

Supplementary Material

Bioinformatic analysis of glycolysis and lactate metabolism genes in head and neck squamous cell carcinoma

Huanyu Jiang^{1,2}, Lijuan Zhou², Haidong Zhang^{1,2}, Zhenkun Yu^{1,2*}

¹School of Medicine, Southeast University, Nanjing 210009, Jiangsu, China

²Department of Otolaryngology Head and Neck Surgery, The Affiliated BenQ Hospital of Nanjing Medical University, Nanjing 210019, Jiangsu, China

Corresponding author: Zhenkun Yu

E-mail: yuzhenkun65@hotmail.com

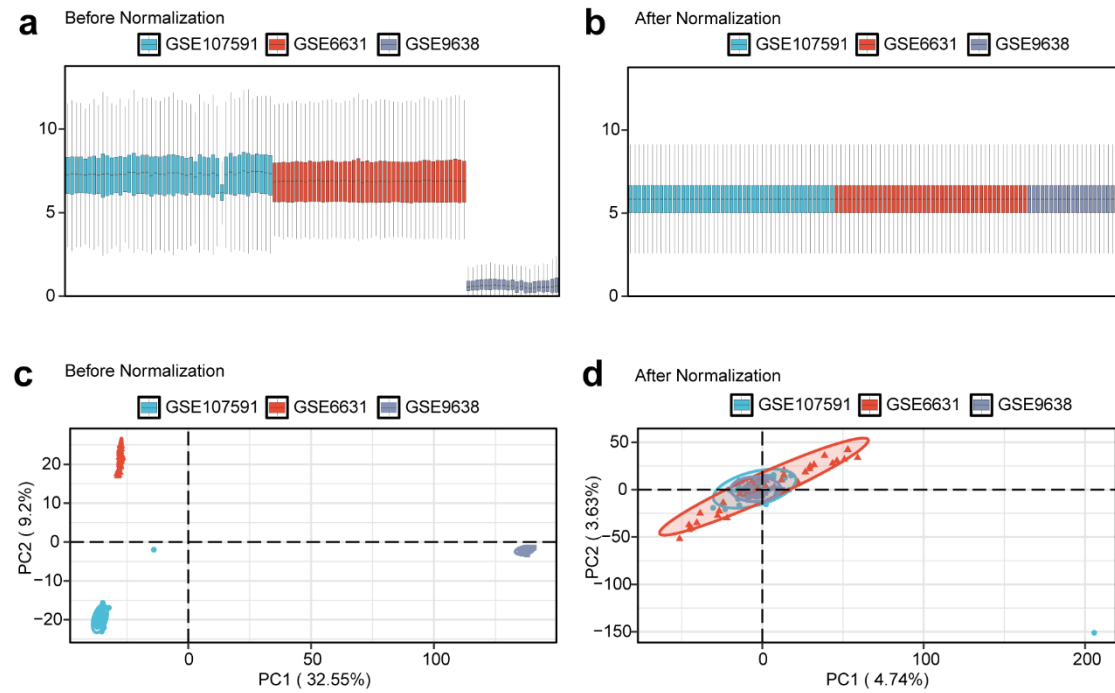


Fig. S1 Batch effects removal of GSE107591, GSE6631, and GSE9638

(a) Boxplots showing the distribution of combined GEO datasets before batch removal. (b) Boxplots showing the distribution of combined GEO datasets after batch effects removal. (c) PCA plot of the datasets before de-batching. (d) PCA map of the combined GEO datasets after batch processing. GEO: Gene Expression Omnibus, PCA: Principal component analysis.

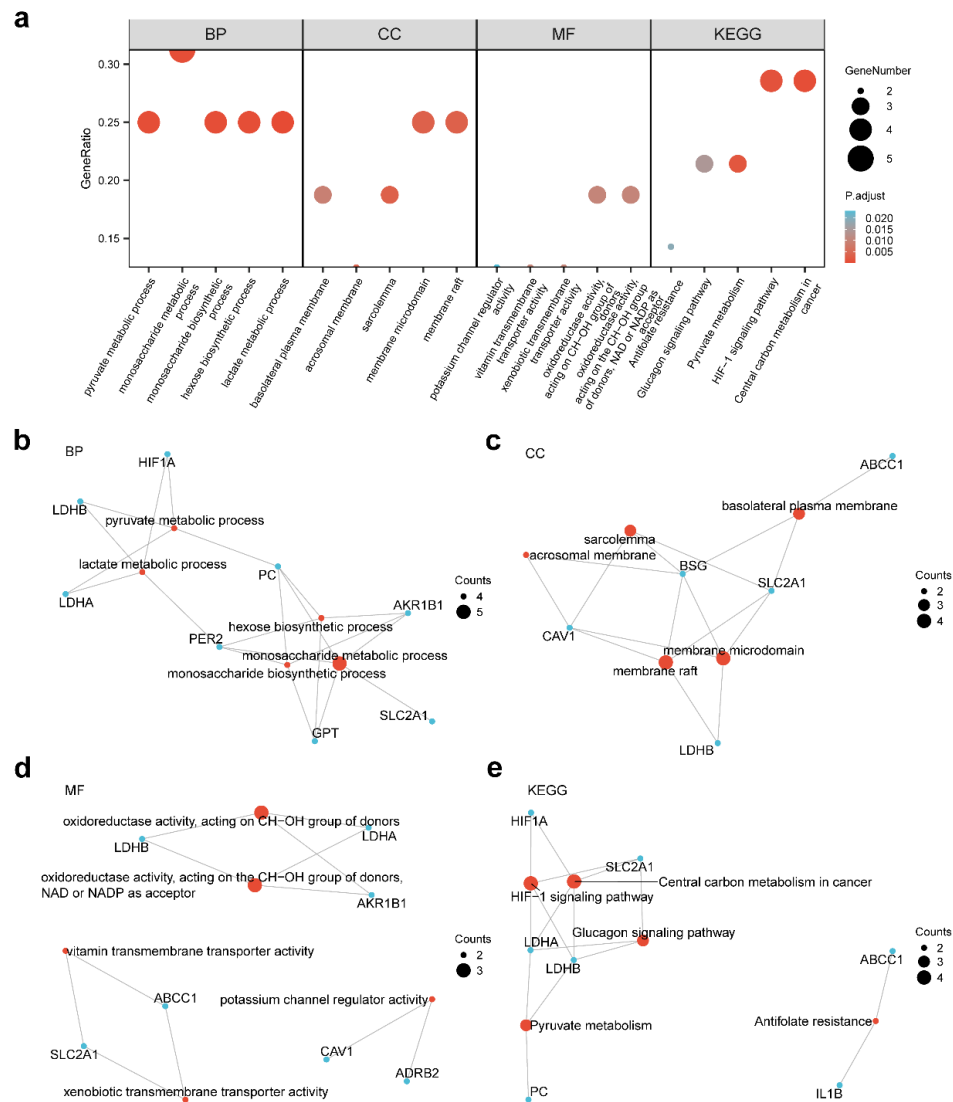


Fig. S2 GO and KEGG enrichment analyses of GALMRDEGs

(a) Bubble plot of GO and KEGG enrichment analyses results of GALMRDEGs. GO and KEGG terms are shown on the abscissa. (b–e) GO and KEGG enrichment analyses results of GALMRDEGs: BP (b), CC (c), MF (d), and KEGG (e). The red nodes represent items, the blue nodes represent molecules, and the lines represent the relationships between items and molecules. The bubble size in the bubble plot represents the number of genes, and the color of the bubble represents the size of the adj.*P*, where the redder the color, the smaller the adj.*P*, and the bluer the color, the larger the adj.*P*. GALMRDEGs: Glycolysis and lactate metabolism-related differentially expressed genes, GO: Gene Ontology, BP: Biological process, CC: Cellular component, MF: Molecular function, KEGG: Kyoto Encyclopedia of Genes and Genomes. KEGG pathway data were obtained with permission from Kanehisa Laboratories (www.kegg.jp).

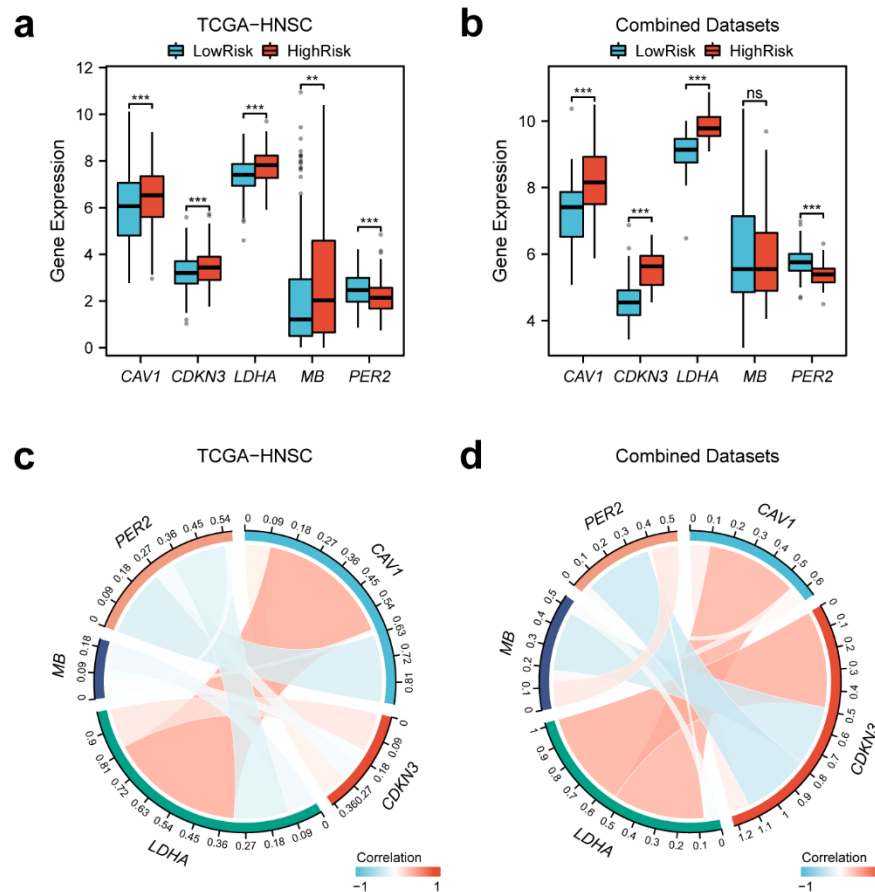


Fig. S3 Analysis of the differences in expression and correlation of key genes

(a) Boxplots for group comparison of key genes in HNSCC from TCGA-HNSC dataset. (b) Group comparison boxplots of key genes in HNSCC from combined GEO datasets. (c) Correlation chord plots of key genes in HNSCC from TCGA-HNSC dataset. (d) Correlation chord diagram of key genes in HNSCC from combined GEO datasets. In the correlation graph, red indicates a positive correlation, blue indicates a negative correlation, and the width of the string shows the strength of the correlation. ns: $P \geq 0.05$, not statistically significant, *: $P < 0.05$, statistically significant, **: $P < 0.01$, highly statistically significant, ***: $P < 0.001$, extremely highly statistically significant. TCGA: The Cancer Genome Atlas, HNSCC: Head and neck squamous cell carcinoma, GEO: Gene Expression Omnibus.

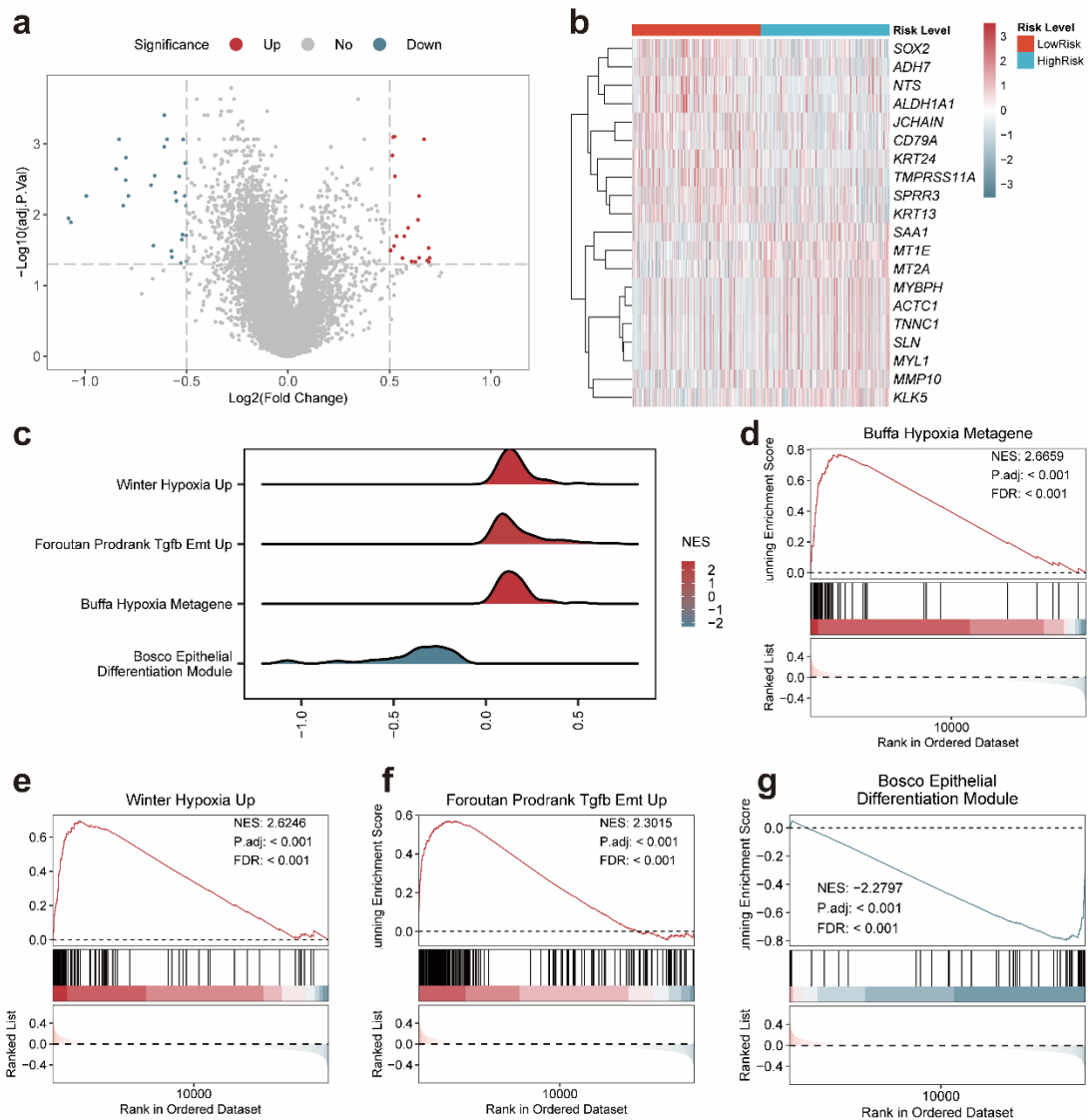


Fig. S4 Differential gene expression analysis and GSEA for TCGA-HNSC

(a–b) Volcano map (a) and heatmap of expression levels (b) of differentially expressed genes between the high- and low-risk groups from TCGA-HNSC dataset. (c) GSEA biological function mountain map display of HNSCC in TCGA-HNSC dataset. (d–g) GSEA showed that HNSCC in TCGA-HNSC dataset was significantly enriched in Buffa Hypoxia Metagene (d), Winter Hypoxia Up (e), Foroutan Prodrank TGF- β EMT Up (f), and Bosco Epithelial Differentiation Module (g). TCGA: The Cancer Genome Atlas, HNSCC: Head and neck squamous cell carcinomas, GSEA: Gene set enrichment analysis.

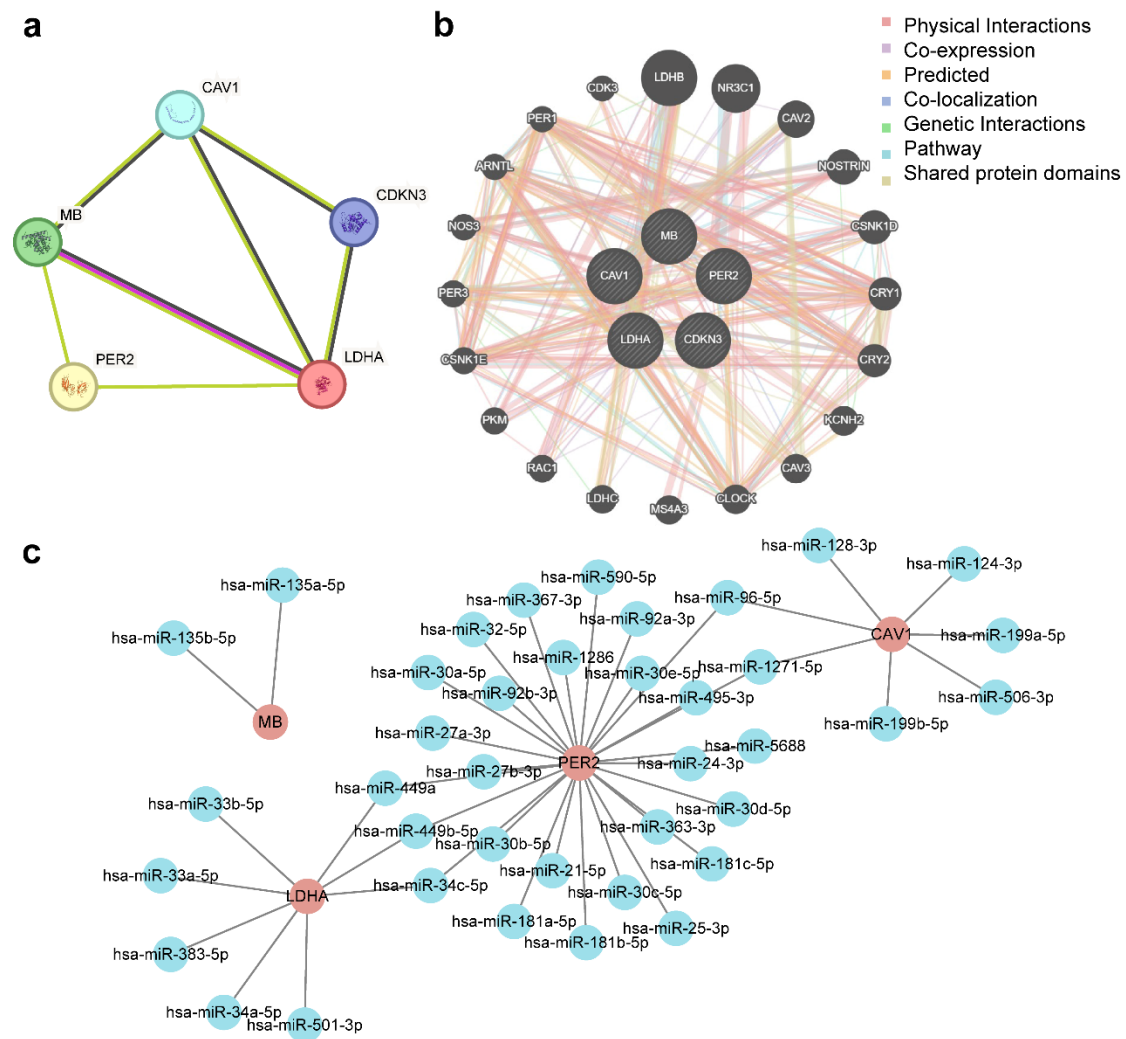


Fig. S5 PPI network analysis

(a) The STRING database was used to develop a PPI network of key genes. (b) GeneMANIA predicts the key functions of similar genes in the network. The circles indicate the key genes and their functionally similar genes, and the colors corresponding to the lines represent their interrelated functions. (c) mRNA-miRNA regulatory network of key genes. Pink represents mRNA, and blue represents miRNA. PPI: Protein-protein interaction, GALMRDEGs: Glycolysis and lactate metabolism-related differentially expressed genes.

Table S1. Results of GO and KEGG enrichment analysis for GALMRDEGs

ONTOLOGY	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
BP	GO:0006089	lactate metabolic process	4/16	17/18903	8.0911E-10	1.0891E-06	5.2634E-07
BP	GO:0019319	hexose biosynthetic process	4/16	86/18903	6.9703E-07	0.00037553	0.0001815
BP	GO:0046364	monosaccharide biosynthetic process	4/16	90/18903	8.37E-07	0.00037553	0.0001815
BP	GO:0005996	monosaccharide metabolic process	5/16	248/18903	1.4491E-06	0.00046826	0.00022631
BP	GO:0006090	pyruvate metabolic process	4/16	108/18903	1.7395E-06	0.00046826	0.00022631
CC	GO:0045121	membrane raft	4/16	326/19869	0.00011079	0.00334613	0.00184773
CC	GO:0098857	membrane microdomain	4/16	327/19869	0.00011211	0.00334613	0.00184773
CC	GO:0042383	sarcolemma	3/16	136/19869	0.00016456	0.00334613	0.00184773
CC	GO:0002080	acrosomal membrane	2/16	30/19869	0.00026101	0.00398041	0.00219798
CC	GO:0016323	basolateral plasma membrane	3/16	229/19869	0.00075736	0.00923976	0.0051022
MF	GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	3/16	126/18432	0.00016367	0.01018113	0.00526767
MF	GO:0016614	oxidoreductase activity, acting on CH-OH group of donors	3/16	138/18432	0.00021411	0.01018113	0.00526767

MF	GO:0042910	xenobiotic transmembrane transporter activity	2/16	31/18432	0.00032372	0.01018113	0.00526767
MF	GO:0090482	vitamin transmembrane transporter activity	2/16	32/18432	0.00034512	0.01018113	0.00526767
MF	GO:0015459	potassium channel regulator activity	2/16	54/18432	0.00098468	0.02323853	0.0120235
KEGG	hsa05230	Central carbon metabolism in cancer	4/14	70/8772	3.5047E-06	0.00034697	0.00028776
KEGG	hsa04066	HIF-1 signaling pathway	4/14	108/8772	1.9783E-05	0.00097925	0.00081214
KEGG	hsa00620	Pyruvate metabolism	3/14	47/8772	5.0353E-05	0.00166165	0.00137808
KEGG	hsa04922	Glucagon signaling pathway	3/14	107/8772	0.00058245	0.01441565	0.01195557
KEGG	hsa01523	Antifolate resistance	2/14	30/8772	0.00100304	0.01883666	0.01562211

BgRatio, background ratio; GO, Gene Ontology; BP, biological process; CC, cell component; MF, molecular function; GALMRDEGs, glycolysis and lactate metabolism-related differentially expressed genes; KEGG, Kyoto encyclopedia of genes and genomes. KEGG pathway data were obtained with permission from Kanehisa Laboratories (www.kegg.jp).

Table S2. GSEA enrichment analysis in high- and low-risk groups in TCGA-HNSC dataset

ID	setSize	enrichmentScore	NES	pvalue	p.adjust	qvalues
BUFFA_HYPOXIA_METAGENE	50	0.770100791	2.66590753	1.00E-10	1.47E-08	1.18E-08
WINTER_HYPOXIA_UP	93	0.694108722	2.624609752	1.00E-10	1.47E-08	1.18E-08
FOROUTAN_PRODRANK_TGFB_EMT_UP	184	0.569764848	2.301489159	1.00E-10	1.47E-08	1.18E-08
BOSCO_EPITHELIAL_DIFFERENTIATION_MODULE	63	-0.797256113	-2.279656654	1.00E-10	1.47E-08	1.18E-08
FOROUTAN_TGFB_EMT_UP	191	0.552181322	2.240532091	1.00E-10	1.47E-08	1.18E-08
PEREZ_TP53_AND_TP63_TARGETS	203	-0.587348461	-1.887437523	1.00E-10	1.47E-08	1.18E-08
BLANCO_MELO_COVID19_BRONCHIAL_EPITHELIAL_CELLS_SARS_COV_2_I NFECTION_UP	141	0.543213983	2.123618465	4.74E-10	6.07E-08	4.88E-08
KEGG_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450	68	-0.722286353	-2.085523355	1.57E-09	1.68E-07	1.36E-07
REACTOME_NEGATIVE_REGULATION_OF_NOTCH4_SIGNALING	54	0.716597072	2.505492931	5.73E-09	4.73E-07	3.81E-07
FOROUTAN_INTEGRATED_TGFB_EMT_UP	120	0.549934296	2.134398501	5.68E-09	4.73E-07	3.81E-07

YAUCH_HEDGEHOG_SIGNALING_PARACRINE_DN	243	0.440813429	1.876356678	8.33E-09	6.73E-07	5.42E-07
KEGG_DILATED_CARDIOMYOPATHY	90	0.599624478	2.260446483	1.57E-08	1.15E-06	9.29E-07
LIANG_SILENCED_BY_METHYLATION_2	57	0.690839561	2.437900302	1.80E-08	1.27E-06	1.02E-06
KAAB_FAILED_HEART_VENTRICLE_DN	44	0.748656764	2.540191646	2.13E-08	1.47E-06	1.19E-06
KEGG_DRUG_METABOLISM_CYTOCHROME_P450	70	-0.682946815	-1.982299728	7.07E-08	4.34E-06	3.50E-06
REACTOME_AUF1_HNRNP_D0_BINDS_AND_DESTABILIZES_MRNA	55	0.687083775	2.407172814	7.44E-08	4.53E-06	3.64E-06
LIN_SILENCED_BY_TUMOR_MICROENVIRONMENT	106	-0.621088	-1.88673074	1.04E-07	6.00E-06	4.83E-06
HELLER_HDAC_TARGETS_SILENCED_BY_METHYLATION_DN	277	-0.514324479	-1.678817567	1.03E-07	6.00E-06	4.83E-06
WINTER_HYPOXIA_DN	50	-0.709250191	-1.955864309	4.05E-07	1.95E-05	1.57E-05

GSEA, gene set enrichment analysis; NES, normalized enrichment score

Table S3. List of miRNAs associated with the key genes

miRNA	mRNA
hsa-miR-96-5p	<i>CAVI</i>
hsa-miR-199a-5p	<i>CAVI</i>
hsa-miR-199b-5p	<i>CAVI</i>
hsa-miR-124-3p	<i>CAVI</i>
hsa-miR-128-3p	<i>CAVI</i>
hsa-miR-506-3p	<i>CAVI</i>
hsa-miR-1271-5p	<i>CAVI</i>
hsa-miR-33a-5p	<i>LDHA</i>
hsa-miR-34a-5p	<i>LDHA</i>
hsa-miR-34c-5p	<i>LDHA</i>
hsa-miR-383-5p	<i>LDHA</i>
hsa-miR-449a	<i>LDHA</i>
hsa-miR-33b-5p	<i>LDHA</i>
hsa-miR-449b-5p	<i>LDHA</i>
hsa-miR-501-3p	<i>LDHA</i>
hsa-miR-135a-5p	<i>MB</i>
hsa-miR-135b-5p	<i>MB</i>
hsa-miR-21-5p	<i>PER2</i>
hsa-miR-24-3p	<i>PER2</i>
hsa-miR-25-3p	<i>PER2</i>
hsa-miR-27a-3p	<i>PER2</i>
hsa-miR-30a-5p	<i>PER2</i>
hsa-miR-32-5p	<i>PER2</i>
hsa-miR-92a-3p	<i>PER2</i>
hsa-miR-96-5p	<i>PER2</i>
hsa-miR-30c-5p	<i>PER2</i>
hsa-miR-30d-5p	<i>PER2</i>
hsa-miR-181a-5p	<i>PER2</i>
hsa-miR-181b-5p	<i>PER2</i>
hsa-miR-181c-5p	<i>PER2</i>
hsa-miR-27b-3p	<i>PER2</i>
hsa-miR-30b-5p	<i>PER2</i>
hsa-miR-34c-5p	<i>PER2</i>
hsa-miR-30e-5p	<i>PER2</i>
hsa-miR-363-3p	<i>PER2</i>
hsa-miR-367-3p	<i>PER2</i>
hsa-miR-449a	<i>PER2</i>
hsa-miR-495-3p	<i>PER2</i>
hsa-miR-92b-3p	<i>PER2</i>
hsa-miR-590-5p	<i>PER2</i>
hsa-miR-449b-5p	<i>PER2</i>
hsa-miR-1271-5p	<i>PER2</i>
hsa-miR-1286	<i>PER2</i>
hsa-miR-5688	<i>PER2</i>

Table S4. Baseline characteristics of patients with HNSCC

Characteristics	Overall
Clinical stage, n (%)	
Stage I	19 (3.9%)
Stage II	95 (19.4%)
Stage III	102 (20.9%)
Stage IV	273 (55.8%)
Pathologic stage, n (%)	
Stage I	25 (5.7%)
Stage II	69 (15.9%)
Stage III	79 (18.2%)
Stage IV	262 (60.2%)
Histological grade, n (%)	
G1	62 (12.8%)
G2	300 (62.1%)
G3	119 (24.6%)
G4	2 (0.4%)
Age, median (IQR)	61 (53, 69)
Sex, n (%)	
Male	369 (73.4%)
Female	134 (26.6%)

HNSCC: Head and neck squamous cell carcinoma, IQR: Interquartile range

Table S5. Overlap of glycolysis-related genes and lactate metabolism-related genes across three GEO datasets (GSE107591, GSE6631, and GSE9638)

Gene Symbol
<i>ABCC1</i>
<i>ADRB2</i>
<i>AKR1B1</i>
<i>BSG</i>
<i>CAVI</i>
<i>CDKN3</i>
<i>CS</i>
<i>DNM1L</i>
<i>FOXMI</i>
<i>FOXO3</i>
<i>GCDH</i>
<i>GCKR</i>
<i>GPT</i>
<i>HAGH</i>
<i>HIF1A</i>
<i>HK2</i>
<i>IL1B</i>
<i>KCNJ5</i>
<i>KRAS</i>
<i>LDHA</i>
<i>LDHB</i>
<i>MB</i>
<i>NPPA</i>
<i>P4HB</i>
<i>PARK7</i>
<i>PC</i>
<i>PER2</i>
<i>PTP4A2</i>
<i>SLC25A12</i>
<i>SLC2A1</i>
<i>SOD1</i>
<i>STAT3</i>
<i>TNF</i>