

# Early Prediction of Lymph Node Metastasis in Head and Neck Squamous Cell Carcinoma Using a Three-Gene Classifier

**Nimita Saji**

St.John's Medical College Hospital

**Angel Jaimi**

St.John's Medical College Hospital

**Lisa Anna Jiju**

St.John's Research Institute

**Madhumathy G Nair**

St.John's Research Institute

**Gayatri Ravikumar**

[gayatri.ravikumar@gmail.com](mailto:gayatri.ravikumar@gmail.com)

St.John's Medical College Hospital

---

## Research Article

**Keywords:** Head and neck squamous cell carcinoma, cervical lymph node metastasis, gene score, IL15, BIRC5, HNRNPA2B1, molecular biomarkers, risk stratification

**Posted Date:** October 11th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-11147646/v1>

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

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

# Abstract

## Purpose

Head and neck squamous cell carcinoma (HNSCC) remains a major global health burden, with cervical lymph node metastasis (CLNM) being a key determinant of recurrence and survival. Accurate identification of occult nodal disease remains challenging, highlighting the need for reliable biomarkers.

## Methods

Candidate genes associated with CLNM and epithelial–mesenchymal transition in HNSCC were identified through literature mining and evaluated in GSE65858 and TCGA datasets. Genes consistently associated with CLNM were used to derive a three-gene score, which was subsequently validated by quantitative RT-PCR in 63 treatment-naïve FFPE HNSCC samples. Associations with clinicopathological parameters were assessed using appropriate statistical tests.

## Results

Literature mining identified 135 genes associated with lymph node metastasis. In GSE65858 (n = 240), 22 genes were significantly associated with nodal status, including 6 upregulated and 16 downregulated genes. TCGA validation (n = 515) identified 11 significantly dysregulated genes. IL15, BIRC5, and HNRNPA2B1 were consistently upregulated and associated with CLNM in both datasets. The three-gene score derived by computing the average gene expression of the 3 genes was significantly associated with CLNM and stage in GSE65858 (p = 0.002) and CLNM and tumor grade in TCGA (p = 0.001). In the institutional cohort, the score was associated with CLNM (p = 0.006), tumor size (p = 0.016), and pathological stage (p < 0.0001). ROC analysis demonstrated an AUC of 0.703 (95% CI, 0.573–0.832; p = 0.002) for discriminating node-positive from node-negative cases.

## Conclusion

A three-gene score comprising IL15, BIRC5, and HNRNPA2B1 showed reproducible associations with CLNM across independent datasets and an institutional cohort. Larger studies are warranted to validate its utility for predicting nodal metastasis and improving risk stratification in HNSCC.

## INTRODUCTION

Head and Neck Squamous Cell Carcinoma(HNSCC) is the 7th most commonly occurring cancer worldwide and ranks second in India. It represents a significant global health burden with approximately 940,000 new cases, and 480,108 deaths reported worldwide ,underscoring its substantial contribution to cancer-related morbidity and mortality despite advances in clinical diagnosis and treatment [1, 2].

Compounding this burden, HNSCC continues to be challenged by its delayed diagnosis, poor clinical outcomes, limited availability of specific biomarkers and expensive therapeutic modalities [3].

The management of HNSCC primarily involves surgical resection, with chemotherapy and radiotherapy employed as adjunctive modalities, particularly in advanced-stage disease. Despite advances in therapeutic strategies, disease outcome remains strongly influenced by the presence of regional lymph node involvement [4, 5].

The five year survival rate is 83% for localized cancer but declines to 64% when it spreads to nearby lymph nodes. This substantial reduction in survival underscores the prognostic significance of regional lymph node involvement, which is a well-established adverse prognostic factor, associated with an increased risk of recurrence and distant metastasis. Accordingly, several clinical and histopathological prognostic factors have been studied in HNSCC, of which cervical lymph node involvement represents a major determinant of disease outcome.

However, the management of the clinically node-negative HNSCC remains challenging. Elective neck dissection in patients without clinical or radiological evidence of cervical lymph node metastasis may expose patients to unnecessary surgical morbidity without providing a clear therapeutic benefit [6]. Therefore, accurate identification of occult cervical lymph node metastasis is crucial for appropriate treatment planning. Conventional approaches, including clinical examination, radiological imaging and sentinel lymph node biopsy, have inherent limitations and may lack adequate sensitivity for detecting metastatic disease [7, 8].

In view of these diagnostic challenges, molecular profiling approaches, particularly gene -expression based biomarkers, offer a potential means of distinguishing tumors with metastatic potential. Despite extensive research into individual biomarkers, the complex nature of HNSCC suggests that a combination of molecular markers may provide greater prognostic value than any single biomarker. Gene-expression signatures have therefore emerged as a promising approach for capturing the molecular alterations associated with tumour progression, metastasis and patient survival. Such signatures may complement established clinicopathological parameters and facilitate improved identification of patients at high risk of disease progression.

Building on this rationale, the primary objective of this study is to identify gene expression signatures related to cervical lymph node metastasis and recurrence in HNSCC, with the aim of predicting early nodal spread, improving prognostic stratification and potentially aiding in treatment options.

## **Material and Methods**

### **In silico analysis of head and neck squamous cell carcinoma patient data:**

The study encompassed a comprehensive literature-mining approach to identify genes associated with lymph node metastasis in head and neck squamous cell carcinoma (HNSCC). The literature search was conducted using the search criteria “head and neck squamous cell carcinoma” and “lymph node metastasis,” with studies published between 2002 and 2025 focusing on genes and proteins implicated in head and neck squamous cell carcinoma and lymph node metastasis considered for gene identification. Genes that were implicated in epithelial-to-mesenchymal transition and lymph node metastasis in HNSCC, were identified from the published literature. The expression of these genes and their association with lymph node metastasis were subsequently analyzed using publicly available datasets.

To analyze the pattern of gene expression in HNSCC, data was downloaded from the Gene Expression Omnibus data portal (<https://www.ncbi.nlm.nih.gov/geo/>). The GEO accession number of the dataset is GSE65858. The datasets comprised microarray-based expression profiles. Quality control procedures were applied to microarray probe-level intensity files. The primary treatment naive specimens, that remained after removing low-quality arrays, duplicate arrays, and arrays from non-HNSCC samples were used for the analysis.

Expression values were log<sub>2</sub>-transformed and normalized using RSN. The relative normalized log transformed microarray data was used for all the analysis. The TCGA data was accessed from the TCGA Research Network: <https://www.cancer.gov/tcga> (accessed on 15 November 2021). Primary treatment naive specimens were used for analysis from the TCGA dataset. Transcriptomic data from TCGA were obtained using high-throughput RNA sequencing (RNA-Seq) technology.

Gene expression analysis was done on GSE65858 dataset comparing various clinical parameters. The clinical parameters included gender, age ( $\geq 50$  years), smoking, and alcohol intake. The pathological stage was classified as early (Stage I and II tumours) and late (Stage III and IV). T1 and T2 tumors were considered “small tumors”. Nodal status was considered as node negative(N0) and node positive(any N other than N0). The association between cervical nodal metastasis and candidate genes identified from the literature in primary HNSCC samples in the TCGA cohort was analysed and compared with the results from GSE65858 dataset.

## **Computation of 3 gene score:**

From the gene expression analysis performed, consistently upregulated 3 genes that were significantly associated with lymph node metastasis in both the datasets ie GSE65858 and TCGA, were selected. An average gene score was computed by calculating the mean expression values of the genes. This score was subsequently used to assess its association with lymph node metastasis status and survival outcomes.

## **Validation of the gene score in predicting lymph node metastasis status in clinical samples of our dataset:**

For experimental validation of the genes identified through in silico analysis, Formalin-Fixed Paraffin-Embedded (FFPE) samples were retrieved from the Pathology archives of XXXXXXXXXXXX. FFPE samples from primary treatment naïve resection biopsies of HNSCCs were accessed. Sixty-three samples were used for molecular analysis. Approval of the study was obtained from the Institutional Ethics Committee under Reference No.196/2024.

## RNA Extraction and Quantitative RT PCR

Real-time quantitative PCR (qPCR) was used to assess the mRNA expression levels. Two 10µ sections were subjected to xylene deparaffinization followed by digestion at 65°C with Tris-EDTA (TE) buffer containing 1% sodium dodecyl sulphate (SDS) and Proteinase K (Qiagen) for overnight digestion for a period of 16 hours. Total RNA was then extracted using TRI Reagent (Sigma-Aldrich) according to the manufacturer's instructions. Quantitation of RNA was done using Qubit RNA High Sensitivity Assay (Invitrogen™, ThermoFisher Scientific). 500 ng of total RNA was then reverse transcribed into cDNA using the ABI high-capacity cDNA archive kit (Applied biosystems, ThermoFisher Scientific) as per the manufacturer's protocol. Primers were designed for various genes. All primers were designed using Primer3 software and obtained from Eurofins, Bangalore. Reference genes were used to normalize any variations through processing and handling methods. The expression level of housekeeping genes was determined using a panel of 3 reference genes – ALAS1, RPLP0 and RPLP27. A total reaction volume of 10µl containing 1ng of cDNA template per qRT-PCR reaction was done in duplicate using SYBR Green chemistry and run on Applied Biosystems (ABI) Step One Real-Time PCR System. Pre-incubation and initial denaturation of template cDNA was performed at 95°C for 10 min, followed by amplification for 45 cycles at 95°C for 15 sec and 60°C for 1 min. Relative expression of test genes were calculated by the  $\Delta$ CT method and expressed as Relative Normalized Units (RNU). The details of primers used for gene expression analysis are given in Supplementary table 1.

## Statistical Analysis

Descriptive statistics were used for all clinical variables. The difference in gene expression levels was evaluated by the Mann–Whitney U test/Kruskal–Wallis test. Correlations were evaluated by Pearsons's rank test. For all tests, a p-value of < 0.05 was statistically significant. All statistical analysis was carried out using the software XLSTAT -version 2024.4.0

## Results

### ***1. In-silico analysis of HNSCC patient data identified genes that are associated with lymph node metastasis in head and neck squamous cell carcinoma:***

The literature mining identified 135 genes which showed association with cervical nodal metastasis in HNSCC. From the GSE65858 dataset, a total of 300 samples were considered for analysis and quality control procedures were applied to microarray probe-level intensity files. A total of 270 tumor arrays

remained after removing low-quality arrays, duplicate arrays, and arrays from non-HNSCC samples. Out of the 270 samples, 240 primary treatment naive specimens were used for the analysis. Association of these 135 identified genes with various clinico- pathological features including nodal metastasis, amongst the 240 primary HNSCC in the GSE65858 data set were studied. The genes that were significantly associated ( $p < 0.05$ ) with the various clinicopathological features are listed in **Table 1**.

Table 1: Genes that were significantly associated with the various clinicopathological parameters in GSE65858 dataset.

| Variable                     | Upregulated genes                                    | Downregulated genes                                                                                         |
|------------------------------|------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|
| Gender (Female)              | PLRKT, CDKN2A, NBN, MBP                              | SERPINA3                                                                                                    |
| Age ( $\geq 50$ )            | VEGFA, CAD, CCL3, BCL2                               | SPRR2B, SH3GL3, WNT5A, TNNC1, MYOT, MAPK1, CDH1                                                             |
| Smoking (smoking: yes)       | EGFR, SPRR2B, CCND1, COPS6, FOXP3, SOD2, CTTN, CDH17 | PLGRKT, CDKN2A, GSN, CCR4                                                                                   |
| Alcohol (Alcohol: yes)       | MMP1, MMP3, VEGFC, CCND1, TGFB1, HIF1A, MSN          | CDKN2A, CCR5, MBP                                                                                           |
| Lymph node status (positive) | IL15, FANCA, USP7, BIRC5, HNRNPA2B1, BCL2            | ITGA3, TNC, CSNK1A1, SPRR2B, THBD, SH3GL3, PTHLH, CSTA, CDH3, ERV3-1, CD44, HIF1A, MMP10, BMP2, MAPK1, TGFA |
| Tumour stage (late stage)    | GSDMB, USP7, CAD, STAC3, FANCA, KMT2A, BCL2          | MMP3, TNC, SCARB2, SPRR2B, SH3GL3, TGFB1, FOXP3, WNT5A, PTGS2, MMP10                                        |
| Tumour size (large)          | LGALS3BP, VEGFA, ANGPT2, TGFA, ADAM17                | SCARB2, CXCL12                                                                                              |

Out of the 135 genes, 22 genes were differentially expressed and correlated with cervical lymph node metastasis status in the analysis done on GSE65858 data set. Of these, 6 genes were upregulated and 16 genes were downregulated. To validate these results, further analysis was performed on the TCGA dataset with 515 primary tumour specimens.

The association between nodal metastasis and 135 candidate genes identified from the literature using data from 515 primary HNSCC samples in the TCGA cohort was analysed. Out of the 22 genes that emerged significant in the previous dataset, 11 genes were significantly associated with nodal metastasis in TCGA as well. Three genes (IL15, HNRNPA2B1, BIRC5) were significantly upregulated in both the datasets, and 8 genes (CSNK1A1, SPRR2B, PTHLH, CSTA, CDH3, HIF1A, MAPK1 and TGFA) were downregulated in both datasets. Both GSE65858 and TCGA datasets were comparable with respect to the distribution of the various clinico-pathological parameters and is shown in **Table 2**.

Table 2: Distribution of clinicopathological parameters in GSE65858 and TCGA datasets.

| Clinico-pathological parameters | GSE65858 dataset<br>(n=240) | TCGA dataset (n=515) |
|---------------------------------|-----------------------------|----------------------|
| Gender                          |                             |                      |
| Male                            | 82%                         | 74%                  |
| Female                          | 18%                         | 26%                  |
| Age                             |                             |                      |
| <50 years                       | 15%                         | 16%                  |
| >=50 years                      | 85%                         | 84%                  |
| Median age                      | 58 years                    | 61 years             |
| Stage Categories                |                             |                      |
| Early stage(Stage I &Stage II)  | 19%                         | 22%                  |
| Late stage(Stage III &Stage IV) | 81%                         | 78%                  |
| Stage                           |                             |                      |
| Stage I: 7%                     | 7%                          | 6%                   |
| Stage II: 12%                   | 12%                         | 16%                  |
| Stage III: 13%                  | 13%                         | 18%                  |
| Stage IVA: 59%                  | 59%                         | 57%                  |
| Stage IVB: 6%                   | 6%                          | 3%                   |
| Stage IVC: 3%                   | 3%                          | 0%                   |
| T categories                    |                             |                      |
| Small tumours(T1&T2)            | 43%                         | 40%                  |
| Large tumours(T3& T4)           | 57%                         | 60%                  |

Since upregulated gene expressions are ideal for computation of a diagnostic score, an average gene score was computed by calculating the average of the gene expression values of 3 genes (IL15, HNRNPA2B1 and BIRC5) that were significantly upregulated. This score was further used for various associations with lymph node metastasis status and survival data. **Figure 1** is a visual representation depicting the positive association between lymph node status and 3 gene score in GSE65858 and TCGA dataset. The 3-gene score also significantly correlated with stage in GSE65858 ( $p=0.002$ ) and higher grade in TCGA ( $p=0.001$ ) as depicted in **Figure 2**.

Figure 1(**Fig 1**) Box plot representation, depicting the positive association between lymph node status and 3-gene score in GSE65858 (a) and TCGA (b) datasets

Figure 2(**Fig 2**) Box plot illustrating the association of the 3-gene score with stage in GSE65858 dataset(a) and grade in TCGA data set(b)

**II. Three gene score is positively associated with lymph node metastasis in HNSCC samples from our hospital case series:**

For molecular validation of the above findings, 63 samples from patients with HNSCC were analysed. RNA was extracted and q-PCR performed for analyzing the gene expression values of the genes that emerged significantly in the in-silico analysis. The clinicopathological features of these samples (n=63) are shown in **Table 3**.

Table 3: Clinicopathological parameters of HNSCC cases from our hospital database included for molecular validation of in-silico analysis.

| Clinicopathological features |          | Total number of cases<br>(n=63) ( %) |
|------------------------------|----------|--------------------------------------|
| Primary site                 |          | Oral cavity (n=58)<br>Larynx(n=5)    |
| Gender                       | Female   | 25%                                  |
|                              | Male     | 75%                                  |
| Lymph node status            | Negative | 57%                                  |
|                              | Positive | 43%                                  |
| Histological Grade           | 1        | 26%                                  |
|                              | 2        | 72%                                  |
|                              | 3        | 2%                                   |
| Pathological stage           | 1        | 14                                   |
|                              | 2        | 30                                   |
|                              | 3        | 17                                   |
|                              | 4        | 24                                   |
|                              | Nx       | 14                                   |
| Age (Mean)                   |          | 52yrs                                |
| Tumor size (Median)          |          | 2.5 cm                               |

As elaborated in the methods, q-PCR was performed for IL15, HNRNPA2B1 and BIRC5. The average gene score was computed by calculating the average of the relative normalized units of the 3 genes.

High gene score demonstrated a positive association with positive lymph node involvement in the HNSCC samples from our hospital case series ( $p = 0.006$ ). Correlation analysis revealed that the 3-gene score was significantly associated with the number of lymph nodes involved ( $\rho = 0.335$ ,  $p = 0.010$ ). Furthermore, the 3-gene score demonstrated significant positive correlations with tumor size ( $\rho = 0.323$ ,  $p = 0.016$ ) and pathological stage ( $\rho = 0.630$ ,  $p < 0.0001$ ), indicating that the gene score increased with a larger tumor size and a higher stage. Histological grade did not show a significant association with the gene score ( $\rho = 0.098$ ,  $p = 0.473$ ). Receiver operating characteristic (ROC) curve analysis was performed to assess the ability of gene score to discriminate between lymph node-positive and lymph node-negative HNSCC cases. The 3-gene score demonstrated a significant discriminatory performance with an AUC of 0.703 (95% CI: 0.573-0.832,  $p = 0.002$ ), indicating a fair ability to predict lymph node metastasis. **Figure 3** is a visual representation of the associations of 3 gene scores with lymph node status validated on our samples.

Binary logistic regression demonstrated that the 3-gene score was significantly associated with lymph node metastasis. The model was statistically significant (likelihood ratio  $\chi^2 = 7.954$ ,  $p = 0.005$ ) and showed good calibration (Hosmer-Lemeshow  $p = 0.814$ ). The 3-gene score independently predicted lymph node positivity ( $\beta = 0.363$ , OR = 1.438, 95% CI: 1.086-1.904,  $p = 0.011$ ).

## Discussion

In the present study, a three-gene signature comprising IL15, HNRNPA2B1 and BIRC5 was identified through analysis of two independent HNSCC datasets with over 700 samples and was significantly associated with lymph node metastasis. Validation in our institutional FFPE cohort demonstrated a significant association of positive lymph node status with the 3-gene score.

Although largely undetectable in normal adult tissues, BIRC5 is over expressed in the majority of malignancies, making it an important indicator of poor clinical prognosis and a promising target for anticancer therapies. It is also known to play a role in migration and invasion. The gene BIRC5 (Baculoviral IAP Repeat Containing 5), also referred to as survivin, belongs to the inhibitor of apoptosis (IAP) protein family and functions as a crucial regulator of cell division and survival. It supports cellular proliferation while inhibiting apoptotic pathways[9].

HNRNPA2B1 (Heterogeneous Nuclear Ribonucleoprotein A2/B1) is a versatile RNA-binding protein involved in multiple aspects of RNA metabolism, including RNA processing, mRNA stability, and nucleocytoplasmic transport. It has an important role in cancer development, progression and epithelial-mesenchymal transition (EMT) in HNSCC. In addition, HNRNPA2B1 contributes significantly to immune regulation, particularly in antiviral defence mechanisms and inflammatory responses [10].

Interleukin-15 (IL-15) is a key immunostimulatory cytokine that promotes the development, survival, and activation of natural killer (NK) cells and CD8 + T cells, thereby enhancing anti-tumor immune responses. However, in some head and neck cancers, the presence of soluble IL-15 receptor alpha (sIL-15R $\alpha$ ) may alter IL-15 signaling and has been associated with increased production of pro-inflammatory cytokines

such as IL-6 and TNF- $\alpha$ , which can contribute to a tumor-promoting inflammatory microenvironment. IL-15 also has a role in immunotherapy, where integrating with CAR T cells leads to complete tumour response and long survival of CAR-T cells[11, 12].

An important finding of the present study is that the three-gene signature demonstrated consistent association with lymph node metastasis across two independent public datasets and an institutional validation cohort, suggesting that the observed relationship is reproducible across different patient populations and experimental platforms. Notably, the gene score also correlated positively with tumor size, pathological stage, and the number of metastatic lymph nodes, indicating that it may reflect a broader biological program associated with tumor progression and metastatic dissemination. The combination of IL15, HNRNPA2B1, and BIRC5 integrates components of immune regulation, RNA processing, and apoptosis resistance, respectively, highlighting the multifactorial nature of metastatic progression in HNSCC. These findings support the concept that lymph node metastasis is driven not by a single molecular pathway but by coordinated alterations involving tumor cell survival, immune interactions, and post-transcriptional gene regulation.

Several limitations should be acknowledged. ROC analysis revealed that the gene score could moderately discriminate between node-positive and node-negative tumors (AUC = 0.703). Although the predictive performance was not sufficiently high to support its use as a standalone diagnostic marker, the findings suggest that the 3-gene score may contribute to a biomarker panel for identifying patients at increased risk of cervical lymph node metastasis. The retrospective nature of both the public datasets and institutional cohort may introduce selection bias. Furthermore, the study did not evaluate functional studies performed to establish direct mechanistic roles of these genes in metastatic progression. Therefore, the current findings require validation in larger prospective multicentric cohorts before clinical implementation.

Interestingly, despite its significant association with lymph node metastasis and tumor progression parameters, the 3-gene score did not show a significant association with overall survival or disease-free survival in our analysis. This finding suggests that while the signature may reflect biological processes related to regional metastatic spread, it may not be sufficient to capture the complex factors that ultimately determine patient outcomes. Survival in HNSCC is influenced by multiple variables, including treatment modality, tumor subsite, HPV status, margin status, comorbidities, and response to therapy, which may diminish the impact of any single molecular signature. Furthermore, lymph node metastasis and survival, although related, represent distinct clinical endpoints, and biomarkers associated with metastatic potential do not necessarily translate into independent prognostic indicators. Therefore, the primary utility of the present 3-gene signature may lie in risk stratification for nodal involvement rather than prediction of long-term survival outcomes. Identifying such specific gene signatures and converting them into IHC panels can improve the diagnosis of cervical lymph node metastasis. Once validated in a larger series of samples, these genes can aid in early prediction of lymph node metastasis and probably recurrence & metastasis, guiding clinical decisions.

# Declarations

1.Disclosures: Nil

2.Financial support(Funding): This study was supported by funds received by Author 2 and 5(Angel Jaimi and guide Dr. Gayatri Ravikumar) from Rajiv Gandhi University of health sciences(RGUHS) as a part of Short-Term UG Research Grant 2025-26 ref: RGHHS/ADV-RES/UG-GRANTS RESEARCH/234/2025-26, dated 22-01-2025 & 06.03.2025, Project code: UG25MED0459

3.Competing interests: The authors have no competing interests to declare that are relevant to the content of this article.

4.Ethical Approval : Ethical approval was obtained .from the Institutional Ethics Committee of St.John's Medical College, under Reference No.196/2024.

5.Consent: Not applicable.

6.Data, Materials and/or Code availability: Not applicable

7.Authors' contribution statements:

1. Nimita Saji and 2.Angel Jaimi(undergraduate students) have done literature review,public data mining, associations of different dysregulated genes with clinicopathological parameters in two independent publicly available dataset.

3.Lisa Anna Jiju and 4.Dr. Madhumathy G Nair have done the molecular experiments in the study.Madhumathy G Nair also assisted the undergraduate students with statistical analysis.

5.Dr. Gayatri Ravikumar was responsible for conceptualization of the study, writing the grant proposal, supervising the project, case identification from pathology database, manuscript preparation and also served as corresponding author.

8. Data availability statement: The study uses public databases and the molecular testing done our institutional cohort of HNSCC cases, the anonymized data for the same may be available as spreadsheet if required.

**Acknowledgments:** We thank the patients whose samples helped us to conduct this study. We also thank Ms.Tripti, for assisting with the molecular work.

**Funding:** This molecular experiments done for the study was supported by research grant awarded to Angel Jaimi, by RGHHS as a part of Short-Term UG Research Grant 2025-26 ref: RGHHS/ADV-RES/UG-GRANTS RESEARCH/234/2025-26, dated 22-01-2025 & 06.03.2025, Project code: UG25MED0459

Conflict of interest : Nil

# References

1. Sun H, Yu M, An Z, Liang F, Sun B, Liu Y, Zhang S. Global burden of head and neck cancer: epidemiological transitions, inequities, and projections to 2050. *Front Oncol.* 2025;15:1665019. doi:10.3389/fonc.2025.1665019.
2. Singh S, Eisenberg R, Hoffman G. An overview and comparative evaluation of head and neck cancer risk factors in India and Australia. *Int J Otolaryngol Head Neck Surg.* 2018;7(5):456-465. doi:10.4236/ijohns.2018.75027.
3. Rivera C. Essentials of oral cancer. *Int J Clin Exp Pathol.* 2015 Sep 1;8(9):11884-94. PMID: 26617944; PMCID: PMC4637760.
4. Ding Z, Li Y, Pan X, Xuan M, Xie H, Wang X. Sentinel lymph node biopsy versus elective neck dissection in squamous cell carcinoma of the oral cavity with a clinicallyN0 neck: Systematic review and meta-analysis of prospective studies. *Head Neck.* (2021) 43:3185–98. doi: 10.1002/hed.26803
5. Arora A, Husain N, Bansal A, Neyaz A, Jaiswal R, Jain K, et al. Development of a new outcome prediction model in early-stage squamous cell carcinoma of the oral cavity based on histopathologic parameters with multivariate analysis: the Aditi-Nuzhat lymph-node prediction score (ANLPS) system. *Am J Surg Pathol.* (2017) 41:950–60. doi: 10.1097/PAS.0000000000000843
6. Afzali, P. & Ward, B. B. Management of the neck in oral squamous cell carcinoma: background, classification, and current philosophy. *Oral Maxillofac. Surg. Clin. North Am.* 31, 69–84 (2019).
7. Lim S-C, Zhang S, Ishii G, Endoh Y, Kodama K, Miyamoto S, et al. Predictive markers for late cervical metastasis in stage I and II invasive squamous cell carcinoma of the oral tongue. *Clin Cancer Res.* (2004) 10:166–72. doi: 10.1158/1078-0432.CCR-0533-3.
8. Liu M, Wang SJ, Yang X, Peng H. Diagnostic efficacy of sentinel lymph node biopsy in early oral squamous cell carcinoma: A meta-analysis of 66 studies. *PLoS One*(2017) 12:e0170322. doi: 10.1371/journal.pone.0170322
9. Wang N, Huang X, Cheng J. BIRC5 promotes cancer progression and predicts prognosis in laryngeal squamous cell carcinoma. *PeerJ.* 2022 Feb 1;10:e12871. doi: 10.7717/peerj.12871. PMID: 35178302; PMCID: PMC8815368.
10. Gupta A, Yadav S, Pt A, Mishra J, Samaiya A, Panday RK, Shukla S. The HNRNPA2B1-MST1R-Akt axis contributes to epithelial-to-mesenchymal transition in head and neck cancer. *Lab Invest.* 2020 Dec;100(12):1589-1601. doi: 10.1038/s41374-020-0466-8. Epub 2020 Jul 15. PMID: 32669614.
11. Patidar M, Yadav N, Dalai SK. Interleukin 15: A key cytokine for immunotherapy. *Cytokine Growth Factor Rev.* 2016 Oct;31:49-59. doi: 10.1016/j.cytogfr.2016.06.001. Epub 2016 Jun 7. PMID: 27325459.
12. Zhou Y, Husman T, Cen X, Tsao T, Brown J, Bajpai A, Li M, Zhou K, Yang L. Interleukin 15 in Cell-Based Cancer Immunotherapy. *Int J Mol Sci.* 2022 Jun 30;23(13):7311. doi: 10.3390/ijms23137311. PMID: 35806311; PMCID: PMC9266896

Figures

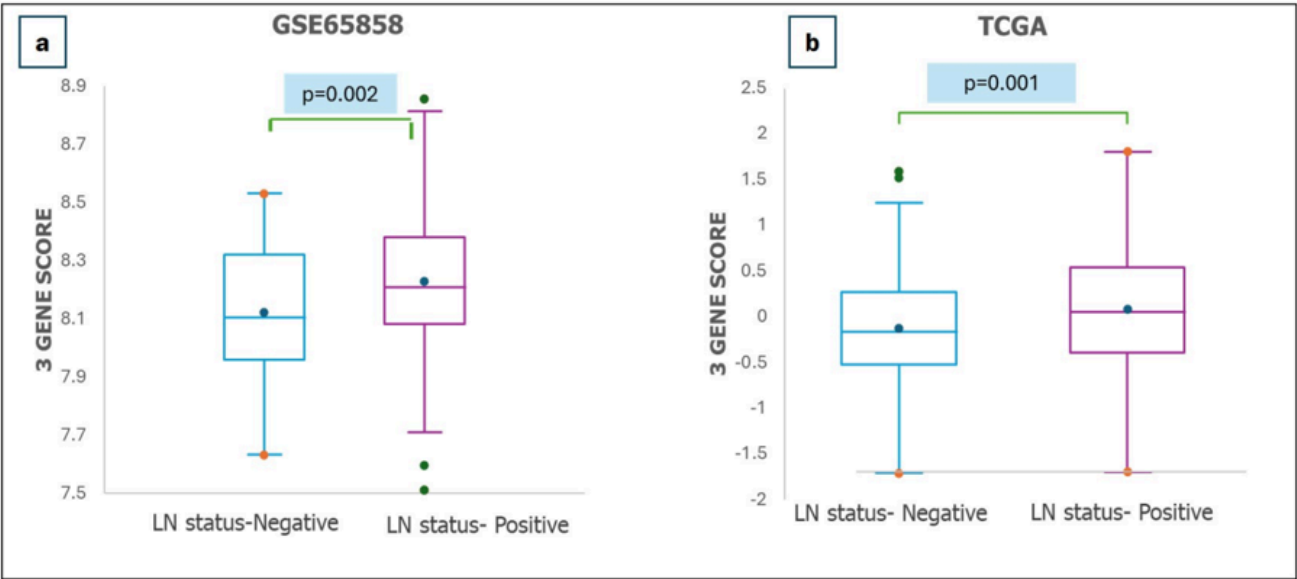


Figure 1

Box plot representation, depicting the positive association between lymph node status and 3-gene score in GSE65858 (a) and TCGA (b) datasets

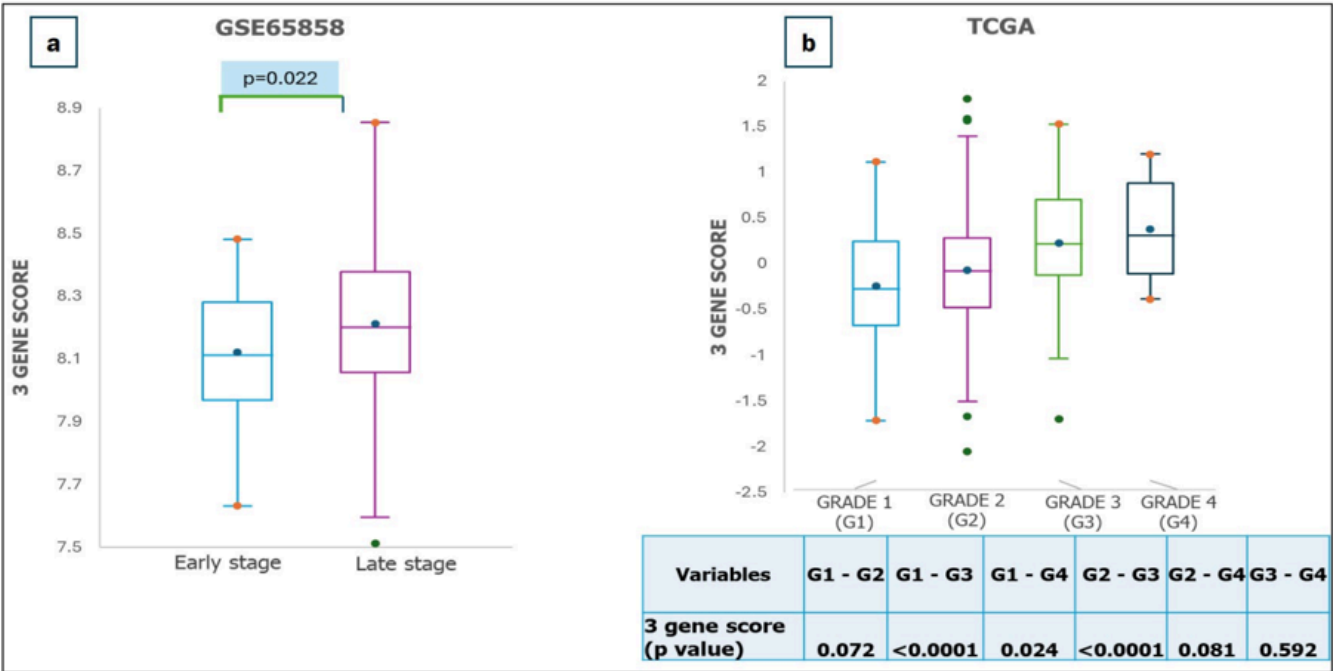


Figure 2

Box plot illustrating the association of the 3-gene score with stage in GSE65858 dataset(a) and grade in TCGA data set(b)

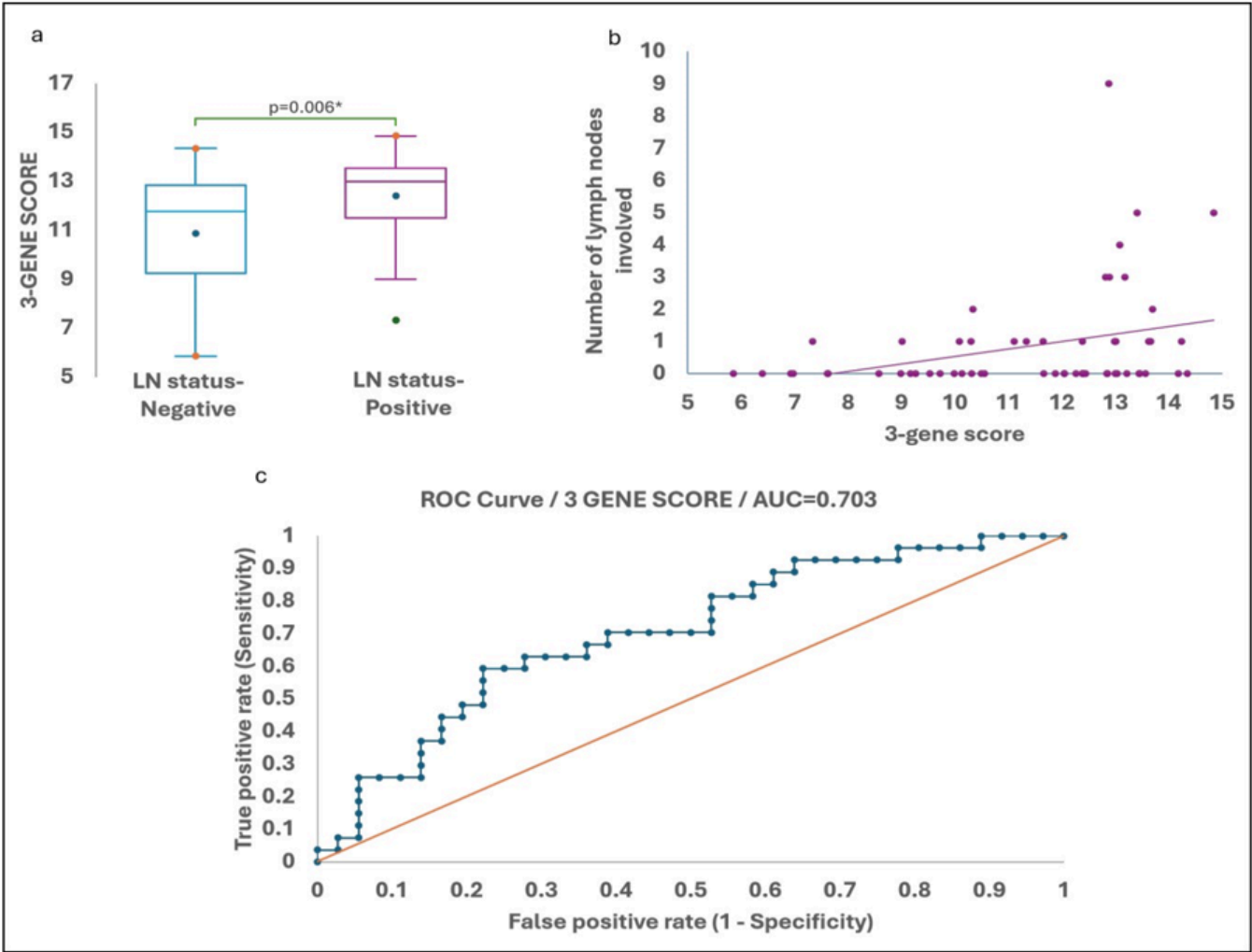


**Figure 3**  
a visual representation of the associations of 3 gene scores with lymph node status validated on our samples.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1Primersequences.docx](#)