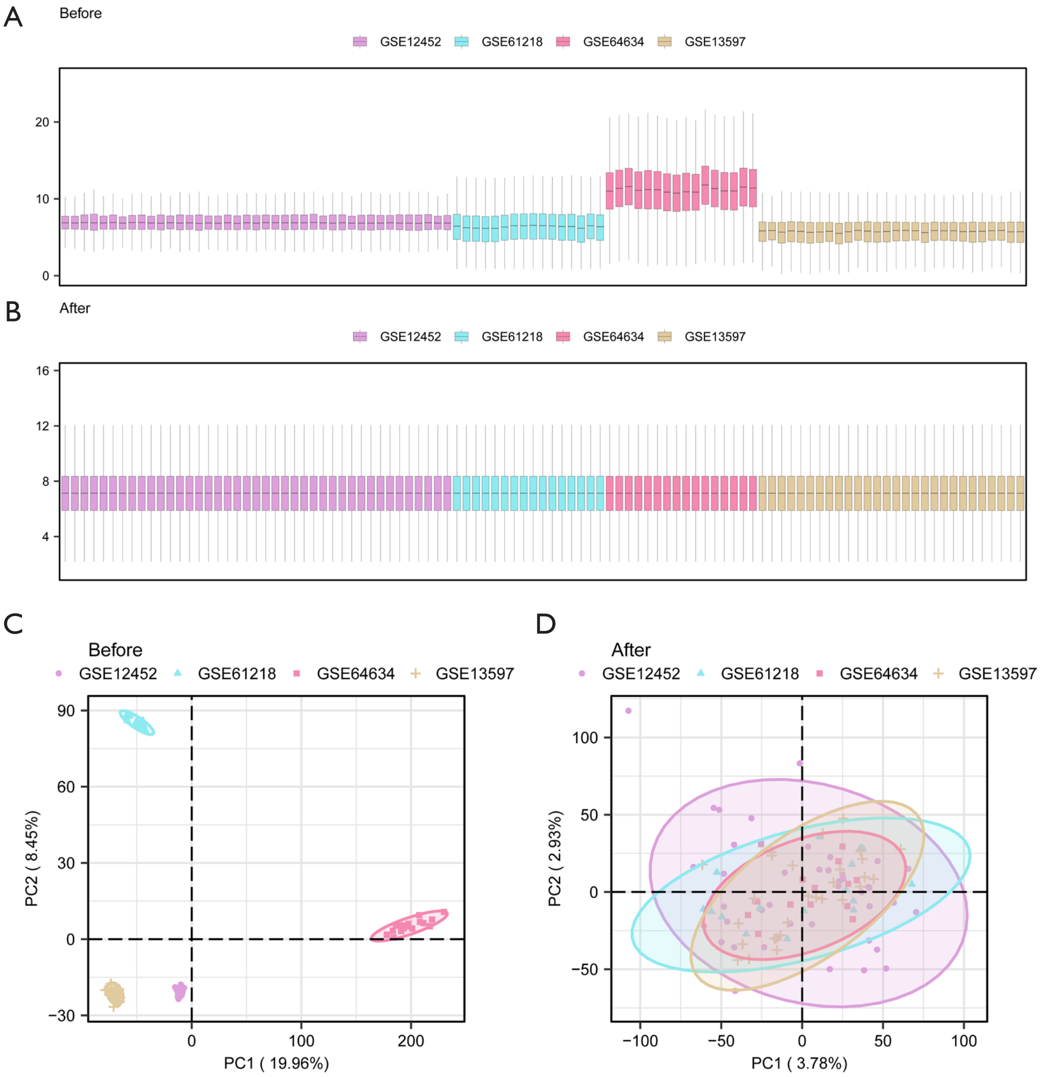

Figure S1 Data preprocessing. Distribution box plots of the integrated GEO dataset (GEO-Combined) before batch effect removal (A) and after batch effect removal (B). PCA plots of the integrated GEO-Combined dataset before batch effect removal (C) and after batch effect removal (D). GEO, Gene Expression Omnibus; PC, principal component; PCA, principal component analysis.

https://dx.doi.org/10.21037/tcr-2025-1263

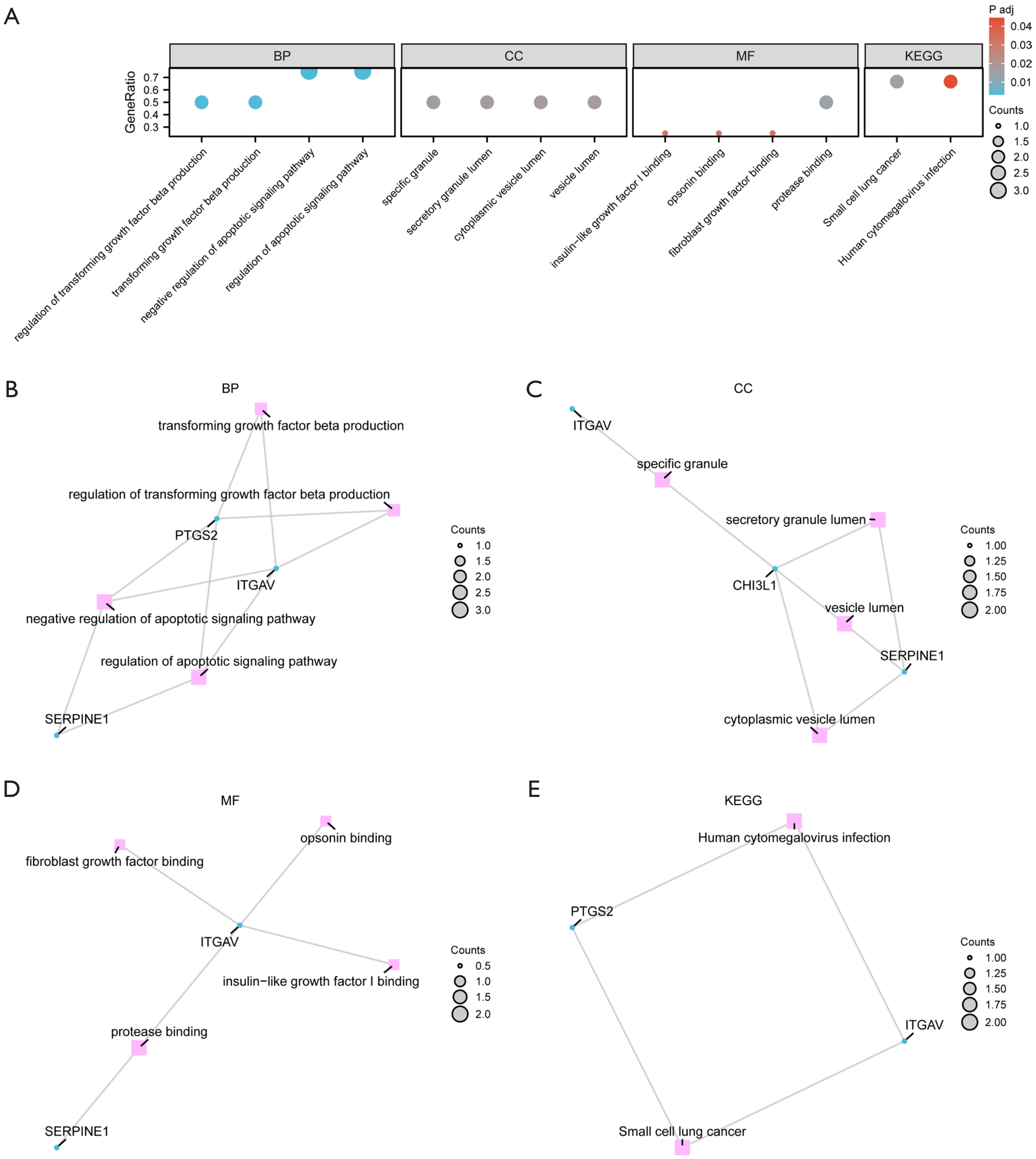

Figure S2 GO and KEGG enrichment analyses of the key genes. (A) Bubble plot depicting the results of the GO and KEGG enrichment analyses for the key genes. Circular network diagrams illustrating the enriched terms in the GO and KEGG analyses for the BP (B), CC (C), MF (D), and KEGG pathways (E). (A) In the bubble plot, the X-axis represents the GO terms, and the Y-axis represents the enrichment factor GeneRatio for the different GO terms. (B-E) In the circular network diagrams, the blue nodes represent specific genes, while the pink nodes represent specific pathways. The selection criteria for the GO enrichment entries were a P value <0.05 and an FDR q value <0.05. BP, biological process; CC, cellular component; FDR, false discovery rate; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function.

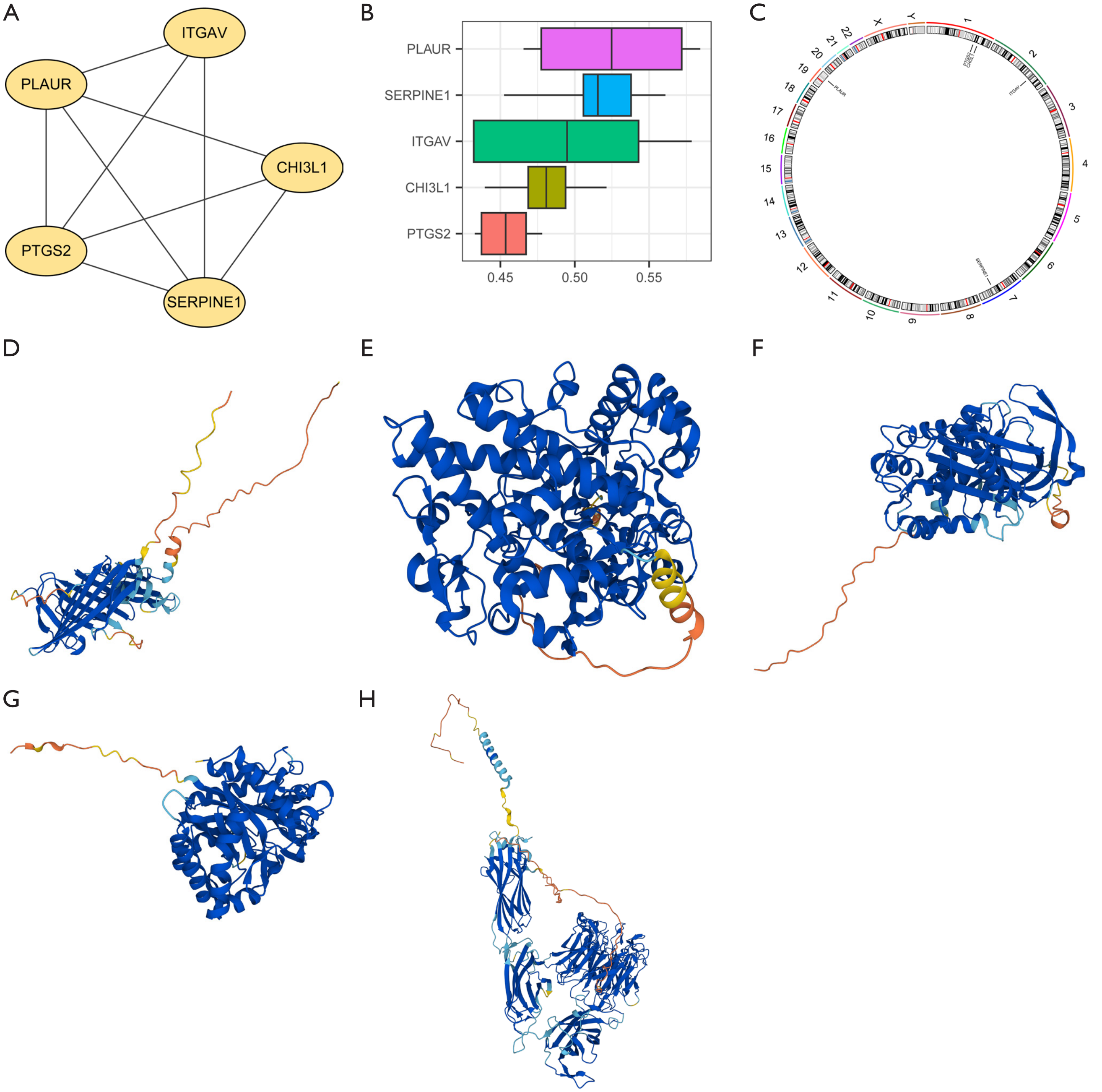

Figure S3 Network analysis of the PPI. (A) PPI network for the key genes. (B) Functional similarity analysis results for the key genes. (C) Chromosomal localization map of the key genes. (D-H) Protein structure representations of the key genes: *PLAUR* (D), *PTGS2* (E), *SERPINE1* (F), *CHI3L1* (G), and *ITGAV* (H). The confidence score for each residue, as generated by the AlphaFold website, ranges from 0 to 100, and is referred to as the pLDDT. Some regions with pLDDT ≤50 may represent isolated unstructured segments, indicating very low model confidence in the red region. The yellow region (50< pLDDT ≤70) indicates that the model confidence is relatively low; the light blue region (70< pLDDT ≤90) corresponds to normal model confidence; while the blue region (pLDDT >90) indicates exceptionally high model confidence. pLDDT, per-residue local distance difference test; PPI, protein-protein interaction.

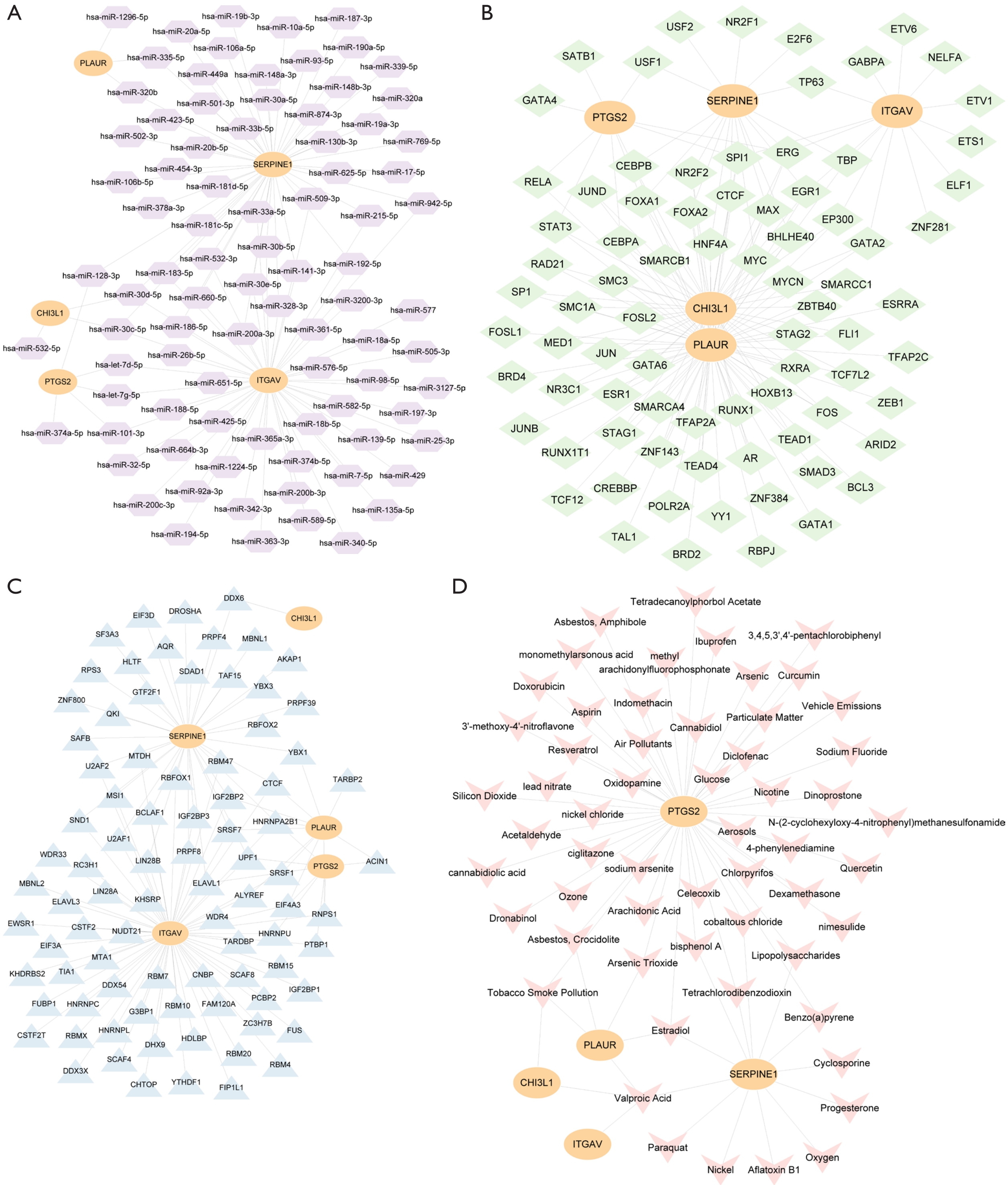

Figure S4 Network of molecular interactions among the key genes. (A) In the mRNA-miRNA interaction network of the key genes, the orange ellipses represent mRNA, and the purple polygons represent miRNA. (B) In the mRNA-TF interaction network, the orange ellipses signify mRNA, while the green diamonds represent TF. (C) In the mRNA-RBP interaction network, the orange ellipses denote mRNA, and the blue triangles represent RBP. (D) In the mRNA-drugs interaction network, the orange ellipses represent mRNA, and the pink inverted triangles represent drugs. miRNA, microRNA; mRNA, messenger RNA; RBP, RNA-binding protein; TF, transcription factor.

Table S1 Clinical data for qRT-PCR of patients with NPC at Fujian Cancer Hospital

Characteristics	Data (n=33)
Age (years), n (%)	
>60	6 (18.18)
≥40 and ≤60	20 (60.61)
<40	7 (21.21)
Clinical T stage, n (%)	
1	2 (6.06)
2	6 (18.18)
3	19 (57.58)
4	6 (18.18)
Clinical N stage, n (%)	
0	2 (6.06)
1	15 (45.45)
2	10 (30.30)
3	6 (18.18)
Clinical TNM stage, n (%)	
II	4 (12.12)
III	18 (54.55)
IVA	10 (30.30)
IVB	1 (3.03)

N, node; NPC, nasopharyngeal carcinoma; qRT-PCR, quantitative real-time polymerase chain reaction; T, tumor; TNM, tumor-node-metastasis.

Table S2 Primers for PLAUR, PTGS2, SERPINE1, CHI3L1, and ITGAV

Primers	Forward primers (5' to 3')	Reverse primers (5' to 3')
qPCR-PLAUR	GAGCTATCGGACTGGCTTGAA	CGGCTTCGGGAATAGGTGAC
qPCR-PTGS2	TAAGTGCGATTGTACCCGGAC	TTTGTAGCCATAGTCAGCATTGT
qPCR-SERPINE1	AGTGGACTTTTCAGAGGTGGA	GCCGTTGAAGTAGAGGGCATT
qPCR-CHI3L1	GAAGACTCTCTTGTCTGTCGGA	AATGGCGGTACTGACTTGATG
qPCR-ITGAV	GCTGTCGGAGATTTCAATGGT	TCTGCTCGCCAGTAAAATTGT

qPCR, quantitative polymerase chain reaction.

Table S3 GSEA results for the NPC Test_Data dataset

Description	Set size	Enrichment score	NES	P value	P adjust	q value
REACTOME_ASSEMBLY_OF_COLLAGEN_FIBRILS_AND_OTHER_MULTIMERIC_STRUCTURES	58	0.740878	2.539856	0.002364	0.039717	0.033844
REACTOME_MET_PROMOTES_CELL_MOTILITY	40	0.749123	2.418214	0.002208	0.039717	0.033844
REACTOME_ANCHORING_FIBRIL_FORMATION	15	0.820237	2.065192	0.002294	0.039717	0.033844
KEGG_HEDGEHOG_SIGNALING_PATHWAY	47	0.614931	2.022627	0.002315	0.039717	0.033844
REACTOME_TRANSCRIPTIONAL_REGULATION_BY_THE_AP_2_TFAP2_FAMILY_OF_TRANSCRIPTION_FACTORS	30	0.661573	2.00407	0.002212	0.039717	0.033844
PID_WNT_SIGNALING_PATHWAY	26	0.674001	1.990689	0.002193	0.039717	0.033844
REACTOME_NEGATIVE_REGULATION_OF_TCF_DEPENDENT_SIGNALING_BY_WNT_LIGAND_ANTAGONISTS	12	0.80554	1.926225	0.002193	0.039717	0.033844
WP_WNT_SIGNALING_IN_KIDNEY_DISEASE	34	0.618199	1.9227	0.002208	0.039717	0.033844
WP_PHOTODYNAMIC_THERAPYINDUCED_NFKB_SURVIVAL_SIGNALING	34	0.614275	1.910496	0.002208	0.039717	0.033844
WP_LNCRNA_IN_CANONICAL_WNT_SIGNALING_AND_COLORECTAL_CANCER	86	0.507904	1.897077	0.002415	0.039717	0.033844
WP_QUERCETIN_AND_NFKB_AP1_INDUCED_APOPTOSIS	14	0.769725	1.895464	0.002299	0.039717	0.033844
WNT_SIGNALING	81	0.510058	1.891739	0.002415	0.039717	0.033844
WP_NCRNAS_INVOLVED_IN_WNT_SIGNALING_IN_HEPATOCELLULAR_CARCINOMA	78	0.507376	1.85772	0.002415	0.039717	0.033844
WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLORECTAL_CANCER	143	0.442109	1.758363	0.002519	0.039717	0.033844
REACTOME_TP53_REGULATES_TRANSCRIPTION_OF_CELL_CYCLE_GENES	49	0.521332	1.721486	0.002347	0.039717	0.033844
WP_METABOLIC_REPROGRAMMING_IN_PANCREATIC_CANCER	42	0.523828	1.701619	0.002268	0.039717	0.033844
WP_METABOLIC_REPROGRAMMING_IN_COLON_CANCER	42	0.522397	1.69697	0.002268	0.039717	0.033844
WP_WNT_SIGNALING_PATHWAY_AND_PLURIPOTENCY	96	0.444265	1.678542	0.002415	0.039717	0.033844
WP_PI3KAKT_SIGNALING_PATHWAY	283	0.386865	1.674012	0.002994	0.040599	0.034595
WP_FOCAL_ADHESION_PI3KAKTMTORSIGNALING_PATHWAY	265	0.382729	1.64259	0.00289	0.039874	0.033978

GSEA, gene set enrichment analysis; NES, normalized enrichment score; NPC, nasopharyngeal carcinoma.

https://dx.doi.org/10.21037/tcr-2025-1263

Table S4 GSVA results for the NPC Test_Data dataset

Pathway	LogFC	AveExpr	*t*	P value	P adjust	B
HALLMARK_ESTROGEN_RESPONSE_EARLY	−0.32033	0.002278	−2.40385	0.030079	0.046999	−4.54831
HALLMARK_APICAL_SURFACE	−0.33104	−0.02936	−2.45185	0.027403	0.045672	−4.46049
HALLMARK_MYOGENESIS	−0.33982	0.047895	−2.45613	0.027176	0.045672	−4.45263
HALLMARK_XENOBIOTIC_METABOLISM	−0.37453	0.048514	−2.55113	0.02257	0.040303	−4.27666
HALLMARK_COAGULATION	−0.3838	−0.00795	−2.82766	0.013049	0.027257	−3.75128
HALLMARK_ESTROGEN_RESPONSE_LATE	−0.38969	−0.00808	−2.92264	0.01079	0.024523	−3.56712
HALLMARK_CHOLESTEROL_HOMEOSTASIS	−0.40014	−0.02754	−2.62221	0.019624	0.03634	−4.14337
HALLMARK_HYPOXIA	−0.41882	−0.04402	−3.76321	0.001978	0.007308	−1.89352
HALLMARK_APICAL_JUNCTION	−0.44302	0.019686	−3.94336	0.001378	0.006262	−1.53134
HALLMARK_DNA_REPAIR	−0.4476	−0.04466	−3.72569	0.002134	0.007308	−1.96898
HALLMARK_UV_RESPONSE_UP	−0.59063	−0.02208	−6.94009	5.64E−06	4.70E−05	4.061298
HALLMARK_G2M_CHECKPOINT	−0.6154	−0.0869	−3.18081	0.006416	0.015276	−3.05949
HALLMARK_WNT_BETA_CATENIN_SIGNALING	−0.65657	−0.02875	−7.71517	1.65E−06	2.31E−05	5.316391
HALLMARK_INTERFERON_ALPHA_RESPONSE	−0.66731	−0.07155	−3.68228	0.002329	0.007308	−2.05628
HALLMARK_NOTCH_SIGNALING	−0.66856	−0.0256	−7.64112	1.85E−06	2.31E−05	5.200014
HALLMARK_E2F_TARGETS	−0.67961	−0.07964	−3.44964	0.003726	0.010958	−2.52339
HALLMARK_TNFA_SIGNALING_VIA_NFKB	−0.71538	−0.03702	−7.67603	1.75E−06	2.31E−05	5.254968
HALLMARK_MYC_TARGETS_V2	−0.74158	−0.14013	−2.76825	0.014691	0.029381	−3.86561
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	−0.77627	−0.02635	−7.18315	3.80E−06	3.80E−05	4.463761
HALLMARK_ANGIOGENESIS	−0.84921	−0.04162	−8.71518	3.78E−07	1.89E−05	6.818019

GSVA, gene set variation analysis; NPC, nasopharyngeal carcinoma.

Table S5 ARDEGs selected by random forest

Gene symbol	Mean decrease accuracy	Mean decrease Gini
FN1	0.0062	0.0948
EPHB6	0.005133	0.0844
ETV7	0.006467	0.0798
PTGS2	0.005633	0.0782
CXCL12	0.006867	0.076
PLAU	0.004967	0.073
COL17A1	0.002167	0.0722
FGF1	0.0053	0.0716
ID2	0.005367	0.0706
THY1	0.002667	0.0682
KITLG	0.008067	0.068
PBK	0.0055	0.0666
TRIM28	0.002667	0.0664
LAMB3	0.004633	0.0662
EDNRB	0.0041	0.0652
ICAM1	0.005867	0.0646
SERPINE1	0.0055	0.0644
PLAUR	0.004333	0.0622
TP63	0.0055	0.061
WNT5A	0.004833	0.06
NME1	0.0023	0.0594
COL4A2	0.003567	0.0592
CCND1	0.004567	0.059
PPP2R5A	0.003833	0.059
C5AR1	0.006567	0.0584
KIF14	0.006333	0.0578
GATA6	0.0039	0.0576
PYCARD	0.004633	0.0574
CDK1	0.004	0.0572
CENPF	0.005	0.0568
CLU	0.003833	0.0566
SGK1	0.002533	0.0564
RPS6KA1	0.0058	0.0558
ITGAV	0.003633	0.0538
PIK3C2B	0.004833	0.0528
MAP3K1	0.002567	0.0526
WNT2	0.0043	0.0506
MAPK3	0.0025	0.0504
ATP2A3	0.0034	0.0502
MMP3	0.0034	0.0502
LYPD1	0.003233	0.0492
HOXA10	0.004133	0.048
PDCD6IP	0.0027	0.0474
EZR	0.002067	0.0474
CASP6	0.0044	0.0472
RACGAP1	0.0057	0.047
TGFBR3	0.002167	0.0468
TGFBR2	0.002	0.0466
E2F1	0.001567	0.0464
CYCS	0.002667	0.0458
CHI3L1	0.002333	0.0448
IRF6	0.002352	0.0446
BMP2	0.002833	0.0444
SMAD1	0.003067	0.0438
LGALS1	0.0039	0.0426
BIRC5	0.0015	0.0426
S100A4	0.001333	0.0422
MMP11	0.002567	0.042
TP53	0.002133	0.0416
LPAR1	0.003167	0.0414
BRCA1	0.001067	0.0408
UBE2S	0.003967	0.0408
CDC25C	0.001067	0.0404
CCR7	0.003167	0.0386
PCNA	0.002567	0.037
XAF1	9.00E-04	0.0368
GNE	0.0022	0.0364
LAMB1	0.006133	0.0362
NOX4	0.0014	0.036

ARDEGs, anoikis-related differentially expressed genes.

Table S6 GO enrichment analysis results of the key genes

Ontology	ID	Description	Gene ratio	Bg ratio	P value	P adjust	q value
BP	GO:2001234	Negative regulation of apoptotic signaling pathway	3/4	230/18,800	7.16E−06	0.00321	0.000701
	GO:0071634	Regulation of transforming growth factor beta production	2/4	39/18,800	2.51E−05	0.003424	0.000748
	GO:2001233	Regulation of apoptotic signaling pathway	3/4	370/18,800	2.98E−05	0.003424	0.000748
	GO:0071604	Transforming growth factor beta production	2/4	43/18,800	3.06E−05	0.003424	0.000748
CC	GO:0042581	Specific granule	2/4	160/19,594	0.000393	0.01652	0.005797
	GO:0034774	Secretory granule lumen	2/4	322/19,594	0.00158	0.017109	0.006003
	GO:0060205	Cytoplasmic vesicle lumen	2/4	325/19,594	0.00161	0.017109	0.006003
	GO:0031983	Vesicle lumen	2/4	327/19,594	0.001629	0.017109	0.006003
MF	GO:0002020	Protease binding	2/4	136/18,410	0.000322	0.014485	0.002711
	GO:0031994	Insulin-like growth factor I binding	1/4	13/18,410	0.002822	0.032181	0.006022
	GO:0001846	Opsonin binding	1/4	16/18,410	0.003472	0.032181	0.006022
	GO:0017134	Fibroblast growth factor binding	1/4	23/18,410	0.004988	0.032181	0.006022

BP, biological process; CC, cellular component; GO, Gene Ontology; MF, molecular function.

Table S7 KEGG enrichment analysis results of the key genes

Ontology	ID	Description	Gene ratio	Bg ratio	P value	P adjust	q value
KEGG	hsa05222	Small cell lung cancer	2/3	92/8,164	0.000374	0.014964	0.006301
	hsa05163	Human cytomegalovirus infection	2/3	225/8,164	0.002227	0.04455	0.018758

KEGG, Kyoto Encyclopedia of Genes and Genomes.

https://dx.doi.org/10.21037/tcr-2025-1263