

## Supplementary Materials

### Methods

#### Patients and specimens

This retrospective translational study based on tumor-infiltrating immune cells (TIIs) quantitation for oral squamous cell carcinoma (OSCC) by using training-testing-validation three independent cohorts. Patients diagnosed as primary OSCC and surgically treated in three independent tertiary referral cancer centers were screened and enrolled. Inclusion criteria for eligible patients were listed as follows, (a) archived samples with adequate tumor components; (b) primary OSCC with no history of prior cancer treatment; (c) patients underwent radical surgery and neck dissection (elective or therapeutic neck dissection as required); (d) epidemiological, clinicopathological and follow-up data available; (e) appropriate patient informed consent. A total number of 392 eligible formalin-fixed paraffin-embedded (FFPE) specimens from three patient cohorts were finally included with the training cohort (130 patients, from Department of Oral and Maxillofacial Surgery, Affiliated Stomatological Hospital, Nanjing Medical University, January 2007-December 2015), the testing cohort (131 patients from Department of Oral and Maxillofacial surgery, First People's Hospital of Lianyungang, January 2008-December 2016), the validation cohort (131 patients from Department of Oral and Maxillofacial surgery, Affiliated Hospital of Jiangnan University, January 2006 to December 2017). Detailed research design and analytical pipeline was elucidated in **Figure 1**. The whole study was reviewed and approved by the Ethic Committee of Affiliated School of Stomatology, Nanjing Medical University, and performed in accordance with the ethical standards of the 1964 Declaration of Helsinki.

#### Immunohistochemical staining (IHC) and quantifications of TIIs in OSCC samples

Upon definitive diagnosis of primary OSCC by senior pathologists based on previous H&E

staining slides, archived samples were screened and included following histopathological confirmation of the existence of typical SCC and invasive margins in each sample. Conventional immunohistochemistry for CD1a, CD3, CD4, CD8, CD11b, CD20, CD45RO, CD57, CD66b, CD163, FOXP3 and PD-1 was performed on 4- $\mu$ m thickness paraffin sections as we previously reported<sup>1 2</sup>. Briefly, paraffin sections were deparaffinized in xylene and rehydrated through graded alcohols. Then, tissue slides were processed in microwave heating in 10mM citrate buffer (pH 6.0) for 15 min for antigen retrieval and 3% H<sub>2</sub>O<sub>2</sub> for endogenous peroxidase inactivation. These sections were incubated with primary antibodies (**Table S8**) at 4°C overnight and secondary antibodies (Maxim, China), further developed with 3,3'-diaminobenzidine, counterstained with hematoxylin, and mounted in non-aqueous mounting media. Negative controls (without primary antibody incubation) were included in each staining run.

These immunohistochemical staining results were independently assessed by two senior oral pathologists who were blinded to the clinical outcome. Localization patterns of CD3<sup>+</sup>, CD4<sup>+</sup>, CD8<sup>+</sup>, CD11b<sup>+</sup>, CD20<sup>+</sup>, CD45RO<sup>+</sup>, CD57<sup>+</sup>, CD66b<sup>+</sup>, CD163<sup>+</sup>, FOXP3<sup>+</sup> and PD-1<sup>+</sup> TILs in samples were divided into two specific areas, center of tumor (CT) and Invasive margin (IM), respectively. As illustrated in **Figure 2A**, the IM was defined as a region of 500 $\mu$ m width on each side of the border between malignant cells and tumor stroma, which was consistent with previous reports<sup>3 4</sup>. At low-power field (100 $\times$ ) under microscope, tissue sections were screened and ten representative fields in each section were selected at 200 $\times$ magnification (0.29 mm<sup>2</sup> per field) under microscope and captured as images using an inverted microscope (Leica, Germany). Densities of these TILs subsets were semi-automatically calculated using ImageJ software (version 2.0) and presented as mean number of cells per field. The X-tile software (version 3.6.1) was utilized to identify the optimal cutoff values for the densities of TILs based on their associations with overall survival (OS)

of patients in the training cohort. Immune densities (cells/mm<sup>2</sup>) of CD3<sup>+</sup>, CD8<sup>+</sup> and CD20<sup>+</sup> TILs were used to calculate the Immunoscore and T and B cell (TB) score as previous reports<sup>4 5</sup>.

### **Prognostic signature based on immune landscape by IHC and machine learning**

Initially, to construct an integrative immune-feature-based prognostic classifier, we first employed the Kaplan-Meier method for OS to select the most useful prognostic features from 24 immune features describing 12 subtypes of TILs in CT and IM in samples from the training cohort. Support vector machine (SVM) algorithm represented as a popular machine learning approach for data classification and function approximation, which trained on a set of labeled patterns to find a hyperplane that divided these samples into two sides<sup>6 7</sup>. Here, we adopted the SVM-recursive feature elimination algorithm to select and rank useful features as previously described<sup>8-10</sup>. In detail, we defined OSCC prognostic prediction as a classification problem where the input was a vector that we called a "pattern" of *n* components which were called "features". The features consisted of the abundances of TILs in CT and IM regions. We limited ourselves to a two-class classification problem (i.e., whether a patient died within 5 years or not). The recursive feature elimination method was applied for features selection and ranking in the training data set. The pruning method was applied to exclude those useless features. We assessed how well an individual feature contributed to the prognostic predication, and then all candidate features were ranked based on their contributions. To assess the contribution of individual feature, we developed a reduced feature set by excluding those useless features. The contribution of each feature was measured based on the performance of SVM trained with this reduced feature set. After ranking, features with the least contribution to the prognostic predication were excluded. The radial basis function kernel was adopted, because our classification was nonlinear. To explore the possibility to stratify different prognostic subgroups of patients based on immune biomarkers using SVM, we developed an

OSCC-SVM classifier in the training cohort and further validated in 2 independent cohorts. Patients on the side of the hyperplane who had better survival were classified into high OSCC-SVM group.

### **Collection and preprocessing of publically available datasets**

The original RNA sequencing data (TPM gene-normalized) of head neck squamous cell carcinoma (HNSCC) samples with corresponding clinical information were downloaded from The Cancer Genome Atlas (TCGA) dataset. The OSCC datasets (328 OSCC samples and 32 normal samples) were further retrieved and enrolled. Another two independent OSCC datasets (GSE41613 and GSE42743) with corresponding clinical information were obtained from the Gene Expression Omnibus (GEO) database<sup>11</sup>. To analyze the efficiency of immunotherapy, the datasets of patients with melanoma treated with anti-PD-1 or anti-CTLA4 therapy (GSE91061, GSE78220) were obtained from GEO database and the datasets of patients with metastatic urothelial cancer treated with anti-PD-L1 agents were retrieved from the R package *IMvigor210CoreBiologies*<sup>12-14</sup>. The “Combat” algorithm was applied to reduce the likelihood of batch effects from non-biological technical biases between different datasets.

### **Estimation of immune cell type fractions and prognostic model development**

Based on marker genes for TILs provided by Bindea et al., single sample Gene Set Enrichment Analysis (ssGSEA) was used to quantify 24 subtypes of TILs, including T lymphocytes, DCs, and NK cells, et al. by using ‘*gsva*’ package<sup>15</sup>. In addition, another deconvolution algorithm named CIBERSORT and LM22 signature matrix were also employed to estimate the proportions of 22 TILs in OSCC. CIBERSORT derives a *P* value for the deconvolution of each sample using Monte Carlo sampling, providing a measure of confidence in the results. Cases with a CIBERSORT output of *P* < 0.05 indicated that the inferred fractions of immune cell populations were accurate<sup>16</sup>. The optimal cutoff values for each TILs subtype from these algorithms were calculated using X-tile software.

Based on these cutoff values, patients were divided into the high-infiltration subgroup and low-infiltration subgroup.

To build a prognostic model from selected immune features, least absolute shrinkage and selection operator (LASSO) regression analyses were performed on quantification data of TILs. The parameters with non-zero coefficients were screened and retained for prognostic model construction. The risk score for each patient was calculated using the formula: Risk score =  $\sum_{i=1}^n \text{coef}_i * x_i$ , where the coef<sub>i</sub> is the coefficient of LASSO regression and x<sub>i</sub> is the enrichment score of each type of immune cell.

### **Consensus clustering of patients based on TILs and characterizations of genomic and transcriptional differences**

Based on the enrichment scores of individual TILs by ssGSEA for each patients, 328 patients with OSCC from TCGA dataset were clustered using 'ConsensusClusterPlus' package<sup>17</sup>. Data about 4 molecular subtypes of OSCC were downloaded from TCGA database. The Alluvial diagram was plotted for patient cluster versus molecular subtypes or pan-SCC subtypes adopted from Dr. Li Ruijing's report using 'ggalluvial' package<sup>18</sup>.

Differentially expressed genes (DEGs) between patient clusters were identified by the following selection criteria ( $|\log_2$  fold change (FC)) $\geq 1$  and false discovery rate (FDR) $< 0.05$ ) using "DESeq2" package and visualized in Heatmaps<sup>19</sup>. The detailed biological functions of DEGs were characterized by Gene ontology (GO) enrichment analyses and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Gene sets of "c2.cp.kegg.v7.1.entrez.gmt", 'c5.all.v7.1.entrez.gmt', and 'h.all.v7.1.entrez.gmt' were downloaded from MSigDB database for running GSEA. All these enrichment analyses were performed by using 'clusterProfiler' package in R<sup>20</sup>. Co-expression clusters were identified by weighted correlation network analysis (WGCNA) R package and

superimposed onto a composite protein-protein interaction (PPI) database of STRING.

The original datasets of mutational profiles in TCGA-OSCC samples were retrieved and processed using “MuTect2 Variant” and “maftools” package<sup>21</sup>. Tumor mutation burden (TMB) was defined as the total number of somatic gene coding errors, base substitutions, and gene insertions or deletions detected across per million bases. Copy number variation (CNV) data were downloaded from GDAC Firehose and subjected to CNV identification using GISTIC 2.0 and visualization by RCircos package<sup>22</sup>. Multiple gene signatures and scores associated with cancer immunity and response to immunotherapy including cytolytic score, IFN- $\gamma$  signature, tertiary lymphoid structures signature score, T-Cell receptor diversity, chemokine gene expression signature, T cell-inflamed gene expression signature and tumor inflammation signature were calculated as previously reported<sup>23-29</sup>.

### Statistical Analyses

Associations between parametric, nonparametric and stratified variables were evaluated using Pearson, Spearman and Fisher's exact test or chi-square test when appropriate. The optimal cutoff values of TILs densities were determined by the X-tile (version 3.6.1) program with the minimum  $P$  value<sup>30</sup>. Overall survival (OS) was defined as the interval between initial surgery and either death or the last follow-up. Disease-free survival (DFS) was calculated as the interval between initial surgery and the presence of local recurrence, metastasis, death or the last follow-up. Rates of survival were calculated by Kaplan-Meier method and compared with log-rank test. SVM was generated using the e1071 and msvmRFE R package. The concordance index (C-index) was generated using the rms R package. Time-dependent receiver operating characteristics (ROC) curves were plotted to define sensitivity and specificity by calculating the area under the curves (AUC) using the timeROC R package. The relative importance of each parameter to survival risk was assessed using the  $\chi^2$

from Harrell's rms R package. Univariate or multivariate regression analyses were performed by Cox proportional hazards regression model. Two tailed  $P$  values  $<0.05$  were considered statistically significant. All statistical analyses were performed with SPSS 24.0 (IBM, Armonk, NY, USA), Graphpad Prism 7 (La Jolla, CA, USA) and R 4.0.1 with diverse packages as indicated.

### Supplementary Figure legends

**Figure S1.** Correlations between two independent examiners to evaluate CD8<sup>+</sup> T cell densities from training (**A,B**), testing (**C,D**) and validation (**E,F**) cohorts.

**Figure S2.** Infiltration patterns of 24 TILs fractions among samples in early (stage I-II) and advanced (stage III-IV) stages from the training cohort.

**Figure S3. Two immune subtypes in OSCC and their associations with survival in patients from the training cohort.**

**(A)** Heatmap showing two immune clusters (immune-hot and immune-cold) defined by 24 immune features consensus expression.

**(B)** Differential representative immune features were observed in the Cluster 1/2 subgroups.

**(C,D)** Kaplan-Meier analyses of OS (**C**) and DFS (**D**) curves based on two immune subtypes in patients from the training cohort. OS: overall survival; DFS: disease-free survival

**Figure S4.** The survival status and profiles of 7 immune features in OSCC patients from training (**A**), testing (**B**) and validation cohorts (**C**) by OSCC-SVM classifier into high and low subgroups.

SVM: support vector machine

**Figure S5.** The Kaplan-Meier analyses of OS in patients from the training cohort stratified by the densities of CD8 IM, CD11b IM, CD11b CT, C20 CT, CD45RO IM, FOXP3 CT and PD-1 CT. OS: overall survival

**Figure S6.** The Kaplan-Meier analyses of DFS in patients from the training cohort stratified by the

densities of CD8 IM, CD11b IM, CD11b CT, CD20 CT, CD45RO IM, FOXP3 CT and PD-1 CT. DFS: disease-free survival

**Figure S7.** The Kaplan-Meier analyses of OS in patients from the testing cohort stratified by the densities of CD8 IM, CD11b IM, CD11b CT, CD20 CT, CD45RO IM, FOXP3 CT and PD-1 CT. OS: overall survival

**Figure S8.** The Kaplan-Meier analyses of DFS in patients from the testing cohort stratified by the densities of CD8 IM, CD11b IM, CD11b CT, CD20 CT, CD45RO IM, FOXP3 CT and PD-1 CT. DFS: disease-free survival

**Figure S9.** The Kaplan-Meier analyses of OS in patients from the validation cohort stratified by the densities of CD8 IM, CD11b IM, CD11b CT, CD20 CT, CD45RO IM, FOXP3 CT and PD-1 CT. OS: overall survival

**Figure S10.** The Kaplan-Meier analyses of DFS in patients from the validation cohort stratified by the densities of CD8 IM, CD11b IM, CD11b CT, CD20 CT, CD45RO IM, FOXP3 CT and PD-1 CT. DFS: disease-free survival

**Figure S11.** The sensitivity and specificity of OSCC-SVM and other clinicopathological parameters in prognostic prediction (OS, the upper panel; DFS, the lower panel) were estimated by 3-year ROC curves in training (**A,D**), testing (**B,E**) and validation cohorts (**C,F**). OS: overall survival; DFS: disease-free survival; ROC: receiver operating characteristics; AUC: area under the curves

**Figure S12.** Relative importance of each parameter to DFS risk using the  $\chi^2$  proportion test for OSCC-SVM and other clinical parameters in training (**A**), testing (**B**) and validation cohorts (**C**). SVM: support vector machine; DFS: disease-free survival

**Figure S13.** Kaplan-Meier analyses of overall survival (OS, upper panel) and disease-free survival (DFS, lower panel) according to OSCC-SVM (**A,B**), Immunoscore (**C,D**) and TB score (**E,F**) in patients from 3 cohorts ( $N=392$ ). SVM: support vector machine; TB: T and B cell

**Figure S14. Profiles of TILs infiltration in OSCC samples estimated by CIBERSORT algorithm.**

**(A)** The proportional bar chart showed the distribution of 22 immune cells in each of TCGA-OSCC sample.

**(B)** The heatmap based on proportions of 22 immune cell types in OSCC and normal samples.

**(C)** Violin plots showing differences of 22 immune cell types in OSCC and normal samples.

**(D)** The abundances of TILs subsets were compared with diverse clinical stages in TCGA-OSCC samples.

**Figure S15. Prognostic models constructed based on TILs inferred via CIBERSORT algorithm.**

**(A)** The coefficient profile plot of 22 immune cells was produced against the log lambda sequence.

**(B)** Tuning parameter (lambda) in the LASSO model was selected via tenfold cross-validation and minimum criteria. Dotted vertical lines were drawn at the optimal values using the minimum criteria and the 1 standard error of the minimum criteria (the 1-SE criteria).

**(C)** The Kaplan-Meier analyses revealed significant associations between 13-immune-features and OS in patients from the training cohort (TCGA-OSCC).

**(D)** The time-dependent ROC curve analyses with 3, 5 years as the defining points were performed to evaluate the predictive values of the 13-immune-features based risk score in the training cohort (TCGA-OSCC).

**(E)** The Kaplan-Meier analyses revealed significant associations between 13-immune-features based risk score and OS in patients from the validation cohort (GES41613 and GSE42743).

**(F)** The time-dependent ROC curve analyses with 3, 5 years as the defining points were performed to evaluate the predictive values of the 13-immune-features based risk score in the validation cohort (GES41613 and GSE42743).

OS: overall survival; ROC: receiver operating characteristics; AUC: area under the curves

**Figure S16.** The ROC curves and AUC compared the performance of TILs-base risk score, PD-L1, PD-1 expression and CTLA4 in prognostic prediction from IMvigor210 **(A)** and GSE78220 **(B)**.

ROC: receiver operating characteristics; AUC: area under the curves

**Figure S17. Two immune subtypes in OSCC and their associations with survival in patients from GSE41613 and GSE42743 cohort.**

**(A)** Consensus clustering matrix for k=2.

**(B)** The Heatmap showing two immune clusters (immune-hot and immune-cold) defined by 24 immune features consensus expression. Tcm cells: central memory T cells; Tem cells: effector memory T cells; Tgd cells: gamma delta T cells; Th1 cells: T helper 1 cells; Th2 cells: T helper 2 cells; DC cells: dendritic cells; pDC: plasmacytoid dendritic cells; iDC: interstitial dendritic cells;

aDC cells: activated dendritic cells; NK cells: natural killer cells; TFH cells: T follicular helper cells;  
TReg cells: regulatory T cells

**(C)** Differential representative immune features were observed in the Cluster 1/2 subgroups.

**(D)** The Kaplan-Meier analyses revealed significant associations between immune subtypes and overall survival (OS) in patients from GSE41613 and GSE42743 cohort.

**Figure S18. Two immune subtypes in OSCC and their associations with survival in patients from GSE65858 cohort.**

**(A)** Consensus clustering matrix for k=2.

**(B)** The Heatmap showing two immune clusters (immune-hot and immune-cold) defined by 24 immune features consensus expression. Tcm cells: central memory T cells; Tem cells: effector memory T cells; Tgd cells: gamma delta T cells; Th1 cells: T helper 1 cells; Th2 cells: T helper 2 cells; DC cells: dendritic cells; pDC: plasmacytoid dendritic cells; iDC: interstitial dendritic cells; aDC cells: activated dendritic cells; NK cells: natural killer cells; TFH cells: T follicular helper cells; TReg cells: regulatory T cells

**(C)** Differential representative immune features were observed in the Cluster 1/2 subgroups.

**(D)** The Kaplan-Meier analyses revealed significant associations between immune subtypes and OS/PFS in patients from GSE65858 cohort. OS: overall survival; PFS: progression-free survival

**Figure S19.** Alluvial diagram revealed associations between our two immune subtypes in OSCC with previously reported immune types in pan-SCC or pan-cancer. IS: immune subtype; C1: wound healing; C2: IFN- $\gamma$  dominant; C3: inflammatory; C4: lymphocyte depleted; C5: immunologically quiet; C6: TGF- $\beta$  dominant

**Figure S20. The GO and KEGG analyses of differentially expressed genes between two immune subtypes.**

**(A,B)** KEGG and GO analyses for the up-regulated genes in Cluster 1 (immune-hot).

**(C,D)** KEGG and GO analyses for the down-regulated genes in Cluster 2 (immune-cold).

**Figure S21.** Results from GSEA analyses for the differentially expressed genes between Cluster 1 (immune-hot) and Cluster 2 (immune-cold).

**Figure S22. PPI network for differentially expressed genes between two immune subtypes.**

**(A)** Diagram of the individual modules and their interactions. The circle size proportional to module size and top 6 modules were selected.

**(B-D)** Module 1 (**B**), Module 4 (**C**) and Module 6 (**D**) were further expanded and KEGG analysis were performed for each module.

**Figure S23. Spectrum of genetic mutations in two immune clusters of OSCC.**

**(A-B)** The overall summary of mutation information in Cluster 1 (**A**, immune-hot) and Cluster 2 (**B**, immune-cold).

**(C-D)** The oncoplot showed the top ten mutated genes and their mutation information in Cluster 1 (**C**, immune-hot) and Cluster 2 (**D**, immune-cold).

**Figure S24. RCircos plot illustrated associations between differentially expressed genes in Cluster 1 and Cluster 2 and copy number variation for chromosome ideogram.**

A chromosome ideogram was shown in the outer ring. The associations between differentially expressed genes in Cluster 1 (immune-hot) and Cluster 2 (immune-cold) and copy number variation were represented in the next outermost ring. Red represented the expression of genes which were positively correlated with copy number variation and blue represented the expression of genes which were negatively correlated with copy number variation. The next ring represented the number of amplified copy number or deleted copy number (red indicated amplified; green indicated deleted). The innermost ring represent the difference of gene expression level between Cluster 1 and Cluster 2 (red indicated Up-regulated genes in Cluster 1(LogFC>1), green indicated Down-regulated genes in Cluster 1(LogFC< -1)). FC: fold change

**Figure S25. Distributions of multiple immune-related signatures in two immune subtypes were compared.**

**(A)** TMB difference between two clusters was compared. TMB: Tumor mutation burden

**(B-H)** The Violin plots showed distributions of multiple immune-related signatures in both clusters, including cytolytic score **(B)**, T-Cell receptor diversity **(C)**, T cell-inflamed gene expression profile signature **(D)**, Tumor inflammation signature **(E)**, chemokine gene expression signature **(F)**, IFN- $\gamma$  signature **(G)**, Tertiary lymphoid structures signature **(H)**.

**Figure S26.** The heatmap showed the expression of major histocompatibility complex (MHC), immunomodulatory molecules and markers of cytolytic activities in both clusters. The expression values were represented by Z scores calculated.

## References

- 1 Zhou C, Diao P, Wu Y, *et al.* Development and validation of a seven-immune-feature-based prognostic score for oral squamous cell carcinoma after curative resection. *Int J Cancer* 2020;146:1152-1163.
- 2 Hu Y, He MY, Zhu LF, *et al.* Tumor-associated macrophages correlate with the clinicopathological features and poor outcomes via inducing epithelial to mesenchymal transition in oral squamous cell carcinoma. *J Exp Clin Cancer Res* 2016;35:12.
- 3 Halama N, Michel S, Kloor M, *et al.* Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. *Cancer Res* 2011;71:5670-5677.
- 4 Pages F, Mlecnik B, Marliot F, *et al.* International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* 2018;391:2128-2139.
- 5 Mlecnik B, Van den Eynde M, Bindea G, *et al.* Comprehensive Intrametastatic Immune Quantification and Major Impact of Immunoscore on Survival. *J Natl Cancer Inst* 2018;110.
- 6 Vapnik VN. An overview of statistical learning theory. *IEEE Trans Neural Netw* 1999;10:988-999.
- 7 Wang HY, Sun BY, Zhu ZH, *et al.* Eight-signature classifier for prediction of nasopharyngeal [corrected] carcinoma survival. *J Clin Oncol* 2011;29:4516-4525.
- 8 Jiang Y, Xie J, Han Z, *et al.* Immunomarker Support Vector Machine Classifier for Prediction of Gastric Cancer Survival and Adjuvant Chemotherapeutic Benefit. *Clin Cancer Res*

- 2018;24:5574-5584.
- 9 Zhu ZH, Sun BY, Ma Y, *et al.* Three immunomarker support vector machines-based prognostic classifiers for stage IB non-small-cell lung cancer. *J Clin Oncol* 2009;27:1091-1099.
  - 10 Jiang Y, Xie J, Huang W, *et al.* Tumor Immune Microenvironment and Chemosensitivity Signature for Predicting Response to Chemotherapy in Gastric Cancer. *Cancer Immunol Res* 2019;7:2065-2073.
  - 11 Lohavanichbutr P, Mendez E, Holsinger FC, *et al.* A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. *Clin Cancer Res* 2013;19:1197-1203.
  - 12 Riaz N, Havel JJ, Makarov V, *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* 2017;171:934-949.e16.
  - 13 Hugo W, Zaretsky JM, Sun L, *et al.* Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* 2016;165:35-44.
  - 14 Mariathasan S, Turley SJ, Nickles D, *et al.* TGFbeta attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554:544-548.
  - 15 Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013;14:7.
  - 16 Newman AM, Liu CL, Green MR, *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;12:453-457.
  - 17 Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 2010;26:1572-1573.
  - 18 Li B, Cui Y, Nambiar DK, *et al.* The Immune Subtypes and Landscape of Squamous Cell

- Carcinoma. *Clin Cancer Res* 2019;25:3528-3537.
- 19 Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;11:R106.
- 20 Yu G, Wang LG, Han Y, *et al.* clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284-287.
- 21 Mayakonda A, Lin DC, Assenov Y, *et al.* Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;28:1747-1756.
- 22 Mermel CH, Schumacher SE, Hill B, *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* 2011;12:R41.
- 23 Dufva O, Polonen P, Bruck O, *et al.* Immunogenomic Landscape of Hematological Malignancies. *Cancer Cell* 2020;38:424-428.
- 24 Ayers M, Lunceford J, Nebozhyn M, *et al.* IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;127:2930-2940.
- 25 Cabrita R, Lauss M, Sanna A, *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* 2020;577:561-565.
- 26 Thorsson V, Gibbs DL, Brown SD, *et al.* The Immune Landscape of Cancer. *Immunity* 2018;48:812-830.e14.
- 27 Saloura V, Izumchenko E, Zuo Z, *et al.* Immune profiles in primary squamous cell carcinoma of the head and neck. *Oral Oncol* 2019;96:77-88.
- 28 Cristescu R, Mogg R, Ayers M, *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018;362:eaar3593.
- 29 DanaHER P, Warren S, Lu R, *et al.* Pan-cancer adaptive immune resistance as defined by the

Tumor Inflammation Signature (TIS): results from The Cancer Genome Atlas (TCGA). *J Immunother Cancer* 2018;6:63.

- 30 Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* 2004;10:7252-7259.