Fall 2015

DNA Repair Deficiency in Huntington's Disease Fibroblasts and Induced Pluripotent Stem Cells

Peter Anthony Mollica
Old Dominion University, pmolli7@aol.com

DNA REPAIR DEFICIENCY IN HUNTINGTON'S DISEASE FIBROBLASTS AND INDUCED PLURIPOTENT STEM CELLS

by

Peter A. Mollica

B.S. December 2011, Old Dominion University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

BIOMEDICAL SCIENCES

OLD DOMINION UNIVERSITY

December 2015

Approved by :

_____

Christopher J. Osgood (Co-Advisor)

_____

Roy C. Ogle (Co-Advisor)

_____

Lesley H. Greene (Member)

ABSTRACT

DNA REPAIR DEFICIENCY IN HUNTINGTON'S DISEASE FIBROBLASTS AND INDUCED PLURIPOTENT STEM CELLS

Peter A. Mollica
Old Dominion University, 2015
Co-Advisors: Dr. Christopher Osgood
Dr. Roy C. Ogle

Mutant huntingtin protein (mhtt)–the protein responsible for cellular dysfunction in Huntington's disease (HD) –is a product of an expanded trinucleotide repeat (TNR) cytosine-adenine-guanine (CAG) sequence in exon 1 of the *huntingtin (HTT)* gene. The pathology of HD has been extensively researched; however, the mechanism by which the disease-causing TNR expansions occur in somatic cells remains elusive. Interestingly, HD has often been referred to a 'DNA repair disease', even though DNA repair dysfunction *in situ* has not been identified. We hypothesized that presence of the mhtt protein affects the expression of DNA repair genes used to address DNA repair, ultimately affecting genome stability, thus providing a possible mechanism for TRN instability. Using quantitative polymerase chain reaction (qPCR) gene arrays for 84 DNA repair genes, we identified 18 DNA repair genes with decreased fold changes between 2- and 3- fold, as well as 11 genes down regulated greater than 3- fold in one HD fibroblast sample relative to a wild-type sample. To ensure our results were not limited to the samples tested, we then increased our number of HD samples and investigated gene expression of *APEX1, BRCA1, RPA1,* and *RPA3* using sensitive TaqMan Gene Expression assays. Further, immunocytochemistry (ICC) analysis validated expression deficiencies at the protein level. These data identify down-regulated genes necessary to maintain stability in the genome of multiple HD affected fibroblast lines. Our data infers that the presence of the toxic mutant huntingtin (mHtt) protein is involved in the DNA repair gene inhibition.

The mutant huntingtin protein (mHtt) produced in HD exhibits a partial gain-of-function in that the hydrophobic, expanded polyglutamine region at the N-termini aggregates to unintended targets, such as transcription factors and histone modifiers. To identify the broad pathway regulating gene expression down-regulation, we investigated epigenetic regulatory mechanisms. Rapid revival of selected DNA repair expression was observed in response to pharmacological hypomethylation treatment, but not to histone modification treatments. This identifies differential methylation patterns occur as a result of mHtt presence. Using capillary electrophoresis fragment analysis to characterize *HTT* TNR gene expansions, our data reveals that intermittent 5-azacytidine treatments induced *HTT* gene stability over 4 population doublings, elucidating methylation patterning involvement in TNR instability.

Furthermore, induced pluripotent stem cells (iPSCs) undergo global epigenetic changes relative to its native cell type that include methylation patterning changes. Upon reprogramming of HD-affected fibroblasts into a pluripotent state we revealed that gene expression was recovered to wild-type levels and was maintained through 20 population doublings. As well, iPSC-HD lines show contraction-biased instability, opposite to expansion-biased instability in native fibroblast cell types. Differentiation of iPSC-HD lines into mesenchymal-like cells (MLCs) further revealed that APEX1 expression remained static, while others retreated to pre-iPSC expression levels. Interestingly, HD-iPSC derived MLCs showed that TNR regions maintained stability in the *HTT* gene pathogenic region, showing no changes in (CAG) repeats. These findings demonstrate that DNA repair gene expression in HD fibroblasts is altered, thus providing insight into the mechanism in which TNR instability persists, ultimately leading to genetic anticipation. This also identifies possible biomarkers that can be used to monitor disease progression and therapeutic treatment success. This study also provides evidence that TNR instability is pharmacologically alterable. This is the first evidence that a DNA repair gene deficiency is present in cells affected by Huntington's disease. More so, our data suggests that mechanisms involved in pluripotency has a protective effect on the pathogenic TNR region of the *HTT* gene.

.

All of my work is dedicated to my wife. She made too many sacrifices on my behalf. Without my wife, Michelle, none of my accomplishments would have been imaginable, much less, achievable. Without her patience and commitment to my dreams, I would be nothing. She sustained me throughout this process and showed exemplary dedication to our relationship and to my completion of graduate school. I love her and I will never be able to fully repay the sacrifices she made.

E hoomau maua kealoha.

ACKNOWLEDGMENTS

Most importantly, I thank God for putting all of these amazing people in my life. I don't deserve the friendships and opportunities He set before me. I would like to thank Dr. Roy Ogle for "adopting" me into his lab in the middle of my project. He had no obligation to allow me to stay a member of his lab and I cannot thank him enough for not canning me. I am also thankful for Dr. Leslie Greene, whom stuck her neck out for me to get me teaching-assistantships to help fund my graduate studies. Initially, the standards she set for me seemed out of reach, but she guided me down the path to reach my goals. Dr. Chris Osgood is a blessing, without whom I would not have been able to navigate the convoluted Ph.D. process. His patience and understanding sustained me in my studies while being able to maintain a healthy family life. I would also like to thank Tammy Subotich. She already knows how vital she has been to my success, and no words I put on paper will ever do that justice. Lastly, I must thank Dr. Robert Bruno and Dr. Patrick Sachs. Without their endless guidance and constant reinforcement, I would surely not be where I am today. I know that being constantly interrupted must have been annoying. I know that being constantly questioned must have been irritating. I know that being constantly hounded with my tedious issues must have been tiresome. The knowledge that they instilled in me is great, and I thank them for that. But mostly, the commitment, dedication, and patience they showed me far surpasses all others. It is rare that you come across people with values like those in the civilian world. They are truly inspiring mentors and real friends. The examples that they have set are far beyond anybody's I have met. They mentored me and coached me throughout my graduate studies, and someday I hope to repay my gratitude. Because of their loyalty and actions they showed me, I would follow them into battle if asked to do so. Until then, I got your six.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

CHAPTER I

INTRODUCTION

Huntington's disease

Huntington's disease (HD) is a rare, neurodegenerative disease that affects 1 in every 30,000 people in the United States. Phenotypic expression of HD is the result of cellular dysfunction, leading to the eventual degeneration of medium spiny neurons of the basal ganglia (Gonitel et al., 2008; Swami et al., 2009). Through the degeneration of striatal neurons in the basal ganglia, affected individuals progressively experience cognitive, emotional and movement impairments (Peyser and Folstein, 1990; Pla et al., 2014). In about 70% of affected individual's onset of symptoms varies based of the size of the TNR expansion within the huntingtin (*HTT*) gene. HD is an autosomal-dominant inherited disease, with individuals who inherit HD generally not presenting with symptoms until the ages of 30 to 50, often after having children of their own.

The first accurate medical description on HD was made in 1872, by an American doctor named George Huntington. Prior to his accurate description, HD was referred to as *chorea*, describing the involuntary dancing propensities of affected individuals. Later, this definition would be expanded to include jerky involuntary movements of the hips, face, and shoulders. Clinically, this term is still predominately used when describing the "dance-like" movements created by patients affected by neurological disorders. Dr. Huntington's unreferenced paper called *On Chorea*, was initially published in Medical and Surgical Reporter of Philadelphia in 1872. His report addressed and highlighted three main peculiarities. He reported that (1) the disease is hereditary, (2) that affected individuals have suicidal and cognition impairments, and (3) that it manifests only in adult life (Huntington, 2003). He hypothesized that chorea, and HD by

---

The journal format for this dissertation is modeled after Neurobiology of Disease

extension, was the result of derangement of the cerebellum, only knowing at that time the involvement of the cerebellum in governing and controlling motor control. After the discovery of the HD-causing gene in 1993, it seemed that a cure would be just around the corner. Twenty-two years later, there is still no cure for HD. At best, pharmacological intervention is initiated upon symptom presentation to alleviate depression and chorea, but is eventually abandoned as the disease progresses and symptoms become unmanageable (Holl et al., 2010; Ondo et al., 2002; Reilmann, 2013; Reilmann et al., 2015; Sagredo et al., 2012; Yavuz et al., 2013). Some research is focusing their efforts toward inhibiting expression of the affected allele, utilizing gene silencing technologies such RNA interference (Cambon and Deglon, 2013; Lombardi et al., 2009; van Bilsen et al., 2008; Watson and Wood, 2012; Zhang et al., 2009). Multiple studies have shown that silencing of the pathogenic mutant huntingtin (*mHTT)* gene restores normal function, and HD-aggregate inclusions disappear (Yamamoto et al., 2000). At the present time, the most promising treatment is cell replacement therapies. Using animal models, replacement of striatal tissue with fetal striatal neural cells has been successful, promoting recovery (Kendall et al., 1998; Palfi et al., 1998). With the recent discovery of reprogramming patient derived, terminally differentiated cells into induced pluripotent stem cells (iPSCs), much attention is focusing on cellular replacement using patient-specific cells (Benraiss and Goldman, 2011; Choi et al., 2014; Crane et al., 2014; Fink et al., 2014; Yu et al., 2007).

A novel characteristic of HD is the genetic anticipation experienced by successive generations who inherit the mutated allele (Langbehn et al., 2004; Langbehn et al., 2010; Ranen et al., 1995). Genetic anticipation is the phenomena by which a diseases' symptoms manifest earlier and more severely in progeny who successively inherit the disease, relative to the affected transmitting parent (Rieger et al., 1991). While this phenomena is primarily associated with neurological and TNR disorders, it has also been reported in other diseases and disorders such as Lynch syndrome, breast cancer, autism, type 2 diabetes mellitus and Graves' disease (Brix et al., 2003; Martinez-Delgado et al., 2011; Peixoto et al., 2006; Rai and Sharif, 2014; Segui et al., 2013; Stupart et al., 2014; Westphalen et al., 2005;

Yaturu et al., 2005). In relation to HD, it is well established that the severity of oms is linearly correlated, with age-at-onset (AAO) experienced being inversely correlated to the size of the (CAG) TNR expansion in the *HTT* gene (Penney et al., 1997; Wexler et al., 2004) (Fig. 1). It has been shown that increased expansions in successive generation's correlates with more severe symptomatic expression and earlier age of onset in > 70% of individuals affected by HD (Lee et al., 2012; Walker, 2007). Interestingly data also reveals that paternal *vs* maternal transmission of the affected allele contributes significantly to TNR length of offspring, with paternal transmission generating larger expansions in offspring (Ranen et al., 1995). Furthermore, data suggests that age of transmission does not affect (CAG) inheritance, but (CAG) repeat length changes are influenced by the paternal (CAG) TNR length (Wheeler et al., 2007). Conversely, the same study determined that there was greater mean (CAG) variability between affected siblings when the disease was transmitted maternally. Studies reveal that relatively larger expansions of the pathogenic *HTT* gene are often present in progeny upon HD inheritance (Ranen et al., 1995; Swami et al., 2009; Wheeler et al., 2007), with the largest rate of expansion occurring early in the disease course (Kennedy et al., 2003).

HD is a dynamic mutation disease, where the number of the pathogenic (CAG) repeats increases the probability of expression of the mutant phenotype. People with fewer than 35 (CAG) repeats are usually not affected by the disease, or symptoms present late in life. Individuals with 35 to 40 repeats are at an increased risk of developing HD. Expansions greater than 40 repeats inevitably lead to HD progression (Bates, 2003).

Symptomatic manifestation of HD is the result of neurodegeneration of striatal neurons in the basal ganglia, specifically medium spiny neurons (MSNs). The basal ganglia is a complex subcortical region in the base of the forebrain. This portion of the brain is a collection of subcortical nuclei that exhibits inhibitory and excitatory output, being involved in multiple process, such as voluntary movement, cognition, and emotion. Anatomically, the basal ganglia is comprised of the striatum (dorsal and ventral), globus pallidus, ventral pallidum, subthalamic

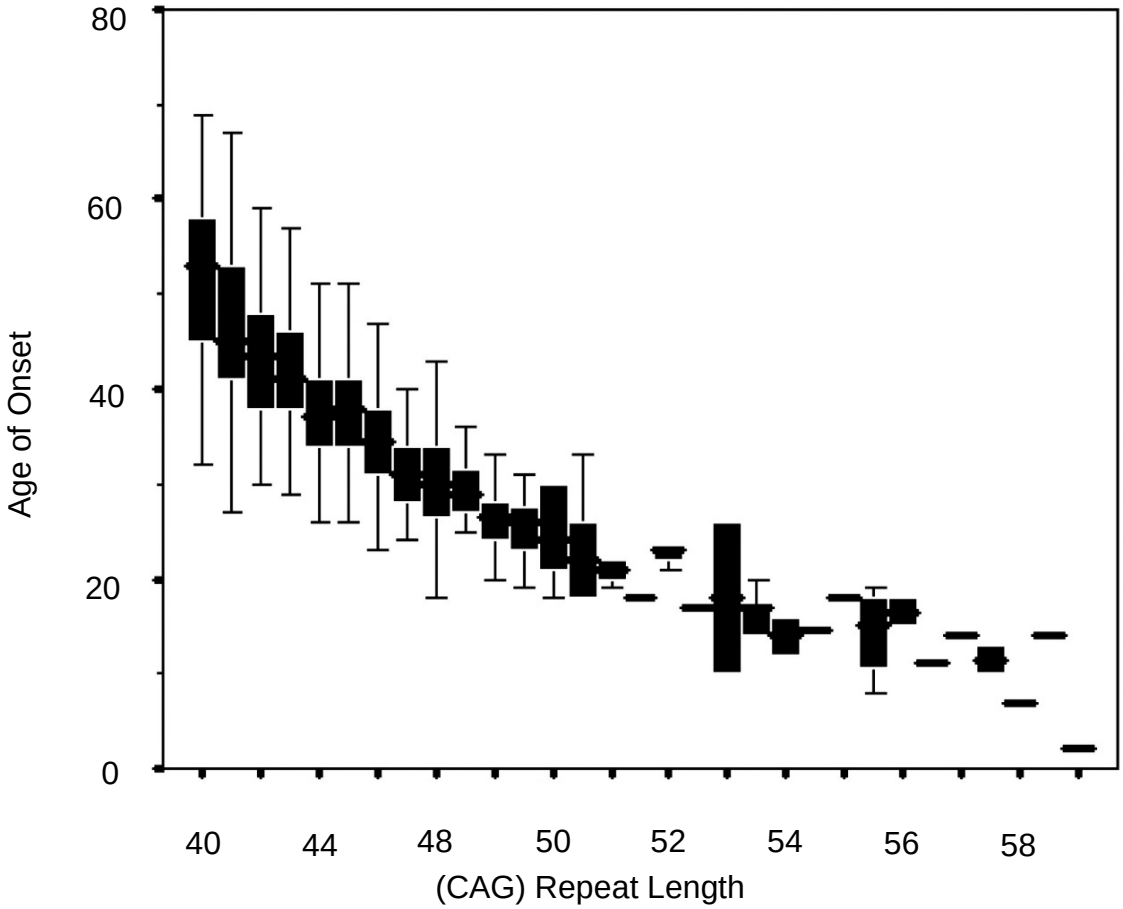


Fig. 1. Box plot of age of onset and repeat length of the longer *HTT* allele. Huntington's disease exhibits genetic anticipation, in which affected individuals present with clinical symptoms earlier in life inversely related to the length of the (CAG) expansion in the *HTT* gene. Figure taken from (Wexler, 2004). Copyright (2004) National Academy of Sciences, U.S.A.

nucleus, and nigra substantia. Functionally, the basal ganglia serves as a "switchboard" in which input is properly routed and regulated. It is this region of the brain that is predominantly affected by HD, and because of its' extensive involvement in coordinated motor control, the degeneration of the basal ganglia causes symptomatic chorea and cognition impairment.

Wild-type and mutant huntingtin protein

The *huntingtin* (*HTT*) gene is located on chromosome 4p16 (Fig. 2). The gene encodes 67 exons producing the huntingtin (Htt) protein, which is ~350kDa (Reiner et al., 2011). The huntingtin protein is ubiquitously expressed throughout tissues types, and is essential for embryonic development, with the most prevalent expression in neuronal tissues (DiFiglia et al., 1995; Duyao et al., 1995; Ferrante et al., 1997; Fusco et al., 1999; Landwehrmeyer et al., 1995; Nasir et al., 1995; Trottier et al., 1995; Velier et al., 1998; Zeitlin et al., 1995). The human *HTT* gene contains trinucleotide repeats (TNR) of cytosine-adenine-guanine (CAG), the codon that encodes the amino acid glutamine. Repetitive repeats of DNA sequences are prone to instability, as witnessed in HD. The disease causing variable region of the protein is located at the N-termini in exon 1. Unaffected individuals have (CAG) repeats between 6 and 35. There are some case where patients with 27 – 35 (CAG) repeats have symptoms manifest with reduced penetrance late in life. Individuals with (CAG) repeats > 36 create a mutant huntingtin (mHtt) protein, and will eventually develop symptoms.

The poly-glutamine (polyQ) amino acid chain in the mutant protein product is known to be responsible for disease manifestation. Due to this characteristic, HD is categorized as a polyglutamine disease. At the present time, nine polyglutamine diseases have been identified. The polyglutamine diseases are Huntington's disease, Dentatorubropallidoluysian atrophy, Spinal and bulbar muscular atrophy, and Spinocerebellar ataxia types 1, 2, 3 (Joseph Machado disease), 6, 7, and 17 (Devys et al., 2001; Masino and Pastore, 2002; Orr, 2001; Saudou et al., 1996; Zoghbi and Orr, 2000). The aggregate-prone protein products created in these diseases are found predominantly in degenerating neurons, such

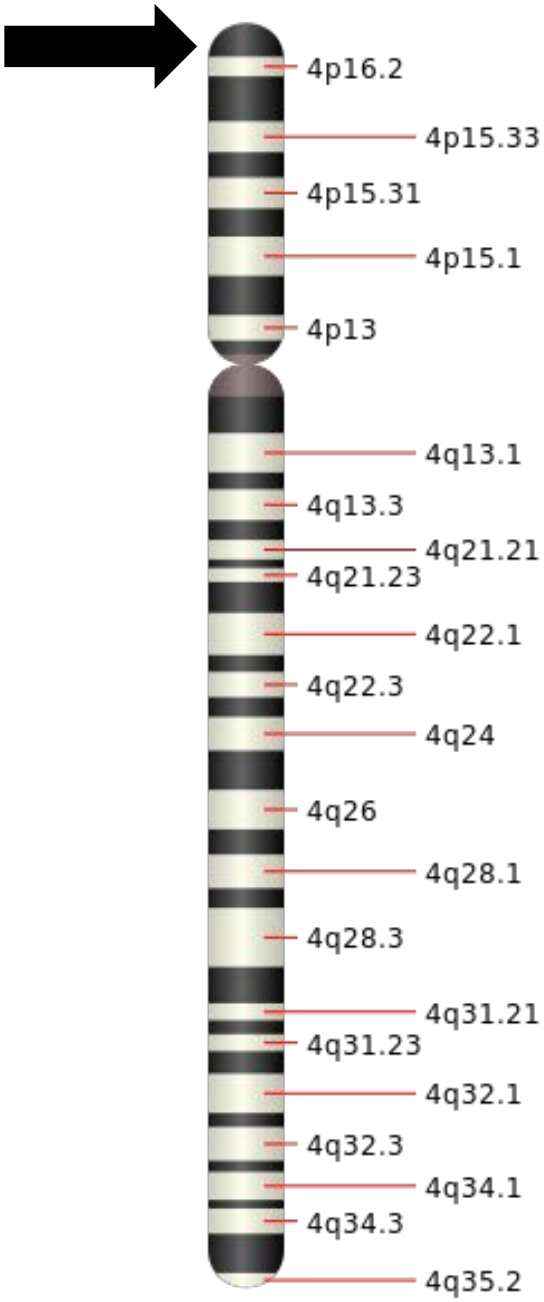


Fig. 2. Location of *HTT* gene in the genome. The *huntingtin* gene is located near the telomeric region of the p-arm (short) of chromosome, indicated by the arrow. Image obtained through Universal Protein Resource (UniProt).

as the spinal cord, midbrain, and cerebellum (Zoghbi and Orr, 1999). The underlying cause of these diseases is similar in pathogenesis, leading to similar, but not identical, symptomatic presentation and progression. Unfortunately, at present time, there is no cure for diseases in this classification, which emphasizes the necessity for continued research.

The structure of the entire Htt protein remains unsolved. Interestingly, this is primarily due not only to its large size, but also due to the pathogenic polyQ region in the N-termini. It has recently been discovered that the polyQ region in Htt proteins have “conformational flexibility”. In other words, the polyQ regions will change its structure, indicated by the Htt17Q trimer structure that has been resolved (Fig. 3) (Kim et al., 2009). In the *HTT* gene, the 3’ side of the pathogenic TNR region is immediately flanked by (CCG)’s repeats, encoding a polyproline (polyP) region. Interestingly, the presence of the polyP region has been shown to decrease the aggregation potential of disease-causing polyQ regions (Bhattacharyya et al., 2006; Dehay and Bertolotti, 2006). Though not fully understood, the absence of the polyP sequence enhances mHtt aggregation. Currently, the three proposed mechanisms addressing polyP region involvement with differential aggregation characteristics are; (1) the presence of the polyP region alters the protein conformation, reducing spatial availability for increased aggregation, (2) the polyP sequence recruits interactors that shield the toxic effects of mHtt, or (3) the polyP residues may solubilize the polyQ region, assisting in its’ stability (Steffan et al., 2004) (Fig. 4). This evidence further complicates HD, as well wild-type patients with large repeats that do not yet constitute HD. It could be hypothesized that the absence of the polyP sequence in a wild-type patient, with (CAG) repeats < 35, could enhance Htt aggregation, which would initiate symptomatic progression.

Downstream of the polyQ and polyP regions of the protein lie multiple consensus sequences called huntingtin, elongation factor3, protein phosphatase 2A, and TOR 1 (HEAT) domains. These HEAT domains facilitate protein-protein interactions (Li et al., 2006; Takano and Gusella, 2002). It has also been shown

that Htt contains conserved amino acid sequences that target it for nuclear export and localization (Bessert et al., 1995; Maiuri et al., 2013; Trottier et al., 1995; Xia et al., 2003; Zheng et al., 2013). It has also been reported that polyQ expansions impair the nuclear export and localization properties (Cornett et al., 2005). These data together reveal the important role of Htt, and its' ability to chaperone proteins thorough the nuclear envelope. As well, it is worth emphasizing that the presence of expanded polyQ regions impair mHtt exportation, thus localizing the toxic mHtt protein within the nucleus. Furthermore, data shows that caspase activity on mHtt results in cleavage of the toxic polyQ-containing N-termini (Li and Li, 2004). This cleavage results in uninhibited aggregate formation, further compromising cellular integrity.

The exact function of wild-type Htt is unknown, but has been shown to be essential for embryonic development, as well as involved in neurotransmission, mitotic spindle formation, apoptosis, and axonal transport (Godin et al., 2010; Goehler et al., 2004; Harjes and Wanker, 2003; Horn et al., 2006; Saudou et al., 1998; Waelter et al., 2001). The expanded polyQ region of mHtt exhibits a toxic, partial gain-of-function. This is evident in the fact that silencing or inactivation of the mHtt gradually allows neuronal recovery (van Bilsen et al., 2008; Yamamoto et al., 2000; Zhang et al., 2009). Importantly, there is data suggesting that the loss of wild-type Htt function can sensitize neuronal cells, amplifying the toxic effects of mHtt present (Cattaneo et al., 2005; Zuccato and Cattaneo, 2007; Zuccato and Cattaneo, 2009). It is evident that disease progression in HD is from the combinatory presence of mHtt, with the loss of Htt.

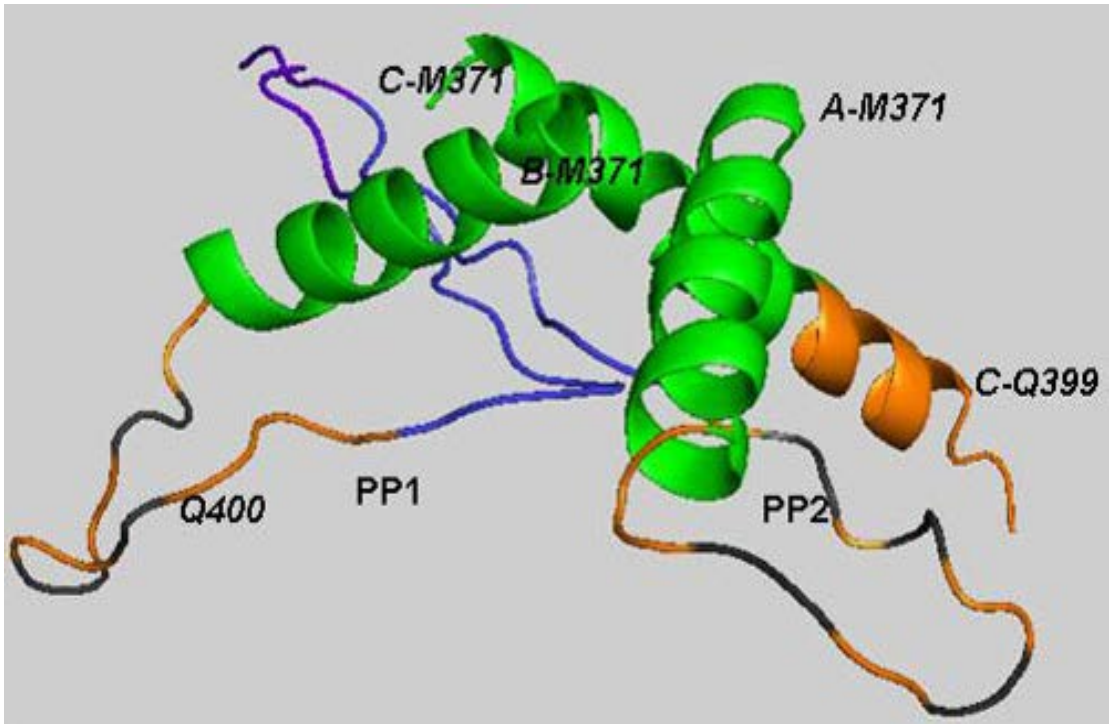


Fig. 3. Resolved structure of the N-termini Htt17Q-trimer. The trimer above represents the conformational flexibility of polyQ sequences. The green and orange segments between structures are polyQ regions, followed by the polyP regions indicated by the abbreviation “PP1, PP2”. Figure taken from (Kim et al., 2009). Copyright Elsevier, U.S.A.

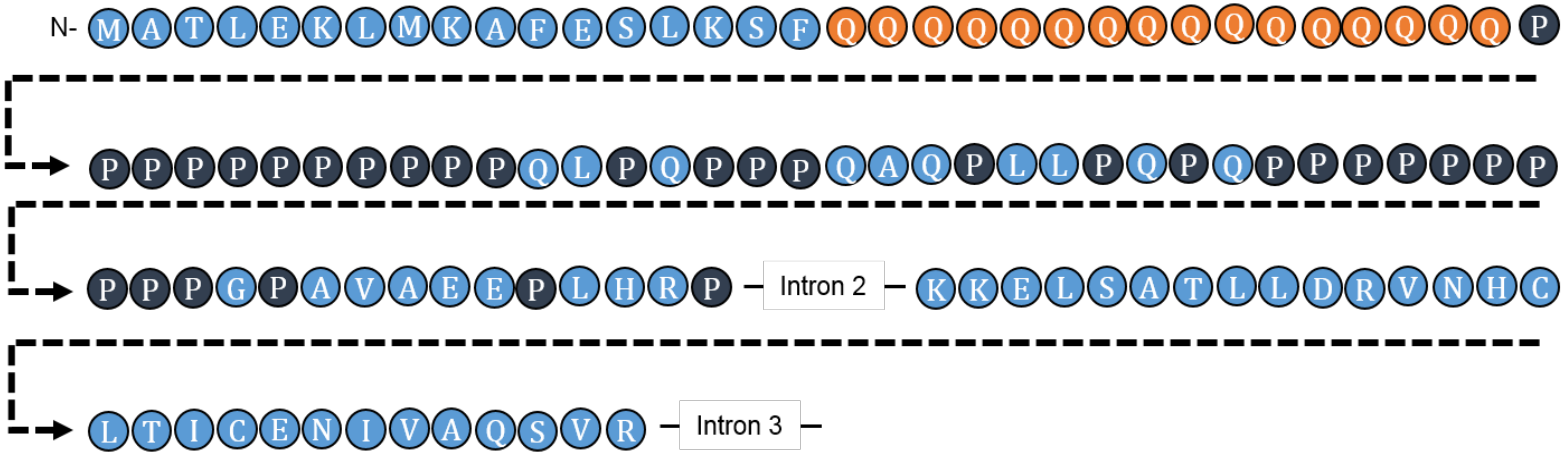


Fig. 4. Amino acid sequence of Htt N-termini showing interruption of intron 2. The depiction identifies the polyQ and polyP regions of the Htt protein in exon 1.

With the discovery that polyQ residues form polar zipper structures, insight was gained that Htt function involved binding of transcription factors that also contained polyQ sequences. Both the wild-type and mutated form of Htt can interact with transcription factors. One of the first molecular observation in HD was the differentially down-regulation of neurotransmitter genes, specifically enkephalin and substance P mRNA (Augood et al., 1996). This finding was followed up with similar gene expression investigations, leading to the hypothesis that mHtt was affecting gene transcription. Throughout the course of HD research, it has been confirmed that mHtt aggregation of unintended targets, combined with decreases in Htt presence, causing differential gene down-regulation, cell toxicity, dysfunctional neurotransmission and cellular dysfunction (Lipinski and Yuan, 2004; Nucifora et al., 2001; Shimohata et al., 2000a; Shimohata et al., 2000b). Even currently, researchers are still elucidating the severity of gene dysregulation in HD.

One of the most researched gene dysregulations caused by mHtt is that of brain derived neurotrophic factor (BDNF). BDNF is a protein that supports the survival of existing neurons and facilitates in the differentiation of neuronal cells, as well as being involved in the formation of newly formed synapses (Acheson et al., 1995; Huang and Reichardt, 2001). The interference on BDNF normal regulation and function is two-fold in the presence of mHtt. First, wild-type Htt seems to be a regulator of BDNF, and the presence of mHtt showed decreased gene expression (Zuccato et al., 2001). Secondly, decreased levels of Htt reduces the huntingtin-mediated transport of the BDNF protein along microtubules (Gauthier et al., 2004). This is an example of just one dysfunction gene affected by mHtt. These two data alone address the extensive repercussions of the presence of mHtt (Fig. 5).

The presence of the expanded mHtt also inhibits efficiency of normal neurotransmitter endocytosis. The wild-type Htt protein forms a complex with huntingtin interacting protein 1 (Hip1), adaptor-related protein complex-2 mu-1 subunit (AP2), and clathrin (CLTA) which is involved in receptor-mediated endocytosis (Singaraja et al., 2002; Waelter et al., 2001). Reductions in the Htt,

decrease the availability to required machinery to maintain appropriate synaptic transmission. Also, the presence of mHtt increases the degradation of AP2 subunits involved in the endocytosis complex (Borgonovo et al., 2013). Exocytosis is also impaired with the presence of mHtt through the depletion of complex II (CPLX2), as well as through the interaction of β-tubulin (TUBB) (Edwardson et al., 2003; Smith et al., 2009).

Furthermore, Htt has been linked to apoptotic regulation. It has been shown that defects in calcium (Ca^{2+}) regulation is a hallmark in HD. mHtt sensitizes inositol 1,4,5-trisphosphate receptor, type 1 (ITPR1) and also stimulates N-methyl-D-aspartate receptor subunit NR1 (NMDAR1), both of which increase Ca^{2+} levels in the cytoplasm and mitochondria (Ciammola et al., 2006; Evert et al., 2000; Li et al., 2000; Liu, 1998; Raymond et al., 2011; Rigamonti et al., 2000; Suzuki et al., 2012). This accumulation of Ca^{2+} within the mitochondria initiates the release of caspase cofactors inducing apoptosis. The involvement of mHtt on the mitochondrial also increases oxidative stress exhibited by HD patients. Post-mortem analysis of HD brains reveals elevated levels of lactate, inferring two possible causes (Jenkins et al., 1993; Koroshetz et al., 1997). The first possibility would implicate that oxygen delivery is inadequate, slowing mitochondrial metabolism resulting in anaerobic respiration. The second possible cause would involve some disorder resulting in increased carbohydrate metabolism where lactic acid formation is favored. Further, some studies in HD- affected mouse models have shown activity decreases in mitochondria complex I, while other complex activity remained unchanged (Parker et al., 1990). Human studies show decreased activity in complex II and complex III, with no changes in complex I and IV (Butterworth et al., 1985; Mann et al., 1990). Even though the mechanism by which mHtt is altering the activity of mitochondrial function, it is evident that the abnormal protein interactions lead to an increase in oxidative stress, resulting in neuronal cell death.

One of the most researched pathogenic consequences of the mHtt protein is the abnormal aggregation patterns seen with transcriptional factors that are required for basal transcription initiation. mRNA expression abnormalities are

shown to arise in part because mHtt products hijack and aggregate important transcriptional regulators and histone acetyl transferases including nuclear protein p53, cyclic AMP-response element (CRE)-binding protein (CREB), specific protein-1 (Sp1), CREB-binding protein (CREB-BP/CBP), TATA-box-binding protein (TBP), and TATA binding protein (TBP)-associated factor ($TAF_{II}130$) (Cha, 2007; Gil and Rego, 2008; Jiang et al., 2006; McCampbell and Fischbeck, 2001). Importantly, this aggregation causes vital deficiencies in HD. This transcriptional inhibition causes decreased expression of a myriad of genes (Dunah et al., 2002; Lipinski and Yuan, 2004; Shimohata et al., 2000a). Interestingly, these transcriptional regulators have been shown to contain polyQ amino acid sequences, for protein-DNA interactions. It is these binding domains that are inadvertently targeted by mHtt. This unintended aggregation disrupts normal gene transcription and is ultimately responsible for some differential gene expression patterns in HD.

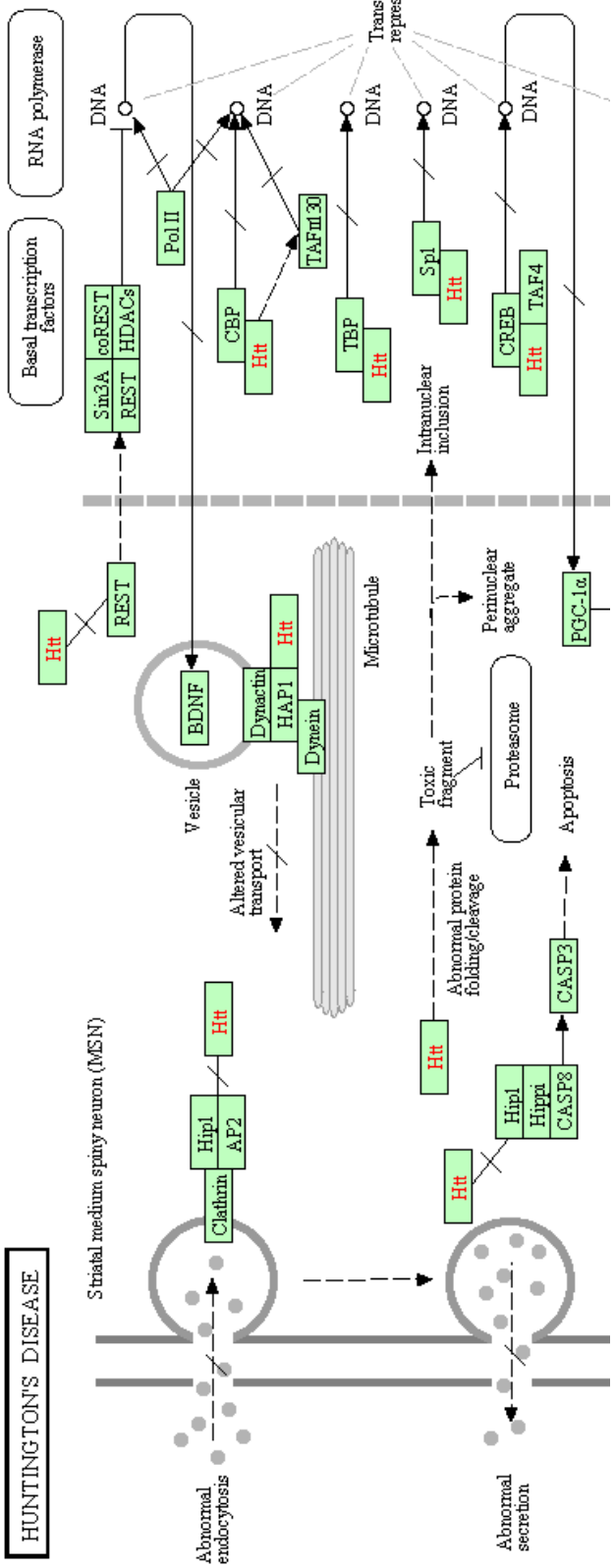


Fig. 5. KEGG Pathway for mutant huntingtin protein cellular interactions. The Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to display known interactions of mHtt in mammalian cells, specifically striatal medium spiny neurons. The above figure displays the multi-tiered dysfunction caused by both the presence of mHtt and the reduction of wild-type Htt proteins. Figure found at *http://www.genome.jp/kegg-bin/show_pathway?hsa05016*.

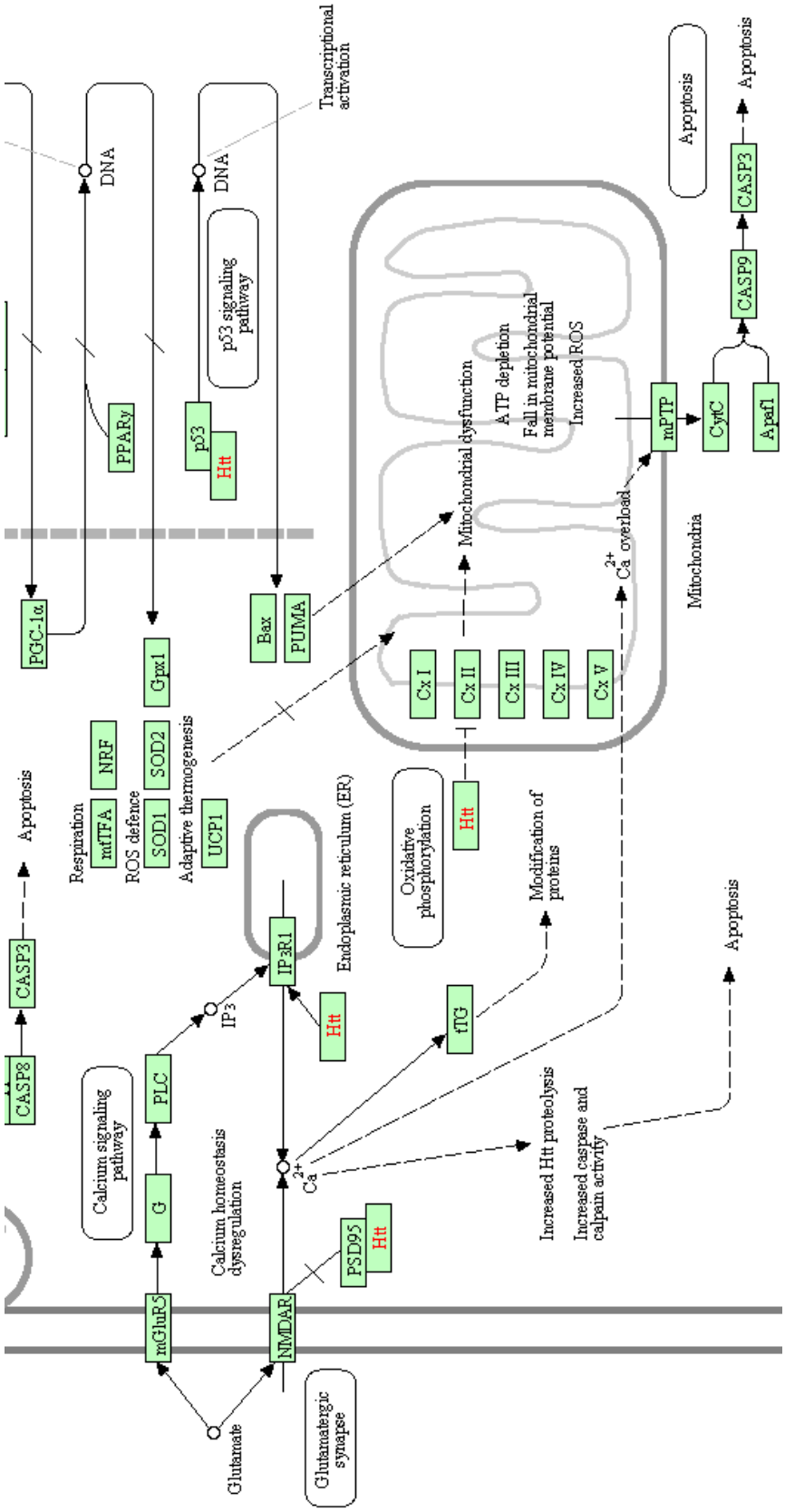


Fig. 5. Continued.

Trinucleotide repeat instability

As mentioned earlier, a key hallmark of HD is the genetic anticipation experienced. Through the expansion of the pathogenic TNR region of the *HTT* gene during transmission, affected individuals experience increased severity of symptoms, while progression occurs earlier in life (Aziz et al., 2012; Langbehn et al., 2004; Langbehn et al., 2010; Norremolle et al., 1995; Ranen et al., 1995; Swami et al., 2009). The cause of genetic anticipation comes from the likelihood that the tandem repeats in the *HTT* gene will expand. Trinucleotide repeat instability is the propensity for a dynamic mutation of repetitive bases in the genome to increase or decrease in size with a probability that either will cause phenotypic changes. Fundamentally, this is the pathogenic basis for the HD. Diseases prone to TNR instability are referred to as *trinucleotide repeat disorders*, and include polyglutamine and non-polyglutamine diseases, with more than half being polyglutamine in pathogenetic in origin. The first of these diseases to be identified was Fragile X syndrome. The diseases characterized by TNR instability have pathogenic tandem repeats ranging from 21 for Spinocerebellar ataxia type 6, to more than 250 in HD (Table 1).

Aside from germ line inheritance and transmission, the (CAG) repeat in HD exhibits somatic instability and undergoes progressive expansion. This characteristic further complicates the identification of the comprehensive model underlying HD. Tissue specific somatic (CAG) instability is present throughout, with the gene showing moderate instability in the liver and kidney, and low instability muscle and heart tissue (Telenius et al., 1994). Interestingly, the greatest expansions are seen in the testis and brain (De Rooij et al., 1995; Telenius et al., 1994). This mosaicism seen in tissues has been linked to new HD mutations arising from previously “unaffected” patients initially exhibiting (CAG) repeats of 33 and 35, ultimately expanding to 45 repeats (Cannella et al., 2005). It is worth noting that *HTT* mosaicism in the brain has been linked to an earlier age of onset, but it is not known how this affects pathogenesis (Reiner et al., 2011). To confound the disease further, researchers must first identify if tissue specific (CAG) repeat

instability is the result of the same mechanism, or if different mechanisms in different tissues are responsible for this mosaicism.

Table 1

Trinucleotide repeat disorders and pathogenic characteristics.

Disorder	Pathogenic Gene	Tandem Repeat Sequence	Normal Repeats	Pathogenic Repeats
DRPLA (Dentatorubropallidoluysian Atrophy)	*ATN1 or DRPLA*	CAG	6-35	49-88
HD (Huntington's disease)	*HTT*	CAG	6-35	36-250
SBMA (Spinal and bulbar muscular atrophy)	*AR*	CAG	9-36	38-62
SCA1 (Spinocerebellar ataxia Type 1)	*ATXN1*	CAG	6-35	49-88
SCA2 (Spinocerebellar ataxia Type 2)	*ATXN2*	CAG	14-32	33-77
SCA3 (Spinocerebellar ataxia Type 3 or Machado-Joseph disease)	*ATXN3*	CAG	12-40	55-86
SCA6 (Spinocerebellar ataxia Type 6)	*CACNA1A*	CAG	4-18	21-30
SCA7 (Spinocerebellar ataxia Type 7)	*ATXN7*	CAG	7-17	38-120
SCA17 (Spinocerebellar ataxia Type 17)	*TBP*	CAG	25-42	47-63
FRAXA (Fragile X syndrome) [α]	*FMR1*	CGG	6 - 53	> 230
FXTAS (Fragile X-associated tremor/ataxia syndrome) [α]	*FMR1*	CGG	6 - 53	55-200
FRAXE (Fragile XE mental retardation) [α]	*AFF2* or *FMR2*	CCG	6 - 35	> 200
FRDA (Friedreich's ataxia)	*FXN* or *X25*	GAA	7 - 34	> 100
DM (Myotonic dystrophy)	*DMPK*	CTG	5 - 37	> 50
SCA8 (Spinocerebellar ataxia Type 8)	*OSCA* or *SCA8*	CTG	16 - 37	110 - 250
SCA12 (Spinocerebellar ataxia Type 12)	*PPP2R2B* or *SCA12*	nnn On 5' end	7 - 28	66 - 78

[α]These syndromes are generally characterized in the clinical setting and referred to as Fragile X-associated disorders (FXD).

There a many proposed mechanisms addressing TNR instability that are mediated by DNA repair, replication, homologous recombination, and transcription. Every process in which the DNA double-helix is denatured has shown to be affected by trinucleotide repeat instability, but the greatest instability has been seen during replication and recombination (Cleary and Pearson, 2003; Lenzmeier and Freudenreich, 2003). This implies that DNA replication is the initiating event in TNR instability. During DNA replication and metabolic processes, the native DNA double-helix is separated to allow the binding of polymerases, transcriptional regulators, stabilizing proteins, and other DNA-binding motif proteins. Through slippage of single-stranded DNA (ssDNA), intermediate secondary structures are created, also known as TNR hairpins or bulky adducts (Kang et al., 1995; Salinas-Rios et al., 2011; Sinden, 2001; Wells et al., 1998). These hairpins exhibit increased stability through Watson-Crick complementarity of cytosine-guanine base pairing. The subsequent processing of replication and ligation will insert newly synthesized tandem repeats, or with a lower probability, delete tandem repeats. The insertion or deletion of newly synthesized tandem repeats is dependent on which strand is undergoes slippage. When slippage occurs on the template strand, newly synthesized DNA will have deleted tandem repeats. Likewise, if slippage occurs on the nascent DNA strand, the subsequent replication products will result in an increase in tandem repeat sequences (Fig. 6). It has been shown that there is greater probability for expansions to occur in TNR regions. It has also been shown that multiple, shorter hairpin formations are more favored, versus fewer, longer hairpins (Avila-Figueroa et al., 2011; Gacy et al., 1995; Rolfsmeier et al., 2001; Yu et al., 1995; Zacharias, 2001). Even though specialized DNA repair mechanisms have evolved to address these specific bulky hairpins, it has been shown that DNA repair acts as both positive and negative effectors of TNR instability (Cleaver et al., 2009; Kovtun et al., 2007; Lin and Wilson, 2012; Liu and Wilson, 2012; Panigrahi et al., 2010; Pinto et al., 2013). Interestingly, the propensity favoring smaller, more frequent hairpins increases the probability of multiple hairpins interacting with each other, known as duplex formation, which ultimately decreases the efficiency of DNA repair

processes (Goula et al., 2012; Lai et al., 2013; Liu et al., 2013; Qiu et al., 2015; Xu et al., 2014). The increase in bulky, localized hairpins is more difficult to resolve than fewer hairpins distributed throughout the strand. Devoid of DNA replication, ironically the largest TNR expansions are detected in mitotically inactive cells (Kennedy et al., 2003; Kennedy and Shelbourne, 2000; Shelbourne et al., 2007). Contrary to logic, the degree of instability and proliferation rate of tissues does not seem to be linked, where slowly proliferating liver tissue show moderate instability and mitotically inactive striatal tissues show high instability (Fortune et al., 2000; Gomes-Pereira et al., 2001; Kennedy and Shelbourne, 2000; Larson et al., 2015; Lia et al., 1998; Mangiarini et al., 1997; Wheeler et al., 1999). With no data showing a direct correlation to tandem repeat instability and proliferation rates, it is concluded that instability occurs by a mechanism(s) independent of replication mediation. As well, the instability shown in arrested cells implicate the involvement of DNA damage repair mechanisms.

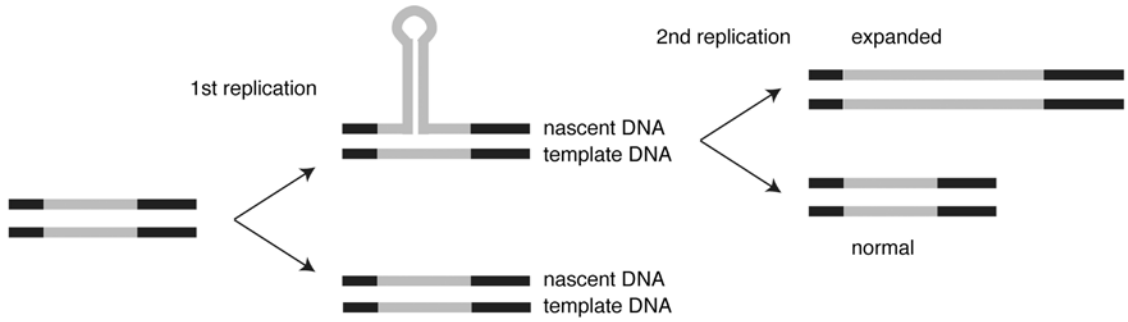
DNA replication and metabolic processes alone did not address tissue-specific (CAG) mosaicism seen in mitotically inactive neuronal cells. During mitotic arrest, the absence of a sister chromatid/chromosome eliminates the likelihood that TNR instability is occurring during homologous recombination. This rules out the chances that striatal neurons are undergoing TNR instability during replication or recombination. The most accepted mechanism of TNR instability in mitotically inactive cells is involving transcription, especially in G1/G0-arrested striatal medium spiny neurons. Until recently, proof of concept was shown only in bacteria cells, and the data did not provide insights into the mechanism. These data was suggestive that there was interplay between replication and transcription, causing the instability (Bowater et al., 1997; Mochmann and Wells, 2004; Parniewski et al., 1999; Schumacher et al., 2001). Again, this mechanism seemed unlikely, as replication is not a factor in nondividing neurons. Furthermore, there has been instability shown in nonproliferating female germ line and sperm precursor cells (Kaytor et al., 1997; Kovtun and McMurray, 2001; McMurray and Kortun, 2003).

Eventually a mechanism was proposed addressing TNR instability in nondividing cells that was replication independent (Mochmann and Wells, 2004).

This mechanism proposed that during the strand separation for transcription, hairpins would form on both opposing strands, and is referred to as transcription-coupled trinucleotide repeat instability. The success by which DNA repair resolved the hairpins was the factor that determined whether an expansion or deletion occurred (Fig. 7). If DNA repair pathways recognized and addressed the hairpins, the result would be deletions of the sequence composing the bulky adduct. But, if cleavage occurred on the strand opposite of the hairpin, the hairpin would eventually resolve itself out of secondary structures. At this point, DNA repair would synthesize missing bases and the product would be an expansion. While this proposed mechanism addresses TNR instability in non-proliferating cells, it is highly probable that more than one mechanism is causing TNR instability in proliferating cells.

Trinucleotide repeat instability forms the basis for genetic anticipation, and thus the pathogenicity for HD. Even though the subject is extremely researched, we have yet to create any comprehensive mechanism. Identifying whether one mechanism is responsible for the mosaicism of repeat instability in tissues of different types, or if different mechanisms are involved for different tissues seems to be the starting point. Most data is pointing to the probability that tissue-mosaicism is mechanistic-specific as well, which further convolutes the problem by identifying all the mechanisms. Unfortunately, the more data gained only presents more questions to ask. As for HD and our current knowledge of cell biology, and knowing disease-affected cells are mitotically inactive, DNA repair likely plays a role and is mediated by transcription. Using proliferating cells in studies to determine TNR instability allows for the analysis of collective responses by the mechanisms in the cell. The advantage of using mitotically active cells provides a platform to determine, not only the mechanisms involved, but also allows for determination of the degree of involvement each mechanism has on TNR regions.

A



B

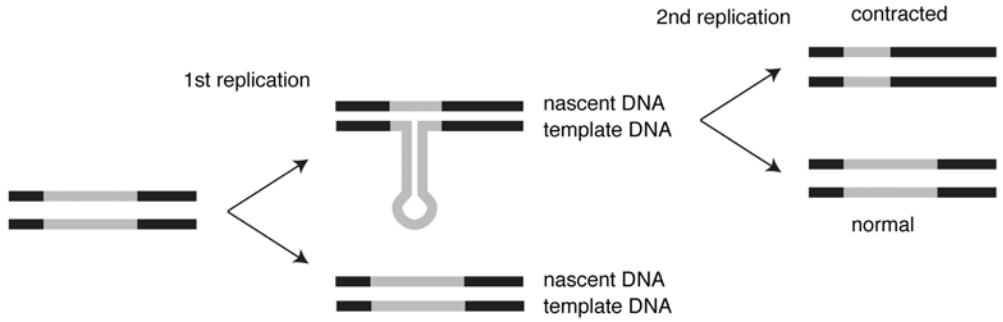


Fig. 6. Mechanism for hairpin formation and trinucleotide repeat instability. Upon polymerase slippage, single-stranded DNA has anneal to itself. This structure is stabilized through Watson-Crick base pairing of cytosine-guanine bonding. Hairpins in nascent DNA result in expansions (A), where hairpin formation in template strands result in contractions (B). Figure from (Lin et al., 2010). Copyright Nature Publishing Group, U.S.A.

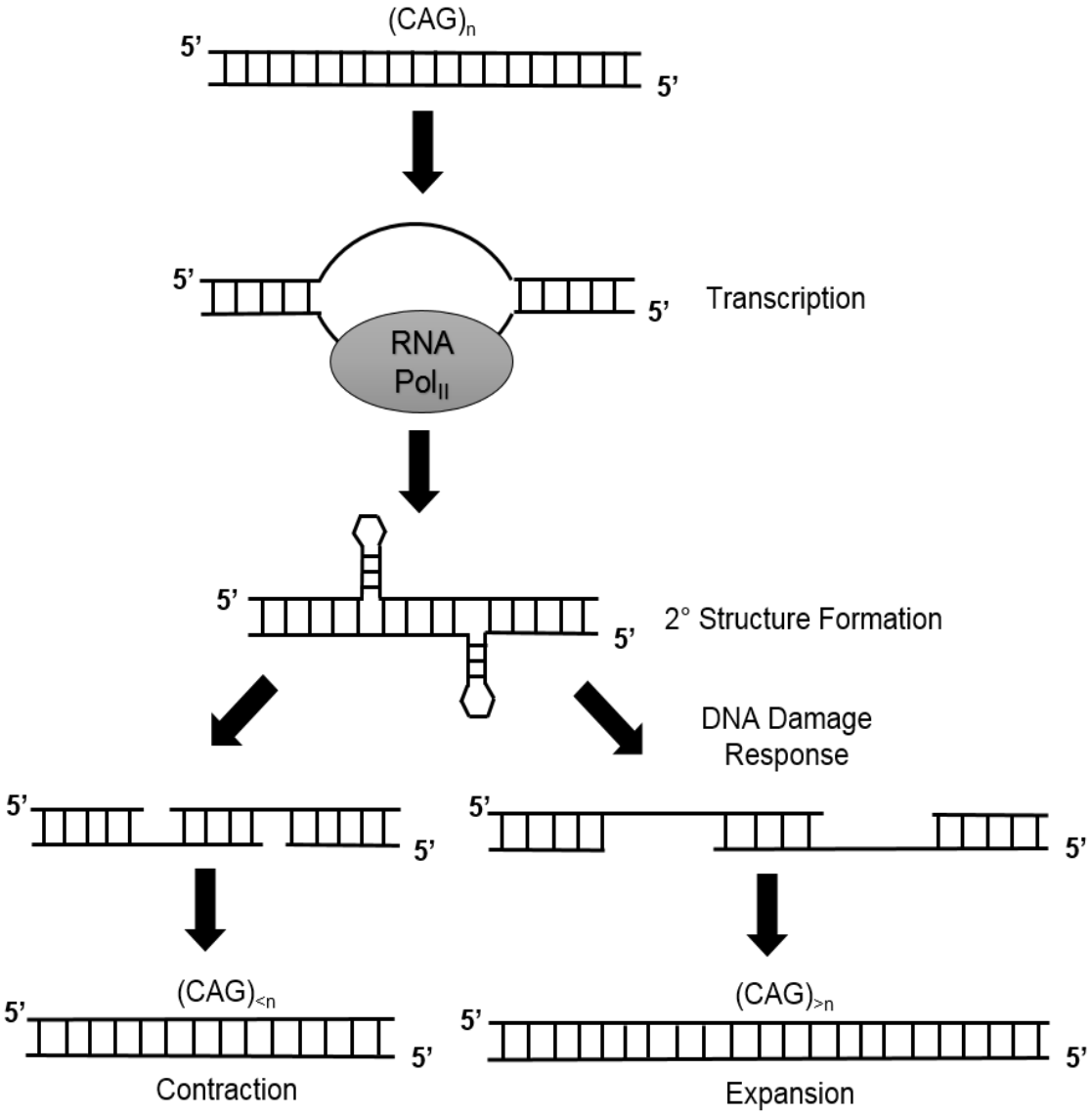


Fig. 7. Pathway for transcription-induced triplet repeat instability. Upon transcription termination, separated DNA strands have the propensity to form hairpin structures, due to their intrastrand complementarity. Correct resolution of the bulky adducts by DNA repair is determined on how they are addressed. Cleavage flanking hairpins themselves results in removal of the complementing base pairs composing the hairpins (bottom-left). Incisions in the opposing DNA backbone, opposite to the hairpin, would allow the eventual denaturing of the hairpin, leading to genome expansions (bottom-right). Figure adapted from (Lin and Wilson, 2007).

DNA repair as effectors of trinucleotide repeat instability

DNA damage repair is necessary to preserve genomic integrity from damage/mutations caused by both endogenous and exogenous factors. Unfortunately, data is emerging that DNA damage repair can inadvertently have negative consequences on genome integrity when addressing alterations, specifically when addressing repeat expansions (Zhao and Usdin, 2015a). In most instances, DNA repair pathways ensure proper maintenance, addressing DNA strand damage induced by endogenous and exogenous mutagens and effectors. However, as shown in trinucleotide repeat diseases, new research is uncovering ways that DNA repair increases the risk of mutations. The magnitude and severity of the negative effects of dysfunctional DNA repair are dependent upon a number of factors, including location of the tandem repeats, size of the repeats, and number of repeats within the allele, as well as the specific repeat sequence and orientation (Fry and Usdin, 2006). Also, physical DNA secondary structure and conformation during DNA modifications process plays an effector of efficient DNA repair. The physical existence and unique properties of tandem repeats in the genome creates a multitude of factors that have challenging characteristics that complicate efficient DNA damage repair.

DNA replication-independent mechanisms suggest that TNR instability in proliferation-inactive cells, such as neurons and some germ line cells, is mediated by DNA repair. This proposal warrants consideration based on the fact that TNR instability has been shown to occur, unnecessary of cells entering the S-phase of the cell cycle. There are predominantly two general proposed mechanisms of replication independent TNR instability. Both mechanisms involve DNA repair responses to bulky adducts and both result in tandem repeat length alterations. The first mechanism infers that DNA repair is initiated, resulting in strand nicks at damaged bases or at oxidative damage. The cleaved single strand now has the propensity to create complementarily-base hairpins which lead to the insertion of tandem repeats. The second general proposal involves the consequence of heteroduplex formation during transcriptional unwinding of the double helix. The formation of aberrant DNA structures then initiate a DNA repair response. In both

of these general mechanisms proposed, TRN instability has been shown to occur. While these mechanisms provide adequate explanations of TNR instability in mitotically inactive cells, they ultimately rely on DNA damage repair mechanisms. As well, these mechanisms do not address somatic tandem repeat expansion mosaicism witnessed in HD.

The most basic process involved in DNA damage repair is mismatch repair (MMR). This repair process is essential to identifying genome regions that are highly repetitive and have a propensity to cause DNA polymerase to slip back, thus synthesizing extra bases on the nascent strand. Upon re-annealing of the two strands, bulges are created (Jalal et al., 2011). The MMR response involves the formation of Msh2/Msh6 and Msh2/Msh3 complexes that form cyclical structures that encompass the DNA strands and detects bulges. Upon detection of an bulky adduct, MutL is recruited to ensure removal of the adduct on the damage strand, and that DNA is resynthesized (Weinberg, 2014). Ironically, more data is coming to light that MMR drives triplet repeat expansions (Castel et al., 2010; Halabi et al., 2012; Kovalenko et al., 2012; Panigrahi et al., 2012; Pinto et al., 2013). Through the use of mouse models, knock-outs of proteins essential to MMR show TNR stability is induced, reducing the rate of expansions (Pinto et al., 2013). Also, there has been evidence that increased oxidative damage inhibits the efficacy of MMR to properly maintain the genome from MMR-related damage (Chang et al., 2002). The combination of these data is difficult to interpret in the context of HD. These data implies that the increased oxidative stress present in HD should inactivate MMR, thus stabilizing TNR regions in the genome. On the other hand, inactivated MMR in HD would exacerbate the formation of bulky adducts causes by slipped polymerases. It is evident that further investigation on the role of MMR, possibly in conjunction with other repair mechanisms, is needed to discover the comprehensive mechanism responsible for TNR instability.

Excision repair pathways are the predominant pathways that maintain genome integrity. These pathways address a myriad of DNA lesions, such as complex intra-strand formations, to base methylations. Four basic steps are involved in excision repair, with each step absolutely necessary for successful

DNA damage repair. The basic steps are damage recognition, excision of damaged bases, synthesis of excised region, and ligation of newly synthesized strands. The response elicited to address this damage is determined by the complexity and extent of the damage recognized. In recent years, it has been made evident that there are two distinct pathways in which excision pathways address DNA damage, known as base excision repair (BER) and nucleotide excision repair (NER). In the context of tandem repeat expansions and HD, components of both excision pathways have been implicated in being effectors of TRN instability.

The most basic excision repair pathway is base excision repair. BER initiates a series of steps involving proteins to repair a recognized non-bulky lesion on a newly synthesized strand. This damage is predominantly from oxidation, deamination, and alkylation to the DNA strand. Unique to this pathway are a class of at least 11 proteins known as glycosylases, which are proteins that recognize specific types of damage. Two sub-pathways of BER exist, in which either 1 nucleotide is replaced, or 2-13 nucleotides are replaced; dependent upon which pathway addresses damage determines which set of proteins are activated to conduct necessary repair. Interestingly, it is shown that knock-out mouse models of glycosylases OGG1 and NEIL1 reduces tandem repeat expansions in somatic cells (Kovtun et al., 2007; Mollersen et al., 2012). The primary function of these proteins is to identify and excise oxidized bases. Importantly, the loss of NEIL1 expression has shown intergenerational reductions in the average size of expansions (Mollersen et al., 2012). Furthermore, apurinic/apyrimidinic endonuclease 1 (APEX1) is gaining more attention in research as its role in TNR instability is becoming more apparent. Recent studies show that APEX1 prevented expansions with its exonuclease activity in hairpin loop-outs and stem-loops (Beaver et al., 2015). Another study showed that oxidative damage to the bases in a hairpin ultimately affected the efficiency of APEX1 to properly address the hairpin (Li et al., 2014). These data further confound the comprehensive model of TNR instability, as revealing multiple components of different mechanisms being effectors of tandem repeat expansions in the genome.

The last excision-based DNA repair process is nucleotide excision repair (NER). NER is a process that targets DNA damage where bulky lesions have been created within the double stranded DNA. Unlike the previously mentioned repair mechanisms, DNA damage caused by ultra-violet radiation is recognized by NER (Friedberg, 2011). The repair of DNA damage by NER involves up to 30 base pairs, and is divided into two pathways: transcription-coupled repair (TCR-NER) and global genome repair (GGR-NER) (Hanawalt et al., 2000). Though only the initial steps of damage recognition differentiate the two sub-pathways, both pathways converge for the remaining steps of incision, repair and ligation (Scharer, 2011). Both pathways ultimately implement the appropriate base pairs.

Global genome repair employs sensing proteins through continuous surveying of the DNA strands during transcribed and non-transcribed DNA. Tight stringency is placed on the replication process to ensure the appropriate implementation of proper nucleotides. DNA-damage binding (DDB) proteins and the xeroderma pigmentosum, complementation group C - *S. cerevisiae* Rad23 homolog (XPC-hHR23B) complex facilitate proper synthesis by verifying the absence of helix distortions created by UV damage. Upon detection of helix abnormalities, the bound complex will initiate the remaining steps of mechanistic repair.

Transcription-coupled repair differs from global genome repair in that it has a severely lower efficiency of repairing DNA damage in non-transcribed DNA, hence the name ‘transcription-coupled repair’. The cell does not rely on the detection of helix distortions by XPC or DDB proteins. This pathway is initiated when the cell has detected prolonged blockage of RNA polymerase II (RNA pol II). In both sub-pathways of NER, once the ultra-violet light induced DNA damage has been detected, transcription factor II H ($TF_{II}H$), composed of XPD and XPB, creates an incision while XPG stabilizes the strands and maintains endonuclease properties. This complex contains other proteins that facilitate helicase activity and single-strand DNA protection.

While extensive research on neurodegenerative-based mouse-models shows that NER, and its components affect tandem repeats, currently no data

shows that aberrant NER machinery is directly responsible for TNR instability in HD pathogenesis. Loss of XPA induces contractions in specific neuronal tissues in SCA1 mouse models (Hubert et al., 2011). Also, Cockayne Syndrome B (CSB), a factor necessary for transcription elongation, induces TNR expansions in FXD mice in both germ line and somatic cells (Zhao and Usdin, 2014; Zhao and Usdin, 2015b). Interestingly, CSB up-regulates OGG1, NEIL1, and APEX1 expression, concluding that it might play an alternate role in BER responses (Muftuoglu et al., 2009; Tuo et al., 2002; Wong et al., 2007).

Double-strand breaks (DSB), and its' subsequent repair introduce severely complicated mechanisms, further confounded with a multitude of sequential protein recruitments. The very nature of double-strand breaks in the genome, homologous or non-homologous, represents damage from normal cellular processes, such as oxidative stress, or it can be in response to severe DNA damage from irradiation, chemical agents, or ultra-violet light. Ultimately, failure to repair DSB's will ultimately lead to cell death. To increase the complexity of repairing DSB's, there can be a combination of types of chromosomal breaks. Fortunately, there are several different mechanisms to address all combinations of DSB's. The simplest instance would be the utilization of homologous recombination, in which same regions of a sister chromosome are used as a template to synthesize regions in a damaged strand. The damaged DNA strand undergoes digestion of the damage, and also of flanking bases to expose complimentary sequences used for annealing to the sister chromatid/chromosome. Upon completion of single-strand overhangs, the strand migrates to its compliment in the sister chromatid/chromosome. After successful sequence pairing, DNA polymerase elongates the damaged strand, pushing the non-bound strand out from the chromosome. The strand from the sister chromosome that has been relocated out then becomes a template for DNA synthesis for the remaining strand that was initially damaged. Upon successful synthesis of damaged strands, ligation occurs, integrating both newly synthesized regions into their respective location in the damaged strand. This process has often been described as "error-free", as the template for synthesis is from an

undamaged, sister chromatid/chromosome, assumed to have an intact and function gene sequence region.

As of now, homologous recombination of DSB repair is the only repair mechanism of this type to be implicated in tandem repeat instability. *S. cerevisiae* Rad52 is a protein that binds the ends of the damaged DSB's and facilitates its' migration and annealing to the complement strand on the sister chromatid/chromosome. In DM1 mouse models, loss of Rad52 results in germ line tandem repeat contractions (Savouret et al., 2003). There is no evidence as of yet that homologous recombination events elicit changes in TNR instability in HD. However, data shows that expansions occur in HD haploid sperm in the absence of sister chromatids/chromosomes (Kovtun and McMurray, 2001). The absence of an undamaged DSB or sister template eliminates DSB repair as a mediator of TNR instability. Interestingly, Rad52 does interact with OGG1 via BER of DNA damage by oxidation, implying that it may have an alternative role in TNR instability (de Souza-Pinto et al., 2009).

To further complicate the pathogenesis of TNR instability in HD, DNA synthesizing machinery has been implicated as having effects on tandem repeat regions. DNA polymerases conduct the physical synthesis of new strands or new sequences inserted into regions where damaged bases were removed. Studies have implicated DNA polymerases β and δ/ε as effecting TNR regions due to polymerase slippage, with ultimately causes "re-synthesis" of previously synthesized template regions (Liu et al., 2009; Liu and Wilson, 2012; Sadeghian and Ochsenfeld, 2015).

It is apparent that TNR instability is, at least partially, a product of aberrant DNA repair processes. Our current knowledge only presents more questions. All four of the general DNA damage repair pathways have been shown to be effectors of TNR instability. Three of the four mechanisms have been validated to be involved in HD or HD models. The fourth mechanism, homologous recombination of DSB's, has only been shown to affect tandem repeat sequences in another TNR disorder. Currently research has not implicated homologous recombination to affect *HTT* gene expansions. Needless to say, the somatic TNR mosaicism only

further complicates research. We not only have to identify the direct mechanism by which HD TNR instability occurs, but there is evidence of cross-interactions between proteins associated with multiple repair pathways. And still, it is not clear if differential somatic TNR instability in each cell/tissue type is the product of the same DNA damage response.

Epigenetics and Huntington's disease

The dysfunction that mHtt causes to cellular pathways is extensive. Direct involvement of mHtt has been adequately researched, with the beginnings of many investigations involving proteins that directly bind and interact with mHtt. Unfortunately there is still a lot of that we don't know. There are reports of differential homeostatic pathways emerging concerning HD, with results only showing that mHtt in involved somehow, but not by direct interaction of known factors. These changes witnessed implies that epigenetic patterning is altered in HD. Although there is much we don't understand about epigenetic mediation of HD, it is firmly understood that the differential epigenetic alterations are facilitated by mHtt involvement.

The most studied epigenetic factor is histone acetylation. Briefly, residues on histone tails can have acetyl groups bound, dispersing the electrostatic charge resulting in relaxation of the DNA strand. The relaxing of the DNA strand promotes gene expression. The opposite, deacetylated residues promote gene down regulation, in that the electrostatic charge is localized, promoting DNA strand tightening around the histone complex (Kurdistani et al., 2004). Some studies have reported differential histone acetylation patterns in HD, while other data contrast those findings (Benn et al., 2009; Bobrowska et al., 2011; Buckley et al., 2010). It is thoroughly shown that the presence of mHtt affects gene expression by prevention of DNA relaxing around the histone complex (Glajch and Sadri-Vakili, 2015; Lee et al., 2013; Wang et al., 2014). The binding and sequestration of CBP by mHtt is one of the first dysfunctional epigenetic changes noticed in HD (Steffan et al., 2000). CBP, referenced earlier for its transcription factor properties, also functions as a histone acetyltransferase (HAT). In the presence of CBP, histone

complex tails are acetylated, relaxing DNA from around the complex. This relaxed state allows transcriptional regulatory factors the physical space needed to bind to regulatory elements (Jiang et al., 2006; Nucifora et al., 2001; Steffan et al., 2000; Taylor et al., 2003). This interaction between mHtt and CBP has been shown in many HD models, recently including neuronal cells from HD patients (Hoshino et al., 2003; Yeh et al., 2013). Further, there are multiple data sets reporting conflicting information concerning mHtt altering the expression and efficacy of histone deacetylases (HDACs), involved in removal of acetyl groups on histone tails to tightly package DNA. Some groups show that some enzymes involved in HDAC class I were overexpressed, but this did not result in increased transcript levels in cortical and striatal tissue .In contrast, HD mouse models showed increases in HDAC class I enzymes and decreases in HDAC class II enzymes, but cortical expression was not changed (Quinti et al., 2010). More so, decreasing some HDAC class I and II enzymes did not have an effect on HD pathological phenotypes (Benn et al., 2009; Bobrowska et al., 2012; Bobrowska et al., 2011; Moumne et al., 2012). These data clearly demonstrate the variation on HAT and HDAC functions that mHtt imparts.

As well as acetylation of histones, another epigenetic regulatory factor that has been shown to be differential in HD is methylation patterning. Cytosine bases to the 5' side of guanine bases (CpG) can undergo a methylation event at the 5^{th} atom of the ring. The accumulation of this modification in localized regions of gene promoters can have profound effects on gene expression. Increased methylation of CpG regions inhibit binding of basal regulatory factors, thus decreasing expression. Also, heavily methylated regions decrease gene expression by recruitment of co-repressor complexes (Jaenisch and Bird, 2003). Initially thought to be a completely repressive mechanism, data is emerging that methylation can also promote gene expression (Suzuki and Bird, 2008).

Genome wide studies of methylation patterns in HD have revealed extensive differences relative to unaffected models. These data reveals that increased methylation patterns are gene specific, with some having increased methylation and some sites have decreased methylation. Further, these data

shows that methylation is not regulated by a stable epigenetic patterning, but instead can be the result of cellular processes (Ng et al., 2013). This study further identifies regions of transcriptional factor SOX2, a protein necessary for pluripotency, as being heavily methylated in HD mouse models.

Furthermore, data suggests that differential methylation of CpG sites occur in HD. Presumptively, increased methylated CpG (meCpG) sites would have negative effects on the genome, in that spontaneous deamination of meCpG sites yield thymine bases. This formation of thymine can result in a point mutation, known as transition mutations. Transition mutations are changes in a purine nucleotide or pyrimidine nucleotide to another purine or pyrimidine nucleotide, respectively. This further complicates HD, in that hyper-methylated regions are prone to increased mutation probabilities, resulting in point mutations.

It is clear that mHtt has a multitude of negative effects on normal cell function. While most pathways that are known to be aberrant have some data showing mechanistic interference by mHtt, the differential epigenetic patterning mechanisms are still coming to light. In all the possible effectors of HD pathogenesis, it is evident that TNR instability in somatic expansions and transmission is most likely a combination of more than one currently defined mechanism, if not involving unknown mechanisms. Clearly, a comprehensive model to define pathogenesis and HD TNR instability would employ the simultaneous analysis of multiple pathways while introducing modifying gene expression assays.

Induced pluripotent stem cells and Huntington's disease

The comprehensive model by which TNR expansions persist in neurodegenerative diseases is incomplete. More so, the mechanism by which TNR expansions resulting in genetic anticipation during gene transmission has continued to confound as well. Research into these cell-specific models has previously been largely held back due to the inability to have access to cell types. Prior to the discovery of iPSC, the only method of investigating disease-specific cells, striatal neurons, required grossly invasive and were often only attainable by

a relatively small portion of researchers. The ability to reprogram readily accessible cell types into induced pluripotent stem cells, and then further target their differentiation into disease-specific cell models *in vitro* provides researchers a platform to investigate cellular mechanisms at a molecular level not previously possible. Since the discovery of iPSC reprogramming, iPSC-derived cell models have permitted experiments that would otherwise have been impossible, such as targeted therapeutics, disease-specific cell characterization, and cell replacement (Benraiss and Goldman, 2011; Castiglioni et al., 2012; Chae et al., 2012; Consortium, 2012). Having the ability to create and investigate disease-specific cells types allows us to refocus attention into determining the inherent pathogenesis of TNR instability. Determining the relationship between DNA damage repair, epigenetic involvement, and tissue-specific TNR instability through differentiation is the key to understanding the mechanism by which pathogenic genetic anticipation occurs.

Recent studies have shown that pluripotent cells have enhanced DNA repair response and efficacy of nonhomologous end joining and nucleotide excision repair, similar to embryonic stem cells (ESCs) (Fan et al., 2011; Luo et al., 2012). Interestingly, iPSCs have also been reported to show increased microsatellite instability, devoid of any known DNA repair gene mutations (Luo et al., 2012). While some studies have shown increases of base excision repair glycosylases (OGG1 and NEIL1) in neural stem and progenitor cells, other data reports down-regulation of BER components XRCC1, LIG1 and LIG3 during terminal differentiation (Hildrestrand et al., 2007; Hildrestrand et al., 2009; Narciso et al., 2007; Rolseth et al., 2008). As mentioned previously, over-expression and under-expression of DNA repair genes have both positive and negative effects on TNR instability. It is likely that these differential gene expression changes could be involved in the TNR instability in HD.

The field of iPSC is a relatively new field, and provides a tool for any researcher group to investigate their niche in their respective disease-specific cell type. In relation to HD, recent data reports 26 differentially expressed genes in HD-iPSC models, relative to wild-type iPSCs (Chae et al., 2012). Interestingly, this

same study identified an increase is genes related to oxidative stress, a known effector of HD pathogenesis. Also, they saw decreases in cytoskeletal proteins involved in neuronal differentiation, underlying the possibility that neurogenesis is aberrant in HD from initial pluripotent cell state. As well, HD-iPSCs show increased lysosomal activity, inferring reversion to pluripotency might alleviate mHtt-induced aggregates (Camnasio et al., 2012). Another study showed that it is possible to correct the expanded (CAG) TNR region with a normal (CAG) region in HD-iPSC using homologous recombination, and further showed that terminal-differentiation of the corrected HD-iPSCs into striatal neurons retained the corrected phenotype (An et al., 2012). This study alone emphasizes the importance of iPSC and disease-specific-cell models.

Rationale of study and hypothesis

Huntington's disease is a neurodegenerative disease that exacerbates itself through trinucleotide instability leading to genetic anticipation. Through cellular dysfunction, multiple processes are interrupted, ultimately leading to striatal neuron loss. Of the cellular processes effected in HD, gene expression through mHtt interference has been shown to reduce expression of a myriad of genes through the inadvertent binding and sequestration of transcription factors and epigenetic modifiers. The disease pathogenesis is poorly understood, as there are currently multiple proposed mechanisms postulated to be responsible for somatic cell tandem repeat instability mosaicism witnessed in HD. Currently it is known that effectors of TNR instability include replication, DNA damage repair, transcription, and endogenous and exogenous stressors. As well, recent finding are implicating epigenetic modifications as having negative effects on TNR stability. DNA damage repair is the central node in all the processes in which TNR expansions are allowed to persist. Thus, it is likely that a combination of these processes work in conjunction with each other to allow the pathogenic TNR region of the *HTT* gene to remain uncorrected.

Our long-term goal is to understand the mechanism by which mHtt affects DNA repair gene expression, and the consequence for differential gene

expression. The objective of this research is to explore the relationship between HD and the DNA repair gene expression. Therefore, it is our hypothesis that TNR instability is exacerbated by DNA damage repair gene dysregulation in HD. Further, we hypothesize that gene expression of selected DNA repair genes is a consequence of mHtt interference on epigenetic patterning. The hypothesis is supported by studies demonstrating the extensive transcriptional dysregulation caused by the presence of mHtt in the cell, and the known effects differential DNA repair regulation can have on TNR instability. The rational is that elucidating novel DNA repair gene regulation, and effectors of the differential gene expression, will enhance our knowledge of comprehensive mechanism for TNR instability and somatic mosaicism seen in HD. Here, we used HD-affected fibroblast cells and iPSCs to identify novel down-regulated DNA repair genes that are essential to genome integrity. We further determined the extent of the witnessed dysregulation and found it to be present in multiple HD samples. Using pharmacological dose treatments to alter epigenetic regulators, we successfully recovered the aberrant gene expression. We also determined that intermittent treatments of hypomethylation agent 5-azacytidine induced TNR stability over prolonged treatments. These data reveals that pharmacological intervention to prevent somatic instability might be possible. Reprogramming HD-fibroblasts into a pluripotent cell-state revealed that DNA repair gene deficiencies were corrected to wild-type levels, elucidating the possibility that global epigenetic changes in conjunction with pluripotency affect DNA repair gene expression in HD. Further, determined that HD-iPSC lines shown contraction-biased instability in early passages, and then shown enhanced stability in late passages. This identifies a cell-state in HD where mechanisms that address TNR regions are functioning with increased efficacy. Then we determined that further differentiation of HD-iPSC lines retain TNR stability. These two experiments provide essential data to identifying the endogenous pathway(s) responsible for TNR instability. The data we present here suggests that TNR instability in HD might, at least in part, be the result of mHtt interference with epigenetic regulatory mechanisms associated with

pluripotency. This information is vital to identifying specific targets to alleviate HD TNR instability.

CHAPTER II

DIFFERENTIALLY EXPRESSED DNA REPAIR GENES IN HUNTINGTON'S DISEASE FIBROBLASTS

Introduction

Huntington's disease (HD) is an autosomal dominant neurodegenerative disease that causes dysarthria, chorea and dementia, as well as emotional and psychiatric disturbances (Peyser and Folstein, 1990; Pla et al., 2014). HD is characterized by an uninterrupted trinucleotide repeat (TNR) expansion of a normal cytosine-adenine-guanine $(CAG)_{9\text{-}34}$ sequence in the N-terminal coding region of the *Huntingtin* gene (*HTT*) on chromosome 4p16 (MacDonald et al., 1993). Normal 350 kDa huntingtin protein (Htt) is ubiquitously expressed in all tissues and cell types, and is involved in numerous processes including vesicular transport, transcriptional regulation, anti-apoptosis, embryonic development, hematopoiesis, mitosis and other essential survival pathways (Dragatsis et al., 2000; Duyao et al., 1995; Gauthier et al., 2004; Marcora et al., 2003; Metzler et al., 2000; Nasir et al., 1995; Rigamonti et al., 2000). Translation of the $(CAG)_n$ sequence produces polyglutamine (polyQ) residues. However, posttranslational modification of an expanded mutant huntingtin protein (mHtt) renders short fragments of polar poly-Q chains that aggregate and bind to each other and create inclusion bodies leading to neuronal degeneration (DiFiglia et al., 1997; Kim et al., 2006; Schaffar et al., 2004).

While neuronal death has been strongly associated with HD, several HD mouse models have near normal numbers of neurons yet exhibit a disease phenotype. Scientists have linked this neuronal death independent phenotype to transcriptional dysregulation (Cha, 2007). Further studies found that mHtt-induced gene expression is altered in HD brain but not limited to neuronal tissue (Cha, 2007). mRNA expression abnormalities are thought to arise in part because Poly-Q aggregate traps (PQATs) hijack and aggregate important transcriptional regulators including nuclear protein p53, cyclic AMP-response element (CRE)-

binding protein (CREB), specific protein-1 (Sp1), CREB-binding protein (CREB-BP/CBP), TATA-box-binding protein (TBP), and TATA binding protein (TBP)-associated factor ($TAF_{II}130$) (Cha, 2007; Gil and Rego, 2008; Jiang et al., 2006; McCampbell and Fischbeck, 2001).

Though not completely understood, it is widely accepted that during somatic DNA replication and modification processes, the (CAG)-TNR expansion involves the production of looped intermediates that can form alternate DNA conformations, such as slipped-strand structures and hairpins (Sinden, 2001). These single-stranded DNA (ssDNA) looped intermediates contain *cis*-stabilized "Watson-Crick" base-pairing of cytosine-guanine through intra-strand complementary base pairs, as well as adenine-adenine mismatched base pairs (Gacy et al., 1995; Kovtun and McMurray, 2008; Mirkin, 2007). *Trans*-factors directly enhance TNR hairpin-structure instability (expansion and contraction), as previously shown with MSH2, MLH1, MLH3, PMS2, and MSH3 (Gomes-Pereira et al., 2004; Halabi et al., 2012; Kovalenko et al., 2012; Lin and Wilson, 2009; Pinto et al., 2013). However, expansion increases during replication do not exclusively address the increased occurrence of TNR expansions observed in quiescent cells affected by neurodegenerative diseases (Gomes-Pereira et al., 2001; Kennedy and Shelbourne, 2000).

Integration of DNA hairpins is facilitated through multiple cellular mechanisms, both dependent and independent of genome replication. Instability has been recorded in the transient separated DNA strands during recombination, repair, and transcription independent of genome replication (Jakupciak and Wells, 2000; Lin et al., 2006; Lin et al., 2009; Lopez Castel et al., 2009; Wells et al., 2005). In addition, epigenetic modifications and R-loops have involvement in transcription-induced (CAG*CTG)-TNR instability. Despite the extensive list of factors involved in TNR instability, a comprehensive model to explain the molecular mechanism of TNR instability in HD has yet to be elucidated.

DNA repair dogma states that non-canonical structures assembled during transcription or replication are identified and addressed appropriately based upon the nature and severity of the lesion. Specifically, mismatch repair, and possibly

other repair processes, should address erroneous strand insertions that create bulky adducts. Some studies have implicated individual DNA repair pathways as molecular mechanisms for TNR instability (Lai et al., 2013; Liu and Wilson, 2012). Currently there is no evidence that HD directly involves DNA repair gene down-regulation. DNA repair pathways are affected by HD. Based on published observations, we hypothesized that PQATs contribute to decreased mRNA levels of essential DNA repair genes, thus facilitating the characteristic trinucleotide expansion and impaired damage responses in HD. Here we measured DNA repair gene expression in a HD-affected fibroblast cell line compared to a cell line established from a healthy relative.

Materials and methods

Cell lines and cell culture

Early passage control fibroblast line GM-02187 and HD-affected fibroblast line GM-03621 was obtained from Coriell Cell Repositories (Camden, NJ). HD patient's genomic DNA previously sequenced was characterized as having 18 and 60 (CAG) repeats within exon 1 of the *HTT* gene. Cells were maintained at 37°C in a humidified incubator with 5% CO_2. Normal human fibroblasts and HD-affected fibroblasts were cultured in Modified Eagle's Medium (MEM) (Life Technologies, Carlsbad, CA) with 15% fetal bovine serum (FBS; Life Technologies) or 1:1 Dulbecco's Modified Eagle's Medium and MEM (Life Technologies) with 10% FBS, respectively as recommended by Coriell. Cell media also contained 1x penicillin/streptomycin and 1x glutamine (Life Technologies). Cells were seeded in T75 cm^2 flasks at 5.0 x 10^4 cells/ml, and medium was changed between 72 and 96 hours of seeding. Cells were passaged every 5-6 days using TrypLE™ Express (Life Technologies) following manufacturer's instructions. Use of human cell lines purchased was deemed exempt from IRB review, as it is impossible to identify cell donors.

RNA isolation and quantification

Upon 70-80% confluency, total cellular RNA was isolated with Direct-zol™ RNA MiniPrep kits (Zymo Research, Irvine, CA) according to the manufacturer's protocol. RNase-free DNase I digestion (50 units) (Roche Applied Science, Indianapolis, IN) was performed as recommended. DNase I Buffer was added to RNA samples at 1x. One µl (2 U) DNase I was added to each 50 µl reaction and incubated at 37 °C for 30 minutes. RNA samples were then extracted with phenol/chloroform to inactivate DNase I. After alcohol precipitation, each RNA sample was brought up to one µg/µL in DNase-free RNase-free water. RNA quantification was determined by UV absorbance at 260 nm (A260 nm) on a NanoDrop 2000 (Thermo Scientific, Wilmington, DE). RNA quality was deemed acceptable with 260/280 ratios around 2.0 and 260/230 ratios slightly higher than 1.8.

First strand synthesis and PCR arrays

One µg of each RNA sample was reverse transcribed into cDNA using RT^2 First Strand Kit (SABiosciences, Valencia, CA) according to manufacturer's instructions. Normal and HD-affected cDNA was amplified using a RT^2 Profiler™ PCR Array for Human DNA Repair and RT^2 SYBR Green PCR Master Mix (PAHS-042) (SABiosciences) to assess mRNA levels. The PCR array provides optimized primers for 84 genes directly involved in human DNA repair, 5 endogenous housekeeping genes, as well as reverse-transcriptase and PCR reaction controls. PCR array experiments were conducted on a MyiQ2™ Real Time PCR Detection System (Bio-Rad, Hercules, CA). Cycling conditions were as follows: 1 cycle of 10 minutes at 95°C, 40 cycles of 15 seconds at 95°C and 1 minute at 60°C.

Selected gene expression levels of *APEX1, BRCA1*, and *MSH5* were independently verified four times using PrimePCR™ SYBR® Green Assays (Bio-Rad) for human *BRCA1* and *MSH5*, with *ACTB* as an endogenous normalizer. One µg of each purified cellular RNA sample was reverse transcribed using iScript (Bio-Rad) at 42°C for 30 minutes in the presence of hexamer and oligo(dT). Negative controls included reactions lacking reverse transcriptase or RNA template. Primer sets for *ACTB* and *APEX1* cDNA were created using Beacon

Designer software (PREMIER Biosoft International, Palo Alto, CA) and synthesized by Integrated DNA Technologies (Coralville, IA). Primer set efficiencies was determined by a dilution series of cDNA in triplicate. A 100 bp *ACTB* fragment was amplified with the following primer set: 5'-AGCCTCGCCTTTGCCGATCC-3' (Tm 62.3 °C) and 5'-ACATGCCGGAGCCGTTGTCG-3' (Tm 62.8 °C). A 317 bp *APEX1* fragment was amplified using the following primer set: 5'-GACAAAGAGGCAGCAGGAGAGG-3' (Tm 59.8 °C) and 5'-GGGCACTGGCGGGAAAGC-3' (Tm 62.7 °C). All DNA repair gene expression was performed by real-time PCR amplification with a MyiQ2™ Real Time PCR Detection System. In detail, each 20 µl qPCR reaction with Prime PCR™ assays contained: 1x gene specific PrimePCR assay mix, 1x SsoAdvanced Universal SYBR® Green Supermix, and 100 ng of cDNA sample. Mastermix for *ACTB* and *APEX1* in a 20 µl reaction included: 1 x SsoAdvanced Universal SYBR® Green Supermix, 0.5 nM of each primer, and 100 ng of cDNA template. Reaction cycling parameters were polymerase activation at 95.0 °C for 120 seconds, followed by 40 cycles of 95.0 °C for 15 seconds and 60.0 °C for 30 seconds. A melt curve analysis was conducted afterward where the temperature was lowered to 65.0 °C, and then increased in increments of 0.5 °C every 5 seconds to 95.0 °C.

Data analysis

HD-affected and control fibroblast RT^2 Profiler PCR Array experiments were carried out four times. Cycle-threshold analysis of targeted genes was initially conducted by online analysis software provided by SABiosciences specifically designed for the RT^2 Profiler™ Human DNA Repair PCR Array. Fold regulation was determined by the ratio of the measured value ($2^{(-\Delta Ct)}$) for an HD-affected sample to the value ($2^{(-\Delta Ct)}$) for the wild-type control. For PCR array independent validation, gene expression levels of *APEX1, BRCA1* and *MSH5* were verified against the initial experiment four times using ACTB to normalize. Comparative analysis was conducted using the $\Delta\Delta C_T$ method (Livak and Schmittgen, 2001). Data are expressed as a mean +/- standard deviation.

Statistical analysis

All experiments were conducted three to five times with sub-sampling of each independent quantitative experiment. Statistical significance was determined by a two-tailed, heteroscedastic student T-test.

Results

Differential expression of DNA repair genes in one HD fibroblast line

Total mRNA of HD affected fibroblasts and unaffected fibroblasts were extracted and analyzed using RT^2 Profiler^{TM} PCR Array for Human DNA Repair. The RT^2 Profiler^{TM} PCR Array allowed evaluation of gene expression for 84 genes vital to successful recognition and repair of DNA damage (Table 2). Each individual well of the assay contained validated primers related to base-excision repair, mismatch repair, nucleotide excision repair, double-strand break repair, single-strand DNA stability, and post-repair modification.

The results show consistent down-regulation of multiple genes involved in base-excision repair, mismatch repair, nucleotide excision repair, and other accessory genes involved in DNA repair (Fig. 8). The scatter plot is of the log base 10 of the hybridization intensity of each gene in the two samples. The lines above and below represent a 3-fold change in gene expression threshold. The middle line represents a fold-change ($2^{-\Delta Ct}$) of 1. Of the 84 genes examined through the assay, 11 genes show significantly deficient levels and 1 gene showed an increase in expression in HD affected cells. Expression levels of 53 genes remained consistently regulated (< 2- fold) between the wild-type sample GM02187 and HD sample GM03621, showing no significant differences, while 18 additional genes showed decreased fold changes between 2 and 3. The 11 points below the bottom line indicate down-regulated (*ATXN3, CCNO, DMC1, EXO1, LIG3, MSH5, NTHL1, RAD18, RPA1, XPC,* and *XRCC6BP1*) genes greater than 3- fold. Up-regulated and down-regulated genes in HD cells and the control cells are summarized in Table 3. The 1 point above the top line indicate up-regulated (*XRCC6*) gene. Statistically significant down-regulation of the *XRCC6BP1, DMC1,* and *RPA1* genes were found ($p < 0.05$). Endogenous housekeeping genes used was *B2M,*

HPRT1, GAPDH, and *ACTB*. A fifth housekeeping gene is provided in the array (*RPL13A*), but was removed from analysis due to extreme inter-assay variability.

Table 2

DNA repair related genes analyzed using the RT^2 Profiler™ Array.

Gene	Description
APEX1	APEX nuclease (multifunctional DNA repair enzyme) 1
APEX2	APEX nuclease (apurinic/apyrimidinic endonuclease) 2
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad3 related
ATXN3	Ataxin 3
BRCA1	Breast cancer 1, early onset
BRCA2	Breast cancer 2, early onset
BRIP1	BRCA1 interacting protein C-terminal helicase 1
CCNH	Cyclin H
CCNO	Cyclin O
CDK7	Cyclin-dependent kinase 7
DDB1	Damage-specific DNA binding protein 1, 127kDa
DDB2	Damage-specific DNA binding protein 2, 48kDa
DMC1	DMC1 dosage suppressor of mck1 homolog, meiosis-specific homologous recombination (yeast)
ERCC1	Excision repair cross-complementing rodent repair deficiency, complementation group 1 (includes overlapping antisense sequence)
ERCC2	Excision repair cross-complementing rodent repair deficiency, complementation group 2
ERCC3	Excision repair cross-complementing rodent repair deficiency, complementation group 3 (xeroderma pigmentosum group B complementing)
ERCC4	Excision repair cross-complementing rodent repair deficiency, complementation group 4
ERCC5	Excision repair cross-complementing rodent repair deficiency, complementation group 5
ERCC6	Excision repair cross-complementing rodent repair deficiency, complementation group 6
ERCC8	Excision repair cross-complementing rodent repair deficiency, complementation group 8
EXO1	Exonuclease 1
FEN1	Flap structure-specific endonuclease 1
LIG1	Ligase I, DNA, ATP-dependent
LIG3	Ligase III, DNA, ATP-dependent
LIG4	Ligase IV, DNA, ATP-dependent
MGMT	O-6-methylguanine-DNA methyltransferase
MLH1	MutL homolog 1, colon cancer, nonpolyposis type 2 *(E. coli)*
MLH3	MutL homolog 3 *(E. coli)*
MMS19	MMS19 nucleotide excision repair homolog *(S. cerevisiae)*
MPG	N-methylpurine-DNA glycosylase

Table 2

Continued.

Gene	Description
MRE11A	MRE11 meiotic recombination 11 homolog A *(S. cerevisiae)*
MSH2	MutS homolog 2, colon cancer, nonpolyposis type 1 *(E. coli)*
MSH3	MutS homolog 3 *(E. coli)*
MSH4	MutS homolog 4 *(E. coli)*
MSH5	MutS homolog 5 *(E. coli)*
MSH6	MutS homolog 6 *(E. coli)*
MUTYH	MutY homolog *(E. coli)*
NEIL1	Nei endonuclease VIII-like 1 *(E. coli)*
NEIL2	Nei endonuclease VIII-like 2 *(E. coli)*
NEIL3	Nei endonuclease VIII-like 3 *(E. coli)*
OGG1	8-oxoguanine DNA glycosylase
PARP1	Poly (ADP-ribose) polymerase 1
PARP2	Poly (ADP-ribose) polymerase 2
PARP3	Poly (ADP-ribose) polymerase family, member 3
PMS1	PMS1 postmeiotic segregation increased 1 *(S. cerevisiae)*
PMS2	PMS2 postmeiotic segregation increased 2 *(S. cerevisiae)*
PNKP	Polynucleotide kinase 3'-phosphatase
POLB	Polymerase (DNA directed), beta
POLD3	Polymerase (DNA-directed), delta 3, accessory subunit
POLL	Polymerase (DNA directed), lambda
PRKDC	Protein kinase, DNA-activated, catalytic polypeptide
RAD18	RAD18 homolog *(S. cerevisiae)*
RAD21	RAD21 homolog *(S. pombe)*
RAD23A	RAD23 homolog A *(S. cerevisiae)*
RAD23B	RAD23 homolog B *(S. cerevisiae)*
RAD50	RAD50 homolog *(S. cerevisiae)*
RAD51	RAD51 homolog *(S. cerevisiae)*
RAD51C	RAD51 homolog C *(S. cerevisiae)*
RAD51B	RAD51 homolog B *(S. cerevisiae)*
RAD51D	RAD51 homolog D *(S. cerevisiae)*
RAD52	RAD52 homolog *(S. cerevisiae)*
RAD54L	RAD54-like *(S. cerevisiae)*
RFC1	Replication factor C (activator 1) 1, 145kDa
RPA1	Replication protein A1, 70kDa
RPA3	Replication protein A3, 14kDa
SLK	STE20-like kinase
SMUG1	Single-strand-selective monofunctional uracil-DNA glycosylase 1
TDG	Thymine-DNA glycosylase
TOP3A	Topoisomerase (DNA) III alpha
TOP3B	Topoisomerase (DNA) III beta
TREX1	Three prime repair exonuclease 1
UNG	Uracil-DNA glycosylase
XAB2	XPA binding protein 2
XPA	Xeroderma pigmentosum, complementation group A

Table 2

Continued.

Gene	Description	Gene	Description
XPC	Xeroderma pigmentosum, complementation group C	*XRCC6BP1*	XRCC6 binding protein 1
XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1	*B2M*[a]	Beta-2-microglobulin
XRCC2	X-ray repair complementing defective repair in Chinese hamster cells 2	*HPRT1*[a]	Hypoxanthine phosphoribosyltransferase 1
XRCC3	X-ray repair complementing defective repair in Chinese hamster cells 3	*RPL13A*[a]	Ribosomal protein L13a
XRCC4	X-ray repair complementing defective repair in Chinese hamster cells 4	*GAPDH*[a]	Glyceraldehyde-3-phosphate dehydrogenase
XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining)	*ACTB*[a]	Actin, beta
XRCC6	X-ray repair complementing defective repair in Chinese hamster cells 6	*HGDC*[b]	Human Genomic DNA Contamination

[α]The RT^2 PCR Profiler Array for Human DNA Repair contains validated and optimized primer sets for 84 genes associated with human DNA repair pathways including base excision repair, nucleotide excision repair, mismatch repair, and single- and double- strand break repair. The array also contains 5 housekeeping genes to normalize expression between tested samples, as well as negative controls to detect contamination. [a]House keeping genes. [b]Human genomic DNA contamination control.

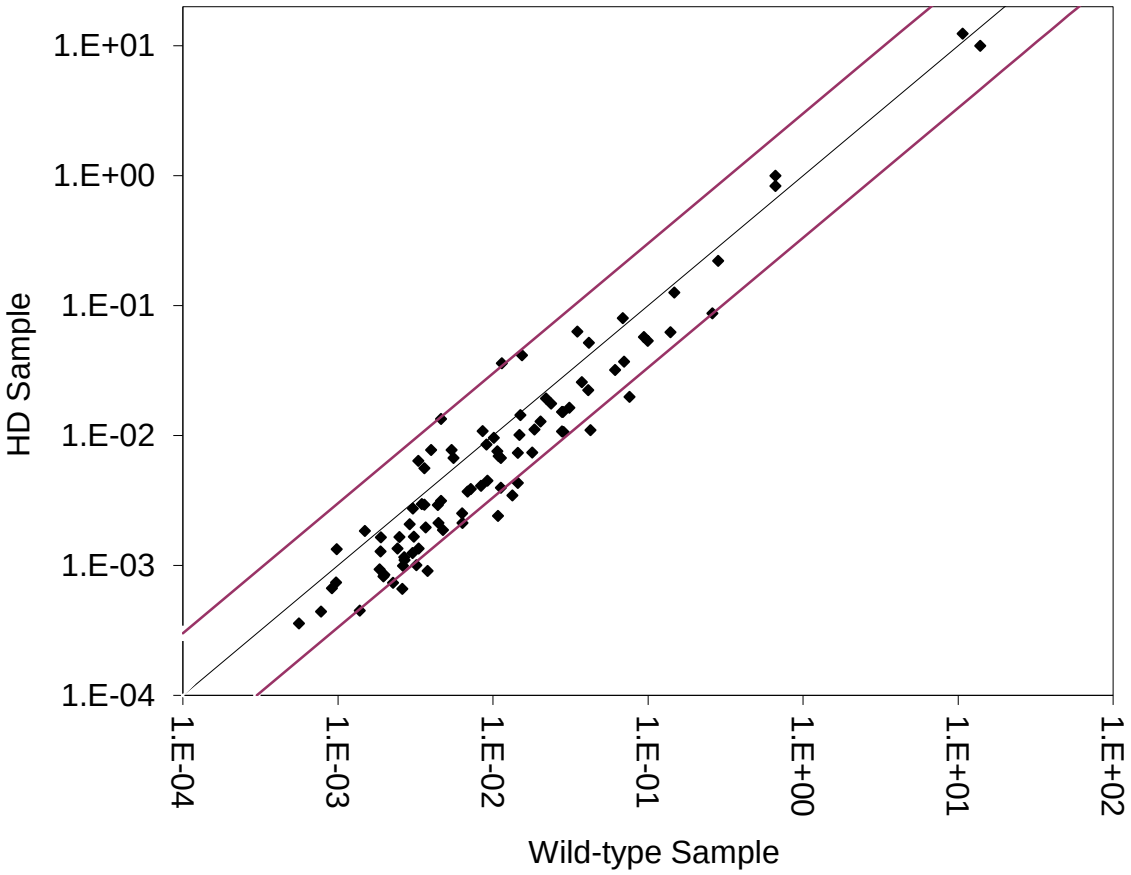


Fig. 8. Expression of 84 DNA repair genes in HD-affected fibroblasts vs. wild-type fibroblasts. Scatter plot of the log base 10 hybridization intensity of each DNA repair gene. The x-axis is wild-type fibroblasts and the y-axis is HD-affected fibroblasts. A fold-change ($2^{-\Delta Ct}$) of 1 is represented by the middle line. The top and bottom lines represent the desired gene expression threshold in fold-change (3). The 11 points below the bottom line represent *ATXN3, CCNO, DMC1, EXO1, LIG3, MSH5, NTHL1, RAD18, RPA1, XPC,* and *XRCC6BP1*. The one point above the top line represents *XRCC6.*

Table 3

RT^2 Profiler Array differentially expressed DNA repair genes and pathway involvement.

Gene	DNA Repair Pathway Involvement	Fold change	95% CI	*P*-value
Down-regulated				
CCNO	Base Excision Repair	0.2235	(0.00001, 0.91)	0.355084
LIG3	Base Excision Repair	0.2990	(0.00001, 0.70)	0.257783
APEX1	Base Excision Repair	0.3744	(0.07, 0.68)	0.106029
BRCA1	Recombination and Nonhomologous End-Joining) (V(D)J recombination)	0.7416	(0.07, 1.42)	0.60676
NTHL1	Base Excision Repair	0.2607	(0.00001, 0.60)	0.159206
XRCC6BP1	Double-Strand Break Repair (Homologous Recombination and Nonhomologous End-Joining) (V(D)J recombination)	0.2536	(0.08, 0.42)	**0.027606**
DMC1	Double-Strand Break Repair (Homologous Recombination)	0.3249	(0.12, 0.53)	**0.020691**
EXO1	Mismatch Repair	0.3138	(0.00001, 0.83)	0.179559
MSH5	Mismatch Repair	0.3238	(0.12, 0.53)	0.076045
ATXN3	Nucleotide Excision	0.2395	(0.02, 0.46)	0.086207
RAD18	Post-replication repair of damaged DNA	0.2585	(0.00001, 0.54)	0.081376
RPA1	Stabilizes ssDNA intermediates	0.2580	(0.09, 0.43)	**0.022726**
Up-regulated				
XRCC6	Double-Strand Break Repair (Nonhomologous End-Joining) (V(D)J recombination)	3.1670	(0.00001, 12.80)	0.991114
Housekeeping genes				
B2M		1.2554	(0.00001, 2.99)	0.769982
HPRT1		0.9514	(0.00001, 2.35)	0.647336
GAPDH		1.1612	(0.08, 2.24)	0.978881
ACTB		0.7210	(0.00001, 2.08)	0.458981

[α]Down-regulated DNA repair genes in HD cell line GM03621 in the RT2 Profiler™ PCR Array for Human DNA Repair. Fold changes are relative to wild-type cell line GM02187. Expression is determined using the $2^{(-\Delta\Delta CT)}$. Bolded p-values represent statistical significance. Statistical significance was determined by a two-tailed, heteroscedastic student T-test.

Independent confirmation of DNA repair gene down-regulation

Upon reviewing the mRNA expression data from the RT^2 Profiler^{TM} PCR Array, *APEX1, BRCA1*, and *MSH5* consistently showed down-regulation (Fig. 9). As per the analysis software provided, the mean $2^{-\Delta Ct}$ of the four experiments is utilized for fold change. *APEX1, BRCA1*, and *MSH5* are arbitrarily set at a fold expression of 1 in the control sample. A trend was observed in the down-regulation of these genes, though statistical significance was inconclusive, the substantial down-regulation warranted further investigation. Therefore *APEX1, BRCA1*, and *MSH5* were selected to be independently verified through additional experiments including new extractions analyzed with real-time PCR using PrimePCR™ Assays (Bio-Rad).

PrimePCR™ Assays provided primers that were verified and optimized to confirm our initial findings from the RT^2 Profiler™ PCR Array. Four independent analyses of *APEX1, BRCA1*, and *MSH5* mRNA levels using PrimePCR™ Assays revealed similar decreases in expression levels in each gene. For each experiment, the individual fold expression ($2^{-\Delta\Delta Ct}$) is averaged. Error bars indicate standard error of fold expression across the four experiments. *ACTB* was used as a housekeeping gene. Statistical analysis revealed that *APEX1* was significantly down-regulated. Although not statistically significant, down-regulation of *BRCA1* and *MSH5* was similar with PrimePCR™ Assays as we had seen previously with the RT^2 Profiler™ PCR Array (Table 4). This consistent down-regulation verifies the results previously acquired with the Profiler PCR Array.

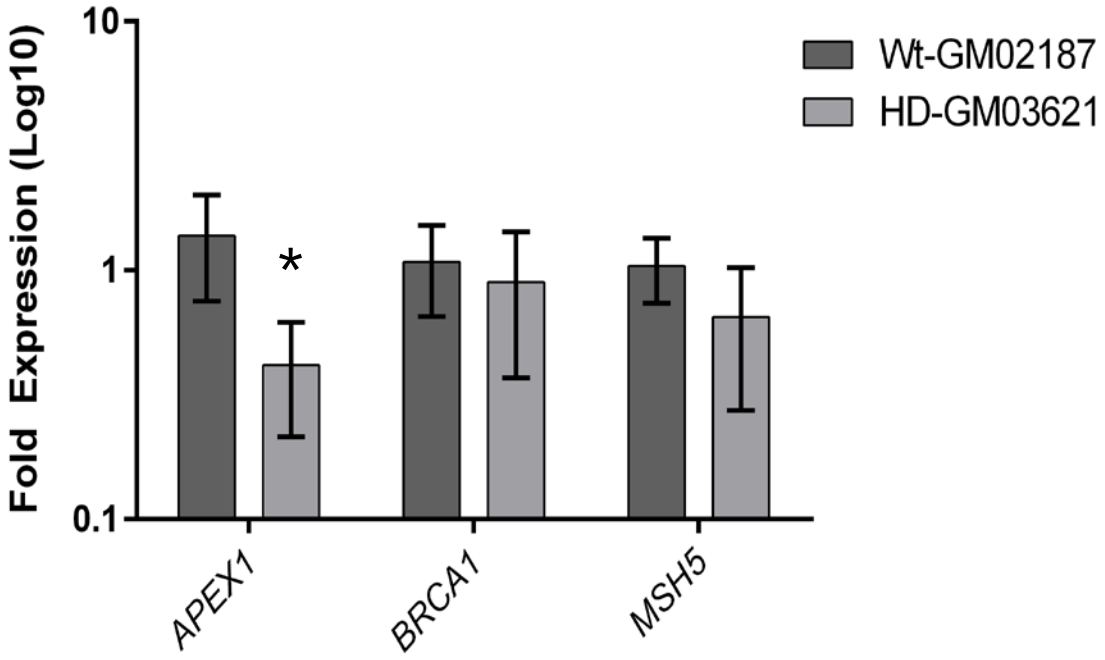


Fig. 9. Verification of *APEX1*, *BRCA1* and *MSH5* down-regulation in HD-affected fibroblasts using PrimePCR™ Assays. The mRNA level is presented in relation to control group, wild-type unaffected fibroblast line GM02187, with mRNA levels of targeted genes arbitrarily set up as 1. Error bars represent the standard deviation of the mean of four independent experiments.

Table 4

Independent verification of selected DNA repair gene expression using PrimePCR™ Assays.

Gene	Fold change	95% CI	*P*-value
Down-regulated			
APEX1	0.3744	(0.07, 0.68)	**0.026029**
BRCA1	0.7416	(0.07, 1.42)	0.60676
MSH5	0.5689	(0.13, 1.04)	0.287329
Housekeeping gene			
ACTB	1	(1.00, 1.00)	0

[α]Validation of gene expression of *APEX1, BRCA1,* and *MSH5* of HD cell line GM03621 using PrimePCR™ Assays. Beta-actin was used as an endogenous normalizing gene. Gene expression was determined using the $2^{(-\Delta\Delta CT)}$ method relative to wild-type cell line GM02187. Statistical significance was determined by a two-tailed, heteroscedastic student T-test.

Discussion

In this preliminary study we discovered HD-affected fibroblasts have an overwhelming number of down-regulated DNA repair genes in the absence of any known DNA repair gene mutations. Using qPCR-based assays we compared the expression of 84 DNA repair genes of HD-affected fibroblasts to unaffected fibroblasts. We initially identified 29 genes down-regulated and 4 genes up-regulated greater than 2-fold. To our knowledge, this is the first report of DNA repair gene down-regulation in HD. Using the RT^2 Profiler^{TM} PCR Array for Human DNA Repair we demonstrate the down-regulation of multiple genes involved in base excision repair, nucleotide excision repair, mismatch repair, double-strand break repair (homologous and nonhomologous end-joining), as well as accessory genes involved in strand stabilization and polymerase activation.

We further verified decreased expression of *APEX1, BRCA1*, and *MSH5* each necessary for successful completion base excision repair, mismatch repair, and double-strand break repair. Using PrimePCR™ Assays we independently verified the expression levels of *APEX1, BRCA1*, and *MSH5*. Both qPCR-based assays yielded similar decreased levels of these genes. Based upon the reduction in mRNA of these genes, attenuated levels of *APEX1, BRCA1*, and *MSH5* proteins are likely to be observed in future studies. Aberrant DNA repair mRNA compromise efficiency of repair machinery to address damaged DNA and non-canonical secondary structures created during DNA modification. Through the inability to address ssDNA loop-outs, TNR instability maintains and expansions persist.

HD is a dynamic mutation disease, where the number of (CAG) repeats increases the probability of expression of the mutant phenotype. People with fewer than 35 (CAG) repeats are not affected by the disease, but individuals with 35 to 40 repeats are at an increased risk of developing HD. Expansions greater than 40 repeats inevitably lead to HD progression (Bates, 2003). Though the mechanism is not fully understood, HD-affected individuals exhibit genetic anticipation caused by subsequent generations of HD-affected individuals developing expanded versions of the mutated allele (Langbehn et al., 2004; Langbehn et al., 2010;

Ranen et al., 1995). Increased expansion in successive generations correlates with more severe symptomatic expression and earlier age of onset in 70% of individuals (Lee et al., 2012; Walker, 2007). Dysfunctional DNA repair pathways would likely have implications on the severity of genetic anticipation experienced in HD.

Translation of the *mHtt* $(CAG)_n$ sequence present in HD patients produces repetitive poly-Q fragments that aggregate, thereby trapping and sequestering key cellular proteins. Effects of disease-causing poly-Q fragments in disruption of transcriptional regulation by binding of poly-Q residues to other proteins with similar amino acid sequences have been extensively reported. Poly-Q aggregate traps (PQATs) hijack important transcriptional activators and disrupt transcriptional pathways including cyclic adenosine monophosphate responsive element-binding protein (CREB), CREB-binding protein (CREB-BP/CBP), transcription factor Sp1 (Sp1), transcription factor II D (TFIID), transcription factor II F (TFIIF), and TATA-box-binding protein (TBP) (Dunah et al., 2002; Kazantsev et al., 1999; Nucifora et al., 2001; Shimohata et al., 2000a; Zhai et al., 2005). Based on the down-regulation detected of *APEX1*, *BRCA1*, and *MSH5* mRNA in HD-affected fibroblasts, we examined each gene's respective promoter regions. Surprisingly this examination revealed APEX1 and BRCA1 are regulated by transcription factors Sp1, CREB and CREB-BP/CBP, while *MSH5* has a Sp1 binding site (Atlas et al., 2001; Grosch and Kaina, 1999; Harrison et al., 1995; Xu et al., 1995). Sequestration of CBP by PQAT involvement is linked to the increased methylation and hypoacetylation of histones in HD (McFarland et al., 2012; Sadri-Vakili et al., 2007). PQAT involvement with CBP, and acetyltransferase activity through the associated CBP/p300 complex, gives rise to the possibility that direct transcriptional regulatory sequestration, histone modification, and chromatin remodeling may partially be an underlying factor for the DNA repair gene down-regulation observed in our study (Gardian et al., 2005; Ng et al., 2013; Ryu et al., 2006; Steffan et al., 2001; Steffan et al., 2000). Though not within the scope of this study, it would be logical to address the transcriptional regulatory machinery

for each of the down-regulated genes in HD-affected cells thoroughly characterize known transcriptional disruption to down-regulated mRNA levels.

CHAPTER III

DNA METHYLTRANSFERASE INHIBITOR 5-AZACYTIDINE RECOVERS DNA REPAIR GENE EXPRESSION AND STABILIZES TRINUCLEOTIDE REPEATS IN HUNTINGTON'S DISEASE FIBROBLASTS

Introduction

Huntington's disease (HD) is a neurodegenerative disease that affects medium spiny neurons in the basal ganglia, resulting in involuntary muscle movements and ultimately leading to death. It is caused by an expansion of cytosine-adenine-guanine (CAG) trinucleotide repeat (TNR) region in exon 1 of the huntingtin (*HTT*) gene (MacDonald et al., 1993). HD is a dynamic mutation disease, where the number of (CAG) repeats increases the probability of expression of the mutant phenotype. People with fewer than 35 (CAG) repeats are not affected by the disease, but individuals with 35 to 40 repeats are at an increased risk of developing HD. Expansions greater than 40 repeats inevitably lead to HD progression (Bates, 2003). Though the mechanism is not fully understood, HD-affected individuals exhibit genetic anticipation caused by subsequent generations of HD-affected individuals developing expanded versions of the mutated allele (Langbehn et al., 2004; Langbehn et al., 2010; Ranen et al., 1995). It has been shown that increased expansions in successive generation's correlates with more severe symptomatic expression and earlier age of onset in 70% of individuals (Lee et al., 2012; Walker, 2007). Interestingly, post-mortem analysis of affected cells and other somatic tissues in HD- and other polyglutamine (poly-Q) disease- mouse models reveals disease-causing gene (CAG) mosaicism, in which (CAG)'s of different tissues have different sized expansions (Fortune et al., 2000; Gomes-Pereira et al., 2001; Kennedy and Shelbourne, 2000; Telenius et al., 1994). As well, a myriad of TNR instability effectors have been identified, such as environmental stresses, chemical inducers, and DNA repair gene expression (Chatterjee et al., 2015; Gomes-Pereira et al., 2004; Gomes-Pereira and Monckton, 2004; Kovalenko et al., 2012; Wheeler et al., 2007). This evidence

alone infers that TNR instability is not only present in germ line cells, but also throughout development and maturation, and thus dependent upon efficacy of DNA repair and its' ability to appropriately respond to endogenous and exogenous DNA damaging agents and modifiers. The purpose of this study was to investigate DNA repair gene expression in HD fibroblasts.

There are multiple possible mechanisms proposed explaining TNR expansions leading to genetic anticipation, such as polymerase slippage and stabile heteroduplex formation. Unfortunately, the comprehensive mechanism of TNR instability eludes researchers. Of the proposed mechanisms, heteroduplex formation and its' inability for successful resolution are the common factor behind TNR instability (Lee and McMurray, 2014). Through intra-stand complimentary base pairing, single-stranded DNA (ssDNA) creates secondary structures, such as hairpins or loop-outs (Sinden, 2001). The ssDNA lesion is stabilized through cytosine-guanine hydrogen bonding (Gacy et al., 1995; Kovtun and McMurray, 2008; Mirkin, 2007). Though this secondary structure meets the criteria for DNA damage as we understand, the lesion is ignored and the result is an incorrect number of CAG repeats. These expansions lead to earlier onset and more severe symptoms during disease progression. Interestingly, expansion in the *HTT* gene has been reported in both mitotically active and quiescent cells. Tissue specific analysis of R6/2 mouse-models show large variability in CAG repeats, with mitotically inactive striatal tissue samples revealing greater CAG expansions in the *HTT* gene (Fortune et al., 2000; Vatsavayai et al., 2007). Prevention of TNR expansions is not only the key to understanding the pathogenesis of HD, but also essential to the development of pharmaceutical treatments.

Pathology of HD involves cellular dysfunction on multiple levels. The elongated gene encodes for repetitive glutamine amino acids, in which these polyglutamine (poly-Q) chains contain hydrophobic properties. A number of studies have suggested that post-translational cleavage of huntingtin protein renders a shorter N-terminal fragment with the poly-Q expansion (Kim et al., 2006; Ratovitski et al., 2009). The mhtt protein and potential fragments created exhibits a toxic gain-of-function, in which poly-Q chains bind unintended targets,

sequestering them from normal function (Nucifora et al., 2001). Poly-glutamine regions of wild-type and mhtt physically interact many DNA-binding proteins (Seredenina et al., 2011). Consequently, expanded poly-Q regions of mHtt have been show to disrupt normal cell functions through interaction with multiple basal transcription factors causing gene expression dysregulation, including CREB, CREBBP, Sp1, TFII30, NRSF/REST, PFC1α, BCL11b, TBP, and others (Dunah et al., 2002; Kazantsev et al., 1999; Li et al., 2002; Seredenina and Luthi-Carter, 2012; Steffan et al., 2000; Zhai et al., 2005).

Another proposed mechanism of HD pathology is the involvement of mhtt on DNA methylation patterning. It is established that gene silencing can be a result of increased DNA- and histone- methylation. ERG-associated protein with SET domain (ESET), a regulator of H3 trimethylation, shows increased expression in HD and can be therapeutically addressed, thus increasing neuronal survival in the R6/2 mouse model (Ryu et al., 2006). Furthermore, extensive changes in DNA methylation is associated with the expression of mHtt. Interestingly, these changes in methylation patterns seem to be sequence- specific to DNA binding motifs (Ng et al., 2013). While it seems that researchers haven't seen direct protein to protein interactions between mHtt and methyltransferases, the probability that either gain-of-function mHtt, or the loss of wild-type Htt, has an impact on epigenetic imprinting.

The detection and identification of methylation as regulatory effects on gene expression involves a multi-step process. Prior to quantitative gene expression assays, such as qPCR, extracted gDNA samples are exposed to bisulfite treatments. This method exploits the ability of unmethylated cytosine residues to undergo conversion to uracil. This conversion leads to subsequently created strands (through amplification processes) having thymine incorporated within the strand in place of the unmethylated cytosine. Methylated cytosine resist the transformation to uracil, thus further copies of methylated cytosine nucleotides maintain the natural cytosine incorporation. Using nucleotide identification techniques, primarily capillary electrophoresis sequencing, researchers are able to accurately detect methylation patterns and levels in genes and regions flanking

genes. The amount of methylation present in regions flanking transcriptional start sites are known to effect transcriptional regulation. It has been previously agreed upon that increased methylation patterning in regions encompassing transcriptional start sites has negative effects upon gene expression. Interestingly, there is some data suggesting that increased methylation patterns can have increased-expression effects as well (Suzuki and Bird, 2008). After successful bisulfite treatments, workflows remain the same using sequencing techniques to accurately identify chemically-replaced cytosines with thymines in bisulfite treated and non-treated samples.

Here we hypothesize that HD fibroblasts would have abnormal expression of DNA repair genes, sensitizing them to (CAG) TNR expansions. We determined that key DNA repair genes were down-regulated in HD fibroblasts compared to wild-type controls. Cell lines pHD-1 and pHD-2 had expansions of 17/42 and 18/44 respectively, and treatment with 5-azacytidine resulted in increased DNA repair expression and stabilization of (CAG) repeats during culture. This is the first study to identify endogenously deficient DNA repair genes in human fibroblasts and is proof of principle that pharmacological intervention could lead to increased (CAG) stability in patients at risk for the disease.

Materials and methods

Cell lines and cell culture

Three previously confirmed HD (GM03621, GM02079, and GM03868), two wild-type (GM02187 and GM02153), and two unknown (GM02191 and GM04022) fibroblast cell lines were obtained from Coriell Cell Repositories (Camden, NJ). Unknown fibroblast samples were obtained from donors of which one parent was confirmed to be affected by HD, but were not diagnosed. Genetic characteristics of exon 1 of the *HTT* gene of each cell line used in our experiments are determined later, and referenced in Table 6. Cell lines were cultured and maintained at 37°C in a humidified incubator with 5% CO_2. Fibroblasts were cultured in Dulbecco's Modified Eagle's Medium (DMEM) (Life Technologies, Carlsbad, CA) with 10% fetal bovine serum (FBS; Life Technologies) as recommended by Coriell. Cell

media also contained 1.0% Antibiotic-Antimycotic (100X) (ABAM; Life Technologies). Cells were seeded in T75 cm^2 flasks at 5.0 x 10^4 cells, and medium was changed between 72 and 96 hours of seeding. Cells were passaged every 5-6 days using TrypLE™ Express (Life Technologies) following manufacturer's instructions. Use of human cell lines purchased was deemed exempt from IRB review, as it is impossible to identify cell donors.

5-Azacytidine and trichostatin-A treatment

Wild-type and HD fibroblasts were grown in culture for 24 hours and then treated with a serial dilution of 5-azatacytidine (Sigma Aldrich; St. Louis, MO) reconstituted in phosphate-buffer saline (PBS) to determine LC_{50} for 120 hours (data not shown). Serial dilution concentrations were 20, 10, 5, 1, 0.5 µM. LC_{50} was determined to be 10 µM. At cell passaging, cell media was supplemented with 5-azacytidine to achieve 10 µM. Cell sample media was replaced daily with supplemented media for 5 days. At the end of 5 days, cells were washed and gDNA and total RNA was collected.

Cell lines pHD-1 and pHD-2 were plated in T-75 flasks and treated intermittently with 10 µM 5-azacytidine over the course of 35 days to reach four population doublings. Treatment schedule was as follows; Pass in normal medium for 24 hours, treat with 10 µM 5-azacytidine for 5 days, allow cells to recover for 3 days in normal medium, repeat with passaging at a 1:4 split ratio. During 5-azacytidne treatment days, cells were washed with PBS 2x before replacing the medium supplemented with 5-azacytidine. Genomic DNA samples were collected at the beginning of each treatment schedule and at the end of the designated time frame.

For histone-deacetylase inhibition (HDACi) treatment, complete fibroblast medium was supplemented with 0.5 µM and 5.0 µM trichostatin-A (TSA) (Sigma Aldrich) reconstituted in DMSO to achieve 0.01% v/v. TSA containing media was applied 16 hours after passaging cells, and allowed to remain for 24 hours. Upon removal of TSA, total RNA was immediately extracted.

RNA isolation, quantification, and first strand synthesis

Upon 70-80% confluency (approximately one million cells), total cellular RNA was isolated with Tri-zol™ (Thermo Fisher, Waltham, MA) following the phenol: chloroform manufacturer's protocol. To ensure pure RNA samples, RNase-free DNase I digestion (50 units) (Roche Applied Science, Indianapolis, IN) was performed as recommended. RNA quantification was determined by UV absorbance at 260 nm (A260 nm) on a NanoDrop 2000 (Thermo Scientific, Wilmington, DE). RNA quality was deemed acceptable with 260/280 ratios around 2.0 and 260/230 ratios slightly higher than 1.8. One µg of each RNA sample was reverse transcribed into cDNA using Platinum™ Taq DNA Polymerase High Fidelity Kit (Thermo Fisher) according to manufacturer's instructions. cDNA samples were further diluted to a concentration of 5-10 ng/µL in dnase-free, rnase-free water.

Quantitative real time PCR

Wild type, HD, and unknown cDNA samples was amplified using TaqMan® Gene Expression Assay (Applied Biosystems, Foster City, CA.) to assess mRNA levels. TaqMan® Gene Expression Assays provided optimized primers and probes specific for DNA repair genes *APEX1* (Hs00959050_g1)*, BRCA1* (Hs01556193_m1)*, RPA1* (Hs00161419_m1), and *RPA3* (Hs01047933_g1)*;* and the endogenous housekeeping gene *ACTB* (Hs99999903_m1)*.* Real-time quantitative PCR experiments were conducted with a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA). Probes consist of 5' terminal reporter dye FAM (6-caroxyfluorescein) and 3' terminal quencher dye NFQ-MGB. TaqMan® Fast Advanced Master Mix (Life Technologies) was used in conjunction with TaqMan® Gene Expression Assays. Complete reactions were conducted in 20 µL volumes and contained 1X TaqMan® Fast Advanced Master Mix, 1X TaqMan® Gene Expression Assay, and 50 ng of cDNA template. All qPCR experiments included NTC and NRT controls to assure results were not from contaminating nucleic acid in either PCR or cDNA reactions. Thermal cycling conditions were as follows: 1 cycle of 2 minutes at 50°C, 1 cycles of 20 seconds

at 95°C, and 40 cycles of 1 second at 95°C and 20 seconds at 60°C. All experiment reactions were independently completed in triplicate. Negative template controls were run on each experiment containing nuclease-free water. All cell lines were tested 3 independent times, with at least 2 population doubling events occurring between sample experiments. Each experimental sample was sub-sampled 3 times during qPCR reactions. Error bars represent the standard deviation of three independent experiments conducted in triplicate. *ACTB* was used as our endogenous housekeeping gene for normalization.

Immunofluorescence of DNA repair genes

Cell samples were seeded onto Lab-Tek microscope chamber slides at a density of $1.5*10^3$ cells/ml. Cells were washed three times with phosphate-buffered saline (PBS). Slides were fixed with 10% formalin for 30 min at room temperature, followed by PBS washing three times. Samples were blocked with 10% goat serum for 60 min at room temperature. Primary antibodies were obtained from ABCAM (Cambridge, MA). APEX1 antibody (ab105081; 1:700), BRCA1 antibody (ab16781; 1:250), RPA1 antibody (ab79398; 1:250), and RPA3 antibody (ab6432; 1:50) were diluted in 1.0% goat serum and applied overnight in a humidified chamber at 4.0°C. Each sample slides contained designated chambers lacking the addition of primary antibodies to provide a negative control against non-specific binding and for background correction. The following day samples were washed three times for 10 minutes with 1x TBS. FITC conjugated secondary anti-rabbit and anti-mouse antibodies (1:1000) in 1.0% goat serum were applied to the samples for 60 min, followed by three washes with 1x TBS containing 0.1% NP-40. DAPI (1:1000) was diluted in TBS and applied for 10 min. Following incubation, slides were washed four times with 1x TBS containing 0.1% NP-40. Fluorescent images were acquired with a Zeiss Axio microscope at 10x magnification. Images were collected in 16 bit, black and white to maximize resolution. False coloring was added by ImageJ. At least three images of each gene/cell line was collected to ensure global capture of the population of cells.

Genomic DNA extractions and purification

Approximately $1.0*10^6$ cells were harvested and gDNA was extracted using DNeasy Blood & Tissue Kit (Qiagen; Hilden, Germany) according to the manufacturer's protocol. Samples were purified and concentrated using 100% ethanol and sodium acetate. All gDNA samples were resuspended in DNase-free water to a concentration of 15-25 ng/µL. Purity was assessed using a Nanodrop 2000c and deemed acceptable with 260/280 ratio's between 1.7 - 1.8.

HTT amplification for fragment analysis

We chose to test these potential biomarkers against 2 unknown fibroblast lines, each having one parent affected by HD but never having been tested for confirmation of being affected by HD. Homemade primers were synthesized by Integrated DNA Technologies. The forward primer sequence was 5'-ATGAAGGCCTTCGAGTCCCTCAAGTC-3', with a carboxyfluorescein at its 5' end, and the reverse was 5'-CGGTGGCGGCTGTTGCTGCTGCTGCTGCTG-3' (Jama et al., 2013). Melting temperatures were 63.5°C and 72.7°C respectively. Amplification reactions were run in 20 µL volumes and contained the following components; 1X FailSafe Premix J (Epicenter, Madison, WI), 1U of Platinum® *Taq* DNA Polymerase High Fidelity (Life Technologies), 1.0 µL of 20-25 ng/µL gDNA sample, and 0.5 µM of each primer. PCR conditions were as follows; 95.0°C for 1 min hot start/initial denaturation, followed by 35 cycles of 94.0°C for 1 minute, 64.0°C for 1 minute, 72.0°C for 2 minutes, and a final extension of 15 minutes at 72.0°C. Each cell line had 10 µL loaded into a 2% w/v agarose gel for electrophoresis separation. Confirmation of correct amplicon size was visualized with ethidium bromide on a UV transilluminator (260 nm). After size confirmation with agarose gel, 1.0 µL of each sample was combined with 1.0 µL of Mapmarker 1000 internal size standard (BioVentures Inc., Murfreesboro, TN) and 9.0 µL Hi-Di Formamide (Life Technologies). Just before sample loading of 10 µL onto capillary electrophoresis plates, each sample was heated to 95.0°C for 2 minutes for denaturation and then immediately placed on ice for 2 minutes. Capillary electrophoresis fragment analysis was conducted on an Applied Biosystems 3130

Genetic Analyzer (Applied Biosystems) with 36 cm arrays and POP-7™ Polymer-CG (Applied Biosystems). Electrokinetic injection of samples was initiated with 12 kV for 16 sec, and then electrophoresed at 15 kV for 1200 sec at 60°C. Raw data was analyzed using GeneMarker Version 2.6.4 software (SoftGenetics LLC, State College, PA) (Jama et al., 2013). We conducted amplified fragment length polymorphism analysis using optimized macros for TNR expansion calling and labeling.

Bisulfite conversion and meCpG analysis

Genomic DNA samples were treated to obtain bisulfite conversion using the EpiTect Fast Bisulfite Kit (Qiagen) according to the manufacturer's protocol. In each reaction, 2 µg of gDNA was converted for each sample. Genomic DNA quality was determined by 1.5% agarose gel electrophoresis. Homemade primers were created to evaluate methylation patterns approximately -500 base pairs upstream of the respective genes' coding start sight using MethPrimer Software (Li and Dahiya, 2002). Oligonucleotide primers and Tm's are listed in Table 5. PCR amplification of bisulfite treated gDNA was conducted using 1U of Platinum Taq DNA Polymerase High Fidelity (Life Technologies), 500 ng of gDNA template, 1X High Fidelity Buffer, 0.4 mM of each dNTP, 2 mM $MgSO_4$, 0.5 µM of each forward and reverse primer, 2 mM betaine, with nuclease-free water in 20 µL reactions. We utilized a touch-up PCR cycling protocol where the annealing temperature was gradually increased after the first and second set of 5 cycles, with annealing temperatures increasing from 53°C to 56°C and finally 60°C for the remaining 30 cycles. PCR cycling conditions were as follows; 95°C for 1 min hot start, followed by a touch-up protocol with increasing annealing temperatures for 1 min, 72°C for 1 min for 40 total cycles. A final extension of 10 minutes at 72°C was applied at the end. Correct amplicons were assessed using gel electrophoresis on a 1.5% agarose gel. Appropriate bands were extracted using QIAquick Gel Extraction Kit (Qiagen). Each extracted amplicon was into the pCR4-TOPO TA vector with the TOPO TA cloning kit for sequencing (Thermo Fisher). Cycle sequencing of fragments was conducted using the M13 forward and reverse

primers provided with the kit. Methylation analysis and representation was conducted using online software *Methylation Plotter* (Mallona et al., 2014).

Data analysis and statistical significance

Messenger RNA expression of fibroblasts was independently conducted three times with each experiment sub-sampled in triplicate. Raw cycle-threshold (Ct) values of each sub-sample were averaged to give one mean Ct. Mean Ct values of tested genes was normalized against the housekeeping gene *ACTB* to determine ΔCt values. Each gene ΔCt value was then normalized to wild-type fibroblast sample GM-02187 to determine ΔΔCt's. Fold expression was decided by $2^{-\Delta\Delta Ct}$. Error bars were determined using the standard deviation of experimental triplicate variation between fold-expression.

Fluorescent image analysis was conducted using *ImageJ* software. *ImageJ* provided unbiased automatic background subtraction for each channel. Cell fluorescence was determined by multiplying the total area of the image by mean background of the image, then subtracted from integrated density. The cell fluorescence of each image was then divided by the number of complete nuclei in the image to give the corrected total cell fluorescence (CTCF). Mean CTCF is represented graphically as normalized values to the mean CTCF of wild-type fibroblast sample GM02187 (Wt-1). Error bars represent the standard deviation of individual CTCF obtained in each image.

Analysis of *HTT* gene expansion raw data was conducted using *GeneMarker* software (State College, PA). Macro's for bin calling were customized using gDNA samples obtained from Coriell Cell Repositories and were graciously created by Mohamed Jama (University of Utah). The macros created algorithms that binned the most prominent fluorescence peaks of a specific wavelength within a -3.0 to +3.0 base pair range, standardized by an internal size standard fluorescence ladder of a specific wavelength.

Data for multiple comparisons were subjected to ANOVA statistical analysis with Dunnett's test post-hoc to determine statistically significant differences

between samples. Two way analysis was conducted using student t-tests. Statistical significance was set as $p<0.05$.

Table 5

APEX1 primer set for bisulfite amplification and capillary electrophoresis sequencing.

Gene (HGNC-ID)	Segment	Forward Primer	Reverse Primer	Tm (°C) Fwd	Tm (°C) Rev
APEX1 (587)	First	ATAATGTGTTGTGTATTTGGTATAA	ATAACAACCTAACTCCTCCTAACCC	51.36	58.05
	Second	TGTTAGGAGTTGTGGAGGTTTTTAT	TTCTCTAAAACACTTAAATCCCCAA	58.12	58.23

[α]Using MethPrimer and Methyl Primer Express software, bisulfite reaction primers were designed and synthesized to target approximately -500 bp upstream of the start sight of each investigated gene. The primers were utilized in promoter amplification.

Results

APEX1, BRCA1, and RPA1 deficient gene expression identified in HD and unknown fibroblasts

To investigate the wide spread cellular effects of mHtt on DNA repair gene expression, we extracted mRNA from confirmed HD fibroblasts and unknown fibroblasts. A total of 7 primary fibroblast lines were obtained from Coriell cell repositories, with 2 unaffected samples (GM02187 and GM02153), 3 previously confirmed HD-affected samples (GM02079, GM03621, and GM03868), and 2 unknown samples (GM02079 and GM04022). Cell lines are in figures are designated “HD-n” if they were previously confirmed to have HD. Cell lines designated “pHD-n” are cell lines in which they have a 50% probability of being affected by HD, but were initially unknown to us. Gene expression analysis revealed statistically significant down-regulation of all three genes in all HD-affected and unknown fibroblast lines (Fig. 10). The mRNA levels presented are in relation to control sample Wt-1 (GM02187), with gene expression levels of targeted genes arbitrarily set as 1.

In addition, we conducted immunocytochemistry of all fibroblast lines and protein content reflected the mRNA reduction (Figs. 11-14). Corrected total cell fluorescence (CTCF) revealed that protein levels in pHD-1 and pHD-2 samples were approximately 23% of APEX1, 22% of BRCA1, 8% of RPA1 and 32% of RPA3 (Fig. 15), and 54% of APEX1, 35% of BRCA1, 46% of RPA1, and 36% of RPA3 (Fig. 16) respectively. Protein levels are relative to the lowest expressing wild-type fibroblast sample. These data demonstrate that DNA repair components is significantly decreased in HD fibroblast samples, possibly having consequences on DNA repair pathways to efficiently address lesions and adducts created during DNA synthesis.

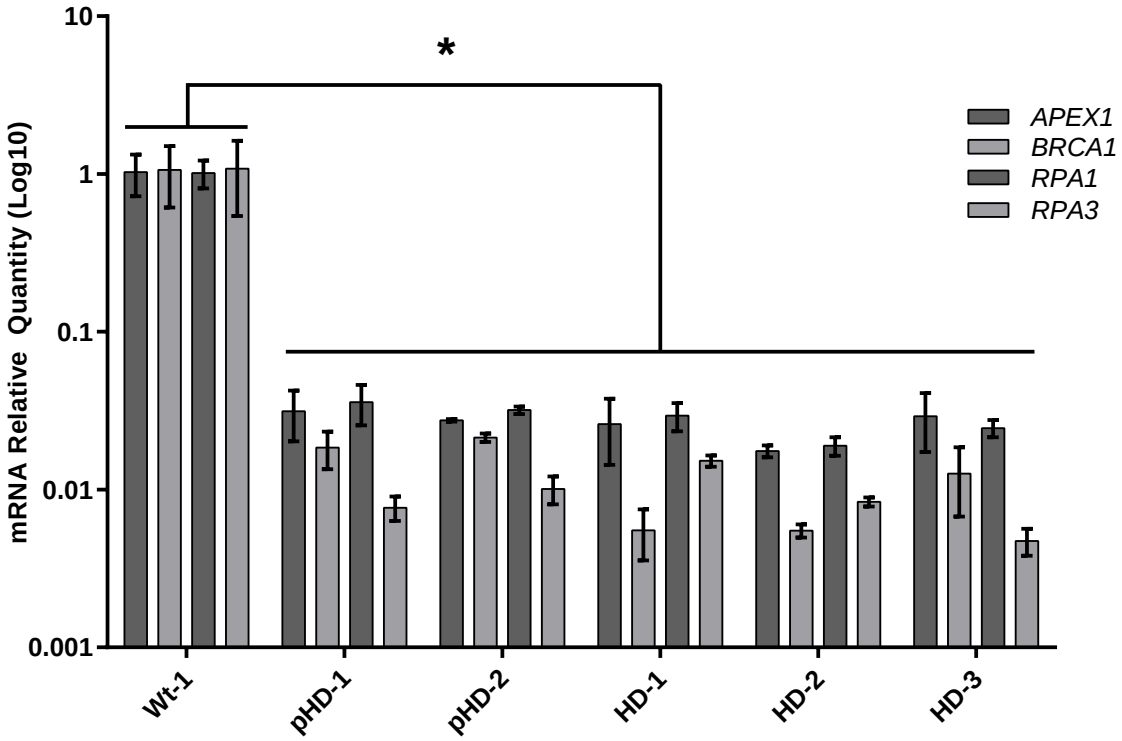


Fig. 10. DNA repair deficiencies in multiple HD fibroblast cell samples. *APEX1, BRCA1, RPA1* and *RPA3* expression in wild-type (Wt-1), potential-HD-affected (pHD), and HD-affected (HD) fibroblasts. Using TaqMan® Gene Expression Assays, we identified HD DNA repair mRNA down-regulation of *APEX1, BRCA1, RPA1,* and *RPA3.* Gene expression is determined by the ($2^{-\Delta\Delta Ct}$) of the mean of each independent triplicate. Error bars represent the standard deviation of the mean between replicate samples. Each tested gene showed significant decreases in expression for all HD and pHD cell lines, relative to the wild-type sample. Statistical significance was determined using a two-tailed, homoscedastic T-Test (*$p<0.001$). Gene expression is normalized to wild-type (Wt-1) fibroblast line GM02187 gene expression.

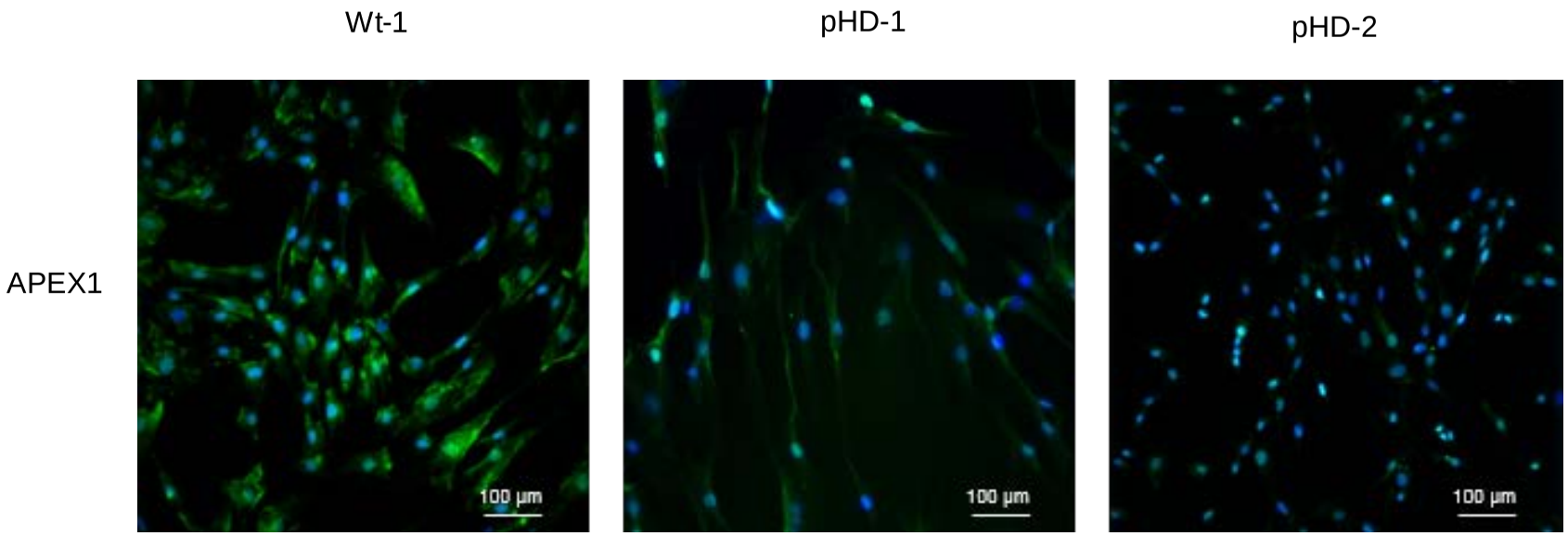


Fig. 11. APEX1 protein deficiencies in HD fibroblast cell lines. The above images are Wt-1 sample GM02187, HD-1 fibroblast sample GM02191, and HD-2 fibroblast sample GM04022. Each cell line was plated on chamber slides and grown for 72 hours. Slides were fixed with 100% methanol for 20 min at -20°C. After fixation, each slide was stained with primary antibodies detecting APEX1 protein. Primary antibodies were counterstained with secondary conjugated Alexa Fluor-488. Nuclei staining was conducted with DAPI. Images were acquired using a Zeiss Axio microscope with a 10X objective.

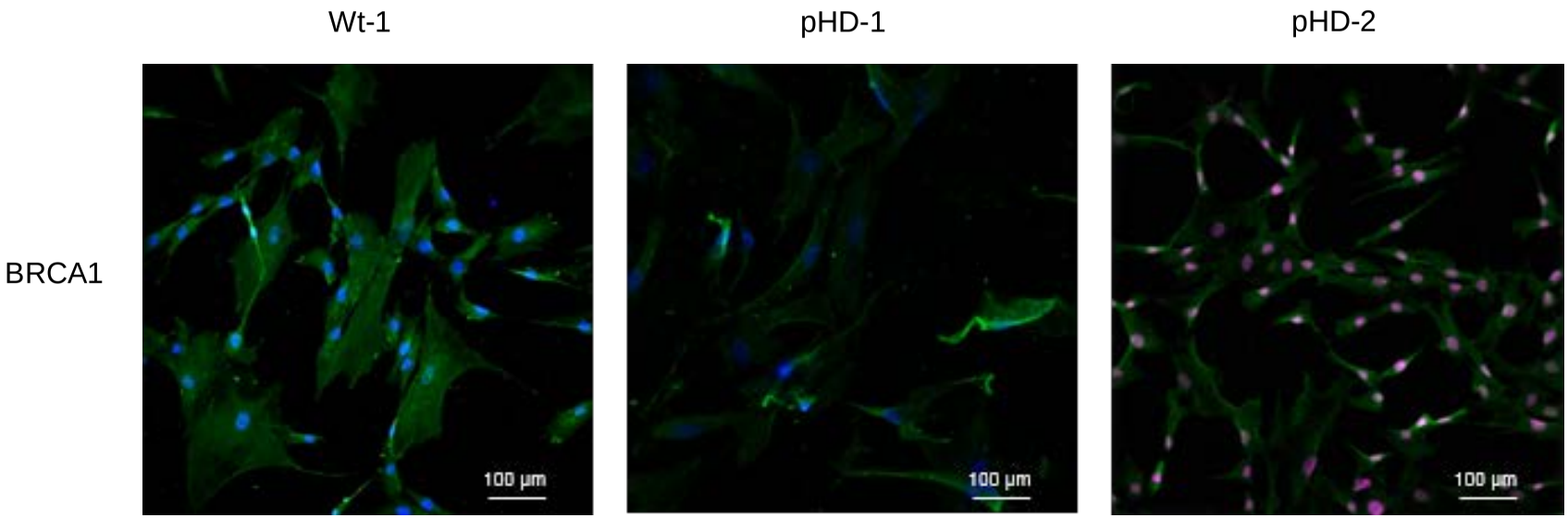


Fig. 12. BRCA1 protein deficiencies in HD fibroblast cell lines. The above images are Wt-1 sample GM02187, HD-1 fibroblast sample GM02191, and HD-2 fibroblast sample GM04022. Each cell line was plated on chamber slides and grown for 72 hours. Slides were fixed with 100% methanol for 20 min at -20°C. After fixation, each slide was stained with primary antibodies detecting BRCA1 protein. Primary antibodies were counterstained with secondary conjugated Alexa Fluor-488. Nuclei staining was conducted with DAPI. Images were acquired using a Zeiss Axio microscope with a 10X objective.

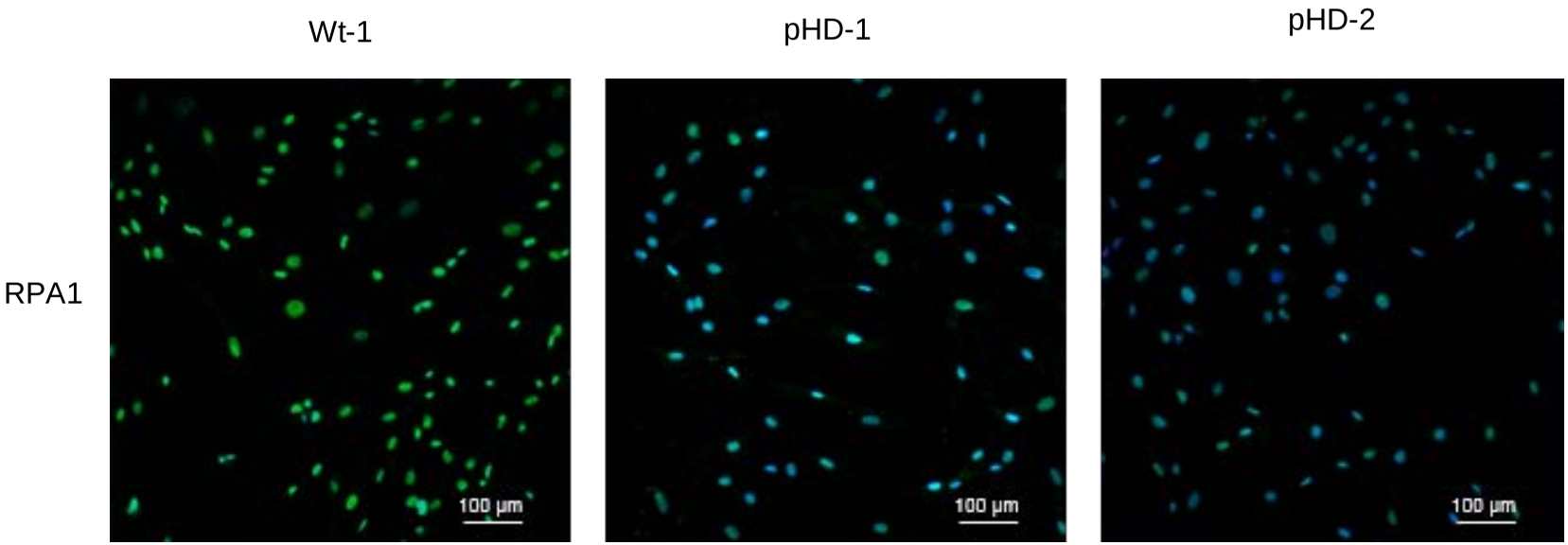


Fig. 13. RPA1 protein deficiencies in HD fibroblast cell lines. The above images are Wt-1 sample GM02187, HDp1 fibroblast sample GM02191, and HD-2 fibroblast sample GM04022. Each cell line was plated on chamber slides and grown for 72 hours. Slides were fixed with 100% methanol for 20 min at -20°C. After fixation, each slide was stained with primary antibodies detecting RPA1 protein. Primary antibodies were counterstained with secondary conjugated Alexa Fluor-488. Nuclei staining was conducted with DAPI. Images were acquired using a Zeiss Axio microscope with a 10X objective.

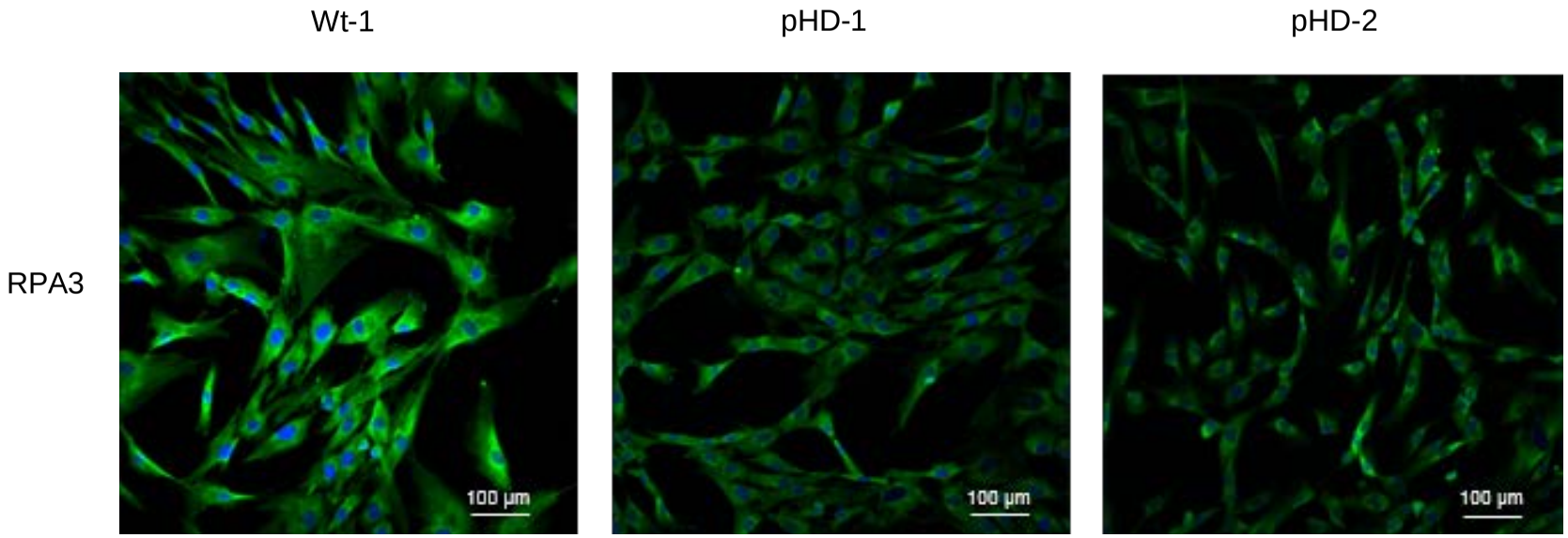


Fig. 14. RPA3 protein deficiencies in HD fibroblast cell lines. The above images are Wt-1 sample GM02187, HD-1 fibroblast sample GM02191, and HD-2 fibroblast sample GM04022. Each cell line was plated on chamber slides and grown for 72 hours. Slides were fixed with 100% methanol for 20 min at -20°C. After fixation, each slide was stained with primary antibodies detecting RPA3 protein. Primary antibodies were counterstained with secondary conjugated Alexa Fluor-488. Nuclei staining was conducted with DAPI. Images were acquired using a Zeiss Axio microscope with a 10X objective.

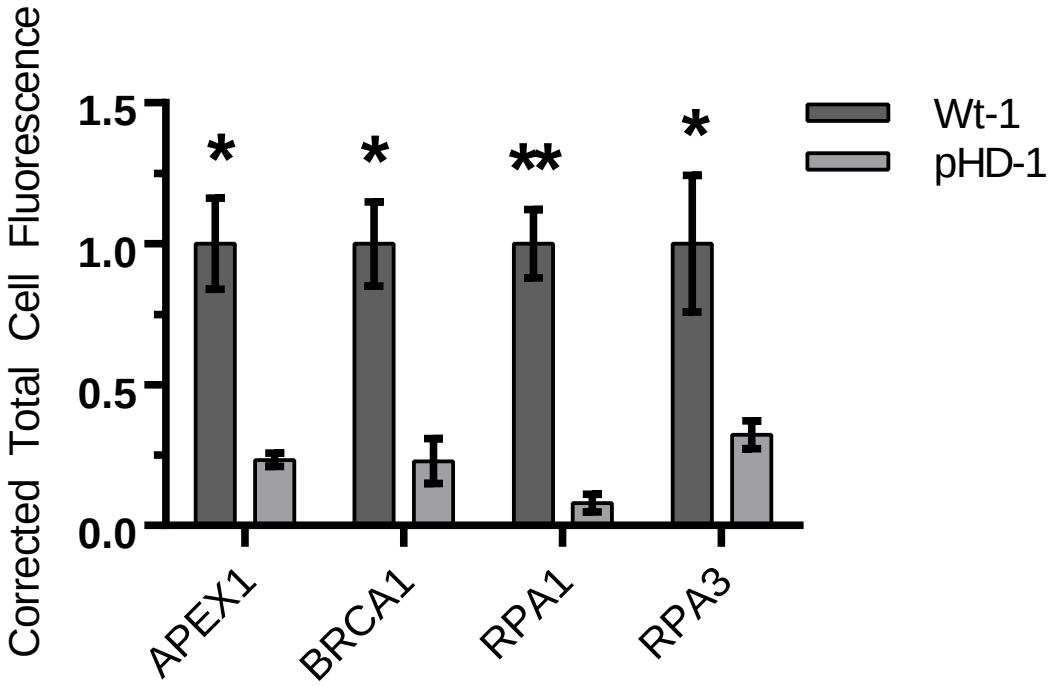


Fig. 15. Immunofluorescence measurement of protein deficiencies in pHD-1 fibroblast cell line. Measuring cell fluorescence, semi-quantitative protein analysis of pHD-1 revealed down-regulation of investigated DNA repair proteins. Total fluorescence of cell area was measured and then normalized against the wild-type sample. Error bars represent the standard deviation of the mean corrected total cell fluorescence. Statistical significance was determined using a two-tailed, homoscedastic T-Test (*p<0.01, **p<0.001).

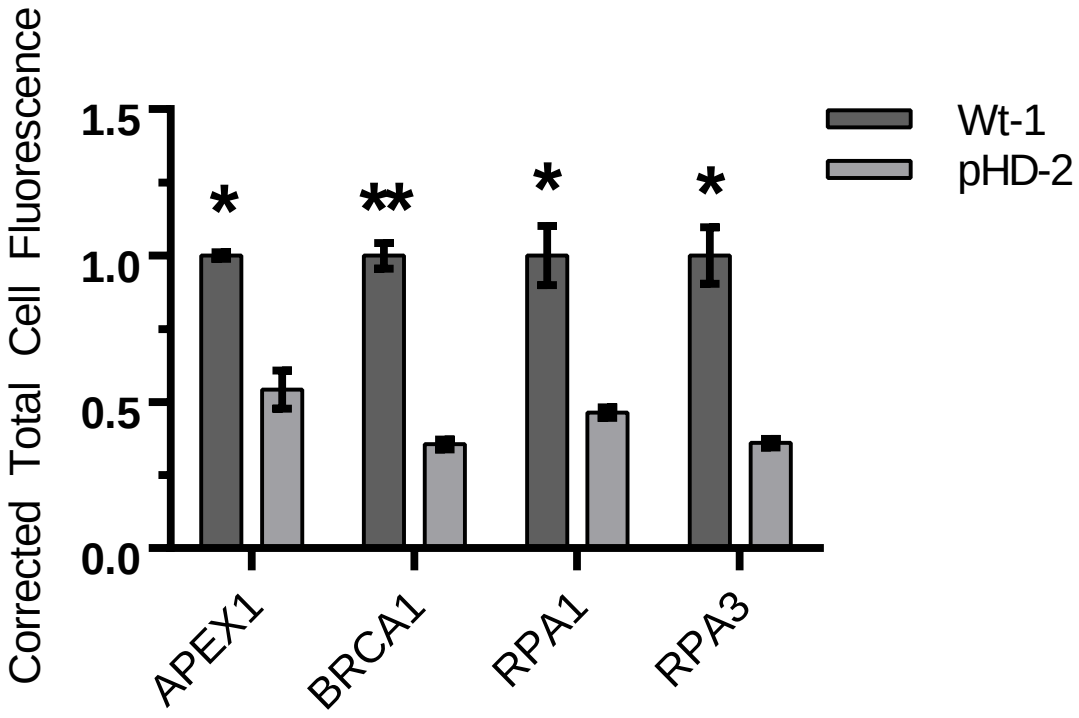


Fig. 16. Immunofluorescence measurement of protein deficiencies in pHD-2 fibroblast cell line. Measuring cell fluorescence, semi-quantitative protein analysis of pHD-2 revealed down-regulation of investigated DNA repair proteins. Total fluorescence of cell area was measured and then normalized against the wild-type sample. Error bars represent the standard deviation of the mean corrected total cell fluorescence. Statistical significance was determined using a two-tailed, homoscedastic T-Test (*p<0.01, **p<0.001).

Unknown fibroblast samples are affected by HTT expansion

Two cell lines were purchased from Coriell Cell Repositories that had a 50% probability of being affected by HD (pHD-1 and pHD-2), but were not previously confirmed as either normal or affected. As positive controls we obtained 6 gDNA samples from various HD-affected and wild-type characterizations. These chosen positive control gDNA samples have previously been tested in multiple laboratories and were sequenced to confirm CAG repeats (Data not shown). To determine HD status of unknown fibroblast lines we conducted capillary electrophoresis fragment analysis on all samples (Figs. 17A-G). Characterization of lines are summarized below (Table 6). Both unknown fibroblast lines GM02191 (pHD-2) and GM04022 (pHD-1) were confirmed to be affected by HD. Fragment analysis revealed that the HTT fragment length varied based upon the pathogenic $(CAG)_n$ expansion. Examination of the fragments revealed unknown samples GM02191 containing 17/42 (CAG) repeats and GM04022 containing 18/44 (CAG) repeats at the initial time of capillary electrophoresis. Combined with data identifying *APEX1, BRCA1, RPA1* and *RPA3* down-regulation, this confirmation of unknown samples being affected by HD identifies therapeutic biomarkers, as well as a possible alternate means for HD diagnosis.

Interestingly, we witnessed TNR instability in previously characterized HD samples HD-1 and HD-2. Previously Coriell Cell Repositories reported *HTT* gene characterization of HD-1 as containing 21/44 (CAG)'s and HD-2 containing 18/44 (CAG)'s. Our initial characterization was conducted after 10 population doublings and revealed *HTT* expansions of 21/45 and 18/48, respectively. No instability was detected in either wild-types samples Wt-1 or Wt-2 after identical population doublings.

A

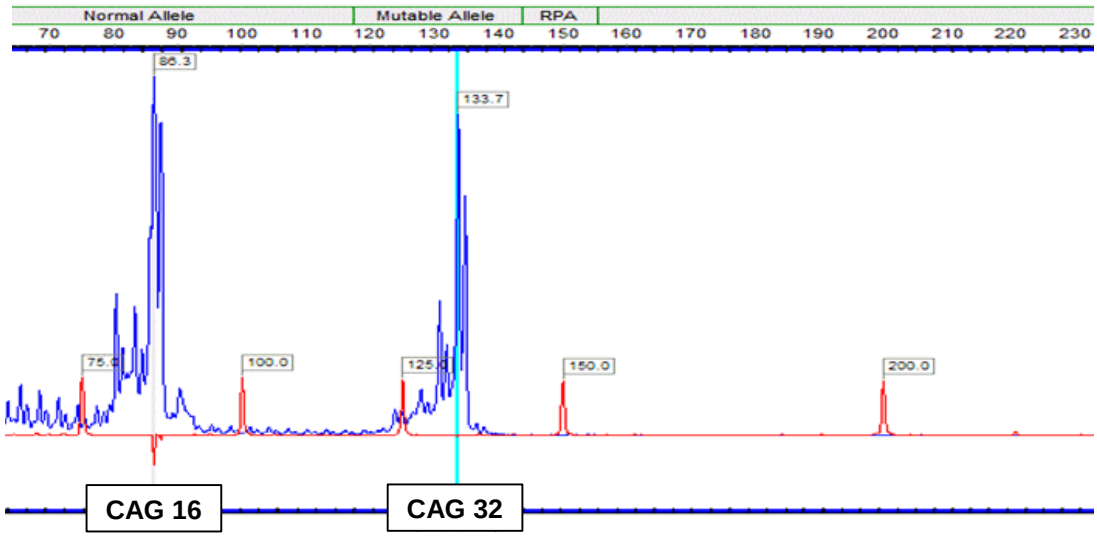


Fig. 17. Fragment analysis characterizing of all fibroblast lines. Using triple-primed polymerase chain reaction in combination with capillary electrophoresis, UNK fibroblast samples show HD (CAG) expansions within the gene. Wild-type samples used in our experiments are (A) Wt-2 GM02153 with 16/32 (CAG)'s, and (B) Wt-1 GM02187 with 17/23 (CAG)'s. HD-confirmed cell lines (C) HD-1 (GM02079) contain 21/45, (E) HD-2 (GM03621) contain 18/60, and (G) HD-3 (GM03868) contain 17/48. Cell lines (D) pHD-2 (GM02191) and (F) pHD-1 (GM04022) were purchased from Coriell Cell Repositories as described as "Unknown – 50% probability of being affected". Fragment analysis reveals both unknown cell lines are in fact HD-affected, with alleles containing 17/42 and 18/44, respectively. Interestingly, TP-PCR also shows TNR instability with a subpopulation of gDNA extracted from HD each sample, relative to their previously published expansions by Coriell. GM02191 shows expansions up to 45 (CAG)'s in the affected allele, and GM04022 show expansions up to 47 (CAG)'s. Alleles containing $(CAG)_{>35}$ are diagnosed HD. ROX1000 MapMarker dye serves as an allelic ladder and is indicated by the red peaks.

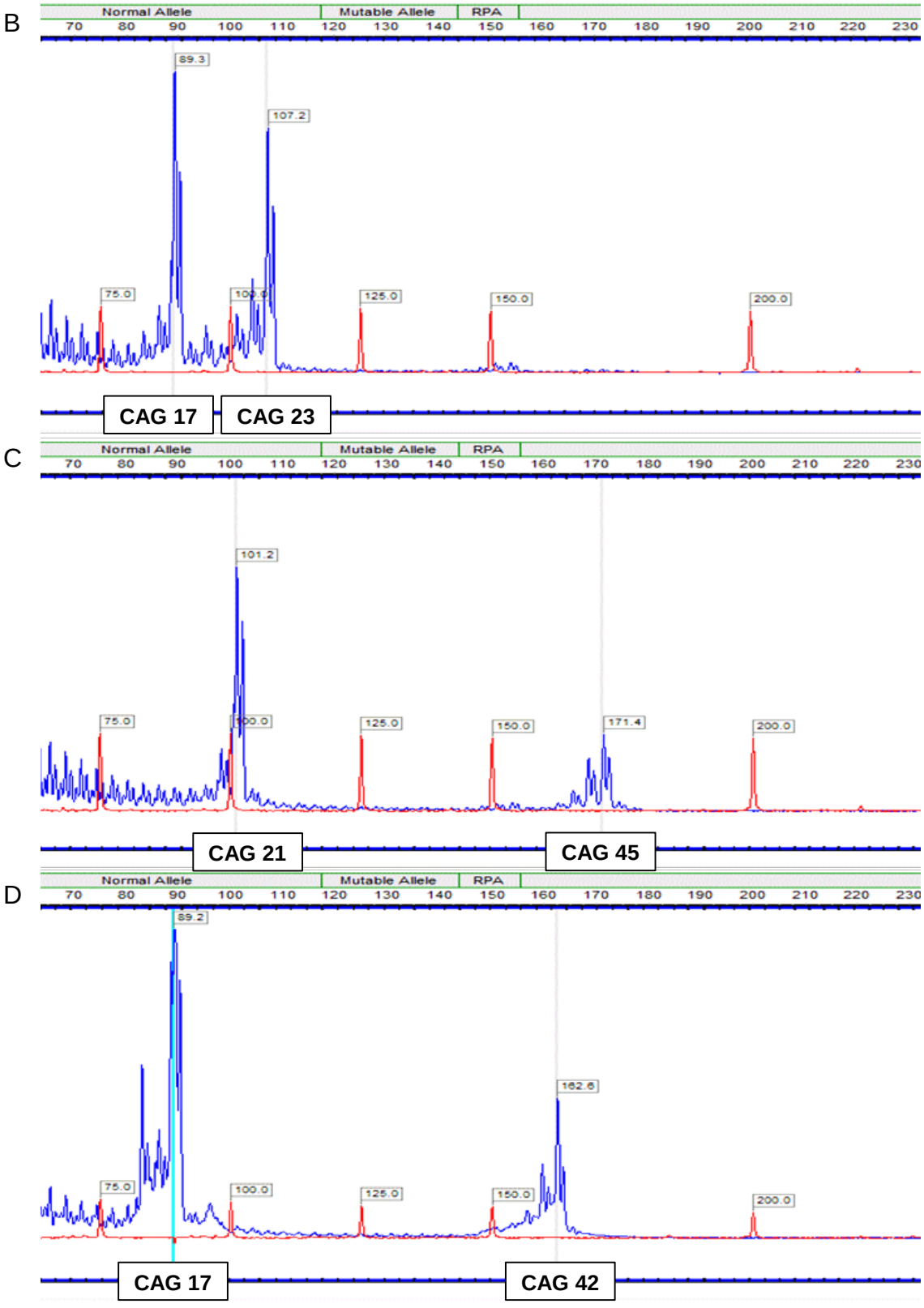


Fig. 17. Continued.

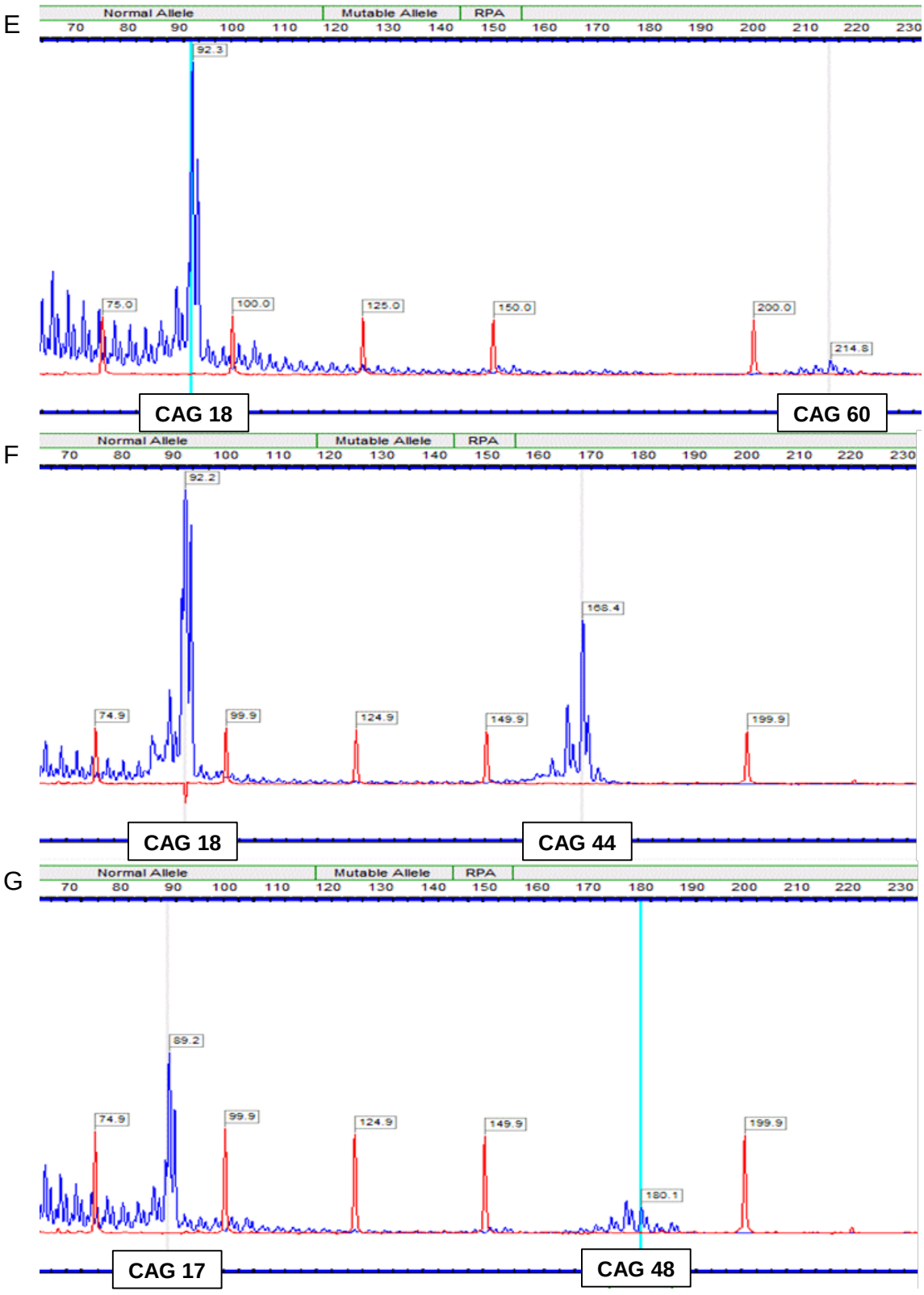


Fig. 17. Continued.

Table 6

Genetic characteristics of HD and wild-type fibroblast lines.

Cell Lines	Experiment Designation	Age (years)	Reported Status	Reported $(CAG)_n$	Confirmed Status	Confirmed $(CAG)_n$
GM02187	Wt-1	60	Wild-type	17-23	Wild-type	17-23
GM02153	Wt-2	40	Wild-type	16-32	Wild-type	16-32
GM02079	HD-1	48	HD	21-44	HD	21-45
GM03621	HD-2	29	HD	18-60	HD	18-60
GM03868	HD-3	45	HD	18-44	HD	18-48
GM04022*	pHD-1	28	Unknown	-	HD	18-44
GM02191*	pHD-2	32	Unknown	-	HD	17-42

[α]HD-affected and wild-type fibroblast samples are all female. All cell lines were purchased from the Coriell Cells Repository. The status of fibroblast lines GM02191 and GM04022 were designated “unknown” upon ordering and initial experiments because they had not yet been characterized and were subsequently confirmed to be HD-affected through capillary electrophoresis fragment analysis. Experimental designation is the abbreviation used through-out the manuscript. The (CAG) expansions shown above represent the characterizations previously published by Coriell Cell Repository, with the exception of our “unknown” sample in which the (CAG) is the first characterization by fragment analysis completed. This finding is denoted by *.

APEX1, BRCA1 and RPA1 gene expression is affected by 5-azacytidine treatments, but not by trichostatin-A

We next wanted to identify the mechanism causing down-regulation in the identified DNA repair genes. As no previous DNA repair gene mutational events have been associated with HD, we hypothesized that epigenetic regulators could be a primary cause of the observed down-regulations. To test for the involvement of histone acetylation and DNA condensation as a primary regulator we treated pHD-1 and pHD-2 samples with 5 µM trichostatin-A (TSA), a potent global de-acetylase inhibitor for 24 h. For controls, the cells lines were treated with 0.01% v/v DMSO for identical time treatment. Quantitative PCR revealed no significant DNA repair gene change in both pHD-1 (Fig. 18A) and pHD-2 (Fig. 18B) in response to TSA treatments, with the vehicle control showing minor fluctuations relative to an untreated control sample.

Similar to TSA treatments, we chose to investigate the effects of 5-azacytidine to determine possible methylation involvement in expression. pHD-1 and pHD-2 samples were seeded and cultured in the presence of culture media supplemented with 10 µM 5-azacytidine for 120 h. Media was replaced daily with freshly made 5-azacytidine to optimize passive demethylation during DNA synthesis and to facilitate the removal of toxic products. After 5 days, mRNA was collected and subjected to qPCR to evaluate gene expression. HD samples GM02191 and GM04022 showed significant up-regulation of *APEX1, BRCA1,* and *RPA1* (Fig. 19A, 19B). Interestingly, HD sample GM02191 did not show a change in the expression of *RPA3* mRNA, however. To further confirm expression, ICC of each cell line was conducted (Fig. 20A, 20B). Corrected total cell fluorescence of ICC validated gene up-regulation, with pHD-1 showing increases in APEX1, BRCA1, RPA1, and RPA3 (Fig. 21A) with fold changes of 1.8, 1.4, 1.2, and 1.4 respectively. pHD-2 showed similar fold increases with 1.6, 1.5, 2.0, and 1.2 (Fig. 21B). Error bars are the SD of the mean CTCF of at least 3 images of each treated/untreated sample. Statistical significance was determined using a two-tailed, unpaired, parametric T-Test (*p<0.05, **p<0.01). Overall, these data suggests the DNA repair gene expression in HD fibroblasts is affected by 5-

azacytidine treatment. Though no definitive research has shown that mHtt interacts with/ or causes hypermethylation, our results indicate that brief treatments of 5-azacytidine will up-regulate DNA repair gene expression in HD fibroblasts. While we recognize that gene expression levels are not returned to wild-type expression levels, these data suggests that hypomethylation by 5-azacytidine is at least partially responsible for the DNA repair gene deficiency we report in this work.

Differential methylation patterning in APEX1 promoter in HD fibroblasts

To confirm 5-azacytidine treatments caused passive hypomethylation, we conducted bisulfite sequencing analysis of ~500 base pairs upstream of the *APEX1* transcriptional start site. We also included ~75 base pairs downstream from the start site. gDNA from 5-azacytidine treated pHD-1 (T-pHD-1), treated pHD-2 (T-pHD-2), untreated pHD-1 (U-pHD-1), and untreated pHD-2 (U-pHD-2) underwent a bisulfite reaction to replace all cytosine's that were not methylated. As well, samples from 5-azacytidine treated Wt-1 (T-Wt-1) and untreated Wt-1 (U-Wt-1) samples were also included. Samples were analyzed using ABI Genetic Analyzer 3130 to determine the presence of meCpG sites. Treated Wt-1 (T-Wt-1) sample and untreated Wt-1 (U-Wt-1) bisulfite sequencing of the promoter revealed no methylation in the -500 to +75 base pair region relative to the transcriptional start site (Figure 22). Treated samples T-pHD-1 and T-pHD-2 showed no methylation patterning, which is expected after treatment of DNA hypomethylation agent 5-azacytidine. Untreated samples U-pHD-1 and U-pHD-2 revealed differential methylation patterns relative to U-Wt-1. U-pHD-1 revealed methylation sites at base pairs -437 and -28. U-pHD-2 showed methylated CpG's at base pairs – 435, -427, -417, -412, -138, -68, -28, and -7, upstream of the transcriptional start sight. Not only does these data confirm that 5-azacytidine treatments caused direct hypomethylation through passive cell replication, but this also brings to light the differential methylation patterns seen in HD-fibroblasts. This differential methylation in the promoter region of *APEX1* elucidates the possible cause of gene down-regulation seen in our gene expression data.

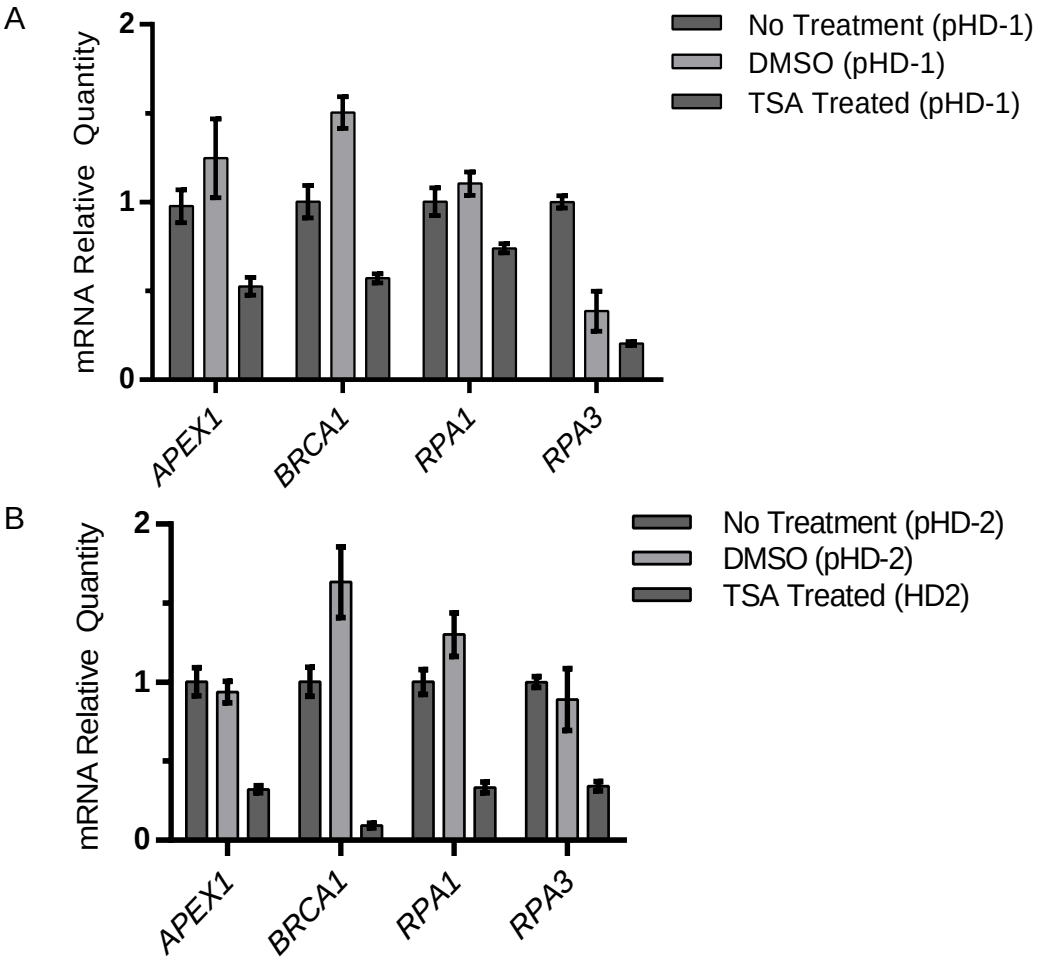


Fig. 18. DNA repair gene mRNA expression with trichostatin-A treatment. (A) pHD-1 fibroblasts and (B) pHD-2 fibroblasts were exposed to 5 µM trichostatin-A for 24h. As well, a separate sample of cells was treated with 0.01% v/v of DMSO as a vehicle control. Gene expression is determined by the mean ($2^{-\Delta\Delta Ct}$) of each independent replicate, relative to the untreated controls. Each tested gene showed alarming decreases in expression for the pHD-1 and pHD-2 fibroblast samples in response to trichostatin-A treatment. Gene expression is normalized to the same untreated-HD fibroblast gene expression. Error bars represent the standard deviation of the mean. Statistical significance was determined using a two-tailed, unpaired, parametric T-Test (*p<0.05, **p<0.01).

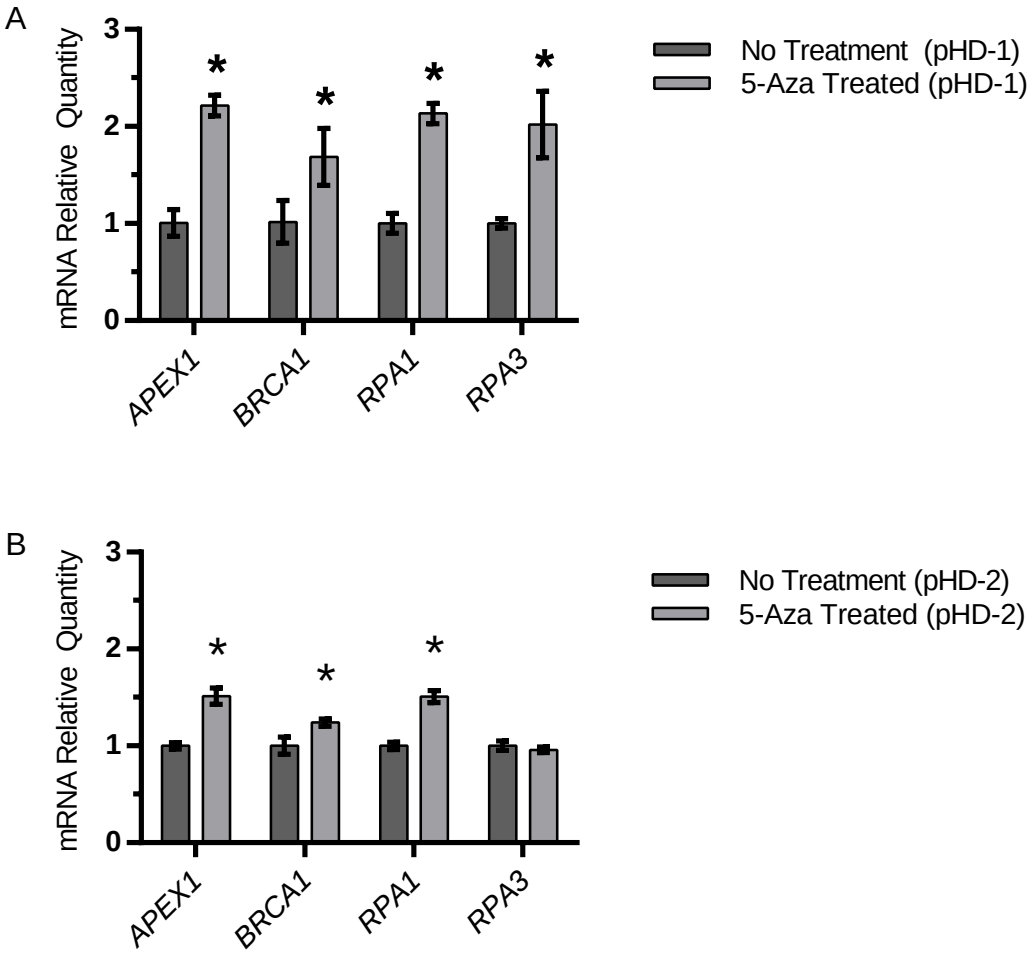


Fig. 19. DNA repair gene mRNA expression with 5-azacytidine treatment. (A) pHD-1 fibroblasts and (B) pHD-2 fibroblasts were exposed to 10 µM 5-azacytidine for 120 h. Media was changed daily to ensure adequate amounts for passive demethylation and removal of toxic products. Each tested gene showed significant increases in mRNA expression for the pHD-1 treated fibroblast line. *APEX1, BRCA1, and RPA1* showed statistically significant increases in pHD-2 treated cells. Gene expression is normalized to the same untreated-HD fibroblast gene expression. Error bars represent the standard deviation of the mean. Statistical significance was determined using a two-tailed, unpaired, parametric T-Test (*p<0.05, **p<0.01).

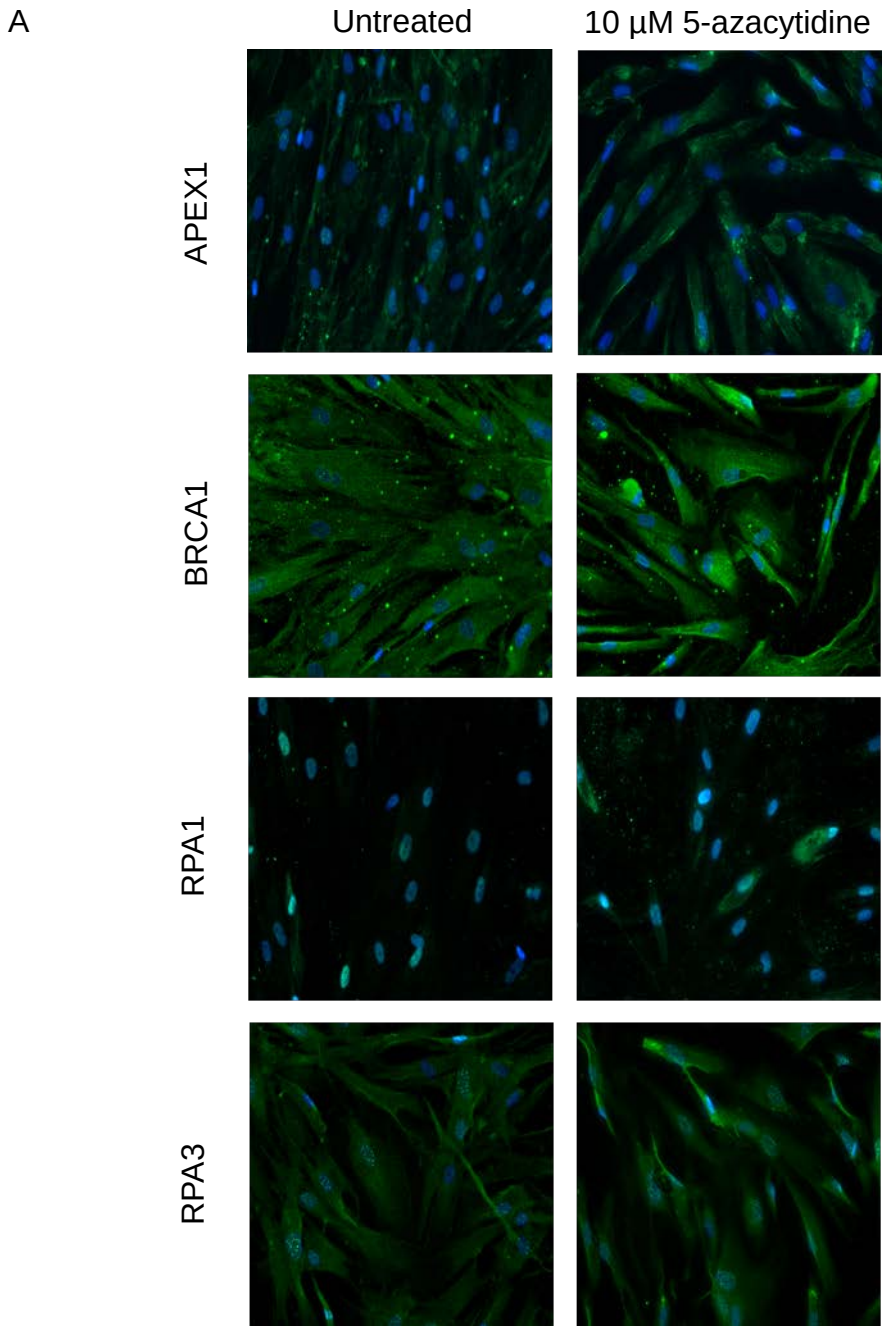


Fig. 20. Immunocytochemistry of DNA repair genes in pHD-1 cells treated with 10 µM 5-azacytidine. pHD-1 (A) and pHD-2 (B) fibroblast lines were exposed to 10 µM 5-azacytidine for 120 h, alongside untreated controls (right and left columns respectively). ICC images show slight increases in fluorescence of DNA repair genes. Nuclei staining was conducted with DAPI. Images were acquired using a Zeiss Axio microscope with a short-working distance 20X objective.

B

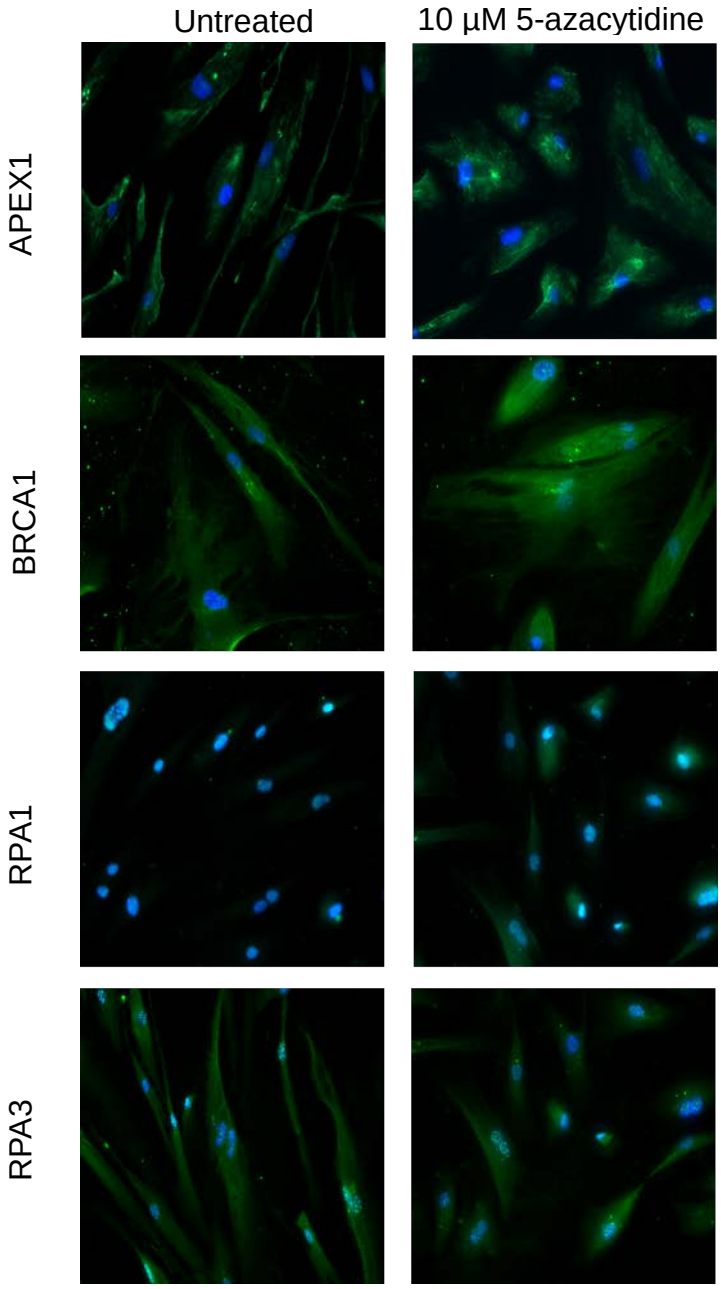


Fig. 20. Continued.

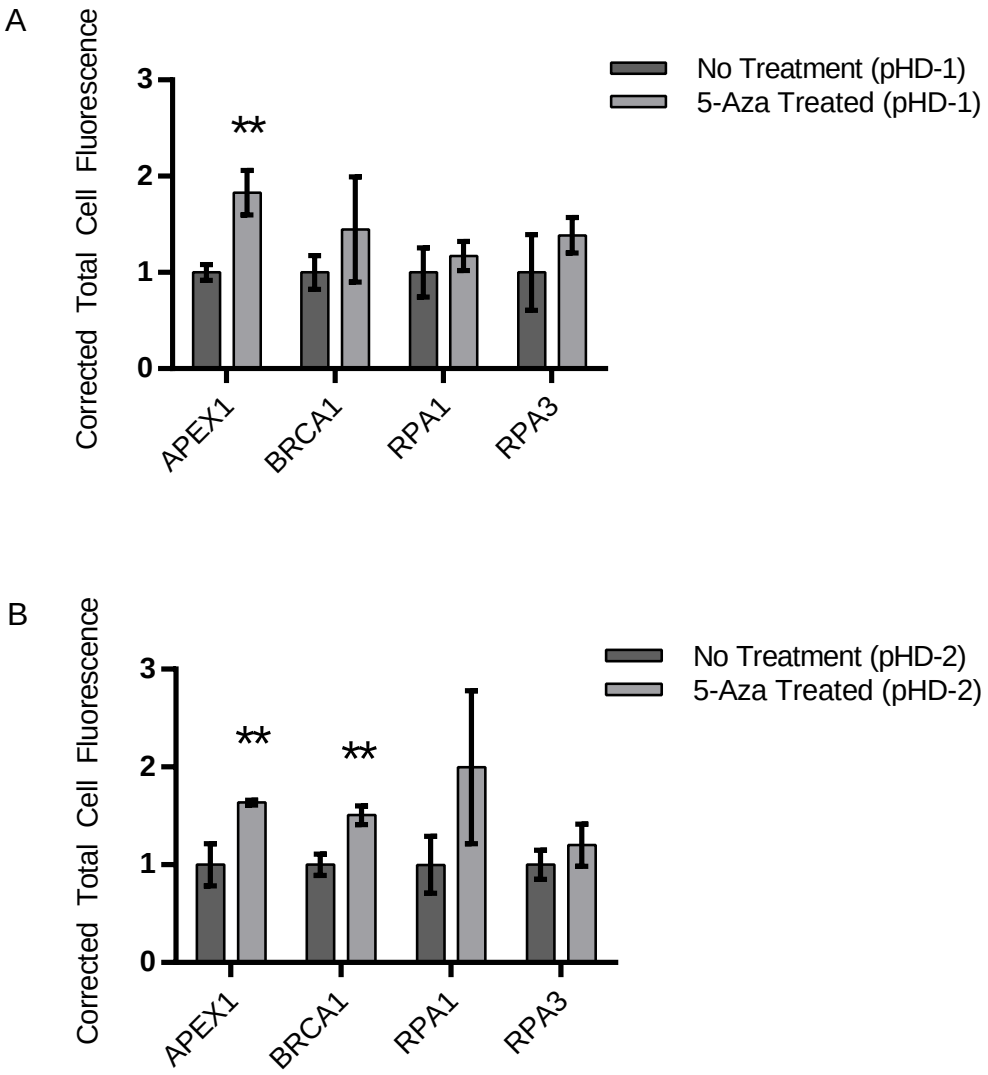


Fig. 21. Immunofluorescence measurement of DNA repair gene protein expression with 5-azacytidine. Protein quantity was determined using corrected total cell fluorescence of the collected cell. Analysis revealed statistically significant up-regulation of APEX1 for 5-azacytidine treated pHD-1 (A) fibroblasts. pHD-2 (B) fibroblasts treated with 5-azacytidine showed statistically significant up-regulation of APEX1 and BRCA1. Bars represent the mean CTCF of at least three pictures. Error bars represent the standard deviation of the mean. CTCF is relative to the untreated sample. Statistical significance was determined using a two-tailed, unpaired, parametric T-Test (*p<0.05, **p<0.01).

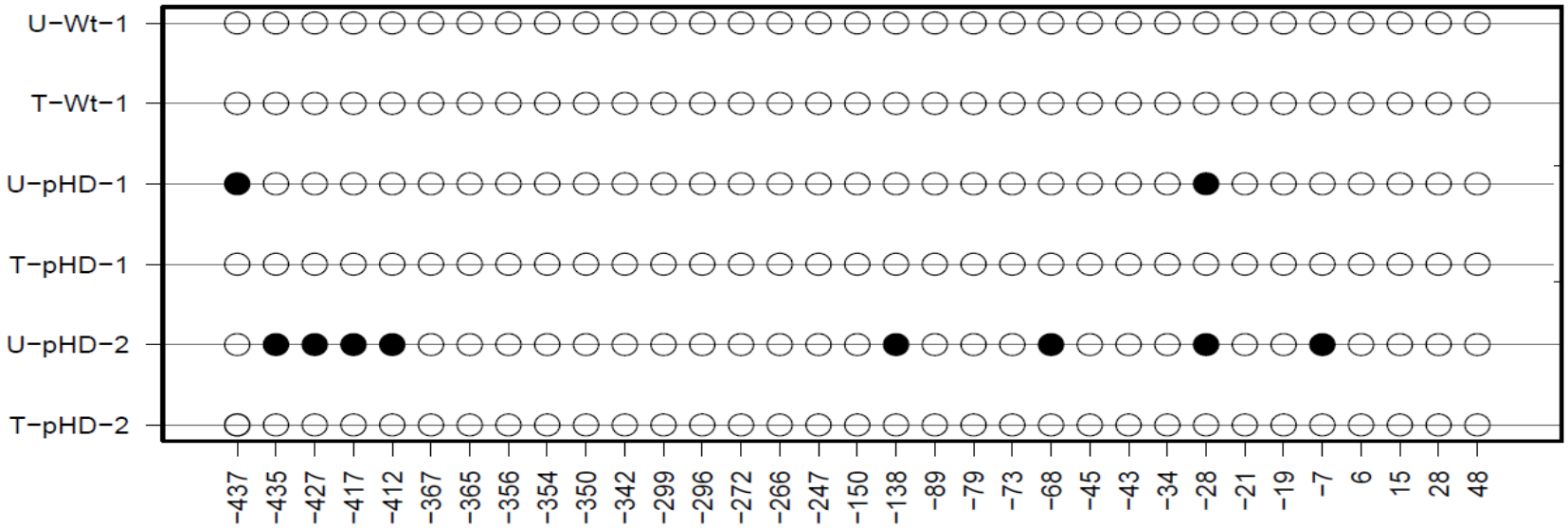


Fig. 22. Differential methylation pattern of *APEX1* CpG sites in promoter region. Genomic DNA from pHD-1 fibroblasts and pHD-2 fibroblasts that were exposed 10 µM trichostatin-A for 120h. Open and closed circles represent unmethylated and methylated CpGs, respectively. Bisulfite sequencing analysis was conducted using homemade primers. As well, gDNA from Wt-1 fibroblasts was analyzed as well. Treated Wt-1 (T-Wt-1) sample and untreated Wt-1 (U-Wt-1) sample promoter analysis revealed no methylation in the -500 base pair region upstream from the transcriptional start site. As well, T-pHD-1 and T-pHD-2 gDNA samples showed no methylation present. Untreated pHD samples (U-pHD-1 and U-pHD-2) showed slight methylation patterning. U-pHD-1 gDNA revealed methylation sites at base pairs -437 and -28. U-pHD-2 showed methylated CpG's at base pairs – 435, -427, -417, -412, -138, -68, -28, and -7, upstream of the transcriptional start sight. Methylation analysis was conducted using online software *Methylation Plotter* found at http://gattaca.imppc.org:3838/methylation_plotter/.

Intermittent 5-azacytidine treatments induces TRN stability in the HTT gene

Cell lines pHD-1 and pHD-2 were plated in T-75 flasks and treated intermittently with 10 µM 5-azacytidine over the course of 35 days to reach four population doublings. Genomic DNA samples were taken of pHD-1 and pHD-2 samples at the initial split before the treatment began, and at the end of the 35 day time-frame. In parallel to the treated samples, we passaged a sub-population of each sample as a non-treated control. pHD-1 sample began the treatment regimen with *HTT* gene affected allele containing 43 (CAG)'s, with a small population of cells showing expansions in the affected allele up to 44 and 45 (Fig. 23). At the end of treatment schedule, the untreated control pHD-1 sample showed expansions of predominantly 44, with a small population of 45 (CAG)'s. The pHD-1 sample treated intermittently with 5-azacytidine showed (CAG) repeats of 43, with a small population of 44. Similar results were witnessed in sample pHD-2. Initial expansion of the affected allele was characterized as having predominantly 41 (CAG)'s (Fig. 24). At the end of the treatment cycles, the untreated pHD-2 sample showed predominantly 42 (CAG)'s with instability within the population up to 45. The treated pHD-2 sample showed the majority of cells had 41 (CAG)'s, with a small subset of the cell population having 42. Representation of both alleles is shown (Fig. 25).

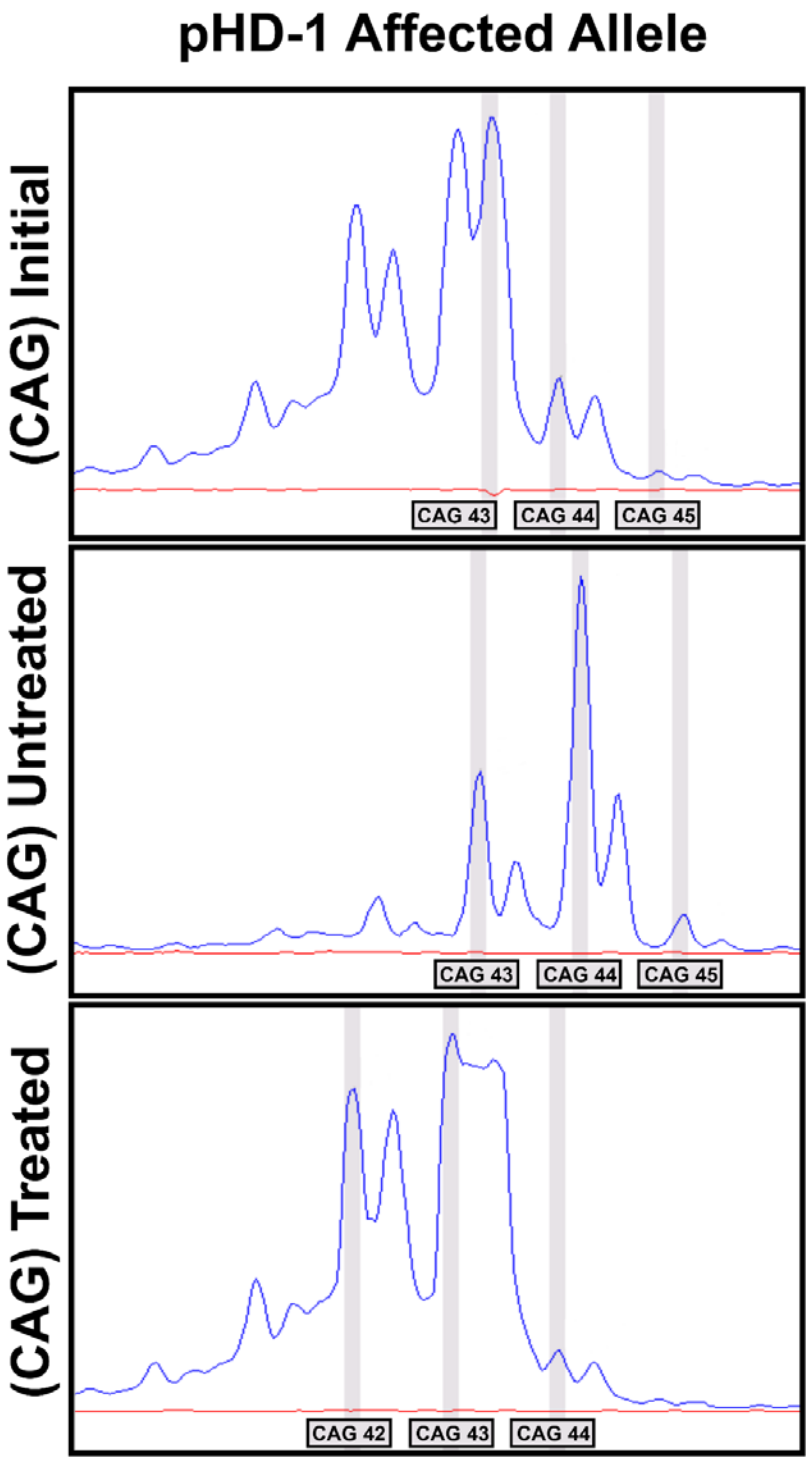


Fig. 23. Fragment analysis revealing 5-azacytidine induced TNR stability in 4 population doublings in pHD-1 fibroblasts.

Fig. 23. Continued. pHD-1 fibroblast cell lines were treated with 10 µM 5-azacytidine periodically over the course of 35 days. The top chromatogram of each image represents the (CAG) expansions at day 0. The middle chromatogram represents the (CAG) expansion of the untreated control at day 35. The bottom chromatogram shows the (CAG) expansion of the 5-azacytidine treated samples. pHD-1 fibroblasts show beginning and treated samples with (CAG) repeats of 17/43, with a small population of cells showing expansions up to 44-45. Untreated pHD-1 show (CAG) an HD expansion of predominantly 44, with a small population of 45.

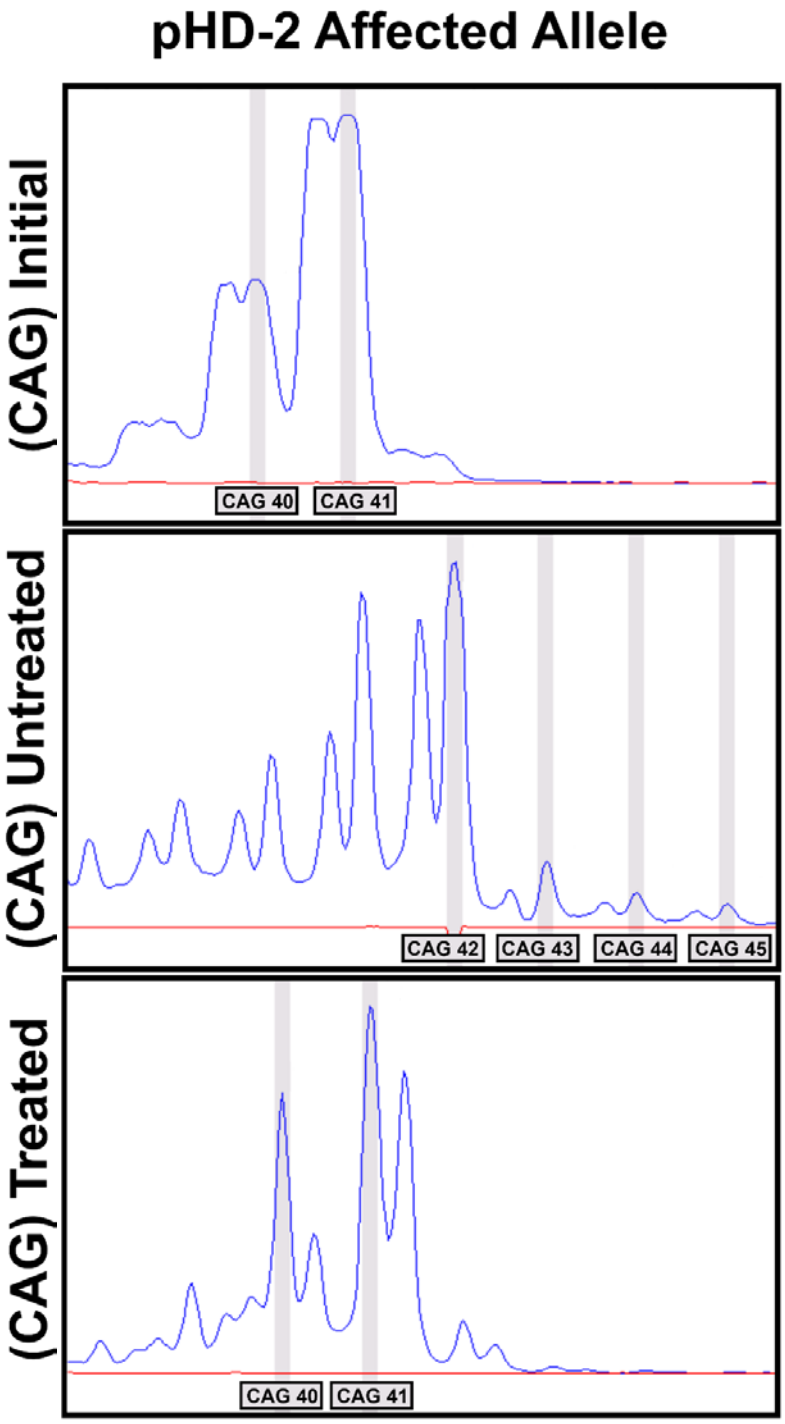


Fig. 24. Fragment analysis revealing 5-azacytidine induced TNR stability in 4 population doublings in pHD-2 fibroblasts.

Fig. 24. Continued. Sample pHD-2 cell were treated with 10 µM 5-azacytidine periodically over the course of 35 days. The top chromatogram of each image shows the (CAG) expansions at day 0. The middle chromatogram represents the (CAG) expansion of the untreated control at day 35. The bottom chromatogram shows the (CAG) expansion of 5-azacytidine treated samples. Treated pHD-2 fibroblasts shows (CAG) repeats of 17/41, similar to the starting sample. Untreated pHD-2 cells show minor TNR instability within the gene with expansions predominantly 17/42, with the affected allele showing a small population of cells expanding up to 45.

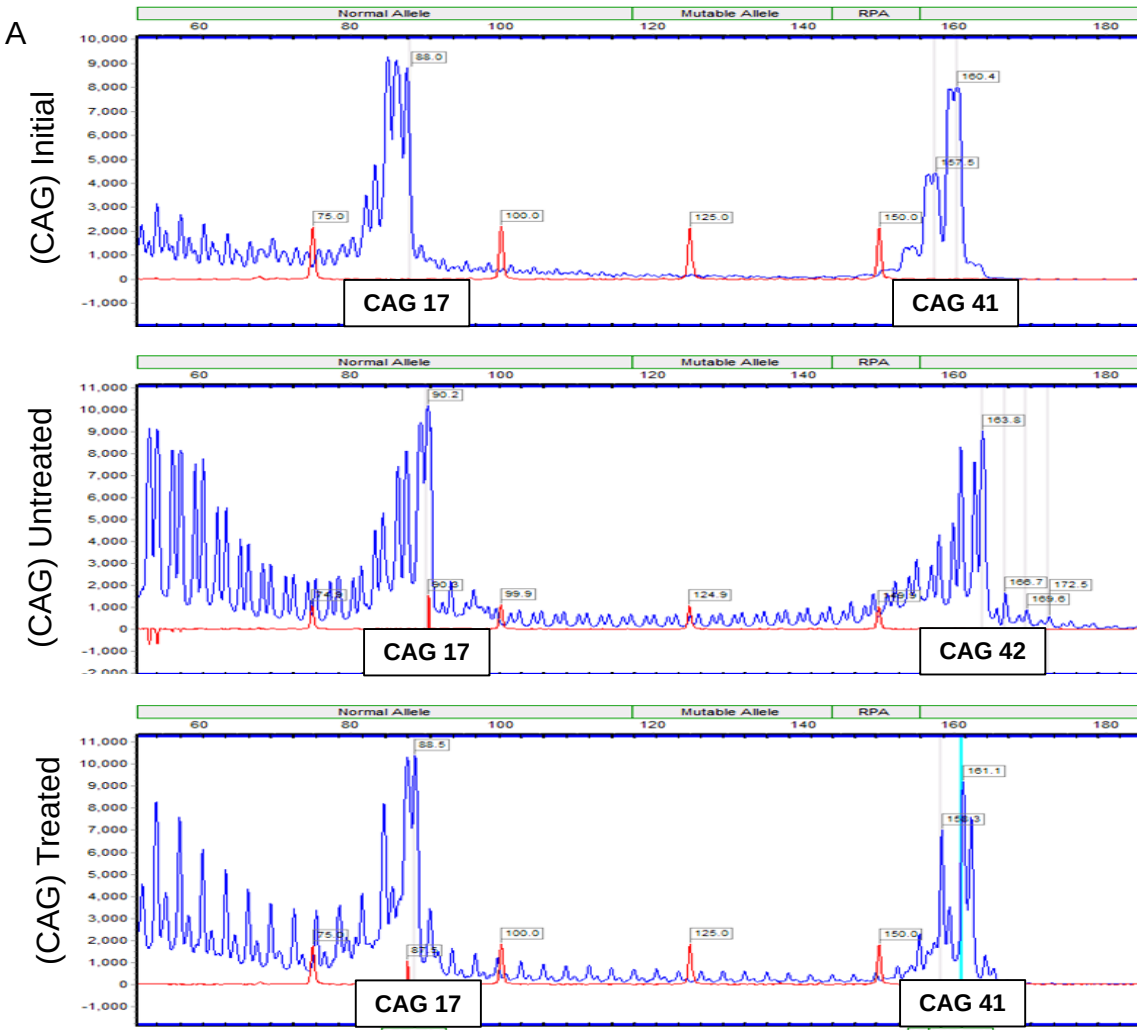


Fig. 25. Chromatograph showing entire triplet-primed PCR amplicons of 5-azacytidine induced TNR stability in 4 population doublings in HD fibroblasts. The top chromatogram of each image represents the (CAG) expansions at day 0. The middle chromatogram represents the (CAG) expansion of the untreated control at day 35. The bottom chromatogram shows the (CAG) expansion of the 5-azacytidine treated samples. (A) pHD-2 treated with 5-azacytidine shows (CAG) repeats of 17/41. Untreated pHD-1 show minor TNR instability with predominantly 17/42, with the affected allele showing a small population of cells expanding up to 45. (B) pHD-1 fibroblasts show the initial and treated samples with (CAG) repeats of 17/43, with a small population of cells showing expansions up to 44-45. Untreated pHD-1 show (CAG) an HD expansion of predominantly 44, with a small population of 45.

B

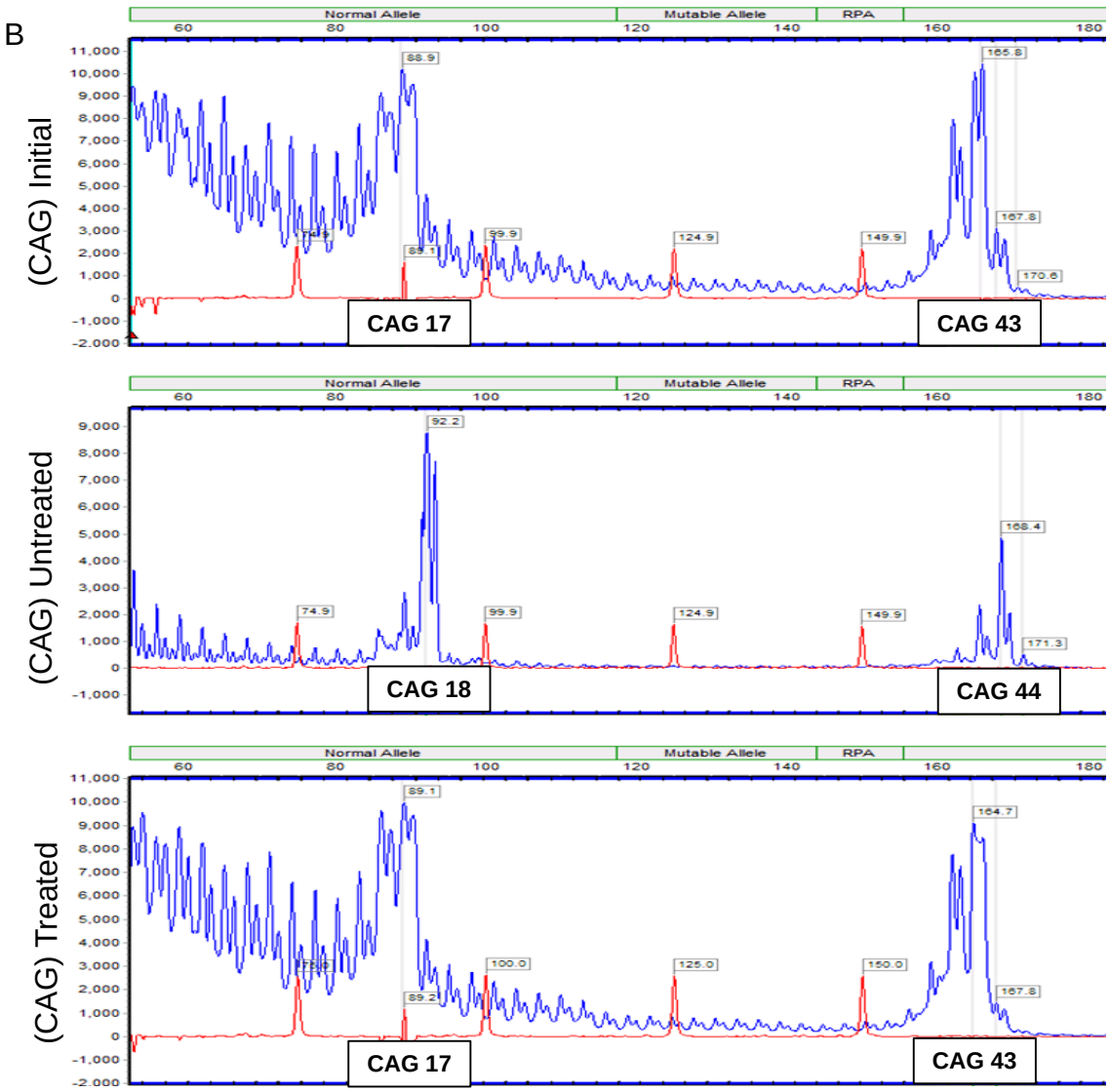


Fig. 25. Continued.

Discussion

In an effort to elucidate the pathogenetic cause of HD (CAG) expansions, we investigated DNA repair gene expression, known effectors of TNR instability. The mhtt protein produced by the expanded TNR region of exon 1 in the *HTT* gene interacts nonspecifically with many unintended targets. The resulting interaction leads to cellular dysfunction on multiple levels including transcriptional dysregulation, mitochondrial dysfunction, altered vesicular transport, ROS defense, abnormal endocytosis and secretion, and increased apoptosis.

We initially sought to determine if HD-affected fibroblasts suffer from decreased DNA repair gene expression, possibly contributing to TNR instability. Through qPCR and quantification of ICC, we found consistent decreases in APEX1, BRCA1, RPA1, and RPA3 in multiple HD fibroblast cell samples. Interestingly, using the decrease gene expression levels witnessed in HD-confirmed cells, we successfully identified two possible samples as being affected by HD (pHD-1 and pHD-2). Furthermore, protein quantification resembled the decreased mRNA levels.

Having seen consistent expression levels in 5 unrelated, HD-confirmed samples, we initially ruled out the possibility of genetic mutations and began to explore probably epigenetic causes. The unintended aggregation of mHtt to other targets is extensive, and has been thoroughly reported to inhibit gene expression because of failure to acetylate histone complexes, thus allowing strand access. We first examined the effect of trichostatin-A (TSA) on selected DNA repair gene expression. TSA has been shown to inhibit histone deacetyltransferase activity, thus allowing histone tails to remain acetylated, relaxing the DNA strand, which in turn allows successful recruitment of transcriptional or replication machinery. Neither sample showed increases with TSA treatment.

While little evidence exists that implicates HD in methylation dysfunction, our next step was determine the possibility that differential methylation patterns could be involved or that a hypomethylation agent would have off-target effects. Using passive, demethylation, we treated cells with the cytidine analogue, 5-azacytidine. Samples treated with 10 μM 5-azacytidine for 120 h revealed that

gene expression levels of the selected DNA repair genes were up-regulated in HD fibroblasts. To further investigate the possibility of methylation-involved pathogenesis of HD, we wanted to determine the effect of 5-azacytidine on HD fibroblasts over prolonged, intermittent exposure. Intriguingly, both fibroblast lines exposed to 5-azacytidine over 4 population doublings retained the initial (CAG) expansion as they had begun, with pHD-1 maintaining predominantly 18/43, and the untreated control expanding to 18/44. Similarly, pHD-2 treated fibroblasts started and retained (CAG) 17/41, with the control ending at 17/42. While the treated concentration of 10 µM was used, a logical course of action is to decrease this concentration below LC_{95} concentrations and determine its' affect over extended treatment periods.

Our results indicate the production of mhtt protein directly or indirectly inhibit the regulation of genes necessary for maintaining genome integrity. The role of BRCA1 in DNA damage is extensive, with involvement in double-strand breaks in DNA and mismatch repair, as well as p53, an upstream activator of DNA repair, initiator of apoptosis, and regulator of the cell cycle (Lin and Benchimol, 1997; Morgan and Kastan, 1997; Ouchi et al., 1998; Schwartz and Rotter, 1998). BRCA1s' direct interaction with mismatch repair proteins, and its' interaction with BRCA1-associated Ring domain protein (BARD1) suggest that it plays a global role in DNA damage repair (Wang et al., 2001). It has also been shown that BRCA1-deficient cells, Rad51 complexes fail to assemble, severely impairing DNA repair (Spiro and McMurray, 2003). There has been very few reports of BRCA1 being involved in HD exacerbation or vice versa.

Replication protein A (RPA) is a heterotrimeric (14-, 32-, and 70-kDa subunit), single-stranded binding protein, vital for the stability of single-stranded DNA during DNA metabolic processes, including transcription, replication, cell cycle advancement, and DNA damage response (Binz et al., 2004; Broderick et al., 2010; Chen and Wold, 2014). RPA70 and RPA14, together RPA32, create an RPA complex, essential to stabilize ssDNA stands, thus preventing premature, intra- and inter-strand re-annealing (Wold, 1997). It has been shown that prolonged overexpression of SSB in *E.coli* (RPA homolog) induce (CAG)·(CTG)

stability (Andreoni et al., 2010). We logically concluded that the reduction in ssDNA stability would allow unintentional secondary structure formation. Bulky adduction creation results in abnormal spatial patterning, which has shown to hinder appropriate DNA damage responses (Goula et al., 2009; Goula et al., 2012; Jeon et al., 2012a).

Base damage is the most prevalent DNA lesion in the human genome, occurring approximately 10^4 times in one cell within 24 h (Lindahl and Nyberg, 1972). DNA base lesions can be caused by both environmental cellular stressors and endogenous factors, resulting in spontaneous hydrolysis (Hoeijmakers, 2007). These apurinic/apyrimidinic (AP) sites are addressed with base excision repair pathways that involve removal of the damaged base, cleavage of the phosphodiester backbone, replacement of the proper base, and eventual ligation of the DNA strand. Specifically, APEX1 is responsible for the backbone cleavage on the 5' site of the AP site. Previous reports have shown that AP sites within (CAG) hairpins are generally resistant to APEX1 incision in constructed domains, where AP sites in fully annealed (CAG)/(CTG) sequences are corrected effectively (Li et al., 2014). More importantly, the N-termini of APEX1 contains reduction/oxidation properties known to maintain optimal transcriptional activity and DNA binding affinity (Bhakat et al., 2009; Li et al., 2012; Luo et al., 2008; Ono et al., 1995; Ramana et al., 1998; Singh and Englander, 2012). Furthermore, emerging research is also linking APEX1 and base excision repair pathways to indirect demethylation of the genome, through repair of mCyt 5' to oxo-guanines (Endutkin and Zharkov, 2015; Santos et al., 2013). In relation to HD, a double dose effect of inefficient cleavage of *HTT* gene hairpins by APEX1 and decreased expression could account for the progressive TNR expansions witnessed as well as the response to 5-azacytidine and the TNR stability reported.

We further identified that 5-azacytidine treatment of HD fibroblasts up-regulates investigated DNA repair gene expression and protein synthesis. As of now, little evidence exists that mhtt directly affects or interacts with epigenetic methylation patterning. However, it is worth noting that there is evidence of extensive changes in methylation patterns in HD, explicitly involving sequence

specific DNA-binding motifs for certain transcription factor families (Ng et al., 2013). As well, we also witnessed TNR stability in HD-affected fibroblast cells with the intermittent exposure to 5-azacytidine. Our results indicate that 5-azacytidine can marginally rescue the DNA repair gene deficiencies we witnessed here, and it stabilizes *HTT* gene instability. This suggests that pharmaceutical intervention is possible in restricting the TNR instability, thus decreasing the genetic anticipation experienced in HD.

In conclusion, our identification of endogenously down-regulated DNA repair genes in HD-affected fibroblasts begs the questions of (1) how does mHtt affect DNA repair gene expression, and (2) is the DNA repair gene expression witnessed further confounding TNR instability experienced in HD, thus allowing genetic anticipation to progress unregulated. While there is overwhelming evidence showing HD and DNA repair as being effectors of each other, no comprehensive model has been determined to clearly explain TNR instability in HD or pathogenesis of sporadic HD. Importantly, the identification of *APEX1, BRCA1, RPA1* and, *RPA3* down-regulation in HD can possibly provide therapeutic targets and provide as biomarkers for disease progression or treatment success. Our observations of gene up-regulation in response to 5-azacytidine infers that the interference of mHtt (or loss of wild-type Htt) is widespread cellularly, and that we are still determining second and third degree down-stream effects. Our observations revealing intermittent treatment of 5-azacytidne and it's' up-regulation of DNA repair genes raises the possibility that mHtt might be interfering with DNA methylation epigenetics. Any dysfunction of methylation pattern maintenance caused by mHtt would clearly alter physiological epigenetics and would have repercussions long before HD neuronal dysfunction presents, especially when considering the extensive involvement that methylation plays on neurogenesis (Covic et al., 2010; Sun et al., 2011; Szulwach et al., 2010). Importantly, the identification of *APEX1, BRCA1, RPA1* and, *RPA3* down-regulation in HD can possibly provide therapeutic targets and provide as biomarkers for disease progression or treatment success. While the passive demethylation by cell division doesn't address mitotically inactive cells (medium

spiny neurons), the stabilization of the (CAG) expansion in exon 1 of the *HTT* gene over 4 population doublings begs the possibility that controlling TNR instability and preventing HD-associated genetic anticipation might involve methylation regulators.

CHAPTER IV

HUNTINGTON'S DISEASE INDUCED PLURIPOTENT STEM CELLS SHOW ENHANCED LEVELS OF DNA REPAIR GENE EXPRESSION AND TRINUCLEOTIDE REPEAT INDUCED STABILITY

Introduction

Huntington's disease (HD) is a genetically inherited neurodegenerative disease. HD is caused by an expansion of cytosine-adenine-guanine (CAG) trinucleotide repeats (TNR) in exon 1 of the *HTT* gene. The disease-causing pathogenic TNR region is prone to instability, evidenced by somatic tissue mosaicism and intergenerational differences during disease transmission. As of yet, there is no cure for HD. At best, clinicians temporarily address disease symptoms during early disease onset, but ultimately abandon pharmacological intervention as HD progresses. Currently, there is extensive research on treating symptomatic progression of HD, with the most promising treatments involving cell-replacement therapies, specifically with the use of induced pluripotent stem cells (iPSCs). While the field of iPSC is relatively young, much emphasis is being placed on mapping the proteome and identifying differential characteristics unique to iPSC, relative to other more studied cell types. This is especially true in the context of diseased-iPSC models.

Reprogramming somatic, terminally differentiated cell types into pluripotent cell types requires simultaneous infection of transcription factors Oct4, Sox2, cMyc, and Klf4. These four transcription factors together have been referred to as the "Yamanaka factors", after the discovering scientist, Dr. Shinya Yamanaka (Takahashi et al., 2007). It is fairly recognized that the Yamanaka factors results in the greatest efficiency and is thus the most conventional method of inducing pluripotency. There are other combinations of transcription factors that can be used to induce pluripotency, being Oct4, Sox2, Nanog, and Lin28A. Unfortunately, the reprogramming process is extremely harsh on somatic cells, with reprogramming-efficiencies only reaching ~0.01%-0.1% of total cells undergoing

successful reprogramming to pluripotency. The extremely low percentage of pluripotent cells during reprogramming is due to the fact that two of the four transcription factors used in this process are oncogenes, namely cMyc and Klf4. These genes have dramatic effects in the cell. Interestingly, they act in positive-feedback loops on themselves and on each other. Also, they reactivate human telomerase reverse transcriptase, or *TERT*, the gene responsible for telomere maintenance. It is this reactivation of telomerase that allow cell division unrestricted by the Hayflick limit. Further, reprogramming to pluripotency has shown dramatic epigenetic changes in cells. As mentioned previously, two of the major epigenetic regulatory mechanisms is methylation patterning and histone modification. While much research is still needed, data shows that a combination of the Yamanaka factors change epigenetic print the cell had previously acquired (Hewitt et al., 2011; Sullivan et al., 2010; Vaskova et al., 2013). We wanted to determine the effect that pluripotency reprogramming, and by extension the global epigenetic shift that occurs, would have on DNA repair gene expression in cells that have dysregulated pathways as the result of epigenetic regulation.

While iPSC hold great promise for personalized cellular replacement therapies, a more immediate use for iPSCs is in the modeling of genetic diseases. For instance, one of the most recognized characteristics of iPSCs is the global epigenetic changes involved in reprogramming. Multiple studies have identified differential regulation of epigenetic regulatory elements and patterns in normal and diseased cells (Carey et al., 2011; Hewitt et al., 2011; Leung et al., 2014; Vaskova et al., 2013). Another key finding is that analysis of Affimetrix data has so far revealed more than 2000 differentially expressed genes in wild-type iPSCs and ESCs, relative to human neonatal foreskin fibroblasts (Cai et al., 2015). Other data has revealed that increased levels of endogenous reactive oxygen species (ROS) exist in iPSC, and that enhanced levels of DNA repair gene expression exist in response to the increased ROS levels (Fan et al., 2011). It is no surprise that the very nature of pluripotency requires different levels of mechanism involvement, and the data reporting dramatic differences in gene expression supports this. In the context of TNR instability and pluripotent cell-states, current data is extremely

limited. However, one study involving Friedreich ataxia (FRDA), another neurodegenerative disease caused by repeat expansions, revealed disease-associated tandem repeat expansions in FRDA-iPSCs, and that DNA mismatch repair protein expression was elevated (Du et al., 2012). Unfortunately, this knowledge is in its infancy and much research still must be done to comprehensively understand mechanisms involving TNR instability and targeted-cell differentiation *in vitro* and *in vivo*. Of equal importance, HD-iPSC models have enabled researchers to study the pathogenicity of HD in disease-affected cell models, striatal medium spiny neurons. In the short time iPSC-reprogramming has been available, research using HD-iPSCs showed that pluripotent cells affected by HD show differentially expressed proteins involved in oxidative stress and apoptotic responses (Chae et al., 2012). Research using HD-iPSCs has also elucidated that HD-iPSC models maintain the disease phenotype after differentiation and transplantation (Consortium, 2012; Jeon et al., 2012b; Juopperi et al., 2012). Interestingly, transplanted human-HD-iPSC into the striatum of mice show delayed HD phenotypes, indicating mechanisms altered in reprogramming affect the onset of HD (Jeon et al., 2014). These data confirm that altered expression exists during pluripotency, and further identifies stages in development where cellular mechanisms are altered, thus effecting cell response efficacy and genome integrity. While extensive research is being conducted using iPSC models, only a small number of investigations are being conducted involving the pathogenic progression of HD using iPSC. Thorough investigation of HD-iPSC, and cells derived from HD-iPSC is needed to create a comprehensive understanding the pathogenic TNR instability characteristic of HD and its progression.

Of the research needed, iPSC models of HD are essential to understanding the pathogenic trinucleotide repeat (TNR) instability characteristic of HD progression and genetic anticipation. Perhaps the most important contribution that iPSC generation has in the research setting is the ability for a research team to characterize mechanisms affecting TNR instability during pluripotency, and then subsequently monitor mechanistic changes during targeted differentiation, while

systematically monitoring repeat changes. Essentially, this experiment model allows for detailed TNR region analysis while mimicking developmental differentiation. This is a key component to elucidating the mechanisms involved in somatic TNR mosaicism witnessed in HD. Also, a highlight to using somatic reprogramming is that it allows for unlimited access to disease-specific cells, which would normally require extensive preparation and planning involving terminally-ill patients affected by HD.

In this study, we begun to dissect DNA repair gene expression during developmental differentiation, thus laying a foundation for whole tissue differential analysis of genes involved in pathways effecting TNR instability. Here we examined four DNA damage repair genes (*APEX1, BRCA1, RPA1,* and *RPA3)* that we have previously shown to be down regulated in HD cells. Reprogramming of two HD fibroblast lines into iPSCs revealed expression of these genes had returned to similar levels as wild-type iPSCs. Also, we report that HD-iPSC models undergo contraction-biased TNR instability in early stages, and then resort to stabilized regions in later stages. Interestingly, further differentiation of HD-iPSC models into mesenchymal-like cells (MLCs) show slightly reduced levels of three of these DNA repair genes (*BRCA1, RPA1,* and *RPA3*), but still maintained expression levels greater than levels in the native-cell state. Further, HD-MLCs maintained TNR stability, showing no evidence of partial population expansions or contractions. These findings suggest that DNA repair gene up-regulation assist in the maintenance of genome fidelity and that subsequent differentiation maintains increased repair gene expression relative to their parental lines, possibly inducing genomic TNR stability.

Materials and methods

Cell lines and cell culture

Two human female HD fibroblast lines (GM02191 and GM04022) were previously purchased from Coriell Cell Repositories (Camden, NJ). Two wild-type iPSC lines were previously generated and confirmed in our lab from breast-derived adipose stem cells (iPSC-WT1 and iPSC-WT2). Fibroblast lines were cultured in

Dulbecco's Modified Eagle's Medium (DMEM) with Glutamax (Life Technologies, Carlsbad, CA), supplemented with 10% premium fetal bovine serum (Life Technologies) and 1% Antibiotic-Antimycotic (100x) (ABAM; Life Technologies). Induced-pluripotent stem cells lines were initially cultured in complete KnockOut DMEM/F-12 medium (Life Technologies). Complete KnockOut Serum media (KOSR) was comprised of 20% KnockOut Replacement Serum, 1x MEM Non-essential Amino Acids (Life Technologies), 10 ng/mL of Basic Fibroblast Growth Factor (Life Technologies), 55 nM β-mercaptoethanol, (Sigma Aldrich; St. Louis, MO) and 1% ABAM. Confirmed iPSC lines were adapted to a feeder-independent protocol utilizing Essential 8 medium (E8) (Thermo Fisher Scientific; Waltham, MA) to minimize spontaneous differentiation. MLCs were differentiated and cultured in mesenchymal stem cell media (MSC) comprised of DMEM with Glutamax, 10% fetal bovine serum, and 1% ABAM. All cell lines were cultured and maintained at 37°C in a humidified incubator with 5% CO_2. Use of human cell lines used was deemed exempt from IRB review, as it is impossible to identify cell donors.

iPSC generation of HD cell lines

To generate stable pluripotent cells, on day 0, 2.5-3*10^5 cells were infected using Sendai viral vectors expressing KOS (Klf4,Oct3/4, Sox2), c-Myc, and Klf4 using the CytoTune™ –iPS 2.0 Sendai Reprogramming Kit (Life Technologies, Carlsbad, CA). Cells were transformed according to the manufacturer's protocol and lot specification (L2120009) of Sendai viral components, using a multiplicity of infection of 5:5:3. On day 7 post-transduction, cells were plated onto irradiated murine-embryonic fibroblasts (iMEFs) (Globalstem; Gaithersburg, MD). After confirmation of pluripotent-positive cells, colonies were manually split and passaged for expansion. Upon initial expansion, all iPSC lines were adapted to a feeder-independent protocol using vitronectin coat plates at a concentration of 0.5 μg/cm^2 and were continuously cultured in E8 for the remainder of our experiments. iPSC lines adapted to E8 were passaged and split at ratios of 1:6 to 1:9 using 0.5 μM EDTA in Ca- and Mg- phosphate buffered saline. Media was replenished on

all pluripotent cell lines every 24 hours. Generation of iPSCs using Sendai virus was conducted in compliance with IBC #13-021.

Confirmation of successful iPSC reprogramming

Candidate iPSCs were initially selected for 30-32 days after infection with positive staining for TRA-1-81 (Stemgent; Lexington, MA) according to the manufacturer's specifications. After manually splitting 8-10 colonies that were positive for TRA-1-81 stain for expansion, the remainder of the colonies were analyzed for alkaline phosphatase (AP) activity using Stemgent Alkaline Phosphatase Staining Kit II (Stemgent) according to the provided protocol. This further allowed us to distinguish stem cells from parental cells and iMEF feeder cells. After successful expansion of candidate iPSCs, quantitative gene expression for pluripotency markers was determined using TaqMan Gene Expression Assays found in Table 7.

Derivation of mesenchymal-like cells from iPSCs

iPSC-MLC derivation was performed using previously published protocols (Kang et al., 2015). Pure pluripotent colonies were isolated through manual selection and were brought to single-cell suspension in DMEM/F12 using StemPro® Accutase® Cell Dissociation Reagent (Thermo Fisher Scientific). Cells were plated onto 0.1% gelatin (room temperature for 2 hours) in E8 for 20-24 hours, and then media was replaced daily with MSC medium for two weeks. Sequential passaging was conducted using TrypLE Express Enzyme (Thermo Fisher Scientific).

Table 7
TaqMan® Gene Expression Assay pluripotency markers.

Gene	Alias	TaqMan® Gene Expression Assay Number
Pluripotency Marker		
TERT	Telomerase reverse transcriptase	Hs00972656_m1
LIN28A	Lin-28 Homolog A (*C. elegans*)	Hs00702808_s1
KLF4	Kruppel-like factor 4	Hs00358836_m1
MYC	Proto-oncogene c-Myc	Hs00153408_m1
POU5F1	POU Class 5 homeobox 1 (Oct3/4)	Hs00742896_s1
NANOG	Nanog homeobox	Hs04260366_g1
Housekeeping Gene		
ACTB	β-Actin	Hs99999903_m1

[α] ACTB was the housekeeping gene used to normalize cellular expression.

Nucleic acid extraction and gene expression

Over the course of multiple passages of newly transduced pluripotent cells and newly differentiated MLCs, gDNA was extracted. Approximately $5*10^5$ cells were collected and gDNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen; Hilden, Germany) according to the manufacturer's protocol. Clean up and concentration of samples was conducted using 100% ethanol and sodium acetate. gDNA samples were resuspended in DNase-free water at concentrations of 15-25 ng/µL. Purity was assessed using a Nanodrop 2000c and deemed acceptable with 260/280 ratio's between 1.8 - 1.9.

Total cellular RNA was isolated using phenol-chloroform protocol with Tri-zol™ (Thermo Fisher Scientific). RNA samples were DNase I (Roche Applied Science, Indianapolis, IN) to ensure purity. Quality and concentration was determined with a Nanodrop 2000c. One µg of each RNA sample was reverse transcribed into cDNA using Platinum™ Taq DNA Polymerase High Fidelity Kit (Thermo Fisher) according to manufacturer's instructions. cDNA samples were further diluted to a concentration of 5-10 ng/µL in dnase-free, rnase-free water.

Gene expression was conducted using TaqMan® Gene Expression Assays for *APEX1* (Hs00959050_g1), *BRCA1* (Hs01556193_m1), *RPA1* (Hs00161419_m1), and *RPA3* (Hs01047933_g1); and the endogenous housekeeping gene *ACTB* (Hs99999903_m1). Real-time quantitative PCR experiments were conducted with a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA). TaqMan® Fast Advanced Master Mix (Life Technologies) was used as amplification master-mix. Complete qPCR reactions included 1x of TaqMan Fast Master Mix, 50 ng of cDNA template, 1 µL of TaqMan Gene Expression Assay (20x), and brought up to 20 µL with water. Thermal cycling conditions were: 1 cycle of 2 minutes at 50°C, 1 cycles of 20 seconds at 95°C, and 40 cycles of 1 second at 95°C and 20 seconds at 60°C. All experiments were completed in triplicate, with each reaction sub-sampled three times. Error bars represent the standard deviation of three independent experiments conducted in triplicate. *ACTB* was used as our endogenous housekeeping gene for normalization.

HTT amplification for fragment analysis

Using homemade primers (Table 5, Chapter III), we analyzed TNR instability using triplet-primed polymerase chain reaction (TP-PCR). This method is previously described in *Chapter III, Materials and methods*. Briefly, a forward and reverse primer was designed and synthesized targeting the pathogenic tandem repeat sequence in the *HTT* gene, with the forward primer containing a bound carboxyfluorescein at its' 5' end. The repetitive (CTG) sequence at the 3' end of the reverse primer allowed for non-specific binding of the primer, while the 5' end specificity ensure complete amplification of only the gene specific sequence. PCR fragments were subsequently subjected to capillary electrophoresis to determine amplicon sizes.

Data analysis and statistical significance

Gene expression was independently conducted three times with each experiment sub-sampled in triplicate. Raw cycle-threshold (Ct) values of each sub-sample were averaged to give one mean Ct. Mean Ct values of tested genes was normalized against the housekeeping gene *ACTB* to determine ΔCt values. Each gene ΔCt value was then normalized to either the wild-type iPSC or MLC line, relative to the experiment comparison, to determine ΔΔCt's. Fold expression was decided by $2^{-\Delta\Delta Ct}$. Error bars were determined using the standard deviation of experimental triplicate variation between fold-expression. Data were subjected to ANOVA statistical analysis with Dunnett's test post-hoc to determine statistically significant differences between samples. Statistical significance was set as $p<0.05$.

Results

Successful reprogramming of HD fibroblasts into iPSCs

To determine the effects of cellular reprogramming on TNR instability of the *HTT* gene and to measure DNA repair gene expression, we generated iPSCs from two HD-affected fibroblast cell lines. At day 30 after initial reprogramming infection, live staining of candidate iPSC colonies was conducted using TRA-1-81.

Both HD1 and HD2 cell formations showed positive staining, while iMEF feeder cells did not stain (Figs. 26A and 27A). Morphological similarities of cell colonies are see in bright-field images, showing high nuclear-cytoplasm ratios as well as defined border formations. After successful manual splitting, the remainder of colonies were analyzed for AP activity. Again, both HD lines revealed increased AP activity, relative to the iMEF feeder layer upon which they were transformed (Figs. 26B and 27B). Gene expression assays revealed that HD1- and HD2-candidate iPSC lines show significant increases in pluripotency markers (Figs. 26C and 27C). These data confirm successful reprogramming of fibroblast lines into induced pluripotent stem cells.

iPSC-HD lines have normalized APEX1, BRCA1, RPA1, and RPA3 gene expression

We have previously shown that fibroblast lines affected by HD have significantly decreased levels of essential DNA damage repair gene expression. Our previous work identified *APEX1, BRCA1, RPA1,* and *RPA3* as being down-regulated in five HD fibroblast lines. Upon successful reprogramming of HD1 and HD2 fibroblast lines into iPSC-HD1 and iPSC-HD2, we conducted qPCR using TaqMan Gene Expression Assays to determine whether the DNA repair gene deficiency was maintained throughout the reprogramming process. Analysis of these gene expression profiles reveals that iPSC-HD1 and iPSC-HD2 have statistically significant up-regulation relative to their native fibroblast expression levels (Fig. 28). Gene expression for *APEX1, BRCA1, RPA1,* and *RPA3* in iPSC-HD1 showed statistically significant increases between fold changes of ~49 - 86. iPSC-HD2 also showed statistically significant increases. Next we wanted to determine if the selected DNA repair gene expression was not only increased relative to the native-cell state, but also to determine if gene expression was returned to wild-type levels. Gene expression analysis revealed that iPSC-HD1 and iPSC-HD2 expressed *APEX1, BRCA1, RPA1,* and *RPA3* at levels consistent with two wild-type iPSC lines (iPSC-Wt1 and iPSC-Wt2), with no statistically

significant differences (Fig. 29). These data suggest that the embryonic-like state found in iPSCs somehow reversed the DNA repair gene dysregulation seen in HD.

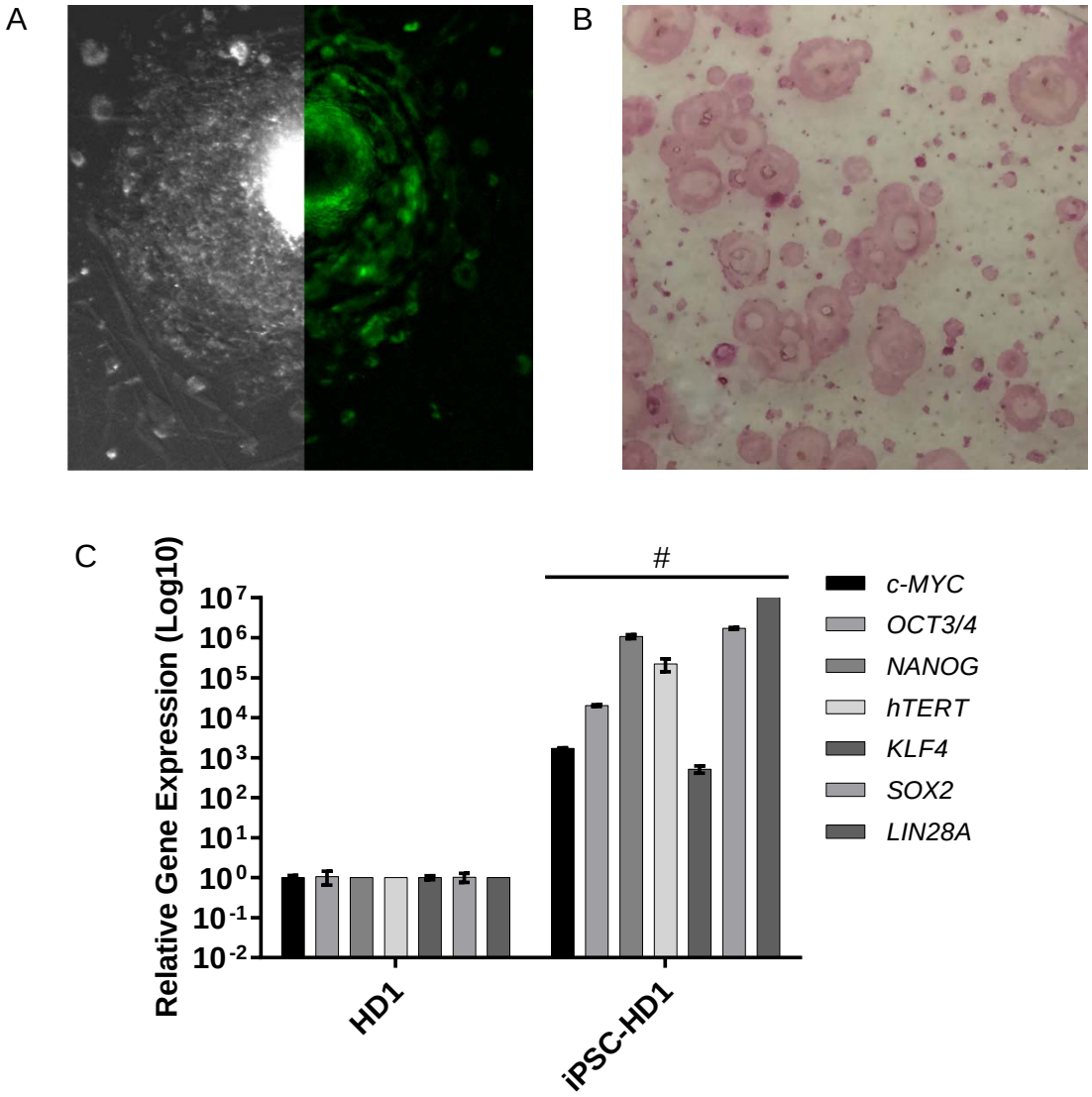


Fig. 26. Successful reprogramming of HD1 into induced pluripotent stem cells. (A) Cell colonies from the initial reprogramming were stained with TRA-1-81 live stain to detect pluripotent cells. Left half of images show phase images of morphological colony formation on MEF feeder-layer, while the right half show positive TRA-18-1 staining of colonies, but reveal negative binding to the feeder layer. (B) Cell colonies were from the reprogramming plate were positive for alkaline phosphatase. (C) Gene expression profile of fibroblast cells lines HD1 and iPSC-HD1 showing significant increases in pluripotency genes *c-Myc, OCT3/4, NANOG, hTERT, KLF4, SOX2,* and *LIN28A*. (#$p<0.0001$)

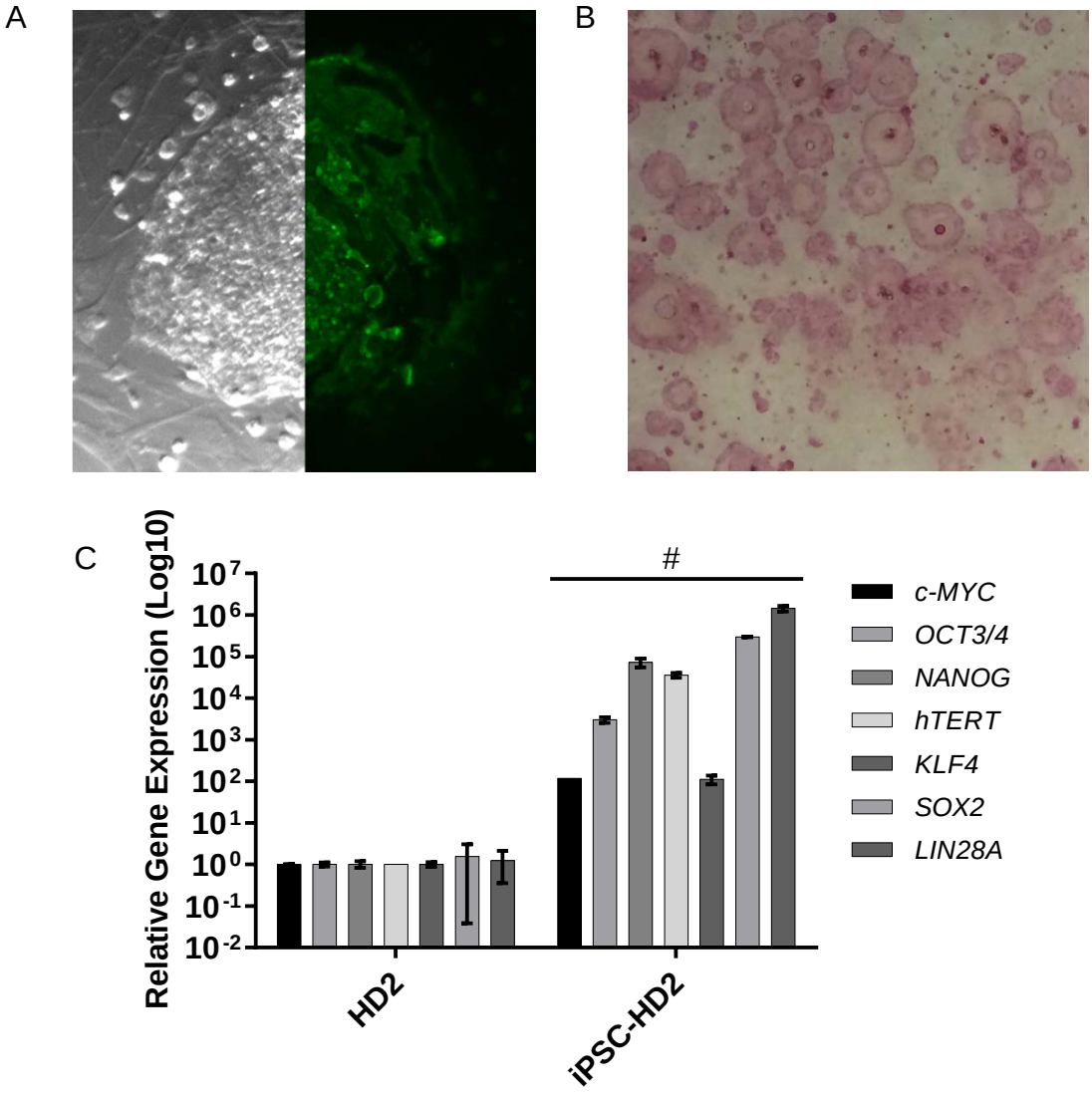


Fig. 27. Successful reprogramming of HD2 into induced pluripotent stem cells. (A) Cell colonies from the initial reprogramming were stained with TRA-1-81 live stain to detect pluripotent cells. Left half of images show phase images of morphological colony formation on MEF feeder-layer, while the right half show positive TRA-1-81 staining of colonies, but reveal negative binding to the feeder layer. (B) Cell colonies were from the reprogramming plate were positive for alkaline phosphatase. (C) Gene expression profile of fibroblast cells lines HD2 and iPSC-HD2 showing significant increases in pluripotency genes *c-Myc, OCT3/4, NANOG, hTERT, KLF4, SOX2,* and *LIN28A*. (#p<0.0001)

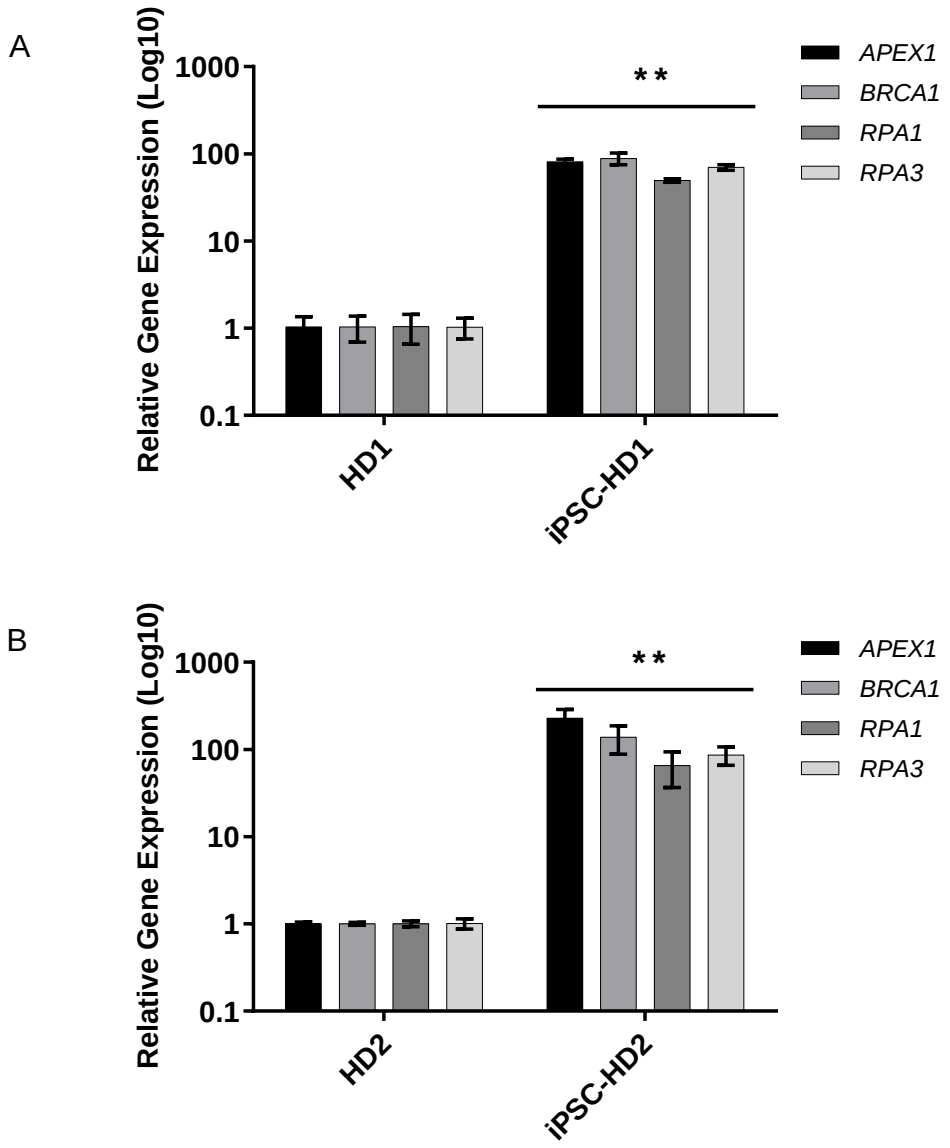


Fig. 28. Up-regulation of *APEX1, BRCA1, RPA1, and RPA3* in HD-iPSC lines. Analysis of iPSC-HD1 (A) and iPSC-HD2 (B) reveal significant up-regulation of APEX1, BRCA1, RPA1, and RPA3 relative to their native fibroblast cell states. This reveals epigenetic changes in pluripotency cell state could be involved in selected DNA repair gene expression. Statistical significance is determined by students t-test (**p<0.01)

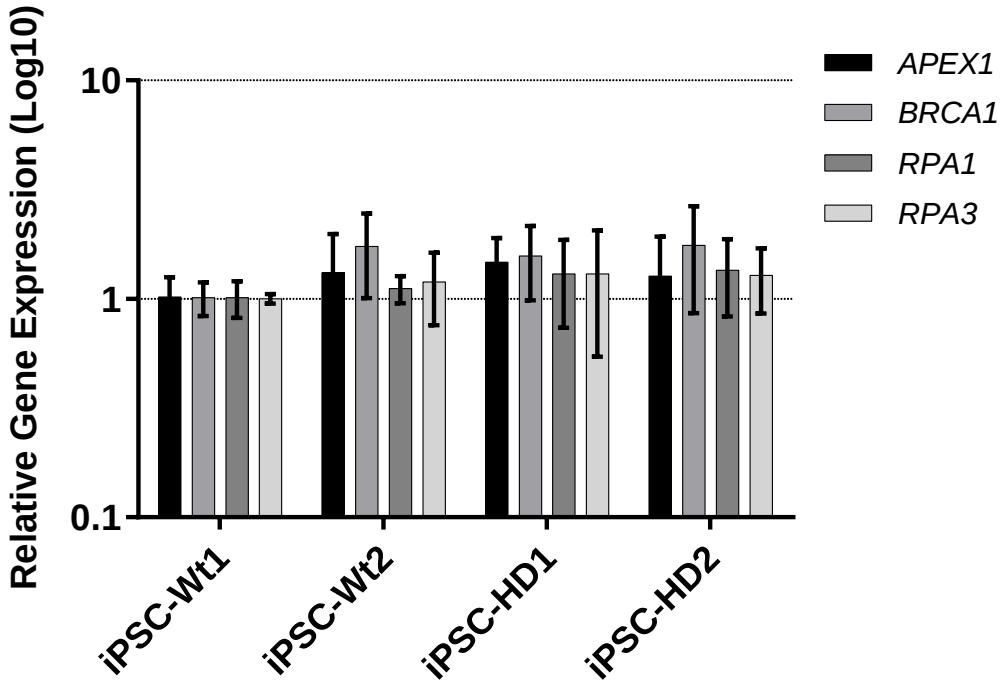


Fig. 29. Gene expression of DNA damage repair genes *APEX1, BRCA1, RPA1, and RPA3* in two HD-iPSC lines. Analysis of selected DNA damage repair genes in two HD-derived iPSC (iPSC-HD1 and iPSC-HD2) lines reveals similar expression levels relative to two wild-type iPSC (iPSC-Wt1 and iPSC-Wt2) lines. This indicates that reprogramming to pluripotent cell states, and the mechanisms involved, are up-regulating these selected genes.

MLC-HD lines exhibit decreases in BRCA1, RPA1, and RPA3 expression

To investigate whether targeted differentiation of iPSC-HD lines retained normalized DNA repair gene expression, we differentiated these lines to MLCs. After 10 population doublings, mRNA was analyzed to determine gene expression of *APEX1, BRCA1, RPA1,* and *RPA3.* Interestingly, both MLC-HD1 and MLC-HD2 did not show changes in *APEX1* expression (Fig. 30). The expression of *RPA1* showed a significant decrease in MLC-HD1, but a non-significant decrease in MLC-HD2. Further, MLC-HD1 showed decreased levels of *BRCA1* and *RPA3*. The maintained regulation of *APEX1* infers that the deficiency-causing mechanism was no longer present in MLC-HD lines. The return to decreased levels of *BRCA1, RPA1,* and *RPA3* suggests that the mechanism causing the gene dysfunction returns upon targeted differentiation.

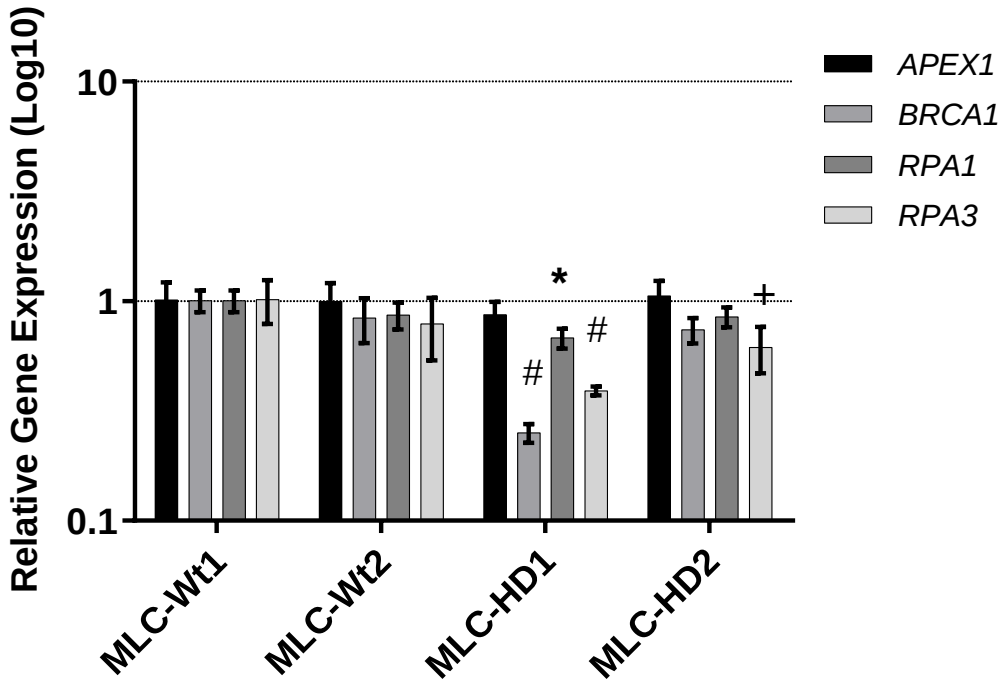


Fig. 30. Differentiation into MLC's reveals down-regulation of *BRCA1, RPA1, and RPA3*. Targeted differentiation of HD-iPSC lines into mesenchymal-like-cells (MLCs) reveals DNA repair gene expression of *BRCA1, RPA1,* and *RPA3* starts to decrease in MLC-HD1 and MLC-HD2 cells lines, indicating epigenetic memory patterns could be affecting selected gene expression. (**p*<0.05, +*p*<0.01, #*p*<0.0001).

Contraction-biased TNR instability in iPSC-HD lines

Genomic DNA samples of iPSC-HD1 and iPSC-HD2 were taken periodically during routine culturing. Approximately one sixth (~$5*10^4$ cells) of the entire population of "passaging cells" was removed for gDNA extraction and subsequent TNR instability characterization. To determine the rate of TNR instability, we collected gDNA samples at every passage up to passage 10. Using TP-PCR and capillary electrophoresis, we witnessed contraction-biased TNR instability within the pathogenic region of the *HTT* gene in both iPSC-HD lines. Characterization of iPSC-HD1 at passage 1 shows the HD-pathogenic allele containing (CAG) repeats of predominantly 44, with the unaffected allele having 18. After 8 population doublings, iPSC-HD1 shows a contraction of the affected allele to 43 (CAG)'s (Fig. 31). This contraction in iPSC-HD1 remained stable through more than 10 passages. Interestingly, iPSC-HD2 showed an identical pattern of contraction and subsequent stability, where it began at 42 (CAG)'s and contracted to 41 at passage 8 (Fig. 32). Furthermore, the affected allele remained stable for the continuation of the experiment, after the initial contraction. However, we also witnessed that both iPSC-HD lines exhibited slight contractions in the unaffected alleles, where iPSC-HD1 contracted from (CAG) repeats of 18 to 17 (Fig. 33), and iPSC-HD2 contracted from 17 to 16 (CAG) repeats (Fig. 34). To the best of our knowledge, this is the first report of unaffected allele instability in any HD model.

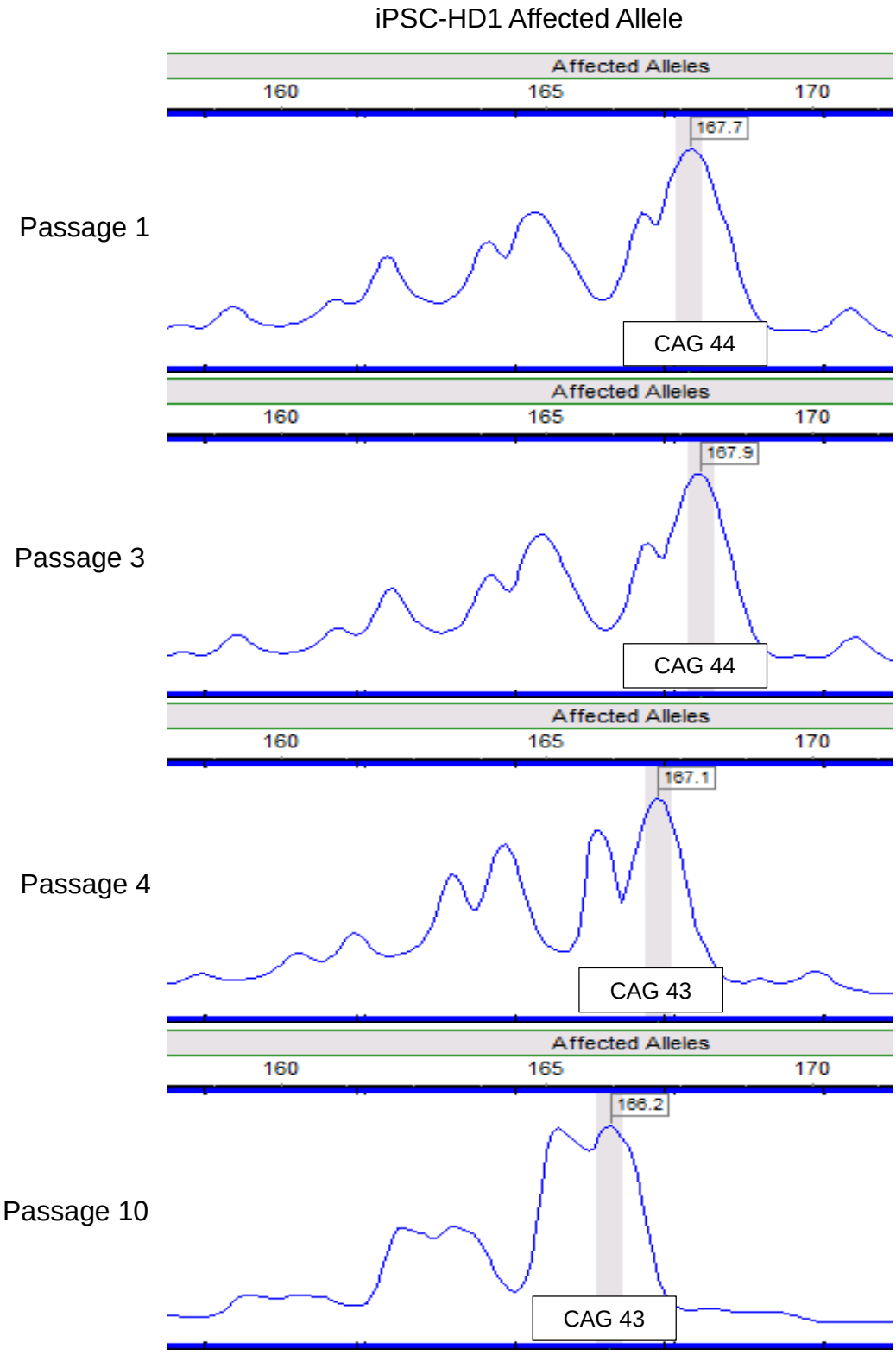


Fig. 31. Capillary electrophoresis fragment analysis reveals contraction-biased TNR instability in the iPSC-HD1 affected allele**.**

Fig. 31. Continued. *Genomic* DNA samples extracted from iPSC-HD1 were characterized for TNR instability in the *HTT* pathogenic region. Samples were taken at passages 1, 3, 4, and 10. Over the course of 10 passages, both cell lines revealed contractions of the TNR region. iPSC-HD1 affected allele shows a decrease from 44 to 43. Interestingly, both cells lines maintained the contracted allele from passage 4 to 10.

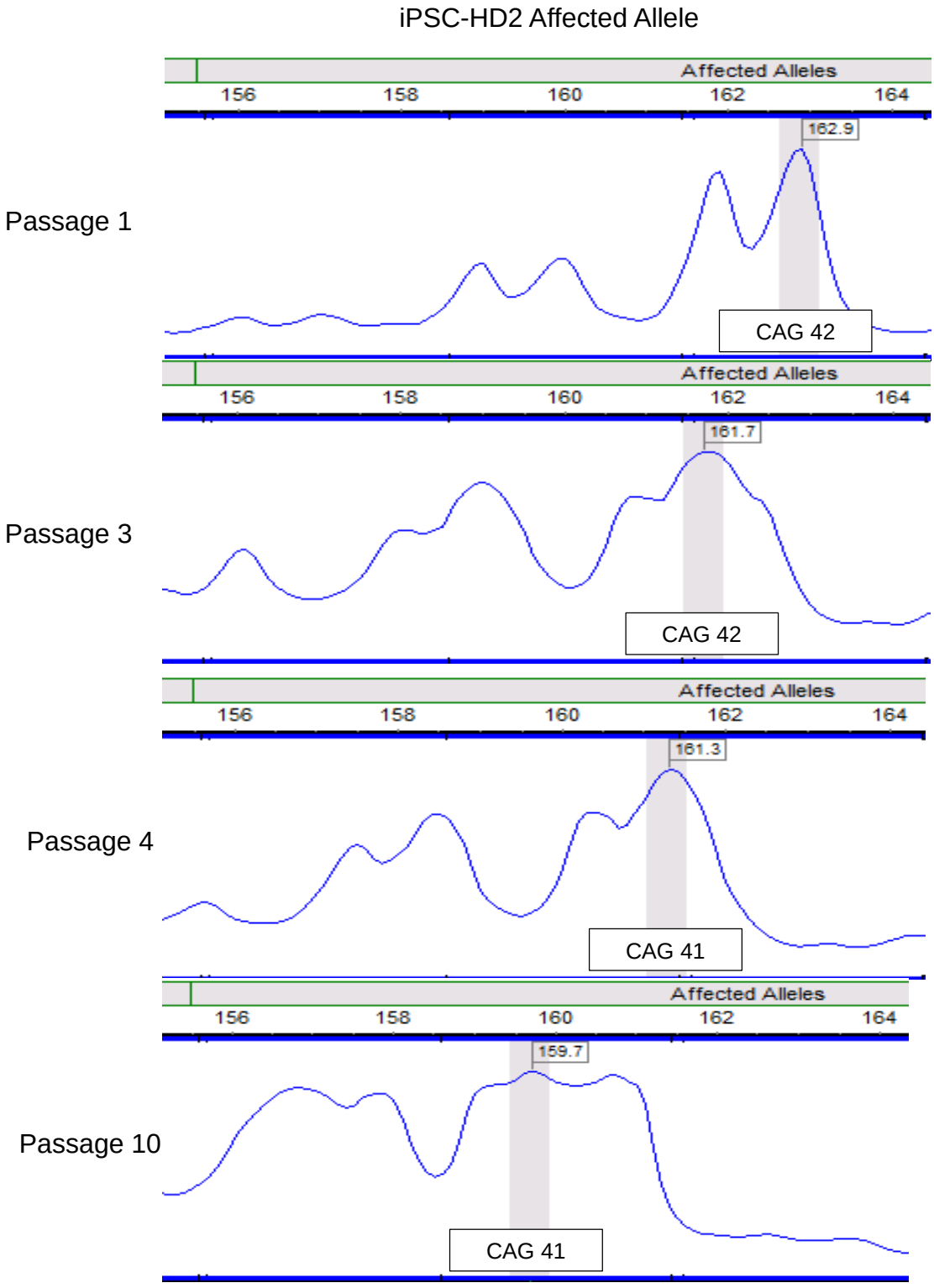


Fig. 32. Capillary electrophoresis fragment analysis reveals contraction-biased TNR instability in the iPSC-HD2 affected allele.

Fig. 32. Continued. Genomic DNA samples extracted from iPSC-HD2 were characterized for TNR instability in the *HTT* pathogenic region. Samples were taken at passages 1, 3, 4, and 10. Over the course of 10 passages, both cell lines revealed contractions of the TNR region. iPSC-HD2 shows a decrease from 42 to 41. Interestingly, both cells lines maintained the contracted allele from passage 4 to 10.

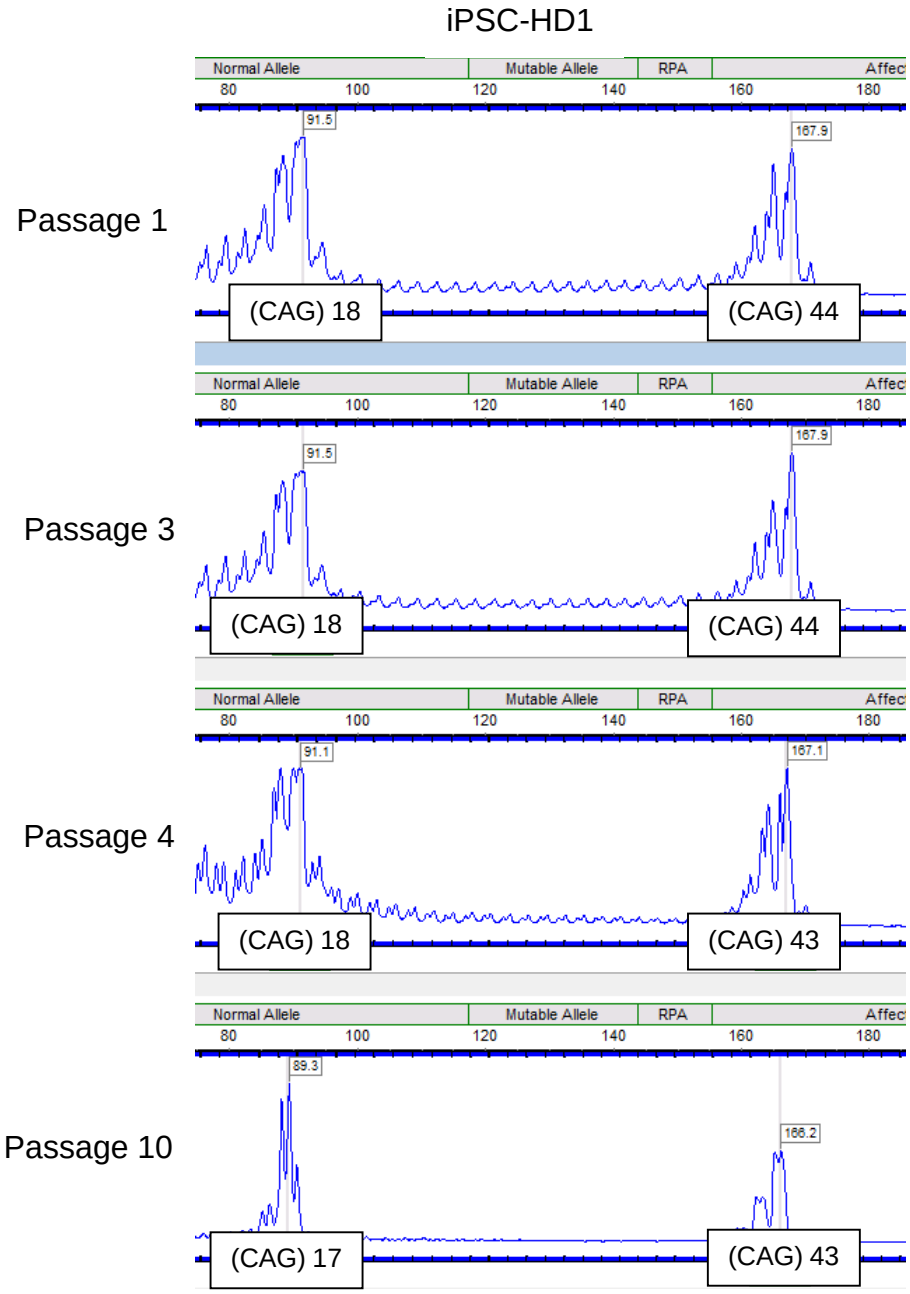


Fig. 33. Fragment analysis reveals contraction-biased TNR instability in both alleles of iPSC-HD1. *Genomic* DNA samples extracted from iPSC-HD1 were characterized for TNR instability in the *HTT* pathogenic region. Samples were taken at passages 1, 3, 4, and 10. As well as the reported contraction of the HD-affected allele, here we see the unaffected allele contract from 18 to 17 (CAG)'s between passage 4 and 10.

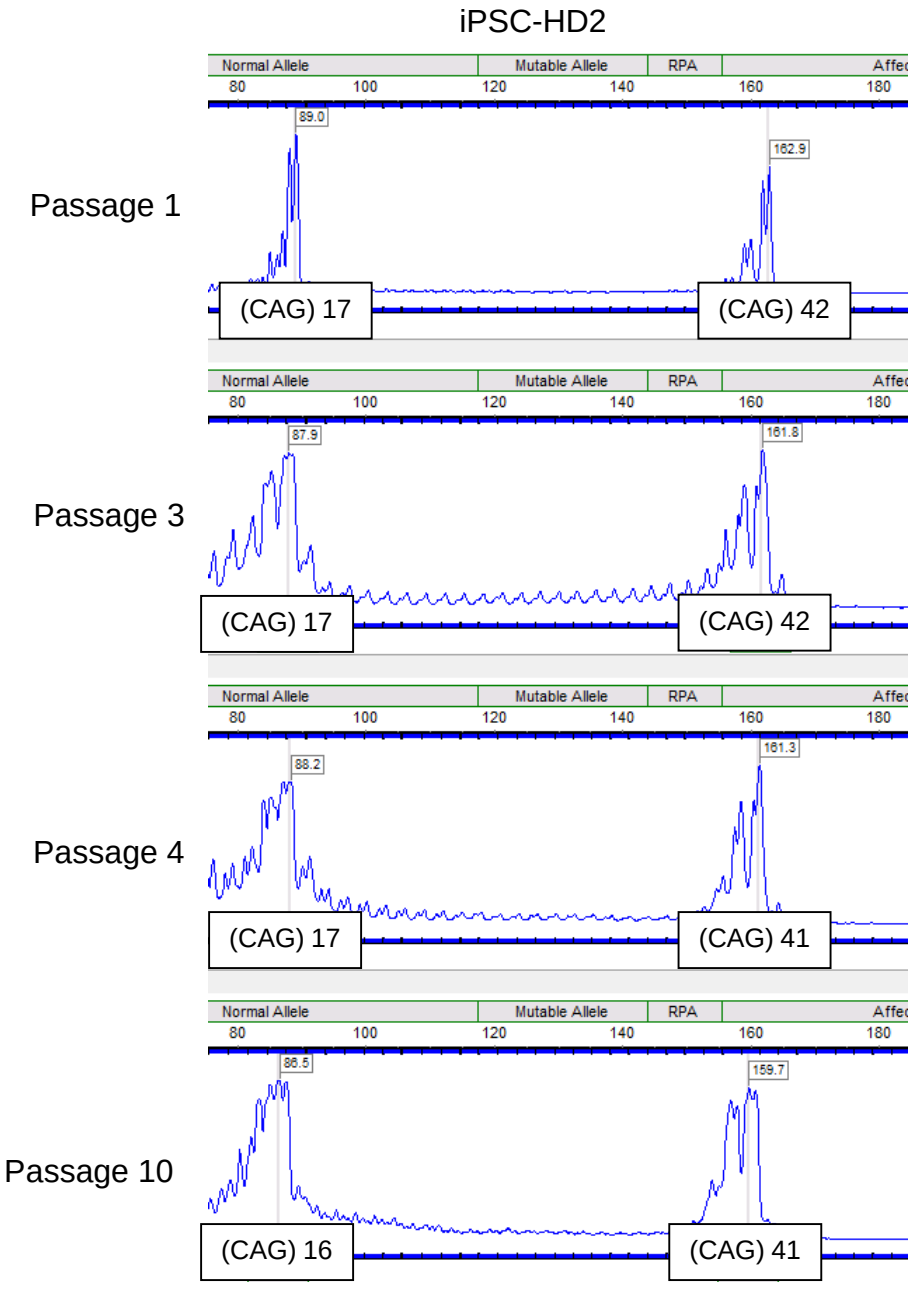
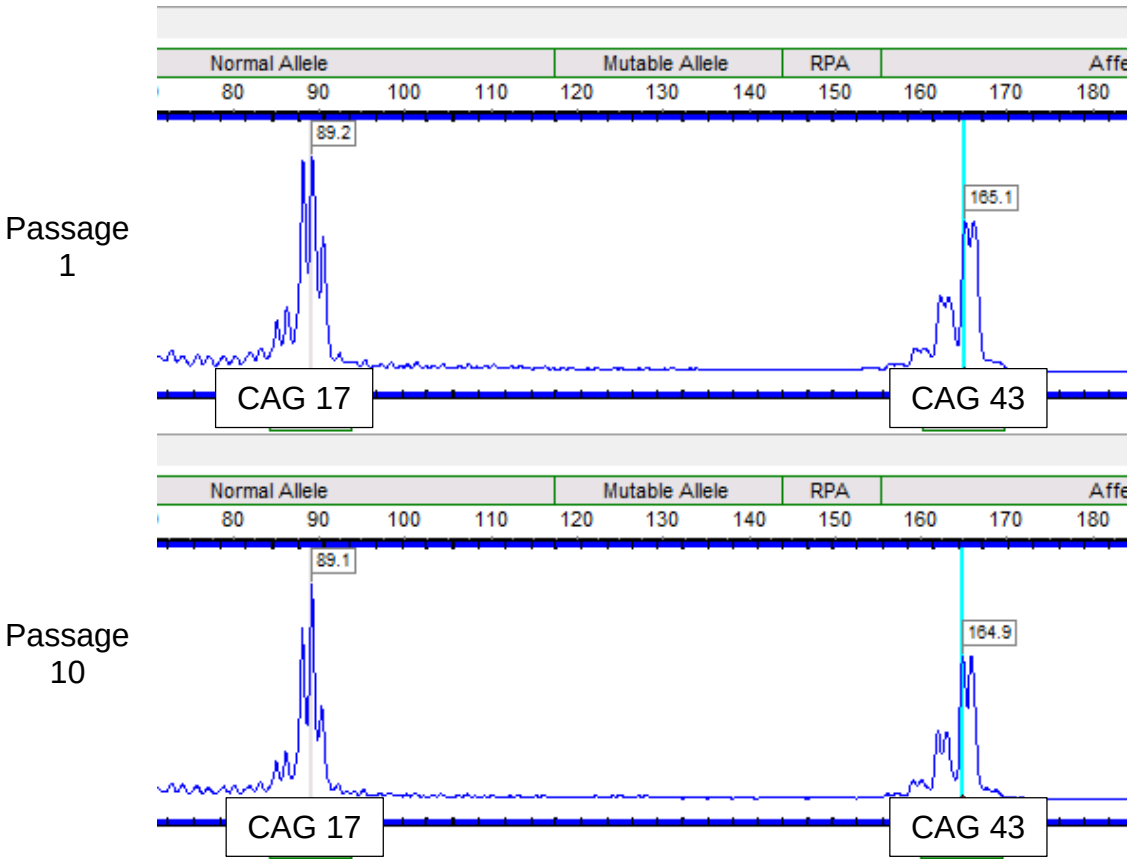


Fig. 34. Fragment analysis reveals contraction-biased TNR instability in both alleles of iPSC-HD2. Genomic DNA samples extracted from iPSC-HD2 were characterized for TNR instability in the *HTT* pathogenic region. Samples were taken at passages 1, 3, 4, and 10. Here we see the unaffected allele contract from 17 to 16 (CAG)'s between passage 4 and 10.

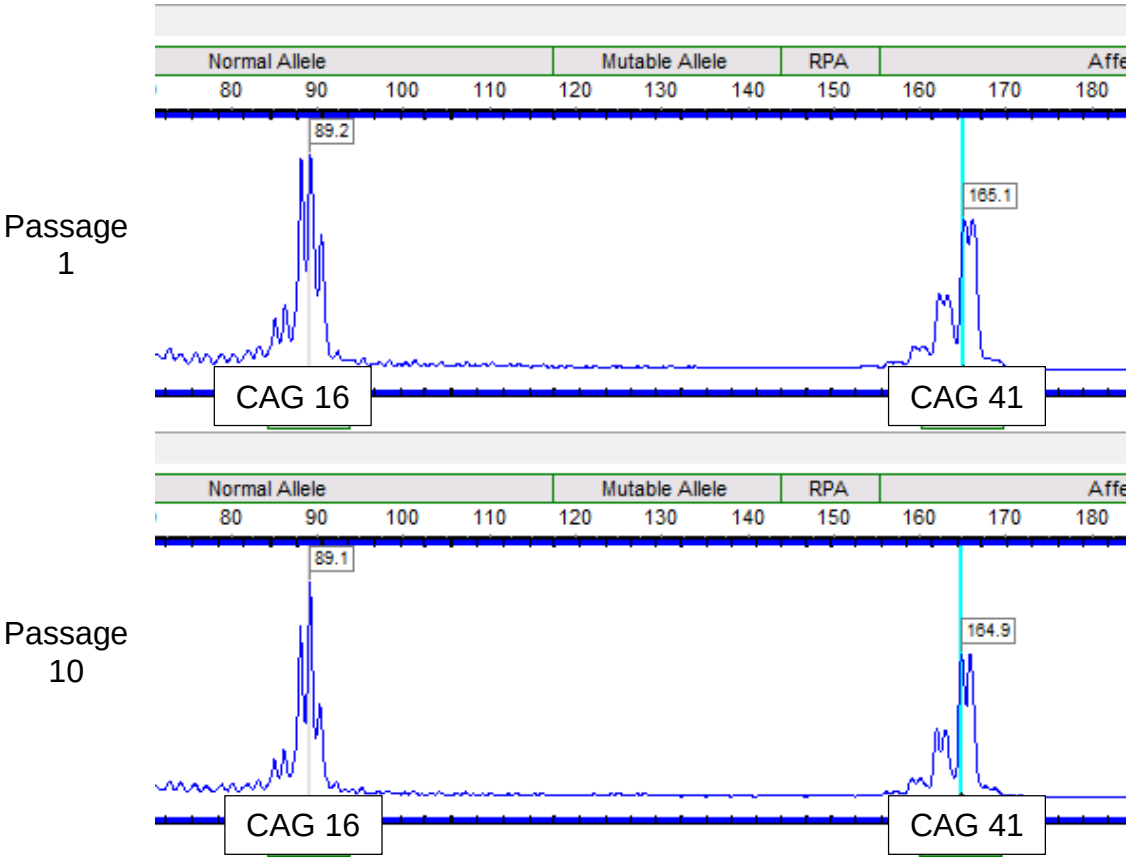
HTT gene TNR stability in iPSC-HD derived MLCs

To determine whether TNR stability was maintained in iPSC-HD derived cells, we further differentiated each line into MLCs. Genomic DNA was collected from newly differentiated MLC-HD lines at passage 1 (start of differentiation), and every passage up to 10. Using TP-PCR and capillary electrophoresis fragment analysis, we characterized the (CAG) repeat length of the *HTT* gene pathogenic region in MLC-HD1 and MLC-HD2. Both newly created MLC lines showed the same repeat expansions in both alleles, relative to their starting *HTT* gene expansions from their native iPSC state. MLC-HD1 started at 17/43 (CAG) repeats and maintained those repeats for 10 population doublings (Fig. 35). MLC-HD2 started at 16/41 (CAG) repeats, and ended at the same sizes after 10 population doublings (Fig. 36). These data infer that differentiation of HD cells from pluripotency to mesodermal lineage does not have destabilizing effects on the TNR region in HD, as we did not witness expansions or contractions in the entire population of MLC cells. As well, we also conclude that cell reprogramming, and further differentiation, increases the stability in TNR sequences in the *HTT* gene.



MLC-HD1 *HTT* Affected and Unaffected Alleles

Fig. 35. Capillary electrophoresis fragment analysis reveals TNR stability in mesenchymal-like cells differentiated from iPSC-HD1. Characterization of *HTT* pathogenic region reveals that the (CAG) expansions did not show instability from the differentiation process or through 10 population doublings. MLC-HD1 started with, and maintained 17/43 (CAG) repeats in both alleles.



MLC-HD2 *HTT* Affected and Unaffected Alleles

Fig. 36. Capillary electrophoresis fragment analysis reveals TNR stability in mesenchymal-like cells differentiated from iPSC-HD2. Characterization of *HTT* pathogenic region reveals that the (CAG) expansions did not show instability from the differentiation process or through 10 population doublings. MLC-HD2 started with, and maintained 16/41 (CAG) repeats in both alleles.

Discussion

We previously reported that HD-affected fibroblast cell lines exhibit decreased expression of DNA repair genes that have been shown to be effectors of TNR instability. We also reported that treatment of HD-fibroblasts with a hypomethylation agent resulted in up-regulation of gene expression, and that intermittent treatment for prolonged periods induced TNR stability, preventing (CAG) expansions. In this study, we wanted to investigate whether the gross epigenetic changes seen during the reprogramming to pluripotency had an effect on DNA repair gene expression in HD (Carey et al., 2011; Hewitt et al., 2011; Leung et al., 2014; Vaskova et al., 2013). Further, we wanted to determine whether mechanisms associated with cellular reprogramming and subsequent differentiation affected TNR stability of the *HTT* gene. Here we report that reprogramming terminally differentiated HD fibroblasts into iPSC resulted in up-regulation of DNA repair genes *APEX1, BRCA1, RPA1,* and *RPA3*. Our results agree with other studies showing that iPSCs have differentially expressed genes, and some showing enhanced regulation of DNA repair gene expression (Cai et al., 2015; Fan et al., 2011). Importantly, we show that during pluripotency, early passaged HD-affected cells undergo contraction-biased instability of the affected gene. Late passage HD-iPSC lines show no changes in TNR regions, thus implying increased stability. Interestingly, these results contrast with data shown in Friedreich ataxia (FRDA) iPSCs, another neurodegenerative disease caused by TNR expansions (Du et al., 2012). Though HD and FRDA are both genetic diseases caused by TNR expansions, they differ in tandem repeat sequences, where repeats in HD consist of (CAG) and FRDA consist of (GAA). The different nucleotide sequence of HD and FRDA, create different mutant products, and these data together suggest that the mutant products are misbehaving in different fashions during pluripotency. The subsequent differentiation of HD-iPSC lines into MLCs revealed that gene expression of *BRCA1, RPA1,* and *RPA3* begin to regress to initial down-regulated levels, as shown in our previous work. However, even though regulation of these genes show decreases, the expression levels remain above expression levels seen in their parental cell-state. This response to

differentiation suggests that re-installed epigenetic patterning is affecting gene expression. Also, HD-MLCs reveal that *HTT* gene sizes of the pathogenic regions remained static through 10 passages. While no mechanism of TNR instability has been comprehensively defined in HD, these data suggest that mechanisms associated with pluripotency and early-differentiated cells act protectively on the TNR region of the *HTT* gene.

Changes associated with reprogramming terminally differentiated cells into a pluripotent cell-state is extensive. Our observations of differential gene expression of DNA repair genes in HD-iPSC lines confirm other reports of altered cell regulatory processes in conjunction with pluripotency. HD-iPSC models have been reported to show similar phenotypic repeat expansions during reprogramming relative to native-state cells used for reprogramming, but it is yet to be reported whether the rate of TNR instability is altered or enhanced during pluripotency (Consortium, 2012). These findings suggest that DNA repair gene up-regulation assist in the maintenance of genome fidelity and that subsequent differentiation maintains repair gene expression briefly, possibly inducing genomic TNR stability. Our observations indicate that dysfunction DNA repair gene regulation in somatic cells affected by HD is a consequence of epigenetic patterning in HD. Importantly, we show that reprogramming of DNA repair deficient cells rescues gene expression to wild-type levels. Our data reveal that pathogenic TNR regions of the toxic *HTT* gene are somehow protected from expansion-biased instability during pluripotency, as evidenced by the contraction-only changes in early passages, as well as the static repeat length in late passages. Furthermore, the protective properties elicited upon the TNR region seem to be retained with targeted differentiation.

CHAPTER V

CONCLUSION

Identifying the mechanism of TNR instability in HD is essential to establish targeted therapeutic treatments to eradicate HD. The collective findings in this work are compelling as they provide the first evidence that mitotically active cells affected by HD have decreased expression of DNA damage repair genes essential to genome integrity and TNR stability. Our hypothesis that mHtt protein in fibroblasts affects DNA repair gene expression was supported by our findings. First, we determined that multiple genes associated with DNA damage repair are differentially expressed in multiple HD cell samples. We then determined that treatment with the hypomethylating agent 5-azacytidine resulted in DNA repair gene up-regulation, as well as inducing TNR stability in multiple HD fibroblast samples. Further, we revealed that reprogramming of HD-fibroblast cells into induced-pluripotent stem cells resulted in up-regulation of select DNA repair gene expression to wild-type expression levels, and contraction-biased TNR instability. Also, we show that HD-iPSCs that undergo differentiation into MLCs retain their starting allele lengths, and they retain wild-type expression levels of *APEX1*, a protein essential for neuronal survivability. Importantly, HD-MLC lines showed enhanced TNR stability. We determined that epigenetic changes in the genome of HD cells determined DNA repair gene expression and effected TNR stability.

These results will have clinical and research relevance for HD on multiple platforms. First, the identification of endogenously deficient DNA damage repair genes provides possible biomarkers to monitor disease progression and therapeutic treatments. This will also provide researchers insight into possible cellular machinery that is absent in HD, thus elucidating to possible dysfunctional repair mechanisms. Data showing differential methylation patterning in HD-affected fibroblasts also reveals the extent to which mHtt interacts and interrupts normal cellular processes. As well, identification of 5-azacytidine as a TNR stabilizing agent has promising implications as a pharmacological agent that

prevents somatic TNR mosaicism, possibly reducing the severity of disease progression over time. Furthermore, the initial contraction-biased instability in conjunction with pluripotency, and the subsequent induced stability from pluripotency to differentiated MLCs suggests that a mechanism associated with reprogramming is activated and is initially maintained in early passages. Future identification of the mechanism involved in this finding will elucidate pharmacological targets for HD treatments.

Although our study has identified a number of DNA repair genes down-regulated in HD, substantial data were found that present new avenues of discovery. Each finding we report here only exponentially increases the number of questions to be asked. After this work, the next steps are to create striatal neurons from Wt-iPSC and HD-iPSC and (1) investigate the extent of the DNA repair deficiency, (2) measure selected DNA repair expression response after 5-azacytidine treatments, and (3) measure TNR instability over prolonged culture. I speculate that similar DNA repair gene deficiencies will be present, thus narrowing the focus of targets causing TNR instability in striatal neurons. Further, our data that shows APEX1 up-regulation in response to 5-azacytidine implies that adjusted treatments might improve the overall survivability of diseased neurons.

Collectively these finding provide insight into the spontaneous (CAG) expansions with the pathogenic region of the *HTT* gene. Further, this work promotes the continued investigation of DNA repair gene expression and HD-iPSC, for use as cell models to determine comprehensive mechanisms involved in TNR instability. Together these data emphasize the complexity of mechanisms involved in TNR somatic mosaicism experienced in HD. Consequently, additional studies must be conducted involving iPSC-derived cells to fully understand TNR instability. However, we have identified that pathways and mechanisms related to pluripotency exhibit protective effects on TNR regions prone to expansions in HD.

REFERENCES

Acheson, A., et al., 1995. A BDNF autocrine loop in adult sensory neurons prevents cell death. Nature. 374**,** 450-3.

An, M. C., et al., 2012. Genetic correction of Huntington's disease phenotypes in induced pluripotent stem cells. Cell Stem Cell. 11**,** 253-63.

Andreoni, F., et al., 2010. Overexpression of the single-stranded DNA-binding protein (SSB) stabilises CAG*CTG triplet repeats in an orientation dependent manner. FEBS Lett. 584**,** 153-8.

Atlas, E., et al., 2001. A CREB site in the BRCA1 proximal promoter acts as a constitutive transcriptional element. Oncogene. 20**,** 7110-4.

Augood, S. J., et al., 1996. Reduction in enkephalin and substance P messenger RNA in the striatum of early grade Huntington's disease: a detailed cellular in situ hybridization study. Neuroscience. 72**,** 1023-36.

Avila-Figueroa, A., et al., 2011. A small unstructured nucleic acid disrupts a trinucleotide repeat hairpin. Biochem Biophys Res Commun. 413**,** 532-6.

Aziz, N. A., et al., 2012. CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. Neurology. 79**,** 952; author reply 952-3.

Bates, G., 2003. Huntingtin aggregation and toxicity in Huntington's disease. Lancet. 361**,** 1642-4.

Beaver, J. M., et al., 2015. AP endonuclease 1 prevents trinucleotide repeat expansion via a novel mechanism during base excision repair. Nucleic Acids Res. 43**,** 5948-60.

Benn, C. L., et al., 2009. Genetic knock-down of HDAC7 does not ameliorate disease pathogenesis in the R6/2 mouse model of Huntington's disease. PLoS One. 4**,** e5747.

Benraiss, A., Goldman, S. A., 2011. Cellular therapy and induced neuronal replacement for Huntington's disease. Neurotherapeutics. 8**,** 577-90.

Bessert, D. A., et al., 1995. The identification of a functional nuclear localization signal in the Huntington disease protein. Brain Res Mol Brain Res. 33**,** 165-73.

Bhakat, K. K., et al., 2009. Transcriptional regulatory functions of mammalian AP-endonuclease (APE1/Ref-1), an essential multifunctional protein. Antioxid Redox Signal. 11**,** 621-38.

Bhattacharyya, A., et al., 2006. Oligoproline effects on polyglutamine conformation and aggregation. J Mol Biol. 355**,** 524-35.

Binz, S. K., et al., 2004. Replication protein A phosphorylation and the cellular response to DNA damage. DNA Repair (Amst). 3**,** 1015-24.

Bobrowska, A., et al., 2012. SIRT2 ablation has no effect on tubulin acetylation in brain, cholesterol biosynthesis or the progression of Huntington's disease phenotypes in vivo. PLoS One. 7**,** e34805.

Bobrowska, A., et al., 2011. Hdac6 knock-out increases tubulin acetylation but does not modify disease progression in the R6/2 mouse model of Huntington's disease. PLoS One. 6**,** e20696.

Borgonovo, J. E., et al., 2013. Mutant huntingtin affects endocytosis in striatal cells by altering the binding of AP-2 to membranes. Exp Neurol. 241**,** 75-83.

Bowater, R. P., et al., 1997. Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in Escherichia coli. Nucleic Acids Res. 25**,** 2861-8.

Brix, T. H., et al., 2003. Preliminary evidence of genetic anticipation in Graves' disease. Thyroid. 13**,** 447-51.

Broderick, S., et al., 2010. Eukaryotic single-stranded DNA binding proteins: central factors in genome stability. Subcell Biochem. 50**,** 143-63.

Buckley, N. J., et al., 2010. The role of REST in transcriptional and epigenetic dysregulation in Huntington's disease. Neurobiol Dis. 39**,** 28-39.

Butterworth, J., et al., 1985. Distribution of phosphate-activated glutaminase, succinic dehydrogenase, pyruvate dehydrogenase and gamma-glutamyl transpeptidase in post-mortem brain from Huntington's disease and agonal cases. J Neurol Sci. 67**,** 161-71.

Cai, Y., et al., 2015. Gene expression of OCT4, SOX2, KLF4 and MYC (OSKM) induced pluripotent stem cells: identification for potential mechanisms. Diagn Pathol. 10**,** 35.

Cambon, K., Deglon, N., 2013. Lentiviral-mediated gene transfer of siRNAs for the treatment of Huntington's disease. Methods Mol Biol. 1010**,** 95-109.

Camnasio, S., et al., 2012. The first reported generation of several induced pluripotent stem cell lines from homozygous and heterozygous Huntington's disease patients demonstrates mutation related enhanced lysosomal activity. Neurobiol Dis. 46**,** 41-51.

Cannella, M., et al., 2005. New Huntington disease mutation arising from a paternal CAG34 allele showing somatic length variation in serially passaged lymphoblasts. Am J Med Genet B Neuropsychiatr Genet. 133B**,** 127-30.

Carey, B. W., et al., 2011. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. Cell Stem Cell. 9**,** 588-98.

Castel, A. L., et al., 2010. Repeat instability as the basis for human diseases and as a potential target for therapy. Nature Reviews Molecular Cell Biology. 11**,** 165-170.

Castiglioni, V., et al., 2012. Induced pluripotent stem cell lines from Huntington's disease mice undergo neuronal differentiation while showing alterations in the lysosomal pathway. Neurobiol Dis. 46**,** 30-40.

Cattaneo, E., et al., 2005. Normal huntingtin function: an alternative approach to Huntington's disease. Nat Rev Neurosci. 6**,** 919-30.

Cha, J. H., 2007. Transcriptional signatures in Huntington's disease. Prog Neurobiol. 83**,** 228-48.

Chae, J. I., et al., 2012. Quantitative proteomic analysis of induced pluripotent stem cells derived from a human Huntington's disease patient. Biochem J. 446**,** 359-71.

Chang, C. L., et al., 2002. Oxidative stress inactivates the human DNA mismatch repair system. Am J Physiol Cell Physiol. 283**,** C148-54.

Chatterjee, N., et al., 2015. Environmental stress induces trinucleotide repeat mutagenesis in human cells. Proc Natl Acad Sci U S A. 112**,** 3764-9.

Chen, R., Wold, M. S., 2014. Replication protein A: single-stranded DNA's first responder: dynamic DNA-interactions allow replication protein A to direct single-strand DNA intermediates into different pathways for synthesis or repair. Bioessays. 36**,** 1156-61.

Choi, K. A., et al., 2014. Stem cell therapy and cellular engineering for treatment of neuronal dysfunction in Huntington's disease. Biotechnol J. 9**,** 882-94.

Ciammola, A., et al., 2006. Increased apoptosis, Huntingtin inclusions and altered differentiation in muscle cell cultures from Huntington's disease subjects. Cell Death Differ. 13**,** 2068-78.

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

Cleaver, J. E., et al., 2009. Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. Nat Rev Genet. 10**,** 756-68.

Consortium, H. D. i., 2012. Induced pluripotent stem cells from patients with Huntington's disease show CAG-repeat-expansion-associated phenotypes. Cell Stem Cell. 11**,** 264-78.

Cornett, J., et al., 2005. Polyglutamine expansion of huntingtin impairs its nuclear export. Nat Genet. 37**,** 198-204.

Covic, M., et al., 2010. Epigenetic regulation of neurogenesis in the adult hippocampus. Heredity (Edinb). 105**,** 122-34.

Crane, A. T., et al., 2014. Use of Genetically Altered Stem Cells for the Treatment of Huntington's Disease. Brain Sci. 4**,** 202-19.

De Rooij, K. E., et al., 1995. Somatic expansion of the (CAG)n repeat in Huntington disease brains. Hum Genet. 95**,** 270-4.

de Souza-Pinto, N. C., et al., 2009. The recombination protein RAD52 cooperates with the excision repair protein OGG1 for the repair of oxidative lesions in mammalian cells. Mol Cell Biol. 29**,** 4441-54.

Dehay, B., Bertolotti, A., 2006. Critical role of the proline-rich region in Huntingtin for aggregation and cytotoxicity in yeast. J Biol Chem. 281**,** 35608-15.

Devys, D., et al., 2001. Pathological mechanisms in polyglutamine expansion diseases. Adv Exp Med Biol. 487**,** 199-210.

DiFiglia, M., et al., 1995. Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. Neuron. 14**,** 1075-81.

DiFiglia, M., et al., 1997. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science. 277**,** 1990-3.

Dragatsis, I., et al., 2000. Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. Nat Genet. 26**,** 300-6.

Du, J., et al., 2012. Role of mismatch repair enzymes in GAA.TTC triplet-repeat expansion in Friedreich ataxia induced pluripotent stem cells. J Biol Chem. 287**,** 29861-72.

Dunah, A. W., et al., 2002. Sp1 and TAFII130 transcriptional activity disrupted in early Huntington's disease. Science. 296**,** 2238-43.

Duyao, M. P., et al., 1995. Inactivation of the mouse Huntington's disease gene homolog Hdh. Science. 269**,** 407-10.

Edwardson, J. M., et al., 2003. Expression of mutant huntingtin blocks exocytosis in PC12 cells by depletion of complexin II. J Biol Chem. 278**,** 30849-53.

Endutkin, A., Zharkov, D., 2015. 115 Cytosine demethylation through off-target base excision repair. J Biomol Struct Dyn. 33 Suppl 1**,** 72-3.

Evert, B. O., et al., 2000. Cell death in polyglutamine diseases. Cell Tissue Res. 301**,** 189-204.

Fan, J., et al., 2011. Human induced pluripotent cells resemble embryonic stem cells demonstrating enhanced levels of DNA repair and efficacy of nonhomologous end-joining. Mutat Res. 713**,** 8-17.

Ferrante, R. J., et al., 1997. Heterogeneous topographic and cellular distribution of huntingtin expression in the normal human neostriatum. J Neurosci. 17**,** 3052-63.

Fink, K. D., et al., 2014. Intrastriatal Transplantation of Adenovirus-Generated Induced Pluripotent Stem Cells for Treating Neuropathological and

Functional Deficits in a Rodent Model of Huntington's Disease. Stem Cells Transl Med. 3**,** 620-31.

Fortune, M. T., et al., 2000. Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. Hum Mol Genet. 9**,** 439-45.

Friedberg, E. C., 2011. Nucleotide excision repair of DNA: The very early history. DNA Repair (Amst). 10**,** 668-72.

Fry, M., Usdin, K., 2006. Human nucleotide expansion disorders. Springer, Berlin New York.

Fusco, F. R., et al., 1999. Cellular localization of huntingtin in striatal and cortical neurons in rats: lack of correlation with neuronal vulnerability in Huntington's disease. J Neurosci. 19**,** 1189-202.

Gacy, A. M., et al., 1995. Trinucleotide repeats that expand in human disease form hairpin structures in vitro. Cell. 81**,** 533-40.

Gardian, G., et al., 2005. Neuroprotective effects of phenylbutyrate in the N171-82Q transgenic mouse model of Huntington's disease. J Biol Chem. 280**,** 556-63.

Gauthier, L. R., et al., 2004. Huntingtin controls neurotrophic support and survival of neurons by enhancing BDNF vesicular transport along microtubules. Cell. 118**,** 127-38.

Gil, J. M., Rego, A. C., 2008. Mechanisms of neurodegeneration in Huntington's disease. Eur J Neurosci. 27**,** 2803-20.

Glajch, K. E., Sadri-Vakili, G., 2015. Epigenetic Mechanisms Involved in Huntington's Disease Pathogenesis. J Huntingtons Dis. 4**,** 1-15.

Godin, J. D., et al., 2010. Huntingtin is required for mitotic spindle orientation and mammalian neurogenesis. Neuron. 67**,** 392-406.

Goehler, H., et al., 2004. A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. Mol Cell. 15**,** 853-65.

Gomes-Pereira, M., et al., 2004. Pms2 is a genetic enhancer of trinucleotide CAG.CTG repeat somatic mosaicism: implications for the mechanism of triplet repeat expansion. Hum Mol Genet. 13**,** 1815-25.

Gomes-Pereira, M., et al., 2001. Mouse tissue culture models of unstable triplet repeats: in vitro selection for larger alleles, mutational expansion bias and tissue specificity, but no association with cell division rates. Hum Mol Genet. 10**,** 845-54.

Gomes-Pereira, M., Monckton, D. G., 2004. Chemically induced increases and decreases in the rate of expansion of a CAG*CTG triplet repeat. Nucleic Acids Res. 32**,** 2865-72.

Gonitel, R., et al., 2008. DNA instability in postmitotic neurons. Proc Natl Acad Sci U S A. 105**,** 3467-72.

Goula, A. V., et al., 2009. Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. PLoS Genet. 5**,** e1000749.

Goula, A. V., et al., 2012. The nucleotide sequence, DNA damage location, and protein stoichiometry influence the base excision repair outcome at CAG/CTG repeats. Biochemistry. 51**,** 3919-32.

Grosch, S., Kaina, B., 1999. Transcriptional activation of apurinic/apyrimidinic endonuclease (Ape, Ref-1) by oxidative stress requires CREB. Biochem Biophys Res Commun. 261**,** 859-63.

Halabi, A., et al., 2012. DNA mismatch repair complex MutSbeta promotes GAA.TTC repeat expansion in human cells. J Biol Chem. 287**,** 29958-67.

Hanawalt, P. C., et al., 2000. Regulation of nucleotide excision repair in bacteria and mammalian cells. Cold Spring Harb Symp Quant Biol. 65**,** 183-91.

Harjes, P., Wanker, E. E., 2003. The hunt for huntingtin function: interaction partners tell many different stories. Trends Biochem Sci. 28**,** 425-33.

Harrison, L., et al., 1995. Characterization of the promoter region of the human apurinic endonuclease gene (APE). J Biol Chem. 270**,** 5556-64.

Hewitt, K. J., et al., 2011. Epigenetic and phenotypic profile of fibroblasts derived from induced pluripotent stem cells. PLoS One. 6**,** e17128.

Hildrestrand, G. A., et al., 2007. The capacity to remove 8-oxoG is enhanced in newborn neural stem/progenitor cells and decreases in juvenile mice and upon cell differentiation. DNA Repair (Amst). 6**,** 723-32.

Hildrestrand, G. A., et al., 2009. Expression patterns of Neil3 during embryonic brain development and neoplasia. BMC Neurosci. 10**,** 45.

Hoeijmakers, J. H., 2007. Genome maintenance mechanisms are critical for preventing cancer as well as other aging-associated diseases. Mech Ageing Dev. 128**,** 460-2.

Holl, A. K., et al., 2010. Combating depression in Huntington's disease: effective antidepressive treatment with venlafaxine XR. Int Clin Psychopharmacol. 25**,** 46-50.

Horn, S. C., et al., 2006. Huntingtin interacts with the receptor sorting family protein GASP2. J Neural Transm. 113**,** 1081-90.

Hoshino, M., et al., 2003. Histone deacetylase activity is retained in primary neurons expressing mutant huntingtin protein. J Neurochem. 87**,** 257-67.

Huang, E. J., Reichardt, L. F., 2001. Neurotrophins: roles in neuronal development and function. Annu Rev Neurosci. 24**,** 677-736.

Hubert, L., Jr., et al., 2011. Xpa deficiency reduces CAG trinucleotide repeat instability in neuronal tissues in a mouse model of SCA1. Hum Mol Genet. 20**,** 4822-30.

Huntington, G., 2003. On chorea. George Huntington, M.D. J Neuropsychiatry Clin Neurosci. 15**,** 109-12.

Jaenisch, R., Bird, A., 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 33 Suppl**,** 245-54.

Jakupciak, J. P., Wells, R. D., 2000. Genetic instabilities of triplet repeat sequences by recombination. IUBMB Life. 50**,** 355-9.

Jalal, S., et al., 2011. DNA Repair: From Genome Maintenance to Biomarker and Therapeutic Target. Clinical Cancer Research. 17**,** 6973-6984.

Jama, M., et al., 2013. Triplet repeat primed PCR simplifies testing for Huntington disease. J Mol Diagn. 15**,** 255-62.

Jenkins, B. G., et al., 1993. Evidence for impairment of energy metabolism in vivo in Huntington's disease using localized 1H NMR spectroscopy. Neurology. 43**,** 2689-95.

Jeon, G. S., et al., 2012a. Deregulation of BRCA1 leads to impaired spatiotemporal dynamics of gamma-H2AX and DNA damage responses in Huntington's disease. Mol Neurobiol. 45**,** 550-63.

Jeon, I., et al., 2014. In Vivo Roles of a Patient-Derived Induced Pluripotent Stem Cell Line (HD72-iPSC) in the YAC128 Model of Huntington's Disease. Int J Stem Cells. 7**,** 43-7.

Jeon, I., et al., 2012b. Neuronal properties, in vivo effects, and pathology of a Huntington's disease patient-derived induced pluripotent stem cells. Stem Cells. 30**,** 2054-62.

Jiang, H., et al., 2006. Depletion of CBP is directly linked with cellular toxicity caused by mutant huntingtin. Neurobiol Dis. 23**,** 543-51.

Juopperi, T. A., et al., 2012. Astrocytes generated from patient induced pluripotent stem cells recapitulate features of Huntington's disease patient cells. Mol Brain. 5**,** 17.

Kang, R., et al., 2015. Mesenchymal stem cells derived from human induced pluripotent stem cells retain adequate osteogenicity and chondrogenicity but less adipogenicity. Stem Cell Res Ther. 6**,** 144.

Kang, S., et al., 1995. Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in E. coli. Nat Genet. 10**,** 213-8.

Kaytor, M. D., et al., 1997. Increased trinucleotide repeat instability with advanced maternal age. Hum Mol Genet. 6**,** 2135-9.

Kazantsev, A., et al., 1999. Insoluble detergent-resistant aggregates form between pathological and nonpathological lengths of polyglutamine in mammalian cells. Proc Natl Acad Sci U S A. 96**,** 11404-9.

Kendall, A. L., et al., 1998. Functional integration of striatal allografts in a primate model of Huntington's disease. Nat Med. 4**,** 727-9.

Kennedy, L., et al., 2003. Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. Hum Mol Genet. 12**,** 3359-67.

Kennedy, L., Shelbourne, P. F., 2000. Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? Hum Mol Genet. 9**,** 2539-44.

Kim, M. W., et al., 2009. Secondary structure of Huntingtin amino-terminal region. Structure. 17**,** 1205-12.

Kim, Y. J., et al., 2006. Lysosomal proteases are involved in generation of N-terminal huntingtin fragments. Neurobiol Dis. 22**,** 346-56.

Koroshetz, W. J., et al., 1997. Energy metabolism defects in Huntington's disease and effects of coenzyme Q10. Ann Neurol. 41**,** 160-5.

Kovalenko, M., et al., 2012. Msh2 acts in medium-spiny striatal neurons as an enhancer of CAG instability and mutant huntingtin phenotypes in Huntington's disease knock-in mice. PLoS One. 7**,** e44273.

Kovtun, I. V., et al., 2007. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. Nature. 447**,** 447-52.

Kovtun, I. V., McMurray, C. T., 2001. Trinucleotide expansion in haploid germ cells by gap repair. Nat Genet. 27**,** 407-11.

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

Kurdistani, S. K., et al., 2004. Mapping global histone acetylation patterns to gene expression. Cell. 117**,** 721-33.

Lai, Y., et al., 2013. Instability of CTG repeats is governed by the position of a DNA base lesion through base excision repair. PLoS One. 8**,** e56960.

Landwehrmeyer, G. B., et al., 1995. Huntington's disease gene: regional and cellular expression in brain of normal and affected individuals. Ann Neurol. 37**,** 218-30.

Langbehn, D. R., et al., 2004. A new model for prediction of the age of onset and penetrance for Huntington's disease based on CAG length. Clin Genet. 65**,** 267-77.

Langbehn, D. R., et al., 2010. CAG-repeat length and the age of onset in Huntington disease (HD): a review and validation study of statistical approaches. Am J Med Genet B Neuropsychiatr Genet. 153B**,** 397-408.

Larson, E., et al., 2015. Age-, tissue- and length-dependent bidirectional somatic CAG*CTG repeat instability in an allelic series of R6/2 Huntington disease mice. Neurobiol Dis. 76**,** 98-111.

Lee, D. Y., McMurray, C. T., 2014. Trinucleotide expansion in disease: why is there a length threshold? Curr Opin Genet Dev. 26**,** 131-40.

Lee, J., et al., 2013. Epigenetic mechanisms of neurodegeneration in Huntington's disease. Neurotherapeutics. 10**,** 664-76.

Lee, J. M., et al., 2012. CAG repeat expansion in Huntington disease determines age at onset in a fully dominant fashion. Neurology. 78**,** 690-5.

Lenzmeier, B. A., Freudenreich, C. H., 2003. Trinucleotide repeat instability: a hairpin curve at the crossroads of replication, recombination, and repair. Cytogenet Genome Res. 100**,** 7-24.

Leung, K. S., et al., 2014. The involvement of DNA methylation and histone modification on the epigenetic regulation of embryonic stem cells and induced pluripotent stem cells. Curr Stem Cell Res Ther. 9**,** 388-95.

Li, L.-C., Dahiya, R., 2002. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 18**,** 1427-1431.

Li, M., et al., 2012. Human AP endonuclease/redox factor APE1/ref-1 modulates mitochondrial function after oxidative stress by regulating the transcriptional activity of NRF1. Free Radic Biol Med. 53**,** 237-48.

Li, M., et al., 2014. APE1 incision activity at abasic sites in tandem repeat sequences. J Mol Biol. 426**,** 2183-98.

Li, S. H., et al., 2002. Interaction of Huntington disease protein with transcriptional activator Sp1. Mol Cell Biol. 22**,** 1277-87.

Li, S. H., et al., 2000. Intranuclear huntingtin increases the expression of caspase-1 and induces apoptosis. Hum Mol Genet. 9**,** 2859-67.

Li, S. H., Li, X. J., 2004. Huntingtin-protein interactions and the pathogenesis of Huntington's disease. Trends Genet. 20**,** 146-54.

Li, W., et al., 2006. Expression and characterization of full-length human huntingtin, an elongated HEAT repeat protein. J Biol Chem. 281**,** 15916-22.

Lia, A. S., et al., 1998. Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative intertissue transcription levels and proliferative capacities. Hum Mol Genet. 7**,** 1285-91.

Lin, Y., Benchimol, S., 1997. p53-mediated cell cycle arrest and apoptosis. Leukemia. 11 Suppl 3**,** 324-6.

Lin, Y., et al., 2010. R loops stimulate genetic instability of CTG.CAG repeats. Proc Natl Acad Sci U S A. 107**,** 692-7.

Lin, Y., et al., 2006. Transcription promotes contraction of CAG repeat tracts in human cells. Nat Struct Mol Biol. 13**,** 179-80.

Lin, Y., et al., 2009. Transcription destabilizes triplet repeats. Mol Carcinog. 48**,** 350-61.

Lin, Y., Wilson, J. H., 2007. Transcription-induced CAG repeat contraction in human cells is mediated in part by transcription-coupled nucleotide excision repair. Mol Cell Biol. 27**,** 6209-17.

Lin, Y., Wilson, J. H., 2009. Diverse effects of individual mismatch repair components on transcription-induced CAG repeat instability in human cells. DNA Repair (Amst). 8**,** 878-85.

Lin, Y., Wilson, J. H., 2012. Nucleotide excision repair, mismatch repair, and R-loops modulate convergent transcription-induced cell death and repeat instability. PLoS One. 7**,** e46807.

Lindahl, T., Nyberg, B., 1972. Rate of depurination of native deoxyribonucleic acid. Biochemistry. 11**,** 3610-8.

Lipinski, M. M., Yuan, J., 2004. Mechanisms of cell death in polyglutamine expansion diseases. Curr Opin Pharmacol. 4**,** 85-90.

Liu, G., et al., 2013. Oligodeoxynucleotide binding to (CTG) . (CAG) microsatellite repeats inhibits replication fork stalling, hairpin formation, and genome instability. Mol Cell Biol. 33**,** 571-81.

Liu, Y., et al., 2009. Coordination between polymerase beta and FEN1 can modulate CAG repeat expansion. J Biol Chem. 284**,** 28352-66.

Liu, Y., Wilson, S. H., 2012. DNA base excision repair: a mechanism of trinucleotide repeat expansion. Trends Biochem Sci. 37**,** 162-72.

Liu, Y. F., 1998. Expression of polyglutamine-expanded Huntingtin activates the SEK1-JNK pathway and induces apoptosis in a hippocampal neuronal cell line. J Biol Chem. 273**,** 28873-7.

Livak, K. J., Schmittgen, T. D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 25**,** 402-8.

Lombardi, M. S., et al., 2009. A majority of Huntington's disease patients may be treatable by individualized allele-specific RNA interference. Exp Neurol. 217**,** 312-9.

Lopez Castel, A., et al., 2009. CTG/CAG repeat instability is modulated by the levels of human DNA ligase I and its interaction with proliferating cell nuclear antigen: a distinction between replication and slipped-DNA repair. J Biol Chem. 284**,** 26631-45.

Luo, L. Z., et al., 2012. DNA repair in human pluripotent stem cells is distinct from that in non-pluripotent human cells. PLoS One. 7**,** e30541.

Luo, M., et al., 2008. Role of the multifunctional DNA repair and redox signaling protein Ape1/Ref-1 in cancer and endothelial cells: small-molecule inhibition of the redox function of Ape1. Antioxid Redox Signal. 10**,** 1853-67.

MacDonald, M. E., et al., 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. Cell. 72**,** 971-83.

Maiuri, T., et al., 2013. The huntingtin N17 domain is a multifunctional CRM1 and Ran-dependent nuclear and cilial export signal. Hum Mol Genet. 22**,** 1383-94.

Mallona, I., et al., 2014. Methylation plotter: a web tool for dynamic visualization of DNA methylation data. Source Code Biol Med. 9**,** 11.

Mangiarini, L., et al., 1997. Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. Nat Genet. 15**,** 197-200.

Mann, V. M., et al., 1990. Mitochondrial function and parental sex effect in Huntington's disease. Lancet. 336**,** 749.

Marcora, E., et al., 2003. Stimulation of NeuroD activity by huntingtin and huntingtin-associated proteins HAP1 and MLK2. Proc Natl Acad Sci U S A. 100**,** 9578-83.

Martinez-Delgado, B., et al., 2011. Genetic anticipation is associated with telomere shortening in hereditary breast cancer. PLoS Genet. 7**,** e1002182.

Masino, L., Pastore, A., 2002. Glutamine repeats: structural hypotheses and neurodegeneration. Biochem Soc Trans. 30**,** 548-51.

McCampbell, A., Fischbeck, K. H., 2001. Polyglutamine and CBP: fatal attraction? Nat Med. 7**,** 528-30.

McFarland, K. N., et al., 2012. Genome-wide histone acetylation is altered in a transgenic mouse model of Huntington's disease. PLoS One. 7**,** e41423.

McMurray, C. T., Kortun, I. V., 2003. Repair in haploid male germ cells occurs late in differentiation as chromatin is condensing. Chromosoma. 111**,** 505-8.

Metzler, M., et al., 2000. Huntingtin is required for normal hematopoiesis. Hum Mol Genet. 9**,** 387-94.

Mirkin, S. M., 2007. Expandable DNA repeats and human disease. Nature. 447**,** 932-40.

Mochmann, L. H., Wells, R. D., 2004. Transcription influences the types of deletion and expansion products in an orientation-dependent manner from GAC*GTC repeats. Nucleic Acids Res. 32**,** 4469-79.

Mollersen, L., et al., 2012. Neil1 is a genetic modifier of somatic and germline CAG trinucleotide repeat instability in R6/1 mice. Hum Mol Genet. 21**,** 4939-47.

Morgan, S. E., Kastan, M. B., 1997. p53 and ATM: cell cycle, cell death, and cancer. Adv Cancer Res. 71**,** 1-25.

Moumne, L., et al., 2012. Genetic knock-down of HDAC3 does not modify disease-related phenotypes in a mouse model of Huntington's disease. PLoS One. 7**,** e31080.

Muftuoglu, M., et al., 2009. Cockayne syndrome group B protein stimulates repair of formamidopyrimidines by NEIL1 DNA glycosylase. J Biol Chem. 284**,** 9270-9.

Narciso, L., et al., 2007. Terminally differentiated muscle cells are defective in base excision DNA repair and hypersensitive to oxygen injury. Proc Natl Acad Sci U S A. 104**,** 17010-5.

Nasir, J., et al., 1995. Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. Cell. 81**,** 811-23.

Ng, C. W., et al., 2013. Extensive changes in DNA methylation are associated with expression of mutant huntingtin. Proc Natl Acad Sci U S A. 110**,** 2354-9.

Norremolle, A., et al., 1995. Correlation between magnitude of CAG repeat length alterations and length of the paternal repeat in paternally inherited Huntington's disease. Clin Genet. 47**,** 113-7.

Nucifora, F. C., Jr., et al., 2001. Interference by huntingtin and atrophin-1 with cbp-mediated transcription leading to cellular toxicity. Science. 291**,** 2423-8.

Ondo, W. G., et al., 2002. Tetrabenazine treatment for Huntington's disease-associated chorea. Clin Neuropharmacol. 25**,** 300-2.

Ono, Y., et al., 1995. Developmental expression of APEX nuclease, a multifunctional DNA repair enzyme, in mouse brains. Brain Res Dev Brain Res. 86**,** 1-6.

Orr, H. T., 2001. Beyond the Qs in the polyglutamine diseases. Genes Dev. 15**,** 925-32.

Ouchi, T., et al., 1998. BRCA1 regulates p53-dependent gene expression. Proc Natl Acad Sci U S A. 95**,** 2302-6.

Palfi, S., et al., 1998. Fetal striatal allografts reverse cognitive deficits in a primate model of Huntington disease. Nat Med. 4**,** 963-6.

Panigrahi, G. B., et al., 2010. Isolated short CTG/CAG DNA slip-outs are repaired efficiently by hMutSbeta, but clustered slip-outs are poorly repaired. Proc Natl Acad Sci U S A. 107**,** 12593-8.

Panigrahi, G. B., et al., 2012. Human mismatch repair protein hMutLalpha is required to repair short slipped-DNAs of trinucleotide repeats. J Biol Chem. 287**,** 41844-50.

Parker, W. D., Jr., et al., 1990. Evidence for a defect in NADH: ubiquinone oxidoreductase (complex I) in Huntington's disease. Neurology. 40**,** 1231-4.

Parniewski, P., et al., 1999. Nucleotide excision repair affects the stability of long transcribed (CTG*CAG) tracts in an orientation-dependent manner in Escherichia coli. Nucleic Acids Res. 27**,** 616-23.

Peixoto, A., et al., 2006. BRCA1 and BRCA2 germline mutational spectrum and evidence for genetic anticipation in Portuguese breast/ovarian cancer families. Fam Cancer. 5**,** 379-87.

Penney, J. B., Jr., et al., 1997. CAG repeat number governs the development rate of pathology in Huntington's disease. Ann Neurol. 41**,** 689-92.

Peyser, C. E., Folstein, S. E., 1990. Huntington's disease as a model for mood disorders. Clues from neuropathology and neurochemistry. Mol Chem Neuropathol. 12**,** 99-119.

Pinto, R. M., et al., 2013. Mismatch repair genes Mlh1 and Mlh3 modify CAG instability in Huntington's disease mice: genome-wide and candidate approaches. PLoS Genet. 9**,** e1003930.

Pla, P., et al., 2014. Mood disorders in Huntington's disease: from behavior to cellular and molecular mechanisms. Front Behav Neurosci. 8**,** 135.

Qiu, Y., et al., 2015. Molecular mechanism of resolving trinucleotide repeat hairpin by helicases. Structure. 23**,** 1018-27.

Quinti, L., et al., 2010. Evaluation of histone deacetylases as drug targets in Huntington's disease models. Study of HDACs in brain tissues from R6/2 and CAG140 knock-in HD mouse models and human patients and in a neuronal HD cell model. PLoS Curr. 2.

Rai, B., Sharif, F., 2014. Genetic anticipation and autism. Indian Pediatr. 51**,** 841-2.

Ramana, C. V., et al., 1998. Activation of apurinic/apyrimidinic endonuclease in human cells by reactive oxygen species and its correlation with their adaptive response to genotoxicity of free radicals. Proc Natl Acad Sci U S A. 95**,** 5061-6.

Ranen, N. G., et al., 1995. Anticipation and instability of IT-15 (CAG)n repeats in parent-offspring pairs with Huntington disease. Am J Hum Genet. 57**,** 593-602.

Ratovitski, T., et al., 2009. Mutant huntingtin N-terminal fragments of specific size mediate aggregation and toxicity in neuronal cells. J Biol Chem. 284**,** 10855-67.

Raymond, L. A., et al., 2011. Pathophysiology of Huntington's disease: time-dependent alterations in synaptic and receptor function. Neuroscience. 198**,** 252-73.

Reilmann, R., 2013. Pharmacological treatment of chorea in Huntington's disease-good clinical practice versus evidence-based guideline. Mov Disord. 28**,** 1030-3.

Reilmann, R., et al., 2015. A randomized, placebo-controlled trial of AFQ056 for the treatment of chorea in Huntington's disease. Mov Disord. 30**,** 427-31.

Reiner, A., et al., 2011. Genetics and neuropathology of Huntington's disease. Int Rev Neurobiol. 98**,** 325-72.

Rieger, R., et al., 1991. Glossary of genetics : classical and molecular. Springer-Verlag, Berlin ; New York.

Rigamonti, D., et al., 2000. Wild-type huntingtin protects from apoptosis upstream of caspase-3. J Neurosci. 20**,** 3705-13.

Rolfsmeier, M. L., et al., 2001. Cis-elements governing trinucleotide repeat instability in Saccharomyces cerevisiae. Genetics. 157**,** 1569-79.

Rolseth, V., et al., 2008. Widespread distribution of DNA glycosylases removing oxidative DNA lesions in human and rodent brains. DNA Repair (Amst). 7**,** 1578-88.

Ryu, H., et al., 2006. ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease. Proc Natl Acad Sci U S A. 103**,** 19176-81.

Sadeghian, K., Ochsenfeld, C., 2015. Unraveling the Base Excision Repair Mechanism of Human DNA Glycosylase. J Am Chem Soc. 137**,** 9824-31.

Sadri-Vakili, G., et al., 2007. Histones associated with downregulated genes are hypo-acetylated in Huntington's disease models. Hum Mol Genet. 16**,** 1293-306.

Sagredo, O., et al., 2012. Cannabinoids: novel medicines for the treatment of Huntington's disease. Recent Pat CNS Drug Discov. 7**,** 41-8.

Salinas-Rios, V., et al., 2011. DNA slip-outs cause RNA polymerase II arrest in vitro: potential implications for genetic instability. Nucleic Acids Res. 39**,** 7444-54.

Santos, F., et al., 2013. Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. Epigenetics Chromatin. 6**,** 39.

Saudou, F., et al., 1996. Polyglutamine expansions and neurodegenerative diseases. Cold Spring Harb Symp Quant Biol. 61**,** 639-47.

Saudou, F., et al., 1998. Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. Cell. 95**,** 55-66.

Savouret, C., et al., 2003. CTG repeat instability and size variation timing in DNA repair-deficient mice. EMBO J. 22**,** 2264-73.

Schaffar, G., et al., 2004. Cellular toxicity of polyglutamine expansion proteins: mechanism of transcription factor deactivation. Mol Cell. 15**,** 95-105.

Scharer, O. D., 2011. Multistep damage recognition, pathway coordination and connections to transcription, damage signaling, chromatin structure, cancer and aging: current perspectives on the nucleotide excision repair pathway. DNA Repair (Amst). 10**,** 667.

Schumacher, S., et al., 2001. Modulation of transcription reveals a new mechanism of triplet repeat instability in Escherichia coli. J Mol Biol. 307**,** 39-49.

Schwartz, D., Rotter, V., 1998. p53-dependent cell cycle control: response to genotoxic stress. Semin Cancer Biol. 8**,** 325-36.

Segui, N., et al., 2013. Telomere length and genetic anticipation in Lynch syndrome. PLoS One. 8**,** e61286.

Seredenina, T., et al., 2011. Decreased striatal RGS2 expression is neuroprotective in Huntington's disease (HD) and exemplifies a compensatory aspect of HD-induced gene regulation. PLoS One. 6**,** e22231.

Seredenina, T., Luthi-Carter, R., 2012. What have we learned from gene expression profiles in Huntington's disease? Neurobiol Dis. 45**,** 83-98.

Shelbourne, P. F., et al., 2007. Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. Hum Mol Genet. 16**,** 1133-42.

Shimohata, T., et al., 2000a. Expanded polyglutamine stretches interact with TAFII130, interfering with CREB-dependent transcription. Nat Genet. 26**,** 29-36.

Shimohata, T., et al., 2000b. Interaction of expanded polyglutamine stretches with nuclear transcription factors leads to aberrant transcriptional regulation in polyglutamine diseases. Neuropathology. 20**,** 326-33.

Sinden, R. R., 2001. Neurodegenerative diseases. Origins of instability. Nature. 411**,** 757-8.

Singaraja, R. R., et al., 2002. HIP14, a novel ankyrin domain-containing protein, links huntingtin to intracellular trafficking and endocytosis. Hum Mol Genet. 11**,** 2815-28.

Singh, S., Englander, E. W., 2012. Nuclear depletion of apurinic/apyrimidinic endonuclease 1 (Ape1/Ref-1) is an indicator of energy disruption in neurons. Free Radic Biol Med. 53**,** 1782-90.

Smith, R., et al., 2009. Mutant huntingtin interacts with {beta}-tubulin and disrupts vesicular transport and insulin secretion. Hum Mol Genet. 18**,** 3942-54.

Spiro, C., McMurray, C. T., 2003. Nuclease-deficient FEN-1 blocks Rad51/BRCA1-mediated repair and causes trinucleotide repeat instability. Mol Cell Biol. 23**,** 6063-74.

Steffan, J. S., et al., 2004. SUMO modification of Huntingtin and Huntington's disease pathology. Science. 304**,** 100-4.

Steffan, J. S., et al., 2001. Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. Nature. 413**,** 739-43.

Steffan, J. S., et al., 2000. The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. Proc Natl Acad Sci U S A. 97**,** 6763-8.

Stupart, D., et al., 2014. Fertility and apparent genetic anticipation in Lynch syndrome. Fam Cancer. 13**,** 369-74.

Sullivan, G. J., et al., 2010. Induced pluripotent stem cells: epigenetic memories and practical implications. Mol Hum Reprod. 16**,** 880-5.

Sun, J., et al., 2011. Epigenetic regulation of neurogenesis in the adult mammalian brain. Eur J Neurosci. 33**,** 1087-93.

Suzuki, M., et al., 2012. Calcium leak through ryanodine receptor is involved in neuronal death induced by mutant huntingtin. Biochem Biophys Res Commun. 429**,** 18-23.

Suzuki, M. M., Bird, A., 2008. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet. 9**,** 465-76.

Swami, M., et al., 2009. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. Hum Mol Genet. 18**,** 3039-47.

Szulwach, K. E., et al., 2010. Cross talk between microRNA and epigenetic regulation in adult neurogenesis. J Cell Biol. 189**,** 127-41.

Takahashi, K., et al., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 131**,** 861-72.

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

Taylor, J. P., et al., 2003. Aberrant histone acetylation, altered transcription, and retinal degeneration in a Drosophila model of polyglutamine disease are rescued by CREB-binding protein. Genes Dev. 17**,** 1463-8.

Telenius, H., et al., 1994. Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet. 6**,** 409-14.

Trottier, Y., et al., 1995. Cellular localization of the Huntington's disease protein and discrimination of the normal and mutated form. Nat Genet. 10**,** 104-10.

Tuo, J., et al., 2002. Functional crosstalk between hOgg1 and the helicase domain of Cockayne syndrome group B protein. DNA Repair (Amst). 1**,** 913-27.

van Bilsen, P. H., et al., 2008. Identification and allele-specific silencing of the mutant huntingtin allele in Huntington's disease patient-derived fibroblasts. Hum Gene Ther. 19**,** 710-9.

Vaskova, E. A., et al., 2013. "Epigenetic memory" phenomenon in induced pluripotent stem cells. Acta Naturae. 5**,** 15-21.

Vatsavayai, S. C., et al., 2007. Progressive CAG expansion in the brain of a novel R6/1-89Q mouse model of Huntington's disease with delayed phenotypic onset. Brain Res Bull. 72**,** 98-102.

Velier, J., et al., 1998. Wild-type and mutant huntingtins function in vesicle trafficking in the secretory and endocytic pathways. Exp Neurol. 152**,** 34-40.

Waelter, S., et al., 2001. The huntingtin interacting protein HIP1 is a clathrin and alpha-adaptin-binding protein involved in receptor-mediated endocytosis. Hum Mol Genet. 10**,** 1807-17.

Walker, F. O., 2007. Huntington's disease. Lancet. 369**,** 218-28.

Wang, F., et al., 2014. Epigenetic modifications as novel therapeutic targets for Huntington's disease. Epigenomics. 6**,** 287-97.

Wang, Q., et al., 2001. Adenosine nucleotide modulates the physical interaction between hMSH2 and BRCA1. Oncogene. 20**,** 4640-9.

Watson, L. M., Wood, M. J., 2012. RNA therapy for polyglutamine neurodegenerative diseases. Expert Rev Mol Med. 14**,** e3.

Weinberg, R. A., 2014. The biology of cancer. Garland Science, New York, NY, US.

Wells, R. D., et al., 2005. Advances in mechanisms of genetic instability related to hereditary neurological diseases. Nucleic Acids Res. 33**,** 3785-98.

Wells, R. D., et al., 1998. Small slipped register genetic instabilities in Escherichia coli in triplet repeat sequences associated with hereditary neurological diseases. J Biol Chem. 273**,** 19532-41.

Westphalen, A. A., et al., 2005. Evidence for genetic anticipation in hereditary non-polyposis colorectal cancer. Hum Genet. 116**,** 461-5.

Wexler, N. S., et al., 2004. Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. Proc Natl Acad Sci U S A. 101**,** 3498-503.

Wheeler, V. C., et al., 1999. Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. Hum Mol Genet. 8**,** 115-22.

Wheeler, V. C., et al., 2007. Factors associated with HD CAG repeat instability in Huntington disease. J Med Genet. 44**,** 695-701.

Wold, M. S., 1997. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu Rev Biochem. 66**,** 61-92.

Wong, H. K., et al., 2007. Cockayne syndrome B protein stimulates apurinic endonuclease 1 activity and protects against agents that introduce base excision repair intermediates. Nucleic Acids Res. 35**,** 4103-13.

Xia, J., et al., 2003. Huntingtin contains a highly conserved nuclear export signal. Hum Mol Genet. 12**,** 1393-403.

Xu, C. F., et al., 1995. Distinct transcription start sites generate two forms of BRCA1 mRNA. Hum Mol Genet. 4**,** 2259-64.

Xu, M., et al., 2014. Base excision repair of oxidative DNA damage coupled with removal of a CAG repeat hairpin attenuates trinucleotide repeat expansion. Nucleic Acids Res. 42**,** 3675-91.

Yamamoto, A., et al., 2000. Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. Cell. 101**,** 57-66.

Yaturu, S., et al., 2005. Preliminary evidence of genetic anticipation in type 2 diabetes mellitus. Med Sci Monit. 11**,** CR262-265.

Yavuz, K. F., et al., 2013. Aripiprazole treatment for choreoathetoid and psychotic symptoms of Huntington's disease. J Neuropsychiatry Clin Neurosci. 25**,** E31.

Yeh, H. H., et al., 2013. Histone deacetylase class II and acetylated core histone immunohistochemistry in human brains with Huntington's disease. Brain Res. 1504**,** 16-24.

Yu, A., et al., 1995. The trinucleotide repeat sequence d(GTC)15 adopts a hairpin conformation. Nucleic Acids Res. 23**,** 2706-14.

Yu, J., et al., 2007. Induced pluripotent stem cell lines derived from human somatic cells. Science. 318**,** 1917-20.

Zacharias, M., 2001. Conformational analysis of DNA-trinucleotide-hairpin-loop structures using a continuum solvent model. Biophys J. 80**,** 2350-63.

Zeitlin, S., et al., 1995. Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. Nat Genet. 11**,** 155-63.

Zhai, W., et al., 2005. In vitro analysis of huntingtin-mediated transcriptional repression reveals multiple transcription factor targets. Cell. 123**,** 1241-53.

Zhang, Y., et al., 2009. Allele-specific silencing of mutant Huntington's disease gene. J Neurochem. 108**,** 82-90.

Zhao, X. N., Usdin, K., 2014. Gender and cell-type-specific effects of the transcription-coupled repair protein, ERCC6/CSB, on repeat expansion in a mouse model of the fragile X-related disorders. Hum Mutat. 35**,** 341-9.

Zhao, X. N., Usdin, K., 2015a. The Repeat Expansion Diseases: The dark side of DNA repair. DNA Repair (Amst). 32**,** 96-105.

Zhao, X. N., Usdin, K., 2015b. The transcription-coupled repair protein ERCC6/CSB also protects against repeat expansion in a mouse model of the fragile X premutation. Hum Mutat. 36**,** 482-7.

Zheng, Z., et al., 2013. An N-terminal nuclear export signal regulates trafficking and aggregation of Huntingtin (Htt) protein exon 1. J Biol Chem. 288**,** 6063-71.

Zoghbi, H. Y., Orr, H. T., 1999. Polyglutamine diseases: protein cleavage and aggregation. Curr Opin Neurobiol. 9**,** 566-70.

Zoghbi, H. Y., Orr, H. T., 2000. Glutamine repeats and neurodegeneration. Annu Rev Neurosci. 23**,** 217-47.

Zuccato, C., Cattaneo, E., 2007. Role of brain-derived neurotrophic factor in Huntington's disease. Prog Neurobiol. 81**,** 294-330.

Zuccato, C., Cattaneo, E., 2009. Brain-derived neurotrophic factor in neurodegenerative diseases. Nat Rev Neurol. 5**,** 311-22.

Zuccato, C., et al., 2001. Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. Science. 293**,** 493-8.

APPENDIX

LIST OF FIGURES WITH PERMISSIONS OBTAINED FROM PUBLISHERS

ELSEVIER LICENSE
TERMS AND CONDITIONS
Oct 16, 2015

This is a License Agreement between Peter Mollica ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.
All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.
Supplier-Elsevier Limited
The Boulevard,Langford Lane
Kidlington,Oxford,OX5 1GB,UK
Registered Company Number-1982084
Customer name-Peter Mollica
Customer address-4064 Bridgehampton Lane
VIRGINIA BEACH, VA 23455
License number-3730620361424
License date-Oct 16, 2015
Licensed content publisher-Elsevier
Licensed content publication-Structure
Licensed content title-Secondary Structure of Huntingtin Amino-Terminal Region
Licensed content author-Mee Whi Kim,Yogarany Chelliah,Sang Woo Kim,Zbyszek Otwinowski,Ilya Bezprozvanny
Licensed content date-9 September 2009
Licensed content volume number-17
Licensed content issue number-9
Number of pages-8
Start Page-1205
End Page-1212
Type of Use-reuse in a thesis/dissertation
Portion-figures/tables/illustrations
Number of figures/tables/illustrations-1
Format-print
Are you the author of this Elsevier article? No
Will you be translating? No
Original figure numbers Figure 1. b.
Title of your thesis/dissertation
DNA REPAIR DEFICIENCY IN HUNTINGTON'S DISEASE FIBROBLASTS AND INDUCED PLURIPOTENT STEM CELLS
Expected completion date Dec 2015
Estimated size (number of pages) 166
Elsevier VAT number GB 494 6272 12
Permissions price 0.00 USD
VAT/Local Sales Tax 0.00 USD / 0.00 GBP
Total 0.00 USD

NATURE PUBLISHING GROUP LICENSE
TERMS AND CONDITIONS

This is a License Agreement between Peter Mollica ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3720920728878
License date	Oct 02, 2015
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Chemical Biology
Licensed content title	Replication-dependent instability at (CTG)•(CAG) repeat hairpins in human cells
Licensed content author	Guoqi Liu, Xiaomi Chen, John J Bissler, Richard R Sinden, Michael Leffak
Licensed content date	Aug 1, 2010
Volume number	6
Issue number	9
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	FIGURE 1: HAIRPIN-INDUCED TRINUCLEOTIDE REPEAT INSTABILITY
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	DNA REPAIR DEFICIENCY IN HUNTINGTON'S DISEASE FIBROBLASTS AND INDUCED PLURIPOTENT STEM CELLS
Expected completion date	Dec 2015
Estimated size (number of pages)	166
Total	0.00 USD

VITA

Peter A. Mollica

Address

Department of Biological Sciences, Old Dominion University, Norfolk, VA, 23529

Education

- Ph.D. Biomedical Sciences, December 2015, Old Dominion University, Norfolk, VA
- Certificate, Molecular Diagnostics, April 2012, Old Dominion University, Norfolk, VA
- B.S. Biological Sciences, December 2011, Old Dominion University, Norfolk, VA

Publications

- **P. Mollica**, R. Ogle, J. Reid, R. Bruno, and P. Sachs. Huntington's disease induced pluripotent stem cells show enhanced levels of DNA repair gene expression and trinucleotide repeat induced stability. (In progress).
- **P. Mollica**, R. Ogle, J. Reid, P. Sachs, and R. Bruno. DNA methyltransferase inhibitor 5-azacytidine recovers CNA repair gene expression and stabilizes trinucleotide repeats in Huntington's disease fibroblasts. (Submitted).
- J. Reid, **P. Mollica**, R. Bruno, and P. Sachs. Optimizing fabrication processes in micro-extrusion-based bioprinting using a 3D-printed micro-extrusion device. (In progress).
- **P. Mollica**, T. Sundin, P. Hentosh. Differentially expressed DNA repair genes in Huntington's disease fibroblasts: Preliminary Analysis. (Submitted).