Comprehensive characterization of the tumor microenvironment for assessing immunotherapy outcome in patients with head and neck squamous cell carcinoma

Jian Zhang[1],*, Xi Zhong[2],*, Huali Jiang[3],*, Hualong Jiang[4],*, Tao Xie[1], Yunhong Tian[1], Rong Li[1], Baiyao Wang[1], Jiexia Zhang[5], Yawei Yuan[1]

[1]Department of Radiation Oncology, Affiliated Cancer Hospital & Institute of Guangzhou Medical University; State Key Laboratory of Respiratory Diseases, Guangzhou Institute of Respiratory Disease, Guangzhou, 510095, P. R. China
[2]Department of Radiation, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, 510095, P. R. China
[3]Department of Cardiovascularology, Tungwah Hospital of Sun Yat-sen University, Dongguan, 523000, P. R. China
[4]Department of Urology, Tungwah Hospital of Sun Yat-sen University, Dongguan, 523000, P. R. China
[5]State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, 510120, P. R. China
*Equal contribution

Correspondence to: Yawei Yuan, Jiexia Zhang; **email:** yuanyawei@gzhmu.edu.cn; zhangjx_gmu@126.com, https://orcid.org/0000-0002-2254-862X

ABSTRACT

The tumor microenvironment (TME) constitutes a complex milieu of cells and cytokines that maintain equilibrium between tumor progression and prognosis. However, comprehensive analysis of the TME and its clinical significance in head and neck squamous cell carcinoma (HNSCC) remains to be unreported. In this study, based on large-scale RNA sequencing data pertaining to single nucleotide variants (SNVs) and copy number variations (CNVs) in HNSCC patients from The Cancer Genome Atlas database, we analysed subpopulations of infiltrating immune cells and evaluated the role of TME infiltration pattern (TME score) in assessing immunotherapy outcome. TME signature genes involved in several inflammation and immunity signalling pathways were observed in the TME score subtype, which were considered immunosuppressive and potentially responsible for significantly worse prognosis. In comparison with SNV- and CNV-mediated tumor mutation burden, TME score can significantly differentiate between high- and low-risk HNSCC and predict immunotherapy outcome. Our data provide clarity on the comprehensive landscape of interactions between clinical characteristics of HNSCC and tumor-infiltrating immune cells. TME score seems to be a useful biomarker that can predict immunotherapy outcome in HNSCC patients.

INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC), one of most common cancers in the world, is an aggressive and frequently lethal head and neck malignancy [1]. Despite advances in diagnostic and therapeutic approaches, the 5-year survival rate of patients with HNSCC is only approximately 50%–60% [2]. HNSCC

is associated with distinct clinical and biological heterogeneity; patients with aggressive disease are managed using cetuximab, an anti-EGFR antibody, but only around 13% metastatic patients respond to such treatment [3, 4]. Phase III clinical trials of programmed death (PD)-1 immune checkpoint inhibitors (ICIs) (pembrolizumab and nivolumab) and PD-ligand (L)1 (durvalumab and avelumab) ICI immunotherapy have been performed in patients with recurrent and metastatic HNSCC. Pembrolizumab combined with platinum and fluorouracil was found to outperform the cetuximab-based platinum and fluorouracil combination in terms of overall survival (median, 13.6 vs. 10.1 months) when administered as the first-line of treatment for recurrent and metastatic HNSCC. Nevertheless, nearly 64% patients still did not respond to the PD-1/PD-L1 ICI immunotherapy, indicating innate, adapted, or quickly acquired resistance to the treatment [5–8]. Unfortunately, little is known about mechanisms underlying the response of patients with HNSCC to immunotherapy.

The tumor microenvironment (TME) is composed of transformed cells, infiltrating immune cells and stromal cellular elements. Tumor-infiltrating immune cells are known to substantially influence therapeutic responses and clinical outcomes [9]. For instance, tumor-associated macrophages and regulatory T cells are associated with pro-tumor functions [10–12], whereas tumor-infiltrating lymphocytes and CD8+ T effector cells have been associated with improved clinical outcomes and better response to immunotherapy [13–15]. However, the mechanisms by which immune infiltration affects immunotherapy outcome in patients with HNSCC remain poorly understood.

With the application of high-throughput sequencing, a comprehensive landscape of genomic and transcriptomic alterations in HNSCC has emerged from The Cancer Genome Atlas (TCGA) database, permitting the analysis of, for instance, DNA methylation, single nucleotide variants (SNVs) and copy number variations (CNVs) [16, 17]. Several recurrent chromosomal region copy number alterations have been found to affect broad and focal chromosomal regions, leading to the activation of multiple candidate driver genes [18]. Alterations in associated pathways, such as p53, mitogen-activated protein kinase, phosphatidylinositol 3′–kinase-Akt and nuclear factor kappa-B signalling pathways, may elicit complementary or synergistic effects that are important in tumorigenesis [19]. However, whether genomic alterations drive corresponding changes in the TME in HNSCC remains to be comprehensively elucidated; moreover, their functional interactions and roles in immunotherapy

outcome and clinical prognosis require further exploration.

Herein based on large-scale RNA sequencing (RNA-seq) data pertaining to patients with HNSCC from TCGA database, we investigated 22 subpopulations of infiltrating immune cells using the CIBERSORT algorithm. We analysed TME infiltration pattern (TME score) and systematically correlated TME phenotypes with genomic characteristics and clinicopathological features of HNSCC. We found TME score to be an effective prognostic biomarker and it seems useful for predicting immunotherapy outcome in patients with HNSCC.

RESULTS

Overview of TME characteristics in HNSCC

To investigate TME characteristics in HNSCC, based on the RNA-seq data of 502 patients with HNSCC from TCGA database, we systematically investigated 22 subpopulations of infiltrating immune cells using CIBERSORT. The distribution ratio of infiltrating immune cells among the samples showed a significant difference (Figure 1A), and the TME cell network depicted a comprehensive landscape of tumor–immune cell interactions, cell lineages and their effects on the overall survival (OS) of patients with HNSCC (Figure 1B; Supplementary Table 1).

Further, we performed unsupervised consensus clustering of all tumor samples for the molecular classification of HNSCC. The optimal number of clusters was determined by the K value. After assessing relative changes in the area under the cumulative distribution function curve and consensus matrix heatmap, we selected a three-cluster solution (K = 3), which showed no appreciable increase in the area under the cumulative distribution function curve (Figure 1C). To gain further insights into the molecular heterogeneity of HNSCC, unsupervised consensus clustering was performed to explore discernible patterns of the TME clusters. Based on the consensus matrix heatmap, three distinct TME-based molecular clusters were identified (Figure 1D). To further clarify the clinical implications of these clusters, we performed Kaplan–Meier curve analyses to elucidate the association between them and clinical prognosis. Notably, survival analysis based on the TME phenotypes indicated that TME cluster 3 (n = 124) was significantly associated with better prognosis and that TME cluster 2 (n = 150) was associated with poorer prognosis (log-rank test, P = 0.0018). Of the 500 patients with HNSCC, 226 belonged to TME cluster 1, which was associated with

intermediate prognosis (log-rank test, P > 0.05) (Figure 1E).

Signature and functional annotation of the TME clusters

To determine the characteristics of immune cells in the three TME clusters, infiltrating immune cell subpopulations were analysed using CIBERSORT. Tumor ploidy and malignant cell purity showed no differences among the three clusters (Supplementary Figure 1A and 1B), whereas the proportion of infiltrating immune cells among the clusters showed

significant differences, particularly that of resting memory CD4+ T cells, M0 macrophages, naïve B cells and plasma cells (Figure 2A). Further, the clusters were subjected to unsupervised hierarchical clustering, and the obtained results revealed that the 22 subpopulations of infiltrating immune cells showed differential patterns among the clusters (Figure 2B). To identify and elucidate differences in infiltrating immune cells, the relative fraction of 22 leukocyte subpopulations in each sample was estimated based on the differential expression of 547 genes [20]. Comparing the TME clusters revealed several differences, including increased levels of naïve B cells, regulatory T cells and
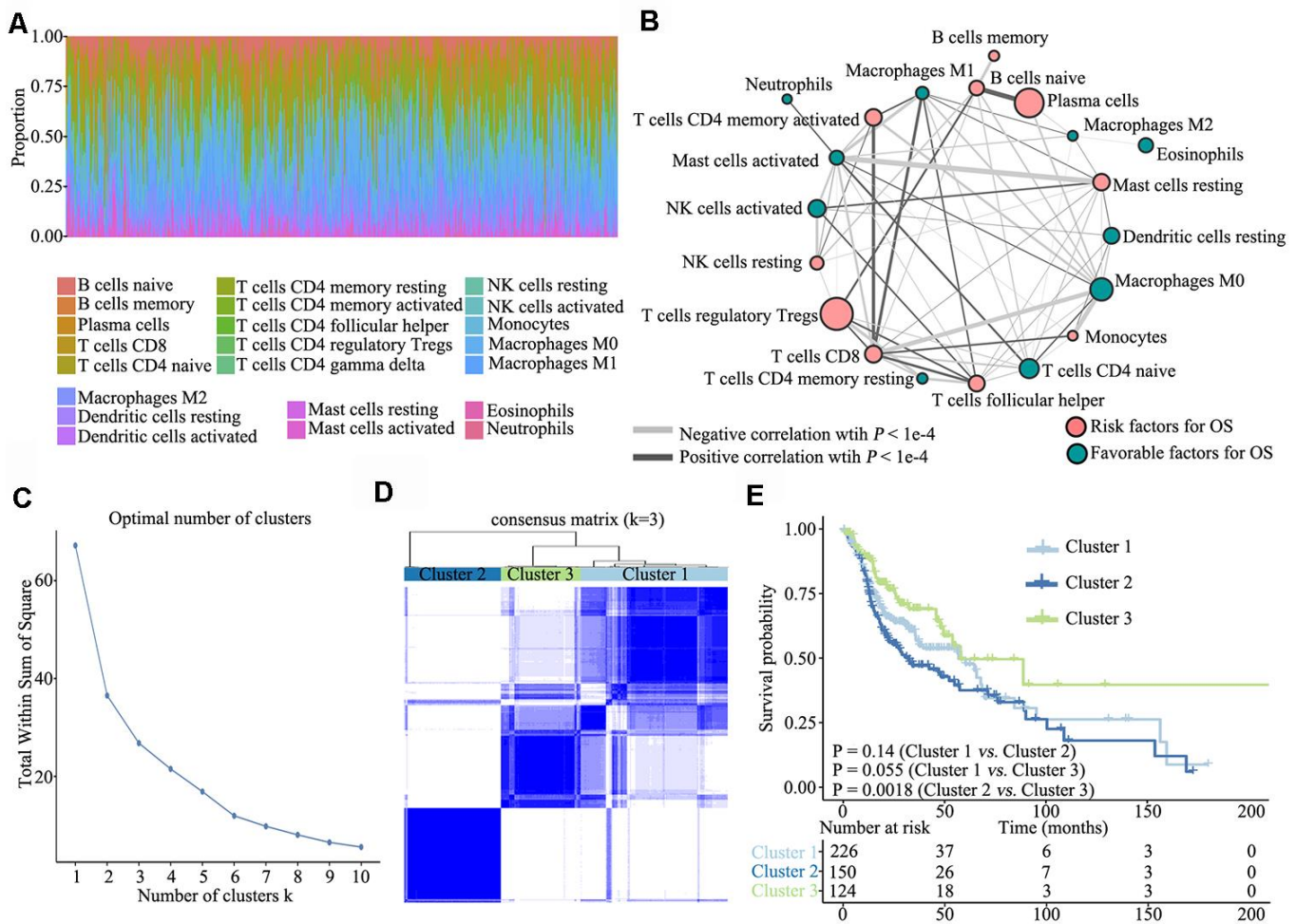
Figure 1. Overview of TME characteristics. (**A**) Relative percentage of each immune cell type in 502 patients with HNSCC from TCGA database. (**B**) Tumor–immune cell interactions. The size of each cell represents the impact of each TME cell type on survival and was calculated using log10 (log-rank test P value). Risk factors for overall survival are indicated in pink, and favourable factors are in green. The lines connecting TME cell types represent cellular interactions. The thickness of the lines represents the strength of correlation, which was estimated using Spearman correlation analysis. Negative correlation is indicated in grey and positive correlation in black. (**C**) The elbow criterion determines the optimal number of TME clusters (K = 3). (**D**) Consensus clustering analysis identification of the three TME clusters (samples, n = 500). The white (consensus value = 0, samples never clustered together) and blue (consensus value = 1, samples always clustered together) heatmap display sample consensus. (**E**) Kaplan–Meier curves for survival probability of the three clusters. Log-rank test was used for data analysis.

eosinophils; in contrast, the levels of resting memory CD4+ T cells, resting NK cells and M2 macrophages were markedly reduced (Figure 2C).

To identify biological characteristics underlying each TME phenotype, differentially expressed genes (DEGs) among the three clusters were analyzed using the limma package. In total, 306, 998 and 676 DEGs were obtained upon three comparisons (TME cluster 1 vs. TME cluster 2, TME cluster 2 vs. TME cluster 3 and TME cluster 1 vs. TME cluster 3). Overall, 622 common DEGs were further screened in the comparisons. Next, we used random forest algorithms for dimension reduction to extract phenotype signatures. Unsupervised hierarchical clustering was based on the expression of the 312 most representative D EGs (Supplementary Table 2). Gene ontology enrichment analysis of TME signature genes was conducted using the R package. Significantly enriched biological processes are summarized in Supplementary Table 3. Analyses of TME signature mRNAs indicated that within the cellular component category, immunoglobulin complex, external side of plasma membrane, T cell receptor complex and receptor complex were significantly annotated (Figure 2D). Further, within the biological process category, antigen receptor-mediated signalling pathway, regulation of lymphocyte activation, positive regulation of cell activation and positive regulation of leukocyte activation were significantly annotated (Figure 2E), and within the molecular function category, pathways involved in antigen binding, immunoglobulin receptor binding and cytokine receptor activity were significantly annotated (Figure 2F). These findings indicated that the immune landscape of the three TME clusters plays a key role in immune regulation.

Clinical characteristics of the TME phenotypes

Next, we systematically classified the TME clusters according to the gene coefficient value and calculated TME score. The maxstat R package was used to identify TME score breakpoint. The distribution of risk scores and overall survival status of patients are shown in Figure 3A and 3B, respectively. TME score was calculated for all patients, and the TME score breakpoint was used to classify them into the high (n = 220) and low (n = 280) TME score groups. High TME score and was associated with favourable outcomes, and low TME score and was associated with poor outcomes (Figure 3C). Kaplan–Meier curve and Cox regression analyses further suggested that patients in the high TME score group had significantly better OS probability than those in the low TME score group [HR 1.401 (1.074–1.827), log-rank test, P = 0.013] (Figure 3D). To

understand the pathway two TME score groups involved in, gene set enrichment analysis was performed, which indicated that allograft rejection (E score = −0.414; P = 0.0091), inflammatory response (E score = −0.388; P = 0.032), interferon-α response (E score = −0.495; P = 0.002), interferon-γ response (E score = −0.450; P = 0.001), reactive oxygen species (E score = −0.468; P = 0.048) and xenobiotic metabolism (E score = −0.390; P = 0.0343) pathways were significantly downregulated in the high TME score group (Figure 3E).

**Somatic mutations of the TME phenotypes**

To detect driver mutations in the TME, we analysed HNSCC-related SNP data and consequently detected 21 putative driver genes, including TP53, TTN, CSMD3, CDKN2A and NOTCH1 (Figure 4A and 4B), which were associated with the TME, using random forest algorithm with 1000 iterations. The TP53 gene showed the highest mutation rate (74% vs. 70% in the high and low TME score groups, respectively), which was consistent with the findings of other studies [21–25]. The somatic mutations of TP53, TTN, CSMD3, CDKN2A, FRG1B, FAT1 and NOTCH1 were all >20% in the high and low TME score groups; in particular, FAT1 was significantly different between high TME score and low TME score groups (P < 0.05) (Figure 4C). We also found that several genes involved in the cell cycle pathway (e.g., TP53, CDKN2A, NOTCH1 and PIK3CA) were frequently altered. Clinical reports have described associations between individual altered genes and response or resistance to ICIs [26, 27]. These mutations may be associated with a change in the TME. Consequently, we further detected and classified candidate somatic mutations, which led to the identification of nine classifications; nonstop, in-frame (insert and delete), frame-shift (insert and delete), translation start, splice, nonsense and mutations were the common somatic mutations (Figure 4D). Mutation type and spectrum analyses showed that single nucleotide mutations, particularly C > T transitions at TpCpW trinucleotide sites, were the predominant type of mutations (Figure 4E, 4F).

Mutational signatures

Mutational signatures in the cancer genome might reflect and help trace DNA damage caused by DNA-damaging agents to which cells have been exposed. Thus, we counted the number of SNVs in the matrix of 96 possible mutations occurring in a trinucleotide context in each HNSCC sample and found that the predominant mutations were C > T and/or C > G transitions at TpCpW trinucleotide sites in both the high and low TME score groups (Figure 5A, 5B).

Using 1,000 iterations of non-negative matrix factorization [28], we then performed cosine similarity analyses to compare mutational signatures in HNSCC with current COSMIC; consequently, six independent and stable mutational signatures were identified. *De novo* signatures identified in the high TME score group were enriched in APOBEC cytidine deaminase signature (COSMIC Signature 13; cosine similarity: 0.842),
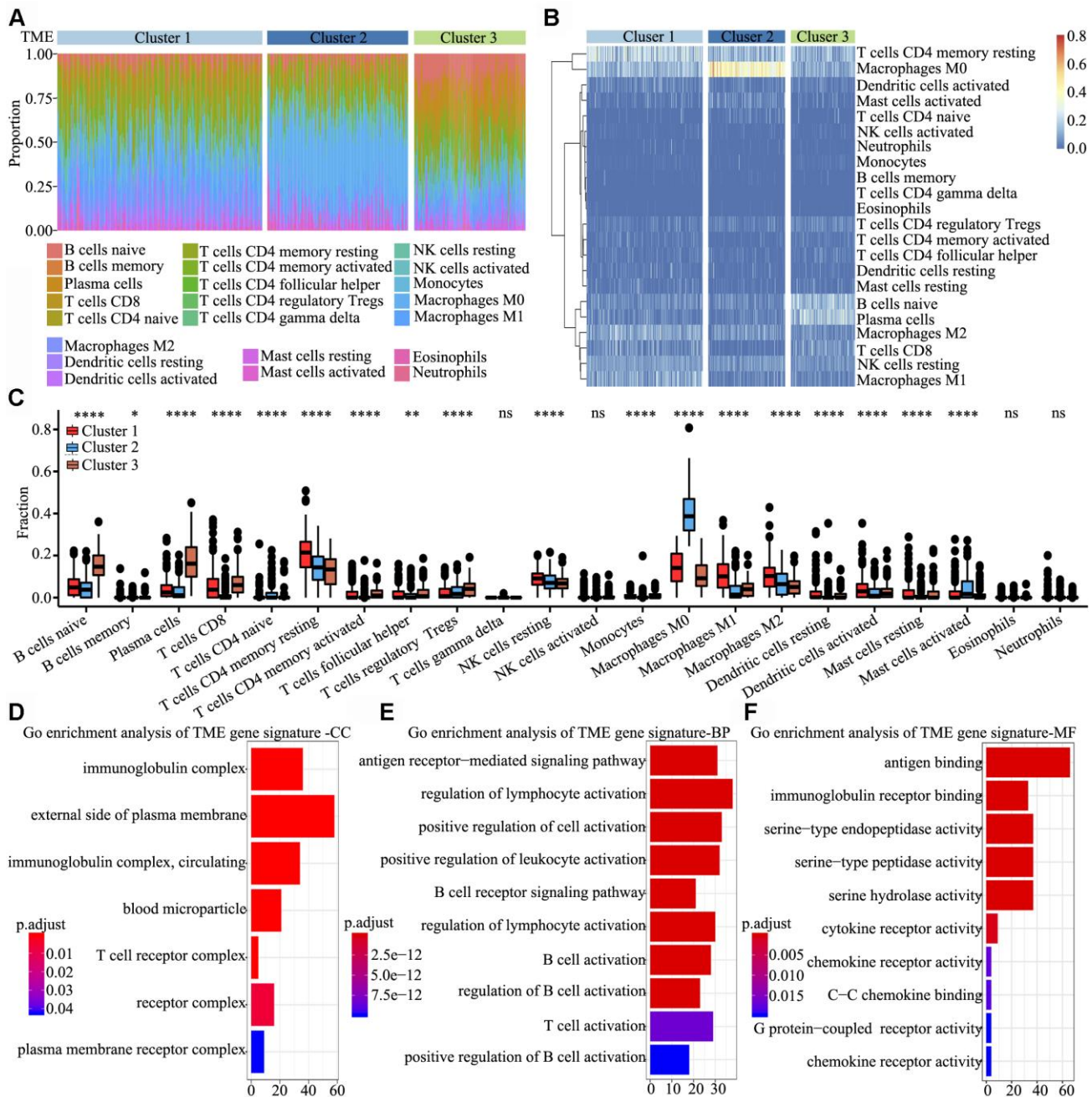

Figure 2. Signature and functional annotation of the TME clusters. (**A**) Relative percentage of each immune cell type in the three TME clusters. (**B**) Unsupervised hierarchical clustering of the clusters. (**C**) Relative populations of TME cells present in the three clusters. Within each group, the scattered dots represent expression values of TME cells. We also plotted the Immunoscore for the three clusters. The thick line represents the median value. The lower and upper ends of the boxes are the 25th and 75th percentiles. The whiskers encompass 1.5 times the interquartile range. Statistical differences in the three clusters were compared using the Kruskal–Wallis test. The range of P values are labelled above each boxplot with asterisks (* P < 0.05, ** P < 0.01, *** P < 0.001, ****P < 0.0001, ns = not significant). (**D–F**) Gene ontology enrichment analysis of TME signature genes in the cellular component (**D**), biological process (**E**) and molecular function (**F**) categories. The x-axis indicates the number of genes within each gene ontology term.

UV exposure (COSMIC Signature 7; cosine similarity: 0.93), tobacco mutagens (COSMIC Signature 4; cosine similarity: 0.917) and spontaneous deamination of 5-methylcytosine (COSMIC Signature 1; cosine similarity: 0.942) (Figure 5C–5F). In contrast, *de novo* signatures identified in the low TME score group were enriched in APOBEC cytidine deaminase signature (COSMIC Signature 13; cosine similarity: 0.809) and defective DNA mismatch repair (COSMIC Signature 6; cosine similarity: 0.849) (Figure 5G, 5H). C > T and C > G mutations at TpCpN trinucleotides were attributed to the overactivity of the AID/APOBEC family of cytidine deaminases. These data indicate that

the overactivity of APOBEC family genes may be involved in HNSCC tumorigenesis and immune outcome.

## Copy number alterations of the TME phenotypes

We observed that 98% HNSCC samples had CNVs at the chromosome arm level, including loss at 3p, 4p, 5q, 8p, 9p, 11q, 13q, 18q and 21q and gain at 1q, 3q, 5p, 8q, 14q, 20p and 20q. Analyses using GISTIC indicated that the chromosome arm level had a significant gain at 8q and 20q in the high TME score group (Figure 6A) and at 8q, 3q and 3p in the low TME score group
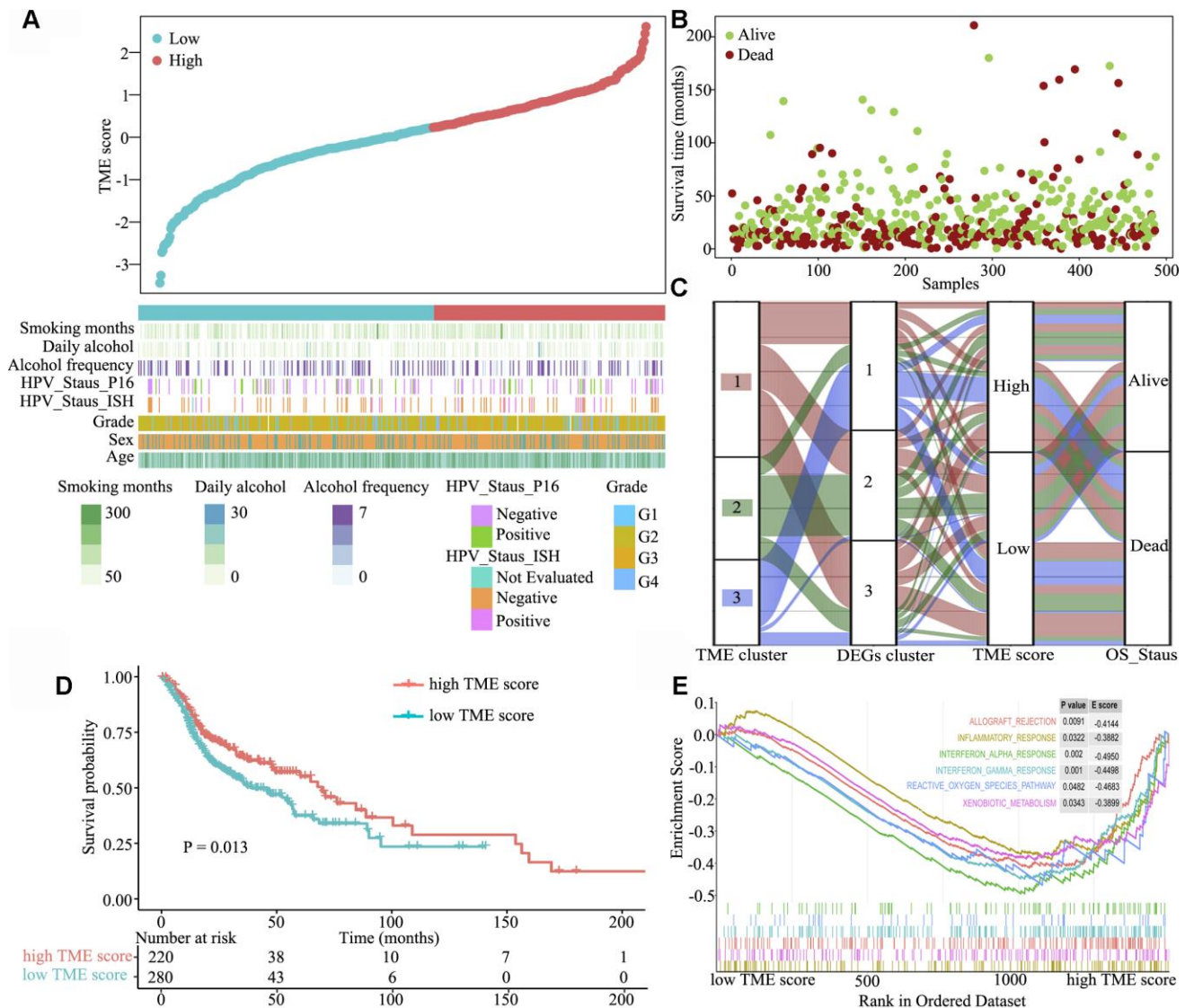
Figure 3. Clinical characteristics of the TME phenotypes. (**A**) TME score distribution in patients with HNSCC. (**B**) Overall survival status of patients with HNSCC. (**C**) Alluvial diagram of TME gene clusters in groups with different TME clusters, DEG clusters, TME scores and survival outcomes. (**D**) Gene set enrichment analysis of hallmark gene sets between the high and low TME score groups. (**E**) Kaplan–Meier curves for the high and low TME score groups. As evident, the high TME score group was associated with better outcomes than the low TME score group (log-rank test, $P < 0.001$).

(Figure 6B); moreover, the chromosome arm level had a significant loss at 8p, 3p and 11q in the high TME score group and at 9p, 5q and 4p in the low TME score group (Figure 6C, 6D). Similar SNV and CNV profiles along with similar RNA expression patterns supported that HNSCC belongs to the "squamous" molecular subtype, as identified by Hoadley et al. [29]. We also identified 13 focal regions with recurrent gain of copy number and

20 regions with recurrent loss of copy number (both $q < 1e{-}4$), in which many genes have previously been identified to be tumor-associated genes (Supplementary Table 4). Among them, three recurrent focal amplifications (11q13.3, 7p11.2 and 3q26.33) (Figure 6E, 6F) and three recurrent focal deletions (9p21.3, 8p23.2 and 4q35.2) (Figure 6G, 6H) were identified for the first time in the TME score groups.
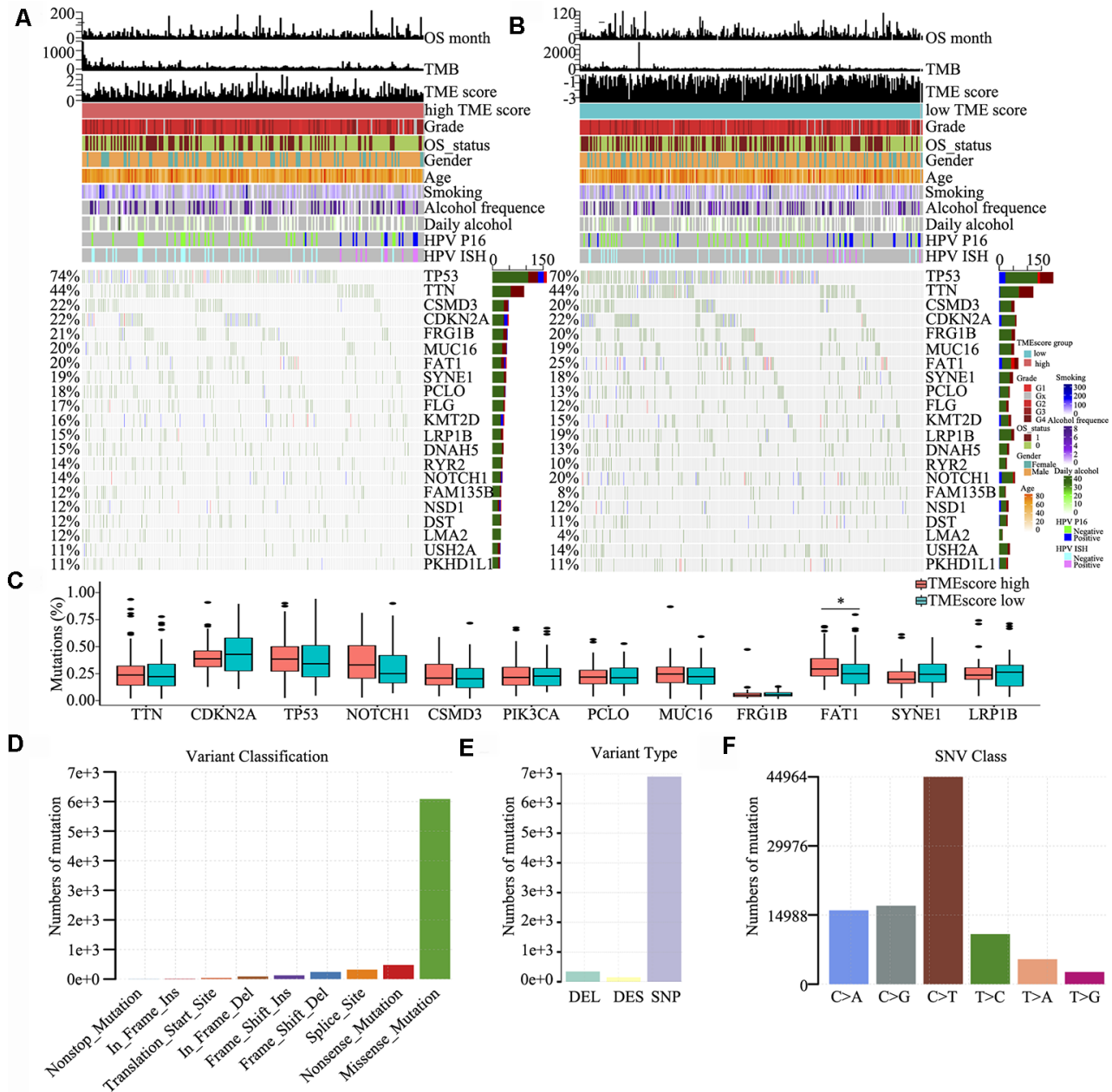
Figure 4. Somatic mutations in HNSCC. (**A**, **B**) Distribution of highly variant mutated genes correlated with TME score groups. The upper bar plot indicates overall survival (OS), TMB and TME score for each patient, whereas the left bar plot shows the mutation frequency of each gene in separate TME score groups [high (**A**) and low (**B**) TME score groups]. TME score, grade, overall survival status, gender, age, smoking, alcohol frequency, daily alcohol, HPV P16 status and HPV ISH status are shown as patient annotations. (**C**) Mutation percentage of common mutated genes in the TME score groups. (**D**) Genome variant classification. (**E**) Genome variant type. (**F**) Single nucleotide variant class.

Integrated genomic landscape of the TME score phenotype

To elucidate the molecular characteristics of HNSCC, according to information pertaining to the TME clusters and TME score groups, as well as tumor purity, malignant cell ploidy, tumor mutation burden (TMB) and clinical information (age, grade and OS status), a comprehensive genomic landscape of HNSCC samples was integrated and has been depicted in Supplementary Table 5. As shown in Figure 7A and Supplementary Figure 1C and 1D, no difference was present between the high and low TME score groups with regard to tumor ploidy and malignant cell purity. However, consistent with TMB, TME score was able to significantly differentiate between high- and low-risk HNSCC. TMB, in concert with PD-L1 expression, is reportedly a useful biomarker for immune checkpoint blockade selection in diverse cancers [30]. Herein ROC analysis showed that in comparison with TMB, TME score is a similar efficacious biomarker to determine the effectiveness of immunotherapy in patients with HNSCC (area under ROC: 0.549 vs. 0.572, P = 0.64) (Figure 7B).

DISCUSSION

HNSCC is a heterogeneous disease of the upper aerodigestive tract, encompassing distinct histological types, different anatomical sites and even HPV+ or EBV+ cancers. Dysfunctional immune cells in patients with recurrent/metastatic HNSCC can be repaired using immunotherapies in combination with conventional treatment methods. In this study, our findings indicated that assessing the immune score via the TME signature provided a potent predictor of survival in patients with HNSCC. Functional analysis of TME signature genes suggested that they are involved in the activation and inhibition of immune responses. Mutational signature and CNV analyses indicated that somatic mutations in tumor DNA or chromosome arm can give rise to neoantigens and mutation-derived antigens, which are recognized and targeted by the immune system, thereby activating the immune response. We report that TME score can act as an immunotherapy biomarker for HNSCC.

Immune checkpoint blockade therapy that impedes the PD-1/PD-L1 and anti-CTLA-4 pathway can increase
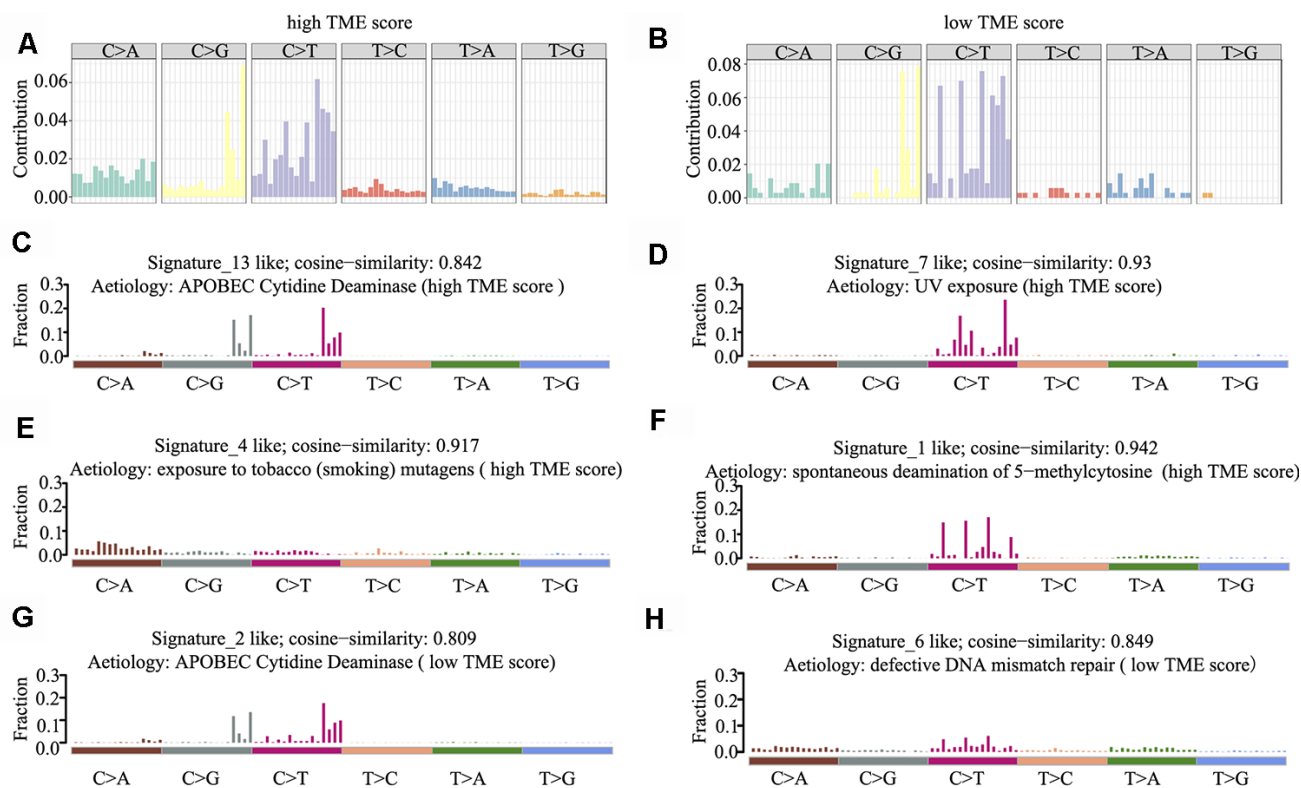
Figure 5. Mutational signature of the TME score groups. (**A**, **B**) Distribution of mutation type frequency in the high (**A**) and low (**B**) TME score groups. (**C**–**H**) Mutational signatures identified in the high (**C**–**F**) and low (**G**–**H**) TME score groups, respectively. The y-axis indicates exposure of 96 trinucleotide motifs to overall signature. The plot title indicates best match against validated COSMIC signatures and cosine similarity value along with the proposed aetiology.
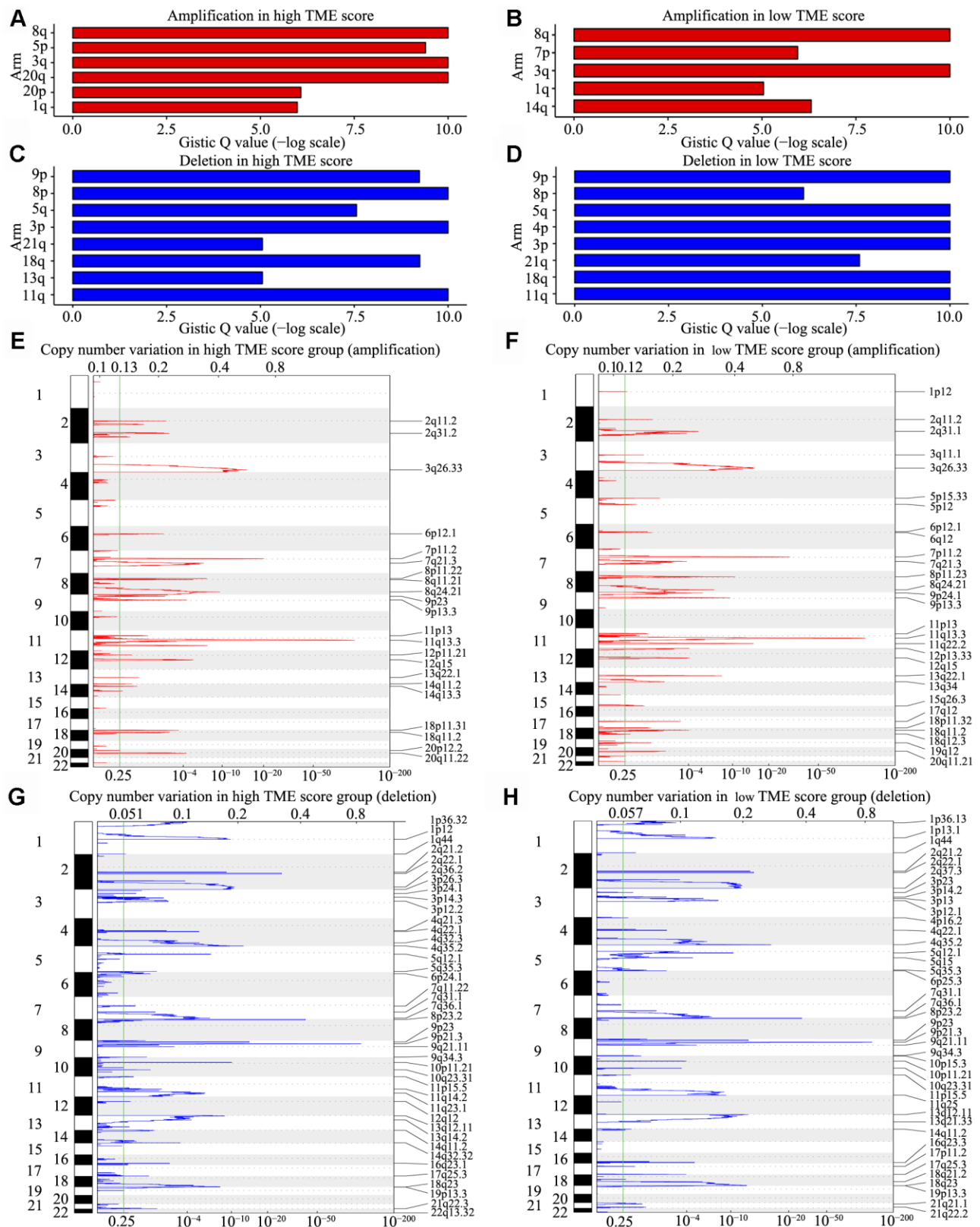
Figure 6. CNV analysis in HNSCC. (**A–D**) CNV at arm level. The bar graphs show the frequency of arm-level CNV amplification (**A**, **B**) and deletion (**C**, **D**), the vertical axis denotes chromosome arms. (**E–H**) CNV at focal regions detected by GISTIC v2·0. Regions of recurrent focal amplifications (**E**, **F**) and focal deletions (**G**, **H**) in the high and low TME score groups are plotted by false discovery rate (x-axis) for each chromosome (y-axis). Dashed lines represent the centromere of each chromosome.

OS in patients with advanced melanoma, non-small-cell lung cancer, urothelial cancer, renal cell carcinoma and other cancer types [31–35]. However, such a response is observed in only a minority of patients. Several studies have reported that PD-1 and PD-L1 expression, microsatellite instability status and mutation load are not robust biomarkers for predicting the benefits of immune checkpoint blockade [36–38]. Thus, it is of utmost importance to screen effective biomarkers for checkpoint immunotherapy. HNSCC, similar to all other cancers, results from a stepwise accumulation of genomic instability, chromosomal aberrations and genetic mutations [39]. Within cancer tissues, arising mutant cells strive for metabolism, avoid immune surveillance and in collaboration with the extracellular matrix, tumor stroma, immune cells, vessel remodelling and diverse immunoinhibitory soluble or membrane-bound cytokines, establish a unique TME [40]. Tekpli et al. reported an independent poor-prognosis subtype of breast cancer defined by a distinct TME [41]. Zeng et al. also reported that TME characteristics could be used to interpret the response of gastric tumors to immunotherapies [42]. Herein our findings provide clarity on the comprehensive landscape of the interactions between the clinical characteristics of HNSCC and tumor-infiltrating immune cells. With the application of computational algorithms, a methodology was established to determine TME score.

Herein we used CIBERSORT to evaluate differential immune cell infiltration in paired HNSCC and adjacent normal tissues, and the results revealed considerable differences in the immune cell fraction in both the intra- and intergroups. Our results also uncover details pertaining to the infiltration of LM22 immune cell subpopulations in HNSCC, with the proportion of macrophages being >45% (M0 = 27%, M2 = 9% and M1 = 9%). CD4+ memory activated T cells, resting NK cells, regulatory T cells, CD8+ T cells, follicular T cells, monocytes, resting mast cells, plasma cells, naïve B cells and memory B cells were risk factors for OS, while neutrophils, activated mast cells, activated NK cells, resting memory CD4+ T cells, naïve CD4+ T cells, M0 macrophages, M2 macrophages and eosinophils were favourable factors for OS. HNSCC samples were clustered into three main clusters: low, intermediate and high immune infiltration. Patients in the high immune infiltration cluster showed the best survival probability than those in the other two clusters.

Integrated analyses revealed that TME score is a prognostic biomarker for HNSCC; in particular, naïve B cells, regulatory T cells and follicular T cells in the high TME score group were associated with an improved outcome, whereas neutrophils and activated mast cells in the low TME score group were associated with poorer outcome. Further, the involvement of several
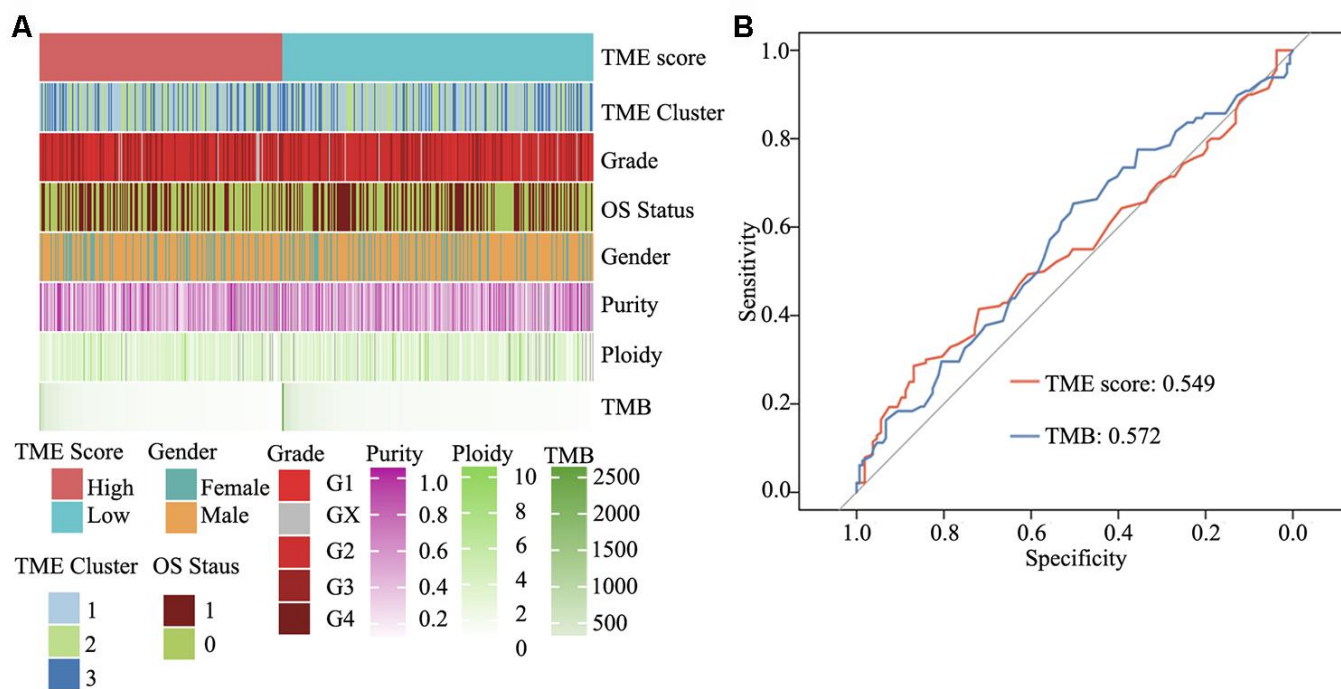
Figure 7. Clinical and integrated genomic landscape of HNSCC with the TME score phenotype. (**A**) Comprehensive genomic landscape of HNSCC. (**B**) Prediction immunotherapy effect of TME score by ROC analysis.

immune activation pathways suggested that TME score is a predictive biomarker to further advance precision immunotherapy of HNSCC. Higher non-synonymous mutational burden has been associated with an improved overall response rate, durable clinical benefits and progression-free survival in patients treated with ICIs [43]. C > T and/or C > G mutations at TpCpN trinucleotides were attributed to the overactivity of the AID/APOBEC family of cytidine deaminases. Although the APOBEC family of proteins might serve as endogenous carcinogenic mutagens [44], they also play crucial roles in the innate immune response to viral infections by modifying the viral genome [45, 46]. Through ROC analysis, we found that TME score showed a predictive value similar to that of TMB, which indicates that it may act as an independent biomarker for immunotherapy.

To summarize, our results provide a comprehensive view of the cellular, molecular and genetic factors associated with TME infiltration patterns and define a potential mechanism by which tumors respond to immunotherapy. We believe that TME score is a useful biomarker that can effectively predict immunotherapy outcome in patients with HNSCC.

MATERIALS AND METHODS

Data acquisition and processing

RNA-seq data (N = 546) and survival data (N = 530) were obtained from TCGA database (https://portal.gdc.cancer.gov/projects/TCGA-HNSC). Patients diagnosed with HNSCC and with clinicopathological and survival information (N = 500) were transformed as original read counts. Next, genes with low expression levels were removed using the filterByExpr function of edgeR. The expression data were then transformed using voom to facilitate the evaluation of immune cell subpopulations by CIBERSORT.

Evaluation of tumor-infiltrating immune cells

To evaluate the number of each type of tumor-infiltrating immune cells, we applied the original CIBERSORT gene signature file LM22, which defines 22 immune cell subpopulations, to analyse datasets pertaining to HNSCC. Gene expression datasets were prepared using standard annotation files and data uploaded to the CIBERSORT web portal (http://cibersort.stanford.edu/), with the algorithm run using the default signature matrix at 1,000 permutations [20].

Consensus clustering of TME-infiltrating patterns

Different TME cell infiltration patterns were grouped using hierarchical agglomerative clustering based on Euclidean distance and Ward's linkage. In addition, we used an unsupervised clustering method (K-means) for dataset analysis to identify TME cell infiltration patterns and to classify patients for further analyses [47]. The consensus clustering algorithm was used to detect the number of clusters in the meta-dataset. This was performed using the *ConsensusClusterPlus* R package and was repeated 1,000 times to ensure the stability of classification [48].

DEGs associated with the TME phenotypes

To identify genes associated with TME cell infiltration patterns in patients with HNSCC, we grouped the TME cluster into three classes: TME cluster 1, 2 and 3. DEGs among these three clusters were determined using the R package limma, which implements an empirical Bayesian approach to estimate changes in gene expression levels using moderated *t*-tests. DEGs among the TME clusters were determined by significance criteria (P < 1e−3 and |log2FC| > 1), as implemented in the R package limma.

TME gene signature analyses

The construction of TME metagenes was performed as follows. Each DEG among TME cluster 1, 2 and 3 was standardized across all samples. An unsupervised clustering method (K-means) was used for DEG analyses to classify patients into three groups for further analyses. The random forest classification algorithm was used for dimension reduction to reduce noise or redundant class-specific DEGs in TME clusters. DEGs among the TME clusters were annotated using the clusterProfiler R package [49]. A consensus clustering algorithm was applied to define the cluster of genes and to calculate the signature score. For gene expression (normalized by RMA or TPM methods) analysis, the expression of each gene in a signature was first transformed into a z-score, and then, principal component analysis (PCA) was performed using the consensus clustering algorithm. Principal component 1 was extracted to serve as the gene signature score. After obtaining the prognostic value of each gene signature score, we used the gene expression grade index method to obtain the TME score of each patient [50]:

$$\text{TME score} = \sum \text{voom}(X) - \sum \text{voom}(Y)$$

wherein X represents the signature score of expression value of positive gene clusters, and Y represents the expression level of gene clusters. Patients with HNSCC were therefore assigned to groups based on high or low TME scores using the cut-off value obtained with the

maxstat R package; the clinical outcomes were further analysed.

Functional and pathway enrichment analysis

Gene ontology terms (cellular component, biological process and molecular function) of TME gene signatures were identified using the clusterProfiler R package [49]. Gene set enrichment analysis of DEGs with high or low TME scores was performed based on the MSigDB database (Broad Institute) [51]. Broad hallmarks and specific pathways of interest from curated gene sets/canonical pathway collection were all included.

Mutational signature analysis

Mutational signature was first identified using the BayesNMF algorithm. The count of somatic mutations was calculated for each type of substitution (96 trinucleotide mutation contexts) to generate a mutational catalogue. We then ran the Bayesian NMF 1,000 times with the hyperparameter for the inverse gamma prior setting to 10 ($a = 10$); the iterations were terminated when the tolerance for convergence was < 10 e–7, and half-normal was selected as "pirors" for this algorithm [52]. Signatures identified following matrix factorization were scaled and the results were compared to the Catalogue of Somatic Mutations in Cancer (COSMIC) signature database. A cosine similarity value was then estimated for the best possible match [53]. For signature enrichment analysis, we used matrix H, containing signature exposures for every sample in every signature. Using K-means clustering, the samples were grouped into r clusters, thereby assigning samples to an identified signature.

For apolipoprotein B-editing catalytic polypeptide-like subunit (APOBEC)-based enrichment analysis, we used the method described by Roberts et al. [44] to estimate an enrichment score, which defined the strength of APOBEC-related mutagenic processes for every tumor sample in Mutation Annotation Format. Briefly, the enrichment of $C > T$ mutations occurring within over all $C > T$ mutations in a given sample was compared to background cytosines and occurring around ±20 bp of mutated bases. We further used this method to identify genes associated with APOBEC enrichment by classifying samples as APOBEC-enriched (enrichment score > 2) and non-APOBEC-enriched (enrichment score < 2), followed by using one-way Fisher's exact tests to identify genes overrepresented among APOBEC-enriched samples.

Analysing CNVs

To identify variant regions that drive cancer pathogenesis, the Genomic Identification of Significant Targets in Cancer (GISTIC) algorithm was used to detect genomic regions manifesting amplifications (copy number > 1) or deletions (copy number < −1) [54]. G-score was determined to evaluate the amplitude of aberrations and frequency of occurrence in variant regions. Briefly, GISTIC v2.0 was used to define the prepared CNV profiles for all genes in the 488 HNSCC samples. False discovery rate q-values were assigned to each variant region. "Peak regions," also known as significantly aberrant regions, indicated the greatest frequency and amplitude of aberrations [55, 56].

Analysing somatic DNA copy number alterations

To elucidate tumor purity and malignant cell ploidy from the somatic DNA copy number alterations, ABSOLUTE was used to detect subclonal heterogeneity and somatic homozygosity and to calculate statistical sensitivity to identify specific aberrations. Briefly, as described by Carter et al. [57], based on the CNV data of somatic DNA copy number alterations, pre-designed cancer karyotypes and somatic mutation frequencies were scored and integrated. The highest score was considered to represent the optimal model, and tumor purity and malignant cell ploidy were then detected using the ABSOLUTE R package limma.

Statistical analysis

Student's t-test and ANOVA were utilized to compare continuous and discrete variables, respectively. Pearson's chi-squared test was used for comparing categorical clinicopathological variables, and survival probability and differences were analysed using log-rank test and the Cox proportional hazards model (multivariate analysis). Statistical analyses were performed using standard R packages (version 3.5.2). $P < 0.05$ indicated statistical significance.

Abbreviations

HNSCC: head and neck squamous cell carcinoma; SNVs: single nucleotide variants; CNVs: copy number variations; TMB: tumor mutational burden; TME: tumor microenvironment; TCGA: The Cancer Genome Atlas database; RNA-seq: RNA sequencing; DEGs: Differentially expressed genes; COSMIC: Catalogue of Somatic Mutations in Cancer; APOBEC: apolipoprotein B-editing catalytic polypeptide-like subunit; GISTIC: Genomic Identification of Significant Targets in Cancer; OS: overall survival.

AUTHOR CONTRIBUTIONS

JZ, XZ, HLiJ and HLoJ performed all experiments, prepared figures and drafted the manuscript. JZ, TX,

YHT, RL and BYW participated in data analysis and interpretation of results. JZ, JXZ and YWY designed the study and participated in data analysis. All authors have read and approved the manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin. 2005; 55:74–108.
   https://doi.org/10.3322/canjclin.55.2.74
   PMID:15761078

2. de Andrade DA, Machiels JP. Treatment options for patients with recurrent or metastatic squamous cell carcinoma of the head and neck, who progress after platinum-based chemotherapy. Curr Opin Oncol. 2012; 24:211–17.
   https://doi.org/10.1097/CCO.0b013e3283510773
   PMID:22498572

3. Er Ö, Colak SG, Ocakoglu K, Ince M, Bresolí-Obach R, Mora M, Sagristá ML, Yurt F, Nonell S. Selective photokilling of human pancreatic cancer cells using cetuximab-targeted mesoporous silica nanoparticles for delivery of zinc phthalocyanine. Molecules. 2018; 23:2749.
   https://doi.org/10.3390/molecules23112749
   PMID:30355983

4. Licitra L, Mesia R, Rivera F, Remenár É, Hitt R, Erfán J, Rottey S, Kawecki A, Zabolotnyy D, Benasso M, Störkel S, Senger S, Stroh C, Vermorken JB. Evaluation of EGFR gene copy number as a predictive biomarker for the efficacy of cetuximab in combination with chemotherapy in the first-line treatment of recurrent and/or metastatic squamous cell carcinoma of the head and neck: EXTREME study. Ann Oncol. 2011; 22:1078–87.
   https://doi.org/10.1093/annonc/mdq588
   PMID:21048039

5. Ferris RL, Blumenschein G Jr, Fayette J, Guigay J, Colevas AD, Licitra L, Harrington K, Kasper S, Vokes EE, Even C, Worden F, Saba NF, Iglesias Docampo LC, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016; 375:1856–67.
   https://doi.org/10.1056/NEJMoa1602252
   PMID:27718784

6. Chow LQ, Haddad R, Gupta S, Mahipal A, Mehra R, Tahara M, Berger R, Eder JP, Burtness B, Lee SH, Keam B, Kang H, Muro K, et al. Antitumor activity of pembrolizumab in biomarker-unselected patients with recurrent and/or metastatic head and neck squamous cell carcinoma: results from the phase ib KEYNOTE-012 expansion cohort. J Clin Oncol. 2016; 34:3838–45.
   https://doi.org/10.1200/JCO.2016.68.1478
   PMID:27646946

7. Seiwert TY, Burtness B, Mehra R, Weiss J, Berger R, Eder JP, Heath K, McClanahan T, Lunceford J, Gause C, Cheng JD, Chow LQ. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. Lancet Oncol. 2016; 17:956–65.
   https://doi.org/10.1016/S1470-2045(16)30066-3
   PMID:27247226

8. Bauml J, Seiwert TY, Pfister DG, Worden F, Liu SV, Gilbert J, Saba NF, Weiss J, Wirth L, Sukari A, Kang H, Gibson MK, Massarelli E, et al. Pembrolizumab for platinum- and cetuximab-refractory head and neck cancer: results from a single-arm, phase II study. J Clin Oncol. 2017; 35:1542–49.
   https://doi.org/10.1200/JCO.2016.70.1524
   PMID:28328302

9. Wu T, Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett. 2017; 387:61–68.
   https://doi.org/10.1016/j.canlet.2016.01.043
   PMID:26845449

10. De Palma M, Lewis CE. Macrophage regulation of tumor responses to anticancer therapies. Cancer Cell. 2013; 23:277–86.
    https://doi.org/10.1016/j.ccr.2013.02.013
    PMID:23518347

11. Nishikawa H, Sakaguchi S. Regulatory T cells in cancer immunotherapy. Curr Opin Immunol. 2014; 27:1–7.

https://doi.org/10.1016/j.coi.2013.12.005
PMID:24413387

12. Noy R, Pollard JW. Tumor-associated macrophages: from mechanisms to therapy. Immunity. 2014; 41:49–61.
https://doi.org/10.1016/j.immuni.2014.06.010
PMID:25035953

13. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pagès C, Tosolini M, Camus M, Berger A, Wind P, Zinzindohoué F, Bruneval P, Cugnenc PH, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006; 313:1960–64.
https://doi.org/10.1126/science.1129139
PMID:17008531

14. Pagès F, Kirilovsky A, Mlecnik B, Asslaber M, Tosolini M, Bindea G, Lagorce C, Wind P, Marliot F, Bruneval P, Zatloukal K, Trajanoski Z, Berger A, et al. In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. J Clin Oncol. 2009; 27:5944–51.
https://doi.org/10.1200/JCO.2008.19.6147
PMID:19858404

15. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, Chmielowski B, Spasic M, Henry G, Ciobanu V, West AN, Carmona M, Kivork C, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. Nature. 2014; 515:568–71.
https://doi.org/10.1038/nature13954 PMID:25428505

16. Leemans CR, Snijders PJ, Brakenhoff RH. The molecular landscape of head and neck cancer. Nat Rev Cancer. 2018; 18:269–82.
https://doi.org/10.1038/nrc.2018.11 PMID:29497144

17. 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 PMID:25631445

18. Gollin SM. Cytogenetic alterations and their molecular genetic correlates in head and neck squamous cell carcinoma: a next generation window to the biology of disease. Genes Chromosomes Cancer. 2014; 53:972–90.
https://doi.org/10.1002/gcc.22214
PMID:25183546

19. Cheng H, Yang X, Si H, Saleh AD, Xiao W, Coupar J, Gollin SM, Ferris RL, Issaeva N, Yarbrough WG, Prince ME, Carey TE, Van Waes C, Chen Z. Genomic and transcriptomic characterization links cell lines with aggressive head and neck cancers. Cell Rep. 2018; 25:1332–45.e5.
https://doi.org/10.1016/j.celrep.2018.10.007
PMID:30380422

20. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12:453–57.
https://doi.org/10.1038/nmeth.3337 PMID:25822800

21. Chang J, Tan W, Ling Z, Xi R, Shao M, Chen M, Luo Y, Zhao Y, Liu Y, Huang X, Xia Y, Hu J, Parker JS, et al. Genomic analysis of oesophageal squamous-cell carcinoma identifies alcohol drinking-related mutation signature and genomic alterations. Nat Commun. 2017; 8:15290.
https://doi.org/10.1038/ncomms15290
PMID:28548104

22. Cueto García L, Cuarón A, Infante O. [Various considerations on the current use of nuclear cardiology studies]. Arch Inst Cardiol Mex. 1987; 57:169–75.
PMID:3038042

23. Chung CH, Guthrie VB, Masica DL, Tokheim C, Kang H, Richmon J, Agrawal N, Fakhry C, Quon H, Subramaniam RM, Zuo Z, Seiwert T, Chalmers ZR, et al. Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. Ann Oncol. 2015; 26:1216–23.
https://doi.org/10.1093/annonc/mdv109
PMID:25712460

24. Gross AM, Orosco RK, Shen JP, Egloff AM, Carter H, Hofree M, Choueiri M, Coffey CS, Lippman SM, Hayes DN, Cohen EE, Grandis JR, Nguyen QT, Ideker T. Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. Nat Genet. 2014; 46:939–43.
https://doi.org/10.1038/ng.3051 PMID:25086664

25. Hedberg ML, Goh G, Chiosea SI, Bauman JE, Freilino ML, Zeng Y, Wang L, Diergaarde BB, Gooding WE, Lui VW, Herbst RS, Lifton RP, Grandis JR. Genetic landscape of metastatic and recurrent head and neck squamous cell carcinoma. J Clin Invest. 2016; 126:169–80.
https://doi.org/10.1172/JCI82066 PMID:26619122

26. Burr ML, Sparbier CE, Chan YC, Williamson JC, Woods K, Beavis PA, Lam EY, Henderson MA, Bell CC, Stolzenburg S, Gilan O, Bloor S, Noori T, et al. CMTM6 maintains the expression of PD-L1 and regulates anti-tumour immunity. Nature. 2017; 549:101–05.
https://doi.org/10.1038/nature23643
PMID:28813417

27. George S, Miao D, Demetri GD, Adeegbe D, Rodig SJ, Shukla S, Lipschitz M, Amin-Mansour A, Raut CP, Carter SL, Hammerman P, Freeman GJ, Wu CJ, et al. Loss of PTEN is associated with resistance to anti-PD-1 checkpoint blockade therapy in metastatic uterine leiomyosarcoma. Immunity. 2017; 46:197–204.
https://doi.org/10.1016/j.immuni.2017.02.001
PMID:28228279

28. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, Boyault S, Burkhardt B, Butler AP, et al, Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, and ICGC PedBrain. Signatures of mutational processes in human cancer. Nature. 2013; 500:415–21.
https://doi.org/10.1038/nature12477 PMID:23945592

29. Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, Leiserson MD, Niu B, McLellan MD, Uzunangelov V, Zhang J, Kandoth C, Akbani R, et al, and Cancer Genome Atlas Research Network. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell. 2014; 158:929–44.
https://doi.org/10.1016/j.cell.2014.06.049
PMID:25109877

30. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, Peters S. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. Ann Oncol. 2019; 30:44–56.
https://doi.org/10.1093/annonc/mdy495
PMID:30395155

31. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC, Akerley W, van den Eertwegh AJ, Lutzky J, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363:711–23.
https://doi.org/10.1056/NEJMoa1003466
PMID:20525992

32. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, Tykodi SS, Sosman JA, Procopio G, Plimack ER, Castellano D, Choueiri TK, Gurney H, et al, and CheckMate 025 Investigators. Nivolumab versus everolimus in advanced renal-cell carcinoma. N Engl J Med. 2015; 373:1803–13.
https://doi.org/10.1056/NEJMoa1510665
PMID:26406148

33. Borghaei H, Paz-Ares L, Horn L, Spigel DR, Steins M, Ready NE, Chow LQ, Vokes EE, Felip E, Holgado E, Barlesi F, Kohlhäufl M, Arrieta O, et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. N Engl J Med. 2015; 373:1627–39.
https://doi.org/10.1056/NEJMoa1507643
PMID:26412456

34. Rosenberg JE, Hoffman-Censits J, Powles T, van der Heijden MS, Balar AV, Necchi A, Dawson N, O'Donnell PH, Balmanoukian A, Loriot Y, Srinivas S, Retz MM, Grivas P, et al. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. Lancet. 2016; 387:1909–20.
https://doi.org/10.1016/S0140-6736(16)00561-4
PMID:26952546

35. Nghiem PT, Bhatia S, Lipson EJ, Kudchadkar RR, Miller NJ, Annamalai L, Berry S, Chartash EK, Daud A, Fling SP, Friedlander PA, Kluger HM, Kohrt HE, et al. PD-1 blockade with pembrolizumab in advanced merkel-cell carcinoma. N Engl J Med. 2016; 374:2542–52.
https://doi.org/10.1056/NEJMoa1603702
PMID:27093365

36. Panda A, Mehnert JM, Hirshfield KM, Riedlinger G, Damare S, Saunders T, Kane M, Sokol L, Stein MN, Poplin E, Rodriguez-Rodriguez L, Silk AW, Aisner J, et al. Immune activation and benefit from avelumab in EBV-positive gastric cancer. J Natl Cancer Inst. 2018; 110:316–20.
https://doi.org/10.1093/jnci/djx213
PMID:29155997

37. Fuchs CS, Doi T, Jang RW, Muro K, Satoh T, Machado M, Sun W, Jalal SI, Shah MA, Metges JP, Garrido M, Golan T, Mandala M, et al. Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 clinical KEYNOTE-059 trial. JAMA Oncol. 2018; 4:e180013.
https://doi.org/10.1001/jamaoncol.2018.0013
PMID:29543932

38. Roh W, Chen PL, Reuben A, Spencer CN, Prieto PA, Miller JP, Gopalakrishnan V, Wang F, Cooper ZA, Reddy SM, Gumbs C, Little L, Chang Q, et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. Sci Transl Med. 2017; 9:eaah3560.
https://doi.org/10.1126/scitranslmed.aah3560
PMID:28251903

39. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med. 2004; 10:789–99.
https://doi.org/10.1038/nm1087
PMID:15286780

40. Ferris RL. Immunology and immunotherapy of head and neck cancer. J Clin Oncol. 2015; 33:3293–304.
https://doi.org/10.1200/JCO.2015.61.1509
PMID:26351330

41. Tekpli X, Lien T, Røssevold AH, Nebdal D, Borgen E, Ohnstad HO, Kyte JA, Vallon-Christersson J, Fongaard M, Due EU, Svartdal LG, Sveli MA, Garred Ø, et al, and OSBREAC. An independent poor-prognosis subtype of breast cancer defined by a distinct tumor immune microenvironment. Nat Commun. 2019; 10:5499.
https://doi.org/10.1038/s41467-019-13329-5
PMID:31796750

42. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, Bin J, Liao Y, Rao J, Liao W. Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. Cancer Immunol Res. 2019; 7:737–50. https://doi.org/10.1158/2326-6066.CIR-18-0436 PMID:30842092

43. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, Miller ML, Rekhtman N, Moreira AL, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015; 348:124–28. https://doi.org/10.1126/science.aaa1348 PMID:25765070

44. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet. 2013; 45:970–76. https://doi.org/10.1038/ng.2702 PMID:23852170

45. Vieira VC, Soares MA. The role of cytidine deaminases on innate immune responses against human viral infections. Biomed Res Int. 2013; 2013:683095. https://doi.org/10.1155/2013/683095 PMID:23865062

46. Stavrou S, Ross SR. APOBEC3 proteins in viral immunity. J Immunol. 2015; 195:4565–70. https://doi.org/10.4049/jimmunol.1501504 PMID:26546688

47. Nidheesh N, Abdul Nazeer KA, Ameer PM. An enhanced deterministic k-means clustering algorithm for cancer subtype prediction from gene expression data. Comput Biol Med. 2017; 91:213–21. https://doi.org/10.1016/j.compbiomed.2017.10.014 PMID:29100115

48. Xu T, Le TD, Liu L, Su N, Wang R, Sun B, Colaprico A, Bontempi G, Li J. CancerSubtypes: an r/bioconductor package for molecular cancer subtype identification, validation and visualization. Bioinformatics. 2017; 33:3131–33. https://doi.org/10.1093/bioinformatics/btx378 PMID:28605519

49. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–87. https://doi.org/10.1089/omi.2011.0118 PMID:22455463

50. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst. 2006; 98:262–72. https://doi.org/10.1093/jnci/djj052 PMID:16478745

51. 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 USA. 2005; 102:15545–50. https://doi.org/10.1073/pnas.0506580102 PMID:16199517

52. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013; 9:e1003118. https://doi.org/10.1371/journal.pcbi.1003118 PMID:23950696

53. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. Cell Rep. 2013; 3:246–59. https://doi.org/10.1016/j.celrep.2012.12.008 PMID:23318258

54. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim 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 PMID:21527027

55. Beroukhim R, Getz G, Nghiemphu L, Barretina J, Hsueh T, Linhart D, Vivanco I, Lee JC, Huang JH, Alexander S, Du J, Kau T, Thomas RK, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proc Natl Acad Sci USA. 2007; 104:20007–12. https://doi.org/10.1073/pnas.0710052104 PMID:18077431

56. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, Barretina J, Boehm JS, Dobson J, Urashima M, Mc Henry KT, Pinchback RM, Ligon AH, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010; 463:899–905. https://doi.org/10.1038/nature08822 PMID:20164920

57. Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, Zack T, Laird PW, Onofrio RC, Winckler W, Weir BA, Beroukhim R, Pellman D, Levine DA, et al. Absolute quantification of somatic DNA alterations in human cancer. Nat Biotechnol. 2012; 30:413–21. https://doi.org/10.1038/nbt.2203 PMID:22544022
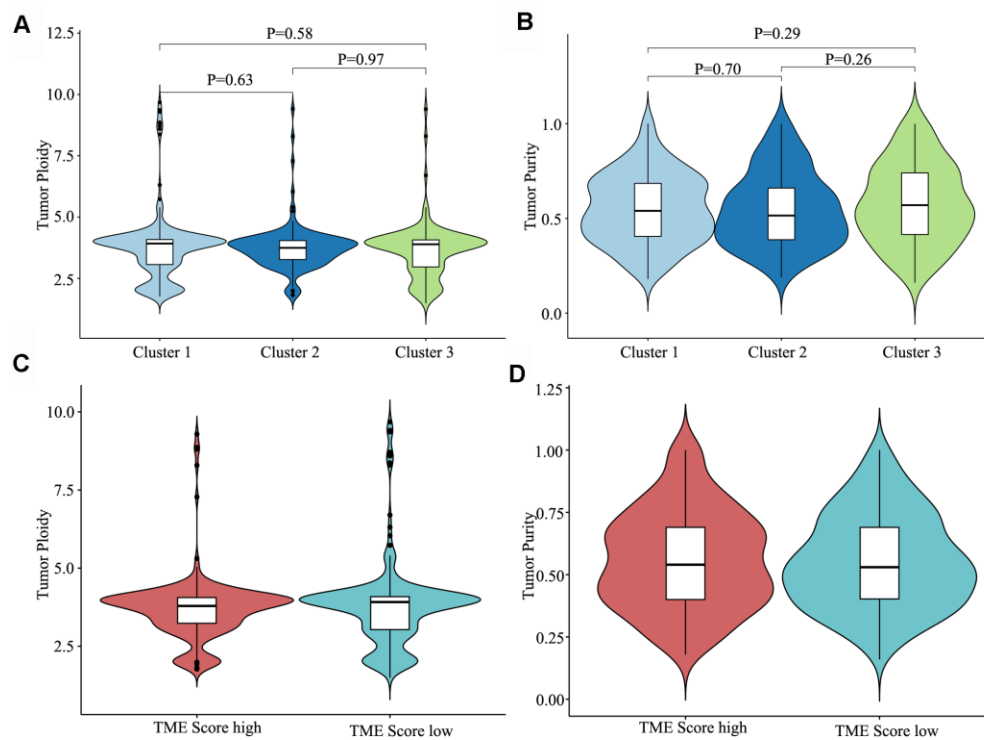
Supplementary Figure 1. Tumor ploidy and malignant cell purity in HNSCC samples. Ploidy (**A**) and purity (**B**) in TME cluster-1, -2 and -3. Ploidy (**C**) and purity (**D**) in the high and low TME score groups.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1 to 5.

Supplementary Table 1. Relationship between different immune cells and survival.

Supplementary Table 2. Differentially expressed genes of intersection in TME cluster-1, -2 and -3.

Supplementary Table 3. Immune-related pathways related to differentially expressed genes.

Supplementary Table 4. Amplification and deletion of chromosomes in the high and low TME score groups.

Supplementary Table 5. Tumor purity and malignant cell ploidy, TMB (tumor mutation burden) and clinical information (age, grade and OS status) pertaining to HNSCC samples.