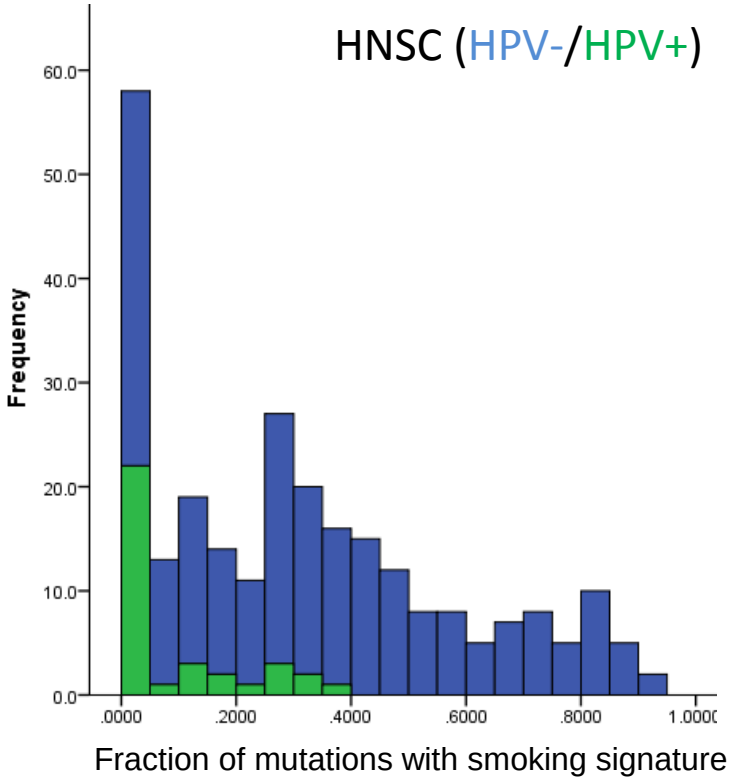
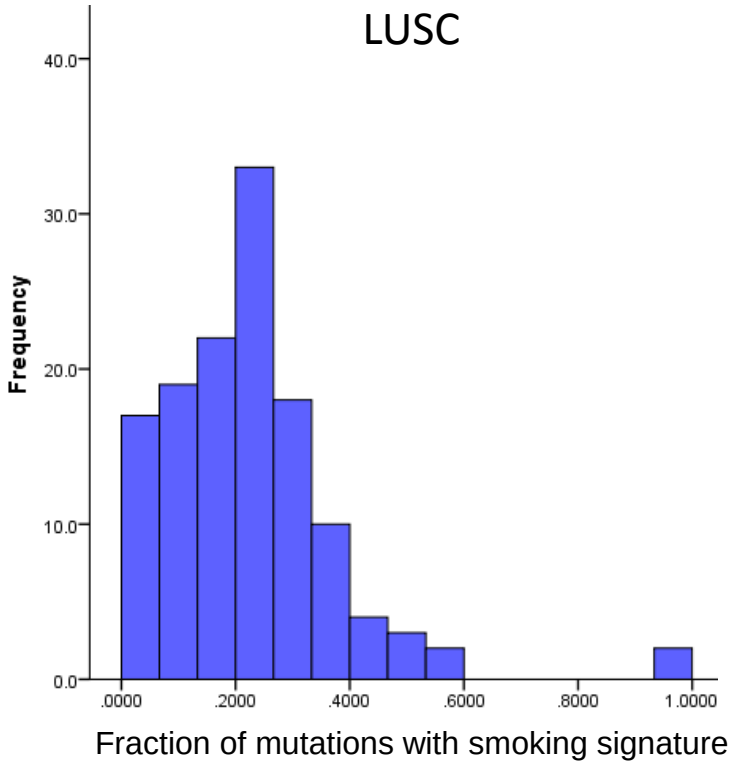
SUPPLEMENTARY FIGURES AND TABLES

Supplementary Figures

Supplementary Figure 1

Distribution of smoking signature





Difference in distributions (HNSC vs LUSC)
p=.016, Wilcoxon rank-sum test

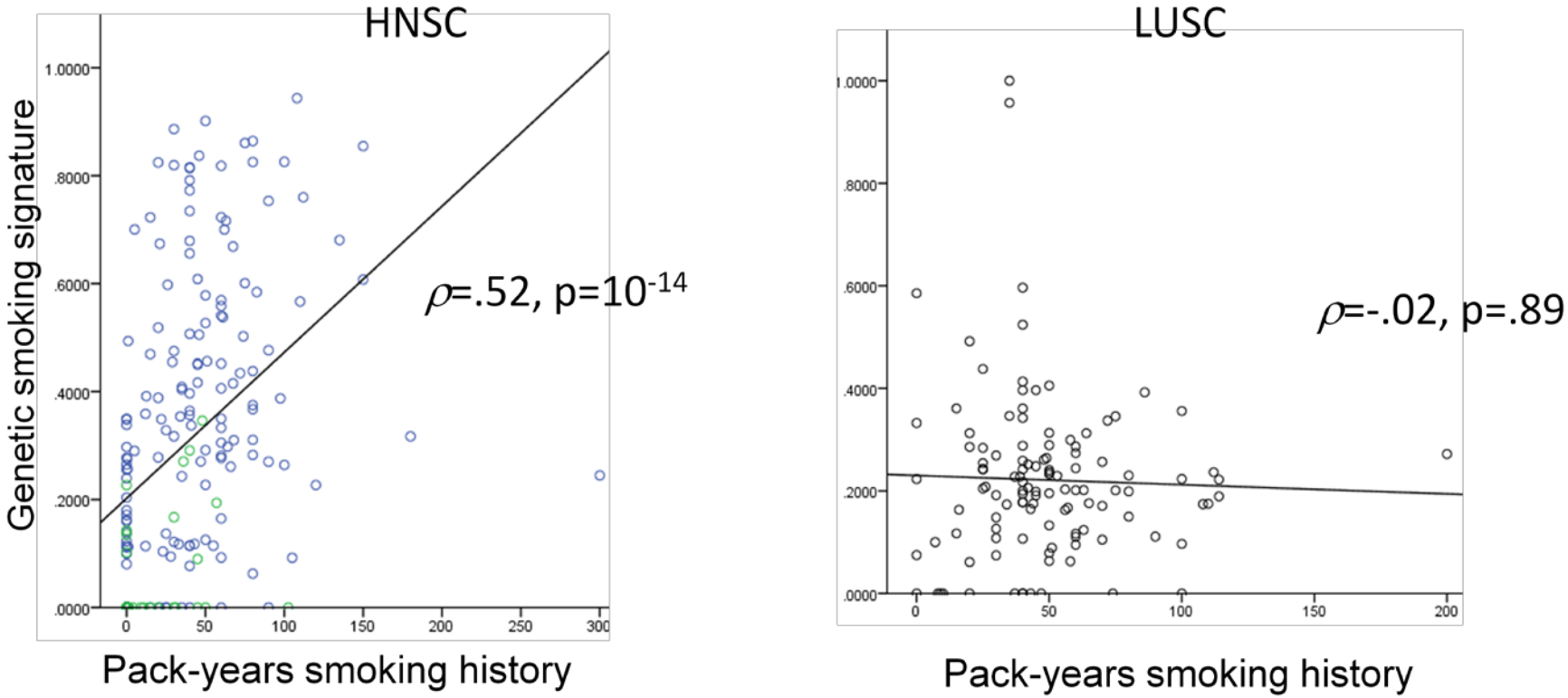
<u>% of cases with smoking signature > 0 (p=.003, Fisher exact test)</u>

LUSC	89.2%
HNSC	76.7%
HPV +	37.1%
HPV -	84.2%

Supplementary Figure 1. Histogram showing the distribution of the mutational smoking signature (signature 4, as obtained from deconstructsigs), in TCGA HNSC and LUSC samples.

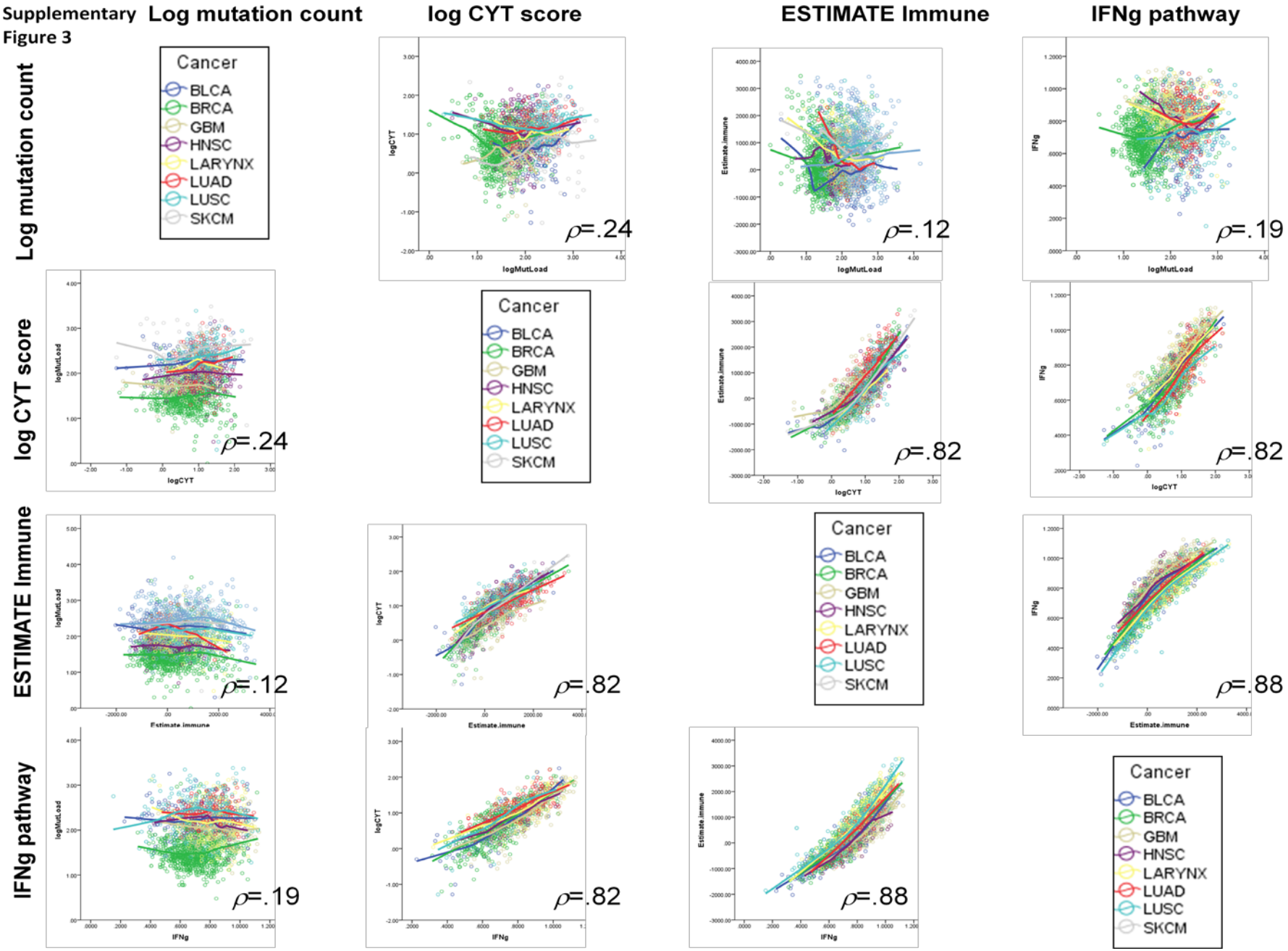
Supplementary Figure 2

Clinical smoking history and mutational smoking signature



Supplementary Figure 2. Association between clinical smoking history (in pack-years, defined as packs per day x years smoking), and the mutational smoking signature. Correlation assessed with non-parametric Spearman's rho, two-tailed.

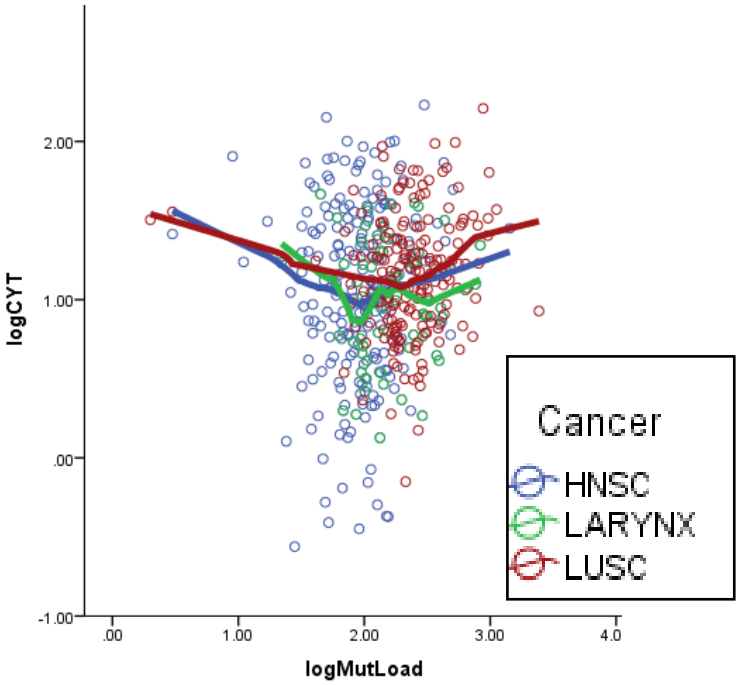
Supplementary Figure 3

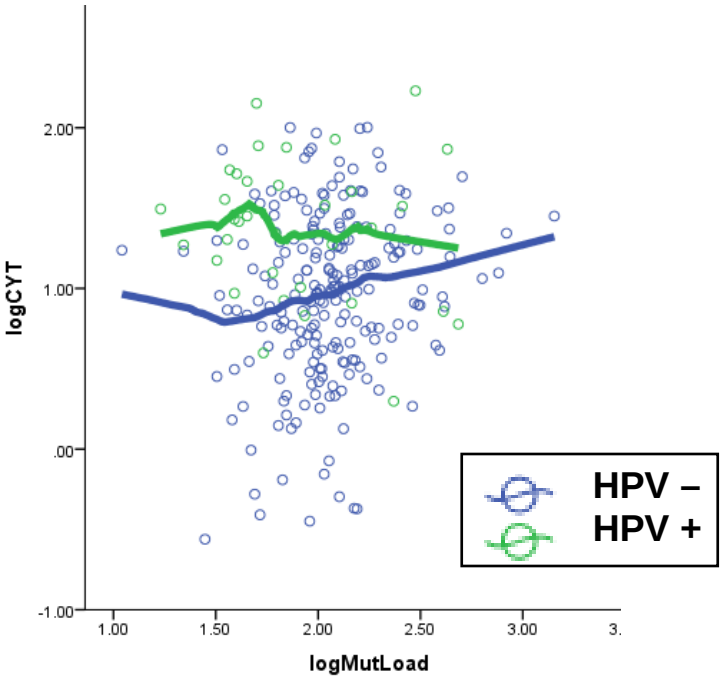


Supplementary Figure 3. Correlations between immunogenomic metrics in 8 cancer types, including $\log_{10}$ (mutation count), $\log_{10}$ (cytolytic score), ESTIMATE ImmuneScore, and enrichment of the REACTOME IFNg pathway, as described in the Methods. Overall correlations are expressed as Spearman's rho. BLCA, bladder cancer; BRCA, breast cancer; GBM, glioblastoma; HNSC, head and neck squamous cell carcinoma (all subsites); LARYNX, HNSC (larynx tumors only); LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; SKCM, cutaneous melanoma.

Supplementary Figure 4

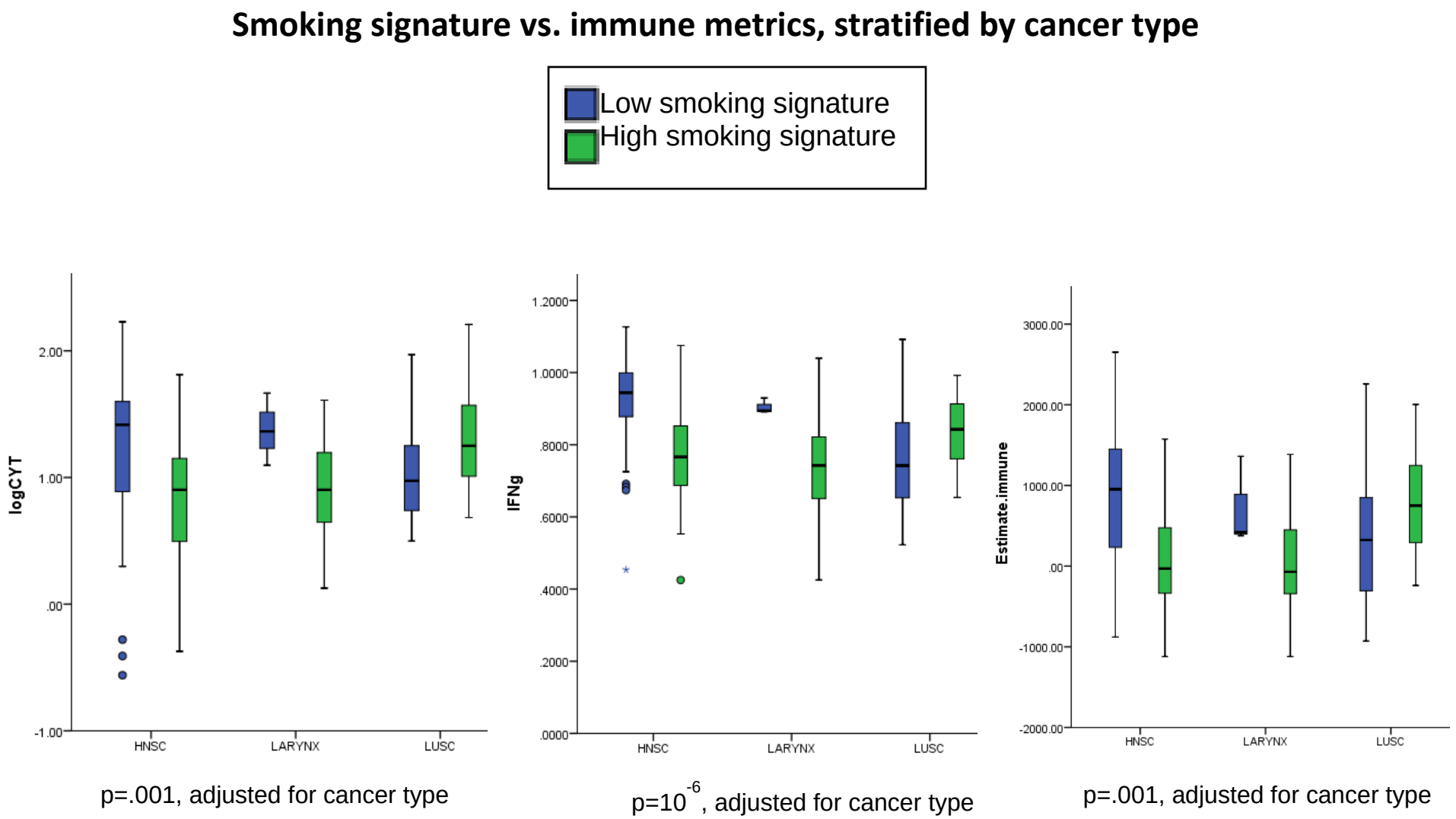
Mutational load vs. cytolytic score





Supplementary Figure 4. Scatterplots showing (**left**) mutational load and CYT (cytolytic) score in HNSC (all subsites), HNSC (larynx only), and LUSC, and (**right**) mutational load and CYT score in HPV+ and HPV- HNSC.

Supplementary Figure 5



p=.001, adjusted for cancer type

$p=10^{-6}$, adjusted for cancer type

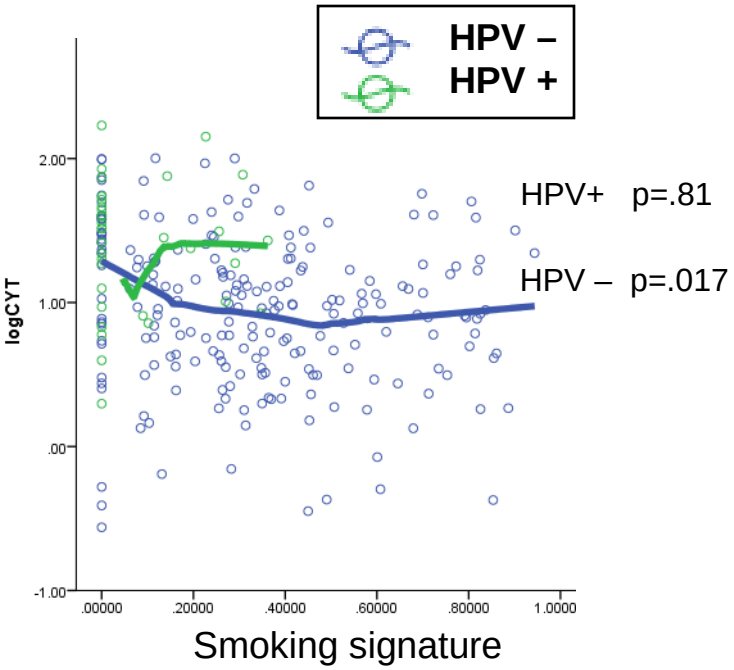
p=.001, adjusted for cancer type

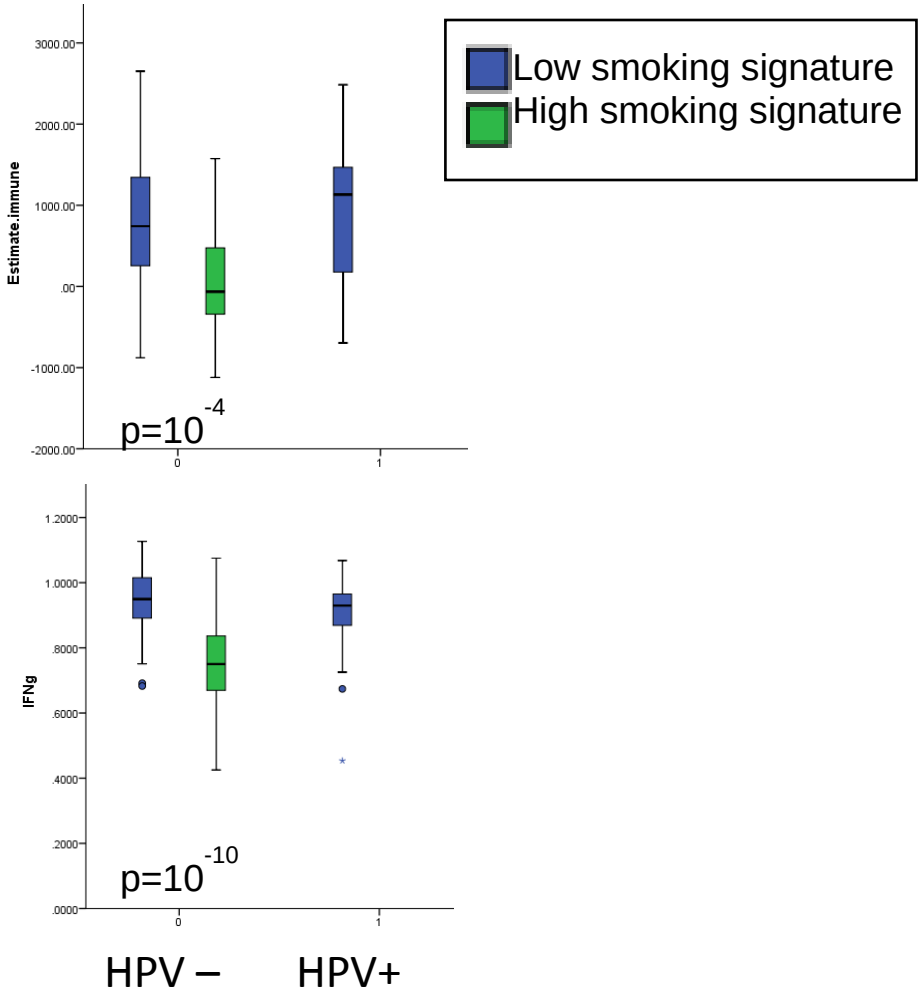
Supplementary Figure 5. Boxplots showing immune metrics and in tumors falling in the lowest and highest quartiles of mutational smoking signature, in HNSC (all subsites), HNSC (larynx only), and LUSC. Boxes represent median (horizontal line) and interquartile range, with whiskers representing minima/maxima and outliers being those samples beyond 1.5x

interquartile range. P values represent hypothesis testing for the smoking status independent variable and cancer site covariate with the labeled immune metric dependent variable, in multivariable least-squares regression.

Supplementary Figure 6

Smoking signature vs. immune metrics, stratified by HPV status

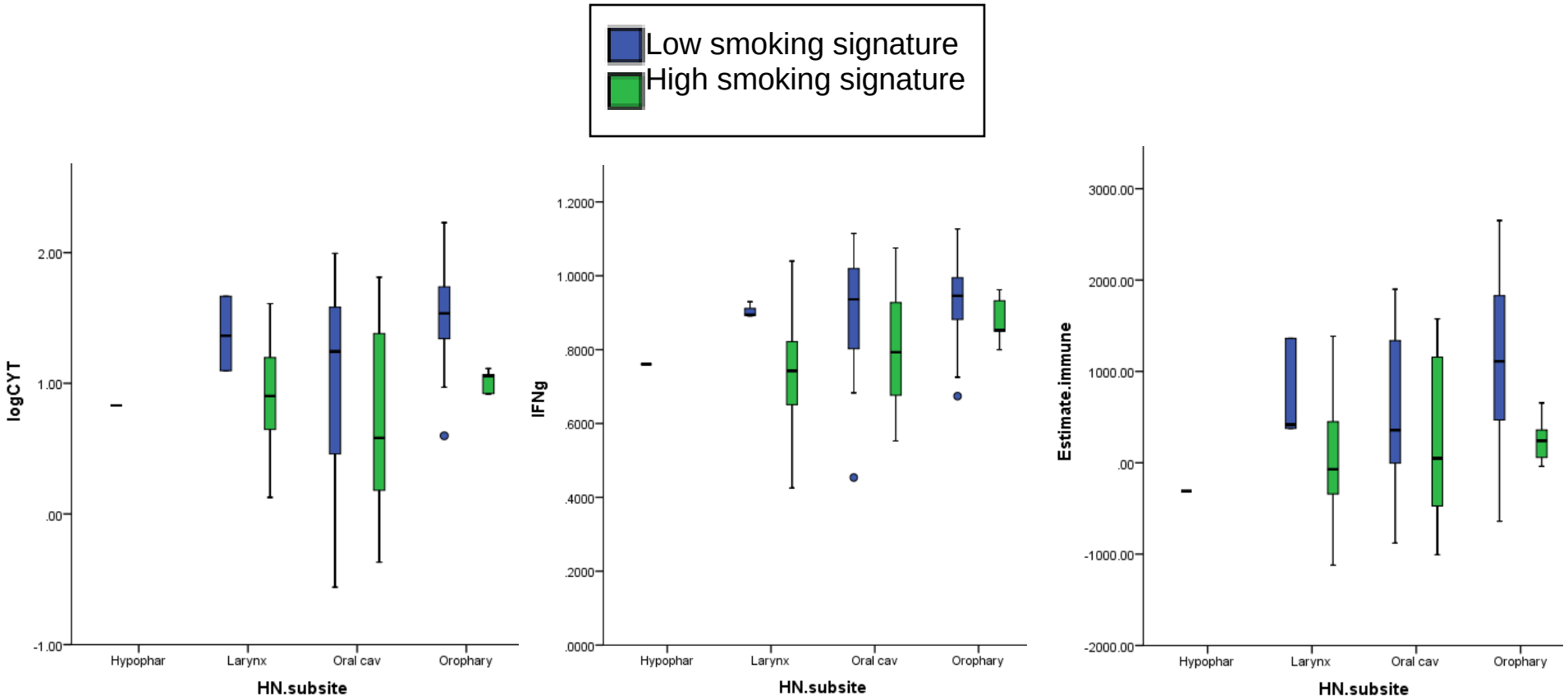




Supplementary Figure 6. (**Left**) scatterplot of mutational smoking signature vs. CYT (cytolytic) score, in HPV+ and HPV- HNSC, where p value represents Spearman correlation. (**Right**) box plots of smoking status vs. ESTIMATE ImmuneScore and IFNg pathway signaling, stratified by HPV status, where p value represents t-test. There were no HPV+ tumors with high smoking signature. HNSC, head and neck squamous cell carcinoma; HPV, human papillomavirus.

Supplementary Figure 7

Smoking signature vs. immune metrics, stratified by HN subsite



p=.001, adjusted for subsite

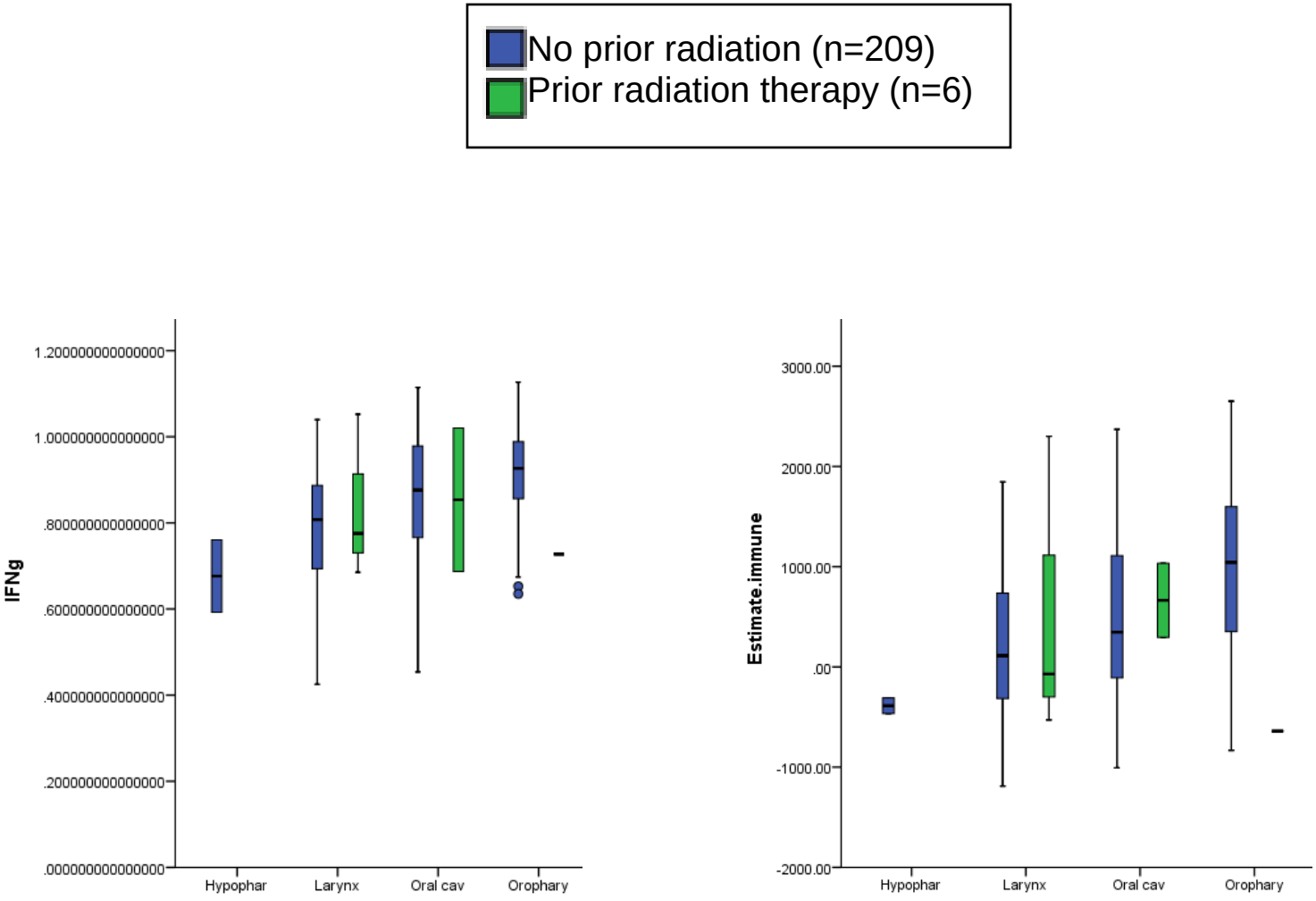
p=10^{-8}, adjusted for subsite

p=10^{-7}, adjusted for subsite

Supplementary Figure 7. Boxplots showing immune metrics in smoking-low and smoking-high HNSC tumors, stratified by anatomic subsite. P values represent hypothesis testing for the smoking status independent variable and subsite covariate with the labeled immune metric dependent variable, in multivariable least-squares regression, demonstrating that smoking is associated with an immunosuppressive phenotype across all head and neck anatomic subsites.

Supplementary Figure 8

Prior radiation therapy vs. immune metrics, stratified by HN subsite



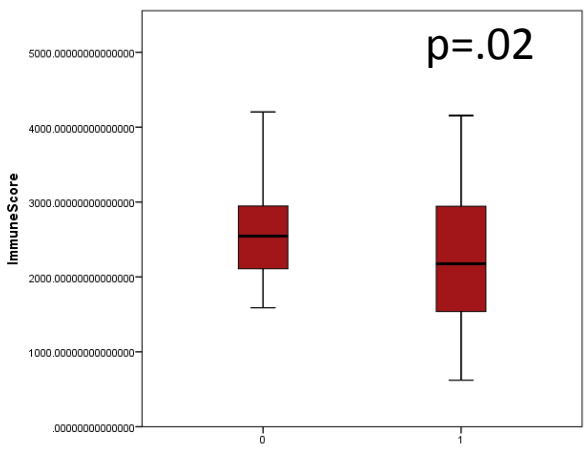
p=.66

p=.77

Supplementary Figure 8. Boxplots showing immune metrics in HNSC tumors stratified by prior radiation therapy. P values represent hypothesis testing for the radiation therapy independent variable and subsite covariate with the labeled immune metric dependent variable, in multivariable least-squares regression. Insufficient number of patients with prior radiotherapy had CYT scores available.

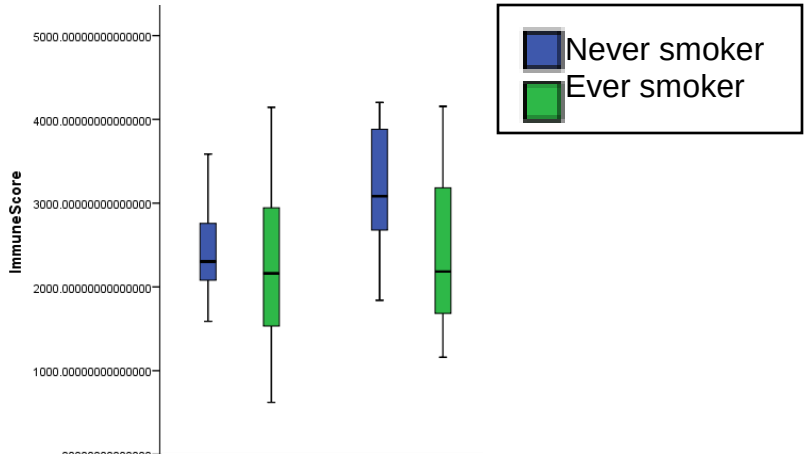
Supplementary Figure 9

ESTIMATE Immune Score and clinical smoking history
HNSC (independent expression dataset)



Never n=27 Ever n=109

Clinical smoking history



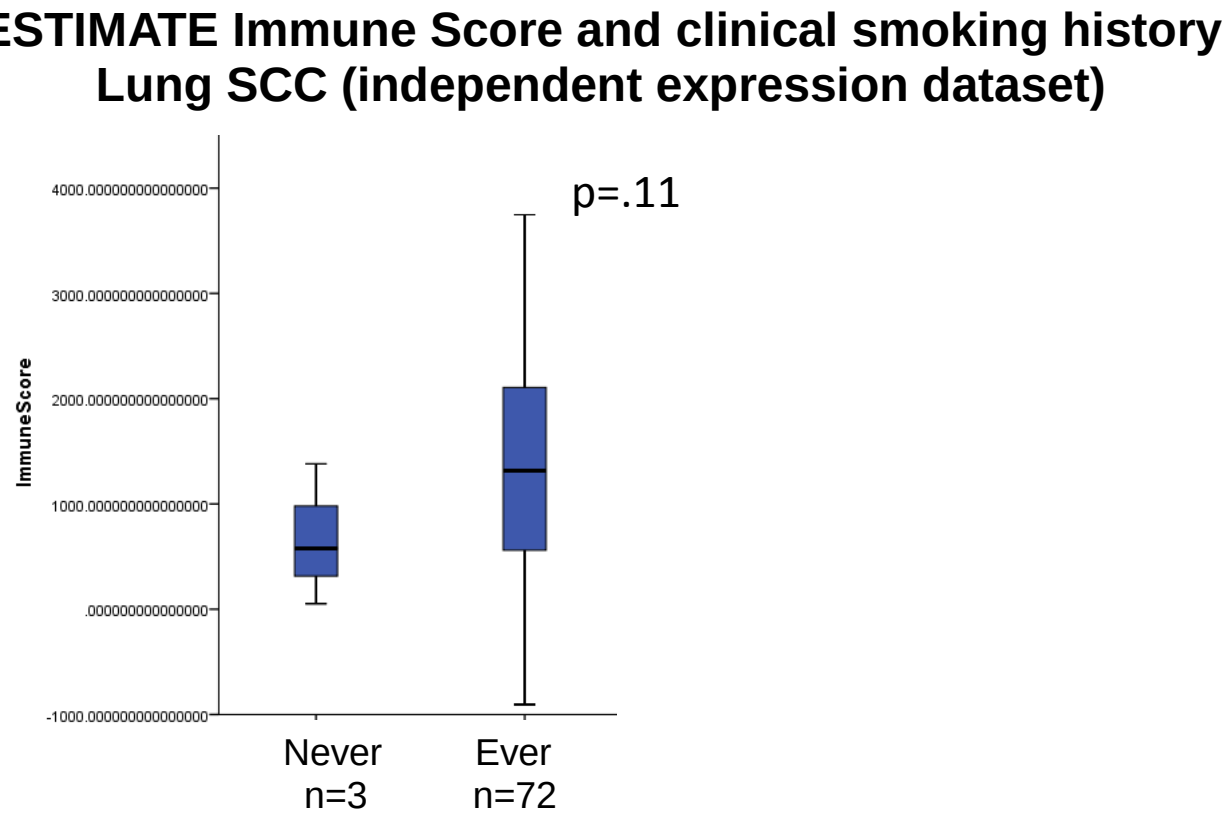
HPV – n=82 HPV + n=14

HPV status

Dataset of Walter et al, PLoS One 2013; PMID 23451093

Supplementary Figure 9. Boxplots showing immune infiltration and clinical smoking history in external HNSC expression dataset of Walter et al (n=138; 136 with smoking history, and 96 with HPV status available) (1), with hypothesis testing performed with single-tailed t-test.

Supplementary Figure 10

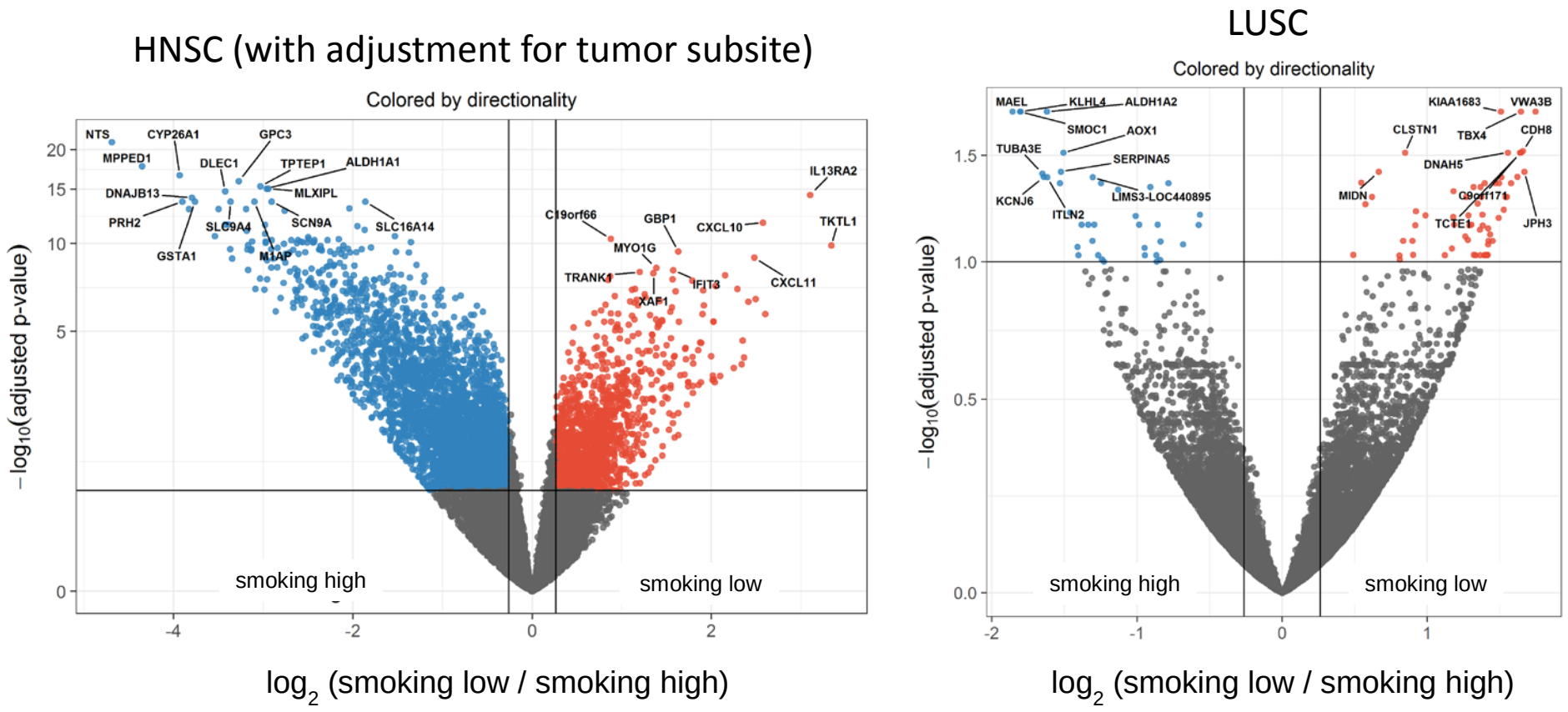


Dataset of Lee et al, Clin Canc Res 2008; PMID 19010856

Supplementary Figure 10. Boxplots showing immune infiltration and clinical smoking history in external LUSC expression dataset of Lee et al (n=75) (2), with hypothesis testing performed with single-tailed t-test.

Supplementary Figure 11

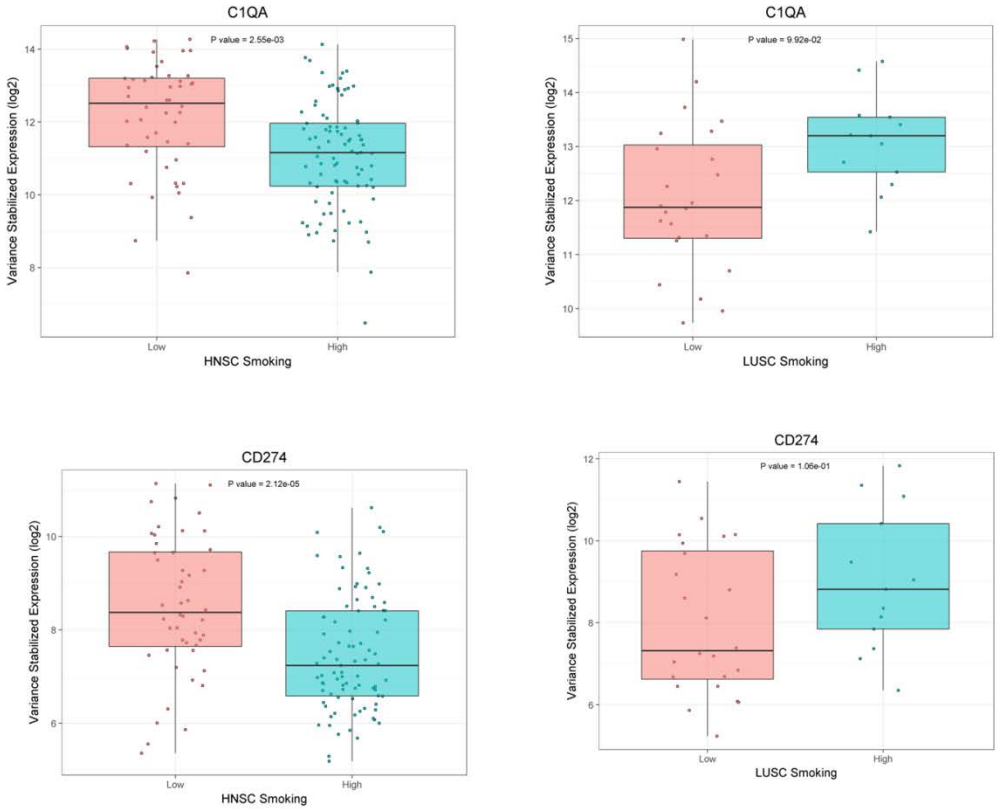
Differentially expressed genes in smoking signature-high and smoking signature-low HNSC and LUSC tumors

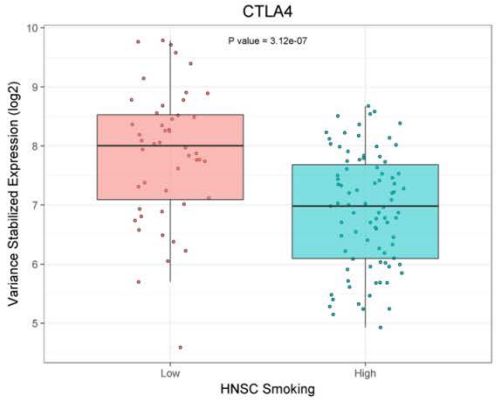


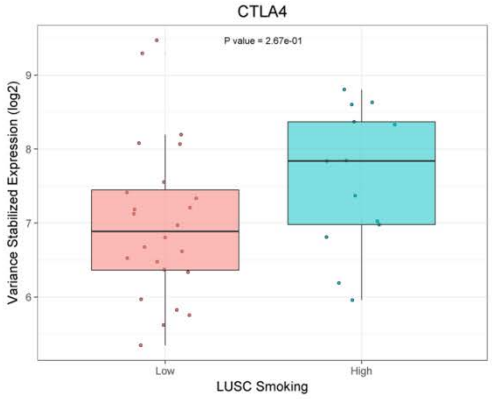
Supplementary Figure 11. Volcano plots of differentially expressed genes (DEGs) in smoking-low vs. smoking-high HNSC and LUSC. DEGs were identified in DESeq as described in the Methods, with HNSC analysis limited to HPV-negative cases only and including anatomic subsite as a covariate. DEGs are plotted with those upregulated in smoking-high tumors on the left, and those upregulated in smoking-low tumors (ie, downregulated in smoking-high tumors) on the right, with $\log_2$ fold change on the x-axis, and $-\log_{10}$ p-value on the y-axis.

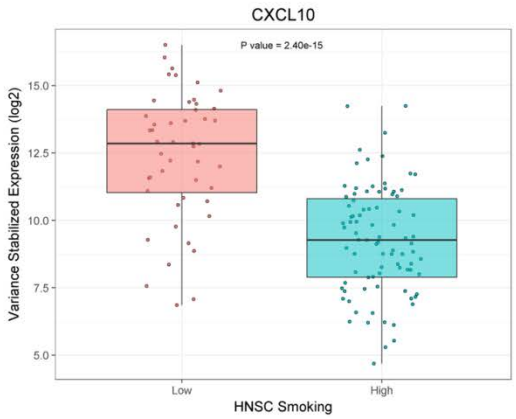
Supplementary Figure 12.

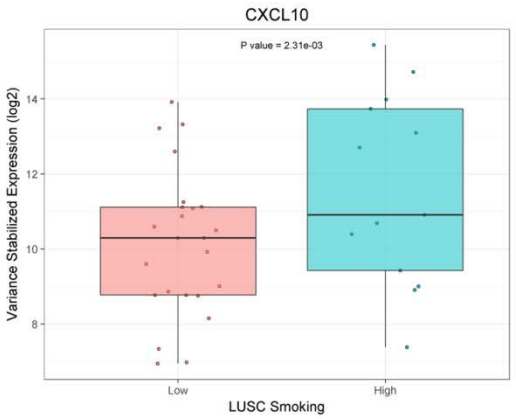
Differentially expressed genes in smoking-high and smoking-low HNSC and LUSC.

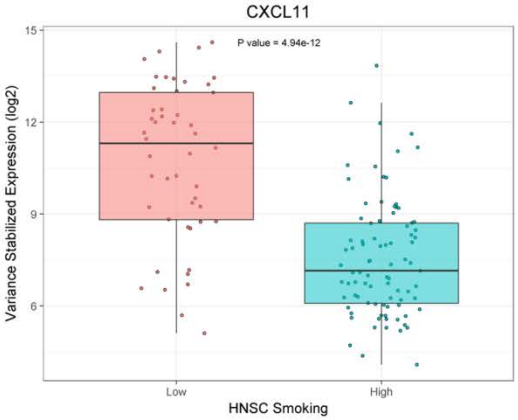


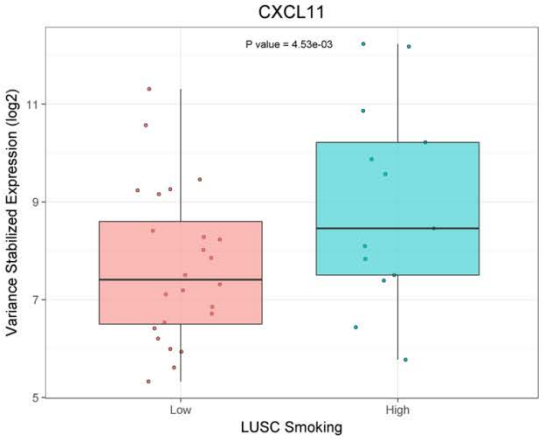

CTLA4
P value = 3.12e-07
Variance Stabilized Expression (log2)
10
9
8
7
6
5
Low
High
HNSC Smoking

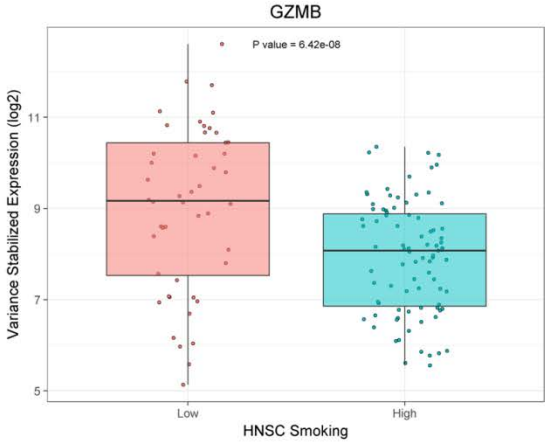

CTLA4
P value = 2.67e-01
Variance Stabilized Expression (log2)
9
8
7
6
Low
High
LUSC Smoking

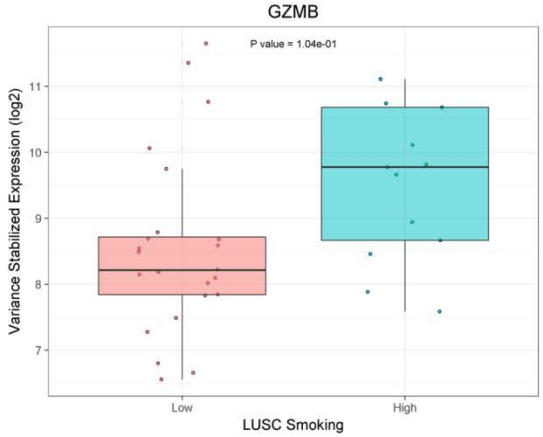

CXCL10
P value = 2.40e-15
Variance Stabilized Expression (log2)
15.0
12.5
10.0
7.5
5.0
Low
High
HNSC Smoking

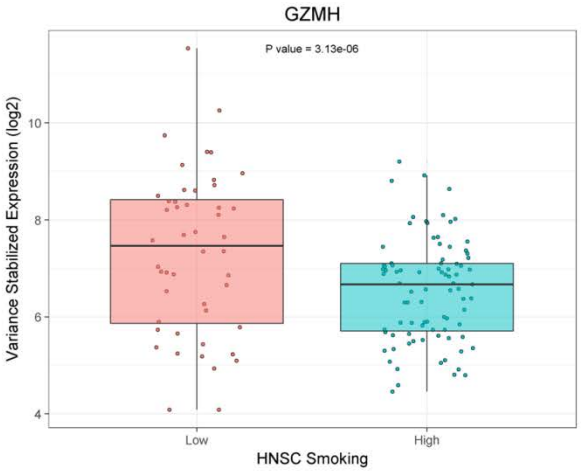

CXCL10
P value = 2.31e-03
Variance Stabilized Expression (log2)
14
12
10
8
Low
High
LUSC Smoking

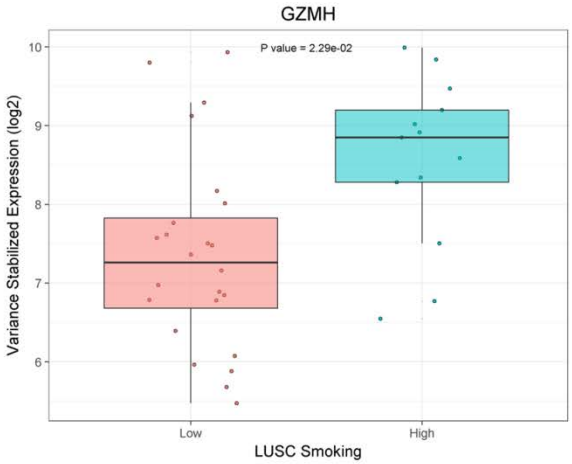

CXCL11
P value = 4.94e-12
Variance Stabilized Expression (log2)
15
12
9
6
Low
High
HNSC Smoking

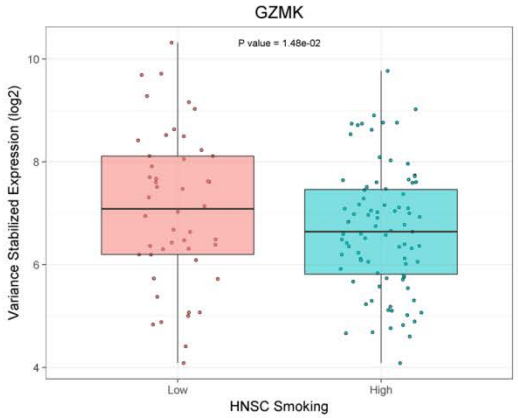

CXCL11
P value = 4.53e-03
Variance Stabilized Expression (log2)
11
9
7
5
Low
High
LUSC Smoking

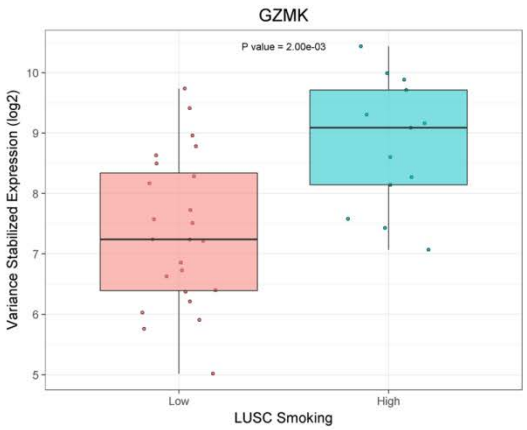

GZMB
P value = 6.42e-08
Variance Stabilized Expression (log2)
11
9
7
5
Low
High
HNSC Smoking

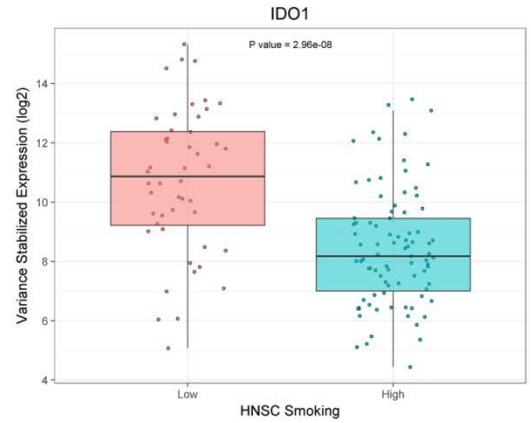

GZMB
P value = 1.04e-01
Variance Stabilized Expression (log2)
11
10
9
8
7
Low
High
LUSC Smoking

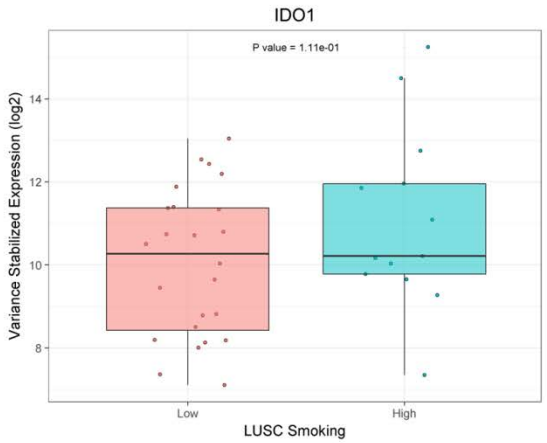

GZMH
P value = 3.13e-06
Variance Stabilized Expression (log2)
10
8
6
4
Low
High
HNSC Smoking

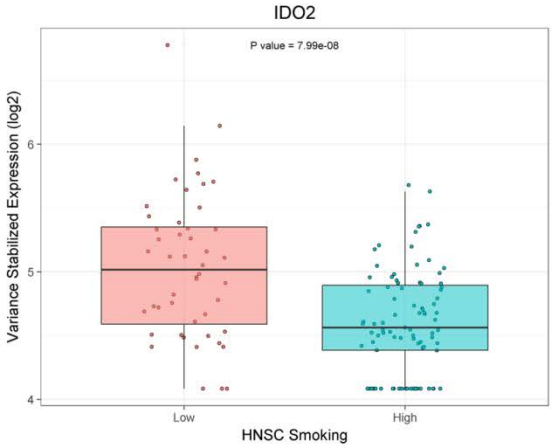

GZMH
P value = 2.29e-02
Variance Stabilized Expression (log2)
10
9
8
7
6
Low
High
LUSC Smoking

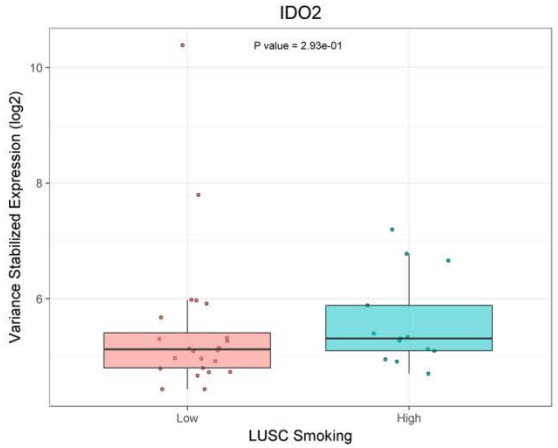

GZMK
P value = 1.48e-02
Variance Stabilized Expression (log2)
10
8
6
4
Low
High
HNSC Smoking

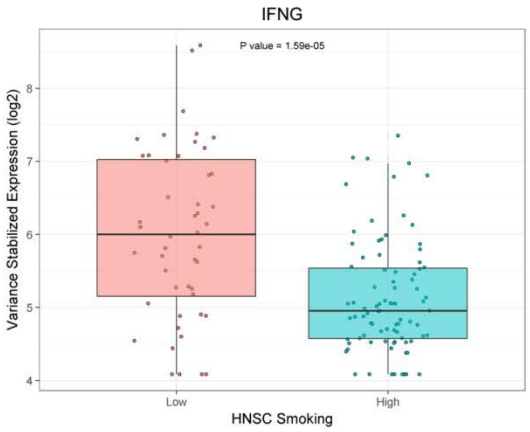

GZMK
P value = 2.00e-03
Variance Stabilized Expression (log2)
10
9
8
7
6
5
Low
High
LUSC Smoking

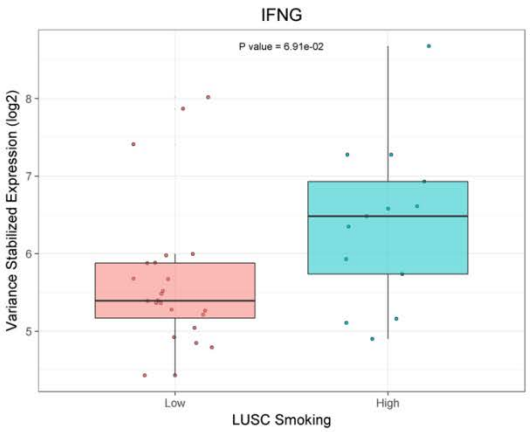

IDO1
P value = 2.96e-08
Variance Stabilized Expression (log2)
14
12
10
8
6
4
Low
High
HNSC Smoking

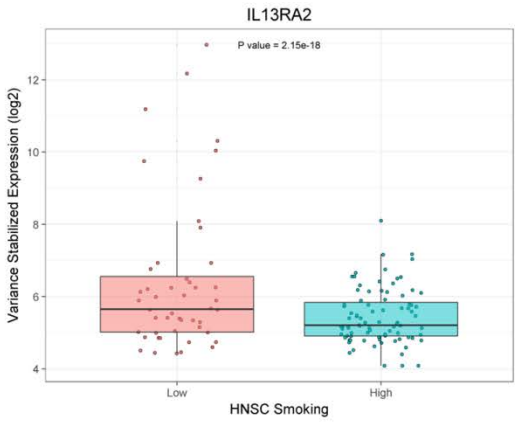

IDO1
P value = 1.11e-01
Variance Stabilized Expression (log2)
14
12
10
8
Low
High
LUSC Smoking

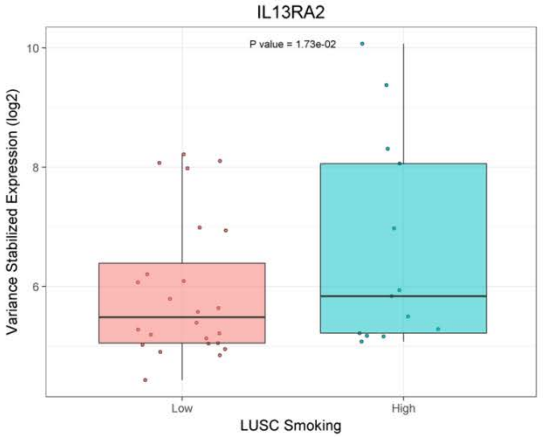

IDO2
P value = 7.99e-08
Variance Stabilized Expression (log2)
6
5
4
Low
High
HNSC Smoking

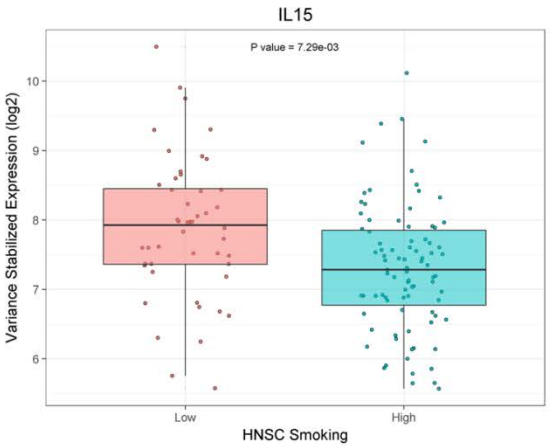

IDO2
P value = 2.93e-01
Variance Stabilized Expression (log2)
10
8
6
Low
High
LUSC Smoking

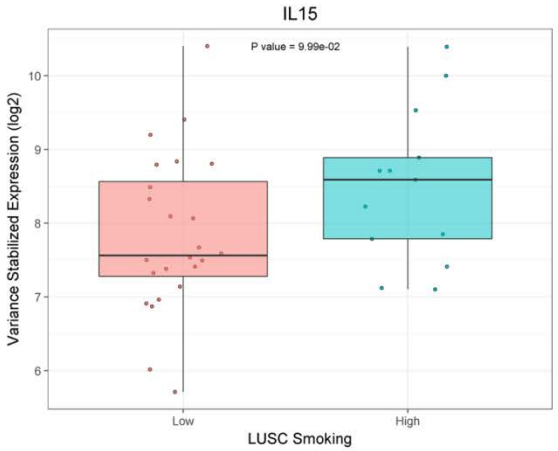

IFNG
P value = 1.59e-05
Variance Stabilized Expression (log2)
8
7
6
5
4
Low
High
HNSC Smoking

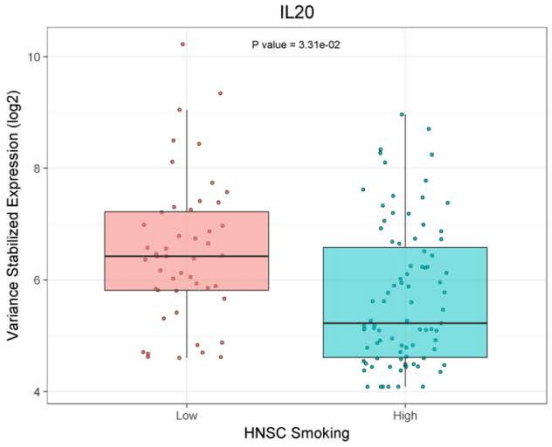

IFNG
P value = 6.91e-02
Variance Stabilized Expression (log2)
8
7
6
5
Low
High
LUSC Smoking

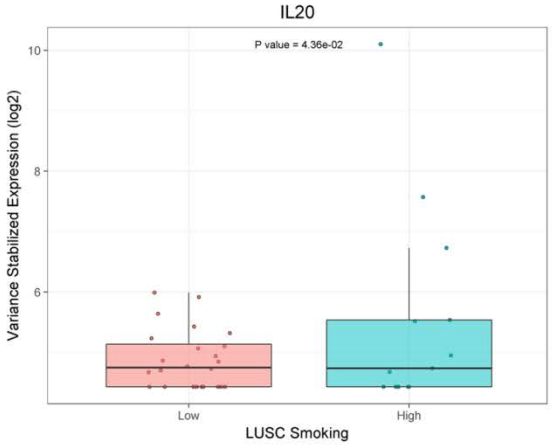

IL13RA2
P value = 2.15e-18
Variance Stabilized Expression (log2)
12
10
8
6
4
Low
High
HNSC Smoking

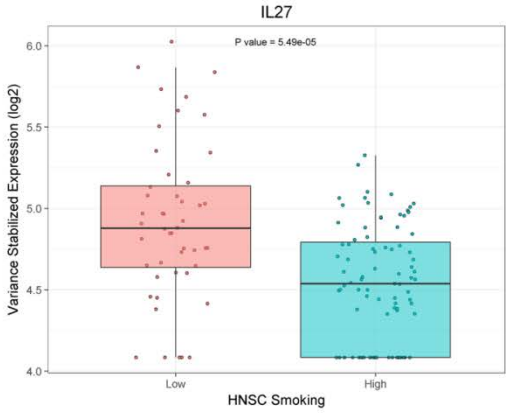

IL13RA2
P value = 1.73e-02
Variance Stabilized Expression (log2)
10
8
6
Low
High
LUSC Smoking

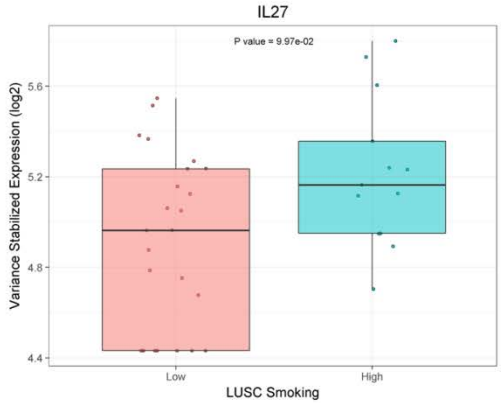

IL15
P value = 7.29e-03
Variance Stabilized Expression (log2)
10
9
8
7
6
Low
High
HNSC Smoking

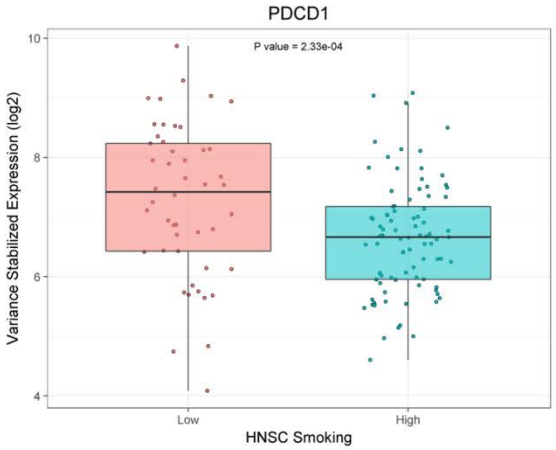

IL15
P value = 9.99e-02
Variance Stabilized Expression (log2)
10
9
8
7
6
Low
High
LUSC Smoking

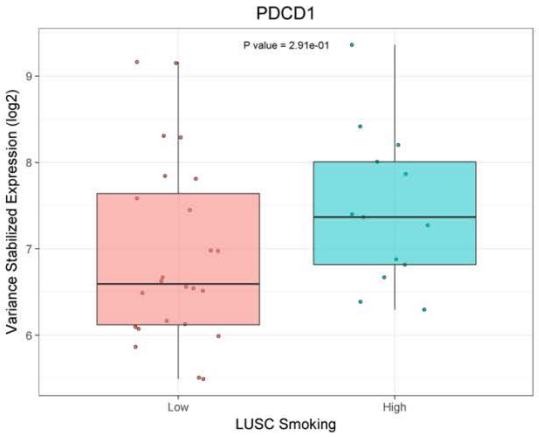

IL20
P value = 3.31e-02
Variance Stabilized Expression (log2)
10
8
6
4
Low
High
HNSC Smoking

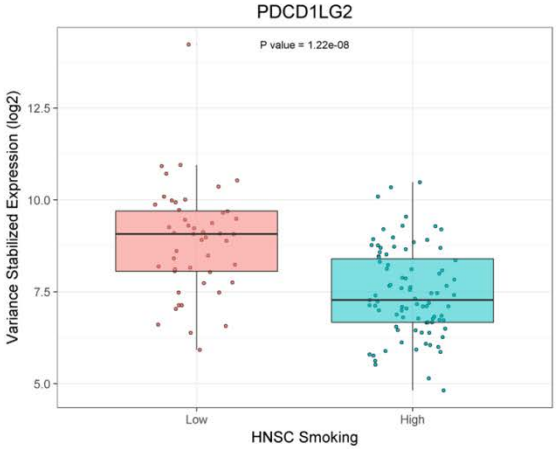

IL20
P value = 4.36e-02
Variance Stabilized Expression (log2)
10
8
6
Low
High
LUSC Smoking

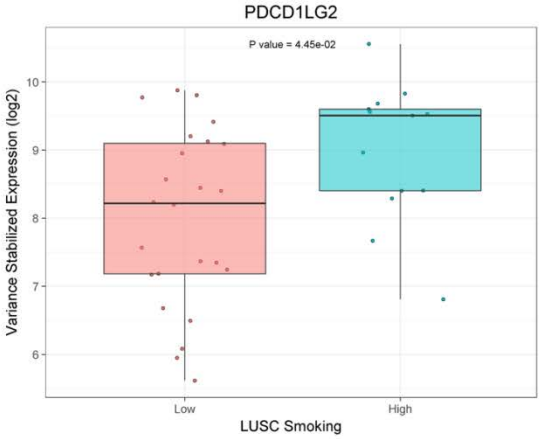

IL27
P value = 5.49e-05
Variance Stabilized Expression (log2)
6.0
5.5
5.0
4.5
4.0
Low
High
HNSC Smoking

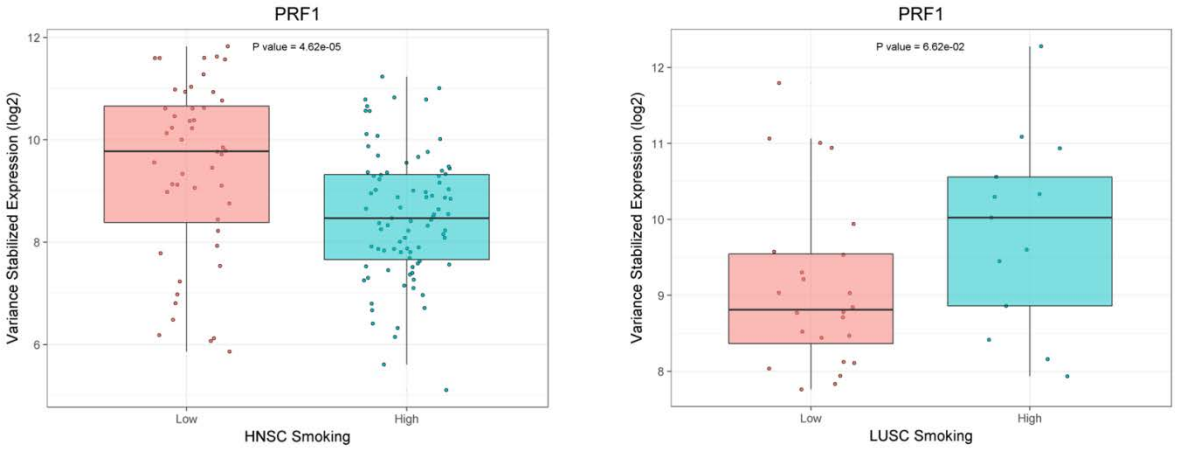

IL27
P value = 9.97e-02
Variance Stabilized Expression (log2)
5.6
5.2
4.8
4.4
Low
High
LUSC Smoking


PDCD1
P value = 2.33e-04
Variance Stabilized Expression (log2)
10
8
6
4
Low
High
HNSC Smoking


PDCD1
P value = 2.91e-01
Variance Stabilized Expression (log2)
9
8
7
6
Low
High
LUSC Smoking


PDCD1LG2
P value = 1.22e-08
Variance Stabilized Expression (log2)
12.5
10.0
7.5
5.0
Low
High
HNSC Smoking


PDCD1LG2
P value = 4.45e-02
Variance Stabilized Expression (log2)
10
9
8
7
6
Low
High
LUSC Smoking



Supplementary Figure 12. Boxplots showing immune genes differentially expressed, with opposite directionality, in HNSC vs. LUSC. Each row represents the indicated gene in HNSC (HPV-negative only) (left) and LUSC (right), stratified by smoking status, with nominal p-value displayed at the top of the graph.

Supplementary Tables

Supplementary Table 1.

Summary of genetic smoking signature (Signature 4) in head and neck and lung squamous cell carcinomas*

Variable	HNSC (all)	HNSC (HPV- only)	LUSC
Total patients with tobacco signature data	287	228	130
Mean (median) followup time, days	847 (536)	808 (499)	939 (686)
Mean (SD) tobacco signature	.30 (.26)	.34 (.26)	.22 (.16)
Patients with low tobacco signature	81 (28%)	48 (21%)	24 (18%)
Mean (SD) tobacco signature	.014 (.03)	.021 (.037)	.033 (.04)
Patients with high tobacco signature	91 (32%)	87 (38%)	13 (10%)
Mean (SD) tobacco signature	.61 (.16)	.61 (.16)	.55 (.20)

HNSCC (with available smoking signature, known HPV and subsite statu	Oral cavity	Oropharynx	Larynx	Hypopharynx	Overlapping	Total
HPV-negative	85	14	71	1	57	228
HPV-positive	8	20	1	1	5	35
Total	93	34	72	2	62	263

*Summary of genetic smoking signature (Signature 4) in head and neck and lung squamous cell carcinomas, expressed as mean with SD (standard deviation), and summary of distribution of HPV+ and HPV- HNSC by anatomic subsite. Not included in lower table: 24 HNSC cases with unknown HPV status and anatomic subsite. HNSC, head and neck squamous cell carcinoma. LUSC, lung squamous cell carcinoma. HPV, human papillomavirus.

Supplementary Table 2. Complete HNSC data, including TCGA ID, clinical, mutational smoking signature, and other immunogenomic metrics.

Available for separate download.

Supplementary Table 3. Complete LUSC data, including TCGA ID, clinical, mutational smoking signature, and other immunogenomic metrics.

Available for separate download.

Supplementary Table 4. Cox regression models of overall survival in HNSC and LUSC, including mutational smoking signature (categorical or continuous), and other covariates.

Covariate	HR	95%CI	P value*
HNSC			
Model with smoking signature alone			
Smoking Signature (High vs Low)	1.50	1.23-1.81	3.30×10^{-6}
Model with smoking signature and HPV status			
Smoking Signature (High vs Low)	1.35	1.11-1.65	0.003
HPV+	0.26	0.06-1.13	0.07
Model with smoking signature, HPV status, and radiation treatment			
Smoking Signature (High vs Low)	1.35	1.11-1.65	0.003
HPV+	0.77	0.43-1.40	0.41
Adjuvant radiotherapy	0.40	0.06-1.18	0.08
LUSC			
Model with smoking signature alone			
Smoking Signature (High vs Low)	1.02	0.71-1.46	0.92
Model excluding 4 patients receiving adjuvant chemotherapy or radiotherapy			
Smoking Signature (High vs Low)	0.83	0.28-2.45	0.73
HNSC and LUSC combined model, with interaction term			
Smoking Signature (High vs Low)	1.47	1.26-1.72	5.12×10^{-5}
Cancer type (LUSC vs HNSC)	2.53	1.10-6.46	0.10
Interaction (cancer type x smoking)	0.68	0.48-0.96	0.03
Alternative models including smoking signature as a continuous variable			
HNSC			
Smoking signature (continuous)	2.50	1.31-4.78	0.006

LUSC			
Smoking signature(continuous)	1.3	0.26-6.44	0.75
Smoking signature as a continuous variable, adjusted for adjuvant therapy			
HNSC			
Smoking signature (continuous)	2.42	1.26-4.64	0.008
Adjuvant radiation therapy	0.68	0.44-1.06	0.09
LUSC			
Smoking signature (continuous)	1.01	0.19-5.40	0.99
Adjuvant radiation therapy	1.77	0.42-7.38	0.43
Adjuvant chemotherapy	0.44	0.14-1.45	0.18

*2-tailed Wald statistic in Cox regression model

HNSC, head and neck squamous cell carcinoma

LUSC, lung squamous cell carcinoma

Supplementary Table 5. Correlations of the genetic smoking signature with other immunogenomic metrics**.**

	Spearman rho	P value*
log (mutation count) vs smoking signature		
HNSCC (all)	0.33	1.01×10^{-7}
Larynx SCC	0.68	6.22×10^{-10}
Lung SCC	0.49	2.80×10^{-9}
log (CYT score) vs smoking signature		
HNSCC (all)	-0.28	4.07×10^{-6}
Larynx SCC	-0.15	0.22
Lung SCC	0.20	0.02
IFNg pathway vs smoking signature		
HNSCC (all)	-0.39	3.20×10^{-11}
Larynx SCC	-0.38	0.001
Lung SCC	0.18	0.047
ESTIMATE Immune Score vs smoking signature		
HNSCC (all)	-0.37	1.29×10^{-10}
Larynx SCC	-0.29	0.01
Lung SCC	0.19	0.03

With adjustment for tumor purity[†]

HNSC	Standardized beta	P value*
Immune Score	-0.27	<.001
ASCAT purity	0.042	0.54
HPV	-0.29	<.001

LUSC	Standardized beta	P value*
Immune Score	0.2	0.05
ASCAT purity	0.01	0.92

Somatic alterations vs. immune metrics

Cancer Type / Gene		ESTIMATE Immune Score	P value[‡]
HNSC			
TP53	wild-type	624	p=.003
	mutated	337	
PIK3CA	wild-type	462	p=.96
	mutated	468	
CASP8	wild-type	443	p=.04
	mutated	820	
CDKN2A	wild-type	780	p=.68
	mutated	815	
NOTCH1	wild-type	453	p=.56
	mutated	537	
FAT1	wild-type	460	p=.84
	mutated	483	

	HLA-A	wild-type	455	p=.22
		mutated	815	
	B2M	wild-type	469	p=.40
		mutated	170	
	HRAS	wild-type	459	p=.60
		mutated	668	
	NFE2L2	wild-type	479	p=.10
		mutated	151	
	FBXW7	wild-type	790	p=.07
		mutated	584	
LUSC				
	TP53	wild-type	360	p=.38
		mutated	626	
	PIK3CA	wild-type	606	p=.70
		mutated	538	
	KEAP1	wild-type	676	p=.05
		mutated	196	
	PTEN	wild-type	602	p=.82
		mutated	652	
	NFE2L2	wild-type	625	p=.56
		mutated	531	
	NOTCH1	wild-type	641	p=.21
		mutated	199	
	HLA-A	wild-type	615	p=.98
		mutated	624	
	CDKN2A	wild-type	646	p=.16
		mutated	431	
	CDKN2A	normal	542	p=.04
		deletion	801	
	CASP8	wild-type	594	p=.04
		mutated	1574	

* two-tailed significance of Spearman correlation

† Linear regression model for tobacco signature, including ESTIMATE Immune Score, ASCAT purity, and HNSC HPV Status as covariates.

‡ two-tailed independent samples t-test, comparing mutated vs. wild-type

Supplementary Table 6. List of differentially expressed genes in HNSC (HPV-negative cases only), comparing smoking-low to smoking-high tumors with FDR-adjusted p-value <.1. A positive fold change indicates upregulation in smoking-low tumors. HNSC tumor subsite was included as a covariate in the determination of DEGs. HPV-positive cases were excluded.

Available for separate download.

Supplementary Table 7. List of differentially expressed genes in LUSC, comparing smoking-low to smoking-high tumors with FDR-adjusted p-value <.1. A positive fold change indicates upregulation in smoking-low tumors.

Available for separate download.

Supplementary Table 8. List of signaling pathways enriched in DEGs in HNSC, as determined by Ingenuity Pathway Analysis, and expressed with –log p-value, directional z-score (where available), and ratio of DEGs in the named pathway.

Available for separate download.

Supplementary Table 9. List of signaling pathways enriched in DEGs in LUSC, as determined by Ingenuity Pathway Analysis, and expressed with –log p-value, directional z-score (where available), and ratio of DEGs in the named pathway.

Available for separate download.

Supplementary Table 10. Differentially expressed genes between smoking-high and smoking-low tumors, with opposite directionality in HNSC vs. LUSC. The gene list has been truncated to those genes with nominal p-value <.1 in at least one cancer type.

Available for separate download.

Supplementary Table 11. Overlap of differentially regulated genes in Supplementary Table 10 with the gene signatures shown to have predictive utility in anti-PD-1 treated patients in Ayers et al (3).

IFN gamma signature (4/6)
IDO1
CXCL10
CXCL9
HLA-DRA
STAT1
IFNG
Expanded immune gene signature (T cell inflamed) (5/18)
CD3D
IDO1
CIITA
CD3E
CCL5
GZMK
CD2
HLA-DRA
CXCL13
IL2RG
NKG7
HLA-E
CXCR6
LAG3
TAGAP
CXCL10
STAT1
GZMB

Supplementary Table 12. Clinical response data in head and neck squamous cell carcinoma patients: 81 patients treated at MSKCC with anti-PD-1 or anti-PD-L1 drugs, with minimum 6 month follow-up.

Response	Clinical smoking history		CB rate		
	No. of Former	No. of Never	Smokers, %	Nonsmokers, %	P*
All HNSC patients (n=81)			33.3	57.1	.03
Clinical benefit (CR/PR or SD)	18	16			
No clinical benefit (POD)	36	12			
HPV-negative patients (n=43)			25.6	66.7	.02
Clinical benefit (CR/PR or SD)	8	8			
No clinical benefit (POD)	23	4			
HPV-positive patients (n=38)			43.4	50.6	.47
Clinical benefit (CR/PR or SD)	10	8			
No clinical benefit (POD)	13	8			

* Fisher exact test, single-tailed

HNSC, head and neck squamous cell carcinoma

HPV, human papillomavirus

CB, clinical benefit; CR, complete response; PR, partial response; POD, progression of disease

REFERENCES

1. Walter V, Yin X, Wilkerson MD, et al. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. *PloS one.* 2013;8(2):e56823.
2. Lee ES, Son DS, Kim SH, et al. Prediction of recurrence-free survival in postoperative non-small cell lung cancer patients by using an integrated model of clinical information and gene expression. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2008;14(22):7397-7404.
3. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *The Journal of clinical investigation.* 2017;127(8):2930-2940.
4. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst.* 2005;97(16):1180-1184.