Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Potential Role of The Upregulation of The Zinc Finger Protein of The Cerebellum 2 In Nasopharyngeal Carcinoma

Yun-Xi Peng
Guangxi Medical University First Affiliated Hospital
Jian Li
Hechi People's Hospital
Hua Ding
Guangxi Medical University First Affiliated Hospital
Ye-Ying Fang
Guangxi Medical University First Affiliated Hospital
Wen Qin
Guangxi Medical University First Affiliated Hospital
Jin-Lian Liao
Guangxi Medical University First Affiliated Hospital
Li Jiang
Guangxi Medical University First Affiliated Hospital
Miao Mo
Guangxi Medical University First Affiliated Hospital
Cai-Xing Lin
Guangxi Medical University First Affiliated Hospital
Ji-Yun Wu
Guangxi Medical University First Affiliated Hospital
Su-Ning Huang
Guangxi Cancer Hospital and Guangxi Medical University Affiliated Cancer Hospital
Zhu Xin Wei (weizhuxin@stu.gxmu.edu.cn)
Guangxi Medical University First Affiliated Hospital

Research

Keywords: zinc finger protein of the cerebellum 2, nasopharyngeal carcinoma, immunohistochemical staining, tissue microarray, weighted gene co-expression network analysis

Posted Date: August 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-762984/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: Nasopharyngeal carcinoma (NPC) is a common type of cancer mainly caused by EB virus infection in southern China and Southeast Asia. Zinc finger protein of the cerebellum 2 has been reported to be dysregulated in numerous cancers. However, its role in NPC has not yet been completely investigated.

Methods: ZIC2 mRNA expression and clinical role of NPC was explored by Gene Expression Omnibus (GEO) databases and tissue microarrays. Immunohistochemical (IHC) staining was used to assessed the protein expression level of ZIC2 in NPC. Moreover, functional enrichment analyses were performed by the overlapping genes of differentially expressed genes in NPC and ZIC2 related co-expressed genes. Cistrome Data Browser (Cistrome DB) and scatter diagrams displayed the possibility of ATAD2 as a target gene of ZIC2. Last, characteristic phenotype was identified by weighted gene co-expression network analysis (WGCNA).

Results: The mRNA expression levels of ZIC2 were significantly increased in NPC within genechips and tissue microarrays (SMD = 2.14; 95% CI = 1.85–2.44). Increased ZIC2 correlated with poor prognosis in NPC patients. The IHC results also showed that ZIC2 protein expression in NPC tissues was remarkably higher than in control tissues. Enrichment analysis showed that genes positively related to ZIC2 are mainly enriched in cell division and in the cell cycle. WGCNA analysis discovered that NPC dataset GSE12452 may be positive correlated with N stage and nagetive correlated with grade.

Conclusion: ZIC2 upregulation may promote NPC tumorigenesis through cancer cell proliferation and cell cycle regulation. ATAD2 may produce a marked effect as the downstream target of ZIC2.

Background

Nasopharyngeal carcinoma (NPC) is a highly invasive and metastatic malignant tumor that originates in the nasopharynx, which is at the back of the nose. According to global cancer statistics, there were 133,354 new NPC cases and 80,008 deaths in 2020 (1). In most parts of the world, the age-standardized incidence rate (ASIR) of NPC is lower than 1.0 per 100,000. However, the onset of NPC significantly varies across global geographic regions; and in Southeast Asia and southern China, the ASIR of NPC is higher than 7.0 per 100,000. Epstein-Barr (EB) virus infection, hereditary factors, and environmental factors appear to contribute to NPC (2). The hidden location and complex anatomical structure of the nasopharynx, coupled with the few early symptoms of the tumor, can easily cause its delayed diagnosis and treatment. In recent years, although radiotherapy technology has significantly improved, the overall treatment efficacy of NPC is still not ideal, and its distant metastasis is as high as 14.1% (3). Therefore, exploring molecular targets related to NPC is necessary to understand the pathogenesis of NPC, to provide a basis for its truly individualized treatment, and to enhance the effect of such treatment.

The zinc finger protein of the cerebellum 2 (ZIC2), which is characterized by a highly conserved C2H2 zinc fingers motif and, when mutated, causes the common and severe malformation *holoprosencephaly* (HPE) (4, 5), functions as a transcriptional regulatory factor that can interact with multiple DNA and proteins (6–8). Previous studies documented that ZIC2 overexpression enhanced the invasion (9, 10), cellular proliferation (11, 12), metastasis (13–15), angiogenesis (10, 16), and cancer stem cell traits (17) in many solid tumors. However, the mechanism of ZIC2 in NPC has seldom been revealed in reports. Jiang et al. (18) and Yi et al.

(19) discovered the overexpression of ZIC2 in NPC through bioinformatics and immunohistochemistry, respectively. Yu et al. (20) clarified that miRNA-129-5p suppresses NPC by targeting ZIC2, and Lv et al. (21) demonstrated that tumor-suppressive miRNA-873 acted as an upstream factor of the downregulation of ZIC2 in NPC. Moreover, Shen et al. (22) explored the regulative effect of HOXA10 on ZIC2 expression in NPC. Although these studies were conducted to investigate the role of ZIC2 in NPC, the mechanisms of such role are not yet precise. In this study, we used bioinformatics analysis to identify new target markers of ZIC2. We also included more samples for immunohistochemistry to enhance our understanding of NPC carcinogenesis and progression.

Materials And Methods

High-throughput data mining and analysis

We obtained microarray datasets, which could be used to compare the ZIC2 expression profile data between normal nasopharynx tissues and NPC tissues, from the GEO datasets ArrayExpress, Oncomine, and SRA. We selected the included microarray datasets according to the following three criteria: (1) All the specimens came from humans; (2) Each gene chip contained the transcription profile of the messenger RNA (mRNA) expression; and (3) Both the NPC and the non-cancerous nasopharynx tissues were provided.

We performed the background correction, normalization, and log2 transformation of the preceding data for processing data with the “limma” package in the R 4.0.3 software environment in order to eliminate the influence of the sequencing depth and the gene length on the expression level. Moreover, we used the NPC-related microarrays for differentially expressed genes (DEGs) analysis through nonspecific filtration via the empirical Bayesian method, in order to facilitate the borrowing of information (|log2 Fold Change| > 1 and adj. $P < 0.05$). From the upregulated or downregulated DEGs that we collected, we chose those that appeared in two or more independent datasets as the DEGs in this study. Similarly, we identified genes with a ZIC2 |correlation coefficient|>0.3, a P-value < 0.05, and a repeat number larger than or equal to 2 as significant co-expressed genes (CEGs). We obtained the overlapping genes of the DLX2 positively correlated CEGs and upregulated DEGs (positive gene set), of the DLX2 negatively related CEGs, and of the downregulated DEGs (negative gene set) for our follow-up research.

In-house immunohistochemical evaluation

The tissue microarrays (TMAs) that contained 154 NPC and 49 chronic nasopharyngeal mucositis tissue specimens (NPC131, NPC241, and NPC482) were provided by Pantomics, Inc. (Richmond, CA). At the same time, the NPC tissues and mucosal inflammation tissues that we derived from 12 patients had been diagnosed with nasopharyngeal carcinoma, and we collected 14 cases of chronic nasopharyngitis from the First Affiliated Hospital of Guangxi Medical University. We assayed the protein level of ZIC2 in the tissues via immunohistochemistry (IHC). First, we prepared the paraffin-embedded tissue specimens into 4μm-thick tissue sections, and then we dewaxed and hydrated them. We boiled these slides for 3 minutes in a pressure cooker to recover the antigen. Then we incubated the samples in a diluted rabbit polyclonal ZIC2 antibody (Biorbyt) at room temperature for 1 hour, and the secondary antibody at 37°C for 1 hour. We performed the rest of the measure operation steps according to the instructions. To ensure the accuracy of the experiment, two

pathologists separately determined the final results of the IHC staining. We evaluated the final regional differences in the staining through the immunoreactivity scores (IRSs), which ranged from 0 to 12 were calculated by multiplying two parameters: the staining intensity and the proportion of the stained tumor cells. We evaluated the staining intensity as follows: 0 for no staining, 1 for weak staining, 2 for moderate staining, and 3 for strong staining. We defined the frequency of the stained cells as follows: 0 for < 10% positive cells; 1 for 10–25% positive cells; 2 for 26–50% positive cells; 3 for 51–75% positive cells; and 4 for more than 75% positive cells. After calculating the IRS value, we assessed the relationship between the ZIC2 protein expression and the clinicopathological features of NPC.

Clinical significance of the statistical analysis of ZIC2 in NPC

After comparing the expression of ZIC2 in a single NPC database, we used the same chip platform to enlarge the sample size for comprehensive analyze of the ZIC2 expression. We evaluated the ordinary ZIC2 expression level and the clinical value of gender and age using the standard mean difference (SMD) strategy with STATA 15.1 software. We used the chi-square and I^2 tests to detect the heterogeneity, and while $I^2 > 50\%$ ($P < 0.05$), a random-effects model would be adopted. Otherwise, a fixed-effect model may be suitable. We assessed the publication bias and the combined quality using Begg's funnel plots. Subsequently, we calculated continuous variables of the ZIC2 expression value in each dataset to true-positive, false-positive, false-negative, and true-negative counts, and discriminated the cut-off values. A summary of the receiver operating characteristic (sROC) curves is presented to thoroughly evaluate the general discriminatory power of ZIC2 between the NPC patients and the control group sample. In addition, we computed the sensitivity and specificity to determine the sensitivity of the positive diagnostic samples and to make specific judgments of the negative samples.

Prediction of the ZIC2 target genes

The upregulated CEGs may increase the expression level with the raised expression of ZIC2 upwards. Therefore, the correlation between ZIC2 and ATAD2 is presented with a correlation scatter diagram generated using GraphPad Prism software 8.0.To determine the target gene of ZIC2, we intersected the score that was greater than or equal to 1 in the Cistrome Data Browser (Cistrome DB; http://cistrome.org/db/#/, a database for analyzing human and mouse genome data using ChIP-seq data and chromatin accessibility data) and the upregulated CEGs that appeared in at least five databases. We used the Integrative Genomics Viewer (IGV) software 2.9.4 to visualize the peak calling results.

Functional enrichment analysis

We analyzed the conjunction gene sets of the differential expression- and ZIC2-related genes via the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. First, we completed the GO and KEGG pathway analysis using R packages ("clusterProfiler," "enrichplot," and "ggplot2"). GO analysis has three main parts: analysis of the biological process (BP), analysis of the cellular component (CC), and analysis of the molecular function (MF). Likewise, we submitted the above two gene sets to the Search Tool for the Retrieval of Interacting Genes (STRING) (http://string-db.org/) database for protein-protein interaction (PPI) network analysis. STRING calculated the tab-separated value between the proteins and referred it to the Cytoscape 3.7.2 software plug-in cytoHubba of the top 10 hub genes. We also performed Gene Set Enrichment Analysis (GSEA) of gene expression profile data to group and classify the known genes according to multiple functional gene sets in order to understand their expression status in

specific functional gene sets. We put the gene expression profile data in the GSE12452 chipset into the GSEA 4.1.0 software for analysis. We used the Gene matrices c5.go.v7.4.symbols.gmt [Gene ontology] and c2.cp.kegg.v7.3.symbols.gmt [Curated] from the Gene sets database separately for the GO and KEGG analysis.

WGCNA

Weighted correlation network analysis (WGCNA) can identify highly collaborative gene sets and candidate biomarker genes or therapeutic targets according to the interlinkages of the gene sets to explore the association between the gene network and the phenotypes concerned. We first selected the GSE12452 expression profile data as analysis objects, and then recognized the outlier sample, using the screening condition height = 90. We used the dynamic cut tree algorithm to cluster the topological gene matrix and identify the network module. After removing this sample, based on the soft threshold power β, we constructed the adjacency matrix to build gene co-expression networks. Next, we transformed the topological overlap matrix (TOM), which compared the weighted correlations between the nodes, using the preceding adjacency matrix to represent the hierarchical clustering results that contained at least 30 genes (minModuleSize = 30). After performing the dynamic cut tree method, we merged these similar modules (abline = 0.25) and finally turned them into gene network modules. We used the module eigengenes (MEs), which is the first representative component of the module, to calculate the correlation with the clinical features in each module. We implemented all these using the WGCNA package of the R 4.0.3 software.

Results

ZIC2 mRNA and protein expression levels in NPC

We included eight gene chip datasets (GSE12452, GSE34573, GSE40290, GSE53819, GSE61218, GSE64634, GSE68799, and GSE126683) in this study that comprised 57 control samples and 157 NPC samples. The TMAs contained 63 chronic nasopharyngeal mucositis tissues and 166 NPC tissues. Violin plots differentiated the NPC tissues from the non-NPC tissues based on their expression values and ROC curves, as shown in Figs. 1 and 2, respectively. To verify the results of the ZIC2 expression level from public databases and microarrays, we evaluated the expression of the ZIC2 protein in a wide variety of NPC tissues and non-NPC tissues to confirm their expression in situ via IHC (Fig. 3). Both of them showed that the expression level of ZIC2 was enhanced in NPC. The basic information for each dataset are shown in Table I. In addition, we screened the DEGs and CEGs of each dataset. Altogether, there were 1,047 upregulated and 1,280 downregulated DEGs of NPC that had no fewer than two gene chip datasets. We also identified 4,321 CEGs that were positively related to DLX2 and 2,906 CEGs that were negatively related to DLX2 in no less than two datasets. The intersection of the DEGs and CEGs separately produced 808 positive gene sets and 859 negative gene sets.

Table : The basic features of high throughput datasets of ZIC2 expression profiling included in this study

Study	Platforms	Country	Control group	Cancer group	Mean0± SD0	Mean1± SD1	T	P
GSE12452	GPL570	USA	10	31	5.03±0.48	7.39±0.99	7.22	0.001
GSE34573	GPL570	UK	4	16	4.04±1..82	6.03±1.39	2.42	0.026
GSE40290	GPL8380	China	8	25	0.13±0.53	0.33±0.22	2.49	0.018
GSE53819	GPL6480	China	18	18	3.00±0.30	4.61±1.66	4.04	0.001
GSE61218	GPL19061	China	6	10	1.22±0.22	3.02±1.05	5.25	0.001
GSE64634	GPL570	China	4	12	4.84±2.21	7.66±1.70	2.68	0.018
GSE68799	GPL11154	China	4	42	0.36±0.61	4.11±4.11	1.81	0.078
GSE126683	GPL16956	China	3	3	11.37±0.31	11.62±0.46	0.78	0.479
NPC131	——	Canada	33	98	1.61±1.27	7.89±1.722	19.24	0.001
NPC241	——	Canada	12	12	1.92±1.00	8.50±0.91	16.95	0.001
NPC482	——	Canada	4	44	2.00±0.82	7.73±1.99	5.66	0.001
In-house	——	China	14	12	0.93±1.07	7.25±2.05	10.06	0.001
Total			120	323				

Mean0± SD0: control group Mean1± SD1: cancer group

Clinical significance of ZIC2 in NPC

We analyzed relevant clinical information on tissue microarrays and performed standardized mean difference (SMD) analysis with regard to age and gender. The results showed no statistical correlation with clinical features (Table). Moreover, our meta-analysis showed that the ZIC2 expression was upregulated in the NPC tissues, unlike in the normal nasopharynx tissues [SMD = 2.14, 95% confidence interval (CI): 1.85–2.44], and there was pronounced heterogeneity (I^2 = 87.9%, P < 0.0001), so we used the random-effect model (Fig. 4A). However, when there was publication bias, the funnel plot was asymmetric, and the distribution was skewed. Begg's funnel plot analysis gave an indication of no significant publication bias, with the P-value = 0.244 (Fig. 4B). Furthermore, sensitivity analysis clearly showed the degree of detection bias (Fig. 4C). Multiple diverse trials with the same index were represented by an ROC curve based on the weight of their odds ratio (OR) in the meta-analysis, called "sROC," the AUC value of which was 0.93 (95% CI: 0.91–0.95) (Fig. 5A). In the Fagan plot, the pretest probability, posttest probability positive, and post-test probability negative were 20%, 51%, and 1%, respectively (Fig. 5B). The corresponding sensitivity and specificity were 0.97 (95% CI: 0.91–0.99) and

0.77 (95% CI: 0.63–0.86), respectively (Fig. 5C), which suggest that ZIC2 demonstrated a significant discriminatory capacity for NPC screening.

Table : The relationship between clinicopathological parameters and ZIC2 protein expression of TMEs and in-house samples

Characteristic	Cases	Mean±SD	T	P
Tissue type				
Control	63	1.54±1.19	25.95	<0.0001
NPC	166	7.84±1.78		
Age (year)				
<50	79	7.71±2.00	1.273	0.205
≥50	74	8.07±1.47		
Gender				
Male	120	7.96±1.71	1.02	0.311
Female	33	7.61±1.95		

The target gene of the transcription factor ZIC2

As a transcription factor, ZIC2 binding to the ABCC4 promoter regulates prostate cancer proliferation (23). This suggests that ZIC2 may also be involved in transcriptional regulation as a transcription factor in NPC. Overall, there were six overlapping genes: RACGAP1, PCNA, ATAD2, RBBP8, HSPE1, and UHRF1. Next, after the presence of DNA peaks in the ZIC2 transcription regulatory regions was re-analyzed, we chose the ATPase family AAA domain containing protein 2 (ATAD2) as the putative target gene (Fig. 6). Then we transferred their correlation in each gene set into a scatter plot.

Enrichment analysis and PPI network construction

We used GO and KEGG to analyze the overall function of ZIC2 in NPC. Regarding the positive gene set related to ZIC2, our GO analysis revealed the importance of the nuclear division, chromosomal region, and ATPase activity in terms of the BP, CC, and MF, respectively, and our KEGG pathway enrichment analysis mainly investigated the human papillomavirus infection (Fig. 7A-B). For the negative gene set, microtubule-based movement, motile cilium, and peptidase regulator activity were the most clustered in the BP, CC, and MF terms, respectively. Our KEGG pathway enrichment analysis showed that the pathways were mainly enriched in the chemokine signaling pathway (Fig. 8A–B). To further explore the potential mechanism of ZIC2, we used the genes in the first three KEGG signaling pathways for the PPI analysis. The analysis identified the top 10 positively related core genes as CDK1, CCNB1, CCNA2, CCNE2, CDC6, MCM2, CDC45, CCNB2, MCM4, and BUB1B (Fig. 9A), and the top 10 negatively related core genes as CD79A, CD22, CD5, CD19, BTK, CD79B, CR2, MS4A1, CD37, and CD72 (Fig. 9B). The results of the GSEA enrichment analysis were similar to those of GO

and KEGG. The most significant pathways in the BP, CC, and MF were the ribonucleoprotein complex biogenesis, chromosomal region, and single-stranded DNA binding, respectively. The GSEA analysis revealed that highly expressed genes are mainly enriched in the ribonucleoprotein complex biogenesis, chromosomal region, and single-stranded DNA binding in the GO gene feature set database and in the cell cycle in the KEGG gene feature set database (Table III).

Table Ⅲ: GSEA analysis of genes positively related to ZIC2 involved in GO and KEGG

Tag	NAME	SIZE	ES	NES	NOM p-val	FDR q-val
GO-BP	ribonucleoprotein_complex_biogenesis	410	2.6169	0.5528	0.0	0.0
GO-BP	DNA _replication	272	2.5399	0.5545	0.0	0.0
GO-CC	chromosomal_region	327	2.7182	0.5844	0.0	0.0
GO-CC	chromosome_centromeric_region	192	2.6781	0.6122	0.0	0.0
GO-MF	single_stranded_DNA_binding	110	2.4374	0.5972	0.0	0.0
GO-MF	DNA_dependent_ATPase_activity	50	2.2545	2.6859	0.0	0.0
KEGG	cell_cycle	122	0.6264	2.5898	0.0	0.0
KEGG	spliceosome	115	0.6063	2.4749	0.0	0.0

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, biological process; CC, cellular component; MF, molecular function; ES, enrichment score; NES, normalized enrichment score; NOM p-val; norminal P value; FDR, false discovery rate

WGCNA analysis

First, we constructed a sample clustering tree and ruled out a sample of GSM312909 because it was above the threshold, as shown in Fig. 11A. In our results, we set the soft threshold at 10 to be able to construct the scale-free network (R^2 = 0.94) (Fig. 11B). Finally, a total of 17 modules were generated based on the 0.25 cut-off value of dissimilarity of the modules and on average hierarchical clustering (Fig. 11C). The relationship between the modules and the clinical traits is shown in Fig. 11D. The red module denotes the positive correlation with the corresponding clinical trait, and the blue module denotes the opposite. The stronger the relevance is, the darker the color is. The illustration shows that the tan module that was positively associated with the N stage was the deepest (cor = 0.38, P = 0.04) (the grey module represents the genes that were not divided into any module), and the green module that was negatively associated with the grade was the deepest (cor = -0.36, P = 0.05).

Discussion

The occurrence and development of malignant tumors involve multiple signaling molecular pathways. In recent years, the genomic changes that are key to the progression of NPC have been revealed as the multiple functional loss mutations in the negative regulator of NF-κB; recurring gene damage, including deletion of the CDKN2A/CDKN2B locus, amplification of CCND1, TP53 mutation, and changes in the PI3K/MAPK signaling

pathway; chromatin modification; and DNA repair (24–27). In addition, some studies have explored molecular markers related to nasopharyngeal carcinoma. For example (28–30), CKS1, P27, and CENP-F have been found to be involved in the regulation of an abnormal cell cycle. Some oncoproteins and tumor suppressor genes, such as RAP2A, PTP4A2, and ECRG4 (31–33), are closely related to cell proliferation. Although many oncogenes and carcinogenic factors have been found to affect the occurrence and development of NPC, the exact mechanism is still unknown.

Many studies recently reported that ZIC2 might regulate the development of various tumors, including prostate cancer (23, 34–36), lung adenocarcinoma (17), breast cancer, clear cell renal cell carcinoma (15, 37, 38), colorectal cancer (14, 39–41), hepatocellular carcinoma (13, 42–44), and cervical cancer (10, 45). All these tumors, except breast cancer, were upregulated. In this study, a total of 7 microarrays, 3 TMEs, and in-house tissues were incorporated, including 120 control samples and 323 NPC samples. ZIC2 was excessively expressed in NPC than in the normal tissues according to the gene chip, tissue chip data, and IHC results of the tissue samples. These results are consistent with those of related studies, and certainly add to our understanding of the expression of ZIC2 in NPC at the human tissue level, which no relevant study had previously investigated. Moreover, by mining data from the microarrays data as mentioned above, we observed that ATAD2 expression is positively correlated to ZIC2 expression. Interestingly, when we searched for the binding position of the ChIP-seq data in the Cistrome DB, we found that ZIC2 and ATAD2 could combine and showed peaks. ATAD2—or ANCCA, CT137, or PRO2000, as it is also called—is a member of the ATPase family, which is associated with various cellular activities by regulating protein complexes and is responsible for ATP binding and hydrolysis (46). Members of the ATPase family of proteins participate in diverse cellular processes that include cell cycle regulation, protein proteolysis and decomposition, organelle biogenesis, and intracellular trafficking (47). Being deficient in normal regulation, either due to the ATAD2 locus amplification or to specific changes in the core members of the transcription mechanism, may cause ATAD2 abnormal activation, which will eventually lead to oncogenesis (48, 49). Moreover, ATAD2 showed expression disturbance in hepatocellular carcinoma, ovarian cancer, stomach cancer, and other cancers (50–52). The discovery of Liu et al., through bioinformatics analysis, that ATAD2 may take part in the tumorigenesis of NPC is noteworthy (53). This also agrees with our findings. Besides, our GO and KEGG enrichment analysis showed that ZIC2 positive relation genes are mainly enriched in cell division and in the cell cycle. Likewise, our GSEA analysis revealed that the upregulation of ZIC2 is associated with the proliferation and division of cells and DNA synthesis, which is consistent with the results of our GO and KEGG analysis. Another crucial finding is that highly expressed genes are also enriched in DNA-dependent ATPase activity in GO-MF, because their putative target gene, ATAD2, functions similarly as it belongs to the ATPase family. In general, the ATPase family may play a role in the preceding cell functional activities, among which is ATAD2.Evidence from function analysis supports this result.

Nevertheless, this study had some limitations. First, it tested only the expression of ZIC2 protein using IHC and TME. Further research should investigate the clinical significance of ZIC2 mRNA expression in NPC; and to confirm the results, in vivo and in vitro experiments are needed. Another limitation of this study is that none of its experiments verified the transcriptional regulation between ZIC2 and ATAD2. Therefore, further analysis is needed. Finally, the results of the WGCNA analysis suggested that the modules in GSE12452 have little

significant correlation with clinical traits. However, this was possibly due to the limited number of samples. More clinical data should bfirste included in the follow-up study.

Conclusion

In this study, ZIC2 is upregulated in NPC tissues and could serve as a potential biomarker for the diagnosis and prognosis of NPC patients. Furthermore, ZIC2 may play a role in tumorigenesis through transcriptionally regulating potential target genes ATAD2 via modulating cancer cell proliferation and cell cycle regulation pathways.

Abbreviations

NPC Nasopharyngeal carcinoa

ASIR Age-standardized incidence rates

ZIC2 Zinc finger protein of the cerebellum 2

HPE Holoprosencephaly

GEO Gene Expression Omnibus

SRA Short Read Archive

DEGs Differentially expressed genes

CEGs Co-expressed genes

TMAs Tissue microarrays

IHC Immunohistochemistry

IRS Immunoreactivity score

SMD Standard mean difference

sROC A summary of the receiver operating characteristic

ATAD2 ATPase family AAA domain containing protein 2

Cistrome DB Cistrome Data Browser

IGV Integrative Genomics Viewer

GO Gene Ontology

KEGG Kyoto Encyclopedia of Genes and Genomes

BP Biological process

CC Cellular component

MF Molecular function

STRING Search Tool for the Retrieval of Interacting Gene

PPI Protein-protein interaction

GSEA Gene Set Enrichment Analysis

WGCNA Weighted correlation network analysis

TOM Topological overlap matrix

MEs Module eigengenes

CI Confidence interval

OR Odds ratio

AUC Area under curve

Declarations

Acknowledgements

Not applicable.

Ethics approval and consent to participate

The current study was approved by the Ethics Committee of the First Affiliated

Hospital of Guangxi Medical University.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analysed during the current study are available in the GEO repository, [https://www.ncbi.nlm.nih.gov/geo/]

Competing interests

The authors declare that they have no competing interests.

Funding

Guangxi Zhuang Autonomous Region Health and Family Planning Commission Self-financed Scientific Research Project (Z2014046) and Guangxi Medical and Health

Appropriate Technology Development and Promotion Application Project (S2017020)

Authors' contributions

PYX, LJ and DH contributed significantly to analysis and manuscript preparation; FYY, QW and LJL helped perform the analysis with constructive discussions; JL, MM and LCX performed the experiment; PYX, WJY, performed the data analyses and wrote the manuscript; WZX, HSN contributed to the conception of the study.

All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal carcinoma. Lancet. 2019;394(10192):64–80.
3. Hutajulu SH, Kurnianda J, Tan IB, Middeldorp JM. Therapeutic implications of Epstein-Barr virus infection for the treatment of nasopharyngeal carcinoma. Ther Clin Risk Manag. 2014;10:721–36.
4. de Almeida IG Jr, Kuratani DK, Gomes LM, Fiegenbaum M, Estima Correia EP, Gazzola Zen PR, et al. Nasal fistula, epidermal cyst and hypernatremia in a girl presenting holoprosencephaly due to a rare ZIC2 point mutation. Eur J Med Genet. 2020;63(2):103641.
5. Savastano CP, El-Jaick KB, Costa-Lima MA, Abath CM, Bianca S, Cavalcanti DP, et al. Molecular analysis of holoprosencephaly in South America. Genet Mol Biol. 2014;37(1 Suppl):250–62.
6. Ali RG, Bellchambers HM, Arkell RM. Zinc fingers of the cerebellum (Zic): transcription factors and co-factors. Int J Biochem Cell Biol. 2012;44(11):2065–8.
7. Nagai T, Aruga J, Minowa O, Sugimoto T, Ohno Y, Noda T, et al. Zic2 regulates the kinetics of neurulation. Proc Natl Acad Sci U S A. 2000;97(4):1618–23.
8. Brown SA, Warburton D, Brown LY, Yu CY, Roeder ER, Stengel-Rutkowski S, et al. Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. Nat Genet. 1998;20(2):180–3.
9. Zhang P, Yang F, Luo Q, Yan D, Sun S. miR-1284 Inhibits the Growth and Invasion of Breast Cancer Cells by Targeting ZIC2. Oncol Res. 2019;27(2):253–60.
10. Wang YF, Yang HY, Shi XQ, Wang Y. Upregulation of microRNA-129-5p inhibits cell invasion, migration and tumor angiogenesis by inhibiting ZIC2 via downregulation of the Hedgehog signaling pathway in cervical

cancer. Cancer Biol Ther. 2018;19(12):1162–73.

11. Inaguma S, Ito H, Riku M, Ikeda H, Kasai K. Addiction of pancreatic cancer cells to zinc-finger transcription factor ZIC2. Oncotarget. 2015;6(29):28257–68.
12. Chen X, Yang S, Zeng J, Chen M. miR–1271–5p inhibits cell proliferation and induces apoptosis in acute myeloid leukemia by targeting ZIC2. Mol Med Rep. 2019;19(1):508–14.
13. Lu SX, Zhang CZ, Luo RZ, Wang CH, Liu LL, Fu J, et al. Zic2 promotes tumor growth and metastasis via PAK4 in hepatocellular carcinoma. Cancer Lett. 2017;402:71–80.
14. Tan Y, Hu Y, Xiao Q, Tang Y, Chen H, He J, et al. Silencing of brain-expressed X-linked 2 (BEX2) promotes colorectal cancer metastasis through the Hedgehog signaling pathway. Int J Biol Sci. 2020;16(2):228–38.
15. Wu CY, Li L, Chen SL, Yang X, Zhang CZ, Cao Y. A Zic2/Runx2/NOLC1 signaling axis mediates tumor growth and metastasis in clear cell renal cell carcinoma. Cell Death Dis. 2021;12(4):319.
16. Jiang Z, Zhang Y, Chen X, Wu P, Chen D. Inactivation of the Wnt/β-catenin signaling pathway underlies inhibitory role of microRNA-129-5p in epithelial-mesenchymal transition and angiogenesis of prostate cancer by targeting ZIC2. Cancer Cell Int. 2019;19:271.
17. Wei-Hua W, Ning Z, Qian C, Dao-Wen J. ZIC2 promotes cancer stem cell traits via up-regulating OCT4 expression in lung adenocarcinoma cells. J Cancer. 2020;11(20):6070–80.
18. Jiang X, Feng L, Dai B, Li L, Lu W. Identification of key genes involved in nasopharyngeal carcinoma. Braz J Otorhinolaryngol. 2017;83(6):670–6.
19. Yi W, Wang J, Yao Z, Kong Q, Zhang N, Mo W, et al. The expression status of ZIC2 as a prognostic marker for nasopharyngeal carcinoma. Int J Clin Exp Pathol. 2018;11(9):4446–60.
20. Yu D, Han GH, Zhao X, Liu X, Xue K, Wang D, et al. MicroRNA-129-5p suppresses nasopharyngeal carcinoma lymphangiogenesis and lymph node metastasis by targeting ZIC2. Cell Oncol (Dordr). 2020;43(2):249–61.
21. Lv B, Li F, Liu X, Lin L. The tumor-suppressive role of microRNA-873 in nasopharyngeal carcinoma correlates with downregulation of ZIC2 and inhibition of AKT signaling pathway. Cancer Gene Ther. 2021;28(1–2):74–88.
22. Shen ZH, Zhao KM, Du T. HOXA10 promotes nasopharyngeal carcinoma cell proliferation and invasion via inducing the expression of ZIC2. Eur Rev Med Pharmacol Sci. 2017;21(5):945–52.
23. Bawa PS, Ravi S, Paul S, Chaudhary B, Srinivasan S. A novel molecular mechanism for a long non-coding RNA PCAT92 implicated in prostate cancer. Oncotarget. 2018;9(65):32419–34.
24. Lin DC, Meng X, Hazawa M, Nagata Y, Varela AM, Xu L, et al. The genomic landscape of nasopharyngeal carcinoma. Nat Genet. 2014;46(8):866–71.
25. Zheng H, Dai W, Cheung AK, Ko JM, Kan R, Wong BW, et al. Whole-exome sequencing identifies multiple loss-of-function mutations of NF-κB pathway regulators in nasopharyngeal carcinoma. Proc Natl Acad Sci U S A. 2016;113(40):11283–8.
26. Li YY, Chung GT, Lui VW, To KF, Ma BB, Chow C, et al. Exome and genome sequencing of nasopharynx cancer identifies NF-κB pathway activating mutations. Nat Commun. 2017;8:14121.
27. Zhang L, MacIsaac KD, Zhou T, Huang PY, Xin C, Dobson JR, et al. Genomic Analysis of Nasopharyngeal Carcinoma Reveals TME-Based Subtypes. Mol Cancer Res. 2017;15(12):1722–32.

28. Xu L, Fan S, Zhao J, Zhou P, Chu S, Luo J, et al. Increased expression of Cks1 protein is associated with lymph node metastasis and poor prognosis in nasopharyngeal carcinoma. Diagn Pathol. 2017;12(1):2.
29. Jiang Q, Yang H, Cheng C, Xiong H, Liu S, Long J, et al. Decreased P27 protein expression is correlated with the progression and poor prognosis of nasopharyngeal carcinoma. Diagn Pathol. 2013;8:212.
30. Cao JY, Liu L, Chen SP, Zhang X, Mi YJ, Liu ZG, et al. Prognostic significance and therapeutic implications of centromere protein F expression in human nasopharyngeal carcinoma. Mol Cancer. 2010;9:237.
31. Lee YE, He HL, Chen TJ, Lee SW, Chang IW, Hsing CH, et al. The prognostic impact of RAP2A expression in patients with early and locoregionally advanced nasopharyngeal carcinoma in an endemic area. Am J Transl Res. 2015;7(5):912–21.
32. Gao Y, Zhang M, Zheng Z, He Y, Zhu Y, Cheng Q, et al. Over-expression of protein tyrosine phosphatase 4A2 correlates with tumor progression and poor prognosis in nasopharyngeal carcinoma. Oncotarget. 2017;8(44):77527–39.
33. Chen JY, Wu X, Hong CQ, Chen J, Wei XL, Zhou L, et al. Downregulated ECRG4 is correlated with lymph node metastasis and predicts poor outcome for nasopharyngeal carcinoma patients. Clin Transl Oncol. 2017;19(1):84–90.
34. Toth R, Schiffmann H, Hube-Magg C, Büscheck F, Höflmayer D, Weidemann S, et al. Random forest-based modelling to detect biomarkers for prostate cancer progression. Clin Epigenetics. 2019;11(1):148.
35. Jiang Z, Zhang Y, Chen X, Wu P, Chen D. Correction to: Inactivation of the Wnt/β-catenin signaling pathway underlies inhibitory role of microRNA-129-5p in epithelial-mesenchymal transition and angiogenesis of prostate cancer by targeting ZIC2. Cancer Cell Int. 2021;21(1):237.
36. Hoogland AM, Böttcher R, Verhoef E, Jenster G, van Leenders GJ. Gene-expression analysis of gleason grade 3 tumor glands embedded in low- and high-risk prostate cancer. Oncotarget. 2016;7(25):37846–56.
37. Shang Z, Zhao T, Ou T, Yan H, Cui B, Wang Q, et al. The level of zinc finger of the cerebellum 2 is predictive of overall survival in clear cell renal cell carcinoma. Transl Androl Urol. 2020;9(2):614–20.
38. Kuleszo D, Lipska-Ziętkiewicz B, Koczkowska M, Zakrzewski K, Grajkowska W, Roszkowski M, et al. Hedgehog signalling network gene status analysis in paediatric intracranial germ cell tumours. Folia Neuropathol. 2019;57(3):227–38.
39. Xu Z, Zheng J, Chen Z, Guo J, Li X, Wang X, et al. Multilevel regulation of Wnt signaling by Zic2 in colon cancer due to mutation of β-catenin. Cell Death Dis. 2021;12(6):584.
40. Nagaraj SH, Reverter A. A Boolean-based systems biology approach to predict novel genes associated with cancer: Application to colorectal cancer. BMC Syst Biol. 2011;5:35.
41. Jeong E, Lee Y, Kim Y, Lee J, Yoon S. Analysis of Cross-Association between mRNA Expression and RNAi Efficacy for Predictive Target Discovery in Colon Cancers. Cancers (Basel). 2020;12(11).
42. Zhu P, Wang Y, He L, Huang G, Du Y, Zhang G, et al. ZIC2-dependent OCT4 activation drives self-renewal of human liver cancer stem cells. J Clin Invest. 2015;125(10):3795–808.
43. Sun L, Lin Y, Wang G, Zhang L, Hu L, Lu Z. Correlation of zinc finger protein 2, a prognostic biomarker, with immune infiltrates in liver cancer. Biosci Rep. 2021;41(1).
44. Sun XJ, Wang MC, Zhang FH, Kong X. An integrated analysis of genome-wide DNA methylation and gene expression data in hepatocellular carcinoma. FEBS Open Bio. 2018;8(7):1093–103.

45. Chan DW, Liu VW, Leung LY, Yao KM, Chan KK, Cheung AN, et al. Zic2 synergistically enhances Hedgehog signalling through nuclear retention of Gli1 in cervical cancer cells. J Pathol. 2011;225(4):525–34.
46. Snider J, Thibault G, Houry WA. The AAA + superfamily of functionally diverse proteins. Genome Biol. 2008;9(4):216.
47. Kedzierska S. [Structure, function and mechanisms of action of ATPases from the AAA superfamily of proteins]. Postepy Biochem. 2006;52(3):330–8.
48. Boussouar F, Jamshidikia M, Morozumi Y, Rousseaux S, Khochbin S. Malignant genome reprogramming by ATAD2. Biochim Biophys Acta. 2013;1829(10):1010–4.
49. Ciró M, Prosperini E, Quarto M, Grazini U, Walfridsson J, McBlane F, et al. ATAD2 is a novel cofactor for MYC, overexpressed and amplified in aggressive tumors. Cancer Res. 2009;69(21):8491–8.
50. Li Q, Qiu J, Yang H, Sun G, Hu Y, Zhu D, et al. Kinesin family member 15 promotes cancer stem cell phenotype and malignancy via reactive oxygen species imbalance in hepatocellular carcinoma. Cancer Lett. 2020;482:112–25.
51. Ge T, Liu T, Guo L, Chen Z, Lou G. MicroRNA-302 represses epithelial-mesenchymal transition and cisplatin resistance by regulating ATAD2 in ovarian carcinoma. Exp Cell Res. 2020;396(1):112241.
52. Nayak A, Kumar S, Singh SP, Bhattacharyya A, Dixit A, Roychowdhury A. Oncogenic potential of ATAD2 in stomach cancer and insights into the protein-protein interactions at its AAA + ATPase domain and bromodomain. J Biomol Struct Dyn. 2021:1–17.
53. Liu K, Kang M, Zhou Z, Qin W, Wang R. Bioinformatics analysis identifies hub genes and pathways in nasopharyngeal carcinoma. Oncol Lett. 2019;18(4):3637–45.

Figures

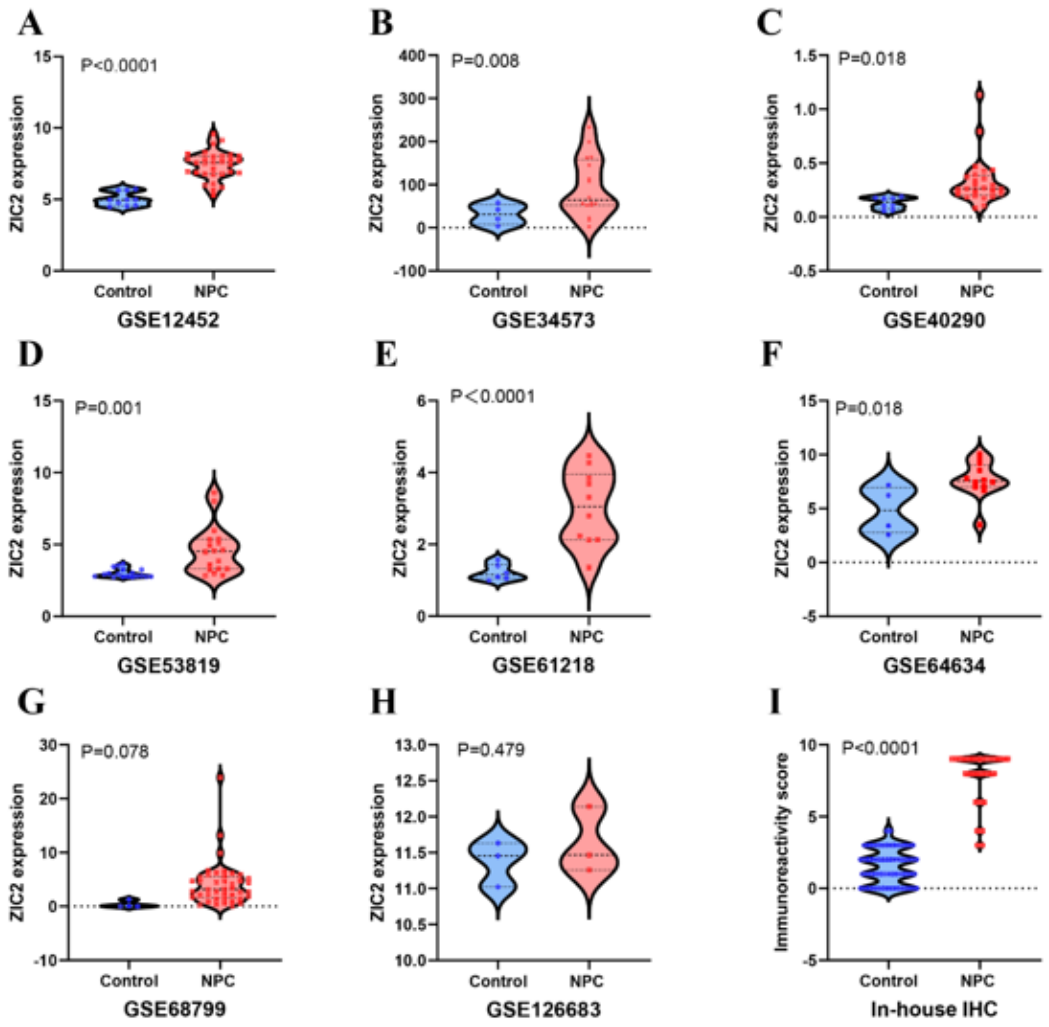

Figure 1

ZIC2 expression levels in NPC tissues based on multiple cohorts. (A-H) Different expression levels of ZIC2 between NPC and non-tumour nasopharynx tissues based on eight microarrays from GEO. (I) Expression level of ZIC2 in NPC and non-tumour nasopharynx tissues based on in-house tissues and TME.

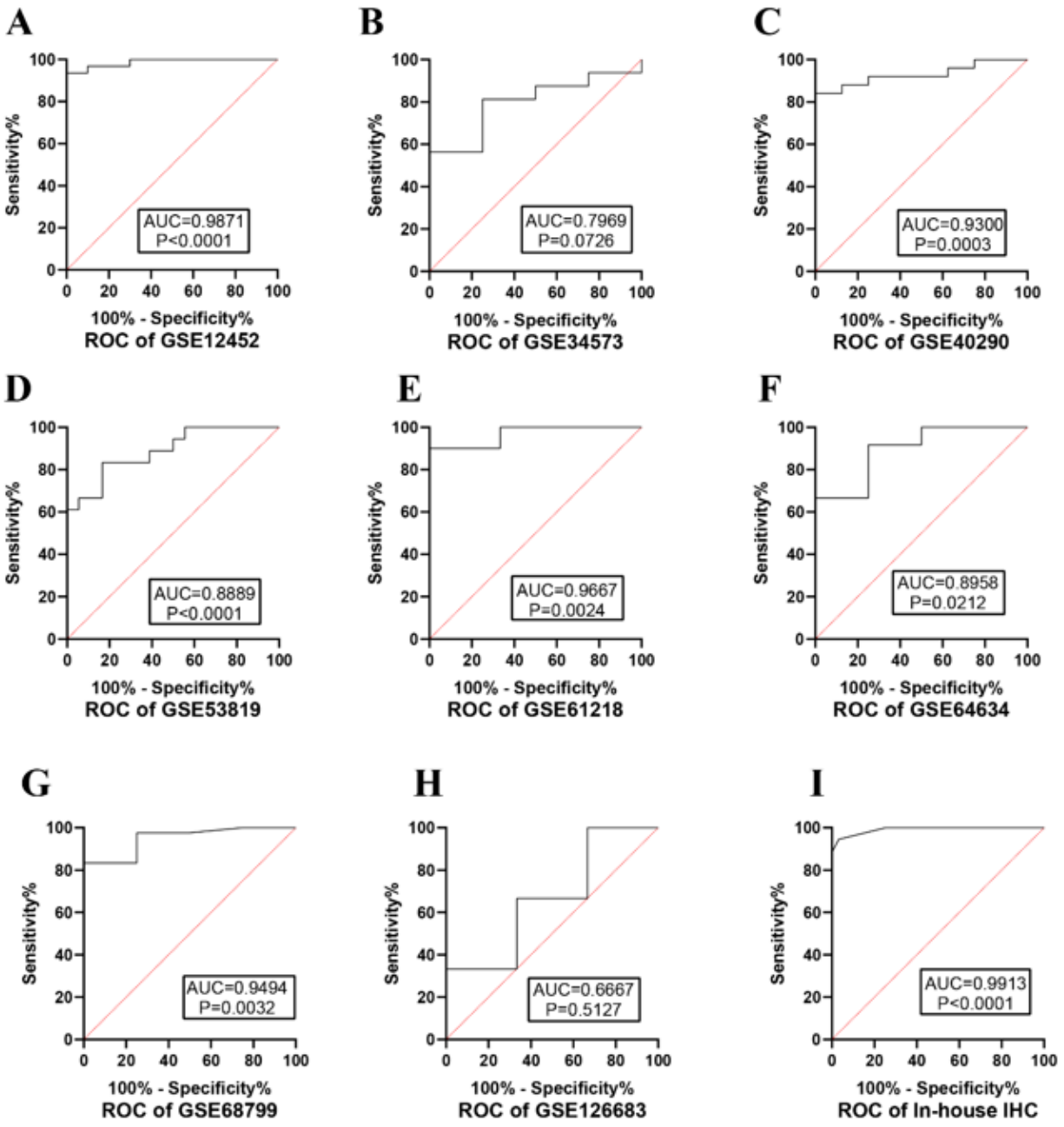

Figure 2

The expression data of ZIC2 and corresponding ROC curves in NPC tissues and normal tissues from GEO datasets and in-house tissues and TME.

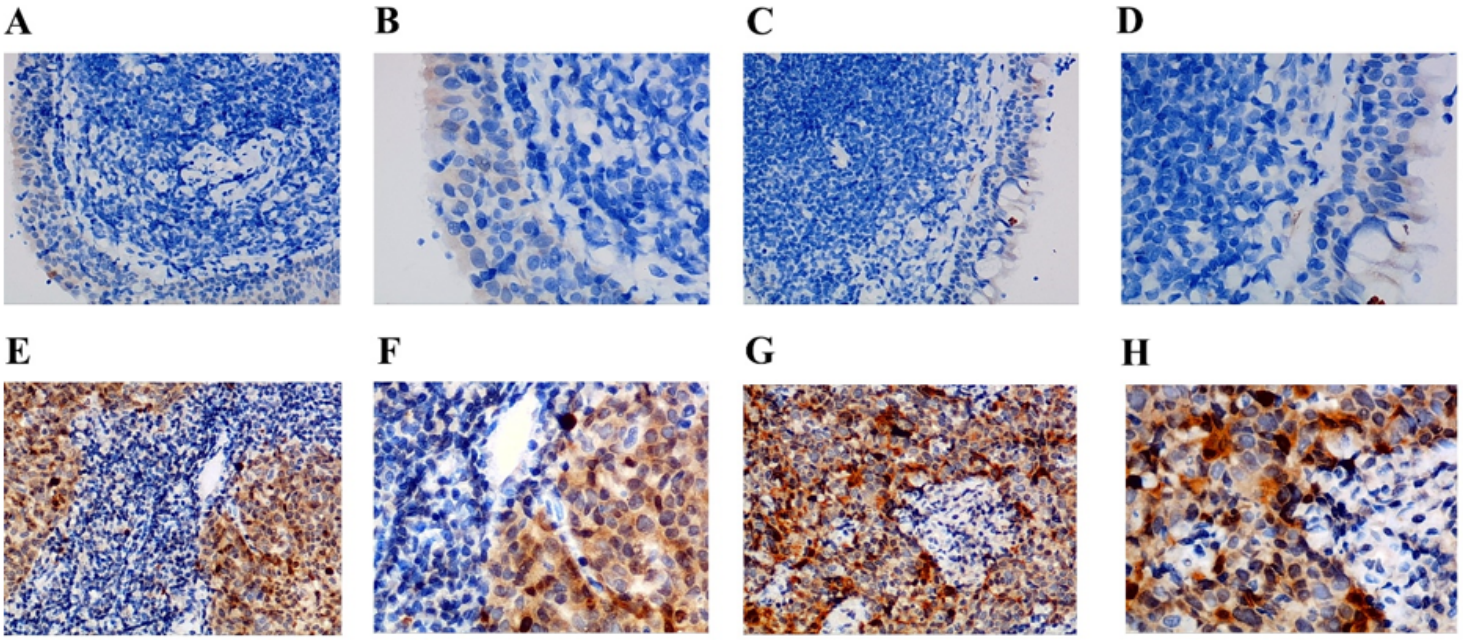

Figure 3

Expression of ZIC2 protein determined by immunohistochemistry in control and NPC tissue. (A-D) Non-NPC tissue. (E-H) NPC tissue. Magnification: 200× (A, C, E, G) ; 400×(B, D, F, H)

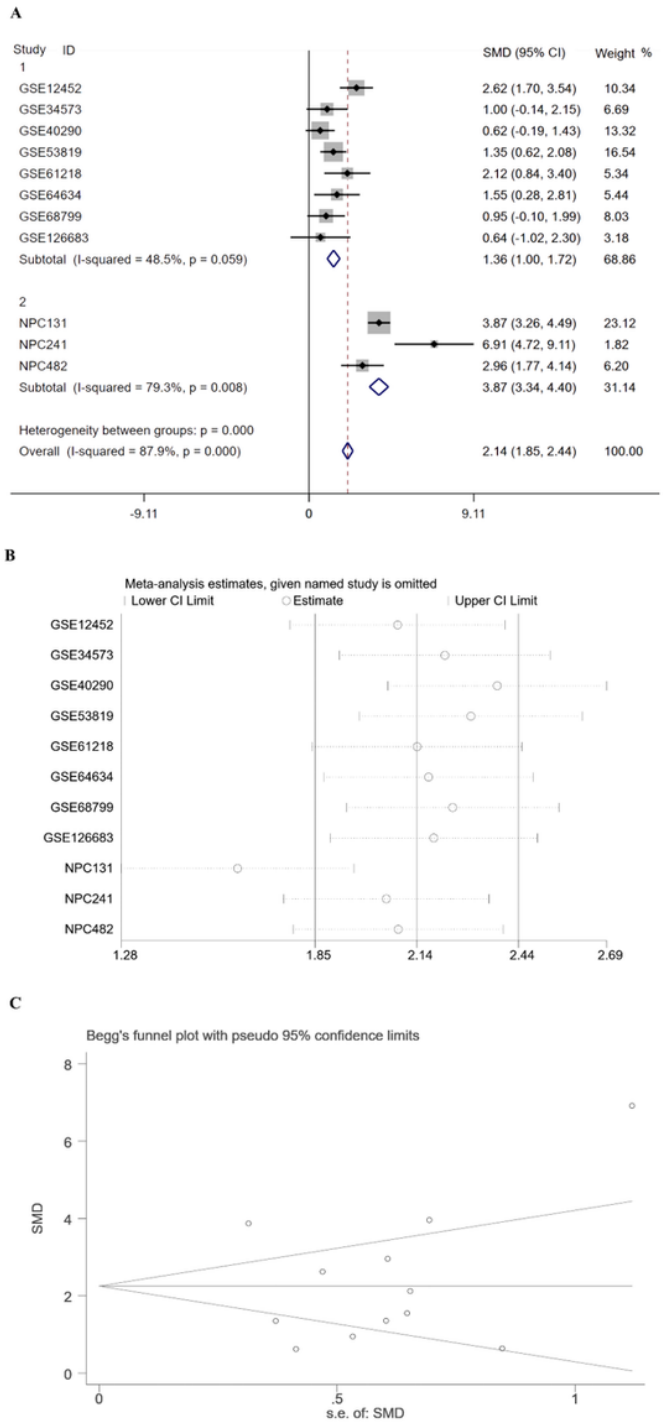

Figure 4

Expression level and survival analysis of ZIC2 based on GEO datasets, IHC and TMEs data. (A) Forest plots for evaluating standard mean difference of ZIC2 between NPC and noncancerous tissues based on GEO datasets. (B) Sensitivity analysis. (C) Corresponding funnel plot for publication bias test after analysis of the expression level of ZIC2 in NPC.

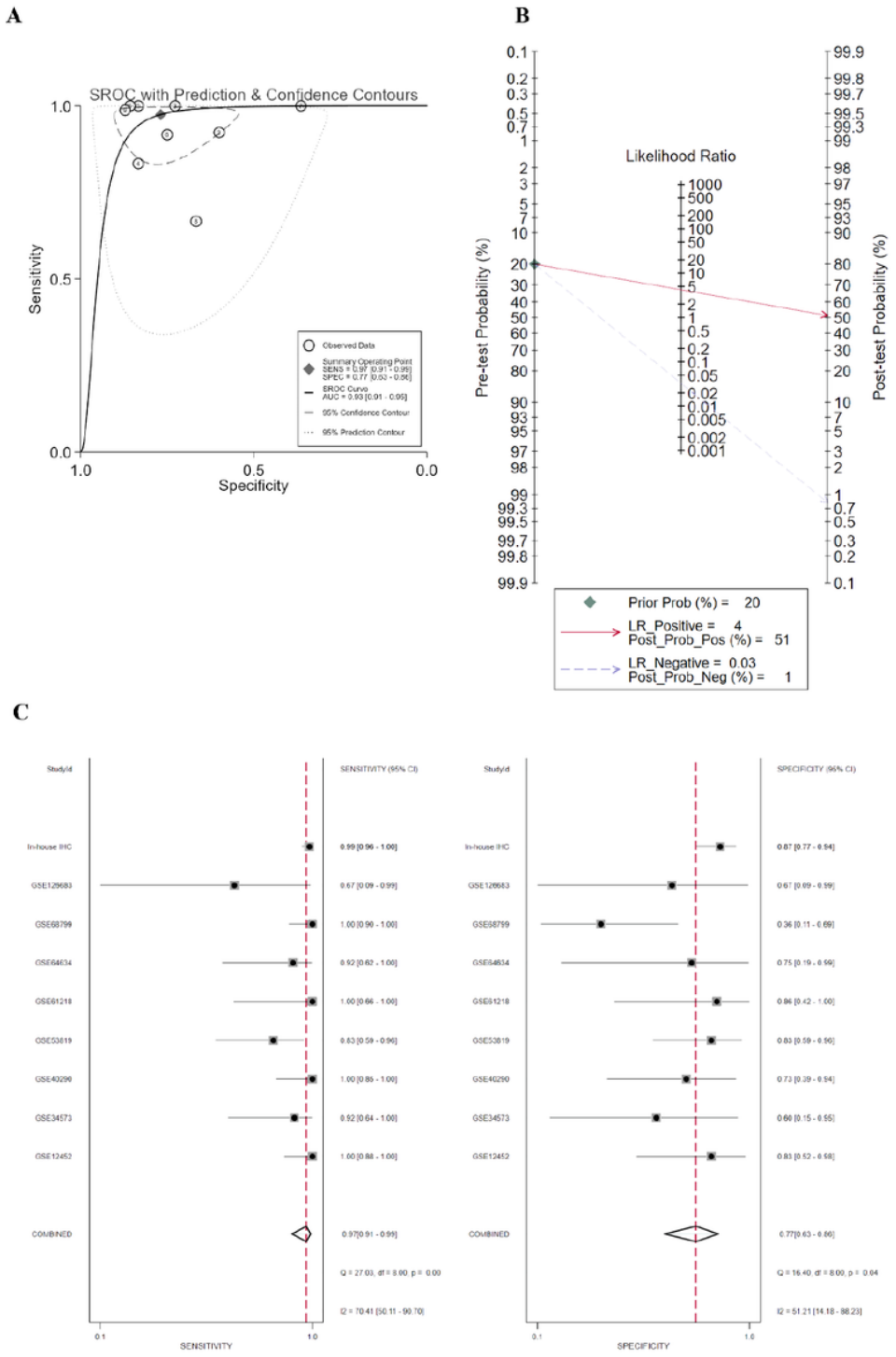

Figure 5

NPC patient diagnostic trial by using ZIC2 as a marker. (A) sROC plot. (B) Fagan's nomogram. (C) Forest plots of comprehensive sensitivity and comprehensive specificity.

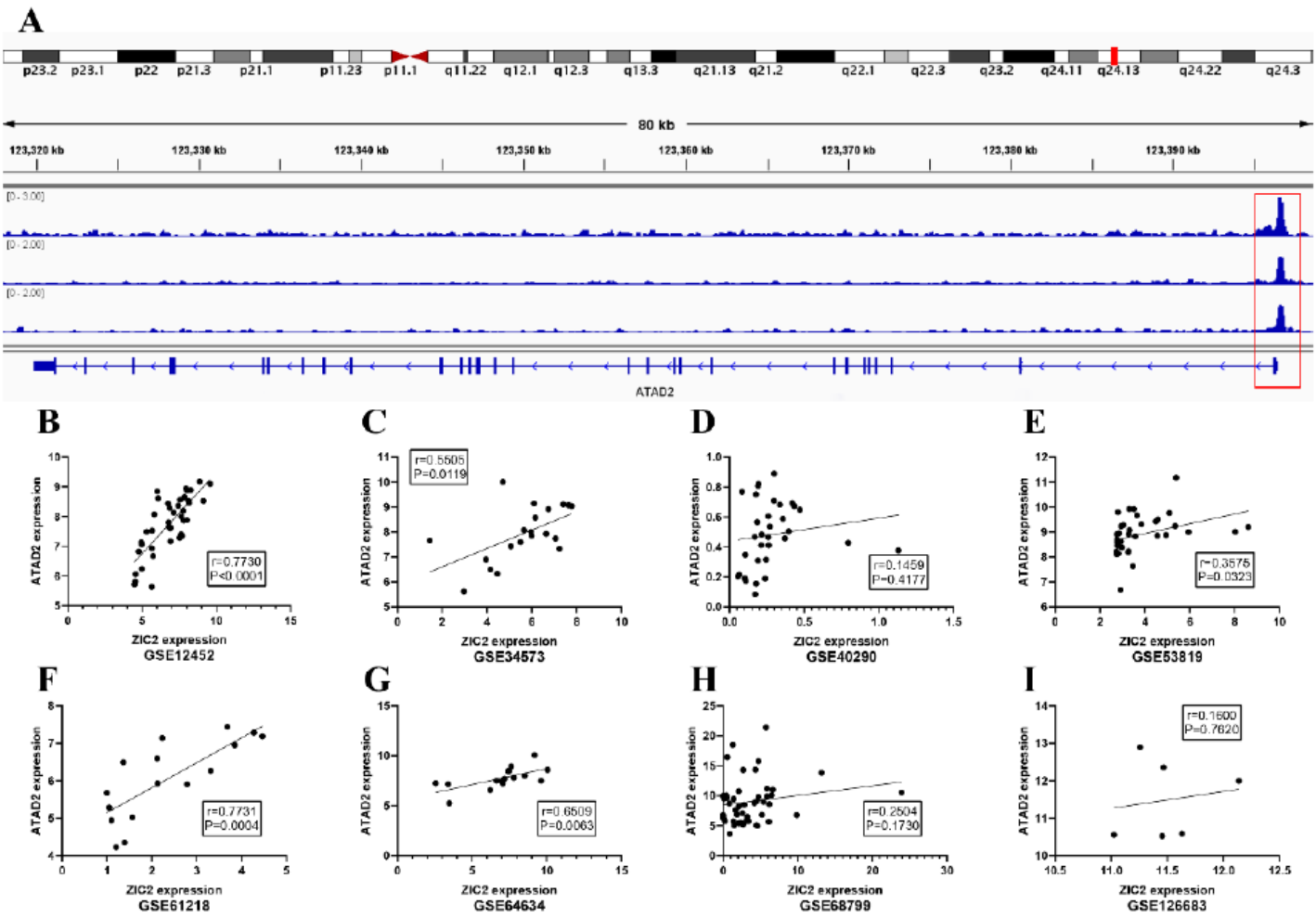

Figure 6

ATAD2 might become the target gene of transcription factor ZIC2. A: B-I: Correlation of ATAD2 expression with ZIC2 expression in GEO database.

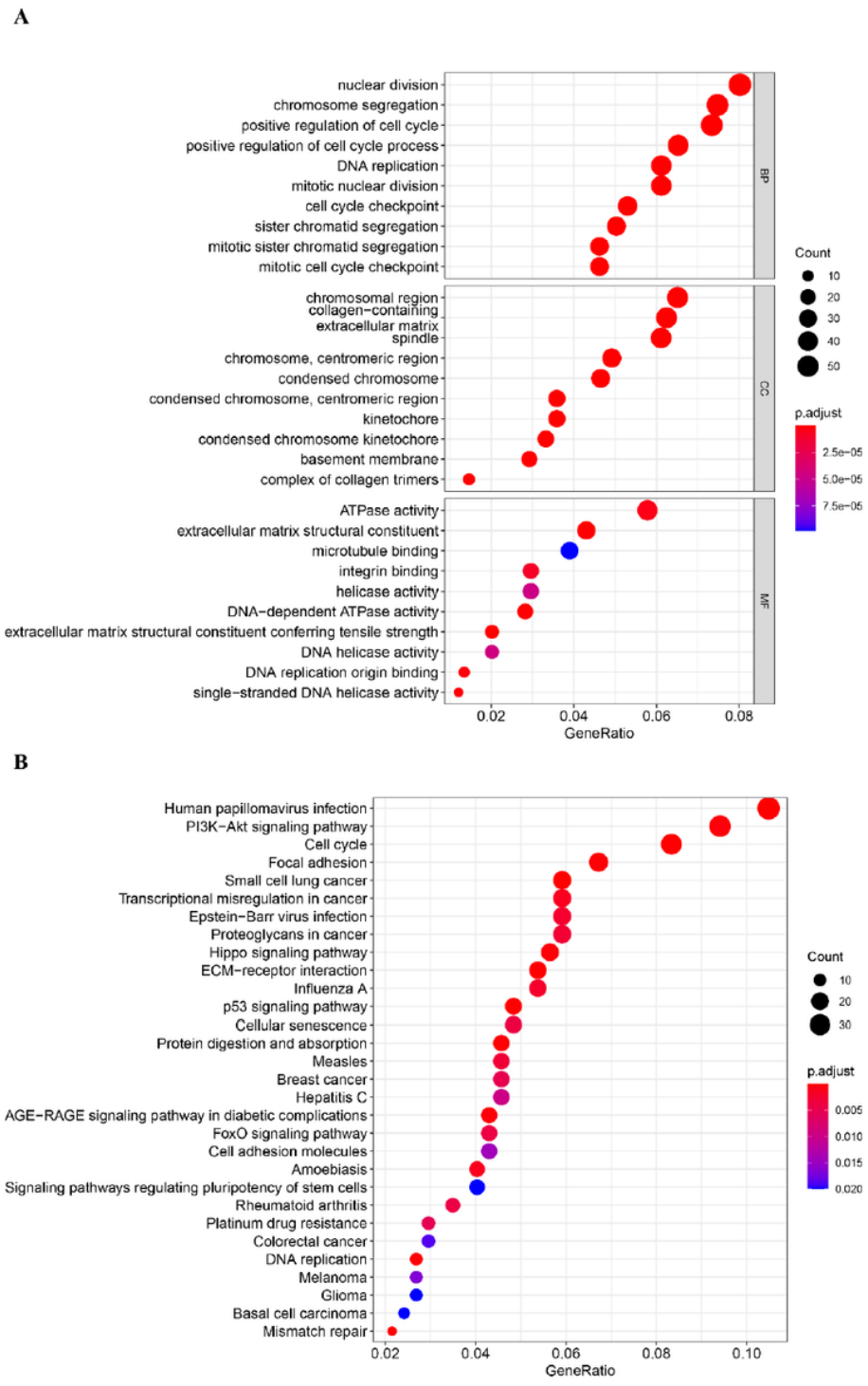

Figure 7

GO and KEGG analysis of intersection genes with up-regulated DEGs and ZIC2 positively related to CEGs. (A) The bubble plots of BP, CC, and MF of GO annotation. (B) The bubble graphs of KEGG annotation.

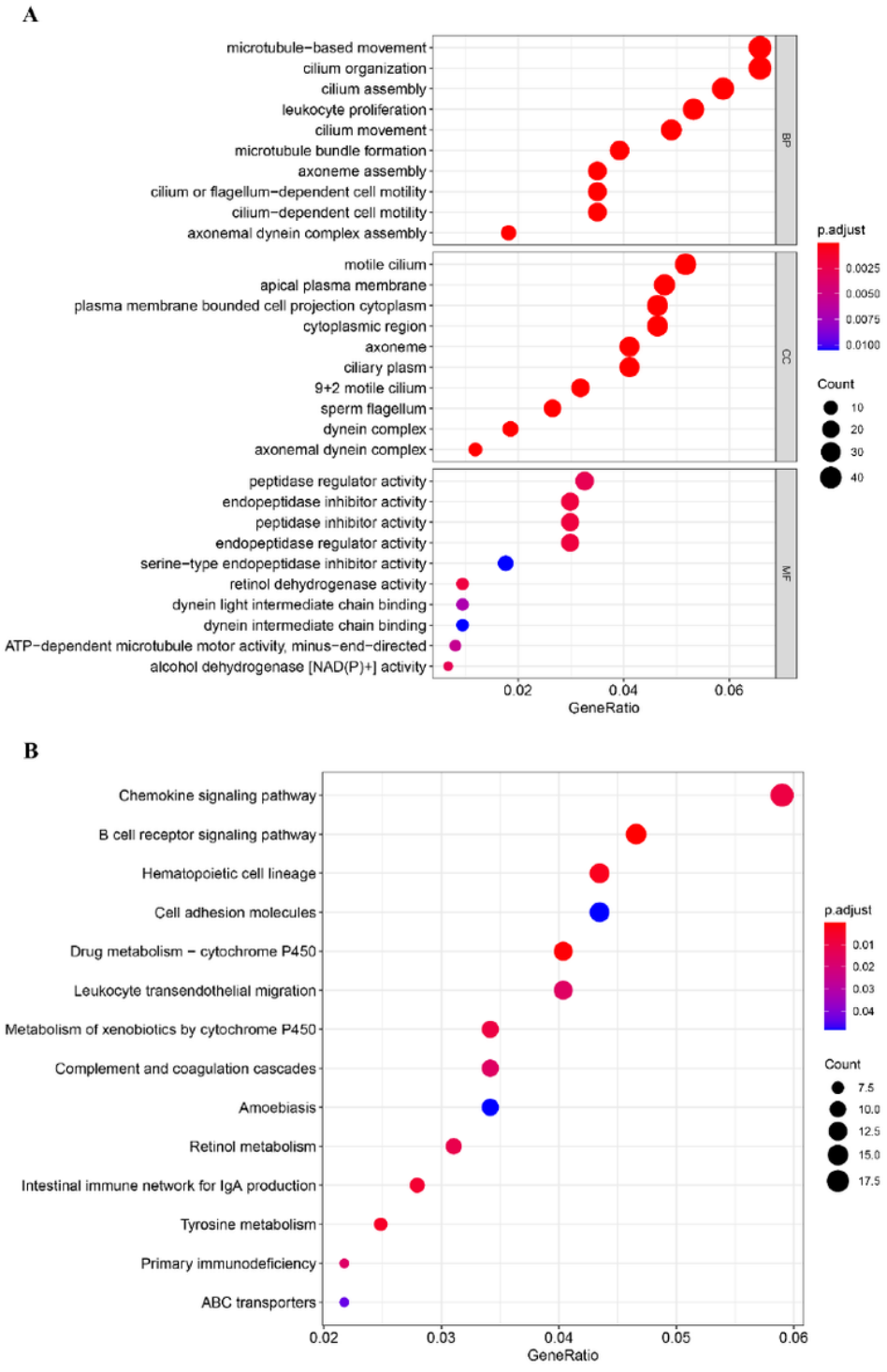

Figure 8

GO and KEGG analysis of intersection genes with down-regulated DEGs and ZIC2 negatively related to CEGs. (A) The bubble plots of BP, CC, and MF of GO annotation. (B) The bubble plots of KEGG annotation.

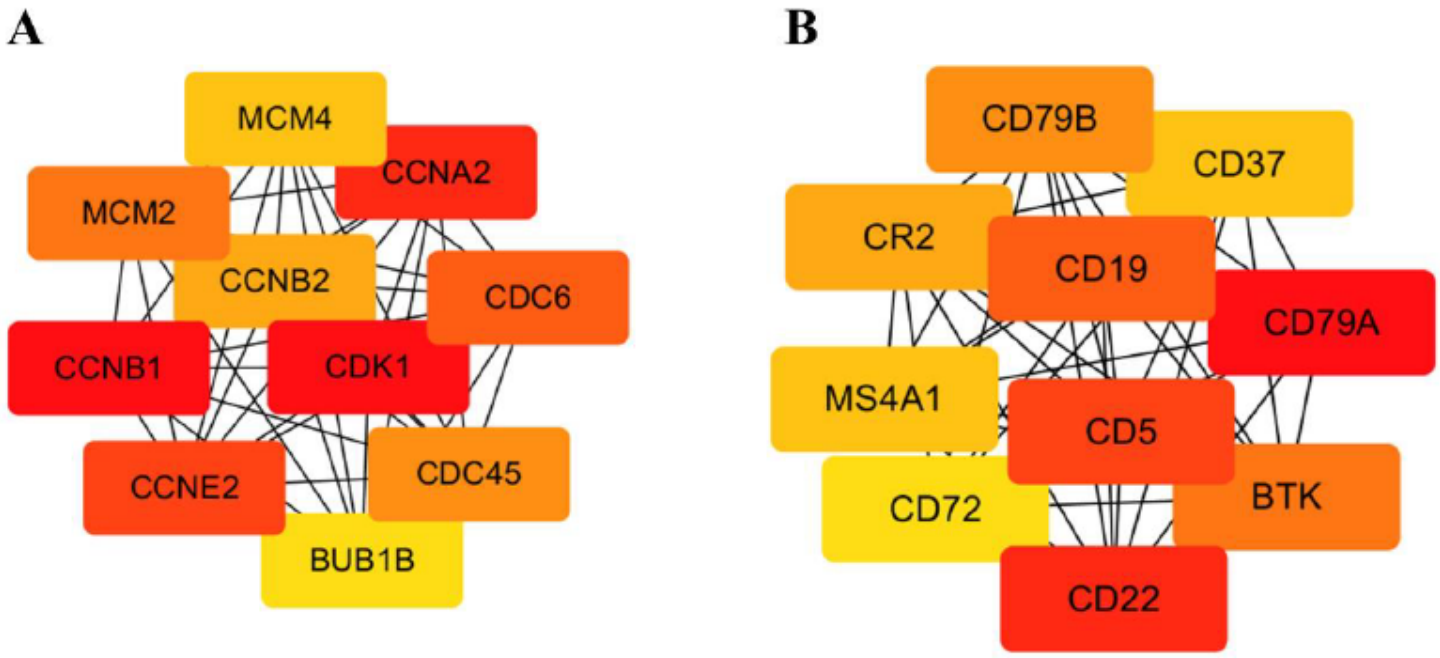

Figure 9

Protein-protein interaction network with ZIC2. (A) PPI network based on the overlapping top 10 hub genes of the first three KEGG pathways of up-regulated DEGs and ZIC2 positively correlated CEGs. (B) PPI network based on the overlapping top 10 hub genes of the first three KEGG pathways of down-regulated DEGs and ZIC2 negatively correlated CEGs.

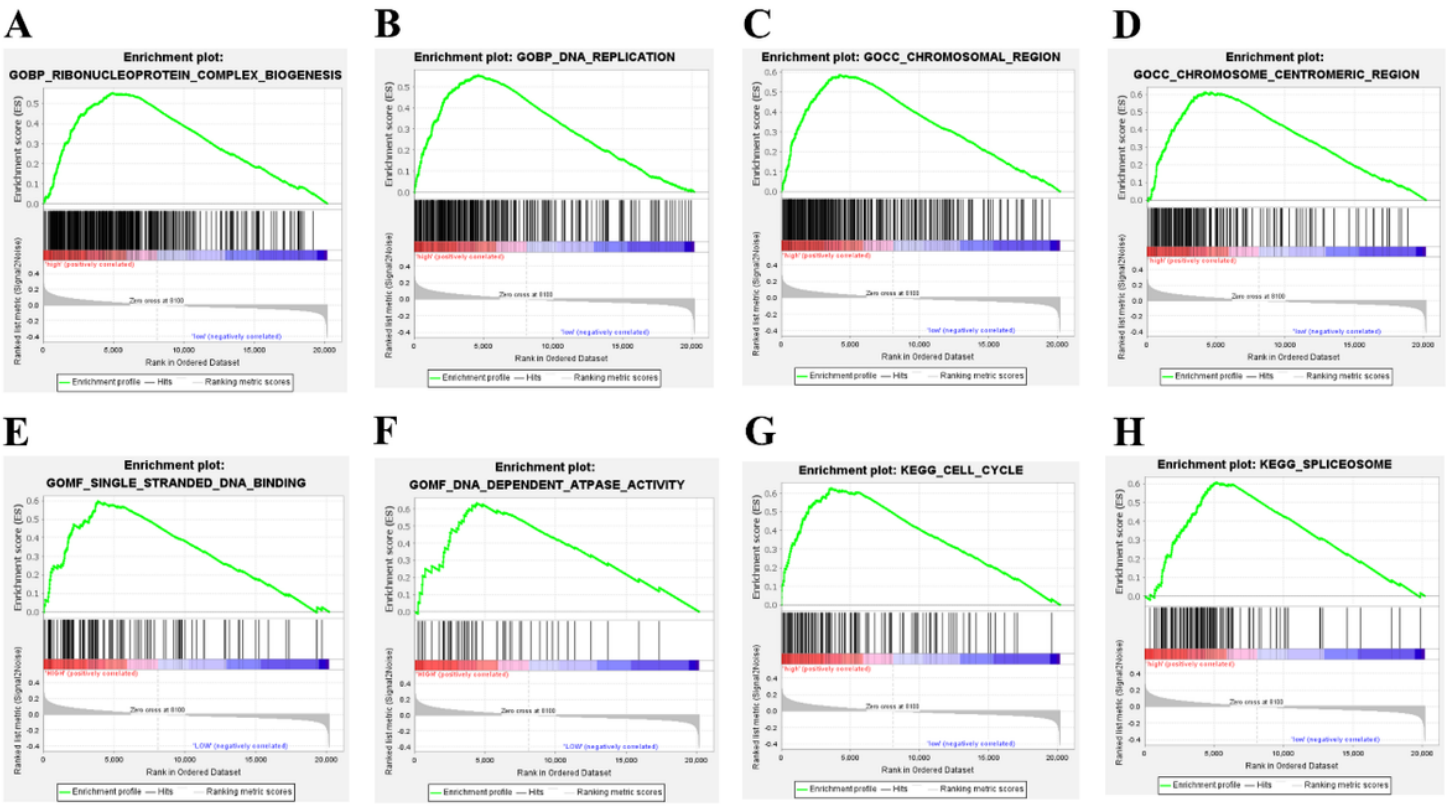

Figure 10

GSEA analysis. (A-F) The GSEA signal pathway enrichment analysis based on GO analysis. (G, H) The GSEA signal pathway enrichment analysis based on KEGG analysis.

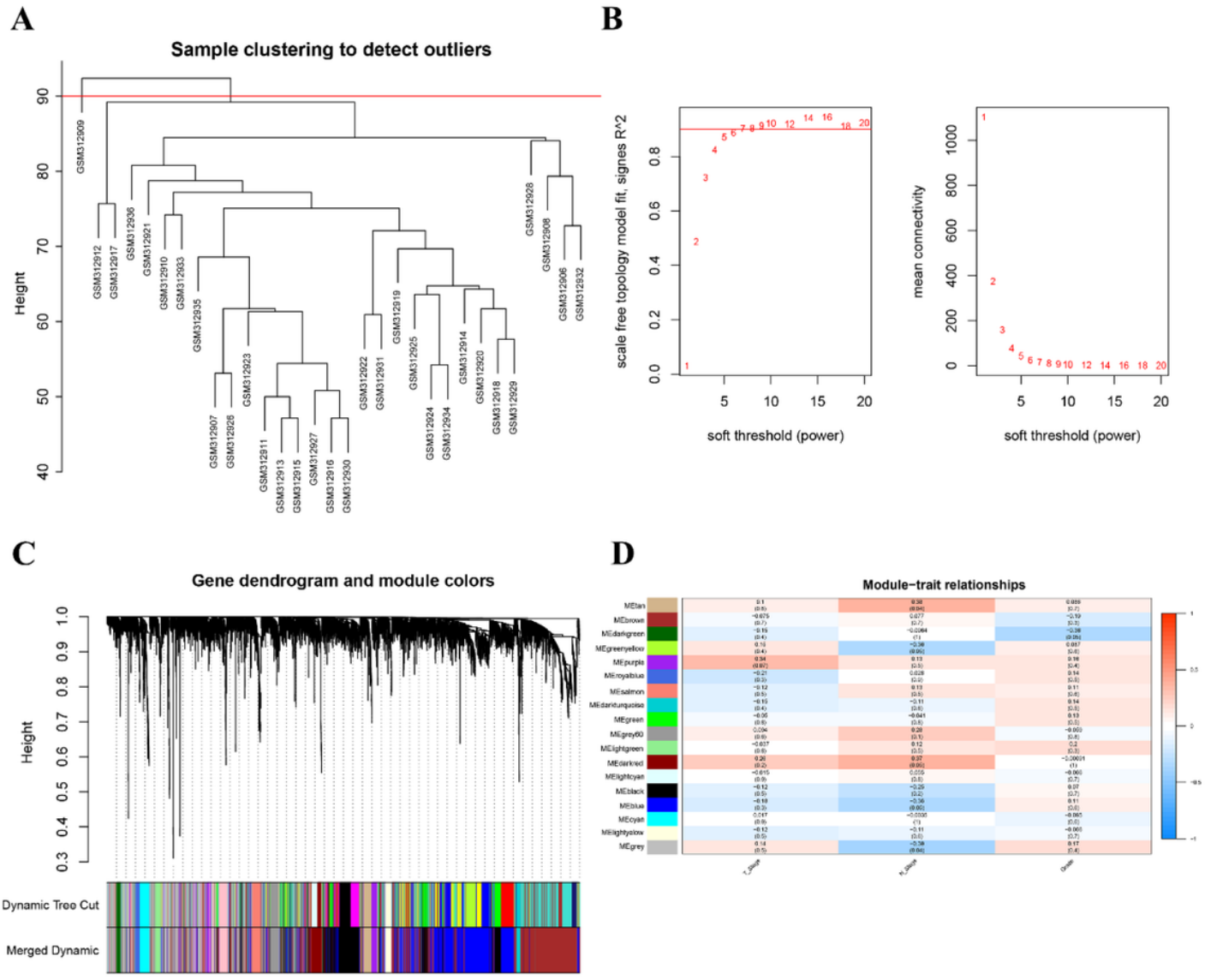

Figure 11

The WGCNA results of cancer samples in GSE12452. (A) The sample clustering tree reflects 30 samples of GSE12452 with a deviated sample of GSM312909. Outlier samples above the red line are excluded. (B) Analysis of scale-free topology (fitting degree R2) and mean connectivity for various soft-threshold values. (C) Module clustering dendrogram. Each branch with same color represents the same co-expression module. (D) Sample dendrogram and trait heatmap. The correlation between clinical traits (T_stage, N_stage and grade) with each module.