UNIVERSITÀ DEGLI STUDI DI MILANO

PhD Course in Molecular and Cellular Biology

XXX Cycle

The huntingtin CAG repeats as 'fine-tuning knob' for protein neural function

Raffaele Iennaco

PhD Thesis

Scientific tutor: Elena Cattaneo

Academic year: 2016-2017

SSD: BIO/14

Thesis performed at University of Milan, Department of Biosciences

Laboratory of Stem Cell Biology and Pharmacology
of Neurodegenerative Diseases

Directed by Prof. Elena Cattaneo

INDEX

ABSTRACT **6**

1. INTRODUCTION **7**

1.1 HTT structure and function **8**

1.1.1 *HTT* gene evolution 8

1.1.2 HTT expression and distribution 11

1.1.3 HTT structure 12

1.1.4 HTT PTMs, proteolysis, and interactors 14

1.1.4 HTT functions 17

HTT in development 17

HTT regulates tissue maintenance and cell morphology 19

HTT in cell survival 20

1.2 CAG trinucleotide repeats **22**

1.2.1 Simple sequence repeats and their biological meaning 22

1.2.2 *HTT* CAG repeats variability in human population 24

1.3 ESCs: Cell-based platform to study HTT functions **26**

1.3.1 Embryonic Stem Cells (ESCs) 26

1.3.2 From ESCs to neurons: neural differentiation protocols 30

1.3.3 Neural rosette formation as *in vitro* model for neurogenesis 32

2. AIM **35**

3. METHODS **37**

4. RESULTS **43**

4.1 HTT rosette formation potential is confined to its N-terminal portion **44**

4.2 Generation of cell lines stably expressing mutated versions of the CAG tract **44**

4.3 PolyQ stretch deletion causes defects in rosette formation **47**

4.4 PolyQ stretch increases rosette formation in a length-dependent manner **48**

4.5 Interruption of polyQ tract purity reduces *in vitro* neurulation **49**

4.6 Pathological HTT polyQ tract impairs neurulation **50**

4.7 PolyQ domain is not involved in the antiapoptotic function of HTT **51**

4.8 Automatization of rosette assay for high-content analysis **54**

4.8.1 Optimization of neural differentiation protocol procedures 54

4.8.2 Automatization of image acquisition and rosette quantification 56

4.8.3 Evaluation of rosette phenotype by automated method 58

5. DISCUSSION AND CONCLUSIONS **61**

6. REFERENCES **64**

ABSTRACT

Introduction. Huntingtin gene (*HTT*) encodes for a high molecular weight protein (HTT) that is essential during gastrulation and in forebrain formation. The N-terminal region of the protein has been recently associated to *in vitro* pro-neurulation activity. This region includes a peculiar CAG repeat stretch whose expansion results in an inherited neurodegenerative disorder named Huntington's disease (HD). The disease itself, its penetrance and age of onset are dependent on the length of CAG stretch that encodes for a polyglutamine (polyQ) tract. Evolutionary studies showed that this polyQ tract appeared for the first time in echinoderms, then increased in length during deuterostome evolution in parallel to the emergence of progressively more complex nervous systems. Despite these indications, the functional significance of the polyQ variability in normal HTT remains to be explored.

Aim. Here we investigate whether the polyQ tract in HTT is biologically active during early neurogenesis. In particular, as a tool to study *in vitro* neurogenesis, we adopted the rosette assay, an *in vitro* assay that measures the ability of ES cell-derived neural progenitors to form radial arrangements of columnar cells (named neural rosettes) that mimic neural tube formation *in vivo*. Additionally, we worked on designing a novel automated method for faster and unbiased quantifications of the rosette phenotype.

Methods. We produced a mouse embryonic stem (mES) cell-based platform for complementation assays with HTT N-terminal portion bearing different polyQ repeats, their deletion and substitutions (0Q, 2Q, 4Q, 7Q, 15Q, 128Q, Q3Q(CAA)Q3, and Q3PQ3). The newly generated cells were exposed to a neural induction protocol to assess their rosette formation potential.

Results. We report that the pro-neurulation activity, but not the antiapoptotic function, of HTT is impaired by the sole polyQ domain deletion from the N-terminal portion of HTT. Moreover, the CAG tract length affects rosette formation potential in a linear fashion. Conversely, either a pathological CAG expansion or a CAG interruption causes loss of HTT N-terminal pro-neurulation activity. Finally, to further characterize this phenotype in an unbiased and quantitative way, we developed and tested a novel automated system, based on high-content imaging approach that allowed us to pinpoint this trait with unparalleled precision and efficiency.

Conclusions. Overall, these results demonstrate that HTT ability to promote rosette formation is linked to a persistent and uninterrupted CAG tract, while HTT antiapoptotic activity lays outside HTT polyQ domain.

1. INTRODUCTION

Huntingtin (HTT) is a high molecular weight protein (348kDa) that is encoded by the *Huntingtin* gene (*HTT*); also known as *IT-15* and maps on chromosome 4. While the human HTT protein (NCBI RefSeq: NP_002102) is 3144 amino acids long, the mouse protein (NP_034544) contains 3120 amino acids. When excluding the differences in their polyglutamine (polyQ) tract and proline-rich domain (PRD), the human and mouse proteins are 91% identical and 95% similar (*Ehrnhoefer et al., 2009*).
An expansion of the CAG repeats (CAGs) within the coding region of the *HTT* results in an elongated stretch of glutamine (Q) in the HTT N-terminus that causes an inherited autosomal dominant neurodegenerative disorder named Huntington's disease, HD (*HD Collaborative Research Group, 1993*). In the non-HD population the CAG triplet is repeated between 9 and 35 times, while an expansion above 36 repeats results in the development of the disease. Rare carriers of 36 to 39 CAGs show a lower penetrance and later onset of the disease compared to those with 40 or more CAGs (*McNeil et al., 1997; Quarrell et al., 2007*). The neuropathology of HD results in a massive brain neurodegeneration characterized by the prevalent dysfunction and loss of specific neurons (efferent medium spiny neurons) in the striatum of the basal ganglia, which is the principal responsible for the typical HD symptoms (*Reiner et al., 1988*). However, corticostriatal projection neurons are also particularly affected and a more widespread degeneration of the brain (including the layers V and VI of cortex) is also reported (*Rosas et al., 2003, 2005, 2008*). HD symptoms include generalized motor dysfunction, psychiatric decline and cognitive disturbances. Although alterations of the central nervous system (CNS) are the main characteristics of the disease, patients also show alterations at the metabolic level, such as immune disturbances, osteoporosis, cardiac failure, weight loss and testicular and skeletal muscle atrophy (*van der Burg et al., 2009*). The symptoms last about 20 years after onset, and death often occurs for fatal aspiration pneumonia. Today there is still no cure for HD.
As HD develops because of the dominant toxic activity of a single mutant allele, the majority of the studies in the field are focussed on investigating the impact of mutant HTT on neuronal and brain function. However, the study of the biology of normal HTT and its function allows a better understanding on how exactly the mutant protein causes the disease and can give new perspectives on future therapeutic targets. Therefore, the laboratory in which this thesis has been performed, together with others, has focused on identifying HTT native functions and its physiological roles. Based on these studies published over the past 20 years, in the first section of this thesis work I describe the literature regarding the evolution, structure and known functions of the wild-type *HTT* gene.

1.1 HTT structure and function

1.1.1 *HTT* gene evolution

The *HTT* gene can be traced back very early in life history, before the divergence between protostomes and deuterostomes (occurred approximately 670 mya according to *Erwin et al., 2011*). Indeed, an ortholog of the *HTT* gene, that carries no CAGs, is found in the very ancient organism *Dictyostelium discoideum* (*Palidwor et al., 2009; Myre et al., 2011;* **Figure 1.1**). *Dictyostelium D. is a* soil-living amoeba that lives mostly as a single-celled separate organism, but during starvation they are able to stream together to form a multicellular organism. Although the inactivation of *HTT* gene in *Dictyostelium* does not compromise cell growth, it leads to the loss of capacity to form a multicellular organism (*Myre et al., 2011*). It has been also shown that the *Dictyostelium HTT* gene inactivation impairs cytokinesis, cell shape, chemotaxis, and homophilic cell-cell adhesion during hypo-osmotic stress (*Myre et al., 2011; Wang et al., 2011*). Our laboratory has shown that *Dictyostelium* HTT is antiapoptotic when the corresponding gene is expressed in mammalian cells (*Lo Sardo et al., 2012*). This function is related to the N-terminal portion of HTT and is conserved in mammals (*Rigamonti et al., 2000, 2001; Zuccato et al., 2010)*.

Most of the studies regarding HTT function in protostomes come from *Drosophila melanogaster*, whose *HTT* gene has 29 exons. *Drosophila* HTT protein was predicted to be composed of 3583 amino acids and, in contrast to mammals, lacks the polyQ and proline-rich domain (PRD) (*Li et al., 1999*). HTT in *Drosophila* is ubiquitously expressed during all stages of fly development (*Li et al., 1999*). Although there is no polyQ tract in the N-terminal portion of *Drosophila* HTT (**Figure 1.1)**, several polyQ stretches are found in other portions of the protein (e.g., *Drosophila yakuba* has 10 and 12 glutamines at position 625 to 634 and 1118 to 1131, respectively) which are lacking in the human HTT, suggesting that in *Drosophilae* lineages independent polyQ tract insertion events have occurred (*Schaefer et al., 2012*). Moreover, bioinformatic analyses of HTT from *Drosophila melanogaster*, *Drosophila pseudoobscura*, *Apis mellifera* (honey bee), and *Tribolium castaneum* (red flour beetle) revealed a more heterogeneous evolution in insect class compared with deuterostomes (*Tartari et al., 2008*). Putative *HTT* null mutant of *Drosophila* shows normal gastrulation (unlike mice), suggesting that in *Drosophila* the protein is not involved in embryogenesis (*Zhang et al., 2009; Zuccato et al., 2010*). These two different phenotypes reveal apparent differences in the mouse/fly embryonic developmental process (*Li et al., 1999; Tartari et al., 2008*) or may be the result of compensatory mechanisms for the loss of the protein in the *HTT* knock-out flies (*Zheng and Joinnides, 2009*). Further studies in adult *HTT* knock-out *Drosophila* did not show any defect in synapse formation, neurotransmission, or axonal transport even if a decreased number of branches of giant fiber neurons were observed (*Zhang et al., 2009*). In contrast, Gunawardena and his group reported the involvement of *Drosophila* HTT in axonal transport and synaptogenesis (*Gunawardena et al., 2003*). Loss of HTT also revealed mobility defects and reduced viability in aged flies, suggesting an important role of the protein in adulthood (*Zhang et al., 2009*).

More recent findings revealed that specific functions of HTT are conserved between flies and mammals. An *in vitro* experiment found that *Drosophila* HTT may functionally replace the missing murine HTT protein, validating the conservation of the protein's function (*Godin et al., 2010*). By using an *in vitro* assay, our laboratory has shown that *Drosophila* HTT has an antiapoptotic role and that this activity is also well conserved in the deuterostome lineage (*Lo Sardo et al., 2012*). Indeed, this function appeared in *Dictyostelium* and has then been maintained during evolution, as showed by the ability of heterologous HTT proteins (from sea urchin, *Ciona*, anphioxus, zebrafish, mouse, and human) to promote cell survival (*Lo Sardo et al., 2012*). The axonal transport function of HTT is also conserved across mammals and flies (*Zala et al., 2013*).

In order to discover relevant huntingtin domains and activities, a series of *HTT* homologues from both protostome and deuterostome branches have been cloned and used to construct a first multiple alignment in the N-terminal region of HTT. In particular, the protein sequence of organisms that represent fundamental turning points in the phylogenetic tree, with regards to the development and evolution of the nervous system, have been reconstructed. The sequence of HTT from echinoderm *Strongylocentrotus purpuratus* (purple sea urchin) (*Tartari et al., 2008*), urochordate *Ciona intestinalis* (tunicate sea squirt) (*Gissi et al., 2006*), and cephalocordate *Branchiostoma floridae* (amphioxus) (*Candiani et al., 2007*), together with HTT from insects and vertebrates have been used to identify the putative HTT domains and to evaluate the evolutionary relevance of the polyQ tract.

The first appearance of the polyQ tract in HTT is in the *Strongylocentrotus purpuratus* that carries an NHQQ sequence in the N-terminal portion. This portion can be considered biologically comparable to the four glutamines found in lower vertebrates (*Tartari et al., 2008;* **Figure 1.1** and **1.2**). Two glutamines are also present in the cephalocordate *Branchiostoma floridae* (*Candiani et al., 2007*), but here they are preceded by two hydrophobic amino acids (alanine and phenylalanine, AF) in contrast with the two polar amino acids (asparagine and histidine, NH) present in the sea urchin. Another difference exists between the N-terminal sequences of the two organisms, which is the presence of a 21 residues tract upstream of the NHQQ of the sea urchin; this tract appears different in composition with respect to the amphioxus. All these findings suggest a putative stabilization of the polyQ tract in the N-terminal portion of the amphioxus HTT due to selective evolutionary pressure (*Candiani et al., 2007; Tartari et al., 2008*). Surprisingly, HTTs of the tunicates *Ciona savigny* and *Ciona intestinalis* have no polyQ tract, and they have probably suffered a deletion in the N-

terminal portion which is shorter than in their vertebrate orthologs (*Gissi et al., 2006*). Although *Ciona* possesses the main chordate traits during neural development, these differences in protein sequence are not unexpected because these urochordates quickly diverged from the other chordates, probably due to adaptation of their specific ecologic niche (*Hughes and Friedman, 2005*).

In vertebrates the length of the polyQ tract increases, from 4 to 35 glutamines (threshold for the healthy human population) (**Figure 1.1** and **1.2**). In particular, 4 glutamines are present in fish, amphibians, reptiles and birds. Six glutamines are found in HTT from the marsupial *Monodelphis domestica* (gray short-tailed opossum), 7 in *Mus musculus* (mouse), 8 in *Rattus norvegicus* (rat), 10 in *Canis familiaris* (dog), 15 in *Bos taurus* (domestic cow), and 18 in *Sus scrofa* (pig) (**Figure 1.1** and **1.2**). Moreover, the analyses of the polyQ tract in non-human primates from 10 different species revealed that the number of glutamines was remarkably consistent between species, ranging from 7 to 16Q (*Rubinsztein et al., 1994*). The sequencing findings of our unpublished work are in line with this polyQ length range in non-human primates (**Figure 1.1** and **1.2)**. In humans, the polyQ tract reaches the longest and most polymorphic length, from 9 to 35.

Early studies suggested that the normal human CAG repeats are not in equilibrium but instead appear to be subject to a mutational bias that causes a gradual expansion of the CAG length (*Rubinsztein et al., 1994*). A full discussion of the CAG repeats in humans and their biological meaning is given in chapter 1.2. As mentioned before, a PRD domain follows the polyQ tract. However, this domain appears exclusively in higher vertebrates and specifically in mammals and thus represents a recent and sudden acquisition of a particular region in the evolution of HTT (**Figure 1.2**). Although many authors propose that this region is able to stabilize the polyQ tract by keeping it soluble (*Steffan et al., 2004*), an open question remains about the function of the PRD.

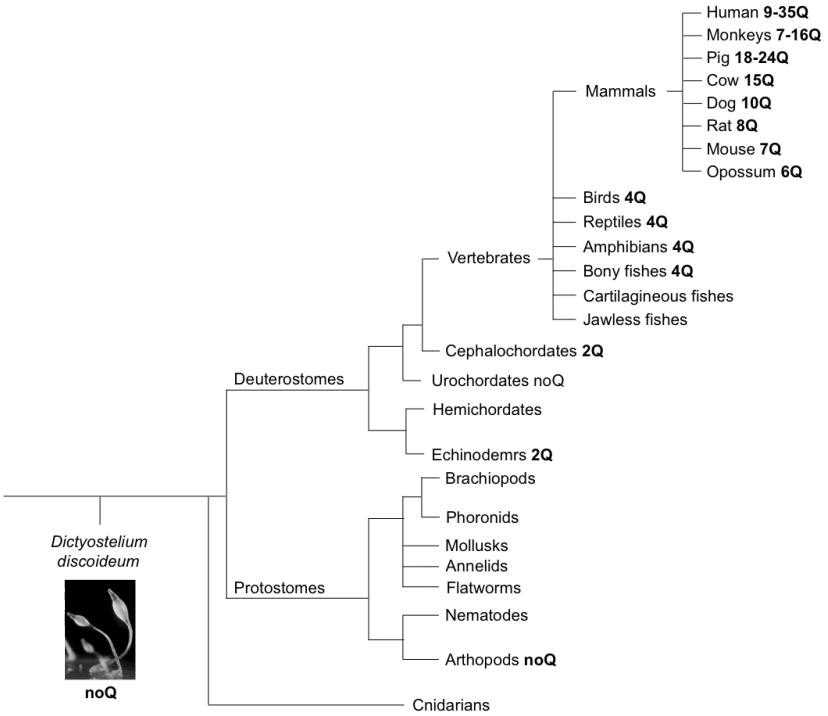


Figure 1.1. HTT polyQ tract evolution along the human lineage. Schematic illustration of a phylogenetic tree of huntingtin polyQ tract from amoeba *Dictyostelium discoideum* to humans. Number of Q repeats within the polyQ tract is reported.

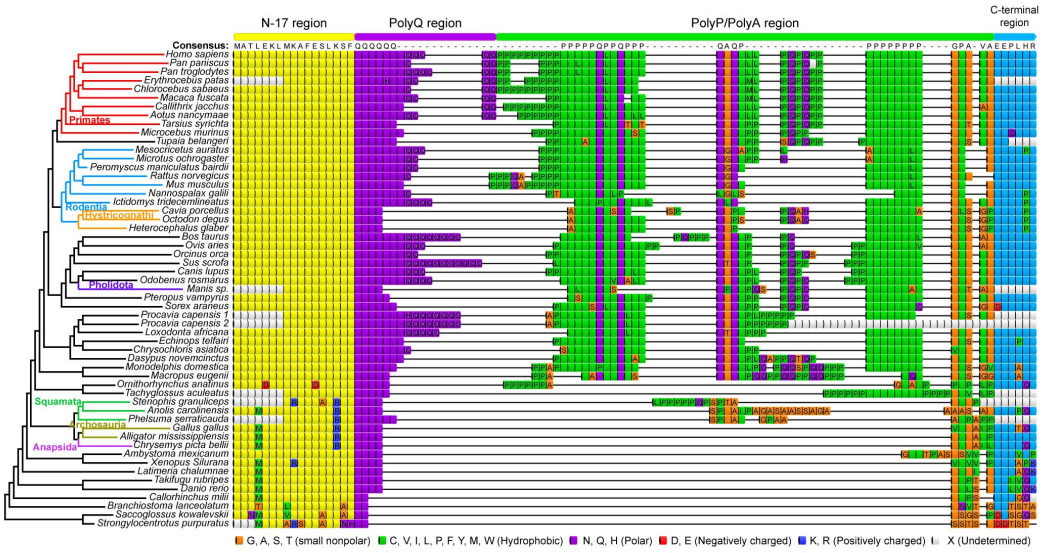


Figure 1.2. Representative multiple sequence alignment of HTT N-terminus in 55 sequences along the human lineage. The polyQ and PRD portions are highlighted in violet and green, respectively. In the N-17 region (yellow) and in the C-terminal region (blue) only amino acids differing from consensus are highlighted according to the colour scheme of Lesk algorithm (reported in the bottom).

1.1.2 HTT expression and distribution

Analysis of HTT expression revealed that both in humans and mice the protein is ubiquitously distributed, even if the highest levels have been detected in CNS and testis (*Strong et al., 1993; Landwehrmeyer et al., 1995; Sharp et al., 1995; Bhide et al., 1996; Sapp et al., 1997; Vonsattel and DiFiglia, 1998*).

HTT gene in human and rodents has two mRNA isoforms (*Htt*). These transcripts (10366 bp and 13711 bp) are generated by alternative cleavage and polyadenylation of the primary transcript; the second transcript, in particular, differs by an additional 3' UTR sequence of 3360 bp (*Lin et al., 1993, 1994*). While the long transcript is predominantly expressed in the brain of both species, the shorter transcript seems to be more widely expressed. Later studies also reported that mutant *HTT* gene is also mis-spliced, generating a short polyadenylated mRNA. This transcript is translated into a pathogenic exon 1 HTT protein (*Sathasivam et al., 2013*). Only very recently, a study reported an alternative splicing for HTT originating protein variants that lack exons 10, 12, 29, and 46 or, alternatively, retain a 57 bp portion of intron 28 or even an additional exon (41b) (*Hughes et al., 2014; Ruzo et al., 2015*). These HTT isoforms are rare, however some of them could be upregulated during development. Moreover, the absence of these exons may lead to changes in HTT susceptibility to cleavage, phosphorylation, and ability to interact with other proteins (*Hughes et al., 2014; Ruzo et al., 2015*).

HTT is expressed from post-fertilization stages, as revealed by an expression study in mouse and human development (*Dragatsis et al., 1998*). Later, *HTT* expression increases in concomitance with the neuronal maturation in the post-natal stage (*Bhide et al., 1996*). Furthermore, evolutionary analyses highlighted that HTT is predominantly present in non-neural tissues in the sea urchin (*Heliocidaris herithrogramma*), while it is ubiquitously distributed in the primitive chordate sea pineapple (*Halocynthia roretzi*) with highest expression level in the nervous system (*Kauffman et al., 2003*). Likewise, in amphioxus (*Branchiostoma floridae*) an increased HTT expression level has been shown in the nervous system, with the strongest level being present in the most anterior part of neural tube (*Candiani et al., 2007*). On the whole, these results suggest that the neural expression of HTT could be a chordate-specific feature.

Within the nervous system, HTT expression is not restricted to the brain regions that degenerate in the HD, indeed HTT is also found in other brain areas. In particular, when analysing mouse and human post-mortem brain, HTT expression was especially abundant in neurons of the telencephalon and thalamus, but had a poor expression in the hypothalamus (*Strong et al., 1993; Landwehrmeyer et al., 1995; Sharp et al., 1995; Bhide et al., 1996; Sapp et al., 1997; Vonsattel and DiFiglia, 1998*). In the telencephalon, HTT-positive neurons are found in the neocortex, especially the pyramidal neurons of layers 3 and 5, which project to the striatum (*Fusco et al., 1999*). Also the hippocampus exhibited high level of HTT expression, particularly the pyramidal neurons of the CA2 and CA3 layers. The same study revealed that the striatum contains scattered large neurons rich in HTT and more numerous medium-sized neurons with a moderate HTT content (*Fusco et al., 1999*). The CAG-expansion mutation of the *HTT* gene seems to have no effect on the regional expression of HTT in the brain (*Aronin et al., 1995; Landwehrmeyer et al., 1995; Schilling et al., 1995; Bhide et al., 1996; Sapp et al., 1997; Vonsattel and DiFiglia, 1998*).

Regarding intracellular localization, HTT is present in the soluble cytoplasmic compartment, as well as in cell bodies, dendrites, and axons (*Sharp et al., 1995; Vonsattel and DiFiglia, 1998*). Further studies identified HTT in brain membrane fractions (*Suopanki et al., 2006*), suggesting an interaction between HTT and lipids. Moreover, HTT interacts with vesicle membrane and microtubules (*DiFiglia et al., 1995; Gutekunst et al., 1995; Sharp et al., 1995)*, where it plays a role in axonal and vesicular transport. HTT is also located to the endoplasmic reticulum (ER) as a consequence of an activity embedded into the first 18 amino acids of HTT N-terminus (*Atwal and Truant, 2008*). This domain is also capable of targeting HTT to late endosomes as well as in autophagic vesicles, implying a putative role of HTT in vesicles trafficking and in the formation of autophagic vacuoles (*Atwal and Truant, 2008*). In addition, HTT localizes to the Golgi complex and perinuclear tubulovesicular membranes (*Velier et al., 1998; Strehlow et al., 2007*) and is associated with mitochondria (*Gutekunst et al., 1998; Hilditch-Maguire et al., 2000; Panov et al., 2002; Choo et al., 2004*). Finally, nuclear localization has been described for wild-type HTT (*Hoogeveen et al., 1993; De Rooij et al., 1996; Sapp et al., 1997; Dorsman et al., 1999; Wilkinson et al., 1999; Kegel et al., 2002*). A dynamic movement of HTT across the nuclear membrane has also been reported. This capacity is associated to the HEAT repeats (*Truant et al., 2006*) and regulated by a nuclear import signal located in the HTT N-terminus (*Atwal et al., 2007; Desmond et al., 2012*) and by a nuclear export signal located in the HTT C-terminus (*Xia et al., 2003*). Finally, HTT localizes at the mitotic spindle pole (*Godin et al., 2010*), where it regulates cell division; this localization seems to be associated to the HTT phosphorylation at serines 13 and 16 (*Godin et al., 2010; Atwal et al., 2011; Maiuri et al., 2013*).

1.1.3 HTT structure

The high molecular weight of HTT protein hampers the use of classical approaches (production of crystals and mass spectrometry studies) for elucidating its structure; therefore, most of the structural findings derive from computational analyses. Most of the HTT studies were focused on the N-terminal region of the protein and specifically on the *HTT* exon 1 (encoding for amino acids 1 to 68, as it contains the expandable polyQ tract). In addition to this tract, exon 1 contains a portion encoding for the first 17 amino acids of HTT (named N17) that precedes the polyQ tract, and a PRD domain that follows the polyQ tract. Both polyQ tract and PRD are polymorphic in humans.

The N17 portion is highly conserved in vertebrates but less in protostomes (*Tartari et al., 2008*). This region acts as a nuclear export signal (NES) and its structure is an amphipathic α-helix (*Atwal et al., 2007*), which is significant for retention in the endoplasmic reticulum (*Atwal et al., 2007; Rockabrand et al., 2007*). Many post-translational modifications (PTMs) are present in the N17, such as acetylation, sumoylation, and ubiquitination at lysines 6, 9, and 15 (K6, K9 and K15) and phosphorylation at serines 13 and 16 (S13 and S16,

respectively) that affect the clearance of HTT and its subcellular localization (**Table 1.1**) (*Atwal et al., 2007; Maiuri et al., 2013; Steffan et al., 2004; Thompson et al., 2009*).

The first three-dimensional structure of a polyQ stretch was revealed by Max Perutz and his team in 1994. They showed that a peptide of 15 Qs was able to adopt a β-sheet structure and that a combination of these units can form the "polar zipper" atomic model, which consists of antiparallel β-strands held together by hydrogen bonds (*Perutz et al., 1994*). They also suggested that the physiological function of the polyQ tract was to bind transcription factors that contain another polyQ region (*Perutz et al., 1994*). The PRD, typical of the mammalian group, is variable and is important for the interaction with proteins that contain tryptophans or Src homology 3 domains (*Schaefer et al., 2012*). It has a proline-proline (PP) helix, a rigid structure that adopts kinked or straight conformation (*Kim et al., 2009*). The PP helix of the PRD functions as stabilizer of the polyQ tract, to avoid a conformational collapse, and this ability may account for a known role played by PRD region in the aggregation/toxicity of mutant HTT (*Bhattacharyya et al., 2006; Darnell et al., 2007; Dehay and Bertolotti, 2006; Duennwald et al., 2006; Kim et al., 2009*). Notwithstanding the above mentioned characteristics, the deletion of the PRD domain *in vivo* has no pronounced effect on mouse behavior (*Neveklovska et al., 2012*). X-ray crystallography resolved the secondary structure of an HTT N-terminal region carrying 17Q. In particular, while the N17 portion forms an α-helical structure, the poly17Q stretch adopts multiple flexible conformations including α-helix, random coil, and extended loop (*Kim et al., 2009*). The conformation of the poly17Q region is influenced by the conformation of neighbouring domains, demonstrating the importance of the native protein context (*Kim et al., 2009*).

The putative roles of these three HTT domains (polyQ, PRD and N17) have been investigated *in vivo*. In particular, the consequence of deleting mouse HTT N17, polyQ and PRD domains, or a combination of them, *in vivo,* support the idea that these domains are dispensable for HTT critical functions during early embryonic development, but are likely more important for HTT functions in later CNS development or maintenance (*Clabough and Zeitlin, 2006; Zheng et al., 2010; Neveklovska et al., 2012; André et al., 2017*).

The rest of the protein (66 exons encoding amino acids from 69 to 3144) is less well characterized. A secondary structure analysis of full-length HTT showed that it is composed of about 40% α-helix, 12%, β-strand and 48% turn and coil (*Huang et al., 2015*). Further studies showed the presence of several HEAT repeats, so called because they are found in Huntingtin, Elongation factor 3, protein phosphatase 2A, and the yeast kinase TOR1. The HEAT repeats are amino acid sequences (about 40aa long) that occur multiple times within a protein, and they are formed of antiparallel α-helices separated by a non-helical portion. These repeated regions are important for protein-protein interaction (*Andrade and Bork, 1995; Neuwald and Hirano, 2000; Andrade et al., 2001; Palidwor et al., 2009*). The HEAT repeats of HTT are well conserved in deuterostomes, suggesting that they are an ancestral feature in HTT evolution. In particular, bioinformatics analyses reported between 16 and 36 HEAT repeats clustered into three to five three-dimensional structures, called α-rod domains, separated by disordered regions (**Figure 1.3**) (*Palidwor et al., 2009; Takano and Gusella, 2002; Tartari et al., 2008; Warby et al., 2008*). The HEAT repeat domains may function as a solenoid-like structure that acts as a scaffold for numerous protein complexes and mediates inter- and intra-molecular interactions. The middle region of HTT (507-1230aa) can bind to the N-terminal (1-506aa) and C-terminal (2721-3144aa) domains of HTT (*Palidwor et al., 2009*); the 507-1230aa domain can also self-associate to form HTT homodimers. Other studies demonstrated that two HTT N-terminal portions (1-416aa and 1-586aa) bind to different C-terminal regions of HTT (1725-2800aa and 2416-3144aa, respectively), and these intramolecular interactions are disrupted upon proteolysis (*El‐Daher et al., 2015; Ochaba et al., 2014*). Very recently a study on the full-length HTT revealed that it is composed of five distinct domains and confirmed that HTT adopts a spherical α-helical solenoid structure where the N-terminal and C-terminal regions fold to contain a circumscribed central cavity. Interestingly, the authors also showed that the polyQ tract expansion increases the α-helical properties of HTT and affects the intramolecular interactions among the domains (*Vijayvargia et al., 2016*). All these observations suggest that HTT can adopt various three-dimensional conformations, depending on its intra-molecular interactions. These interactions may also involve

other proteins to form complexes, and in support of this, numerous HTT interactors have been found (see paragraph 1.1.4). Consistently, purified HTT can adopt up to 100 structurally distinguishable conformations (*Seong et al., 2010*).

Finally, other functional motifs able to regulate HTT function or localization have been identified (*Xia et al., 2003*), such as the NES site at position 2397-2406aa.

1.1.4 HTT PTMs, proteolysis, and interactors

HTT post-translational modifications (PTMs). HTT undergoes multiple PTMs, including phosphorylation, acetylation, palmitoylation, ubiquitination, and sumoylation (**Figure 1.3**; **Table 1.1**). The role of these PTMs has been analysed mostly in the context of the mutant polyQ protein. HTT PTMs are critical for HTT stability, cellular localization, modulation of protein-protein interactions, regulation of vesicle transport and autophagy. Moreover, these modifications could also impact on the function of other proteins (*Zuccato et al., 2010; Saudou and Humbert, 2016*). Several PTMs have a therapeutic relevance, as they modulate toxicity of mutant HTT, and result in phenotypic changes in animal models of HD (*Graham et al., 2006; Gu et al., 2009*). Most of the reported PTMs in HTT are localized in the four PEST domains, so called because they are amino acid sequences enriched in proline (P), glutamate (E), serine (S), and threonine (T) (*Warby et al., 2008*). For a complete list of HTT residues subjected to PTMs see **Table 1.1**.

HTT proteolysis. Huntingtin protein is subjected to proteolysis at several sites by a variety of proteases some of which are unknown. The proteolytic sites are in PEST domains that are mostly found in HTT disordered regions (**Figure 1.3**) (*Warby et al., 2008*). The first reported proteases able to cleave HTT are the caspases (*Goldberg et al., 1996; Wellington et al., 1998; Hermel et al., 2004*), followed by the calpains (*Gafni and Ellerby, 2002; Kim et al., 2001, 2003; Gafni et al., 2004*), cathepsins, in particular aspartyl proteases (*Lunkes et al., 2002; Kim et al., 2006*) and MMP10, a metalloproteinase (*Miller et al., 2010*).

HTT proteolysis is relevant in the protein stability and homeostasis. Furthermore, caspases activity has been associated to synaptic plasticity, memory, and learning (*Li et al., 2010*), it is therefore possible that proteolytic cleavages of HTT protein could be involved in the regulation of these processes. A recent study showed that the proteolysis of wild-type HTT may inactivate some of its functions through the production of toxic non-polyQ C-terminal fragments (*El‐Daher et al., 2015*). Nevertheless, no proteolytic events of wild-type HTT were reported in normal individuals. On the contrary, in HD brains a specific increase of the proteolytic activity was described with a consequent accumulation of small N-terminal HTT fragments. Importantly, the toxicity of these fragments is due to their translocation into the nucleus (*Benn et al., 2005; Graham et al., 2006; Saudou et al., 1998*). It has also been proposed that PTMs and proteolysis work in a cooperative manner to modulate several protein functions (*Tan et al., 2010*).

HTT interactors. Many studies aimed at discovering putative partners of HTT (*Harjes and Wanker, 2003*; *Goehler et al., 2004*; *Kaltenbach et al., 2007*; *Ratovitski et al., 2012*; *Culver et al., 2012*). These studies led to more than 350 wild-type HTT potential partners. In a few cases, the functional relationship between wild-type HTT and its interacting proteins has been established. For instance, HTT interacts with PSD95 (*Sun et al., 2001*), with Huntingtin-Associated Protein 1 (HAP1) and the p150 subunit of dynactin (*Gauthier et al., 2004*), with Mixed-Lineage Kinase 2 (MLK2) (*Marcora et al., 2003*), and with Huntingtin-Interacting Protein 1 (HIP1) (*Gervais et al., 2002*). Another relevant example is that wild-type HTT can promote the expression of BDNF by interacting through REST-interacting LIM domain protein (RILP) and HAP1 with the repressor element-1 transcription factor/neuron-restrictive silencer factor (REST/NRSF) in the cytoplasm; in this way, this complex does not translocate into the nucleus and is prevented from binding to the repressor element 1/neuron-restrictive silencer element (RE1/NRSE) located in the promoter of the BDNF gene and other neuronal genes (*Shimojo, 2008; Zuccato et al., 2003*). HTT large size and its stability, together with the high number of partners identified,

argue in favour of HTT being a scaffold protein. For example, HTT scaffolds Dynein/Dynactin to regulate several cellular processes such as vesicular transport, cell division and ciliogenesis (**Figure 1.4**) (*Saudou and Humbert, 2016*). Noteworthy, it has been shown that few proteins interact with the internal or the C-terminal part of HTT. This finding is in line with intramolecular interactions observed between the N-terminal and C-terminal regions (*El‐Daher et al., 2015; Li et al., 2006; Ochaba et al., 2014; Palidwor et al., 2009*). Indeed, as a result of these interactions, HTT can adopt a closed conformation and allow only a few limited interactions between the C-terminal region and other proteins.

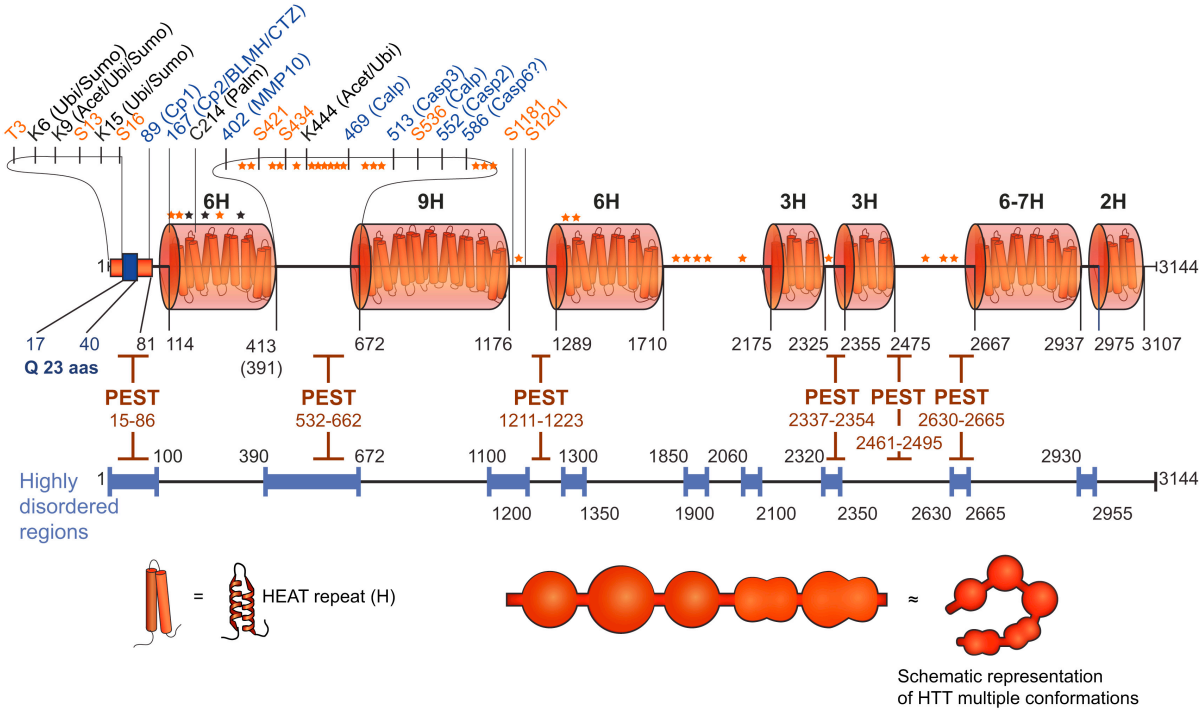


Figure 1.3. Cartoon of the human HTT protein sequence (from *Saudou and Humbert, 2016*)**.** Amino acid (aa) positions in orange: phosphorylation sites identified by mass spectrometry and further confirmed by other approaches. Amino acid positions in black: sites of indicated modifications. Amino acid positions in blue: cleavage sites. Cp1, cleavage site by unknown protease; Cp2/BLMH/CTZ, cleavage by Bleomycin hydrolase or cathepsin Z; Casp6? is a confirmed cleavage site: while initially identified as a caspase 6 site, caspase 6 may not be the cleaving protease *in vivo*. Orange and black stars indicate, respectively, phosphorylation and acetylation sites identified by mass spectrometry only with no further confirmations. Numbers arranged linearly on the 2D structure correspond to the limits of the indicated domains (HEAT, PEST, and Highly disordered regions). H indicates the number of predicted HEAT repeats organized in larger domains and PEST regions are proteolysis-sensitive domains. Highly disordered regions correspond to predicted disordered regions using PONDR, DisEMBL, and NORSp. Bottom: Schematic HTT with spheres that correspond to stretches of HEAT repeats. HTT could adopt multiple conformations according to *Seong et al., 2010*. Ubi, ubiquitin; Sumo, sumoyl; Acet, acetyl; Palm, palmitoyl; MMP10, metalloproteinase 10; Calp, calpain; Casp3/2/6, caspase 3/2/6.

Residue	Modification	Enzyme	Function	Analysis	References
Thr 3	Phos			MS/Ab	Aiken et al. (2009), Huang et al. (2015)
Lys 6	Ubi/Sumo			IV	Steffan et al. (2004)
Lys 9	Acet/Ubi/Sumo			MS/IV	Cong et al. (2011), Steffan et al. (2004)
Ser 13	Phos	IKK	K9/Ubi/Sumo	MS/Ab	Thompson et al. (2009)
Lys 15	Ubi/Sumo			IV	Steffan et al. (2004)
Ser 16	Phos	IKK	K9/Ubi/Sumo	MS/Ab	Thompson et al. (2009)
Ser 116	Phos			MS	Watkin et al. (2014)
Ser 120	Phos			MS	Watkin et al. (2014)
Lys 178	Acet			MS	Cong et al. (2011)
Cys 214	Palm			MS/IV-Palm	Yanai et al. (2006)
Lys 236	Acet			MS/Ab	Cong et al. (2011)
Thr 271	Phos			MS	Watkin et al. (2014)
Lys 345	Acet			MS/Ab	Cong et al. (2011)
Ser 417	Phos	Akt/PKA?		MS	Huang et al. (2015), Watkin et al. (2014)
Ser 419	Phos	Akt/PKA?		MS	Huang et al. (2015), Moritz et al. (2010)
Ser 421	Phos	Akt/SGK1/PP2B	Axonal transport	MS/Ab	Huang et al. (2015), Humbert et al. (2002), Moritz et al. (2010), Pardo et al. (2006), Rangone et al. (2004), Schilling et al. (2006), Watkin et al. (2014)
Ser 431	Phos			MS/Ab	Dong et al. (2012), Huang et al. (2015), Watkin et al. (2014)
Ser 432	Phos			MS/Ab	Dong et al. (2012), Huang et al. (2015)
Ser 434	Phos	Cdk5	Casp-3 site 513	Ab/MS	Huang et al. (2015), Luo et al. (2005), Watkin et al. (2014)
Ser 438	Phos			MS	Huang et al. (2015)
Lys 444	Acet/Ubi	CBP/HDAC1	autophagy	MS/Ab	Cong et al. (2011), Jeong et al. (2009)
Ser 457	Phos			MS	Watkin et al. (2014)
Ser 459	Phos			MS	Watkin et al. (2014)
Ser 461	Phos			MS	Watkin et al. (2014)
Ser 464	Phos			MS	Watkin et al. (2014)
Ser 465	Phos			MS	Watkin et al. (2014)
Ser 466	Phos			MS	Watkin et al. (2014)
Ser 487	Phos			MS	Watkin et al. (2014)
Thr 488	Phos			MS	Watkin et al. (2014)
Ser 491	Phos			MS	Watkin et al. (2014)
Ser 536	Phos	PKC?	Calp 536 site	MS	Schilling et al. (2006)
Ser 642	Phos	Akt?		MS	Huang et al. (2015), Moritz et al. (2010)
Ser 644	Phos			MS	Huang et al. (2015)
Ser 645	Phos			MS	Huang et al. (2015)
Ser 1181	Phos	Cdk5	axonal transport	MS/Ab	Anne et al. (2007), Huang et al. (2015), Schilling et al. (2006)
Ser 1197	Phos			MS	Huang et al. (2015)
Ser 1201	Phos	Cdk5	axonal transport	MS/Ab	Anne et al. (2007), Huang et al. (2015), Schilling et al, (2006)
Ser 1351	Phos			MS	Huang et al. (2015)
Tyr 1357	Phos			MS	Huang et al. (2015)
Ser 1866	Phos			MS	Huang et al. (2015)
Ser 1868	Phos			MS	Huang et al. (2015)
Thr 1872	Phos			MS	Huang et al. (2015)
Ser 1876	Phos			MS	Huang et al. (2015)
Ser 2076	Phos	ERK1?		MS	Schilling et al. (2006)
Thr 2337	Phos			MS	Huang et al. (2015)
Ser 2550	Phos			MS	Huang et al. (2015)
Ser 2653	Phos	ERK1?		MS	Huang et al. (2015), Schilling et al. (2006)
Ser 2657	Phos	GSK3?		MS	Huang et al. (2015), Schilling et al. (2006)

Table 1.1. HTT Post-translational Modifications (adapted from *Saudou and Humbert, 2016*)**.** The identification of the modifications was performed using mass spectrometry (MS), specific antibodies against the modified sites (Ab), and *in vitro* assays (IV; IV-Palm: in vitro palmitate assay) as indicated. Thr, threonine; Lys, lysine; Ser, serine; Cys, cysteine; Tyr, tyrosine; Phos, phosphorylation; Ubi, ubiquitination; Sumo, sumoylation; Acet, acetylation; Palm, palmitoylation.

1.1.4 HTT functions

The numerous physiological outputs of wild-type HTT functions are relevant both in the developing and in mature organism; these outputs at an integrated level (tissue or whole organism) may correspond to a limited set of functions at the molecular level. A current question in the field is whether the neurodegeneration observed in the adult HD brain may a consequence of developmental defects.
Here, we will address HTT functions at an integrated level by analysing its role in embryonic development, tissue maintenance and cell morphology, and cell survival. Some HTT functions have also been elucidated at the molecular level, such as the involvement of HTT in vesicular transport, cell division, ciliogenesis, endocytosis, vesicle recycling, and endosomal trafficking, autophagy, gene transcription regulation, and synaptic activity (*Saudou and Humbert, 2016; Liu and Zeitlin, 2017*).

HTT in development

Huntingtin was shown to be essential for embryonic development, since the inactivation of *HTT* gene is lethal in mice. In particular, HTT is critical for gastrulation (*Duyao et al., 1995; Nasir et al., 1995; Zeitlin et al., 1995*) and for CNS formation (*White et al., 1997; Auerbach et al., 2001*). Although the studies of the HTT functions *in vivo* during gastrulation are hindered by the early lethality of knock-out embryos, a role for HTT in neurogenesis is emerged, and severe alterations of the brain development were shown (*White et al., 1997*).

Gastrulation. The constitutive inactivation of the *HTT* gene by targeting exon 1 or 5 is lethal in mice between embryonic day 8.5 and 10.5; this stage is shortly after the onset of gastrulation and the formation of early nervous system (*Duyao et al., 1995; Nasir et al., 1995; Zeitlin et al., 1995*). This seems to be a direct consequence of an increased apoptosis in the embryonic ectoderm. Moreover, it has been shown that chimeric embryos derived by the injection of *HTT* null embryonic stem (ES) cells into wild-type host blastocysts are rescued from lethality (*Dragatsis et al., 1998*). In contrast, when wild-type ES cells are injected into *HTT* null blastocysts, the chimeric embryos die shortly after gastrulation, demonstrating that HTT acts in the extra-embryonic tissues, whereas the epiblast and its derivatives are affected secondarily (*Dragatsis et al., 1998*). More specifically, an altered nutritive function of the extra-embryonic tissues in the absence of HTT has been discovered (*Dragatsis et al., 1998*). In line with this observation, extra-embryonic tissues lacking HTT reveal defects in iron transport (*Dragatsis et al., 1998*). Coherent with this idea, *HTT* knock-down analyses in zebrafish model showed a deficiency in iron utilization and development (*Lumsden et al., 2007*).

Neural tube formation. HTT was found to be critical in the formation of nervous system. The first evidence came from an experiment in which mice expressing less than 50% wild-type HTT showed defects in the precursor of the epiblast, the structure that will give rise to the neural tube, and malformations of the cortex and striatum. Additionally, these mice died shortly after birth (*White et al., 1997; Auerbach et al., 2001*). In a further experiment it was shown that the reduction of HTT levels in zebrafish generates defects in the formation of most of the anterior regions of the neural plate, in the telencephalic progenitor cells and in the preplacodal tissue (*Henshall et al., 2009*). This shows the HTT is involved in neural tube formation at the post-gastrulation stage.
Evidence that HTT is involved in neurulation comes from experiments with *HTT* null mouse embryonic stem (mES) cells as well as *HTT* knock-down zebrafish embryo. In particular, wild-type ES cells subjected to neural differentiation protocol give rise to 3D structures named neuroepithelial rosettes (*Ying et al., 2003a*). These neural structures emerge at days 7-8 of the neural differentiation protocol and are known to recapitulate many aspects of neural tube development and early stages of the neurulation process (*Elkabetz et al., 2008; Abranches et al., 2009*). HTT-depleted mES cells are less capable to form neuroepithelial rosettes (*Lo Sardo et al., 2012*). Further investigations showed that the N-terminal fragment of HTT, but not its C-terminal portion, completely rescued neurulation in complementation assays performed in HTT-depleted ES cells (*Tartari et al.,*

2008; Lo Sardo et al., 2012). In addition to *in vitro* studies, experiments in zebrafish support the idea that HTT is implicated in neurulation. Indeed, embryos injected with *Htt*-morpholino showed an altered neural tube structure and defects in cellular adhesion and apicobasal polarity (*Lo Sardo et al., 2012*).

The pro-neurulation function of HTT is specific of the deuterostome group. In fact, complementation assays using HTT-depleted ES cells expressing heterologous HTT fragments showed that the ability to sustain neurulation was almost absent in *Dictyostelium* and *Drosophila*, subtle in sea urchin and *Ciona* (lower deuterostomes), whereas it was more evident with HTT fragments from the cephalochordate *Branchiostoma floridae* and even stronger with HTT from fish and mammals (*Lo Sardo et al., 2012*). Given that the amphioxus represents a critical step in the transition from invertebrates to vertebrates and in the acquisition of a more complex nervous system (*Holland et al., 2004*), the presence of pro-neurulation activity by HTT in this organism suggests a crucial role of this protein in the evolutionary dawn of the nervous system in chordates. The appearance of the HTT polyQ tract in echinoderms, and its subsequent stabilization in chordates, points in the direction of a putative implication of this tract in driving the transition toward vertebrate nervous system complexity.

One of the more characterized pathways by which HTT regulates neurulation includes two proteins implicated in the cell adhesion mechanism (*Lo Sardo et al., 2012*). The first protein is N-cadherin that is specifically required for the morphology of neural tube (*Detrick et al., 1990; Fujimori et al., 1990; Radice et al., 1997; Lele et al., 2002; Hong and Brewster, 2006; Chalasani and Brewster, 2011*); the second one is the metalloprotease ADAM10, an activator of N-cadherin protein (*Reiss et al., 2005*). In addition, it has been shown that synapse-associated protein 97 (SAP97) is responsible for driving ADAM10 to the postsynaptic membrane (*Marcello et al., 2007*). A recent work demonstrated that HTT binds ADAM10, and this interaction prevents the formation of ADAM10/SAP97 complex, regulating N-cadherin cleavage during neurulation (*Lo Sardo et al., 2012*). The absence of HTT increases the formation of ADAM10/SAP97 complex and causes an altered cleavage of N-cadherin, leading to neurulation defects as showed by the reduced ability to form *in vitro* neuroepithelial rosettes of the HTT-depleted mES cells, previously mentioned. It has been also found that in zebrafish embryos the morphological neural tube defects due to HTT loss-of-function is rescued by the pharmacological inhibition of ADAM10 (*Lo Sardo et al., 2012*). Another indication that validates the involvement of N-cadherin in the neurulation pathway comes from experiments where N-cadherin zebrafish mutants show neurulation defects (*Lele et al., 2002*). In particular, by analysing the expression patterns of the floor plate together with markers of apicobasal polarity, it has been shown that *HTT* knock-down in zebrafish embryos phenocopies N-cadherin zebrafish mutants (*Lo Sardo et al., 2012*). This finding confirms that HTT controls neuroepithelial cell adhesion and morphogenesis of the neural tube by influencing ADAM10/N-cadherin functionality.

Overall these results support the idea that HTT may regulate critical steps during the nervous system formation, particularly in neurulation, and that this function may have had an impact on the evolution of deuterostomes.

Cortical neurogenesis. HTT plays a role also during the later phases of brain development, in particular the protein is relevant for cortical neurogenesis. *In vivo* inactivation of huntingtin by RNA interference or by ablation of the *HTT* gene affects spindle orientation and cell fate of cortical progenitors of the ventricular zone in mouse embryos (*Godin et al., 2010*). More specifically, HTT was found on the spindle pole and, by recruiting dynein/dynactin complex and other molecular complexes (such as the NUclear Mitotic Apparatus, NUMA), it controls mitotic spindle orientation (*Godin et al., 2010*). This function is also present in *Drosophila*, indeed the specific disruption of *Drosophila* HTT in neuroblast precursors leads to spindle misorientation; suggesting that this function is conserved during evolution. Moreover, *Drosophila* HTT restores spindle misorientation in mammalian cells (*Godin et al., 2010*).

Cortical neuroblast migration. Another role of HTT regards the migration of cortical neuroblasts from the ventricular zone to the cortical plate (*Tong et al., 2011*). It was reported that the knock-down of HTT expression in neuroepithelial cells of neocortex results in disturbed cell migration, reduced proliferation, and increased cell

death. In the cerebellum, however, HTT knock-down results in cell death but not in perturbed migration. The timing of knock-down during early development is also an important variable. Based on these results, a spatial and temporal requirement for HTT expression in neural development is evident (*Tong et al., 2011*). Finally, the phenotypic abnormalities of impaired migration and cell death were only found in transfected cells while neighboring non-transfected cells showed no defects, implicating a non-cell autonomous mechanism (*Tong et al., 2011*).

Neural identity. Studies of mutant mice expressing low levels of HTT and analyses in chimeric mice generated by blastocyst injection of *HTT* null ES cells confirm that wild-type HTT plays an important role later in development as well, specifically in forebrain formation.
Early studies reported a specific role of HTT for the development of cortical and striatal neurons; in particular, HTT appears to be required for proliferation or survival of neurons in some brain regions during early development. Indeed, HTT-depleted cells preferentially colonize hypothalamus, midbrain, and hindbrain and are scarce in cortex, striatum, thalamus, and Purkinje cell layer of the cerebellum suggesting that neurons in these regions may strictly require HTT to develop and/or survive normal (*Reiner et al., 2001*). Further analyses of chimeric embryos at E12.5 showed ongoing degeneration of HTT-depleted cells specifically in the striatum, cortex, and thalamus, corroborating the view that neuroblasts in these areas need to synthesize HTT if they are to progress in development and differentiation (*Reiner et al., 2001, 2003*). Finally, it has been discovered that neurons, throughout the brain, require HTT for long-term survival (*Reiner et al., 2001*). These results are in line with the data collected using conditional *HTT* knock-out mice, in which the loss of HTT late in development or shortly after birth impacts the survival of cortical and striatal neurons in the adult brain (*Dragatsis et al., 2000*). Another mechanism by which HTT impacts on brain development is the regulation of motile cilium biogenesis (*Keryer et al., 2011*). In particular, conditional inactivation of the *HTT* gene in Wnt1 cell lineages results in congenital hydrocephalus, implicating HTT in the regulation of cerebral spinal fluid (CSF) homeostasis (*Dietrich et al., 2009; Keryer et al., 2011*).

HTT regulates tissue maintenance and cell morphology

In parallel to the role of HTT during brain development, recent studies implicate HTT is a regulator of tissue maintenance. Indeed, depletion of HTT, *in vivo*, results in a decreased pool and specification of basal and luminal progenitors in the mammary epithelium. This leads to altered mammary morphogenesis probably due to mitotic spindle misorientation (*Elias et al., 2014*). HTT is also relevant for the establishment of apical polarity and, consequently, the depletion of HTT from luminal cells *in vivo* alters mouse ductal morphogenesis together with impaired lumen formation (*Elias et al., 2015*).

Further evidences of HTT roles in the regulation of epithelial morphogenesis have been found. In particular, the previous described involvement of HTT in the cell adhesion mechanism, as showed by aberrant distribution of the tight junction protein zona occludens 1 (ZO1) in *HTT* null ES cells and *Htt* morpholino zebrafish embryos, is one example (*Lo Sardo et al., 2012*). Moreover, ES cells and neurons with lowered HTT levels have shown reduced amounts of mRNAs of adherence proteins (*Strehlow et al., 2007*). Consistent with these observations, mice with reduced HTT expression in testis show spermatogenesis defects that result in disorganized seminiferous tubules with fewer spermatocytes and round spermatids (*Dragatsis et al., 2000*). Another recent evidence reports that HTT is involved in the maintenance of tight junctions, in particular it has been shown that reduced amounts of HTT causes a greater epithelial-to-mesenchymal transition in mammary tumour cells (*Thion et al., 2015*).

Finally, HTT may also impact on histogenesis and organogenesis by regulating metabolism. *In vivo* experiments showed that mice overexpressing wild-type HTT have an increased body weight as well as enlarged heart, spleen, liver, kidneys, and lungs. These findings are the result of a HTT dose-dependent modulation of the insulin-like growth factor 1 (IGF-1) expression (*Pouladi et al., 2010; Van Raamsdonk et al., 2006*).

HTT in cell survival

Several *in vitro* and *in vivo* studies highlighted the pro-survival (antiapoptotic) properties of wild-type HTT. A first set of experiments showed that expression of wild-type HTT in cell lines and primary cultures of neurons protects against cell death. In particular, striatal cells overexpressing wild-type HTT were resistant to serum deprivation or exposure to 3-nitropropionic acid (a toxin that gives a similar pattern of HD neurodegeneration when injected into animals) (*Rigamonti et al., 2000, 2001*). Further analyses discovered that this antiapototic activity is embedded in the HTT N-terminus, since both full-length normal HTT and its 548 amino acid N-terminus were able to protect cells against cell death (*Rigamonti et al., 2000*). Interestingly, normal HTT is protective also when the cell death is induced by mutant HTT itself; indeed, it was shown that wild-type HTT can protect non-neuronal cell lines (*Ho et al., 2001*) and primary striatal neurons (*Leavitt et al., 2006*) from apoptosis resulting from mutant HTT. Conversely, depletion of HTT renders cells more vulnerable to cell death in line with an increased level of caspase-3 activity (*Zhang et al., 2006; Lo Sardo et al., 2012*). The antiapoptotic function of HTT seems to be conserved during evolution as the expression of heterologous HTT, including HTT from *Drosophila melanogaster* and *Dictyostelium discoideum*, in HTT-null mES cells exhibited anti-apoptotic activity.

These *in vitro* observations were supported by *in vivo* studies in mouse. Earlier analysis in a heterozygous *HTT* knock-out mouse model revealed apoptotic cell death in several areas, cognitive dysfunction, and behavioural abnormalities (*Nasir et al., 1995; O'Kusky et al., 1999*). Moreover, neuronal-specific inactivation of the *HTT* gene in the postnatal forebrain generated apoptotic cells in the hippocampus, cortex, and striatum, as well as a lack of axon fibers (*Dragatsis et al., 2000*). Analysis of a mouse model in which HTT is selectively deleted from adult mature cortical and hippocampal neurons demonstrated defects in both the survival and the dendritic arborization of newborn hippocampal neurons (*Pla et al., 2013*). These findings associate endogenous wild-type HTT with adult hippocampal neurogenesis. Furthermore, the controlled deletion of wild-type HTT had anxiogenic-like effects, suggesting that some of the mood disorders (very common among HD patients) may be caused by an alteration of normal HTT function in the hippocampus and cortex (*Pla et al., 2013*). Apoptotic cell death was also found in *Htt*-morpholino knock-down zebrafish embryos, in which an increased caspase-3 activity and a severe underdevelopment of the CNS were observed (*Diekmann et al., 2009*). More recently, a non-coding SNP (Single Nucleotide Polymorphism) variant in the *HTT* promoter was found (*Bečanović et al., 2015*). This genetic variant on the wild-type *HTT* allele reduces its expression, and is associated with earlier age of onset in HD patients (*Bečanović et al., 2015*), supporting the idea that HTT levels can impact cell survival.

The antiapoptotic/pro-survival activities of HTT are very pertinent to HD. Indeed, one of the most prominent histopathological features of post-mortem HD brain samples is the severe neuronal atrophy of several brain regions; this massive atrophy may cause a reduction of up to 30% in brain weight (*Rosas et al., 2008*). The cell death associated with the presence of mutant HTT is recapitulated in cellular and mouse models (*Gray et al., 2008; Saudou et al., 1998*).

The first attempts to decipher the molecular details of how normal HTT promotes survival focused on the apoptotic machinery. Wild-type HTT prevents the formation of a functional apoptosome complex and, consequently, blocks the activation of caspase-3 (*Rigamonti et al., 2000*) and caspase-9 (*Rigamonti et al., 2001*). HTT can also interact with active caspase-3 and inhibit its activity (*Zhang et al., 2006*). In addition, HTT blocks the formation of caspase-activating complex HIPPI, hampering pro-caspase-8 activation (*Hackam et al., 2000; Gervais et al., 2002)*. A further analysis indicated that HTT is a substrate for Akt, a serine/threonine-specific protein kinase that activates a pro-survival pathway by stimulating the expression of pro-survival genes and repressing pro-apoptotic genes (*Rangone et al., 2004*). Another mechanism by which HTT promotes cell survival is exemplified by the nature of the cortico-striatal connection. The striatum does not produce BDNF and depends almost exclusively on the BDNF delivered by cortico-striatal afferences (*Baquet et al., 2004*). In this context, HTT favours the transcription of BDNF (*Zuccato et al., 2001*) and promotes the axonal transport and delivery of vesicles containing BDNF to the cortico-striatal synapse (*Gauthier et al., 2004*). When BDNF is released at cortico-striatal synapses, it promotes endocytosis and the

retrograde transport of TrkB receptors (Tropomyosin receptor kinase B) to striatal cell bodies where they activate survival signalling (*Liot et al., 2013*). TrkB binds to and colocalizes with HTT and dynein. It has been shown that silencing HTT reduces vesicular transport of TrkB in striatal neurons and, additionally, that HTT polyQ expansion in HD alters the binding of TrkB-containing vesicles to microtubules and reduces axonal transport (*Liot et al., 2013*). HTT, thus, ensures both the anterograde delivery of BDNF to the cortico-striatal synapse and also the retrograde transport of BDNF-TrkB endosomes along striatal dendrites towards the cell bodies. Therefore, given the central role of HTT in these processes and, consequently, in the survival of striatal neurons, it is not surprising that these neurons are the first to degenerate in HD.

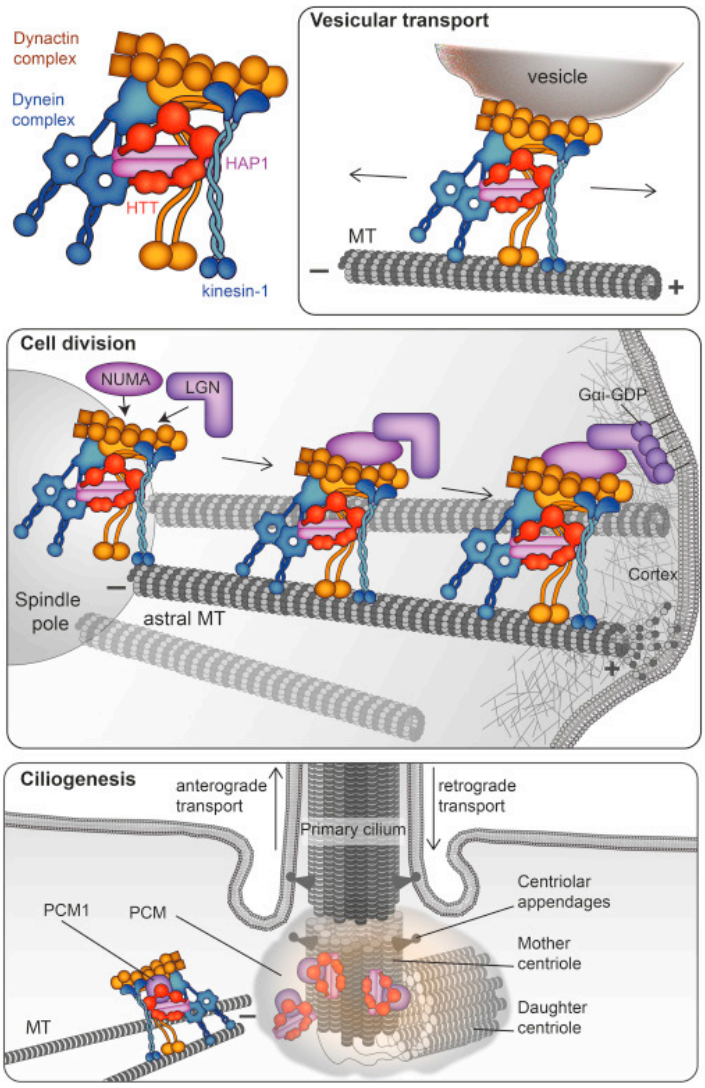


Figure 1.4. Huntingtin scaffolds Dynein/Dynactin to regulate several cellular processes (from *Saudou and Humbert, 2016*). Top: HTT controls the transport of organelles, in both anterograde and retrograde directions, and in axons and dendrites within neurons. Middle: During mitosis, HTT is important for spindle pole assembly and also regulates the kinesin1-dependent trafficking of dynein/dynactin/NUMA/LGN to the cell cortex. Bottom: HTT mediates the dynein/dynactin/HAP1-dependent transport of proteins to the pericentriolar material, including PCM1 protein that is required for ciliogenesis. MT, microtubules; PCM, pericentriolar material.

1.2 CAG trinucleotide repeats

1.2.1 Simple sequence repeats and their biological meaning

The HTT CAG repeats are one example of a broader group of tandem repeats, also called simple sequence repeats (SSRs), present in the human genome. These SSRs have been first described in the early '80s and are present in many eukaryotic genomes (from yeast to humans) (*Hamada et al., 1982*). In addition, many human genes possess multiple SSRs. These repeats have a propensity for "slippage mutations", which increase or decrease the number of repeats without altering the content of the sequence (*Richards and Sutherland, 1994*). The abundant polymorphism, which results from repeated slippage events, has become the basis for DNA fingerprinting, lineage analysis and gene mapping. This mutability, together with the apparent lack of information contained in such repetitive "genetic stutters", appeared once to preclude any possibility of critical function for the SSRs. Indeed, much of the literature on SSR polymorphism has assumed (*Ellegren, 2004*) that SSRs are genetic 'junk,' supplying only 'neutral' variation with no appreciable effect on phenotype. On the contrary, since Hamada and colleagues first established that altering the number of repeated dinucleotides could affect gene activity (*Hamada et al., 1982*) and according to their genome localization (gene promoters, introns, and exons), many examples of repeat-number-related effects came to light.

Interestingly, at least 20 different neurological disorders (such as HD, Spinocerebellar ataxia, and Myotonic dystrophy) are caused by expanded SSRs (reviewed in *Nithianantharajah and Hannan, 2007; Orr and Zoghbi, 2007*). Such diseases are frequently characterized by "genetic anticipation", an hereditary tendency toward further expansion of pathological repeat alleles in each generation, leading to earlier onset and accelerated disease progression in subsequent generations. In addition to their effects at pathological level, SSR alleles can also contribute to normal variation in brain and behavioural traits (*Fondon et al., 2008*). Indeed, many studies implicate SSRs not just in diseases but also in circadian rhythmicity, socio-sexual interaction, aggression, cognition and personality. SSRs have been suggested also to affect neuronal differentiation, brain development and even behavioural evolution (*Fondon et al., 2008*). In particular, SSR polymorphism contributes to normal quantitative genetic variation in traits ranging from cell adhesion in yeast to skeletal morphology in dogs (*Kashi et al., 1997; Kashi and King, 2006; Fondon et al., 2008*). Three best-studied examples of the relationships among SSR polymorphism, phenotype and molecular function involve three genes influencing animal behaviour: *PER* gene in *Drosophila melanogaster*, *AVPR1A* gene in *Microtus sp.* (field mice) and *SLC6A4* gene in primates.

First, *Drosophila* PER protein contains a tract of threonine-glycine and serine-glycine dipeptide repeats encoded by a hexanucleotidic SSR in exon 5. It has been shown that flies with 20 repeats resulted in reduced sensitivity of circadian rhythm to temperature fluctuations compared to those with a shorter 17 repeats (*Sawyer et al., 1997*). Moreover, in Europe, northern Africa and Australia 20-repeat alleles were more frequent in flies sampled at higher latitudes, suggesting that natural selection discriminates among alleles of different lengths based on their ability to compensate the temperature in regions exposed to greater temperature fluctuations (*Kyriacou et al., 2008*). Interspecies comparisons, including gene transplantation experiments, also support a functional role of the *PER* SSRs; in particular, these studies revealed a coevolutionary compensation between the effect of SSRs and flanking sequence variants (*Peixoto et al., 1998*).

Secondly, several converging lines of evidence identify functional SSR variation in the 5'- regulatory region of the *AVPR1A* (Arginine Vasopressin Receptor 1A) gene in *Microtus ochrogaster* (prairie voles) as a significant contributor to social behaviour such as selective partner preference and offspring care (*Hammock and Young, 2005*). In particular, comparison between prairie voles whose parents carried longer than average SSR alleles and those whose parents had shorter alleles revealed differences in V1aR (Vasopressin receptor 1A) protein distribution in the brain as well as alterations in vasopressin-dependent social behaviour (*Hammock and Young, 2005*). Additionally, mice that are transgenic for the prairie vole variant of the gene

have a neuroanatomical pattern of receptor binding that is similar to that of the prairie vole, and exhibit increased affiliative behaviour after injection with arginine vasopressin. These data indicate that the pattern of *AVPR1A* gene expression in the brain may be functionally associated with species-specific social behaviours in mammals (*Young et al., 1999*). Indeed, whereas SSR length correlates with brain and behavioural traits among individual prairie voles, phylogenetic analysis of other vole species has established that SSR in *AVPR1A* gene is not by itself a reliable marker for complex social behaviour which depends on multiple brain circuits, environmental context and many genes (*Hammock and Young, 2005; Fink et al., 2006*).

Thirdly, one of the best examples in humans involves a polymorphic SSR (designated 5-HTTLPR) located in the upstream promoter region of the serotonin transporter gene, *SLC6A4*. A classical study showed that the shorter of two common alleles, which yields lower levels of gene expression in transcription reporter assays, was correlated with increased neuroticism, tension and harm avoidance (*Lesch et al., 1996*). However, many subsequent studies have failed to corroborate those associations or have otherwise presented conflicting findings (*Sen et al., 2004; Munafò et al., 2005*). Additionally, individuals carrying two long alleles of this SSR appear to be somewhat protected from the negative effects of significant life stress (*Caspi et al., 2003*). Notably, gene-environment interactions are an important consideration in human association studies and can be used to clarify conflicting findings. Furthermore, a similar (but not identical) polymorphism (rh5-HTTLPR) was discovered in the rhesus macaque (*Macaca mulatta*) serotonin transporter gene. Like in humans, the shorter of these polymorphism alleles has a lower gene expression *in vitro* and was associated with reduced levels of the serotonin metabolite 5-HIAA (*Bennett et al., 2002*), increased distress during assessment (*Champoux et al., 2002*), and greater inhibition by anxiety-provoking novel environments (*Bethea et al., 2004*). Curiously, early dispersal of male offspring from their natal group has also been associated with the shorter allele (*Trefilov et al., 2000*). For a more complete list of candidate-gene sites where SSR polymorphisms were associated with human behaviour see **Table 1.2**.

Another feature of the SSR repeats is the somatic mosaicism found in several neurodegenerative repeat-expansion diseases (*Telenius et al., 1994; Ueno et al., 1995; Tanaka et al., 1996; Gonitel et al., 2008*). However, numerous observations suggest the presence of SSR somatic variations also during normal brain development where they can exert a direct impact on gene function (*Hamada et al., 1984; Trifonov, 1989; Gerber et al., 1994; Kashi et al., 1997; King et al., 1997; Comings, 1998; King and Soller, 1999; Fondon and Garner, 2004; Kashi and King, 2006*). In spinocerebellar ataxia 2 (SCA2), it is not the length of the inherited allele *per se* but the susceptibility of the SSRs to hyperexpansion, and the subsequent somatic variation, which are responsible of the pathogenicity. This pathogenicity can be reduced by the stabilizing influence of repeat interruptions as shown by the discovery of individuals who possessed expanded alleles but who never developed the symptoms of this otherwise fully penetrant disease. Closer investigations revealed the presence of silent slippage-suppressing CAA interruptions within the CAG repeat sequence (*Choudhry et al., 2001*). Additional examples of the stabilization of otherwise pathological repeat alleles by repeat interruptions have been reported (*Sobczak and Krzyzosiak, 2004; Mulvihill et al., 2005; Matsuura et al., 2006*). Interestingly, it was also found that the presence of interruptions in SCA10 repeat expansion indicates a significant risk for the epilepsy phenotype (*McFarland et al., 2014*), suggesting that, in some instances, repeat interruptions can contribute to additional phenotypes beyond the typical disease repertoire. As mentioned above, the mitotic slippage could also produce somatic variation in normal conditions. Non-pathological somatic SSR mutations occur, indeed, during several processes of organogenesis. For example, the SSR slippage mutations in the coding region of the melanocortin receptor gene during skin development cause black spotting in domestic red pigs (*Kijas et al., 2001*). Another example, is represented by the somatic hypermutation of antibody genes (*Beale and Iber, 2006*). Finally, the enrichment of SSRs in genes involved in neuronal differentiation and function (*Karlin et al., 2002*) raises the intriguing possibility that somatic mutation of SSRs might be a normal, perhaps even essential, component of brain development (*Nithianantharajah and Hannan, 2007*).

Overall the characteristics of SSRs have prompted speculation that their mutability could play an important and potentially beneficial role in evolution (*Kashi et al., 1997; King et al., 1997; King and Soller, 1999; Fondon and Garner, 2004; Kashi and King, 2006; King and Kashi, 2007; Kovtun and McMurray, 2008; King,*

2012). Thus, the triplet repeat-expansion diseases represent only the pathological extreme of a much more general mutational process, which also contributes to normal brain function and development.

Gene	SSR type and motif	Affected trait	Evidence
Serotonin transporter SERT (SLC6A4), 5-HTTLPR	Noncoding, 5′ promoter; 44 base-pair motif	Anxiety-related traits	*In vitro* and *in vivo* assays; inconsistent association studies; moderated by life stress; similar results in rhesus macaque
Serotonin transporter SERT (SLC6A4)	Noncoding, intron; 17 base-pair motif	Susceptibility to bipolar disorder, response of SERT to lithium	Association; *in vitro* assays
Dopamine receptor DRD4	Coding; 48 base-pair motif	Novelty-seeking behaviours	Association
Dopamine transporter DAT1 (SLC6A3)	Noncoding, 3′ UTR; 40 base-pair motif	Attention deficit hyperactivity disorder; episodic memory formation	*In vitro* and *in vivo* assays; association
DNA-binding protein Jarid2	Coding; tetranucleotide motif	Schizophrenia	Association
Androgen receptor	Coding; CAG repeat	Cognitive function	Association
α2b-adrenoceptor ADRA2B	Coding; imperfect glutamic acid repeat	Emotional memory	Association; *in vitro* assays
Arginine vasopressin receptor AVPR1a	Noncoding, upstream; dinucleotide repeats	Altruism, other social behaviours	Association; *post mortem* gene expression

Table 1.2. Several aspects of human behaviour are associated with repeat-number variation (adapted from *Fondon et al., 2008*)**.**

1.2.2 *HTT* CAG repeats variability in human population

The discovery of the gene responsible of HD allowed many researchers to focus on its CAG repeat tract, not only at the pathological level but also in normal conditions. Early epidemiological studies show that *HTT* CAG repeats in the human gene are not in equilibrium but are subjected to an ongoing process of expansion. The CAGs in the normal human population vary between 9 and 35 repetitions with the most common CAG size being 17 repeats (*Kremer et al., 1994*). Moreover, the degree of spread of the CAG repeats is biased toward the high CAG repeats (*Rubinsztein et al., 1994*).

More recent findings indicate that pathological *HTT* alleles continually mutate from a pool of intermediate alleles (IAs), which are healthy genes with a number of repetitions between 27 and 35 (*Goldberg et al., 1993; ACMG/ASHG HD Genetic Testing Working Group, 1998; Costa et al., 2003*). IAs were first identified in unaffected family members, parents and siblings, of individuals with apparently *de novo* HD mutations. *De novo* mutations for HD result from CAG repeat instability, which expands an intermediate allele into the HD range (*Goldberg et al., 1993; Myers et al., 1993; Almqvist et al., 2001*). One out of 17 people carry this intermediate allele (*Semaka et al., 2013*). The factors currently known to influence the risk of intermediate allele instability include the CAG repeat size, the sex and age of the transmitting parent, the family history and the HD gene sequence and haplotype (*Semaka et al., 2006*). In particular, the finding that new mutations for HD arise from IAs (27-35 CAGs) and not from normal alleles (<27 CAGs) supports a role for CAG repeat size in repeats instability (*Chong et al., 1997*). Additionally, it has been shown that expanded IAs are preferentially

transmitted by males with advanced paternal age (mean 36.7 years) (*Goldberg et al., 1993*). Conversely, there are no documented cases of maternal intermediate allele expansion resulting in an affected offspring (*Kremer et al., 1995*). However, the fact that substantial expansions of maternally transmitted alleles with as few as 36 repeats have occurred provides evidence that expansions of large maternal intermediate alleles into the disease range are theoretically possible (*Laccone and Christian, 2000*). Spermatogenesis has been postulated to account for the observed difference between male and female germline repeat expansion (*Kremer et al., 1995; Telenius et al., 1995*). During spermatogenesis, there are frequent rounds of DNA replication throughout adult life; on the contrary oocytes form prior to birth. It was speculated that numerous rounds of DNA replication in spermatogenesis increases the likelihood of CAG repeat expansion in paternal transmission, as this process may provide a greater opportunity for a large expansion to occur (*Goldberg et al., 1993; Chong et al., 1997; Pearson et al., 2005*). Alternatively, during spermatogenesis there may be a higher opportunity to accumulate numerous small expansions that incrementally increase the CAG size into the HD range (*Goldberg et al., 1993; Chong et al., 1997; Pearson et al., 2005*). Furthermore, the context of family history allows to categorize intermediate alleles as either general population intermediate alleles (GP:IA) or new mutation intermediate alleles (NM:IA). GP:IAs are randomly ascertained from the general population and are coincidentally identified within the context of a family history of HD. Conversely, NM:IAs are ascertained from new HD mutation families. Familial transmission studies and examination of the CAG size heterogeneity in sperm analyses showed that, while IA repeat expansion into the HD range is more common in new mutation families, the likelihood that offspring of GP:IA carriers inherit a CAG size in the disease expanded range is extremely low, if not negligible (*Goldberg et al., 1995; Chong et al., 1997; Kelly et al., 1999*). Another important factor that influences the risk of intermediate allele expansion is the genetic variability near its repeat tract (*cis*-elements) (*Almqvist et al., 1994; Cleary et al., 2002; Cleary and Pearson, 2003, 2005; Pearson et al., 2005*). Accordingly, detailed haplotypes using numerous single nucleotide polymorphisms (SNPs) located across the *HTT* gene were constructed (*Warby et al., 2009b*). In particular, these findings suggest that in Caucasians, CAG repeat expansion occurs primarily in two haplogroup variants (named A1 and A2 haplotypes). Thus, one can hypothesize that IAs found on these high-risk haplotypes might be more prone to repeat instability. The current data also argue, and further support, that *cis*-elements have a crucial predisposing influence on CAG instability in *HTT* (*Warby et al., 2009b*). Interestingly, a significant correlation was found between the CAG repeat length of the maternal and paternal allele in the *HTT* gene among healthy subjects, suggesting an assortative mating (*Nopoulos et al., 2011a*). It was suggested that this might be among the mechanisms at the origin of the evolutionary pressure toward longer CAG alleles reported in the human population.

Several human studies suggest “positive effects” in subjects carrying more CAG repeats in *HTT* gene. One study conducted on 278 normal subjects revealed an increase of grey matter within the pallidum with increasing long CAG repeats in the normal range (*Mühlau et al., 2012*), concluding that CAG size influences normal brain structure. Secondly, pre-manifest adult HD patients having a CAG length in the pathological range, but that are distant from the expected age of onset, appear to perform better than control subjects in perceptual sensitivities tests, with a clear benefit from long CAGs in the gene (*Beste et al., 2012*). Thirdly, an unpublished study by P. Nopoulos on pre-manifest HD positive children (12 years of age) carrying up to 44 CAG repeats show better visual and motor skills compared to normal healthy controls (oral communication, CHDI Therapeutic Conference 2013). Finally, HD patients carrying an increased number of repeats in the normal allele exhibited less severe cognitive symptoms (*Aziz et al., 2009*). All these studies will require further independent validation. In particular, one recent study has disputed the conclusion by Aziz and his group (*Lee et al., 2012*), however only motor and not cognitive deficits were analysed.

Collectively, these lines of evidence are in agreement with the hypothesis that the different number of CAGs in the *HTT* gene may act as a genetic modulator of brain function and/or behaviour and correlates with aspects of cognitive performance in a length-dependent manner from low to borderline CAG in the normal range.

1.3 ESCs: Cell-based platform to study HTT functions

1.3.1 Embryonic Stem Cells (ESCs)

In this thesis work we investigated the role of the polyQ tract in HTT neural function by exploiting *in vitro* differentiation of mES cells towards a neural fate that represents a powerful and easy tool to study neurulation *in vitro*. Therefore, in this section I will illustrate the main features of mouse Embryonic Stem Cells (mESCs), describing the extraordinary properties of these cells in being able to give rise to differentiated derivatives of the three primary germ layers and highlighting also some differences with human Embryonic Stem Cells (hESCs).

The three main features that identify the ESCs are: pluripotency (the capability of differentiating into tissues derived from all three germ layers), self-renewal (maintenance of an undifferentiated state) and unlimited proliferation (*Smith, 2001; Niwa, 2007; Nichols and Smith, 2011*). The capacity of stem cells to differentiate into specialized cell types and be able to give rise to any mature cell type is referred to as potency. Potency of the stem cell specifies the differentiation potential and, according to such potential, it is possible to classify: totipotent, multipotent, pluripotent, oligopotent, and unipotent cells.

(i) Totipotent cells can differentiate into embryonic and extraembryonic cell types. Such cells can construct a complete, viable organism. These cells are produced from the fusion of an egg and sperm cell. The only totipotent cells are the fertilized egg and the cells produced by the first few divisions of the fertilized egg are also totipotent. Totipotent cells give rise to somatic stem/progenitor cells and primitive germline stem cells (*Weissman, 2000*).

(ii) Pluripotent stem cells are the descendants of totipotent cells and can differentiate into all cells derived from any of the three germ layers, but they are not able to generate the extraembryonic trophoblast. These pluripotent cells are characterized by self-renewal and a differentiation potential for all cell types of the adult organism (*Horie et al., 2011*). Embryonic stem cells fall under this category, as they are able to generate all cell types of the body *in vivo* and in culture.

(iii) Multipotent stem cells can differentiate into a limited number of cells, only those of a closely related family of cells. For example, the bone marrow contains adult hematopoietic stem cells, multipotent stem cells that give rise to all the cells of the blood but not to other types of cells.

(iv) Oligopotent stem cells can differentiate into only a few cells, such as lymphoid or myeloid stem cells. The corneal epithelium is a squamous epithelium (*Singh et al., 2010*) that is constantly renewing and contains oligopotent stem cells (*Majo et al., 2008*).

(v) Unipotent cells can produce only one cell type, their own, but have the property of self-renewal, which distinguishes them from non-stem cells. Most epithelial tissues self-renew throughout adult life due to the presence of unipotent progenitor cells (*Blanpain et al., 2007*). Spermatogonial cells are an example of unipotent progenitor cells, as they can only give rise to sperm cells (*Jaenisch and Young, 2008*).

The first mES cells were successfully derived directly from mouse blastocysts by Evans and Kaufman from the University of Cambridge (*Evans and Kaufman, 1981*) and by Martin from the University of California, in 1981 (*Martin, 1981*). Some years later, also human ES cell lines were derived (*Thomson et al., 1998*). The culture conditions for retaining pluripotency included a feeder layer of mitotically inactivated mouse fibroblasts and fetal calf serum that was batch-selected to promote proliferation and retain an undifferentiated phenotype. The initiative to derive pluripotent cell lines from early embryos was inspired by studies on teratocarcinoma cells (*Kleinsmith and Pierce, 1964*).

Mouse ES cells are derived directly from the inner cell mass (ICM) of the blastocyst between embryonic day (E) 3.5 and E4.5. At E4.5, the blastocyst contains three cell types: the trophectoderm, the hypoblast and the epiblast. While the trophectoderm and the hypoblast contribute to extraembryonic tissues, the epiblast gives rise to all cell types of the developing embryo (*Osorno and Chambers, 2011*). This pre-implantation epiblast cell population contains cells that have the ability to differentiate into derivatives of all three somatic lineages and the germline. Generally, ES cells require extrinsic growth factors to maintain their pluripotency in culture.

These extrinsic growth factors act on different signalling pathways to regulate intrinsic transcription factor networks to sustain ES cells in the undifferentiated state. The main pathways are: LIF/JAK/STAT3, BMP signalling, TGF-β/activin/nodal signalling, and FGF/MEK signalling (*Han et al., 2013*) (**Figure 1.5**).

Leukaemia inhibitory factor (LIF) (*Hirai et al., 2011*) and bone morphogenetic protein 4 (BMP4) (*Qi et al., 2004*). were found to be necessary for the retention of mES cells in a pluripotent state *in vitro*. Upon LIF binding, the LIF receptor (LIFR) recruits the transmembrane glycoprotein 130 (gp130) to form a heterodimer which subsequently activates Janus kinase (JAK) through transphosphorylation (*Niwa et al., 1998*). Activated JAK then phosphorylates gp130, creating a docking site to bind the SH2 domain of Signal Transducers and Activators of Transcription 3 (STAT3) (*Ihle and Kerr, 1995; Stahl et al., 1995; Hemmann et al., 1996; Gerhartz et al., 1996*). Once STAT3 binds to the gp130 docking site, JAK then phosphorylates the recruited STAT3. Phosphorylated STAT3 forms a homodimer, which subsequently translocates into the nucleus, where it binds to gene enhancers to regulate target gene expression (*Auernhammer and Melmed, 2000; Reich and Liu, 2006; Chen et al., 2008*). Although the LIF/JAK/STAT3 pathway has been well documented to maintain pluripotency of mouse ES cells in the presence of serum, the mechanisms by which activated STAT3 functions in this regard are poorly understood. A more recent study identified 718 STAT3-bound genomic sites that were co-occupied by pluripotency transcription markers (Oct4, Sox2 and Nanog) by using chromatin immunoprecipitation sequencing (ChIP-seq) (*Chen et al., 2008*). In addition, it was demonstrated that knocking down STAT3-target genes induces activation of endodermal and mesodermal genes, supporting the conclusion that STAT3 prevents mESC differentiation by suppressing lineage-specific genes (*Bourillot et al., 2009*). In mES cells, LIF can substitute MEF feeder layers in maintaining pluripotency in the presence of animal serum. However, in serum-free cultures, LIF is insufficient to block neural differentiation and maintain pluripotency. Interestingly, the LIF receptor and gp130 are also expressed in hESCs; however, LIF is unable to maintain the pluripotent state of hESCs, suggesting that mES and hES cells require distinct signalling mechanisms to regulate their pluripotency (*Dahéron et al., 2004*).

Bone Morphogenetic Protein (BMP) is a member of the TGF-β superfamily (*Sebald et al., 2004*). BMP ligands bind to BMP receptors (BMPRs), the activated receptors phosphorylate BMP-responsive SMAD1/5/8 (Small Mother Against Decapentaplegic) molecules, which subsequently form a complex with SMAD4 and translocate into nucleus to regulate target gene expression. In particular, SMAD complex activates inhibitors of differentiation (Id) genes, which block neural differentiation by antagonizing neurogenic transcription factors as well as inhibiting ERK (extracellular signal–regulated kinase) proteins (*Ying et al., 2003b*). In mES cells BMP was shown to be able to replace serum (*Ying et al., 2003b*), therefore, exogenous LIF in combination with BMP4 proteins can maintain the pluripotency of mouse ES cells in the absence of MEFs and serum. On the contrary, BMP was shown to promote human ES cells differentiation to trophoblasts, while inhibiting BMP signalling with the BMP antagonist, Noggin, sustained the undifferentiated state of human ES cells (*Xu et al., 2002, 2005*). In line with these findings, dorsomorphin and DMH1, BMP inhibitors, were shown to promote long-term self-renewal and pluripotency of human ES cells, presumably by inhibiting BMP induced extraembryonic lineage differentiation (*Yu et al., 2008; Hao et al., 2010, 2011; Gonzalez et al., 2011*).

The TGF-β/activin/nodal pathway consists of molecules belonging to the TGF-β superfamily, which are able to sustain human ES self-renewal and pluripotency (*James et al., 2005; Vallier et al., 2005; Xiao et al., 2006*). The molecular mechanisms of this pathway implicate the activation of SMAD2/3, which induces (in combination with SMAD4) the expression of the pluripotent transcription factor Nanog (*Xu et al., 2008*). In contrast to its important role in maintaining hESC pluripotency, the TGF-β/Activin/Nodal signalling is not essential for pluripotency of mESCs. Although this pathway was shown to be active in undifferentiated mESCs as assessed by phosphorylation of SMAD2/3, inhibition of SMAD2/3 phosphorylation by the inhibitor SB431542 had no effect on the undifferentiated state of mESCs (*James et al., 2005*). However, the TGF-β/Activin/Nodal signalling may play a role in mESC proliferation. Indeed, a recent study showed that

inhibition of TGF-β/Activin/Nodal signalling by SMAD7 or SB431542 dramatically decreased mESC proliferation without effect on their pluripotency (*Ogawa et al., 2007*).

Finally, another crucial pathway for ESC pluripotency maintenance is FGF/MEK signalling. Although the role of exogenous FGFs (Fibroblast Growth Factors) in human ES cell has been known since a long time, the molecular mechanisms by which they function remain unclear. FGFs signal by binding to FGF receptors (FGFRs) activates multiple signalling cascades, including Mitogen-Activated Protein Kinases (MAPKs), the Janus kinase/signal transducer and activator of transcription (JAK/STAT), phosphatidylinositol 3-kinase (PI3K) and phosphoinositide phospholipase C (PLCg) pathway (*Dailey et al., 2005*). However, several studies have highlighted the FGFs contribution to the maintenance of human ES cells mainly through the FGF/MEK pathway (*Kang et al., 2005; Li et al., 2007*). FGF signalling in mESCs has also been extensively investigated. Mouse ESCs genetically deficient in FGF4 and extracellular-signal regulated kinase 2 (ERK2) differentiate inefficiently. These results can be reproduced using inhibitors of FGF receptor and ERK, suggesting that the blockage of the FGF/MEK signalling pathway may promote mESC pluripotency (*Burdon et al., 1999; Kunath et al., 2007; Stavridis et al., 2007*). Although inhibition of the FGF/MEK pathway can attenuate ES cell differentiation, it is insufficient to support mES cells self-renewal. Combination of the MEK inhibitor PD0325901 with the Glycogen synthase kinase-3 (GSK-3) inhibitor CHIR99021 (known as 2i) can efficiently sustain the pluripotency of mouse ES cells in the absence of exogenous cytokines (*Ying et al., 2008; Smith and Ying, 2015*). Several groups demonstrated that improvement of mouse ES cell pluripotency by inhibition of GSK-3 occurred via Wnt/β-catenin signalling, whereas many others argued that GSK3 was likely to exert β-catenin independent effects in ES cells (*Aubert et al., 2002; Kielman et al., 2002; Sato et al., 2004; Ogawa et al., 2006; Pereira et al., 2006; Takao et al., 2007; Ying et al., 2008; Wray et al., 2010*).

Overall these indications reveal that even if both human and mouse ESCs are derived from blastocyst-stage embryos, maintenance of their pluripotency requires different biological signals. In general, mouse ESCs maintain their pluripotency by activating LIF/STAT3 and BMP signalling, while human ESCs require TGF-β/Activin/Nodal and FGF/MEK pathways. Interestingly, several pathways, such as BMP and FGF/MEK, have completely opposite effects on maintaining the pluripotency in these two different models. Indeed, activation of BMP signalling and inhibition of the FGF/MEK pathway promote mouse ES self-renewal, whereas inhibition of BMP signalling and activation of FGF/MEK pathway sustain human ESC pluripotency. These distinct signalling effects on pluripotency may reflect intrinsic differences between mouse and human ESCs. More recent studies have established that conventional human ES cells do not represent the “ground or naïve state” of stemness, but rather a more developmentally mature “primed state” resembling mouse epiblast stem cells (mEpiSCs) found in the post-implantation, pre-gastrulation stage of embryos (*Tesar et al., 2007; Silva et al., 2008, 2009; Bao et al., 2009; Nichols and Smith, 2009; Hanna et al., 2010*). Indeed, conventional hESCs exhibit numerous similarities to the mouse EpiSCs (E5.5-E7.5) over mouse ESCs (E3.5) (see **Table 1.3**).

Pluripotency and self-renewal efficiency is governed by a gene regulatory network centred around three main transcription factors: Oct4, Sox2 and Nanog (*Chambers and Tomlinson, 2009*). In 1998, Austin Smith's group clarified the importance of Oct4 in the maintenance of a pluripotent state (*Nichols et al., 1998*). Oct4 is expressed in pluripotent cells and is specifically required for cells that become allocated to the interior of the blastocyst to acquire a pluripotent identity. The expression level of Oct4 is a critical determinant of the phenotype of ES cells and derivatives. In line with the *in vivo* phenotype, deletion of Oct4 results in differentiation of cells to a trophectodermal type. Sox2 is a member of the Sry-related HMG box family of transcription factors that interact with DNA through binding to the minor groove. Sox2 shares many of the same DNA targets as Oct4, with many of the characterized target sites being composites of the non-palindromic Oct/Sox recognition sequences (*Masui et al., 2007; Chambers and Tomlinson, 2009*). Nanog is a homeobox-containing transcription factor whose DNA-binding domain adopts a typical three helix structure that binds to the TAAT DNA sequence (*Chambers et al., 2003, 2007*). *In vivo* analysis of Nanog-null embryos indicates that the latter is essential for the specification of the pluripotent epiblast similarly to the requirement

for Oct4. Genomic studies revealed that Oct4, Sox2, and Nanog frequently bind to the same regulatory regions in undifferentiated mouse and human ESCs, and that these binding sites are often in close proximity to one another (*Loh et al., 2006; Tesar et al., 2007; Mathur et al., 2008; Sharov et al., 2008*). These results indicate that Oct4, Sox2, and Nanog may physically interact with each other and, in some cases, coordinately regulate target genes. Additionally, it was reported that these combinatorial Oct4/Sox2/Nanog binding sites were more conserved between mouse and human ES cells compare to those individual (*Boyer et al., 2005; Loh et al., 2006; Wang et al., 2006; Masui et al., 2007; Göke et al., 2011*).

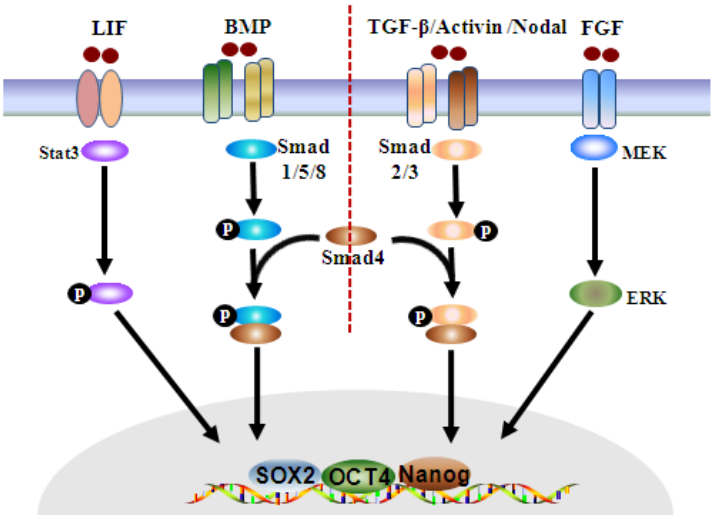


Figure 1.5. Signalling pathways required to support pluripotency in mouse/human ES cells (from *Han et al., 2013*). Exogenous growth factors signal through distinct signalling pathways to regulate transcription factors for ES cell pluripotency.

Property	mESCs	mEpiSCs	hESCs	hiPSCs
Morphology	domed	flattened	flattened	flattened
Clonogenicity	high (single cells)	low (clumps)	low (clumps)	low (clumps)
Response to LIF/Stat3	self-renewal	none	none	none
Response to Activin/bFGF	differentiation	self-renewal	self-renewal	self-renewal
Response to BMP	self-renewal	differentiation	differentiation	differentiation
XX status	XaXa	XaXi	XaXi	XaXi
Teratoma	yes	yes	yes	yes
Chimaera	yes	no	ND	ND

Table 1.3. Comparison of the properties of mouse ES cells (mESCs), mouse epiblast stem cells (mEpiSCs), human ES cells (hESCs) and human iPS cells (hiPSCs) (adapted from *Han et al., 2013*).

1.3.2 From ESCs to neurons: neural differentiation protocols

Based upon their proliferative and pluripotential properties, ESCs are an attractive alternative to fetal or adult-derived central nervous system tissue and a useful tool for studying the mechanisms underlying embryonic development and cell differentiation. In recent years, several research groups have developed experimental protocols to induce ES cells to differentiate into numerous cell types. Among these protocols, particular attention has been paid to neural induction protocols, which allow the conversion of ESCs into neural cells.

Based on the work in non-mammalian species, particularly *Xenopus laevis* and chick, a "default model" hypothesis emerged for neural specification, proposing that in the absence of cell-cell signalling, ectodermal cells will adopt a neural fate (*Muñoz-Sanjuán and Brivanlou, 2002*). Several studies support this hypothesis for ESC neural differentiation, as well. For example, a study showed that when mouse ES cells are grown in the absence of feeder layers or exogenous factors (i.e. at very low cell density under which factors become limiting), roughly 70% of the cells die within 24 hours; of the remaining cells, the majority (82%) expresses the NESTIN (a specific marker of neural progenitors, *Lendahl et al., 1990*), indicating that the default fate of dissociated mouse ES cells is neural (*Tropepe et al., 2001*). However, many other studies in adherent monoculture support an alternative hypothesis as they sustain that, although the elimination of inductive signals for alternative fates is sufficient for ES cells to develop into neural precursors, this process is not a simple default pathway but requires autocrine FGF (*Li et al., 2001; Ying and Smith, 2003; Ying et al., 2003a*). This second hypothesis is also supported by observations in *Xenopus* and chick, where positive factors are required for neural conversion, and once the neural transition is achieved, signals are necessary to promote the survival and proliferation of these committed cells (*Stern, 2005*). In conclusion, the mechanism of neuroectoderm formation from pluripotent founder cells is controversial, and remains heavily debated.

In general, two types of experimental approaches for differentiating towards an *in vitro* neural fate have been proposed: Embryoid Bodies (EBs) and monolayer cultures (see **Table 1.4** adapted from *Cai and Grabel, 2007* that documents the properties of six principal protocols using ESCs to derivate Neural Stem Cells, NSCs).

(i) EBs (3D cultures): mimicking neural specification during embryogenesis. Initial studies demonstrated that the generation of neural cell types from mouse ESCs were based upon the formation of intermediates called embryoid bodies. This method tried to recapitulate the multistep process of neural development that occurs in the embryo. When mouse ESCs are removed from their feeder layer, and placed in suspension culture in the absence of the growth factor LIF, they form aggregates, which within 2-4 days consist of an outer layer of hypoblast-like cells (extraembryonic visceral endoderm) surrounding an epiblast-like core (primitive ectoderm) (*Rathjen et al., 2002; Cai and Grabel, 2007*). The EB core continues to express the ESC marker Oct4 and begins to express the primitive ectoderm marker FGF5. Between day 6 and 8, the core undergoes cavitation and forms an inner epithelial layer. Cells within this layer can be committed to definitive ectoderm, characterized by Sox2 and Otx2 expression. Subsequently, a re-organization into a columnar epithelium, resembling the neural tube (rosette), is accompanied by the expression of neurectoderm-specific markers such as Sox1 and Six3 (*Rathjen et al., 2002; Maye et al., 2004*). Moreover, it has been shown that retinoic acid causes a massive increase of the yield of neural lineage cells generated by EBs protocol (*Guan et al., 2001; Gottlieb, 2002*).

(ii) Monolayer culture: depriving ESCs of both cell-cell interactions and signals by culture in serum-free medium. Given the evidence for a default pathway of neural induction, in which the absence of signals promotes this lineage specification, several approaches for promoting neural differentiation of ESCs include a selection step of culture in a serum-free, nutrient-poor neurobasal medium. Indeed, direct differentiation of ESCs to neural stem cells (NSCs) can be obtained under serum-free conditions in monolayer culture at moderate cell densities. This differentiation protocol was established by the Smith's laboratory using the

mouse ESC line 46C. Cells are plated at around 1×10^4 cells/cm^2 on gelatin-coated dishes in the absence of LIF in serum-free defined medium (N2B27 medium). Within 5 days, up to 75% of the cells become NESTIN+ and form neural rosettes in which cells elongate and align radially, in a manner reminiscent of the neural tube formation, as well as ectoderm and neuroectoderm differentiation in ESC embryoid bodies (*Ying and Smith, 2003; Ying et al., 2003a*). Thus, regardless of the adopted protocol, the signature of neural progenitors in culture is the appearance of the neural rosettes.

Protocols	1. RA induction (4−/4+)	2. MEDII CM induction	3. Serum-free selection	4. Stromal co-culture	5. Low density clonal neurosphere	6. Monolayer Serum-free
Mouse ESC lines used	D3, CCE	E14, D3	J1, CJ7, D3, R1	ESC: CJ7, AB2.2,E14,ESB5 ntES:C4, C15,C16,CN1,2, CT2	R1	46C (E14 derived) 15 clones of ESC
Coculture, CM, or other factors	None	HepG2 conditioned medium	None	Stromal cell (MS5,S17,PA6 etc)	LIF	none
Initial plating density	Not quantified >1×10^5 cells/ml	1×10^5 cells/ml	2-2.5$\times10^4$cells/cm^2	50 cells/cm^2	1-20 cells/microwell	0.5-1.5x10^4 cells/cm^2
Culture type	Suspension	Suspension	Suspension + adherent	Adherent	Suspension	Adherent
EB formation	Yes	Modified	Yes	No	No	No
Serum or serum replacement	10% FBS+ 10% newborn calf serum	10% FBS	10% FBS for EB then no serum	15% serum replacement	No serum	No serum
Days to reach NSC peak	8	7	10-12	6	3 (4 hrs in PBS)	5
% NSC at peak	39% neuron-like Cells[b]	Nearly 100% 95.7% NCAM+	>80%	High, not quantified	100%	75%
NSC marker	βIII tubulin	Sox1, Sox2, nestin	nestin	nestin, NCAM, Musashi	nestin	Sox1, nestin
NSC Regional identity	NA	Otx1 (fore- and midbrain) En1, En2 (midbrain)	Otx1 (fore- and midbrain) En1 (midbrain)	No specific regional identity	Emx2 (forebrain) HoxB1 (hindbrain)	NA
Differentiation potential of derived NSC	NA	Neurons, glia (>95%), neural crest	Neurons (Map2) Astrocytes (GFAP) Oligodentrocytes (O4)	Neurons (dopaminergic, serotonin, GABAergic and motor neurons with high efficiency), glia	Neurons (Map2) Astrocytes (GFAP) Oligodendrocytes (O4)	Neurons (GABA, TH) Astrocytes (GFAP) Oligodendrocytes (CNPase)
Other lineages	Many other lineages present	None	Non-neural lineages selected against in serum-free medium	None	primitive endoderm present (GATA4), no mesoderm or definitive endoderm	Non-neural cell types and Oct4+ cells present
Key references:	Bain et al., 1995 Bibel et al., 2004	Rathjen et al., 2002	Okabe et al. 1996	Barberi et al., 2003 Kawasaki et al., 2000	Tropepe et al., 2001 Smukler et al., 2006	Ying et al., 2003

Table 1.4. Comparison of protocols using ESCs to derive NSC (adapted from *Cai and Grabel, 2007*)**.** ESCs, embryonic stem cells; NSCs, neural stem cells; RA, retinoic acid; CM, conditioned medium; NA, not applicable; GFAP, glial fibrillary acidic protein; FBS, fetal bovine serum; NCAM, nerve cell adhesion molecule; En1-2, Engrailed 1-2; MAP2, microtubule associated protein 2; GABA, γ-aminobutyric acid.

1.3.3 Neural rosette formation as *in vitro* model for neurogenesis

The development of the nervous system can be divided roughly into three processes; neural induction, neurulation, and regional specification. Cues taken from the embryo during each of these processes have been useful in establishing methods for *in vitro* neural differentiation. In particular*, in vitro* rosette formation of NSCs is widely and reliably used as a model system for reproducing *in vivo* neurogenesis. NSCs derived from ESCs, generated from both adherent and embryoid body-intermediate approaches, typically arrange themselves radially on a flat surface, eventually forming a central lumen. This radial, floral-like arrangement is known as "rosette". Neural rosettes have a structural and functional similarity to the embryonic neural tube (**Figure 1.6**). Morphological similarities are not the only common feature, indeed both rosettes and neural tube display localized zones of proliferation and can be patterned by signalling molecules and growth factors, suggesting that their formation and differentiation are governed by similar mechanisms (*Wilson and Stice, 2006; Elkabetz et al., 2008*). Interestingly, multicellular rosettes have been recently appreciated as important cellular intermediates during the formation of other organ systems, such as kidney tubule elongation, pancreatic branching morphogenesis, *Drosophila* eye development and many others (*Harding et al., 2014*). Several studies have also revealed that the cytoskeletal rearrangements responsible for rosette formation appear to be conserved. By contrast, the extracellular cues that trigger these rearrangements *in vivo* are more diverse and less well understood (*Harding et al., 2014*).

Rosettes were identified at the end of the 18th century by James Homer Wright (*Lee et al., 2002*) and Simon Flexner (*Flexner, 1891*). These structures were described as a precise histologic architectural pattern detected within specific nervous system tumours (neuroblastoma, medulloblastoma, retinoblastoma, ependymoma and pineocytoma). In the early 1960s, while Pierce and co-workers were investigating the potency of embryonal carcinoma cells that back then were used as analogue for early embryo cells in differentiation studies, some rosettes-reminiscent structures were identified in embryo bodies and were correlated for the first time to neural development (*Pierce and Verney, 1961*). In 1975, Martin and Evans (*Martin and Evans, 1975*) obtained similar results from *in vitro* experiments performed on embryonal carcinoma cells and in 1991 Kawata and colleagues described the *in vitro* neural rosette formation obtained from a clonal human teratocarcinoma cell line (*Kawata et al., 1991*). In 1995, for the first time post-mitotic neurons from ES cells were generated (*Bain et al., 1995*); soon after, in 1996, Ronald McKay and co-workers optimised the protocol to achieve up to 95% NESTIN+ cells in the proliferation phase and >60% MAP2+ cells after differentiation. According to this protocol, by 5-7 days a large proportion of the surviving NESTIN+ cells developed a small elongated shape and were organised in a rosette-like structure (*Okabe et al., 1996*). In 2001, Zhang and co-workers reported the appearance of neural rosettes in one of the first neural differentiation experiment with human embryonic stem cells (*Zhang et al., 2001*). More recently, many researchers explored whether the neural rosettes and the radial arrangement of neural progenitors in the neural tube are produced by similar mechanisms.

Below is a summary of these studies and their findings:

(i) During neural differentiation mouse ESC-derived NSCs exhibit distinct bipolar cell shape, a morphological characteristic of radial glia. Meanwhile, radial glia were recognized as the NSCs of the embryonic neural tube, which give rise to the variety of neurons and glia of the nervous system (*Noctor et al., 2001*). Mouse ESC-derived NSCs were found to express proteins characteristic of radial glia, i.e. RC2 and brain lipid binding protein (BLBP) (**Figure 1.6**) (*Bibel et al., 2004*). Human ESC-derived rosette NSCs also demonstrate radial glia-like properties (*Elkabetz et al., 2008*). Moreover, global gene expression and immunocytochemistry analysis performed using human ESC-derived rosette NSCs illustrate the capacity of these cells to give rise to multiple neuronal subtypes of both the central and peripheral nervous systems (*Elkabetz et al., 2008*). Glial subtypes, such as astrocytes (*Elkabetz and Studer, 2008*) and oligodendrocytes (*Hatch et al., 2009*), have also been produced from rosette NSCs.

(ii) The neural tube consists of a radially arranged neuroepithelium surrounding a central lumen. The side of the epithelial cell layer facing the lumen (ventricles) is considered apical, while the side facing the outer pial surface is considered basal. Apicobasal polarity is quickly established following the closure of the neural tube, partially due to the asymmetric distribution of proteins involved in adherent and tight junction formation (*Miyata, 2007*). The neural progenitors that compose the rosettes show the same antigenic characteristics of the cells in the neural tube in formation (*Abranches et al., 2009*). In fact, to recognize neural progenitors during differentiation, immunocytochemical analysis can be performed using the antibody directed against cytoplasmic protein NESTIN. This is an intermediate filament protein (IF) type VI (*Michalczyk and Ziman, 2005; Guérette et al., 2007*), expressed mainly in the cytoskeleton of neural progenitor cells, where it is involved in the radial growth of axon. Additionally, the tight junction zona occludens 1 (ZO1) protein is expressed exclusively at the apical surface of cells surrounding the lumen of the neural tube. Its localization coincides with expression of the neuroepithelial adhesion marker N-cadherin (N-Cad) (**Figure 1.6**) (*Marthiens and ffrench-Constant, 2009*). This asymmetric expression pattern is also observed in ESC-derived neural rosettes, where ZO1 protein is observed initially at tight junctions between cultured ESCs. However, at the onset of neural differentiation, this protein is rapidly redistributed and restricted to the centre of emerging rosettes (*Elkabetz et al., 2008*). Similarly to *in vivo* immunohistochemical analysis of the neural tube, ZO1 and N-Cad also show significant overlap in domains of expression *in vitro* at the luminal side of the neural rosette, where a correct cell polarity was established (**Figure 1.6**). Three major protein complexes are involved in establishing cell polarity: crumbs (CRB), partition-defective (PAR) and discs large (DLG) complexes (*Margolis and Borg, 2005*). CRB complex is required to determine the apical membrane and consists of the transmembrane CRB protein and the cytoplasmic proteins PALS1 and PATJ (*Margolis and Borg, 2005*). PAR complex, on the other hand, contributes in defining the apico-lateral membrane and includes PAR3, PAR6, aPKC and CDC42 proteins (*Margolis and Borg, 2005*). Finally, the DLG complex delineates the base-lateral plasma membrane and consists of Scribble, Lgl (Lethal giant larvae) and Dlg (Discs large) proteins (*Margolis and Borg, 2005*). These three complexes have an antagonistic action and are the crucial players in the regulation of the cell polarity, following a spatial-temporal order and interacting with different proteins of cellular adhesion and cytoskeleton (*Margolis and Borg, 2005*).

(iii) A work from Lorenz Studer's research group focused on the identity of hESC-derived neural rosettes, finding that by default such neural rosettes adopt polarized neuroepithelial structures of anterior CNS fate including expression of Foxg1b (BF1) (*Elkabetz et al., 2008*). Indeed, in the absence of extrinsic patterning cues neural rosettes acquire markers of anterior neural ectoderm. However, neural rosette stage cells can also undergo anteroposterior specification and, after specific treatments with sonic hedgehog (SHH) and retinoic acid (RA) or SHH and FGF8, spinal motoneurons and midbrain precursors were obtained, respectively (*Elkabetz et al., 2008*). Moreover, the differentiation of neural rosette progeny in the presence of ventral and dorsal patterning cues such as SHH or Wnt3A led to the induction of markers compatible with ventral forebrain fate and the emergence of GABA+ neurons and cells expressing dorsal markers such as Msx1. The anterior CNS bias of neural rosettes is reminiscent of the default model postulated in classical studies of *Xenopus* CNS development in which it was found that anterior CNS fates are established first and are followed by caudal transformation in response to secreted signals (*Elkabetz et al., 2008*).

(iv) A well-known feature of neural tube cells is the interkinetic nuclear migration (INM), a precise movement of cell nuclei according to their cell cycle stage (*Sauer and Walker, 1959*). Nuclei within rosettes undergo similar migration. Nuclei undergoing DNA synthesis were located at the basal side of the cells (visualized by BrdU incorporation), whereas mitotic nuclei were always confined to the luminal side, as shown by phosphohistone H3 (PH3) antibody staining. PH3 and BrdU stainings never overlapped, indicating a highly controlled spatial confinement, which normally occurs in the proliferating neural tube (*Abranches et al., 2009*).

(v) In the embryonic neuroepithelium, the Notch pathway controls the rate at which proliferating neural progenitors commit to differentiation. When Notch activity is inhibited, precocious neuronal differentiation is

usually observed. Similarly, depletion of Notch signalling in rosette cultures causes rosette progenitors to embark on neuronal differentiation (*Abranches et al., 2009*).

Overall, the results indicate that neuroepithelial rosette cultures recapitulate several aspects of embryonic neural tube development and early mammalian neurulation events. They, therefore, represent a good experimental system to study early stages of neural development *in vitro* encompassing neural induction and neurulation.

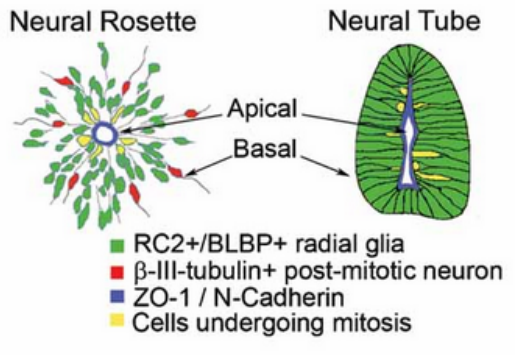


Figure 1.6. ESC-derived neural rosettes *in vitro* bear striking resemblance to the neural tube (from *Germain et al., 2010*). Apicobasal polarity is similar between rosettes and the neural tube, with tight junctions (indicated by ZO1 expression, blue) at the apical surface forming a lumen. Radially arranged progenitors in the rosette, as well as the neural tube, express RC2 and BLBP (green).

2. AIM

Aim

This thesis work aims at investigating whether the polyQ domain in HTT is functionally relevant during *in vitro* neurogenesis. A mES cell-based platform for complementation assays with HTT fragments mutated or deleted in the CAG tract has been set up to this purpose. Subsequently, these cell lines were subjected to neural differentiation and analysed by a rosette assay. This *in vitro* assay measures the ability of ES cell-derived neural progenitors to form neural rosettes, which are radial arrangements of columnar cells expressing many of the proteins found during neural tube formation *in vivo.* The work is divided in two phases.

(1) In the first phase, we generated mES cell lines bearing N-terminal HTT carrying different polyQ tract variants. In particular, the selected CAG repeat variations were the deletion, substitutions, and length modification (both in the normal and pathological range). The newly generated cell lines were then tested in a neural induction protocol to evaluate their rosette formation capacity, by measuring (i) the percentage of neural progenitor cells organized in rosettes, (ii) rosette size and (iii) rosette lumen size.

(2) Secondly, we worked on the rosette assay procedure in order to develop a novel automated system to perform unbiased and quantitative assessment of the rosette phenotypes *in vitro*.

3. METHODS

Cell lines. We used mouse ES cell lines expressing the wild-type *Hdh* gene ($Hdh^{+/+}$) or being homozygous knockout for the gene ($Hdh^{-/-}$), in which both alleles of the *Hdh* gene were inactivated by deletion of exons 4 and 5 (*Duyao et al., 1995*). All complemented cells were derived by the transfection of pCAG in $Hdh^{-/-}$ cells. Both $Hdh^{+/+}$ and $Hdh^{-/-}$ were analysed for karyotypic abnormalities (Q-banding) and they showed several common chromosomal alterations at the cell population level (the most frequent were duplications of chromosomes 1, 8, 11 and 14) (*Rebuzzini et al., 2008*).

Mouse ES cell culture. ES cells were maintained in Glasgow minimal essential medium supplemented with 10% heat-inactivated fetal bovine serum (vol/vol, EuroClone, REF ECS0186L), 0.1mM β-mercaptoethanol (Gibco, REF 31350-010), 100μM non-essential amino acids (Gibco, REF 11140-035, 1mM sodium pyruvate (Gibco, REF 11360-039), 2mM l-glutamine, 100U/mL penicillin, 100μg/mL streptomycin (EuroClone, REF ECB3001D) and 1,000U/mL murine leukemia inhibitor factor (LIF, ESGRO) (Millipore, REF ESG1107) in gelatinized tissue culture flasks. Cells were passaged every 2 days after dissociation with 0.05% trypsin-EDTA (vol/vol) (Gibco, REF 15400-054).

Plasmids. The constructs, encoding for the N-terminal portions of mouse (Mm) or human (Hs) HTT, were designed according to previously published HTT multiple sequence alignment (*Tartari et al., 2008*) and cloned in pCAG plasmids. Below are reported the newly generated plasmids:

- pCAG Mm 0Q
- pCAG Mm 2Q
- pCAG Mm 4Q
- pCAG Mm 7Q
- pCAG Mm Q3PQ3
- pCAG Mm Q3Q(CAA)Q3

pCAG Hs 15Q and pCAG Hs 128Q were previously generated in the laboratory.

Bacteria transformation. Under a sterile environment, 1μL of re-suspended plasmid (40ng/μL) has been added to a single 50μL vial containing DH5α bacteria. After 30 minute of ice incubation, tubes are transferred in a heating block at 42°C for 30 seconds and then for 2 minutes in ice again (step of thermal shock required to make membrane permeable to DNA). Then, 250μL of S.O.C. solution (w/v: 2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM MgCl2, 10mM MgSO4, 20mM glucose, Invitrogen) are added and cells are incubated for 1 hour at 37°C. At the end of incubation, bacteria are plated on warmed LB-agar plates (Luria-Bertani Broth w/v: 1% tryptone, 0.5% yeast extract, 10mM NaCl, 1.5% agar) containing ampicillin. Thanks to the ampicillin resistance gene present in the transformed plasmids, only bacteria that have incorporated it will grow in presence of ampicillin. Plates are incubated over-night at 37°C, to make transformed bacteria able to grow and to form single colonies.

DNA maxi-preparation. QIAGEN Plasmid Maxi kit has been used according to the manufacturer's instructions. Single bacterial colonies were picked from a selective plate and inoculated in 4 mL of Luria-Bertani Broth (LB) (w/v: 1% tryptone, 0.5% yeast extract, 10mM NaCl) medium containing the ampicillin. Bacteria solutions were incubated over-night at 37°C with vigorous shaking. 300μL of this starter culture were inoculated in 300mL of LB medium containing the appropriate selective antibiotic. Bacteria were grown at 37°C for 12-16 hours with vigorous shaking. Bacterial cells have been harvested by centrifugation at 4500xg for 30 min at 4°C. Pellet was re-suspended in 10mL of Buffer P1. 10 mL of Buffer P2 was added for cell lysis. The tube was mixed thoroughly by vigorously inverting it 4-6 times, and incubated at room temperature (15-25°C) for 5 min. 10mL of chilled Buffer P3 were added, mixed immediately and thoroughly by vigorously inverting 4-6 times, and incubated on ice for 20 minutes. Tubes were centrifuged at 10,000xg for 30 minutes at 4°C. During centrifugation, QIAGEN-tip 500 was equilibrated by applying 10mL Buffer QBT (750mM NaCl; 50mM MOPS, pH 7.0; 15% isopropanol (v/v); 0.15% Triton Xa-100 (v/v)), and allowed the column to empty by gravity flow. The supernatant was applied to the equilibrated QIAGEN-tip and allowed to enter the resin by gravity flow. QIAGEN-tip has been washed 2 × 30mL with Buffer QC (1.0M NaCl; 50mM MOPS, pH 7.0;

15% isopropanol (v/v)). DNA was eluted with 15mL Buffer QF (1.25M NaCl; 50mM Tris-HCl, pH 8.5; 15% isopropanol (v/v)). DNA was then precipitated by adding 10.5mL of room-temperature isopropanol. Mixed and centrifuged immediately at 10,000xg for 30 min at 4°C. The supernatant was removed and DNA pellet washed with 5ml of room-temperature 70% ethanol, and centrifuged at 10,000xg for 10 minutes. The supernatant was removed and the pellet air-dried for 5-10 minutes. DNA was re-dissolved in 150-200µL of H_2O. At the end of this maxi-preparation, extracted DNA was used to transfect mES cells after DNA quality the control and quantification.

Plasmid DNA quality control and quantification. The plasmid DNA obtained after maxi-preparation was quantified with Nanodrop 1000 (Thermo Fisher Scientific) spectrophotometer. The same analysis also allows to evaluate the purity of the plasmid DNA solution. To verify that the constructs, encoding for N-terminal fragments of HTT, were properly inserted in the pCAG plasmids with their correct length, an enzyme digestion was performed using the restriction enzymes NotI and XhoI. The pCAG plasmid has two restriction sites specifically recognized by this enzyme: one site is immediately upstream of the N548 fragment while the second one is downstream of it. If the N-terminal fragment has been inserted properly, the enzymatic digestion will produce two fragments of about 6.4Kb and 1.7Kb. We then performed DNA electrophoretic on 1% agarose gel in order to separate the fragments by their size. The gel was visualized through a transilluminator (Molecular Imager Gel DocTM Xr Imaging System, Bio-rad) and images were analysed with Quantity One (Bio-rad) software.

DNA transfections. *Hdh*$^{-/-}$ ES cells were plated at a density of 2 x 10^4 cells per cm^2. After 24 h, Lipofectamine 2000 (Invitrogen) was used to transfect the cells with different expression vectors (10µg). Cells were selected for effective transfection with puromycin (2µg/ml) 24 h after lipofection, and for 10 days. Western blot and immunocytochemistry were performed to determine the expression of the transgene.

Mouse ES cell monolayer differentiation. ES cells were dissociated and plated onto 0.1% gelatin-coated tissue culture dishes at a density of 2.5-4 x 10^4 cells per cm^2 in N2B27 medium. Medium was renewed every 2 days. N2B27 medium was a 1:1 mixture of DMEM/F12 and Neurobasal medium containing 1:200 N2 supplement (Gibco, REF 17502-048), 1:100 B27 supplement (Gibco, REF 17504-044) and 0.1mM β-mercaptoethanol (Gibco, REF 31350-010). Under this monoculture condition, in the absence of LIF, mES cells lose pluripotent status and commit to a neural fate over 5-6 days (*Ying and Smith, 2003; Ying et al., 2003a*). **Figure 3.1** recapitulates specifications of the protocol applied.

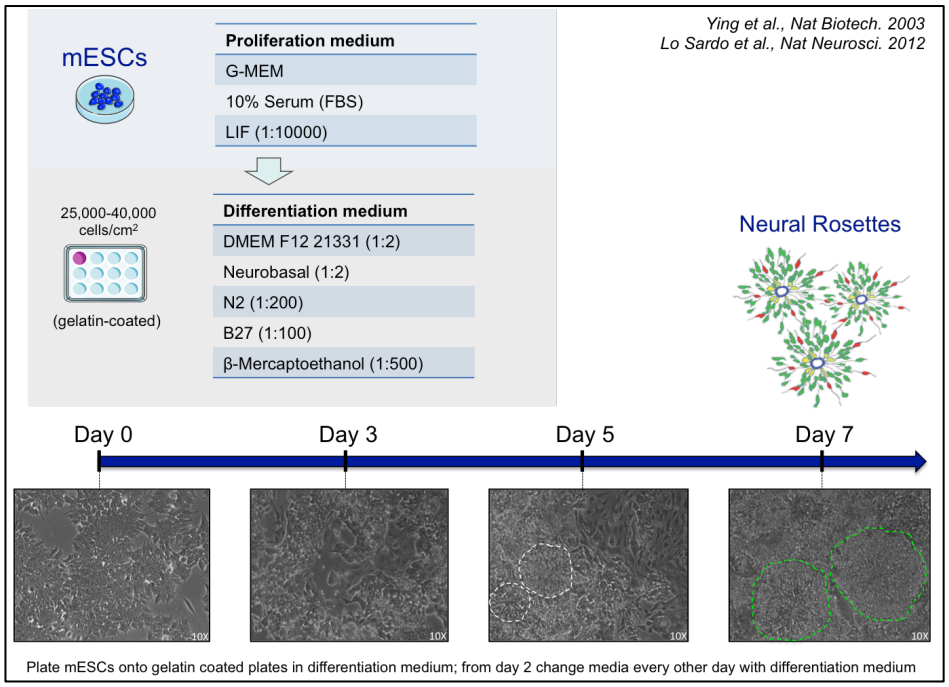


Figure 3.1. Schematic representation of the neural differentiation protocol.

Immunocytochemistry. Cells were fixed in 4% paraformaldehyde for 15 min at room temperature (20-23°C). They were subsequently permeabilized and blocked in blocking buffer containing PBS with 0.5% Triton X-100 (vol/vol) and with 5% fetal bovine serum (vol/vol) for 1 h at room temperature. Primary antibodies were diluted in blocking buffer and incubated overnight at 4 °C. After three washes in PBS, appropriate secondary antibodies, conjugated to Alexa fluorophores 488 or 568 (Molecular Probes, Invitrogen) and were diluted 1:500 in blocking solution, and incubated for 1 h at room temperature. Cells were incubated for 10 min with Hoechst 33258 (5µg/mL, Molecular Probes, Invitrogen) for nuclear counterstaining. Images were acquired with a Leica DMI 6000B microscope (equipped with LAS-AF imaging software) and processed with the software ImageJ (US National Institutes of Health).

Antibodies. The following antibodies (Abs) and dilutions were used for immunocytochemistry and western blot:

Monoclonal Ab for HTT (Millipore, Mab2166; immunofluorescence, 1:800; western blot, 1:1000)
Monoclonal Ab for NESTIN (Millipore, Mab353; immunofluorescence, 1:200)
Polyclonal Ab for ZO1 (Life Technologies, REF 402300; immunofluorescence, 1:200)
Monoclonal Ab for OCT4 (SantaCruz, SC-5279; immunofluorescence, 1:100)
Polyclonal Ab for NANOG (Abcam, ab21624; immunofluorescence, 1:100)
Polyclonal Ab for PALS1 (Life Technologies, sc-33831; immunofluorescence 1:200)
Polyclonal Ab for αPKC (Life Technologies, sc-216; immunofluorescence 1:200)
Polyclonal Ab for PAR3 (Millipore, cod. 07-330; immunofluorescence 1:200)

Protein lysates and western blot. Cells were lysed in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.1% SDS, 1% nonidet P40, 0.5% sodium deoxycholate, wt/vol) with 1 mM PMSF and protease inhibitor (Thermo Scientific, REF 1861281). Lysates were cleared by centrifugation at 12,000g and 4 °C for 30 min. The resulting supernatant was collected. Protein concentration was determined with the Pierce-BCA Protein assay kit (Thermo Scientific REF 23225) and 30µg were loaded on a 7.5% SDS-PAGE gel. Separated proteins were transferred to a nitrocellulose membrane, blocked with 5% non-fat dry milk (wt/vol, Bio-rad, REF 170-6404) in Tris-buffered saline (TBS) with 0.1% Tween-20 (vol/vol), and incubated with primary antibody at

room temperature for 3 h. After washing, filters were incubated for 1 h at room temperature with a secondary antibody (peroxidase conjugate, Bio-rad, 1:3000) and then washed three times with TBS and 0.1% Tween-20. The Clarity Western ECL Substrate (Bio-rad, REF 170-5061) was used to visualize immunoreactive bands by chemiluminescence detection with ChemiDoc MP Imaging System (Bio-rad).

Rosette quantification. Rosettes were quantified in cell cultures on day 7 of neural differentiation after staining for NESTIN and ZO1. Twelve images of random fields for each cell line were acquired. Each experiment was independently repeated at least three times. Rosette quantification consists of measuring three parameters: *(i)* percentage of NESTIN+ cells inside the rosette, area of each *(ii)* rosette and *(iii)* lumen. **Figure 3.1** shows the pipeline developed using the ImageJ software. For the measurement of the percentage of NESTIN+ cells inside rosettes, we firstly calculated the total area occupied by NESTIN+ cells. Second, the shape of each rosette was outlined and the area occupied by the NESTIN staining was considered as the area occupied by NESTIN+ cells inside the rosettes. The remaining NESTIN+ signal was selected as the area occupied by NESTIN+ cells outside the rosettes. For rosette and lumen size, the contour of each rosette (as highlighted by the marker NESTIN) and that of each rosette's lumen (as highlighted by marker ZO1) was manually outlined and two lists of regions of interest (ROIs) were measured to calculate rosette and lumen mean area, respectively.

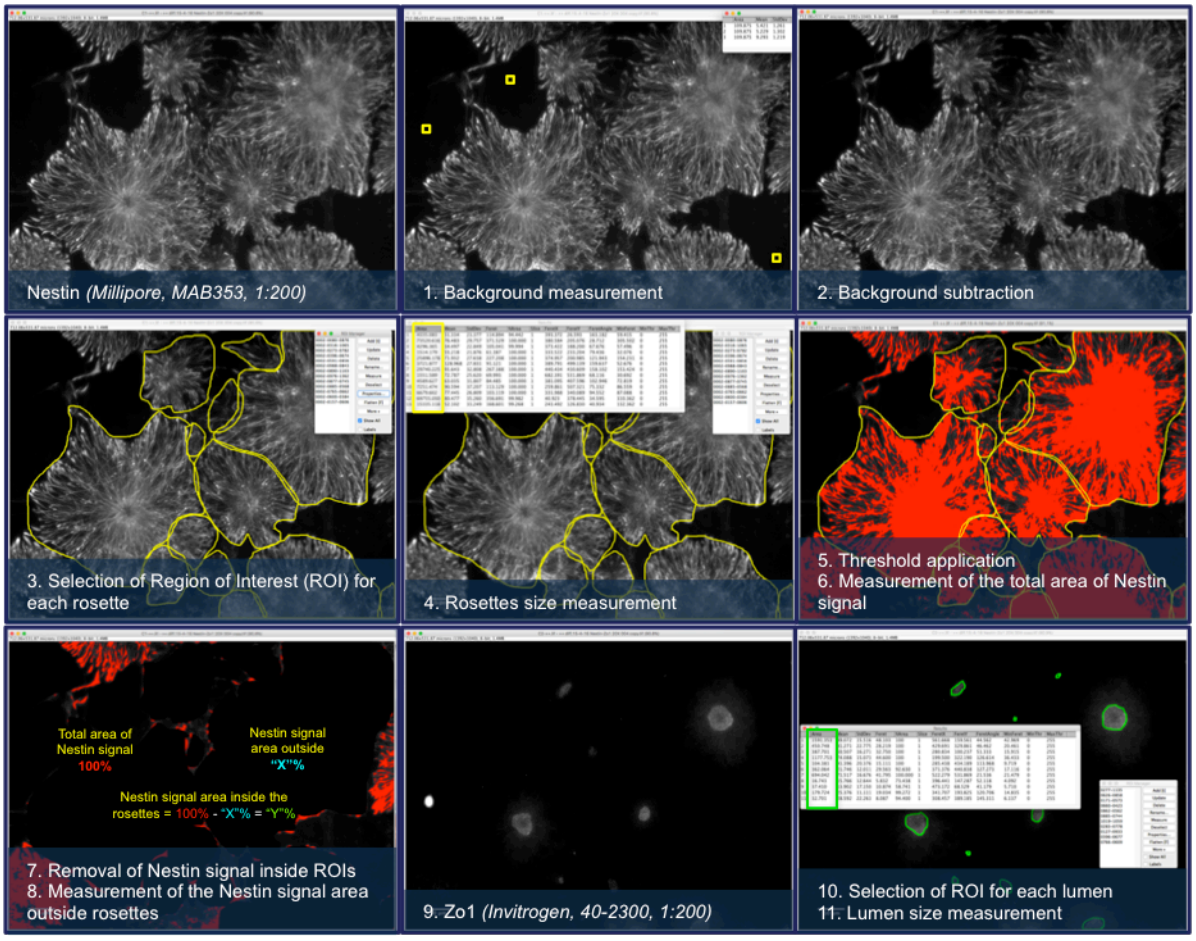


Figure 3.2. Step-by-step rosette quantification (manual method).

Apoptosis Assay. Cells were seeded in triplicate in two 96-well plates, at 16 x 10^3 cells per well. After 24-h incubation at 37 °C, cells received 100µl of freshly prepared GMEM+LIF without serum. 48 h after plating, 100µl of CellTiter Reagent (Promega, REF G7570) was added to all wells. After 10 min of incubation at room temperature, luminescence was measured with a Veritas-MicroplateLuminometer (Turner Biosystems) to detect cell viability. Caspase-Glo 3/7 Reagent (100µl, Promega, REF G8091) was added to all parallel wells. After 30 min of incubation at room temperature, luminescence was recorded with a Veritas-MicroplateLuminometer (Turner Biosystems) to detect Caspase-3/7 activity. Each experiment was independently repeated at least three times. Results are presented as the ratio of caspase-3/7 luminescence and cell viability luminescence recorded for each cell line.

Statistical analyses. The difference between all cell lines was tested using generalized linear mixed models assuming a binomial error distribution and a logit link-function (for the variable: % NESTIN+ cells inside rosettes) or a Gaussian error distribution and an identity link-function (for the variables: rosette and lumen size). In modelling the percentage NESTIN+ cells we considered experiment and well as random effect factors. In modelling rosette and lumen size we considered experiment, well and field as random effect factors. Post hoc comparisons between cell lines were performed by applying Bonferroni procedure. The difference between cell lines in the Caspase-3 activity was tested in linear mixed models assuming a Gaussian error distribution and an identity link-function. The random effect of the experiment was included in the model. Post hoc comparisons between cell lines were performed by applying the Bonferroni procedure.

4. RESULTS

4.1 HTT rosette formation potential is confined to its N-terminal portion

Previous studies have shown that HTT is important for rosette formation. In fact, HTT depletion in *Hdh*$^{-/-}$ cells leads to a reduction in rosette formation and size while complementation with the N-terminal 548aa portion of the protein is able to restore proper rosette formation and size in the same cells (*Lo Sardo et al., 2012*).
In the first part of this thesis work I further tested the robustness of the data already published in *Lo Sardo et al., 2012*. To this aim I have prepared replicate experiments in which I evaluated the rosette formation potential, by measuring *(i)* the percentage of NESTIN+ cells organized in rosettes, *(ii)* the rosette size and *(iii)* the rosette lumen size. The results obtained confirmed that *Hdh*$^{-/-}$ neural progenitors cells have a drastically reduced capacity to generate proper rosettes and lumens (see column 2 and 3 of **Table 4.1, 4.2, 4.3, 4.4** for the number of experimental replicates). We then confirmed that the expression of the N-terminal portion of HTT alone in *Hdh*$^{-/-}$ cells (7Q cells) completely rescues rosette formation by promoting the organization of high size mature rosettes with large and well-defined lumens (see column 5 of **Table 4.1, 4.2** and column 4 of **Table 4.3** for the number of experimental replicates).

4.2 Generation of cell lines stably expressing mutated versions of the CAG tract

In order to determine whether the HTT polyQ domain is implicated in rosette formation, we performed a series of variations in this area and tested their effect in a *Hdh*$^{-/-}$ cell-based complementation assay. N-terminal HTT portions carrying selected polyQ variations were cloned into a pCAG plasmid, a mammalian expression vector that contains a strong CAG promoter ensuring stable and continued expression of the transgenes in mES cells (**Figure 4.1a,b**). Specifically, we generated the following N-terminal HTT variants:

(i) a mouse N-terminal HTT construct deleted of the polyQ region;

(ii) a mouse N-terminal HTT construct carrying either 2 or 4 glutamines. We decided to test the 2Q-stretch because this is the first polyQ tract appearing in deuterostome evolution, in echinoderms. On the other hand, we opted for the 4Q-stretch in order to mimic the second step of polyQ evolution i.e. the 4Qs found in lower vertebrates (birds, reptiles, fishes and amphibians);

(iii) a mouse N-terminal HTT construct in which synonymous (CAG replaced with CAA, encoding for a glutamine) and non-synonymous (CAG replaced with CCG, encoding for a proline) substitutions were introduced in the 4th position of the CAG tract. We decided to use a proline interruption since several proteins harbouring a polyQ tract also carry prolines. Moreover, since proline has different physical-chemical properties with respect to glutamine, and provides rigidity to the polypeptide chain of the protein, we predicted this mutation could significantly alter the structure of the polyQ tract. On the contrary, the replacement of a CAG codon with a CAA (synonymous substitution) should not affect the biological function of the tract, as the amino acidic sequence remains unaltered.

Schematic diagrams of HTT constructs bearing different polyQ mutations are reported in **Figure 4.2a**. All pCAG vectors were transfected in *Hdh*$^{-/-}$ mES cells and after 10 days of puromycin selection we pooled selected cells. Transgene expression was assessed by western blot and immunocytochemistry analysis using MAB2166 antibody that recognizes the N-terminal portion of HTT. Western blot analysis detected expression of the transgenes at the expected size in the transfected *Hdh*$^{-/-}$ cell lines, whereas the control cell lines (untransfected *Hdh*$^{+/+}$ and *Hdh*$^{-/-}$ cells) showed no transgene expression as expected (**Figure 4.2b,c**). Immunocytochemistry analysis corroborated transgene expression and revealed that HTT fragments are confined to the cell cytoplasm **(Figure 4.2d)**. We have also evaluated the pluripotency state of these newly generated cell lines and found that OCT4 and NANOG expression were similar between untransfected and transgene expressing cells (**Figure 4.3**).

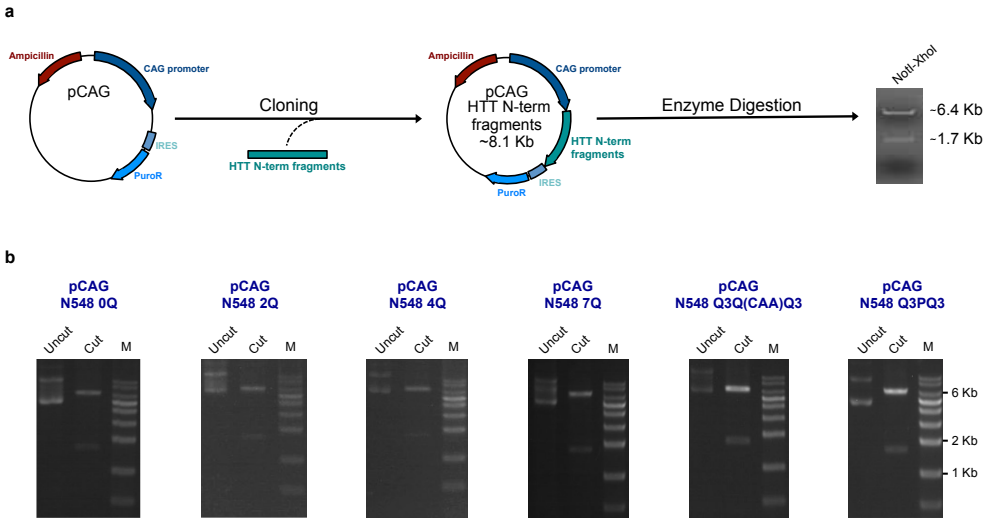


Figure 4.1. Generation of pCAG plasmids with the *HTT* sequences encoding for different HTT N-terminal fragments. (a) Schematic representation of the experimental strategy used to generate pCAG plasmids. **(b)** Gel electrophoresis of the pCAG plasmid products (pCAG N548 0Q, -2Q, -4Q, -7Q, Q3Q(CAA)Q3, and –Q3PQ3) digested by NotI and XhoI enzymes. For each gel, lane 1: pCAG plasmid product (Uncut); lane 2: pCAG plasmid digested product (Cut); lane 3: 1Kb DNA Ladder (M).

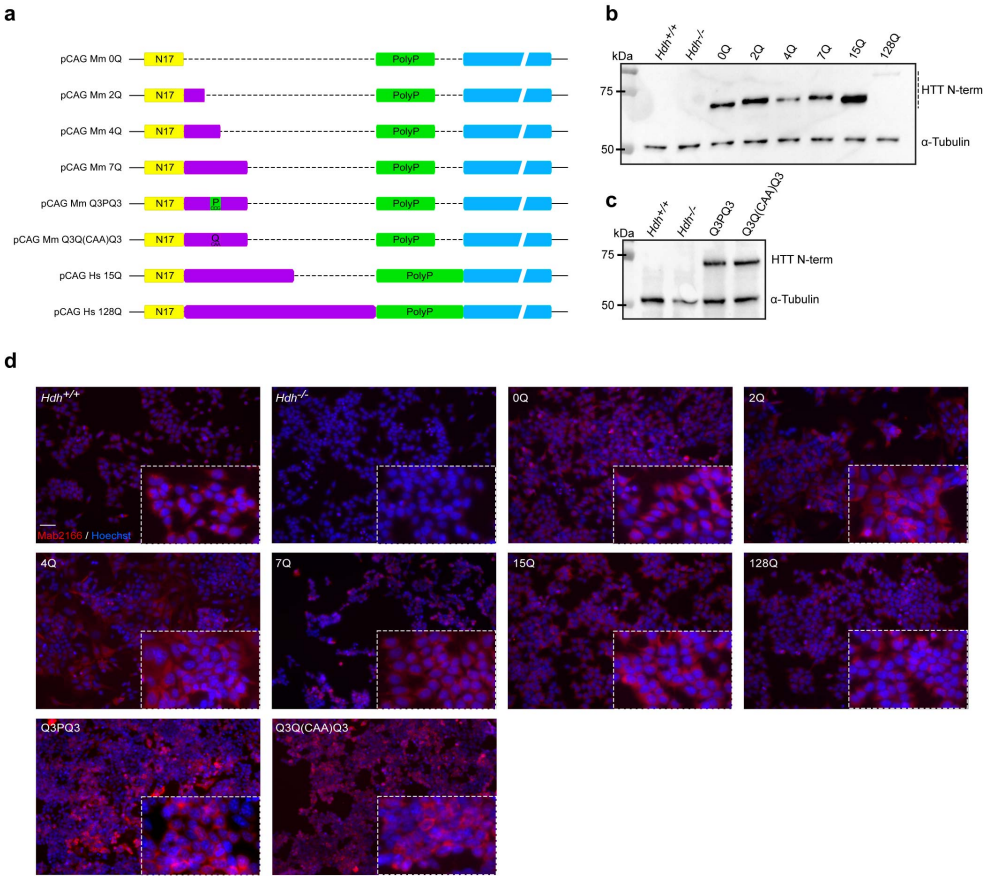


Figure 4.2. *Hdh*$^{-/-}$ cells overexpressing HTT N-terminal fragments with mutations in the polyQ tract. (a) Schematic representation of expression vectors carrying murine and human HTT N-termini (Mm, *Mus musculus*; Hs, *Homo sapiens*) with different polyQ lengths/modifications. **(b,c,d)** MAB-2166 (anti-HTT) Western blot (b,c) and immunostaining (d) in self-renewal conditions of *Hdh*$^{-/-}$ cell lines that overexpress murine or human HTT N-termini with different polyQ lengths/substitutions. The scale bars correspond to 50µm. Insets are shown at 2X magnification.

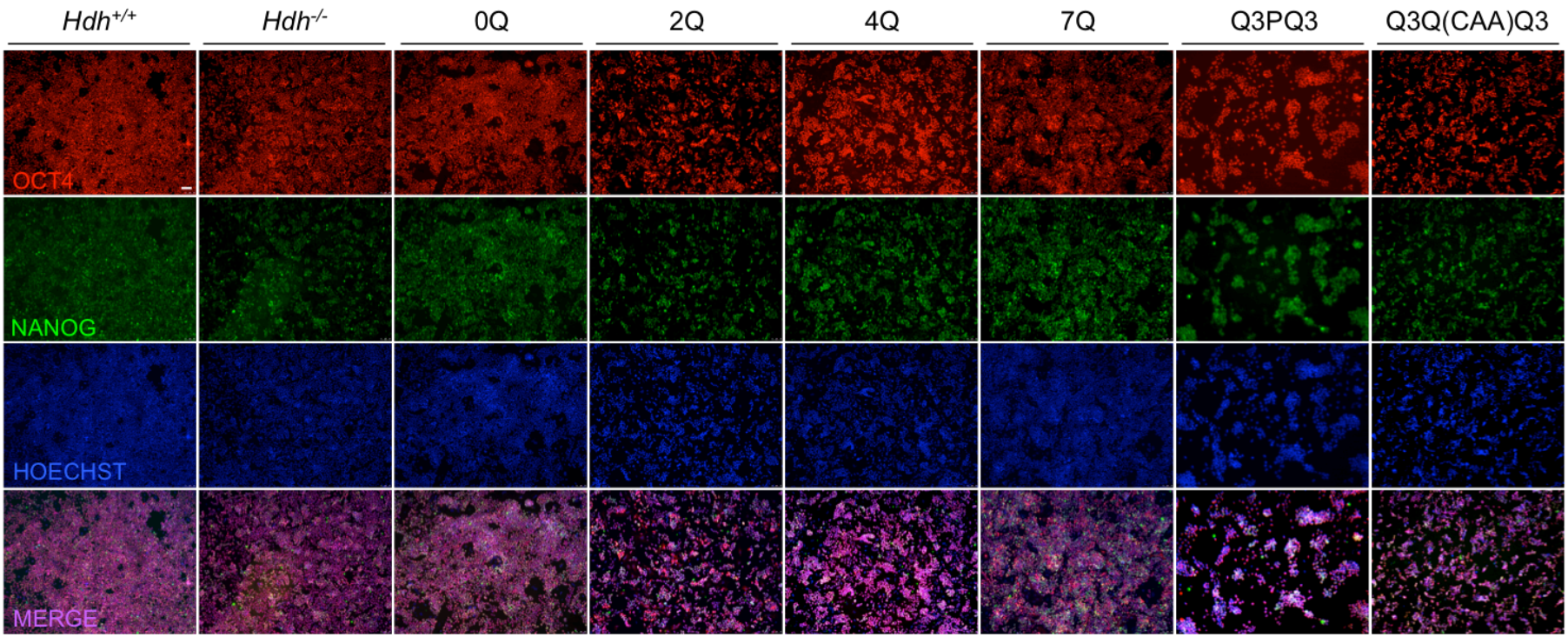


Figure 4.3. Immunocytochemical analysis of pluripotency markers. ES cells fixed in self-renewal conditions and stained for OCT4, NANOG, and HOECHST. The scale bar corresponds to 75µm.

4.3 PolyQ stretch deletion causes defects in rosette formation

To investigate the possible role of the polyQ tract in neurulation, we tested rosette formation potential in polyQ-depleted cells (0Q cells) by using *Hdh*$^{+/+}$ and 7Q cells as positive controls while *Hdh*$^{-/-}$ cells were used as negative controls. Differentiated cells were fixed at day 7 and stained with antibodies against NESTIN and ZO1 markers (**Figure 4.4a,b;** see **Figure 3.1** for a schematic representation of neural differentiation protocol). NESTIN is a type VI intermediate filament protein that marks the intermediated filaments of the cytoskeleton of the neural cells during the radial growth of their axons (*Lendahl et al., 1990*), while ZO1 is a protein present within the occluding junctions which identifies the rosette lumen (*Marthiens and ffrench-Constant, 2009*).

In order to test the activity of polyQ-deleted N-terminal HTT upon expression in *Hdh*$^{-/-}$ cells we exposed the cells to neural induction and monitored the appearance of rosettes over time. In **Figure 4.4c,d,e** 0Q cells exhibit an impairment of rosette formation, as judged by the percentage of NESTIN+ cells inside rosettes as well as smaller rosette and lumen areas (**Figure 4.4c,d,e**). In particular, we found that in six independent experiments, the percentage of NESTIN+ cells inside the rosettes was significantly reduced in 0Q cells (47.5±6.7%) compared to *Hdh*$^{+/+}$ and 7Q control cell lines (89.2±1.3% and 87.8±2.8%, respectively). Moreover, the mean rosette area, calculated on 271 images from a total of 2938 rosettes, was approximately 50% smaller in 0Q cells compared to *Hdh*$^{+/+}$ and 7Q cells. We also evaluated the lumen size of the rosettes, as it mirrors the enhanced or reduced numbers of neural progenitors that compose the rosettes. By measuring a total of 1241 lumens we report a significant reduction of the mean lumen size in 0Q cells, where the average area was about 3 times smaller than the one measured in *Hdh*$^{+/+}$ and 7Q cultures. See **Table 4.1** for the number of experimental replicates.

Our *in vitro* results suggest that the polyQ tract may have played a role in the appearance and/or evolution of a nervous system, as demonstrated by the lower percentage of neural progenitors organized in rosettes as well as by the smaller size of rosettes and lumens in 0Q cultures.

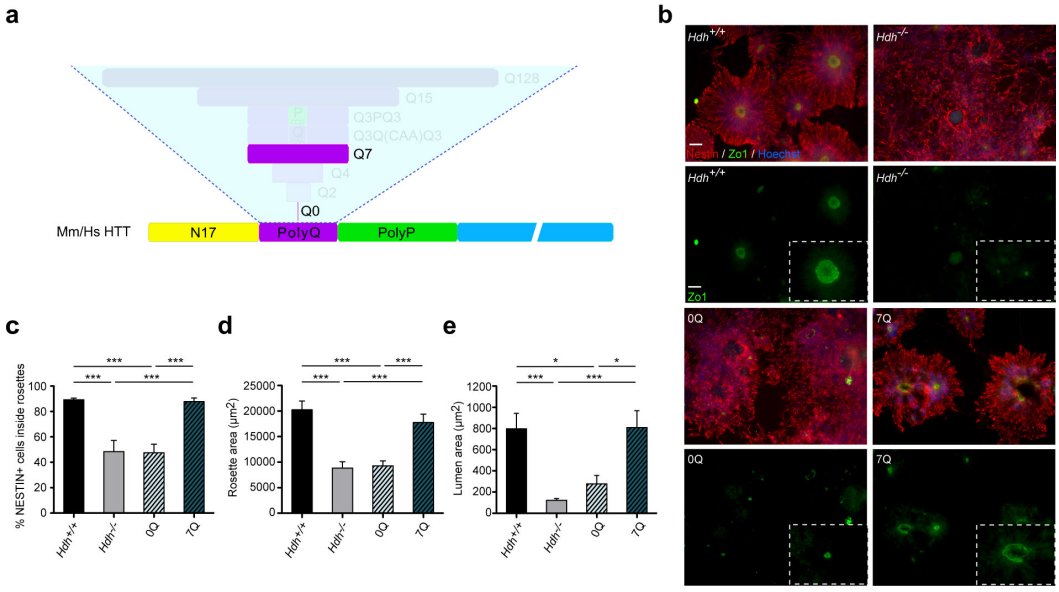


Figure 4.4. Effects of HTT polyQ tract deletion in an ES cell-derived neural rosette assay mimicking early neurulation. (**a**) Diagram of the different HTT N-terminal portions assessed, in which the polyQ tract modifications (with or without polyQ) are highlighted (**b**) Representative images of rosette/lumen phenotype in *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 0Q and 7Q ES cells stained for NESTIN and ZO1 at day 7 of neural induction. (**c,d,e**) Percentage of NESTIN+ cells inside rosettes, rosette and lumen mean area in *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 0Q, and 7Q ES cells exposed to neural induction. Data are expressed as mean ± SEM. See **Table 4.1** for the number of biological replicates. All pairwise statistical comparisons were run by applying generalized linear mixed models with a *post hoc* Bonferroni correction (*: P<0.05; ***: P<0.001). The scale bars correspond to 50µm. Insets are shown at 2X magnification.

4.4 PolyQ stretch increases rosette formation in a length-dependent manner

Bioinformatics studies established that the length of the polyQ tract increases throughout the evolution of deuterostomes, in concomitance with the emergence of progressively more complex nervous systems (*Tartari et al., 2008*). On this basis, we speculated that the effect of HTT polyQ tract on rosette formation could be Q-dependent, and that more Q in HTT could lead to a progressively more important rescue in the rosette phenotype in *Hdh*$^{-/-}$ cells (described in 4.3). To test this hypothesis, cell lines carrying a mouse HTT N-terminal transgene bearing 0Q, 2Q, 4Q or 7Q were differentiated in an unique subset of replicate experiments (**Table 4.2**). Our data showed that the length of Q stretch linearly covaries with rosette formation potential (**Figure 4.5**). In fact, an increase in CAG length leads to a higher percentage of NESTIN+ cells organized in rosettes, as well as to larger lumen and to an increase in rosette area (P value < 0.001 in all three parameters measured; See **Table 4.2** for the number of experimental replicates).

These results suggest that the number of Q in HTT acts as a fine modulator of the pro-neurulation function of the protein.

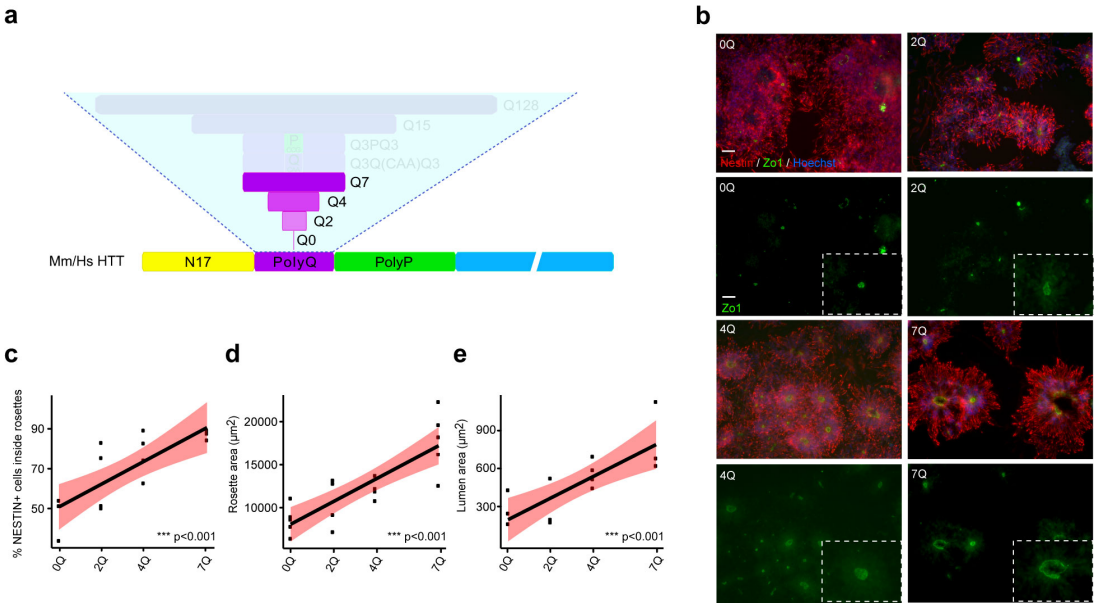


Figure 4.5. Effects of HTT polyQ tract length in an ES cell-derived neural rosette assay mimicking early neurulation. (**a**) Diagram of different HTT N-terminal portions assessed, in which the polyQ tract modifications (polyQ tract length) are highlighted (**b**) Representative images of rosette/lumen phenotype in 0Q, 2Q, 4Q and 7Q ES cells stained for NESTIN and ZO1 at day 7 of neural induction. (**c,d,e**) Regression lines and confidence intervals between Q length and rosette/lumen phenotypic variables in neuralized 0Q, 2Q, 4Q and 7Q ES cells. See **Table 4.2** for the number of biological replicates. All pairwise statistical comparisons were run by applying generalized linear mixed models with a *post hoc* Bonferroni correction (***: P<0.001). The scale bars correspond to 50µm. Insets are shown at 2X magnification.

4.5 Interruption of polyQ tract purity reduces *in vitro* neurulation

We next investigated whether an interruption in the CAG tract tarnishes rosette formation potential. Specifically, we performed differentiation experiments with cells bearing a CCG substitution that causes a replacement from a glutamine to proline (Q3PQ3 cells) or a CAA synonymous substitution that leads to no amino acid replacement (Q3Q(CAA)Q3 cells) in the 4th position of the polyQ tract. Our complementation assays showed that the presence of a proline in the mouse polyQ tract in the 4th aa position causes a reduced percentage of neural progenitors organized in rosettes with respect to control *Hdh*$^{+/+}$ and 7Q cultures (**Figure 4.6c**, columns 1, 3, and 4**)**. We also found that a proline interruption of the CAG tract leads to the formation of smaller rosettes and lumens compared to *Hdh*$^{+/+}$ and 7Q cells (**Figure 4.6d,e**, columns 1, 3, and 4**).** By contrast, a CAA synonymous substitution in 4th position of the mouse CAG stretch did not affect the rosette phenotype in all three parameters assessed (**Figure 4.6c,d,e**, column 5). See **Table 4.3** for the number of experimental replicates.

Overall, these findings indicate that preservation of an unaltered polyQ tract is important for rosette formation.

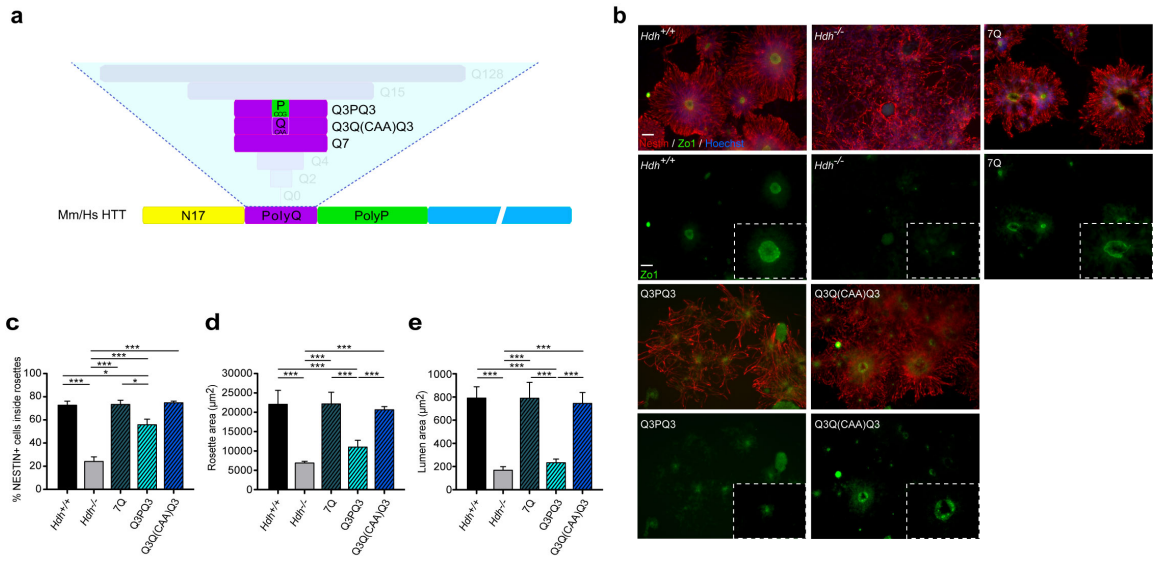


Figure 4.6. Effects of HTT polyQ tract interruption in an ES cell-derived neural rosette assay mimicking early neurulation. (**a**) Diagram of different HTT N-terminal portions assessed, in which the polyQ tract modifications (polyQ tract interruptions) are highlighted (**b**) Representative images of rosette/lumen phenotype in *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 7Q, Q3PQ3, and Q3Q(CAA)Q3 ES cells stained for NESTIN and ZO1 at day 7 of neural induction. (**c,d,e**) Percentage of NESTIN+ cells inside rosettes, rosette and lumen mean area in *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 7Q, Q3PQ3, and Q3Q(CAA)Q3 ES cells exposed to neural induction. Data are expressed as mean ± SEM. See **Table 4.3** for the number of biological replicates. All pairwise statistical comparisons were run by applying generalized linear mixed models with a *post hoc* Bonferroni correction (*: P<0.05; ***: P<0.001). The scale bars correspond to 50µm. Insets are shown at 2X magnification.

4.6 Pathological HTT polyQ tract impairs neurulation

To assess the activity of a pathologically expanded polyQ tract in the same assay, we exposed mES cells bearing a human N-terminal HTT construct with a pathological 128Q stretch (128Q cells) to neural induction (**Figure 4.7a,b)**. We found that the percentage of NESTIN+ cells inside the rosettes was drastically reduced in 128Q cells (29.6±8%), and to the same extent as observed in *Hdh*$^{-/-}$ cultures (24.0±9%), (**Figure 4.7c**, columns 2 and 4) but was preserved in mES cells bearing human HTT N-terminal fragment with normal 15Q stretch (15Q cells, 80.6±4%), as observed in control *Hdh*$^{+/+}$ cells (74.7±6%), (**Figure 4.7c**, columns 1 and 3). Rosette and lumen sizes were also significantly reduced by approximately 2-fold in 128Q cells (as occurred in *Hdh*$^{-/-}$ cells) compared to 15Q cells (**Figure 4.7d,e**, columns 2, 3 and 4). Data are expressed as mean ± SEM (see **Table 4.4** for the number of experimental replicates). On the whole, these findings showed that a pathological CAG expansion (128Q) affects HTT pro-neurulation function, suggesting that, during evolution, a CAG expansion beyond a certain threshold may have been detrimental.

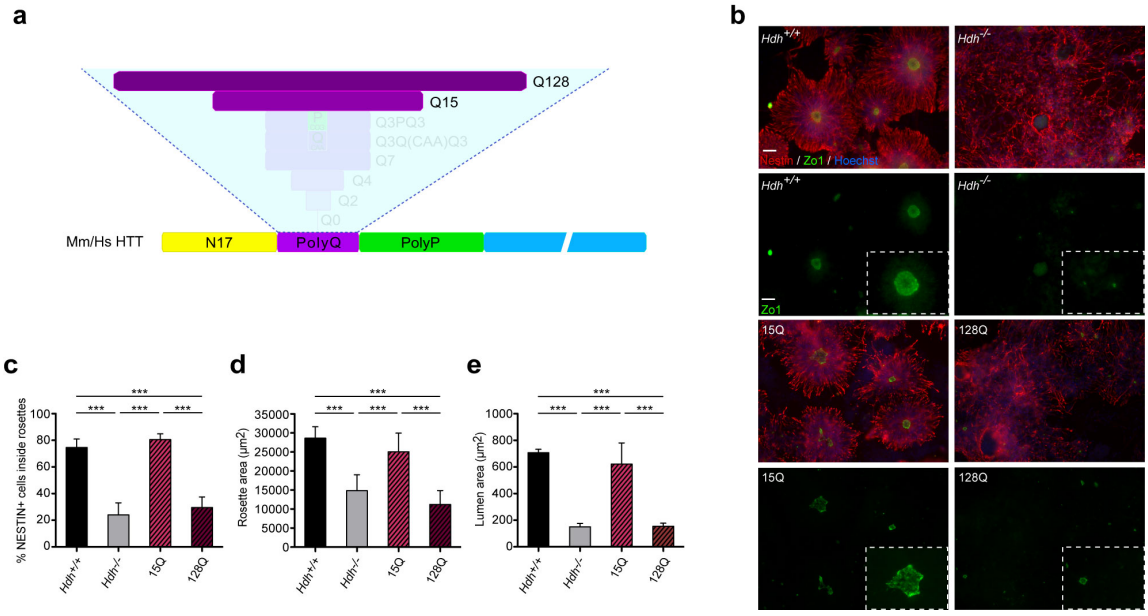


Figure 4.7. Effects of human HTT pathological polyQ tract in an ES cell-derived neural rosette assay mimicking early neurulation. (**a**) Diagram of different HTT N-terminal portions assessed, in which the polyQ tract modifications (polyQ tract with normal and pathological length) are highlighted (**b**) Representative images of rosette/lumen phenotype in *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 15Q, and 128Q ES cells stained for NESTIN and ZO1 at day 7 of neural induction. (**c,d,e**) Percentage of NESTIN+ cells inside rosettes, rosette and lumen mean area in *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 15Q, and 128Q ES cells exposed to neural induction. Data are expressed as mean ± SEM. See **Table 4.4** for the number of biological replicates. All pairwise statistical comparisons were run by applying generalized linear mixed models with a *post hoc* Bonferroni correction (***: P<0.001). The scale bars correspond to 50µm. Insets are shown at 2X magnification.

4.7 PolyQ domain is not involved in the antiapoptotic function of HTT

Since HTT protein and its N-terminal portion were previously associated with an antiapoptotic function (*Lo Sardo et al., 2012*), we evaluated whether *Hdh*$^{-/-}$ cells complemented with N-terminal HTT fragment with 0Q were protected from apoptosis under serum starvation. We first verified the previous data by confirming that Caspase-3 activity is increased in HTT-depleted (*Hdh*$^{-/-}$) cells compared to *Hdh*$^{+/+}$ (P value < 0.001; **Figure 4.8** columns 1 and 2), while the exogenous expression of HTT N-terminus with its normal Q in *Hdh*$^{-/-}$ cells (7Q cells) prevents cell death (**Figure 4.8** column 4). Notably, level of Caspase-3 activity observed in 0Q cells was similar to that occurred in *Hdh*$^{+/+}$ and 7Q cells (**Figure 4.8** column 1, 3 and 4). Therefore, the polyQ tract in HTT N-terminus is not required for the antiapoptotic function of the protein. Overall, our findings demonstrate that while the pro-neurulation function of HTT is linked to a persistent CAG tract, HTT antiapoptotic activity lays within other domains of the HTT N-terminus.

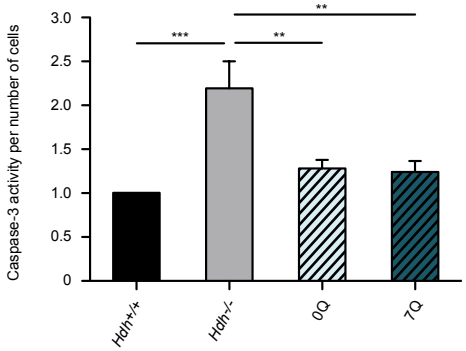


Figure 4.8. Effect of the polyQ tract deletion on the HTT anti-apoptotic function. Caspase-3 activity in proliferating *Hdh*$^{+/+}$, *Hdh*$^{-/-}$, 0Q and 7Q ES cells exposed to serum deprivation condition. Data are the mean ± SEM, n=10 independent experiments. Statistical pairwise comparisons were performed using a generalized linear mixed model with post hoc Bonferroni correction (**: P<0.01, ***: P<0.001).

% NESTIN+ cells inside rosettes

EXP	Hdh +/+	Hdh -/-	0Q	7Q
1	90,72	56,97	33,74	84,16
2	83,50	72,17	75,80	77,16
3	92,41	24,02	35,00	88,94
4	90,07	55,94	35,00	92,16
5	90,74	19,37	51,76	96,97
6	87,92	61,99	53,87	87,54
Mean	89,23	48,41	47,53	87,82
SD	3,16	21,53	16,50	6,79
SEM	1,29	8,79	6,74	2,77

Rosette area (μm^2)

EXP	Hdh +/+	Hdh -/-	0Q	7Q
1	18343,2	6818,0	8594,2	16203,4
2	14945,0	12202,6	6375,1	12537,2
3	27287,3	6173,4	11055,8	18181,0
4	18233,6	8937,8	12940,3	-
5	21792,2	5839,2	8899,8	22283,7
6	20951,9	12937,9	7768,3	19595,3
Mean	20258,87	8818,17	9272,22	17760,11
SD	4200,93	3108,38	2361,76	3664,35
SEM	1715,02	1268,99	964,18	1638,75

Lumen area (μm^2)

EXP	Hdh +/+	Hdh -/-	0Q	7Q
1	691,2	88,7	160,5	620,6
2	1138,0	148,9	428,7	679,9
3	458,3	151,0	244,4	1126,5
4	903,1	97,3	-	-
Mean	797,63	121,45	277,84	808,99
SD	290,64	33,10	137,20	276,52
SEM	145,32	16,55	79,21	159,65

Table 4.1. Subset of experiments to evaluate the effect of HTT polyQ tract deletion in rosette assay. *Top:* independent experiments analysed to measure the percentage of NESTIN+ cells inside rosette; *middle:* independent experiments analysed to measure the rosette size; *bottom:* independent experiments analysed to measure the lumen size.

% NESTIN+ cells inside rosettes

EXP	0Q	2Q	4Q	7Q
1	33,74	50,01	62,61	84,16
2	-	51,24	74,08	-
3	51,16	75,27	89,08	88,94
4	53,87	82,89	82,59	87,54
Mean	46,26	64,85	77,09	86,88
SD	10,93	16,73	11,44	2,45
SEM	6,31	8,36	5,72	1,42

Rosette area (μm^2) (g)

EXP	0Q	2Q	4Q	7Q
1	8594,2	9137,5	11866,0	16203,4
2	-	12770,5	12175,6	-
3	6375,1	-	-	12537,2
4	11055,8	7149,6	10760,5	18181,0
5	8899,8	-	-	22283,7
6	7768,3	13147,6	13694,6	19595,3
Mean	8538,62	10551,31	12124,19	17760,11
SD	1713,48	2900,36	1210,38	3664,34
SEM	766,29	1450,18	605,19	1638,74

Lumen area (μm^2) (h)

EXP	0Q	2Q	4Q	7Q
1	160,5	173,8	443,5	620,6
2	-	174,7	515,2	-
3	428,7	195,9	586,8	679,9
4	244,4	521,3	694,0	1126,5
Mean	277,84	266,43	559,86	808,99
SD	137,20	170,22	106,88	276,52
SEM	79,21	85,11	53,44	159,65

Table 4.2. Subset of experiments to evaluate the effect of HTT polyQ tract length in rosette assay. *Top:* independent experiments analysed to measure the percentage of NESTIN+ cells inside rosette; *middle:* independent experiments analysed to measure the rosette size; *bottom:* independent experiments analysed to measure the lumen size.

	EXP	Hdh +/+	Hdh -/-	7Q	Q3PQ3	Q3Q(CAA)Q3
% NESTIN+ cells inside rosettes	1	63,99	31,44	82,65	39,35	-
	2	58,90	45,90	68,89	76,63	-
	3	64,04	10,14	73,26	-	-
	4	66,41	11,31	82,63	56,03	-
	5	81,59	30,22	62,15	48,30	77,17
	6	65,11	35,04	-	42,08	-
	7	81,31	26,41	52,92	-	70,66
	8	93,61	5,45	81,60	70,41	73,01
	9	80,52	27,73	71,04	63,49	74,66
	10	71,50	16,77	85,23	50,75	78,58
	Mean	72,70	24,04	73,37	55,88	74,82
	SD	11,01	12,76	10,90	13,35	3,17
	SEM	3,48	4,03	3,63	4,72	1,42
	EXP	Hdh +/+	Hdh -/-	7Q	Q3PQ3	Q3Q(CAA)Q3
Rosette area (μm^2)	1	16614,0	5388,2	29834,0	10121,7	-
	2	15866,6	6529,8	-	8174,7	22407,0
	3	35614,8	7384,9	18162,5	16377,0	20691,4
	4	19409,0	8091,2	16565,40	13508,4	21041,4
	5	22765,4	6995,2	24093,7	6820,5	18477,5
	Mean	22053,97	6877,88	22163,91	11000,45	20654,31
	SD	8051,02	1010,37	6052,81	3918,75	1629,05
	SEM	3600,52	451,85	3026,40	1752,52	814,53
	EXP	Hdh +/+	Hdh -/-	7Q	Q3PQ3	Q3Q(CAA)Q3
Lumen area (μm^2)	1	719,2	145,7	1144,6	145,9	-
	2	651,1	234,9	-	282,2	988,5
	3	718,3	109,1	549,4	322,1	788,2
	4	1176,8	250,5	598,2	223,9	658,0
	5	692,1	103,6	866,2	192,9	543,1
	Mean	791,48	168,79	789,60	233,42	744,47
	SD	217,14	69,62	274,62	70,09	191,05
	SEM	97,11	31,14	137,31	31,34	95,53

Table 4.3. Subset of experiments to evaluate the effect of the purity interruption of the HTT polyQ tract in rosette assay. *Top:* independent experiments analysed to measure the percentage of NESTIN+ cells inside rosette; *middle:* independent experiments analysed to measure the rosette size; *bottom:* independent experiments analysed to measure the lumen size.

	EXP	Hdh +/+	Hdh -/-	15Q	128Q
% NESTIN+ cells inside rosettes	1	88,37	15,28	68,29	11,33
	2	64,01	4,28	84,84	39,93
	3	63,99	31,44	87,24	21,53
	4	82,23	45,17	82,04	45,44
	Mean	74,65	24,04	80,60	29,56
	SD	12,55	17,97	8,48	15,88
	SEM	6,27	8,98	4,24	7,94
	EXP	Hdh +/+	Hdh -/-	15Q	128Q
Rosette area (μm^2)	1	34195,0	22478,9	30928,3	18500,2
	2	23884,4	13832,2	28933,5	7106,7
	3	27803,8	8186,0	15214,8	8047,8
	Mean	28627,75	14832,38	25025,56	11218,22
	SD	5204,45	7198,75	8554,68	6323,90
	SEM	3004,79	4156,20	4939,04	3651,11
	EXP	Hdh +/+	Hdh -/-	15Q	128Q
Lumen area (μm^2)	1	722,8	124,9	548,3	110,1
	2	637,3	96,6	571,9	138,5
	3	760,9	215,9	1061,8	216,6
	4	708,5	163,9	307,2	154,9
	Mean	707,34	150,33	622,31	155,00
	SD	51,70	51,69	316,46	45,04
	SEM	25,85	25,85	158,23	22,52

Table 4.4. Subset of experiments to evaluate the effect of the pathological HTT polyQ tract in rosette assay. *Top:* independent experiments analysed to measure the percentage of NESTIN+ cells inside rosette; *middle:* independent experiments analysed to measure the rosette size; *bottom:* independent experiments analysed to measure the lumen size.

4.8 Automatization of rosette assay for high-content analysis

Identification and analysis of rosette formation are tedious and subjective procedures due to the acquisition of random fields and the manual quantification of rosettes and lumens. Thus, phenotype quantification could be directly influenced by the operator visual bias. For this reason we designed a novel automated method based on an image analysis algorithm for unbiased and quantitative assessment of rosette phenotypes. The relevance of this effort is to make our rosette assay a high-content imaging-based approach for classifying a larger and more diverse panel of phenotypes during *in vitro* neurulation. To this aim we focused, in the first step, on adopting the neural differentiation setup for high-content analyses and, subsequently, on developing an image analysis pipeline for rosette quantification.

4.8.1 Optimization of neural differentiation protocol procedures

Initially, we worked on the scaling down of the multi-well plate format, a parameter essential for high-content analysis. Moreover, the multi-well plates adopted have to be compatible with the automated equipment and imaging analysis. In the manual rosettes assay we used normal polystyrene 12-well plates coated with 0.1% gelatin. Polystyrene plates are ideal for cell attachment and subsequent differentiation, but they exhibit reduced optical properties compared to glass bottoms or polymer coverslip (optical plastic) plates. To overcome this issue, we tested several multi-well plates and, on the basis of the capacity of the cells to adhere and uniformly disperse in the well, we selected optical plastic 24-well plates (IBIDI, cod. 250210). These plates, in addition to permit cell adhesion, possess high-quality plastic bottoms ideal for image acquisition and analysis. Additionally, in order to allow ES cells to differentiate toward monodispersed neural rosettes (this is a condition more suitable for the automatic quantification), we also tested different basement membrane matrices and cell density conditions. In particular, after testing different substrate conditions (without substrate, with gelatin 0.05%, 0.1%, 0.2% and with matrigel substrate) and cell density (from $1x10^4/cm^2$ to $1{,}5x10^5/cm^2$) we selected gelatin 0.1% and a cell density of $2.5\text{-}3.5x10^4$ cells/cm^2 as the optimal neural differentiation condition.

For immunostaining of the neural rosettes in the automated assay, we used NESTIN (to identify neural progenitors) and ZO1 (to identify the tight junctions in the rosette lumen) markers, as in the manual method. However, when we tested new lumen markers to better identify the area of rosette apical site we revealed that among ZO1, PALS1, PAR3, and αPKC (four polarity markers targeting the tight junctions, described in 1.3.3) the antibodies against ZO1 and PALS1 were more suitable for our purposes due to their reduced aspecific binding (see **Figure 4.9a**) compared to the anti-PAR3 and -αPKC antibodies. Moreover, anti-PALS1 antibody gave a more intense and defined staining of the lumen compared to the antibody against ZO1 (see **Figure 4.9b**). Therefore, we decided to use PALS1 as a new marker for the identification of lumens in both manual and automated rosettes assay.

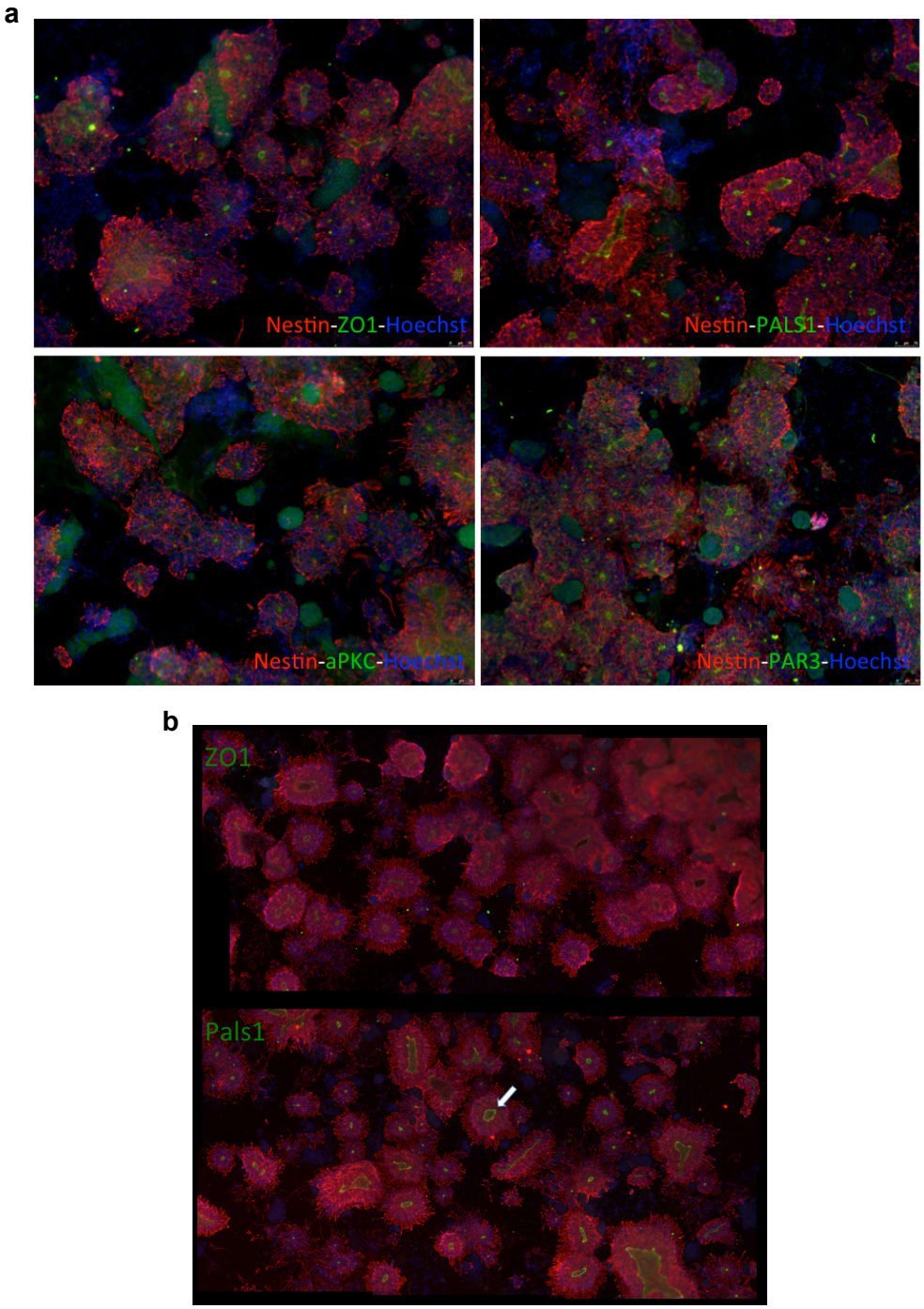


Figure 4.9. Comparison of four different antibodies against ZO1, PALS1, aPKC and PAR3 polarity markers. (a) Immunocytochemistry analysis on the *Hdh*$^{+/+}$ cell line at day 7 of neural differentiation carried out using antibodies against ZO1, Pals1, aPKC and PAR3 polarity markers (green) and against NESTIN (red). All fields were acquired with Leica DMI 6000 microscope at magnification 10X**. (b)** Representative image stitching of immunocytochemistry analysis on the *Hdh*$^{+/+}$ cell line at day 7 of neural differentiation carried out using antibodies against ZO1, Pals1 polarity markers and against NESTIN marker. The arrow indicates an example of intense and defined lumen staining.

4.8.2 Automatization of image acquisition and rosette quantification

An automated image acquisition pipeline was applied by using the IN Cell Analyzer 1000. This instrument is an automated cellular and subcellular imaging system for fast, automated multi-wavelength imaging and analysis in fixed and live cells. The powerful, easy-to-use IN Cell Analyzer 1000 enhances the productivity of cell-based approaches for high-content cellular analysis, by combining system flexibility with superior image and data quality. In particular, we have set up an acquisition of 40 random fields for each well at 10X magnification, in contrast with 4 random fields acquired during the manual method by using Leica DMI 6000 microscope.

A specific protocol for automated quantification was set up with IN Cell Developer Toolbox software. Given that rosettes are complex structures defined by multiple components (NESTIN+ area that covers the scaffold of the rosette, ZO1+ area that corresponds to the central lumen, and Hoechst positive nuclei that represent the number of cells in the rosette) and present heterogeneous sizes and shapes, a multi-step protocol is required. Indeed, for the protocol to work, the putative rosette must meet a series of criteria, each of which is identified in a single step of the pipeline. For convenience, we can divide the protocol in nine steps able to identify several target sets: steps from 1 to 3 are relevant for the identification, by Hoechst staining, of the rosette cells' nuclei; steps from 4 to step 7 permit to identify the lumens, using ZO1 as a marker; steps from 5 to 9 lead to rosette identification, using NESTIN as a marker. Below we report the individual steps of the pipeline set up.

Step 1: Identification of large Hoechst positive areas. In particular, here we proceeded with *object segmentation* followed by some image-processing algorithms, named *post-processing* steps, among which *dilation* (the effect of this operation is to gradually enlarge the boundaries of regions of the target set), *targets filled* (with which the software fills the holes inside the identified objects), and *sieve* (used to refine a segmented image by removing objects or artifacts belonging to a particular size range). On the whole, these *post-processing* operations allow to identify big clusters of nuclei that we will use in step 6 to separate the rosettes clumped together.

Step 2: Identification of Hoechst positive cells. Here we applied *object segmentation* followed by *erosion, a* function that strongly contracts the region's outer boundary of the target set in order to obtain the "seeds of the signal". These signal seeds provide a mask that we will apply in the next step.

Step 3: These are *post-processing* steps that include *sieve*, *erosion*, and *clump-breaking* that separates the clumps of objects using the mask generated in Step 2. After this, we will obtain the proper nuclei identification of the cells that forms the rosettes.

Step 4: Identification of lumens. We perform an *object segmentation* step followed by *sieve* and *dilation* as *post-processing* steps. Moreover, *SD-levels* (standard deviation of pixel densities within the target) and *Dens-levels* (mean signal level value of the pixels contained within the target) were registered to establish the proper values to use as *acceptance criteria.* Only the ZO1+ areas that satisfy the acceptance criteria values are recognized as lumens.

Step 5: Identification of total NESTIN signal area. We performed the detection process named *intensity segmentation* followed by *sieve* as *post-processing* step in order to identify the whole NESTIN signal area.

Step 6: Identification and split of the putative NESTIN+ rosettes. We performed *intensity segmentation* followed by *sieve post-processing* step and subsequently we applied *clump-breaking* function that separates the clumps of rosettes using the mask of large Hoechst positive areas identified in Step 1. We thus obtain the single putative NESTIN+ rosette.

Since the final targets of interest (neural rosettes) are composed by multiple elements, we performed a series of *target linking* operations (Steps 7 to 9) in order to create a link between targets identified and to pinpoint real rosettes that satisfy all criteria described above.

Step 7: Rosettes and lumen association. We performed an operation of *one to one* link between ZO1+ lumens identified in Step 4 and NESTIN+ rosettes identified in Step 6. By using 75% of overlap between ZO1+ lumens and NESTIN+ rosettes as criterion, we selected only one possible ZO1+ lumen for each single

rosette. After this step the number of all putative lumens identified in Step 4 decreases, ensuring a more precise identification of real lumens associated with rosettes.

Step 8: Rosette and nuclei association. We performed a *one to many* link between nuclei identified in Step 3 and NESTIN+ rosettes identified in Step 6. By using 80% of overlap between nuclei and NESTIN+ rosettes as criterion, we identified the nuclei that belong to each single rosette.

Step 9: Identification of real rosettes with nuclei and lumens. By using a *composed one to one* link between previously-linked target sets in Step 7 and 8, we identify only the rosettes that present both a lumen and the nuclei. Those are the rosettes that we consider as real rosettes and on which we focused our attention for the analysis.

Figure 4.10 recapitulates all steps, described above, and demonstrates the efficiency of the pipeline in recognizing neural rosettes and their associated parameters. Particularly, from this multi-step analysis we get a series of parameters related to the three targets identified (nuclei in Step 3; lumens and rosettes in Step 7 and 9, respectively) as well as parameters from several combinations of these three targets. Indeed, while the manual rosette assay quantified with ImageJ analysis software is able to provide only 3 parameters (percentage of NESTIN+ cells inside rosette, rosette size and lumen size) through the automated rosette assay we increase the number of parameters measured from 3 to 13. In particular, we have the possibility to measure:

a. Neural induction (neural conversion, total NESTIN+ signal area over Dapi+ cells)
b. Percentage of NESTIN+ cells inside rosettes
c. Number of rosettes
d. Rosette size (area)
e. Rosette perimeter
f. Rosette shape (ratio between rosette area and perimeter)
g. Rosette cell density (number of Dapi+ cells inside a rosette over its area)
h. Lumen formation (total ZO1+ signal area over total NESTIN+ signal area)
i. Lumen size (area)
j. Lumen perimeter
k. Lumen fiber length (total length within a single fibrous shape)
l. Lumen max chord (maximum distance across the inside of a target)
m. Lumen form factor (estimate of circularity)

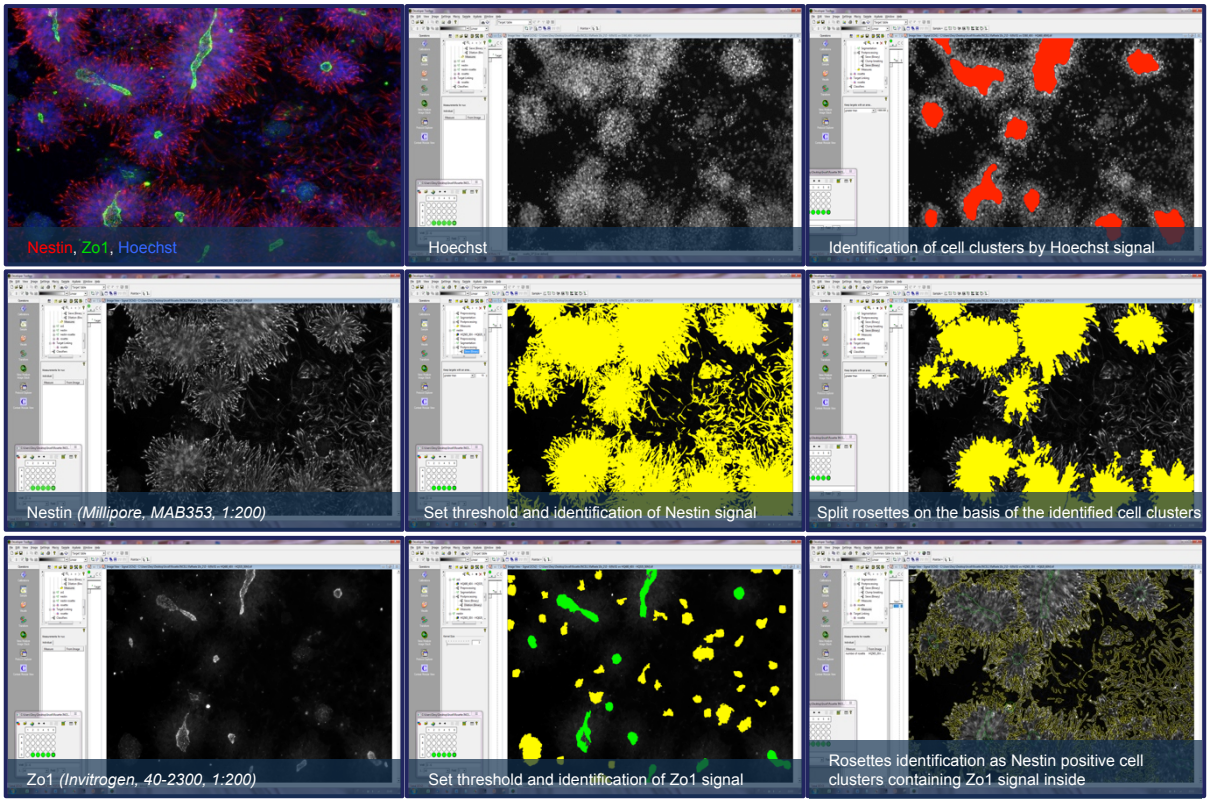
Additionally, the automated rosette assay allows to increase 10-20 times the amount of data collected with respect to manual counting. Indeed, the automated system allows to measure the parameters from about 1000-2000 rosettes for each experiment and for each cell line, while the manual quantification allows to collect data from approximately 100 rosettes for each cell line.

On the whole, this unbiased approach permits to further improve rosette acquisition/quantification and to strongly increase the rosette measurements dataset, leading to development of an automated assay for high-content analysis of *in vitro* neurulation.

4.8.3 Evaluation of rosette phenotype by automated method

In order to test the new automated method and validate the previous manually quantified rosette phenotypes, we performed 4 independent neural differentiation experiments with $Hdh^{+/+}$ and $Hdh^{-/-}$ cells. We confirmed the defects in rosette formation in $Hdh^{-/-}$ cells with respect to $Hdh^{+/+}$ by assessing the *percentage of NESTIN+ cells inside rosettes* ($P<0.001$; **Figure 4.11b**). Moreover, the mean *rosette area* was approximately 40% smaller in $Hdh^{-/-}$ cells compared to $Hdh^{+/+}$ cells, a result that mirrors the previously acquired manual data ($P<0.001$; **Figure 4.11d**). We also evaluated the lumen of the rosettes, and we confirmed a significant reduction of the mean *lumen size* that was approximately 3 times smaller in $Hdh^{-/-}$ cells compared to $Hdh^{+/+}$ cells ($P<0.001$; **Figure 4.11i**). Overall these findings corroborated the ones obtained with manual quantification.

As described before, the automated rosette quantification allows to measure a series of additional parameters. In particular, we observed that *neural induction, number of rosettes*, *rosette perimeter, lumen formation*, *lumen perimeter*, *lumen fiber length* and *lumen max chord were* significantly reduced in $Hdh^{-/-}$ cultures compared to $Hdh^{+/+}$ cultures (all P values were <0.001; **Figure 4.11a,c,e,h,j,k,l,** respectively). Instead, there is a significant increase in the *lumen form factor* parameter in $Hdh^{-/-}$ cells with respect to $Hdh^{+/+}$ cells ($P<0.01$; **Figure 4.11m**). No significant difference was detected in the *rosette shape* and *rosette cell density* parameters (**Figure 4.11f,g**). Besides validating the rosette phenotype in $Hdh^{-/-}$ cells, the automated procedure can extend the spectrum of parameters that we can assess in the rosette analysis, therefore increasing exponentially the power of our analyses.



b

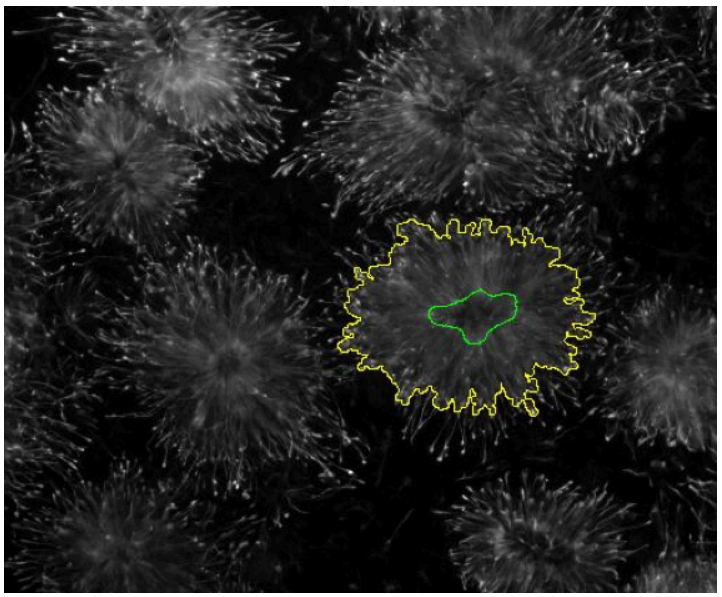


Figure 4.10. Automated rosette quantification method. (a) Sequence of the steps developed for the automated rosette quantification by IN Cell Developer Toolbox software. **(b)** Example of automated rosette identification by IN Cell Developer Toolbox software. Yellow line indicates the counter of the rosette; green line indicates the counter of the lumen.

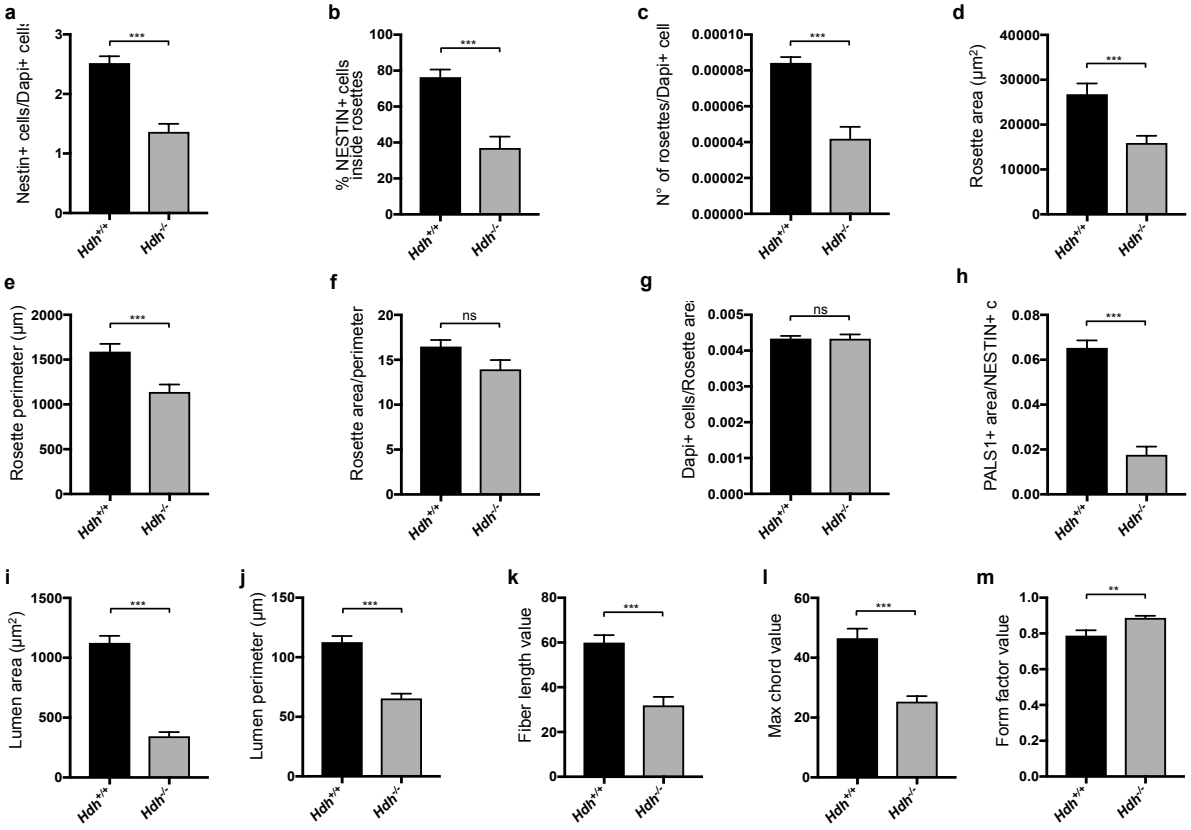


Figure 4.11. Quantification of rosette phenotype by automated method. Quantification of rosette and lumen parameters in *Hdh*$^{+/+}$ and *Hdh*$^{-/-}$ cells by using the automated procedure: (**a**) neural induction, (**b**) percentage of NESTIN+ cells inside the rosettes, (**c**) number of rosettes, (**d**) rosette size, (**e**) rosette perimeter, (**f**) rosette shape, (**g**) rosette cell density, (**h**) lumen formation, (**i**) lumen size, (**j**) lumen perimeter, (**k**) lumen fiber length, (**l**) lumen max chord, and (**m**) lumen form factor. Data are the mean ± SEM, n=4 independent experiments (*T-test*; **: P<0.01; ***: P<0.001; ns: not significant).

5. DISCUSSION AND CONCLUSIONS

Previous studies showed that HTT is important for rosette formation as cultures of *Hdh*$^{-/-}$ mES cells are severely deficient in rosettes formation and only show small sized rosettes (*Lo Sardo et al., 2012*). Our complementation assays showed that constitutive expression of exogenous normal full length mouse HTT or its N-terminal 548aa portion, in *Hdh*$^{-/-}$ cells restores proper rosette formation and size (*Lo Sardo et al., 2012*). While these findings have associated the ability to form neural rosettes to the HTT N-terminal portion, here we show, for the first time, a direct link between the HTT polyQ tract and the capacity of the cells to form proper rosettes. Indeed, the deletion of the sole Q domain from HTT N-terminus leads to a reduction in the neural rosette formation potential, as demonstrated by the quantification of the percentage of neural progenitors organized in rosettes as well as by the size of the rosettes and their lumens (Figure 4.4). Although we provide evidences of a role of the HTT polyQ tract during *in vitro* neurulation, an open question still remains on the functional role of the *wild-type* polyQ tract *in vivo* during the early phases of neural development. Studies carried out by S. Zeitlin's group have shown that the *Hdh*$^{\Delta Q\text{-}\Delta Q}$ knock-in mice are born with normal Mendelian frequency with no gross phenotypic differences with respect to control littermates. However, they exhibit subtle behavioural phenotypes in motor and cognitive tests (*Clabough and Zeitlin, 2006*). This may be an indication of other unknown molecular defects or of the presence of *in vivo* compensatory effect that still need to be investigated. Therefore, the rosette assay captures only a small snapshot of what happens *in vivo* and may not give the whole picture of the complex biological systems surrounding nervous system development. However, it can zoom in and analyse with a high degree of accuracy the binary connection between HTT function and its evolutionary role in nervous system development.

We have also observed that modulation of polyQ length regulates neural progenitor organization in a dose-dependent manner, since its elongation from 0 to 7Q improves the ability of ES cell-derived neural progenitors to form rosettes. The finding that rosette formation potential linearly covaries with CAG length (Figure 4.5) suggests that the number of glutamines in the polyQ domain may act as a 'fine-tuning knob' of the HTT-dependent rosette formation mechanisms. These data are in line with the hypothesis of Fondon and colleagues, that states that the trinucleotide repeats could play a relevant and potentially beneficial role in evolution. These repeats, by virtue of their special mutational and functional properties, modify genes with which they are associated, thus generating the genetic variation that underlies adaptive evolution (*King et al., 1997; Fondon et al., 2008*). More specifically, the mutational expansion of certain triplet repeats is implicated not just in hereditary neurodegenerative disease but also in normal brain development and behavioural traits (*Fondon et al., 2008*). Our observation of a polyQ length-dependent neurulation activity is also in line with previous indications derived from complementation experiments with HTT homologs bearing CAG tracts that increase in length during human lineage evolution. These findings indicated that cells expressing heterologous HTT fragments from progressively more recently evolved species show an increased rosette formation potential (*Lo Sardo et al., 2012*). However, on the basis of data from *Lo Sardo et al., 2012* we cannot directly link the ability to modulate neural rosette formation with polyQ length, as it is not the only element to differ in the tested heterologous HTT fragments. On the contrary, in this thesis, we exclusively used mouse HTT fragments bearing 0, 2, 4 and 7Q thus focusing on the effect of the sole polyQ tract. This allows us to assign a direct relationship between polyQ length and rosette formation potential.

As the polyQ tract is conserved throughout evolution (*Tartari et al., 2008*), we tested whether an amino acid interruption in the polyQ sequence is compatible with HTT neurogenic potential. We observed that a CAA synonymous substitution does not affect the rosette formation phenotype (Figure 4.6). In contrast, the P interruption in the same position generates ES cell derived neural progenitors cultures with fewer and smaller rosette and lumen area (Figure 4.6). These findings underline the need to preserve the polyQ tract unaltered, as the interruption of its purity by a proline results in a reduction of HTT pro-neurulation potential.

Our experiments also show that a 128Q domain (bearing a pathological polyQ length) in HTT N-terminus determines a reduced rosette formation potential, with a significant decrease in the ability of neural progenitors to self-organize in rosettes together with a reduction of lumen and rosette areas (Figure 4.7). Collectively, the results described here suggest that there may be a neurodevelopmental component in Huntington Disease

(HD). This is in agreement with data from Molero and colleagues (*Molero et al., 2016*) in which the expression of mutant HTT exclusively until postnatal day 21 in HD mice displayed similar profiles of impairments to mice expressing mutant HTT throughout life, advancing the hypothesis that neurodevelopmental aberrations may play important roles in HD pathogenesis and progression (*Molero et al., 2016*). Accordingly, smaller intracranial volume in individuals with prodromal HD provides a further *in vivo* evidence for abnormal neurodevelopment in HD (*Nopoulos et al., 2011b*).

Finally, we show that HTT molecules carrying either 0 or 7Q upon expression in $Hdh^{-/-}$ cells are all equally capable of conferring protection from apoptosis after serum starvation (Figure 4.8). These results revealed that, while the pro-neurulation function of HTT is linked to the CAG repeat tract, the antiapoptotic activity of HTT lies within other domains of its N-terminus.

In conclusion, in this study we show, for the first time, that the stretch of Q in HTT is biologically active during early neurogenesis since its deletion, its pathological expansion or its interruption directly impairs HTT pro-neurulation activity, while its elongation in the normal range (from 0 to 7Q) progressively increases the neurogenic power of this protein.

The main readout of our experiments is the quantitative and qualitative evaluation of rosette formation, however the current methodology is a time consuming procedure subject to an examiner's bias. By adopting an automated image acquisition procedure with the IN Cell Analyzer 1000 we developed a cell-based high-content analysis platform to better assess a different spectrum of parameters surrounding rosette formation (Figure 4.10 and 4.11). Moreover, the same platform could be also exploited to investigate novel HTT phenotypic targets.

In this thesis we worked with HTT null ES cells overexpressing several N-terminal portions of the protein. In order to further assess the functional relevance of these domains in a more physiological condition, we are currently generating knock-in mES cell line bearing a Recombinase Mediated Cassette Exchange (RMCE) by CRISPR-Cas9 technology. This RMCE cassette spans the promoter, the first exon and part of the first intron of the *HTT* gene allowing efficient insertions of any mutated *HTT* exon 1 in the endogenous locus. Therefore, this cassette has two main advantages compared to the overexpressing fragments used up till now: the first is that it allows expression of the full-length HTT containing the variants of interest; the second, is that it integrates in the endogenous locus and therefore is exposed to the natural genomic environment of HTT allowing a better reproduction of the HTT aura and its physiological expression. In a future series of analyses we also plan to investigate other potential contributions of the polyQ tract in diverse HTT activities as well as the role of the N17 and PRD domains in neural and non-neural HTT functions.

6. REFERENCES

Abranches, E., Silva, M., Pradier, L., Schulz, H., Hummel, O., Henrique, D., and Bekman, E. (2009). Neural Differentiation of Embryonic Stem Cells In Vitro: A Road Map to Neurogenesis in the Embryo. PLOS ONE *4*, e6286.

ACMG/ASHG HD Genetic Testing Working Group (1998). The American College of Medical Genetics/American Society of Human Genetics Huntington Disease Genetic Testing Working Group. Laboratory guidelines for Huntington disease genetic testing. Am J Hum Genet *62*, 1243–1247.

Almqvist, E., Andrew, S., Theilmann, J., Goldberg, P., Zeisler, J., Drugge, U., Grandell, U., Tapper-Persson, M., Winblad, B., and Hayden, M. (1994). Geographical distribution of haplotypes in Swedish families with Huntington's disease. Hum. Genet. *94*, 124–128.

Almqvist, E.W., Elterman, D.S., MacLeod, P.M., and Hayden, M.R. (2001). High incidence rate and absent family histories in one quarter of patients newly diagnosed with Huntington disease in British Columbia. Clin. Genet. *60*, 198–205.

Andrade, M.A., and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. Nat. Genet. *11*, 115–116.

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. J. Mol. Biol. *309*, 1–18.

André, E.A., Braatz, E.M., Liu, J.-P., and Zeitlin, S.O. (2017). Generation and Characterization of Knock-in Mouse Models Expressing Versions of Huntingtin with Either an N17 or a Combined PolyQ and Proline-Rich Region Deletion. J Huntingtons Dis *6*, 47–62.

Aronin, N., Chase, K., Young, C., Sapp, E., Schwarz, C., Matta, N., Kornreich, R., Landwehrmeyer, B., Bird, E., and Beal, M.F. (1995). CAG expansion affects the expression of mutant Huntingtin in the Huntington's disease brain. Neuron *15*, 1193–1201.

Atwal, R.S., and Truant, R. (2008). A stress sensitive ER membrane-association domain in Huntingtin protein defines a potential role for Huntingtin in the regulation of autophagy. Autophagy *4*, 91–93.

Atwal, R.S., Xia, J., Pinchev, D., Taylor, J., Epand, R.M., and Truant, R. (2007). Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity. Hum. Mol. Genet. *16*, 2600–2615.

Atwal, R.S., Desmond, C.R., Caron, N., Maiuri, T., Xia, J., Sipione, S., and Truant, R. (2011). Kinase inhibitors modulate huntingtin cell localization and toxicity. Nat Chem Biol *7*, 453–460.

Aubert, J., Dunstan, H., Chambers, I., and Smith, A. (2002). Functional gene screening in embryonic stem cells implicates Wnt antagonism in neural differentiation. Nat. Biotechnol. *20*, 1240–1245.

Auerbach, W., Hurlbert, M.S., Hilditch-Maguire, P., Wadghiri, Y.Z., Wheeler, V.C., Cohen, S.I., Joyner, A.L., MacDonald, M.E., and Turnbull, D.H. (2001). The HD mutation causes progressive lethal neurological disease in mice expressing reduced levels of huntingtin. Hum. Mol. Genet. *10*, 2515–2523.

Auernhammer, C.J., and Melmed, S. (2000). Leukemia-inhibitory factor-neuroimmune modulator of endocrine function. Endocr. Rev. *21*, 313–345.

Aziz, N.A., Jurgens, C.K., Landwehrmeyer, G.B., EHDN Registry Study Group, van Roon-Mom, W.M.C., van Ommen, G.J.B., Stijnen, T., and Roos, R. a. C. (2009). Normal and mutant HTT interact to affect clinical severity and progression in Huntington disease. Neurology *73*, 1280–1285.

Bain, G., Kitchens, D., Yao, M., Huettner, J.E., and Gottlieb, D.I. (1995). Embryonic stem cells express neuronal properties in vitro. Dev. Biol. *168*, 342–357.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-

implantation epiblast to pluripotent embryonic stem cells. Nature *461*, 1292–1295.

Baquet, Z.C., Gorski, J.A., and Jones, K.R. (2004). Early Striatal Dendrite Deficits followed by Neuron Loss with Advanced Age in the Absence of Anterograde Cortical Brain-Derived Neurotrophic Factor. J. Neurosci. *24*, 4250–4258.

Beale, D., and Iber, R. (2006). Somatic evolution of antibody genes. In The Implicit Genome (Caporale, L.H., Ed.), (Oxford University Press), pp. 177–190.

Bečanović, K., Nørremølle, A., Neal, S.J., Kay, C., Collins, J.A., Arenillas, D., Lilja, T., Gaudenzi, G., Manoharan, S., Doty, C.N., et al. (2015). A SNP in the HTT promoter alters NF-κB binding and is a bidirectional genetic modifier of Huntington disease. Nat. Neurosci. *18*, 807–816.

Benn, C.L., Landles, C., Li, H., Strand, A.D., Woodman, B., Sathasivam, K., Li, S.-H., Ghazi-Noori, S., Hockly, E., Faruque, S.M.N.N., et al. (2005). Contribution of nuclear and extranuclear polyQ to neurological phenotypes in mouse models of Huntington's disease. Hum Mol Genet *14*, 3065–3078.

Bennett, A.J., Lesch, K.P., Heils, A., Long, J.C., Lorenz, J.G., Shoaf, S.E., Champoux, M., Suomi, S.J., Linnoila, M.V., and Higley, J.D. (2002). Early experience and serotonin transporter gene variation interact to influence primate CNS function. Mol. Psychiatry *7*, 118–122.

Beste, C., Wascher, E., Dinse, H.R., and Saft, C. (2012). Faster perceptual learning through excitotoxic neurodegeneration. Curr. Biol. *22*, 1914–1917.

Bethea, C.L., Streicher, J.M., Coleman, K., Pau, F.K.-Y., Moessner, R., and Cameron, J.L. (2004). Anxious behavior and fenfluramine-induced prolactin secretion in young rhesus macaques with different alleles of the serotonin reuptake transporter polymorphism (5HTTLPR). Behav. Genet. *34*, 295–307.

Bhattacharyya, A., Thakur, A.K., Chellgren, V.M., Thiagarajan, G., Williams, A.D., Chellgren, B.W., Creamer, T.P., and Wetzel, R. (2006). Oligoproline effects on polyglutamine conformation and aggregation. J. Mol. Biol. *355*, 524–535.

Bhide, P.G., Day, M., Sapp, E., Schwarz, C., Sheth, A., Kim, J., Young, A.B., Penney, J., Golden, J., Aronin, N., et al. (1996). Expression of normal and mutant huntingtin in the developing brain. J. Neurosci. *16*, 5523–5535.

Bibel, M., Richter, J., Schrenk, K., Tucker, K.L., Staiger, V., Korte, M., Goetz, M., and Barde, Y.-A. (2004). Differentiation of mouse embryonic stem cells into a defined neuronal lineage. Nat Neurosci *7*, 1003–1009.

Blanpain, C., Horsley, V., and Fuchs, E. (2007). Epithelial Stem Cells: Turning over New Leaves. Cell *128*, 445–458.

Bourillot, P.-Y., Aksoy, I., Schreiber, V., Wianny, F., Schulz, H., Hummel, O., Hubner, N., and Savatier, P. (2009). Novel STAT3 target genes exert distinct roles in the inhibition of mesoderm and endoderm differentiation in cooperation with Nanog. Stem Cells *27*, 1760–1771.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev. Biol. *210*, 30–43.

van der Burg, J.M.M., Björkqvist, M., and Brundin, P. (2009). Beyond the brain: widespread pathology in Huntington's disease. Lancet Neurol *8*, 765–774.

Cai, C., and Grabel, L. (2007). Directing the differentiation of embryonic stem cells to neural stem cells. Dev. Dyn. *236*, 3255–3266.

Candiani, S., Pestarino, M., Cattaneo, E., and Tartari, M. (2007). Characterization, developmental expression and

evolutionary features of the huntingtin gene in the amphioxus Branchiostoma floridae. BMC Dev. Biol. *7*, 127.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., et al. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science *301*, 386–389.

Chalasani, K., and Brewster, R.M. (2011). N-cadherin-mediated cell adhesion restricts cell proliferation in the dorsal neural tube. Mol. Biol. Cell *22*, 1505–1515.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311–2322.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell *113*, 643–655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Champoux, M., Bennett, A., Shannon, C., Higley, J.D., Lesch, K.P., and Suomi, S.J. (2002). Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. Mol. Psychiatry *7*, 1058–1063.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008a). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Chong, S.S., Almqvist, E., Telenius, H., LaTray, L., Nichol, K., Bourdelat-Parks, B., Goldberg, Y.P., Haddad, B.R., Richards, F., Sillence, D., et al. (1997). Contribution of DNA sequence and CAG size to mutation frequencies of intermediate alleles for Huntington disease: evidence from single sperm analyses. Hum. Mol. Genet. *6*, 301–309.

Choo, Y.S., Johnson, G.V.W., MacDonald, M., Detloff, P.J., and Lesort, M. (2004). Mutant huntingtin directly increases susceptibility of mitochondria to the calcium-induced permeability transition and cytochrome c release. Hum. Mol. Genet. *13*, 1407–1420.

Choudhry, S., Mukerji, M., Srivastava, A.K., Jain, S., and Brahmachari, S.K. (2001). CAG repeat instability at SCA2 locus: anchoring CAA interruptions and linked single nucleotide polymorphisms. Hum. Mol. Genet. *10*, 2437–2446.

Clabough, E.B.D., and Zeitlin, S.O. (2006). Deletion of the triplet repeat encoding polyglutamine within the mouse Huntington's disease gene results in subtle behavioral/motor phenotypes in vivo and elevated levels of ATP with cellular senescence in vitro. Hum. Mol. Genet. *15*, 607–623.

Cleary, J.D., and Pearson, C.E. (2003). The contribution of cis-elements to disease-associated repeat instability: clinical and experimental evidence. Cytogenet. Genome Res. *100*, 25–55.

Cleary, J.D., and Pearson, C.E. (2005). Replication fork dynamics and dynamic mutations: the fork-shift model of repeat instability. Trends Genet. *21*, 272–280.

Cleary, J.D., Nichol, K., Wang, Y.-H., and Pearson, C.E. (2002). Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nat. Genet. *31*, 37–46.

Comings, D.E. (1998). Polygenic inheritance and micro/minisatellites. Mol. Psychiatry *3*, 21–31.

Costa, M. do C., Magalhães, P., Ferreirinha, F., Guimarães, L., Januário, C., Gaspar, I., Loureiro, L., Vale, J., Garrett, C., Regateiro, F., et al. (2003). Molecular diagnosis of Huntington disease in Portugal: implications for genetic counselling and clinical practice. Eur. J. Hum. Genet. *11*, 872–878.

Culver, B.P., Savas, J.N., Park, S.K., Choi, J.H., Zheng, S., Zeitlin, S.O., Yates, J.R., and Tanese, N. (2012). Proteomic Analysis of Wild-type and Mutant Huntingtin-associated Proteins in Mouse Brains Identifies Unique Interactions and Involvement in Protein Synthesis. J. Biol. Chem. *287*, 21599–21614.

Dahéron, L., Opitz, S.L., Zaehres, H., Lensch, M.W., Lensch, W.M., Andrews, P.W., Itskovitz-Eldor, J., and Daley, G.Q. (2004). LIF/STAT3 signaling fails to maintain self-renewal of human embryonic stem cells. Stem Cells *22*, 770–778.

Dailey, L., Ambrosetti, D., Mansukhani, A., and Basilico, C. (2005). Mechanisms underlying differential responses to FGF signaling. Cytokine Growth Factor Rev. *16*, 233–247.

Darnell, G., Orgel, J.P.R.O., Pahl, R., and Meredith, S.C. (2007). Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. J. Mol. Biol. *374*, 688–704.

De Rooij, K.E., Dorsman, J.C., Smoor, M.A., Den Dunnen, J.T., and Van Ommen, G.J. (1996). Subcellular localization of the Huntington's disease gene product in cell lines by immunofluorescence and biochemical subcellular fractionation. Hum. Mol. Genet. *5*, 1093–1099.

Dehay, B., and Bertolotti, A. (2006). Critical role of the proline-rich region in Huntingtin for aggregation and cytotoxicity in yeast. J. Biol. Chem. *281*, 35608–35615.

Desmond, C.R., Atwal, R.S., Xia, J., and Truant, R. (2012). Identification of a karyopherin β1/β2 proline-tyrosine nuclear localization signal in huntingtin protein. J. Biol. Chem. *287*, 39626–39633.

Detrick, R.J., Dickey, D., and Kintner, C.R. (1990). The effects of N-cadherin misexpression on morphogenesis in Xenopus embryos. Neuron *4*, 493–506.

Diekmann, H., Anichtchik, O., Fleming, A., Futter, M., Goldsmith, P., Roach, A., and Rubinsztein, D.C. (2009). Decreased BDNF Levels Are a Major Contributor to the Embryonic Phenotype of Huntingtin Knockdown Zebrafish. J. Neurosci. *29*, 1343–1349.

Dietrich, P., Shanmugasundaram, R., Shuyu, E., and Dragatsis, I. (2009). Congenital hydrocephalus associated with abnormal subcommissural organ in mice lacking huntingtin in Wnt1 cell lineages. Human Molecular Genetics *18*, 142–150.

DiFiglia, M., Sapp, E., Chase, K., Schwarz, C., Meloni, A., Young, C., Martin, E., Vonsattel, J.P., Carraway, R., and Reeves, S.A. (1995). Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. Neuron *14*, 1075–1081.

Dorsman, J.C., Smoor, M.A., Maat-Schieman, M.L., Bout, M., Siesling, S., van Duinen, S.G., Verschuuren, J.J., den Dunnen, J.T., Roos, R.A., and van Ommen, G.J. (1999). Analysis of the subcellular localization of huntingtin with a set of rabbit polyclonal antibodies in cultured mammalian cells of neuronal origin: comparison with the distribution of huntingtin in Huntington's disease autopsy brain. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *354*, 1061–1067.

Dragatsis, I., Efstratiadis, A., and Zeitlin, S. (1998). Mouse mutant embryos lacking huntingtin are rescued from lethality by wild-type extraembryonic tissues. Development *125*, 1529–1539.

Dragatsis, I., Levine, M.S., and Zeitlin, S. (2000). Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. Nat. Genet. *26*, 300–306.

Duennwald, M.L., Jagadish, S., Muchowski, P.J., and Lindquist, S. (2006). Flanking sequences profoundly alter polyglutamine toxicity in yeast. Proc. Natl. Acad. Sci. U.S.A. *103*, 11045–11050.

Duyao, M.P., Auerbach, A.B., Ryan, A., Persichetti, F., Barnes, G.T., McNeil, S.M., Ge, P., Vonsattel, J.P., Gusella, J.F., Joyner, A.L., et al. (1995). Inactivation of the mouse Huntington's disease gene homolog Hdh. Science *269*, 407–410.

Ehrnhoefer, D.E., Butland, S.L., Pouladi, M.A., and Hayden, M.R. (2009). Mouse models of Huntington disease: variations on a theme. Dis Model Mech *2*, 123–129.

El‐Daher, M.-T., Hangen, E., Bruyère, J., Poizat, G., Al‐Ramahi, I., Pardo, R., Bourg, N., Souquere, S., Mayet, C., Pierron, G., et al. (2015). Huntingtin proteolysis releases non‐polyQ fragments that cause toxicity through

dynamin 1 dysregulation. The EMBO Journal *34*, 2255–2271.

Elias, S., Thion, M.S., Yu, H., Sousa, C.M., Lasgi, C., Morin, X., and Humbert, S. (2014). Huntingtin Regulates Mammary Stem Cell Division and Differentiation. Stem Cell Reports *2*, 491–506.

Elias, S., McGuire, J.R., Yu, H., and Humbert, S. (2015). Huntingtin Is Required for Epithelial Polarity through RAB11A-Mediated Apical Trafficking of PAR3-aPKC. PLOS Biology *13*, e1002142.

Elkabetz, Y., and Studer, L. (2008). Human ESC-derived neural rosettes and neural stem cell progression. Cold Spring Harb. Symp. Quant. Biol. *73*, 377–387.

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N.D., Tabar, V., and Studer, L. (2008). Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. Genes Dev. *22*, 152–165.

Ellegren, H. (2004). Microsatellites: simple sequences with complex evolution. Nat. Rev. Genet. *5*, 435–445.

Erwin, D.H., Laflamme, M., Tweedt, S.M., Sperling, E.A., Pisani, D., and Peterson, K.J. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science *334*, 1091–1097.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Fink, S., Excoffier, L., and Heckel, G. (2006). Mammalian monogamy is not controlled by a single gene. Proc. Natl. Acad. Sci. U.S.A. *103*, 10956–10960.

Flexner (1891). A peculiar glioma (neuroepithelioma?) of the retina. Bulletin of the Johns Hopkins Hospital.

Fondon, J.W., and Garner, H.R. (2004). Molecular origins of rapid and continuous morphological evolution. Proc. Natl. Acad. Sci. U.S.A. *101*, 18058–18063.

Fondon, J.W., Hammock, E.A.D., Hannan, A.J., and King, D.G. (2008). Simple sequence repeats: genetic modulators of brain function and behavior. Trends Neurosci. *31*, 328–334.

Fujimori, T., Miyatani, S., and Takeichi, M. (1990). Ectopic expression of N-cadherin perturbs histogenesis in Xenopus embryos. Development *110*, 97–104.

Fusco, F.R., Chen, Q., Lamoreaux, W.J., Figueredo-Cardenas, G., Jiao, Y., Coffman, J.A., Surmeier, D.J., Honig, M.G., Carlock, L.R., and Reiner, A. (1999). Cellular localization of huntingtin in striatal and cortical neurons in rats: lack of correlation with neuronal vulnerability in Huntington's disease. J. Neurosci. *19*, 1189–1202.

Gafni, J., and Ellerby, L.M. (2002). Calpain activation in Huntington's disease. J. Neurosci. *22*, 4842–4849.

Gafni, J., Hermel, E., Young, J.E., Wellington, C.L., Hayden, M.R., and Ellerby, L.M. (2004). Inhibition of calpain cleavage of huntingtin reduces toxicity: accumulation of calpain/caspase fragments in the nucleus. J. Biol. Chem. *279*, 20211–20220.

Gauthier, L.R., Charrin, B.C., Borrell-Pagès, M., Dompierre, J.P., Rangone, H., Cordelières, F.P., De Mey, J., MacDonald, M.E., Lessmann, V., Humbert, S., et al. (2004). Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. Cell *118*, 127–138.

Gerber, H.P., Seipel, K., Georgiev, O., Höfferer, M., Hug, M., Rusconi, S., and Schaffner, W. (1994). Transcriptional activation modulated by homopolymeric glutamine and proline stretches. Science *263*, 808–811.

Gerhartz, C., Heesel, B., Sasse, J., Hemmann, U., Landgraf, C., Schneider-Mergener, J., Horn, F., Heinrich, P.C., and Graeve, L. (1996). Differential activation of acute phase response factor/STAT3 and STAT1 via the cytoplasmic domain of the interleukin 6 signal transducer gp130. I. Definition of a novel phosphotyrosine motif mediating STAT1 activation. J. Biol. Chem. *271*, 12991–12998.

Germain, N., Banda, E., and Grabel, L. (2010). Embryonic stem cell neurogenesis and neural specification. J. Cell. Biochem. *111*, 535–542.

Gervais, F.G., Singaraja, R., Xanthoudakis, S., Gutekunst, C.-A., Leavitt, B.R., Metzler, M., Hackam, A.S., Tam, J., Vaillancourt, J.P., Houtzager, V., et al. (2002). Recruitment and activation of caspase-8 by the Huntingtin-interacting protein Hip-1 and a novel partner Hippi. Nat. Cell Biol. *4*, 95–105.

Gissi, C., Pesole, G., Cattaneo, E., and Tartari, M. (2006). Huntingtin gene evolution in Chordata and its peculiar features in the ascidian Ciona genus. BMC Genomics *7*, 288.

Godin, J.D., Colombo, K., Molina-Calavita, M., Keryer, G., Zala, D., Charrin, B.C., Dietrich, P., Volvert, M.-L., Guillemot, F., Dragatsis, I., et al. (2010). Huntingtin is required for mitotic spindle orientation and mammalian neurogenesis. Neuron *67*, 392–406.

Goehler, H., Lalowski, M., Stelzl, U., Waelter, S., Stroedicke, M., Worm, U., Droege, A., Lindenberg, K.S., Knoblich, M., Haenig, C., et al. (2004). A Protein Interaction Network Links GIT1, an Enhancer of Huntingtin Aggregation, to Huntington's Disease. Molecular Cell *15*, 853–865.

Göke, J., Jung, M., Behrens, S., Chavez, L., O'Keeffe, S., Timmermann, B., Lehrach, H., Adjaye, J., and Vingron, M. (2011). Combinatorial binding in human and mouse embryonic stem cells identifies conserved enhancers active in early embryonic development. PLoS Comput. Biol. *7*, e1002304.

Goldberg, Y.P., Kremer, B., Andrew, S.E., Theilmann, J., Graham, R.K., Squitieri, F., Telenius, H., Adam, S., Sajoo, A., and Starr, E. (1993). Molecular analysis of new mutations for Huntington's disease: intermediate alleles and sex of origin effects. Nat. Genet. *5*, 174–179.

Goldberg, Y.P., McMurray, C.T., Zeisler, J., Almqvist, E., Sillence, D., Richards, F., Gacy, A.M., Buchanan, J., Telenius, H., and Hayden, M.R. (1995). Increased instability of intermediate alleles in families with sporadic Huntington disease compared to similar sized intermediate alleles in the general population. Hum. Mol. Genet. *4*, 1911–1918.

Goldberg, Y.P., Nicholson, D.W., Rasper, D.M., Kalchman, M.A., Koide, H.B., Graham, R.K., Bromm, M., Kazemi-Esfarjani, P., Thornberry, N.A., Vaillancourt, J.P., et al. (1996). Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract. Nat. Genet. *13*, 442–449.

Gonitel, R., Moffitt, H., Sathasivam, K., Woodman, B., Detloff, P.J., Faull, R.L.M., and Bates, G.P. (2008). DNA instability in postmitotic neurons. PNAS *105*, 3467–3472.

Gonzalez, R., Lee, J.W., Snyder, E.Y., and Schultz, P.G. (2011). Dorsomorphin promotes human embryonic stem cell self-renewal. Angew. Chem. Int. Ed. Engl. *50*, 3439–3441.

Gottlieb, D.I. (2002). Large-scale sources of neural stem cells. Annu. Rev. Neurosci. *25*, 381–407.

Graham, R.K., Deng, Y., Slow, E.J., Haigh, B., Bissada, N., Lu, G., Pearson, J., Shehadeh, J., Bertram, L., Murphy, Z., et al. (2006). Cleavage at the caspase-6 site is required for neuronal dysfunction and degeneration due to mutant huntingtin. Cell *125*, 1179–1191.

Gray, M., Shirasaki, D.I., Cepeda, C., André, V.M., Wilburn, B., Lu, X.-H., Tao, J., Yamazaki, I., Li, S.-H., Sun, Y.E., et al. (2008). Full-Length Human Mutant Huntingtin with a Stable Polyglutamine Repeat Can Elicit Progressive and Selective Neuropathogenesis in BACHD Mice. J. Neurosci. *28*, 6182–6195.

Gu, X., Greiner, E.R., Mishra, R., Kodali, R., Osmand, A., Finkbeiner, S., Steffan, J.S., Thompson, L.M., Wetzel, R., and Yang, X.W. (2009). Serines 13 and 16 are critical determinants of full-length human mutant huntingtin induced disease pathogenesis in HD mice. Neuron *64*, 828–840.

Guan, K., Chang, H., Rolletschek, A., and Wobus, A.M. (2001). Embryonic stem cell-derived neurogenesis. Retinoic acid induction and lineage selection of neuronal cells. Cell Tissue Res. *305*, 171–176.

Guérette, D., Khan, P.A., Savard, P.E., and Vincent, M. (2007). Molecular evolution of type VI intermediate filament proteins. BMC Evolutionary Biology *7*, 164.

Gunawardena, S., Her, L.-S., Brusch, R.G., Laymon, R.A., Niesman, I.R., Gordesky-Gold, B., Sintasath, L., Bonini, N.M., and Goldstein, L.S.B. (2003). Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila. Neuron *40*, 25–40.

Gutekunst, C.A., Levey, A.I., Heilman, C.J., Whaley, W.L., Yi, H., Nash, N.R., Rees, H.D., Madden, J.J., and Hersch, S.M. (1995). Identification and localization of huntingtin in brain and human lymphoblastoid cell lines with anti-fusion protein antibodies. Proc. Natl. Acad. Sci. U.S.A. *92*, 8710–8714.

Gutekunst, C.A., Li, S.H., Yi, H., Ferrante, R.J., Li, X.J., and Hersch, S.M. (1998). The cellular and subcellular localization of huntingtin-associated protein 1 (HAP1): comparison with huntingtin in rat and human. J. Neurosci. *18*, 7674–7686.

Hackam, A.S., Yassa, A.S., Singaraja, R., Metzler, M., Gutekunst, C.-A., Gan, L., Warby, S., Wellington, C.L., Vaillancourt, J., Chen, N., et al. (2000). Huntingtin Interacting Protein 1 Induces Apoptosis via a Novel Caspase-dependent Death Effector Domain. J. Biol. Chem. *275*, 41299–41308.

Hamada, H., Petrino, M.G., and Kakunaga, T. (1982). A Novel Repeated Element with Z-DNA-Forming Potential is Widely Found in Evolutionarily Diverse Eukaryotic Genomes. Proceedings of the National Academy of Sciences of the United States of America *79*, 6465–6469.

Hamada, H., Seidman, M., Howard, B.H., and Gorman, C.M. (1984). Enhanced gene expression by the poly(dT-dG).poly(dC-dA) sequence. Mol Cell Biol *4*, 2622–2630.

Hammock, E.A.D., and Young, L.J. (2005). Microsatellite Instability Generates Diversity in Brain and Sociobehavioral Traits. Science *308*, 1630–1634.

Han, G., Wang, H., and Hao, J. (2013). Molecular Mechanisms of Embryonic Stem Cell Pluripotency.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. U.S.A. *107*, 9222–9227.

Hao, J., Ho, J.N., Lewis, J.A., Karim, K.A., Daniels, R.N., Gentry, P.R., Hopkins, C.R., Lindsley, C.W., and Hong, C.C. (2010). In vivo structure-activity relationship study of dorsomorphin analogues identifies selective VEGF and BMP inhibitors. ACS Chem. Biol. *5*, 245–253.

Hao, J., Sawyer, D.B., Hatzopoulos, A.K., and Hong, C.C. (2011). Recent Progress on Chemical Biology of Pluripotent Stem Cell Self-renewal, Reprogramming and Cardiomyogenesis. Recent Pat Regen Med *1*, 263–274.

Harding, M.J., McGraw, H.F., and Nechiporuk, A. (2014). The roles and regulation of multicellular rosette structures during morphogenesis. Development *141*, 2549–2558.

Harjes, P., and Wanker, E.E. (2003). The hunt for huntingtin function: interaction partners tell many different stories. Trends Biochem. Sci. *28*, 425–433.

Hatch, M.N., Nistor, G., and Keirstead, H.S. (2009). Derivation of high-purity oligodendroglial progenitors. Methods Mol. Biol. *549*, 59–75.

HD Collaborative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. Cell *72*, 971–983.

Hemmann, U., Gerhartz, C., Heesel, B., Sasse, J., Kurapkat, G., Grötzinger, J., Wollmer, A., Zhong, Z., Darnell, J.E., Graeve, L., et al. (1996). Differential activation of acute phase response factor/Stat3 and Stat1 via the cytoplasmic domain of the interleukin 6 signal transducer gp130. II. Src homology SH2 domains define the specificity of stat factor activation. J. Biol. Chem. *271*, 12999–13007.

Henshall, T.L., Tucker, B., Lumsden, A.L., Nornes, S., Lardelli, M.T., and Richards, R.I. (2009). Selective neuronal requirement for huntingtin in the developing zebrafish. Hum. Mol. Genet. *18*, 4830–4842.

Hermel, E., Gafni, J., Propp, S.S., Leavitt, B.R., Wellington, C.L., Young, J.E., Hackam, A.S., Logvinova, A.V., Peel, A.L., Chen, S.F., et al. (2004). Specific caspase interactions and amplification are involved in selective neuronal vulnerability in Huntington's disease. Cell Death Differ. *11*, 424–438.

Hilditch-Maguire, P., Trettel, F., Passani, L.A., Auerbach, A., Persichetti, F., and MacDonald, M.E. (2000). Huntingtin: an iron-regulated protein essential for normal nuclear and perinuclear organelles. Hum. Mol. Genet. *9*, 2789–2797.

Hirai, H., Karian, P., and Kikyo, N. (2011). Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor. Biochem. J. *438*, 11–23.

Ho, L.W., Brown, R., Maxwell, M., Wyttenbach, A., and Rubinsztein, D.C. (2001). Wild type huntingtin reduces the cellular toxicity of mutant huntingtin in mammalian cell models of Huntington's disease. Journal of Medical Genetics *38*, 450–452.

Holland, L.Z., Laudet, V., and Schubert, M. (2004). The chordate amphioxus: an emerging model organism for developmental biology. Cell. Mol. Life Sci. *61*, 2290–2308.

Hong, E., and Brewster, R. (2006). N-cadherin is required for the polarized cell behaviors that drive neurulation in the zebrafish. Development *133*, 3895–3905.

Hoogeveen, A.T., Willemsen, R., Meyer, N., de Rooij, K.E., Roos, R.A., van Ommen, G.J., and Galjaard, H. (1993). Characterization and localization of the Huntington disease gene product. Hum. Mol. Genet. *2*, 2069–2073.

Horie, M., Ito, A., Kawabe, Y., and Kamihira, M. (2011). A Genetically Engineered STO Feeder System Expressing E-Cadherin and Leukemia Inhibitory Factor for Mouse Pluripotent Stem Cell Culture. Journal of Bioprocessing & Biotechniques.

Huang, B., Lucas, T., Kueppers, C., Dong, X., Krause, M., Bepperling, A., Buchner, J., Voshol, H., Weiss, A., Gerrits, B., et al. (2015). Scalable production in human cells and biochemical characterization of full-length normal and mutant huntingtin. PLoS ONE *10*, e0121055.

Hughes, A.L., and Friedman, R. (2005). Loss of ancestral genes in the genomic evolution of Ciona intestinalis. Evol. Dev. *7*, 196–200.

Hughes, A.C., Mort, M., Elliston, L., Thomas, R.M., Brooks, S.P., Dunnett, S.B., and Jones, L. (2014). Identification of novel alternative splicing events in the huntingtin gene and assessment of the functional consequences using structural protein homology modelling. J. Mol. Biol. *426*, 1428–1438.

Ihle, J.N., and Kerr, I.M. (1995). Jaks and Stats in signaling by the cytokine receptor superfamily. Trends Genet. *11*, 69–74.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567–582.

James, D., Levine, A.J., Besser, D., and Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. Development *132*, 1273–1282.

Kaltenbach, L.S., Romero, E., Becklin, R.R., Chettier, R., Bell, R., Phansalkar, A., Strand, A., Torcassi, C., Savage, J., Hurlburt, A., et al. (2007). Huntingtin Interacting Proteins Are Genetic Modifiers of Neurodegeneration. PLOS Genetics *3*, e82.

Kang, H.B., Kim, J.S., Kwon, H.-J., Nam, K.H., Youn, H.S., Sok, D.-E., and Lee, Y. (2005). Basic fibroblast growth factor activates ERK and induces c-fos in human embryonic stem cell line MizhES1. Stem Cells Dev. *14*, 395–401.

Karlin, S., Brocchieri, L., Bergman, A., Mrazek, J., and Gentles, A.J. (2002). Amino acid runs in eukaryotic proteomes

and disease associations. Proc. Natl. Acad. Sci. U.S.A. *99*, 333–338.

Kashi, Y., and King, D.G. (2006). Simple sequence repeats as advantageous mutators in evolution. Trends Genet. *22*, 253–259.

Kashi, Y., King, D., and Soller, M. (1997). Simple sequence repeats as a source of quantitative genetic variation. Trends Genet. *13*, 74–78.

Kauffman, J.S., Zinovyeva, A., Yagi, K., Makabe, K.W., and Raff, R.A. (2003). Neural expression of the Huntington's disease gene as a chordate evolutionary novelty. J. Exp. Zool. B Mol. Dev. Evol. *297*, 57–64.

Kawata, M., Sekiya, S., Kera, K., Kimura, H., and Takamizawa, H. (1991). Neural rosette formation within in vitro spheroids of a clonal human teratocarcinoma cell line, PA-1/NR: role of extracellular matrix components in the morphogenesis. Cancer Res. *51*, 2655–2669.

Kegel, K.B., Meloni, A.R., Yi, Y., Kim, Y.J., Doyle, E., Cuiffo, B.G., Sapp, E., Wang, Y., Qin, Z.-H., Chen, J.D., et al. (2002). Huntingtin Is Present in the Nucleus, Interacts with the Transcriptional Corepressor C-terminal Binding Protein, and Represses Transcription. J. Biol. Chem. *277*, 7466–7476.

Kelly, T.E., Allinson, P., McGlennen, R.C., Baker, J., and Bao, Y. (1999). Expansion of a 27 CAG repeat allele into a symptomatic Huntington disease-producing allele. Am. J. Med. Genet. *87*, 91–92.

Keryer, G., Pineda, J.R., Liot, G., Kim, J., Dietrich, P., Benstaali, C., Smith, K., Cordelières, F.P., Spassky, N., Ferrante, R.J., et al. (2011). Ciliogenesis is regulated by a huntingtin-HAP1-PCM1 pathway and is altered in Huntington disease. J Clin Invest *121*, 4372–4382.

Kielman, M.F., Rindapää, M., Gaspar, C., van Poppel, N., Breukel, C., van Leeuwen, S., Taketo, M.M., Roberts, S., Smits, R., and Fodde, R. (2002). Apc modulates embryonic stem-cell differentiation by controlling the dosage of beta-catenin signaling. Nat. Genet. *32*, 594–605.

Kijas, J.M., Moller, M., Plastow, G., and Andersson, L. (2001). A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs. Genetics *158*, 779–785.

Kim, M., Roh, J.-K., Yoon, B.W., Kang, L., Kim, Y.J., Aronin, N., and DiFiglia, M. (2003). Huntingtin is degraded to small fragments by calpain after ischemic injury. Exp. Neurol. *183*, 109–115.

Kim, M.W., Chelliah, Y., Kim, S.W., Otwinowski, Z., and Bezprozvanny, I. (2009). Secondary structure of Huntingtin amino-terminal region. Structure *17*, 1205–1212.

Kim, Y.J., Yi, Y., Sapp, E., Wang, Y., Cuiffo, B., Kegel, K.B., Qin, Z.-H., Aronin, N., and DiFiglia, M. (2001). Caspase 3-cleaved N-terminal fragments of wild-type and mutant huntingtin are present in normal and Huntington's disease brains, associate with membranes, and undergo calpain-dependent proteolysis. Proc Natl Acad Sci U S A *98*, 12784–12789.

Kim, Y.J., Sapp, E., Cuiffo, B.G., Sobin, L., Yoder, J., Kegel, K.B., Qin, Z.-H., Detloff, P., Aronin, N., and DiFiglia, M. (2006). Lysosomal proteases are involved in generation of N-terminal huntingtin fragments. Neurobiol. Dis. *22*, 346–356.

King, D.G. (2012). Evolution of simple sequence repeats as mutable sites. Adv. Exp. Med. Biol. *769*, 10–25.

King, D.G., and Kashi, Y. (2007). Mutability and evolvability: Indirect selection for mutability. Heredity *99*, 123–124.

King, D.G., and Soller, M. (1999). Variation and Fidelity: The Evolution of Simple Sequence Repeats as Functional Elements in Adjustable Genes. In Evolutionary Theory and Processes: Modern Perspectives, (Springer, Dordrecht), pp. 65–82.

King, D.G., Soller, M., and Kashi, Y. (1997). Evolutionary tuning knobs. Endeavour *21*, 36–40.

Kleinsmith, L.J., and Pierce, G.B. (1964). MULTIPOTENTIALITY OF SINGLE EMBRYONAL CARCINOMA CELLS. Cancer Res. *24*, 1544–1551.

Kovtun, I.V., and McMurray, C.T. (2008). Features of trinucleotide repeat instability in vivo. Cell Res. *18*, 198–213.

Kremer, B., Goldberg, P., Andrew, S.E., Theilmann, J., Telenius, H., Zeisler, J., Squitieri, F., Lin, B., Bassett, A., Almqvist, E., et al. (1994). A Worldwide Study of the Huntington's Disease Mutation: The Sensitivity and Specificity of Measuring CAG Repeats. New England Journal of Medicine *330*, 1401–1406.

Kremer, B., Almqvist, E., Theilmann, J., Spence, N., Telenius, H., Goldberg, Y.P., and Hayden, M.R. (1995). Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. Am J Hum Genet *57*, 343–350.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

Kyriacou, C.P., Peixoto, A.A., Sandrelli, F., Costa, R., and Tauber, E. (2008). Clines in clock genes: fine-tuning circadian rhythms to the environment. Trends Genet. *24*, 124–132.

Laccone, F., and Christian, W. (2000). A Recurrent Expansion of a Maternal Allele with 36 CAG Repeats Causes Huntington Disease in Two Sisters. Am J Hum Genet *66*, 1145–1148.

Landwehrmeyer, G.B., McNeil, S.M., Dure, L.S., Ge, P., Aizawa, H., Huang, Q., Ambrose, C.M., Duyao, M.P., Bird, E.D., and Bonilla, E. (1995). Huntington's disease gene: regional and cellular expression in brain of normal and affected individuals. Ann. Neurol. *37*, 218–230.

Leavitt, B.R., van Raamsdonk, J.M., Shehadeh, J., Fernandes, H., Murphy, Z., Graham, R.K., Wellington, C.L., and Hayden, M.R. (2006). Wild-type huntingtin protects neurons from excitotoxicity. Journal of Neurochemistry *96*, 1121–1129.

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., et al. (2012). CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. Neurology *78*, 690–695.

Lee, R.E., Young, R.H., and Castleman, B. (2002). James Homer Wright: a biography of the enigmatic creator of the Wright stain on the occasion of its centennial. Am. J. Surg. Pathol. *26*, 88–96.

Lele, Z., Folchert, A., Concha, M., Rauch, G.-J., Geisler, R., Rosa, F., Wilson, S.W., Hammerschmidt, M., and Bally-Cuif, L. (2002). parachute/n-cadherin is required for morphogenesis and maintained integrity of the zebrafish neural tube. Development *129*, 3281–3294.

Lendahl, U., Zimmerman, L.B., and McKay, R.D. (1990). CNS stem cells express a new class of intermediate filament protein. Cell *60*, 585–595.

Lesch, K.-P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., Petri, S., Benjamin, J., Müller, C.R., Hamer, D.H., and Murphy, D.L. (1996). Association of Anxiety-Related Traits with a Polymorphism in the Serotonin Transporter Gene Regulatory Region. Science *274*, 1527–1531.

Li, J., Wang, G., Wang, C., Zhao, Y., Zhang, H., Tan, Z., Song, Z., Ding, M., and Deng, H. (2007). MEK/ERK signaling contributes to the maintenance of human embryonic stem cell self-renewal. Differentiation *75*, 299–307.

Li, W., Serpell, L.C., Carter, W.J., Rubinsztein, D.C., and Huntington, J.A. (2006). Expression and Characterization of Full-length Human Huntingtin, an Elongated HEAT Repeat Protein. J. Biol. Chem. *281*, 15916–15922.

Li, X., Chen, Y., Schéele, S., Arman, E., Haffner-Krausz, R., Ekblom, P., and Lonai, P. (2001). Fibroblast growth factor signaling and basement membrane assembly are connected during epithelial morphogenesis of the embryoid body. J. Cell Biol. *153*, 811–822.

Li, Z., Karlovich, C.A., Fish, M.P., Scott, M.P., and Myers, R.M. (1999). A putative Drosophila homolog of the Huntington's disease gene. Hum. Mol. Genet. *8*, 1807–1815.

Li, Z., Jo, J., Jia, J.-M., Lo, S.-C., Whitcomb, D.J., Jiao, S., Cho, K., and Sheng, M. (2010). Caspase-3 activation via mitochondria is required for long-term depression and AMPA receptor internalization. Cell *141*, 859–871.

Lin, B., Rommens, J.M., Graham, R.K., Kalchman, M., MacDonald, H., Nasir, J., Delaney, A., Goldberg, Y.P., and Hayden, M.R. (1993). Differential 3' polyadenylation of the Huntington disease gene results in two mRNA species with variable tissue expression. Hum. Mol. Genet. *2*, 1541–1545.

Lin, B., Nasir, J., MacDonald, H., Hutchinson, G., Graham, R.K., Rommens, J.M., and Hayden, M.R. (1994). Sequence of the murine Huntington disease gene: evidence for conservation, alternate splicing and polymorphism in a triplet (CCG) repeat [corrected]. Hum. Mol. Genet. *3*, 85–92.

Liot, G., Zala, D., Pla, P., Mottet, G., Piel, M., and Saudou, F. (2013). Mutant Huntingtin Alters Retrograde Transport of TrkB Receptors in Striatal Dendrites. J. Neurosci. *33*, 6298–6309.

Liu, J.-P., and Zeitlin, S.O. (2017). Is Huntingtin Dispensable in the Adult Brain? J Huntingtons Dis *6*, 1–17.

Lo Sardo, V., Zuccato, C., Gaudenzi, G., Vitali, B., Ramos, C., Tartari, M., Myre, M.A., Walker, J.A., Pistocchi, A., Conti, L., et al. (2012). An evolutionary recent neuroepithelial cell adhesion function of huntingtin implicates ADAM10-Ncadherin. Nat Neurosci *15*, 713–721.

Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. *38*, 431–440.

Lumsden, A.L., Henshall, T.L., Dayan, S., Lardelli, M.T., and Richards, R.I. (2007). Huntingtin-deficient zebrafish exhibit defects in iron utilization and development. Hum Mol Genet *16*, 1905–1920.

Lunkes, A., Lindenberg, K.S., Ben-Haïem, L., Weber, C., Devys, D., Landwehrmeyer, G.B., Mandel, J.-L., and Trottier, Y. (2002). Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions. Mol. Cell *10*, 259–269.

Maiuri, T., Woloshansky, T., Xia, J., and Truant, R. (2013). The huntingtin N17 domain is a multifunctional CRM1 and Ran-dependent nuclear and cilial export signal. Hum Mol Genet *22*, 1383–1394.

Majo, F., Rochat, A., Nicolas, M., Jaoudé, G.A., and Barrandon, Y. (2008). Oligopotent stem cells are distributed throughout the mammalian ocular surface. Nature *456*, 250–254.

Marcello, E., Gardoni, F., Mauceri, D., Romorini, S., Jeromin, A., Epis, R., Borroni, B., Cattabeni, F., Sala, C., Padovani, A., et al. (2007). Synapse-associated protein-97 mediates alpha-secretase ADAM10 trafficking and promotes its activity. J. Neurosci. *27*, 1682–1691.

Marcora, E., Gowan, K., and Lee, J.E. (2003). Stimulation of NeuroD activity by huntingtin and huntingtin-associated proteins HAP1 and MLK2. Proc Natl Acad Sci U S A *100*, 9578–9583.

Margolis, B., and Borg, J.-P. (2005). Apicobasal polarity complexes. Journal of Cell Science *118*, 5157–5159.

Marthiens, V., and ffrench-Constant, C. (2009). Adherens junction domains are split by asymmetric division of embryonic neural stem cells. EMBO Rep *10*, 515–520.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U.S.A. *78*, 7634–7638.

Martin, G.R., and Evans, M.J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. Proc Natl Acad Sci U S A *72*, 1441–1445.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat. Cell Biol. *9*, 625–635.

Mathur, D., Danford, T.W., Boyer, L.A., Young, R.A., Gifford, D.K., and Jaenisch, R. (2008). Analysis of the mouse embryonic stem cell regulatory networks obtained by ChIP-chip and ChIP-PET. Genome Biol. *9*, R126.

Matsuura, T., Fang, P., Pearson, C.E., Jayakar, P., Ashizawa, T., Roa, B.B., and Nelson, D.L. (2006). Interruptions in the Expanded ATTCT Repeat of Spinocerebellar Ataxia Type 10: Repeat Purity as a Disease Modifier? Am J Hum Genet *78*, 125–129.

Maye, P., Becker, S., Siemen, H., Thorne, J., Byrd, N., Carpentino, J., and Grabel, L. (2004). Hedgehog signaling is required for the differentiation of ES cells into neurectoderm. Dev. Biol. *265*, 276–290.

McFarland, K.N., Liu, J., Landrian, I., Zeng, D., Raskin, S., Moscovich, M., Gatto, E.M., Ochoa, A., Teive, H.A.G., Rasmussen, A., et al. (2014). Repeat interruptions in spinocerebellar ataxia type 10 expansions are strongly associated with epileptic seizures. Neurogenetics *15*, 59–64.

McNeil S.M., Novelletto A., Barnes G., et al. (1997). Reduced penetrance of the Huntington's disease mutation. Hum. Mol. Genet. 5, 6, 775-779

Michalczyk, K., and Ziman, M. (2005). Nestin structure and predicted function in cellular cytoskeletal organisation. Histol. Histopathol. *20*, 665–671.

Miller, J.P., Holcomb, J., Al-Ramahi, I., de Haro, M., Gafni, J., Zhang, N., Kim, E., Sanhueza, M., Torcassi, C., Kwak, S., et al. (2010). Matrix metalloproteinases are modifiers of huntingtin proteolysis and toxicity in Huntington's disease. Neuron *67*, 199–212.

Miyata, T. (2007). Asymmetric cell division during brain morphogenesis. Prog. Mol. Subcell. Biol. *45*, 121–142.

Molero, A.E., Arteaga-Bracho, E.E., Chen, C.H., Gulinello, M., Winchester, M.L., Pichamoorthy, N., Gokhan, S., Khodakhah, K., and 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. *113*, 5736–5741.

Mühlau, M., Winkelmann, J., Rujescu, D., Giegling, I., Koutsouleris, N., Gaser, C., Arsic, M., Weindl, A., Reiser, M., and Meisenzahl, E.M. (2012). Variation within the Huntington's Disease Gene Influences Normal Brain Structure. PLoS One *7*.

Mulvihill, D.J., Nichol Edamura, K., Hagerman, K.A., Pearson, C.E., and Wang, Y.-H. (2005). Effect of CAT or AGG interruptions and CpG methylation on nucleosome assembly upon trinucleotide repeats on spinocerebellar ataxia, type 1 and fragile X syndrome. J. Biol. Chem. *280*, 4498–4503.

Munafò, M.R., Clark, T., and Flint, J. (2005). Does measurement instrument moderate the association between the serotonin transporter gene and anxiety-related personality traits? A meta-analysis. Mol. Psychiatry *10*, 415–419.

Muñoz-Sanjuán, I., and Brivanlou, A.H. (2002). Neural induction, the default model and embryonic stem cells. Nat. Rev. Neurosci. *3*, 271–280.

Myers, R.H., MacDonald, M.E., Koroshetz, W.J., Duyao, M.P., Ambrose, C.M., Taylor, S. a. M., Barnes, G., Srinidhi, J., Lin, C.S., Whaley, W.L., et al. (1993). De novo expansion of a (CAG)n repeat in sporadic Huntington's disease. Nat Genet *5*, 168–173.

Myre, M.A., Lumsden, A.L., Thompson, M.N., Wasco, W., MacDonald, M.E., and Gusella, J.F. (2011). Deficiency of Huntingtin Has Pleiotropic Effects in the Social Amoeba Dictyostelium discoideum. PLOS Genetics *7*, e1002052.

Nasir, J., Floresco, S.B., O'Kusky, J.R., Diewert, V.M., Richman, J.M., Zeisler, J., Borowski, A., Marth, J.D., Phillips, A.G., and Hayden, M.R. (1995). Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. Cell *81*, 811–823.

Nestor, C.E., and Monckton, D.G. (2011). Correlation of inter-locus polyglutamine toxicity with CAG•CTG triplet repeat expandability and flanking genomic DNA GC content. PLoS ONE *6*, e28260.

Neuwald, A.F., and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. Genome Res. *10*, 1445–1452.

Neveklovska, M., Clabough, E.B.D., Steffan, J.S., and Zeitlin, S.O. (2012). Deletion of the huntingtin proline-rich region does not significantly affect normal huntingtin function in mice. J Huntingtons Dis *1*, 71–87.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

Nichols, J., and Smith, A. (2011). The origin and identity of embryonic stem cells. Development *138*, 3–8.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379–391.

Nithianantharajah, J., and Hannan, A.J. (2007). Dynamic mutations as digital genetic modulators of brain development, function and dysfunction. Bioessays *29*, 525–535.

Niwa, H. (2007). How is pluripotency determined and maintained? Development *134*, 635–646.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. *12*, 2048–2060.

Noctor, S.C., Flint, A.C., Weissman, T.A., Dammerman, R.S., and Kriegstein, A.R. (2001). Neurons derived from radial glial cells establish radial units in neocortex. Nature *409*, 714–720.

Nopoulos, P., Epping, E.A., Wassink, T., Schlaggar, B.L., and Perlmutter, J. (2011a). Correlation of CAG repeat length between the maternal and paternal allele of the Huntingtin gene: evidence for assortative mating. Behav Brain Funct *7*, 45.

Nopoulos, P.C., Aylward, E.H., Ross, C.A., Mills, J.A., Langbehn, D.R., Johnson, H.J., Magnotta, V.A., Pierson, R.K., Beglinger, L.J., Nance, M.A., et al. (2011b). Smaller intracranial volume in prodromal Huntington's disease: evidence for abnormal neurodevelopment. Brain *134*, 137–142.

Ochaba, J., Lukacsovich, T., Csikos, G., Zheng, S., Margulis, J., Salazar, L., Mao, K., Lau, A.L., Yeung, S.Y., Humbert, S., et al. (2014). Potential function for the Huntingtin protein as a scaffold for selective autophagy. Proc. Natl. Acad. Sci. U.S.A. *111*, 16889–16894.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem. Biophys. Res. Commun. *343*, 159–166.

Ogawa, K., Saito, A., Matsui, H., Suzuki, H., Ohtsuka, S., Shimosato, D., Morishita, Y., Watabe, T., Niwa, H., and Miyazono, K. (2007). Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. J. Cell. Sci. *120*, 55–65.

Okabe, S., Forsberg-Nilsson, K., Spiro, A.C., Segal, M., and McKay, R.D. (1996). Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. Mech. Dev. *59*, 89–102.

O'Kusky, J.R., Nasir, J., Cicchetti, F., Parent, A., and Hayden, M.R. (1999). Neuronal degeneration in the basal ganglia and loss of pallido-subthalamic synapses in mice with targeted disruption of the Huntington's disease gene. Brain Research *818*, 468–479.

Orr, H.T., and Zoghbi, H.Y. (2007). Trinucleotide repeat disorders. Annu. Rev. Neurosci. *30*, 575–621.

Osorno, R., and Chambers, I. (2011). Transcription factor heterogeneity and epiblast pluripotency. Philos Trans R Soc Lond B Biol Sci *366*, 2230–2237.

Palidwor, G.A., Shcherbinin, S., Huska, M.R., Rasko, T., Stelzl, U., Arumughan, A., Foulle, R., Porras, P., Sanchez-Pulido, L., Wanker, E.E., et al. (2009). Detection of Alpha-Rod Protein Repeats Using a Neural Network and Application to Huntingtin. PLoS Comput Biol *5*.

Panov, A.V., Gutekunst, C.-A., Leavitt, B.R., Hayden, M.R., Burke, J.R., Strittmatter, W.J., and Greenamyre, J.T. (2002). Early mitochondrial calcium defects in Huntington's disease are a direct effect of polyglutamines. Nat. Neurosci. *5*, 731–736.

Pearson, C.E., Nichol Edamura, K., and Cleary, J.D. (2005). Repeat instability: mechanisms of dynamic mutations. Nat. Rev. Genet. *6*, 729–742.

Peixoto, A.A., Hennessy, J.M., Townson, I., Hasan, G., Rosbash, M., Costa, R., and Kyriacou, C.P. (1998). Molecular coevolution within a Drosophila clock gene. Proc Natl Acad Sci U S A *95*, 4475–4480.

Pereira, L., Yi, F., and Merrill, B.J. (2006). Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. Mol. Cell. Biol. *26*, 7479–7491.

Perutz, M.F., Johnson, T., Suzuki, M., and Finch, J.T. (1994). Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. Proc. Natl. Acad. Sci. U.S.A. *91*, 5355–5358.

Pierce, G.B., and Verney, E.L. (1961). An in vitro and in vivo study of differentiation in teratocarcinomas. Cancer *14*, 1017–1029.

Pla, P., Orvoen, S., Benstaali, C., Dodier, S., Gardier, A.M., David, D.J., Humbert, S., and Saudou, F. (2013). Huntingtin Acts Non Cell-Autonomously on Hippocampal Neurogenesis and Controls Anxiety-Related Behaviors in Adult Mouse. PLOS ONE *8*, e73902.

Pouladi, M.A., Xie, Y., Skotte, N.H., Ehrnhoefer, D.E., Graham, R.K., Kim, J.E., Bissada, N., Yang, X.W., Paganetti, P., Friedlander, R.M., et al. (2010). Full-length huntingtin levels modulate body weight by influencing insulin-like growth factor 1 expression. Human Molecular Genetics *19*, 1528–1538.

Qi, X., Li, T.-G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., and Zhao, G.-Q. (2004). BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. Proc. Natl. Acad. Sci. U.S.A. *101*, 6027–6032.

Quarrell, O.W.J., Rigby, A.S., Barron, L., Crow, Y., Dalton, A., Dennis, N., Fryer, A.E., Heydon, F., Kinning, E., Lashwood, A., et al. (2007). Reduced penetrance alleles for Huntington's disease: a multi‐centre direct observational study. J Med Genet *44*, e68.

Radice, G.L., Rayburn, H., Matsunami, H., Knudsen, K.A., Takeichi, M., and Hynes, R.O. (1997). Developmental defects in mouse embryos lacking N-cadherin. Dev. Biol. *181*, 64–78.

Rangone, H., Poizat, G., Troncoso, J., Ross, C.A., MacDonald, M.E., Saudou, F., and Humbert, S. (2004). The serum- and glucocorticoid-induced kinase SGK inhibits mutant huntingtin-induced toxicity by phosphorylating serine 421 of huntingtin. Eur. J. Neurosci. *19*, 273–279.

Rathjen, J., Haines, B.P., Hudson, K.M., Nesci, A., Dunn, S., and Rathjen, P.D. (2002). Directed differentiation of pluripotent cells to neural lineages: homogeneous formation and differentiation of a neurectoderm population. Development *129*, 2649–2661.

Ratovitski, T., Chighladze, E., Arbez, N., Boronina, T., Herbrich, S., Cole, R.N., and Ross, C.A. (2012). Huntingtin protein interactions altered by polyglutamine expansion as determined by quantitative proteomic analysis. Cell Cycle *11*, 2006–2021.

Rebuzzini, P., Neri, T., Mazzini, G., Zuccotti, M., Redi, C.A., and Garagna, S. (2008). Karyotype analysis of the euploid cell population of a mouse embryonic stem cell line revealed a high incidence of chromosome abnormalities that varied during culture. Cytogenet. Genome Res. *121*, 18–24.

Reich, N.C., and Liu, L. (2006). Tracking STAT nuclear traffic. Nat. Rev. Immunol. *6*, 602–612.

Reiner, A., Albin, R.L., Anderson, K.D., D'Amato, C.J., Penney, J.B., and Young, A.B. (1988). Differential loss of striatal projection neurons in Huntington disease. Proc. Natl. Acad. Sci. U.S.A. *85*, 5733–5737.

Reiner, A., Del Mar, N., Meade, C.A., Yang, H., Dragatsis, I., Zeitlin, S., and Goldowitz, D. (2001). Neurons lacking huntingtin differentially colonize brain and survive in chimeric mice. J. Neurosci. *21*, 7608–7619.

Reiner, A., Dragatsis, I., Zeitlin, S., and Goldowitz, D. (2003). Wild-type huntingtin plays a role in brain development and neuronal survival. Mol. Neurobiol. *28*, 259–276.

Reiss, K., Maretzky, T., Ludwig, A., Tousseyn, T., de Strooper, B., Hartmann, D., and Saftig, P. (2005). ADAM10 cleavage of N-cadherin and regulation of cell-cell adhesion and beta-catenin nuclear signalling. EMBO J. *24*, 742–752.

Richards, R.I., and Sutherland, G.R. (1994). Simple repeat DNA is not replicated simply. Nat. Genet. *6*, 114–116.

Rigamonti, D., Bauer, J.H., De-Fraja, C., Conti, L., Sipione, S., Sciorati, C., Clementi, E., Hackam, A., Hayden, M.R., Li, Y., et al. (2000). Wild-type huntingtin protects from apoptosis upstream of caspase-3. J. Neurosci. *20*, 3705–3713.

Rigamonti, D., Sipione, S., Goffredo, D., Zuccato, C., Fossale, E., and Cattaneo, E. (2001). Huntingtin's Neuroprotective Activity Occurs via Inhibition of Procaspase-9 Processing. J. Biol. Chem. *276*, 14545–14548.

Rockabrand, E., Slepko, N., Pantalone, A., Nukala, V.N., Kazantsev, A., Marsh, J.L., Sullivan, P.G., Steffan, J.S., Sensi, S.L., and Thompson, L.M. (2007). The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis. Hum. Mol. Genet. *16*, 61–77.

Rosas, H.D., Koroshetz, W.J., Chen, Y.I., Skeuse, C., Vangel, M., Cudkowicz, M.E., Caplan, K., Marek, K., Seidman, L.J., Makris, N., et al. (2003). Evidence for more widespread cerebral pathology in early HD: an MRI-based morphometric analysis. Neurology *60*, 1615–1620.

Rosas, H.D., Hevelone, N.D., Zaleta, A.K., Greve, D.N., Salat, D.H., and Fischl, B. (2005). Regional cortical thinning in preclinical Huntington disease and its relationship to cognition. Neurology *65*, 745–747.

Rosas, H.D., Salat, D.H., Lee, S.Y., Zaleta, A.K., Pappu, V., Fischl, B., Greve, D., Hevelone, N., and Hersch, S.M. (2008). Cerebral cortex and the clinical expression of Huntington's disease: complexity and heterogeneity. Brain *131*, 1057–1068.

Rubinsztein, D.C., Amos, W., Leggo, J., Goodburn, S., Ramesar, R.S., Old, J., Bontrop, R., McMahon, R., Barton, D.E., and Ferguson-Smith, M.A. (1994). Mutational bias provides a model for the evolution of Huntington's disease and predicts a general increase in disease prevalence. Nat Genet *7*, 525–530.

Ruzo, A., Ismailoglu, I., Popowski, M., Haremaki, T., Croft, G.F., Deglincerti, A., and Brivanlou, A.H. (2015). Discovery of Novel Isoforms of Huntingtin Reveals a New Hominid-Specific Exon. PLOS ONE *10*, e0127687.

Sapp, E., Schwarz, C., Chase, K., Bhide, P.G., Young, A.B., Penney, J., Vonsattel, J.P., Aronin, N., and DiFiglia, M. (1997). Huntingtin localization in brains of normal and Huntington's disease patients. Ann. Neurol. *42*, 604–612.

Sathasivam, K., Neueder, A., Gipson, T.A., Landles, C., Benjamin, A.C., Bondulich, M.K., Smith, D.L., Faull, R.L.M., Roos, R.A.C., Howland, D., et al. (2013). Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. Proc. Natl. Acad. Sci. U.S.A. *110*, 2366–2370.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat. Med. *10*, 55–63.

Saudou, F., and Humbert, S. (2016). The Biology of Huntingtin. Neuron *89*, 910–926.

Saudou, F., Finkbeiner, S., Devys, D., and Greenberg, M.E. (1998). Huntingtin Acts in the Nucleus to Induce Apoptosis but Death Does Not Correlate with the Formation of Intranuclear Inclusions. Cell *95*, 55–66.

Sauer, M.E., and Walker, B.E. (1959). Radioautographic study of interkinetic nuclear migration in the neural tube. Proc. Soc. Exp. Biol. Med. *101*, 557–560.

Sawyer, L.A., Hennessy, J.M., Peixoto, A.A., Rosato, E., Parkinson, H., Costa, R., and Kyriacou, C.P. (1997). Natural variation in a Drosophila clock gene and temperature compensation. Science *278*, 2117–2120.

Schaefer, M.H., Wanker, E.E., and Andrade-Navarro, M.A. (2012). Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks. Nucleic Acids Res. *40*, 4273–4287.

Schilling, G., Sharp, A.H., Loev, S.J., Wagster, M.V., Li, S.H., Stine, O.C., and Ross, C.A. (1995). Expression of the Huntington's disease (IT15) protein product in HD patients. Hum. Mol. Genet. *4*, 1365–1371.

Sebald, W., Nickel, J., Zhang, J.-L., and Mueller, T.D. (2004). Molecular recognition in bone morphogenetic protein (BMP)/receptor interaction. Biol. Chem. *385*, 697–710.

Semaka, A., Creighton, S., Warby, S., and Hayden, M.R. (2006). Predictive testing for Huntington disease: interpretation and significance of intermediate alleles. Clin. Genet. *70*, 283–294.

Semaka, A., Kay, C., Doty, C.N., Collins, J.A., Tam, N., and Hayden, M.R. (2013). High frequency of intermediate alleles on Huntington disease-associated haplotypes in British Columbia's general population. Am. J. Med. Genet. B Neuropsychiatr. Genet. *162B*, 864–871.

Sen, S., Burmeister, M., and Ghosh, D. (2004). Meta-analysis of the association between a serotonin transporter promoter polymorphism (5-HTTLPR) and anxiety-related personality traits. Am. J. Med. Genet. B Neuropsychiatr. Genet. *127B*, 85–89.

Seong, I.S., Woda, J.M., Song, J.-J., Lloret, A., Abeyrathne, P.D., Woo, C.J., Gregory, G., Lee, J.-M., Wheeler, V.C., Walz, T., et al. (2010). Huntingtin facilitates polycomb repressive complex 2. Hum. Mol. Genet. *19*, 573–583.

Sharov, A.A., Masui, S., Sharova, L.V., Piao, Y., Aiba, K., Matoba, R., Xin, L., Niwa, H., and Ko, M.S.H. (2008). Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. BMC Genomics *9*, 269.

Sharp, A.H., Loev, S.J., Schilling, G., Li, S.H., Li, X.J., Bao, J., Wagster, M.V., Kotzuk, J.A., Steiner, J.P., and Lo, A. (1995). Widespread expression of Huntington's disease gene (IT15) protein product. Neuron *14*, 1065–1074.

Shimojo, M. (2008). Huntingtin regulates RE1-silencing transcription factor/neuron-restrictive silencer factor (REST/NRSF) nuclear trafficking indirectly through a complex with REST/NRSF-interacting LIM domain protein (RILP) and dynactin p150 Glued. J. Biol. Chem. *283*, 34880–34886.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. *6*, e253.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722–737.

Singh, V., Sinha, R.J., Sankhwar, S.N., B, Mehrotra, A., Ahmed, N., and Mehrotra, S. (2010). Squamous Cell Carcinoma of the Kidney Ã¢Â€Â“ Rarity Redefined: Case Series with Review of Literature. Journal of Cancer Science & Therapy.

Smith, A.G. (2001). Embryo-derived stem cells: of mice and men. Annu. Rev. Cell Dev. Biol. *17*, 435–462.

Smith, A.G., and Ying, Q.L. (2015). Culture medium containing kinase inhibitors, and uses thereof (Google Patents).

Sobczak, K., and Krzyzosiak, W.J. (2004). Patterns of CAG repeat interruptions in SCA1 and SCA2 genes in relation to

repeat instability. Hum. Mutat. *24*, 236–247.

Stahl, N., Farruggella, T.J., Boulton, T.G., Zhong, Z., Darnell, J.E., and Yancopoulos, G.D. (1995). Choice of STATs and other substrates specified by modular tyrosine-based motifs in cytokine receptors. Science *267*, 1349–1353.

Stavridis, M.P., Lunn, J.S., Collins, B.J., and Storey, K.G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development *134*, 2889–2894.

Steffan, J.S., Agrawal, N., Pallos, J., Rockabrand, E., Trotman, L.C., Slepko, N., Illes, K., Lukacsovich, T., Zhu, Y.-Z., Cattaneo, E., et al. (2004). SUMO modification of Huntingtin and Huntington's disease pathology. Science *304*, 100–104.

Stern, C.D. (2005). Neural induction: old problem, new findings, yet more questions. Development *132*, 2007–2021.

Strehlow, A.N.T., Li, J.Z., and Myers, R.M. (2007). Wild-type huntingtin participates in protein trafficking between the Golgi and the extracellular space. Hum. Mol. Genet. *16*, 391–409.

Strong, T.V., Tagle, D.A., Valdes, J.M., Elmer, L.W., Boehm, K., Swaroop, M., Kaatz, K.W., Collins, F.S., and Albin, R.L. (1993). Widespread expression of the human and rat Huntington's disease gene in brain and nonneural tissues. Nat. Genet. *5*, 259–265.

Sun, Y., Savanenin, A., Reddy, P.H., and Liu, Y.F. (2001). Polyglutamine-expanded huntingtin promotes sensitization of N-methyl-D-aspartate receptors via post-synaptic density 95. J. Biol. Chem. *276*, 24713–24718.

Suopanki, J., Götz, C., Lutsch, G., Schiller, J., Harjes, P., Herrmann, A., and Wanker, E.E. (2006). Interaction of huntingtin fragments with brain membranes--clues to early dysfunction in Huntington's disease. J. Neurochem. *96*, 870–884.

Takano, H., and Gusella, J.F. (2002). The predominantly HEAT-like motif structure of huntingtin and its association and coincident nuclear entry with dorsal, an NF-kB/Rel/dorsal family transcription factor. BMC Neurosci *3*, 15.

Takao, Y., Yokota, T., and Koide, H. (2007). Beta-catenin up-regulates Nanog expression through interaction with Oct-3/4 in embryonic stem cells. Biochem. Biophys. Res. Commun. *353*, 699–705.

Tan, J.-A.T., Song, J., Chen, Y., and Durrin, L.K. (2010). Phosphorylation-Dependent Interaction of SATB1 and PIAS1 Directs SUMO-Regulated Caspase Cleavage of SATB1. Mol. Cell. Biol. *30*, 2823–2836.

Tanaka, F., Sobue, G., Doyu, M., Ito, Y., Yamamoto, M., Shimada, N., Yamamoto, K., Riku, S., Hshizume, Y., and Mitsuma, T. (1996). Differential pattern in tissue-specific somatic mosaicism of expanded CAG trinucleotide repeat in dentatorubral-pallidoluysian atrophy, Machado-Joseph disease, and X-linked recessive spinal and bulbar muscular atrophy. Journal of the Neurological Sciences *135*, 43–50.

Tartari, M., Gissi, C., Lo Sardo, V., Zuccato, C., Picardi, E., Pesole, G., and Cattaneo, E. (2008). Phylogenetic comparison of huntingtin homologues reveals the appearance of a primitive polyQ in sea urchin. Mol. Biol. Evol. *25*, 330–338.

Telenius, H., Kremer, B., Goldberg, Y.P., Theilmann, J., Andrew, S.E., Zeisler, J., Adam, S., Greenberg, C., Ives, E.J., and Clarke, L.A. (1994). Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat. Genet. *6*, 409–414.

Telenius, H., Almqvist, E., Kremer, B., Spence, N., Squitieri, F., Nichol, K., Grandell, U., Starr, E., Benjamin, C., and Castaldo, I. (1995). Somatic mosaicism in sperm is associated with intergenerational (CAG)n changes in Huntington disease. Hum. Mol. Genet. *4*, 189–195.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Thion, M.S., McGuire, J.R., Sousa, C.M., Fuhrmann, L., Fitamant, J., Leboucher, S., Vacher, S., Montcel, D., Tezenas, S., Bièche, I., et al. (2015). Unraveling the Role of Huntingtin in Breast Cancer Metastasis. J Natl Cancer Inst *107*.

Thompson, L.M., Aiken, C.T., Kaltenbach, L.S., Agrawal, N., Illes, K., Khoshnan, A., Martinez-Vincente, M., Arrasate, M., O'Rourke, J.G., Khashwji, H., et al. (2009). IKK phosphorylates Huntingtin and targets it for degradation by the proteasome and lysosome. J. Cell Biol. *187*, 1083–1099.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tong, Y., Ha, T.J., Liu, L., Nishimoto, A., Reiner, A., and Goldowitz, D. (2011). Spatial and temporal requirements for huntingtin (Htt) in neuronal migration and survival during brain development. J Neurosci *31*, 14794–14799.

Trefilov, A., Berard, J., Krawczak, M., and Schmidtke, J. (2000). Natal dispersal in rhesus macaques is related to serotonin transporter gene promoter variation. Behav. Genet. *30*, 295–301.

Trifonov, E.N. (1989). The multiple codes of nucleotide sequences. Bull. Math. Biol. *51*, 417–432.

Tropepe, V., Hitoshi, S., Sirard, C., Mak, T.W., Rossant, J., and Kooy, D. van der (2001). Direct Neural Fate Specification from Embryonic Stem Cells. Neuron *30*, 65–78.

Truant, R., Atwal, R., and Burtnik, A. (2006). Hypothesis: Huntingtin may function in membrane association and vesicular trafficking. Biochem. Cell Biol. *84*, 912–917.

Ueno, S., Kondoh, K., Kotani, Y., Komure, O., Kuno, S., Kawai, J., Hazama, F., and Sano, A. (1995). Somatic mosaicism of CAG repeat in dentatorubral-pallidoluysian atrophy (DRPLA). Hum. Mol. Genet. *4*, 663–666.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J. Cell. Sci. *118*, 4495–4509.

Van Raamsdonk, V., M, J., Gibson, W.T., Pearson, J., Murphy, Z., Lu, G., Leavitt, B.R., and Hayden, M.R. (2006). Body weight is modulated by levels of full-length Huntingtin. Hum Mol Genet *15*, 1513–1523.

Velier, J., Kim, M., Schwarz, C., Kim, T.W., Sapp, E., Chase, K., Aronin, N., and DiFiglia, M. (1998). Wild-type and mutant huntingtins function in vesicle trafficking in the secretory and endocytic pathways. Exp. Neurol. *152*, 34–40.

Vijayvargia, R., Epand, R., Leitner, A., Jung, T.-Y., Shin, B., Jung, R., Lloret, A., Singh Atwal, R., Lee, H., Lee, J.-M., et al. (2016). Huntingtin's spherical solenoid structure enables polyglutamine tract-dependent modulation of its structure and function. Elife *5*, e11184.

Vonsattel, J.P., and DiFiglia, M. (1998). Huntington disease. J. Neuropathol. Exp. Neurol. *57*, 369–384.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364–368.

Wang, Y., Steimle, P.A., Ren, Y., Ross, C.A., Robinson, D.N., Egelhoff, T.T., Sesaki, H., and Iijima, M. (2011). Dictyostelium huntingtin controls chemotaxis and cytokinesis through the regulation of myosin II phosphorylation. Mol Biol Cell *22*, 2270–2281.

Warby, S.C., Montpetit, A., Hayden, A.R., Carroll, J.B., Butland, S.L., Visscher, H., Collins, J.A., Semaka, A., Hudson, T.J., and Hayden, M.R. (2009b). CAG Expansion in the Huntington Disease Gene Is Associated with a Specific and Targetable Predisposing Haplogroup. Am J Hum Genet *84*, 351–366.

Warby, S.C., Doty, C.N., Graham, R.K., Carroll, J.B., Yang, Y.-Z., Singaraja, R.R., Overall, C.M., and Hayden, M.R. (2008). Activated caspase-6 and caspase-6-cleaved fragments of huntingtin specifically colocalize in the nucleus. Hum. Mol. Genet. *17*, 2390–2404.

Weissman, I.L. (2000). Stem Cells: Units of Development, Units of Regeneration, and Units in Evolution. Cell *100*,

157–168.

Wellington, C.L., Ellerby, L.M., Hackam, A.S., Margolis, R.L., Trifiro, M.A., Singaraja, R., McCutcheon, K., Salvesen, G.S., Propp, S.S., Bromm, M., et al. (1998). Caspase cleavage of gene products associated with triplet expansion disorders generates truncated fragments containing the polyglutamine tract. J. Biol. Chem. *273*, 9158–9167.

White, J.K., Auerbach, W., Duyao, M.P., Vonsattel, J.P., Gusella, J.F., Joyner, A.L., and MacDonald, M.E. (1997). Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. Nat. Genet. *17*, 404–410.

Wilkinson, F.L., Nguyen, T.M., Manilal, S.B., Thomas, P., Neal, J.W., Harper, P.S., Jones, A.L., and Morris, G.E. (1999). Localization of rabbit huntingtin using a new panel of monoclonal antibodies. Brain Res. Mol. Brain Res. *69*, 10–20.

Wilson, P.G., and Stice, S.S. (2006). Development and differentiation of neural rosettes derived from human embryonic stem cells. Stem Cell Rev *2*, 67–77.

Wray, J., Kalkan, T., and Smith, A.G. (2010). The ground state of pluripotency. Biochem. Soc. Trans. *38*, 1027–1032.

Xia, J., Lee, D.H., Taylor, J., Vandelft, M., and Truant, R. (2003). Huntingtin contains a highly conserved nuclear export signal. Hum Mol Genet *12*, 1393–1403.

Xiao, L., Yuan, X., and Sharkis, S.J. (2006). Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells. Stem Cells *24*, 1476–1486.

Xu, R.-H., Chen, X., Li, D.S., Li, R., Addicks, G.C., Glennon, C., Zwaka, T.P., and Thomson, J.A. (2002). BMP4 initiates human embryonic stem cell differentiation to trophoblast. Nat. Biotechnol. *20*, 1261–1264.

Xu, R.-H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. Nat. Methods *2*, 185–190.

Xu, R.-H., Sampsell-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., et al. (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. Cell Stem Cell *3*, 196–206.

Ying, Q.-L., and Smith, A.G. (2003). Defined Conditions for Neural Commitment and Differentiation. In Methods in Enzymology, (Academic Press), pp. 327–341.

Ying, Q.-L., Stavridis, M., Griffiths, D., Li, M., and Smith, A. (2003a). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat. Biotechnol. *21*, 183–186.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003b). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281–292.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Young, L.J., Nilsen, R., Waymire, K.G., MacGregor, G.R., and Insel, T.R. (1999). Increased affiliative response to vasopressin in mice expressing the V1a receptor from a monogamous vole. Nature *400*, 766–768.

Yu, P.B., Hong, C.C., Sachidanandan, C., Babitt, J.L., Deng, D.Y., Hoyng, S.A., Lin, H.Y., Bloch, K.D., and Peterson, R.T. (2008). Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. Nat. Chem. Biol. *4*, 33–41.

Zala, D., Hinckelmann, M.-V., and Saudou, F. (2013a). Huntingtin's function in axonal transport is conserved in Drosophila melanogaster. PLoS ONE *8*, e60162.

Zeitlin, S., Liu, J.P., Chapman, D.L., Papaioannou, V.E., and Efstratiadis, A. (1995). Increased apoptosis and early

embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. Nat. Genet. *11*, 155–163.

Zhang, S., Feany, M.B., Saraswati, S., Littleton, J.T., and Perrimon, N. (2009). Inactivation of Drosophila Huntingtin affects long-term adult functioning and the pathogenesis of a Huntington's disease model. Dis Model Mech *2*, 247–266.

Zhang, S.C., Wernig, M., Duncan, I.D., Brüstle, O., and Thomson, J.A. (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat. Biotechnol. *19*, 1129–1133.

Zhang, Y., Leavitt, B.R., Raamsdonk, J.M. van, Dragatsis, I., Goldowitz, D., MacDonald, M.E., Hayden, M.R., and Friedlander, R.M. (2006). Huntingtin inhibits caspase‐3 activation. The EMBO Journal *25*, 5896–5906.

Zheng, Q., and Joinnides, M. (2009). Hunting for the function of Huntingtin. Dis Model Mech *2*, 199–200.

Zheng, S., Clabough, E.B.D., Sarkar, S., Futter, M., Rubinsztein, D.C., and Zeitlin, S.O. (2010). Deletion of the Huntingtin Polyglutamine Stretch Enhances Neuronal Autophagy and Longevity in Mice. PLOS Genetics *6*, e1000838.

Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B.R., Goffredo, D., Conti, L., MacDonald, M.E., Friedlander, R.M., Silani, V., Hayden, M.R., et al. (2001). Loss of Huntingtin-Mediated BDNF Gene Transcription in Huntington's Disease. Science *293*, 493–498.

Zuccato, C., Tartari, M., Crotti, A., Goffredo, D., Valenza, M., Conti, L., Cataudella, T., Leavitt, B.R., Hayden, M.R., Timmusk, T., et al. (2003). Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled neuronal genes. Nat. Genet. *35*, 76–83.

Zuccato, C., Valenza, M., and Cattaneo, E. (2010). Molecular mechanisms and potential therapeutical targets in Huntington's disease. Physiol. Rev. *90*, 905–981.