Sriwijaya Journal of Otorhinolaryngology e-ISSN 2987-131X
https://phlox.or.id/index.php/sjorl Vol. 4 No. 1 (2026)

Original Research Article

Secondary Cross-Platform Transcriptomic Reanalysis of Laryngeal Squamous Cell Carcinoma Identifies Reproducible Tumor-Associated Programs but No Independently Validated Prognostic Biomarker

Rachmat Hidayat[1,*], Mustafa Mahmud[2]

[1]*Department of Medical Biology, Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia*
[2]*Department of Thoracic Surgery, CMHC Research Center, Palembang, Indonesia*

ARTICLE INFO

Article history:
Received: *February 25, 2026*
Accepted: *March 29, 2026*
Published: *April 30, 2026*

Keywords:
Bioinformatics
Differential gene expression
Laryngeal squamous cell carcinoma
Prognosis
Transcriptomics

***Corresponding author:**
Rachmat Hidayat
E-mail: Dr.rachmat.hidayat@gmail.com
ORCID: 0000-0001-5770-1201

DOI:
https://doi.org/10.59345/sjorl.v4i1.314

ABSTRACT

Background: Public transcriptomes can reveal reproducible laryngeal squamous cell carcinoma (LSCC) programs, but differential expression alone does not establish prognosis.

Objective: To evaluate paired tumor-associated expression changes and test whether any gene met independent prognostic criteria.

Methods: Depositor-processed GSE127165 RNA-sequencing data from 57 paired LSCC and adjacent-mucosa specimens were analyzed with patient-blocked limma-trend after predefined filtering of $\log_2$(FPKM + 1) values. Differentially expressed features required Benjamini–Hochberg false-discovery rate (FDR) <0.05 and absolute paired mean difference ≥1. Functional over-representation used g:Profiler. Expression replication used 23 paired TCGA larynx-labeled cases. Prognostic associations were assessed continuously in GSE27020 training (n=59) and validation (n=50) cohorts and 116 TCGA larynx-labeled tumors, with progression-free interval primary and overall survival secondary.

Results: Of 11,908 tested features, 230 were upregulated and 205 downregulated. Excluding three quality-control-associated pairs retained 418 primary calls; quantile normalization retained 396. Among 327 measurable discovery features in TCGA, 184 replicated at FDR <0.05 with the same direction; directional concordance was 89.0% and effect correlation was ρ=0.774. Keratinization, extracellular-matrix, adhesion, and inflammatory programs were over-represented. No gene satisfied the operational prognostic criterion. *SERPINE1* was exploratory: training HR=2.09, validation HR=1.53, and TCGA progression-free interval HR=1.28; the modified Knapp–Hartung pooled inference was non-significant.

Conclusion: LSCC showed reproducible bulk-tissue expression programs, but no single differentially expressed gene qualified as an independently validated prognostic biomarker. These findings support biological prioritization, not clinical decision-making.

All authors have reviewed and approved the final version of the manuscript.

1. Introduction

Laryngeal cancer imposes a substantial global burden and remains a major cause of cancer-related morbidity because both the disease and its treatment can compromise voice, swallowing, airway protection, and social functioning. The Global Cancer Observatory estimated 193,937 new cases and 96,619 deaths worldwide in 2024, with marked geographical and sex-related variation.[1] Most malignant laryngeal tumors are squamous cell carcinomas arising within a mucosal field exposed to tobacco smoke, alcohol, occupational agents, and other interacting carcinogenic influences. Disease behavior is influenced by anatomical subsite, tumor and nodal stage, margin status, extranodal extension, comorbidity, treatment modality, and competing mortality. Nevertheless, patients with apparently similar clinicopathological features can experience very different patterns of local control, recurrence, organ preservation, and survival.[2]

The biological heterogeneity of LSCC is partly obscured when it is grouped with squamous carcinomas from the oral cavity, oropharynx, and hypopharynx. Head-and-neck squamous cell carcinomas share recurrent disruption of tumor-suppressor, cell-cycle, oxidative-stress, and squamous-differentiation programs, but anatomical sites differ in etiological exposures, human papillomavirus prevalence, lymphatic behavior, treatment, and prognosis. A recent LSCC-specific genomic study of 135 AACR Project GENIE samples

further demonstrated heterogeneous somatic alterations across primary and metastatic disease.[3] A larynx-restricted approach is therefore preferable when the goal is to infer LSCC-specific diagnostic or prognostic biology.

Transcriptome profiling provides a composite readout of malignant cells, stromal components, immune infiltrates, extracellular matrix, and residual epithelium. Recent single-cell and functional immune-microenvironment studies in LSCC have shown that epithelial, fibroblast, stem-like, and myeloid populations can contribute distinct expression programs.[4,5] Differential-expression analysis can therefore identify reproducible tumor-associated states, prioritize candidate mechanisms, and generate markers for subsequent assay development, but bulk-tissue signals cannot be assigned to one cellular compartment without additional localization. Moreover, a gene that differs between tumor and non-tumor tissue is not automatically prognostic. Prognostic utility additionally requires a time-to-event endpoint, an adequate number of events, evaluation of confounding, correction for multiple comparisons, and independent validation in patients who were not used for discovery. The distinction is particularly important in high-dimensional studies, where hundreds or thousands of gene–outcome associations are examined with comparatively few events.

Several recent bioinformatics studies have proposed LSCC prognostic signatures based on four messenger RNAs,[6] a 10-gene recurrence score,[7] RNA-binding proteins,[8] seven long non-coding RNAs,[9] immune-related genes,[10] ferroptosis-associated long non-coding RNAs,[11] co-expression hubs,[12] neutrophil-extracellular-trap genes,[13] and lactate-related genes.[14] Hong et al. reused GSE127165, TCGA-LSCC, and GSE27020 to develop an immune-gene recurrence signature,[10] while Li et al. combined GSE127165 and TCGA in differential-expression, network, and hub-gene analyses.[12] These reports establish biological potential but also show how heavily a small number of public cohorts have been reused. Furthermore, both a 2025 HNSCC analysis[15] and a 2026 systematic review and meta-analysis[16] associated elevated *SERPINE1* with adverse outcomes, making a de novo novelty claim untenable. The 2026 review reported associations with OS and DFS but not consistently with progression-related or disease-specific endpoints, making endpoint-specific confirmation especially important.[16] Our targeted literature update through 22 July 2026 was intended to position this reanalysis and was not a formal systematic review. Closely related studies and the incremental contribution of the present analysis are detailed in Table 1.

Table 1. Positioning against closely related public-dataset studies.

Study	Datasets / scope	Feature strategy	Validation and result	Distinction from present reanalysis
Zhang et al. 2021 [6]	Public LSCC GEO and TCGA cohorts	Four-mRNA prognostic signature	Internal and public-cohort evaluation	Composite model; not paired DEG-to-single-gene replication
Liu et al. 2022 [7]	Multiple public laryngeal-cancer sources	10-gene recurrence score	Risk-score evaluation across three datasets	Signature development rather than paired cross-platform DEG replication
Hong et al. 2023 [10]	GSE127165, TCGA-LSCC, GSE27020	Immune-related genes; Cox/LASSO	TCGA training; GSE27020 risk-score validation	Reuses the same cohorts; present work tests locked single-gene evidence and retains non-validation
Li et al. 2024 [12]	GSE127165 and TCGA-LSCC	DEG, WGCNA, PPI hub genes	Overlapping ECM/focal-adhesion biology	Present contribution is paired replication, attrition reporting, and prognostic evidence grading
Yan et al. 2026 [16]	HNSCC studies; systematic review/meta-analysis	*SERPINE1* evidence synthesis	Adverse OS/DFS; less consistent progression-related evidence	Broader anatomical scope; supports endpoint-qualified, non-novel interpretation
Present secondary reanalysis	GSE127165, GSE27020, larynx-labeled TCGA	Paired DEG; locked training probes; continuous Cox	184/327 expression replications; zero independently validated prognostic genes	Reproducibility-focused negative validation with complete-case small-*k* uncertainty

Adjacent mucosa introduces an additional interpretive issue. It is clinically relevant and controls for host, exposure, and tissue context, but it is not equivalent to healthy laryngeal mucosa from cancer-free individuals. Molecular, epigenetic, microbial, inflammatory, and cell-composition changes may extend beyond the visible tumor boundary; recent integrated methylome–transcriptome work in laryngeal cancer illustrates how these biological layers can be coupled.[17] Consequently, tumor-versus-adjacent contrasts identify tumor-associated expression changes within a cancer-bearing field, not necessarily cancer-specific alterations in the strict diagnostic sense.

We designed this secondary reanalysis to answer two linked but distinct questions. First, which expression changes and bulk-tissue biological programs are reproducibly associated with LSCC tissue? Second, do any resulting differentially expressed genes show sufficiently consistent time-to-event associations to qualify as prognostic biomarkers? The incremental contribution is the combination of patient-paired discovery, independent paired cross-platform replication, locked-probe survival validation, explicit identifier attrition, small-*k* uncertainty analysis, and

retention of a negative prognostic conclusion. It is not discovery of extracellular-matrix biology, inflammatory involvement, or *SERPINE1*. The study was reported with reference to REMARK principles.

2. Methods

2.1 Study design and analysis hierarchy

This was a retrospective secondary analysis of de-identified public datasets. Four distinct components were conducted: paired differential-expression discovery in GSE127165; directional functional over-representation; cross-platform expression replication in paired TCGA-HNSC larynx-labeled samples; and prognostic evaluation in GSE27020 training, GSE27020 validation, and TCGA. The hierarchy prevented tumor–mucosa evidence from being misrepresented as prognostic validation.

The primary differential-expression threshold was BH FDR < 0.05 with an absolute paired mean $\log_2$(FPKM + 1) difference of at least 1. GSE27020 DFS was defined by the source GEO record. TCGA PFI was primary because it includes progression or recurrence events; it may also include new tumor events or cancer-related death and is not interchangeable with DFS. OS was secondary. Expression was modeled continuously per one standard deviation; median-split curves were descriptive only.

The predefined operational biomarker-grading rule required discovery significance, same-direction TCGA expression replication at FDR < 0.05, training prognostic FDR < 0.05, validation $p < 0.05$, TCGA PFI $p < 0.05$, and a consistent HR direction. The conjunction was not prospectively registered and is not formal global family-wise type-I-error control. BH families were separately defined for 11,908 discovery features, measurable discovery genes within each TCGA endpoint, 262 GSE27020 candidates within each cohort, and 251 complete three-cohort meta-analytic candidates. Validation estimates were generated for the full candidate set rather than only training-selected genes.

2.2 Data sources and cohort selection

GSE127165 contained Illumina HiSeq 4000 RNA sequencing of 57 LSCC tissues and 57 paired tumor-adjacent mucosal tissues. The GEO record described hg19-based processing and Cuffdiff v2.1.1 gene-level FPKM quantification; no sequencing-batch variable was deposited. Identifiers beginning with "C" and "P" denoted carcinoma and paired mucosa, and pairing was verified for all 57 patients. Raw reads are available through SRA SRP186837, but the primary reanalysis used the exact deposited FPKM matrix. GEO supplied stage, tobacco, and alcohol fields, but not age, sex, laryngeal subsite, grade, treatment, margin, HPV/p16, or follow-up.

GSE27020 contained Affymetrix Human Genome U133A profiles from primary operable laryngeal cancers. The deposited series annotations included 59 training and 50 separate validation patients treated surgically with or without radiotherapy; DFS/recurrence was the endpoint. Age, grade, DFS, and cohort membership were available in the series matrix. Sex, subsite, TNM/stage, HPV/p16, margin, and detailed treatment were unavailable in the deposited series annotations.

TCGA-HNSC phenotype, curated survival, and TCGA.HNSC.sampleMap/HiSeqV2 expression matrices were retrieved from UCSC Xena. HiSeqV2 contains gene-level RSEM-normalized expression transformed as $\log_2(x + 1)$ by Xena. Cases were restricted to "Larynx" in anatomic_neoplasm_subdivision and to primary tumors or solid-tissue normals; ICD-O site and histology were audited after construction. One sample per patient and tissue was retained. There were 117 clinical tumors and 116 with expression; 115 were C32.9 and one C32.1, while 114 were conventional and two basaloid-type squamous carcinomas. This cohort is larynx-labeled but predominantly subsite-unspecified. Twenty-three solid-tissue-normal specimens had matched tumors. PFI and OS came from the curated resource. TCGA was not used for GSE127165 candidate selection. Dataset roles are detailed in Table 2, clinical reporting boundaries are detailed in Table 3, and cohort assembly is shown in Figure 1.

Table 2. Public cohorts and their analytic roles.

Role	Dataset	Design / platform	Use in this study
DEG discovery	GSE127165	57 paired LSCC–adjacent mucosa; Illumina HiSeq 4000; FPKM	Paired DEG, PCA, QC sensitivity, GO/pathway analysis
DFS prognosis	GSE27020 training	59 primary laryngeal tumors; Affymetrix HG-U133A; 13 events	Candidate screening using continuous Cox models
DFS validation	GSE27020 validation	50 independent primary laryngeal tumors; Affymetrix HG-U133A; 21 events	Independent array-based prognostic validation
Cross-platform replication	TCGA-HNSC larynx pairs	23 paired primary tumors and solid-tissue normals; Xena normalized RNA expression	Replication of discovery DEG direction and FDR
PFI / OS validation	TCGA-HNSC larynx tumors	116 primary tumors with expression; 42 PFI events; 50 deaths	Primary PFI and secondary OS Cox analyses

2.3 RNA-sequencing preprocessing and quality assessment

The GSE127165 matrix contained 32,752 rows: 26,876 ENSG identifiers and 5,876 XLOC features. Version suffixes were removed before matching, and approved symbols were mapped using the HGNC complete-set export retrieved 22 July 2026. Features were retained when mean FPKM exceeded 0.5 and FPKM exceeded 0.5 in at least 20% of samples. The filtered matrix contained 11,908 features—11,167 ENSG and 741 XLOC—not 11,908 unambiguously mapped genes.

Because the public matrix contained FPKM rather than counts, negative-binomial procedures were not applied. Values were transformed as $\log_2(\text{FPKM} + 1)$ and modeled with limma-trend. The tissue coefficient is a paired mean $\log_2(\text{FPKM} + 1)$ difference, not a count-derived fold change. Robust empirical Bayes moderation was used for continuous expression, but FPKM remains composition-sensitive and cannot recover count uncertainty. A distributional sensitivity analysis applied quantile normalization to filtered $\log_2(\text{FPKM} + 1)$ values and repeated the paired model. This cannot replace raw-count reprocessing but tests dependence on an explicit between-sample normalization step.

Sample-level QC included total FPKM, median expression, zero fraction, missingness, and descriptive PCA of the 1,000 most variable transformed features. Robust z scores used the median and median absolute deviation. Three samples—C52, P46, and P56—met predefined extreme-QC rules. The primary model retained all samples. Sensitivity analysis removed both members of each associated pair, leaving 54 pairs.

Table 3. REMARK-oriented clinical characteristics and reporting availability.

Cohort	N	Age / sex	Clinical profile	Outcome / follow-up	Principal missingness or boundary
GSE127165	57 pairs	Not deposited	Stage I/II/III/IV: 20/10/15/12; tobacco yes 54; alcohol yes 32	Tumor–mucosa expression only	Subsite, grade, treatment, margins, HPV/p16, and follow-up unavailable
GSE27020 training	59	62 y (IQR 55–69); sex unavailable	Grade 1/2/3: 19/28/10; 2 missing; surgery ± radiotherapy	DFS: 13 events; reverse-KM follow-up 50 mo	Subsite, TNM/stage, margins, HPV/p16, and detailed treatment unavailable
GSE27020 validation	50	64 y (IQR 59.25–71.5); sex unavailable	Grade 1/2/3: 23/21/6; surgery ± radiotherapy	DFS: 21 events; reverse-KM follow-up 30 mo	Same limitations; separate patients but same source study/platform
TCGA larynx-labeled	116	62 y (IQR 56–68); 96 male/20 female	Stage I–II 13; III–IV 88; 15 missing; C32.9 115/116; conventional 114, basaloid 2	PFI 42 events, follow-up 36.0 mo; OS 50 deaths, follow-up 44.6 mo	Predominantly subsite-unspecified; treatment heterogeneous; adjusted complete cases n=101

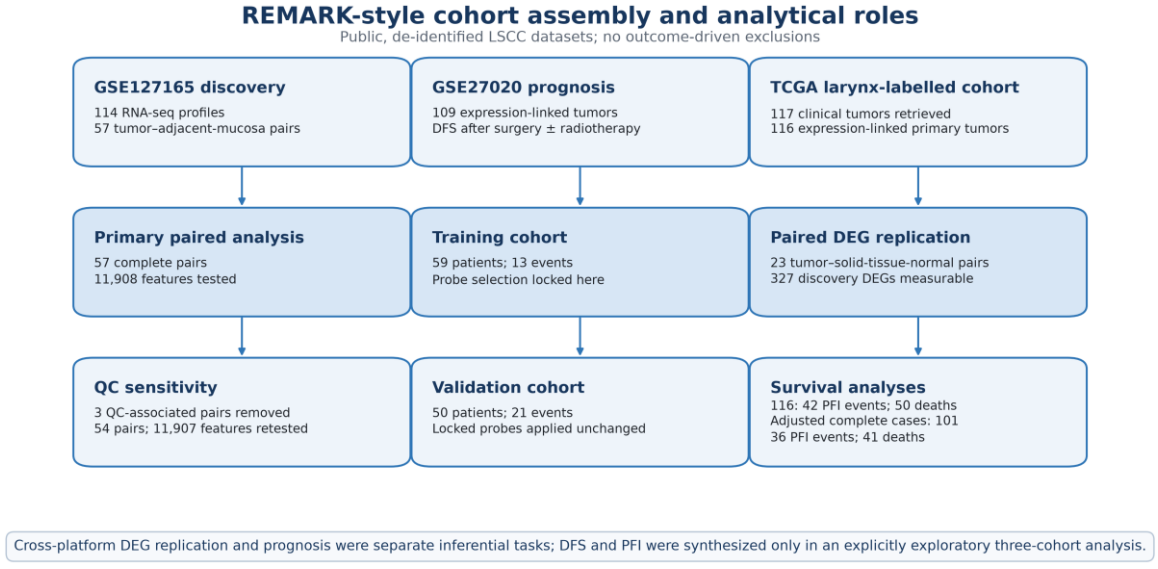

Figure 1. Cohort assembly and analytical roles. Differential-expression replication and prognosis were separate inferential tasks.

2.4 Differential-expression modeling

The primary design included patient and tissue status, preserving within-patient comparison. Limma-trend with robust empirical Bayes moderation was fitted to $\log_2$(FPKM + 1). *P* values were BH-adjusted across 11,908 features. Features with FDR < 0.05 and an absolute paired mean $\log_2$(FPKM + 1) difference of at least 1 were called differentially expressed. Positive values denoted tumor upregulation.

The pair-exclusion sensitivity independently recomputed the expression filter, then repeated modeling and BH correction. Robustness was summarized by retained primary calls and Spearman effect correlation across commonly tested features.

2.5 Functional enrichment analysis

Upregulated and downregulated lists were analyzed separately through g:Profiler. All 11,908 tested identifiers were submitted as custom background. The service recognized 190/230 upregulated and 185/205 downregulated identifiers and used an effective domain of 10,621. Results used g:Profiler e114_eg62_p19_27110d83, Ensembl 114/GRCh38.p14, GO class release 23 January 2026, and KEGG release 15 March 2026. Service-adjusted $p < 0.05$ defined significant over-representation. This was not evidence of causal activation or cell-type-specific signaling.

2.6 TCGA paired replication of differential expression

Discovery features with approved symbols and exact Xena matches were evaluated in 23 paired TCGA larynx-labeled tumor–solid-tissue-normal cases. Xena values were modeled with patient and tissue status using robust limma-trend. BH correction was applied across measurable discovery genes. Replication required FDR < 0.05 and the same effect sign; directional concordance and Spearman effect correlation were also calculated. Because processed scales differ, rank concordance—not slope equivalence—was the target.

2.7 Microarray preprocessing and prognostic analysis in GSE27020

GSE27020 probes were mapped with official GPL96 annotation. When multiple probes mapped to one gene, the representative probe was selected by largest interquartile range in the 59-patient training cohort only and applied unchanged to validation. This locked rule changed 21 of 262 probes compared with all-sample IQR selection and avoided using validation expression to define features. The deposited scale was retained; missing expression was removed before within-cohort standardization.

Within each cohort, Cox models related continuous standardized expression to DFS. Outputs included sample size, events, coefficient, standard error, HR, 95% CI, Wald p, apparent univariable concordance, BH FDR, and scaled-Schoenfeld proportional-hazards p value. No cut point was optimized. No retained GSE27020 survival time was non-positive. Median follow-up was estimated by reverse Kaplan–Meier.

2.8 TCGA prognostic validation and multivariable follow-up

For TCGA, one primary-tumor expression profile per larynx patient was joined to curated PFI and OS records. Candidate genes were standardized within the larynx cohort. Separate univariable Cox models were fitted for PFI and OS, and BH correction was applied within each endpoint. Models required at least 30 evaluable patients, eight events, and nonzero expression variance.

SERPINE1 was selected as an illustrative exploratory candidate before TCGA outcome inspection because of its cross-GSE direction, matrix-remodeling relevance, and prior literature.[15,16] Complete-case models included standardized *SERPINE1*, age per ten years, sex, and stage I–II versus III–IV. Stage-stratified models were accompanied by stage-covariate sensitivity models. Schoenfeld diagnostics were applied to all univariable and adjusted models. Fifteen patients lacked stage; adjusted models retained 101 patients with 36 PFI events and 41 deaths, while the early-stage stratum had four PFI events and seven deaths. Two basaloid-type cases were retained as laryngeal squamous carcinomas but could not support subtype estimation.

2.9 Cross-cohort meta-analysis and uncertainty sensitivity

Standardized-expression log HRs were synthesized only for genes with complete GSE27020 training, GSE27020 validation, and TCGA PFI estimates. Eleven of 262 genes lacked one TCGA estimate and were excluded from the principal 251-gene family. Fixed effects were retained for comparability; REML random effects were the principal heterogeneity-aware sensitivity model. Standardization aligned units but did not remove heterogeneity in platform, case mix, treatment, follow-up, or endpoint. Q, I^2, and REML τ^2 were reported.

Modified Knapp–Hartung inference used $k - 1$ degrees of freedom and variance inflation bounded below by one. Fixed and REML calculations were validated with metafor. BH correction was applied separately across 251 fixed-effect, REML, and modified Knapp–Hartung *p* values. The two GSE27020 subsets share a source study and possible laboratory and eligibility effects; they were treated as separate patient cohorts, not fully independent studies. Pooled significance did not override failed independent validation.

2.10 Statistical software, reproducibility, and ethics

Analyses used R 4.5.3 with data.table, limma, survival, metafor, ggplot2, ggrepel, httr, and jsonlite. A master workflow executes GEO discovery, locked-probe prognosis, TCGA replication, survival analyses, and final meta-analysis. An input manifest records accessions, URLs, retrieval date, byte size, and SHA-256 checksums; session information and complete outputs are retained. Thresholds were predefined for

this analysis and were not relaxed. The reproducibility package is prepared for permanent public deposition; its persistent identifier is reported in the Availability of data and materials statement.

All datasets were public and de-identified; no new participants were recruited, identifiable records accessed, or intervention performed.

3. Results

3.1 Cohort assembly and expression quality

All 57 GSE127165 tumors had matched tumor-adjacent mucosa. Of 32,752 rows, 11,908 were tested after filtering. Stage was I in 20, II in 10, III in 15, and IV in 12; 54 patients were recorded as tobacco users and 32 as alcohol users. Age, sex, subsite, treatment, and follow-up were not deposited. Descriptive PCA showed a tumor–mucosa displacement with overlap; PC1 and PC2 explained 34.9% and 18.7% of variance. PCA did not constitute a technical-quality guarantee. Cohort assembly is shown in Figure 1, and the PCA structure is shown in Figure 2.

QC metrics flagged C52 because of unusually high total FPKM and an elevated zero fraction, P46 because of a low median expression value and elevated zero fraction, and P56 because of high total FPKM. No sample had a missing paired partner. Retaining all samples in the primary model avoided post hoc decisions based on visual appearance; the pair-exclusion analysis provided a direct assessment of their influence.

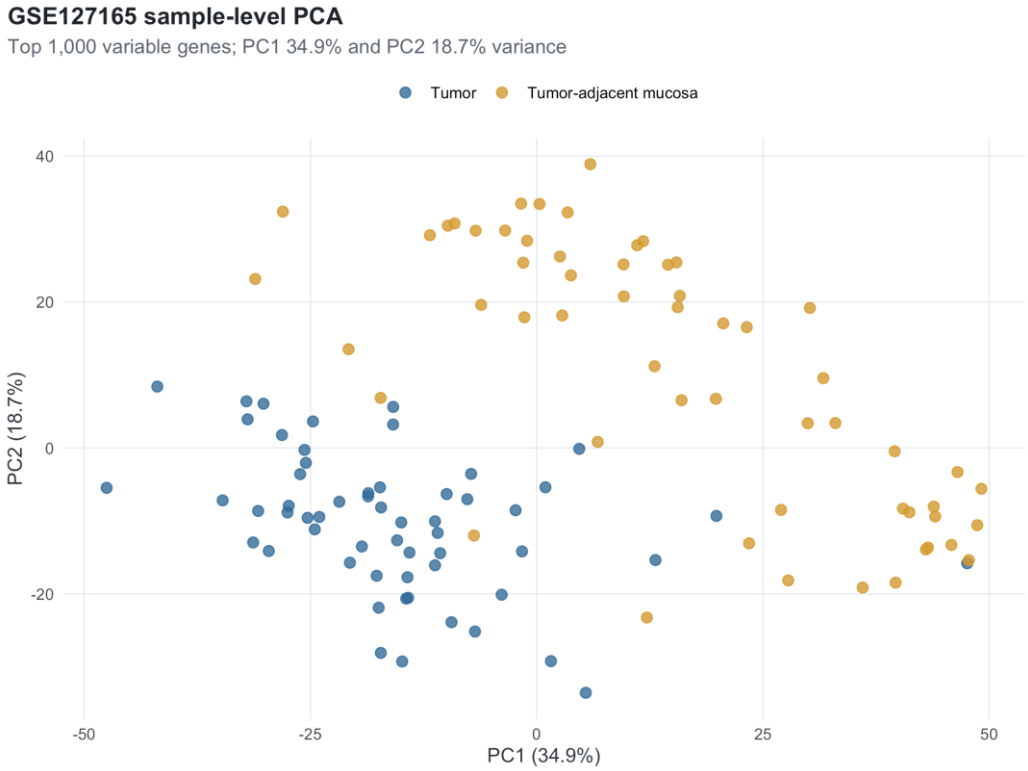

Figure 2. Principal-component analysis of GSE127165 after $\log_2$(FPKM + 1) transformation. The plot is descriptive and not a technical-quality guarantee.

3.2 Differential-expression landscape

The primary analysis identified 435 differentially expressed features: 230 upregulated and 205 downregulated. These comprised 413 ENSG and 22 XLOC features; 368 mapped to approved HGNC symbols. Statistically supported increases included *HOXB7* (paired mean $\log_2$(FPKM + 1) difference 1.12; FDR 8.32×10^{-19}), *MMP1* (3.66; 2.16×10^{-18}), *COL4A1* (1.26; 6.00×10^{-18}), *FEN1* (1.12; 9.24×10^{-18}), *SERPINE1* (2.16; 2.42×10^{-17}), *FSCN1*, *BIRC5*, *FOXM1*, *PSMD2*, and *KPNA2*. Large effects included *MMP1*, *KRT16*, *SPRR2G*, *KRT17*, *S100A7*, *KRT6B*, *LCE3D*, *KRT14*, *COL1A1*, *KRTDAP*, *S100A7A*, and *SPP1*.

Strongly supported decreases included *TJP3* (−1.08; FDR 3.33×10^{-22}), *FAM3D* (−2.04; 3.86×10^{-22}), *VSIG2*, *SH3BGRL2*, *KLHDC7A*, *CYP4B1*, *GABRP*, *AGFG2*, *TSPAN6*, and *CRISP3*. Large negative effects included *PRR4*, *PRH2*, *PRB3*, *PRH1*, *SCGB3A1*, *PIGR*, *LYZ*, *KRT4*, *STATH*, *BPIFB1*, *MUC5B*, and *ZG16B*. These bulk-tissue changes combined loss of differentiated mucosal and secretory features with proliferative, keratinizing, and matrix-rich states. The full effect–significance distribution is shown in Figure 3.

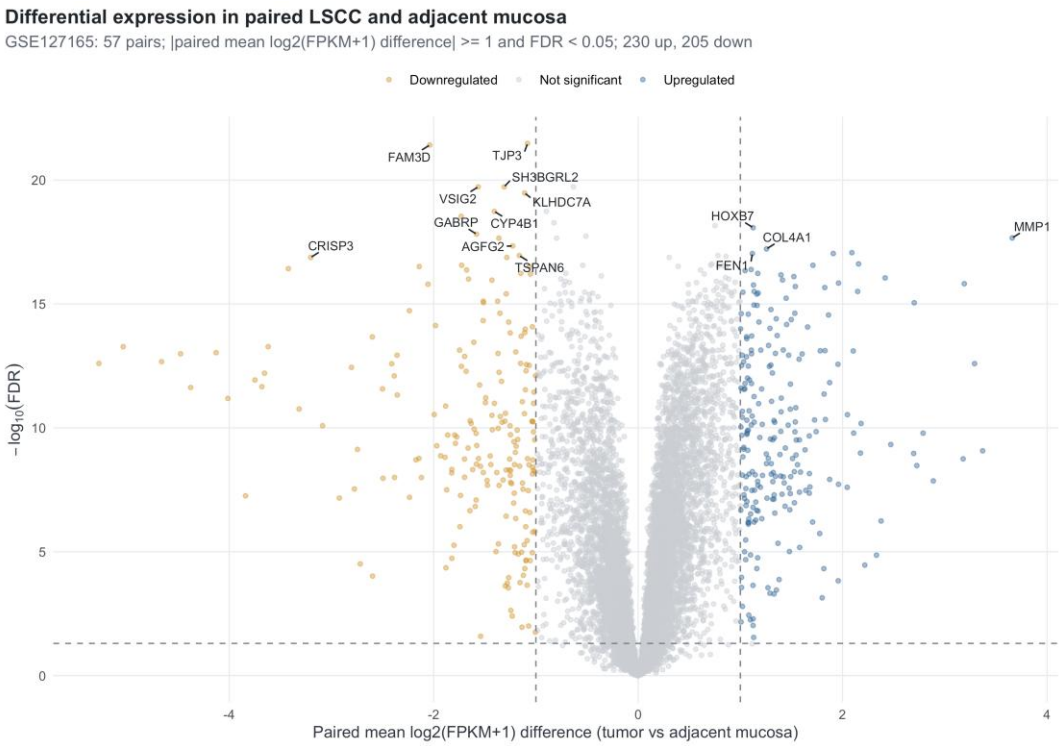

Figure 3. Differential-expression volcano plot. Thresholds were BH FDR <0.05 and absolute paired mean $\log_2$(FPKM + 1) difference ≥1.

3.3 Sensitivity to QC-associated pairs

Removal of pairs 52, 46, and 56 left 54 pairs. Independent refiltering retained 11,907 features and identified 220 upregulated and 206 downregulated features. Of 435 primary calls, 418 remained significant; paired effects correlated at ρ = 0.998. Quantile-normalization sensitivity identified 193 upregulated and 228 downregulated features, retained 396 primary calls, and produced effect correlation ρ = 0.979. Thus, direction and overall biological signal were stable, although exact threshold counts depended partly on normalization.

3.4 Functional programs represented by LSCC DEGs

Among 230 submitted upregulated features, 190 were recognized; among 205 downregulated features, 185 were recognized. The effective background was 10,621 rather than the 11,908 submitted identifiers. Upregulated features were over-represented in keratinocyte differentiation, keratinization, cornified-envelope formation, extracellular-matrix localization, integrin signaling, focal adhesion, and cadherin-related terms. Cornified envelope had adjusted p = 4.50×10^{-34}; ECM–receptor interaction, 1.59×10^{-9}; protein digestion and absorption—mainly matrix/collagen genes—3.15×10^{-8}; IL-17 signaling, 1.38×10^{-6}; and integrin signaling, 1.88×10^{-6}.

Downregulated genes were enriched for epithelial and tissue differentiation, intermediate-filament organization, humoral or antimicrobial defense, extracellular and secretory compartments, salivary secretion (adjusted p = 2.58×10^{-6}), cornified-envelope formation (4.26×10^{-6}), drug metabolism through cytochrome P450 (0.00165), and arachidonic-acid metabolism (0.0299). The presence of cornified-envelope-related signals in both directions reflected different subsets of structural genes rather than a mathematical contradiction. Tumor tissue upregulated a hyperproliferative squamous keratin program while losing elements of normal differentiated mucosa. Direction-specific enrichment results are detailed in Figure 4.

3.5 Independent replication of differential expression in TCGA larynx pairs

Of 435 discovery features, 368 mapped to approved symbols, 327 matched Xena, and 108 were lost through absent/ambiguous symbol mapping or matrix coverage. In 23 paired larynx-labeled cases, 184 genes replicated at TCGA FDR < 0.05 with the same direction. Directional concordance was 89.0% and Spearman effect correlation was ρ = 0.774.

Replicated genes included strong reductions in *HPGD*, *KRT4*, *MAL*, *TMPRSS11B*, *CRNN*, *KRT78*, *CLCA4*, *MUC21*, *CRISP3*, *CAPN14*, *PADI1*, and *AQP3*, and increases in *COL4A1*, *PTHLH*, *CHST2*, *SPP1*, and *FST*. For example, *SPP1* had a paired mean $\log_2$(FPKM + 1) difference of 2.42 in GSE127165 and a 5.58-unit paired Xena difference, with TCGA FDR 9.32×10^{-6}. Cross-platform effect concordance and replication status are shown in Figure 5.

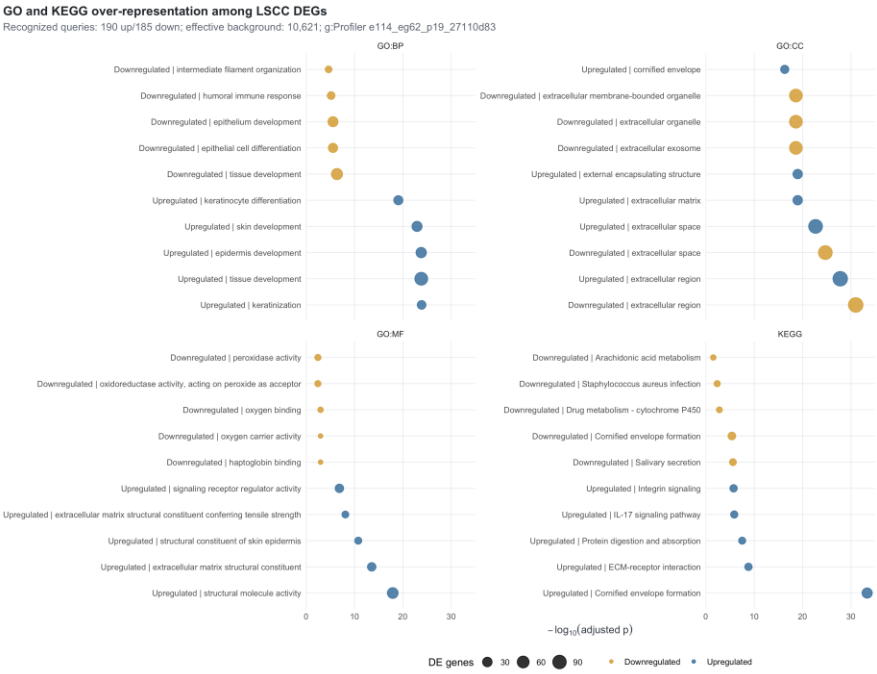

Figure 4. Gene Ontology and pathway over-representation. Effective recognized queries were 190 upregulated and 185 downregulated features against a background of 10,621.

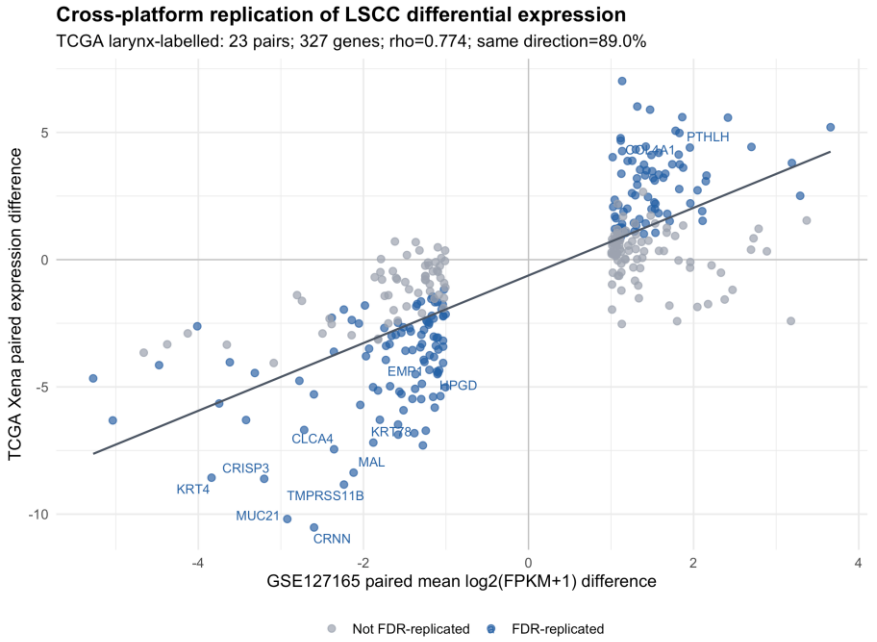

Figure 5. Cross-platform replication for 327 discovery-selected genes. Spearman ρ=0.774 and directional concordance=89.0%; processed scales differ.

3.6 Prognostic screening in GSE27020

GSE27020 included 59 training patients with 13 events and 50 validation patients with 21 events. Median ages were 62 (IQR 55–69) and 64 (59.25–71.5) years; reverse-Kaplan–Meier median follow-up was 50 and 30 months. Training-only probe selection yielded 262 candidates and changed 21 representative probes relative to all-sample IQR selection. No gene passed training FDR plus nominal validation, nominal $p < 0.05$ in both cohorts with consistent direction, or the full operational grading rule.

For illustrative *SERPINE1*, training DFS HR was 2.09 (95% CI 1.303–3.360; p=0.00224; FDR=0.445) and validation HR was 1.53 (0.985–2.382; p=0.0582; FDR=0.797). Schoenfeld p values were 0.238 and 0.440. *LAMB3*, *CST1*, *ADM*, *TGFBI*, *SERPINH1*, *HSPB6*, *INHBA*, *CENPF*, *IGFBP3*, *PLAU*, and *SPP1* were research-prioritization signals only.

3.7 TCGA larynx prognostic evaluation

The TCGA cohort comprised 116 expression-linked primary tumors. Median age was 62 years (IQR 56–68); 96 patients were male and 20 female. Stage was I–II in 13 and III–IV in 88, with 15 missing. Radiotherapy was recorded as yes for 70, no for 23, and missing for 23; margin was negative in 84, positive in nine, close in four, and missing in 19. There were 42 PFI events and 50 deaths. Reverse-Kaplan–Meier median follow-up was 36.0 months for PFI and 44.6 months for OS.

For *SERPINE1*, primary PFI HR was 1.284 (95% CI 0.917–1.797; p=0.145; FDR=0.416; Schoenfeld p=0.266), so TCGA did not independently validate a progression-related association. Secondary OS HR was 1.489 (1.096–2.024; p=0.0110; FDR=0.445; Schoenfeld p=0.767). In 101 complete cases, stage-stratified adjusted HRs were 1.285 for PFI (0.865–1.909; p=0.215) and 1.863 for OS (1.250–2.776; p=0.00223). Stage-covariate sensitivity HRs were 1.217 and 1.726, respectively. Global proportionality tests for stage-covariate models were significant (PFI p=0.008; OS p=0.034), emphasizing instability rather than clinical independence.

3.8 Three-cohort meta-analysis and evidence classification

Eleven of 262 candidates lacked one TCGA estimate and were not pooled. Among 251 complete three-cohort genes, 23 had fixed-effect FDR < 0.05 and 20 had REML normal-theory FDR < 0.05; 21 fixed-effect signals had a consistent direction. None survived modified Knapp–Hartung correction and none satisfied the operational independent-validation rule. Because GSE27020 DFS and TCGA PFI are related but non-identical, all pooling remained exploratory.

SERPINE1 fixed-effect HR was 1.518 (95% CI 1.202–1.916; p = 0.000445; FDR = 0.0186; I^2 = 26.5%). REML HR was 1.542 (1.169–2.034; τ^2 = 0.0165; p = 0.00220; FDR = 0.0427). Modified Knapp–Hartung retained the REML point estimate but widened the CI to 0.839–2.834 (p = 0.0922; FDR = 0.760). The individual and pooled estimates are detailed in Table 4 and visualized in Figure 6. This illustrates why normal-theory pooling cannot substitute for independent validation.

Table 4. *SERPINE1* prognostic evidence across cohorts and sensitivity analyses.

Analysis	Endpoint	HR per 1 SD	95% CI	p / FDR	Interpretation
GSE27020 training	DFS	2.093	1.303–3.360	p=0.00224; FDR=0.445	Nominal only; PH p=0.238
GSE27020 validation	DFS	1.532	0.985–2.382	p=0.0582; FDR=0.797	Not significant; PH p=0.440
TCGA larynx-labeled	PFI	1.284	0.917–1.797	p=0.145; FDR=0.416	Primary TCGA endpoint not validated; PH p=0.266
TCGA larynx-labeled	OS	1.489	1.096–2.024	p=0.0110; FDR=0.445	Nominal secondary association; PH p=0.767
Stage-stratified adjusted TCGA	OS	1.863	1.250–2.776	p=0.00223; n=101/41 deaths	Exploratory complete-case model
Fixed-effect pooled	DFS/PFI	1.518	1.202–1.916	p=0.000445; FDR=0.0186	Exploratory normal-theory summary
REML random effects	DFS/PFI	1.542	1.169–2.034	p=0.00220; FDR=0.0427	τ^2=0.0165; k=3
REML modified Knapp–Hartung	DFS/PFI	1.542	0.839–2.834	p=0.0922; FDR=0.760	Not significant with small-k uncertainty

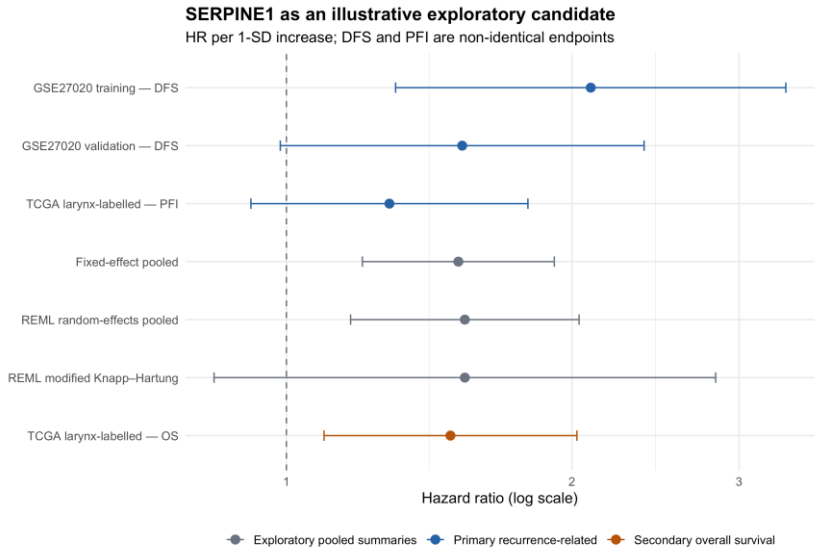

Figure 6. *SERPINE1* as an illustrative exploratory candidate. Individual DFS/PFI estimates, fixed-effect, REML, modified Knapp–Hartung, and secondary OS estimates are separated.

4. Discussion

4.1 Principal findings

This study produced two levels of evidence. The bulk-tissue transcriptomic phenotype was strong, internally robust, and reproducible: 435 features met the primary threshold, 418 remained after independent refiltering of QC sensitivity pairs, 396 remained after quantile-normalization sensitivity, and 184 of 327 measurable genes replicated in 23 paired TCGA cases. Prognostic evidence was weaker. No gene satisfied corrected independent validation and none remained significant under modified Knapp–Hartung correction.

The central scientific message is therefore not that transcriptomics failed, but that diagnostic/tumor-associated reproducibility and prognostic reproducibility are different statistical problems. Expression contrasts benefited from paired designs with 57 and 23 within-patient comparisons, strong biological effects, and many expressed genes. Prognostic analyses had 13, 21, and 42 recurrence/progression events, respectively, while testing hundreds of correlated candidates. Even biologically plausible genes could produce compatible hazard directions without enough information for corrected independent significance.

4.2 Squamous differentiation and keratinization

The DEG profile captured a profound reorganization of epithelial differentiation. Upregulation of *KRT14*, *KRT16*, *KRT17*, *KRT6B*, *S100A7*/*S100A7A*, SPRR-family genes, and cornified-envelope pathways is consistent with a hyperproliferative, stress-associated squamous program. At the same time, reductions in *KRT4*, *CRNN*, *MAL*, *TJP3*, mucosal secretory genes, and salivary-defense programs indicated loss of normal differentiated epithelial functions. The apparently bidirectional enrichment of cornified-envelope biology reflected different gene subsets: tumor tissue did not simply become "more differentiated," but adopted an abnormal keratinization state while losing components of healthy mucosal identity. A recent multi-dataset LSCC analysis likewise identified extracellular-matrix and mucin-related transcriptional programs.[18]

HOXB7, *FOXM1*, *FEN1*, *BIRC5*, and *KPNA2* supported increased proliferative and genome-maintenance activity. *FOXM1* and *BIRC5* frequently mark cycling, therapy-resistant epithelial populations, whereas *FEN1* participates in DNA replication and repair. Bulk expression cannot establish whether these genes are uniformly expressed across malignant cells or concentrated in highly proliferative subclones. Nevertheless, their coordinated increase across paired tissue comparisons argues against an isolated-marker explanation.

The strong downregulation of *PRR4*, *PRH1*/2, *PRB3*, *STATH*, *PIGR*, *LYZ*, *BPIFB1*, *MUC5B*, and *ZG16B* suggested loss of secretory and innate-defense functions. Some of these transcripts may originate from glandular, immune, or differentiated epithelial compartments rather than malignant cells. Their reduction can therefore reflect cell-composition changes in addition to transcriptional repression within carcinoma cells. Recent LSCC studies that integrated public expression data with immunohistochemistry for *ZNF71* [19] or evaluated *VCP*/p97 protein expression clinically[20] underscore why tissue localization and protein-level confirmation are needed. Single-cell or spatial data would be required to distinguish cell-composition shifts from within-cell transcriptional repression.

4.3 Extracellular matrix, invasion, and inflammatory signaling

Extracellular matrix remodeling was one of the most consistent themes. *MMP1* showed a paired mean $\log_2$(FPKM + 1) increase of 3.66, while *COL1A1*, *COL4A1*, *SERPINH1*, *TGFBI*, *FN1*, *LAMB3*, *PLAU*, *SERPINE1*, and *SPP1* linked malignant tissue to collagen organization, matrix turnover, adhesion, and stromal interaction. Recent LSCC transcriptomic network studies independently emphasized extracellular-matrix, basement-membrane, and focal-adhesion programs.[12,18] Our paired cross-platform results extend those observations from selected hub genes to a reproducible bulk-tissue program.

MMP1 can facilitate collagen degradation and invasion, but its bulk-tissue expression may arise from carcinoma cells, cancer-associated fibroblasts, or both. *COL1A1*, *FN1*, *SERPINH1*, and *TGFBI* similarly point toward fibroblast-rich or desmoplastic tissue states. *SPP1* may reflect tumor-associated macrophages as well as malignant cells. Therefore, an "ECM pathway" should not be interpreted as a single tumor-cell-autonomous mechanism. Rather, it represents a multicellular tissue phenotype that may affect invasion, treatment response, and immune exclusion.

IL-17 signaling enrichment was compatible with the broader inflammatory and stromal patterns reported in recent LSCC transcriptomic and immune-microenvironment studies.[5,10,13] IL-17-associated genes can be expressed by immune, epithelial, and stromal compartments and may integrate inflammatory recruitment with epithelial stress and matrix remodeling. The present over-representation analysis does not measure cytokine activity directly. Protein assays, phosphoproteomics, spatial localization, or perturbation experiments would be required to establish whether IL-17 signaling is a driver, a response, or a correlate of LSCC tissue remodeling.

4.4 Cross-platform reproducibility of tumor-associated genes

The TCGA paired analysis guarded against cohort-specific findings. GSE127165 and TCGA differed in geography, processing, collection, normalization, and symbol coverage. Despite this, 89.0% of measurable discovery genes had the same direction and 184 replicated at corrected significance. The effect

correlation of 0.774 supports rank concordance across processed scales, not numerical equivalence of their slopes.

Several large-effect genes were not represented or exactly matched in the older Xena matrix, explaining why only 327 of 435 could be tested. Failure to test is not evidence of biological non-replication. Conversely, non-significance among measurable genes may reflect limited power in 23 TCGA pairs rather than true inconsistency. Directional concordance complements FDR counts and reduces overinterpretation of an arbitrary threshold.

The paired design controlled stable patient-level factors, including germline background and many exposure-related characteristics. It did not control regional tissue composition or field effects. Adjacent mucosa may already harbor molecular alterations, and TCGA "solid tissue normal" specimens are also tumor-adjacent. Agreement between the two datasets therefore establishes reproducibility of tumor-versus-adjacent contrasts, not performance against truly healthy community controls.

4.5 Prognostic evidence and the case of SERPINE1

SERPINE1 is an illustrative exploratory candidate, not the study's positive biomarker result. It was strongly upregulated, nominally associated with training DFS, borderline in validation, non-significant for primary TCGA PFI, and nominally associated with secondary OS after multiplicity correction failed. *SERPINE1* encodes plasminogen activator inhibitor-1 and links fibrinolysis, matrix turnover, adhesion, and tumor–stroma signaling. Prior HNSCC research[15] and the 2026 systematic review[16] already connected higher expression with adverse OS and DFS but not consistently with progression-related endpoints. Our result is therefore an endpoint-qualified larynx-focused reanalysis, not discovery.

The endpoint pattern matters. The TCGA PFI hazard ratio was in the adverse direction but not significant, whereas OS associations were stronger. A true recurrence-specific biomarker should preferably reproduce for recurrence-like endpoints. OS can incorporate recurrent disease, treatment effects, second primary tumors, cardiopulmonary mortality related to smoking, and other causes. The apparent OS association may be clinically relevant, but it cannot repair the missing TCGA PFI validation.

The stage-stratified adjusted OS association for *SERPINE1* was stronger than the univariable association. This could reflect negative confounding, cohort imbalance, or selection induced by complete-case analysis. It should not be interpreted as definitive independence. Only 20 TCGA patients were female, 15 lacked pathological-stage grouping, detailed treatment covariates were incomplete, and the adjusted analysis retained 101 patients. The unexpectedly lower estimated hazard among males further suggests that covariate effects in this selected cohort should not be generalized.

The meta-analysis illustrates another methodological point. Fixed-effect and REML normal-theory procedures labeled *SERPINE1* significant after FDR correction, yet modified Knapp–Hartung inference did not. With three estimates, two arising from one source study, uncertainty in between-cohort variance, dependence, and endpoint comparability is substantial. We therefore classify *SERPINE1* as hypothesis-generating, not validated.

4.6 Comparison with previous prognostic signatures

Recent LSCC studies have reported four-mRNA, recurrence-score, RNA-binding-protein, long-non-coding-RNA, immune-gene, hub-gene, neutrophil-extracellular-trap, and lactate-related prognostic models.[6–14] Our question was different: whether genes selected because they were differentially expressed in GSE127165 individually predicted outcome across cohorts. A multigene score can perform even when no component is individually significant after multiplicity correction, because correlated weak effects can combine. Our negative single-gene results therefore do not directly validate or invalidate those composite models.

Biomarker development in LSCC has also expanded beyond bulk risk scores to serum miRNA signatures,[21] paired tissue–serum miR-449a assessment,[22] longitudinal serum Nov/uPAR measurements,[23] HSP27/HSP70 protein expression,[24] and combined single-cell/bulk T-cell analyses.[25] These primary studies demonstrate diverse assay and validation strategies, but none constitutes external validation of the present DEG-to-single-gene prognostic pipeline. The incremental contribution here is paired modeling, cross-platform paired replication, locked validation preprocessing, transparent attrition, and retention of negative prognostic evidence. Models that select, tune, dichotomize, and evaluate within modest reused cohorts remain vulnerable to optimism.

Our study deliberately did not construct a new high-dimensional risk score. With only 34 total DFS events across GSE27020 and 42 TCGA PFI events, developing and validating a multivariable expression signature would create a high risk of optimism. The more informative contribution was to quantify which tumor-associated genes replicated biologically and how far the prognostic evidence survived increasingly conservative analyses.

4.7 Methodological strengths

Strengths included paired patient modeling; predefined thresholds; independent refiltering after pair-level QC exclusion; a quantile-normalization sensitivity analysis; direction-specific enrichment with reported effective denominators; training-only locked probe selection; continuous survival modeling; full univariable proportional-hazards diagnostics; larynx-labeled TCGA restriction; primary PFI selection; complete-case three-cohort pooling; and REML plus modified Knapp–Hartung uncertainty. These safeguards prevented nominal pooled significance from being mislabeled as validation.

The master workflow reproduces GEO and TCGA analyses in sequence and records checksums, versions, complete outputs, and the empty strict-candidate table. A permanent archived public release is required for independent access; a local package alone does not satisfy publication-level availability.

4.8 Limitations

The study has important limitations. GSE127165 provided FPKM rather than counts. Log-transformed FPKM with limma-trend and quantile-normalization sensitivity can evaluate robustness of the deposited matrix, but cannot reconstruct alignment, counting, compositional normalization, or count-level mean–variance uncertainty. A fixed-reference raw-read reanalysis remains preferable.

Bulk tissue mixes malignant epithelial cells, fibroblasts, endothelial cells, immune infiltrates, glands, and normal epithelium. Consequently, DEG and pathway results cannot be assigned to specific cell types. Tumor purity and stromal composition were not incorporated. Deconvolution could help but would introduce reference-dependent assumptions; single-cell or spatial transcriptomics would provide stronger localization.

Adjacent tissue was not healthy control tissue. Field cancerization, inflammation, and proximity to tumor may attenuate or reshape contrasts. The paired design improves control of patient-level variables but does not establish diagnostic specificity in cancer-free individuals. A future diagnostic study should include laryngeal mucosa from non-cancer controls and account for smoking, inflammation, benign lesions, and anatomical subsite.

Clinical annotations differed. GSE27020 used DFS after surgery with or without radiotherapy, whereas TCGA PFI can include progression, recurrence, new tumor events, or cancer-related death. TCGA treatment was heterogeneous and incompletely recorded. Standardization aligned HR units but did not make endpoints or populations exchangeable; pooled analyses were exploratory and cannot be generalized to definitive radiotherapy, concurrent chemoradiotherapy, organ-preservation protocols, salvage laryngectomy, or recurrent/metastatic therapy.

The prognostic cohorts were small relative to the candidate set. Training events numbered 13, validation events 21, and TCGA PFI events 42. BH adjustment correctly imposed a high evidentiary bar, but power was limited. A gene can be truly associated yet fail independent significance. Conversely, pooled significance can arise from weak correlated signals and does not prove clinical usefulness.

Multivariable TCGA analysis was limited to *SERPINE1* and a small covariate set. Fifteen patients were excluded from complete-case models; early-stage event counts were very small; global proportionality was violated in stage-covariate sensitivity models. Moreover, 115/116 tumors were C32.9, so glottic–supraglottic transportability cannot be assessed. Detailed T/N, extranodal extension, treatment, HPV/p16, smoking intensity, performance status, and comorbidity could not be modeled without overfitting. Calibration, time-dependent discrimination, decision benefit, assay reproducibility, and incremental value were not assessed.

Finally, a comprehensive systematic review and formal novelty assessment were outside the computational analysis. The targeted literature search identified prior *SERPINE1* and LSCC signatures, so "novel biomarker" terminology is not justified. The findings therefore emphasize cross-platform replication, evidence grading, and the gap between DEG discovery and prognostic validation.

4.9 Clinical and research implications

The replicated transcriptomic programs provide a rational shortlist for mechanistic and translational work. Matrix and adhesion genes could be evaluated by multiplex immunohistochemistry or spatial transcriptomics to determine whether expression is tumor-cell, fibroblast, macrophage, or endothelial in origin. Keratinization programs could be related to histological differentiation, subsite, smoking, treatment response, and recurrence pattern. IL-17-associated signals could be tested with cytokine measurements and spatial immune profiling.

For prognosis, the next step should be a confirmatory study rather than another unconstrained public-dataset signature. Any small research-prioritization set—including *SERPINE1*, *SERPINH1*, *IGFBP3*, *LAMB3*, *PLAU*, *CALML5*, and alternatives—must be locked before outcome analysis and should not be called a panel for clinical use. A GEO/TCGA-independent, larynx-subsite-resolved cohort should evaluate postoperative recurrence-risk refinement beyond a locked clinical model, with calibration, decision utility, proportionality, and assay reproducibility.

An external validation cohort should also distinguish local recurrence, regional recurrence, distant metastasis, second primary cancer, and non-cancer death. Competing-risk methods may be necessary, especially in tobacco-exposed populations with cardiopulmonary comorbidity. If *SERPINE1* remains associated with OS but not recurrence, its role may be as a marker of systemic risk, aggressive tissue ecology, or comorbidity-linked biology rather than a recurrence-specific tumor biomarker.

5. Conclusion

This secondary cross-platform reanalysis identified a reproducible bulk-tissue LSCC phenotype characterized by abnormal squamous differentiation, loss of mucosal and secretory features, matrix remodeling, adhesion, and inflammatory over-representation. Prognostic evidence was not independently validated: no candidate met the operational rule or survived modified Knapp–Hartung multiplicity correction. The study supports biological reproducibility and research prioritization, not a novel prognostic biomarker or any change to surveillance,

adjuvant therapy, organ preservation, or systemic treatment.

Declarations

Ethics approval and consent to participate

This study used only public, de-identified GEO and TCGA data and involved no participant contact or intervention. No new consent was obtained. Institutional review-board approval was not required for this secondary analysis of publicly available, de-identified data.

Consent for publication

Not applicable.

Availability of data and materials

GSE127165 and GSE27020 are available through NCBI GEO. TCGA-HNSC expression, phenotype, and curated survival data are available through UCSC Xena and the TCGA Clinical Data Resource. The reproducibility package contains the master workflow, component scripts, executed notebook, input manifest with SHA-256 checksums, inclusion lists, session information, complete differential-expression and enrichment tables, survival outputs, proportional-hazards diagnostics, meta-analysis results, and the empty strict-candidate table. The reproducibility package will be deposited in a permanent public repository on publication.

Competing interests

The authors declare that they have no competing interests.

Funding

No specific funding was received.

Authors' contributions

R.H.: conceptualization and clinical interpretation; M.M.: methodology, software, formal analysis, and visualization; R.H. and M.M.: data curation and validation. All named authors confirm source and code verification, interpretation, manuscript revision, and accountability for the final work.

Acknowledgements

The authors acknowledge the patients and investigators who contributed data to GEO and TCGA, and the maintainers of NCBI GEO, UCSC Xena, Bioconductor, HGNC, and g:Profiler.

Use of artificial intelligence tools

None.

6. References

1. Sung H, Filho AM, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. CA Cancer J Clin. 2026;76(4):e70090. doi:10.3322/caac.70090.
2. Kim DH, Kim SW, Han JS, et al. The prognostic utilities of various risk factors for laryngeal squamous cell carcinoma: a systematic review and meta-analysis. Medicina (Kaunas). 2023;59(3):497. doi:10.3390/medicina59030497.
3. Hsia B, Bitar GA, Tran NS, et al. Genomic profiling of laryngeal squamous cell carcinoma reveals novel biomarkers for precision medicine. J Pers Med. 2025;16(1):2. doi:10.3390/jpm16010002.
4. Li Y, Lin C, Chu Y, et al. Characterization of cancer stem cells in laryngeal squamous cell carcinoma by single-cell RNA sequencing. Genomics Proteomics Bioinformatics. 2024;22(4):qzae056. doi:10.1093/gpbjnl/qzae056.
5. Zhu X, Heng Y, Ma J, et al. Prolonged survival of neutrophils induced by tumor-derived G-CSF/GM-CSF promotes immunosuppression and progression in laryngeal squamous cell carcinoma. Adv Sci (Weinh). 2024;11(46):e2400836. doi:10.1002/advs.202400836.
6. Zhang C, Shen B, Chen X, et al. Identification of a prognostic 4-mRNA signature in laryngeal squamous cell carcinoma. J Cancer. 2021;12(19):5807–5816. doi:10.7150/jca.47557.
7. Liu Y, Gao Z, Peng C, et al. Construction of a 10-gene prognostic score model of predicting recurrence for laryngeal cancer. Eur J Med Res. 2022;27(1):249. doi:10.1186/s40001-022-00829-2.
8. Shen Y, Chen H, Huang Q, et al. Transcriptomic signature associated with RNA-binding proteins for survival stratification of laryngeal cancer. Aging (Albany NY). 2022;14(16):6605–6625. doi:10.18632/aging.204234.
9. Zheng X, Gao W, Zhang Z, et al. Identification of a seven-lncRNAs panel that serves as a prognosis predictor and contributes to the malignant progression of laryngeal squamous cell carcinoma. Front Oncol. 2023;13:1106249. doi:10.3389/fonc.2023.1106249.
10. Hong X, Li Y, Lv Q, et al. Identification of an immune-related genes signature to predict risk of recurrence for patients with laryngeal squamous cell carcinoma. Int J Immunopathol Pharmacol. 2023;37:3946320231172075.doi:10.1177/03946320231172075.
11. Dong Y, Zhao K, Qiu Q, et al. Ferroptosis-associated lncRNA prognostic signature predicts prognosis and immune response in laryngeal squamous carcinoma. Cell Mol Biol (Noisy-le-grand). 2023;69(12):223–231. doi:10.14715/cmb/2023.69.12.36.
12. Li H, Bo S, Guo Y, et al. Identification of hub genes and key modules in laryngeal squamous cell carcinoma. Transl Cancer Res. 2024;13(7):3771–3782. doi:10.21037/tcr-24-104.
13. Wu G, Jin R, Liao J, et al. Molecular subtype and prognostic model of laryngeal squamous cell carcinoma based on neutrophil extracellular trap-related genes. Transl Cancer Res. 2025;14(3):1772–1785. doi:10.21037/tcr-24-1531.
14. Liu J, Wang YT, Zeng LL, et al. Lactate-related genes signature as a novel prognostic landscape in laryngeal squamous cell carcinoma: insights from 156 machine learning algorithms and in vitro validation. Int J Biol Macromol. 2025;323(Pt 1):147098. doi:10.1016/j.ijbiomac.2025.147098.
15. Zhu C, Liu H, Li Z, et al. Prognostic significance and therapeutic potential of *SERPINE1* in head and neck squamous cell carcinoma. Cancer Med. 2025;14(2):e70605. doi:10.1002/cam4.70605.

16. Yan S, Li X, Zhu C, et al. The prognostic value of *SERPINE1* in clinical outcomes in head and neck squamous cell carcinoma: a systematic review and meta-analysis. Technol Cancer Res Treat. 2026;25:15330338261435460.
doi:10.1177/15330338261435460.
17. Yuan X, Lau HC, Huang H, et al. Integrative methylome and transcriptome analysis reveals epigenetic regulation of Fusobacterium nucleatum in laryngeal cancer. Microb Genom. 2024;10(3):001221.
doi:10.1099/mgen.0.001221.
18. Mo BY, Li GS, Huang SN, et al. The underlying molecular mechanism and identification of transcription factor markers for laryngeal squamous cell carcinoma. Bioengineered. 2021;12(1):208–224.
doi:10.1080/21655979.2020.1862527.
19. Jiang FC, Luo JY, Dang YW, et al. Downregulation of zinc finger protein 71 in laryngeal squamous cell carcinoma tissues and its potential molecular mechanism and clinical significance: a study based on immunohistochemistry staining and data mining. World J Surg Oncol. 2022;20(1):359.
doi:10.1186/s12957-022-02823-8.
20. Seuthe IMC, Hunsicker-Biederbeck HC, Ruwe M, et al. Valosin containing protein (*VCP*/p97) expression in laryngeal squamous cell carcinoma: clinical evaluation and implications for targeted therapy. Pathol Oncol Res. 2025;31:1612202. doi:10.3389/pore.2025.1612202.
21. Falco M, Tammaro C, Cossu AM, et al. Identification and bioinformatic characterization of a serum miRNA signature for early detection of laryngeal squamous cell carcinoma. J Transl Med. 2024;22(1):647.
doi:10.1186/s12967-024-05385-3.
22. Gharib AF, Alsalmi O, Alhuthali HM, et al. miR-449a: a novel biomarker for diagnosis, prognosis, and treatment response in locally advanced laryngeal squamous cell carcinoma. Oncol Res. 2026;34(3):21.
doi:10.32604/or.2025.073051.
23. Winkelmann AC, Wendler O, Schuster J, et al. Early detection and longitudinal follow-up of non-invasive biomarkers for laryngeal squamous cell carcinoma: first results of the BEAL-Study. Cancer Med. 2025;14(10):e70961. doi:10.1002/cam4.70961.
24. Borowczak J, Łaszczych D, Czyżnikiewicz A, et al. Low expression of HSP27 and HSP70 predicts poor prognosis in laryngeal squamous cell carcinoma. J Cancer Res Clin Oncol. 2025;151(10):264.
doi:10.1007/s00432-025-06309-4.
25. Pan X, Wang J, Liu F, et al. Tumor microenvironment T-cell heterogeneity and prognostic model construction in laryngeal squamous cell carcinoma by single-cell and bulk RNA sequencing. Comput Methods Biomech Biomed Engin. 2025:1–31.
doi:10.1080/10255842.2025.2570515.