SUPPLEMENTARY FIGURES



Fig S1. A. eQTL-score schematic. SNPs alleles are oriented so that they are all in alignment in terms of direction of effects on gene expression. Number of alleles affecting gene expression are then summed into a continuous score **B.** Pearson correlation of Gene level eQTL-scores in TCGA and Discovery cohorts. P<=0.05 is indicated by an X.



Fig S2. A-C. Recursive feature elimination analysis. Mean squared error for each model is shown for a range of total number of features included in a given model. Optimal model size is indicated by a dashed red line. **D-E.** Pearson correlation matrices of features respective to germline and somatic models.



Fig S3. A-C. SHAP derived nonlinear feature importance beeswarm plots of germline, somatic, and composite models. **D-E.** Correlation of immunogenicity features with TMB. **F-G.** Coefficients of features from linear regression-based germline and somatic models

bioRxiv preprint doi: https://doi.org/10.1101/2024.01.12.575430; this version posted January 15, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Fig S4. A-C. Pearson correlation of three different IC-Index scores with TMB. **D.** Purity vs ploidy across all patients, colored by response with pearson correlation shown. **E-G.** ROC plots comparing the performance of composite IC-Index to TMB and clinical predictors of ICB response. **H-I.** ROC comparing the performance of composite IC-Index to transcriptomic predictors of ICB response **J.** Histogram of composite IC-Index across patients, colored by ICB response status.



Fig S5. (A-C). Measures of TIME infiltrates stratified by high vs low composite IC-Index scores. Ttests were used to compare means between groups. **(D-E)** Confusion matrices of somatic and germline IC-Indices (cutoff >=5) vs TIME score (cutoff=median).



Fig S6. Difference in feature importance rankings of composite model features in linear (left) versus nonlinear (right) models.



Fig S7. A. Response rates by reliance groupings across discovery patients. Chi squared tests were performed between groups. **B.** T_{FH} cell infiltration estimates stratified by MHC-reliance grouping and ICB response. Mann-Whitney U tests were used to compare means between groups. **C.** Waterfall plot of ratio of well-presented MHC-II to MHC-I neoantigens across discovery patients. **D.** MHC-I presentation pathway damage as a proportion of total number of mutations split by MHC Reliance grouping. T-tests were used to compare groups. **E.** Proportion of patients with any damage to MHC-I antigen presentation pathway split by MHC Reliance grouping. Chi squared tests were performed between groups.



Fig S8. A. MHC-I reliant patients in TCGA with melanoma, non-small cell lung cancer or renal cell carcinoma stratified by CD4/CD8 T Cell infiltration ratio. **B.** MHC-II reliant patients in TCGA with melanoma, non-small cell lung cancer or renal cell carcinoma stratified by CD4/CD8 T Cell infiltration ratio. P Values generated via log-rank test



Fig S9. A. Univariate analysis of potential checkpoint inhibitors in the discovery samples. **B.** Proportion of patients responding in each MHC reliance group when discovery samples are partitioned according to low versus high expression of 7 different checkpoint genes. **C.** Univariate analysis of potential checkpoint inhibitors in the validation samples. **D.** Proportion of patients responding in each MHC reliance group when validation samples are partitioned according to low versus high expression of 7 different checkpoint inhibitors in the validation samples. **D.** Proportion of patients responding in each MHC reliance group when validation samples are partitioned according to low versus high expression of 7 different checkpoint genes



Fig S10. A. Checkpoint expression and neoantigen levels by MHC Reliance category in discovery samples. **B.** Checkpoint expression and neoantigen levels by MHC Reliance category in validation samples. **C.** Pearson correlation between various checkpoint genes across discovery samples. **D.** Pearson correlation between various checkpoint genes across validation samples.



Fig S11. A-B. Heatmap of CIBERSORTx derived immune infiltration estimates for discovery **(A)**, and validation **(B)** samples split by response and MHC Reliance category. Mann-Whitney U tests were performed between responders and nonresponders of each MHC Reliance category. P value of <=0.05 are marked with an asterisk.