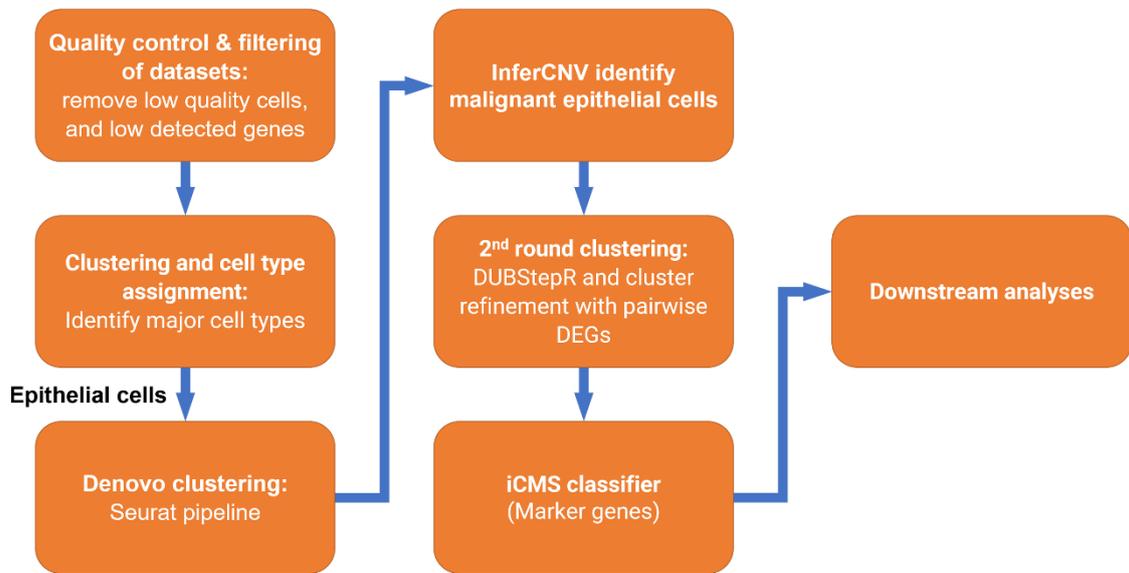
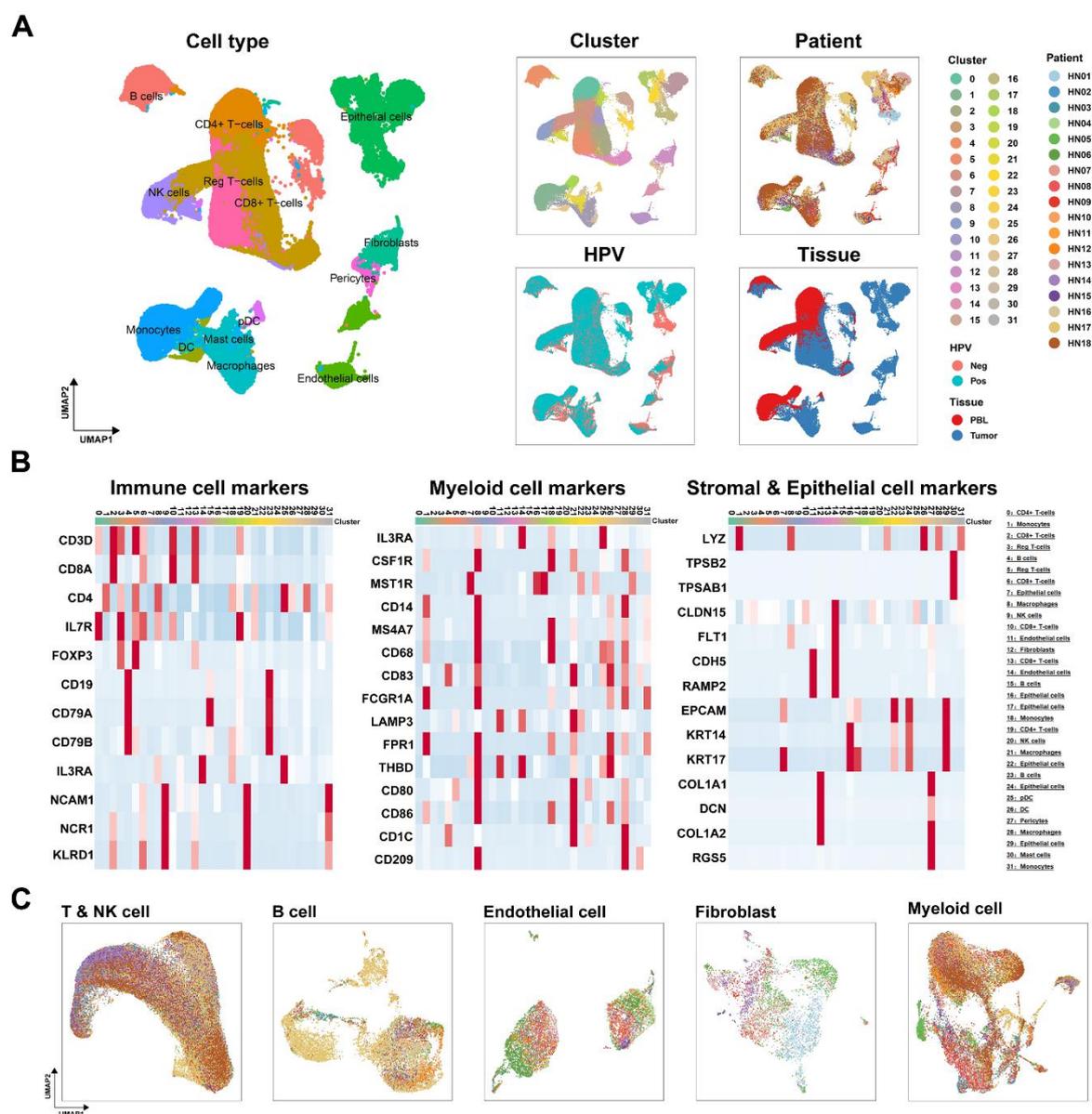


Supplementary Figure 1



Supplementary Figure 1. Flowchart for data collection and quality control.

Supplementary Figure 2



Supplementary Figure 2. Identification of major cell types in HNSCC samples.

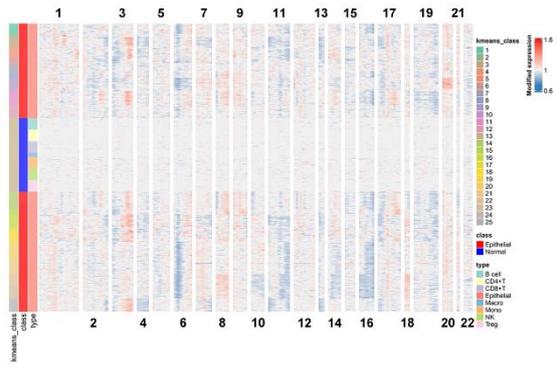
A. The major cell types were identified and annotated by clusters, patients, HPV (negative, positive), and tissues (PBL, tumor).

B. Heatmap depicted the expression of marker genes used for cell type annotations.

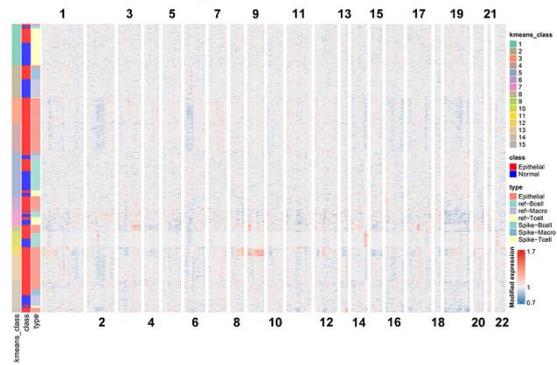
C. Major cell types were annotated by patients.

HPV, human papillomavirus; PBL, peripheral blood lymphocytes.

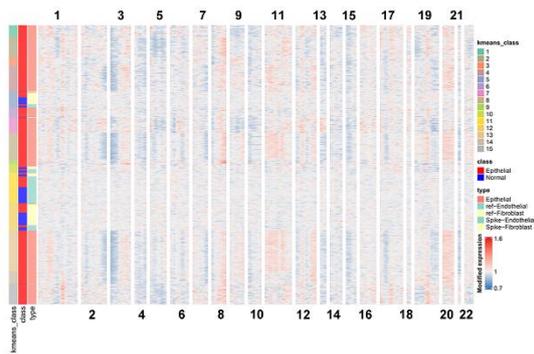
Supplementary Figure 3 GSE164690



SNOW study

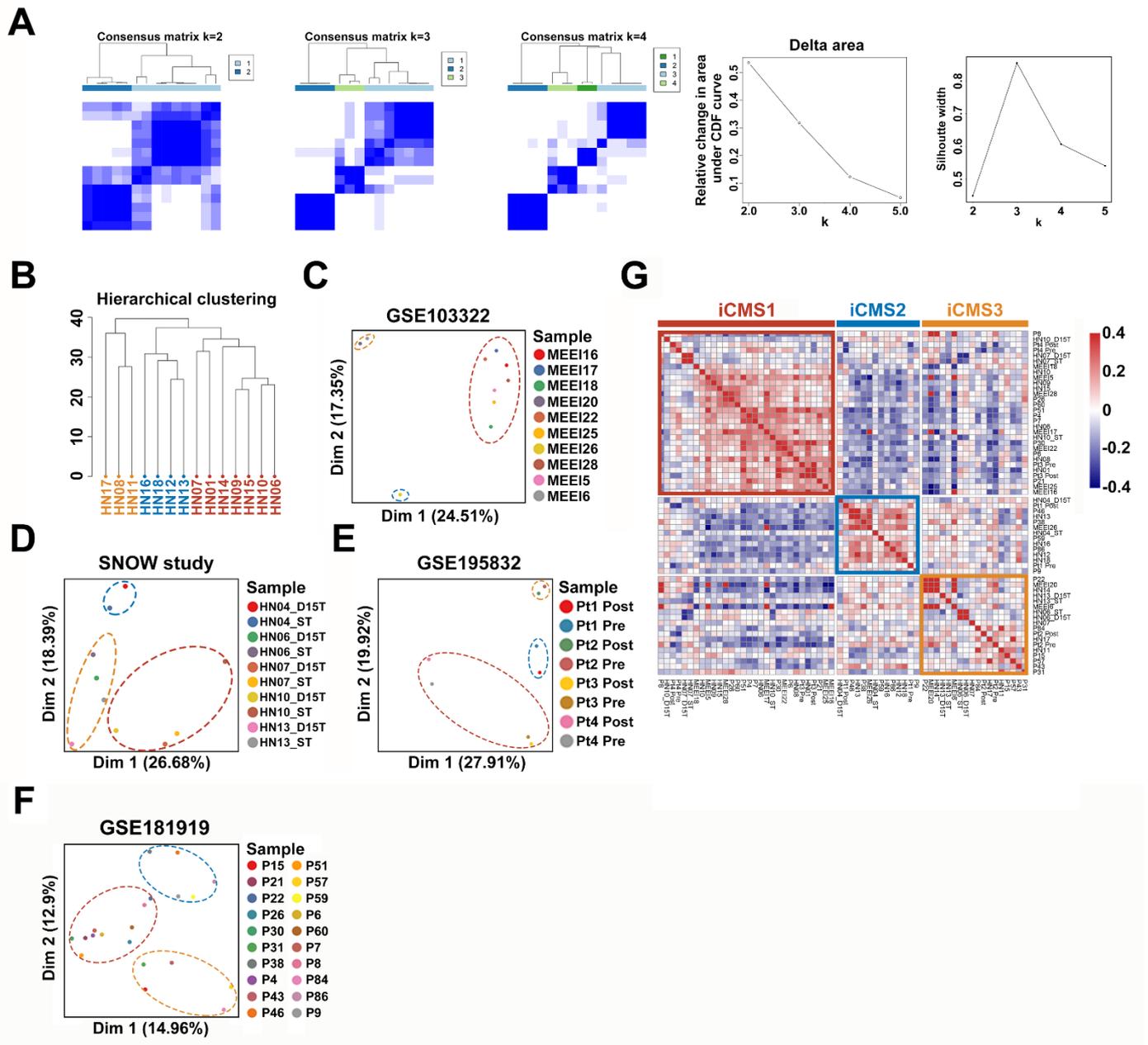


GSE195832



Supplementary Figure 3. Identification of malignant epithelial cells in scRNA-seq datasets by inferring CNV.

Supplementary Figure 4



Supplementary Figure 4. Sub-clustering of malignant epithelial cells from different cohorts.

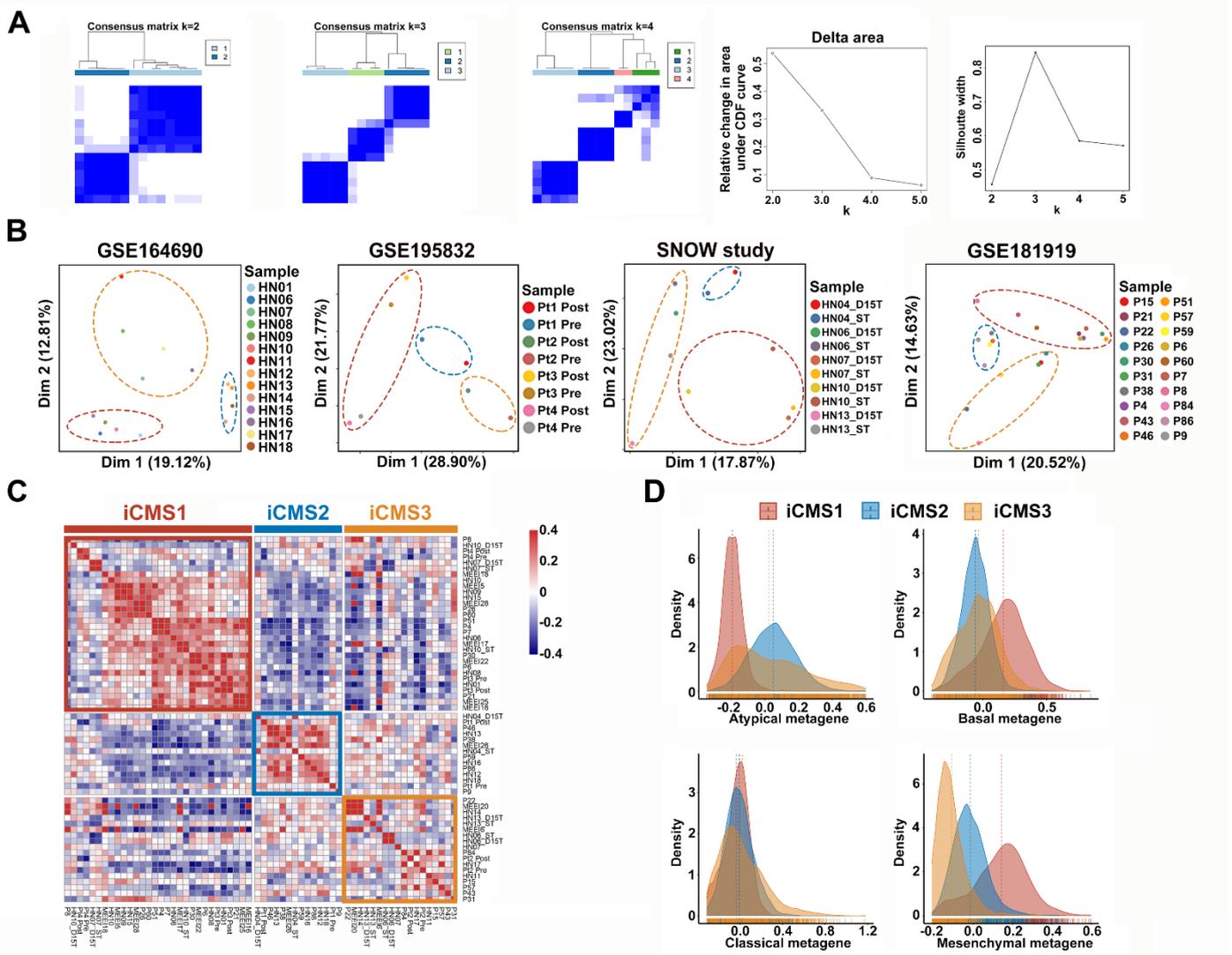
A. Consensus clustering of 14 patient samples from GSE164690 identified optimal k for classification using iCMS classifier genes.

B. Hierarchical clustering of 14 patient samples from GSE164690.

C-F. PCA of epithelial patient-specific pseudobulk transcriptomes for each cohort (**C.** GSE103322, **D.** SNOW study, **E.** GSE195832, **F.** GSE181919).

G. Correlation heatmap indicated the highly similarity between samples in same iCMS subtype using pooled cohorts (Spearman's correlation, r value).

Supplementary Figure 5



Supplementary Figure 5. Reproducing of iCMS classification in independent cohorts.

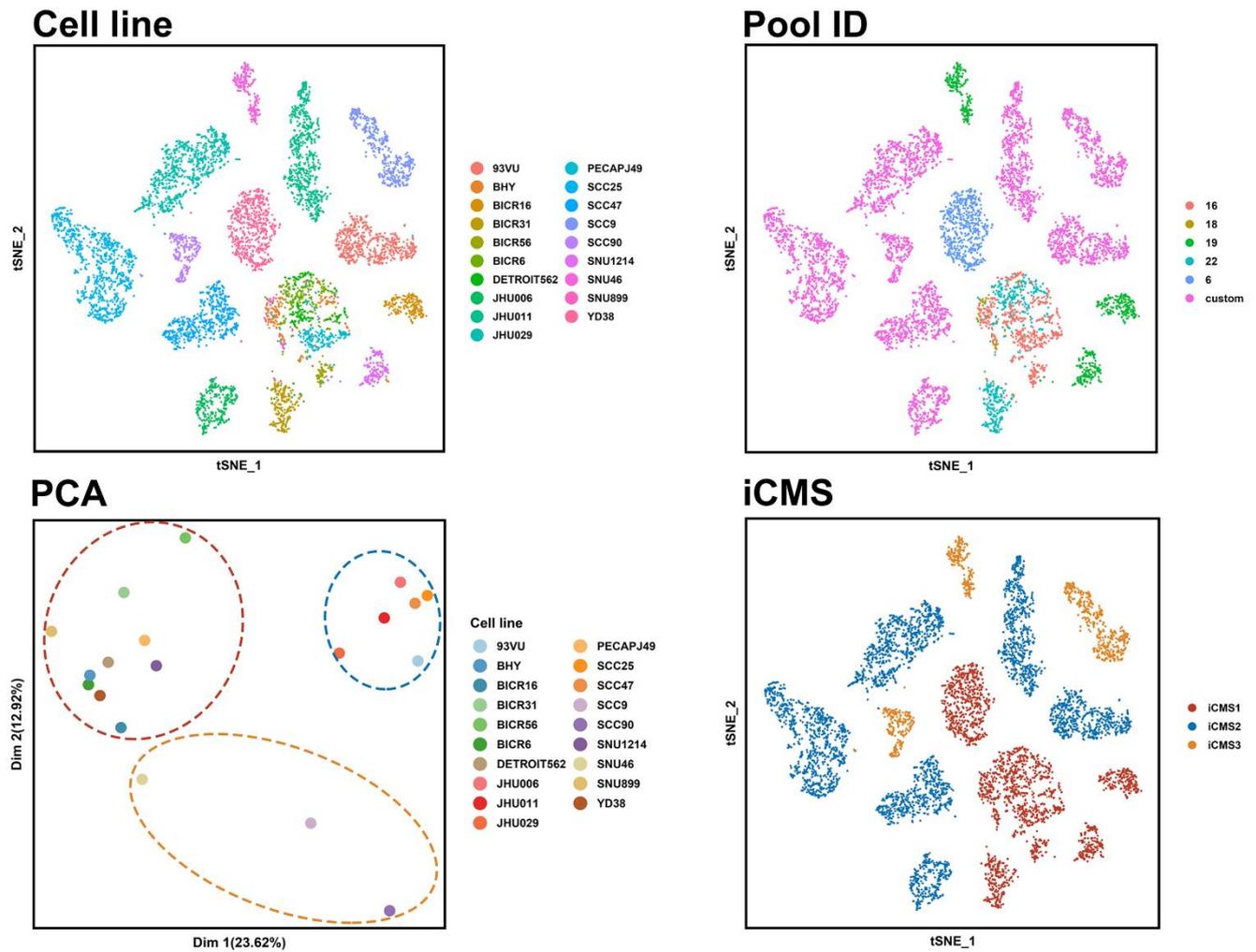
A. Consensus clustering of 14 patient samples from GSE164690 identified optimal k for classification using genes identified from GSE103322.

B. PCA of epithelial patient-specific pseudo-bulk transcriptomes for each cohort using the same gene set identified from GSE103322.

C. Correlation heatmap indicated the highly similarity between samples in same iCMS subtype using pooled cohorts (Spearman's correlation, r value).

D. Associations between iCMS and canonical 4 transcriptional subtypes of HNSCC were shown. The metagene scores for each cell were calculated by averaging the scaled expressions within the same transcriptional subtype.

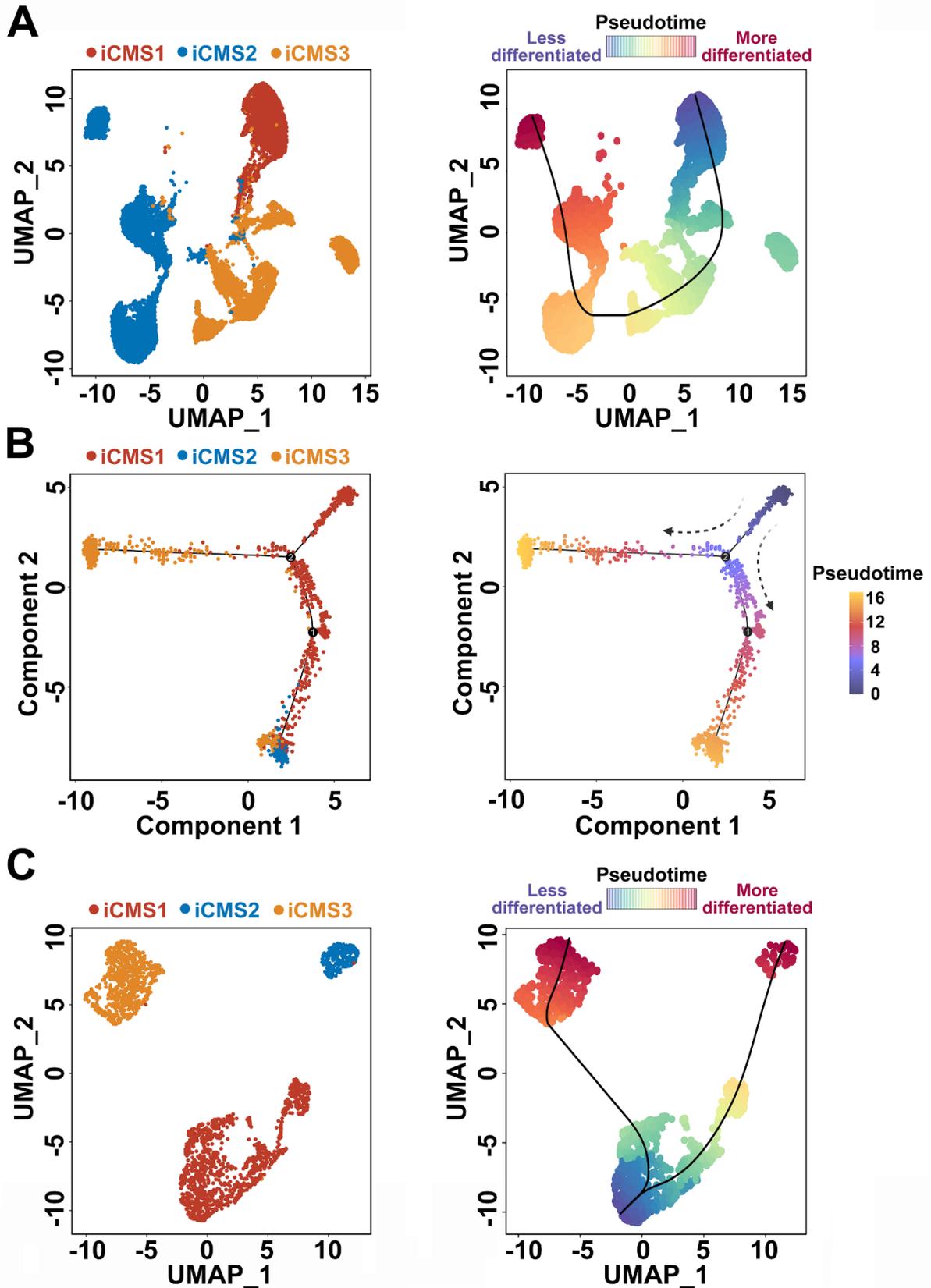
Supplementary Figure 6



Supplementary Figure 6. The iCMS classification is recapitulated in HNSCC cell lines.

Cells were annotated by cell line, pool ID and iCMS, which were classified by PCA.

Supplementary Figure 7



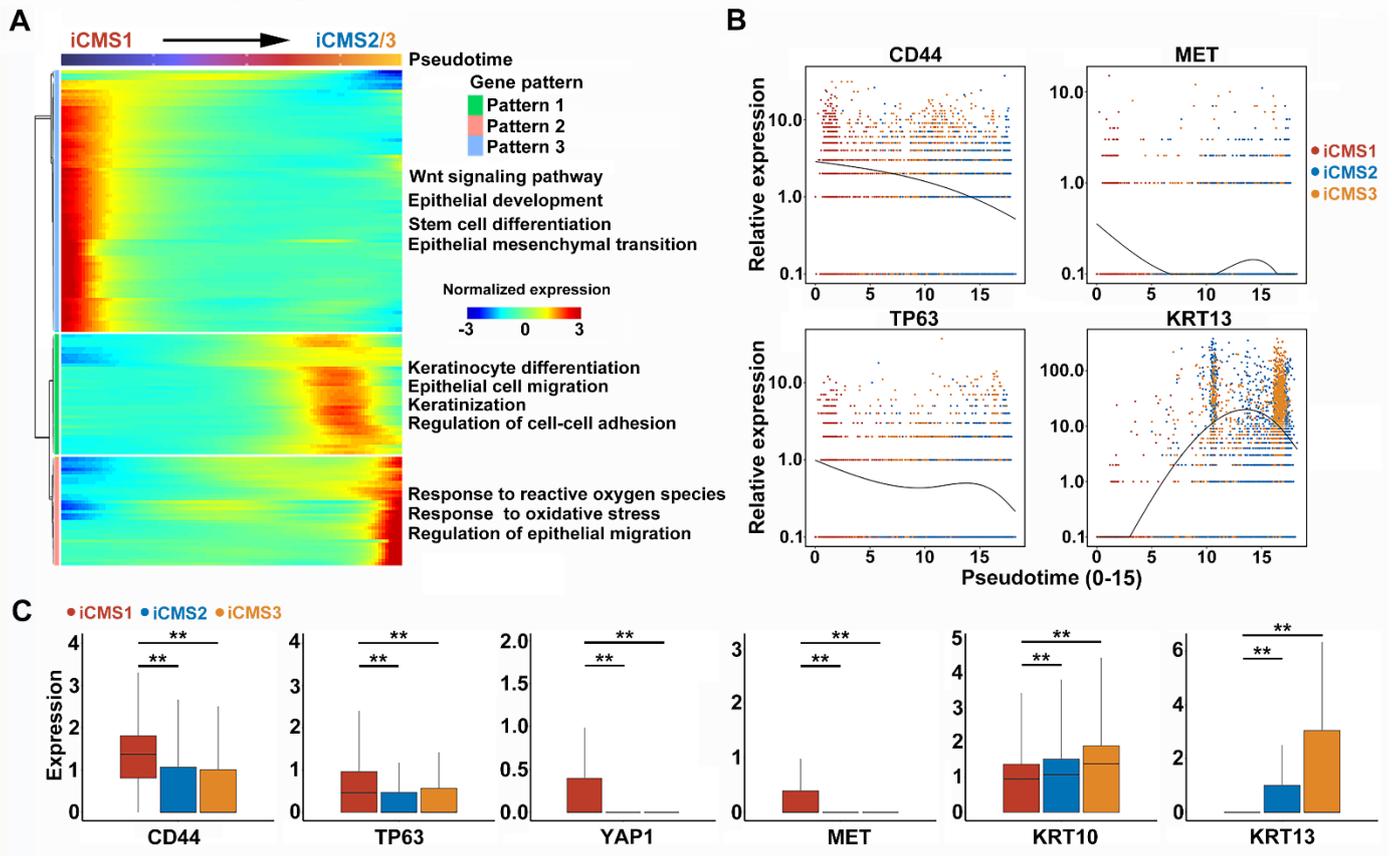
Supplementary Figure 7. Construction of malignant epithelial cell differentiation trajectory by pseudotime analyses.

A. Pseudotime of iCMS malignant epithelial cells in GSE164690 inferred by Slingshot. Left panel,

UMAP representation of iCMS subgroups. Right panel, pseudotime trajectory was plotted and cells were colored by pseudotime.

B, C. Pseudotime of iCMS malignant epithelial cells in GSE103322 inferred by Monocle2 (**B**) and Slingshot (**C**). Left panel, cells colored by iCMS subtypes. Right panel, pseudotime trajectories were plotted and cells were colored by pseudotime.

Supplementary Figure 8



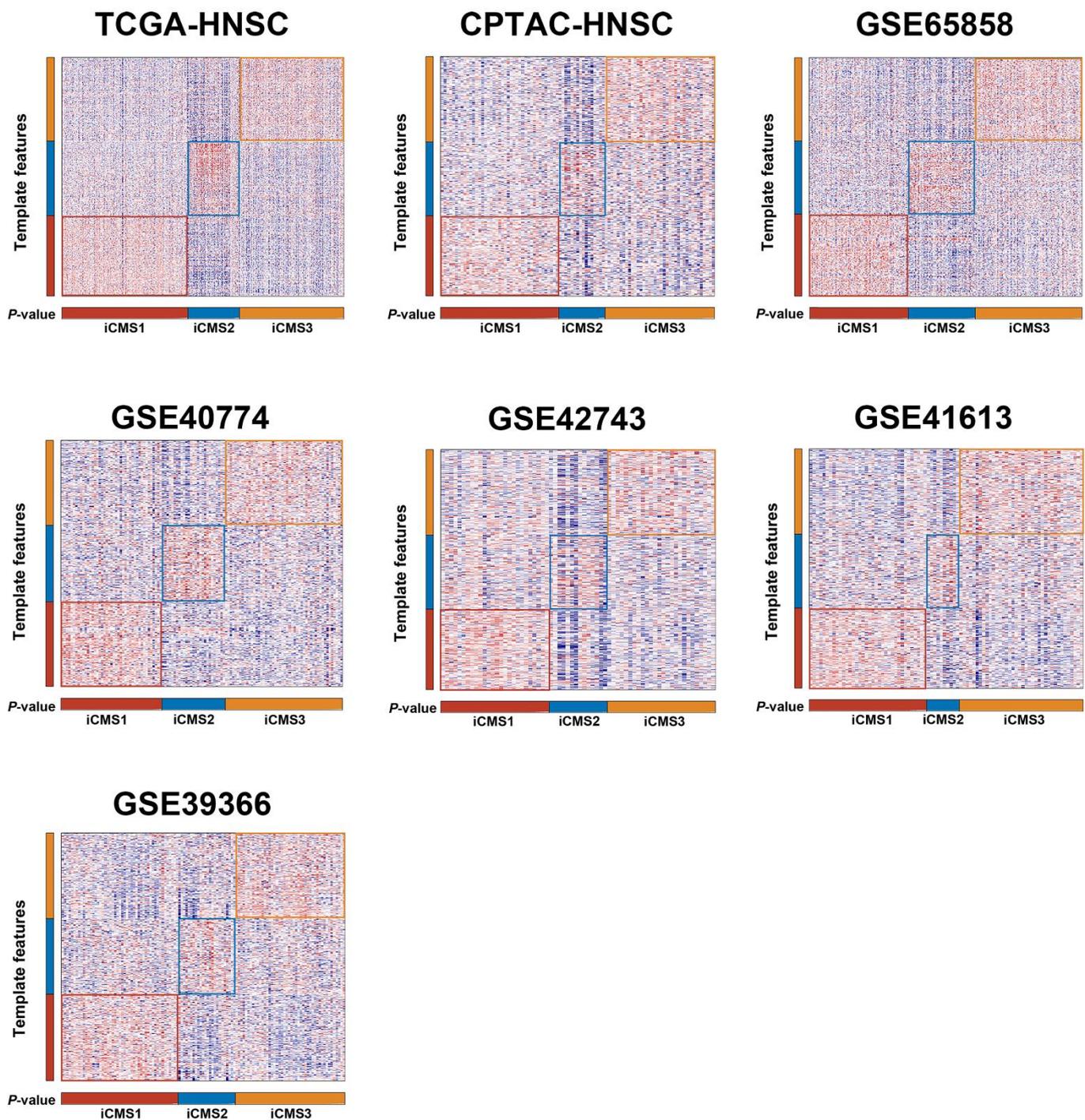
Supplementary Figure 8. Differential gene expression profiles along pseudotime trajectory.

A. Differentially expressed genes along pseudotime trajectory were clustered and functionally annotated.

B. Loess regression-smoothed genes expression of stemness (TP63, CD44, MET) and epithelial differentiation (KRT13) markers along the pseudotime trajectory.

C. Differences in expression patterns of stemness (TP63, CD44, MET, YAP1) and epithelial differentiation (KRT10, KRT13) markers among iCMS malignant epithelial subsets by Wilcoxon rank-sum test.

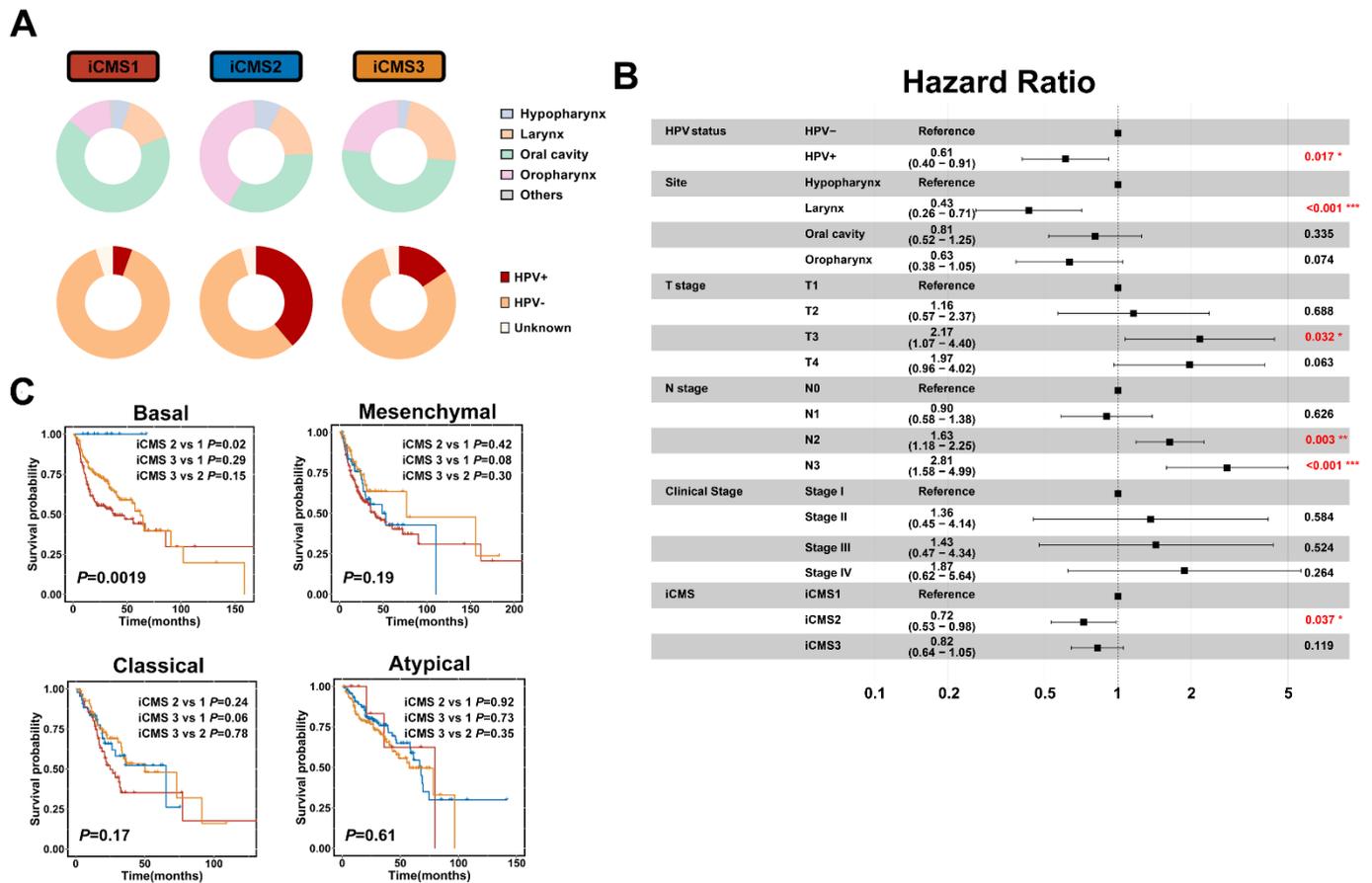
Supplementary Figure 9



Supplementary Figure 9. The iCMS classification is applied to multiple HNSCC bulk transcriptomes.

Correlation matrix displayed the Pearson correlation significance (P -value) calculated across the HNSCC samples in each indicated bulk transcriptomes dataset (TCGA-HNSC, CPTAC-HNSC, GSE65858, GSE40774, GSE42743, GSE41613 and GSE39366, respectively).

Supplementary Figure 10



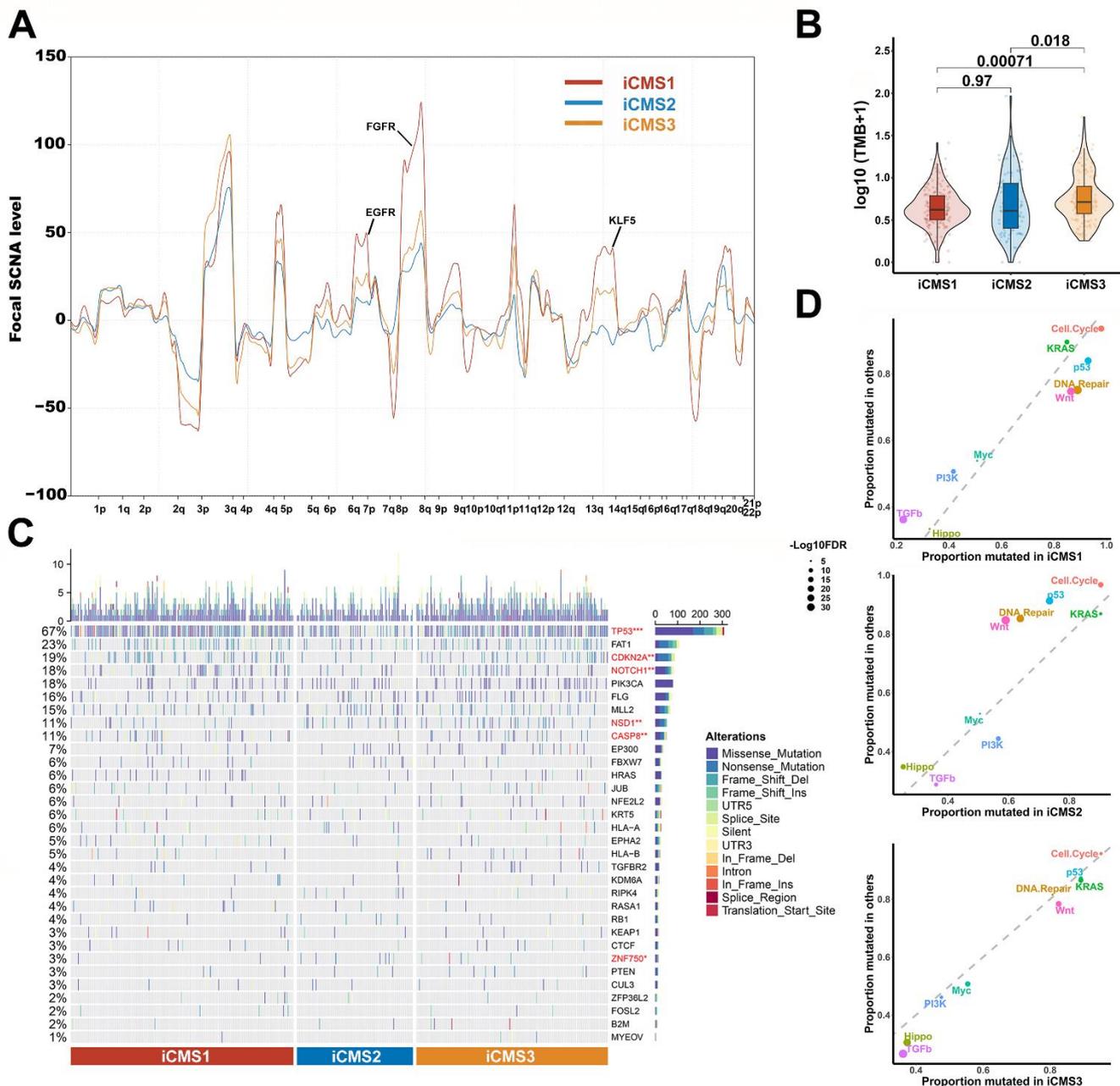
Supplementary Figure 10. The iCMS subtype is identified as an independent prognostic factor.

A. Distributions of anatomical sites and HPV status among iCMS patient subpopulations in TCGA-HNSC dataset.

B. Multivariate Cox regression analysis identified iCMS subtype was identified as an independent prognostic factor after adjusting for HPV status, anatomical site, clinical stage, and other clinicopathological parameters.

C. Kaplan-Meier curves showed that among these canonical transcriptional subtypes, patients were further categorized into subgroups with superior or inferior survival (Log-rank test).

Supplementary Figure 11



Supplementary Figure 11. Genomic features of iCMS.

A. CNV in chromosome arm-level among three iCMS subtypes were plotted and representative genes involved in these CNV regions were highlighted.

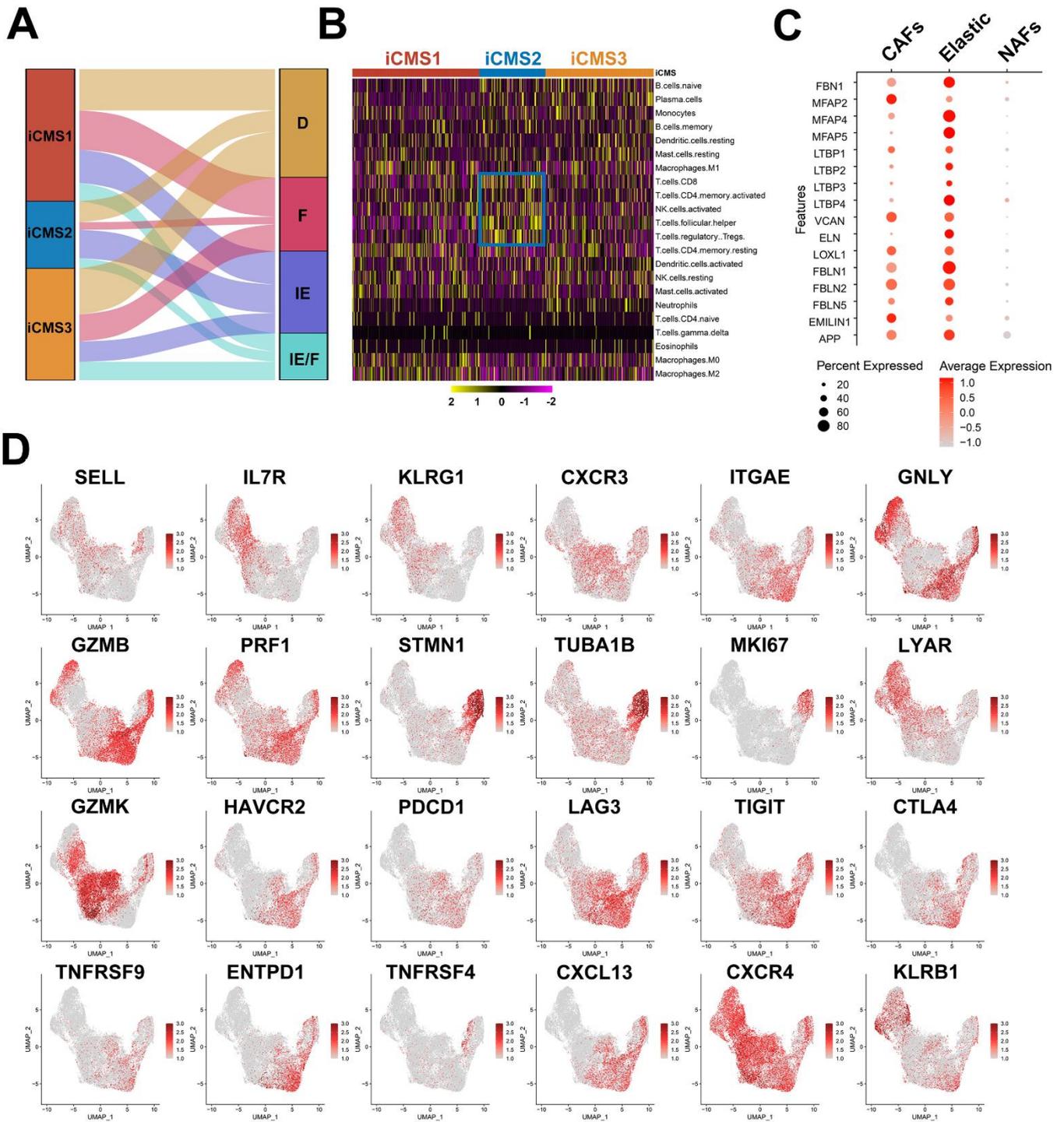
B. TMB was calculated and compared among iCMS1-3 (Wilcoxon rank-sum test).

C. Detailed distributions of driver mutations among iCMS1-3 in TCGA-HNSC dataset were shown and compared.

D. Scatter plots of ratios of TCGA-HNSC samples with mutations in selected pathways. Dot size corresponds to FDR by two-sided Chi-square test with BH-correction.

CNV, copy number variation; TMB, tumor mutational burden.

Supplementary Figure 12



Supplementary Figure 12. Characterization of TME among iCMS.

A. Sankey plot showed the relationship between iCMS subtypes and pan-cancer TME subtype (MFP) defined by Alexander Bagaev in 2021 (1). D: Desert; F: Fibrotic; IE: Immune-Enriched Non-Fibrotic; IE/F: Immune-Enriched Fibrotic.

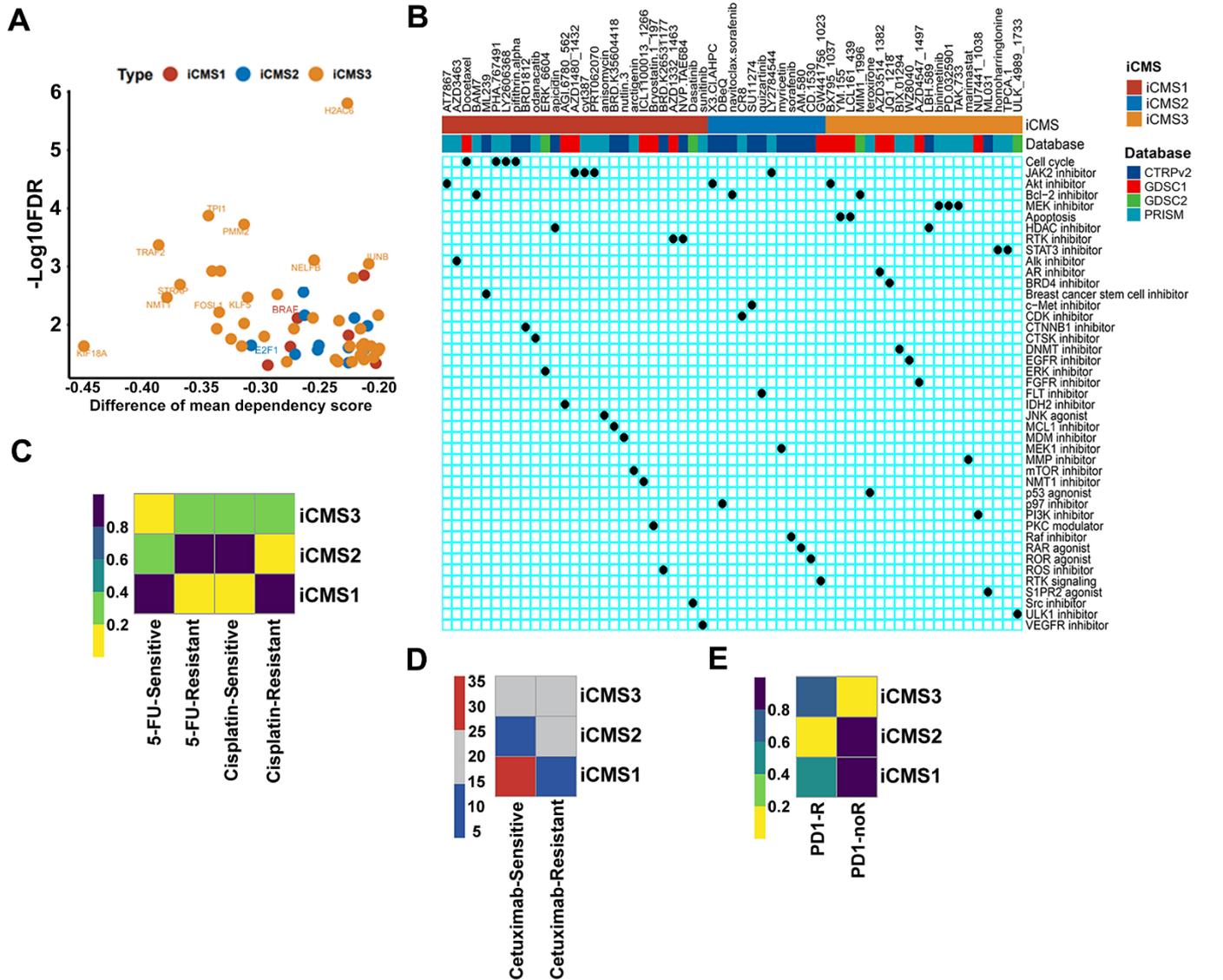
B. Heatmap depicted the immune infiltration among iCMS subtypes calculated by xCell.

C. Dot plot displayed the marker genes of fibroblast subpopulations.

D. Marker genes of CD8+ T-cells were shown.

TME, tumor microenvironment

Supplementary Figure 13



Supplementary Figure 13. Drug repurposing analyses based on iCMS.

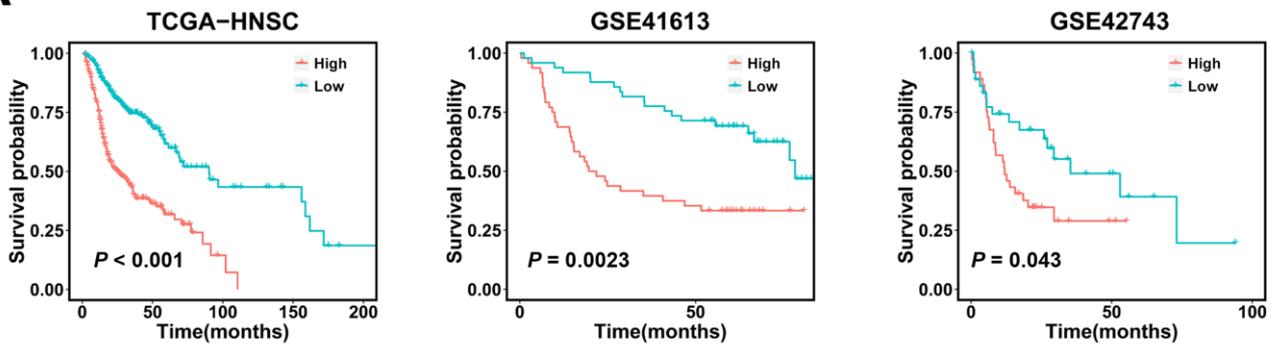
A. Dot plot indicated gene dependencies among iCMS subtypes. Only genes with FDR < 0.05 were shown.

B. Four drug sensitivity databases (CTRPv2, GDSC1, GDSC2, PRISM) were used to discover therapeutic vulnerabilities in iCMS. Agents from these four databases were pooled and annotated by their targeted pathways.

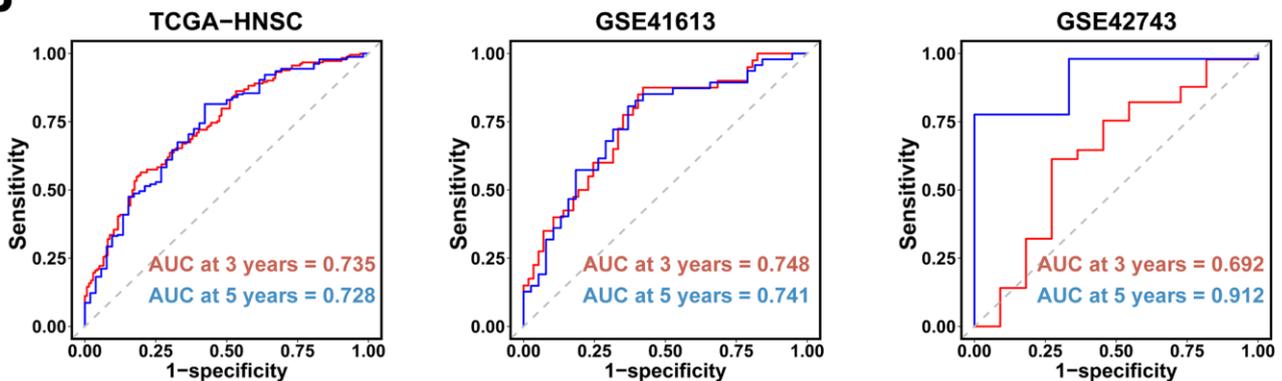
C-E. The responsiveness to chemotherapy (5-FU and Cisplatin, **C**), targeted therapy (cetuximab, **D**) and immunotherapy (anti-PD-1, **E**) were compared among iCMS by SubMap analyses.

Supplementary Figure 14

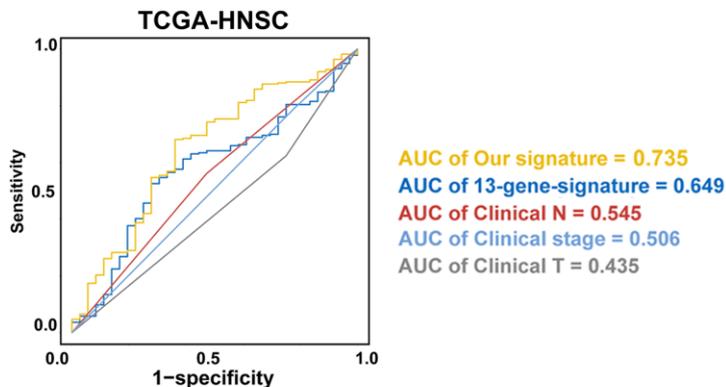
A



B



C



Supplementary Figure 14. Development and validation of a novel prognostic signature.

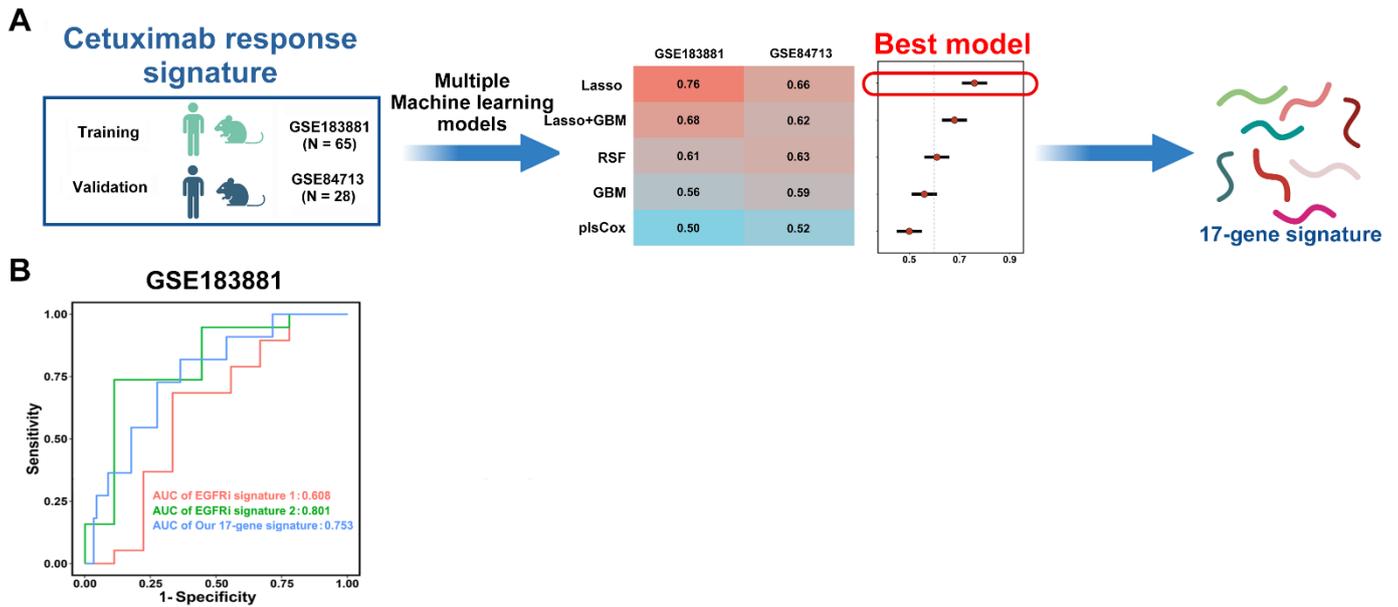
A. Kaplan-Meier plots of our 7-gene prognostic signature in training (TCGA-HNSC), validation (GSE41613, GSE42743, Log-rank test).

B. The predictive performance of this 7-gene signature was assessed by ROC analyses with end points at 3-year and 5-year in three HNSCC cohorts (TCGA-HNSC, GSE42743 and GSE41613), respectively.

C. ROC analyses compared the performance of our 7-gene prognostic signature, selected clinicopathological parameters and a previously published 13-gene signature in TCGA-HNSC dataset (2).

ROC, receiver operating characteristic.

Supplementary Figure 15

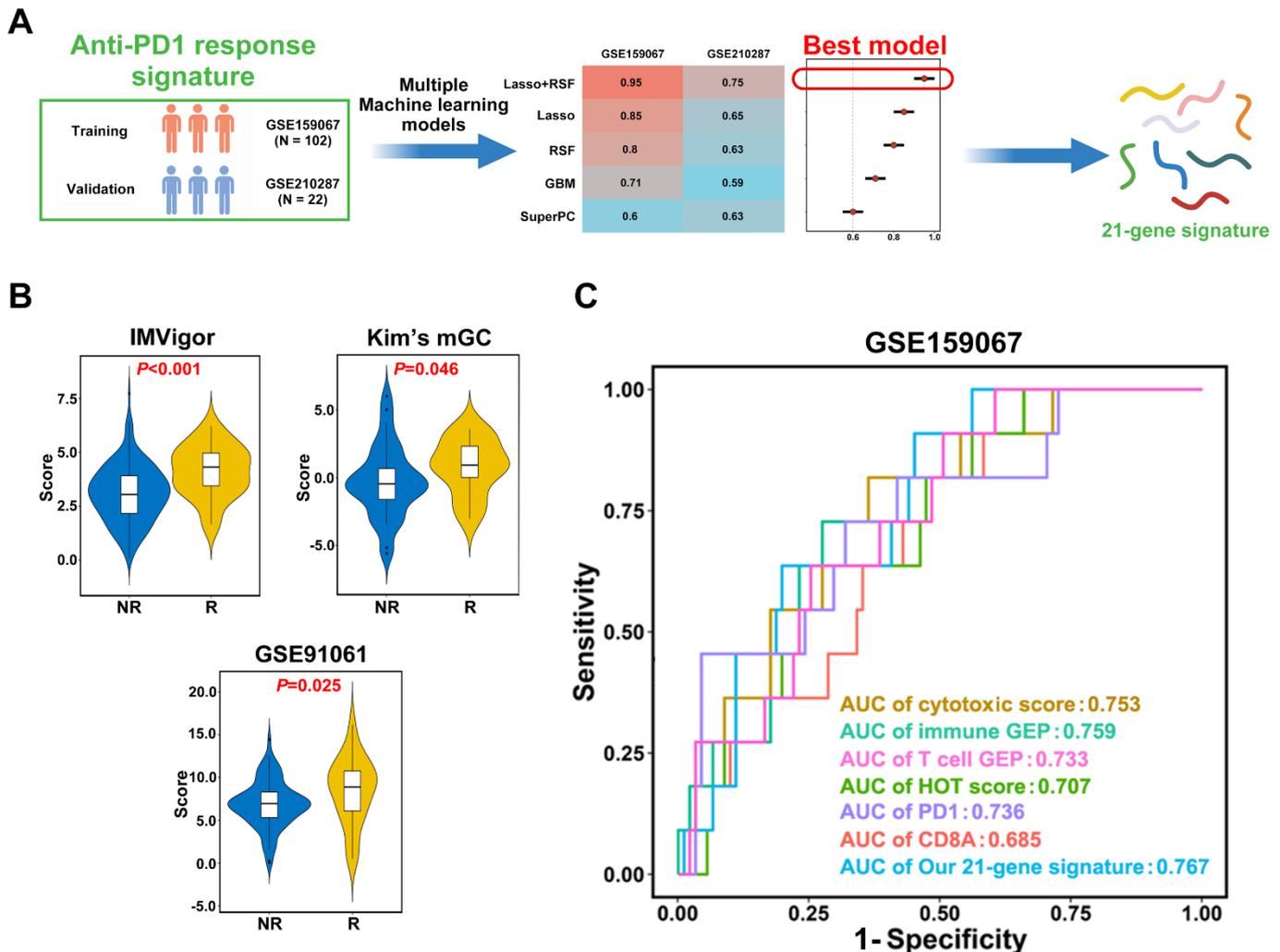


Supplementary Figure 15. Development of a novel 17-gene signature for Cetuximab treatment response.

A. Flowchart of analytic pipeline used for signature construction by machine learning. Ten machine learning algorithms (Lasso, GBM, Ridge, survivalSVM, superPC, plsRcox, CoxBoost, Enet, StepCox, RSF) were utilized and the optimal model was determined based on C-Index values in each predictive model.

B. ROC analyses compared the performance of our 17-gene signature and two previously published gene signatures for EGFRi sensitivity in GSE183881 dataset. EGFRi signature 1: Goodspeed A, et al (3); EGFRi signature 2: Byers LA, et al (4).

Supplementary Figure 16



Supplementary Figure 16. Development of a novel 21-gene signature for anti-PD-1 treatment response.

A. Flowchart of analytic pipeline used for signature construction by multiple machine learning approaches.

B. This 21-gene signature was applicable in bladder (IMVigor210, Powles T, et al.) (5), gastric (Kim's mGC, Kim ST, et al.) (6) cancer and melanoma (GSE91061, Riaz N, et al.) (7) patients treated with anti-PD-1 into response or non-response subgroups. Detailed patient treatment response was retrieved from original papers and compared with Wilcoxon rank-sum test test.

C. ROC analyses compared the performance of our 21-gene signature, 4 previously published gene signatures and PD1/CD8A expression for anti-PD-1 sensitivity in GSE159076 dataset. Cytotoxic score: Rooney MS, et al (8); immune GEP: Roh W, et al (9); T cell GEP: Huang AC, et al (10); HOT

score: Foy JP, et al(11).

Reference

1. Bagaev A, Kotlov N, Nomie K, Svekolkina V, Gafurov A, Isaeva O, *et al.* Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* 2021;**39**(6):845-65 e7 doi 10.1016/j.ccell.2021.04.014.
2. Lohavanichbutr P, Mendez E, Holsinger FC, Rue TC, Zhang Y, Houck J, *et al.* A 13-gene signature prognostic of HPV-negative OSCC: discovery and external validation. *Clin Cancer Res* 2013;**19**(5):1197-203 doi 10.1158/1078-0432.CCR-12-2647.
3. Goodspeed A, Jean A, Theodorescu D, Costello JC. A Gene Expression Signature Predicts Bladder Cancer Cell Line Sensitivity to EGFR Inhibition. *Bladder Cancer* 2018;**4**(3):269-82 doi 10.3233/BLC-170161.
4. Byers LA, Diao L, Wang J, Saintigny P, Girard L, Peyton M, *et al.* An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin Cancer Res* 2013;**19**(1):279-90 doi 10.1158/1078-0432.CCR-12-1558.
5. Powles T, Duran I, van der Heijden MS, Loriot Y, Vogelzang NJ, De Giorgi U, *et al.* Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* 2018;**391**(10122):748-57 doi 10.1016/S0140-6736(17)33297-X.
6. Kim ST, Cristescu R, Bass AJ, Kim KM, Odegaard JI, Kim K, *et al.* Comprehensive molecular characterization of clinical responses to PD-1 inhibition in metastatic gastric cancer. *Nat Med* 2018;**24**(9):1449-58 doi 10.1038/s41591-018-0101-z.
7. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, *et al.* Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. *Cell* 2017;**171**(4):934-49 e16 doi 10.1016/j.cell.2017.09.028.
8. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 2015;**160**(1-2):48-61 doi 10.1016/j.cell.2014.12.033.
9. Roh W, Chen PL, Reuben A, Spencer CN, Prieto PA, Miller JP, *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**(379) :eaah3560 doi 10.1126/scitranslmed.aah3560.
10. Huang AC, Orlowski RJ, Xu X, Mick R, George SM, Yan PK, *et al.* A single dose of neoadjuvant PD-1 blockade predicts clinical outcomes in resectable melanoma. *Nat Med* 2019;**25**(3):454-61 doi 10.1038/s41591-019-0357-y.
11. Foy JP, Karabajakian A, Ortiz-Cuaran S, Bousageon M, Michon L, Bouaoud J, *et al.* Immunologically active phenotype by gene expression profiling is associated with clinical benefit from PD-1/PD-L1 inhibitors in real-world head and neck and lung cancer patients. *Eur J Cancer* 2022;**174**:287-98 doi 10.1016/j.ejca.2022.06.034.