



UNIVERSITAT DE
BARCELONA

Function of the histone demethylase PHF8 in neural progenitor cells and glial differentiation

Simona Iacobucci

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Doctoral program in Genetics
Faculty of Biology, University of Barcelona

Research performed at the
Institute of Molecular Biology of Barcelona, CSIC

Function of the histone demethylase PHF8 in neural progenitor cells and glial differentiation

*A thesis submitted by Simona Iacobucci to obtain the doctoral degree by
the University of Barcelona*

Author:
Simona Iacobucci

Directed by:
Dr. María A. Martínez Balbás

Tutored by:
Dr. Pedro Martínez Serra

Barcelona 2021

Contents

List of abbreviations	vi
Introduction	
1. Chromatin	
1.1. <i>Structure and organization</i>	1
1.2. <i>DNA methylation</i>	3
1.3. <i>Histones and post translational modifications: H4K20me1 and H3K9me2</i>	4
1.4. <i>Lysine methyltransferases (KMTs) and demethylases (KDMs)</i>	7
1.5. <i>JmJC domain histone demethylases</i>	9
1.6. <i>Demethylases of KDM7 family</i>	11
1.7. <i>Chromatin and metabolism</i>	12
2. PHF8	
2.1. <i>PHF8 and transcription</i>	15
2.2. <i>PHF8 function in cell cycle and DNA repair</i>	17
2.3. <i>PHF8 and metabolism</i>	19
2.4. <i>PHF8 role in neurodevelopment</i>	20
2.5. <i>PHF8 and X-linked intellectual disabilities</i>	22
2.6. <i>PHF8 and cancer</i>	23
3. Epigenetic mechanisms in neurodevelopment	
3.1. <i>Mouse cortex as a model of study</i>	24
3.2. <i>Neural stem cells as an in vitro model of study</i>	25
3.3. <i>Neurogenesis</i>	27
3.4. <i>Epigenetic mechanisms in neurogenesis</i>	28
3.5. <i>Astroglialogenesis</i>	30
3.6. <i>Epigenetic mechanisms in astroglialogenesis</i>	31
4. Astrocytes	
4.1. <i>Astrocytes and synaptogenesis</i>	33
4.2. <i>Astrocytes and intellectual disabilities</i>	36
Objectives	39
Material and methods	
Reagents	40

1.1. <i>Plasmids</i>	40
1.2. <i>Antibodies</i>	40
1.3. <i>Primers</i>	42
2. Cell culture and differentiation	
2.1. <i>NSCs</i>	43
2.2. <i>HEK 293T</i>	43
2.3. <i>HeLa</i>	43
2.4. <i>Astrocytes differentiation</i>	43
2.5. <i>Cocultures: neurons- astrocytes</i>	44
3. Genetic manipulation of growing cells	44
3.1. <i>Calcium phosphate transfection</i>	44
3.2. <i>Lentiviral transduction</i>	44
4. Molecular biology procedures	45
4.1. <i>Genomic DNA extraction</i>	45
4.2. <i>Phenol chloroform extraction and ethanol precipitation</i>	45
4.3. <i>RNA extraction (for RNA seq and RT)</i>	46
4.4. <i>DNAse treatment</i>	46
4.5. <i>Reverse transcription</i>	46
4.6. <i>qPCR</i>	47
4.7. <i>Mini and Maxi preparations of DNA</i>	47
4.8. <i>DNA electrophoresis</i>	48
4.9. <i>Total protein extraction</i>	48
4.10. <i>Protein quantification by Bradford</i>	49
4.11. <i>SDS-Page electrophoresis</i>	49
4.12. <i>Western blot</i>	49
4.13. <i>MACS isolation of astrocytes</i>	50
4.14. <i>Indirect immunofluorescence and analysis of synapses</i>	50
4.15. <i>Growth curve</i>	51
4.16. <i>Cell cycle analysis by flow cytometry</i>	51
4.17. <i>Luciferase assay</i>	52
4.18. <i>ChIPseq</i>	52
4.19. <i>RNAseq</i>	53
5. Gas chromatography- mass spectrometry (GC-MS)	53
6. Electrophysiology	54

7. Bioinformatic methods	
7.1. <i>ChIP seq analysis</i>	55
7.2. <i>Gene ontology analysis</i>	55
7.3. <i>Gene expression omnibus accessions</i>	55
7.4. <i>ChIP-seq capture obtaining</i>	56
7.5. <i>RNA seq analysis</i>	56
7.6. <i>Heatmap construction</i>	56
8. Statistical analysis	57
8.1. <i>Sample size</i>	57
8.2. <i>Standard deviation and standard error of the mean</i>	57
8.3. <i>Student's t-test</i>	57
8.4. <i>ANOVA and Tukey's multiple comparisons test</i>	57
8.5. <i>Mann-Whitney Rank Sum Test</i>	58

Results

1. Characterization of PHF8 function in astrocytes	59
1.1. <i>PHF8 expression during astrocyte differentiation</i>	59
1.2. <i>PHF8 regulates transcription during astrocyte differentiation</i>	65
1.3. <i>PHF8 binds to astrogenic and synaptogenic genes</i>	73
1.4. <i>PHF8 depleted NSCs differentiate into defective astrocytes</i>	80
1.5. <i>PHF8 depletion impairs neuronal synapses</i>	87
1.6. <i>PHF8 maintains low levels of H4K20me1/3 at astrogenic and synaptogenic genes</i>	93
1.7. <i>PHF8 histone demethylase activity preserves astrocytic transcriptional program</i>	97
2. Characterization of PHF8 function in neural stem cells	101
2.1. <i>PHF8 depleted neural stem cell suffer delay in cell cycle</i>	101
2.2. <i>Analysis of PHF8 depleted neural stem cell transcriptional profile</i>	106
2.3. <i>PHF8 depletion impairs serine biosynthesis</i>	110
2.4. <i>PHF8 directly regulates serine biosynthesis genes</i>	115

2.5. <i>PHF8 cooperates with c-Myc to regulate transcription of serine biosynthesis genes</i>	115
---	-----

Discussion

1. Regarding the function of PHF8 in astrocytes	119
1.1. <i>PHF8 levels are regulated during astrocyte differentiation through Notch signaling</i>	119
1.2. <i>PHF8 directly regulates a key astrogenic gene: Nfia and its depletion impairs astrocytes differentiation</i>	120
1.3. <i>PHF8 directly regulates synaptogenic genes and its depletion in astrocytes causes aberrant synaptogenesis</i>	122
1.4. <i>PHF8 prevents ectopic heterochromatin formation at astrogenic and synaptogenic genes</i>	123
2. Regarding the function of PHF8 in neural stem cells	124
2.1. <i>PHF8 regulates cell cycle and metabolic genes expression</i>	124
2.2. <i>PHF8 depletion impairs serine metabolism</i>	125
2.3. <i>PHF8 regulates the expression of serine biosynthesis genes cooperating with c-MYC</i>	128

Conclusions	129
--------------------	-----

Bibliography	130
---------------------	-----

Appendix	150
-----------------	-----

List of abbreviation

A: ampere
ASD: autism spectrum disorder
ATP: adenosine triphosphate
 α KG: alpha-ketoglutarate
BMP: bone morphogenetic proteins
bps: base pairs
BSA: bovine serum albumin
BrdU: bromodeoxyuridine
ChIP: chromatin immunoprecipitation
ChIP-seq: chromatin immunoprecipitation sequencing
CNS: central nervous system
CTD: carboxy-terminal domain
CTR: control
DIV: days *in vitro*
DNA: deoxyribonucleic acid
DNTP: deoxynucleotide triphosphate
dpf: day-post-fertilization
EdU: 5-Ethynyl-2'-deoxyuridine
EGF: epidermal growth factor
E/I: excitatory/ inhibitory
EMT: epithelial to mesenchymal transition
EP: electroporation / electroporated
ESCs: embryonic stem cells
FBS: fetal bovine serum
FC: fold change
FGF: fibroblast growth factor
GC: gas chromatography
GEO: gene expression omnibus
GF: growth factor
Glu: glutamic acid
GO: gene ontology
gRNA: guide-RNA
GTF: general transcription factors
H: histone
H1-hESCs: human embryonic stem cells
HAT: histone acetyl transferase
HDAC: histone deacetylase
HKMT: histone lysine methyl transferase
HP1: heterochromatin protein 1

ID: Intellectual disabilities
IGV: Integrative Genomics Viewer
IFN γ : interferon gamma
JmJc: JumonjiC
kb: kilobase
KD: knockdown
KDMs: Lysine demethylases
KMTs: Lysine methyltransferase
KO: knockout
LAD: lamina-associated domains
LB: lysogeny broth
LTP: long term potentiation
MACS: magnetic- activated cell sorting
mEPSCs: miniature excitatory postsynaptic currents
MS: mass spectrometry
mPTP: mitochondrial permeability transition pore
Mut: mutant
NGS: next generation sequencing
NSCs: neural stem cells
PBS: phosphate buffered saline
PcG: Polycomb group of proteins
PCR: polymerase chain reaction
PNK: T4-polynucleotide kinase
Poly-A: polyadenylic
PTMs: post-translational modifications
RAR: retinoic acid receptor
rep: replicate
RG: radial glial cells
RMT: arginine methyltransferase
RNA-seq: RNA sequencing
RNAPII: RNA-polymerase II
ROIs: Regions Of Interest
ROS: *reactive oxygen species*
RPM: reads per million
RT: retro transcription
SD: standard deviation
SDS: sodium dodecyl sulfate
Ser: serine
SEM: standard error of the mean
Shh: Sonic hedgehog
shRNA: short hairpin RNA

SVZ: subventricular zone
TAD: topologically associating domain
TBE: Tris-borate-EDTA
TBP: TATA-box binding protein
TET: ten eleven translocations
TFs: transcription factors
TGF β : transforming growth factor beta
TSS: transcriptional start site
U2OS: human osteosarcoma cells
UPR: unfolded protein response
V: volts
VZ: ventricular zone
WT: wild type
XLID: X-linked intellectual disability

Introduction

The development of nervous system depends on the coordination between signaling pathways, specific transcription factors and epigenetic regulators, which all together orchestrate the gene expression program of each cell type. We explored the relevance of specific epigenetic mechanisms in regulating astrocytes differentiation and neural stem cells (NSCs) function. Specifically, we investigated how the histone demethylase PHF8 regulates astrocytes and NSCs state.

1. Chromatin

1.1. *Structure and organization*

In eukaryotic cells the DNA, around 2 meters long, must be packaged into chromatin to fit inside the nucleus [1], that have a diameter of 5- 10 μm . The repeating structural unit of chromatin is the nucleosome which is composed of the core particle and linker DNA [2]. The core particle consists of 147 base pairs (bps) of DNA wrapped around an octamer of histones, with two copies of each H2A, H2B, H3 and H4. The linker DNA is bound by the histone H1, that holds the cores together [3]. The repeating nucleosomes form a flexible 10 nm fiber which represents the first level of DNA compaction; in this way the DNA is about seven times shorter than it would be without the histones [4]. In the traditional view the next level of compaction is the 30 nm chromatin fiber as nucleosomes and linker DNA are coiled between them (Figure I1) [5, 6]. This is currently being challenged by the fluid-like model that views the chromatin as a dynamic structure based on the irregular 10-nm fiber [7, 8]. However, chromatin topology during interphase is not randomly organized, it shapes structures of growing organization called loops, TADs, compartments and chromosomal territories [9] (Figure I2). Chromatin forms long-range interactions in which two distant DNA segments are brought close to each other forming a loop. These loops can be long few kilobases until more than 100 mega bases and they can be really stable over time or not [10]. Over loops, chromatin is organized in TADs which are mega base-sized topologically associating domains observed in Hi-C. TADs indicate high frequency interactions between loci within a domain, and low frequency contacts between loci in different domains [11].

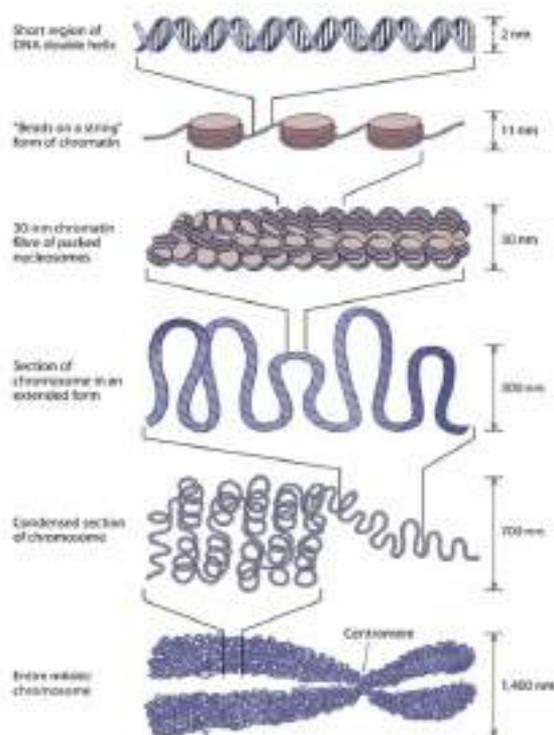


Figure 11: State of compaction of chromatin; from nucleosomes to the 30 nm fiber and higher-ordered structure. Adapted from [12].

The contact maps obtained in the original Hi-C study showed a very characteristic plaid-like pattern with alternating blocks of enriched and depleted interaction frequencies [13]. On this basis, it was suggested that chromatin is subdivided into two DNA sets named compartments A and B: loci from one set preferentially contact other loci from the same set. A compartment correlates with gene density, transcriptional activity, chromatin accessibility, and activating chromatin marks; B compartment correlates with lamina-associated domains (LADs) and late replication timing, which suggests a proximity to the nuclear periphery [14]. Finally, it has been defined the chromosome territories that occupy preferential positions within the nuclei and relative to each other, those chromosomal territories can be observed both by microscopic and chromosome conformation capture techniques (also named 3C methods) [15, 16].

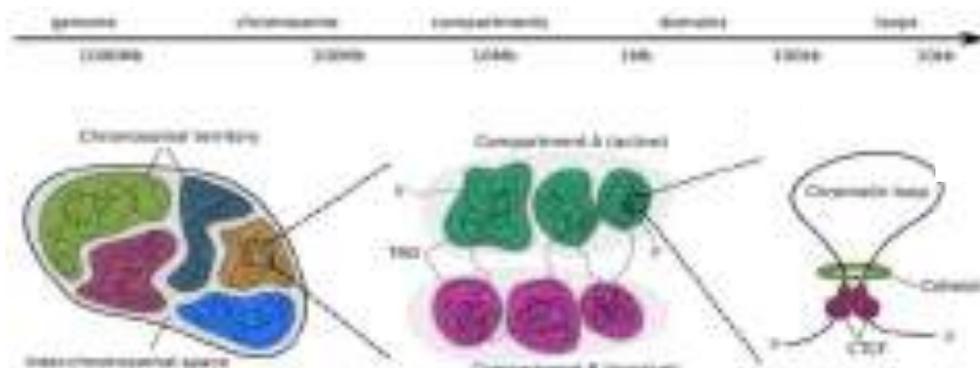


Figure 12: Hierarchical genome organization at different scales: chromatin loops, compartments A and B and chromosomal territory. Adapted from [9].

1.2. DNA methylation

DNA methylation is a major epigenetic modification of vertebrate genomes associated with transcriptional repression [17]. The most striking feature of DNA methylation patterns are the CpG islands, unmethylated GC-rich regions with high densities of CpG positioned at the 5' ends of many genes [18]. A proportion of CpG islands become methylated during development, therefore the associated promoter become stably silent [19]. Many enzymes can modify the DNA methylation pattern with different mechanisms to maintain, gain or lose DNA methylation. The maintenance of methylation reproduces DNA methylation pattern between cell generations through the DNMT1 (DNA methyltransferase 1) that methylates those new CpGs whose partners on the parental strand already carry a methyl group [20]. De novo methylation is mediated by DNMT3A and DNMT3B [21], which are highly expressed in embryonic cells when the majority of de novo methylation events occur. DNA methylation contributes to transcriptional repression as the methyl group can directly interfere in the binding of proteins to DNA sequence, hence some factors fail to bind DNA. DNA methylation also induces the recruitment of other factors such as MBD1, MBD2, MBD3 (methyl-CpG-binding domain protein), and MeCP2 (methyl-CpG binding protein 2) that promote transcriptional repression [22].

DNA demethylation can be both active or passive; active DNA demethylation removes or modifies the methyl group from 5-methylcytosine (5mC), passive DNA demethylation consists in the loss of 5mC during successive rounds of replication in the absence of

functional DNA methylation machinery. Many factors can promote DNA demethylation: DNA cytosine deaminases, DNA glycosylases, DNA repair factors and even DNA methyltransferases [23]. In mammals, active DNA demethylation is achieved through ten eleven translocation (TET) enzymes that oxidize 5-methylcytosine [24, 25]. 5-hydroxymethylcytosine is a key point in demethylation as it can be passively consumed through DNA replication or actively reverted to cytosine through iterative oxidation and base excision repair [26].

In mammalian genes, DNA methylation adds stability to the repression of transcription when located at the start sites of genes, it also regulates other regions such as enhancers and insulators, and it is essential to embryonic development [27].

1.3. Histones and post translational modifications: H4K20me1 and H3K9me2

Histone are basic proteins highly conserved across organisms and composed by five major families: H1/H5, H2A, H2B, H3 and H4. Each core histone contains two common regions; the “histone fold” and the “histone tail”[28]. The histone fold is responsible for the formation of stable H2A–H2B and H3–H4 dimers, while the N-terminal tail is a flexible region that protrude from the core and interact with DNA [29]. The residues of the tails are targets of post-translational modifications and, in this way, they regulate transcription, replication, recombination and DNA repair. Histone modifications include acetylation, methylation and ubiquitination on lysine, methylation and citrullination on arginine, and phosphorylation on serine, threonine and tyrosine (Figure I3). Acetylation and methylation on specific lysine residues are especially involved in epigenetic gene regulation [30]. Lysine acetylation and deacetylation is a dynamic process carried out by histone acetyl transferases (HATs) and histone deacetylases (HDACs). HATs utilize the acetyl CoA as cofactor to catalyze the addition of an acetyl group to the ϵ -amino group of the lysine. This reaction weakens the interaction between histones and DNA, so HATs usually work as coactivators and HDACs as corepressors [31]. Major acetylation sites on histone H3 include K9, K14, K18, K23 and K27, and correlate with transcriptional activation, as they are localized to transcription start sites and enhancers of genes.

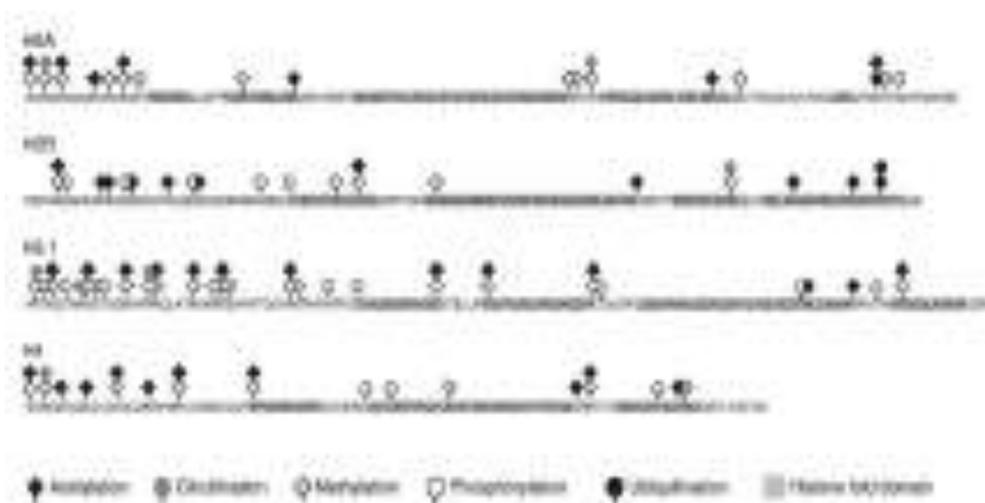


Figure I3: Modifications on core histones. Adapted from [30].

While, the correlation of methylation on H3 with transcriptional activation or repression depends on its level and the specific residue involved (Figure I4); trimethylation of H3K4 is associated with activation while trimethylation of H3K9 is associated with repression, for example. Histone methylation does not change the electrical charge of the amino acid, but it has functional consequences [32]. It consists in the addition of $-CH_3$ groups to lysines or arginines, through histone lysine methyl transferases (HKMTs) or arginine methyltransferases (RMTs). Lysines can be mono-methylated (me1), di-methylated (me2) and tri-methylated (me3) and different states of methylation correlate with transcriptional activation or repression depending on the genomic localization. Among all the histones modifications, we were particularly interested in H4K20me1 and H3K9me2. H4K20me1 regulates diverse cellular processes: DNA damage response, transcriptional regulation, mitotic condensation and DNA replication. H4K20me1 levels are highly regulated during the cell cycle; low during G_1 phase, resulting in a very low level of H4K20me1 in the beginning of S phase, it accumulates during S and G_2 phases resulting in a peak in M phase [33]. H4K20me1 is enriched both at promoters and at the gene bodies [34], and it has been associated both with activation and repression of transcription. It is essential during development and alterations in H4K20me1 deposition are associated with a variety of diseases ranging from cancer to developmental disorders, like the Meier-Gorlin syndrome [35].

Modification of histone	Mono-methylation	Di-methylation	Tri-methylation	Acetylation
H2AK5				Activation
H2AK7				Activation
H2AK9				Activation
H2AK13				Activation
H2BK5				Activation
H2BK11				Activation
H2BK15				Activation
H2BK20				Activation
H2BK29				Activation
H3R2	Activation			
H3K4	Activation	Activation	Activation	
H3K9	Activation/repression	Repression	Activation/repression	
H3K14				Activation
H3R17	Activation			
H3K18				Activation
H3K25				Activation
H3R26	Activation			
H3K27	Activation	Repression	Repression	Activation
H3K36		Activation	Activation	
H3K36				Activation
H3K79	Activation	Repression	Repression	
H3K115				Activation
H4R3		Activation		
H4K5				Activation
H4K8				Activation
H4K12				Activation
H4K16				Activation
H4K20	Activation/repression	Repression	Repression	Activation
H4K39	Repression			
H4K91				Activation

Figure 14: Major modifications of histone and their genetic regulations. Adapted from[36].

It has been proposed also that the role of H4K20me1 may be context dependent. For example, the presence of H4K20me1 at highly expressed genes can be affected by the co-occurrence of neighboring H4K16 acetylation (H4K16ac), which has an established role in activating gene expression [37]. The H4K16ac is a posttranslational modification involved in DNA damage repair [38] and is enriched on active enhancers [39]. Experiments in embryonic stem cells (ESCs) show that, upon differentiation, H4K16ac is reduced at TSS, although the global abundance of the modification seems unchanged.

H4K20me2, as H4K20me1, has been shown to play a role in the cell cycle control, particularly marking points of origin for DNA replication. It recruits the DNA replication licensing machinery through the ORC1-BAH domain [40] and it is enriched at sites of DNA damage, recruiting 53BP1 to DNA double strand breaks [41].

While, H4K20me1 and H4K20me2 are involved in DNA replication and DNA damage repair, H4K20me3 is a hallmark of silenced

heterochromatic regions. It is highly enriched at pericentric heterochromatin, telomeres, imprinted regions and repetitive elements, indicating that this modification is involved in transcriptional silencing [42]. H3K9me1 is enriched at promoters and 5' UTRs, with decreasing levels at coding regions of active genes and minimal enrichment at non-genic regions. H3K9me1 could act as a buffer between activation and repression by allowing rapid methylation and demethylation [43].

H3K9me2 demarcates heterochromatin, particularly the non-genic regions. H3K9me2 is prevalent in gene deserts, pericentromeric and subtelomeric regions, while it is almost depleted at active genes [44]. The H3K9me2 distribution occurs in large tracts of several megabases enclosing both non-coding and gene containing DNA. H3K9me2 domains strongly correlate with Lamin B1 binding. Mapping of Lamin B1 domains demonstrated that inactive genes are preferentially located at the nuclear periphery in association with the nuclear lamina [45]. H3K9me2 is also critical in neurodevelopment [46]; it increases across the genome as cells differentiate and acquire lineage specificity [47].

H3K9me3 correlates with repressed genes and can be located at the promoters, coding regions of genes and non-genic regions [48]. *Chromatin immunoprecipitation* (ChIP)-sequencing analyses demonstrated that H3K9me3 is prevalent at many non-genic regions including the repetitive satellite DNA, centromeric and pericentromeric DNA and long terminal repeats of transposons [49]. H3K9me3 exists in large blocks combined to H4K20me3, especially in pericentromeric chromatin; both H3K9me3 and H4K20me3 are critical for the formation of senescence associated heterochromatin foci (SAHF) [50]. Studies of knockout mice for SUV39H1 and SUV39H2, the HMT responsible for these modification (see below), proved that loss of the H3K9me3 mark results in genomic instability and cancer predisposition.

1.4. *Lysine methyltransferases (KMTs) and demethylases (KDMs)*

Lysine methyltransferases

Methylation of lysines on histone and non-histone proteins is generated by protein lysine methyltransferases (also known as 'writers') and removed by protein lysine demethylases (also known as 'erasers'). Lysine methylation facilitates protein-protein interactions, regulates protein- DNA interactions and can affect the stability and subcellular localization of proteins. In humans, there are two domains with annotated lysine methyltransferase activity: the SET domain and the seven-beta-strand domain [51]. Exist 55 SET-domain-containing proteins; half are active

KMTs, one (SETD3) is a histidine methyltransferase, and the enzymatic activities of the remainder are not known. The three different methyl states of H4K20 are generated by three distinct SET KMTs: SETD8, SUV4-20H1, and SUV4-20H2.

In humans, exist five H3K9 methyltransferases with different catalytic activities and target genes: SUV39H1, SUV39H2, SETDB1, G9a and GLP (Figure 15)[52].

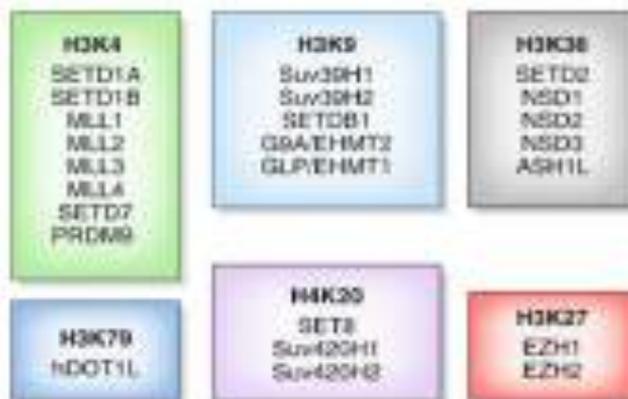


Figure 15: Human histone KMTs categorized by their established substrate specificity. Adapted from [51].

Lysine demethylases

Many histone lysine demethylases have been discovered and they cover most of the lysine methylation sites: H3K4, H3K9, H3K27, H3K36 and H4K20 (Figure 16). The KDMs are divided into two families based on sequence conservation and catalytic mechanism. Lysine demethylation mediated by FAD dependent amine oxidases (KDM1s) uses a flavin adenine dinucleotide to catalyze the demethylation. The second family is formed by proteins containing the Jumonji C (JmjC) domain. These enzymes are Fe (II) dependent and they catalyze the demethylation of mono, di and trimethylated lysines using 2-oxoglutarate and oxygen. The reaction converts the methyl group in the lysine to a hydroxymethyl group, subsequently released as formaldehyde. More recently, it has been discovered the existence of other two proteins that act as demethylases. hHR23A, a well-known DNA repair protein [53], and hHR23B were identified as histone H4K20 demethylases performing a screening of cDNA library containing nuclear proteins. Overexpression of hHR23A reduces the levels of H4K20me_{1/2/3} in cells. *In vitro*, it specifically demethylates H4K20me_{1/2/3} and generates formaldehyde. The

enzymatic activity requires the cofactors Fe (II) and α -ketoglutarate and the Ubiquitin-Associated domain [54].

FAMILY	KDM	OTHER NAMES	HISTONE SUBSTRATE
LSD	KDM1A	KDM1, KIAA0601, LSD1, AOF1	H3K4me3/me1, H3K9me2/me1
	KDM1B	LSD2, AOF2, Obox1192	H3K4me3/me1, H3K9me2/me1
JMJC	KDM2A	JHDM1A, KIAA1064, COX2, FBX7, FBX11	H3K36me3/me1
	KDM2B	JHDM1B, COX2, FBX12, FBX15, POC2	H3K4me3, H3K36me2
	KDM3A	JHDM3A, JADC1, JAD3A, KIAA0742, TSCA	H3K9me3/me1
	KDM3B	JHDM3B, JADC2, KIAA1040, Obox7	H3K9me3/me1
	JMJDC1	JHDM1C, KIAA1395, TRFR	H3K9me3/me1
	KDM4A	JHDM4A, JADCA, KIAA0877	H3K27me3/me2, H3K36me3/me2, H1.4K26me3
	KDM4B	JHDM4B, JADCB, KIAA0876	H3K27me3/me2, H3K36me3/me2, H1.4K26me3
	KDM4C	JHDM4C, JADDC, KIAA0780, GASC1	H3K27me3/me2, H3K36me3/me2, H1.4K26me3
	KDM4D	JHDM4D, JADDD, FLJ10281	H3K27me3/me2, H1.4K26me3/me2
	KDM5A	JHDM5A, RBP2	H3K4me3/me2
	KDM5B	JHDM5B, RLL1	H3K4me3/me2
	KDM5C	JHDM5C, SMC5	H3K4me3/me2
	KDM5D	JHDM5D, SMC7	H3K4me3/me2
	KDM6A	LTX	H3K27me3/me2
	KDM6B	JLDC, KIAA0346	H3K27me3/me2
	LTX		
KDM7A	JHDM7A, KIAA1718, KDM7	H3K9me3/me1, H3K27me3/me1	
PHF1	JHDM7B, KIAA1111, ZNF422	H3K9me3/me1, H3K27me1	
PHF2	JHDM7C, KIAA0662	H3K9me3/me1, H3K27me3	

Figure I6: Histone demethylases and their substrate specificities. Adapted from [55].

1.5. *JmJC* domain histone demethylases

JumonjiC domain-containing proteins is a class of histone demethylases that directly reverse histone methylation through an oxidative reaction that requires Fe (II) and α -ketoglutarate as cofactors (Figure 17) [56]. These proteins are evolutionarily conserved in species spanning from yeast to human. They are involved in many physiological and pathological processes, such as embryonic stem cell renewal [57], neural stem cell differentiation [58], X-linked mental retardation [59], tumors [60] and metabolic gene expression [61].

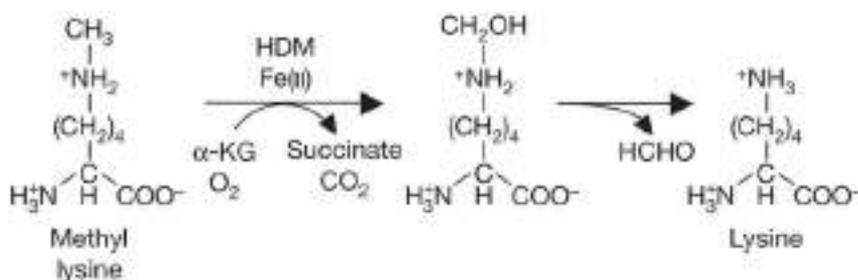


Figure 17: Chemical reaction of lysine demethylation by a JmjC-domain containing KDM. Adapted from [56].

In humans 24 JmjC-domain containing genes encode for proteins with demethylase activity. Their function is further distinguished by the combination of other conserved domains including the PHD, Tudor, CXXC, FBOX, ARID, LRR, and JmjN domains. Based on sequence homologies and structural similarities, these 24 JmjC-domain containing demethylases can be categorized into seven protein subfamilies (Figure 18) [62].



enzymatically inactive due to this. However, it has been demonstrated that PHF2 becomes active when it is phosphorylated by the kinase PKA [67].

KDM7 family are not able to demethylate trimethylated histones since their active centers are too small to accommodate trimethylated lysines [68]. It has been shown that the trimethylation of H3K4, bound by the PHD domain, enhances PHF8 activity toward H3K9me2 [69]. On the contrary, little is known about the properties of the C-terminal halves of KDM7 proteins. They do not contain any known protein domains, and the homology between the three human proteins is low in this region compared to that of PHD and JmjC. Anyway, it was demonstrated that direct association of PHF8 with the carboxyterminal domain of RNA polymerase II largest subunit strongly depends on this part of the protein [70]. The C-terminal domains contain nuclear localization signals, and both PHF2 and PHF8 exhibit several phosphorylation sites which appear to be important for the regulation of their activity. These enzymes are involved in multiple pathologic processes, including cancers and intellectual disabilities [71].

1.7. Chromatin and metabolism

Metabolites are not only the fuel of the cells; they can influence gene expression through chromatin regulation and control cellular mechanisms like proliferation and differentiation. On the other hand, epigenetic factors regulate the expression of enzymes involved in metabolic programs that will generate metabolites that, in turn, will affect the activity of chromatin regulators. It is considerable that many metabolic intermediates are substrates or cofactors of enzymes like DNA and histone methyltransferases/ demethylases or histone acetyltransferases/ deacetylases. Thus, chromatin modifications are especially responsive to metabolic inputs and genes encoding metabolic enzymes are, in turn, directly regulated by epigenetic enzymes (Figure 110) [72].

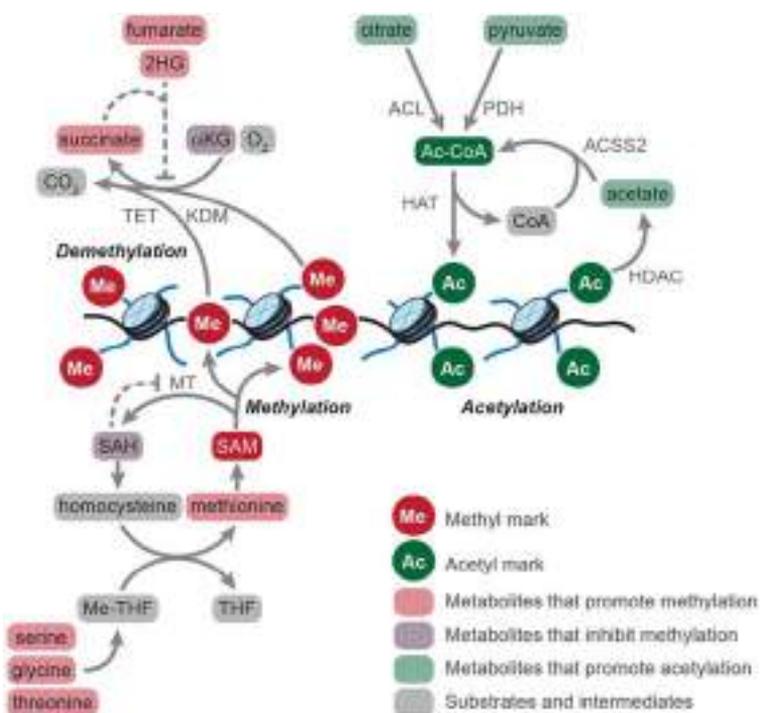


Figure 110: Metabolites are substrates of chromatin modification enzymes; methyltransferases transfer methyl groups from S-adenosylmethionine to histones and DNA. Lysine demethylases and ten-eleven translocation enzymes catalyze demethylation of histone and DNA, respectively, using alpha-ketoglutarate (α KG) [72].

In connection with the topics covered in this thesis, it is worth considering that JmjC-domain-containing histone demethylases like PHF8 are in fact α -ketoglutarate (α KG) dependent dioxygenases. Thus, the interplay between the metabolic pathways that regulate and consume α KG collectively influences the activity of JmjC-domain histone demethylases. This has many implications in stem cells and cancer state; it has been demonstrated that an increase of α KG induce the loss of repressive chromatin modifications and promote self-renewal [73], whereas cancer cells, which have depletion of extracellular glutamine and a concomitant decrease in intracellular α KG levels, show increase in repressive histone methylation [74]. Many metabolic pathways altogether define the state of a cell and highly proliferative cells like neural stem cells and many cancer cells that often require exogenous supply of amino acids for optimal growth, largely depend on non-essential amino acids [75]. In particular, serine has been revealed as an important amino acid in proliferative cells.

Serine can be taken up into the cell using a number of different transporters or can be synthesized by the cell. The process of *de novo* synthesis is particularly important in the brain, where there is a high demand for D-serine as a neurotransmitter, but the transport of plasma serine is complicated due to the blood–brain barrier [75]. Serine is synthesized through SSP, which begins with the glycolytic intermediate 3-phosphoglycerate (3-PG). 3-PG is converted to 3-hydroxypyruvate by the action of the enzyme, phosphoglycerate dehydrogenase (PHGDH) and then 3-hydroxypyruvate (3P-pyruvate) takes part in a transamination reaction with glutamate catalyzed by the enzyme, phosphoserine aminotransferase (PSAT1), resulting in production of phosphoserine (3P-serine) and alpha-ketoglutarate. Phosphoserine is dephosphorylated by the action of phosphoserine phosphatase (PSPH) and produces serine. Serine is mutually converted into glycine by serine hydroxymethyltransferases (SHMTs), which have two isoforms; SHMT1 in cytoplasm and SHMT2 in mitochondria (Figure I11) [76].

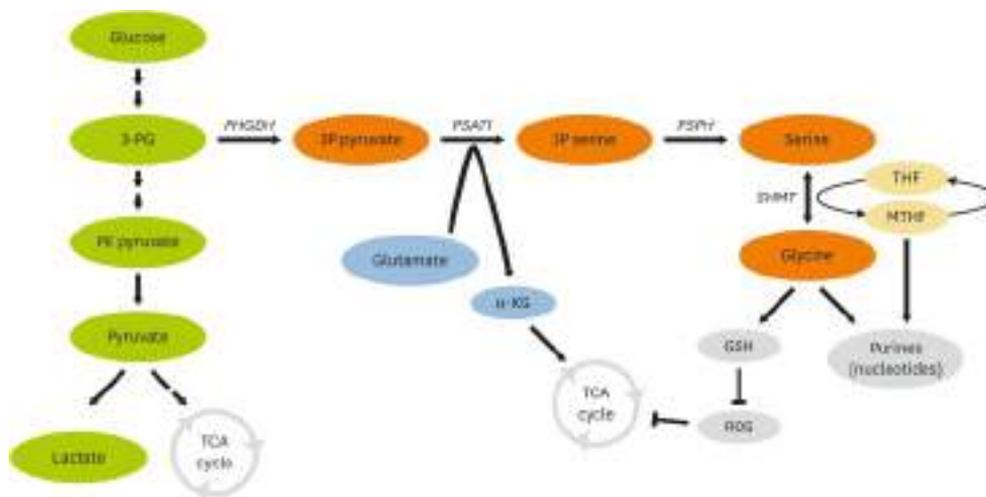


Figure I11: Serine biosynthesis pathway. Adapted from [76].

Serine metabolism is linked to one-carbon metabolism, which influences epigenetic patterns through production of SAM. SAM derived from methionine is the major methyl donor in cellular methyl transfer process including DNA/RNA methylation.

There is limited evidence to show that epigenetic modifiers directly regulate SSP. However, a few studies recently suggested several possibilities [77]. Firstly, the histone methyltransferase EHMT2 (G9a)

regulates SSP. Loss of EHMT2 by activity inhibition or silencing decreases expression of SSP-related genes, *PHGDH*, *PSAT1*, *PSPH*, and *SHMT1/2*, by reducing mono-methylation and increasing di-methylation at histone H3 lysine 9 (H3K9). In turn, the suppressed SSP reduces the concentration of serine and glycine, leading to cell death. Thus, EHMT2 in cancer provides serine and glycine to support cell proliferation by increasing expression of SSP-related genes. Another epigenetic modifier that regulates SSP is KDM4C. KDM4C is a histone demethylase, targeting histone H3K9. It acts on H3K9 tri-methylation at the promoter of the *ATF4* gene and activates the expression [78].

2. PHF8

2.1. *PHF8 and transcription*

PHF8 is a ubiquitously expressed nuclear protein whose dysfunction is implicated in many cancers and neurodevelopment diseases. Human PHF8 consists of 1060 amino acids (the Jmj-C domain covers the residues 231 to 387), while murine Phf8 encodes a 1023 amino acid protein and shares 95% of identity with human PHF8. Studying PHF8 localization in the genome of HeLa cells, it has been observed that PHF8 and H3K4me3 tend to colocalize at the transcription start sites (TSS) of genes. When PHF8 binds H3K4me3 promoters through its PHD domain, it recruits RNA polymerase II and activates transcription [70]. Fortschegger K., et al. propose a model by which increased H3K4me3 levels lead to PHF8 recruitment to gene promoters and the association with RNA polymerase II stabilizes the preinitiation complex formation, leading to enhanced transcription (Figure I12). It has been demonstrated that PHF8 regulates also ribosomal RNA transcription; PHF8 interacts with RNA polymerase I and activates rDNA transcription. In this way it promotes cell growth and proliferation in human osteosarcoma cells (U2OS). The authors of the paper mentioned propose that PHF8 cooperates with H3K4 methyltransferases to establish the transcriptionally active state and prevent the invasion of the repressive mark H3K9me2 [69].

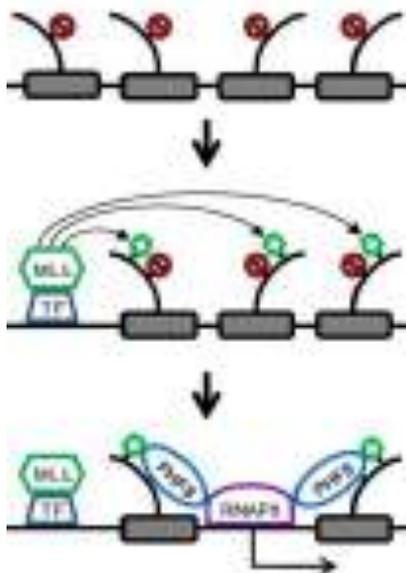


Figure I12: Model for PHF8 coactivator function. Inactive chromatin carries repressive chromatin marks; upon induction, transcription factors (TF) bind to the transcription start site, and bring in H3K4 methylation complexes MLL. PHF8 binds to H3K4me3 marks, removes H4K20me and/or H3K9me2 and helps basal transcription factors to recruit RNA Polymerase II, coactivating transcription. Adapted from [70].

They also show that a point mutation in the JmjC domain of PHF8 (F279S) abolished the demethylase activity towards rRNA genes. The mutation F279S has been found in families with X-linked mental retardation, suggesting that a dysfunction in the JmjC domain may lead to intellectual disabilities.

Although the best-known role for PHF8 is in gene activation, others data in the literature demonstrate that it can contribute to the transcriptional silencing. Through bioinformatic analysis, PHF8 has been shown to co-localize with the corepressor REST/NSRF [79] and H4K20me1 is a mark of transcriptional activation when located in the gene body, thus PHF8, demethylating H4K20me1, could act as repressor. Using affinity purifications and mass spectrometry, our laboratory showed that PHF8 interacts with SIN3A and HDAC1. Before interferon gamma ($\text{IFN}\gamma$) stimulation, PHF8 is bound to a subset of $\text{IFN}\gamma$ -responsive promoters in association with HDAC1 and SIN3A; in this way it keeps the promoters in a silent state maintaining low levels of H4K20me1. Upon $\text{IFN}\gamma$ treatment, PHF8 is phosphorylated by ERK2 and removed from the promoters; in this way H4K20me1 levels increase activating transcription [80].

2.2. PHF8 function in cell cycle and DNA repair

Many histone demethylases shape the epigenome regulating important nuclear processes such as cell cycle progression and DNA repair [81] (Figure I13). As H4K20me1 is an essential histone mark regulating cell cycle progression, PHF8 has a great impact in cell cycle.

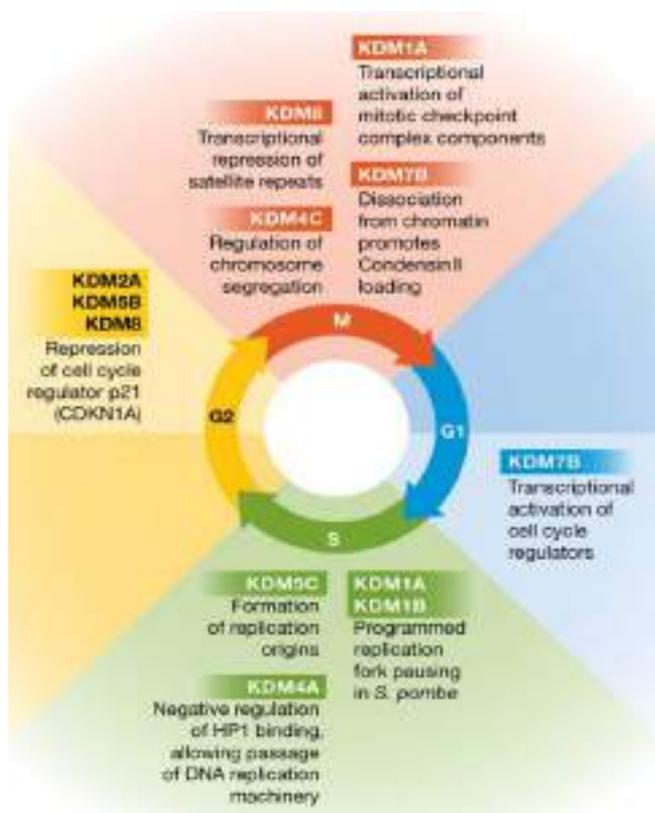


Figure I13: KDMs contribute to the establishment of chromatin states that are required for the expression of important cell cycle regulators, DNA replication, segregation of chromosomes, and genomic stability during cell division. Adapted from [81].

PHF8 controls G1/S transition removing the H4K20me1 mark at TSS of a subset of E2F1- regulated gene and the absence of PHF8 causes a delay in G1–S transition. Proper PHF8 release from chromatin in prophase is necessary for the accumulation of H4K20me1 that therefore can load the condensin II complex [82].

In contrast to the paper mentioned above, Lim H.J. et al. showed that PHF8 protein levels are regulated during the cell cycle and are especially high during G₂ phase and mitosis. Purifying PHF8 by mass spectrometry,

the authors identified many subunits of the anaphase-promoting complex (APC) interacting with PHF8. APC is an E3 ubiquitin ligase that marks cell cycle proteins for degradation and regulates post translationally PHF8 itself. PHF8 high levels in G₂ phase permit the transcriptional regulation of genes involved in G₂/M transition, so PHF8 loss leads to prolonged G₂ phase and defective mitosis (Figure I14). Hence, they conclude that PHF8 plays a relevant role in transcriptional activation of key G₂/M genes [83].

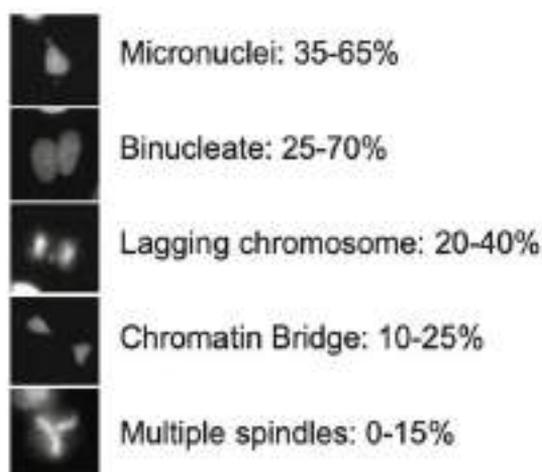


Figure I14: Images showing examples of defective mitosis in PHF8 RNAi cells, with the range of percentages of these defects being observed across all the PHF8 shRNAs. Adapted from [83].

Another paper showed that cyclin E-CDK2, that plays a critical role in G₁/S transition, phosphorylates PHF8 to stimulate its demethylase activity. In this way it promotes cell cycle progression and transcription of cyclin E, E2F3, and E2F7 [84]. In *C. elegans*, two PHF8 homologs, JMJD- 1.1 and JMJD-1.2, have been shown to play a role in genomic stability [85]. The authors of the paper produced mutants of both homologs and exposed them to various types of DNA damage. They found that both mutants had hypersensitivity to interstrand DNA crosslinks, while only one mutant resulted in hypersensitivity to double- strand DNA breaks. The authors speculate that the higher level of heterochromatin in the mutant worms, both before and after DNA damage, is one of the main factors inhibiting DNA repair, and that JMJD- 1.1 influences homologous recombination by relaxing heterochromatin structure or indirectly regulating the expression of genes affecting DNA repair. More recently it has been demonstrated that PHF8 interacts with TopBP1-binding protein, also known as ATR

activator (ATR is a master regulator of the DNA-damage response). The PHF8/TopBP1 interaction is regulated during cell cycle and mediated by CK2 kinase that phosphorylates PHF8 at Ser854. Thus, PHF8pSer854 stabilizes TopBP1 protein levels and ensures replication fork restart, a crucial mechanism of recovery from replication stress that maintains genome stability [86].

2.3. PHF8 and metabolism

Two publications revealed a role for PHF8 in metabolic processes. The first showed that PHF8 is involved in the reprogramming of somatic cells mediated by mitochondrial permeability transition pore (mPTP). The authors demonstrated that, during reprogramming, mitochondrial *reactive oxygen species* (ROS), associated with mPTP opening, lead to PHF8 upregulation and consequent H3K9me2 and H3K27me3 demethylation of pluripotent genes (Figure I15). PHF8 catalyzes lysine demethylation through an oxidative reaction that requires alpha-ketoglutarate (α -KG) as cofactor, and α -KG levels are elevated upon the opening of mPTP, so α -KG may contribute to increase PHF8 activity [87].

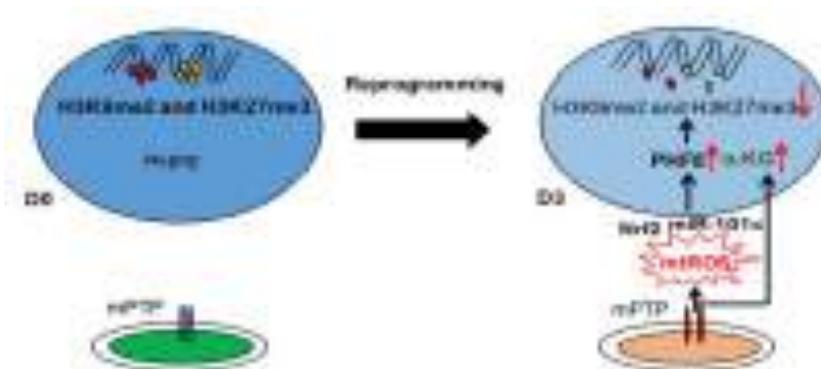


Figure I15: Scheme showing the role of mPTP opening during the early phase of reprogramming. Adapted from [87].

In the second paper, the authors showed that in *C. elegans* *jmjd-1.2* and *jmjd-3.1* coordinate the transcriptional response to mitochondrial stress. *Jmjd-1.2* is homolog to the mammalian PHF8 and can demethylase H3K9/K23/K27me2, while *jmjd-3.1* is homolog to JMJD3 and UTX-1.

Mitochondrial stress, if it's mild, can have beneficial effects on the lifespan as it activates the unfolded protein response (UPR^{mt}), which is a stress signaling mechanism that ensure mitochondrial homeostasis. In *C. elegans* mild mitochondrial stress during larval development delays aging and maintains UPR^{mt} signaling, suggesting that an epigenetic mechanism modulates both longevity and mitochondrial proteostasis. The authors showed that *jmjd-1.2* and *jmjd-3.1* are necessary for the induction of UPR^{mt} and the extension of lifespan mainly removing repressive H3K27me2/3 mark [88].

2.4. *PHF8* role in neurodevelopment

Nonsense and missense mutations of PHF8 have been linked with X-linked intellectual disability (XLID) [89, 90], hence, it is of interest to understand how PHF8 can affect neurodevelopment. A paper from 2010 showed that PHF8 interacts with and functions as a coactivator for retinoic acid receptor (RAR). Knockdown of PHF8 in mouse embryonic carcinoma cells impairs retinoic acid-induced neuronal differentiation. The overexpression of wild-type PHF8, but not of the catalytic F279S mutant, forces cells to neuronal differentiation [91]. Zebrafish PHF8 regulates cell survival and it is found mainly in the head region at 1 day-post-fertilization (dpf) and in the jaw of the embryo at 3 dpf. The authors of the paper showed that PHF8 is critical in brain and jaw development (Figure I16), which suggests the involvement of PHF8 mutations in craniofacial deformities and mental retardation also in human [92].

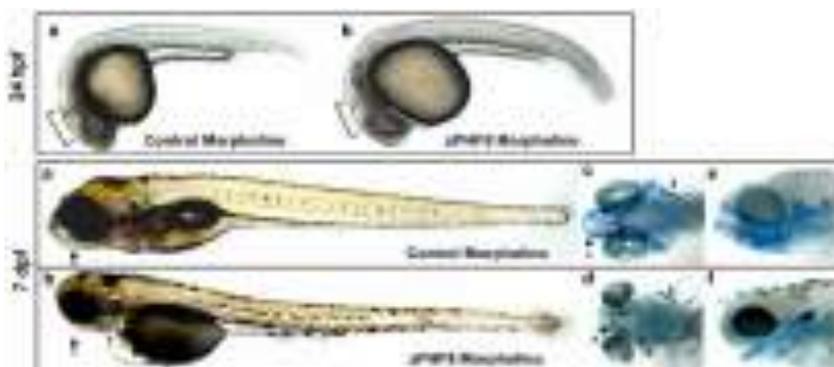


Figure I16: Zebrafish embryos injected at the 1 cell stage with 250uM of control or zPHF8 morpholino (MO). At 24 hours post fertilization (hpf), brain development was delayed. At 7 days post fertilization (dpf) PHF8 MO embryos displayed craniofacial developmental abnormalities including stunted lower jaw. Adapted from [92].

Kleine-Kohlbrecher et al. demonstrated that PHF8 interacts with ZNF711, a protein related to X-linked intellectual disabilities (XLID) too, in human neuroblastoma cells. ZNF711 binds to a subset of PHF8 target genes, including another XLID gene JARID1C.

The *C. elegans* PHF8 homolog is highly expressed in neurons and its inactivation leads to uncoordinated locomotion in *C. elegans* followed by a strong global increase in H3K9me2 and H3K27me2 marks [93].

In 2012, our laboratory demonstrated that PHF8 controls the expression of genes involved in cell adhesion and cytoskeleton organization such as RhoA, Rac1 and GSK3b. A lack of PHF8 in neurons results in a disorganized actin cytoskeleton, impaired cell adhesion and deficient neurite outgrowth as it leads to down-regulation of cytoskeleton genes [94].

Further, it has also been shown that PHF8 can regulate neuronal activity-dependent gene transcription. The authors of this paper used treatments to increase synaptic activity and cause long term potentiation (LTP) and observed a raise of nuclear levels of both PHF8 and TIP60 proteins, specifically in the neurons where was successfully induced LTP. Consistent with the increase in PHF8, they found that LTP transiently downregulated the PHF8 substrate H3K9me2 [95].

More recently, it has been generated a knockout allele for Phf8 in mice to examine the consequences of Phf8 loss in development and mouse behaviour. Surprisingly, Phf8 deficient mice neither displayed global developmental defects nor signs of cognitive impairment. However, the animals reported a striking resiliency to stress-induced anxiety and depression-like behaviour. The authors of the paper observed misregulation of serotonin signalling in the prefrontal cortex of Phf8 deficient mice and identified the serotonin receptors Htr1a and Htr2a as direct targets of PHF8 [96].

Another group reported that Phf8 knockout mice display impaired learning and memory, and impaired hippocampal long-term potentiation (LTP), even without gross morphological defects. They also showed that mTOR signaling pathway is hyperactive in the hippocampus of Phf8 knockout mouse. They conclude that demethylation of H4K20me1 by Phf8 results in transcriptional suppression of RSK1 and homeostasis of mTOR signaling. Indeed, pharmacological suppression of mTOR signaling with rapamycin recovers the weakened LTP and the cognitive deficits (Figure 117)[97].

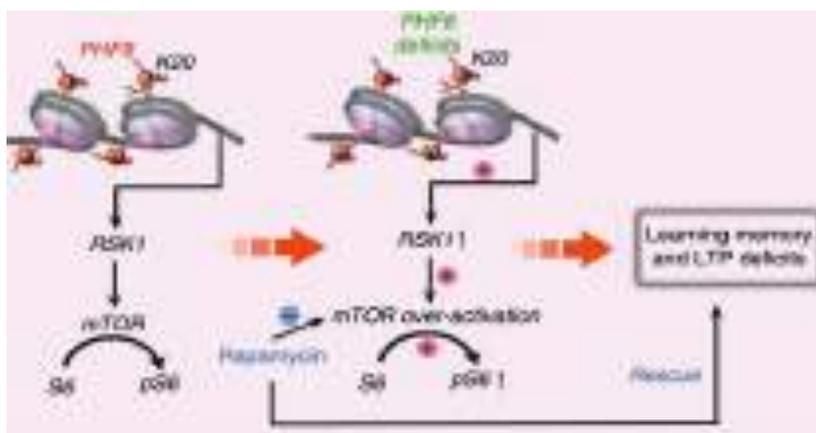


Figure I17: Working model describing the mechanisms underlying the Phf8 deletion induced cognitive impairment. Adapted from[97].

2.5. *PHF8 and X-linked intellectual disabilities*

X-linked intellectual disability (XLID) includes more than one hundred inherited syndromes caused by mutations of genes on X chromosome. XLID affects 1,6/1000 males, and 2,4/1000 females are carriers of mutations [98]. XLIDs are divided in non-syndromic forms (IDX), in which intellectual disability is the only clinical manifestation, and syndromic forms (IDXs), in which intellectual disability is associated with biochemical abnormalities, neurological features, and detectable physical signs (skeletal abnormalities and facial dysmorphism).

The first demonstration that PHF8 was directly involved in XLID came in 2005 when Laumonnier et al. showed that PHF8 mutations in two unrelated families were associated with XLID and cleft lip/palate (CL/P) [89]. They concluded that PHF8 is involved in midline formation and its catalytic domain is crucial as the two truncating mutations found in the two families occur nearby and in the JmjC domain of PHF8. Importantly, one of the truncating mutations was also found in a family with Siderius-Hamel CL/P syndrome [99]. Later, it was identified a novel de novo nonsense mutation (p.K177X), that results in premature truncation of PHF8 protein with loss of the JmjC domain and five NLS [100].

A missense mutation c.836C>T in the JmjC domain (found in a Finnish family with multiple-affected male patients [90]) changes phenylalanine to serine (F279S) and confirms the relevance of the catalytic domain activity.

The patient's phenotype is characterized by mild mental retardation, dysmorphic features, unilateral or bilateral cleft lip and cleft palate.

2.6. *PHF8 and cancer*

Numerous studies establish a relevant role for PHF8 in cancer; PHF8 functions as an oncogene in many types of cancer and its overexpression is associated with poor prognosis.

In breast cancer was reported the interplay between PHF8 and HER2 signaling. In HER2-positive breast cancers PHF8 levels are elevated by HER2 and, in turn, PHF8 regulates the expression of HER2 [101]. Interestingly, genome-wide gene expression analysis revealed that PHF8 overexpression induces an epithelial-to-mesenchymal transition (EMT)-like process, inducing the upregulation of SNAI1 and ZEB1. PHF8 pushes the transcriptional activation of SNAI1, by TGF- β signaling, demethylating H3K9me2. It has also been demonstrated that PHF8 is upregulated and positively correlates with MYC in breast cancer. MYC regulates the expression of PHF8 post transcriptionally through miR-22 that directly targets and inhibits PHF8 expression. The authors conclude that PHF8 contributes to MYC-induced cell proliferation and to the expression of EMT-related genes (Figure I18) [102].

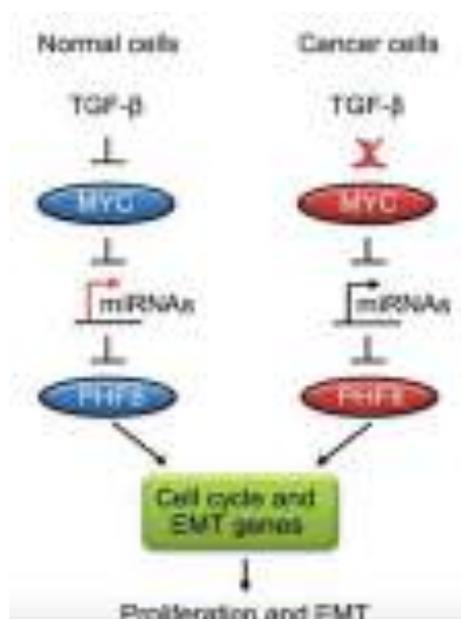


Figure I18: Schematic illustration of the regulations and functions of PHF8 in the context of TGF- β and MYC signaling. Adapted from [102].

PHF8 has been demonstrated to be upregulated also in human prostate cancer (PCa); its depletion induces PCa cell apoptosis by activating proapoptotic proteins and inactivating antiapoptotic ones [103]. In various prostate cancer cell lines PHF8 interacts with and functions as an activity-dependent androgen receptor (AR) coactivator and is induced by hypoxia. Knockdown of hypoxia-inducible factor HIF2 α or HIF1 α abolishes PHF8 expression, indicating that the HIF/PHF8/AR axis could serve as a potential biomarker for PCa and a therapeutic target too [104].

In gastric cancer (GC), PHF8 overexpression results in a poor prognosis for patients; it interacts with β -catenin and binds to the promoter of Vimentin, leading to the activation of its transcription. Vimentin is involved in cancer initiation and progression, including EMT and metastatic spread of cancer. In addition, *Helicobacter pylori*, the most important risk factor for GC, induces PHF8 expression [105].

Tumor tissues from patients with colorectal cancer (CRC) show increased PHF8 expression and its levels correlates with tumor-node-metastasis stage. PHF8 downregulation in CRC cells inhibits proliferation and migration, and promotes apoptosis [106].

PHF8 upregulation is common in hepatocellular carcinoma (HCC) tissues and correlates with worse survival. PHF8 downregulation suppresses cell growth, migration, invasion and autophagy. Its depletion abolishes the expression of SNAI1, Vimentin, N-cadherin, and leads to increased E-cadherin level [107, 108].

3. Epigenetic mechanisms in neurodevelopment

3.1. *Mouse cortex as a model of study*

Mice are one of the most common models to understand brain development; during fetal and postnatal development, brain generates neuronal and glial cells with diverse cellular phenotypes. Intrinsic and extrinsic signals cooperate to determine if neural progenitors continue to proliferate or start to differentiate; loss and gain-of-function experiments have been an instrument to identify these cues, leading to the acknowledgment of the neocortical development. The developing mouse

cortex contains 90% of neural cells (progenitors, excitatory neurons, interneurons, astrocytes, and oligodendrocytes) and 10% of non-neural cells (microglia and endothelial cells). The adult mouse cortex comprises the same classes of cells, but in different proportions depending on the region [109]. Cortical neurogenesis in mice embryos begins at day 11 (E11), when cortical progenitors populate the proliferative zones of the dorsal telencephalon, the ventricular and the subventricular zones (VZ and SVZ). Those progenitors are initially organized in a pseudostratified neuroepithelium and go through successive steps of maturation that progressively restrict their fate to generate all cortical cell types.

In human, the cerebral cortex is characterized by a six-layer organization in which new born neurons accumulate according to an inside-out sequence. Every area provides distinct functions from motor and sensory to cognitive processing. Neuronal composition and identity in each layer are specific and differ between cortical areas; they depend on the start of neurogenesis, cell cycle output during embryonic development and cell death in early postnatal stages [110].

3.2. *Neural stem cells as an in vitro model of study*

Neural stem cells (NSCs) are self-renewing multipotent cells existing in the developing cortex and in specific regions of adult mammalian central nervous system (CNS) [111]. They can generate neurons, astrocytes and oligodendrocytes depending on specific intrinsic and extrinsic signals (Figure 119). *In vivo* NSCs exist in niches that support self-renewal and regulate the balance between symmetrical self-renewal and asymmetrical division. In the past, isolation, purification and expansion of NSCs from their niches has been a challenge as the factors required to their maintenance were not understood. The epidermal growth factor (EGF) and the basic fibroblast growth factor (FGF2) were identified as key players; they sustain NSCs prolonged division and maintain their properties. We currently know that NSCs can be grown as neurospheres or monolayer system.

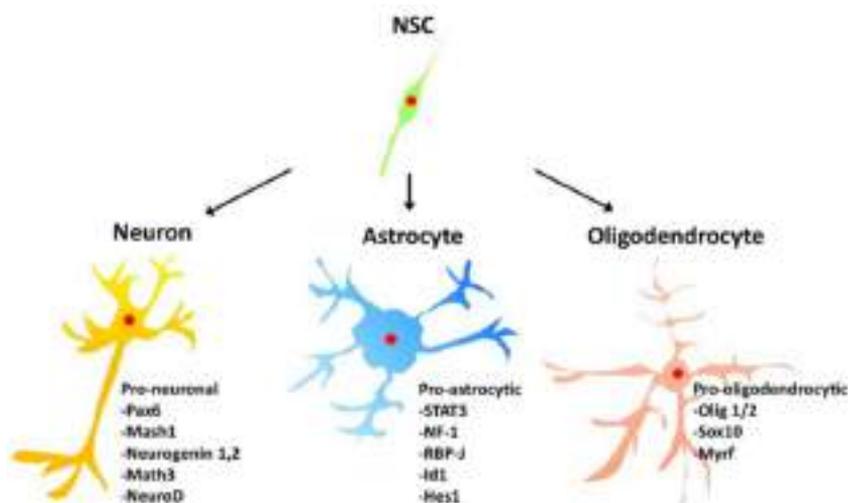


Figure 119: Neural stem cells differentiation into the major neural cell types (neurons, astrocytes, and oligodendrocytes) and are characterized by transcription factors specific of each lineage. Adapted from [112].

In this thesis, we took advantage of the monolayer system to maintain NSCs in culture and, when necessary, differentiate them. Neural stem cells can be efficiently expanded as adherent, clonal and uniform cell lines by exposure to EGF and FGF2; the cells divide symmetrically and retain their tripotential differentiation capacity [113]. *In vitro* NSCs are similar to forebrain neurogenic radial glial cells (RG) and when stimulated they can differentiate to neurons, astrocytes and oligodendrocytes [113, 114]. The cultures on monolayer avoid lineage restriction and spontaneous differentiation, but the use of the growth factors might alter the transcription of those cells [115]. The main differences between the gene expression phenotype of NSCs *in vitro* and *in vivo* concern regional identity and neuronal differentiation, which is clearly limited in *in vitro* cultures. The transcriptional profile of RGs *in vivo* is heterogeneous and this is also due to different positional signals. Instead, NSCs lose their original competence to generate specific neuronal subtypes of cells *in vitro*. In conclusion NSCs *in vitro*, even with their limitations, are a good and applicable model to better understand progenitors state and function [116].

3.3. Neurogenesis

Mouse neurogenesis starts when neuroepithelial progenitors appear and expand the number of progenitors through proliferation. The neuroepithelial progenitors then turn into radial glial cells (RG), which will differentiate into neurons and glial cells generating the cortex [117]. Neuronal differentiation starts when some radial glial cells begin to divide asymmetrically to generate another identical cell and a more differentiated one. In the forebrain, RG will give rise to other progenitor cells with a more restrictive fate: intermediate cells called basal progenitors (BPs). In the subsequent stages of neurogenesis, more progenitors go through symmetric terminal divisions in which both daughter cells differentiate. In this way, the expansion of the progenitors progressively slows and stops [118]. Many studies have investigated how signaling, transcription and epigenetics integrates to permit the switch from progenitors to differentiated cells. In both vertebrates and invertebrates, proneural genes control neurogenesis through Notch signaling which starts the expression of downstream differentiation genes [119]. The proneural genes encode basic helix-loop-helix (bHLH) transcriptional activators, which heterodimerize and bind to E boxes, in the promoter of target genes, activating the transcription of those genes. In the mouse cortex the key proneural genes are Neurogenin1 (Ngn1), Neurogenin2 (Ngn2), and Achaete-scute homolog1 (Ascl1). Ngn2 can regulate both Ngn1 (positively) and Ascl1 (negatively) transcription in cortical progenitors [120, 121]. Ngn2 and Ascl1 have been shown to directly control Delta-like protein1 (Dll1) (Figure 120) [122]. The transition from neuroepithelial cells to RG coincides with the Notch signaling, detected by the expression of the ligand Dll1, and the downstream transcription factors Hes1 and Hes5 in the dorsal telencephalon [123]. Notch is required to determine RG identity but deletion of Hes1 and Hes5 or RBPJ, the Notch effector transcription factor, did not block the appearance of RG [124], indicating that other pathways like Neuregulin 1 (Nrg1) and Wnt are involved in the process. Nrg1 is expressed in the developing cortex and signals through the receptors ErbB2 and ErbB4 to promote RG identity meanwhile suppressing RG differentiation to astrocytes. Gain and loss-of-function studies have demonstrated that Wnt stimulates proliferation and self-renewal of RG progenitors in early stages of cortex development. Later, Wnt signaling also promotes the maturation of RG cells into BPs and their proliferation regulating N-myc and the proneural gene Ngn1 [125].

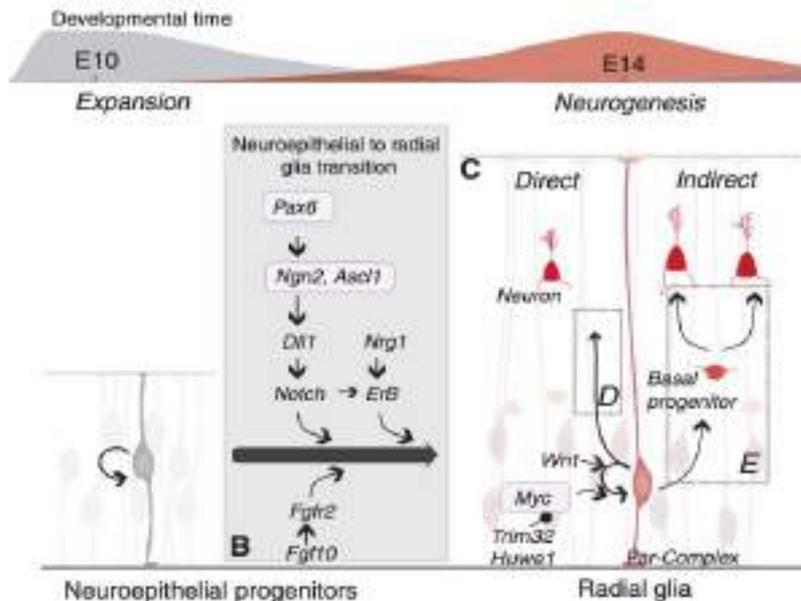


Figure I20: Molecular pathways regulating the onset, progression, and termination of neurogenesis in the mouse cerebral cortex. Adapted from [126].

3.4. Epigenetic mechanisms in neurogenesis

Epigenetic modifications are crucial to guide NSCs differentiation and lineage commitment as they coordinate the expression of transcriptional regulators in space and time. Such modifications are produced mainly by DNMTs and KMTs; for example, by DNMT1 and DNMT3 that, maintaining DNA methylation, regulate the division of neural progenitor cells. Mutations in any of the Dnmt genes in mice leads to drastic developmental abnormalities and embryonic or early postnatal lethality [127]. Deletion of Dnmt1 in NSCs causes hypomethylation and de-repression of genes necessary for neuronal differentiation. A study in mice from 2014 resolved the neuronal DNA methylome and showed that 75% of DNA methylation occurs at CpG sites, with the rest occurring at CpH sites (non-CpG sites). CpH methylation occurs *de novo* during neuronal maturation in both mice and humans [128, 129]. The knockdown of DNMT3A in neurons causes loss of methylation at many CpH sites, but not at CpG sites, suggesting that neuronal CpH methylation is more dynamic and actively maintained by DNMT3A.

As I mentioned before, histone modifications are largely involved in neurogenesis: in ESCs the housekeeping genes hold marks for transcription initiation like H3K4me3, while developmental genes are marked by active (H3K4me3) and repressive (H3K27me3) modifications [130]. This specific state of the chromatin, which holds marks of activation and repression, is described as the “bivalent” state or “poised state” [131]. Regulation of developmental and pluripotency genes is closely linked to Polycomb group proteins (PcG) activity. PcG are in charge of H3K27me3 deposition at genes of cell lineage commitment and they control the neurogenic to astrogenic transition modulating the expression of *Ngn1* [132]. During cell differentiation the repressive marks (H3K9me3 and H3K27me3), which in ESCs cover around 4% of the genes, expand to 12%–16% [133]. Deletion of Enhancer of Zeste homolog 2 (*Ezh2*), the enzyme responsible of H3K27 methylation, in NSCs results in a global loss of H3K27me3, de-repression of a large set of neuronal genes and impaired neuronal differentiation [134]. While many genes are silenced during differentiation, others need to be expressed, in fact Polycomb activity is antagonized by other signaling like the retinoic acid pathway, which induces *Jmjd3*, an H3K27me3 demethylase. Enhanced expression of *Jmjd3* promotes demethylation of several neuronal genes such as *Ngn1*, *Doublecortin (Dcx)*, *NK2 homeobox 2 (Nkx2.2)* and *Dlx5* [135]. *Jmjd3* is also essential to fully activate TGF β -responsive enhancers during neural development [136]. Thus, during neurogenesis, several transcription factors display transient changes in histone methylation at their promoters. For example, the *Tbr2* gene, which encodes a transcription factor involved in the generation of BPs [137], changes from a repressive configuration marked by H3K27me3 in proliferative RG to an active configuration marked by H3K4me3 in RG undergoing neurogenic divisions (Figure I21) [138]. Trithorax group proteins (TrxG) also instruct neurogenesis; the histone methyltransferases Mll1 antagonizes H3K27me3 deposition on the promoter of neurogenic transcription factors like *Dlx2* [139]. Moreover, many HDACs modulate histone deacetylation during the neurogenic process, for example, HDAC2 is upregulated during the differentiation of NSCs into neurons [140].



Figure I21: During neurogenesis, cell specific expression of genes is regulated by transcription factors that are submitted to the control of histone methylation by TrxG (H3K4me3) and PcG (H3K27me3) proteins. Adapted from [141].

3.5. Astrogliogenesis

Around stage E18.5 of mouse cortical development, RG stop producing neurons and the first astrocyte emerges. During this process, most RG release their apical attachment and lose the radial processes acquiring gradually the astrocytic morphology. In that moment the output of Notch signaling undergoes a dramatic shift. In early cortical development Notch signaling promotes the RG progenitor state (as described above), at later stages Notch activation promotes astrocyte differentiation and blocks neurogenesis [142, 143]. The key transducers of Notch signaling, which are the intracellular domain of Notch receptor and the RBPJ transcription factor, form a transcriptional activation complex on the glial fibrillary acidic protein (GFAP) promoter. It has been demonstrated that cultured neural progenitors, derived from RBPJ mutant ESCs, show a significant delay in astrocyte development *in vitro*. The main sources of Notch signaling during corticogenesis are young neurons and BPs, which express the Notch ligand Dll1. Notch signaling in cortical progenitors activates the Nuclear Factor IA (NFIA), which is the driving transcription factor for astrocytes specification (Figure I22) [144].

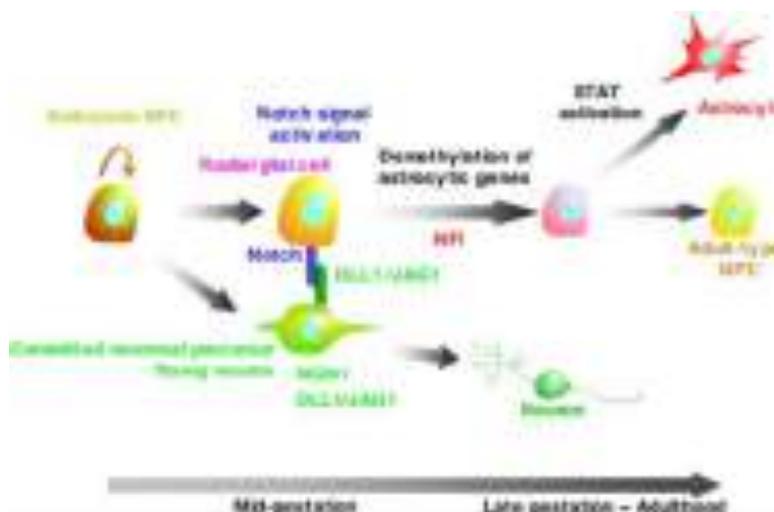


Figure I22: Representation of Notch activation of NSCs to differentiate into astrocytes. Adapted from [145].

NFIA not only is crucial in astrocytes differentiation but it is also expressed throughout maturation to maintain astrocyte function [146]. Sox9 also regulates early gliogenesis inducing NFIA during astrocytes differentiation [147]. Others pro-astrocytic signals are the Jak/Stat pathway [148] and BMP signaling. Early Jak/Stat and BMP expression promotes neurogenesis, whereas late expression drives astrocytes development. When BMP4 is overexpressed in cortical neurons, radial glial cells tend to differentiate into astrocytes [149]. Moreover, it has been demonstrated that BMP signaling contributes to glial cell maturation *in vivo*; when the signaling is inactivated deleting *Bmpr1a* and *Bmpr1b* in mice neural tube, P0 mutant mice exhibit a 25-40% decrease in GFAP in the spinal cord [150]. NFIA is necessary also in adults, indeed its conditional deletion in mice astrocytes causes morphological and physiological alterations. The phenotype is especially evident in the hippocampus where the loss of astrocytic NFIA results in impaired interaction with neurons, weakening of learning and memory skills. In conclusion NFIA loss in mice astrocytes impairs synaptic plasticity [146].

3.6. *Epigenetic mechanisms in astrogliogenesis*

NSCs enter the gliogenic phase from late-gestation to perinatal periods differentiating into astrocytes and oligodendrocytes. Murine NSCs from different developmental stages (E11.5, 14.5 and 18.5) have been used to determine the DNA methylome; the results demonstrate the existence of consecutive waves of global DNA methylation/demethylation that regulate the sequential generation of neurons, astrocytes and oligodendrocytes [151]. Indeed, astrocyte-specific promoters undergo dramatic changes in DNA methylation during the switch from neurogenic to gliogenic phase. These changes explain why NSCs are responsive to astrogenic signaling or not in different stages of neural development. In early stages of development NSCs are not able to undergo astrocytic differentiation when exposed to the leukemia inhibitory factor (LIF), even though they express functional LIF receptors [152]. Later, at stage E14, cultures of murine NSCs quickly differentiate into astrocytes upon LIF stimulation. Another example is the Jak/Stat signaling pathway; many glial promoters, responsive to Stat transcription factors, are specifically methylated during the neurogenic period, preventing Stat binding [153]. At the end of

neurogenesis those promoters are demethylated, also through Notch signaling, so the genes can be transcribed.

Notch signaling is the main signaling pathway involved; NFIA both induces demethylation of GFAP promoter and promotes dissociation of DNMT1 from GFAP promoter. In this way the STAT binding site is de-repressed [145]. Further, STAT3 is able to recruit the transcriptional co-activators CBP/p300, which acetylates H3K9 and H3K14 on GFAP promoter. These histone modifications promote H3K4 trimethylation and the recruitment of RNA polymerase II, thus activating gene transcription [154]. Methylation has an important role also in the expression of the immature astrocytic marker S100B. At early stages methyl CpG binding protein 2 (MECP2), which recruits HDACs and other corepressors, binds the promoter of S100b preventing its gene expression. At E14 a specific cytosine residue within the promoter region is demethylated and MECP2 can no longer bind to this site allowing S100b transcription (Figure I23)[155]. During astrocyte differentiation, LIF acts synergistically with RA and through histone H3 acetylation it activates GFAP promoter [156]. RA receptors (RAR) form complexes with retinoid X receptors (RXR) and bind to RA response elements in the promoter regions of target genes. When the RA ligand is absent, RAR/RXR associates with transcriptional repressors, leading to gene silencing by recruitment of HDACs. Conversely, binding of RA enables the release of HDACs from the RAR/RXR complex and the recruitment of HAT co-activators [157].



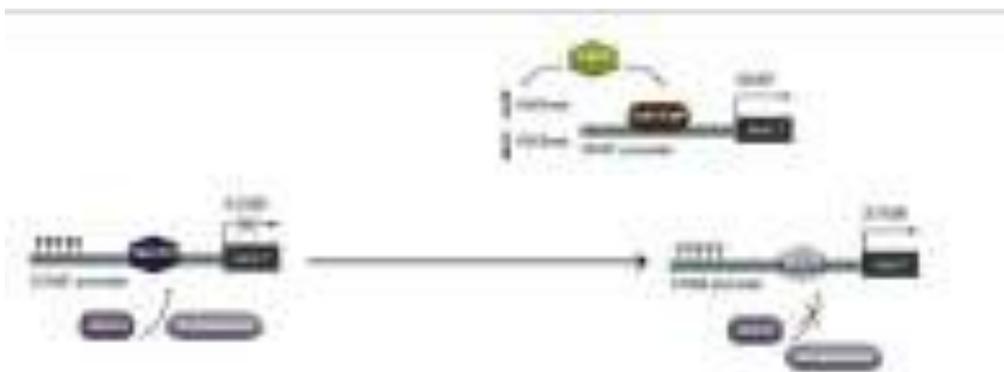


Figure I23: Examples of major transcriptional and epigenetic events driving gene expression of two main astrocytic markers, GFAP and S100b. Adopted from [158].

4. Astrocytes

4.1. Astrocytes and synaptogenesis

Neuronal synapses only form after astrocytes generation, at the same time of neuronal branching and maturation [159]. A functional synapse is formed by the presynaptic terminal of a neuron (containing neurotransmitter vesicles), a post synaptic density of another neuron (where are located the receptors for neurotransmitters) and many astrocyte's processes [160]. Synapse formation begins during the first postnatal (P) week, peaks at P14, and stabilizes at P21 to P28, concurrent with synapse elimination and the refinement of circuits (Figure I24) [161]. Astrocytes are the most abundant glial cell population in the adult brain and they control many processes; blood-brain barrier maintenance, ion homeostasis, nutrients and metabolites regulation. Importantly they can regulate synaptic transmission through the uptake of glutamate [162], they interact with neurons actively stimulating or eliminating neuronal synapses [163] and they release "gliotransmitters" through calcium signals [164]. The active communication between astrocytes and neurons is fundamental in memory, sleep, circadian rhythms and neuronal survival [165]. Conditional deletion of astrocytes in adult mice results in neuronal loss and severe motor deficits [166]. Many papers have shown that neurons cultured in isolation make only few synapses, whereas when

neurons are cultured with astrocytes or astrocyte- secreted factors the synapse formation highly increases.

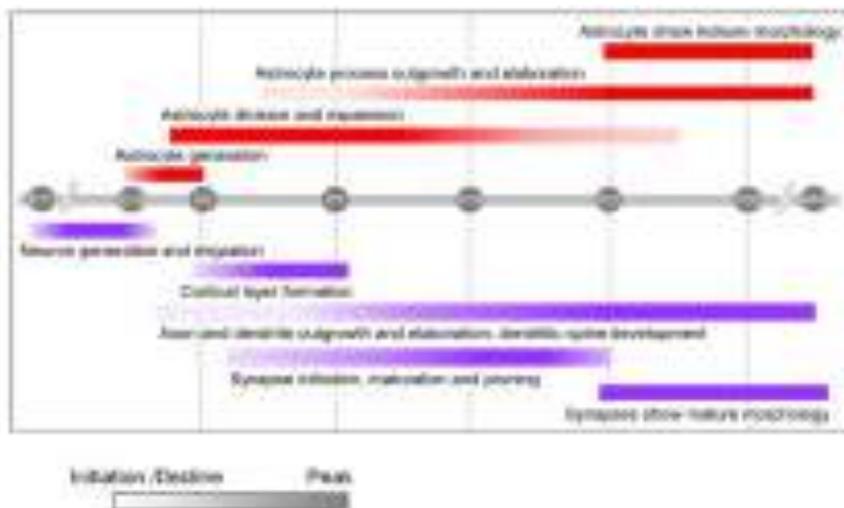


Figure 124: Timeline of key developmental processes in the rodent cortex from embryonic stages to the end of the first month of life. Adapted from [159].

At the first postnatal week cortical astrocytes are still dividing [167], have an immature morphology and express the synapse promoting factors thrombospondins (TSPs) [168] and glypicans (GPCs) [169]. TSP1/2 bind to neuronal receptors, such as $\alpha 2\delta-1$ (the receptor for the anti-epileptic and analgesic drug gabapentin) to induce presynaptic differentiation and clustering of synaptic proteins [170]. GPC4 and 6 are astrocyte-secreted signals and act increasing the surface level and clustering of the GluA1 subunit of AMPA glutamate receptor (AMPA) [169].

The neuronal and astrocytic proteins associated to synapse include: proteins that regulate the presynaptic vesicle transport and release, proteins in the postsynaptic density, neurotransmitter receptors of both excitatory and inhibitory synapses, and proteins that are secreted by astrocytes at synapses. Different astrocyte-secreted proteins can regulate various stages of synaptogenesis such as initiation (first postnatal week) or maturation (second- third postnatal week). For example, Hevin, an astrocyte- secreted protein that promote synapse formation, is low at P1 and peaks at P10–15, staying high in adulthood, so it plays a role in the synapse maturation. Conversely, the astrocyte secreted specific inhibitor of Hevin shows low expression at P10, peaks at P15, and is downregulated in the adult [171]. Hevin (Sparcl1) functions

by bridging two neuronal cell adhesion molecules, Neurexin 1a (Nrxn1a) and Neuroligin 1B (Nlgn1B), across the synapse and promotes the formation of both pre and postsynaptic specializations (Figure I25) [172]. The astrocyte-secreted glypican 4 also participates to the organization of active synaptic connections by coordinating both pre and post synaptic neurons. It induces the formation of active excitatory synapses recruiting AMPA glutamate receptors to the postsynaptic cell surface [173]. Synaptic plasticity is controlled by several glial mechanisms too; astrocytes regulate NMDA receptors through vesicular release of D-serine that acts as co-agonist for NMDA receptors [174]. In addition to their role in regulating synaptic strength and plasticity, glial cells actively refine circuits through pruning and phagocytosis of unnecessary and weak synapses. Indeed astrocytes express phagocytosis machinery and eliminate synapses through MEGF10 and MERTK pathways in an activity-dependent manner (Figure I23) [175]. The above mentioned are only some of the many factors which permit the intricate communication between neurons and astrocytes and have key roles in synapse formation.

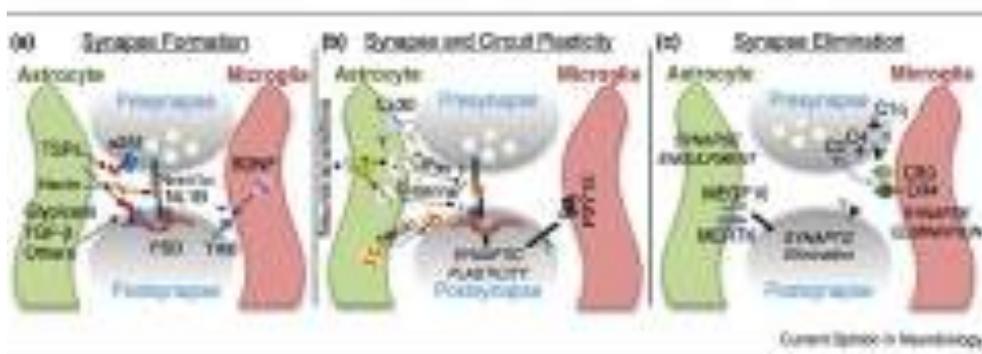


Figure I25: Synaptogenesis is controlled by several astrocyte factors which roles in synapse maturation, synapse and circuit plasticity and synapse elimination. Adapted from [176].

4.2. Astrocytes and intellectual disabilities

The clinical spectrum of intellectual disabilities (ID) varies extensively and is estimated to affect from 1% to 3% of the population. The causes of ID are heterogeneous and include genetic and/or environmental factors that

influence the development and function of the CNS during the pre, peri, or postnatal period. The anomalies result both in the formation and functions of synaptic circuits.

RNA sequencing technologies of brain cells [177] revealed that most ID genes are expressed not only in neurons but also in astrocytes (about 70% of ID genes)[178]. For example, Fragile X mental retardation protein (FMRP) and Methyl-CpG Binding Protein 2 (MeCP2), whose genes are mutated in Fragile x syndrome (FXS) and Rett syndrome (RS) respectively, are both expressed in astrocytes. In astrocytes, the FMRP doesn't have the same role of neuronal FMRP; it regulates mGluR5 and glutamate transporter (GLT1) expression, so it determines the glutamate uptake in astrocytes [179]. MeCP2 protein, which binds to methylated DNA, regulates the expression of different sets of genes in astrocytes through chromatin structural changes [180]. MECP2 is a well-known transcription factor that controls gene expression through regulation of epigenetic markers [181]. It is expressed in many tissues and, although the disease is generally attributed to neuronal dysfunctions, glial MECP2 plays a pathophysiological role. It has been shown that MECP2-null astrocytes are unable to support the normal dendritic ramification of wild-type neurons growing in culture [182]. Deletion of MECP2 in glia negatively influences neurons, and its re-expression in astrocytes significantly improves mice locomotion, improve respiratory anomalies and prolongs mice lifespan [183].

In most ID, astrocytes display an abnormal state; many astrocytic markers, such as GFAP and S100 β are altered and astrocytes exhibit a less complex morphology with thicker branches, suggesting astrocytic reactivity [184]. Rett patients present increased GFAP expression compared with controls, possibly because MeCP2 directly represses GFAP expression [185]. Increase in GFAP, but also S100 β expression, was also found in astrocytic cultures of MeCP2-deficient mice [186]. The modulation of GFAP expression is not still clear in the case of FXS, as both increased or no change in GFAP expression have been reported in patients [187, 188]. It has been demonstrated that loss in expression of the Fragile X mental retardation 1 protein (FMRP) from astrocytes is associated with delayed dendrite maturation and improper synapse formation. Indeed, cultured astrocytes isolated from an *Fmr1* knockout (*Fmr1* KO) mouse model of FXS displayed a significant decrease in thrombospondin-1 (TSP-1) protein expression compared to the wild type (WT) astrocytes. TSP-1, as mentioned above, is an important astrocyte-

secreted protein that is involved in the regulation of spine development and synaptogenesis [189].

The importance of glial cells in the pathophysiology of autism spectrum disorders (ASDs) is due to the association between ASDs and many genes related to glial cell activation belonging to immune and inflammatory categories [190]. In physiological conditions, inflammatory processes in the developing brain are controlled by homeostatic mechanisms that limit the inflammation induced by the environment [191]. In the brain these surveillance mechanisms are principally controlled by microglia and astrocytes, which are crucial in controlling that inflammatory processes efficiently remove pathogens. Thus, microglia and astrocytes dysfunctions can cause chronic neuroinflammation. Some reports have shown an increased levels of pro-inflammatory cytokines in the post-mortem brains of subjects with ASDs [192], leading to the hypothesis that chronic neuroinflammation can be crucial in ASDs [193].

To summarize, astrocytes are involved in the formation and maturation of neuronal networks by acting at multiple levels: dendritic growth, synaptogenesis, synapse maintenance, and pruning [38, 39]. These regulation of synaptic networks (mediated by astrocytes) is altered in ID. The alterations can be manifested by changes in astrocytes complexity and reactivity, synthesis and secretion of factors that support neuronal growth, expression of receptors and transporters which permit glutamate sensing and clearance, and activity of astrocytic channels that regulates K^+ buffering [194] (Figure I26). Astrocytes indeed have high permeability to potassium ions and are able to spatially redistribute the local excess of synaptic potassium through gap junctions in neighboring regions. For this, transient increases in extracellular potassium levels caused by neuronal activity are buffered by astrocytes, which prevents network hyperactivity. The balance between excitatory and inhibitory transmission (E/I) is commonly altered in FXS and RS [195]. This effect may due to changes in the number of synapses or ionic homeostasis that finally lead to alteration of neurotransmitters and altered E/I balance.

Objectives

Taking into consideration the current knowledge about the subjects treated, several goals were proposed for this PhD thesis in order to further understand PHF8 function in both astrocytes and neural stem cells biology. To do that, we determined the following objectives:

1. To investigate the role of PHF8 histone demethylase during astrocytes differentiation.
 - Analyse the PHF8-mediated transcriptional profile in astrocytes.
 - Determine the chromatin bound regions of PHF8 in astrocytes.
 - Elucidate astrocytes phenotype upon PHF8 depletion.
 - Examine astrocytic PHF8 function during synaptogenesis.
 - Determine the molecular mechanism responsible for the PHF8-associated phenotype in astrocytes.

2. To elucidate the function of PHF8 in neural stem cells.
 - Examine neural stem cells phenotype upon PHF8 depletion.
 - Determine the PHF8-mediated transcriptional profile in neural stem cells.
 - Analyse the metabolic impact of PHF8 depletion.

Materials and methods

1. Reagents

1.1. Plasmids

Plasmid	Origin/Provider
pCMV-VSVG	Dr. Timothy Thomson
pCMV-GAL-POL	Dr. Timothy Thomson
pLKO-Control (CAACAAGATGAAGAGCACC)	Sigma
pLKO-PHF8-1 (GCAGGTAAATGGGAGAGGTTT)	Sigma
pLKO-PHF8-2 (GCAGGTAAATGGGAGAGGTT)	Sigma
pGL2- Hes5promoter	Addgene
pRL-TK Renilla	Promega
NICD-pCIG	Dr Elisa Martí
Pinducer hPHF8	Cloned by Claudia Navarro
Pinducer hPHF8 H247A	Cloned by Claudia Navarro

1.2. Antibodies

Antibody target	Provider and reference	Dilution used
PHF8	Abcam, ab36068	Western Blot 1:1000, ChIP 1:500, Immunocytochemistry 1:500
H4K20me1	Abcam, ab9051	ChIP 2ug/ml Immunocytochemistry 1:500
H4K20me3	Abcam, ab9053	ChIP 2ug/ml Immunocytochemistry 1:500
H3K9me2	Abcam, ab1220	ChIP 2ug/ml Immunocytochemistry 1:500
H3K9me3	Abcam, ab8898	Immunocytochemistry 1:500
H3K4me3	Abcam, ab8580	Immunocytochemistry 1:500
H3K27me3	Millipore, 07449	Immunocytochemistry 1:500
NESTIN	Abcam, ab5968	Immunocytochemistry 1:500
GFAP	Dako, z0334	Immunocytochemistry 1:500
OLIG2	Merck, AB9610	Immunocytochemistry 1:500
GLAST	Millipore, 06-570	Immunocytochemistry 1:500

β -TUBULIN III (TUJ1)	Covance, MMS-435P	Immunocytochemistry 1:500
BASSOON	Synaptic Systems, 141004	Immunocytochemistry 1:500
SHANK2	Synaptic Systems, 162202	Immunocytochemistry 1:500
SHANK3	Synaptic Systems, 162304	Immunocytochemistry 1:500
AQUAPORIN 4	Abcam, ab125049	Immunocytochemistry 1:500
EAAT2	Abcam, ab41621	Immunocytochemistry 1:500
NG2 Chondroitin Sulfate Proteoglycan	Merck, AB5320	Immunocytochemistry 1:500
IBA-1	ThermoFisher, PA5-27436	Immunocytochemistry 1:500
GPR17	Cayman chemical, 10136	Immunocytochemistry 1:500
α -TUBULIN	Millipore, MAB3408	Western Blot 1:10000
Unspecific IgGs		ChIP, same dilution than the specific IgG
Anti- rabbit IgG HRP	Amersham #GENA934	Western Blot 1:10000
Anti-mouse IgG HRP	Amersham #NA9310	Western Blot 1:10000
Anti- rabbit IgG IRDye	LI-COR #926-32221	Western Blot 1:5000
Anti-mouse IgG IRDye	LI-COR #926-32210	Western Blot 1:5000
Anti- Rabbit Fluor 555	Alexa #A31572	Immunocytochemistry 1:1000
Anti- Mouse Fluor 488	Alexa #A21202	Immunocytochemistry 1:1000

1.3. Primers

Primer sequences	Primer Forward	Primer Reverse
Gene expression		
Phf8	GCATACTGGAGAACCGAGAG	CGAGATTTCAAAGCAGGGTC
Nfia	CCTCCAACCACATCAACAGAAG	GTACCAGGACTGTGTCTGTTG
Gfap	AGAAAGGTTGAATCGCTGGAG	CTGTGAGGTCTGGCTTGG
S100b	TACTCGGACACTGAAGCCAG	CCCGGAGTACTGGTGGAAG
Olig2	GCTTAGATCATCCCTGGGGC	AGATCATCGGGTTCTGGGGA
Gapdh	ATGTTTCGTCATGGGTGTG	CCTTCCACGATACCAAAGTTG
Hes5	CTACCTGAAGCACAGCAAAG	AGCTTCATCTGCGTGTGC
Sparcl1	TGGATTACTTCGGAGCTTGC	GCTTTTCATTGAGATAGCCGC
Cdk5rap2	CACGTCCAGACAGTCTCTTTG	CGCCTTAATTTTACCTCTTCCG
Ncam1	CATGTGCATCGCTGTAAACC	TCATGGTTTGGAGTCCGTTT
Gpc6	ATTGCCCTACACCATCTGC	TCAGCCCATCGTTCATGATC
Cntnap2	CCCATGTCTTCAGCCACTG	CGATGACCCCTCCAATGATAG
Nlgn1	TTAGGTGATAATGACGGTGCTG	GATCACATTGCCATAGCTTGC
Nrxn1	AGATGTCCACCTCAATCATGG	AATGTCCTCATCGTCACTGG
Tsc1	AGTTCTTGAATAGGCAGCTCC	CTCTAGCTCTTCCGATATGCAG
Mmd2	TCCAACCTCTACTTCCTGTCC	GCAGTGCTCTACCATCCTG
Zcchc24	GGACACTACATTAAGGACTGCC	TCCCACTCATCCATTTTCTCTTG
Mbp	ACCCAAGATGAAAACCCAGTAG	CCTCCGTAGCCAAATCCTG
Plp1	CTCCAACCTTCTGTCCATCTG	TGAGTTTAAGGACGGCGAAG
Kdm5b	CCTCATATTTACTCTCCCTTCTC	GTAAGTAGAGTCTGATAAAGCTCCTG
ChIP – qPCR		
Nfia TSS	AGCCTGTCATGGGAAATC	ATCAATGGTGTGAGAAAGGT
Nfia -400	TGCAAAGTCTCTTTCAAGCACA	ATCCAATCTAACCCGAGC
Sparcl1 intra	GTGTTAGTGTTCCCTTCCGT	AGAGGAACTCATGAACAGTCAA
Cntnap2 intra	TGCACACACAACATATTCCAC	CACACTCATCCAGATCAATAACTA
Phf8 TSS	TGTTTACCATATCTCTCCACCC	GTTGTAGGAGATTCAAAGCAATCA
Arid1b intra	AGGGAGAGATTCTTAGTCCAT	TTTCATTCAAACGACCGCA
Olig2 TSS	TGCTGCCTCCACCCA	GCTCGGTCTGTAATAAGCAT
Nfia intra	ACCACTGTATGTCTGTGC	CTTCCACTTGGGTTTGTTC
Ncam1 intra	TTCCAGCAAACACTGCAC	AGAAGTCCAATAGTATGCCTGA
Cdk5rap2 intra	CAGTTCGGAGGTCAAAGG	TGATGACTTGAGTTTGTATCCC
Jarid1c TSS	TTCCGCCAATGAAATGAACTAT	TCCCTTATTTGGAGGTGGT

2. Cell culture and differentiation

2.1. NSCs

Mouse neural stem cells (NSCs) were dissected from cerebral cortices of C57BL/6J mouse fetal brains (E12.5) and cultured in poly-D-lysine (5µg/ml, 2 hours 37°C) and laminin (5 µg/ml 37°C, 4 hours 37°C) precoated dishes following the previous published procedures [196]. NSCs were grown in medium containing equal parts of DMEM F12 (without Phenol Red, Gibco) and Neurobasal medium (Gibco) with Penicillin/Streptomycin, Glutamax (1%), N2 and B27 supplements (Gibco), sodium pyruvate (1 mM), non-essential amino acids (0.1 mM), Heparin (2 mg/l), HEPES (5 mM), bovine serum albumin (25 mg/l) and β-mercaptoethanol (0.01 mM) as previously described [197]. Fresh recombinant human Epidermal Growth Factor (EGF) (R&D systems) and Fibroblast Growth Factor (FGF) (Invitrogen) to 20 ng/ml and 10 ng/ml final concentrations respectively were added to the media. Under these conditions, NSCs maintain the ability to self-renew and to originate a wide range of differentiated neural cell types [196, 198].

2.2. HEK 293T cells

In order to produce lentivirus, HEK293T cells were used, as they have a high degree of transfectability. These cells derive from human embryonic kidney transformed with the large T antigen of the SV40 virus [199]. HEK293T cells were cultured in DMEM (Gibco #41965-062) supplemented with 10% of fetal bovine serum (FBS) (Gibco #10270106) and 1% of Penicillin/Streptomycin [200].

2.3. HeLa cells

HeLa cells came from the cervix of a young lady called Henrietta Lacks and were the first human cells to be continuously grown in culture. They were grown in DMEM/F-12 supplemented with 10% of fetal bovine serum (FBS) (Gibco #10270106) and 1% of Penicillin/Streptomycin. Cells were kept in incubators at 37 °C, with a partial pressure of CO₂ of 5% [201].

2.4. Astrocytes differentiation

For NSCs differentiation into astrocytes, the medium was replaced by the astrocytic medium (used for differentiating NSCs to astrocytes) containing

DMEM/F-12, 5% N2, 1% Glutamax, 10% FBS and FGF 10 ng/ml; fresh astrocytic medium was supplied every 2 days.

2.5. *Cocultures: neurons- astrocytes*

Primary neuronal cultures were obtained from the hippocampus of 18-day-old fetal C57BL/6 wild type mice (Charles River), of either sex [202]. Dissociated cells were plated onto previously differentiated astrocytes (10 days differentiation) in multiwell plates of 12 wells at seeding density $0,3 \times 10^6$ and maintained in Neurobasal medium supplemented with 2% B27 (Life Technologies, Carlsbad CA, USA), Penicillin/Streptomycin 1%, L-glutamine 0,5mM and glutamate 12 μ M. Cultures were maintained in standard conditions at 37°C and 5% CO₂. After 3 days *in vitro*, the medium was partially replaced by fresh medium.

3. Genetic manipulation of growing cells

3.1. *Calcium phosphate transfection*

This method allows the delivery of plasmids into the cells through the endocytosis of calcium phosphate precipitates that contain DNA molecules stuck on the surface. To generate the precipitates, the DNA, diluted in 0.25M CaCl₂, was included into a mixture containing 250mM NaCl, 9mM KCl, 1.5mM Na₂HPO₄, 10mM glucose and 50mM Hepes pH 7.12 under vortexing conditions. After 10 minutes at room temperature, the mixture was added to the growing medium and 6 hours later, new medium was supplied to avoid cell stress due to acidity.

3.2. *Lentiviral transduction*

This delivery system is highly efficient both in NSCs and Hela and permits to obtain up to a 90% of genetically modified cells. It was applied for the transduction of shRNAs that enabled the knocking down of PHF8 or the transduction of vectors for PHF8 overexpression. The procedure consisted in three steps: lentiviral production, lentiviral transduction and selection. Lentiviral particles were produced in HEK293T cells by cotransfecting in four 10 cm plates the DNA encoding the shRNAs

(pLKO.1-Control, pLKO.1-PHF8-1, pLKO.1-PHF8-2) or the cDNA of PHF8 WT and catalytic mutant (pINDUCER20-hPHF8, pINDUCER hPHF8 H247A [203]) together with pCMV-VSVG and pCMV-GAG-POL plasmids that encode the viral capsid and transcriptional machinery respectively. After 24 and 48 hours, supernatants containing lentiviral particles were collected and centrifuged in a sucrose bed at 57000xg during 2 hours. Then, supernatant was removed and viral particles were resuspended in NSCs or HeLa medium. These particles can be stored at -80°C or be immediately used for infection. Transduction of the previously produced lentivirus consisted in the addition of the medium containing viral particle to the receptor cells. Approximately, one production allowed to infect 1.5×10^6 cells. One day after infection, cells were selected with the correspondent antibiotic, that for pLKO.1 plasmid was puromycin at a concentration of 2µg/mL and for pINDUCER20 plasmids containing neomycin resistance was G-418 at a concentration of 1mg/mL. After 48 hours cells were considered “selected” and the knocking down of the particular protein was assessed by RT-qPCR or Western Blot.

4. Molecular biology procedures

4.1. Genomic DNA extraction

Approximately 6×10^6 cells were lysed in 200µL of buffer (10mM Tris-HCl, 10mM EDTA, 10mM NaCl and 0.5% SDS). Then, lysates were incubated with 0.5mg/mL of proteinase K (Sigma #P2308) during 1 hour at 50°C and with 1mg/mL of RNase A (Fermentas # EN0531) during 2 hours at 50°C. The resulting mixtures of DNA and proteins were subjected to phenol-chloroform extraction to purify the DNA.

4.2. Phenol chloroform extraction and ethanol precipitation

This procedure serves to purify DNA from complex protein-DNA mixtures. It is based in the different affinity of DNA and proteins for the organic solvent phenol. First, 1 volume of phenol was added to the mixture and the samples were mixed by vortexing. Then they were centrifuged at maximum speed during 3 minutes, so the aqueous and the organic phases were separated, and the aqueous phase containing the DNA was

moved to a clean tube. That step was repeated with chloroform and the aqueous phase obtained was subjected to ethanol precipitation. Ethanol precipitation followed phenol-chloroform extraction in order to concentrate DNA. An amount of 0.1 volume of sodium acetate 3M and 1 volume of cold ethanol were added to the sample, then the tubes were incubated at -80°C to favor precipitation. After 1 hour, samples were centrifuged at maximum speed during 20 minutes at 4°C . Finally, the pellet was dried and the DNA was resuspended in H_2O UltraPure.

4.3. RNA extraction (for RNA-seq and RT)

Trizol reagent (Invitrogen #15596018) has been used to extract RNA. Approximately 3×10^6 cells were lysed with 1mL of Trizol. Then, 200 μL of chloroform were added and after centrifugation at maximum speed during 5 minutes the upper phase of the tube was collected and precipitated by adding 800 μL of isopropanol. It was followed by another centrifugation at maximum speed during 10 minutes, supernatant was discarded and pellet was washed with 70% ethanol. After the last centrifugation of 5 minutes at maximum speed, pellet was resuspended in H_2O UltraPure.

4.4. DNase treatment

After RNA extraction, samples were treated with DNase to eliminate potential contamination of genomic DNA in the resuspended RNA. For this treatment, DNA-free Kit (Ambion # AM1906) was used. The kit allows DNA removal from RNA samples due to the inert beads that serve as inhibition agent. The protocol consisted in the addition of 0.1 volumes of 10X buffer to the RNA sample which was incubated with 1 μL of DNase during 30 minutes at 37°C . Next, 0.1 volumes of inhibition agent were added and after 2 minutes, the tube was centrifuged at 10000 rcf during 1.5 minutes and pure RNA was obtained by collecting the supernatant.

Then, RNA was quantified using a Biodrop device and quality was evaluated with 260/280 and 260/230 ratios. Besides, an agarose gel was run to check RNA integrity before proceeding with other techniques.

4.5. Reverse transcription (RT)

To quantify gene expression RT-qPCR experiments were performed. This technique allows the quantification of the RNA levels in the different tested

conditions. RT of mRNA was performed with 1µg of RNA, using the High Capacity cDNA RT kit (Invitrogen #4368814). The protocol consisted in the incubation of the RNA with random hexamers, dNTPs and a retrotranscriptase in a thermocycler following these parameters: 25°C 10 minutes, 37°C 120 minutes and 85°C 5 minutes. The complementary DNA (cDNA) generated stored at -20°C or -80°C.

4.6. qPCR

After RT of the RNA, I used quantitative polymerase chain reaction (qPCR) to quantify cDNA. qPCR experiments are based on the quantification of the emitted fluorescence by a fluorophore that binds DNA as PCR proceeds. A higher presence of cDNA results in more emitted fluorescence and vice versa. qPCR reactions were manually set-up in a volume of 10µL using SYBR Green Kit (Roche #4887352001). Reactions were carried out in 96-well plates in a LightCycler 480 (Roche) with the following cycling conditions (95°C 5 minutes, 40 cycles of 95°C 5 minutes - 60°C 10 seconds - 72°C 20 seconds, melting curve 95°C 5 seconds - 65°C 1 minute - 97°C). Specific primer pairs were designed spanning exon-exon junctions of a region conserved between splice variants to avoid the amplification of genomic DNA. To validate qPCR results, non-template controls were run and standard curves with every new primer pair were checked, so that only primers with an efficiency of 95% or higher were used. qPCR data were analysed using the $2^{-\Delta\Delta CT}$ method. Outliers were defined as values that differ more than 0.5 cycles from the other two wells in the triplicates. When identified, outliers were discarded. If the non-template controls were Ct=37 or lower reaction was repeated. In order to normalize qPCR results, Gapdh gene was used as a reference gene because after testing different housekeeping genes it has been the most constant and reliable one between conditions.

4.7. Mini and Maxi preparations of DNA

In molecular biology, solutions of DNA containing around 300 ng/µL of DNA are known as “minipreps”, likewise, solutions of DNA containing approximately 1µg/µL or more are known as “maxipreps”. To obtain these solutions of DNA from bacteria transformation, 5mL or 500mL of Luria Broth (LB) were inoculated with either an isolated colony or 5mL of miniculture. After overnight growth, the bacterial mass was subjected to

DNA purification with the alkaline lysis method following the protocol and using the buffers of the QIAprep Spin Miniprep Kit (for minipreps, QIAGEN # 27106) or the QIAGEN Plasmid Maxi Kit (for maxipreps, QIAGEN # 12165). The alkaline lysis method had three steps: resuspension, lysis and neutralization. The protocol consisted in the sequential addition of three buffers (P1, P2 and P3) corresponding to the three mentioned steps (Buffer P1: 100µg/mL RNase A, 50mM Tris-HCl, 10mM EDTA pH 8.0; buffer P2: 200mM NaOH, 1% SDS; buffer P3: KAc 3M, pH 5.5), then the lysate was passed through a column that specifically retains DNA, and after washes, DNA was eluted and precipitated with isopropanol. Finally, the DNA was washed with 70% ethanol and after drying it was resuspended in H₂O UltraPure.

4.8. DNA electrophoresis

This technique was used to purify and visualize DNA prior or after other applications. First, an agarose gel of the desired percentage was prepared by mixing agarose with TBE (Tris-Borate-EDTA) buffer (45 mM Tris, 45mM boric acid and 1 mM EDTA). After heating in the microwave until the agarose was dissolved, the solution was chilled and Redsafe reagent was added (Intron #21141). Redsafe is a reagent that permits the visualization of nucleic acids due to the green fluorescence that emits upon DNA and RNA binding. Then, this mixture was solidified in an electrophoretic chamber and after addition of TBE buffer, the DNA and RNA samples containing orange-glycerol were loaded on a gel. Gels were typically run at 80V and the visualization was acquired by using an UV-transilluminator.

4.9. Total protein extraction

To perform protein extractions, it was used a buffer called RIPA (Radioimmunoprecipitation assay buffer), which is highly astringent and breaks cytoplasmic as well as nuclear membranes (150mM NaCl, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% SDS and 50mM Tris, pH 8.0 and protease inhibitors). Approximately, 500µL of buffer were added to 6×10⁶ cells, but the volume could vary according to concentration requirements. This suspension was incubated on ice during 20 minutes and it was centrifuged at maximum speed during 10 minutes at 4°C. The supernatant contained the protein extract.

4.10. Protein quantification by Bradford

In order to measure the total amount of protein present in a sample, the Bradford method was performed. This method is based on the reaction occurring between the proteins in the sample and the Coomassie Brilliant Blue G-250, a reagent that change its color depending on the protein concentration [204]. To measure the concentration, 1 μ L of the protein extract was mixed with 1mL of Bradford solution (Bio-Rad # 5000001) and, after 3 minutes, the colorimetric reaction was measured in a spectrometer. Using the absorbance value of the sample and a bovine serum albumin (BSA) calibrate line, the concentrations of the samples were obtained.

4.11. SDS-Page electrophoresis

This classic procedure developed by Laemmli [205] serves to separate proteins in a gel according to their sizes. The sodium dodecyl sulfate (SDS) detergent provides with net negative charge to all the proteins in solution, thus ensuring that the migration occurs exclusively according to their sizes. To prepare the samples, the protein extracts were mixed with Laemmli buffer (375mM Tris-HCl, 9% SDS, 50% Glycerol, 0.03% Bromophenol blue) and 5% of β - mercaptoethanol and they were heat during 5 minutes at 95°C. Then, the samples were loaded in a polyacrylamide gel (Page) that was formed by two parts: the stacking and the resolving. The stacking gel had a pH of 6.8 and a polyacrylamide concentration of 5%. On the other hand, the resolving gel had a pH of 8.8 and a variable polyacrylamide concentration depending on the sizes of the proteins to resolve. After setting up the chamber with the gel, everything was covered with SDS-Page running buffer (25mM Tris-HCl, 192mM glycine, 0.1% SDS) and the power was set on at 25mA until all the sample had run through the gel.

4.12. Western blot

After SDS-Page electrophoresis, the Western Blot [206] was performed to specifically detect a protein in the sample that was resolved in the gel. This method consisted in setting up a multilayered cassette in which from positive side to negative side were placed a sponge, Whatmann paper, a nitrocellulose membrane, the gel proceeding from the SDS-Page, Whatman paper and another sponge. This cassette was introduced in a

chamber and the set was covered by Transfer buffer which was composed of 25mM Tris-HCl, 192mM glycine, 0.05% SDS and 10% methanol. After chamber, cassette and buffer were set, the power source was turned on at 80 V during 90 minutes. At the end, proteins were transferred to the nitrocellulose membrane. To specifically identify a protein in the nitrocellulose membrane, first it was blocked with milk 5% in PBS-Tween 0.1% during 1 hour at room temperature. Then, the membrane was washed three times with PBS-Tween 0.1% and was incubated with specific primary antibodies, usually overnight, at 4°C. Next, the membrane was washed again with PBS-Tween 0.1% (3 times) and incubated with a secondary antibody during 1 hour at room temperature. At that point the Western blot was prepared to be revealed. Both chemiluminescent and fluorometric methods were used to reveal immunoblots. Chemiluminescence was used to detect proteins that were less abundant; in that case, a secondary antibody bound to horseradish peroxidase (HRP) was adopted. On the other hand, fluorescence was more useful to reveal abundant proteins or proteins whose antibodies were really sensitive; for this method, secondary antibodies bound to a fluorophore were used. The luminol-based Immobilon Western kit (Millipore #WBKLS0500) was used to reveal membranes with chemiluminescence in an automated processor in a dark room. In the case of fluorometric method the revealing was performed in a LI-COR Odyssey scanner. Once visualized the detected proteins, they were analysed using ImageJ.

4.13. MACS isolation of astrocytes

Purified astrocytes were isolated from E18 embryos or P2 or P7 mouse whole brains by magnetic-activated cell sorting (MACS) (Miltenyi Biotec) with anti-GLAST (ACSA 2) MicroBeads according to the manufacturer's instructions. Cells were lysed with a buffer containing 1% SDS, 2 mM EDTA pH 7.4, 10 mM Tris-HCl pH 7.4 and proteases inhibitors. The absence of GFAP immunoreactivity in cells isolated at E18 suggests that GLAST-positive cells are still progenitors, in line with postnatal astroglialogenesis in the mouse brain.

4.14. Indirect immunofluorescence and analysis of synapses

For this protocol, NSCs, grown in a coverslip, were fixed for 20 minutes at room temperature in 4% of paraformaldehyde and permeabilized with PBS-Triton X-100 0.1% before blocking at room temperature for 1 hour in

1% BSA (Bovine serum albumin). Then, the coverslip was incubated overnight at 4°C with primary antibodies. After washing the samples three times with PBS-Triton X-100 0.1%, cells were incubated for 3 hours at room temperature with Alexa-conjugated secondary IgG antibodies and 0.1ng/μL 4',6-diamidino-2-phenylindole (DAPI) (ThermoFisher #D1306). These antibodies emit fluorescence that can be detected in the fluorescence microscope. Images were captured by a Leica SP5 confocal microscope using LAS-AF software and quantification was achieved by ImageJ counting cells in randomly located fields. In the case of the analysis of synapses in coculture experiments, the percentage of synapses with pre- and post-synaptic terminals was determined in a single confocal plane using ImageJ software. Regions Of Interest (ROIs) were drawn by using Freehandline tool on Tubβ3 images and the length of the segments was measured using the analyse function. Bassoon, Shank2 and double-positive puncta were counted by generating merge images Bassoon/Shank2/ Tubβ3. Synapse density was calculated dividing the number of puncta by the ROI length.

4.15. Growth curve

Cells were plated in multi wells of 6 and counted at 0, 24, 48 and 72 hours in duplicates with the Invitrogen Countess II that contains advanced autofocusing and counting algorithms, allowing to count quickly and accurately.

4.16. Cell cycle analysis by flow cytometry

This experiment was performed in collaboration with the group of Alejandro Vaquero in IDIBELL (Bellvitge Biomedical Research Institute). Click-iT® EdU Imagin kit (Invitrogen) was used. The first step was to prepare 2X 5-Ethynyl-2'-deoxyuridine (EdU) solution 20uM in culture medium from the 10mM stock solution. The preheated 2X EdU solution was added to the cells so that the final concentration is 1X EdU. The cells were incubated with EdU for 45 minutes (incubation time can vary depending on cell type). Subsequently, cells were collected and fixed with 70% ethanol and they were stored at -20°C for a minimum overnight and a maximum of 1 week. To stain the cells, samples were centrifuged 5 minutes at 2000 rpm, washed with a solution composed by 1ml PBS, 1% FBS and 2mM EDTA and centrifuged again 5 minutes at 2000 rpm. A second wash was performed and followed by centrifugation at 2000 rpm. The samples were permeabilized using 1ml PBS, 1% FBS, 2mM EDTA

and 0.5% triton and incubated 20 minutes at room temperature. After washing twice, Click-iT buffer was added as indicated in the kit protocol. An amount of 100 μ L of reaction per sample was added and incubated 30 minutes at room temperature, protecting the samples from light. Finally, the samples were analysed by flow cytometry.

4.17. Luciferase assay

Human HEK 293T cells were transfected by standard calcium phosphate coprecipitation with the luciferase reporter constructs and renilla for transfection efficiency normalization. Cells were harvested after 48 hours. Luciferase and renilla activities were measured using the Dual Luciferase Reporter Assay System (Promega).

4.18. ChIP and ChIP-seq

The chromatin immunoprecipitation procedure consists on the immunoprecipitation of proteins directly or indirectly bound to the chromatin and posterior identification of the DNA bound using qPCR (for a specific region), or sequencing (for all the regions bound in the genome). For every experiment, 6×10^6 NSCs were fixed with 1% of formaldehyde during 10 minutes, the fixation was stopped by adding 0.125M of glycine during 5 minutes. Then, cells were lysed in 1% SDS lysis buffer (1% SDS, 10mM EDTA pH8.0, 50mM Tris-HCl pH8.1) and sonicated to obtain fragments of around 300 bps of DNA, ensuring resolution in the DNA detection. The sonication step was performed in a Bioruptor sonicator with variable parameters, due to the inconsistency of the Bioruptor performance; example parameters are 30 cycles of 30 seconds on and 30 seconds off at high potency. Next, chromatin was purified by centrifugation at maximum speed during 10 minutes and the recovered supernatant was used for a sonication test to evaluate the correct size of the chromatin fragments in an agarose gel. Once chromatin was properly shredded, the immunoprecipitation step followed by diluting the chromatin tenfold with immunoprecipitation buffer (1% Triton X-100, 2mM EDTA, 150mM NaCl and 20mM Tris-HCl pH8.0) and adding the specific antibody for the protein of interest. In parallel another reaction was run using an unspecific IgG. After overnight incubation at 4°C the antibody-protein complexes were captured using Magna ChIP magnetic beads (Millipore #16-661) during 4 hours at 4°C. Next, antibody-protein-DNA complexes were sequentially washed with buffers TSEI (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0 and 150mM NaCl), TSEII (0.1% SDS,

1% Triton X-100, 2mM EDTA, 20mM Tris-HCl pH8.0 and 500mM NaCl), TSEIII (0.25M LiCl, 1% NP40, 1% sodium deoxycholate, 1mM EDTA and 10mM Tris-HCl pH8.0) and TE (10mM Tris-HCl pH 8.0 and 1mM EDTA) and eluted using elution buffer (1% SDS, 0,1M NaHCO₃) during 15 minutes. At that point, samples were subjected to decrosslinking overnight at 65°C to recover the DNA. Subsequently, the samples were treated with 1mg/mL of RNase A during 30 minutes at 37°C and with proteinase K at 55°C during 2 hours. Finally, DNA was purified by phenol-chloroform extraction followed by ethanol precipitation to perform qPCR. Usually DNA was resuspended in 50µL of H₂O UltraPure. Lastly, ChIP DNA was analysed by qPCR with SYBR Green (Roche) in a LightCycler 480 PCR system (Roche) using specific primers. Percentage of input was used for the quantification of the immunoprecipitated material with respect to the total starting chromatin. PHF8 antibody (Abcam, ab36068) was used for immunoprecipitation. In the case of ChIP-seq, DNA was purified by GenElute Mammalian Genomic DNA Purification Kit (G1N350-1KT) and the libraries were prepared and sequenced in a HiSeq 2000 sequencer (Illumina).

4.19.RNA-seq

RNA was extracted using High pure RNA isolation kit from Roche followed by DnaseI treatment from two biological independent samples. Libraries were prepared using the TruSeq Stranded Total RNA Sample Preparation kit with Ribo-Zero Human/Mouse/Rat Kit (Illumina, RS-122-2201/2) according to the manufacturer's protocol. Briefly, 500 ng of total RNA was used for ribosomal RNA depletion. Then, ribosomal RNA depleted RNA was fragmented for 4.5 min. The remaining steps of the library preparation were followed according to the instructions. The libraries were analysed using Agilent DNA 1000 chip to estimate the quantity and check size distribution, and they were quantified by qPCR using the KAPA Library Quantification Kit (Roche, 07960204001) prior to amplification with Illumina's cBot. The libraries were sequenced on Illumina High HiSeq 2500 with paired-end 50 base pair long reads.

5. Gas chromatography- mass spectrometry (GC-MS)

GC-MS metabolomic analysis was performed at the University of Barcelona. Control and PHF8 KD HeLa cells were harvested by trypsin

digestion, transferred to a microfuge tube, and frozen on dry ice. To extract the metabolites, 1 ml of iced 80% methanol in H₂O was added and vortexed during 1 minute, then the samples were put in an ultrasonic bath for 5 minutes. The process of extraction was repeated twice, then samples were exposed to infrared lamp and, once dried, they were derivatized adding 100 µL of MTBSTFA for 1 hour at 70°C. Four biological replicate samples ($3,5 \times 10^6$ /sample) were analysed for each condition. GC-MS analysis was performed with a Shimadzu GC-MS QP2010 gas chromatograph-mass spectrometer. Data were collected and recorded using the Shimadzu software GCMS solution version 2.54.

6. Electrophysiology

During my thesis I spent three months in the laboratory of Claudia Verderio (CNR Institute of Neuroscience, Italy) to perform some experiments in collaboration with her team; the electrophysiology experiments were performed by the postdoc Martina Gabrielli.

Whole-cell voltage clamp recordings were performed using a MultiClamp 700A amplifier (Molecular Devices) and a 1320A Digidata (Molecular Devices), coupled to a pCLAMP 10 Software (Molecular Devices), and using an inverted Axiovert 200 microscope (Zeiss). Miniature excitatory post-synaptic currents (mEPSC) were recorded from DIV 12-14 neurons in Krebs-Ringer's HEPES solution (KRH) (125 mM NaCl, 5 mM KCl, 1.2 mM MgSO₄, 1.2 mM KH₂PO₄, 2 mM CaCl₂, 6 mM D-glucose, and 25 mM HEPES/NaOH, pH 7.4), supplemented with 1 µM TTX and 20 µM bicucullin. Experiments were performed at room temperature (20–25 °C), setting the holding potential at – 70 mV and using the following internal solution: Potassium Gluconate (130 mM KGluc, 10 mM KCl, 1 mM EGTA, 10 mM Hepes, 2 mM MgCl₂, 4 mM MgATP, 0.3 mM Tris-GTP; pH 7.4, adjusted with KOH). Recording pipettes were pulled from patch-clamp borosilicate capillary glass (World Precision Instruments) to a tip resistance of 3–5 MΩ using a two-stage vertical puller (Narishige). Traces were sampled at 10 kHz and filtered at 2 kHz. Series resistance was monitored during recording. mEPSC were detected offline using Clampfit software (Molecular Devices) setting a threshold of 7 pA.

7. Bioinformatic methods

Along my doctoral thesis I collaborated with the team of Dr. Xavier de la Cruz (Vall d' Hebron Institute of Research VHIR, Barcelona) for bioinformatic analysis, which were performed by the PhD student Natalia Padilla. Nonetheless, some analysis has been performed by me using bioinformatic tools like Gene ontology [1, 207] and Heatmapper [208].

7.1. *ChIP-seq analysis*

For PHF8 ChIP-seq, 50 base pairs sequences were mapped to the *Mus musculus* genome release 10 (mm10) using Bowtie2 [209], files were filtered to remove duplicates and peaks were called using MACS [210] with an effective genome size of 1870000000 and a p-value of 0.001 for PHF8 ChIP-seq. The Bioconductor package ChIPseeker [211] was used to annotate the genes of each peak. Specifically, the function `annotatePeak` matches peaks with genomic features extracted from mm10 (UCSC) and calculates the proportion of peaks matching each feature. ChIP-seq data have been deposited in the GEO database under the accession GSE141969.

7.2. *Gene ontology analysis*

Gene ontology (GO) was used to perform GO enrichment analysis [212]; this service connects to the analysis tool from the PANTHER Classification System, which is maintained with GO annotations. It uses the 'overrepresentation test' which takes the input list (and a 'reference' list), and performs a statistical test for over- and under representation. It uses Fisher's exact test by default and calculates the false discovery rate (FDR), that is designed to control the false positive rate in the statistical test results.

7.3. *Gene expression omnibus accessions*

In this thesis we used previously published genome wide data which are deposited in the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). Accession numbers are indicated in the following table.

Dataset	Accession number
RNA-seq from day E10.5 mouse forebrain	GSE88173
RNA-seq from mouse neurons	GSM1269905, GSM1269906
RNA-seq from mouse oligodendrocytes	GSM1269912, GSM1269911

PHF8 expression profiling by array in HeLa	GSE38175
PHF8 ChIP-seq in HeLa	GSM520381
PHF8 ChIP-seq in H1-hESC	GSM1003509
c-Myc ChIP-seq in HeLa	GSM935320
H3K4me3 ChIP-seq	GSM566169
H4K20me1 ChIP-seq	GSM558474
H3K9me2 ChIP-seq	GSM1846169
H3K9me3 ChIP-seq	GSM566171

7.4. ChIP-seq capture obtaining

Along the Results section I will include captions of the ChIP-seq signal of different proteins to observe specific genomic coordinates. These captions were obtained by loading bigwig files of the different experiments in the IGV genome browser from the Broad Institute [213] or in the UCSC browser from the University of California Santa Cruz [214].

7.5. RNA-seq analysis

Alignment was performed using the Spliced Transcripts Alignment to a Reference software (STAR) [215]. The assignment of aligned reads to genes was performed using HTSeq [216] and DESeq2 was used to assess differential expression analysis [217]. The DESeq2 method for differential analysis of count data use shrinkage estimation for dispersions and fold changes to improve stability and interpretability of estimates. This enables a more quantitative analysis focused on the strength rather than the mere presence of differential expression. RNA-seq data have been deposited in the GEO database under the accession GSE141970.

7.6. Heatmap construction

To generate heatmap, Heatmapper web server was used. The “expression function” within Heatmapper allows the user to view expression data such as that from RNA-seq. It permits to choose the method for computing hierarchical clustering and the method for computing distance between rows and columns. To create the heatmaps, it was used the Euclidean method in which the distance is computed as the length of the line segment connecting two values.

8. Statistical analysis

In this last subsection I will provide details on which statistical tests have been applied to assess the reproducibility and significance of the results.

8.1. *Sample size*

As a general rule, experiments have been performed in triplicate. In specific cases like validation of ChIP-seq and RNA-seq data by qPCR, the number of replicates were two. The electrophysiology experiments were performed in three independent experiments with a total n=34 (17 control; 17 PHF8 KD condition).

8.2. *Standard deviation and standard error of the mean*

Along the experimental work, the graphics corresponding to experiments that fit a linear model have been represented as the mean. Error bars correspond to the standard deviation (SD) in the case of indirect immunofluorescences and RT-qPCR assays, and to the standard error of the mean (SEM) in the case of ChIPs. The numeric values have been calculated with Microsoft Excel software or Prism- GraphPad.

8.3. *Student's t-test*

To assess the significance of the results that follow a linear model, it was performed the Student's t-test. It was established that an experiment is statistically significant when within a 95% of confidence the result represented a true hypothesis. Asterisks represent the different p-values resulting from this test and calculation of the values was carried out with Microsoft Excel software. * p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001.

8.4. *ANOVA and Tukey's multiple comparisons test*

To assess the significance of differences in the immunofluorescence experiments with three or more conditions it was applied the one-way analysis of variance (ANOVA) followed by Tukey's multiple comparisons test [218]. The ANOVA was used to evaluate whether there was any evidence that the means of the populations differ. When the ANOVA led to the conclusion that there was evidence that the group means differ, the

Tukey multiple comparison test was applied. This test compares the difference between each pair of means with appropriate adjustment for the multiple testing.

8.5. Mann-Whitney Rank Sum Test

Mann-Whitney Rank Sum test was used in the electrophysiology experiments; it is a nonparametric test of the null hypothesis that it is equally likely that a randomly selected value from one population will be less than or greater than a randomly selected value from a second population. It can be used to investigate whether two independent samples were selected from populations having the same distribution [219].

Results

1. Characterization of PHF8 function in astrocytes

As mentioned in the introduction, previous results from our laboratory showed that PHF8 histone demethylase activity is essential for proper neurites outgrowth in mouse primary cortical neurons [94]. However, very little is known about the role of PHF8 in glial lineage specification and function. One of the aims of my thesis has been to figure out if PHF8 contributes to astrocyte differentiation and unveil the impact of its depletion on astrocyte development and function.

1.1. Study of PHF8 expression during astrocyte differentiation

Analysis of public available data of human [220] and mouse [221] neural cells showed that astrocytes express high levels of PHF8 in both species (Figure R1). We found especially interesting that in human the highest levels of PHF8 are found in fetal astrocytes; it suggests a potential role for PHF8 in early events of astrocytes differentiation.

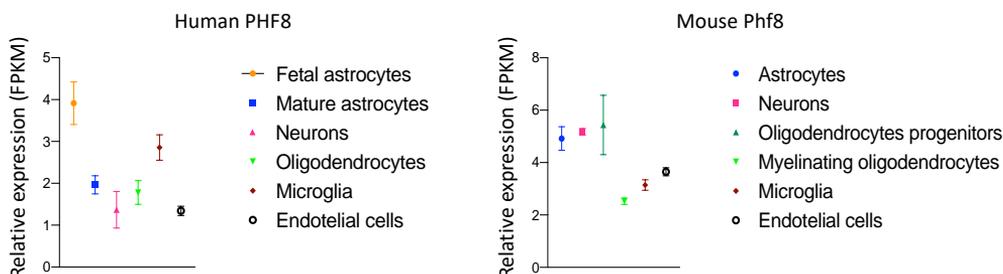


Figure R1: Figure shows public data of human (right panel) and mouse (left panel) RNA-seq experiments. Relative expression is shown by fragments per kilobase of transcript per million reads mapped (FPKM). Data are available at <http://www.brainrnaseq.org>.

To evaluate the relevance of PHF8 during astrocyte differentiation, we employed NSCs from cortices of mouse embryos at E12.5 [197, 222] (see methods), we differentiated them to astrocytes following the protocol described in methods. Briefly, once obtained the NSCs, the proliferating medium (containing EFG and EGF) was replaced by the astrocytic one containing DMEM/F-12, N2, Glutamax, FBS and FGF (Figure R2).

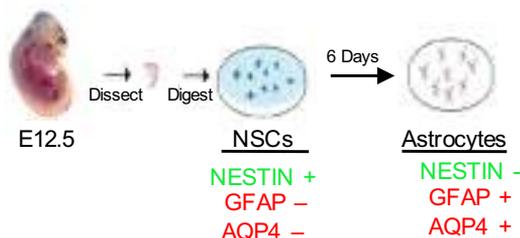


Figure R2: Schematic picture of the model used to study PHF8 function during astrocyte differentiation and the main markers characterizing NSCs and astrocytes.

To test the efficiency of the differentiation protocol from NSCs into astrocytes, we analysed the expression of the progenitor marker NESTIN, a cytoskeletal intermediate filament characterizing neural stem cells, and the expression of two astrocytic protein: the astrocyte associated glial fibrillary acidic protein (GFAP), which encodes one of the major intermediate filament proteins of astrocytes, and Aquaporin4 (AQP4), the predominant water channel expressed by astrocytes [223]. The results in Figure R3 show that after 6 days in astrocytic medium 100% of NSCs lost the expression of the progenitor marker NESTIN and acquired high levels of GFAP and AQP4.

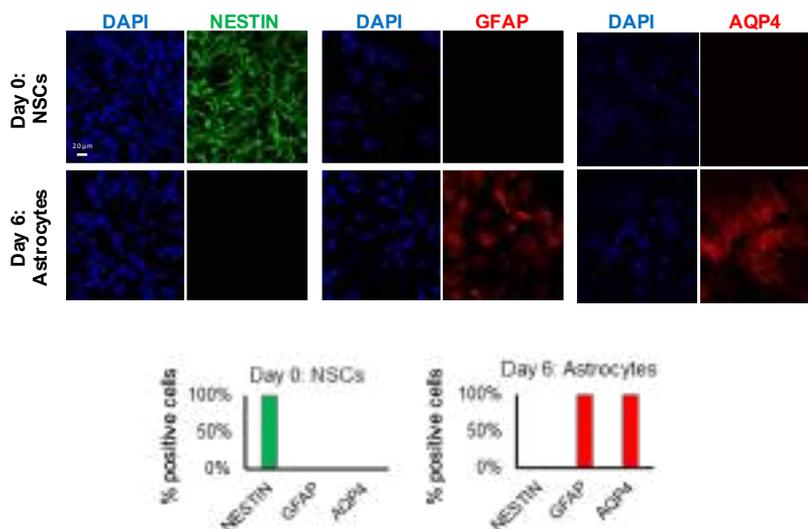


Figure R3: Immunostaining assays using NESTIN, GFAP, AQP4 antibodies and DAPI of NSCs (day 0) and astrocytes (day 6). Quantification indicating the percentage of positive cells is shown on the bottom of the Figure.

We also checked the levels of β -Tubulin III (TuB β 3), a microtubule protein of the tubulin family found almost exclusively in neurons, during the differentiation protocol. Although some TuB β 3- expressing cells were detected at day 3 and 5, they disappeared at 6 days of differentiation which is the timepoint in which we performed most of the experiments, as we observed that almost all cells differentiated to astrocytes (Figure R4).

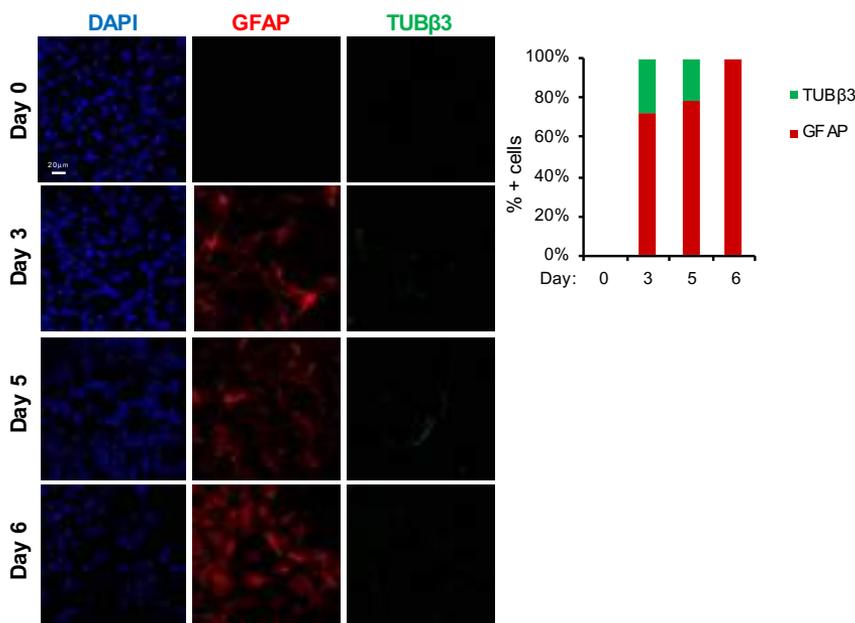


Figure R4: Immunostaining assays using GFAP and TUB β 3 antibodies and DAPI of NSCs maintained in astrocytic differentiation medium at day 0, 3, 5 and 6. Quantification indicating the percentage of positive cells is shown on the right side of the Figure.

To further characterize the differentiation process, we assessed by qPCR the expression of *Nfia*, the master transcription factor that drives the onset of astrogenesis, *S100b*, a well characterized marker of astrocytes, and *Gfap* during differentiation. We observed progressive increase in the expression of the three astrocytic genes but not of *Olig2*, a transcription factor of oligodendroglial cells which we used as a negative control (Figure R5).

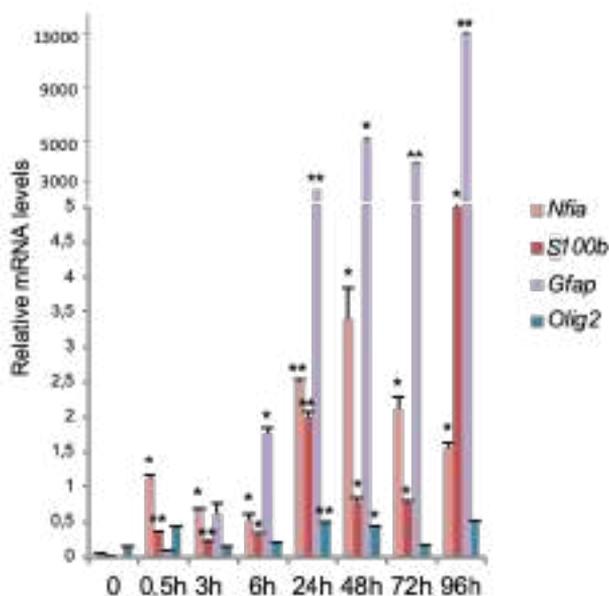


Figure R5: NSCs were maintained in astrocytic differentiation medium for the indicated times. Total RNA was prepared and the levels *Nfia*, *S100b*, *Gfap* and *Olig2* genes were determined by qPCR. Expression values were normalized to the housekeeping gene *Gapdh*. Figure shows values relative to time 0.

Then, we performed immunostaining assays of H4K20me1, H4K20me3, H3K9me2 and H3K27me3 histone modifications to analyse the changes in the chromatin during the differentiation from NSCs to astrocytes. We observed that these marks associated to heterochromatin formation increased during the differentiation process (Figure R6).

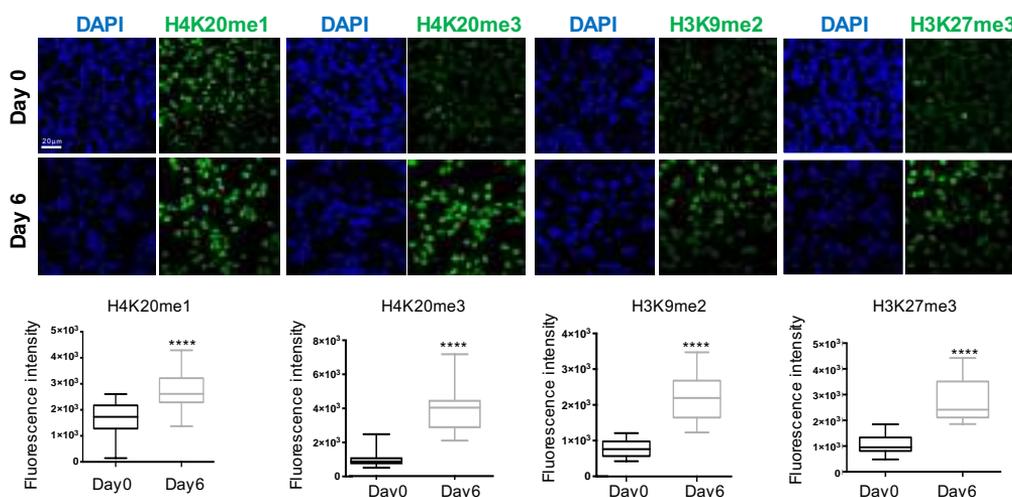


Figure R6: Immunostaining assays using the antibodies against H4K20me1/3, H3K9me2 and H3K27me3 histone marks and DAPI in NSCs (day 0) and astrocytes (day 6). The quantification is shown on the bottom of the Figure.

As we noticed that PHF8 was highly expressed in astrocytes (Figure R1), we analysed its expression in astrocytes derived from NSCs and all along the differentiation process. Result in Figure R7 shows a clear increase of PHF8 protein levels at early differentiation time (8h) and in fully differentiated astrocytes (6 days upon differentiation).

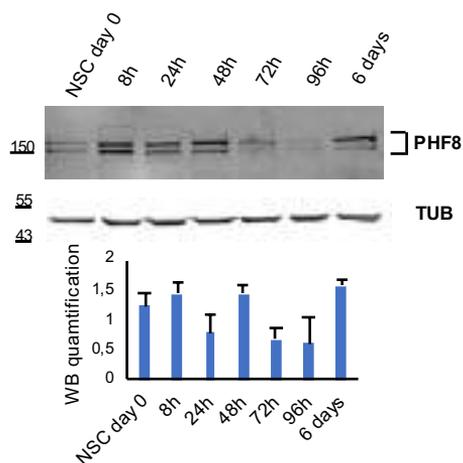


Figure R7: NSCs were maintained in astrocytic differentiation medium for the indicated times. The levels of PHF8 were determined by immunoblot with PHF8 antibody. The PHF8 protein levels normalized to TUBULIN are shown on the bottom of the Figure.

Then we confirmed that PHF8 is also upregulated *in vivo* during astrocytes differentiation. We performed Western blot assays of MACS-isolated astrocytes (GLAST-positive cells) from early development (E18) and intact postnatal mouse brain (P2 and P7) (Figure R8).

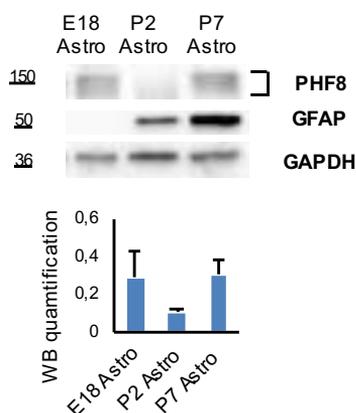


Figure R8: PHF8 Western blot assays of MACS-isolated astrocytes (GLAST-positive from E18 and postnatal mouse brain (P2 and P7). GAPDH antibody was used as loading control. The PHF8 protein levels normalized to GAPDH are shown on the bottom of the Figure.

Next, we moved onward to identify which signaling pathway could be responsible for *Phf8* full expression. Since Notch, TGF β , and BMP signals are involved in astrocyte differentiation, we activated these pathways and analysed the *Phf8* mRNA levels by qPCR. To do that we transfected or treated 293T cells (due to the difficulty to transfect NSCs) with one of the following components; the Notch intracellular domain (NICD), Tgf β , BMP4 and BMP7. The results in Figure R9 show a clear *Phf8* induction after transfection of the NICD which, once activated, regulates Notch target genes, meaning that *Phf8* is indeed regulated by Notch signaling.

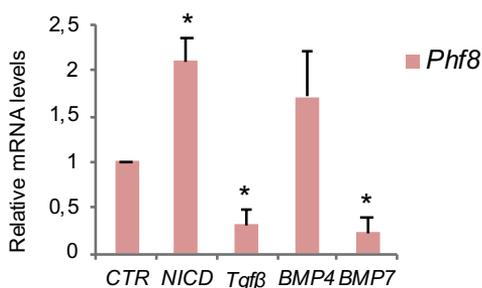


Figure R9: HEK 293T cells were transfected or treated with NICD, Tgf β , BMP7, BMP4 or empty vector (CTR) as indicated. Expression values were normalized to the housekeeping gene *Gapdh*. Figure shows values relative to empty vector (CTR).

As Notch pathway is crucial to induce astrocyte differentiation [224, 225], we analysed whether PHF8 could modulate Notch activity. To test it, we overexpressed *Phf8* upon Notch activation (by NICD transfection) and analysed the expression of the well-known Notch-target, *Hes5*, by qPCR. Results in Figure R10 show that PHF8 facilitated Notch target activation at short times (12h) (Figure R10 left), while it inhibited Notch target genes expression at longer times (48h) (Figure R10 right).

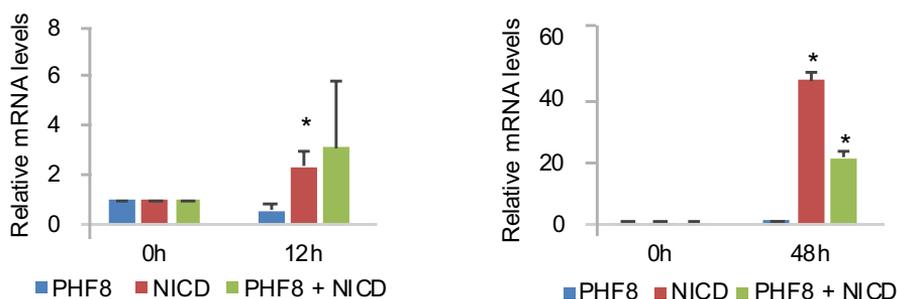


Figure R10: HEK 293T cells were transfected with a vector expressing *Phf8* together or not with *Nicd*. Total mRNA was purified and the *Hes5* levels at 0h, 12h and 48h were established by qPCR. Figure shows values relative to time 0h. Expression values were normalized to the housekeeping gene *Gapdh*.

To make this data stronger we used a luciferase reporter vector fused to *Hes5* promoter and activated Notch pathway (through NICD transfection) both in HEK 293T cells that overexpressed PHF8 (Figure R11 left) and in cells in which we depleted PHF8 by shRNA (Figure R11 right). Interestingly, performing luciferase assay, we observed that PHF8 represses Notch target activation, which is in accordance with the results obtained by qPCR at 48h (Figure R10 right).

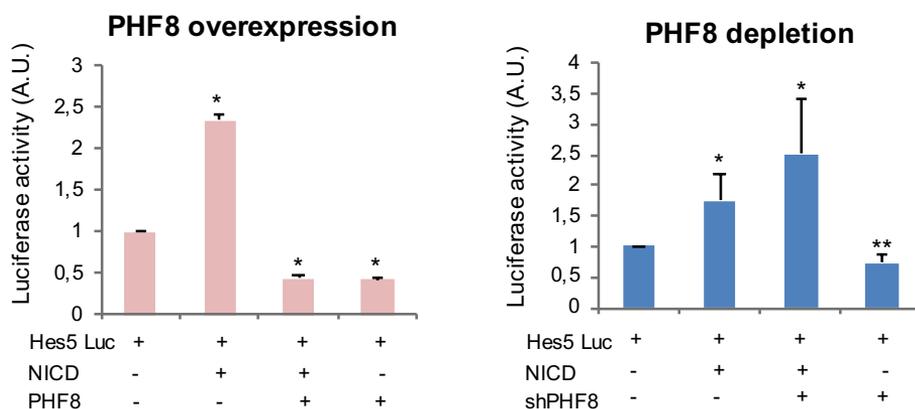


Figure R11: HEK 293T cells were transfected with a *Hes5* promoter fused to luciferase reporter vector alone or together with NICD, PHF8 or both, as indicated. *Hes5* promoter activity was quantified by the luciferase activity (left panel). In the right panel shPHF8 vector was used instead PHF8 expressing plasmid. Luciferase activity values were normalized to the *Renilla* levels used as a transfection internal control.

1.2. *PHF8* regulates transcription during astrocyte differentiation

To acquire further knowledge into PHF8's function during astrocytic differentiation, we analysed the PHF8-dependent transcriptional profile by RNA-sequencing. To this end, NSCs were transduced with a lentivirus containing a control shRNA or a specific PHF8 shRNA that efficiently decreased PHF8 protein levels (Figure R12).

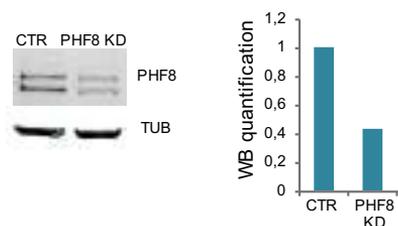


Figure R12: NSCs were infected with lentivirus expressing shRNA control or shRNA specific for PHF8. Total protein extracts were prepared and the PHF8 levels were determined by immunoblot.

Next, control (CTR) and PHF8-depleted (PHF8 KD) NSCs were differentiated into astrocytes, following the protocol described in methods (Figure R13 left). After 6 days in differentiation medium, we confirmed that astrocytes from PHF8 KD NSCs exhibited decreased PHF8 expression compared to those derived from control NSCs by qPCR (Figure R13 right).

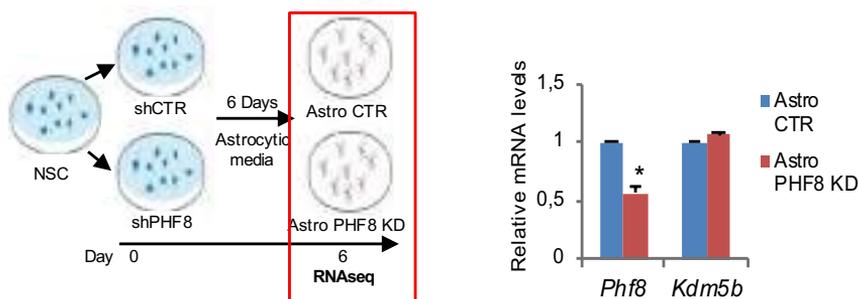


Figure R13: Schematic view of the experiment to study PHF8 transcriptional profile in astrocyte (top panel). In the bottom panel NSCs were infected with lentivirus expressing shRNA control or shRNA specific for PHF8. After 6 days in astrocytic medium, total RNA was purified and the *Phf8* levels were determined by qPCR. Expression values were normalized to the housekeeping gene *Gapdh*. Figure shows values relative to Astro CTR.

To identify the PHF8-dependent transcriptional profile, we purified total RNA from two control (Astro CTR) and two PHF8-depleted (Astro PHF8 KD) astrocytes samples and performed poly-A RNA sequencing in the CRG/CNAG genomics unit. In Figure R14 are represented the clustered heatmap depicting Pearson correlation (left panel) and the principal component analysis plot of the two samples (right panel) respectively.

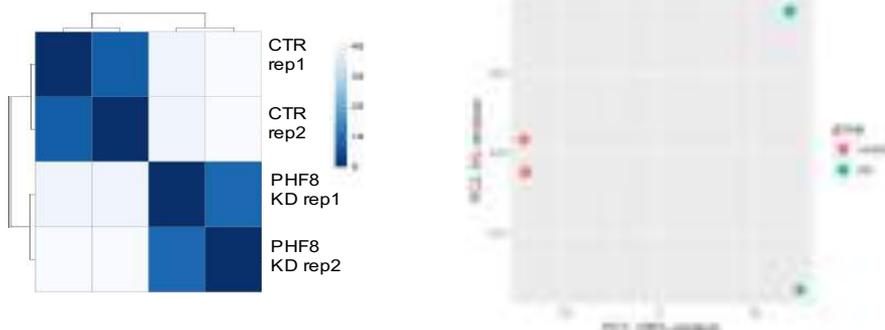


Figure R14: Clustered heatmap showing Pearson correlation of the two samples Astro PHF8 KD and Astro CTR based on read coverage within genomic regions (left).

Principal component analysis plot of normalized RNA-seq read counts. PC1 shows 100% of the total variance and separates treated samples from control samples (right).

The transcriptional profiles of two control (Astro CTR) and two PHF8-depleted (Astro PHF8 KD) astrocytes samples showed that 4987 transcripts significantly changed their expression [\log_2 fold change (FC) $>0,5$ and (FC) $<-0,5$ and p-value $<0,08$] in two biological independent experiments (Figure R15).

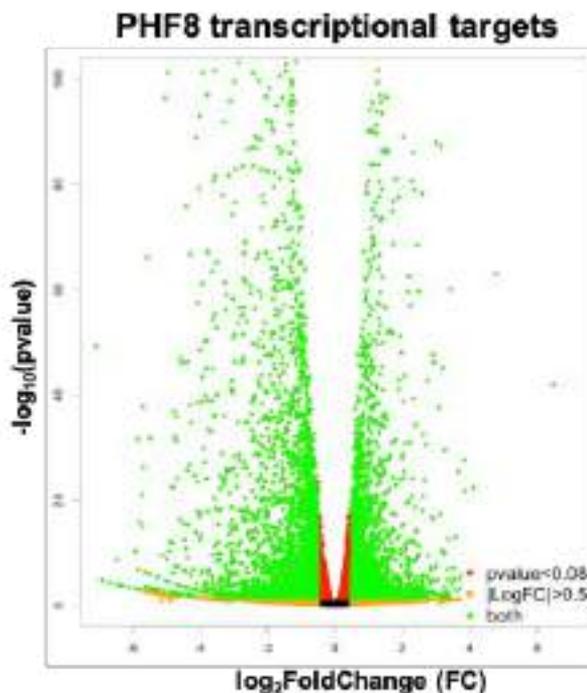


Figure R15: Volcano plot represents PHF8 transcriptional targets identified by RNA-seq in Astro CTR and Astro PHF8 KD. The green dots represent all the genes with p-value $<0,08$ and \log_2 fold change (FC) $>0,5$ and (FC) $<-0,5$.

In the Figure R16 is represented the heatmap showing the top 30 regulated genes identified by RNA-seq in Astro CTR and Astro PHF8 KD.

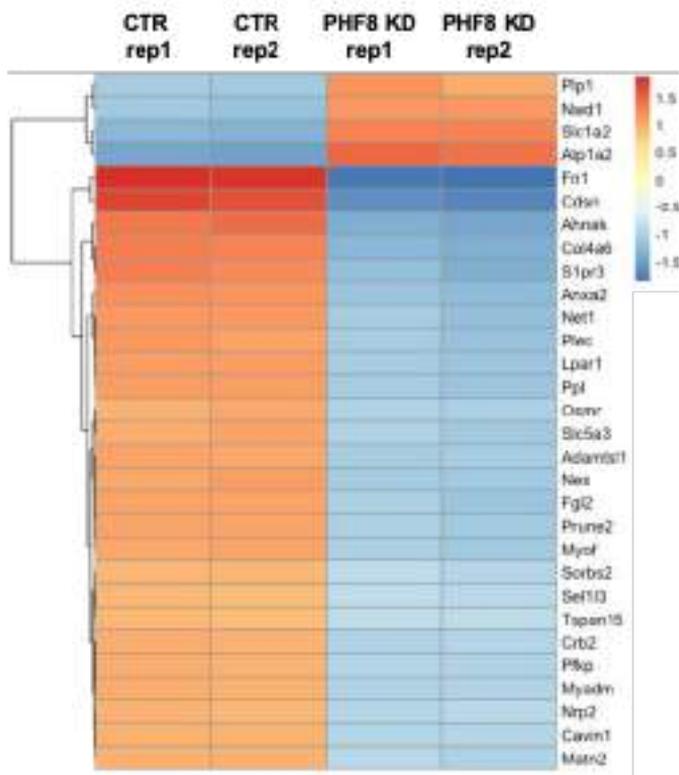


Figure R16: Heatmap showing the top 30 regulated genes identified by RNA-seq in Astro CTR and Astro PHF8 KD.

Of the 4986 differentially expressed transcripts upon PHF8 depletion, 2899 (58%) were downregulated and 2087 (42%) upregulated (Figure R17 left). We observed that, when we shifted the \log_2 fold change from 0,8 to higher values (Figure R17 right), the percentage of downregulated genes tends to increase, suggesting that PHF8 mainly acts as an activator in astrocytes.

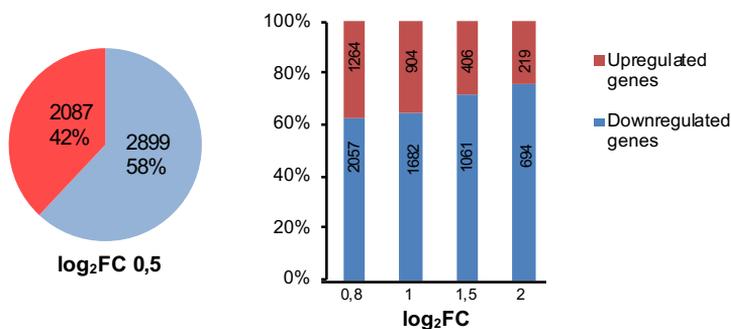
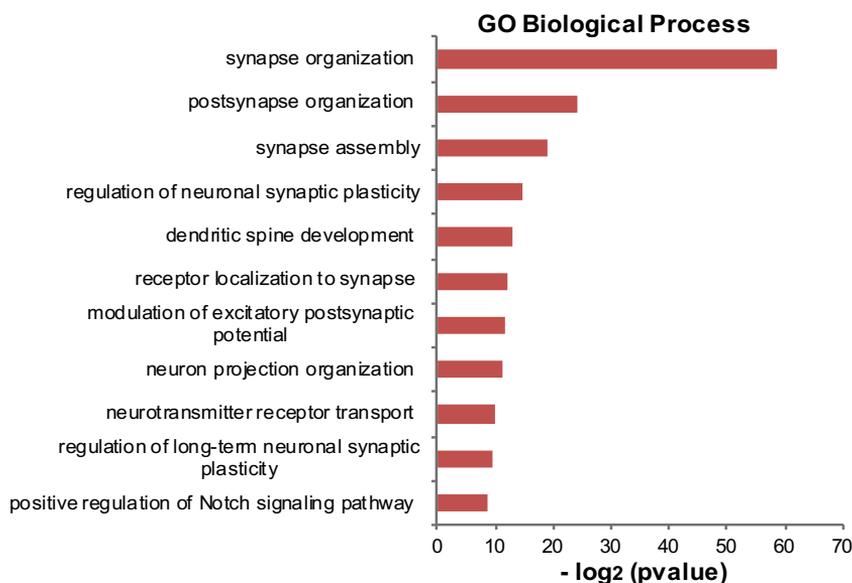


Figure R17: Graph depicting the percentage of upregulated and downregulated genes in the astro PHF8 KD compared to astro CTR with p-value <0,08 and classified by increasing log₂ fold change (FC).

To further characterize the differences between control and PHF8 KD astrocytes, we performed a gene ontology (GO) enrichment analysis of regulated genes to identify those biological processes and cellular components most sensitive to PHF8 depletion. The analysis shown in Figure R18 revealed enrichment in the categories of genes involved in synapse formation and maturation, postsynaptic density membrane and, in general, synapse organization including the astrocytic genes *Gpc4*, *Sparc*, *Thbs1*, *Nrxn1*, *Pcdh8*, *Sdc4* [171-173, 226].



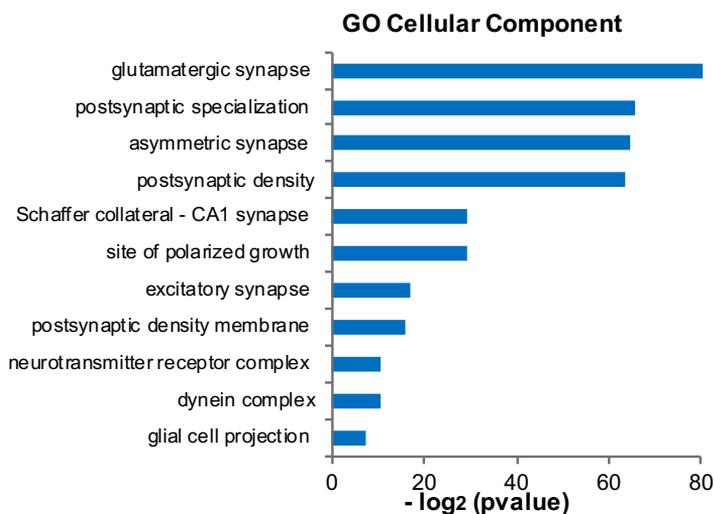


Figure R18: Gene ontology analysis showing the Biological Process and Cellular Component of the PHF8 regulated genes (p -value <0.08 and \log_2 FoldChange >0.5 and <-0.5) was performed using as a background the whole *Mus musculus* genome.

In addition, a master regulator of astrocytes differentiation, *Nfia* [144], was expressed at lower level in PHF8 KD Astro compared to CTR, pointing out that PHF8 may have a role in astrocytic differentiation (Figure R19).

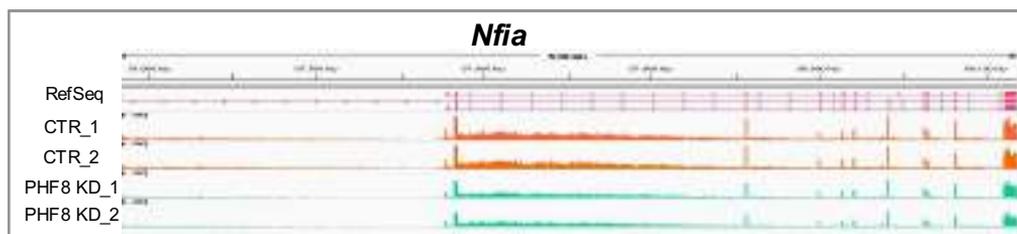


Figure R19: IGV capture showing *Nfia* RNA levels in Astro CTR and Astro PHF8 KD.

In Figure R20 are represented IGV captures showing RNA levels of genes expressed in astrocytes and involved in synaptogenesis such as *Thbs1*, *Sparc* and *Gpc4*. These genes are less expressed in Astro PHF8 KD compared to Astro CTR.

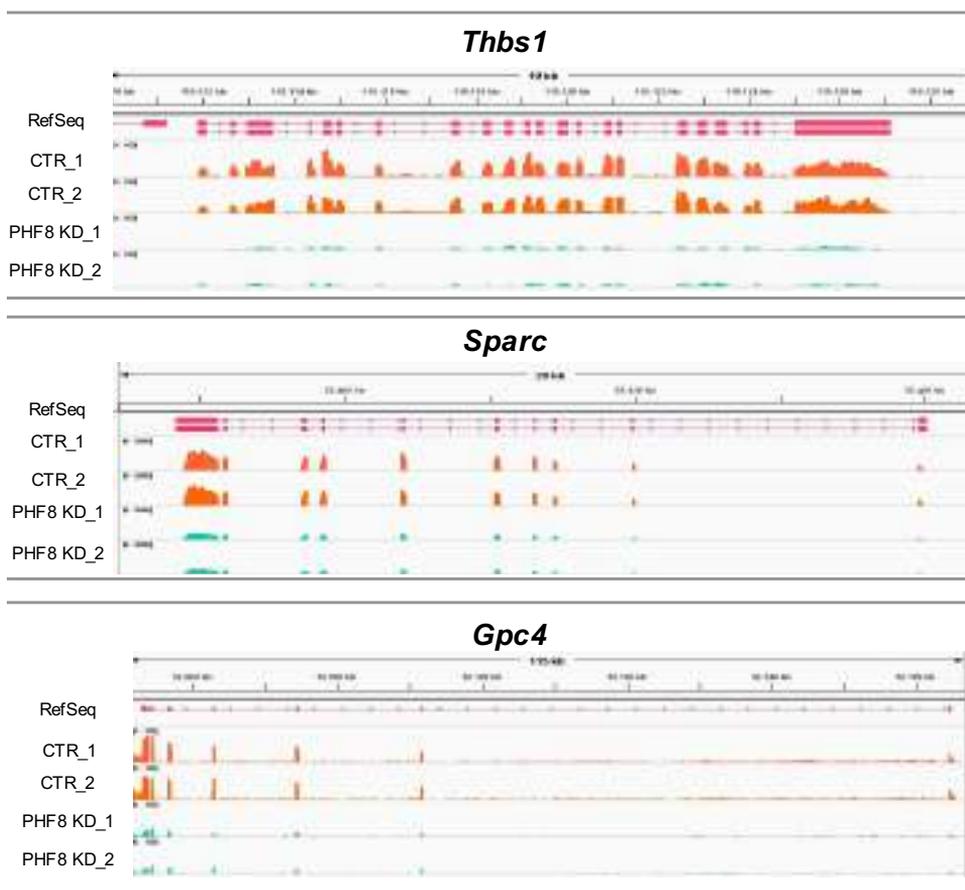


Figure R20: IGV captures showing RNA levels in the two replicates of Astro CTR and Astro PHF8 KD.

Noteworthy, genes involved in synapse formation and maturation were both downregulated and upregulated in PHF8-depleted astrocytes even if the downregulated ones had higher log₂ fold changes. Importantly, when we compared a published list of astrocytic genes, which have a role in neuron-astrocyte interplay at synapses [227], with the differentially expressed transcripts found in the RNA-seq, we detected that 57% of those key genes in synaptogenesis were affected in PHF8 KD astrocytes (Figure R21).

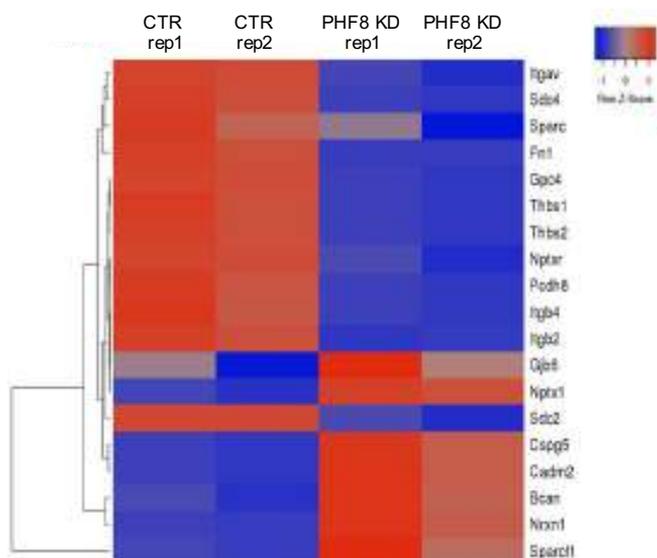


Figure R21: Heatmap showing synapse related genes identified by RNA-seq experiment in the two biological replicates of Astro PHF8 KD and CTR.

Interestingly, in the GO Biological process analysis of the RNA-seq (Figure R18) we found the category of positive regulation of Notch signaling pathway and this pathway is crucial to induce astrocyte differentiation. So, we checked some Notch targets genes in the RNA-seq (*Hes5*, *Dll3*, *Dll1*, *Notch3*, *Rbpj*) and we actually found that those transcripts were misregulated in PHF8 KD astrocytes (Figure R22).

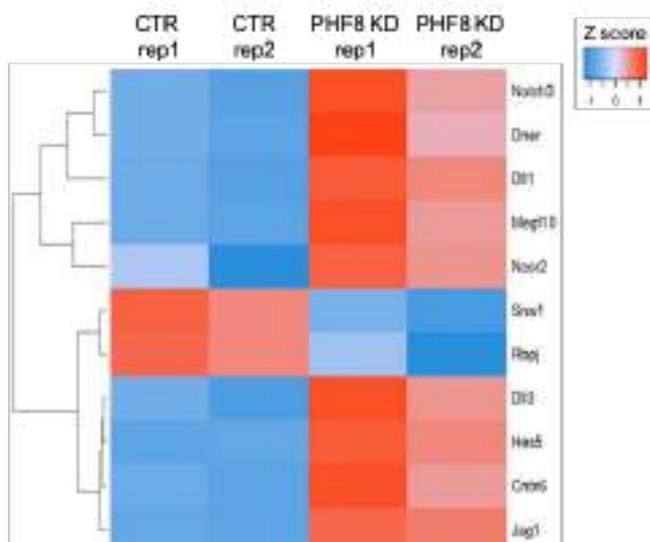


Figure R22: Heatmap showing some Notch target genes identified by RNA-seq in the two biological replicates of Astro PHF8 KD and CTR.

Finally, we validated those results by measuring the expression of some genes by qPCR of control and PHF8 KD astrocytes (Figure R23).

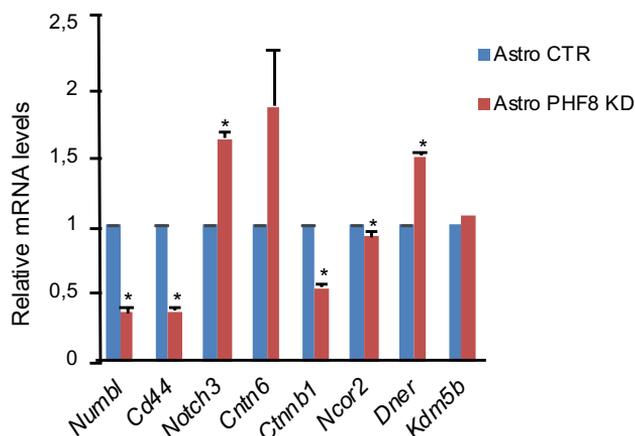


Figure R23: Notch related genes in PHF8 KD vs CTR astrocytes identified in the RNA-seq experiment was validated by qPCR assays. Values were normalized to the housekeeping gene *Gapdh*.

In summary these data suggest that PHF8 facilitates the transcription of the master regulator *Nfia* and of many synaptogenic genes expressed by astrocytes (*Thbs1*, *Sparc*, *Gpc4*), moreover it regulates Notch target genes that are crucial in the induction of astrogenesis.

1.3. PHF8 binds to astrogenic and synaptogenic genes

We next determined PHF8 biological substrates to better understand PHF8's contribution to gene regulation. To do that we differentiated NSCs to astrocytes during 6 days and performed chromatin immunoprecipitation followed by sequencing in the CRG/CNAG services (Figure R24).

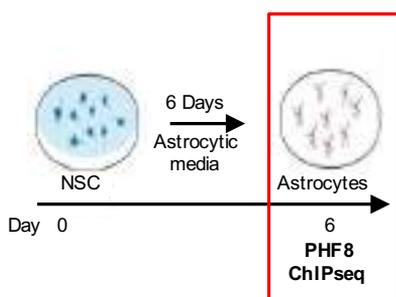


Figure R24: Scheme of the ChIP-seq experiment to identify the PHF8's binding sites in astrocytes.

Upon normalization to the input, 8401 peaks (p-value 0,001) were detected in ChIP-seq data for PHF8. The analysis of the genomic

distribution of those peaks revealed that 46,8% of PHF8 peaks are located in distal intergenic regions. The remaining 53% were located along the genome and particularly enriched at introns (Figure R25).

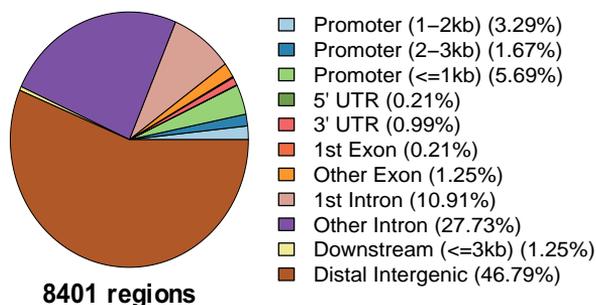


Figure R25: Genomic distribution of PHF8 ChIP-seq peaks in astrocytes.

Then, PHF8 ChIP-seq results were validated by qPCR analysis of 7 randomly chosen genes (Figure R26).

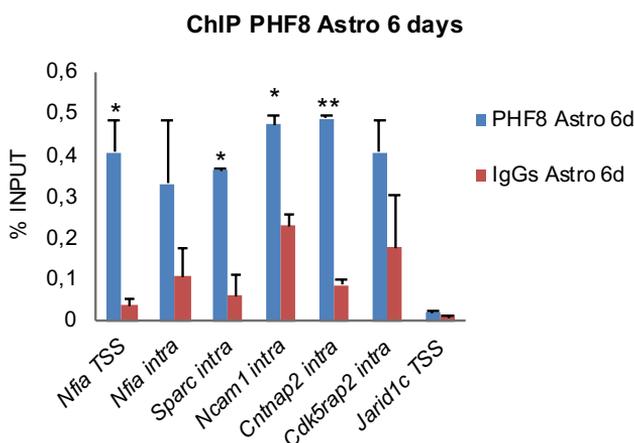


Figure R26: The levels of PHF8 at the indicated genes in astrocytes were determined by ChIP-qPCR. Data from qPCR were normalized to the input.

Moreover, we demonstrated the binding of PHF8 to genes essential for astrocyte differentiation and function (*Nfia*, *Sparc*) at early stages of differentiation (day 1) by ChIP qPCR too (Figure R27). The result indicates that, at initial phases, PHF8 is yet bound to key astrogenic genes for the differentiation process.

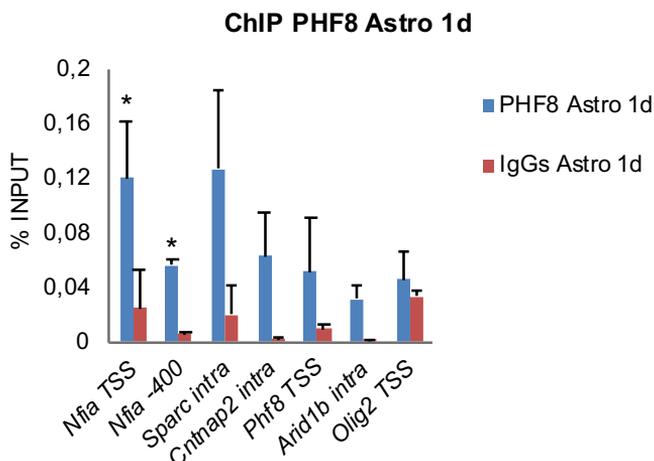


Figure R27: The levels of PHF8 at the indicated genes after 1 day in astrocyte differentiation media were determined by ChIP-qPCR assays. Data from qPCR were normalized to the input.

Interestingly, the PHF8 genomic distribution in postmitotic astrocytes was noticeably different if compared to the previously described distribution in a line of human embryonic stem cells (H1-hESCs). In fact, in H1-hESCs the majority of the PHF8 binding sites reside at promoters, around 65% (Figure R28).

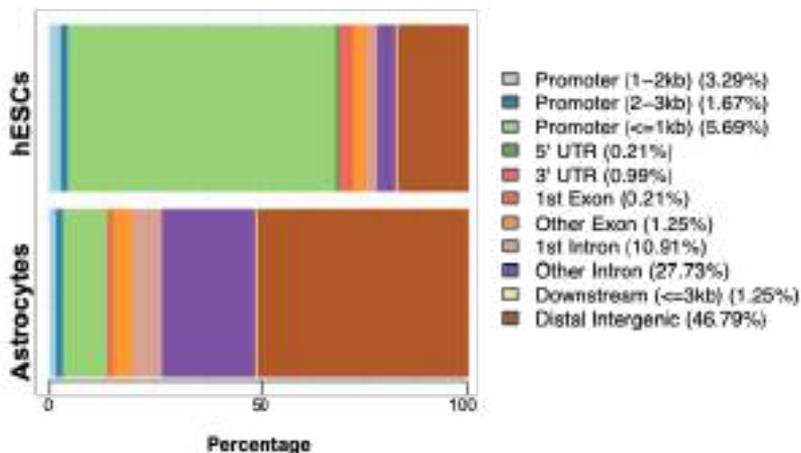


Figure R28: Diagram depicting the differential genomic distribution of PHF8 in ESCs and astrocytes.

Comparing the PHF8 genomic location in astrocytes with H1-hESCs, we observed that those peaks at promoters in H1-hESCs corresponded to the ones at introns and intergenic regions in astrocytes (Figure R29).

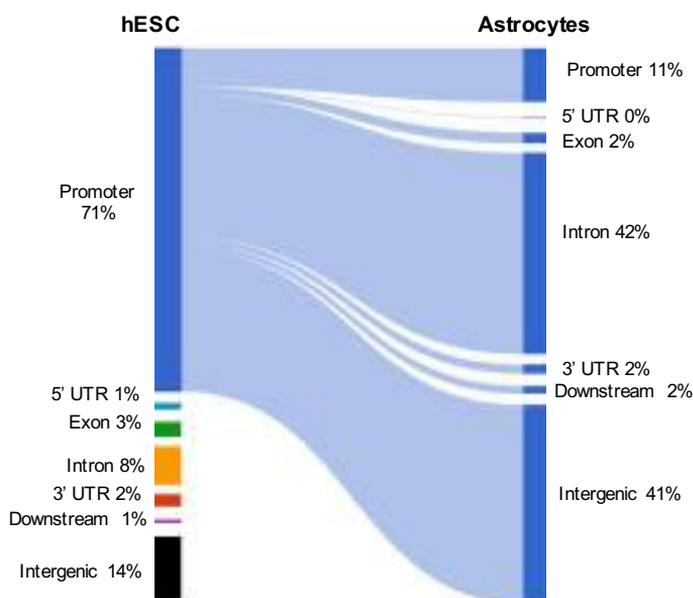
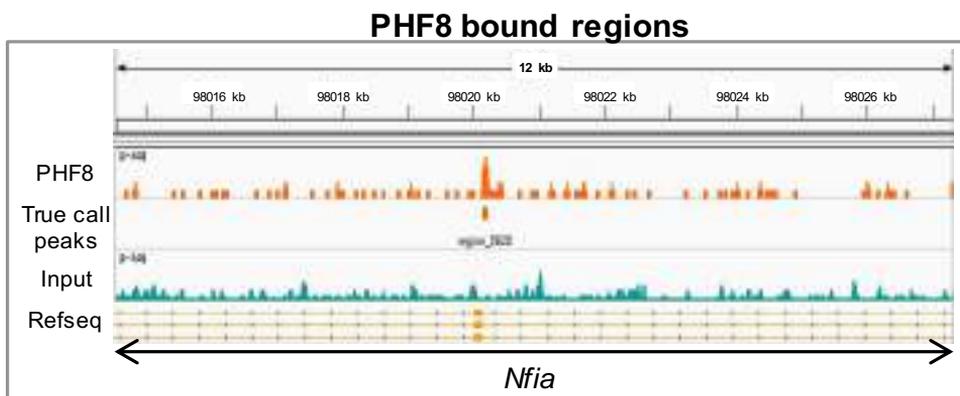


Figure R29: Diagram depicting that the PHF8 peaks on promoters in H1-hESCs are located mainly in introns and intergenic regions in astrocytes.

In Figure R30 are represented IGV captures from the ChIP-seq showing PHF8 peaks in *Nfia* and *Frdm3* genes in astrocytes.



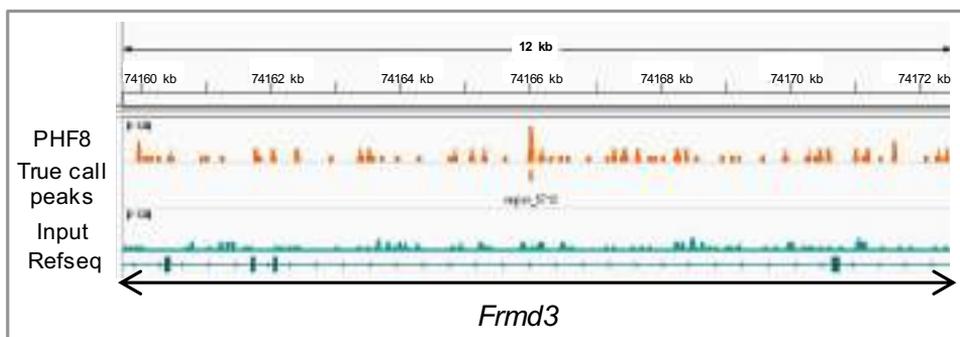


Figure R30: IGV captures showing PHF8 peaks in *Nfia* and *Frmd3* genes in astrocytes.

Gene ontology analysis of PHF8-bound regions indicated that PHF8 was associated to genes involved in neural development, particularly in the processes of astrogenesis (*Nfia*) and synaptogenesis (*Sparc*, *Gpc4*), synaptic transmission and assembly (Figure R31).

GO Biological Process

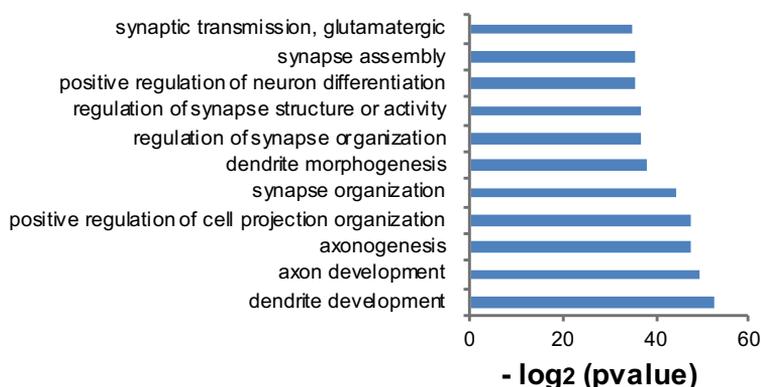


Figure R31: Gene ontology analysis showing Biological Process of PHF8-bound genes using as a background the whole *Mus musculus* genome.

To better understand how PHF8 is targeted to chromatin, we performed Homer motif enrichment analysis. We identified that one of the most statistically significant predicted PHF8 binding site was RBPJ1 DNA binding motif. Importantly, RBPJ1 is a transcription factor of the Notch signaling pathway which interacts with NICD in response to Notch activation (Figure R32).

Motif	Factor	p value
	Rbpj1	1e-4

Figure R32: Motif enrichment analysis of PHF8 ChIP-seq peaks in astrocytes using Homer known motif showing an enriched motif.

Interestingly, RBPJ1 motif and PHF8 binding were identified also at *Phf8* gene (Figure R33), suggesting that PHF8 is itself a target of Notch signaling (see also Figure R9).

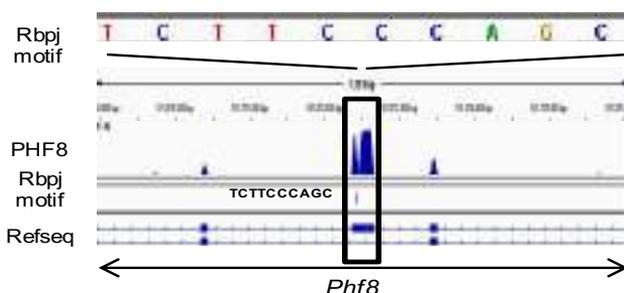


Figure R33: IGV capture showing PHF8 peaks and RBPJ1 binding motif in *Phf8* gene in astrocytes.

To better understand which are those regions bound by PHF8 and RBPJ1, we performed gene ontology analysis of the RBPJ1 DNA binding motif found in PHF8 ChIP-seq. The results showed that they bind genes involved in synapse assembly and function and neuron development (Figure R34).

GO Biological Process

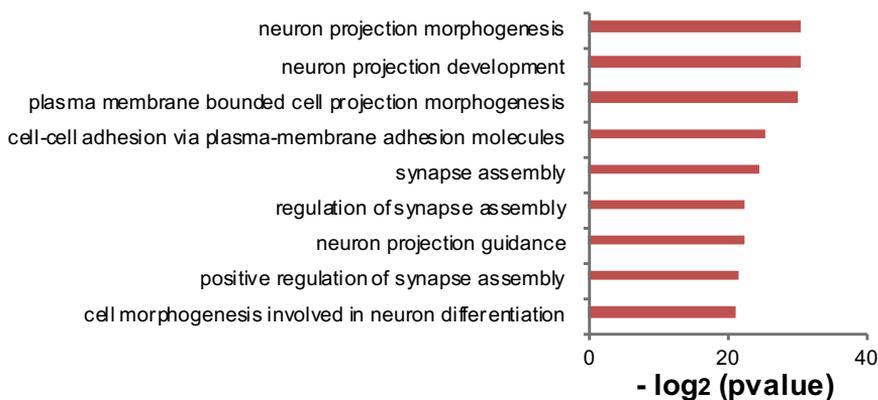


Figure R34: Gene ontology analysis showing the PHF8 bound genes enriched in RBPJ1 motif.

Next, we identified the PHF8 direct transcriptional targets by comparing the genes bound by PHF8 in the ChIP-seq experiment (4254) with the transcriptional profile ($\log_2FC > 0,5$ and $\log_2FC < -0,5$ and $p\text{-value} < 0,08$, 4987 transcripts). Among the genes bound by PHF8, 867 (20,3%) showed a PHF8-dependency for transcriptional regulation in the RNA-seq experiment (Figure R35).

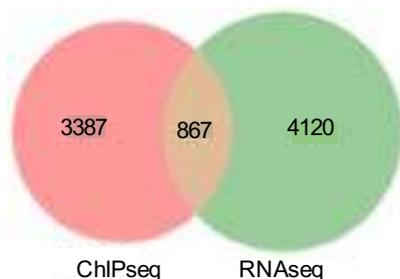


Figure R35: Venn diagram showing overlapping between PHF8 bound regions and PHF8 transcriptional targets.

Gene ontology analysis of the PHF8-direct target genes showed that the most enriched terms were related with astrogenic differentiation, synapse assembly and trans-synaptic signaling (Figure R36).

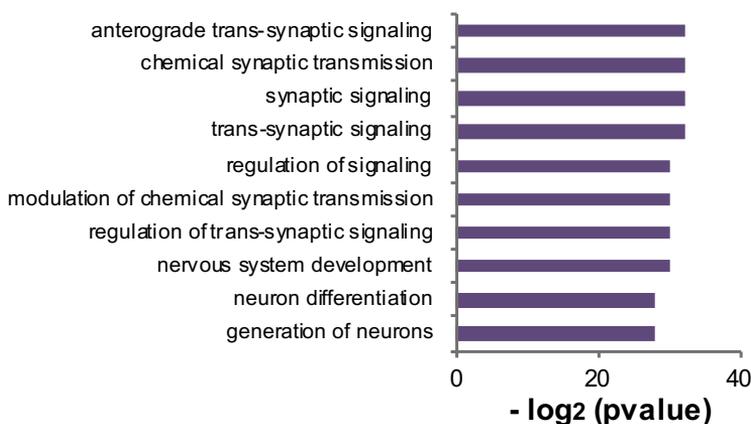


Figure R36: Gene ontology analysis showing biological process of the PHF8-direct target genes.

Altogether, these data support a model in which astrogenic PHF8 directly regulates astrocytes differentiation and synaptogenesis.

1.4. *PHF8* depleted NSCs differentiate into defective astrocytes

As we detected that PHF8 controls the expression of genes involved in astrocyte differentiation, especially *Nfia*, we explored whether PHF8 depletion impairs proper astrocyte differentiation (Figure R37).

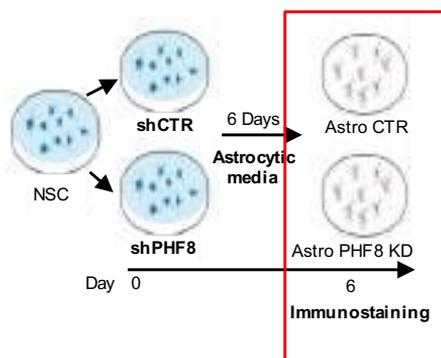
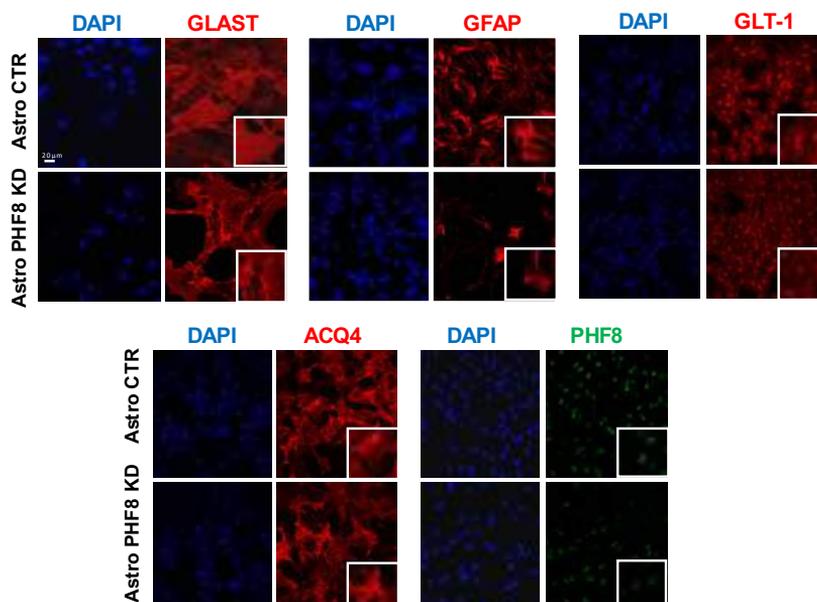


Figure R37: Scheme representing control and PHF8-depleted NSCs that were differentiated during 6 days to generate Astro CTR and Astro PHF8 KD before performing immunostaining assay.

We performed immunostaining assays for the well-known astrocytic markers glutamate aspartate transporter 1 (GLAST), GFAP, the glutamate aspartate transporter II (GLT-1, also known as EAAT2) and ACQ4. The results in Figure R38 revealed that PHF8 depletion caused a decrease in GFAP protein expression, with no considerable alterations in other astrocytic markers.



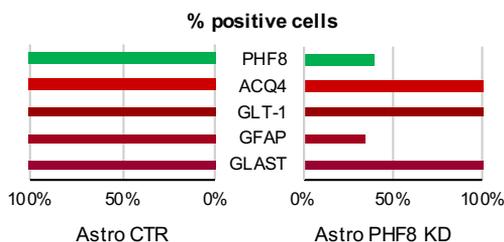


Figure R38: Control and PHF8-depleted NSCs were differentiated to astrocytes during 6 days to generate Astro CTR and Astro PHF8 KD respectively. Cells were fixed and stained with GLAST, GFAP, GLT-1, ACQ4 and PHF8 antibodies and DAPI. The % of cells expressing these markers in each population is shown on the bottom.

To better understand the phenotype of PHF8 KD astrocytes that seemed to be somehow defective, we compared the transcriptional profile of NSCs (GSE88173) with Astro CTR and Astro PHF8-KD. The cells resulting from differentiation upon PHF8 depletion had the major astrocyte signatures although they misexpressed some transcripts such as *Gfap*, *Aquaporin-4* and *Pbxip1* (Figure R39).

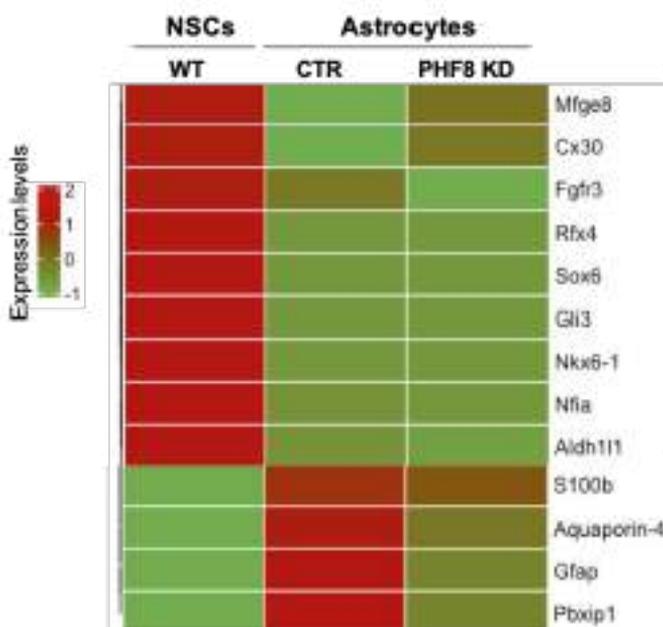


Figure R39: Heat map showing astrocytic gene expression identified by RNA-seq in PHF8 KD and CTR astrocytes compared to NSCs (GSE88173).

Moreover, some PHF8-defective astrocytes exhibited features of other neural cell lineages; we tested their immunoreactivity for oligodendrocytes (OLIG2, NG2 and GPR17), neurons (TUB β III), microglia (IBA1) and neural stem cells (NESTIN) markers. We obtained that a % of Astro PHF8 KD was positive for TUB β III (13%) or OLIG2 (33%) (Figure R40). Notably, under our differentiation conditions, 12% of Astro CTR expressed OLIG2, a percentage that was higher in Astro PHF8 KD (33%). However, the lack of signal for other oligodendrocyte markers (GPR17) and the presence of astrocytic markers (GLAST, GLT-1 and ACQ4) (Figure R38) indicate that they were not oligodendrocytes.

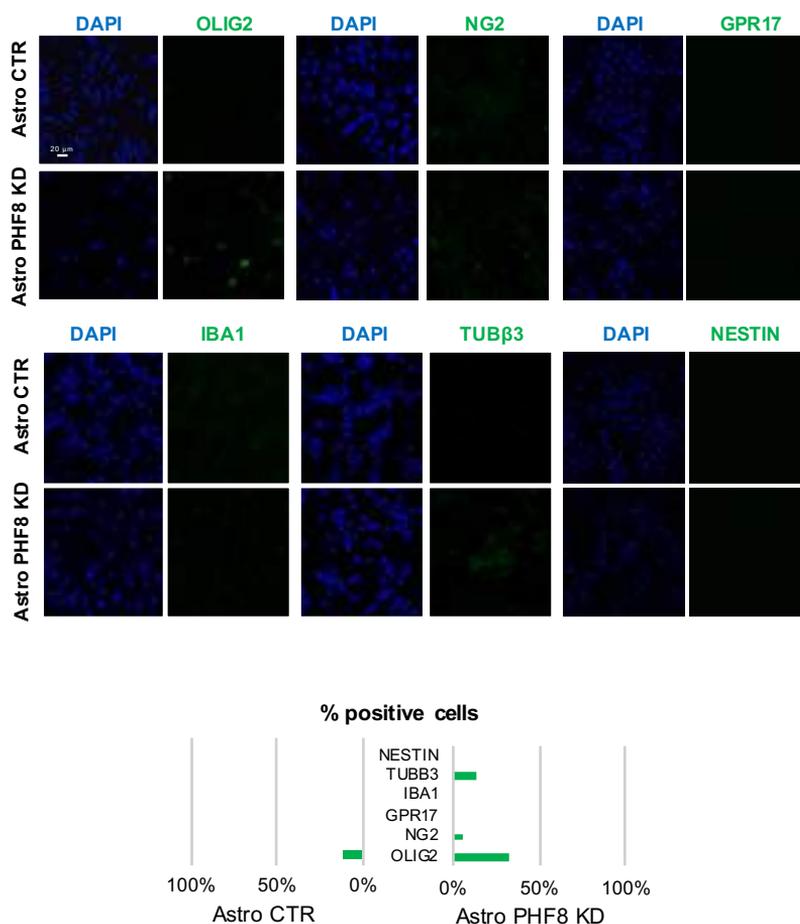


Figure R40: Control and PHF8-depleted NSCs were differentiated into astrocytes during 6 days to generate Astro CTR and Astro PHF8 KD. Cells were fixed and stained with OLIG2, NG2, GPR17, IBA1, TUB β 3 and NESTIN antibodies and DAPI. The % of cells expressing the markers is shown at the right of the figure.

Then we compared the transcriptional profiles of astrocytes CTR and PHF8 KD with previous published RNA-seq of oligodendrocytes and neurons (Figure R41). We found that both CTR and PHF8 KD astrocytes were clearly different, at transcriptional level, from oligodendrocytes or neurons, reinforcing the idea that PHF8-depleted cells differentiated into defective astrocytes.

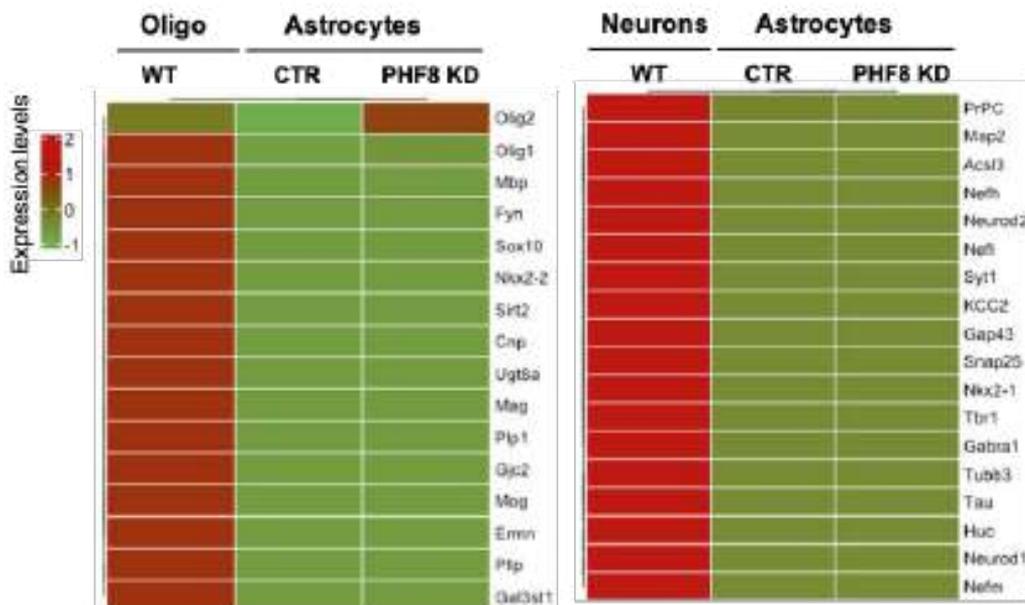


Figure R41: Heat maps showing oligodendrocytic and neuronal related gene expression identified by RNA-seq in CTR and PHF8 KD astrocytes compared to oligodendrocytes (GSM1269912, GSM1269911) and neurons (GSM1269905, GSM1269906) respectively.

As we observed that all PHF8 KD astrocytes properly expressed GLAST, GLT-1 and AQP4 but, at the same time, misexpressed oligodendrocytes and neuronal markers, we performed double immunostaining. The result in Figure R42 shows that a small % of Astro PHF8 KD expressed multiple lineage markers: TUB β III and GLAST; OLIG2 and GLAST.

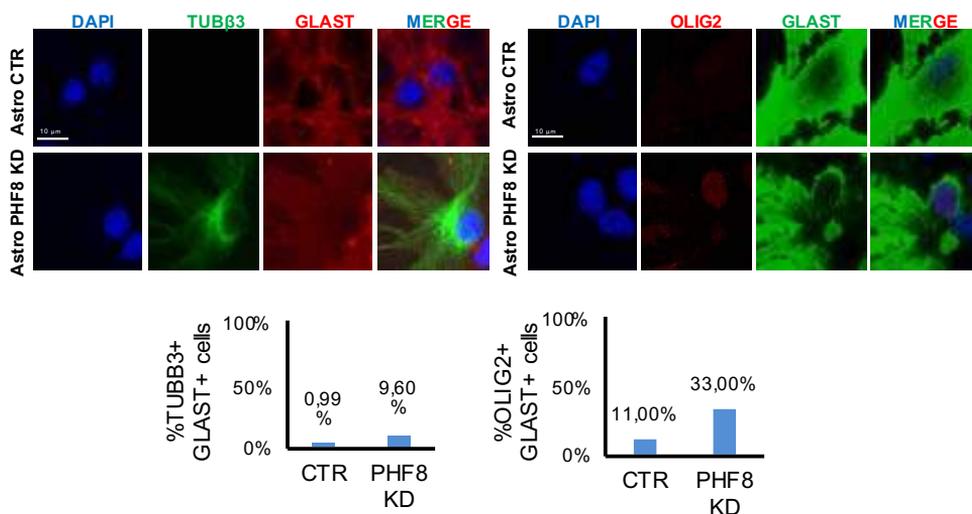


Figure R42: Control and PHF8-depleted NSCs were differentiated to astrocytes during 6 days to generate Astro CTR and Astro PHF8 KD respectively. Cells were fixed and stained with GLAST and TUBβ3 or GLAST and OLIG2. The % of cells expressing the markers is shown at the bottom of the figure.

Then we investigated astrocytes PHF8-KD capacity to resume proliferation and maintain the stem state upon differentiation. To do that, we compared the transcriptional profile of NSCs with astrocytes CTR and PHF8-KD. Figure R45 shows that Astro PHF8-KD did not express proliferation nor stemness related genes. Moreover, no signal for the progenitor marker NESTIN was observed in PHF8 depleted astrocytes by immunostaining assays (Figure R43).

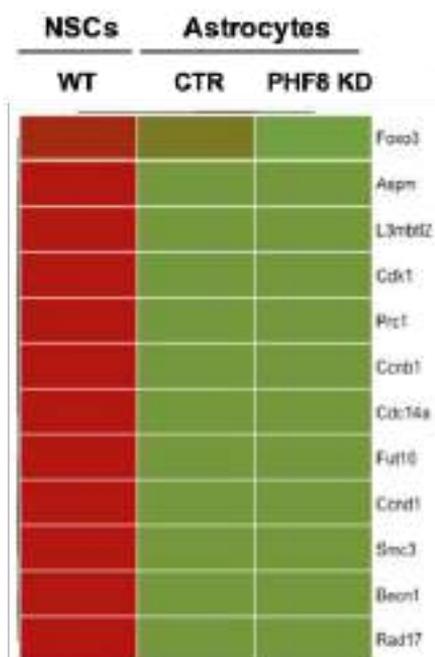


Figure R43: Heat maps showing cell cycle and stemness related gene expression identified by RNA-seq in PHF8 KD and CTR astrocytes compared to NSCs (GSE88173).

Finally, the ability of Astro PHF8 KD to proliferate was analysed too; data in Figure R44 reveal that PHF8 KD astrocytes lost the ability to enter the cell cycle as CTR under the differentiation conditions used in the study.

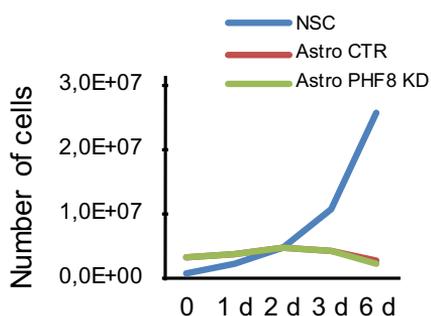


Figure R44: Growth curve showing the proliferation rate of NSCs growing in expansion medium and Astro CTR or Astro PHF8 KD growing in astrocyte differentiation medium during 6 days.

Next, we investigated the contribution of *Phf8* to astrocytic fate, by establishing a neural stem cell line that overexpressed PHF8 in an inducible manner and cultured the cells in a medium without growth factors (Figure R45). In this differentiation media NSCs are able to differentiate into neurons, astrocytes and oligodendrocytes [228].

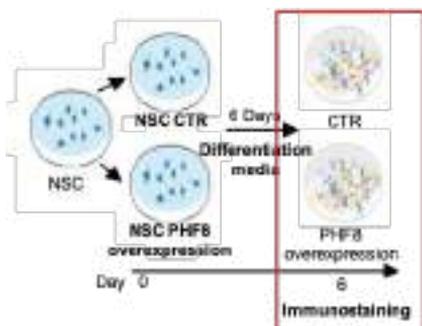


Figure R45: Picture representing control and PHF8-overexpressing NSCs that were differentiated in a medium without growth factor during 6 days before performing the immunostaining assay.

After 6 days in the medium without growth factors, the cells were fixed and the percentage of cells expressing astrocytic markers were determined by immunostaining assay. When *Phf8* overexpression was induced we observed a clear increase in both, the percentage of GFAP expressing cells and the relative protein levels compared to CTR condition, which has basal *Phf8* level (Figure R46).

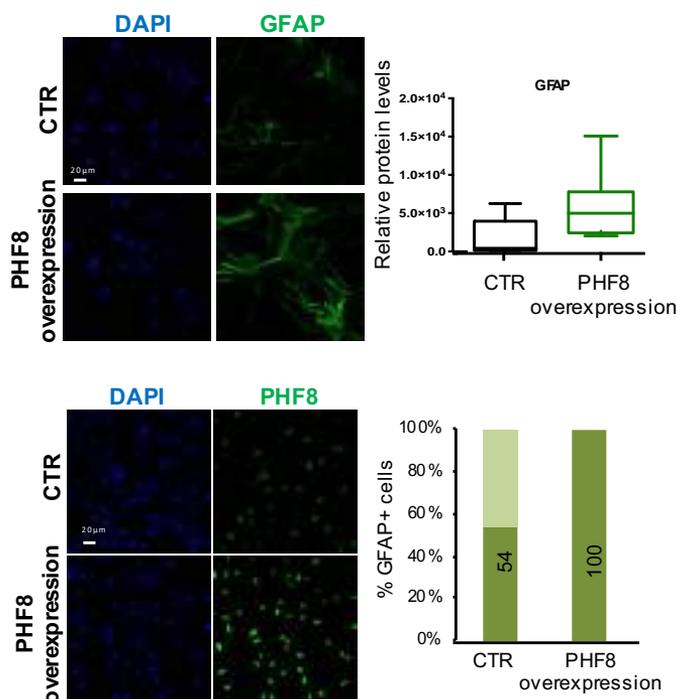


Figure R46: Control and PHF8-overexpressing NSCs were maintained in medium without growing factors for 6 days. The % of GFAP-expressing cells and the relative level of GFAP were determined by immunostaining assays.

On the contrary, when *Phf8* overexpression was induced in the medium without growth factors, we could not appreciate any changes in the expression of OLIG2 compared to CTR cells (Figure R47).

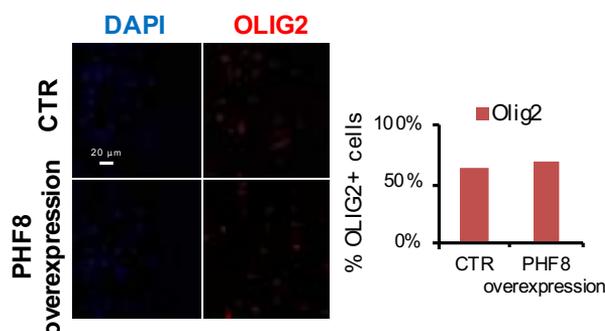


Figure R47: Control and PHF8-overexpressing NSCs were maintained in medium without growing factors during 6 days. The % of OLIG2 expressing cells is depicted on the right.

In sum, these results indicate that PHF8 promotes neural stem cells differentiation towards astrocytes and that PHF8-depleted NSCs differentiate into defective astrocytes.

1.5. *PHF8* depletion impairs neuronal synapses

As PHF8 depletion led to profound defects in synaptogenic gene expression, we next investigated the function of astrocytic PHF8 in synapse formation. To this purpose, we dissected primary hippocampal neurons from E18 mice (in collaboration with Claudia Verderio's lab) and cultured them on PHF8-depleted or control astrocytes to evaluate the density and function of excitatory synapses. In order to do that, control or PHF8-depleted NSCs were first differentiated to astrocytes during 10 days. Then, primary neurons were plated on the differentiated astrocytes and maintained in co-culture for 14 days, to finally analyse the density of excitatory synapses and measure basal synaptic transmission (Figure R48).

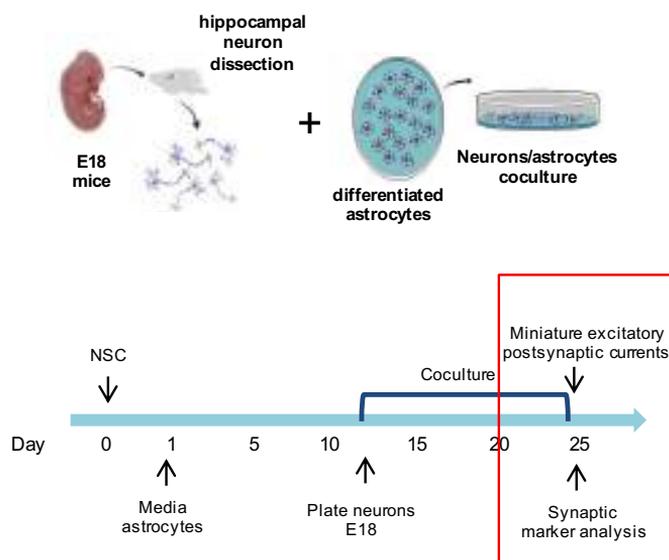


Figure R48: Schematic representation of the neurons/astrocytes co-cultures experiments.

During this process, the maturation of neurons was analysed by immunofluorescence assays in the co-cultures using TUB β III antibody. As expected, we observed a progressive enlargement of the neuron cell bodies and maturation of the dendritic tree over time, at 7, 11 and 14 days *in vitro* (DIV) (Figure R49).

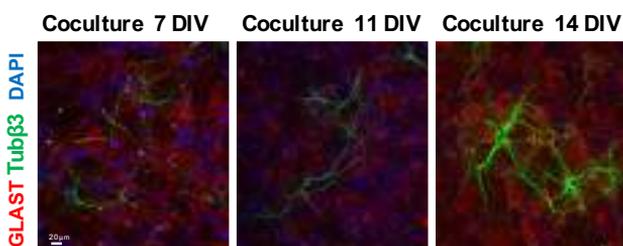


Figure R49: Immunostaining showing the GLAST, TUB β 3 and DAPI staining in co-cultures at 7, 11 and 14 DIV.

To prove the involvement of astrocytic PHF8 in synapse formation, we analysed the density of excitatory synapses. Immunostaining for the presynaptic active zone marker BASSOON and the postsynaptic density marker SHANK2 along the dendrites showed a significant decrease in the density of both pre- and post-synaptic puncta as well as of juxtaposed

pre- and post-synaptic terminals in neurons co-cultured with PHF8 KD astrocytes compared to neurons co-cultured with control astrocytes (Figure R50).

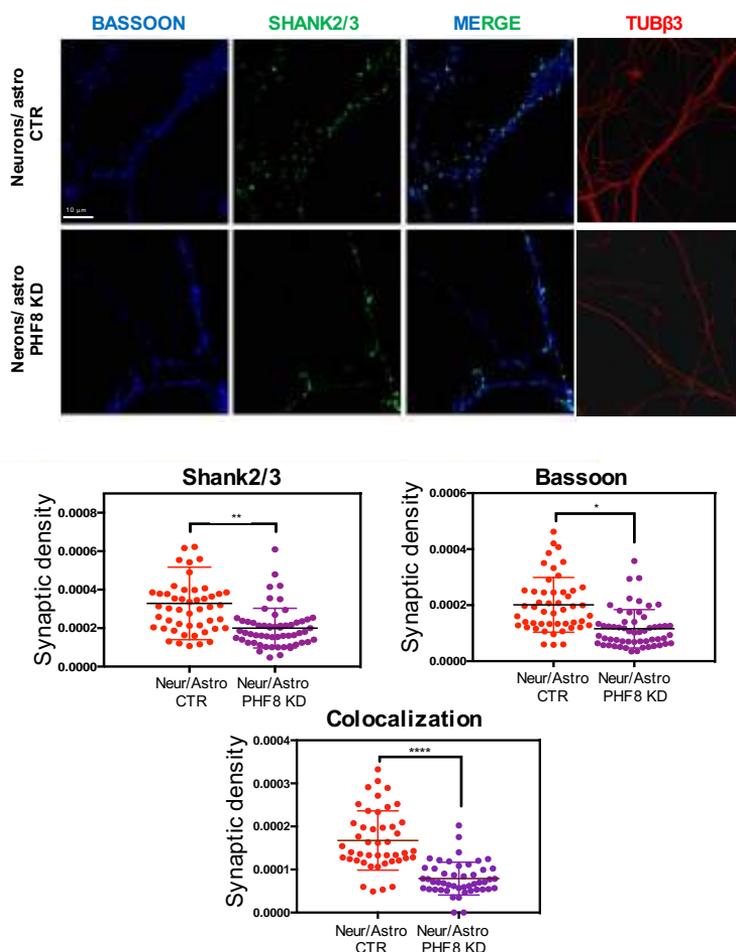


Figure R50: Immunostaining showing BASSOON and SHANK2/3 levels in neuron cultured on astrocytes CTR or PHF8 KD. The synaptic density was calculated dividing the number of puncta by the dendrite length. Quantification is at the bottom.

As further control, we examined synaptic density in neurons cultured in the absence of supporting astrocytes. The analysis showed a similar decrease in the density of BASSOON positive presynaptic terminals in purified neurons and neurons co-cultured with PHF8 depleted astrocytes compared to neurons co-cultured with control astrocytes. The same held

true for SHANK2 positive postsynaptic terminals and BASSON/SHANK2 colocalizing puncta (Figure R51).

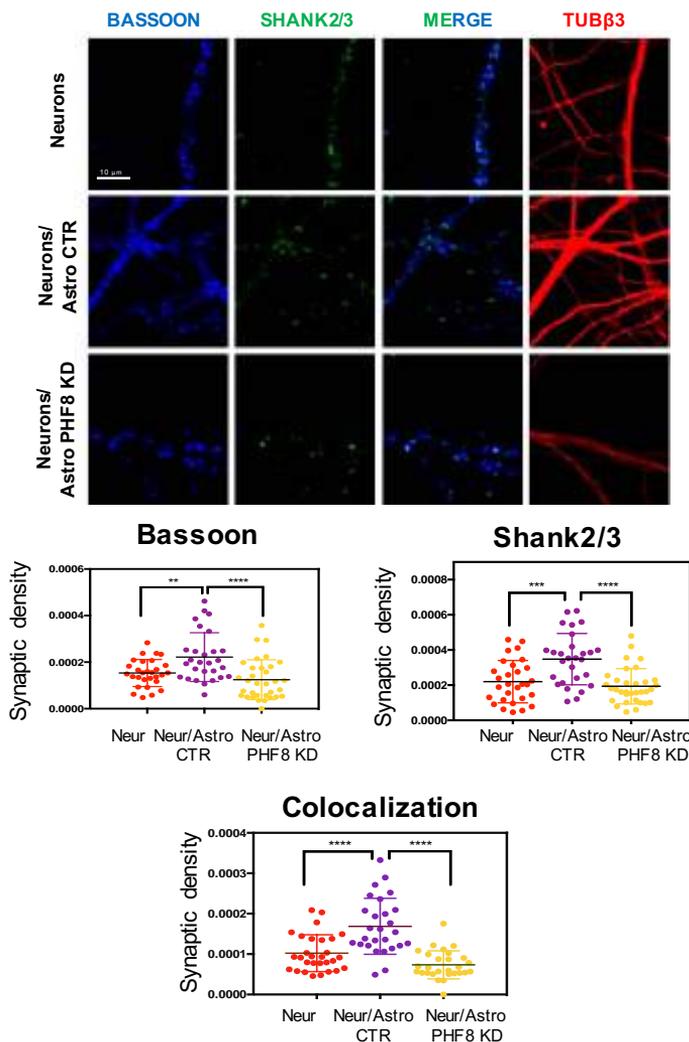


Figure R51: Immunostaining assay showing BASSON and SHANK2/3 levels in neuron cultured without astrocytes, on Astro CTR or Astro PHF8 KD. The synaptic density was calculated dividing the number of puncta by the dendrite length. Quantification is shown in the bottom of the Figure.

Then, we checked the maintenance of astrocyte identity after 25 days in culture by immunostaining assays using AQP4, GLAST, TUB β III and OLIG2 markers (Figure R52). Results show that astrocytes CTR and PHF8 KD maintain their characteristic, as observed in previous experiments (Figure R38 and R40).

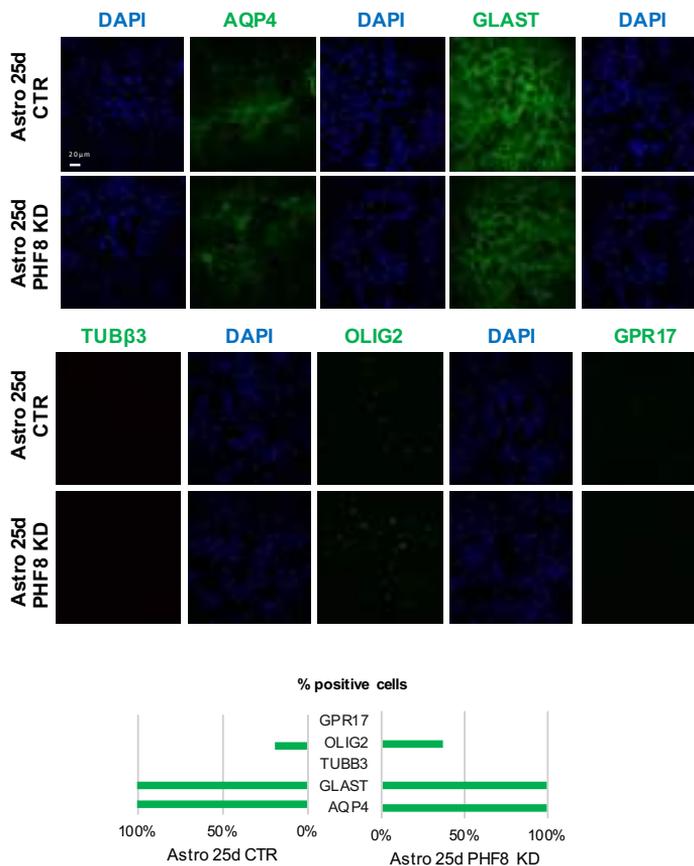


Figure R52: Immunostaining assay showing the AQP4, GLAST, TUBβ3, OLIG2, GPR17 and DAPI signals in CTR and PHF8 KD astrocytes maintained 25 days in culture. The images are representative of two biological independent experiments. Quantification showing the % of cells expressing the indicated markers is depicted on the bottom.

To check neuron's state, we analysed the major and minor cell body diameters of neurons in cocultures using immunofluorescence confocal images of TUBβ3 and DAPI. We found no significant differences in body diameters of neurons cultured on CTR or PHF8 KD astrocytes (Figure R53).

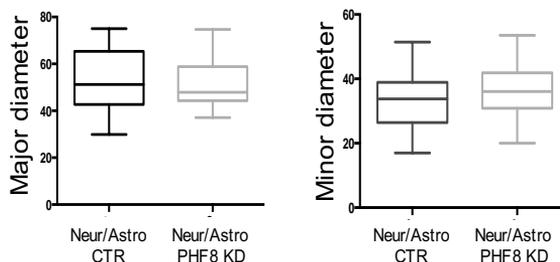


Figure R53: Analysis of cell body diameter (major, left and minor, right) extrapolated from immunofluorescence confocal for TUB β 3 and DAPI.

To functionally evaluate the impact of astrocytic PHF8 depletion on synaptic transmission, we measured miniature excitatory postsynaptic currents (mEPSCs) through whole-cell patch-clamp electrophysiological recordings on 14-day-old neurons co-cultured with control or PHF8-depleted astrocytes. mEPSCs analysis revealed a significant decrease in both the frequency and the amplitude of the miniature excitatory events upon astrocytic PHF8 depletion. The result indicates a clear reduction in the strength of synaptic transmission in neurons cultured on PHF8 KD astrocytes (Figure R54).

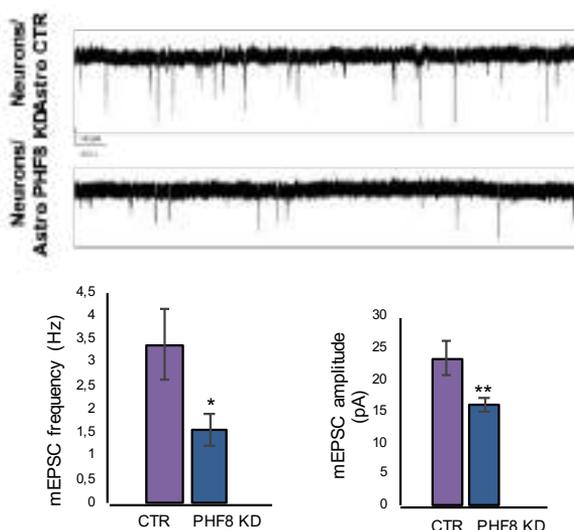


Figure R54: Representative traces of mEPSCs from neuron cultured on Astro CTR or PHF8 KD and histogram showing the mean frequency ($P= 0.042$ Mann-Whitney Rank Sum Test) and amplitudes of mEPSCs ($P= 0.009$ Mann-Whitney Rank Sum Test).

Altogether, the above data demonstrate that astrocytic PHF8 deficiency induces profound alterations in the formation and function of excitatory synapses *in vitro*.

1.6. PHF8 maintains low levels of H4K20me1/3 at astrogenic and synaptogenic genes

Previous studies have shown that H4K20me1 is the main substrate of PHF8 demethylating activity [82]. Thus, we evaluated the impact of PHF8 depletion on H4K20me1 global levels in astrocyte by immunofluorescence analysis. We found a slight increase in H4K20me1 levels in astrocytes PHF8 KD compared to CTR (Figure R55), as previously demonstrated in other cellular contexts [82].

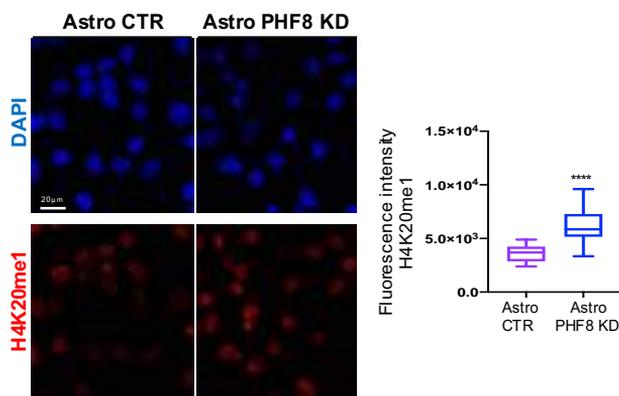


Figure R55: Astro CTR and Astro PHF8 KD were immunostained using H4K20me1 antibody. Quantification of the fluorescence intensity is shown on the right side.

As H4K20me1 mark is a substrate for the histone methyltransferase SUV20H1, we also tested the levels of H4K20me3 and found a clear increase in H4K20me3 heterochromatic mark in astrocytes PHF8 KD. Interestingly, both the intensity and the number H4K20me3 foci raised upon PHF8 depletion (Figure R56).

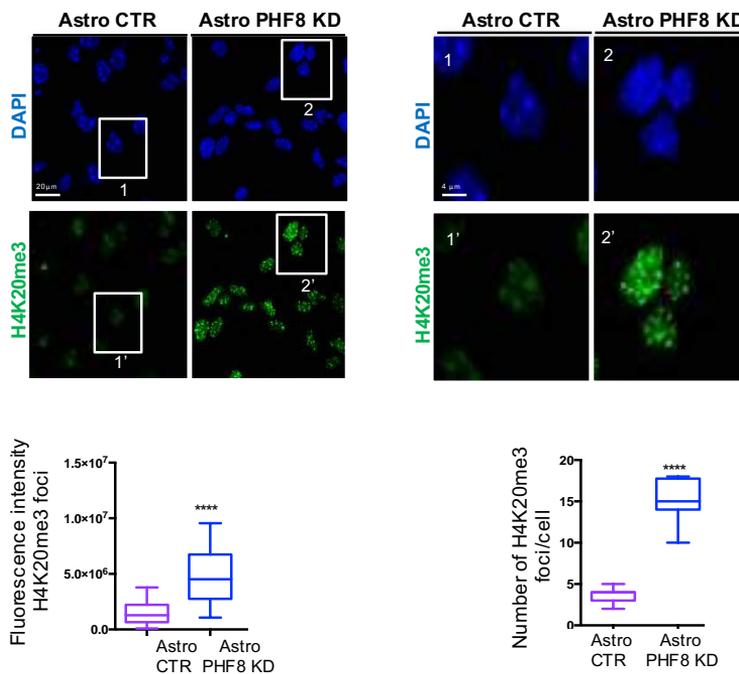


Figure R56: Astro CTR and Astro PHF8 KD were immunostained using H4K20me3 antibody. Quantification of the fluorescence intensity and number of H4K20me3 foci/cell are shown is shown on the bottom of the Figure.

We detected no changes in other histone marks associated to transcriptional activation such as H3K4me3 (Figure R57).

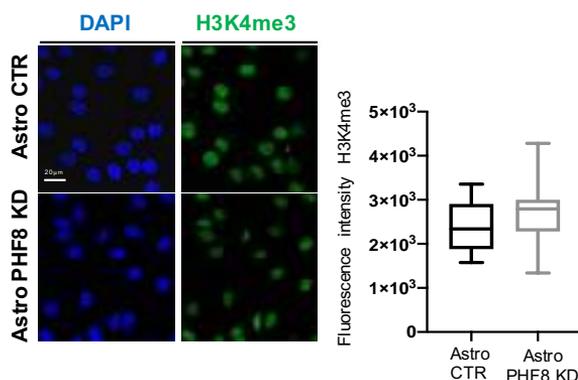


Figure R57: Astro CTR and Astro PHF8 KD were stained using H3K4me3 antibody. Quantification of the fluorescence intensity is shown is shown on the right side of the Figure.

Next, we investigated whether PHF8 is important to prevent the accumulation of H4K20me1 at PHF8-regulated genes, defined as direct

targets, during astrocyte differentiation. To do it, we chose two PHF8-target regions identified in the ChIP-seq experiment and essential for astrogenesis (*Nfia*) and synaptogenesis (*Sparc*) and tested the effect of PHF8-depletion on H4K20me1 levels both in NSCs and control or PHF8-depleted astrocytes by ChIP-qPCR assays. We noticed a clear decrease in H4K20me1 mark upon astrocyte differentiation (comparing NSCs and Astro CTR); no significant changes at the *Olig2* promoter (a non PHF8-target used as a negative control) were observed. Interestingly, H4K20me1 decrease did not take place in PHF8-depleted astrocytes, meaning that PHF8 catalytic activity is involved in their regulation (Figure R58).

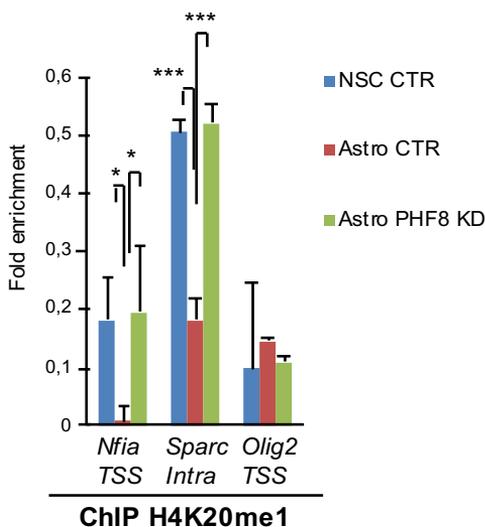


Figure R58: The levels of H4K20me1 in NSCs, Astro CTR, and Astro PHF8 KD were determined by ChIP-qPCR at the indicated genes. *Olig2* TSS region was used as negative control. “Intra” refers to intragenic region identified in the PHF8 ChIP-seq experiment.

The decrease in H4K20me1 mark noticed upon astrocyte differentiation (Figure R6) well correlated with gene activation (Figure R59), both comparing NSCs with Astro CTR, and Astro CTR with Astro PHF8 KD confirming that PHF8-depleted astrocytes do not properly express the key transcription factor *Nfia* and the synapse-associated astrocytic protein *Sparc*. *Olig2* was used as a negative control as it is not expressed in astrocytes.

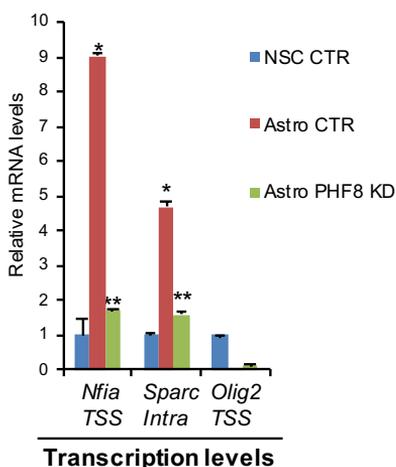


Figure R59: Expression levels of the indicated genes in NSCs, astro CTR, and astro PHF8 KD were determined by qPCR. Values were normalized to the housekeeping gene *Gapdh*, and Figure shows values relative to NSC CTR. *Olig2* mRNA was used as negative controls.

We also analysed the consequences of PHF8 depletion on H4K20me3 levels, as we observed an increase of this mark at global levels. As depicted in Figure R60, a clear increase in H4K20me3 was observed in PHF8 KD cells in *Nfia* and *Sparc* genes in CHIP-qPCR experiments.

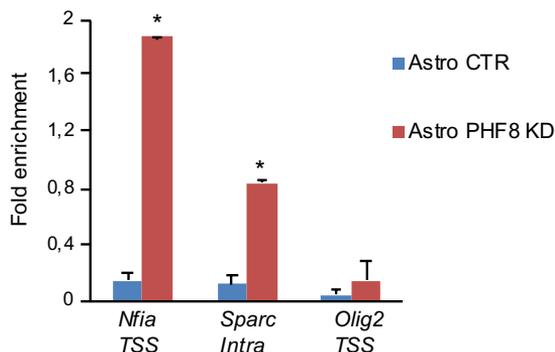


Figure R60: The levels of H4K20me3 in Astro CTR, and Astro PHF8 KD were determined by CHIP-qPCR at the indicated genes. *Olig2* TSS region was used as negative control. Data from qPCR were normalized to the input and expressed as fold enrichment over the data obtained in shCTR. “Intra” refers to intragenic region identified in the PHF8 CHIP-seq experiment.

Given that PHF8 also demethylases H3K9me2 histone mark [68, 93] we checked if PHF8 acts on H3K9me2 in the regions identified by PHF8 CHIP-seq. Figure R61 shows that PHF8 is not responsible for keeping low

levels of H3K9me2 at the analysed genes in astrocytes, as no increase in H3K9me2 was observed upon PHF8 depletion.

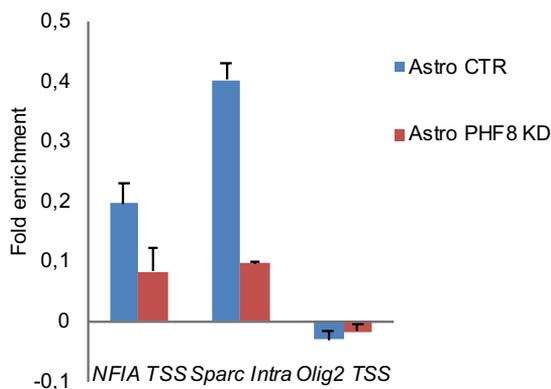


Figure R61: The levels of H3K9me2 in Astro CTR, and Astro PHF8 KD were determined by CHIP-qPCR at the indicated genes. Olig2 TSS region was used as negative control. Data from qPCR were normalized to the input and expressed as fold enrichment over the data obtained in CTR.

1.7. *PHF8 histone demethylase activity preserves astrocytic transcriptional program*

Given that changes on H4K20me1/3 levels correlated with transcriptional changes of genes involved in astrocytic differentiation and synapses, we assessed the role of PHF8 HDM activity on the observed phenotypes. In order to do that, we established PHF8 KD NSCs cell lines that overexpressed either PHF8 WT or PHF8 mutant lacking HDM activity (mutant H247A) (Figure R62).

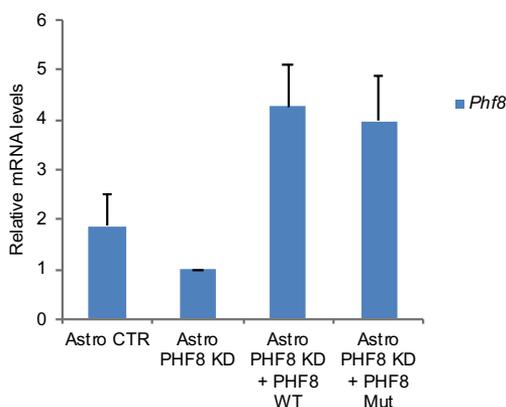


Figure R62: PHF8 WT or mutant (H247A) were expressed in PHF8 KD Astro; the level of *Phf8* was determined by qPCR and compared to the level in PHF8 KD astrocytes. Expression values were normalized to the housekeeping gene *Gapdh*.

We rescued the defects on astrocyte differentiation by analyzing GFAP protein in differentiated cells. We found that GFAP expression is actually recovered upon PHF8 WT overexpression but not upon the catalytic mutant overexpression (Figure R63).

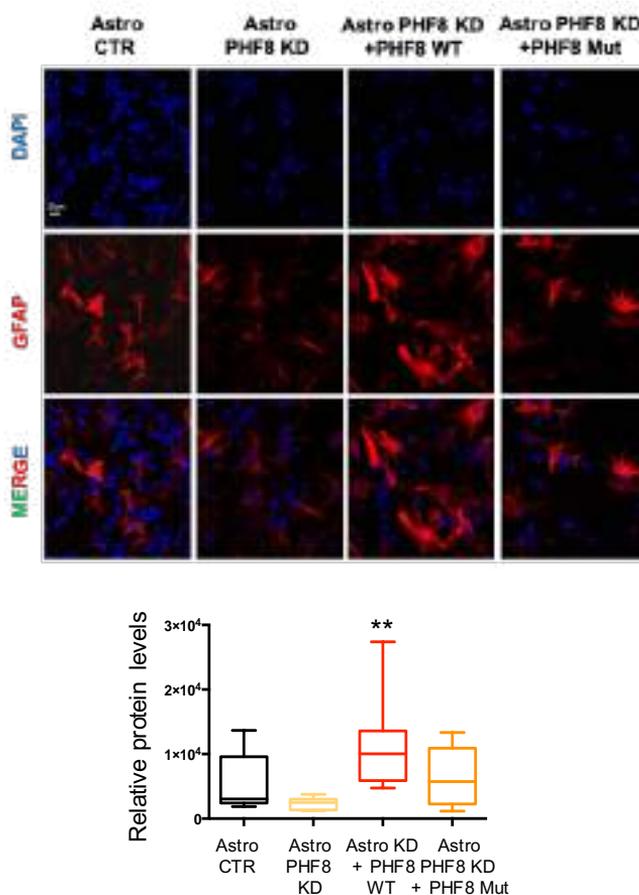


Figure R63: Control and PHF8-depleted NSCs were differentiated to astrocytes during 6 days, then the expression of PHF8 WT or the catalytic mutant (H247A) (Mut) in PHF8 KD Astro were induced by doxycycline addition and differentiated during 6 days. The cells were immunostained using GFAP antibody and DAPI.

Next, we tested the role of the catalytic activity in the control of synaptic and astrocytic gene expression. qPCR experiment demonstrated that

PHF8 WT, but not the mutant, was able to rescue the expression levels of the tested astrocytic and synaptogenic genes regulated by PHF8 during differentiation (Figure R64).

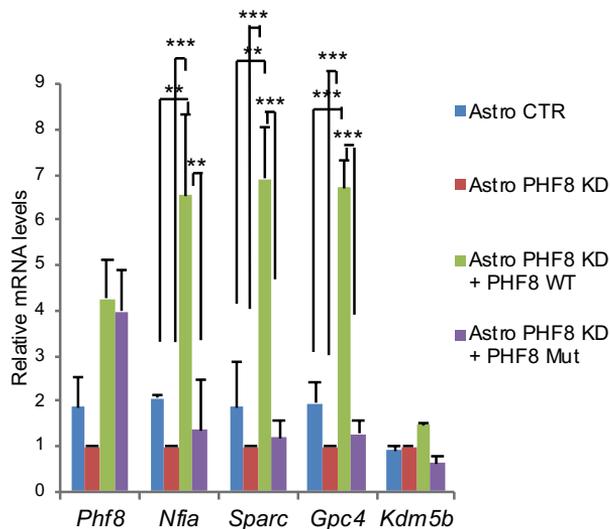


Figure R64: PHF8 WT and mutated at the catalytic domain (H247A) were expressed in Astro PHF8 KD; then, the expression levels of the indicated genes were determined by qPCR. Expression values were normalized to the housekeeping gene *Gapdh*, and Figure shows values relative to Astro PHF8 KD. *Kdm5b* mRNA was used as negative control.

Finally, the importance of PHF8 HDM activity was demonstrated by the rescue of H4K20me3 levels after overexpression of PHF8 WT, but not the catalytic mutant (Figure R65). Notably, in the case of overexpression of the PHF8 catalytic mutant an apparent increase in the H4K20me3 intensity was observed, enlightening the importance of PHF8's enzymatic activity in preventing heterochromatin marks accumulation during astrocyte differentiation.

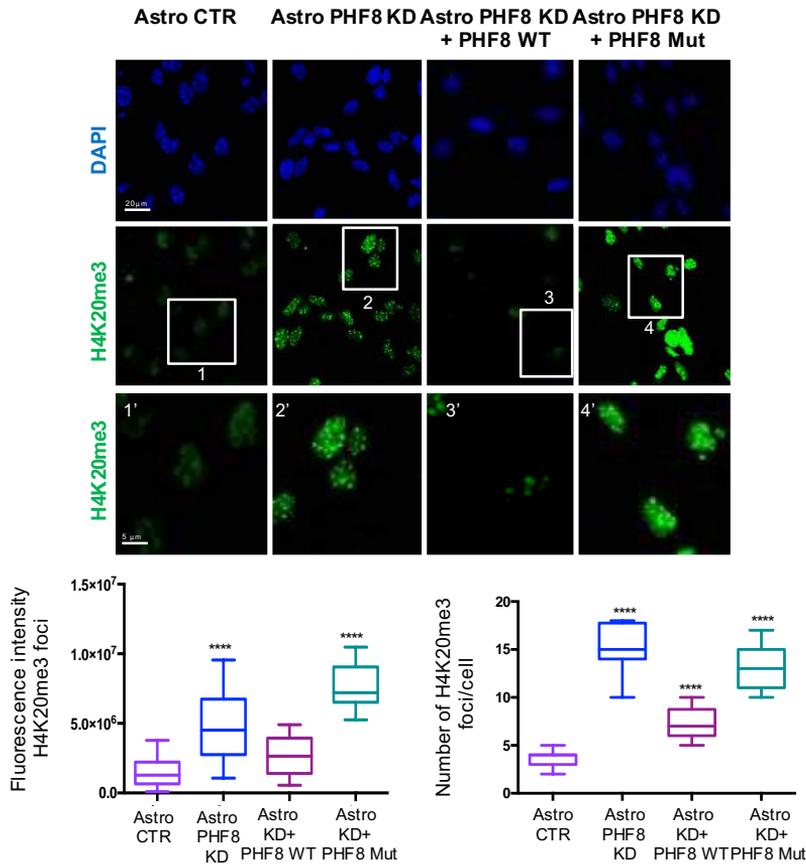


Figure R65: PHF8 WT and the catalytic mutant were expressed in Astro PHF8 KD; then, the levels of H4K20me3 were established by immunostaining assays. Zoom in (scale bar 5 μm) showing H4K20me3 foci. The images are representative of three biological independent experiments. Fluorescence intensity of H4K20me3 and H4K20me3 foci number were determined using ImageJ. Boxplots represent the quantification of the fluorescence intensities as well as the number of H4K20me3 foci/cell in each population.

Taken together, these data indicate that the major role of PHF8 in astrocytes is to demethylate H4K20me1, preventing subsequent ectopic heterochromatin formation, at those genes essential in the differentiation process and during synaptogenesis.

2. Characterization of PHF8 function in neural stem cells

2.1. PHF8 depleted neural stem cells suffer delay in cell cycle

To evaluate the impact of PHF8 depletion on NSCs, cells were transduced with lentivirus containing the control shRNA (CTR) or the specific PHF8 shRNA that efficiently decreased PHF8 levels. We confirmed that NSCs PHF8 KD line efficiently decreases PHF8 protein levels (Figure R12). To analyse the role of PHF8 in NSCs, we first examined the consequences of its depletion in cell proliferation. PHF8 KD NSCs exhibited a striking decrease in cell growth compared to CTR cells (Figure R66).

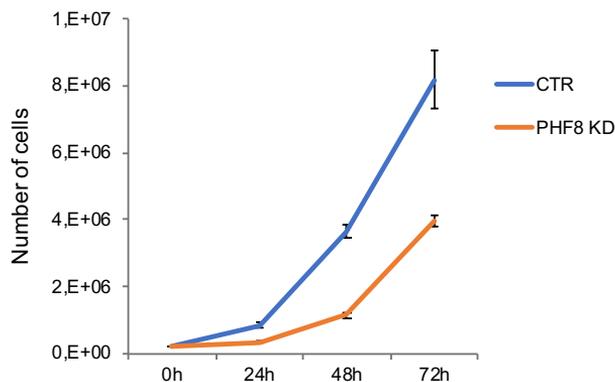


Figure R66: Growth curve showing the proliferation rate of NSCs infected with lentivirus expressing shRNA control (CTR) or shRNA for PHF8 (PHF8 KD) from 0 to 72h.

Moreover, flow-cytometry analysis (FACs) demonstrated a delay in G1/S transition (G1 CTR 39,9%, PHF8 KD 48,7%), and a slight delay in G2/M (G2 CTR 8,22%, PHF8 KD 12%) upon PHF8 depletion (Figure R67). IdU incorporation analysis coupled with FACs showed that PHF8 KD cells has a shorter S phase, particularly mid S phase than control NSCs (Figure R67).

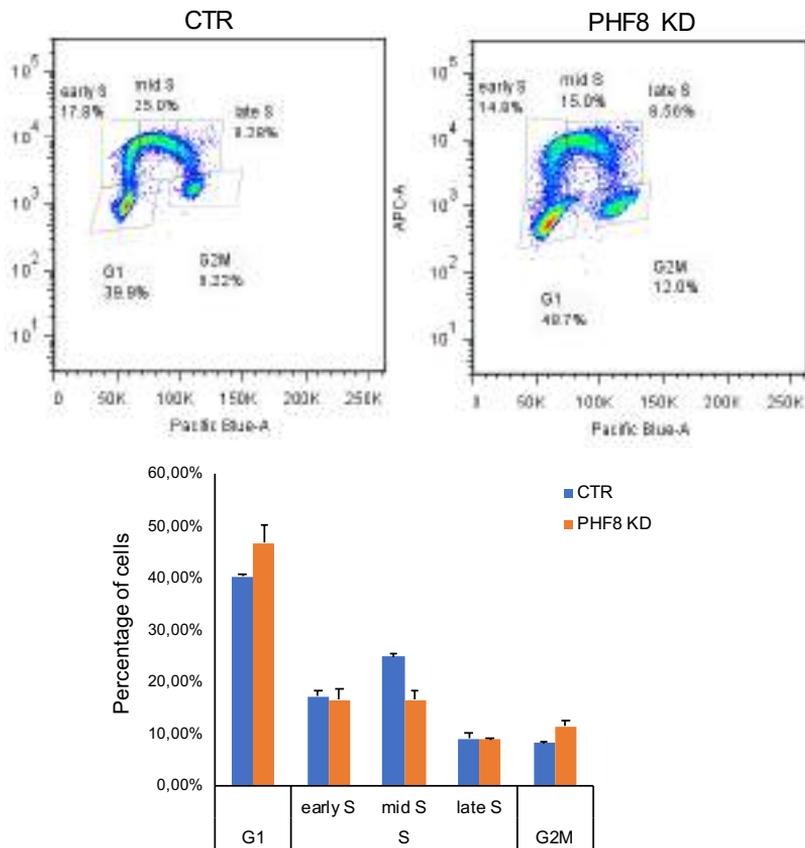
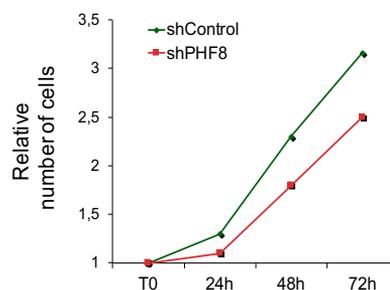


Figure R67: Flow-cytometry analysis of NSCs CTR and PHF8-depleted cells previously stained with 5-Ethynyl-2'-deoxyuridine (EdU). Graph showing quantification is on the bottom of the Figure.

The phenotype observed in the neural stem cells PHF8 KD largely resembles the one reported in HeLa PHF8 KD from previous work in our laboratory (Figure R68) [229] and others [82, 83].



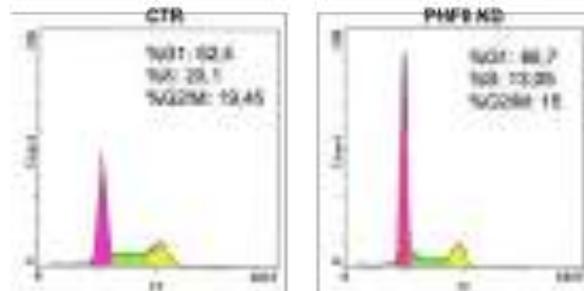
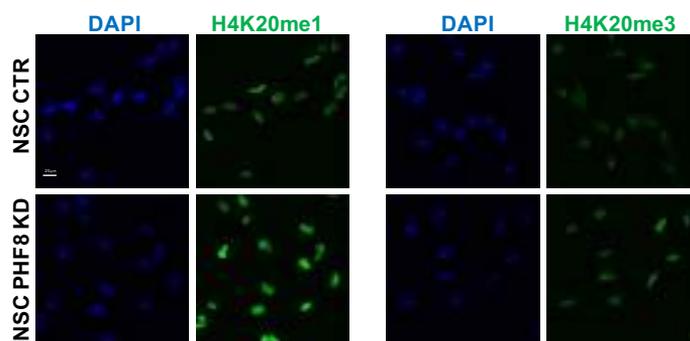


Figure R68: Growth curve showing the proliferation rate of HeLa infected with lentivirus expressing shRNA control (CTR) or shRNA for PHF8 (PHF8 KD) from 0 to 72h (top panel). Flow-cytometry analysis of HeLa CTR and PHF8-depleted cells previously stained with bromodeoxyuridine (BrdU) (bottom panel). Adapted from [229].

To better characterize NSCs phenotype upon PHF8 depletion we assessed the changes in the histone methylation status performing immunostaining assays of H4K20me1, H4K20me3, H3K9me2 and H3K9me3 histone modifications. It has been demonstrated that PHF8 demethylases mainly H4K20me1 and H3K9me2, however we also checked the levels of H4K20me3 and H3K9me3, as mono- and di- serves as substrates for trimethylation. We observed that PHF8 KD NSCs show increased levels of both H4K20me1 and 3, even if the highest increment is in the H4K20me1. H3K9me2 slightly decreased in PHF8 KD cells, instead H3K9me3 levels were higher in PHF8 KD compared to CTR NSCs (Figure R69).



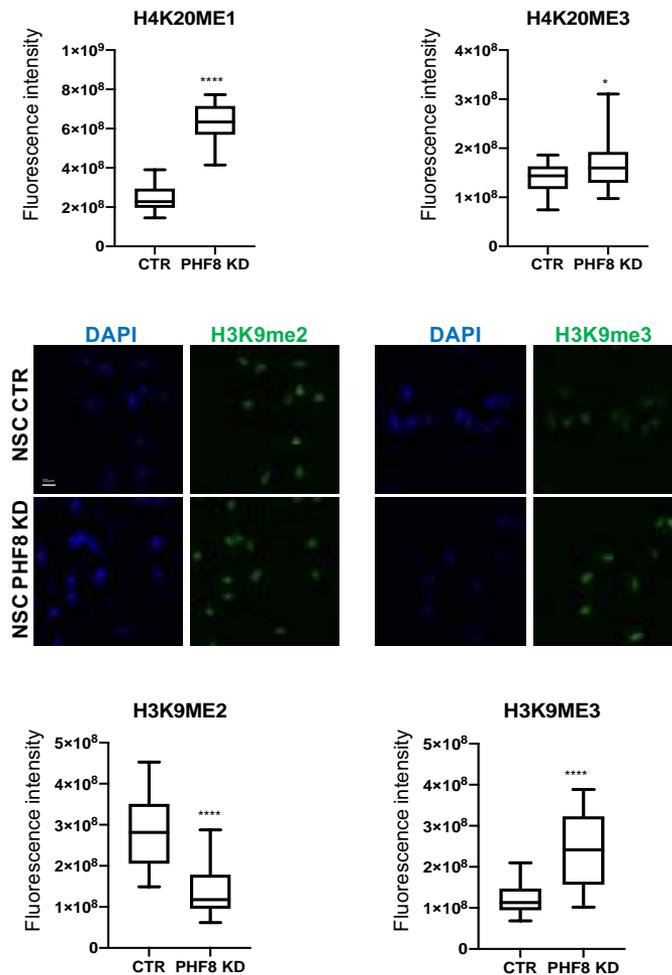


Figure R69: NSCs CTR and NSCs PHF8 KD were immunostained using H4K20me1, H4K20me3, H3K9me2 and H3K9me3 antibodies. Quantification of the fluorescence intensity is shown at the bottom of the Figure.

As we observed an increase in H3K9me3 levels in PHF8 KD NSCs, we decided to perform immunofluorescence assay of heterochromatin binding protein α (HP1 α), which is known to interact with H3K9me3. The Figure R70 shows that in PHF8 depleted cells there is accumulation of HP1 α compared to CTR NSCs.

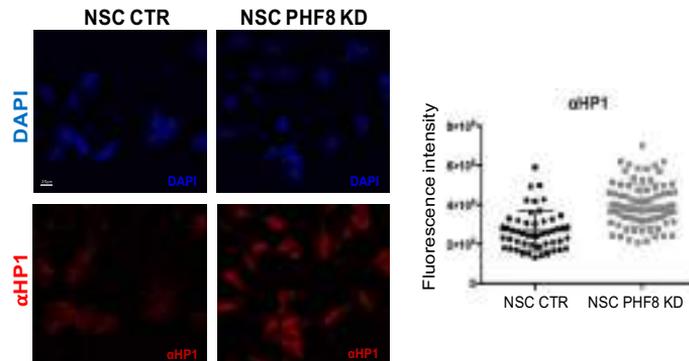


Figure R70: NSCs CTR and PHF8 KD were immunostained using HP1 α antibody. Quantification of the fluorescence intensity is shown at the right of the Figure.

Moreover, we noticed an increased frequency of chromosome segregation defects, in particular, we could appreciate anaphase chromatin bridges and multinuclear cells in PHF8-depleted cells compared to control NSCs (Figure R71).

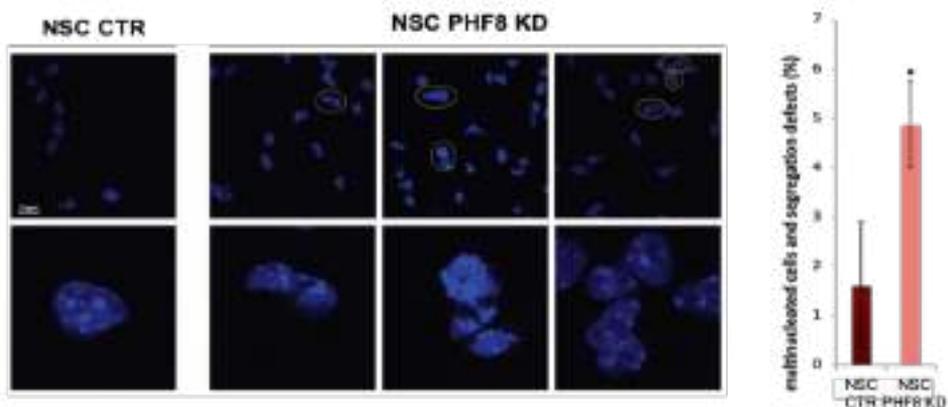


Figure R71: Formation of multi-nucleated cells and segregation defects are shown for NSC CTR and PHF8 KD. Graph on the right part of the Figure represents the percentage of multinucleated cells.

Those defects have been found also in other cell lines upon PHF8 depletion [83], indicating that PHF8 is important to maintain genome stability.

2.2. Analysis of PHF8 depleted NSC transcriptional profile

To deeply understand the function of PHF8 in NSCs we determined the PHF8-dependent transcriptional profile by RNA-sequencing. To do that we purified total RNA from control and two PHF8-depleted NSCs samples and performed RNA sequencing of mRNA transcripts in the CRG/CNAG genomics unit. In Figure R72 is shown the clustered heatmap depicting Pearson correlation based on read coverage and the principal component analysis plot of CTR and NSC PHF8 KD samples.

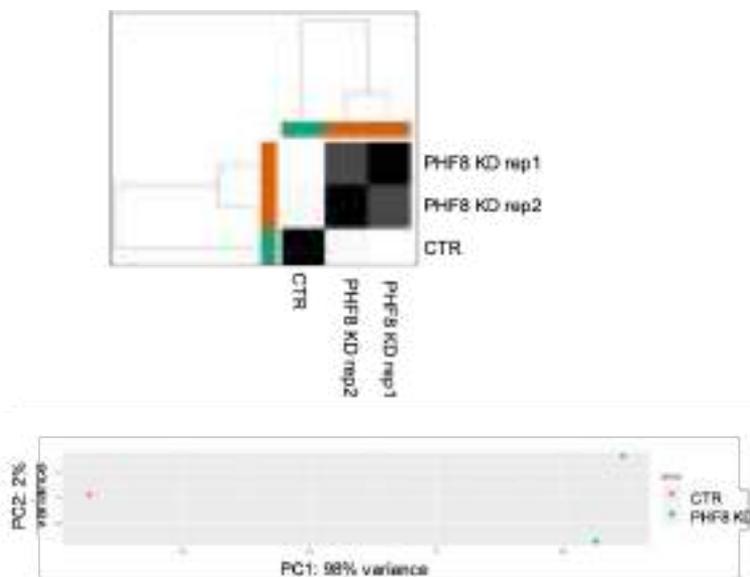


Figure R72: Clustered heatmap depicting Pearson correlation and the principal component analysis of the NSC PHF8 KD and CTR RNA-seq samples based on read coverage within genomic regions.

We examined two biological PHF8 KD NSCs samples compared with NSCs CTR and found 5946 transcripts that significantly changed their expression upon PHF8 depletion [\log_2 fold change (FC) $>0,5$ and (FC) $<-0,5$ and p -value $<0,08$]. (Figure R73).

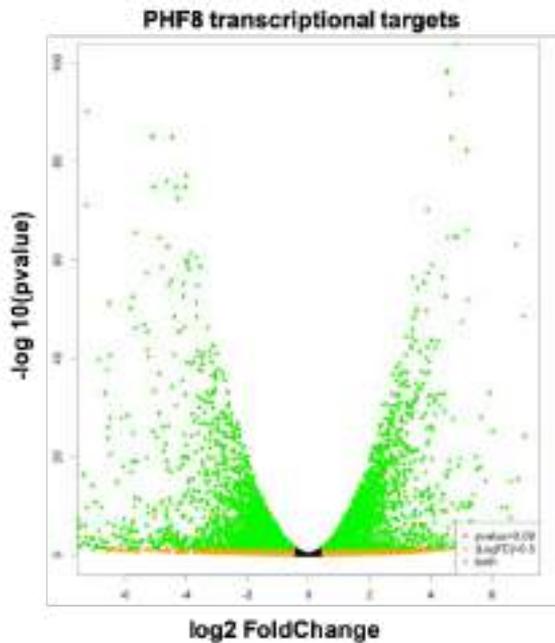
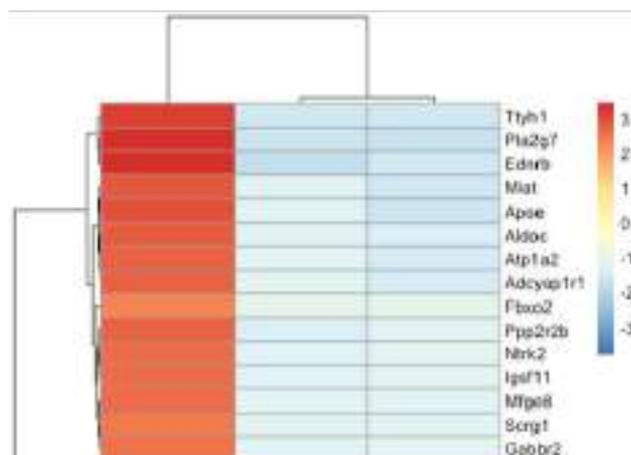


Figure R73: Volcano plot represents PHF8 transcriptional targets identified by RNA-seq in NSCs CTR and NSCs PHF8 KD.

In the Figure R74 is represented the heatmap showing the top 50 regulated genes differentially expressed in NSCs CTR compared to NSCs PHF8 KD and identified by RNA-seq [\log_2 fold change (FC) $>0,5$ and (FC) $<-0,5$ and p-value $<0,08$].



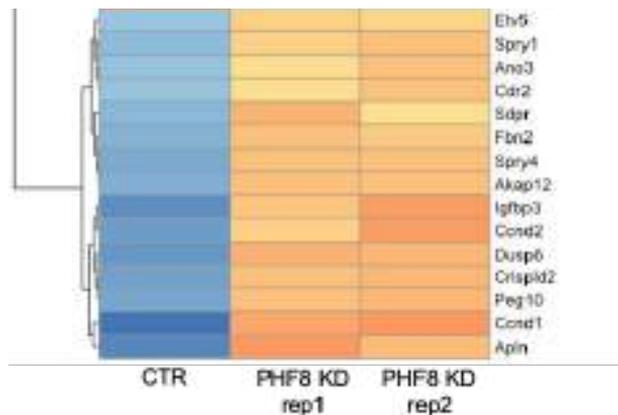


Figure R74: Heatmap showing the top 50 regulated genes identified by RNA-seq in NSC CTR and NSC PHF8 KD. Two biological replicates of shPHF8 cells were used for RNA-seq. All the genes displayed have p -value < 0.08 and \log_2 FoldChange > 0.5 and < -0.5 .

Among the 5946 differentially expressed transcripts, 2947 (49,6%) were downregulated and 2999 (50,4%) were upregulated upon PHF8 depletion (Figure R75, left panel). We observed that when we shifted the \log_2 fold change to 1 and 1,5 (Figure R75, right panel) the percentage of downregulated and upregulated genes did not change significantly.

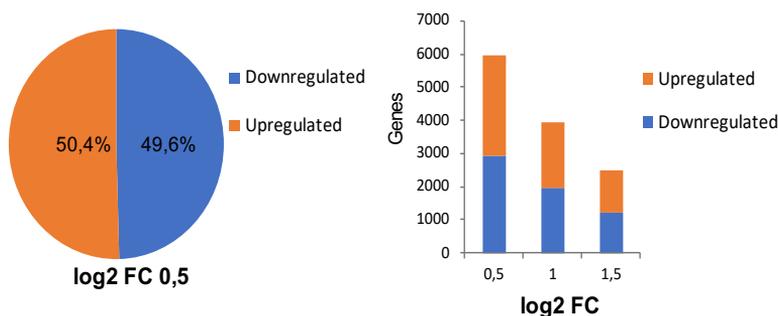


Figure R75: Graph depicting the percentage of upregulated and downregulated genes in the NSC PHF8 KD compared to CTR with p -value $< 0,08$ and classified by increasing \log_2 of fold change (FC).

Actually, data previously obtained in our laboratories confirmed that PHF8 can function both as an activator and as a repressor in HeLa too. The results of the microarray gene expression of CTR and PHF8 KD HeLa

(Figure R76) [94] are consistent with the RNA-seq analysis of CTR and PHF8 KD neural stem cells.

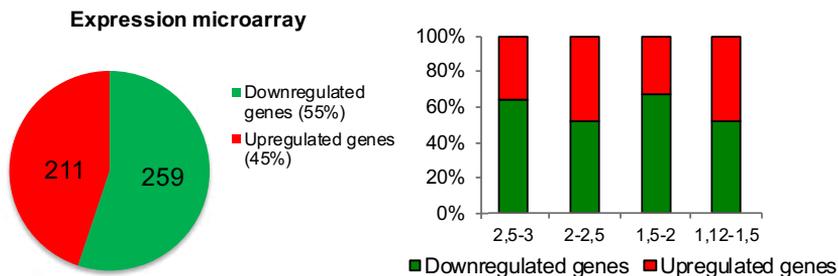


Figure R76: Graph depicting the percentage of upregulated and downregulated genes in the HeLa PHF8 KD compared to CTR with \log_2 fold change >1 and <-1 , and classified by increasing \log_2 fold change. Adapted from [94].

Enrichment analysis of the gene ontology terms over the differentially transcribed genes showed that the most enriched categories were associated especially with DNA replication (Atrx, Ccdc88a, Lig3, Polb, Polg), mitotic sister chromatid segregation (Chmp2a, Chmp2b, Chmp4c), G1/S transition (Cdk4), as well as centromere complex assembly (H3f3a, H3f3b) (Figure R77). Interestingly also appeared some categories related to metabolism: nucleoside monophosphate biosynthetic process, chondroitin sulfate proteoglycan metabolic process and folic acid-containing compound metabolic process.

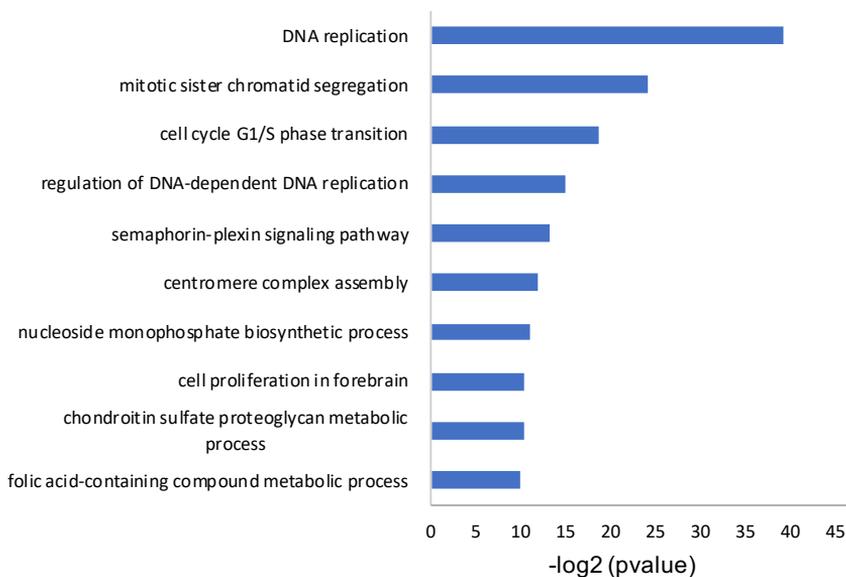


Figure R77: Gene ontology analysis showing the Biological Process of the PHF8 regulated genes ($p\text{-value} < 0.08$ and $\log_2\text{FoldChange} > 0.5$ and < -0.5) was performed using as a background the whole *Mus musculus* genome.

2.3. *PHF8* depletion impairs serine biosynthesis

As we found categories related to metabolism (Figure R77), we classified the differentially transcribed genes involved in different metabolic processes like nucleoside monophosphate biosynthetic compound, amino acids, glycine, pyruvate and serine. The Figure R78 shows that serine metabolism was the most enriched category among the metabolic terms.

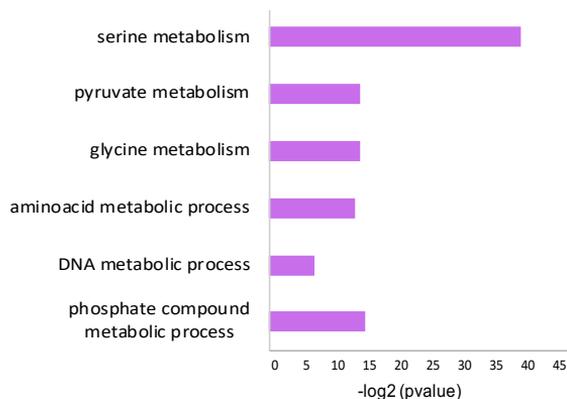


Figure R78: Graph representing the percentage of gene differentially regulated in the RNA-seq of NSC CTR and PHF8 KD and belonging to different GO categories related to metabolic processes.

Thus, we represented (Figure R79) the top regulated genes in the RNA-seq related to serine metabolism (the majority of those genes are downregulated in PHF8 KD NSCs). We noticed that some of those genes are in common with other metabolic pathways too, for example the metabolism of the alpha-ketoglutarate (αKG), or the glutamate.

A closer analysis to the transcripts related to metabolism indicated that also many amino acids transporters were misregulated in PHF8 KD NSCs compared to CTR cells (Figure R80).

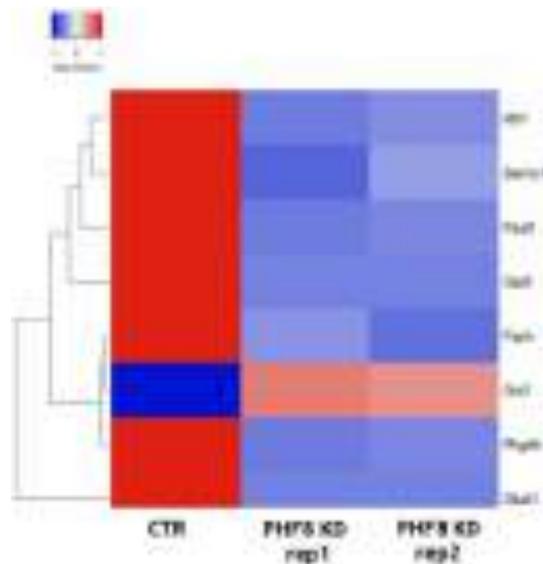


Figure R79: Heatmap showing some serine metabolism related genes identified by RNA-seq experiment in the samples of NSCs PHF8 KD and CTR. All the genes showed p -value <0.08 and \log_2 FoldChange >0.5 and <-0.5 .

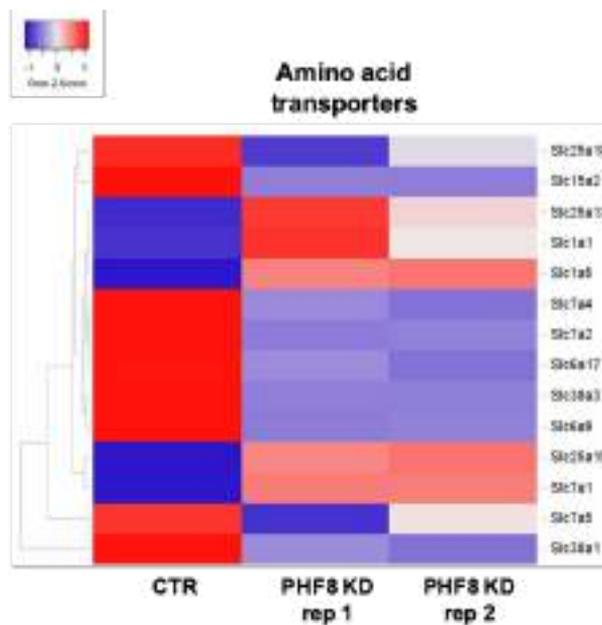


Figure R80: Heatmap showing differentially expressed amino acids transporters identified by RNA-seq experiment in the samples of NSC PHF8 KD and CTR. All the genes showed p -value <0.08 and \log_2 FoldChange >0.5 and <-0.5 .

We confirmed by qPCR that two key genes in the serine synthesis metabolism, PHGDH and PSAT1, were downregulated in NSCs PHF8 KD compared to CTR (Figure R81).

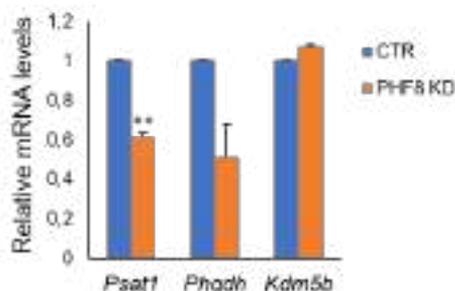


Figure R81: Graph showing PHGDH and PSAT1 mRNA levels in the NSCs CTR and PHF8 KD.

Interestingly, PSAT1 was also downregulated in HeLa PHF8 KD compared with CTR as shown in a previously published microarray from our laboratory (Figure R82) [94]. This result and the similarities observed between PHF8 KD NSC and PHF8 KD HeLa considering both the phenotype (Figure R68) and the transcriptional profile (Figure R76) suggest that PHF8 may have the same function in the two cell lines.



Figure R82: Cluster diagram of genes showing Fold Change of PSAT1 in the microarray. Adapted from [94].

As we observed that the main enzymes related to serine, glutamate and glycine metabolism were downregulated (Figure R79), we decided to actually evaluate the concentration of those amino acids in CTR and PHF8 KD HeLa. In order to do that we performed a gas-chromatography mass spectrometry experiment in the chromatography unit of the Barcelona University (UB); the results in Figure R83 showed a reduction in serine (Ser) and glutamic acid (Glu) concentration in PHF8 KD HeLa compared to CTR cells.

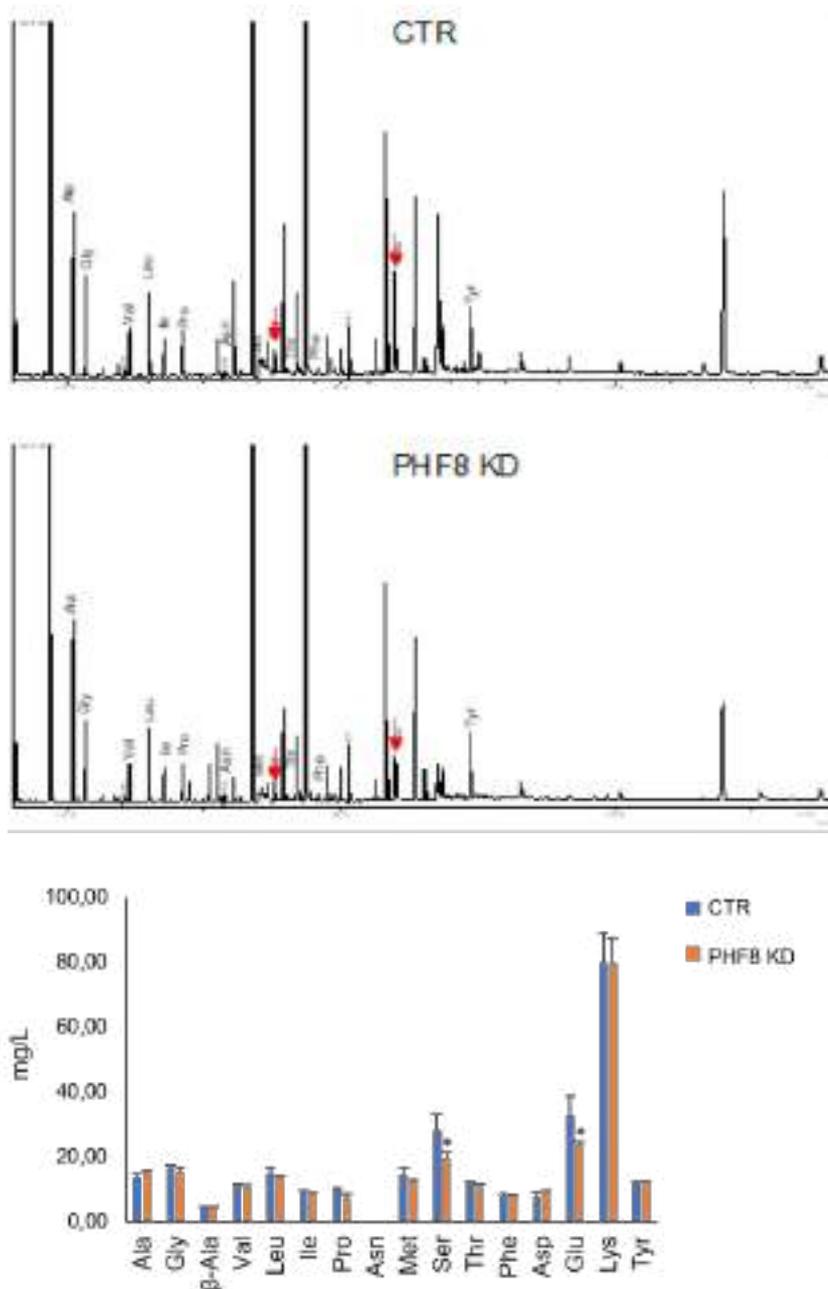


Figure R83: Representative chromatograms of CTR and PHF8 KD HeLa metabolites measured by gas chromatography- mass spectrometry. Quantifications of three CTR and three PHF8 KD replicates are shown at the bottom of the Figure.

Knowing that PHF8 KD cells suffer a depletion of serine, that it is essential for proliferation, we hypothesized that serine addition could rescue PHF8 KD cell growth phenotype. So, we cultured HeLa CTR and PHF8 KD in media depleted of serine and, when we added serine to both of them, only cells PHF8 KD were able to grow faster (Figure R84).

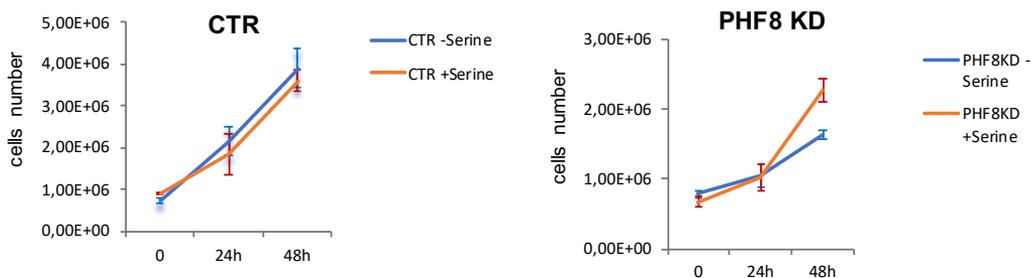


Figure R84: Growth curve showing the proliferation rate of NSCs PHF8 KD grown in media supplemented or not with serine from 0 to 48 h.

We also checked whether glutamine or alpha ketoglutarate (α KG), which can also be regulated by the serine biosynthesis pathway, could rescue PHF8 KD phenotype. Interestingly, when we added glutamine or α KG to HeLa PHF8 KD, cells were not affected indicating that no one of them was able to rescue the PHF8 KD cell cycle delay (Figure R85).

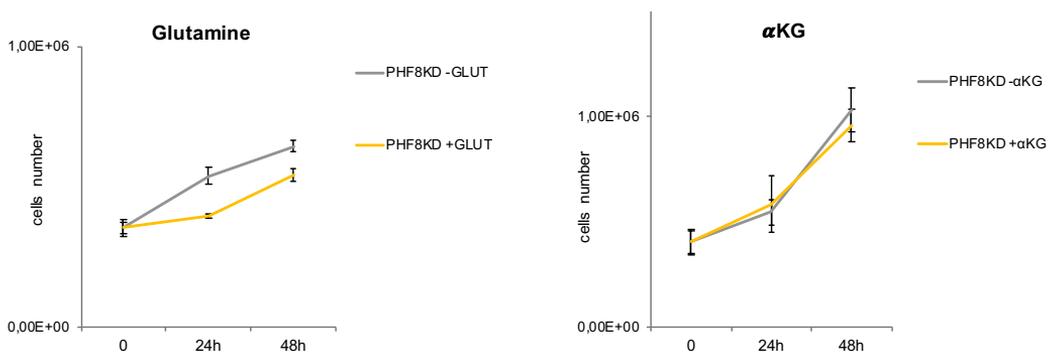


Figure R85: Growth curve showing the proliferation rate of NSCs PHF8 KD grown in media supplemented or not with alpha-ketoglutarate (α KG) and glutamine from 0 to 48h.

The serine synthesis pathway is especially induced in absence of serine, so we tested if the induction of PSAT1 and PHGDH (the first two genes essential in the serine biosynthesis) occurs in CTR and PHF8 KD HeLa

in absence of serine. Results in Figure R86 show that the two genes are induced in CTR cells but not in PHF8 KD HeLa in absence of serine indicating that PHF8 depletion impairs the proper expression of two major enzymes involved in serine synthesis.

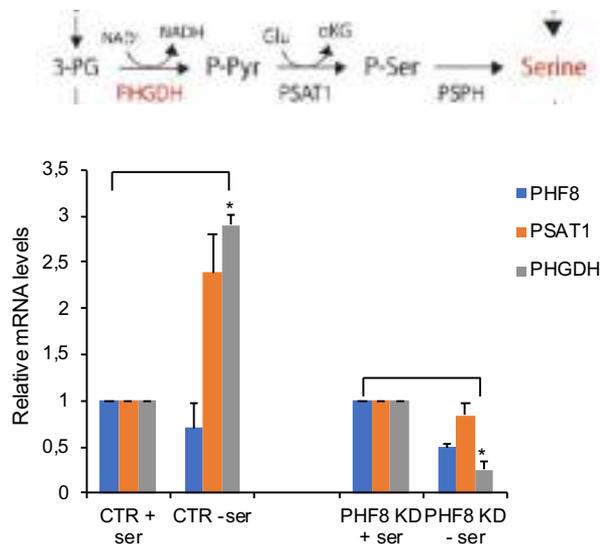


Figure R86: HeLa CTR and PHF8 KD were grown in media with or without serine. PHF8, PSAT1 and PHGDH levels were determined by qPCR.

2.4. *PHF8 directly regulates serine biosynthesis genes*

To check if PHF8 is bound to the metabolic enzymes responsible for serine biosynthesis that are transcriptionally downregulated in PHF8 KD cells (both in NSCs (FigureR81) and HeLa (FigureR82)), we analysed previously published PHF8 ChIP-seq performed in HeLa cells and compared PHF8 peaks with some histone modifications associated to active or repressed transcription. We found that PHF8 actually was bound to PSAT1 and PHGDH TSS and it colocalized with high levels of H3K4me3 (an activation mark) and low levels of H4K20me1 and H3K9me2 (the histone marks demethylated by PHF8) and low levels of H3K9me3 (a repression mark) in HeLa (Figure R87).

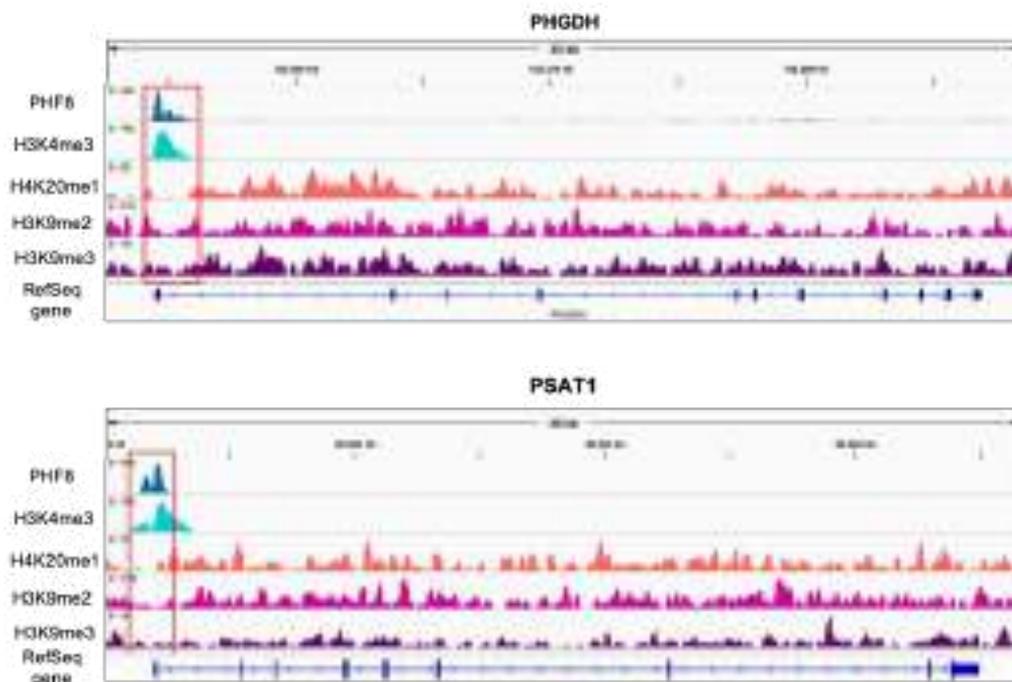


Figure R87: IGV captures showing PHF8 peaks and ChIP-seq profiles of H3K4me3, H4K20me1, H3K9me2, and H3K9me3 in PHGDH and PSAT1 genes.

2.5. *PHF8 cooperates with c-MYC to regulate transcription of serine biosynthesis genes*

Next, we sought to investigate how PHF8 is targeted to serine biosynthesis gene promoters. We analysed the promoter sequences of metabolic genes whose expression is affected by the depletion of PHF8 (identified in the RNA-seq) using Pscan tool [230]. Binding sites for NRF1, ATF4, E2F4 and MYC family among others were identified. As c-Myc transcription factor has been reported to regulate serine biosynthesis gene expression [231], and PHF8 cooperates with c-MYC in transcriptional activation [70], we investigated whether it contributes to PHF8-mediated transcriptional regulation of PHGDH and PSAT1.

To do so, we compared previously published PHF8 ChIP-seq in HeLa with the c-MYC ChIP-seq in the same cell line. Data in Figure R88 show the

colocalization of PHF8 and c-MYC binding sites at PHGDH and PSAT1 TSS.

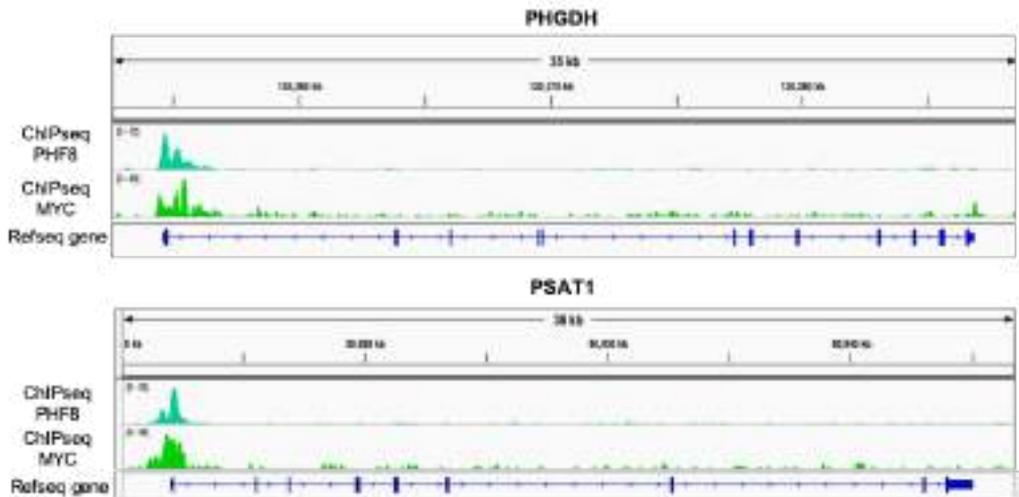


Figure R88: IGV captures showing PHF8 peaks and ChIP-seq profiles of c-MYC at PSAT1 and PHGDH genes.

Next, we analysed if PHF8 and c-MYC colocalized genome wide (Figure R89). We observed a good overlap of the two proteins. 41,6% of c-MYC bound promoters were also bound by PHF8. The results indicate that PHF8 and c-MYC could cooperate to regulate transcription, in particular to regulate serine biosynthesis genes.

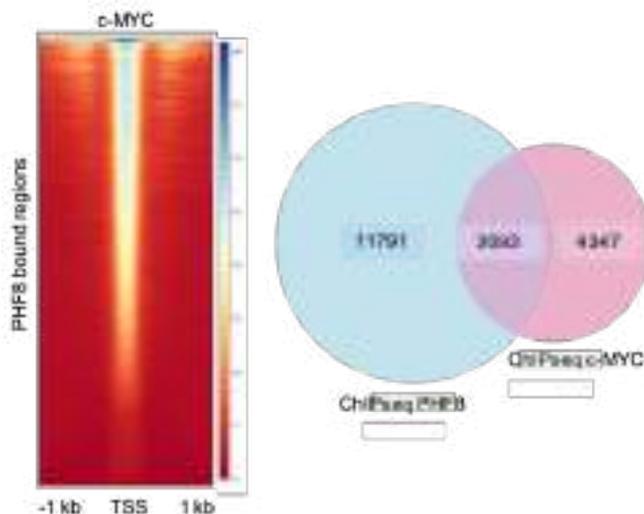


Figure R89: Heatmap depicting PHF8 binding to promoters bound by c-MYC 1 kb around TSS (left). Venn diagram showing overlap between PHF8 and c-MYC bound regions (right).

Altogether, these data demonstrate that PHF8 depletion in NSCs and HeLa affects the serine biosynthesis pathway leading to low levels of serine in PHF8 depleted cells. PHF8, probably cooperating with c-MYC, regulates two major enzymes involved in serine biosynthesis demethylating H4K20me1 on PHGDH and PSAT1 TSS.

Discussion

In this section I will comment on my results considering the current bibliography. The discussed topics are ordered according to their appearance in the Results section.

1. Regarding the function of PHF8 in astrocytes

1.1. *PHF8 levels are regulated during astrocyte differentiation through Notch signaling*

Knowing that PHF8 mRNA levels are especially high in fetal astrocytes, one of the first experiments that we performed was to check PHF8 expression both at RNA and protein levels during astrocyte differentiation. The results indicated that PHF8 is modulated during astrocyte differentiation *in vitro*. In addition, we isolated embryonic and postnatal mice astrocytes to evaluate PHF8 protein levels. So, we could demonstrate that PHF8 is actually modulated *in vivo* during early stages of astrocyte differentiation. These data suggest a potential contribution of PHF8 to both astrocyte differentiation and function. Then, we decided to investigate which was the signaling pathway that modulates PHF8 levels during astrocyte differentiation. It is well known that Notch signaling is crucial in astrocyte differentiation [224, 225], so we analysed if Notch and other signaling could induce PHF8. We obtained that Notch induces PHF8 overexpression and, interestingly, that PHF8 levels regulate Notch target genes like the well-known transcription factor *Hes5*. Those data led us to hypothesize about the existence of a regulatory feedback loop between PHF8 and Notch signaling. A support to this hypothesis came from the bioinformatic analysis of the PHF8 ChIP-seq, which identified the predicted PHF8 binding sites in astrocyte genome. We obtained that one of the most statistically significant binding sites was RBPJ1 DNA binding motif. RBPJ1 is an essential effector of Notch signaling pathway; when Notch is activated it stimulates RBPJ1 to bind specific regions of DNA, thus controlling the activity of genes related to neural development and, in this case, astrocyte differentiation [225]. Interestingly, RBPJ1 motif and PHF8 binding were identified at *Phf8* gene supporting the idea that PHF8, in addition to regulate Notch targets, is itself a target (Figure D1).

Moreover, RBPJ1-PHF8 interaction has been previously identified in another cellular context showing that PHF8 is involved in Notch target genes activation [232]. Performing GO analysis of the PHF8 and RBPJ1 bound regions, we detected that they bind to genes related to assembly and function of synapses in astrocytes. Moreover, we discovered that, among the genes differentially expressed between astrocytes control and PHF8 KD, there were many Notch target genes like *Hes5*, *Dll3*, *Dll1*, *Cd44*. Those results, together with the finding that *Nfia* (another Notch target gene [145, 233]) is downregulated, indicate a close interplay between Notch signaling and PHF8 during astrocyte differentiation. To conclude, we revealed that *Phf8* is itself a Notch target, suggesting the existence of a regulatory feedback mechanism responsible for *Nfia* transcriptional control that allow the proper astrocyte differentiation.

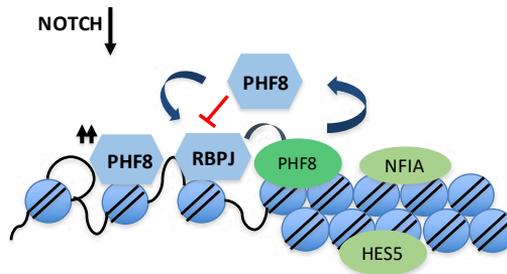


Figure D1: Model depicting the regulatory feedback loop between PHF8 and Notch signaling.

1.2. *PHF8 directly regulates a key astrogenic gene: Nfia and its depletion impairs astrocyte differentiation*

NFIA is known to be the master regulator of astrocyte differentiation as it activates the expression of astrocyte-specific genes [144, 233, 234]. We demonstrated that NFIA is a PHF8 transcriptional target by RNA-seq and CHIP-seq experiments; and consequently, PHF8 depletion impairs proper NFIA expression in astrocytes. We observed that PHF8 KD astrocytes are somehow defective; they do not express properly the glial fibrillary acid protein (GFAP) but they express other typical astrocytic markers. This phenotype well resembles the one found in mice *Nfia* knock-out that show normal expression of astrocyte markers but decreased levels of GFAP in the cortex and the hippocampus [235]. Many studies have demonstrated that NFIA occupies and regulates the GFAP promoter prior to

the induction of astrocyte differentiation [145, 233, 236]. Thus, the downregulation of *Nfia* observed in PHF8 depleted astrocytes could be responsible of the phenotype observed in PHF8 KD defective astrocytes. We also observed that in PHF8 depleted astrocytes many genes related to synapse formation and maturation are downregulated. A recent publication showed that *Nfia* loss in astrocytes leads to diminution of the synaptic function [146], so NFIA depletion, caused by PHF8 depletion, could, in part, be responsible of the defects observed in synapses too. However, we suggest that it is more likely that PHF8 directly regulates those synaptic genes in astrocytes, as we demonstrate by RNA-seq, ChIP-seq and rescue experiments, and that its depletion lead to impaired synaptogenesis in neurons/astrocytes co-cultures (Figure D2).

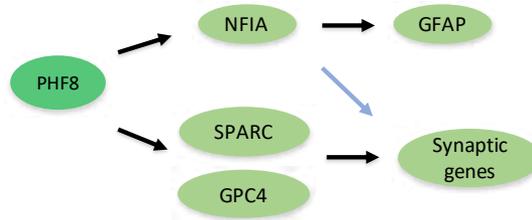


Figure D2: Model depicting PHF8 regulation of NFIA and synaptic genes.

We defined the astrocytes PHF8 KD defective not only because they don't express GFAP properly but also because they misexpress some genes typical of oligodendrocyte lineage. Upon PHF8 depletion a subpopulation of astrocytes expressed high levels of the marker OLIG2 and low levels of NG2. Interestingly, it has been demonstrated an inverse correlation between OLIG2 levels and GFAP expression in some astrocyte subtypes [237]. They have been observed astrocytes expressing OLIG2 in the grey matter of the mouse spinal cord, thalamus and forebrain [238-241]. Thus, it could be possible that PHF8 depletion redirects the differentiation towards those astrocyte subtype, but it wouldn't explain the presence of the NG2 marker. NG2 chondroitin sulfate proteoglycan is one of the most reliable and widely-used markers for oligodendrocyte progenitor cells (OPCs) [242-244]. During development, NG2 cells give rise to oligodendrocytes but some remain as NG2-glia that persist into adulthood; *in vitro* studies showed the bipotential ability of oligodendrocyte progenitor cells to differentiate into astrocytes and oligodendrocytes, however there are no demonstration of the ability of NG2-glia to differentiate into astrocytes *in vivo* [245]. Considering those data, we can both hypothesize that PHF8 KD astrocytes misexpress NG2

or that those NG2 positive cells are an immature population of glial cells with the potential to differentiate into oligodendrocytes.

1.3. *PHF8 directly regulates synaptogenic genes and its depletion causes aberrant synaptogenesis*

We demonstrated that genes involved in synapse formation and maturation are directly regulated by PHF8 in astrocytes. In the last years, it became clear the importance of astrocytes that function by promoting synapse formation, pruning aberrant synapse and regulating synaptic plasticity [246, 247]. For example, *Sparc* and *Gpc4*, which are both downregulated upon PHF8 depletion in astrocytes, participate to the organization of active synapse connections coordinating pre and post synaptic neurons. They induce the formation of active excitatory synapses recruiting AMPA glutamate receptors to the postsynaptic cell surface [172, 173]. We demonstrated that wild-type neurons cultured on astrocytes PHF8 KD show a reduced expression of synaptic proteins and a clear reduction in the strength of synaptic transmission. We can conclude that depletion of astrocytic PHF8 causes decreased density and strength of excitatory synapses in neurons-astrocytes cocultures *in vitro*. This phenotype is particularly relevant if we consider that PHF8 mutations in the catalytic domain have been associated to X-linked intellectual disabilities (XLID) [90, 248, 249] and that we could rescue the transcription of synaptic genes by the overexpression of PHF8 WT but not of the catalytic mutant. Previous studies have shown that *Phf8* KO mice are deficient in learning and memory [97] and they are resistant to anxiety- and depression-like behaviors [96]. However, only slight transcriptional changes were observed in neurons of those *Phf8* KO mice and the effects on glial cells were not evaluated. Now, we demonstrate that PHF8 induces profound transcriptional changes in astrocyte synaptic genes, which in turn, strongly affect synaptic transmission. Thus, it should be considered the possibility that PHF8 depletion especially affects astrocytes differentiation and function to further understand the complex mechanisms causing XLID.

1.4. *PHF8 prevents ectopic heterochromatin formation at astrogenic and synaptogenic genes*

We also determined the molecular mechanism behind PHF8-mediated transcriptional changes. Our results demonstrated that PHF8 maintains the expression of *Nfia* and synaptic genes by keeping low levels of the H4K20me1 histone mark. In fact, PHF8 KD in astrocytes caused elevated H4K20me1 levels at synaptic genes, and it correlated with a weak transcription of those genes. These data agree with previous studies demonstrating that depletion of PHF8 in neurons resulted in down regulation of cytoskeleton genes by increasing H4K20me1 levels [94]. Moreover, we showed that upon PHF8 depletion the H4K20me3, a histone mark typical of heterochromatin, became upregulated both at global (immunostaining experiment) and at local levels (ChIP-qPCR). Thus, PHF8 demethylase activity is fundamental to maintain the H4K20me1/H4K20me3 equilibrium on those genes that participate to astrocytes differentiation and synapse maturation. H4K20me3 upregulation in PHF8 KD astrocytes suggests that H4K20me1 is being used as a substrate for the HMTs Suv420H1/2 to generate H4K20me3. Finally, the importance of PHF8 HDM activity was demonstrated by the rescue of H4K20me3 levels after the overexpression of PHF8 WT, but not of the catalytic mutant. Interestingly, it has been published that another demethylase of the KDM7 family, PHF2, limits the accumulation of the heterochromatic mark H3K9me3 at the promoters of cell cycle genes [250].

We believe that the results obtained in this first part of the thesis, implement our knowledge about the role of PHF8 in astrocyte differentiation and synaptic formation, suggesting that PHF8 may be a key regulator of astrogliogenesis and synaptogenesis (Figure D3). These findings can be important to understand the mechanisms of XLID, in which PHF8 plays a role, even if it is still unknown the exact mechanism. Moreover, they suggest that H4K20me1/H4K20me3 imbalance can significantly contribute to neurodevelopmental disorders, highlighting the necessity to investigate the mechanisms of crosstalk between epigenetics, development, and diseases.

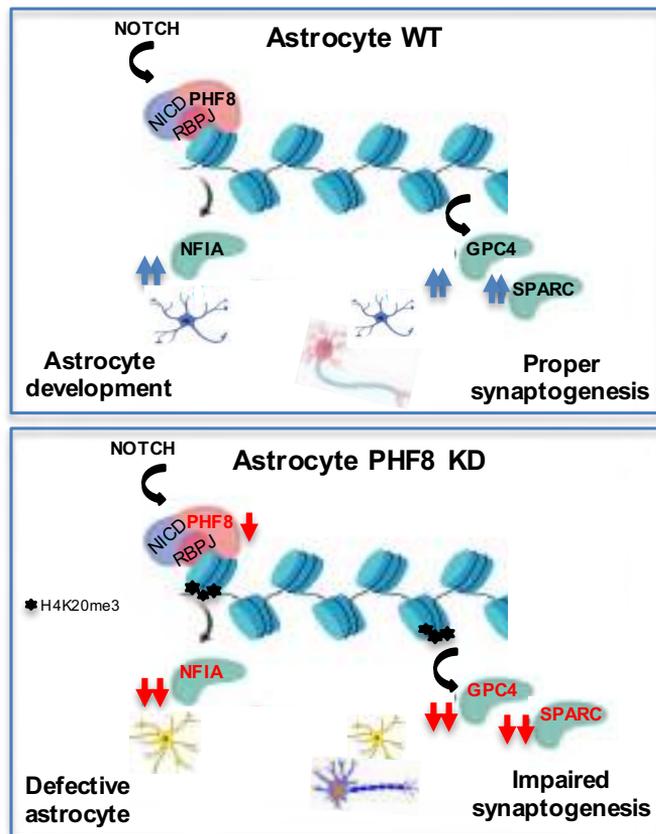


Figure D3: Model depicting the contribution of PHF8 to astrocyte differentiation and function. PHF8 directly regulates the expression of the master regulator of astrocyte differentiation *Nfia* as well as genes involved in synapses. Depletion or alterations of PHF8 catalytic activity lead to defective astrocytes that are deficient in synaptic function.

2. Regarding the function of PHF8 in neural stem cells

2.1. PHF8 regulates cell cycle and metabolic genes expression

To investigate PHF8 function in NSCs, we first analysed PHF8 KD cells phenotype. As demonstrated in other cell lines [70], we found that NSCs PHF8 KD suffer cell cycle delay; flow-cytometry experiment showed a delay in G1/S transition and consequently delayed cell growth. Similar

results were obtained for another member of the KDM7 demethylase family, PHF2 [250] *in vitro* and *in vivo*, suggesting that the demethylase of KDM7 family could promote neural progenitor proliferation. Those data were reinforced by the RNA-seq experiment, in which we found that actually many genes related to cell cycle were differentially expressed in NSCs CTR compared to PHF8 KD cells. The GO analysis revealed that, beyond the category of cell cycle, DNA replication and cell proliferation, there were others related to the metabolic processes. It was particularly interesting the discovery of categories related to cell metabolism, as still very little is known about PHF8 and metabolism [87, 88], especially in mammals. Considering those results, we concluded that PHF8 contributes to cell cycle progression by facilitating the expression of cell cycle and metabolism related genes. We also noticed that the H4K20me1, which is the main histone mark demethylated by PHF8, and two others histone marks typical of heterochromatin, H4K20me3 and H3K9me3, were upregulated in PHF8 KD NSCs. H4K20me3 upregulation is probably caused by elevated H4K20me1 levels that can be used by the HMTs Suv420H1/2. Instead, we still don't know if H3K9me3 upregulation is a direct or indirect effect of PHF8 depletion. We could speculate that the elevated H3K9me3 levels observed in NSCs PHF8 KD are an effect of the cell cycle arrest, which could cause senescence in those cells. Another possibility is that PHF8 depletion, downregulating PSAT1, directly affects α -ketoglutarate levels that regulate H3K9me3 levels, as it has been demonstrated in embryonic stem cells [251]. Moreover, we showed that PHF8 depletion in NSCs induces nuclei defects like chromatin bridges and multinuclear cells, as it was shown in cancer cells too [83]. In fact, it has been recently published that PHF8 ensures the replication fork restart permitting the recovery from replication stress and thus maintaining genome stability [86].

2.2. PHF8 depletion impairs serine metabolism

Indications that de novo serine metabolism could be affected in PHF8 depleted NSCs came from the RNA-seq analysis, in which we found that two key enzymes in serine biosynthesis, PHGDH and PSAT1, were both downregulated in PHF8 KD NSCs compared to CTR cells. We confirmed that PSAT1 is downregulated in PHF8 KD HeLa too, suggesting that PHF8 depletion could impair serine metabolism in different cell types or, at least, in high proliferating cells like NSCs and cancer cells. Serine is an important 1 carbon donor to the folate cycle, contributes to the synthesis of nucleotides, amino acid, phospholipids and glutathione, which is

essential for redox buffering [252]. The novo serine synthesis is one of many side branches of glycolysis, allowing to convert glucose-derived carbons into serine (Figure D4).

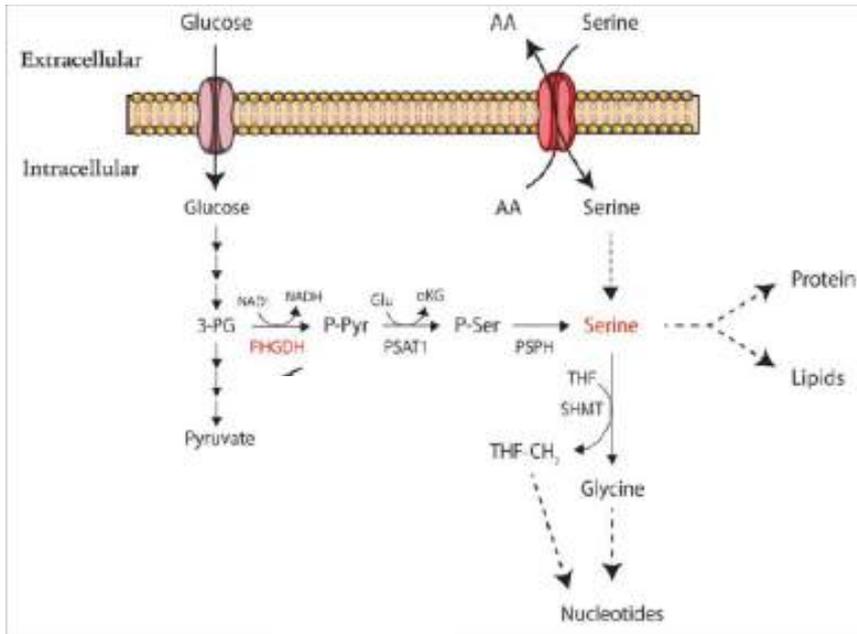


Figure D4: Serine concentrations stimulates glutathione (GSH) synthesis and feeds into the one-carbon metabolic network (1CMet). Adapted from [253].

A key experiment demonstrating that actually PHF8 KD suffers serine deprivation came from the GC-MS results. It also revealed that glutamic acid concentration is lower in PHF8 KD cells compared to CTR. We noticed that some genes related to glutamate (the negative ion form of glutamic acid) metabolism were misexpressed in the RNA-seq too (Glud1, Gpt2 and Got1). Glutamate is synthesized from α KG in the tricarboxylic acid cycle (TCA) by glutamate dehydrogenase (Gdh) with NH_3 and NADPH. Glutamine is synthesized from glutamate by glutamine synthetase (Gs) with NH_3 and ATP. L-Glutamine is a nutritionally semi-essential amino acid that plays an important role in promoting protein synthesis and generating the antioxidant glutathione. Extracellular L-glutamine crosses the plasma membrane, through a transport system, and can be converted into α -ketoglutarate through two pathways called the glutaminase I and II. The α -ketoglutarate is an intermediate metabolite in the TCA and plays a key role in cell proliferation. It is generated by glutaminase (Gls) and glutamate dehydrogenase 1 (Glud1) during glutamine metabolism [254]. Moreover, it is produced by the

isocitrate dehydrogenase (Idh) during the tricarboxylic acid cycle, and by transaminases like Psat1 during the glycolysis- branched seine synthesis [255](Figure D5).

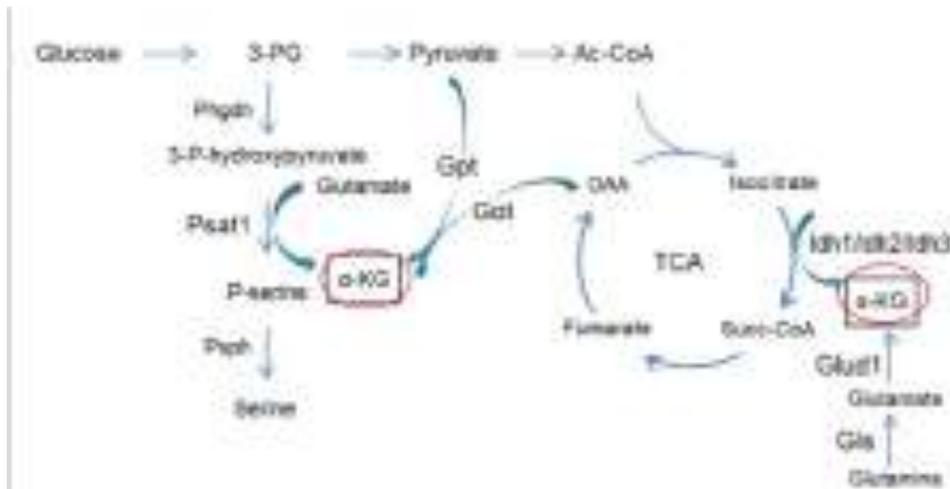


Figure D5: Phosphoserine aminotransferase 1 (Psat1), can regulate changes in α -ketoglutarate concentration. α -ketoglutarate is also produced during TCA and through Glut1. Adapted from [251].

Thus, we tried to rescue the cell proliferation delay defects of PHF8 KD cells administrating serine, glutamine or α -ketoglutarate. It is important to consider that α -ketoglutarate is the cofactor of that JmjC-domain-containing histone demethylases, like PHF8, so its concentration can be also relevant to determine PHF8 activity. We obtained that only serine could rescue PHF8 KD cell cycle delay, even if it was not able to restore the proliferation rate at the levels of CTR cells, indicating that other mechanisms actually affect PHF8 KD cell growth. Neither glutamine nor α -ketoglutarate were able to rescue the PHF8 KD cell cycle delay suggesting that PHF8 depleted cells do not depend on their concentrations for growing, at least in the conditions we tested. As I previously mentioned, it has been demonstrated that α -ketoglutarate regulates H3K9me3, so it could be possible that the elevated levels of H3K9me3 and H4K20me3, observed in NSCs PHF8 KD, are caused by a low concentration of α -ketoglutarate that impairs demethylases activity.

It has been reported that de novo serine metabolism functions as feedstock for 1 carbon donation and that serine deprivation impairs cell growth in cancer cells [255]. Another interesting study relates mutations

in enzymes critical for serine metabolism, PHGDH and PSAT1, with microcephaly [256]. Microcephaly is usually associated with XLID [257] and defects in serine metabolism, due to PHF8 depletion, impairs neural stem cells growth. Thus, PHF8 depletion and serine deprivation, could be a new mechanism to further study in order to better understand XLID.

2.3. PHF8 regulates the expression of serine biosynthesis genes cooperating with c-MYC

Analysing previously published ChIP-seq data, we could show that PHF8 binds serine biosynthesis gene promoters, PHGDH and PSAT1. PHF8 binding correlates with low levels of H4K20me1 and H3K9me2 and transcription activation. Besides, we investigated how PHF8 is targeted to serine biosynthesis gene promoters. c-Myc, which actually has been described to regulate serine biosynthesis genes [231], appeared among the binding site sequences most enriched in those metabolic genes differentially expressed in PHF8 KD NSCs compared to CTR cells. It is known that PHF8 cooperates with c-Myc in transcriptional activation [70], so it is plausible that PHF8 cooperates with it to directly demethylates H4K20me1 on PHGDH and PSAT1 TSS, as indicated by the overlap of their ChIP-seq peaks too. These data are very preliminary and they need further investigations, but we consider that they could be very promising as they connect cell cycle defects, due to PHF8 depletion, with serine metabolism, providing a new avenue to be explored in order to better understand XLID.

Conclusions

1. PHF8 levels are modulated during astrocytes differentiation.
2. PHF8 is a regulator of Notch signaling pathway.
3. PHF8 directly regulates *Nfia*, a master transcription factor of astrocyte differentiation.
4. PHF8 knock down NSCs differentiate into defective astrocytes
5. PHF8 directly regulates synaptogenic genes like *Sparc* and *Gpc4* in astrocytes.
6. Depletion of astrocytic PHF8 causes decreased density and strength of excitatory synapses in neurons-astrocytes cocultures *in vitro*.
7. PHF8 demethylates H4K20me1 at astrogenic and synaptogenic genes to facilitate their transcriptional activation, preventing ectopic heterochromatin formation.
8. PHF8 demethylase activity is crucial for proper expression of astrogenic and synaptogenic genes.
9. PHF8 regulates cell cycle and metabolic genes transcription in NSCs and it is essential for proliferation.
10. PHF8 depletion impairs serine metabolism both in NSCs and HeLa cells.
11. PHF8 knock down HeLa cells show low concentration of serine.
12. Serine addition in PHF8 knock down HeLa cells partially rescue PHF8 KD proliferation defects.
13. PHF8 binds and demethylates H4K20me1 serine biosynthesis gene promoters.

Bibliography

1. *The Gene Ontology Resource: 20 years and still GOing strong.* Nucleic Acids Res, 2019. **47**(D1): p. D330-d338.
2. Kornberg, R.D., *Chromatin Structure: A Repeating Unit of Histones and DNA.* Science, 1974. **184**(4139): p. 868.
3. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 Å resolution.* Nature, 1997. **389**(6648): p. 251-260.
4. Li, G. and D. Reinberg, *Chromatin higher-order structures and gene regulation.* Current Opinion in Genetics & Development, 2011. **21**(2): p. 175-186.
5. Syed, S.H., et al., *Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome.* Proc Natl Acad Sci U S A, 2010. **107**(21): p. 9620-5.
6. Li, G. and P. Zhu, *Structure and organization of chromatin fiber in the nucleus.* FEBS Lett, 2015. **589**(20 Pt A): p. 2893-904.
7. Ausió, J., *The shades of gray of the chromatin fiber: recent literature provides new insights into the structure of chromatin.* Bioessays, 2015. **37**(1): p. 46-51.
8. Prieto, E.I. and K. Maeshima, *Dynamic chromatin organization in the cell.* Essays Biochem, 2019. **63**(1): p. 133-145.
9. Szalaj, P. and D. Plewczynski, *Three-dimensional organization and dynamics of the genome.* Cell Biol Toxicol, 2018. **34**(5): p. 381-404.
10. Kadauke, S. and G.A. Blobel, *Chromatin loops in gene regulation.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2009. **1789**(1): p. 17-25.
11. Nora, E.P., et al., *Spatial partitioning of the regulatory landscape of the X-inactivation centre.* Nature, 2012. **485**(7398): p. 381-5.
12. Jansen, A. and K.J. Verstrepen, *Nucleosome positioning in Saccharomyces cerevisiae.* Microbiology and molecular biology reviews : MMBR, 2011. **75**(2): p. 301-320.
13. Lieberman-Aiden, E., et al., *Comprehensive mapping of long-range interactions reveals folding principles of the human genome.* Science, 2009. **326**(5950): p. 289-93.
14. Ryba, T., et al., *Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types.* Genome Res, 2010. **20**(6): p. 761-70.
15. Nagano, T., et al., *Single-cell Hi-C reveals cell-to-cell variability in chromosome structure.* Nature, 2013. **502**(7469): p. 59-64.

16. Cremer, T., et al., *Chromosome territories--a functional nuclear landscape*. *Curr Opin Cell Biol*, 2006. **18**(3): p. 307-16.
17. Bogdanović, O. and R. Lister, *DNA methylation and the preservation of cell identity*. *Curr Opin Genet Dev*, 2017. **46**: p. 9-14.
18. Bird, A.P., et al., *Non-methylated CpG-rich islands at the human alpha-globin locus: implications for evolution of the alpha-globin pseudogene*. *Embo j*, 1987. **6**(4): p. 999-1004.
19. Jaenisch, R., et al., *DNA methylation, retroviruses, and embryogenesis*. *Journal of Cellular Biochemistry*, 1982. **20**(4): p. 331-336.
20. Bestor, T.H., *Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain*. *Embo j*, 1992. **11**(7): p. 2611-7.
21. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases*. *Nature Genetics*, 1998. **19**(3): p. 219-220.
22. Bird, A.P. and A.P. Wolffe, *Methylation-Induced Repression*; *Belts, Braces, and Chromatin*. *Cell*, 1999. **99**(5): p. 451-454.
23. Nabel, C.S., S.A. Manning, and R.M. Kohli, *The curious chemical biology of cytosine: deamination, methylation, and oxidation as modulators of genomic potential*. *ACS Chem Biol*, 2012. **7**(1): p. 20-30.
24. Kohli, R.M. and Y. Zhang, *TET enzymes, TDG and the dynamics of DNA demethylation*. *Nature*, 2013. **502**(7472): p. 472-479.
25. Rasmussen, K.D. and K. Helin, *Role of TET enzymes in DNA methylation, development, and cancer*. *Genes & development*, 2016. **30**(7): p. 733-750.
26. Wu, X. and Y. Zhang, *TET-mediated active DNA demethylation: mechanism, function and beyond*. *Nature Reviews Genetics*, 2017. **18**(9): p. 517-534.
27. Jones, P.A., *Functions of DNA methylation: islands, start sites, gene bodies and beyond*. *Nature Reviews Genetics*, 2012. **13**(7): p. 484-492.
28. Mariño-Ramírez, L., et al., *Histone structure and nucleosome stability*. *Expert Review of Proteomics*, 2005. **2**(5): p. 719-729.
29. Iwasaki, W., et al., *Contribution of histone N-terminal tails to the structure and stability of nucleosomes*. *FEBS Open Bio*, 2013. **3**: p. 363-369.

30. Kimura, H., *Histone modifications for human epigenome analysis*. Journal of Human Genetics, 2013. **58**(7): p. 439-445.
31. Yang, X.J. and E. Seto, *HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention*. Oncogene, 2007. **26**(37): p. 5310-8.
32. Cornett, E.M., et al., *Lysine Methylation Regulators Moonlighting outside the Epigenome*. Mol Cell, 2019. **75**(6): p. 1092-1101.
33. van Nuland, R. and O. Gozani, *Histone H4 Lysine 20 (H4K20) Methylation, Expanding the Signaling Potential of the Proteome One Methyl Moiety at a Time*. Molecular & Cellular Proteomics, 2016. **15**(3): p. 755.
34. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-37.
35. Bicknell, L.S., et al., *Mutations in the pre-replication complex cause Meier-Gorlin syndrome*. Nature genetics, 2011. **43**(4): p. 356-359.
36. Kim, Y.Z., *Altered Histone Modifications in Gliomas*. Brain Tumor Res Treat, 2014. **2**(1): p. 7-21.
37. Kapoor-Vazirani, P., J.D. Kagey, and P.M. Vertino, *SUV420H2-mediated H4K20 trimethylation enforces RNA polymerase II promoter-proximal pausing by blocking hMOF-dependent H4K16 acetylation*. Molecular and cellular biology, 2011. **31**(8): p. 1594-1609.
38. Krishnan, V., et al., *Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice*. Proceedings of the National Academy of Sciences, 2011. **108**(30): p. 12325.
39. Taylor, G.C.A., et al., *H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction*. Genome Research, 2013. **23**(12): p. 2053-2065.
40. Kuo, A.J., et al., *The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier–Gorlin syndrome*. Nature, 2012. **484**(7392): p. 115-119.
41. Hsiao, K.-Y. and C.A. Mizzen, *Histone H4 deacetylation facilitates 53BP1 DNA damage signaling and double-strand break repair*. Journal of Molecular Cell Biology, 2013. **5**(3): p. 157-165.
42. Schotta, G., et al., *A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin*. Genes Dev, 2004. **18**(11): p. 1251-62.

43. Barski, A., et al., *High-Resolution Profiling of Histone Methylations in the Human Genome*. Cell, 2007. **129**(4): p. 823-837.
44. Rosenfeld, J.A., et al., *Determination of enriched histone modifications in non-genic portions of the human genome*. BMC Genomics, 2009. **10**(1): p. 143.
45. Guelen, L., et al., *Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions*. Nature, 2008. **453**(7197): p. 948-951.
46. Horton, J.R., et al., *Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases*. Nature Structural & Molecular Biology, 2010. **17**(1): p. 38-43.
47. Epsztejn-Litman, S., et al., *De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes*. Nat Struct Mol Biol, 2008. **15**(11): p. 1176-1183.
48. Vakoc, C.R., et al., *Histone H3 Lysine 9 Methylation and HP1 Are Associated with Transcription Elongation through Mammalian Chromatin*. Molecular Cell, 2005. **19**(3): p. 381-391.
49. Regha, K., et al., *Active and Repressive Chromatin Are Interspersed without Spreading in an Imprinted Gene Cluster in the Mammalian Genome*. Molecular Cell, 2007. **27**(3): p. 353-366.
50. Narita, M., et al., *Rb-Mediated Heterochromatin Formation and Silencing of E2F Target Genes during Cellular Senescence*. Cell, 2003. **113**(6): p. 703-716.
51. Husmann, D. and O. Gozani, *Histone lysine methyltransferases in biology and disease*. Nature Structural & Molecular Biology, 2019. **26**(10): p. 880-889.
52. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases*. Nature, 2000. **406**(6796): p. 593-9.
53. Bergink, S., et al., *Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex*. J Cell Biol, 2012. **196**(6): p. 681-8.
54. Cao, X., et al., *Histone H4K20 Demethylation by Two hHR23 Proteins*. Cell Rep, 2020. **30**(12): p. 4152-4164.e6.
55. Fueyo, R., M.A. García, and M.A. Martínez-Balbás, *Jumonji family histone demethylases in neural development*. Cell Tissue Res, 2015. **359**(1): p. 87-98.
56. Tsukada, Y., et al., *Histone demethylation by a family of JmjC domain-containing proteins*. Nature, 2006. **439**(7078): p. 811-6.

57. Loh, Y.H., et al., *Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells*. *Genes Dev*, 2007. **21**(20): p. 2545-57.
58. Jepsen, K., et al., *SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron*. *Nature*, 2007. **450**(7168): p. 415-9.
59. Iwase, S., et al., *The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases*. *Cell*, 2007. **128**(6): p. 1077-88.
60. Hsia, D.A., et al., *KDM8, a H3K36me2 histone demethylase that acts in the cyclin A1 coding region to regulate cancer cell proliferation*. *Proc Natl Acad Sci U S A*, 2010. **107**(21): p. 9671-6.
61. Tateishi, K., et al., *Role of Jhdm2a in regulating metabolic gene expression and obesity resistance*. *Nature*, 2009. **458**(7239): p. 757-61.
62. Klose, R.J., E.M. Kallin, and Y. Zhang, *JmjC-domain-containing proteins and histone demethylation*. *Nat Rev Genet*, 2006. **7**(9): p. 715-27.
63. Labbé, R.M., A. Holowatyj, and Z.-Q. Yang, *Histone lysine demethylase (KDM) subfamily 4: structures, functions and therapeutic potential*. *American journal of translational research*, 2013. **6**(1): p. 1-15.
64. Klose, R.J., E.M. Kallin, and Y. Zhang, *JmjC-domain-containing proteins and histone demethylation*. *Nature Reviews Genetics*, 2006. **7**(9): p. 715-727.
65. Li, H., et al., *Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF*. *Nature*, 2006. **442**(7098): p. 91-5.
66. Fortschegger, K. and R. Shiekhhattar, *Plant homeodomain fingers form a helping hand for transcription*. *Epigenetics*, 2011. **6**(1): p. 4-8.
67. Baba, A., et al., *PKA-dependent regulation of the histone lysine demethylase complex PHF2-ARID5B*. *Nature Cell Biology*, 2011. **13**(6): p. 668-675.
68. Horton, J.R., et al., *Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases*. *Nat Struct Mol Biol*, 2010. **17**(1): p. 38-43.
69. Feng, W., et al., *PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation*. *Nat Struct Mol Biol*, 2010. **17**(4): p. 445-50.

70. Fortschegger, K., et al., *PHF8 targets histone methylation and RNA polymerase II to activate transcription*. Mol Cell Biol, 2010. **30**(13): p. 3286-98.
71. Chaturvedi, S.S., et al., *Structure-function relationships in KDM7 histone demethylases*. Adv Protein Chem Struct Biol, 2019. **117**: p. 113-125.
72. Intlekofer, A.M. and L.W.S. Finley, *Metabolic signatures of cancer cells and stem cells*. Nature metabolism, 2019. **1**(2): p. 177-188.
73. Carey, B.W., et al., *Intracellular α -ketoglutarate maintains the pluripotency of embryonic stem cells*. Nature, 2015. **518**(7539): p. 413-6.
74. Pan, M., et al., *Regional glutamine deficiency in tumours promotes dedifferentiation through inhibition of histone demethylation*. Nature Cell Biology, 2016. **18**(10): p. 1090-1101.
75. Yang, M. and K.H. Vousden, *Serine and one-carbon metabolism in cancer*. Nature Reviews Cancer, 2016. **16**(10): p. 650-662.
76. Kim, H. and Y.J. Park, *Links between Serine Biosynthesis Pathway and Epigenetics in Cancer Metabolism*. Clinical nutrition research, 2018. **7**(3): p. 153-160.
77. Ding, J., et al., *The histone H3 methyltransferase G9A epigenetically activates the serine-glycine synthesis pathway to sustain cancer cell survival and proliferation*. Cell Metab, 2013. **18**(6): p. 896-907.
78. Zhao, E., et al., *KDM4C and ATF4 Cooperate in Transcriptional Control of Amino Acid Metabolism*. Cell Rep, 2016. **14**(3): p. 506-519.
79. Wang, J., et al., *PHF8 and REST/NRSF co-occupy gene promoters to regulate proximal gene expression*. Sci Rep, 2014. **4**: p. 5008.
80. Asensio-Juan, E., et al., *The histone demethylase PHF8 is a molecular safeguard of the IFN γ response*. Nucleic Acids Res, 2017. **45**(7): p. 3800-3811.
81. Dimitrova, E., A.H. Turberfield, and R.J. Klose, *Histone demethylases in chromatin biology and beyond*. EMBO reports, 2015. **16**(12): p. 1620-1639.
82. Liu, W., et al., *PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression*. Nature, 2010. **466**(7305): p. 508-12.

83. Lim, H.J., et al., *The G2/M regulator histone demethylase PHF8 is targeted for degradation by the anaphase-promoting complex containing CDC20*. Mol Cell Biol, 2013. **33**(21): p. 4166-80.
84. Sun, L., et al., *Cyclin E-CDK2 protein phosphorylates plant homeodomain finger protein 8 (PHF8) and regulates its function in the cell cycle*. J Biol Chem, 2015. **290**(7): p. 4075-85.
85. Lee, C., et al., *A PHF8 homolog in C. elegans promotes DNA repair via homologous recombination*. PLoS One, 2015. **10**(4): p. e0123865.
86. Feng, H., et al., *CK2 kinase-mediated PHF8 phosphorylation controls TopBP1 stability to regulate DNA replication*. Nucleic Acids Research, 2020. **48**(19): p. 10940-10952.
87. Ying, Z., et al., *Short-Term Mitochondrial Permeability Transition Pore Opening Modulates Histone Lysine Methylation at the Early Phase of Somatic Cell Reprogramming*. Cell Metab, 2018. **28**(6): p. 935-945 e5.
88. Merkwirth, C., et al., *Two Conserved Histone Demethylases Regulate Mitochondrial Stress-Induced Longevity*. Cell, 2016. **165**(5): p. 1209-1223.
89. Laumonier, F., et al., *Mutations in PHF8 are associated with X linked mental retardation and cleft lip/cleft palate*. J Med Genet, 2005. **42**(10): p. 780-6.
90. Koivisto, A.M., et al., *Screening of mutations in the PHF8 gene and identification of a novel mutation in a Finnish family with XLMR and cleft lip/cleft palate*. Clin Genet, 2007. **72**(2): p. 145-9.
91. Qiu, J., et al., *The X-linked mental retardation gene PHF8 is a histone demethylase involved in neuronal differentiation*. Cell Res, 2010. **20**(8): p. 908-18.
92. Qi, H.H., et al., *Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development*. Nature, 2010. **466**(7305): p. 503-7.
93. Kleine-Kohlbrecher, D., et al., *A functional link between the histone demethylase PHF8 and the transcription factor ZNF711 in X-linked mental retardation*. Mol Cell, 2010. **38**(2): p. 165-78.
94. Asensio-Juan, E., C. Gallego, and M.A. Martínez-Balbás, *The histone demethylase PHF8 is essential for cytoskeleton dynamics*. Nucleic Acids Res, 2012. **40**(19): p. 9429-40.
95. Oey, N.E., et al., *A Neuronal Activity-Dependent Dual Function Chromatin-Modifying Complex Regulates Arc Expression*. eNeuro, 2015. **2**(1).

96. Walsh, R.M., et al., *Phf8 loss confers resistance to depression-like and anxiety-like behaviors in mice*. Nat Commun, 2017. **8**: p. 15142.
97. Chen, X., et al., *Phf8 histone demethylase deficiency causes cognitive impairments through the mTOR pathway*. Nat Commun, 2018. **9**(1): p. 114.
98. Stevenson, R.E. and C.E. Schwartz, *Clinical and molecular contributions to the understanding of X-linked mental retardation*. Cytogenet Genome Res, 2002. **99**(1-4): p. 265-75.
99. Siderius, L.E., et al., *X-linked mental retardation associated with cleft lip/palate maps to Xp11.3-q21.3*. Am J Med Genet, 1999. **85**(3): p. 216-20.
100. Abidi, F., et al., *A novel mutation in the PHF8 gene is associated with X-linked mental retardation with cleft lip/cleft palate*. Clin Genet, 2007. **72**(1): p. 19-22.
101. Liu, Q., et al., *Contribution of synergism between PHF8 and HER2 signalling to breast cancer development and drug resistance*. EBioMedicine, 2020. **51**: p. 102612.
102. Shao, P., et al., *Histone demethylase PHF8 promotes epithelial to mesenchymal transition and breast tumorigenesis*. Nucleic Acids Res, 2017. **45**(4): p. 1687-1702.
103. Ma, Q., et al., *The histone demethylase PHF8 promotes prostate cancer cell growth by activating the oncomiR miR-125b*. Onco Targets Ther, 2015. **8**: p. 1979-88.
104. Tong, D., et al., *The HIF/PHF8/AR axis promotes prostate cancer progression*. Oncogenesis, 2016. **5**(12): p. e283.
105. Li, S., et al., *Histone demethylase PHF8 promotes progression and metastasis of gastric cancer*. Am J Cancer Res, 2017. **7**(3): p. 448-461.
106. Lv, Y., et al., *Histone demethylase PHF8 accelerates the progression of colorectal cancer and can be regulated by miR-488 in vitro*. Mol Med Rep, 2017. **16**(4): p. 4437-4444.
107. Zhou, W., et al., *PHF8 upregulation contributes to autophagic degradation of E-cadherin, epithelial-mesenchymal transition and metastasis in hepatocellular carcinoma*. J Exp Clin Cancer Res, 2018. **37**(1): p. 215.
108. Ye, H., et al., *PHF8 Plays an Oncogene Function in Hepatocellular Carcinoma Formation*. Oncol Res, 2019. **27**(5): p. 613-621.

109. Tasic, B., et al., *Adult mouse cortical cell taxonomy revealed by single cell transcriptomics*. *Nature Neuroscience*, 2016. **19**(2): p. 335-346.
110. Borello, U. and A. Pierani, *Patterning the cerebral cortex: traveling with morphogens*. *Current Opinion in Genetics & Development*, 2010. **20**(4): p. 408-415.
111. Altman, J., *Are New Neurons Formed in the Brains of Adult Mammals?* *Science*, 1962. **135**(3509): p. 1127.
112. Park, S.-Y. and J.-S. Han, *Phospholipase D1 Signaling: Essential Roles in Neural Stem Cell Differentiation*. *Journal of Molecular Neuroscience*, 2018. **64**(3): p. 333-340.
113. Conti, L., et al., *Niche-Independent Symmetrical Self-Renewal of a Mammalian Tissue Stem Cell*. *PLOS Biology*, 2005. **3**(9): p. e283.
114. Goffredo, D., et al., *Setting the conditions for efficient, robust and reproducible generation of functionally active neurons from adult subventricular zone-derived neural stem cells*. *Cell Death & Differentiation*, 2008. **15**(12): p. 1847-1856.
115. Reynolds, B.A. and R.L. Rietze, *Neural stem cells and neurospheres—re-evaluating the relationship*. *Nature Methods*, 2005. **2**(5): p. 333-336.
116. Conti, L. and E. Cattaneo, *Neural stem cell systems: physiological players or in vitro entities?* *Nat Rev Neurosci*, 2010. **11**(3): p. 176-87.
117. Tiberi, L., P. Vanderhaeghen, and J. van den Aamele, *Cortical neurogenesis and morphogens: diversity of cues, sources and functions*. *Current Opinion in Cell Biology*, 2012. **24**(2): p. 269-276.
118. Dennis, D., et al., *Forebrain neurogenesis: From embryo to adult*. *Trends in developmental biology*, 2016. **9**(1): p. 77-90.
119. Huang, C., J.A. Chan, and C. Schuurmans, *Proneural bHLH genes in development and disease*. *Curr Top Dev Biol*, 2014. **110**: p. 75-127.
120. Fode, C., et al., *A role for neural determination genes in specifying the dorsoventral identity of telencephalic neurons*. *Genes & development*, 2000. **14**(1): p. 67-80.
121. Nieto, M., et al., *Neural bHLH Genes Control the Neuronal versus Glial Fate Decision in Cortical Progenitors*. *Neuron*, 2001. **29**(2): p. 401-413.

122. Castro, D.S., et al., *Proneural bHLH and Brn proteins coregulate a neurogenic program through cooperative binding to a conserved DNA motif*. *Dev Cell*, 2006. **11**(6): p. 831-44.
123. Hatakeyama, J., et al., *Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation*. *Development*, 2004. **131**(22): p. 5539-50.
124. Imayoshi, I., et al., *Essential roles of Notch signaling in maintenance of neural stem cells in developing and adult brains*. *J Neurosci*, 2010. **30**(9): p. 3489-98.
125. Viti, J., A. Gulacsi, and L. Lillien, *Wnt regulation of progenitor maturation in the cortex depends on Shh or fibroblast growth factor 2*. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 2003. **23**(13): p. 5919-5927.
126. Guillemot, F., et al., *Molecular mechanisms of cortical differentiation*. *European Journal of Neuroscience*, 2006. **23**(4): p. 857-868.
127. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development*. *Cell*, 1999. **99**(3): p. 247-57.
128. Guo, J.U., et al., *Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain*. *Nature Neuroscience*, 2014. **17**(2): p. 215-222.
129. Lister, R., et al., *Global Epigenomic Reconfiguration During Mammalian Brain Development*. *Science*, 2013. **341**(6146): p. 1237905.
130. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells*. *Cell*, 2006. **125**(2): p. 315-26.
131. Boheler, K.R., *Stem cell pluripotency: a cellular trait that depends on transcription factors, chromatin state and a checkpoint deficient cell cycle*. *J Cell Physiol*, 2009. **221**(1): p. 10-7.
132. Scheper, W. and S. Copray, *The molecular mechanism of induced pluripotency: a two-stage switch*. *Stem Cell Rev Rep*, 2009. **5**(3): p. 204-23.
133. Hawkins, R.D., et al., *Distinct epigenomic landscapes of pluripotent and lineage-committed human cells*. *Cell Stem Cell*, 2010. **6**(5): p. 479-91.
134. Pereira, J.D., et al., *Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in*

- the cerebral cortex*. Proc Natl Acad Sci U S A, 2010. **107**(36): p. 15957-62.
135. Jepsen, K., et al., *SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron*. Nature, 2007. **450**(7168): p. 415-419.
 136. Fueyo, R., et al., *Lineage specific transcription factors and epigenetic regulators mediate TGF β -dependent enhancer activation*. Nucleic Acids Research, 2018. **46**(7): p. 3351-3365.
 137. Arnold, S.J., et al., *The T-box transcription factor Eomes/Tbr2 regulates neurogenesis in the cortical subventricular zone*. Genes & development, 2008. **22**(18): p. 2479-2484.
 138. Albert, M., et al., *Epigenome profiling and editing of neocortical progenitor cells during development*. The EMBO journal, 2017. **36**(17): p. 2642-2658.
 139. Lim, D.A., et al., *Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells*. Nature, 2009. **458**(7237): p. 529-533.
 140. MacDonald, J.L. and A.J. Roskams, *Histone deacetylases 1 and 2 are expressed at distinct stages of neuro-glial development*. Developmental Dynamics, 2008. **237**(8): p. 2256-2267.
 141. Albert, M. and W.B. Huttner, *Epigenetic and Transcriptional Pre-patterning—An Emerging Theme in Cortical Neurogenesis*. Frontiers in Neuroscience, 2018. **12**(359).
 142. Chambers, C.B., et al., *Spatiotemporal selectivity of response to Notch1 signals in mammalian forebrain precursors*. Development, 2001. **128**(5): p. 689-702.
 143. Grandbarbe, L., et al., *Delta-Notch signaling controls the generation of neurons/glia from neural stem cells in a stepwise process*. Development, 2003. **130**(7): p. 1391-402.
 144. Deneen, B., et al., *The Transcription Factor NFIA Controls the Onset of Gliogenesis in the Developing Spinal Cord*. Neuron, 2006. **52**(6): p. 953-968.
 145. Namihira, M., et al., *Committed Neuronal Precursors Confer Astrocytic Potential on Residual Neural Precursor Cells*. Developmental Cell, 2009. **16**(2): p. 245-255.
 146. Huang, A.Y.-S., et al., *Region-Specific Transcriptional Control of Astrocyte Function Oversees Local Circuit Activities*. Neuron, 2020.
 147. Kang, P., et al., *Sox9 and NFIA coordinate a transcriptional regulatory cascade during the initiation of gliogenesis*. Neuron, 2012. **74**(1): p. 79-94.

148. Kamakura, S., et al., *Hes binding to STAT3 mediates crosstalk between Notch and JAK–STAT signalling*. *Nature Cell Biology*, 2004. **6**(6): p. 547-554.
149. Gomes, W.A., M.F. Mehler, and J.A. Kessler, *Transgenic overexpression of BMP4 increases astroglial and decreases oligodendroglial lineage commitment*. *Developmental Biology*, 2003. **255**(1): p. 164-177.
150. See, J., et al., *BMP signaling mutant mice exhibit glial cell maturation defects*. *Molecular and cellular neurosciences*, 2007. **35**(1): p. 171-182.
151. Sanosaka, T., et al., *DNA Methylome Analysis Identifies Transcription Factor-Based Epigenomic Signatures of Multilineage Competence in Neural Stem/Progenitor Cells*. *Cell Rep*, 2017. **20**(12): p. 2992-3003.
152. Molne, M., et al., *Early cortical precursors do not undergo LIF-mediated astrocytic differentiation*. *J Neurosci Res*, 2000. **59**(3): p. 301-11.
153. Fan, G., et al., *DNA methylation controls the timing of astroglialogenesis through regulation of JAK-STAT signaling*. *Development*, 2005. **132**(15): p. 3345-56.
154. Cheng, P.Y., et al., *Interplay between SIN3A and STAT3 mediates chromatin conformational changes and GFAP expression during cellular differentiation*. *PLoS One*, 2011. **6**(7): p. e22018.
155. Namihira, M., K. Nakashima, and T. Taga, *Developmental stage dependent regulation of DNA methylation and chromatin modification in a immature astrocyte specific gene promoter*. *FEBS Letters*, 2004. **572**(1): p. 184-188.
156. Asano, H., et al., *Astrocyte differentiation of neural precursor cells is enhanced by retinoic acid through a change in epigenetic modification*. *Stem Cells*, 2009. **27**(11): p. 2744-52.
157. Chatagnon, A., et al., *RAR/RXR binding dynamics distinguish pluripotency from differentiation associated cis-regulatory elements*. *Nucleic acids research*, 2015. **43**(10): p. 4833-4854.
158. Pavlou, M.A.S., et al., *Transcriptional and epigenetic mechanisms underlying astrocyte identity*. *Prog Neurobiol*, 2019. **174**: p. 36-52.
159. Farhy-Tselnicker, I. and N.J. Allen, *Astrocytes, neurons, synapses: a tripartite view on cortical circuit development*. *Neural Development*, 2018. **13**(1): p. 7.
160. Ventura, R. and K.M. Harris, *Three-dimensional relationships between hippocampal synapses and astrocytes*. *The Journal of*

- neuroscience : the official journal of the Society for Neuroscience, 1999. **19**(16): p. 6897-6906.
161. Li, M., et al., *Synaptogenesis in the developing mouse visual cortex*. Brain Res Bull, 2010. **81**(1): p. 107-13.
 162. Parpura, V., et al., *Glutamate-mediated astrocyte-neuron signalling*. Nature, 1994. **369**(6483): p. 744-7.
 163. Barres, B.A., *The Mystery and Magic of Glia: A Perspective on Their Roles in Health and Disease*. Neuron, 2008. **60**(3): p. 430-440.
 164. Nedergaard, M., *Direct signaling from astrocytes to neurons in cultures of mammalian brain cells*. Science, 1994. **263**(5154): p. 1768-71.
 165. Halassa, M.M., et al., *Astrocytic Modulation of Sleep Homeostasis and Cognitive Consequences of Sleep Loss*. Neuron, 2009. **61**(2): p. 213-219.
 166. Schreiner, B., et al., *Astrocyte Depletion Impairs Redox Homeostasis and Triggers Neuronal Loss in the Adult CNS*. Cell Rep, 2015. **12**(9): p. 1377-84.
 167. Ge, W.-P., et al., *Local generation of glia is a major astrocyte source in postnatal cortex*. Nature, 2012. **484**(7394): p. 376-380.
 168. Christopherson, K.S., et al., *Thrombospondins Are Astrocyte-Secreted Proteins that Promote CNS Synaptogenesis*. Cell, 2005. **120**(3): p. 421-433.
 169. Allen, N.J., et al., *Astrocyte glypicans 4 and 6 promote formation of excitatory synapses via GluA1 AMPA receptors*. Nature, 2012. **486**(7403): p. 410-414.
 170. Eroglu, C., et al., *Gabapentin receptor alpha2delta-1 is a neuronal thrombospondin receptor responsible for excitatory CNS synaptogenesis*. Cell, 2009. **139**(2): p. 380-392.
 171. Kucukdereli, H., et al., *Control of excitatory CNS synaptogenesis by astrocyte-secreted proteins Hevin and SPARC*. Proc Natl Acad Sci U S A, 2011. **108**(32): p. E440-9.
 172. Singh, S.K., et al., *Astrocytes Assemble Thalamocortical Synapses by Bridging NRX1alpha and NL1 via Hevin*. Cell, 2016. **164**(1-2): p. 183-196.
 173. Farhy-Tselnicker, I., et al., *Astrocyte-Secreted Glypican 4 Regulates Release of Neuronal Pentraxin 1 from Axons to Induce Functional Synapse Formation*. Neuron, 2017. **96**(2): p. 428-445.e13.

174. Henneberger, C., et al., *Long-term potentiation depends on release of D-serine from astrocytes*. *Nature*, 2010. **463**(7278): p. 232-6.
175. Chung, W.-S., et al., *Astrocytes mediate synapse elimination through MEGF10 and MERTK pathways*. *Nature*, 2013. **504**(7480): p. 394-400.
176. Stogsdill, J.A. and C. Eroglu, *The interplay between neurons and glia in synapse development and plasticity*. *Curr Opin Neurobiol*, 2017. **42**: p. 1-8.
177. Zeisel, A., et al., *Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq*. *Science*, 2015. **347**(6226): p. 1138-42.
178. Lelieveld, S.H., et al., *Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability*. *Nat Neurosci*, 2016. **19**(9): p. 1194-6.
179. Higashimori, H., et al., *Astroglial FMRP-dependent translational down-regulation of mGluR5 underlies glutamate transporter GLT1 dysregulation in the fragile X mouse*. *Hum Mol Genet*, 2013. **22**(10): p. 2041-54.
180. Yasui, D.H., et al., *MeCP2 modulates gene expression pathways in astrocytes*. *Mol Autism*, 2013. **4**(1): p. 3.
181. Chao, H.T. and H.Y. Zoghbi, *MeCP2: only 100% will do*. *Nat Neurosci*, 2012. **15**(2): p. 176-7.
182. Ballas, N., et al., *Non-cell autonomous influence of MeCP2-deficient glia on neuronal dendritic morphology*. *Nat Neurosci*, 2009. **12**(3): p. 311-7.
183. Lioy, D.T., et al., *A role for glia in the progression of Rett's syndrome*. *Nature*, 2011. **475**(7357): p. 497-500.
184. Nguyen, M.V., et al., *MeCP2 is critical for maintaining mature neuronal networks and global brain anatomy during late stages of postnatal brain development and in the mature adult brain*. *J Neurosci*, 2012. **32**(29): p. 10021-34.
185. Forbes-Lorman, R.M., J.R. Kurian, and A.P. Auger, *MeCP2 regulates GFAP expression within the developing brain*. *Brain Research*, 2014. **1543**: p. 151-158.
186. Okabe, Y., et al., *Alterations of Gene Expression and Glutamate Clearance in Astrocytes Derived from an MeCP2-Null Mouse Model of Rett Syndrome*. *PLOS ONE*, 2012. **7**(4): p. e35354.
187. Pacey, L.K., et al., *Persistent astrocyte activation in the fragile X mouse cerebellum*. *Brain Behav*, 2015. **5**(10): p. e00400.

188. Higashimori, H., et al., *Selective Deletion of Astroglial FMRP Dysregulates Glutamate Transporter GLT1 and Contributes to Fragile X Syndrome Phenotypes In Vivo*. The Journal of neuroscience : the official journal of the Society for Neuroscience, 2016. **36**(27): p. 7079-7094.
189. Cheng, C., S.K.M. Lau, and L.C. Doering, *Astrocyte-secreted thrombospondin-1 modulates synapse and spine defects in the fragile X mouse model*. Molecular Brain, 2016. **9**(1): p. 74.
190. Voineagu, I., et al., *Transcriptomic analysis of autistic brain reveals convergent molecular pathology*. Nature, 2011. **474**(7351): p. 380-384.
191. Serhan, C.N. and J. Savill, *Resolution of inflammation: the beginning programs the end*. Nat Immunol, 2005. **6**(12): p. 1191-7.
192. Siniscalco, D., et al., *Inflammation and Neuro-Immune Dysregulations in Autism Spectrum Disorders*. Pharmaceuticals (Basel, Switzerland), 2018. **11**(2): p. 56.
193. Estes, M.L. and A.K. McAllister, *Immune mediators in the brain and peripheral tissues in autism spectrum disorder*. Nat Rev Neurosci, 2015. **16**(8): p. 469-86.
194. Cresto, N., et al., *Do Astrocytes Play a Role in Intellectual Disabilities?* Trends in Neurosciences, 2019. **42**(8): p. 518-527.
195. Benavides-Piccione, R., et al., *On dendrites in Down syndrome and DS murine models: a spiny way to learn*. Prog Neurobiol, 2004. **74**(2): p. 111-26.
196. Currele, D.S., et al., *Culture of mouse neural stem cell precursors*. J Vis Exp, 2007(2): p. 152.
197. Estaras, C., et al., *Genome-wide analysis reveals that Smad3 and JMJD3 HDM co-activate the neural developmental program*. Development, 2012. **139**(15): p. 2681-91.
198. Pollard, S.M., et al., *Adherent neural stem (NS) cells from fetal and adult forebrain*. Cereb Cortex, 2006. **16** **Suppl 1**: p. i112-20.
199. Graham, F.L., et al., *Characteristics of a human cell line transformed by DNA from human adenovirus type 5*. J Gen Virol, 1977. **36**(1): p. 59-74.
200. Blanco-Garcia, N., et al., *Autoacetylation regulates P/CAF nuclear localization*. J Biol Chem, 2009. **284**(3): p. 1343-52.
201. Masters, J.R., *HeLa cells 50 years on: the good, the bad and the ugly*. Nat Rev Cancer, 2002. **2**(4): p. 315-9.

202. Antonucci, F., et al., *Microvesicles released from microglia stimulate synaptic activity via enhanced sphingolipid metabolism*. The EMBO journal, 2012. **31**(5): p. 1231-1240.
203. Meerbrey, K.L., et al., *The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo*. Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3665-70.
204. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Anal Biochem, 1976. **72**: p. 248-54.
205. He, F., *Laemmli-SDS-PAGE*. Bio-protocol, 2011. **1**(11): p. e80.
206. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications*. Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4350-4.
207. Ashburner, M., et al., *Gene ontology: tool for the unification of biology*. The Gene Ontology Consortium. Nat Genet, 2000. **25**(1): p. 25-9.
208. Babicki, S., et al., *Heatmapper: web-enabled heat mapping for all*. Nucleic Acids Res, 2016. **44**(W1): p. W147-53.
209. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2*. Nat Methods, 2012. **9**(4): p. 357-9.
210. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS)*. Genome Biol, 2008. **9**(9): p. R137.
211. Yu, G., L.G. Wang, and Q.Y. He, *ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization*. Bioinformatics, 2015. **31**(14): p. 2382-3.
212. Mi, H., et al., *PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools*. Nucleic Acids Res, 2019. **47**(D1): p. D419-d426.
213. Robinson, J.T., et al., *Integrative genomics viewer*. Nat Biotechnol, 2011. **29**(1): p. 24-6.
214. Karolchik, D., et al., *The UCSC Table Browser data retrieval tool*. Nucleic Acids Res, 2004. **32**(Database issue): p. D493-6.
215. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics, 2013. **29**(1): p. 15-21.
216. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data*. Bioinformatics, 2015. **31**(2): p. 166-9.
217. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.

218. Keselman, H.J. and J.C. Rogan, *A Comparison of the Modified-Tukey and Scheffe Methods of Multiple Comparisons for Pairwise Contrasts*. Journal of the American Statistical Association, 1978. **73**(361): p. 47-52.
219. Kerby, D.S., *The Simple Difference Formula: An Approach to Teaching Nonparametric Correlation*. Comprehensive Psychology, 2014. **3**: p. 11.IT.3.1.
220. Zhang, Y., et al., *Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse*. Neuron, 2016. **89**(1): p. 37-53.
221. Zhang, Y., et al., *An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex*. The Journal of Neuroscience, 2014. **34**(36): p. 11929.
222. Sun, Y., et al., *Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms*. Cell, 2001. **104**(3): p. 365-76.
223. Szu, J.I. and D.K. Binder, *The Role of Astrocytic Aquaporin-4 in Synaptic Plasticity and Learning and Memory*. Frontiers in integrative neuroscience, 2016. **10**: p. 8-8.
224. Ge, W., et al., *Notch signaling promotes astroglial gene activation via direct CSL-mediated glial gene activation*. J Neurosci Res, 2002. **69**(6): p. 848-60.
225. Martini, S., et al., *A critical role for Sox9 in notch-induced astroglial gene activation and stem cell maintenance*. Stem Cells, 2013. **31**(4): p. 741-51.
226. Cheng, C., S.K. Lau, and L.C. Doering, *Astrocyte-secreted thrombospondin-1 modulates synapse and spine defects in the fragile X mouse model*. Mol Brain, 2016. **9**(1): p. 74.
227. Hillen, A.E.J., J.P.H. Burbach, and E.M. Hol, *Cell adhesion and matricellular support by astrocytes of the tripartite synapse*. Prog Neurobiol, 2018. **165-167**: p. 66-86.
228. Jori, F.P., et al., *RB and RB2/P130 genes cooperate with extrinsic signals to promote differentiation of rat neural stem cells*. Mol Cell Neurosci, 2007. **34**(3): p. 299-309.
229. Juan, E.A., *Caracterización funcional de PHF8, miembro de la familia Jumonji de demetilinas de histonas*, in PhD thesis. 2011, University of Barcelona: Barcelona.
230. Zambelli, F., G. Pesole, and G. Pavesi, *Pscan: finding over-represented transcription factor binding site motifs in sequences*

- from co-regulated or co-expressed genes*. Nucleic acids research, 2009. **37**(Web Server issue): p. W247-W252.
231. Sun, L., et al., *cMyc-mediated activation of serine biosynthesis pathway is critical for cancer progression under nutrient deprivation conditions*. Cell Research, 2015. **25**(4): p. 429-444.
232. Yatim, A., et al., *NOTCH1 nuclear interactome reveals key regulators of its transcriptional activity and oncogenic function*. Mol Cell, 2012. **48**(3): p. 445-58.
233. Piper, M., et al., *NFIA controls telencephalic progenitor cell differentiation through repression of the Notch effector Hes1*. J Neurosci, 2010. **30**(27): p. 9127-39.
234. Shu, T., et al., *Abnormal development of forebrain midline glia and commissural projections in Nfia knock-out mice*. J Neurosci, 2003. **23**(1): p. 203-12.
235. das Neves, L., et al., *Disruption of the murine nuclear factor I-A gene (Nfia) results in perinatal lethality, hydrocephalus, and agenesis of the corpus callosum*. Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(21): p. 11946-11951.
236. Cebolla, B. and M. Vallejo, *Nuclear factor-I regulates glial fibrillary acidic protein gene expression in astrocytes differentiated from cortical precursor cells*. J Neurochem, 2006. **97**(4): p. 1057-70.
237. Cai, J., et al., *A crucial role for Olig2 in white matter astrocyte development*. Development, 2007. **134**(10): p. 1887.
238. Barnabé-Heider, F., et al., *Origin of new glial cells in intact and injured adult spinal cord*. Cell Stem Cell, 2010. **7**(4): p. 470-82.
239. Griemsmann, S., et al., *Characterization of Panglial Gap Junction Networks in the Thalamus, Neocortex, and Hippocampus Reveals a Unique Population of Glial Cells*. Cereb Cortex, 2015. **25**(10): p. 3420-33.
240. Ohayon, D., et al., *Sulfatase 2 promotes generation of a spinal cord astrocyte subtype that stands out through the expression of Olig2*. Glia, 2019. **67**(8): p. 1478-1495.
241. Tatsumi, K., et al., *Olig2-Lineage Astrocytes: A Distinct Subtype of Astrocytes That Differs from GFAP Astrocytes*. Front Neuroanat, 2018. **12**: p. 8.
242. Nishiyama, A., et al., *Co-localization of NG2 proteoglycan and PDGF α -receptor on O2A progenitor cells in the developing rat brain*. Journal of Neuroscience Research, 1996. **43**(3): p. 299-314.

243. Dawson, M.R., J.M. Levine, and R. Reynolds, *NG2-expressing cells in the central nervous system: are they oligodendroglial progenitors?* J Neurosci Res, 2000. **61**(5): p. 471-9.
244. Stallcup, W.B., *The NG2 proteoglycan: past insights and future prospects.* J Neurocytol, 2002. **31**(6-7): p. 423-35.
245. Nishiyama, A., Z. Yang, and A. Butt, *Astrocytes and NG2-glia: what's in a name?* J Anat, 2005. **207**(6): p. 687-93.
246. Clarke, L.E. and B.A. Barres, *Emerging roles of astrocytes in neural circuit development.* Nature reviews. Neuroscience, 2013. **14**(5): p. 311-321.
247. Liddelow, S. and B. Barres, *SnapShot: Astrocytes in Health and Disease.* Cell, 2015. **162**(5): p. 1170-1170.e1.
248. Loenarz, C., et al., *PHF8, a gene associated with cleft lip/palate and mental retardation, encodes for an Nepsilon-dimethyl lysine demethylase.* Hum Mol Genet, 2010. **19**(2): p. 217-22.
249. Qiao, Y., et al., *Autism-associated familial microdeletion of Xp11.22.* Clin Genet, 2008. **74**(2): p. 134-44.
250. Pappa, S., et al., *PHF2 histone demethylase prevents DNA damage and genome instability by controlling cell cycle progression of neural progenitors.* Proc Natl Acad Sci U S A, 2019. **116**(39): p. 19464-19473.
251. Hwang, I.Y., et al., *Psat1-Dependent Fluctuations in α -Ketoglutarate Affect the Timing of ESC Differentiation.* Cell Metab, 2016. **24**(3): p. 494-501.
252. Yang, M. and K.H. Vousden, *Serine and one-carbon metabolism in cancer.* Nat Rev Cancer, 2016. **16**(10): p. 650-62.
253. Mullarky, E., et al., *A novel small-molecule inhibitor of 3-phosphoglycerate dehydrogenase.* Mol Cell Oncol, 2016. **3**(4): p. e1164280.
254. DeBerardinis, R.J. and T. Cheng, *Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer.* Oncogene, 2010. **29**(3): p. 313-324.
255. Locasale, J.W., *Serine, glycine and one-carbon units: cancer metabolism in full circle.* Nature Reviews Cancer, 2013. **13**(8): p. 572-583.
256. Acuna-Hidalgo, R., et al., *Neu-Laxova syndrome is a heterogeneous metabolic disorder caused by defects in enzymes of the L-serine biosynthesis pathway.* Am J Hum Genet, 2014. **95**(3): p. 285-93.

257. Seltzer, L.E. and A.R. Paciorkowski, *Genetic disorders associated with postnatal microcephaly*. Am J Med Genet C Semin Med Genet, 2014. **166c**(2): p. 140-55.

Appendix

I would like to include the publications in which I participated during my doctoral thesis.

- Iacobucci S, Padilla N, Gabrielli M, Navarro C, Lombardi M, Verderio C, de la Cruz X, Martínez-Balbás MA.
PHF8 histone demethylase regulates astrocytes differentiation and function.
Unpublished
- Pappa S, Padilla N, Iacobucci S, Vicioso-Mantis M, de la Campa E, Navarro C, Marcos E, de la Cruz X, Martínez-Balbás MA.
PHF2 histone demethylase prevents DNA damage and genome instability by controlling cell cycle progression of neural progenitor cells.
PNAS, (2019) 116 (39), 19464–19473
- Fueyo R, Iacobucci S, Pappa S, Estarás C, Lois S, Vicioso-Mantis M, Navarro C, Cruz- Molina S, Reyes JC, Rada-Iglesias Á, de la Cruz X, Martínez-Balbás MA.
Lineage specific transcription factors and epigenetic regulators mediate TGFβ-dependent enhancer activation.
Nucleic Acids Research, (2018) 7, 3351–3365
- Elena Asensio-Juan1, Raquel Fueyo, Stella Pappa, Simona Iacobucci, Carmen Badosa, Sergi Lois, Miriam Balada, Laia Bosch-Presegue, Alex Vaquero, Sara Gutierrez, Carme Caelles, Carme Gallego, Xavier de la Cruz, and Marian A. Martínez-Balbás.
The histone demethylase PHF8 is a molecular safeguard of the IFN gamma response. Nucleic Acids Research, (2017) 45, 3800-3811.