A myriad of miRNA variants in control and Huntington’s disease brain regions detected by massively parallel sequencing

Eulàlia Martí1,2,*, Lorena Pantano1,2, Mónica Bañez-Coronel1,2, Franc Llorens3,4, Elena Miñones-Moyano1,2, Sílvia Porta5, Lauro Sumoy6, Isidre Ferrer4,5 and Xavier Estivill1,2,7,*

1Genetic Causes of Disease Group, Genes and Disease Program, Centre for Genomic Regulation (CRG), 2Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), 3Molecular and Cellular Neurobiotechnology, Institute of Bioengineering of Catalonia and Department of Cell Biology, University of Barcelona, 4Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), 5Institut Neuropatologia, Servei Anatomia Patològica, IDIBELL-Hospital Universitari de Bellvitge, Universitat de Barcelona, Barcelona, Catalonia, 6Institute of Predictive and Personalized Medicine of Cancer, Badalona and 7Experimental and Health Sciences Department, Pompeu Fabra University, Barcelona, Catalonia, Spain

Received February 19, 2010; Revised June 8, 2010; Accepted June 9, 2010

ABSTRACT

Huntington disease (HD) is a neurodegenerative disorder that predominantly affects neurons of the forebrain. We have applied the Illumina massively parallel sequencing to deeply analyze the small RNA populations of two different forebrain areas, the frontal cortex (FC) and the striatum (ST) of healthy individuals and individuals with HD. More than 80% of the small-RNAs were annotated as microRNAs (miRNAs) in all samples. Deep sequencing revealed length and sequence heterogeneity (IsomiRs) for the vast majority of miRNAs. Around 80–90% of the miRNAs presented modifications in the 3′-terminus mainly in the form of trimming and/or as nucleotide addition variants, while the 5′-terminus of the miRNAs was specially protected from changes. Expression profiling showed strong miRNA and isomiR expression deregulation in HD, most being common to both FC and ST. The analysis of the upstream regulatory regions in co-regulated miRNAs suggests a role for RE1-Silencing Transcription Factor (REST) and P53 in miRNAs downregulation in HD. The putative targets of deregulated miRNAs and seed-region IsomiRs strongly suggest that their altered expression contributes to the aberrant gene expression in HD. Our results show that miRNA variability is a ubiquitous phenomenon in the adult human brain, which may influence gene expression in physiological and pathological conditions.

INTRODUCTION

Huntington’s disease (HD) is an autosomal dominant neurodegenerative disease caused by a CAG repeat expansion in the HTT (huntingtin) gene that encodes HTT. Although HTT is ubiquitously expressed, patients with HD show predominantly central nervous system (CNS) manifestations, including abnormal motor symptoms, cognitive decline and psychiatric disturbances. HD is characterized by widespread mRNA misregulation, especially in the striatum and cortical regions (1). Such deregulation is the result, in part, of aberrant nuclear localization of the transcriptional repressor RE1-Silencing Transcription Factor (REST) (2,3). This REST nuclear translocation leads to increased transcriptional repression of BDNF, a REST target gene, therefore compromising striatal neuronal survival. In addition to protein-coding genes, recent reports have shown that REST controls the expression of neuronal large (macroRNAs) and small [microRNAs (miRNAs)] non-coding RNAs (4,5), with various roles in gene expression modulation. Therefore, aberrant expression of non-coding RNAs may have indirect consequences in HD gene expression deregulation.
miRNAs are small non-coding RNAs that participate in the post-transcriptional regulation of gene expression through sequence complementarity to open reading frame (ORF) and 3′ untranslated regions (UTRs) of messenger RNAs (mRNAs). This binding leads to target mRNA degradation or translational inhibition. miRNAs play important roles in organogenesis, metabolism and neuronal development. Many miRNAs are selectively expressed in the brain and appear to regulate fundamental neuronal processes, including the maintenance of the neuronal transcriptome, dendrite outgrowth and synaptic plasticity (6). In addition, miRNAs have been shown to be essential for neuronal survival and alterations in miRNA function have been involved in a number of neurodegenerative processes (7,8). These alterations comprise aberrant expression of miRNAs in the brain of patients with neurodegenerative diseases (9–15) and single-nucleotide polymorphisms (SNPs) (16) or a point mutation (14) in the target 3′ UTR NA binding site.

Two main findings point to the possibility that small-RNA silencing-dependent mechanisms participate in HD neuropathology. First, increased repression by REST leads to changes in the expression of specific neuronal miRNAs in HD patients and HD mouse models (17,18). Second, HTT interacts with Argonaute proteins, which are key members of the RNA-induced silencing complex (RISC). This interaction with the mutant HTT leads to a reduced reporter gene silencing activity compared with the wild-type HTT (19). Overall, these data strongly suggest that alterations in miRNA-mediated post-transcriptional regulation could be a mechanism contributing to mRNA expression deregulation in HD.

Although alterations in miRNA expression have been reported in brain samples of HD patients, a deep characterization of miRNA profiling and sequence modifications is lacking. The available high-throughput sequencing technologies allow detection of the current reference miRNA sequences residing in the miRBase and novel miRNAs. This approach has also the potential to uncover variations in the mature miRNA length, and enzymatic modifications such as RNA editing and 3′ nucleotide additions. The Illumina sequencing platform yields several million sequences from a single flow cell lane, providing a high resolution. In the present study, we used Illumina deep sequencing to extensively profile and identify disease-specific changes of small non-coding RNAs that occur in the frontal cortex (FC) and the striatum (ST), two brain areas affected in HD. Our results show a strong deregulation of miRNA expression in HD and highlight the potential importance of the large number of modified mature miRNA sequences in the physiology and pathology of the adult brain.

METHODS

Human brain samples

Patients with HD had suffered from a familial degenerative disease characterized by involuntary muscle movement, difficulties with speech, trouble swallowing and walking and excessive blinking. This was followed by abnormal behavior and cognitive impairment, continuous falls and hypokinesia, leading to terminal akinetic-rigid disorder. Genetic studies showed increased CAG repetitions in the Htt gene in affected relatives or an increased number of CAG nucleotides in affected individuals. Neuroimaging [computed tomography (CT) or magnetic resonance imaging (MRI)] scans disclosed severe atrophy of the caudate and putamen and moderate to severe cortical atrophy during the course of the disease. Based on the clinical and genetic data, a diagnosis of HD was made during life in every case.

Brain samples corresponding to the frontal cortex (FC) and the striatum (ST) (dorsal caudate) of HD patients and controls were obtained from the Institute of Neuropathology and the University of Barcelona Brain Bank, after the informed consent of the patients or their relatives and the approval of the local ethics committee had been given (Supplementary Table S1). Cases with and without clinical neurological disease were processed in the same way, following the same sampling and staining protocols. Gross atrophy was seen in HD cases. At autopsy, half of each brain was fixed in 4% buffered formalin, whereas the other half was cut in coronal sections 1 cm thick, frozen on dry ice, and stored at −80°C until use. The postmortem delay was between 3 and 17 h. The neuropathological study was carried out on dewaxed 4-μm-thick paraffin sections of the frontal lobe (area 8), primary motor, primary sensory, parietal, temporal superior, temporal inferior, anterior cingulated, anterior insular and primary and associative visual cortices; entorhinal cortex and hippocampus; head of the caudate, medial caudate, putamen and globus pallidus; medial and posterior thalamus; subthalamus; Meynert nucleus; amygdala; midbrain (two levels);pons and medulla oblongata; and cerebellar cortex and dentate nucleus. The sections were stained with hematoxylin and eosin and Klüver–Barrera and processed for immunohistochemistry to glial fibrillary acidic protein (GFAP), CD68 and Lecquericum esculentum lectin for microglia; Aβ-amyloid; pan-tau, AT8 tau and phosphorylation-specific tau Thr181, Ser202, Ser214, Ser262, Ser396 and Ser422; zB-crystallin; z-synuclein; and ubiquitin.

The neuropathological examination in HD cases revealed severe atrophy of the caudate and putamen, the atrophy being more marked in the dorsal border, together with cerebral cortical atrophy and enlargement of the lateral ventricles. This was accompanied by marked neuronal loss and astrocytic gliosis in the caudate and putamen, globus pallidus and cerebral cortex. Individual neurons in the cerebral cortex and striatum exhibited ubiquitin-positive intranuclear inclusions typical of diseases with CAG triplet expansions and ubiquitin-positive neuritic processes in the same regions. After neuropathological examination, HD cases were categorized as stage 4 following Vonsattel classification (20,21).

RNA extraction and sequencing

Total RNA was isolated from ~10 mg of the different brain areas using the miRNeasy kit (Qiagen) that
preserves small RNAs. Equal amounts of total RNA of two equivalent patients and brain areas were mixed, giving rise to a total of four samples (C-FC, HD-FC, C-ST and HD-ST) showing RNA integrity numbers (RIN) between 6 and 7. A total of 5 μg of small-RNA-enriched total RNA was processed for Illumina sequencing. Briefly, the processing by Illumina consisted of the following successive steps (22,23): acrylamide gel purification of RNA bands corresponding to the size range 18–30 nt, ligation of 5′ and 3′ adapters to the RNA in two separate subsequent steps each followed by acrylamide gel purification, cDNA synthesis followed by acrylamide gel purification, and a final step of polymerase chain reaction (PCR) amplification to generate template library for cluster generation on a flow cell. After hybridization of the sequencing primer libraries were sequenced it using a GAII instrument (Illumina). The reads generated were 36-nt long and were analysed using SeqBuster pipeline (24).

Pre-analysis of human brain sequencing data

Pre-analysis consists in the recognition and removal of the adapter and the annotation of the sequences. Pre-analysis was performed using a stand-alone version provided in http://estivill_lab.crg.es/seqbuster (24). Out of a total of several million reads, we discarded any reads without a minimum of 10-nt linker subsequence directly adjoining the insert, showing two or less mismatches. Following the application of this filter, populations of 5 914 073; 5 561 686; 5 125 116 and 5 855 660 reads were recovered in C-FC, C-ST, HD-FC and HD-ST samples, respectively. Taking into account the unique sequences with more than three counts, the total number of reads were 5 827 761; 5 466 923; 5 041 667 and 5 759 841. Then sequences were mapped to human pre-miRNA and mature miRNA databases provided in the miRBase (http://microrna.sanger.ac.uk/sequences/, Release 14), and mRNA, ncRNA, repeats and genome databases available at (http://hgdownload.cse.ucsc.edu/goldenPath/hg18/bogZips/), using Mega BLAST. For precursor miRNA annotation the following parameters were configured (24): one mismatch, 3 nt in the 3′-addition variants and the priority degree equal to 3. For miRNA and miRNA* annotation the following parameters were configured: one mismatch, 3 nt in the 3′ or 5′-trimming variants, 3 nt in the 3′-addition variants, a priority degree to 1 and 2 for miRNA and miRNAs* databases, respectively, and the parental database was the precursor miRNA database. These options permitted annotations of the following types of alignment: (i) perfect match, where the sequence is completely identical to the reference sequence; (ii) trimming at the 3′-end of the reference miRNA sequence, which is a miRNA variant several nucleotides shorter or longer that matches to the mature or precursor reference sequence, respectively; (iii) trimming at the 5′-end of the sequence, an analogous case focused on the 5′-end of the miRNA; (iv) nucleotide additions at the 3′-end of the sequence and (v) nucleotide substitutions, showing nucleotide changes with respect to the reference sequence. The parameters for the alignment to the mRNA and genome databases allowed as much as one mismatch and up to three nucleotide additions in the 3′-terminus. The priority parameter was 4, 5, 6 and 7 for mRNA, ncRNA, repeats and genome databases, respectively (24).

Analysis of miRNA variability (IsomiRs)

For deep characterization of miRNA, variants we applied several filters. First, the sequences considered in the analysis presented a frequency above 3. Second, 10 was chosen as the ’Contribution Cut-Off’ parameter, meaning that every isomiR considered in the analysis contributes by more than a 10% to the total number of variants annotated in the same miRNA locus. Third, we applied the Z-score option to exclude sequencing errors as the possible cause of the nucleotide changes observed in some variants (24). Annotated miRNAs and IsomiRs with the corresponding counts are available at http://estivill_lab.crg.es/Marti_etal/. Four files are available containing the data set of control FC (C_FC), control ST (C_ST), HD FC (HD_FC) and HD FC (HD_ST) samples.

Novel miRNA prediction

The ‘human miRNA prediction’ package in the SeqBuster pipeline (24) uses the algorithm described in the microPred pipeline (25) that is based on a non-comparative computational method for the effective identification of pre-miRNAs among the hairpin secondary structures predicted from the human genome. A test input file has to be provided containing all sequences for which the analysis is going to be performed.

SNP analysis in miRNAs

The ‘SNP analysis’ package is offered by SeqBuster in the ‘Download’ section of the web server to be run in the local machine (24). This package contains genomic coordinates for available SNPs (http://hgdownload.cse.ucsc.edu/downloads.html). The algorithm produces a table showing all the sequences that have a nucleotide substitution event at the same position where an SNP has been detected.

Differential expression analysis

The sequencing performance in the samples to be compared was evaluated using the ‘Sequencing Capacity’ package in the ‘Basic Analysis’ module of SeqBuster with default parameters (24). In this package, the Wilcoxon test was applied to determine statistically significant differences in the frequency distribution between samples. Two types of differential expression analyses were performed, that included different types of sequences: (i) the different miRNA loci were containing all the sequences mapping to each locus; and (ii) 5′-trimming and nucleotide substitution IsomiRs affecting the seed region of the miRNA. To perform these analyses several options and filters were applied. First, the frequency was normalized to reads per million (RPM). Second, the Z-test (26) was applied to show statistical significance in the differential expression. Third, the Benjamini and Hochberg method was applied to correct the P-value
assigned by the Z-test. Fourth, in analyzing the seed region variants, 10 was chosen as the ‘Contribution Cut-Off’ value (24). Fifth, we applied the Z-score option to exclude sequencing errors as the possible cause of the nucleotide changes observed in the isomiR (27). In the output resulting from the analysis we only showed sequences with a total count contribution above 50, considered as the sum of the frequencies of the two samples being compared.

Microarray preparation and statistical analysis

Agilent Human miRNA microarrays (G4470B), containing 13,737 probes, were used to analyze independent samples from the FC and the ST of four control individuals and four HD patients (Supplementary Table S1). Briefly, 500 ng of total RNA from each sample were chemically labeled by dephosphorylating using calf intestinal alkaline phosphatase (CIP) and ligating Cyanine3-pCp by a T4-RNA ligase using Agilent miRNA Complete Labeling and Hyb Kit (p/n5190-0456). Labeled samples were dried and resuspended in 18 μl of nuclease-free water and co-hybridized with in situ hybridization buffer (Agilent) for 20 h at 55°C and washed at room temperature for 5 min in Gene Expression Wash Buffer 1 (Agilent) and 5 min at 37°C in Gene Expression Wash Buffer 2 (Agilent). The images were generated on a confocal microarray scanner (G2505B, Agilent) at 5 μm resolution and quantified using Feature Extraction (Agilent). Extracted log2-transformed intensities were quantile normalized to make all data comparable.

To assess differential expression, significance analysis of microarrays (SAM) was used (28). Results of SAM analysis were corrected for multiple testing according to the false discovery rate (FDR) method (29). Probes with FDR-adjusted P-value below 5% and additionally a fold change exceeding 1.2 in absolute value were selected as the relevant ones. Final relative expression values were computed by taking the median log2 ratio of the respective probes per each miRNA. All statistical analyses were performed with Bioconductor project (version 2.3) in the R statistical environment (30).

Quantitative real-time PCR

A total of 14 differentially expressed miRNAs found concordant both by sequencing and microarray analyses data were confirmed by real-time quantitative reverse transcription PCR (qRT-PCR) using TaqMan® MicroRNA assays (Applied Biosystems) in striatal samples of five control individuals and five HD patients. Total RNA was used as a template to generate specific first-strand complementary DNA (cDNA) for each miRNA using Taqman specific miRNA retrotranscription kit. qRT-PCR was performed following manufacturer’s instructions using the ABI PRISM 7900HT sequence detection system (Applied Biosystems). Relative quantification was calculated with the 2 −ΔΔCt method (31) using the snRNA, RNU58B as an endogenous control. Statistical significance and fold change were calculated using linear mixed-effects models (LMM) that contemplate random effects accounting for the different sources of variation: technical (qRT-PCRs) and biological replicates (32).

Functional enrichment analysis

We used the TargetScan algorithm that predicts biological targets by searching for the presence of conserved 8-mer and 7-mer sites that match the seed region of each isomiR (33,34). We used the TargetScan custom option (www.targetscan.org) to predict mRNA targets of the seed region IsomiRs differentially expressed between libraries. The mRNA targets for the corresponding differentially expressed reference miRNAs were identified through the SeqBuster ‘Target prediction’ module using the TargetScan algorithm. Since the cooperative action of multiple miRNAs can be multiplicative and sometimes synergistic (33), miRNAs with more predicted target sites for co-expressed IsomiRs or reference-miRNAs should be more drastically affected. Therefore, we considered targets predicted by three or more HD-FC or HD-ST co-regulated IsomiRs or reference miRNAs. Then, Ingenuity pathway analysis (IPA) was used to explore subsets of genes exclusively targeted by HD-co-regulated IsomiRs and those affected by the IsomiRs and reference-miRNAs. In the IPA filters, we only considered the brain and CNS cell lines as target tissues. The P-value associated with a biological process is calculated with the right-tailed Fisher’s exact test, considering the number of functions/pathways/lists eligible molecules that participate in that annotation, the total number of knowledge base molecules known to be associated with that function, the total number of functions/pathways/lists eligible molecules and the total number of genes in the Reference Set (IPA tutorial).

Analysis of the co-regulated miRNA promoter regions

We compared and characterized the promoter regions of miRNAs with similar expression patterns, using the algorithm described by Blanco et al. (35). The package requires a test file containing a list of the co-regulated miRNAs and a list with all the expressed miRNAs. To identify a possible significant enrichment of cis-regulatory elements in co-regulated miRNAs the algorithm assigns a P-value using the permutation-based simulation.

RESULTS

Overview of small RNA sequencing and annotation

In the present study, we chose to characterize small RNAs in the FC and the dorsal caudate nucleus of the ST, two brain areas strongly affected in HD that showed a differential pattern of neuronal loss. Total RNA was isolated from these brain areas in two HD patients and two nonaffected individuals showing similar quality standards (see ‘Methods’ section). Equivalent amounts of RNA were mixed from the FC or the ST, which resulted in four small RNA libraries from the FC and the ST of two patients with HD and two nonaffected individuals.

Sequencing of small RNAs yielded 6,788,286; 5,934,411; 6,495,715; and 6,776,627 unfiltered sequence reads from
miRNA, passenger strand of the miRNA duplex that is less stable; sense strand; Sn, small non-coding RNAs; tss, transcription start site

Table 1. Distribution of the different types of sequences in control and Huntington disease brain samples

<table>
<thead>
<tr>
<th>Database</th>
<th>Data</th>
<th>C-FC</th>
<th>HD-FC</th>
<th>C-ST</th>
<th>HD-ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>miRNAs</td>
<td>Frequency</td>
<td>5035</td>
<td>324</td>
<td>4019</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>20925</td>
<td>19683</td>
<td>23198</td>
<td>21102</td>
</tr>
<tr>
<td>miRNA*</td>
<td>Frequency</td>
<td>21544</td>
<td>27039</td>
<td>6729</td>
<td>23193</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>840</td>
<td>829</td>
<td>405</td>
<td>837</td>
</tr>
<tr>
<td>Pre-miRNAs</td>
<td>Frequency</td>
<td>19061</td>
<td>19283</td>
<td>18189</td>
<td>24084</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>1036</td>
<td>1347</td>
<td>1083</td>
<td>1654</td>
</tr>
<tr>
<td>mRNAs</td>
<td>Frequency</td>
<td>100651</td>
<td>91681</td>
<td>126949</td>
<td>87436</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>4582</td>
<td>3857</td>
<td>4671</td>
<td>4108</td>
</tr>
<tr>
<td>miRNAs</td>
<td>Predicted frequency</td>
<td>108921</td>
<td>77386</td>
<td>111211</td>
<td>70278</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>1094</td>
<td>846</td>
<td>1083</td>
<td>883</td>
</tr>
<tr>
<td>tRNAs</td>
<td>Frequency</td>
<td>0</td>
<td>225</td>
<td>103</td>
<td>280</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>0</td>
<td>45</td>
<td>23</td>
<td>56</td>
</tr>
<tr>
<td>Genome</td>
<td>Frequency</td>
<td>871</td>
<td>1051</td>
<td>321</td>
<td>1403</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>87</td>
<td>208</td>
<td>57</td>
<td>253</td>
</tr>
<tr>
<td>Unmapped</td>
<td>Frequency</td>
<td>541389</td>
<td>805763</td>
<td>644896</td>
<td>978547</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>32105</td>
<td>34594</td>
<td>37649</td>
<td>43573</td>
</tr>
<tr>
<td>Total</td>
<td>Frequency</td>
<td>582761</td>
<td>504166</td>
<td>5466923</td>
<td>5759841</td>
</tr>
<tr>
<td></td>
<td>Unique sequences</td>
<td>60669</td>
<td>61466</td>
<td>68168</td>
<td>72475</td>
</tr>
</tbody>
</table>

Sequences showing a frequency >3 are considered. miRNA*, passenger strand of the miRNA duplex that is less stable; +, sense strand; Sn, small non-coding RNAs; tss, transcription start site small RNA.

count FC (C-FC), HD-FC (HD-FC), control ST (C-ST) and HD ST (HD-ST), respectively. Any sequence observed more than three times was considered a reliable representation of a small RNA molecule. After removal of reads containing ambiguous base calls, unique sequences remained between 60 669 (C-FC), 61 416 (HD-FC) 68 168 (C-ST) and 72 475 (HD-ST) (Table 1). After mapping to the genome the total of unique sequences represented between 5 000 000 and 6 000 000 reads. Most reads mapped the genome, with <9–16% of the sequences not being identified (Table 1).

Sequences were annotated based on the overlap with publicly available genome annotations, including miRNA, tRNAs, rRNAs, other small RNAs and genomic repeats. miRNAs were the most abundant class of small RNAs (80–85%), spanning the entire range of expression. Sequences deriving from 372–392 distinct miRNA genes were identified in the different libraries, showing a sequence length distribution with a sharp peak in the 22 nt length (Supplementary Figure S1). Let-7 was the most abundant family of miRNAs, with sequence counts up to 600 000. The top conserved miRNAs occupied 40–50% of the total miRNA reads (Supplementary Table S2). However, human-specific miRNAs were expressed at a significantly lower level, constituting around 0.1% of the total number of reads. This agrees with previous reports showing that the miRNA abundance is linked to the extent of conservation (36).

To identify putative new miRNAs among the sequences mapped onto the genome, we used the microPred pipeline (25). This is a noncomparative computational tool developed for the identification of miRNAs from a pool of sequenced RNA transcripts resulting from deep sequencing. Between 846 and 1094 different sequences in each sample were found to statistically fulfill the structural features that define a miRNA, including the ability to form hairpin secondary structures that showed Dicer processing features. Approximately two-thirds of the sequences corresponded to slight length variations of a single representative. Therefore, 226–375 new putative miRNAs were found in each sample, being 85 common to all samples (Supplementary Table S3). Functional studies are required to demonstrate that these new putative miRNAs can regulate gene expression through the RISC silencing machinery.

miRNA variability is ubiquitous in control and HD brain samples

Variants of the mature miRNA have been referred to as IsomiRs (37). In the present study, 20 000–23 000 different variants (frequency >3) were found in every sample. Furthermore, most of the sequences mapping onto miRNA database (reference miRNAs and IsomiRs) showed from 3–50 counts (Supplementary Figure S2). It has been proposed that much of the miRNA modifications can be explained by variability in either Dicer1 or Drosha cleavage positions at the 5'–end or 3'–end of the miRNA-precursor (5'– or 3'–trimming variants) during miRNA biogenesis (37). Another type of variability is related with pri-miRNA post-trancriptional editing as a consequence of adenosine or cytidine desaminase activities (nucleotide substitution variants), which can be detected at different positions of the mature miRNA (38–42). Besides, nucleotide additions at the 3'–end of the mature miRNA (3'-addition variants) have also been reported as the most common form of miRNA enzymatic modification (37,41). The ~380 miRNAs detected in every sample displayed IsomiRs and only two and five miRNAs remained invariable in FC and ST libraries respectively (Table S4). The frequency of these invariable miRNAs was very low, suggesting that the lack of variants may be due to the low abundance rather than a selective protection from modifications.

In all the libraries, variability was detected mainly as the result of modifications in the cleavage position at the 3'–end of the pri-miRNA. In this case, a considerable proportion of miRNAs presented 2 or more 3'–trimming variants (Figure 1A). However, for the 5'–trimming variants, additions in the 3'–end of the mature miRNA, or single-nucleotide substitutions with respect to the reference sequence, the majority of the miRNAs presented IsomiRs with only one change.

Taking into account the possible physiological significance of an isomiR, it should be considered how the new variant affects the regulation of a target mRNA. miRNAs interact with their mRna targets by imperfect base pairing following a set of rules that have been identified in experimental and bioinformatic analyses (43,44). These include perfect binding of the 5'–seed (nt 2–8), mismatches at the central region of the miRNA (nt 10–12) and a reasonable complementarity with the 3' half of the miRNA, specially positions 13–16. It should be stressed that these requirements are not completely understood as target
Prediction algorithms have very high false-negative and false-positive rates. However, the requirements for target recognition favor the possibility that the identified miRNA variants target new mRNAs. In addition, since the variant and the reference miRNA sequences are very similar, it is likely that both small RNAs compete for a number of target mRNAs. Thus, to evaluate the possible relevance of each variant, we analyzed the frequency of each isomiR with respect to that of reference miRNA (Figure 1B). The majority of the nucleotide substitution or 5'-trimming variants showed a low frequency compared to that of the reference miRNAs. However, variants affecting the 3'-terminus of miRNAs, especially the 3'-trimming variants, showed a variable proportion with respect to reference miRNAs. These results highlight the possibility that the deregulation in the expression of nucleotide substitution and 5'-trimming IsomiRs could have more dramatic physiological consequences. The large abundance of some of the variants compared with the correspondent reference miRNA (over 80%) indicates that these IsomiRs should be considered as the reference miRNA, at least in the human brain. In some cases, such as miR-1291, miR-320c and miR-1260, the reference miRNA sequence was not even detected (see examples in Supplementary Table S5).

To gain insights into the possible relevance of the IsomiRs as small silencing RNAs we evaluated the presence of such variants in Ago immunoprecipitates (Ago-IP). We analyzed publicly available sequencing small-RNAs data sets (GEO database at http://www.ncbi.nlm.nih.gov/geo/) of human HEK293 Ago1- and Ago2-IP and mouse embryonic brain Ago1-IPs. We detected the four types of IsomiRs in all the Ago-IPs (Supplementary Figure S3). Furthermore, some of the IsomiRs found in the Ago-IPs were also detected in the sequenced brain samples (see examples in Supplementary Figure S3). In addition, the 3'-trimming and the 3'-addition variants were the most abundant, similarly to the analysis of total small-RNAs in the human brain samples (Figure 1). These results show that IsomiRs are incorporated into the Ago silencing complexes and strongly support the idea that the variants are functional.

We analyzed the positions and nucleotides involved in the different types of variants. Most 3'-trimming variants

![Figure 1](https://academic.oup.com/nar/article-abstract/38/20/7219/1306682/7224Nucleic%20Acids%20Research%2C%202010%2C%20Vol.%2038%2C%20No.%2020?download=true)
involved positions −1 (1 nt upstream) and +1 (1 nt downstream) of the 3′-end of the reference miRNA, while the majority of the 5′-trimming variants involved the position +1 of the 5′-terminus of the reference miRNA, matching the miRNA precursor (Figure 2A and B). Different nucleotides at different positions were preferentially involved in 5′- or 3′-trimming variants, which suggests a predilection of the dicing machinery for a nucleotide in a selective position. In addition, the vast majority of 3′-trimming variants consisted in a single A or U addition to the 3′-end of the mature miRNA (Figure 2C), in accordance with other reports (37,41).

An increasing number of studies have revealed nucleotide substitutions as a possible source of IsomiRs, which have been shown as the result of post-transcriptional modifications or RNA-editing, rather than sequencing technical artifacts (45–48). Illumina sequencing error rates have been previously calculated for every position of the sequence and every type of nucleotide change, in genomic libraries (27). Therefore, in the analysis of the nucleotide substitution variants, we only considered the nucleotide modifications significantly different from these error rates (see ‘Methods’ section). Under these conditions, ~35% of the miRNAs detected in each library presented nucleotide substitution variants (Figure 1). Table S6 shows the list of the nucleotide modifications at different positions of the reference miRNA, commonly detected in the four sequenced samples. We analyzed if these nucleotide substitution variants could correspond to SNPs. To this end, we compared the chromosome coordinates of all available SNPs with those of the nucleotide substitution variants. We found that miR-219-5p
presented variants (frequency \( \geq 10 \)) that matched with an SNP in dbSNP; however, the frequency of these IsomiRs was very low compared with the reference miRNA (Supplementary Table S7). This miRNA presented additional patterns of nucleotide substitution at the same positions. This suggests that the vast majority of the nucleotide substitution IsomiRs are likely post-transcriptional.

Nucleotide modifications affecting the seed region (nucleotides 2–8) are of particular interest given the essential role of this fragment in target gene recognition and expression regulation. The majority of the miRNAs with IsomiRs in the seed region, presented nt-modifications in positions 5–8 (Figure 3). In the seed region, A to G nucleotide changes (potential A to I editing) were observed in a number of miRNAs. These were particularly abundant at position 5 of the miR-411 and miR-379, as previously described (47). Other common examples included an A to G change in position 6 of all miRNAs of the cluster miR-376a, 376-b and 376-c, in agreement with previous reports (40). miR-320b and miR-320c are interesting examples to highlight, since they present nucleotide substitution variants at several positions of the mature miRNA (Supplementary Table S6). The detection of the majority of miR-320c variants was confirmed in the re-analysis of small RNAs sequencing data of human embryonic stem cells [(37); data not shown]. In addition to the nucleotides belonging to the seed region, other strongly invariable positions belong to the 3\(^\prime\) half of the miRNA, whose complementarity with target mRNAs is an important determinant for interaction stabilization (44) (Figure 3).

**miRNA transcriptome deregulation in HD samples**

To contrast small RNA expression between biologically comparable samples, similar sequencing efficiencies should be considered. To evaluate the sequencing performance, the total number of counts for every unique read in each sample was ordered according to decreasing frequencies (Supplementary Figure S4). The frequency distribution of the different reads showed no statistical differences between the samples (Wilcoxon test, \( P > 0.05 \)), making them comparable for differential expression analysis.

Figure 4 shows an overall view of the expression levels of all the different sequences (gray spots) mapping onto diverse miRNA loci (black spots) in HD-FC (\( y \)-axis) and HD-ST (\( x \)-axis). A considerable number of miRNAs loci were commonly deregulated in HD-FC and HD-ST (black spots in the upper right and lower left squares, respectively). To identify the miRNAs significantly deregulated in HD, we chose the \( Z \)-test (26), since it is the most restrictive test used for evaluation of differentially expressed sequences in high-throughput sequencing methodologies. Furthermore, to control the FDR the \( P \)-value assigned by the statistical test was corrected applying the Benjamini and Hochberg method (29). A total of 102 upregulated miRNAs [ratio >1.2; \( P <0.05 \); normalized frequency (HD-FC + C-FC) >50] and 58 downregulated miRNAs [ratio <−1.2 and \( P <0.05 \); normalized frequency (HD-FC+C-FC) >50] were detected in the HD-FC library. A similar approach in the HD-ST, showed 69 significantly upregulated miRNAs and 62 downregulated miRNAs (Supplementary Tables S8 and S9). The
majority (55–80%) of the HD deregulated miRNAs were common between FC and ST samples. Interestingly, the HD-FC presented an increased number of upregulated miRNAs (2.5-fold) that were not significantly altered in the HD-ST.

To identify the miRNA most likely relevant in HD we performed miRNA arrays of FC and ST in four control individuals and four HD patients. Considering an FDR-adjusted P-value below 5% and a fold expression change above 1.2 or below \(-1.2\) in absolute value, 19 upregulated and 14 downregulated miRNAs, common to the FC and the ST were confirmed (Table 2). Selected miRNAs (seven upregulated and seven downregulated) were further validated using Taqman assays in ST samples in five control individuals and five HD patients. It should be taken into consideration that these techniques, differing from sequencing strategies, are most favorable for the detection of the reference miRNAs and may not be optimal for the detection of IsomiRs.

The differential expression analysis of the sequenced samples confirmed the expression pattern of 9 out of 11 miRNA (Supplementary Table S10), whose expression was altered in the FC of HD patients in two previous studies (17,18). In addition, the majority of the reported miRNAs deregulated in other neurodegenerative disorders (Alzheimer’s disease, Parkinson’s disease, Prion disease and spinocerebellar ataxia type 1) were altered in the HD sequenced samples (Supplementary Table S11). Furthermore, a fraction of these miRNAs showed an expression deregulation pattern similar to that reported in the other neurodegenerative diseases, indicating that they may participate in gene expression deregulation underlying common processes in neurodegeneration.

When studying the correlation between the expression pattern of all IsomiRs and the corresponding reference miRNA our analysis showed that the majority (80%) of HD-upregulated 3'- or 5'-IsomiRs displayed an upregulation of the corresponding reference miRNA. A similar result was obtained when considering the downregulated reference IsomiRs (Supplementary Table S12). These results suggest that the mechanisms modulating the degree of expression of the IsomiRs and the corresponding miRNAs are parallel in most cases. However, a few sequences mapping on the same miRNA were discordant, presenting the reference miRNA and the isomiR an opposite expression pattern in HD. The majority of these discordant IsomiRs were 3'-variants (3'-trimming and 3'-addition variants). Whether selective forms of a miRNA are relevant in HD and how these variants are selectively deregulated in HD deserve further research.

Since the seed region is essential in the recognition and expression modulation of target miRNAs we analyzed HD-deregulated 5'-trimming variants and nucleotide substitution variants involving any nucleotide change in positions 2–8 of the reference miRNA (Table 3). A total of 23 seed-region IsomiRs were deregulated in HD FC and/or ST, that corresponded to 18 different miRNAs. Of these, miR-935, miR-139-3p, miR-1307, miR-1308, miR-411 and miR-142-5p variants can be considered the only representatives among all the IsomiRs mapping onto the same locus. The rest of the seed-region deregulated IsomiRs showed a varying proportion (10–50%) with respect to that of the reference or the most abundant variant mapping onto the same locus.

The REST transcription factor mis-function has been considered to play a major role in HD (2–3). We analyzed whether miRNAs that are potentially regulated by REST showed differential expression in our sequencing libraries. For this analysis we considered miRNAs with known motifs for REST (4,17,49,50), which were detected at frequencies above 10 in all the samples (a total of 30 miRNAs). It is worth to mention that 16 (ST) and 19 (FC) out of the 30 REST modulated
miRNAs, were significantly deregulated in the HD samples (Table 4). Interestingly, miR-9 in HD-FC and miR-29b in HD-ST presented discordant up- and downregulated variants, which highlights the putative relevance of the different types of IsomiRs in the context of HD. Considering all the deregulated REST target miRNAs, the number of HD downregulated miRNAs was twice the amount that would be expected by chance, both in HD-FC (\( P < 0.004 \), Bootstrapping method) and HD-ST (\( P = 0.003 \)). Given that REST is a transcriptional repressor, the enrichment in repressed REST target miRNAs agrees with a major role of this transcription factor in miRNA dysregulation in HD. To address whether other transcription factors could explain miRNA gene expression deregulation we characterized the promoter regions of co-regulated miRNAs (35). The results showed that in addition to REST, p53 regulatory

<table>
<thead>
<tr>
<th>MiRNA</th>
<th>Expression pattern in HD (sequencing data)</th>
<th>Expression pattern in HD (miRNA arrays)</th>
<th>Expression pattern in HD (miRNA qRT-PCR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-let-7g</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
</tr>
<tr>
<td>hsa-miR-100</td>
<td>UP</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-101</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-106b</td>
<td>UP</td>
<td>UP</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-1250</td>
<td>UP</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-126</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-126*</td>
<td>UP</td>
<td>NA</td>
<td>DOWN</td>
</tr>
<tr>
<td>hsa-miR-1308</td>
<td>UP</td>
<td>NA</td>
<td>DOWN</td>
</tr>
<tr>
<td>hsa-miR-143</td>
<td>UP</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-145</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-146a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-148b</td>
<td>UP</td>
<td>UP</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-151-3p</td>
<td>UP</td>
<td>UP</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-151-5p</td>
<td>UP</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-15a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-15b</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-16</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
</tr>
<tr>
<td>hsa-miR-17</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
</tr>
<tr>
<td>hsa-miR-181a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-181a*</td>
<td>UP</td>
<td>UP</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-193b</td>
<td>UP</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-197</td>
<td>UP</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-1974</td>
<td>UP</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-199b-3p</td>
<td>UP</td>
<td>UP</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-199b</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-196</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-204</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-20a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-219-2-3p</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-219-5p</td>
<td>UP</td>
<td>UP</td>
<td>UP</td>
</tr>
<tr>
<td>hsa-miR-22</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-23a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-27b</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-27</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-29k*</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-30a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-30b</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-30c</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-30e</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-338-3p</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-33a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-33</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-363</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-364</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-374a*</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-451</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-451</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-484</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-486-5p</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-547-3p</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-644</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-887</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-92a</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-93</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-99b</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-let-7a</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-let-7c</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-let-7d</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-let-7e</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-103</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-107</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-1224-5p</td>
<td>DOWN</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-124</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-127-3p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-128</td>
<td>DOWN</td>
<td>DOWN</td>
<td>DOWN</td>
</tr>
<tr>
<td>hsa-miR-1301</td>
<td>DOWN</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-1307</td>
<td>DOWN</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-139-3p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>DOWN</td>
</tr>
<tr>
<td>hsa-miR-181d</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-193b*</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-199a-3p</td>
<td>DOWN</td>
<td>NA</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-221</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-222</td>
<td>DOWN</td>
<td>DOWN</td>
<td>DOWN</td>
</tr>
<tr>
<td>hsa-miR-323-3p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-330-3p</td>
<td>DOWN</td>
<td>UP</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-338-3p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-340-5p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-382</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-383</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-409-5p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-423-5p</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-432</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-433</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-485-3p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-485-5p</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-92b*</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-495</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-543</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-598</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-708</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-760</td>
<td>DOWN</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>hsa-miR-95</td>
<td>DOWN</td>
<td>DOWN</td>
<td>–</td>
</tr>
</tbody>
</table>

The differential expression in the sequenced samples considers miRNAs commonly deregulated in HD-FC and HD-ST > 1.2 and < -1.2. The reference miRNA and all the variants mapping on the same miRNA locus are considered in this analysis. Only sequences with a normalized frequency cutoff above 50 (norm freq C-FC + norm freq HD-FC > 50) and (norm freq C-ST + norm freq HD-ST > 50) were included. In the miRNA arrays, a fold change above 1.2 and below –1.2 in absolute value was considered. The results showing a false discovery rate-adjusted \( P \)-value (FDR) below % are highlighted in black, and those with a \( P > 10% \) are shown in light gray. Selected miRNAs (seven upregulated and seven downregulated) were chosen for validation by RT-PCR. The significantly deregulated miRNAs (\( P < 0.01 \)) are highlighted in black. (--) Means that no deregulation was found. (NA) Means that the miRNA was not available in the array. In bold, the miRNAs with a more consistent deregulation pattern using different techniques in additional samples are shown.
motifs are significantly enriched in the promoter region of the downregulated miRNAs in the FC (Table S13).

Putative targets of HD deregulated miRNAs and seed-region IsomiRs

We focused on the possible biological pathways affected by HD-deregulated miRNAs and seed-region IsomiRs. In this analysis we included miRNA/IsomiRs commonly upregulated or downregulated in the FC and ST (Table 3 and Supplementary Tables S8 and S9). We first identified putative targets for the differentially expressed miRNAs and IsomiRs using the TargetScan 5.1 algorithm that predicts biological targets of miRNAs or IsomiRs by searching for the presence of conserved 8-mer and 7-mer sites that match the seed region of each miRNA/IsomiR (33,34). Since the cooperative action of multiple miRNAs can be multiplicative and sometimes synergic (33), mRNAs with more predicted target sites for co-expressed miRNAs/IsomiRs should be more drastically affected. Therefore, we considered targets predicted by three or more HD-upregulated or HD-downregulated miRNAs/IsomiRs. Based on these parameters we identified 45 (HD-enriched) and 40 (HD-diminished) significantly deregulated seed regions that corresponded to 55 and 43 miRNA/IsomiRs, respectively. The number of targets common to a minimum of three co-regulated miRNAs/IsomiRs was 1922 for HD-enriched and 1244 for HD-diminished miRNAs/IsomiRs. It is worth noting that the two lists of genes showed a significant overlap.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Type of variant</th>
<th>Seed region</th>
<th>Norm freq HD-FC</th>
<th>Norm freq CFC</th>
<th>HD-FC (P &lt; 0.05)</th>
<th>Norm freq HD-ST</th>
<th>Norm freq CST</th>
<th>HD-ST (P &lt; 0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>has-miR-101</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>UACAGUA</td>
<td>3972</td>
<td>2421</td>
<td>UP</td>
<td>3258</td>
<td>1955</td>
</tr>
<tr>
<td>has-miR-124</td>
<td>nt- substitution</td>
<td>IsomiR</td>
<td>AAAGCAGA</td>
<td>75</td>
<td>48</td>
<td>UP</td>
<td>26</td>
<td>70</td>
</tr>
<tr>
<td>has-miR-124</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AAAGGCA</td>
<td>751</td>
<td>928</td>
<td>DOWN</td>
<td>196</td>
<td>503</td>
</tr>
<tr>
<td>has-miR-126</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AGGCAG</td>
<td>139</td>
<td>125</td>
<td>–</td>
<td>41</td>
<td>84</td>
</tr>
<tr>
<td>has-miR-130</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>GUACCGG</td>
<td>163</td>
<td>131</td>
<td>–</td>
<td>189</td>
<td>83</td>
</tr>
<tr>
<td>has-miR-130</td>
<td>nt- substitution</td>
<td>IsomiR</td>
<td>UCGGCGG</td>
<td>47</td>
<td>82</td>
<td>DOWN</td>
<td>47</td>
<td>65</td>
</tr>
<tr>
<td>has-miR-139-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>GGGAGCGAG</td>
<td>226</td>
<td>667</td>
<td>DOWN</td>
<td>56</td>
<td>367</td>
</tr>
<tr>
<td>has-miR-140-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>ACCAGG</td>
<td>5751</td>
<td>4051</td>
<td>UP</td>
<td>3396</td>
<td>3703</td>
</tr>
<tr>
<td>has-miR-142-5p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AUAAAGG</td>
<td>67</td>
<td>38</td>
<td>UP</td>
<td>46</td>
<td>49</td>
</tr>
<tr>
<td>has-miR-192</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>GACCUAU</td>
<td>0</td>
<td>51</td>
<td>DOWN</td>
<td>36</td>
<td>37</td>
</tr>
<tr>
<td>has-miR-1974</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>UGGUUGU</td>
<td>126</td>
<td>51</td>
<td>UP</td>
<td>94</td>
<td>69</td>
</tr>
<tr>
<td>has-miR-1974</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>GGUUGA</td>
<td>45</td>
<td>24</td>
<td>UP</td>
<td>47</td>
<td>33</td>
</tr>
<tr>
<td>has-miR-1979</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>CCAACUGU</td>
<td>42</td>
<td>27</td>
<td>UP</td>
<td>56</td>
<td>34</td>
</tr>
<tr>
<td>has-miR-199a-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>CACAGU</td>
<td>4</td>
<td>71</td>
<td>DOWN</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>has-miR-199b-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>CACAGU</td>
<td>66</td>
<td>0</td>
<td>UP</td>
<td>49</td>
<td>10</td>
</tr>
<tr>
<td>has-miR-323-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>CACAUAC</td>
<td>104</td>
<td>197</td>
<td>DOWN</td>
<td>70</td>
<td>143</td>
</tr>
<tr>
<td>has-miR-330-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>CAAAGCA</td>
<td>907</td>
<td>1138</td>
<td>DOWN</td>
<td>761</td>
<td>1017</td>
</tr>
<tr>
<td>has-miR-330-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AAGCACA</td>
<td>240</td>
<td>323</td>
<td>DOWN</td>
<td>186</td>
<td>222</td>
</tr>
<tr>
<td>has-miR-342-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AAGCACA</td>
<td>70</td>
<td>50</td>
<td>–</td>
<td>0</td>
<td>62</td>
</tr>
<tr>
<td>has-miR-411</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AGUAGAC</td>
<td>85</td>
<td>88</td>
<td>–</td>
<td>29</td>
<td>53</td>
</tr>
<tr>
<td>has-miR-485-3p</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>GUAUAC</td>
<td>22</td>
<td>47</td>
<td>DOWN</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>has-miR-935</td>
<td>5’ trimming</td>
<td>IsomiR</td>
<td>AGUACC</td>
<td>365</td>
<td>296</td>
<td>UP</td>
<td>100</td>
<td>169</td>
</tr>
</tbody>
</table>

The seed region of the miRNAs in gray should be considered as the only representatives among all the isomiRs mapping in the same locus; for these miRNAs no reference seed region was sequenced.
suggesting that some genes can be putatively targeted by sets of unique HD up and downregulated miRNAs. The IPA tool was used for the subsets of genes affected by the HD-enriched or HD-downregulated miRNAs/IsomiRs. The most enriched biological and disease processes and canonical pathways amongst the HD-enriched or HD-diminished miRNA/IsomiR targets are provided in Table 5. Some of the top functional groups, including genetic disorder, neurological disease, developmental disorders, cell-to-cell signaling and

<table>
<thead>
<tr>
<th>Locus</th>
<th>miR/isomiR</th>
<th>HD-FC versus C-FC (P &lt; 0.05)</th>
<th>HD-ST versus C-ST (P &lt; 0.05)</th>
<th>REST modulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-124</td>
<td>Ref.</td>
<td>–</td>
<td>DOWN</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-124</td>
<td>NA</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-124</td>
<td>s:10AC</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-124</td>
<td>s:19UC</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-124</td>
<td>s:8AC</td>
<td>–</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-124</td>
<td>tr5(down)U</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-29a</td>
<td>NA</td>
<td>UP</td>
<td>UP</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-29a</td>
<td>Ref.</td>
<td>UP</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-29b</td>
<td>NA</td>
<td>UP</td>
<td>UP</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-29b</td>
<td>Ref.</td>
<td>UP</td>
<td>UP</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-9</td>
<td>NA</td>
<td>UP</td>
<td>–</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-9</td>
<td>Ref.</td>
<td>DOWN</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-132</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-132</td>
<td>Ref.</td>
<td>UP</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-212</td>
<td>NA</td>
<td>UP</td>
<td>–</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-139-5p</td>
<td>NA</td>
<td>UP</td>
<td>DOWN</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-139-5p</td>
<td>Ref.</td>
<td>–</td>
<td>DOWN</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-139-5p</td>
<td>tr5(up)GG</td>
<td>UP</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-139-5p</td>
<td>tr5(down)U</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-133a</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>EMSA (17)</td>
</tr>
<tr>
<td>hsa-miR-203</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>EMSA/ChIP (17)</td>
</tr>
<tr>
<td>hsa-miR-204</td>
<td>NA</td>
<td>UP</td>
<td>UP</td>
<td>EMSA/ChIP (17)</td>
</tr>
<tr>
<td>hsa-miR-204</td>
<td>Ref.</td>
<td>UP</td>
<td>UP</td>
<td>EMSA/ChIP (17)</td>
</tr>
<tr>
<td>hsa-miR-330-3p</td>
<td>NA</td>
<td>–</td>
<td>DOWN</td>
<td>EMSA/ChIP (17)</td>
</tr>
<tr>
<td>hsa-miR-330-3p</td>
<td>Ref.</td>
<td>–</td>
<td>DOWN</td>
<td>EMSA/ChIP (17)</td>
</tr>
<tr>
<td>hsa-miR-330-3p</td>
<td>tr5(up)G</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-330-3p</td>
<td>tr5(up)GC</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-346</td>
<td>NA</td>
<td>–</td>
<td>DOWN</td>
<td>EMSA/ChIP (17) RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-218</td>
<td>NA</td>
<td>–</td>
<td>DOWN</td>
<td>EMSA/ChIP (17)</td>
</tr>
<tr>
<td>hsa-miR-95</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-95</td>
<td>Ref.</td>
<td>DOWN</td>
<td>DOWN</td>
<td>Chromatin IP (49)</td>
</tr>
<tr>
<td>hsa-miR-153</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-455</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>RE1 motive (50)</td>
</tr>
<tr>
<td>hsa-miR-146b-5p</td>
<td>NA</td>
<td>DOWN</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-1179</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-1301</td>
<td>NA</td>
<td>DOWN</td>
<td>DOWN</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-1301</td>
<td>Ref.</td>
<td>DOWN</td>
<td>DOWN</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-129-3p</td>
<td>NA</td>
<td>UP</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-129-3p</td>
<td>Ref.</td>
<td>UP</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-129-3p</td>
<td>s:20UC</td>
<td>UP</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-129-5p</td>
<td>NA</td>
<td>DOWN</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-129-5p</td>
<td>Ref.</td>
<td>DOWN</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-7</td>
<td>NA</td>
<td>DOWN</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-7</td>
<td>Ref.</td>
<td>DOWN</td>
<td>UP</td>
<td></td>
</tr>
<tr>
<td>hsa-miR-137</td>
<td>NA</td>
<td>–</td>
<td>DOWN</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-137</td>
<td>Ref.</td>
<td>–</td>
<td>DOWN</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-184</td>
<td>Ref.</td>
<td>–</td>
<td>DOWN</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-1224-5p</td>
<td>NA</td>
<td>DOWN</td>
<td>DOWN</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-940</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-375</td>
<td>NA</td>
<td>–</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-328</td>
<td>NA</td>
<td>UP</td>
<td>–</td>
<td>RE1 motive (4)</td>
</tr>
<tr>
<td>hsa-miR-328</td>
<td>Ref.</td>
<td>UP</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

(-) no differential expression was found in the present study; Ref, reference miRNA; tr5, 5′-trimming variants; (up) and (down) followed by the indicated nucleotides show the short sequence involved in the trimming variants, upstream and downstream of the corresponding reference miRNA terminus, respectively; s, shows a nucleotide substitution variant.

The position of the nucleotide and the pattern of the modification is shown next. NA, considers, for differential expression analysis, all the sequences mapping onto a specific miRNA locus different from the reference miRNA or the trimming variants, when these are shown in a separate row. miRNAs in gray did not show significant expression changes.
Here, we report the first unbiased analysis of small RNA populations of two different brain areas of healthy individuals and individuals with HD. Deep sequencing of small RNAs from the human brain revealed length and sequence heterogeneity for the vast majority of miRNAs, indicating that the miRNA transcriptome is more complex than previously described. However, the fact that the proportion of the different types of variants was found to be similar in all brain samples agrees with the concept that the molecular mechanisms involved in the generation of these variants are not altered in HD.

miRNA variability (IsomiRs) has recently been described using several large-scale sequencing strategies in plants (45,46), mouse and human stem cells (37,51), and human brain samples (47). Our analysis revealed that IsomiRs are also present in human and mouse Ago-IPs, suggesting that they are functional as gene expression modulators. We have found 20,000–23,000 different variants with a frequency >3 in every brain sample. Our study permits several generalizations regarding the types of miRNA variants in the human brain, irrespective of HD neuropathology. Around 80–90% of the miRNAs presented modifications in the 3’-terminus either in the form of trimming and/or as nucleotide addition variants.

### Table 5. Top biological functions and canonical pathways for predicted targets of HD-deregulated miRNAs and seed region-isomiRs: IPA analysis

<table>
<thead>
<tr>
<th>Biological/molecular function</th>
<th>P-value</th>
<th>No. of Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease and disorders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic disorder</td>
<td>1.11E-03</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>Psychological disorders</td>
<td>1.11E-03</td>
<td>1.11E-03</td>
</tr>
<tr>
<td>Developmental disorder</td>
<td>1.02E-02</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>Skeletal and muscular disorders</td>
<td>2.29E-02</td>
<td>2.29E-02</td>
</tr>
<tr>
<td>Cellular movement</td>
<td>7.59E-05</td>
<td>4.26E-02</td>
</tr>
<tr>
<td>Cell morphology</td>
<td>1.77E-04</td>
<td>2.82E-02</td>
</tr>
<tr>
<td>Cellular assembly and organiz</td>
<td>2.20E-04</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>Cell-to-cell signalling and interaction</td>
<td>3.05E-04</td>
<td>2.75E-02</td>
</tr>
<tr>
<td>Cellular growth and proliferation</td>
<td>3.88E-04</td>
<td>2.29E-02</td>
</tr>
<tr>
<td>Nervous system development and function</td>
<td>2.00E-05</td>
<td>4.19E-02</td>
</tr>
<tr>
<td>Organ development</td>
<td>3.18E-04</td>
<td>2.29E-02</td>
</tr>
<tr>
<td>Tissue morphology</td>
<td>6.08E-03</td>
<td>2.82E-02</td>
</tr>
<tr>
<td>Organ morphology</td>
<td>2.02E-02</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>Embryonic development</td>
<td>2.29E-02</td>
<td>2.29E-02</td>
</tr>
<tr>
<td>Axonal guidance signaling</td>
<td>3.13E-07</td>
<td>79/320</td>
</tr>
<tr>
<td>TGF-b signaling</td>
<td>3.94E-07</td>
<td>27/70</td>
</tr>
<tr>
<td>Glialoma signaling</td>
<td>4.11E-07</td>
<td>32/91</td>
</tr>
<tr>
<td>Molecular mechanisms of cancer</td>
<td>6.41E-07</td>
<td>75/304</td>
</tr>
<tr>
<td>B cell receptor signaling</td>
<td>7.57E-07</td>
<td>38/120</td>
</tr>
<tr>
<td>Nervous system development and function</td>
<td>2.00E-05</td>
<td>4.19E-02</td>
</tr>
<tr>
<td>Organ development</td>
<td>3.18E-04</td>
<td>2.29E-02</td>
</tr>
<tr>
<td>Tissue morphology</td>
<td>6.08E-03</td>
<td>2.82E-02</td>
</tr>
<tr>
<td>Organ morphology</td>
<td>2.02E-02</td>
<td>2.02E-02</td>
</tr>
<tr>
<td>Embryonic development</td>
<td>2.29E-02</td>
<td>2.29E-02</td>
</tr>
</tbody>
</table>

### DISCUSSION

*Table 5.* Top biological functions and canonical pathways for predicted targets of HD-upregulated miRNAs and isomiRs:

<table>
<thead>
<tr>
<th>Biological/molecular function</th>
<th>P-value</th>
<th>No. of Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease and disorders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genetic disorder</td>
<td>9.57E-04</td>
<td>4.75E-02</td>
</tr>
<tr>
<td>Developmental disorder</td>
<td>8.64E-03</td>
<td>4.36E-02</td>
</tr>
<tr>
<td>Genetic disorder</td>
<td>2.11E-02</td>
<td>2.11E-02</td>
</tr>
<tr>
<td>Skeletal and muscular disorders</td>
<td>4.56E-02</td>
<td>4.56E-02</td>
</tr>
<tr>
<td>Cellular movement</td>
<td>1.27E-04</td>
<td>4.40E-02</td>
</tr>
<tr>
<td>Cellular development</td>
<td>7.85E-04</td>
<td>4.42E-02</td>
</tr>
<tr>
<td>Cell-to-cell signalling and interaction</td>
<td>9.57E-04</td>
<td>2.43E-02</td>
</tr>
<tr>
<td>Cell death</td>
<td>1.41E-03</td>
<td>3.24E-02</td>
</tr>
<tr>
<td>Cellular growth and proliferation</td>
<td>1.51E-03</td>
<td>2.42E-02</td>
</tr>
<tr>
<td>Nervous system development and function</td>
<td>8.05E-05</td>
<td>4.56E-02</td>
</tr>
<tr>
<td>Organ development</td>
<td>8.05E-05</td>
<td>8.05E-05</td>
</tr>
<tr>
<td>Tissue development</td>
<td>8.02E-04</td>
<td>8.64E-03</td>
</tr>
<tr>
<td>Embryonic development</td>
<td>1.97E-03</td>
<td>2.11E-02</td>
</tr>
<tr>
<td>Tissue morphology</td>
<td>6.39E-03</td>
<td>4.22E-02</td>
</tr>
</tbody>
</table>

interaction, cell growth and proliferation, nervous system development and function and organ development were common between the two groups of miRNAs/IsomiRs, which is in agreement with the significant overlap of the predicted mRNA targets. HD was identified as one of the significant canonical pathways (P = 6.6E-03) for the predicted HD-downregulated miRNAs/IsomiRs (Figure S5). Interestingly, 16 out of the 30 predicted targets involved HD-canonical pathway are significantly deregulated in HD, as shown in the study by Hodges et al. (1). Ten out of these 16 mRNAs were upregulated (Supplementary Table S14), suggesting that the downregulation of specific miRNAs could directly contribute to the aberrant increase in the expression of genes relevant in HD pathways. In line with this, the amount of predicted targets of the downregulated miRs/IsomiRs that overlaps with the upregulated mRNAs (1) was significantly enriched (9%, P = 0.004, Bootstraping method) compared with the predicted targets of the randomly selected miRs/IsomiRs expressed in the brain. The TarBase of experimentally supported targets for several miRNAs (http://diana.cslab.ece.ntua.gr/tarbase/) confirmed a similar idea (Table S15). Considering the available validated targets of the HD-deregulated miRNAs, the number of overlapping HD-upregulated genes in the Hodges et al. (1) study (a total of 62) was clearly increased compared with that of the HD-downregulated genes (a total of 7).

To gain insights into the biological pathways possibly affected by isomiR expression deregulation, we run a specific IPA for the predicted targets of the seed-region-HD-deregulated IsomiRs (Table 3). The results revealed a significant enrichment in genes involved in neurological diseases for both HD-enriched and HD-diminished genes relevant in HD pathways. In line with this, the HD canonical pathway was significantly enriched for the HD-downregulated seed region isomiR targets. These results underline the possible contribution of the IsomiRs in the modulation of HD related pathways.
Furthermore, considering all the sequences mapping onto the same miRNA locus, the 3′-terminus variants presented a higher proportion with respect to the reference miRNA. This is consistent with the idea that variability in the 3′-terminus of the miRNA is more permissive, having minor consequences in gene expression regulation. In contrast, our analysis displayed a decreased proportion of miRNAs with heterogeneity at the 5′-terminus, showing the majority of these variants a negligible abundance as compared with that of the reference sequence. This suggests that the 5′ terminus of the miRNAs is specially protected from variations, which agrees with the crucial role of Watson–Crick base pairing of the 5′-seed region of the miRNA with the 3′ UTR of the mRNA, for gene targeting (43,44).

In accordance with previous reports (41,52,53), variability in the 3′-terminus was detected as an A or U single nucleotide extension. These modifications could influence mRNA expression regulation, since 3′-end pairing has been suggested to contribute to target recognition, particularly when sites have weaker miRNA seed matches (54,55). In plants, it has been shown that addition of A residues to the 3′-end plays a negative role in miRNA degradation (56). In addition, it has been proposed that the combined effects of 5′ deletions and 3′ U extensions in *Oryza sativa* and *Arabidopsis thaliana* can alter the specificity by which miRNAs associate with different Argonaute proteins (45). Whether short nucleotide extensions in animal miRNAs influence their stability or the silencing mechanism needs to be proven. However, recent data show selective stability for different miRNAs in human cells that depends on regions of the 3′-terminus (57). Differences in the stability could provide an explanation for the discordant expression deregulation of some 3′-trimming variants and the corresponding reference miRNAs.

Our analysis highlights significant nucleotide modifications along the reference mature miRNAs. Several lines of evidence argue against RT-PCR and sequencing errors as contributors to sequence discrepancies with respect to the reference miRNAs. First, the frequencies of nucleotide modifications were remarkably higher compared to the estimates attributable to Illumina sequencing errors (27). Second, we observed a positional non-randomness of nucleotide changes along the length of the miRNA seen in all libraries. Third, analogous nucleotide modifications were found in four independent libraries. Fourth, several nucleotide substitution variants found in the brain have been reported in independent sequencing experiments in human samples. Finally, nucleotide changes, insertions and deletions have been reported in previous studies using different sequencing strategies.

Nucleotide substitutions have been described in few miRNAs as pri-miR precursor editing changes from A to G attributed in part to A to I deaminations, which lead to a repression in the maturation of the miRNA (38–41,45,58). Edited adenosines have been also described in mature miRNAs (38–40). Our results not only have confirmed the existence of these variants, but also show that an A to G change is one of the most common types of nucleotide substitution events in the seed region. However, a considerable proportion of brain miRNAs presented nucleotide changes that were distinct from the classical A to I editing. Common nucleotide substitutions in brain miRNAs included U to G, C to A, G to A and G to U changes. These patterns of nucleotide modifications have been recently shown in miRNAs and tRNAs of plants (45) and in the let family of miRNAs in different mouse cell lines (48). In animals, these modifications in the mature miRNA may result in alternative base pairing between the variant and the target mRNA, possibly affecting the efficiency and/or quality of gene expression regulation (59). This has been shown human miR-376a that presents an A to G editing event in the seed region that affects gene targeting (40). Another example has been recently reported in the mouse let-7 family of miRNAs, as several types of nucleotide modifications at the ninth position of mmu-let7a result in an increase of the stability of the predicted miRNA:mRNA duplexes (48).

We have found that the detected nucleotide substitution events are not randomly distributed along the sequence. The less variable positions present key roles in gene target recognition and silencing. For instance, scarce miRNAs presented nucleotide substitutions at positions 2–4 of the miRNA within the 5′-seed region (nt 2–8). A contiguous and perfect base pairing of the miRNA nucleotides 2–8, representing the seed region, is the most stringent requirement for efficient target recognition (44,54). In addition, low variability was detected at positions 14–16, contained in an anchor site (nt 13–16) that has been proposed to hold important determinants for miRNA:mRNA association. However, positions 9–12 of the miRNA were amongst the more variable, which could contribute to the imperfect miRNA:mRNA pairing detected in the central region of the miRNA. For gene expression repression, bugles or mismatches must be present in the central region of the miRNA:mRNA duplex (9–12 of the miRNA), precluding the Argonaute-mediated endonucleolytic cleavage of mRNA (44). All together, the present analysis strongly suggests a biological function for this sequence plasticity in miRNAs, which may have broad implications in mRNA targeting, stability and/or gene expression regulation mechanism.

HD is characterized by widespread changes in neuronal gene expression (1,60). Recent studies have shown that the miRNA transcriptome is also perturbed in HD (17,18) affecting miR-9-9*, miR-124a, miR-132, miR-29b, miR-29a, miR-330, miR-17, miR-196a, miR-222, miR-485 and miR-486. Although further miRNA sequencing of HD brain samples will provide a comprehensive picture of the myriad of small RNAs in this disorder, our results confirm the deregulation pattern for 9 out of the 11 miRNAs, but also suggested that miRNA expression deregulation is more extensive in HD, affecting a considerable number of miRNAs and IsomiRs. Alterations in the miRNA/isomiR transcriptome may contribute to the aberrant expression of downstream genes detected in HD. In line with this, the putative targets of the HD-deregulated miRNAs and/or seed-region IsomiRs denote that these variants could participate in HD pathogenesis. It is worth pointing out that HD canonical pathway was among the significantly enriched pathways...
for the HD-downregulated miRs and/or IsomiRs. In addition, the study by Hodges et al. (1) shows that many HD-downregulated genes may underlie a defect in the capacity to maintain axonal projections. In agreement, our analysis revealed that axonal guidance signaling and actin and cytoskeleton signaling were among the highly significantly enriched pathways for the predicted targets of the upregulated miRNAs.

A large number of miRNAs presented the same deregulation profile in the two studied areas, which show a different pattern of neuronal loss, similarly to what has been reported in gene expression analysis of different brain areas in HD (1). Therefore, shared deregulated miRNAs may lie beneath gene expression deregulation linked to the common neuropathological changes, including neuronal degeneration, astrocitic gliosis and moderate microglial proliferation. Since advanced stages of the disease are considered in the present study, it is feasible that many changes in the miRNA transcriptome are not linked to HD-primary episodes, but may be reactive events to secondary-common processes occurring in neurodegeneration, such as oxidative stress. In line with this, several miRNAs present a similar pattern of deregulation as that described in other neurodegenerative disorders. In addition, at advanced stages some of the HD-related transcriptional modifications, especially gene expression downregulation, could be caused by changes in the relative number of neurons linked to selective neurodegeneration rather than changes in specific neuronal miRNA genes. However, a previous study shows similar HD-deregulated miRNAs in data sets of brain homogenates and laser capture microdissection that ensures similar numbers of neurons in control and HD samples (1). These data suggest that miRNA expression deregulation could be explained, at least in part, by transcriptional changes occurring in the remaining cells. The region-specific miRNA transcriptome deregulation may account for selective pathophysiological changes in HD, including the preferential affectionation of the gamma-aminobutyric acid (GABA)ergic spiny neurons in the ST and the projection neurons in the FC.

Some miRNAs that are in chromosomal clusters exhibited similar expression patterns in HD. For instance, miR-221 and miR-222 were both downregulated in HD. In addition, we noted that in some cases, miRNA members of the same family exhibited similar expression trends, suggesting their co-regulation in HD. For instance, let-7a, let-7c, let-7d and let-7e were downregulated in HD-FC and HD-ST. In another example, miR-30a, miR-30b, miR-30c and miR-30e were all upregulated in HD-FC and HD-ST. These results are in agreement with the idea of a relevant role of certain miRNA families in HD pathology. An interesting opened question is the understanding of the mechanisms accounting for the co-regulation of non-clustered miRNAs of the same family in pathological processes.

Consistent with both the deregulation of miRNA and mRNA gene expression in HD, a number of transcriptional modulators are affected by mutant HTT (61). There is strong evidence that the abnormal function of the transcriptional repressor REST plays a major role in gene expression deregulation in HD (2,3). BDNF is a REST-modulated neuronal gene whose loss in the striatum of HD patients contributes to the medium spiny neuronal loss and clinical manifestations of the disease (3,62). It is interesting that miR-30a, which has been shown to target BDNF (63), is significantly upregulated in both HD-FC and HD-ST brain samples. Therefore, in addition to REST, miR-30a upregulation could contribute to BDNF expression inhibition.

Some neuronal miRNAs modulated by REST (miR-124, miR-29a, miR-29b, miR-9/9* miR-132 and miR-330-3p) are abnormally expressed in the FC of HD patients (5,17,18). Taking into account all brain-expressed miRNAs showing experimental or predictive evidences for REST modulation (REST-miRNAs) our analysis identified a larger number of deregulated REST-miRNAs. Furthermore, these miRNAs were found to be enriched among those downregulated in HD brain regions. Overall, our results agree with the concept that REST dysfunction generally alters the miRNA transcriptome in HD, which may influence the expression of downstream targets. In addition, we identified p53 tumor suppressor as a transcription factor with a putative role in HD-miRNA deregulation. P53 is found in polyglutamine aggregates both in vivo and in vitro, suggesting an hypofunction of p53 in HD (64). Therefore, p53 may contribute to miRNA expression downregulation in HD.

In summary, this study shows that miRNA variability is a ubiquitous phenomenon in the adult human brain that may influence the mechanism of gene silencing or gene targeting. How the IsomiRs are selectively generated and the possible specific role that these variants have in development, organ physiology or pathological processes, such as HD, opens a new perspective in the field of miRNAs.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We thank the staff of the Ultrasquencing Unit of the CRG for the small-RNA libraries preparation and sequencing on the Illumina/Solexa platform, and the Microarray Unit for the labelling, hybridization and analysis of the Agilent arrays. We also thank Georgia Escaramis for the qPCR data statistical analysis.

FUNDING
The Spanish Ministry of Health ‘Fondo de Investigaciones Sanitarias (PI081367) and “Instituto de Salud Carlos III” (CIBERESP); the Sixth Framework Programme of the European Commission through the SIROCCO integrated project LSHG-CT-2006-037900; the Spanish Ministry of Science and Innovation (SAF2008-00357); Spanish Ministry of Health “Programa Miguel Servet” (to E.M.); Spanish Ministry of Health “Contrato Postdoctoral Sara Borrell” (M.B.-C. and S.P.); Spanish
REFERENCES


