

Multimodal genomic markers predict immunotherapy response in the head and neck squamous cell carcinoma

Gaurav Kumar^{1,5}, Andrew P. South^{2,5}, Joseph M. Curry^{3,5}, Alban Linnenbach^{2,5}, Larry A Harshyne^{4,5}, Adam Ertel^{1,5}, Paolo Fortina^{1,5}, Adam Luginbuhl^{3,5}

Affiliations:

¹Department of Cancer Biology, Thomas Jefferson University, Philadelphia, PA, USA;

²Department of Dermatology, Thomas Jefferson University, Philadelphia, PA, USA; ³Department of Otolaryngology, Thomas Jefferson University, Philadelphia, PA, USA; ⁴Department of Medical Oncology, Thomas Jefferson University, Philadelphia, PA, USA; ⁵Sidney Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, USA

Running Title: Genomic predictive markers of immunotherapy in HNSCC

Correspondence: Gaurav Kumar, PhD,
Department of Cancer Biology, BLSB 1009
233 South 10th Street
Philadelphia, PA-19107, USA
Phone no.- 217-955-7531
Email: Gaurav.kumar@jefferson.edu

ABSTRACT

Immune checkpoint blockade (ICB) therapy has had a major impact on the clinical management of head and neck squamous cell carcinoma (HNSCC). However, clinical responses to ICB are observed in only a fraction of patients. In the present paper, we used multimodal approaches to a) evaluate the utility of existing ICB biomarkers including tumor mutation burden (TMB), microsatellite instability (MSI), and a T-cell specific signature in predicting therapy response in HNSCC using TCGA cohort, and; b) identify a novel molecular signature to predict ICB therapy response using an ICB clinical trial of HNSCC. Our results confirm previous reports showing TMB efficacy as a biomarker and its outcome can be influenced by age, tumor sub-site, and smoking status; and that High-TMB and high-MSI tumors are associated with T-cell signature, and better survival probability in the HNSCC. We go on to demonstrate that High-TMB and high-MSI utilizes cell-cycle/cell proliferation processes for their molecular functionality; and identify a novel Cell Proliferation ICB-therapy Predicting (CPIP) signature capable of predicting ICB therapy response in HNSCC, retrospectively, where traditional biomarkers of ICB response were insufficient. In summary, the present study defines strategies and novel signatures that can be appropriately used for patient selection for ICB therapy that can improve the clinical outcomes of HNSCC patients.

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is an aggressive, genetically complex disease characterized by malignant and uncontrolled growth of cells at various sites, including the oral cavity, larynx, oropharynx, hypopharynx, paranasal sinuses, and nasal cavity. Globally, HNSCC is the sixth most common malignancy and accounts for over 650,000 new cases annually with a mortality of approximately 50-60% per year. By far, the commonest site of cancer within this region is the oral cavity, and the commonest type of tumor is squamous cell carcinoma (SCC), which accounts for over 90% of all malignancies at this site [1]. In the United States, HNSCC accounts for 4% of all cancers, with nearly 50,000 Americans diagnosed annually and nearly 11,000 succumbing to their disease [2]. Long-term tobacco use, consumption of alcohol, and infection with high-risk types of Human Papilloma Virus (HPV), are considered the main oncogenic drivers [3].

Immunotherapy, aiming to boost natural defenses to eliminate malignant cells, is a monumental breakthrough for cancer treatment and has revolutionized the field of oncology [4]. It eradicates the tumor cells by inducing or enhancing anti-tumor immune responses in active or passive ways, including adoptive cell transfer (ACT) and immune checkpoint blockades (ICB) [4, 5]. Although immunotherapies demonstrate durable clinical responses, their efficacies vary, and only subsets of patients benefit from them [6, 7]. To augment the clinical benefits of ICB therapies to HNSCC patients, identification of molecular markers/signatures, and approaches predicting therapy response is extremely urgent. These signatures and markers will enable predictions about which patients will respond most successfully to ICB therapies. Identifying those patients who are likely to respond will not only improve their clinical outcomes but also inform about mechanisms of therapy resistance and highlight future therapeutic targets and designs.

Among the emerging biomarkers for ICB response are PD-1/PD-L1 axis, T-cell inflamed signature, tumor mutational burden (TMB), and microsatellite instability (MSI). While their role in predicting ICB response in multiple cancers such as non-small-cell lung cancer (NSCLC), colon cancer, and melanoma [14,18] is well-established, their value in HNSCC is an evolving area of research. At present, majorly, TMB and PD-1/PD-L1 axis are utilized to predict ICB therapy response in HNSCC. However, due to their inconsistency in predicting responses, more robust signatures and approaches are urgently needed [8-12]. On that front, [13, 14] demonstrated success in predicting ICB therapy response in HNSCC using a T-cell inflamed signature. However, as each HNSCC patient can have a variable baseline immune-reactivity due to their variable phenotype and associated risk factors, immune-based predictive molecular signatures might not be sufficient to predict the ICB therapy response. Hence, more robust

markers/approaches are needed. In the present study, we identified a novel molecular signature that is rather than fully relying on immune-response, is based on the cell-proliferation processes. Extracted using multimodal approaches (genomic and transcriptomic data), this signature will not only predict the immunotherapy success, but will also lead to elucidation of mechanisms of tumor-intrinsic and immune mechanism in HNSCC. Additionally, we identified/confirmed the role of TMB and MSI in predicting ICB response, along with their association with external factors, as HNSCC is a clinically, histologically, and molecularly distinct cancer. Collectively, the present study will not only congregate the conclusions of existing predictive markers of ICB response, but will also provide a novel signature that can be implemented for ICB therapy success in HNSCC.

RESULTS

Tumor mutation burden analysis: As association of TMB with HNSCC using TCGA-HNSC dataset has been studied before [15-17], we first confirmed our results on mutation profile of TCGA-HNSC, and association of TMB with survival prognosis. Here, we found our observations to be in accordance with these studies [15, 16, 18]. Next, we examined the influence of long-term tobacco use, consumption of alcohol, and infection with the HPV16 virus on TMB outcome in the TCGA-HNSC cohort (n=506). We observed that TMB outcome is significantly influenced by the smoking status (smokers: 3.60 mutations/MB vs. non-smokers: 2.14 mutations/MB; $p=0.003$) but not by HPV status (HPV-positive: 2.80 mutations/MB vs. HPV-negative: 2.43 mutations/MB; $p=0.80$) as shown in **Fig. 1A**. Since, it is also believed that age, sex, and tumor subsites can also influence the efficacy of ICB therapy response, we next evaluated the influence of these factors on TMB outcome. Here, we observed that older patients (>60 years) tend to have higher TMB rate as compared to young patients (<45 years) (old: 3.85 mutations/Mb vs. young: 1.67 mutations/Mb; $p=0.002$) (**Fig. 1A**). Although, associations of TMB with gender and race was observed to be non-significant, tumor subsite was observed to be highly correlated with the TMB outcome. Here, tumors obtained from larynx were observed to have significantly higher TMB as compared to the ones from oral cavity and/or oropharynx as shown previously by our group [19] and further expanded in **Table 1**.

It is believed that high TMB tumors contain more neoantigens, that can trigger infiltration of CD8+ T cells stimulating broad antitumor immune responses associated with better clinical outcomes. To evaluate the influence of TMB on regulating the infiltration of immune-specific cells, we observed the correlation of TMB with CD8+ T-cells, T cells, and immune cells infiltration. Here, we observed CD8+ T cells, T cells infiltration and immune cells infiltration to be positively correlated with TMB, albeit non-significant (**Fig. 1B**).

Microsatellite instability analysis: The MSI score was observed (n=506) to be a mean of 0.19 for TCGA-HNSC cohort. We used this as a cut-off to define high-MSI and low-MSI tumors (**Fig. 2A**). Here also, first we evaluated the influence of external risk factors on MSI status, and found no significant associations of HPV status, age, sex, and tumor subsite as shown in **Table 2**. Next, correlation of MSI with CD8+ T-cells, T cells, and immune cells infiltration was observed. Here, we observed a positive and in fact significant correlation of MSI status with CD8+ T cells, T cells infiltration and immune cells infiltration (**Fig. 2B**). Finally, impact of MSI on patient's survival was observed, where, we found patients with higher MSI have better (although non-significant)

survival probability as compared to the patients with lower MSI (Chi-square=3.6 on 1 degrees of freedom; log rank p-value=0.06) as shown in **Fig. 2C**.

T-cell specific molecular signature: TMB and T-cell-inflamed signature by gene expression profile (GEP) are individually believed to be associated with better prediction of ICB therapy response in multiple cancers, hence, we next tested the correlation of TMB with T-cell specific molecular signature in HNSCC. Here, RNA expression data for high and low TMB TCGA-HNSC cohort was obtained and their ranked gene list was utilized to perform a gene-set enrichment analysis (GSEA) against T-cell specific molecular signature. We observed that high TMB tumors have up-regulated T-cell specific molecular signature in the TCGA-HNSC cohort as shown in **Fig.3A**. This confirms the relationship of TMB and T-cell signature, and their utility in predicting ICB therapy response in HNSCC.

TMB and MSI related molecular signature: Next, to identify the molecular mechanism of TMB and MSI status, we analyzed the transcriptomic profile of high vs. low TMB and MSI patients. Here, differential expression analysis between high-TMB (n=134) and low-TMB (n=360) was performed, where, we observed 3326 genes to be up-regulated, and 2939 to be down-regulated in high-TMB patients. Similarly, for MSI status, we observed 7310 genes to be up-regulated (n=86), and 3920 (n=408) to be down-regulated in high-MSI patients as shown in **Fig. 3B**. To evaluate the biological processes being implicated by high TMB and MSI status, we performed GSEA analysis using hallmarks of cancer signatures, where we observed cell-cycle and progression processes including MYC, G2M and E2F to be up-regulated in both high TMB and MSI HNSC tumors as shown in **Fig 3C**.

Cell Proliferation ICB-therapy Predicting signature: It is known that due to variable levels of immune-response in individuals, T-cell specific signatures alone are insufficient to predict ICB therapy response, hence additional signatures/markers are imperative for better immunotherapy response prediction. Hence, we utilized the differentially expressed genes from high-TMB and high-MSI tumors, along with highly ranked genes from the cell-cycle/proliferation processes, to identify a signature that can be further used to predict ICB response in HNSCC. We obtained a unique list of 31 genes related to cell-cycle/progression processes that can be utilized to predict the ICB therapy response in HNSCC, we termed this list as Cell Proliferation ICB-therapy Predicting (CPIP) molecular signature. To validate the CPIP signature efficacy, we utilized RNA sequencing data of pre-therapeutic tumors from phase-II neoadjuvant trial of nivolumab (PD1-

inhibitor) in HNSCC. Originally, this cohort had lower expression of PD1/PDL-1/CD8A/IFNG genes (**Fig 4A**), and T-cell specific molecular signature in the responder group (**Fig. 4B**). To retrospectively predict the ICB therapy response in this cohort, we ranked the expressed genes from this cohort to perform GSEA against our CPIP signature. Here, we found that responder patients had high-expression of the CPIP signature prior to the therapy as compared to the non-responders. In this specific cohort, out of a total list of 31 genes, 12 genes (AKR1C3, CDX2, CRLF1, DDX4, FGF19, GRM1, NR5A1, PEG10, PRAME, SOHLH1, TRPM1, and TRPM3) were observed to be driving the ICB therapy response prediction (**Fig. 4C**). As this signature is derived from genes included in functional processes of TMB and MSI, both these endpoints were observed to be highly correlated with the CPIP signature (**Fig. 5A**). Further, as expected, all genes constituting this signature was observed to be in a strong correlation with each other (**Fig. 5B**). Next, to evaluate the mechanism of how CPIP signature genes are correlating with T-cell activation response, we performed a network analysis using normalized expression of genes of CPIP signature, and CD8a from our clinical ICB cohort. Here, we observed that CD8a is directly or indirectly related with majority of the genes of CPIP signature (**Fig. 6**). Moreover, we observed that in this cohort CD8a is negatively correlated (blue color edges) with the linked genes of CPIP signature. This demonstrates that CPIP signature can be successfully used to predict the ICB therapy response where patients have lower immune profile.

DISCUSSION

In the present paper, we followed a rationale, step-wise approach to characterize the multimodal genomic markers to predict ICB therapy response in HNSCC.

This study confirms that TMB, and MSI status have strong scientific base as a predictor of immunotherapy response in HNSCC. TMB is a strong genomic biomarker that has been tested in multiple cancers, including HNSCC to predict ICB therapy response. However, since HNSCC is associated with several risk factors, their influence on TMB is not fully studied. Here, we have observed that TMB outcome can be influenced by several external factors known to play role in HNSCC pathogenesis. One of the major factors influencing the TMB is long-term smoking. Tobacco smoke is a complex mixture of chemicals among which at least 60 are carcinogens. Many of these are thought to cause cancer by inducing DNA damage which, if misreplicated, leads to an increased burden of somatic mutations and hence an elevated chance of acquiring “driver” mutations in cancer genes [15]. Tobacco smoking has been previously associated with TMB in NSCLC [20], lung adenocarcinoma [21], along with HNSCC [15, 19]. Our results are in accordance with the previous results, emphasizing the value of utilizing smoking as a confounding factor, when TMB is used as a predictive biomarker in HNSCC. Additionally, TMB is believed to be a site-specific biomarker with important spatial and histologic differences, as we have found strong association of TMB with tumors located in larynx as opposed to the ones located in oral cavity and oropharynx. Our results are in accordance with [19, 22] in HNSCC and NSCLC. As per the demographic correlation, we observed age to have a strong association with the TMB load. Although, mutations accumulate continuously in somatic cells with increasing age, it is unclear whether the age of patients is a determining factor in deriving benefits from immunotherapy. From our results, it is likely that age of the HNSCC patients can influence the TMB outcome. Our results are in accordance with [23-25], where the authors found high correlation of age and TMB in patients of liver, bile duct, colorectal, pancreatic, melanoma, esophageal, HNSCC, and lung cancer.

Recently [18] demonstrated HPV-positive tumors with significantly lower TMB as compared to HPV-negative HNSCC. Although, we observed similar trend in our analysis, but we observed the differences to be non-significant. This may be due to utilizing different cut-off values to define high and low TMB. While, we used the mean TMB/Mb as a cut-off, [18] used only top 25% patients as high-TMB and bottom 25% as low-TMB. Discordance like this provides a strong case to define a universal cut-off value for each cancer type. It should be noted here that generally >10 mutations/Mb is defined as a high-TMB tumor but as suggested by our data and others that

each cancer has different mutation rate [15], this cut-off cannot be used for all cancer types to define a high TMB tumor.

MSI status has been mainly utilized to predict ICB therapy response in colorectal cancer. However, after its implementation of predicting PD-1 response in HNSCC [26], more large-scale studies are needed to verify its reliability and stability in clinical applications. In our study, we observed high utility of MSI as a predictor of immunotherapy response. It was observed to have better association with the T-cells activation compared to TMB. This is due to the fact that high TMB/MSI tumors harbor both a higher rate of mutant neo-antigens and a higher number of somatic mutations (including frameshift mutations) that triggers recognition of the tumor by the immune system and result in higher infiltration of T cells leading to improved response to immunotherapies [27]. Interestingly, none of the major external factors were observed to be influenced in our study. Additional molecular studies might be required to further confirm this finding. Using mean cut-off value, MSI survival statistics between high-MSI vs. low-MSI was observed to be on borderline for significance. This might be due to the use of mean as the cut-off. Previous studies [28] have used higher MSI (>4) as a cut-off and have obtained better and significant survival probabilities.

Due to the complex nature of cancer, and imperfection in measurement of the existing predicting biomarkers, new markers and approaches are needed to improve the ICB therapy response in HNSCC. On novel markers front, we identified a molecular signature, CPIP that can successfully predict the ICB therapy response, where existing predicting markers (PD1/PDL-1 axis and T-cell specific signatures) were ineffective. As shown in the results, our Nivolumab treated HNSCC cohort had lower expression of these traditional markers and no existing immune-based marker was clearly predicting the ICB response. Here, our CPIP signature successfully predicted the therapy response, retrospectively. As cell cycle/proliferation process is one of the hallmarks in cancer and is associated with somatic mutations, generation of mutant neo-antigens, and activation of immune-response, it could serve to obtain a prediction on ICB therapy response. Our CPIP signature includes 31 cell cycle/proliferation related genes and based on our network analysis utilizing normalized counts of expression, we observed that at least in our cohort CD8a is negatively correlated with the genes in CPIP signature. Thus, the present study demonstrates that this signature would be very useful to predict the response of the patients that have overall lower immune signature.

On novel approaches front, this study demonstrated that utilizing a gene signature is more robust for therapy prediction instead of using a specific gene. Typically, expression of PD1 and/or PD-L1 axis is used to predict the ICB therapy response, however, they are observed to be

imperfect. This inconsistency may be due to several reasons: 1) no specific cut-off value: they are continuous marker with a range of expression levels that complicates its evaluation, 2) technical variation: different assays can generate variable scoring criteria and different thresholds; 3) intratumoral variation: they are a heterogeneous marker, with intratumoral and temporal variation of expression [29]. In the present study, we rather used a molecular signature that is more robust to predict the therapy response. It has already been demonstrated by [13, 14], that a molecular signature is better than a single gene/marker to predict the immunotherapy response.

We note here that further validation of CPIP signature using other published ICB trials, where the immune-signature is not sufficient to predict the therapy response would be ideal. However, access to those studies might be challenging. In future, we plan to expand our cohort and implement machine-learning algorithms to utilize this signature in individual patient settings for ICB therapy prediction in clinical settings.

CONCLUSION:

Our data implicate that TMB, MSI, T-cell specific signatures are strong markers that can be implemented in the HNSCC as a predictive marker for immunotherapy. Our novel CPIP signature can be utilized as a predictive molecular signature for ICB therapy in HNSCC, where immune-signature cannot predict the therapy response.

METHODS

TCGA-Data source: The whole-exome sequencing (WES) data in level 3-curated Mutation Annotation Format (MAF) file, RNA sequencing data as raw count matrix, and clinical data for 506 patients of HNSC from The Cancer Genome Atlas (TCGA) was obtained using 'TCGAbiolinks' package in R/Bioconductor [30, 31]. The MSISensor data for the TCGA-HNSC was obtained from [28].

Tumor mutation burden:

The TMB load for each cancer type was analyzed by calculating the total number of somatic nonsynonymous mutations per coding area of a tumor exome, described by [32]. The mean TMB was used as the cut-off to determine high vs. low TMB. Mutational analysis was observed using maftools package in R/Bioconductor [33].

Immune landscape analysis:

The immune landscape data including CD8+ T-cells, TIS score, and IIS score was obtained from [6]. Here, TIS was defined as the mean of the standardized values for CD8+ T, T central and effector memory, Th1, Th2, Th17, and Treg cells. IIS was defined as the mean of the standardized values for macrophages, DC subsets, B cells, cytotoxic cells, eosinophils, mast cells, neutrophils, NK cell subsets, and all T cell subsets excluding T gamma delta and T follicular helper cells. T-cell-inflamed gene list was obtained from [14], but was further modified to include additional T-cell-related genes from [6].

Correlation and survival analysis:

Correlations of HPV status, smoking status, age, gender, race, tumor location, CD8+ T cells, TIS, and IIS was observed using 'ggpubr' package in R/Bioconductor. Pearson method was utilized for the correlations.

RNA-sequencing data analysis:

The RNA-seq data of pre-therapeutic tumors from phase-II neoadjuvant trial of nivolumab (PD1-inhibitor) was obtained as raw count matrix from Squamous cell carcinoma (SCC) group in Sidney Kimmel Cancer Center, Thomas Jefferson University, PA, USA. Response was characterized as pathologic treatment effect where by each tumor was graded by two pathologists to the degree of treatment effect within the primary tumor [34]. Overall, we had sequencing data from n=13 patients, where, n=5 were responders, and n=8 were non-responders. Before differential

expression, batch effects or sample heterogeneity was tested using iSeqQC [35]. Differential gene expression was performed between responder and non-responder using the DESeq2 package [36] in R/Bioconductor. Genes were considered differentially expressed (DE) if they had adjusted p value ≤ 0.05 and absolute fold change ≥ 2 . All plots were constructed using R/Bioconductor.

Functional annotations:

The DESeq2 test statistic was used as a ranking metric to perform Gene Set Enrichment Analysis (GSEA) in pre-ranked mode, with genes having zero base mean or “NA” test statistic values filtered out to avoid providing numerous duplicate values. GSEA was performed against custom gene list (T-cell specific genes and CPIP signature), and hallmarks of cancer from MSigDB collections [37]. DAVID [38] was performed on list of DEGs between high vs low TMB and high vs low MSI. Network analysis was performed using igraph [39], where normalized expression counts from Nivolumab HNSCC cohort was used as input. Here, Pearson correlation was used to perform the correlation distances.

REFERENCES:

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA Cancer J Clin* 2018, **68**(6):394-424.
2. Hutchinson MND, Mierzwa M, D'Silva NJ: **Radiation resistance in head and neck squamous cell carcinoma: dire need for an appropriate sensitizer.** *Oncogene* 2020, **39**(18):3638-3649.
3. Alsahafi E, Begg K, Amelio I, Raulf N, Lucarelli P, Sauter T, Tavassoli M: **Clinical update on head and neck cancer: molecular biology and ongoing challenges.** *Cell Death Dis* 2019, **10**(8):540.
4. Zhang Y, Zhang Z: **The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications.** *Cell Mol Immunol* 2020, **17**(8):807-821.
5. Jiang X, Li L, Li Y, Li Q: **Molecular Mechanisms and Countermeasures of Immunotherapy Resistance in Malignant Tumor.** *J Cancer* 2019, **10**(7):1764-1771.
6. Mandal R, Senbabaoglu Y, Desrichard A, Havel JJ, Dalin MG, Riaz N, Lee KW, Ganly I, Hakimi AA, Chan TA *et al*: **The head and neck cancer immune landscape and its immunotherapeutic implications.** *JCI Insight* 2016, **1**(17):e89829.
7. Signorelli D, Giannatempo P, Grazia G, Aiello MM, Bertolini F, Mirabile A, Buti S, Vasile E, Scotti V, Pisapia P *et al*: **Patients Selection for Immunotherapy in Solid Tumors: Overcome the Naive Vision of a Single Biomarker.** *Biomed Res Int* 2019, **2019**:9056417.
8. Maleki Vareki S, Garrigos C, Duran I: **Biomarkers of response to PD-1/PD-L1 inhibition.** *Crit Rev Oncol Hematol* 2017, **116**:116-124.
9. Ren D, Hua Y, Yu B, Ye X, He Z, Li C, Wang J, Mo Y, Wei X, Chen Y *et al*: **Predictive biomarkers and mechanisms underlying resistance to PD1/PD-L1 blockade cancer immunotherapy.** *Mol Cancer* 2020, **19**(1):19.
10. Spencer KR, Wang J, Silk AW, Ganesan S, Kaufman HL, Mehnert JM: **Biomarkers for Immunotherapy: Current Developments and Challenges.** *Am Soc Clin Oncol Educ Book* 2016, **35**:e493-503.
11. Yarchoan M, Albacker LA, Hopkins AC, Montesion M, Murugesan K, Vithayathil TT, Zaidi N, Azad NS, Laheru DA, Frampton GM *et al*: **PD-L1 expression and tumor mutational burden are independent biomarkers in most cancers.** *JCI Insight* 2019, **4**(6).
12. Yi M, Jiao D, Xu H, Liu Q, Zhao W, Han X, Wu K: **Biomarkers for predicting efficacy of PD-1/PD-L1 inhibitors.** *Mol Cancer* 2018, **17**(1):129.
13. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, Albright A, Cheng JD, Kang SP, Shankaran V *et al*: **IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade.** *J Clin Invest* 2017, **127**(8):2930-2940.
14. Ott PA, Bang YJ, Piha-Paul SA, Razak ARA, Bennouna J, Soria JC, Rugo HS, Cohen RB, O'Neil BH, Mehnert JM *et al*: **T-Cell-Inflamed Gene-Expression Profile, Programmed Death Ligand 1 Expression, and Tumor Mutational Burden Predict Efficacy in Patients Treated With Pembrolizumab Across 20 Cancers: KEYNOTE-028.** *J Clin Oncol* 2019, **37**(4):318-327.

15. Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, Totoki Y, Fujimoto A, Nakagawa H, Shibata T *et al*: **Mutational signatures associated with tobacco smoking in human cancer**. *Science* 2016, **354**(6312):618-622.
16. Wang X, Li M: **Correlate tumor mutation burden with immune signatures in human cancers**. *BMC Immunol* 2019, **20**(1):4.
17. Wu HX, Wang ZX, Zhao Q, Chen DL, He MM, Yang LP, Wang YN, Jin Y, Ren C, Luo HY *et al*: **Tumor mutational and indel burden: a systematic pan-cancer evaluation as prognostic biomarkers**. *Ann Transl Med* 2019, **7**(22):640.
18. Zhang L, Li B, Peng Y, Wu F, Li Q, Lin Z, Xie S, Xiao L, Lin X, Ou Z *et al*: **The prognostic value of TMB and the relationship between TMB and immune infiltration in head and neck squamous cell carcinoma: A gene expression-based study**. *Oral Oncol* 2020, **110**:104943.
19. South AP, den Breems NY, Richa T, Nwagu U, Zhan T, Poojan S, Martinez-Outschoorn U, Johnson JM, Luginbuhl AJ, Curry JM: **Mutation signature analysis identifies increased mutation caused by tobacco smoke associated DNA adducts in larynx squamous cell carcinoma compared with oral cavity and oropharynx**. *Sci Rep* 2019, **9**(1):19256.
20. Chae YK, Davis AA, Raparia K, Agte S, Pan A, Mohindra N, Villafior V, Giles F: **Association of Tumor Mutational Burden With DNA Repair Mutations and Response to Anti-PD-1/PD-L1 Therapy in Non-Small-Cell Lung Cancer**. *Clin Lung Cancer* 2019, **20**(2):88-96 e86.
21. Nagahashi M, Sato S, Yuza K, Shimada Y, Ichikawa H, Watanabe S, Takada K, Okamoto T, Okuda S, Lyle S *et al*: **Common driver mutations and smoking history affect tumor mutation burden in lung adenocarcinoma**. *J Surg Res* 2018, **230**:181-185.
22. Stein MK, Pandey M, Xiu J, Tae H, Swensen J, Mittal S, Brenner AJ, Korn WM, Heimberger AB, Martin MG: **Tumor Mutational Burden Is Site Specific in Non-Small-Cell Lung Cancer and Is Highest in Lung Adenocarcinoma Brain Metastases**. *JCO Precision Oncology* 2019(3):1-13.
23. Yu Z, Wang J, Feng L, Yang X, Qi Q, Li W, Zhang X, Ge M, Qin H: **Association of tumor mutational burden with age in solid tumors**. *Journal of Clinical Oncology* 2020, **38**(15_suppl):e13590-e13590.
24. Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Abascal F, Hall MWJ, Cagan A, Murai K, Mahbubani K, Stratton MR *et al*: **Somatic mutant clones colonize the human esophagus with age**. *Science* 2018, **362**(6417):911-917.
25. Zhang Y, Lin A, Li Y, Ding W, Meng H, Luo P, Zhang J: **Age and Mutations as Predictors of the Response to Immunotherapy in Head and Neck Squamous Cell Cancer**. *Front Cell Dev Biol* 2020, **8**:608969.
26. Tardy MP, Di Mauro I, Ebran N, Refae S, Bozec A, Benezery K, Peyrade F, Guigay J, Sudaka-Bahadoran A, Badoual C *et al*: **Microsatellite instability associated with durable complete response to PD-L1 inhibitor in head and neck squamous cell carcinoma**. *Oral Oncol* 2018, **80**:104-107.
27. Anagnostou V, Niknafs N, Marrone K, Bruhm DC, White JR, Naidoo J, Hummelink K, Monkhurst K, Lalezari F, Lanis M *et al*: **Multimodal genomic features predict outcome of immune checkpoint blockade in non-small-cell lung cancer**. *Nature Cancer* 2020, **1**(1):99-111.

28. Ding L, Bailey MH, Porta-Pardo E, Thorsson V, Colaprico A, Bertrand D, Gibbs DL, Weerasinghe A, Huang KL, Tokheim C *et al*: **Perspective on Oncogenic Processes at the End of the Beginning of Cancer Genomics**. *Cell* 2018, **173**(2):305-320 e310.
29. Melendez B, Van Campenhout C, Rorive S, Remmelink M, Salmon I, D'Haene N: **Methods of measurement for tumor mutational burden in tumor tissue**. *Transl Lung Cancer Res* 2018, **7**(6):661-667.
30. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I *et al*: **TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data**. *Nucleic Acids Res* 2016, **44**(8):e71.
31. Mounir M, Lucchetta M, Silva TC, Olsen C, Bontempi G, Chen X, Noushmehr H, Colaprico A, Papaleo E: **New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx**. *PLoS Comput Biol* 2019, **15**(3):e1006701.
32. Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, Ni A, Novik JB, Mangarin LMB, Abu-Akeel M *et al*: **Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer**. *Cancer Cell* 2018, **33**(5):843-852 e844.
33. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP: **Maftools: efficient and comprehensive analysis of somatic variants in cancer**. *Genome Res* 2018, **28**(11):1747-1756.
34. Merlino DJ, Johnson JM, Tuluc M, Gargano S, Stapp R, Harshyne L, Jr., Leiby BE, Flanders A, Zinner R, Axelrod R *et al*: **Discordant Responses Between Primary Head and Neck Tumors and Nodal Metastases Treated With Neoadjuvant Nivolumab: Correlation of Radiographic and Pathologic Treatment Effect**. *Front Oncol* 2020, **10**:566315.
35. Kumar G, Ertel A, Feldman G, Kupper J, Fortina P: **iSeqQC: a tool for expression-based quality control in RNA sequencing**. *BMC Bioinformatics* 2020, **21**(1):56.
36. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. *Genome Biol* 2014, **15**(12):550.
37. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles**. *Proc Natl Acad Sci U S A* 2005, **102**(43):15545-15550.
38. Huang da W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources**. *Nat Protoc* 2009, **4**(1):44-57.
39. Csardi G NT: **The igraph software package for complex network research**. *InterJournal, Complex Systems* 2006:1695.

Table 1: TMB association with external factors

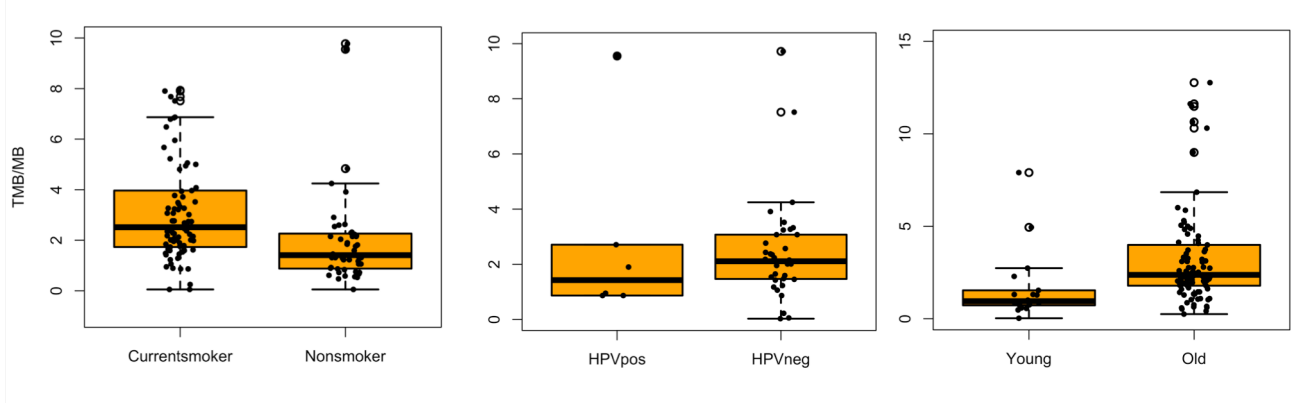
		TMB (mutations/Mb)	<i>P-value</i>
Sex	Male	3.32	
	Female	3.11	<i>0.60</i>
Race	White	3.35	
	Black or African American (AA)	2.74	<i>0.10</i>
Site	Larynx	4.97	<i>larynx vs OC: 0.00001</i>
	Oral Cavity (OC)	2.48	<i>larynx vs Oropharynx: 0.0002</i>
	Oropharynx	2.43	<i>OC vs Oropharynx: 0.91</i>

Table 2: MSI association with external factors

		MSI score	<i>P-value</i>			MSI score	<i>P-value</i>
Age	Young	0.14		Sex	Male	0.17	
	Old	0.20	<i>0.55</i>		Female	0.13	<i>0.64</i>
Race	White	0.16		Site	Larynx	0.14	<i>larynx vs OC: 0.94</i>
	Black or African America (AA)	0.11	<i>0.24</i>		Oral Cavity (OC)	0.13	<i>larynx vs Oropharynx: 0.35</i>
Smoking	Life-long non-smokers	0.25			Oropharynx	0.27	<i>OC vs Oropharynx: 0.35</i>
	Current smokers	0.18	<i>0.64</i>	HPV status	Positive	0.82	
					Negative	0.06	<i>0.37</i>

FIGURE 1

A



B

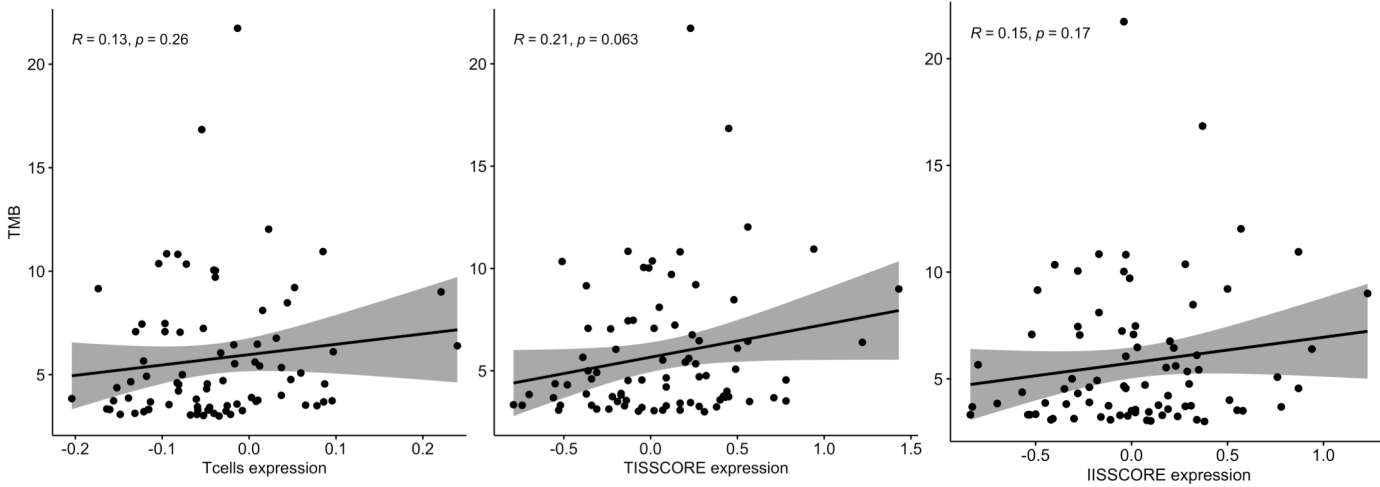


Figure 1: TMB profile in TCGA-HNSC: A) Tobacco smoking status showing significant difference in TMB score between current smokers vs. non-smokers; HPV status showing non-significant difference in TMB status between HPV-positive vs. HPV-negative patients, and; age showing significant difference in TMB score between young (<45 years) vs. old (>60 years); B) Immune-infiltration score (T cells, T-cell infiltration score (TIS), and Immune-cell infiltration score (IIS)) related to TMB is shown to be positively correlated (although non-significant) with higher TMB.

FIGURE 2

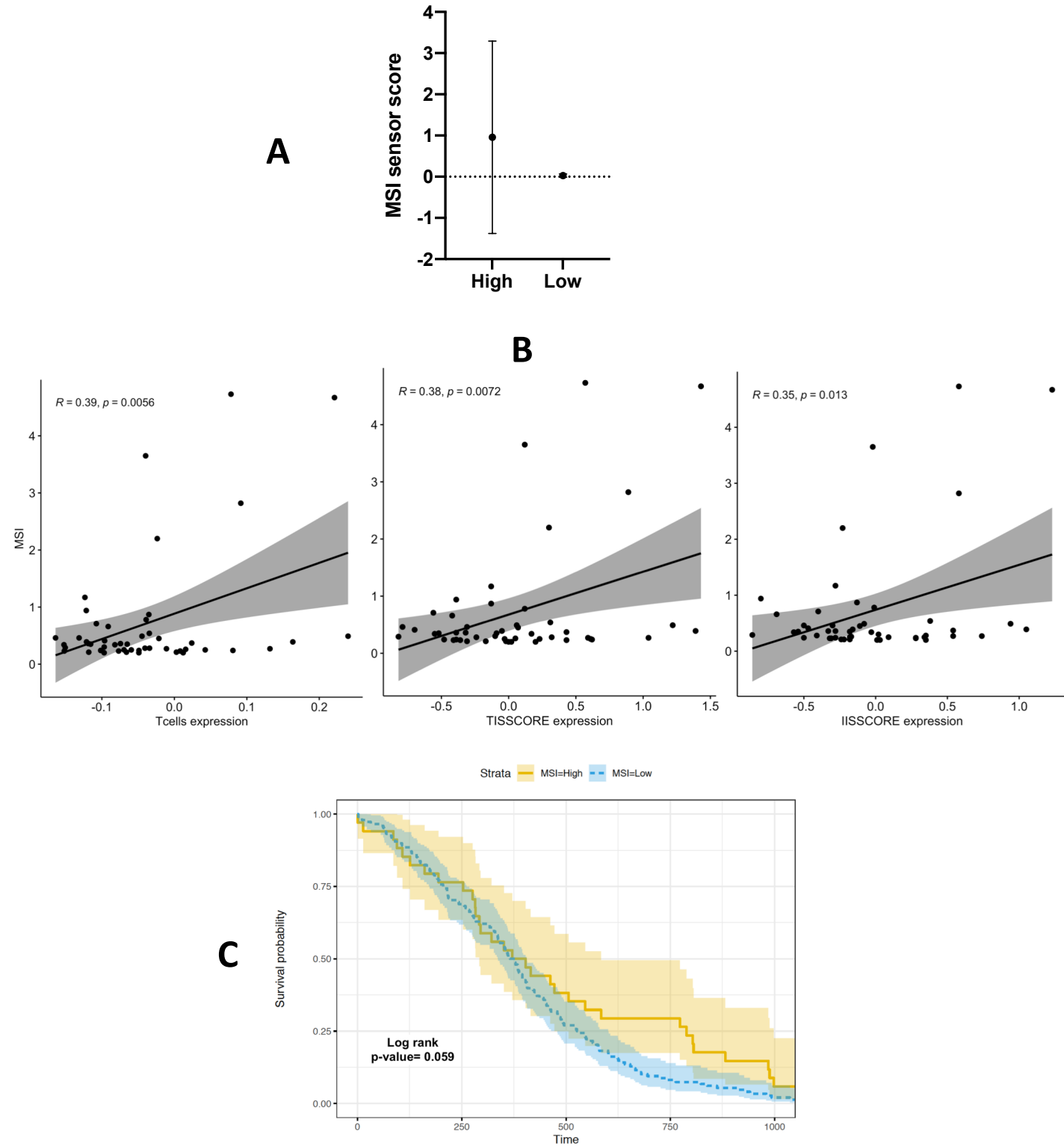
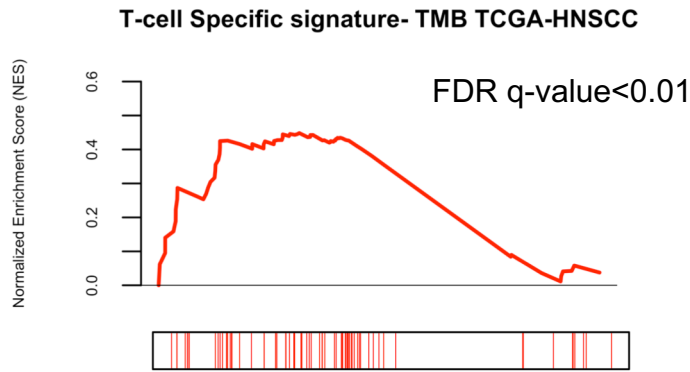


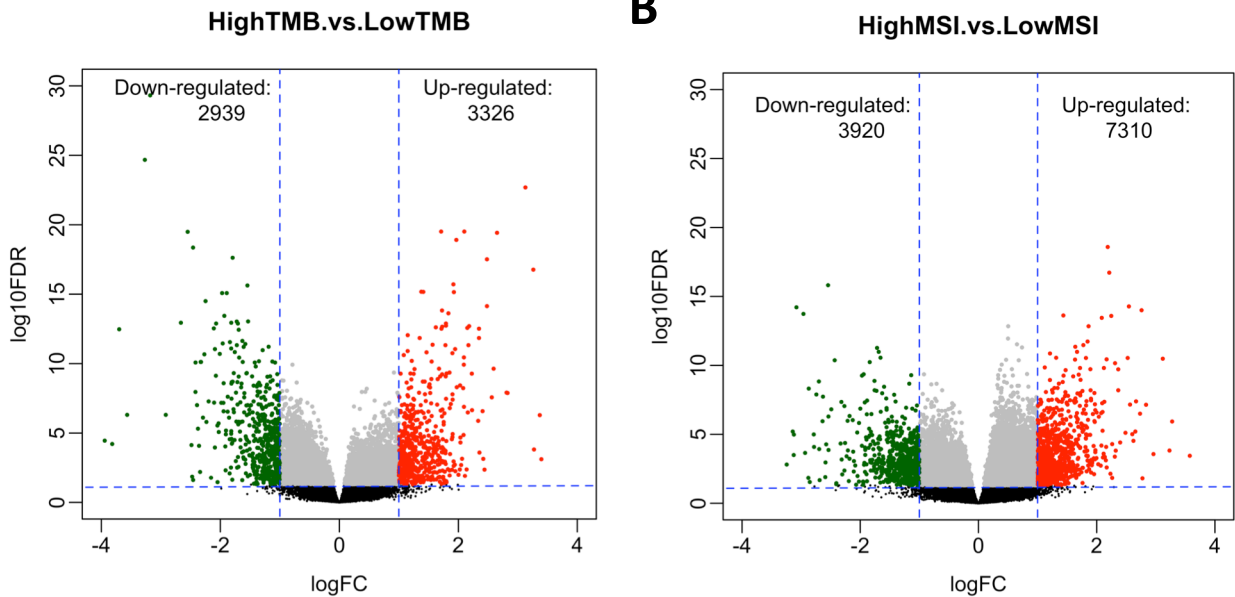
Figure 2: MSI profile in TCGA-HNSC: A) MSI sensor score between high vs low MSI patients; B) Immune-infiltration score (T cells, TIS, and IIS) related to MSI is shown to be positively correlated with higher MSI; C) Kaplan-Meier curve shows better (although non-significant) survival probability of high-MSI HNSCC patients (High-MSI HNSCC patients are represented by the yellow line, and low-MSI patients by the blue line).

FIGURE 3

A



B



C

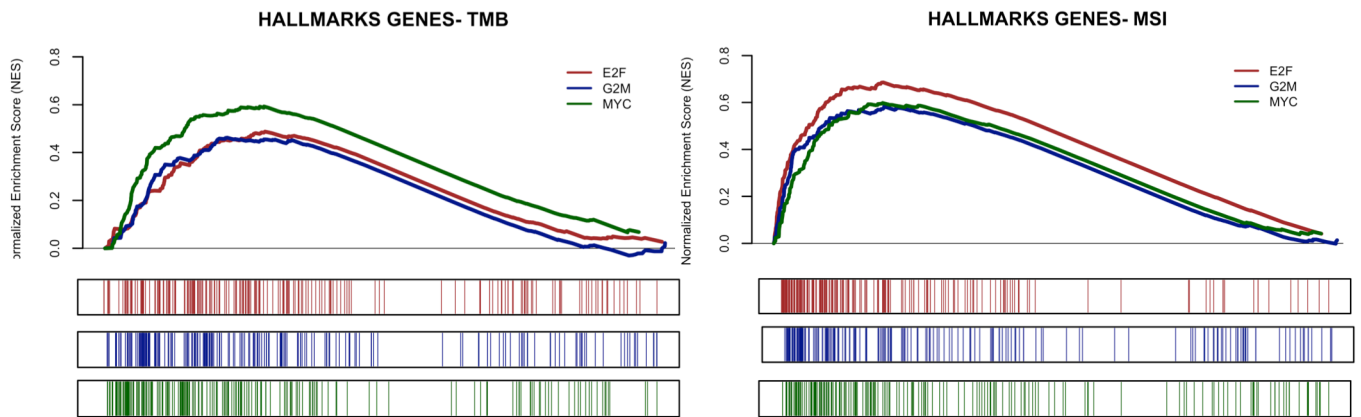


Figure 3: Transcriptomic profile in TCGA-HNSC: A) Gene set enrichment of T-cell specific molecular signature demonstrating high-TMB TCGA-HNSC tumors have high T-cell specific signature; B) Number of differentially expressed genes between high-TMB vs. low-TMB, and high-MSI vs. low-MSI; C) Gene set enrichment of hallmarks of cancer showing up-regulation of cell-cycle and progression processes in high-TMB/MSI tumors.

FIGURE 4

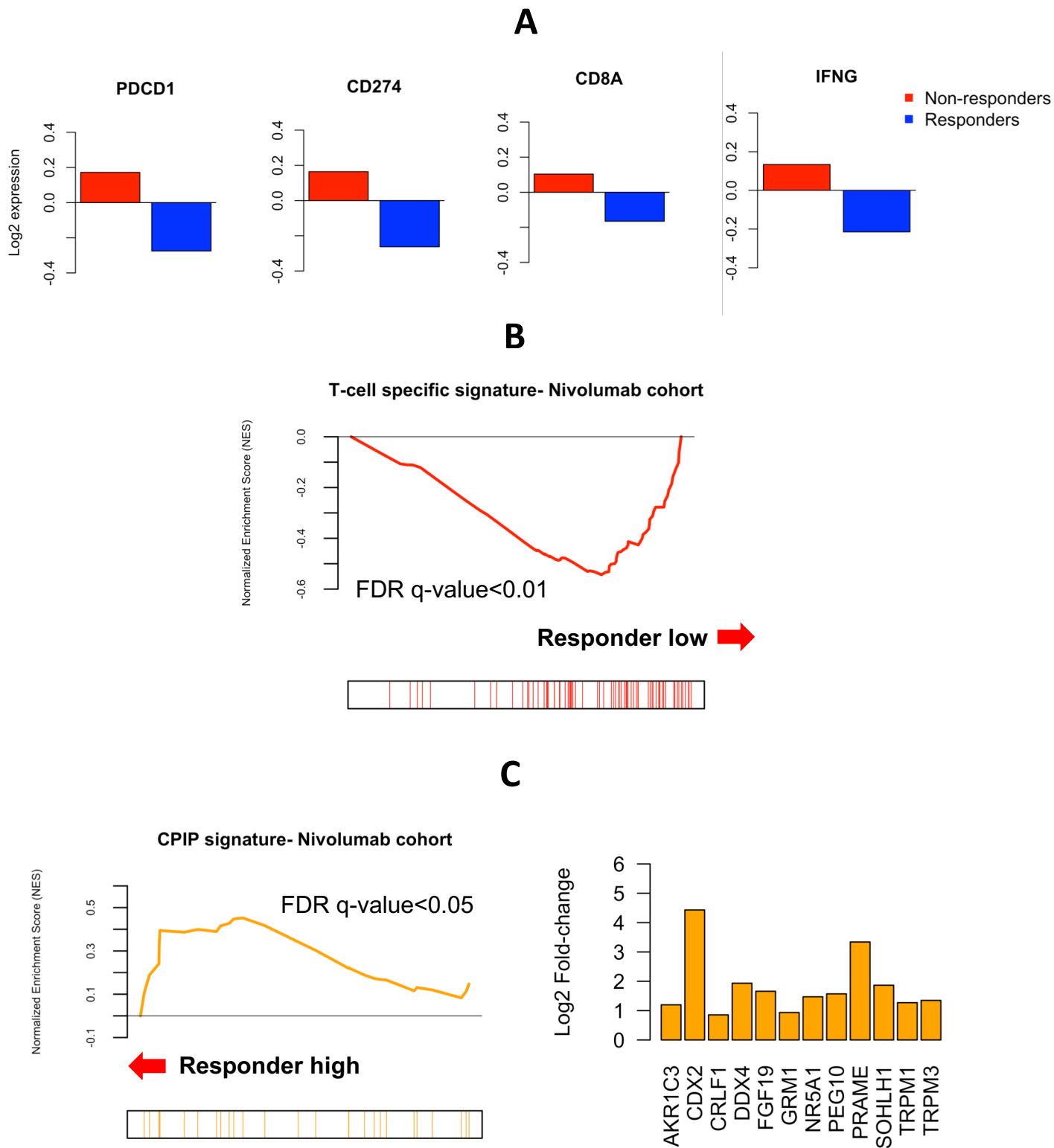
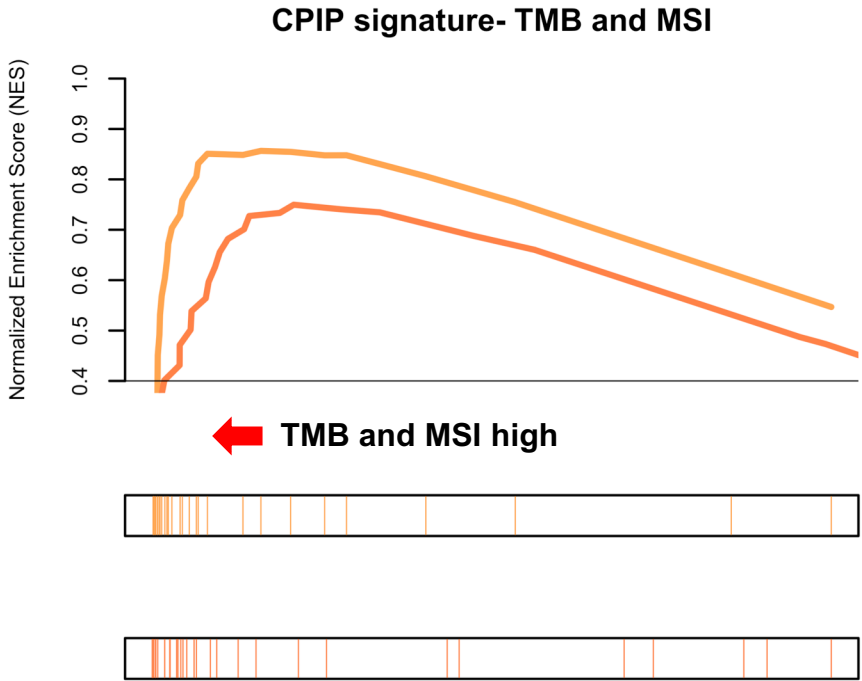


Figure 4: CPIP signature in Nivolumab cohort: A) Reduced log2 expression of established markers of ICB therapy response in the responders from Nivolumab treated HNSCC cohort; B) Gene set enrichment of T-cell specific molecular signature demonstrating responders have low T-cell specific signature; C) Gene set enrichment of CPIP molecular signature demonstrating responders have high CPIP molecular signature, along with log2 fold change of the genes in the leading edge of the analysis.

FIGURE 5

A



B

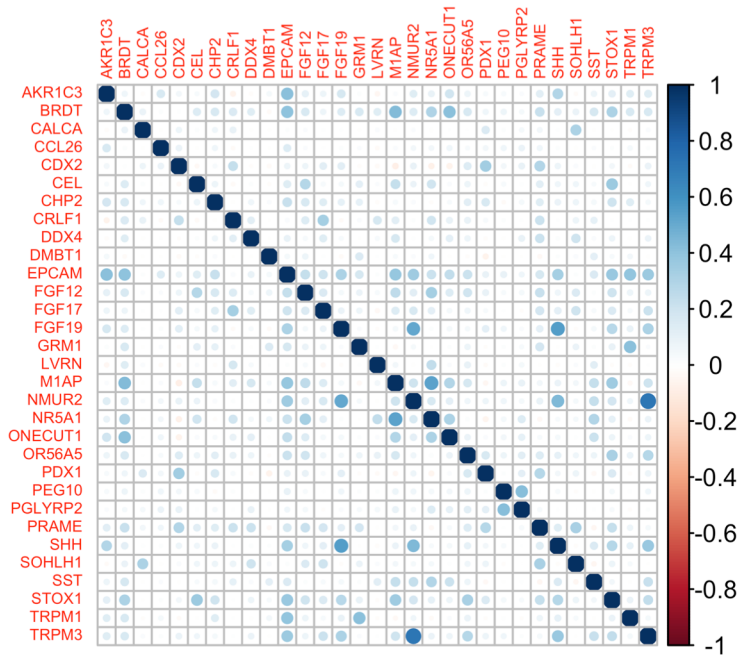


Figure 5: CPIP signature correlations: A) Gene set enrichment of CPIP molecular signature highly correlated with high TMB and MSI tumors in TCGA-HNSC, B) Correlation of all 31 genes of CPIP showing strong correlations between each other.

FIGURE 6

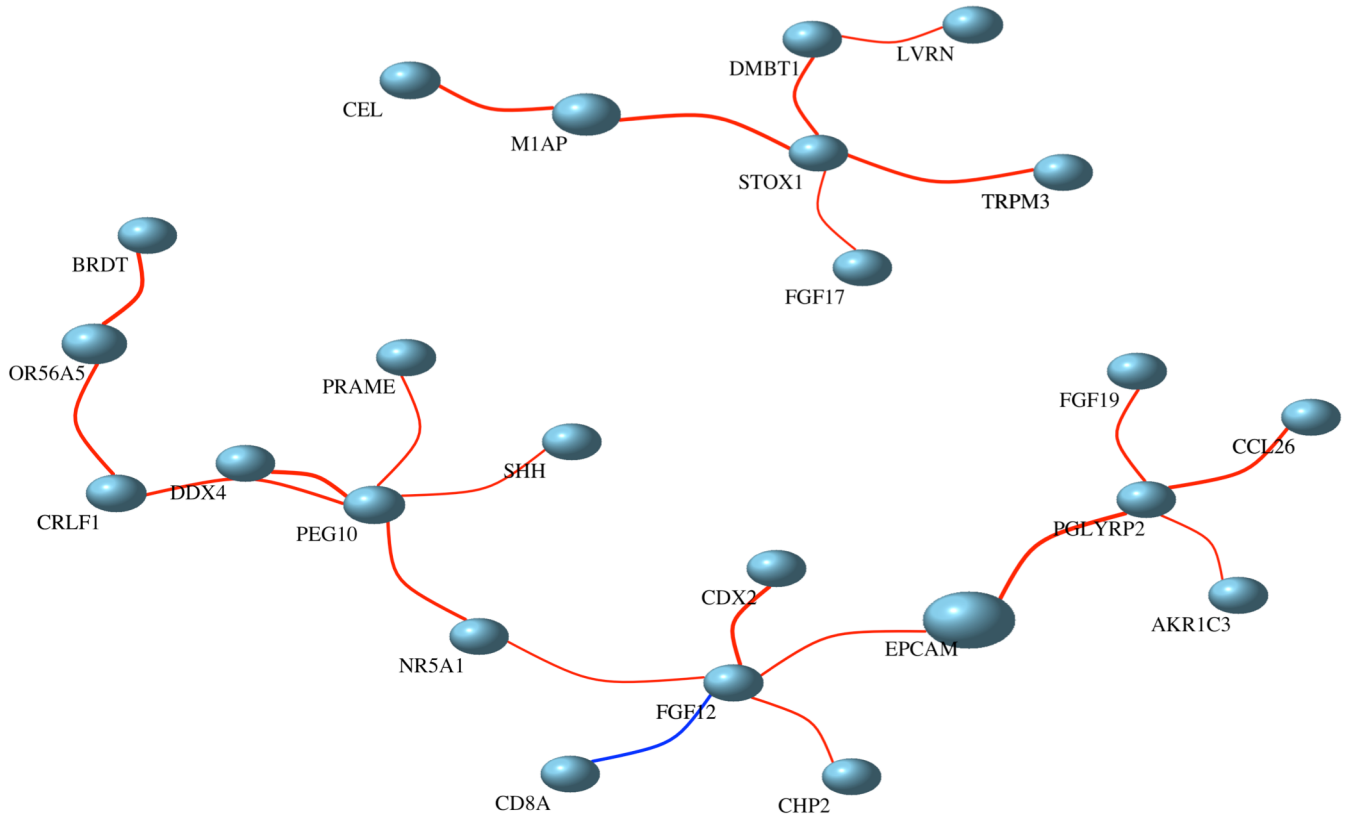


Figure 6: Network analysis of CPIP signature: Network analysis showing interactions of CD8a with the genes in CPIP signature. Here, size refers to their importance in the network due to connection with other genes. Different colors of clusters mean closeness of the related genes. Red color of the edges mean positive correlation and blue color shows negative correlations.