The relationship of microRNAs to clinical features of Huntington's and Parkinson's disease
BOSTON UNIVERSITY
SCHOOL OF MEDICINE

Dissertation

THE RELATIONSHIP OF MICRORNAS TO CLINICAL FEATURES OF
HUNTINGTON’S AND PARKINSON’S DISEASE

by

ANDREW G HOSS
B.A., Binghamton University, 2006
M.S., Tufts University, 2010

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
2016
Approved by

First Reader

Richard H. Myers, Ph.D.
Professor of Neurology

Second Reader

Jeanne C. Latourelle, Ph.D.
Assistant Professor of Neurology
DEDICATION

I would like to dedicate this work to my wife, Jenn - without your constant encouragement, love and self-sacrifice, this process would never have been possible 😊
ACKNOWLEDGMENTS

Adam, Jeanne and Rick - I could not have done this without you.
THE RELATIONSHIP OF MICRORNAS TO CLINICAL FEATURES OF HUNTINGTON’S AND PARKINSON’S DISEASE

ANDREW G HOSS
Boston University School of Medicine, 2016
Major Professor: Richard H. Myers, Ph.D., Professor of Neurology

ABSTRACT

MicroRNAs (miRNAs) represent a major system of post-transcriptional regulation, by either preventing translational initiation or by targeting messenger RNA transcripts for storage or degradation. miRNA deregulation has been reported in neurodegenerative disorders, such as Huntington’s disease (HD) and Parkinson’s disease (PD), which may impact gene expression and modify disease progression and/or severity. To assess the relationship of miRNA levels to HD, small RNA sequence analysis was performed for 26 HD and 36 non-disease control samples derived from human prefrontal cortex. 75 miRNAs were differentially expressed in HD brain as compared to controls at genome-wide significance (FDR q<0.05). Among HD brains, nine miRNAs were significantly associated with the extent of neuropathological involvement in the striatum and three of these significantly related to a continuous measure of striatal involvement, after statistical adjustment for the contribution of HD gene length. Five miRNAs were identified as having a significant, inverse relationship to age of motor onset, in particular, miR-10b-5p, the mostly strongly over-expressed
miRNA in HD cases. Although prefrontal cortex was the source of tissue profiled in these studies, the relationship of miR-10b-5p levels to striatal involvement in the disease was independent of cortical involvement. In blood plasma from 26 HD, 4 asymptomatic HD gene carriers and 8 controls, miR-10b-5p levels were significantly elevated in HD as compared to non-diseased and preclinical HD subjects, demonstrating that miRNA alterations associated with diseased brain may be detected peripherally. Using small RNA sequence analysis for 29 PD brains, 125 miRNAs were identified as differentially expressed at genome-wide significance (FDR q<0.05) in PD versus controls. A set of 29 miRNAs accurately classified PD from non-diseased brain (93.9% specificity, 96.6% sensitivity, 4.8% absolute error). In contrast to HD, among PD cases, miR-10b-5p was significantly decreased and had a significant, positive association to onset age independent of age at death. These studies provide a detailed miRNA profile for HD and PD brain, identify miRNAs associated with disease pathology and suggest miRNA changes observed in brain can be detected in blood. Together, these findings support the potential of miRNA biomarkers for the diagnosis and assessment of progression for neurodegenerative diseases.
# TABLE OF CONTENTS

DEDICATION ...................................................................................................................... iv  
ACKNOWLEDGMENTS ........................................................................................................ v  
ABSTRACT ........................................................................................................................... vi  
TABLE OF CONTENTS ....................................................................................................... viii  
LIST OF TABLES .................................................................................................................. xi  
LIST OF FIGURES ............................................................................................................... xii  
LIST OF ABBREVIATIONS ................................................................................................. xiv  

Chapter 1: INTRODUCTION ............................................................................................. 1  
  Background ....................................................................................................................... 1  
  Neuropathology .................................................................................................................. 2  
  Molecular pathogenesis of HD ......................................................................................... 3  
Parkinson’s disease ............................................................................................................ 6  
  Background ....................................................................................................................... 6  
  Genetic heterogeneity ........................................................................................................ 8  
Preclinical and progressive biomarkers for neurodegenerative disease................. 9  
  Huntington’s disease biomarkers .................................................................................... 9  
  Parkinson’s disease biomarkers ..................................................................................... 13  
Role of miRNAs in neurodegenerative diseases ....................................................... 16  
Hypothesis and Specific Aims ....................................................................................... 21  

Chapter 2: microRNAs located in the Hox gene clusters are implicated in Huntington’s disease pathogenesis .......................................................... 24  
  Abstract ............................................................................................................................. 24  
  Introduction ....................................................................................................................... 25
LIST OF TABLES

Table 1. HD brain samples analyzed for mRNA-seq, miRNA-seq and RT-qPCR validation of miR-10b-5p. ................................................................. 27

Table 2. Control brain samples analyzed for mRNA-seq, miRNA-seq and RT-qPCR validation of miR-10b-5p ................................................................. 28

Table 3. Differentially expressed miRNAs from miRNA-seq ................................................................. 29

Table 4. 22 differentially expressed targets of miR-10b-5p, miR-196a, miR-196b, miR-1247 and miR-615-3p. ........................................................................ 40

Table 5. Differential expression of Hox cluster genes in HD. ................................................................. 41

Table 6. Summary of the brain samples used for miRNA-sequence analysis. ........................................ 77

Table 7. Summary of the samples used for the study ........................................................................ 111

Table 8. Comparison of miR-10b-5p, miR-486-5p, miR-132-3p, and miR-363-3p levels in plasma between control and HD patients ........................................ 114

Table 9. Summarized sample information ......................................................................................... 125
LIST OF FIGURES

Figure 1. Schematic of miRNA biogenesis. ................................................................. 17
Figure 2. Proposed HD miRNA mechanisms. ............................................................. 18
Figure 3. Proposed PD miRNA mechanisms. ............................................................. 20
Figure 4. miR-196a-5p, miR-10b-5p and miR-615-3p were found significantly
differentially expressed in Huntington's disease ............................................... 29
Figure 5. Neuron counts from prefrontal cortical tissue homogenate ..................... 31
Figure 6. miR-10b-5p expression in control, Parkinson's disease and
Huntington's disease prefrontal cortex .............................................................. 33
Figure 7. miR-10b-5p overexpressing PC12 Q73 cells exhibit reduced
cytotoxicity ............................................................................................................. 35
Figure 8. Differentially expressed miRNAs in HD are located in Hox gene
clusters .................................................................................................................... 43
Figure 9. Characterization of miRNA in Huntington's disease brain ..................... 79
Figure 10. Nine miRNAs are associated with Vonsattel grade .............................. 81
Figure 11. Association of clinical features to HD CAG repeat size ....................... 84
Figure 12. miR-10b is associated with age of onset and striatal involvement ......... 86
Figure 13. CAG-adjusted clinical features of HD show patterns of association
with miRNA expression .......................................................................................... 88
Figure 14. Gene ontology terms are similar for mRNA targets of clinically relevant
deregulated miRNAs .............................................................................................. 91
Figure 15. miR-10b-5p has an ordinal association with HD stage ....................... 115
Figure 16. miRNA changes related to Parkinson's disease. 129
Figure 17. miRNA profile for Parkinson's disease with dementia. 130
Figure 18. miR-10b-5p levels are associated with motor onset age in both Parkinson's and Huntington's disease. 132
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>BA9</td>
<td>Brodmanns Area 9</td>
</tr>
<tr>
<td>CAG</td>
<td>Cytosine – adenine – guanine</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>FDR</td>
<td>False discovery rate</td>
</tr>
<tr>
<td>GO</td>
<td>Gene Ontology</td>
</tr>
<tr>
<td>HD</td>
<td>Huntington’s disease</td>
</tr>
<tr>
<td>HTT</td>
<td>Huntingtin</td>
</tr>
<tr>
<td>H-V score</td>
<td>Hadzi-Vonsattel score</td>
</tr>
<tr>
<td>mHTT</td>
<td>Mutant HTT</td>
</tr>
<tr>
<td>miRNA</td>
<td>MicroRNA</td>
</tr>
<tr>
<td>MSN</td>
<td>Medium spiny neurons</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>PDD</td>
<td>Parkinson’s disease with dementia</td>
</tr>
<tr>
<td>PMI</td>
<td>Post-mortem interval</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver operator characteristic</td>
</tr>
<tr>
<td>SNCA</td>
<td>Alpha-synuclein</td>
</tr>
<tr>
<td>UHDRS</td>
<td>Unified HD rating scale</td>
</tr>
<tr>
<td>UPDRS</td>
<td>Unified PD rating scale</td>
</tr>
</tbody>
</table>
Chapter 1: INTRODUCTION

Background

Huntington’s disease (HD) is a neurodegenerative disease caused by a cytosine-adenine-guanine (CAG) repeat expansion in the HTT gene located on chromosome 4p16.3 [1]. HD is a chronic, progressive, and fatal disease, with motor, cognitive and psychiatric symptoms. Most individuals diagnosed with HD first present symptoms at middle age, beginning with the emergence of insidious movement abnormalities, which progresses into abrupt, nonrandom, involuntary choreiform movements. Chorea is a key symptom of HD and defines disease diagnosis. Psychological problems, such as depression or psychosis are commonly observed [2]. Cognitive decline and loss of executive function in advanced stage HD is inevitable. Survival from onset to death averages 17 to 20 years.

The disease is autosomal dominant transmitted, affecting an estimated four to eight people per 100,000 in Western countries [3]. The normal range of the CAG repeat region of HD gene is 8 to 26 repeats, where 40 or more repeats is pathogenic [4, 5]. Individuals with an intermediate repeat range between 27 and 35 [6], and range of reduced repeats between 36 and 39, may also develop symptoms [7]. The CAG repeat size is inversely related to age of onset, where individuals with longer CAG repeat sizes are predicted to have earlier onset than individuals with short CAG repeat sizes [8] (see Figure 1.1). Although all HD is
caused by the same mutation, variability among individuals with HD is widely seen, with different degrees of motor and cognitive symptoms.

**Neuropathology**

The key neuropathologic feature of HD is selective neuronal loss of medium spiny neurons (MSNs) in the caudate nucleus and putamen, substructures of the striatum. Striatal atrophy begins at the tail of the caudate, which progresses in a rostral to caudal direction. Neuronal loss in the head of the caudate, near the lateral ventricle, and in the putamen, at the region closest to the internal capsule, occurs medially [9, 10]. Several imaging-based studies have identified a strong relationship between clinical features and the extent of brain atrophy in HD [11], indicating the loss of the inhibitory GABAergic MSNs in the striatum is associated with the motor aspects of the disease. Volumetric measurements of the caudate and putamen by MRI suggest that atrophy in the striatum begins 9 to 20 years prior to diagnosis [12].

By clustering neuropathological evaluations of 41 brain regions (rated 0-4, “0” being absent of change, “4” representing the highest degree of change), quantitative scores were created to represent the combined involvement of the disease within the striatal and cortical brain regions [13]. Termed the “Hadzi-Vonsattel” score, 523 brains from the Myers lab were scored by a neuropathologist and assessed for their relationship to disease features. Hadzi-Vonsattel striatal score was found to be strongly related to the size of the CAG expansion, where larger repeats are associated with more severe involvement at
time of death [10]. Cortical atrophy, however, did not relate to the extent of striatal involvement nor to CAG repeat size [10]. Variability in the relative to the involvement of cortical and subcortical regions likely influences the wide range of motor versus cognitive and psychiatric symptoms recognized among HD patients. While both cortical and subcortical regions are involved in HD [9, 14, 15], the cortex is considerably less involved than the striatum. The mechanisms that lead to patient-to-patient variability in cortical and CAG-independent striatal neurodegeneration remain unknown.

**Molecular pathogenesis of HD**

The HD mutation produces a neurotoxic, mutant HTT protein (mHTT). The polyglutamine tract expansion caused by the HD mutation exists within the N-terminal region of the HTT protein [4], which leads to aggregate formation in the brains of HD patients [16] as well as in HD *in vitro* models [17].

Cytosolic and intranuclear inclusions containing mHTT fragments in neurons of the striatum and cerebral cortex are a histopathological hallmark of HD [18]. In post-mortem HD brains, extensive aggregation is observed in the striatum and in sub-layers of the cerebral cortex, and to a lesser extent, in the basal ganglia, hippocampus, and cerebellum [18]. Striatal MSNs and cortical pyramidal interneurons exhibit extensive aggregation, where aggregates initially localize to the soma and dendritic processes and later as the disease progresses, intranuclear formation occurs [18].
Aberrant nuclear localization of mHTT fragments precedes cell death in neurons [19]. However, the process by which preferential death of striatal MSNs occurs is not fully understood. HD pathogenesis is likely caused by a gain-of-function of mHTT, where the polyglutamine repeat expansion contributes to the misfolding and aggregate form of HTT, resulting in a protein with new, toxic function [20] or by the loss of wildtype HTT function through a loss-of-function or dominant-negative effect.

Gain-of-function is the most well accepted hypothesis for HD pathogenesis. Transcription factors, such as CREB-binding protein (CBP) [21], specificity protein 1 (SP1) [22], TATA-binding protein (TBP) [23] and p53 [21], and transcriptional repressors such as REST/NSRF [24], spuriously bind to mHTT intranuclear inclusions impairing their normal activity, leading to transcriptional dysregulation [25]. The transcription factor nuclear factor kappa-light-chain-enhancer of activated B cells (NFkB), which is particularly related to immune function, has shown altered activity in HD due to direct, aberrant interactions with soluble mHTT [26, 27].

The HD gene encodes a large, 348 kDa cytosolic protein (HTT) that is widely distributed throughout the body, with its highest levels in the brain [28, 29]. The protein structure of HTT suggests it is a multifunctional scaffolding protein, with multiple HEAT repeats and hydrophobic alpha-helices that assist in protein-protein interactions, and a nuclear export signal [30] which aids in nuclear transport. HTT directly associates with the microtubule motor dynenin [31], and
indirectly binds kinesin-1 through Huntingtin-associated protein 1 [32],
modulating the bidirectional transport of endocytic and exocytic vesicles along
cytoskeletal tracks [33]. Within RNA transport granules, HTT colocalizes with
Argonaute 2, a component of the RNA-induced silencing complex, indicating it
may play a role in localized translational control through the silencing of RNA
transcription [34, 35]. HTT is critical for nervous system development [36] and
embryonic lethal in knockout huntingtin mice [37]. HTT may also play a role in
cytokine production in monocytes and macrophages under normal and HD
settings [38].

HD is associated with the disruption of an assortment of cellular
processes, impairing cell adhesion[39], mitochondrial energetics [40], cellular
transport [41], autophagic processing [42], ubiquitin–proteasome system [43],
synaptic plasticity [44] and enhancement of glutamate-related excitotoxicity [45].
HTT cleavage is mediated by caspases to produce an amino-terminal polyQ
containing fragment [46], and mHTT is more susceptible to cleavage than its
normal counterpart. The N-terminal region is cleaved by caspases and calpains
to produce N-terminal fragments in both normal and HD cells, where only the
expanded polyQ fragments of HD self-associate and seed aggregates [47–49].
Activation of caspases induces apoptosis; thus, overactive caspases have
deleterious effects on already stressed neurons.
Parkinson’s disease

Background

Parkinson’s disease (PD) is the second most common neurodegenerative disease, characterized by resting tremor, rigidity, bradykinesia and postural instability [50]. The loss of dopamine-producing neurons in the substantia nigra par compacta, a substructure of the basal ganglia, is a hallmark of PD neuropathology. Reduction of dopaminergic neurons results in dopamine depletion, improper regulation of the nigrostriatal pathways, consequently hypokinesia.

Neuronal cytoplasmic inclusions, termed Lewy bodies, and Lewy neurites found in neuronal processes, are observed throughout the brain [51]. Lewy bodies predominantly consist of α-synuclein protein, and to a lesser extent, ubiquitin, calbindin, and tubulin [51]. α-synuclein is a small, 140 amino-acid protein found in high abundance at the presynaptic terminal of neurons, representing up to 1% of all cytosolic protein found in the brain [52]. A predictable topography of pathological changes in the central nervous system is observed, as annotated by Braak [53], beginning in olfactory bulb and medulla oblongata, then progressing rostrally from the medulla to the pons, then to the substantia nigra and midbrain, limbic structures, and neocortical structures.

α-synuclein has the ability to generate β-sheet structures under specific physiological conditions, creating amyloid-like fibrils. Soluble, oligomeric α-synuclein can transition into spherical, string-like, protofibrils, which gradually
assemble into insoluble, neurotoxic fibrils [51]. Injection of protofibrils into PD mouse models results in widespread aggregate deposition in the brain [54], which indicates transmission of the protofibrils through neural connectivity.

The mechanism by which dopaminergic neurons preferentially degenerate likely occurs through several converging pathways. α-synuclein accumulation is neurotoxic. Oligomeric α-synuclein may lead to dopamine leakage of neurotransmitter filled vesicles into the cytosol, which could lead to oxidative stress-induced death [55]. Cytosolic dopamine contributes to oxidative stress and cellular damage, by disrupting glutathione homeostasis, enhancing calcium influx, increasing calcium-mediated calpain activation and supporting in free radical production [56]. Neurotoxins, such as 6-hydroxydopamine (6-OHDA), MPTP, and the pesticide paraquat, have been shown to destroy DA neurons, and individuals that have exposure to these chemicals display Parkinsonism [57]. These toxins infer with complex I of the electron transport chain, causing the increase of free radicals and halting mitochondrial metabolism, leading to cell death [58]. In PD, impaired energetics by mitochondrial dysfunction may be aggravated by abnormal proteostasis, as normal clearance of proteins by chaperone-mediate autophagy or other proteolytic mechanisms may be impaired [59]. Both neuronal and glial sources have been implicated in the contribution of oxidative stress, and chronic neuroinflammatory activity from dying cells or inclusion bodies may play a role in the degenerative process [60].
Genetic heterogeneity

The underlying mechanism of α-synuclein misfolding and insoluble aggregation is poorly understood. However, the discovery of genetic forms of PD caused by mutations of the α-synuclein gene (SNCA) [61-67], have led to its identification and insight into PD pathogenic mechanisms. Multiplication of SNCA (duplication and triplication of the gene) leads a dominant forms of PD, and SNCA levels of gene expression inversely related to age of onset of PD [68]. A missense substitution (A53T) impacts protein clearance and also leads to dominant inheritance.

Mendelian and complex forms of PD [61] may provide novel insight into disease etiology and pathological mechanisms. Mutations of the leucine-rich repeat kinase 2 gene (LRRK2) represent the most common inherited form of PD [68]. 13 other Mendelian forms have been identified, and this number will continue to grow as genome-wide methods are continually applied [68]. Many of these genes, including parkin, DJ-1, and PINK, are linked to respiration and mitochondrial function.

Due to the occurrence of reduced penetrance, variable expressivity, and environmental phenocopy effects, pedigrees often lack Mendelian patterns of inheritance. Genome-wide association studies illustrate the complex etiology of PD and genetic heterogeneity of the disease. The results from the mega-meta PD GWAS over (~13,000 PD, 80,000 control), identified 24 loci, and the second most significant locus, microtubule-associated protein tau (MAPT), also encodes
for the protein which when phosphorylated, is present in aggregates in tauopathies and Alzheimer’s disease [61]. However, GWAS variants only account for an estimated 30% of familial and 5% of sporadic, non-familial cases [69, 70].

**Preclinical and progressive biomarkers for neurodegenerative disease**

**Huntington’s disease biomarkers**

Although genetic testing allows for reliable detection of HD (by sequencing of the CAG repeat of the HD gene), there remains a need for biomarkers of disease progression in premanifest individuals. The mechanism by which premanifest HD gene carriers transition to manifest HD remains unknown. Slowing disease progression may delay the onset of disease symptoms.

Clinical progression of HD can be measured by a number of clinical tests. The Unified Huntington Disease Rating Scale (UHDRS) is the gold standard, which assesses of the degree of motor, cognitive, and behavioral involvement as well as total functional capacity (TFC) and degree of independence [2]. Although UHDRS may be able to determine the extent of deterioration of executive functioning (as measured by TFC), its utility is limited to symptomatic, early and mid-stage HD patients.

As HD progresses, the brain loses significant volume, and these early structural changes are observable using magnetic resonance imaging (MRI) in premanifest individuals. A 50% reduction in volume of the putamen and 28% reduction of the caudate nucleus has been observed in early stage HD patients.
TRACK-HD, a large (n=366) cross-sectional study aimed to discover baseline predictors of HD progression, observed volumetric changes in premanifest HD across a 36 month period, and in early symptomatic HD patients, these changes correlated with progression [72]. Results from PREDICT-HD, another large HD cohort, estimated that volumetric brain changes might be detectable up to 15–20 years prior to diagnosis [73]. The severity of HD clinical manifestations relate to the loss of neurons, but also to neuronal dysfunction and synaptic changes as measured by functional MRI (fMRI). Increased activation of the dorsolateral prefrontal cortex was observed in presymptomatic individuals as they approached disease onset, independent of cortical atrophy [74]. Increased neuroinflammation, measured by microglial activation [75, 76] or proinflammatory cytokines (such as TNF-α, IL-6, IL-8) in blood has been observed in premanifest and early stage HD [77].

Altered dopamine signaling may play a role in HD pathogenesis. Positron emission tomography (PET) based imaging using radiolabeled striatal dopamine receptors (D1, D2) and well as dopamine transporter (DAT), have shown decreased postsynaptic receptor density and binding in asymptomatic HD gene carriers and HD patients [78]. Brain metabolism can also be measured using PET. Reduced glucose uptake was first observed in mice with neostriatal lesions, using 14C-2-deoxyglucose (DG) autoradiography [79] and later, glucose hypometabolism in human neostriatum was shown to preceded degeneration and correlate with motor impairments [80-83].
Increased levels of leukocyte 8-hydroxydeoxyguanosine (8-OHdG), a marker of free radical oxidative damage to DNA, have been observed in the caudate and parietal cortex [84], serum [85], and leukocytes [86]. Increased 8-OHdG levels were observed in premanifest individual in longitudinal studies, with higher levels in those with larger repeats or older ages [87]. However, these results were not repeatable in an independent cohort, assayed by two separate using two assay protocols (LCECA, LC-MSMS) [88]. Phosphodiesterase 10A (PDE10A) using $[^{18}F]$MNI-659, PET imaging tracer selective for the striatum, has shown promise in preliminary cross-sectional and longitudinal studies, where reduced PDE10A levels correlate with clinical measures of disease severity [89].

Although precise and accurate anatomical or functional endpoints can be measured using neuroimaging, poor reproducibility can result due to the lack of standardization. Imaging is susceptible to unreliability across centers, due to differences in data acquisition, the handling of motion artifacts, scanner models, image processing, the software for normalization, the reference for spatial mapping and overall analysis [90].

Ideally, assays that predict the continuum of HD would be derived from peripheral biofluids, due to cost effectiveness, simplicity and ease of clinical adoption. As immunoassays improve, mHTT may be a tractable biomarker for HD progression. The extent of accumulation of mHTT protein in peripheral fluids such as sorted blood cells [91] and cerebrospinal fluid [92], can be measured
using high-resolution immunoassays. In cerebrospinal fluid, using single molecular counting for mHTT, mHTT was detectable only in HD gene carriers, was three times higher in symptomatic HD as compared to premanifest HD gene carriers, and was associated with of cognitive and motor dysfunction in symptomatic HD patients [92]. However, these results, while encouraging, are relatively new and yet to be reproduced.

Several studies describe the early and chronic elevation of proinflammatory cytokines in HD blood cells but these signals do not provide the resolution needed to track disease progression [31-33]. Researchers have attempted to use a candidate gene approach, in particular in probing inflammatory cytokine genes, with little success [34,35].

Genome-wide genomic profiling is an attractive option for HD diagnostics, as these tests have the ability to measure cellular state with high-resolution. A set candidate genes identified in genome-wide studies can be easily standardized with PCR-based tests or as DNA sequencing technologies continue to improve, a set of genes could efficiently tested with a single assay. H2A histone family, member Y (H2AFY) mRNA was found to overexpressed in HD prefrontal cortex and HD patients compared to controls, and was recent validated by an independent study [93], which used RNA-sequencing of peripheral blood cells to identify 167 genes that related to clinical motor scores [93]. Although these results lacked concordance to a previous blood study [94], the concordance of H2AFY mRNA is encouraging. Discrepancies could be due to the
technologies used (microarray versus sequencing), the susceptibility of RNA to nucleased and the relatively short half-life of messenger RNA, estimated from 15 minutes to 17 hours *in vitro* [36]. Other representative genomic markers, such as miRNAs, with high stability and long half-lives [38-40] may be more useful for HD prediction.

**Parkinson’s disease biomarkers**

PD diagnostics apply similar strategies as in HD, however isolating preclinical PD patients is difficult. Additionally, the diagnosis of idiopathic PD is encumbered by the heterogeneous presentation of symptoms, distinctive clinical subtypes (PD, PDD), and variable rates of progression. 30% of PD patients have PD-related dementia [95]. Monogenic, familial forms only represent a small sample of PD patients.

Three phases have been proposed – preclinical, premotor, and motor phases of PD [96]. In the preclinical phase, no symptoms are present. Nonmotor manifestations are detectable in the premotor phase. Prodromal or premotor markers include olfactory dysfunction [97], neuropsychiatric, and autosomal changes such as gastrointestinal complications [98], or sleep abnormalities (excessive sleepiness [99], idiopathic REM sleep behavior disorder [100]). Over 90% of patients with PD have olfactory dysfunction [97], and using the Brief Smell Identification Test in 2,267 healthy males, longitudinal tests found the odds of PD development within 4 years in the lowest quartile of olfactory acuity to be 5.2 times that observed in the highest two quartiles [101]. Clinical measures of
PD progression of motor symptoms and cognitive decline can be assessed by the Unified Parkinson's Disease Rating Scale (UPDRS), a battery of tests comprised of six tests, measuring cognitive health, complications of therapy, extent of motor involvement, level of independence as severity of PD symptoms as measured by the Hoehn and Yahr scale [102]. These measures are used for diagnosis, clinical evaluation and monitoring of the rate of functional decline in PD.

The onset of DA loss precedes motor symptoms. Image based analysis using radioligands for DA transporter to assess the integrity of the dopaminergic system have been successful in some instances of differentiating the degree of PD severity [103]. In a longitudinal 3-center observational study with 37 months of prospective follow-up, a 17-fold increased for the relative risk for developing PD was observed among those with transcranial sonography substantia nigra hyperechogenicity [104], occurring in over 90% of PD patients, compared to 10% of healthy older adults. Histopathological evidence from gastrointestinal tract neurons, from colonic submucosa of PD patients, was positive for α-synuclein accumulation [105]. Biochemical screening methods for Lewy body related proteins in peripheral fluids like CSF and blood might act as high-resolution biomarkers. In postmortem CSF from PD and control subjects, assaying six proteins (fibrinogen, transthyretin, apolipoprotein E, clusterin, apolipoprotein A-1, and glutathione-S-transferase-Pi), using two-dimensional difference gel electrophoresis, PD samples differed by 1.5-fold in protein concentration [106].
Comparing protein levels in the CSF of 42 PD and 69 controls, in PD, increased α-synuclein levels were associated with progression of motor symptoms and cognitive decline over 2 years, lower levels of Aβ was associated with worsening of performance on delayed memory recall, and high levels of phosphorylated tau were associated with worsening in motor symptoms [107].

Genome-wide transcriptomics have been used for PD biomarker discovery. In an RNA-seq study of blood, asymptomatic individuals with the G2019S LRRK2 mutation were compared to idiopathic PD and health controls, to identify genes that may be specifically altered in LRRK-related PD [1]. In a study of 105 total patients, microarrays were used to create a detail transcriptomic profile of early-stage PD peripheral blood, and though differentially expressed genes were identified, classification of PD based on RNA expression failed in validation [2]. Reanalysis of these data identified a set of 10 genes that may predict PD using PCA [3]. Levels of seven brain-relate genes were measured in blood samples from 62 early-stage PD patients and 64 controls, and using a stepwise multivariate logistic regression analysis, five genes were identified as disease predictors of PD (sensitivity 90.3, specificity 89.1, ROC AUC=0.96, validated AUC=0.95) [4]. After induction of deep brain stimulation, RNA changes were observed in PD leukocytes [108, 109]. Using network analysis from four independent microarray studies, hepatocyte nuclear factor 4 alpha (HNF4A) and polypyrimidine tract binding protein 1 (PTBP1) were significantly altered in blood from two independent cohorts, and based on their levels, exhibited significant
longitudinal changes over the course of the 3 year follow-up period and classified PD with 90% sensitivity and 80% specificity [110]. mRNA biomarkers in blood may be informative for identifying PD-related genes, but due to their inherent lability and patient-to-patient variability, other markers, such as microRNAs, may be representative of mRNA changes but more biochemically stable.

**Role of miRNAs in neurodegenerative diseases**

MicroRNAs (miRNAs) are small non-coding RNAs, 18-25 nucleotides in length, which bind complementary sites within the 3’ untranslated regions of messenger RNA. Through the assembly of the RNA-induced silencing complex (RISC), the duplexed transcript signals for degradation by endonuclease cleavage, or preventing translation of the gene product by stalling protein synthesis machinery [111] (see Figure 1). As translational regulators of mRNA expression, miRNAs can directly control cell fate decisions, and consequently, miRNAs are involved in numerous aspects of embryonic and neurodevelopment [112].
Figure 1. Schematic of miRNA biogenesis. Primary transcripts are transcribed and processed by DROSHA, exported from the nucleus by Exportin 5, trimmed by DICER and loaded into RISC to become active silencers. Insoluble aggregates can affect normal miRNA mechanisms, by increasing mRNA content due to spurious protein binding or impairing miRNA complex activity.

Dysregulation of miRNAs has been linked to neurological and neurodegenerative disorders [113] and several studies have explored the role of miRNAs in HD (see Figure 2). In targeted study, Parker et al 2008 observed miRNAs predicted to target RE1-silencing transcription factor (REST), a transcriptional repressor which has enhanced function as mHTT fails to prevent its nuclear translocation and subsequent downstream repressive effects of neuroprotective genes [114]. Marti et al [115] performed miRNA-sequencing for two pooled HD samples and two pooled control samples, and while a number of miRNA differences were observed, differential expression could not be performed. Altered expression of miRNAs, quantified using microarray technology, has been reported in cellular models of HD [114, 116, 117] and in mouse models of HD [117-120], which either focus on REST/CoREST-related
miRNAs or miRNAs which are predicted to target *HTT*. Most telling may be the interaction of HTT with important miRNA-related machinery. Wildtype HTT was found to colocalize with Argonaute 2, the endonuclease required for RNA-mediated gene silencing by the RNA-induced silencing complex (RISC), within neuronal transport RNA granule [121, 122], suggesting HTT may facilitate miRNA-mediated mRNA degradation and/or localized translation of specific mRNAs.

Figure 2. Proposed HD miRNA mechanisms.

Kim et al. 2007 was the first to highlight the importance of miRNA function within the dopaminergic system [123]. Dicer, the enzyme responsible for the cleavage of primary miRNA into their mature, active form, was genetically removed exclusively in post-mitotic, dopaminergic neurons. Dicer-null neurons in these transgenic mice into lived into adulthood, but later in life, mice displayed PD-like phenotypes and dopamine-expressing neurons were progressively lost.
In vitro, these neurons were unable to maturate but were effectively rescued with the addition of small endogenous RNAs [123]. While specific miRNAs were altered in the midbrain of these mice, subsequent genetic studies have found no association of these miRNAs or their target to PD [124]. However, miRNAs have been observed to modulate of alpha-synuclein protein expression, and single nucleotide polymorphisms (SNPs) in the 3'UTR of SNCA [125], as well as SNCA 3'UTR length, may be associated with PD, revealing a potential role of SNCA-binding miRNAs in PD [126-128]. Several miRNAs that we reported altered in PD brain may have an indirect relationship to PD-related genes (see Figure 3).

Genetic forms of PD include mutations within the α-synuclein gene (SNCA), as well in Leucine-rich Repeat Kinase 2 (LRRK2), one of the most common causes of familial PD [61]. miR-34b/c was identified as being altered in brain from microarray, however genome-wide changes were not reported [129]. A number of PD blood studies have been performed with little concordance [130-134], although miRNA in blood of PD patients after Levodopa treatment showed suggestive alterations [135]. In Burgos et al. 2014, small RNA sequencing was performed for blood serum and CSF from 67 PD and 78 control postmortem subjects [136], a number of changes were reported, but without information from brain, were difficult to interpret.
Commonly, miRNA biomarker studies have been initiated in blood or other peripheral biofluids, and their relationship to neuropathological involvement has not been considered. As noted above, both HD and PD are known to have physiological complications such as gastrointestinal problems, weight loss, sleep disorders, as well as autonomic and sensory dysfunctions. It is likely that some of these changes produce alterations in blood that are not indicative of the underlying neurodegenerative process and thus would not predict the neurologically dependent disease processes. By using primarily affected tissue, such as prefrontal cortex, from postmortem patient samples as an initial miRNA screening approach, candidate miRNAs can be selected for studying the periphery, decreasing noise and irrelevant signals such as physiological changes which do not explain disease progression. A comprehensive evaluation of the correspondence of brain miRNA and miRNA detectible in biofluids is the logical first in the effective detection of HD and PD biomarkers.
Biochemically, miRNAs are relatively stable. The half-life of the majority miRNA has been predicted to be on average 5 days [137]. Plasma miRNAs have been observed unaffected by extreme heat or pH, long-time storage at room temperature, or multiple freeze-thaw cycles [38-40]. The small size of miRNA allow higher potential for neuropathologically specific miRNA to cross the blood-brain barrier in exosomes [138] and circulate stably in peripheral fluids as cell free molecules [139]. The ability of miRNA to circulate provides a unique opportunity to monitor the neurological disease processes from the periphery.

miRNA sequencing offers high-resolution detection over a wide dynamic range of miRNA levels to create a detailed miRNA profile. miRNA sequence analysis would provide the most accurate and comprehensive miRNAs signature in HD or PD, to identify neuropathologically and potentially clinically relevant miRNA for biomarker discovery.

**Hypothesis and Specific Aims**

**Hypothesis**

We hypothesize miRNAs are altered in Huntington’s and Parkinson’s disease brain. As key post-transcriptional regulators, miRNAs associated with disease may also help describe neuropathology or disease features, such as the age of motor onset. Due to their stability and extracellular transport, brain-related miRNA changes may be observed in the blood, and therefore function as potential disease classifiers or predictors of disease phenotypes.
Specific Aims

**Aim 1: Characterize miRNA expression in HD brain.** We will perform Illumina small non-coding next-generation sequencing on 42 Brodmann Area 9 (BA9) prefrontal cortex samples (16 HD, 26 controls), and combine this cohort with our initial study of 12 HD and 9 controls to identify disease-related miRNA. Additional samples will provide statistical power and confidence needed to find differentially expressed miRNA, as well as model relationships of clinical features of HD, such as age of disease onset, age at death, and extent of neurodegeneration in striatum and prefrontal cortex, to miRNA expression.

**Aim 2: Evaluate blood plasma miRNAs as circulating biomarkers for HD disease progression.** We will assess whether brain-specific miRNA changes are peripherally observable. Using blood plasma-derived RNA collected from a cohort of 26 HD, 4 asymptomatic HD gene carriers and 8 controls, we will measure levels of miR-10b-5p, miR-486-5p, miR-132-3p and miR-363-3p, four miRNAs with associations to the extent of striatal involvement and/or onset age. We will compare the levels across the three groups to identify whether miRNA levels associate with disease stage.

**Aim 3: Characterize miRNA levels in Parkinson’s disease brain.** We will perform sequencing in 29 PD prefrontal cortex and using our existing controls, compare the two groups to identify genome-wide miRNA alterations in PD brain. We will use classification analysis to find a miRNA profile for PD, and
identify miRNAs that relate to disease phenotypes, such as onset age and dementia.
CHAPTER 2: MICRORNAS LOCATED IN THE HOX GENE CLUSTERS ARE
IMPlicated IN HUNTINGTON’S DISEASE PATHOGENESIS

Citation:

Abstract

Transcriptional dysregulation has long been recognized as central to the pathogenesis of Huntington’s disease (HD). MicroRNAs (miRNAs) represent a major system of post-transcriptional regulation, by either preventing translational initiation or by targeting transcripts for storage or for degradation. Using next-generation miRNA sequencing in prefrontal cortex (Brodmann Area 9) of twelve HD and nine controls, we identified five miRNAs (miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-615-3p and miR-1247-5p) up-regulated in HD at genome-wide significance (FDR q-value <0.05). Three of these, miR-196a-5p, miR-196b-5p and miR-615-3p, were expressed at near zero levels in control brains. Expression was verified for all five miRNAs using reverse transcription quantitative PCR and all but miR-1247-5p were replicated in an independent sample (8HD/8C). Ectopic miR-10b-5p expression in PC12 HTT-Q73 cells increased survival by MTT assay and cell viability staining suggesting increased expression may be a protective response. All of the miRNAs but miR-1247-5p are located in intergenic regions of Hox clusters. Total mRNA sequencing in the same samples identified fifteen of 55 genes within the Hox cluster gene regions
as differentially expressed in HD, and the Hox genes immediately adjacent to the four Hox cluster miRNAs as up-regulated. Pathway analysis of mRNA targets of these miRNAs implicated functions for neuronal differentiation, neurite outgrowth, cell death and survival. In regression models among the HD brains, huntingtin CAG repeat size, onset age and age at death were independently found to be inversely related to miR-10b-5p levels. CAG repeat size and onset age were independently inversely related to miR-196a-5p, onset age was inversely related to miR-196b-5p and age at death was inversely related to miR-615-3p expression. These results suggest these Hox-related miRNAs may be involved in neuroprotective response in HD. Recently, miRNAs have shown promise as biomarkers for human diseases and given their relationship to disease expression, these miRNAs are biomarker candidates in HD.

Introduction

Huntington's disease (HD) (OMIM: 143100) is an inherited neurodegenerative disorder characterized by involuntary movement, dementia, and changes in personality. HD is transmitted as an autosomal dominant disorder, for which an expansion of a CAG trinucleotide repeat within the coding region of the huntingtin gene (HTT) is the disease causing mutation [1]. The CAG repeat codes for a polyglutamine domain in the Htt protein and results in neuronal cell death predominantly affecting the caudate nucleus and putamen although neuronal loss is widespread in the HD brain [2,3]. While the biological processes leading to neurodegeneration in HD are poorly understood,
transcriptional dysregulation has long been proposed as central to the pathogenesis of HD. Widespread alterations in gene expression have been reported [4] and several studies suggest that gene expression may be altered at one or more of the stages of RNA processing, translation, protein post-translational modification or trafficking [5,6].

MicroRNAs (miRNAs) are small non-coding RNAs that function as translational regulators of mRNA expression. miRNAs may inhibit gene expression either by repressing translation, or by targeting mRNA for either storage or degradation [7]. Recently, dysregulation of miRNAs has been linked to neurological and neurodegenerative disorders [8] and several studies have explored the role of miRNAs in HD. Marti et al [9] performed miRNA-sequencing for two pooled HD samples and two pooled control samples and reported altered expression for a large number of miRNAs. Altered expression of miRNAs, quantified using microarray technology, has been reported in cellular models of HD [10-12] and in mouse models of HD [12-15] but a comprehensive study of miRNA and mRNA expression obtained through next-generation sequencing technology in human HD samples has not been performed.

In order to investigate (1) the presence of altered miRNA expression and (2) the potential role of miRNAs on the altered mRNA expression seen in HD, we performed both miRNA-sequencing and mRNA sequencing, using Illumina massively parallel sequencing in twelve HD and nine neurologically normal control brains. To our knowledge this is the first genome-wide quantification of
miRNA expression comparing human HD and control brain, and the first to combine total miRNA expression with total mRNA expression obtained through massively parallel sequencing.

**Results**

**Selection of prefrontal cortex and BA9.** While the striatum is the region most heavily involved neuropathologically in HD [3], 80% to 90% of the neurons in that region will have degenerated by the time of death. These changes, together with the presence of reactive astrocytosis, alter the cellular composition of the striatum. In contrast, cortical involvement in HD is well defined [2,16] and while it does not experience dramatic neuronal degeneration, cortical neurons are known to exhibit the effects of protein aggregation and nuclear inclusion bodies characteristic of the disease. Therefore, we selected the prefrontal cortex for these studies.

**Table 1. HD brain samples analyzed for mRNA-seq, miRNA-seq and RT-qPCR validation of miR-10b-5p.**

All of the HD samples passed mRNA-seq QC. Scale of neuron loss: 0=absent, 1=mild, 2=moderate

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>miRNA-seq</th>
<th>RT-qPCR</th>
<th>PMI (hr.)</th>
<th>RIN or RQN</th>
<th>Death age</th>
<th>Onset age</th>
<th>Duration (yr.)</th>
<th>CAG repeat size</th>
<th>Neuron Loss in Neocortical Gray Matter</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD-01</td>
<td>Passed</td>
<td>Y</td>
<td>37</td>
<td>7.1</td>
<td>55</td>
<td>44</td>
<td>11</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>HD-02</td>
<td>Passed</td>
<td>Y</td>
<td>6</td>
<td>7.5</td>
<td>69</td>
<td>63</td>
<td>6</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>HD-03</td>
<td>Passed</td>
<td>Y</td>
<td>21</td>
<td>7</td>
<td>71</td>
<td>52</td>
<td>19</td>
<td>43</td>
<td>1</td>
</tr>
<tr>
<td>HD-05</td>
<td>Passed</td>
<td>Y</td>
<td>19</td>
<td>6.9</td>
<td>48</td>
<td>25</td>
<td>23</td>
<td>48</td>
<td>2</td>
</tr>
<tr>
<td>HD-06</td>
<td>Passed</td>
<td>Y</td>
<td>NA</td>
<td>6.2</td>
<td>40</td>
<td>34</td>
<td>6</td>
<td>51</td>
<td>1</td>
</tr>
<tr>
<td>HD-07</td>
<td>Passed</td>
<td>Y</td>
<td>8</td>
<td>8.5</td>
<td>72</td>
<td>55</td>
<td>17</td>
<td>41</td>
<td>1</td>
</tr>
<tr>
<td>HD-08</td>
<td>Passed</td>
<td>Y</td>
<td>21</td>
<td>7.4</td>
<td>43</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>49</td>
</tr>
<tr>
<td>HD-09</td>
<td>Passed</td>
<td>Y</td>
<td>4</td>
<td>7.8</td>
<td>68</td>
<td>45</td>
<td>23</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>HD-10</td>
<td>Passed</td>
<td>Y</td>
<td>6</td>
<td>8.3</td>
<td>59</td>
<td>35</td>
<td>24</td>
<td>46</td>
<td>1</td>
</tr>
<tr>
<td>HD-12</td>
<td>Passed</td>
<td>Y</td>
<td>13</td>
<td>6</td>
<td>68</td>
<td>52</td>
<td>16</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>HD-13</td>
<td>Passed</td>
<td>N</td>
<td>25</td>
<td>6.1</td>
<td>57</td>
<td>40</td>
<td>17</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>HD-14</td>
<td>Passed</td>
<td>Y</td>
<td>11</td>
<td>7.3</td>
<td>48</td>
<td>38</td>
<td>10</td>
<td>45</td>
<td>1</td>
</tr>
<tr>
<td>Mean</td>
<td>-</td>
<td>-</td>
<td>15.48</td>
<td>7.18</td>
<td>58.17</td>
<td>43.91</td>
<td>15.64</td>
<td>45.17</td>
<td>0.875</td>
</tr>
</tbody>
</table>
Table 2. Control brain samples analyzed for mRNA-seq, miRNA-seq and RT-qPCR validation of miR-10b-5p. All the control samples passed mRNA-seq QC. RIN = RNA Integrity Number, RQN = RNA Quality Number, PMI = Postmortem Interval

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>miRNA-seq</th>
<th>RT-qPCR</th>
<th>PMI (hr.)</th>
<th>RIN or RQN</th>
<th>Death age</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-14</td>
<td>Passed</td>
<td>Y</td>
<td>21</td>
<td>8</td>
<td>79</td>
</tr>
<tr>
<td>C-21</td>
<td>Passed</td>
<td>Y</td>
<td>26</td>
<td>7.3</td>
<td>76</td>
</tr>
<tr>
<td>C-29</td>
<td>Passed</td>
<td>Y</td>
<td>13</td>
<td>6.4</td>
<td>93</td>
</tr>
<tr>
<td>C-31</td>
<td>Passed</td>
<td>Y</td>
<td>24</td>
<td>7.3</td>
<td>53</td>
</tr>
<tr>
<td>C-32</td>
<td>Passed</td>
<td>Y</td>
<td>24</td>
<td>8.3</td>
<td>57</td>
</tr>
<tr>
<td>C-33</td>
<td>Passed</td>
<td>Y</td>
<td>15</td>
<td>7.5</td>
<td>43</td>
</tr>
<tr>
<td>C-35</td>
<td>Failed PCA</td>
<td>N</td>
<td>21</td>
<td>7.6</td>
<td>46</td>
</tr>
<tr>
<td>C-36</td>
<td>Passed</td>
<td>Y</td>
<td>17</td>
<td>7.5</td>
<td>40</td>
</tr>
<tr>
<td>C-37</td>
<td>Failed PCA</td>
<td>N</td>
<td>28</td>
<td>8.3</td>
<td>44</td>
</tr>
<tr>
<td>C-38</td>
<td>Passed</td>
<td>Y</td>
<td>20</td>
<td>7.7</td>
<td>57</td>
</tr>
<tr>
<td>C-39</td>
<td>Passed</td>
<td>Y</td>
<td>15</td>
<td>7.3</td>
<td>80</td>
</tr>
<tr>
<td>Mean</td>
<td>-</td>
<td>-</td>
<td>20.36</td>
<td>7.49</td>
<td>60.73</td>
</tr>
</tbody>
</table>

*Five miRNAs are up-regulated in HD.* After removing sample outliers using principal component analysis filtering, we identified five out of 1,417 detected mature miRNA species as differentially expressed between twelve HD and nine control prefrontal cortex samples using the R statistical package DESeq (Tables 1, 2 and 3; Figure 4). All five miRNAs were significantly up-regulated in HD. The largest effect between conditions was seen for miR-10b-5p, with a 28.41 fold increased expression in HD relative to control samples (mean control expression=915.81; mean HD expression=26,020.05, Figure 4). miR-1247-5p was expressed at moderate levels in both control (mean=49.44) and HD brain (mean=102.01). Three of the miRNAs, miR-196a-5p (mean control
expression=1.47; mean HD expression=27.49), miR-196b-5p (mean control expression=2.49; mean HD expression=11.01) and miR-615-3p (mean control expression=1.09, mean HD expression=6.66), had near zero expression levels in all nine control samples.

Table 3. Differentially expressed miRNAs from miRNA-seq
* FDR-adjusted q-value

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Control expression</th>
<th>HD expression</th>
<th>Fold Change</th>
<th>p-value</th>
<th>q-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-196a-5p</td>
<td>1.47</td>
<td>27.49</td>
<td>18.66</td>
<td>2.05E-10</td>
<td>2.91E-07</td>
</tr>
<tr>
<td>miR-10b-5p</td>
<td>915.81</td>
<td>26020.05</td>
<td>28.41</td>
<td>1.99E-08</td>
<td>1.41E-05</td>
</tr>
<tr>
<td>miR-615-3p</td>
<td>1.09</td>
<td>6.66</td>
<td>6.09</td>
<td>2.73E-05</td>
<td>1.29E-02</td>
</tr>
<tr>
<td>miR-1247-5p</td>
<td>49.44</td>
<td>102.01</td>
<td>2.06</td>
<td>7.67E-05</td>
<td>2.72E-02</td>
</tr>
<tr>
<td>miR-196b-5p</td>
<td>2.49</td>
<td>11.01</td>
<td>4.41</td>
<td>9.77E-05</td>
<td>2.77E-02</td>
</tr>
</tbody>
</table>

Figure 4. miR-196a-5p, miR-10b-5p and miR-615-3p were found significantly differentially expressed in Huntington’s disease
miR-10b-5p, miR-1247-5p, miR-196a-5p, miR-196b-5p, and miR-615-3p were identified as differentially expressed in Huntington’s disease prefrontal cortex compared to non-neurological disease controls by Illumina miRNA-sequencing. Normalized expression values quantified from DESeq analysis are shown on the y-axis. miR-196a-5p, miR-196b-5p and miR-615-3p were essentially not expressed in control samples, while the mean HD expression was 27.49, 11.01 and 6.66 respectively. miR-1247-5p was expressed at moderate levels in both control (mean=49.44) and HD brain (mean=102.01). miR-10b-5p was expressed in control (mean=915.81) and highly expressed in HD brain (mean=26,020.05). For miRNA, *p<0.05 and ***p<0.001, as determined by DESeq, followed by the Benjamini-Hochberg multiple comparison correction. (HD=Huntington’s disease).
**Validation and replication of miRNA findings.** miRNA expression differences were orthogonally validated using the Exiqon miRCURY LNA™ technology for reverse transcription quantitative PCR (RT-qPCR) in eleven of twelve sequenced HD samples and nine control samples originally studied for miRNA-seq. All five miRNAs were confirmed to be significantly up-regulated in HD, consistent with our miRNA-sequencing findings.

To replicate our findings in an independent sample set, we performed RT-qPCR in an additional eight control and eight HD prefrontal cortical samples. Four out of five miRNA (miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-615-3p) were confirmed as significantly increased in expression in HD.

**Similar proportion of neurons in HD and control cortical brain homogenate samples.** HD is characterized by progressive cortical atrophy, with
recognizable neuropathologic abnormalities in the neocortical gray matter [2,16-20] (Table 1). To address whether miRNA expression changes in HD may be due to altered ratios in brain cell-type abundance, such as a change in the ratio of neurons to glial cells, we compared the number of neuronal and non-neuronal nuclei across conditions. Suspensions of cell nuclei of prefrontal cortex from 28 HD cases and 19 controls were immunocytochemically labeled with anti-NeuN, a neuron-specific nuclear antigen, followed by flow cytometric analysis. The mean and range of NeuN+ ratios for controls and cases were not significantly different (t=1.67, p-value=0.10; Figure 5), suggesting cortical neuron loss in the BA9 area in HD is relatively modest and does not account for the dramatic alterations in miRNA levels reported here.

Figure 5. Neuron counts from prefrontal cortical tissue homogenate. No significant difference is observed when comparing ratios of NeuN+ counts to total events quantified by flow cytometry.
**Increased miR-10b-5p expression is not observed in Parkinson’s disease (PD).** To establish whether miR-10b-5p change is a generalized response to neurodegeneration, we evaluated this miRNA in PD prefrontal cortex. While cortical neuronal loss is variable in PD, both PD and HD are neurodegenerative and caused by protein inclusions. We selected PD prefrontal cortex samples that exhibited reported neuron loss on their neuropathological evaluation (n=6) and PD samples without reported cortical neuronal loss (n=8). From total RNA, RT-qPCR was performed for miR-10b-5p. No difference was seen in miR-10b-5p expression when stratifying PD based on the extent of neuron loss (t=0.59, p-value=0.58). Additionally, no significant difference in HD miR-10b-5p expression from qPCR was observed when stratifying HD cases based on a measure of cortical neuron loss (f=0.28, p-value=0.76).

Next, the relative expression of miR-10b-5p in PD was compared to all nineteen HD and eighteen control samples assayed. While no significant difference in miR-10b-5p expression was observed between control and PD samples (q=0.05, p=0.99), a significant difference was seen in HD compared to PD (q=7.30, p<0.0001; **Figure 6**), suggesting increased miR-10b-5p expression, independent of neuron loss, is not a generalized response to neurodegeneration.
Figure 6. miR-10b-5p expression in control, Parkinson's disease and Huntington's disease prefrontal cortex.

Up-regulation of miR-10b-5p was confirmed in HD by performing RT-qPCR, comparing nineteen Huntington’s disease prefrontal cortex samples to eighteen non-neurological disease control samples (***p<0.001) or fourteen Parkinson’s disease samples (***p<0.001). ∆∆C_{T} values of miR-10b-5p in PD and HD as compared to controls are shown on the y-axis. The absence of up-regulation in PD frontal cortex suggests that up-regulation of miR-10b-5p may be HD specific. (C_{T}= cycle threshold; RT-qPCR=reverse transcription quantitative PCR; PD=Parkinson’s disease; HD=Huntington’s disease)

Ectopic miR-10b-5p expression protects HD cell lines from polyglutamine-mediated cytotoxicity. To determine the functional importance of miR-10b-5p up-regulation in HD, we ectopically expressed miR-10b-5p in PC12 Q73 cells. These cell stably expressed huntingtin fragment derived from exon 1 (1-90), contain a pathogenic, 73 long polyglutamine repeat and a MYC epitope for protein identification. PC12 cells have been shown to terminally differentiate and form neural processes upon nerve growth factor (NGF) treatment [21], and HD models of these cells have been highly characterized,
exhibiting phenotypic changes such as aggregate formation and polyglutamine-dependent cell death [22-26].

PC12 Q73 cells were transfected with miR-10b-5p mimic or a negative control mimic, cel-miR-67-3p, after 48 hours post-differentiation. Cell survival was quantified using a MTT cell viability assay 48 hours post-transfection. Increased survival, though modest (53.9% versus 48.2%), was statistically higher for cells transfected with miR-10b-5p compared to cells transfected with negative control miRNA (q=4.58, p-value<0.0001; Figure 7). The enhanced survival via ectopic miR-10b-5p expression was further substantiated in experiments using viable fluorescent cell staining, where miR-10b-5p transfected cells showed increased cell viability over cells transfected with negative control miRNA (t=2.381, p-value=0.018).

Thus, miR-10b-5p may play a protective role in enhancing cell survival during stress. To model stress, we treated miRNA transfected cells with 1uM MG 132, a potent proteasome inhibitor that increases huntingtin aggregation and cellular apoptosis in PC12 HD cell lines [27]. As expected, MG 132 treated cells had reduced cell viability as compared to untreated cells (cel-miR-67-3p, q=6.52, adjusted p-value<0.0001; miR-10b-5p, q=10.88, adjusted p-value<0.0001). However, MG 132 treated miR-10b-5p transfected PC12 Q73 cells exhibited improved survival over those transfected with negative control miRNA (q=3.728, adjusted p-value=0.045). No statistical difference was observed when comparing miR-10b-5p levels with MG 132 treatment to cel-miR-67-3p without treatment,
(q=2.95, adjusted p-value=0.16), suggesting miR-10b-5p may enhance survival in times of cellular stress.

**Figure 7. miR-10b-5p overexpressing PC12 Q73 cells exhibit reduced cytotoxicity.**
PC12 cells expressing huntingtin exon 1 with a polyglutamine expansion spanning 73 repeats were transfected with miR-10b-5p or cel-miR-67-3p as a negative control. On day 3 post-differentiation, a subset of cells were treated with 1 uM MG 132. A MTT assay was used to measure cell viability after four days post differentiation. On the Y-axis, the viability percentage was calculated from the initial cell count. Error bars represent SEM. (****p<0.0001; **p<0.001 *p<0.05)

**miRNA expression is related to clinical variables in HD.** RNA sequence count data may be non-normally distributed [28], and tests of normality for miRNA expression levels in HD found that miR-10b-5p was negatively skewed (see Materials and Methods). Therefore, to test the relationship of miRNA expression to clinical variables such as CAG repeat size, age at onset of
motor symptoms, disease duration and age at death, as well as to the sample quality information for RIN/RQN (RNA integrity number/RNA quality number), we applied a step-wise backwards selection, negative binomial regression model.

Age at onset, duration and age at death are inter-dependent and could not be simultaneously included in the models. Furthermore, age at onset and age at death were strongly correlated with each other (Pearson $r=0.85$, p-value=$5e^{-04}$) and both were correlated with CAG repeat size ($r=-0.84$, p-value=$6e^{-04}$, and $r=-0.89$, p-value=$1e^{-04}$ respectively) while duration was not correlated with age at onset, age at death or CAG repeat size in this sample. To determine which variables best modeled the relationship of the miRNAs to clinical variables, we compared the Akaike information criterion (AIC) for each variable (onset age, death age and duration) in regression analyses that adjusted for the effect of CAG repeat size. Of these three variables, duration was found to have the poorest fit with each of the five miRNAs and therefore we report analyses containing age at onset and age at death.

Among the HD brains, CAG repeat size, age at onset and age at death were all independently found to have a negative association with miR-10b-5p (CAG, $\beta=-0.18$, p-value=$2.7e^{-05}$; onset, $\beta=-0.05$, p-value=$1.9e^{-05}$; death, $\beta=-0.07$, p-value=$6.8e^{-07}$). CAG repeat size and age at onset were found to be independently, negatively related to miR-196a-5p (CAG, $\beta=-0.15$, p-value=$1.7e^{-02}$; onset, $\beta=-0.07$, p-value=$1.4e^{-03}$). Age at death was significantly related to miR-615-3p expression ($\beta=-0.03$, p-value=$0.0045$) and age at onset was
associated with miR-196b-5p (β=-0.04, p-value=9e-04). No association to any clinical features was seen for miR-1247-5p. In order to fully evaluate whether there was any effect of disease duration on the observed relationships to the clinical features, duration was added back into final models. No substantial changes to the effect estimates were observed with the addition of duration to any of the models.

None of the miRNA levels was related to post-mortem interval in either control or HD case samples. The essentially null level of expression in controls prevented meaningful assessment of the relationship of miR-196a-5p, miR-196b-5p and miR-615-3p with clinical variables, in particular age at death, or sample variables, post-mortem interval (PMI), or RIN/RQN. Analysis of miR-10b-5p showed no association to age at death (β=-0.002, p-value=0.60), or PMI (β=-0.014, p-value=0.31), but did show association with RIN/RQN (β=0.54, p-value=7.2e-05) in controls. miR-1247-5p showed association with later age at death (β=-0.013, p-value=0.024) in controls.

*Expression of miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p are correlated.* Among the twelve HD samples, the levels of four out of the five significantly differentially expressed miRNAs (miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-615-3p) were strongly correlated with each other, (Spearman r range 0.71-0.88; p range 0.0002-0.01). miR-1247-5p was not significantly correlated with these miRNAs (Spearman r range 0.13-0.51; p range 0.09-0.70). Because the values of miR-615-3p and miR-196a-5p were essentially zero in the
control samples, correlations among the miRNAs were not performed for controls.

**mRNA targets of miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p may have similar functions.** Watson-Crick base-pairing between nucleotide position 2 through 8 on the mature miRNA, termed the ‘seed region,’ and the 3’ untranslated region (3’ UTR) of target mRNA determine the recognition, specificity and efficiency of miRNA silencing [29]. Seed sequences differ for miR-10b-5p (ACCCUGU), miR-615-3p (CCGAGCC) and miR-1247-5p (CCCGUCC) suggesting these miRNA have different targets, while miR-196a-5p and miR-196b-5p share a seed sequence (AGGUAGU) and only differ by a single base difference in mature miRNA sequence.

Targets of the five miRNAs were obtained from miRWalk (http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/index.html), a repository of experimentally validated miRNA targets curated from literature and online resources [30]. miRWalk targets of miR-196a, miR-196b and miR-1247 were not strand specific. The miRWalk database contained 84 unique targets for miR-10b-5p, 80 for miR-196a, 40 for miR-196b, two for miR-1247 and twelve for miR-615-3p. Since miR-1247 had just two validated targets, it was removed from analysis.

Four target genes (DICER1, HOXA7, HOXB4, HOXD1) were shared across all four miRNAs. miR-10b-5p shared eleven targets with miR-196a-5p (HOXB8, COX8A, HOXA10, NPC1, FLT3, AKT1, NPM1, DROSHA, AGO2, NFYC, PAX7), and one with miR-615-3p (MAPK8). miR-196a and miR-196b
shared 28 targets. In all, eleven of the 167 unique validated targets were Hox cluster genes (*HOXA1, HOXA7, HOXA9, HOXA10, HOXB4, HOXB7, HOXB8, HOXC8, HOXD1, HOXD4, HOXD10*).

To understand the influence these miRNAs may be having on shared biological processes, targets of each miRNA were analyzed using IPA Core Analysis. To find overlap in biological functions and canonical pathways of each miRNA and its targets, the IPA Core Comparison Analysis tool was used. After correcting for multiple comparisons, targets of miR-10b-5p, miR-196a, miR-196b and miR-615-3p shared significant overlap in 33 biological functions; the top three functional categories were “*Cell Death and Survival,*” (Benjamini-Hochberg adjusted p-value, range=3.5e-07 – 1.5e-04), “*Nervous System Development and Function*” (range=1.5e-07 – 1.5e-03) and “*Cellular Assembly and Organization*” (range=2.5e-05 – 1.7e-03). Twelve pathways were shared among all four sets of miRNA targets, including “*Huntington’s Disease Pathway*” (range=7.6e-04 – 8.1e-03), (Gene set=AKT1, BAX, CAPSN1, CLTC, CREB1, EGFR, HDAC9, JUN, MAPK8).
Table 4. 22 differentially expressed targets of miR-10b-5p, miR-196a, miR-196b, miR-1247 and miR-615-3p.
* FDR-adjusted q-value for 167 targets of the five miRNAs.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>miRNA Location</th>
<th>Mean Control Expression (n=9)</th>
<th>Mean HD Expression (n=12)</th>
<th>Fold Change</th>
<th>p-value</th>
<th>q-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>SERPINE1</td>
<td>miR-10b-5p 7q22.1</td>
<td>22.91</td>
<td>140.82</td>
<td>6.15</td>
<td>3.03E-11</td>
<td>5.06E-09</td>
</tr>
<tr>
<td>CDKN1A</td>
<td>miR-196a 6p21.2</td>
<td>336.73</td>
<td>841.75</td>
<td>2.5</td>
<td>1.58E-04</td>
<td>1.22E-02</td>
</tr>
<tr>
<td>HOXD4</td>
<td>miR-10b-5p 2q11.1</td>
<td>1.74</td>
<td>18.33</td>
<td>10.51</td>
<td>2.38E-04</td>
<td>1.22E-02</td>
</tr>
<tr>
<td>ANXA3</td>
<td>miR-10b-5p 4q21.21</td>
<td>259.93</td>
<td>553.71</td>
<td>2.13</td>
<td>2.92E-04</td>
<td>1.22E-02</td>
</tr>
<tr>
<td>TWIST1</td>
<td>miR-10b-5p 7p21.2</td>
<td>43.43</td>
<td>105.16</td>
<td>2.42</td>
<td>5.63E-04</td>
<td>1.72E-02</td>
</tr>
<tr>
<td>CD33</td>
<td>miR-196a, miR-196b 19q13.3</td>
<td>16.58</td>
<td>46.63</td>
<td>2.81</td>
<td>6.75E-04</td>
<td>1.72E-02</td>
</tr>
<tr>
<td>DIO3</td>
<td>miR-1247 14q32</td>
<td>10.89</td>
<td>41.93</td>
<td>3.85</td>
<td>7.29E-04</td>
<td>1.72E-02</td>
</tr>
<tr>
<td>MMP2</td>
<td>miR-10b-5p 16q13-q21</td>
<td>58.67</td>
<td>137.83</td>
<td>2.35</td>
<td>8.26E-04</td>
<td>1.72E-02</td>
</tr>
<tr>
<td>MMP9</td>
<td>miR-10b-5p 20q11.2-q13.1</td>
<td>5.32</td>
<td>17.33</td>
<td>3.26</td>
<td>9.33E-04</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>HOXA10</td>
<td>miR-196a, miR-196b 7p15.2</td>
<td>1.06</td>
<td>17.06</td>
<td>16.12</td>
<td>1.21E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>RHOD</td>
<td>miR-10b-5p 11q14.3</td>
<td>12.71</td>
<td>37.96</td>
<td>2.99</td>
<td>1.23E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>COL1A1</td>
<td>miR-196a 17q21.33</td>
<td>30.19</td>
<td>220.28</td>
<td>7.3</td>
<td>1.31E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>HLA-E</td>
<td>miR-10b-5p 6p21.3</td>
<td>3703.47</td>
<td>7769.76</td>
<td>2.1</td>
<td>1.34E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>PPARA</td>
<td>miR-10b-5p 22q13.31</td>
<td>444.7</td>
<td>865.02</td>
<td>1.95</td>
<td>1.53E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>PAX6</td>
<td>miR-196a 11p13</td>
<td>693.52</td>
<td>1337.23</td>
<td>1.93</td>
<td>1.62E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>EGFR</td>
<td>miR-10b-5p 7p12</td>
<td>784.95</td>
<td>1762.88</td>
<td>2.25</td>
<td>1.66E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>HOXB7</td>
<td>miR-196a 17q21.33</td>
<td>1.63</td>
<td>6.99</td>
<td>4.28</td>
<td>2.83E-03</td>
<td>2.78E-02</td>
</tr>
<tr>
<td>PLAUR</td>
<td>miR-10b-5p 19q13</td>
<td>56.15</td>
<td>119.67</td>
<td>2.13</td>
<td>3.65E-03</td>
<td>3.38E-02</td>
</tr>
<tr>
<td>HOXD10</td>
<td>miR-10b-5p 2q31.1</td>
<td>1.25</td>
<td>9.33</td>
<td>7.45</td>
<td>4.73E-03</td>
<td>3.96E-02</td>
</tr>
<tr>
<td>RUNX1</td>
<td>miR-10b-5p 21q22.3</td>
<td>87.69</td>
<td>224.88</td>
<td>2.56</td>
<td>4.74E-03</td>
<td>3.96E-02</td>
</tr>
<tr>
<td>SOX2</td>
<td>miR-10b-5p 3q26.3-q27</td>
<td>1963.76</td>
<td>3492.72</td>
<td>1.78</td>
<td>5.32E-03</td>
<td>4.23E-02</td>
</tr>
<tr>
<td>KRT5</td>
<td>miR-196a 12q13.13</td>
<td>113.74</td>
<td>51.99</td>
<td>-2.19</td>
<td>6.00E-03</td>
<td>4.55E-02</td>
</tr>
</tbody>
</table>

**mRNA targets of differentially expressed miRNAs are differentially expressed.** Total mRNA-sequencing was performed in the same brain samples as miRNA-sequencing to examine whether gene expression was affected by miRNA up-regulation. Of the 169 unique gene targets for the five differentially expressed miRNAs, 167 were detected using mRNA-sequencing. 22 mRNA targets were significantly differentially expressed between the HD and control prefrontal cortex samples (False Discovery Rate (FDR) adjusted q-value=0.05}
after adjusting for 167 comparisons). Only one gene (keratin 5, *KRT5*) was down-regulated in HD (*Table 4*), and four of these target genes were located in the Hox clusters (*HOXD4, HOXA10, HOXB7 and HOXD10*).

**Table 5. Differential expression of Hox cluster genes in HD.**
* FDR-adjusted q-value for the 55 genes in the four Hox clusters

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mean Control Expression (n=9)</th>
<th>Mean HD Expression (n=12)</th>
<th>Fold Change</th>
<th>p-value</th>
<th>q-value *</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOXA11</td>
<td>1.06</td>
<td>8.2</td>
<td>7.75</td>
<td>3.96E-05</td>
<td>2.07E-03</td>
</tr>
<tr>
<td>HOXA5</td>
<td>1.06</td>
<td>7.63</td>
<td>7.21</td>
<td>1.03E-04</td>
<td>2.07E-03</td>
</tr>
<tr>
<td>HOXD8</td>
<td>1.15</td>
<td>7.84</td>
<td>6.8</td>
<td>1.13E-04</td>
<td>2.07E-03</td>
</tr>
<tr>
<td>HOXD4</td>
<td>1.74</td>
<td>18.33</td>
<td>10.51</td>
<td>2.38E-04</td>
<td>3.22E-03</td>
</tr>
<tr>
<td>HOXB9</td>
<td>1.06</td>
<td>9.2</td>
<td>8.69</td>
<td>2.93E-04</td>
<td>3.22E-03</td>
</tr>
<tr>
<td>HOXA10</td>
<td>1.06</td>
<td>17.06</td>
<td>16.12</td>
<td>1.21E-03</td>
<td>1.11E-02</td>
</tr>
<tr>
<td>HOXC6</td>
<td>1.15</td>
<td>6.16</td>
<td>5.34</td>
<td>1.62E-03</td>
<td>1.27E-02</td>
</tr>
<tr>
<td>HOXA11-AS</td>
<td>1.25</td>
<td>7.39</td>
<td>5.9</td>
<td>2.49E-03</td>
<td>1.71E-02</td>
</tr>
<tr>
<td>HOXB7</td>
<td>1.63</td>
<td>6.99</td>
<td>4.28</td>
<td>2.83E-03</td>
<td>1.73E-02</td>
</tr>
<tr>
<td>HOXA13</td>
<td>1.45</td>
<td>8.74</td>
<td>6.03</td>
<td>4.07E-03</td>
<td>2.24E-02</td>
</tr>
<tr>
<td>HOXD10</td>
<td>1.25</td>
<td>9.33</td>
<td>7.45</td>
<td>4.73E-03</td>
<td>2.36E-02</td>
</tr>
<tr>
<td>HOXD1</td>
<td>55.91</td>
<td>22.8</td>
<td>-2.45</td>
<td>8.55E-03</td>
<td>3.92E-02</td>
</tr>
<tr>
<td>HOXC10</td>
<td>1.36</td>
<td>8.04</td>
<td>5.9</td>
<td>1.06E-02</td>
<td>4.14E-02</td>
</tr>
<tr>
<td>HOXC4</td>
<td>3.57</td>
<td>10.77</td>
<td>3.02</td>
<td>1.12E-02</td>
<td>4.14E-02</td>
</tr>
<tr>
<td>HOTAIRM1</td>
<td>3.52</td>
<td>12.52</td>
<td>3.56</td>
<td>1.13E-02</td>
<td>4.14E-02</td>
</tr>
</tbody>
</table>

*miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p expression is related to Hox cluster gene expression.* Four of the five up-regulated miRNAs are located intergenic to Hox gene clusters (*Figure 8*). Because of gene duplication, miR-196a is derived from both the HOXB and HOXC clusters; miR-10b is located in the HOXD cluster and miR-615 is found in the HOXC cluster.
A total of 55 genes (40 protein-coding genes, eleven antisense transcripts, three functional IncRNAs and one pseudogene) are located in the four Hox clusters [33,34]. To evaluate evidence for a general regional up-regulation of Hox cluster genes, an expression analysis of the mRNA-sequence data was performed for all annotated genes within the Hox loci (Table 5). Fifteen out of 55 genes within the Hox loci were differentially expressed in HD. Fourteen Hox genes were significantly up-regulated (FDR-adjust q-value<0.05, mean fold-change = 6.73, range 3.02 to 16.12) and a single Hox gene was down-regulated (HOXD1, FDR-adjust q-value=3.92e-02, fold change = -2.45). The majority of differentially expressed Hox genes (13 out of 15) were essentially unexpressed in controls.

The genes adjacent to the four differentially expressed miRNAs were highly expressed. Two genes immediately adjacent to miR-10b-5p were significantly up-regulated in HD (HOXD4, FDR-adjusted q=3.22e-03; HOXD8, FDR-adjusted q=2.07e-03), (Figure 8). HOXB9 (FDR-adjusted q-value=3.22e-03) immediately downstream of miR-196a-1 and HOXC10 (FDR-adjusted q-value=4.14e-02) immediately upstream of miR-196a-2 were also up-regulated. Furthermore, all three Hox genes located upstream of miR-196b were significantly up-regulated in HD (HOXA10, FDR-adjusted q-value=1.11e-02; HOXA11, FDR-adjusted q-value=2.07e-03; HOXA13, FDR-adjusted q-value=2.24e-02). HOXC6 (FDR-adjusted q-value=1.27e-02) immediately upstream of miR-615 was also up-regulated.
Figure 8. Differentially expressed miRNAs in HD are located in Hox gene clusters.
Schematic representation of Hox clusters. Hox genes are represented as numbered boxes (labeled 1-13), miRNA are represented by triangles and other genes in the regions (functional IncRNA, PRAC) are represented by rectangles. Antisense transcripts and pseudogenes are not pictured. Nineteen genes within Hox cluster regions were found significantly differentially expressed in HD prefrontal cortex using mRNA-sequencing (FDR-adjusted p-value<0.05). Four miRNAs, one IncRNA, and fourteen Hox genes were significantly up-regulated in HD (indicated by red), many of which are adjacent to differentially expressed miRNAs. A single Hox gene (HOXD1) was down-regulated in HD (indicated by blue) (HD=Huntington’s disease).

Discussion

Up-regulation of expression for five miRNAs in HD brain. We report a next-generation sequencing study of small RNAs, identifying 1,417 mature miRNA species in the prefrontal cortex (Brodmann Area 9) of twelve HD and nine control brains. Five of these, miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-615-3p and
miR-1247-5p, were up-regulated in HD at genome-wide significance (FDR q-value <0.05), and three of these five, miR-196a-5p, miR-196b-5p and miR-615-3p, were expressed at near zero levels in the control brains. Up-regulation of miR-10b-5p was validated in the miRNA-sequencing samples and confirmed in an independent replication sample set. Several studies implicating a role for miRNAs in HD have been performed, although, to our knowledge this is the first genome-wide quantification of miRNA expression comparing individual human HD and control brain samples.

Packer et al. [11], studying an array of 365 mature miRNAs, had previously reported miR-196a-5p to be significantly increased by nearly six-fold in Brodmann Area 4 of HD grade 1 brains. Recently, a study by Cheng et al. [13] found increased miR-196a expression suppressed mutant HTT expression in both HD neuronal cell models and HD transgenic mouse models. These findings suggest increased expression of miR-196a may be an adaptive response, promoting neuronal survival and may have therapeutic implications for HD. Miyazaki et al. [35] studied miR-196a in spinal and bulbar muscular atrophy (SBMA), a neurodegenerative disease caused by a similar polyglutamine repeat expansion in the androgen receptor (AR) gene. They found increased miR-196a expression via adeno-associated virus vector-mediated delivery reduced AR mRNA levels leading to improved neurological function in transgenic SBMA mouse models. Together, these findings suggest a neuroprotective role for miR-196a and its targets and possible therapeutic implications across multiple
polyglutamine-expansion neurodegenerative diseases. miR-196a-5p and miR-10b-5p were among the 56 miRNAs found to be elevated in response to mutant HTT over-expression in undifferentiated NT2 cells [36]. According to the miRNA search program “PubmiR,” [37] miR-196b-5p, miR-1247-5p and miR-615-3p have not been previously reported in HD miRNA studies.

A number of past studies have examined miRNA levels in HD, HD transgenic mice or cellular models; however, we did not replicate the results obtained in these studies. Gaughwin et al. [36] reported miR-34b elevated in plasma samples in HD, but we found neither miR-34b-3p nor miR-34b-5p to be altered in HD brain at genome-wide levels. We were not able to confirm any of the miRNAs reported in past microarray studies that examined targeted subsets of miRNAs, including the nine miRNAs reported as down-regulated in two mouse models of HD (YAC128 and R6/2) by Lee et al. [14] using a 567 miRNA microarray or the 38 miRNAs with altered expression in HD transgenic mice in a 382 miRNA microarray [15]. Johnson et al. [10-12] reported miR-29a and miR-330 to be significantly up-regulated in HD samples, neither of which was found to be altered in this study [10]. In a RT-qPCR study comparing 90 miRNAs in mouse Hdh (Q111/Q111) striatal cells to control mice [12,38], none of the 27 reported differentially expressed miRNAs was different at genome-wide levels in our study. The most commonly reported altered miRNA in HD studies, miR-132, has been reported as both down-regulated [10,14,39] and up-regulated [11], but was not differentially expressed in our study.
While some of the lack of concordance may be a consequence of the differences between human and animal models of HD, it is also likely that some of the differences are a consequence of the different technologies employed by these studies. Microarrays may have different levels of detection for some miRNAs from that seen by miRNA sequencing. Finally, nearly all of the studies employ microarray methods. Microarrays that study only 365 (e.g. Packer et al. [11],) to 567 miRNAs (e.g. Lee et al. [14]) are not performing as many contrasts and thus do not adjust for as many contrasts as our genome wide analysis (e.g. 1,417 miRNAs detected) demands.

**miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p implicate Hox cluster genes.** Four (miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p) of the five differentially expressed miRNAs are related to Hox cluster genes as follows: (1) these four are located in intergenic regions of the Hox clusters, (2) eleven Hox genes are validated targets of these four miRNAs, (3) Hox genes adjoining differentially expressed miRNAs are differentially expressed and (4) multiple Hox cluster genes are differentially expressed in HD versus control brains (Table 4).

Of the eleven Hox gene targets, eight did not differ in their expression across condition. A single target, *HOXD1* was seen to be down-regulated in HD (FC=-2.45). *HOXD1* is a reported target of four of the five miRNAs [40] which may explain its repression in HD.
Three Hox gene targets were up-regulated in HD (HOXB7, HOXD4, HOXD10). It is possible these up-regulated Hox genes share similar regulatory mechanisms, as the increased miRNA expression does not produce the expected miRNA-mediated gene silencing and suppress the observed up-regulation of the miRNA target genes. Coevolution of Hox genes and Hox-related miRNAs may further suggest that they share regulatory elements or mechanisms [41]. Furthermore, Hox genes and related miRNAs have been observed to have similar patterns of transcriptional activation and both are activated by retinoic acid [42-46]. Although miR-10b-5p has been validated as targeting HOXD4, they may exhibit patterns of co-expression. Specifically, Phua et al. [45] report miR-10b and HOXD4 are temporally co-expressed during neurodifferentiation. Here, we see a similar up-regulation and co-expression pattern in HD, where miR-10b and HOXD4 are both highly expressed.

Hox genes are a family of transcription factors that contribute to major morphological changes during embryonic development and are required for anterior-posterior body axis in bilaterally developing species [47]. They are highly involved in most aspects of early development, and are prominently expressed in the developing brain [48]. Hox-related miRNAs may also follow similar spatio-temporal patterns of expression during embryogenesis [49].

Hox genes are regulated by retinoic acid but also other factors, including basic fibroblast growth factor [50], steroid hormones [51,52] and polycomb repressive complex group [53]. Polycomb group (PcG) proteins assemble into
large silencing complexes and control histone-modifying activity. Hox genes are repressed by PcG complexes, specifically Polycomb Repressive Complex 2 (PRC2), which trimethylates histone H3 at lysine 27 (H3K27me3) [53].

Seong et al [54] observed knockout huntingtin mouse embryos lacked repression of *HOXB1*, *HOXB2*, and *HOXB9* and showed diminished global H3K27me3, while a knock-in expanded repeat mouse exhibited increased H3K27me3 signal, suggesting mutant huntingtin may alter proper PRC2 activity. These findings raise the possibility that the increased expression of miRNAs and Hox genes reported here are related to enhanced H3K27me3 or impaired PcG repression.

However, the role of Hox in the adult, HD brain is still unclear. Increased transcriptional activity of Hox may be compensatory, helping to preserve or re-establish cell polarity, or an indirect result of impaired epigenetic regulation.

**miR-10b-5p response in HD may be protective.** To functionally validate our miRNA-sequencing findings, we chose to assess miR-10b-5p. We believed this miRNA to be the most biologically active of the differentially expressed miRNAs. miR-10b-5p had the highest basal expression levels and the highest fold change between conditions. Additionally, miR-10b-5p levels were not increased in PD, a comparable protein aggregate, neurodegenerative disease, nor in PD samples with pathology in the prefrontal cortex equivalent to HD.

To determine whether miR-10b-5p had a protective or deleterious effect on neuron viability, we ectopically expressed miR-10b-5p in terminally differentiated PC12 Q73 cells. Since the levels the five differentially expressed
miRNA were up-regulated, we felt overexpression of miR-10b-5p best represented the phenotype observed in HD brain.

We reported increased miR-10b-5p expression enhanced the survival of PC12 Q73 cells. Furthermore, we found that increased miR-10b-5p expression enhanced survival in the presence of apoptosis-inducing compound, MG 132. In this experiment, survival in cells with increased miR-10b-5p expression was comparable to that of unchallenged cells and significantly greater than untreated cells exposed to toxin. These finding provide support for the hypothesis that increased miR-10b-5p may be a neuroprotective response to the expanded polyglutamine repeat seen in HD and speaks to the role of this microRNA in the pathology of HD.

**miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p have overlapping biological functions.** pathway analysis, we showed that miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-3p targeted genes are predicted to be involved in apoptosis as well as nervous system development and function. In neuroblastoma SH-SY5Y cell lines, miR-10a, miR-10b and miR-615-5p expression levels significantly increased during all-trans-retinoic-acid (ATRA) treatment, indicating miR-10a/b and miR-615-5p may have a role in neurodifferentiation [44]. SH-SY5Y cells treated with antisense miR-10a or miR-10b had impaired neurite outgrowth and morphology but did not show changes in overall cell proliferation [44]. miR-10a and miR-10b were highly expressed in SK-N-BE, LAN5 and SH-SY5Y cell lines during ATRA treatment and ectopic
expression of miR-10ab mirrored the phenotype of the ATRA treatment [42]. Taken together, these studies implicate these miRNAs in neuron differentiation, migration, and outgrowth.

In our past studies [16], we found increased neurite outgrowth in HD prefrontal cortex. Relative to controls, HD pyramidal neurons had a significantly increased number of primary dendritic segments, increased total dendritic length, and more dendritic branches than control neurons. Here, we report four miRNAs that have been observed in cell models to present a similar phenotype. It is possible that increased expression of these miRNAs and related targets represent an adaptive response of neurons stressed by a toxic expanded polyglutamine protein fragment.

*miR-10b-5p, miR-196a-5p, miR-196b-5p and miR-615-5p are related to HD pathogenesis.* Four of the five up-regulated miRNAs showed association to clinical features of HD (CAG repeat size, age of motor onset and age at death for miR-10b-5p; CAG repeat size and age at onset for miR-196a-5p, age at onset for miR-196b-5p and age at death for miR-615-3p). Due to the near zero level of expression in controls, it was not possible to assess the relationship of miR-196a-5p, mir-196b-5p and miR-615-3p to age at death, but miR-10b-5p was not correlated with age at death in controls. Thus, the increased expression of these miRNAs did not appear to be related to normal aging, but rather a component of gene regulation and transcription in the context of neurodegeneration. A growing body of literature points to the presence of toxic effects of the HD gene.
substantially before the onset of symptoms, perhaps from the time of conception [55-57].

Because age at death represents the lifetime exposure of the individual to the effects of the HD gene, we hypothesize that the association of miR-10b-5p and miR-615-3p with age at death may represent the lifetime exposure to the effects of the HD mutation. If the relationship of altered miRNA expression to age at death supports the view that the HD gene may have a life-long effect among expanded CAG-repeat carriers, this raises the possibility that the HD mutation may influence neuronal development in the developing brain through the action of one or more of these miRNAs and Hox cluster genes.

**Target genes of over-expressed miRNAs show increased expression in HD.** We report five miRNAs as being highly up-regulated in HD and though our expectation was to see the mRNA targets of these miRNAs as decreased, we observe increased expression of many of their shared mRNA targets. We believe these effects are not attributable to differences in cell populations studied, since flow cytometric analysis measuring neuron abundance found no significant difference across condition. Rather, we hypothesize positive miRNA-mRNA target relationships are a result of HD-specific alterations in mRNA processing. Translation is a highly dynamic process. Cytoplasmic mRNA actively engaged in translation can cycle to a non-translated state and accumulate in stress granules or processing bodies (P-bodies). During cellular stress, mRNA can be sequestered to P-bodies or stress granules, to stall translation through
translational repression machinery or miRNA silencing, until stress conditions have been resolved [7,58-60]. P-bodies may also serve an important role in RNA transport. Because neurons are highly polarized, cytoplasmic transport of mRNA is essential for localized translation to discrete regions of the cell. During transport, it is believed that mRNAs are silenced by miRNA, upon rapid exchange at the synapse [60-62].

In HD cortical neurons, excitotoxicity, oxidative damage, aberrant gene expression and energetic defects lead to stress conditions and in response, cells may sequester mRNA to P-bodies and stress granules. Among the 55 Hox locus genes studied, only one of the fifteen significantly differentially expressed genes is down regulated (Table 4). Thus, the increased levels of most of the validated gene targets of these four miRNAs may be reactionary, as they are sequestered to P-bodies for storage as part of a protective process to enhance cell viability [7].

To the best of our knowledge, no study has addressed the role of P-bodies or stress granules in HD. However, it was observed in live cortical neurons that wildtype huntingtin co-localized in P-bodies, specifically in neuronal RNA granules, along with Argonaute 2, the endonuclease required for RNA-mediated gene silencing by the RNA-induced silencing complex (RISC) [63,64]. Therefore, it is reasonable to suggest mutant huntingtin may impair miRNA-mediated mRNA degradation and/or localized translation of specific mRNAs.
There is evidence that miRNA-mRNA regulatory mechanisms may be altered in other neurodegenerative diseases as well. In a joint examination of miRNA-mRNA expression in Alzheimer’s disease (AD) and control prefrontal cortex, an overwhelming number of miRNA to mRNA targets were found to be positive correlated. Genomic variants in TDP-43 and FUS, genes that encode stress granule proteins, were found to cause familial Amyotrophic lateral sclerosis [65,66] and several other stress granule proteins (TIA-1, G3BP) may also be pathogenic [67].

**miRNAs as potential biomarkers in HD.** These studies suggest potential relationships of these miRNAs to CAG repeat expansion, age at onset or age at death. If these findings hold up on further examination, these miRNAs may hold potential to provide insight into important biological and disease expression for HD. miRNA are extremely stable. The half-life of the majority miRNAs has been predicted to be on average five days and plasma miRNAs have been found to be stable after being subjected to high heat, extreme pH, long-time storage at room temperature, or multiple freeze-thaw cycles [68-70]. If these miRNAs cross the blood-brain barrier and can be detected at reasonable levels in serum/plasma from mutant HD gene carriers, they may serve as biomarkers of disease expression.

**Materials and Methods**

**Sample Information.** Frozen brain tissue from prefrontal cortex Brodmann Area 9 (BA9) was obtained from the Harvard Brain and Tissue
Twelve Huntington’s disease (HD) samples and eleven neurologically-normal control samples were selected for the study (Table 1). The HD subjects had no evidence of Alzheimer or Parkinson disease (PD) comorbidity based on neuropathology reports. For microscopic examination, 16 tissue blocks were systematically taken and histologically assessed as previously described [3]. All samples were male. HD samples and controls were not different for postmortem interval (PMI) (t=1.07, p=0.30), RNA integrity number (RIN) (t=0.83, p=0.41) or death age (t=0.40, p=0.69). CAG repeat size was known for all HD samples and onset age and disease duration was unknown for a single sample (Table 1). Eight additional HD, nine control and fourteen PD cases were studied as part of validation and replication studies, and were obtained from the HBTRC and the Sun Health Research Institute Sun City, Arizona.

**RNA extraction.** Total RNA, for all samples studied, was isolated using QIAzol Lysis Reagent and purified using miRNeasy MinElute Cleanup columns (Qiagen Sciences Inc., Germantown, MD). RNA quality for sequencing was assessed using either Agilent’s BioAnalyzer 2100 system and RNA 6000 Nano Kits to find RNA Integrity Number (RIN) or Agilent 2200 TapeStation and DNA ScreenTape assay RNA Quality Number (RQN; Agilent, Foster City, CA). Both methods calculate the area under the peak for 18S and 28S RNA as a ratio of total RNA as well as the relative height of the 18S and 28S peaks to determine
RNA quality [71]. The RIN/RQN values were similar for the twelve HD and eleven control specimens studied for miRNA and mRNA (t=0.95, p=0.36).

**Illumina miRNA sequencing (miRNA-seq).** For each brain sample, 1 ug of RNA was used to construct sequencing libraries using Illumina’s TruSeq Small RNA Sample Prep Kit, according to the manufacturer’s protocol (Illumina, San Diego, CA). In brief, small RNA molecules were adapter-ligated, reverse transcribed, PCR amplified and gel purified to generate the library. Multiplexed samples were equimolarly pooled into sets of eight samples per flowcell lane and sequenced using 1x50 bp single-end reads on Illumina’s HiSeq 2000 system at Tufts University sequencing core facility (http://tucf-genomics.tufts.edu/). Demultiplexing and FASTQ file generation (raw sequence read plus quality information in Phred format) were done using Illumina’s Consensus Assessment of Sequence and Variation (CASAVA) pipeline.

**Primary processing of Illumina miRNA-seq reads.** Sequence read quality was evaluated using the FASTQ quality filter module from the FASTX-toolkit version 0.0.13 (http://hannonlab.cshl.edu/fastx_toolkit/), and only those reads with at least 80% of the base calls above Q20 (Phred score) were retained. The 3’ adapter sequence (5’-TGGAATTCTCGGGTGCCAAGG-3’) was removed from all reads using the FASTA/Q clipper module from the FASTX-toolkit. A minimum length threshold of 15 nucleotides was set for clipped reads because miRNAs of this length will contain the seed sequence. To avoid redundancy amongst identical read species, the reads were collapsed using the FASTA/Q
collapser module from FASTX-toolkit to generate a FASTA file of only the unique read species.

**Alignment and mapping of miRNA-seq reads.** Quality-filtered, 3’ adapter-clipped reads were aligned to the UCSC human reference genome (build hg19) using Bowtie version 0.12.3 [72]. Alignment parameters were set to allow for no mismatch alignments and no limits on multiple mapping instances. Multiple-mapped identical sequences were summed for a single count for that annotated mature miRNA. The default settings were used for all other alignment options.

The miRNA aligned data are available on ArrayExpress as follows:

**Experiment name:** RNA-seq of micro RNAs (miRNAs) in Human prefrontal cortex to identify differentially expressed miRNAs between Huntington’s Disease and control brain samples. **ArrayExpress accession:** E-MTAB-2206, **Release date:** 2014-01-30 [https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-2206/](https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-2206/)

**miRNA abundance estimation.** Aligned reads that overlapped with the human miRNA annotation version 19 from miRBase (http://www.mirbase.org/ftp.shtml) were identified using default BEDTools’ `IntersectBed` functionality [73]. To select for mature miRNA reads, sequences more than 27 bases in length were removed. Only those reads for which the aligned 5’ start-nucleotide matched exactly to the 5’ start-nucleotide of the annotated miRNA were retained for the analysis. After filtering, collapsed read counts were summed per annotated mature miRNA.
**miRNA differential expression.** The R [http://www.R-project.org](http://www.R-project.org) package DESeq version 1.10.1 [28] was used to perform the differential expression analysis between HD and control samples using the read counts generated for each sample as described above. miRNAs with zero read counts across all case and control samples were removed from analysis. To accommodate the analysis of miRNAs with read counts of zero for some samples, a pseudo-count of one was added to all raw counts for every miRNA across all the samples, prior to performing DESeq’s estimateSizeFactors and estimateDispersions functions with default options. DESeq assumes that count data follow a negative binominal distribution and factors in technical and biological variance when testing for differential gene expression between groups. DESeq’s function, estimateSizeFactors, was used to obtain normalization factors for each sample and to normalize miRNA read counts.

The normalized counts were evaluated by principal component analysis (PCA) with the FactoMineR R package for all HD and control samples. The samples identified to be three or more standard deviations away from the mean on the first or second principal component were considered outliers and were removed from analysis. The first two principal components were used because they each explained more than 10% of the variance, while the remaining principal components explained less than 10% of the variance. Two control samples (C-35 and C-37) were identified as outliers based on PCA analysis.
miRNA differential expression analysis was performed with DESeq's nbinomTest function for the remaining nine control and twelve HD samples. All analyses were performed on DESeq normalized counts.

**miRNA quantitative PCR.** miRNA were assayed using Exiqon’s miRCURY LNA Universal RT miRNA PCR following the manufacturer’s protocol (Exiqon Inc, Denmark). In brief, reactions were incubated for 60 min at 42°C followed by heat-inactivation of reverse transcription for 5 min at 95°C and stored at 4°C. After cDNA synthesis, samples were diluted to 0.2 ng/ul in water. Brain samples were assayed using Exiqon ExiLENT SYBR Green master mix and LNA primer sets containing UniRT and miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-615-3p or miR-1247. Reference primer hsa-SNORD48 PCR/UniRT was used for brain samples; U6 snRNA for cell lines. Samples were run in triplicate for each primer set in 384-well format (5 ul PCR Master mix, 1ul PCR primer mix, 4 ul 0.2 ng cDNA). Reactions were cycled using Applied Biosystems 7900HT Fast Real-Time PCR System using manufacturer’s instructions (Life Technologies, Carlsbad, CA). For analysis, threshold cycle (C_T) was generated by ABI SDS v2.4 software. C_T values for triplicate wells were normalized by average RNU48 value for brain or U6 for cells. miRNA fold change was calculated using the 2^{-\Delta\Delta C_T} method [74].

**Neuron abundance quantification.** 0.5-1.0 g of tissue in 5 ml of lysis buffer was homogenized using a dounce tissue grinder. Lysates were transferred to ultracentrifugation tubes, loaded on top of sucrose solution and centrifuged at
24,400 RPM for 2.5 hr at 4°C (Beckman Coulter, Pasadena, CA; L8-70 M with SW80 rotor). Nuclei pellets were resuspended in 500 ul PBS and incubated at 4°C in a staining solution containing 0.72% normal goat serum, 0.036% BSA, 1:1200 anti-NeuN (Millipore, Germany), 1:1400 Alexa488 goat anti-mouse secondary antibody (Life Technologies, Carlsbad, CA), for 45 min. Flow cytometry was performed at the Boston University Medical School Flow Cytometry Core Lab on a FACS Vantage SE flow cytometer.

**Illumina messenger RNA sequencing (mRNA-seq).** For each brain sample, 1 ug of RNA was used to construct sequencing libraries using Illumina’s TruSeq RNA Sample Prep Kit according to the manufacturer’s protocol. In brief, mRNA molecules were polyA selected, chemically fragmented, randomly primed with hexamers, synthesized into cDNA, 3’ end-repaired and adenylated, sequencing adapter ligated and PCR amplified. Each adapter-ligated library contained one of twelve TruSeq molecular barcodes. Multiplexed samples were equimolarly pooled into sets of three samples per flowcell lane and sequenced using 2x100 bp paired-end reads on Illumina’s HiSeq 2000 system at Tufts University sequencing core facility (http://tucf-genomics.tufts.edu/). Demultiplexing and FASTQ file generation were accomplished using Illumina’s CASAVA pipeline.

**Primary processing of Illumina mRNA-seq reads.** Forward and reverse sequencing reads were independently quality-filtered using the FASTQ quality filter module from the FASTX-toolkit version 0.0.13
with the same criteria as that applied for the processing of the miRNA-seq reads. Reads failing the quality threshold, as well as their corresponding mate reads, were removed.

**Alignment and mapping of mRNA-seq reads.** Quality-filtered paired-end reads were aligned to the UCSC human reference genome (build hg19) using TopHat version 2.0.4 [75,76]. This version of TopHat incorporates the Bowtie version 2.0.0.7 algorithm to perform the alignment [72] as well as SAMtools version 0.1.18.0 for alignment file formatting [77]. For efficient read mapping, TopHat requires the designation of the mean and standard deviation of the distance between paired-end reads, the read inner-distance. To estimate the appropriate read inner-distance, we aligned a subset of 5 million reads from four HD and four control samples to the Ensembl human reference transcriptome (release 66) using Bowtie version 2.0.0.7. Using the CollectInsertSizeMetrics function from picardTools version 1.76 [http://sourceforge.net/projects/picard/files/picard-tools/], we estimated the average mean inner-distance per condition and subsequently applied these values for the TopHat alignment; 22 for HD samples 25 for controls respectively, (the current TopHat default setting is 20). To account for read variability, the standard deviation for inner-distance was set to 100. The number of allowed splice mismatches was set to 1. Default settings were used for all other alignment options.
**mRNA gene abundance estimation.** Gene expression quantification was performed using htseq-count version 0.5.3p9 (http://www-huber.embl.de/users/anders/HTSeq) and the GENCODE version 14 annotation gtf file as reference (http://www.gencodegenes.org/releases). Intersection non-empty mode and unstranded library type were specified as parameters for htseq-count. Default settings were used for all other options.

**mRNA differential expression analysis.** The mRNA differential expression analysis between HD and control samples was performed using DESeq version 1.10.1 [28]; the workflow was the same as described for the miRNA differential expression analysis. No outliers were found based on the PCA of the DESeq-normalized count data. The nbinomTest function was run for eleven control samples and twelve HD samples to assess differentially expressed genes. Multiple comparison adjustment for multiple testing with the Benjamini-Hochberg correction was used to control for false discovery rate. For Hox gene differential expression analysis, 55 comparisons were used. Genes located within HOX-gene containing regions were queried through the Ensembl database (release 72), interfacing through the R package BiomaRt [78,79]. Genes that were between HOXA1-HOXA13, HOXB1-HOXB13, HOXC4-HOXC13 and HOXD1-HOXD13 start sites were regarded as “Hox genes.” For miRNA target differential expression, 154 comparisons were used for Benjamini-Hochberg correction.
**miRNA-mRNA target analysis.** Information on experimentally validated miRNA targets of miR-10b-5p, miR-196a-5p and miR-615-3p were extracted from the miRWalk “Validated Targets” module [30]. Strand specificity was preserved. Targets for miR-196a-1 and miR-196a-2 were merged for analysis. IPA Core Analysis (analysis.ingenuity.com) was run as nervous system and CNS cell line specific across all species, using target gene lists imported from miRWalk output. “Bio Functions” and “Canonical Pathway” analyses were used. Right-tailed Fisher’s Exact Tests were run through IPA software and p-values with FDR-adjusted q-values (p<0.05) were considered significant. Biological functions across the 3 significant miRNA were compared using the IPA Core Comparison Analysis tool. Benjamini-Hochberg Multiple Testing Correction p-values (p<0.05) were considered significant.

**Linear modeling of miRNA relationship to clinical covariates.** To account for the non-normality in the miRNA data, negative binominal general linear regressions were performed using Proc genmod in SAS. DESeq normalized counts were rounded to the nearest integer before running the model. To test the normality of gene expression data, Shapiro-Wilk tests were performed. Differentially expressed miRNA data trended as non-normally distributed in HD (miR-10b-5p, p=0.04; miR-196a-5p, p=0.05; miR-615-3p, p=0.06), but not in controls (miR-10b-5p, p=0.71; miR-196a-5p and miR-615-3p were essential zero).
**Generation of transgenic cell lines.** PC12 (rat adrenal gland phaeochromocytoma) cells were grown at 37°C and 5% CO₂ in Dulbecco’s modified Eagle’s medium (DMEM; Life Technologies, Carlsbad, CA) with 20% fetal bovine serum (FBS; Atlanta Biologicals, Flowery Branch, GA), 100 units/ml penicillin and 100 units/ml streptomycin (Life Technologies, Carlsbad, CA). pcDNA3.1mycC expressing human huntingtin fragment (1-90) containing 73 polyglutamine repeats (Coriell Institute; CHDI-90000034) was used for stable transfection. Cells were seeded to 70% confluency and grown overnight. 15 µl of Attractene Transfection Reagent (Qiagen, Gaithersburg, MD) was added to 4 µg plasmid DNA diluted in 300 µl Opti-MEM (Life Technologies, Carlsbad, CA). Cells were grown in complete media and selected for four weeks using 500 mg/ml G418 (Life Technologies, Carlsbad, CA). To create monoclonal cultures, single colonies were isolated using dilution cloning, picked with filter paper, grown in a 6-well plate and screened for transgenic expression by Western blot analysis using mouse Anti- c-Myc (Novex, R950-25, Life Technologies, Carlsbad, CA).

**Cell differentiation and miRNA overexpression.** 96-well culture plates were seeded with 10,000 cells per well. For differentiation, culture medium was replaced with medium composed of DMEM with 0.5% FBS, 100 mg/ml G418, 100 units/ml penicillin and 100 units/ml streptomycin and 100 ng/ml nerve growth factor (R&D Systems, Minneapolis, MN). After 48 hr, miRNA was transfected into HD cells using 0.25 ul Lipofectamine 2000 (Life Technologies, Carlsbad, CA) and 6.25 pmol miR-10b-5p or miRIDIAN microRNA Mimic Negative Control #1 (cel-
miR-67-3p, Thermo Scientific, Waltham, MA) per well, following manufacturer's protocol. miR-10b-5p overexpression was verified using qPCR.

**Cell viability assays.** For MTT assays, 1 uM MG 132 (Tocris Bioscience, United Kingdom) was added to select wells containing 10,000 cells per well at 72 hr post-differentiation. Cell viability was assessed at 96 hr post-differentiation. Following manufacturer’s protocol, CellTiter 96 Non-Radioactive Cell Proliferation Assay kit (Promega; Madison, WI) was used to determine cell number. Cells were incubated for 1.5 hr at 37°C and 5% CO₂ with MTT dye solution. Undifferentiated HD cells were serially diluted across a 96-well plate to create a standard curve for cell number calculation. Absorbance was measured using Bio-Tek Synergy H1 spectrophotometer at 540 nm for miR-10b-5p transfected wells, with MG 132 (n=44) and without MG 132 (n=35) and cel-miR-67-3p transfected wells with MG 132 (n=40) and without MG 132 (n=40). One-way ANOVA way used for statistical analysis. For cell viability staining, miR-10b-5p and negative control mimic were transfected after 48 hours of differentiation in 12-well culture plate with 4 replicates each, 250,000 cells per well. Molecular Probes Neurite Outgrowth Staining Kit (Life Technologies, Carlsbad, CA) was used according to manufacturer's protocol. Using Bio-Tek Synergy H1 microplate reader, fluorescent area scans were taken at 530 nm excitation/590 nm emission with a 5x5 matrix per well.
References


CHAPTER 3: MIR-10B-5P EXPRESSION IN HUNTINGTON’S DISEASE BRAIN RELATES TO AGE OF ONSET AND THE EXTENT OF STRIATAL INVOLVEMENT


Abstract

**Background.** MicroRNAs (miRNAs) are small non-coding RNAs that recognize sites of complementarity of target messenger RNAs, resulting in transcriptional regulation and translational repression of target genes. In Huntington’s disease (HD), a neurodegenerative disease caused by a trinucleotide repeat expansion, miRNA dysregulation has been reported, which may impact gene expression and modify the progression and severity of HD.

**Methods.** We performed next-generation miRNA sequence analysis in prefrontal cortex (Brodmann Area 9) from 26 HD, 2 HD gene positive, and 36 control brains. Neuropathological information was available for all HD brains, including age at disease onset, CAG-repeat size, Vonsattel grade, and Hadzi-Vonsattel striatal and cortical scores, a continuous measure of the extent of neurodegeneration. Linear models were performed to examine the relationship of miRNA expression to these clinical features, and messenger RNA targets of associated miRNAs were tested for gene ontology term enrichment.
**Results.** We identified 75 miRNAs differentially expressed in HD brain (FDR q-value <0.05). Among the HD brains, nine miRNAs were significantly associated with Vonsattel grade of neuropathological involvement and three of these, miR-10b-5p, miR-10b-3p, and miR-302a-3p, significantly related to the Hadzi-Vonsattel striatal score (a continuous measure of striatal involvement) after adjustment for CAG length. Five miRNAs (miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-10b-3p, and miR-106a-5p) were identified as having a significant relationship to CAG length-adjusted age of onset including miR-10b-5p, the mostly strongly over-expressed miRNA in HD cases. Although prefrontal cortex was the source of tissue profiled in these studies, the relationship of miR-10b-5p expression to striatal involvement in the disease was independent of cortical involvement. Correlation of miRNAs to the clinical features clustered by direction of effect and the gene targets of the observed miRNAs showed association to processes relating to nervous system development and transcriptional regulation.

**Conclusions.** These results demonstrate that miRNA expression in cortical BA9 provides insight into striatal involvement and support a role for these miRNAs, particularly miR-10b-5p, in HD pathogenicity. The miRNAs identified in our studies of postmortem brain tissue may be detectable in peripheral fluids and thus warrant consideration as accessible biomarkers for disease stage, rate of progression, and other important clinical characteristics of HD.
**Background**

Huntington’s disease (HD) is an inherited disorder caused by a CAG trinucleotide repeat expansion in **HTT** which leads to progressive motor and cognitive impairment due to the gradual loss of neurons within striatal and cortical brain regions [1]. Although monogenic, HD displays remarkable variation in clinical expression, most readily observed by the range in age at clinical onset as determined by the manifestation of motor symptoms, varying from age 4 years to age 80 [2]. While onset age is unequivocally related to the size of the expanded CAG repeat, with longer repeats leading to earlier onset, only 50% to 70% of the variation can be attributed to repeat size [3,4]. The remaining variation is highly heritable (**h^2 = 0.56**), suggesting a strong role for genes that modify disease progression [3].

MicroRNAs (miRNAs) are small non-coding RNAs that negatively regulate the expression of genes in a sequence-specific manner, binding to the 3′-untranslated region (3′UTR) to initiate cleavage or translational repression of target transcripts [5,6]. miRNAs influence a diverse range of cellular processes [7] and consequently, their altered expression may lead to or influence disease-related pathological phenotypes, or reveal unknown aspects of the disease process. In the central nervous system (CNS), miRNAs are abundant, as brain-
specific miRNAs assist in various neuronal processes such as synaptic
development, maturation and plasticity [8,9]. Altered miRNA expression has been
observed in diseases of the CNS, particularly in age-dependent
neurodegenerative diseases, which suggests that the expression of miRNAs may
contribute to neuropathogenesis [10,11].

In HD, the dysregulation of miRNAs has been reported in HD in vitro
models, transgenic HD animals and human HD brain [12-24]. We hypothesize
that post-transcriptional regulation by miRNAs plays a role in modifying the
progression and severity of HD. Recently, we completed a study of miRNA
expression obtained through next-generation sequencing technology in human
HD and control brain samples to investigate the presence of altered miRNA
expression in HD and its role in transcriptional dysregulation [13]. The original
study provided sample power to detect large miRNA changes, but the sample
size was not sufficient to detect more subtle changes in miRNA expression and
did not represent a wide enough range of HD pathology to investigate
relationships to clinical features of HD. Therefore, to follow-up on these findings,
we have sequenced small RNAs in an additional 16 HD brains, two of which are
gene positive asymptomatic Vonsattel grade 0 cases, and 27 control samples, for
a combined study of 28 HD and 36 control samples. The increased sample size
enables the detection of significantly altered miRNAs with lower levels of
differential expression as well as more comprehensive characterization of the
relationship of these miRNAs to relevant clinical features of the disease,
including the age at motor onset of the disease, disease duration (the time between onset and death), age at death and extent of pathological involvement in the striatum and cerebral cortex. A deeper understanding of the global miRNA expression in HD may elucidate pathogenic mechanisms of disease progression in HD and suggest new therapeutic targets.

Results

Differential expression analysis highlights disrupted miRNA expression in HD brain. To evaluate the relationship of miRNA expression to salient clinical and pathological features of HD, we profiled miRNA expression using small RNA-sequencing of prefrontal cortex (Brodmann Area 9) of 26 symptomatic HD and 36 control samples (see Table 6). Although the striatum is the most affected brain region in HD, differences in miRNA expression between HD and unaffected controls, independent of cellular composition, would be difficult to assess due to the extent of neuron loss and the increase of reactive astrocytosis in HD striatal tissue [25]. Therefore, prefrontal cortex, which exhibits hallmark characteristics of HD pathology [26], relates to striatal involvement (Pearson r = 0.44, p < 2e-16) [27], but experiences less extreme changes than the striatum [28,29], was used for sequencing. In addition, previous studies have found no difference in cell counts between HD and controls from similar BA9 brain samples [13,30].

The HD samples consisted of Grade 2 (n = 4), Grade 3 (n = 15), and Grade 4 (n = 7) brains as determined by Vonsattel grade, an assessment of
striatal involvement classified as 0 through 4 in order of the severity of neuropathological involvement [25]. Sequenced samples were also among the 523 HD brains characterized by the recently established measure of pathological involvement termed the Hadzi-Vonsattel score (H-V score), which independently characterizes both striatal and cortical pathological involvement in each brain [28]. While Vonsattel grading and H-V striatal score are closely related, (Pearson r = 0.90, measured using 346 HD brains), H-V scores are a continuous metric and therefore more amenable to adjustment of covariates such as CAG repeat size in modeling of neuropathological involvement and independently assesses striatal and cortical involvement. H-V scores ranged from 0–4, where 0 indicates no detectable neuropathological involvement and 4 indicates severe neuropathological involvement. Samples from symptomatic individuals had striatal scores ranging 1.43–3.82 and cortical scores ranging from 0.40–2.36 (see Table 6). Additionally, two Grade 0 brains (both with CAG repeat expansions of 42 repeats) were small-RNA sequenced and analyzed separately from the 26 HD brains used in differential expression analysis. Grade 0 brains were neuropathologically normal and asymptomatic at the time of death (see Table 1).

Table 6. Summary of the brain samples used for miRNA-sequence analysis.
After processing sequencing data to remove sequencing artifacts, normalize using variance stabilization transformation, and adjust for batch effects (see Methods), 938 miRNAs were reliably quantified and 75 of these were significantly differentially expressed in HD versus control brains after adjusting for multiple comparisons (FDR q-value < 0.05, see Appendix Table 1). In HD, 46 miRNAs were identified as significantly up-regulated and 29 as down-regulated in their expression. Hox-related miRNAs had the most extreme, positive fold changes, where miR-10b-5p was 3.9 log2 fold increased, miR-196a-5p was 2.4 log2 fold increased, miR-615-3p was 1.6 log2 fold increased, miR-10b-3p was 1.5 log2 fold increased, and miR-196b-5p was 1.3 log2 fold increased (see Figure 9, see Table 7). Both the 5′ and 3′ mature miRNAs were differentially expressed for eight miRNA precursors (miR-10b, miR-129, miR-1298, miR-142, miR-144, miR-148a, miR-302a, and miR-486). In HD and controls, most 5′-3′ miRNA pairs were positively correlated in their expression, with the exception of miR-1298 in HD and miR-10b and miR-302a in controls.
Appendix Table 1. Differentially expressed miRNAs in Huntington’s disease prefrontal cortex

Figure 9. Characterization of miRNA in Huntington's disease brain. Volcano plot of 75 significantly differentially expressed miRNA after FDR-adjustment for 938 comparisons. Points labeled red were up-regulated in HD and points labeled as blue were down-regulated in HD. Hox-related miRNA points are labeled and represent the top differentially expressed miRNA in HD.

To confirm our previously published findings, we re-analyzed the twelve HD and nine controls samples from our original study using our updated sequence analysis pipeline (see Methods) and then used the newly sequenced samples, consisting of 14 HD and 27 control brains, as a replication set. Fourteen miRNAs were significantly differentially expressed (FDR q-value < 0.05) in the original set using the updated analysis pipeline, compared to five differentially expressed
miRNA in the original study. Fourteen differentially expressed miRNAs were significantly differentially expressed in the replication set and thirteen of these fourteen were significant in the combined sequence analysis. As previously reported, Hox-related miRNA, including miR-10b-5p, were among the most strongly differentially expressed across all three studies (see Appendix Table 1).

Firefly Bioworks microRNA assay, a multiplexed, particle-based technology using flow cytometry to measure miRNA levels, was used to quantify and orthogonally validate miRNA differential expression from sequencing (see Methods). A subset of 21 controls and 15 HD samples from the sequencing study were selected for the assay. Sixteen miRNAs with moderately high expression levels were selected for testing and an additional six miRNAs were used as input normalizers. miR-10b-5p was confirmed as significant after correcting for multiple corrections (p-value = 3.0e-10, q-value = 6.6e-9). Seven out of sixteen miRNAs assayed approached but did not reach significance after adjustment in this subset (unadjusted p-value < 0.05, miR-10b-5p, miR-194-5p, miR-223-3p, miR-132-3p, miR-144-5p, miR-148a-3p, miR-486-5p). Eight of the remaining nine miRNAs that failed to achieve significance had the same direction of effect. These results were consistent with the reduced power available from this subset.

**Nine miRNAs relate to Vonsattel grade.** To explore the relationship of miRNA expression to principal clinical aspects of the disease, we next modeled the expression of the 75 differentially expressed miRNAs to the Vonsattel grade of neuropathological involvement. Analysis of variance (ANOVA) was performed
to compare the expression of the 75 differentially expressed miRNAs across Vonsattel grade in all 28 (Grade 0–4) HD gene-positive and control brains. 65 miRNA were found to be significant in the ANOVA (FDR-adjusted q-value < 0.05), indicating differential expression may be driven by the difference of controls to specific grades. Next, ANOVA was performed exclusively in HD brains to find whether miRNA differences exist across Vonsattel grades. Nine miRNAs were significant in both ANOVA tests after adjusting for multiple comparisons, indicating a significant difference in the expression of these miRNAs across Vonsattel grades (both FDR q-values < 0.05). Last, pairwise comparisons of each grade with the control group were performed using post-hoc Tukey’s HSD (honestly significant different) tests to find specific groups that significantly differed from one another. **Figure 10** highlights the nine miRNAs that are associated with grade in order of statistical significance from the ANOVA inclusive of control brains in the test. In **Figure 10**, significant differences across grade and control groups as determined by Tukey HSD are denoted by letters (a-d) in the grey banner above each boxplot, whereby groups with different letters are significantly different from one another while those which share letters are not.

**Figure 10. Nine miRNAs are associated with Vonsattel grade.**
In HD brains, expression of differentially expressed miRNA was compared across Vonsattel grades 0–4. Boxplots represent nine FDR-significant miRNAs (FDR q < 0.05, adjusted for 75 contrasts) associated with Vonsattel grade by analysis of variance (ANOVA). X-axes represent Vonsattel grade, classified 0–4 in order of the severity of striatal involvement and Y-axes show the VST expression values after batch correction. Significant differences across grades and controls are denoted by letters in the grey banner above the boxplot, labeled
a-d. Groups with different letters are significantly different from one another while those with the same letter are not, after correcting for multiple comparisons. For example, group “a” would be significantly different from group “b” and “c.” Conditions represented by multiple letters indicate no significant difference among those groups. For example, group “ab” would not be significantly different than groups “a” and “b,” but would be different group “c.”

Several patterns in the relationship of grade to miRNA expression were observed. First, the expression of miR-10b-5p was significant in nearly all comparisons; pairwise contrasts between all grades as well as with the control group were different except for grade 0, although grade 0 was different than grades 2, 3 and 4 (see Figure 10A). Second, the expression of miRNAs in grade 0 brains was rarely different than controls, with the exception of miR-200c-3p, where its expression in grade 0 brains was significantly lower than both controls and grades 2–4 brains (see Figure 10G). Third, the expression of miRNAs in grade 3 and 4 brains appeared relatively similar to one another, with the exception of miR-10b-5p, as mentioned above, and miR-4488, where grade 3 brains were significantly lower than all other groups (see Figure 10D). Although not significant in the HD-only ANOVA, significant pairwise differences between grade 3 and 4 were observed for miR-1298-5p (Bonferroni q-value = 3.6e-2) and miR-615-3p (Bonferroni q-value = 2.2e-2).

To assess the sensitivity and specificity of miR-10b-5p for predicting HD, area under the curve (AUC) values were calculated using receiver operating characteristic curves (ROC). When comparing HD to controls to predict HD, the AUC was 99.47% (95% confidence level was 98.46%-100%). In a comparison of asymptomatic HD to HD to predict HD status, the AUC was 98.08% (95% confidence level was 92.75%-100%) and comparing asymptomatic HD to controls, the AUC was 84.72% (95% CI: 71.09%-98.36%).
miRNA expression relates to striatal involvement and age of onset in HD. To further elucidate the meaning of the associations of the miRNAs to HD, we examined the relationship between the 75 differentially expressed miRNAs and other salient features of the disease (age at motor onset, disease duration, age at death, and H-V scores of striatal and cortical involvement). To avoid confounding the analysis of these clinical features by the known, strong relationship between HTT CAG repeat size and disease pathology and onset [4,28,31,32], CAG-adjusted residuals were calculated for all continuous clinical traits (see Figure 11).

Figure 11. Association of clinical features to HD CAG repeat size. CAG-adjusted residuals for onset age, death age, duration, H-V striatal score and H-V cortical score were computed from data derived from 346 HD brain samples with CAG repeat sizes <56 from Hadzi et al. [28]. Red dots represent samples studied in these analyses.
Using linear regression analysis and applying FDR-adjustment for the 75 comparisons, three miRNAs (miR-10b-5p, miR-10b-3p, miR-302a-3p) were observed to have a significant relationship to CAG-adjusted striatal score (FDR q-values = 2.28e-2). All three were significant in the analysis of miRNA expression to Vonsattel grade (see above). Additionally, five miRNAs were identified as having significant association to CAG-adjusted age of onset (miR-10b-5p, FDR q-value = 3.49e-3; miR-196a-5p, FDR q-value = 1.32e-2; miR-196b-5p, FDR q-value = 1.71e-2; miR-10b-3p, FDR q-value = 1.71e-2; miR-106a-5p, FDR q-value = 1.71e-2). **Figure 12** highlights the relationship of miR-10b to CAG-adjusted striatal score and onset, where both 3p and 5p mature sequences of miR-10b were the only miRNA species to have significant, linear
association to these two clinical features independent of CAG effect. No FDR-significant relationships of miRNA to disease duration or death age were observed.

**Figure 12. miR-10b is associated with age of onset and striatal involvement.**

In 26 Vonsattel grade 2, 3 and 4 HD brains, both mature miR-10b sequences (-5p and -3p) have FDR-significant relationships to clinical features of HD. Y-axes show the variance stabilizing transformation expression values after batch correction. Grade 0 cases are not included, as they have neither onset age nor Hadzi-Vonsattel (H-V) striatal score.

A. Scatterplot of miR-10b-5p expression and CAG-adjusted H-V striatal score. B. Scatterplot of miR-10b-3p expression and CAG-adjusted H-V striatal score. C. Scatterplot of miR-10b-5p expression and CAG-adjusted onset age. D. Scatterplot of miR-10b-3p expression and CAG-adjusted onset age.
No significant relationship of the expression of the 75 differentially expressed miRNA to CAG-adjusted cortical score was observed, although nominal associations were seen. In order to account for the potential impact of cortical involvement on the relationship of miRNA expression to striatal involvement, we performed a multivariate regression analysis modeling miRNA expression to striatal H-V score while correcting for cortical H-V score. After
CAG-adjusted cortical score correction, CAG-adjusted striatal score remained significant (miR-10b-5p p-value = 0.04, miR-10b-3p p-value = 0.01, miR-302a-3p p-value = 0.005).

Last, to characterize the patterns of association of miRNAs to clinical features, Pearson coefficients of the correlation of the expression of the differentially expressed miRNAs to five CAG-adjusted features (onset age, disease duration, death age, striatal score and cortical score) were hierarchically clustered. Grade 0 and controls samples were not included in these analyses. Correlation coefficients rather than beta coefficients were used in order to standardize the direction of effect. Here, we observed differentially expressed miRNAs with correlation p-values < 0.05 clustered into distinguishable patterns of association to clinical variables (see Figure 13). Differentially expressed miRNAs increased in HD compared to controls tended to have negative correlations with onset and death, and positive correlations with striatal and cortical score. Conversely, differentially expressed miRNAs with negative relative fold changes had positive correlations with onset and death, and negative correlations with striatal and cortical scores.

Figure 13. CAG-adjusted clinical features of HD show patterns of association with miRNA expression.
CAG-adjusted measures of onset age, disease duration, death age, Hadzi-Vonsattel (H-V) striatal and cortical score were correlated with differentially expressed miRNAs in HD brains. miRNAs with at least one nominal p-value < 0.05 are shown. Pearson correlation coefficients and features were independently hierarchically clustered. Red boxes indicate positive correlations and blue boxes indicate negative correlations. Seven miRNAs in the left section are down-regulated in HD and the ten miRNAs in the right section are up-
regulated. Unsupervised clustering separated miRNA by their direction of fold change.

**Targets of HD-related miRNAs are associated with nervous system development and transcriptional regulation.** To attempt to understand the potential functional impact of miRNA dysregulation in HD, gene ontology enrichment was performed using predicted targets for miRNAs that correlated with clinical features and were suitably annotated in Targetscan (twelve miRNAs in total). 5712 unique mRNA targets for miRNAs with positive fold change in HD (miR-106a/302a-5p, miR-196a/miR-196b, miR-302a-3p, miR-363, miR-10b, miR-615-3p), and 6572 mRNA targets for negative fold change in HD (miR-129-3p, miR-129-5p, miR-132-3p, miR-4449, miR-4488, miR-490-5p) were found using Targetscan [33], and stratified by fold change for gene ontology term (GO) enrichment analysis. Using TopGO’s weight algorithm with Fisher’s Exact Test for gene ontology term enrichment and a weighted p-value cutoff less than \( p < 0.05 \), 354 GO Biological Processes, 86 GO Molecular Functions and 62 GO
Cellular Component terms for mRNA targets of down-regulated miRNA were significant. 260 GO Biological Processes, 78 GO Molecular Functions, 48 GO Cellular Component terms for mRNA targets of up-regulated miRNA were significant.

To make these long lists of GO terms more intelligible, terms were summarized using semantic similarity measures to remove gene-set and GO term redundancy (see Methods).

Targets of up- and down-regulated miRNAs had substantial overlap in their overall function. Three of the top twenty collapsed GO Biological Processes terms were shared between the two sets of targets (see Figure 14A). These terms were “nervous system development,” “Fc-epsilon receptor signaling pathway,” and “proteasome − mediated ubiquitin − dependent protein catabolic process.” “Nervous system development” was the most significant term in both sets (Up $p = 8.5e-5$, Down $p = 9.9e-7$). The top enriched term was “positive regulation of transcription, DNA-templated”, (N = 1678, $p = 2.7e-4$) for the positive gene set and “synaptic transmission”, (N = 3166, $p = 3.4e-6$) for the negative gene set. Of the 78 up-regulated Molecular Function terms and 86 down-regulated terms, fifteen terms were the same (see Figure 14B). Top terms were included “sequence-specific DNA binding transcription factory activity”, “sequence-specific DNA binding” and “calcium ion binding”. Though shared between the two groups, “transcription factor binding” was enriched higher in down-regulated miRNAs than positive ones. For GO Cellular Component, six
terms were the same between the two gene sets. These terms included “nucleus” and “cytoplasm” as well as “cell junction” (see Figure 14C).

**Figure 14. Gene ontology terms are similar for mRNA targets of clinically relevant deregulated miRNAs.**

**Figure 14A** illustrates the overlap in GO Biological Processes between targets of increased miRNA (in orange) and decreased miRNA (in blue) in HD. The x-axis shows the number of gene ontology terms that fall within a given semantic term set, and the y-axis lists the top twenty enriched terms for each set of miRNA targets. Darker colored points represent terms with higher significance and the size of the points represents the union of all genes that fall within a given the term. A number of terms, including “nervous system development” as well as terms relating to transcriptional regulation are shared across up- and down-regulated miRNA target groups. The similarity targets of up-regulated miRNA (in orange) and down-regulated miRNA (in blue) for GO Molecular Function are seen in **Figure 14B** and for GO Cellular Component in **Figure 14C**.

**Discussion**

In a next-generation sequence analysis of small non-coding RNAs in 26 HD and 36 control brains we detected 938 miRNAs and 75 of these were
differentially expressed. All five miRNAs reported as differentially expressed in our previous study (miR-10b-5p, miR-196a-5p, miR-196b-5p, miR-615-3p and miR-1247-5p) were significantly differentially expressed in this study [13]. These results were independently validated in the 41 (14 HD and 27 control brains) newly studied brains (see Appendix Table 1), and support the presence and robust up-regulation of Hox-related miRNAs in HD brain [13]. The increased number of differentially expressed miRNAs is likely due to an increase in sample size. Increasing our sample size (from N = 21 to N = 62) enhanced the statistical power to detect additional miRNAs with smaller but significant changes in miRNA expression. We believe these miRNA signals are not attributed to a change in BA9 architecture, influenced by neuronal cell death or reactive glial response, because cell numbers between HD and controls from the same brain samples were indistinguishable [13,30].

Dysregulation of several miRNAs from our study have been observed in HD in other contexts. Concordant with our findings, miR-132-3p down-regulation in human HD parietal cortical tissue [15] and in brains of R6/2 and YAC128 HD mouse models has been observed [15,18]. miR-132 is highly enriched in the brain [34,35] and its expression has been shown to affect neuron morphogenesis and enhance neurite outgrowth by suppressing the GTPase-activating protein p250GAP (p250GAP/RICS) [36]. Another target of miR-132 is acetylcholinesterase (ACHE), which encodes an enzyme responsible for the breakdown of the neurotransmitter acetylcholine at the neural synapse [37].
Acetylcholinesterase is critically involved in cognition, and acetylcholinesterase inhibitors are FDA-approved for the treatment of cognitive impairments in Alzheimer’s disease [37]. Thus, decreased miR-132 levels may negatively impact brain health, through the dysregulation of p250GAP (limiting its suppression) and ACHE (indirectly decreasing acetylcholine levels).

Differentially expressed miRNAs may also target HTT transcripts as a response to mutant HTT to reduce HTT transcriptional levels and limit toxicity. miRNAs that target the HTT 3'UTR and reduce HTT transcript levels in vitro, miR-148a-5p, miR-150-5p and miR-214-5p, were significantly up-regulated in their expression [21,38]. Although miR-196a does not directly target HTT [33], increased miR-196a expression was observed in a primate model of HD and its over-expression in vitro and in animal transgenic models suppressed mutant HTT expression [24]. The miRNAs with the largest effect in our study, miR-10b-5p, putatively targets HTT by binding to two 3'UTR sites (both 7mer-1A seed, positions 2742–2748 and 3301–3307) and may reduce expression of HTT although it is not clear whether or not this would be neuroprotective [33]. However, miR-10b-5p also targets brain-derived neurotrophic factor (BDNF) [39], a growth factor required for the survival and differentiation of striatal neurons [40]. BDNF has been extensively studied in HD [41], as normal huntingtin protein is reported to up-regulate BDNF levels, while mutant huntingtin impairs BDNF protein abundance which may consequently lead to death of striatal neurons [42]. Because of the potential biological importance of BDNF, and the possibility that
miR-10b-5p may diminish translation of BDNF, over-expression of miR-10b-5p might be harmful to neuronal cells. However, in Hoss et al. [13], we showed that ectopic expression of miR-10b-5p in PC12 cells expressing a mutant huntingtin fragment enhanced cell survival [13], and miR-10b-5p has been observed to facilitate neurodifferentiation [43]. Given its high levels of differential expression, strong relationship to striatal involvement and age at onset, more research into miR-10b-5p is justified to understand its role in the pathogenesis of HD, its potential as a biomarker of disease progression and its potential as a therapeutic target.

The cell type most responsible for miRNA changes cannot be determined from these data. Tissue homogenate was used for sequencing, so the source of miRNA signal is likely both neuronally and non-neuronally derived. To determine the miRNA cellular specificity in the brain, Jovicic et al. [16] measured miRNA expression in cultured neurons, oligodendrocytes, microglia and astrocytes to find miRNAs enriched for each cell type. Based upon this study, miRNAs found to be specifically enriched in neuronal cultures (miR-129-3p, miR-129-5p, miR-132, miR-135b, miR-431, miR-433) were all down-regulated in our study whereas miRNAs enriched in microglial cultures (miR-126-5p, miR-126-3p, miR-141, miR-142-3p, miR-142-5p, miR-150, miR-200c and miR-223) were all up-regulated. According to these enrichment categories, microglial activation miRNAs do not relate to clinical features of the disease. Conversely, three neuronal-related miRNAs, miR-129-3p/5p and miR-132, were associated with
pathological involvement (see Figure 13). Therefore, we hypothesize that
differential expression of those miRNAs related to neuron function may also relate to the HD pathology.

The relationship of the expression of miRNAs with Vonsattel grade suggests expression changes may occur early in the disease process (see Figure 10). Many of these miRNA changes appear present ordinal trends with an increase (miR-10b-5p, miR-10b-3p, miR-302a, miR-196a-5p, miR-196b-5p) or decrease (miR-663b, miR-4488, miR-4449) in their expression across grade. In particular, miR-10b-5p was significantly different across all groups, with the exception of the asymptomatic grade 0 brains and we believe this is an issue of statistical power. It is possible that the expression of these miRNAs may relate to HTT aggregation or proteasomal degradation, as intranuclear inclusions are observed in pre-symptomatic HD [44] and the density of aggregate formation continues over the course of the disease. Three miRNAs (miR-10b-3p/5p, miR-302a) related to H-V striatal score, independent of the CAG repeat expansion size and for miR-10b-5p, independent of cortical involvement. These results suggest the relationship of miRNA expression to striatal involvement in the disease is independent of cortical involvement, which is a critical finding, because prefrontal cortex was the source of tissue profiled in these studies.

Based on correlation (see Figure 13), up-regulated miRNAs clustered together based on their relationships to clinical features. Generally, these miRNAs had strong, positive associations to striatal and cortical H-V scores,
weak positive association with disease duration and strong negative associations to onset and death age. Down-regulated miRNAs clustered together as well but were less defined in their relationships to clinical features. Most down-regulated miRNAs were inversely associated with H-V scores and duration, opposite to up-regulated miRNAs. These patterns suggest that decreasing up-regulated miRNAs and increasing down-regulated miRNA may be beneficial. However, it remains to be determined which altered miRNAs are compensatory and potentially neuroprotective and which are pathological and neurotoxic. Furthermore, it is unknown whether these changes are consequential, revealing important molecular aspects of the disease process, or are simply innocent by-products.

However, using target analysis and GO term enrichment, we observed predicted targets of both up- and down-regulated miRNAs shared many of the same biological processes and overall systems relating to “nervous system development.” Both sets contained several transcriptional regulation related terms (transcriptional regulation, DNA-dependent or RNA polymerase II, chromatin remodeling, post-transcriptional gene regulation, chromatin remodeling, etc.). Both sets of genes contained terms on metabolism, apoptosis, metal-binding and ubiquitin. Disruption to any of these systems may affect neuron health. Overall, these finding imply both up- and down-regulated miRNAs many be part of the same or similar biological pathways.
**Conclusions**

Our findings identify many miRNA alterations in HD brain and a large number of these are related to clinical manifestations of the disease, where the signal is independent of the size of the CAG repeat expansion. The study of Grade 0 cases suggests that miR-10b-5p expression changes may occur pre-symptomatically. Up- and down-regulated miRNAs may target genes in similar biological systems, and these genes are involved in transcriptional regulation, neuronal development and other important aspects surrounding neuron function. These miRNAs represent attractive candidates for predicting onset age and overall health of the striatum in HD. Studies pursuing these miRNAs as potential biomarkers for HD are in progress, as miRNAs may be detectable in peripheral fluids [45] and thus have potential to function as accessible biomarkers for disease stage, rate of progression, treatment efficacy and other important clinical characteristics of HD.

**Methods**

**Sample information.** Frozen brain tissue from prefrontal cortex Brodmann Area 9 (BA9) was obtained from the Harvard Brain and Tissue Resource Center McLean Hospital, Belmont MA, Banner Sun Health Research Institute, Sun City, Arizona [46] and Human Brain and Spinal Fluid Resource Center VA, West Los Angeles Healthcare Center, Los Angeles, CA. 26 Huntington’s disease (HD) samples, 2 asymptomatic HD gene carriers, and 36 neurologically and neuropathologically normal control samples were selected for
the study. HD subjects had no evidence of other neurological disease based on neuropathological examination. HD samples and controls were not different in postmortem interval (PMI) (p-value = 0.69), RNA integrity number (p-value = 0.08) or gender (p-value = 0.51) but differed in ages at death (HD mean age = 59.5, control mean age = 68.6; p-value = 0.01) (see Table 1). Asymptomatic HD samples did not differ in age at death (mean age = 67.5) in comparison to HD or control samples (control p-value = 0.92; HD p-value = 0.40). Information on CAG genotype, onset age, death age, disease duration, Vonsattel grade, Hadzi-Vonsattel striatal and cortical scores for HD samples can be found on GEO.

Total RNA was isolated using QIAzol Lysis Reagent and purified using miRNeasy MinElute Cleanup columns. RNA quality for sequencing was assessed using Agilent’s BioAnalyzer 2100 system and RNA 6000 Nano Kits to determine RNA Integrity Number or Agilent 2200 TapeStation and Agilent DNA ScreenTape assay RNA Quality Number. For each brain sample, 1 ug of RNA was used to construct sequencing libraries using Illumina’s TruSeq Small RNA Sample Prep Kit, according to the manufacturer’s protocol, and sequenced using 1x51nt single-end reads on Illumina’s HiSeq 2000 system at Tufts University (http://tucf-genomics.tufts.edu/) or the Michigan State sequencing core facility (http://rtsf.natsci.msu.edu/genomics/).

**miRNA sequence analysis.** Reads were quality filtered, removing reads below 80% Q20, using FASTX-toolkit FASTQ quality filter (version 0.0.13.2, http://hannonlab.cshl.edu/fastx_toolkit/). Adapter sequence (5’-
TGAAGTTCTCGGGTGCCAAGG-3') was removed from the 3’ end of all reads using cutadapt 1.2.1 (http://code.google.com/p/cutadapt/) and reads less than 15 nucleotides in length were discarded [47]. Reads were collapsed using FASTX-toolkit FASTA/Q collapse. Reads were aligned to the UCSC human reference genome (build hg19) using Bowtie version 1.0.0, using no mismatch alignments and a limit of 200 multiple mapping instances [48]. Aligned reads that overlapped with the human miRNA annotation, miRBase version 20, (http://www.mirbase.org/ftp.shtml) were identified using BEDTools IntersectBed [49]. Reads longer than 27 bases were removed. miRNA reads were counted if ±4 nucleotides from the mature, annotated 5’ start coordinates. Reads that mapped to multiple locations, represented by a single mature miRNA, were recorded as a single miRNA count. Multi-mapped reads represented by multiple mature miRNA annotations were discarded. R version 3.1.0 and Bioconductor 2.1.4 version were used for differential expression analysis. DESeq2 version 1.40.0 was used for estimation of library size and correction, as well as variance-stabilizing transformation (VST) [50,51]. miRNAs with a mean less than 2 raw read counts across all samples were removed. Batch effect was corrected using ComBat with default options through the Bioconductor package sva 3.10 [52,53]. All samples were included in VST and batch correction. Using 36 controls and 26 HD grades 2–4, differential expression analysis was performed with LIMMA version 3.20.8 [54,55], adjusting for age at death in the model. Q-values were FDR-adjusted for 938 comparisons. The unprocessed fastq files, normalized
miRNA counts and results from miRNA differential expression analysis have been deposited in NCBI’s Gene Expression Omnibus [56], and are accessible through GEO Series accession number GSE64977 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE64977).

**Firefly miRNA assay.** A panel of 16 differentially expressed miRNAs with moderate to high expression (miR-10b-5p, miR-194-5p, miR-223-3p, miR-132-3p, miR-144-5p, miR-148a-3p, miR-486-5p, miR-363-3p, miR-199a-5p, miR-16-2-3p, miR-142-3p, miR-34c-5p, miR-129-5p, miR-433-3p, miR-885-5p, miR-346) and six stably expressed miRNAs in sequencing (miR-9-5p, miR-92a-3p, miR-98-5p, miR-101-3p, miR-151a-3p, miR-338-3p) was used for validation. In a 96-well filter plate, Firefly Multimix (Firefly BioWorks, www.fireflybio.com) was incubated with 25ul Hybridization Buffer and 25ul total RNA at a concentration of 1 ng/ul at 37°C for 60 minutes. After rinsing to removing unbound RNA, 75ul of Labeling Buffer was added to each well, and the plate was incubated for 60 minutes at room temperature. Adapted-modified miRNAs were released from the particles using 90°C water, and PCR amplified using a fluorescently-label primer set. PCR product was hybridized to fresh Firefly Multimix for 30 minutes at 37°C and re-suspended in Run Buffer for readout. Particles were scanned on an EMD Millipore Guava 8HT flow cytometer. Raw output was background subtracted, normalized using the geometric mean of the six normalizer miRNAs and log-transformed. LIMMA version 3.20.8 [54] was used to calculate significance.
**HD feature analysis.** For analysis of miRNA expression to Vonsattel grade, Tukey HSD statistics and compact letter display were generated by the multcomp R package [57]. CAG-adjusted age of onset was calculated using the logarithmic model from Djousse et al. 2003 [4]. Hadzi-Vonsattel striatal and cortical scores were measured in 523 HD brain samples as previously described [28]. Samples with greater than 55 repeats or missing CAG information were excluded from analysis, leaving 346 samples. To provide robust residual estimates for the subset of samples included in the sequencing project, H-V striatal score, H-V cortical score, death age and disease duration features were corrected for CAG size by modeling each feature to CAG size within the HD dataset (N = 346) and the residuals from the model were extracted for each sample (see Figure 11) [28]. VST-batch corrected counts were used for all subsequent analyses. CAG-adjusted residuals and miRNA expression relationships were analyzed using linear regressions. Covariates (PMI, RIN, age at death) were not included in linear models, as neither PMI nor RIN were determined to have an effect on the outcome of the results. Age at death could not be included in the analysis due to the relationship of age at death and HD clinical pathology. Q-values were FDR-adjusted for 75 differentially expressed miRNA contrasts for linear regressions were reported. For the cluster analysis in Figure 13, Pearson correlations for miRNA expression to clinical feature were performed and those miRNAs with $p$-values < 0.05, without adjustment for multiple comparisons, were reported. Pearson correlation
coefficients were hierarchically clustered using Euclidean distance and unsupervised complete clustering method through the R-package pheatmap version 0.7.7.

**Target prediction and gene ontology enrichment.** Targets can, release 6.2 [33] was used to select mRNA targets of miRNAs with at least one relationship to clinical feature. Fourteen miRNAs were available on Targets can and twelve miRNAs had unique seed sequences. Targets were removed with total context scores $\geq -0.1$. miRNAs with positive fold change in HD (miR-106a/302a-5p, miR-196a/miR-196b, miR-302a-3p, miR-363, miR-10b, miR-615-3p), and negative fold change in HD (miR-129-3p, miR-129-5p, miR-132-3p, miR-4449, miR-4488, miR-490-5p) were stratified for gene ontology term (GO) enrichment analysis. GO term enrichment for “biological processes,” “molecular function,” and “cellular component,” was performed using topGO [58] with the “weight01” algorithm and Fisher statistic within the R statistical environment. A weighted Fisher $p$-value $< 0.05$ threshold was used to select significant GO enrichment. Significant terms were collapsed by semantic similarity using the program REVIGO [59], with $p$-value included for each term and using the “Small (0.5)” similarity setting. The union of genes from REVIGO “parent” terms was calculated using topGO’s genes.in.term function.

References


39. Varendi K, Kumar A, Harma MA, Andressoo JO: miR-1, miR-10b, miR-155, and miR-191 are novel regulators of BDNF. Cellular and molecular life sciences: CMLS 2014


CHAPTER 4: STUDY OF PLASMA-DERIVED MIRNAS MIMIC DIFFERENCES IN HUNTINGTON’S DISEASE BRAIN


Introduction

Huntington’s disease (HD) is an autosomal dominant inherited movement disorder, caused by an expanded CAG trinucleotide repeat sequence in the huntingtin gene [1]. The emergence of abnormal, choreiform movements, often accompanied by neurophysiological, psychiatric or cognitive impairments [2], defines disease onset and generally occurs around middle age [3]. Neurodegeneration precedes clinical diagnosis. As many as half of the neurons in the caudate nucleus in the striatum are lost by the onset of motor signs of disease [4], and volumetric changes in the striatum occur as early as two decades prior to predicted onset age [5]. Because motor and cognitive impairments correlate with the neuroanatomical changes in the striatum [6], to prevent neuronal loss and thereby prevent or delay disease onset, therapeutic intervention would ideally occur prior to HD manifestation.

While predictive genetic testing can reliable detect HD, the lack of validated biomarkers for HD progression limits the evaluation of preventive and early-stage disease-modifying therapies. Current progressive measures for prodromal and
early-stage HD patients rely on rating of functional decline, which are susceptible to inter-rater variability and limited sensitivity [2].

Large, multicenter, longitudinal studies comparing large cohorts of asymptomatic HD gene carriers and early-stage HD to healthy controls have used a battery of clinical and neuroimaging based assessments, aimed at identifying robust quantitative measures [7, 8]. These studies, among others, have identified a number of image-based biomarkers that may relate to HD progression, such as morphometric changes [7, 9], elevation in the glial cell marker myo-inositol [10], the reduction in the neuronal integrity marker N-acetyl aspartate [10], and recently, the medium spiny neuron marker PDE10A [11]. Although neuroimaging biomarkers are encouraging, prior to clinical adoption, problems in reproducibility across imaging facilities from technical and analytical inconsistencies must first be addressed [12].

Alternatively, the disease mechanisms that are observed in the brain may be detectable in the blood. Research measuring oxidative stress using 8OHdG levels [13], mutant HTT accumulation [14], inflammatory markers [15], and genome-wide RNA changes [16-18], have uncovered a number of concordant changes between brain and blood. While promising, the clinical utility of these measures is yet to be reported.

Recently, we completed comprehensive evaluations of altered miRNA levels obtained through next-generation sequencing technology in human HD and control prefrontal cortex samples to investigate their role in transcriptional
dysregulation in this disease [19]. We identified 75 miRNAs significantly altered in HD and several of these showed significant associations to HTT CAG repeat-adjusted age at motor onset, or extent of striatal neuropathological involvement [20], including miR-10b-5p, associated with both. Furthermore, among asymptomatic HD gene carriers, levels of miR-10b-5p were distinguishable from both the low expression observed in normal controls and higher levels seen among symptomatic HD patients, suggesting a relationship between cortical levels of miR-10b-5p and disease stage.

Because brain-derived miRNAs may pass through the blood-brain-barrier by exosome transport [21], and because they are remarkably stable [22], the miRNAs identified in our studies of postmortem brain tissue may be detectable in peripheral fluids [23-25], and thus potentially provide accessible biomarkers for disease stage and rate of progression in clinical trials.

The first step in the evaluation of clinical utility is to determine whether the observed HD brain-related miRNA alterations may also be detected in peripheral samples. To evaluate the biomarker potential for miRNAs, we compared the levels of four miRNAs related to HD clinical features in postmortem brains (miR-10b-5p, miR-486-5p, miR-132-3p, and miR-363-3p) in HD, asymptomatic and healthy control plasma samples.

**Methods**

Study participants (n=38) were recruited through the Boston University Neurological Associates (BUNA) and Tewksbury State Hospital from 2012-2014,
with appropriate IRB approval and consent (Protocol Number H-31052) (see Table 7). No significant differences were observed in the age or sex distribution between any of the 3 groups (HD cases, asymptomatic HD carriers, and controls). A trained phlebotomist drew blood by standard practice. BD Vacutainer CPT Mononuclear Cell Preparation Tubes containing 0.1 mL sodium citrate anticoagulant and 0.1 M Ficoll medium were used to isolate plasma from 8 mL of whole blood.

Table 7. Summary of the samples used for the study.

<table>
<thead>
<tr>
<th>Condition</th>
<th>N</th>
<th>Age</th>
<th>Onset age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>8</td>
<td>46.1 ± 13.5</td>
<td></td>
<td>3M, 5F</td>
</tr>
<tr>
<td>Asymptomatic</td>
<td>4</td>
<td>42.5 ± 28.7</td>
<td></td>
<td>1M, 3F</td>
</tr>
<tr>
<td>HD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manifest HD</td>
<td>26</td>
<td>53.0 ± 8.7</td>
<td>47.6 ± 9.9</td>
<td>11M, 15F</td>
</tr>
</tbody>
</table>

One mL of plasma was used for RNA extraction. To minimize platelet contamination [26], residual platelets were removed by centrifugation, collecting supernatant after spinning for 5 min at 16,000 x g. 0.22 um filtration was used to remove heterogeneous, phospholipid membrane bound microparticles, 0.05-1.5 um in size, shed from platelets and other blood cells [27]. RNA was extracted using Qiazol and miRNeasy RNA isolation kit from Qiagen, according to manufacturer’s protocol. RNA purity and abundance was assessed by spectrophotometry.

The four miRNAs included in this study were selected based on the following criteria: (1) All four were altered at genome-wide significant levels in HD brain [20] (up in HD: miR-10b-5p, miR-486-5p, miR-363-3p; down in HD: miR-
(2) All four were readily abundant in both brain and blood [28], and (3) candidate miRNAs showed nominal association in the cortical study (p<0.05) to CAG-adjusted clinical HD features (onset: miR-10b-5p, miR-486-5p miR-363-3p; striatal neuropathological involvement: miR-10b-5p, miR-132-3p).

Exiqon miRCURY LNA Universal RT miRNA PCR was used to assay miR-10b-5p (cat. 205637), miR-486-5p (cat. 204001), miR-132-3p (cat. 204129), and miR-363-3p (cat. 204726), following the manufacturer’s protocol. UniSp6 synthetic spike-in was used to evaluate cDNA efficiency. Following cDNA synthesis, samples were diluted to 0.2 ng/ul in RNAse free water. Both SNORD44 (cat. 203902) and miR-451a (cat. 204737) were used for normalization. For quantitative PCR (qPCR), samples were assayed in triplicate across three 384-well plates, using Applied Biosystems 7900HT Real-Time PCR System. For analysis, threshold cycle (C\text{T}) values for triplicate wells were normalized by average RNU44 and miR-451a values. Extreme outlier wells and samples (standard deviations above 10) were removed. miRNA levels were calculated using the ΔΔC\text{T} method [29].

Linear regression analyses predicting -ΔΔC\text{T} were used to test the association between miRNA levels and disease status in HD cases and controls for the four miRNA. One-tailed tests were used to test the a priori hypothesis of consistent direction of effects as observed in the brain study, thus any relationships inconsistent with the cortical findings would not be identified as significant. Bonferroni correction, assuming four comparisons was applied.
For miR-10b-5p, we further examined the relationship between asymptomatic HD gene carriers and HD cases and controls separately using linear regression, again using a one-tailed test of the previously observed relationships. In addition, cumulative logit models using the “ordinal” package in R [30], were applied to test an ordinal relationship between controls, asymptomatic gene carriers and HD cases using a Chisq likelihood-ratio test. Finally, we examined the relationship of miR-10b-5p to age of motor onset in cases using linear regression.

Results

All four miRNAs were detected in plasma at reasonable levels (miR-10b-5p average \( C_T = 33.8 \), miR-132-3p average \( C_T = 32.4 \), miR-363-3p average \( C_T = 32.2 \), miR-486-5p average \( C_T = 25.7 \)). After calculating -\( \Delta \Delta C_T \), levels of candidate miRNAs in the 26 HD patients were compared to the 8 controls to test whether miRNAs alterations in HD plasma resembled changes in observed in HD brain. We observed increased levels of miR-10b-5p (one-sided \( p = 6.77e^{-3}, \beta = 2.39 \)) and miR-486-5p (one-sided \( p = 0.044, \beta = 1.44 \)) between HD and control subjects. These changes are consistent with the changes observed in the brain [20] where miR-10b-5p and miR-486-5p were also increased (see Table 8). Levels of miR-132-3p in plasma, though not significant, were lower in HD, consistent with the results in brain (one-sided \( p = 0.92, \beta = -0.62 \)). miR-363-3p levels were not altered in blood, nor consistent with changes observed in HD brain (one-sided \( p = 0.99, \beta = -7.93e^{-2} \)).
Table 8. Comparison of miR-10b-5p, miR-486-5p, miR-132-3p, and miR-363-3p levels in plasma between control and HD patients. miR-10b-5p and miR-486-5p are significantly increased in HD blood in same direction as changes observed in HD brain. miR-132-5p and miR-363-3p were not altered in HD blood.

<table>
<thead>
<tr>
<th>miRNA</th>
<th>Bonferroni corrected one-sided q-value</th>
<th>-ΔΔC\textsubscript{T} beta</th>
<th>Direction of effect blood to brain</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-10b-5p</td>
<td>0.0068</td>
<td>2.39E+00</td>
<td>+ / +</td>
</tr>
<tr>
<td>miR-486-5p</td>
<td>0.044</td>
<td>1.44E+00</td>
<td>+ / +</td>
</tr>
<tr>
<td>miR-132-3p</td>
<td>0.92</td>
<td>-6.22E-01</td>
<td>- / -</td>
</tr>
<tr>
<td>miR-363-3p</td>
<td>0.2</td>
<td>-7.93E-02</td>
<td>- / +</td>
</tr>
</tbody>
</table>

One of the four candidate miRNAs, miR-10b-5p, also showed significant association to age at motor onset and to disease stage in the brain study. Therefore, we focused next on miR-10b-5p and its relationship to HD symptomology. We did not see consistency in the relationship between miR-10b-5p and onset age in blood (one-sided p= 0.99, positive effect) compared to that observed in HD brain (negative effect) although a positive association of miR-10b-5p levels in blood to onset age, opposite to that seen in brain, may exist (two-sided p=3.20e-3, beta=0.13).

Although there was no statistical difference between asymptomatic HD patients and controls in miR-10b-5p levels, (one-sided p= 0.24, \( \beta = 0.53 \)), miR-10b-5p levels were significantly elevated from asymptomatic gene carriers to HD patients (one-sided p= 0.049, \( \beta = 1.15 \)), and on average, miR-10b-5p levels were lowest in control subjects (mean=0.00), higher in asymptomatic HD gene carriers.
(mean=0.75), and highest in HD patients (mean=2.38). Using a cumulative logistic regression to test the ordinal relationship of miRNA levels between controls, asymptomatic HD gene carriers and manifest HD patients found a significant positive association with miR-10b-5p levels (one-sided p= 2.05e-4, cumulative OR =2.27) (see Figure 15), concordant with the direction of the effect observed in HD prefrontal cortex.

Figure 15. miR-10b-5p has an ordinal association with HD stage. Boxplot of miR-10b-5p levels (-ΔΔCt) for control (in white), asymptomatic HD gene carrier (in grey), and manifest HD (in black) subjects. miR-10b-5p had a significant, upward, ordinal association to disease stage (one-sided p= 2.05e-4, cumulative OR =2.27). Regression analysis revealed differences between asymptomatic and HD patients (one-sided p= 0.049, β= 1.15), and differences between controls and HD (p= 3.39e-3, Bonferroni-corrected one-sided p= 6.77e-3, β= 2.39), but not difference between controls and asymptomatic HD patients (one-sided p= 0.24, β= 0.53).
Discussion

The results from this study confirm that two candidate miRNAs, miR-10b-5p and miR-486-5p, are increased in both brain and blood in HD. Levels of miR-10b-5p were significant elevated HD patients compared to asymptomatic HD gene carriers, and while the miR-10b-5p were not significantly different in asymptomatic subjects compared to controls, the ordinal relationship of controls, asymptomatic HD and manifest HD suggest premanifest HD miRNAs changes may exist.
In a previous study, miR-34b was observed significantly increased in both an HD cell model and in plasma from premanifest HD patient [31]. However, miR-34b alterations were not detected in early or late-stage HD and we did not observe a relationship of this miRNA to HD in brain, so it is difficult to interpret the utility of this finding [20].

While we believe our study is beneficial for clinical biomarker discovery, but because of study design limitations, it should be interpreted as a proof-of-concept experiment. We believe increasing our sample size will provide the statistical power for detection of asymptomatic HD gene carriers and controls miR-10b-5p differences. More importantly, our study requires an independent validation. Research of the DNA damage marker, 8-OHdG, in blood, and its relationship to HD progression [13, 32] highlights the importance of biomarker validation, as recent studies have yet to reproduce these results [32].

Commonly, the inherit variation of data from human samples, coupled with the lack of standardization in biological procedures, results in failed cross-validation. To minimize technical inconsistencies due to the vanishingly small RNA quantity in plasma [28], RNA can be directly measured using with RNA extraction-free protocols. Firefly Bioworks, which allows for the multiplex analysis of up to 68 miRNAs, as well as recent advancements in extraction-free, low input, library preparation protocols for miRNA-sequencing, may aid in the orthogonal validation of potential miRNA biomarkers. Moreover, these assays allow for larger input normalization panels (as opposed to miR-451a/SNORD44), as well
as the measurement of miRNA alterations that were detected in HD prefrontal cortex that could not be assessed in this study [20].

A longitudinal study of miR-10b-5p is critical for assessing the relationship of miRNA levels to disease progression within the same subjects, as evaluated by motor and cognitive sections of United Huntington’s Disease Ratings Scale (UHDRS). The relationship of miRNA levels to functional ratings may be potentially confounded by CAG repeat length, as we observed a relationship of CAG repeat length to miR-10b-5p levels in brain. Therefore, to establish accurate prediction of progressive based on miRNA levels, statistical correction of the CAG length should be applied.

Blood-based biochemical assays are minimally invasive and relatively simple compared to neuroimaging-based diagnostics. Here, we demonstrate the strength of this study and overall approach, built from alterations observed in HD cortex. The miRNA differences in HD blood may resemble changes observed in HD brain, which provides greater feasibility that these alterations are representative of the disease process, rather than physiological response to HD-associated weight loss, which is independent to neurodegeneration and other disease features [33].

References


CHAPTER 5: DISTINCT MICRORNA ALTERATIONS OBSERVED IN PARKINSON’S DISEASE PREFRONTAL CORTEX

Citation: Hoss AG, Labadorf A, Beach TG, Latourelle JC, Myers R. microRNA profiles in Parkinson’s disease prefrontal cortex. Frontiers in Aging Neuroscience. In submission.

Abstract

Objective. The goal of this study was to examine the microRNA (miRNA) profile of Parkinson’s disease (PD) frontal cortex as compared to normal control brain, allowing for the identification of PD specific signatures as well as the study of disease-related phenotypes, such as onset age or dementia.

Methods. Small RNA sequence analysis was performed from prefrontal cortex for 29 PD samples and 33 control samples. After sample QC, normalization and batch correction, linear regression was used to identify miRNAs altered in PD, and a PD classifier was developed using weighted voting class prediction. The relationship of miRNA levels to onset age and PD with dementia (PDD) was also characterized in case-only analyses.

Results. 125 miRNAs were differentially expressed in PD at a genome-wide level of significance (FDR q<0.05). A set of 29 miRNAs classified PD from non-diseased brain (93.9% specificity, 96.6% sensitivity). The majority of differentially expressed miRNAs (105/125) showed an ordinal relationship from control, to PD without dementia (PDN), to PDD. Among PD brains, 36 miRNAs classified PDD from PDN (sensitivity =81.2%, specificity =88.9%). Among
differential expressed miRNAs, miR-10b-5p had a positive association with onset age (q=4.7e-2).

**Conclusions.** Based on cortical miRNA levels, PD brains were accurately classified from non-diseased brains. Additionally, the PDD miRNA profile exhibited a more severe pattern of alteration among those differentially expressed in PD. To evaluate the clinical utility of miRNAs as potential clinical biomarkers, brain-related miRNA alterations, in particular, miR-10b-5p, must be further characterized and tested in peripheral biofluids.

**Introduction**

Parkinson’s disease (PD) is a progressive movement disorder, characterized clinically by resting tremor, rigidity, bradykinesia and postural instability [1]. Motor symptoms are accompanied by the loss of dopamine-producing neurons in the substantia nigra pars compacta, and associated with widespread deposition of cytoplasmic protein inclusions, largely composed of insoluble α-synuclein, throughout the brain [2].

PD subtypes can be separated based on distinct, clinical phenotypes. Approximately one-third of patients experience dementia (PDD), which has significant ramifications for quality of life and burden of care [3, 4]. Additionally, there is wide variation in the age of motor onset (ranging from age 20 to age 90) [5], with young-onset occurring before age 50 and representing 5-10% of PD cases [6]. While neuropathological hallmarks contribute to the degeneration of the nigrostriatal dopaminergic system, the etiology, clinical heterogeneity and
fundamental pathological mechanisms by which preferential neuronal loss occurs in PD are largely unknown.

The microRNA (miRNA) profile of diseased brains may offer insight into the molecular and pathological mechanisms that occur in PD. miRNAs are short, noncoding RNAs that inhibit translation through sequence-specific binding of the 3’-untranslated region (3’UTR) of target messenger RNAs (mRNAs). miRNAs can transcriptionally regulate a set or multiple sets of genes simultaneously, and in the brain, their regulatory effects have profound effects on neuronal development, differentiation and maturation [7]. Deregulation of miRNAs has been implicated in neurodegenerative diseases [8], and several studies suggest miRNAs may impact PD pathogenesis [9, 10].

In this study, we performed genome-wide miRNA sequencing of 29 PD cases and 33 neuropathologically-normal controls from prefrontal cortex. As opposed to microarray technologies that can only detect specific sequences that have been chosen a priori, high-throughput sequencing provides a relatively unbiased quantification of miRNA molecule abundance. We compared miRNA levels between PD and controls and based on their miRNA profile, classified PD from non-diseased brains. Among the PD brains, we identified miRNAs associated with the presence or absence of dementia and age of motor onset, as well as identified a set of miRNAs altered in both PD and Huntington’s disease, which may be relevant to the pathological processes that occur in the PD brain.
Methods

Sample information and small RNA sequencing. Frozen brain tissue from prefrontal cortex Brodmann Area 9 (BA9) for 29 PD samples and 33 control samples was obtained from the Banner Sun Health Research Institute, Sun City, Arizona [11], Harvard Brain and Tissue Resource Center McLean Hospital Belmont MA, and Human Brain and Spinal Fluid Resource Center VA West Los Angeles Healthcare Center, Los Angeles, CA. PD samples had no evidence of Alzheimer’s disease comorbidity based on neuropathological examination. Sample information can be found in Table S1 and summarized in Table 1. The medical charts of all 29 PD samples were reviewed to obtain information regarding clinical diagnoses of dementia (PDD n=11, no evidence of dementia (PDN) =18). 21 subjects had information on the age of onset of motor symptoms. All samples were male. Differences in covariates were tested assuming unequal variance. No difference in postmortem interval (PMI) (p-value =0.10) or RNA integrity number (p-value = 0.08) was observed between PD and controls. PD and controls differed in ages at death (PD mean age = 77.6, control mean age = 68.1; p-value = 3.2e-3). Potential confounding by age at death was assessed in subsequent analyses. PDN and PDD samples did not differ in PMI (p-value =0.62), RNA integrity number (p-value =0.27), age at onset (p-value =0.24), duration (p-value =0.44), or age at death (p-value =0.28).

Total RNA was isolated using QIAzol Lysis Reagent and purified using miRNeasy MinElute Cleanup columns. Samples were prepared using Illumina’s
TruSeq Small RNA Sample Prep Kit, according to the manufacturer’s protocol, and sequenced on Illumina’s HiSeq 2000 system with 1x51nt single-end reads at Tufts University and the Michigan State sequencing core facility.

**Table 9. Summarized sample information.**
Ages at death were significantly different between controls and PD, as indicated (\*\* \( p < 0.01 \) compared to controls). No differences between PDN and PDD were observed.

<table>
<thead>
<tr>
<th>Type</th>
<th>N</th>
<th>Motor Onset</th>
<th>Disease duration</th>
<th>Age at death</th>
<th>Post mortem interval</th>
<th>RNA integrity number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>33</td>
<td>68.1 ± 14.8</td>
<td>15.0 ± 8.7</td>
<td>7.6 ± 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All diagnosed PD</td>
<td>29</td>
<td>66.5 ± 9.8</td>
<td>10.5 ± 6.5</td>
<td>77.6 ± 9.0 **</td>
<td>11.1 ± 9.7</td>
<td>7.3 ± 0.7</td>
</tr>
<tr>
<td>PD, Non demented</td>
<td>18</td>
<td>64.1 ± 7.2</td>
<td>11.5 ± 6.4</td>
<td>76.1 ± 8.9</td>
<td>11.9 ± 9.2</td>
<td>7.2 ± 0.8</td>
</tr>
<tr>
<td>PD with dementia</td>
<td>11</td>
<td>69.8 ± 12.2</td>
<td>9.2 ± 6.7</td>
<td>79.9 ± 9.0</td>
<td>9.9 ± 10.9</td>
<td>7.5 ± 0.5</td>
</tr>
</tbody>
</table>

**Statistical analysis.** Reads were processed and counted as described in Hoss et al. 2015 [12]. R version 3.1.0, Bioconductor version 2.1.4, and DESeq2 version 1.40.0 were used for variance stabilizing transformation (VST) of count data. Batch correction was applied using ComBat using sva 3.10 [13] and LIMMA version 3.20.8 [14] was used for differential expression analysis of PD cases and controls. All PD samples were from a single batch, therefore DESeq2 normalized, VST counts from PD samples without batch correction were used for the PD-only analyses relating miRNA levels to clinical features. Differential expression analysis was performed with and without adjustment for age at death. After multiple comparisons correction using a false discovery rate [15], FDR-adjusted \( q \)-values<0.05 were reported as significant.
To further evaluate the differential expression patterns between PD cases and controls, unsupervised, Ward hierarchical clustering by Euclidian distance was applied using the heatmap2 function in the gplots R package [16].

Supervised, predictive modeling of case status was performed using the GenePatterns WeightedXVoting module using 29 PD-associated miRNAs with large effects [17].

To determine if miRNAs were associated with PD in the presence or absence of dementia, VST counts for PDN and PDD were compared using LIMMA, adjusting for age at death, and p-values were FDR-adjusted for the number of comparisons. miRNAs nominally associated with PDD were used to classify PD patients with and without dementia, as described above, using the weighted voting method. To determine beta estimates relative to control samples, PDN and PDD were separately compared to controls using LIMMA. In addition, cumulative logit models using the “ordinal” package in R [18], were applied to test whether an ordinal relationship existed across control, PDN and PDD samples. FDR-corrected Chisq likelihood-ratio tests were used to determine significant miRNA associations.

Linear models were used to model the relationship between age of motor onset and miRNA levels among the 21 PD samples with onset data. Tests were performed genome-wide and exclusively among the set of differentially expressed miRNAs. FDR-adjusted q-values<0.05 were reported. Models were run with and without adjustment for age at death.
Finally, to determine whether overlap in miRNA alterations exist between PD and HD brain, results from PD differential expression analysis were compared to those of our previously published Huntington’s disease (HD) study [12], which contain the same control brains. HD data was accessed from NCBI’s Gene Expression Omnibus, series accession number GSE64977 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE64977) and analyzed using the same bioinformatics approach for differential expression analysis as described above (Bioconductor version 2.1.4, DESeq2 version 1.40.0, ComBat sva 3.10, LIMMA version 3.20.8).

Results

miRNA levels are altered in Parkinson’s disease compared to non-disease brains. To identify miRNA differences in PD as compared to non-disease subjects, miRNA sequence analysis was performed in prefrontal cortex (Brodmann Area 9) for 33 controls and 29 idiopathic PD samples (see Table 1). Results of differential expression analyses, correcting for sequencing batch effects and with and without adjustment for age at death are shown in Table S2. 125 miRNAs were significantly altered in PD after adjusting for age at death (FDR $q$-value<0.05, see Table S2). Unadjusted results were similar, but confounding by age of death was observed for some miRNA, so adjusted results are reported here. Most miRNA alterations were moderate, with 77% of the differentially abundant miRNAs (96/125) within a ±0.5 log fold change (LFC).
The levels of 64 miRNAs were down-regulated, whereas the levels of 61 miRNAs were up-regulated in PD relative to controls.

Next, we used classification models to investigate whether the levels of PD-related miRNAs in brain could accurately assign disease status. To select the most informative miRNAs, we filtered on effect size (LFC>0.5 or LFC< -0.5). After filtering, 29 PD-related miRNAs were used in an unsupervised hierarchical cluster analysis. Samples clustered based on disease status with the exception of five PD which clustered with the controls, (see Figure 16A). To further assess whether miRNA levels could differentiate PD and control samples, disease status was predicted using a weighted voting classification with leave-one-out cross-validation. This model is tested by iteratively leaving one sample out, creating a training model by assigning a weighted linear combination based on the levels of the 29 miRNAs, and testing this model on the left out sample. Here, only three errors were observed using 29 miRNAs (two Type I errors, one Type II error), with 93.9% specificity, 96.6% sensitivity and an absolute error rate of 4.8% (see Figure 16B). Both Type I errors were called with low confidence (9.8-10.1%).

**miRNA alterations in Parkinson’s disease with dementia.** To assess whether miRNA differences specific to the PDD subtype were distinguishable from a generalized PD response, we performed a differential expression analysis comparing PDN to PDD using normalized VST count data from 18 PDN and 11 PDD samples. We observed no genome-wide significant (q<0.05) miRNAs associated with dementia in PD, with or without adjustment for age or disease
duration, after multiple correction testing (see Table S3). Even when limiting to the 125 differential expressed miRNAs we saw no significant differences between PDN and PDD. We however noted stronger directions of effect in PDD when separately comparing PDN versus control and PDD versus control, suggesting PDD may represent a more severe version of the PD miRNA profile spectrum.

**Figure 16. miRNA changes related to Parkinson's disease.**

A. Heatmap of 29 miRNAs differentially expressed between PD and control prefrontal cortex samples with log fold changes (LFC) greater than 0.5 or less than -0.5. Scaled level values are color-coded according to the legend on the right. The dendrogram on the left depicts hierarchical clustering based on level. The top dendrogram depicts clustering based on the miRNA signal from each sample. The top bar indicates disease status (blue: control, yellow: PD, non-demented (PDN), orange: PD with dementia (PDD)). B. Disease prediction using 29 miRNAs. Scores less than zero were called as non-diseased whereas scores above zero were called as PD. Blue circles = control, yellow boxes = PD, orange triangles = PDN.
To test whether PDD had increased miRNA alterations in comparison to PDN, we created an ordered categorical variable (controls, PDN, and PDD) and tested the association of this variable to genome-wide miRNA levels. 105 of the 125 differentially expressed miRNAs had a significant ordinal association (q<0.05) (see Table S4, and Figure 2A), indicating that in the majority of differentially expressed miRNAs in PD, PDD samples exhibit larger differences than PDN cases as compared to controls for the same miRNAs.

**Figure 17. miRNA profile for Parkinson's disease with dementia.**
A. Line plot for the 125 DE miRNAs in PD. The counts were scaled using Z-transformation, and the means were calculated for each miRNA for each condition (control, PDN, PDD). The line colors correspond to the beta estimates from the ordinal regression analysis, where blue=negative ordered relationship and red=positive ordered relationship. B. PDN/PDD class prediction, using the 36 nominally significant miRNAs from PDN/PDD differential expression analysis. Four errors were observed among the 29 samples studied.

We further investigated the clinical utility of these miRNA profiles for the assessment of dementia using classification analyses (WeightedXVoting) [17].
The 36 nominally significant miRNAs (p<0.05) identified in the PDN/PDD comparison from LIMMA were used to classify disease state (See Figure 2B), though with more limited accuracy than the PD-control model (absolute error rate =13.8%, sensitivity =81.2%, specificity =88.9%). Four miRNA features overlap between the control/PD and PDN/PDD models (miR-132-3p, hsa-miR-132-5p, hsa-miR-145-5p, hsa-miR-212-5p).

**miR-10b-5p levels are associated with the onset of motor symptoms in both Parkinson’s and Huntington’s disease.** To understand whether deregulated miRNAs were specific to PD, or a general response to the neurodegenerative process, we compared miRNA that were significantly altered in PD to those significantly altered in Huntington’s disease (HD). 21 miRNAs were found differentially expressed in both PD and HD experiments, and of these, only two miRNAs had opposite directions of effect between diseases (miR-10b-5p, miR-320b).

Within the PD case sample, we tested the association of age of motor onset of PD with miRNA levels. Although we did not observe significance in a genome-wide analysis, restricting our study to the 125 significantly differentially abundant PD miRNAs revealed miR-10b-5p to have a significant, positive association to onset age (beta=0.040, q-value=4.7e-2, model $r^2=0.49$, see Figure 3A). Adding death age increased the magnitude of the effect estimate (death adjusted age of onset beta=0.049, p-value onset =3.2e-3, p-value death =0.40), although did not stand up to multiple comparisons corrections (q-value
onset =0.19). While miR-10b-5p is significantly decreased in PD, miR-10b-5p was observed in our previous HD cortical miRNA study [19] to be massively increased in HD in comparison to controls (see Figure 3B). Intriguingly, PD and HD also exhibit opposite effects with regard to onset age, where miR-10b-5p has a strong, negative relationship to age of onset in HD ($r^2=0.64$; $r^2=0.39$ after accounting for the contribution of HD gene repeat length) and a strong, positive effect in PD (see Figure 3C).

**Figure 18.** miR-10b-5p levels are associated with motor onset age in both Parkinson's and Huntington's disease. 

A. Comparison of miR-10b-5p level in PD and HD. *p<0.05, **p<0.01, ***p<0.001, p-values adjusted for genome-wide comparisons. B. Scatterplot of miR-10b-5p levels to motor onset age in Parkinson’s disease (PD). In PD, miR-10b-5p levels exhibit a positive association to onset. C. Scatterplot of miR-10b-5p levels to motor onset age in Huntington’s disease (HD). In HD, miR-10b-5p levels exhibit a negative association to onset.
Discussion

*Parkinson’s disease related miRNAs.* In this study, we identified 125 miRNAs altered at genome-wide levels in PD prefrontal cortex using next-generation sequencing. This is the largest miRNA sequencing analysis performed in PD versus control brain samples (29 versus 33 respectively), the first to provide a detailed miRNA PD profile, classify brains by miRNA levels and to evaluate the relationship of miRNA levels in brain to relevant clinical features.

Reduced levels of miR-133b [9], miR-34b and miR-34c [20], and elevated levels of autophagy-related miRNAs, were previously reported [21], and while these miRNAs were detectable in our study, we did not observe significant changes in their levels. These discrepancies are likely a consequence of the different brain regions that were studied (midbrain versus prefrontal cortex), and the assay technologies that were used to profile miRNA levels (reverse transcriptase quantitative PCR [9, 21], and microarray [20] versus miRNA-sequencing).

Several miRNAs that we report altered in PD brain may interact with PD-related genes. Monogenic forms of PD include mutations within the α-synuclein gene (*SNCA*), Leucine-rich Repeat Kinase 2 (*LRRK2*), one of the most common causes of familial PD [22] and glucocerebrosidase (*GBA*). While we did not observe alterations of *SNCA*-targeting miRNAs, miR-7 and miR-153 [23, 24], two miRNAs shown be regulated by *LRRK2* (let-7i-3p/5p and miR-184 [10]) and one
miRNA experimentally shown to target LRRK2 expression (miR-1224 [25]), were observed to be down-regulated in PD.

Glucocerebrosidase (GBA) deficiency is associated with PD [26]. We observed miR-127-5p, which has been shown to reduce GBA activity [27], to be down-regulated in PD brains, and miR-16-5p which has been shown to correspond to enhanced GBA protein levels [27], was found to up-regulated in brain in our study. It is noteworthy to observe LRRK2-related miRNAs, as none of the PD brains in our study had LRRK2 mutations. This may support a role of LRRK2 and GBA in PD, independent of that produced by the known mutations in these genes.

Classification based on miRNA abundance. We were able to classify PD based on the levels of 29 miRNAs with less than a 5% error rate. Although this classification was performed using postmortem brain samples, we believe this may be relevant for PD biomarker discovery, particularly if these miRNAs are peripherally detectable. We reasonably differentiated PD subtypes (PDN/PDD) based on miRNA levels, and we observed a pattern of increased changes in the PDD samples relative to the PDN samples in the set of altered miRNAs. We observed that the majority of differentially expressed miRNAs had an ordinal relationship to controls and PD cases stratified by the presence or absence of dementia, suggesting PDD may represent a more severe alteration of the PD miRNA profile.
**PD-related miRNA changes in biofluids.** Our study in PD brains identified profiles of miRNAs that distinguish PD from controls, which if also observed in peripheral biofluids, such as blood or cerebrospinal fluid (CSF), could be valuable in the evaluation of PD diagnosis, prognosis, or progression. The small size (~22 nucleotides) of miRNA may allow for neuropathologically altered miRNA to cross the blood-brain barrier in exosomes [28] and circulate stably in peripheral fluids as cell-free molecules [29]. Although there was no overlap of miRNAs in brain to changes observed in most PD blood studies [30-34], we did observe increased levels in one (miR-29a-3p) of three miRNAs previously reported as increased in blood of PD patients after Levodopa treatment [35]. In Burgos et al. 2014, small RNA sequencing was performed for blood serum and CSF from 67 PD and 78 control subjects [36], five miRNAs were found significantly altered in PD serum and 17 were significantly altered in CSF. Of these 22 miRNAs, five showed consistent overlap with our cortical findings, with one from serum (miR-1294) and four from CSF (miR-132-5p, miR-127-3p, miR-212-3p, miR-1224-5p). Thus, miRNAs detected in CSF may have a stronger relationship with brain miRNAs levels than those detected in serum.

**Common miRNA changes in PD and HD.** We previously reported 75 miRNAs altered in HD prefrontal cortex [12], and when comparing miRNAs altered in both PD and HD, 21 miRNAs were observed deregulated in both diseases. Among the miRNAs with concordant changes in PD and HD, a several miRNAs correlated with various HD clinical features, such as the extent of striatal
degeneration and duration of the disease [12]. These miRNAs may represent a generalized, neurodegenerative response in the prefrontal cortex, which relate to severity and/or progression across these diseases. However, it is important to recognize that the results from the PD and HD studies are not independent, as they were both analyzed using the same 33 control brains.

Of the two miRNAs with discordance between PD and HD, miR-10b-5p emerged due to its relationship to onset in both diseases. miR-10b-5p is markedly increased in HD in comparison to controls, and has a negative association to age of onset for HD, with higher levels of miR-10b-5p corresponding to early onset age. In contrast, miR-10b-5p is significantly decreased in PD, and has a positive association to onset age, where higher miRNA levels correspond to later onset ages. In a separate Alzheimer’s disease (AD) study, examining miRNA levels in prefrontal cortex, miR-10b-5p levels were significantly reduced in AD [37]. However, at the very earliest stages of AD, miR-10b-5p levels clustered with up-regulated miRNAs whereas at early to middle stages, miR-10b-5p levels appeared to decline [37]. The relationship of miR-10b-5p to these three age-related, neurodegenerative diseases suggests a complicated pattern of miR-10b-5p alteration in response to the neurodegenerative or pathologic protein aggregation processes.

**Conclusion**

This study provides evidence that miRNA levels are altered in PD prefrontal cortex. These changes are sufficiently consistent that diseased brains
can be discriminated with high confidence from non-diseased brains based on the level of 29 miRNAs. PDD may represent a more severe profile of PD related miRNAs than PDN. 21 miRNAs changes were similar between PD and HD, with the exception of miR-10b-5p, which had opposite direction of effects to disease association and to motor onset age in the two diseases. Further characterization of miR-10b-5p in the neurodegenerative disease context is warranted to better understand if it has clinical potential as a biomarker for disease progress or to identify potential therapeutic targets.

References


Conclusion

Summary of Aims

miRNAs are altered in HD brain, several of which relate to clinical phenotypes, such as onset age and striatal degeneration. This effect is CAG-independent, and likely independent of both death age and cortical involvement. In particular, miR-10b-5p levels have the strongest associations to these features, where increased levels in HD brain predict earlier onset age and more striatal involvement. Levels of miR-10b-5p are increased in asymptomatic and grade 2 brains, indicating the levels of this miRNA are increased pre-clinically or within the early stages of the disease process.

In plasma, we found two miRNAs (miR-10b-5p and miR-486-5p) that were significantly altered. These same miRNAs exhibited nominal association (p<0.05) to CAG-adjusted onset (miR-10b-5p and miR-486-5p) and to CAG-adjusted striatal involvement (miR-10b-5p) in the HD cortical study. Levels of both miRNAs were elevated in both HD plasma and brain. miR-10b-5p levels were observed to be slightly elevated in asymptomatic HD and significantly increased in manifest HD compared to healthy controls, which suggests miR-10b-5p changes occur in HD, and may increase in plasma during the transition from premotor to manifest HD.

We identified significant differences in miRNA levels in PD brain and based on a set of deregulated miRNAs, we were able to classify PD and non-diseased brains with high accuracy. The miRNA profile of PDD suggests PDD
has a more extreme molecular phenotype than PDN. A comparison with our previous Huntington’s disease (HD) study, which used the same set of control brains, revealed a set of miRNAs altered in both PD and HD. When testing for the association of differentially expressed miRNAs to age at onset, miR-10b-5p had a significant, positive relationship. In contrast to HD, miR-10b-5p was slightly decreased in PD and increased levels of miR-10b-5p were observed in individuals with late-onset.

**Overall conclusion**

These studies provide a detailed miRNA profile for HD and PD brain, identify miRNAs associated with disease pathology and suggest miRNA changes observed in brain can be detected in blood. Together, these findings support the potential of miRNA biomarkers for the diagnosis and assessment of progression for neurodegenerative diseases.
APPENDIX

Table S2.1. miRNA RT-qPCR validation study results.

Table S2.2. Sample information for eight Huntington’s disease brains used for RT-qPCR replication study.

Table S2.3. Sample information for eight control brains used for RT-qPCR replication study.

Table S2.4. miRNA RT-qPCR replication study results.

Table S2.5. Sample information for fourteen Parkinson’s disease brains used for RT-qPCR replication study.

Table S2.6. Read statistics for miRNA-sequence analysis.

Table S2.7. Mean and standard deviation inner-distance estimates for TopHat2 alignment.

Table S2.8. Read statistics for mRNA-sequence analysis.

Table S3.1. Sample information for 36 control brains used for miRNA-sequencing.

Table S3.2. Sample information for 28 Huntington’s disease brains used for miRNA-sequencing.

Table S3.3. Read statistics for miRNA-sequence analysis.

Table S3.4. Correlation of differentially expressed miRNA precursor pairs.

Table S3.5. Summary statistics for Firefly BioWorks assay.

Table S3.6. Linear regression analysis modeled the relationship of miRNA expression and Vonsattel grade.

Table S5.1. Detailed sample information.
Missing information is designated as “NA”.

Table S5.2. Results from miRNA differential expression analysis in Parkinson’s disease.


60. Dexter DT, Jenner P: **Parkinson disease: from pathology to molecular disease mechanisms.** *Free radical biology & medicine* 2013, **62**:132-144.


71. Harris GJ, Pearlson GD, Peyser CE, Aylward EH, Roberts J, Barta PE, Chase GA, Folstein SE: **Putamen volume reduction on magnetic resonance**


CURRICULUM VITAE

ANDREW HOSS

PERSONAL
Birth year: 1984
Home address: 203-205 Elm Street #3, Cambridge, MA 02139
Phone: 617-913-2875
Email: andrewghoss@gmail.com

EDUCATION
Ph.D. Genetics & Genomics, Boston University School of Medicine, Boston, MA, June 2015 (expected)
M.S. Biology, Tufts University, 2011
B.A. Biological Sciences & Psychology, Binghamton University, 2006

RESEARCH EXPERIENCE
BOSTON UNIVERSITY, DEPARTMENT OF NEUROLOGY, Boston, MA

Project #1: Neurodegenerative disease transcriptomic profiling
• Utilized small RNA and large RNA sequence analysis from human brain to identify biomarkers of Huntington’s and Parkinson’s disease progression
• Extensive background in bioinformatics analysis (Tuxedo, bedtools, samtools, fastx toolkit, novoalign, miRDeep, R/Bioconductor) and data interpretation (differential expression profiling, multivariate statistics, predictive modeling, regression analysis)
• Independently established a funded collaboration with Firefly Bioworks to develop their novel flow cytometer-based miRNA detection platform
• Created in vitro Huntington’s disease models and custom image analysis scripts using ImageJ to assess cell death and the extent of aggregate formation for functional validation of miRNAs

Project #2: Characterization of cyclin-g associated kinase (GAK) in Parkinson’s disease
• Evaluated the GAK locus using targeted sequencing in 425 familial Parkinson’s disease cases to identify the functional variants ultimately responsible for the association to the disease

WORK EXPERIENCE
BROAD INSTITUTE OF MIT AND HARVARD, Cambridge, MA

Process Development Associate I/II, Special Projects Group, 2009 – 2011

- Led R&D and process improvement efforts in Ion Torrent/Life Technologies collaboration
- First external user of Ion Torrent, performing over 1000 sequencing runs prior to commercial release
- Helped patent on-bead paired-end sequencing by reducing the invention to practice
- Maintained platform infrastructure by managing protocols, databases, reagent supply chains, software and machine upgrades
- Frequently applied molecular biology techniques and various nucleic acid quantification methods for sequencing chemistry optimization


- Participated in the evaluation, implementation and scale-up of 454 FLX/Titanium technologies for numerous long-term initiatives including the Human Microbiome Project
- Expert knowledge of NGS workflows and data analysis for 454, Helicos, Illumina and SOLiD platforms
- Scripted liquid handling robots and engineered several plate-based modules to streamline the 454 sequencing process

Process Technician Assistant, Core Sequencing, 2007

- Operated automated laboratory equipment to clone, amplify, purify, and prepare DNA libraries for cycle sequencing reactions
- Experienced in cloning, transformation, colony plating and picking techniques
- Prepared quality control validations and maintained LIMS records
- Executed process development projects to improve ultra-high throughput Sanger sequencing

OTHER EXPERIENCE

- LEANGEN PARTNERS, Cambridge, MA, Consultant, 2011
- CORNELL UNIVERSITY, Ithaca, NY, Field Technician, Dr. Anne Clark Laboratory, 2006
- BINGHAMTON UNIVERSITY, Teaching Assistant, Organic Chemistry I & II, 3 semesters, 2006
- BINGHAMTON UNIVERSITY, Undergraduate Assistant, Fetal Alcohol Laboratory, 2004
CERTIFICATIONS

• Six Sigma Green Belt (final project to improve sequence quality saved $136,000/year)
• JMP ANOVA
• SAS I & II

PRESENTATIONS AND AWARDS

• FIREFLY FRONTIERS GRANT, Firefly Bioworks, 2013
• BUSM Evans Day, Poster Award, 2013
• BUSM Genome Science Institute, Poster Award, 2014
• BUSM Genome Science Institute, Poster Award, 2015
• Society for Neuroscience, Poster Presentation, 2013
• Huntington’s Study Group, Poster Selection, 2014
• Advances in Genome Biology and Technology Meeting, Poster Selection, 2011
• Advances in Genome Biology and Technology Meeting, Poster Selection, 2008

PUBLICATIONS


Hoss, AG, Labadorf A, Hadzi TC, Beach TG, Myers RH, Latourelle JC. Distinct microRNA alterations observed in Parkinson’s disease prefrontal cortex. Frontiers of Aging Neuroscience. In Submission.

PATENTS