtools* Bioconductor package version 1.8.10 (Mayakonda et al., 2018) was used to somatic variant exploration.

R version 3.5.0 was used for data analysis and plotting.

## **Gene expression and DNA methylation data integration**

Gene expression (RNA-seq) and DNA methylation (EPIC array) data integration were performed using the kernel-based method implemented in *mixKernel* version 0.7 CRAN package (Mariette and Villa-Vialaneix, 2018). 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 the 2,000 highly variable genes (based on IQR value) was used. For DNA methylation dataset, the beta-value matrix including the 5,000 highly variable genes (based on IQR value) was used. The kernel matrix of each datasets was computed using the *compute.kernel()* function using the “*linear*” kernel method. The two kernel matrices were combined using the *combine.kernels()* function with the “*full-UMKL*” method. 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 version 4.0.2 environment.

## **Single cell RNA sequencing (scRNA-seq) (Appendix Figure 2)**

### **Tissue processing and cell population enrichment**

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 mcg/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/μl. Tissues were processed within 2 hours after tumor resection and loaded in 10x Chromium instrument within 4 hours.

### **Single cell RNA sequencing**

Sample preparation 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 10000 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- 3min; 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 26x98 bp as sequencing mode, targeting at least 50 000 reads par cell.

### **scRNA-seq cell filtering**

Single sample adaptive filtering strategy was applied to remove "bad cells". The number of detected genes and the percentage of mitochondrial RNA were considered as filtering criteria. This strategy assumes that the sample has good quality with a majority of tumor cells. The observation of cells by microscopy during cell counting and before loading in 10x Chromium instrument support this hypothesis.

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 (Appendix Figure 3). 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. These cells are considered as damaged since their cytoplasmic membrane may be leaked leading to the loss of the cytoplasmic mRNA while the mitochondrial mRNA still retained in the mitochondria. The 3 samples are integrated since their respective 'bad cells' are separately removed.

### **scRNA-seq data integration**

scRNA-seq data from the 3 different samples were integrated using the approach described in (Stuart et al., 2019) and implemented in Seurat version 3. Briefly, the algorithm consists of performing pairwise canonical correlations analysis (CCA), finding anchors (which are set of cells that are considered to be biologically similar between the two datasets) in the CCA space using mutual nearest neighbors (MNNs) and then adjusting the two datasets based on these anchors. Anchors are scored using shared nearest neighbors (SNN) based approach.

The CCA and anchors finding were performed using the *FindIntegrationAnchors()* function of Seurat v3 with default settings. The data integration itself was performed using the *IntegrateData()* function of Seurat v3 with default settings. Data integration quality was assessed by plotting the cell cycle phases on the integrated data (Appendix Figure 4).

### **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., 2019b) was used to perform this multiple clustering and to pick the best combination of number of PCs and resolution. Combinations are assessed using gap statistic and the performance of differentially expressed genes to classify the cells into their respective cluster (Please refer to (Chen et al., 2019b) for more details).

In addition, iterative clustering was also applied by removing cells belonging to well defined cluster and re-running the clustering with the remaining cells (Appendix Figure 5). This allow to find the maximum possible clusters in our dataset.

All these analyses converge to the identification of 10 clusters that can be obtained using a 12 PCs and a resolution equals to 0.2. In other words, the clustering run using 12 PCs and resolution 0.2 recapitulate all the possible clusters identified throughout the analyses performed above.

UMAP was used to visualize the clustering result in reduced two dimensions. UMAP was conducted using the *RunUMAP()* function based on 12 PCs with default value for all other parameters.

### **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

In order to identify the specifically activated transcription factors (TFs) for each cluster, we conducted a gene regulatory network analysis using the SCENIC framework (Aibar et al., 2017). Basically, the analysis consists of 3 major steps. First, genes that are co-expressed with each TF are identified by regression-based network inference. Second, modules which are composed by a TF and their targets are created based on 3 criteria: the top number of targets for each TF, the top number of regulators for each target, and a percentile score of targets. The third step consists of keeping only, for each module, the direct targets using a motif enrichment analysis resulting to a list of regulons which are a TFs and their direct targets. For more details about these steps, please refer to Sande et al., 2020 (Van de Sande et al., 2020).

The *pyscenic* program version 0.10.3 was used to conduct the analysis. First, a *loom* file containing the raw counts of all 3 samples were generated. The *arboreto\_with\_multiprocessing.py* script (available in <https://github.com/aertslab/pySCENIC/blob/master/src/pyscenic/cli/>) with GRNboost2 algorithm (Moerman et al., 2019) were used to launch the first step of SCENIC analysis which consists of the identification of TF-targets interactions based on co-expression and to create modules. The TFs list (containing 1,839 human TFs) downloaded from <https://github.com/aertslab/pySCENIC/blob/master/resources/> on August 7th, 2020 were used. To select the direct targets in each modules and to create regulons, the motif enrichment analysis was performed using the *ctx* subcommand of *pyscenic*. The *cisTarget* database (<https://resources.aertslab.org/cistarget/>) and the human motif annotation table downloaded from <https://ressources.aertslab.org/cistarget/motif2tf/> on August 12th 2020 were used for this analysis. To estimate the activity of a specific regulon in each cell, an area under the curve (AUC) based score was computed using the *aucell* subcommand of *pyscenic*. To assess the specificity of a regulon in a given cluster, a regulon specificity score (RSS) was computed for each cell cluster and for each regulon using the *regulon\_specificity\_scores()* function of *pyscenic*.

Python version 3.6.11 was used for the analysis.

## scRNA-seq trajectory inference (TI) and pseudotime analysis

Trajectory inference analyses were conducted using reverse graph embedding approaches. For the sake of confirmation, two algorithms were used: 1) one implemented in the ePiGraph version 1.0.0 (Albergante et al., 2020) R package and based on the elastic principal graph

(Gorban and Zinovyev, 2005) and 2) the other implemented in Monocle3 version 0.2.3.0 (Cao et al., 2019) R package and based on the Partition-based graph abstraction (PAGA)(Wolf et al., 2019).

For the elPiGraph 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, a cds object was first built using the *new\_cell\_data\_set()* function. Data were then processed with *preprocess\_cds()* function using PCA and 12 number of PCs. UMAP dimensionality reduction was then applied with *reduce\_dimension()* function. Cell clustering was performed using *cluster\_cells()* function before inferring graph with *learn\_graph()* function.

TI analysis 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., 2020) tool using the dynamical modelling. Spliced and unspliced matrices were generated separately for each sample in loom file using the velocito tool version 0.17.17 (La Manno et al., 2018). The 3 loom files corresponding to the 3 samples were imported in python as anndata object and combined into a single anndata object using the scanpy python library version 1.6.0 *anndata.concat()* function. The proportion of spliced and unspliced transcripts are visualized using the *pl.proportions()* method of scVelo. Gene filtering and data normalization were performed using respectively *pp.filter\_genes()* and *pp.normalize\_per\_cell()* methods. RNA velocity was estimated using *tl.velocity()* function with 'dynamical' mode. The embedding UMAP coordinates generated by Monocle3 was used to visualize the result.

RNA velocity analysis was performed inside python version 3.7.8 environment.

### **scRNA-seq ligand-receptor interaction analysis**

Ligand-receptor interaction analysis was launched using the cellPhoneDB framework consisting of a database and a statistical tool (Efremova et al., 2020). CellPhoneDB database includes ligands, receptors and their interactions which take into account the subunit architecture of the ligands and the receptors. The framework works at cluster level. The average

expression of each ligand and each receptor were computed for each cluster. An average expression for a specific ligand-receptor interaction between two clusters correspond to the lowest average of the couple ligand receptor. A p-value is computed using the null-distribution of mean generated using clusters of cells assigned randomly. To conduct this ligand-receptor interaction analysis, the *cellphonedb* binary (version 2.1.7) was run on the integrated expression matrix including all genes. The resulting table (mean and p-value for each ligand-receptor and for each couple of clusters) was imported in R for filtering and visualization. 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.

# Discussion

### **Integrative anatomical-molecular analysis**

To date, different studies in human have attempted to correlate ATRT molecular subgroups with predominant anatomical location but they roughly described ATRT location as supra or infratentorial, failing to provide precise information about the putative tissue of origin. We assumed that the precise anatomical description could improve the understanding of ATRT development. Consequently, we provide for the first time so far, a detailed description of different anatomical locations and performed integrative analysis considering this new layer of information along with molecular data. Importantly, our analysis showed 4 distinct anatomico-molecular subgroups that allow us to further investigate the putative cells of origin. To note, the existence of subclusters within ATRT-SHH (by DNA methylation) (Ho et al., 2020; Johann et al., 2016) and within ATRT-MYC (by transcriptomics) (Chun et al., 2019) subgroups has been previously suggested. Interestingly, Johann et al. showed subclusters within ATRT-SHH subgroup corresponding to supratentorial location and an infratentorial location while Chun et al. observed after NMF analysis two further subgroups within Group 1 tumors (MYC-ATRT). However, both failed to demonstrate molecular or clinical evidence supporting the existence of further biologically relevant subgroups. Our study demonstrates the existence of a biologically consistent new ATRT subgroup, that is, interestingly, close to the ATRT-SHH at DNA methylation level and close to the ATRT-MYC at transcriptome level. Furthermore, we show that, in addition to its specific location, the immune signatures present in this subgroup could be the reason to shift to ATRT-MYC at transcriptome level whereas at DNA methylation level their neuronal lineage brings them in the ATRT-SHH subgroup.

### **CNCS-ATRT and NCC**

We demonstrate that a subset of CNCS-ATRT (MYC/MYC) showed extra-axial locations with some of them presenting typical location of schwannomas and identified two cases with NF2 mutation in this subgroup. ATRT arising from different cranial nerves have been previously reported in children and young adults (Beschoner et al., 2006; Oh et al., 2015; Rizzo et al., 2012; Verma and Morriss, 2008; Wang et al., 2015; Wu, 2020; Wykoff et al., 2008). These tumors were often misdiagnosed as schwannoma, due to their typical location. In line with our observations, others authors have reported ATRT located in the cranial nerves showing both neurinoma and rhabdoid components (Bertrand et al., 2018; Rizzo et al., 2012; Rossi et al., 2018). This led us to hypothesize that these tumors originate from a « rhabdoid » transformation of an initially benign nervous tumor. Therefore, we investigated whether the NF2 and Hippo pathway might be significantly involved CNCS-ATRT (data not shown) but our results showed

that this relationship is anecdotal. Vitte et al. investigated the mechanisms by which loss of *Smadcb1* lead to schwannomas versus RT by inducing *Smadcb1* and *NF2* deletion in NCC and showed that *Smadcb1* alone is not sufficient for the occurrence of schwannomas, but requires the additional bi-allelic loss of *NF2*. However, the possibility of a schwannoma caused by loss of *NF2* being secondarily transformed into RT by loss of *SMARCB1* remains to be experimentally demonstrated.

Schwannomas and meningiomas, frequently associated with *NF2*, originate from tissues derived from the NCC. The NCC are a transient multipotent cell population, highly migratory, that originate in the dorsal part of the neural tube. Depending on their axial level of origin and migratory pathways, NCC adopt different fates and contribute to various tissues and organs (PNS, bone and cartilage, meninges). The cranial nerves are derived from the cranial NCC. On the other hand, schwannomatosis, a condition that can be caused by germline mutations in *SMARCB1* or *LTZR1*, lead also to schwannomas located in the cranial nerves but much less frequently than in *NF2*. In addition, the involvement of cranial nerves (CNVIII) occurs in *LZTR1*-related schwannomatosis and is anecdotal in *SMARCB1*-related schwannomatosis. Of note, germline *SMARCB1* mutations in patients with rhabdoid tumors, compared with *SMARCB1*-related schwannomatosis, are more likely truncating, centrally placed and involving multiple exons deletions (Smith et al., 2014).

Recent work supports the origin of rhabdoid tumors in NCC (Custers et al., 2021; Vitte et al., 2017). Interestingly, by targeting neural crest cells, the P0-CreC:*Smadcb1*F/F mouse (Vitte et al., 2017) developed tumors in the cranial nerves and meninges, recapitulating the phenotype of CNCS-ATRT subgroup but failed to obtain ATRT at other locations. Furthermore, this model correlated at transcriptomic level with our CNCS-ATRT (MYC/MYC) subgroup, supporting an origin in the NCC for this subgroup. However, these authors argued that this model give rise to RTs recapitulating all the ATRT molecular subgroups, which is not in full agreement with our results. The correlation of P0-CreC:*Smadcb1*F/F with human samples showed that this model is much less correlated with the other anatomical-molecular ATRT subgroups, suggesting that they are less probably related with NCC lineage. In the other hand, Jarno et al., by comparison of somatic mutations shared by the tumor and the surrounding normal tissues, showed that RT are phylogenetically related to neural crest derived tissues but their work is based on 2 ECRT samples only, and didn't include ATRT. Nevertheless, this result reinforces the hypothesis of NCC as lineage of origin for CNCS-ATRT (MYC/MYC) since it has been demonstrated similarities between ECRT and MYC -ATRT (Chun et al., 2019) as well as in our RTs cohort (data not shown).

Finally, we could hypothesize that in our Rosa26 mouse model we targeted earlier NCC as P0-Cre: Smarcb1 F/F model since Rosa26 locus is ubiquitously expressed and we obtained MYC tumors when we inactivated Smarcb1 from E6 to E10. These tumors were located in the meningeal region and were well correlated with CNCS-ATRT (MYC/MYC), reinforcing once again the extraparenchymal origin of these tumors.

### **BG/IV-ATRT and the GE origin**

BG/IV-ATRT (SHH/MYC) subgroup showed a particular location intricately linked to the subventricular zone and the basal ganglia. Interestingly, the differential expression analysis brings out relevant genes involved in forebrain development and more precisely, the ganglionic eminence, an embryologically consistent link with the location of these tumors. Although the neuronal origin of these tumors has been previously suggested, our analysis integrating anatomical localization and molecular profiling allows us to refine the characterization of these tumors to differentiate them from ATRTs located in the upper cerebellum (CAL-ATRT) and to bring to light the overexpression of genes characteristic of the ganglionic eminence, pointing for the first time to this area as the origin of these tumors. Moreover, the anatomical correlation with Rosa26: SHH tumors, which also express GE markers, encourages us to further develop a model targeting NSC similarly as those developed for other embryonal tumors (Neumann et al., 2017), but attempting to specifically target the neuronal GE progenitors. In addition, BG/IV-ATRT (SHH/MYC) showed inflammatory immune response and this response seems to be different from the CNCS-ATRT (MYC/MYC) subgroup. Previously, it has been reported that ATRT showed different immune response according to the molecular subgroup, with MYC-ATRT subgroup showing the higher immune infiltration scores (Chun et al., 2019; Leruste et al., 2019). However, these results are partially discordant in terms of different immune response according to ATRT-subgroup with recently published work, especially for ATRT-SHH (Melcher et al., 2020). The comparison of the results leads us to suspect that the discordances may be resolve in the light of our study since the molecular subgroup assignment was done by RNAseq for the first studies and by DNA methylation in the last one. Finally, our results also stress the importance of immune response to characterize ATRT subgroups.

### **CAL-ATRT: from anatomical location to single cell interaction**

The newly described anatomical-molecular classification allow us to identify the CAL-ATRT (SHH/SHH) subgroup and lead us to perform single cell analysis of three human samples coming from this location. Up to now, current knowledge points out that pediatric brain tumors are caused by genetic alterations that impair normal differentiation programs and retain cells in a self-renewing, progenitor-like phenotype (Jessa et al., 2019). Our results show that the opposite is possible, i.e. that cells with some degree of commitment to neuronal differentiation undergo a neuronal repression and acquire stem cell properties. The fact that tumor cells can originate by dedifferentiation from differentiated cells types has been already described (Cobaleda et al., 2007; Friedmann-Morvinski et al., 2012; Liu et al., 2011; Schwitalla et al., 2013). Furthermore, epigenetic dysregulation and dedifferentiation has been described as mechanisms involved in cancer development (Friedmann-Morvinski and Verma, 2014; Nakano et al., 2019; Yamada et al., 2014). We identified a path from NPL1 to UD cells that was characterized by the loss of expression of neurogenic transcription factors and neuronal differentiation genes and the progressive increase expression of stem cell and pluripotency genes markers. Previous studies in glioblastoma have observed similar kinetic expression, consisting in the loss of expression of differentiation marker genes (Tuj1 and GFAP) and progressive increase expression of stem cell markers (Nestin, Sox2) (Friedmann-Morvinski et al., 2012). Interestingly, we found that UD clusters, the core of tumor cells, were characterized by the activity of specific TF related to epithelia-mesenchymal transition (EMT) and mesodermal commitment pathway, stem cell and pluripotency marker. The EMT generates cells with stem cell properties and shares common transcription players in tumor dedifferentiation (Mani et al., 2008). Interestingly, RT are described by pathologist as undifferentiated tumor and the “rhabdoid morphology” appear in a large group of unrelated neoplasms suggesting that this morphology is a phenotypic de-differentiation end point of tumors (Morgan et al., 2000). Finally, this dedifferentiated phenotype has been related to SMARCB1 loss in multiple cancer types (Agaimy et al., 2015; Agaimy et al., 2016a; Agaimy et al., 2016b; Agaimy et al., 2017; Karnezis et al., 2016). All these results are perfectly consistent with our results showing the dedifferentiation of NPL1 cluster into UD clusters.

Finally, our study shows a possible role of NOTCH signaling in this process. NOTCH signaling is involved in many biological processes including cell proliferation, cell differentiation, neuronal development, apoptosis as well as in tumor development (Bazzoni and Bentivegna, 2019; Parmigiani et al., 2020; Siebel and Lendahl, 2017; Teodorczyk and Schmidt, 2014; Than-Trong et al., 2018) . Although the involvement of the NOTCH pathway has been previously

described in the ATRT (Ho et al., 2020; Johann et al., 2016; Torchia et al., 2016), to our knowledge we are the first to demonstrate it at the single cell level highlighting its intra-tumoral molecular mechanism. This finding should be confirmed by in-vivo experiment by, for example, inactivating NOTCH pathway in ATRT cell lines. Targeting NOTCH signaling could be therefore a potential treatment at least for this subgroup of ATRT. Furthermore, we point out the involvement of WNT, FGFR and EPHB pathways that could be also important for future therapies. Understanding the interplay between these signaling pathways may be a major clue to apprehend ATRT development.

# Perspectives

## **The Nestin Mouse Model**

We have identified a subgroup of human ATRT located in the basal ganglia and intraventricular region (BG/IV-ATRT (SHH/MYC)), that reminds Rosa26-Neuronal (//SHH) mouse ATRT resulting from E6-E7 *Smarb1* inactivation. These tumors developed from brain tissue close to SVZ. We aim to develop a genuine mouse model of ATRT that reproduces as closely as possible human ATRT from this region. This new model should overcome also the problems of the Rosa26CreERT2: *Smarb1*<sup>fl<sub>ox</sub>/fl<sub>ox</sub></sup> model, i.e. the inability to control the ATRT subtype. For these reasons members of my team worked to develop a new model, the *Smarb1*<sup>F/F</sup>; NestinCre mouse model, based on the inactivation of *Smarb1* in cells expressing Nestin, a neural stem cell marker. Of note, Nestin is expressed early in development (E7) in several proliferative zone of CNS in mice.

During the development of this model, we observed that *Smarb1*<sup>F/F</sup>; NestinCre<sup>+</sup> mice died shortly after birth (Figure 11A) and were ingested by the mother. When we performed a caesarean at E18 we observed the expected ratio in all litters (Figure 11B). Careful examination of embryo brains at E18 showed massive brain malformation, consisting in a severe hypoplasia of the cortex and disorganized layers. No tumor was observed but we identified regions of cell hyperproliferations, restricted to the ventral SVZ, at the level of the ganglionic eminences, and in the hypothalamus (Figure 11 C, D). These pups died and had no time to develop tumors. To circumvent the early death of these pups, we performed heterotopic engraftment of pup brains in the flank of syngenic mice. This leads to *Smarb1* deficient tumors with a clear rhabdoid phenotype. We have therefore hypothesized that actively proliferating cells observed may be the tumor initiating cells. In order to better characterize these candidate cells-of-origin, we have planned different approaches: ex-vivo characterization of brain and single cell RNAseq analysis.

### **Ex-vivo characterization of *Smarb1*<sup>F/F</sup>; NestinCre<sup>+</sup> E18 brains**

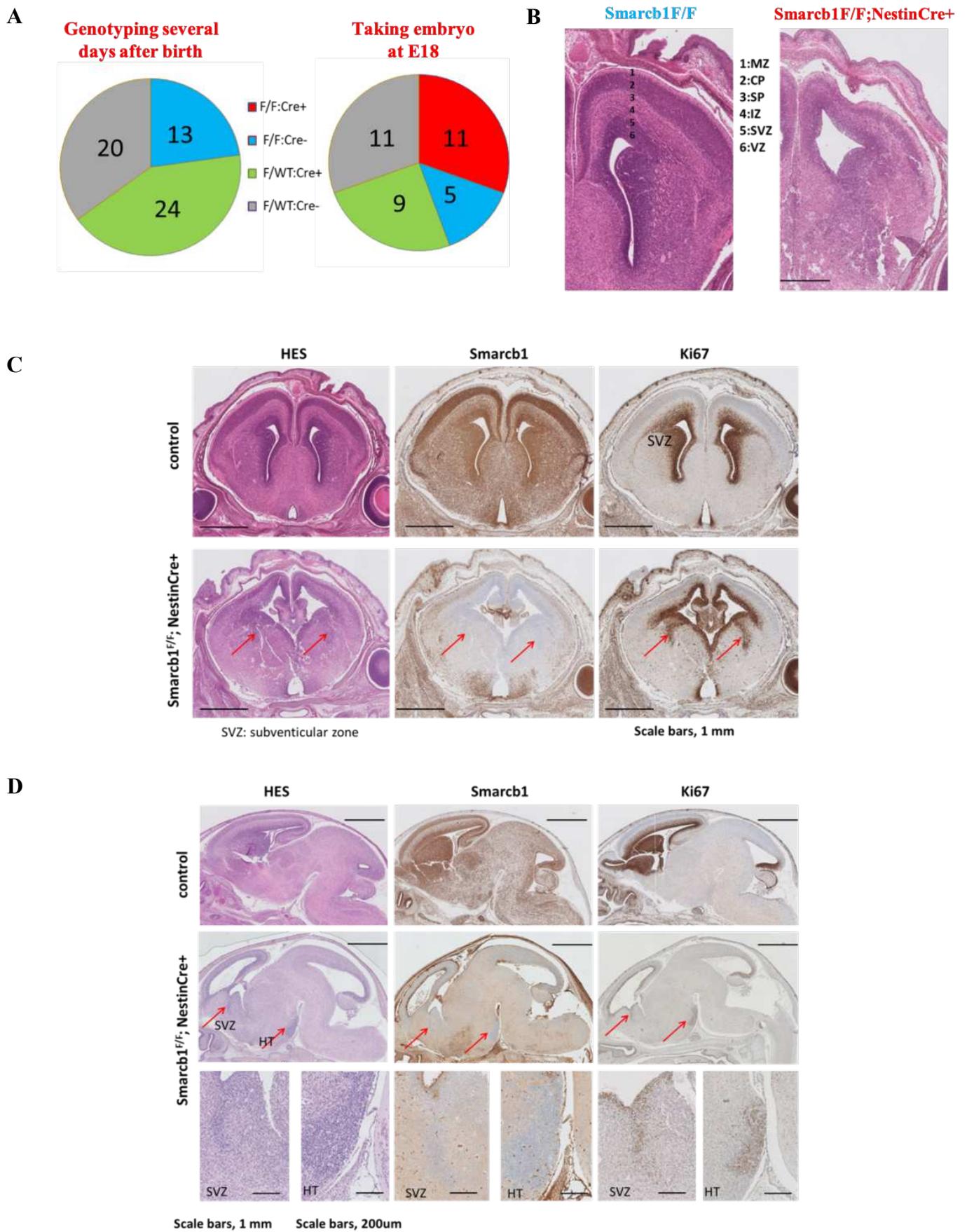
We have already performed brain dissection of *Smarb1*<sup>F/F</sup>; NestinCre<sup>+</sup> brains at 3 different gestational ages (E14, E16, E18). Since we observed no macroscopic differences between mutants and controls at E14 and E16, we focused on E18. We have generated multiples brain slides and immunofluorescence staining were performed (Sox 2, early progenitor marker and pH3, proliferative cell marker). We observed an enrichment of Sox2<sup>+</sup> progenitors in the ganglionic eminence with high proliferation, together with a high proliferation rate assessed by pH3 (Figure 12). Interestingly, we have observed a decrease of Sox2<sup>+</sup> progenitor cells in the

cortex, that could be due to cells death, rapidly differentiating progenitors, or more likely, a lack of cells derived from ganglionic eminence neural progenitors.

### **Single Cell RNAseq analysis of Nestin model E18 brains**

We have started by performing single cell RNAseq with Chromium 10x of the whole brain of Smarcb1 F/F (n=2) and Smarcb1F/F; NestinCre+ (n=2) in order to investigate how the loss of Smarcb1 impact cellular diversity in embryonic brain development. Single cell analysis of individual and merged samples has allowed us to identify different neuronal and non-neuronal populations that were present in both control and mutant samples; however, we failed to identify cells suspected to be the cells of origin (data not shown). The main limitation of this preliminary approach was the lack of specificity since whole brain is included and the number of samples. To overcome these limitations, we have planned to increase the number of samples and to enrich on cell of interest by dissecting the area of interest from brain slides. These analyses are ongoing.

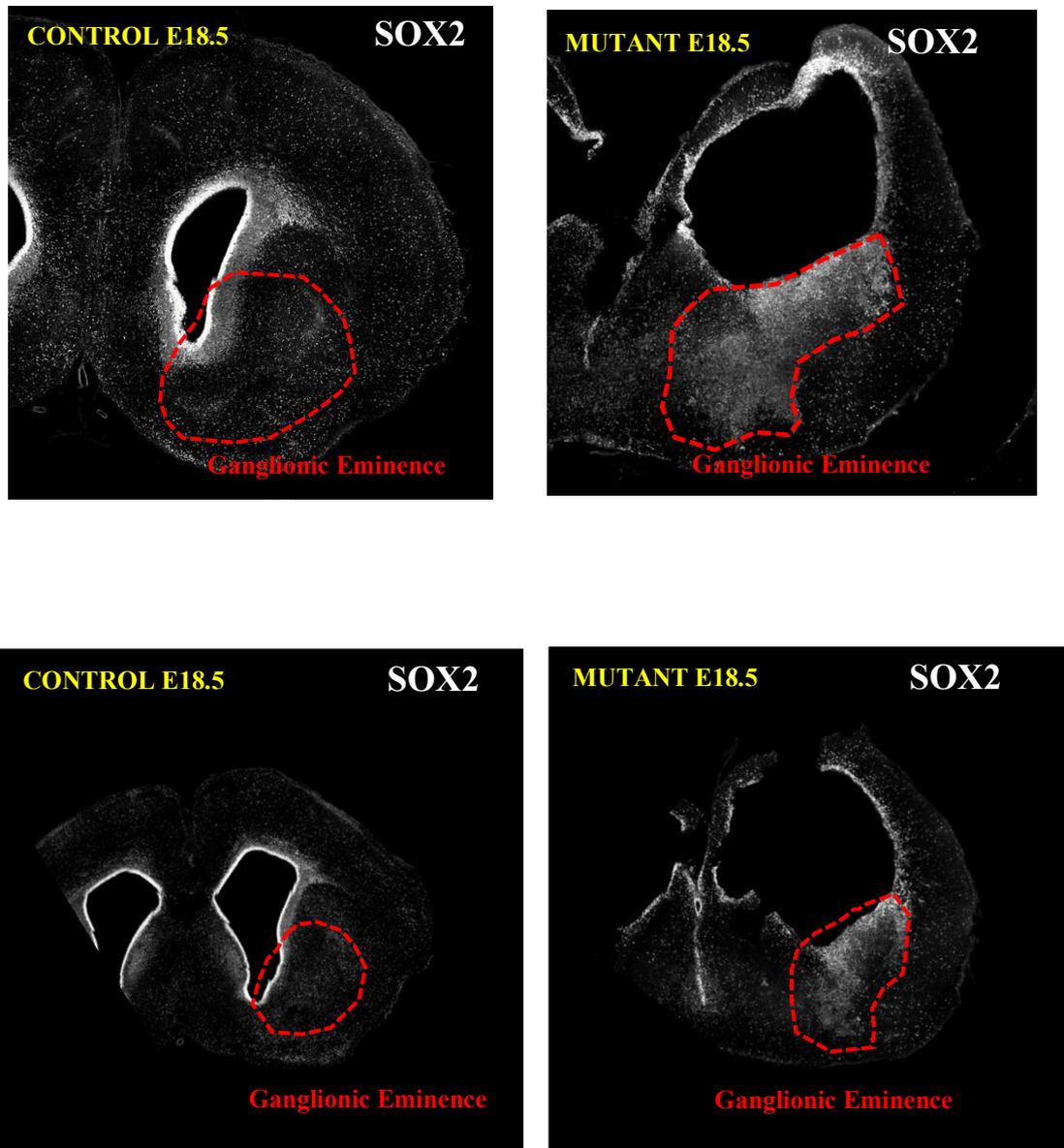
**Figure 11.**



**Figure 11. Smarcb1 inactivation in neural stem cells generates foci of hyper proliferation in highly restricted areas of neural stem cell niches**

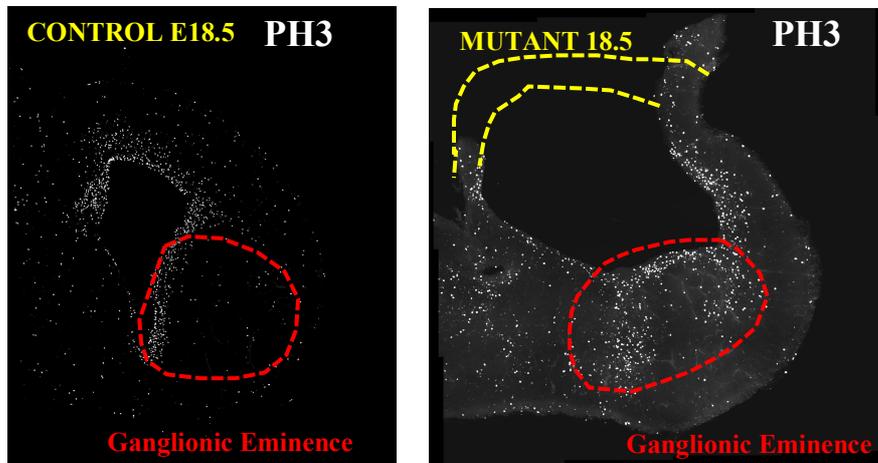
- A. Ratio of the various genotypes in newborn mice (left circle) and at E18 (mendelian ratio, right circle).
- B. Severe cortical hypoplasia with disorganized layers (normal control at the left, Nestin-Cre:Smarcb1 flox/flox brains at the right).
- C. Coronal sections of normal brain (first line) and Nestin-Cre:Smarcb1 flox/flox brains (second line); HES on the left column, Baf47 in the middle and ki67 at the right; arrows indicate the abnormal proliferating zones.
- D. Sagittal sections of normal brain (first line) and Nestin-Cre:Smarcb1 flox/flox brains (second line and third line); HES on the left column, Baf47 in the middle and ki67 at the right ; arrows indicate the abnormal proliferating zones; third line: zoom on the proliferating zones in the sub ventricular region and in the hypothalamic niche.

Figure 12A. Ex-vivo characterization of Nestin mice brains.



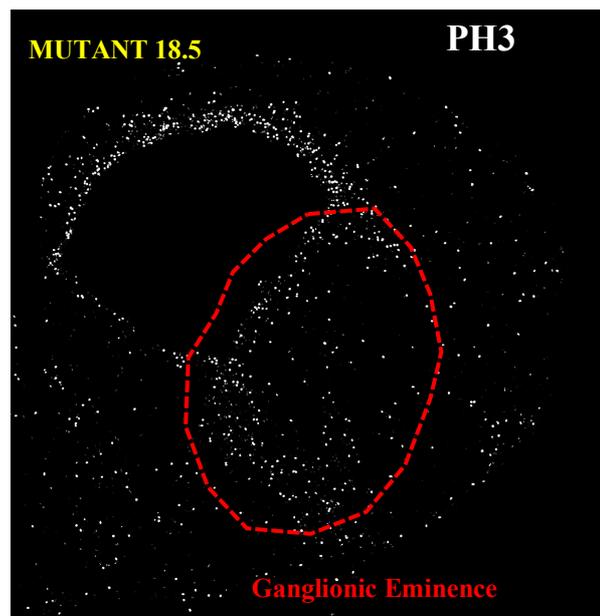
Immunofluorescence staining brain at E18. The red dash line highlights the ganglionic eminence region. In the mutant, the SOX2-positive cells show an abnormal distribution pattern beyond the immediate SVZ (subventricular zone).

**Figure 12B. Ex-vivo characterization of Nestin mice brains**

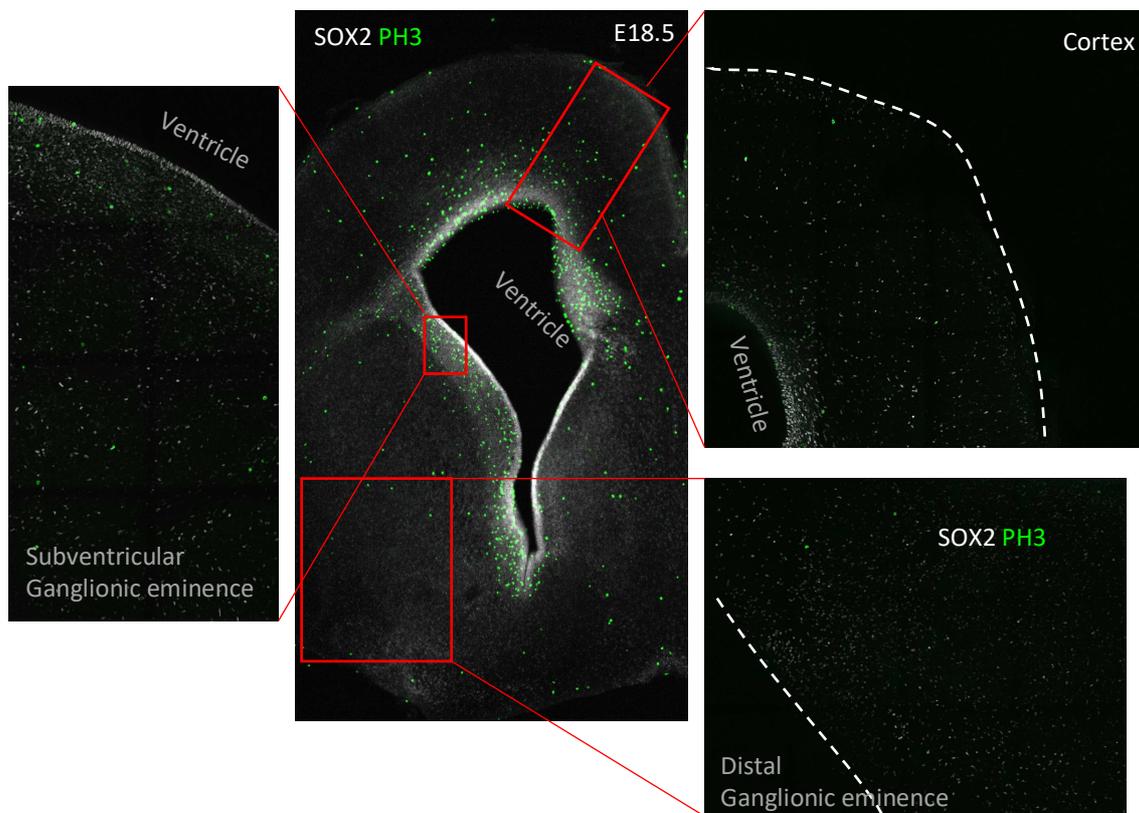


Immunofluorescence staining brain at E18. The red dash line highlights the ganglionic eminence region. The mutant brain is broken because of its fragility; the yellow dashed line shows the silhouette before the breakage. In the control, PH3-positive cells (white) are mostly in the immediate subventricular zone. In the mutant, PH3-positive cells go beyond the immediate subventricular zone and extend throughout the ganglionic eminence.

Mutant brain slide showing PH3-positive cells (white) with abnormal distribution in the ganglionic eminence region.

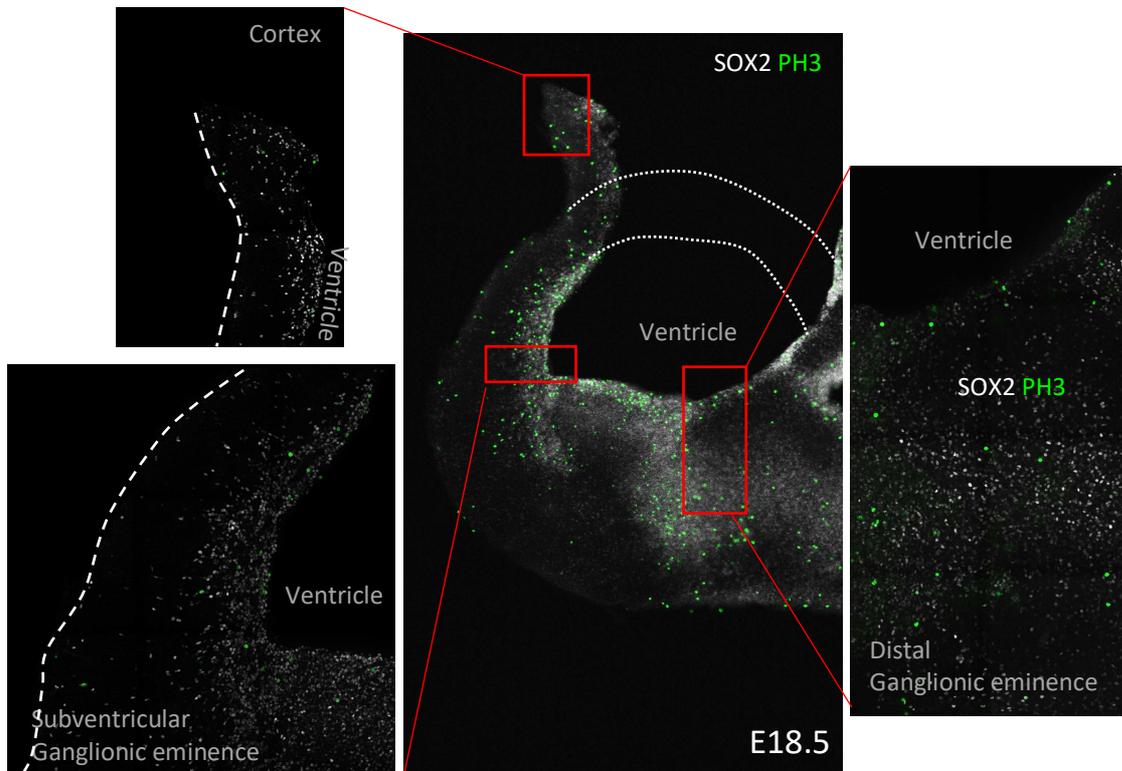


**Figure 12C. Ex-vivo characterization of Nestin mice brains: CONTROL (in detail).**



Immunofluorescence staining brain at E18. In the control, SOX2 (white) and PH3-positive cells (green) are mostly in the immediate subventricular zone.

**Figure 12D. Ex-vivo characterization of Nestin mice brains: MUTANT (in detail).**



Immunofluorescence staining brain at E18. This figure shows an abnormal distribution of SOX2/PH3 positive cells in the ganglionic eminence region of the mutant. The SOX2 (white) and PH3-positive (green) cells go beyond the immediate subventricular zone and extend throughout the ganglionic eminence.

## Bibliography

- Agaimy, A., Haller, F., Frohnauer, J., Schaefer, I. M., Ströbel, P., Hartmann, A., Stoehr, R. and Klöppel, G.** (2015). Pancreatic undifferentiated rhabdoid carcinoma: KRAS alterations and SMARCB1 expression status define two subtypes. *Mod. Pathol.* **28**, 248–260.
- Agaimy, A., Bertz, S., Cheng, L., Hes, O., Junker, K., Keck, B., Lopez-Beltran, A., Stockle, M., Willich, B. and Hartmann, A.** (2016a). Loss of expression of the SWI/SNF complex is a frequent event in undifferentiated/dedifferentiated urothelial carcinoma of the urinary tract. *Virchows Arch.* **469**, 321–330.
- Agaimy, A., Daum, O., Bruno, M., Lichtmanegger, I., Michal, M. and Hartmann, A.** (2016b). SWI/SNF Complex-deficient Undifferentiated/Rhabdoid Carcinomas of the Gastrointestinal Tract. *Am J Surg Pathol* **40**, 544–553.
- Agaimy, A., Cheng, L., Egevad, L., Feyerabend, B., Hes, O., Keck, B., Pizzolitto, S., Sioletic, S., Wullich, B. and Hartmann, A.** (2017). Rhabdoid and Undifferentiated Phenotype in Renal Cell Carcinoma. *Am. J. Surg. Pathol.* **41**, 253–262.
- Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J. C., Geurts, P., Aerts, J., et al.** (2017). SCENIC: Single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086.
- Albergante, L., Mirkes, E., Bac, J., Chen, H., Martin, A., Faure, L., Barillot, E., Pinello, L., Gorban, A. and Zinovyev, A.** (2020). Robust and scalable learning of complex intrinsic dataset geometry via EIPiGraph. *Entropy* **22**,.
- Assenov, Y., Müller, F., Lutsik, P., Walter, J., Lengauer, T. and Bock, C.** (2014). Comprehensive analysis of DNA methylation data with RnBeads. *Nat. Methods* **11**, 1138–1140.
- Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., et al.** (2015). A global reference for human genetic variation. *Nature* **526**, 68–74.
- Azim, E., Jabaudon, D., Fame, R. M. and MacKlis, J. D.** (2009). SOX6 controls dorsal progenitor identity and interneuron diversity during neocortical development. *Nat. Neurosci.* **12**, 1238–1247.
- Azzarelli, R., Simons, B. D. and Philpott, A.** (2018). The developmental origin of brain

- tumours: a cellular and molecular framework. *Development* **145**, dev162693.
- Bachmann, C., Nguyen, H., Rosenbusch, J., Pham, L., Rabe, T., Patwa, M., Sokpor, G., Seong, R. H., Ashery-Padan, R., Mansouri, A., et al.** (2016). mSWI/SNF (BAF) Complexes Are Indispensable for the Neurogenesis and Development of Embryonic Olfactory Epithelium. *PLoS Genet.* **12**, 1–29.
- Baker, S. J., Ellison, D. W. and Gutmann, D. H.** (2016). Pediatric gliomas as neurodevelopmental disorders. *Glia* **64**, 879–895.
- Bartelheim, K., Nemes, K., Seeringer, A., Kerl, K., Buechner, J., Boos, J., Graf, N., Dürken, M., Gerss, J., Hasselblatt, M., et al.** (2016). Improved 6-year overall survival in AT/RT – results of the registry study Rhabdoid 2007. *Cancer Med.* **5**, 1765–1775.
- Bazzoni, R. and Bentivegna, A.** (2019). Role of Notch Signaling Pathway in Glioblastoma Pathogenesis. *Cancers (Basel)*. **11**, 1–25.
- Bergen, V., Lange, M., Peidli, S., Wolf, F. A. and Theis, F. J.** (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414.
- Bertrand, A., Rondenot, C., Masliah-Planchon, J., Leblond, P., de la Fourchardière, A., Pissaloux, D., Aït-Raïs, K., Lequin, D., Jouvét, A., Freneaux, P., et al.** (2018). Rhabdoid component emerging as a subclonal evolution of paediatric glioneuronal tumours. *Neuropathol. Appl. Neurobiol.* **44**, 224–228.
- Beschorner, R., Mittelbronn, M., Koerbel, A., Ernemann, U., Thal, D. R., Scheel-Walter, H. G., Meyermann, R. and Tatagiba, M.** (2006). Atypical teratoid-rhabdoid tumor spreading along the trigeminal nerve. *Pediatr. Neurosurg.* **42**, 258–263.
- Biegel, J. A., Zhou, J. Y., Rorke, L. B., Stenstrom, C., Wainwright, L. M. and Fogelgren, B.** (1999). Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. *Cancer Res.* **59**, 74–79.
- Brinkman, T. M., Ness, K. K., Li, Z., Huang, I. C., Krull, K. R., Gajjar, A., Merchant, T. E., Klosky, J. L., Partin, R. E., Olsson, I. T., et al.** (2018). Attainment of functional and social independence in adult survivors of pediatric CNS tumors: A report from the st jude lifetime cohort study. *J. Clin. Oncol.* **36**, 2762–2769.
- Brownlee, P. M., Meisenberg, C. and Downs, J. A.** (2015). The SWI/SNF chromatin remodelling complex: Its role in maintaining genome stability and preventing tumorigenesis. *DNA Repair (Amst)*. **32**, 127–133.
- Bruschi, M., Grill, J. and Guerrini-Rousseau, L.** (2021). Pediatric brain tumors as a developmental disease. *Curr. Opin. Oncol. Publish Ah*, 1–7.

- Buckles, G. R., Thorpe, C. J., Ramel, M. C. and Lekven, A. C.** (2004). Combinatorial Wnt control of zebrafish midbrain-hindbrain boundary formation. *Mech. Dev.* **121**, 437–447.
- Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., et al.** (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol. Cell* **6**, 1287–1295.
- Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., et al.** (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502.
- Capper, D., Jones, D. T. W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D. E., et al.** (2018). DNA methylation-based classification of central nervous system tumours. *Nature* **555**, 469–474.
- Cenik, B. K. and Shilatifard, A.** (2021). COMPASS and SWI/SNF complexes in development and disease. *Nat. Rev. Genet.* **22**, 38–58.
- Chapman, H., Waclaw, R. R., Pei, Z., Nakafuku, M. and Campbell, K.** (2012). The homeobox gene *Gsx2* controls the timing of oligodendroglial fate specification in mouse lateral ganglionic eminence progenitors. *Dev.* **140**, 2289–2298.
- Chen, Y. J. J., Vogt, D., Wang, Y., Visel, A., Silberberg, S. N., Nicholas, C. R., Danjo, T., Pollack, J. L., Pennacchio, L. A., Anderson, S., et al.** (2013). Use of “MGE Enhancers” for Labeling and Selection of Embryonic Stem Cell-Derived Medial Ganglionic Eminence (MGE) Progenitors and Neurons. *PLoS One* **8**,
- Chen, Y., Zhang, H., Xu, Z., Tang, H., Geng, A., Cai, B., Su, T., Shi, J., Jiang, C., Tian, X., et al.** (2019a). A PARP1-BRG1-SIRT1 axis promotes HR repair by reducing nucleosome density at DNA damage sites. *Nucleic Acids Res.* **47**, 8563–8580.
- Chen, Y. C., Suresh, A., Underbayev, C., Sun, C., Singh, K., Seifuddin, F., Wiestner, A. and Pirooznia, M.** (2019b). IKAP-Identifying K mAjor cell Population groups in single-cell RNA-sequencing analysis. *Gigascience* **8**, 1–8.
- Cheng, S. W. G., Davies, K. P., Yung, E., Beltran, R. J., Yu, J. and Kalpana, G. V.** (1999). c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. *Nat. Genet.* **22**, 102–105.
- Chi, S. N., Zimmerman, M. A., Yao, X., Cohen, K. J., Burger, P., Biegel, J. A., Rorke-Adams, L. B., Fisher, M. J., Janss, A., Mazewski, C., et al.** (2009). Intensive multimodality Treatment for children with newly diagnosed CNS atypical teratoid rhabdoid tumor. *J. Clin. Oncol.* **27**, 385–389.

- Chun, H.-J. E., Johann, P. D., Milne, K., Zapatka, M., Buellesbach, A., Ishaque, N., Iskar, M., Erkek, S., Wei, L., Tessier-Cloutier, B., et al.** (2019). Identification and Analyses of Extra-Cranial and Cranial Rhabdoid Tumor Molecular Subgroups Reveal Tumors with Cytotoxic T Cell Infiltration. *Cell Rep.* 1–17.
- Cobaleda, C., Jochum, W. and Busslinger, M.** (2007). Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* **449**, 473–477.
- Curry, R. N. and Glasgow, S. M.** (2021). The Role of Neurodevelopmental Pathways in Brain Tumors. *Front. Cell Dev. Biol.* **9**, 1–27.
- Custers, L., Khabirova, E., Coorens, T. H. H., Oliver, T. R. W., Calandrini, C., Young, M. D., Braga, F. A. V., Ellis, P., Mamanova, L., Segers, H., et al.** (2021). Somatic mutations and single-cell transcriptomes reveal the root of malignant rhabdoid tumours. *Nat. Commun.* 1–11.
- De Mori, R., Severino, M., Mancardi, M. M., Anello, D., Tardivo, S., Biagini, T., Capra, V., Casella, A., Cereda, C., Copeland, B. R., et al.** (2019). Agenesis of the putamen and globus pallidus caused by recessive mutations in the homeobox gene *GSX2*. *Brain* **142**, 2965–2978.
- Del Baldo, G., Carta, R., Alessi, I., Merli, P., Agolini, E., Rinelli, M., Boccuto, L., Milano, G. M., Serra, A., Carai, A., et al.** (2021). Rhabdoid Tumor Predisposition Syndrome: From Clinical Suspicion to General Management. *Front. Oncol.* **11**, 1–7.
- Dennis, D. J., Han, S. and Schuurmans, C.** (2019). bHLH transcription factors in neural development, disease, and reprogramming. *Brain Res.* **1705**, 48–65.
- Dho, Y. S., Kim, S. K., Cheon, J. E., Park, S. H., Wang, K. C., Lee, J. Y. and Phi, J. H.** (2015). Investigation of the location of atypical teratoid/rhabdoid tumor. *Child's Nerv. Syst.* **31**, 1305–1311.
- Di Giovannantonio, L. G., Salvio, M. Di, Omodei, D., Prakash, N., Wurst, W., Pierani, A., Acampora, D. and Simeone, A.** (2014). *Otx2* cell-autonomously determines dorsal mesencephalon versus cerebellum fate independently of isthmic organizing activity. *Development* **141**, 377–388.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R.** (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.
- Donkelaar** (2015). Development of the Basal Ganglia and the Basal Forebrain. *Brain Mapp. An Encycl. Ref.* **2**, 357–365.
- Efremova, M., Vento-Tormo, M., Teichmann, S. A. and Vento-Tormo, R.** (2020).

- CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexes. *Nat. Protoc.* **15**, 1484–1506.
- Eze, U. C., Bhaduri, A., Haeussler, M., Nowakowski, T. J. and Kriegstein, A. R.** (2021). Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat. Neurosci.* **24**, 584–594.
- Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., Krogsdam, A., Loncova, Z., Posch, W., Wilflingseder, D., et al.** (2019). Correction to: Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data (Genome Medicine (2019) 11 (34) DOI: 10.1186/s13073-019-0638-6). *Genome Med.* **11**, 1–20.
- Fossey, M., Li, H., Afzal, S., Carret, A. S., Eisenstat, D. D., Fleming, A., Hukin, J., Hawkins, C., Jabado, N., Johnston, D., et al.** (2017). Atypical teratoid rhabdoid tumor in the first year of life: the Canadian ATRT registry experience and review of the literature. *J. Neurooncol.* **132**, 155–162.
- Friedmann-Morvinski, D. and Verma, I. M.** (2014). Dedifferentiation and reprogramming: Origins of cancer stem cells. *EMBO Rep.* **15**, 244–253.
- Friedmann-Morvinski, D., Bushong, E. A., Ke, E., Soda, Y., Marumoto, T., Singer, O., Ellisman, M. H. and Verma, I. M.** (2012). Dedifferentiation of neurons and astrocytes by oncogenes can induce gliomas in mice. *Science (80-. ).* **338**, 1080–1084.
- Frühwald** (2020). Age and DNA methylation subgroup as potential independent risk factors for treatment stratification in children with atypical teratoid/rhabdoid tumors. *Neuro. Oncol.* **22**, 1006–1017.
- Fulp, C. T., Cho, G., Marsh, E. D., Nasrallah, I. M., Labosky, P. A. and Golden, J. A.** (2008). Identification of Arx transcriptional targets in the developing basal forebrain. *Hum. Mol. Genet.* **17**, 3740–3760.
- Gadd, S., Sredni, S. T., Huang, C. C. and Perlman, E. J.** (2010). Rhabdoid tumor: Gene expression clues to pathogenesis and potential therapeutic targets. *Lab. Investig.* **90**, 724–738.
- Gasser, E., Johannssen, H. C., Rüllicke, T., Zeilhofer, H. U. and Stoffel, M.** (2016). Foxa1 is essential for development and functional integrity of the subthalamic nucleus. *Sci. Rep.* **6**, 1–15.
- Gilbertson, R. J. and Gutmann, D. H.** (2007). Tumorigenesis in the brain: Location, location, location. *Cancer Res.* **67**, 5579–5582.
- Gorban, A. and Zinovyev, A.** (2005). Elastic principal graphs and manifolds and their

- practical applications. *Comput. (Vienna/New York)* **75**, 359–379.
- Gu, Z., Eils, R. and Schlesner, M.** (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849.
- Guidi, C. J., Sands, A. T., Zambrowicz, B. P., Turner, T. K., Demers, D. A., Webster, W., Smith, T. W., Imbalzano, A. N. and Jones, S. N.** (2001). Disruption of *Ini1* Leads to Peri-Implantation Lethality and Tumorigenesis in Mice. *Mol. Cell. Biol.* **21**, 3598–3603.
- Han, Z. Y., Richer, W., Fréneaux, P., Chauvin, C., Lucchesi, C., Guillemot, D., Grison, C., Lequin, D., Pierron, G., Masliah-Planchon, J., et al.** (2016). The occurrence of intracranial rhabdoid tumours in mice depends on temporal control of *Smarcb1* inactivation. *Nat. Commun.* **7**, 10421.
- Hanks, M., Wurst, W., Anson-Cartwright, L., Auerbach, A. B. and Joyner, A. L.** (1995). Rescue of the *En-1* mutant phenotype by replacement of *En-1* with *En-2*. *Science (80- )*. **269**, 679–682.
- Hibi, M. and Shimizu, T.** (2012). Development of the cerebellum and cerebellar neural circuits. *Dev. Neurobiol.* **72**, 282–301.
- Hirata, H. and Tomita, K.** (2001). *Hes1* and *Hes3* regulate maintenance of the isthmic organizer and development of the mid/hindbrain. *EMBO J.* **20**, 4454±4466.
- Ho, B., Johann, P. D., Johann, P. D., Johann, P. D., Grabovska, Y., De Dieu Andrianteranagna, M. J., De Dieu Andrianteranagna, M. J., Yao, F., Frühwald, M., Hasselblatt, M., et al.** (2020). Molecular subgrouping of atypical teratoid/rhabdoid tumors - A reinvestigation and current consensus. *Neuro. Oncol.* **22**, 613–624.
- Holdhof, D., Schoof, M., Hellwig, M., Holdhof, N. H., Niesen, J. and Schüller, U.** (2020). HGFAP-Positive Stem Cells Depend on *Brg1* for Proper Formation of Cerebral and Cerebellar Structures. *Cereb. Cortex* **30**, 1382–1392.
- Holdhof, D., Johann, P. D., Spohn, M., Bockmayr, M. and Safaei, S.** (2021a). Atypical teratoid / rhabdoid tumors ( ATRTs ) with SMARCA4 mutation are molecularly distinct from SMARCB1 - deficient cases. *Acta Neuropathol.* **141**, 291–301.
- Holdhof, D., Schoof, M., Al-Kersh, S., Spohn, M., Kresbach, C., Göbel, C., Hellwig, M., Indenbirken, D., Moreno, N., Kerl, K., et al.** (2021b). *Brahma*-related gene 1 has time-specific roles during brain and eye development. *Development* **148**, 196147.
- Isakoff, M. S., Sansam, C. G., Tamayo, P., Subramanian, A., Evans, J. A., Fillmore, C. M., Wang, X., Biegel, J. A., Pomeroy, S. L., Mesirov, J. P., et al.** (2005). Inactivation of the *Snf5* tumor suppressor stimulates cell cycle progression and cooperates with *p53* loss in oncogenic transformation. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17745–17750.

- Jagani, Z., Mora-Blanco, E. L., Sansam, C. G., McKenna, E. S., Wilson, B., Chen, D., Klekota, J., Tamayo, P., Nguyen, P. T. L., Tolstorukov, M., et al.** (2010). Loss of the tumor suppressor Snf5 leads to aberrant activation of the Hedgehog-Gli pathway. *Nat. Med.* **16**, 1429–1434.
- Jakovcevski, I. and Zecevic, N.** (2005). Olig transcription factors are expressed in oligodendrocyte and neuronal cells in human fetal CNS. *J. Neurosci.* **25**, 10064–10073.
- Jessa, S., Blanchet-Cohen, A., Krug, B., Vladoiu, M., Coutelier, M., Faury, D., Poreau, B., De Jay, N., Hébert, S., Monlong, J., et al.** (2019). Stalled developmental programs at the root of pediatric brain tumors. *Nat. Genet.* **51**, 1702–1713.
- Johann, P. D., Erkek, S., Zapatka, M., Kerl, K., Buchhalter, I., Hovestadt, V., Jones, D. T. W., Sturm, D., Hermann, C., Segura Wang, M., et al.** (2016). Atypical Teratoid/Rhabdoid Tumors Are Comprised of Three Epigenetic Subgroups with Distinct Enhancer Landscapes. *Cancer Cell* **29**, 379–393.
- Joyner, A. L., Herrup, K., Auerbach, B. A., Davis, C. A. and Rossant, J.** (1991). Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the En-2 homeobox. *Science (80- )*. **251**, 1239–1243.
- Kadoch, C. and Crabtree, G. R.** (2015). Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, 1–18.
- Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J. and Crabtree, G. R.** (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601.
- Karnezis, A. N., Hoang, L. N., Coatham, M., Ravn, S., Almadani, N., Tessier-Cloutier, B., Irving, J., Meng, B., Li, X., Chow, C., et al.** (2016). Loss of switch/sucrose non-fermenting complex protein expression is associated with dedifferentiation in endometrial carcinomas. *Mod. Pathol.* **29**, 302–314.
- Kelly, S. M., Raudales, R., He, M., Lee, J. H., Kim, Y., Gibb, L. G., Wu, P., Matho, K., Osten, P., Graybiel, A. M., et al.** (2018). Radial Glial Lineage Progression and Differential Intermediate Progenitor Amplification Underlie Striatal Compartments and Circuit Organization. *Neuron* **99**, 345-361.e4.
- Kieran** (2012). Absence of Oncogenic Canonical Pathway Mutations in Aggressive Pediatric Rhabdoid Tumors. *Pediatr Blood Cancer* **59**, 1155–1157.
- Kim, J. K., Huh, S.-O., Choi, H., Lee, K.-S., Shin, D., Lee, C., Nam, J.-S., Kim, H., Chung, H., Lee, H. W., et al.** (2001). Srg3, a Mouse Homolog of Yeast SWI3, Is Essential for

- Early Embryogenesis and Involved in Brain Development. *Mol. Cell. Biol.* **21**, 7787–7795.
- Klochendler-Yeivin, A., Fiette, L., Barra, J., Muchardt, C., Babinet, C. and Yaniv, M.** (2000). The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression. *EMBO Rep.* **1**, 500–506.
- Kosho, T., Miyake, N. and Carey, J. C.** (2014a). Coffin-Siris syndrome and related disorders involving components of the BAF (mSWI/SNF) complex: Historical review and recent advances using next generation sequencing. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **166**, 241–251.
- Kosho, T., Okamoto, N., Imai, Y., Ohashi, H., van Erde, A. M., Chrzanowska, K., Clayton-Smith, J., Kingston, H., Mari, F., Aggarwal, S., et al.** (2014b). Genotype-phenotype correlation of coffin-siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **166**, 262–275.
- Krumlauf, R.** (2016). *Hox Genes and the Hindbrain. A Study in Segments*. 1st ed. Elsevier Inc.
- La Manno, G., Gyllborg, D., Codeluppi, S., Nishimura, K., Salto, C., Zeisel, A., Borm, L. E., Stott, S. R. W., Toledo, E. M., Villaescusa, J. C., et al.** (2016). Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. *Cell* **167**, 566–580.e19.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., et al.** (2018). RNA velocity of single cells. *Nature* **560**, 494–498.
- Lafay-Cousin, L., Hawkins, C., Carret, A. S., Johnston, D., Zelcer, S., Wilson, B., Jabado, N., Scheinemann, K., Eisenstat, D., Fryer, C., et al.** (2012). Central nervous system atypical teratoid rhabdoid tumours: The Canadian Paediatric Brain Tumour Consortium experience. *Eur. J. Cancer* **48**, 353–359.
- Langer, L. F., Ward, J. M. and Archer, T. K.** (2019). Tumor suppressor SMARCB1 suppresses super-enhancers to govern hESC lineage determination. *Elife* **8**, 1–23.
- Lê Cao, K. A., Boitard, S. and Besse, P.** (2011). Sparse PLS discriminant analysis: Biologically relevant feature selection and graphical displays for multiclass problems. *BMC Bioinformatics* **12**, 0–16.
- Lee, R. S., Stewart, C., Carter, S. L., Ambrogio, L., Cibulskis, K., Sougnez, C., Lawrence, M. S., Auclair, D., Mora, J., Golub, T. R., et al.** (2012). A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J. Clin. Invest.* **122**, 2983–2988.
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'Donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., et al.** (2016).

- Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291.
- Lemon, B., Inouye, C., King, D. S. and Tjian, R.** (2001). Selectivity of chromatin-remodelling cofactors for ligand-activated transcription. *Nature* **414**, 924–928.
- Leruste, A., Tosello, J., Ramos, R. N., Tauziède-Espariat, A., Brohard, S., Han, Z. Y., Beccaria, K., Andrianteranagna, M., Caudana, P., Nikolic, J., et al.** (2019). Clonally Expanded T Cells Reveal Immunogenicity of Rhabdoid Tumors. *Cancer Cell* **36**, 597–612.
- Lessard, J., Wu, J. I., Ranish, J. A., Wan, M., Winslow, M. M., Staahl, B. T., Wu, H., Aebbersold, R., Graef, I. A. and Crabtree, G. R.** (2007). An Essential Switch in Subunit Composition of a Chromatin Remodeling Complex during Neural Development. *Neuron* **55**, 201–215.
- Li, H.** (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **00**, 1–3.
- Li, B. K., Al-Karmi, S., Huang, A. and Bouffet, E.** (2020). Pediatric embryonal brain tumors in the molecular era. *Expert Rev. Mol. Diagn.* **20**, 293–303.
- Liu, C., Sage, J. C., Miller, M. R., Verhaak, R. G. W., Hippenmeyer, S., Vogel, H., Foreman, O., Bronson, R. T., Nishiyama, A., Luo, L., et al.** (2011). Mosaic analysis with double markers reveals tumor cell of origin in glioma. *Cell* **146**, 209–221.
- Lobe, C. G.** (1997). Expression of the helix-loop-helix factor, Hes3, during embryo development suggests a role in early midbrain-hindbrain patterning. *Mech. Dev.* **62**, 227–237.
- Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., Hawkins, C., Ng, H. K., Pfister, S. M., Reifenberger, G., et al.** (2021). The 2021 WHO classification of tumors of the central nervous system: A summary. *Neuro. Oncol.* **23**, 1231–1251.
- Love, M. I., Huber, W. and Anders, S.** (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21.
- Mani, S. A., Guo, W., Liao, M. J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., Brooks, M., Reinhard, F., Zhang, C. C., Shipitsin, M., et al.** (2008). The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. *Cell* **133**, 704–715.
- Marathe, H. G., Mehta, G., Zhang, X., Datar, I., Mehrotra, A., Yeung, K. C. and de la Serna, I. L.** (2013). SWI/SNF Enzymes Promote SOX10-Mediated Activation of Myelin Gene Expression. *PLoS One* **8**, 1–14.
- Mariette, J. and Villa-Vialaneix, N.** (2018). Unsupervised multiple kernel learning for

- heterogeneous data integration. *Bioinformatics* **34**, 1009–1015.
- Marino, S. and Gilbertson, R. J.** (2021). Harnessing brain development to understand brain tumours. *Dev.* **148**, 1–4.
- Martinez, S., Andreu, A., Mecklenburg, N. and Echevarria, D.** (2013). Cellular and molecular basis of cerebellar development. *Front. Neuroanat.* **7**, 1–12.
- Masliah-Planchon, J., Bièche, I., Guinebretière, J. M., Bourdeaut, F. and Delattre, O.** (2015). SWI/SNF chromatin remodeling and human malignancies. *Annu. Rev. Pathol. Mech. Dis.* **10**, 145–171.
- Mathur, R. and Roberts, C. W. M.** (2018). SWI/SNF (BAF) Complexes: Guardians of the Epigenome. *Annu. Rev. Cancer Biol.* **2**, 413–427.
- Matsumoto, K., Nishihara, S., Kamimura, M., Shiraishi, T., Otaguro, T., Uehara, M., Maeda, Y., Ogura, K., Lumsden, A. and Ogura, T.** (2004). The prepattern transcription factor *Irx2*, a target of the FGF8/MAP kinase cascade, is involved in cerebellum formation. *Nat. Neurosci.* **7**, 605–612.
- Matsumoto, S., Banine, F., Struve, J., Xing, R., Adams, C., Liu, Y., Metzger, D., Chambon, P., Rao, M. S. and Sherman, L. S.** (2006). Brg1 is required for murine neural stem cell maintenance and gliogenesis. *Dev. Biol.* **289**, 372–383.
- Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C. and Koeffler, H. P.** (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* **28**, 1747–1756.
- McInnes, L., Healy, J. and Melville, J.** (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.
- Melcher, V., Graf, M., Interlandi, M., Moreno, N., de Faria, F. W., Kim, S. N., Kastrati, D., Korbanka, S., Alfert, A., Gerß, J., et al.** (2020). Macrophage-tumor cell interaction promotes ATRT progression and chemoresistance. *Acta Neuropathol.* **139**, 913–936.
- Méndez-Maldonado, K., Vega-López, G. A., Aybar, M. J. and Velasco, I.** (2020). Neurogenesis From Neural Crest Cells: Molecular Mechanisms in the Formation of Cranial Nerves and Ganglia. *Front. Cell Dev. Biol.* **8**, 635.
- Mittal, P. and Roberts, C. W. M.** (2020). The SWI/SNF complex in cancer — biology, biomarkers and therapy. *Nat. Rev. Clin. Oncol.* **17**, 435–448.
- Miyoshi, G., Hjerling-Leffler, J., Karayannis, T., Sousa, V. H., Butt, S. J. B., Battiste, J., Johnson, J. E., Machold, R. P. and Fishell, G.** (2010). Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. *J. Neurosci.* **30**, 1582–1594.

- Mora-Blanco, E. L., Mishina, Y., Tillman, E. J., Cho, Y. J., Thom, C. S., Pomeroy, S. L., Shao, W. and Roberts, C. W. M.** (2014). Activation of  $\beta$ -catenin/TCF targets following loss of the tumor suppressor SNF5. *Oncogene* **33**, 933–938.
- Moreno, N., Schmidt, C., Ahlfeld, J., Pöschl, J., Dittmar, S., Pfister, S. M., Kool, M., Kerl, K. and Schüller, U.** (2014). Loss of Smarc proteins impairs cerebellar development. *J. Neurosci.* **34**, 13486–13491.
- Morgan, M. B., Stevens, L., Patterson, J. and Tannenbaum, M.** (2000). Cutaneous epithelioid malignant nerve sheath tumor with rhabdoid features: A histologic, immunohistochemical, and ultrastructural study of three cases. *J. Cutan. Pathol.* **27**, 529–534.
- Nagl, N. G., Zweitzig, D. R., Thimmapaya, B., Beck, G. R. and Moran, E.** (2006). The c-myc gene is a direct target of mammalian SWI/SNF-related complexes during differentiation-associated cell cycle arrest. *Cancer Res.* **66**, 1289–1293.
- Nakamura, H., Katahira, T., Matsunaga, E. and Sato, T.** (2005). Isthmus organizer for midbrain and hindbrain development. *Brain Res. Rev.* **49**, 120–126.
- Nakano, M., Kikushige, Y., Miyawaki, K., Kunisaki, Y., Mizuno, S., Takenaka, K., Tamura, S., Okumura, Y., Ito, M., Ariyama, H., et al.** (2019). Dedifferentiation process driven by TGF-beta signaling enhances stem cell properties in human colorectal cancer. *Oncogene* **38**, 780–793.
- Nakayama, R. T., Pulice, J. L., Valencia, A. M., McBride, M. J., McKenzie, Z. M., Gillespie, M. A., Ku, W. L., Teng, M., Cui, K., Williams, R. T., et al.** (2017). SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nat. Genet.* **49**, 1613–1623.
- Narayanan, R., Pirouz, M., Kerimoglu, C., Pham, L., Wagener, R. J., Kiszka, K. A., Rosenbusch, J., Seong, R. H., Kessel, M., Fischer, A., et al.** (2015). Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. *Cell Rep.* **13**, 1842–1854.
- Nemes, K. and Frühwald, M. C.** (2018). Emerging therapeutic targets for the treatment of malignant rhabdoid tumors. *Expert Opin. Ther. Targets* **00**, 14728222.2018.1451839.
- Neumann, J. E., Wefers, A. K., Lambo, S., Bianchi, E., Bockstaller, M., Dorostkar, M. M., Meister, V., Schindler, P., Korshunov, A., Von Hoff, K., et al.** (2017). A mouse model for embryonal tumors with multilayered rosettes uncovers the therapeutic potential of Sonic-hedgehog inhibitors. *Nat. Med.* **23**, 1191–1202.
- Nguyen, H., Sokpor, G., Pham, L., Rosenbusch, J., Stoykova, A., Staiger, J. F. and Tuoc,**

- T.** (2016). Epigenetic regulation by BAF (mSWI/SNF) chromatin remodeling complexes is indispensable for embryonic development. *Cell Cycle* **15**, 1317–1324.
- Nowak, J., Nemes, K., Hohm, A., Vandergrift, L. A., Hasselblatt, M., Johann, P. D., Kool, M., Frühwald, M. C. and Warmuth-Metz, M.** (2018). Magnetic resonance imaging surrogates of molecular subgroups in atypical teratoid/rhabdoid tumor. *Neuro. Oncol.* **20**, 1672–1679.
- Oh, C. C., Orr, B. A., Bernardi, B., Garré, M. L., Rossi, A., Figà-Talamanca, L., Robinson, G. W. and Patay, Z.** (2015). Atypical teratoid/rhabdoid tumor (ATRT) arising from the 3rd cranial nerve in infants: a clinical-radiological entity? *J. Neurooncol.* **124**, 175–183.
- Olave, I., Wang, W., Xue, Y., Kuo, A. and Crabtree, G. R.** (2002). Identification of a polymorphic, neuron-specific chromatin remodeling complex. *Genes Dev.* **16**, 2509–2517.
- Oruetxebarria, I., Venturini, F., Kekarainen, T., Houweling, A., Zuijderduijn, L. M. P., Mohd-Sarip, A., Vries, R. G. J., Hoeben, R. C. and Verrijzer, C. P.** (2004). p16INK4a Is Required for hSNF5 Chromatin Remodeler-induced Cellular Senescence in Malignant Rhabdoid Tumor Cells. *J. Biol. Chem.* **279**, 3807–3816.
- Parisian, A. D., Koga, T., Miki, S., Johann, P. D., Kool, M., Crawford, J. R. and Furnari, F. B.** (2020). SMARCB1 loss interacts with neuronal differentiation state to block maturation and impact cell stability. *Genes Dev.* **34**, 1316–1329.
- Parker, H. J. and Krumlauf, R.** (2020). *A Hox gene regulatory network for hindbrain segmentation*. 1st ed. Elsevier Inc.
- Parker, H. J., Pushel, I. and Krumlauf, R.** (2018). Coupling the roles of Hox genes to regulatory networks patterning cranial neural crest. *Dev. Biol.* **444**, S67–S78.
- Parmigiani, E., Taylor, V. and Giachino, C.** (2020). Oncogenic and Tumor-Suppressive Functions of NOTCH Signaling in Glioma. *Cells* **9**, 2304.
- Partanen, J.** (2007). FGF signalling pathways in development of the midbrain and anterior hindbrain. *J. Neurochem.* **101**, 1185–1193.
- Peretz, Y., Eren, N., Kohl, A., Hen, G., Yaniv, K., Weisinger, K., Cinnamon, Y. and Seladonfeld, D.** (2016). A new role of hindbrain boundaries as pools of neural stem/progenitor cells regulated by Sox2. *BMC Biol.* **14**, 1–20.
- Peukert, D., Weber, S., Lumsden, A. and Scholpp, S.** (2011). Lhx2 and Lhx9 determine neuronal differentiation and compartment in the caudal forebrain by regulating Wnt signaling. *PLoS Biol.* **9**, e1001218.
- Qi, W., Wang, R., Chen, H., Wang, X., Xiao, T., Boldogh, I., Ba, X., Han, L. and Zeng, X.** (2015). BRG1 promotes the repair of DNA double-strand breaks by facilitating the

- replacement of RPA with RAD51. *J. Cell Sci.* **128**, 317–330.
- Reddy, A. T., Reddy, A. T., Reddy, A. T., Strother, D. R., Judkins, A. R., Burger, P. C. and Fouladi, M.** (2020). Efficacy of High-Dose Chemotherapy and 3-D Conformal Radiation for Atypical Teratoid / Rhabdoid Tumor: A Report from the Children ' s Oncology Group Trial. *J. Clin. Oncol.* **38**, 117.
- Rizzo, D., Fréneaux, P., Brisse, H., Louvrier, C., Lequin, D., Nicolas, A., Ranchère, D., Verkarre, V., Jouvét, A., Dufour, C., et al.** (2012). SMARCB1 deficiency in tumors from the peripheral nervous system: A link between schwannomas and rhabdoid tumors? *Am. J. Surg. Pathol.* **36**, 964–972.
- Roberts, C. W. M. and Galusha, S. A.** (2000). Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice Charles. **97**, 13796–13800.
- Roberts, C. W. M., Leroux, M. M., Fleming, M. D. and Orkin, S. H.** (2002). Highly penetrant , rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. **2**, 415–425.
- Rohart, F., Gautier, B., Singh, A. and Lê Cao, K. A.** (2017). mixOmics: An R package for ' omics feature selection and multiple data integration. *PLoS Comput. Biol.* **13**, 1–19.
- Rorke, L. B., Packer, R. J. and Biegel, J. A.** (1996). Central nervous system atypical teratoid/rhabdoid tumors of infancy and childhood: Definition of an entity. *J. Neurosurg.* **85**, 56–65.
- Rossi, S., Brenca, M., Zanatta, L., Trincia, E., Guerriero, A., Pizzato, C., Fiorindi, A., Viscardi, E., Giangaspero, F., Maestro, R., et al.** (2018). A pediatric intra-axial malignant smarcb1-deficient desmo-plastic tumor arising in meningioangiomas. *J. Neuropathol. Exp. Neurol.* **77**, 883–889.
- Roychoudhury, K., Salomone, J., Qin, S., Cain, B., Adam, M., Steven Potter, S., Nakafuku, M., Gebelein, B. and Campbell, K.** (2020). Physical interactions between Gsx2 and Ascl1 balance progenitor expansion versus neurogenesis in the mouse lateral ganglionic eminence. *Development* **147**, 185348.
- Santen, G. W. E., Kriek, M. and van Attikum, H.** (2012). SWI/SNF complex in disorder switching from malignancies to intellectual disability. *Epigenetics* **7**, 1219–1224.
- Schmitt, A., Gofferje, V., Weber, M., Meyer, J., Mössner, R. and Lesch, K. P.** (2003). The Brain-Specific Protein MLC1 Implicated in Megalencephalic Leukoencephalopathy with Subcortical Cysts Is Expressed in Glial Cells in the Murine Brain. *Glia* **44**, 283–295.
- Schneppenheim, R., Frühwald, M. C., Gesk, S., Hasselblatt, M., Jeibmann, A., Kordes, U., Kreuz, M., Leuschner, I., Subero, J. I. M., Obser, T., et al.** (2010). Germline

- Nonsense Mutation and Somatic Inactivation of SMARCA4/BRG1 in a Family with Rhabdoid Tumor Predisposition Syndrome. *Am. J. Hum. Genet.* **86**, 279–284.
- Schrey, D., Carceller Lechón, F., Malietzis, G., Moreno, L., Dufour, C., Chi, S., Lafay-Cousin, L., von Hoff, K., Athanasiou, T., Marshall, L. V., et al.** (2016). Multimodal therapy in children and adolescents with newly diagnosed atypical teratoid rhabdoid tumor: individual pooled data analysis and review of the literature. *J. Neurooncol.* **126**, 81–90.
- Schwitalla, S., Fingerle, A. A., Cammareri, P., Nebelsiek, T., Göktuna, S. I., Ziegler, P. K., Canli, O., Heijmans, J., Huels, D. J., Moreaux, G., et al.** (2013). Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152**, 25–38.
- Shain, A. H. and Pollack, J. R.** (2013). The Spectrum of SWI/SNF Mutations, Ubiquitous in Human Cancers. *PLoS One* **8**, 55119.
- Shen, R. and Seshan, V. E.** (2016). FACETS: Allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res.* **44**, 1–9.
- Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. and Sirotkin, K.** (2001). dbSNP: The NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311.
- Siebel, C. and Lendahl, U.** (2017). Notch signaling in development, tissue homeostasis, and disease. *Physiol. Rev.* **97**, 1235–1294.
- Smith, M. J., Wallace, A. J., Bowers, N. L., Eaton, H. and Evans, D. G. R.** (2014). SMARCB1 mutations in schwannomatosis and genotype correlations with rhabdoid tumors. *Cancer Genet.* **207**, 373–378.
- Sokpor, G., Xie, Y., Rosenbusch, J. and Tuoc, T.** (2017). Chromatin remodeling BAF (SWI/SNF) complexes in neural development and disorders. *Front. Mol. Neurosci.* **10**, 1–22.
- Sredni, S. T. and Tomita, T.** (2015). Rhabdoid tumor predisposition syndrome. *Pediatr. Dev. Pathol.* **18**, 49–58.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., Hao, Y., Stoeckius, M., Smibert, P. and Satija, R.** (2019). Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21.
- Sturm, G., Finotello, F., Petitprez, F., Zhang, J. D., Baumbach, J., Fridman, W. H., List, M. and Aneichyk, T.** (2019). Comprehensive evaluation of transcriptome-based cell-type

- quantification methods for immuno-oncology. *Bioinformatics* **35**, i436–i445.
- Szu, J., Wojcinski, A., Jiang, P. and Kesari, S.** (2021). Impact of the Olig Family on Neurodevelopmental Disorders. *Front. Neurosci.* **15**, 1–19.
- Teodorczyk, M. and Schmidt, M. H. H.** (2014). Notching on cancer’s door: Notch signaling in brain tumors. *Front. Oncol.* **4**, 1–14.
- Than-Trong, E., Ortica-Gatti, S., Mella, S., Nepal, C., Alunni, A. and Bally-Cuif, L.** (2018). Neural stem cell quiescence and stemness are molecularly distinct outputs of the notch3 signalling cascade in the vertebrate adult brain. *Dev.* **145**,.
- Torchia, J., Golbourn, B., Feng, S., Ho, K. C., Sin-Chan, P., Vasiljevic, A., Norman, J. D., Guilhamon, P., Garzia, L., Agamez, N. R., et al.** (2016). Integrated (epi)-Genomic Analyses Identify Subgroup-Specific Therapeutic Targets in CNS Rhabdoid Tumors. *Cancer Cell* **30**, 891–908.
- Tsikitis, M., Zhang, Z., Edelman, W., Zagzag, D. and Kalpana, G. V.** (2005). Genetic ablation of Cyclin D1 abrogates genesis of rhabdoid tumors resulting from *Ini1* loss. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12129–12134.
- Tuoc, T. C., Boretius, S., Sansom, S. N., Pitulescu, M. E., Frahm, J., Livesey, F. J. and Stoykova, A.** (2013). Chromatin Regulation by BAF170 Controls Cerebral Cortical Size and Thickness. *Dev. Cell* **25**, 256–269.
- Turrero García, M. and Harwell, C. C.** (2017). Radial glia in the ventral telencephalon. *FEBS Lett.* **591**, 3942–3959.
- Upadhyaya, S. A., Robinson, G. W., Onar-Thomas, A., Orr, B. A., Johann, P., Wu, G., Billups, C. A., Tatevossian, R. G., Dhanda, S. K., Srinivasan, A., et al.** (2021). Relevance of molecular groups in children with newly diagnosed atypical teratoid rhabdoid tumor: Results from prospective St. Jude multi-institutional trials. *Clin. Cancer Res.* **27**, 2879–2889.
- Van de Sande, B., Flerin, C., Davie, K., De Waegeneer, M., Hulselmans, G., Aibar, S., Seurinck, R., Saelens, W., Cannoodt, R., Rouchon, Q., et al.** (2020). A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat. Protoc.* **15**, 2247–2276.
- Verma, A. and Morriss, C.** (2008). Atypical teratoid/rhabdoid tumor of the optic nerve. *Pediatr. Radiol.* **38**, 1117–1121.
- Versteeg, I.** (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**, 203–206.
- Vitte, J., Gao, F., Coppola, G., Judkins, A. R. and Giovannini, M.** (2017). Timing of

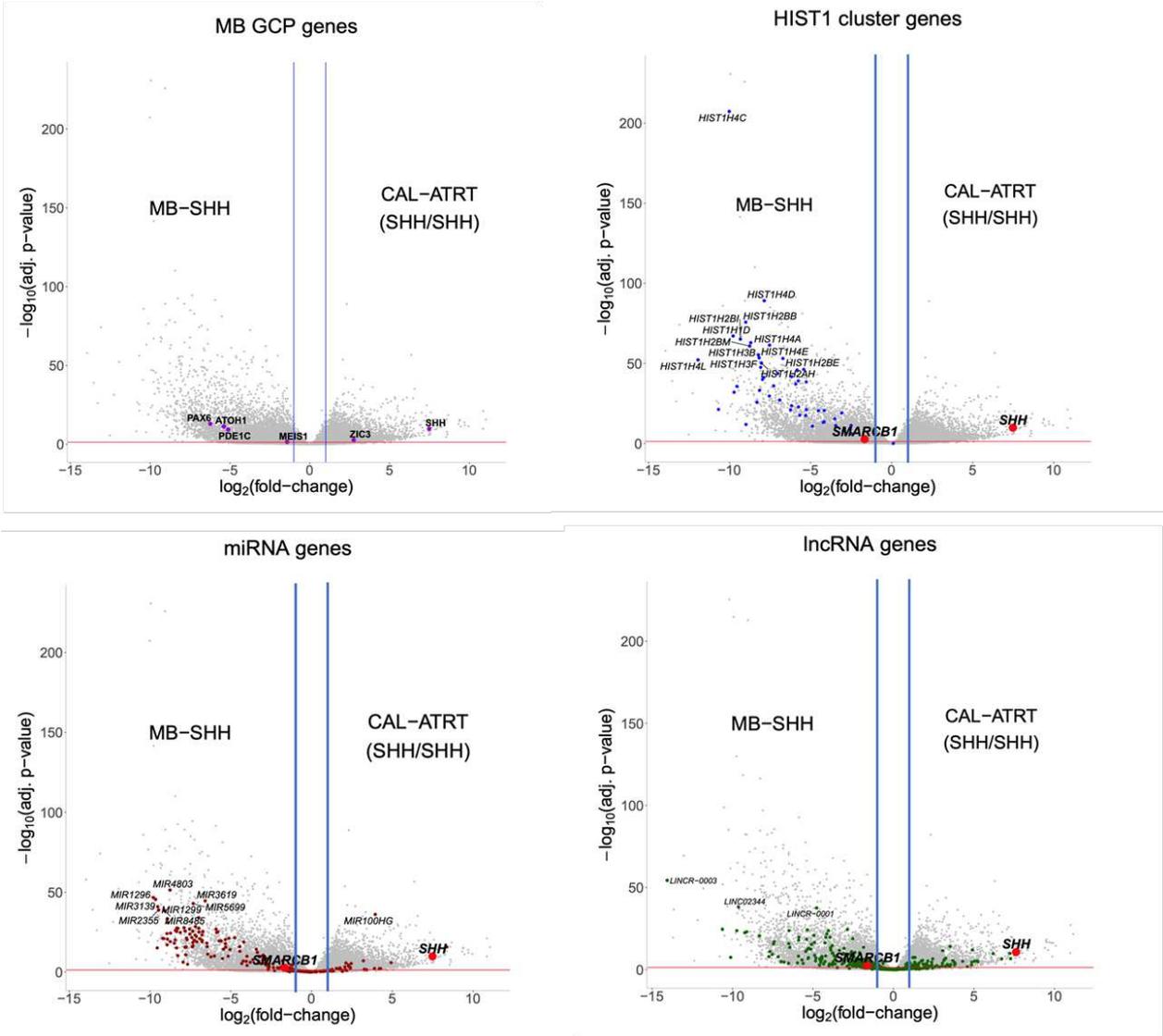
- Smrcb1 and Nf2 inactivation determines schwannoma versus rhabdoid tumor development. *Nat. Commun.* **8**, 1–12.
- Vogel-Ciernia, A., Matheos, D. P., Barrett, R. M., Kramár, E. A., Azzawi, S., Chen, Y., Magnan, C. N., Zeller, M., Sylvain, A., Haettig, J., et al.** (2013). The neuron-specific chromatin regulatory subunit BAF53b is necessary for synaptic plasticity and memory. *Nat. Neurosci.* **16**, 552–561.
- Wang, K., Li, M. and Hakonarson, H.** (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, 1–7.
- Wang, X., Liu, X., Lin, Z., Chen, Y., Wang, P. and Zhang, S.** (2015). Atypical teratoid/rhabdoid tumor (AT/RT) arising from the acoustic nerve in a young adult: A case report and a review of literature. *Med. (United States)* **94**, 1–4.
- Wang, X., Lee, R. S., Alver, B. H., Haswell, J. R., Wang, S., Mieczkowski, J., Drier, Y., Gillespie, S. M., Archer, T. C., Wu, J. N., et al.** (2017). SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat. Genet.* **49**, 289–295.
- Weider, M., Küspert, M., Bischof, M., Vogl, M. R., Hornig, J., Loy, K., Kosian, T., Müller, J., Hillgärtner, S., Tamm, E. R., et al.** (2012). Chromatin-Remodeling Factor Brg1 Is Required for Schwann Cell Differentiation and Myelination. *Dev. Cell* **23**, 193–201.
- Wilkerson, M. D. and Hayes, D. N.** (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573.
- Wilson, B. G. and Roberts, C. W. M.** (2011). SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* **11**, 481–492.
- Wilson, B. G., Wang, X., Shen, X., McKenna, E. S., Lemieux, M. E., Cho, Y. J., Koellhoffer, E. C., Pomeroy, S. L., Orkin, S. H. and Roberts, C. W. M.** (2010). Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer Cell* **18**, 316–328.
- Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., Rajewsky, N., Simon, L. and Theis, F. J.** (2019). PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 1–9.
- Wonders, C. P. and Anderson, S. A.** (2006). The origin and specification of cortical interneurons. *Nat. Rev. Neurosci.* **7**, 687–696.
- Wright, C. F., McRae, J. F., Clayton, S., Gallone, G., Aitken, S., FitzGerald, T. W., Jones, P., Prigmore, E., Rajan, D., Lord, J., et al.** (2018). Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with

- developmental disorders. *Genet. Med.* **20**, 1216–1223.
- Wu, P.-S.** (2020). 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.
- Wu, J. I., Lessard, J., Olave, I. A., Qiu, Z., Ghosh, A., Graef, I. A. and Crabtree, G. R.** (2007). Regulation of Dendritic Development by Neuron-Specific Chromatin Remodeling Complexes. *Neuron* **56**, 94–108.
- Wykoff, C. C., Lam, B. L., Brathwaite, C. D., Biegel, J. A., McKeown, C. A., Rosenblum, M. K., Allewelt, H. B. and Sandberg, D. I.** (2008). Atypical teratoid/rhabdoid tumor arising from the third cranial nerve. *J. Neuro-Ophthalmology* **28**, 207–211.
- Yamada, Y., Haga, H. and Yamada, Y.** (2014). Concise Review: Dedifferentiation Meets Cancer Development: Proof of Concept for Epigenetic Cancer. *Stem Cells Transl. Med.* **3**, 1182–1187.
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P. W., Levine, D. A., et al.** (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**,.
- Yu, Y., Chen, Y., Kim, B., Wang, H., Zhao, C., He, X., Liu, L., Liu, W., Wu, L. M. N., Mao, M., et al.** (2013). Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell* **152**, 248–261.
- Yuzwa, S. A., Borrett, M. J., Innes, B. T., Voronova, A., Ketela, T., Kaplan, D. R., Bader, G. D. and Miller, F. D.** (2017). Developmental Emergence of Adult Neural Stem Cells as Revealed by Single-Cell Transcriptional Profiling. *Cell Rep.* **21**, 3970–3986.
- Zhao, Y. N., Li, K., Han, X. S., Pan, Y. W., Laneve, P., Rea, J. and Caffarelli, E.** (2021). The mechanism of non-coding RNAs in medulloblastoma (Review). *Front. Pediatr.* **7**, 1–7.



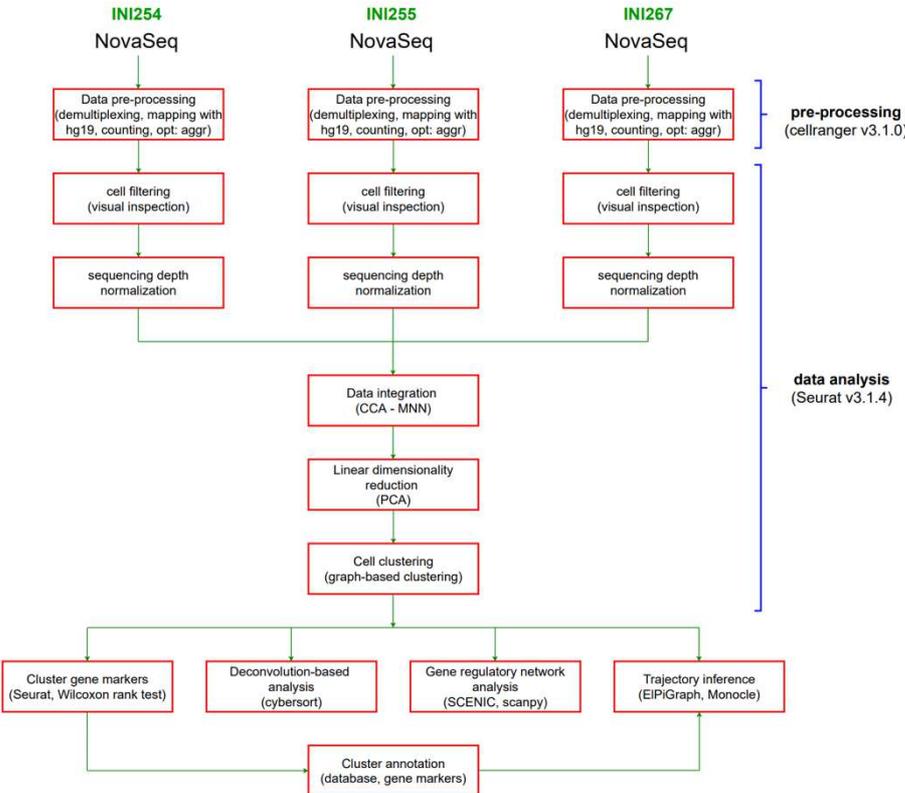
# Appendix Figures

**Appendix Figure 1. Differential expression analyses CAL-ATRTR versus SHH-MB**

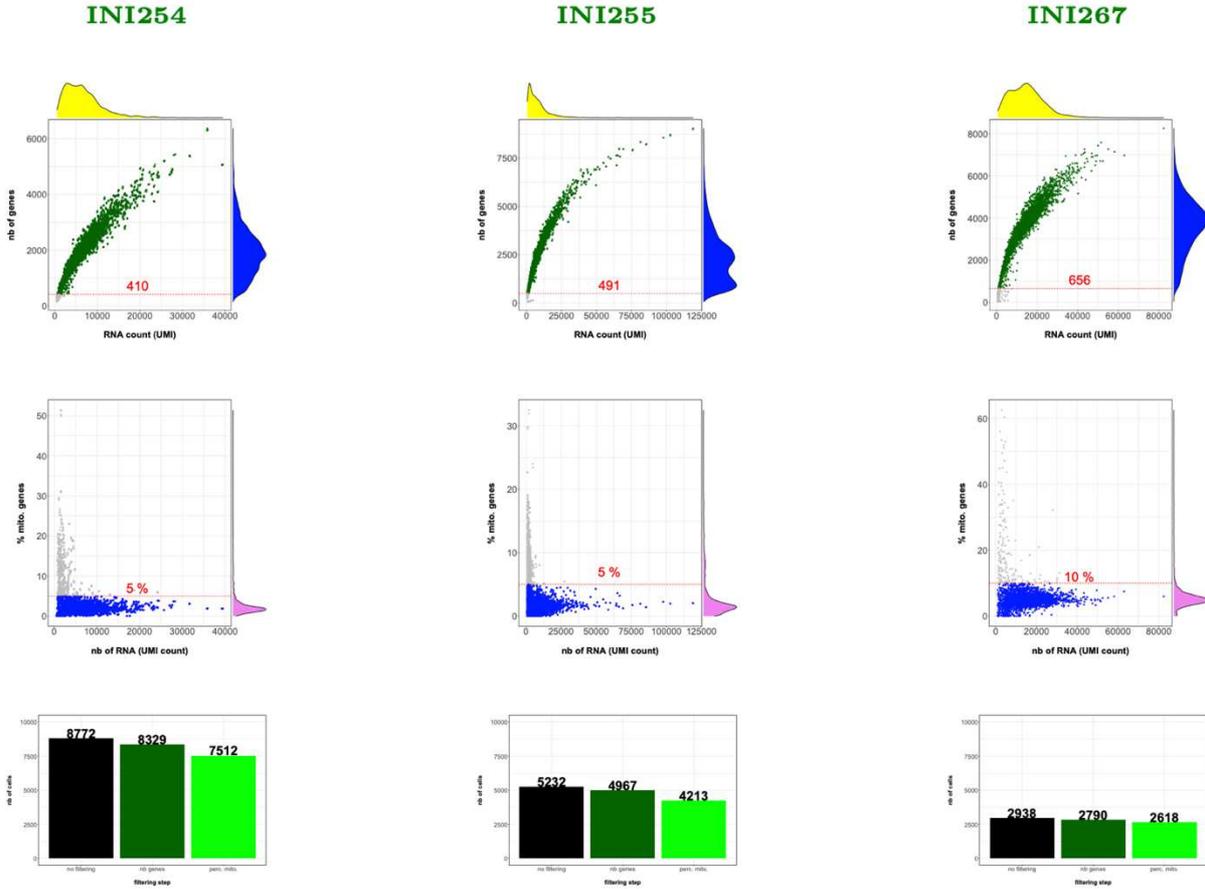


Volcano Plot showing the result of differential expression analysis between CAL-ATRTR and SHH-MB. Genes expressed in GCP, HIST1 genes, miRNA genes and lncRNA genes are labelled (A-D).

# Appendix Figure 2. Overview Single-cell RNAseq Methods

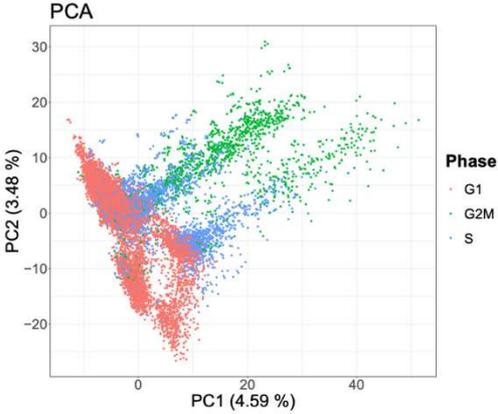
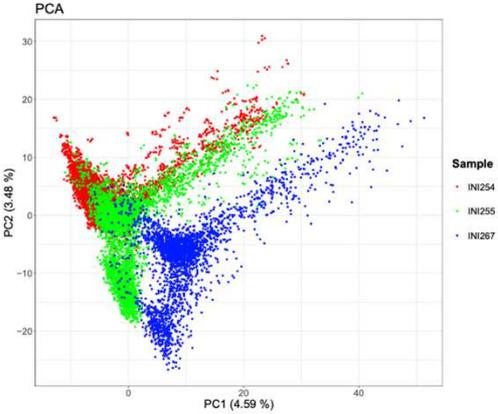


Appendix Figure 3. Single-cell filtering

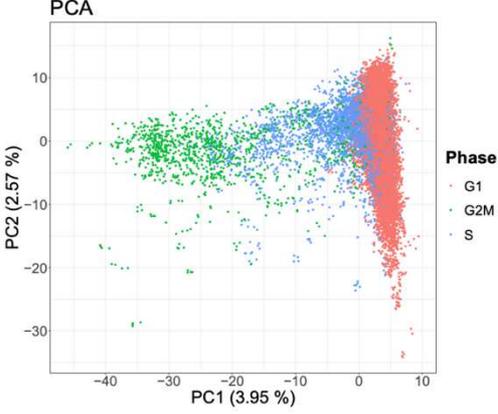
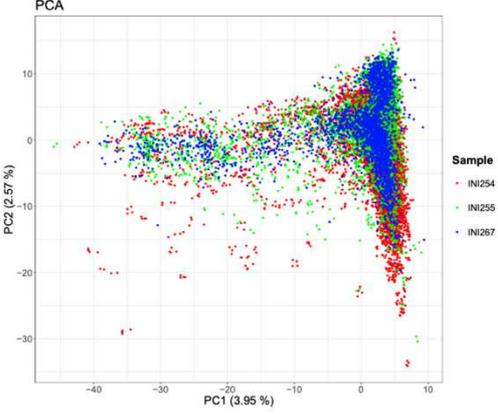


**Appendix Figure 4. Data integration quality assessment**

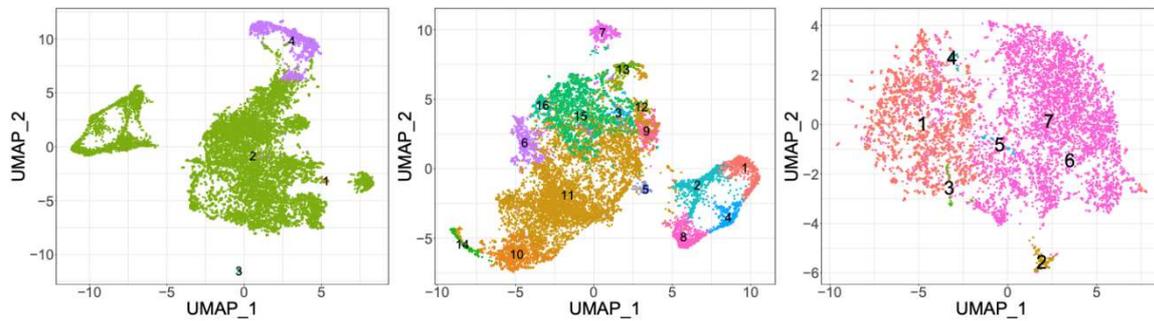
**Before**



**After**



## Appendix Figure 5. Iterative clustering by IKAP



The analysis with the BestIkap found 4 clusters (left UMAP). First, we performed cell types identification of cluster 1, 3 and 4. We removed the cells from cluster 1, 3 and 4 and then we re-do IKAP clustering with the remaining cells with unknown identity (cluster 2). Then, we analyzed the best ikap of the new object containing only the unknown cells of ikap1 (cluster 2 of the left UMAP). The Ikap2 analysis found 16 clusters (middle UMAP). We identified most of them, already found with other methods of clustering, but not the largest clusters (11 and 15). We therefore performed Ikap3 analysis of the unidentified clusters (11 and 15). The Ikap3 analysis (right UMAP) found again the 2 big clusters and 5 very small clusters. These 5 small clusters corresponded to residual cells from cell types/clusters previously identified. The two largest clusters (1,7) found in this ikap3 analysis had previously been consistently identified by all other clustering methods used, suggesting that they are true biological clusters.

# **Appendix Tables**

**Appendix Table 1. Liste of genes resulting from sPLS-DA.**

ene	comp1	comp2	comp3
CFAP74	0	0	-0.0202005365533264
GABRD	0	0	-0.0464628047756835
ARHGEF16	0	0	-0.0205265773096529
TEKT2	0.00356393790377508	0	0
CACHD1	0.0619215499494486	0	0
SLC44A5	0.0283817497303871	0	0
OLFM3	0	0.0710771098425677	0
SLC22A15	0	0.0904433481973893	0
MTMR11	0	-0.102730676200029	0
CASQ1	0	-0.0369887111573251	0
VANGL2	0.0109474613731316	0	0
TSTD1	0	0	-0.11172538584585
ILDR2	0.0134371082667456	0	0
CCDC181	0.0349081987953862	0	0
F5	0	0	-0.0622351802765879
B3GALT2	0	-0.0753031000710873	0
SHISA4	0	-0.0522817549467765	0
SLC26A9	0	-0.0108941730041158	0
MARC1	0.0967265833136825	0	0
FAM110C	0.213106127955983	0	0
LOC400940	0.0247997075691612	0	0
CAPN13	0	0	-0.0630238942068122
DGUOK-AS1	0	-0.102051908110151	0
SLC9A2	0	-0.00201808943486204	0
HOXD9	-0.0313451204247391	0	0
HOXD8	-0.0934414853206658	0	0
HOXD-AS2	-	0	0
	0.00151224531739126		
HOXD4	-0.0785649692655096	0	0
LINC01116	-0.0540288181596093	0	0
ZNF385B	0	0	-0.00112348603013649
COL5A2	-0.1358114170617	0	0
CASP8	-0.132845493789	0	0
SCG2	0	-0.196889970398037	0
KCNJ13	0	0	-0.342703547177429
TWIST2	0	0.0310319269179024	0
LRRN1	0.00384034495201572	0	0
OXTR	0	-0.0662287975452553	0
WNT7A	0.089016397663876	0	0
OSBPL10	0	0	-0.0431472938139615

CDCP1	0	0	-0.198255609184493
SLC6A20	0	0	-0.121561392850441
MITF	0	0	-0.00998092885023373
CADM2	0	-0.0670631943403869	0
IGSF11	0	-0.159128578230232	0
MAATS1	0.119063443599098	0	0
SEMA5B	0.0802301862046072	0	0
KLF15	0.0540993771292053	0	0
NME9	0.00312516517706991	0	0
CLRN1	0	-0.26879900749631	0
PPM1L	0.105004469604244	0	0
TNIK	0.0388142705540554	0	0
ATP13A5	0	0	-0.125991986930163
ATP13A4-AS1	0	0	-0.0168891183086517
TFRC	0	-0.205766262631477	0
RASL11B	0	0.0131377635131027	0
SPP1	0	-0.0353698675010098	0
DKK2	0	0.0131130558229762	0
HHIP-AS1	0	0.0267431778889431	0
DCHS2	0.0466651606640444	0	0
RAPGEF2	0	-0.0815471299322213	0
HELT	0	0	-0.281218539598993
ANKRD37	0	0	-0.143765747690501
TTC23L	0.0736323658309058	0	0
OSMR	0	-0.0531830996907512	0
CDC20B	0	0	-0.0858844231808351
MCIDAS	0	0	-0.187711614543145
CCNO	0	0	-0.176774389084366
ADGRV1	0.265872814882594	0	0
CDO1	0.0207101430617783	0	0
FBXL21	0	0.000345124090160303	0
PCDHB8	0	-0.0969316243626543	0
HTR4	0	-0.0215664149987743	0
GALNT10	0	0	-0.0425915301988897
WWC1	0.074588253203229	0	0
NSG2	0.140969577755098	0	0
RNF182	0.0147908578387507	0	0
DCDC2	0	0	-0.0662372758723662
KIAA0319	0.0790206226993169	0	0
C6orf223	0	0	-0.20551494949008
CYP39A1	0.0694692895203235	0	0
KLHL31	0	0.124245720337424	0
ANKRD6	0.149435048478206	0	0

SLC22A16	0	0	-0.158890535640587	0
FABP7	0	-0.0913764253682411		0
C6orf118	0.133737693918588	0		0
PCLO	0	0.0151938193938579		0
BAIAP2L1	0	0.0495277368058063		0
KLF14	0	0.0654383847540908		0
PTN	0	-0.03277110324126		0
CLDN23	0	-0.263265569245728		0
DLC1	0	-0.0594067767683222		0
PSD3	0.00374286595521964	0		0
TNFRSF10A	0	-0.0314461522679056		0
FZD3	0.0101763372066693	0		0
NRG1	0	-0.241337383438873		0
RAB11FIP1	0	0	-0.272565878178802	
EYA1	0	0.025697498784019		0
AARD	0	0.036186947585317		0
COLEC10	0	0	-0.135988116823859	
ENPP2	0	0	-0.00362842285135613	
OC90	0	0	-0.244948753218724	
ARC	0	-0.224535181005891		0
MSMP	0	-0.226478680981886		0
IGFBPL1	0.0820153903567099	0		0
AQP7P3	0	-0.0772957800896528		0
AQP7P1	0	-0.0874966201049436		0
LOC102724580	0	-0.0858550380492216		0
BANCR	0	0	-0.18518891349782	
HSD17B3	0	-0.0288799579349813		0
CTSV	0	0	-0.0785377226328385	
FAM225A	-0.0729491981134011	0		0
RGS3	0	-0.149611672386409		0
PAEP	0	-0.0189579658388314		0
NACC2	0	-0.0183704351421454		0
UAP1L1	0	-0.012502922234597		0
TUBBP5	0	0.00787730869198353		0
SHROOM2	0.0132053822348307	0		0
GPM6B	0	-0.0870951857449852		0
NYX	0	0	-0.138920117306013	
PLP2	-0.0519699697969731	0		0
MIR325HG	0	-0.0676602085967656		0
NUP62CL	0.0691301266439607	0		0
HTR2C	0	0	-0.098419093830928	
PLXNB3	0	-0.0471858134877046		0
MYO3A	0.183022119822423	0		0

FZD8	0	-0.0160121038436596	0
TMEM72	0	0	-0.0233313671821323
GPRIN2	0	0.0342066480872751	0
CHAT	0	-0.0544748782384834	0
LRRTM3	0	-0.128473662584448	0
PLAU	0	-0.120770314006533	0
RGR	0	-0.0339866042284005	0
MYOF	0	0	-0.260533353102443
PLCE1	0	-0.0644660454513726	0
LOXL4	0	-0.0661910626535242	0
SH3PXD2A	-0.0339168823645554	0	0
NANOS1	0	0.0057165936713845	0
CPXM2	0	0	-0.19548914039069
TCERG1L	0.074026454491262	0	0
PRAP1	0	0	-0.0500221208430859
H19	-0.164325364479522	0	0
PTH	0	0	-0.0381815012587207
SOX6	0	-0.055735760258123	0
CD82	0	-0.0588016004694076	0
RHOD	0	0	-0.0777812942003884
MAP6	0.0927656348065143	0	0
TYR	0	0	-0.265771844923468
CADM1	0	-0.0480924254358967	0
GRAMD1B	0.0505624803089068	0	0
LINC02387	0	0	-0.0846105957121068
CPNE8	0	0	-0.106103513074003
DHH	0	-0.102229807340012	0
LIMA1	0	-0.0701687996476539	0
KRT75	0	-0.011936916367136	0
HOXC13	-0.215324391250104	0	0
HOTAIR	-0.0316118524341433	0	0
HOXC11	-0.0574518021438738	0	0
HOXC-AS3	-0.0813403317319698	0	0
HOXC10	-0.192717601988815	0	0
HOXC9	-0.130096991281496	0	0
HOXC8	-0.0993057695835197	0	0
HOXC6	-0.179574010316691	0	0
HOXC5	-0.0728674187808542	0	0
AGAP2-AS1	0	-0.0303457834369571	0
RMST	0.117646433290302	0	0
MIR135A2	0.0742717486723684	0	0
ASCL1	0.103661850446058	0	0
TRPV4	-0.0433633439461686	0	0

MSI1	0.185099221978621	0	0
VPS37B	0	-0.0383291145268732	0
LINC00943	0	-0.039993356164721	0
PCDH9	0	-0.049888299692776	0
SOX1-OT	0.199226656682731	0	0
SOX1	0.29621141841776	0	0
FOXG1	0	-0.108419759017052	0
LINC01551	0	-0.218952527948588	0
TRIM9	0	-0.167062745731079	0
OTX2-AS1	0.066953955587401	0	0
FUT8-AS1	0	-0.0676365487215212	0
ACTN1	0	-0.0106556012902143	0
CCDC177	0.0488982389370924	0	0
ISM2	0	0	-0.0204749673709625
KCNK10	0.150561372311663	0	0
TUNAR	0.138251981889417	0	0
SMAD6	0	0.0582294416371613	0
LRRC49	0.00186338797862604	0	0
INSYN1	0	0	0.140893367393846
RGMA	0.0359284965535415	0	0
PYCARD	-0.0324780366918536	0	0
TOX3	0.206754062640283	0	0
MT3	0	-0.0517366323475997	0
FBXL8	-0.0638025244089886	0	0
BCO1	0	0	-0.101112427531018
HSD17B2	0	0	-0.150549337277704
RTN4RL1	0	0.0394142852208422	0
SLC13A5	0	-0.095937371829271	0
TOM1L2	0	-0.0577712210605358	0
KRT16P3	0	-0.040608940820178	0
SEZ6	0	-0.0145880391359106	0
WNT9B	0.130822091589753	0	0
ABCC3	0	-0.0298760407062387	0
MRC2	0	-0.0758660053043581	0
CASC17	0	-0.279925173464946	0
SOX9	0	-0.00362164729595135	0
RNF157	0	-0.00975282850940485	0
UTS2R	0	0	-0.00591171641826056
METRNL	0	0	-0.058115756746708
ARHGAP28	0	0.0453485454551099	0
MTCL1	0	-0.104723984909878	0
APCDD1	0.0405405930087368	0	0
MC5R	0.0290748997550884	0	0

<b>PHLPP1</b>	0	-0.211433153063013	0
<b>ICAM5</b>	-0.0375909427502295	0	0
<b>IL27RA</b>	- 0.00977024439135653	0	0
<b>JAK3</b>	- 0.00109165026318065	0	0
<b>HOMER3-AS1</b>	0	0	-0.0136359043631753
<b>LINC01224</b>	0.119515690392219	0	0
<b>ZNF536</b>	0	0.0392571767288096	0
<b>MIA</b>	0	-0.00921684423631845	0
<b>PPM1N</b>	-0.0388871895647432	0	0
<b>IGFL3</b>	0	-0.00741121618856963	0
<b>EMP3</b>	-0.0183518823204536	0	0
<b>SLC17A7</b>	-0.0888169639417604	0	0
<b>RCN3</b>	-0.0417086510259372	0	0
<b>MYH14</b>	0	-0.139407190729997	0
<b>LENG9</b>	-0.143714314939507	0	0
<b>CDC42EP5</b>	-0.187264563371131	0	0
<b>TGM6</b>	0	0	-0.0138201899634101
<b>ACSS1</b>	0	-0.057814298317406	0
<b>CDH22</b>	0	0.0293669087037099	0
<b>PTGIS</b>	0	0.158763412212702	0
<b>CDH4</b>	0	-0.0231063423322867	0
<b>STMN3</b>	0	0	0.123820422563198
<b>MYT1</b>	0.0536497720486797	0	0
<b>CLIC6</b>	0	0	-0.198773857633616
<b>NEFH</b>	-0.043617015940047	0	0
<b>CDC42EP1</b>	-0.14696873087521	0	0
<b>APOBEC3C</b>	-0.0806881890131919	0	0
<b>MLC1</b>	0	-0.171272495387662	0
<b>PANX2</b>	0	-0.0706532517538317	0

**Titre :** Analyse intégrative clinico-moléculaire et séquençage ARNseq de cellule-uniques afin d'identifier l'origine des tumeurs rhabdoïdes térétoïdes atypiques

**Résumé :** Les tumeurs rhabdoïdes térétoïdes atypiques (ATRT) sont des tumeurs malignes rares et agressives du système nerveux central (SNC) affectant les nourrissons et les jeunes enfants et caractérisées par une inactivation bi-allélique du gène suppresseur de tumeur SMARCB1 dans un génome par ailleurs très simple. Sur la base du profil de méthylation et d'expression, des études récentes ont mis en évidence la diversité moléculaire de ces tumeurs, qui sont maintenant divisées en au moins trois sous-groupes, à savoir les tumeurs dites MYC-, TYR- et SHH-ATRT. Cette diversité moléculaire suggère l'existence de diverses cellules d'origine pour chaque sous-type, bien qu'aucune de ces cellules d'origine n'ait été identifiée à l'heure actuelle. L'analyse du profil d'expression a été faiblement informative sur le tissu d'origine mais a donné des indications sur le lignage (neurogénique et mélanogénique pour les sous-types SHH et TYR respectivement) et sur certains gènes et voies de développement exprimés de manière récurrente (SHH et voie NOTCH pour le sous-type SHH ; groupes HOX pour le sous-type MYC).

Dans ce travail, nous avons fait une description détaillée des différentes localisations anatomiques des ATRT afin d'obtenir des informations précises sur le tissu d'origine. Ensuite, nous avons réalisé une analyse intégrative de la localisation anatomique avec les données transcriptomiques et épigénétiques. Cette analyse nous a permis de décrire et de caractériser 4 sous-groupes anatomico-moléculaires d'ATRT : les tumeurs des nerfs crâniens, du cortex et de la moelle épinière, CNCS-ATRT ; les tumeurs des ganglions de la base et intraventriculaires, BG/IV-ATRT ; les tumeurs du lobe antérieur du cervelet, CAL-ATRT ; et les tumeurs du pédoncule cérébelleux moyen et du vermis cérébelleux inférieur, MCP/ICV-ATRT. Ensuite, nous avons cherché à étudier la pertinence de nos modèles murins chez l'homme. Nous avons constaté que le modèle de souris Rosa26 : MYC récapitule parfaitement les caractéristiques moléculaires et les phénotypes cliniques d'un sous-ensemble d'ATRT humains (CNCS-ATRT) et soutient l'hypothèse d'une origine dans les cellules de la crête neurale. Les tumeurs Rosa26 : SHH sont phénotypiquement similaires aux BG/IV-ATRT et sont caractérisées par l'expression de marqueurs génétiques des éminences ganglionnaires. Enfin, les tumeurs CAL-ATRT humaines sont caractérisées par l'expression de gènes typiques de la région entre le cerveau moyen et le cerveau postérieur pendant le développement embryonnaire. Bien que nous ne disposions pas d'un modèle de souris récapitulant cette localisation, l'analyse intégrative nous a permis d'identifier le sous-groupe CAL-ATRT et nous a conduits à effectuer une analyse RNAseq cellule-unique de trois échantillons humains provenant exclusivement de cette localisation. Nous avons consécutivement démontré au niveau cellule-unique, pour la première fois à ce jour, la dédifférenciation des cellules progénitrices neuronales et l'implication de la voie Notch dans la transformation maligne des CAL-ATRT.

**Mots clés :** ATRT sous-groupes ; corrélation anatomique ; analyse intégrative ; RNAseq cellule-unique; voie de signalisation Notch

**Title :** Integrative Clinico-Molecular Analysis and Single-Cell RNA sequencing To Unravel The Origin Of Atypical Teratoid Rhabdoid Tumors

**Abstract :** Atypical teratoid rhabdoid tumors (ATRT) are rare and aggressive malignancies of the central nervous system (CNS) affecting infants and young children and characterized by a biallelic inactivation of SMARCB1 tumor suppressor gene in an otherwise very simple genome. Based on methylation and expression profiling, recent studies have pointed out the molecular diversity of these tumors, that are now divided in at least three subgroups, i.e. the so-called MYC-, TYR-, and SHH-ATRTs. This molecular diversity is suggestive of various cells of origin for each sub-type, though none of these putative cells of origin is undoubtedly identified at 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).

In this work, we have provided a detailed description of different anatomical ATRT locations to obtain precise information about the putative tissue of origin. Then, we have performed integrative analysis of anatomical location, transcriptomic and epigenetic data. This analysis has allowed us to describe and characterize 4 distinct anatomico-molecular ATRT subgroups: cranial nerves, cerebral cortex and spinal cord tumors, CNCS-ATRT; Basal ganglia and Intraventricular tumors, BG/IV-ATRT; cerebellar anterior lobe tumors, CAL-ATRT; middle cerebellar peduncle and inferior cerebellar lobe tumors, MCP/ICV-ATRT. Next, we have sought to investigate the relevance of our mouse models for human disease. We have found that the Rosa26: MYC mouse model perfectly recapitulates molecular features and clinical phenotypes of a subset of human ATRT (CNCS-ATRT) and support the hypothesis of an origin in neural crest cells. The Rosa26: SHH tumors are phenotypically similar to the BG\_IV-ATRT and are characterized by the expression of gene markers from the ganglionic eminences. Finally, the CAL-ATRT are characterized by the expression of midbrain-hindbrain boundary (MHB) genes. Although we don't have a faithful mouse model for this location, the integrative analysis has allowed us to identify the CAL-ATRT subgroup and lead us to perform single cell analysis of three human samples exclusively from this location. Consecutively, we have demonstrated at the single-cell level, by the first time so far, the dedifferentiation of neuronal progenitor cells and the involvement of Notch pathway in the malignant transformation of CAL-ATRT.

**Keywords :** ATRT subgroups; Anatomical Correlation; Integrative Analysis; Single Cell RNAseq; Notch signalling pathway