

NF- κ B over-activation portends improved outcomes in HPV-associated head and neck cancer

Travis P. Schrank^{1,*}, Andrew C. Prince^{1,*}, Tejas Sathe^{2,3}, Xiaowei Wang^{4,5}, Xinyi Liu^{4,5}, Damir T. Alzhanov¹, Barbara Burtness^{6,9}, Albert S. Baldwin^{6,9}, Wendell G. Yarbrough^{1,7,8,#} and Natalia Issaeva^{1,7,8,#}

¹Department of Otolaryngology/Head and Neck Surgery, UNC, Chapel Hill, NC 27599, USA

²Department of Surgery, Otolaryngology, Yale, New Haven, CT 06519, USA

³Current address: Department of Surgery, Columbia University, New York, NY 10032, USA

⁴Department of Pharmacology and Bioengineering, University of Illinois at Chicago, Chicago, IL 60612, USA

⁵Bioinformatics Core, University of Illinois Cancer Center, Chicago, IL 60612, USA

⁶Department of Medicine, Yale School of Medicine, New Haven, CT 06510, USA

⁷Lineberger Cancer Center, UNC, Chapel Hill, NC 27599, USA

⁸Department of Pathology and Laboratory Medicine, UNC, Chapel Hill, NC 27599, USA

⁹Yale Cancer Center, Yale School of Medicine, New Haven, CT 06510, USA

*These authors contributed equally to this work

#Senior authors

Correspondence to: Wendell G. Yarbrough, **email:** dell@med.unc.edu
Natalia Issaeva, **email:** natalia.issaeva@med.unc.edu

Keywords: HPV; head and neck cancer; CYLD; TRAF3; NF- κ B

Received: March 23, 2022

Accepted: May 03, 2022

Published: May 24, 2022

Copyright: © 2022 Schrank et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Evolving understanding of head and neck squamous cell carcinoma (HNSCC) is leading to more specific diagnostic disease classifications. Among HNSCC caused by the human papilloma virus (HPV), tumors harboring defects in *TRAF3* or *CYLD* are associated with improved clinical outcomes and maintenance of episomal HPV. *TRAF3* and *CYLD* are negative regulators of NF- κ B and inactivating mutations of either leads to NF- κ B overactivity. Here, we developed and validated a gene expression classifier separating HPV+ HNSCCs based on NF- κ B activity. As expected, the novel classifier is strongly enriched in NF- κ B targets leading us to name it the NF- κ B Activity Classifier (NAC). High NF- κ B activity correlated with improved survival in two independent cohorts. Using NAC, tumors with high NF- κ B activity but lacking defects in *TRAF3* or *CYLD* were identified; thus, while *TRAF3* or *CYLD* gene defects identify the majority of tumors with NF- κ B activation, unknown mechanisms leading to NF- κ B activity also exist. The NAC correctly classified the functional consequences of two novel *CYLD* missense mutations. Using a reporter assay, we tested these *CYLD* mutations revealing that their activity to inhibit NF- κ B was equivalent to the wild-type protein. Future applications of the NF- κ B Activity Classifier may be to identify HPV+ HNSCC patients with better or worse survival with implications for treatment strategies.

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is a devastating disease that impairs fundamental tissues involved in respiration, phonation, and digestion. It is categorized into two discrete diseases based on etiology:

human papillomavirus (HPV) negative HNSCC, which is primarily caused by exposure to ethanol and tobacco, and HPV-associated (HPV+) HNSCC [1]. These forms of HNSCC have contrasting clinical, epidemiological, and histological features [2–4] with HPV+ HNSCC occurring in a younger population with less or no smoking history [5,

6]. HPV-mediated cancer arises primarily in the reticulated epithelia of the oropharynx (e.g., tonsils, base of tongue), whereas HPV-negative HNSCC is found at all subsites (e.g., oral cavity, larynx) [2]. Unfortunately, the global incidence of HPV+ HNSCC is increasing, and for at least a decade, HPV has caused more head and neck cancers than uterine cervical cancers annually in the United States [7, 8].

Since HPV+ HNSCC has only recently been recognized as a distinct clinicopathological entity [9], management of HNSCC has been driven by escalating therapies to improve cancer control in the more treatment-resistant HPV-negative HNSCC [2, 6]. While oncologic outcomes for HPV+ HNSCC are generally favorable, treatment paradigms developed for HPV-negative disease burden many survivors of HPV+ HNSCC with lifelong debilitating treatment-associated side effects [10]. On the other hand, ~30% of HPV+ HNSCC patients exhibit a more aggressive disease course and suffer recurrence [11, 12]. Therefore, there is a growing clinical demand to develop robust stratification tools to accurately identify patients with good or poor prognosis and that could be used to personalize treatment.

TRAF3 belongs to the TRAF family of proteins that are known as intracellular adaptors and E3 ubiquitin ligases mediating receptor-based signaling [13]. TRAF3 polyubiquitinates and degrades NF- κ B-inducing kinase (NIK) restraining non-canonical NF- κ B signaling. The deubiquitinating enzyme Cyldromatosis (CYLD) is a tumor suppressor that was found to be mutated in familiar cylindromatosis, a condition associated with benign skin tumors. CYLD mediates deubiquitination of the NF- κ B essential modulator (NEMO) thus inhibiting canonical NF- κ B signaling [14, 15]. A cross talk between canonical and non-canonical NF- κ B signaling suggests that TRAF3 and CYLD affect both NF- κ B pathways.

Somatic defects in the NF- κ B inhibitors *TRAF3* and *CYLD* are found in ~30% of HPV+ HNSCC tumors [1, 16, 17]. These gene defects are uncommon in uterine cervical cancer and HPV-negative HNSCC. While frequent *TRAF3* or *CYLD* inactivating mutations are found in B cell lymphomas, where constitutive NF- κ B activity is known to play a key survival role [18–20], these mutations are rarely found in solid tumors [16]. Exceptions with more frequent *TRAF3* and *CYLD* mutations include two virally-associated cancers, HPV+ HNSCC and Epstein-Barr virus-associated nasopharyngeal carcinoma (NPC) [21–23]. Although initial studies focused on NF- κ B activity as a defense against viral infections, further investigation revealed more nuances with some viruses, like EBV and HIV, depending on NF- κ B activity to support viral replication and viral gene expression [24–27]. The correlation between TRAF3 and CYLD alterations and the lack of classic oncogenic HPV integration events, suggests

that HPV may similarly exploit NF- κ B activity in HNSCC to be able to maintain extrachromosomal HPV genomic material.

The power of multi-variable models and/or multi-omic approaches can be harnessed to improve tumor subtyping [28–31]. For example, an RNA expression-based PARP inhibitor outcome prediction model in ovarian cancer outperformed BRCA1/2 mutational status in predicting treatment response [30]. In the present study, transcriptional differences between tumors with and without *TRAF3* and *CYLD* defects formed the basis for a novel classification of HPV+ HNSCC. Based on established roles of *TRAF3* and *CYLD* as inhibitors of NF- κ B, it was expected that the resultant classifier would segregate tumors on the basis of NF- κ B activity. Gene set enrichment analysis confirmed that the classifier identified tumors with high or low NF- κ B activity and, relative to TRAF3 and CYLD defects, this NF- κ B Activity Classifier (NAC) improved identification of tumors with good and poor survival. Among TCGA specimens, two novel missense mutations in *CYLD* were identified: N300S and D618A [16]. To understand the implications of these point mutations, we used the NAC and correlated results with a cell-based assay to evaluate their effect on NF- κ B transcriptional activity.

To improve on genomic classification, we designed this study to provide a foundation for development of NF- κ B related, RNA based classification strategies to better identify HPV+ HNSCC patients with good or poor prognosis that could potentially aid in future efforts towards treatment personalization.

RESULTS

Development of the NF- κ B activity classifier (NAC)

We previously reported that TRAF3 and CYLD alterations in a subset of HPV+ oropharyngeal squamous cell carcinoma (OPSCC) tumors correlated with NF- κ B activation and improved survival [16]. NF- κ B is a family of inducible transcription factors that play a role in innate and adaptive immune response; constitutively active NF- κ B is a well-known oncogene in various cancer types that increases cell proliferation, migration, invasion and metastasis while inhibiting apoptosis [32–34]. NF- κ B activation, induced by carcinogens or oncogenic viruses, was found in head and neck tumors and cells [35]. Given the variable role that NF- κ B plays in HPV+ OPSCC tumorigenesis, we hypothesized that tumor groups based on NF- κ B related gene expression may correlate with treatment outcome, considering that tumors lacking defects in *TRAF3* and *CYLD* may have unrecognized mechanisms driving constitutive NF- κ B activation. TCGA expression data were first grouped by the presence of a known *TRAF3* or *CYLD* defect and the top 100

differentially expressed genes identified (Figure 1A). As anticipated, gene set enrichment analyses demonstrated a high enrichment score (>0.3) for NF-κB target genes (Figure 1B, Grey) and several notable NF-κB target genes were differentially expressed – *TRAF2*, *NF-κB2*, *BIRC3*, and *MAP3K14*.

Machine learning techniques were used to refine the signature resulting in a set of 50 key genes dubbed the NF-κB Activity Classifier Gene Signature (Supplementary Table 1). Using the NF-κB Activity Classifier (nearest centroid), all tumors were then given a final classification to identify tumors with high NF-κB activity (Figure 1, track 1, green). As may be expected based on unknown mechanisms of NF-κB activation, some additional samples

without inactivating alterations (deep deletion, nonsense/frameshift mutation) in either *TRAF3* or *CYLD* (Figure 1A, track 3 – burnt orange) were included in the NF-κB active group.

In order to identify a set of tumors with equivalently high activation of NF-κB, as observed with destructive nonsense or frameshift mutations in *TRAF3* or *CYLD*, we also defined a more stringent threshold of NF-κB activation, based on the lowest classifier score observed for the highest confidence destructive alterations (nonsense or frameshift) of *TRAF3* or *CYLD* (see Figure 1, track 2 – green brown). Notably, 6 tumors included in this “highly active” NF-κB group also were found to be without deep deletion, frameshift/nonsense

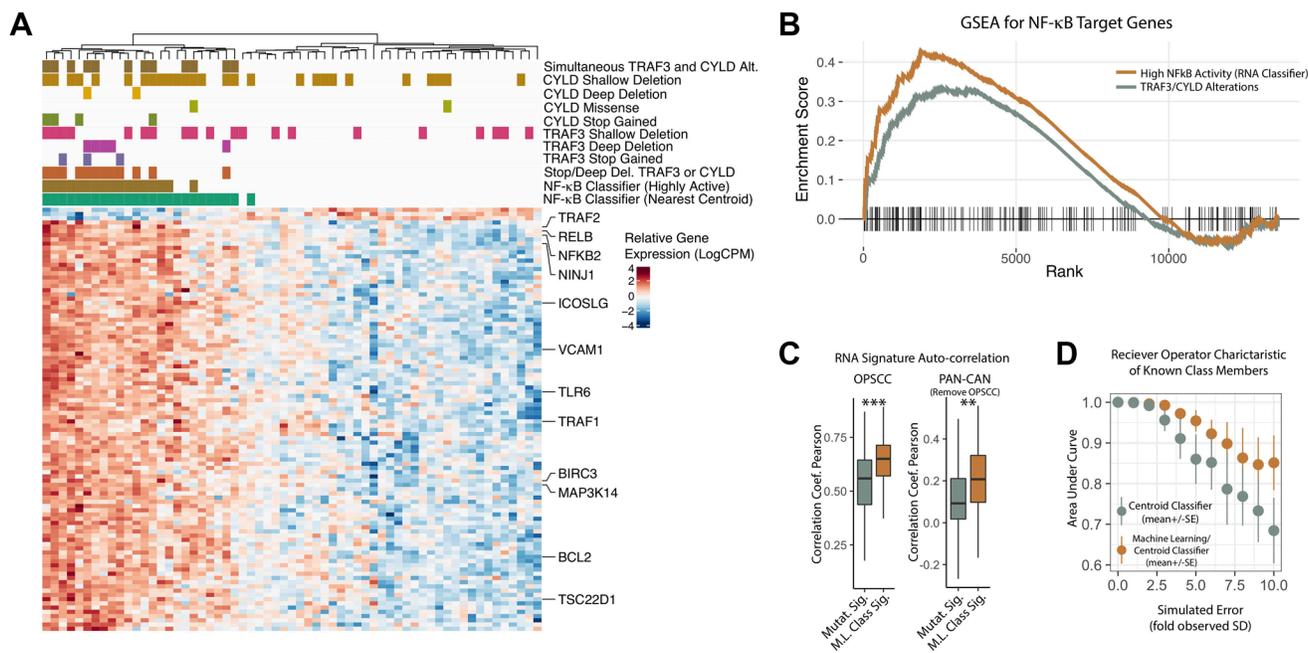


Figure 1: Development of an NF-κB activity related RNA expression classifier. (A) Heatmap of RNA Expression Changes Associated with TRAF3/CYLD Alterations and Deletions. Normalized \log_2 (read counts per million), color scaled by row. Columns– Tumor Samples, organized by unguided clustering. Rows – Top 100 genes by *p*-value differentially expressed between high-confidence NF-κB active vs. inactive tumors (see methods for details). Row annotation – Known NF-κB target genes curated from literature review. Column Annotation Details: Track 1 (green) - RNA classifier (“NF-κB active”) based on nearest centroid. Track 2 (green brown) - RNA classifier (“NF-κB highly active”) based on minimal classifier score identified for TRAF3/CYLD nonsense or frameshift mutation bearing tumors. Track 3 (orange) – Tumor contains a frameshift, nonsense, or deep deletion in TRAF3 or CYLD. Track 4 (purple) - Tumor contains a frameshift or nonsense mutation in TRAF3. Track 5 (lavender) - Tumor contains a deep deletion in TRAF3. Track 6 (pink) - Tumor contains a shallow deletion in TRAF3. Track 7 (army green) - Tumor contains a frameshift or nonsense mutation in CYLD. Track 8 (lime green) - Tumor contains a missense mutation in CYLD. Track 9 (yellow) - Tumor contains a deep deletion in CYLD. Track 10 (mustard) - Tumor contains a shallow deletion in CYLD. Track 11 (dark brown) – Tumor contains any alteration in both TRAF3 and CYLD. Shallow Deletion – Gistic copy-number score = -1, Deep Deletion – Gistic copy-number score = -2, Stop Gained – frameshift or nonsense mutation. Missense – missense or in frame indel. Stop/Deep Del. – Any one of nonsense, frameshift, or deep deletion. (B) Gene Set Enrichment Analysis for NF-κB Target Genes. All available genes after data filtering (see methods) were ranked according to signal-to-noise ratio when comparing the two groups of tumors. The MiSigDB Hallmark TNFA/NFκB gene set was tested for enrichment. NF-κB highly active – tumors were defined according to RNA based classifications (see methods); these were compared to all other tumors in the study cohort. NF-κB Pathway Alteration – Any missense, nonsense, frameshift, deep deletion in TRAF3 and/or CYLD; these were compared to all other tumors in the study cohort. Lines – enrichment score values. Dashed Line – maximum achieved enrichment score (NFκB high activity only). Vertical Hashes – rank positions of the test gene set (Hallmark NF-κB). (C) Auto-correlation of RNA Gene Set before and after the machine learning (ML) procedure. (D) Classifier Performance of Gene Sets before and after ML improvement, with increasing (simulated) error of measurement. Performance determined by area under the receiver operating characteristic curve. ****P* value < 5×10^{-4} , ***P* value < 5×10^{-3} .

mutation of *TRAF3* or *CYLD*, bolstering the utility of an RNA based approach to identify NF- κ B activated HPV+ HNSCC tumors.

All tumors harboring concurrent alterations (including shallow deletions) in both *TRAF3* and *CYLD* were found to be in the NF- κ B active group (Figure 1A, track 11 - brown), and two of these tumors were included in the “highly active” NF- κ B group. These data suggest an intriguing hypothesis that combinations of more subtle changes simultaneously effecting both *TRAF3* and *CYLD* might also contribute to NF- κ B activation.

RNA-based classification strengthens the association with NF- κ B target gene expression

To determine if the NF- κ B Activity Classifier enhanced correlation with NF- κ B target genes relative to groupings based on *TRAF3/CYLD* alterations, we performed gene set enrichment analysis using *TRAF3/CYLD* (missense, nonsense, frame shift) and the highly active NF- κ B classification as determined by the NAC. This analysis demonstrated significant enrichment for the Hallmark NF- κ B target gene set for both *TRAF3/CYLD* and highly active NF- κ B classifiers (p -value < 0.01); however, stratification using the NF- κ B Activity Classifier demonstrated stronger enrichment (Figure 1B).

Machine learning (ML) improves NF- κ B gene set properties and classifier robustness

Auto-correlation, or compactness, is a desirable feature of RNA expression signatures since loss of compactness when applied to new datasets can limit their diagnostic utility [36]. To begin determining compactness of the NF- κ B signature auto-correlation was examined. Pearson correlation coefficients were improved after the machine learning procedure, both in the HNSCC tumors used for deriving the gene set; as well as across all tumor types included in the TCGA pan-cancer atlas (Figure 1C). Since clinical expression datasets might be expected to have more error compared to TCGA, we also considered how robust our classifications were to increasing noise of measurement. To examine this, we calculated the area under the receiver-operator characteristic curve (AUC) for the original and ML improved classifier with increasing levels of (random) simulated error applied to the RNA expression data. The ML-improved classifier had higher AUC values at higher levels of noise, and maintained a median AUC of >0.95 even with a five-fold increase in error as compared to the original RNA data from TCGA (Figure 1D). Taken together these analyses illustrate the favorable properties of our NF- κ B signature, as well as a high-degree of robustness of the nearest centroid classifications based on these genes.

Weighted gene correlation network analysis identifies an NF- κ B associated gene expression module in HPV+ HNSCC

As *TRAF3* and *CYLD* have other molecular functions in addition to inhibiting NF- κ B, and to determine the relationship of the NAC to other biological aspects, we performed weighted gene correlation network analysis (WGCNA). To render required processor times tractable, only the 13,000 most highly expressed genes were included in the WGCNA analysis, excluding 2 of the 50 classifier genes. This unguided discovery approach identified 7 sets (or modules) of highly autocorrelated genes; the relative size and correlative dissimilarity between the modules are displayed in Figure 2A. These modules were then screened for (hypergeometric) enrichment of the established hallmark gene sets from the MiSig database (Figure 2C). Interestingly, one module (“yellow”) was found to be most associated with NF- κ B target gene expression by both p -value and fraction of module genes in the test signature (Figure 2C). Of note, no other modules were enhanced for NF- κ B targets. Furthermore, 47 of 48 signature genes included in the WGCNA analysis were found to be in the “yellow” module (Figure 2B, Supplementary Table 2 for WGCNA modules, and Supplementary Table 3 for hypergeometric enrichment analysis). The “yellow” module was also associated with early estrogen receptor signaling (Figure 2C).

Expression-based classification improves correlation with survival

Clinical outcomes for the TCGA HPV+ HNSCC cohort were assessed with PFI, available for all TCGA samples [37]. Kaplan-Meier survival curves were created for samples stratified by the presence of a *TRAF3* or *CYLD* genomic alteration (Figure 3A) and using the NF- κ B Activity Classifier (Figure 3B). In both cases, a survival advantage was apparent for this distinct disease phenotype. However, the NF- κ B Activity Classifier was associated with a larger hazard ratio (HR = 6.8) and statistically significant difference in PFI (p = 0.01) (Figure 3A, 3B). Although fewer tumors (n = 57) were annotated for recurrence-free survival (RFS), classification of NF- κ B active tumors using the NAC also correlated with improved RFS (Supplementary Figure 1, p -value = 0.006).

NF- κ B activity correlates with HPV viral integration status

We previously reported that somatic alterations in *TRAF3* and *CYLD* were associated with lack of viral integration in HPV+ HNSCC. To examine if our RNA-based estimates of NF- κ B activity also correlated with viral integration, we first determined integration based

on discordant read pair mapping - sequences that mapped to both the human and HPV viral genomes. Tumors were only considered integrated if multiple discordant read pairs mapped to similar areas of the human and viral genomes [38]. The ratio of expression of viral genes *E6* and *E7* to *E1* and *E2* has been used as a surrogate marker for integration [39], however, in our hands the ratio of *E6/E7* to *E2/E5* was more correlated to integration identified by discordant read pairs (Figure 3C). Comparison of RNA-based NF-κB activity (classifier scores) demonstrated a strong relationship to viral integration status, with episomal tumors having much higher median NF-κB activity (Figure 3D, p -value < 0.001).

NF-κB activity correlates with patient outcome in an independent validation dataset

To validate the prognostic value of the NAC, we queried the literature for suitable datasets, finding one study with RNAseq data and clinical annotation (Supplementary Table 4, [40]). Since somatic mutational data was not available in this RNA expression dataset, we applied single-sample gene set enrichment analysis (ssGSEA) to score each tumor for NF-κB activity using the NAC (Figure 4A). Interestingly, NAC gene signature ssGSEA scores were distributed in a bimodal pattern,

enabling empiric classification of tumors based on a simple threshold roughly dividing the two distributions (Figure 4A). Recurrence-free survival (RFS) analysis demonstrated improved survival for the NF-κB active group (Figure 4B).

NF-κB activity classifier RNA signature maintains favorable properties in an independent validation dataset

To investigate the relationship to of the NF-κB activity gene signature to global variability in (human) gene expression, we performed principal component (PC) analysis (Figure 4C–4D). NF-κB activity groups were not strongly correlated with the principal component associated with the greatest degree of variability in the dataset (PC1). Among the 10 top principal components, only PC3 (and to a lesser degree PC2), were associated with the NF-κB activity groups (Figure 4C, 4D). Taken together, these results suggest that variability in the expression of the NF-κB activity gene signature is specific, and not simply a reflection of gross data variability. Principal component (PC3) and NAC gene signature ssGSEA scores were strongly correlated (Figure 4D inset, Pearson's Rho = -0.63, p -value = 5×10^{-12}), which suggests that expression of NF-κB activity signature genes

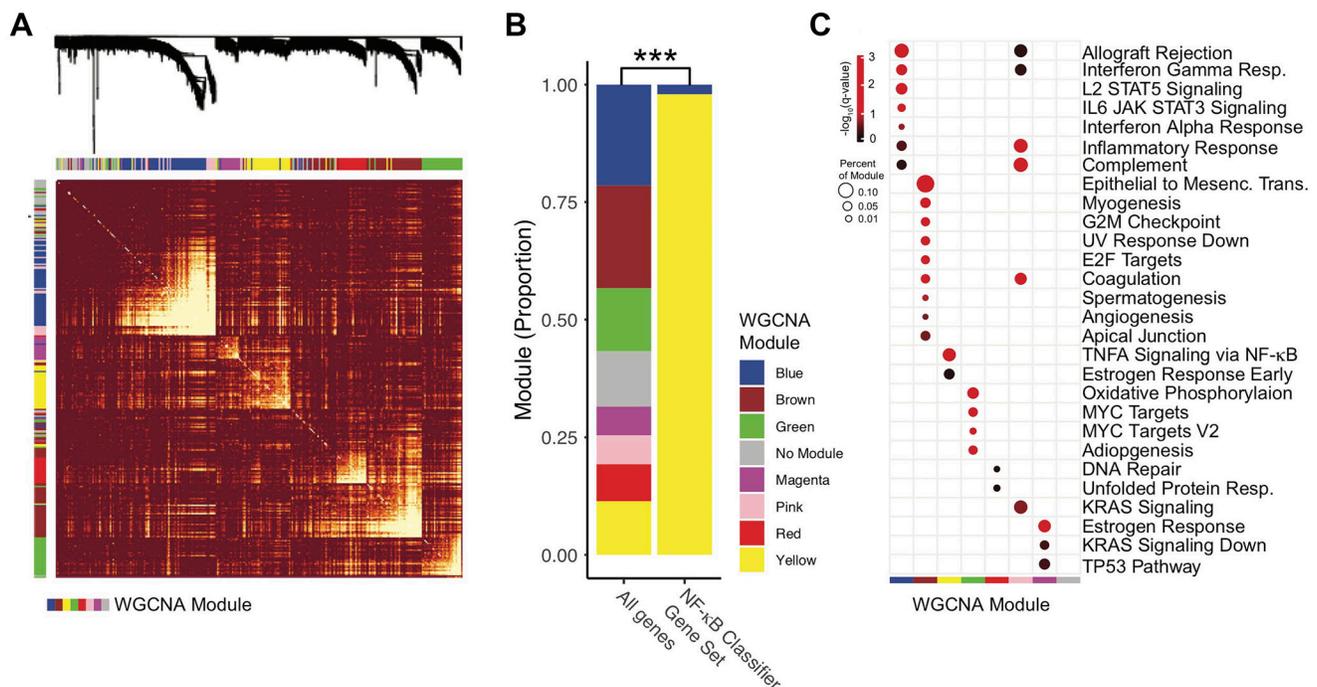


Figure 2: Characterization of the NF-κB activity classifier genes with weighted gene correlation network analysis (WGCNA). Only modules with more than 250 and less than 5000 genes were analyzed. (A) Expression Dissimilarity matrix with clustering dendrogram. For clarity, a subset of 1500 genes are displayed. Warmer colors (red) represent higher degrees of dissimilarity. Row and Column Annotations – WGCNA gene expression modules, colors correspond to module name, as in panel C. (B) Proportion of Genes by WGCNA module. NF-κB Classifier Gene Set – Gene set (50 genes) used in the NF-κB activity classifier. All genes – Genes analyzed by WGCNA but not included in the NF-κB activity classifier. P -value represent chi-squared test. *** p -value < 0.0001. (C) Hypergeometric Enrichment Plot. Identified WGCNA modules were screened for enrichment in Hallmark Gene Sets from MiSigDB. Red-black color scale indicates adjusted p -value ($-\log_{10}[q\text{-value}]$). Only results with $q < 0.05$ were displayed. Percent of module genes in Hallmark gene set is represented by point size. Q -values represent hypergeometric enrichment as reported by the EnrichR R package.

can be reliably identified independent of scoring metric, which is a key feature of high-quality gene signatures [36].

CYLD missense mutants are not associated with loss of function

Stratification of tumors by the NF-κB Activity Classifier found that only one of the two identified CYLD

missense mutations was associated with increased NF-κB activity (Figure 1, track 8 – lime green). Considering the missense mutation in the “highly active” NF-κB group had concurrent shallow deletions in both *TRAF3* and *CYLD*, we evaluated the functional consequences of the *CYLD* missense mutations. To test *CYLD* activity, we developed *CYLD* knockout cells and confirmed loss of *CYLD* expression and activation of NF-κB (Figure

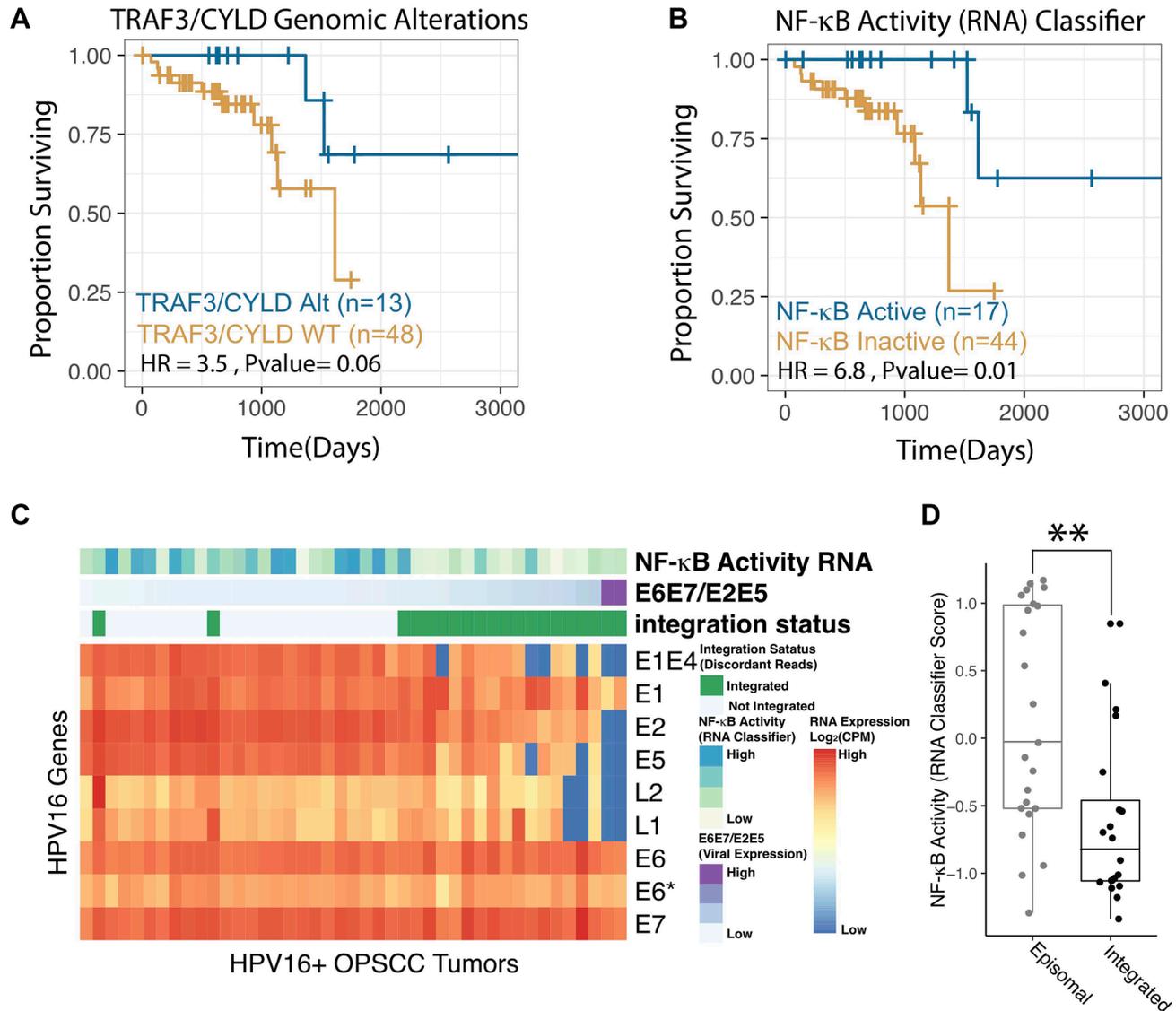


Figure 3: NF-κB activity classifier correlates with patient outcomes and viral integration status. (A–B) Kaplan-Meier Analysis of Progression Free Interval of HPV+ HNSCC. *P*-values represent log-rank test. HR – Hazard Ratio. NF-κB Active – Highly NF-κB active tumors by RNA expression as defined according to the RNA based classifier (see methods), these were compared to all other tumors (NF-κB Inactive) in the study cohort. TRAF3/CYLD Alt – Any missense, nonsense, frameshift, deep deletion in TRAF3 and/or CYLD, these were compared to all other tumors (TRAF3/CYLD WT) in the study cohort. See Supplementary Figure 1 for grossly similar recurrence-free survival results. (C) Heatmap of HPV16 Viral Gene Expression for 61 HPV16+ OPSCC tumors included in the TCGA. Columns – tumors. Rows – HPV16 viral genes. Column Annotations: NF-κB activity RNA - nearest classifier score, higher values are more proximal to the NF-κB active centroid. E6E7/E2E5 Ratio – $[E6 \text{ expression (raw counts)} + E7 \text{ expression (raw counts)}] / [E2 \text{ expression (raw counts)} + E5 \text{ expression (raw counts)}]$. The columns are organized by this metric which is reported to strongly correlated with viral genomic integration. Integration Status – HPV viral integration status as determined by the ViFi pipeline. (D) Box Plot comparing NF-κB activity in integrated and episomal tumor groups. Integration as assigned by ViFi. NF-κB activity – Raw NF-κB classifier scores as in panel C. $**p < 0.001$.

5A, 5B). Site-directed mutagenesis was used to recreate observed mutations (Figure 5C) and activity of mutant proteins to inhibit an NF- κ B reporter was compared to wild-type CYLD in CYLD knockout cells. As expected, CYLD knockout cells showed significantly elevated NF- κ B activity compared to parental cells (Figure 5D). Interestingly, both N300S and D618A mutant CYLD proteins were as efficient in inhibiting NF- κ B transcriptional activity as wild-type CYLD (Figure 5D). These data suggest that N300S and D618A CYLD missense mutations are not inactivating mutations and are not responsible for NF- κ B activation.

DISCUSSION

HNSCC is increasing global incidence due to human papillomavirus and continued consumption of carcinogens [2, 7, 10]. In contrast to HPV-negative HNSCC, HPV-mediated tumors are more susceptible to contemporary treatment paradigms, which also leads to

improved patient survival [41]. However, HPV+ HNSCC survivors are frequently burdened with significant side effects including pain; neck muscle stiffness; dry mouth; and difficulty with speech, eating/drinking, and breathing. Efforts to reduce these significant quality-of-life effects have triggered multiple trials of treatment de-escalation. In these trials, patients are selected for deintensified treatment based on patient factors like smoking status, histological characteristics following an ablative procedure, or response to induction chemotherapy [42]. Given that methods to identify patients for deintensified therapy are imperfect, our improved classifiers may serve as prognostic biomarker to help clinicians with therapeutic decisions.

Recent work examined genomic characteristics of the tumor that could be used prior to treatment to prognostically stratify patients. Somatic mutations or deletions in *TRAF3* or *CYLD* identified a subset of HPV+ HNSCC associated with improved outcome [1, 16, 17]. Increasing evidence demonstrates these somatic mutant

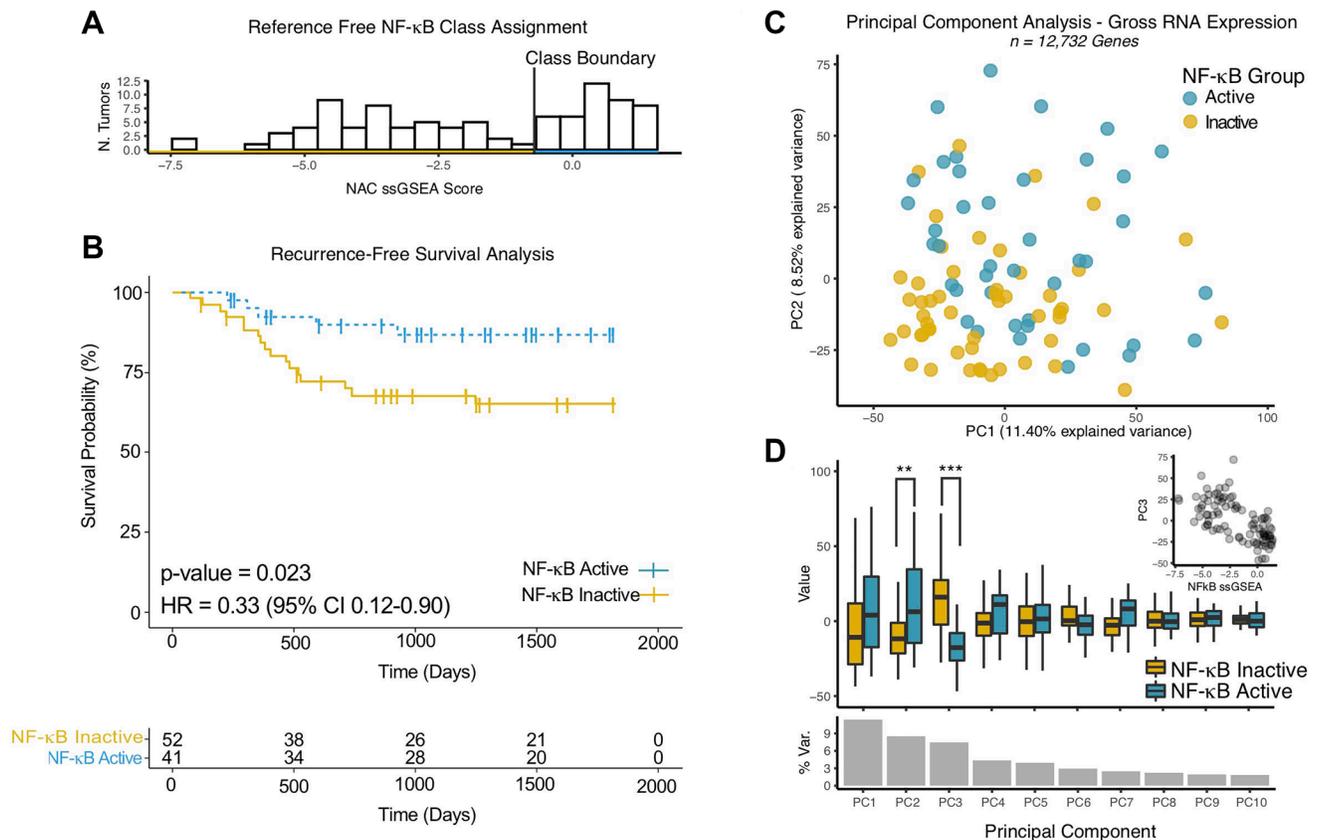


Figure 4: NF- κ B activity classifier gene expression is cohesive and correlates with patient outcomes in an independent validation cohort. (A) Histogram of single- sample (ss)GSEA Scores for NF- κ B activity classifier genes for each tumor in the validation cohort. Class Boundary – an empiric threshold based on the bimodal distribution of scores to assign (binary) NF- κ B activity status. (B) Kaplan-Meier Analysis of Recurrence Free Survival of HPV+ HNSCC. *P*-values represent log-rank test. HR – Hazard Ratio. NF- κ B Active/Inactive – NF- κ B active tumors by RNA expression as defined according to the ssGSEA scores for NF- κ B activity classifier genes determined for each tumor as in panel A. (C) Scatter plot of tumors based on gross RNA expression in principle component space, the top two principal components are displayed. Colors - NF- κ B activity groups as in panel A. (D) Box Plot of principle component values comparing NF- κ B activity groups. *P*-values represent Wilcoxon Rank-sum test. ***p*-value < 0.001, ****p*-value < 5×10^{-9} . % Var. – Percentage of total variance explained by the individual principal component. Inset – Scatter plot of NF κ B ssGSEA scores vs. PC3.

tumors identify a distinct clinical entity given notable molecular, histopathologic, and outcome differences [3, 16, 43]. Regarding function, TRAF3 is a ubiquitin ligase that regulates numerous receptor pathways, ultimately functioning to negatively regulate both canonical and non-canonical NF- κ B pathways [44]. Similarly, CYLD inhibits the NF- κ B pathway in its role as a deubiquitinase [45]. Inactivation of TRAF3 or CYLD results in activation of NF- κ B producing robust downstream effects as demonstrated by significant RNA expression changes amongst mutant TRAF3/CYLD tumors (Figure 1) [46].

NF- κ B was thought to protect cells from viruses through induction of immune response genes; however, it is now apparent that many viruses rely on or even induce aberrant NF- κ B activity to promote host cell survival and proliferation, supporting viral lifecycle and gene expression. Previous work revealed that NF- κ B overactivation favors carcinogenesis with EBV and HIV-mediated disease with a fundamental role of constitutive NF- κ B signaling in EBV tumorigenesis [22, 24–27]. When aberrantly activated, NF- κ B stabilizes the EBV episome, while suppressing the lytic cycle [22, 24, 47]. We found that in HPV+ HNSCC TCGA cohort increased NF- κ B activity significantly correlated with the absence of HPV integrations (Figure 3C and 3D). Whether constitutively active NF- κ B supports the presence of HPV episomes or inhibits HPV integrations in human genome remains to be investigated and currently is studied in our laboratory.

Current knowledge of HPV-induced carcinogenesis is largely derived from study of uterine cervical cancer with the classical model showing persistent infection followed by HPV genome integration leading to increased expression of HPV oncoproteins [48]. The absence of

HPV integration in a substantial portion of HNSCC coupled with constitutive NF- κ B activation, as we show here (Figure 3), suggests that HPV carcinogenesis in the upper aerodigestive tract may be driven by maintenance of episomal HPV. Interestingly, HPV genome integration has consistently associated with worse survival in these tumors [39, 49, 50].

Recent finding revealed that ER expression correlated with improved survival in HPV+ HNSCC [51]. Initial studies found that ER expression and signaling inhibited NF- κ B through estrogen stabilization of I κ B α [52]. Later investigations unveiled that ER signaling enhanced NF- κ B activity in macrophages and T cells, suggesting that the interaction between ER and NF- κ B may depend on cellular context [53, 54]. Given that both ER expression and loss of *TRAF3* portend improved prognosis in HPV+ HNSCC, description that ER- α stimulation depletes cells of TRAF3 via ubiquitination provides a potential mechanistic connection of these findings [55]. As far as we are aware, the crosstalk between NF- κ B and ER is not described in the presence of HPV and in HNSCC. Although our work cannot determine causality, the WGCNA analysis suggests a positive correlation between ER signaling and NF- κ B activity in HPV+ HNSCC, with the “yellow” module being enriched for both NF- κ B and early estrogen response genes. Also, the nearest neighbor (relative to “yellow”) “magenta” module was enriched for estrogen response genes (Figure 2A and 2C).

Use of multi-variable predictor models is gaining recent clinical traction since these tools provide a more comprehensive assessment of the intratumoral environment [28–30]. In our case, we hypothesized that

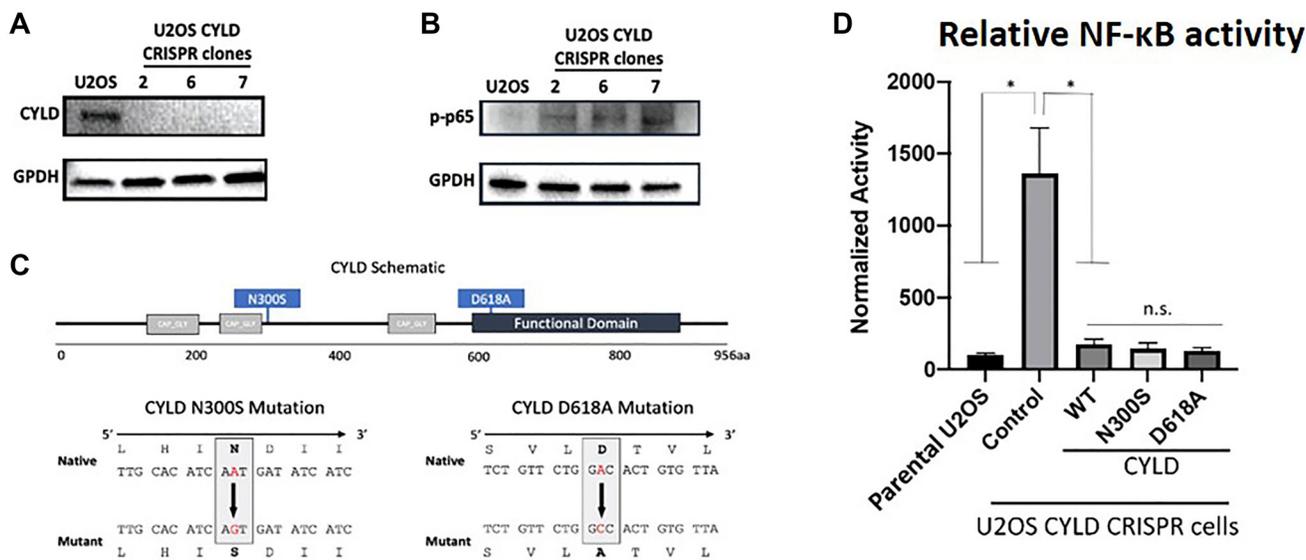


Figure 5: Expression of CYLD (A), pp65 (B) and GPDH in U2OS parental and CYLD CRISPR clones as determined by immunoblotting. (C) Schematic representations of CYLD protein and schema of CYLD N300S and D618A mutant constructions. (D) NF- κ B reporter activity in U2OS parental, U2OS CYLD CRISPR (control) cells, or U2OS CYLD CRISPR cells transiently transfected with wild-type or mutant CYLD constructs. *t*-test was used to compare U2OS to other conditions. ** -- adjusted *p*-value (Bonferroni correction) < 0.05.

undefined alterations in addition to TRAF3 or CYLD gene defects are in play to activate NF- κ B in HPV+ HNSCC. Querying only *TRAF3* or *CYLD* defects would be blind to these alternative NF- κ B activating strategies leading to imperfect tumor classification. Indeed, the NF- κ B Activity Classifier identified several NF- κ B active tumors excluded by genomic analysis of *TRAF3/CYLD* (Figure 1A). Tumors with deep deletions in either *TRAF3* or *CYLD*, or a truncating mutation proximal to the proteins' functional domain, were consistently included in the "active" NF- κ B category. Conversely, tumors with isolated shallow deletions tended to be in the NF- κ B "inactive" category. However, the NF- κ B Activity Classifier identified many samples in the NF- κ B "active" category that do not follow this clear-cut pattern, in particular identifying that simultaneous shallow deletion of *TRAF3* and *CYLD* in a tumor correlated with NF- κ B activity. The finding that all tumors with shallow co-occurring deletions in both *TRAF3* and *CYLD* were included in the NF- κ B "active" group suggests a functional interaction of *TRAF3* and *CYLD* in these tumors. We interrogated tumors without inactivating alterations in TRAF3/CYLD in the NF- κ B active group for mutations of genes known to influence the NF- κ B pathway; indeed, one tumor contained missense mutation in the *MAP3K14* (NIK), and there was a nonsense mutation in the *NF κ BIA*, as well as a nonsense mutation in *TRAF2* in two additional tumors (Supplementary Table 5). However, we were unable to detect additional mutations in well known NF- κ B regulators in the rest of tumors most likely due to the complex nature of the NF- κ B pathway. On the other hand, our direct testing revealed that missense mutations of *CYLD* found in HPV+ HNSCC do not lose ability to regulate NF- κ B (Figure 5). One tumor with the D618A *CYLD* mutation was classified as NF- κ B highly active, but this tumor also harbored simultaneous shallow *TRAF3* and *CYLD* deletions. Accuracy of the NF- κ B Activity Classifier to identify NF- κ B activity in HPV+ HNSCC was suggested through its improved correlation with patient outcome compared to segregating tumors based on *TRAF3* or *CYLD* defects. From the biological perspective, this finding also supports the notion that NF- κ B activation and related changes in gene expression may be the key factor determining the biological differences previously reported for *TRAF3/CYLD* mutant HPV+ HNSCC.

Our previous work identified the potential value of *TRAF3* and *CYLD* gene defects to predict outcomes in HPV+ HNSCC [16]. Herein, we demonstrate that an RNA-based classifier trained on tumors harboring these mutations may improve prognostic classification (Figure 3A, 3B, Figure 4B and Supplementary Figure 1). As clinical algorithms for treatment de-escalation are not presently informed by prognostic biomarkers, the possibility of an RNA-based approach for determining NF- κ B related prognostic groups is quite relevant. Furthermore, RNA-based gene expression profiling has

the potential to synthesize disparate observations related to prognosis in HPV+ OPSCC. Specifically, other groups have found that ER-alpha expression is prognostic [56] and we find that ER signaling is correlated with NF- κ B activity (Figure 2C). Similarly, we find that NF- κ B activity assessed by RNA expression is highly related to viral integration status which has also been put forward as a prognostic marker in HPV+ OPSCC [39]. Future work will be needed to optimize RNA-based biomarkers which represent the full prognostic potential of all relevant pathways, including NF- κ B signaling, ER signaling, and viral oncogene expression, but such a synthetic approach is likely possible based on the correlations between these transcriptional pathways we have identified.

Although application of gene expression sets from translational and experimental studies has only limited success to date, our analyses support the biological and clinical utility of the gene set we have developed. The NF- κ B related gene signature and classifier developed in this work demonstrate desirable properties suggesting that they are translatable across multiple cohorts and RNA quantification technologies. Using TCGA data set, we confirmed the robustness of RNA-based classifications in the presence of high levels of noise (Figure 1D). The NF- κ B RNA gene set was highly auto-correlated and distinct from other transcriptional programs in HPV+ HNSCC (Figure 1C, Figure 2). Using a second cohort, we validated the utility of our gene set outside of the original training data (Figure 4). In the validation cohort, a bimodal expression of the NF- κ B gene signature (Figure 4A) suggests that indeed two biological groups (NF- κ B high and low) are a feature of HPV+ HNSCC, and these groups also correlated with RFS in second data set (Figure 4B). Furthermore, the NF- κ B gene signature expression was not correlated to 8/10 top principal components demonstrating that the gene set does not simply report gross (transcriptome wide) changes in gene expression. Conversely, the very strong correlation to PC3 suggests that gene set remains compact when applied to new data sets and can likely be quantified by many metrics (Figure 4C, 4D).

This report validates and expands on our findings that significant expression changes related to NF- κ B activity occur in the subset of HPV+ HNSCC tumors marked by *TRAF3* or *CYLD* mutations. We are planning future studies investigating the importance of "long-tail" mutations in the NF- κ B pathway which might further illuminate the origins of NF- κ B dysregulation in HPV+ HNSCC.

A major discovery in the recent past is that HPV associated HNSCC have improved survival compared to tobacco associated tumors. This finding, coupled with advancements in tumor genomic analysis, definitively established HPV+ and HPV-negative HNSCC as distinct tumors. Similarly, we noted genomic differences amongst subclasses of HPV+ HNSCC and found that defects

in *TRAF3* and *CYLD* correlated with survival. Here we present data that these subclasses may also be identified by direct assessment of NF- κ B activity; as demonstrated by gene expression differences highlighted by the NF- κ B Activity Classifier. Since clinicians are exploring therapeutic deintensification for HPV+ HNSCC, identifying patients with good or poor prognosis using the NF- κ B Activity Classifier may be useful to guide therapeutic decisions.

MATERIALS AND METHODS

Data acquisition

Only de-identified, publicly available clinical and genomic data were utilized for this study. Per-gene quantified mRNA read count data, as well as per-gene discretized Gistic2 copy-number analysis data for the Cancer Genome Atlas [57] HNSCC, were downloaded from the Broad Firehose Portal [58]. In this work, we consider a Gistic score of -2 synonymous with deep deletion, and Gistic score of -1 synonymous with a shallow deletion. Gistic uses a dynamic segmentation algorithm to define chromosomal arm level (-1) and deeper focal deletions (-2) based on per tumor thresholds [59]. Clinical data for the TCGA HNSCC cohort were acquired from Liu et al., [37]. Variant calls were downloaded using the R TCGAbiolinks [60] package; calls performed with VarScan [61] were used for all analyses. TCGA RNA sequencing BAM files were downloaded from dbGaP, with NIH request #99293-1 for project #27853: “Prognostic signature in head and neck cancer” (PI – N.I.).

Cohort selection and inclusion criteria

RNA assigned HPV status from the Firehose clinical annotations were used to assign HPV status, only HPV positive tumors were included [62]. Tumors with TP53 mutations or deep deletions were excluded from the analysis. Anatomic subsites from the oropharynx, tonsil, and base of tongue were included, and nearby subsites of the hypopharynx and oral tongue considering HPV+ TP53 wild-type tumors were likely an oropharyngeal primary. Tumors from more distal sites (e.g., larynx, alveolar ridge, maxilla) were excluded. A total of 61 patients met these criteria.

Bioinformatics

RNA read count data was preprocessed by filtering low expression genes to obtain an approximately Gaussian distribution of Log_2CPM values. Filtered read count data were then normalized using the trimmed means of M values methods provided in the R edgeR package [63]. The Limma-voom pipeline was used for

all subsequent differential expression analysis [64]. Classifiers used the nearest centroid method, and were defined and cross validated using the R cancerclass package [65].

To construct a high-performance RNA-based classifier for NF- κ B activity in HPV+ HNSCC, we employed a centroid classifier, trained on high confidence class members. Preliminary groups of NF- κ B active and inactive tumors were assigned by mutational status. Specifically, all tumors with deep deletions (Gistic value = -2) or mutations (missense, nonsense, frame shift) in the NF- κ B regulator genes *TRAF3* and *CYLD* were considered NF- κ B active, and other tumors inactive. An initial differential expression was performed between these preliminary groups, and a classifier defined, based on the top 100 genes ranked by p -value. High confidence class members were defined as having correct initial assignment and having RNA expression values very similar to the class-defining average of expression (less than 0.25% of the inter-centroid distance). The gene set and classifications were then improved with a machine learning (filtering) procedure, in which tumors initially misclassified or were more than 0.25% away from a centroid were temporarily removed (filtered). Then the filtered data were then used for differential expression and construction of a final classifier. The top 50 genes (by p -value) were selected for this final classifier based on lack of improvement in the receiver operator characteristic with the addition of more genes. Adjusted p -values (multiple comparison correction per the LIMMA package) were calculated and reported. This final classifier had perfect performance on leave-one-out-cross validation. All tumors in the HPV+ HNSCC cohort were then classified according to this final classifier (nearest centroid method) for correlation with clinical and genomic data. Sample classifications were further tuned by setting an empiric threshold for NF- κ B activity at the distance of the frameshift or nonsense *TRAF3/CYLD* mutation farthest from the NF- κ B active centroid.

To identify potentially biologically relevant autocorrelated gene sets or gene expression modules [36], the WGCNA algorithm was applied to the above-described RNA expression data, filtered to the top $\sim 13,000$ genes to limit computational intensity. (WGCNA: an R package for weighted correlation network analysis [66]. Default parameters according to recommendations from the WGCNA package authors were used unless otherwise noted. The soft threshold network was constructed calculating a scale-free topology fit index for powers ranging from 4–20. The final scale-free network was constructed with soft power set to 6.

Raw RNAseq reads were analyzed for evidence of viral integration using the ViFi package [38]. Viral genes expression was also quantified using Salmon [67] and the HPV16 A1 genotype, RefSeq NC_001526.4.

Survival analysis

Clinical data, specifically progression-free interval (PFI), were extracted from Liu et. al. across the full cohort ($n = 61$) [37]. We note that the values for PFI from Liu et al., were very similar or identical (but included four more cases) when compared to recurrence-free survival (RFS) data available from Broad Firehose Portal [58]. Survival statistics were generated with the R survival package (v3.2-7) and visualized with the R survminer package (0.4.8). p -values represent log-rank test.

Gene set enrichment analysis

Ranked gene lists were created using the signal to noise ratio for the change in expression between two groups of interest as defined in the popular GSEA software package distributed by the Broad Institute [68, 69]. Hallmark signatures from the MiSigDB were used as gene sets of interest [70]. GSEA testing and related multiple comparison testing were performed with the R fgsea package [71]. Hypergeometric (gene ontology) enrichment analysis was performed for the derived WGCNA modules using the EnrichR package with default parameters [72]. All results were corrected for multiple comparisons by the EnrichR pipeline, and adjusted p -values were considered significant if adjusted $p < 0.05$.

Evaluating the TCGA mutational landscape

The *TRAF3/CYLD* mutational loci and type were assessed across HPV+ HNSCC tumors. *TRAF3* genetic alterations were predominantly deep deletions as well as two truncations; these alterations preclude translation of the TRAF3 ubiquitin ligase enzymatic domain resulting in this NF- κ B overactive phenotype. Similarly, *CYLD* alterations included deep deletions and truncations occurring prior to its de-ubiquitinase functional domain [1]. In both cases, protein loss of function is evident, leading to unchecked NF- κ B activation. However, two novel *CYLD* missense mutations (N300S and D618A) with unknown functional significance were discovered, demanding further functional appraisal.

Modeling the novel CYLD missense mutations

Employing the QuikChange II-E Site-Directed Mutagenesis Kit (Agilent #200523) per the manufacture's protocol, a wild-type Flag-HA-CYLD expression vector [73]. (Addgene #22544) was mutated to reflect the two novel *CYLD* missense mutations, N300S and D618A. Synthetic forward and reverse oligonucleotide primers (Sigma-Aldrich) were designed to harbor the desired point mutation with high CYLD binding affinity in the region of interest.

To create the N300S *CYLD* mutation, forward primer ACATCAGTGATATCATCCCAGCTTTAT and reverse primer GCAATAGAATTGTACTTTCAACACACG were used. To develop the D618A *CYLD* mutation, ggggtctaagtaacacagtgccagaacagaactaaaagc and gcttttagttctgtctggccactgtgttacttagaccc were used for the forward and reverse primers, respectively. Sanger sequencing performed by Eton Bioscience (San Diego, CA, USA) confirmed targeted mutation success.

Creation of CYLD knockout mammalian cells

Co-transfection of *CYLD* CRISPR/Cas KO (Santa Cruz #sc-400882-KO-2) and *CYLD* HDR (Santa Cruz #sc-400882-HDR-2) plasmids were used per manufacture's protocol to develop *CYLD* knockout U2OS cells. U2OS was chosen as the parental cell based on known wild-type *TP53* and *Rb* expression, characteristic of HPV+ HNSCC disease [74]. Cells were grown in 5% CO₂ at 37°C in DMEM (Genesee #25-501N) supplemented with 10% FBS (Genesee #25-514H) and 1% each of penicillin-streptomycin (Genesee #25-512), non-essential amino acids (Genesee #25-536), and glutamine (Genesee #25-509). KO *CYLD* cell media was further supplemented with 1 μ g/ml puromycin (InvivoGen ant-pr-1) used to select for CRISPR-Cas9 clones. Confirmation of *CYLD* knockdown was performed with Western blot and a luciferase NF- κ B functional assay.

Western blot

Cells were collected by trypsinization and lysed in radioimmunoprecipitation assay (RIPA lysis buffer (Sigma) with the addition of protease inhibitors (Roche) and phosphatase inhibitors (Sigma) for 15 minutes on ice. Lysates were then mechanically homogenized with an 18-gauge syringe and insoluble material was removed by centrifugation at 14,000 rpm for 15 minutes at 4°C. Protein concentration was determined using Qubit assay (Invitrogen). Twenty micrograms of total protein were mixed with 2X loading Laemmli buffer (Biorad) supplemented with DTT (Sigma) and incubated for 10 minutes at 95°C. Proteins were separated in 4% to 20% Tris-glycine polyacrylamide gels (Mini-PROTEAN; Bio-Rad) and electrophoretically transferred onto polyvinylidene fluoride membranes. Membranes were blocked with 3% BSA in PBS and incubated with primary antibodies against *CYLD* (Santa Cruz) and phospho-p65 (Cell Signaling) as well as control primary antibodies against GAPDH (Santa Cruz). Secondary antibodies were conjugated with horseradish peroxidase (Cell Signaling). After sequential washes in TBST buffer, a chemiluminescent HRP substrate was applied to the membrane and signals were immediately visualized using a ChemiDoc Bio-Rad imager.

***In vitro* NF-κB functional evaluation**

U2OS and U2OS CYLD KO cells were plated in a 96 well plate at 5×10^4 cells/100 μ l/well. After 24 hours, cells were co-transfected with a 3κB-conA-luciferase expression vector (a generous gift from Dr. Neil Perkins of the University of Dundee, Dundee, UK) and either a CYLD wild-type, CYLD N300S, CYLD D618A, or an empty expression vector using a lipofectamine 2000 (Thermo Fisher #11668030) system per manufacturer's protocol. Forty-eight hours following transfection, cells were lysed and luciferin was applied per manufacturer's protocol (Promega #E1501). Luciferase activity was measured using Promega GloMax Explorer.

Abbreviations

HNSCC: head and neck squamous cell carcinoma; HPV: human papilloma virus; HPV+: human papilloma virus associated; NAC: NF-κB Activity Classifier; NPC: nasopharyngeal carcinoma; ML: Machine Learning; AUC: receiver-operator characteristic curve; WGCNA: weighted gene correlation network analysis; ssGSEA: single-sample gene set enrichment analysis; ER: estrogen receptor; PC: principal component; PCA: principal component analysis; OPSCC: oropharyngeal squamous cell carcinoma.

Data availability

Raw TCGA data were obtained from NCBI dbGaP (the Database of Genotypes and Phenotypes) Authorized Access system with dbGaP permission.

Author contributions

Conceptualization, Travis P. Schrank, Wendell G. Yarbrough, and Natalia Issaeva; Data curation, Andrew Prince, Tejas Sathe and Damir T. Alzhanov; Formal analysis, Travis P. Schrank and Andrew Prince; Funding acquisition, Wendell G. Yarbrough and Natalia Issaeva; Investigation, Travis P. Schrank, Tejas Sathe, Xiaowei Wang, Xinyi Liu, and Damir T. Alzhanov; Methodology, Travis P. Schrank, Natalia Issaeva, and Damir T. Alzhanov; Resources, Wendell G Yarbrough and Natalia Issaeva.; Supervision, Natalia Issaeva; Writing – original draft, Travis P. Schrank and Andrew Prince; Writing – review & editing, Travis P. Schrank, Wendell G. Yarbrough, Barbara Burtness, Albert S Baldwin, and Natalia Issaeva. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

We thank Yale Department of Surgery, Otolaryngology and the Department of Otolaryngology/

Head and Neck Surgery and Lineberger Comprehensive Cancer Center at the University of North Carolina at Chapel Hill for support of these studies.

CONFLICTS OF INTEREST

Authors have no conflicts of interest to declare.

FUNDING

This work was supported by the NIH: NIDCD (T32 DC005360), 1U01DE029754-01 (NIDCR), and Yale Head-Neck SPORE (1P50CA236762-01A1) DRP. TPS was supported by NIH K08DE029241-01A1.

REFERENCES

1. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015; 517:576–82. <https://doi.org/10.1038/nature14129>. [PubMed]
2. Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE, Grandis JR. Head and neck squamous cell carcinoma. *Nat Rev Dis Primers*. 2020; 6:92. <https://doi.org/10.1038/s41572-020-00224-3>. [PubMed]
3. Williams EA, Montesion M, Alexander BM, Ramkissoon SH, Elvin JA, Ross JS, Williams KJ, Glomski K, Bledsoe JR, Tse JY, Mochel MC. CYLD mutation characterizes a subset of HPV-positive head and neck squamous cell carcinomas with distinctive genomics and frequent cylindroma-like histologic features. *Mod Pathol*. 2021; 34:358–70. <https://doi.org/10.1038/s41379-020-00672-y>. [PubMed]
4. Gillison ML, Akagi K, Xiao W, Jiang B, Pickard RKL, Li J, Swanson BJ, Agrawal AD, Zucker M, Stache-Crain B, Emde AK, Geiger HM, Robine N, et al. Human papillomavirus and the landscape of secondary genetic alterations in oral cancers. *Genome Res*. 2019; 29:1–17. <https://doi.org/10.1101/gr.241141.118>. [PubMed]
5. Pytynia KB, Dahlstrom KR, Sturgis EM. Epidemiology of HPV-associated oropharyngeal cancer. *Oral Oncol*. 2014; 50:380–86. <https://doi.org/10.1016/j.oraloncology.2013.12.019>. [PubMed]
6. Pan C, Issaeva N, Yarbrough WG. HPV-driven oropharyngeal cancer: current knowledge of molecular biology and mechanisms of carcinogenesis. *Cancers Head Neck*. 2018; 3:12. <https://doi.org/10.1186/s41199-018-0039-3>. [PubMed]
7. Shiboski CH, Schmidt BL, Jordan RC. Tongue and tonsil carcinoma: increasing trends in the U.S. population ages 20–44 years. *Cancer*. 2005; 103:1843–49. <https://doi.org/10.1002/cncr.20998>. [PubMed]
8. Viens LJ, Henley SJ, Watson M, Markowitz LE, Thomas CC, Thompson TD, Razzaghi H, Saraiya M. Human

- Papillomavirus-Associated Cancers - United States, 2008-2012. *MMWR Morb Mortal Wkly Rep.* 2016; 65:661–66. <https://doi.org/10.15585/mmwr.mm6526a1>. [PubMed]
9. Herrero R, Castellsagué X, Pawlita M, Lissowska J, Kee F, Balaram P, Rajkumar T, Sridhar H, Rose B, Pintos J, Fernández L, Idris A, Sánchez MJ, et al, and IARC Multicenter Oral Cancer Study Group. Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study. *J Natl Cancer Inst.* 2003; 95:1772–83. <https://doi.org/10.1093/jnci/djg107>. [PubMed]
 10. Chaturvedi AK, Engels EA, Anderson WF, Gillison ML. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol.* 2008; 26:612–19. <https://doi.org/10.1200/JCO.2007.14.1713>. [PubMed]
 11. Burtneß B, Harrington KJ, Greil R, Soulières D, Tahara M, de Castro G Jr, Psyrrri A, Basté N, Neupane P, Bratland Å, Fuereder T, Hughes BGM, Mesía R, et al, and KEYNOTE-048 Investigators. Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. *Lancet.* 2019; 394:1915–28. [https://doi.org/10.1016/S0140-6736\(19\)32591-7](https://doi.org/10.1016/S0140-6736(19)32591-7). [PubMed]
 12. Fakhry C, Zhang Q, Nguyen-Tan PF, Rosenthal D, El-Naggar A, Garden AS, Soulieres D, Trotti A, Avizonis V, Ridge JA, Harris J, Le QT, Gillison M. Human papillomavirus and overall survival after progression of oropharyngeal squamous cell carcinoma. *J Clin Oncol.* 2014; 32:3365–73. <https://doi.org/10.1200/JCO.2014.55.1937>. [PubMed]
 13. Pedros C, Altman A, Kong KF. Role of TRAFs in Signaling Pathways Controlling T Follicular Helper Cell Differentiation and T Cell-Dependent Antibody Responses. *Front Immunol.* 2018; 9:2412. <https://doi.org/10.3389/fimmu.2018.02412>. [PubMed]
 14. Fochi S, Bergamo E, Serena M, Mutascio S, Journo C, Mahieux R, Ciminale V, Bertazzoni U, Zipeto D, Romanelli MG. TRAF3 Is Required for NF-κB Pathway Activation Mediated by HTLV Tax Proteins. *Front Microbiol.* 2019; 10:1302. <https://doi.org/10.3389/fmicb.2019.01302>. [PubMed]
 15. Shi JH, Sun SC. Tumor Necrosis Factor Receptor-Associated Factor Regulation of Nuclear Factor κB and Mitogen-Activated Protein Kinase Pathways. *Front Immunol.* 2018; 9:1849. <https://doi.org/10.3389/fimmu.2018.01849>. [PubMed]
 16. Hajek M, Sewell A, Kaech S, Burtneß B, Yarbrough WG, Issaeva N. TRAF3/CYLD mutations identify a distinct subset of human papillomavirus-associated head and neck squamous cell carcinoma. *Cancer.* 2017; 123:1778–90. <https://doi.org/10.1002/ncr.30570>. [PubMed]
 17. Cui Z, Kang H, Grandis JR, Johnson DE. CYLD Alterations in the Tumorigenesis and Progression of Human Papillomavirus-Associated Head and Neck Cancers. *Mol Cancer Res.* 2021; 19:14–24. <https://doi.org/10.1158/1541-7786.MCR-20-0565>. [PubMed]
 18. Annunziata CM, Davis RE, Demchenko Y, Bellamy W, Gabrea A, Zhan F, Lenz G, Hanamura I, Wright G, Xiao W, Dave S, Hurt EM, Tan B, et al. Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. *Cancer Cell.* 2007; 12:115–30. <https://doi.org/10.1016/j.ccr.2007.07.004>. [PubMed]
 19. Keats JJ, Fonseca R, Chesi M, Schop R, Baker A, Chng WJ, Van Wier S, Tiedemann R, Shi CX, Sebag M, Braggio E, Henry T, Zhu YX, et al. Promiscuous mutations activate the noncanonical NF-kappaB pathway in multiple myeloma. *Cancer Cell.* 2007; 12:131–44. <https://doi.org/10.1016/j.ccr.2007.07.003>. [PubMed]
 20. Ahmed Z, Afridi SS, Shahid Z, Zamani Z, Rehman S, Aiman W, Khan M, Mir MA, Awan FT, Anwer F, Iftikhar R. Primary Mediastinal B-Cell Lymphoma: A 2021 Update on Genetics, Diagnosis, and Novel Therapeutics. *Clin Lymphoma Myeloma Leuk.* 2021; 21:e865–75. <https://doi.org/10.1016/j.clml.2021.06.012>. [PubMed]
 21. Mirghani H, Mortuaire G, Armas GL, Hartl D, Aupérin A, El Bedoui S, Chevalier D, Lefebvre JL. Sinonasal cancer: Analysis of oncological failures in 156 consecutive cases. *Head Neck.* 2014; 36:667–74. <https://doi.org/10.1002/hed.23356>. [PubMed]
 22. Chung GT, Lou WP, Chow C, To KF, Choy KW, Leung AW, Tong CY, Yuen JW, Ko CW, Yip TT, Busson P, Lo KW. Constitutive activation of distinct NF-κB signals in EBV-associated nasopharyngeal carcinoma. *J Pathol.* 2013; 231:311–22. <https://doi.org/10.1002/path.4239>. [PubMed]
 23. Li Y, Shi F, Hu J, Xie L, Zhao L, Tang M, Luo X, Ye M, Zheng H, Zhou M, Liu N, Bode AM, Fan J, et al. Stabilization of p18 by deubiquitylase CYLD is pivotal for cell cycle progression and viral replication. *NPJ Precis Oncol.* 2021; 5:14. <https://doi.org/10.1038/s41698-021-00153-8>. [PubMed]
 24. Santoro MG, Rossi A, Amici C. NF-kappaB and virus infection: who controls whom. *EMBO J.* 2003; 22:2552–60. <https://doi.org/10.1093/emboj/cdg267>. [PubMed]
 25. Li YY, Chung GT, Lui VW, To KF, Ma BB, Chow C, Woo JK, Yip KY, Seo J, Hui EP, Mak MK, Rusan M, Chau NG, et al. Exome and genome sequencing of nasopharynx cancer identifies NF-κB pathway activating mutations. *Nat Commun.* 2017; 8:14121. <https://doi.org/10.1038/ncomms14121>. [PubMed]
 26. Eliopoulos AG, Dawson CW, Mosialos G, Floettmann JE, Rowe M, Armitage RJ, Dawson J, Zapata JM, Kerr DJ, Wakelam MJ, Reed JC, Kieff E, Young LS. CD40-induced growth inhibition in epithelial cells is mimicked by Epstein-Barr Virus-encoded LMP1: involvement of

- TRAF3 as a common mediator. *Oncogene*. 1996; 13:2243–54. [PubMed]
27. Imbeault M, Ouellet M, Giguère K, Bertin J, Bélanger D, Martin G, Tremblay MJ. Acquisition of host-derived CD40L by HIV-1 in vivo and its functional consequences in the B-cell compartment. *J Virol*. 2011; 85:2189–200. <https://doi.org/10.1128/JVI.01993-10>. [PubMed]
 28. Miyamoto DT, Lee RJ, Kalinich M, LiCausi JA, Zheng Y, Chen T, Milner JD, Emmons E, Ho U, Broderick K, Silva E, Javadi S, Kwan TT, et al. An RNA-Based Digital Circulating Tumor Cell Signature Is Predictive of Drug Response and Early Dissemination in Prostate Cancer. *Cancer Discov*. 2018; 8:288–303. <https://doi.org/10.1158/2159-8290.CD-16-1406>. [PubMed]
 29. Pitroda SP, Pashtan IM, Logan HL, Budke B, Darga TE, Weichselbaum RR, Connell PP. DNA repair pathway gene expression score correlates with repair proficiency and tumor sensitivity to chemotherapy. *Sci Transl Med*. 2014; 6:229ra42. <https://doi.org/10.1126/scitranslmed.3008291>. [PubMed]
 30. McGrail DJ, Lin CC, Garnett J, Liu Q, Mo W, Dai H, Lu Y, Yu Q, Ju Z, Yin J, Vellano CP, Hennessy B, Mills GB, Lin SY. Improved prediction of PARP inhibitor response and identification of synergizing agents through use of a novel gene expression signature generation algorithm. *NPJ Syst Biol Appl*. 2017; 3:8. <https://doi.org/10.1038/s41540-017-0011-6>. [PubMed]
 31. Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics*. 2009; 25:2906–12. <https://doi.org/10.1093/bioinformatics/btp543>. [PubMed]
 32. Taniguchi K, Karin M. NF- κ B, inflammation, immunity and cancer: coming of age. *Nat Rev Immunol*. 2018; 18:309–24. <https://doi.org/10.1038/nri.2017.142>. [PubMed]
 33. Dolcet X, Llobet D, Pallares J, Matias-Guiu X. NF- κ B in development and progression of human cancer. *Virchows Arch*. 2005; 446:475–82. <https://doi.org/10.1007/s00428-005-1264-9>. [PubMed]
 34. Natarajan V, Komarov AP, Ippolito T, Bonneau K, Chenchik AA, Gudkov AV. Peptides genetically selected for NF- κ B activation cooperate with oncogene Ras and model carcinogenic role of inflammation. *Proc Natl Acad Sci U S A*. 2014; 111:E474–83. <https://doi.org/10.1073/pnas.1311945111>. [PubMed]
 35. Zhang J, Chen T, Yang X, Cheng H, Späth SS, Clavijo PE, Chen J, Silvin C, Issaeva N, Su X, Yarbrough WG, Annunziata CM, Chen Z, Van Waes C. Attenuated TRAF3 Fosters Activation of Alternative NF- κ B and Reduced Expression of Antiviral Interferon, TP53, and RB to Promote HPV-Positive Head and Neck Cancers. *Cancer Res*. 2018; 78:4613–26. <https://doi.org/10.1158/0008-5472.CAN-17-0642>. [PubMed]
 36. Dhawan A, Barberis A, Cheng WC, Domingo E, West C, Maughan T, Scott JG, Harris AL, Buffa FM. Guidelines for using sigQC for systematic evaluation of gene signatures. *Nat Protoc*. 2019; 14:1377–400. <https://doi.org/10.1038/s41596-019-0136-8>. [PubMed]
 37. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV, Omberg L, Wolf DM, Shriver CD, et al, and Cancer Genome Atlas Research Network. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell*. 2018; 173:400–16.e11. <https://doi.org/10.1016/j.cell.2018.02.052>. [PubMed]
 38. Nguyen ND, Deshpande V, Luebeck J, Mischel PS, Bafna V. ViFi: accurate detection of viral integration and mRNA fusion reveals indiscriminate and unregulated transcription in proximal genomic regions in cervical cancer. *Nucleic Acids Res*. 2018; 46:3309–25. <https://doi.org/10.1093/nar/gky180>. [PubMed]
 39. Koneva LA, Zhang Y, Virani S, Hall PB, McHugh JB, Chepeha DB, Wolf GT, Carey TE, Rozek LS, Sartor MA. HPV Integration in HNSCC Correlates with Survival Outcomes, Immune Response Signatures, and Candidate Drivers. *Mol Cancer Res*. 2018; 16:90–102. <https://doi.org/10.1158/1541-7786.MCR-17-0153>. [PubMed]
 40. Liu X, Liu P, Chernock RD, Kuhs KAL, Lewis JS Jr, Luo J, Gay HA, Thorstad WL, Wang X. A prognostic gene expression signature for oropharyngeal squamous cell carcinoma. *EBioMedicine*. 2020; 61:102805. <https://doi.org/10.1016/j.ebiom.2020.102805>. [PubMed]
 41. Mehta V, Yu GP, Schantz SP. Population-based analysis of oral and oropharyngeal carcinoma: changing trends of histopathologic differentiation, survival and patient demographics. *Laryngoscope*. 2010; 120:2203–12. <https://doi.org/10.1002/lary.21129>. [PubMed]
 42. Petar S, Marko S, Ivica L. De-escalation in HPV-associated oropharyngeal cancer: lessons learned from the past? A critical viewpoint and proposal for future research. *Eur Arch Otorhinolaryngol*. 2021; 278:4599–603. <https://doi.org/10.1007/s00405-021-06686-9>. [PubMed]
 43. Shinriki S, Jono H, Maeshiro M, Nakamura T, Guo J, Li JD, Ueda M, Yoshida R, Shinohara M, Nakayama H, Matsui H, Ando Y. Loss of CYLD promotes cell invasion via ALK5 stabilization in oral squamous cell carcinoma. *J Pathol*. 2018; 244:367–79. <https://doi.org/10.1002/path.5019>. [PubMed]
 44. Guven-Maiorov E, Keskin O, Gursoy A, VanWaes C, Chen Z, Tsai CJ, Nussinov R. TRAF3 signaling: Competitive binding and evolvability of adaptive viral molecular mimicry. *Biochim Biophys Acta*. 2016; 1860:2646–55. <https://doi.org/10.1016/j.bbagen.2016.05.021>. [PubMed]
 45. Mathis BJ, Lai Y, Qu C, Janicki JS, Cui T. CYLD-mediated signaling and diseases. *Curr Drug Targets*. 2015; 16:284–94. <https://doi.org/10.2174/1389450115666141024152421>. [PubMed]
 46. Chen T, Zhang J, Chen Z, Van Waes C. Genetic alterations in TRAF3 and CYLD that regulate nuclear factor κ B and interferon signaling define head and neck cancer subsets

- harboring human papillomavirus. *Cancer*. 2017; 123:1695–98. <https://doi.org/10.1002/cncr.30659>. [PubMed]
47. Young LS, Dawson CW. Epstein-Barr virus and nasopharyngeal carcinoma. *Chin J Cancer*. 2014; 33:581–90. <https://doi.org/10.5732/cjc.014.10197>. [PubMed]
 48. Hebner CM, Laimins LA. Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity. *Rev Med Virol*. 2006; 16:83–97. <https://doi.org/10.1002/rmv.488>. [PubMed]
 49. Nulton TJ, Kim NK, DiNardo LJ, Morgan IM, Windle B. Patients with integrated HPV16 in head and neck cancer show poor survival. *Oral Oncol*. 2018; 80:52–55. <https://doi.org/10.1016/j.oraloncology.2018.03.015>. [PubMed]
 50. Veitía D, Liuzzi J, Ávila M, Rodríguez I, Toro F, Correnti M. Association of viral load and physical status of HPV-16 with survival of patients with head and neck cancer. *Ecancermedicalscience*. 2020; 14:1082. <https://doi.org/10.3332/ecancer.2020.1082>. [PubMed]
 51. Kano M, Kondo S, Wakisaka N, Wakae K, Aga M, Moriyama-Kita M, Ishikawa K, Ueno T, Nakanishi Y, Hatano M, Endo K, Sugimoto H, Kitamura K, et al. Expression of estrogen receptor alpha is associated with pathogenesis and prognosis of human papillomavirus-positive oropharyngeal cancer. *Int J Cancer*. 2019; 145:1547–57. <https://doi.org/10.1002/ijc.32500>. [PubMed]
 52. Zang YC, Halder JB, Hong J, Rivera VM, Zhang JZ. Regulatory effects of estriol on T cell migration and cytokine profile: inhibition of transcription factor NF-kappa B. *J Neuroimmunol*. 2002; 124:106–14. [https://doi.org/10.1016/s0165-5728\(02\)00016-4](https://doi.org/10.1016/s0165-5728(02)00016-4). [PubMed]
 53. Calippe B, Douin-Echinard V, Laffargue M, Laurell H, Rana-Poussine V, Pipy B, Guéry JC, Bayard F, Arnal JF, Gourdy P. Chronic estradiol administration in vivo promotes the proinflammatory response of macrophages to TLR4 activation: involvement of the phosphatidylinositol 3-kinase pathway. *J Immunol*. 2008; 180:7980–88. <https://doi.org/10.4049/jimmunol.180.12.7980>. [PubMed]
 54. Hirano S, Furutama D, Hanafusa T. Physiologically high concentrations of 17beta-estradiol enhance NF-kappaB activity in human T cells. *Am J Physiol Regul Integr Comp Physiol*. 2007; 292:R1465–71. <https://doi.org/10.1152/ajpregu.00778.2006>. [PubMed]
 55. Wang C, Huang Y, Sheng J, Huang H, Zhou J. Estrogen receptor alpha inhibits RLR-mediated immune response via ubiquitinating TRAF3. *Cell Signal*. 2015; 27:1977–83. <https://doi.org/10.1016/j.cellsig.2015.07.008>. [PubMed]
 56. Koenigs MB, Lefranc-Torres A, Bonilla-Velez J, Patel KB, Hayes DN, Glomski K, Busse PM, Chan AW, Clark JR, Deschler DG, Emerick KS, Hammon RJ, Wirth LJ, et al. Association of Estrogen Receptor Alpha Expression With Survival in Oropharyngeal Cancer Following Chemoradiation Therapy. *J Natl Cancer Inst*. 2019; 111:933–42. <https://doi.org/10.1093/jnci/djy224>. [PubMed]
 57. Lerner SP, Weinstein J, Kwiatkowski D, Kim J, Robertson G, Hoadley KA, Akbani R, Creighton C, and TCGA Muscle Invasive Bladder Cancer Analysis Working Group. The Cancer Genome Atlas Project on Muscle-invasive Bladder Cancer. *Eur Urol Focus*. 2015; 1:94–95. <https://doi.org/10.1016/j.euf.2014.11.002>. [PubMed]
 58. Deng M, Brägelmann J, Kryukov I, Saraiva-Agostinho N, Perner S. Firebrowser: an R client to the Broad Institute's Firehose Pipeline. *Database (Oxford)*. 2017; 2017:baw160. <https://doi.org/10.1093/database/baw160>. [PubMed]
 59. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011; 12:R41. <https://doi.org/10.1186/gb-2011-12-4-r41>. [PubMed]
 60. Mounir M, Lucchetta M, Silva TC, Olsen C, Bontempi G, Chen X, Noushmehr H, Colaprico A, Papaleo E. New functionalities in the TCGAblinks package for the study and integration of cancer data from GDC and GTEx. *PLoS Comput Biol*. 2019; 15:e1006701. <https://doi.org/10.1371/journal.pcbi.1006701>. [PubMed]
 61. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012; 22:568–76. <https://doi.org/10.1101/gr.129684.111>. [PubMed]
 62. Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, Zhu J, Haussler D. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol*. 2020; 38:675–78. <https://doi.org/10.1038/s41587-020-0546-8>. [PubMed]
 63. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010; 26:139–40. <https://doi.org/10.1093/bioinformatics/btp616>. [PubMed]
 64. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014; 15:R29. <https://doi.org/10.1186/gb-2014-15-2-r29>. [PubMed]
 65. Jan B, Kosztyla D, von Törne C, Stenzinger A, Darb-Esfahani S, Dietel M, Denkert C. cancerclass: An R Package for Development and Validation of Diagnostic Tests from High-Dimensional Molecular Data. *J Stat Softw*. 2014; 59:1–19. <https://doi.org/10.18637/jss.v059.i01>.
 66. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008; 9:559. <https://doi.org/10.1186/1471-2105-9-559>. [PubMed]
 67. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017; 14:417–19. <https://doi.org/10.1038/nmeth.4197>. [PubMed]

68. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005; 102:15545–50. <https://doi.org/10.1073/pnas.0506580102>. [PubMed]
69. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, Houstis N, Daly MJ, Patterson N, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003; 34:267–73. <https://doi.org/10.1038/ng1180>. [PubMed]
70. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015; 1:417–25. <https://doi.org/10.1016/j.cels.2015.12.004>. [PubMed]
71. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv*. 2021. <https://doi.org/10.1101/060012>.
72. Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, Jeon M, Ma'ayan A. Gene Set Knowledge Discovery with Enrichr. *Curr Protoc*. 2021; 1:e90. <https://doi.org/10.1002/cpz1.90>. [PubMed]
73. Sowa ME, Bennett EJ, Gygi SP, Harper JW. Defining the human deubiquitinating enzyme interaction landscape. *Cell*. 2009; 138:389–403. <https://doi.org/10.1016/j.cell.2009.04.042>. [PubMed]
74. Toots M, Ustav M Jr, Männik A, Mumm K, Tamm K, Tamm T, Ustav E, Ustav M. Identification of several high-risk HPV inhibitors and drug targets with a novel high-throughput screening assay. *PLoS Pathog*. 2017; 13:e1006168. <https://doi.org/10.1371/journal.ppat.1006168>. [PubMed]