UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI BIOSCIENZE

DOTTORATO DI RICERCA
DBMC
IN BIOLOGIA MOLECOLARE E CELLULARE

UNIVERSITÀ DEGLI STUDI DI MILANO
Scuola di Dottorato in Biologia molecolare e cellulare
XXXIII Ciclo

Investigating the immediate consequences of normal and mutant HTT loss in HD-hESC through the dTAG system.

Manuel Cernigoj

PhD Thesis

Scientific tutor: Prof. Elena Cattaneo

Academic year: 2019-2020

SSD: BIO/11, BIO/14

Thesis performed at the Laboratory of Stem Cell Biology and Pharmacology of Neurodegenerative Diseases. Department of Biosciences, University of Milan

and Istituto Nazionale di Genetica Molecolare.

Table of Contents

Abstract

Huntington's Disease (HD) is a neurodegenerative disorder cause by a dominant CAG triplet repeat expansion (poly-glutamine(Q) in the protein) in the first exon of the Huntingtin gene (*HTT*). In the healthy population, the number of CAG repeats range between 9 and 35, while an expansion above 36 CAG repeats causes the manifestation of the pathology later in life. Numerous studies indicate that the pathological CAG length causes the poly-Q expanded protein to acquire a toxic function which ultimately kills the neurons[1–10]. Other studies show that several disfunctions associated with the presence of the mutant protein can be phenocopied in cells and mice deleted of the healthy gene, suggesting that a loss of function mechanism may also contributes to HD[11–13]. Thus the disease seems to be the result of a two-component pathological mechanism; first a loss-of-function (LOF) due to reduced normal HTT physiological activity, and second a gain-of-function (GOF) due to mutant HTT toxic effects[14]. Being able to discriminate between LOF and GOF phenotypes is especially relevant when considering the ongoing gene silencing clinical studies aiming at reducing HTT level either in an allele-specific or non-allele-specific manner[15–17]. On top of that, transcriptional alterations have been reported both in HD patients and HD mouse models[18], and several evidence have linked HTT to gene expression[1,19–27]. However, these data refer to situations in which a transcriptional balance has already been established following the genetic perturbation. No data is available regarding the mechanisms acutely implemented by the cell immediately after perturbation to re-establish transcriptional balance in response to those changes. For these reasons, in this project we aimed at generating a human embryonic stem cell (hESC) HD model in which normal or mutant HTT can be rapidly and efficiently removed upon exposure to a small molecule. This is made possible by a degradation tag (dTAG)[28] that, when fused to HTT, induces a proteosome-mediated degradation upon exposure to a cell permeable ligand. Accordingly, the first part of the project consists in the generation of two dTAG-hESC-HD lines in which either the normal or the mutant allele were targeted by Cas9-assisted genome editing followed by the assessment of the efficiency and degradation kinetic of tagged HTT in both cell lines. These experiments revealed complete tagged HTT loss between one and two hours of treatment with a ligand concentration of 10^{-7} M. Moreover, no difference was observed when comparing degradation kinetics of normal and mutant HTT in self-renewing dTAG-hESC lines. We aim to use these lines to study the immediate transcriptomic changes and the potential compensatory mechanisms established by the cells in response to normal or mutant HTT protein depletion. Therefore, a first RNA-seq experiment was performed to investigate over time transcriptional changes during dTAG-HTT degradation. This experiment revealed no substantial transcriptomic changes between normal and the HTT-depleted state of self-renewing dTAG-hESC HD lines. We are now looking into the transcriptional changes driven by HTT degradation in hESC *in vitro* derived neurons that can provide useful information on the biosafety of the ongoing HTT-lowering approaches for HD treatment.

Abstract (Italian)

La malattia di Huntington è una malattia neurodegenerativa causata da una mutazione genica dominante, che consiste in un'espansione di triplette CAG (poli-glutammina (Q) nella proteina) presenti nel primo esone del gene Huntingtin (*HTT*). Nella popolazione sana, il numero di ripetizioni CAG è compreso tra 9 e 35, mentre un'espansione sopra le 36 ripetizioni CAG causa il manifestarsi della patologia in età adulta. Numerosi studi indicano che la lunghezza patologica del tratto CAG fa sì che la proteina con poli-Q espanso acquisisca una funzione tossica, che porta alla morte delle cellule neuronali[1–10]. Altri studi hanno dimostrato che, gli stessi fenotipi riscontrati in presenza della proteina mutante, sono stati riscontrati anche in cellule e topi in cui gene sano è stato inattivato, suggerendo la presenza di un meccanismo di tipo loss-of-function in HD[11–13]. Perciò, la malattia sembra essere il risultato di un meccanismo patologico a due componenti; la prima è la loss-of-function (LOF) a causa della ridotta attività fisiologica dell'HTT normale, la seconda è la gain-of-function (GOF) dovuta agli effetti tossici dell'HTT mutante[14]. Essere in grado di discriminare tra fenotipi di tipo LOF e GOF è particolarmente importante quando si prendono in considerazione gli studi clinici sul silenziamento genico attualmente in corso e volti a ridurre il livello di HTT in modo allele-specifico o non-allele-specifico[15–17]. Inoltre, diverse alterazioni trascrizionali sono state riscontrate sia nei pazienti HD che nei modelli murini HD e diversi lavori hanno ricollegato l'HTT all'espressione genica[1,19–27]. Tuttavia, questi dati si riferiscono a situazioni in cui è già stato stabilito un equilibrio trascrizionale a seguito della perturbazione genetica. Non sono disponibili dati sui meccanismi implementati acutamente dalla cellula immediatamente dopo la perturbazione per ristabilire l'equilibrio trascrizionale in risposta a tali cambiamenti. Per queste ragioni, in questo progetto abbiamo mirato a generare un modello di staminali embrionali umane (hESC) HD in cui l'HTT normale o mutata possono essere rimosse rapidamente ed efficacemente, in seguito all'esposizione ad una small-molecule. Ciò è reso possibile da un tag di degradazione (dTAG)[28], il quale, quando fuso all'HTT, induce una degradazione proteosoma-mediata in risposta all'esposizione ad un ligando permeabile alle cellule. La prima parte del progetto consiste nella generazione di due linee dTAG-hESC-HD, in cui l'allele normale o mutante sono stati ingegnerizzati geneticamente tramite il sistema CRISPR/Cas9, e successivamente, nella valutazione dell'efficienza e della cinetica di degradazione dell'HTT-taggata, in entrambe le linee cellulari. Questi esperimenti hanno rivelato una completa rimozione di HTT-taggata nel corso di una-due ore di trattamento utilizzando una concentrazione di ligando pari a 10^{-7} M. Inoltre, non è stata osservata alcuna differenza in termini di cinetica di degradazione tra HTT normale e mutante confrontando le due linee dTAG-hESC in self-renewal. Intendiamo dunque utilizzare queste linee per studiare i cambiamenti trascrizionali immediati e i potenziali meccanismi compensatori attuati dalla cellula in risposta alla deplezione della proteina HTT, che sia essa normale o mutante. Pertanto, è stato eseguito un primo esperimento RNA-seq per studiare i cambiamenti trascrizionali nel tempo durante la degradazione di dTAG-HTT. Questo esperimento non ha rivelato cambiamenti sostanziali nel trascrittoma delle linee HD dTAG-hESC in self-renewal confrontando le cellule in cui l'HTT è stata rimossa con le cellule in cui i livelli di HTT sono rimasti inalterati. Siamo ora in procinto di esaminare i cambiamenti trascrizionali provocati dalla degradazione dell'HTT nei neuroni derivati *in vitro* da hESC, i quali potranno fornire informazioni utili sulla bio-sicurezza degli approcci di riduzione dell'HTT, in corso per il trattamento della malattia di Huntington.

Aim

Several works in the literature have linked HTT to transcriptional regulation. Indeed, HTT appears to interact with different transcription factors providing a scaffold to mediate interactions with the basal transcriptional machinery. Moreover, data derived from post-mortem brains of HD patients and transgenic mouse models compared to their respective healthy controls point to the presence of a transcriptional dysregulation in HD. All these data were collected from cells or organisms having no functional protein or a mutant version of the protein a long time after the genetic perturbation has occurred. During this period of time, cells transcriptome must have undertaken several adjustments before reaching a steady-state, effectively preventing the observation of the immediate mechanisms put in place by the cell in response to the lack of healthy HTT or to the presence of its mutant version. In fact, no previous work was able to investigate the immediate consequences to normal or mutant HTT deprivation. Hence, the main objective of this thesis work is to investigate immediate transcriptional response to normal and mutant HTT loss. To achieve this several steps are necessary:

1. Genetic engineered of a HD hESC line to insert a protein degradation tag (dTAG) either on the normal allele (dTAG21Q) or on the expanded allele (dTAG56Q)
2. Determine the efficacy and kinetics of the system in both dTAG cell lines.
3. Investigate the immediate transcriptional response to wt or mut HTT degradation in both self-renewing and hESC-derived neurons

1 | Introduction

1.1 Huntington's disease

1.1.1 Pathological features

Huntington's disease (HD) is an autosomal dominant neurodegenerative disorder that arises upon the abnormal expansion of a CAG repeat region residing in the first exon of the *IT-15* gene. The expanded CAG tract is translated into a pathogenic poly-glutamine (poly-Q) starch that alters the physiological conformation of the protein Huntingtin (HTT), impairing its function and making it prone to form aggregates. This condition, over time, leads to progressive neurodegeneration with the medium spiny neurons (MSNs) in the corpus striatum being the most affected cell type[29]. In the healthy population, the CAG tract displays between 9 and 35 repetitions with an average of 17-20 repetitions[30]. Individuals with more than 36 CAGs will likely face the manifestation of the disease around the third-fourth decade of their life. Initially, patients start to experience involuntary movements and motor dysfunctions. These are often shown together with mental and emotional symptoms. In late-stage HD, the patients manifest rapid cognitive decline, weight loss, and even speech difficulties[31]. Normally, after about 15 to 20 years from diagnosis, the patients succumb to pneumonia, heart failure, or other complications. At present, there is no disease-modifying treatment available for HD. However, huntingtin lowering through antisense oligonucleotides (ASOs)[32] is emerging as a promising approach to treat this difficult condition, given the positive results achieved in several preclinical experiments[16,33,34] . Still, this approach presents some limitations. The patients must be treated chronically with the ASOs to sustain the effect over time and the method of administration is fairly invasive. Additionally, current ASOs are not selective for the pathologic allele but inducing the reduction of expression of the non-expanded allele as well, thus, potential adverse effects induced by excessive reduction of normal HTT must be taken into consideration.

1.1.2 Huntingtin protein

The genetic product of the *IT-15* gene is a protein named HTT. It is a large protein composed of 3,144 amino acids (348 kDa), a size that poses a challenge for structural analysis. Thankfully, its full-length three-dimensional structure in complex with the HTT-associated protein 40 (HAP40) has been recently resolved by cryo-electron microscopy shading light on the stereological conformation of the protein[35]. HTT is composed mainly of α-helices, most of which are arranged in HEAT repeats or other tandem repeats. The acronym HEAT comes from four proteins that were originally found to contain this repeated motif, that is Huntingtin, elongation factor 3, the A subunit of protein phosphatase 2A (PP2A) and the signalling kinase TOR1[36]. The HEAT motif consists of two alpha-helices (α- and β-helices), arranged in an antiparallel fashion and connected by a short linker (turn). Conventionally, multiple HEAT motifs are stacked with each other to form a two-layer curved structure. In this structure, each helix forms weak hydrophobic interactions with the helices nearby. Thanks to these weak interactions, HEAT repeats result highly flexible and elastic having the potential to undergo large conformational changes by either interacting with other proteins or responding to external forces or environmental changes[37] . These properties allow HTT to interact with many different partners by adopting various three-dimensional conformations to support various biological processes.

HTT is composed of N-terminal and C-terminal domains containing multiple HEAT repeats arranged to form a solenoid-like structure and a third domain forming a bridge of tandem repeats that connect the other two domains (Fig. 1).

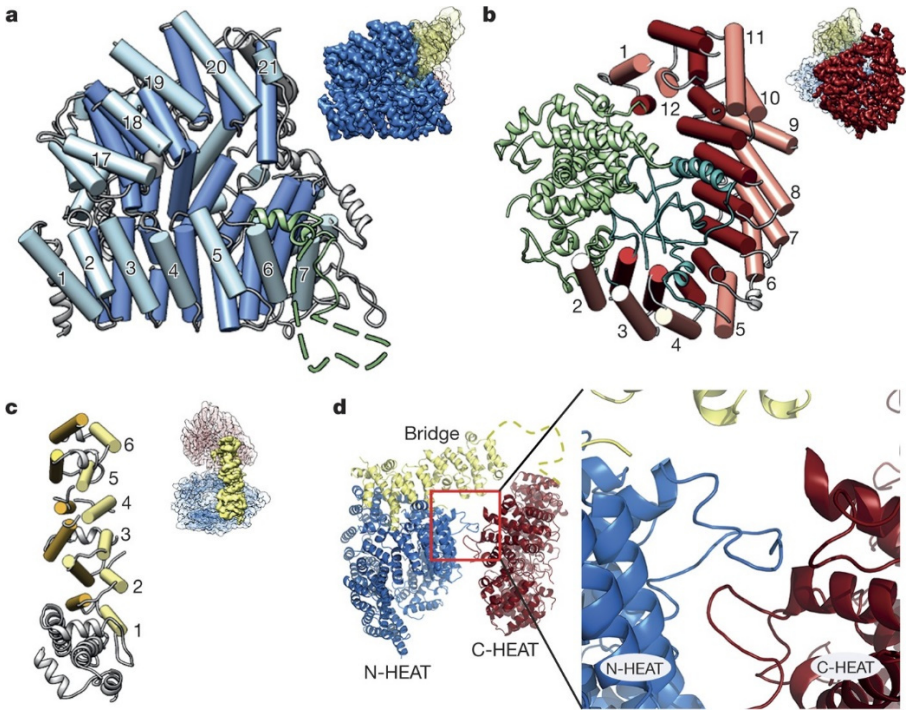

Fig. 1 | Tridimensional structure of HTT obtained by cryo-EM.
a. N-terminal HEAT domain. **b.** C-terminal HEAT domain. **c.** Bridge domain. In **a-c** helices forming part of tandem repeats are shown as rods in similar colors, and other helices as ribbons. At the top right of **a-c**, the part of Htt shown in each panel is highlighted in the Htt-HAP40. **d.** Back view of Htt highlighting the interaction region (inset) between loops of N- and C-HEAT. The unresolved sequence at the C-terminus of the bridge domain is shown as a yellow dashed line[35].

Thus, HTT tertiary structure consists of 3 major domains.

N-HEAT domain:

The N-terminal HEAT domain (Fig. 1 a) is composed of 21 HEAT repeats arranged as a one-and-a-half-turn right-handed superhelix. This domain contains two putative membrane-binding regions. The first is located in the first 17 residues that may form an amphipathic helix. The second lies in the region between residues 168 and 366, in this region the HEAT repeats 2-4 are positively charged, potentially stabilizing the membrane-binding thanks to ionic interaction. The N-HEAT domain also contains a large unstructured portion that projects outwards and is localized between HEAT repeats 6 and 7. This insertion may modulate protein-protein interactions and proteolytic accessibility thanks to the phosphorylation sites it contains. This portion also contains various proteolytic cleavage sites. Nonetheless, since this region is tightly packed, it is unlikely that its proteolytic cleavage can generate N-terminal fragments.

C-HEAT domain:

The C-terminal HEAT domain (Fig. 1 b) is shaped as an elliptical ring made of 12 HEAT repeats. It contains two insertions, the first is localized between HEAT repeat 1 and 2 and consists of 12 helices, whereas the second insertion separates HEAT repeats 2 and 3 and is mostly unstructured. Both of them are exposed on the concave surface of the C-terminal domain and may shield this side of the ring from protein–protein interactions.

Bridge domain:

The bridge domain (Fig. 1 c) is much smaller compared to the other two. Its structure is mostly α-helical but with no HEAT motif. It contains a portion of six tandem helical repeats flanked by five non-repeat helices and a flexible portion at the C-terminus.

HTT post-translational modifications:

HTT is subjected to proteolysis at several sites by a variety of proteases, including caspases, calpains, cathepsins and aspartyl proteases[38–45]. HTT cleavage sites are mostly localized in a disorganized region rich in proline (P), glutamic acid (E), serine (S), and threonine (T) known as PEST sites[46]. *In vitro* tests demonstrated no difference in terms of proteolysis susceptibility between normal and mutant HTT[39]. Nevertheless, in the brain of HD patients, HTT proteolysis was found increased and led to the accumulation of short N-terminal fragments containing the expanded polyQ stretch both in the cytoplasm and in the nucleus. These fragments are reported to be toxic for the cell since their translocation in the nucleus triggers transcriptional dysregulation as well as defects in autophagy, signal transduction and both calcium and mitochondrial homoeostasis. HTT is also subjected to various post-translational modifications, including phosphorylation, acetylation, ubiquitylation, sumoylation and palmitoylation. In HD, some of these modifications can modulate mutant HTT toxicity. For instance, acetylation of mutant HTT can increase its clearance by the autophagic pathway. Similarly, phosphorylation at S13 and S16 promotes clearance of both wild-type and mutant polyQ-HTT reducing its toxicity[47–50]. Conversely, phosphorylation of mutant HTT at S434 or S536 reduces its proteolysis and the generation of toxic fragments. HTT phosphorylation also regulates many of its physiological functions. For instance, the phosphorylation state of S421 and S1181/S1201 regulates microtubules-dependent intracellular transport of organelles. Overall, a variety of post-translational modification affects HTT protein behaviour and regulation.

1.1.3 Biological processes regulated by HTT

As previously mentioned, HTT flexible structure allows it to interact with different partners and participate in many biological processes concerning cellular dynamics (adhesion, cytoskeleton, endocytosis and trafficking), metabolism, gene expression and signal transduction[29,51,52] . In all these processes, HTT appears to play a scaffolding role. It binds to multiple members of different pathways tethering them into complexes and helping localize them to specific areas of the cell such as the plasma membrane, the cytoplasm, the nucleus, the Golgi, the endosomes and the mitochondria. However, in presence of the HD mutation, many HTT-regulated processes are affected at different levels. Among these, some of the most relevant for disease progression involves vesicular trafficking along the axon, protein clearance and gene transcription.

Axonal vesicle-transport:

The axonal transport is responsible for the movement of synaptic vesicles, lipids, proteins, trophic factors and even organelles to and from the soma of the neurons along the microtubules. The transport can occur in two directions: when the cargos are moving from the soma to the synapse is named anterograde transport, when the movement is in the opposite direction is named retrograde transport (Fig. 2 A). These movements are possible thanks to motor proteins, bio-molecular machines that produce movement upon energy consumption. Kinesin is the motor protein responsible for anterograde transport, whereas the dynein-dynactin complex is responsible for retrograde transport. HTT is involved in the regulation of both the anterograde and retrograde axonal transport at different levels (Fig. 2 B). HTT interacts with HAP1 which in turn bind to the $p150^{Glued}$ subunit of the dynactin complex. This association influences the mobility and/or cargo binding to dynein. Thus, HTT and HAP1 serve as scaffolding proteins to stabilize the dynein-

dynactin complex and enabling the transport of the cargo along the axon[53–61] . Moreover, HTT can bind and maintain GAPDH on the vesicles being transported along the microtubules allowing GAPDH to supply energy to speed up the vesicles transport[62]. Finally, HTT also controls the directionality of the axonal transport through its phosphorylation state on S421. In fact, when HTT is in its phosphorylated state at this particular site, it recruits kinesin-1 to the dynactin complex, promoting the anterograde transport. Conversely, unphosphorylated HTT is not able to bind kinesin-1, thus favouring retrograde transport[54,63–65].

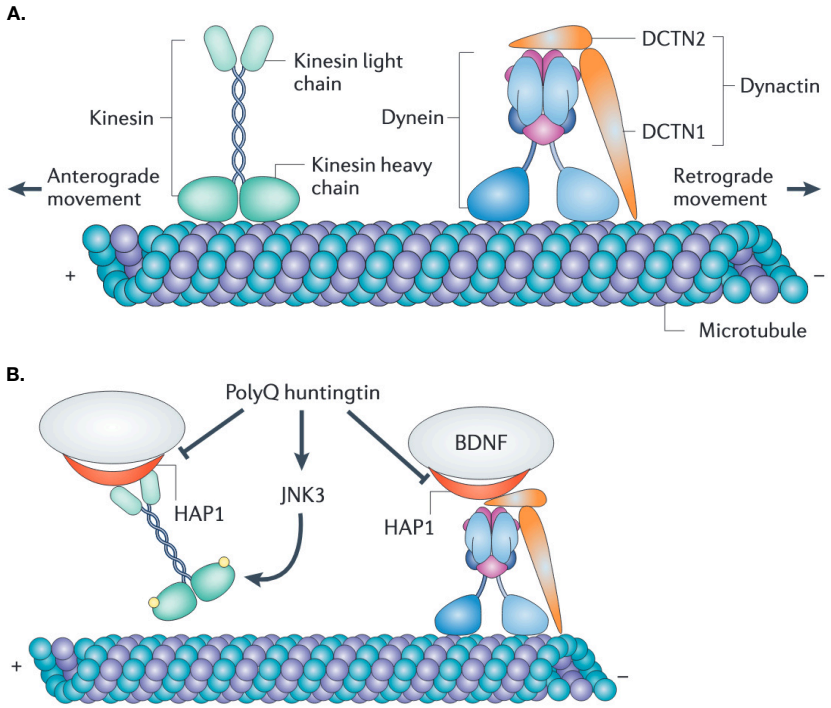

Fig. 2 | The axonal transport machinery.
A. Kinesin and dynein are responsible for the axonal transport of most cargoes. Kinesin is composed of two heavy chains and two light chains and transport cargoes in the anterograde direction towards the plus end of microtubules. Dynein complexes are composed of different dynein chains and several dynactin subunits. Those complex move cargoes in the retrograde direction towards the minus end of the microtubules. Both kinesin and dynein have a glomerular motor domain in their heavy subunits that binds to microtubules and hydrolyses ATP to propel cargoes along the microtubule rails. **B.** In HD, polyQ huntingtin impairs both retrograde and anterograde axonal transport. Huntingtin-associated protein 1 (HAP1) is an adaptor protein that interacts with kinesin light chains and dynactin subunit 1 (DCTN1). PolyQ huntingtin disrupts HAP1-mediated axonal transport, including that of brain-derived neurotrophic factor (BDNF). PolyQ huntingtin also promote the activation of c-Jun N-terminal kinase 3 (JNK3) which in turn phosphorylates (yellow circles) the kinesin motor domain, preventing its binding on the microtubules.

Protein clearance:

Intracellular protein clearance can occur through two different pathways: autophagy and the ubiquitin-proteasome system (UPS). The **autophagic process** is fundamental for the clearance of damaged organelles and elimination of toxic or aggregated protein by lysosomal degradation. The process is made of multiple steps (Fig. 3 A). The first step is the initiation, in which the isolation membrane is formed. The second step is the elongation of the isolation membrane that grows in size to encircle the protein cargos. The third step is the cargo loading of the isolation membrane mediated by LC3. Next, the membrane closes on itself forming the autophagosome, this step is named maturation. At this point, the autophagosome is transported along the microtubules to reach the lysosome for the fusion step. Finally, the newly formed autophagolysosome proceeds with the degradation of the protein cargos.

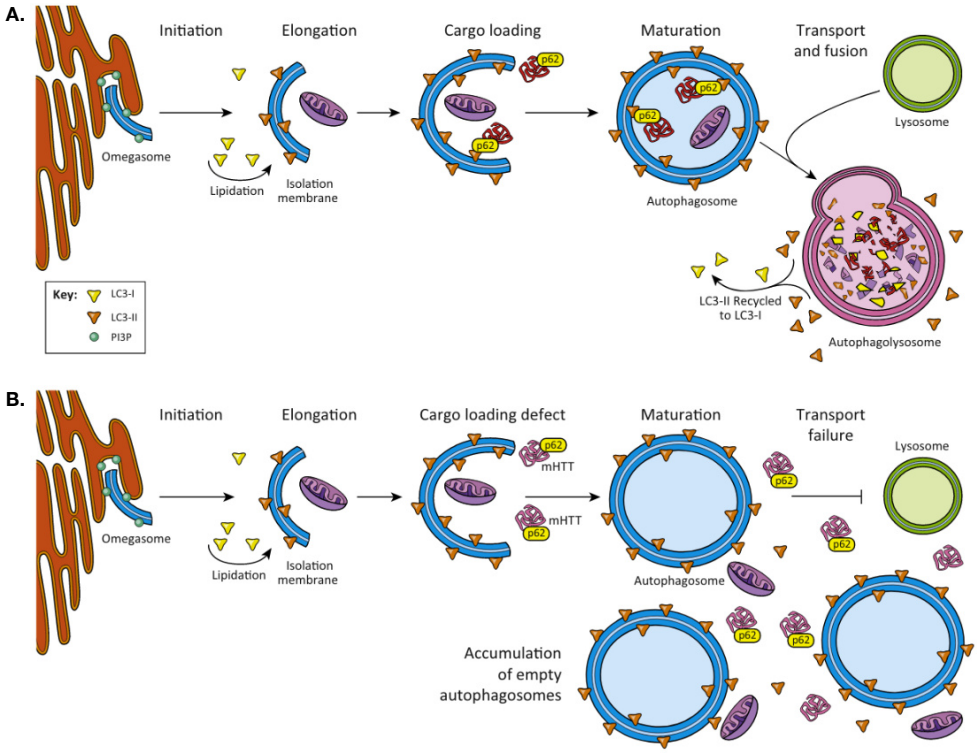

Fig. 3 | Autophagic alterations in HD.
A. The normal autophagic pathway involves the formation of the isolation membrane that is loaded with the cargos to be degraded. After maturation, the autophagosome is transported along the microtubules to the lysosome with which it fuses generating the autophagolysosome. **B.** In HD several steps of this process are affected, including the cargo loading, the transport and the fusion of the autophagosome with the lysosome[66].

Although autophagy was initially thought to act as a bulk degradation mechanism, it has recently been demonstrated that autophagy can degrade specific cellular components like mitochondria, ribosomes, ER, or aggregated proteins such as mutant HTT. This type of autophagy is named selective autophagy. In HD the autophagic flux is dysregulated and HTT seems to play a direct role in controlling the autophagic process (Fig. 3 B). Indeed, HTT depleted cells display lower basal rates of autophagic degradation and, as emerged from human and rodent HD samples, autophagosome number is increased in HD compared to wt, whereas the autophagic flux is fairly similar to the wt condition. Deeper analyses demonstrated that most of the autophagosomes formed in HD are empty or contain mixed cargo composition. This is due to a defect at the cargo recognition step. In addition, muHTT impairs autophagosome mobility, therefore preventing its fusion with the lysosome and further increase the number of empty autophagosomes. This means that the enhanced autophagic flux and formation of the autophagosomes together with the defective cargo recognition and mobility lead to an accumulation of empty autophagosomes, causing a feedback loop of upregulation of a defective process. Mutant HTT also interferes with the autophagic process by sequestering mTOR, further upregulating the autophagic pathway. In addition, as mentioned earlier, the cargo loading is mediated by LC3 which recognize and bind to specific peptide sequences named LC3-interacting repeat (LIR) present on the cargo. *In silico* analysis predicted 11 LIR motifs within full-length HTT, however, only one of them is present in the N-terminal part of the protein. Thus, N-terminal fragments produced from HTT proteolytic cleavage are poorly loaded into the autophagosome, leading to an even larger accumulation of HTT N-terminal fragments, which are toxic in the HD context[1–6].

The other system dedicated to maintaining protein homeostasis is the **ubiquitin-proteasome system (UPS)**. Briefly, this system is composed of three enzymes (E1, E2, E3) that act in sequential order to prepare the protein for the degradation. The process consists in covalent attachment of a

poly-Ubiquitin chain that marks the protein for degradation by the proteasome. The system will be discussed more in details in section 1.3.3. The role of this system in HD hasn't been fully elucidated yet. Nevertheless, the accumulation of poly-Ubiquitin chains was observed in brains of both HD mouse models and HD patients. Despite these observations, *in vivo* studies in different mouse models of HD, revealed no alterations in the functionality of the UPS system. This apparent controversy was resolved by Ortega and colleagues who demonstrated that the UPS system results transiently dysfunctional after mutant HTT expression. This condition is rapidly reverted once inclusion bodies (IBs) start to form suggesting a protective role of the IBs against proteotoxic stress. The impairment of the UPS seems to be indirectly triggered by mutant HTT in different ways (Fig. 4). First, HTT gradually saturates HSP40 and HSP70 chaperones decreasing the pool of soluble chaperones. This leads to an accumulation of other chaperone clients that are sent to the proteasome for degradation. This cause overloading of proteasome activity that generates stress and causes the failure of the system. Second, mitochondrial dysfunctions, that are also present in HD, imply a reduction in ATP levels which in turn affect many energy-dependent pathways, including protein clearance systems.

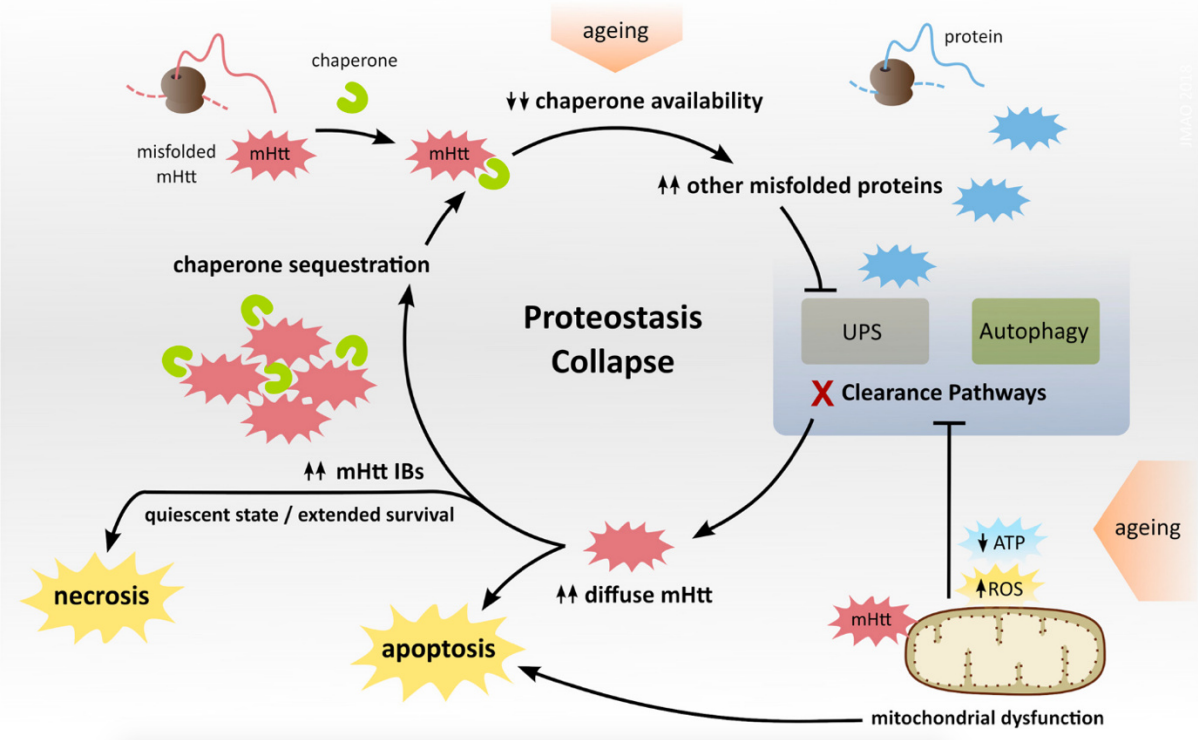

Fig. 4 | Proteostasis collapse in HD.
mHTT gradually titrates chaperones and in combination with ageing effects reduce the pool of available chaperones. This increase the quantity of other misfolded proteins which saturates the proteasome and cause UPS dysfunctions[8,9]. In addition, mitochondrial dysfunction causes an increase of ROS and a decrease in ATP production, impairing energy-dependent protein clearance pathways. Proteostasis collapse favors mHTT accumulation thereby promoting apoptosis and HTT aggregation into inclusion bodies (IB). IBs initially have a protective role against proteotoxicity but in the long term they disrupt cellular homeostasis, potentially leading to necrosis[10] .

Gene transcription:

HTT is largely cytoplasmatic but it is also present in the nucleus and plenty of evidence linked HTT with a large number of transcription factors (TFs). First, motives similar to the poly-Q tract of HTT are present in some transcription factors where they serve as a mediator of the binding between the TF and transcriptional regulators. Second, HTT interacts with several TFs, such as NeuroD, p53, CBP, CtBP, CA150, NF-kB, REST/NRSF[1,19–24] . The interaction with REST is particularly interesting

since NRSF negatively regulates genes that are essential for neuronal development. Among them, the brain-derived neurotrophic factor (BDNF) gene is of particular interest since it is fundamental for MSNs survival and has been found dysregulated in HD[19] . BDNF needed by MSNs is mostly produced and provided by cortical neurons. In HD, however, mutant HTT affects BDNF transcription in cortical neurons, leading to a reduction of BDNF production. HTT also binds to SP1 and the co-activator TAFII130 that function as a general transcription factor[25,26]. This indicates that HTT might act as a scaffold protein that mediates the interaction of SP1 with the basal transcriptional machinery. Other evidence indicates that HTT might also modulate gene repression by binding to the N-CoR–Sin3a repressor complex. This repressor blocks the transcription from ligand-activated nuclear receptors, for instance, the retinoic acid receptors[67–70] . Besides, HTT is a regulator of chromatin remodelling as it modulates PRC2-mediated histone-methylation. PRC2 can mark DNA regions to become transcriptionally silent through methylation of the histone H3 on the lysine 27 (H3K27me3). HTT regulates PRC2 activity thanks to the interaction with two of its subunits, Ezh2 and Suz12[27]. Finally, HTT also binds to nuclear receptors such as LXR-α, VDR, TRα1 and PPARγ, adding another layer of regulation to HTT-mediated gene transcription[71].

Other biological process regulated by HTT

In addition to those discussed above, HTT is also involved in many other biological processes like clathrin-mediated endocytosis, synaptic organization and apoptosis. The role of HTT is again mostly to provide a scaffold to facilitate the interaction between other proteins.

1.2 Modeling Huntington's Disease.

Huntington's Disease (HD) is a monogenic dominant disorder where the mutation on a single gene (HTT) is sufficient to cause the pathology. This is advantageous in terms of disease modelling. A wide variety of in vitro and in vivo models have been generated in the past years often based on transgenic rodents or rodents-derived neurons expressing polyQ-expanded HTT. These animal and cellular models have been particularly useful to disclose many aspects of the disease, however, the differences between rodents and humans, including the increased complexity of the central nervous system72 and the difference in the non-pathological number of CAG repeats (9 to 35 in the humans versus 7 in the mouse), limit the possibility to investigate many pathological mechanisms of the disease. Therefore, the generation of human in vitro models is fundamental to inform on human-specific disease mechanisms. Among those, induced pluripotent stem cells (iPSC) and genetically engineered human embryonic stem cells (hESC) are the most valuable for modelling and screening purposes thanks to their ability to proliferate indefinitely and to specialize into all possible embryonic cell types. In fact, many HD-mediated alterations have been described in PSCs by different studies. Self-renewing stem cells already recapitulate some of the disease hallmarks, such as CAG instability, mitochondrial dysfunctions, apoptosis, oxidative stress, autophagy and signalling alterations73–80. However, since HD is an adult-onset disease affecting mostly neuronal cells, it is far more relevant to investigate disease mechanisms within the relevant cellular context. To our advantage, pluripotent stem cells can be instructed to become neurons by providing them with specific molecular stimuli. The combination, timing and intensity of those stimuli are crucial to obtain the desired cell type. Various in vitro differentiation protocols have been developed during the last years taking inspiration from the cellular events that make up the embryo during development. Among those, our laboratory has published a protocol for the generation of authentic medium spiny neurons (MSNs) representing the most affected neuronal type in HD. Such protocol couples neural induction via BMP/TGF-β inhibition with exposure to the developmental morphogens controlling dorso-ventral and anterior-posterior axes such as sonic hedgehog (SHH) and the WNT signalling inhibitor dickkopf1 (DKK1). In the first phase of the differentiation, dual

SMAD inhibition is applied to downregulate BMP and TGF-β pathways and force the cells out of the pluripotent state and acquire a neuroectodermal identity. In the second phase, SHH drives neuroectodermal cells towards a ventral telencephalic identity and, concomitantly, DKK1 pushes the cells towards rostral fate. Fine tuning of these processes is able to instruct cells with specific positional identity to become progenitors of the lateral ganglionic eminence (LGE), the region that gives rise to the striatum and to MSNs. In the third phase, already committed progenitors are left to autonomously complete the differentiation in a stimuli free medium, supported by neurotrophic factors, such as N2, B27, retinoic acid and BDNF. During this procedure, the expression of specific markers of differentiation is monitored to ensure that the correct differentiation route is followed. In particular, neural induction is confirmed upon PAX6 expression and OCT4 downregulation. Afterwards, by the end of specification and patterning cells should result $FOXG1^{+}/GSX2^{+}$, with ASCL1 appearing shortly after. Finally, as the cells progress towards the end of differentiation, they should express markers such as CTIP2, FOXP1, FOXP2, MAP2 and GABA. Ultimately, cells that reach authentic MSN identity are marked by the co-expression of CTIP2 and DARPP32. Thanks to neuronal differentiation protocols, many other HD-related alterations were identified and affect, among others, property like neurogenic potential, proliferation, cell cycle progression and the electrophysiological property of the derived neurons11,74,81–86. All these evidence point to a developmental deficit occurring in presence of mutant-HTT that affect the stem cells biology, the differentiation process and ultimately the generated neurons. For these reasons, it is of great advantage to investigate HD pathogenesis during the differentiation of these cell models, to spot early developmental processes contributing to the pathology.

1.3 Approaches to study protein function

To study a gene's function, traditional genetic approaches often rely on its inactivation through genetic Knock-Out (KO) that completely abolishes the ability of the locus to generate a functional protein. Thanks to this approach, the functional element of the gene is removed from the cellular system and manifestation of a phenotype can be attributed to the targeted gene. Information regarding the missing functionality can be extracted from the observed phenotype and this allows to pave the way toward the understanding of the biological function of the gene under investigation. Generally, the restoration of the activity through ectopic expression of a functional copy of the gene can revert the observed phenotype. As the central dogma of the molecular biology states, the genetic information is encoded in the DNA (gene) than transcribed in (RNA) to be finally translated into an active molecule (protein) that performs the biological function[72]. Thus, the activity of the gene under study can be suppressed at three different levels, as described in the following paragraphs.

1.3.1 Disruption at the DNA level (gene KO)

A DNA double-strand break (DSB) force the cell to activate one of its DNA repair mechanisms of which the most important ones are Non-Homologous End Joining (NHEJ) and Homologous Recombination (HR) (Fig. 5). The NHEJ pathway is the most active and it recruits the DNA binding component KU70/80 to the damaged ends to protect them from exonucleases. Then, after a step of DNA end processing, the ligase IV enzyme is recruited to seal the damaged DNA. This process is not error-free and often few nucleotides are lost or random nucleotides are inserted at the damaged site. The HR pathway is the most reliable since it uses the DNA from the homologous chromosome or sister chromatid as a template to re-write the lost information. Here, the damaged ends are subjected to resection, a process that trims one of the two strands generating protruding ends. The missing DNA sequence is then *ex novo* synthesized using the homologous DNA as a

template. This pathway, however, is poorly active throughout the cell cycle, except for the mitotic phase. Traditionally, homologous recombination was the preferred method to generate a gene knockout. More recently, site-specific nucleases have been leveraged for gene KO with higher efficiency since they rely on the NHEJ pathway to repair the DNA instead of the HR. Currently, the CRISPR-Cas9 system is mostly preferred for the generation of genetic KO thanks to its flexibility and ease of use (see paragraph 1.4). Still, this kind of approach presents many limitations. Genetic KO is irreversible unless a functional copy of the gene is used to restore the initial condition. In addition, a compensatory mechanism to bypass the loss-of-function of the target gene might be established once the cell line is generated. For instance, depending on the strategy design, the target DNA sequence or the resulting insertion/deletion obtained upon Cas9 cleavage may still allow for a truncated form of the target protein to be present and retain part of the original functionality. Moreover, the nucleases might generate off-target cleavage in a different genomic region, thus potentially affecting unrelated coding sequences therefore altering the phenotype under investigation. Most importantly, essential genes cannot be analysed with this approach since the cell would die before any phenotype can be observed.

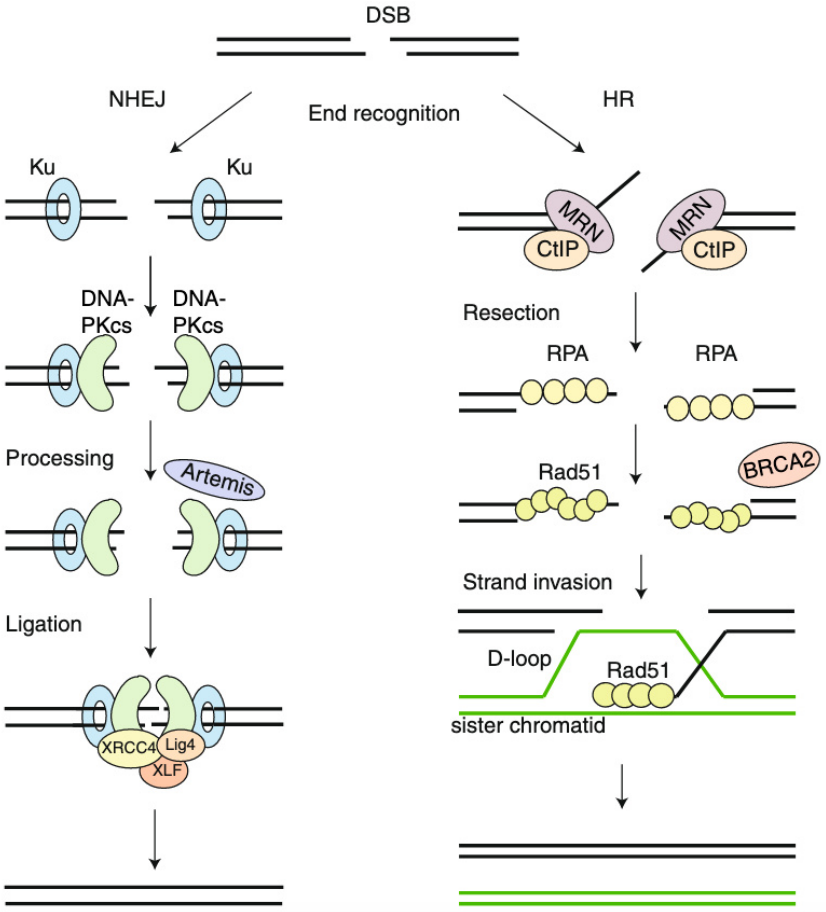

Fig. 5 | NHEJ and HR
NHEJ) Upon double strand break, the DNA free ends are recognized by the Ku70/80 heterodimer, which recruits DNA-PKcs. If the ends are incompatible, nucleases such as Artemis can trim the ends. The XRCC4-DNA Ligase IV-XLF ligation complex seals the break. **HR)** The MRN-CtIP-complex starts resection on the breaks to generate single stranded DNA (ssDNA). The ssDNA is first coated by RPA, which is subsequently replaced by Rad51 with the help of BRCA2. These Rad51 nucleoprotein filaments mediate strand invasion on the homologous template. Extension of the D-loop and capture of the second end lead to repair[73].

1.3.2 Silencing of the mRNA (RNAi)

The second strategy aims at interfering with the gene expression at the RNA level and is based on the RNA-interference (RNAi) mechanism. RNAi relies on RNA-mediated RNA-degradation in which an RNA molecule (siRNA, miRNA or shRNA) recognizes and binds to the target RNA triggering its elimination mediated by the RNAi-induced silencing complex (RISC) (Fig. 6). This approach can be applied in a much shorter time-frame respect to the generation of a KO cell line and it can also be time-controlled, up to some degree. Additionally, it allows the targeting of essential gene products broadening the range of possible targets. It can also be used to target specific isoforms based on unique sequences as the target of the iRNA strategy or it can be used to target gene families based on shared target sequences. However, off-target effects are also possible and complete silencing is unlikely since silencing efficiency depends on the protein turnover. On top of that, this approach may require prolonged treatment to sustain the silencing effect over time.

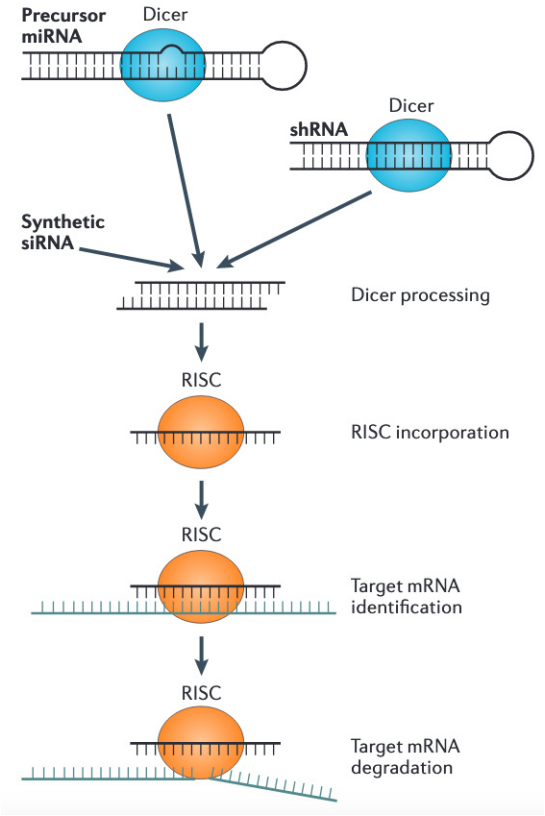

Fig. 6| RNA silencing scheme
Dicer generates microRNAs (miRNAs) from RNA hairpins encoded by miRNA genes. The Dicer machinery also processes exogenous double-stranded sequences containing loops called short hairpin RNAs (shRNAs). Synthetic siRNAs bypass Dicer processing and can directly associate with the RNA-induced silencing complex (RISC) to mediate recognition of target mRNAs through base-pair complementarity. Argonaute 2 in the RISC complex enzymatically cleaves the target mRNA, leading to target gene silencing[74].

1.3.3 Degradation of the protein (Targeted proteolysis)

Recently, a set of tools to silence gene expression *via* protein degradation has widened the range of possibilities to investigate protein functions. The concept is to hijack the ubiquitin-proteasome system (UPS) to efficiently and rapidly eliminate the target protein. The UPS acts based on the covalent attachment of Ubiquitin (Ub) to primarily lysine residues on the target protein to mark it for destruction. This process is named ubiquitination and take place thanks to the activity of three

enzymes, E1, E2 and E3 (Fig. 7). E1 is responsible for the first step in which it consumes ATP to self-ubiquitinate, activating the Ub monomer. E2 transfers the activated Ub onto one of its own lysine residues in a process named conjugation. Finally, E3 already complexed with the target protein mediates the ligation of Ub from the E2 to the target protein.

Targeted proteolysis is rapidly becoming popular with several different systems designed to leverage this process to remove the protein of interest. Each of these systems involves the recruitment of the target protein to an E3 ligase to mark it for degradation. This approach is advantageous because is independent of all the regulating processes occurring at the DNA or RNA level that can bypass the inactivation or lead to inconsistent degradation. In addition, as for the RNAi, targeted proteolysis is suitable to target essential genes since the removal of the gene of interest can be triggered on-demand in a time-controlled manner. On the other hand, some of this system can be time-consuming to develop and some are limited by the availability of targeting components and E3 ligases.

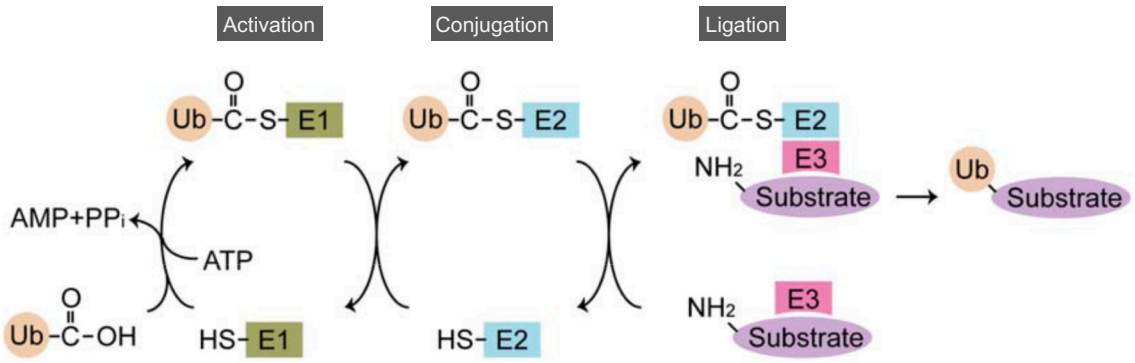

Fig. 7 | Schematic representation of the ubiquitination reaction
Ub is activated at its C-terminal glycine residue in an ATP-dependent manner by E1. ATP-hydrolyzation reaction grants the energy to bind Ub to a specific cysteine residue of E1. Ub is then transferred to an active site cysteine residue of E2, preserving the high energy thioester bond. The substrate is recognized by E3, which also recruits the E2–Ub complex. Finally, Ub is linked by its C-terminus to a lysine residue on the substrate protein[75].

The systems for targeted proteolysis can be categorized into two major categories:

1.3.3.1 Degrons

The systems in this category rely on the fusion of a degron sequence with destabilizing function to the protein of interest (POI)

a. AiD

The AiD system relies on the usage of a hormone of the auxin family to promote ubiquitination and UPS-mediated degradation (Fig. 8 a). In plants, the auxin hormone indole-3-acetic acid (IAA) can bind to the F-box transport inhibitor response 1 (TIR1) to promote its association with the SCF-RING E3 ligase. Whilst animal cells lack the auxin response system, the SCF-mediated degradation pathway is conserved. To adapt this system for targeted degradation, it has been developed the IAA17 degron peptide sequence. The setup of this system requires the integration of two genetic elements in the host genome. One is the DNA sequence for IAA17 must be integrated into the gene that produces the POI in order to generate a fusion protein; the other is the *TIR1* gene since the protein it produces is essential for binding to the IAA hormone. Thus, when the IAA hormone is provided to the cell, all the components are in place to trigger targeted degradation of the POI.

c. Shld/UnaG

Shld and UnaG are two separate systems, both based on destabilizing degrons added to the POI that constitutively promote the degradation of the fusion protein. Those degrons, however, can be stabilized upon the addition of a ligand that stops the degradation (Fig. 8 b). In the Shld system, FKBP12 degron is fused to the POI to constantly mediate its degradation. Upon addition of Shld-1 ligand, the fusion protein is stabilized, and the POI can escape degradation and go back to its function. UnaG is a fluorescent protein from *Unagi* eel that requires bilirubin (BR) as a cofactor to fluoresce. It has been engineered to act as a destabilizing degron in absence of BR, whereas, upon BR administration UnaG is stabilized, therefore escaping degradation.

d. SMASh Tag

SMASh Tag use the hepatitis C virus (HCV) protease non-structural protein 3 (NS3) fused to its cofactor NS4 to act as a destabilizing degron (Fig. 8 c). To confer reversibility to the system is sufficient to link the NS3/4 to the POI through an NS3 cleavage site. In this way, the proteolytic activity of NS3 separates NS3/4 from the POI leading to degradation of the NS3/4 alone. When a specific protease inhibitor like asunaprevir is provided to the system, it blocks the cleavage and the POI fused to NS3/4 can be eliminated.

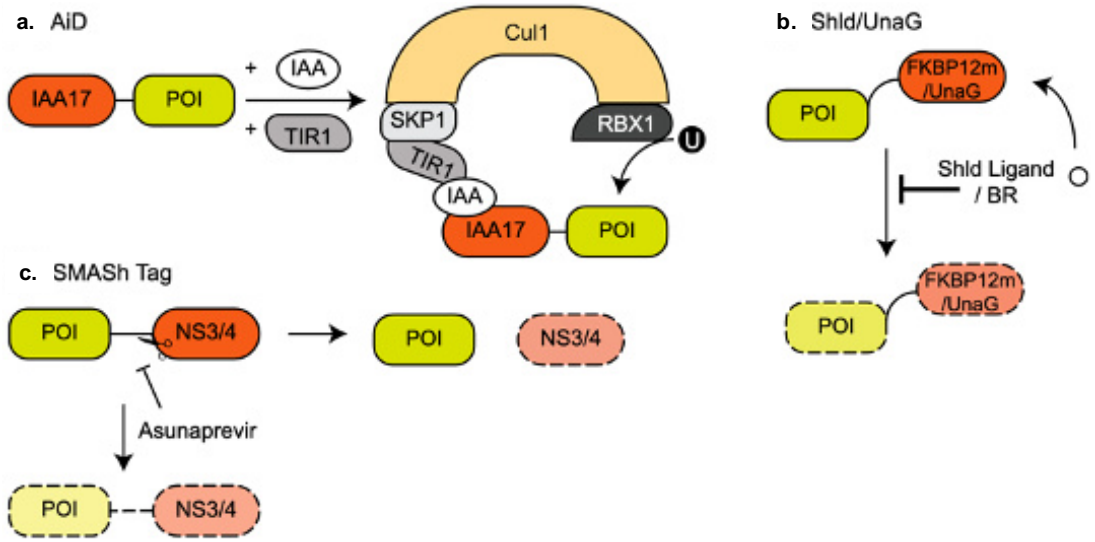

Fig. 8 | Tools for targeted proteolysis: Degrons
a. Auxin-inducible degron (AiD): the POI is fused to an IAA17 degron. Expression of TIR1 and addition of auxin (IAA) results in the recruitment of the IAA17-fused POI to TIR1, which in turn is recruited to the SKP1–Cul1–RBX1 complex, resulting in POI ubiquitination by RBX1. **b.** Shld and UnaG degrons: in the absence of the stabilizing Shld ligand or bilirubin (BR), the POI is degraded when fused to the destabilizing FKBP12 or UnaG mutants, respectively. Addition of the respective compound stabilizes the fusion protein. **c.** SMASh-Tag: the POI is fused to NS3/4 through an NS3 cleavage site. NS3 is cleaved from the POI and degraded, while the native POI is stable (right arrow). Upon addition of the NS3 inhibitor asunaprevir, NS3 cleavage is inhibited and the entire fusion protein is degraded (down arrow)[76].

1.3.3.2 High-affinity binders

In this category antibodies, camelid derived nanobody or small molecules are used to target the POI to an E3 enzyme for ubiquitination.

a. TRIM Away

This system is based on the TRIM21-mediated intracellular antiviral response. TRIM contains a RING-box, which has E3 activity and a C-terminal PRYSPRY domain, which has a high affinity to the constant F_c portion of IgGs. Thus, IgGs that have targeted a

viral particle can be co-endocytosed by the cell triggering TRIM21 activity that recognizes the antibody-virus complex and mediates its ubiquitination and degradation. The system has been adapted as protein degradation tool by expressing the TRIM21 in the target cell, in this context the delivery of the POI-specific antibody can trigger POI ubiquitination and rapid elimination (Fig. 9 a).

b. AdPROM/DeGradFP/ZIF1

Affinity-directed PROtein Missile (AdPROM) together with DeGradFP and ZIF1 are three systems based on a polypeptide binder that drive the degradation of the POI thanks to its affinity for a receptor protein that can recruit the ubiquitin ligase, either directly or with adaptors (Fig. 9 b). AdPROM requires the Von Hippel Lindau (VHL) protein, which serves as a receptor for the CUL2 ubiquitin ligase, that needs to be fused to a polypeptide binder able to recognise the POI. Camelid nanobodies or synthetic binders, known as monobodies, can be used for this purpose. Once the VHL-binder expression is established, the fusion protein can bind the POI and CUL2 through Elongin B and C adaptor proteins leading to degradation of the POI. deGradFP use the same approach with a different receptor (F-box domain of Slmb protein) to bind a different ubiquitin ligase (SCF E3) and uses nanobodies against GFP to target GFP-fused POIs. ZIF1 instead is based on the SOCS-Box protein ZIF, able to recognize and degrade the POI tagged with a ZF1 motif.

c. Hydrophobic tagging

Hydrophobic tagging (HyT) utilize the surveillance system of molecular chaperones Hsp70 and Hsp90 to recruit the co-chaperone E3 ligase CHIP (Fig. 9 c). This degron system works thanks to a small molecule containing a POI-binding moiety and a hydrophobic group that is recognized by the chaperone proteins. Once the binding occurs, the chaperones recruit the E3 ligase CHIP and send the tagged protein to the proteasome.

d. PROTAC/SNIPER

PROteolysis-TArgeting Chimeras (PROTACs), unlike most of the other protein degradation systems, does not require prior genetic modification of the cell. It uses small bivalent molecules to bind simultaneously to the POI and the ubiquitin ligase that mediates the degradation (Fig. 9 d).

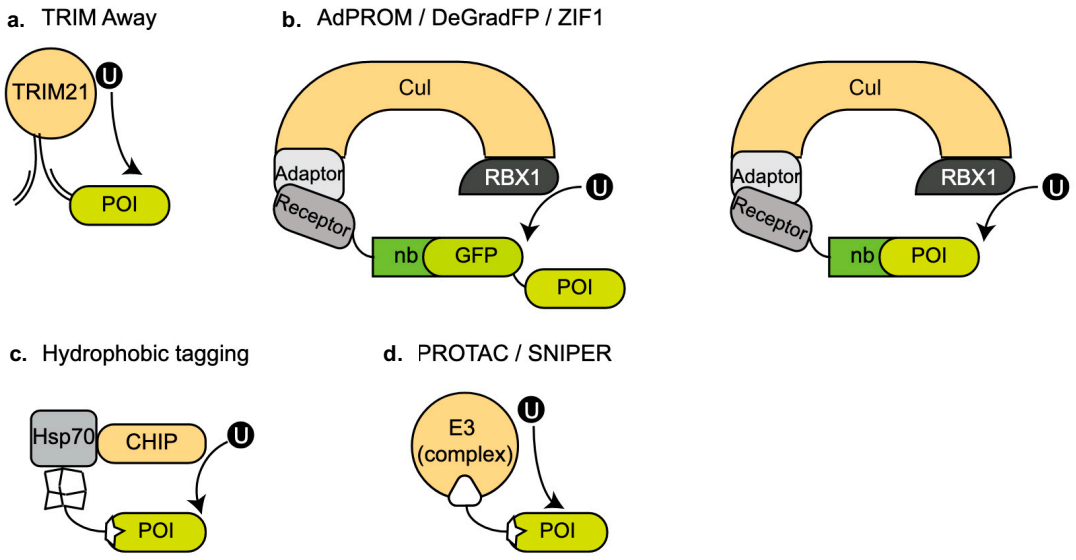

Fig. 9 | Tools for targeted proteolysis: High-affinity binders
a. TRIM-away: the FC region of POI-specific antibodies, once injected into cells, is recognized by TRIM21, while the F_{ab} binds the POI. TRIM21 marks the POI for degradation through ubiquitination. **b.** AdPROM/DeGradFP/ZIF1: a cullin substrate receptor is fused to a target-specific or GFP-specific nanobody and expressed in cells. While the nanobody (nb) recognizes the POI/GFP, the receptor is recruited to the cullin–RBX–adaptor complex. Presentation of the POI to RBX1 results in ubiquitination and degradation of the POI. AdPROM and ZIF1 utilize the CUL2 system with EloB/C as adap-tors, while employing VHL and ZIF1 as substrate receptors, respectively. DeGradFP works with the CUL1 system, using SKP1 as an adaptor, and Slmb F-box as the substrate receptor. **c.** Hydrophobic tagging: a synthetic molecule with a POI-binding moiety binds to the target protein. The fused hydrophobic adamantyl group leads to ubiquitination and degradation of the POI, presumably through the recruitment of the molecular chaperone Hsp70 and its co-chaperone, the E3 ligase CHIP. **d.** PROTAC/SNIPER: a synthetic molecule with two warheads interacts with an E3 ligase on one end and a protein of interest on the other. Spatial proximity allows for ubiquitination of the POI which is subsequently degraded[76].

1.3.3.3 PROTAC evolution

This can be considered as an additional category that lies in between the previous two since it is based on high-affinity binders, but it also implies a genetic modification of the POI. In particular, polypeptide tags are fused to the POI to overcome the main limitation of conventional PROTACs, the inability to target the majority of the proteins and the time-consuming process to develop selective binders against new targets.

a. **HaloPROTACs**
HaloPROTACs use high-affinity binders to link the HALO-POI fusion protein to the VHL receptor and rapidly induce degradation of the HALO-POI fusion protein. The major limitation of this system, however, is the size of the HALO tag itself that can potentially alter the target protein functionality by limiting its interactions. It is also possible for the HALO tag to be targeted for ubiquitination, thus altering protein levels of the basal condition.

b. **dTAG-13**
Nabet and colleagues. recently described a similar system that uses the smaller FKBP12(F36V) peptide (dTAG) as a tag for the high-affinity binder dTAG-13 to link the dTAG-POI to the CRBN receptor of the E3 ligase. dTAG-13 mechanism and features will be described in detail in the next section.

1.3.4 The dTAG system for dynamic protein degradation

As previously mentioned, it was recently presented "*The dTAG system for immediate and target-specific protein degradation*"[28]. The relevance of their work is to provide to the scientific community a widely applicable tool to study the immediate consequence of protein loss. As shown in Fig. 10 FKBP12(F36V) mutant peptide is fused to the target protein to form a chimeric protein. The mutant allowed the selection of a highly specific ligand (dTAG-13) that binds only the mutant and not FKBP12 wt. Upon addition of the heterobifunctional dTAG-13 molecule, it hooks to one end the chimeric protein while the other end contacts the CRBN E3 ligase complex. Once dTAG-13-mediated dimerization occurs, the chimeric protein gets polyubiquitinated by the E3 ligase and marked for UPS-dependent degradation. To develop this platform, the authors screened different candidate degraders by varying regio-isomerisation, linker length and composition of candidate ligands, to identify the most selective and potent molecule, which ended up being the dTAG-13. Moreover, as the authors described, dTAG-13 at the concentration of 100 nM was able to completely degrade the target protein within 1 h upon administration. The system was tested for the degradation of different proteins such as BRD4, KRAS, EZH2, HDAC1, MYC and PLK1 to

confirm its broad application for the targeting of both, nuclear and cytoplasmic proteins. Moreover, dTAG fusion with KRAS, that is a GTPase that participate in the RAS/MAPK signalling pathway, was used to prove the utility of the system to analyse immediate cellular response to degradation. Furthermore, the authors also proved the applicability of the system for *in vivo* studies. An engineered cell line expressing tagged luciferase was transplanted in the mouse model and the systemic administration of the dTAG ligand efficiently induced a loss of the bioluminescent signal. Nevertheless, no evidence about blood-brain-barrier permeability of the dTAG-13 ligand were provided during this study. Finally, *in vitro* and *in vivo* experiments revealed rapid recovery of physiological levels of the tagged protein upon ligand withdrawal, proving the reversibility of the system.

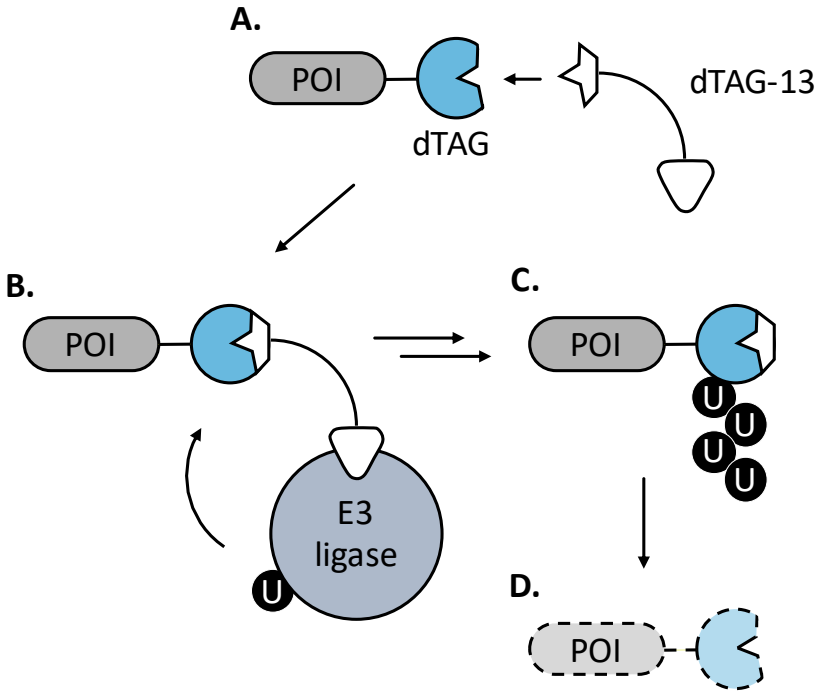

Fig. 10 | Schematic representation of the dTAG degradation mechanism
A. The fusion protein chimera of the target protein and the dTAG bind to the heterobifunctional dTAG-13 ligand. **B.** Dimerization of the protein chimera with the CRBN E3 ligase complex. **C.** The fusion protein is poly-Ubiquitinated (Ub). **D.** proteasome-mediated degradation occurs[28].

1.4 Genome Editing

Genetic engineering techniques play an important role for this thesis work, as they have allowed the generation of embryonic stem lines expressing HTT tagged with the degradation tag that allows to study cellular responses following rapid elimination of the tagged allele. Genome Editing (GE) refers to a compound of molecular technologies that allow tailored modification of genomic DNA. GE is a form of genetic engineering in which DNA is inserted, deleted, modified or replaced in a specific position of the genome of the target cell. Early genetic engineering techniques date back to the 1970s allowing to introducee new genetic elements into living organisms. The major drawback of these procedures was the random nature by which the DNA was inserted into the host's genome, potentially impairing or altering the functionality of other genes. The randomness of these approaches was due to the lack of homology sequences that directed the DNA to a specific region in the host genome. This limitation was overcome with the early gene targeting methods which allowed the targeting to a specific genomic region thanks to the use of homologous sequences as a substrate of the homologous recombination DNA repair pathway. The frequency of

exogenous DNA integration through homologous recombination is extremely low, especially if no double-strand break (DSB) is present in the homology region. Fortunately, the discovery of more and more specific nucleases dramatically increased the efficiency of this process.

Meganucleases:

Meganucleases are considered to be the most specific naturally occurring restriction enzymes. Several hundred of these enzymes have been identified in a large number of organisms, both prokaryotes and eukaryotes. Each meganuclease recognizes a specific sequence (homing site) with a length between 14 and 40 bp, as a result, this site generally occurs only once in any given genome. On one hand, the long recognition sequence ensures very high specificity. On the other hand, it heavily limits the flexibility of the system, hindering the spectrum of targetable genes.

Zinc-finger nucleases:

Zinc fingers (ZFs) are small protein structural motives characterized by the coordination of one or more zinc ions (Zn^{2+}) to stabilize the folding of the protein itself. ZF motives have been identified in many different proteins, consequently named Zinc-finger proteins (ZFPs). Zinc fingers were first identified as a DNA-binding motif in transcription factors, later on they were found to bind also RNA, protein, and lipid substrates. ZFNs for genome editing are chimeric protein deriving from the fusion of different Zinc-Finger motives with the DNA-cleaving domain of the FokI nickase. Several ZF motives exist, each of them can bind to a specific trinucleotide (NNN) present in the DNA. Therefore, ZFN can be engineered by combining different ZF motives to match any target sequence. Although, this has been challenging and time-consuming, nowadays the combination of automated cloning strategies and synthetic biology have made the process easier and more efficient. For DNA editing purposes, two FokI nickases are required on the opposite strands of the target DNA to produce a DSB, thus the combination of two ZFNs to cleave the target sequence grant an increased specificity. Nevertheless, the production of an efficient ZFN may require several rounds of testing and optimization.

TALENs:

Transcription activator-like effectors (TALE) are proteins produced by the bacteria of the *Xanthomonas* genus. These proteins can bind promoter sequences in the host cell and activate the expression of genes that aid bacterial infection. TALEs contain a central repeat domain of ~34 amino acids that mediate DNA binding at the single-nucleotide level. TALENs for gene editing are chimeric protein deriving from the fusion of the FokI nickase with a set of TALE domains that govern the DNA binding specificity. TALEN system provides higher flexibility when compared to the ZFN because facilitates the design of the intended protein based on the specific target site. On top of that, the targeting efficiency of TALENs in most of the cases results higher compared to that of ZFNs. Nevertheless, the effort required to build up a new TALEN for each target sequence is similar to the one faced by ZFPs and it still represents a limiting factor of this system.

CRISPR/Cas9:

The advent of the CRISPR/Cas9 system revolutionised the GE field thanks to its major advantages over all the system described above. The CRISPR system is a two-component system which has always been used by the bacteria as an “adaptive” immune response against invading phages (Jennifer A. Doudna, and Emmanuelle Charpentier, Science 2014). During the first contact between the phage and the bacterial host, a short sequence of the viral genome - the protospacer - is integrated into a specific region of the bacterium genome within a specific cluster. This region contains palindromic sequences separated by spacer sequences therefore it has been named Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR). Each spacer corresponds to

a fragment of each viral genome with which the bacterium came into contact. In response to a second infection from the same viral species, the spacer sequence and the adjacent palindromic repeat are transcribed to produce the CRISPR-RNA (crRNA). This RNA is processed by nucleases and anneals to another RNA molecule, the tracrRNA, in correspondence of the palindromic portion. The tracrRNA is needed for the crRNA binding with the Cas9 nuclease. The crRNA complexed with the tracrRNA is named guide RNA (gRNA). The gRNA binds to the Cas9 nuclease, that in turn can bind to the protospacer sequence of the invading virus thanks to its sequence complementarity. The Cas9 nuclease bound to the gRNA and the viral DNA can generate a DSB. However, the cleavage occurs only if the Cas9 is stabilized by the binding to a trinucleotide sequence situated right downstream of the protospacer sequence. This sequence takes the name of Protospacer Adjacent Motif (PAM). Under these conditions, Cas9 cleaves the DNA 3 bp upstream the PAM, leading to viral genome degradation. This powerful system was adapted to edit the genome of mammalian cells thanks to the production of custom synthetic gRNA, designed to target the sequence of interest in the genome. The CRISPR/Cas9 system allows to easily target new DNA sequence of interest in a time and cost-effective fashion, simple oligonucleotides synthesis and *in vitro* testing. It offers the ability to target almost the totality of the genomic region of interest with the only limitation being the PAM requirement. Moreover, contrarily to the other systems, is suitable for multiplexed targeting by providing multiple gRNA combined with the Cas9 nuclease. All these benefits come at the expense of a slightly higher rate of off-target DNA cleavage that has to be taken into account in each round of editing.

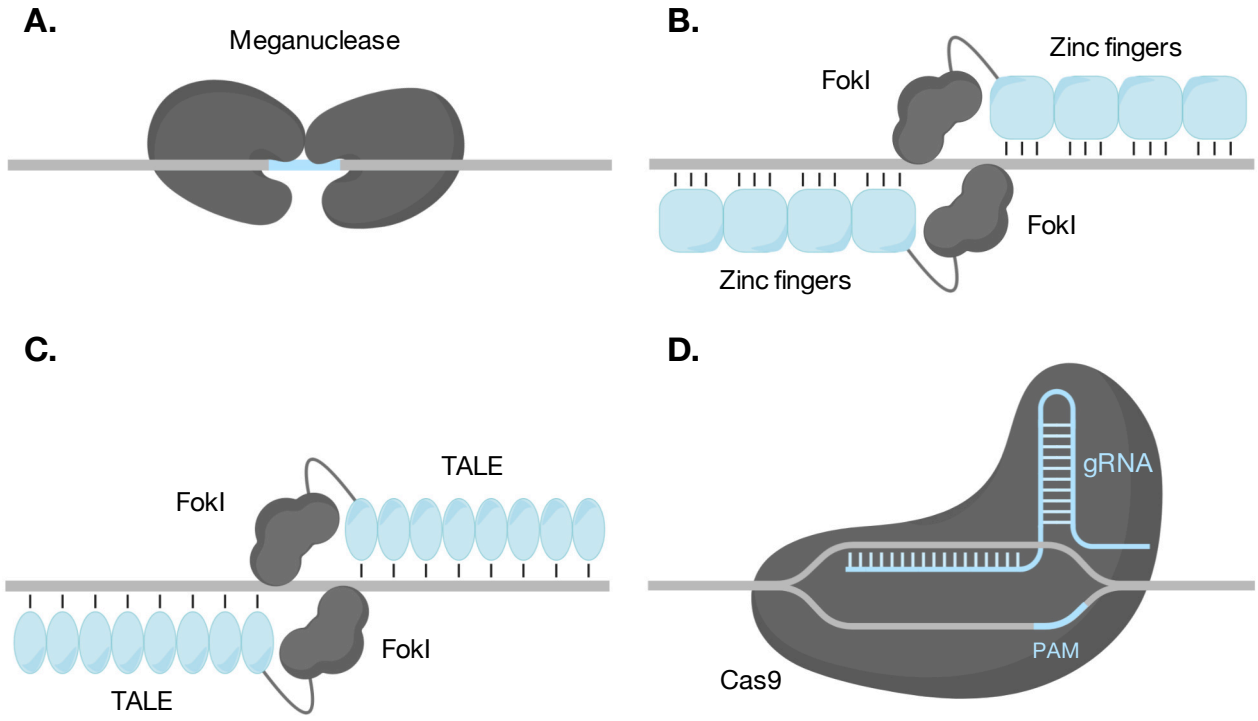

Fig. 11 | Targeted nucleases used for genome engineering approaches
In **A.-D.** the DNA-cleaving enzyme is marked in black whereas the DNA-binding element is marked in light blue. **A.** Meganuclease: characterized by a long homing site which ensure great specificity at the expense of the flexibility. **B.** ZFNs: composed of different ZN motives and a FokI nuclease display a high specificity and decent flexibility **C.** TALENs: composed of several TALE domains linked to the FokI nuclease. Respect to the ZFNs have higher specificity, higher efficiency and better flexibility. **D.** CRISPR/Cas9: the DNA-binding element is an RNA molecule that can be programmed to achieve incredible flexibility, the only limitations are the specificity and the PAM requirements.

In conclusion, in the introduction, I have presented the major technologies available at the moment to introduce genetic tools for functional studies. In this thesis, I have applied these technologies to investigate the immediate transcriptional response to HTT removal.

2 | Results and Discussion

2.1 dTAGging of HTT locus in an HD model of human embryonic stem cells for acute HTT degradation

To investigate the transcriptional response to acute HTT degradation, we decided to leverage a genetic strategy based on the insertion of a degradation tag (dTAG) in the HTT locus through Cas9-assisted gene editing. The dTAG system allows us to rapidly trigger the degradation of HTT upon the administration of a cell-permeable molecule in the culture medium.

2.1.1 Description of the editing procedure

The parental cell line we started from was the RUES2-58Q clone 25.2[77], a cell line we received from Ali Brivanlou's lab at Rockefeller University. This cell line has been edited with the CRISPR/Cas9 system to insert an expanded CAG tract in one of the two endogenous HTT alleles (heterozygosis). The resulting genotype of the HTT locus is 24Q / 58Q (Fig. 12). Brivanlou's lab also provided us with a list of the gRNA they screened for the editing on the HTT locus. Among them, we selected the most convenient for our editing strategy. We nucleofected these cells with the selected gRNA conjugated with recombinant Cas9 to form a ribonucleoprotein (RNP) complex and the donor template plasmid to integrate the dTAG cassette into either of the endogenous alleles. To do this, we used two alternative donor plasmids (Fig. 12) carrying normal (dTAG21Q donor) or expanded allele (dTAG56Q donor). We used different poly-Q length compare to the endogenous alleles to be able to discriminate by PCR the endogenous sequences from the synthetic alleles. This is essential to distinguish between "pure" clones, therefore deriving from a single cell or mixed population that may show the morphology typical of clonal colonies but are in reality mosaics. This step is crucial, since, in our experience, hESC have the tendency to stick one to another making it difficult to obtain colonies derived from a single cell (pure clones). Post-nucleofection we plated the cells at low density in 6 cm Petri dishes. As soon as the cells recovered and formed small colonies, we applied the antibiotic selection for 5 consecutive days. After the selection period, we were confident that all the cell that didn't uptake the donor would have died. Thus, we proceeded with the isolation of the antibiotic-resistant clones in a 96-well plate. Once each clone filled the well, we transferred it into the 24-well format and subjected it to a second round of antibiotic treatment to rule out the possibility of carrying on transiently resistant clones. Finally, we ended up with few positive clones that were expanded and used to generate a frozen cell stock, after the collection of a pellet for the PCR screening of the correct integration of the dTAG cassette.

2.1.2 Screening of positive clones

The collected pellets were used for genomic DNA extraction and purification. We designed PCR reactions to discriminate clones that correctly integrated the dTAG cassette in the HTT locus from clones that originated from random integration of the donor plasmid. To verify the correct dTAG integration we designed two independent PCR reaction, one amplifying the genomic region spanning the right-homology arm (R-HA) junction and one amplifying the region spanning the left-homology arm (L-HA) junction. Graphical representation of these PCRs is visible in (Fig. 13 A). We first used the PCR across the R-HA as an initial screening, obtaining four positive clones out of 24 tested for the dTAG21Q integration (Fig. 13 B) and two clones out of 32 tested for dTAG56Q integration (Fig. 13 C). PCR products were purified, and Sanger sequenced in both directions to verify the sequence (Fig. 14). The samples positive for this PCR were then used for the PCR across

the L-HA for further validation of the editing (Fig. 13 F). Sequencing data again confirmed the expected sequence (Fig. 14).

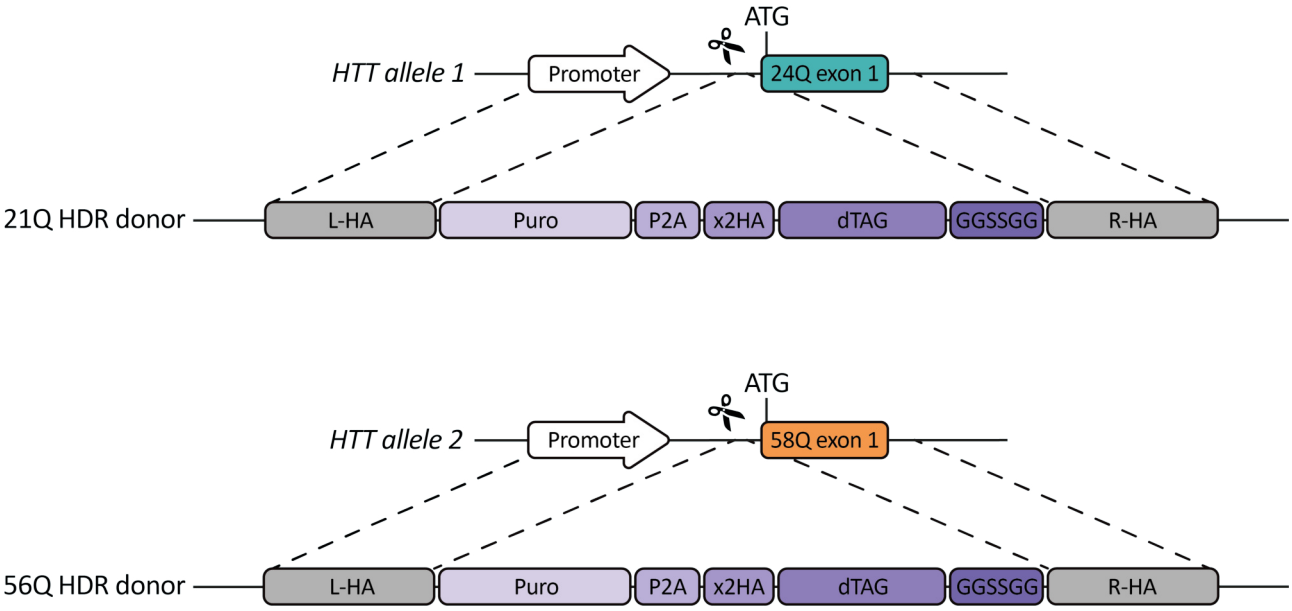

Fig. 12 | Graphical illustration of the homology mediated knock-in of the dTAG cassette in the 24Q or 58Q HTT allele.
The two HTT alleles present in the parental line are depicted as HTT allele 1 and 2. The 21Q donor was used to knock-in the dTAG cassette selectively in the HTT allele 1, whereas the 56Q donor was used to knock-in the dTAG cassette selectively in the HTT allele 2. The dTAG cassette contains a promoter-less Puromycin resistance gene for antibiotic selection of the cells that have integrated the cassette. The P2A viral peptide allow to separate the Puro from the downstream aminoacidic sequence due to the failure in the formation of a peptide bond. Downstream, fused to HTT there are two copy of the HA epitope for efficient antibiotic detection, dTAG protein and a GGSSGG flexible linker.

2.1.3 Quality Controls

The clones selected form the screening were subjected to multiple quality controls (QCs). First, we performed a third PCR to amplify the region across the gRNA binding site that includes the CAG repeats tract (Fig. 13 A). Thanks to the pre-designed diversity in CAG length of the donor templates (21Q and 56Q) relative to both the endogenous alleles of the RUES2 lines used (24Q and 58Q), this PCR allows us to exclude mosaics derived from a mixed population of cells from pure clones. In our experience, we came to realize that this event is not so uncommon, especially in human ESC. Most likely, this event takes place at the moment of cell plating post-nucleofection. Since the hESC have the tendency to stick one to another, the hypothesis is that one cell that receives all the editing components may stick to a cell that doesn't. If the second cell receives nothing or the RNP complex alone, it is likely that it would die due to the antibiotic selection. Conversely, if the second cell receives the donor plasmid alone (no RNP) and integrates it randomly, it may become antibiotic-resistant and thus forming a colony with the first cell carrying the correct editing. This colony, although made of two different starting cells, can be easily mistaken as a colony coming from a single clone. In the case of a pure clone, this PCR should produce a band corresponding to the unedited allele corresponding to the 56Q (405 bps) for the dTAG21Q clones and the 24Q (306 bps) for the dTAG56Q clones. Whereas, in the case of a mixed colony the PCR should produce a doublet of bands, corresponding to the two endogenous alleles (24Q / 58Q). On the basis of this PCR, we excluded dTAG21Q c26, since it showed a duplet of bands similar to the unedited clones, and dTAG21Q c39 since it displayed no amplification (Fig. 13 D). Thus, we selected clones 30 and 33

for the dTAG21Q line and clones 01 and 27 for the dTAG56Q line since the band indicating mosaicism was absent in these clones (Fig. 13 D-E).

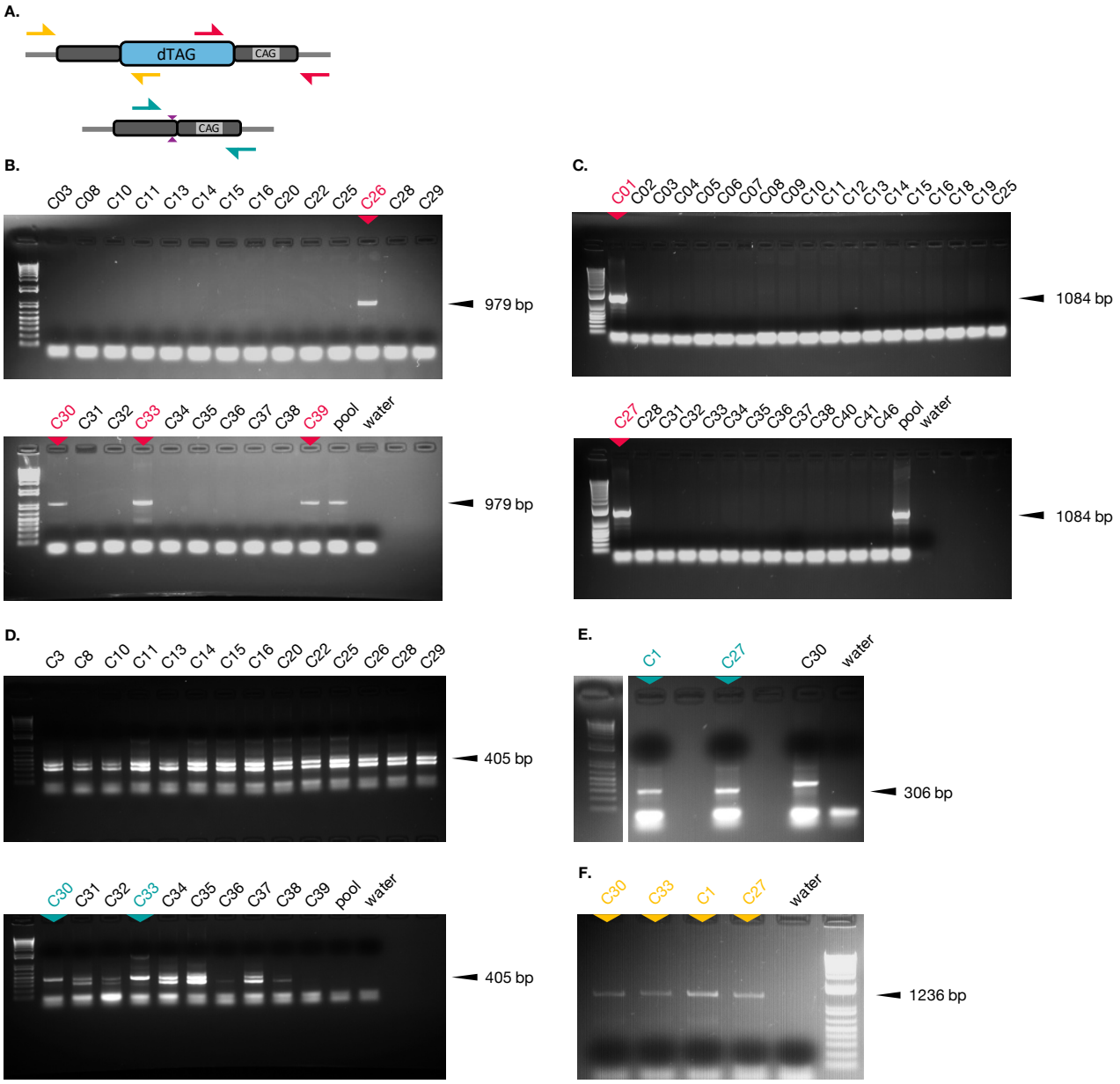

Fig. 13 | PCR screening of dTAG-positive clones.
A. Graphical representation of PCRs designed for dTAG-positive clones screening. **B-C.** Electrophoretic separation of the PCR spanning the R-HA for the B. dTAG21Q and C. dTAG56Q clones. This PCR was used for the initial screening of the dTAG-positive clones. **D-E.** Electrophoretic separation of the PCR amplifying the untargeted HTT allele 2 for the dTAG21Q and allele 1 for the dTAG56Q. This PCR allow the exclusion of potential mosaics derived from a mixed cell population and can detect the presence of InDels eventually generated by the DNA repair mechanism triggered by Cas9 cleavage. **F.** Electrophoretic separation of the PCR spanning the L-HA for the dTAG21Q clones on the left and dTAG56Q on the right.

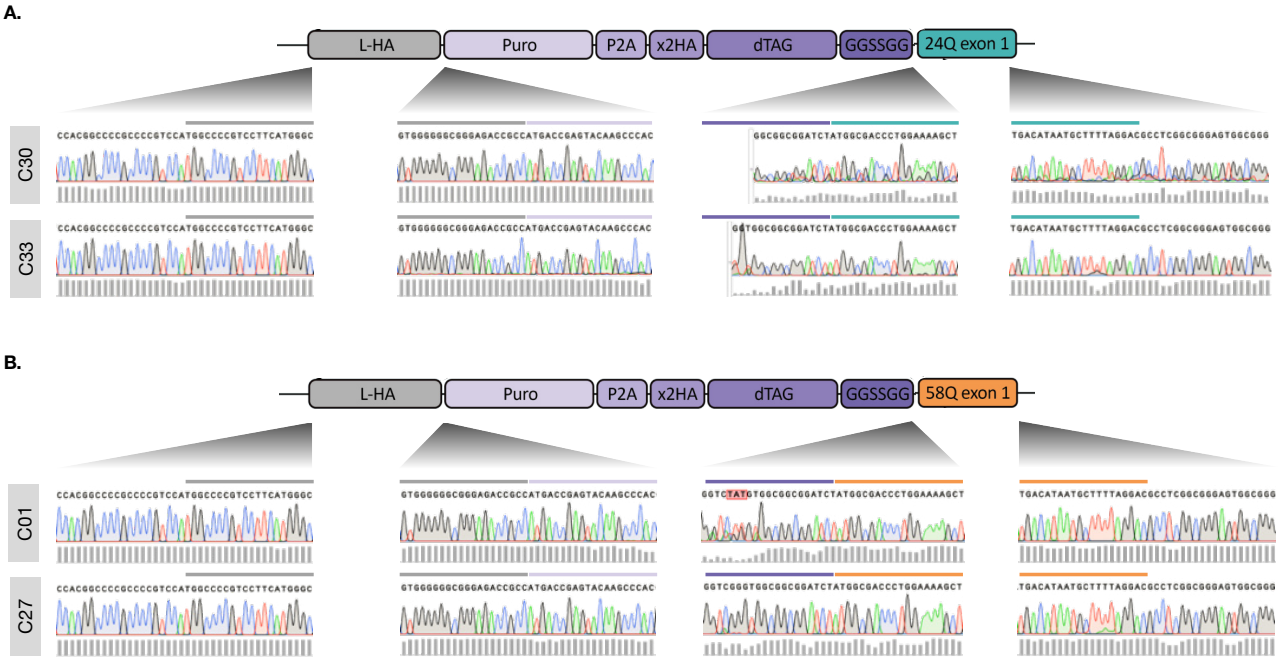

Fig. 14 | Sanger sequencing of PCR products.
A. Chromatogram and base score relative to the sequencing of PCR across the L-HA (left) and R-HA (right) for dTAG21Q c30 and c33. No mismatches present. **B.** Chromatogram and base score relative to the sequencing of PCR across the L-HA (left) and R-HA (right) for dTAG56Q c01 and c27. No mismatches present.

In addition, this same PCR can be used to evaluate the presence of insertion/deletions (InDels) in the untargeted allele that can occur as the result of RNP-mediated double-strand break in absence of homologous recombination. To do this, we used the TIDE (Tracking of InDels by DEcomposition) web tool that is able to decompose Sanger chromatograms to give an estimation of the amount of insertion or deletion generated in the sequence by the editing process. TIDE software aligns the gRNA recognition sequence to the edited sample and to an unedited control. Subsequently, the software attempt to decompose the chromatogram of the edited sample based on the information provided by the unedited sequence and the expected cleavage site. The algorithm returns an estimation of the relative abundance and number of insertions or deletions in the edited sample compared to the unedited control. The software also provides the statistical significance of each insertion or deletion identified. We tested dTAG21Q c30 and 33 together with dTAG56Q c1 and 27, comparing the sequencing data to the unedited parental line RUES2-58Q clone 25.2. As shown in Fig. 15 all of the tested clones displayed no significant InDels in the proximity of the Cas9 cleavage site, suggesting that the editing procedure didn't alter the sequence in the wt unedited allele.

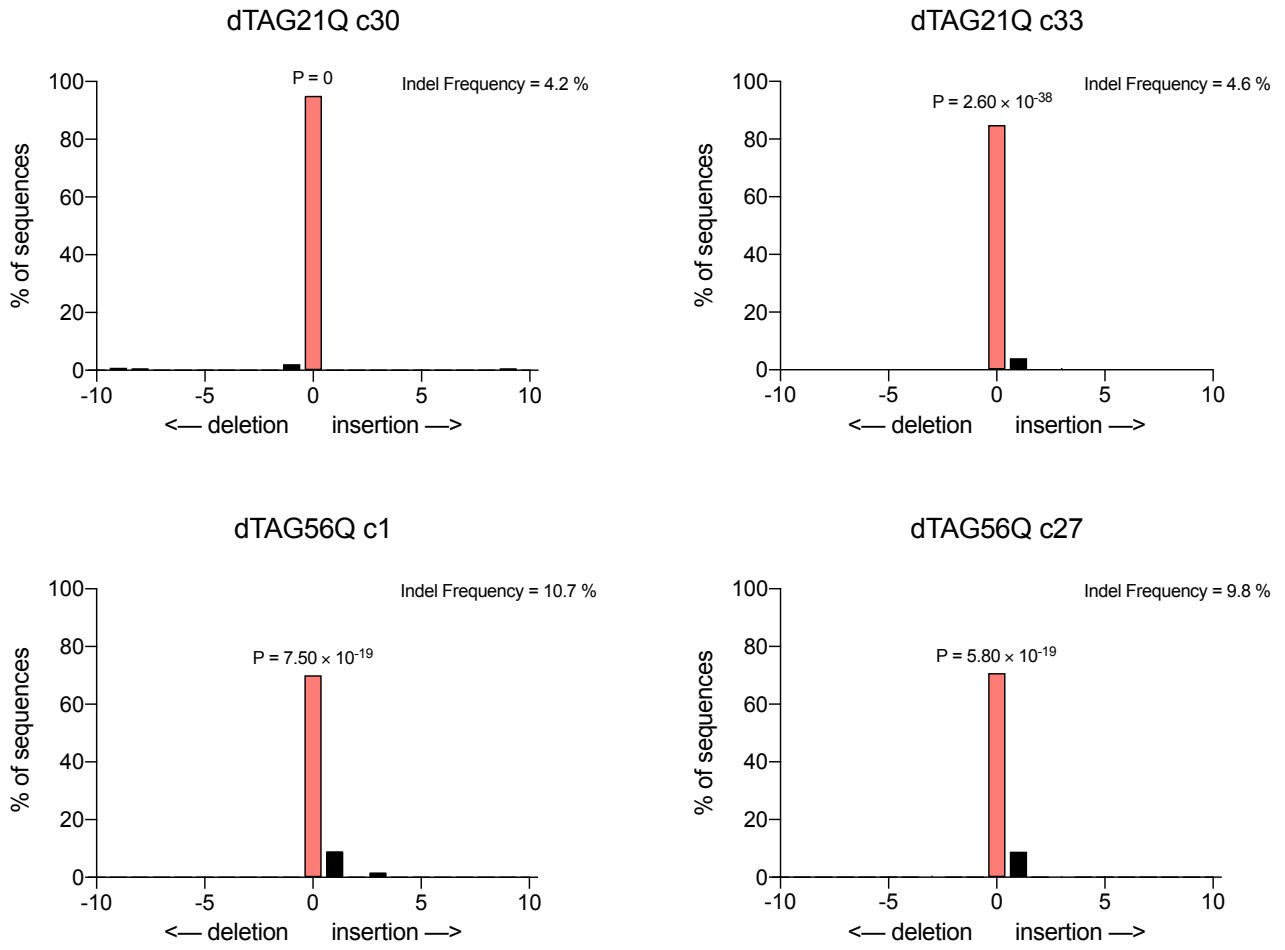

Fig. 15 | Tracking of InDels by Decomposition (TIDE) analysis.
Decomposition plots resulting from TIDE analysis. These plots display the percentage of sequence that matches the control, corresponding to the histogram at x = 0 and the percentage of sequence that matches the control considering 1 to 9 bases removed (deletion range) or inserted (insertion range). A P value is associated to each bar of the histogram and the bar is colored in pink if the P value is significative.

Second, we analysed the karyotype of the four selected clones to address the presence of some of the most common chromosomal abnormalities detected in cell culture. We used the Q-banding service provided by ISENET Biobanking to investigate if important chromosomal alterations such as polyploidy, aneuploidy, translocations, inversions, rings, and copy number changes were present in our clones. Fortunately, all of the clones displayed a normal karyotype, not presenting any of these alterations (Fig. 16). Finally, we performed real-time qPCR to check the expression of the major pluripotency genes and the expression level of the *HTT* gene in the edited clones compared to the unedited parental line. We analysed five different genes involved the transcriptional regulation of pluripotent stem cells, such as OCT4, SOX2, FUT4, KLF4 and KLF5. In addition, we also analysed the mRNA level of HTT to check if the editing procedure altered HTT transcription. From these analyses, we detected no significant difference in the expression of the five pluripotency marker genes between the dTAG lines and the unedited parental line (Fig. 17). We also detected no alteration in the amount of HTT mRNA in the dTAG lines compared to the unedited line. Overall, these data indicate that the editing procedure didn't affect the pluripotency of the generated dTAG lines and that the dTAG cassette didn't alter HTT transcription.

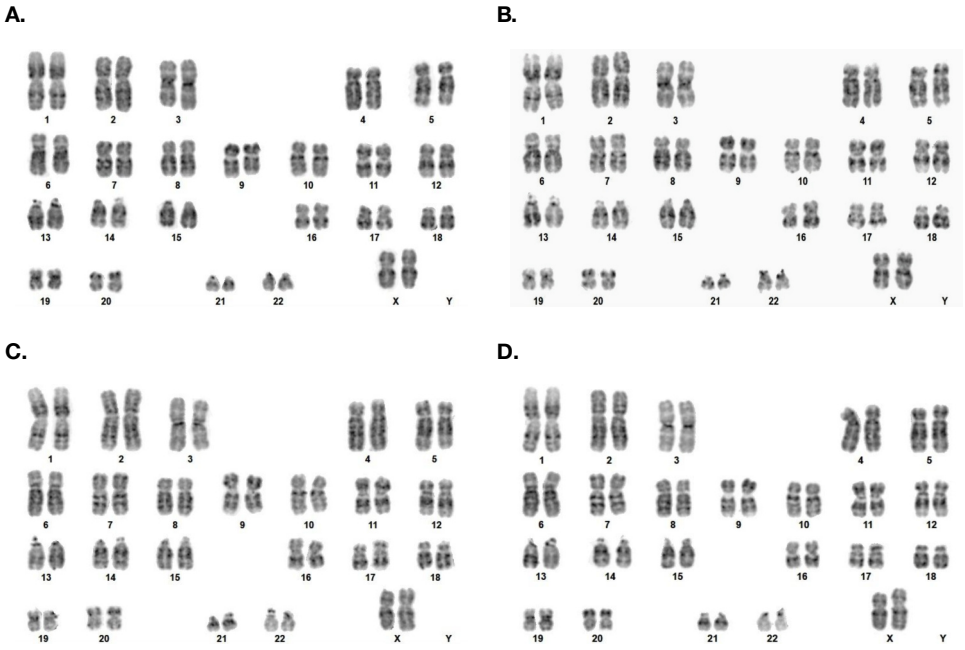

Fig. 16 | Q-banding analysis performed on the dTAG lines.
Images of the quinacrine stained chromosomes for dTAG21Q clones **A.** 30 **B.** 33 and dTAG56Q clones **C.** 01 and **D.** 27. All the clones present a normal female karyotype.

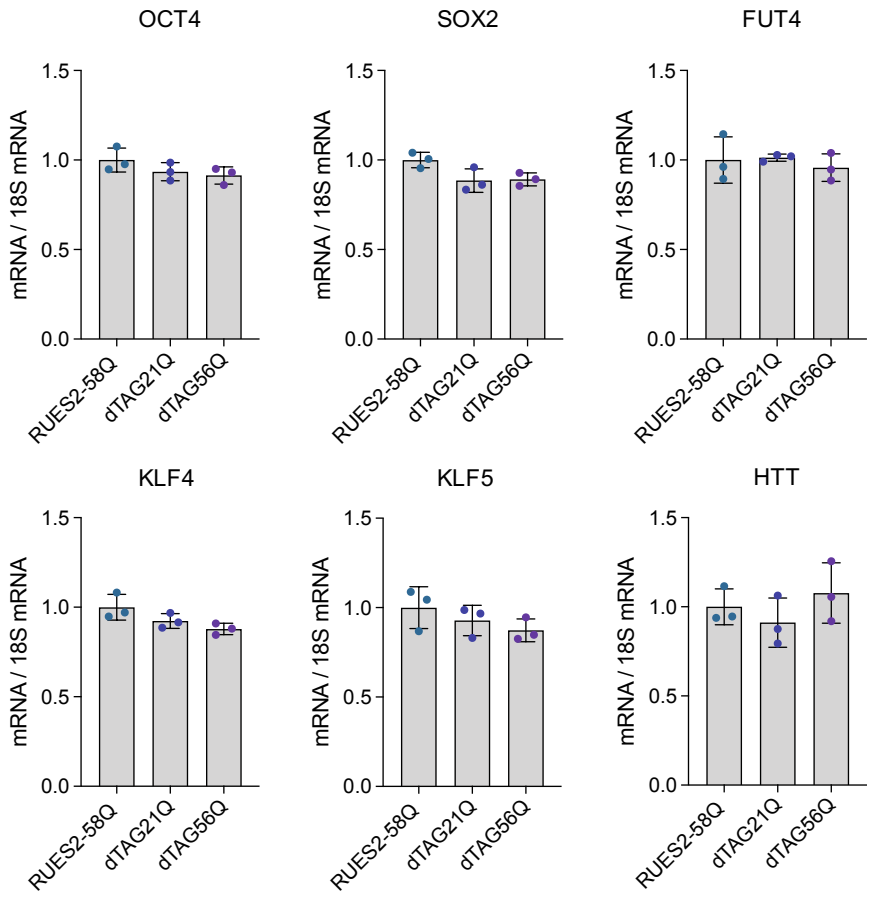

Fig. 17 | Gene expression analysis of the major pluripotency markers and HTT.
Relative mRNA abundance of the indicated genes. one-way ANOVA; n = 3 biological experiments, data are represented as mean ± SD.

2.1.4 System validation

Since the selected clones passed all the quality controls, we moved to the functional validation of the dTAG system. To do so, we treated the dTAG21Q c30 that we had in culture with two different λ concentration [1 μM] or [0.1 μM] for 1, 5, or 25 hours and collected samples to measure HTT protein amount by WB. As a control, the same cells were treated with an equivalent volume of DMSO which is the solvent of the λ molecule. As an additional control, we also treated the unedited parental line RUES2-58Q c25.2 with the highest concentration of λ [1 μM]. Western blot analysis performed on these samples confirmed the functionality of the dTAG system to induce complete degradation of the tagged HTT allele already after 1 h of treatment (Fig. 18). In fact, when looking at the amount of Total HTT we observed a ~ 50% reduction in the treated samples compared to the t0 (Fig. 18 E-F.). This effect is observed at both the tested concentrations and is induced by the λ molecule since the DMSO treatment caused no reduction of total HTT (Fig. 18 G). Similarly, the λ [1 μM] treatment on the RUES2-58Q c25.2 induced no significant alteration of total HTT levels (Fig. 18 H). By looking at the HA-tagged HTT the functionality of the system is even more evident since this detection is restricted to the dTAG allele only (Fig. 18 J-K.). Indeed, for both the concentration, HA-tagged HTT signal is completely shut down between 1 and 5 hours of λ treatment.

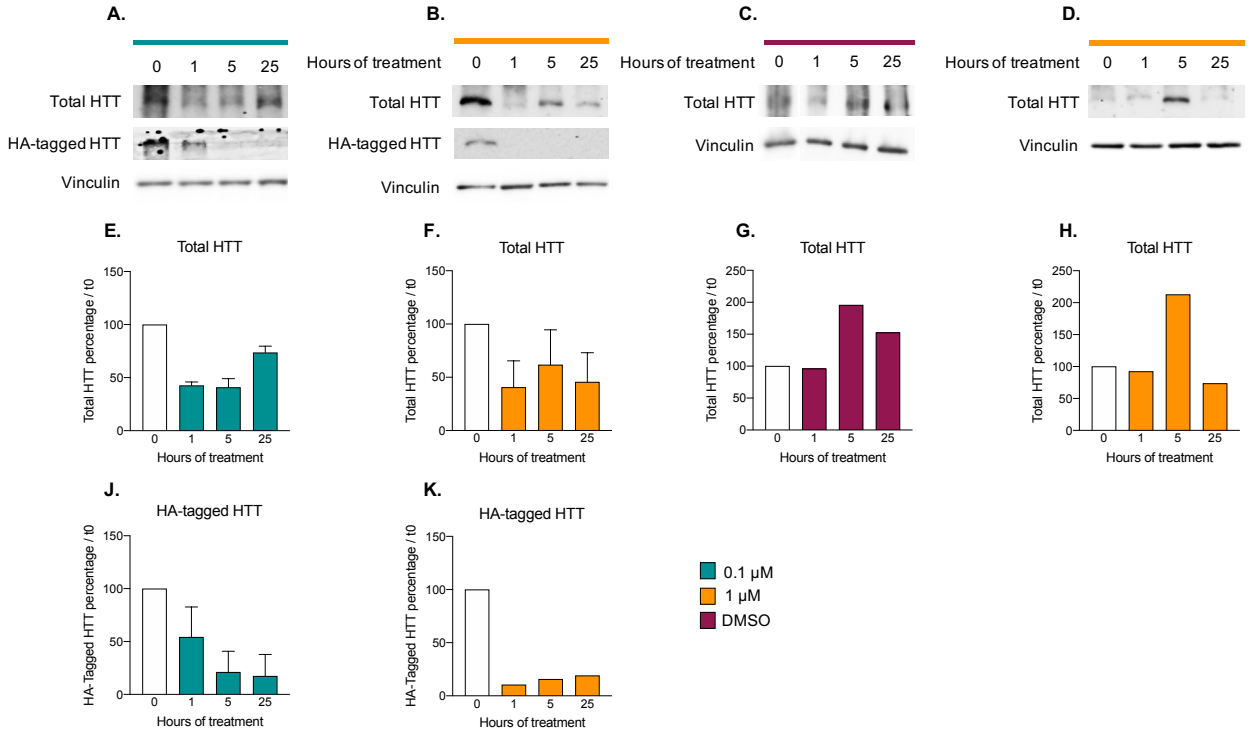

Fig. 18 | Validation of the dTAG-mediated HTT degradation by western blot.
A-D. Western blot images of dTAG21Q c30 treated with λ **A.** 0.1 μM or **B.** 1 μM and RUES2-58Q c25.2 parental line treated with **C.** DMSO or **D.** λ 1 μM. Top row immunoblot of Total HTT (anti-HTT antibody), middle row immunoblot of HA-dTAG-HTT (anti-HA antibody) and bottom row immunoblot of Vinculin (anti-Vinculin antibody). **E-J.** Densitometric quantifications of **E.** total and **J.** dTAG-HTT of dTAG21Q c30 treated with λ 0.1 μM. **F-K.** Densitometric quantifications of **F.** total and **K.** dTAG-HTT of dTAG21Q c30 treated with λ 1 μM. **G-H.** Densitometric quantifications of total HTT in RUES2-58Q c25.2 parental line treated with **G.** DMSO **J.** λ 1 μM. n = 1 biological experiments, two technical replicates, data are represented as mean ± SD.

We then set to confirm this data by immunocytochemistry (ICC) on all the four clones we obtained from the screening. Fist, to set up the total and tagged HTT staining acquisition and quantification, we performed ICC on the RUES2-58Q c25.2 parental line and the dTAG21Q c30, also including

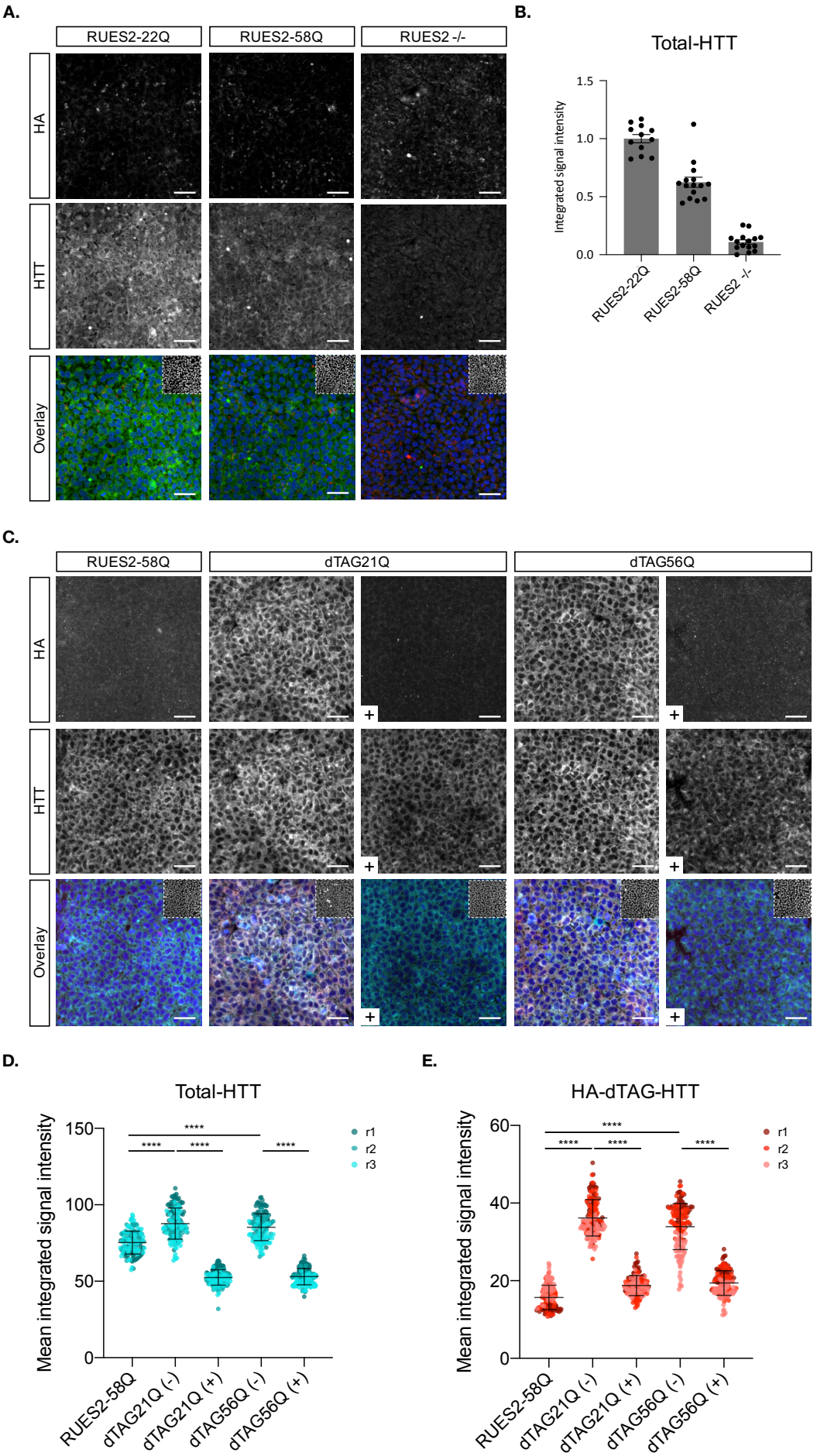
A.
RUES2-22Q
RUES2-58Q
RUES2 -/-
HA
HTT
Overlay
B.
Total-HTT
Integrated signal intensity
1.5
1.0
0.5
0.0
RUES2-22Q
RUES2-58Q
RUES2 -/-
C.
RUES2-58Q
dTAG21Q
dTAG56Q
HA
HTT
Overlay
+
D.
Total-HTT
Mean integrated signal intensity
150
100
50
0
****
****
****
****
r1
r2
r3
RUES2-58Q
dTAG21Q (-)
dTAG21Q (+)
dTAG56Q (-)
dTAG56Q (+)
E.
HA-dTAG-HTT
Mean integrated signal intensity
60
40
20
0
****
****
****
****
r1
r2
r3
RUES2-58Q
dTAG21Q (-)
dTAG21Q (+)
dTAG56Q (-)
dTAG56Q (+)

Fig. 19 | Validation of the dTAG-mediated HTT degradation by immunofluorescence
A. Immunofluorescence images HA (red) and HTT (green) in RUES2-22Q, 58Q and -/-, Scale bars: 50 µm **B.** Integrated intensity quantification of HTT fluorescent signal performed with CellProfiler. n = 1 biological experiment; 15 random fields, data are represented as mean ± SD. **C.** Immunofluorescence images HA (red) and HTT (cyan) in RUES2-58Q (parental) treated with DMSO for 24 h, dTAG21Q and dTAG56Q treated with DMSO or λ (+) for 48 h, Scale bars: 50 µm. **D-E.** Integrated intensity quantification of **D.** HTT and **E.** HA (dTAG-HTT) fluorescent signal performed with CellProfiler. **** P < 0.0001 two-way ANOVA, Tukey's multiple comparisons test; n = 3 biological experiments (r1, r2, r3); 50 random fields, data are represented as mean ± SD.

RUES2-22Q c30 and RUES2 -/-. With this experiment, we verified the specificity of the anti-HA antibody as we detected signal only in the dTAG line and no signal in all the other conditions (Fig 19 A). Moreover, as can be seen in Fig. 19 A-B we observed that the anti-HTT antibody is suitable to detect a small difference in HTT protein amount as we were able to detect a lower HTT signal in the HD cell line (RUES2-58Q c25.2) respect to the wt line (RUES2-22Q c30). This may be an indication of a hypomorphic phenotype of the 58Q allele respect to the 22Q allele causing a weaker signal when staining the cells for total HTT. However, to confirm this hypothesis, further investigation is required. Having seen the good sensitivity of HTT/HA quantification by ICC, we set up an assay to measure HTT/HA degradation. We treated both dTAG21Q and dTAG56Q lines for 24 h with either λ [10^{-7} M] or with the equivalent volume of DMSO as control. In parallel, we also treated the parental unedited line (RUES2-58Q c25.2) with DMSO as reference for HTT and HA signal intensity (Fig 19 C). As expected, we observed a ~ 50% reduction of total HTT signal when comparing the λ treatment with the DMSO treatment for both dTAG lines (Fig. 19 D). Second, the reduction is instead ~ 100% when considering the HA signal only, derived exclusively from the tagged allele (Fig. 19 E). Therefore, this experiment confirmed once again the ability of this system to trigger the complete degradation of the target protein upon λ exposure.

2.2 Identification of the lowest λ dose for rapid and complete HTT degradation

Once we proved the functionality of the dTAG system to induce HTT degradation, we aimed at determining the lowest λ concentration to achieve fast and complete degradation of tagged HTT. To do so, we needed to derive the λ dose-response curve in our cell lines. We could not use the data described by Nabet and colleagues[28] since the dose-response curve is dependent on the degradation rate of the tagged protein and, likely, different proteins will have a different pace of degradation. Therefore, we treated our dTAG lines for 1 h with serial dilutions of λ, starting from 10^{-6} M down to 10^{-10} M with 10-fold dilution steps. During these experiments, for both lines, we treated independently each clone. In this way, we could establish if one of the two clones could have different responsiveness to λ. Western blot analysis of the collected samples show a dose-dependent lowering of both total HTT and tagged-HTT signals, as expected (Fig. 20 A). It is noteworthy that the dTAG56Q line shows a doublet of bands differently form the 21Q line. This is the result of the molecular weight difference between wt and mut HTT that in this line accounts for about 5 kDa. Moreover, the dTAG56Q line carries an additional 14 kDa due to the tags (2×HA-dTAG) on the mutant allele, bringing the difference to 19 kDa that is detectable in a 5% acrylamide gel. Conversely, in the dTAG21Q line, the wt HTT with the addition of 14 kDa becomes heavier than the mutant allele with a difference of 9 kDa. This difference does not allow to resolve the two bands with conventional acrylamide gels. We then quantified tagged-HTT signal to derive the dose-response curve for each clone. As a result of this analysis, we noticed no significant difference between the two clones of each line, therefore, we decided to treat them as distinct biological replicates and derive the dose-response curve considering the average of the two clones. As shown in Fig. 20 B, the two lines present very similar curves with a calculated IC_{50} of $1{,}92 \times 10^{-8}$ M and

1,41 × 10^{-8} M for dTAG21Q and dTAG56Q lines, respectively. Most importantly, we estimated the lowest dose to achieve approximately complete degradation in 1 h to be 1 × 10^{-7} M. All subsequent experiments were conducted using this dose.

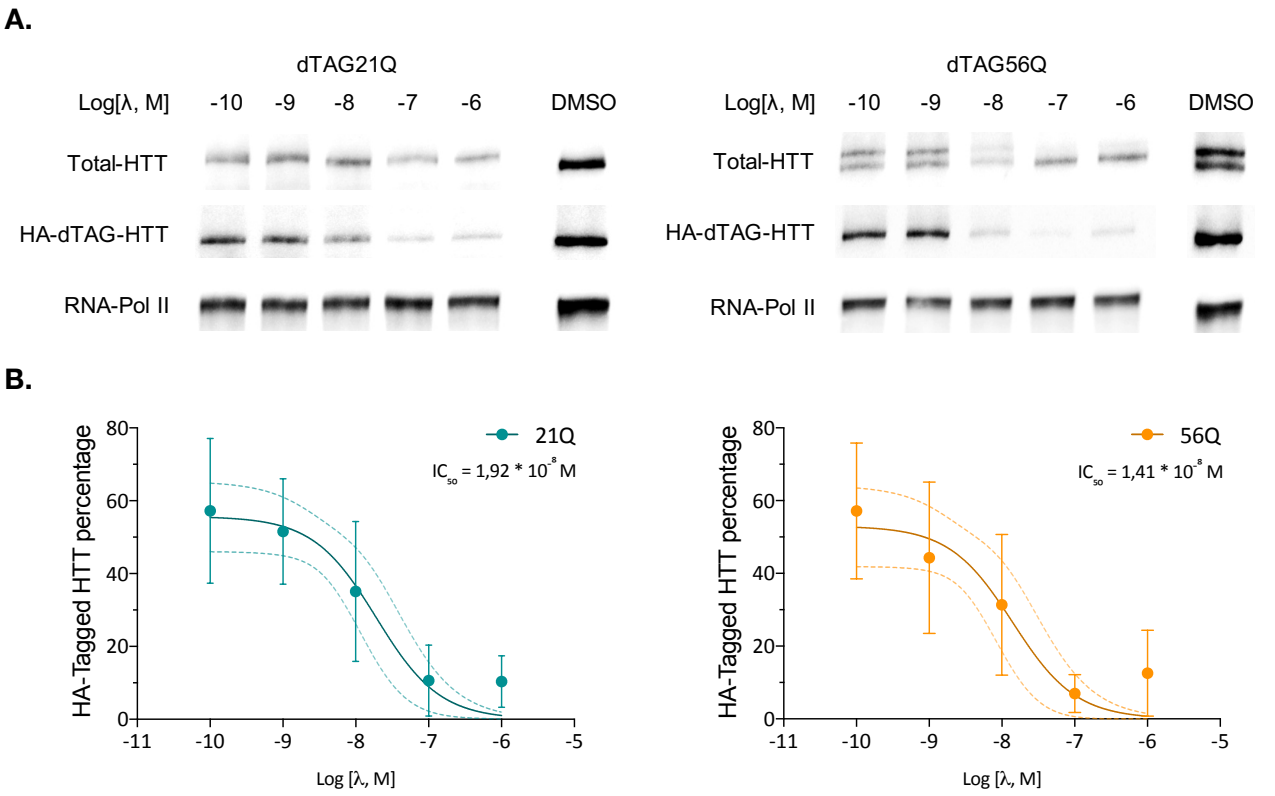

Fig. 20 | Dose-dependent dTAG-HTT degradation.
A. Western blot images of dTAG21Q (left) and dTAG56Q (right) treated with 10-fold increasing concentration of λ for 1 h **B.** Best-fitting dose-response curves calculated on densitometric quantifications data. n = 3 biological experiments, data are represented as mean of two independent clones ± SD (n = 6 data point analyzed for each concentration).

2.3 Protein-degradation kinetic of normal- and mutant-HTT

To analyze the transcriptional response to normal- or mutant-HTT degradation in a time-dependent manner, we next wanted to retrieve precise information about degradation dynamics in our dTAG lines. Thus, we treated our dTAG lines for 5 h with λ [1 × 10^{-7} M], collecting samples every 30 minutes in the first 3 h and then 4 and 5 h after treatment. We treated independently each clone for both lines, again to verify whether differences in the degradation dynamics may exist. Western blot analysis of these samples revealed a very rapid degradation that causes the totality of the tagged HTT to be eliminated between 1.5 and 2 h of treatment (Fig 21). The same result can be deducted from the quantification of the tagged HTT. Furthermore, the data points obtained from the quantifications allowed us to fit the degradation curves and calculate the kinetics of degradation. These curves follow a non-linear one-phase exponential decay model. The fitted curves allow us to derive various parameters related to HTT degradation dynamics. These parameters are reported in Tab. 22. As can be seen from the left graph in figure 21 B, considering only the tagged HTT, the fitted curves calculated from the experimental data are superimposable. This indicates that there are no significant differences in terms of degradation rate between the wt and mutant HTT. Observing instead the graph on the right in fig. 21 B, we can appreciate significant differences. Indeed, the two fitted curves seem to diverge one from the other, in terms of slope and of plateau point. However, the slope should be the same as we have already established that there is no

difference in terms of degradation rate between dTAG21Q HTT and dTAG56Q HTT. In fact, by performing the extra sum-of-squares F test the difference result to be not significant. The main divergence, therefore, appears to be at a steady-state, which could explain why the slope of the two curves appears different. In particular, the steady-state of the dTAG21Q line is around 30% of the initial protein amount and the F test confirmed that is significantly different from the expected value of 50% considering the depletion of a single allele. Since the steady-state of the dTAG21Q line represents the percentage of mutant HTT left after degradation of the wt HTT, a value lower than 50% could suggest the presence of a hypomorphic phenotype of the mutant allele with respect to the tagged wt allele. If this were true, however, the opposite phenomenon should be observed in the dTAG56Q line where we should expect to have a steady-state value of around 70% of the initial amount, assuming that the mutant allele actually contributes 30% of the total HTT. However, this does not seem to be the case indicating that there is some other phenomenon affecting these observations. Moreover, as shown in Fig. 21 B, the quantification of total HTT after λ treatment of the dTAG21Q line returns a reduction of 50% and is therefore in contrast with what was observed by western blot. This discrepancy may be the result of the different nature on these two assays, even though they both rely on the same antibody. Indeed, ICC is a technology that allows for the detection of the totality of the target, including both the soluble and insoluble fraction. Conversely, WB mainly allows the quantification of the relative abundance of the target within the soluble fraction. Since is difficult to establish the relative proportion of these two components, the contradiction between these two means of quantifications might be explained by the presence of a variable fraction of insoluble protein between normal and expanded HTT. If this were the case, a 50% reduction would be expected by ICC, as the antibody would also bind to the insoluble fraction. Whereas, in the western blot a portion of the insoluble fraction may fail to migrate into the acrylamide gel, resulting in a lower amount of mutant HTT after degradation. To rule out this hypothesis, additional experiments would be required.

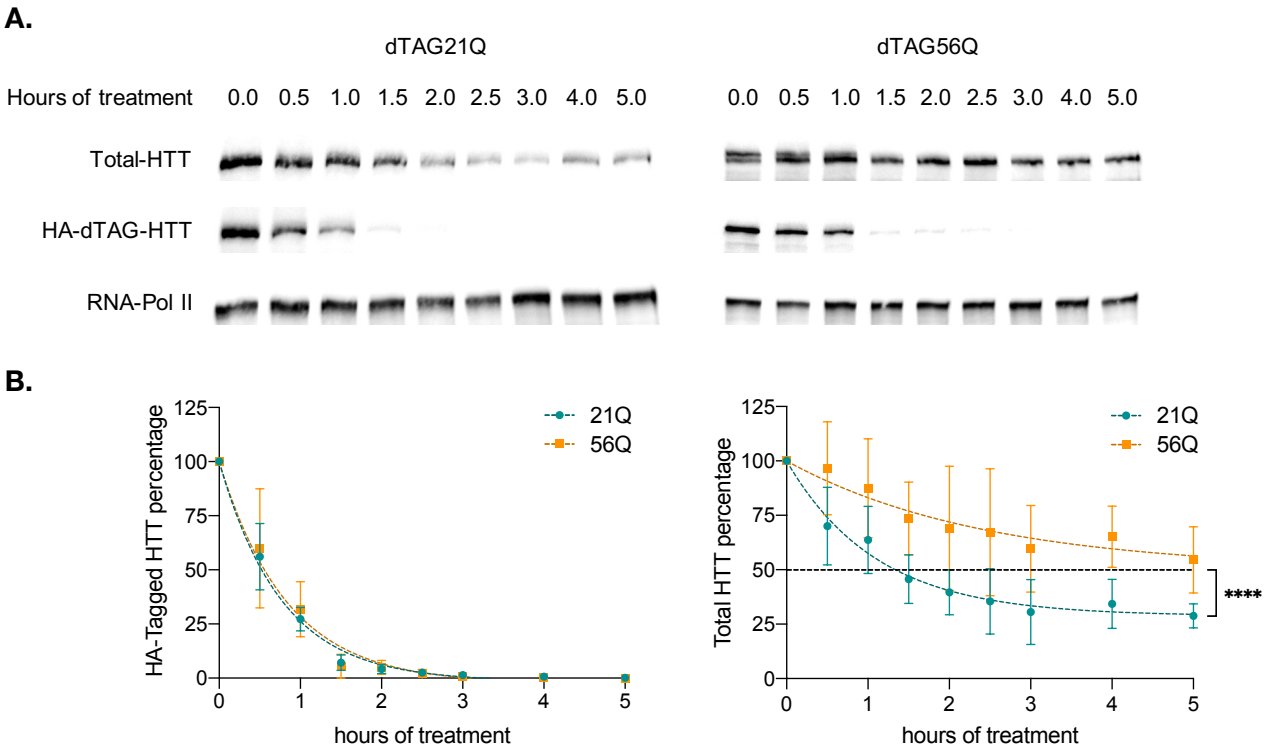

Fig. 21 | Protein-degradation dynamics of normal- and mutant-HTT.
A. Western blot images of dTAG21Q (left) and dTAG56Q (right) treated with λ 1×10^{-7} M for the indicated time period **B.** Best-fitting exponential decay curves calculated on densitometric quantification data. **** P < 0.0001 Extra sum-of-squares F test; n = 3 biological experiments, data are represented as mean of two independent clones ± SD (n = 6 data point analyzed for each concentration).

Non-linear fit of total HTT

Parameters	*21Q*	*56Q*
Y0	100	100
Plateau	28,75	48,01
K	0,905	0,383
Half Life	0,7659	1,81
Tau	1,105	2,611
Span	71,25	51,99

Non-linear fit of HA-dTAG-HTT

Parameters	*21Q*	*56Q*
Y0	100	100
Plateau	-1,466	-2,574
K	1,296	1,199
Half Life	0,5348	0,5779
Tau	0,7715	0,8337
Span	101,5	102,6

Tab. 22 | Parameters for protein-degradation dynamics of normal- and mutant-HTT.
Parameters calculated from the best-fit curves of the densitometric quantification data. **Y0**: Y value when X (time) is zero. It is expressed in the same units as Y. **Plateau**: Y value at infinite times, expressed in the same units as Y. **K**: rate constant, expressed in reciprocal of the X axis time units. **Tau**: time constant, expressed in the same units as the X axis. It is computed as the reciprocal of K. **Half-life**: is computed as ln(2)/K. Is in the time units of the X axis. **Span**: difference between Y0 and Plateau, expressed in the same units as your Y values.

2.4 Validation of the differentiation potential of the dTAG lines.

Since Huntington's disease mainly affects neurons, our main goal is to investigate the HTT functions in this cell type. For this reason, our laboratory has developed a neuronal differentiation protocol to generate medium spiny neurons (MSNs) starting from hESCs[78]. This protocol includes 3 phases (Fig. 23 A); the first phase is the neural induction and has the aim of pushing the hESCs out of pluripotency to become neuroectoderm; the second phase is the patterning specification in which morphogens are used to specify the cells towards a ventral telencephalic progenitor identity; the third phase is the terminal differentiation in which morphogens are removed and BDNF, a neurotrophic factor, is administered, allowing the progenitors to differentiate towards MSNs. It is therefore fundamental for our future analyses to establish whether the two generated dTAG lines are able to correctly differentiate and generate ventral telencephalic neurons. With this objective, we have exposed to our differentiation protocols the two dTAG lines and a non-pathological line as a differentiation benchmark. To monitor the differentiation progression, we acquired phase-contrast images throughout the duration of the protocol. As can be seen in Fig. 23 B before neural induction the cells present the classic embryonic stem cells morphology. After 8 days of differentiation, the organization typical of neural rosettes is already visible, indicating a successful neuroectodermal commitment. From day 16, cells begin to detach in some regions of the well and axonic extensions are exposed. Lastly, during the course of terminal differentiation, the outgrowth of neurites and processes from cell bodies become evident and neurofilaments form dense networks. In addition, during the differentiation protocol, we collected samples for gene expression analysis in order to monitor over time the correct acquisition of differentiation markers and the downregulation of pluripotency-linked genes. In particular, we monitored the expression of OCT4 and PAX6 to verify the correct response to neural induction. OCT4 is one of the key transcription factors involved in the maintenance of the pluripotency state in the ESC while PAX6 is expressed in the neuroectoderm. As expected, OCT4 was correctly downregulated while PAX6 was induced, indicating a correct transition to neuroectoderm. However, at day 4 of differentiation, we observed a significant difference in PAX6 mRNA when comparing the dTAG lines with the RUES2-22Q line. This is not

surprising, since both dTAG lines are HD and our lab as previously reported a differentiation delay in HD lines[11]. We also monitored the expression of FOXG1, GSX2 and ASCL1 to verify the acquisition of the ventral telencephalic progenitor identity. FOXG1 is expressed exclusively in the telencephalon while GSX2 and ASCL1 are transcription factors expressed by the neural progenitors of the ganglionic eminences in the ventral telencephalon. In particular, GSX2 is expressed earlier in differentiation by inducing the expression of ASCL1[79–81].

As shown in figure 24 D, all lines show a progressive induction of FOXG1 indicating the acquisition of the telencephalic identity. GSX2 and ASCL1 are absent until day 8 and have a peak of expression on day 26 indicating that positional identity along the anterior-posterior axis occurs earlier than identity acquisition on the dorsal-vertebral axis. We also monitored the expression of CTIP2 and DARPP32 whose co-expression determines the medium spiny neurons (MSNs) identity. CTIP2 is a key transcription factor for striatum development while DARPP32 encodes a signal transduction protein that responds to dopaminergic and glutamatergic signalling. As shown in Fig 24 H-I both genes are also expressed in hECS and are downregulated in response to neural induction. No increase in the expression of DRAPP32 was detected neither at day 26, suggesting that, at this time point, the neurons still have to progress in their terminal differentiation process. Conversely, CTIP2 expression increases at day 26 of differentiation, indicating a correct differentiation towards striatal neurons. We also observed a significative difference in the amount of CTIP2 mRNA between the RUES2-22Q line and the dTAG lines. This is in line with the data previously reported by our lab[11] showing fewer CTIP2 positive cells in the HD line compared to the wt at late stages of differentiation. Finally, we also monitored HTT expression throughout the differentiation. We observed a significant increase in the amount of HTT messenger RNA between day 0 and day 26 in RUES2-22Q and dTAG21Q lines. Overall these data suggest that both dTAG lines have comparable differentiation ability respect to the RUES2-22Q and that around day 26 neuronal progenitors are acquiring the identity of striatal neurons as they are starting to express CTIP2 but the expression of DARPP32 has not started, yet.

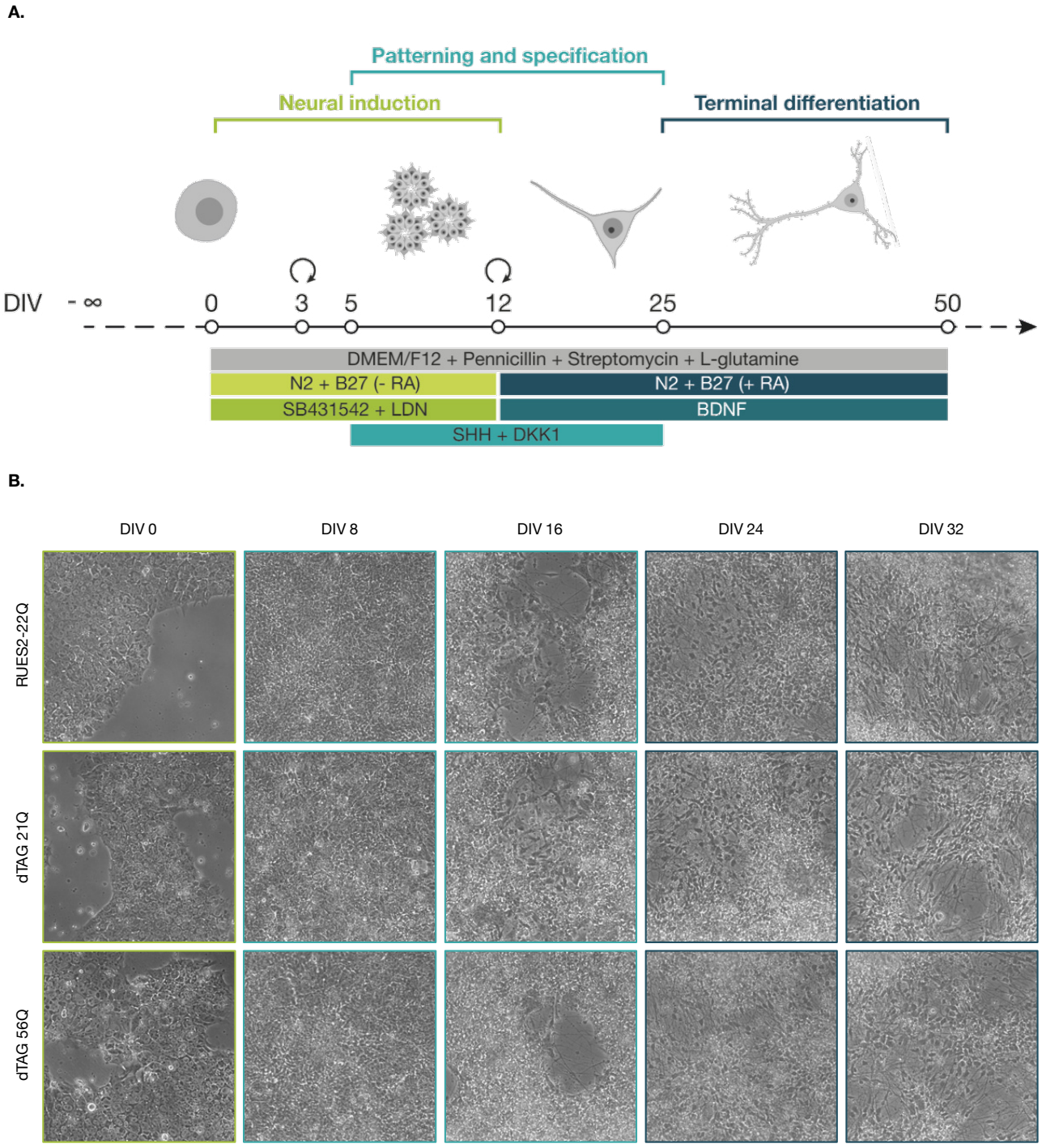

Fig. 23 | Neuronal differentiation of dTAG lines.
A. Neuronal differentiation protocol[78] **B.** Phase contrast images of RUES2-22Q, dTAG21Q and dTAG56Q exposed to the neuronal differentiation protocol for the indicated time period.

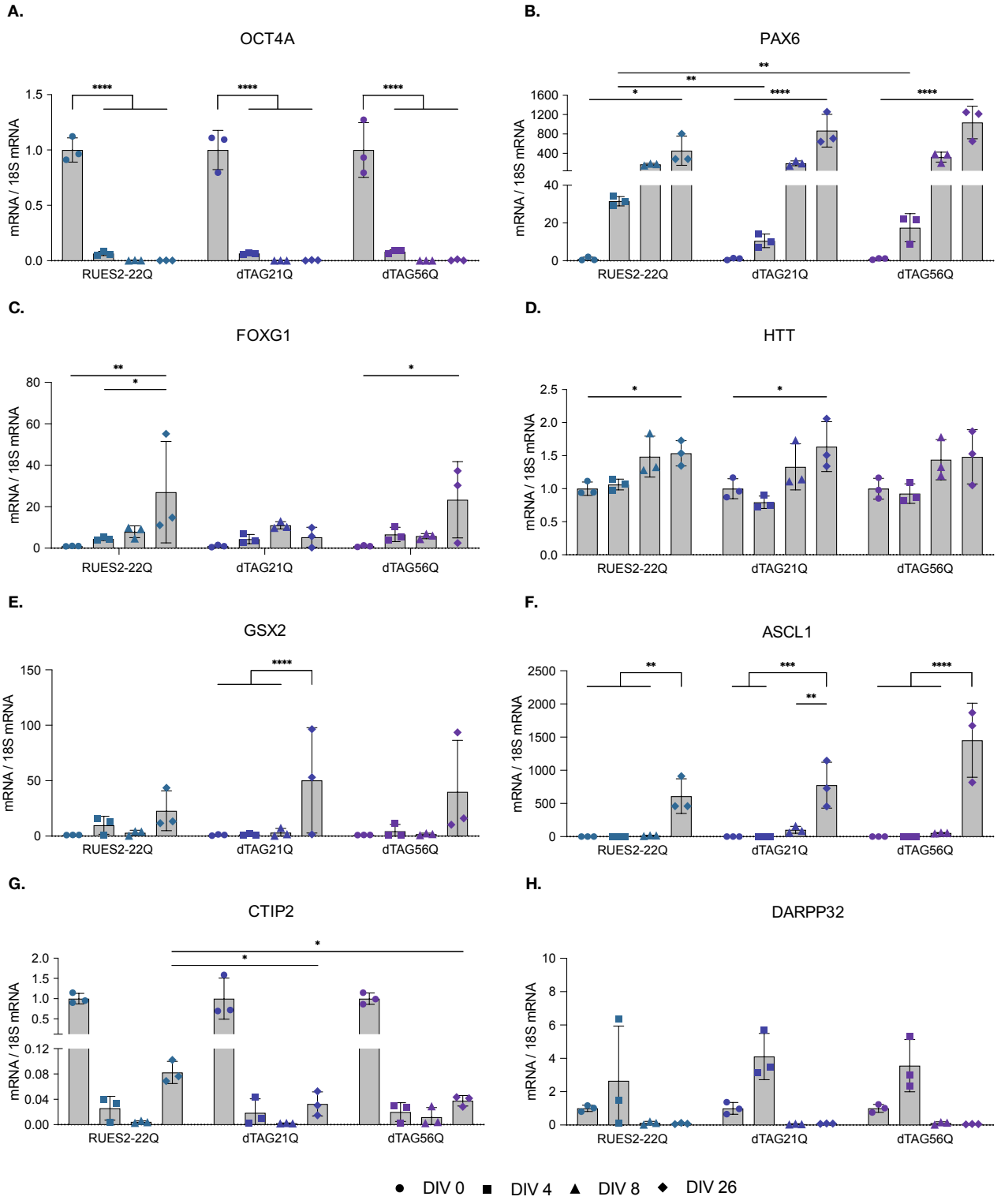

Fig. 24 | Gene expression profiles of dTAG lines differentiation.
A-H. Relative mRNA abundance of the indicated genes throughout the differentiation. DIV 0: circles, DIV 4: squares, DIV 8: triangles, DIV 26: diamonds. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$ one-way ANOVA, Tukey's multiple comparisons test; n = 3 biological experiments, data are represented as mean ± SD.

2.5 Transcriptomic response to wt- or mut-HTT acute depletion

As previously mentioned, the main goal of this work is to investigate the immediate effects triggered in HD cells during acute removal of wt or mutant HTT proteins. The data collected in the previous experiments served to reveal ideal conditions for HTT degradation in our cell lines. Thanks to these, we were able to define the time window and time points in which to investigate the consequences of rapid loss of either wt or mutant HTT. We started by analyzing the transcriptional response to HTT removal in our dTAG lines in self-renewal state, with the aim of repeating the analysis also in neurons generated from them after differentiation. We started at the pluripotent state to investigate whether HTT might have a role already in the hESC since its deprivation in the first stages of the embryonic development causes embryonic lethality[82–85]. Moreover, the experimental design at this stage is fairly simple allowing us to test the system and set up the pipelines for the analysis in a homogeneous cell population with reduced batch-to-batch variability. Conversely, hESC-derived neurons will have a high heterogeneity due to different cell response to the differentiation stimuli and higher batch-to-batch variability.

2.5.1 Transcriptional response to HTT depletion in self-renewing hES dTAG lines

After addressing the best conditions for HTT degradation, we set to investigate the transcriptional changes induced by acute depletion of the protein. Based on the degradation dynamics observed in the previous experiments (Fig. 21 B) we decided to treat our hES dTAG lines for 48 h with either λ [1×10^{-7} M] or with an equivalent volume of DMSO as control. During this temporal window, we collected samples at different time points: 0 h, 2 h, 4 h, 8 h and 48 h for both treatments. For this experiment, each clone was treated independently. After RNA extraction and processing, an equal amount of sample deriving from each clone of the same line was pooled matching the time point and the treatment (Fig. 25 A). Finally, pooled RNA samples were analyzed for bulk RNA-seq. We immediately performed a first quality check of the experiment. To this aim, we imported in R the dataset and assigned categorical variables corresponding to the genotype (21Q, 56Q), the treatment (ligand, DMSO, NT), the time point (0, 2, 4, 8, 48) and the experimental replicate (1, 2, 3). After filtration of not expressed genes and normalization of the read counts, we performed a principal component analysis (PCA) on the whole dataset to check if the samples cluster based on one of the above variables and to eventually spot samples that clusters differently from all the others. This PCA revealed the presence of an outlier, consisting of the second replicate at the 0 h time point of the dTAG21Q line (Fig. 25 B). Since this sample differed substantially from all the other samples in the dataset, we were forced to remove it from analyzes. After removing the above sample, we repeated the PCA and this time the samples separated into clusters based on the experimental time

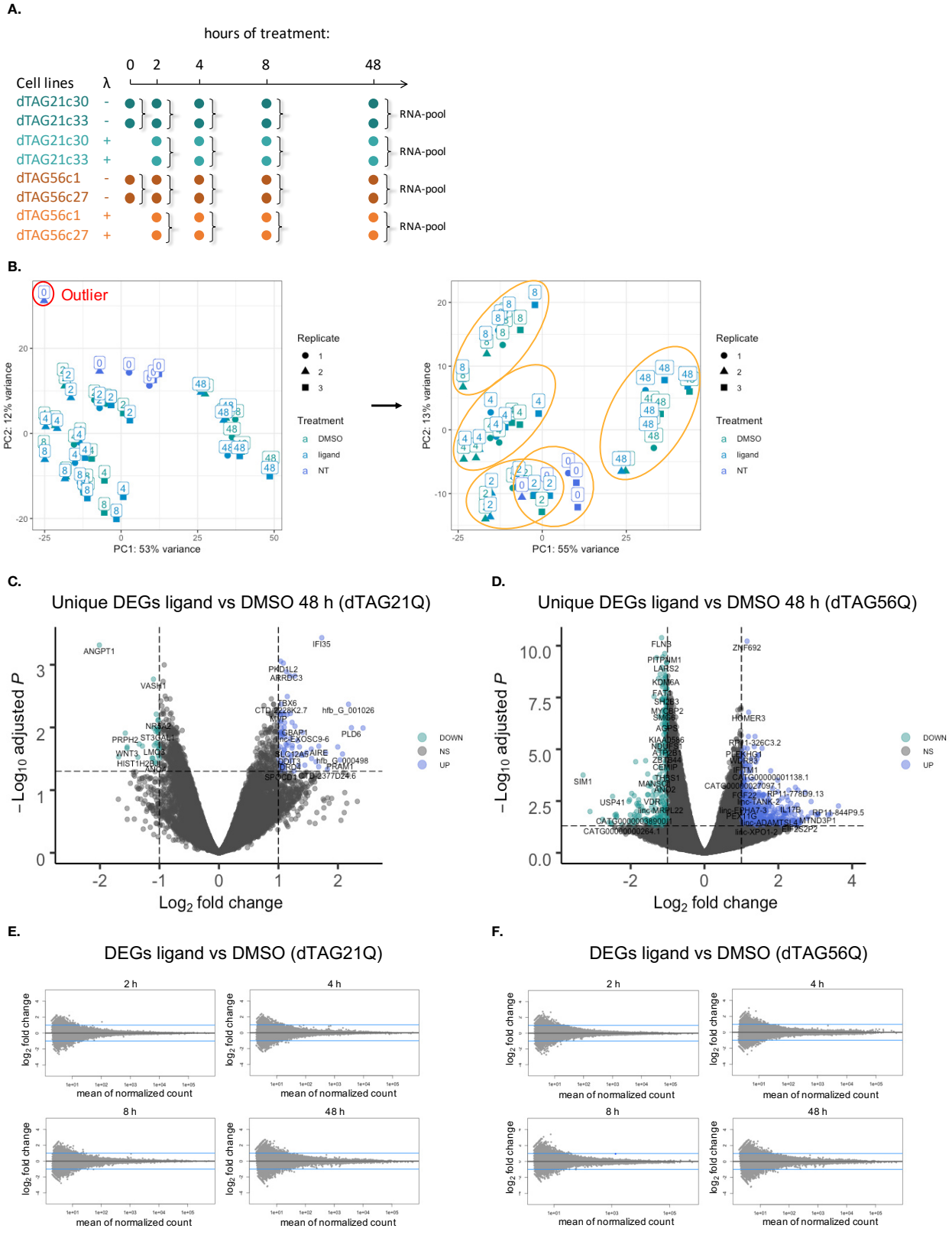

Fig. 25 | Transcriptional response to HTT depletion in hES dTAG lines.
A. Schematic representation of samples collection. **B.** PCA plots of the whole dataset prior (left) and post (right) outlier removal (cell line genotype is not marked). **C.** Volcano plot of the DEGs between 0h and 48 h resulting unique for the ligand treatment after filtration of the DEGs shared with DMSO treatment (dTAG21Q line). **D.** Volcano plot of the DEGs between 0h and 48 h resulting unique for the ligand treatment after filtration of the DEGs shared with DMSO treatment (dTAG56Q line). **E.** MA plots of the DEGs between ligand and DMSO treatments at the indicated time points (dTAG21Q line). **F.** MA plots of the DEGs between ligand and DMSO treatments at the indicated time points (dTAG56Q line).

point. Based on this PCA, we can already deduce several things. The first is that the greatest variance is given by the effect of time. In fact, the first principal component PC1 which explains 55% of the variance separates the 48 h time point from all the others and the PC2 which explains 13% of the variance separates the other time points. The second is that batch to batch effect does not appear to be present since the variance separating matched samples belonging to different experiments is much smaller compared to the variance separating the time points. The third is that the effect due to the treatment seems to be very limited or in any case much smaller compared to the effects of time and experimental replicate. Subsequently, we analyzed the differentially expressed genes (DEGs) comparing the 0 h starting point with each time following point of λ or DMSO treatment in a series of pairwise comparisons that can be classified in two different groups.

- "Dynamic comparisons" referring to comparisons that examine first the variable of time and then filter the DEGs unique for a given treatment
- "Static comparisons" directly comparing DEGs between the two treatments at the same time point.

We initially compared 0 h vs 48 h as we expect to have the largest difference between these samples. After obtaining the lists of DEGs between 0 and 48 h for the two treatments, we filtered the DEGs in order to obtain those unique to the treatment with λ. This operation was done for both the dTAG21Q and dTAG56Q lines (Fig. 25 C-D). During these analyzes, we noticed that the number of DEGs between 0 and 48 h treated with λ was about 10 times greater in the case of the dTAG56Q line compared to the dTAG21Q line. This observation could have been indicative of a difference between the degradation of wt HTT, in the case of the dTAG21Q, and mutant HTT, in the case of the dTAG56Q. However, the fact that the same difference was also found when comparing the number of DEGs after DMSO treatments between the two lines, made us think of a possible intrinsic difference between the two lines that is amplified by the treatments. For this reason, the next step in our analysis was to directly compare the samples at 0 h between the two lines. Interestingly, we found no substantial differences between the two lines thus further investigations are needed to understand the factors influencing this observation. In the meantime, we also directly compared the λ treatments at 48 h against the DMSO treatments at 48 h to verify if the previously obtained DEGs emerged from this comparison, as well. Surprisingly, in both the dTAG21Q line and dTAG56Q line, by directly comparing the 48 h time points of the two treatments, no differentially expressed genes emerged. Therefore, it seems that the genes identified as DEGs unique to λ in the previous 0-48 h comparisons were not significative. To check if this was due to a problem in the 48 h samples, we extended the analysis by performing direct comparisons across all other time points. All comparisons showed no DEGs suggesting that the DEGs unique to λ treatment were just false positives and that there is no difference due to the presence of the ligand in either line (Fig. 25 E-F). This implies that the removal of wt or mutant HTT in hESCs does not induce any significant transcriptional changes in our system in the analyzed time window. From these data, we can conclude that depletion of normal or mutant HTT does not affect transcription in self-renewing hESCs. These data may also indicate that heterozygous depletion of either wt or mutant HTT has no significant effect on hESC homeostasis suggesting that dTAG / - cell lines might be required to exclude compensatory effects of the endogenous HTT alleles. Additionally, HTT upregulation during neuronal differentiation of the dTAG lines (Fig. 24 D) suggests that HTT may have a much more relevant role during neuronal specification and maturation. Hence the next and final step will be to perform the transcriptomic analyses on the hESC-derived neurons during terminal differentiation.

3 | Conclusions and future perspectives

In this project, we successfully engineered HD hESCs to study the immediate transcriptional response to wt or mutant HTT removal. To do so, we harnessed the recently published[28] dTAG system to control HTT degradation. The dTAG cassette was integrated by CRISPR-assisted HDR at the N-terminus of normal or expanded HTT, generating a fusion protein. The heterobifunctional cell-permeable small molecule dTAG-13 (here called λ) administered to the cells acts as a bridge and binds to one side of the dTAG-HTT protein and on the other side to the E3 ubiquitin ligase enzyme, promoting proteosome-mediated degradation. We successfully obtained two edited clones from each targeting process, two clones carrying the engineered normal allele (dTAG21Q) whereas two clones have the mutant allele, targeted (dTAG56Q). We thus confirmed by PCR-sanger sequencing the absence of DNA-sequence mutations in the targeted locus and in the untargeted endogenous allele. In addition, TIDE analyses on the sequencing chromatograms excluded the presence of InDels in the proximity of the gRNA cleavage site of the endogenous untargeted allele. This indicates that the expression of the untargeted allele is preserved and that the same gRNA can be used once again to either perform an additional round of targeting for the generation of a homozygous dTAG line or to inactivate the second allele to remove the contribution of endogenous HTT since this gRNA drives cleavage of the HTT sequence in a coding region. This would allow the generation of a tagged cell line over a null background that could reveal subtle loss-of-function phenotypes. One of the main concerns during the generation of a targeted clonal cell line, is the possibility that not all the cells in the "clone" carry the modification. We, therefore, excluded the presence of mosaicism in the selected clones through a PCR designed to amplify the untargeted HTT allele being able to confirm two clones of each line to be pure and moved forward with the validation. The functionality of the degradation system was validated by western blot analysis and immunocytochemistry quantifications. The selected clones presented a normal karyotype and no alteration in the expression of HTT or pluripotency-linked genes. Furthermore, neuronal differentiation experiments with the dTAG lines validated their ability to differentiate towards striatal neurons therefore showing proficiency for the neuroectodermal lineage. Indeed, both dTAG lines presented similar expression profiles of the major striatal differentiation markers as compared to the RUES2-22Q wt line used as differentiation benchmark. Notably, PAX6 upregulation was slightly delayed in the dTAG-HD lines compared to the RUES2-22Q wt. A delay in CTIP2 expression in the dTAG-HD lines seems also to be present since at day 26 they exhibit a weaker CTIP2 expression compared to the RUES2-22Q wt. These data are in line with the work published in 2018 by the lab showing that HD mutation in iPS lines affects early telencephalic induction and late neural identity during *in vitro* differentiation[11]. Additionally, HTT expression was significantly increased in the late stages of differentiation, advocating for a fundamental role of HTT in cell differentiation. Subsequently, we investigate the minimum concentration of λ required to achieve fast and complete dTAG-HTT degradation obtaining λ dose-response curves for all dTAG clones and estimating 10^{-7} M as minimum λ concentration to achieve complete degradation in 1 h of treatment. These data are in line with what reported by Nabet and colleagues[28] confirming the fast degradation kinetics of the system. Next, we derived total- and tagged- HTT degradation curves for both dTAG lines using the identified λ concentration. We reported no difference in term of degradation kinetic between normal and mutant dTAG-HTT. Interestingly, the percentage of total HTT measured at the steady-state after degradation in the dTAG21Q resulted significantly lower than the expected 50 %, potentially suggesting a hypomorphic phenotype of the mutant allele over the wt that should theoretically account for half the cell HTT. The hypothesis of mutant HTT hypomorphism is not new as many works in the literature reported similar results in murine and *in vitro* model of HD as well as in HD patients[86–88]. However, since dTAG56Q line is not supporting what observed in the

dTAG21Q line, further experiments are required to better investigate normal and mutant HTT expressions. Finally, we used our engineered cell lines to investigate the transcriptomic response to normal or mutant HTT depletion over time in the hESC. RNA-seq data were analysed with two diverse pipelines, one more sensitive and one more stringent. Since the latter addresses only genes that are significantly different between the two treatments, we decided to consider only data obtained with this pipeline. These data indicate that knockdown of either normal or expanded HTT in the hESC is not affecting transcriptional homeostasis. This has two major implications. First, mutant HTT is not affecting the transcriptome of self-renewal hESC since its absence or its hemizygous presence were not significantly different from the respective controls (dTAG56Q, λ, 48 h vs dTAG56Q, NT, 0 h and dTAG21Q, λ, 48 h vs dTAG21Q, NT, 0 h respectively). Second, in the hESC the presence of one functional HTT allele, being it normal or mutant, is sufficient to preserve cell homeostasis (dTAG56Q, λ, 48 and dTAG21Q, λ, 48 h respectively). This does not exclude that inactivation of both alleles would have relevant consequences for hESC homeostasis. To challenge this hypothesis, we plan to generate dTAG / null lines to increase the sensitivity of the system by eliminating potential compensatory effects of the second allele.

Currently, we are finalizing ongoing HTT degradation experiments after neuronal differentiation and are looking forward to analyse the transcriptional response to normal or mutant HTT degradation in mature neurons. We expect to detect a significant difference in gene expression from these experiments since HTT is especially important for neuronal fate acquisition as also confirmed by gene expression analysis during neuronal differentiation. These data will be particularly relevant to have a better understanding of the biological consequences of HTT-lowering approaches in the context of the gene silencing therapeutic strategies currently under clinical development, such as the ASO (antisense oligonucleotide) -mediated silencing of mutant HTT. ASO therapies led to positive results in the preclinical tests and the data that are being collected in phase I/II-III clinical trials seem to be really promising for HD treatment. Indeed, Roche (Ionis) is currently testing its non-allele selective ASO (tominersen) in phase III since 2019 and is planning for prodromal stage trials. They observed dose-dependent reduction of HTT amount in the cerebrospinal fluid of the patients. They are now evaluating the efficacy by tracking the levels of neurofilament ad biomarker to monitor HD progression. On the other side, Wave's allele-selective ASOs (WVE-120101 and WVE-120102) are also being evaluated in phase Ib/IIa trials and were successfully able to lower mutant Huntingtin amount in the spinal fluid in a dose-dependent manner. Therefore, we hope to provide precious information on the biosafety of these HTT-lowering approaches that are close to become available as therapy for HD treatment. Overall, we were able to insert the dTAG system in our hESC line, the cells responded well to the λ treatment showing rapid and efficient degradation of dTAG-HTT already with 10^{-7} M dose. We assessed the degradation kinetic in both our lines and performed the first RNA-seq experiment in the hPSC which revealed no significant transcriptomic alteration due to normal- or mutant-HTT knockdown. The second RNA-seq experiment on the hPSC-derived neurons is currently ongoing and we are looking forward to analyse the results, hoping to provide useful information about the consequence of HTT lowering approaches in the neurons.

4 | Materials and Methods

4.1 Cell culture

The RUES2 lines used in this thesis are hESC isolated at the Rockefeller University and edited with CRISPR/Cas9 system to replace the first exon of HTT with pathological variants carrying different CAG lengths (Ruzo et al.) or a physiological variant as control. These lines grow on Geltrex (Thermo Fisher Scientific) coated dishes and in feeder-free mTeSR1 medium (StemCell Technology). Cell split occurs once a week at approximately 70-80% confluence. ReleSR (StemCell Technology) is used to detach the cells without dissociating the cells inside colonies. After approximately 5 min of incubation cell have their adhesion to the plate weakened. ReleSR can be removed and fresh mTeSR1 is added to mechanically lift off the cells to be split. Cells batch are frozen in a cold freezing solution composed of Knockout serum replacement (Life technologies) with the addition of 10% DMSO, 10 μM ROCK inhibitor (Y-27632) in 1ml freezing vials at -150°C using the freezing boxes (CoolCell). Cells are rapidly thawed in a water bath at 37°C, diluting immediately the freezing solution with 9 volumes of warm phosphate saline buffer PBS. Cells in solution are centrifuged at 1200 x g for 3 minutes, then resuspended and plated in mTeSR1 with 10 μM ROCK inhibitor. Each batch is kept in culture for a maximum of 3 months before thawing a new batch.

Medium/Reagent	Manufacturer
Accutase	Millipore
B27 supplement (50X)	Life Technologies
B27 supplement w/o RA (50X)	Life Technologies
BDNF	PrepoTech
Cultrex Reduced Growth Factor	Trevigent
DKK-1	PrepoTech
DMEM-F12 (#21331)	ThermoFisher
DMSO	Goldbio
Geltrex	ThermoFisher
GLUTAMAX (100X)	ThermoFisher
Knockout Serum Replacement (SR)	ThermoFisher
LDN	Evotec
mTeSR1	StemCell Technologies
N2 supplement (100x)	Life Technologies
PBS	Euroclone
Penicilin Streptomycin (Pen/Strep)	Euroclone
SB431542	Evotec
SHH(C25II)-N	Tocris
ReleSR	StemCell Technologies
ROCK Inhibitor Y-27632	StemCell Technologies
Trypan Blue Stain (0.4%)	Life Technologies

Tab. 26 | Cell culture Media and reagents

4.2 Editing design

The dTAG lines generated in this thesis work derive from RUES2-58Q clone 25.2[77]. The design of the dTAG cassette was the same used by the author of the 2018 publication[28]. for the N-terminal targeting except that our cassette was adapted for HDR-mediated integration instead of MMEJ-mediated integration. To do so, 800 bp DNA sequence upstream and downstream of the first HTT codon were selected as left and right homology arms respectively. In addition, to target either the wt or the mutant HTT allele, two slightly different versions of the cassette were designed. This is due to the fact that the right homology arm (R-HA) include the poly-Q repeat region. Thus a 21Q R-HA cassette was used to target the wt allele, whereas a 56Q R-HA cassette was used to target the mutant allele. The two dTAG cassettes were synthesized and cloned in a pUC-57 vector by Genscript, generating the two donor plasmids used in the editing experiments. To integrate the dTAG cassette we used the CRISPR/Cas9 system to induce a double-strand break (DSB) in the 5' UTR region of the HTT gene. The HTT_preQ_sgRNA17 (Invitrogen TrueGuideSyn) used for the editing (Fig. 27) was selected among a list of gRNAs provided by the group that generated the RUES2 HD allelic series. All the gRNAs in the list were already screened for good cleavage efficiency so we selected the one with the cleavage site closer to the first HTT codon to maximize the editing efficiency.

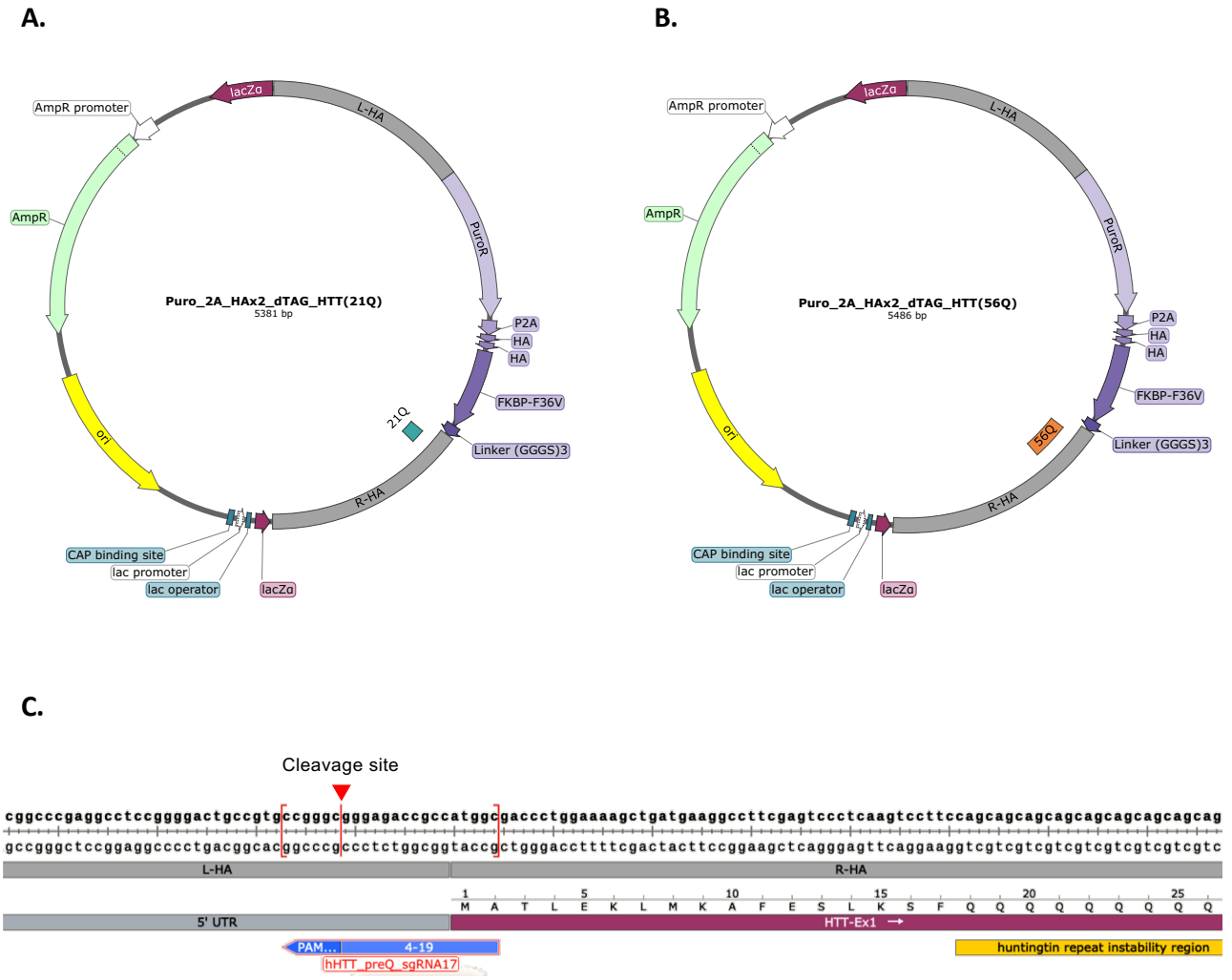

Fig. 27 | Editing design.

4.3 dTAG cell lines generation

The dTAG lines were generated by genome editing from the RUES2-58Q clone 25.2. The parental line was nucleofected to convey the DNA cleaving machinery, composed of Cas9 protein and the sgRNA as a ribonucleoprotein complex (RNP), along with the HDR-compatible donor plasmid (Fig. 27 Puro_2A_HAx2_dTAG_HTT(21Q) or Puro_2A_HAx2_dTAG_HTT(56Q)).

4.3.1 Nucleofection

This protocol uses the Lonza hES nucleofection kit and requires the preparation of the following reagents:

The Lonza nucleofection buffer is composed of two solutions:

1. *Mix 82 µl of nucleofector solution with 18 µl of supplement (volumes for one reaction).*

The sgRNA was annealed *in vitro* starting from the crRNA (HTT_preQ_sgRNA17) and the tracRNA (IDT Alt-R tracrRNA). The protocol for the annealing is the following:

2. *Mix equimolar amounts of crRNA and tracrRNA (60 pmol each) into 5 µl final volume with duplex buffer in a 200 µl PCR tube.*
3. *In thermocycler, incubate for 5 min at 95°C and then let the reaction to cool down at RT. This is the ready-to-use sgRNA solution.*

Subsequently, Cas9 protein (IDT Alt-R S.p. Cas9 Nuclease V3) is mixed with the previously generated sgRNA *in vitro*:

4. *Mix 60 pmol of sgRNA solution (2.) with 40 pmol of Cas9 in 10 µl LONZA nucleofection buffer at RT for 10-15 min.*

After the 15 min incubation, the donor plasmid is added to the mix:

5. *Add 1 µg of donor plasmid to the RNP mix (Puro_2A_HAx2_dTAG_HTT(21Q) or Puro_2A_HAx2_dTAG_HTT(56Q))*

The following steps can be performed during the 15 min incubation (4.)

6. *Lift-off the cells using 1.5ml Accutase (Sigma-Aldrich) with 10 µM Rho-kinase Inhibitor (RI)*
7. *Wait for the cells to detach spontaneously*
8. *Block with 4 volume PBS*
9. *Centrifuge at 1200 rpm for 3'*
10. *Resuspend the cells in mTeSR1 (4 ml for each starting plate) and count*
11. *Aliquot 2 × 10^6 cells for each reaction in a new 15 ml tube and pellet the cells*
12. *Prepare 3 ml of warm mTeSR1 with 10 µM RI and 1 µM SCR7 (ligase IV inhibitor) for each reaction.*
13. *Place 2.5 ml of warm mTeSR1 into a coated 6 cm Petri dish.*

The following steps can be performed after step. (5.)

14. *Resuspend the cell pellets with 90 µl nucleofection buffer and transfer them to the RNP-DNA mix.*
15. *Pipette up and down 3 times to homogenize the suspension.*
16. *Transfer the suspension in a nucleofector cuvette and shock the cells using program B-016*.

17. *Block shocked suspension with 500 µl warm mTeSR1 with 10 µM RI and 1 µM SCR7 to inhibit NHEJ and then transfer the cell suspension into the 6 cm petri with 2.5 ml of medium (13.) using the provided plastic Pasteur.*
18. *Next day replace the medium with freshly made 3 ml mTeSR1 RI and SCR7 until the day after.*

4.3.2 Clones isolation and expansion

Two days after nucleofection the medium is replaced with mTeSR1 containing 300 ng/ml of Puromycin. The antibiotic selection is kept for 5-7 days. By that time several clones in the plate should be visible. Once clones grow to a size of hundreds of cells, they can be scraped out and transferred into a coated 96 well plate using a P20 pipette. Starting from the next day, the medium is replaced every day with fresh mTeSR1. Once the clones fill up the 96 well, they are split in a 24 well plate. Starting from the day after the split, the antibiotic selection is applied for further 4 days. This usually further cleanup for non-resistant clones. When clones are ready to be split, half of the cells are transferred in a 6 well plate to be frozen once the well is filled and half is pelleted for genotyping.

4.3.3 PCR screening

Clones having the correct integration of the dTAG cassette were screened based on a PCR reaction amplifying selectively the genomic region across the right homology arm (R-HA) generated by the integration of the dTAG cassette (Tab 28 Set 1). Thus one primer anneals to the dTAG cassette that can be present as donor plasmid randomly integrated into the genome or derive from correct integration of the cassette in the HTT locus. The second primer anneals to the genomic region right downstream the R-HA and is always present whether the editing event occurred or not. Thus DNA amplification occurs only if the dTAG cassette is correctly integrated into the HTT locus. The same approach was used to validate the correct integration by amplification of the region across the left homology arm (Tab 28 Set 2).

4.4 Quality controls

4.4.1 Genetic quality controls

To exclude the presence of mosaicisms in the analyzed clones a third PCR was performed (Tab 28 Set 3). This PCR amplifies from the 5’ UTR to the CAG repeat region in HTT exon 1. Thus, it allows to discriminate the amplified allele based on the length of the CAG repeat and it allows to check for the presence of InDels in the untargeted endogenous allele due to Cas9 activity. In case of mosaics, the PCR produced a doublet of bands corresponding to the two endogenous alleles, in case of a pure clone a single band corresponding to the untargeted endogenous allele is visible. The same PCR was used for the TIDE analysis. Amplification products were purified from the agarose gel and Sanger sequenced by Eurofins Genomics. Chromatogram data (.ab1) from the sequencing were uploaded to the TIDE website (https://tide.nki.nl/). The left boundary parameter was adjusted to take into account the distance of the cleavage site from the start of the sequence, other parameters were not modified.

Primer Name	Sequence (5' -> 3')	Tm	Application	Set
dTAG-hHTT HDR FW1	CTCTCGTCTTCGATGTGGAGCT	62.1	Screening PCR	1
dTAG-hHTT HDR RV1	CAAATCTCAGGCTGTTTTAAGTGCC	61.3	Screening PCR	
dTAG-hHTT HDR FW2	CCAGCTCACTGGGGGTG	60.0	Screening PCR	2
dTAG-hHTT HDR RV2	GTCCGGATCGACGGTGTG	60.5	Screening PCR	
End-all FW	GGTCCAAGATGGACGGCCGC	65.5	PCR / TIDE	3
End-all RV	GGCGGCTGAGGAAGCTGAGGA	65.7	PCR / TIDE	
hOCT4A FW	AGGCTCTGAGGTGTGGGGGA	60.0	qPCR	4
hOCT4A RV	AGGCGGCTTGGAGACCTCTCA	60.0	qPCR	
hPAX6 FW	ATGGCAGCTGTGTGTGACACT	60.0	qPCR	5
hPAX6 RV	GTGGAATTGGTTGGTAGACAC	60.0	qPCR	
hFOXG1 FW	TGTTGACTCAGAACTCGCTGG	60.0	qPCR	6
hFOXG1 RV	CTGCTCTGCGAAGTCATTGAC	60.0	qPCR	
hHTT-EX3 FW	CGCTATGGAACTTTTTCTGCTGTG	60.0	qPCR	7
hHTT-EX4 RV	CTGTAACCTTGGAAGATTAGAATCCATC	60.0	qPCR	
hGSX2 FW	CTCGCTCATCATCAAGGACA	60.0	qPCR	8
hGSX2 RV	AGTGCAGGTGCGAAGTGAC	60.0	qPCR	
hASCL1 FW	TGTGCAAAAGCAGTGGGCTC	60.0	qPCR	9
hACL1 RV	TATTGGGGTGGGGGCTACTG	60.0	qPCR	
hCTIP2 FW	GATGCCAGAATAGATGCCGG	60.0	qPCR	10
hCTIP2 RV	CGCCACACTGCTTCCTTTTG	60.0	qPCR	
hDARPP32 FW	CTGAGGACCAAGTGGAAGAC	60.0	qPCR	11
hDERPP32 RV	CGCCACACTGCTTCCTTTTG	60.0	qPCR	
h18S FW	CGGCTACCACATCCAAGGAA	60.0	qPCR	12
h18S RV	GCTGGAATTACCGCGGCT	60.0	qPCR	

Tab. 28 | Primer list

4.4.2 Karyotype

The karyotype of the edited clones was checked with the Q-banding assay by ISENET Biobanking. The analysis consists of a treatment with quinacrine dihydrochloride to stain chromosomes. Quinacrine is a fluorescent dye whose fluorescence is enhanced in A-T rich regions and is quenched in G-C rich regions producing a specific banding pattern in each chromosome. For the analysis, cells were seeded in T25 cell culture geltrex coated flasks. Cells were sent to the company at approximately 70 % confluence as recommended.

4.5 RNA-based assays

4.5.1 RNA extraction and quality controls

RNA was extracted from thawed cell pellets using the TRIzol reagent (Life Technologies) following the manufacturer's instructions. Extracted and purified RNA samples were quantified with the Nanodrop 1000 (Thermo Fisher Scientific) and diluted to 500 ng/µL. One µL of each sample was loaded into a 1 % agarose gel to check for RNA integrity and for the absence of DNA contamination. If DNA contamination was present in at least one sample of the experimental group, all samples in the group were treated with Ambion DNA-freeTM DNase (Invitrogen) for DNA degradation. After eventual DNase treatment, 1 µL of each treated sample was loaded into a 1 % agarose gel to confirm DNA degradation.

4.5.1 Reverse-transcription (RT)

DNA-free RNA (1 μg) was reverse-transcribed to cDNA by using the iScript cDNA Synthesis Kit (Bio-Rad). RT is performed in a thermocycler using the following protocol:

Step	Temp	Time
Priming	25°c	5’
Reverse transcription	42°c	30’
Rt inactivation	50°c	15’
	85°C	5’
Hold	4°c	Infinite

Tab. 29 | Reverse transcription protocol

After RT, cDNA samples were diluted 1:60 for the qPCR of the reference gene (h18S) and 1:20 for the qPCR of all other genes.

4.5.2 Real-time qPCR

Real-time qPCRs were performed using the TaqMan assay (Thermo Fisher scientific) to analyse the key pluripotency related genes or the SsoFastTM EvaGreen® Supermix 2X (Bio-Rad) to monitor gene expression during differentiation. All qPCR reactions were performed using a CFX96TM Real-Time System (Bio-Rad) coupled with a C1000TM Thermal Cycler (Bio-Rad) and following the manufacturer instructions. TaqMan assays cat. no. and primers used in with the EvaGreen Supermix are listed in table 30 and 28 respectively. Raw data were analysed by CFX Maestro Software (BioRad).

Target gene	Assay ID
hOCT4	Hs01570480_s1
hSOX2	Hs01053049_s1
hFUT4	Hs01106466_s1
hKLF4	Hs00358836_m1
hKLF5	Hs00156145_m1
h18S	Hs99999901_s1

Tab. 30| TaqMan assays list

4.5.3 Bulk RNA-seq

Samples for bulk RNA-seq were diluted to a final concentration of approximately 50 ng/μL. Besides, each sample was quality checked by Nanodrop 1000 (Thermo Fisher scientific) to have an OD 260/280 ratio ≥ 1.9 and an OD 260/230 ratio ≥ 1.7 according to the requirements indicated by Eurofins Genomics for sample submission. Each sample was assigned to a barcode and shipped in dry ice to Eurofins for Bulk RNA-seq.

4.6 Western blot

Thawed cell pellets were resuspended and lysed in RIPA buffer (Table 31) supplemented with HaltTM Protease and Phosphatase Inhibitor Cocktail 1 mM (Thermo Fisher Scientific) and PMSF 1 mM (Sigma-Aldrich).

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

Tab. 31 | RIPA buffer recipe

Lysis reactions were kept in ice for 30 min and then centrifuged at 12'000 x g for 30 min. Supernatant was transferred in a new tube and protein concentration was determined by the BCA Protein Assay (Thermo Fisher Scientific). A total protein amount of 35 ug was desaturated at 95°C with 8% sample buffer for 10 min and then loaded into 5 % SDS-PAGE gels. After 3 to 4 hours of electrophoretic separation, proteins were transferred onto a nitrocellulose membrane using the Trans-Blot Turbo System (BioRad) for 20 min. Membrane were rinsed with TBS-T 0.1% and blocked for 1 h at RT in TBS-T 0.1% containing 5% Bovine Serum Albumin (Sigma-Aldrich). Membranes were incubated with primary antibodies overnight at 4°C. Primary and secondary antibodies are listed in table 32. Three washes in TBS-T 0.1% were performed before 1 h incubation at RT with HRP-conjugated secondary antibody (Biorad). Three washes in TBS-T 0.1% were performed before setting up the chemiluminescent reaction. Finally, HRP-signal was detected with a chemiluminometer (ChemiDoc, Bio-Rad) using ClarityTM Western ECL Substrate (Bio-Rad) following manufacturer's instructions. Densitometric quantification was performed using the ImageLab software (Biorad).

Antigen	Species	Manufactorer	#	Dilution	Application	Conjugate
Huntingtin (D7F7)	Rabbit	Cell Signaling	5656	1:1000	WB	-
				1:500	IF	-
HA-Tag	Rabbit	Cell Signaling	3724	1:1000	WB	-
HA-Tag	Mouse	Cell Signaling	2367	1:1000	WB	-
				1:500	IF	-
Vinculin	Mouse	Sigma-Aldrich	9131	1:1000	WB	-
RNA-Polymerase II	Mouse	Abcam	5408	1:1000	WB	-
Rabbit IgG	Goat	Life Technologies	A11008	1:500	IF	Alexa Fluor 488
Mouse IgG	Goat	Life Technologies	A11004	1:500	IF	Alexa Fluor 568
Rabbit IgG	Goat	Biorad	1706515	1:3000	WB	HRP
Mouse IgG	Goat	Biorad	1706516	1:3000	WB	HRP

Tab. 32 | Antibody list

4.7 Immunocytochemistry

Cell were rinsed with RT PBS and fixed with ice-cold 4% PFA for 15 min at 4°C. Three washes in PBS were performed before either storing the plates at 4°C or proceeding with immunostaining. In the first case plates are stored in storage solution (PBS with 1% NaN_3) at 4°C (before proceeding with the immunostaining plates were rinsed in PBS to remove storage solution). In the second case cells were permeabilized in PBS with 0.5% Triton-X-100 for 10 min at RT and directly incubated with blocking solution (PBS with 2.5% normal goat serum (NGS, Vector) and 0,25% Triton X-100) for 1 h at RT. Then cells were incubated overnight at 4°C with primary antibodies diluted in blocking solution. Primary and secondary antibodies are listed in table XXX. Next, plates were washed three times in PBS with 0,1% Triton X-100 before 1 h incubation at RT with Alexa Fluor-conjugated secondary antibodies diluted 1:500 in blocking solution. Cells were incubated for additional 10 min with 0.1 µg/mL Hoechst (Invitrogen) in PBS. Finally, three wash steps were performed in PBS with 0,1% Triton X-100 before image acquisition. At least 20 40X randomized fields were acquired for each well using the InCell 6000 automated microscope. At least three wells for each line were acquired for each biological replicates. Signal intensity for HTT and HA stains were quantified with CellProfiler software (Fig. 33).

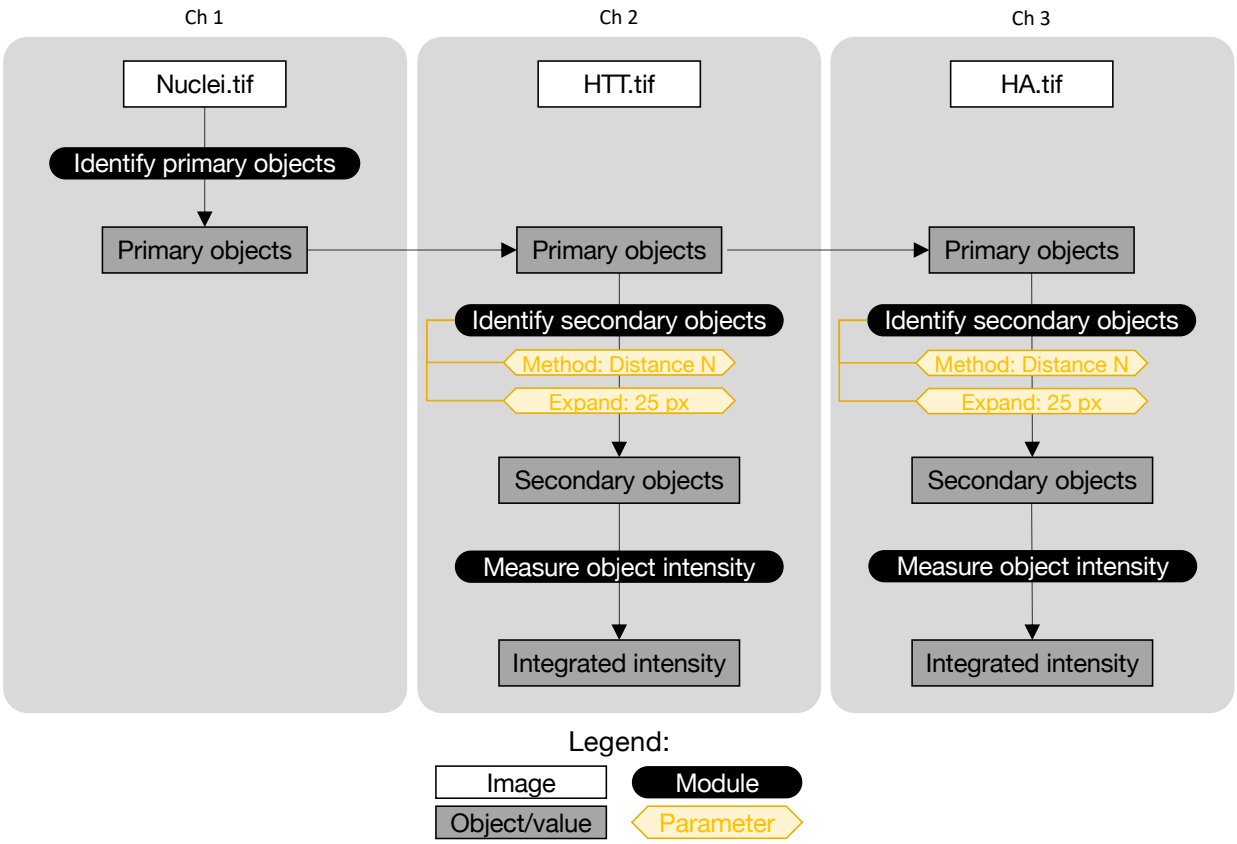

Fig. 33 | CellProfiler pipeline scheme

4.8 Striatal differentiation of RUES2 lines

Neuronal differentiations of the RUES2 hES lines (control or dTAG) were performed according to the protocol described by Delli Carri et al. (2013). The protocol consist in three major steps each of them with a specific goal:

STEP 1. Neural induction (from day 0 to day 11-12) pushes the cells out of pluripotency towards neuroectodermal fate (this is done through the dual SMAD inhibition).

STEP 2. Patterning and specification (from day 5 to day 25) provides the instructions to the cells to acquire ventral telencephalic progenitor identity.

STEP 3. Terminal differentiation (from day 25 to the end of the protocol) supplies the neural progenitors with neurotrophic factors and let them spontaneously differentiate towards striatal neurons.

Since STEP 1. and 2. overlap, four different medium composition are used throughout the protocol. The composition of each medium is reported in the following table:

Supplement:	Final concentration			
Penicillin	100 U/ml	100 U/ml	100 U/ml	100 U/ml
Streptomycin	100 µg/ml	100 µg/ml	100 µg/ml	100 µg/ml
L-glutamine	2 mM	2 mM	2 mM	2 mM
N2	1 %	1 %	1 %	1 %
B27	1 %	1 %	-	-
B27 + RA	-	-	2 %	2 %
SB431542	10 µM	10 µM	-	-
LDN193189	500 nM	500 nM	-	-
SHH	-	200 ng/µL	200 ng/µL	-
DKK1	-	100 ng/µL	100 ng/µL	-
BDNF	-	-	30 ng/µL	30 ng/µL
Day	**0-5**	**5-12**	**12-24**	**25**

Fig. 34 | Differentiation media

Cells are plated at day -2 in mTeSR1 with 10 µM RI onto Geltrex coated 6-well plates at a density of $6{,}4 \times 10^5$ cells/cm^2. At day -1 the medium is replaced with fresh mTeSR1 medium without RI. At day 0 cell should have reached a 70-80 % confluence and mTeSR1 is replaced with DMEM-F12 (day 0-5). Medium is replaced with freshly prepared day 0-5 medium every day. At day 3-4 cells should have reached 100% confluence and should be much smaller compared to day 0. At this stage cells are splitted as clumps with ReLeSR with a 1:3 ratio forming a 240 µL drop in geltrex coated 6-well plates. This ensure that the high cell density is maintained even after the split. Once cells have attached onto the plate, additional medium is added to fill the well surface. At day 5, day 0-5 medium is replaced with day 5-12 medium to start the patterning and specification thanks to the presence of morphogens. Around day 7-10 neuroepithelial rosettes should be visible, this is an indication of a correct response to the neural induction. At day 12 cells are detached as single cells in Accutase and replated on Cultrex reduced growth factor coated 6-well plates as highly dense drops ($8{,}0 \times 10^3$ cells/µL). Concomitantly, day 5-12 medium is replaced with day 12-24 medium (removing the SMAD inhibitors and adding neurotrophic factors). From now on, ¼ of total medium volume is replaced every other day with fresh medium. At day 25, day 12-24 medium is replaced with day 25 medium removing morphogens and letting the neural progenitors progress through terminal differentiation to acquire striatal neurons identity.

4.9 Bioinformatic analysis of RNA-seq data

RNA-seq data were quality controlled and trimmed before being imported in STAR Aligner for reads mapping to the human genome (hg 38). SATR also count the number of reads mapped to each gene generating as output the dataframe containing the read count value for each gene in the Human genome in each sample. This dataframe was imported in R for the analyses. Information concerning genotype (dTAG 21Q or dTAG56Q), treatment (ligand or DMSO control), time point (0 h (untreated), 2 h, 4 h, 8 h, 48 h) and biological replicate were extracted from the dataframe and stored in vectors. For each analysis, the selected dataset was normalized using DESeq2 to take into account the technical variability of the RNA-seq and to minimize it to highlight biological difference. Besides, genes having sum of read count across all samples smaller than 10 were excluded from the analysis. Principal component analysis (PCA) was performed on the whole dataset to check for sample clustering. Since the sample corresponding to dTAG 21Q untreated replicate 2 was significantly different from all other samples it was considered as an outlier and excluded from the analysis. Afterwards, all samples were tested for differential gene expression by performing pairwise comparisons. At first, one line at a time was considered for the analysis, starting from the dTAG21Q then proceeding with the dTAG56Q. Subsequently both line were considered to compare the two lines in the same condition (treatment, time point). Two different pipelines were used for differential expression (DE) analysis (Fig. 35):

Pipeline A (more sensitive):

1. DE analysis between ligand or DMSO at a given time point and untreated at 0 h
2. Two new datasets containing differentially expressed genes (DEGs) between ligand or DMSO treatment and untreated are obtained as outputs
3. The dataset of DEG for the DMSO treatment is compared with the one for ligand treatment and the common genes are filtered out from the latter
4. The output is a dataset containing only DEGs that are unique for ligand treatment

Pipeline B (more stringent):

1. DE analysis between ligand and DMSO at a given time point
2. The output is a dataset containing DEGs between ligand and DMSO treatment at the analyzed time point

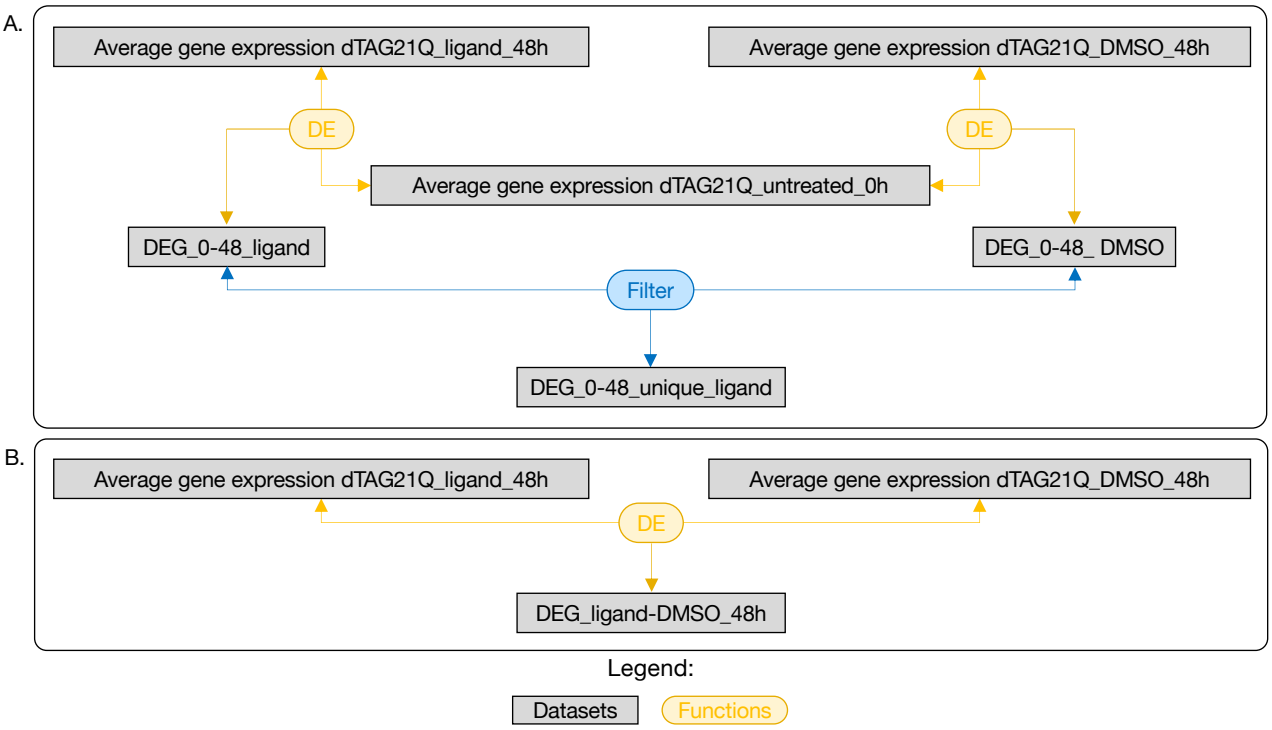

Fig. 35 | RNA-seq analysis pipeline scheme

5 | References

1. Kegel, K. B. *et al.* Huntingtin expression stimulates endosomal-lysosomal activity, endosome tubulation, and autophagy. *J Neurosci* **20**, 7268–7278 (2000).
2. Petersén A, null *et al.* Expanded CAG repeats in exon 1 of the Huntington's disease gene stimulate dopamine-mediated striatal neuron autophagy and degeneration. *Hum Mol Genet* **10**, 1243–1254 (2001).
3. Martinez-Vicente, M. *et al.* Cargo recognition failure is responsible for inefficient autophagy in Huntington's disease. *Nat Neurosci* **13**, 567–576 (2010).
4. Ravikumar, B. *et al.* Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nat Genet* **36**, 585–595 (2004).
5. Wong, Y. C. & Holzbaur, E. L. F. The regulation of autophagosome dynamics by huntingtin and HAP1 is disrupted by expression of mutant huntingtin, leading to defective cargo degradation. *J Neurosci* **34**, 1293–1305 (2014).
6. Kalvari, I. *et al.* iLIR: A web resource for prediction of Atg8-family interacting proteins. *Autophagy* **10**, 913–925 (2014).
7. Rui, Y.-N. *et al.* Huntingtin functions as a scaffold for selective macroautophagy. *Nat Cell Biol* **17**, 262–275 (2015).
8. Hipp, M. S. *et al.* Indirect inhibition of 26S proteasome activity in a cellular model of Huntington's disease. *J Cell Biol* **196**, 573–587 (2012).
9. Hipp, M. S., Park, S.-H. & Hartl, F. U. Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends Cell Biol* **24**, 506–514 (2014).
10. Soares, T. R., Reis, S. D., Pinho, B. R., Duchen, M. R. & Oliveira, J. M. A. Targeting the proteostasis network in Huntington's disease. *Ageing Res Rev* **49**, 92–103 (2019).
11. Conforti, P. *et al.* Faulty neuronal determination and cell polarization are reverted by modulating HD early phenotypes. *PNAS* **115**, E762–E771 (2018).
12. Keryer, G. *et al.* Ciliogenesis is regulated by a huntingtin-HAP1-PCM1 pathway and is altered in Huntington disease. *J Clin Invest* **121**, 4372–4382 (2011).
13. McKinstry, S. U. *et al.* Huntingtin Is Required for Normal Excitatory Synapse Development in Cortical and Striatal Circuits. *J. Neurosci.* **34**, 9455–9472 (2014).
14. Ross, C. A. & Tabrizi, S. J. Huntington's disease: from molecular pathogenesis to clinical treatment. *The Lancet Neurology* **10**, 83–98 (2011).
15. Drouet, V. *et al.* Sustained effects of nonallele-specific Huntingtin silencing. *Ann Neurol* **65**, 276–285 (2009).
16. Kordasiewicz, H. B. *et al.* Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. *Neuron* **74**, 1031–1044 (2012).
17. Boudreau, R. L. *et al.* Nonallele-specific silencing of mutant and wild-type huntingtin demonstrates therapeutic efficacy in Huntington's disease mice. *Mol Ther* **17**, 1053–1063 (2009).
18. Valor, L. M. Transcription, Epigenetics and Ameliorative Strategies in Huntington's Disease: a Genome-Wide Perspective. *Mol Neurobiol* **51**, 406–423 (2015).
19. Zuccato, C. *et al.* Huntingtin interacts with REST/NRSF to modulate the transcription of NRSE-controlled

neuronal genes. *Nature Genetics* **35**, 76–83 (2003).

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

21. Steffan, J. S. *et al.* The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. *Proc Natl Acad Sci U S A* **97**, 6763–6768 (2000).

22. McCampbell, A. *et al.* CREB-binding protein sequestration by expanded polyglutamine. *Hum Mol Genet* **9**, 2197–2202 (2000).

23. Holbert, S. *et al.* The Gln-Ala repeat transcriptional activator CA150 interacts with huntingtin: neuropathologic and genetic evidence for a role in Huntington's disease pathogenesis. *Proc Natl Acad Sci U S A* **98**, 1811–1816 (2001).

24. Takano, H. & Gusella, J. F. 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 (2002).

25. Dunah, A. W. *et al.* Sp1 and TAFII130 Transcriptional Activity Disrupted in Early Huntington's Disease. *Science* **296**, 2238–2243 (2002).

26. Li, S.-H. *et al.* Interaction of Huntington disease protein with transcriptional activator Sp1. *Mol Cell Biol* **22**, 1277–1287 (2002).

27. Seong, I. S. *et al.* Huntingtin facilitates polycomb repressive complex 2. *Hum Mol Genet* **19**, 573–583 (2010).

28. Nabet, B. *et al.* The dTAG system for immediate and target-specific protein degradation. *Nat Chem Biol* **14**, 431–441 (2018).

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

30. Kremer, B. *et al.* A worldwide study of the Huntington's disease mutation. The sensitivity and specificity of measuring CAG repeats. *N Engl J Med* **330**, 1401–1406 (1994).

31. Kirkwood, S. C., Su, J. L., Conneally, P. & Foroud, T. Progression of symptoms in the early and middle stages of Huntington disease. *Arch Neurol* **58**, 273–278 (2001).

32. Tabrizi, S. J. *et al.* Targeting Huntingtin Expression in Patients with Huntington's Disease. *N Engl J Med* **380**, 2307–2316 (2019).

33. Stanek, L. M. *et al.* Antisense oligonucleotide-mediated correction of transcriptional dysregulation is correlated with behavioral benefits in the YAC128 mouse model of Huntington's disease. *J Huntingtons Dis* **2**, 217–228 (2013).

34. Southwell, A. L. *et al.* Huntingtin suppression restores cognitive function in a mouse model of Huntington's disease. *Sci Transl Med* **10**, (2018).

35. Guo, Q. *et al.* The cryo-electron microscopy structure of huntingtin. *Nature* **555**, 117–120 (2018).

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

37. Yoshimura, S. H. & Hirano, T. HEAT repeats - versatile arrays of amphiphilic helices working in crowded environments? *J Cell Sci* **129**, 3963–3970 (2016).

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

39. Goldberg, Y. P. *et al.* Cleavage of huntingtin by apopain, a proapoptotic cysteine protease, is modulated by the polyglutamine tract. *Nat Genet* **13**, 442–449 (1996).

40. Hermel, E. *et al.* Specific caspase interactions and amplification are involved in selective neuronal vulnerability in Huntington's disease. *Cell Death Differ* **11**, 424–438 (2004).

41. Kim, Y. J. *et al.* 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 (2001).

42. Lunkes, A. *et al.* Proteases acting on mutant huntingtin generate cleaved products that differentially build up cytoplasmic and nuclear inclusions. *Mol Cell* **10**, 259–269 (2002).

43. Miller, J. P. *et al.* Matrix metalloproteinases are modifiers of huntingtin proteolysis and toxicity in Huntington's disease. *Neuron* **67**, 199–212 (2010).

44. Ratovitski, T. *et al.* Mutant huntingtin N-terminal fragments of specific size mediate aggregation and toxicity in neuronal cells. *J Biol Chem* **284**, 10855–10867 (2009).

45. Tebbenkamp, A. T. N. *et al.* Analysis of proteolytic processes and enzymatic activities in the generation of huntingtin n-terminal fragments in an HEK293 cell model. *PLoS One* **7**, e50750 (2012).

46. Warby, S. C. *et al.* Activated caspase-6 and caspase-6-cleaved fragments of huntingtin specifically colocalize in the nucleus. *Hum Mol Genet* **17**, 2390–2404 (2008).

47. Atwal, R. S. *et al.* Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity. *Hum Mol Genet* **16**, 2600–2615 (2007).

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

49. Steffan, J. S. *et al.* SUMO modification of Huntingtin and Huntington's disease pathology. *Science* **304**, 100–104 (2004).

50. Thompson, L. M. *et al.* IKK phosphorylates Huntingtin and targets it for degradation by the proteasome and lysosome. *J Cell Biol* **187**, 1083–1099 (2009).

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

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

53. Caviston, J. P., Ross, J. L., Antony, S. M., Tokito, M. & Holzbaur, E. L. F. Huntingtin facilitates dynein/dynactin-mediated vesicle transport. *Proc Natl Acad Sci U S A* **104**, 10045–10050 (2007).

54. Colin, E. *et al.* Huntingtin phosphorylation acts as a molecular switch for anterograde/retrograde transport in neurons. *EMBO J* **27**, 2124–2134 (2008).

55. Engelender, S. *et al.* Huntingtin-associated protein 1 (HAP1) interacts with the p150Glued subunit of dynactin. *Hum Mol Genet* **6**, 2205–2212 (1997).

56. Gauthier, L. R. *et al.* Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. *Cell* **118**, 127–138 (2004).

57. Gunawardena, S. *et al.* Disruption of axonal transport by loss of huntingtin or expression of pathogenic polyQ proteins in Drosophila. *Neuron* **40**, 25–40 (2003).

58. Li, S. H., Gutekunst, C. A., Hersch, S. M. & Li, X. J. Interaction of huntingtin-associated protein with dynactin P150Glued. *J Neurosci* **18**, 1261–1269 (1998).

59. McGuire, J. R., Rong, J., Li, S.-H. & Li, X.-J. Interaction of Huntingtin-associated protein-1 with kinesin light chain: implications in intracellular trafficking in neurons. *J Biol Chem* **281**, 3552–3559 (2006).

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

61. Twelvetrees, A. E. *et al.* Delivery of GABAARs to synapses is mediated by HAP1-KIF5 and disrupted by mutant huntingtin. *Neuron* **65**, 53–65 (2010).

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

63. Humbert, S. *et al.* The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt. *Dev Cell* **2**, 831–837 (2002).

64. Rangone, H. *et al.* The serum- and glucocorticoid-induced kinase SGK inhibits mutant huntingtin-induced toxicity by phosphorylating serine 421 of huntingtin. *Eur J Neurosci* **19**, 273–279 (2004).

65. Pardo, R. *et al.* Inhibition of calcineurin by FK506 protects against polyglutamine-huntingtin toxicity through an increase of huntingtin phosphorylation at S421. *J Neurosci* **26**, 1635–1645 (2006).

66. Martin, D. D. O., Ladha, S., Ehrnhoefer, D. E. & Hayden, M. R. Autophagy in Huntington disease and huntingtin in autophagy. *Trends Neurosci* **38**, 26–35 (2015).

67. Boutell, J. M. *et al.* Aberrant interactions of transcriptional repressor proteins with the Huntington's disease gene product, huntingtin. *Hum Mol Genet* **8**, 1647–1655 (1999).

68. Heinzel, T. *et al.* A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression. *Nature* **387**, 43–48 (1997).

69. Alland, L. *et al.* Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression. *Nature* **387**, 49–55 (1997).

70. Yohrling, G. J., Farrell, L. A., Hollenberg, A. N. & Cha, J.-H. J. Mutant huntingtin increases nuclear corepressor function and enhances ligand-dependent nuclear hormone receptor activation. *Mol Cell Neurosci* **23**, 28–38 (2003).

71. Futter, M. *et al.* Wild-type but not mutant huntingtin modulates the transcriptional activity of liver X receptors. *J Med Genet* **46**, 438–446 (2009).

72. Cobb, M. 60 years ago, Francis Crick changed the logic of biology. *PLoS Biol* **15**, e2003243 (2017).

73. Brandsma, I. & Gent, D. C. Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome Integr* **3**, 9 (2012).

74. Wittrup, A. & Lieberman, J. Knocking down disease: a progress report on siRNA therapeutics. *Nat Rev Genet* **16**, 543–552 (2015).

75. Nakamura, N. The Role of the Transmembrane RING Finger Proteins in Cellular and Organelle Function. *Membranes (Basel)* **1**, 354–393 (2011).

76. Röth, S., Fulcher, L. J. & Sapkota, G. P. Advances in targeted degradation of endogenous proteins. *Cell Mol Life Sci* **76**, 2761–2777 (2019).

77. Ruzo, A. *et al.* Chromosomal instability during neurogenesis in Huntington's disease. *Development* **145**, (2018).

78. Delli Carri, A. *et al.* Human pluripotent stem cell differentiation into authentic striatal projection neurons. *Stem Cell Rev Rep* **9**, 461–474 (2013).

79. Corbin, J. G., Gaiano, N., Machold, R. P., Langston, A. & Fishell, G. The Gsh2 homeodomain gene controls multiple aspects of telencephalic development. *Development* **127**, 5007–5020 (2000).

80. Toresson, H., Potter, S. S. & Campbell, K. Genetic control of dorsal-ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2. *Development* **127**, 4361–4371 (2000).

81. Yun, K., Potter, S. & Rubenstein, J. L. Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon. *Development* **128**, 193–205 (2001).

82. Duyao, M. P. *et al.* Inactivation of the mouse Huntington's disease gene homolog Hdh. *Science* **269**, 407–410 (1995).

83. Nasir, J. *et al.* Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell* **81**, 811–823 (1995).

84. Zeitlin, S., Liu, J. P., Chapman, D. L., Papaioannou, V. E. & Efstratiadis, A. Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nat Genet* **11**, 155–163 (1995).

85. Woda, J. M. *et al.* Inactivation of the Huntington's disease gene (Hdh) impairs anterior streak formation and early patterning of the mouse embryo. *BMC Dev Biol* **5**, 17 (2005).

86. Evers, M. M. *et al.* Making (anti-) sense out of huntingtin levels in Huntington disease. *Mol Neurodegener* **10**, 21 (2015).

87. Kumar, A. *et al.* Allelic series of Huntington's disease knock-in mice reveals expression discorrelates. *Hum Mol Genet* **25**, 1619–1636 (2016).

88. Pouladi, M. A. *et al.* Full-length huntingtin levels modulate body weight by influencing insulin-like growth factor 1 expression. *Hum Mol Genet* **19**, 1528–1538 (2010).

6 | Appendix

6.1 Summary of first year work

6.1.1 Aims and introduction

In my first year of PhD, I worked on a different project aiming at generating a multicolour fluorescent human embryonic stem cell (hESC) line reporter for some of the key genes for striatal differentiation. The ability to detect the expression of these genes in live, differentiating cells is extremely valuable. In first place, by sorting fluorescent reporter cells it is possible to enrich the population of committed ventral telencephalic progenitors, thus dramatically improving the purity of neural progenitors employed for cell replacement experiments. Moreover, reporter lines constitute a useful tool to monitor cell progression and response to high-throughput morphogens modulation during the differentiation. These data will suggest the best culture conditions to increase the final yield of Medium Spiny Neurons (MSNs).

6.1.2 Introduction

Striatal neurons derived from human pluripotent stem cells (hPSCs) exposed to *in vitro* differentiation represent an amenable cell model of Huntington's Disease. hPSCs are induced to differentiate into ventral telencephalic progenitors through the combination of dual-SMAD inhibition and SHH and WNT modulation. After 50 days of differentiation, these cells eventually differentiate into electro-physiologically active and mature neurons, with 10-15% of them resembling authentic Medium Spiny Neurons (MSNs) identified by the co-expression of DARPP32, CTIP2 and GAD67. To monitor the differentiation process, the progressive expression of specific markers genes is routinely checked. Among these, the activation of specific transcription factors is fundamental for the acquisition of the correct striatal identity. GSX2 is the first striatal marker to be expressed; it is a transcription factor active in the neural progenitors proliferating in the ventral telencephalon. During the formation of the striatum, GSX2 is expressed in the Ventricular Zone (VZ) of the Lateral Ganglionic Eminence (LGE) where it defines (together with the transcription factor PAX6) the pallium-subpallium boundary, thus becoming one of the earliest markers for striatal progenitors. Germ-line inactivation of Gsx2 in mice severely compromises striatal formation and striatal neuron identity. ASCL1 is another transcription factor transiently expressed after and under the control of GSX2. ASCL1 also promotes the biosynthesis of the neurotransmitter GABA, committing the neuronal progenitor to a GABAergic fate. During development, ASCL1 is present in the Ventricular and Subventricular zone (VZ and SVZ). Later in differentiation, mature striatal neurons begin to express CTIP2, a transcription factor that appears at the early post-mitotic stages of MSN maturation in the mantle zone of the LGE. It has a central role in the MSNs maturation and organization and is fundamental for the striosomes-matrix architecture. Another specific marker of the mature MSN identity is DARPP32, a bifunctional signal transduction molecule. Stimulation of the glutamatergic and dopaminergic receptor regulates its phosphorylation and functions as a kinase or phosphatase inhibitor. Our lab has developed an *in vitro* differentiation protocol able to drive hPSCs identity towards MSN fate. To improve this protocol, a better understanding of the molecular dynamics that come into play during differentiation is needed. To this purpose, it would be of great advantage a multicolour reporter cell line for some of the key genes involved in the formation of the striatum, such as GSX2, ASCL1, CITIP2, DARPP32. Each gene can be tagged with a fluorescent reporter thanks to the CRISPR/Cas9 system. The double-strand-brake generated by this system is processed by one of the two main DNA repair pathways: the Homologous Directed

Repair (HDR) and the Non-Homologous End Joining (NHEJ). During the repairing process, these pathways can integrate exogenous DNA sequences, if a suitable donor template is provided. In the first case, homologous sequences flanking the exogenous DNA are required for its integration. The NHEJ instead can integrate any linearized DNA fragment. The recently described Homology Independent Targeted Integration (HITI) approach take advantage of this repair pathway and of the property of the CRISPR/Cas9 system to exert control on the orientation of the integrated DNA. This is achieved by flanking the exogenous DNA sequence with two copy of the gRNA target sequence used to cleave the genome but in reversed orientation. Reporter cell lines generated through either of these approaches can be used to purify the population of "bona fide" striatal progenitor from the heterogeneous population that exist in vitro during differentiation and to monitor in real-time the progression and efficiency of fate determination.

6.1.3 Results

In the first year of my PhD, I started working on the generation of the fluorescent reporter cell lines using CRISPR/Cas9 technology. We decided to target the C-terminus of four different genes involved in striatal differentiation to express a fluorescent protein under the control of the endogenous promoter without disrupting endogenous expression of the locus. Among them, DARPP32 was selected for the first round of editing since previous immunoblot data confirmed its expression already in the hESCs. Thus, facilitating the functional validation step. In this first round of editing, two different approaches were compared: the standard HDR-based and the recently described HITI-based editing.

For the HDR-based approach, a donor plasmid (fig. 1b) carrying the tdTomato fluorescent reporter was designed. This donor features a 2A peptide to separate the tdTomato from the DARPP32 protein, a nuclear localization signal (NLS) to confine its expression in the nucleus and a floxed Balsticidin antibiotic selection. The antibiotic resistance is crucial to enrich HDR-mediated exogenous DNA integration since it is an extremely rare event in the hESCs.

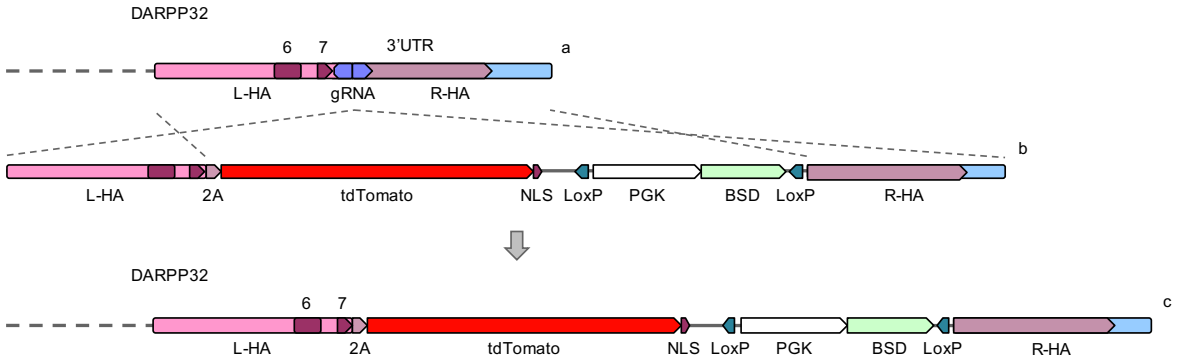

Figure 1 *HDR-mediated editing of DARPP32 C-terminus;* ***a*** *C-terminus of DARPP32;* ***b*** *HDR donor plasmid;* ***c*** *C-terminus of DARPP32 post-editing*

After the screening of best-performing gRNA in HEK293 cells performed using the Surveyor assay (data not shown), H9 (human embryonic stem cells) were edited delivering the Cas9/gRNA as RNP complex together with the donor plasmid as a template by nucleofection. When cells recovered, antibiotic selection was applied to enrich the clones with proper integration. After antibiotic selection, only three clones were obtained, and PCR analysis revealed that two of them carried the recombination cassette correctly integrated at the C-terminus of DARPP32 while the third originated from a random integration event of the entire plasmid conferring resistance to the antibiotic (data not shown). The two positive clones were expanded and used for the functional validation of the reporter.

For the HITI-based approach a donor carrying the EGFP fluorescent reporter was designed (fig 2b). This donor features the IRES sequence to drive a secondary translation of the reporter protein, NLS to confine EGFP expression in the nucleus, a floxed Balsticidin antibiotic selection and two copies of the gRNA target sequence flanking the cassette. These sequences are in opposite orientation compared to the one present in the genome, allowing to control the orientation of the integrated cassette.

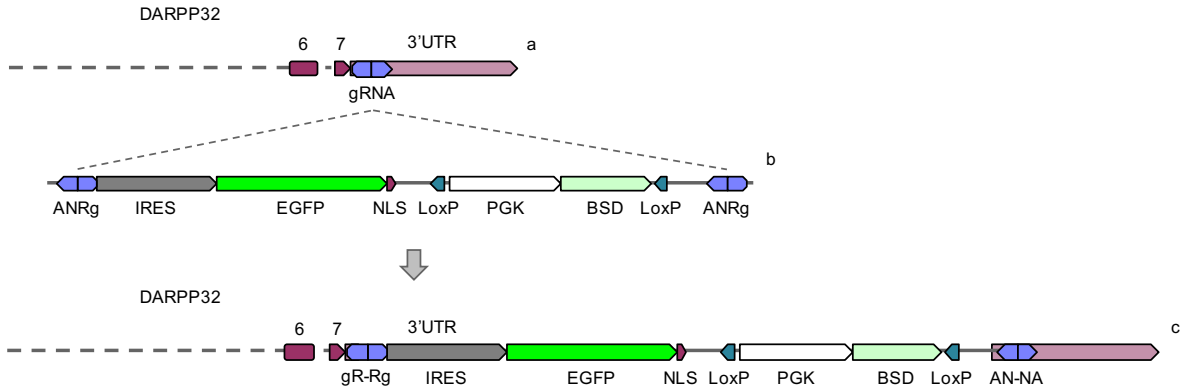

Figure 2 *HITI-mediated editing of DARPP32 C-terminus;* ***a*** *C-terminus of DARPP32;* ***b*** *HITI donor plasmid;* ***c*** *C-terminus of DARPP32 post-editing*

After antibiotic selection, a surprisingly high number of clones was obtained. Among them, 60 were isolated and analysed by PCR. Two PCR reactions were designed to amplify the left and the right junction (LJ and RJ respectively) between the genome and the cassette. PCR analysis performed on each clone revealed that 33.8% of the clones were positive for the LJ-PCR, 61% were positive for the RJ-PCR and 23.7% were positive for both (data not shown). Afterwards, each amplification product was analysed by sanger sequencing to detect InDles frequency. From this analysis emerged that all of the sequenced clones contained InDels and that InDels size and position in the LJ was less variable compared to RJ. Among the clones carrying smaller InDels, four were selected to measure EGFP transcript expression by qPCR. This analysis revealed that the EGFP mRNA was detectable only in 2 out of 4 clones (data not shown). Subsequently, western blot analysis was performed to validate the presence of the reporter protein in the pluripotent clones generated either by HITI (EGFP) or HDR (tdTomato). As expected, DARPP32 protein was detectable in all clones, however, both EGFP and tdTomato were absent (fig. 3b. left panel; data for HITI clones not shown). To address whether this might be due to the negative influence of the antibiotic resistance cassette on the upstream target gene we also treated the two HDR clones with CRE-GFP to remove the antibiotic resistance by CRE-lox recombination. GFP-positive cells were sorted and the excision of the antibiotic resistance cassette was verified by PCR (data not shown). Afterwards, the western blot analysis was repeated on the two HDR-derived clones and on the same clones after excision of the antibiotic resistance. Unfortunately, tdTomato was not detectable even post-excision of the antibiotic resistance gene (fig. 3b right panel).

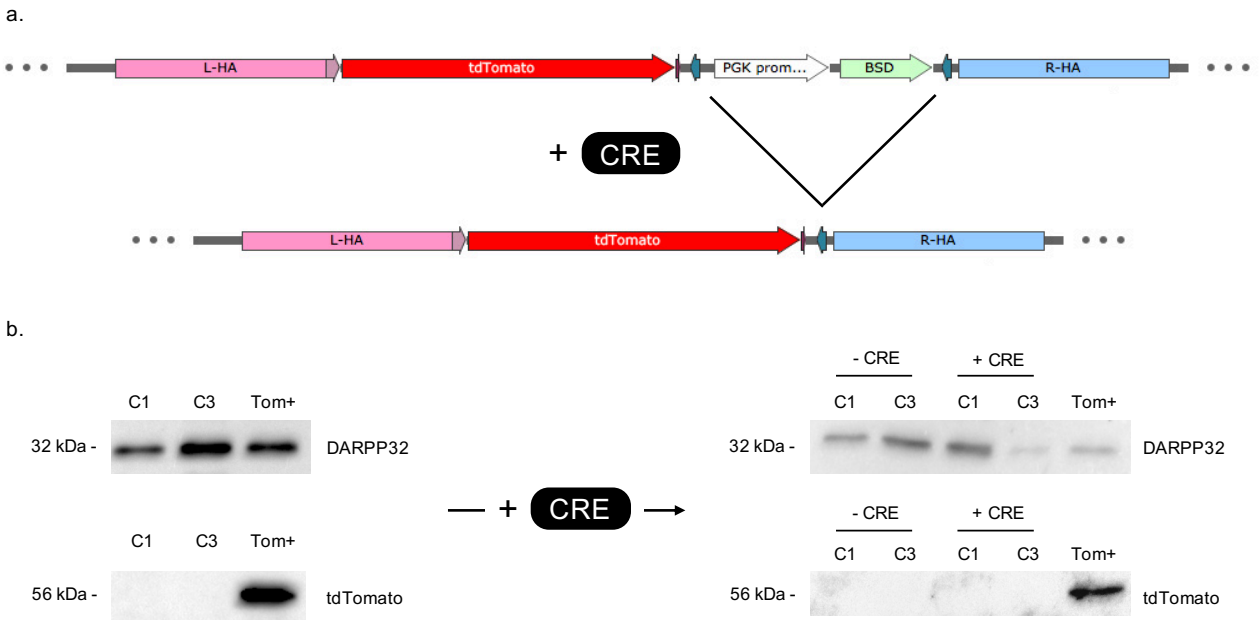

Figure 3 *Western blot analysis of DARPP32-Tom clones;* ***a.*** *Scheme of the edited locus before (top) and after (bottom) CRE recombinase treatment.* ***b.*** *Western blot images of DARPP32 (top) and tdTomato (bottom) in the edited clones and in the positive control (Tom +) before (left) and after (right) CRE recombinase treatment.*

Finally, HDR-derived clone and H9 control cells were subjected to striatal differentiation to address if the lack of reporter expression can be due to the pluripotency epigenetic state or can be linked to the genomic locus. The expression of the target gene and its reporter gene was evaluated by qPCR at different time points during differentiation. As expected, the H9 control progressed correctly toward striatal identity displaying a peak in DARPP32 expression at D44, whereas the edited clone displayed weaker expression of DARPP32 at the same time point, indicating a suboptimal differentiation (fig 4a.). Still, the reporter expression appears to progressively increase as differentiation progresses (fig 4b.). This preliminary data suggest that the lack of reporter expression in pluripotency may be due to a silencing mechanism of the reporter gene linked to that state.

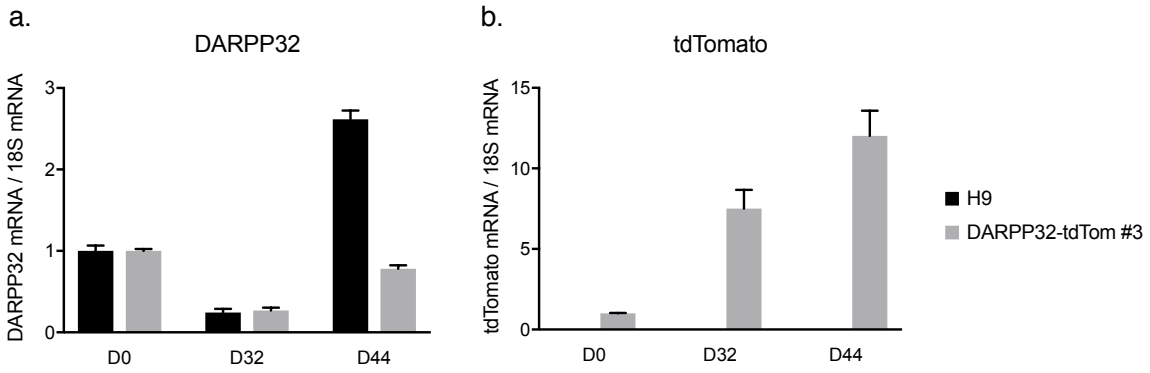

Figure 4 *Differentiation experiment of H9 control vs H9 DARPP32-tdTomato edited clone 3;* ***a*** *DRAPP32 mRNA;* ***b*** *tdTomato mRNA. Data represent the mean of two technical replicates ± SEM.*

6.1.4 Conclusions

The data obtained during the first year of PhD indicate that HITI can be a promising approach for genome engineering due to its high editing efficiency, though InDels at the genome-cassette junctions were observed in all screened clones. This data are in contrast with what reported in the

manuscript published by Belmonte's group in 2016. Still, further experiments are required to assess if this discrepancy is due to suboptimal experimental conditions or to the different cellular systems. Regardless, qPCR analysis performed on the clones obtained by HITI- or HDR-based editing revealed a weak expression of the reporter mRNAs in the pluripotent state. This was associated with the absence of the reporter proteins as observed by western blot analysis. Nevertheless, during *in vitro* differentiation of the HDR-edited clone 3, an increasing trend in the reporter expression along the differentiation was observed, although poor striatal specification was achieved in comparison to the H9 control line. Unfortunately, the amount of cell material was not sufficient to investigate the presence of the reporter protein by western blot analysis.

6.1.5 Change of PhD project

Shortly after the end of the first year of my PhD, additional data emerged on the fluorescent reporter for GSX2, which was also under development. The GSX2 reporter line, exposed to striatal differentiation, displayed few cells detectable by flow cytometry at the stage when GSX2 expression should be present. However, the overall proportion of GFP expressing cells resulted much lower than expected considering the frequency of GSX2 positive cells that we normally detect at this stage (data not shown). Further experiments suggested that this phenomenon seems to be due to epigenetic silencing of the edited allele during the differentiation process. Indeed, chromatin demethylation approaches (such as azacytidine treatments) were able to slightly improve the overall frequency of GFP positive cells (data not shown). In this context, an increasing interest in investigating the immediate consequence of HTT (wt or HD) removal developed in the lab. For these reasons, I decided to change the topic of my PhD project focusing on the development of a cellular system for acute wt or mutant HTT removal.

6.2 Contribution to published articles

A CRISPR-strategy for the generation of a detectable fluorescent hESC reporter line (WAe009-A-37) for the subpallial determinant GSX2.

Besusso D, Cossu A, Mohamed A, Cernigoj M, Codega P, Galimberti M, Campus I, Conforti P, Cattaneo E.

Stem Cell Res. 2020 Oct 1;49:102016. doi: 10.1016/j.scr.2020.102016. Online ahead of print. PMID: 33039807

The aim of this work was to generate by genome editing a GSX2-GFP reporter hESC line. This line offers the possibility to leverage the presence of fluorescent signal to isolate by fluorescent activated cell sorting subpallial-committed progenitors. Therefore allowing to bypass the major limitation of current *in vitro* differentiation protocols, that is their poor efficiency, by enriching for the desired progenitor identity. This expedient can be implemented to enrich in specific subpopulations for cell replacement based therapies. In particular, my contribution to this work was to participate in the generation and screening of the RC17 lines. I also contributed to the dCas9-VPR activation experiments and to the assay to determine GT2A cleavage efficiency (Fig. 1 E-F).

Stem Cell-Derived Human Striatal Progenitors Innervate Striatal Targets and Alleviate Sensorimotor Deficit in a Rat Model of Huntington Disease.

Besusso D, Schellino R, Boido M, Belloli S, Parolisi R, Conforti P, Faedo A, Cernigoj M, Campus I, Laporta A, Bocchi VD, Murtaj V, Parmar M, Spaiardi P, Talpo F, Maniezzi C, Toselli MG, Biella G, Moresco RM, Vercelli A, Buffo A, Cattaneo E.

Stem Cell Reports. 2020 May 12;14(5):876-891. doi: 10.1016/j.stemcr.2020.03.018. Epub 2020 Apr 16. PMID: 32302555

The aim of this work was to set up cell replacement in Huntington's Disease animal models using Stem Cell-derived Human striatal progenitors and to assess the therapeutic efficacy of this approach. To this aim, H9 human embryonic stem cells were exposed to an *in vitro* striatal differentiation procedure, modified from Delli Carri et al. (2013). Ventral telencephalic progenitors were isolated after 20 DIV to be grafted unilaterally in the striatum of quinolinic acid (QA)-lesioned athymic adult rats. The transplanted animals were followed up for 2 months and then sacrificed for neuroanatomical analyses. Grafted cells survival, innervation and functionality was evaluated by immunohistochemistry, rabies virus-mediated synaptic tracing, and ex vivo electrophysiology. Data demonstrate that transplanted cells are able to project to the appropriate target structures, and receive synaptic contact from both host and graft cells. Additionally, significant improvement in sensory-motor tasks were recorded up to 2 months post-transplant. My contribution to this work was to produce the lentiviruses expressing the rabies helper construct under control of the human Synapsin promoter. These viruses were used to infect the cells before transplantation providing the all the elements necessary for the Monosynaptic Tracing (Fig. 4, Supplementary Fig. S4).

Allele-specific silencing as treatment for gene duplication disorders: proof-of-principle in autosomal dominant leukodystrophy.

Giorgio E, Lorenzati M, Rivetti di Val Cervo P, Brussino A, Cernigoj M, Della Sala E, Bartoletti Stella A, Ferrero M, Caiazzo M, Capellari S, Cortelli P, Conti L, Cattaneo E, Buffo A, Brusco A.

Brain. 2019 Jul 1;142(7):1905-1920. doi: 10.1093/brain/awz139. PMID: 31143934

The aim of this work was to implement the allele-specific silencing by RNA interference (ASP-siRNA) strategy to target gene duplication linked disorders to avoid excessive downregulation that can potentially result harmful for the cells. The effectiveness of this approach was demonstrated in the context of Adult-onset autosomal dominant leukodystrophy (ADLD). The disease is caused by a duplication of the laminB1 (LMNB1) gene and lead to demyelination of the central nervous system. The most efficient ASP-siRNAs were tested in ADLD patient-derived fibroblasts. Following ASP-siRNAs treatment, both LMNB1 mRNA and protein levels were restored close to control levels. In addition, ADLD-specific phenotypes were rescued murine oligodendrocytes overexpressing human LMNB1, and neurons directly reprogrammed from patients' fibroblasts. I contributed to this work with the assembly of the constructs expressing ASCL1, BRN2 and two shRNA for REST silencing into a single plasmid (LV-ABR plasmid), with the production of the correspondent lentiviral particles and with their functional validation (Supplementary Fig. 1 A,B,C). LV-ABR was used for the direct reprogramming of primary human fibroblasts from healthy donors and ADLD patients (Fig. 5)