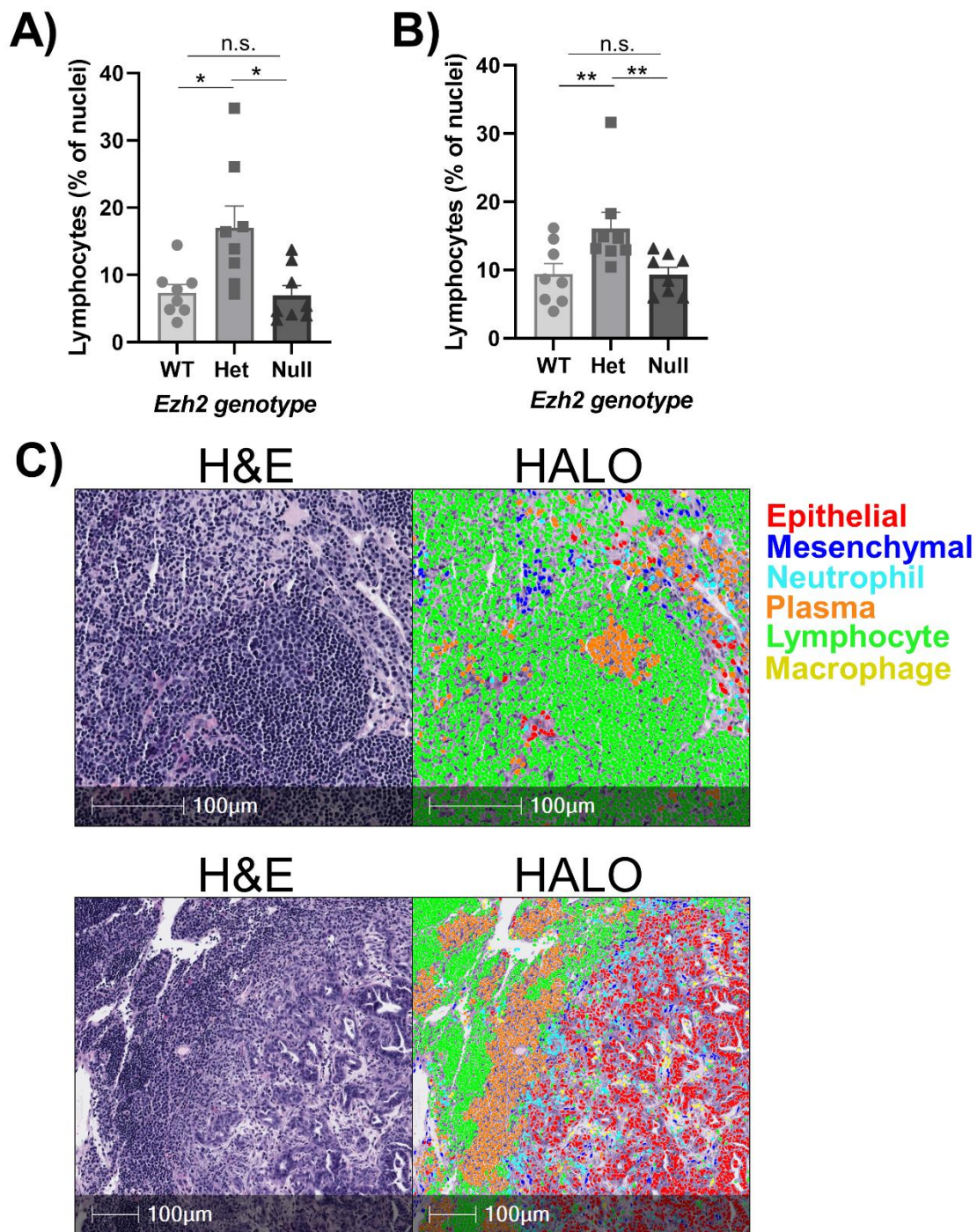


Using artificial intelligence to identify tumor microenvironment heterogeneity in non-small cell lung cancers

DuCote et al., Supplementary Information

Supplementary Figures 1-3

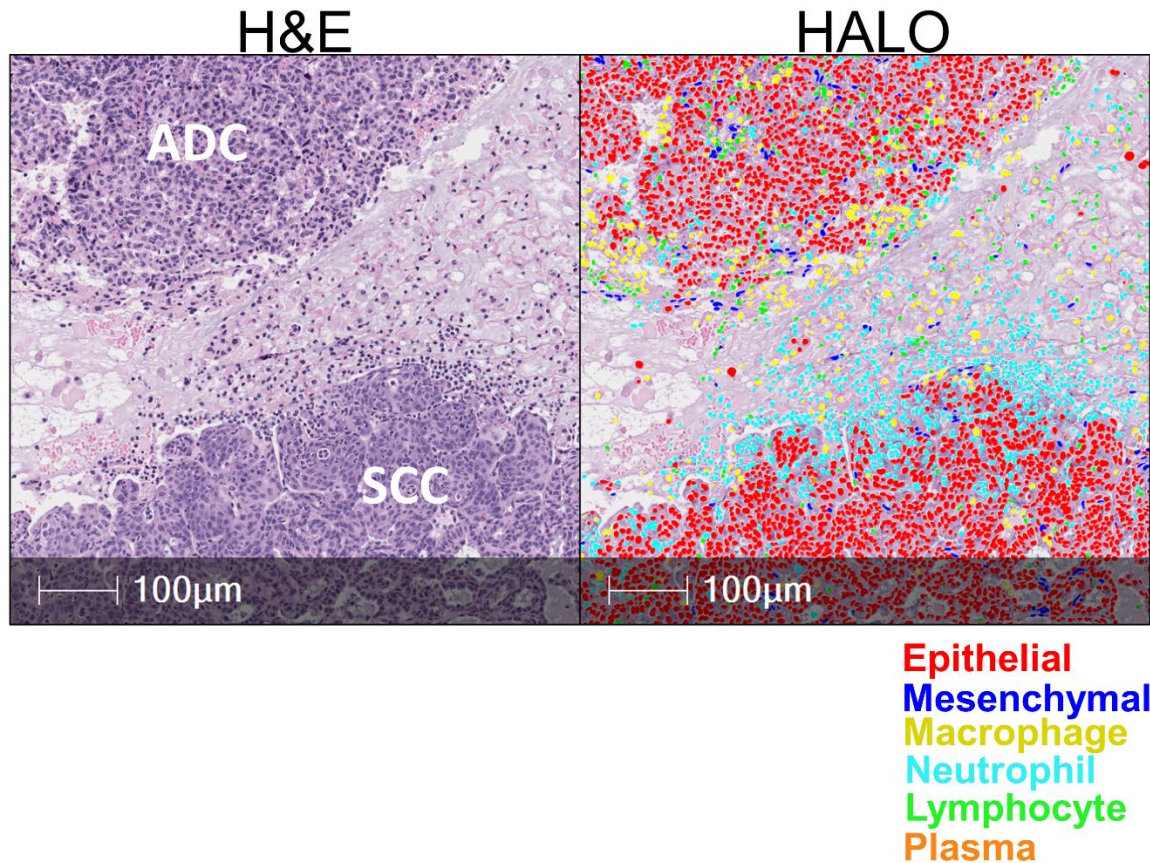
Supplementary Table 1



Supplementary Figure 1: HALO AI nuclear phenotyper algorithm accurately identifies cell types

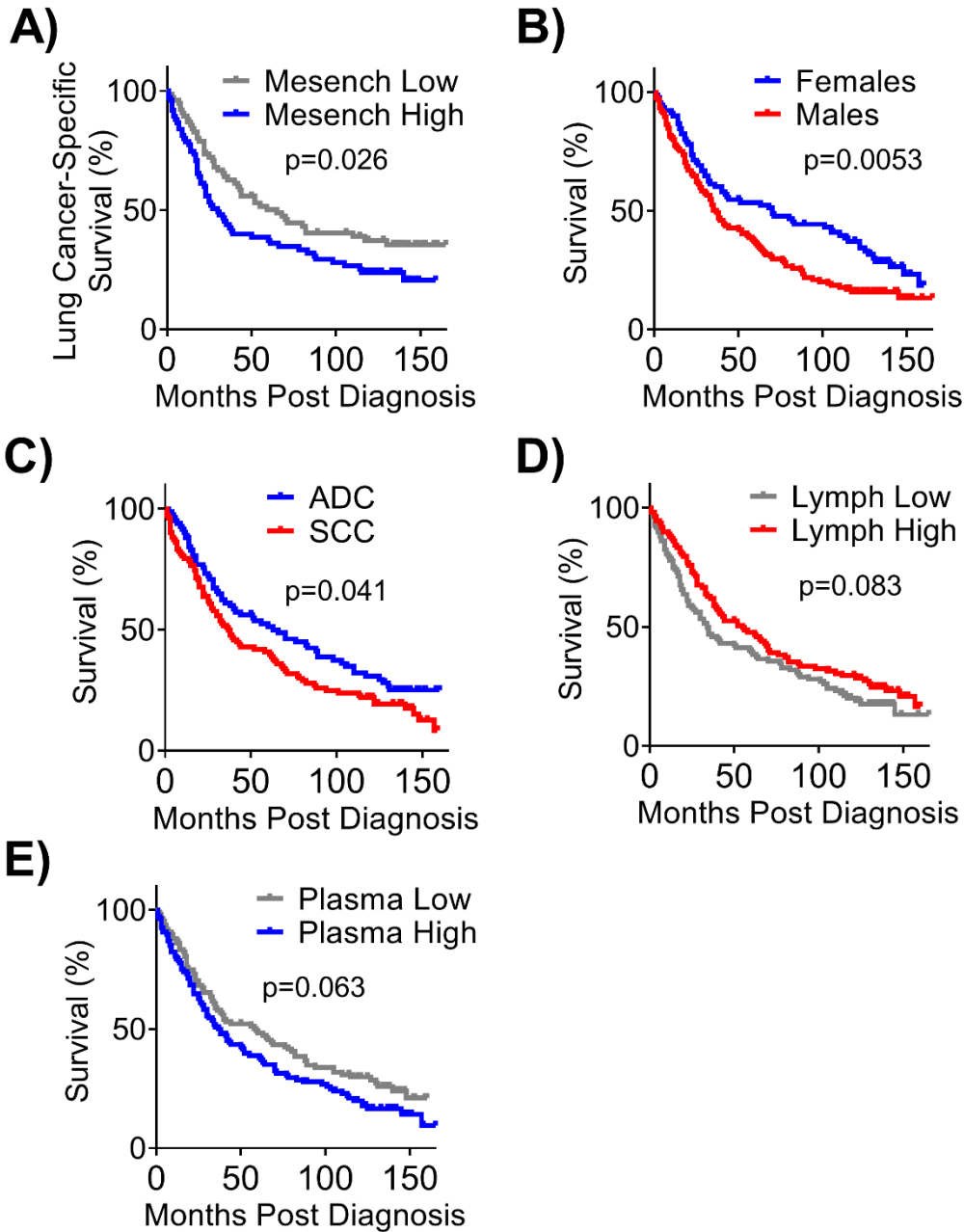
A+B) Bar graph depicting the percentage of lymphocytes identified by IHC quantification (**A**), or by the HALO AI nuclear phenotyper algorithm (**B**), mean +/- SEM is graphed. mean +/- SEM is graphed, n=8 mice per *Ezh2* genotype and 24 mice total, * indicates p<0.038 and ** indicates

$p < 0.013$ by one-way ANOVA with multiple comparisons and Holm-Šídák's multiple comparisons test. **C)** H&E-stained section of murine lymph node (left), with different cell types identified by the HALO AI nuclear phenotyper algorithm (right). **D)** H&E-stained section of murine lymph node with metastasized adenocarcinoma (left), with different cell types identified by the HALO AI nuclear phenotyper algorithm (right).



Supplemental Figure 2: Neutrophil infiltration predominates in genetically-defined murine lung squamous cell carcinomas

H&E-stained section of murine $KRAS^{G12D}/Lkb1$ -null tumors (left), with different cell types identified by the HALO AI nuclear phenotyper algorithm (right). Adenocarcinoma and squamous cell carcinoma in close proximity demonstrate the selective recruitment of neutrophils to the squamous cell carcinoma tumor (bottom), and the relatively more abundant macrophages in the adenocarcinoma tumor (top).



Supplemental Figure 3: Lymphocytes negatively correlate with neutrophils and mesenchymal cells, and mesenchymal cells predict poor prognosis in human samples

A-E) Kaplan-Meier survival plots between tumors divided into two groups using the indicated marker segregated at median, or segregated by histology, p value shown is Mantel-Cox LogRank test, n=216 patients for all plots, except n=185 for adenocarcinoma vs squamous cell carcinoma.

Supplementary Table 1 – Cox Regression Modeling of Key Phenotypes
Human Tissue Microarray n=185

Univariate Analysis								
	Est. Coeff.	Standard Error	Wald Statistic	Degrees of Freedom	p value	Odds Ratio	95.0% CI	
							Lower	Upper
Mesenchymal High vs Low	0.349	0.167	4.383	1	0.036	1.418	1.023	1.967
Sex Male vs Female	0.434	0.169	6.560	1	0.010	1.543	1.107	2.150
Age High vs Low	0.022	0.009	6.766	1	0.009	1.023	1.006	1.040
Histology ADC vs SCC	-0.342	0.169	4.089	1	0.043	0.711	0.510	0.990
Multivariate Analysis								
	Est. Coeff.	Standard Error	Wald Statistic	Degrees of Freedom	p value	Odds Ratio	95.0% CI	
							Lower	Upper
Mesenchymal High vs Low	0.236	0.169	1.946	1	0.163	1.266	0.909	1.763
Sex Male vs Female	0.408	0.172	5.604	1	0.018	1.504	1.073	2.109
Age High vs Low	0.023	0.009	7.076	1	0.008	1.023	1.006	1.040
Histology ADC vs SCC	-0.288	0.170	2.872	1	0.090	0.749	0.537	1.046