

# Regulatory network underlying alterations of gene expression in early death of Huntington's disease

**Zhike Zhou**

First Affiliated Hospital of China Medical University

**Fenqin Chen**

First Affiliated Hospital of China Medical University

**Shanshan Zhong**

First Affiliated Hospital of China Medical University

**Yi Zhou**

University of Georgia

**Rongwei Zhang**

First Affiliated Hospital of China Medical University

**Kexin Kang**

First Affiliated Hospital of China Medical University

**Xiaoqian Zhang**

First Affiliated Hospital of China Medical University

**Mei Zhao**

Shengjing Hospital of China Medical University

**Ying Xu** (✉ [xyn@uga.edu](mailto:xyn@uga.edu))

University of Georgia

**Chuansheng Zhao** (✉ [cszhao@mail.cmu.edu.cn](mailto:cszhao@mail.cmu.edu.cn))

First Affiliated Hospital of China Medical University <https://orcid.org/0000-0002-4100-1259>

---

## Research article

**Keywords:** early death, Huntington's disease, hub gene, network

**Posted Date:** June 4th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-32034/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

The aim of this study is to explore the genomic mechanism of early death in Huntington's disease (HD). We identified 10160 and 1511 differentially expressed genes (DEGs) from comparisons of HD versus control and early versus late death, respectively. On the basis of 922 overlapped DEGs among them, six functional modules were established by weight gene correlation network analysis. The turquoise module most strongly related to overall survival of HD was significantly enriched in GABAergic synapse, retrograde endocannabinoid signaling and neuroactive ligand-receptor interaction. Low expression of five DEGs (CA10, WSB2, ACTR3B, PCDH19 and GABRB3) with highest degree of connectivity and non-zero regression coefficients were identified as hub genes, based on which the LASSO model exactly predicted the occurrence of early death in HD. Furthermore, the network analysis revealed indirect interaction of CA10 with the central hub gene GABRB3, which was involved in the pathway of GABAergic synapse. Compared with high expression of hub genes, their low expression appeared shorter overall survival of HD patients according to Kaplan-Meier survival analysis. Consequently, low expression of CA10, WSB2, ACTR3B, PCDH19 and GABRB3 were possibly associated with early death of HD. The likelihood of low expression of GABRB3 regulated by CA10 in GABAergic synapse contributed to the pathogenesis of early death in HD.

## Background

Huntington's disease (HD) is an autosomal dominant, progressive neurodegenerative disorder characterized by neuropsychiatric disturbances, cognitive impairment and motor dysfunction [1]. This clinical entity is caused by CAG trinucleotide repeat expansion in huntingtin gene, resulting in neuronal dysfunction and necrosis by a mutant huntingtin protein [2]. The onset of HD is generally occurring in mid-adult, with an average longevity of proximate 60 years old [3]. However, it is not only a condition of adult-onset disease, but also be symptomatic in juvenile – there are cases of early-death or late-onset accompanied by atypical clinical symptoms, such as Westphal variant with severe developmental retardation in children, as well as very mild symptoms in elderly patients [4,5]. A fundamental but unresolved issue emerges regarding what exactly determines the progression and mortality in patients with HD, which likely shed light on the anticipation of this incurable disease to be not untreatable.

Despite decades of research into the mutant huntingtin protein since the discovery, yet its function remains elusive and, most patients ultimately suffer from intolerable disabilities and early death [6,7]. A growing emphasis on prominent group of genes encoding huntingtin-interacting proteins, provides valuable clues to illustrate huntingtin functions and the pathophysiology of HD [8]. In case of Slc12A5, which is highly enriched and down-regulated in huntingtin proteome [9,10], disrupting cytoskeletal interactions and promoting developmental apoptosis of cortical projection neurons [11]. Moreover, targeting Gpr52 reduces the levels of mutant huntingtin and rescues HD-related phenotypes by the cyclic adenosine monophosphate-dependent signaling pathway in cellular and mouse models [12]. Notably, these studies only highlighted the involvement of genes in the onset of HD, rather than intensive investigation of the genetic effect on prognosis or mortality. As the inherited certainty was taken into

account, we sought to identify causative genes associated with diversities of overall survival in HD to immediately target the early death mechanism of the primary disorder.

## Materials And Methods

### Data resources

The gene expression profile of GSE33000 dataset were collected in the Gene Expression Omnibus database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). [13]. This dataset was generated using the Rosetta/Merck Human 44k 1.1 microarray from postmortem prefrontal cortex samples between 157 HD patients and 157 non-demented controls with matched genotype and clinical data. The data of gene expression profile were normalized by application of the *normalizeBetweenArrays* function in the *limma* package of R software version 3.6.2 [14].

### Gene set enrichment analysis (GSEA)

The biological process (BP) Gene Ontology terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that were potentially related to HD and early death in GSE33000 dataset were filtrated by GSEA analysis [15,16]. The statistic of default weight was used to conduct the permutation for 1000 times and normalized  $P < 0.05$  was considered as significant enrichment. Thereafter, visualization of enrichment data in GSEA analysis were accomplished by adapting packages of *ClusterProfiler*, *ggplot2*, *enrichplot* and *GSEABase* in R.

### Identification of differentially expressed genes (DEGs)

The median time of overall survival was identified as the cut-off point in order to dichotomize subjects into early death and late death groups. The differential expression analysis was carried out using the *limma* package of R software, with  $P < 0.05$  adjusted by the false discovery rate considering statistically significant [14,17]. The *lmFit* and *eBayes* functions of *limma* package were utilized to obtain the DEGs for HD versus control group and early death versus late death cohort, respectively.

### Weight gene correlation network analysis (WGCNA)

WGCNA was conducted based on the intersection genome of DEGs between HD/control group and early/late death cohort. The outlier samples were excluded to guarantee the dependability of network after plotting the clustering dendrogram by *hclust* function. During the module construction, the value of soft thresholding power was filtered using the *pickSoftThreshold* function, and the appropriate power value of 9 was determined by *sft\$powerEstimate* function. The co-expression network was performed by the *WGCNA* package, with each module designating a unique color label and the minimum size setting to 30 [18]. Thereafter, topological overlap matrix was calculated to ascertain modules by hierarchically clustering eigengenes [19], and the network heatmap of the interaction association between modules was visualized using *TOMplot* package. Each module in the heatmap was represented by a different color for the horizontal and vertical axis. The yellow brightness in the unit of intermediate connection indicated the

connectivity degree of different modules. Additionally, functional enrichment analysis of each module was conducted by utilizing the *clusterProfiler* package.

### Identification of hub genes and construction of protein-protein interaction (PPI) network

In WGCNA, hub genes referring to the highest degree of connectivity in a module were used to determine the biological significance of the module. The absolute value of Pearson's correlation was adapted to analyze the relationship between intramodular connectivity and clinical trait, which were measured by module membership (MM) and gene significance (GS), respectively [20]. An eigengene that met the criteria of Pearson's correlation ( $MM \geq 0.9$  and  $GS \geq 0.5$ ) was defined as a candidate hub gene. The pairwise scatterplot matrix of the relationship between candidate hub genes was quantitatively analyzed with *gpairs* package. Based on the STRING database (Search Tool for the Retrieval of Interacting Genes, <https://www.string-db.org/>) [21], DEGs linked to the interaction of hub genes with early death and HD were extracted to construct the PPI network of the clinically significant module. The *cytoscape* software was used to visualize the global regulatory network, intersection pathways of DEGs and GABAergic synapse pathway [22].

### Analysis of receiver operating characteristic (ROC) curve and Kaplan-Meier survival curve

Least absolute shrinkage and selection operator (LASSO) method was applied to eliminate factors with relatively insufficient contribution and select the most useful predictors as the hub genes for early death of HD [23]. The *glmnet* package was used to construct the LASSO model according to the expression profile of candidate hub genes. The regression coefficient of LASSO analysis was calculated to establish the model index of each sample, with the gene expression value weighted by the following formula:  $\text{index} = \text{ExpGene1} * \text{Coef1} + \text{ExpGene2} * \text{Coef2} + \text{ExpGene3} * \text{Coef3} + \dots$  (Coef: gene regression coefficient; Exp: gene expression value). The ROC curve was plotted using the *pROC* package and the diagnostic performance of LASSO model in predicting early death of HD was estimated by the area under the curve (AUC). Generally, an AUC value of 100% indicated a complete prediction, while 50% represented a random selection. In addition, analysis of time-dependent Kaplan-Meier survival curve was conducted by the *survival* package [24]. All P values were bilateral with those less than 0.05 considered statistically significant.

## Results

### Identification of differentially expressed genes

Based on the clinical data of all subjects, the mean overall survival of 157 HD patients ( $55.85 \pm 14.85$ ) was significantly shorter than that of 157 control samples ( $63.52 \pm 9.91$ ;  $P = 1.144e-007$ ) (Figure 1A). After eliminating the unavailable genes, a total of 19414 genes were included in this study. Among them, 10160 DEGs with 4885 down-regulated and 5275 up-regulated were identified between HD and controls (Figure 1B); meanwhile, 1511 DEGs with 829 down-regulated and 682 up-regulated were established in early death compared with late death cohort (Figure 1C). Finally, we extracted the intersection genome of

922 DEGs that were commonly down-regulated or up-regulated in both HD/control and early/late death group. The plot of heatmap summarized the top 25 down-regulated and up-regulated DEGs of intersection genome that were related to HD and early death (Figure 1D).

#### Module associated with overall survival of HD

Six modules with unique colors were generated by WGCNA in the trait of overall survival in HD (Figure 2A). As shown in Figure 2B, the green module (correlation coefficient = 0.26,  $P = 3e-06$ ) and yellow module (correlation coefficient = 0.29,  $P = 1e-07$ ) were positively correlated with overall survival of HD; whilst the turquoise module (correlation coefficient = -0.51,  $P = 1e-22$ ), blue module (correlation coefficient = -0.34,  $P = 6e-10$ ) and brown module (correlation coefficient = -0.34,  $P = 1e-09$ ) were negatively associated with overall survival of HD. The heatmap of network (Figure 2C) demonstrated no significant difference in interactions among different modules, indicating the high-scale personalization between modules and the relative independence of eigengenes in each module. The scatter diagram of the correlation between GS and MM was illustrated (Figure 2D), of which the turquoise module (correlation coefficient = 0.5,  $P = 4.6e-27$ ) was the most relevant module to the trait of overall survival in HD.

#### Functional enrichment analysis of modules and the LASSO model of hub genes

The analysis of functional enrichment revealed that turquoise module DEGs (Figure 3A) was involved in BP of membrane potential regulation, nervous system process and chemical synaptic transmission; as well as KEGG pathways (Figure 3B) enriched by GABAergic synapse, retrograde endocannabinoid signaling, morphine addiction and neuroactive ligand-receptor interaction. According to  $MM > 0.9$  and  $GS > 0.5$ , seven DEGs were identified as candidate hub genes (CA10, WSB2, NETO2, LPPR4, ACTR3B, PCDH19 and GABRB3), all of which were positively interacted with each other (Figure 3C). The LASSO model was constructed based on the expression profile of candidate hub genes (Figure 3D). Two (NETO2 and LPPR4) of them with zero regression coefficients were eliminated, whereas the remaining five DEGs (CA10, WSB2, ACTR3B, PCDH19 and GABRB3) were identified as hub genes ( $\lambda_{\min} = 0.01264$ ). The index formula of hub gene-based model was generated as follows:  $\text{index} = \text{CA10} * (-0.9456) + \text{WSB2} * 0.2320 + \text{ACTR3B} * 1.1667 + \text{PCDH19} * 0.3986 + \text{GABRB3} * 1.4177$ .

#### PPI network and GSEA

The DEGs of turquoise module were extracted to construct the PPI network of global regulation based on STRING database (Figure 4A). Furthermore, we found that GABRB3 was jointly involved in the enrichment of KEGG pathways including GABAergic synapse, neuroactive ligand-receptor interaction, morphine addiction and retrograde endocannabinoid signaling (Figure 4B). As shown in figure 4C, CA10 regulated the interactions of GABRB3 with GABRA1, GNB5 and KCNJ6 in the pathway of GABAergic synapse. Gene set enrichment analysis revealed the BP of HD (Figure 5A) and early death (Figure 5B), as well as the KEGG pathways of HD (Figure 5C) and early death (Figure 5D).

#### Performance evaluation of AUC and Kaplan-Meier survival analysis

The LASSO model based on hub genes was used to estimate the AUC values of ROC curve (CA10: AUC = 84.2%, WSB2: AUC = 84.6%, ACTR3B: AUC = 73.6%, PCDH19: AUC = 73.2%, GABRB3: AUC = 70.9%) (Figure 6A). The Kaplan-Meier survival analysis showed that the overall survival of HD patients with low expression of hub gene was significantly shorter than those with high hub gene expression (CA10:  $P = 6.335e-11$ , WSB2:  $P = 3.58e-08$ , ACTR3B:  $P = 1.979e-05$ , PCDH19:  $P = 0.001158$ , GABRB3:  $P = 2.99e-05$ ) (Figure 6B).

## Discussion

During the current study, GSEA was performed to explore the expression profile of 19414 genes belonging to the BP of synapse assembly, synaptic vesicle cycle and neurotransmitter transport; as well as the KEGG pathways of phagosome, neuroactive ligand-receptor interaction and retrograde endocannabinoid signaling. It is worth noting that these genes in the processes and pathways were enriched in HD and early death individuals. Subsequently, a comprehensive characterization of regulatory network for the intersection genome of DEGs between HD and early death was integrated; hereby, the excavation of key target genes and interactive pathways from functional modules with respect to the trait of overall survival might contribute to elucidate the genome-scale pathogenesis of early death in HD patients.

Generally consistent with the results of GSEA, functional enrichment analysis exhibited DEGs of turquoise module participating in BP related to membrane potential regulation, nervous system process and chemical synaptic transmission; as well as KEGG pathways regarding GABAergic synapse, neuroactive ligand-receptor interaction and retrograde endocannabinoid signaling. Convincing previous evidence demonstrated that a series of huntingtin-interacting proteins associated with the deficits of GABAergic synapse yielded the onset of HD [25]. Additionally, endocannabinoids modulated the synaptic transmission through the pathway of retrograde endocannabinoid signaling, resulting in cognitive impairment and movement dysfunction partially similar to the symptoms of HD [26]. Based on  $MM > 0.9$  and  $GS > 0.5$  in WGCNA, seven DEGs including CA10, WSB2, NETO2, LPPR4, ACTR3B, PCDH19 and GABRB3 were identified as potential pathogenic factors for early death and HD. Among them, PCDH19 has been found to coordinate neurogenesis, neuronal migration and synaptic connections in mammalian cortex [27,28]. Low expression of PCDH19 promoted over-differentiation of progenitor cells, thus to premature neuronal apoptosis and structural defects of synapse; contrastively, high expression of PCDH19 improved abnormal neurogenesis, fine-scale spatial organization of excitatory neurons and synaptic junction in the cortex [29]. ACTR3B, also known as actin related protein 3B, had implications in the nucleation and assembly of branched actin networks [30]. It bound to DNA double-strand fractures and, in conjunction with actin, facilitating clustering and homologous directed repair at the breakage [31]. There was evidence that the disturbance in ACTR3B expression was account for abnormal neuronal morphology and synaptic growth, as well as alterations in synapse activity [32-34], which was likely related to the pathogenesis of early death in HD.

Glaringly, in the global network analysis, GABRB3 was identified as a central hub gene at a site with genome-wide significant linkage to overall survival of HD. An observation from high-throughput flow

cytometry assay revealed that GABRB3 was a critical gene for early development of central nervous system, and the functional impairment of GABRB3 was prone to neurodevelopment-related disorders [35]. Specifically, it was increasingly expressed in the embryonic brain, with its encoded  $\beta 3$  subunit involved in the biological processes of stem cell differentiation, assembly and trafficking of GABA receptor [36,37]. The inhibitory synaptic transmission mediated by GABA receptor were susceptible to the accumulation of mutant huntingtin, leading to early degeneration of GABAergic neurons in the striatum of HD patients [38]. Indeed, many of the manifestations that were emerged early in HD, such as hippocampal-dependent learning and memory impairment, were considerably attributed to the weakened inhibitory GABAergic transmission verified in a HD mouse model [25]. This happened to coincide with the results of our network and GABAergic synapse pathway that low expression of the central hub gene GABRB3, interacting with other DEGs (e.g., GABRA1, GNB5 and KCNJ6), were included in the intersection genome between HD and early death, suggesting that GABRB3-mediated GABAergic synapse was probably a crucial pathway for early death of HD. With the exception of GABAergic synapse, the module-pathway network showed that GABRB3 also participated in neuroactive ligand-receptor interaction, morphine addiction and retrograde endocannabinoid signaling. The potential roles of these pathways in the overall survival of HD was not fully understood, and therefore deserved further exploration with an expectation of offering a GABRB3-related therapeutic target for this incurable disease.

Less is known about the properties of carbonic anhydrase CA10, rather than a non-catalytic activity of extracellular protein that has been described for two decades in relation to carbonic anhydrase [39]. Nevertheless, recent study confirmed that CA10, an evolutionarily conserved ligand of neurexins in the surface transposome, served as a functional adaptor to enhance the expression of neurexin proteins, alter post-translational modifications and improve their surface levels [40]. It mediated the interaction between neurexins and certain postsynaptic molecules, thereby facilitating the formation of novel trans-synaptic complexes [40]. Evidence in zebrafish using morpholino-mediated knockdown of CA10 showed developmental retardation, aberrant behavior and early death under unknown mechanisms of CA10 deficiency [41]. Herein, the network analysis exhibited multiple indirect interactions of CA10 with the center hub gene GABRB3, suggesting that GABRB3-mediated GABAergic transmission was possibly modulated by CA10 through its adaptor function. In episodic ataxia type-1 families, carbonic anhydrase inhibitors decreasing the excitability of GABAergic interneurons by intracellular alkalinization were used to alleviate the clinical symptoms of cerebellar ataxia [42]. Correspondingly, this would provide us with a feasible implication to the intervention of early death of HD by modulating CA10 to maintain postsynaptic potentials of GABAergic inhibition.

Visualization of the global regulatory network showed the down-regulation of candidate hub genes that were strongly related to the clinical trait of overall survival in HD patients. Further matrix of pairwise scatterplot confirmed a positively mutual correlation between them, implying that the expression of a gene subsided with the attenuation of the other. This might be illustrated by the involvement of candidate hub genes underlying the pathogenesis of early death in HD, particularly CA10, which indirectly interacted with GABRB3 in the pathway of GABAergic synaptic transmission [29,34,38,41]. Five genes (CA10, WSB2, ACTR3B, PCDH19 and GABRB3) with non-zero regression coefficients were selected to conduct the ROC

analysis, of which the AUC values were all greater than 70%, supporting that they were hub genes and could be served as potential predictors for early death in HD worthy of clinical application [43]. In addition, the verification of the Kaplan-Meier survival curve demonstrated that, compared with the high expression of CA10, WSB2, ACTR3B, PCDH19 and GABRB3, their low expression was associated with shorter overall survival of HD patients. It might be an intriguing hypothesis that the improvement of low expression of the hub genes was prospectively used for targeted treatment or prevention to HD patients with a high risk of early death. Since our prediction was based on an analysis of post-mortem samples, subsequent in vivo model experiments are required to validate the underlying mechanisms of the hub genes involved in the early death of HD.

## Conclusions

Low expressions of CA10, WSB2, ACTR3B, PCDH19 and GABRB3 have important implications in the pathogenesis of early death in HD. GABAergic synapse is potentially a crucial pathway by which CA10 interacts intimately with GABRB3 to yield early death of HD. The results emerging from our study stand poised to extend a fundamental guidance, in theory, for prognosis prediction and a promising intervention into the target of GABRB3 modulated by CA10 for early death of HD.

## Abbreviations

AUC: area under the curve; BP: biological process; Coef: gene regression coefficient; DEGs: differentially expressed genes; Exp: gene expression value; GEO: Gene Expression Omnibus; GS: gene significance; GSEA: Gene set enrichment analysis; HD: Huntington's disease; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: Least absolute shrinkage and selection operator; MM: module membership; PPI: protein-protein interaction; ROC: receiver operating characteristic; STRING: Search Tool for the Retrieval of Interacting Genes; WGCNA: weight gene correlation network analysis.

## Declarations

### Acknowledgments

All the authors of this study should be appreciated sincerely.

### Authors' contributions

Conceptualization: ZKZ, YX and CSZ; Methodology: ZKZ, FQC, SSZ and YZ; Investigation: SSZ, RWZ, XQZ, KXK and MZ; Writing-Original draft: ZKZ, FQC, MZ; Writing-Review and Editing: ZKZ, YX, CSZ; Funding Acquisition: ZKZ, CSZ. All authors revised and approved the final version of the manuscript after review.

### Funding

The study is supported by the National Natural Science Foundation of China (No. 81372104); the Natural Science Foundation of Liaoning Province (No. 20180540150); the Shenyang Population and Health Technical Critical Special Project (No. F16-206-9-01); the Program of the Distinguished Professor of Liaoning Province (Chuansheng Zhao); Guidance plan for key research and development plans of Liaoning Province (No. 2019JH8/10300002).

### **Availability of data and materials**

The datasets utilized and/or analyzed in the present study are available from the Gene Expression Omnibus (GEO) repository GSE33000 in the National Center for Biotechnology Information (NCBI). The BioProject identifier of the samples is PRJNA146519.

### **Ethics approval**

Not applicable.

### **Consent for publication**

All authors read and approved the final version of the manuscript.

### **Competing interests**

The authors declare no conflict of interest in this work.

## **References**

1. Fischbeck KH, Wexler NS. Oligonucleotide Treatment for Huntington's Disease. *N Engl J Med.* 2019;380:2373-2374.
2. Tabrizi SJ, Leavitt BR, Landwehrmeyer GB, Wild EJ, Saft C, Barker RA, et al. Targeting Huntingtin Expression in Patients with Huntington's Disease. *N Engl J Med.* 2019;380:2307-2316.
3. Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR, et al. Huntington disease. *Nat Rev Dis Primers.* 2015;1:15005.
4. Dolgin E. When Huntington's disease comes early. *Nature.* 2018;557:S44-S45.
5. Töpper R, Schwarz M, Lange HW, Hefter H, Noth J. Neurophysiological abnormalities in the Westphal variant of Huntington's disease. *Mov Disord.* 1998;13:920-8.
6. Brody H. Huntington's disease. *Nature.* 2018;557:S35.
7. Schapira AH, Olanow CW, Greenamyre JT, Bezdard E. Slowing of neurodegeneration in Parkinson's disease and Huntington's disease: future therapeutic perspectives. *Lancet.* 2014;384:545-55.
8. Li SH, Li XJ. Huntingtin-protein interactions and the pathogenesis of Huntington's disease. *Trends Genet.* 2004;20:146-54.

9. Culver BP, Savas JN, Park SK, Choi JH, Zheng S, Zeitlin SO, et al. Proteomic analysis of wild-type and mutant huntingtin-associated proteins in mouse brains identifies unique interactions and involvement in protein synthesis. *J Biol Chem*. 2012;287:21599-614.
10. Shirasaki DI, Greiner ER, Al-Ramahi I, Gray M, Boontheung P, Geschwind DH, et al. Network organization of the huntingtin proteomic interactome in mammalian brain. *Neuron*. 2012;75:41-57.
11. Mavrovic M, Uvarov P, Delpire E, Vutskits L, Kaila K, Puskarjov M, et al. Loss of non-canonical KCC2 functions promotes developmental apoptosis of cortical projection neurons. *EMBO Rep*. 2020;21:e48880.
12. Yao Y, Cui X, Al-Ramahi I, Sun X, Li B, Hou J, et al. A striatal enriched intronic gpcr modulates huntingtin levels and toxicity. *Elife*. 2015;4:e05449.
13. Narayanan M, Huynh JL, Wang K, Yang X, Yoo S, McElwee J, et al. Common dysregulation network in the human prefrontal cortex underlies two neurodegenerative diseases. *Mol Syst Biol*. 2014;10:743.
14. Ritchie ME, Phipson B, Wu D, Hu Y, Law C, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
15. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34:267-73.
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545-50.
17. Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15:R29.
18. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
19. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL. Hierarchical organization of modularity in metabolic networks. *Science*. 2002;297:1551-5.
20. Fredlund E, Staaf J, Rantala JK, Kallioniemi O, Borg A, Ringnér M, et al. The gene expression landscape of breast cancer is shaped by tumor protein p53 status and epithelial-mesenchymal transition. *Breast Cancer Res*. 2012;14:R113.
21. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45:D362-D368.
22. Otasek D, Morris JH, Bouças J, Pico AR, Demchak B. Cytoscape Automation: empowering workflow-based network analysis. *Genome Biol*. 2019;20:185.
23. Sinha P, Delucchi KL, McAuley DF, O'Kane CM, Matthay MA, Calfee CS. Development and validation of parsimonious algorithms to classify acute respiratory distress syndrome phenotypes: a secondary analysis of randomised controlled trials. *Lancet Respir Med*. 2020;8:247-257.

24. Jackson CH. flexsurv: A Platform for Parametric Survival Modeling in R. *J Stat Softw.* 2016;70:i08.
25. Dargaei Z, Bang JY, Mahadevan V, Khademullah CS, Bedard S, Parfitt GM, et al. Restoring GABAergic inhibition rescues memory deficits in a Huntington's disease mouse model. *Proc Natl Acad Sci U S A.* 2018;115:E1618-E1626.
26. Castillo PE, Younts TJ, Chávez AE, Hashimoto Y. Endocannabinoid signaling and synaptic function. *Neuron.* 2012;76:70-81.
27. Pederick DT, Homan CC, Jaehne EJ, Piltz SG, Haines BP, Baune BT, et al. Pcdh19 Loss-of-Function Increases Neuronal Migration In Vitro but is Dispensable for Brain Development in Mice. *Sci Rep.* 2016;6:26765.
28. Pederick DT, Richards KL, Piltz SG, Kumar R, Mincheva-Tasheva S, Mandelstam SA, et al. Abnormal cell sorting underlies the unique X-linked inheritance of PCDH19 epilepsy. *Neuron.* 2018;97:59-66.e5.
29. Lv X, Ren SQ, Zhang XJ, Shen Z, Ghosh T, Xianyu A, et al. TBR2 coordinates neurogenesis expansion and precise microcircuit organization via Protocadherin 19 in the mammalian cortex. *Nat Commun.* 2019;10:3946.
30. Allison SJ. ARP3 in podocyte development. *Nat Rev Nephrol.* 2019;15:124.
31. Hurst V, Shimada K, Gasser SM. Nuclear Actin and Actin-Binding Proteins in DNA Repair. *Trends Cell Biol.* 2019;29:462-476.
32. Hotulainen P, Llano O, Smirnov S, Tanhuanpää K, Faix J, Rivera C, et al. Defining mechanisms of actin polymerization and depolymerization during dendritic spine morphogenesis. *J Cell Biol.* 2009;185:323-39.
33. Rocca DL, Martin SP, Jenkins EL, Hanley JG. Inhibition of Arp2/3-mediated actin polymerization by PICK1 regulates neuronal morphology and AMPA receptor endocytosis. *Nature Cell Biology.* 2008;10:259-271.
34. Stanishneva-Konovalova TB, Kelley CF, Eskin TL, Messelaar EM, Wasserman SA, Sokolova OS, et al. Coordinated autoinhibition of F-BAR domain membrane binding and WASp activation by Nervous Wreck. *Proc Natl Acad Sci U S A.* 2016;113:E5552-61.
35. Shi YW, Zhang Q, Cai K, Poliquin S, Shen W, Winters N, et al. Synaptic clustering differences due to different GABRB3 mutations cause variable epilepsy syndromes. *Brain.* 2019;142:3028-3044.
36. Andäng M, Hjerling-Leffler J, Moliner A, Lundgren TK, Castelo-Branco G, Nanou E, et al. Histone H2AX-dependent GABA(A) receptor regulation of stem cell proliferation. *Nature.* 2008;451:460-4.
37. Laurie DJ, Wisden W, Seeburg PH. The distribution of thirteen GABAA receptor subunit mRNAs in the rat brain. III. Embryonic and postnatal development. *J Neurosci.* 1992;12:4151-72.
38. Rikani AA, Choudhry Z, Choudhry AM, Rizvi N, Ikram H, Mobassarah NJ, et al. The mechanism of degeneration of striatal neuronal subtypes in Huntington disease. *Ann Neurosci.* 2014;21:112-4.
39. Hewett-Emmett D, Tashian RE. Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families. *Mol Phylogenet Evol.* 1996;5:50-77.

40. Sterky FH, Trotter JH, Lee SJ, Recktenwald CV, Du X, Zhou B, et al. Carbonic anhydrase-related protein CA10 is an evolutionarily conserved pan-neurexin ligand. *Proc Natl Acad Sci U S A*. 2017;114:E1253-E1262.
41. Aspatwar A, Tolvanen ME, Ojanen MJ, Barker HR, Saralahti AK, Bäuerlein CA, et al. Inactivation of *ca10a* and *ca10b* genes leads to abnormal embryonic development and alters movement pattern in zebrafish. *PLoS One*. 2015;10:e0134263.
42. Herson PS, Virk M, Rustay NR, Bond CT, Crabbe JC, Adelman JP, et al. A mouse model of episodic ataxia type-1. *Nat Neurosci*. 2003;6:378-83.
43. Zheng P, Yang J, Li Y, Wu J, Liang W, Yin B, et al. Gut Microbial Signatures Can Discriminate Unipolar from Bipolar Depression. *Adv Sci (Weinh)*. 2020;7:1902862.

## Figures

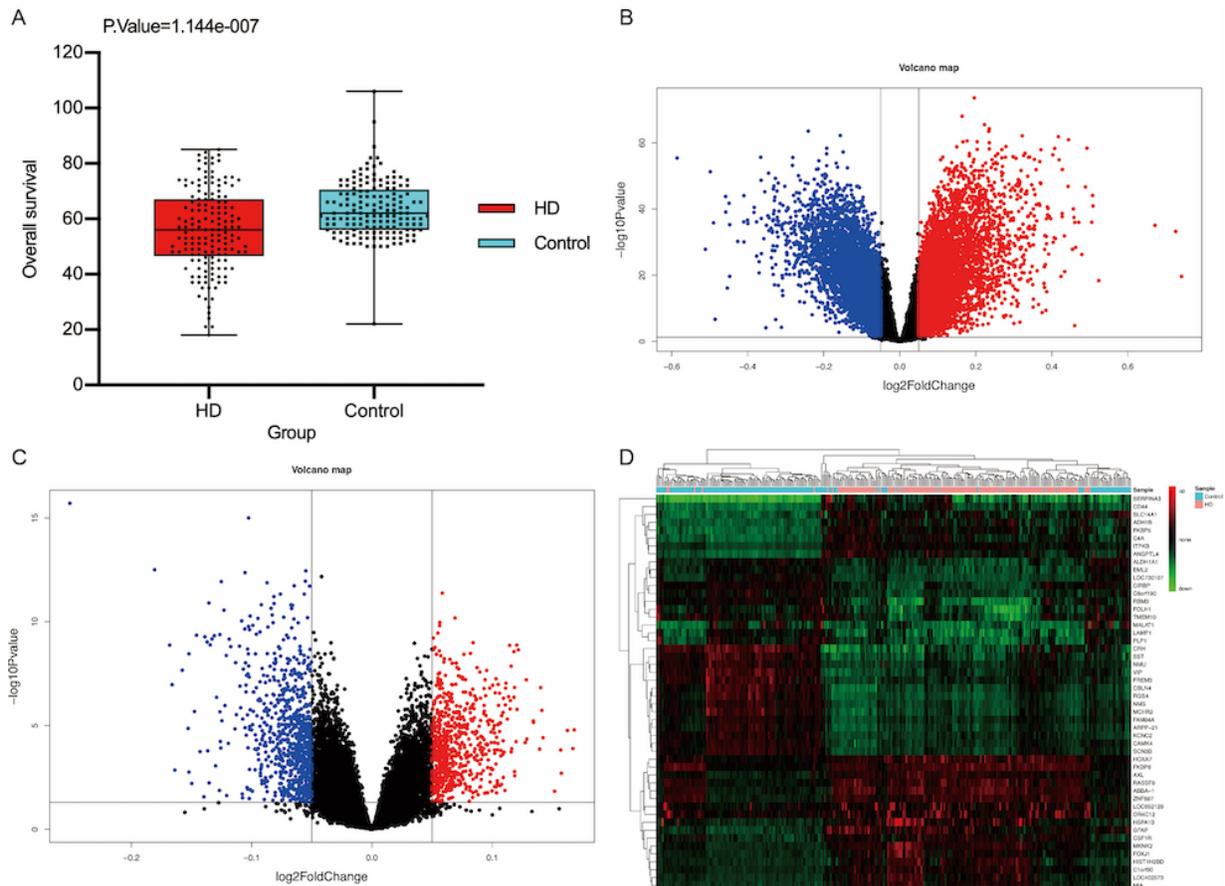
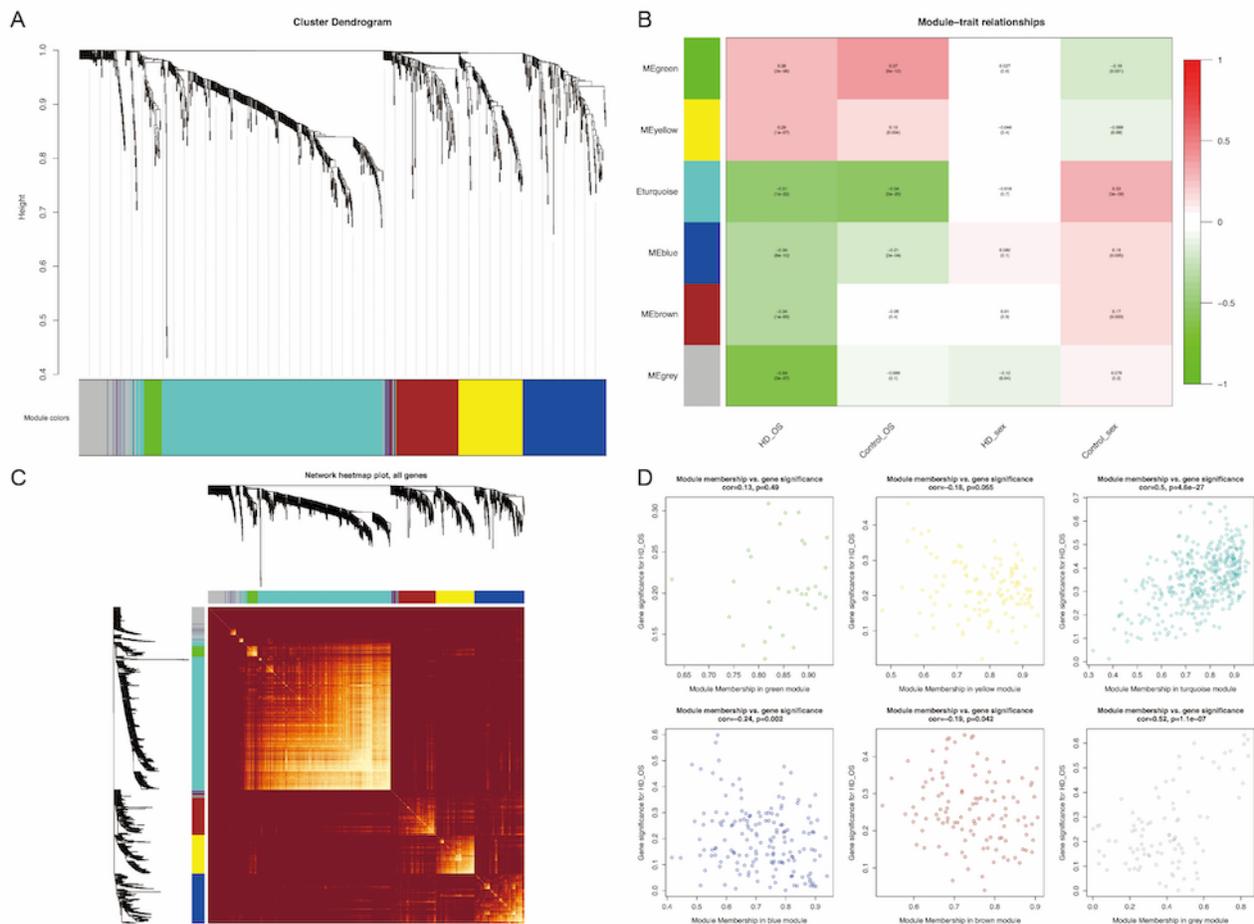


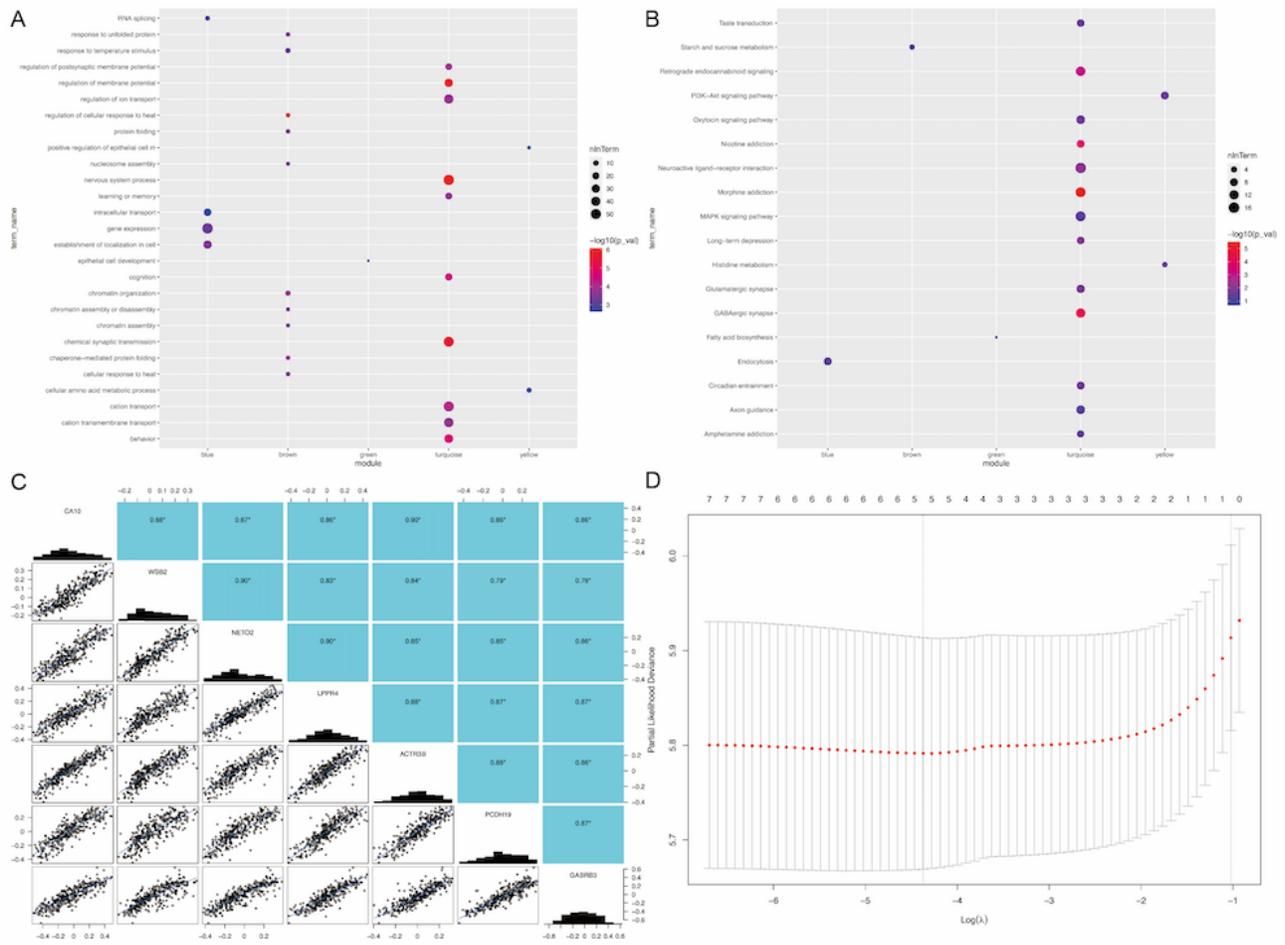
Figure 1

Differential expression gene analysis. The comparison of mean overall survival between HD and controls (A). The volcano plot of the HD/control (B) and early/late death (C): red, blue and black represent up-regulated, down-regulated and non-significantly differentially expressed genes, respectively. The heatmap of top 25 up-regulated and down-regulated genes (D). HD: Huntington's disease.



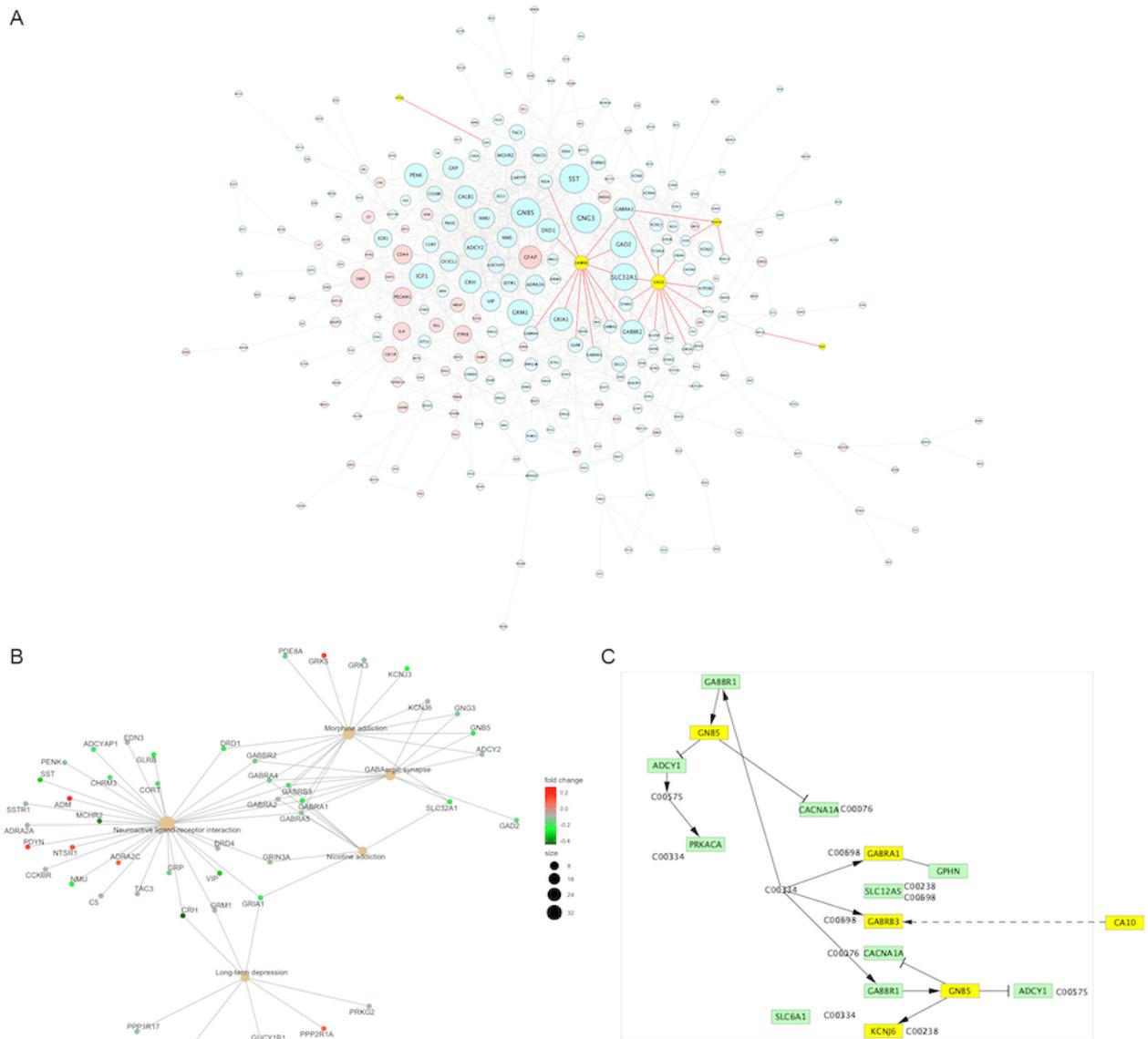
**Figure 2**

Weighted correlation network analysis. The cluster dendrogram with six modules of different colors (A): grey represents non-clustering genes. The correlation heatmap between modules and clinical traits (B): red and green indicate positive and negative correlations with traits, respectively. The network heatmap of the interaction association between modules (C): yellow brightness in the middle represents the connectivity degree of different modules. The scatter plot of eigengenes in each module on the correlations between module membership and gene significance (D). HD: Huntington's disease, OS: overall survival.



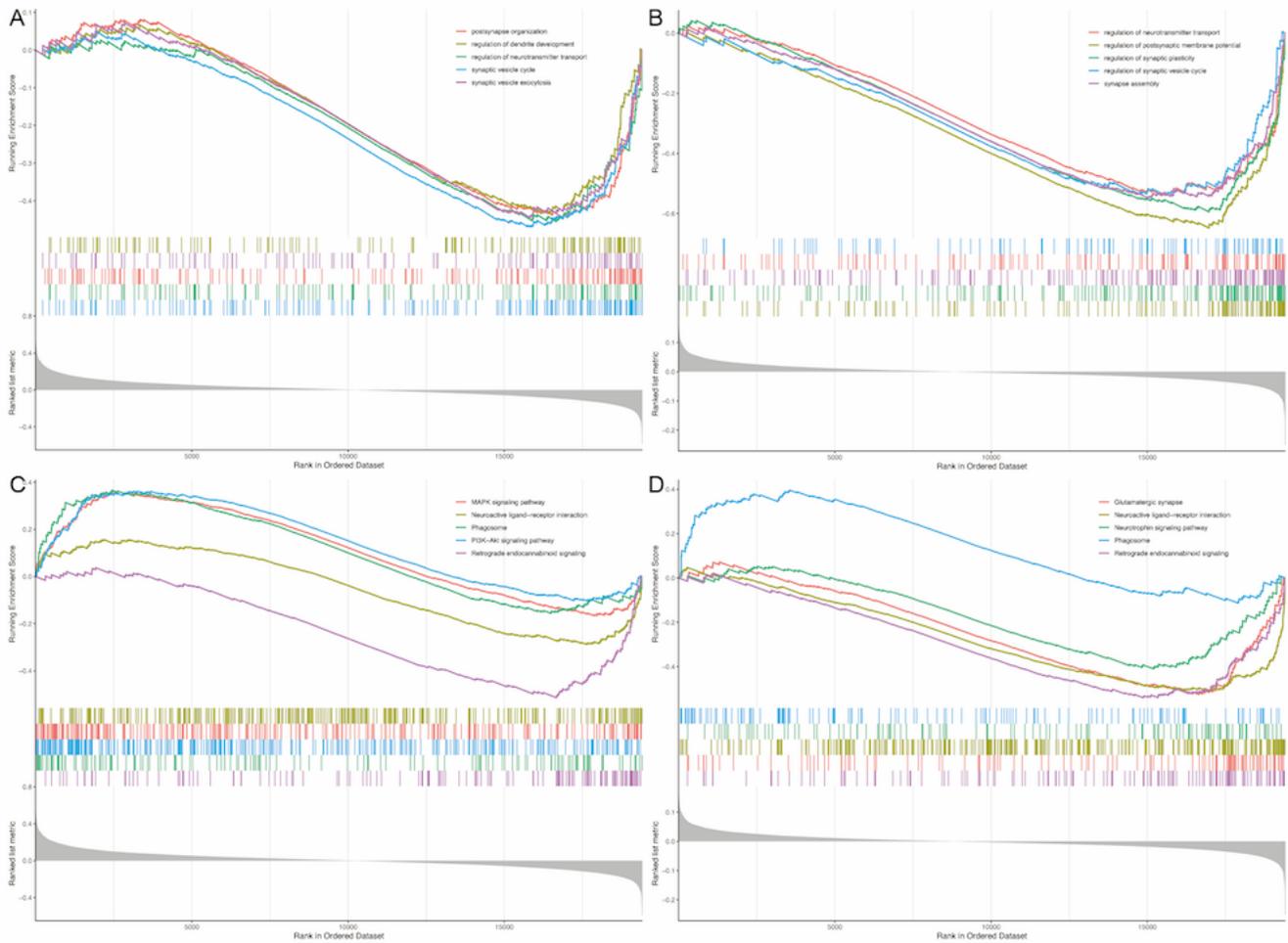
**Figure 3**

Functional enrichment and LASSO model. Enrichment analysis of modules on biological processes (A) and KEGG pathways (B): enrichment significance rises from blue to red and dot size represents the number of eigengenes. The pairwise scatterplot matrix of the relationship between candidate hub genes (C): blue indicates positive correlation. The LASSO model based on candidate hub genes (D). LASSO: Least absolute shrinkage and selection operator, KEGG: Kyoto Encyclopedia of Genes and Genomes.



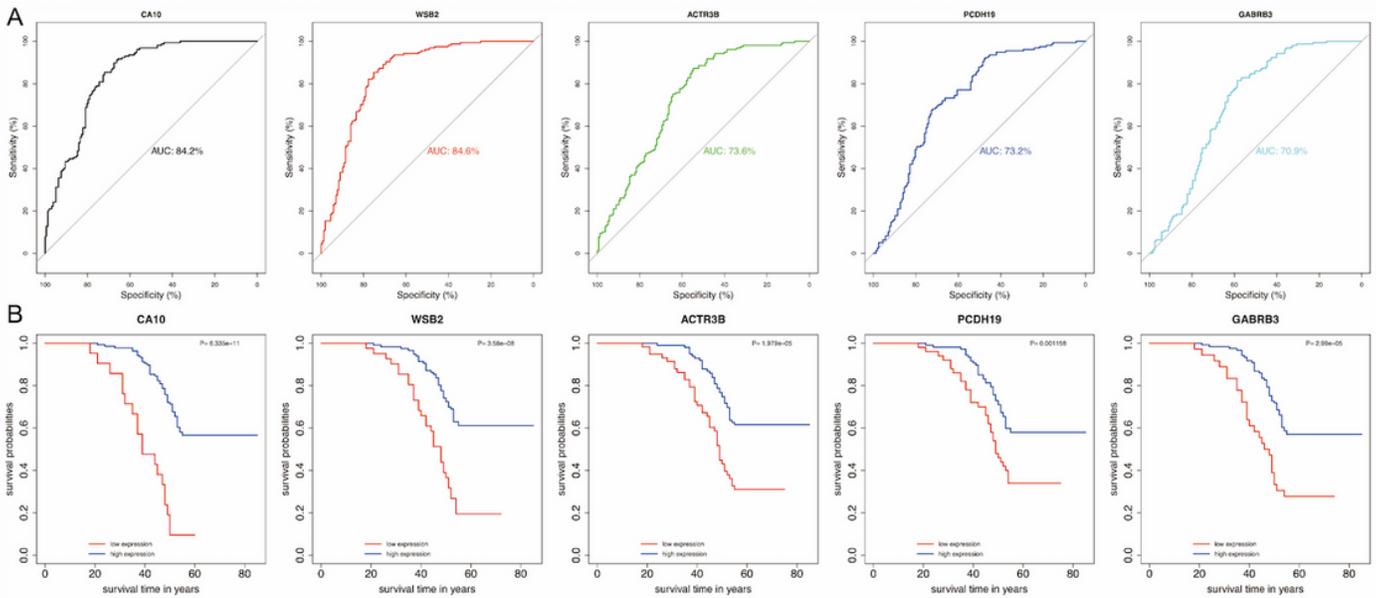
**Figure 4**

Network and GABAergic synapse pathway. The regulatory network of turquoise module (A): yellow nodes indicate hub genes and node size is related to the degree of gene connectivity; red represents high expression of genes, while blue and yellow represents low expression. The model-pathway regulatory network (B): gene expression rises from blue to red; dot size of pathway indicates the number of eigengenes. The pathway of GABAergic synapse (C): yellow represents the eigengenes included in turquoise module.



**Figure 5**

Gene set enrichment analysis. Biological processes enriched in HD (A) and early death (B). KEGG pathways enriched in HD (C) and late death (D). HD: Huntington's disease, KEGG: Kyoto Encyclopedia of Genes and Genomes.



**Figure 6**

Performance evaluation of AUC and Kaplan-Meier survival analysis. ROC curves analysis of hub genes (A). Survival analysis of hub genes (B): red and blue represents low and high expression of genes, respectively. AUC: area under the curve, ROC: receiver operating characteristic.