Clin. Lab. 2026;72:XXX-XXX

ORIGINAL ARTICLE

Differential Expression of Glucocorticoid-Related Genes and Immune Cell Infiltration in Nasopharyngeal Carcinoma

BiaoLi Long [1], GuoHua Hu [2]

[1] *Department of Otolaryngology, The Affiliated Dazu's Hospital of Chongqing Medical University, Chongqing, China*
[2] *Department of Otorhinolaryngology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China*

SUMMARY

Background: **Effective therapies are lacking for nasopharyngeal carcinoma (NPC), a malignancy with high incidence and frequent recurrence in Southeast Asia. Therefore, novel approaches for effective NPC treatment developed based on a deeper understanding of the molecular mechanisms underlying its pathogenesis are urgently needed. To this end, this study investigated glucocorticoid-related genes associated with NPC.**
Methods: **In this study, bioinformatic analyses were performed using datasets from the Gene Expression Omnibus, and 780 differentially expressed genes (DEGs) were identified, including 280 upregulated and 500 downregulated genes.**
Results: **Among the results of the bioinformatic analyses, enrichment analysis of these glucocorticoid-related DEGs revealed their strong associations with critical biological processes, including monocyte chemotaxis, lymphocyte regulation, and the IL-17 signaling pathway. Meanwhile, nine core genes of interest were identified by protein-protein interaction network analysis. Immune cell infiltration analysis illustrated variations in the abundance of immune cell types within the tumor microenvironment in NPC.**
Conclusions: **This study revealed the important role of glucocorticoid-related genes in promoting NPC progression. Our findings provide a robust foundation for developing targeted therapeutic strategies for patients with NPC based on innovative biomarkers.**
(Clin. Lab. 2026;72:xx-xx. DOI: 10.7754/Clin.Lab.2025.250529)

Correspondence:
GuoHua Hu
Department of Otorhinolaryngology
The First Affiliated Hospital of
Chongqing Medical University
Youyi Road, Yuzhong District
Chongqing, 400016
China
Email: 2h4841@hospital.cqmu.edu.cn

KEYWORDS

nasopharyngeal carcinoma, glucocorticoid-related genes, bioinformatics, differential expression, immune infiltration, Gene Expression Omnibus datasets, protein-protein interaction network

LIST OF ABBREVIATIONS

NPC - Nasopharyngeal carcinoma
DEGs - Differentially expressed genes
GRDEGs - Glucocorticoid-related differentially expressed genes
GRGs - Glucocorticoid-related genes
PPI - Protein-protein interaction
GEO - Gene Expression Omnibus
EPC - Edge percolated component
MNC - Maximum neighborhood component
BP - Biological process

Manuscript accepted July 4, 2025

CC - Cell component
MF - Molecular function

INTRODUCTION

Nasopharyngeal carcinoma (NPC) originates in the nasopharynx. It has a high incidence, particularly in Southeast Asia, threatening the health of individuals and straining financial resources. Epidemiological research has indicated that the annual incidence of NPC in endemic areas exceeds 30 cases per 100,000 people. Because the early symptoms of NPC are not easily identified, most patients with NPC are diagnosed at later stages, resulting in ineffective treatment and lower survival rates [1].

Effective therapies for NPC, which is difficult to cure, are lacking because of NPC´s aggressiveness and recurrence. Therefore, the molecular mechanisms underlying the pathogenesis of NPC and novel biomarkers need to be identified to support the development of novel treatments.

In a previous investigation, the key regulatory pathways in NPC, particularly those regulating the immune cell infiltration and interactions with the tumor microenvironment, were discovered through bioinformatic analysis [2]. Moreover, the interaction between glucocorticoids and immune modulation, especially the crucial role of glucocorticoid-related differentially expressed genes (GRDEGs) in the regulation of NPC-associated immune evasion, remains largely unknown [3].

To this end, this study focused on the molecular mechanisms of glucocorticoid-related genes associated with NPC progression. In this study, several bioinformatic tools were used to analyze GRDEGs in NPC, including differential expression analysis, enrichment analysis, and protein-protein interaction (PPI) networks, and datasets from the Gene Expression Omnibus (GEO) were analyzed. Through our findings, the biological importance of the identified GRDEGs in NPC progression was uncovered; thereby they might be recognized as potential novel therapeutic targets for improving clinical outcomes [4,5].

MATERIALS AND METHODS

Data resource

We retrieved data on human nasopharyngeal tissues from the GSE12452 [6] and GSE53819 [7] datasets using the GEOquery package (version 2.70.0) in R software [8,9]. The information included in the GPL570 and GPL6480 platforms within the GEO datasets is presented in Table 1. The GSE12452 dataset included 31 cases of NPC and 10 controls, whereas the GSE53819 dataset included 18 cases of NPC and 18 controls. All of these cases were incorporated into this study.

Table 1. Information on the GEO microarray chip.

	GSE12452	GSE53819
Platform	**GPL570**	**GPL6480**
Species	***Homo sapiens***	***Homo sapiens***
Tissue	**Nasopharyngeal tissue**	**Nasopharyngeal tissue**
Samples in the NPC group	**31**	**18**
Samples in the control group	**10**	**18**
Reference	**PMID: 17119049**	**PMID: 24763226**

The GeneCards database provides comprehensive information about the human genome [10], specifically those related to glucocorticoid-related genes (GRGs). In this research, we searched for "glucocorticoid" and only kept the protein-coding genes with relevance score > 4 as GRGs. In total, 492 GRGs were obtained. Additionally, by using "glucocorticoid" as a keyword in a PubMed search [11], 532 GRGs were obtained after removing duplicates. The GSE12452 and GSE53819 datasets, comprising 49 NPC and 28 normal nasopharyngeal tissue samples, were obtained from the GEO database and processed using the R package sva (version 3.50.0) [12]. Moreover, the R package limma [13] (version 3.58.1) was used to integrate the GEO datasets (combined dataset) and perform operations such as probe standardization and annotation.

Principal component analysis (PCA) [14], a method for reducing data dimensionality, was performed to examine the expression matrices both before and after this process to evaluate the efficacy of batch effect removal. Through this method, the feature vectors (components) from high-dimensional data were extracted and subsequently used to transform the data into a low-dimensional form. This enables these features to be presented in 2D or 3D graphs.

GRDEGs in NPC

Samples from the combined GEO datasets were categorized into NPC and control groups. Genetic difference analysis between these two groups was performed using the R package limma (version 3.58.1). Differentially expressed genes (DEGs) were selected using $|logFC| > 1$ and adj. $p < 0.05$ as the criteria. If $|logFC|$ was greater than 1 and adj. p was lower than 0.05, the genes were identified as upregulated genes. Genes with $|logFC| < -1$ and adj. $p < 0.05$ were regarded as downregulated genes. The Benjamini-Hochberg procedure was utilized to calculate the false discovery rate (FDR) and adj. p. The analysis results of DEGs were visualized via volcano plots created by the R package ggplot2 (version 3.4.4). To select the NPC-associated GRDEGs, in the variance analysis, all genes in the combined GEO datasets that satisfied $|logFC| > 1$ and adj. $p < 0.05$ were obtained.

All of the selected GRDEGs were intersected; their Venn diagram and heatmap were created using the pheatmap package (version 1.0.12) in R.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses

GO [15] and KEGG analyses [16] were performed in our study. In the analysis of GRDEGs, their biological term classifications and the enrichment of gene clusters were performed using the R package clusterProfiler (version 4.10.0) [17]. The criteria for entry screening were adj. $p < 0.05$ and FDR (q-value) < 0.25. The Benjamini-Hochberg procedure was used to calculate the FDR and adj. p.

Gene set enrichment analysis (GSEA)

GSEA, a computational method integrated in clusterProfiler, was used in our study to interpret gene expression data [18]. In GSEA, genes are first ranked according to their differential expression between the two biological states, followed by calculation of the enrichment scores for predefined gene sets. Their statistical significance was evaluated by permutation testing to identify biologically relevant pathways or functions.

Based on their logFC values, DEGs in the combined GEO dataset were identified, followed by calculation of the enrichment scores for the predefined gene sets to determine their statistical significance. Molecular Signatures Database, especially the gene set "c2.cp.all.v2022.1.hs.symbols.gmt [all canonical pathways](3050)," was used for GSEA with the following parameters: random seed, 2022; number of calculations, 1,000; and each gene set contained a minimum of 10 genes and a maximum of 500 genes [19]. The screening criteria for GSEA were adj. $p < 0.05$ and q-value < 0.25. The Benjamini-Hochberg method was employed to calculated adj. p.

Exploring PPI networks and hub genes

PPI networks are schematic representations of the physical associations among proteins within a cell. STRING serves as both a knowledgebase and a software tool that can be used to search for known PPIs and to predict potential ones [20].

In our study, a PPI network related to GRDEGs was constructed using the STRING database following minimum interaction coefficient > 0.9 or minimum required interaction score for low confidence = 0.9 as the criterion.

Additionally, five algorithms, namely Closeness, Edge Percolated Component (EPC), Maximum Neighborhood Component (MNC), Degree, and Maximal Clique Centrality (MCC), integrated in the CytoHubba plugin of Cytoscape software [21,22] were used to calculate the scores of GRDEGs for constructing the relevant PPI network. Moreover, the top 10 GRDEGs were selected by their scores. Finally, the selected GRDEGs were intersected, and their Venn diagram was created. The genes obtained from the intersection were regarded as glucocorticoid-related hub genes.

Construction of mRNA–miRNA and mRNA-TF regulatory networks

The regulation of gene expression, either target genes or regulatory RNA (ranging from mRNA to non-coding RNAs such as miRNA), is essential for biological development and evolution. In this study, the starBase database was used to explore the interaction between miRNAs and target genes to understand the mechanism by which miRNA mediates the regulation of glucocorticoid-related hub genes [23]. Through this method, an mRNA-miRNA regulatory network was developed. The contents of this network were visualized using Cytoscape software.

Another comprehensive database (hTFtarget [24]) related to the regulation of human TF and their target genes (mRNAs) was utilized to investigate the TFs associated with glucocorticoid-related hub genes. The map of the mRNA–TF regulatory network was developed using Cytoscape software.

Differential expression of hub genes and their receiver operating characteristic (ROC) curves

To compare the differences in glucocorticoid-related hub gene expression between the NPC and control groups, the R package pROC was used to plot the ROC curves of these hub genes, after which the area under the ROC curve (AUC) was calculated to evaluate the diagnostic effect of the hub genes on the occurrence of NPC. The AUC ranged from 0.5 to 1, and the diagnostic accuracy based on the AUC was categorized as follows: 0.5 - 0.7, low accuracy; 0.7 - 0.9, moderate accuracy, and > 0.9, high accuracy.

Immune infiltration analysis

In this study, the composition and abundance of immune cells within a cell mixture were estimated using CIBERSORT25, a tool to deconvolute the transcriptome expression matrix based on linear support vector regression [25]. Immune infiltration analysis was performed using CIBERSORT algorithm along with the LM22 feature gene matrix. After removing the data with an immune cell enrichment score greater than zero, the immune cell infiltration matrix of the combined dataset was obtained.

By comparing the differences in the level of immune cell infiltration between the NPC and control groups, the correlations among different immune cells and the correlations between hub genes and immune cells were determined by R software. The pheatmap package was used to generate the correlation heatmaps, in which $r < 0.3$ implied no correlation, $r = 0.3 - 0.5$ indicated a weak correlation, $r = 0.5 - 0.8$ represented a moderate correlation, and $r > 0.8$ denoted a strong correlation.

Statistical analysis

In this research, R software (version 4.2.2) was utilized for data processing and analysis. For continuous variables, the data are presented as the mean ± SD. The Wilcoxon rank-sum test was employed to compare the differences between the NPC and control groups. Unless otherwise specified, Spearman's correlation analysis was used to compute the correlation coefficients between different molecules, and the full results were considered significant when $p < 0.05$.

RESULTS

Merging two NPC datasets into a combined GEO dataset

With the R package sva, we merged data of GSE12452 and GSE53819 into a combined GEO dataset. These two datasets within the GEO database are associated with NPC progression. Before merging, the batch effect was eliminated through data filtering. As presented in Figure 1A and B, gene expression pattern in the datasets differed between before and after batch effect removal. The distribution of low-dimensional features in the datasets before and after batch effect removal was also compared using PCA. The PCA plots are presented in Figure 1C and D. Both the distribution boxplots and PCA plots demonstrate the effectiveness of batch effect removal.

Identification of GRDEGs in NPC

The genetic differences in the combined GEO dataset, including NPC and control groups, were evaluated using limma. A total of 780 DEGs in this combined dataset were identified according to the criteria of $|logFC| > 1$ and adj. $p < 0.05$. In the variance analysis, 280 candidate DEGs were identified using $|logFC| > 1$ and adj. $p < 0.05$. Eventually, 26 GRDEGs were obtained (for details, see Table S2). The volcano plots generated by ggplot2 (version 3.4.4) are presented in Figure 2A. The FDR and adj. p were calculated using the Benjamini-Hochberg procedure. All of the selected GRDEGs were intersected, and their Venn diagram (Figure 2B) and heatmap (Figure 2C) were created using the pheatmap package (version 1.0.12) in R.

GO and KEGG pathway enrichment analysis of GRDEGs

To evaluate the functions of the identified GRDEGs associated with NPC, GO and KEGG pathway enrichment analyses were performed. These analyses included functional evaluations in terms of biological process (BP), cell components (CC), and molecular function (MF). Among the 26 selected GRDEGs, the enriched BP terms were related to the regulation of steroid biosynthesis and lipid biosynthesis. The detailed results are presented in Table 2. The enriched CC terms were related to the regulation of components such as the endoplasmic reticulum lumen, organelle outer membrane, platelet alpha-granule lumen, and outer membrane. The enriched MF terms were related to activities such as cytokine, cytokine-receptor binding, receptor ligand, signaling receptor activator, chemokine, and the chemotaxis of monocytes and lymphocytes.

Furthermore, based on the results of KEGG analysis, the following aspects were also enriched: the IL-17 signaling pathway, cytokine-cytokine receptor interaction, rheumatoid arthritis, viral protein interaction with cytokine and cytokine receptor, and the AGE–RAGE signaling pathway in diabetic complications.

The results of GO and KEGG analyses are illustrated as bubble plots in Figure 3A, whereas the GO-demonstrated network maps in Figure 3B - E include BP, CC, MF, and the biological pathways. In these plots and maps, the lines represent the relevant molecules and the annotations of the corresponding entries. A larger node indicates a higher number of entries. Notably, the genes were mostly enriched in the activities of signaling receptor activators, receptor ligands, and cytokine.

GSEA

To assess the conformity of the combined GEO dataset, GSEA was performed. Figure 4A presents the relationship between the affected CCs and the involved MFs. The detailed results are presented in Table 3. Genes within the combined GEO dataset were significantly enriched in several biological functions and signaling pathways, such as transcriptional regulation by TP53, PI3K/Akt signaling pathway, functions and pathways related to the termination of O-glycan biosynthesis, and the drug metabolism cytochrome P450 (Figure 4B - E).

Construction of PPI networks and identification of hub genes

A PPI network associated with 26 GRDEGs was constructed using the STRING database (Figure 5A). The scores of relevant GRDEGs were calculated using five algorithms integrated in the CytoHubba plugin of Cytoscape software, including MCC (Figure 5B), MNC (Figure 5C), Degree (Figure 5D), and EPC (Figure 5E). Based on these scores, 13 GRDEGs were selected: *CCL4*, *IFNG*, *IL1A*, *CXCL10*, *CCL2*, *SERPINE1*, *BMP2*, *FN1*, *SPP1*, *MMP3*, *MMP1*, *SRD5A2*, and *AKR1C3.* Among these, the top 10 GRDEGs were further chosen. Nine genes obtained from this intersection were regarded as glucocorticoid-related hub genes in NPC: *CXCL10*, *IL1A*, *IFNG*, *CCL2*, *CCL4*, *FN1*, *SPP1*, *MMP3*, and *BMP2*.

Differential expressions of hub genes and their ROC curves

The expression differences of the aforementioned nine glucocorticoid-related hub genes between the NPC and control groups were clarified. The results indicated that the expression of all nine hub genes significantly differed between the two groups ($p < 0.001$, Figure 6A).

The ROC curves of these hub genes were plotted using pROC (Figure 6B - D), whereas the AUCs, which re-

Table 2. Results of GO and KEGG enrichment analyses of GRDEGs.

Ontology	ID	Description	Gene ratio	Bg ratio	p	adj. p	q-value
BP	GO:0002548	Monocyte chemotaxis	5/26	70/18,800	3.83e-08	2.31e-05	1.25e-05
BP	GO:1901623	Regulation of lymphocyte chemotaxis	4/26	26/18,800	4.21e-08	2.31e-05	1.25e-05
BP	GO:0050810	Regulation of steroid biosynthetic process	5/26	72/18,800	4.42e-08	2.31e-05	1.25e-05
BP	GO:0006694	Steroid biosynthetic process	6/26	175/18,800	1.18e-07	3.69e-05	1.99e-05
BP	GO:0046890	Regulation of lipid biosynthetic process	6/26	175/18,800	1.18e-07	3.69e-05	1.99e-05
CC	GO:0005788	Endoplasmic reticulum lumen	4/26	311/19,594	7.06e-04	3.96e-02	2.97e-02
CC	GO:0031968	Organelle outer membrane	3/26	232/19,594	3.48e-03	5.00e-02	3.76e-02
CC	GO:0031093	Platelet alpha granule lumen	2/26	67/19,594	3.55e-03	5.00e-02	3.76e-02
CC	GO:0019867	Outer membrane	3/26	234/19,594	3.57e-03	5.00e-02	3.76e-02
MF	GO:0005125	Cytokine activity	8/26	235/18,410	8.02e-10	1.34e-07	7.43e-08
MF	GO:0048018	Receptor ligand activity	8/26	489/18,410	2.40e-07	1.49e-05	8.26e-06
MF	GO:0030546	Signaling receptor activator activity	8/26	496/18,410	2.67e-07	1.49e-05	8.26e-06
MF	GO:0008009	Chemokine activity	4/26	49/18,410	6.34e-07	2.65e-05	1.47e-05
MF	GO:0005126	Cytokine receptor binding	6/26	272/18,410	1.77e-06	5.90e-05	3.28e-05
KEGG	hsa04657	IL-17 signaling pathway	7/25	94/8,164	8.69e-09	1.03e-06	7.68e-07
KEGG	hsa04060	Cytokine-cytokine receptor interaction	8/25	295/8,164	1.67e-06	9.95e-05	7.39e-05
KEGG	hsa05323	Rheumatoid arthritis	5/25	93/8,164	7.64e-06	3.03e-04	2.25e-04
KEGG	hsa04061	Viral protein interaction with cytokine and cytokine receptor	4/25	100/8,164	2.20e-04	5.07e-03	3.77e-03
KEGG	hsa04933	AGE-RAGE signaling pathway in diabetic complications	4/25	100/8,164	2.20e-04	5.07e-03	3.77e-03

flect the diagnostic accuracy, were calculated. The AUCs of several hub genes such as *IL1A*, *IFNG*, *CCL2*, *CCL4* (Figure 6E), *SPP1* (Figure 6F), and *BMP2* (Figure 6G) were lower than 0.9, indicating that these genes have low accuracy for NPC diagnosis.

However, *CXCL10*, *FN1*, and *MMP3* might be more valuable for NPC diagnosis because of their higher AUCs (AUC > 0.9).

Immune infiltration analysis using CIBERSORT

As presented in the bar chart in Figure 7A, the proportions of different cell types and profiles of gene expression were deconvoluted from bulk RNA sequencing data using CIBERSORT software. The abundance of 22 immune cell types between the NPC and control groups in the combined GEO dataset is also presented in Figure 7A.

A group comparison plot (Figure 7B) demonstrated the differences in the abundance of 22 infiltrating immune cells between the two groups. The results illustrated that mast cells could be activated by immune cells. Statistically significant differences in the abundance of two subpopulations of T cells (CD4 memory-activated T cells and CD8 T cells) were found between the two groups ($p < 0.05$, Figure 7A). Additionally, statistically significant differences were also observed in certain immune subpopulations between the two groups, such as M0 macrophages, resting mast cells, memory B cells, naive B cells, M1 macrophages, neutrophils, activated NK cells, and plasma cells, resting CD4 memory T cells, and T follicular helper cells (all $p < 0.001$). As presented in Figure 7C and D, the correlation heatmap generated using pheatmap indicates the correlations among different immune cells and between hub genes. In addition, the correlation heatmap presented in Figure 7C demonstrates the abundance of 13 infiltrating immune cell subpopulations, among which memory B cells were positively correlated with resting mast cells ($r = 0.44$), whereas resting mast cells displayed the strongest negative correlation with activated mast cells

Table 3. Results of GSEA of the combined dataset.

ID	Set size	Enrichment score	NES	p	adj. p	q-value
REACTOME_TRANSCRIPTIONAL_ REGULATION_BY_TP53	329	0.46	2.03	1.00e-10	9.87e-09	7.91e-09
REACTOME_TRANSCRIPTIONAL_ REGULATION_BY_TP53	329	0.46	2.03	1.00e-10	9.87e-09	7.91e-09
REACTOME_TERMINATION_OF_O_ GLYCAN_BIOSYNTHESIS	21	0.77	2.04	9.36e-05	1.60e-03	1.29e-03
KEGG_DRUG_METABOLISM_ CYTOCHROME_P450	57	0.68	2.18	3.38e-07	1.46e-05	1.17e-05
REACTOME_COLLAGEN_DEGRADATION	61	0.77	2.68	1.00e-10	9.87e-09	7.91e-09
REACTOME_CELL_CYCLE_CHECKPOINTS	236	0.63	2.66	1.00e-10	9.87e-09	7.91e-09
REACTOME_ASSEMBLY_OF_COLLAGEN_FI BRILS_AND_OTHER_MULTIMERIC_ STRUCTURES	59	0.77	2.66	1.00e-10	9.87e-09	7.91e-09
REACTOME_COLLAGEN_FORMATION	80	0.73	2.66	1.00e-10	9.87e-09	7.91e-09
NABA_COLLAGENS	42	0.80	2.59	1.00e-10	9.87e-09	7.91e-09
REACTOME_COLLAGEN_CHAIN_ TRIMERIZATION	42	0.80	2.59	1.00e-10	9.87e-09	7.91e-09
WP_RETINOBLASTOMA_GENE_IN_CANCER	86	0.69	2.58	1.00e-10	9.87e-09	7.91e-09
PID_SYNDECAN_1_PATHWAY	46	0.78	2.57	1.00e-10	9.87e-09	7.91e-09
REACTOME_COLLAGEN_BIOSYNTHESIS_ AND_MODIFYING_ENZYMES	59	0.73	2.54	1.00e-10	9.87e-09	7.91e-09
REACTOME_G2_M_CHECKPOINTS	123	0.64	2.51	1.00e-10	9.87e-09	7.91e-09
PID_INTEGRIN1_PATHWAY	65	0.71	2.50	1.00e-10	9.87e-09	7.91e-09
REACTOME_CELL_CYCLE_MITOTIC	464	0.55	2.48	1.00e-10	9.87e-09	7.91e-09
REACTOME_NON_INTEGRIN_MEMBRANE_ ECM_INTERACTIONS	56	0.72	2.47	1.61e-10	1.53e-08	1.23e-08
REACTOME_MITOTIC_SPINDLE_ CHECKPOINT	101	0.64	2.43	1.00e-10	9.87e-09	7.91e-09
REACTOME_ACTIVATION_OF_ATR_IN_ RESPONSE_TO_REPLICATION_STRESS	36	0.76	2.42	4.81e-09	3.21e-07	2.57e-07
REACTOME_ACTIVATION_OF_THE_PRE_ REPLICATIVE_COMPLEX	33	0.78	2.39	5.40e-09	3.50e-07	2.81e-07

(r = -0.75).

Finally, correlation bubble plots were used to illustrate the correlations between hub genes and the abundance of infiltrating immune cells (Figure 7D). The results indicated that the *CXCL10* gene had the strongest positive correlation with M1 macrophages (r = 0.75), whereas the *FN1* gene had the greatest negative correlation with memory B cells (r = -0.70).

DISCUSSION

NPC is a type of head and neck malignancy arising from the epithelial tissue of the nasopharynx. Per data from the US National Cancer Institute, fewer than one case of NPC is diagnosed per 100,000 people globally each year. Nevertheless, NPC is endemic in certain regions such as Southeast Asia, North Africa, and the Arctic, where higher incidence rates are reported [25]. Because of its subtle early symptoms, NPC is often diagnosed at advanced stages, resulting in poor treatment outcomes and low survival rates. Current therapeutic methods including radiotherapy, chemotherapy, and surgery are constrained by the aggressive nature and high recurrence rates of NPC. Hence, there is an urgent need for novel biomarkers and therapeutic targets [4,10].

Through sequencing technologies and microarrays, the gene expression status can be accurately evaluated, thereby allowing exploration of the molecular mechanisms underlying the initiation and development of NPC and improving the diagnosis, treatment, and prognosis of this cancer.

Because reports have examined the impact of GRGs on

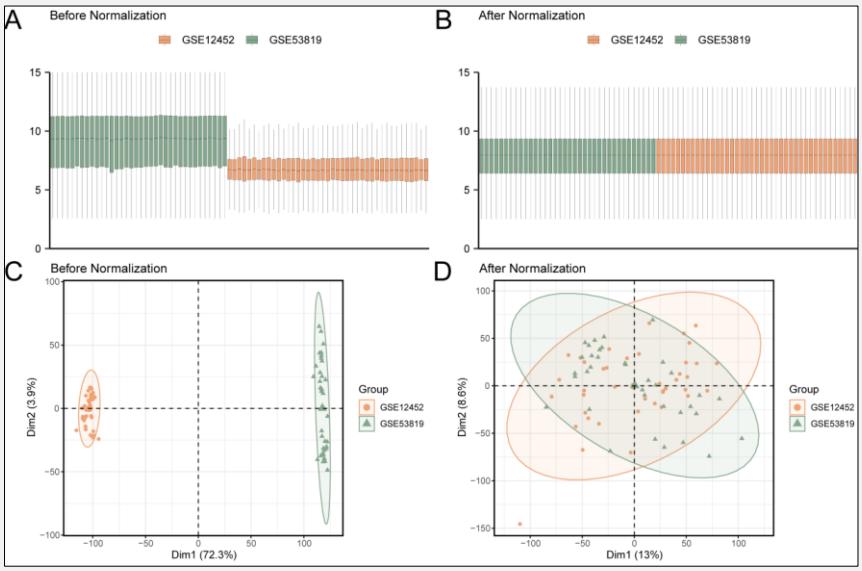

Figure 1. Batch effect removal in the GSE12452 and GSE53819 datasets.

A Boxplot of the combined GEO dataset distribution before batch removal. B Post-batch integrated GEO dataset (combined dataset) distribution boxplots. C PCA plot of the datasets before batch removal. D The PCA map of the combined GEO dataset after batch processing. The GSE12452 dataset is denoted in orange, and the GSE53819 dataset is presented in green.

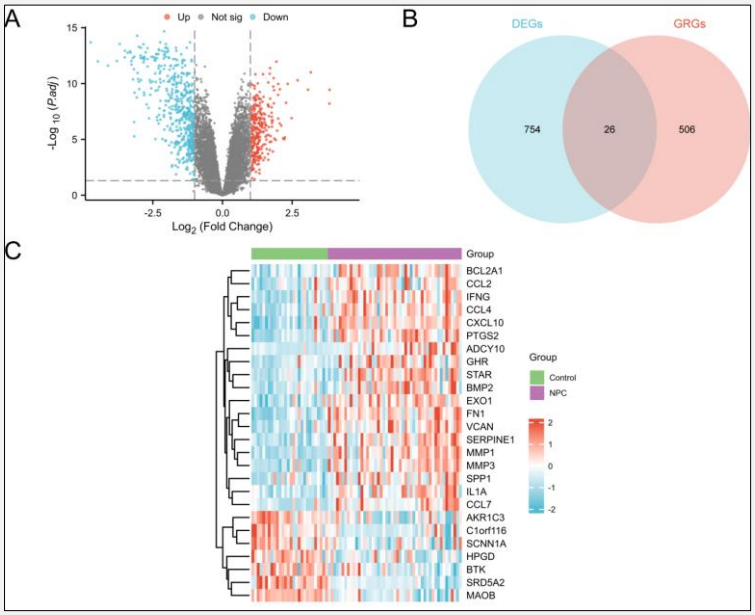

Figure 2. Differential gene expression analysis.

A Integrated GEO dataset (combined datasets) in the volcanic diagram analysis of DEGs in the control and NPC groups. B Venn diagram of DEGs and GRGs in the integrated GEO datasets (combined datasets). C Heatmap of GRDEGs in the integrated GEO datasets (combined datasets). Green denotes the control group, and purple denotes the NPC group. In the heatmap, red represents high expression, and blue represents low expression.

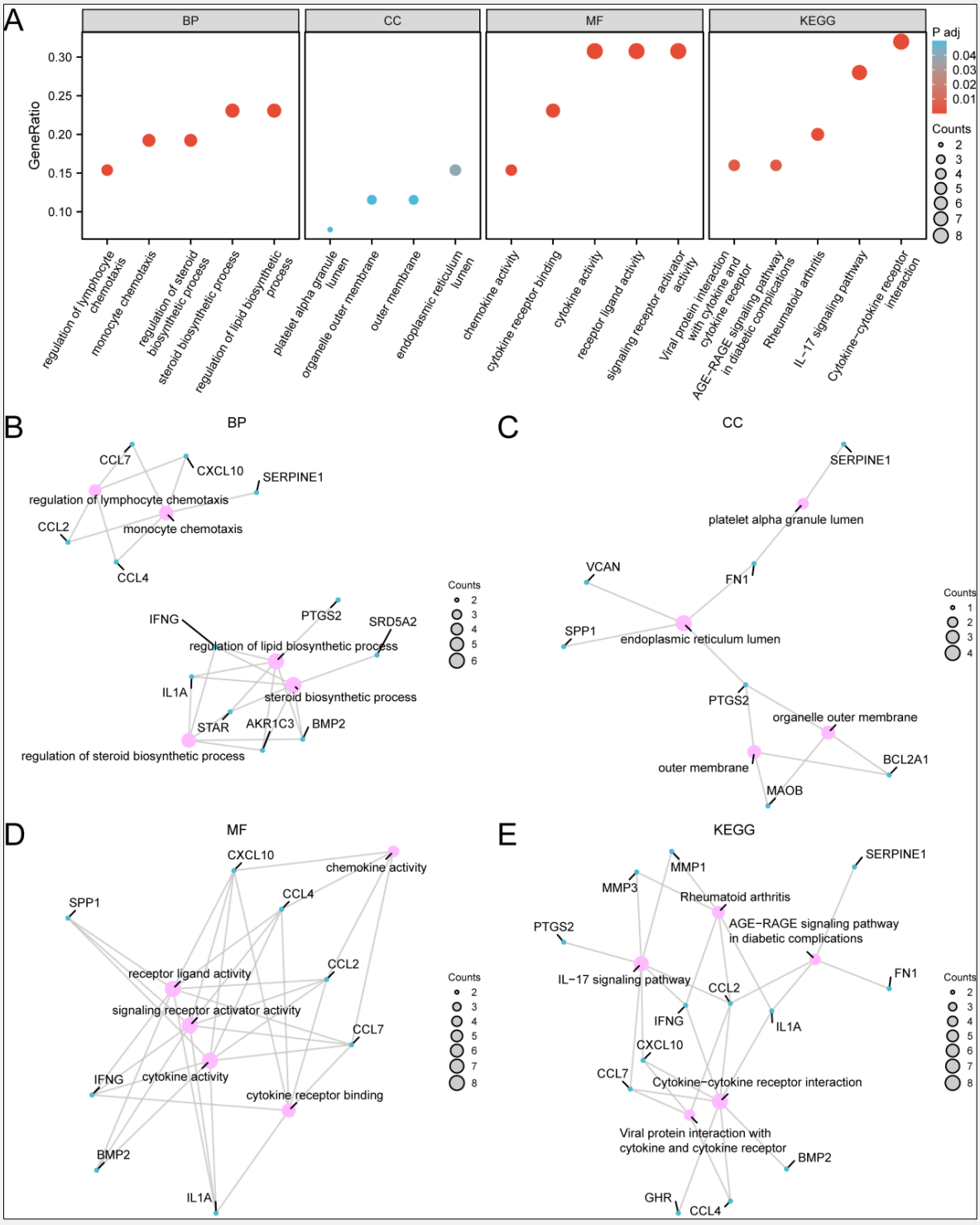

Figure 3. GO and KEGG enrichment analyses of GRDEGs.

A Bubble plot of the results of GO and KEGG enrichment analysis of GRDEGs. The abscissa presents GO terms and KEGG terms.
B - E Network diagrams presenting the results of GO and KEGG enrichment analyses of GRDEGs for BP (B), CC (C), MF (D), and KEGG pathways (E). Pink nodes represent items, blue nodes represent molecules, and lines represent the relationship between items and molecules. The bubble size in the bubble plot stands for number of genes, and various bubble color stands for the magnitude of adj. p (red, low; blue, high). The screening criteria for GO and pathway (KEGG) enrichment analysis were adj. p < 0.05 and q-value < 0.25, and the p-value was corrected using the Benjamini-Hochberg procedure.

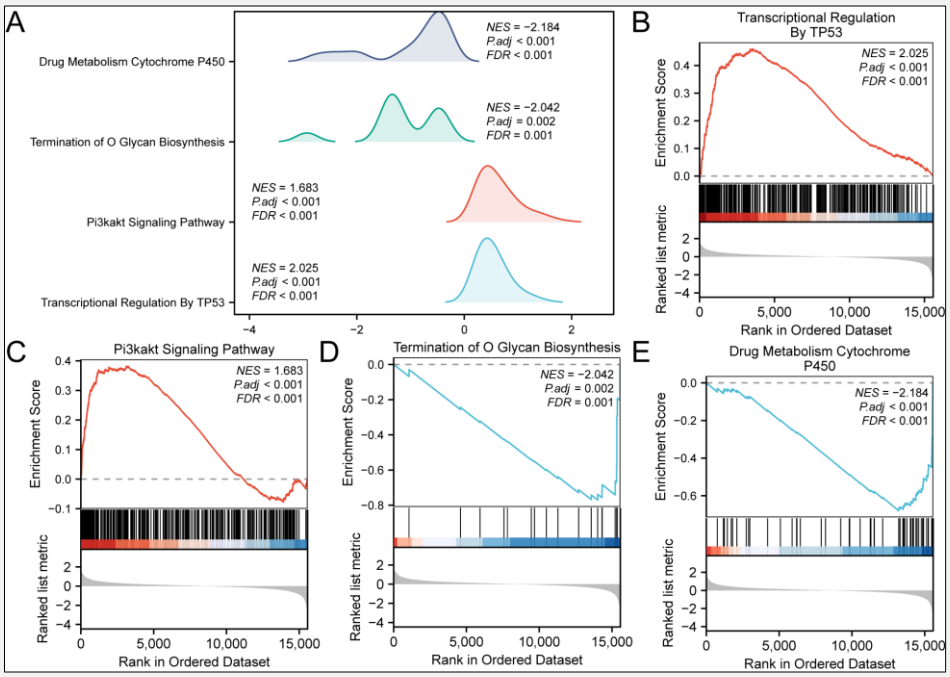

Figure 4. GSEA for the combined GEO dataset.

A GSEA presents four biological functions as mountain maps for the combined GEO dataset.
B - E GSEA revealed that all genes were significantly enriched in transcriptional regulation by TP53 (B), the PI3K/Akt signaling pathway (C), the termination of O-glycan biosynthesis (D), and the drug metabolism enzyme cytochrome P450 (E). The screening criteria for GSEA were adj. p < 0.05 and q-value < 0.25, and the p-value was corrected using the Benjamini-Hochberg procedure.

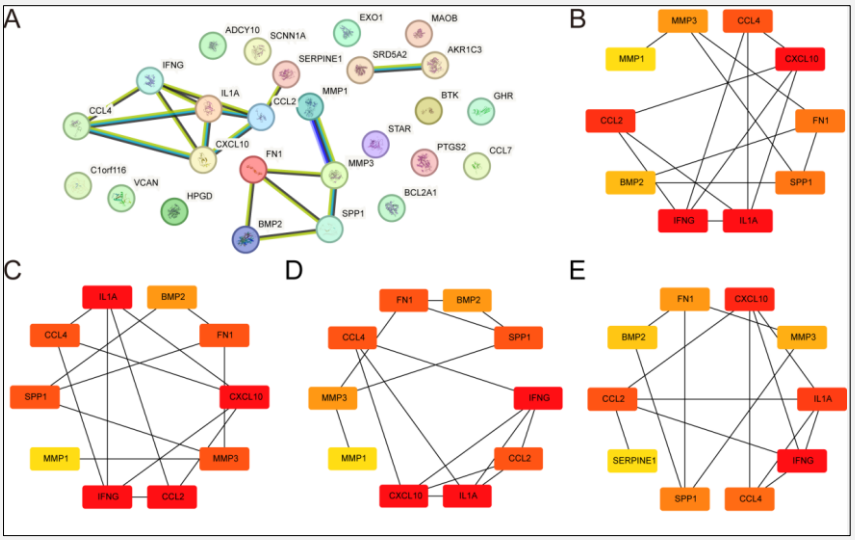

Figure 5. Construction of the PPI network and hub gene analysis.

A PPI network of GRDEGs calculated using the STRING database, only interacting proteins are displayed. B - F PPI network for the top 10 GRDEGs calculated using algorithms included in the CytoHubba plug-in, including MCC (B), MNC (C), Degree (D), EPC (E).

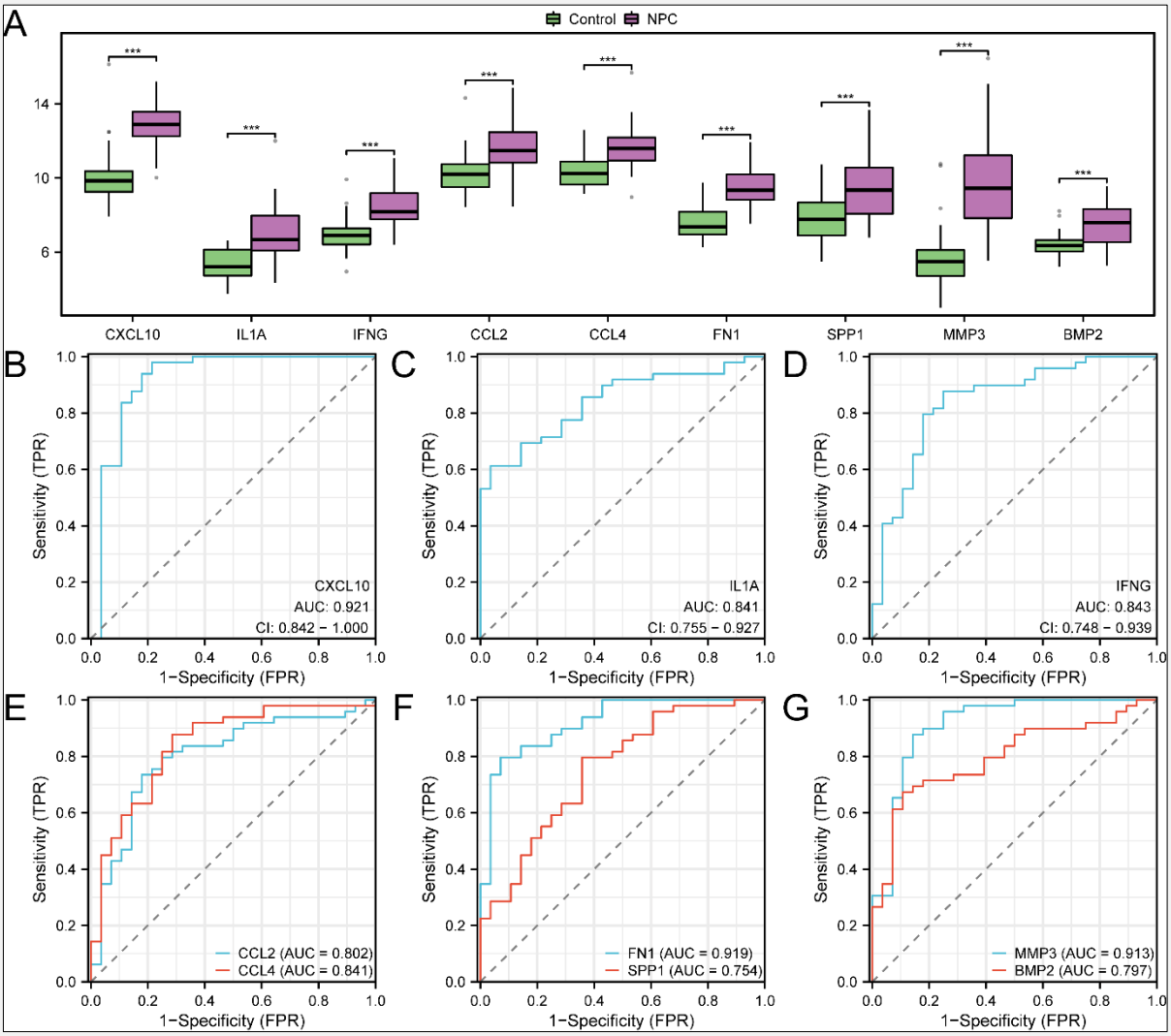

Figure 6. Differential expression validation and ROC curve analysis.

A Group comparison diagram of hub genes between the NPC and control groups. B - G ROC curves of hub genes, including *CXCL10* (B), *IL1A* (C), *IFNG* (D), *CCL2*/*CCL4* (E), *FN1*/*SPP1* (F), and *MMP3*/*BMP2* (G), in the combined GEO dataset. *** p < 0.001. AUC > 0.5 indicates that the expression of the molecule tends to promote the occurrence of the event. The diagnostic effect increases as the AUC approaches 1. An AUC of 0.7 - 0.9 indicates moderate accuracy, and an AUC exceeding 0.9 indicates high accuracy. TPR true positive rate, FPR false positive rate. Green represents the control group, and purple represents the NPC group.

the development of NPC, this study explored GRDEGs related to the pathogenesis of NPC. Bioinformatic tools were used in this study to retrieve data from a combined GEO dataset, identify the GRDEGs related to the pathogenesis of NPC, and investigate the substantial roles of these genes in immune regulation and their interactions within the tumor microenvironment. Through this approach, our understanding of NPC biology was improved, which could facilitate the development of therapeutic approaches. The bioinformatics analyses in this study included differential expression, pathway enrichment, and PPI network analyses, all of which demonstrated the potential role of GRDEGs in NPC progression, providing avenues for future research to improve early diagnosis and treatment strategies [26,27].

In our study, 780 DEGs were identified, including 280 upregulated and 500 downregulated genes. Some of these genes might emerge as biomarkers and therapeutic targets for NPC treatment according to their biological functions in NPC development. The identified DEGs

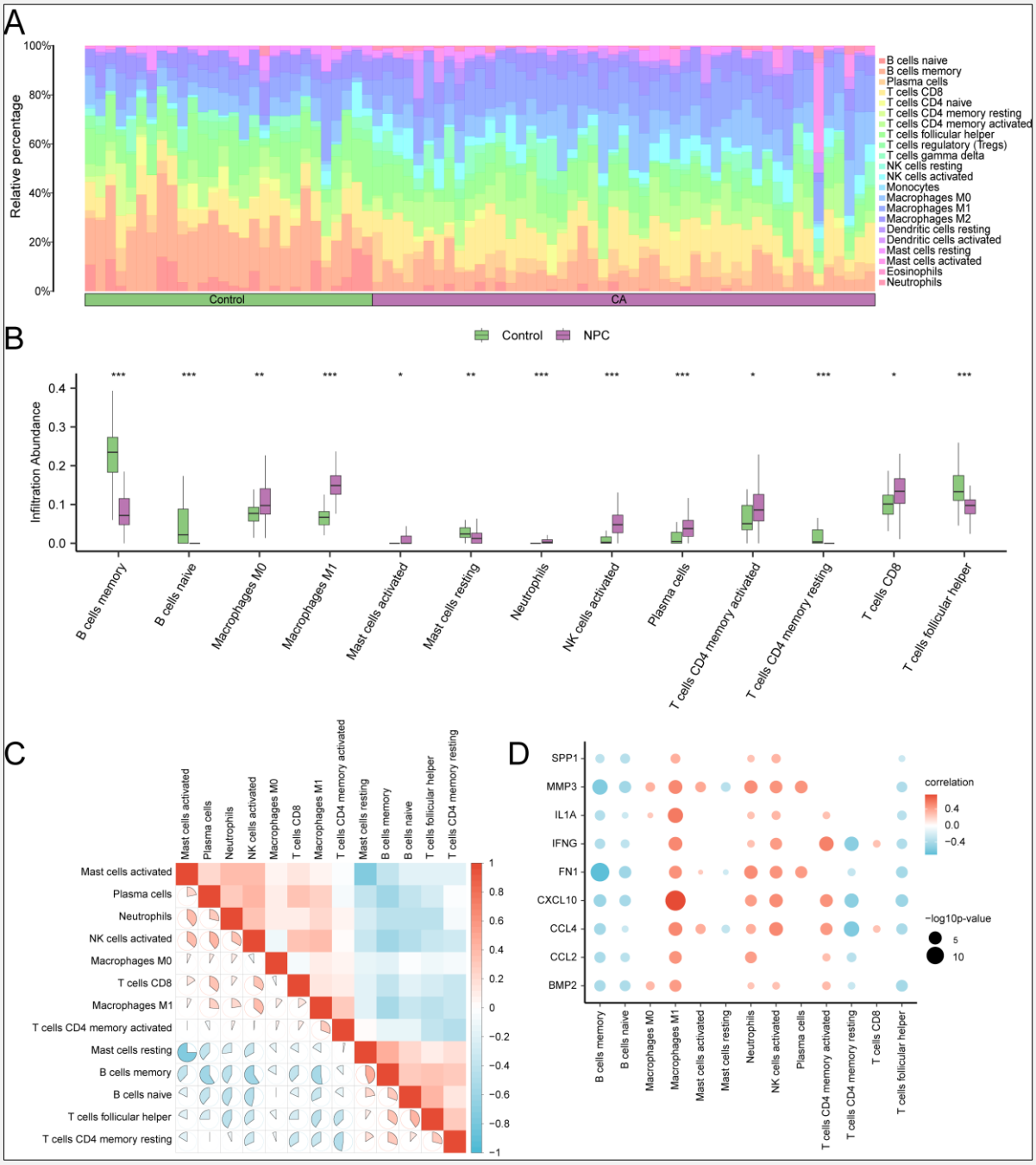

Figure 7. Immune infiltration analysis using the single-sample gene set enrichment analysis algorithm.

A - B Bar graph (A) and group comparison graph (B) presenting the proportions of immune cells in the NPC and control groups. Green control group, purple NPC group. C Heatmap (in the group comparison plots) presenting the correlation between the abundance of infiltrating immune cells that significantly differed between the groups in the combined GEO dataset. D Heatmap of the correlation of glucocorticoid-related hub genes with the abundance of seven infiltrating immune cells in the combined GEO dataset. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. $r < 0.3$ indicates no correlation, $r = 0.3 - 0.5$ indicates a weak correlation, $r = 0.5 - 0.8$ indicates a moderate correlation, and $r > 0.8$ indicates a strong correlation.

participate in tumor progression and metastasis or mediate the loss of tumor suppressor functions, suggesting a need for future research to investigate the roles of these genes in NPC pathology and develop novel therapeutic approaches targeting these pathways. In addition, combining these findings with clinical data could improve our understanding of NPC, thereby facilitating the development of precision medicine approaches that improve patient outcomes through customized treatment strategies [28-30].

By analyzing DEGs related to NPC, 26 potential GRDEGs, such as BCL2A1, CCL2, BTK, and SRD5A2, were identified. Among them, BTK, CD72, PTPN6, and VAV1 were identified as independent prognostic predictors in patients with NPC [31]. Our study also revealed that the expression of these genes was lower in the NPC group than in the control group, in line with previous research (Figure 3) [31].

Aberrant expression of TP53, a tumor suppressor gene, could promote the malignant transformation of cells and tumorigenesis through cell cycle and apoptosis dysregulation and genomic instability. It is necessary to perform enrichment analysis of the TP53-mediated transcriptional regulation of GRDEGs to explore therapeutic interventions in NPC. Furthermore, the PI3K/AKT signaling pathway plays an important role in cell survival and proliferation. This pathway has become a significant target in cancer therapy given its function in promoting tumor cell growth and drug resistance. Because both the TP53 and PI3K/AKT pathways have important roles in the progression of NPC, novel therapies that specially target both pathways could significantly enhance treatment efficacy and reduce drug resistance in NPC treatment [4,27,32]. This is in line with the results of GSEA, offering a distinct direction for future research.

In the PPI network, nine hub genes were identified, and these genes are associated with 13 GRDEGs and are involved in the tumor microenvironment in NPC. In future research, it is necessary to investigate the functions of these hub genes to assess their potential as biomarkers or targets for personalized therapies. Moreover, evaluating the expression of these hub genes in relation to patients' prognosis could offer valuable insights into their clinical relevance [5,28,33].

Differences in the distributions of immune cell populations between the NPC and control groups, especially activated mast cells and CD4 memory T cells, were revealed by immune infiltration analysis. These findings emphasize the complexity of the immune landscape in NPC, suggesting that changes in immune cell composition have crucial roles in tumor development and progression. As the interactions between immune cells and the identified hub genes might influence treatment efficacy, developing innovative combination therapies that leverage the immune system could enhance treatment efficacy in NPC [34-36].

Several of the identified hub genes in this study have been implicated in the pathogenesis of NPC. *CXCL10* has been recognized as a risk factor for NPC, demonstrating a strong association with disease incidence and exhibiting potential diagnostic utility. Moreover, *CXCL10* could serve as a predictive biomarker for NPC development [37]. *IL1A* is significantly overexpressed in patients with NPC, and it is considered a key pyroptosis-related gene that contributes to both disease diagnosis and subtype classification, thereby offering guidance for therapeutic strategies [38]. Similarly, *IFNG*, a gene involved in programmed cell death, has been proposed as a promising biomarker for the diagnosis and treatment of NPC [39].

Downregulation of *CCL2* has been found to significantly inhibit tumor growth in prostate cancer models both *in vitro* and *in vivo*, and silencing of this gene reduces osteoclast formation induced by conditioned medium from prostate cancer cells. In NPC, *Notch1* knockdown has been reported to influence CCL2 expression, suggesting a regulatory interaction [40]. Elevated *CCL4* expression is closely linked to the development of NPC, and circulating *CCL4* could represent a non-invasive biomarker for evaluating tumor invasiveness, cell signaling interactions, and therapeutic response monitoring [41].

FN1 plays an important regulatory role in apoptosis-related gene networks within NPC, and it has been implicated in tumor progression, supporting its potential as a therapeutic target [42]. Using the CellChat R package, *SPP1* was identified as a key mediator of intercellular communication between high- and low-risk NPC groups [43]. *MMP3*, a known downstream oncogene, is transcriptionally activated in NPC and strongly associated with disease progression [44]. Finally, *BMP2* has been found to counteract the tumor-suppressive effects of F-box and WD repeat domain containing 7 overexpression in NPC cells. Transcriptional activation of *BMP2* is promoted by homeobox A10, which binds directly to its promoter region [45].

Despite the merits of this study, its limitations must be addressed. Sole bioinformatic analysis without experimental validation might have decreased the robustness of our findings. The reliance on publicly available datasets could have produced potential biases, including batch effects that could obscure the true biological variations associated with GRDEGs. Another issue is the small sample size of the combined GEO dataset, which might have reduced the reliability of the data used for analysis. Therefore, larger cohorts are needed to verify our findings. Finally, the clinical relevance of the potential biomarkers and therapeutic targets remains unknown because of the absence of clinical correlation analyses.

CONCLUSION

The differential expression of GRGs in NPC was found by this study. The roles of GRDEGs in the pathogenesis of NPC were also evaluated, including their associated biological pathways. Through our findings, novel po-

tential biomarkers and therapeutic targets were identified, and these results could support the development of novel diagnostic and therapeutic strategies for NPC in the future.

Acknowledgment:
The authors thank the Dazu Hospital affiliated with Chongqing Medical University for its support of this study. They also thank *Medjaden* Company for the scientific editing of this article.

Source of Funds:
Open access funding was provided by Dazu Hospital affiliated with Chongqing Medical University. This work was supported by the Science and Technology Bureau of Dazu District, Chongqing (grant No.: DZKJ2 024JSYJ-KWXM1009) and the Chongqing Education Commission (grant No.: KJQN202400404).

Availability of Data and Materials:
The data can be obtained from BiaoLi Long via email (150449@hospital.cqmu.edu.cn).

Declaration of Interest:
The authors declare no competing interests.

References:

1. Reffai A, Mesmoudi M, Derkaoui T, et al. Epidemiological Profile and Clinicopathological, Therapeutic, and Prognostic Characteristics of Nasopharyngeal Carcinoma in Northern Morocco. Cancer Control 2021;28:10732748211050587. (PMID: 34664512)

2. Chen P, Liu B, Xia X, Huang P, Zhao J. Current progress in immunotherapy of nasopharyngeal carcinoma. Am J Cancer Res 2023;13(4):1140-7. (PMID: 37168337)

3. Gondhowiardjo SA, Adham M, Rachmadi L, et al. Immune cells markers within local tumor microenvironment are associated with EBV oncoprotein in nasopharyngeal cancer. BMC Cancer 2022; 22(1):887 (PMID: 35963999)

4. Xia T, Ji Y, Lu YN, Xie HJ, You YW, You B. [Autophagy promotes recurrence of nasopharyngeal carcinoma via inducing the formation of dormant polyploid giant cancer cells]. Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi 2022;57(9):1102-9 (PMID: 36177565)

5. Tai J, Park J, Han M, Kim TH. Screening Key Genes and Biological Pathways in Nasopharyngeal Carcinoma by Integrated Bioinformatics Analysis. Int J Mol Sci 2022;23(24):15701. (PMID: 36555343)

6. Dodd LE, Sengupta S, Chen IH, et al. Genes involved in DNA repair and nitrosamine metabolism and those located on chromosome 14q32 are dysregulated in nasopharyngeal carcinoma. Cancer Epidemiol Biomarkers Prev 2006;15(11):2216-25. (PMID: 17119049)

7. Bao YN, Cao X, Luo DH, et al. Urokinase-type plasminogen activator receptor signaling is critical in nasopharyngeal carcinoma cell growth and metastasis. Cell Cycle 2014;13(12):1958-69. (PMID: 24763226)

8. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics 2007;23(14):1846-7. (PMID: 17496320)

9. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets-update. Nucleic Acids Res 2013; 41(Database issue):D991-5. (PMID: 23193258)

10. Stelzer G, Rosen N, Plaschkes I, et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. Curr Protoc Bioinformatics 2016; 54:1.30.1-1.30.33 (PMID: 27322403)

11. Chantzichristos D, Svensson PA, Garner T, et al. Identification of human glucocorticoid response markers using integrated multi-omic analysis from a randomized crossover trial. Elife 2021;10: e62236. (PMID: 33821793)

12. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28(6):882-3 (PMID: 22257669)

13. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47. (PMID: 25605792)

14. Ben Salem K, Ben Abdelaziz A. Principal Component Analysis (PCA). Tunis Med 2021;99(4):383-9 (PMID: 35244921)

15. Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res 2019;47(D1):D419-D26. (PMID: 30407594)

16. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28(1):27-30. (PMID: 10592173)

17. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284-7. (PMID: 22455463)

18. Wu MC, Lin X. Prior biological knowledge-based approaches for the analysis of genome-wide expression profiles using gene sets and pathways. Stat Methods Med Res 2009;18(6):577-93. (PMID: 20048386)

19. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27(12):1739-40. (PMID: 21546393)

20. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47(D1):D607-D613. (PMID: 30476243)

21. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014;8 Suppl 4(Suppl 4):S11. (PMID: 25521941)

22. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498-504. (PMID: 14597658)

23. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 2014; 42(Database issue):D92-7. (PMID: 24297251)

24. Zhang Q, Liu W, Zhang HM, et al. hTFtarget: A Comprehensive Database for Regulations of Human Transcription Factors and Their Targets. Genomics Proteomics Bioinformatics 2020;18(2): 120-8. (PMID: 32858223)

25. Petersson F. Nasopharyngeal carcinoma: a review. Semin Diagn Pathol 2015;32(1):54-73. (PMID: 25769204)

26. Zheng M, Ren Y, Jing L, Cheng M, Lin J, Yu Y. Nasopharyngeal carcinoma cell screening based on nuclear targeting Surface-Enhanced Raman Scattering (SERS) detection. Anal Chim Acta 2024;1316:342864. (PMID: 38969411)

27. Wen JY, Chen G, Li JD, et al. Downregulated miR-150-5p in the Tissue of Nasopharyngeal Carcinoma. Genet Res (Camb) 2022; 2022:2485055. (PMID: 36118276)

28. Hsu WL, Tao J, Fu S, et al. Kinetics of EBV antibody-based NPC risk scores in Taiwan NPC multiplex families. Int J Cancer 2024; 155(8):1400-8. (PMID: 38822730)

29. Ouyang Y, Jin YB, Chen XP, et al. STIL is upregulated in nasopharyngeal carcinoma tissues and promotes nasopharyngeal carcinoma proliferation, migration and invasion. Neoplasma 2020; 67(1):37-45. (PMID: 31607137)

30. Luo WJ, He SW, Zou WQ, et al. Epstein-Barr virus microRNA BART10-3p promotes dedifferentiation and proliferation of nasopharyngeal carcinoma by targeting ALK7. Exp Biol Med (Maywood) 2021;246(24):2618-29. (PMID: 34424090)

31. Lu X, Chen X, Wang X, Qing J, Li J, Pan Y. Construction of lncRNA and mRNA co-expression network associated with nasopharyngeal carcinoma progression. Front Oncol 2022;12:965088. (PMID: 35957889)

32. Luo W. Nasopharyngeal carcinoma ecology theory: cancer as multidimensional spatiotemporal "unity of ecology and evolution" pathological ecosystem. Theranostics 2023; 13(5):1607-31. (PMID: 37056571)

33. Wan F, Zhang H, Hu J, et al. Mesenchymal Stem Cells Inhibits Migration and Vasculogenic Mimicry in Nasopharyngeal Carcinoma Via Exosomal MiR-125a. Front Oncol 2022;12:781979. (PMID: 35251967)

34. Gondhowiardjo SA, Handoko, Adham M, et al. Tumor microenvironment predicts local tumor extensiveness in PD-L1 positive nasopharyngeal cancer. PLoS One 2020;15(3):e0230449. (PMID: 32191754)

35. Poluan RH, Sudigyo D, Rahmawati G, et al. Transcriptome Related to Avoiding Immune Destruction in Nasopharyngeal Cancer in Indonesian Patients Using Next-Generation Sequencing. Asian Pac J Cancer Prev 2020;21(9):2593-601. (PMID: 32986357)

36. Wang L, Wang D, Zeng X, et al. Exploration of spatial heterogeneity of tumor microenvironment in nasopharyngeal carcinoma via transcriptional digital spatial profiling. Int J Biol Sci 2023; 19(7):2256-69. (PMID: 37151882)

37. Kai J, Huang H, Su J, et al. Identification of shared immune infiltration characteristic molecules in dermatomyositis and nasopharyngeal carcinoma using bioinformatics: Traits in dermatomyositis and nasopharyngeal cancer. Skin Res Technol 2024;30(8): e13871. (PMID: 39081134)

38. Wang Y, Zou Y, Chen X, Wang X, Zheng H, Ye Q. Relevance of pyroptosis-associated genes in nasopharyngeal carcinoma diagnosis and subtype classification. J Gene Med 2024;26(1):e3653. (PMID: 38282154)

39. Yan J, Wu L, Zheng M, et al. Exploring programmed cell death-related biomarkers and disease therapy strategy in nasopharynxgeal carcinoma using transcriptomics. Front Biosci (Landmark Ed) 2024;29(7):240. (PMID: 39082346)

40. Guo H, Wang F, Diao Y, et al. Knockdown of Notch1 inhibits nasopharyngeal carcinoma cell growth and metastasis via downregulation of CCL2, CXCL16, and uPA. Mol Carcinog 2019; 58(10):1886-96. (PMID: 31270884)

41. Liang C, Kan J, Wang J, et al. Nasopharyngeal carcinoma-associated inflammatory cytokines: Ongoing biomarkers. Front Immunol 2024;15:1448012. (PMID: 39483474)

42. Li D, Bao L, Liu S, et al. Identification and validation of molecular features of the anoikis gene-related hub genes in nasopharynxgeal carcinoma. Appl Biochem Biotechnol 2025;197(3):2066-92. (PMID: 39666232)

43. Chen X, Ding Q, Lin T, et al. An immune-related prognostic model predicts neoplasm-immunity interactions for metastatic nasopharyngeal carcinoma. Front Immunol 2023;14:1109503. (PMID: 37063853)

44. Wang D, Zuo S, Ge J, et al. circTP63-N suppresses the proliferation and metastasis of nasopharyngeal carcinoma via engaging with HSP90AB1 to modulate the YAP1/hippo signaling pathway. Sci China Life Sci 2025;68(3):689-705. (PMID: 39754006)

45. Zhong Q, Wang Z, Kang H, Wu R. Molecular mechanism of FBXW7-mediated ubiquitination modification in nasopharyngeal carcinoma cell proliferation in vitro and in vivo. Pathol Res Pract 2023;244:154056. (PMID: 36989847)