UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in Biologia molecolare e cellulare

XXXIV Ciclo

# hES cells to study early neurodevelopment in Huntington's disease

## Silvia Brocchetti

PhD Thesis

**Scientific tutor: Prof. Elena Cattaneo**

**Coordinator of the PhD School: Prof. Martin Kater**

Academic year: 2020-2021

SSD: BIO/14

Thesis performed at the Laboratory of Stem Cell Biology and Pharmacology of Neurodegenerative Diseases. Department of Biosciences, University of Milan and Istituto Nazionale di Genetica Molecolare.

# Table of Contents

# Abstract

Huntington Disease (HD) is a genetic neurodegenerative disease caused by a CAG expansion in the gene encoding for the Huntingtin (HTT) protein. The progressive death and atrophy of specific striatal GABAergic projection neurons named Medium Spiny Neurons (MSNs) lead to motor, cognitive and psychiatric dysfunctions in HD patients. Several studies are pointing to the crucial functions of HTT during early development, advocating to a developmental component of HD. Based on these observations, we employed the isogenic series of HD human embryonic stem cells (RUES2) to study the effects of mutant HTT on neuronal differentiation in a fixed genetic background.

Firstly, by exposing the control RUES2 cell line to ventral telencephalic differentiation we observed the propensity of the cells to differentiate toward the medial ganglionic eminence (MGE) compared to other human embryonic stem cell (hESC) lines. Moreover, HD RUES2 displayed aberrant MGE cell fate acquisition. Also, these cells recapitulate known HD phenotypes, such as reduced expression of BDNF and NEUROD1, and increased cleavage of N-cadherin. In particular, HD lines exhibit a defect in the transition from pluripotency toward neuroectodermal fate as documented by the persistency of the pluripotent marker OCT4 and by reduced upregulation of the neuroectodermal marker PAX6.

Considering the fundamental function of Polycomb group proteins (PcGs) in regulating cell fate identity and differentiation and the crucial role of epigenetics in HD pathogenesis, we started investigating whether these complexes were affected in our cell system. Differences in the number and size of the three-dimensional PcG foci structures emerged between control and HD lines, suggesting a potential link between HTT and these epigenetic complexes. Moreover, level of H3K9me3 histone modification was affected in HD lines during the differentiation. Overall, these preliminary data suggest that the presence of the mutation in the HD gene causes alterations in the regulation of both constitutive and facultative heterochromatin already in pluripotency. This hypothesis will be further tested through experiments of ChIP-seq analysis, in which we expect to identify target genes whose expression may be dysregulated in the early stages of neural development.

# Abstract (Italian)

La Corea di Huntington (MH) è una malattia neurodegenerativa genetica causata dall'espansione del tratto CAG nel gene codificante la proteina Huntingtina (HTT). La morte progressiva di specifici neuroni GABAergici striatali, chiamati Medium Spiny Neurons (MSNs) è responsabile delle disfunzioni motorie, cognitive e psichiatriche dei pazienti MH. Numerosi studi stanno evidenziando le cruciali funzioni dell'HTT durante le prime fasi dello sviluppo, sostenendo la presenza di una componente evolutiva della MH.

Abbiamo quindi impiegato la serie isogenica di linee hES RUES2 HD per studiare gli effetti di diverse lunghezze CAG patologiche in un background genetico costante.

Innanzitutto, esponendo la linea RUES2 al differenziamento telencefalico ventrale abbiamo osservato la propensione a differenziare verso l'eminenza gangliare mediale (MGE). Inoltre, la linea HD risulta incapace di acquisire un destino MGE. Queste linee inoltre ricapitolano fenotipi HD noti, come la ridotta espressione di BDNF e NEUROD1 e una maggiore scissione della proteina sinaptica N-cadherina. In particolare, è risultato alterato il passaggio dalla pluripotenza al destino neuroectodermico; poiché, le linee HD mostrano una persistente espressione di OCT4 insieme a ridotti livelli del marker neuroectodermico PAX6.

Considerando il ruolo fondamentale delle proteine dei gruppi Polycomb (PcG) nella regolazione dell'identità cellulare e del differenziamento, abbiamo caratterizzato questi complessi nel nostro sistema cellulare. Sono emerse differenze nel numero e nella dimensione delle strutture tridimensionali dei PcG tra la linea controllo e HD, suggerendo un potenziale legame tra la proteina HTT e questi complessi epigenetici. Inoltre, le linee HD RUES2 presentano un'alterazione anche a livello della modificazione istonica H3K9me3 durante il differenziamento. In conclusione, questi dati preliminari hanno evidenziato modificazioni nella regolazione dell'eterocromatina costitutiva e facoltativa, suggerendo un possibile coinvolgimento della proteina HTT nella regolazione epigenetica. Questa ipotesi sarà ulteriormente verificata mediante analisi ChIP-seq, per identificare geni bersaglio la cui espressione potrebbe essere alterata durante le prime fasi dello sviluppo neurale.

# Aim

The crucial role exerted by HTT during the early stages of fetal development has been ascertained by several studies. Is still matter of debate whether muHTT could alter early neurodevelopmental programs and therefore give rise to neuronal populations susceptible in adulthood to environmental stressors and toxins.

This PhD project aimed at identifying the molecular defects occurring during early neural cell fate specification produced by the HD mutation, which could later impact on disease onset. For this purpose, we employed a collection of isogenic hES RUES2 HD lines which constitutes a unique genetic toolkit, since it allows to study the effects of different pathologic CAG lengths in a fixed genetic background with a direct correlation between genotype and phenotype.

The aims of this thesis were to:

I. to test the effect of the mutation on the differentiation of hES RUES2 lines exposed to a striatal differentiation protocol mimicking human ventral telencephalon development to identify pathogenic phenotypes;

II. to elucidate the impact of muHTT on cell fate specification by exploring the first steps in the differentiation protocol;

III. to study the effect of muHTT on chromatin structure and gene expression during differentiation and in pluripotency. In particular, we studied the Polycomb group of proteins (PcGs), as these are responsible for the deposition of the repressive histone mark H3K27me3, and of the counterpart H3K4me3 generated by the Trithorax MLL2/COMPASS complex. In addition to these complexes controlling facultative heterochromatin, we also investigated the constitutive heterochromatin regulation mediated by H3K9me3.

# 1 | Introduction

## 1.1 Huntington's Disease

Huntington's Disease (HD) is a dominant inherited neurodegenerative disorder caused by a mutation in the CAG tract of the gene encoding the Huntingtin (HTT) protein. The prevalence of HD is of 10 to 12 cases per 100.000 in the Caucasian population and is characterized by midlife onset. The progressive death and atrophy of specific GABAergic projection neurons in the striatum from caudate nucleus and putamen of the basal ganglia, the Medium Spiny Neurons (MSNs), are responsible for symptoms as personality disturbance, choreiform movements and cognitive decline typical of HD [1,2]. This neuronal degeneration provokes a reduction of the striatum volume together with thinning of the cortex [3,4].

An abnormal expansion of the CAG repeats within the coding region of HTT protein was attributed in 1993 as the genetic cause of HD. HTT is a large protein of 3,144 amino acids characterized by a very well-studied N-terminal region containing the expandable polyglutamine (polyQ) tract. Healthy individuals carry a polyQ tract ranging from 10 to 35 repeats. The expansion of the number of CAG to 35 or more is responsible for the generation of the mutant form of Huntingtin (muHTT) resulting in HD. Patients with 36 to 39 repeats have a reduced penetrance and later onset of the disorder, whereas a length exceeding 40 CAG causes full penetrance HD. Generally, there is a direct relationship between the age of onset of the disease and the length of the CAG expansion, as illustrated in the Figure 1.1 below [5].
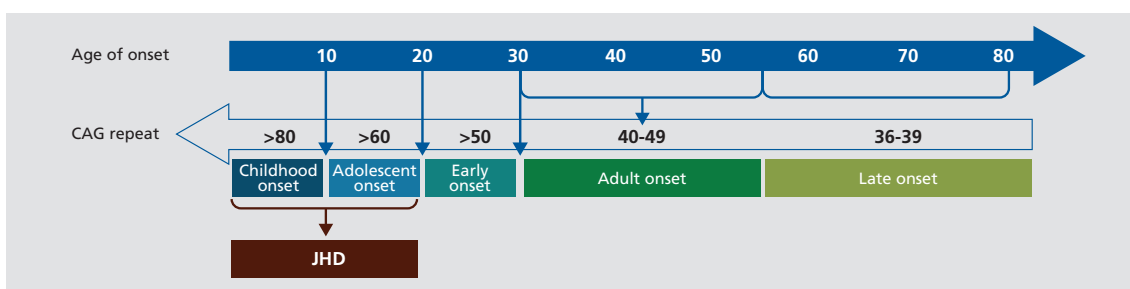


*Figure 1.1. Relationship between the length of the CAG tract and the age of onset of Huntington's disease. Figure from Nopoulos et al., (2016). JHD: Juvenile Huntington's disease.*

HTT is broadly expressed in different human tissues and it is particularly enriched in the nervous system. Here, HTT plays multiple functions, including the regulation of molecular and cellular mechanisms occurring during CNS development, as neurogenesis, synaptic development and identity, together with the maintenance of neural function in adults [6].

In the last decade, increasing number of studies performed on HD cellular systems, mouse models and HD patients highlighted the possibility that the HTT mutation may cause dysfunctions already during neurodevelopment [7–14]. This hypothesis is supported by the presence of deregulated neurodevelopmental genes and neuronal degeneration long time before the onset of HD symptoms, both in humans and mice models [15–20]. This evidence raises the issue of whether these neurodevelopmental abnormalities originate the premanifest symptoms in HD. The researchers found that the children with risk of developing HD present smaller head size, denoting brain growth deficit [22]. Moreover, prodromal HD children present also lower body index (BMI), whereas juvenile HD is characterized by body weight deficit implying developmental anomalies caused by mitochondrial dysfunction [22]. Furthermore, neuropathological analyses of brains from prodromal HD patients, proved the presence of signs of developmental pathology [23]. Besides, brain alterations were observed by neuroimaging studies decades prior disease [24–27]. The hypothesis of an aberrant development in HD has also been recently sustained by a study that compares tissue fragments from control and HD human fetal brain at different gestational weeks, and in which the authors identified abnormal developmental features [28].

## 1.2 Huntingtin and development

In addition to the numerous functions of HTT in the adult brain, HTT is crucial also during embryogenesis and development [30]. In particular, the absence of HTT causes severe defects during gastrulation leading to embryonic death in mice around E7.5 - 8.5 [11,12,30]. This embryonic lethality may be rescued by expressing the wild-type HTT at the 10-15% of its normal level. Even after restoring normal HTT expression level, epiblast defects, faulty neural tube formation and progressive cortical and striatal architectural abnormalities occurred later in life [31,32].

Furthermore, in Molero et al. (2016) study by generating conditional knock-in mouse model of HD expressing muHTT exclusively during embryogenesis, up to post-natal day 21, were recapitulated typical biochemical and behavioral features of HD as impairments in pluripotency and stem-cell mediated striatal neurogenesis genes [10,13,33].

Additional proof of the role played by HTT during early neural development was reported in Nguyen et al. work (2013). The authors found that muHTT impairs development of the striatum at multiple stages, such as during neural induction and early neurogenesis. Nevertheless, whether the expanded polyQ mutation confers a gain or loss of function to the normal HTT protein remains still a matter of debate.

In line with this study, additional evidences showed that loss of HTT during the differentiation of embryonic stem cells (ES) impairs proper neuroepithelial rosette structure formation, a cellular organization recapitulating *in vitro* neural tube development [35–37]. Furthermore, several works on human induced pluripotent stem cells (iPSCs) models have supported the neurodevelopmental features observed in HD. HD iPSCs exhibited disease phenotype along differentiation and dysregulated transcription of neuronal development genes, as alterations in striatal-specific and TGF-β signaling genes [38], and defective neural progenitors' maturation [39–42].

In particular, the HD iPSC Consortium (2012) observed by differentiating neurons *in vitro* from iPSCs derived from HD patients that cells bearing muHTT presented an altered developmental speed, metabolism, electrophysiological properties, cell adhesion and, high cell death [8].

Furthermore, altered corticogenesis was observed using HD-iPSCs differentiation since the cells reported altered transcriptomic profiles, morphology and delayed electrophysiological maturation [42,43].

The essential function of HTT during cortical neurogenesis was observed through the alteration of mitotic spindle orientation of cortical progenitors after RNAi-mediated silencing of Huntingtin. The altered gene expression of HTT, resulted in cell fate impairment in the ventricular zone of mouse embryos [7]. Moreover, HTT is important also in the regulation of neuronal cell migration, as for neuroblasts movement from the ventricular zone to the cortical plate. The knockdown of HTT in

neuroepithelial cells of neocortex of mouse embryo led to alteration in neuroblasts migration, reduction in cell proliferation and an increase of cell mortality during the specific time window of early development [44]. HTT silencing at later stage of development caused less severe effects suggesting that the specific HTT expression timing is a fundamental variable in the pathogenesis. Finally, HTT exerts its control also on cortical and striatal neurons specification during development, as in absence of these proteins the neurons are unable to correctly migrate and colonize these cerebral areas [46].

## 1.3 Epigenetics of Huntington's disease

An increasing number of studies are suggesting that altered epigenetic modifications, including a vast group of heritable molecular modifications of DNA or associated proteins as noncoding RNAs regulation, might contribute to HD pathogenesis (Figure 1.2). Also, the onset and progression of HD in addition to the genetic defects are believed to be dependent on environmental factors. In the last decades, raised up the hypothesis that altered transcriptional and chromatin modulation might significantly contribute to Huntington's disease pathogenesis.

Although HTT is mainly localized in the cytoplasm, this protein is present also in the nucleus where it interacts with multiple transcription factors (TFs), multi-subunit complexes and the DNA itself [46–48]. Therefore, dysregulated interactions of muHTT with TFs and proteins of the epigenetic machinery could impact on transcriptional alterations frequently observed in HD and consequently affect mechanisms and pathways fundamental for proper neuronal physiology.
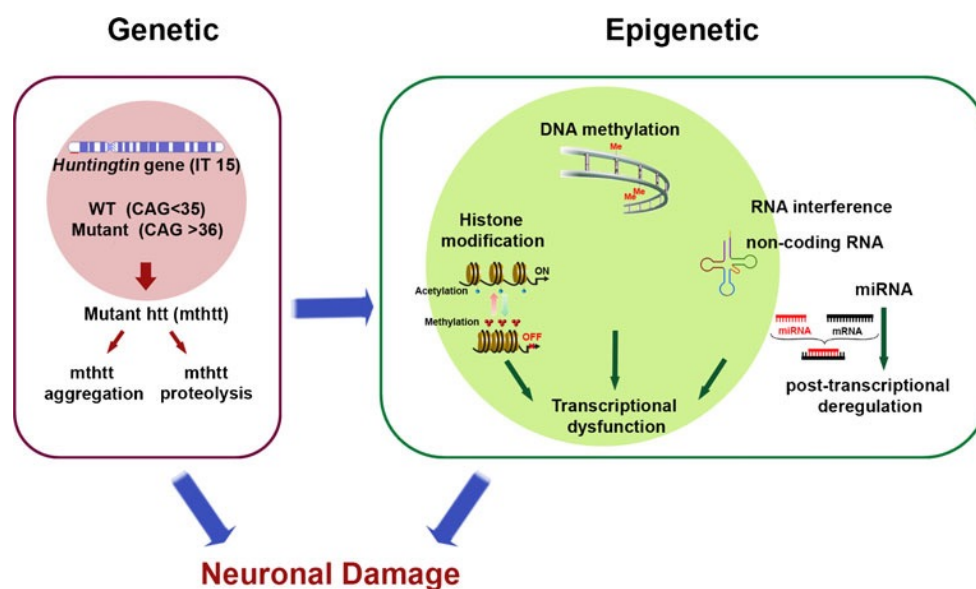


*Figure 1.2. Relationship between the length of the CAG tract and the age of onset of Huntington's disease. The methylation state of neuronal genes promoter is altered in HD. Moreover, the transcriptional alteration in HD might be associated to post-transcriptional modifications of histone and abnormal nucleosome dynamics. Altered activity of noncoding RNAs and microRNAs can deregulate gene expression at transcriptional and post-transcriptional levels. Significant epigenetic alterations might be caused by both aggregates and fragments of muHTT, leading to synaptic and neuronal damage [50].*

### 1.3.1 PRC2 and Trithorax complexes in HD

Several studies sustain the role of wild-type and muHTT on histone methylation [47,49,50], (Figure 1.3). HTT is a well-known facilitator of the activity of the multi-subunit complex PRC2 in the nucleus, as reported in Seong et al. (2009). This study reported that both PRC2 subunits Ezh2 and Suz12 co-immunoprecipitated with HTT in wild-type mouse ES cells (mESCs). In addition, they observed an

increase of PRC2 activity in parallel to the expansion of the polyQ tract with a consequent enrichment in H3K27me3. Conversely in the absence of HTT the activity of PRC2 significantly decreased, therefore leading to a significant reduction of H3K27me3.
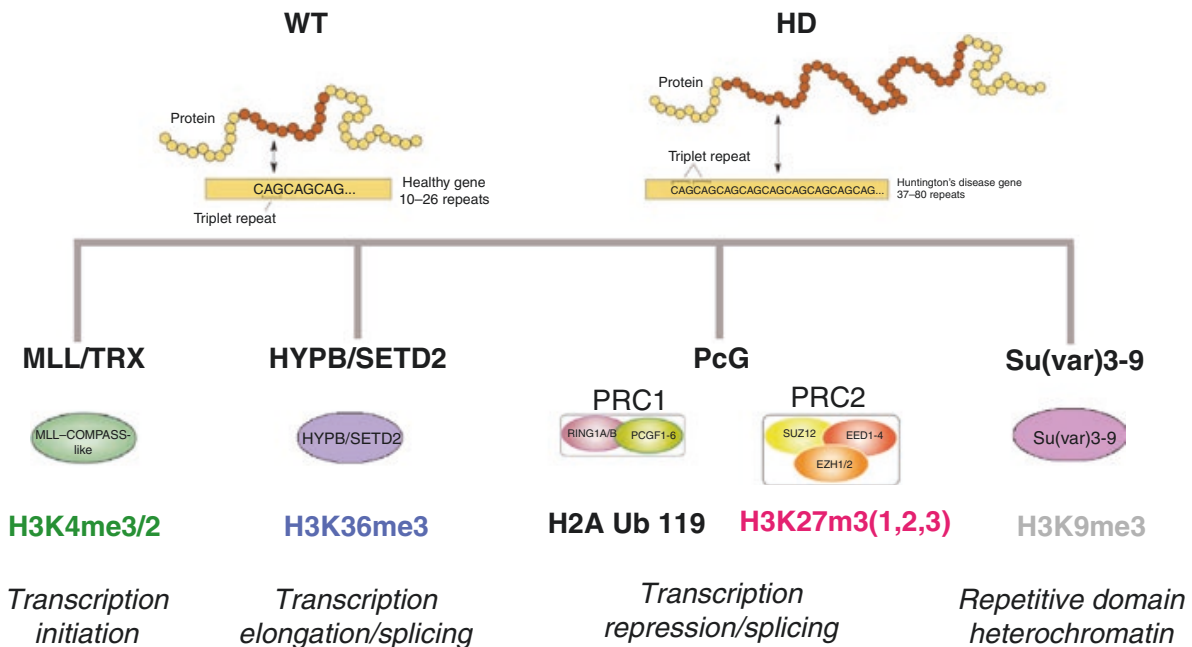


*Figure 1.3. Schematic representation of the chromatin remodeling complexes that are directly or indirectly regulated by wild-type or mutant huntingtin. (1) MLL/TRX complex di- or tri- methylates specifically the H3K4 acting mainly on transcription initiation; (2) the H3K36 methyltransferase HYPB/ SETD2 regulates transcriptional elongation and splicing; (3) the Polycomb group proteins (PcGs) are divided into the two main complexes PRC1, that ubiquitinates the histone H2A, and PRC2 that mono-, di- or tri- methylates the H3K27 residues, together they regulate transcriptional repression and splicing regulation; (4) the H3K9 is trimethylated by the Su(var)3-9 complex hence regulating repetitive domain of heterochromatin. Figure modified from Barlett et al., (2017).*

This relationship among HTT and PRC2 was confirmed also by another study published by Biagioli et al. (2015), where alterations of genome-wide pattern of H3K27me3 were observed in mESCs expressing muHTT. In this work, the data collected demonstrated that HTT is required in pluripotent mESCs for the correct deposition of H3K27me3 marks at a subgroup of bivalent domains (BD). On the contrary, BDs in neural progenitors are fundamental for the proper maintenance and appropriate removal of the H3K27me3 mark. Differentiating several mouse knock-in lines and KO cells into EBs and performing ChIP-seq analysis, they found a reduction in the deposition of the H3K27me3 at bivalent loci in absence of HTT, conversely in presence of muHTT an enrichment of H3K4me3 generated by the Trithorax complex. Therefore, in presence of muHTT, as PRC2 is a direct target of the HTT protein [49], Polycomb activity might be altered and consequently the transcriptional regulation as well.

Another interesting evidence regarding the regulation of PRC2 in adult neurons is provided by Von Schimmelmann and colleagues (2016). They observed in MSNs deficient for PRC2 an upregulation

of PRC2 target genes encoding for death-promoting proteins and transcription factors, that are normally suppressed in these cells. Most of these genes are under the control of bivalent chromatin regions, whose role, which might be beneficial, in fully mature neurons remains still to be elucidated. This work suggests that although most H3K27me3 target genes in MSNs are insensitive to the deficiency of PRC2, the transcriptional silencing of the bivalent genes in adult neurons is selectively controlled by PRC2. Therefore, loss of normal PRC2 regulation in MSNs may lead to fatal and progressive neurodegeneration caused by the upregulation of bivalent PRC2 target genes.

Nevertheless, the presence of the expanded polyQ alters not only PRC2 activity, and consequently the H3K27me3 levels, but also the H3K4me3 enrichment [47].

Fundamentally, lowering of H3K4me3 levels leads to a reduced gene transcription of target genes, as observed *in vivo* in human HD *post-mortem* brains, and in both cortex and striatum of R6/2 mice [50]. These phenotypes sustain that the effects of muHTT on the regulation of chromatin conformation might contribute to HD pathogenesis. Similarly, Dong et al. (2015) identified hundreds of genes with different H3K4me3 levels in human HD *post-mortem* brains. Of note, they found enrichment in HD-distal peaks of H3K4me3 corresponding to genes for chromatin remodelers and transcription factors, as EZH2 and SUZ12, which are the core subunits of the PRC2 complex. Hence, transcriptional dysregulation in HD might not be dependent only on alterations in PRC2 pattern but also on H3K4me3 enrichment.

In a very recent work, Pear et al. group (2020) investigated whether the direct binding of HTT or muHTT on specific locations on chromatin might contribute to the transcriptional dysregulation in HD. By chromatin immunoprecipitation and ChIP-seq experiments they observed occupancy of HTT in the mouse genome at specific locations, comprising thousands of loci that are differentially occupied in striatal tissue from a knock-in mouse model of HD (B6. *HttQ$^{111/+}$)* compared to wild type controls. In particular, they observed in *HttQ$^{111/+}$* mice increased occupancy of HTT peaks near genes down-regulated in HD, and depletion close to up-regulated genes, supporting the hypothesis of a possible relation among transcriptional dysregulation observed in HD and HTT-chromatin interactions. Moreover, they observed in the loci with increased HTT occupancy a positive correlation with H3K27me3 enrichment and a negative correlation with H3K4me3 levels. This evidence supports the hypothesis that alterations of normal HTT-chromatin regulatory interactions lead to transcriptional perturbations in HD responsible for both up- and down- regulation of certain genes. Additionally, they detected a strong enrichment of HTT peaks in H3K9me3-positive regions of constitutive heterochromatin. This interaction is in line with what already reported in literature [54].

HTT regulates in hESCs the trimethylation of H3K9 through the binding with the ATF7IP-SETDB1 complex. This binding prevents the interaction with other heterochromatin regulators and

transcriptional repressors, thus keeping low levels of H3K9me3 in pluripotency [55]. This regulation is perturbed in presence of the polyglutamine expansion, whereas loss of HTT induces global increase in H3K9me3 enrichment at genetic loci for transcriptional regulators of neuronal differentiation. The first evidence correlating H3K9me3 with HD came from Ryu et al. (2006) study, where increased levels of H3K9me3 and the H3K9 methyl-transferase – a ERG-associated protein carrying SET domain (ESET) – were observed in the neocortex and striatum of HD patients and also in the striatum of R6/2 HD mice. Moreover, reducing the trimethylation of Lys9 by pharmacological treatment, the authors observed an improved motor performance, delayed neuronal atrophy, extended lifespan, and increased body weight in the R6/2 mice.

According to these data, Lee et al. (2013) confirmed a direct correlation among alterations in gene transcription in HD and chromatin remodeling mediated by H3K9me3. They performed H3K9me3-ChIP genome-wide sequencing together with RNA-seq to identify genes deregulated by trimethylated histone H3K9 (H3K9me3)- dependent heterochromatin in stable striatal HD cell lines (STHdhQ7/7 and STHdhQ111/111 cells). They identified that high enrichment in HD striatal cells of H3K9me3-dependent chromatin condensation led to deregulation of cellular mobility, neuronal differentiation pathways and neuronal synaptic transmission, as the cholinergic receptor M1 (CHRM1). In addition, muHTT affects the CHRM1 gene expression, a key muscarinic receptor transducing $Ca^{2+}$ signaling, leading to synaptic failure and striatal dysfunctions typical of HD.

Further evidence linking HTT with the regulation of the H3K9me3 mark was provided by Irmak et al. (2018) study. They found that HTT binds with ATF7IP, a regulator of SETDB1 that promotes the trimethylation of H3K9. HTT binding to the ATF7IP- SETDB1 prevents H3K9me3 enrichment and consequently maintains lower level of this modifications in hESCs. In absence of HTT, the ATF7IP nuclear localization is altered and the interaction with other epigenetic regulators and transcriptional repressors is favored, thus leading to H3K9me3 enrichment in neuronal differentiation genes (e.g., ASCL2, GBX1). Although these genes are correctly expressed at low levels in hESCs, this altered epigenetic conformation results in defective neural induction during differentiation. In parallel, polyQ expansion impairs the interaction between muHTT and ATF7IP and elicits increased H3K9 trimethylation in HD-iPSCs. Equally, knockdown of ATF7IP ends in H3K9me3 decline and amelioration of gene expression alterations in neurons derived from HD-iPSCs. Taken together, all these data highlight that different chromatin remodeling complexes and histone-modifying enzymes are not regulated as single entities but are part of super molecular complexes that control transcription through the coordination of their repressive or activating functions [58]. Therefore, based on the data already present in literature, the HTT protein seems to take part on this finely tuned regulation.

# 1.4 LGE development and *in vitro* striatal differentiation

The embryonic telencephalon consists in the dorsal telencephalon (pallium), that generates the neocortex, and the ventral telencephalon (subpallium), where originate the striatum, the olfactory bulb, the globus pallidus and the population of GABAergic interneurons that migrate toward the cortex [59]. The rapid migration of postmitotic neurons to the subpallium leads to the formation of three intraventricular bulges forming the septum, the medial and lateral ganglionic eminences (MGE/LGE), together referred as the whole ganglionic eminence (WGE). From the most ventral eminence in the MGE originate the amygdaloid body and the globus pallidus, whereas the most dorsal LGE gives rise to the caudate and putamen, that constitute the striatum [66, 67]. In the neural epithelium surrounding the telencephalon are present two proliferative areas where are generated the neural precursors, the ventricular zone (VZ), next to the lateral ventricles, and the subventricular zone (SVZ) which spans from the basal region of the VZ [62]. From the proliferative zones of the ventral telencephalon originate and migrate the projection neurons (as the MSNs) and the GABAergic interneurons that populate the adult brain. In particular, the striatal projection neurons originated in the LGE constitute the 90% of the total neurons of this area. The interneurons born in the LGE populate then the cortex, the olfactory bulb and the striatum, while MGE-derived neurons travel toward the cortex, the globus pallidus and the striatum [68, 69].

The antero-posterior and rostro-caudal fate specification of the forebrain is reached through the spatial and temporal expression of different cooperating pathways, as it is orchestrated by Sonic hedgehog (SHH), Retinoic Acid, WNT, BMP and TGFβ signaling pathways. Building on knowledge of striatal development, many scientists have worked on MSNs neural differentiation protocols.

The differentiation of pluripotent stem cells (iPSCs and ESCs) into MSNs consists in three main steps that resemble *in vitro* striatum development, which are defined as neural induction, regional patterning, and terminal differentiation.

Although there is a wide variety of protocols available, the protocol developed in Cattaneo's lab enables the differentiation of hESCs and iPSCs to generate authentic telencephalic progenitors and finally MSNs [64]. This protocol was developed on the basis of the *in vivo* analyses performed on human striatum development [71], which examined the specific spatiotemporal markers transition (Fig.1.4). This spatiotemporal expression map showed that the VZ progenitors expressed OTX2, Forkhead Box G1 (FOXG1), GS homeobox 2 (GSX2) and Achaete-scute homolog 1 (ASCL1), while the SVZ was defined by ASCL1, Insulin gene enhancer binding protein Isl-1 (ISL1) and EBF Transcription Factor 1 (EBF1). This study revealed a unique signature characteristic of early striatal neurons at 8 – 11 weeks, indeed they co-expresses ISL1, Forkhead box protein P1 (FOXP1), Forkhead box protein P2 (FOXP2), B-cell lymphoma/leukemia 11B (BCL11B, alias CTIP2), EBF1, GABA, Dopamine, and

cAMP-regulated phosphoprotein (DARPP-32) and interestingly Homeobox protein Nkx-2.1 (NKX2.1). These data provided the basis for the development of an *in vitro* striatal differentiation protocol and the characterization of neural progenitors' maturation into authentic MSNs [64]. Notably, a new recent study from Cattaneo's laboratory has deepened by scRNA-seq the characterization of the different cellular states that populate the LGE region by defining a specific network of regulatory genes and transcription factors [66].
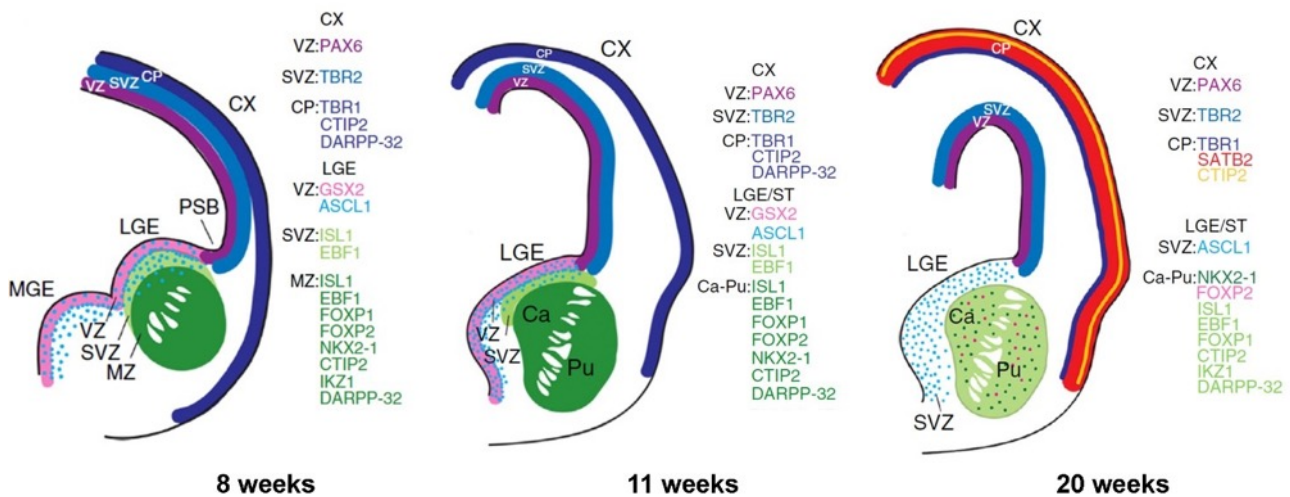


*Figure 1.4. Expression map along early human fetal development, at 8, 11 and 20 weeks of gestation. LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence; CX, neocortex; CP, cortical plate; PSB, pallial-subpallial boundary; VZ, ventricular zone; SVZ subventricular zone; MZ, mantle zone; Ca, caudate; Pu, putamen. Figure modified from Onorati et al., (2014).*

This work generated a new single-cell atlas comprising both protein-coding transcripts and long intergenic noncoding RNAs (lincRNAs) along early human fetal development. This scRNA-seq analysis supported the hypothesis that D1- and D2-MSNs derive from a common pre-MSN precursor and that the different gradient of expression of specific MSN markers as SIX3, SP9 and ISL1 define afterwards their D1- or D2- identity [66]. Moreover, a new set of TFs that was not previously associated with LGE development have been identified, including OTX1 in apical progenitors, VAX1 for basal progenitors, POU2F2 for both D1- and D2- MSNs, NANOG in D2-MSNs, and finally FOXO1 for D1-MSNs [66].

### 1.4.1 Neural induction from human pluripotent stem cells

The development of the nervous system in vertebrates is triggered by a series of signals emitted by an "organizing" region of the early embryo during gastrulation. This phenomenon, defined as neural induction, was discovered for the first time by the embryologists Hilde Mangold and Hans Spemann (1924) who worked with amphibian embryos. Analyzing the molecular bases underlying neural induction, it has been possible to direct neural fate acquisition by suppressing BMP and TGFβ signaling in the ectoderm (Fig.1.5 A). This finding has coined the notion of the so-called "default"

mode of neural induction [73]. Recent studies have demonstrated that a highly efficient neural induction from both human ESCs and iPSCs can be achieved by pharmacological blockage BMP/TGFβ signaling [74]. This method is named "dual SMAD inhibition" and induces a rapid and complete neural conversion by using two inhibitory factors: Noggin antagonizes BMP signaling, whereas SB431542 mainly acts as a suppressor of TGFβ signaling.
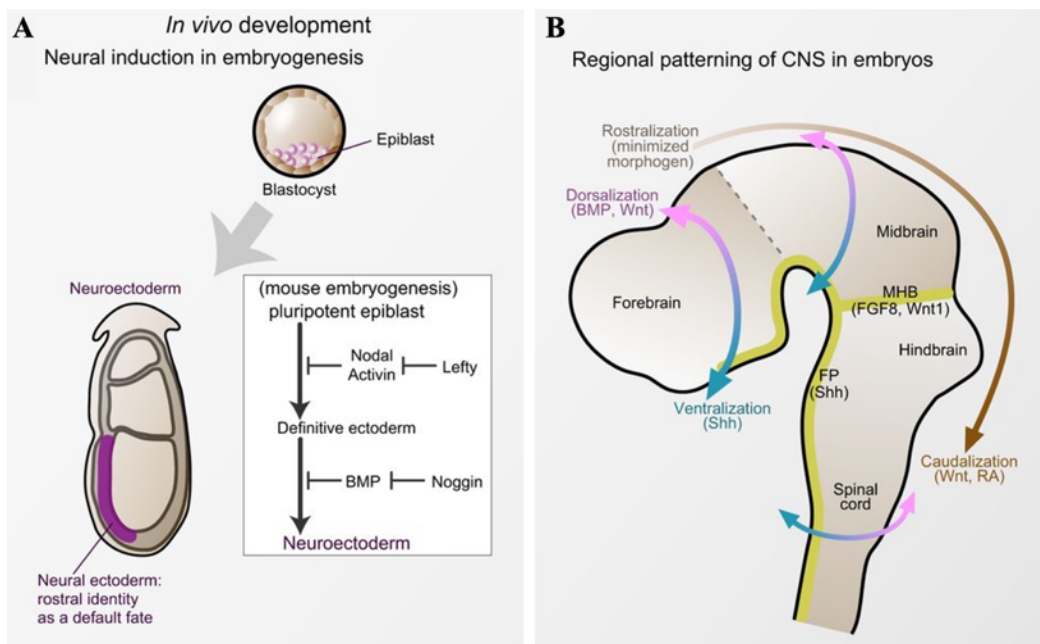


**Figure 1.5. Model of neural induction and regional patterning. A.** *Neural induction is defined a "default" pathway during differentiation, begins in vivo with the inhibition of BMP/Nodal signaling and generate neural cells with rostral identity.* **B.** *The regional patterning in the central nervous system is determined along the rostro-caudal and dorso-ventral axes by the action of morphogens derived from different organizing centers (depicted in yellow), the floor plate (FP) and the midbrain-hindbrain boundary (MHB). Figure modified from Suzuki and Vanderhaeghen (2015).*

### 1.4.2 Patterning and specification

The second step of striatal differentiation after neural induction is the so-called "patterning and specification". In this phase the cells are pushed toward the acquisition of an antero-ventral telencephalic identity through the synergistic activity of the two morphogens Sonic Hedgehog (SHH) and Dickkopf-related protein 1 (DKK1) Figure 1.5. [75]. Sonic hedgehog (Shh) ligand *in vivo*, once secreted by the notochord, generates a gradient along the ventral tube inducing different ventral cell identities through a concentration-dependent mechanism [76]. Therefore, the acquisition of a ventral telencephalic fate requires SHH, whereas low levels of this morphogen favor the transition towards dorsal fate. DKK1 is an inhibitor of the WNT/β-catenin pathway co-receptor LRP6 [71], not enabling the signal transduction through the interaction with the WNT receptor Frizzled. In this way, DKK1 action prevents the expression of dorsal telencephalic genes and moreover, facilitates SHH action by downregulating the expression of GLI3, a downstream effector of SHH [75]. The combination of the

two morphogens SHH and DKK-1 resembles the *in vivo* specification of neuroectodermal cultures into premature telencephalic progenitors, which are characterized by the expression of FOXG1. The differentiation of LGE-derived progenitors in the basal ganglia into GABAergic neurons requires the production of retinoic acid (RA) in the embryonic forebrain [72]. For this reason, during the patterning and specification phase is important the introduction of RA into the medium.

## 1.4.3 Terminal differentiation

The last step of differentiation consists in the commitment of striatal progenitors into fully functional mature MSNs. This purpose is achieved by adding into the medium the Brain-derived neurotropic factor (BDNF) to allow neuronal survival and maturation. At this stage there is a progressive expression of MSNs markers as FOXP1, FOXP2 and finally are generated $DAPP32^+/CTIP2^+$ striatal neurons. These cells are physiologically active, as observed *in vivo* in humans [71,72], and 50% of this population express the Microtubule-associated protein 2 (MAP2). As a result of the ventral specification of RA, 12% of the $MAP2^+$ neurons co-express the Glutamate decarboxylase (GAD65/67), DARPP32 and CTIP2, indicating the correct achievement of MSNs in culture.

# 1.5 *In vitro* models of Huntington disease

Investigations into HD molecular pathology have primarily relied on murine tissues [73]; nevertheless, these models only partially mimic or recapitulate human disease pathophysiology [74].

With the introduction of the preimplantation genetic diagnosis (PDG) was possible the establishment of human ES cells with genetic disorders, as Huntington's disease [75]. Human ESCs with their unlimited capacity for self-renewal and pluripotency represent a scalable source of specialized cells usable in different biomedical applications like disease modeling, pharmacological screening, toxicity testing and cell therapies.

### 1.5.1 RUES2, a new engineered allelic series of isogenic hESCs

The advent of CRISPR/Cas9 genome editing favor the engineering of an isogenic panel of hESC lines carrying different CAG tracts [76] (Figure 1.6). This new cellular model recapitulated some previously reported neuronal phenotypes and revealed chromosomal instability during forebrain neurogenesis in a CAG length-dependent manner [76]. These cell lines, named Rockefeller University Embryonic Stem Cell Line 2 (RUES2) [77,78] were engineered from the parental line through the introduction of a targeted gene mutation combining CRISPR/Cas9 [79–81] with a PiggyBac transposition [82].
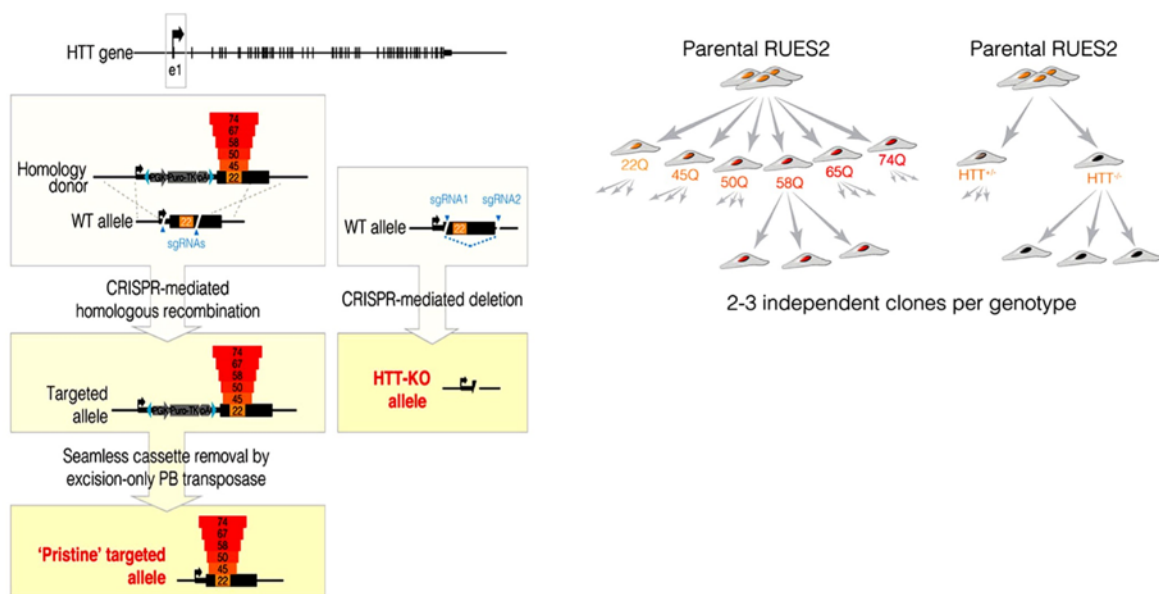


***Figure 1.6. Generation of the isogenic allelic series of RUES2 lines***. *Genome-editing strategy to generate the isogenic hESC clones with increasing length of the CAG tract in exon 1 by CRISPR/Cas9- targeted homologous recombination and a selection cassette flanked by ePiggyBac (ePB) terminal repeats. By transfecting an excision-only piggyBac transposase plasmid, the ePB selection cassette was removed to generate a "pristine" polyQ expanded allele. Figure modified from Ruzo et al., (2018).*

In one allele they introduced five different polyQ tracts (Q45, 50, 58, 67 and 74) to generate the RUES2-Q45, RUES2-Q50, RUES2-Q58, RUES2-Q67 and RUES2-Q74 and therefore model a broad spectrum of CAG expansion that can be observed in HD patients. Using the same editing approach, a sixth isogenic RUES2-Q22, was generated as a control line reinserting the normal Q22tract to verify that the editing procedure did not influence any readout.

The isogenic HD-hESC lines constitute a powerful genetic tool to study HD as it provides a direct correlation between the mutation and the phenotypes in a fixed genetic background.

# 1.6 From pluripotency toward differentiation

The complex network regulating pluripotency and self-renewal of ESCs depends mainly on OCT4 and SOX2. These proteins together regulate the expression of many genes essential for the maintenance of the primitive state and their own expression too. Several target genes are controlled by the synergic activity of these two factors that present a juxtaposition of OCT4 and SOX2 binding element [83]. A third important transcription factor that maintains the pluripotent state and controls the proliferation of undifferentiated ES cells preventing their differentiation is NANOG [98,99].

These three proteins constitute a fundamental transcriptional network responsible for the maintenance of pluripotency and self-renewal of the stem cells. Several genes promoters are co-occupied and regulated by OCT4, SOX2 and NANOG and some of these encode for homeodomain TFs important for regional patterning, development and regulation of the main signaling pathways regulating ESCs [86].

### 1.6.1 OCT4: a master regulator of pluripotency and early embryonic development

The transcription factor OCT4, also known as OCT3 or OCT3/4, represents the master regulator of pluripotency and of several mechanisms crucial in early stages of development. OCT4 is encoded by the POU5F1 gene and belongs to the POU family of proteins, characterized by the presence of two DNA-binding transactivation domains linked by a variable region [87,88]. The POU5F1 gene can generate through the alternative splicing of the intron 1 three different mRNA isoforms, OCT4A, OCT4B and OCT4B1 [89] (Figure 1.7). The OCT4A protein is generated from the OCT4A mRNA; the OCT4B mRNA is translated into the three different protein isoforms OCT4B-265, OCT4B-190, and OCT4B-164 [104]. The expression of the mRNA isoform OCT4A is critical to sustain stem cell self-renewal. On the contrary, the relevance of the OCT4B1 mRNA including the exon 2c, which contains a stop codon, is unclear. The OCT4B-190 is reported to be expressed in response to cell stress, whereas OCT4B-265 is upregulated under genotoxic stress through the p53 signaling pathway [104]. OCT4 protein through its POU domain recognizes and binds the octamer sequence ATGCAAA (or the inverse complement DNA sequence) present on the promoter or enhancer of its target genes.
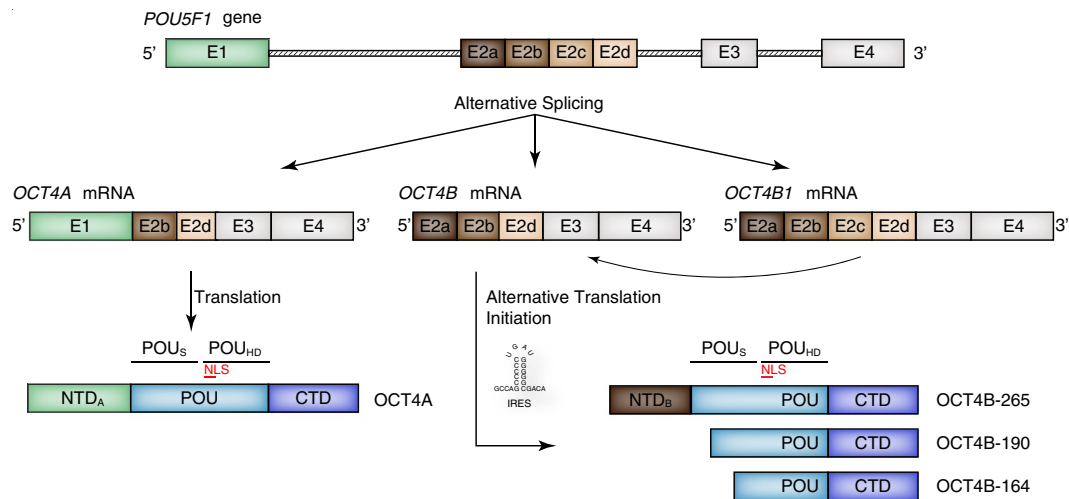
***Figure 1.7. Schematic representation of the alternative splicing of the human OCT4 gene mRNA and protein products.*** *From the OCT4 mRNA are generated three different isoform OCT4A, OCT4B, and OCT4B1. From the OCT4A mRNA is generated the OCT4A protein and OCT4B mRNA is translated into the three different protein isoforms OCT4B-265, OCT4B-190, and OCT4B-164. The N-terminal domain color corresponds to the exon of origin. The colored boxes are the exons while the dashed boxes are the introns* [100].

## 1.6.2  OCT4 regulation and function in early development and pluripotency

The POU5F1 gene contains three distinct *cis*-regulatory elements: the proximal promoter (PP), the proximal enhancer (PE) and the distal enhancer (DE), which controls the expression of OCT4 during the various phases of embryogenesis and development [91]. Therefore, the two enhancers are used in different stages of embryogenesis: the PE is activated in the epiblast whereas the DE is activated in the ICM, primordial cells and ESCs [106]. In human ES there is a crosstalk between the enhancers of the OCT4 gene and distal genomic regions. Gao et al. (2013) found in the interactome of POU5F1 enrichment for several transcription factors binding sites and active epigenetic marks, suggesting the formation of chromatin loops and cooperative regulation of transcription to control and maintain pluripotency.

OCT4 promoter in pluripotent ESCs is hypomethylated and the nucleosomes are characterized by the presence of highly acetylated histone H3 on Lysine 9 and Lysine 14 residues, whereas the Lysine 4 on histone H3 is trimethylated [94]. Once the cells undergo differentiation the locus is subjected to a series of epigenetic modifications resulting in the repression of OCT4 expression. In particular, this silencing is mediated by the deacetylation of the histones H3 and the G9-mediated methylation of the Lysine 9 on the histone H3, that induces the compaction and condensation of chromatin by the Heterochromatin protein 1 (HP1) [94]. Finally, OCT4 promoter is permanently methylated by the activity of the two *de novo* methyltransferase DNMT3A and DNMT3B, hence the gene is silenced [109].

### 1.6.3 Mutant Huntingtin affects neural induction

Recently, a paper published from Cattaneo's group reported that specific aspects of human neurodevelopment are affected by muHTT, such as neural progenitor specification and its early commitment [42]. By exposing nonintegrating HD and control iPSC lines carrying different CAG repeats (60, 109 and 180 CAG), to striatal differentiation protocol [64,65] they observed that long pathological CAG expansions lead to failure in neuro-ectodermal fate acquisition (Figure 1.8). Monitoring the neural induction phase of differentiation, HD-iPS lines with a large Q (Q60 and Q109) manifested delayed exit from pluripotency, which resulted to be more pronounced in the presence of extremely long CAG repeats (Q180 lines). Indeed, these cells presented higher expression of the pluripotency marker OCT4 and lower expression of the neuroectodermal fate determinant PAX6. Furthermore, the lines that remained OCT4 positive were committed to a mesodermal fate even after neural induction. This phenotype led to defective ventral telencephalic HD progenitors' specification and subsequent reduction in DARPP32[+/]CTIP2[+] MSN neurons. Finally, HD-iPSC lines presented defects in neural rosette formation and disrupted cortical organoids organization. This work illustrated how the *in vitro* neuronal differentiation could be affected in a CAG-length dependent manner by a postponed exit from pluripotency (Figure 1.8 D), a consequential delay in neuroectodermal commitment, and a reduced ability to fully reach (ventral) striatal neurons specification.
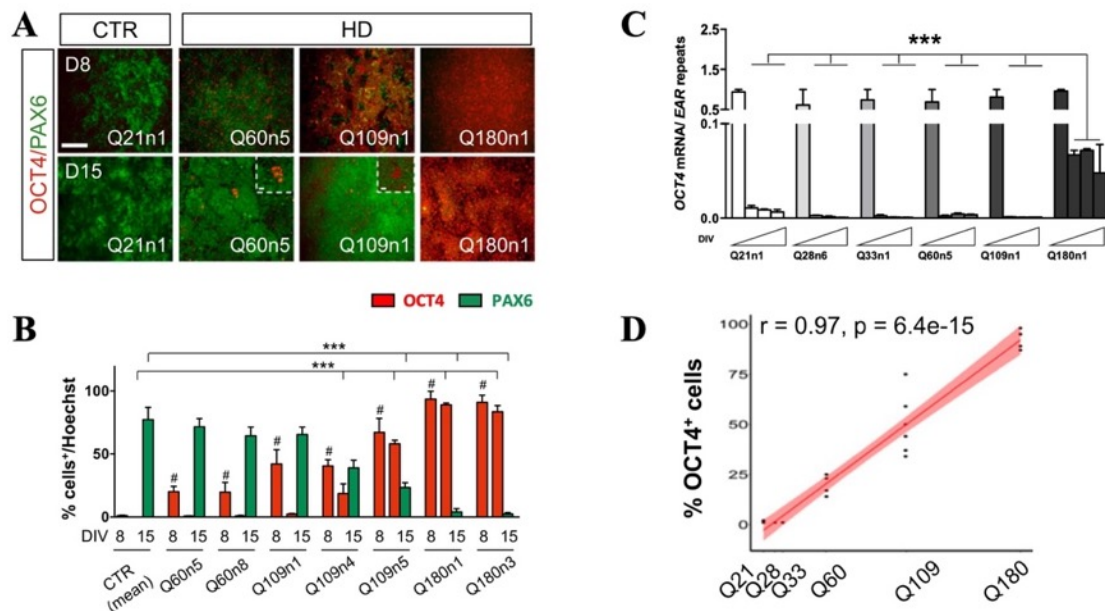


***Figure 1.8. Defective neural induction of HD iPSC cells (Figure modified from Conforti et al. 2018). A.*** *Oct4 and Pax6 immunocytochemistry at day 8 and day 15 of differentiation in CTRL (Q21n1) and HD lines (Q60n5, Q109n1, Q180n1).* ***B.*** *Percentage of OCT4+ (red bars) and PAX6+ (green bars) cells counted from the immunocytochemistry at day 8 and day 15.* ***C.*** *Gene expression analysis by qPCR for OCT4 mRNA in CTRL and HD lines along differentiation at day 0, day 8, day 15 and day 30.* ***D.*** *Plot representing the Pearson correlation among OCT4 and CAG length for all the HD lines and clones (Q60n5, Q60n8, Q109n1, Q109n4, Q109n5, Q180n1, and Q180n3).*

In a recent study from Thompson laboratory [96], the authors observed an upregulation of OCT4 in neuronal cultures derived from HD iPSCs at the end of differentiation. In this case, they identified at day 37 of differentiation a population of approximately 22% of OCT4$^+$ cells by immunocytochemistry. In addition, performing RNA-seq analysis they found a significant upregulation of OCT4 mRNA in HD lines, whereas all the other pluripotency markers analyzed were not dysregulated. Moreover, by RNA-seq and ChIP-seq analysis for the H3K4me3 histone mark they observed that almost half of the transcriptionally upregulated genes in HD had higher enrichment of H3K4me3 at their transcription starting sites (TSS). Therefore, taken together, all these pieces of evidence emphasize the necessity to consider not only the neurodegenerative processes underlying HD but also the possibility of a neurodevelopmental component of this disorder. Indeed, as a consequence of early dysfunctions occurring during neurodevelopment, neuronal homeostasis in adulthood may be affected, leading to increased cellular susceptibility to environmental stressors and toxins that normally would not favor cell death [14,97].

# 1.7 Epigenome complexity and regulation

The term "epigenetics" was introduced for the first time by Dr. Waddington in 1942 [112] to identify the biological events that are not described by genetic principles. These epigenetic signatures are ascribed to chromatin, the ultrastructure of DNA that through its modification affects the functionality of the genome. Epigenome modifications occur at the very early stages of development and differentiation and specific levels of chromatin complexity [99]. The eukaryotic genome to accommodate into a very confined space is packaged inside the nucleus through a highly ordered three-dimensional architecture (Figure 1.9).
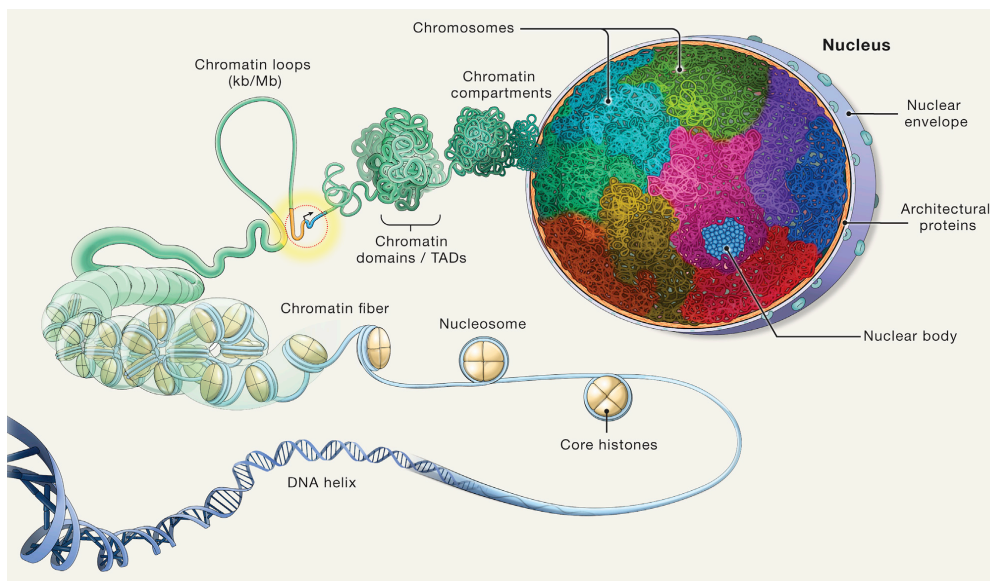


*Figure 1.9. The complexity of the human genome architecture. The starting point of genome organization are the nucleosomes, made of DNA wrapped around an octamer of four pair of histones that generates the chromatin fiber. This structure folds into loops that bring genes promoters (gold/blue) close to their upstream regulatory elements, such as enhancers (yellow) in order to control their expression. The chromatin fiber folds to form chromatin domains named TADs, that associating to each other generate chromatin compartments. Each chromosome occupies a specific volume inside the nucleus, identified as chromosome territory. Inside the nucleus, proteins aggregates and RNA are organized in structures named nuclear bodies (NBs, blue) [153].*

### 1.7.1 Histone modifications: the first level of epigenome complexity

Inside the nucleus the DNA associated to several proteins and RNA molecules is packaged into a complex and highly organized structure defined as chromatin. The basic unit of this structure is the nucleosome: DNA coiled around the four histone proteins H2A, H2B, H3 and H4 [100]. The nucleosome units are separated by a DNA linker that is bound to the histone H1 or its variants.

Several modifications targeting the amino and carboxy-protruding histone tails are critical for regulating gene activation or repression [99]. The best-characterized are the acetylation, methylation and phosphorylation [101]. Diverse histone modifications can coexist in the same histone tails, therefore exerting a specific biologic function by compacting or decondensing chromatin [102]. In fact, chromatin can be classified in two groups depending on its condensation: euchromatin and heterochromatin.

Euchromatin is the decondensed form enriched of expressed genes and therefore carrying H3 and H4 acetylation and H3K4 methylation [103,104]. Conversely, the condensed form is the heterochromatin that is further subdivided into "facultative" and "constitutive". The constitutive heterochromatin is typical of the pericentric and telomeric regions that are critical for chromosome folding, and is characterized by H3K9me3 mark that contributes to chromatin closure and gene silencing [105]. On the other hand, facultative heterochromatin constitutes a dynamic structure, as it can adopt open or compact conformations depending on the action of epigenetic regulators within certain temporal and spatial contexts. This chromatin form presents typically hypoacetylated histones, H4K20me1, H3K9me2, H3K27me3, H2AK110ub1[106].

Methylation on the side chain of lysine (K) residues on histone proteins is the best characterized PTM, as it is stable, versatile, and cross talks with many other modifications. Lysine methylation can be associated to both transcriptionally active and inactive chromatin. Methylated H3K4, H3K36, and H3K79 are related to euchromatin and therefore active genes, whereas the methylation on H3K9, H3K27 and H3K20 result in heterochromatin and inactive genes. The methylation is catalyzed by KTMs (histone lysine methyltransferases) that target specifically certain loci and are all characterized by the well-conserved SET. This domain, that is usually at the C-terminal binds the AdoMet (S-Adenosyl-L- methionine) cofactor to transfer the methyl group to a target lysine residue [58]. The opposite reaction, the demethylation, is mediated by the group of the Jumonji demethylases (KDM5-A/JARID1A, KDM5-C/JARID1C, KDM6A/UTX) through an oxidative reaction and radical attack involving Fe(II) and a-ketoglutarate [107–109].

**1.7.2 Chromatin modifiers: the second level of epigenome complexity**

Once histones are modified by specific epigenetic factors, a unique chromatin environment is generated and recruits particular TFs. Ultimately, these sequence-specific DNA-binding proteins, (Figure 1.10) employ chromatin-modifying enzymes creating enhancer and promoter states that can be either permissive or repressive for transcription [110].
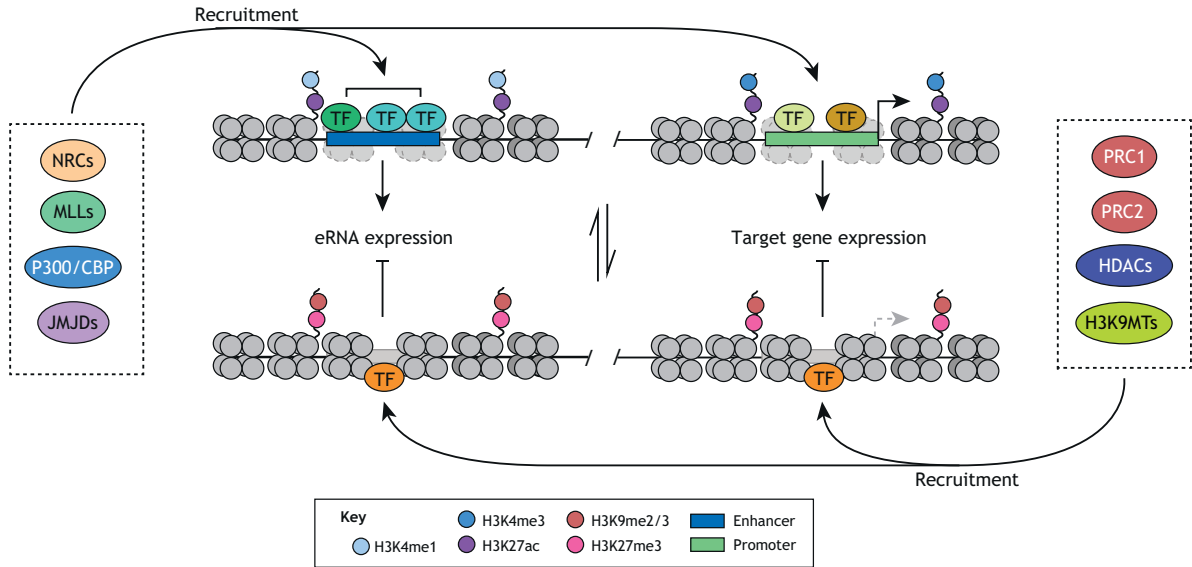
**Figure 1.10. Enhancers and promoter activity are regulated by different chromatin-modifying enzymes recruited to chromatin.** *Several enzymes can be recruited through the binding of several TFs. The modifications catalyzed by these enzymes correlate with enhancer and gene activity and include: the nucleosome remodeling complexes (NRCs), MLLs, P300/CBP and JMJDs [120].*

Modifications correlated with enhancer and gene repression are catalyzed by enzymes as the Polycomb Repressive Complexes 1 and 2 (PRC1/2), Histone deacetylases (HDACs) and H3K9 methyltransferases (H3K9MTs). Epigenetic signatures of active enhancers and promoters are typically associated with H3K27ac signal and local chromatin accessibility. Moreover, the H3K4me1 and H3K4me3 modifications are carried by active enhancers and promoters. On the contrary, H3K9 and H3K27 methylation are typical of inactive enhancers and promoters and therefore of chromatin inaccessibility [125].

### 1.7.2.1 Trithorax and Polycomb group proteins (PcGs)

The Trithorax group (TrxG) of proteins is very complex and heterogeneous (comprises the SWI/SNF complex and COMPASS family) as its action in transcriptional activation counteracts the repression mediated by the PcG proteins (Figure 1.11).

The central core of the complex, referred as COMPASS (sometimes WARD), includes WDR5, ASH2, RBBP5, and DPY3 factors. This part additionally comprises HCF1, WDR82, and the DNA-binding protein CXXC1 (CFP1) and mediated the trimethylation of the Lys4 on the histone H3 [111]. There are other COMPASS-like group of proteins, as the MLL1/MLL2 COMPASS that contain MENIN and HCF1, where the MLL2 and MLL1 are present in a mutually exclusive manner. In

particular, MLL1 is responsible for the trimethylation of a subset of genes (e.g. *HOX* genes), whereas MLL2 is required for the methylation of the bivalent promoters in hESCs [112].
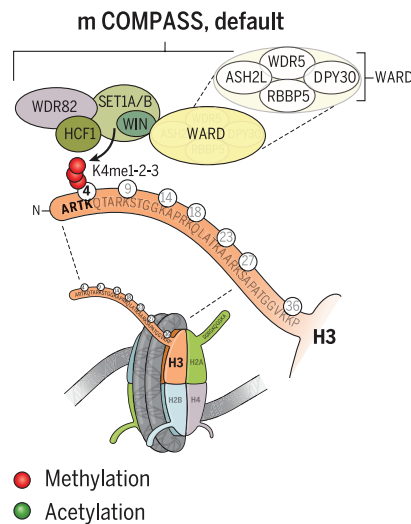


*Figure 1.11. The Trithorax group and the COMPASS family of histone H3K4 methylases. Mammalian COMPASS (mCOMPASS) composition and its mono-, di-, and trimethylation of lysine 4 on histone H3 [121].*

The best studied epigenetic regulator is the family of Polycomb group proteins (PcG proteins). The Polycomb Repressive Complex 1 and 2 (PRC1, PRC2) are two-well characterized epigenetic repressors. PRC1 contributes to chromatin compaction by histone H2A mono-ubiquitination, whereas PRC2 catalyzes tri-methylation of the histone H3 lysine 27 (H3K27me3) inducing the formation of heterochromatin and consequently transcriptional repression. This complex is composed of two lysine methyltransferases, Ezh1 and Ezh2, and two non-enzymatic subunits, as Suz12 and EED which are necessary for the correct enzymatic activity and all components assembly [113]. The mechanism of PRC2 recruitment to DNA is not yet clear in mammals, it is known that the Jumonji/ARID domain-containing protein Jarid2 and the members of the Polycomb-like family, the Pcl proteins, might be responsible for PRC2 recruitment to target genes [129-132] (Figure 1.12).

Likewise, PRC1 can recruit PRC2 [118–120] and the recruitment of both these two complexes can be facilitated also by noncoding RNAs [121] and histone marks [122,123]. Therefore, the recruitment of PRC2 to chromatin is based on different mechanisms and it is probably locus-specific [99]. In mammals, most of the genes important for lineage commitment, development and differentiation are finely controlled by the PcG proteins [124–126].

The PRC1 complex is bigger than PRC2 and more heterogeneous in its composition. The central components are RING1A and RING1B (RING1A/B), that are E3 ubiquitin ligases responsible for the monoubiquitylation of the lysine 119 of histone H2A (H2AK119ub).
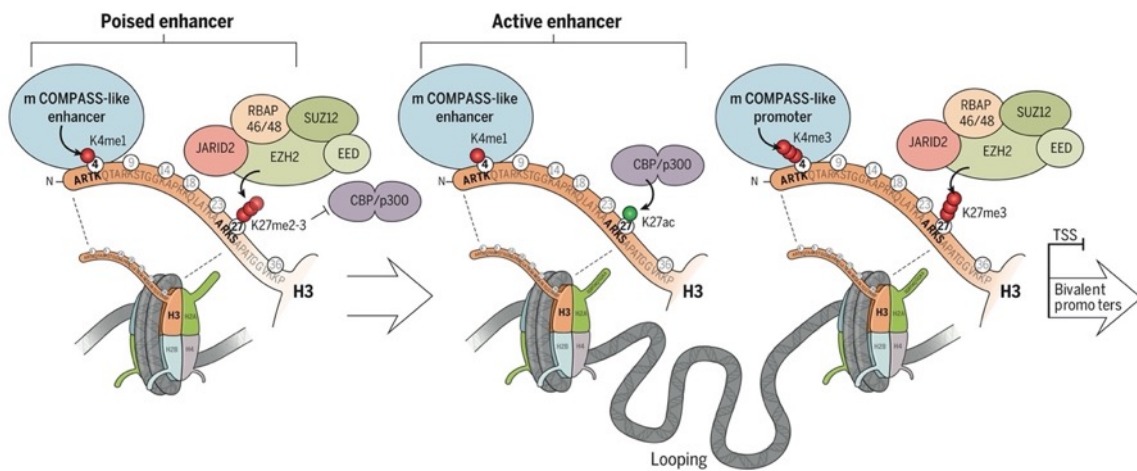
**Figure 1.12. Schematic representation of the PRC2 complex**. *(Left) PRC2 acts at enhancer regions, where it can di- and tri- methylate Lys27 preventing the activity of the histone acetyl- transferases (HATs) CBP/p300. PRC2 through the enzymatic subunit EZH2 catalyzes the tri-methylation of the Lysine 27 on histone H3. H3K27me3 recruits PRC1 that ubiquitylates H2AK119 promoting chromatin compaction and gene silencing. (Right) Bivalent promoters are marked by the H3K27me3 by PRC2 and the H3K4me3 by the MLL2/COMPASS [121].*

This histone modification leads to chromatin compaction and transcriptional silencing. PRC1 can be present into two different forms, named canonical and noncanonical (Figure 1.13).
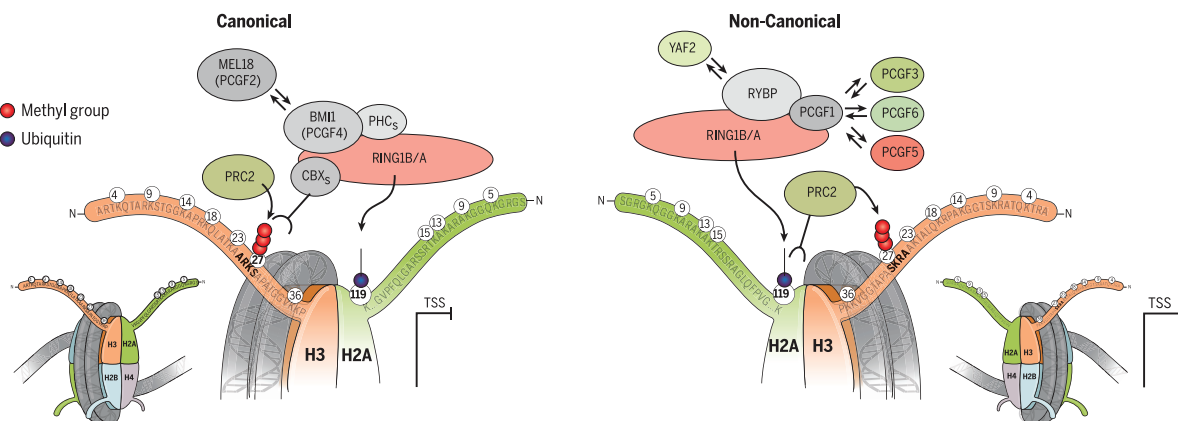


**Figure 1.13. Schematic representation of PRC1 complex**. *(Left) Canonical PRC1 containing the CBXs proteins, recruited by the H3K27me3 deposited by PRC2. The catalytic subunits RING1A/B monoubiquitylates the Lys119 on H2A. (Right) Noncanonical PRC1 complex containing RYBP and any of the different PCGF proteins (PCGF1/3/5/6). Through the deposition of the H2AK119ub mark is recruited PRC2 [121].*

The canonical form contains the chromobox homolog (CBX) protein that recognized H3K27me3 residues and therefore is recruited after PRC2 activity. This complex also includes the protein BMI1 (PCGF4) that is necessary for the correct assembly of the canonical PRC1 and RING1A/B catalytic activity. The noncanonical PRC1 comprises RYBP, that is common to all noncanonical variants, and any of the various form of the PCGF proteins. Noncanonical PRC1 binding to chromatin is

30

independent from PRC2 but through the monoubiquitylation of Lys119 can recruit PRC2. The different biochemical composition of the canonical and noncanonical PRC1 complex have been well clarified, the different molecular and biological activities remain to be elucidated, together with the perspective cooperation at specific loci [127].

**1.7.2.2 H3K9 Methyltransferases**

The formation of constitutive heterochromatin is mediated by di- and tri- methylation of the H3K9 residue. This histone modification is catalyzed by a family of methyltransferases characterized through a SET-domain. The enzymes SETB1, SUV39H1 and SUV39H2 contribute to both H3K9me2 and H3K9me3 methylation [128,129], whereas the proteins GLP and G9a (named also EHMT1 and EHMT2, respectively) catalyze H3K9me1 and H3K9me2 [130–132]. In particular, SUV39H1 and SUV39H2 are critical for pericentromeric heterochromatin formation, whereas G9a mono- and di-methylates H3K9 residues of euchromatin [131,133,134]. Then, the H3K9me2 and H3K9me3 are recognized and bound by the Heterochromatin Protein 1 (HP1 and HP3 are the isoforms present in mammals). HP1 can self-oligomerize and recruit other repressive histone modifiers, thus contributing to the compaction and spreading of heterochromatin [135–137].

**1.7.3 Chromatin high ordered architecture: the third level of epigenome complexity**

The chromatin fiber can fold through self-interactions and generate loops that favor the activation or repression of genes [138] (Figure 1.14). These structures facilitate the interaction of regulatory elements, typically enhancers with promoters, that are far apart from 10 to several hundred kb in size. Then chromatin loops fold into the so-called chromatin domains, named Topological Associated Domains (TADs). TADs typically comprise a central region with frequent local connections where are located the tissue-specific genes that is flanked by less interacting boundary regions, where are resident the housekeeping genes [138]. The ATP-dependent protein cohesin determines the formation of chromatin loops through its motor activity [139,140], whereas the boundaries of the extruded-loop structures are defined by the CTCF proteins by defining the location of cohesin [141] (Figure 1.14 A). The association of different TADs generates chromatin compartments. Chromosomes occupy a specific positions and volume within the nucleus and these regions are referred as chromosomal territories (CTs) [142], (Figure 1.14 B).
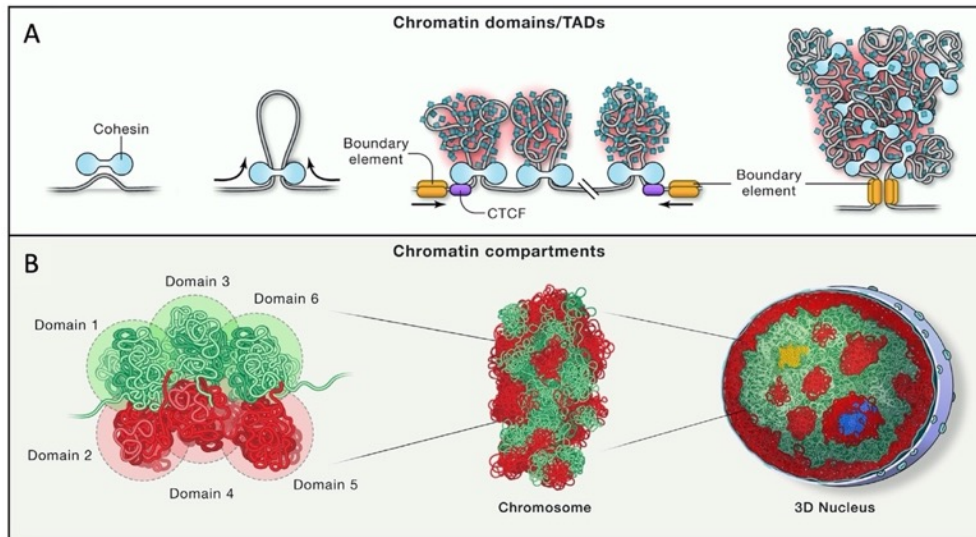
*Figure 1.14. Self-organization of chromatin. (A)* *Chromatin loops fold into TADs through the loop extraction mediated by the cohesion motor (light blue). The boundaries (gold) of these structures are defined by the CTCF protein (purple) by defining the position of cohesion.* *(B)* *The association of multiple homotypic domains (depicted in green and red, 1-6) generate chromatin compartments that finally constitute the chromosome. Interactions between the various homotypic domains of chromosomes and the 3D structure of the nucleus constitute blocks of heterochromatin (red) and euchromatin (green)[153].*

## 1.7.4 The nuclear environment: the fourth level of epigenome complexity

The nuclear space is dense of macromolecules non-homogeneously distributed that constitute distinct nuclear bodies (NBs). These nuclear niches are very dynamic and result from protein-protein interactions that impact on gene regulation [143,144], as transcription (working as transcriptional factories), replication (operating as replication foci) or repression (acting as repressive bodies). The PcG bodies or foci are an example of nuclear bodies that are often associated to facultative pericentric heterochromatin. PcG foci are nuclear compartments that can be present in variable number and size [145,146] and likely include PcG-bound loci that interact both in *cis* and in *trans*, which are located in proximity [147–152]. The folding properties and the compaction of the chromatin-associated to PcGs is exclusive as compared with other compartments in the nucleus; furthermore, these nuclear bodies are variable among the various cell types.

In mammals, PcG foci are normally confined in certain nuclear compartments, as are mainly located in the center of the nucleus without interacting with the nuclear lamina (Figure 1.15).
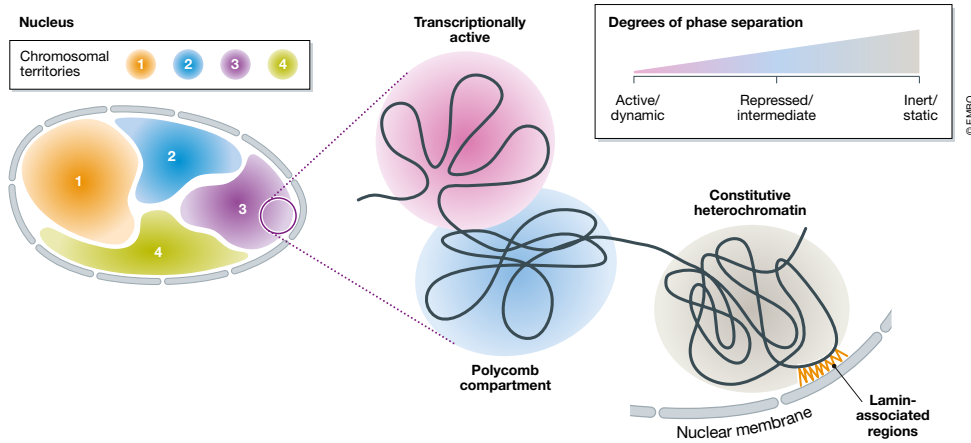
*Figure 1.15. Chromatin organization inside the nucleus. Inside each chromosomal territory are present different degrees of phase separation. This process results from homotypic interactions among proteins and allow the formation and maintenance of chromatin-chromatin interactions and domains, as euchromatin and heterochromatin 163. The transcriptionally active regions interplay with RNA polymerase, chromatin modifying enzymes and TFs. Constitutive heterochromatin is the less dynamic and more phase separated as it interacts with the nuclear lamina and heterochromatic factors. PcG foci constitute a unique compartment generated by PRC1 and PRC2 interactions, where there are intermediate chromatin dynamics 164.*

This is the principal feature that distinct chromatin folded with PcGs to active and constitutively inactive chromatin [153,154]. These *cis* and *trans* interactions among PcG-bound loci are mainly mediated by PRC1 [149,155], whereas PRC2 is responsible for PRC1 recruitment to its target sites through the trimethylation of the H3K27 residues [148] (Figure 1.16 A). In particular, both local and long-range interactions among PcG-bound loci are mediated by the multimerization through the sterile-alpha motif (SAM) domains of the PRC1 complex [146,156] (Figure 1.16 B).
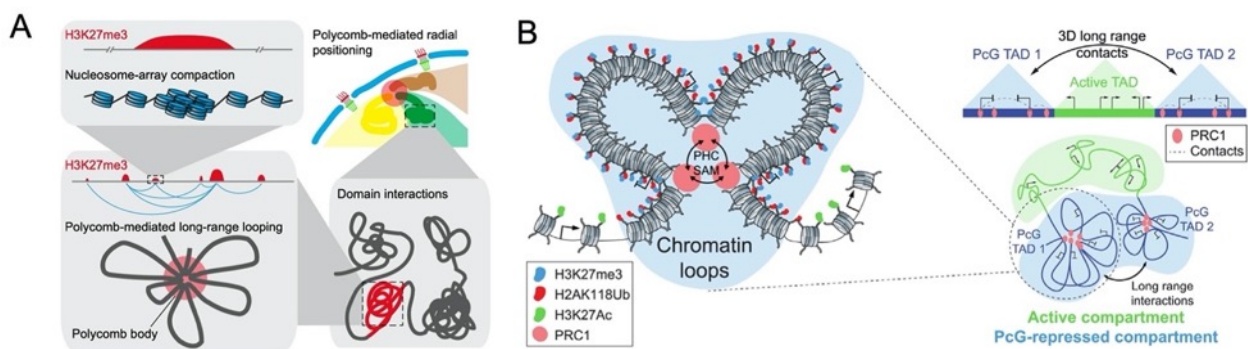


*Figure 1.16. Schematic representation of Polycomb bodies structure. (A) Cartoon depicting the PcG foci localization in the nucleus and the hierarchical 3D organization depending on the interplay among the deposition of H3K27me3 residues by PRC2 and PRC1 recruitment. Figure from Rada-Iglesias, Grosveld, and Papantonis (2018). (B) The genome is subdivided into active (green) and repressed (blue) compartments, organized into Topological Associated Domains (TADs) depicted as triangles. The 3D organization of PcGs derives from the generation of long-range contacts (black arrow) between multi-looped PcG TADs 1 and 2 (blue triangles) that segregate active TADs (in green). Finally, the polymerization of PHC-SAM domains of PRC1 allows the connection between PRC1-bound sites and the formation of the final 3D structure 170.*

### 1.7.5 Regulation of gene expression and nuclear architecture: the nuclear lamina

In the nuclei, the inner nuclear membrane is firmly connected with a meshwork of intermediate filaments, primarily of type V referred as lamins [157]. Lamins are subdivided into A-type lamins (lamin A and C) and the B-type lamins (lamin B1 and laminB2) [176,177]. A-type lamins are regulated during development, as are almost absent in the early embryo [160,161], whereas are mainly expressed in lineage-committed progenitor cells and fully differentiated cells [162,163]. Furthermore, A-type lamins are mostly located in the perinuclear and nucleoplasmatic compartments. Otherwise, B-type lamins are ubiquitously expressed and tightly associated to the nuclear membrane. These proteins are fundamental epigenetic regulators, as the nuclear lamina interacts with specific regions of the genome at specific sequences named Lamina Associated Domains (LADs). These domains are enriched of H3K9me2, H3K9me3 and H3K27me3 histone modifications and therefore maintain the gene expression repressed [164,165].

# 1.8 Chromatin remodeling during embryonic development and stem cell differentiation

Embryonic stem cells are distinguished by the presence of an hyperdynamic and open chromatin structure, as genome plasticity is necessary for the maintenance of pluripotency and self-renewal [166,167]. Chromatin morphology in hESCs is unique, as heterochromatin is ordered in a small number of large domains, that with differentiation are reduced in size and increase in number [168–171]. The chromatin proteins, as the linker histone 1 (H1) and the core histones are more loosely bound to chromatin. In pluripotent stem cells also HP1 and Lamin B are less tightly bound [166,172] (Figure 1.17 A). Lamin A in pluripotent stem cells is absent and starts to be present in the nuclear lamina during early differentiation [173].
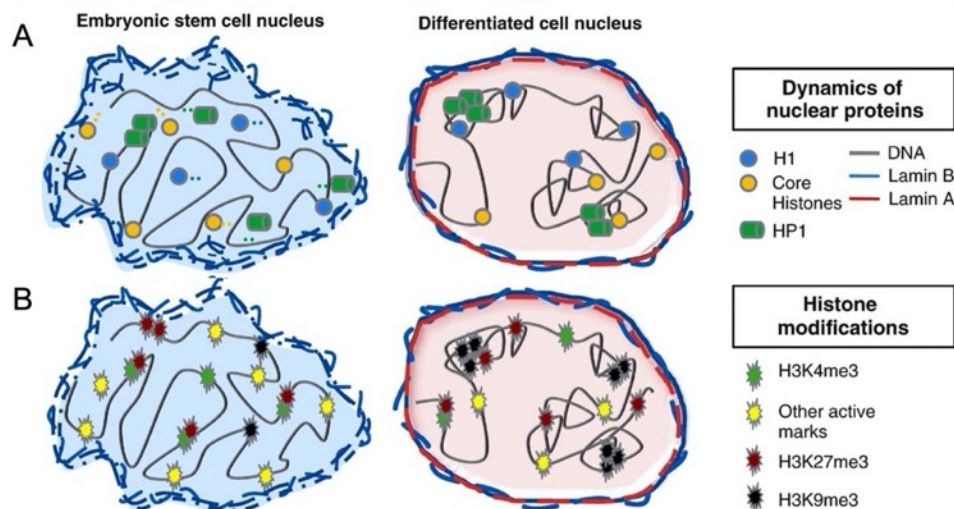


*Figure 1.17. Schematic representation of the key features of ESCs and differentiated cells. (A) ESCs present more dynamic interactions among several nuclear and chromatin proteins, as H1 linker protein (blue circles), core histones (yellow circles) and HP1 (green). The nuclear envelope is surrounded by Lamin B. Differentiated cells present also Lamin A in the nucleus and the proteins are more tightly bound to each other. (B) ESCs carry predominantly active histone modifications, whereas differentiated cells are characterized by repressive histone marks as H3K9me3 [190].*

Moreover, ESCs present high rate of active histone modifications, as H3K14ac, H3K4me3, H3K36me2 and H3K36me3 [166,174], whereas differentiated cells are characterized by repressive marks, as H3K9me3 that accumulates in well-defined foci [175] (Figure 1.17 B).

Differentiated neuronal progenitor cells (NPCs) present heterochromatin foci hinting the formation and maturation of heterochromatin during ESCs differentiation [176]. During differentiation the genes responsible for stemness, such as Oct4, Nanog and Klf4 interact more frequently with the nuclear lamina [177]. Interestingly, embryonic stem cells present a unique histone signature on the chromatin

domains bound by the PcG proteins, often associated to developmental genes, referred as "bivalent domains" (BD), (Figure 1.1.8).
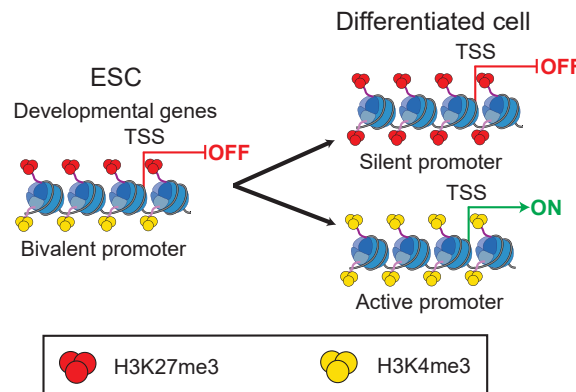


*Figure 1.1.8. Schematic representation of the distribution of bivalent domains in ESCs. In self-renewing cells BDs maintain developmental genes poised. These genes are resolved upon differentiation depending on the c activated by losing H3K27me3 or silenced by H3K4me3 loss. Figure modified from Blanco et al.*

This term was coined by Be[.....]l. [.....]enti[.....]ific modifica[.....] characterized by the presence of [.....] and [.....] at th[.....]gions in mous[.....] combination of both "repressiv[.....]iva[.....] was [.....]ghly enriched[.....] compared to differentiated cell[.....]d t[.....]ci [.....]TFs related t[.....] development and lineage specifi[.....] ref[.....]ssi[.....] of developme[.....] controlled by the synchronized activity of the PcG proteins and the Trithorax group catalyzes the tri-methylation of the Lys4 on histone H3. Bivalent domains are mainly pre rich regions and specifically at promoter regions that are "poised" for activation, where downstream the transcription starting site by the RNA polymerase II ready for a rapid response to distinct developmental inputs.

BDs are normally resolved during differentiation into H3K4me3 or H3K27me3 regions depending on the expression pattern related to a specific cell fate. This process is facilitated by histone demethylases that remove specifically H3K4me3, as JMJD3 and UTX [196-199], according to their essential function in development and differentiation [181,184] (Figure 1.19).

As already explained, in embryonic stem cells gene silencing is mediated mostly by facultative heterochromatin, through the repressive action of the H3K27me3 histone mark, mainly by the presence of bivalent domains. This modification is not permanent and tight as the H3K9me3 that generates constitutive heterochromatin and mediate transcriptional repression in differentiated cells [174,185].
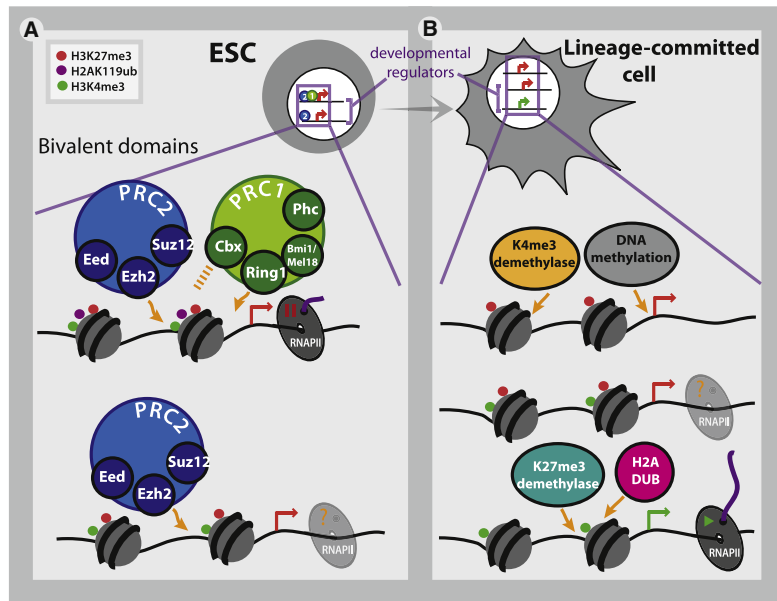
*Figure 1.19. Schematic representation of bivalent domains regulation in ESCs and lineage-committed cells. **A.** ESCs are characterized by the presence of BDs that present both H3K27me3 and H3K4me3 marks. Moreover, these cells present colocalization of PRC1 and PRC2 at some promoters, whereas PRC2 alone targets specifically other ones. In presence of both complexes the promoters are poised by the presence of the RNA Polymerase II. **B.** Once cells start their commitment many bivalent domains are resolved depending on the expression state of the gene. This is possible through the action of the histone demethylase JMJD3 and UTX and possibly histone H2A deubiquitylates (H2A DUB) to activate PcG target genes and facilitate H3K27me3 and H2AK119Ub1 removal* [135]*.*

## 1.8.1    The pivotal role of PRC1/PRC2 during development

The Polycomb repressive complexes are fundamental during early development, since the absence of PRC2 activity causes embryonic lethality in mice [205-208], whereas the loss of PRC1 activity leads to less severe phenotypes that appear later in development [190–195]. PRC2 function during embryonic development has been elucidated via generating knockout mice for the various components. In the table below, from Surface, Thornton, and Boyer (2010), are reported the main loss of function phenotypes related to certain PcG proteins both *in vitro* and *in vivo* (Table 1).

The variability of these phenotypes probably reveals not only how PRC2 is fundamental during development but also the high degree of overlapping functions of PcG homologs. New evidence on PRC2 roles in development and in cell fate commitment in mammals came from studies on embryonic stem cells. PcG proteins were found to control a large group of developmental and signaling genes in hESCs, such as Hox gene clusters and members of the gene families *Fox, Dlx, Pou, Irx, Pax, Sox, Tbx, Wnt* are regulated by PcG proteins [215-217].

These genes are repressed to prevent differentiation by PcG proteins in proliferating ESCs, although these cells are able to maintain their self-renewal and pluripotent genes in absence of these proteins [218-220]. Thus, PcG proteins are essential for cell fate transition as they dynamically regulate genes that can potentially be activated or silenced upon differentiation.

| Subunit | Phenotype In Vitro (ESC) | Phenotype In Vivo | Key References |
|---|---|---|---|
| **PRC2** | | | |
| Suz12 | De-repression of target genes; global loss of H3K27me3 and-me2, decrease in Ezh2 protein levels. Embryoid bodies lack proper structure. | Lethal at early postimplantation stage. Die ~E7.5–E8.5 | Pasini et al., 2004, 2007; Jung et al., 2010 |
| Eed | Target genes are de-repressed; genome-wide decrease in H3K27me1, H3K27me2, and H3K27me3. Decrease in Ezh2 protein levels. Eed null ESCs fail to differentiate properly in vitro, but can contribute to chimeras. | Disrupted axial patterning; Fail to properly gastrulate and to produce embryonic mesoderm. Die ~E8.5. | Faust et al., 1995; Montgomery et al., 2005; Chamberlain et al., 2008; Shumacher et al., 1996; Leeb et al., 2010 |
| Ezh2 | Fail to abolish H3K27me1 and H3K27me3 at some genes. Null ESCs fail to undergo mesoendoderm differentiation, but phenotype is less severe than Eed null, because of a partial redundancy with Ezh1, particularly at developmental genes. | Lethal at early postimplantation stage. Die ~E7.5–E8.5. | Shen et al., 2008; O'Carroll et al., 2001 |
| Jarid2 | Global H3K27 methylation unaffected; fail to properly differentiate. | Incompletely penetrant; neural, cardiac, liver, and hematopoietic defects. Die ~E11.5-15.5, depending on strain background. | Jung et al., 2005; Motoyama et al., 1997; Takeuchi et al., 1995; Shen et al., 2009; |
| Pcl2/Mtf2 | Upregulated pluripotency regulators; fail to properly differentiate. | Mice are viable, but have growth defects. Incompletely penetrant; posterior homeotic transformation. | Walker et al., 2010; Wang et al., 2007 |
| **PRC1** | | | |
| Ring1B | ESC lines are unstable and display a high propensity toward differentiation; de-repression of target genes; embryoid body formation is abnormal. Some ESC lines can differentiate into all three germ layers. | Developmental arrest in early gastrulation similar to PRC2 components. Die ~E10.5. | Voncken et al., 2003; van der Stoop et al., 2008; Leeb and Wutz, 2007; Leeb et al., 2010 |
| Ring1A | Ring1A/Ring1B double knockout ESCs lose ESC identity and fail to self-renew after several passages. | Mice are viable. Anterior transformation and axial skeletal patterning abnormalities in both heterozygote and homozygote mutants. | del Mar Lorente et al., 2000; Endoh et al., 2008 |
| Bmi1 | | Mice are viable, but with posterior homeotic transformation and severe immunodeficiency. | van der Lugt et al., 1996; Akasaka et al., 2001 |
| M33 (Cbx family) | | Mice are viable, but have growth defects; homeotic transformation; and increased sensitivity to retinoic acid during development. | Coré et al., 1997 |
| Rae28 (Phc family) | | Some null mice are viable. More severe phenotypes include perinatal lethality, posterior skeletal transformations, and neural crest defects. | Takihara et al., 1997 |

*Table 1. Summary of the loss of function phenotypes both in vitro and in vivo of certain PcG proteins occurring during early development. Figure modified from Surface, Thornton, and Boyer (2010).*

### 1.8.2   Importance of H3K9me2/me3 deposition for normal embryonic development

Several loss-of-function studies conducted in mouse embryos demonstrated the requirement of H3K9me2/3 deposition to complete development. In particular, knockout for the proteins G9a and GLP causes embryonic lethality in mice together with significant morphological abnormalities and altered chromatin organization and gene expression [131,132].

Moreover, embryonic lethality is caused also by the homozygous knockout of the protein SETDB1, earlier than G9 and GLP mutants, since it occurs around implantation time and leads also to defects in inner cell mass growth [202]. Double knockout mice for SUV39H1 or SUV39H2 causes genome

instability and therefore leads to sub-Mendelian ratio of newborns and prenatal lethality [203]. The same effects are caused by HP1 β loss, probably due to defective neuromuscular junctions and cerebral cortex development [204]. These diverse lethal phenotypes related to the various classes of protein factors and methyltransferases associated to H3K9me2/3 suggest their different contributions during development [54]. Early lineage commitment is regulated by G9a, GLP, and SETDB1 [94,132,205–209], whereas SUV39H1/H2 are responsible for the maintenance of fully differentiated cell identity [210–212] and genome stability [203].

### 1.8.3 H3K9me3 deposition: crucial regulator of pluripotency exit and maintenance of cell identity

The OCT4 gene is highy expressed during gametogenesis and embryonic stem cells [213,214] and is progressively and irreversibly silenced *in vivo* upon embryo implantation or *in vitro* with differentiation, as other crucial pluripotency-associated genes, including Nanog, Stella and Rx-1. This transcriptional repression results from increased deposition of H3K9 methylation by SET-containing methyltranseferases, as GLP and G9a [94,209,215]. In particular, *OCT4* undergoes a multistep inactivation process that involves G9a methylation. This methyltransferase prevents OCT4 reactivation in differentiated cells when returned in pluripotency culture conditions [94], and finally triggers DNA methylation at the promoter through the *de novo* methyltransferases Dnmt3a/b.

# 2 | Results and Discussion

## 2.1 The effects of mutant HTT on neuronal differentiation of HD RUES2 lines

We studied the specific effects of the Huntingtin CAG expansion during differentiation by exploiting the isogenic series of HD human embryonic stem cells (RUES2), which in fact allows to directly correlate the effects of different pathological CAG lengths in a fixed genetic background.

### 2.1.1 The RUES2 line is prone to acquire an MGE-like signature upon exposure to striatal differentiation

Firstly, we evaluated the potential of the parental RUES2 line to acquire a striatal-MSN cell identity. Therefore, we exposed the cells to a striatal differentiation protocol [71, 72].

To study the propensity of the RUES2 parental line to differentiate toward the dorso-ventral axis, we differentiated this control line in parallel to H9 hES cells, as are already known to respond correctly to the protocol employed. Therefore, we analysed by high content qPCR the expression for several LGE and MGE markers from DIV15 to DIV50. The H9 line correctly displayed the acquisition of an LGE fate by the progressive upregulation of ISL1, EBF1, CTIP2, DARPP32, and DLX6 (Fig 2.1 A), mimicking the transition from VZ to SVZ to mantle zone. Inversely, the expression analysis of the RUES2 parental line revealed a downregulation of the LGE transcripts (Fig 2.1 A) in parallel to the upregulation of MGE transcripts, as NKX2.1, SOX6 and SPON1 at DIV25 and NKX6.2, LHX8 and OLIG2 at DIV50 (Fig 2.1 B).
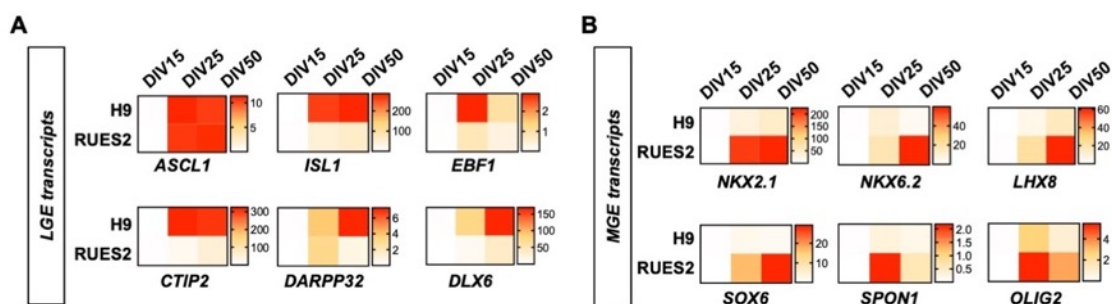


*Figure 2.1. High content gene expression analysis for LGE and MGE markers in differentiated RUES2 cells. Heat map displaying the expression levels of high-content qPCR analysis (Biomark) for the H9 and RUES2 lines at DIV15, DIV25, and DIV50 of the LGE (A) and MGE (B) transcripts. Student's t-test, \* p< 0,05 \*\* p< 0,01 \*\*\* p<0,001 \*\*\*\* p<0,0001. N = 3 biological replicates.*

**2.1.2 Mutant HTT impairs differentiation and MGE cell fate acquisition of HD RUES2**

The effect of muHTT on differentiation and cell fate specification has been reported in several studies employing human neurons derived from PSCs, as MSNs that are mainly affected in HD [38,40–43,76,96,216,217]. Although the parental RUES2 line is predisposed toward the MGE, we wanted to verify whether in the presence of muHTT the differentiation was affected. For this purpose, we employed the isogenic edited control line carrying Q22 and the HD lines with 50 and 58 CAGs [85].
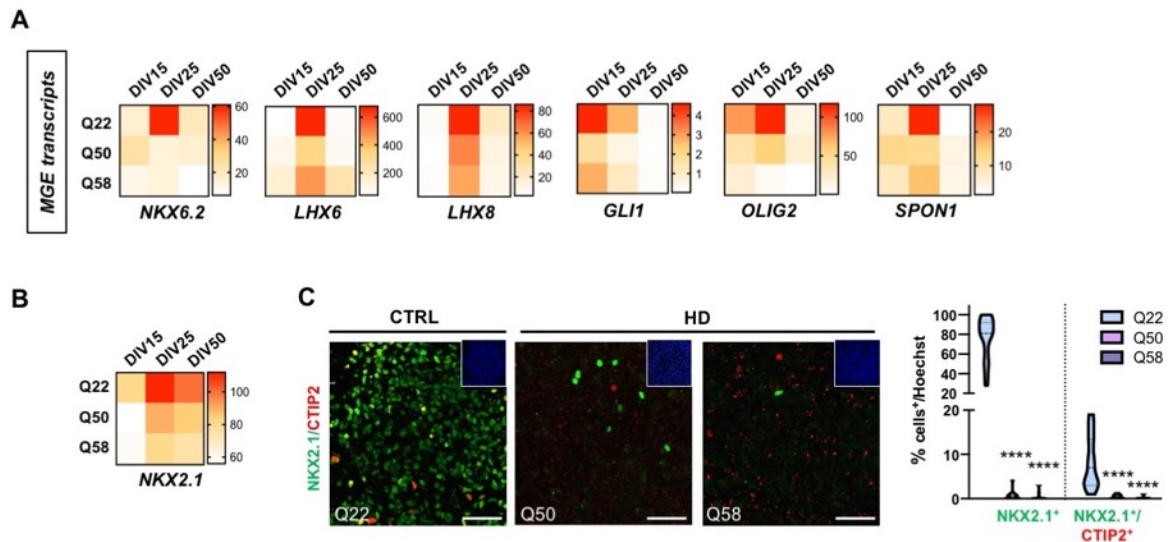


*Figure 2.2. **Characterization of MGE cell fate acquisition in HD RUES2 lines differentiation. (A)** Heat map displaying high-content qPCR analysis (Biomark) for the RUES2 control Q22, HD Q50, and HD Q58 lines at DIV15, DIV25, and DIV50 of the MGE transcripts. N = 3 biological replicates. **(B)** Gene expression analysis from high content qPCR for the RUES2 control Q22, HD Q50, and HD Q58 lines at DIV15, DIV25, and DIV50 of NKX2.1. N = 3 biological replicates. **(C)** Immunofluorescence analysis for the RUES2 control Q22, HD Q50, and HD Q58 lines at DIV30 for NKX2.1 (green) and CTIP2 (red). Representative confocal images, 40X, zoom = 1.7. Scale bar, 50 μm. Top right, Hoechst inset. Quantification of control Q22, HD Q50, and HD Q58 RUES2 lines for NKX2.1+ and NKX2.1+/CTIP2+ cells for the by CellProfiler software (Version 2.1.1). N = 3 biological replicates. ANOVA one-way, ****p < 0.0001.*

Gene expression analysis of the main MGE transcripts pointed out that muHTT negatively affects MGE determination in HD RUES2 lines, as the mRNA level of NKX6.2, LHX6, LHX8, GLI1, OLIG2, SPON1 and GSX1 were reduced compared to control Q22 (Figure 2.2. A). Moreover, while the NKX2.1 transcript starts to be upregulated from DIV15 in control Q22 line, the HD displayed a delayed expression with lower levels along all the differentiation (Figure 2.2. B). NKX2.1 is not only a crucial marker for MGE development, but also for LGE formation when it is co-expressed with ISL1 and CTIP2 [72,233]. Therefore, we checked this marker also by immunofluorescence (Figure 2.2.C). The ventral identity acquisition of the control line was confirmed by the presence of only 10% of NKX2.1+/CTIP2+ progenitors, with a complete loss of this cell population in the HD lines. Besides, the control Q22 line correctly displayed a gradual increase of NKX2.1+ cells along the differentiation,

achieving a peak of 80% at DIV50. On the contrary, we did not detect this cell population in the HD RUES2 lines (Figure 2.2.C).

## 2.2. Focus on neuronal induction to elucidate the defects of cell fate specification of HD RUES2 lines

### 2.2.1 Monitoring the OCT4 to PAX6 transition during neural induction

Firstly, we monitored by immunofluorescence the transition from pluripotency to neuroectoderm by measuring the total number of cells positive for the pluripotency marker OCT4 and the neuroectodermal marker PAX6, at DIV0, DIV4 and DIV8. For this analysis, we employed the isogenic series of hES RUES2 HD lines, including the control Q22 together with HD lines carrying Q50, Q58 and Q72, that we received from Prof. Brivanlou at the Rockefeller University through the CHDI Foundation in the context of a collaboration [76].

The results obtained confirmed phenotypes already identified in the HD iPS cells [43], giving proof of the robustness of this isogenic cellular system. All RUES2 lines were pluripotent, as shown by the OCT4 staining displayed at DIV0 (Figure 2.3 A). Upon exposure to the dual SMAD inhibition, the control Q22 line correctly acquired the neuroectodermal fate presenting an upregulation of PAX6[+] and a progressive OCT4 down-regulation at DIV4 and mainly DIV8. In contrast, all HD lines tested revealed at DIV4 a significant persistence of OCT4[+] cells, together with a reduction of PAX6[+] compared to the control Q22 cultures. At DIV8 the Q58 exhibited still a significant persistence of OCT4[+] cells in respect to the control Q22 and HD Q74 line, in parallel to significant reduction in PAX6[+] cells compared with the control Q22 line.

To further check the validity of these immunofluorescence data, we performed qPCR analysis on the same lines confirming the defective down-regulation of OCT4 in the HD lines during neural induction (Figure 2.3 B left panel). Especially, the HD lines with longer CAG tracts, the Q58 and Q74, presented a significantly higher transcript level at DIV4 compared to the control Q22 line.
Moreover, we monitored the expression profile of PAX6 in the same time-window, and also for this marker the immunofluorescence data were confirmed. In pluripotency, PAX6 mRNA was correctly not detected in both control and HD lines, conversely at DIV4 and at DIV8 the HD lines displayed a significant reduction of the transcript compared to the Q22 lines (Figure 2.3 B right panel).
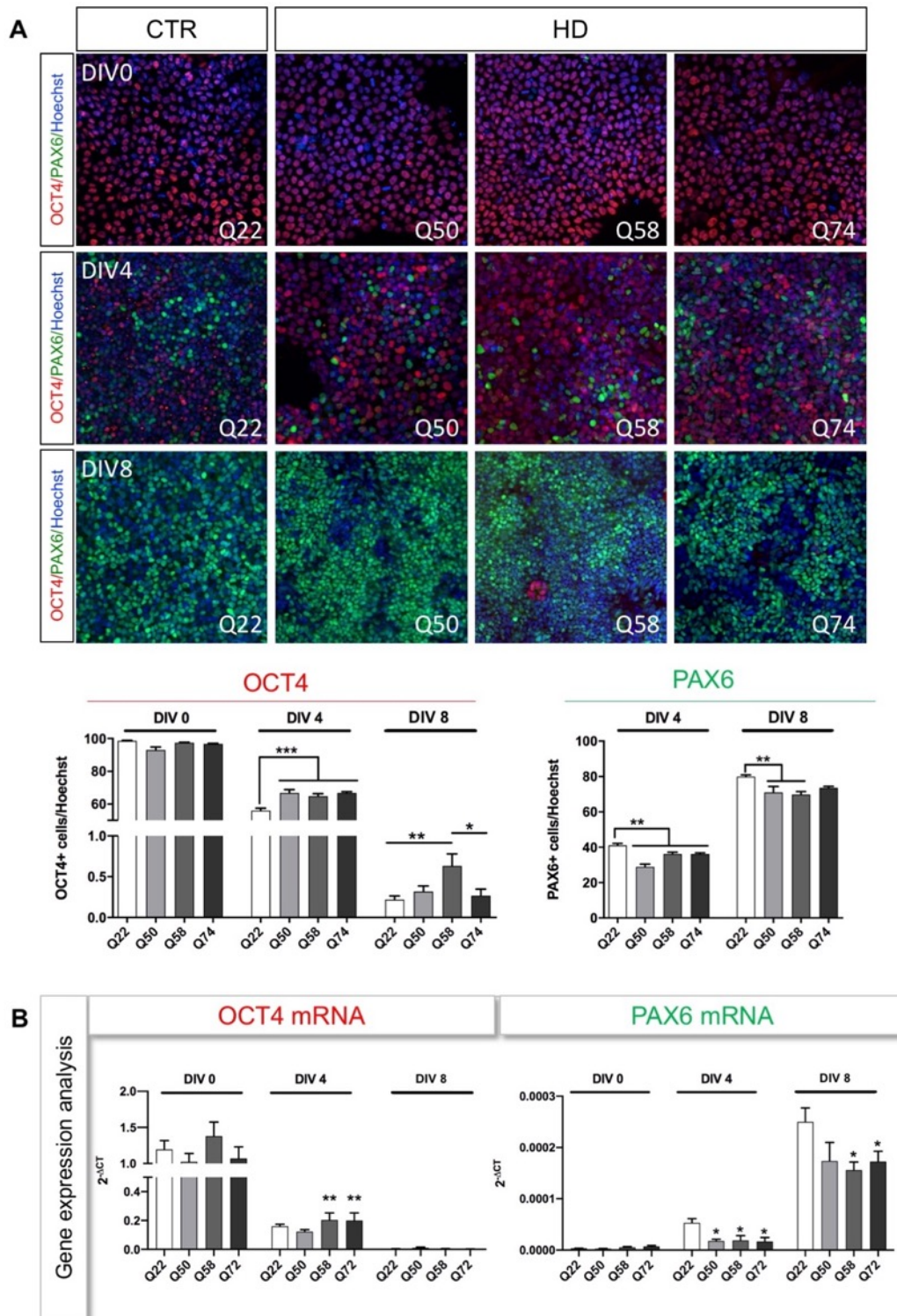
***Figure 2.3. Monitoring in the RUES2 series OCT4/PAX6 transition during the neural induction. (A)***
*Immunofluorescence analysis and respective quantification of RUES2 Q22 (average of two clones), Q50 (average of two clones), Q58 (average of two clones), Q74 lines at DIV0, DIV4 and DIV8 for the pluripotency marker OCT4 (red) and the neuroectodermal marker PAX6 (green). N=4 differentiation experiments. Images acquired by InCell 6000, 40X, N=20 images for each line/differentiation. 20 fields for well for a total of 160 images for line/time points. ANOVA one way, Tukey post-test, \*p<0,05; \*\*p<0,01; \*\*\*p<0,001. **(B)** qPCR analysis of RUES2 Q22 (average of two clones), Q50 (average of two clones), Q58 (average of two clones), Q74 lines at DIV0, DIV4 and DIV8 for OCT4 and PAX6. Total mRNA level normalized to 18s housekeeping transcript. Data are represented as mean ± SEM. ANOVA one way, Tukey post-test, \*p<0,05; \*\*p<0,01; \*\*\*p<0,001. N=4 differentiation experiments. RUES2 lines tested: N=8, RUES2 20CAG cl.30, RUES2 20CAG cl.66, RUES2 48CAG cl 16.3, RUES2 48CAG cl.20.4, RUES2 56CAG cl 21.1, RUES2 56CAG cl 25., RUES2 72CAG cl 12.*

A recent study from the Thompson laboratory observed an analogous OCT4 upregulation at the end of differentiation of HD iPSCs [96]. In this case, the authors identified the presence of 22% of OCT4$^+$ cells by immunofluorescence at DIV37 of differentiation. In addition, they performed RNA-seq analysis and they found a significant upregulation of OCT4 mRNA in HD lines, whereas all the other pluripotency markers were not dysregulated.

Overall, these data point out that muHTT interferes with the OCT4 gene transcription in pluripotency, and with the proper acquisition of a neuroectodermal fate.


To exclude that the defective OCT4/PAX6 transition could be only a faulty response to the striatal differentiation protocol, we evaluated the capacity of the RUES2 HD lines to correctly exit from pluripotency when they were exposed also to other protocols. Therefore, we decided to employ for this experiment the control Q22 and one representative HD line the Q58 to test three different "induction" conditions (Figure 2.4). First, in the condition A cultures were exposed to the conventional neural induction characterized by the dual SMAD inhibition (by addition of SB431542 and LDN to the medium). Second, in the condition B cells were induced to neuronal differentiation without the dual SMAD inhibitors, (referred in the figure as "no dual SMAD"). Finally, in the last condition C cultures were exposed to BMP4 and ACTIVIN to drive the cells toward a meso-endodermal fate acquisition mimicking the protocol of Teo et al. (2011). We performed immunofluorescence for each condition to monitor the transition from pluripotency to the various differentiation fates looking at OCT4 and PAX6 positive cells at DIV0, DIV4 and DIV8.

Cultures in condition A exhibited the same phenotype described before (Figure 2.3 A), where the control Q22 line correctly moved from pluripotency to neuroectoderm downregulating OCT4 and upregulating PAX6 (Figure 2.4 A left panel). Conversely, in the HD Q58 line there was still a consistent number of OCT4$^+$ cells and a limited population of PAX6$^+$ cells at DIV4. This reduction was maintained also at DIV8. After exposure to the second protocol (condition B) the resulting scenario partially mimicked the condition A. The absence of SB431542 and LDN delayed the transition also in the control Q22 line, since at DIV8 OCT4$^+$ cells were still detectable and a small population of PAX6$^+$ cells was observed.
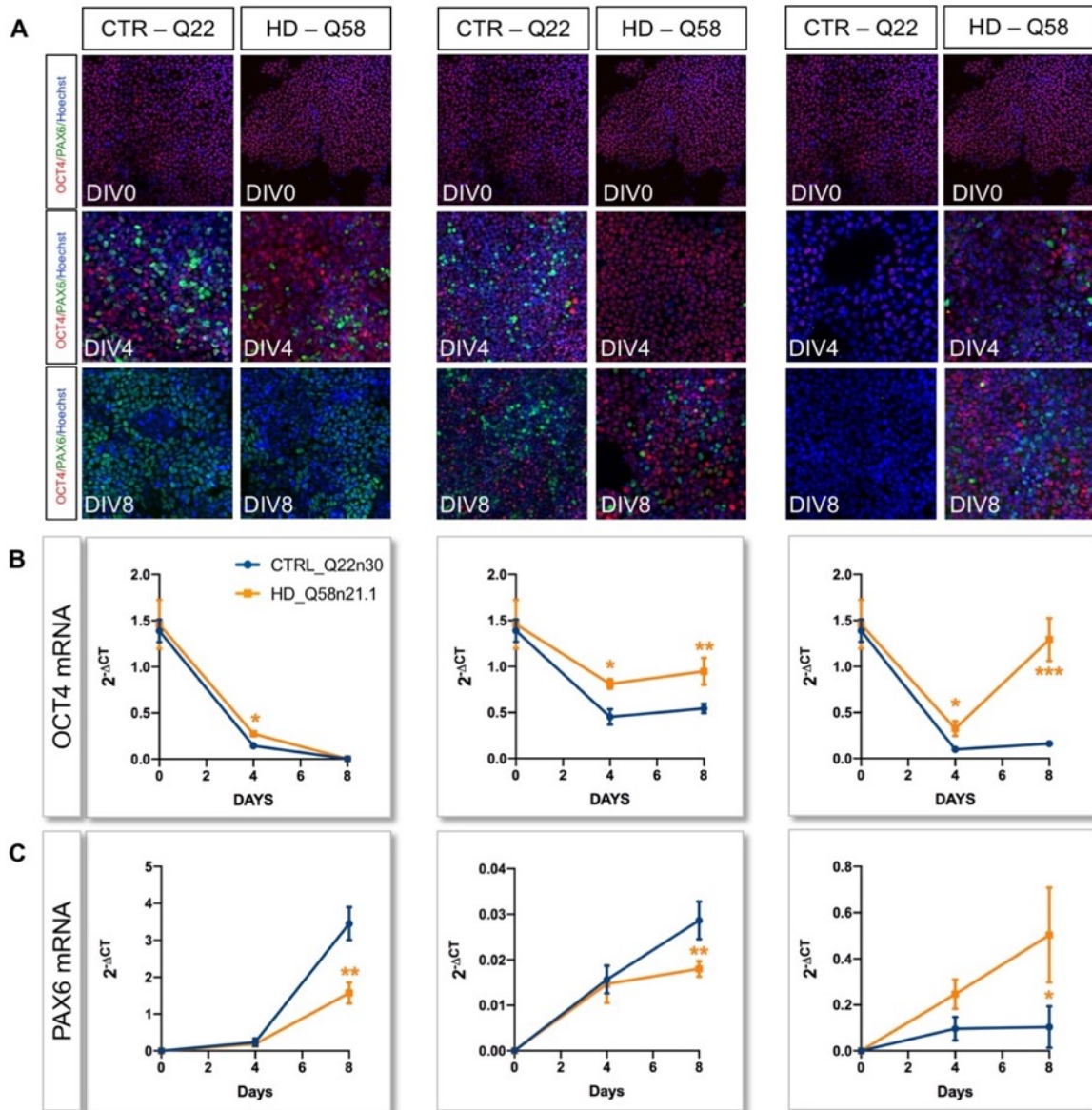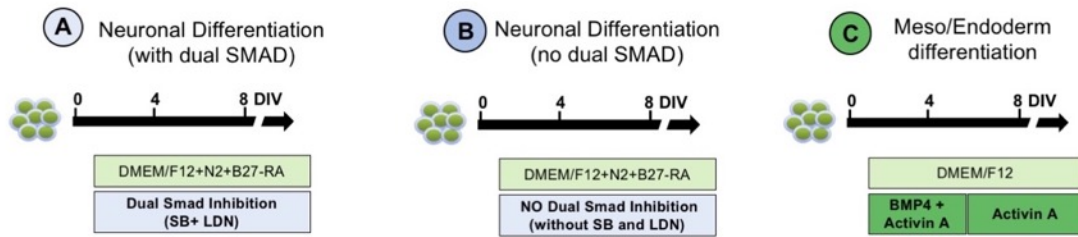
***Figure 2.4. Differentiation of control and HD RUES2 lines in different conditions to test the effect of OCT4/PAX6 transition.*** *On top are schematized the three differentiating conditions. Condition A: equal to normal dual SMAD inhibition, when SB431542 and LDN are added to the differentiation medium (DMEM/F12 with N2 and B27-RA). Condition B: in absence of dual SMAD inhibition but in presence of the differentiation medium (DMEM/F12 with N2 and B27-RA). Condition C: meso-endodermal differentiation by addition of BMP4 and ACTIVIN in DMEM/F12 medium. **(A)** Immunofluorescence analysis of control Q22 and HD Q58 lines at DIV0, DIV4 and DIV8 for OCT4 (red) and PAX6 (green). **(B)** qPCR analysis of control Q22 (blue) and HD Q58 (orange) lines at DIV0, DIV4 and DIV8 for OCT4 and PAX6 **(C)**. Total mRNA level normalized to 18s housekeeping transcript. Data are represented as mean ± SEM. ANOVA one way, Tukey post-test, \*p<0,05; \*\*p<0,01; \*\*\*p<0,001. N=2 biological replicates.*

In presence of muHTT, the phenotype observed in the condition A was worsened, as at DIV4 were present in the culture exclusively OCT4[+] cells, that prevailed also at DIV8 when just few PAX6[+] cells

emerged (Figure 2.4 A central panel). Finally, when the control Q22 culture was driven toward to the meso-endodermal fate by exposure to BMP4 and ACTIVIN (condition C), we detected a progressive reduction of the OCT4$^+$ population and complete absence of PAX6$^+$ cells, as expected from this protocol (Figure 2.4 A right panel). On the contrary, the Q58 line displayed a persistent population of OCT4$^+$ cells along the induction and an unexpected number of PAX6$^+$ cells from DIV4 of the differentiation.

In parallel to the immunofluorescence, we performed gene expression analysis for OCT4 and PAX6 transcripts. In the condition A, we detected a reduction of OCT4 mRNA both in control and HD line during the time, although at DIV4 the Q58 displayed significant higher expression of OCT4 compared to the control Q22 in line with the IC data (Figure 2.4 B left graph).

The same phenotype was shown by the HD line when subjected to neural differentiation without dual SMAD inhibition, as OCT4 mRNA was maintained upregulated both at DIV4 and at DIV8 in respect to the control line Q22 (Figure 2.4 B central graph). Finally, when the HD culture was exposed to the condition C the upregulation of OCT4 transcript was at DIV4 and finally exacerbated at DIV8 of the differentiation, in line with the data observed by immunofluorescence (Figure 2.4 B right graph).

Moreover, we evaluated the expression profile of the neuroectodermal marker PAX6 in the three conditions. Consistently with the previous data, we detected a significant reduction of PAX6 mRNA level at DIV8 in the Q58 line when the culture was exposed to the conditions A and B (Figure 2.4 C left and central graph). Lastly, in presence of BMP4 and ACTIVIN we measured an unexpected upregulation in PAX6 expression in the HD line, indicating defects in cell fate acquisition.

Overall, these data suggest that the persistency of OCT4 is not caused by the differentiation protocol employed but originate from the effects of muHTT on the correct transition toward a specific cell fate.

### 2.2.2 Known HD phenotypes are recapitulated by HD RUES2 during *in vitro* differentiation

Next, we looked at the effect of muHTT at subsequent stages of differentiation. Neural rosettes constitute a cellular arrangement that recapitulates *in vitro* the neural tube development [38]. The presence of smaller lumen size in HD-hiPSCs at DIV15 of striatal differentiation was described by Conforti et al. (2018). In addition, HD-RUES2 lines differentiated into cortical neurons by Ruzo et al. (2018) with a default neural induction protocol displayed neural rosettes with normal lumen size at DIV19.

We exploited the marker PALS1 and NCAD for the identification of the rosettes' lumen at the end of neural induction of *in vitro* striatal differentiation. By immunofluorescence analysis we observed in the Q50 and Q58 HD lines a significant reduction in the rosette lumen area (Figure 2.5 A), indicating an altered ability of the HD RUES2 lines to form these neural structures.
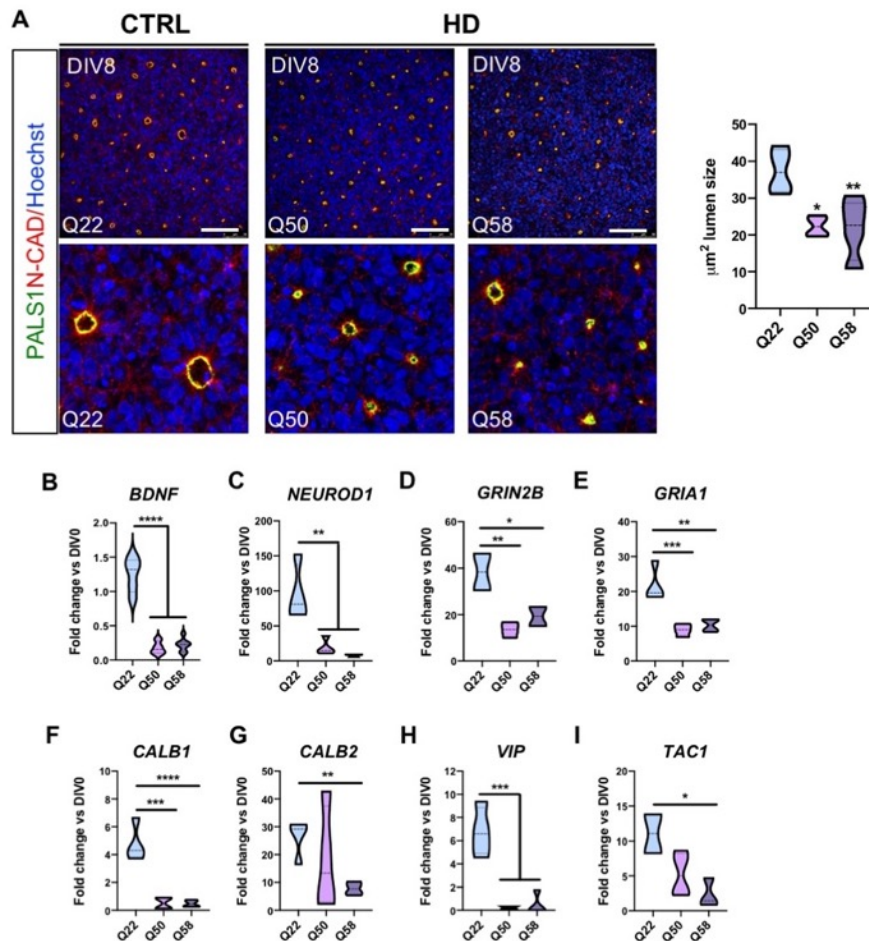


*Figure 2.5. Study of known HD phenotypes in control and HD RUES2 lines during the differentiation. (A) Immunofluorescence for the RUES2 control Q22, HD Q50, and HD Q58 lines of neural rosette formation at DIV8 by N-CAD+/PALS1+. Confocal images, 40X. Scale bar, 50 μm (crops of N-CAD+/PALS1+ of the same images). Counts of rosette lumen sizes by CellProfiler software (Version 2.1.1). ANOVA one way, Tukey post-test, \*p < 0.05, \*\*p < 0.01, one-way N = 3 biological replicates. (B) qPCR analysis for RUES2 control Q22, HD Q50, and HD Q58 lines (total mRNA level normalized to 18s housekeeping transcript) at DIV30 for transcripts known to be altered in mouse and cellular HD models, as BDNF (B), NEUROD1 (C), GRIN2B (D), GRIA1 (E), CALB1 (F), CALB2 (G), VIP (H) and TAC1 (I). Data are represented as mean ± SEM. ANOVA one way, Tukey post-test, \*p<0,05; \*\*p<0,01; \*\*\*p<0,001. N=3 biological experiments.*

Then, we tested whether neurons from RUES2 HD cells recapitulate known HD phenotypes. By high-content qPCR analysis we investigated whether transcripts involved in neurodevelopment and known to be affected in HD could be altered also in this *in vitro* system.

BDNF is a neurotrophin that controls fundamental functions in the CNS, as the regulation of synaptic maturation, activity and plasticity [235]. Moreover, BDNF controls striatal neurons survival and cortico-

striatal synapses activity [236]. In the cortex the expression of BDNF is impaired in absence of HTT and in presence of an expanded polyQ stretch [237-239].

In line with this evidence, neurons derived from HD RUES2 Q50 and HD Q58 displayed a significant downregulation of BDNF mRNA level compared to control line (Figure 2.5 B). Furthermore, HD-derived neurons exhibited a significant reduction in NEUROD1 mRNA, a neurodevelopment marker controlled by the REST/NRSF repressor complex. It is known that muHTT represses gene transcription of several neuronal genes associated with aberrant development and neurogenesis by REST/NRSF both *in vitro* and *in vivo*[41].

Equally, HD RUES2 Q50 and Q58 exhibited significantly reduced levels of NEUROD1 mRNA at DIV30 of differentiation area (Figure 2.5 C). The same downregulation was detected in Q50 and Q58 also for the developmental genes *GRIN2B* and *GRIA1* [225], together with CALB1 and CALB2, two calcium-binding protein expressed in the first stages of development [242], area (Figure 2.5 D-G). In addition, we measured also the expression level of several genes involved in the central nervous system development. Firstly, the neuropeptide vasoactive intestinal polypeptide (VIP) exhibited reduced level in HD Q50 and Q58 derived neurons, as reported in R6/2 mice [242] (Figure 2.5 H). Secondly, we looked at the Tachykinin Precursor 1 (TAC1), which is a striatal transcript whose expression is supported by BDNF [243]. Similarly, to the previous transcripts analyzed we observed reduced TAC1 mRNA level in our RUES2 HD lines (Figure 2.5 I). All these data indicate that also in this new isogenic cellular platform muHTT induces an impairment of MGE specification and terminal differentiation together with known HD phenotypes, such as persistency in OCT4 expression during neural induction [43], altered neural rosettes formation and reduced expression of the neurodevelopmental regulators NEUROD1 and BDNF [42].

## 2.3 Investigation of the early effects of muHTT on the epigenome

Delving into the first phase of the differentiation, the defective transition from pluripotency toward the neuroectodermal fate emerged as the first significant phenotype. After verifying that the persistency of OCT4 was not caused by a failure in protein degradation, we wondered whether the persistence of OCT4 and delayed expression of PAX6 during the neural induction might result from an altered epigenetic regulation. Mutant HTT interacting with proteins and TFs of the epigenetic machinery could induce early transcriptional alterations and affect neural development contributing later in the adult life to disease pathogenesis.

To identify a putative mechanism responsible for the defect observed in the HD RUES2 lines to exit from pluripotency and initiate the differentiation, we directed our attention to the epigenetic complexes involved in the regulation of chromatin during this crucial phase.

As abundantly explained in the paragraph 1.8 of the introduction section, there are several chromatin-modifying enzymes that binding to specific chromatin regions are permissive or repressive for transcription, as the Polycomb Repressive Complexes 1 and 2 (PRC1/2), the Trithorax group (TrxG), the Histone deacetylases (HDACs) and H3K9 methyltransferases (H3K9MTs).

Polycomb group of proteins (PcG) are fundamental in the regulation of lineage commitment of ESCs by controlling at specific time points the expression of a key set of developmental genes [125].

Particularly, knock out animals for the EED subunit (EED KO mESCs) of the PRC2 complex displayed maintenance of OCT4, NANOG and SOX2 expression along differentiation without affecting the pluripotency, further demonstrating the role of this complex in the exit from pluripotency [229].

Starting from these considerations, we decided to investigate potential alteration in the PRC2 complex in our cellular model along the neural induction.

### 2.3.1 Characterization of PRC2 during neural differentiation of RUES2 lines

Firstly, we evaluated the protein and expression level of the PRC2 subunits by western blot (WB) and qPCR. By WB we looked at the main molecular components of the PRC2 machinery, such as Jarid2, Suz12, Ezh2 and EED (Figure 2.6 A). For this analysis we employed the control line with Q22 (clones Q22cl30 and Q22cl66), the Q50 (clones Q50cl16.3 and Q50cl20.4), Q58 (clones Q58cl21.1 and Q58cl25.2) and finally the Q74cl12.
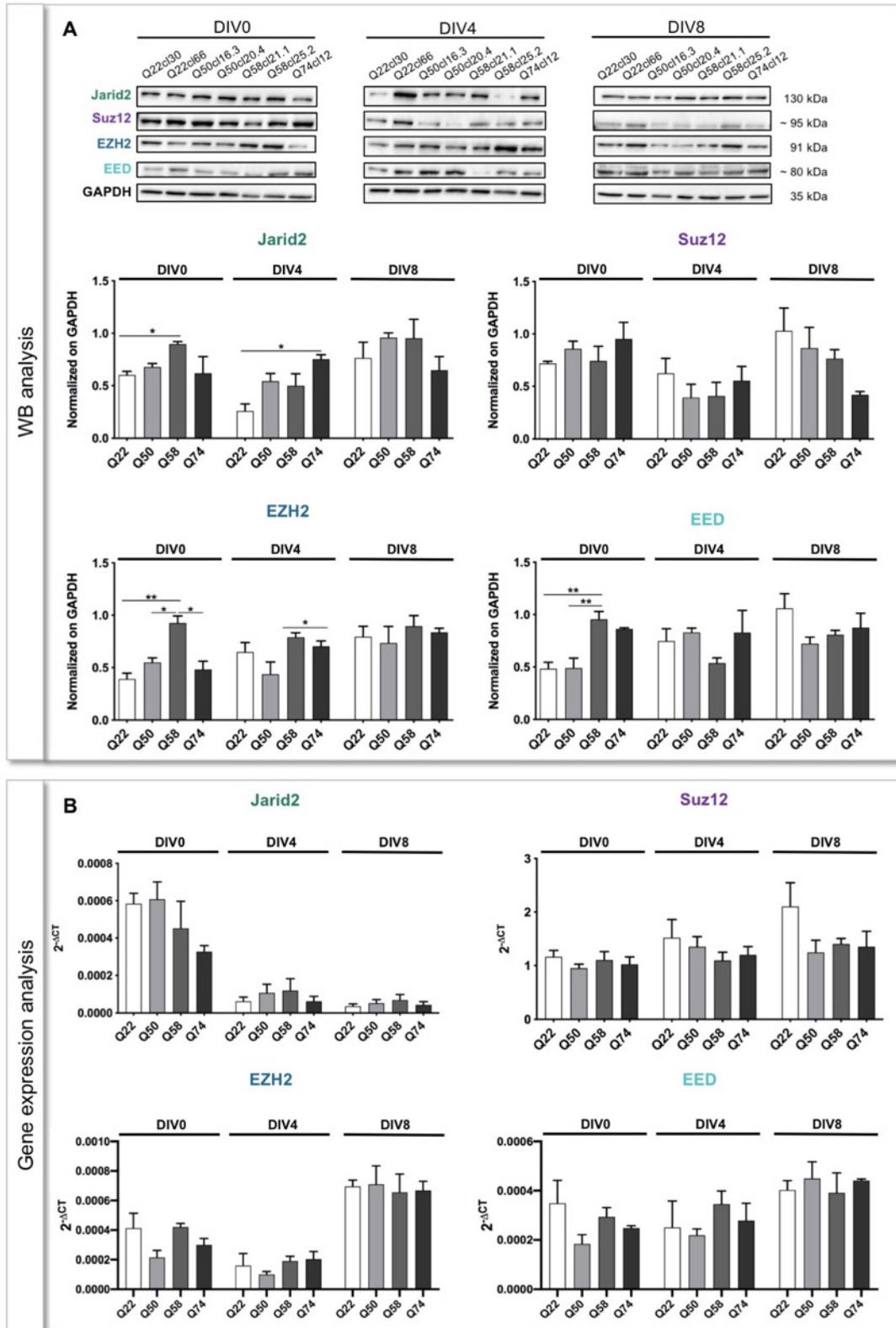
***Figure 2.6. Investigation of PRC2 component in the RUES2 series during neural induction. (A)*** *Western blot and relative densitometric analysis for the RUES2 Q22 (average of two clones), Q50 (average of two clones), Q58 (average of two clones), Q74 lines at DIV0, DIV4 and DIV8 of the main PRC2 subunits of the complex, as Jarid2, Suz12, EZH2 and EED.* ***(B)*** *qPCR analysis for the same lines reported in (A) at DIV0, DIV4 and DIV8 of the main PRC2 subunits of the complex, as Jarid2, Suz12, EZH2 and EED. Total mRNA level normalized to 18s housekeeping transcript. Data are represented as mean ± SEM. ANOVA one way, Tukey post-test, \*p<0,05; \*\*p<0,01; \*\*\*p<0,001. N=4 biological replicates.*

Globally the analysis showed a significant difference at the protein level in the HD Q58 lines at DIV0. In particular, the Q58 displayed higher protein level of Jarid2, EZH2 and EED compared to the controls Q22 (Figure 2.6 A). Upon the beginning of the neural induction, we did not observe significant differences both at DIV4 and at DIV8.

Next, we proceeded the characterization of the PRC2 machinery by gene expression analysis. During the neural induction we did not find any difference among the lines (Figure 2.6 B).

Overall, we did not observe major changes between the lines in protein and transcript level of the PRC2 machinery suggesting that the PRC2 complex structure per se is not affected by the presence of muHTT. This result is in line with the data reported by Biagioli et al. (2015), where they did not describe differences in the transcript level of PRC2 components.

## 2.3.2 Characterization of PcG foci in the nuclei of RUES2 during the differentiation

Besides the fundamental regulatory activity exerted by PRC2 during the first steps of neural development, several studies are pointing to understand how PRC2 and HTT are connected. First, normal HTT was found to facilitate the tri-methyltransferase activity of PRC2 and the timing and duration of this interaction is essential for normal embryonic development [49]. Second, in the presence of muHTT, PRC2 activity has been observed further enhanced [47]. Moreover, PRC2 deficient MSNs presented a delay in the expression of definite target genes coding for specific neuronal-type and other cellular transcriptional regulators advising that HTT might participate in PRC2 activity and that PRC2 is implicated in MSN survival and differentiation [54].

The PcG proteins associated to histone marks form inside the nucleus microscopically discrete foci named "PcG bodies" [230]. Inside these tridimensional structures, chromatin acquires unique packaging properties [231]. This 3D arrangement is mediated by PRC2, that after the trimethylation of the Lys27 on the histone H3 recruits PRC1 at these marked nucleosomes, leading to chromatin compaction. Particularly, the closure of the chromatin fiber is mediated by the generation of long-range interactions that are stabilized by both PRC1 and PRC2 [232].

Assuming that HTT might impact on PRC2 and PRC1 activity and therefore influence the crucial function of PcGs in regulating facultative heterochromatin, we monitored by immunofluorescence the number and the size of the PcG foci inside the nuclei of our RUES2 cellular system.

These experiments were performed in control Q22 and Q58 at DIV0, DIV4 and DIV8. We decided to test the Q58 line because it displayed the most distinct phenotype in the analyses described above.

To begin with, we started by characterizing the catalytic subunit EZH2 of the PRC2 complex, which is responsible for the trimethylation of the Lys27, the first step of chromatin compaction (Figure 2.7).
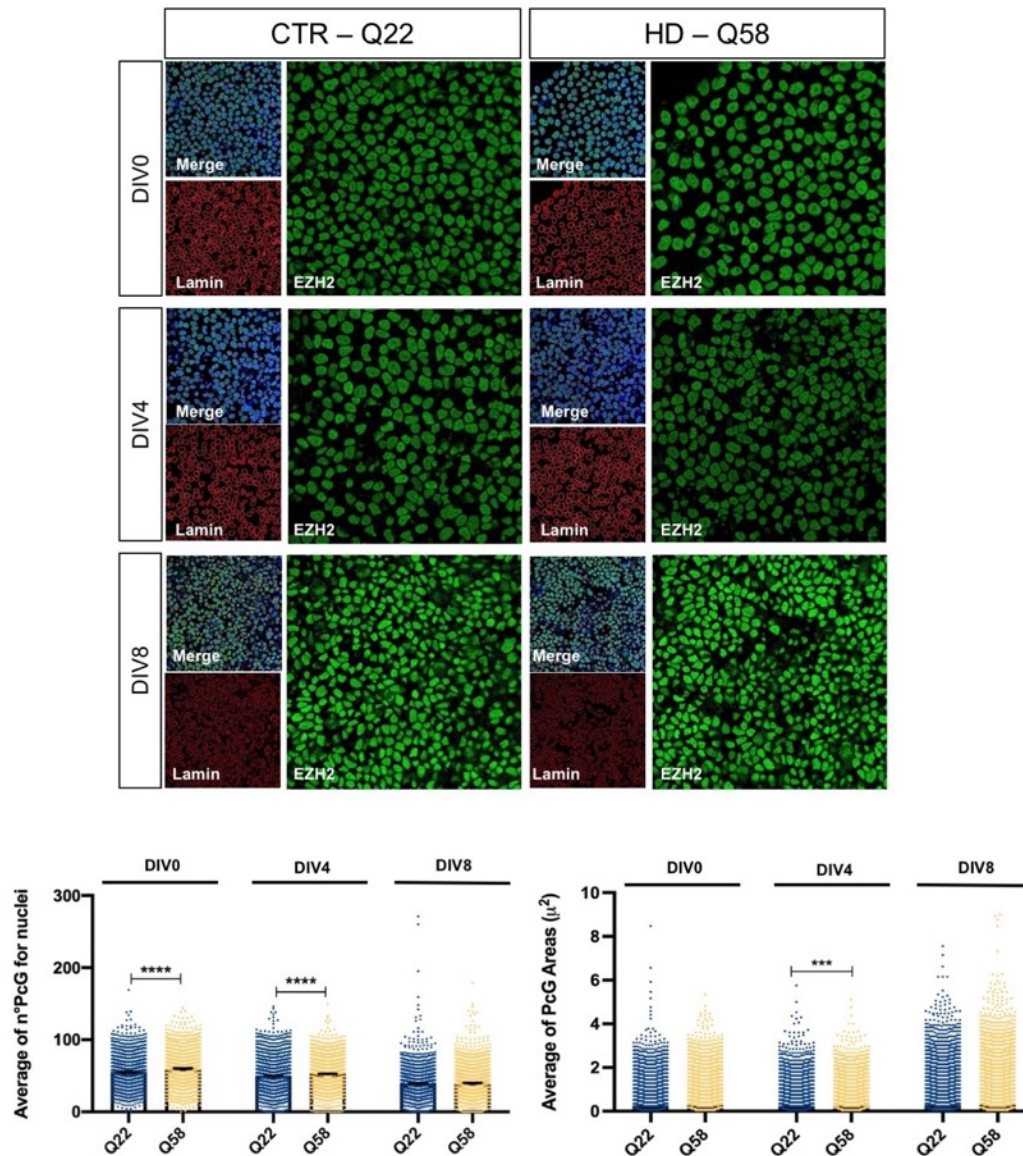


***Figure 2.7. Characterization in RUES2 lines of EZH2 PcG bodies during neural induction.*** *Immunofluorescence analysis and relative quantification of PcG area and number of PcG per nuclei for control Q22 and HD Q58 RUES2 lines at DIV0, DIV4 and DIV8 for EZH2 (green), LaminB1 (red) and Hoechst (blue). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For PcG number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For PcG Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1500-2000 counted nuclei per line.*

Thus, we performed immunofluorescent analysis for EZH2, coupled to LaminB1 that marks the nuclear lamina surrounding the nucleus. The LaminB1 together with the Hoechst staining allows the identification of the nuclei and afterwards of the PcG foci. This quantitative analysis was conducted by exploiting the automated pipeline published in Cesarini et al., (2015), (See Materials and Methods, paragraph 4.4.2). Thanks to this approach, we were able to evaluate the area of EZH2 dots inside the nucleus and the number of each spot marking the EZH2 proteins.

In normal conditions, the number of EZH2 foci is expected to reduce during the time, as with the differentiation is reinforced the repression of non-lineage specific genes, in parallel to the de-repression and finally the activation of lineage-specific genes.

Firstly, in our cellular system, we observed that in the Q22 line the number of PcGs decreased along the differentiation, from DIV0 to DIV8, whereas in the Q58 line it remained constant over time (Figure 2.7).

Regarding the area of the PcG foci we should expect instead an increase in their size during the time to stabilize the silencing of non-lineage specific genes. At DIV4 we correctly observed an increase in EZH2 foci dimension in Q22 cells, whereas the Q58 line displayed a significant reduction in the size compared to the control.


As formerly mentioned, the histones marked by PRC2 catalytic activity are recognized by the PRC1 complex that ubiquitinates the Lys119 on histone H2. Finally, PRC1 induces transcriptional repression mediating the *cis* and *trans* long-range interactions [149,155] and the multimerization through its SAM domains, that ultimately leads to the 3D formation of the PcG bodies [146,156].

Therefore, we performed immunostaining also for the PRC1 catalytic subunit, RING1, quantifying the area and the total number of these spots inside the nuclei (Figure 2.8). We stained the control Q22 and the HD Q58 line along the neural induction, at DIV0, DIV4 and DIV8. In this case both parameters related to RING1 changed their profile in presence of muHTT along the differentiation revealing more clearly a relation between the HTT protein and PcGs. The Q58 line in contrast to the control, displayed at DIV0 a significant lower amount of PcG bodies, and this number augmented during the time instead of being reduced. In addition, the pluripotent HD lines displayed also smaller foci in respect to the control and during the differentiation the size of RING1 bodies significantly increased in respect to the Q22.

Taken together the quantitative analysis for EZH2 and RING1 subunits we can identify two different profiles along the neural induction. Especially, in pluripotency we observed an increased amount of EZH2 in the Q58 line in parallel to a significant reduction for the RING1 subunit. Conversely, at the end of neural induction we did not observe differences for EZH2 but an augmented number and size of RING1 foci.
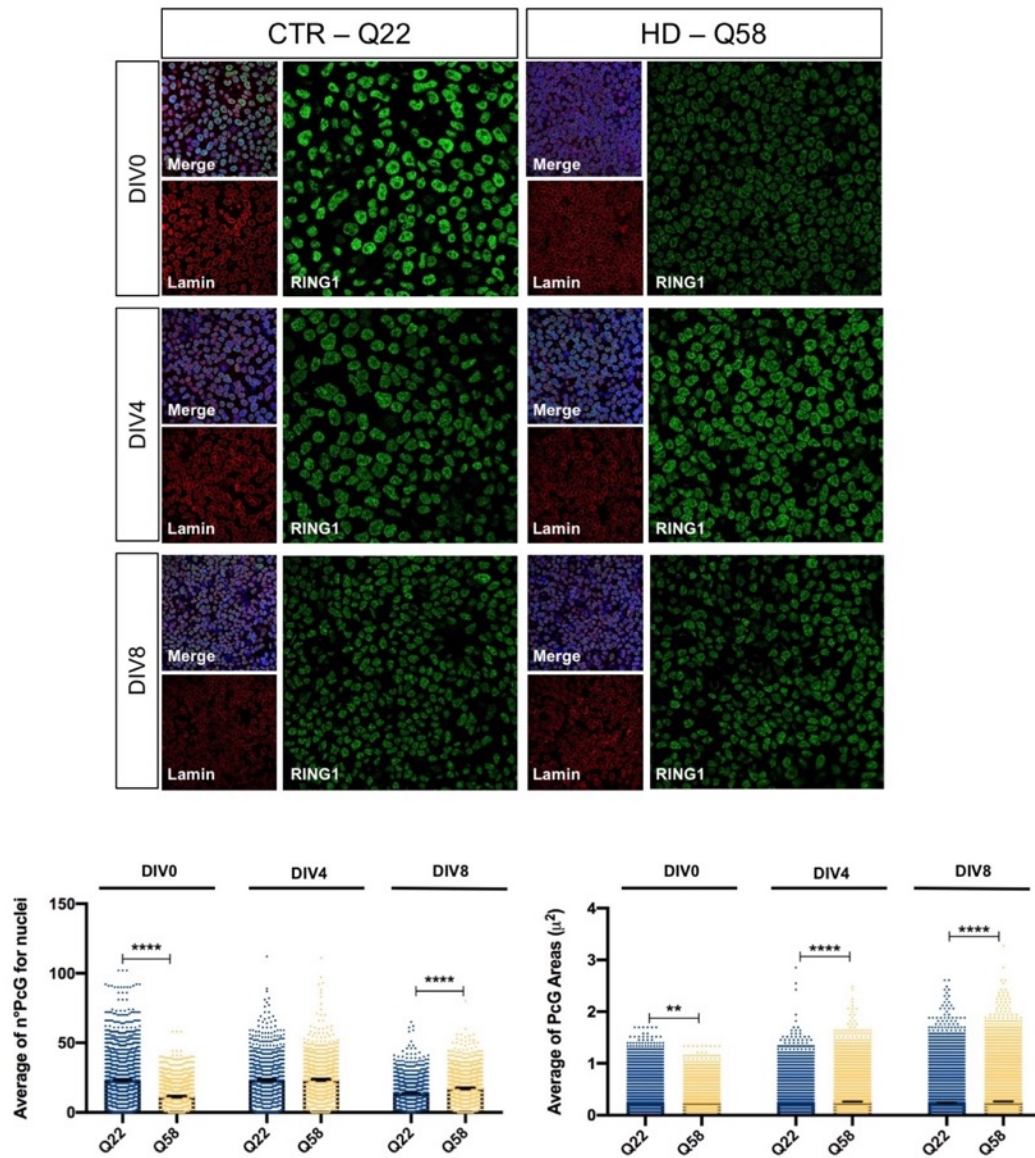
***Figure 2.8. Characterization in RUES2 lines of RING1 PcG bodies during neural induction.*** *Immunofluorescence analysis and relative quantification of PcG area and number of PcG per nuclei for control Q22 and HD Q58 RUES2 lines at DIV0, DIV4 and DIV8 for RING1 (green), LaminB1 (red) and Hoechst (blue). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For PcG number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For PcG Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p< 0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1500-2000 counted nuclei per line.*

These results suggest that muHTT impacts on the regulation of the facultative heterochromatin, influencing differently PRC1 and PRC2 complexes and therefore leading to a dysregulation of gene transcription.

## 2.4 Altered epigenetic regulation of OCT4 and PAX6 in HD RUES2 line

Since the above analyses revealed some differences linked to the PRC1 and PRC2 complexes, we wondered whether this could impact on chromatin conformation.

For these reasons, in collaboration with Dr. Chiara Lanzuolo (Institute of Biomedical Technologies (ITB) – CNR, Chromatin and Nuclear architecture Laboratory at INGM) we performed ChIP analysis for the H3K27me3 and H3K4me3 modifications on Q22 and Q58 in proliferation, at DIV4 and DIV8. The trimethylated Lys4 is generated by the Trithorax MLL2/COMPASS (KMT2B; WBP7) group, a complex of proteins that targets the same genes of PcG and antagonize its function. The H3K4me3 and H3K27me3 residues, related to gene activation and repression respectively, co-localize in regions defined as "bivalent domains" or "bivalent promoters". This peculiar epigenetic mark characterizes several genes involved in the regulation of development and lineage specification [178]. Notably, most of the bivalent domains are co-occupied by the PRC1 complex [197]. The simultaneous presence of activating and repressing marks keeps the promoter poised, meaning that is ready to be expressed and transcribed by the RNA polymerase II. Bivalent promoters are tightly interconnected in ESCs through long-range interactions mediated by PRC1 and PRC2 [234,235]. Trithorax MLL2 is crucial for the regulation of these long-range connections and maintains the genes in a transcriptionally accessible state [236]. The balance among Polycomb complexes and MLL2/ COMPASS regulates the transcriptional accessibility of bivalent genes ensuring a robust and well-timed induction of developmental gene expression which might be crucial for proper lineage specification during early development [236].

Remarkably, Biagioli et al., (2015) observed that in presence of muHTT the level of H3K27me3 decreased at bivalent loci in parallel to an altered enrichment of H3K4me3. Additionally, in the recent study from Thompson laboratory [106] was detected an upregulation of OCT4 expression at DIV37 of differentiation and by ChIP-seq they found that genes abnormally upregulated at the end of differentiation presented increased methylation of the Lysin 4 on Histone H3.

Therefore, we accomplished ChIP analysis for two biological replicates for these specific histone modifications, H3K27me3 and H3K4me3. We performed gene expression analysis by qPCR to test the efficiency of the chromatin immunoprecipitation for H3K4me3 and H3K27me3, checking the enrichment on the regulatory regions of two specific genes that we found altered in our cellular system, OCT4 and PAX6.

The expression of the POU5F1 gene promoter, coding for OCT4, is regulated by three upstream regions, the distal enhancer (DE), the proximal enhancer (PE) and proximal promoter (PP). The two enhancers (PE and DE) are activated differentially in a cell stage-specific manner.
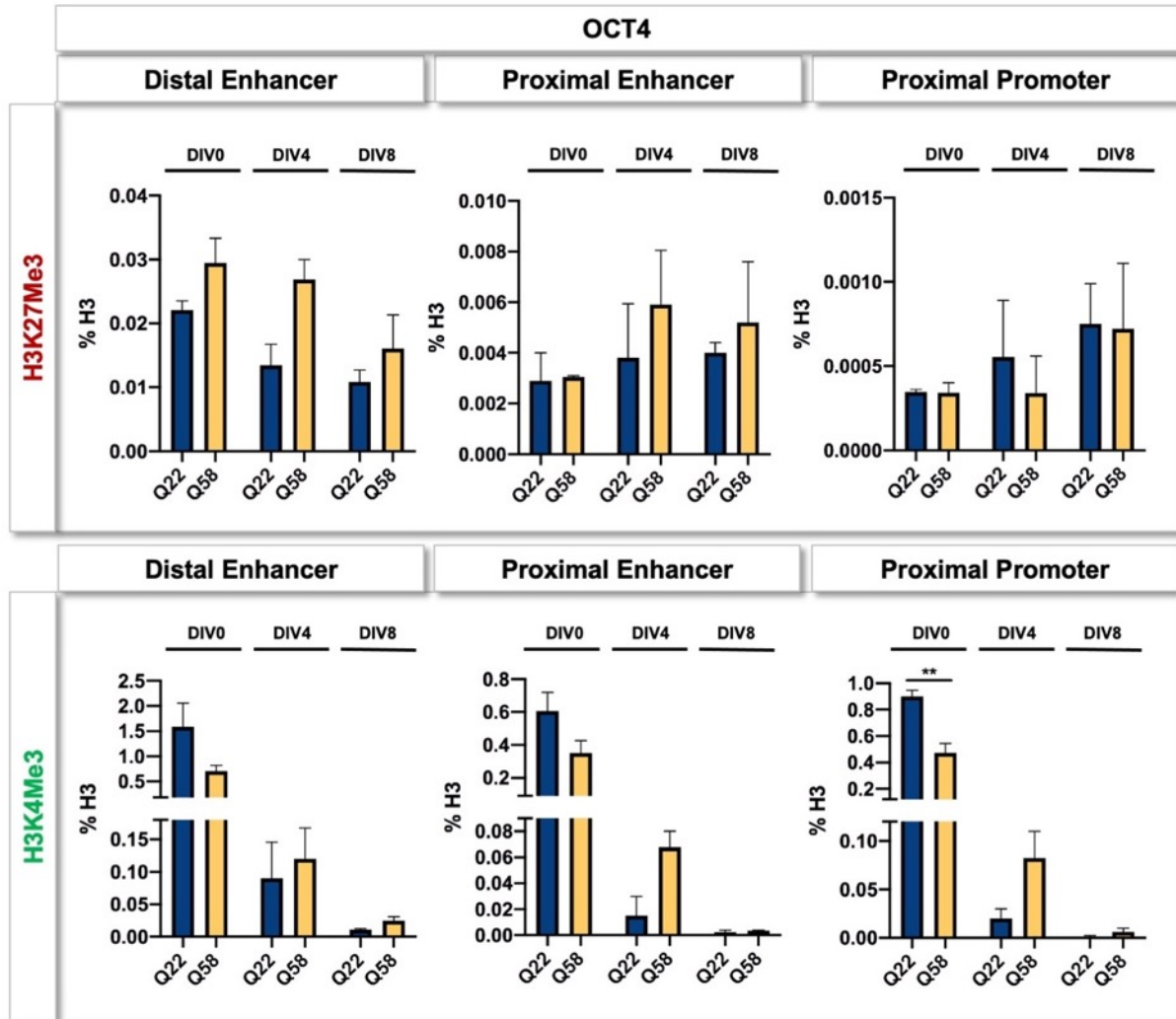


*Figure 2.9. Quality check by gene expression analysis for H3K27me3 and H3K4me3 ChIP on RUES2 lines of OCT4 regulatory regions. qPCR on Q22 and Q58 lines at DIV0, DIV4 and DIV8 on the three regulatory regions upstream the OCT4 promoter (PP, PE, and DE) both for the H3K27me3 and H3K4me3 ChIP. Data are normalized on H3/INPUT ratio. Data are represented as mean ± SEM. ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates.*

Starting from the ChIP for the H3K27me3, we observed both in the control Q22 and Q58 lines over time a reduced enrichment at the level of the distant enhancer (Figure 2.9 top panel). Regarding the proximal enhancer the H3K27me3 level remained stable in both lines, whereas in the proximal promoter gradually increased over time in control and HD lines.

Concerning the H3K4me3 mark, all three regulatory regions presented the same prompt reduction along the neural induction on the POU5F1 gene, therefore this modification seems to be responsible for the transcriptional drop occurring with the beginning of the differentiation. In pluripotency, the

HD Q58 lines showed in correspondence of the proximal promoter a significant decrease in H3K4me3 enrichment compared to the Q22. This reduction in presence of muHTT does not impact on transcription, whereas the delayed lowering in H3K4me3 enrichment at DIV4 in the Q58 line results in the persistent transcription of the OCT4 gene. At DIV8 the HD line seems to recover and decrease H3K4me3 enrichment. This trend characterized both the proximal and distal enhancer (Figure 2.9 bottom panel). To completely repress OCT4 expression is necessary the presence of the high H3K27me3 levels, therefore we could investigate whether later on during the differentiation are present altered level of this histone modification in the HD line.

Obviously, without the support of statistics on two biological replicates, we can only advance speculations and hypothesis from this gene expression analysis, as it represented a first quality control before the final sequencing.

Although we did not observe significant differences for the H3K27me3 ChIP on the upstream regulatory regions of OCT4, the profiles characterizing H3K4me3 enrichment along the neural induction seem to sustain the data previously described concerning the defective exit of pluripotency observed in the HD lines (See paragraph 2.2.1).

Based in these observations we will perform ChIP-seq analysis for both H3K27me3 and H3K4me3 on the control Q22 and HD Q58 lines to identify whether the expression of other genes might be dysregulated during the early stages of neural development. We will investigate not only the cells during the neural induction, at DIV4 and DIV8, but also in the pluripotent state.

Apart from the fact that in pluripotency the control and HD cultures appear to be identical, and we did not find differences at the transcriptional and protein level in respect to OCT4, the analysis performed on the PcG bodies displayed that these cells present epigenetic differences already at this time point for PRC1 and PRC2. Therefore, the reduction in H3K4me3 in the Q58 line at DIV0 might suggest the presence of an early altered epigenetic configuration that manifests later with alterations at the transcriptional and protein levels with the beginning of the differentiation. Moreover, the augmented trimethylation of the Lys4 at the level of the proximal promoter and enhancer detected at DIV4 is consistent with the persistent expression of OCT4 and higher protein level observed in the RUES2 HD lines (Figure 2.3). Finally, at DIV8 was still detectable enrichment of H3K4me3 in the HD line at the PP, PE, and DE being still in line with the immunofluorescence analysis of OCT4 in presence of the polyQ expansion (Figure 2.3B).

To further corroborate this hypothesis, after the ChIP for the trimethylated H3K27 and H3K4, we quantified by qPCR the enrichment of these histone marks on PAX6 promoter (Figure 2.10). From these results, we can appreciate that although the level of H3K27 is subtly reduced in pluripotent HD cells, this is sufficient to maintain silenced PAX6 gene, probably in association with lower H3K4me3

enrichment. At DIV8 the Q58 showed a significant increase in H3K4me3 enrichment (Figure 2.10), in contrast to the reduced transcription detected in presence of the mutation at this time point (Figure 2.3). From these data we can only assume that the presence of the polyQ expansion the epigenetic regulation is dysfunctional and impacts on H3K4me3 dynamic along the differentiation, as the profile in the control line changed and progressively decreased over time whereas in the HD line remained constant.
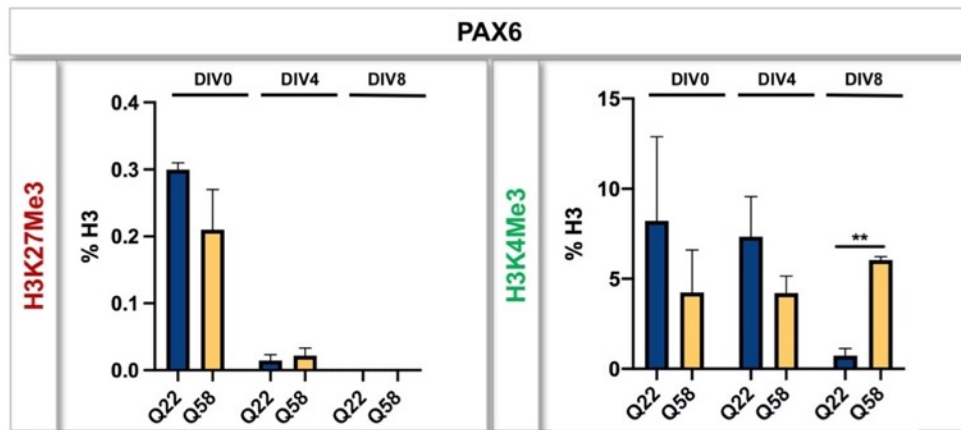


*Figure 2.10. Quality check by gene expression analysis for H3K27me3 and H3K4me3 ChIP on RUES2 lines for PAX6 promoter. qPCR on Q22 and Q58 lines at DIV0, DIV4 and DIV8 PAX6 promoter both for the H3K27me3 and H3K4me3 ChIP. Data are normalized on H3/INPUT. Data are represented as mean ± SEM. ANOVA One way, Tukey post-test, **p<0,01; ***p<0.001; ****p<0.0001. N=2 biological replicates.*

Taken together, the different H3K4me3 enrichment in the Q58 line on OCT4 and PAX6 regulatory regions might suggest the presence of abnormalities in chromatin conformation explaining the inability of these cells to respond correctly to external stimuli, compromising the exit from pluripotency and the acquisition of neuroectodermal fate commitment. In particular, muHTT does not appear to influence intrinsically the pluripotent stem cell state but it seems to alter the epigenomic environment, that once the differentiation process starts to lead to alterations at the transcriptional and protein level. Overall, we can suppose that in presence of muHTT cells present a dysfunctional epigenetic signature that could develop pathologically at a later time.

These data definitely need further investigation and the ChIP-seq analysis for these two crucial histone marks will point out whether are present others key genes with altered epigenetic regulation. This experiment could explain the phenotype observed during the neural induction and its perspective long-term effects.

# 2.5 Dysregulation of constitutive heterochromatin might be involved in the early defects observed in the HD RUES2 lines

The results achieved until now suggest that muHTT might interfere in the correct regulation of facultative heterochromatin, affecting PRC1, PRC2 and Trithorax complex activity.

The next step consisted in verifying whether the regulation of constitutive heterochromatin might also be affected by the polyQ expansion. This hypothesis is supported by evidence already present in the literature, considering that altered levels of the H3K9me3 histone modification have been found in post-mortem brains and in HD mice [56]. Moreover, abnormal enrichment of H3K9me3 and concomitant chromatin condensation were reported to lead to dysregulation of cellular mobility, neuronal differentiation, and synaptic transmission [60]. Normally in ESCs, this repressive mark is maintained at low levels and silences mostly genes related to neural differentiation. Finally, Irmak et al. (2018) revealed that alteration of this epigenetic conformation led to a defective neural induction.

### 2.5.1 Monitoring H3K9me3 repressive mark in RUES2 lines during the neural induction

Considering the data reported in literature, we decided to monitor the levels of the H3K9Me3 in our control Q22 and Q58 lines during the neural induction. To do this, we performed immunoassay for the H3K9me3 together with the nuclear protein LaminB1 and the Hoechst staining. As for the previous analyses on the PcG bodies, it was possible to quantify the area and the number of this nuclear staining. In the control line we observed a gradual reduction of H3K9me3 dots over time, whereas their size increased until DIV4 (Figure 2.11). Similar dynamics were displayed also by the Q58 line all along the neural induction, accompanied by a significant increase of H3K9me3 dots for nuclei, especially at DIV0, whereas the area of this dot augmented only at DIV8 in respect to the control line (Figure 2.11).

These preliminary data on H3K9me3 agrees with what has already been observed in literature, demonstrating that this cellular system can reproduce *in vitro* phenotypes that were observed *in vivo* and in other HD systems.

Furthermore, the HD RUES2 line seems to be different from the control already in pluripotency, not only at the level of facultative but also constitutive heterochromatin, thus highlighting the presence of a general altered epigenetic conformation. This could explain the different response to external stimuli and the early HD phenotypes that later impact on differentiation. Overall, these data are in line with our hypothesis of an early epigenetic dysregulation, that makes them more susceptible and vulnerable in the adult brain.
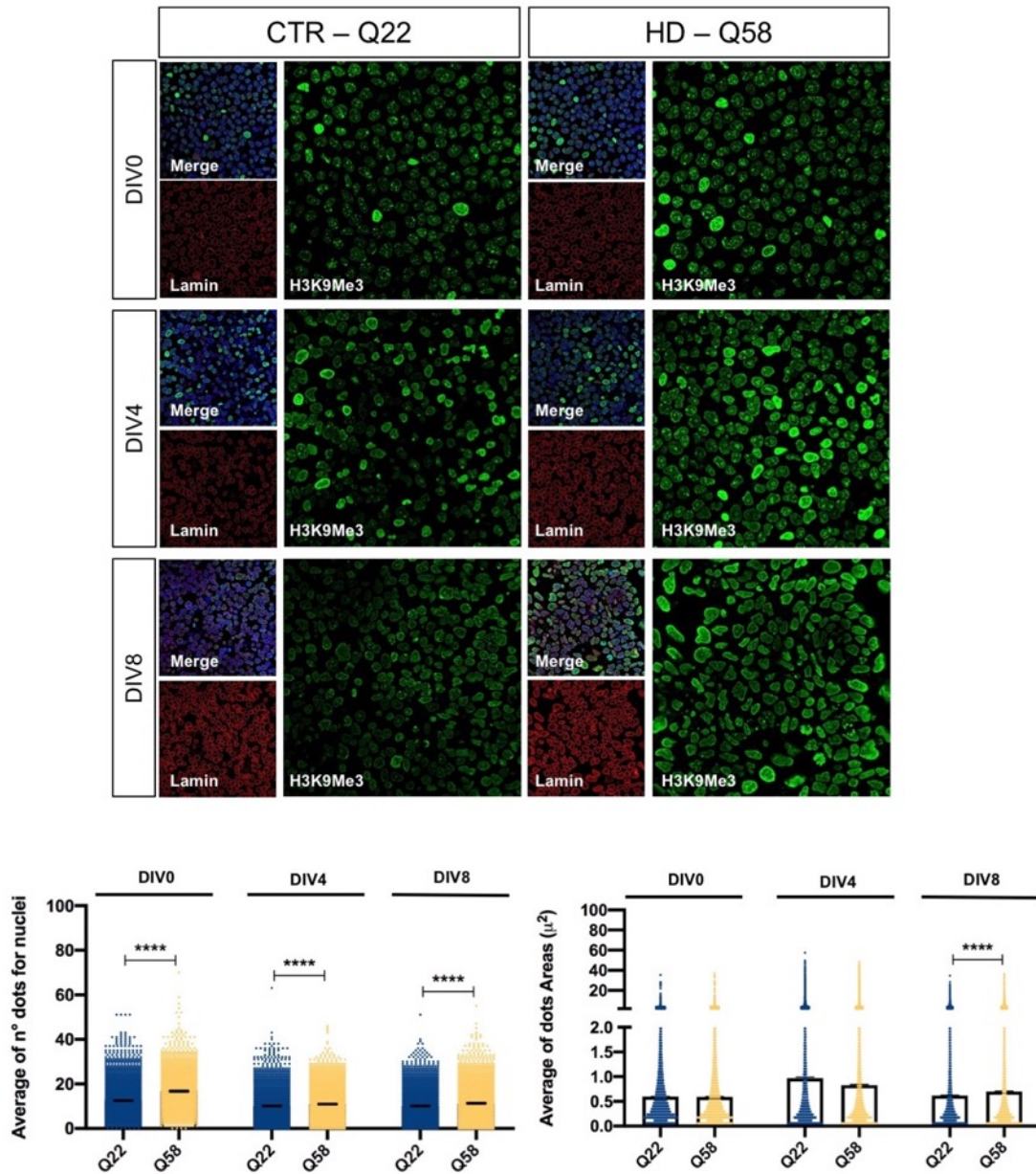
***Figure 2.11. Monitoring in RUES2 lines H3K9me3 histone modification during neural induction.***
*Immunofluorescence analysis and relative quantification of dots area and number of dots per nuclei for control Q22 and HD Q58 RUES2 lines at DIV0, DIV4 and DIV8 for H3K9me3 (green), LaminB1 (red) and Hoechst (blue). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For H3K9me3 dots number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For H3K9me3 dots Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1500-2000 counted nuclei per line.*

## 2.5.2 Investigation of H3K9me3 impact at the genome level by ChIP analysis

After this characterization, we decided to perform ChIP analysis also for the H3K9me3 histone mark on our target genes, OCT4 and PAX6 at DIV0, DIV4 and DIV8 (Figure 2.12).
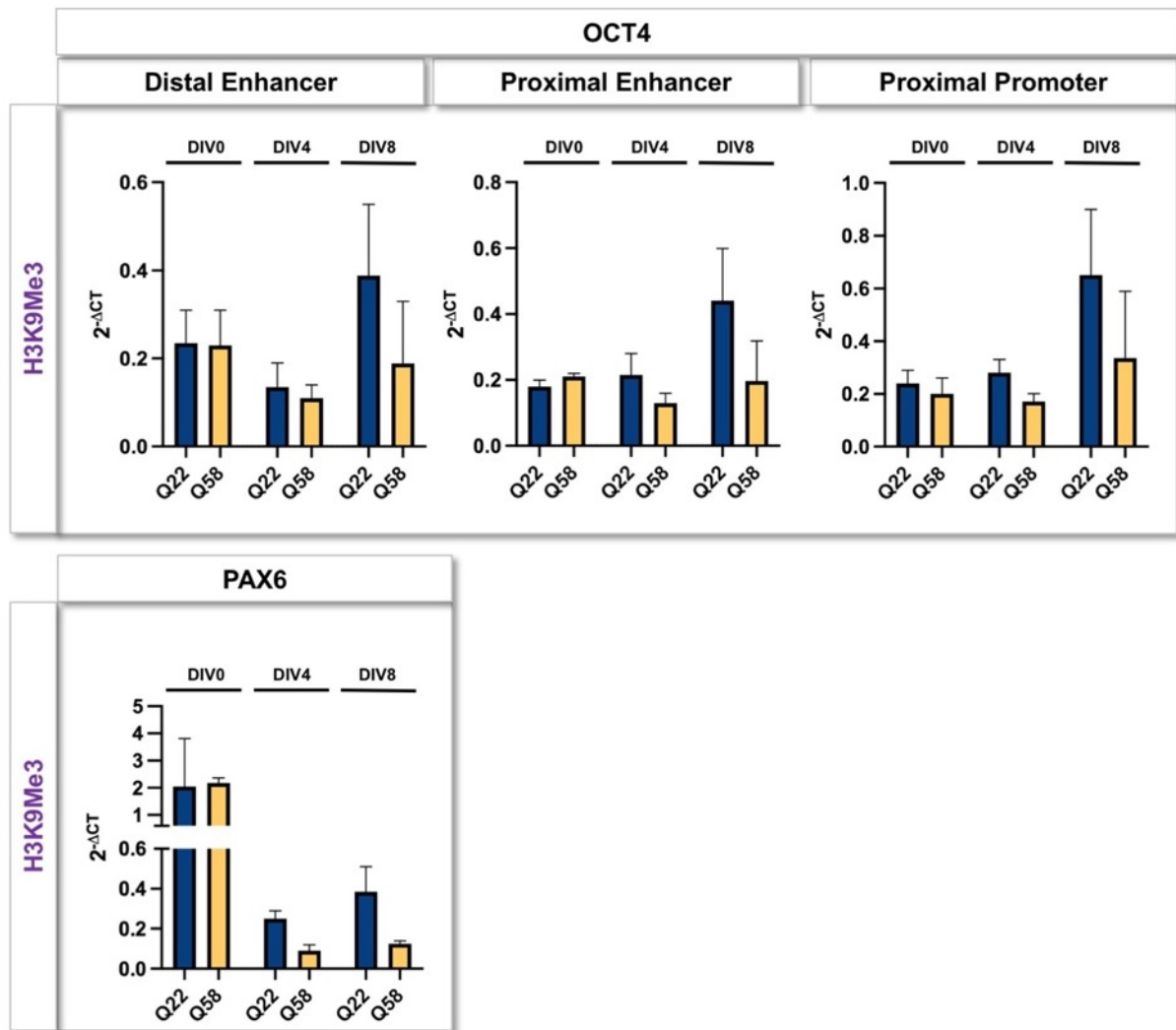


*Figure 2.12. Quality check by gene expression analysis for H3K9me3 ChIP on RUES2 lines for OCT4 and PAX6 regulatory regions. qPCR on Q22 and Q58 lines at DIV0, DIV4 and DIV8 on the three regulatory regions upstream the OCT4 promoter (PP, PE, and DE) and PAX6 for H3K9me3 ChIP. Data are normalized on H3/INPUT. Data are represented as mean ± SEM. ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates.*

Starting from OCT4, we did not observe significant differences among the lines in self-renewal and at DIV4. At DIV8 we found an increase of the H3K9me3 enrichment on all three regulatory regions. Apparently, the increase of H3K9me3 seems to be more pronounced in the control compared to the HD line, that seems not to be able to generate constitutive heterochromatin in presence of muHTT. This pattern reflects the transcriptional profile of OCT4, since it must be definitively switched off at the end of neural induction and the H3K9me3 could contribute to its repression by inducing chromatin compaction.

Regarding the PAX6 promoter, we correctly observed both in the control and HD line an opposite profile during the time, as upon the beginning of the dual SMAD inhibition the level of H3K9me3 enrichment dropped drastically at DIV4 and DIV8 in respect to DIV0 (Figure 2.12).

Considering the differences observed by immunofluorescence and the ChIP data on OCT4 and PAX6, the global ChIP-seq analysis will help in identifying which genes may be under the control of an altered repressive H3K9me3 profile, together with H3K27me3 and H3K4me3. Consequently, it will be possible to verify whether the regulation of facultative and/or constitutive chromatin is altered in HD and hence could contribute to the onset of the early phenotypes observed *in vitro* and later to the pathogenesis of the disease.

## 2.6 Effects of loss of HTT on the epigenome

Since in the presence of muHTT we observed variations in the number and size of PcG bodies together with variations in H3K9me3 levels, we wanted to investigate the effects of the total absence of wild type HTT on the epigenome.

First of all, it is necessary to underline that the KO line, as the HD RUES2 lines, displayed a persistence expression of the OCT4 transcript and higher protein level at DIV4 and DIV8 (data not shown). Moreover, in absence of wild type HTT we detected a delayed expression of PAX6, as well as reduced protein level, at DIV4 and DIV8 (Figure 2.3), (data not shown).

Next, we evaluated the number and size of EZH2 foci during the neural induction (Figure 2.13). In absence of HTT we observed progressive lowering of EZH2 foci over time as in the control, although at DIV0 the KO presented a significant higher number of PcGs, as observed for HD line. Regarding the PcG size, the KO line did not display a gradual increase over time, as for the control and HD line, and only at DIV8 we quantified significantly bigger PcG bodies compared to the HD line (Figure 2.13).
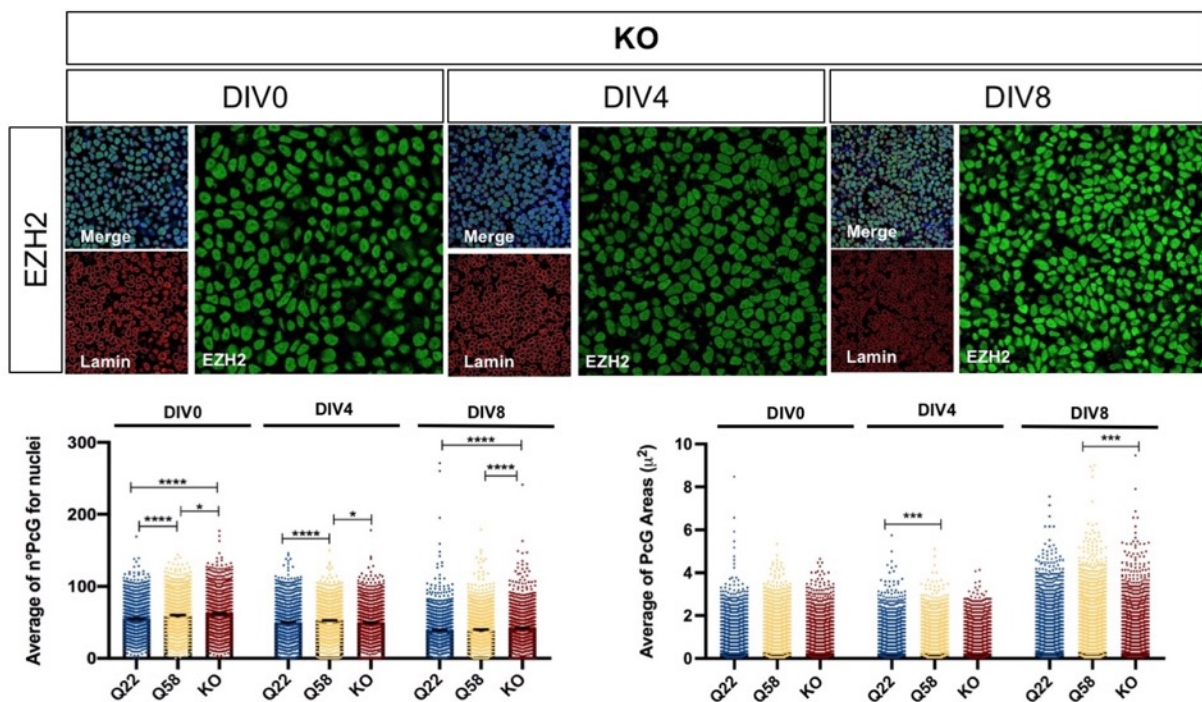


***Figure 2.13. Monitoring in RUES2 KO line EZH2 PcGs during neural induction.*** *Immunofluorescence analysis and relative quantification of dots area and number of dots per nuclei for KO RUES2 line at DIV0, DIV4 and DIV8 for EZH2 (green), LaminB1 (red) and Hoechst (blue). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For PcG number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For EZH2 dots Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1500-2000 counted nuclei per line.*

Concerning the PRC1 complex, the KO line presented fewer RING1 foci in respect to the control line all along the differentiation, from pluripotency to DIV8 (Figure 2.14), revealing a completely

different dynamic from the previous one of EZH2 (Figure 2.13). The main difference emerged at DIV0, where in absence of HTT the number of PcG quantified was significantly lower compared to the control but higher than the mutant Q58 line. Related to the PcG area, the KO line, as the control and mutant Q58 line, displayed an increase in PRC1 size until DIV4, whereas at DIV8 presented significantly smaller bodies in respect to the control line.
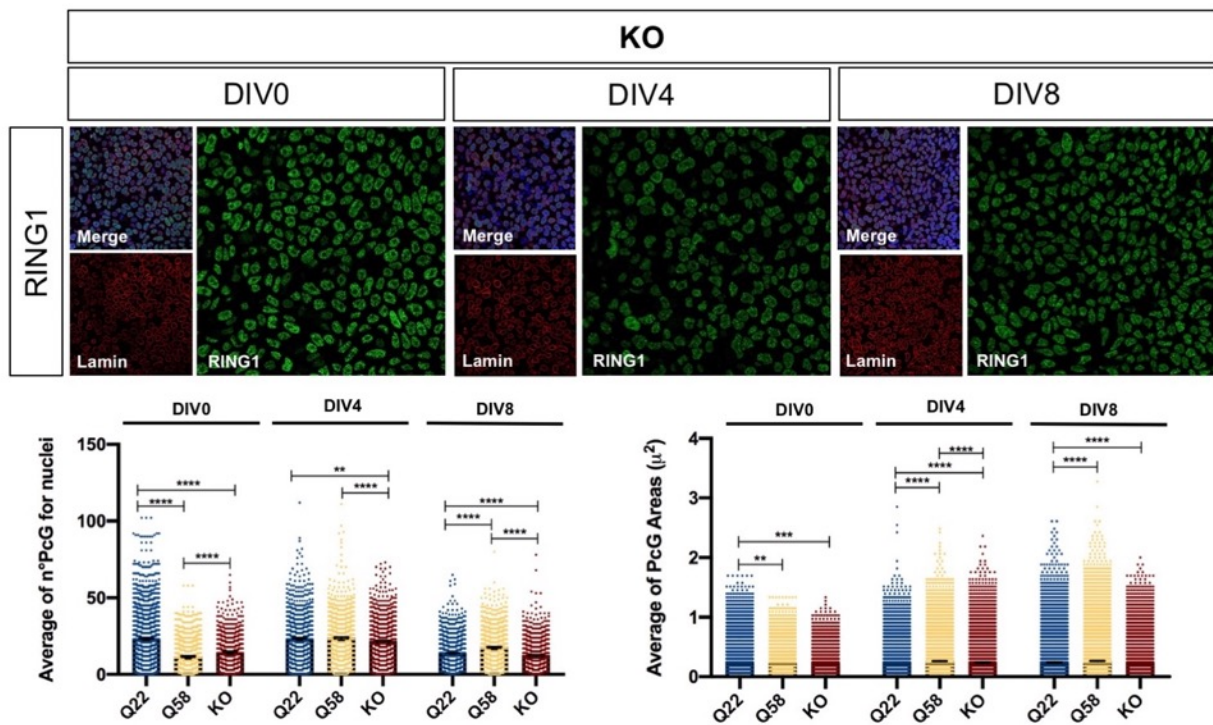


*Figure 2.14. Monitoring in RUES2 KO line RING1 PcGs during neural induction. Immunofluorescence analysis and relative quantification of dots area and number of dots per nuclei for KO RUES2 line at DIV0, DIV4 and DIV8 for RING1 (green), LaminB1 (red) and Hoechst (blue). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For PcG number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For RING1 dots Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1500-2000 counted nuclei per line.*

Lastly, we evaluated the H3K9me3 histone modification in the KO line (Figure 2.15). In relation to the number of H3K9me3 dots, the absence of wild type HTT caused a progressively reduction over time as the control line. The main difference emerged at the pluripotent state when the KO line, as in presence of the polyQ expansion, displayed significantly higher number of H3K9me3 dots in respect to the control. While in the HD line the number remained higher throughout the differentiation, in the KO line at DIV4 and DIV8 decreased significantly. Also, in relation to the dots area, the profile of the KO line is in line with the control Q22, as the size increased until DIV4 and later is reduced. At DIV0 and DIV4 the cells displayed bigger dots compared to the Q22 line, whereas at DIV8 the area decreased significantly in respect to both control and HD lines.

In conclusion, based on these results it is difficult to determine whether the phenotypes observed can be classified as gain or loss of function, because the profile of both the HD and KO lines changes

during the time. We can hypothesize that HTT participates in the regulation of facultative and constitutive heterochromatin since the normal pattern is altered both in absence of wild type HTT or in presence of the mutant form.
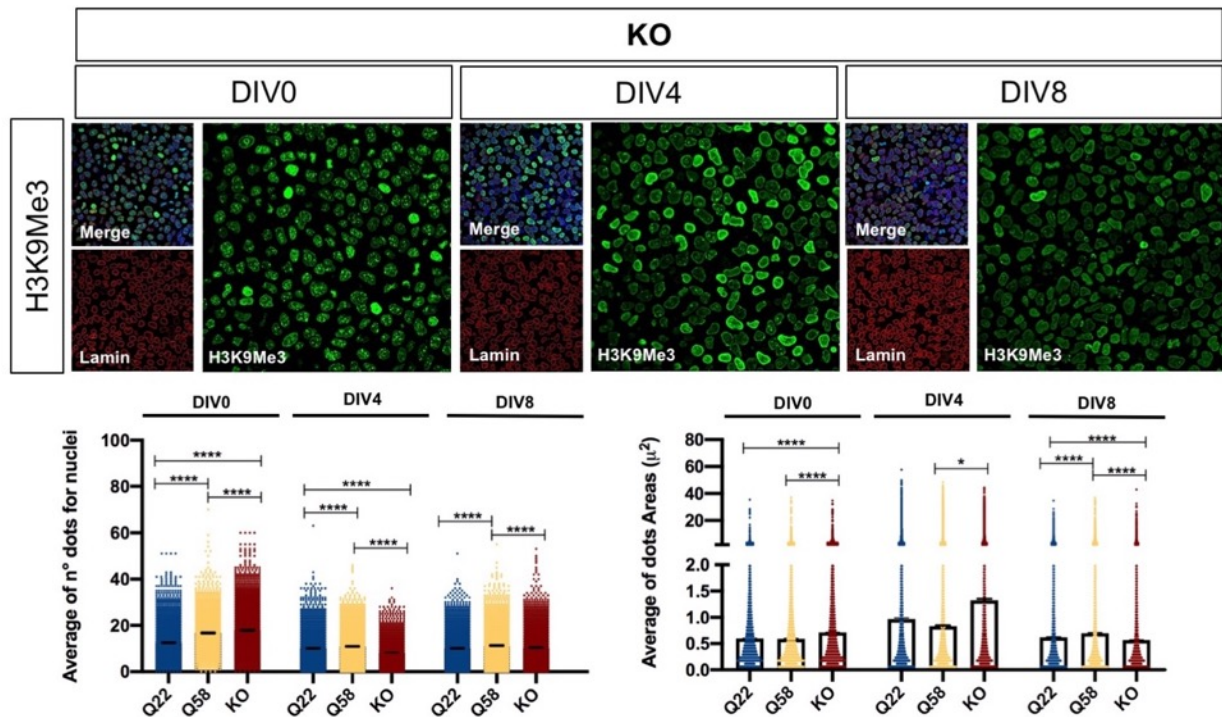


**Figure 2.15. Monitoring in KO RUES2 line H3K9me3 histone modification during neural induction.**
*Immunofluorescence analysis and relative quantification of dots area and number of dots per nuclei for KO RUES2 line at DIV0, DIV4 and DIV8 for H3K9me3 (green), LaminB1 (red) and Hoechst (blue). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For H3K9me3 dots number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For H3K9me3 dots Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1500-2000 counted nuclei per line.*

## 2.7 Characterization of nuclear morphology in RUES2 lines

The nucleus is an essential part of the cell, since it houses the entire genome and maintains its three-dimensional structure, thus controlling gene transcription and therefore cell behaviour [237]. The structure of the nuclear envelope is determined by the nuclear lamina, that connects the chromatin domains to the nuclear periphery and determine the position of nuclear envelope proteins (Stephens et al., 2018). Moreover, nuclear integrity and rigidity is also determined by the histone modification state and the major contributing factors are chromatin and lamins.

The study of Gasset-Rosa et al., (2017) has highlighted some age-related cellular characteristics in the cortex and striatum of HD mouse models, which displayed a decrease in nucleo-cytoplasmic transport, a reduction of integrity of the nuclear envelope and an accumulation of DNA double-strand breaks. Especially in these brain regions, these phenotypes were markedly accelerated in a dose- and age-dependent manner, as muHTT accumulating in the nuclei led to the disruption of the nuclear envelope and partially sequestered factors essential for the nucleocytoplasmic transport.

For these reasons, we aimed at identifying whether the general alterations in facultative and constitutive heterochromatin observed in our HD and KO RUES2 line could impact on nuclear morphology or they are the result of alterations in the nuclear structure.

As mentioned before, PcG foci quantification was determined by the combination of the nuclear staining LaminB1 and Hoechst. These two stainings together allowed the characterization of two parameters, the area, and the circularity of the nuclei.

We performed this quantification during the neural induction (DIV0, DIV4 and DIV8) for the control Q22, Q58 and KO lines in parallel to the analysis of the PcG bodies for EZH2 and RING1 and dots for the H3K9me3.

Normally, during the differentiation cells tend to become smaller and less circular. We observed exactly this profile in the control line, whereas in absence of wild type HTT we did not detect variations during the time. Interestingly, at DIV0 the nuclei of Q58 line were more circular compared to both the control and KO line, and this difference was lost over time (Figure 2.16). Considering the nuclear size parameter, at DIV0 both the HD and KO lines showed bigger nuclei compared to the control and this difference was maintained at the end of neural induction at DIV8 (Figure 2.16). Overall focusing on the pluripotency data, we detected, alterations in PRC1 and PRC2 foci, in constitutive heterochromatin and finally on nuclear morphology, both in absence of wild type HTT or in presence of muHTT.
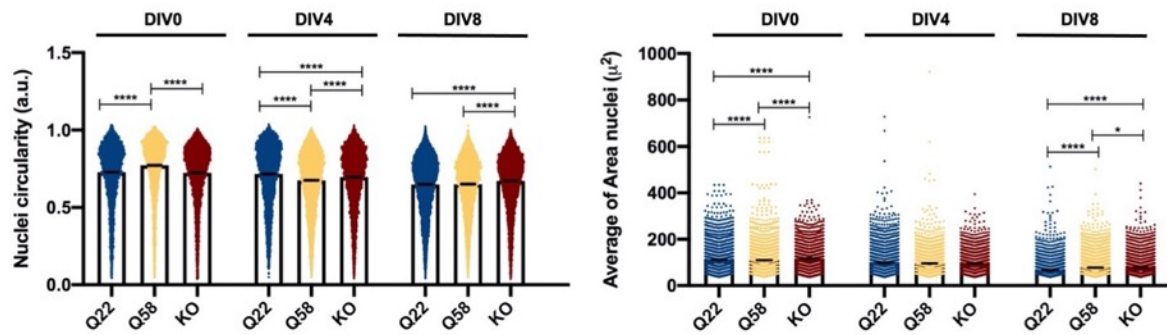
***Figure 2.16****. **Characterization of nuclear morphology of RUES2 lines along neural induction.*** *Immunofluorescence analysis and relative quantification of nuclear circularity and nuclear area for control Q22 and HD Q58 RUES2 lines at DIV0, DIV4 and DIV8. Analysis performed on LaminB1 (red) and Hoechst (blue) signals of images represented in Fig.2.7/2.8/2.13 for the control Q22 and HD Q58 lines, and Fig. 2.15/.2.16/2.17 for the KO RUES2 line. Box plot data are represented as mean ± SEM. For PcG number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For PcG Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 9.000-10.000 counted nuclei per line.*

All together these phenotypes are typical of senescent cells, therefore we wondered whether muHTT or HTT loss could activate an early senescence program in our HD and KO RUES2 lines.

Although senescence has long been linked with aging, recent studies have revealed roles in embryonic development, regeneration and disease [239–241]. Moreover, there is evidence in literature linking senescence to PRC2 activity, chromatin conformation and nuclear morphology [256].

Based on this knowledge, we could assume that muHTT or HTT loss could activate an early senescence program in our HD and KO RUES2 lines. This observation needs to be further corroborated.

## 2.8 Could the early phenotypes impact on the adult neuron?

All the experiments performed up to this point were conducted during the first phase of differentiation, the neural induction. This early time window has been selected with the aim of identifying the origin of the defective OCT4/PAX6 transition in presence of muHTT, that could be responsible for the faulty specification and terminal differentiation observed in the HD RUES2 lines.

Furthermore, we wanted to investigate the hypothesis that HD could originate at the beginning of development. Mainly, we hypothesize that muHTT causes an altered epigenetic conformation early in development and this could impact on the response to the differentiation stimuli and therefore contribute to the pathogenesis of the disease. Therefore, premature epigenetic alterations might ultimately lead to the formation of faulty striatal neurons, which in adulthood would be more fragile and vulnerable and then undergo degeneration.

For this purpose, we characterized our Q22, Q58 an KO cell lines also at the end of the differentiation, at DIV40, performing the same type of analyses and quantifying the parameters related to PcG foci considered during the neural induction.

Firstly, looking at EZH2, the catalytic subunit of PRC2, we observed a significant increase in PcG area in the HD Q58 compared to the control (Figure 2.17). At DIV40, in absence of HTT occurred a significant reduction in PcG bodies number compared to both HD and control lines. Notably, this result is totally different from the previous data obtained at the beginning of the differentiation, as the number of EZH2 foci was higher both in the HD and KO lines (Figure 2.7-Figure 2.13). Although these data are not easy to interpret, we might suppose that during the differentiation was selected a subpopulation of cells inside the HD and KO culture, that over time has tried to reach the same epigenetic condition of the control line.

In addition, we also checked the subunit RING1 to analyze the foci of the PRC1 complex. In this case, in presence of muHTT and in absence of the wild type HTT protein the area and the number of PcG bodies significantly increased compared to the control Q22 line (Figure 2.17). Also, in this case these results are in complete opposition to what observed in pluripotency, where the number and size of foci were significantly lower in the HD Q58 and KO lines compared to the control (Figure 2.8-Figure 2.14). Here as well, we can hypothesize that a group of cells in the culture of HD and KO lines might have been selected during the differentiation, explaining the opposite profile observed at the two extreme time points.
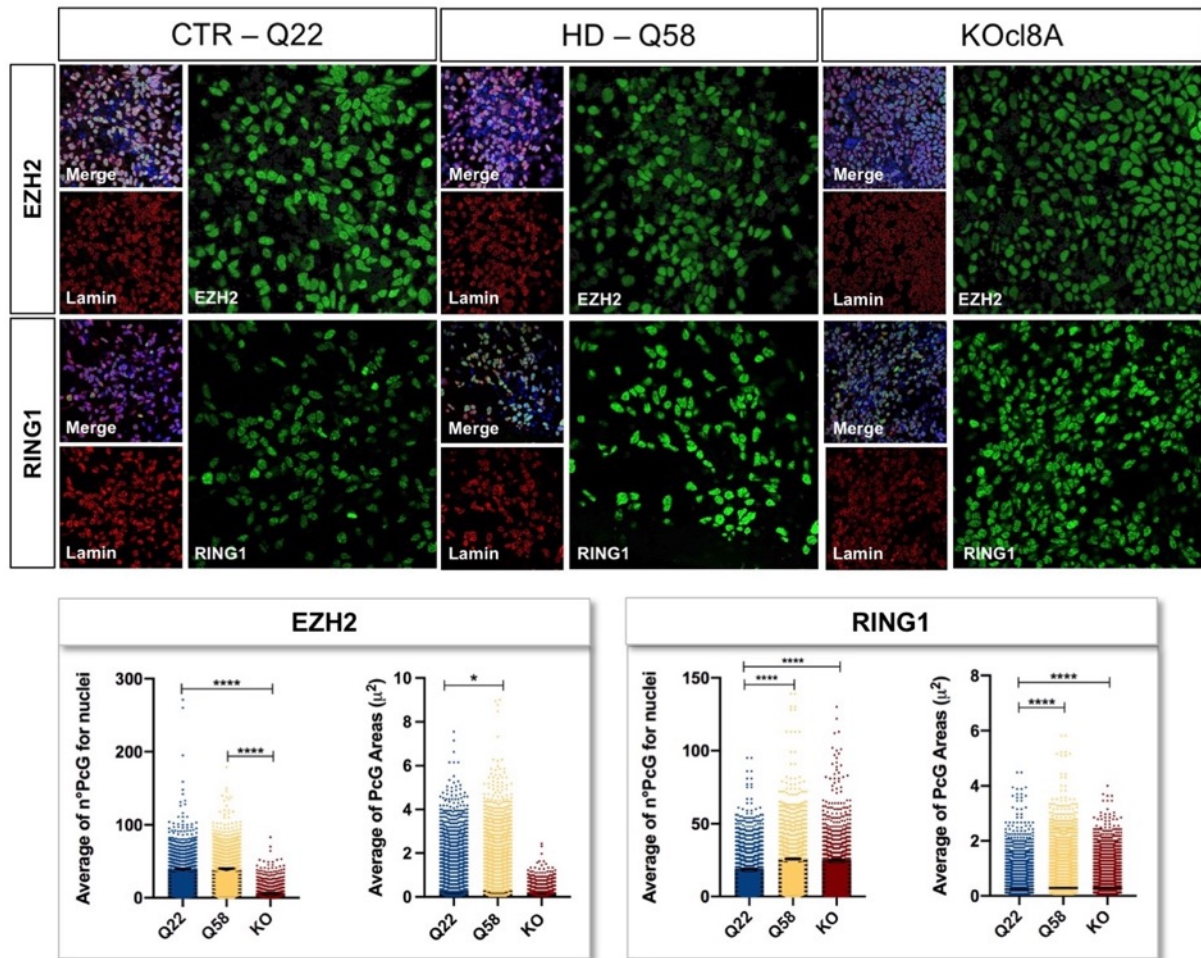
*Figure 2.17. Analysis on RUES2 lines of EZH2 and RING1 PcGs at the end of differentiation. Immunofluorescence analysis and relative quantification for control Q22, HD Q58 and KO RUES2 lines at DIV40 for EZH2 and RING1 PcG bodies (green) and LaminB1 (red). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For PcG number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For PcG Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1.000 counted nuclei per line.*

To sum up, the results obtained during the differentiation for PRC1 and PRC2 foci suggest that these complexes are highly variable over time, as the profile of the quantifications has been completely reversed from DIV0 to DIV40, and this variation could be due to the selection of a certain cell population during the differentiation.

In parallel to this, we decided to test whether the repressive histone mark H3K9me3 also varied at the end of the differentiation between the control, the HD and the KO line.

As previously performed, we have quantified at DIV40 the number of dots per nuclei and the area of each dot. Although we did not observe any difference related to the area, we identified a significant increase in the number of H3K9me3 dots in the Q58 line compared to the control Q22 and the KO line (Figure 2.18).
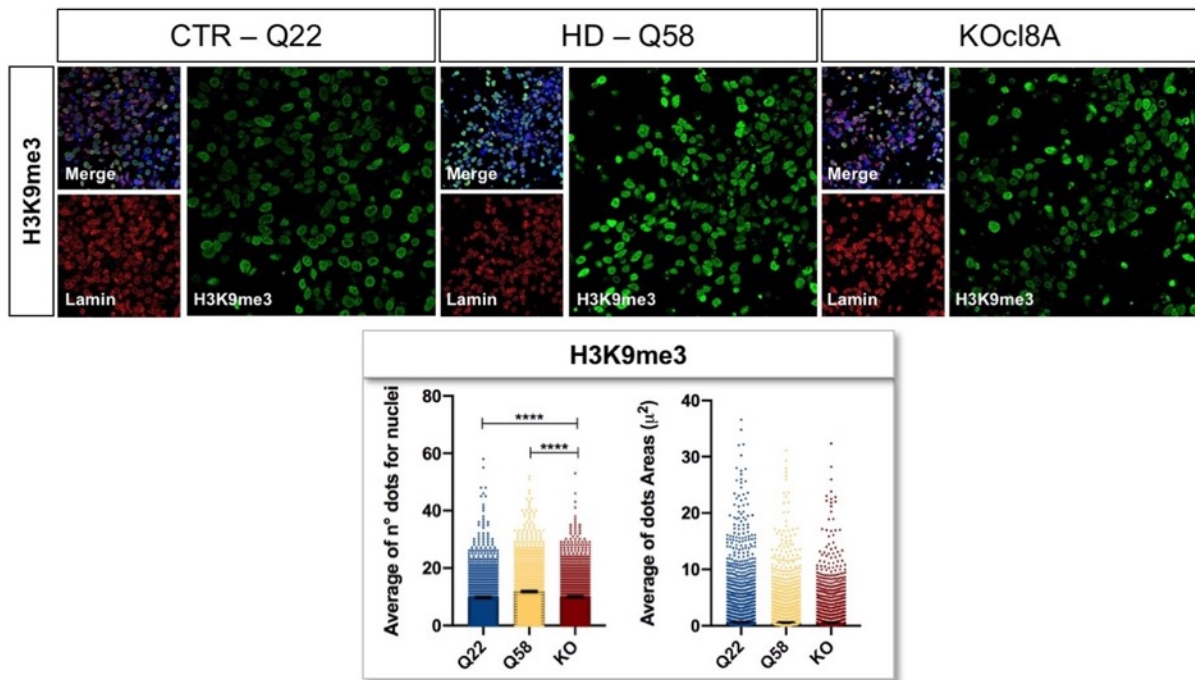
***Figure 2.18. Analysis on RUES2 lines of H3K9me3 histone modification at the end of differentiation.***
*Immunofluorescence analysis and relative quantification for control Q22, HD Q58 and KO RUES2 lines at DIV40 for H3K9me3 (green) and LaminB1 (red). Confocal images acquired with 63X, single z-stack. Box plot data are represented as mean ± SEM. For H3K9me3 dots number, ANOVA One way, Tukey post-test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. For H3K9me3 dots Area, Kruskal-Wallis test, \*\*p<0,01; \*\*\*p<0.001; \*\*\*\*p<0.0001. N=2 biological replicates, 1.000 counted nuclei per line.*

In light of this final result, the HD Q58 line has exhibited throughout the differentiation an increased number of H3K9me3 dots in respect to the control Q22 and KO line (Figure 2.11-Figure 2.15). Conversely, the KO line showed the same profile of the HD line in pluripotency but starting the neural induction and during differentiation, the level of H3K9me3 has always remained comparable to the control line (Figure 2.15 – Figure 2.18).

In this case, we might suppose that augmented level of the H3K9me3 repressive mark observed in presence of the polyQ expansion, represents a gain of function mechanism related to the HTT protein, in line to what already present in literature [55–57].

Moreover, the presence of a variation at the level of the number of H3K9me3 dots without alterations in the area might suggest that also in this case probably occurred the selection of a specific cell population during the differentiation.

The H3K9me3 histone modification is crucial not only in the context of constitutive heterochromatin, but also in the regulation of cell-type specific facultative heterochromatin and therefore the repression of lineage-specific genes [244,245]. This suggests that chromatin reorganization mediated by H3K9me3 is fundamental not only for the maintenance of the cell identity but also for the transition occurring during development, terminal differentiation and maturation [246,247].

Based on this evidence and our data, the results obtained throughout the differentiation, for both H3K9me3 histone modification and PcG bodies, suggest that the early alterations affecting HD lines might be caused by an altered epigenetic status linked to a gain of function mechanism of the mutant HTT protein that interferes with crucial processes controlling of chromatin regulation. In particular, we hypothesize that the cells acquire an altered epigenetic status at the early stages of development, that ultimately leads to an altered neuronal identity transcriptional program.

# 3. Conclusions

In this thesis project, we investigated the hypothesis that the HD mutation may cause early developmental defects which could later impact in adulthood, when the disease manifests.

First, by leveraging a new isogenic human ES cellular platform[76], we found that control parental lines are biased in the acquisition of MSN striatal commitment when exposed to a stepwise striatal differentiation protocol, differentiating into MGE-like neurons instead of the LGE neurons. This biased cell fate acquisition is suggested by the concurrent expression of typical markers of MGE, such as NKX2.1, LHX6 and LHX8, and of GABAergic interneurons, as GAD67 [1]. As already observed for the LGE [38,40,42,96,217] and cortical neurons [7-9] differentiation, the presence of muHTT interfered also in MGE cell fate acquisition altering the expression of known MGE markers. Moreover, in this work we demonstrated that the lines with expanded polyQ recapitulated known HD phenotypes, including the reduced expression of the neurodevelopmental regulators NEUROD1 and BDNF [41] and the abnormal rosettes formation during neural induction [5,11]. Finally, we reported that the CAG expansion in the HTT gene specifically interferes with the down-regulation of the OCT4 pluripotency gene, and with the proper acquisition of a PAX6$^+$ neuroectodermal fate.

Altogether these results indicate that muHTT precludes normal neuronal fate acquisition, suggesting a possible connection between muHTT and abnormal neural development in HD, as already observed in other *in vitro* and *in vivo* models [250].

One key finding of this work is the demonstration of an early defect in cell differentiation which is manifested already during the transition from pluripotency to neural cells formation. In fact, upon exposure to neural induction media, HD lines exhibited persistency of OCT4 expression – which is associated with the pluripotent cell stage - compared to control lines, in parallel to lower than expected expression of the neuroectodermal marker PAX6, hence confirming the inability of HD cells to convert to neural cells. Previous studies demonstrated that allele-specific down-regulation of muHTT induced a significant reversion of this defect in HD-iPS lines [42], confirming that the latter was directly dependent on muHTT. After verifying that the persistency of OCT4 was not caused by a failure in protein degradation, we investigated whether muHTT could alter neuronal differentiation at the epigenetic level.

Based on the evidence reported in literature, we started to study the Polycomb repressor complex, given its role in the early stages of *in vivo* development and *in vitro* differentiation [198–200]. By monitoring the number and size of the "PcG bodies", generated by the association of PRC2 and PRC1 to histone marks[230], we observed different dynamics in EZH2 and RING1 foci organization in the

presence of muHTT. Remarkably, pluripotent HD RUES2 lines displayed an increased number of EZH2 foci concomitant to a significant reduction of RING1 subunits. Inversely, at the end of neural induction the same cells exhibited increased number and size of RING1 foci, whereas there was no difference in EZH2 compared to the control.

These data suggest that muHTT could affect facultative heterochromatin by influencing both PRC1 and PRC2 activity already in pluripotency, therefore leading to a dysregulation in gene transcription. In addition to the Polycomb complexes, we considered also the H3K9me3 histone modification, a key marker of constitutive heterochromatin. We found that the Q58 line displayed a significant increase of H3K9me3 dots for nuclei over time, starting from DIV0, indicating that muHTT may act also on constitutive heterochromatin. These data are in accordance with the literature and demonstrate that HD RUES2 lines can reproduce phenotypes observed *in vivo* and in other HD systems.

Afterwards, we investigated specifically the possibility that the upstream regulatory regions of the OCT4 and PAX6 genes may present alterations in H3K27me3, H3K4me3 and H3K9me3 by ChIP experiments. No differences were found for H3K27me3 at both OCT4 and PAX6 loci, whereas the methylation state of the H3K4 residue was altered during the neural induction for both genes. Concerning the H3K9me3, we found alterations on PAX6 promoter in HD, whereas the regulation of OCT4 was similar among the lines.

These preliminary results further sustain that muHTT influences the epigenetic status of the cells through the alteration of the chromatin state and consequently of gene transcription. Notably, we observed that these modifications were already present in self-renewal although this did not cause an alteration in the pluripotency state, but led to evidence of transcriptional alterations detectable at the beginning of the induction.

To better understand the global impact of muHTT at the epigenetic level, we are currently performing ChIP-seq analysis for H3K27me3, H3K4me3 and H3K9me3 during the differentiation, from DIV0 to DIV40. In this way it will be possible to appreciate whether the early alterations observed in our cell system persist also in mature neurons. The data collected at DIV40 on the PcG bodies revealed alterations in HD also at the end of differentiation. Moreover, to understand whether the changes observed in histone modifications are caused by gain of function of muHTT or loss of function of wild-type HTT, we will include also the HTT KO RUES2 line in the ChIP-seq experiment.

To conclude, it is known that neuronal specification, maturation and survival are driven by transcriptional programs that are established during early neuronal development and persist in the adult brain. Our preliminary results support the hypothesis that muHTT causes epigenetic changes

already in pluripotent stem cells. This phenomenon may generate an altered epigenetic signature that ultimately will impact on the homeostasis of adult neurons, supporting the hypothesis of an early developmental component in HD.

# 3 | Materials and Methods

## 4.1 Cell culture

Cells were cultured in sterile conditions under a sterile laminar flow hood and were maintained in humified incubators at 37 °C with 5% $CO_2$.

### 4.1.1 Cell lines

**Table 1.2: Cell lines**

| RUES 2 Allelic Series | Cell line Code | Source/generated by |
|---|---|---|
| RUES2 20CAG cl.30 | (CHDI-90001585-2) | Coriell |
| RUES2 20CAG cl.66 | (CHDI-90002887-2) | |
| RUES2 48CAG cl 16.3 | (CHDI-90001586) | |
| RUES2 48CAG cl.20.4 | (CHDI-90001587) | Brivanlou Ali (Rockfeller University) |
| RUES2 56CAG cl 21.1 | (CHDI-90001589) | |
| RUES2 56CAG cl 25.2 | (CHDI-90001591) | |
| RUES2 72CAG cl 12 | (CHDI-90002877-1) | Coriell |
| RUES2 KO cl 8A | (CHDI-90002878-1) | |

### 4.1.2 Reagents and media for cell culture work

All cell culture reagents and media were prepared under sterile conditions or sterilized before usage.

**Table 1.3: Media and cell culture solutions**

| Medium/Reagent | Manufacturer | Medium/Reagent | Manufacturer |
|---|---|---|---|
| Accutase | Millipore | Knockout Serum Replacement (SR) | ThermoFisher |
| B27 supplement (50X) | Life Technologies | mTESR | StemCell Technologies |
| B27 supplement w/o RA (50X) | Life Technologies | N2 supplement (100x) | Life Technologies |
| Cultrex Reduced Growth Factor | Trevigent | PBS | Euroclone |
| DMEM-F12 | ThermoFisher | Penicilin Streptomycin (Pen/Strep) | Euroclone |
| DMSO | Goldbio | ReLeSR | StemCell Technologies |
| Geltrex | ThermoFisher | Trypan Blue Stain (0.4%) | Life Technologies |
| GLUTAMAX (100X) | ThermoFisher | | |

**Table 1.4: Cell culture additives.** The following reagents were purchased as powder and resuspended in an appropriate solvent to create a stock solution.

| Reagent | Manufacturer | Concentration (stock) |
|---|---|---|
| BDNF[2] | PrepoTech | 20 ng/ml |
| DKK-1[2] | PrepoTech | 100 ng/ml |
| LDN[1] | Evotec | 0.5 μM |
| ROCK Inhibitor Y-27632[2] | StemCell Technologies | 10 μM |
| SB431542[2] | | 10 μM |
| SHH(C25II)-N[2] | Tocris | 200 ng/ml |

[1] stored at 4 °C; [2] stored at -20 °C.

**Table 1.4: Cell culture media for self-renewing hES RUES2**

| Maintenance medium | hES freezing medium |
|---|---|
| 90% mTeSR1 basal | 70% mTeSR1 medium |
| 10% mTeSR 1 supplement | 20% Knockout SR |
| 1% Pen/Strep | 10% DMSO |
| | 1% ROCK Inhibitor |

**Table 1.5: Cell culture media for differentiating hES RUES2**

| Neural induction (Day0-Day4) | | Neural induction (Day5-Day10) | | Patterning (Day11-Day25) | | Terminal differentiation (Day 25-30) | |
|---|---|---|---|---|---|---|---|
| 100% | DMEM F12 | 100% | DMEM F12 | 100% | DMEM F12 | 100% | DMEM F12 |
| 1% | Pen/Strep | 1% | Pen/Strep | 1% | Pen/Strep | 1% | Pen/Strep |
| 1% | Gmax | 1% | Gmax | 1% | Gmax | 1% | Gmax |
| 1% | N2 | 1% | N2 | 1% | N2 | 1% | N2 |
| 1% | B27 w/o RA | 2% | B27 | 2% | B27 | 2% | B27 |
| 0.1% | SB431542 | 1% | DKK-1 | 1% | DKK-1 | 0.5% | BDNF |
| 0.1% | LDN | 0.5% | SHH | 0.5% | SHH | | |
| | | | | 0.5% | BDNF | | |

**Table 1.5: Cell culture coatings**

| **Geltrex-coating** |
|---|
| Incubation with 1:100 ratio (120-180 μg/mL) in PBS for 30 min at 37 °C |
| **Cultrex Reduced Growth Factor** |
| Incubation with Cultrex Reduced Growth Factor diluted 1:50 (240/360 μg/mL) in ice cold PBS 1X and incubated for 30 min at 37 °C |

### 4.1.3 Maintenance of hES RUES2 lines

Self-renewing hES cells were cultured on 1:100 Geltrex-coated plates 60mm dishes in mTeSR1 medium, which was changed daily. RUES2 lines were split once every week at 80% confluence as clumps with ReLeSR dissociation solution at 1:8 - 1:10 ratio. The karyotype for each line/clone was checked every 3 months with Q-banding analysis by ISENET Biobanking. For the analysis, cells

were seeded in T25 cell culture geltrex coated flasks. Cells were analysed when approximately 70 % confluence was reached.

### 4.1.4 Cryopreservation and thawing of cells

For cryopreservation, cells were resuspended in DMSO-containing freezing medium (Table 1.4) and transferred to cryovials, which were then placed in freezing containers and shifted to a -80 °C freezer. For long-term storage, vials were transferred to -150 °C freezer. For thawing, cryovials with frozen cells were warmed-up in a 37 °C water bath until only a small frozen clump remained. The cell suspension was then immediately transferred to PBS with ROCK inhibitor, pelleted by centrifugation and suspended in mTESR with ROCK inhibitor (1:1000).

### 4.1.5 Striatal differentiation of hES RUES2 into MSNs

Differentiation of hES RUES2 into Medium Spiny Neurons (MSNs) was performed as described by Delli Carri et al. (2013). This protocol consists of three major steps including neural induction (from day 0 to day 10), patterning (from day 5 to day 25) and terminal differentiation (from day 25 to day 30-50).
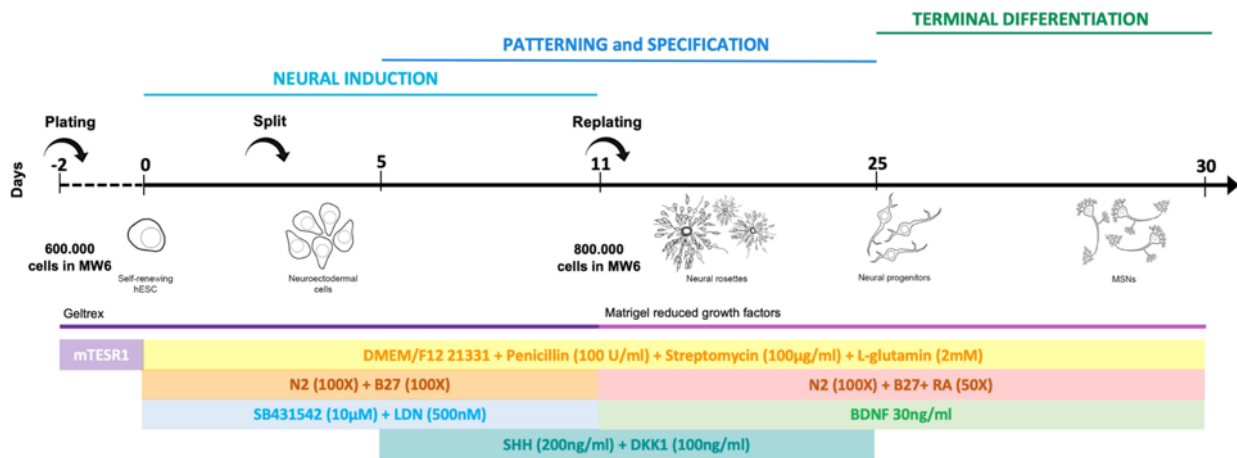


**Fig. 4.1: Scheme of the striatal differentiation protocol of RUES2 HD lines**

RUES2 were plated at a density of $0.6 \times 10^5$ cells/cm$^2$ on Geltrex-coated dishes in mTeSR1 medium with 10 μM ROCK inhibitor. Cells started to be cultured in neural induction medium when the cell layer reached 70% confluence. This medium included DMEM/F12 supplemented with N2 and B27, 10 μM SB431542 and 500 nM of LDN193189 was replaced every day. At day 3 cells were split as clumps with ReLeSR with a ratio of 1:3 MW6 cells. Starting from day 5 the neural induction medium was supplemented with 200 ng/mL SHHC-25II and 100 ng/mL DKK1. At day 11 the cell population was detached by Accutase and replated on Cultrex-reduced growth factor at a density of $8 \times 10^5$/well in a MW6 plate. From day 12, SB431542 and LDN193189 were removed and the patterning medium,

supplemented with N2, B-27 Supplement with retinoic acid 1X, SHH, DKK-1 and BDNF 20 ng/mL, was partially changed every second day. Finally, the cells were differentiated in terminal differentiation medium by adding 30 ng/mL BDNF together with N2 and B27.

## 4.2 RNA-based expression analyses

### 4.2.1 RNA extraction and DNase treatment

RNA isolation was performed to quantify RNA-expression level by quantitative PCR once reverse transcription (qRT-PCR) was performed. RNA was extracted from cells harvested by Accutase using the TRIzol Reagent (Life Technologies, Cat. No. 15596018) following manufacturer's instructions. The purified RNA was quantified by Nanodrop 1000 (Thermo Fisher Scientific) and diluted to reach the concentration of 500 ng/µL. Quality and RNA integrity were evaluated loading 1µL of each sample into a 1% agarose gel. In presence of a potential DNA contamination, samples were treated using the Ambion® DNA-free DNase Treatment and Removal Reagents (Invitrogen, Cat. No. AM1906).

### 4.2.2 RT-PCR and qPCR

The isolated RNA (500 µg) was reverse-transcribed (RT) to cDNA by using the iScript cDNA Synthesis Kit (Bio-Rad, Cat. No. 1708891) in a Thermocycler using the protocol reported in Table 1.6.

**Table 1.6: Cycling conditions for RT**

| Step | Temperature | Time |
|------|-------------|------|
| **Priming** | 25°C | 5' |
| **Reverse transcription** | 42°C | 30' |
| **Rt inactivation** | 50°C | 15' |
| | 85°C | 5' |
| **Hold** | 4°C | Infinite |

All real-time PCR reactions were prepared in a volume of 15 µl containing 50 ng cDNA and SsoFastTM EvaGreen® Supermix 2X (Bio-Rad, Cat. No. 1725204). qPCR was performed using a CFX96TM Real-Time System (Bio-Rad) coupled with a C1000TM Thermal Cycler (Bio-Rad). The final data were analyzed by CFX Manager Software (BioRad). The qRT-PCR primers used are listed on Table 1.7.

**Table 1.7: Primers used for qPCR analysis**

| Primer Name | Sequence 5'→3 | Tm |
|---|---|---|
| h18S FW | CGGCTACCACATCCAAGGAA | 60.0 |
| H18s RV | GCTGGAATTACCGCGGCT | 60.0 |
| hOCT4A FW | AGGCTCTGAGGTGTGGGGGA | 60.0 |
| hOCT4A RV | AGGCGGCTTGGAGACCTCTCA | 60.0 |
| hPAX6 FW | ATGGCAGCTGTGTGTGACACT | 60.0 |
| hPAX6 RV | GTGGAATTGGTTGGTAGACAC | 60.0 |
| hEZH2 FW | AAAGATCCTGTGCAGAAAAA | 57.3 |
| hEZH2 RV | CCTCTTCTGTCAGCTTCATC | 57.3 |
| hEZH1 FW | TACATGTCCAGCTTCCTCTT | 55.3 |
| hEZH1 RV | GACCACTTTGGCATAACAGT | 55.3 |
| hEED FW | TGACGAGAACAGCAATCCAG | 57.3 |
| hEED RV | GGTGTATCAGGGCGTTCAGT | 59.4 |
| hSUZ12 FW | AGATGGGGAATATGAAGTAGCCA | 58.9 |
| hSUZ12 RV | GAGTGAACTGCAACGTAGGT | 57.3 |
| hJARID2 FW | GTCCCCTTTTGCAATCAGCA | 57.3 |
| hJARID2 RV | TCCCATCACTGTCATCGTATTTCT | 59.3 |

# 4.3 Protein-based Assays

### 4.3.1 Protein extraction and quantification

Pelleted cells were mechanically homogenized in RIPA buffer (Table 1.8) added with Halt™ Protease and Phosphatase Inhibitor Cocktail 1 mM (Thermo Fisher Scientific) and PMSF 1 mM (Sigma-Aldrich) in 1:100 ratio. Protein concentration was determined by Thermo Scientific Pierce BCA Protein Assay (Thermo Fisher Scientific).

**Table 1.8: RIPA buffer composition**

| RIPA buffer | |
|---|---|
| Tris HCl (pH 8) | 50 mM |
| EDTA | 1 mM |
| NaCl | 150 mM |
| Deoxycholic acid sodium salt | 0.5% |
| Nonidet P40 | 1% |
| SDS | 0.1% |
| PMSF | 1:100 |
| Protease inhibitor cocktail | 1:100 |

### 4.3.2 Western Blot

A protein amount of 20 to 60 mg was loaded per track into 6% to 12% of SDS-PAGE gels. After the run, proteins were transferred onto a nitrocellulose membrane (Biorad) using the Trans-Blot® Turbo™ System (BioRad). The efficacy of protein transfer was checked by Ponceau staining.

Membranes were blocked for 1 hour at RT in blocking solution containing TBS-T 0.1% with 5% non-fat dry milk (Biorad) or Bovine Serum Albumin in TBS-T and incubated overnight at 4°C with the primary antibody (Table 1.9). After three washing steps in TBS-T 0.1%, nitrocellulose membranes were incubated for 1h at RT with an HRP-linked secondary antibody (dilution 1:3000, Biorad). Finally, the membrane was again washed three times with TBS-T and the HRP-signal was detected with a chemiluminometer (ChemiDoc, Bio-Rad) using ClarityTM Western ECL Substrate (Bio-Rad) in accordance to the manufacturer's instructions. The densitometric analyses were performed by ImageLab software.

**Table 1.9: Antibodies for WB**

| Antibody | Species | Manufacturer | Code | Dilution | Conjugate |
|----------|---------|--------------|------|----------|-----------|
| **EZH2** | Rabbit | BD | 612667 | 1:1000 | - |
| **SUZ12** | Rabbit | Active Motif | 39357 | 1:500 | - |
| **Jarid2** | Rabbit | AbCam | ab5408 | 1:1000 | - |
| **EED** | Rabbit | AbCam | ab4469 | 1:200 | - |
| Rabbit IgG | Goat | Biorad | 1706515 | 1:3000 | HRP |
| Mouse IgG | Goat | Biorad | 1706516 | 1:3000 | HRP |

## 4.4 Immunocytochemistry

Cells were fixed with ice-cold 4% PFA (incubation for 20 min at 4°C) and washed three times with DPBS. In order to store the plates at 4°C without proceeding directly with the immunostaining, 1% Sodium azide (PBSaz) was added. Cells were permeabilized with 0.5% Triton-X-100 in PBS for 10 minutes and soon after incubated in blocking solution (5% normal goat serum (NGS, Vector) for 1h at RT.

### 4.4.1 Immunocytochemistry for OCT4 and PAX6

Afterwards, cells were incubated overnight at 4°C with the respective primary antibodies diluted in solution containing 2.5% NGS and 0,25% Triton X-100. After two washing steps in PBS, Alexa Fluor®-conjugated secondary antibodies (listed in Table 1.10) (Molecular Probe, Life Technologies) were diluted 1:500 in PBS and mixed with 0.1 μg/mL Hoechst (Invitrogen) to counterstain the nuclei. The images were acquired for three biological replicates using InCell 6000 images 40X objective. N=20 images for each line/differentiation for a total of 160 images for line/time points. Cell counting was performed using CellProfiler 2.2.1 software in order to quantify automatically the different images using a specific pipeline. Different pipelines were built using algorithms that allow the identification and measurement of biological objects with specific features. For certain images or staining when CellProfiler analysis was not possible the cell counting was performed manually by ImageJ Software.

### 4.4.2 Immunocytochemistry for PcGs, histone modifications and Lamin

Firstly, cells were incubated overnight at 4°C with the nuclear antibody (H3K9me3, RING1, EZH2) in solution containing 2.5% NGS and 0,25% Triton X-100. After two washing steps in PBS/0,1% Triton X-100, the second primary antibody for LaminB1 was added for 3h at RT. The wells were washed two times with PBS/0,1% Triton X-100 before adding the Alexa Fluor®-conjugated secondary antibodies (listed in Table 1.10) (Molecular Probe, Life Technologies) were diluted 1:500 in PBS. Finally, 0.1 µg/mL Hoechst (Invitrogen) was added for 10' at RT to counterstain the nuclei. At least 10 to 15 pictures were acquired on a Leica TCS SP5 Confocal Laser Scanning Microscope (Leica Microsystems), using a 63X oil immersion objective.

**Table 1.10: Antibodies for IF**

| Antibody | Species | Manufacturer | Code | Dilution | Conjugate |
|---|---|---|---|---|---|
| OCT3/4 | Mouse | SantaCruz | | | Alexa Fluor 568 |
| PAX6 | Rabbit | Covance | 901302 | | Alexa Fluor 488 |
| LMNB1 | Mouse | Merck | AMAB91251 | 1:500 | Alexa Fluor 647 |
| H3 Tri-Methyl (K9) | Rabbit | AbCam | ab8898 | 1:500 | Alexa Fluor 488 |
| RING1B (D22F2) | Rabbit | CellSignaling | 5694 | 1:500 | Alexa Fluor 488 |
| EZH2 (D2C9) | Rabbit | Cell Signaling | 5246 | 1:500 | Alexa Fluor 488 |

The quantification of the PcG foci was performed with an automated pipeline described in Cesarini et al. (2015). The pipeline is based on these main steps: loading of Hoechst- or LaminB- stained nuclei and PcGs immunofluorescence, conversion of the images into grayscale, segmentation by IdentifyPrimaryObject algorithm using a two-class Otsu global threshold method, identification of nuclei regions on Hoechst and LaminB discarding the images on the border, measurement of the PcGs intensity, size and shape of the nuclei (Fig.4.2).
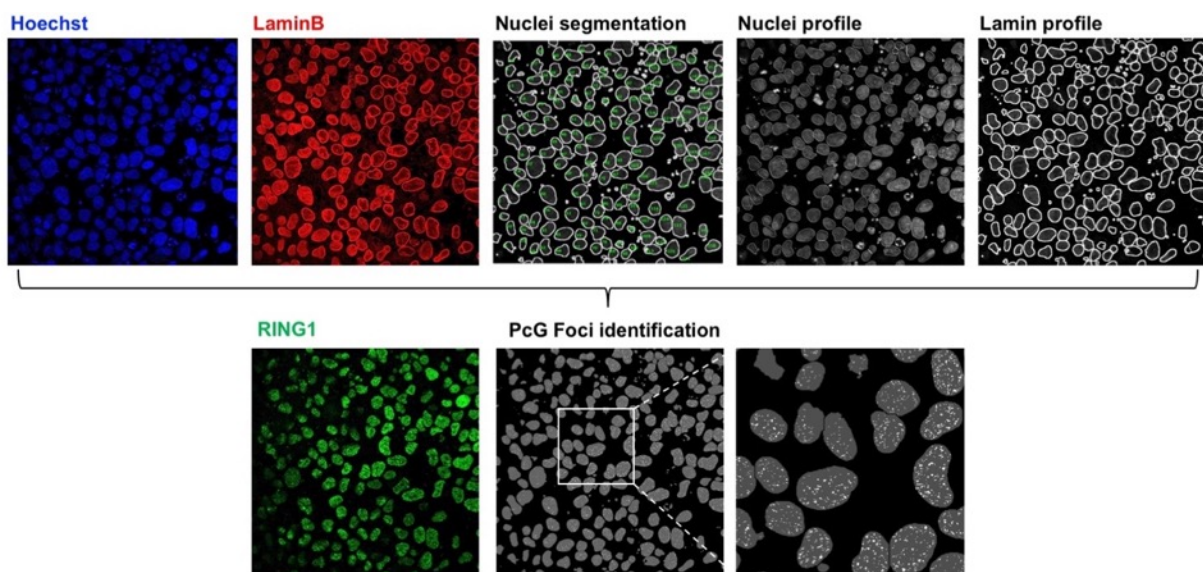


**Fig. 4.2: Scheme of the pipeline for PcG Foci identification**

PcG localization was analyzed in respect to the nuclear periphery dividing each nuclear area in 3 concentric regions of the same size using matlab (bwmorph shrink operation). The number of PcG proteins in each area was measured by counting the pixels contained.

# 4.5 Chromatin Immunoprecipitation

Firstly, we plated self-renewing hES cells on Geltrex-coated 60mm dishes (3 plates for each line) with mTESR1 medium, which was changed daily until was reached confluence to perform the ChIP on pluripotent cell (9 millions of cells). For the analysis on differentiated cells we plated hES cells on Geltrex-coated MW6 plates according to the striatal differentiation protocol (see Paragraph 4.1.5). For the ChIP were collected 5 wells at Day 4, 12 wells at Day8, and 24 wells at Day40 from MW6 plates.

**Cross-linking and cell harvesting**

The hES cells were harvested after 5 minutes of incubation with Accutase and resuspended in a solution made of 9V of culture medium and 1V of formaldehyde solution (50mM Hepes-KOH pH7.5, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 11% formaldehyde - in H2O). The cells were cross-linked with the formaldehyde solution for 10 min at RT in mild agitation. The fixation was quenched by addition of 125 mM Glycine for 5 minutes at RT in mild agitation. The suspension was centrifuged for 20 minutes at 1'200 g at 4°C. The pellet was washed twice with cold 20%FBS/PBS. Then, the cells were counted in order to freeze rapidly units of 3 million cells pellets in 1.5mL Eppendorf tubes using a dry ice-ethanol bath. The samples were stored for at least one night at -80°C until sonication. The pellets were resuspended in 130 µL of sonication buffer (10mM TrisHcl pH 8.0, 2mM EDTA, 0,5% SDS, 1X PMSF, 1X protease inhibitors), to perform the sonication in Covaris M220 focused-ultrasonicator using snap cap microTUBEs with the conditions reported in Table 1.11.

**Table 1.11: Sonication settings**

| Time Point | Peak power | Duty Factor | Cycles/burst | Duration | Cell Line |
|---|---|---|---|---|---|
| Day0 | 75.0 | 15.0 | 250 | 15 min | Q22 |
| | | | | 15 min + 5 min | Q58 |
| Day4 | 75.0 | 15.0 | 250 | 15 min | Q22 |
| | | | | 15 min + 5 min | Q58 |
| Day8 | 75.0 | 15.0 | 250 | 15 min | Q22 |
| | | | | 15 min | Q58 |
| Day40 | 75.0 | 15.0 | 250 | 15 min | Q22 |
| | | | | 15 min | Q58 |

A small aliquot of chromatin was denatured and purified to control the fragmentation by DNA high sensitivity kit for Bioanalyzer (Agilent, 5067-4626). Four volumes of equilibration buffer (10mM

Tris pH8, 175mM NaCl, 1.25% Triton X-100, 0.125% DOC, 1.4mM EDTA, 1X PMSF, 1X protease inhibitors) were added after sonication and the samples were centrifuged for 10' at 4°C at 14000g. The surnatant was transferred to a new tube, quantified and the volume adjusted with IP buffer (10mM TrisHCL pH 8.0, 140 mM NaCl, 1mM EDTA, 0.1% SDS, 0.1%DOC, 1% Triton X-100, 1X PMSF, 1X protease inhibitors). From each sample the 3% was preserved at 4°C as input. Overnight incubation on the rotating wheel at 4°C of the remaining chromatin with the antibody of interest.

G beads (Life Technologies, 1004D) were washed two times in 0.1% BSA/IP buffer, resuspended in the same buffer and incubated overnight on the rotating wheel at 4°C. The next day 15μL of beads were added to each sample and incubated for 2h on the rotating wheel at 4°C. The bound complexes were washed at 4°C twice in IP buffer, twice in high-salt IP buffer (500mM NaCl instead of 140mM), twice in RIPA-LiCl buffer (10mM Tris-HCl pH 8.0, 1mM EDTA, 250mM LiCl, 0.5% DOC, 0.5% NP-40, 1X PMSF, 1X protease inhibitors) and once in 10mM Tris-HCl pH 8.0. Cross-linking was reversed at 65°C overnight in elution buffer (10 mM Tris pH 8, 5 mM EDTA, and 0,4% SDS), and DNA was extracted from beads by standard phenol/chloroform extraction, precipitated, and resuspended in 31 μl of 10-mM Tris, pH 7.5.

**Table 1.12: Antibodies for ChIP**

| Antibody | Species | Manufacturer | Code |
|---|---|---|---|
| Histone H3 | Rabbit | Abcam | ab1791 |
| H3 Tri-Methyl (Lys27) | Rabbit | Merck Millipore | #07-449 |
| H3 Tri-Methyl (Lys4) | Rabbit | Merck Millipore | #07-473 |
| H3 Tri-Methyl (K9) | Rabbit | Abcam | ab8898 |

**Analysis in ChIP-qPCR**

Immunoprecipitated DNA was analyzed in duplicates by real-time PCR starting from a 5 μL of diluted DNA (1:10 dilution from the template DNA) using SsoFastTM EvaGreen® Supermix 2X (Bio-Rad, Cat. No. 1725204). The efficiency of the qPCR primers designed for specific genomic regions was tested generating a standard curve by running six-point serial dilutions (1:3) of a sample of known concentration and plotting the corresponding Ct values to generate a standard curve. qPCR was performed using a CFX96TM Real-Time System (Bio-Rad) coupled with a C1000TM Thermal Cycler (Bio-Rad). The final data were analyzed by CFX Manager Software (BioRad). The qRT-PCR primers used are listed on Table 1.13.

**Table 1.13: Primers for ChIP-qPCR**

| Primer Name | Sequence 5'→3 | Tm |
|---|---|---|
| **hOCT4-PP FW** | AGTCTGGGCAACAAAGTGAGA | 57.9 |
| **hOCT4-PP RV** | AGAAACTGAGGCGAAGGATG | 57.3 |
| **hOCT4-PE FW** | TCTGTTTCAGCAAAGGTTGGG | 57.9 |
| **hOCT4-PE RV** | TTGGTCCCTACTTCCCCTTCA | 59.8 |
| **hOCT4-DE FW** | GAGGATGGCAAGCTGAGAAA | 57.3 |
| **hOCT4-DE RV** | CTCAATCCCCAGGACAGAAC | 59.4 |
| **hPAX6 FW** | AAAACCCCAACCAAACAAAA | 51.1 |
| **hPAX6 RV** | GCAATAAAAATAAAGCGAGAAGAAA | 54.8 |

## 4.6 Statistical analysis

Quantitative data were generated in biological replicates. All results presented as graphs show mean + SEM (standard error of the mean), which was computed using the GraphPad Prism software. Statistical significance, unless otherwise stated, was analyzed by two-tailed Student's t-test for control and experimental conditions, and $p \leq 0.05$ was considered to be statistically significant.

# 4 | References

1.  Reiner, A. *et al.* Differential loss of striatal projection neurons in Huntington disease. *Proc. Natl. Acad. Sci. U. S. A.* (1988) doi:10.1073/pnas.85.15.5733.

2.  Kirkwood, S. C., Su, J. L., Conneally, P. M. & Foroud, T. Progression of symptoms in the early and middle stages of Huntington disease. *Arch. Neurol.* (2001) doi:10.1001/archneur.58.2.273.

3.  Aylward, E. H. *et al.* Longitudinal change in regional brain volumes in prodromal Huntington disease. *J. Neurol. Neurosurg. Psychiatry* (2011) doi:10.1136/jnnp.2010.208264.

4.  Rosas, H. D. *et al.* Cerebral cortex and the clinical expression of Huntington's disease: Complexity and heterogeneity. *Brain* (2008) doi:10.1093/brain/awn025.

5.  O., Q., K.L., O., O., B. & M., S. The prevalence of juvenile huntington9s disease: A review of the literature and meta-analysis. *PLoS Curr.* (2012).

6.  Saudou, F. & Humbert, S. The Biology of Huntingtin. *Neuron* (2016) doi:10.1016/j.neuron.2016.02.003.

7.  Godin, J. D. *et al.* Huntingtin Is Required for Mitotic Spindle Orientation and Mammalian Neurogenesis. *Neuron* (2010) doi:10.1016/j.neuron.2010.06.027.

8.  Mattis, V. B. *et al.* Induced pluripotent stem cells from patients with huntington's disease show CAG repeat expansion associated phenotypes. *Cell Stem Cell* (2012) doi:10.1016/j.stem.2012.04.027.

9.  Lopes, C. *et al.* Dominant-negative effects of adult-onset huntingtin mutations alter the division of human embryonic stem cells-derived neural cells. *PLoS One* (2016) doi:10.1371/journal.pone.0148680.

10. Molero, A. E. *et al.* Impairment of developmental stem cell-mediated striatal neurogenesis and pluripotency genes in a knock-in model of Huntington's disease. *Proc. Natl. Acad. Sci. U. S. A.* (2009) doi:10.1073/pnas.0912171106.

11. Nasir, J. *et al.* Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell* (1995) doi:10.1016/0092-8674(95)90542-1.

12. Zeitlin, S., Liu, J. P., Chapman, D. L., Papaioannou, V. E. & Efstratiadis, A. Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nat. Genet.* (1995) doi:10.1038/ng1095-155.

13. Molero, A. E. *et al.* Selective expression of mutant huntingtin during development recapitulates characteristic features of Huntington's disease. *Proc. Natl. Acad. Sci. U. S. A.* (2016) doi:10.1073/pnas.1603871113.

14. Mehler, M. F. & Gokhan, S. Mechanisms underlying neural cell death in neurodegenerative diseases: Alterations of a developmentally-mediated cellular rheostat. *Trends in Neurosciences* (2000) doi:10.1016/S0166-2236(00)01705-7.

15. Cummings, D. M. *et al.* Aberrant cortical synaptic plasticity and dopaminergic dysfunction in a mouse model of Huntington's disease. *Hum. Mol. Genet.* (2006) doi:10.1093/hmg/ddl224.

16. Hodges, A. *et al.* Regional and cellular gene expression changes in human Huntington's disease brain. *Hum. Mol. Genet.* (2006) doi:10.1093/hmg/ddl013.

17. Kuhn, A. *et al.* Mutant huntingtin's effects on striatal gene expression in mice recapitulate changes observed in human Huntington's disease brain and do not differ with mutant huntingtin length or wild-type huntingtin dosage. *Hum. Mol. Genet.* (2007) doi:10.1093/hmg/ddm133.

18. Labadorf, A. *et al.* RNA sequence analysis of human huntington disease brain reveals an extensive increase in inflammatory and developmental gene expression. *PLoS One* (2015) doi:10.1371/journal.pone.0143563.

19. Milnerwood, A. J. *et al.* Early development of aberrant synaptic plasticity in a mouse model of Huntington's disease. *Hum. Mol. Genet.* (2006) doi:10.1093/hmg/ddl092.

20. Schippling, S. *et al.* Abnormal Motor Cortex Excitability in Preclinical and Very Early Huntington's Disease. *Biol. Psychiatry* (2009) doi:10.1016/j.biopsych.2008.12.026.

21. Lee, J. K. *et al.* Measures of growth in children at risk for Huntington disease. *Neurology* (2012) doi:10.1212/WNL.0b013e3182648b65.

22. Tereshchenko, A. *et al.* Abnormal Weight and Body Mass Index in Children with Juvenile Huntington's Disease. *Journal of Huntington's Disease* (2015) doi:10.3233/JHD-150152.

23. Gomez-Tortosa, E. *et al.* Quantitative neuropathological changes in presymptomatic Huntington's disease. *Ann. Neurol.* (2001) doi:10.1002/1531-8249(200101)49:1<29::AID-

ANA7>3.3.CO;2-2.

24. Paulsen, J. S. *et al.* Brain structure in preclinical Huntington's disease. *Biol. Psychiatry* (2006) doi:10.1016/j.biopsych.2005.06.003.

25. Nopoulos, P. *et al.* Morphology of the cerebral cortex in preclinical Huntington's disease. *Am. J. Psychiatry* (2007) doi:10.1176/appi.ajp.2007.06081266.

26. Paulsen, J. S. *et al.* Striatal and white matter predictors of estimated diagnosis for Huntington disease. *Brain Res. Bull.* (2010) doi:10.1016/j.brainresbull.2010.04.003.

27. Nopoulos, P. C. *et al.* Smaller intracranial volume in prodromal Huntington's disease: Evidence for abnormal neurodevelopment. *Brain* (2011) doi:10.1093/brain/awq280.

28. Barnat, M. *et al.* Huntington's disease alters human neurodevelopment. *Science (80-. ).* (2020) doi:10.1126/science.aax3338.

29. Wiatr, K., Szlachcic, W. J., Trzeciak, M., Figlerowicz, M. & Figiel, M. Huntington Disease as a Neurodevelopmental Disorder and Early Signs of the Disease in Stem Cells. *Molecular Neurobiology* (2018) doi:10.1007/s12035-017-0477-7.

30. Duyao, M. P. *et al.* Inactivation of the mouse huntington's disease gene homolog Hdh. *Science (80-. ).* (1995) doi:10.1126/science.7618107.

31. White, J. K. *et al.* Huntington is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nat. Genet.* (1997) doi:10.1038/ng1297-404.

32. Auerbach, W. The HD mutation causes progressive lethal neurological disease in mice expressing reduced levels of huntingtin. *Hum. Mol. Genet.* (2001) doi:10.1093/hmg/10.22.2515.

33. Arteaga-Bracho, E. E. *et al.* Postnatal and adult consequences of loss of huntingtin during development: Implications for Huntington's disease. *Neurobiol. Dis.* (2016) doi:10.1016/j.nbd.2016.09.006.

34. Nguyen, G. D., Molero, A. E., Gokhan, S. & Mehler, M. F. Functions of Huntingtin in Germ Layer Specification and Organogenesis. *PLoS One* (2013) doi:10.1371/journal.pone.0072698.

35. Lo Sardo, V. *et al.* An evolutionary recent neuroepithelial cell adhesion function of huntingtin implicates ADAM10-Ncadherin. *Nat. Neurosci.* (2012) doi:10.1038/nn.3080.

36. Elkabetz, Y. *et al.* Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev.* (2008)

doi:10.1101/gad.1616208.

37. Abranches, E. *et al.* Neural differentiation of embryonic stem cells in vitro: A road map to neurogenesis in the embryo. *PLoS One* (2009) doi:10.1371/journal.pone.0006286.

38. Ring, K. L. *et al.* Genomic Analysis Reveals Disruption of Striatal Neuronal Development and Therapeutic Targets in Human Huntington's Disease Neural Stem Cells. *Stem Cell Reports* (2015) doi:10.1016/j.stemcr.2015.11.005.

39. Mattis, V. B. *et al.* HD iPSC-derived neural progenitors accumulate in culture and are susceptible to BDNF withdrawal due to glutamate toxicity. *Hum. Mol. Genet.* (2014) doi:10.1093/hmg/ddv080.

40. HD iPSC Consortium *et al.* Developmental alterations in Huntington's disease neural cells and pharmacological rescue in cells and mice. *Nat. Neurosci.* (2017) doi:10.1038/nn.4532.

41. HD iPSC Consortium *et al.* Induced pluripotent stem cells from patients with huntington's disease show CAG repeat expansion associated phenotypes. *Cell Stem Cell* (2012) doi:10.1016/j.stem.2012.04.027.

42. Conforti, P. *et al.* Faulty neuronal determination and cell polarization are reverted by modulating HD early phenotypes. *Proc. Natl. Acad. Sci. U. S. A.* (2018) doi:10.1073/pnas.1715865115.

43. Mehta, S. R. *et al.* Human Huntington's Disease iPSC-Derived Cortical Neurons Display Altered Transcriptomics, Morphology, and Maturation. *Cell Rep.* (2018) doi:10.1016/j.celrep.2018.09.076.

44. Tong, Y. *et al.* Spatial and temporal requirements for huntingtin (Htt) in neuronal migration and survival during brain development. *J. Neurosci.* (2011) doi:10.1523/JNEUROSCI.2774-11.2011.

45. Reiner, A. *et al.* Neurons lacking huntingtin differentially colonize brain and survive in chimeric mice. *J. Neurosci.* (2001) doi:10.1523/jneurosci.21-19-07608.2001.

46. Benn, C. L. *et al.* Huntingtin modulates transcription, occupies gene promoters in vivo, and binds directly to DNA in a polyglutamine-dependent manner. *J. Neurosci.* (2008) doi:10.1523/JNEUROSCI.2126-08.2008.

47. Biagioli, M. *et al.* Htt CAG repeat expansion confers pleiotropic gains of mutant huntingtin function in chromatin regulation. *Hum. Mol. Genet.* (2015) doi:10.1093/hmg/ddv006.

48. Zhai, W., Jeong, H., Cui, L., Krainc, D. & Tjian, R. In vitro analysis of huntingtin-mediated transcriptional repression reveals multiple

transcription factor targets. *Cell* (2005) doi:10.1016/j.cell.2005.10.030.

49. Seong, I. S. *et al.* Huntingtin facilitates polycomb repressive complex 2. *Hum. Mol. Genet.* (2009) doi:10.1093/hmg/ddp524.

50. Vashishtha, M. *et al.* Targeting H3K4 trimethylation in Huntington disease. *Proc. Natl. Acad. Sci. U. S. A.* (2013) doi:10.1073/pnas.1311323110.

51. Von Schimmelmann, M. *et al.* Polycomb repressive complex 2 (PRC2) silences genes responsible for neurodegeneration. *Nat. Neurosci.* (2016) doi:10.1038/nn.4360.

52. Dong, X. *et al.* The role of H3K4me3 in transcriptional regulation is altered in Huntington's disease. *PLoS One* (2015) doi:10.1371/journal.pone.0144398.

53. Pearl, J. R. *et al.* Altered Huntingtin-Chromatin Interactions Predict Transcriptional and Epigenetic Changes in Huntington's Disease. *bioRxiv* 2020.06.04.132571 (2020) doi:10.1101/2020.06.04.132571.

54. Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends in Genetics* (2016) doi:10.1016/j.tig.2015.11.001.

55. Irmak, D. *et al.* Mechanism suppressing H3K9 trimethylation in pluripotent stem cells and its demise by polyQ-expanded huntingtin mutations. *Hum. Mol. Genet.* (2018) doi:10.1093/hmg/ddy304.

56. Ryu, H. *et al.* ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease. *Proc. Natl. Acad. Sci. U. S. A.* (2006) doi:10.1073/pnas.0606373103.

57. Lee, J. *et al.* Epigenetic regulation of cholinergic receptor M1 (CHRM1) by histone H3K9me3 impairs Ca2+ signaling in Huntington's disease. *Acta Neuropathol.* (2013) doi:10.1007/s00401-013-1103-z.

58. Barlett, Pishva, E., Rutten, B. P. F. & van den Hove, D. *Neuroepigenomics in Aging and Disease. Advances in Experimental Medicine and Biology* (2017).

59. Budday, S., Steinmann, P. & Kuhl, E. Physical biology of human brain development. *Frontiers in Cellular Neuroscience* (2015) doi:10.3389/fncel.2015.00257.

60. Herculano-Houzel, S. The human brain in numbers: A linearly scaled-up primate brain. *Frontiers in Human Neuroscience* (2009) doi:10.3389/neuro.09.031.2009.

61. Stiles, J. & Jernigan, T. L. The basics of brain development. *Neuropsychology Review* (2010) doi:10.1007/s11065-010-9148-4.

62. WINDLE, W. F. & CLEARWATERS, P. Development of the nervous system. *Prog. Neurol. Psychiatry* (1946) doi:10.5005/jp/books/11957_1.

63. Liu, H. & Zhang, S. C. Specification of neuronal and glial subtypes from human pluripotent stem cells. *Cellular and Molecular Life Sciences* (2011) doi:10.1007/s00018-011-0770-y.

64. Jain, M., Armstrong, R. J. E., Barker, R. A. & Rosser, A. E. Cellular and molecular aspects of striatal development. *Brain Res. Bull.* (2001) doi:10.1016/S0361-9230(01)00555-X.

65. STURROCK, R. R. A COMPARATIVE QUANTITATIVE AND MORPHOLOGICAL STUDY OF AGEING IN THE MOUSE NEOSTRIATUM, INDUSIUM GRISEUM AND ANTERIOR COMMISSURE. *Neuropathol. Appl. Neurobiol.* (1980) doi:10.1111/j.1365-2990.1980.tb00204.x.

66. Deacon, T. W., Pakzaban, P. & Isacson, O. The lateral ganglionic eminence is the origin of cells committed to striatal phenotypes: neural transplantation and developmental evidence. *Brain Res.* (1994) doi:10.1016/0006-8993(94)90526-6.

67. Campbell, K. Dorsal-ventral patterning in the mammalian telencephalon. *Current Opinion in Neurobiology* (2003) doi:10.1016/S0959-4388(03)00009-6.

68. O'Rahilly, R. & Müller, F. *The Embryonic Human Brain: An Atlas of Developmental Stages, Third Edition. The Embryonic Human Brain: An Atlas of Developmental Stages, Third Edition* (2005). doi:10.1002/0471973084.

69. Delli Carri, A. *et al.* Human Pluripotent Stem Cell Differentiation into Authentic Striatal Projection Neurons. *Stem Cell Rev. Reports* (2013) doi:10.1007/s12015-013-9441-8.

70. Onorati, M. *et al.* Molecular and functional definition of the developing human striatum. *Nat. Neurosci.* (2014) doi:10.1038/nn.3860.

71. Bocchi, V. D. *et al.* The coding and long noncoding single-cell atlas of the developing human fetal striatum. *Science (80-. ).* (2021) doi:10.1126/science.abf5759.

72. Suzuki, I. K. & Vanderhaeghen, P. Is this a brain which i see before me? Modeling human neural development with pluripotent stem cells. *Development (Cambridge)* (2015) doi:10.1242/dev.120568.

73.    Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* (2009) doi:10.1038/nbt.1529.

74.    Li, X. J. *et al.* Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development* (2009) doi:10.1242/dev.036624.

75.    Chamberlain, C. E., Jeong, J., Guo, C., Allen, B. L. & McMahon, A. P. Notochord-derived Shh concentrates in close association with the apically positioned basal body in neural target cells and forms a dynamic gradient during neural patterning. *Development* (2008) doi:10.1242/dev.013086.

76.    Bafico, A., Liu, G., Yaniv, A., Gazit, A. & Aaronson, S. A. Novel mechanism of Wnt signalling inhibition mediated by Dickkopf-1 interaction with LRP6/Arrow. *Nat. Cell Biol.* (2001) doi:10.1038/35083081.

77.    Chatzi, C., Brade, T. & Duester, G. Retinoic acid functions as a key gabaergic differentiation signal in the basal ganglia. *PLoS Biol.* (2011) doi:10.1371/journal.pbio.1000609.

78.    Rubenstein, J. L. R., Shimamura, K., Martinez, S. & Puelles, L. REGIONALIZATION OF THE PROSENCEPHALIC NEURAL PLATE. *Annu. Rev. Neurosci.* (1998) doi:10.1146/annurev.neuro.21.1.445.

79.    Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat. Protoc.* (2012) doi:10.1038/nprot.2012.116.

80.    Kaye, J. A. & Finkbeiner, S. Modeling Huntington's disease with induced pluripotent stem cells. *Molecular and Cellular Neuroscience* (2013) doi:10.1016/j.mcn.2013.02.005.

81.    Sayed, N., Liu, C. & Wu, J. C. Translation of Human-Induced Pluripotent Stem Cells from Clinical Trial in a Dish to Precision Medicine. *Journal of the American College of Cardiology* (2016) doi:10.1016/j.jacc.2016.01.083.

82.    Verlinsky, Y. *et al.* Human embryonic stem cell lines with genetic disorders. *Reprod. Biomed. Online* (2005) doi:10.1016/S1472-6483(10)60810-3.

83.    Ruzo, A. *et al.* Chromosomal instability during neurogenesis in huntington's disease. *Dev.* (2018) doi:10.1242/dev.156844.

84.    James, D., Noggle, S. A., Swigut, T. & Brivanlou, A. H. Contribution of human embryonic stem cells to mouse blastocysts. *Dev. Biol.* (2006) doi:10.1016/j.ydbio.2006.03.026.

85.    Rosa, A., Spagnoli, F. M. & Brivanlou, A. H. Mesendodermal Fate Specification via Species-Specific Target Selection. *Dev. Cell* (2009) doi:10.1016/j.devcel.2009.02.007.

86.    Cong, L. *et al.* Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., … Zhang, F. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science (New York, N.Y.). *Science* (2013) doi:10.1126/science.1231143.Multiplex.

87.    Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-. ).* (2012) doi:10.1126/science.1225829.

88.    Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science (80-. ).* (2013) doi:10.1126/science.1232033.

89.    Lacoste, A., Berenshteyn, F. & Brivanlou, A. H. An Efficient and Reversible Transposable System for Gene Delivery and Lineage-Specific Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell* (2009) doi:10.1016/j.stem.2009.07.011.

90.    Reményi, A., Schöler, H. R. & Wilmanns, M. Combinatorial control of gene expression. *Nature Structural and Molecular Biology* (2004) doi:10.1038/nsmb820.

91.    Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* (2003) doi:10.1016/S0092-8674(03)00392-1.

92.    Takahashi, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* (2003).

93.    Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* (2005) doi:10.1016/j.cell.2005.08.020.

94.    Tantin, D. Oct transcription factors in development and stem cells: Insights and mechanisms. *Dev.* (2013) doi:10.1242/dev.095927.

95.    Zhao, F. Q. Octamer-binding transcription factors: Genomics and functions. *Front. Biosci.* (2013) doi:10.2741/4162.

96.    Wang, X. & Dai, J. Concise review: Isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem Cells* (2010) doi:10.1002/stem.419.

97.    Gao, Y. *et al.* The novel function of OCT4B Isoform-265 in genotoxic stress. *Stem Cells* (2012) doi:10.1002/stem.1034.

98.    Yeom, Y. I. I. *et al.* Germline regulatory element of

Oct-4 specific for the totipotent cycle of embryonal cells. *Development* (1996).

99. Shi, G. & Jin, Y. Role of Oct4 in maintaining and regaining stem cell pluripotency. *Stem Cell Research and Therapy* (2010) doi:10.1186/scrt39.

100. Gao, F., Wei, Z., An, W., Wang, K. & Lu, W. The interactomes of POU5F1 and SOX2 enhancers in human embryonic stem cells. *Sci. Rep.* (2013) doi:10.1038/srep01588.

101. Feldman, N. *et al.* G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat. Cell Biol.* (2006) doi:10.1038/ncb1353.

102. Li, J.-Y. *et al.* Synergistic Function of DNA Methyltransferases Dnmt3a and Dnmt3b in the Methylation of Oct4 and Nanog. *Mol. Cell. Biol.* (2007) doi:10.1128/mcb.01380-07.

103. Smith-Geater, C. *et al.* Aberrant Development Corrected in Adult-Onset Huntington's Disease iPSC-Derived Neuronal Cultures via WNT Signaling Modulation. *Stem Cell Reports* (2020) doi:10.1016/j.stemcr.2020.01.015.

104. Ramocki, M. B. & Zoghbi, H. Y. Failure of neuronal homeostasis results in common neuropsychiatric phenotypes. *Nature* (2008) doi:10.1038/nature07457.

105. Waddington, C. H. The epigenotype. Endeavour 1942; 1:18-20. *Repr. Int. J. Epidemiol.* (2011).

106. Bianchi, A. & Lanzuolo, C. Into the chromatin world: Role of nuclear architecture in epigenome regulation. *AIMS Biophysics* (2015) doi:10.3934/biophy.2015.4.585.

107. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* (1997) doi:10.1038/38444.

108. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* (2007) doi:10.1016/j.cell.2007.02.005.

109. Bernstein, B. E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* (2005) doi:10.1016/j.cell.2005.01.001.

110. Nakamura, T. *et al.* ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol. Cell* (2002) doi:10.1016/S1097-2765(02)00740-2.

111. Milne, T. A. *et al.* MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol. Cell* (2002) doi:10.1016/S1097-2765(02)00741-4.

112. Huisinga, K. L., Brower-Toland, B. & Elgin, S. C. R. The contradictory definitions of heterochromatin: Transcription and silencing. *Chromosoma* (2006) doi:10.1007/s00412-006-0052-x.

113. Trojer, P. & Reinberg, D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol. Cell* **28**, 1–13 (2007).

114. Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* (2004) doi:10.1016/j.cell.2004.12.012.

115. Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M. & Krebs, C. Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc. Natl. Acad. Sci. U. S. A.* (2006) doi:10.1073/pnas.0604005103.

116. Ozer, A. & Bruick, R. K. Non-heme dioxygenases: Cellular sensors and regulators jelly rolled into one? *Nature Chemical Biology* (2007) doi:10.1038/nchembio863.

117. Gökbuget, D. & Blelloch, R. Epigenetic control of transcriptional regulation in pluripotency and early differentiation. *Development (Cambridge)* (2019) doi:10.1242/dev.164772.

118. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by polycomb and compass families. *Science* (2016) doi:10.1126/science.aad9780.

119. Denissov, S. *et al.* Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Dev.* (2014) doi:10.1242/dev.102681.

120. Ketel, C. S. *et al.* Subunit Contributions to Histone Methyltransferase Activities of Fly and Worm Polycomb Group Complexes. *Mol. Cell. Biol.* (2005) doi:10.1128/mcb.25.16.6857-6868.2005.

121. Peng, J. C. *et al.* Jarid2/Jumonji Coordinates Control of PRC2 Enzymatic Activity and Target Gene Occupancy in Pluripotent Cells. *Cell* (2009) doi:10.1016/j.cell.2009.12.002.

122. Shen, X. *et al.* Jumonji Modulates Polycomb Activity and Self-Renewal versus Differentiation of Stem Cells. *Cell* (2009) doi:10.1016/j.cell.2009.12.003.

123. Li, G. *et al.* Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev.* (2010) doi:10.1101/gad.1886410.

124. Pasini, D. *et al.* JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. *Nature* (2010) doi:10.1038/nature08788.

125. Kalb, R. *et al.* Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* (2014)

doi:10.1038/nsmb.2833.

126. Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* (2014) doi:10.1016/j.cell.2014.05.004.

127. Cooper, S. *et al.* Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep.* (2014) doi:10.1016/j.celrep.2014.04.012.

128. Názer, E. & Lei, E. P. Modulation of chromatin modifying complexes by noncoding RNAs in trans. *Current Opinion in Genetics and Development* (2014) doi:10.1016/j.gde.2013.11.019.

129. Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* (2009) doi:10.1038/nature08398.

130. Xu, C. *et al.* Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). *Proc. Natl. Acad. Sci. U. S. A.* (2010) doi:10.1073/pnas.1008937107.

131. Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* (2006) doi:10.1038/nature04733.

132. Surface, L. E., Thornton, S. R. & Boyer, L. A. Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* (2010) doi:10.1016/j.stem.2010.08.004.

133. Posfai, E. *et al.* Polycomb function during oogenesis is required for mouse embryonic development. *Genes Dev.* (2012) doi:10.1101/gad.188094.112.

134. Morey, L., Aloia, L., Cozzuto, L., Benitah, S. A. & Di Croce, L. RYBP and Cbx7 Define Specific Biological Functions of Polycomb Complexes in Mouse Embryonic Stem Cells. *Cell Rep.* (2013) doi:10.1016/j.celrep.2012.11.026.

135. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* (2000) doi:10.1038/35020506.

136. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J. SETDB1: A novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* (2002) doi:10.1101/gad.973302.

137. Tachibana, M., Sugimoto, K., Fukushima, T. & Shinkai, Y. SET Domain-containing Protein, G9a, is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3. *J. Biol. Chem.* (2001) doi:10.1074/jbc.M101914200.

138. Tachibana, M. *et al.* G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* (2002) doi:10.1101/gad.989402.

139. Tachibana, M. *et al.* Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* (2005) doi:10.1101/gad.1284005.

140. Peters, A. H. F. M. *et al.* Partitioning and Plasticity of Repressive Histone Methylation States in Mammalian Chromatin. *Mol. Cell* (2003) doi:10.1016/S1097-2765(03)00477-5.

141. Rice, J. C. *et al.* Histone Methyltransferases Direct Different Degrees of Methylation to Define Distinct Chromatin Domains. *Mol. Cell* (2003) doi:10.1016/S1097-2765(03)00479-9.

142. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* (2001) doi:10.1038/35065138.

143. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* (2001) doi:10.1038/35065132.

144. Canzio, D. *et al.* Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. *Mol. Cell* (2011) doi:10.1016/j.molcel.2010.12.016.

145. Pombo, A. & Dillon, N. Three-dimensional genome architecture: Players and mechanisms. *Nature Reviews Molecular Cell Biology* (2015) doi:10.1038/nrm3965.

146. Fudenberg, G. *et al.* Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* (2016) doi:10.1016/j.celrep.2016.04.085.

147. Mirny, L. A., Imakaev, M. & Abdennur, N. Two major mechanisms of chromosome organization. *Current Opinion in Cell Biology* (2019) doi:10.1016/j.ceb.2019.05.001.

148. Dekker, J. & Mirny, L. The 3D Genome as Moderator of Chromosomal Communication. *Cell* (2016) doi:10.1016/j.cell.2016.02.007.

149. Cremer, T. *et al.* Chromosome territories - a functional nuclear landscape. *Current Opinion in Cell Biology* (2006) doi:10.1016/j.ceb.2006.04.007.

150. Dundr, M. Nuclear bodies: Multifunctional companions of the genome. *Current Opinion in Cell Biology* (2012) doi:10.1016/j.ceb.2012.03.010.

151. Zhu, L. & Brangwynne, C. P. Nuclear bodies: The emerging biophysics of nucleoplasmic phases. *Current Opinion in Cell Biology* (2015) doi:10.1016/j.ceb.2015.04.003.

152. Cheutin, T. & Cavalli, G. Progressive polycomb assembly on H3K27me3 compartments generates Polycomb bodies with developmentally regulated motion. *PLoS Genet.* (2012) doi:10.1371/journal.pgen.1002465.

153. Wani, A. H. *et al.* Chromatin topology is coupled to Polycomb group protein subnuclear organization. *Nat. Commun.* (2016) doi:10.1038/ncomms10291.

154. Denholtz, M. *et al.* Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* (2013) doi:10.1016/j.stem.2013.08.013.

155. Joshi, O. *et al.* Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. *Cell Stem Cell* (2015) doi:10.1016/j.stem.2015.11.010.

156. Schoenfelder, S. *et al.* Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat. Genet.* (2015) doi:10.1038/ng.3393.

157. Cruz-Molina, S. *et al.* PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell* (2017) doi:10.1016/j.stem.2017.02.004.

158. Eagen, K. P., Aiden, E. L. & Kornberg, R. D. Polycomb-mediated chromatin loops revealed by a subkilobase-resolution chromatin interaction map. *Proc. Natl. Acad. Sci. U. S. A.* (2017) doi:10.1073/pnas.1701291114.

159. Kundu, S. *et al.* Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Mol. Cell* (2017) doi:10.1016/j.molcel.2017.01.009.

160. Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* (2014) doi:10.1016/j.cell.2014.11.021.

161. Vieux-Rochas, M., Fabre, P. J., Leleu, M., Duboule, D. & Noordermeer, D. Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear domain. *Proc. Natl. Acad. Sci. U. S. A.* (2015) doi:10.1073/pnas.1504783112.

162. Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* (2017) doi:10.1016/j.cell.2017.09.043.

163. Isono, K. *et al.* SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. *Dev. Cell* (2013) doi:10.1016/j.devcel.2013.08.016.

164. van Steensel, B. & Furlong, E. E. M. The role of transcription in shaping the spatial organization of the genome. *Nature Reviews Molecular Cell Biology* (2019) doi:10.1038/s41580-019-0114-6.

165. Goldman, A. E., Maul, G., Steinert, P. M. & Yang, H. Y. Keratin-like proteins that coisolate with intermediate filaments of BHK-21 cells are nuclear lamins. *Proc. Natl. Acad. Sci. U. S. A.* (1986) doi:10.1073/pnas.83.11.3839.

166. Mckeon, F. D., Kirschner, M. W. & Caput, D. Homologies in both primary and secondary structure between nuclear envelope and intermediate filament proteins. *Nature* (1986) doi:10.1038/319463a0.

167. Rober, R. A., Sauter, H., Weber, K. & Osborn, M. Cells of the cellular immune and hemopoietic system of the mouse lack lamins A/C: Distinction versus other somatic cells. *J. Cell Sci.* (1990).

168. Solovei, I. *et al.* LBR and lamin A/C sequentially tether peripheral heterochromatin and inversely regulate differentiation. *Cell* (2013) doi:10.1016/j.cell.2013.01.009.

169. Guilly, M. N., Bensussan, A., Bourge, J. F., Bornens, M. & Courvalin, J. C. A human T lymphoblastic cell line lacks lamins A and C. *EMBO J.* (1987) doi:10.1002/j.1460-2075.1987.tb02715.x.

170. Stewart, C. & Burke, B. Teratocarcinoma stem cells and early mouse embryos contain only a single major lamin polypeptide closely resembling lamin B. *Cell* (1987) doi:10.1016/0092-8674(87)90634-9.

171. Collas, P., Lund, E. G. & Oldenburg, A. R. Closing the (nuclear) envelope on the genome: How nuclear lamins interact with promoters and modulate gene expression. *BioEssays* (2014) doi:10.1002/bies.201300138.

172. Meuleman, W. *et al.* Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res.* (2013) doi:10.1101/gr.141028.112.

173. Meshorer, E. *et al.* Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell* (2006) doi:10.1016/j.devcel.2005.10.017.

174. Jørgensen, H. F. *et al.* The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells. *Genome Biol.* (2007) doi:10.1186/gb-2007-8-8-r169.

175. Aoto, T., Saitoh, N., Ichimura, T., Niwa, H. & Nakao,

M. Nuclear and chromatin reorganization in the MHC-Oct3/4 locus at developmental phases of embryonic stem cell differentiation. *Dev. Biol.* (2006) doi:10.1016/j.ydbio.2006.04.450.

176. Kobayakawa, S., Miike, K., Nakao, M. & Abe, K. Dynamic changes in the epigenomic state and nuclear organization of differentiating mouse embryonic stem cells. *Genes to Cells* (2007) doi:10.1111/j.1365-2443.2007.01063.x.

177. Park, S. H. *et al.* Ultrastructure of human embryonic stem cells and spontaneous and retinoic acid-induced differentiating cells. *Ultrastruct. Pathol.* (2004) doi:10.1080/01913120490515595.

178. Meshorer, E. & Misteli, T. Chromatin in pluripotent embryonic stem cells and differentiation. *Nature Reviews Molecular Cell Biology* (2006) doi:10.1038/nrm1938.

179. Bhattacharya, D., Talwar, S., Mazumder, A. & Shivashankar, G. V. Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during Drosophila embryogenesis. *Biophys. J.* (2009) doi:10.1016/j.bpj.2008.11.075.

180. Constantinescu, D., Gray, H. L., Sammak, P. J., Schatten, G. P. & Csoka, A. B. Lamin A/C Expression Is a Marker of Mouse and Human Embryonic Stem Cell Differentiation. *Stem Cells* (2006) doi:10.1634/stemcells.2004-0159.

181. Efroni, S., Melcer, S., Nissim-Rafinia, M. & Meshorer, E. Stem cells do play with dice: A statistical physics view of transcription. *Cell Cycle* (2009) doi:10.4161/cc.8.1.7216.

182. Bártová, E. *et al.* Epigenome and chromatin structure in human embryonic stem cells undergoing differentiation. *Dev. Dyn.* (2008) doi:10.1002/dvdy.21773.

183. Mattout, A. & Meshorer, E. Chromatin plasticity and genome organization in pluripotent embryonic stem cells. *Current Opinion in Cell Biology* (2010) doi:10.1016/j.ceb.2010.02.001.

184. Peric-Hupkes, D. *et al.* Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation. *Mol. Cell* (2010) doi:10.1016/j.molcel.2010.03.016.

185. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* (2006) doi:10.1016/j.cell.2006.02.041.

186. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* (2007) doi:10.1038/nature06145.

187. Hong, S. H. *et al.* Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl. Acad. Sci. U. S. A.* (2007) doi:10.1073/pnas.0707292104.

188. Lan, F., Nottke, A. C. & Shi, Y. Mechanisms involved in the regulation of histone lysine demethylases. *Current Opinion in Cell Biology* (2008) doi:10.1016/j.ceb.2008.03.004.

189. Min, G. L. *et al.* Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science (80-. ).* (2007) doi:10.1126/science.1149042.

190. Xiang, Y. *et al.* JMJD3 is a histone H3K27 demethylase. *Cell Res.* (2007) doi:10.1038/cr.2007.83.

191. Cloos, P. A. C., Christensen, J., Agger, K. & Helin, K. Erasing the methyl mark: Histone demethylases at the center of cellular differentiation and disease. *Genes and Development* (2008) doi:10.1101/gad.1652908.

192. Smith, A. Formative pluripotency: The executive phase in a developmental continuum. *Dev.* (2017) doi:10.1242/dev.142679.

193. Faust, C., Lawson, K. A., Schork, N. J., Thiel, B. & Magnuson, T. The Polycomb-group gene eed is required for normal morphogenetic movements during gastrulation in the mouse embryo. *Development* (1998).

194. Faust, C., Schumacher, A., Holdener, B. & Magnuson, T. The eed mutation disrupts anterior mesoderm production in mice. *Development* (1995).

195. Pasini, D., Bracken, A. P., Jensen, M. R., Denchi, E. L. & Helin, K. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J.* (2004) doi:10.1038/sj.emboj.7600402.

196. O'Carroll, D. *et al.* The Polycomb-Group GeneEzh2 Is Required for Early Mouse Development. *Mol. Cell. Biol.* (2001) doi:10.1128/mcb.21.13.4330-4336.2001.

197. Akasaka, T. *et al.* Mice doubly deficient for the polycomb group genes Mel18 and Bmi1 reveal synergy and requirement for maintenance but not initiation of Hox gene expression. *Development* (2001).

198. Coré, N. *et al.* Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development* (1997).

199. del Mar Lorente, M. *et al.* Loss- and gain-of-function mutations show a Polycomb group function for Ring1A in mice. *Development* (2000).

200. Endoh, M. *et al.* Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**, 1513 LP – 1524 (2008).

201. Takihara, Y. *et al.* Targeted disruption of the mouse homologue of the Drosophila polyhomeotic gene leads to altered anteroposterior patterning and neural crest defects. *Development* (1997).

202. Van Der Lugt, N. M. T., Alkema, M., Berns, A. & Deschamps, J. The Polycomb-group homolog Bmi-1 is a regulator of murine Hox gene expression. *Mech. Dev.* (1996) doi:10.1016/S0925-4773(96)00570-9.

203. Boyer, L. A., Mathur, D. & Jaenisch, R. Molecular control of pluripotency. *Current Opinion in Genetics and Development* (2006) doi:10.1016/j.gde.2006.08.009.

204. Ku, M. *et al.* Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* (2008) doi:10.1371/journal.pgen.1000242.

205. Lee, T. I. *et al.* Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells. *Cell* (2006) doi:10.1016/j.cell.2006.02.043.

206. Chamberlain, S. J., Yee, D. & Magnuson, T. Polycomb Repressive Complex 2 Is Dispensable for Maintenance of Embryonic Stem Cell Pluripotency. *Stem Cells* (2008) doi:10.1634/stemcells.2008-0102.

207. Pasini, D., Bracken, A. P., Hansen, J. B., Capillo, M. & Helin, K. The Polycomb Group Protein Suz12 Is Required for Embryonic Stem Cell Differentiation. *Mol. Cell. Biol.* (2007) doi:10.1128/mcb.01432-06.

208. Shen, X. *et al.* EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency. *Mol. Cell* (2008) doi:10.1016/j.molcel.2008.10.016.

209. Dodge, J. E., Kang, Y.-K., Beppu, H., Lei, H. & Li, E. Histone H3-K9 Methyltransferase ESET Is Essential for Early Development. *Mol. Cell. Biol.* (2004) doi:10.1128/mcb.24.6.2478-2486.2004.

210. Peters, A. H. F. M. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* (2001) doi:10.1016/S0092-8674(01)00542-6.

211. Aucott, R. *et al.* HP1 -β is required for development of the cerebral neocortex and neuromuscular junctions. *J. Cell Biol.* (2008) doi:10.1083/jcb.200804041.

212. Bilodeau, S., Kagey, M. H., Frampton, G. M., Rahl, P. B. & Young, R. A. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* (2009) doi:10.1101/gad.1837309.

213. Lohmann, F. *et al.* KMT1E mediated H3K9 methylation is required for the maintenance of embryonic stem cells by repressing trophectoderm differentiation. *Stem Cells* (2010) doi:10.1002/stem.278.

214. Yuan, P. *et al.* Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. *Genes Dev.* (2009) doi:10.1101/gad.1831909.

215. Yeap, L.-S., Hayashi, K. & Surani, M. A. ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. *Epigenetics Chromatin* (2009) doi:10.1186/1756-8935-2-12.

216. Epsztejn-Litman, S. *et al.* De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat. Struct. Mol. Biol.* (2008) doi:10.1038/nsmb.1476.

217. Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* (2012) doi:10.1016/j.cell.2012.09.045.

218. Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* (2012) doi:10.1038/nature10953.

219. Matoba, S. *et al.* Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. *Cell* (2014) doi:10.1016/j.cell.2014.09.055.

220. Brehm, A., Ovitt, C. E. & Schöler, H. R. Oct-4: More than just a POUerful marker of the mammalian germline? in *APMIS* (1998). doi:10.1111/j.1699-0463.1998.tb01326.x.

221. Pesce, M. & Schöler, H. R. Oct-4: Control of totipotency and germline determination. *Mol. Reprod. Dev.* (2000) doi:10.1002/(SICI)1098-2795(200004)55:4<452::AID-MRD14>3.0.CO;2-S.

222. Liu, N. *et al.* Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability. *Genes Dev.* (2015) doi:10.1101/gad.254425.114.

223. Ooi, J. *et al.* Unbiased Profiling of Isogenic Huntington Disease hPSC-Derived CNS and Peripheral Cells Reveals Strong Cell-Type Specificity of CAG Length Effects. *Cell Rep.* **26**, 2494-2508.e7 (2019).

224. Xu, X. *et al.* Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. *Stem Cell Reports* **8**, 619–633 (2017).

225. Elias, L. A. B., Potter, G. B. & Kriegstein, A. R. A Time and a Place for Nkx2-1 in Interneuron Specification and Migration. *Neuron* (2008) doi:10.1016/j.neuron.2008.08.017.

226. Teo, A. K. K. *et al.* Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes Dev.* (2011) doi:10.1101/gad.607311.

227. Zuccato, C. & Cattaneo, E. Brain-derived neurotrophic factor in neurodegenerative diseases. *Nature Reviews Neurology* (2009) doi:10.1038/nrneurol.2009.54.

228. Zuccato, C. & Cattaneo, E. Role of brain-derived neurotrophic factor in Huntington's disease. *Progress in Neurobiology* (2007) doi:10.1016/j.pneurobio.2007.01.003.

229. Zuccato, C. *et al.* Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. *Science (80-. ).* (2001) doi:10.1126/science.1059581.

230. Zuccato, C. *et al.* Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. *Nat. Genet.* (2003) doi:10.1038/ng1219.

231. Zuccato, C. *et al.* Systematic assessment of BDNF and its receptor levels in human cortices affected by Huntington's disease. *Brain Pathol.* (2008) doi:10.1111/j.1750-3639.2007.00111.x.

232. Endele, S. *et al.* Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat. Genet.* (2010) doi:10.1038/ng.677.

233. Al-Jaberi, N., Lindsay, S., Sarma, S., Bayatti, N. & Clowry, G. J. The early fetal development of human neocortical GABAergic interneurons. *Cereb. Cortex* (2015) doi:10.1093/cercor/bht254.

234. Fahrenkrug, J., Popovic, N., Georg, B., Brundin, P. & Hannibal, J. Decreased VIP and VPAC2 receptor expression in the biological clock of the R6/2 Huntington's disease mouse. *J. Mol. Neurosci.* (2007) doi:10.1385/jmn/31:02:139.

235. Xie, Y., Hayden, M. R. & Xu, B. BDNF overexpression in the forebrain rescues Huntington's disease phenotypes in YAC128 mice. *J. Neurosci.* (2010) doi:10.1523/JNEUROSCI.1637-10.2010.

236. Obier, N. *et al.* Polycomb Protein EED is Required for Silencing of Pluripotency Genes upon ESC Differentiation. *Stem Cell Rev. Reports* (2015) doi:10.1007/s12015-014-9550-z.

237. Cmarko, D., Verschure, P. J., Otte, A. P., van Driel, R. & Fakan, S. Polycomb group gene silencing proteins are concentrated in the perichromatin compartment of the mammalian nucleus. *Journal of Cell Science* (2003) doi:10.1242/jcs.00225.

238. Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* (2016) doi:10.1038/nature16496.

239. Pachano, T., Crispatzu, G. & Rada-Iglesias, A. Polycomb proteins as organizers of 3D genome architecture in embryonic stem cells. *Briefings in Functional Genomics* (2019) doi:10.1093/bfgp/elz022.

240. Cesarini, E. *et al.* Lamin A/C sustains PcG protein architecture, maintaining transcriptional repression at target genes. *J. Cell Biol.* (2015) doi:10.1083/jcb.201504035.

241. Mas, G. & Di Croce, L. The role of Polycomb in stem cell genome architecture. *Current Opinion in Cell Biology* (2016) doi:10.1016/j.ceb.2016.09.006.

242. Schoenfelder, S. *et al.* The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* (2015) doi:10.1101/gr.185272.114.

243. Mas, G. *et al.* Promoter bivalency favors an open chromatin architecture in embryonic stem cells. *Nat. Genet.* (2018) doi:10.1038/s41588-018-0218-5.

244. Stephens, A. D. *et al.* Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Mol. Biol. Cell* (2018) doi:10.1091/mbc.E17-06-0410.

245. Gasset-Rosa, F. *et al.* Polyglutamine-Expanded Huntingtin Exacerbates Age-Related Disruption of Nuclear Integrity and Nucleocytoplasmic Transport. *Neuron* (2017) doi:10.1016/j.neuron.2017.03.027.

246. Muñoz-Espín, D. *et al.* XProgrammed cell senescence during mammalian embryonic development. *Cell* (2013) doi:10.1016/j.cell.2013.10.019.

247. Storer, M. *et al.* XSenescence is a developmental mechanism that contributes to embryonic growth and patterning. *Cell* (2013) doi:10.1016/j.cell.2013.10.041.

248. Rhinn, M., Ritschka, B. & Keyes, W. M. Cellular senescence in development, regeneration and disease. *Dev.* (2019) doi:10.1242/dev.151837.

249.   Ito, T., Teo, Y. V., Evans, S. A., Neretti, N. & Sedivy, J. M. Regulation of Cellular Senescence by Polycomb Chromatin Modifiers through Distinct DNA Damage- and Histone Methylation-Dependent Pathways. *Cell Rep.* (2018) doi:10.1016/j.celrep.2018.03.002.

250.   Wang, A. S., Ong, P. F., Chojnowski, A., Clavel, C. & Dreesen, O. Loss of lamin B1 is a biomarker to quantify cellular senescence in photoaged skin. *Sci. Rep.* (2017) doi:10.1038/s41598-017-15901-9.

251.   Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* (2010) doi:10.1016/j.stem.2010.03.018.

252.   Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* (2013) doi:10.1016/j.cell.2012.12.033.

253.   Fadloun, A., Eid, A. & Torres-Padilla, M. E. Mechanisms and Dynamics of Heterochromatin Formation During Mammalian Development. Closed Paths and Open Questions. in *Current Topics in Developmental Biology* (2013). doi:10.1016/B978-0-12-416027-9.00001-2.

254.   Burton, A. & Torres-Padilla, M. E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nature Reviews Molecular Cell Biology* (2014) doi:10.1038/nrm3885.

255.   Conforti, P. *et al.* RUES2 hESCs exhibit MGE-biased neuronal differentiation and muHTT-dependent defective specification hinting at SP1. *Neurobiol. Dis.* (2020) doi:10.1016/j.nbd.2020.105140.

256.   Vezzoli, E. *et al.* Inhibiting pathologically active ADAM10 rescues synaptic and cognitive decline in Huntington's disease. *J. Clin. Invest.* (2019) doi:10.1172/JCI120616.

# 6 | Appendix

## 6.1 Contributions to published articles

**"RUES2 hESCs exhibit MGE-biased neuronal differentiation and muHTT-dependent defective specification hinting at SP1."**

*Paola Conforti[1,2], Dario Besusso[1,2], <u>Silvia Brocchetti[1,2]</u>, Ilaria Campus[1,2], Claudio Cappadona[1,2], Maura Galimberti[1,2], Angela Laporta[1], Raffaele Iennaco[1,2], Riccardo Rossi[2], Vittoria Bocchi D.[1,2], Elena Cattaneo[1,2*]. Neurobiology of Disease, (2020).*

The aim of this work was to characterize and evaluate the neuronal differentiation potential of the collection of isogenic human embryonic stem cells (hESCs), the RUES2 lines, carrying different pathological CAG lengths in the HTT gene. RUES2 lines resulted to be biased towards the medial ganglionic eminence (MGE) and in particular the HD-RUES2 cells exhibited an altered MGE transcriptional signature together with typical HD phenotypes. My goal in this project was to identify potential muHTT-related interactors whose activity is shared by the different lineage-specific signaling pathways that cooperates during neurogenesis. All results and analysis are under the paragraph "*De novo* motif discovery analysis identifies SP1 as a potential common regulator in HD neural differentiation". The figures produced by this bioinformatics analysis are Figure 6 and Supplementary Figure 8 and Figure 9.

I contributed to the *in vitro* studies and specifically *in silico de novo* motif discovery analysis. I wrote the section of the bioinformatic analysis, in terms of results, methods and discussion, and I was involved in the revision and final approval of the manuscript before publishing.

## 6.2 Published articles

In the next section are the published articles described above.
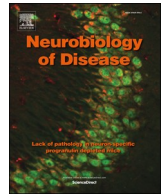
ELSEVIER

# RUES2 hESCs exhibit MGE-biased neuronal differentiation and muHTT-dependent defective specification hinting at SP1

Paola Conforti[a,b], Dario Besusso[a,b], Silvia Brocchetti[a,b], Ilaria Campus[a,b], Claudio Cappadona[a,b], Maura Galimberti[a,b], Angela Laporta[a], Raffaele Iennaco[a,b], Riccardo L. Rossi[b], Vittoria Bocchi Dickinson[a,b], Elena Cattaneo[a,b,*]

[a] *Laboratory of Stem Cell Biology and Pharmacology of Neurodegenerative Diseases, Department of Biosciences, University of Milan, 20122 Milan, Italy*
[b] *Istituto Nazionale Genetica Molecolare, Romeo ed Enrica Invernizzi, Milan 20122, Italy*

## ABSTRACT

RUES2 cell lines represent the first collection of isogenic human embryonic stem cells (hESCs) carrying different pathological CAG lengths in the HTT gene. However, their neuronal differentiation potential has yet to be thoroughly evaluated. Here, we report that RUES2 during ventral telencephalic differentiation is biased towards medial ganglionic eminence (MGE). We also show that HD-RUES2 cells exhibit an altered MGE transcriptional signature in addition to recapitulating known HD phenotypes, with reduced expression of the neurodevelopmental regulators NEUROD1 and BDNF and increased cleavage of synaptically enriched N-cadherin. Finally, we identified the transcription factor SP1 as a common potential detrimental co-partner of muHTT by *de novo* motif discovery analysis on the LGE, MGE, and cortical genes differentially expressed in HD human pluripotent stem cells in our and additional datasets. Taken together, these observations suggest a broad deleterious effect of muHTT in the early phases of neuronal development that may unfold through its altered interaction with SP1.

## 1. Introduction

For many years after the discovery of the abnormally elongated CAG trinucleotides in the huntingtin gene (*HTT*), resulting in a polyglutamine (polyQ) stretch in the HTT protein, as the genetic cause of Huntington's disease (HD), the pathogenesis was studied primarily in vivo using genetic models (Farshim and Bates, 2018). These studies provided valuable information on how mutant HTT (muHTT) causes progressive degeneration of the medium-spiny striatal neurons (MSNs), in addition to the dysfunction and progressive loss of neurons in the cerebral cortex. Mutant HTT is known to form intracellular oligomers and aggregates, and is thought to have a toxic gain-of-function that interferes with many cellular and biological processes (Rataj-Baniowska et al., 2015), causing transcriptional, metabolic, and axonal transport defects and excitotoxicity in neurons. Glial cell function is also affected (Osipovitch et al., 2019). The mutation in CAG repeat length varies in patients, with longer CAG repeats correlating with earlier HD onset (Gusella and MacDonald, 2006; Lee et al., 2012a, 2012b). Accordingly, in an attempt to recapitulate this correlation between the pathological

CAG sizes and the clinical symptoms observed in patients, knock-in genetic models have been generated by inserting a progressively increasing CAG repeat, ranging from 50 (White et al., 1997) to 150 (Lin et al., 2001), in the endogenous HTT gene. Knock-in models develop behavioral deficits at a very early age, even before neuropathology, though none of these models exhibit neuronal death or gliosis, even in very old mice (Lucas and Ortega, 2011). These results demonstrate that neuronal dysfunction precedes neuronal death, which is in agreement with the presence of subtle clinical symptoms and cortical thinning in HD patients decades before the appearance of relevant motor signs (Nanetti et al., 2018; Sampedro et al., 2019; Smith et al., 2000).

The advent of human pluripotent stem cell (hPSC) technologies and neuronal differentiation methods have added valuable information to HD research. Disease-relevant human MSNs and cortical neurons can be obtained from hPSCs carrying different sizes of CAG repeats in *HTT* exon 1 and used to identify disease-perturbed regulatory networks during the course of neuronal differentiation *in vitro* and in mature neurons. Some studies have employed patient-derived human induced pluripotent stem cells (hiPSCs), which have the advantage of carrying

the mutation in the context of the genome of the donor in association with the corresponding clinical information. Another relevant cell model is represented by an engineered allelic series of isogenic human embryonic stem cell (hESC) lines carrying a gradual increase in CAG repeat length (Ruzo et al., 2018). Common evidence from all of these studies is the demonstration that hPSCs carrying the HD mutation exhibit early neuronal defects. In particular, one study found that the expanded CAG negatively impacts the transition from the pluripotent stem cell state to neuroectoderm, as well as the formation of polarized neuroepithelial structures (Conforti et al., 2018). In other studies, neuruloid morphogenesis (Haremaki et al., 2019) and cytokinesis (Ruzo et al., 2018) were severely affected, with giant cells appearing in the dish and resulting in developmental delay. Altered neuronal gene transcription, electrophysiology, metabolism, and cleavage of critical cell adhesion molecules that are enriched in synapses (e.g., N-cadherin, NCAD) have also been reported, consistent with impaired neurogenesis (Conforti et al., 2018; HD iPSC Consortium, 2012, 2017). Defects in the differentiation of MSNs, cortical progenitor homeostasis and cytoarchitecture in 3D organoids (Conforti et al., 2018), and transcriptional alterations in HD cortical neurons have also been described (Mehta et al., 2018). These data indicate that the expression of muHTT in human cells exposed to cortical and striatal differentiation protocols causes delayed maturation and altered morphology of several neuronal subtypes, implying early abnormalities of neurological development, at least in vitro. Such impaired neurological development would compromise neuronal homeostasis in adulthood, leading to greater cellular vulnerability to advanced life stressors (Ramocki and Zoghbi, 2008).

Evidence mostly from animal models indicates that HTT has multiple roles during neurogenesis; therefore, its mutation may cause dysfunctions, starting during embryogenesis (Godin et al., 2010; Lopes et al., 2016; HD iPSC Consortium, 2012; Molero et al., 2009, 2016). Furthermore, expression of polyQ-expanded HTT up to postnatal day 21 in mice has been found to recapitulate some of the typical biochemical and behavioral HD symptoms observed in animals expressing muHTT throughout life (Molero et al., 2009, 2016). The possibility of altered brain development is also suggested by evidence of a smaller intracranial brain volume (Nopoulos et al., 2011) and smaller head size (Lee et al., 2012a) in HD carriers decades before predicted motor onset. This possibility seems to find support in a recent study that compared tissue fragments from control and HD human fetal brains of various ages and concluded in favor of abnormal neuronal development, though the scarcity of data and lack of relevant evidence from HD samples raises some concerns (Barnat et al., 2020). Detailed analysis of cells and recognizable brain tissue is indispensable for establishing whether the basic mechanisms of neural development are altered in the presence of an HD mutation. In this context, hPSCs carrying the HD mutation and exposed to neuronal differentiation can help paint the picture.

Here, we report that, upon exposure to a striatal differentiation protocol (Delli Carri et al., 2013), the hESC line H9 and hiPSC line KOLF2 up-regulate markers of the lateral ganglionic eminences (LGE), the brain region that gives rise to the MSNs of the striatum during development, whereas the RUES2 parental line acquires a significantly more ventralized identity, as indicated by up-regulation of medial ganglionic eminence (MGE) transcripts. Presence of an expanded CAG in the RUES2 line negatively affects the MGE transcriptional signature, as well as levels of other neurodevelopment transcripts, including BDNF. These data, together with the literature, show that the detrimental effect of muHTT in hPSC differentiation is not lineage-specific, but affects multiple neuronal types. *In silico* analysis highlighted transcription factor (TF) SP1 as a potential common co-partner of adverse muHTT activity during neuronal differentiation. Taken together, these observations suggest a broad deleterious effect of muHTT during the early stages of neuronal development that may unfold through its altered interaction with SP1.

## 2. Material and methods

### 2.1. Human cell cultures

RUES2 hESC lines derived and kindly provided by Prof. Ali Brinvalou's Laboratory (Rockefeller University) and H9 hESCs (WiCell Research Institute) were cultured in mTeSR1 (Supplement Fig. 1A). Proliferating RUES2 cells grew in adhesion on Geltrex™ (Thermo Fisher Scientific), whereas H9 cells were maintained on Cultrex coated plates (Trevigen). KOLF2 hiPSCs were kindly provided by the Sanger Institute's Human Induced Pluripotent Stem Cell Initiative (HipSci) project. Pluripotent KOLF2 cells were grown in TeSR-E8 (Voden) on tissue culture plates coated with Synthemax II-SC substrate (Corning) under feeder-free conditions (Supplement Fig. 1A). Additional control (Q21n1) and HD (Q60n5, Q109n1) hiPSCs generated and characterized in previous studies (HD iPSC Consortium, 2017; Conforti et al., 2018) were cultured in mTESR1 medium (Voden) and plated on Cultrex (Trevigen).

All cell lines used in this study were maintained under sterile conditions and regularly tested for mycoplasma (Eurofins). The karyotype of each line/clone was monitored every 3 months by Q-banding analysis (ISENET).

### 2.2. Striatal differentiation

HD and control hPSC lines were exposed to the striatal differentiation protocol published by Delli Carri et al. (2013). Briefly, cells were plated at a density of $0.6 \times 10^5$ cells/cm$^2$ on Cultrex-coated plates (120–180 μg/ml) in complete mTeSRTM1 medium and expanded for 2 days. When cultures reached 70% confluence, cells were exposed to neural induction medium [DMEM-F12 (Gibco), N2 supplement (Gibco), B-27 without retinoic acid (Gibco), 10 μM SB431542, and 500 nM LDN193189] for 12 days. Neural induction lasts 12 days and is triggered by the dual SMAD inhibitor compounds SB431542 (a blocker of the TGF-β pathway) and LDN193189 (an inhibitor of the BMP pathway) (Chambers et al., 2009). These two reagents were provided by CHDI Foundation (NY, USA). Patterning and specification of the cells towards the ventral telencephalic fate lasts from day 5 to day 25. In this time window 200 ng/ml recombinant human SHH (PrepoTech) and 100 ng/ml DKK-1 (PrepoTech) were added. At DIV15, cells were detached by Accutase single cell dissociation and replated at a density of $2.5 \times 10^4$ cells/cm$^2$ on Matrigel-coated plates (240–360 μg/ml). Neuronal maturation lasts from day 25 to day 50 and is achieved in the presence of 30 ng/ml BDNF (PrepoTech,) in DMEM-F12 plus N-2 and B-27 supplements (Gibco).

### 2.3. Western blot

At specific time points of differentiation, cell cultures were collected and lysed in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.1% SDS, 1% NP40) supplemented with 1 mM PMSF and Halt Protease & Phosphatase Inhibitor Cocktail (Thermo Scientific). The protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo Scientific). Equal amounts of total proteins were loaded on 7.5% or 10% SDS-PAGE gels and transferred to a nitrocellulose membrane (Bio-Rad) using the Trans-Blot® Turbo™ System (Bio-Rad). Membranes were blocked for 1 h at room temperature (RT) in 0.1% TBS-T with 5% non-fat dry milk (Bio-Rad) and incubated overnight at 4 °C with the following primary antibodies: MAP2a/b (mouse, 1:500; Beckton Dickinson), CTIP2 (rat, 1:500; Abcam), GAD67 (mouse, 1:1000; Millipore), DARPP32 (rabbit, 1:500; Abcam), N-CAD (mouse, 1:500; Beckton Dickinson), and ADAM10 (rabbit, 1:1000; Abcam). Membranes were washed thrice in 0.1% TBS-T and incubated for 1 h at RT with the appropriate HRP-I conjugated secondary antibodies (dilution 1:3000, Bio-Rad). Following three washes, immunoreactive bands

were detected using Clarity™ western ECL Substrate (Bio-Rad) in accordance with the manufacturer's instructions. GAPDH (rabbit, 1:5000; Abcam) was used to normalize expression. Acquisitions were performed using a ChemiDoc MP imaging system (Bio-Rad) and densitometric analysis using ImageLab software.

### 2.4. Immunocytochemistry

Cell cultures were fixed with ice-cold 4% paraformaldehyde for 15 min and washed thrice with phosphate-buffered saline (PBS) at specific times. Next, cells were permeabilized with 0.5% Triton X-100 in PBS for 10 min and subsequently blocked with 5% normal goat serum (NGS, Vector) for 1 h at RT. Cells were incubated overnight at 4 °C with primary antibodies diluted in solution containing 2.5% NGS and 0.25% Triton X-100. Appropriate Alexa Fluor®-conjugated secondary antibodies (Molecular Probe, Life Technologies) were diluted 1:500 in PBS and mixed with 0.1 μg/ml Hoechst (Invitrogen) to counterstain the nuclei. Images were acquired on a Leica TCS SP5 Confocal Laser Scanning Microscope (Leica Microsystems) using a 40× oil immersion objective with a zoom of 1.7. CellProfiler software (version 2.1.1) was used to quantify positive cells.

The following primary antibodies were used: OCT3/4 (mouse, 1:100; Santa Cruz), Ki67 (rabbit, 1:500; Abcam), p27 (mouse, 1:1000; Cell Signaling), PALS1 (rabbit, 1:500; Cell Signaling), N-CAD (mouse, 1:800; Beckton Dickinson), GSX2 (rabbit, 1:250; Millipore), ASCL1 (mouse, 1:1000; Beckton Dickinson), FOXG1 (rabbit, 1:1000; Diatech), NKX2.1/TTF1 (rabbit, 1:2500; Abcam), CTIP2 (rat, 1:1000; Abcam), DARPP32 (rabbit, 1:250; Abcam), TBR1 (rabbit, 1:1000; Abcam), TBR2 (rabbit, 1:1000; Abcam).

### 2.5. Real-time qPCR and biomark analysis

RNA was prepared using Trizol reagent according to the manufacturer's instructions (Thermo Fisher Scientific). After treating RNA with the DNA-free Kit (Ambion, Invitrogen) to remove contaminating DNA, 500 μg RNA was retrotranscribed using the iScript cDNA Synthesis Kit (Bio-Rad). Conventional qPCR was performed in a 15 μl volume containing 50 ng cDNA and SsoFast™ EvaGreen® Supermix 2× (Bio-Rad) in a CFX96 Real-Time System (Bio-Rad). Gene expression analysis was performed using this method for the BDNF total transcript (forward: 5′-TAACGGCGGCAGACAAAAGA-3′; reverse: 5′-GAAGTATTGCTTCA GTTGGCCT-3′) and reference gene 18S (forward: 5'-CGGCTACCACAT CCAAGGAA-3′; reverse: 5'-GCTGGAATTACCGCGGCT-3′).

All other gene expression analyses were carried out using Fluidigm 96.96 dynamic arrays (Fluidigm Corporation, CA, USA) according to the manufacturer's instructions. Pre-amplification reactions were prepared by mixing PreAmp Master Mix, the pooled TaqMan assay 0.2×, and nuclease-free water. The panel of all TaqMan probes is reported in Supplementary Table 2. Briefly, for each sample, 3.75 μl of PreAmp Mix was aliquoted in a 96-well PCR plate and 1.25 μl of newly retrotranscribed cDNA added to reach a total volume of 5 μl. The pre-amplification was performed in a GeneAmp PCR System 9700 (Applied Biosystems). The 96.96 dynamic array IFC Chips (Fluidigm) were primed with the IFC controller fluid (Fluidigm). After priming, the dynamic array was loaded with 5 μl of each assay and sample mix into the appropriate inlets of the primed chip and loaded with the IFC controller. After loading, the chip was placed in the Biomark instrument and qPCR performed.

### 2.6. De novo motif discovery analysis

The gene IDs (obtained from NCBI) of the selected genes were submitted to RSAT retrieve sequences (http://rsat.sb-roscoff.fr/retrieve-seq_form.cgi) to obtain the 3000 bp upstream the genes. The searching parameters were set as follows: in mandatory options "Single Organism_ *Homo sapiens* GRCh38, in advanced options "Prevent

overlap" unselected and "Mask repeats" selected. Then, the output was loaded in two *de novo* web accessible motif discovery tools: MEME (http://meme-suite.org/tools/meme) and RSAT Oligo Analysis (http://rsat.sb-roscoff.fr/retrieve-seq_form.cgi). MEME parameters were set as follow: the number of motifs that MEME should find was set to 12, the background was set to 1st order model of sequences, minimum width of 6 and maximum width of 12, the other parameters were as default. Oligo Analysis parameters were: *H. sapiens* was selected as organism, the estimate from input sequence was set to Markov model (higher order of dependencies) order 2, max matrices 12, min site weight 6. The motifs identified were submitted to STAMP (http://www.benoslab.pitt.edu/stamp/) to identify the 10 best matches to each of the discovered motif in JASPAR v2010 and TRANSFAC v11.3 databases.

The motifs identified in the 4 different output (MEME_JASPAR, MEME_TRANSFAC, Oligo_JASPAR, Oligo_TRANSFAC) were compared and only the TFs represented in at least 2 different motif discovery tools with high score were considered. Each predicted TFs was examined looking for literature evidences in respect to the biological role and disease context. To further investigate and confirm the predicted interactions we submitted the list of the genes of a specific area (i.e LGE) and the respective TFs found to public pathway and interactions database Pathway Commons (https://www.pathwaycommons.org). Many interactions were confirmed, and the networks obtained were imported in Cytoscape 3.7.1 (http://cytoscape.org/) and enriched with the predicted interactions.

### 2.7. Statistical analysis

All statistical analyses were performed using Prism (GraphPad Software). One-way ANOVA with Tukey's post-hoc test was performed for all biological experiments including HD and control lines. Student *t*-test was used to compare only two group conditions. The numbers in each individual experiment and tests used are described in the figure legends. $P < 0.05$ was considered significant.

## 3. Results

### 3.1. Diverse hPSC lines present different propensities to respond to a striatal MSN differentiation protocol

To investigate the neuronal differentiation potential of the parental RUES2 line and, subsequently, its HD derivatives, we exposed cells to a stepwise striatal differentiation protocol that mimics human ventral telencephalon development (Delli Carri et al., 2013; Onorati et al., 2014) (Fig. S1A), together with H9 and KOLF2 cell lines (clone C1; Fig. S1B). These three cell lines all exhibited a normal karyotype (Fig. S1C).

First, we compared the H9 and RUES2 cell proliferative state by evaluating the proportion of Ki67$^+$ and p27$^+$ cells at DIV0, 15, 25, and 50 (Figs. 1A, B). The percentage of Ki67$^+$ cells decreased over time with an increase in p27$^+$ post-mitotic neurons, suggesting that both lines exited the cell cycle (Figs. 1C, D) and acquired neuronal morphologies with similar timing (Fig. 1A and S1D). Next, we investigated the ability of the lines to acquire a striatal-MSN cell identity. Similar levels of telencephalic marker FOXG1 were found in both lines at DIV15 (Figs. 1E, F). In contrast, when counting the number of GSX2 and ASCL1-positive cells at DIV25, the RUES2 culture was significantly reduced in GSX2 or ASCL1-positive and GSX2$^+$/ASCL1$^+$ double-positive progenitors (Figs. 1G, H). Co-expression of CTIP2 with DARPP32 qualifies cells *in vitro* as being bona-fide MSNs (Besusso et al., 2020). Immunocytochemistry and cell counting revealed that RUES2 cells did not express DARPP32 at the end of the differentiation, as indicated by the absence of CTIP2$^+$/DARPP32$^+$ neurons (Figs. 1I, J). Western blot analysis further confirmed this observation. Although the RUES2 line behaved similar to the H9 line in terms of MAP2, CTIP2, and GAD67 protein levels, DARPP32 was almost undetectable at the end of differentiation (Figs. 1K, L and S1E). The KOLF2 hiPS line exposed to the
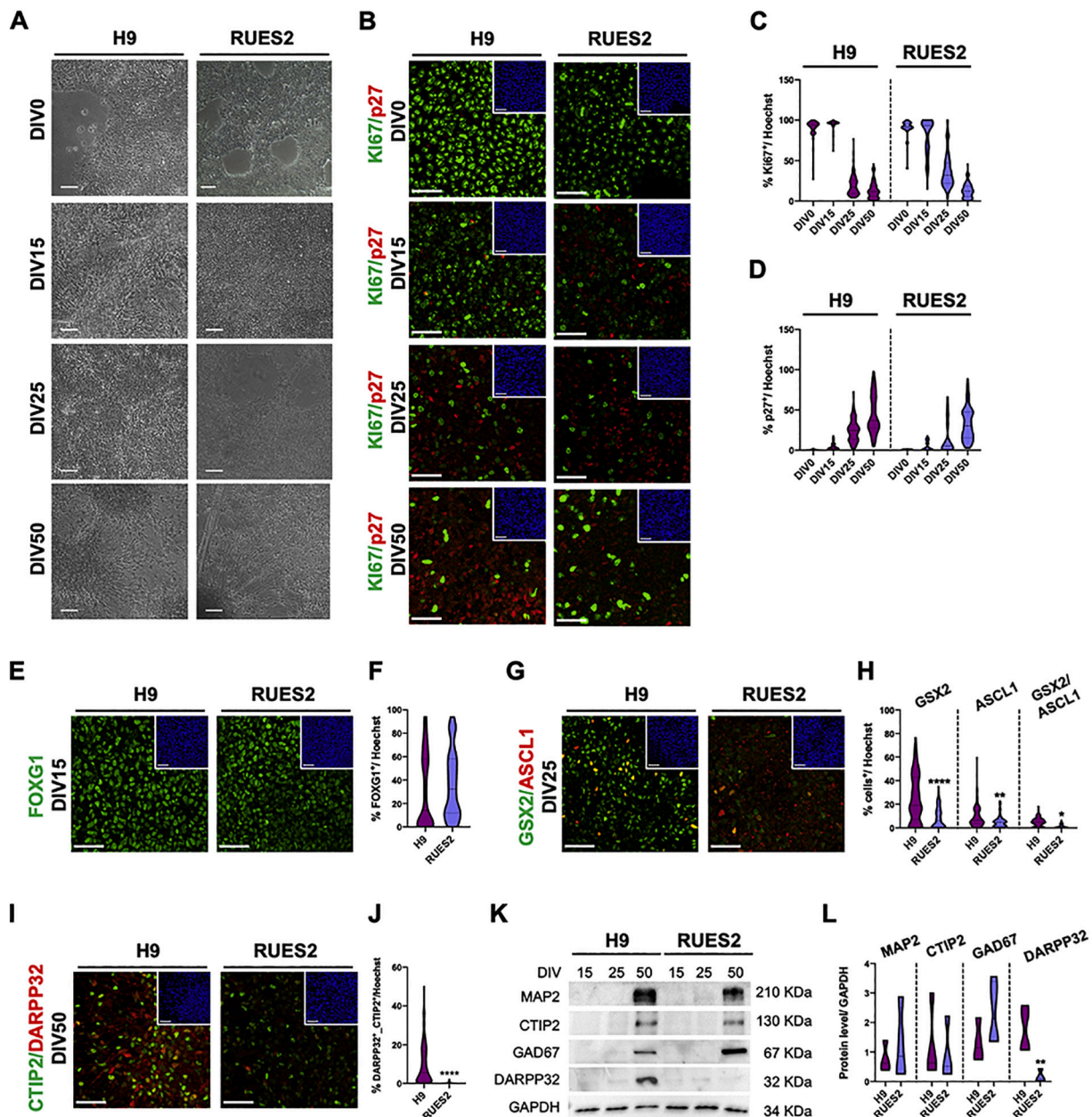
**Fig. 1.** Different hPSC lines present different propensities to respond to striatal differentiation.

(A) Cell morphology of the H9 and RUES2 lines was monitored during differentiation (DIV0, DIV15, DIV25, and DIV50) by phase contrast microscopy. Representative phase contrast images (4× objective) are shown for each time point with an additional 10× objective image at DIV50 to highlight the presence of neurite outgrowth, neurofilaments, and processes from cell bodies at the end of differentiation. Scale bar, 200 μm.

(B) Immunocytochemistry of Ki67 and p27 at DIV0, DIV15, DIV25, and DIV50 of differentiation for the H9 and RUES2 lines. Scale bar, 50 μm. Top right, Hoechst inset.

(C-D) Counts of Ki67$^+$ and p27$^+$ cells for the H9 and RUES2 lines by CellProfiler software (version 2.1.1). $N = 4$ biological replicates for H9 and $N = 3$ for RUES2.

(E) Immunocytochemistry of FOXG1 at DIV15 of differentiation for the H9 and RUES2 lines. Scale bar, 50 μm. Top right, Hoechst inset.

(F) Counts of FOXG1$^+$ cells for the H9 and RUES2 lines by CellProfiler software (version 2.1.1). N = 4 biological replicates for H9 and $N = 3$ for RUES2.

(G) Immunocytochemistry of GSX2 and ASCL1 at DIV25 for the H9 and RUES2 lines. Scale bar, 50 μm. Top right, Hoechst inset.

(H) Counts of GSX2$^+$, ASCL1$^+$, and GSX2$^+$/ASCL1$^+$ cells for the H9 and RUES2 lines by CellProfiler software (version 2.1.1). N = 4 biological replicates for H9 and $N = 3$ for RUES2. Data are presented as mean ± SEM (H9 GSX2$^+$ DIV25: 24.61% ± 1.73%, RUES2 DIV25: 8.6% ± 0.8%; ASCL1$^+$ DIV25: 9.19% ± 0.73%, RUES2 DIV25: 5.12% ± 0.45%; GSX2$^+$/ASCL1$^+$ DIV25: 4.85% ± 0.31%, RUES2 DIV25: 1.21% ± 0.13%) ****$p < 0.0001$, **$p < 0.01$, *$p < 0.05$ using Student *t*-test to compare the two group conditions.

(I) Immunocytochemistry of CTIP2 and DARPP32 at the end of differentiation for the H9 and RUES2 lines. Scale bar, 50 μm. Top right, Hoechst inset.

(J) Counts of double-positive neurons (CTIP2$^+$/DARPP32$^+$ cells) for the H9 and RUES2 lines by CellProfiler software (version 2.1.1). N = 4 biological replicates for H9 and $N = 3$ for RUES2. Data are presented as mean ± SEM (H9 DARPP32$^+$: 18.5% ± 0.11% and CTIP2$^+$/DARPP32$^+$ 9.46% ± 0.07%). ****$p < 0.0001$, Student *t*-test.

(K) Western Blot analysis of MAP2, CTIP2, GAD67, and DARPP32 at DIV15, DIV25, and DIV50 of differentiation for the H9 and RUES2 lines. MAP2, CTIP2, GAD67, and DARPP32 protein levels were normalized to GAPDH. $N = 3$ biological replicates.

(L) Violin plot representing the densitometric analysis performed on western blot results. N = 3 biological replicates. Student t-test was used to compare each group condition, **$p < 0.01$.
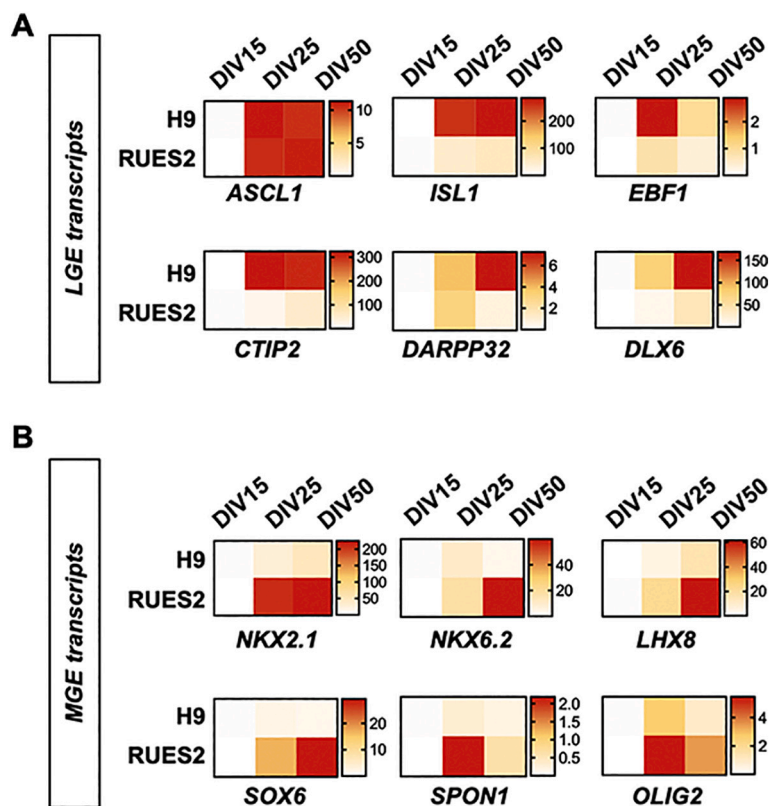
**Fig. 2.** RUES2 lines acquired an MGE-like profile after exposure to a striatal differentiation protocol.

**(A)** Heat map showing the expression levels in high-content qPCR analysis (Biomark) of the LGE at DIV15, DIV25, and DIV50 of the H9 and RUES2 lines. N = 3 biological replicates. *ISL1:* DIV25 *p < 0.05, DIV50 **p < 0.01; *EBF1:* DIV25 **p < 0.01; *CTIP2:* DIV25 **p < 0.01; *DARPP32:* DIV50 ***p < 0.001; *DLX6:* DIV25 **p < 0.01, DIV50 ****p < 0.0001, Student t-test. **(B)** Heat map showing the expression levels in high-content qPCR analysis (Biomark) of the MGE markers at DIV15, DIV25, and DIV50 of the H9 and RUES2 lines. N = 3 biological replicates. *NKX2.1:* DIV25 and DIV50 ****p < 0.0001; *SOX6:* DIV25 *p < 0.05, DIV50 ****p < 0.0001; *SPON1:* DIV25 **p < 0.01; *NKX6.2:* DIV50 ***p < 0.001; *LHX8:* DIV50 ***p < 0.001; *OLIG2:* DIV25 and DIV50 *p < 0.05, Student t-test.

same protocol was similar to H9 (Fig. S1F, G), as both GSX2$^+$/ASCL1$^+$ progenitors were present at DIV25 and CTIP2$^+$/DARPP32$^+$/GAD67$^+$ neurons were counted at the end of the protocol (Fig. S1G). DARPP32 and GAD67 were already detected at DIV25 of differentiation in the KOLF2 line (Fig. S1H, I). These data indicate that hPSC lines exhibit a distinct propensity to respond to the same ventral (striatal) differentiation protocol.

### 3.2. RUES2 line acquired an MGE-like expression profile after exposure to striatal differentiation

To investigate whether RUES2 and H9 lines are differentially specified along the dorso-ventral axis, we performed high-content qPCR analysis on select LGE and MGE markers from DIV15 to DIV50. Analysis of the ventral marker ASCL1 did not show any differential expression at the time points investigated (Fig. 2A; Fig. S2A, B). By monitoring the expression of LGE markers, we confirmed that H9 cells acquire an LGE cell fate through the progressive up-regulation of ISL1, EBF1, CTIP2, DARPP32, and DLX6 (Fig. 2A). The expression profile in differentiating H9 cells recapitulated the developmental transition from VZ to SVZ to mantle zone of the LGE as expected (Delli Carri et al., 2013). In contrast, this set of transcripts was significantly down-regulated in the RUES2 parental line at DIV50 (Fig. 2A; Fig. S2A). The RUES2 line acquired a more ventral MGE-like identity as indicated by the expression of the MGE transcripts NKX2.1, NKX6.2, LHX8, SOX6, SPON1 and OLIG2 (Fig. 2B). Of these mRNAs, NKX2.1, SOX6, and SPON1 were significantly up-regulated starting from DIV25, whereas NKX6.2, LHX8, and OLIG2 were significantly upregulated only at DIV50 (Fig. 2B; Fig. S2B).

In contrast to the H9 line, the RUES2 line did not exhibit any TBR2$^+$/CTIP2$^+$ and TBR1$^+$/CTIP2$^+$ cell populations typically associated with cortical development, further supporting the tendency of this line to acquire a more ventral identity (Fig. S3A, B). Gene expression analysis of the cortical markers TBR1, NEUROD1, and DATCH1 confirmed that these transcripts were all reduced in the RUES2 line compared to the H9 line (Fig. S3C). KOLF2 cells behaved

similarly to H9 cells and expressed transcripts associated with MSN differentiation (Fig. S3D).

### 3.3. Changes in the SHH and WNT signaling pathways contribute to MGE regional identity

In humans, SHH and WNT signaling have opposite effects on regional specification of the ventral and dorsal telencephalon, respectively (Ma et al., 2019), whereas their combination sustains NKX2.1 expression, a transcript that is found in the MGE (Chi et al., 2016) and is up-regulated in differentiating RUES2 cells (Fig. 2B). Therefore, we looked at the main components of these signaling pathways in H9, RUES2, and KOLF2 cultures from DIV0 to DIV10. High-content gene expression analysis revealed no differences between the three cell lines in the mRNA levels of SHH (Fig. 3A; Fig. S4A), Indian Hedgehog (IHH), or the transmembrane protein Smoothened (SMO) (Fig. S4B). However, the RUES2 line presented significant up-regulation of the receptors of HH signaling, PTCH1 and PTCH2, and the transcriptional effectors GLI2 and GLI3 over H9 and KOLF2 cells (Fig. 3A; Fig. S4A), suggesting hyperactivation of SHH signaling in this line. (See Fig. 4.)

Acquisition of a ventral fate requires inhibition of dorsalizing WNT signaling. Therefore, we evaluated the expression of Dickkopf WNT Signaling Pathway Inhibitor 1 (DKK1), an inhibitor of WNT that is added to the medium starting from DIV5 to facilitate the acquisition of a ventral identity. As reported previously (Fasano et al., 2010), DKK1 expression peaked in H9 at DIV4, followed by progressive decline (Fig. 3B; Fig. S4A). In contrast, the RUES2 line maintained low levels of this transcript for the entire induction phase, whereas KOLF2 cells were similar to H9 cells (Fig. 3B). In addition, the RUES2 line exhibited a significant up-regulation of WNT ligands associated with the development of MGE interneurons, such as WNT1, WNT3A, and WNT5A (Chi et al., 2016; Zhang and Zhang, 2010) starting from DIV8 (Fig. 3B; Fig. S4A), whereas WNT8B, which is typically expressed in cortical neurons (Hasenpusch-Theil et al., 2017), was more highly expressed in H9 cells than RUES2 cells (Fig. S4C), supporting our previous analysis that the RUES2 line has a more ventralized identity than
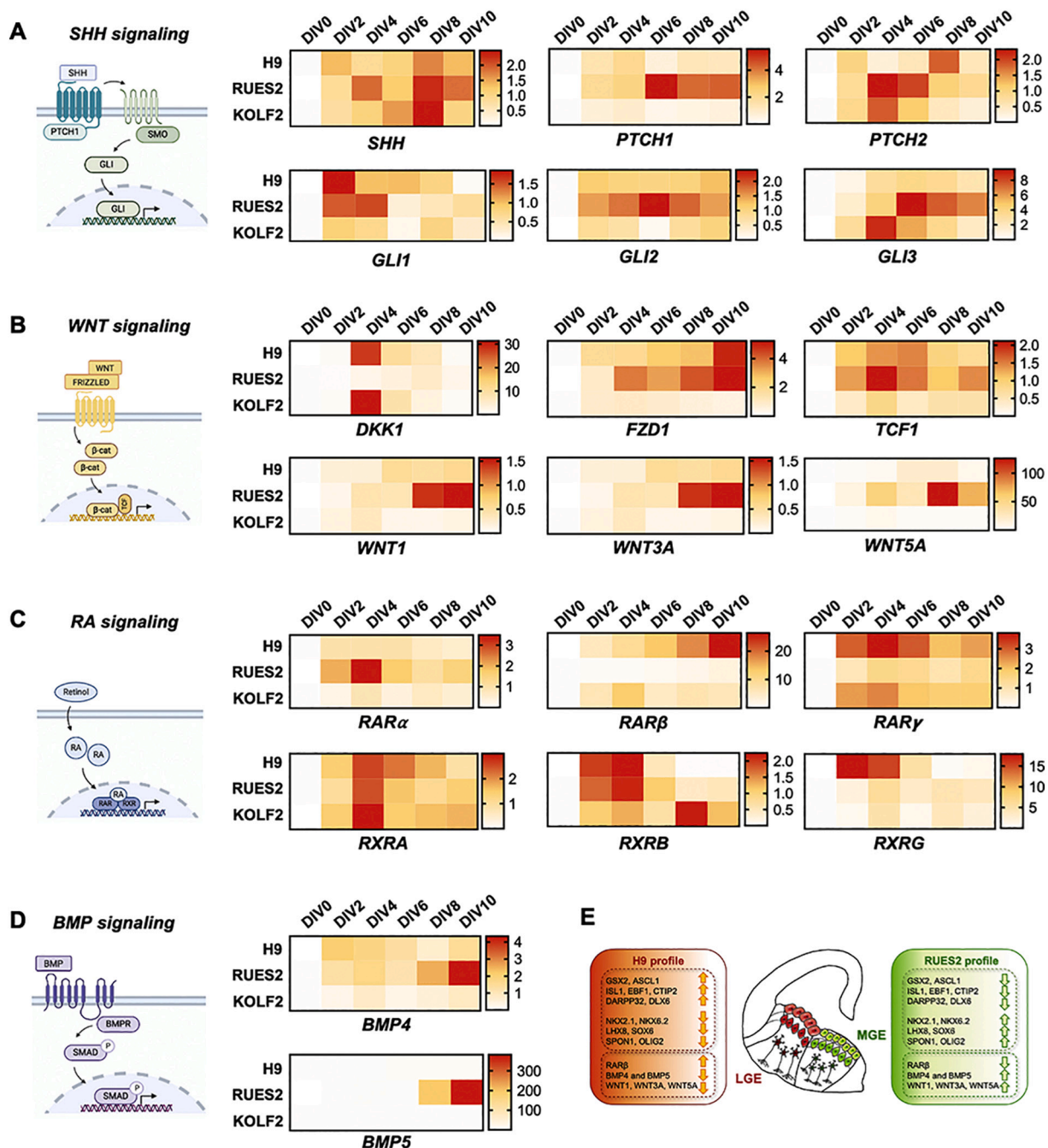
**Fig. 3.** Changes in the SHH and WNT signaling pathways contribute to defining MGE regional identity acquisition.
Summary diagram of signaling pathway components and heat map showing the expression levels in high-content qPCR analysis (Biomark) of **(A)** SHH, **(B)** WNT, **(C)** RA, and **(D)** BMP signaling components at DIV0, DIV2, DIV4, DIV6, DIV8, and DIV10 of differentiation of the H9, RUES2, and KOLF2-C1 lines. N = 3 biological replicates. H9 vs. RUES2, *PTCH1:* DIV6 **p < 0.01, *PTCH2:* DIV4 **p < 0.01, *DKK1:* DIV4 *p < 0.05, *WNT1:* DIV10 *p < 0.05, *WNT3A:* DIV10 **p < 0.01, *WNT5A:* DIV10 **p < 0.01; *RARB:* DIV10: ****p < 0.0001; *BMP4:* DIV10 *p < 0.05, *BMP5:* DIV10 *p < 0.05, one-way ANOVA with Tukey's post-hoc test.
(E) Schematic representation of the expression profile characteristics of the H9 and RUES2 lines with respect to the LGE, MGE, and signaling transcripts analyzed by high-content gene expression analysis.

the H9 line (Fig. 2B). Although we observed different profiles in the expression of WNT receptors FZD1, FZD2, FZD3 and intracellular signaling TCF1 from DIV0 to DIV10, no significant differences were found between the H9 and RUES2 parental lines at any time-points (Figs. 3B and S4C).

LGE-derived GABAergic neurons in the basal ganglia require retinoic acid for their development (Chatzi et al., 2011; Rataj-Baniowska et al., 2015). Therefore, we measured the expression of retinoic acid receptors RAR and RXR in all cell lines during neural induction. Though the mRNA levels of RARα, RARγ, and RXR did not differ between the RUES2 and H9 lines, RARβ was significantly reduced at all time points in RUES2 cells but

progressively increased in H9 cells over time, reaching a peak at DIV10 (Fig. 3C; Fig. S4A). This finding is in line with the propensity of H9 cells to differentiate into MSNs. RUES2 cells also exhibit strong and unique upregulation of BMP4 and BMP5 ligands compared to H9 and KOLF2 cells, possibly contributing to acquisition of an MGE interneuron identity (Mukhopadhyay et al., 2009) (Fig. 3D; Fig. S4A).

Taken together, these data indicate that H9 and RUES2 human cell lines exposed to the same striatal differentiation protocol acquire different cell identities, possibly due to the activation of different signaling pathways in response to the applied morphogens (Fig. 3E).

### 3.4. Aberrant telencephalic and MGE cell fate acquisition in isogenic RUES2 carrying different pathological CAG lengths

Several reports have demonstrated that muHTT affects the differentiation and cell fate specification of the human PSC-derived neurons preferentially affected in HD (i.e., the striatal MSNs originating in the LGE and cortical neurons), (Conforti et al., 2018; HD iPSC Consortium 2012, 2017; Ooi et al., 2019; Ring et al., 2015; Ruzo et al., 2018; Smith-Geater et al., 2020; Xu et al., 2017; Ruzo et al., 2018; Mehta et al., 2018). By exploiting the MGE differentiation bias of RUES2, we tested whether muHTT has a similar detrimental effect on this lineage. For these studies, we employed the isogenic HD RUES2 lines with 50 and 58
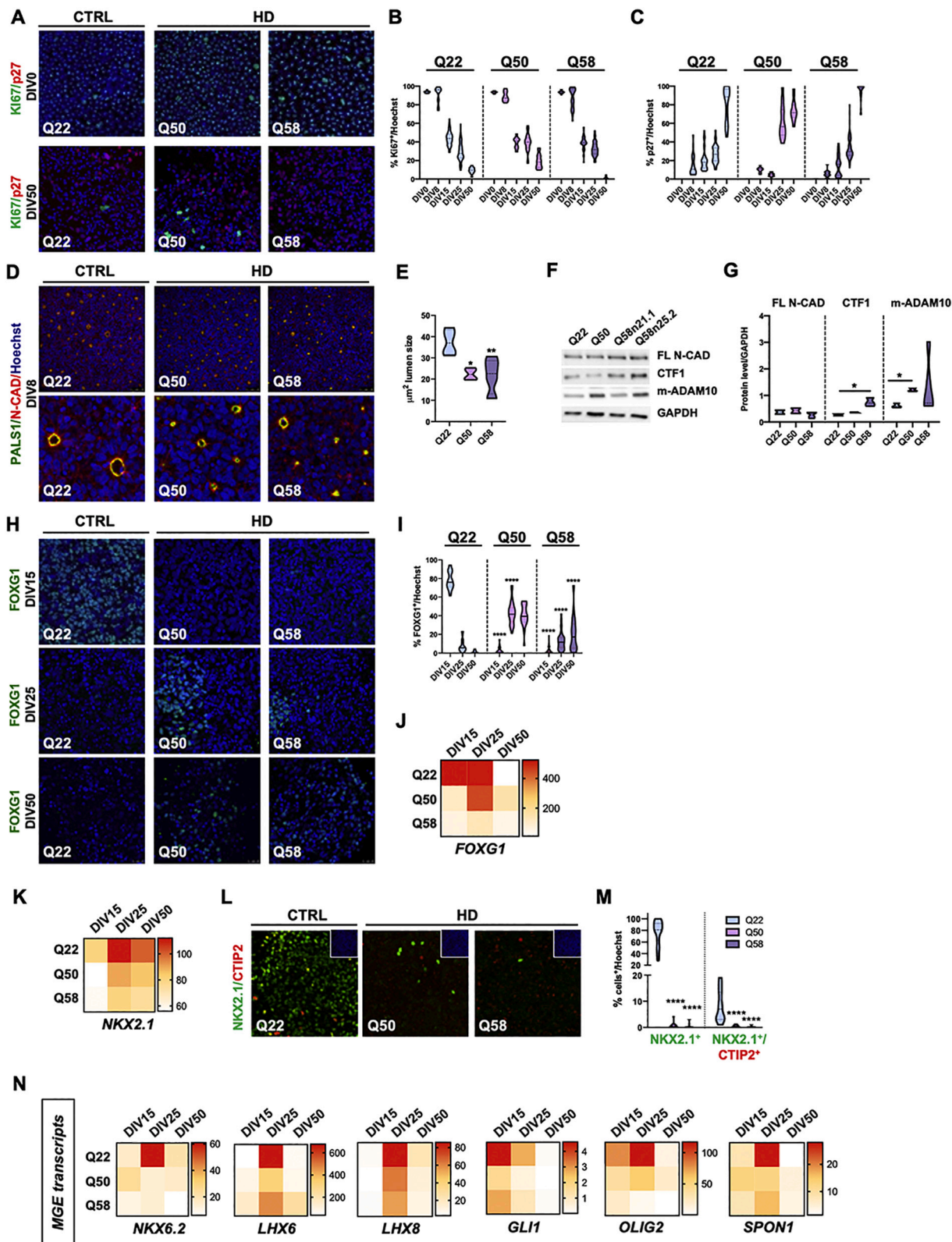


*(caption on next page)*

**Fig. 4.** muHTT interferes with telencephalic and MGE cell fate acquisition in HD RUES2 lines.

(A) Immunocytochemistry of Ki67 and p27 at DIV0 and DIV50 of differentiation of the RUES2 control Q22, HD Q50, and HD Q58 lines. Confocal images, 40×. Scale bar, 50 μm.

(B−C) Counts of Ki67$^+$ and p27$^+$ cells for the Q22, Q50, and Q58 RUES2 lines by CellProfiler software (Version 2.1.1). N = 3 biological replicates. Data are presented as mean ± SEM.

(D) Immunocytochemistry of neural rosette formation by N-CAD$^+$/PALS1$^+$ at DIV8 for the RUES2 control Q22, HD Q50, and HD Q58 lines. Confocal images, 40×. Scale bar, 50 μm (crops of N-CAD$^+$/PALS1$^+$ of the same images).

(E) Counts of rosette lumen sizes by CellProfiler software (Version 2.1.1). N = 3 biological replicates. *p < 0.05, **p < 0.01, one-way ANOVA with Tukey's post-hoc test.

(F) Western blot analysis of the full-length N-cadherin (FL N-CAD), CTF1 fragment, and metalloprotease ADAM10 (mADAM10) at DIV8 of differentiation for the RUES2 control Q22, HD Q50, and HD Q58 lines. The FL N-CAD, CTF1, and mADAM10 protein levels in each line were normalized to GAPDH. N = 3 biological replicates.

(G) Violin plot representing the densitometric analysis performed on western blot results from three biological differentiation experiments. *p < 0.05, one-way ANOVA with Tukey's post-hoc test comparing each group condition.

(H) Immunocytochemistry of FOXG1 at DIV15, DIV25, and DIV50 of differentiation for the RUES2 control Q22, HD Q50, and HD Q58 lines. Confocal images, 40×. Scale bar, 50 μm.

(I) FOXG1$^+$ cell counts for the control Q22, HD Q50, and HD Q58 RUES2 lines by CellProfiler software (Version 2.1.1). N = 3 biological replicates. Q50 vs. Q22 and Q58 vs. Q22, ****p < 0.0001, one-way ANOVA with Tukey's post-hoc test.

(J) Heat map showing the expression levels in high-content qPCR analysis (Biomark) of *FOXG1* at DIV15, DIV25, and DIV50 of differentiation for the RUES2 control Q22, HD Q50, and HD Q58 lines.

(K) Heat map showing the expression levels in high-content qPCR analysis (Biomark) of the MGE marker *NKX2.1* at DIV15, DIV25, and DIV50 for the RUES2 control Q22, HD Q50, and HD Q58 lines. N = 3 biological replicates.

(L) Immunocytochemistry of NKX2.1 and CTIP2 at DIV30 for the RUES2 control Q22, HD Q50, and HD Q58 lines. Confocal images, 40×, zoom = 1.7. Scale bar, 50 μm. Top right, Hoechst inset.

(M) Counts of NKX2.1$^+$ and NKX2.1$^+$/CTIP2$^+$ for the control Q22, HD Q50, and HD Q58 RUES2 lines by CellProfiler software (Version 2.1.1). N = 3 biological replicates. ****p < 0.0001, one-way ANOVA.

(N) Heat map showing the expression levels in high-content qPCR analysis (Biomark) of the MGE markers *NKX6,2, LHX6, LHX8, GLI1, OLIG2,* and *SPON1* at DIV15, DIV25, and DIV50 for the RUES2 control Q22, HD Q50, and HD Q58 lines. N = 3 biological replicates.

---

glutamines and the edited control line carrying Q22 (Fig. S1B) (Ruzo et al., 2018). Normal karyotypes (Fig. S1C) and pluripotency markers were found in all lines (Figs. S5A-M). RUES2 control Q22, HD Q50, and HD Q58 had no significant differences in the number of Ki67$^+$ and p27$^+$ cells from DIV15 to DIV50 (Fig. 4A), suggesting that muHTT does not affect cell cycle exit or neuronal maturation in these cells (Fig. 4B, C), as indicated by phase contrast images (Fig. S5N), immunocytochemistry for TUBB3 and MAP2, and gene transcription analysis for TUBB3, MAP2, and DCX mRNAs (Fig. S5O-R).

Neural rosettes are a stereotypical cell organization of an acquired neuronal phenotype, mimicking the end of neuroectodermal induction. Previous studies have shown that HD-hiPSCs have a smaller lumen size at DIV15 of striatal differentiation (Conforti et al., 2018), whereas HD-RUES2 lines exposed to a default neural induction protocol generate typical neural rosettes with no significant difference (Ruzo et al., 2018). The same HD cellular system exhibits a gradual reduction in lumen size when differentiated into ectodermal neuruloids (Haremaki et al., 2019). Using PALS1 and NCAD markers for rosette lumen identification at DIV8 of *in vitro* differentiation, we confirmed that both the Q50 and Q58 HD RUES2 lines have a significant reduction in rosette lumen area (RUES2 Q22: 37.31 μm$^2$ ± 5.34; RUES2 Q50: 22.36 μm$^2$ ± 2.66; RUES2 Q58: 21.6 μm$^2$ ± 7.61; Fig. 4D, E). Western blot analysis for NCAD confirmed an increase in CTF1 cleavage fragment in HD Q58 cells (two clones; Fig. 4F, G), without changes in NESTIN and NCAD mRNA levels (Fig. S5S, T), which is indicative of an altered capacity of the HD RUES2 cells for cell adhesion and rosette formation.

Next, we monitored the effects of muHTT on ventral telencephalic specification by measuring the acquisition of markers related to a ventral identity. The control Q22 line presented the expected expression profile for FOXG1, with 80 ± 9% of cells immunoreactive at DIV15 and decreasing with time as expected (Fig. 4H, I). In contrast, significantly fewer FOXG1$^+$ cells were found in both HD lines at the same time point (3% ± 2.3 and 4% ± 3 in Q50 and Q58, respectively), with the positive cell peak shifting several weeks for the Q50 line, which remained much lower than in control cells, whereas HD Q58 cultures had a maximum of 11 ± 9% and 15 ± 13% FOXG1$^+$ cells at DIV25 and DIV50, respectively (Fig. 4H, I). Gene expression analysis performed from DIV15 to DIV50 confirmed a reduction in

FOXG1 mRNA in the RUES HD line versus control (Fig. 4 J) and, consequently, a further reduction in the percentage of GSX2$^+$/ASCL1$^+$ ventral progenitors at DIV25 and CTIP2$^+$ at DIV50 (Fig. S6A-D). These data indicate that muHTT also causes a delay in telencephalic fate acquisition in ventralized RUES2 cells.

After observing that the RUES2 line acquired a more ventral fate than the H9 line, we explored the impact of muHTT on MGE transcripts. Although NKX2.1 has been linked to MGE development, it is co-expressed with ISL1 and CTIP2, suggesting that NKX2.1 plays a role in LGE formation (Elias et al., 2008; Onorati et al., 2014). QPCR analyses showed up-regulation of NKX2.1 mRNA in control Q22 cells starting from DIV15 (Fig. 4 K), whereas the HD RUES lines exhibited delayed expression of this transcript, which remained at levels below control for the entire period of differentiation (Fig. 4 K). Similarly, in control Q22 cultures, NKX2.1$^+$ cells increased over time, reaching 80% at DIV50; in both HD lines, we observed a complete loss of this cell population (Fig. 4 L, M). In the RUES2 Q22 line, we found that only 10% of NKX2.1$^+$ cells co-expressed CTIP2, confirming that this hESC line has greater ventral identity under these experimental conditions (Fig. 4 L, M). No NKX2.1$^+$/CTIP2$^+$ double-positive progenitors were found in the RUES2 HD lines. Gene transcription analysis performed on a larger set of MGE markers, including NKX6.2, LHX6, LHX8, GLI1, OLIG2, SPON1, and GSX1 corroborated these results, suggesting a negative effect of muHTT in MGE fate determination (Fig. 4 N).

### 3.5. Isogenic RUES2 lines recapitulate HD phenotypes

To further assess the ability of RUES2 lines to recapitulate known HD phenotypes, we performed a high-content qPCR analysis for transcripts involved in neurodevelopment that are known to be affected by muHTT. REST/NRSF is part of a repressor complex that acts as a central negative regulator of neuronal gene transcription (Ballas and Mandel, 2005; Chong et al., 1995; Schoenherr and Anderson, 1995). Excessive nuclear access of REST/NRSF in HD cells and tissues has been found to reduce the transcription of several neuronal genes, including BDNF (Zuccato et al., 2011; Zuccato and Cattaneo, 2009). Consistently, our qPCR of *BDNF* in neurons derived from RUES2 HD and control lines showed significant downregulation in both HD Q50 and HD Q58 cells
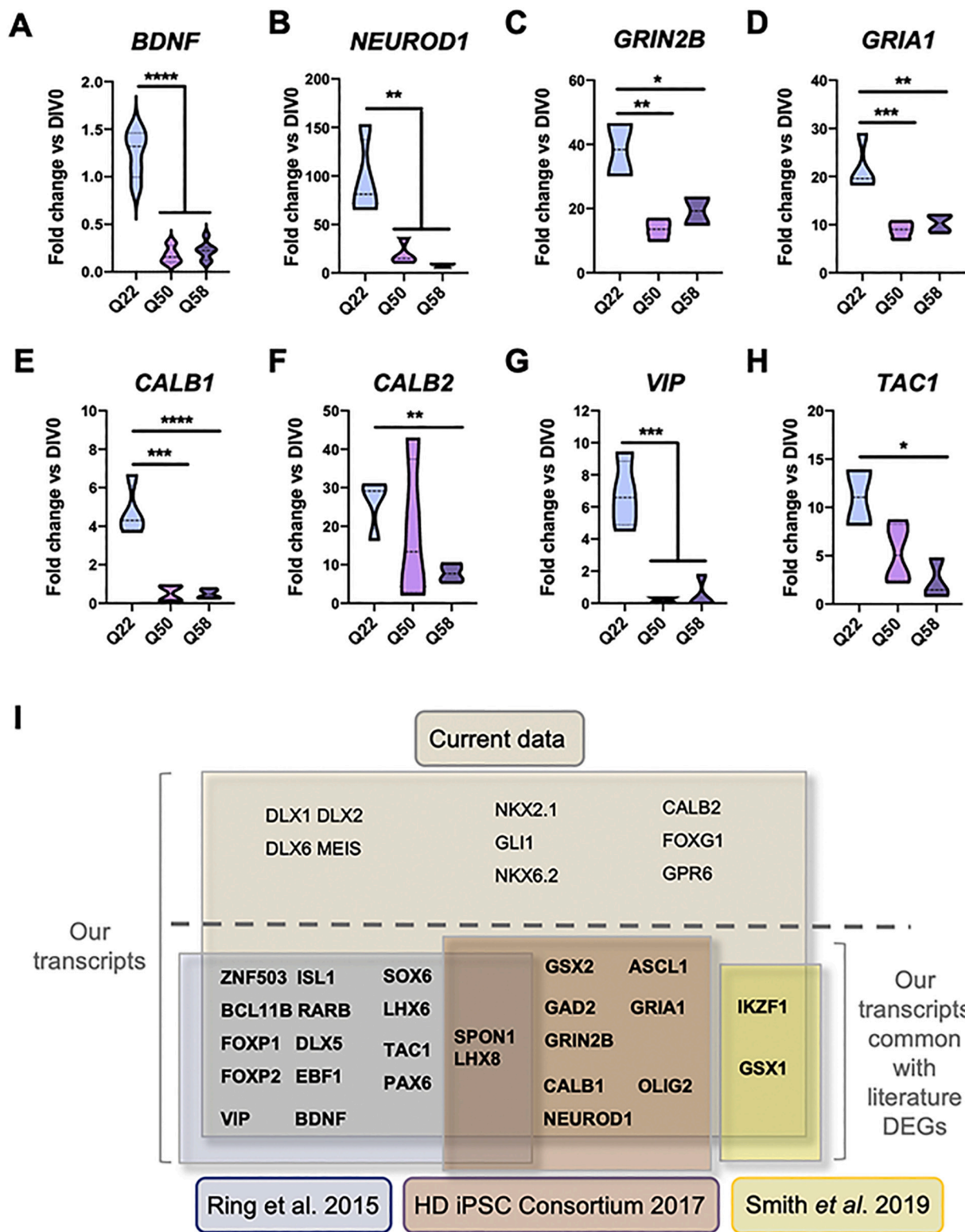
**Fig. 5.** Isogenic RUES2 lines recapitulate the HD phenotype.

(A-H) QPCR of transcripts known to be altered in mouse and cellular HD models at DIV30 for the RUES2 control Q22, HD Q50, and HD Q58 lines (total mRNA level normalized to 18 s housekeeping transcript). N = 3 biological replicates. *BDNF:* Q50 vs. Q22, ****$p < 0.0001$; Q58 vs. Q22, ****$p < 0.0001$. *NEUROD1:* Q50 vs. Q22, **$p < 0.01$; Q58 vs. Q22, **$p < 0.01$. *GRIN2B:* Q50 vs. Q22, **$p < 0.01$; Q58 vs. Q22, *$p < 0.05$. *GRIA1:* Q50 vs. Q22, ***$p < 0.001$. Q58 vs. Q22, **$p < 0.01$. *CALB1:* Q50 vs. Q22, ***$p < 0.001$; Q58 vs. Q22, ****$p < 0.0001$. *CALB2:* Q58 vs. Q22, **$p < 0.001$. *VIP:* Q50 vs. Q22, ***$p < 0.001$; Q58 vs. Q22, ***$p < 0.001$. *TAC1:* Q58 vs. Q22, *$p < 0.05$, one-way ANOVA with Tukey's post-hoc test.

(I) Schematic diagram showing our differential transcripts (current data). Transcripts we found to be in common with differentially expressed genes (DEGs) from studies comparing control and HD lines are showed in bold.

compared to control Q22 cells (Fig. 5A). BDNF levels were also reduced in HD-iPS lines upon exposure to the same differentiation protocol and to cortical differentiation (Shi et al., 2012) (Fig. S7A).

NEUROD1 is controlled by REST/NRSF and has been implicated in aberrant developmental and adult neurogenesis in R6/2 mice and HD iPSC-derived neurons (HD iPSC Consortium, 2017). Both HD RUES2 Q50 and Q58 lines exhibit a significant reduction in NEUROD1 mRNA at DIV30 of differentiation (Fig. 5B). Developmental genes *GRIN2B* and *GRIA1* (Endele et al., 2010) were also reduced in Q50 and Q58 RUES2-derived neurons (Fig. 5C, D), as were CALB1 and CALB2, two calcium binding proteins expressed early in development (Al-Jaberi et al., 2015) (Fig. 5E, F). Other transcripts known to be altered in HD and linked to central nervous system development were also investigated. The neuropeptide vasoactive intestinal polypeptide (VIP), which is highly expressed in the central nervous system, is reduced in R6/2 mice (Fahrenkrug et al., 2007) and significantly reduced in our HD RUES2-derived neurons (Fig. 5G). Similarly, the striatal transcript TAC1, the expression of which is sustained by BDNF (Xie et al., 2010), was reduced in the RUES2 HD lines (Fig. 5H).

To address whether the differentially expressed genes (DEGs) were shared by other HD *in vitro* systems, we interrogated available RNAseq datasets from Ring et al. (2015), the HD iPSC Consortium (2017), and Smith-Geater et al. (2020) and found that several LGE genes were common to all studies, including certain MGE transcripts. In particular, SPON1 and LHX8 were differentially expressed in our analysis and in the Ring et al. (2015) and the HD iPSC Consortium datasets (Fig. 5I). This comparison revealed that common genes were similarly affected in diverse HD hPSC lines exposed to different striatal differentiation protocols, showing the ability of the RUES2 model system to recapitulate mouse and human HD phenotypes.

### 3.6. De novo motif discovery analysis identifies SP1 as a potential common regulator in HD neural differentiation

As muHTT-driven dysregulation appears to affect multiple neuronal lineages (Fig. 6A) (Conforti et al., 2018; HD iPSC Consortium, 2012, 2017; Ooi et al., 2019; Ring et al., 2015; Ruzo et al., 2018; Smith-Geater et al., 2020; Xu et al., 2017), we attempted to identify common regulators of muHTT toxicity by comparing HD-related DEGs using transcriptional data from the above-mentioned studies (Ring et al., 2015; HD iPSC Consortium, 2017; Smith-Geater et al., 2020).

First, we looked at whether the upstream regions of the DEGs characterizing HD hPSCs undergoing LGE, MGE, and cortical differentiation *in vitro* were enriched for specific motifs of TFs that may be affected by muHTT. We employed two different *de novo* motif discovery tools to identify the common TF binding motifs: Multiple EM For Motif Elicitation (MEME) and Oligo Analysis (RSAT) (Bailey et al., 2009; Nguyen et al., 2018) (Fig. S8A). Next, we compared the sequence of the identified motifs with a series of databases of known TF binding sites using STAMP (Mahony and Benos, 2007), a toolkit for comparing DNA motifs that relies on the JASPAR (Bryne et al., 2008) and TRANSFAC (Matys et al., 2003) databases (Fig. S78A; SI1, SI2, SI3). Only the TFs represented by both motif discovery tools were considered further (Fig. S8B). As shown in Fig. 6B, the most significant LGE-specific TFs included SP1 and LHX3, followed by CREB1, KLF4, and SPI1.

The TFs SP1 and RREB1 had the highest scores in the same analysis for MGE and cortex (Fig. 6B). Specifically, the RREB1 protein recognizes and binds RAS-responsive elements (RREs) in gene regulatory regions, and potentiates the transcriptional activity of NEUROD1 (Ray et al., 2014). Lastly, the analysis identified RAP1 as MGE-specific; the protein regulates several signaling pathways affecting cell proliferation and adhesion, as well as the neuronal response to dopamine (Zhang et al., 2018) (Fig. 6B).

Notably, using both databases, STAMP analysis revealed SP1, SPI1, and KLF4 as common putative TFs potentially regulating the DEGs identified in the LGE, MGE, and cortex (Fig. 6B). In particular, SP1

reached the highest significant score in the analysis of signaling pathways (Fig. S9A, B; Table S1 and SI4).

Next, in an attempt to identify gene networks potentially affected by the aberrant SP1-HTT interaction, the DEGs identified by the in silico analyses and the predicted TFs were submitted to the protein interaction database Pathway Commons (http://www.pathwaycommons.org). Using Cytoscape for graphical representation, we found that the DEGs and TFs were connected in an intricate network by both direct binding and transcriptional regulation. The analysis highlighted the direct interaction between SP1 and HTT in all areas analyzed (Fig. 6C-E; Table S1).

## 4. Discussion

As hPSCs can efficiently differentiate into functional neurons through processes that recapitulate in vivo development, the use of small molecules and morphogens to drive hPSCs towards the acquisition of a coordinated antero/posterior or ventral/dorsal fate allows the generation of specific neurons essential for modeling neurodegenerative diseases. Patterning during *in vitro* neural differentiation is regulated by retinoic acid and the FGF and BMP families, with a prominent role of SHH signaling in ventralization (Germain et al., 2013; Li et al., 2009; Liu et al., 2013; Maroof et al., 2013; Nicholas et al., 2013) and WNT signaling in both dorsalization and caudalization (Elkabetz et al., 2008; Kirkeby et al., 2012; Li et al., 2009).

Here, the RUES2 line presented a different expression profile in SHH, WNT, RA, and BMP signaling compared to the H9 and KOLF2 lines. The hyperactivation of SHH signaling, together with the upregulation of BMP4 and BMP5 ligands and the complete loss of RARβ expression, suggests that the RUES2 lines are more ventralized than the H9 and KOLF2 lines. In line with this data, Strano et al. (2020) demonstrated that the patterning variation of different hESCs is associated with differences in SHH and WNT signaling dynamics and that early differences in regional gene expression were predictive of late-stage fate acquisition (Strano et al., 2020). Earlier studies have reported marked differences in differentiation propensities among different human embryonic stem cell lines (Osafune et al., 2008), further pointing to the relevance of deriving multiple lines for lineage-specific differentiation (Wu et al., 2007; Hu et al., 2010).

In our experiments, control RUES2 cells exposed to the striatal differentiation protocol exhibited impairment in the acquisition of an MSN striatal fate, compared to H9 and KOLF2 cells. In particular, the RUES2 line failed to express the MSN regulator DARPP32 while preferentially upregulating MGE-enriched transcripts. This is suggested by simultaneous expression of NKX2.1, LHX6 and LHX8 and other transcripts qualifying MGE GABAergic interneurons, such as GAD67. Leveraging the RUES2 MGE-biased differentiation potential, we show that muHTT interferes with MGE cell fate acquisition, as already demonstrated for LGE (Ring et al., 2015; HD iPSC Consortium, 2017; Xu et al., 2017; Conforti et al., 2018; Smith-Geater et al., 2020) and cortical neurons (Conforti et al., 2018; Ruzo et al., 2018; Mehta et al., 2018; Ooi et al., 2019). MGE GABAergic interneurons are essential for neural circuit function (Tremblay et al., 2016), and their loss or dysfunction is implicated in several human psychiatric disorders, including autism, schizophrenia, and epilepsy (Marín, 2012). Our data suggest that muHTT negatively influences MGE-interneuron differentiation, opening up the possibility that MGE derivatives may make a greater contribution to HD than previously thought.

Our observations also point to the presence of common muHTT-related interactors whose activity is shared by the different lineage-specific signaling pathways that interact during neurogenesis. SP1 is a TF found in all mammalian cell types that plays a central role in several human diseases, including HD. Several lines of evidence link SP1 to HTT, demonstrating their direct interaction in transgenic HD mouse brain, striatal HD cells, and human HD brain (Dunah et al., 2002). In particular, the N-terminal portion of muHTT interacts more strongly
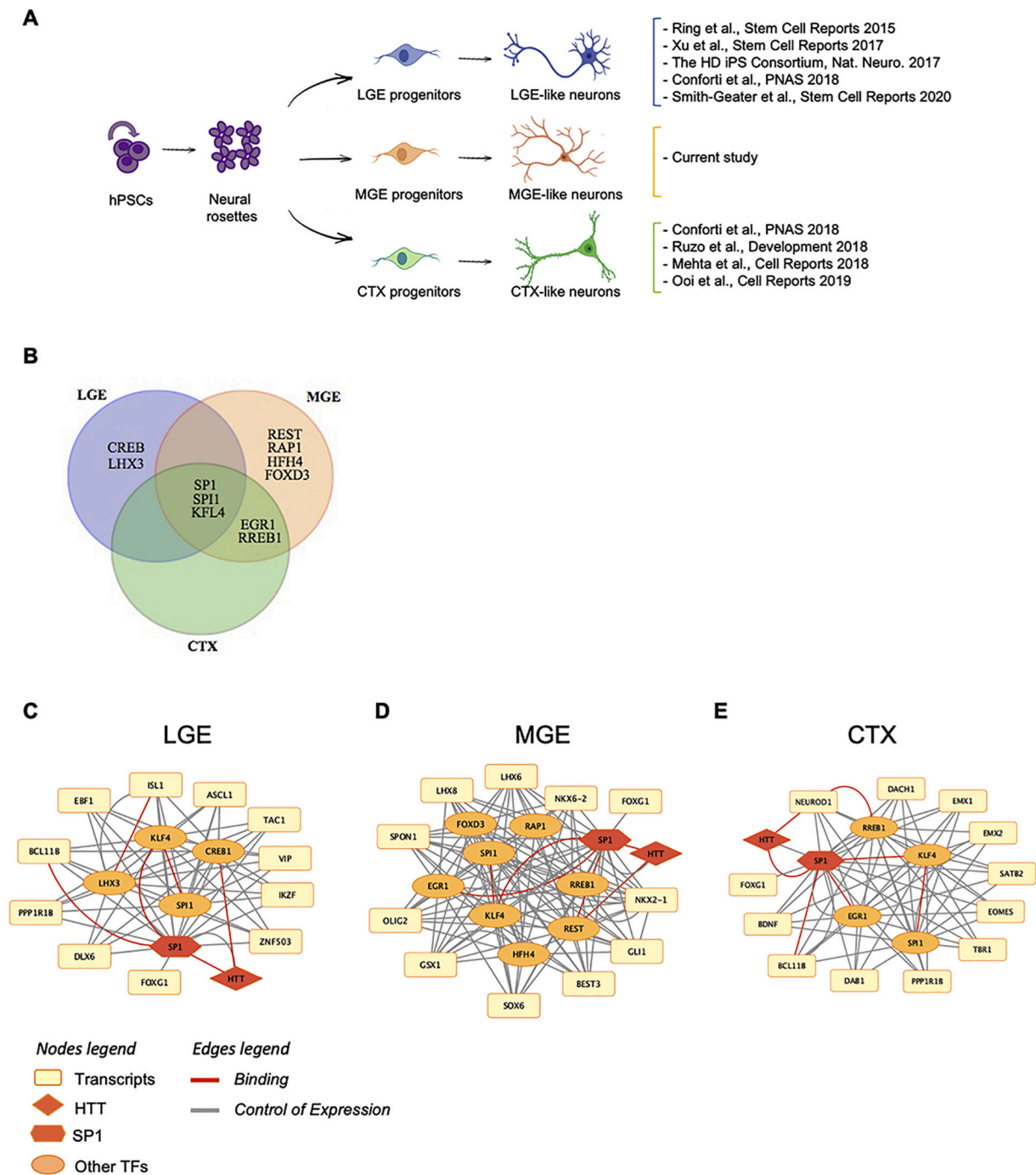
**Fig. 6.** *De novo* motif discovery analysis identified Sp1 as a potential common regulator in neural differentiation.
(A) Schematic representation of the papers reported in the literature differentiating iPSCs and hESCs in different neuronal lineages (LGE, MGE, and CTX). The presence of muHTT may impact the ability to differentiate into multiple neuronal types, and this constitutes the rationale for the *de novo* motif discovery analysis.
(B) Venn diagram of the unique and shared LGE, MGE, and cortical predictive transcription factors (TFs) identified by *de novo* motif discovery analysis.
(C-E) Cytoscape networks for LGE, MGE, and CTX, representing the predicted interactions obtained by *de novo* motif discovery analysis enriched with the molecular and genetic interactions and gene regulation networks generated by Pathway Commons (https://apps.pathwaycommons.org/interactions). The TFs submitted to the *de novo* motif discovery analysis specific for each brain area are depicted in yellow, and in orange are the common regulators. Sp1 is highlighted in red because it was the most represented and relative to HTT. The binding interactions are reported in red and controlled expression in grey.

with SP1 than non-pathologic HTT, resulting in cytoplasmic retention of the TF in HD and reduced nuclear SP1, leading to reduced occupancy of the promoters of SP1-responsive genes, such as DRD2 (Chen-Plotkin et al., 2006; Dunah et al., 2002), REST (Ravache et al., 2010), and HTT itself (Wang et al., 2012).

A second transcriptional regulator that potentially plays a role in modulating HD-related gene signaling is SPI1 or PU.1, which encodes a

protein involved in the regulation of several cellular processes, including cell differentiation, cell growth, apoptosis, DNA damage response, and chromatin remodeling. In particular, SPI1 controls a variety of myeloid genes that are important for brain inflammation and neuroimmunology. Crotti et al. (2014) found that the presence of muHTT in microglia promotes cell-autonomous pro-inflammatory transcriptional activation by increasing the expression and activation of the

transcription of SPI1.

A third TF that emerged from the analysis is Krueppel-like factor 4 (KLF4), an important regulator of early embryonic development, as it contributes to maintaining self-renewal and preventing hESC differentiation. KLF4 is also a powerful negative regulator of the expression of multiple smooth muscle cell (SMC) genes through its activation by SP1 (Deaton et al., 2009; Zhang et al., 2012).

The data from our *in-silico* analysis showed that, in the LGE context, LHX3 is a potential common mediator of signaling. Interestingly, SP1 controls and interacts with LHX3 (Yaden et al., 2006). This gene encodes a protein that, by acting as a transcriptional activator, regulates pituitary development and motor neuron specification through the interaction with ISL1 (Erb et al., 2017; Lee et al., 2012a, 2012b, 2012c; Seo et al., 2015; Thaler et al., 2002).

We also found LGE-specific alterations in the transcription of the cAMP response element-binding protein (CREB), which has already been associated with HD as a possible contributor to neuronal cell death and dysfunction (Wyttenbach, 2001). The CREB pathway involves the activity of TAFII130, a coactivator found in polyQ aggregates that, when overexpressed, can rescue HD reduction in CREB-dependent transcription (Shimohata et al., 2000). Dunah et al. (2002) reported that the expanded polyQ of muHTT impairs the soluble association of TAFII130 with SP1 through direct interference in SP1 binding to the DNA.

The discovery of SP1 as the most common predictive TF of the *de novo* motif discovery analysis is supported by several lines of evidence in the literature, highlighting SP1 as a possible common player in the abnormal neuronal gene transcription profiles identified thus far in all studies that have employed HD-hPSC lines, suggesting SP1 as a potential detrimental co-partner of muHTT during neuronal differentiation.

Overall, this work highlights the different propensity of various hPSCs to acquire a specific cell fate when exposed to the striatal differentiation protocol. The RUES2 line exhibited MGE-biased neuronal differentiation that is compromised by muHTT. *In silico de novo* motif discovery analysis identified the TF SP1 as a putative common regulator of the DEGs specific for different cerebral areas. Therefore, we hypothesize that, in the presence of muHTT, the broad transcriptional activity of SP1 is compromised and this interferes with cell fate commitment.

## Acknowledgments

## Author contributions

P.C., D.B., and E.C. designed the research; P.C., S.B., and A.L. performed the *in vitro* studies; S.B. performed the bioinformatics *in silico* analysis with contributions from R.R.; I.C. helped with WB experiments.; M.G. contributed to cell line maintenance and characterization; C.C. and V.D.B. contributed to the analyses of the high-content qPCR data; R.I. quantified the lumen area; and P.C., D.B., S.B., and E.C. wrote the paper.

## Credit Authors Statement

P.C., D.B., and E.C. designed the research.

P.C., S.B., and A.L. performed the *in vitro* studies.

S.B. performed the bioinformatics in silico analysis with contributions from R.R.

I.C. helped with WB experiments.

M.G. contributed to cell line maintenance and characterization.

C.C. and V.D.B. contributed to the analyses of the high-content qPCR data.

R.I. quantified the lumen area.

P.C., D.B., S.B., and E.C. wrote the paper.

## Declaration of Competing Interest

The authors declare no competing interests.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nbd.2020.105140.

## References

Al-Jaberi, N., Lindsay, S., Sarma, S., Bayatti, N., Clowry, G.J., 2015. The early fetal development of human neocortical GABAergic interneurons. Cereb. Cortex. https://doi.org/10.1093/cercor/bht254.

Ballas, N., & Mandel, G. (2005). The many faces of REST oversee epigenetic programming of neuronal genes. In Current Opinion in Neurobiology. https://doi.org/10.1016/j.conb.2005.08.015.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., Noble, W.S., 2009. MEME suite: tools for motif discovery and searching. Nucleic Acids Res. https://doi.org/10.1093/nar/gkp335.

Barnat, M., Capizzi, M., Aparicio, E., Boluda, S., Wennagel, D., Kacher, R., Kassem, R., Lenoir, S., Agasse, F., Bra, B.Y., Liu, J.P., Ighil, J., Tessier, A., Zeitli, S.O., Duyckaerts, C., Dommergues, M., Durr, A., Humbert, S., 2020. Huntington's disease alters human neurodevelopment. Science. https://doi.org/10.1126/science.aax3338.

Besusso, D., Schellino, R., Boido, M., Belloli, S., Parolisi, R., Conforti, P., Faedo, A., Cernigoj, M., Campus, I., Laporta, A., Bocchi, V.D., Murtaj, V., Parmar, M., Spaiardi, P., Talpo, F., Maniezzi, C., Toselli, M.G., Biella, G., Moresco, R.M., ... Cattaneo, E., 2020. Stem cell-derived human striatal progenitors innervate striatal targets and alleviate sensorimotor deficit in a rat model of Huntington disease. Stem Cell Rep. https://doi.org/10.1016/j.stemcr.2020.03.018.

Bryne, J.C., Valen, E., Tang, M.H.E., Marstrand, T., Winther, O., Da piedade, I., Krogh, A., Lenhard, B., Sandelin, A., 2008. JASPAR, the open access database of transcription factor-binding profiles: New content and tools in the 2008 update. Nucleic Acids Res. https://doi.org/10.1093/nar/gkm955.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., Studer, L., 2009. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nature Biotechnology. https://doi.org/10.1038/nbt.1529.

Chatzi, C., Brade, T., Duester, G., 2011. Retinoic acid functions as a key gabaergic differentiation signal in the basal ganglia. PLoS Biol. https://doi.org/10.1371/journal.pbio.1000609.

Chen-Plotkin, A.S., Sadri-Vakili, G., Yohrling, G.J., Braveman, M.W., Benn, C.L., Glajch, K.E., DiRocco, D.P., Farrell, L.A., Krainc, D., Gines, S., MacDonald, M.E., Cha, J.H.J., 2006. Decreased association of the transcription factor Sp1 with genes downregulated in Huntington's disease. Neurobiol. Dis. https://doi.org/10.1016/j.nbd.2005.11.001.

Chi, L., Fan, B., Zhang, K., Du, Y., Liu, Z., Fang, Y., Chen, Z., Ren, X., Xu, X., Jiang, C., Li, S., Ma, L., Gao, L., Liu, L., Zhang, X., 2016. Targeted differentiation of regional ventral Neuroprogenitors and related neuronal subtypes from human pluripotent stem cells. Stem Cell Rep. https://doi.org/10.1016/j.stemcr.2016.09.003.

Chong, J.A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J.J., Zheng, Y., Boutros, M.C., Altshuller, Y.M., Frohman, M.A., Kraner, S.D., Mandel, G., 1995. REST: A mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell. https://doi.org/10.1016/0092-8674(95)90298-8.

Conforti, P., Besusso, D., Bocchi, V.D., Faedo, A., Cesana, E., Rossetti, G., Ranzani, V., Svendsen, C.N., Thompson, L.M., Toselli, M., Biella, G., Pagani, M., Cattaneo, E., 2018. Faulty neuronal determination and cell polarization are reverted by modulating HD early phenotypes. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1715865115.

Crotti, A., Benner, C., Kerman, B.E., Gosselin, D., Lagier-Tourenne, C., Zuccato, C., Cattaneo, E., Gage, F.H., Cleveland, D.W., Glass, C.K., 2014. Mutant Huntingtin promotes autonomous microglia activation via myeloid lineage-determining factors. Nat. Neurosci. https://doi.org/10.1038/nn.3668.

Deaton, R.A., Gan, Q., Owens, G.K., 2009. Sp1-dependent activation of KLF4 is required for PDGF-BB-induced phenotypic modulation of smooth muscle. Am. J. Physiol. Heart Circ. Physiol. https://doi.org/10.1152/ajpheart.01230.2008.

Delli Carri, A., Onorati, M., Castiglioni, V., Faedo, A., Camnasio, S., Toselli, M., Biella, G., Cattaneo, E., 2013. Human pluripotent stem cell differentiation into authentic striatal projection neurons. Stem Cell Rev. Rep. https://doi.org/10.1007/s12015-013-9441-8.

Dunah, A.W., Jeong, H., Griffin, A., Kim, Y.M., Standaert, D.G., Hersch, S.M., Mouradian, M.M., Young, A.B., Tanese, N., Krainc, D., 2002. Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. Science. https://doi.org/10.1126/science.1072613.

Elias, L.A.B., Potter, G.B., Kriegstein, A.R., 2008. A time and a place for Nkx2-1 in interneuron specification and migration. Neuron. https://doi.org/10.1016/j.neuron.2008.08.017.

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N.D., Tabar, V., Studer, L., 2008. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev. https://doi.org/10.1101/gad.1616208.

Endele, S., Rosenberger, G., Geider, K., Popp, B., Tamer, C., Stefanova, I., Milh, M., Kortüm, F., Fritsch, A., Pientka, F.K., Hellenbroich, Y., Kalscheuer, V.M., Kohlhase, J., Moog, U., Rappold, G., Rauch, A., Ropers, H.H., Von Spiczak, S., Tönnies, H., ... Kutsche, K., 2010. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. Nat. Genet. https://doi.org/10.1038/ng.677.

Erb, M., Lee, B., Seo, S.Y., Lee, J.W., Lee, S., Lee, S.K., 2017. The isl1-lhx3 complex promotes motor neuron specification by activating transcriptional pathways that enhance its own expression and formation. ENeuro. https://doi.org/10.1523/ENEURO.0349-16.2017.

Fahrenkrug, J., Popovic, N., Georg, B., Brundin, P., Hannibal, J., 2007. Decreased VIP and VPAC2 receptor expression in the biological clock of the R6/2 Huntington's disease mouse. J. Mol. Neurosci. https://doi.org/10.1385/jmn/31:02:139.

Farshim, P.P., Bates, G.P., 2018. Mouse models of Huntington's disease. Methods Mol. Biol. https://doi.org/10.1007/978-1-4939-7825-0_6.

Fasano, C.A., Chambers, S.M., Lee, G., Tomishima, M.J., Studer, L., 2010. Efficient derivation of functional floor plate tissue from human embryonic stem cells. Cell Stem Cell. https://doi.org/10.1016/j.stem.2010.03.001.

Germain, N.D., Banda, E.C., Becker, S., Naegele, J.R., Grabel, L.B., 2013. Derivation and isolation of NKX2.1-positive basal forebrain progenitors from human embryonic stem cells. Stem Cells Dev. https://doi.org/10.1089/scd.2012.0264.

Godin, J.D., Colombo, K., Molina-Calavita, M., Keryer, G., Zala, D., Charrin, B.E.C., Dietrich, P., Volvert, M.L., Guillemot, F., Dragatsis, I., Bellaiche, Y., Saudou, F., Nguyen, L., Humbert, S., 2010. Huntingtin is required for mitotic spindle orientation and mammalian neurogenesis. Neuron. https://doi.org/10.1016/j.neuron.2010.06.027.

Gusella, J.F., MacDonald, M.E., 2006. Huntington's disease: seeing the pathogenic process through a genetic lens. Trends Biochem. Sci. https://doi.org/10.1016/j.tibs.2006.06.009.

Haremaki, T., Metzger, J.J., Rito, T., Ozair, M.Z., Etoc, F., Brivanlou, A.H., 2019. Self-organizing neuruloids model developmental aspects of Huntington's disease in the ectodermal compartment. Nat. Biotechnol. https://doi.org/10.1038/s41587-019-0237-5.

Hasenpusch-Theil, K., Watson, J.A., Theil, T., 2017. Direct Interactions Between Gli3, Wnt8b and Fgfs Underlie Patterning of the Dorsal Telencephalon. Cereb. Cortex (New York, N.Y. : 1991). https://doi.org/10.1093/cercor/bhv291.

HD iPSC Consortium, Mattis, V.B., Svendsen, S.P., Ebert, A., Svendsen, C.N., King, A.R., Casale, M., Winokur, S.T., Batugedara, G., Vawter, M., Donovan, P.J., Lock, L.F., Thompson, L.M., Zhu, Y., Fossale, E., Atwal, R.S., Gillis, T., Mysore, J., Li, J.H., ... Arjomand, J., 2012. Induced pluripotent stem cells from patients with huntington's disease show CAG repeat expansion associated phenotypes. Cell Stem Cell. https://doi.org/10.1016/j.stem.2012.04.027.

HD iPSC Consortium, Lim, R.G., Salazar, L.L., Wilton, D.K., King, A.R., Stocksdale, J.T., Sharifabad, D., Lau, A.L., Stevens, B., Reidling, J.C., Winokur, S.T., Casale, M.S., Thompson, L.M., Pardo, M., Díaz-Barriga, A.G.G., Straccia, M., Sanders, P., Alberch, J., Canals, J.M., ... Svendsen, C.N., 2017. Developmental alterations in Huntington's disease neural cells and pharmacological rescue in cells and mice. Nat. Neurosci. https://doi.org/10.1038/nn.4532.

Hu, B.-Y., Weick, J.P., Yu, J., Ma, L.-X., Zhang, X.-Q., Thomson, J.A., Zhang, S.-C., 2010. Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.0910012107.

Kirkeby, A., Grealish, S., Wolf, D.A., Nelander, J., Wood, J., Lundblad, M., Lindvall, O., Parmar, M., 2012. Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. Cell Rep. https://doi.org/10.1016/j.celrep.2012.04.009.

Lee, S., Cuvillier, J.M., Lee, B., Shen, R., Lee, J.W., Lee, S.K., 2012a. Fusion protein Isl1-Lhx3 specifies motor neuron fate by inducing motor neuron genes and concomitantly suppressing the interneuron programs. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1114515109.

Lee, J.K., Mathews, K., Schlaggar, B., Perlmutter, J., Paulsen, J.S., Epping, E., Burmeister, L., Nopoulos, P., 2012b. Measures of growth in children at risk for Huntington disease. Neurology. https://doi.org/10.1212/WNL.0b013e3182648b65.

Lee, J.M., Ramos, E.M., Lee, J.H., Gillis, T., Mysore, J.S., Hayden, M.R., Warby, S.C., Morrison, P., Nance, M., Ross, C.A., Margolis, R.L., Squitieri, F., Orobello, S., Di Donato, S., Gomez-Tortosa, E., Ayuso, C., Suchowersky, O., Trent, R.J.A., McCusker, E., ... Gusella, J.F., 2012c. CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. Neurology. https://doi.org/10.1212/WNL.0b013e318249f683.

Li, X.J., Zhang, X., Johnson, M.A., Wang, Z.B., LaVaute, T., Zhang, S.C., 2009. Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. Development. https://doi.org/10.1242/dev.036624.

Lin, C.H., Tallaksen-Greene, S., Chien, W.M., Cearley, J.A., Jackson, W.S., Crouse, A.B., Ren, S., Li, X.J., Albin, R.L., Detloff, P.J., 2001. Neurological abnormalities in a knock-in mouse model of Huntington's disease. Hum. Mol. Genet. https://doi.org/10.1093/hmg/10.2.137.

Liu, Y., Liu, H., Sauvey, C., Yao, L., Zarnowska, E.D., Zhang, S.C., 2013. Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. Nat. Protoc. https://doi.org/10.1038/nprot.2013.106.

Lopes, C., Aubert, S., Bourgois-Rocha, F., Barnat, M., Rego, A.C., Déglon, N., Perrier, A.L., Humbert, S., 2016. Dominant-negative effects of adult-onset huntingtin mutations alter the division of human embryonic stem cells-derived neural cells. PLoS One. https://doi.org/10.1371/journal.pone.0148680.

Lucas, J.J., Ortega, Z., 2011. Animal Models for Huntington ' S Disease. (Animal Models for Neurodegenerative Disease).

Ma, L., Wang, Y., Hui, Y., Du, Y., Chen, Z., Feng, H., Zhang, S., Li, N., Song, J., Fang, Y., Xu, X., Shi, L., Zhang, B., Cheng, J., Zhou, S., Liu, L., Zhang, X., 2019. WNT/NOTCH pathway is essential for the maintenance and expansion of human MGE progenitors. Stem Cell Rep. 12 (5), 934–949. https://doi.org/10.1016/j.stemcr.2019.04.007.

Mahony, S., Benos, P.V., 2007. STAMP: A web tool for exploring DNA-binding motif similarities. Nucleic Acids Res. https://doi.org/10.1093/nar/gkm272.

Marín, O., 2012. Interneuron dysfunction in psychiatric disorders. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn3155.

Maroof, A.M., Keros, S., Tyson, J.A., Ying, S.W., Ganat, Y.M., Merkle, F.T., Liu, B., Goulburn, A., Stanley, E.G., Elefanty, A.G., Widmer, H.R., Eggan, K., Goldstein, P.A., Anderson, S.A., Studer, L., 2013. Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell. https://doi.org/10.1016/j.stem.2013.04.008.

Matys, V., Fricke, E., Geffers, R., Gößling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., Kloos, D.U., Land, S., Lewicki-Potapov, B., Michael, H., Münch, R., Reuter, I., Rotert, S., Saxel, H., Scheer, M., ... Wingender, E., 2003. TRANSFAC®: transcriptional regulation, from patterns to profiles. Nucleic Acids Res. https://doi.org/10.1093/nar/gkg108.

Mehta, S.R., Tom, C.M., Wang, Y., Bresee, C., Rushton, D., Mathkar, P.P., Tang, J., Mattis, V.B., 2018. Human Huntington's disease iPSC-derived cortical neurons display altered Transcriptomics, morphology, and maturation. Cell Rep. https://doi.org/10.1016/j.celrep.2018.09.076.

Molero, A.E., Gokhan, S., Gonzalez, S., Feig, J.L., Alexandre, L.C., Mehler, M.F., 2009. Impairment of developmental stem cell-mediated striatal neurogenesis and pluripotency genes in a knock-in model of Huntington's disease. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.0912171106.

Molero, A.E., Arteaga-Bracho, E.E., Chen, C.H., Gulinello, M., Winchester, M.L., Pichamoorthy, N., Gokhan, S., Khodakhah, K., Mehler, M.F., 2016. Selective expression of mutant huntingtin during development recapitulates characteristic features of Huntington's disease. Proc. Natl. Acad. Sci. U. S. A. https://doi.org/10.1073/pnas.1603871113.

Mukhopadhyay, A., McGuire, T., Peng, C.Y., Kessler, J.A., 2009. Differential effects of BMP signaling on parvalbumin and somatostatin interneuron differentiation. Development. https://doi.org/10.1242/dev.034439.

Nanetti, L., Contarino, V.E., Castaldo, A., Sarro, L., Bachoud-Levi, A.C., Giavazzi, M., Frittoli, S., Ciammola, A., Rizzo, E., Gellera, C., Bruzzone, M.G., Taroni, F., Grisoli, M., Mariotti, C., 2018. Cortical thickness, stance control, and arithmetic skill: An exploratory study in premanifest Huntington disease. Parkinsonism Relat. Disord. https://doi.org/10.1016/j.parkreldis.2018.02.033.

Nguyen, N.T.T., Contreras-Moreira, B., Castro-Mondragon, J.A., Santana-Garcia, W., Ossio, R., Robles-Espinoza, C.D., Bahin, M., Collombet, S., Vincens, P., Thieffry, D., Van Helden, J., Medina-Rivera, A., Thomas-Chollier, M., 2018. RSAT 2018: regulatory sequence analysis tools 20th anniversary. Nucleic Acids Res. https://doi.org/10.1093/nar/gky317.

Nicholas, C.R., Chen, J., Tang, Y., Southwell, D.G., Chalmers, N., Vogt, D., Arnold, C.M., Chen, Y.J.J., Stanley, E.G., Elefanty, A.G., Sasai, Y., Alvarez-Buylla, A., Rubenstein, J.L.R., Kriegstein, A.R., 2013. Functional maturation of hPSC-derived forebrain interneurons requires an extended timeline and mimics human neural development. Cell Stem Cell. https://doi.org/10.1016/j.stem.2013.04.005.

Nopoulos, P.C., Aylward, E.H., Ross, C.A., Johnson, H.J., Magnotta, V.A., Juhl, A.R., Pierson, R.K., Mills, J., Langbehn, D.R., Paulsen, J.S., 2011. Cerebral cortex structure in prodromal Huntington disease. Neurobiology of Disease. https://doi.org/10.1016/j.nbd.2010.07.014.

Onorati, M., Castiglioni, V., Biasci, D., Cesana, E., Menon, R., Vuono, R., Talpo, F., Laguna Goya, R., Lyons, P.A., Bulfamante, G.P., Muzio, L., Martino, G., Toselli, M., Farina, C., A Barker, R., Biella, G., Cattaneo, E., 2014. Molecular and functional definition of the developing human striatum. Nat. Neurosci. 17 (12), 1804–1815. https://doi.org/10.1038/nn.3860.

Ooi, J., Langley, S.R., Xu, X., Utami, K.H., Sim, B., Huang, Y., Harmston, N.P., Tay, Y.L., Ziaei, A., Zeng, R., Low, D., Aminkeng, F., Sobota, R.M., Ginhoux, F., Petretto, E., Pouladi, M.A., 2019. Unbiased profiling of isogenic Huntington disease hPSC-derived CNS and peripheral cells reveals strong cell-type specificity of CAG length effects. Cell Rep. 26 (9), 2494–2508.e7. https://doi.org/10.1016/j.celrep.2019.02.008.

Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., Melton, D.A., 2008. Marked differences in differentiation propensity among human embryonic stem cell lines. Nat. Biotechnol. 26, 313–315. https://doi.org/10.1038/nbt1383.

Osipovitch, M., Asenjo Martinez, A., Mariani, J.N., Cornwell, A., Dhaliwal, S., Zou, L., Chandler-Militello, D., Wang, S., Li, X., Benraiss, S.J., Agate, R., Lampp, A., Benraiss, A., Windrem, M.S., Goldman, S.A., 2019. Human ESC-derived chimeric mouse models of Huntington's disease reveal cell-intrinsic defects in glial progenitor cell differentiation. Cell Stem Cell. https://doi.org/10.1016/j.stem.2018.11.010.

Ramocki, M.B., Zoghbi, H.Y., 2008. Failure of neuronal homeostasis results in common neuropsychiatric phenotypes. Nature. https://doi.org/10.1038/nature07457.

Rataj-Baniowska, et al., 2015. Retinoic acid receptor β controls development of striato-nigral projection neurons through FGF-dependent and meis1-dependent mechanisms.

J. Neurosci. 14467–14475. https://doi.org/10.1523/JNEUROSCI.4440-15.2016. Correction.

Ravache, M., Weber, C., Mérienne, K., Trottier, Y., 2010. Transcriptional activation of REST by Sp1 in huntington's disease models. PLoS One. https://doi.org/10.1371/journal.pone.0014311.

Ray, S.K., Li, H.J., Metzger, E., Schule, R., Leiter, A.B., 2014. CtBP and associated LSD1 are required for transcriptional activation by NeuroD1 in gastrointestinal endocrine cells. Mol. Cell. Biol. https://doi.org/10.1128/mcb.01600-13.

Ring, K.L., An, M.C., Zhang, N., O'Brien, R.N., Ramos, E.M., Gao, F., Atwood, R., Bailus, B.J., Melov, S., Mooney, S.D., Coppola, G., Ellerby, L.M., 2015. Genomic analysis reveals disruption of striatal neuronal development and therapeutic targets in human Huntington's disease neural stem cells. Stem Cell Rep. https://doi.org/10.1016/j.stemcr.2015.11.005.

Ruzo, A., Croft, G.F., Metzger, J.J., Galgoczi, S., Gerber, L.J., Pellegrini, C., Wang, H., Fenner, M., Tse, S., Marks, A., Nchako, C., Brivanlou, A.H., 2018. Chromosomal instability during neurogenesis in huntington's disease. Development (Cambridge). https://doi.org/10.1242/dev.156844.

Sampedro, F., Martínez-Horta, S., Perez-Perez, J., Horta-Barba, A., Martin-Lahoz, J., Alonso-Solís, A., Corripio, I., Gomez-Anson, B., Kulisevsky, J., 2019. Widespread increased diffusivity reveals early cortical degeneration in Huntington disease. Am. J. Neuroradiol. https://doi.org/10.3174/ajnr.A6168.

Schoenherr, C.J., Anderson, D.J., 1995. The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science 267(5202) (5202), 1360–1363. https://doi.org/10.1126/science.7871435.

Seo, S.Y., Lee, B., Lee, S., 2015. Critical roles of the LIM domains of Lhx3 in recruiting Coactivators to the motor neuron-specifying Isl1-Lhx3 complex. Mol. Cell. Biol. https://doi.org/10.1128/mcb.00335-15.

Shi, Y., Kirwan, P., Smith, J., Robinson, H.P.C., Livesey, F.J., 2012. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. Nat. Neurosci. https://doi.org/10.1038/nn.3041.

Shimohata, T., Nakajima, T., Yamada, M., Uchida, C., Onodera, O., Naruse, S., Kimura, T., Koide, R., Nozaki, K., Sano, Y., Ishiguro, H., Sakoe, K., Ooshima, T., Sato, A., Ikeuchi, T., Oyake, M., Sato, T., Aoyagi, Y., Hozumi, I., ... Tsuji, S., 2000. Expanded polyglutamine stretches interact with TAF(II)130, interfering with CREB-dependent transcription. Nat. Genet. https://doi.org/10.1038/79139.

Smith, M.A., Brandt, J., Shadmehr, R., 2000. Motor disorder in Huntington's disease begins as a dysfunction in error feedback control. Nature. https://doi.org/10.1038/35000576.

Smith-Geater, C., Hernandez, S.J., Lim, R.G., Adam, M., Wu, J., Stocksdale, J.T., Wassie, B.T., Gold, M.P., Wang, K.Q., Miramontes, R., Kopan, L., Orellana, I., Joy, S., Kemp, P.J., Allen, N.D., Fraenkel, E., Thompson, L.M., 2020. Aberrant development corrected in adult-onset Huntington's disease iPSC-derived neuronal cultures via WNT Signaling modulation. Stem Cell Rep. https://doi.org/10.1016/j.stemcr.2020.01.015.

Strano, A., Tuck, E., Stubbs, V. E., & Livesey, F. J. (2020). Variable Outcomes in Neural Differentiation of Human PSCs Arise from Intrinsic Differences in Developmental Signaling Pathways. Cell Reports. https://doi.org/10.1016/j.celrep.2020.107732.

Thaler, J.P., Lee, S.K., Jurata, L.W., Gill, G.N., Pfaff, S.L., 2002. LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions. Cell. https://doi.org/10.1016/S0092-8674(02)00823-1.

Tremblay, R., Lee, S., Rudy, B., 2016. GABAergic interneurons in the Neocortex: from cellular properties to circuits. Neuron. https://doi.org/10.1016/j.neuron.2016.06.033.

Wang, R., Luo, Y., Ly, P.T.T., Cai, F., Zhou, W., Zou, H., Song, W., 2012. Sp1 regulates human Huntingtin gene expression. J. Mol. Neurosci. https://doi.org/10.1007/s12031-012-9739-z.

White, J.K., Auerbach, W., Duyao, M.P., Vonsattel, J.P., Gusella, J.F., Joyner, A.L., MacDonald, M.E., 1997. Huntington is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. Nat. Genet. https://doi.org/10.1038/ng1297-404.

Wyttenbach, A., 2001. Polyglutamine expansions cause decreased CRE-mediated transcription and early gene expression changes prior to cell death in an inducible cell model of Huntington's disease. Hum. Mol. Genet. https://doi.org/10.1093/hmg/10.17.1829.

Xie, Y., Hayden, M.R., Xu, B., 2010. BDNF overexpression in the forebrain rescues Huntington's disease phenotypes in YAC128 mice. J. Neurosci. https://doi.org/10.1523/JNEUROSCI.1637-10.2010.

Xu, X., Tay, Y., Sim, B., Yoon, S.-I., Huang, Y., Ooi, J., Utami, K.H., Ziaei, A., Ng, B., Radulescu, C., Low, D., Ng, A.Y.J., Loh, M., Venkatesh, B., Ginhoux, F., Augustine, G.J., Pouladi, M.A., 2017. Reversal of phenotypic abnormalities by CRISPR/Cas9-mediated gene correction in Huntington disease patient-derived induced pluripotent stem cells. Stem Cell Rep. 8 (3), 619–633. https://doi.org/10.1016/j.stemcr.2017.01.022.

Yaden, B.C., Garcia, M., Smith, T.P.L., Rhodes, S.J., 2006. Two promoters mediate transcription from the human LHX3 gene: involvement of nuclear factor I and specificity protein 1. Endocrinology. https://doi.org/10.1210/en.2005-0970.

Zhang, X.Q., Zhang, S.C., 2010. Differentiation of neural precursors and dopaminergic neurons from human embryonic stem cells. Methods Mol. Biol. https://doi.org/10.1007/978-1-60761-369-5_19. (Clifton, N.J.).

Zhang, X.H., Zheng, B., Gu, C., Fu, J.R., Wen, J.K., 2012. TGF-β1 downregulates AT1 receptor expression via PKC-δ-mediated Sp1 dissociation from KLF4 and smad-mediated PPAR-γ association with KLF4. Arterioscler. Thromb. Vasc. Biol. https://doi.org/10.1161/ATVBAHA.111.244962.

Zhang, L., Zhang, P., Wang, G., Zhang, H., Zhang, Y., Yu, Y., Zhang, M., Xiao, J., Crespo, P., Hell, J. W., Lin, L., Huganir, R. L., & Zhu, J. J. (2018). Ras and Rap Signal Bidirectional Synaptic Plasticity via Distinct Subcellular Microdomains. Neuron. https://doi.org/10.1016/j.neuron.2018.03.049.

Zuccato, C., Cattaneo, E., 2009. Brain-derived neurotrophic factor in neurodegenerative diseases. In Nature Reviews Neurology. https://doi.org/10.1038/nrneurol.2009.54.

Zuccato, C., Marullo, M., Vitali, B., Tarditi, A., Mariotti, C., Valenza, M., Lahiri, N., Wild, E.J., Sassone, J., Ciammola, A., Bachoud-Lèvi, A.C., Tabrizi, S.J., Di Donato, S., Cattaneo, E., 2011. Brain-derived neurotrophic factor in patients with Huntington's disease. PLoS One. https://doi.org/10.1371/journal.pone.0022966.