Research Article

Molecular Systems Biology

A meta-analysis of transcriptomic profiles from Huntington's disease patients substantiates the pathophysiological role of CDC42, NFY, DLX1 and PRMT3

Manuel Seefelder[1] and Stefan Kochanek[1]

[1] Department of Gene Therapy, Ulm University, D-89081 Ulm, Germany

Corresponding author: Manuel Seefelder. E-mail: manuel.seefelder@uni-ulm.de

Abstract (175 words)

Description of robust transcriptomic alterations in Huntington's disease is essential to identify targets for biochemical studies and drug development. Here, we analysed publicly available transcriptome data from the brain and blood of 449 HD patients and 212 healthy controls. We identified 737 and 661 genes with robustly altered mRNA levels in the brain and blood of HD patients, respectively. In the brain, a subnetwork of 320 genes strongly correlated with HD and was enriched in transport-related genes. Bioinformatical analysis of this subnetwork highlighted CDC42, PAK1, NFY, DLX1, HMGN3, and PRMT3. Moreover, we found that CREB1 can regulate 78.0 % of genes whose mRNA levels correlated with HD in the blood of patients. Our meta-analysis indicates that alterations in protein transport, metabolism, transcriptional regulation, and CDC42-mediated functions are central features of HD. Further our data substantiate the role of transcriptional regulators that have not been reported in the context of HD such as DLX1, HMGN3 and PRMT3 and strongly suggest transcriptional dysregulation of NFY and its target genes across tissues.

Keywords: Huntingtin / neurodegeneration / weighted gene co-expression network analyses / robust rank aggregation analysis

Introduction

Huntingtin (HTT) functions in diverse cellular processes such as autophagy, endocytosis, vessicle transport, and transcriptional regulation (Saudou & Humbert, 2016). A triplet repeat expansion in exon 1 of the HTT gene results in the expansion of an N-terminal polyglutamine tract and causes Huntington's disease (HD) (The Huntington's Disease Collaborative Research Group, 1993). Clinically, a progressive loss of motor functions, cognitive impairment, and psychiatric symptoms such as depression and anxiety (Ross & Tabrizi, 2011) characterises HD. Besides neurological symptoms, HD patients suffer from a plethora of non-neuronal symptoms such as cardiac failure, muscle atrophy, impaired glucose tolerance, osteoporosis, weight loss, and testicular atrophy (van der Burg et al, 2009).

Expansion of the N-terminal polyQ tract impairs the multi-faceted function of HTT and its interaction with numerous other proteins (Shirasaki et al, 2012; Zuccato & Cattaneo, 2014). Mutant huntingtin (mHTT), for instance, induces the activation of microglia, leading to increased secretion of interleukin-1β (IL-1β), tumour necrosis factor-alpha (TNF) and increased levels of reactive oxygen and nitrogen species (Crotti & Glass, 2015). Tabrizi et al. and Fan and Raymond showed that mHTT impairs the glutamate uptake in astrocytes leading to excitotoxicity (Tabrizi et al, 1999; Fan & Raymond, 2007). Aberrant splicing of the mHTT mRNA results in the formation of a truncated HTT exon-1 protein forming nuclear and cytoplasmic inclusions (Gipson et al, 2013). R6/2 mice with a knock-in of exon-1 of human HTT show a more severe disease progression than mouse models with a knock-in of full-length mutant HTT (Gipson et al, 2013). The study of several huntingtin-interaction partners and their impaired function in HD further suggested impaired trafficking of clathrin-coated and non-coated vesicles in HD patients (Velier et al, 1998; Harjes & Wanker, 2003). Transcriptomic studies of HD patients, cell lines, and mouse models expressing mHTT observed transcriptional dysregulation of a plethora of genes (Borovecki et al, 2005; Runne et al, 2007; Hu et al, 2011; Narayanan et al, 2014; Labadorf et al, 2015; Labadorf et al, 2017; Agus et al, 2019), although the identified differentially expressed genes (DEGs) differed strongly among the studies. Several mechanisms behind the broad transcriptional dysregulation such as altered expression of enhancer RNAs (Le Gras et al, 2017), sequestration of transcription factors (e.g. CREB1, TBP, or mSin3a) (Cha, 2000; McCampbell et al, 2000; Steffan et al, 2000; Sugars & Rubinsztein, 2003), or the sequestration of proteins such as the muscleblind-like splicing regulator 1 (MBNL1), nucleolin, and proteins of the small interfering RNA (siRNA) machinery (Nalavade et al, 2013; Martí, 2016) have been

discussed.

As noted before, previously published analyses of transcriptomic profiles from HD patients yielded varying results. Since a thorough knowledge of pathological mechanisms behind HD is essential for the design of further biochemical studies and development of therapies, we performed a meta-analysis of publicly available transcriptomic data from HD patients to identify genes altered in several studies. Within our meta-analysis, we found 661 or 737 genes with robustly altered mRNA levels in the blood and brain of HD patients, respectively. Strongly suggesting that dysfunction in protein transport and metabolism are central in HD, we identified by weighted gene co-expression network analysis a subnetwork of 320 genes, enriched in genes functioning in protein transport, that strongly correlated with HD in the brain. Additionally, we identified the cell division cycle 42 (CDC42), p21 (CDC42 / RAC1) Activated Kinase 1 (PAK1), 14-3-3 protein eta (YWHAH), and protein phosphatase-2 catalytic subunit α (PP2CA) as hub genes of this subnetwork. Transcription factor enrichment analysis (TFEA) highlighted distal-less homeobox 1 (DLX1), high mobility group nucleosomal binding domain 3 (HMGN3), and protein arginine methyltransferase 3 (PRMT3) in this subnetwork. Similarly, a subnetwork of 118 genes, including genes coding for constituents of the Arp2/Arp3 complex, were significantly altered in the blood of HD patients. Strikingly, 78.0% of the genes in this blood subnetwork were direct or indirect targets of the CREB1.

Results

Transcriptional changes in the brain of HD patients

Since neurological and neuropsychiatric symptoms are the pathognomonic features of HD and possess a high disease burden for HD patients, several transcriptomic studies investigated transcriptional changes in the brain of HD patients. In this meta-analysis, we included three published transcriptomic studies using post-mortem brain tissue from the prefrontal cortex (NCBI accession number: GSE33000 and GSE64810) (Narayanan et al, 2014; Labadorf et al, 2015) and the caudate nucleus of prodromal HD patients (NCBI accession number GSE129473) (Agus et al, 2019). As described above, this meta-analysis aims at identifying promising candidates for further functional studies and improving our understanding of transcription factors and mechanism which are mainly affected in the brain of HD patients.

To identify genes with significantly altered mRNA levels in those three studies, we determined differentially expressed genes for each study separately, ranked them after their

absolute Z-ratio and performed a robust rank aggregation analysis (RRA). Thereby, we identified 737 differentially expressed genes (RRA score < 0.05) that were among the most altered genes in the analysed datasets (Additional file 1).

Based on all genes identified by RRA (Additional file 1), we performed a weighted gene co-expression analysis (WGCNA) to identify gene modules, i.e. clusters of highly correlated genes. Adjacency and the topological overlap matrix (TOM) for the gene network were calculated with a soft-thresholding power β of 14.5 for which the WGCNA network satisfies the criterion of scale-free topology (R^2 = 0.85) (Additional file 2). By module-based clustering with the diagonal, varying volume, and shape model (VVI), we identified nine modules of which the module eigengenes (first principal component) of seven modules (black, blue, red, brown, magenta, and turquoise) statistically significantly correlated with disease state (HD patients versus healthy individuals) as determined by correlation analysis (Figure 1A). Corroborating a potential link between genes in the black, brown and turquoise modules with HD, we found a positive and statistically significant correlation between module membership, defined as the probability that a gene belongs to this module, and gene significance, defined as the correlation between the expression values and the trait of interest, of 0.46 (p = 0.004), 0.47 (p = 1e-04), and 0.41 (p = 0.008), respectively (Figure 1A). Additionally, genes belonging to the black and brown module showed a high mean gene significance (Figure 1B). In contrast, indicating that transcriptomic alterations of genes of the turquoise modules were less pronounced than the genes of the black and brown module, we observed a low mean gene significance of genes of this module. Corroborating the results of our network analysis, we found that most of the identified genes or proteins are known to interact with several other proteins belonging to the same module. For instance, according to the network analysis in GeneMania, 94.05% and 94.11% of the genes are known to be co-expressed in humans. Likewise, network analysis using the STRING database (confidence cut-off = 0.4) (Szklarczyk et al, 2019) showed that 34 of 53 proteins (64.2%) of the black and 36 of 81 proteins (44.4%) of the brown module interact with at least one other protein (data not shown).

Based on the clustering analysis of the modules eigengenes (Figure 1C) and the similarity of the eigengene adjacency (Figure 1D), we grouped the observed modules in three meta-modules: the first meta-module (M1) consisted of the black, blue, magenta, and red module, the second meta-module (M2) consisted of the brown, green and turquoise module, and the third meta-module (M3) consisted of the yellow and pink module. Combining the identified WGCNA modules to meta-modules and subsequent analysis of this meta-modules

demonstrated a high correlation with HD (correlation r = 0.5, p-value = 1e-38), a positive correlation between gene significance and module membership of 0.73 (p-value = 2e-54), and the highest mean gene significance of all meta-modules for M1 (Figure 2A and B). Corroborating the importance of genes belonging to M1, the eigengene and adjacency of M1 clustered together with the disease state (Figure 2C and D). According to STRING, protein-protein interactions were strongly enriched in M1 (826 observed edges, 602 expected edges, and p < 1.0e-16) at minimal interaction confidence of 0.4. While the eigengene of M2 correlated with HD, we could not observe a positive correlation between gene significance and module membership and, therefore, did not analyse this meta-module further.

Owing to the high similarity between modules of the same meta-module, we performed the transcription factor enrichment analysis (TFEA) and the subsequent GO-term enrichment analysis on the meta-modules M1 instead of single modules. Characterising functional enrichment of genes belonging to the M1 subnetwork, we performed enrichment analyses against the gene ontology (GO) and Reactome database. Using the GO database, we found an enrichment of genes involved in protein transport (GO: 0015031, 12.8 % of genes in M1 and FDR = 0.046) (Figure 3A) such as the Ras-related proteins 11A (RAB11A), 2A (RAB2A), 14 (RAB14), syntaxin-7 (STX7), syntaxin-12 (STX12), or the sorting nexin 3 (SNX3) in M1. Among the three datasets, the 41 proteins that belonged to M1 and function in protein transport processes showed a strong up-regulation in HD patient samples with a median Z-ratio of 1.51 (Figure 3B). To also include genes with lower Z-ratios into the functional enrichment analyses, irrespective of their module membership, we additionally conducted gene set enrichment analysis (GSEA) of the three datasets independently. Strongly suggesting that the alteration of genes involved in protein transport may be relevant in HD, we found a strong enrichment of proteins involved in the co-translational protein targeting to the membrane in all three datasets by GSEA (Figure 3E). Additionally, genes functioning in cellular metabolic processes (GO:0044237, 56.3 % of genes in M1 and FDR = 0.023), cellular respiration (GO:0045333, 3.4 % of genes in M1 and FDR = 0.027) and translation (GO:0006412, 5.6 % of genes in M1, and FDR = 0.023) were statistically significantly enriched in the M1 subnetwork. Similarly, when using the Reactome database, we observed a strong enrichment of proteins involved in protein metabolism, gene expression (transcription) and post-translational protein modification (Figure 3D). Using the network enrichment analysis test (NEAT), we found a highly statistically significant over-enrichment of 38 KEGG pathways in M1 (Additional file 3). Among these enriched pathways were “SNARE interaction in vesicular transport” (adjusted p-value = 1.74e-33), “RNA transport” (adjusted

p-value = 2.10E-217), and “mRNA surveillance pathway” (adjusted p-value = 9.63E-60) (Additional file 3).

To identify highly connected genes within the M1 subnetwork, we computed hub genes, i.e. genes with high intramodular connectivity, gene significance and module membership (Table 1, Additional file 4). Among these hub genes was the cell division cycle 42 protein (CDC42) (Z-ratios: 1.72; 1.83; 1.45 in GSE33000, GSE129473 and GSE64810 respectively), a membrane-associated small GTPase that interacts with several effector proteins and thereby regulates cell migration (Modzelewska et al, 2006), the bipolar attachment of spindle microtubules to kinetochores (Oceguera-Yanez et al, 2005), the extension and maintenance of the formation of filopodia, the dedicator of cytokinesis 10 (DOCK10) mediated spine formation (Gauthier-Campbell et al, 2004), and the structural plasticity of dendritic spines (Gauthier-Campbell et al, 2004). Further corroborating the importance of CDC42 in the subnetwork correlating with HD, CDC42 was additionally central in the M1 protein-protein interaction network, constructed using the STRING database (Additional file 5). Together with CDC42, we identified 18 proteins with altered mRNA levels in all datasets that were directly connected with CDC42 according to the STRING database. Further, CDC42 interacts with other identified hub genes such as the P21/Cdc42/Rac1-Activated Kinase 1 (PAK1) (Manser et al, 1994), 14-3-3 protein eta (YWHAH), or the protein phosphatase 2 catalytic subunit α (PPP2CA). mRNA levels of the CDC42 small effector 2 (CDC42SE2), that functions downstream of CDC42, was upregulated in the brain of HD patients (Z-ratios: 2.17; 0.97; 1.2). Additionally, CDC42 can interact with the CDC42-interacting protein 4 (CIP4), also known as thyroid hormone receptor interactor 10 (TRIP10) and HTT (Holbert et al, 2003; Hsu et al, 2011). that was not robustly dysregulated in our meta-analysis (Z-ratios -1.35, -0.03, 0.51). Besides the interaction of CIP4 / TRIP10 with HTT and CDC42, CIP4 / TRIP10 can interact with the vesicle-associated membrane protein 2 (VAMP2) and 7 (VAMP7) which are linked with other genes robustly altered in the brain of HD patients such as the vesicle-associated membrane protein 1 (VAMP1) and the Ras-related Protein Rab-14 (RAB14) (Lu & Wilson, 2016).

In addition, we identified PAK1 (Z-ratios: 2.07; -0.42; 0.299), which can interact with CDC42, as a hub gene in the M1 subnetwork. Additionally, the PAK1 Interacting Protein 1 (PAK1IP1), which inhibits the activation of PAK1 by CDC42 through its interaction with the N-terminus of PAK1 (Xia et al, 2001), was upregulated in the brain of HD patients in all studies (Z-ratios: 1.88; 1.10; 1.98). Since PAK1, as well as PAK2 and PAK3, belong to the group A PAKs (Harms et al, 2018), we also analysed mRNA levels of other group A PAKs.

While PAK2 was not robustly altered in the brain of HD patients (Z-ratios: -1.43; 1.31; 0.7), PAK3 mRNA levels (Z-ratios: 1.98; 0.46; 0.70) were slightly elevated in the brain of HD patients, although it was not identified by RRA (p = 0.10). Both CDC42 and PAK1 are interacting with another hub gene, the protein phosphatase 2 catalytic subunit alpha (PPP2CA) (Z-ratios: 1.98; 0.74; 1.91) (Figure 4B), which is an important phosphatase for microtubule-associated proteins. PPP2CA was additionally the most central protein in the network analysis of the M1 protein-protein interaction network (Additional file 5).

Broad transcriptional dysregulation upon HD was often linked to direct interaction of mHTT with proteins of the small interfering RNA (siRNA) machinery (Cha, 2000; McCampbell et al, 2000; Steffan et al, 2000; Sugars & Rubinsztein, 2003; Nalavade et al, 2013; Martí, 2016) and different transcriptional regulators such as CREB1, TBP, mSin3a, MBNL1, nucleolin, histone deacetylases (HDACs), or the DNA methyltransferase 1 (DNMT1). Therefore, we performed a transcription factor enrichment analysis (TFEA) of the M1 subnetwork to define which transcription factors would best explain the observed alterations in the brain of HD patients. Analysis of the target genes of the top five TFEA hits, the mitochondrial transcription termination factor 3 (MTERF3), Myb/SANT DNA binding domain containing 4 with coiled-coils (MSANTD4), small nuclear RNA activating complex polypeptide 5 (SNPAC5), zinc finger protein 833 (ZNF833), and thymocyte nuclear protein 1 (THYN1), showed that most of their target genes was upregulated in the brain of HD patients (Figure 5 and Additional file 6). mRNA levels of MTERF3, MSANTD4, SNPAC5, ZNF833 and THYN1 were not consistently altered in the brain of HD patients. Alongside with the mRNA levels of their target genes, mRNA levels of the distal-less homeobox 1 (DLX1) (regulates 11.9 % of M1, TFEA rank = 6, Z-ratios: 2.21; 0.16; 1.04), protein arginine methyltransferase 3 (PRMT3) (regulates 11.6 % of M1, TFEA rank = 14, Z-ratios: 0.46; 2.65; 1.11), and nuclear transcription factor Y subunit β (NFYB) (regulates 29.7 % of M1, TFEA rank = 24, Z-ratios: 1.05; 2.39; 2.02) were robustly upregulated in the brain of HD patients (Figure 5 and Additional file 6). Additionally, mRNA levels of the high mobility group nucleosomal binding domain 3 (HMGN3) (regulates 24.1 % of M1, TFEA rank = 13, Z-ratios: -0.91; 1.90; 1.77), that was additionally ranked high in the TFEA, appeared to be upregulated in only two of three studies and slightly downregulated in the other study (Additional files 1 and 6). Strikingly, we noted that transcriptions factors that have previously been shown to be affected by mHTT such as CREB1 (rank 182) or TBP (rank 208) were ranked low or could not be detected at all (mSin3a) by TFEA.

Besides the effects of mHTT on transcription factors, previous publications indicated that the

dysregulation of epigenetic modifiers such as DNMT1 or histone deacetylases (HDACs) (Federspiel et al, 2019; Siebzehnrübl et al, 2018; Moreno et al, 2016) might contribute to the broad transcriptional dysregulation in HD. In our meta-analysis, HDCA2 (Z-ratios: 0.69, 1.84, 1.86) and HDAC9 (Z-ratios: 2.26, 0.88, 1.42) were upregulated in the brain of HD patients, while the histone deacetylase 5 (HDAC5) mRNA levels were decreased (Z-ratios: -0.18; -2.17; -2.02). By RRA, we did not identify DNMT1, DNMT3A, or DNMT3B as robustly altered genes. Nonetheless, we noted a downregulation of DNMT1 (Z-ratio 0.48; - 1.35, -1.72) and DNMT3A (Z-ratio: 0.54, -1.38, -2.20) in datasets from Agus et al. 2019 and Labadorf et al., while both, DNMT1 and DNMT3A, were slightly upregulated in the larger dataset from Narayanan et al., 2014. DNMT3B (Z-ratio: - 0.76; -0.31; -0.79) was mostly unaltered in the brain of HD patients in all datasets.

Taken together, mRNA levels of 320 genes of the M1 co-expression network strongly correlated with HD and genes involved in protein transport and metabolism were enriched in this co-expression subnetwork (Figure 3). CDC42, PAK1, YWHAH and PP2CA were identified as hub genes of this network. Especially substantiating on the relevance of CDC42, CDC42 has been indirectly linked with HD before (Holbert et al, 2003) and can indirectly or directly interact with other identified hubs and 18 other proteins with robustly altered mRNA levels in the brain of HD patients. Further, the TFEA of the M1 subnetwork highlighted DLX1, NFY and HMGN3 as potential transcriptional regulators whose function might be affected in the brain of HD patients. Additionally, the mRNA levels of several epigenetic modifiers such as HDAC2, HDAC9, DNMT1, and DNMT3A were altered in at least two of the three studies using HD brain samples.

Transcriptional changes in the blood of HD patients

As afore-described, HTT is ubiquitously expressed and HD symptoms are not confined to the central nervous system (van der Burg et al, 2009; Saudou & Humbert, 2016). Hence, we additionally analysed transcriptomic studies of blood samples from HD patients and healthy controls. Borovecki et al. 2005 (GSE1751) analysed the transcriptomic profile of twelve symptomatic and five presymptomatic HD patients in comparison to 14 healthy controls, whereas Hu et al. 2011 (GSE24520) included venous cellular whole blood samples from 6 healthy controls and 8 HD patients. Transcriptomic profiles of lymphocytes from 12 moderate stage HD patients and 10 age-matched healthy controls were analysed by Runne et al. 2007 (GSE8762).

By robust rank aggregation (RRA), we identified 661 genes differentially expressed upon HD

among the three datasets ($p < 0.05$) (Additional file 1). Based on those 661 genes with a soft-thresholding power $\beta = 19.5$ (scale-free topology R2 = 0.87) (Additional file 2) and subsequent module-based clustering with the diagonal, equal volume, varying shape (EVI) model, we identified nine WGCNA modules. Of these modules, the module eigengene of three modules (brown, pink, and yellow) statistically significantly correlated with the disease state (healthy individuals versus HD) (Figure 6A). While genes of the brown module showed a negative correlation between module membership and gene significance, genes of the pink and yellow modules showed a positive correlation (Figure 6A). Corroborating the importance of genes of the pink and yellow modules, genes of these modules showed the highest mean gene significance of all identified modules (Figure 6B). Further, 94.12 % and 86.55 % of the genes belonging to either the pink or yellow module were annotated by GeneMania to be co-expressed in humans. Owing to the low distance between the modules eigengenes of the pink and yellow modules (Figure 6C) and the highly similar adjacency of these modules (Figure 6D), we combined these modules for further downstream analysis and will further refer to this module as the blood meta-module (MB).

Like the enrichment of the meta-module M1 identified in the brain samples, we found a strong enrichment of proteins involved in transport (FDR = 0.03 and 35.6 % of all genes in MB) and metabolic processes (FDR = 0.03 and 66.9 % of all genes in MB) in the MB subnetwork (Figure 7A and B). Consistent with the enrichment of proteins involved in protein transport, we found a strong enrichment of proteins localised to endosome membranes (FDR = 0.002 and 9.3 % of genes in MB) (Figure 7C). Among the 11 proteins localised to endosome membranes were the vesicle-associated membrane protein 7 (VAMP7) (Z-ratios -1.47; 0.99; 3.3), a paralog of VAMP2 that also displayed altered mRNA levels in the brain, and the sorting nexin 10 (SNX10) that is involved in membrane trafficking and protein sorting (Worby & Dixon, 2002). Additionally, we observed dysregulation of the actin-related proteins ACTR2/Arp2 (Z-ratio: -0.95; -1.75; 3.15), ACTR3/Arp3 (-1.92; -1.62; 2.78), and ARPC5 (-1; -1.87; 2.24) that together with ARBC1A, ARBC1B, ARPC3, and ARPC4 form a seven-subunit protein complex playing an essential role in the regulation of the actin cytoskeleton (Molinie & Gautreau, 2018). By analysis of the MB protein-protein interaction network, ACTR2 was the most important hub protein, which strongly substantiates on its pathophysiological relevance in HD. Further indicating that transcriptional dysregulation of regulators of the actin cytoskeleton may be important in HD, other actin-related proteins such as ACTR3B (Z-ratios: 1.3; 1.30; 1.62), ACTR6 (1.38; 1.54; 1.88) and ACTR10 (Z-ratios: 1.64; 1.23; 1.34) were robustly upregulated in the blood of HD patients according to this

meta-analysis. The alteration of mRNA levels of constituents of the Arp2/3 complex and actin-related proteins further substantiates on the identification of CDC42 as an important hub gene in the brain since CDC42 can activate the Arp2/3 complex through Wiskott-Aldrich syndrome proteins (Watson et al, 2017) such as WAS, WASF1, WASF2, WASF3, and WASL that appeared to be neither robustly altered in the brain nor in the blood of HD patients.

To identify transcription factors that can regulate the transcription of genes belonging to MB, we performed a TFEA (Additional file 6). The CGG Triplet Repeat Binding Protein 1 (CGGBP1), zinc finger protein 654 (ZNF654), forkhead box N2 (FOXN2), and the specificity protein 3 transcription factor (SP3) were ranked at the top in the TFEA. Strikingly, transcription of these transcriptional regulators is regulated by the cAMP-responsive element-binding protein-1 (CREBP1). CREB1 was also ranked high (rank 15) in the TFEA and can regulate the transcription of genes coding for constituents of the Arp2/3 complex and hub genes of the MB subnetwork such as membrane-associated ring-CH-type finger 7 (MARCH 7), pumilio RNA binding family member 2 (PUM2), survival motor neuron domain containing 1 (SMNDC1), or zinc finger DHHC-type palmitoyltransferase 17 (ZDHHC17), also known as the huntingtin-interacting protein 14 (HIP14). CREB1 together with the other enriched transcription factors regulated by CREB1 (TET2, SP3, RLF, CGGBP1, ZNF148, FOXN2, ZNF654, ZBTB11, and ZNF770) regulates the transcription of 92 from 118 (78.0 %) genes belonging to MB (Additional file 7).

Taken together, the enrichment of proteins localised to endosome membranes further corroborates the above-described alteration of protein-transport-related genes in the brain of HD patients. The dysregulation of several constituents of the Arp2/3 complex, that is activated by CDC42, substantiates on the relevance of actin cytoskeleton dysregulation in HD.

52 genes were differentially regulated in the blood and brain of HD patients

As noted before, the ubiquitous expression of HTT (Saudou & Humbert, 2016) and the clinical manifestation of HD outside the central nervous system (van der Burg et al, 2009) indicates that transcriptomic changes caused by polyQ expansion of HTT may be not confined to nervous tissues. According to our meta-analysis, 52 genes were dysregulated in brain- and blood-derived samples from HD patients (Additional file 1).

Based on the genes/proteins altered in the blood and brain of HD patients, we constructed a protein-protein interaction network (Figure 8) to investigate the relationship between the

identified genes. Furthermore, to identify transcription factors that might explain the observed transcriptomic alterations, we conducted a transcription factor enrichment analysis. While 82.4 % of the proteins were annotated as co-expressed by GeneMania (data not shown), 48.1 % were interacting with a least one other protein according to the STRING databases (interaction cutoff: 0.4).

In our meta-analysis, we noted that mRNA levels of the zinc finger DHHC-Type palmitoyltransferase 13 (ZDHHC13), i.e. huntingtin-interacting protein 14 (HIP14), and the zinc finger DHHC-Type palmitoyltransferase 17 (ZDHHC17), i.e. huntingtin-interacting protein 14-like protein (HIP14L) were dysregulated in the blood and brain of HD patients. Previous studies linked the altered interaction between mHTT and ZDHHC17 and ZDHHC13 with altered regulation of the striatal N-Methyl-D-Aspartate Receptor (NMDA) trafficking (Kang et al, 2019).

TFEA analysis of the 52 genes altered in the blood and brain of HD samples highlighted the enrichment of NFY target genes (rank 3) (Additional file 7). NFY is a trimeric complex of proteins coded by the *NFYA*, *NFYB*, and *NFYC* genes. Besides the high ranking of *NFYB* in the combined dataset (rank 3), targets of *NFYB* were also enriched in the M1 (brain, rank 24) and the blood (rank 48) (Additional file 7). Moreover, NFYB (Z-ratios in the brain datasets: 1.05, 2.39, 2.02) mRNA-levels and *NFYB* target genes (mean Z-ratio: 1.62) were also increased in the brain of HD patients (Figure 5 and Additional file 1).

HMGN3, NFY and CDC42 mRNA were additionally altered in the striatum of YAC128 and R6/2 mice and predictive for HD

To further substantiate on the relevance of identified hub genes or transcriptional regulators such as CDC42, PAK1, YWHAH, DLX1, HMGN3, or NFY, we analysed transcriptomic alterations in the striatum of R6/2 (Miyazaki et al, 2020) and YAC128 mice (Becanovic et al, 2010) (Additional file 8). Furthermore, we assessed how accurate control and HD mice can be discriminated based on mRNA levels of these genes (Additional file 9 and 10).

Indicating that NFY may play a role in HD across tissues, TFEA of the subnetworks M1 and MB and the combined dataset highlighted NFY. Like in human HD patients, NFYA (Z-ratios: 1.75, 0.77, 1.93 in 12-month-old YAC128, and 24-month-old YAC128 mice respectively) and NFYB (Z-ratios: 1.31, 0.45, 1.38) mRNA levels were elevated in R6/2 and YAC128 mice. Additionally, control and HD mice could be well discriminated based on NFYA (AUC = 0.86; 95% CI = [0.71, 1]), and NFYB (AUC = 0.79; 95% CI = [0.60, 0.98]) mRNA levels (Additional files 8 - 10).

In contrast, HMGN3 was only highlighted by TFEA of the M1 module, correlating with HD in the brain, and not in the blood datasets which may imply that dysregulation of HMGN3 may be confined to the brain In R6/2 and YAC128 mice, HMGN3 mRNA levels were elevated (Z-ratios: 1.89, 0.78, 2.06 in R6/2) in HD mice and control and HD mice could be well discriminated based on the HMGN3 mRNA levels. (AUC = 0.89; 95% CI = [0.76, 1]) (Additional files 8 - 10).

In concert with a more pronounced HD phenotype in 24-month YAC128 and R6/2 mice than in 12-month-old YAC128 mice, the increase of HMGN3, NFYA, and NFYB mRNA levels positively correlated with the age of YAC128. This raises the possibility that HMGN3, NFYA, and NFYB mRNA levels might be utilised as markers for disease progression and severity. However, further investigations on the usability of those genes as biomarkers in larger patient cohorts are required. Further, it should be clarified whether alteration of mRNA levels HMGN3, NFY, their target genes are specific for HD or whether they are present in other neurodegenerative diseases.

DLX1 and PRMT3 mRNA levels, which were both also highlighted by TFEA, were only elevated in the striatum of R6/2 mice but appeared to be unaffected in YAC128 mice. Like in the brain samples of HD patients, we observed a robust downregulation of DNMT3A (Z-ratios: - 1.24, -1.79, - 1.14) in the striatum of R6/2 and YAC128 mice and the mice could be discriminated based on DNMT3A levels (AUC = 0.83; 95% CI = [0.66, 0.99]). In disagreement with the analysis of brain samples from HD patients, DNMT1 and DNMT3B levels were not consistently altered in the striatum of R6/2 and YAC128 mice.

Further corroborating the importance of CDC42 dysregulation in HD, CDC42 mRNA levels were elevated in the striatum of R6/2 and YAC128 (Z-ratios: 1.24, 1.18, 1.88). Additionally, control and HD mice could be discriminated based on CDC42 mRNA levels (AUC = 0.85; 95% CI = [0.68, 1]). mRNA level of PAK1, that was identified as a hub gene in the M1 subnetwork and correlated with HD, was merely elevated in YAC128 mice (Z-ratios: 1.47, 2.08), but was mostly unaffected in R6/2 mice (Z-ratio: - 0.45).

Discussion

In our meta-analysis, we intended to identify by RRA, WGCNA and network analysis robust transcriptomic changes underlying HD. To this end, we included transcriptomic studies analysing different human brain regions and tissues from symptomatic and prodromal HD patients. Thereby, we identified subnetworks of 320 (M1) or 118 (MB) genes with robustly

altered mRNA levels in the brain and blood of HD patients, respectively. Network analysis of differentially expressed genes in the brain highlighted CDC42, PAK1, YWHAH, and PP2CA as hub genes of the M1 subnetwork. In the blood, we identified a subnetwork of 118 genes, including genes coding for several constituents of the Arp2/3 complex that is activated by CDC42. TFEA highlighted the relevance of several already described (e.g. CREB1 and NFY) or novel (e.g. DLX1, PRMT3 and HMGN3) transcription factors that may play a role in HD. In conclusion, our analysis suggests that dysregulation of transcription factors and epigenetic modifiers, cellular metabolism, actin cytoskeleton and SNARE complex proteins play an important role in the pathology of HD (Figure 8).

As noted before, the pathology of HD is neither confined to certain brain regions nor the brain (van der Burg et al, 2009). A good HD therapy should, therefore, target a gene or protein that is not exclusively altered in a certain brain region or tissue. Hence, we analysed RNA data from different brain regions and blood samples in our meta-analysis, although an increased interstudy variability, reducing the sensitivity with which differentially expressed genes are identified, may argue against the combined analysis of different brain regions. Furthermore, certain limitations for the interpretation and combined analysis of transcriptomic data from different studies should be considered: although authors of the original publications strictly controlled RNA quality before RNA sequencing or microarray analysis, small changes in RNA quality might impair transcript quantification and subsequently also the results of this meta-analysis. Second, post-mortem samples from HD and healthy controls can only provide insights into transcriptional changes at the end stage of HD that do not necessarily reflect changes at disease onset or during disease progression. Third, neurodegeneration in the brain of HD patients in late disease stages poses the risk that some of the observed alterations are caused by changes in tissue composition. Bearing the danger of altered tissue composition as a confounding factor in mind, we also included the dataset from Agus et al. 2019 who analysed early-stage, prodromal HD patients in which neuronal loss was less pronounced (Agus et al, 2019).

Several studies have shown that HTT and its interactors such as the huntingtin-associated protein 1 (HAP1) or the huntingtin-interacting protein 1 (HIP1) participate in protein transport and the organisation of the cytoskeleton (Saudou & Humbert, 2016). RNAi-mediated silencing of the huntingtin-interacting protein 1 related (HIP1R), also known as huntingtin-interacting protein 12 (HIP12), for instance, led to the stable association of clathrin-coated structures and their endocytic cargo to dynamin, actin, the Arp2/3 complex, and cortactin (Engqvist-Goldstein et al, 2004; Le Clainche et al, 2007). Furthermore, HAP1 regulates

synaptic vesicle exocytosis (Mackenzie et al, 2016) and neuronal endocytosis through its interaction with the Sec23 homolog A, COPII coat complex component (SEC23A) and the clathrin light chain B (Mackenzie et al, 2017). In line with previous findings that HTT and its interactors regulate the cytoskeleton and transport processes and substantiating that impairments of these functions contribute to the pathophysiology of HD, we identified CDC42 as a hub gene in the M1 subnetwork that was highly correlated with HD in transcriptomic studies of post-mortem brain samples. Moreover, CDC42 mRNA levels were elevated in the brain of HD patients and in the striatum of R6/2 and YAC128 mice. Besides, mRNA levels of several constituents of the Arp2/Arp3 complex (ACTR2/Arp2: blood; ACTR3/Arp3: blood and brain; ACTR3B: brain; ARPC5: blood; ACTR6: brain), that interacts with the HIP1R-cortactin complex and is activated by CDC42, were altered in HD patients. mRNA levels of VAMP1, an indirect CDC42 interactor, were robustly upregulated in all studies using HD brain samples (Additional file 1). mRNA levels of VAMP2 and VAMP7 were additionally altered in the HD blood samples (Additional file 1), although these alterations were not consistent across the different studies. VAMPs are major constituents of protein complexes involved in the docking and fusion of vesicles (Jena, 2011). These complexes are comprised of VAMPs, other syntaxins, the synaptosome associated protein 25 (SNAP-25), the N-ethylmaleimide-sensitive factor-like protein (NSF), the NSF-attachment proteins alpha (NAPA, SNAPA), beta (NAPB / SNAPB), gamma (NAPG / SNPAG), and SNAP receptors (SNARE). Corroborating our finding that the mRNA levels of VAMPs were altered in the brain and blood of HD patients and the striatum of R6/2 and YAC128 mice (this study), VAMP2 protein levels were also increased in striatal synaptosomes of Hdh140Q/140Q mice (Valencia et al, 2013). Contrary, protein levels of other proteins involved in the docking and fusion of vesicles such as SNAP-25 or rabphilin 3a were reduced in the post-mortem cortex of HD patients (Smith et al, 2007). Additionally, analysis of 175Q-HTT knock-in mice demonstrated altered levels of other proteins involved in synaptic function (SNAP-25, Rab3A, and PSD95), axonal transport, and microtubules (dynein, dynactin, and KIF3A) (Smith et al, 2014). Additionally, HTT can interact with SNAP25, and the SNAP25-associated proteins syntaxin 1A (STX1A) and calcium voltage-gated channel auxiliary subunit alpha2delta 1 (CACNA2D1) (Kaltenbach et al, 2007). Raising the possibility that HTT and its abundant interactor, the huntingtin-associated protein 40 (HAP40), plays a role in docking and fusion of synaptic vesicles, we previously found that constituents of this complex, the N-ethylmaleimide-sensitive factor attachment proteins alpha (NAPA, SNAPA), beta (NAPB / SNAPB), and gamma (NAPG / SNPAG) are the closest homologs of HAP40

(Seefelder et al, 2020).

Besides proteins of the SNARE complex, we identified other proteins linked with CDC42 in our network analysis of transcriptomic data from brain tissue of HD patients. For instance, PAK1, identified as a hub gene in the subnetwork M1, and its interactor PAK1IP1 showed robust upregulation in HD patients. The PAK proteins, PAK1 and PAK3, are central regulators of neuronal development and activating PAK1 mutations were aetiologic for secondary macrocephaly, developmental delay, ataxic gait and seizures in two unrelated patients (Harms et al, 2018). Double knock-out of PAK1/PAK3 in mice affected brain size and structure (Huang et al, 2011). Linking PAKs with HTT, Luo and Rubinszstein showed a physical interaction between HTT and PAK1 (Luo & Rubinsztein, 2009) and siRNA-mediated silencing of PAK1 and PAK2 reduced mutant HTT toxicity and aggregation (Luo et al, 2008; Luo & Rubinsztein, 2009). PAK2 knock-down in the murine striatal cell line STHdh(Q111) also reduced mutant HTT toxicity (Tourette et al, 2014).

As afore-mentioned, we found that 78 % of the genes of the subnetwork MB, which strongly correlated with HD in the blood (Figure 6), were directly or indirectly regulated by CREB1 (Additional files 6 and 7). CREB1, a leucine zipper transcription factor, activates the transcription of genes upon binding to the cAMP-response element (CRE). Steffan et al. previously showed that the CREB1-binding protein (CBP), a transcriptional coactivator of CREB1, can interact with HTT (Steffan et al, 2000). In a transcriptomic study of subcutaneous adipose tissue obtained from HD patients, CREB1 target genes were enriched, and the CREB1 mRNA levels were significantly increased (McCourt et al, 2015). Substantiating the physiological importance of CREB1, the double knock-out of CREB-1 and the cAMP-responsive element modulator (CREM) in mice ($Creb1^{Nescre}Crem^{-/-}$) led to severe neuronal loss during brain development and perinatal death (Mantamadiotis et al, 2002). A conditional, postnatal knock-out of *Creb1* and *Crem*, showed considerable atrophy in the striatum and hippocampus and a dystonic phenotype (Mantamadiotis et al, 2002). Indicating that the loss-of-function of CREB1 in mice can partly be compensated by CREM, neither the loss of CREB1 nor CREM alone induced neurodegeneration in mice (Mantamadiotis et al, 2002). While we also found a strong enrichment of CREB1 target genes in our MB subnetwork by TFEA, CREB1 mRNA levels appeared to be unaltered in all analysed datasets. As afore-mentioned, TFEA of the M1 subnetwork consisting of genes with robustly altered mRNA levels in the brain of HD patients did not highlight CREB1. The finding that CREB1 targets were not enriched in the M1 subnetwork (Additional file 6) in combination with the data from Mantamadiotis et al., raises the possibility that CREB1 function may be affected in

the brain of HD patients, but the CREB1 dysfunction is compensated by CREM or other transcription factors. In contrast, CREB1 dysfunction might not be compensated outside the brain due to the lack of detectable CREM protein expression in blood cells (Uhlén et al, 2015). Another transcription factor whose transcription is controlled by CREB1 and whose target genes were enriched in the MB subnetwork is SP3. The dual treatment of R6/2 mice with mithramycin, inhibiting SP3, and cystamine reduced the hypertrimethylation of histone H3 and extended their overall survival over 40% (Ryu et al, 2006). The alteration of the activity or mRNA levels of CREB1 might also partly explain observed alterations in mRNA levels of CDC42 and constituents of the Arp2/3 complex (this meta-analysis) since CREB can regulate their transcription.

The TFEA, performed in this study, also highlighted the enrichment of NFY target genes in the M1 and MB subnetwork as well as in the dataset of genes affected in both tissues. In the brain of HD patients, mRNA levels of NFYB and its target genes were elevated (Figure 5, Additional file 7). Further corroborating that the dysregulation of NFY and its target genes may be important in HD, van Hagen et al. found an enrichment of NFY target genes among a gene cluster that was differentially expressed in rat PC12 cells expressing the exon 1 of human 74Q-HTT (van Hagen et al, 2017). Aggregates of mutant HTT can interact *in vitro* and in the mouse brain with NFY and thereby reduce transcription of the NFY target gene HSP70 (Yamanaka et al, 2008). On the other hand, our meta-analysis may indicate that the transcriptional activity and transcription of NFY might be increased in the brain of HD patients, R6/2 and YAC128 mice. With regards to the currently available data on the role of NFY in HD, we cannot exclude the possibility that elevated mRNA levels of NFY and its target genes are caused by a compensatory mechanism to restore NFY function. This scenario may explain why transcription of NFY targets was reduced in one study (Yamanaka et al, 2008), while we observed increased transcription of NFY and its targets.

Besides NFY and CREB-1, TFEA (Additional file 6) of the M1 subnetwork of genes altered in the brain of HD patients highlighted DLX1, PRMT3, and HMGN3 that may be involved in astrocyte maturation (Lee et al, 2003). To our knowledge, this is the first study which indicates a potential role of DLX1, PRMT3, and HMGN3 dysfunction in HD. As noted above, DLX1 mRNA levels were only upregulated in the brain of HD patients in two studies, while it appeared to be unaffected in the third study analysing the caudate nucleus of prodromal HD patients. Additionally, DLX-1 mRNA levels were elevated in the striatum of R6/2 mice, while it was unaltered in the striatum of YAC128 mice. Besides the role of DLX1 in the adult brain, DLX1 plays also an important role in brain development; DLX1, together

with NOLZ-1 and DLX2, regulates the migration of striatal neurones to the dorsal or ventral striatum and the identity of striatal projection neurones (Chen et al, 2020). Chen et al. also demonstrated that the knock-out of NOLZ-1, also known as zinc finger protein 503 (ZNF503), in mice led to an upregulation of DLX1/2 and an aberrant neuronal migration from the dorsal to the ventral striatum (Chen et al, 2020). Demonstrating that elevated DLX1/DLX2 levels were causative for the aberrant neuronal migration, restoration of the altered DLX1/DLX2 levels in NOLZ-1 knock-out mice rescued the aberrant neuronal migration (Chen et al, 2020). A conditional DLX1 knock-out in cortical interneurons in mice reduced the excitatory input, fewer excitatory synapses and hypoplastic dendrites (Pla et al, 2018) which substantiated on the relevance of DLX1 beyond the striatum. Additionally, DLX1 knock-down in interneurons enhanced dendritic growth through neuropilin-2 and PAK3 (Dai et al, 2014), that was also slightly upregulated in the brain of HD patients (Z-ratios: 1.98; 0.46; 0.70) according to this meta-analysis. PRMT3, a protein methyltransferase whose mRNA levels were elevated in the brain of HD patients and the striatum of R6/2 mice (this study) but neither in the blood nor the striatum of YAC128 mice, is essential for dendritic spine maturation in the rat hippocampus (Miyata et al, 2010) and neuronal development (Ikenaka et al, 2006). Due to a more rapid disease progression and disease onset in R6/2 than YAC128 mice (Menalled et al, 2009; Farshim & Bates, 2018), the elevation of DLX-1 and PRMT3 mRNA levels in the striatum of R6/2 but not in YAC128 mice raises the possibility that the dysregulation of DLX-1 and PRMT3 occurs in later disease stages and is more pronounced upon expression of the HTT-exon1 fragment.

Conclusion

Here, we identified, by RRA and WGCNA, subnetworks of 320 (M1) and118 (MB) genes with robustly altered mRNA levels in the brain and blood of HD patients, resp. In the brain, CDC42, PAK1, YWHAH, and PP2CA were highlighted as hub genes of the M1 subnetwork (Additional file 4), that appears to be enriched in genes functioning in protein transport (Figure 3). Additionally, in the blood, we identified a subnetwork of 118 genes, including genes coding for several constituents of the Arp2/3 complex that is activated by CDC42. TFEA (Figure 5 and Additional file 6) highlighted the relevance of CREB1 in the pathology of HD since the transcription of 78.0 % of genes altered in the blood of HD patients were directly or indirectly regulated by CREB1. Furthermore, DLX1, PRMT3, HMGN3 and NFY target genes were enriched in the identified modules. HMGN3, NFYA, NFYB, and CDC42 mRNA levels were additionally altered in R6/2 and YAC128 mice (Additional file 8) and

could be used to discriminate between control and HD mice (Additional files 9 and 10). Indicating that the upregulation of DLX1 and PRMT3 transcription may occur in later disease stages, DLX-1 and PRMT3 mRNA levels were merely elevated in R6/2 mice but not in YAC128 mice that show a less severe HD phenotype than R6/2 mice.

Our results strongly suggest that abnormal protein transport, cytoskeletal organization, and transcriptional regulation might be central features in the pathophysiology of HD (figure 8). Furthermore, our study substantiates the role of CDC42, previously identified HTT interactors (e.g. PAK1, and PAK2) and transcriptional regulators (e.g. CREB1 and NFY) which have been reported to be sequestered to mutant HTT aggregates. Most interestingly, our data indicate a potential pathophysiological role of DLX-1, HMGN3 and PRMT3 in HD that have not been reported before.

Methods

Retrieval and tiding of datasets

In our meta-analysis, we analysed transcriptomic studies that were published in a peer-reviewed journal and whose raw data were publicly available. Furthermore, we excluded transcriptomic studies with less than eight samples from HD patients. To analyse post-mortem brain tissue, we retrieved data from the Gene Ontology Omnibus (GEO) database of the National Center for Biotechnology Information (NCBI) with the accession number GSE33000 (Narayanan et al, 2014), GSE129473 (Agus et al, 2019), and GSE64810 (Labadorf et al, 2015). Additionally, for the analysis of blood samples from HD patients, raw data were retrieved from the GEO database with the accession numbers GSE1751 (Borovecki et al, 2005), GSE24250 (Hu et al, 2011)(Hu et al, 2011), and GSE8762 (Runne et al, 2007). If genes were measured by several probes, the average of all probes for the respective genes was used. In our meta-analysis, we excluded the samples from presymptomatic HD patients due to a low patient number.

Missing data in the dataset GSE33000 were imputed by sequential and random hot-deck imputation as implement in the R-package VIM (Kowarik & Templ, 2016) since we assumed missing at random after graphical analysis of missing values by the R-function matrixplot (VIM package) (Templ et al, 2012). We normalised raw transcript-levels by quantile-normalisation using the R function normalize.quantiles.robust from the package preprocessCore (Bolstad, 2019) and, afterwards, converted them into Z-Scores.

Robust rank aggregation analysis (RRA)

For the robust rank aggregation (RRA), we computed Z-ratios according to the method proposed by Cheadle et al., 2003 (Cheadle et al, 2003) and ranked them after their absolute Z-ratio. The sorted transcript lists were analysed with RRA, as implemented in the R package RobustRankAggreg (Kolde et al, 2012). We included transcripts with a score < 0.05 in the further downstream analysis and performed clustering analysis and plotting of the heatmaps with the function heatmap.2 implemented in the R-package gplots (version 3.0.3) (Warnes et al, 2020).

Generation of weighted correlation networks

For the weighted correlation network analysis (WGCNA), the signed co-expression networks were build using the R-package WGCNA (Langfelder & Horvath, 2008). Correlation between genes was computed by biweight midcorrelation (Langfelder & Horvath, 2008) to compute adjacency matrices. Based on the scale-free criterion (Zhang & Horvath, 2005), we set the power parameter β and computed the topological overlap measure (TOM) and the corresponding dissimilarity matrices (1 – TOM). Genes were clustered by model-based clustering of the dissimilarity matrix as implemented by Scrucca et al. 2016 (Scrucca et al, 2016). Correlation of module eigengenes with disease state and between gene significance and module membership were calculated by Pearson's product-moment correlation as implemented in R (Langfelder et al, 2008; R Core Team, 2020). We adjusted p-values for multiple testing with the method described by Yekutieli & Benjamini, 2001 (Yekutieli & Benjamini, 2001). Genes with a gene significance score above 0.3, module membership above 0.7, and intramodular connectivity that is larger than the 8th percentile of all genes were identified as hubs.

Enrichment Analysis

Gene set enrichment analysis (GSEA) was conducted with the algorithm implemented in the STRING database (Szklarczyk et al, 2019) for each dataset separately and results were combined by ranking enriched terms after their enrichment scores and aggregation by RRA (Kolde et al, 2012). Gene ontology (GO) and Reactome term enrichment of genes from different subnetworks were performed with the algorithms implemented by the STRING database (Szklarczyk et al, 2019) in Cytoscape (version 3.7.2) (Shannon et al, 2003).

Based on the pathway annotations in the KEGG database (Kanehisa & Goto, 2000) and the protein-protein interaction data from STRING (version 11.0) (Szklarczyk et al, 2019), we

performed a network enrichment analysis test (NEAT) (Signorelli et al, 2016) as implemented in the R package ´neat´. Transcription factor enrichment analysis (TFEA) was performed with ChIP-X Enrichment Analysis 3 (CheA3) (Keenan et al, 2019) using the mean rank as the metric. Furthermore, we calculated the mean Z-ratios of transcription factor targets with R-scripting language (R Core Team, 2020) and plotted the results using ggplot2 (Wickham, 2016).

Network analysis

Protein-protein interaction networks (PPIN) were retrieved from the STRING database (Szklarczyk et al, 2019) using a confidence level cut-off of 0.4 and the Cytoscape software (version 3.7.2) (Shannon et al, 2003). Top 50 hubs of the PPIN were computed with the cytoHubba plug-in (Chin et al, 2014) using the betweenness, bottleneck, closeness, clustering coefficient, degree, DMNC, EcCentricity, EPC, MCC, radiality and stress scoring methods. Results of different scoring methods were aggregated by RRA (Kolde et al, 2012) to increase the robustness of the prediction. We included hub proteins with an RRA score below 0.05 in the further analysis and combined the lists of hubs in the co-expression network, identified by WGNA, and the protein-protein interaction network by RRA (Kolde et al, 2012).

Confirmation of hub genes in HD mouse models

We analysed two datasets with transcriptomic data of HD mouse models to confirm if the identified hub genes and transcriptional regulators were additionally altered in independent datasets. In these studies, transcriptomic alterations in the striatum of R6/2 (Miyazaki et al, 2020) (NCBI accession number: GSE113929) or YAC128 (NCBI accession number: GSE19677) (Becanovic et al, 2010) were analysed. Since Becanovic et al. 2010 used YAC128 mice at the age of 12 and 24 months, we separated the samples according to the age of mice and analysed them as independent datasets. Data tidying and computation of Z-ratios were performed as described in the sections “retrieval and tiding of datasets” and “Robust rank aggregation analysis (RRA)”.

Classification analysis of selected genes (ACTR2, ACTR3, ARPC5, CDC42, CREB1, DLX1, DNMT1, DNMT3A, DNMT3B, HDAC2, HDAC5, HMGN3, NFYA, NFYB, NFYC, PAK1, PRMT3, VAMP2, VAMP7, YWHAH, ZDHHC13, ZDHHC17), identified in the human datasets, was performed, as implemented in the R-package pROC (Robin et al, 2011), to compute the area under the curve (AUC) of the respective receiver-operator characteristics (ROC). 95 % confidence intervals of AUCs were calculated by bootstrapping with 10,000 replicates and genes with confidence intervals for the AUC above 0.5 were considered

capable to discriminate between control and HD mice since the classification model is statistically significantly better than a random classification model.

Acknowledgements

We acknowledge funding of this work by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation – project number 412854449). We thank Andreas Neueder for discussions on the results of this meta-analysis. Furthermore, we thank Robin Nilson for improving the design of figure 8.

Author contributions

MS planned and performed bioinformatical analyses of transcriptomic data. MS and SK wrote and revised the manuscript.

Data Availability

All used transcriptomic data are accessible from the Gene Ontology Omnibus (GEO) database of the National Center for Biotechnology Information (NCBI) with the accession numbers given in the methods section. R scripts used for the analysis are made available upon reasonable request to the corresponding author.

Conflict of interest

The authors declare that they have no competing interests.

References

Agus F, Crespo D, Myers RH, Labadorf A (2019) The caudate nucleus undergoes dramatic and unique transcriptional changes in human prodromal Huntington's disease brain. *BMC medical genomics* **12**: 137

Becanovic K, Pouladi MA, Lim RS, Kuhn A, Pavlidis P, Luthi-Carter R, Hayden MR, Leavitt BR (2010) Transcriptional changes in Huntington disease identified using genome-wide expression profiling and cross-platform analysis. *Human molecular genetics* **19**: 1438–1452

Borovecki F, Lovrecic L, Zhou J, Jeong H, Then F, Rosas HD, Hersch SM, Hogarth P, Bouzou B, Jensen RV, Krainc D (2005) Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 11023–11028

Cha JH (2000) Transcriptional dysregulation in Huntington's disease. *Trends in neurosciences* **23**: 387–392

Cheadle C, Vawter MP, Freed WJ, Becker KG (2003) Analysis of Microarray Data Using Z Score Transformation. *The Journal of Molecular Diagnostics* **5**: 73–81

Chen S-Y, Lu K-M, Ko H-A, Huang T-H, Hao JH-J, Yan Y-T, Chang SL-Y, Evans SM, Liu F-C (2020) Parcellation of the striatal complex into dorsal and ventral districts. *Proceedings of the*

691 *National Academy of Sciences of the United States of America* **117**: 7418–7429

692 Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y (2014) cytoHubba: identifying
693 hub objects and sub-networks from complex interactome. *BMC systems biology* **8 Suppl 4**: S11

694 Crotti A, Glass CK (2015) The choreography of neuroinflammation in Huntington‘s disease.
695 *Trends in immunology* **36**: 364–373

696 Dai X, Iwasaki H, Watanabe M, Okabe S (2014) Dlx1 transcription factor regulates dendritic
697 growth and postsynaptic differentiation through inhibition of neuropilin-2 and PAK3 expression.
698 *The European journal of neuroscience* **39**: 531–547

699 Engqvist-Goldstein AEY, Zhang CX, Carreno S, Barroso C, Heuser JE, Drubin DG (2004)
700 RNAi-mediated Hip1R silencing results in stable association between the endocytic machinery
701 and the actin assembly machinery. *Molecular biology of the cell* **15**: 1666–1679

702 Fan MMY, Raymond LA (2007) N-methyl-D-aspartate (NMDA) receptor function and
703 excitotoxicity in Huntington‘s disease. *Progress in neurobiology* **81**: 272–293

704 Farshim PP, Bates GP (2018) Mouse Models of Huntington‘s Disease. *Methods in molecular*
705 *biology (Clifton, N.J.)* **1780**: 97–120

706 Gauthier-Campbell C, Bredt DS, Murphy TH, El-Husseini AE-D (2004) Regulation of
707 dendritic branching and filopodia formation in hippocampal neurons by specific acylated protein
708 motifs. *Molecular biology of the cell* **15**: 2205–2217

709 Gipson TA, Neueder A, Wexler NS, Bates GP, Housman D (2013) Aberrantly spliced HTT, a
710 new player in Huntington‘s disease pathogenesis. *RNA biology* **10**: 1647–1652

711 Harjes P, Wanker EE (2003) The hunt for huntingtin function: interaction partners tell many
712 different stories. *Trends in Biochemical Sciences* **28**: 425–433

713 Harms FL, Kloth K, Bley A, Denecke J, Santer R, Lessel D, Hempel M, Kutsche K (2018)
714 Activating Mutations in PAK1, Encoding p21-Activated Kinase 1, Cause a Neurodevelopmental
715 Disorder. *American journal of human genetics* **103**: 579–591

716 Holbert S, Dedeoglu A, Humbert S, Saudou F, Ferrante RJ, Néri C (2003) Cdc42-interacting
717 protein 4 binds to huntingtin: neuropathologic and biological evidence for a role in Huntington‘s
718 disease. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 2712–2717

719 Hsu C-C, Leu Y-W, Tseng M-J, Lee K-D, Kuo T-Y, Yen J-Y, Lai Y-L, Hung Y-C, Sun W-S,
720 Chen C-M, Chu P-Y, Yeh K-T, Yan PS, Chang Y-S, Huang TH-M, Hsiao S-H (2011) Functional
721 characterization of Trip10 in cancer cell growth and survival. *Journal of biomedical science* **18**: 12

722 Hu Y, Chopra V, Chopra R, Locascio JJ, Liao Z, Ding H, Zheng B, Matson WR, Ferrante RJ,
723 Rosas HD, Hersch SM, Scherzer CR (2011) Transcriptional modulator H2A histone family,
724 member Y (H2AFY) marks Huntington disease activity in man and mouse. *Proceedings of the*
725 *National Academy of Sciences of the United States of America* **108**: 17141–17146

726 Huang W, Zhou Z, Asrar S, Henkelman M, Xie W, Jia Z (2011) p21-Activated kinases 1 and 3
727 control brain size through coordinating neuronal complexity and synaptic properties. *Molecular*
728 *and cellular biology* **31**: 388–403

729 Ikenaka K, Miyata S, Mori Y, Koyama Y, Taneda T, Okuda H, Kousaka A, Tohyama M
730 (2006) Immunohistochemical and western analyses of protein arginine N-methyltransferase 3 in
731 the mouse brain. *Neuroscience* **141**: 1971–1982

732 Jena BP (2011) Role of SNAREs in membrane fusion. *Advances in experimental medicine and*
733 *biology* **713**: 13–32

734 Kaltenbach LS, Romero E, Becklin RR, Chettier R, Bell R, Phansalkar A, Strand A, Torcassi
735 C, Savage J, Hurlburt A, Cha G-H, Ukani L, Chepanoske CL, Zhen Y, Sahasrabudhe S, Olson J,
736 Kurschner C, Ellerby LM, Peltier JM, Botas J et al (2007) Huntingtin interacting proteins are
737 genetic modifiers of neurodegeneration. *PLoS genetics* **3**: e82

738 Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids*

739 *research* **28**: 27–30

740 Kang R, Wang L, Sanders SS, Zuo K, Hayden MR, Raymond LA (2019) Altered Regulation of
741 Striatal Neuronal N-Methyl-D-Aspartate Receptor Trafficking by Palmitoylation in Huntington
742 Disease Mouse Model. *Frontiers in synaptic neuroscience* **11**: 3

743 Keenan AB, Torre D, Lachmann A, Leong AK, Wojciechowicz ML, Utti V, Jagodnik KM,
744 Kropiwnicki E, Wang Z, Ma'ayan A (2019) ChEA3: transcription factor enrichment analysis by
745 orthogonal omics integration. *Nucleic acids research* **47**: W212-W224

746 Kolde R, Laur S, Adler P, Vilo J (2012) Robust rank aggregation for gene list integration and
747 meta-analysis. *Bioinformatics (Oxford, England)* **28**: 573–580

748 Kowarik A, Templ M (2016) Imputation with the R Package VIM. *J. Stat. Soft.* **74**

749 Labadorf A, Choi SH, Myers RH (2017) Evidence for a Pan-Neurodegenerative Disease
750 Response in Huntington's and Parkinson's Disease Expression Profiles. *Frontiers in molecular*
751 *neuroscience* **10**: 430

752 Labadorf A, Hoss AG, Lagomarsino V, Latourelle JC, Hadzi TC, Bregu J, MacDonald ME,
753 Gusella JF, Chen J-F, Akbarian S, Weng Z, Myers RH (2015) RNA Sequence Analysis of Human
754 Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental
755 Gene Expression. *PloS one* **10**: e0143563

756 Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network
757 analysis. *BMC bioinformatics* **9**: 559

758 Le Clainche C, Pauly BS, Zhang CX, Engqvist-Goldstein AEY, Cunningham K, Drubin DG
759 (2007) A Hip1R-cortactin complex negatively regulates actin assembly associated with
760 endocytosis. *The EMBO journal* **26**: 1199–1210

761 Le Gras S, Keime C, Anthony A, Lotz C, Longprez L de, Brouillet E, Cassel J-C, Boutillier A-
762 L, Merienne K (2017) Altered enhancer transcription underlies Huntington's disease striatal
763 transcriptional signature. *Scientific reports* **7**: 42875

764 Lee SS, Seo HS, Choi SJ, Park HS, Lee JY, Lee K-H, Park JY (2003) Characterization of the
765 two genes differentially expressed during development in human fetal astrocytes. *Yonsei medical*
766 *journal* **44**: 1059–1068

767 Lu R, Wilson JM (2016) Rab14 specifies the apical membrane through Arf6-mediated
768 regulation of lipid domains and Cdc42. *Scientific reports* **6**: 38249

769 Luo S, Mizuta H, Rubinsztein DC (2008) p21-activated kinase 1 promotes soluble mutant
770 huntingtin self-interaction and enhances toxicity. *Human molecular genetics* **17**: 895–905

771 Luo S, Rubinsztein DC (2009) Huntingtin promotes cell survival by preventing Pak2 cleavage.
772 *Journal of cell science* **122**: 875–885

773 Mackenzie KD, Lim Y, Duffield MD, Chataway T, Zhou X-F, Keating DJ (2017) Huntingtin-
774 associated protein-1 (HAP1) regulates endocytosis and interacts with multiple trafficking-related
775 proteins. *Cellular signalling* **35**: 176–187

776 Mackenzie KD, Lumsden AL, Guo F, Duffield MD, Chataway T, Lim Y, Zhou X-F, Keating
777 DJ (2016) Huntingtin-associated protein-1 is a synapsin I-binding protein regulating synaptic
778 vesicle exocytosis and synapsin I trafficking. *Journal of neurochemistry* **138**: 710–721

779 Manser E, Leung T, Salihuddin H, Zhao ZS, Lim L (1994) A brain serine/threonine protein
780 kinase activated by Cdc42 and Rac1. *Nature* **367**: 40–46

781 Mantamadiotis T, Lemberger T, Bleckmann SC, Kern H, Kretz O, Martin Villalba A, Tronche
782 F, Kellendonk C, Gau D, Kapfhammer J, Otto C, Schmid W, Schütz G (2002) Disruption of
783 CREB function in brain leads to neurodegeneration. *Nature genetics* **31**: 47–54

784 Martí E (2016) RNA toxicity induced by expanded CAG repeats in Huntington's disease.
785 *Brain pathology (Zurich, Switzerland)* **26**: 779–786

786 McCampbell A, Taylor JP, Taye AA, Robitschek J, Li M, Walcott J, Merry D, Chai Y, Paulson

787 H, Sobue G, Fischbeck KH (2000) CREB-binding protein sequestration by expanded
788 polyglutamine. *Human molecular genetics* **9**: 2197–2202

789 Menalled L, El-Khodor BF, Patry M, Suárez-Fariñas M, Orenstein SJ, Zahasky B, Leahy C,
790 Wheeler V, Yang XW, MacDonald M, Morton AJ, Bates G, Leeds J, Park L, Howland D, Signer
791 E, Tobin A, Brunner D (2009) Systematic behavioral evaluation of Huntington‘s disease
792 transgenic and knock-in mouse models. *Neurobiology of disease* **35**: 319–336

793 Miyata S, Mori Y, Tohyama M (2010) PRMT3 is essential for dendritic spine maturation in rat
794 hippocampal neurons. *Brain research* **1352**: 11–20

795 Miyazaki H, Yamanaka T, Oyama F, Kino Y, Kurosawa M, Yamada-Kurosawa M, Yamano R,
796 Shimogori T, Hattori N, Nukina N (2020) FACS-array-based cell purification yields a specific
797 transcriptome of striatal medium spiny neurons in a murine Huntington disease model. *The*
798 *Journal of biological chemistry* **295**: 9768–9785

799 Modzelewska K, Newman LP, Desai R, Keely PJ (2006) Ack1 mediates Cdc42-dependent cell
800 migration and signaling to p130Cas. *The Journal of biological chemistry* **281**: 37527–37535

801 Molinie N, Gautreau A (2018) The Arp2/3 Regulatory System and Its Deregulation in Cancer.
802 *Physiological reviews* **98**: 215–238

803 Nalavade R, Griesche N, Ryan DP, Hildebrand S, Krauss S (2013) Mechanisms of RNA-
804 induced toxicity in CAG repeat disorders. *Cell death & disease* **4**: e752

805 Narayanan M, Huynh JL, Wang K, Yang X, Yoo S, McElwee J, Zhang B, Zhang C, Lamb JR,
806 Xie T, Suver C, Molony C, Melquist S, Johnson AD, Fan G, Stone DJ, Schadt EE, Casaccia P,
807 Emilsson V, Zhu J (2014) Common dysregulation network in the human prefrontal cortex
808 underlies two neurodegenerative diseases. *Molecular systems biology* **10**: 743

809 Oceguera-Yanez F, Kimura K, Yasuda S, Higashida C, Kitamura T, Hiraoka Y, Haraguchi T,
810 Narumiya S (2005) Ect2 and MgcRacGAP regulate the activation and function of Cdc42 in
811 mitosis. *The Journal of cell biology* **168**: 221–232

812 Pla R, Stanco A, Howard MA, Rubin AN, Vogt D, Mortimer N, Cobos I, Potter GB, Lindtner
813 S, Price JD, Nord AS, Visel A, Schreiner CE, Baraban SC, Rowitch DH, Rubenstein JLR (2018)
814 Dlx1 and Dlx2 Promote Interneuron GABA Synthesis, Synaptogenesis, and Dendritogenesis.
815 *Cerebral cortex (New York, N.Y. : 1991)* **28**: 3797–3815

816 R Core Team (2020) *R: A Language and Environment for Statistical Computing,* Vienna, Austria

817 Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M (2011) pROC: an
818 open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics* **12**: 77

819 Ross CA, Tabrizi SJ (2011) Huntington‘s disease: from molecular pathogenesis to clinical
820 treatment. *The Lancet Neurology* **10**: 83–98

821 Runne H, Kuhn A, Wild EJ, Pratyaksha W, Kristiansen M, Isaacs JD, Régulier E, Delorenzi
822 M, Tabrizi SJ, Luthi-Carter R (2007) Analysis of potential transcriptomic biomarkers for
823 Huntington‘s disease in peripheral blood. *Proceedings of the National Academy of Sciences of the United*
824 *States of America* **104**: 14424–14429

825 Ryu H, Lee J, Hagerty SW, Soh BY, McAlpin SE, Cormier KA, Smith KM, Ferrante RJ (2006)
826 ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington‘s disease.
827 *Proceedings of the National Academy of Sciences of the United States of America* **103**: 19176–19181

828 Saudou F, Humbert S (2016) The Biology of Huntingtin. *Neuron* **89**: 910–926

829 Scrucca L, Fop M, Murphy TB, Raftery A (2016) mclust 5: Clustering, Classification and
830 Density Estimation Using Gaussian Finite Mixture Models. *The R Journal* **8**: 289

831 Seefelder M, Alva V, Huang B, Engler T, Baumeister W, Guo Q, Fernández-Busnadiego R,
832 Lupas AN, Kochanek S (2020) The evolution of the huntingtin-associated protein 40 (HAP40) in
833 conjunction with huntingtin. *BMC Evolutionary Biology* **20**

834 Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B,

835 Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular
836 interaction networks. *Genome research* **13**: 2498–2504

837 Shirasaki DI, Greiner ER, Al-Ramahi I, Gray M, Boontheung P, Geschwind DH, Botas J,
838 Coppola G, Horvath S, Loo JA, Yang XW (2012) Network organization of the huntingtin
839 proteomic interactome in mammalian brain. *Neuron* **75**: 41–57

840 Signorelli M, Vinciotti V, Wit EC (2016) NEAT: an efficient network enrichment analysis test.
841 *BMC bioinformatics* **17**: 352

842 Smith GA, Rocha EM, McLean JR, Hayes MA, Izen SC, Isacson O, Hallett PJ (2014)
843 Progressive axonal transport and synaptic protein changes correlate with behavioral and
844 neuropathological abnormalities in the heterozygous Q175 KI mouse model of Huntington‘s
845 disease. *Human molecular genetics* **23**: 4510–4527

846 Smith R, Klein P, Koc-Schmitz Y, Waldvogel HJ, Faull RLM, Brundin P, Plomann M, Li J-Y
847 (2007) Loss of SNAP-25 and rabphilin 3a in sensory-motor cortex in Huntington‘s disease.
848 *Journal of neurochemistry* **103**: 115–123

849 Steffan JS, Kazantsev A, Spasic-Boskovic O, Greenwald M, Zhu YZ, Gohler H, Wanker EE,
850 Bates GP, Housman DE, Thompson LM (2000) The Huntington‘s disease protein interacts with
851 p53 and CREB-binding protein and represses transcription. *Proceedings of the National Academy of*
852 *Sciences of the United States of America* **97**: 6763–6768

853 Sugars KL, Rubinsztein DC (2003) Transcriptional abnormalities in Huntington disease.
854 *Trends in genetics : TIG* **19**: 233–238

855 Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M,
856 Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering C von (2019) STRING v11: protein-
857 protein association networks with increased coverage, supporting functional discovery in
858 genome-wide experimental datasets. *Nucleic acids research* **47**: D607-D613

859 Tabrizi SJ, Cleeter MW, Xuereb J, Taanman JW, Cooper JM, Schapira AH (1999) Biochemical
860 abnormalities and excitotoxicity in Huntington‘s disease brain. *Annals of neurology* **45**: 25–32

861 Templ M, Alfons A, Filzmoser P (2012) Exploring incomplete data using visualization
862 techniques. *Adv Data Anal Classif* **6**: 29–47

863 The Huntington‘s Disease Collaborative Research Group (1993) A novel gene containing a
864 trinucleotide repeat that is expanded and unstable on Huntington‘s disease chromosomes. *Cell* **72**:
865 971–983

866 Tourette C, Li B, Bell R, O‘Hare S, Kaltenbach LS, Mooney SD, Hughes RE (2014) A large
867 scale Huntingtin protein interaction network implicates Rho GTPase signaling pathways in
868 Huntington disease. *The Journal of biological chemistry* **289**: 6709–6726

869 Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å,
870 Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA-K,
871 Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T et al (2015) Proteomics. Tissue-based
872 map of the human proteome. *Science (New York, N.Y.)* **347**: 1260419

873 Valencia A, Sapp E, Kimm JS, McClory H, Ansong KA, Yohrling G, Kwak S, Kegel KB,
874 Green KM, Shaffer SA, Aronin N, DiFiglia M (2013) Striatal synaptosomes from
875 Hdh140Q/140Q knock-in mice have altered protein levels, novel sites of methionine oxidation,
876 and excess glutamate release after stimulation. *Journal of Huntington‘s disease* **2**: 459–475

877 van der Burg JMM, Björkqvist M, Brundin P (2009) Beyond the brain: widespread pathology
878 in Huntington‘s disease. *The Lancet Neurology* **8**: 765–774

879 van Hagen M, Piebes DGE, Leeuw WC de, Vuist IM, van Roon-Mom WMC, Moerland PD,
880 Verschure PJ (2017) The dynamics of early-state transcriptional changes and aggregate formation
881 in a Huntington‘s disease cell model. *BMC genomics* **18**: 373

882 Velier J, Kim M, Schwarz C, Kim TW, Sapp E, Chase K, Aronin N, DiFiglia M (1998) Wild-
883 type and mutant huntingtins function in vesicle trafficking in the secretory and endocytic

884 pathways. *Experimental neurology* **152**: 34–40

885 Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler
886 M, Magnusson A, Moeller S, Schwartz M, Venables B (2020) *gplots: Various R Programming Tools for*
887 *Plotting Data*

888 Watson JR, Owen D, Mott HR (2017) Cdc42 in actin dynamics: An ordered pathway
889 governed by complex equilibria and directional effector handover. *Small GTPases* **8**: 237–244

890 Wickham H (2016) *ggplot2: Elegant graphics for data analysis.* Springer, Cham

891 Worby CA, Dixon JE (2002) Sorting out the cellular functions of sorting nexins. *Nature reviews.*
892 *Molecular cell biology* **3**: 919–931

893 Yamanaka T, Miyazaki H, Oyama F, Kurosawa M, Washizu C, Doi H, Nukina N (2008)
894 Mutant Huntingtin reduces HSP70 expression through the sequestration of NF-Y transcription
895 factor. *The EMBO journal* **27**: 827–839

896 Yekutieli D, Benjamini Y (2001) The control of the false discovery rate in multiple testing
897 under dependency. *Ann. Statist.* **29**: 1165–1188

898 Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network
899 analysis. *Statistical applications in genetics and molecular biology* **4**: Article17

900 Zuccato C, Cattaneo E (2014) Normal function of huntingtin. In *Huntington's disease,* Bates G,
901 Tabrizi S, Jones L (eds) pp 243–273. Oxford, New York: Oxford University Press

902

Figures

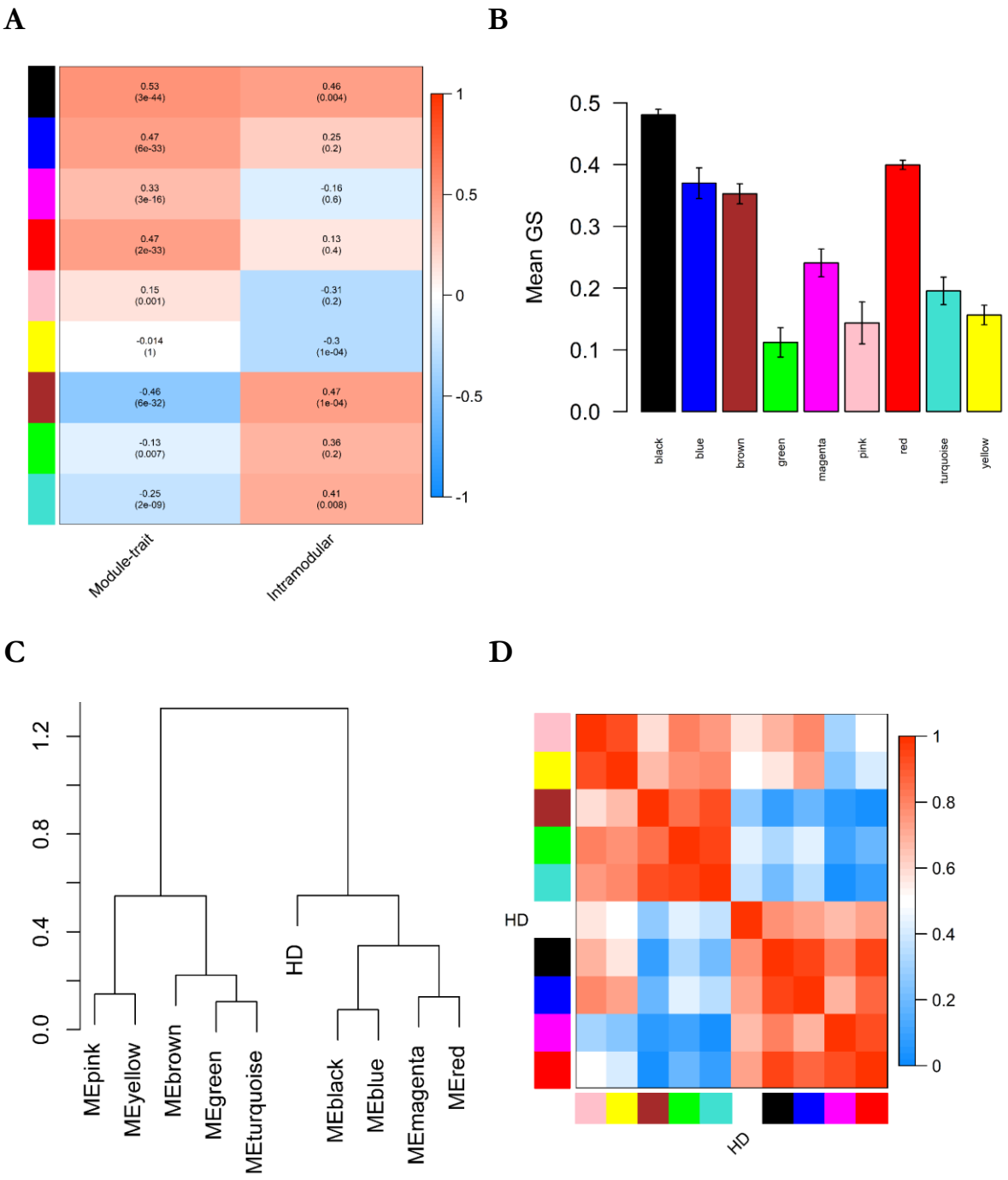

Figure 1: Result of WGCNA analysis and the association of the disease state with module eigengenes. A. Heatmap showing the correlation between module and disease state or between gene significance and module membership. P-values adjusted after Benjamini & Yekutieli (Yekutieli & Benjamini, 2001) are given in brackets. B: Mean gene significance of each module. Error bars depict the 95% confidence interval. C. Dendrogram showing hierarchical clustering of module eigengenes. D. Eigengene adjacency heatmap.

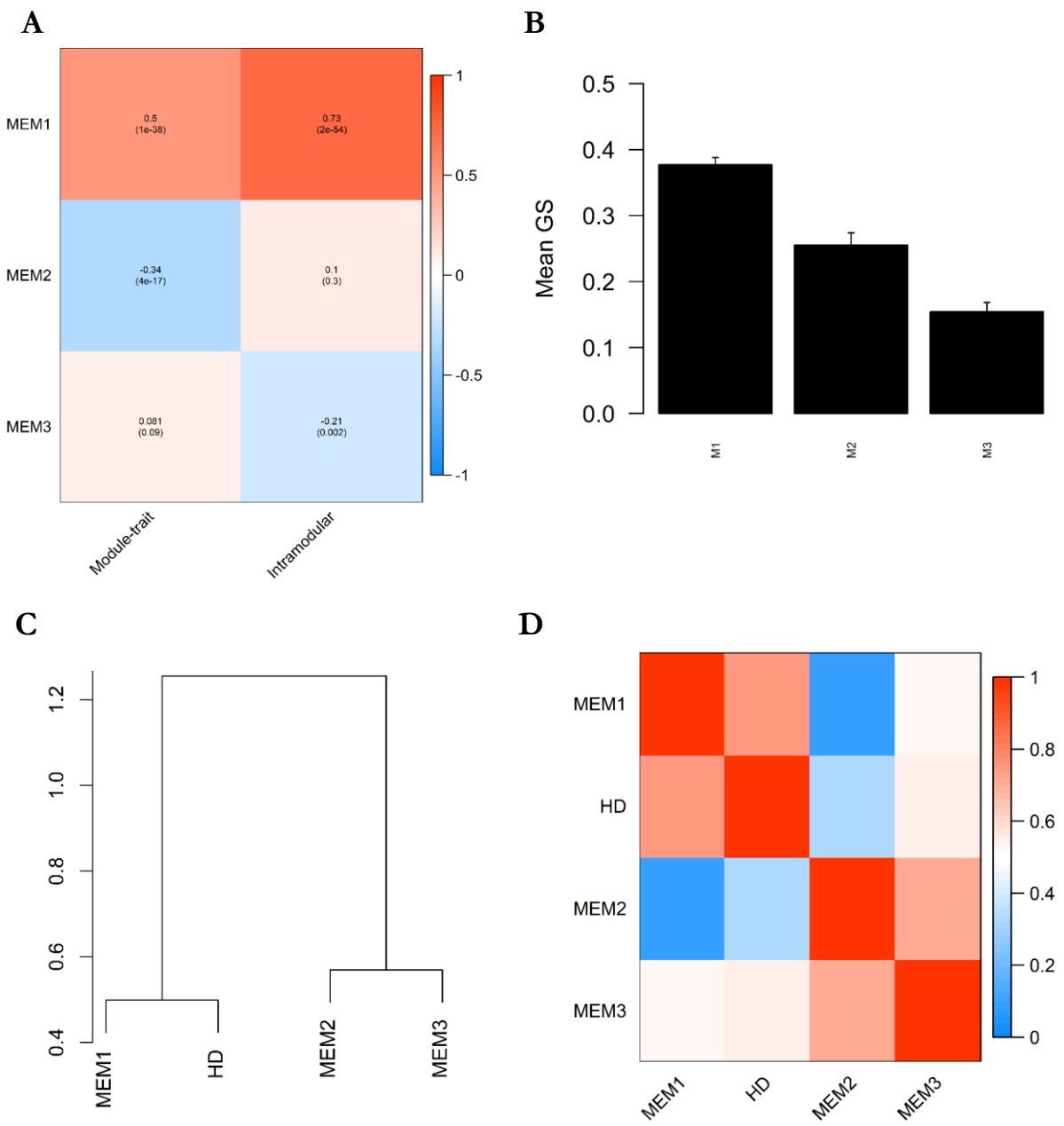

Figure 2: Correlation of WGCNA meta-modules and the association of the disease state with module eigengenes. A. Heatmap showing the correlation between the meta-module eigengenes (MEM1, MEM2, and MEM3), and disease state (HD) or between gene significance (GS) and meta-module membership. P-values adjusted after Benjamini & Yekutieli (Yekutieli & Benjamini, 2001) are given in brackets. B: Mean gene significance (GS) of each meta-module. Error bars depict the 95% confidence interval. C. Dendrogram showing hierarchical clustering of module eigengenes. D. Eigengene adjacency heatmap.

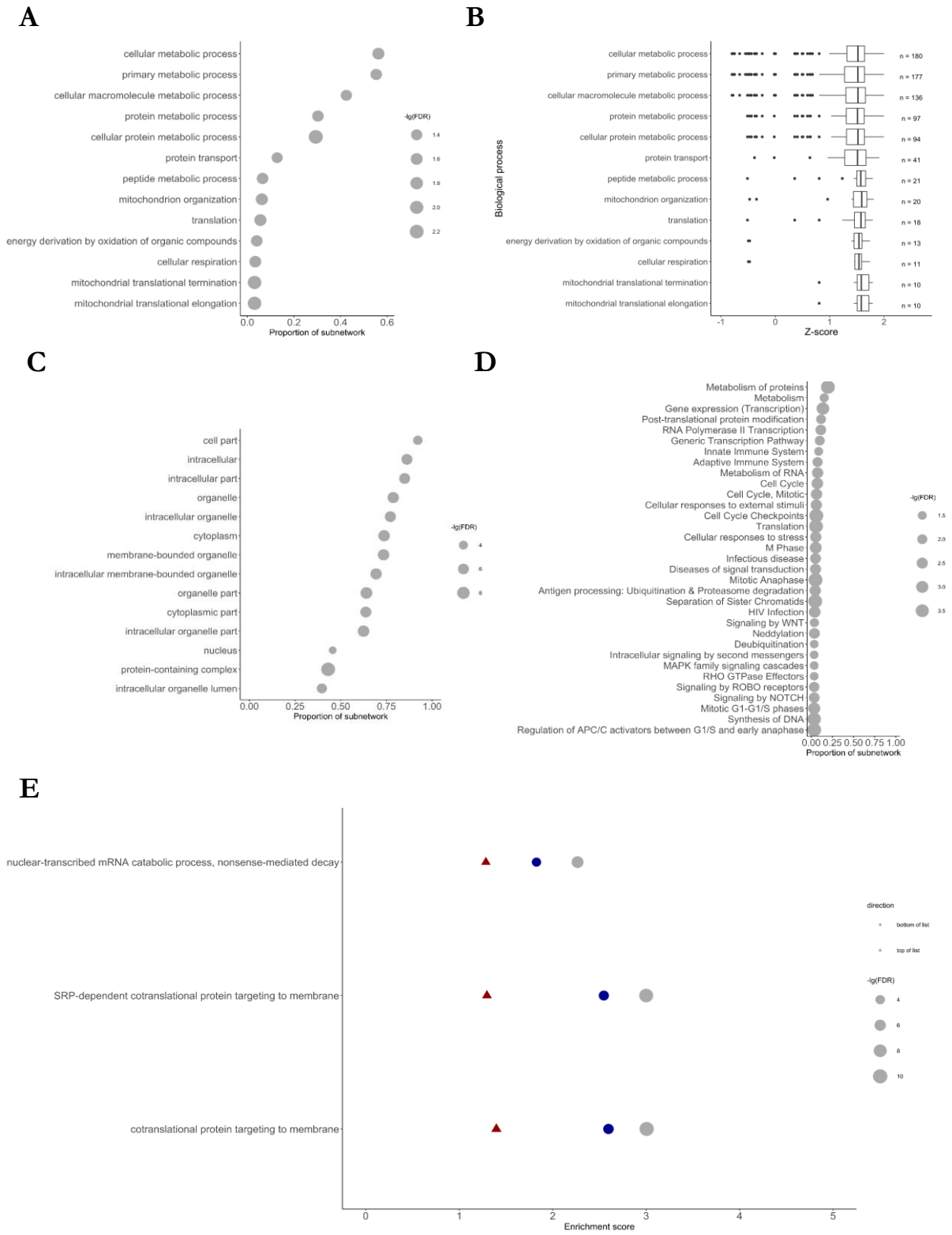

Figure 3: Enrichment analysis of M1 subnetwork. A: Gene ontology (GO) term enrichment analysis for biological processes. B: Mean Z-ratios of genes belonging to the enriched biological processes. C: Gene ontology (GO) term enrichment analysis for cellular compartments. D: Enrichment of Reactome pathways. The size of the circles depicts the negative decadic logarithm of the false-discovery rate (FDR). E: Result of gene set enrichment analysis against the gene ontology database. Only gene sets identified in each dataset (red: GSE33000; blue: GSE129473; grey: GSE64810) are shown and sets were ordered according to the score of the RRA analysis.

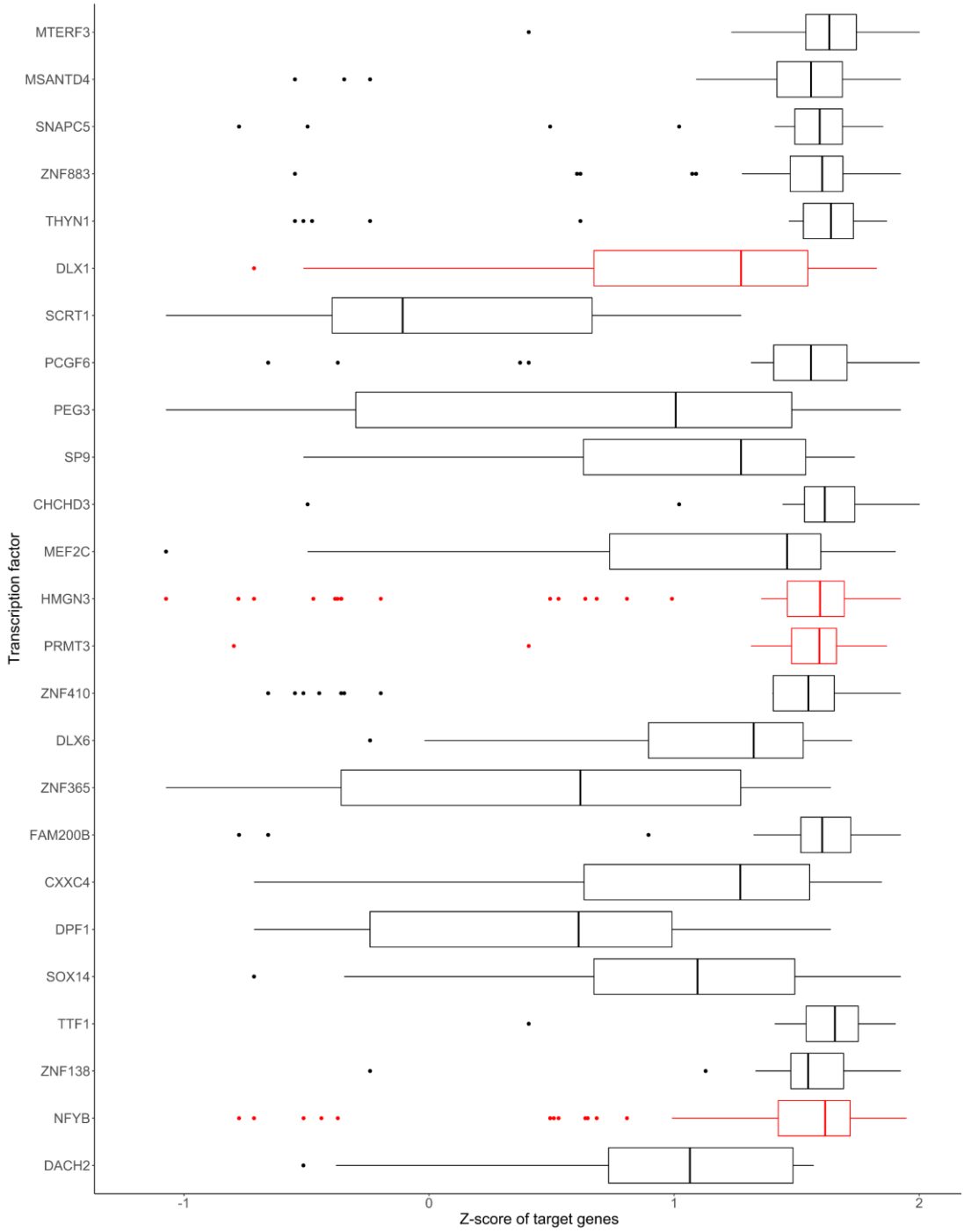

Figure 4: Z-Scores of target genes from the top 25 enriched transcription factors. Enrichment of transcription factors was computed with the Chea3 tool and the Z-ratios for each gene of M1 controlled by the respective transcription factor was averaged for the three different studies. Transcription factors were ordered based on the result of the transcription factor enrichment analysis with the best-ranked transcription factor at the top of the graph and transcription factors with robustly altered mRNA levels were depicted in red.

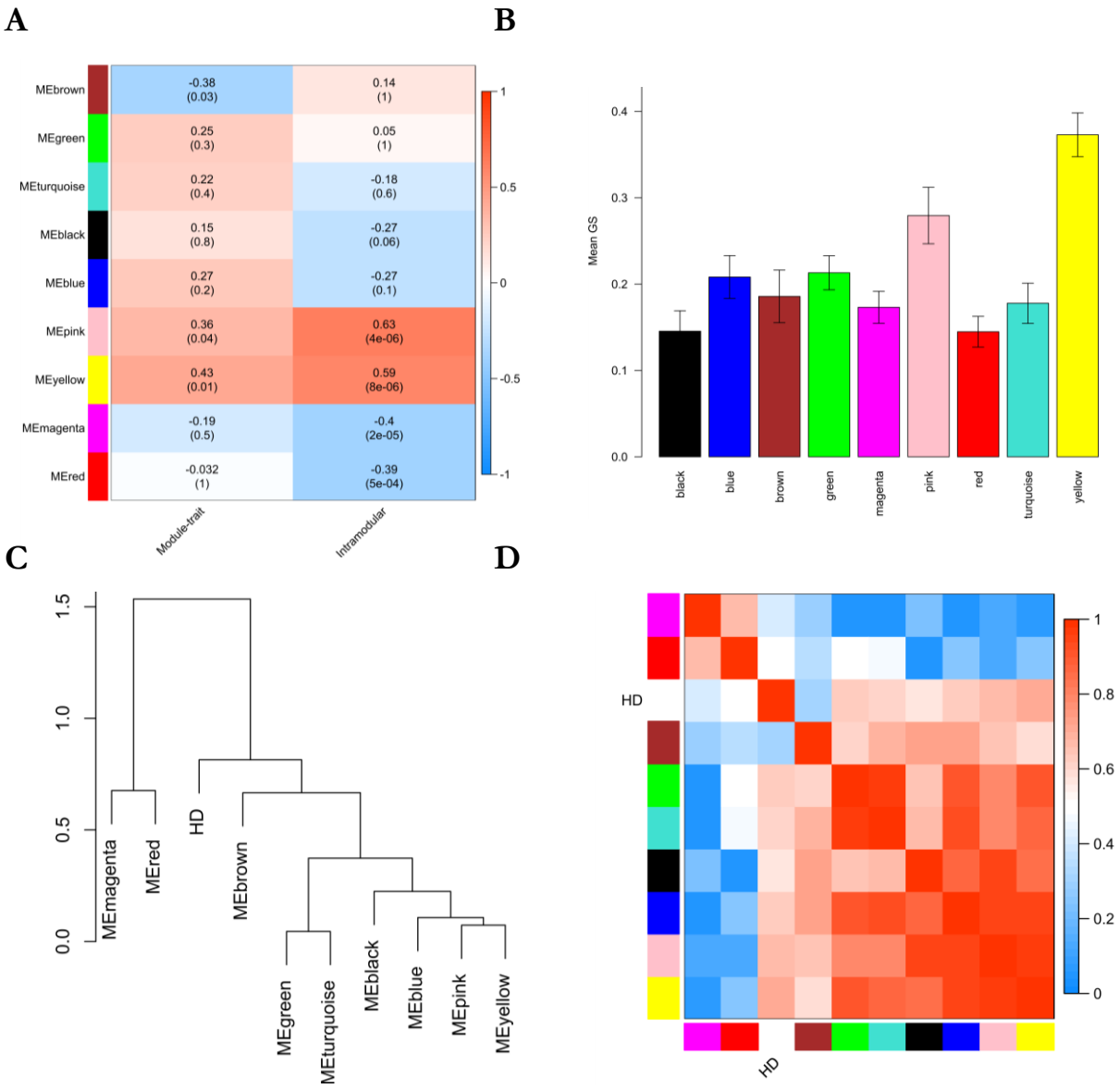

Figure 5: Result of WGCNA analysis and the association of the disease state with module eigengenes. A. Heatmap showing the correlation between module and disease state or between gene significance and module membership. P-values adjusted after Benjamini & Yekutieli (Yekutieli and Benjamini, 2001) are given in brackets. B: Mean gene significance of each module. Error bars depict the 95% confidence interval. C. Dendrogram showing hierarchical clustering of module eigengenes. D. Eigengene adjacency heatmap.

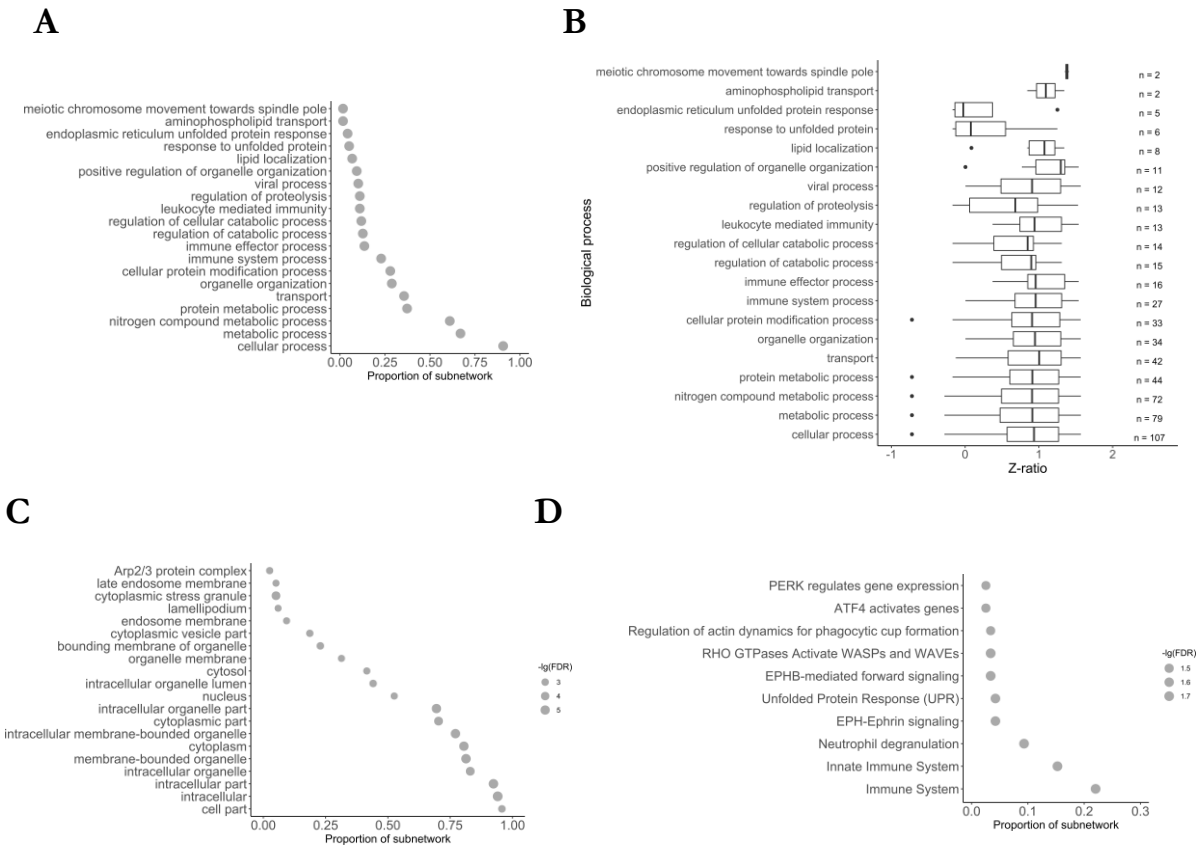

Figure 6: Enrichment analysis of M1 meta-module. A: Gene ontology (GO) term enrichment analysis for biological processes. B: Mean z-ratios of genes belonging to the enriched biological processes. C: Gene ontology (GO) term enrichment analysis for cellular compartments. D: Enrichment of Reactome pathways.

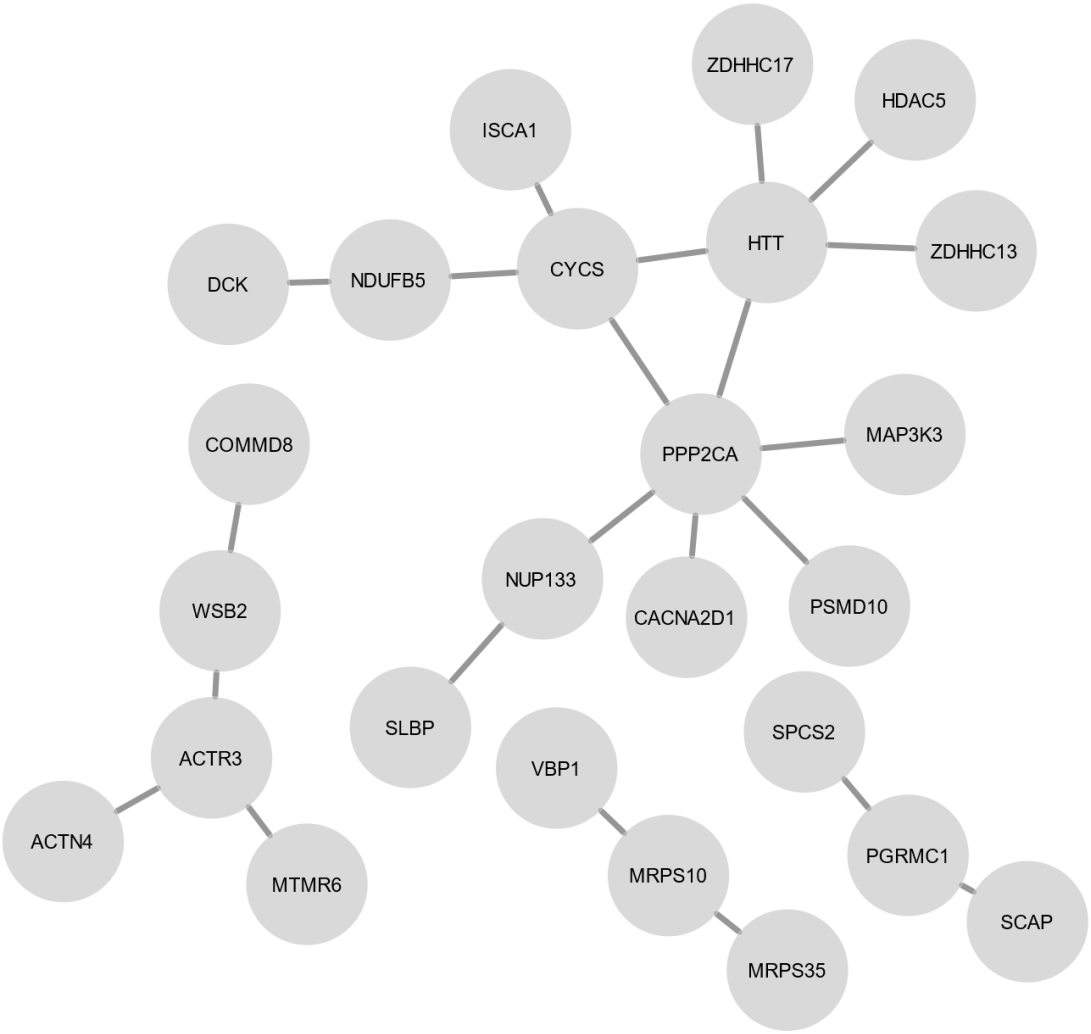

Figure 7: Protein-protein interaction map of proteins with altered mRNA levels in the blood and brain of HD patients. Protein-protein interaction map was constructed the STRING database and plotted with Cytoscape. For illustration purposes, we added HTT to the protein interaction map, although mRNA HTT were not robustly altered.

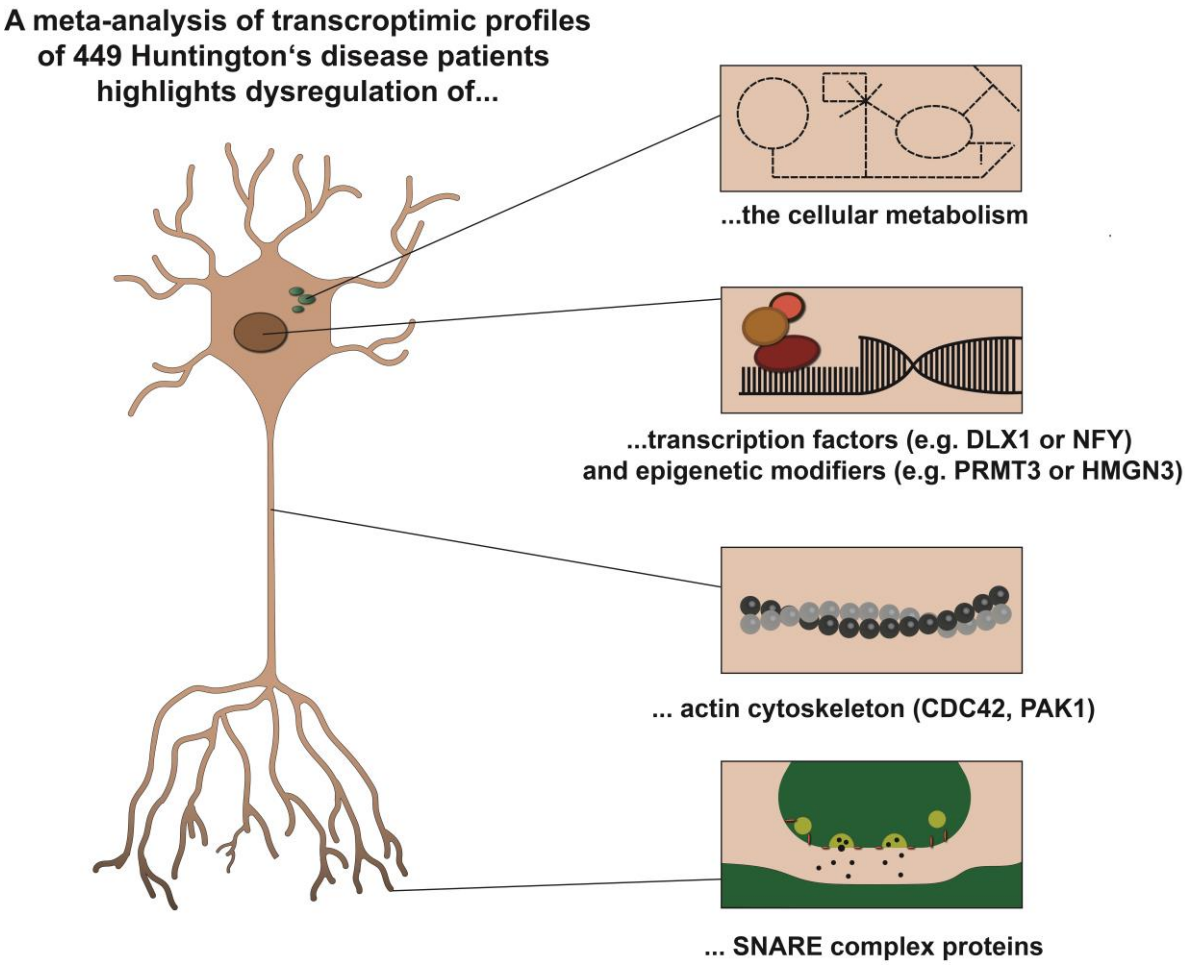

Figure 8: Schematic illustration of transcriptomic alterations in the brain of HD patients.

Tables

Table 1: Top ten hub genes of meta-module M1. Hub genes were determined based on their intramodular connectivity, module membership and gene significance as described in the methods section. A positive Z-score shows upregulation in HD samples, while a negative Z ratio shows a downregulation. The complete list of identified hub genes can be retrieved from the supplement. IC: intramodular connectivity; MM: module membership; GS: gene significance.

Gene	IC	MM	GS	Z-ratio GSE33000	Z ratio GSE129473	Z ratio GSE64810
C3orf14	60.22	0.96	0.48	2.01	0.46	1.77
ATP6AP2	59.31	0.95	0.42	1.72	1.05	1.77
ISCA1	57.93	0.95	0.47	1.96	0.75	1.77
B3GALNT1	57.11	0.95	0.44	1.80	1.27	1.50
POLR2K	54.55	0.94	0.49	1.98	1.44	2.30
PAK1	54.65	0.92	0.44	2.07	- 0.42	0.30
ACP1	53.75	0.94	0.45	1.82	1.22	1.76
CDC42	53.60	0.94	0.42	1.72	1.83	1.45
EID1	53.22	0.93	0.44	1.75	1.11	2.20
RCN2	52.24	0.93	0.43	1.75	1.48	1.95

Additional files

- **Additional file 1 (Excel file, .xlsx):** Genes with robustly altered mRNA levels in the brain and blood of HD patients.
- **Additional file 2 (Excel file, .xlsx):** Selection of the soft-thresholding power β.
- **Additional file 3 (Excel file, .xlsx):** Result of neat analysis
- **Additional file 4 (Excel file, .xlsx):** Identified WGNA hub genes.
- **Additional file 5 (Excel file, .xlsx):** Hubs of protein-protein interaction networks.
- **Additional file 6 (Excel file, .xlsx):** Results of transcription factor enrichment analysis.
- **Additional file 7 (Excel file, .xlsx):** CREB-1 transcription factor subnetwork.
- **Additional file 8 (Excel file, .xlsx):** Alteration of mRNA levels in the striatum of R6/2 and YAC128 mice.
- **Additional file 9 (Excel file, .xlsx):** Results of ROC analysis of selected genes and transcriptional regulators.
- **Additional file 10 (Portable Document Format, .pdf):** ROC curves of selected hub genes and transcriptional regulators.