The role of BDNF in Huntington Disease: A targeted analysis of 12 microarray studies

Ronald Xie1; Sharon Yang1; Felix Ma1; Eric Y. Zhao, BSc2,3

Citation: UBCMJ. 2016: 8.1 (17-22)

Abstract

Objectives Huntington disease (HD) is a common hereditary neurodegenerative disorder. Pathogenesis is strongly associated with mutation of the protein huntingtin (HTT). Brain-derived neurotrophic factor (BDNF) is an essential growth factor in neurons and is downregulated in HD. This study focuses on the RE1-silencing transcription factor (REST)/BDNF pathway and provides statistical analysis on expression levels of many genes involved in this pathway in HD and normal subjects.

Methods Twelve recent microarray studies were systematically selected from the Gene Expression Omnibus (GEO). Over-representation analysis was performed on all assayed genes using the Database for Annotation, Visualization and Integrated Discovery (DAVID). Detailed analysis of genes involved in BDNF expression, delivery, and response was performed, and Fischer’s combined probability test was applied to combine findings across the 12 selected studies.

Results Our findings suggest downregulation of BDNF expression in HD-affected patients compared to controls. Analysis of the gene expressions of REST and AKT2 suggests that BDNF expression may be negatively correlated with REST expression and positively correlated with AKT2 expression.

Conclusions Our analysis demonstrates a systematic approach for the use of publicly available microarray data in the analysis of heritable diseases. Our findings suggest that changes in BDNF expression in HD may play a role in HD pathogenesis.

Introduction

Huntington disease (HD) is an autosomal dominant Mendelian disorder characterized by chorea, motor deficits, and cognitive changes. It results from CAG triplet repeat expansion in the huntingtin gene (HTT), with penetrance dependent on repeat number.1–3 The age of onset varies greatly, but averages at roughly 40 years.4

Brain-derived neurotrophic factor (BDNF) is important in synaptic plasticity and neuronal survival in the striatum, a subcortical nucleus of the forebrain with an important role in HD. Downregulation of the gene BDNF has been implicated in HD pathogenesis,5–7 and BDNF-promoting therapies have been proposed for treatment of HD.8,9 This paper investigates the regulatory axis of BDNF expression and its downstream actions. As depicted in Figure 1, translation of mutant huntingtin protein (mHTT) from just one elongated HTT allele is sufficient to cause disruption of many key molecular pathways. One such pathway involves the RE1-silencing transcription factor (REST) which represses the expression of BDNF by binding to the neuron-restrictive silencer element (NRSE) on DNA.10 Normally, the HTT-bound product of REST fails to localize to the nucleus, thereby promoting expression of BDNF by disinhibition. In HD, the important interaction between mHTT and REST is lost, resulting in BDNF downregulation.

However, recent work has suggested that activation of Tropomyosin receptor kinase B (TrkB)-activated signalling cascades may be deficient in HD rather than BDNF delivery.11,12 TrkB is a BDNF-sensitive tyrosine kinase receptor encoded by the gene NTRK2 and has shown involvement in other neurodegenerative disorders such as Alzheimer’s Disease.13,14 It is an upstream activator of key pathways, including the Ras-Raf-Mek-Erk pathway and the PI3K-Akt pathway.15

Downregulation of TrkB may cause inactivation of downstream effectors despite sufficient expression, delivery and localization of BDNF. This raises the potential for a model of “BDNF insensitivity” due to TrkB dysregulation, rather than BDNF insufficiency alone. The Akt pathway, also known as the protein kinase B pathway, has been shown to regulate apoptosis, protein synthesis, and protein degradation. Activation of the Akt pathway has also been shown to be involved in HD16 and has been linked to neuroprotective effects.17,18 Therefore, it is important to distinguish between the causes of downregulation of Akt pathway in HD by examining the relative gene expression of BDNF, its receptor NTRK2, and that of other involved genes.

Leveraging publicly available data is an accessible and informative approach when exploring answers to well-defined biological questions. We propose a general approach for the use of publicly available microarray data and provide a proof of concept by applying it to the question surrounding the roles of BDNF and NTRK2 in HD. Microarray technology enables high-throughput quantification of gene expression at specific genomic loci. Publicly available microarray data, such as those found in the Gene Expression Omnibus (GEO), can be leveraged to rapidly test hypotheses. In this study, we aim to explore the gene expression underlying BDNF delivery and response systems in HD to determine whether they provide support for a model of BDNF insufficiency or BDNF insensitivity. We examined 11 genes relevant to BDNF delivery and response across 12 microarray studies, which were systematically selected from the Gene Expression Omnibus (GEO).19

The 11 genes of interest included genes encoding REST, proteins from the PI3K-Akt pathway, and receptors known to directly interact with BDNF, namely the Trk receptor family and the nerve growth factor receptor (NGFR). These genes were shortlisted based on local pathway analysis to provide preliminary evidence on the question of BDNF insufficiency vs insensitivity, while not diluting statistical power due to the relatively small number of available microarray studies. We analyzed their differential expression and employed multiple

---

1Faculty of Science, The University of British Columbia, Vancouver BC Canada
2Genome Sciences Centre, The BC Cancer Agency, Vancouver BC Canada
3MD/PhD Training Program, Faculties of Medicine and Graduate Studies, The University of British Columbia, Vancouver BC Canada

Correspondence to Eric Y. Zhao (eyzhao@alumni.ubc.ca)
Gene expression analysis

The expression of 11 genes involved in BDNF production, delivery, and response were extracted from data files for the selected HD and control groups of each study. For each study, a student’s t-test was performed to assess the statistical significance of the expression change between wild type and HD affected subjects. The p-values obtained from each study were then combined using the Fisher’s combined probability hypothesis testing to observe population level changes using data from the 12 shortlisted microarray studies. We also performed exploratory over-representation analysis of differentially expressed gene clusters. Our findings agree with previous work and raise questions for future study. As publically available high throughput data becomes more readily available, our approach may guide future studies investigating differential gene expression in HD.

Methods

Data usage and inclusion criteria

Data for this study were taken exclusively from publicly available data repositories. This study therefore did not require ethics approval. Data were downloaded from GEO via their online interface at http://www.ncbi.nlm.nih.gov/geo/. Raw microarray values were parsed and processed using a custom script (parsegeo.py, see supplemental data). Processed data and all subsequent statistical analyses can also be downloaded in the supplemental data for this paper.

GEO studies were shortlisted by searching for the term “Huntington Disease” and including all studies on Homo sapiens and Mus musculus species. Microarray data and sample preparation methodologies from the resulting 17 studies were carefully reviewed for inclusion in our analysis. Of the 17 studies, a total of five were excluded. One human study (GDS4541) was excluded as the tissue studied was whole blood rather than brain. Four mouse studies (GDS2391, GDS4542, GDS4533, and GDS3178) did not have study groups sufficiently relevant to HD. Of the remaining twelve studies, nine were collected from M. musculus and 3 from H. sapiens. Due to the small number of human studies, only M. musculus studies were used for statistical analyses, whereas H. sapiens studies were referenced afterwards to check for concordance of results with findings from the M. musculus studies. In each study, the untreated control group was compared with the HD model group. In studies containing more than one classification of HD, the model producing the most severe phenotype was chosen. Table 1 reports the study IDs and group names for the selected HD and control cohorts, which can be used to replicate this study using data available publicly at http://www.ncbi.nlm.nih.gov/geo/. After these exclusions, a total of 108 HD affected subjects and 118 normal controls across the 12 studies remained.

Table 1 | Selected studies and the names of study groups utilized in this meta-analysis. All data are publicly available via the Gene Expression Omnibus at http://www.ncbi.nlm.nih.gov/geo/, and can be identified using their GEO Study Accessions and available metadata.

<table>
<thead>
<tr>
<th>GEO Study Accession ID</th>
<th>Organism</th>
<th>HD Group Name</th>
<th>Control Group Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDS2169</td>
<td>M. musculus</td>
<td>Nuclear polyQ (n = 11)</td>
<td>Control (n = 24)</td>
</tr>
<tr>
<td>GDS2464</td>
<td>M. musculus</td>
<td>Mutant huntingtin transgene (n = 4)</td>
<td>Wild type (n = 4)</td>
</tr>
<tr>
<td>GDS2911</td>
<td>M. musculus</td>
<td>Untreated (n = 3)</td>
<td>Wild type (n = 6)</td>
</tr>
<tr>
<td>GDS2912</td>
<td>M. musculus</td>
<td>R6/1 transgenic (n = 9)</td>
<td>Wild type (n = 9)</td>
</tr>
<tr>
<td>GDS3620</td>
<td>M. musculus</td>
<td>YAC128 (n = 10)</td>
<td>Wild type (n = 8)</td>
</tr>
<tr>
<td>GDS3621</td>
<td>M. musculus</td>
<td>YAC128 (n = 10)</td>
<td>Wild type (n = 8)</td>
</tr>
<tr>
<td>GDS3935</td>
<td>M. musculus</td>
<td>Hdh GAC knock-in Q111/Q111 (n = 12)</td>
<td>Wild type (n = 12)</td>
</tr>
<tr>
<td>GDS4534</td>
<td>M. musculus</td>
<td>Hdh Q111/111 (n = 6)</td>
<td>Hdh+/+ (n = 6)</td>
</tr>
<tr>
<td>GDS717</td>
<td>M. musculus</td>
<td>R6/2 (n = 7)</td>
<td>Wild type (n = 6)</td>
</tr>
<tr>
<td>GDS1331</td>
<td>H. sapiens</td>
<td>Symptomatic HD (n = 12)</td>
<td>Healthy Control (n = 14)</td>
</tr>
<tr>
<td>GDS1332</td>
<td>H. sapiens</td>
<td>Symptomatic HD (n = 12)</td>
<td>Healthy Control (n = 14)</td>
</tr>
<tr>
<td>GDS2887</td>
<td>H. sapiens</td>
<td>Moderate HD (n = 12)</td>
<td>Healthy Control (n = 10)</td>
</tr>
</tbody>
</table>

Figure 1 | The mechanism of REST-BDNF signalling involves localization of REST to the cytoplasm by binding to the Huntingtin protein, Htt. Mutant Htt (mHtt) fails to localize REST to the cytoplasm, allowing it to inhibit the binding of the transcription complex at the NRSE. Mutations in HTT are thereby associated with decreased expression of BDNF.
Figure 2 | Dot plots of differential expressions of the 6 most notable genes across 12 studies of Huntington disease (HD). Studies were systematically selected from the Gene Expression Omnibus. Microarray data from the 12 selected studies was used to quantify differential expression. Error bars were calculated via error propagation and reflect 95% confidence intervals. When considered independently, any interval not containing 0 implies a statistically significant difference between HD and control groups within the given study. Fisher’s combined probability test was used to combine the p-values of individual studies and calculate an aggregate chi squared and p-value for the differential expressions of each gene. Human studies are shown in yellow, and mouse studies in black. Human studies were not used in statistical analysis.
test to obtain one p-value for each gene. The Fisher's combined probability test is an omnibus test that can be suitably applied in meta-analyses to determine whether a global null hypothesis can be rejected given the results from multiple independent studies. This test differs from other common multiple hypothesis testing methods, such as the ANOVA f-test, due to its robustness against systematic differences across studies arising from usage of different materials, protocols, and microarray sensitivities, and other factors, as long as the tested global hypothesis can be applied to the datasets of each individual study in the same way. Consequently, even if an included study is addressing another hypothesis altogether, as long as the tested global hypothesis can be applied, data from the experiment can be leveraged for meta-analysis. Furthermore, it is not necessary to assume homogeneity of variance across the tested null hypotheses as in the ANOVA f-test. This especially applies to instances where sample sizes differ substantially (often the case for publically available microarray data), where the homogeneity of variance assumption falls apart for ANOVA f-test. Taken together, using the Fisher's combined probability test allows the inclusion more publically available datasets and provides a meaningful way to evaluate a global null hypothesis from datasets produced by independent individual studies.

As the Fisher's combined probability test does not account for the directionality of findings, the p-values of the studies showing change opposite to the hypothesized direction were assumed to be 1.0 (the maximum possible p-value). This allows for simplification of the calculation without relaxing the conditions for significance, yielding a conservative estimate of the final product. The percentage change in the expression level of each gene for each study was calculated and plotted to further visualize the result of the Fisher's combined probability tests. Error bars for percent changes in expression were determined by error propagation calculation and represent 95% t-distribution confidence intervals.

Whereas expression changes for individual samples is commonly reported as fold changes, the mean expression changes reported here are reported as percentage changes. This is because our reported expression values arise from mean values of studies containing three to twelve HD samples and four to twenty four controls each. By the central limit theorem, the standard error of the mean is substantially less variable than the standard deviation of a single subject. This study aims to identify population level trends in HD, which are much smaller in magnitude than drivers of individual disease cases.

**Over-representation analysis of gene expression change**

For every gene assayed in the included microarray experiments, the natural logarithm of the fold change in gene expression was calculated and averaged across studies. Any gene with a fold change of greater than two-times was included. Over-representation analysis was performed using the Functional Annotation Clustering tool in DAVID with the inclusive set of all genes assayed by the microarrays as the background. Upregulated and downregulated genes were analyzed as two separate sets.

**Results**

Differential expression was observed in BDNF, AKT2, and NGFR, however only BDNF displayed concordance in the direction of change between studies. NTRK3 also displayed a concordant trend toward upregulation, but without statistical significance (p = 0.062). Figure 2 shows the 95% confidence interval of percentage change for each gene in each study.

**Figure 3 | Scatter plots of mean percent change in expression in HD in nine mouse studies reveal a positive trend between BDNF and AKT2 (r = 0.58, p = 0.11) and a significant negative correlation between BDNF and REST (r = -0.63, p = 0.029).**

**Differential expression of BDNF**

The nine murine studies examined displayed decreased mean BDNF expression in HD when compared with normal controls. Downregulation was concordant in seven of the nine studies. When observed individually, decreased BDNF expression in one study (GDS2912) was statistically significant (p = 0.001), and two other studies (GDS2169 and GDS3620) were nearly significant (p = 0.076 and p = 0.075 respectively). The studies showed a range from 2% to 42% decrease in BDNF expression with a median decrease of 9.9%. Two studies displayed increased BDNF expression of 5-6%. Fisher's combined test suggests BDNF downregulation when combining our nine murine studies (p = 0.022). Although changes in REST expression were not statistically significant, the change in mean BDNF expression displayed a significant negative correlation (p = 0.029, r = -0.66) with REST expression change (Figure 3).

**Expression changes in NTRK3, PTEN and the AKT pathway**

NTRK3 expression displayed a near-significant upregulated trend (p = 0.062) with concordance in six out of nine murine studies, with a range of 2.5% to 16.5% increased expression and a median increase of 5.5%.

PTEN expression showed upregulation in six out of nine studies ranging from 2.5% to 15.4%, with a median increase of 5.6%. Fisher's combined probability test returned a p value of 0.114. The low chi-
squared value is mostly driven by the high variability of three studies (GDS2464, GDS2912, GDS3620, see supplemental data for a graph of PITE:N expression).

Of the three AKT family genes profiled, AKT2 was the only one which displayed statistically significant differential expression in HD (p = 0.016). Expression was decreased in five out of nine murine studies, ranging from 4% to 35% (median 8%) downregulation. The other four studies showed small increases in expression in HD ranging from 0.17% to 1.3%. NGFR, a negative regulator of the Akt pathway, displayed decreased expression in HD (p = 0.015). However, the results were discordant and appear to be driven by one study (GDS2464).

Results of over-representation analysis of differentially expressed genes

Functional clusters of upregulated and downregulated genes suggest changes in a variety of functions in HD-affected cells (Table 2). Key upregulated processes of interest include metal binding and carbohydrate metabolism. Downregulated processes include immunospecific binding and secretory granule formation among others.

Discussion

Cross-study analysis of the expression in the BDNF and Akt signalling cascades has yielded findings informative of the role of these key pathways in the pathogenesis of HD. Downregulation of BDNF and a negative trend between BDNF and REST expression changes were observed, which agrees with the canonical model of decreased BDNF production in HD. A nearly significant trend was observed in NTRK3, but with small effect size. TrkC, the protein encoded by NTRK3, is not thought to interact directly with BDNF, but does bind to neurotrophic ligands and act on the Akt pathway.

Despite extensive research surrounding the canonical model of BDNF downregulation in HD, recent published work by Plotkin et al. suggests that its delivery to the neurons of the striatum and activation of receptors are normal. Our results did not show any significant changes in expression for NTRK2. Furthermore, PITE:N, an inhibitor of the PI3K-Akt pathway downstream of TrkB, was found to be upregulated with near statistical significance, both of which are concordant with the findings of Plotkin’s study and alludes to the proposed hypothesis that BDNF-related signal transduction may be impaired in HD. However, our findings suggest that the possibility of decreased BDNF expression is not to be discounted. Either hypothesis is supported by evidence that the Akt cascade may be downregulated in HD, and further study is warranted to describe how Akt activation varies with BDNF expression. However, it is worth noting that expression of FASL, a pro-apoptotic factor whose transcription is inhibited by the Akt pathway, was not found to be induced in HD. The genes AKT1 and AKT3 also showed no significant downregulation. These results may potentially weaken our findings relating to the downregulation of Akt2. They may be caused by lower than optimal sample size or more interestingly, may imply another layer of complexity in the interplay between BDNF, Akt and downstream effectors. Ultimately, our most striking finding is the downregulation of BDNF. Even though only three out of the nine studies individually possess confidence intervals not crossing 0, as a whole data combined across the multiple studies revealed an overall statistically significant difference, with strong concordance between studies. This was observed with high effect size and correlation with REST expression. Not only is BDNF downregulation observed in seven of nine murine studies, it is also observed in all three human studies examined (GDS1331, GDS1332, and GDS2887), which hints at a possible common mechanism in human disease.

As nuclear levels of REST are regulated post-translationally via binding to HTT, it is unclear why REST expression would be increased. This study uncovers preliminary evidence of REST overexpression in HD as a potential contributing mechanism of BDNF downregulation which should be replicated and further investigated. Our findings in NGFR, AKT2, and NTRK3 are suggestive of changes in HD, but fail to conclusively demonstrate that transcription-level changes in these genes play active roles in pathogenesis. Further investigation is necessary with greater sample size in order to uncover potential effects.

ORA suggested many affected functions in HD. The upregulation of metal binding proteins points to potential toxicity arising from the binding of metals in neuronal cells. The downregulation of immune targeting processes is in line with the role of mHTT in impairing immune cell migration. Of particular note to this study, the downregulation of genes involved in secretory granule formation may suggest a role of decreased delivery of BDNF among other cell products to their targets.

This paper lays out an approach for the systematic cross-study analysis of a selected set of genes. A limitation of this analysis is its focus on mouse models, which may not demonstrate direct transferability to human disease. Therefore, these results warrant further study in the setting of human HD. Furthermore, the nine mouse studies included utilize different mouse models to represent diseased states, and so may vary significantly in their gene expression profiles. For example, R6/1 and R6/2 mice (used in studies GDS2912 and GDS717) employ truncated mHTT N-terminal fragments with 116 and 144 CAG repeats, respectively (although R6/2 mice can develop up to 250 CAG repeats over time). On the other hand, HdhQ111 mice (used in GDS4534 and GDS3935) employ a full-length HTT knock-in with 111 CAG repeats.

Table 2 | Upregulated and downregulated functional gene clusters based on over-representation analysis using DAVID. A set of significantly upregulated and downregulated genes were used as input for functional classification, and the most significant five upregulated and downregulated clusters are shown.

<table>
<thead>
<tr>
<th>Downregulated Clusters</th>
<th>Enrichment Score</th>
<th>Benjamini P-value Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulation of cell growth, size, transcription, and metabolism</td>
<td>3.197</td>
<td>1.48 x 10^-6 – 0.97</td>
</tr>
<tr>
<td>Immunospecific binding and plasma membrane proteins</td>
<td>2.498</td>
<td>4.3 x 10^-4 – 0.92</td>
</tr>
<tr>
<td>Histone proteins, nucleosome organization, and DNA packing</td>
<td>2.248</td>
<td>7.9 x 10^-5 – 0.99</td>
</tr>
<tr>
<td>Signal transduction: Rhodopsin like G-protein coupled receptors</td>
<td>1.634</td>
<td>2.3 x 10^-4 – 0.57</td>
</tr>
<tr>
<td>Secretory granules and vesicle formation</td>
<td>1.432</td>
<td>7.9 x 10^-4 – 1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Upregulated Clusters</th>
<th>Enrichment Score</th>
<th>Benjamini P-value Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innate immune antimicrobial defense: lysozyme, hydrolase etc.</td>
<td>11.06</td>
<td>1.35 x 10^-21 – 0.85</td>
</tr>
<tr>
<td>Cation/transition metal binding: Binding of zinc ions in particular.</td>
<td>4.175</td>
<td>7.83 x 10^-7 – 0.089</td>
</tr>
<tr>
<td>Carbohydrate metabolism: Glycoside hydrolase family.</td>
<td>3.704</td>
<td>4.05 x 10^-5 – 0.079</td>
</tr>
<tr>
<td>Sterile alpha motif.</td>
<td>3.504</td>
<td>5.00 x 10^-4 – 0.089</td>
</tr>
<tr>
<td>Ribosomal protein subunit.</td>
<td>1.945</td>
<td>6.02 x 10^-5 – 0.96</td>
</tr>
</tbody>
</table>
YAC128 mice (used in GDS3620 and GDS3621) fall into a third class, possessing full-length human HTT with 128 repeats but in a transgenic rather than endogenous model. Different modeled disease stages and CAG repeat lengths may account for some variability in the effect sizes of the included studies. Mouse strains also possess different genetic backgrounds and may express different diseased phenotypes from similar genetic alterations. Standardization is often a challenge, given the phenotypic variability that exists even within a single strain, such as in R6/2 mice.26

However, a potential strength of our methodology is in the synthesis of findings from different models irrespective of disease stage. This strength will be increasingly pronounced with larger numbers of available studies. Due to small sample sizes, we intend for this study to serve primarily as a proof of concept for our methodological approach. In order to allow for interpretation of findings, p-values were not adjusted for multiple hypothesis testing. When a Benjamini-Hochberg adjustment is performed, the findings derived from Fisher's combined probability test are not statistically significant at the 95% confidence level. Since the Benjamini-Hochberg adjustment is often necessary to reduce the probability of false positive findings, we approached interpretation of these results with caution and considered multiple sources of evidence including statistical analysis, concordance between studies, and correlation data when drawing conclusions. More importantly, this paper highlights a generalizable methodological approach which may be applied to larger study sets with success as a convenient screen for early hypothesis generation and testing.

HD remains an incurable disease with a heavy associated social burden, affecting 5 to 10 per 100 000 people worldwide.27 In this study, we investigated the expression of many key genes implicated in the pathogenesis of HD. Changes in expressions of these genes may elucidate possible explanations for the loss of BDNF-dependent post-synaptic plasticity. Our findings provide evidence for the role of impaired BDNF expression, delivery, and response in HD pathogenesis, in coordination with the REST-BDNF system. They also hint at concomitant downregulation of AKT2, and thereby possible involvement of the Akt pathway, which has known roles in HD. These findings support the canonical model of BDNF downregulation and insufficiency. The use of gene expression analysis across studies remains a powerful method for investigating the molecular biology of well-characterized diseases. We have utilized these methods to shed light on processes in HD which inform on cell function, disease formation, and pathology.

Supplementary files are available online at ubcmj.com

References


