

Co-expression Network Identification and Clinical Prognostic Evaluation of Hub Genes in Head and Neck Squamous Cell Carcinoma

Jialian Feng

Department of Otolaryngology, Head and Neck Surgery, Zhujiang Hospital

Baoai Han

Public Laboratory, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer

Chaosheng Yu

Department of Otolaryngology, Head and Neck Surgery, Zhujiang Hospital, Southern Medical University, Guangzhou, 510000, China

Congxiang Shen

Department of Otolaryngology, Head and Neck Surgery, Zhujiang Hospital, Southern Medical University, Guangzhou, 510000, China

Zhong Wen (✉ wenzhong60@163.com)

Southern Medical University Zhujiang Hospital <https://orcid.org/0000-0003-3591-3523>

Research

Keywords: Weighted gene co-expression network analysis, head and neck squamous cell carcinoma, prognosis, co-expression, The Cancer Genome Atlas

Posted Date: September 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-77378/v1>

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

[Read Full License](#)

Abstract

Background: According to statistics, even with active treatment, the recurrence rate of HNSCC is at 40%-50%. Head and neck cancer remains a challenge for otolaryngologists. Therefore, the identification of new biomarkers is an urgent need for the diagnosis, treatment, and prognosis of malignant tumors of the head and neck.

Methods: In this study, transcriptome data from The Cancer Genome Atlas (TCGA) database and the Gene Expression Omnibus (GEO) database were used to identify differentially expressed genes (DEGs). Weighted gene co-expression network analysis (WGCNA) was performed to identify gene modules and hub genes related to head and neck squamous cell carcinoma (HNSCC). Protei-Protei interaction(PPI) network and Cytoscape software were used to analyze the protein interaction network. HNSCC clinical data from the TCGA and Gene Expression Profile Interactive Analysis 2 databases were used to analyze the survival rate of hub genes, and the correlation between hub genes and tumor stage was further analyzed.

Results: A total of 2836 and 570 DEGs were identified from the TCGA expression data and GEO gene chip datasets, respectively. We found that the green module had the highest correlation with HNSCC. A total of 15 hub genes were also identified. In the Human Protein Atlas database, we found that thioredoxin reductase 1 (TXNRD1) was overexpressed in HNSCC tumors compared with normal tissues at the transcriptional level. Survival analysis also suggested that TXNRD1 was a poor prognostic factor for HNSCC.

Conclusion: Our results indicate that TXNRD1 is very likely to be identified as a potential biomarker and target for HNSCC. However, further research is required to fully reveal its role in HNSCC pathogenesis as well as its value as a prognostic biomarker.

Background

Malignant tumors from the mouth, nasopharynx, oropharynx, hypopharynx, and larynx are collectively referred to as head and neck cancer (HNC). Among them, head and neck squamous cell carcinoma (HNSCC) is the most common[1]. HNC is the sixth most common type of cancer worldwide, with 550,000 people diagnosed and about 380,000 dying from the disease annually[2].

Smoking and drinking are the main risk factors for the development of head and neck malignancies[3]. Moreover, the Epstein-Barr virus and human papilloma virus are related to nasopharyngeal carcinoma and oropharyngeal carcinoma, respectively[4, 5]. Furthermore, the choice of treatment options for HNSCC varies according to the anatomical location of the tumor, tumor stage, patient's age, health status, and preexisting comorbidities.

Surgery, including open and minimally invasive, is considered the standard treatment for most oropharyngeal cancers and early laryngeal cancers. Locally advanced HNSCC is treated with surgery and

definitive radiotherapy, usually accompanied by platinum-based chemotherapy[5].

Overexpression of the epidermal growth factor receptor in tumor tissue has led to the use of cetuximab, a first-line chemotherapy treatment for locally advanced HNSCC and metastatic/recurrent diseases[6]. Immunotherapy has also shown efficacy against HNSCC[7]. Both nivolumab and pembrolizumab have been approved for relapse/metastatic HNSCC in the second line[8], although the effective rate of anti-programmed death-1 monotherapy is still only at 15%[9, 10]. It is difficult to achieve accurate early detection, which may be the most important reason for the high mortality of patients with HNSCC. Therefore, there is an urgent need to develop an effective means of early detection, diagnosis, and treatment to improve the treatment of HNSCC.

In contrast to the limitations of traditional experimentation, the development of microarray and sequencing technology provides an excellent tool and platform for cancer research, with the application of big data bioinformatics rapidly developing[11–14]. The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases provide a large amount of relevant data for cancer research[15, 16]. Moreover, the R package weighted gene co-expression network analysis (WGCNA) can be used as a data exploration tool or gene screening (ranking) method to find clusters (modules) of highly correlated genes[17]. This algorithm has been widely used to find transcription level biomarkers[18–21]. For instance, SHI *et al.* analyzed the GEO data matrix to understand the pathogenesis of lung squamous cell carcinoma and revealed that CCNB1, CEP55, FOXM1, MKI67, and TYMS were potential biomarkers or therapeutic targets[22]. Zhou *et al.* used the GSE62452 and TCGA data sets for survival and regression analyses and identified 10 hub genes closely related to the progression of pancreatic cancer[23].

In the present study, we downloaded HNSCC clinical and gene expression data from TCGA database, including 31 normal tissue samples and 415 HNSCC tissue samples and gene expression profiles from the GEO database (GSE23036). R software was used to screen differentially expressed genes (DEGs) between HNSCC and normal tissue samples. After using WGCNA for gene module identification, intersection was performed to obtain 15 intersection genes. Subsequently, enrichment analysis, protein-protein interaction (PPI) network construction, and Cytoscape version 3.8.0 were used to illustrate significant correlations between DEGs. Finally, the identification, verification, and analysis of hub genes in DEGs revealed the prognostic and clinical value of the thioredoxin reductase 1 (TXNRD1) regulatory network involved in HNSCC.

Methods

Data collection and preprocessing

In this study, transcriptome data of HNSCCs, such as those of the oropharynx, larynx, nasopharynx, hypopharynx, tonsil, base of the tongue, and floor of the mouth, were downloaded from the TCGA database. These included 31 normal tissue samples and 415 HNSCC samples and a total of 14495 genes. Perl software version 5.30.1 was used for data preprocessing and conversion of gene IDs into the

gene names to obtain a new matrix for subsequent analysis. The clinical data of 440 patients with HNSCC were downloaded to further validate our results from the TCGA. At the same time, the GSE23036 dataset executed on the Affymetrix HG-U133A2.0 chip included 63 cases of HNSCC pretreated biopsy tissues and 5 cases of normal head and neck mucosal tissue. Expression data from the GSE23036 dataset and the TCGA database were used to identify hub genes.

Screening DEGs

The R software version 3.6.3(A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.) was used to analyze the selected data. The "edgeR" package was used to process transcriptome data. Then, "limma" package was used to screen DEGs in the transcriptome data from the TCGA database and the GSE23036 dataset. The cutoff standard was set as $|\log FC| < 1$ and the corrected P value was set at < 0.05 . With $geneNum = 50$ as condition, a heat map of differential genes was drawn using the "pheatmap" package. A volcano map was drawn using the "ggplot2" package according to differential gene expression (up-regulation, down-regulation, indifferent genes).

Co-expression network construction

WGCNA is an algorithm based on high-throughput gene expression profiles[24, 25]. It is most suitable for gene co-expression network analysis and is used for cancer identification, revealing correlation between genes and finding significantly related gene modules. The \log_2 (data = $\log_2[\text{data} + 1]$) value was taken from the TCGA and GEO data, and then the "WGCNA" and "limma" packages were used to construct a scale-free co-expression network for HNSCC. An appropriate power value was chosen to convert adjacent relationships into a topological overlap matrix (TOM). To classify genes with similar expression profiles into gene modules, an average linkage hierarchical clustering was performed on gene dendrograms based on the TOM-based difference measurement method. Then, a one-step function "block module" was used for network construction and module detection. BlockwiseModules were implemented with the following parameters: $maxBlockSize = 2000$, $minModuleSize = 50$ and $networkType = "unsigned."$ Finally, modules with the highest positive or negative correlations with normal tissues and HNSCC were identified for intersection, and a Venn diagram was drawn using the "VennDiagram" package.

Function and pathway enrichment analysis

To study a complete set of hub gene function annotations, the functional enrichment and interaction network analysis tool was used to analyze the Gene Ontology (GO) term enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The top 10 items with $P < 0.05$ and enrichment degree > 2 were selected, and the "Bioconductor-enrichplot" software package was used in R for enrichment analysis, with the output being a representative visualization.

PPI network construction

By searching the STRING(The STRING resource is available online at <https://string-db.org/>), a comprehensive score of ≥ 0.4 was defined as the critical point, and a correlation P value < 0.05 was used

as the output condition to obtain PPI network diagrams and tables for the next analysis. Cytoscape software was used to visualize the PPI network, CytoHubba was used to find the core genes, and the maximal clique centrality (MCC) method was used in constructing the gene network.

Relationship between the Hub gene and HNSCC prognosis

The Kaplan-Meier plotter is an online tool that can predict the prognostic value of cancer-related genes based on their expression levels. In this study, patients were divided into high and low expression groups according to the expression level of hub genes. The hazard ratio and log-rank P value of the 95% confidence interval were calculated, and hub genes with significant prognostic values were selected for further clinical comparison. Then, the HNSCC overall survival was calculated.

The Gene Expression Profile Interactive Analysis (GEPIA: <http://gepia.cancer-pku.cn/>) is an interactive web server for analyzing RNA sequencing expression data from TCGA and genotype organizations. Using data from the GEPIA database, the HNSCC disease-free survival (DFS) was analyzed. Moreover, we evaluated the correlation between selected hub gene expression and clinical TNM staging of HNSCC, with $P < 0.05$ indicating statistical significance. Finally, the Human Protein Atlas (HPA) database (link <https://www.proteinatlas.org/>) was used to further verify the expression of hub gene transcription and translation levels in HNSCC and normal tissues.

Results

Screening of DEGs

The workflow of our study is shown in Fig. 1. We comprehensively analyzed HNSCC expression and clinical data from the TCGA cohort and gene expression profile (GSE23036) from the GEO database. A total of 31 normal tissue samples, 415 tumor tissue samples, and 14495 genes from the TCGA database were included in this study. Among them, there were 2836 DEGs, 1265 upregulated genes in HNSCC, and 1571 downregulated genes in HNSCC ($\log_{2}FC < 1$, $P < 0.05$). GeneNum = 50 was set, and a differential gene heatmap was drawn, which intuitively reflected the distribution of gene differential expression (Fig. 2a). A volcano map was drawn according to gene expression (Fig. 2b). A total of 12548 genes in the GEO database were included in the study, and a total of 570 genes showed expression differences. Among them, 236 genes were upregulated, and 334 genes were downregulated in HNSCC ($\log_{2}FC < 1$, $P < 0.05$). Figure 2c presents the heatmap of these DEGs, and Figs. 2c-d illustrate the volcano maps.

Co-expression network

Expression values of the 14495 genes obtained from the TCGA database were used to construct the co-expression network, and DEGs with similar expression patterns were placed into the module (Fig. 3a). Twelve modules were screened out based on the modules and normal tissues and HNSCC histological grade. Figure 3b depicts the heat map of the modules as well as the traits data. Similarly, expression values of the 12548 genes from the gene expression profile (GSE23036) were used to construct a co-expression network, and DEGs with similar expression patterns were placed into the module to obtain the

co-expression network and a heat map of the module and trait data (Figs. 3c-d). Using the WGCNA results of the TCGA data, the green module with the most significant positive correlation with normal tissues was selected, and a green module from the GEO data was also selected. The two sets of DEGs were screened through the TCGA and GEO databases. Then, the intersection genes obtained from these four datasets were used to identify 15 disease-related candidate hub genes for further research (Fig. 4a).

Then, based on the STRING database, we applied critical values of reliability > 0.4 and connectivity > 10 (node/edge) to identify 24 genes as central nodes, and constructed a PPI network (Fig. 4b). The results were imported using Cytoscape for visualization (Fig. 4c), and the 15 hub genes were visualized using CytoHubba according to the MCC ranking method (Fig. 4d, Table 1).

Table 1 The result of cytoHubaa analysis of 15 core genes

Gene-name	MCC
KMO	721
AFMID	720
IDO1	720
IDO2	720
TD02	720
HAAO	720
KYNU	720
CFB	66
C3	66
CFI	55
CR1	48
CD46	30
CD55	24
TXN	19
TXNRD1	18

Functional enrichment analysis

To explore the enrichment pathway of disease-related genes, GO and KEGG function enrichment analyses were performed. KEGG results showed that the genes were significantly enriched in 8 pathways (Table 2), of which the main enrichment was in influenza A ($P < 0.05$, Figs. 5a-b). Results of GO functional analysis showed that these 15 disease-related genes were enriched in 30 functional pathways (Table 3), among which genes were significantly enriched in three main functions: sulfur compound metabolic process, dendrite terminus, and tetrapyrrole binding (Figs. 5c-d). The corrected P values of all genes were < 0.05 , indicating statistical significance.

Table 2
GO results

ONTOLOGY	Description	pvalue	gene list
BP	cellular biogenic amine metabolic process	0.000872965	TD02/MOXD1
BP	sulfur compound metabolic process	0.002987745	MSRA/ELOVL6/TD02
BP	acyl-CoA metabolic process	0.00313613	ELOVL6/TD02
BP	thioester metabolic process	0.00313613	ELOVL6/TD02
BP	long-chain fatty acid metabolic process	0.003374555	CYP2J2/ELOVL6
BP	unsaturated fatty acid metabolic process	0.003435455	CYP2J2/ELOVL6
BP	cellular amine metabolic process	0.00468966	TD02/MOXD1
BP	amine metabolic process	0.005347254	TD02/MOXD1
BP	nucleoside bisphosphate metabolic process	0.005498848	ELOVL6/TD02
BP	ribonucleoside bisphosphate metabolic process	0.005498848	ELOVL6/TD02
CC	dendrite terminus	0.009194177	COBL
CC	GABA-A receptor complex	0.013411113	GABRP
CC	GABA receptor complex	0.014112313	GABRP
CC	axonal growth cone	0.019705279	COBL
CC	apical plasma membrane	0.020762972	SORBS2/MAL
CC	secretory granule lumen	0.02113157	CFD/TCN1
CC	cytoplasmic vesicle lumen	0.02327239	CFD/TCN1
CC	vesicle lumen	0.023401052	CFD/TCN1
CC	apical part of cell	0.029495694	SORBS2/MAL
CC	cell leading edge	0.032242597	SORBS2/COBL
MF	tetrapyrrole binding	0.00018355	CYP2J2/TD02/TCN1
MF	monooxygenase activity	0.002698307	CYP2J2/MOXD1
MF	heme binding	0.004950061	CYP2J2/TD02
MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	0.006800222	CYP2J2/MOXD1
MF	serine-type endopeptidase activity	0.006883195	CFD/TMPRSS2

ONTOLOGY	Description	pvalue	gene list
MF	structural constituent of myelin sheath	0.00788484	MAL
MF	cobalamin binding	0.00788484	TCN1
MF	fatty acid synthase activity	0.008670141	ELOVL6
MF	serine-type peptidase activity	0.008824974	CFD/TMPRSS2
MF	serine hydrolase activity	0.00920163	CFD/TMPRSS2

Table 3
KEGG results

Pathway	pvalue	geneID
Influenza A	0.008812016	TMPRSS2/IL33
Fatty acid elongation	0.023292031	ELOVL6
Biosynthesis of unsaturated fatty acids	0.023292031	ELOVL6
Linoleic acid metabolism	0.02499871	CYP2J2
Nicotine addiction	0.034339825	GABRP
Tryptophan metabolism	0.036029941	TD02
Ovarian steroidogenesis	0.043604125	CYP2J2
Fatty acid metabolism	0.048625193	ELOVL6

Identification of prognostic biomarkers

After identifying hub genes, we analyzed 15 genes using Kaplan-Meier analysis. Only the TXNRD1 results of Kaplan-Meier analysis were statistically significant ($P < 0.05$) (Fig. 6a), and the results showed that TXNRD1 is an oncogene. For further verification, we searched the GEPIA2 database. According to gene expression, patients were divided into high and low expression groups, and the HNSCC DFS was analyzed. Only the TXN and KMO results were statistically significant ($P < 0.05$) (Fig. 6b-c). Finally, we studied the link between tumor staging and hub genes (Fig. 7). Although no significant difference in TXN was detected, the gene expression level showed an upward trend consistent with increasing tumor stages. Finally, for in-depth analysis, the HPA database was searched for verification. TXNRD1 was not expressed in normal tissues, whereas it was moderately expressed in HNCSS, which suggested that TXNRD1 gene overexpression in HNSCC may be related to poor prognosis of the disease. In addition, hub genes, including TXN, were also verified, and details are shown in Fig. 8.

Discussion

Head and neck malignancies are the sixth most common cancers in the world, accounting for 21.2% of systemic malignant tumors, 47% of male malignant tumors, and 2.5% of female malignant tumors[1]. Moreover, squamous cell carcinoma is the most common histological type (97%) among these malignancies. The patients' pronunciation, vision, hearing, breathing, and social communication become deeply affected by these diseases, with their quality of life becoming greatly reduced in addition to a heavy social burden[26–28]. In recent years, many diagnostic molecular markers related to HNSCC have been identified[29–32], but it is difficult to achieve accurate early detection, which may be the most important reason for the high mortality of patients with HNSCC. Therefore, there is an urgent need to develop an effective means of early detection, diagnosis, and treatment to improve the treatment of HNSCC. In this study, a gene expression matrix was used to construct the co-expression network and PPI network. A total of 15 hub genes were identified using CytoHubba, and were analyzed for functional enrichment. The correlation between hub genes and HNSCC prognosis was also evaluated. Three genes (TXNRD1, TXN, and KMO) were identified to be related to poor prognosis of HNSCC, and TXNRD1 was identified as a potential prognostic biomarker among 15 other genes. Kaplan-Meier analysis results and HPA database search results indicated that TXNRD1 is an oncogene of HNSCC. Based on these findings, TXNRD1 was selected as the best potential biomarker for detecting and predicting the prognosis of HNSCC.

Abnormal activation of cell signaling pathways and metabolism drives an increase in oxidative stress in cancer cells. Due to an imbalance in the production and removal of reactive oxygen species (ROS), cells suffer from oxidative stress[33, 34]. In the process of oxygen metabolism, molecules containing chemically active oxygen are generated. The two main types of ROS include free radicals (such as superoxide or hydroxyl) and non-radicals. As a signaling molecule, ROS can activate carcinogenic pathways. The thioredoxin system is the main category of mammalian antioxidant systems[35]. It comprises thioredoxin, thioredoxin reductase, and nicotinamide adenine dinucleotide phosphate and reduced (NADPH)[36, 37]. In contrast, TXNRD is an NADPH-dependent dimeric selenase-containing flavin adenine dinucleotide domain. It belongs to the nucleotide-disulfide oxidoreductase family and is divided into TXNRD1, TXNRD2, and TXNRD3. The mediation between NADPH and TXN in the thioredoxin system is maintained only by the thioredoxin-specific enzyme TXNRD. TXNRD is a selenium-dependent enzyme and an NADPH-dependent flavoprotein. It can reduce oxidized TXN to the dithiol form to maintain cell redox homeostasis. The thioredoxin system is activated in cancer, and TXNRD1 is overexpressed in many solid cancers, such as hepatocellular carcinoma[38, 39], gastric cancer[40], colorectal cancer[41], and prostate cancer[42]. A research team revealed that TXNRD1 inhibition can be used as an anti-cancer treatment[38]. Inhibiting TXNRD1 has also been suggested to have a damaging effect on normal tissues. Stafford *et al.* proposed that specific TXNRD1 inhibitors impaired the growth and survival of tumor xenografts and syngeneic mouse tumors[43]. At the same time, specific TXNRD1 inhibitors have little toxicity to the mitochondria, and specific TXNRD1 inhibitors are promising as new anti-tumor drugs.

In summary, our study was the first attempt to use the WGCNA algorithm to identify the expression of hub genes in HNSCC as well as perform a related prognostic analysis. We identified a total of 15 hub genes, and through multiple verifications using bioinformatics, TXNRD1 was identified as a potential biomarker and target for HNSCC. However, this study has certain limitations. More sample data are needed to verify these results, and the specific mechanism of TXNRD1 in HNSCC needs to be further investigated. Moreover, there was a lack of detailed clinical information (such as staging, lymph node metastasis, gender, age) for HNSCC patients involved. Therefore, our future research will be based on the mechanism of action of hub genes and their potential functions.

Conclusions

We identified DEGs and meaningful gene modules in HNSCC. Valuable hub genes were screened out (TXNRD1, TXN, and KMO) in HNSCC tissues. Survival and clinical correlation analyses and HPA data indicated that TXNRD1 is related to the poor prognosis of HNSCC. However, further research is required to fully reveal the role of these genes in the pathogenesis of HNSCC and to verify their value as prognostic biomarkers.

Declarations

Ethics approval and consent to participate

Consent for publication

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by grants from the National Natural Science Foundation of China Youth Fund (81402456) and the Science and Natural Science Foundation of Guangdong province, China (2015A030313255).

Authors' contributions

The contribution list of all authors of the manuscript is as follows:

Conception and design: Jialian Feng, Baoai Han

Development of methodology: Congxiang Shen, Zhong Wen

Analysis and interpretation of data: Jialian Feng, Chaosheng Yu

Writing, review, and/or revision of the manuscript: Jialian Feng, Baoai Han

Study supervision: Congxiang Shen, Zhong Wen

Acknowledgements

Not applicable

References

1. Alam MS, Siddiqui SA, Perween R. Epidemiological profile of head and neck cancer patients in Western Uttar Pradesh and analysis of distributions of risk factors in relation to site of tumor. *J Cancer Res Ther.* 2017;13(3):430–5.
2. Fitzmaurice C, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015. *JAMA Oncology.* 2017;3(4):524.
3. Jethwa AR, Khariwala SS. Tobacco-related carcinogenesis in head and neck cancer. *Cancer Metastasis Rev.* 2017;36(3):411–23.
4. Jamieson SMF, et al., Evofosfamide for the treatment of human papillomavirus-negative head and neck squamous cell carcinoma. *JCI insight*, 2018. 3(16).
5. Na'Ara S, et al. Outcome of patients undergoing salvage surgery for recurrent nasopharyngeal carcinoma: a meta-analysis. *Ann Surg Oncol.* 2014;21(9):3056–62.
6. Tahara M, et al. Phase II trial of combination treatment with paclitaxel, carboplatin and cetuximab (PCE) as first-line treatment in patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck (CSPOR-HN02). *Ann Oncol.* 2018;29(4):1004–9.
7. Moskovitz J, Moy J, Ferris RL. Immunotherapy for Head and Neck Squamous Cell Carcinoma. *Current Oncology Reports*, 2018. 20(2).
8. Economopoulou P, Kotsantis I, Psyrri A. The promise of immunotherapy in head and neck squamous cell carcinoma: combinatorial immunotherapy approaches. *ESMO Open.* 2016;1(6):e000122.
9. Gildener-Leapman N, Ferris RL, Bauman JE. Promising systemic immunotherapies in head and neck squamous cell carcinoma. *Oral Oncol.* 2013;49(12):1089–96.
10. Rasmussen J, et al. PO-106 Intratumor heterogeneity of PD-L1 expression in Head and Neck squamous cell carcinoma. *Radiother Oncol.* 2019;132:53–4.
11. Wlodarczyk A, et al. A novel microglial subset plays a key role in myelinogenesis in developing brain. *EMBO J.* 2017;36(22):3292–308.

12. Liu J, et al., Establishment of a novel cell cycle-related prognostic signature predicting prognosis in patients with endometrial cancer. *Cancer Cell International*, 2020. 20(1).
13. Han B, et al. Subtype-specific risk models for accurately predicting the prognosis of breast cancer using differentially expressed autophagy-related genes. *Aging (Albany NY)*. 2020;12(13):13318–37.
14. Wang S, et al. Prognostic and predictive value of immune/stromal-related gene biomarkers in renal cell carcinoma. *Oncol Lett*. 2020;20(1):308–16.
15. Jia D, et al. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging (Albany NY)*. 2018;10(4):592–605.
16. Lee H, et al. A pan-cancer analysis of the clinical and genetic portraits of somatostatin receptor expressing tumor as a potential target of peptide receptor imaging and therapy. *EJNMMI Res*. 2020;10(1):42.
17. Ai D, et al. Colorectal Cancer Prediction Based on Weighted Gene Co-Expression Network Analysis and Variational Auto-Encoder. *Biomolecules*. 2020;10(9):1207.
18. Ding M, et al. A comprehensive analysis of WGCNA and serum metabolomics manifests the lung cancer-associated disordered glucose metabolism. *J Cell Biochem*. 2019;120(6):10855–63.
19. Liu Z, et al. Identification of surrogate prognostic biomarkers for allergic asthma in nasal epithelial brushing samples by WGCNA. *J Cell Biochem*. 2019;120(4):5137–50.
20. Feltrin AS, et al. Assessment of complementarity of WGCNA and NER1 results for identification of modules associated to schizophrenia spectrum disorders. *PLoS One*. 2019;14(1):e0210431.
21. Qin D, et al., A Circulating miRNA-Based Scoring System Established by WGCNA to Predict Colon Cancer. *Anal Cell Pathol (Amst)*, 2019. 2019: p. 1571045.
22. Shi Y, et al. Identification of key genes and evaluation of clinical outcomes in lung squamous cell carcinoma using integrated bioinformatics analysis. *Oncology letters*. 2019;18(6):5859–70.
23. Zhou Z, et al. Ten hub genes associated with progression and prognosis of pancreatic carcinoma identified by co-expression analysis. *International Journal of Biological Sciences*. 2018;14(2):124–36.
24. Qian J, et al. Analysis of lncRNA-mRNA networks after MEK1/2 inhibition based on WGCNA in pancreatic ductal adenocarcinoma. *J Cell Physiol*. 2020;235(4):3657–68.
25. Chen S, et al. Identification of crucial genes in abdominal aortic aneurysm by WGCNA. *PeerJ*. 2019;7:e7873.
26. Polverini PJ, D'Silva NJ, Lei YL. Precision Therapy of Head and Neck Squamous Cell Carcinoma. *J Dent Res*. 2018;97(6):614–21.
27. Moskovitz JM, Ferris RL. Tumor Immunology and Immunotherapy for Head and Neck Squamous Cell Carcinoma. *J Dent Res*. 2018;97(6):622–6.
28. Solomon I, et al., Neuroendocrine Factors and Head and Neck Squamous Cell Carcinoma: An Affair to Remember. *Dis Markers*, 2018. 2018: p. 9787831.

29. Huang C, et al. Development and validation of radiomic signatures of head and neck squamous cell carcinoma molecular features and subtypes. *EBioMedicine*. 2019;45:70–80.
30. Wilkins OM, et al. MicroRNA-Related Genetic Variants Associated with Survival of Head and Neck Squamous Cell Carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2019;28(1):127–36.
31. Kanazawa T, et al. Promoter methylation of galanin receptors as epigenetic biomarkers for head and neck squamous cell carcinomas. *Expert Rev Mol Diagn*. 2019;19(2):137–48.
32. Wei GG, et al. Drug repositioning in head and neck squamous cell carcinoma: An integrated pathway analysis based on connectivity map and differential gene expression. *Pathol Res Pract*. 2019;215(6):152378.
33. Yang B, Chen Y, Shi J. Reactive Oxygen Species (ROS)-Based Nanomedicine. *Chem Rev*. 2019;119(8):4881–985.
34. Freinbichler W, et al. Highly reactive oxygen species: detection, formation, and possible functions. *Cell Mol Life Sci*. 2011;68(12):2067–79.
35. McLoughlin MR, et al., TrxR1, Gsr, and oxidative stress determine hepatocellular carcinoma malignancy. *Proceedings of the National Academy of Sciences*, 2019. 116(23): p. 11408–11417.
36. Roshanzadeh A, et al. Real-time monitoring of NADPH levels in living mammalian cells using fluorescence-enhancing protein bound to NADPHs. *Biosens Bioelectron*. 2019;146:111753.
37. Iyanagi T. Molecular mechanism of metabolic NAD(P)H-dependent electron-transfer systems: The role of redox cofactors. *Biochim Biophys Acta Bioenerg*. 2019;1860(3):233–58.
38. Lee D, et al. Induction of Oxidative Stress Through Inhibition of Thioredoxin Reductase 1 Is an Effective Therapeutic Approach for Hepatocellular Carcinoma. *Hepatology*. 2019;69(4):1768–86.
39. Tuo L, et al., PCK1 Downregulation Promotes TXNRD1 Expression and Hepatoma Cell Growth via the Nrf2/Keap1 Pathway. *Frontiers in Oncology*, 2018. 8.
40. Lim JY. Thioredoxin and thioredoxin-interacting protein as prognostic markers for gastric cancer recurrence. *World J Gastroenterol*. 2012;18(39):5581.
41. Hughes DJ, et al., Expression of Selenoprotein Genes and Association with Selenium Status in Colorectal Adenoma and Colorectal Cancer. *Nutrients*, 2018. 10(11).
42. Gerstenberger JP, et al. Selenoprotein and antioxidant genes and the risk of high-grade prostate cancer and prostate cancer recurrence. *Prostate*. 2015;75(1):60–9.
43. Stafford WC, et al. Irreversible inhibition of cytosolic thioredoxin reductase 1 as a mechanistic basis for anticancer therapy. *Sci Transl Med*. 2018;10(428):eaaf7444.

Figures

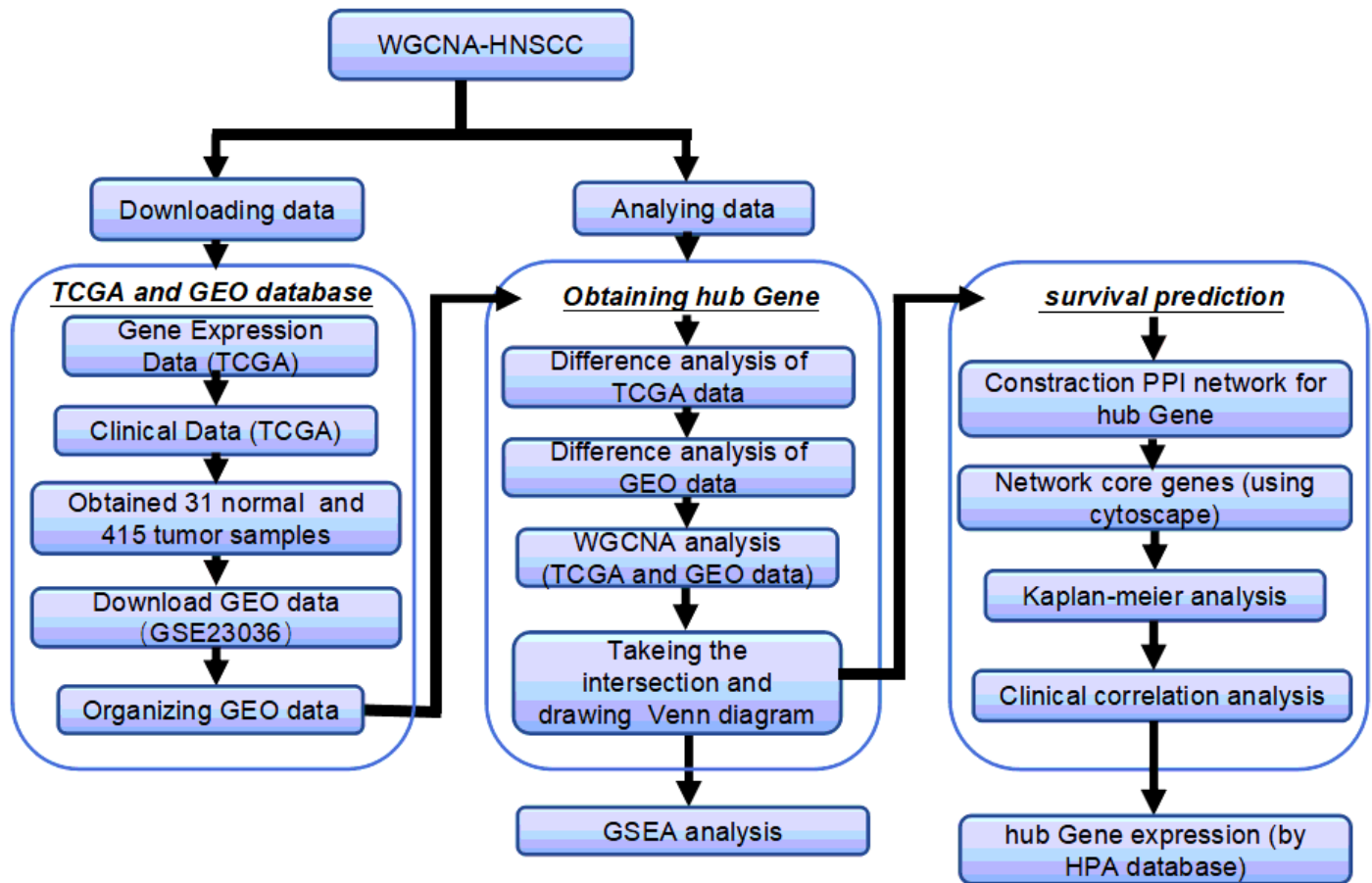


Figure 1

Flow chart of this study.

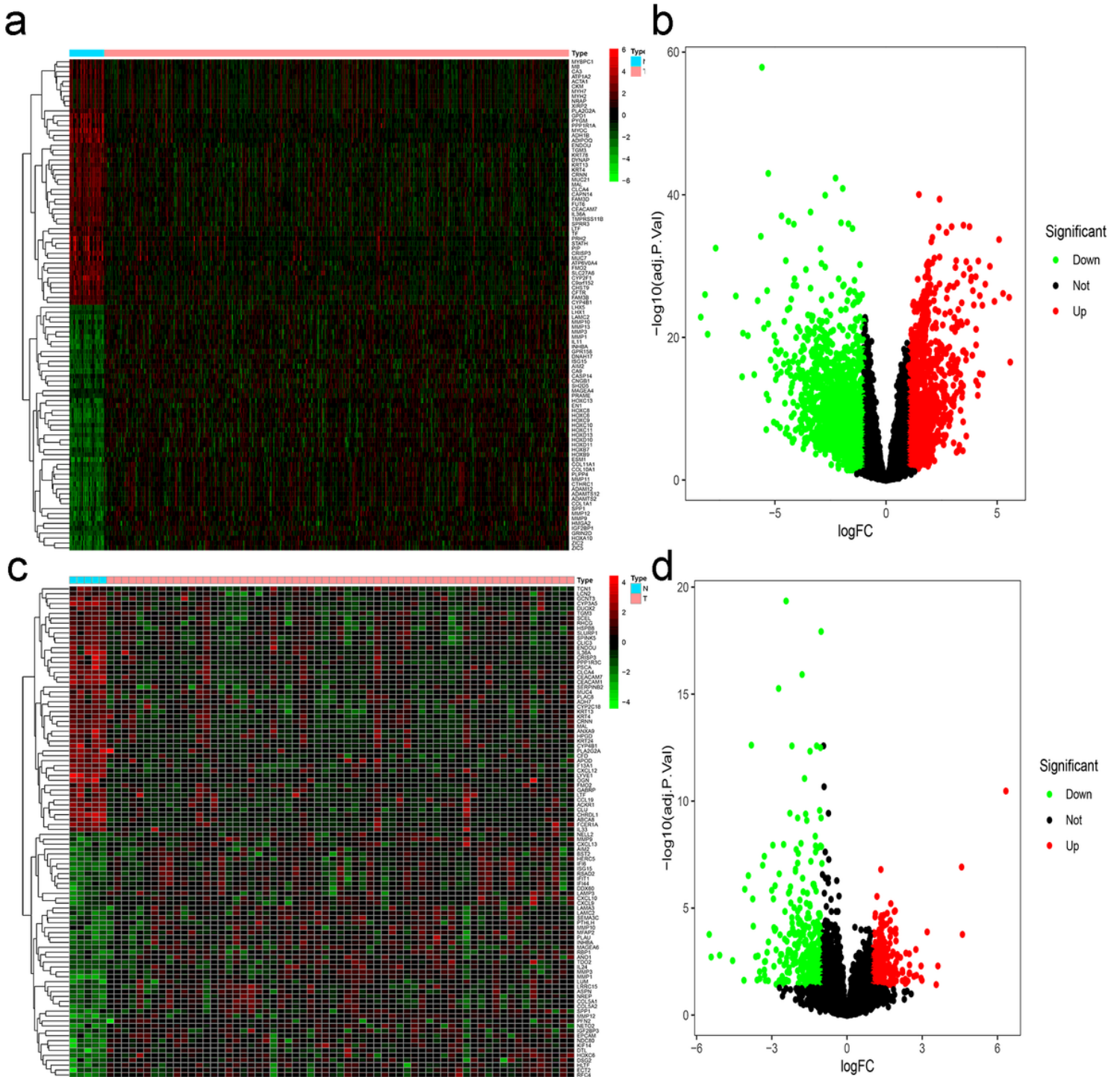


Figure 2

Identification of differentially expressed genes. Heat-map and volcano-map of differentially expressed genes in head and neck squamous cell carcinoma (data from TCGA database (a)-(b), data from GEO database (c)-(d))

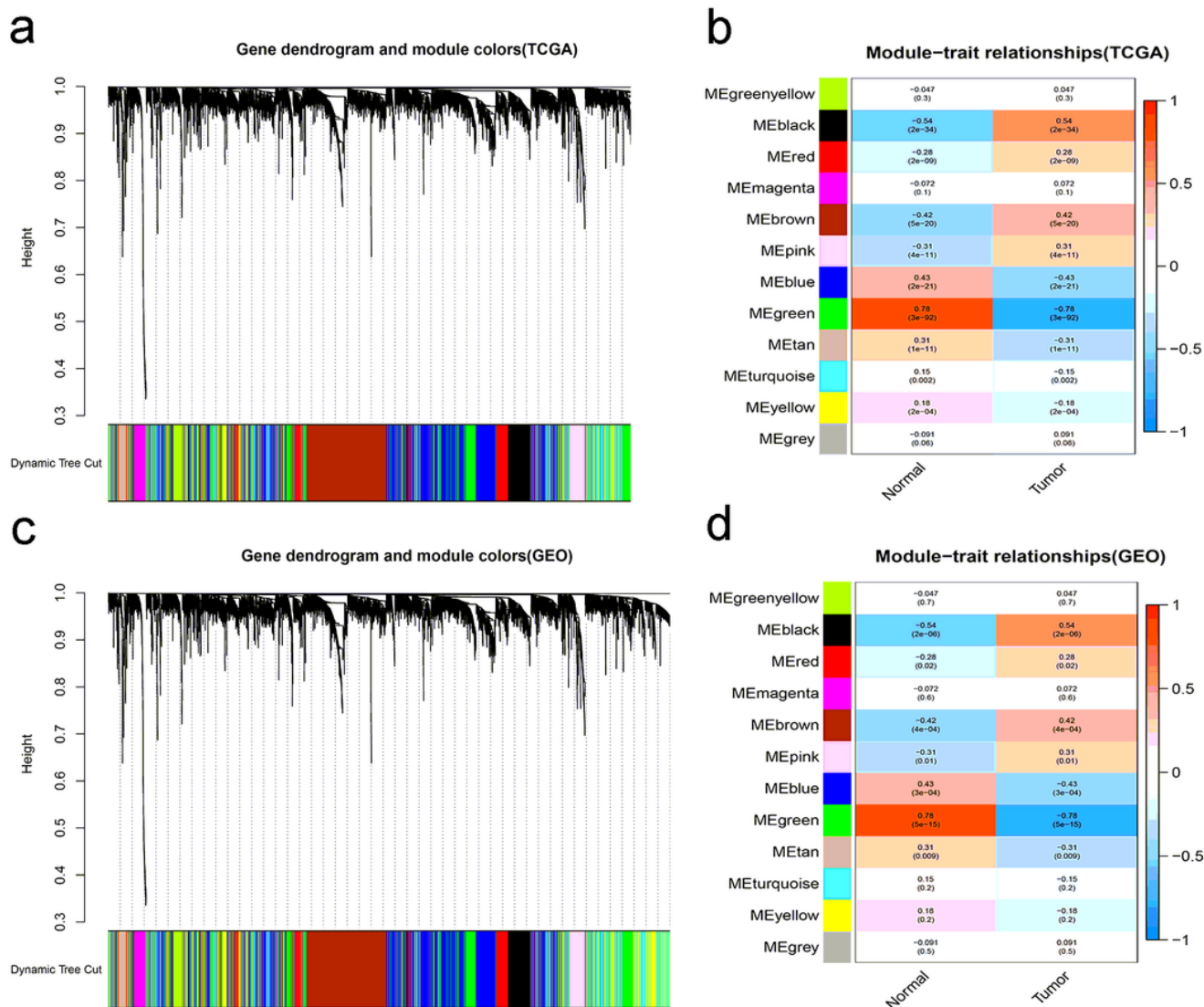


Figure 3

Weighted gene co-expression network analysis. (a),(c) Genes are divided into different modules through hierarchical clustering, and different colors represent different modules. (b),(d) Heat map shows the correlation between the intrinsic genes of the module and normal tissues and HNSCC.

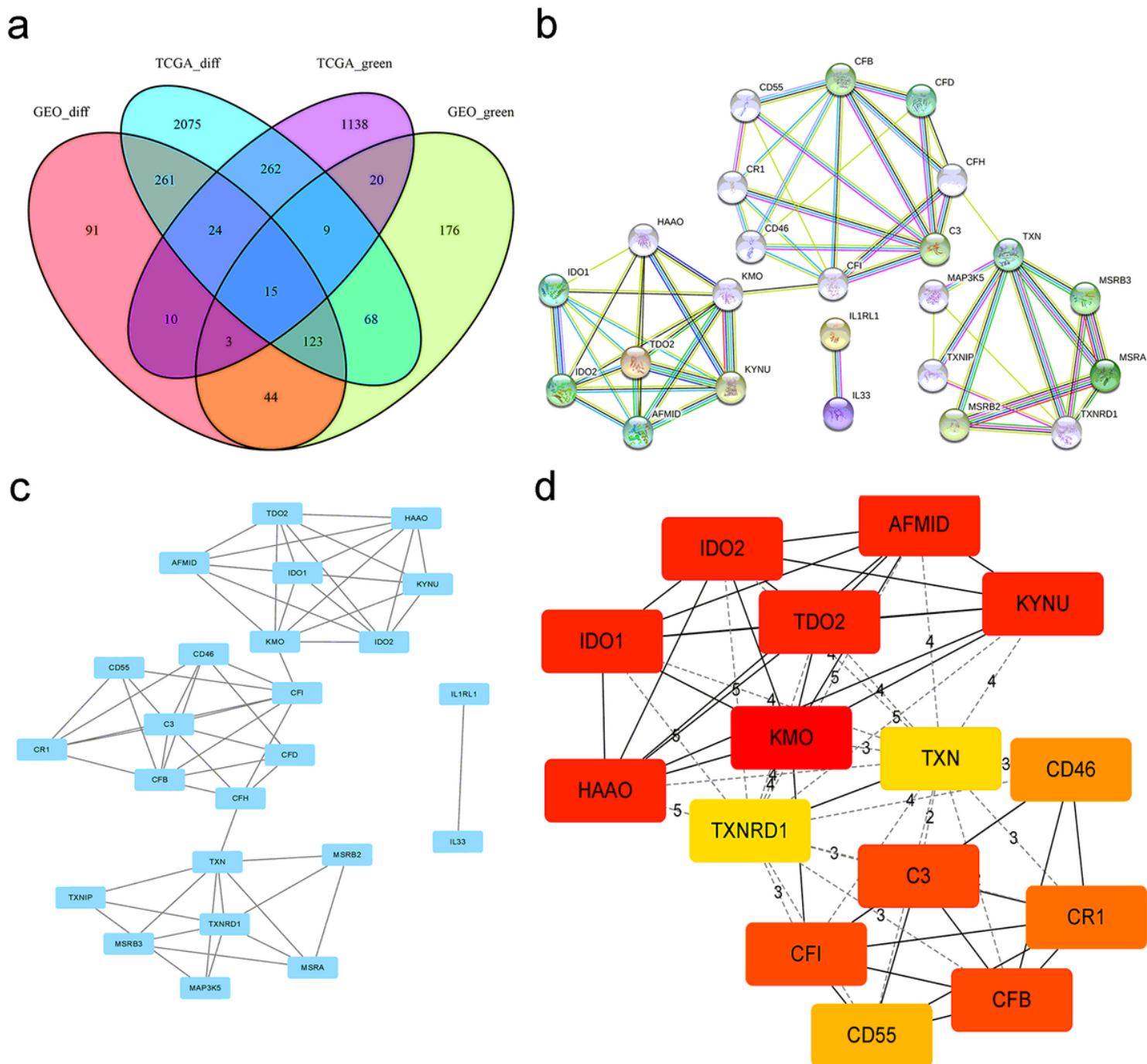


Figure 4

Identification of hub genes and protein-protein networks(PPI). (a) Identification of HNSCC disease-related genes. (b) Protein-protein interaction network of 15 disease-related genes. (c) Visual analysis of protein-protein interaction network by Cytoscape. (d) Identification of real hub genes.

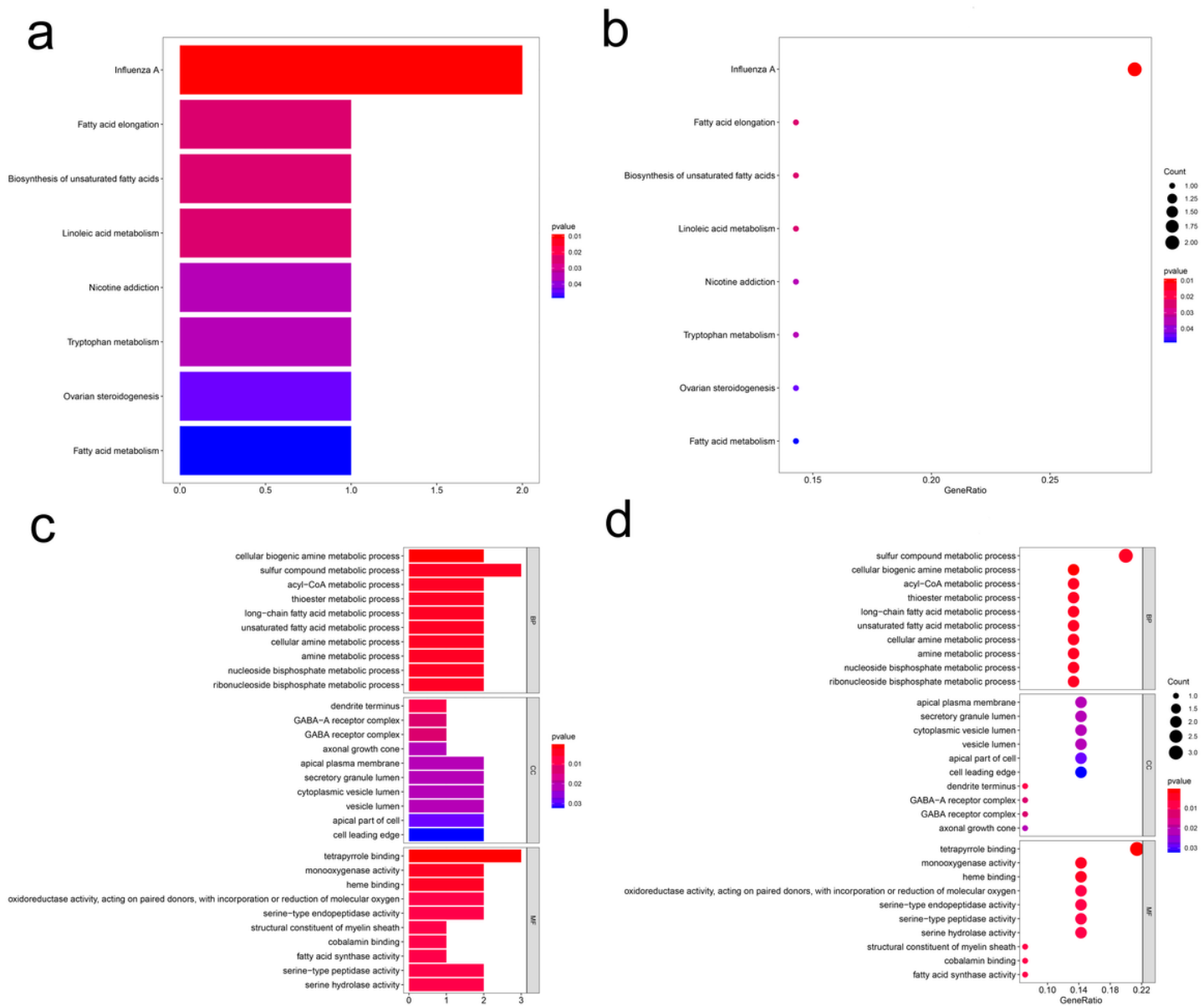


Figure 5

Gene set enrichment analysis of disease-related genes. (a) , (b) Gene Ontology analysis of disease-related genes. (c) (d) KEGG pathway enrichment analysis for disease-related genes. The x-axis shows the P-value of each term and the y-axis shows the GO and KEGG pathway terms.

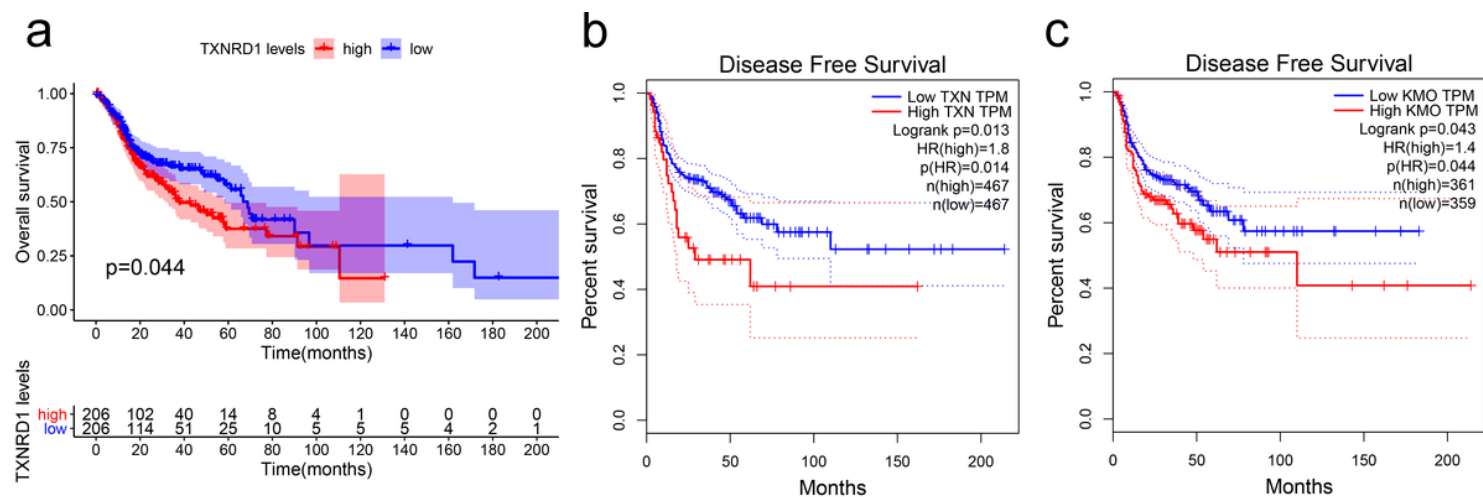


Figure 6

Survival analysis of TXNRD1(a), TXN(b) and KMO(c). Red lines represent high expression of the real hub genes and blue lines represent low expression.

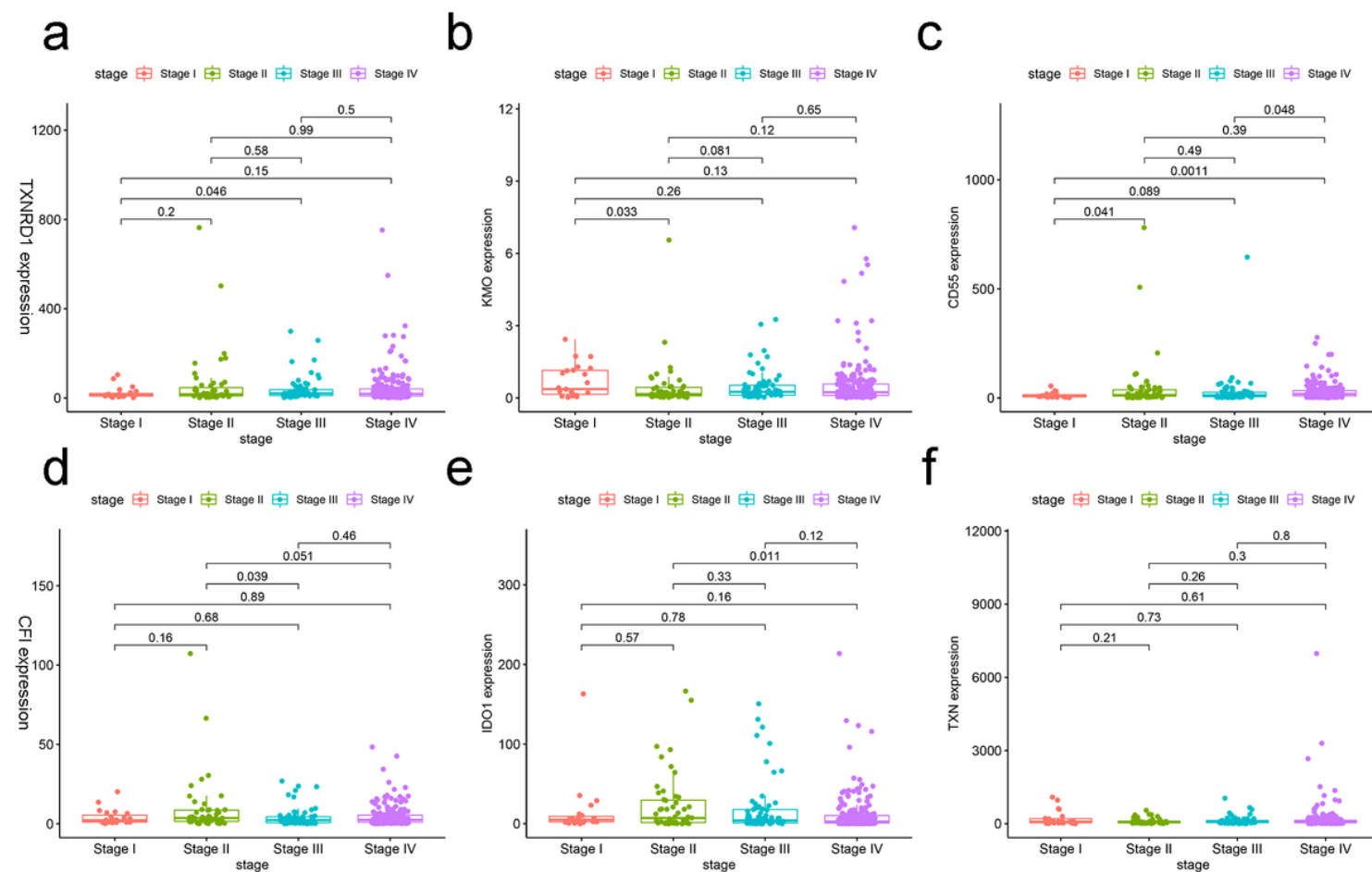


Figure 7

Correlation analysis of hub gene and clinical stage. TXNRD1(a), KMO(b), CD55(c), CFI(d), IDO1(e), TXN(f)

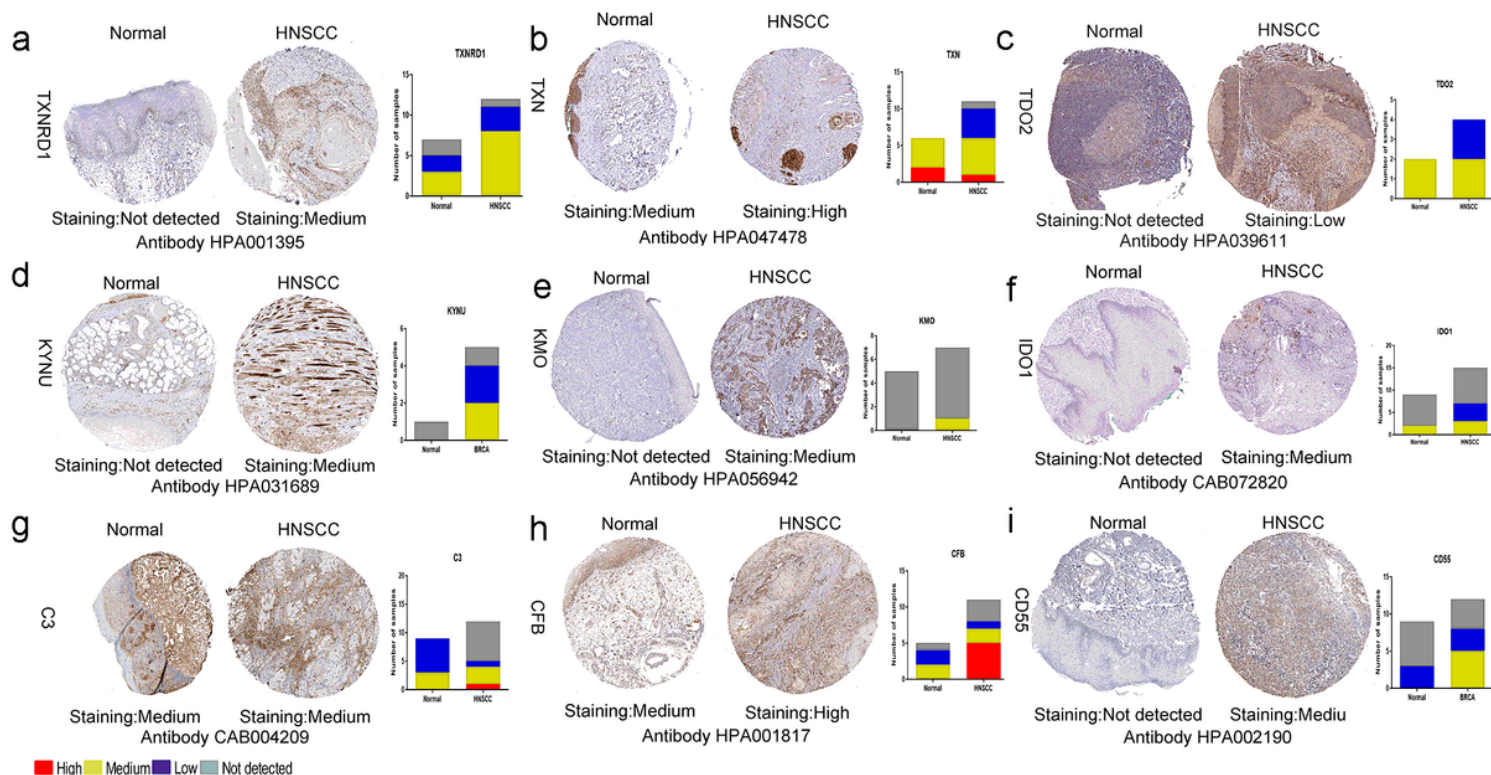


Figure 8

Validation the expression of real hub genes on transcriptional and translational level by The Human Protein Atlas database(HPA). TXNRD1(a),TXN(b),TDO2(c),KYNU(d),KMO(e),IDO1(f),C3(g),CFB(h),CD55(i).