Wild-type huntingtin participates in protein trafficking between the Golgi and the extracellular space

Anne N.T. Strehlow1, Jun Z. Li1,2 and Richard M. Myers1,2,*

1Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305-5120, USA and 2Stanford Human Genome Center, 975 California Avenue, Palo Alto, CA 94304, USA

Received October 9, 2006; Revised and Accepted December 9, 2006

Huntington disease (HD) is an autosomal dominant neurodegenerative disease caused by an expanded CAG trinucleotide repeat in the first exon of the HD gene, which results in a toxic polyglutamine stretch within huntingtin, the protein it encodes. Understanding the normal function of this essential protein is vital to understanding the root of the disease, yet despite more than a decade of investigation, its role in the cell remains elusive. Identifying the subcellular localization of huntingtin and understanding its effects on global gene expression are critical to this endeavor. While most reports agree that huntingtin is predominantly a cytoplasmic protein, conflicting distribution patterns have been demonstrated at the subcellular level. Here, we examine wild-type huntingtin’s localization in cultured cells by expressing the full-length human protein tagged with enhanced green fluorescent protein (EGFP) within its unspliced genomic context. In fibrosarcoma and neuroblastoma cells, huntingtin shows discrete punctate, perinuclear localization overlapping largely with the trans-Golgi and cytoplasmic clathrin-coated vesicles, implicating huntingtin in vesicle trafficking. To determine whether huntingtin is involved in trafficking a specific subset of proteins, we measured changes in global transcription levels in embryonic stem cells and neurons lacking huntingtin. Huntingtin null neurons exhibit a significant reduction in transcripts encoding proteins destined for the extracellular space, many of which are components of the extracellular matrix or involved in cellular adhesion, receptor binding and hormone activity. Together, these findings support a role for huntingtin in the intracellular trafficking of proteins required for the construction of the extracellular matrix.

INTRODUCTION

Huntington disease (HD) is a progressive neurodegenerative disorder characterized by progressive motor, cognitive and behavioral dysfunction caused by the inheritance of a single mutated copy of the HD gene (1,2). The disease phenotype ensues when the naturally polymorphic CAG repeat within the first exon of the HD gene is larger than 36 residues, translating a toxic polyglutamine stretch within the protein it encodes, huntingtin (2). In humans, the length of the CAG repeat region is inversely correlated to the age-at-onset of the disease as well as its progression (3,4). Both normal and mutant huntingtins are widely expressed throughout the adult brain and peripheral tissues, yet selective degeneration of GABAergic medium spiny projection neurons occurs in the caudate and putamen of HD patients (2,5–7). Huntingtin also plays a crucial and indispensable role during embryogenesis, as demonstrated by the finding from multiple groups that mouse embryos lacking huntingtin die at approximately embryonic day 7.5 (8–10). Although the initiation of gastrulation is not inhibited, neurulation fails and the somites, nodes and notochord are not formed (8).

Despite its identification more than a decade ago, the function of huntingtin remains largely unclear. Insights have come from studies investigating huntingtin’s interacting partners, subcellular localization and effects on gene expression. Dozens of proteins have been shown to interact with huntingtin, the bulk of which exhibit enhanced binding to the mutant version of the protein. Only a handful prefer associating the wild-type huntingtin (11–23). Many residents of the
nucleus, including nearly a dozen transcription factors, such as CBP [cAMP response element-binding protein (CREB)-binding protein] and Sp1, are among the demonstrated binding partners of huntingtin (24,25). These associations appear to greatly affect the transcriptional profile of HD-affected cells and brain tissue, evidenced by altered global gene expression on microarrays (24,26–29). However, interactions between mutant huntingtin and ubiquitous transcription factors cannot account for gene- and tissue-specific effects of mutant huntingtin. Thus, it is possible that the specific vulnerability of the striatum in HD results from the depletion of wild-type huntingtin or loss of a normal interaction or activity.

Analysis of both the subcellular localization of wild-type huntingtin and its impact on gene expression can provide clues to its potential function. Biochemical and immunohistochemical evidence has shown that wild-type huntingtin in neurons is primarily in the cytoplasm (7,30,31). However, some support exists for huntingtin’s presence within the nucleus (21,32–34). Conflicting associations between huntingtin and a variety of organelles within the cell have been shown: microtubules, Golgi complex, vesicles, axon terminals, endoplasmic reticulum, centrosomes and mitochondria (7,30–32,35–38). The degree of association between the protein and each of these cellular substructures varies widely between studies, as disparate cellular and animal model systems, antibodies and experimental methods have yielded disparate results. The absence of a firm consensus makes accurately predicting huntingtin’s activity in the cell challenging.

We investigated the discrete localization of wild-type huntingtin protein with fluorescently labeled protein expressed from its intact gene on a bacterial artificial chromosome (BAC) transfected into cells. By exploiting a homologous recombination technique with a BAC clone spanning the entire human HD locus, we were able to express full-length wild-type huntingtin tagged with enhanced green fluorescent protein (EGFP) within its natural genomic context. As the endogenous promoter, regulatory elements and splice sites directing the protein’s expression may be crucial in influencing its proper subcellular location, this approach enabled such regulation. As we show here, fluorescent microscopy and immunocytochemistry in various cell types define a punctate, perinuclear localization pattern. The protein’s distribution overlaps largely with the trans-Golgi and cytoplasmic clathrin-coated vesicles, implicating huntingtin in vesicle trafficking between the Golgi complex and plasma membrane.

Applying genome-wide gene expression arrays, we sought to test further whether huntingtin is involved in trafficking a specific subset of proteins from the Golgi. By taking advantage of the viability of both Hdh null mouse embryonic stem (ES) cells and neurons that are formed from these cells by differentiation in culture, we were able to examine the effect of huntingtin-deficiency on global gene expression throughout the differentiation process from ES cells to neurons, up to 10 days post-differentiation, and identify pathways disrupted by the absence of huntingtin. We found that cells lacking huntingtin exhibit a significant reduction in transcripts encoding proteins destined for the extracellular space, many of which are components of the extracellular matrix or involved in cellular adhesion, receptor binding and hormone activity. Furthermore, lysosomal activity and apoptosis are increased, suggestive of elevated cellular stress in Hdh null cells. Together, these findings further support a role for huntingtin in intracellular trafficking, more specifically, transport of a specific subset of proteins required for the construction of the extracellular matrix.

RESULTS

The general strategy for targeted BAC modification

BACs are useful tools for examining genes within their genomic context. By exploiting linear DNA substrate within the bacterial cell, allowing the DNA to be efficiently recombined in vivo (40–42).

We obtained the BAC (RPCI-11 399e10 human male) spanning the human HD locus located at 4p16.3 from the Children’s Hospital Oakland Research Institute (CHORI) (GenBank accession no. AC005516). The HD-containing BAC (HDBAC) contains 166 113 bp of HD sequence from the transcription start site through exon 67; 3268 bp upstream genomic sequence; 30 967 bp downstream genomic sequence and 8823 bp of pBACe3.6 backbone vector sequence (43). We confirmed the integrity of the HDBAC by restriction enzyme analysis (data not shown).

We generated a panel of BACs for investigating huntingtin localization and expression by inserting several cassettes into the HDBAC, within the open reading frame as well as the flanking sequence. To visualize huntingtin by fluorescence, we targeted EGFP to both the 5’- and 3’-ends of HD just following the initiating methionine codon and immediately prior to the transcription stop codon, respectively. To enable the selection of stable transformants in bacterial cells, we inserted a zeocin cassette by targeting nearly 30 kb downstream of the coding region of the HD transcript. The most distant location was chosen to avoid the interruption of any functional regulatory elements immediately distal to HD. As huntingtin expression is modest in most cell types, we replaced huntingtin’s endogenous promoter in one construct with elongation factor I (EF1) (Fig. 1).

Homologous recombination in the BAC requires two successive insertion events. The primary event involves the insertion of a tetracycline resistance (TetR) cassette in the location where the desired final cassette (EGFP, zeocin and EF1) will ultimately reside. We identified recombinants by their ability to grow in the presence of tetracycline. The secondary event involves the replacement of the TetR cassette with the desired cassette (EGFP, zeocin and EF1), resulting in tetracycline-sensitive recombinants (40,42).

To enable homologous recombination, we incorporated 42 bp of target sequence at the 5’ end of each forward and reverse primer. For each cassette to be inserted in the HDBAC, we used two sets of primers, one set for the primary
insertion of the Tet<sup>®</sup> cassette and a second set for the successive insertion of the EGFP, zeocin or EF1 cassette. Following each recombination event, we confirmed integration by PCR, XhoI restriction digest and sequencing of the purified BAC. Restriction digests verified that rearrangements had not taken place during the modification process (Fig. 2C).

**Wild-type huntingtin exhibits punctate, perinuclear localization polarized to a single side of the nucleus**

We analyzed the subcellular distribution of huntingtin following transient transfection of the HDBACs into human HT1080 fibrosarcoma and U87 neuronal/glioblastoma/astrocytoma...
cells. We used fluorescent microscopy to visualize the subcellular distribution of wild-type huntingtin 24, 48, 72 and 96 h post-transfection, enabling adequate time for sufficient expression off the BAC.

Huntingtin expression was variable, but generally low, even after 96 h post-transfection. Given the increased size of the vector, splicing requirements and less-robust promoter activity, we expected and witnessed considerably lower expression off the BAC compared to the positive EGFP plasmid control (44). While no expression was visible at 24 h post-transfection, faint fluorescence gradually became discernable at 48 h, with 72 and 96 h post-transfection yielding the brightest cells. Few of the cells fluoresced as brightly as cells transfected with an EGFP positive control vector. In more modestly fluorescing HT1080 cells, huntingtin-EGFP expression localized around the perimeter of the nucleus in small, punctated structures (Fig. 3A and C). It also appeared to be polarized to a single side of the nucleus. Huntingtin disseminated throughout the cytoplasm in more brightly fluorescing cells, yet remained more highly concentrated around the nucleus with a gradual lessening of expression towards the plasma membrane. In all cases, we noted an absence of full-length wild-type huntingtin in the nucleus, evidenced by a void in the region of nuclear DAPI staining. We observed the same expression pattern in the U87 neuroblastoma cells (Fig. 3B and D).

Transiently transfected HT1080 cells expressing huntingtin from the EF1 high-expressing mammalian promoter on HDBAC/EF1/EGFP(C)/Zeo exhibited considerably higher expression than those expressing the protein from the endogenous promoter. We noted modest expression visible as early as 24 h post-transfection, with a gradual intensifying of fluorescence in the following days. Similar to expression from the endogenous promoter, huntingtin localization was enhanced in the perinuclear region, slanted to a single side of the nucleus (Fig. 3E). However, at all four time-points, we observed signal throughout the entire cytoplasm, but specifically lacking in the nucleus. The lack of expression restricted around the nucleus may simply be due to EF1-regulated over-expression of the protein, saturating the perinuclear region and causing huntingtin to disseminate more quickly away from the cell’s hub. Again, we saw the same expression pattern in the U87 neuroblastoma cells (Fig. 3F).

Given the identical expression patterns for both the N-terminally and C-terminally tagged huntingtin, we propose that the EGFP tag does not interfere with the protein localization or function, and thus represents accurate assessment of its location within the cell. Fluorescent protein expression did not affect cell viability.

**Huntingtin co-localizes with the Golgi and clathrin-coated vesicles in HT1080 and U87 cells**

To determine whether huntingtin’s punctate, perinuclear expression is consistent with the distribution of other subcellular structures, we performed immunocytofluorescence assays on HT1080 and U87 cells transiently transfected with HDBAC/EGFP(C)/Zeo by using a panel of antibodies that recognize discrete substructures in the cell: microtubules, the Golgi and clathrin-coated vesicles.

Staining for the Golgi zone in both U87 and HT1080 cells expressing fluorescently tagged huntingtin indicated that both huntingtin and the Golgi have similar punctate, perinuclear-enhanced expression co-localizing to a single side of the nucleus (Fig. 4A–D, M–P). In all cases, the protein localization was skewed to the same side of the nucleus. In higher fluorescent cells, the regions of most intense huntingtin expression correlated with the regions of Golgi distribution. Co-localization was also apparent between huntingtin and clathrin-coated vesicles. Anti-clathrin showed an enhanced juxtanuclear distribution, with gradually lessening expression towards the plasma membrane, similar to cells highly fluorescent for huntingtin (Fig. 4E–H, Q–T). In cells expressing moderate levels of huntingtin, the biased perinuclear huntingtin fluorescence matches the most intense staining of clathrin.
Staining for tubulin in U87 cells expressing EGFP-tagged huntingtin indicated that huntingtin’s perinuclear-enhanced expression does not strictly co-localize with tubulin’s spindly star-like distribution (Fig. 4I–L). We found that regions of highest huntingtin expression near the nucleus are the regions relatively devoid of tubulin labeling. As huntingtin overlaps with the Golgi, some of this may be due to the Golgi’s location alongside the centrosome, the organizing...
center from which microtubules polymerize outward (45). Therefore, the void in tubulin staining overlapping the dense Golgi complex is consistent with intense huntingtin fluorescence.

The demonstrated codistribution of full-length wild-type huntingtin expressed from the human HDDBAC with both the Golgi complex and clathrin-coated vesicles in both transiently transfected fibrosarcoma cells and neuroblastoma cells support a role for wild-type huntingtin in the endocytotic pathway.

**Hdh null neurons exhibit more transcriptional changes than do Hdh null ES cells**

As subcellular localization hinted at a role for huntingtin in intracellular transport, we sought to determine whether huntingtin is involved in trafficking a specific subset of proteins. To test this hypothesis, we analyzed changes in global gene expression in Hdh null ES cells and Hdh null differentiated neurons—cultured cell lines that lack huntingtin expression—at 6, 8 and 10 days post-differentiation compared to similar wild-type control cells (these cells are described in Materials and Methods). For each time-point, three biological replicates were subjected to separate culturing, differentiation and total RNA extraction. We reserved aliquots of the RNA samples for quantitative PCR (qPCR) confirmation of a select number of transcripts (see Confirmation of Bead Chip Array Findings below), using the remainder for hybridization to Illumina Sentrix MouseRef-8 BeadChip expression arrays. We compared average normalized fluorescence measurements for the three Hdh null replicates to the average normalized fluorescence measurements for the three wild-type replicates at each time-point. We assessed gene expression differences by using the paired Student’s t test statistic, enabling the fold-change difference between the two measurements to be calculated. This method allowed us to identify mRNAs that were comparatively increased or decreased in the Hdh null ES cells and differentiated neurons relative to controls. mRNAs that met more than 2-fold expression elevation or repression for at least one time-point at $P \leq 0.05$ were explored further. Using the 2-fold stringency cutoff, a cumulative total of 472 transcripts on the array platform demonstrated a significant difference between Hdh null and wild-type cells across all four time-points, representing ~2% of the probes included on the array. The number of genes with altered expression in cells lacking huntingtin was comparatively fewer (80) in the ES cells compared to the differentiated neuronal cells. The number of mRNAs exhibiting more than 2-fold change gradually increased as the neuronal differentiation process progressed, with the array detecting 199, 207 and 249 transcripts at day 6, 8 and 10, respectively.

**Global assessment of gene expression changes in Hdh null ES cells and neurons**

We observed temporal patterns of gene expression among the 24,045 transcripts on the array. Clustering the transcripts by related expression with Cluster and TreeView software generated large blocks of genes following similar profiles in undifferentiated ES cells and differentiated neurons (Fig. 5A). Comparing the number of transcripts with differential expression across all four time-points, this global perspective highlights the relative paucity of mRNAs in Hdh null ES cells (UD) with altered expression, suggesting that at this very early stage the cells are less affected—at the level of transcription—by the absence of huntingtin than are differentiated neurons. This is in agreement with the known expression levels of huntingtin, which indicate that ES cells express low, but detectable, levels of huntingtin compared to the relatively high concentrations observed in the brain (6, 8, 9, 46, 47). It is also apparent that although many genes exhibit an elevation or repression in expression across several time-points or
Eighty unique transcripts showed a greater than 2-fold elevated or repressed expression in Hdh null embryonic stem cells compared to wild-type embryonic stem cells (P < 0.05). The transcripts fit into several classes of interest: protein degradation, extracellular matrix, cell cycle and patterning/developmental regulation. Up arrows indicate induction and down arrows indicate repression. Numbers represent transcript fold-change.

**Table 1. Transcripts with differential expression between Hdh null embryonic stem cells and wild-type embryonic stem cells grouped according to functional activity**

<table>
<thead>
<tr>
<th>Patterning/development</th>
<th>Lefty/Lefty</th>
<th>Orthodenticle homolog 2</th>
<th>Pem (Rhox5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>¶ 2.70</td>
<td>Lefty/Lefty</td>
<td>Orthodenticle homolog 2</td>
<td>Pem (Rhox5)</td>
</tr>
<tr>
<td>¶ 3.81</td>
<td>Orthodenticle homolog 2</td>
<td>Orthodenticle homolog 2</td>
<td>Pem (Rhox5)</td>
</tr>
<tr>
<td>¶ 4.54</td>
<td>Orthodenticle homolog 2</td>
<td>Orthodenticle homolog 2</td>
<td>Pem (Rhox5)</td>
</tr>
</tbody>
</table>

**Protein degradation**

† 7.24  
*Cops8*—Constitutive polymorphic (COP9) homologue subunit 8: subunit of the 19S lid of the 26S proteasome for degradation of ubiquitinated proteins; positively regulates E3 ubiquitin ligases

† 4.36  
*Psmdh*—Proteasome 26S subunit: subunit of the 26S proteasome for degradation of ubiquitinated proteins

† 2.63  
*Manba*—Mannosidase, beta A; lysosomal: glycoprotein catabolism in the lysosome

† 2.05  
*Ifi30*—Interferon gamma-inducible protein 30: disulfide oxidoreductase activity in the lysosome

† 1.69  
*Ube2l1*—Ubiquitin-conjugating enzyme E2L: catalyzes the covalent attachment of ubiquitin-like protein SUMO-1 to other proteins

**Extracellular matrix**

† 2.13  
*B3galnt6*—UDP-Gal:betaGal beta 1,3-galactosyltransferase; polypeptide 6: synthesis of carbohydrate moieties for proteins essential to basement membranes

¶ 2.74  
*Adam23*—A disintegrin and metallopeptidase domain 23: cell—cell and cell—matrix interactions

¶ 3.74  
*Col4a1*—Procollagen, type IV; Alpha 1: structural component of basement membranes conferring tensile strength

¶ 3.99  
*Col4a2*—Procollagen, type IV; Alpha 2: structural component of basement membranes conferring tensile strength

**Cell division**

† 8.71  
*Plk1*—Polo-like kinase 1: involved in cell division

† 3.50  
*Cdc5*—Coiled-coil domain containing 5: regulator of spindle function during cell division

† 2.28  
*Aurka*—Aurora kinase C: organizes microtubules in relation to the centrosome/spindle pole during mitosis

† 3.03  
*Pttg1*—Pituitary tumor-transforming 1: chromosome segregation during mitosis

Eighty unique transcripts showed a greater than 2-fold elevated or repressed expression in Hdh null embryonic stem cells compared to wild-type embryonic stem cells (P < 0.05). The transcripts fit into several classes of interest: protein degradation, extracellular matrix, cell cycle and patterning/developmental regulation. Up arrows indicate induction and down arrows indicate repression. Numbers represent transcript fold-change.

**Hdh null ES cells exhibit diminished expression of vital patterning genes and increased lysosomal activity**

Eighty transcripts fit the criterion of at least a 2-fold change in expression with P ≤ 0.05 [91% of these 80 transcripts (75) have P ≤ 0.01 and 83% (66) have P ≤ 0.001] in Hdh null ES cells compared to wild-type controls. A quarter of these transcripts represent genes of unknown or inferred function based on homology, with the majority of this subset identified as RIKEN cDNA clones. The remaining genes fall into several functional categories: patterning, extracellular matrix, protein degradation and cell cycle (Table 1).

Table 1. Transcripts with differential expression between Hdh null embryonic stem cells and wild-type embryonic stem cells grouped according to functional activity

<table>
<thead>
<tr>
<th>Patterning/development</th>
<th>Lefty/Lefty—Left right determination factor 1: morphogen involved in left–right axis patterning</th>
<th>Orthodenticle homolog 2: transcription factor involved in anterior/posterior axis patterning</th>
<th>Pem (Rhox5)—Reproductive homeobox 5: transcription factor involved in patterning/marker of extraembryonic tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>¶ 2.70</td>
<td>Lefty/Lefty—Left right determination factor 1: morphogen involved in left–right axis patterning</td>
<td>Orthodenticle homolog 2: transcription factor involved in anterior/posterior axis patterning</td>
<td>Pem (Rhox5)—Reproductive homeobox 5: transcription factor involved in patterning/marker of extraembryonic tissue</td>
</tr>
<tr>
<td>¶ 3.81</td>
<td>Orthodenticle homolog 2: transcription factor involved in anterior/posterior axis patterning</td>
<td>Orthodenticle homolog 2: transcription factor involved in anterior/posterior axis patterning</td>
<td>Pem (Rhox5)—Reproductive homeobox 5: transcription factor involved in patterning/marker of extraembryonic tissue</td>
</tr>
<tr>
<td>¶ 4.54</td>
<td>Orthodenticle homolog 2: transcription factor involved in anterior/posterior axis patterning</td>
<td>Orthodenticle homolog 2: transcription factor involved in anterior/posterior axis patterning</td>
<td>Pem (Rhox5)—Reproductive homeobox 5: transcription factor involved in patterning/marker of extraembryonic tissue</td>
</tr>
</tbody>
</table>

Eighty unique transcripts showed a greater than 2-fold elevated or repressed expression in Hdh null embryonic stem cells compared to wild-type embryonic stem cells (P < 0.05). The transcripts fit into several classes of interest: protein degradation, extracellular matrix, cell cycle and patterning/developmental regulation. Up arrows indicate induction and down arrows indicate repression. Numbers represent transcript fold-change.

Eighty transcripts fit the criterion of at least a 2-fold change in expression with P ≤ 0.05 [91% of these 80 transcripts (75) have P ≤ 0.01 and 83% (66) have P ≤ 0.001] in Hdh null ES cells compared to wild-type controls. A quarter of these transcripts represent genes of unknown or inferred function based on homology, with the majority of this subset identified as RIKEN cDNA clones. The remaining genes fall into several functional categories: patterning, extracellular matrix, protein degradation and cell cycle (Table 1).

Three genes with key roles in patterning the developing embryo exhibited differential expression in Hdh null ES cells compared to wild-type controls. *Lefty1* (*LeftB*), a morphogen required for proper left-right axis asymmetry of developing organ systems, was down-regulated. *Lefty1* is also known to be involved in restricting Nodal localization in the developing embryo without affecting its transcript levels. In accordance, no change in *Nodal* mRNA expression was detected on our array. Similarly, *Otx2*, involved in anterior-posterior patterning and neuronal development, and *Pem* (*Rhox5*), a marker of extraembryonic tissue, were both repressed. Our findings are in agreement with a recently published non-quantitative in situ hybridization approach, which demonstrated a down-regulation of *Lefty*, *Otx2* and *Pem* mRNA in E7.5 *Hdh* knock-out mice. The same study also indicated that Nodal was dispersed away from its normal restricted region of localization, though its expression level remained unchanged (48). The observed dysregulation of these transcripts is intriguing. Although their skewed expression has little effect on the viability of *Hdh* ES cells that remain healthy in culture after many successive splittings, the absence of huntingtin in the early stages of differentiation and development is devastating to the embryo, ultimately resulting in early prenatal lethality (8–10).

We noted an elevation in transcript levels of several genes involved in protein degradation or lysosomal function in Hdh null ES cells. *Cops8* and *Psmd8*, components of the 26S proteasome involved in the degradation of ubiquitinated proteins, and *Ube2l1*, a ubiquitin-conjugating enzyme, were increased several-fold over control cells. Glycoprotein catabolism and disulfide oxidoreductase activity within the lysosome may also be elevated, evidenced by the increased mRNA levels of *Manba* and *Ifi30*, respectively.

Several genes (*Plk1*, *Cdc5*, *Aurka* and *Pttg1*) with known involvement in cell cycle regulation exhibited altered expression in Hdh null ES cells, suggestive of impaired cell division. Once again, misregulation may have little influence on cells in the undifferentiated state, but more serious repercussions may result during the repetitive cell cycling required during development, perhaps indicated by the sustained elevation of *Plk1* and *Cdc5* throughout the entire neuronal differentiation process.

We also identified four transcripts with involvement in the formation of the extracellular matrix: two co-regulated sheet-forming collagens (*Col4a1* and *Col4a2*) found in basement
membranes, *Adam23*, a metallopeptidase involved in cell-matrix interactions, and *B3galts6*, required for the synthesis of proteins essential to the basement membrane. Although these functional categories do not represent statistically significant gene ontology (GO) groupings, likely due to the limited number of transcripts with differential expression in *Hdh* null undifferentiated cells, their early misregulation in ES cells suggests they function as precursors or instigators of increased cellular dysfunction in more specialized cell types. It may be important to note that altered transcript levels do not necessarily imply altered protein levels due to possible compensatory mechanisms by the cell.

**Deficiencies in extracellular proteins in *Hdh* null neurons suggests faulty formation of the extracellular matrix and interrupted receptor binding and signaling**

Hundreds of genes exhibited differential gene expression between *Hdh* null and wild-type cells through the progression from ES cells into neuronal cells. A large subset of the transcripts remained down-regulated throughout the entire time-course with a lesser number exhibiting sustained elevation from day 6 to day 10 post-differentiation. Several significant GO categories can be pulled from the data, indicating the possible disruption of several functional pathways in the absence of huntingtin. Using GoStat software to identify biological, molecular and physiological classes of genes, we found that several over-represented gene groups: extracellular space (*P* = 1.77e−13), receptor binding (*P* = 0.000291), hormone activity (*P* = 0.00101), enzyme inhibitor activity (*P* = 0.00101), extracellular matrix (*P* = 0.0025), structural components conferring tensile strength (*P* = 0.0326), retinol/retinal/retinoid/isoprenoid binding (*P* = 0.0114, *P* = 0.0363, *P* = 0.0435, *P* = 0.0435 respectively) and lysosome activity (*P* = 0.0379). By further investigating the function of detected transcripts via NCBI and Expasy, several genes involved in cell adhesion, axonogenesis, action within the synaptic region, transport and apoptosis were readily apparent (Fig. 6).

**Extracellular space.** The most striking observation is the widespread down-regulation of transcripts whose protein products make their way to the extracellular space (see Fig. 6 for the complete list). Many of the transcripts encode for secreted proteins that function outside the cell’s interior, such as *ApoA1* (cholesterol transport), *Sfrp1* (Wnt signaling) and *Ltf* (iron binding and transport), and proteins integral to the plasma membrane, including *Fxyd6* and *Prom1*. A large subset of this broader class includes genes involved in the construction of the extracellular matrix, scaffolding to the plasma membrane and cell adhesion. Six transcripts from the collagen family (*Coll4a1, Coll4a2, Coll6a1, Coll6a3, Coll16a1* and *Coll18a1*) were down-regulated in *Hdh* null neurons. As collagen fibers are crucial for strengthening and organizing the matrix, their diminished presence is suggestive of a deficit in structural architecture and integrity of the cell’s exterior. Likewise, the transcripts of proteins (*Adams2, Mmp2, Adams4* and *Dcn*) required for the cleavage of pro-collagen prior to fibril assembly and matrix degradation, thus mediating its organization and plasticity, are similarly decreased. Perhaps in response to the down-regulation of matrix metalloproteinases, the repression of several metalloproteinase inhibitors (*Serpinb9f* and *Timp3*) was also detected on the arrays. The array also detected diminished expression for a dozen adhesion molecules (*Tgfb1, Mifap4, Adam23, Tgm2, Emn3, Spp1, Cldn6, Thbs2, Epb41l3, Tpm2, Mucedh1 and Prelp*) involved in mediating cell–cell and cell–matrix interactions in differentiated neurons lacking huntingtin. Insufficient amounts of these binding proteins hint at a weakened ability to anchor cells to one another or tether them to the extracellular meshwork in the absence of huntingtin.

Interestingly, *Hs3st5a1, B3galts6, Chst1* and *Mgat3*, enzymes that function in the Golgi to synthesize carbohydrate moieties for proteins essential to basement membranes, proteoglycans and other components of the extracellular matrix, showed elevated expression in *Hdh* null cells. Increased attempts at Golgi glycosylation for hyaluronan and proteoglycans (*Hapl3, Dcn, Agrin, Gpc1, Thbs2*)—despite their reduced expression—may be indicative of a compensatory mechanism by the cell in an attempt to increase production of these vital structural components.

**Receptor binding, hormone activity and signaling.** The down-regulation of transcripts involved in cellular adherence and the construction and binding activities of the extracellular matrix may also interfere with proper receptor function, thus interrupting signaling and hormonal pathways. The vast majority of transcripts involved in receptor binding and hormonal activity detected on the arrays showed diminished expression in *Hdh* null neurons. In addition, six genes involved in the Wnt signaling pathway [*En1, Sfrp1, Dkk3, Wisp1, IGF-II* and *Csnk1e (CK1e)*], including several Wnt target genes, demonstrated similar down-regulation (Fig. 6) (49–52). The Wnt signaling molecules play an important role in early brain development, and recent evidence has demonstrated that these molecules require the presence of heparan sulfate proteoglycans (HSPGs) to regulate their pathways (53,54). The depletion or disruption of HSPGs may be responsible for initiating secondary signal transduction defects in *Hdh* null neuronal cells.

**Axonogenesis and activity with the synapse.** mRNAs with roles in axon guidance and neurite extension are subject to dysregulation in *Hdh* null neurons (Fig. 6). A large set of transcripts [*Sox21, Nav1, Nfatc4, Ptn (HB-GAM), Lgal5, St6gal1, Sema3f, Srgap2*] is indicative of diminished or aberrant outgrowth and extension. The arrays detected a similar dysregulation of transcripts encoding for proteins functioning within the synapse, specifically the down-regulation of mRNAs (*Adra2b, Rab17, Kcnq4, Csnk1e*) involved in the regulation of neurotransmitter release or uptake at the synaptic terminals. While no single neurotransmitter signaling pathway was over-represented on the arrays in the absence of huntingtin, the altered transcription of a group of synapse-enriched transcripts (*Clic6, Pdelb, Slec1a1, Kcnq4, Csnk1e*) indicates interruption of dopamine signaling or potential excitotoxicity via increased activity of glutamate receptors. Furthermore, *Pdelb* and *Csnk1e* are enriched within the striatum, the brain region most susceptible to damage in HD, suggesting that the loss of huntingtin function may contribute to the
**Figure 6.** Significant GO categories between *Hdh* null and wild-type cells. Each colored panel represents the average normalized fluorescence signals from the BeadChips of the *Hdh* null samples compared to wild-type samples. The transcripts fit into several classes of interest: extracellular space, extracellular matrix/cell adhesion, receptor binding activity, lysosomal activity, enzyme inhibitor activity, growth and proliferation, axonogenesis and action within the synapse, Wnt signaling and transport. (elevated expression *Hdh* null cells—yellow, repressed expression in *Hdh* cells—blue).
excitotoxic damage observed in HD patients. In response to
the potential oxidative injury, Gsta1 and Gsta2, detoxification
proteins that protect cells following excess glutamate
exposure, showed up-regulation in Hdh null neurons.

Increased lysosomal and apoptotic activity and decreased
cell growth and proliferation in Hdh null neurons

**Lysosome activity.** A subset of proteins misregulated in Hdh
null neuronal cells has known involvement in lysosomal func-
tion and protein degradation. Four genes (Caps8, Psmd8,
Ube21 and Fhxo44) encoding participants in the ubiquitin–
proteasome pathway maintained elevated transcript levels
throughout the neuronal differentiation. Likewise, Gaa,
Manba and Fuca1, all involved in the breakdown of glyco-
proteins within the lysosome, showed an enhanced expression.
Increased lysosomal activity may be indicative of an over-
representation of misdirected or misfolded proteins within
the cell’s cytoplasm, which are ultimately targeted for
destruction.

**Apoptosis.** Cells expressing the polyglutamine-expanded
version of huntingtin have been shown to be more susceptible
to cell death, evidenced by massive neurodegeneration in
later stages of Huntingtin disease and transcriptional acti-
vation of apoptosis-inducing genes in affected tissues and
cells (5,55,56). Embryos lacking huntingtin also exhibit
pyknotic nuclei and enhanced apoptosis prior to death
at embryonic day E7.5, suggesting that both presence of
polyglutamine-expanded huntingtin and the loss of wild-type
huntingtin lead to cell death at the organismal level (9,10).
We noted a similar presence of apoptosis-related transcripts
in our array, detecting elevated expression of several cell
death-inducing transcripts [Tnx11, Prdx2, Casp3, Pdc8,
Ern2 (Inv2)]. Although arrays are not strictly quantitative,
the transcripts showed a general rise in expression in Hdh
null cells over the length of the time-course, culminating
in their highest levels at day 10 post-differentiation. Furth-
more, the arrays detected a repression of transcripts encoding
protective proteins, such as Gpx3, that guards against
oxidative damage, and the cathepsin proteinase inhibitor,
Serpin3g.

**Growth, proliferation and differentiation.** Despite the ability
to maintain Hdh null ES cells and differentiated neurons
in culture, huntingtin is vital for growth and development
given that mouse embryos lacking the protein are underde-
veloped and smaller in size (9,10). We confirmed this essential
role for huntingtin on our arrays, detecting the general
down-regulation of genes involved in growth, proliferation
and cellular differentiation (Fig. 6). A subset of these
transcripts appear to function specifically within the brain:
Serpinfl, Emp3, Tcfap2c, Dlx3 and Elavl4. Many general
growth-related or mitogenic mRNAs (Ghrh, Sgcb3a1,
Cdkn1c, Sp6, Gkn1, Igf2) also exhibited reduced expression.
Conversely, Hdh null neurons consistently overexpressed
Ccng2, a negative regulator of cell proliferation and cell
cycle progression.

**Iron regulation and intracellular transport
disrupted in Hdh null neurons**

**Transport.** Iron transport has been previously shown to be
askew in Hdh null cells and mouse embryos (57,58). We
also detected transcripts hinting at aberrant metal ion transport
in cells lacking huntingtin (Fig. 6). We observed increases in
two transcripts (Picalm, Cubn) involved in iron uptake at the
plasma membrane, as well as repression of three others
(Sle40a1, Hamp, Ltf) responsible for the regulation of iron
export and binding. The increased uptake of iron into the
cells coinciding with the diminished ability to expel it may
result in iron overload and toxicity in Hdh null cells. Increased
amounts of iron in the brain has been linked to the occurrence
of neurodegenerative disease (59,60). This is in agreement
with the reported neurodegeneration observed adult mice
lacking huntingtin strictly within the forebrain (61). The
arrays also indicated the misregulation of transcripts involved
in zinc and copper transport.

In addition to the transport of iron across the plasma mem-
brane, we noted the dysregulation of several transcripts
involved in intracellular trafficking in Hdh null cells. Although
our localization results hint at a role for huntingtin in intra-
cellular transport, the absence of the wild-type protein was
not assumed to affect the transcription of transport-related pro-
teins. However, both Dna1c1, a component of the microtubule-
base dynein motor protein, and Rab17, a regulator of
receptor-mediated transcytosis and endosome recycling,
showed reduced expression throughout the differentiation
process. With increased expression in Hdh, null neurons
were Abca5, a lysosomal ATPase involved in transport, and
Gdi3, an inhibitor of vesicle docking on the Golgi. Collect-
ively, these findings indicate a possible impairment of
vesicle transport throughout the cell.

**Confirmation of microarray findings**

To confirm the changes in expression detected by the arrays,
we performed qPCR analyses on 12 transcripts, for genes
that were both up-regulated and down-regulated to varying
degrees (Fig. 7). Like the RNA samples hybridized to the
arrays, the measurements from qPCR were average values
from the three original biological replicates and a pooled
sample. The expression of each of the 12 transcripts were
evaluated or repressed in the same manner detected on the array,
although the fold-change difference between Hdh null and
wild-type cells was often higher with qPCR. We observed a
similar fold-change discrepancy in previous experiments
within the lab, and it is likely attributable to binding saturation
on the array and/or increased sensitivity and specificity of
qPCR. The Pearson’s r-value was measured at 0.98 between
the two methods, thus confirming the accuracy and specificity
of the array data with regard to differences in transcript level.

**DISCUSSION**

By using novel approaches, we linked huntingtin’s perinuclear
subcellular localization with a potential role in trafficking a
discrete set of proteins between the Golgi and the extracellular
space. We homologously recombined an EGFP fluorescent tag
Confirmation of a dozen transcripts indicates that array results (white bars) can be
demonstrated to be correlated with qPCR data from the same experiments.
Threshold curves (four replicates) indicate a repression [Camk2b (e) and
Col4a1 (d)] or elevation [Mcm6 (e) and Card4 (f)] of expression in Hdh
null ES cells for several transcripts. Threshold curve data were used to calcu-
late the fold-change difference between Hdh null and wild-type cells. (B) Con-
firmation of a dozen transcripts indicates that array results (white bars) can be
replicated by qPCR (black bars). Pearson’s correlation score ($r = 0.98$) con-
firms that the array data and qPCR data are highly correlated.

**Figure 7.** Confirmation of array findings: qPCR measurements of selected
transcripts exhibiting altered expression in Hdh null compared to wild-type
ES cells. (A) qPCR values for a random sample of 12 transcripts were normal-
ized to positive control transcripts—GAPDH (a) and $\beta$-actin (b)—which
demonstrate very similar expression between Hdh null and wild-type cells.
Threshold curves (four replicates) indicate a repression [Camk2b (e) and
Col4a1 (d)] or elevation [Mcm6 (e) and Card4 (f)] of expression in Hdh
null ES cells for several transcripts. Threshold curve data were used to calcu-
late the fold-change difference between Hdh null and wild-type cells. (B) Con-
firmation of a dozen transcripts indicates that array results (white bars) can be
replicated by qPCR (black bars). Pearson’s correlation score ($r = 0.98$) con-
firms that the array data and qPCR data are highly correlated.

into a BAC carrying the chromosomal region spanning the
entire human HD gene. Expression of huntingtin-EGFP in cul-
tured fibrosarcoma and neuroblastoma cells indicates that full-
length, wild-type huntingtin localizes to the perinuclear region
in a punctate vesicular-like pattern polarized to a single side of
the nucleus. Immunocytochemistry demonstrates that this
pattern mimics the distribution of the Golgi and clathrin-coated
cytosolic vesicles, which also show enhanced localization
around the nucleus. We did not observe fluorescence within
the nucleus. Our results fit well with a subset of previously published
reports of both the localization of huntingtin itself, as well as the distribution of several of the protein’s binding part-
ers. Early reports using immunocytochemistry detected hun-
tingtin diffusely throughout the cytoplasm, with evidence of
its association with a variety of intracellular organelles, includ-
ing the microtubules, endoplasmic reticulum, mitochondria,
vesicular structures and Golgi complex (7,31,35). Several
reports also identified huntingtin within the nucleus, with one
report suggesting that huntingtin may shuttle between
the nucleus and the cytoplasm (32,33,62). Follow-up studies
suggest that huntingtin’s localization is more restricted,
focused in a discrete region encircling the nucleus. However,
only selective antibodies have detected this pattern (38,63).
As different antibodies have elicited a range of localization pat-
terns, likely due to cross-reactivity and their unique affinities
for huntingtin, our use of a fluorescent tag bypassed the require-
ment for antibodies. With this alternative approach, our localization
studies of full-length, wild-type huntingtin under the control of its endogenous local regulatory elements corroborate
the latter perinuclear findings.

Analysis of the subcellular localization of huntingtin pro-
vides some clues to its potential function. The combined distri-
bution pattern from the earlier reports prompted speculation of
huntingtin’s role in vesicle trafficking of the secretory and
dendritic pathways (7,31,35,38,63). Our observed perinuclear
co-localization of huntingtin with the Golgi and clathrin-coated
vesicles meld well with several others lines of evidence invol-
viting huntingtin in vesicular transport. Huntingtin has been
shown to associate with various proteins that play a role in
intracellular trafficking, including HAP1, HIP1, HIP14, $\alpha$-adaptin C (HYPJ), SH3-containing GRB-like protein 3
(SH3GL3), Fip2, PACSIN1 and Cdc42-interacting protein 4
(CIP4) (16,18,19,64–68).

To determine if huntingtin may be involved in trafficking a
specific subset of proteins, we evaluated changes in global
transcription levels in ES cells and neurons lacking huntingtin.
We identified genes with aberrant expression using data from Illumina BeadChip arrays, which contain probes for more than
24 000 human transcripts, enabling us to identify classes of
genes that fit into distinct functional networks and that
require huntingtin for normal activity. Our data suggest that
the disruption in transcripts encoding crucial components of
the extracellular space, primarily structural building blocks
and modifiers of the extracellular matrix, serve as the primary
functional disturbance in Hdh null cells. The down-
regulation of many collagens, adhesion molecules and metal-
lloproteinases may ultimately contribute to the downstream
interference of receptor binding and signaling pathways,
most of which were also decreased in transcript levels in the
mutant cells. Likewise, we detected misregulation of many
transcripts with roles in axon or dendrite extension or that
function within the synapse in Hdh null cells compared to
control cells. We also noted an increase in transcript levels
of genes involved in lysosomal and apoptotic-related activity.
These results fall in line with reports of increased cell death
and neurodegeneration in mouse embryos and adult brains
lacking huntingtin expression (8–10,61,69).

Gene expression of Hdh null ES cells appeared minimally
affected by the absence of huntingtin. This is not wholly unex-
pected given their prolonged viability in culture (8–10).
However, we detected a down-regulation of several transcripts
involved in early patterning of the developing embryo and
reconfirmed this observation by qPCR. These results are in
accordance with recent in situ hybridization results demonstr-
ing misregulation of these same mRNAs (48). Although
it has been shown previously that Hdh null ES cells can be dif-
ferentiated into largely functional neurons, we indicate here
that these neurons exhibit a more definitive dysfunction at
the transcriptional level, made apparent by the altered expres-
sion of many classes of genes (46).

This study provides the first evidence that the absence of hun-
tingtin expression results in a disruption of many components
of the extracellular matrix and basal lamina. These supporting
structures play pivotal roles in the maintenance and activity
of the cell, influencing localized activation of signaling path-
ways, internal cytoskeletal rearrangements and organization
of the matrix into specialized basement membranes. Depending
on the cell type and the context, these interactions drive cell
fate, cell migration and cellular architecture, such as polariz-
ation and process extension (70). Thus, the down-regulation
of key components of the matrix may ultimately affect the
cell’s overall stability and its ability to interact with neighbor-
ing cells. Such intercellular communication is particularly
important for the activity and viability of neurons.
As the matrix plays a supportive role in signal transduction, the repressed expression of many matrix components may contribute to the observed interference of receptor binding and signaling pathways, most of which demonstrated a similar repression. Of notable interest, several mRNAs of the Wnt signaling pathway showed diminished expression in *Hdh* null neuronal cells. Accumulating evidence demonstrates that the signaling activity (ligand levels, distribution, receptor-interactions and expression of downstream target genes) of the Wnts rely upon the presence of HSPGs, cell surface and extracellular matrix macromolecules that consist of long, sulfated sugar moieties attached to a protein core (53,54). Furthermore, just as HSPGs are major components of the mammalian central nervous system, Wnts have been shown to play important roles in early brain development (71,72). While these signaling disruptions may not represent primary defects of huntingtin loss, they may be suggestive of the secondary effects of a weakened extracellular matrix and our observed repressed expression of agrin and glypican, two key HSPGs. These pathways have not been previously implicated in HD or huntingtin-related activity, and thus provide new avenues for future research.

In addition to their role in signal transduction, proteoglycans have been implicated in many important neuronal functions, such as synaptic remodeling, neural plasticity, axon pathfinding and neurite outgrowth (71). Likewise, brain matrix metalloproteinases, which are involved in the modification of the extracellular matrix, have also been shown to be important regulators of many of these same biological processes (73). We detected many dysregulated transcripts involved in synaptic activity (endocytosis and neurotransmitter release) and axon and neurite extension on our arrays. The differential expression of decorin (chondroitin sulphate proteoglycan) and agrin and glypican (HSPGs), as well as many proteoglycan-binding proteins (i.e. Prolargin, HB-CAM, Collagen XVIII) indicates that repression of these cell-surface and extracellular matrix proteins, as well as its remodeling, may negatively influence these neuronal processes in *Hdh* null differentiated neurons. Although differentiated neurons lacking huntingtin appear normal, small morphological variations are possible and may not affect their short-term viability in culture. Given the critical role axonogenesis and neurite outgrowth play in the development of the brain, the dysregulation of many of the brain-enriched transcripts may help explain the early neurological defects and disorganization observed in mice embryos lacking huntingtin.

The statistically significant decreased expression of proteins whose actions are in the extracellular space in *Hdh* null differentiated neurons suggest that there is a defect in intracellular transport in the absence of huntingtin. Our results coincide with a role for huntingtin in vesicular trafficking and intracellular transport. More specifically, huntingtin may participate in trafficking a subset of proteins destined for the extracellular matrix, transporting them from the Golgi, where they are modified, to the plasma membrane via clathrin-coated vesicles. We hypothesize that inability of these proteins to reach the cell’s exterior could result in their build-up within the cytoplasm, and in the cell’s attempt to achieve homeostasis, their expression is down-regulated. The wayward proteins are eventually targeted for degradation in the lysosome, accounting for the increase in transcripts involved in lysosomal activity and protein degradation.

This study cannot rule out the possibility that huntingtin functions as a transcription factor or indirectly modulates transcription. Our localization studies suggest that the native protein is restricted to the cytoplasm, a finding that conflicts with a handful of other reports (32,33). We attribute this incongruity to the specificity of the two differing techniques. Directly observing EGFP-tagged huntingtin off the BAC maintains its endogenous regulation and eliminates variable binding capabilities and cross-reactivity of unique antibody epitopes. Thus, given its localization within the cytoplasm, huntingtin would have to mediate gene expression outside the nucleus, possibly by sequestering transcription factors within the cytoplasm. While huntingtin has been shown to bind to a variety of known transcription factors—NcoR, CBP, Sp1, TAF1, TBP, RAP30, p231HBP/HYPB, REST, CtBP—the vast majority of these associated factors exhibit enhanced binding to the mutant version of huntingtin, bringing into question whether these associations are functional with wild-type huntingtin in vivo under normal physiological conditions (7,11,74). Wild-type huntingtin appears to bind to the transcriptional co-repressor REST in the cytoplasm of neuronal cells, possibly preventing REST from entering the nucleus and restricting expression of its target genes, including transcription of bone-derived neurotrophic factor (BDNF) (23,75). One might expect the sequestration of REST within the cytoplasm to be lost in neurons lacking huntingtin, resulting in a down-regulation of BDNF. However, we did not detect any change in the mRNA levels of BDNF at day 6, 8 or 10 post-differentiation. No change would be expected in *Hdh* null ES cells, as REST localizes to the nucleus in non-neuronal cells. It may be important to note that the array probe spanned a region of the 5’ untranslated region upstream of Exon II, where REST regulation of the neurotrophin occurs.

This work refines the role of wild-type huntingtin within the cell. Our results suggest its specific involvement in the transport of extracellular proteins to the cell’s exterior, indicating that many of the effects of huntingtin’s absence may be secondary to this disruption. Our findings open the door for future investigation into the link between huntingtin and proteoglycan distribution, as well as its role in signal transduction.

**MATERIALS AND METHODS**

**Bacterial cells and BACs**

A BAC—clone 399e10 from the RPCI-11 Human Male BAC Library—spanning the entire human *HD* gene as well as 3268 bp upstream and 31 009 bp downstream was acquired from CHORI (76). The human DNA is cloned in the pBACe3.6 (8823 bp) backbone and selected in a growth media containing 20 μg/ml chloramphenicol (Cam) (Sigma, Cat. no. C-4881).

A temperature-sensitive recombination-proficient *E. coli* strain, EL250, was used for all BAC modification steps. The EL250 strain carries a defective λ prophage that enables targeted recombination while protecting the electroporated linear DNA substrate (42).
BAC DNA preparation

For transfection in EL250 E. coli, BAC DNA was prepared using the Qiagen Plasmid Mega Kit according to the standard protocol (Qiagen, Cat. no. 12183). One liter BAC cultures were grown overnight at 37°C in LB medium containing 20 μg/ml chloramphenicol. The BAC DNA pellet was air dried for ~15 min before being dissolved in 500 μl TE-buffer, pH 8.0 for 20 min. Samples were stored at −30°C till use. XhoI digests (New England BioLabs, Cat. no. R0146S) were performed to confirm the integrity and purity of the BAC DNA using ~400 ng purified BAC DNA in 25 μl total reaction volume incubated at 37°C for 4 h. Gels were electrophoresed overnight at 25 V in 0.5% agarose.

Transforming EL250 cells with the BAC clone

In order to modify the BAC, it was first transformed into the EL250 bacterial cell line to enable temperature-sensitive homologous recombination. Overnight, 3 ml cultures were grown from fresh single EL250 colonies at 32°C. Starter cultures were diluted 50-fold in 100 ml LB medium and grown at 32°C until OD (A600) reached mid-log phase (~0.4). The culture was incubated on ice for 20 min, divided into five pre-chilled 50 ml tubes and centrifuged at 5000 r.p.m. for 10 min at room temperature (Beckman JA10 rotor). The supernatant was removed, and cells were resuspended in 3 ml cold, sterile Millipore water before being aliquoted into two 1.5 ml microfuge tubes on ice. Tubes were centrifuged at full-speed for 5 min at 4°C using a benchtop centrifuge. After rinsing and spinning the pellets two additional times with 1.5 ml cold, sterile water, a final spin was performed to remove all traces of supernatant. The pellets were resuspended in 35 μl 10% glycerol, snap-frozen on dry ice and stored at −80°C until needed. This large culture prepared a total of 10 reactions.

Transformation of the BAC was performed by electroporation of 500 ng circular, supercoiled BAC into 35 μl electrocompetent cells thawed on ice in pre-chilled 0.1 cm gap cuvettes (BIORAD, Cat. no. 165–2089) using a Bio-Rad Gene Pulser (BIORAD, Cat. no. 12183). One liter BAC cultures were grown from fresh single colonies at 32°C in LB medium supplemented with 20 μg/ml Cam, diluted 50-fold in 100 ml LB (20 μg/ml Cam) and grown at 32°C until OD (A600) reached mid-log phase (0.3–0.4).

Preparation of linear DNA cassettes with homology ends

A series of linear cassettes were created for recombination into the HDBAC: EGFP for both the 5’-end and 3’-end of the HD gene, zeocin-resistance (Zeo+) selectable marker, EF1 mammalian promoter and the tetracycline-resistance (TetR) selectable marker. Standard PCR conditions were used to amplify linear DNA fragments using the Finnzyme Polymerase system (DyNAzyme EXT DNA Polymerase, Cat. no. F-505L).

A series of TetR cassettes 2,451 bp in length were made for the initial ‘Positive Selection’ homologous recombination events, in which the gene encoding tetracycline-resistance was inserted into locations where it would ultimately be replaced by EGFP, Zeo+, or EF1. The secondary ‘negative selection’ homologous recombination events replacing the TetR cassette with EGFP, Zeo+ or EF1 were selected for by tetracycline-sensitivity. Chimeric primers were designed containing two domains: 42 bp homologous to flanking regions on the target BAC where the cassette was to be inserted and 21 bp for amplifying the cassette from plasmid template. Therefore, two sets of chimeric primers were designed for the insertion of EGFP, SV40-Zeo+ and EF1. Whereas the BAC homology arms are identical in both primer pairs for a single cassette, one pair contains 21 bp for amplifying TetR and one pair contains 21 bp for amplifying the replacement cassette. The linear DNA products were amplified using standard PCR conditions with 3.75 mM MgCl2 and 5% DMSO.

For all primers listed, nucleotides in italics are homologous to the targeted BAC sequence and those in roman type are homologous to insertion cassettes. For each positive selection event, the tetracycline-resistant cassette was amplified using primers 5’TGTTTGTGATTCCCTTGCGAGACGGGACGAGGAGCCGAGG and 5’TTCAGCCAGGGCTGAAGCTGCAGCGCCTAGA for 3’-end of the cassette with EGFP, Zeo+ and EF1.

Preparation of recombination-proficient, electroporation-competent, EL250 cells

Overnight cultures of the EL250 bacterial cell line carrying the HDBAC were grown from fresh single colonies at 32°C in an LB medium supplemented with 20 μg/ml Cam, diluted 50-fold in 100 ml LB (20 μg/ml Cam) and grown at 32°C until OD (A600) reached mid-log phase (0.3–0.4).

Human Molecular Genetics, 2007, Vol. 16, No. 4
The cultures were immediately transferred to a 43°C shaking waterbath for 20 min to induce expression of recombination proteins: Beta, Exo and Gam. The culture was swirled in ice water for 20 min to cool the culture quickly, divided into five prechilled 50 ml tubes and centrifuged at 5000 r.p.m. for 10 min at 4°C (Beckman JA10 rotor). The supernatant was removed and the cells were resuspended in 3 ml cold, sterile Millipore water before being aliquoted into two 1.5 ml microfuge tubes on ice. Tubes were centrifuged at full-speed for 5 min at 4°C in a benchtop centrifuge. After rinsing and spinning the pellets two additional times with 1.5 ml cold, sterile water, a final spin was performed to remove all traces of supernatant. The pellets were resuspended in 35 μl 10% glycerol, snap-froze on dry ice, and stored at −80°C until needed. This protocol was adapted from Yu et al. (77). For BACs with an integrated Zeo cassette, 25 μg/ml zeocin (Invitrogen, Cat. no. 46-0509) was added to the medium.

Cell transformation with linear cassettes for homologous recombination

**Step I—Positive selection.** Transformation of the linear DNA cassettes were performed by electroporation using 100–300 ng linear Tet cassette DNA into 35 μl of heat-induced EL250 electrocompetent cells thawed on ice. The cells were transferred into pre-chilled 0.1 cm gap cuvettes and electroporated with a BioRad gene pulser set at 1.75 kV, 25 μF and 200 ohms. One ml of SOC medium was added immediately following electroporation. Cells were incubated at 32°C for 90 min with shaking and 500 μl of cell culture was spread on two large LB agar plates supplemented with Cam (20 μg/ml) and Tet (7.5 μg/ml). This protocol was adapted from Lee et al. (42). Tetracycline-resistant colonies are picked 36 h post-transformation for confirmation of cassette integration.

**Step II—Negative selection.** Transformation of replacement cassettes (EGFP, Zeo* or EF1) was performed by electroporation using 100–300 ng linear replacement cassette DNA in 35 μl of heat-induced EL250 competent cells carrying the Tet*-BAC thawed on ice. Following electroporation as described above, cells were incubated at 32°C for 90 min with shaking and 500 μl were spread on two large LB agar plates (9 g agar, 5 g tryptone, 2.5 g yeast extract, 5 g sodium chloride, 5 g NaH2PO4 H2O, 8 ml chlorotetracycline (6.3 mg/ml), 156 μl Cam (20 μg/ml), 125 μl Zeo (100 mg/ml) and 500 ml water) and incubated at room temperature overnight. The media was supplemented with zeocin (25 μg/ml) if a zeocin cassette had been previously recombined. The faint lawn was gently scraped from the surface of the agar into 10 ml LB broth (10⁻¹ dilution) using a disposable cell scraper (Costar, Cat. no. 3010). Using glass beads, 100 μl of 10⁻³ and 10⁻⁴ dilutions were plated on several similar LB agar plates supplemented with Cam (6.25 μg/ml), chlorotetracycline (50 μg/ml) and fresh fusaric acid (12 or 24 μg/ml) (Sigma, Cat. no. F-6513). As a control to estimate titer of viable cells, 100 μl of a 10⁻⁷ dilution was plated on LB agar plates supplemented with Cam (6.25 μg/ml) only. The plates were incubated at 32°C for 48 h. Individual tetracycline-sensitive colonies were picked for screening. This protocol was adapted from Yang et al. (40) and Lee et al. (42).

Confirmation of recombinants. After both Steps I and II, recombinants were screened and verified by three methods: PCR amplification around the region of cassette insertion, XhoI restriction digest and sequencing.

To PCR around the integration site, the potential-positive BACs were purified using the Qiagen R.E.A.L. Prep 96 Plasmid kit (QIAGEN, Cat. no. 26171). Two ml LB cultures supplemented with 20 μg/ml Cam were inoculated with single Tet colonies and incubated for 16 h at 32°C with shaking at 150 r.p.m. in a Precision shallow form shaking bath. Instructions were followed precisely, with the exception that all inversion steps were eliminated to avoid cross-contamination between neighboring colonies. DNA pellets were air-dried for 20 min and redissolved in 35 μl 10 mM Tris-CI, pH 8.5 by incubating overnight at room temperature. PCR was carried out with 3.75 mM Mg²⁺, 5% DMSO and 0.75 μl purified BAC DNA template from 96-well BAC prep using primers flanking the insertion site.

Clones with proper integration were prepped using the Qiagen Mega Prep Kit as described above. To ensure that the BAC did not undergo rearrangement during recombination, a XhoI restriction digest using 300–400 ng purified BAC DNA in 25 μl total reaction volume confirmed the expected digest pattern. Rearrangement-free clones were sequenced around the region of the BAC encompassing the inserted cassette using the ABI 377 Sequencer.

**Cultured cell lines**

Human U87 neuronal/glioblastoma/astrocytoma cells were grown in Dulbecco’s Modified Eagle’s minimum essential media (Gibco, Cat. no. 10569–010) with GlutaMax I, high Glucose, 110 mg/L sodium pyruvate and pyridoxine-HCl, supplemented with 10% fetal bovine serum (HyClone, Cat. no. SH30071.03) and 1% penicillin/streptomycin (Gibco, Cat. no. 15140–122). Human HT1080 fibrosarcoma cells were grown in Eagle’s DMEM supplemented with 4 mM L-glutamine and Earle’s BSS, 1.5 g/l sodium bicarbonate, 0.1 mM non-essential amino acids, 4.5 g/l L-Glucose, 10% fetal bovine serum and 1% penicillin/streptomycin. Cells were grown at 37°C in 5% CO₂ in 25 cm² culture flasks and split 1:4 (U87) or 1:20 (Ht1080) upon 85% confluency using 0.25% Trypsin-EDTA (Gibco, Cat. no. 25200–056). Cells were fed every 2 days following split.

**Liposomal transfection of eukaryotic cells with BAC DNA**

Chamber slides were coated with 700 μl Poly-d-Lysine (Invitrogen, Cat. no. P-6407) diluted to 0.1 mg/ml and incubated for 24 h at room temperature. Prior to plating, the Poly-d-Lysine was removed and slides were washed three times with tissue culture-grade distilled water. Cells were trypsinized at 80% confluency from one 25 ml flask with 0.25% Trypsin/EDTA and harvested by centrifugation. The cells were resuspended in 5 ml culture media. Cell density was counted using Trypan Blue and 2.5 × 10⁻⁴ to 7.5 × 10⁻⁴ cells were plated per well in four-well chamber slides. Additional media was added for a total of 800 μl, and cells were grown at 37°C in 5% CO₂ overnight. Depending on the number of cells plated, cell density was approximately
30–50% confluent after an overnight incubation. The following day, 3:2 FuGene 6 Reagent:DNA complex reactions were prepared using 3 μl FuGene (Roche, Cat. no. 1814075)-1 μg PBS negative control/1 μg EGFP-C1 positive control/2 μg BAC DNA (HDBAC/EGFP(C)/Zeocin, HDBAC/EGFP(N)/Zeocin or HDBAC/EF1/EGFP(C)/Zeocin) in a total reaction volume of 100 μl OPTI-MEM lacking serum (Gibco, Cat. no. 31985-070). Preparation of complexes followed the standard protocol. After a 30–40 min incubation, the complex mixture was added drop-wise into each well and gently swirled to mix. Media was changed daily. Cells were incubated for 24–96 h before fluorescence was captured.

**Fluorescent microscopy**

Cells were visualized under the fluorescent microscope (Zeiss AxioSkop2) at 400× magnification. To fix cells, chamber slides were washed three times with room temperature PBS and incubated in 750 μl 4% paraformaldehyde for 30 min at room temperature in the dark. Paraformaldehyde was removed, the slides were washed three times with PBS, and cells were incubated with DAPI diluted 1:1000 in PBS for 5 min. Vectashield mounting media (Vector Labs, Cat. no. H-1000) and cover slips were applied.

**Fluorescent immunocytochemistry**

Cells were fixed with paraformaldehyde as described above. Paraformaldehyde was removed, the slides were washed three times with PBS, and the cells were incubated in 500 μl 5% goat serum blocking buffer for 1 h. Blocking buffer was removed and cells were incubated in 500 μl primary antibody diluted in PBS for 1 h: anti-β-tubulin (final concentration diluted 1:500 in PBS, 2 μg/ml) (monoclonal, clone KMX-1, Chemicon, Cat. no. MAB3408), anti-Golgi zone (final concentration diluted 1:30 in PBS) (monoclonal, clone 371-4, Chemicon, Cat. no. MAB1271) and anti-clathrin heavy chain (final concentration diluted 1:100 in PBS, 2.5 μg/ml) (BD Biosciences, Cat. no. 610499). Primary antibody was removed, cells were washed four times with PBS and 500 μl Alexa Fluor 555 goat anti-mouse IgG secondary antibody (final concentration diluted 1:400 in PBS, 5 μg/ml) (Molecular Probes, Cat. no. A-21422) diluted in PBS was added to each well. Cells were incubated for 1 h. Secondary antibody was removed, cells were washed four times with PBS and incubated with DAPI (final concentration diluted 1:1000 in PBS) for 5 min. Following four washes with PBS, mounting media and coverslips were applied. All incubations were performed in the dark at room temperature.

**Growth of ES cells**

Mouse ES cells were grown on inactivated hygromycin-resistant 129 mouse embryonic fibroblasts (MEFs) [Cell and Molecular Technology (Specialty Media), Cat. no. PMEF-H]. One day prior to plating or splitting the ES cells, a vial of MEFs was thawed and plated into four 25 cm tissue culture flasks (~1.3 × 10^6 cells/flask) coated with 0.2% gelatin (Sigma, Cat. no. G1393) and allowed to settle overnight in Dulbecco’s minimum essential media (Gibco, Cat. no. 11960-069) supplemented with high glucose, lacking pyridoxine hydrochloride, L-Glutamine and sodium pyruvate, 10% fetal bovine serum (Gibco, Cat. no. 16000-044), 1% 100× penicillin/streptomycin (Gibco, Cat. no. 15140-122) and 1% 100× GlutaMax (Gibco, Cat. no. 35050-061).

**Hdh** null mouse ES cells (courtesy of Scott Zeitlin) and wild-type EmbryoMax Strain 129/SVEV ES cells [Cell and Molecular Technology (Specialty Media), Cat. no. CMTI-1] were plated in the MEF-containing flasks in Dulbecco’s minimum essential media supplemented with high glucose, lacking pyridoxine hydrochloride, L-Glutamine and sodium pyruvate, 10% fetal bovine serum, 10% neonatal calf serum (Gibco, Cat. no. 16010-159), 1% 100× penicillin/streptomycin, 0.1% ESGRO Leukemia Inhibitory Factor (LIF) (Chemicon, Cat. no. ESG1106) and 7 × 10^-6% β-mercaptoethanol (Sigma, Cat. no. M7522). Upon 85% confluency, the media was removed from flasks and cells were washed with warm PBS before incubating in 1 ml warm 0.05% Trypsin-EDTA (Gibco, Cat. no. 25300-054) at 37°C for 10 min. Cells were removed by pipetting with 5 ml media and centrifuged to pellet cells and remove Trypsin-EDTA. After removal of the supernatant, cells were resuspended in 15 ml warm growth media and each 25 cc flask was replated into a single new gelatin-coated 75 cc flask. Splitting reduced MEFs to ~4.6 × 10^5 cells/flask. ES cells were not fed the following day post-plating, after which the media was replaced daily. ES cells were either collected at the undifferentiated state for array analysis or differentiated into neurons (see below).

To collect cells in the undifferentiated state, media and cells were removed from three 75 cc flasks upon 85% confluency for both Hdh null and wild-type ES cells, as described above, and replated in 0.2% gelatin-coated flasks for 1 h to remove feeder cells. The supernatant containing the stem cells was siphoned off and spun down. One ml TRIzolReagent (Gibco, Cat. no. 15596-018) was added to the cell pellets and allowed to incubate for 5 min before dislodging and homogenizing the cells by pipetting. Cell suspension was collected and transferred into a QIAshredder (Qiagen, Cat. no. 79654) in 600 μl aliquots. After spinning for 2 min, the eluate was collected in a 1.7 ml tube and stored at −80°C until RNA preparation.

**Neuronal differentiation of ES cells**

For differentiation into neurons, cells were prepared as described above. For each time-point, the media, ES cells and MEFs were removed from three 75 cc flasks each for both Hdh null and wild-type cell cultures, as described above. The cell suspension was replated in 0.2% gelatin-coated flasks for 1 h to remove feeder cells, the supernatant was siphoned off and 1.5 × 10^6 cells were transferred in single-cell suspension onto non-coated bacteriologic dishes, cultured in the presence of differentiation medium containing Dulbecco’s minimum essential media, 10% fetal bovine serum, 10% neonatal calf serum, 2 mM glutamine, 1 mM sodium pyruvate, 0.1 mM non-essential amino acids, 0.1 mM β-mercaptoethanol and penicillin/streptomycin. Under these conditions, embryoid bodies were formed and cultured for 4 days in the absence of, then followed for 4 days in the presence
of $0.5 \times 10^{-6}$M all-trans-retinoic acid (Sigma, Cat. no. R2625).
Subsequently, 8-day EBs were transferred onto 0.1% gelatinized
tissue culture plates and cultured for an additional 6, 8 and 10
days (three plates each) prior to collection. Dishes were
washed with PBS to remove the media and 1 ml of Trizol
was added to each dish. Cells incubated in the TRIzol® reagent for
5 min to enable the dislodging and homogenization of the cells
by pipetting. Cell suspension was collected and transferred
into a QIAshredder in 600 µl aliquots. After spinning for
2 min, the eluate was collected in a 1.7 mL tube and stored at
$-80^\circ$C until RNA preparation.

RNA extraction from mouse ES cells and neuronal
cells for microarray analysis

The RNA preparation for all samples—undifferentiated ES
cells and differentiated neurons—were completed together to
avoid variation in sample preparation. TRIzol®-homogenized
samples were thawed on ice before adding 200 µl of chloro-
form per 1 ml of TRIzol® to each tube (~240 µl/tube).
Samples were mixed thoroughly and allowed to settle for 5
min at room temperature before centrifuging at 3500 r.p.m.
for 15 min at 4ºC in a tabletop centrifuge (IEC Centra CL3).
The aqueous phase was transferred to a new 14 ml round-
bottom tube. Total RNA was collected according to manu-
facturer’s instructions using the RNAeasy Mini Kit (Qiagen,
Cat. no. 74104). RNA was eluted in 150 µl RNase-free
water. The concentrations were calculated before storing RNA
at $-80^\circ$C until RNA amplification step. A total of 24 samples
(six undifferentiated ES cells and 18 differentiated neurons)
were prepared.

Illumina chip labeling and hybridization

We used the Illumina Sentrix® MouseRef-8 Expression Bead-
Chips (Illumina, Cat. no. BD-26–201), covering over 24 000
well-annotated mouse RefSeq transcripts, for our gene
expression comparisons. A single round of RNA amplifica-
tion is required prior to hybridization of the RNA to the array. The
Illumina TotalPrep RNA Amplification Kit (Ambion, Cat. no.
IL1791-1) generated biotinylated, amplified RNA. The manu-
facturer’s recommended instructions were followed with only
slight deviations. To synthesize First Strand cDNA by reverse
transcription, we used 250 ng total RNA from each sample
collected above. Following the Second Strand cDNA synthesis
and cDNA purification steps, the in vitro transcription to syn-
thesize cRNA was prepared and allowed to incubate overnight
for 15 h. The cRNA was purified according to instructions and
eluted in 100 µl nuclease-free water. A total of 850 ng of RNA
was aliquoted into tubes and brought up in a total volume
of 42.5 µl before concentrating samples by vacuum centrifugation
for 1 h. Samples were resuspended in 11.3 µl RNase-free water
and kept on ice until hybridization. For hybridization to
three Illumina Sentrix® MouseRef-8 BeadChips, the protocol
using included reagents were followed precisely. Samples
were arranged on the chips so that the three biological replicates
of each time point were spread across three separate arrays,
thus controlling for possible chip—chip variation. Chips were
read by the Illumina BeadStation 500× and collected data
was reviewed for analysis.

Statistical analysis

The raw data were processed by the Illumina BeadStudio Data
Analysis Software, which generated the averaged intensity
signals across ~30 beads for each transcript. We imported
the intensity data for all 24 000 transcripts across 24
samples into R-values, log transformed with base of 2, and
quantile normalized the data to generate an equal intensity dis-
tribution across the 24 samples (78). Using the normalized
data, we calculated a Two-way ANOVA with Strain (Hdh
null, wild-type) and Stage (undifferentiated, neuronal day 6,
neuronal day 8, neuronal day 10) as factors, to identify trans-
scripts showing strain-dependent or stage-dependent expression
patterns. Two-group comparisons using the Student’s t test
identified, separately, transcripts that were differentially
expressed between Hdh null cells and 129 wild-type cells at
each of the four stages: undifferentiated ES cells, neuronal
day 6, neuronal day 8 and neuronal day 10. For each of the
four comparisons, we recorded the t scores, P-values and fold-
change (expressed in log scale). This data was further filtered,
with a stringency of 2-fold elevation or repression in Hdh
null cells compared to wild-type cells at $P \leq 0.05$. Up-regulated
expression was color-coded yellow, and down-regulated
expression was color-coded blue.

qPCR confirmation of differentially expressed transcripts

Select transcripts demonstrating at least a 2-fold increase or
2-fold decrease in gene expression in Hdh null ES cells com-
pared to wild-type ES cells were chosen for confirmation using
real-time PCR analysis. Primers for qPCR were designed to
exonic regions of each gene, generating amplicons 60–80 bp
in length with similar melting temperatures (62°C). After con-
firming that the amplicons aligned to the proper genomic
region and did not form secondary structure characteristics
using the University of Santa Cruz In-Silico PCR program,
primers were tested by Real-Time using mouse genomic
DNA (50, 5 and 0.5 ng) to verify that they produced gene-
specific products of an appropriate size and that proper stan-
dard curves were generated based on observed threshold
cycle numbers (Tc). GAPDH and β-actin housekeeping
genes were used as positive controls.

qPCR used the same RNA preparations used for the original
array experiments. For each sample, 4 µg of RNA was treated
with amplification grade DNase I following manufacturer’s
protocol (Invitrogen, Cat. no. 18068-015). The treated RNA
was then converted to cDNA using Invitrogen’s SuperScript
III First-Strand Synthesis System for RT-PCR (Invitrogen,
Cat. no. 18080-051) following the manufacturer’s protocol.
Separate cDNA preparations amplified with oligo(dT)$_{12-18}$
and random hexamers were combined prior to real-time PCR.

The following primers (Operon) were used for amplification
of mouse cDNA (all primers are listed 5’ to 3’): GAPDH—
ACATCAAGAAGGTGGTGAACAG/GACAACCTGGTCC
TCAGTGAGC; β-actin—CAAGTACTGTGTGGATCT
GGTG/ACATCTGCTGGAAGGTGGACAG;
Mcm6 (accession
no. NM_008567.1)—TGAACAACCTTCTCAATGATC
Hdh (accession
no. NM_172729.1)—GACTCCAAGTTCATGCTGTGC
Mest (accession
no. NM_008567.1)—TGAAAACACTTCTCAATGATC
GTC/CTCTGAGGAATCTGAGACAGA; Card4 (accession
no. NM_172729.1)—GACTTACGTTCACCGGGTG
TA/TGTTGATATAGGTTCCCTCAGC; Mest (accession
no. NM_008567.1).
To calculate the corresponding fold-change between the values to normalize for differences in cDNA sample aliquots. The normalized Tc value for the DTc value was determined by subtracting the average normalized null and wild-type expression levels for each transcript, the null ES cells, Doug Mortlock and Greg Barsh for providing the EL250 recombination-proficient cells and TetBH2.4 #5 plasmid, respectively, Greg Barsh, Patrick Brown and Devin Absher for helpful discussions, and Larissa Tsavaler, Patrick Collins, Nathan Trinklein and Loan Nguyen for technical assistance and support. This research was supported by an award from the Wills Foundation and Training Grant T32 GM 07790-25 from the General Medical Sciences Institute at the National Institutes of Health.

Conflict of interest statement. The authors declare that they have had no involvements that might raise the question of bias in the work reported or in the conclusions, implications or opinions stated.

REFERENCES


