Volume 24, No 4 | International Journal of Radiation Research, October 2026

Identification and validation of candidate biomarkers in oral squamous cell carcinoma by whole-transcriptome sequencing and RT-qPCR, with exploratory CT radiomics analysis

Y. Zhang[1,2#], X. Li[3#], P. Hao[2], G. Zhang[2], Y. Wang[2], R. Gao[2], Z. Wei[2], C. Yin[2*]

[1]Chongqing Key Laboratory of Development and Utilization of Genuine Medicinal Materials in Three Gorges Reservoir Area, Chongqing, China
[2]College of Medical Technology, Chongqing Three Gorges Medical College, Chongqing, China
[3]Stomatological Hospital of Chongqing Medical University, Chongqing, China

► ***Original article***

***Corresponding author:**
Chongzhi Yin, Ph.D.,
***E-mail:** ZYL6892@163.com*

Received: February 2026
Final revised: June 2026
Accepted: June 2026

Int. J. Radiat. Res., October 2026; 24(4): 1207-1214

DOI: 10.61186/ijrr.24.4.35

Keywords: *Mouth neoplasms, squamous cell carcinoma of head and neck, biomarkers, tomography, x-ray computed, radiomics.*

Co-first authors.

ABSTRACT

Background: Oral squamous cell carcinoma (OSCC) is frequently diagnosed at advanced stages, and reliable noninvasive biomarkers remain limited. Transcriptomic profiling and CT-based radiomics are increasingly used to characterize tumor heterogeneity and may provide complementary molecular and imaging information in OSCC. ***Materials and Methods:*** Transcriptome sequencing was performed on three paired OSCC and adjacent normal tissues to identify differentially expressed genes (DEGs), followed by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. Fifteen additional paired samples were used for validation by real-time quantitative polymerase chain reaction (RT-qPCR). In a separate exploratory subanalysis, preoperative contrast-enhanced CT images from OSCC patients were manually segmented, and radiomic features were extracted to assess potential radiogenomic associations. ***Results:*** A total of 832 DEGs were identified, including 416 upregulated and 416 downregulated genes. Enrichment analyses showed that DEGs were mainly involved in cell cycle- and salivary secretion-related pathways. RT-qPCR confirmed significant differential expression of AQP5, LYZ, STATH, CDC7, and MCM5 (all $P < 0.05$), whereas CCND2 and E2F1 were not significant (both $P > 0.05$); exact P values are provided in the Results section and tables. In the exploratory CT analysis, high CDC7 expression was associated with irregular tumor margins and lymph node metastasis. ***Conclusion:*** AQP5, LYZ, STATH, CDC7, and MCM5 may serve as candidate biomarkers for OSCC. Exploratory CT radiomics findings suggest that imaging phenotypes may provide complementary information for molecular stratification, but these associations require validation in larger cohorts.

INTRODUCTION

Oral squamous cell carcinoma (OSCC) is among the most prevalent malignancies in the head and neck region, representing over 90% of oral cancers. OSCC poses significant risks to patient survival and quality of life, even when treated with surgery, radiotherapy, or chemotherapy. Despite improvements in surgical and adjuvant therapies, overall survival rates remain suboptimal, primarily due to late-stage diagnosis and tumor heterogeneity. Consequently, early detection is essential for improving patient outcomes [1-3].

In recent years, extensive efforts have been made to identify reliable biomarkers for early detection of OSCC. Transcriptomic studies have revealed that differentially expressed genes (DEGs) play a crucial role in tumor progression, metastasis, and prognosis. Several candidate biomarkers, including *AQP5, LYZ*, and *MCM* family genes, have been reported to be associated with OSCC pathogenesis [4-6].

Previous studies have suggested that two major functional axes contribute to oral squamous cell carcinoma (OSCC) development and progression: salivary secretion-related pathways and cell-cycle regulatory mechanisms. Among salivary secretion-related genes, aquaporin 5 (*AQP5*) has been shown to promote OSCC cell migration and invasion through dysregulated water transport and intracellular signaling. Similarly, lysozyme (*LYZ*) and statherin (*STATH*), both components of the oral fluid proteome, have been identified as differentially expressed in OSCC patients and have been proposed as potential salivary biomarkers for early detection [7].

On the other hand, cell-cycle-related biomarkers including *CDC7*, *MCM5*, *CCND2*, and the transcription factor *E2F1* play established roles in DNA replication initiation and G1/S phase transition. *E2F1*, in particular, functions as a master regulator of cell-cycle progression and has been implicated in OSCC proliferation when its expression is deregulated. Collectively, these findings indicate that both salivary secretion-related genes and cell-cycle-related genes may be involved in OSCC pathogenesis; however, their validation across independent tissue samples

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]

and their relationship with imaging phenotypes remain insufficiently characterized [8, 9].

In parallel, imaging-based approaches, particularly computed tomography (CT) and radiomics, have gained attention for their ability to capture tumor heterogeneity and noninvasively characterize tumor phenotypes [10]. Recent studies suggest that integrating imaging features with molecular data can improve diagnostic accuracy and provide insights into imaging-genomic correlations [5, 11- 13]. CT-based tumor imaging is clinically essential in OSCC management because it supports the assessment of tumor extent, local invasion, depth of invasion (DOI), nodal involvement, and treatment planning. As demonstrated by Blatt *et al.*, CT and MRI currently remain the reference standards for evaluating the primary tumor site and cervical lymph node status, while recent advances in radiomics enable the extraction of high-dimensional quantitative features from standard-of-care CT images that are imperceptible to the human eye. In the context of OSCC, CT-derived radiomic features have shown significant correlations with cervical lymph node metastasis and underlying transcriptomic alterations [14-17]. For instance, Jin *et al.* demonstrated that CT radiomic features correlate with differentially expressed RNAs involved in salivary secretion signaling pathways, providing direct evidence of a radiogenomic link in OSCC. Therefore, linking CT-derived phenotypes with molecular alterations may improve the biological interpretation of imaging findings and provide a rationale for future imaging-guided risk stratification in OSCC. This integrative perspective aligns with the scope of IJRR, where tumor imaging and radiation treatment planning represent core research priorities [18, 19]. However, most previous studies have focused either on molecular biomarkers or imaging features independently. The relationship between external phenotypic features, such as lip dynamics, internal radiological characteristics, and underlying gene expression profiles remains poorly understood [20-22].

The novelty of this study lies in its integrative design, combining whole-transcriptome screening, RT-qPCR validation, and exploratory CT-based imaging assessment in OSCC. Rather than evaluating molecular biomarkers and imaging findings separately, this study explored whether validated gene-expression alterations, particularly in salivary secretion- and cell-cycle-related pathways, show preliminary associations with CT-derived tumor phenotypes. This radiogenomic perspective may provide a hypothesis-generating basis for future noninvasive molecular risk stratification and imaging -guided assessment in OSCC. Accordingly, this study aimed to identify and validate candidate molecular biomarkers for OSCC through whole-transcriptome sequencing and RT-qPCR, and to investigate whether CT radiomic features demonstrate preliminary associations with validated molecular subtypes.

MATERIALS AND METHODS

Study design and patients

This retrospective observational study included patients with histopathologically confirmed oral squamous cell carcinoma (OSCC) who underwent primary surgical resection at hospital of Chongqing Three Gorges Medical College, Chongqing 404120, China. Eligible patients were adults with previously untreated OSCC and available paired tumor and adjacent non-tumor tissue samples. Patients were excluded if they had received neoadjuvant chemotherapy, radiotherapy, immunotherapy, or other anticancer treatment before surgery, or if tissue quality was insufficient for RNA extraction. Paired tumor and adjacent non-tumor tissues were collected intraoperatively. Three paired samples were used for whole-transcriptome sequencing as the discovery cohort, and an independent set of 15 paired samples was used for RT-qPCR validation. Clinical and pathological data, including age, sex, tumor site, pathological stage, lymph node status, and histological grade, were retrieved from medical records.

The current study was performed in line with the Helsinki Declaration. The study protocol was approved by the Chongqing Three Gorges Medical College, Chongqing 404120, China, Approval No. [20210003], date of approval [202103]). Written informed consent was obtained from all subjects before their recruitment in the study.

RNA extraction

Total RNA was extracted from frozen OSCC tumor tissues and matched adjacent non-tumor tissues using TRIzol Reagent (Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's protocol. Approximately 50 mg of frozen tissue was ground in liquid nitrogen and homogenized in 1 mL TRIzol Reagent. After incubation at room temperature for 5 min, 0.2 mL chloroform was added. The mixture was shaken vigorously for 15 s, incubated at room temperature for 3 min, and centrifuged at 12,000 × g for 15 min at 4°C. The upper aqueous phase was transferred to a new RNase-free tube and mixed with 0.5 mL isopropanol. After incubation at room temperature for 10 min, the samples were centrifuged at 12,000 × g for 10 min at 4°C. The RNA pellet was washed twice with 75% ethanol, air-dried, dissolved in RNase-free water, and stored at −80°C until further use. RNA concentration and purity were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA samples with A260/A280 ratios of 1.8–2.1 were considered acceptable. RNA integrity was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies,

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]

Santa Clara, CA, USA), and samples with RIN ≥ 7.0 were used for transcriptome sequencing. RNA samples of sufficient purity and integrity were used for RT-qPCR validation.

Whole-transcriptome sequencing and bioinformatics analysis

For discovery analysis, total RNA from three paired OSCC tumor tissues and matched adjacent non-tumor tissues was used for whole-transcriptome sequencing. RNA quality was assessed before library construction, and samples with adequate purity and integrity were used for sequencing. RNA-seq libraries were prepared using an Illumina-compatible library preparation according to the manufacturer's instructions. The libraries were sequenced on an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA) using a paired-end 150-bp sequencing strategy. Raw reads were processed to remove adaptor sequences, low-quality reads, and reads containing excessive unknown bases. The resulting clean reads were aligned to the human reference genome GRCh38 using HISAT2 version 2.2.1 (Johns Hopkins University, Baltimore, MD, USA). Sequencing quality, read distribution, and mapping characteristics were assessed using RSeQC version 4.0.0 and Qualimap version 2.2.1. Gene-level counts were generated using featureCounts version 2.0.1 based on Ensembl gene annotation. Differential expression analysis between OSCC tumor tissues and matched adjacent non-tumor tissues was performed using the DESeq2 package version 1.34.0 in R version 4.1.0.

Differential expression analysis between OSCC tumor tissues and matched adjacent non-tumor tissues was performed using the DESeq2 package. Genes with |log2 fold change| > 1 and false-discovery-rate-adjusted P value < 0.05 were considered differentially expressed. Nominal P values and FDR-adjusted P values were reported for representative DEGs in the Results section and tables.

Functional enrichment analysis of differentially expressed genes was performed using Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway databases. Protein-protein interaction analysis was performed using the STRING database and visualized using Cytoscape. Candidate genes for RT-qPCR validation were selected according to differential expression level, pathway enrichment results, and biological relevance to OSCC, especially salivary secretion-related and cell-cycle-related pathways [23,24].

RT-qPCR validation

Candidate genes identified from whole-transcriptome sequencing were validated in an independent cohort of 15 paired OSCC tumor tissues and adjacent non-tumor tissues by RT-qPCR. The selected genes included *AQP5, LYZ, STATH, CDC7, MCM5, CCND2, and E2F1* because they were involved in salivary secretion- and cell-cycle-related pathways and showed differential expression in the sequencing analysis.

Total RNA was reverse-transcribed into cDNA using the PrimeScript RT reagent Kit (Takara Bio Inc., Shiga, Japan) according to the manufacturer's instructions. RT-qPCR was performed using TB Green Premix Ex Taq II (Takara Bio Inc., Shiga, Japan.)

RT-qPCR was performed using TB Green Premix Ex Taq II (Takara Bio Inc., Shiga, Japan) on an ABI 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). GAPDH was used as the internal reference gene. Primer sequences for *AQP5, LYZ, STATH, CDC7, MCM5, CCND2, E2F1*, and *GAPDH* are listed in table 1.

GAPDH was used as the internal reference gene. Each RT-qPCR reaction was performed in triplicate in a final volume of 20 μL. The amplification protocol consisted of initial denaturation at 95°C for 30 s, followed by 40 cycles at 95°C for 5 s and 60°C for 30 s. Melting-curve analysis was performed to confirm amplification specificity, and no-template controls were included in each run. Relative expression levels were calculated using the 2−ΔΔCt method. For each sample, ΔCt was calculated as Ct (target gene) − Ct (GAPDH), and ΔΔCt was calculated as ΔCt (OSCC tumor tissue) − ΔCt (adjacent non-tumor tissue). Relative expression was then expressed as $2^{-\Delta\Delta Ct}$.

Table 1. Primer sequences used for RT-qPCR validation.

Gene	Accession No.	Forward primer (5′–3′)	Reverse primer (5′–3′)
AQP5	NM_001651	TACGGTGTGGCAC-CGCTCAATG	AGTCAGTG-GAGGCGAAGATGCA
LYZ	NM_000239	ACTACAATGCTG-GAGACAGAAGC	GCACAAGCTACAG-CATCAGCGA
STATH	NM_003154	TTTTGCGTAGAATT-GGAAGATTCGG	AATCATGTCCTG-CAGTTACTGATG
CDC7	NM_003503	GGAAAACTGCCAG-TTCTTGCCC	GGCACTTTGTCAA-GACCTCTGG
MCM5	NM_006739	GACTTACTCGCCGAG GAGACAT	TGCTGCCTTTCCCAG ACGTGTA
CCND2	NM_001759	GAGAAGCTGTCTCT GATCCGCA	CTTCCAGTT-GCGATCATCGACG
E2F1	NM_005225	GGACCTG-GAAACTGACCATCAG	CAGTGAGGTCTC ATAGCGTGAC
GAPDH	NM_002046	GTCTCCTCTGACTTC AACAGCG	ACCACCCTGTT-GCTGTAGCCAA

Abbreviations: RT-qPCR, real-time quantitative polymerase chain reaction; OSCC, oral squamous cell carcinoma; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

CT acquisition and exploratory radiomics analysis

For the exploratory imaging subanalysis, preoperative contrast-enhanced CT images were collected from OSCC patients with available RT-qPCR validation data. CT examinations were performed before surgical resection using a multidetector CT scanner. The scanning parameters were as follows: tube voltage, 120 kVp; tube current, automatic tube current modulation; matrix size, 512 × 512; slice thickness, 1.0–5.0 mm; and soft-tissue reconstruction

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]

kernel. Contrast-enhanced CT images were obtained after intravenous administration of iodinated contrast medium. Tumor regions of interest were manually delineated on contrast-enhanced CT images by two radiologists experienced in head-and-neck imaging and were reviewed by a third senior radiologist to resolve discrepancies. Segmentation was performed using 3D Slicer version 5.2.2 or equivalent radiological segmentation software. The radiologists were blinded to the RT-qPCR results during image segmentation.

Before radiomic feature extraction, CT images were resampled to isotropic 1 × 1 × 1 mm^3 voxels, and gray-level discretization was performed using a fixed bin width of 25 HU. Radiomic features were extracted using PyRadiomics version 3.0.1 implemented in Python. Extracted features included shape, first-order intensity, and texture features, including gray-level co-occurrence matrix, gray-level run-length matrix, gray-level size-zone matrix, neighboring gray-tone difference matrix, and gray-level dependence matrix features.

The CT analysis was designed as an exploratory radio genomic sub analysis. Validated RT-qPCR genes were used as molecular variables, and patients were divided into high- and low-expression groups according to the median relative expression level of each gene. CT-derived qualitative phenotypes, including tumor margin, border definition, and cervical lymph node status, were compared between expression groups. Because of the small sample size, no predictive model was developed, and the analysis was interpreted as hypothesis-generating.

Statistical analysis

Statistical analyses were performed using SPSS Statistics version 17.0 (IBM, Armonk, NY, USA). Continuous variables are presented as mean ± standard deviation. Pairwise comparisons between OSCC and adjacent normal tissues were performed using paired t-tests when the data were normally distributed, or the corresponding nonparametric test when they were not. Functional enrichment analysis was performed using GO and KEGG pathway databases as described in the bioinformatics analysis section. For comparisons involving more than two groups, ANOVA was used when appropriate. Categorical associations in the CT subanalysis were evaluated using chi-square or Fisher's exact test, depending on sample size and expected frequencies. A two-sided P value < 0.05 was considered statistically significant. Because this was an exploratory retrospective study, no formal sample-size calculation was performed.

RESULTS

OSCC differential gene expression analysis

Whole-transcriptome sequencing was performed on three paired OSCC tumor tissues and adjacent non-tumor tissues.

Using the predefined threshold of |log2 fold change| > 1 and nominal P value <0.05, a total of 832 differentially expressed genes (DEGs) were identified, including 416 upregulated and 416 downregulated genes in OSCC tissues compared with adjacent non-tumor tissues. Representative DEGs, together with their log2 fold changes and exact P values, are summarized in table 2. Because of the exploratory nature of the discovery cohort, these results were further prioritized for RT-qPCR validation. The overall expression pattern of representative DEGs across the discovery cohort is shown in the clustered heatmap (figure 1). A volcano plot summarizing the distribution of all DEGs according to fold change and statistical significance is presented in figure 2. The volcano plot highlights the balanced DEG profile reported in this study, consisting of 416 upregulated and 416 downregulated genes, and identifies the principal candidate genes selected for downstream validation.

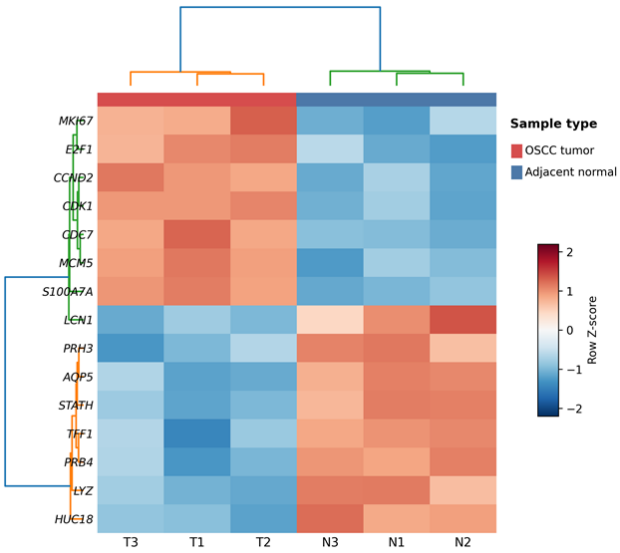


Figure 1. Clustered heatmap of representative differentially expressed genes in OSCC and adjacent non-tumor tissues.

Table 2. Differentially expressed genes.

Gene id	GeneName	log2FC/ FC score	−log10 (P value)	result
RBSG0001	*TFFI*	-15	7.756	Downregulated
RBSG0002	*PRB4*	-15	8.406	Downregulated
RBSG0003	*LCN1*	-14	2.886	Downregulated
RBSG0004	*S100A7A*	+16	3.26	Upregulated
RBSG0005	*HUC18*	-17	4.36	Downregulated
RBSG0006	*PRH3*	-18	6.458	Downregulated

Abbreviations: OSCC, oral squamous cell carcinoma; log2FC, log2 fold change. Positive log2FC indicates upregulation in OSCC tissues, whereas negative log2FC indicates downregulation in OSCC tissues compared with adjacent non-tumor tissues. The −log10(P value) column is shown for ranking and visualization purposes; statistical significance was determined using the original P values. If available, FDR-adjusted P values should be reported in the final differential-expression table.

The heatmap shows the expression pattern of representative differentially expressed genes (DEGs) identified by whole-transcriptome sequencing in

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]

three paired oral squamous cell carcinoma (OSCC) tumor tissues and matched adjacent non-tumor tissues. Each row represents a gene and each column represents an individual sample. Expression values are shown after row-wise normalization to emphasize relative expression differences across samples. Red indicates relatively higher expression and blue indicates relatively lower expression. Hierarchical clustering was used to visualize the separation between OSCC tumor tissues and adjacent non-tumor tissues and to illustrate the overall expression pattern of genes prioritized for downstream validation.

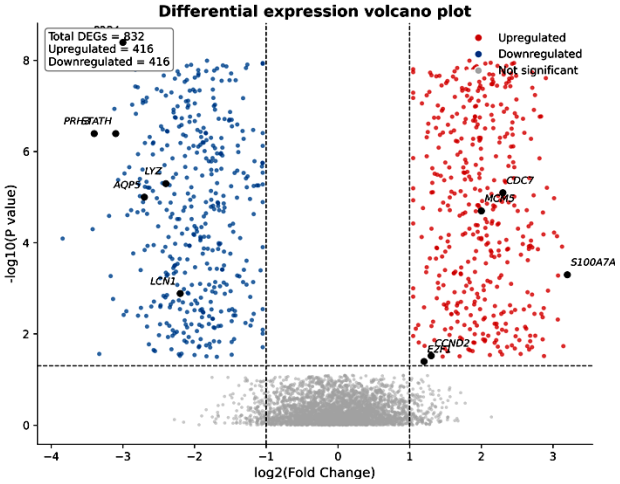


Figure 2. Volcano plot of differentially expressed genes identified by RNA-seq. The volcano plot summarizes the differential gene expression analysis between OSCC tumor tissues and matched adjacent non-tumor tissues. Each dot represents one gene. The x-axis indicates log2 fold change, and the y-axis indicates −log10(P value). Genes located on the right side of the plot represent upregulated genes in OSCC tissues, whereas genes located on the left side represent downregulated genes. Dashed vertical lines indicate the fold-change threshold, and the horizontal dashed line indicates the statistical significance threshold. Upregulated and downregulated DEGs are displayed in different colors, while non-significant genes are shown separately. Candidate genes selected for RT-qPCR validation, including AQP5, LYZ, STATH, CDC7, MCM5, CCND2, and E2F1, are highlighted to show their position within the global RNA-seq differential expression profile.

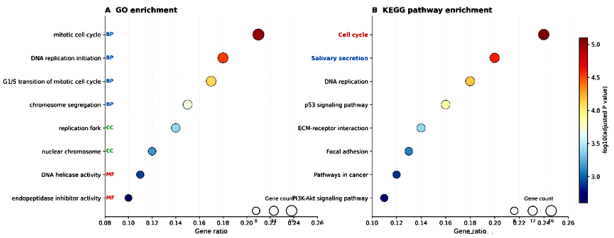


Figure 3. Gene Ontology and KEGG pathway enrichment analysis of differentially expressed genes.

GO and KEGG enrichment analysis

Among these enriched pathways, the salivary secretion pathway and cell-cycle pathway were selected for further validation because they were biologically relevant to OSCC progression and included genes with marked differential expression. Accordingly, *AQP5, LYZ, STATH, CDC7, MCM5, CCND2,* and *E2F1* were selected for RT-qPCR validation.

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to identify biological processes and pathways associated with the DEGs detected by whole-transcriptome sequencing. Panel A presents representative GO enrichment terms, including biological process, cellular component, and molecular function categories. Panel B presents representative KEGG pathway enrichment results. Bubble size indicates the number of DEGs enriched in each term or pathway, and color intensity represents enrichment significance. The salivary secretion and cell-cycle pathways were highlighted because they were biologically relevant to OSCC progression and contained genes selected for RT-qPCR validation, including *AQP5*, *LYZ*, *STATH*, *CDC7*, *MCM5*, *CCND2*, and *E2F1*.

To provide a clearer biological interpretation of the enriched pathways, a pathway-level interaction map was generated for the candidate genes selected for RT-qPCR validation (figure 4). The map was organized into two functional modules. The salivary secretion module included *AQP5*, *LYZ*, and *STATH*, which were downregulated in OSCC tissues and are related to water transport, antimicrobial defense, lubrication, and salivary secretory homeostasis. The cell-cycle / DNA replication module included *CCND2, E2F1, CDC7, and MCM5*, which are involved in G1/S transition, transcriptional activation of replication-related genes, replication-origin activation, and DNA synthesis. This pathway-level representation clarifies the biological context of the candidate genes and their inferred functional consequences in OSCC.

RT-qPCR verification

RT-qPCR validation was performed in an independent cohort of 15 paired OSCC tumor and adjacent non-tumor tissue samples. The RT-qPCR results were consistent with the sequencing data for five of the seven selected candidate genes. *AQP5, LYZ, and STATH* were downregulated in OSCC tissues compared with matched adjacent non-tumor tissues, whereas *CDC7 and MCM5* were upregulated (all $P<0.05$). Reduced expression was observed in 66.67% of samples for *AQP5* (10/15), 73.33% for *LYZ* (11/15), and 73.33% for *STATH* (11/15). Increased expression was observed in 80.00% of samples for *CDC7* (12/15) and 66.67% for *MCM5* (10/15). In contrast, *CCND2* and *E2F1* did not show statistically significant differential expression in the validation cohort ($P>0.05$). Increased expression was observed in 46.67% of samples for *CCND2* (7/15) and 46.67% for *E2F1* (7/15), indicating that these two genes were not consistently dysregulated in the validation cohort. These findings suggest that *AQP5, LYZ, STATH, CDC7*, and *MCM5* represent the most reproducible candidate biomarkers in the present dataset, whereas *CCND2 and E2F1* should be interpreted as pathway-

[Downloaded from ijrr.com on 2026-11-23]
[DOI: 10.61186/ijrr.24.4.35]

associated genes rather than validated biomarkers (Figure 5).

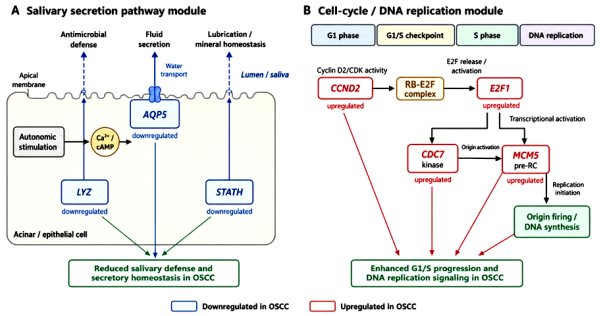


Figure 4. Pathway-level interpretation of candidate genes identified in OSCC. This schematic pathway figure illustrates the biological relationships among the candidate genes selected from the transcriptomic and enrichment analyses. Panel A shows the salivary secretion pathway module. In this module, AQP5, LYZ, and STATH are shown as downregulated genes associated with impaired water transport, antimicrobial defense, lubrication, and salivary secretory homeostasis. Panel B shows the cell-cycle / DNA replication pathway module. In this module, CCND2, E2F1, CDC7, and MCM5 are positioned within the G1/S transition and DNA replication initiation pathway. The figure indicates the inferred direction of gene interactions and pathway-level consequences, including reduced salivary defense and enhanced G1/S progression and DNA replication signaling in OSCC. Blue indicates downregulated genes, red indicates upregulated genes, and green boxes indicate inferred pathway-level biological outcomes.

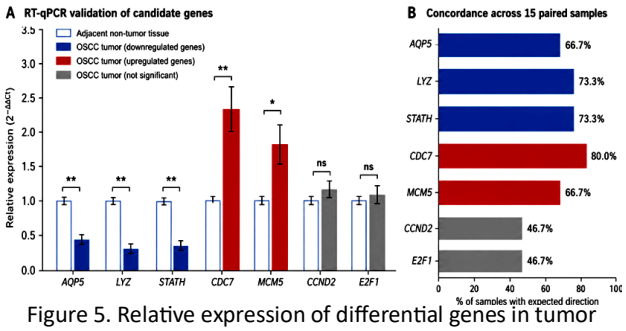


Figure 5. Relative expression of differential genes in tumor tissue. RT-qPCR was performed in an independent validation cohort of 15 paired OSCC tumor and adjacent non-tumor tissue samples to validate the candidate genes identified by transcriptome sequencing. Panel A shows the relative expression levels of AQP5, LYZ, STATH, CDC7, MCM5, CCND2, and E2F1 in OSCC tumor tissues compared with matched adjacent non-tumor tissues. Relative expression was calculated using the 2−ΔΔCt method, with GAPDH used as the internal reference gene. AQP5, LYZ, and STATH showed reduced expression in OSCC tissues, whereas CDC7 and MCM5 showed increased expression. CCND2 and E2F1 did not show statistically significant differential expression in the validation cohort. Panel B shows the percentage of paired samples demonstrating the expected direction of expression change for each gene. Error bars represent variability across samples. Significance labels indicate statistical comparison between paired OSCC and adjacent non-tumor tissues: *P < 0.05; **P < 0.01; ns, not significant.

Exploratory CT radiogenomic analysis

In the exploratory CT radiogenomic subanalysis, patients were categorized into high- and low-expression groups according to the median relative expression level of each validated gene. Higher *CDC7* expression showed a descriptive trend toward irregular tumor margins and cervical lymph node metastasis on preoperative contrast-enhanced CT. In addition, lower *AQP5* expression was more frequently observed in tumors with poorly defined borders. Because of the small sample size, these findings should be interpreted as preliminary and hypothesis-generating rather than definitive radiogenomic associations.

DISCUSSION

Oral squamous cell carcinoma is a biologically heterogeneous malignancy that develops through a multistep process of genetic and epigenetic alterations in the oral mucosa. OSCC progression is driven by dysregulated signaling pathways, abnormal cell-cycle control, impaired DNA damage responses, altered apoptosis, and changes in the tumor microenvironment [25-27]. Transcriptome sequencing technology can quickly and comprehensively analyze differential expression across various RNA types. In this study, whole-transcriptome sequencing followed by RT-qPCR validation identified a set of candidate biomarkers that may contribute to the molecular stratification of OSCC [21, 22, 28, 29].

The salivary secretion-related findings of the present study should be interpreted in the context of the complex relationship between tumor-tissue transcriptomics and salivary biomarker profiles. This pattern is biologically plausible because salivary biomarkers have long been investigated in oral cancer, and both lysozyme and statherin are components of the salivary defense system. *AQP5* has previously been reported to be overexpressed in OSCC by immunohistochemistry, whereas our transcriptomic findings suggest that its expression may vary by sample type, assay platform, and tumor context. Similarly, discrepancies between our tissue-based results for *LYZ* and salivary studies in OSCC may reflect the difference between tumor tissue transcription and salivary protein abundance. By contrast, the downregulation of *STATH* is consistent with prior reports showing reduced statherin levels in saliva from patients with precancerous and cancerous oral lesions, supporting its potential as a biomarker in oral carcinogenesis [30-32]. The cell-cycle-associated genes *CDC7* and *MCM5* were upregulated in our cohort and confirmed by RT-qPCR, consistent with prior reports implicating both genes in OSCC progression [33].

CDC7 has previously been proposed as an OSCC biomarker and has also been associated with replication-stress tolerance and DNA damage response signaling in oral cancer, whereas *MCM5* has

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]

been reported to be overexpressed in OSCC and may have value in early diagnosis or assessment of malignant potential. In contrast, although *CCND2* and *E2F1* were identified at the pathway level, their RT-qPCR results were not statistically significant in our validation cohort; therefore, they should be interpreted as pathway-associated signals rather than confirmed biomarkers in the present study [34-37].

While the exploratory CT radiomics sub-analysis suggested an association between high *CDC7* expression and abnormal tumor margins and lymph node metastasis, low *AQP5* expression and undefined borders, these results align with recent literature on the use of CT-based radiomic biomarkers in OSCC cases. Therefore, based on our results, we hypothesize that cancer imaging characteristics can complement gene expression data [38, 39]. Nonetheless, as the current study used a small sample for the exploratory analysis, these findings need to be validated in larger samples. Several limitations have to be taken into account. First, the small size of the discovery cohort increases the risk of selection bias and reduces the reliability of the results. Second, the RT-qPCR validation sample size was limited, and the validation cohort was drawn from a single center. Third, the exploratory nature of CT analysis means these findings should not yet be regarded as a radiogenomic model. Fourth, as the study did not include external validation or survival analysis, it is premature to conclude on the prognostic significance of the identified biomarkers.

CONCLUSION

In conclusion, this study proposed the following genes as potential biomarkers for OSCC: *AQP5, LYZ, STATH, CDC7*, and *MCM5*. In addition, a novel radiomics approach using CT images was introduced, indicating that the imaging phenotype can serve as an additional means for biomarker identification.

Acknowledgments: The authors would like to thank all patients who participated in this study and the clinical staff of Chongqing Three Gorges Medical College for their assistance in sample collection and data acquisition. The authors also acknowledge the support of the laboratory team for their technical assistance.

Funding: This work was supported by the Natural Science Foundation of Chongqing Science and Technology Bureau (No. CSTB2023NSCQ-LMX0017), the Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN202202702, KJQN202102712), the Key Project of Chongqing Key Laboratory of Development and Utilization of Genuine Medicinal Materials in Three Gorges Reservoir Area (No. Sys20210003), the Natural Science Major Project of Chongqing Three Gorges Medical College (No. XJ2021000702), the Project of Traditional Chinese Medicine inheritance and innovation team of Chongqing Health Commission, and the Youth Top Talent Project of Chongqing Three Gorges Medical College.

Conflicts of interest: None.

Ethical considerations: The current study was performed in line with the Helsinki Declaration. The study protocol was approved by the Chongqing Three Gorges Medical College, Chongqing 404120, China, Approval No. [20210003], date of approval [202103]). Written informed consent was obtained from all subjects before their recruitment in the study.

Author contributions: Conception and design were done by Y.Z. and X.L. Data acquisition and sample preparation were carried out by P.H. and G.Z. Bioinformatics and statistics analysis were done by Y.W. and R.G. Data interpretation and validation experiments were performed by Z.W. and C.Y. Y.Z. was responsible for writing the first draft, and all authors revised the paper and finalized it.

Use of artificial intelligence: The authors declare that no artificial intelligence tools were used in the generation, analysis, or interpretation of the data. AI-assisted tools were only used for language editing and improvement of manuscript readability, without affecting the scientific content.

REFERENCES

1. Gaw G and Gribben M (2022) Can we detect biomarkers of oral squamous cell carcinoma from saliva or mouth swabs? *Evid Based Dent,* ***23****(1): 32-33*.
2. Johnson DE, Burtness B, Leemans CR, Lui VWY, *et al.* (2019) Head and neck squamous cell carcinoma. *J Clin Oncol,* ***37****(1): 67-83*.
3. Biological consequences of cancer radiotherapy in the context of oral squamous cell carcinoma (2019) *Front Oncol,* ***9****: 415*.
4. Wen W, Li R, Wang D, Chai L (2024) Exploring the prognostic landscape of oral squamous cell carcinoma through mitochondrial damage-related genes. *BMC Med Genomics,* ***17****: 274*.
5. Itonaga T, Mikami R, Zama T, Okada Y, *et al.* (2024) Impact of diffusion kurtosis imaging as an early prognostic factor following radiotherapy in patients with head and neck squamous cell carcinoma. *Int J Radiat Res,* ***22****(3): 603-608*.
6. Gong H, Liu Z, Yuan C, Luo Y, *et al.* (2024) Identification of cuproptosis-related lncRNAs with the significance in prognosis and immunotherapy of oral squamous cell carcinoma. *Comput Biol Med,* ***171****: 108198*.
7. Zhou Y and Liu Z (2023) Saliva biomarkers in oral disease. *Clin Chim Acta,* ***548****: 117503*.
8. Bastías D, Maturana A, Marín C, Martínez R, Niklander SE (2024) Salivary biomarkers for oral cancer detection: an exploratory systematic review. *Int J Mol Sci,* ***25****(5): 2634*.
9. Cheng Y, Wang L, Zhang S, Jian W, *et al.* (2025) The investigation of NFκB inhibitors to block cell proliferation in OSCC cells lines. *Curr Med Chem,* ***32****(33): 7314-7326*.
10. Ye Z, Zhang Y, Liang Y, Lang J, *et al.* (2022) Cervical cancer metastasis and recurrence risk prediction based on deep convolutional neural network. *Curr Bioinform,* ***17****: 164-173*.
11. Liu Y, He M, Tang H, *et al.* (2024) Single-cell and spatial transcriptomics reveal metastasis mechanism and microenvironment remodeling of lymph node in osteosarcoma. *BMC Med,* ***22****: 200*.

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]

12. Pfaehler E, Schindele A, Dierks A, Busse C, *et al.* (2025) Value of PET radiomic features for diagnosis and recurrence prediction of newly diagnosed oral squamous cell carcinoma. *Sci Rep,* ***15****: 17475*.
13. Wang Y, Xu Y, Song J, Liu X, *et al.* (2024) Tumor cell-targeting and tumor microenvironment-responsive nanoplatforms for the multimodal imaging-guided photodynamic/photothermal/ chemodynamic treatment of cervical cancer. *Int J Nanomedicine,* ***19****: 5837-5858*.
14. Blatt S, Ziebart T, Krüger M, Pabst AM (2016) Diagnosing oral squamous cell carcinoma: How much imaging do we really need? A review of the current literature. *J Craniomaxillofac Surg,* ***44****(5): 538-549*.
15. Michcik A, Polcyn A, Garbacewicz Ł, Wach T, *et al.* (2023) Efficacy of computed tomography and magnetic resonance imaging in the assessment of depth of invasion in oral squamous cell carcinoma: analysis of 125 clinical cases. *Diagnostics (Basel),* ***13****(23): 3578*.
16. Jiang R, Yin X, Yang P, Cheng L, *et al.* (2024) A transformer-based weakly supervised computational pathology method for clinical-grade diagnosis and molecular marker discovery of gliomas. *Nat Mach Intell,* ***6****(8): 876-891*.
17. Yang Y, Yan JW, Luo Y, Wu S, *et al.* (2025) An optimal imaging modality based on different predictive risk factors for evaluating local staging of rectal cancer preoperatively: a comparative study of MRI and endorectal ultrasound. *Quant Imaging Med Surg,* ***15*** *(11): 10516-10525*.
18. Jin N, Qiao B, Zhao M, Li L, *et al.* (2023) Predicting cervical lymph node metastasis in OSCC based on computed tomography imaging genomics. *Cancer Med,* ***12****(18): 19260-19271*.
19. Mirzaeiyan M, Najafizadeh N, Saeb M, Shahbazi-Gahrouei D (2024) Dosiomics-based comparison of dose distributions in nasopharyngeal cancer patients: 3D-CRT versus Tomotherapy. *Int J Radiat Res,* ***22****(2): 419-425*.
20. Almangush A, Alabi RO, Pirinen M, Mäkitie A, Leivo I (2024) Accumulating evidence from meta-analyses of prognostic studies on oral cancer: towards biomarker-driven patient selection. *BMC Cancer,* ***24****: 1517*.
21. Corti A, Lenoci D, Corino VD, *et al.* (2025) Interplay between MRI radiomics and immune gene expression signatures in oral squamous cell carcinoma. *Sci Rep,* ***15****: 9821*.
22. Kim DW, Jeon H, Ki Y, et al. (2025) Therapeutic dose effects of target-volume changes during head and neck radiotherapy [J]. *International Journal of Radiation Research, 23(3):635-641.*
23. Gan X (2025) GraphService: Topology-aware constructor for large-scale graph applications. *ACM Trans Archit Code Optim,* ***22****(1): 2*.
24. Gan X (2025) TianheGraph: Topology-aware graph processing. ACM Trans Archit Code Optim.
25. Ning Q, Yang T, Guo X, Huang Y, *et al.* (2023) CHB patients with rtA181T-mutated HBV infection are associated with higher risk hepatocellular carcinoma due to increases in mutation rates of tumour suppressor genes. *J Viral Hepat,* ***30****(12): 951-958*.
26. Wang Y, Wang Q, Yang TW, Yin JM, *et al.* (2023) Analysis of immune and inflammatory microenvironment characteristics of noncancer end-stage liver disease. *J Interferon Cytokine Res,* ***43****(2): 86-97*.
27. Li F, Han Y, Ou F, Deng L, *et al.* (2026) Spatial heterogeneity of MDSCs mediated by ANXA1-FPRs signaling drives immune suppression in OSCC progression. *Nat Commun,* ***17****(1): 2535*.
28. Zhu Z, Liu H, Xu B. (2023) Digital 320-slice volume computed tomography scanning in the diagnosis of gastrointestinal lipoma [J]. *International Journal of Radiation Research,* ***21****(2):349-352.*
29. Pérez-Sayáns M, Somoza-Martín JM, Barros-Angueira F, *et al.* (2009) Genetic and molecular alterations associated with oral squamous cell cancer. *Oncol Rep,* ***22****(6): 1277-1282*.
30. Ishimoto S, Wada K, Usami Y, *et al.* (2012) Differential expression of aquaporin 5 and aquaporin 3 in squamous cell carcinoma and adenoid cystic carcinoma. *Int J Oncol,* ***41****(1): 67-75*.
31. Meyer P and Zechel T (2001) Quantitative studies of lysozyme and phosphohexose isomerase enzymes in mixed saliva in oral squamous epithelial carcinoma. *HNO,* ***49****(8): 626-629*.
32. Mendola JD, Cronin-Golomb A, Corkin S, Growdon JH (1995) Prevalence of visual deficits in Alzheimer's disease. *Optom Vis Sci,* ***72****(3): 155-167*.
33. Peng X, Liu Z, Luo C, Sun R, *et al.* (2025) PSMD12 promotes hepatocellular carcinoma progression by stabilizing CDK1. *Front Immunol,* ***16****: 1581398*.
34. Cheng AN, Jiang SS, Fan CC, *et al.* (2013) Increased Cdc7 expression is a marker of oral squamous cell carcinoma and contributes to resistance to DNA-damaging agents. *Cancer Lett,* ***337****(2): 218-225*.
35. Jin S, Ma H, Yang W, *et al.* (2018) Cell division cycle 7 is a potential therapeutic target in oral squamous cell carcinoma and is regulated by E2F1. *J Mol Med (Berl),* ***96****(6): 513-525*.
36. Kim YD, Jeon EH, Kim YS, *et al.* (2015) Molecular genetic study of novel biomarkers for early diagnosis of oral squamous cell carcinoma. *Med Oral Patol Oral Cir Bucal,* ***20****(2): e167-e179*.
37. Yu SY, Wang YP, Chang JY, *et al.* (2014) Increased expression of MCM5 is associated with aggressive progression and poor prognosis of oral squamous cell carcinoma. *J Oral Pathol Med,* ***43*** *(5): 344-349*.
38. Ling X, Alexander GS, Molitoris J, *et al.* (2023) Identification of CT-based radiomic biomarkers for overall survival prediction in oral cavity squamous cell carcinoma. *Sci Rep,* ***13****: 21774*.
39. Ma Q, Ren J, Wang R, Yuan Y, *et al.* (2024) Predicting response to immunotherapy in oral squamous cell carcinoma via a CT-based radiomics model. *BMC Med Imaging,* ***24****: 266*.

[Downloaded from ijrr.com on 2026-11-23]

[DOI: 10.61186/ijrr.24.4.35]