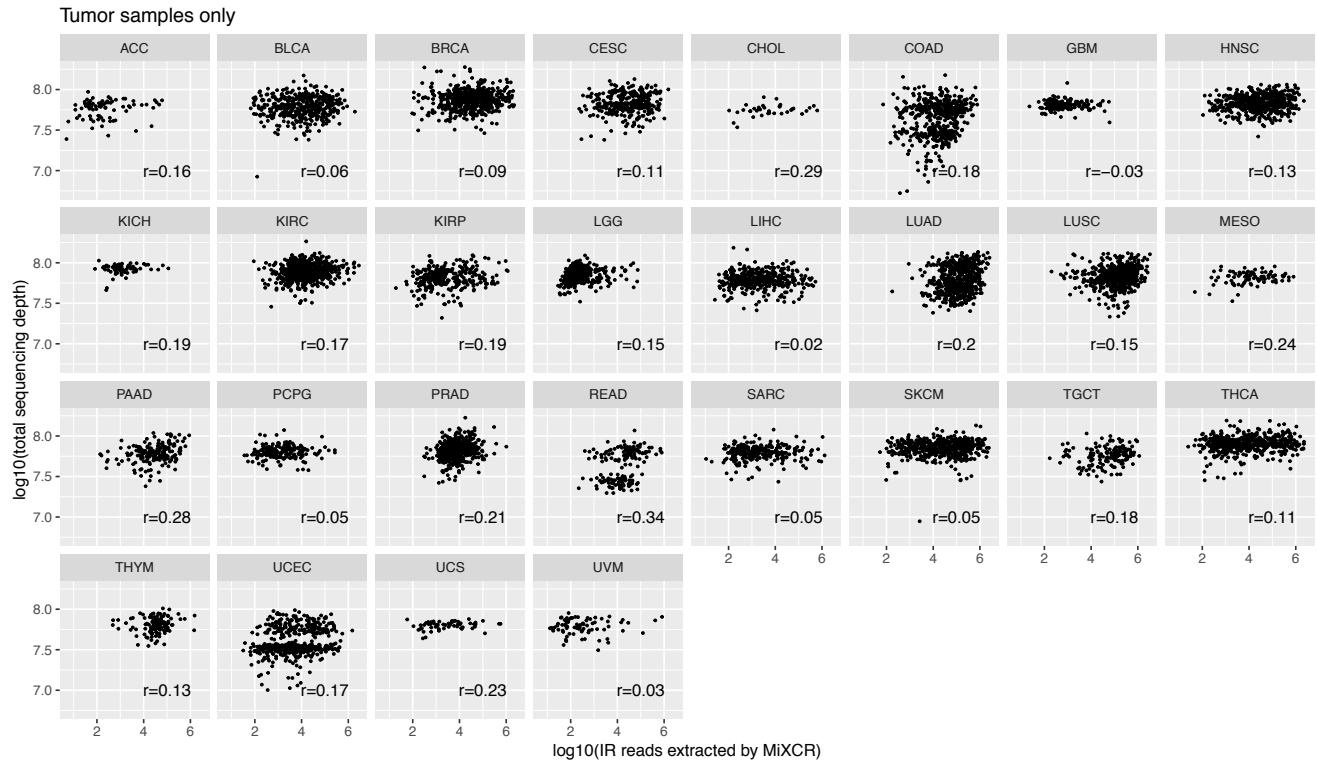
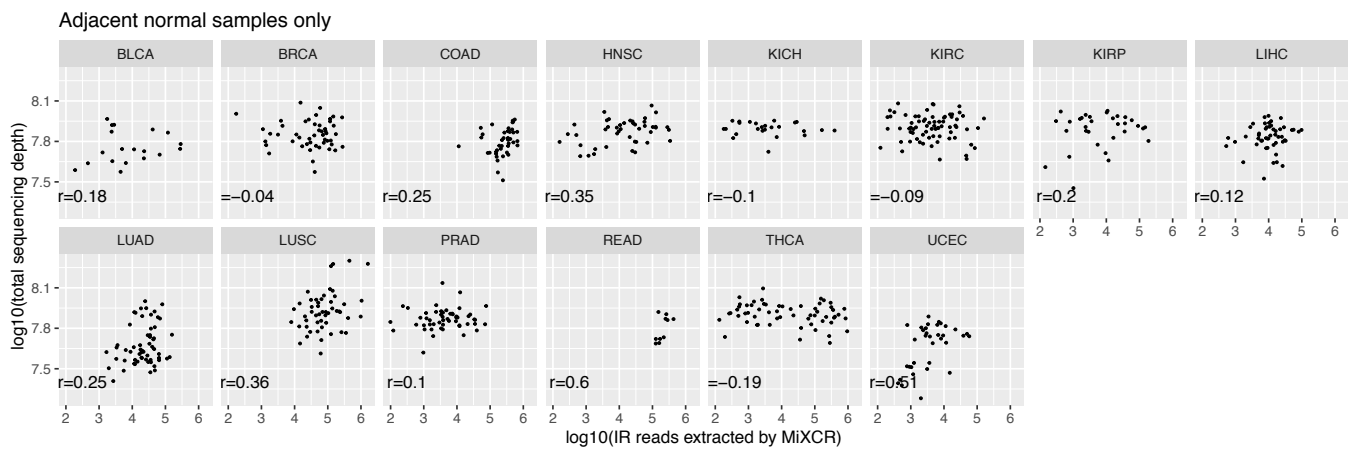


Supplementary Figure 1

A

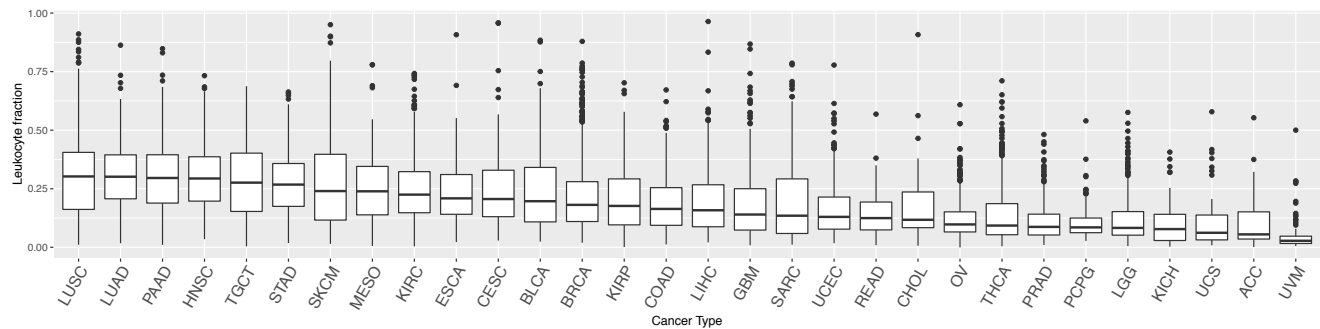


B



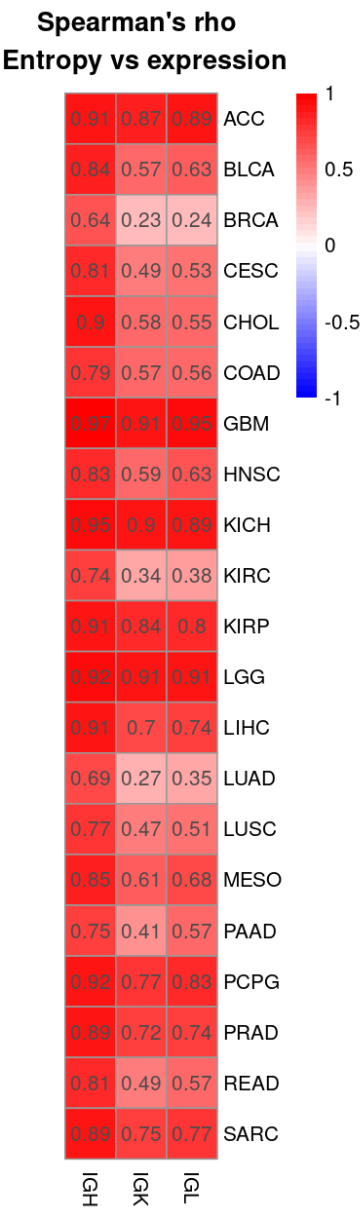
(a) Plot showing the number of BCR reads extracted by MiXCR (x-axis) and the total number of reads in each sample (y-axis) for the tumor samples in each tumor type. (b) Plot showing the number of BCR reads extracted by MiXCR (x-axis) and the total number of reads in each sample (y-axis) for the adjacent normal samples in each tumor type (only tumor types with at least 10 adjacent normal samples are shown).

Supplementary Figure 2



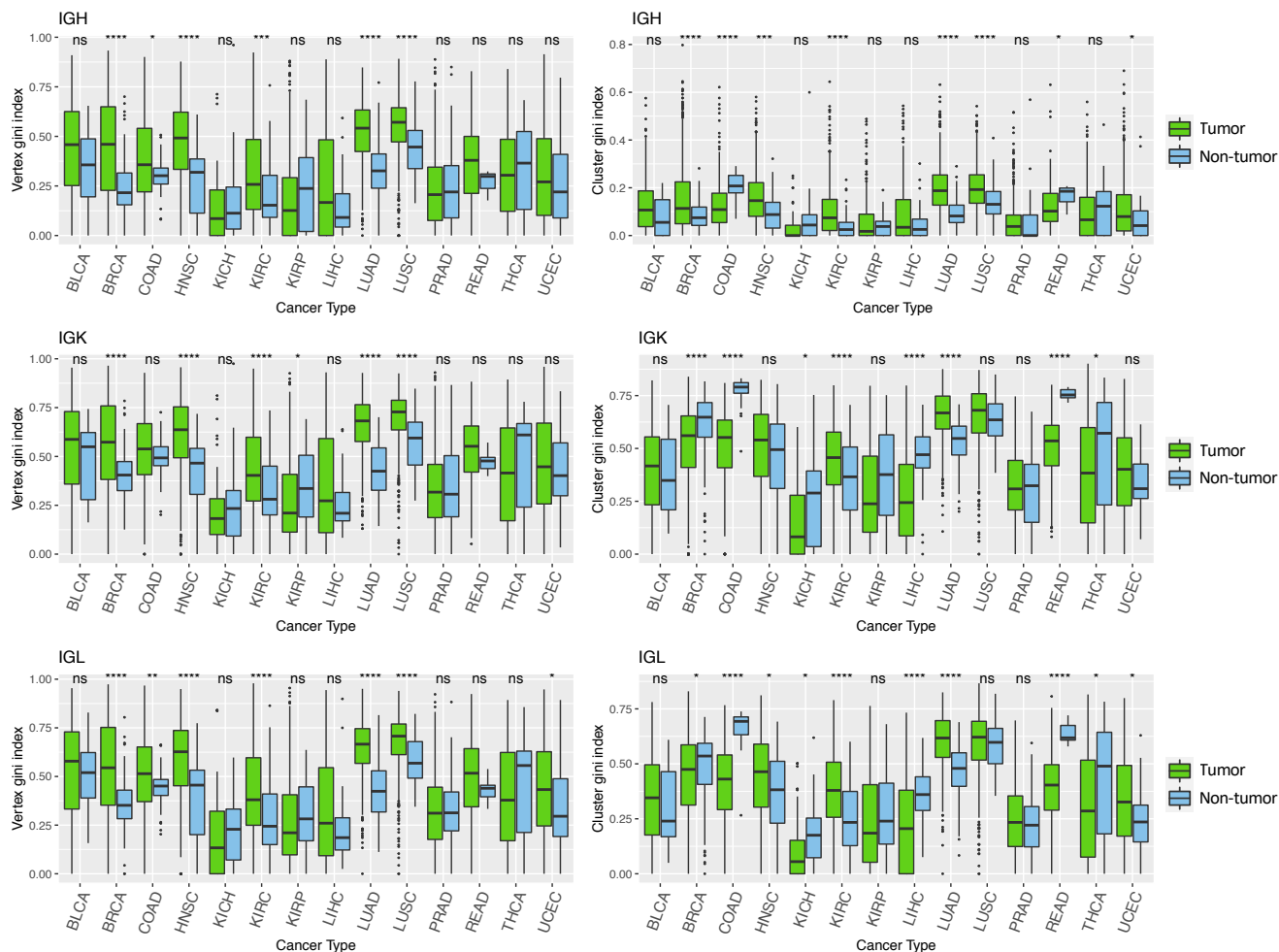
Boxplots showing the leukocyte fraction of each sample across tumor types. Tumor types are ordered by median leukocyte fraction from high to low.

Supplementary Figure 3



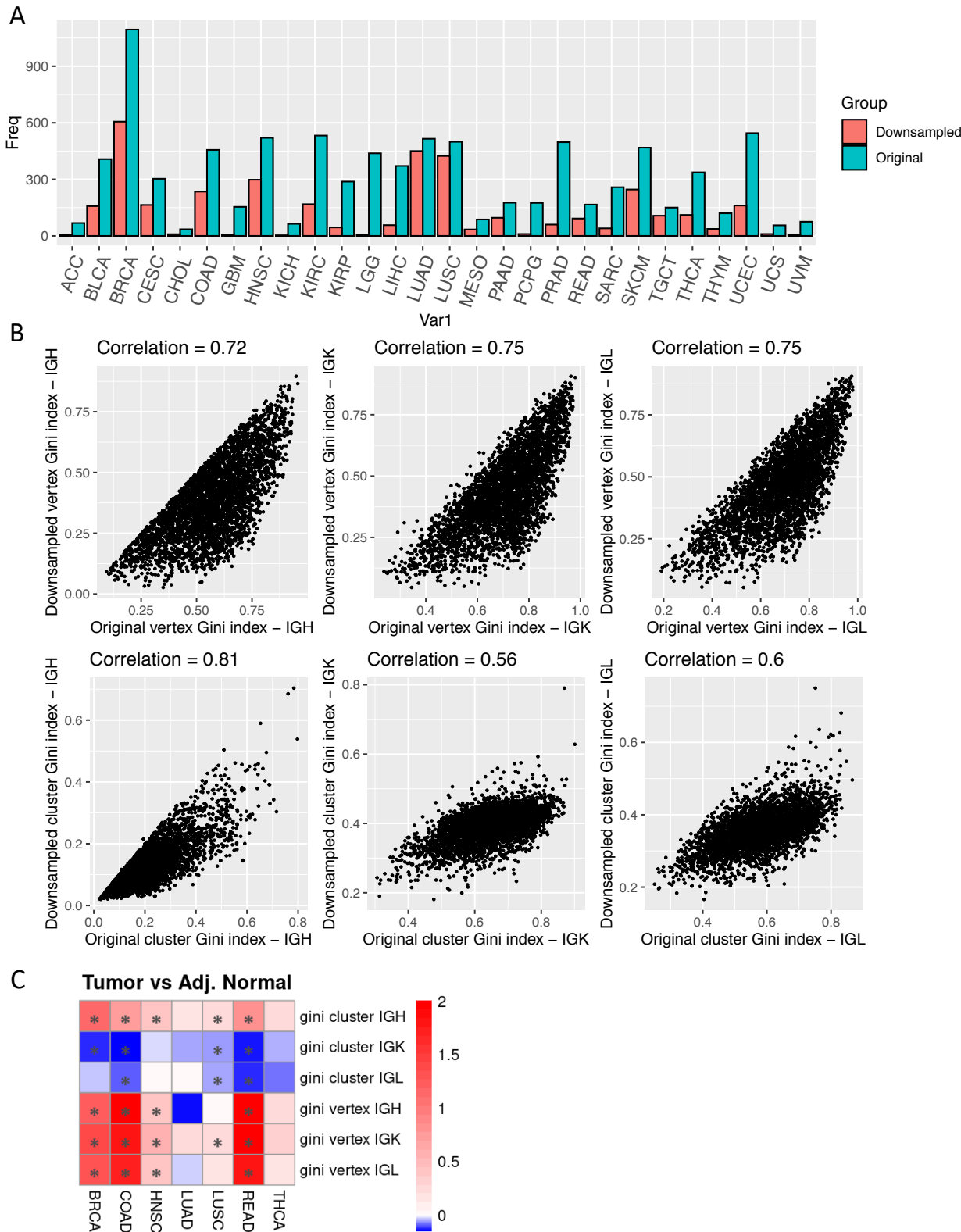
Heatmap showing Spearman’s correlation coefficient for Shannon entropy and expression (e.g. the number of IGH/IGK/IGL reads divided by the total number of reads in the sample). The value in each cell is the correlation coefficient.

Supplementary Figure 4



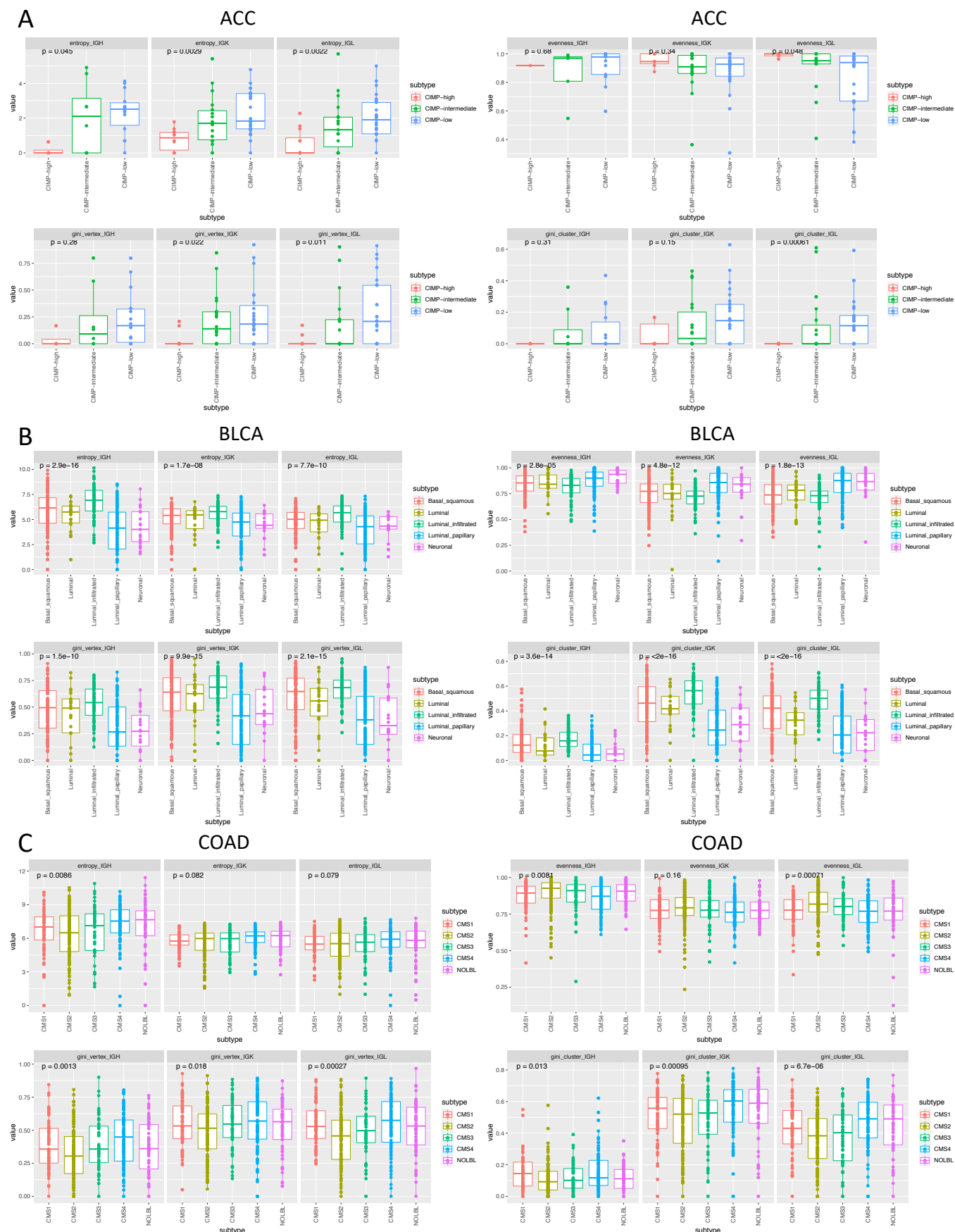
The boxplots show the vertex/cluster Gini indexes for tumors (green) and adjacent non-tumor samples (blue) for the 14 tumor types with at least 10 adjacent non-tumor samples. Statistical significance was calculated using the Wilcoxon rank-sum test. Significant p-values are indicated by symbols above the box plots with one star corresponding to $p\text{-value} \leq 0.05$, two stars corresponding to $p\text{-value} \leq 0.01$, three stars corresponding to $p\text{-value} \leq 0.001$, and four stars corresponding to $p\text{-value} \leq 0.0001$.

Supplementary Figure 5



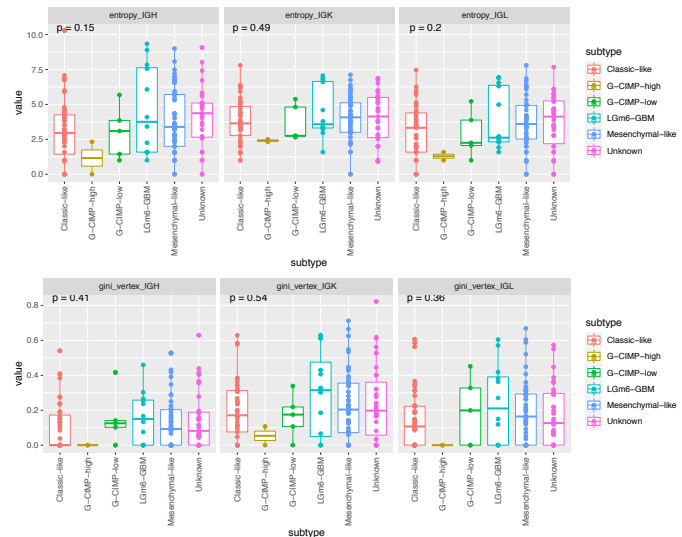
(a) Barplot showing the number of samples used in the original analysis (turquoise) and in the downsampled analysis (red). The downsampled analysis removed samples with fewer than 500 IGH, IGK, and IGL reads. (b) Plots showing the original vertex or cluster Gini indexes (x-axis) versus the downsampled indexes (y-axis). (c). Heatmap showing the log2 fold ratio between the mean tumor value and the mean adjacent normal value. Statistical significance was calculated using the Wilcoxon rank-sum test and comparisons with FDR < 0.05 are marked by an asterisk.

Supplementary Figure 6

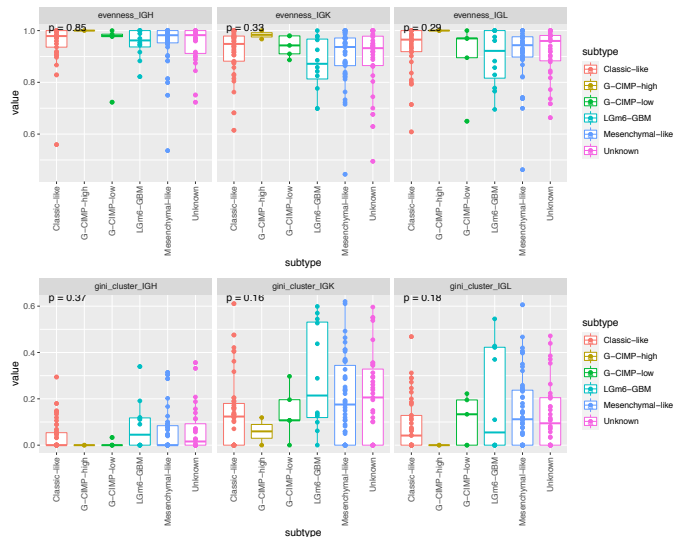


D

GBM

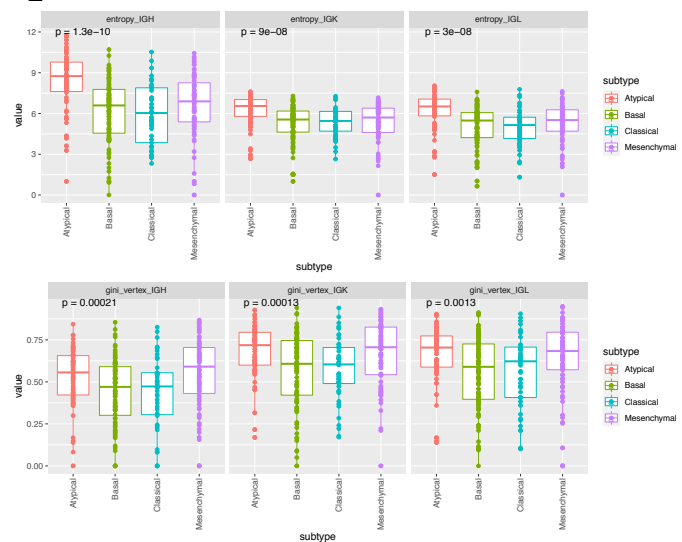


GBM

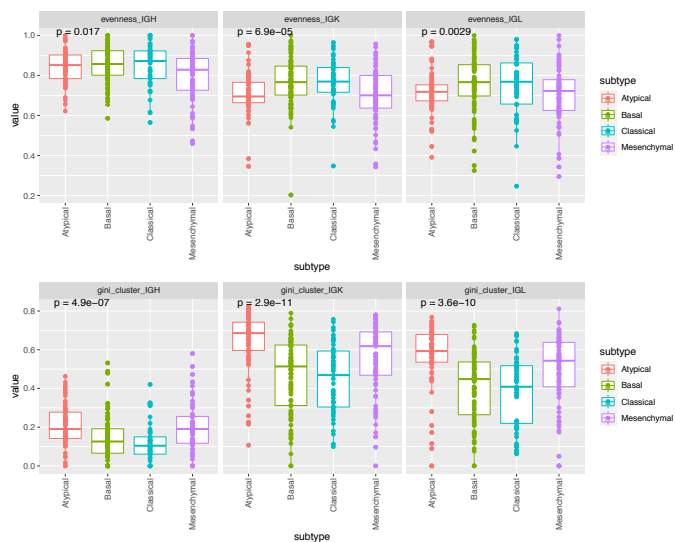


E

HNSC

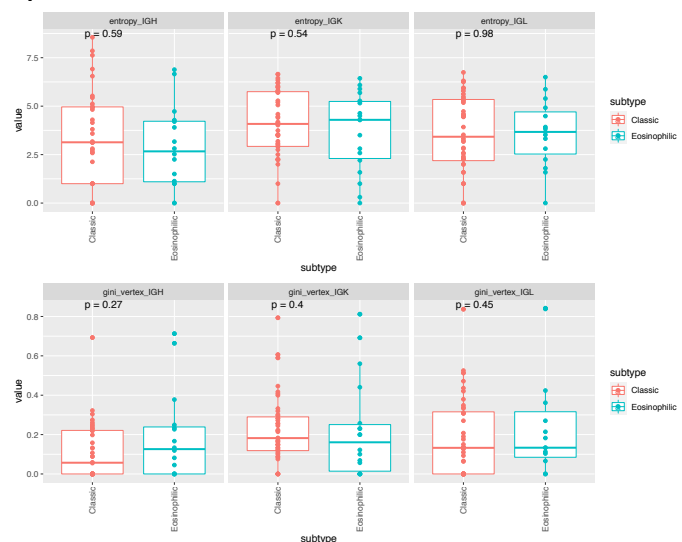


HNSC

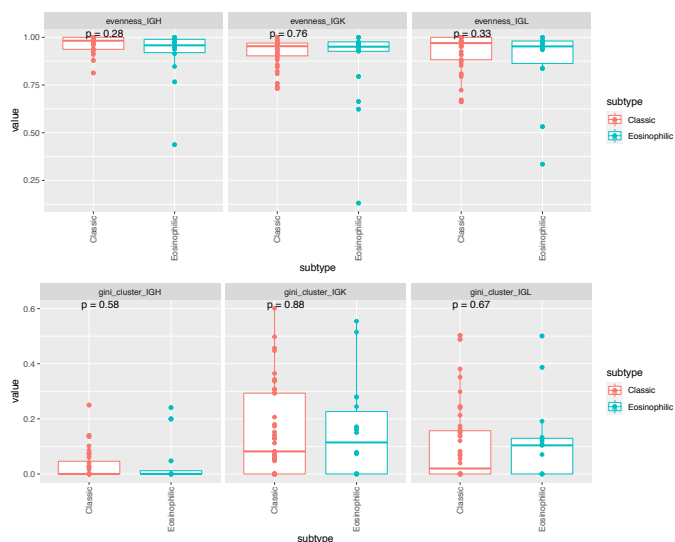


F

KICH

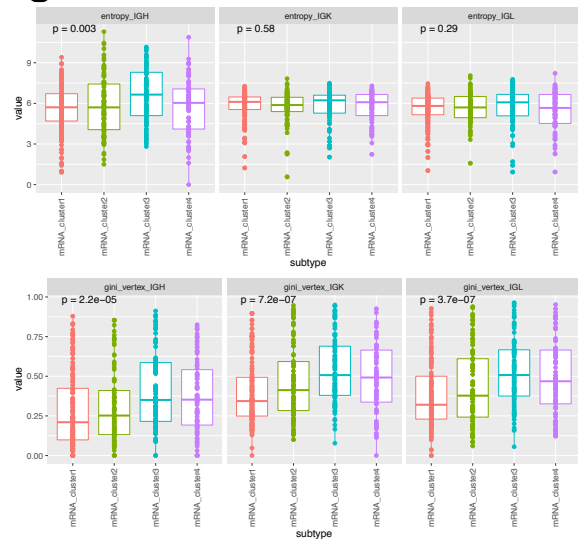


KICH

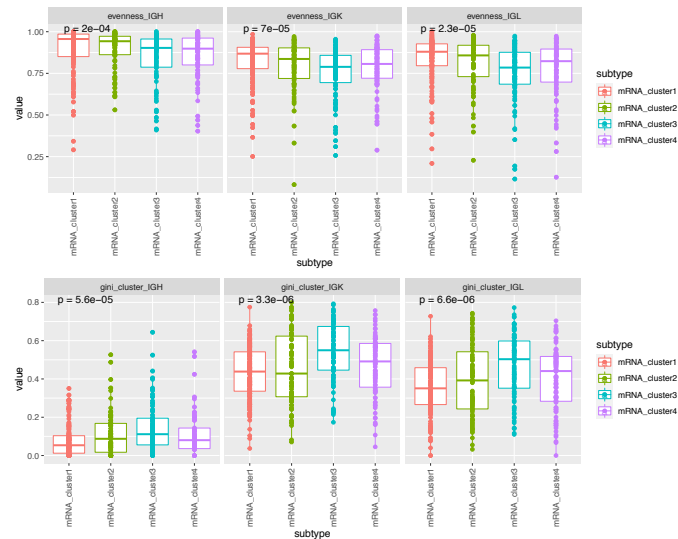


G

KIRC

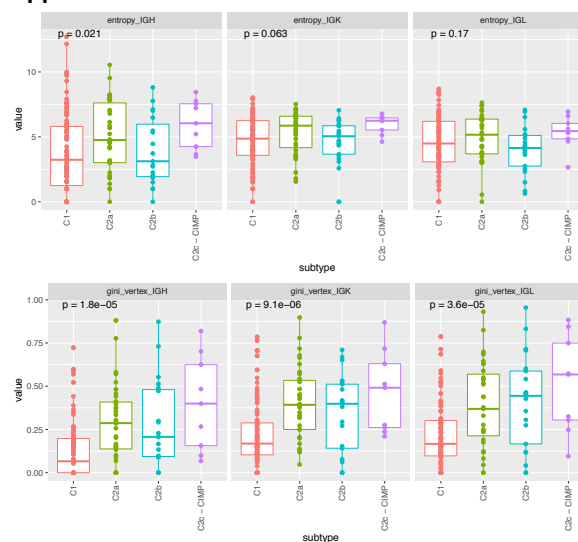


KIRC

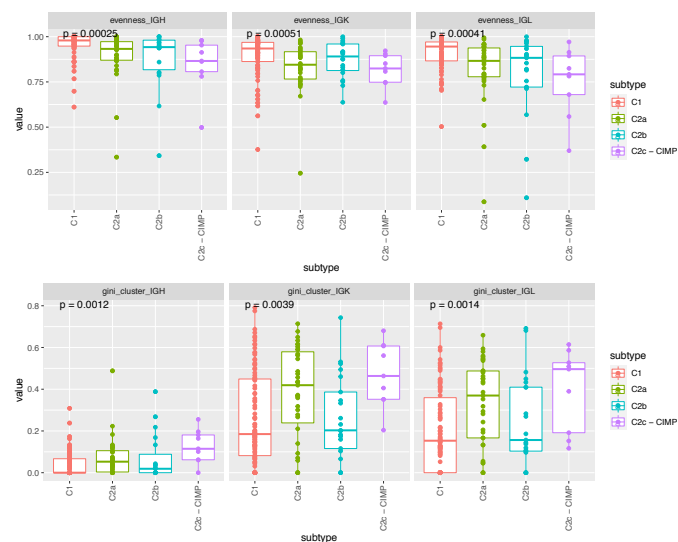


H

KIRP

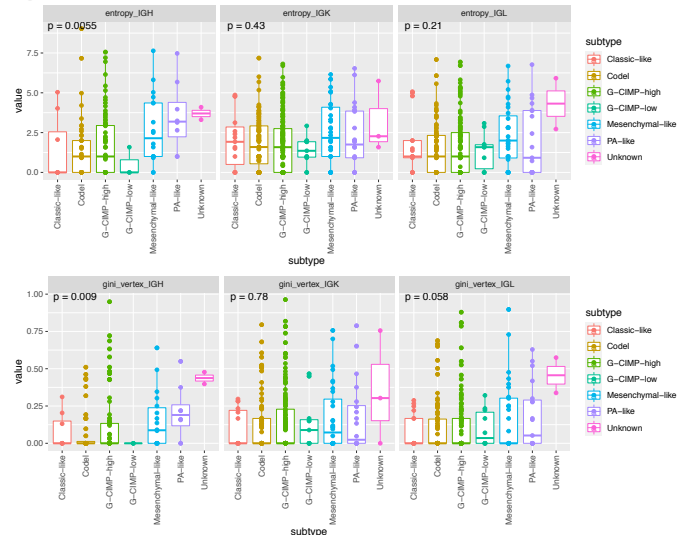


KIRP

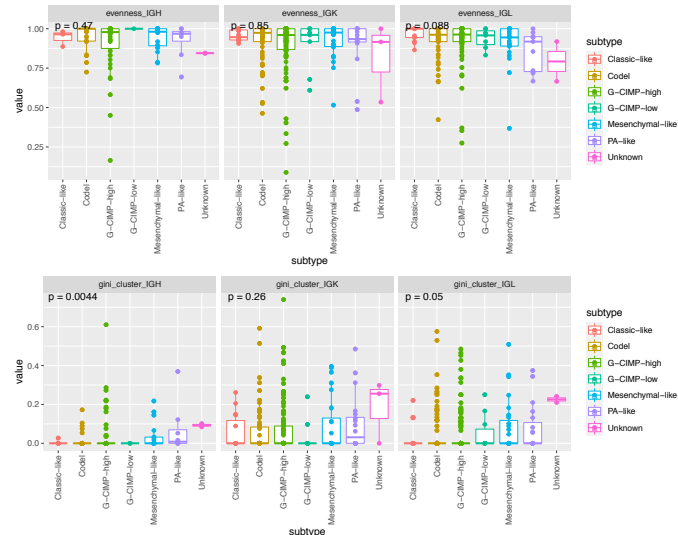


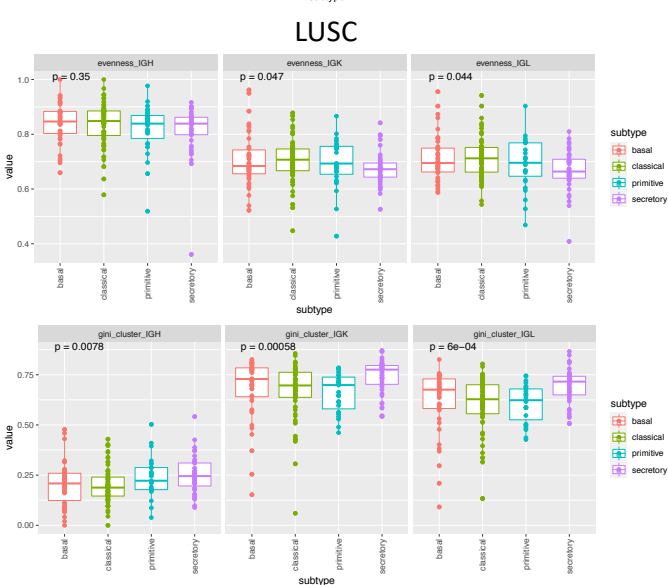
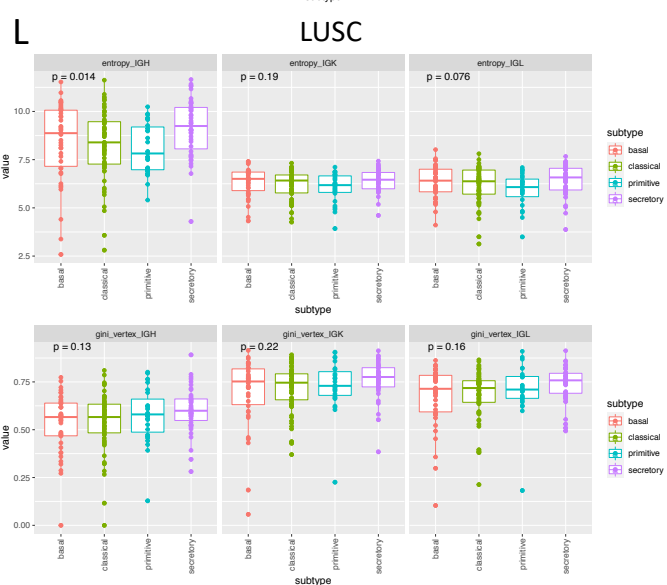
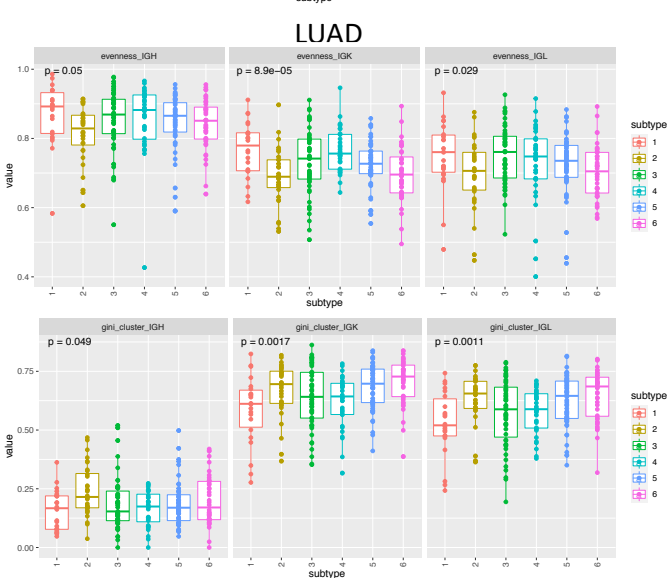
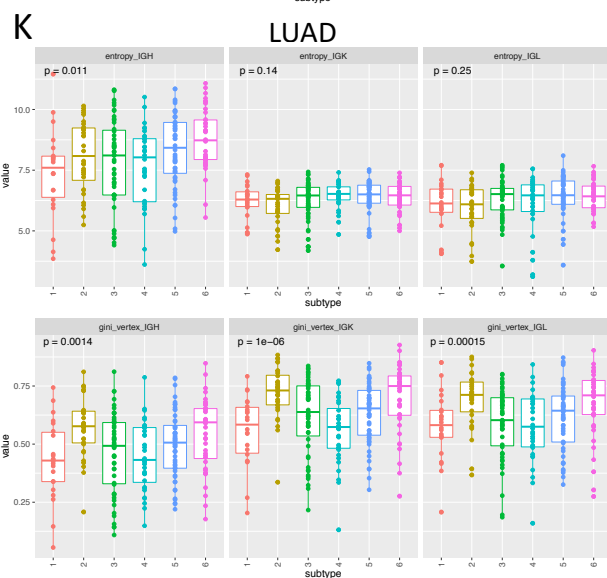
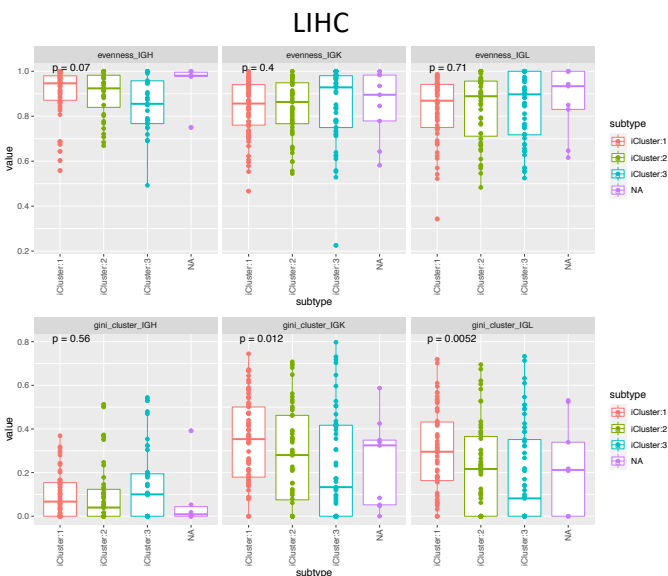
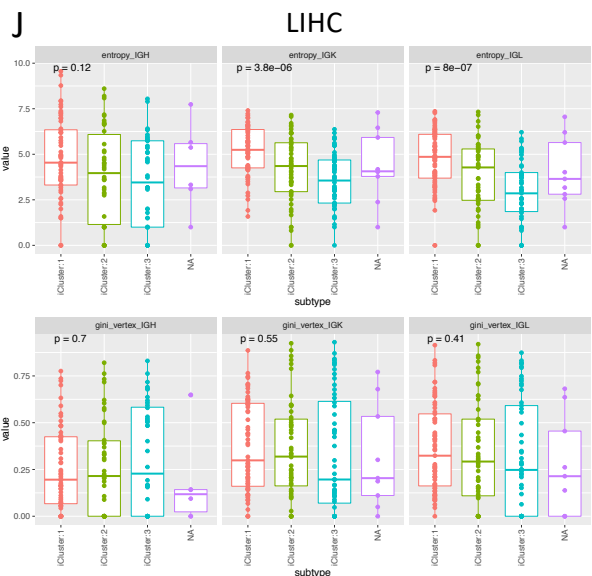
I

LGG



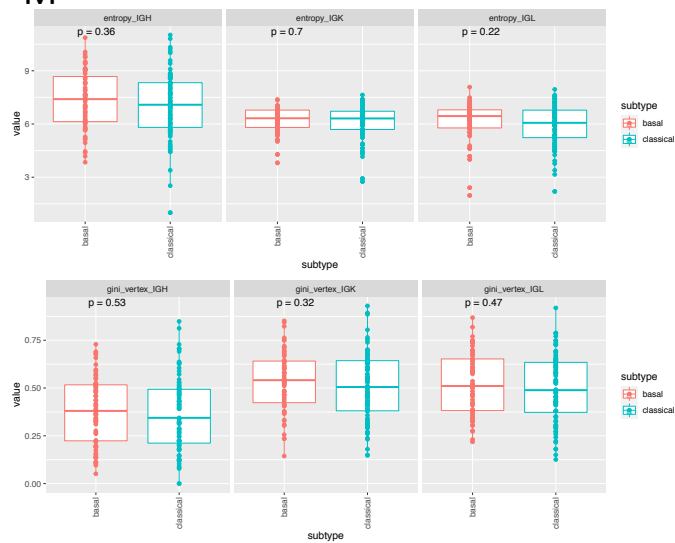
LGG



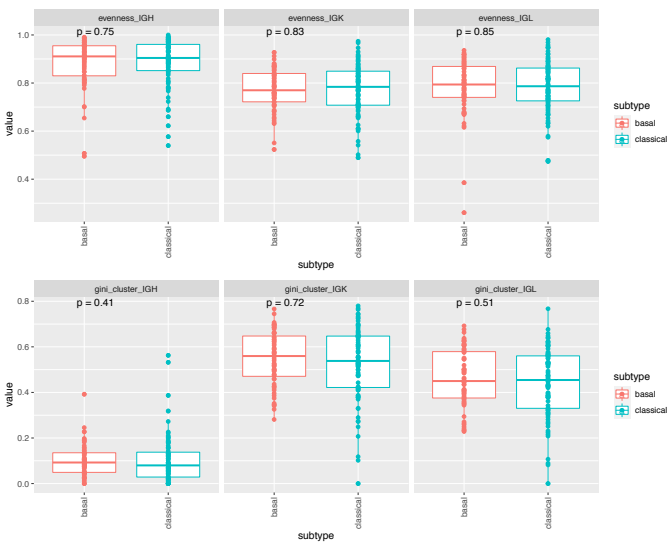


M

PAAD

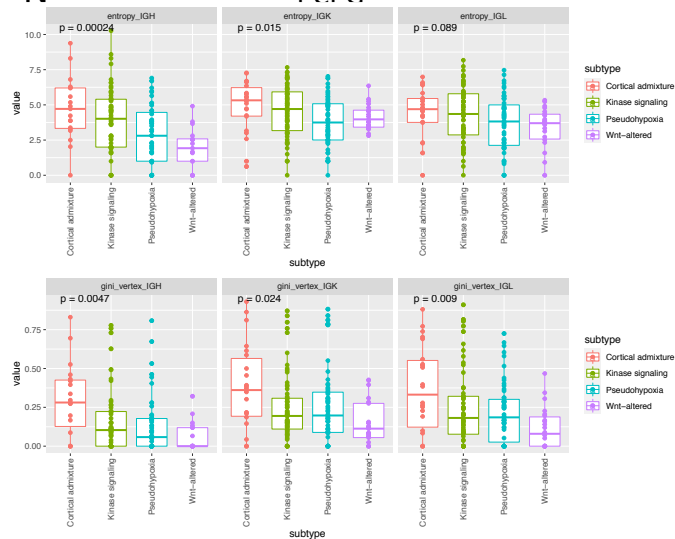


PAAD

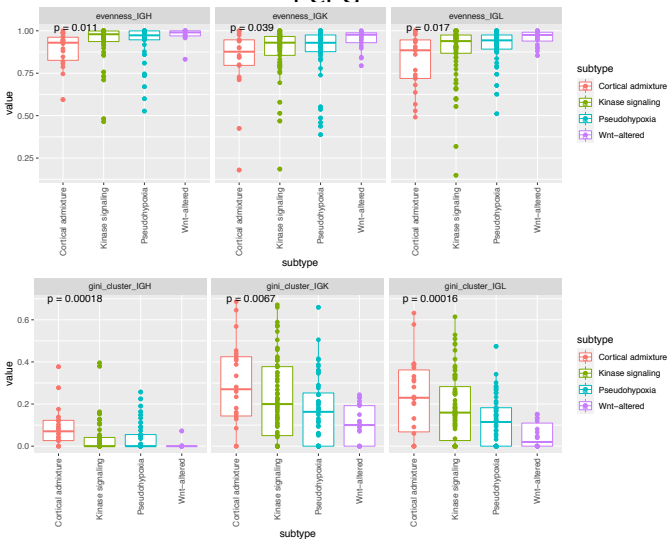


N

PCPG

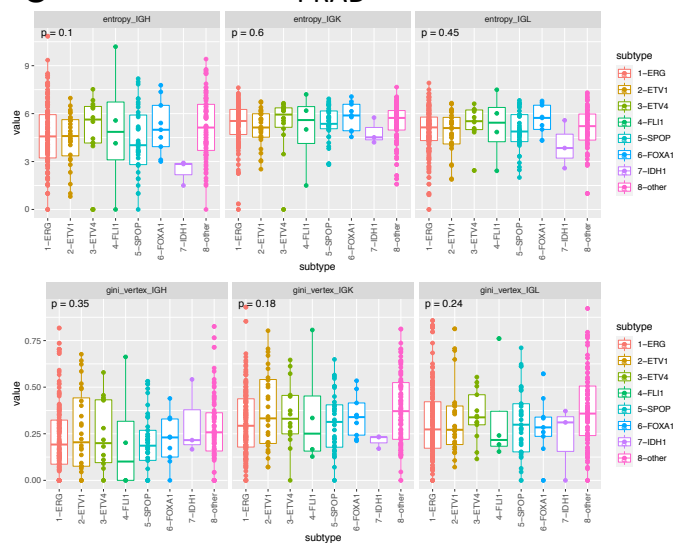


PCPG

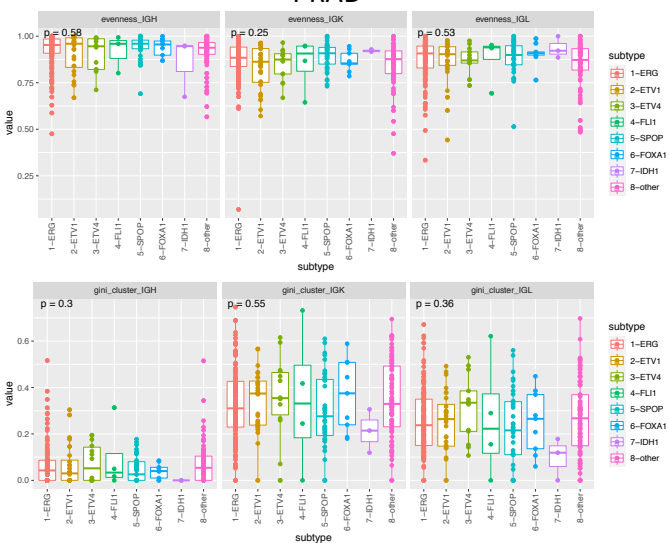


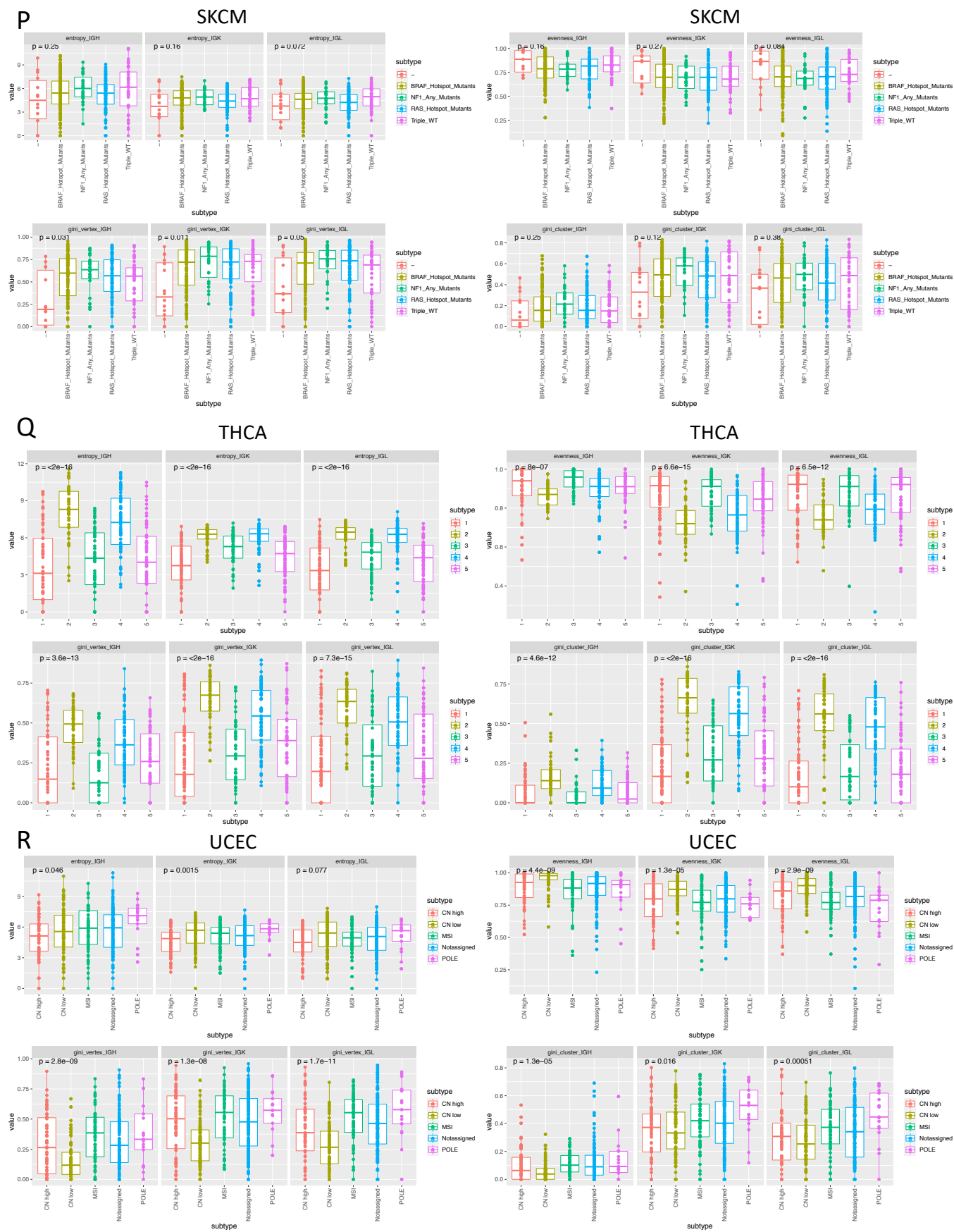
O

PRAD



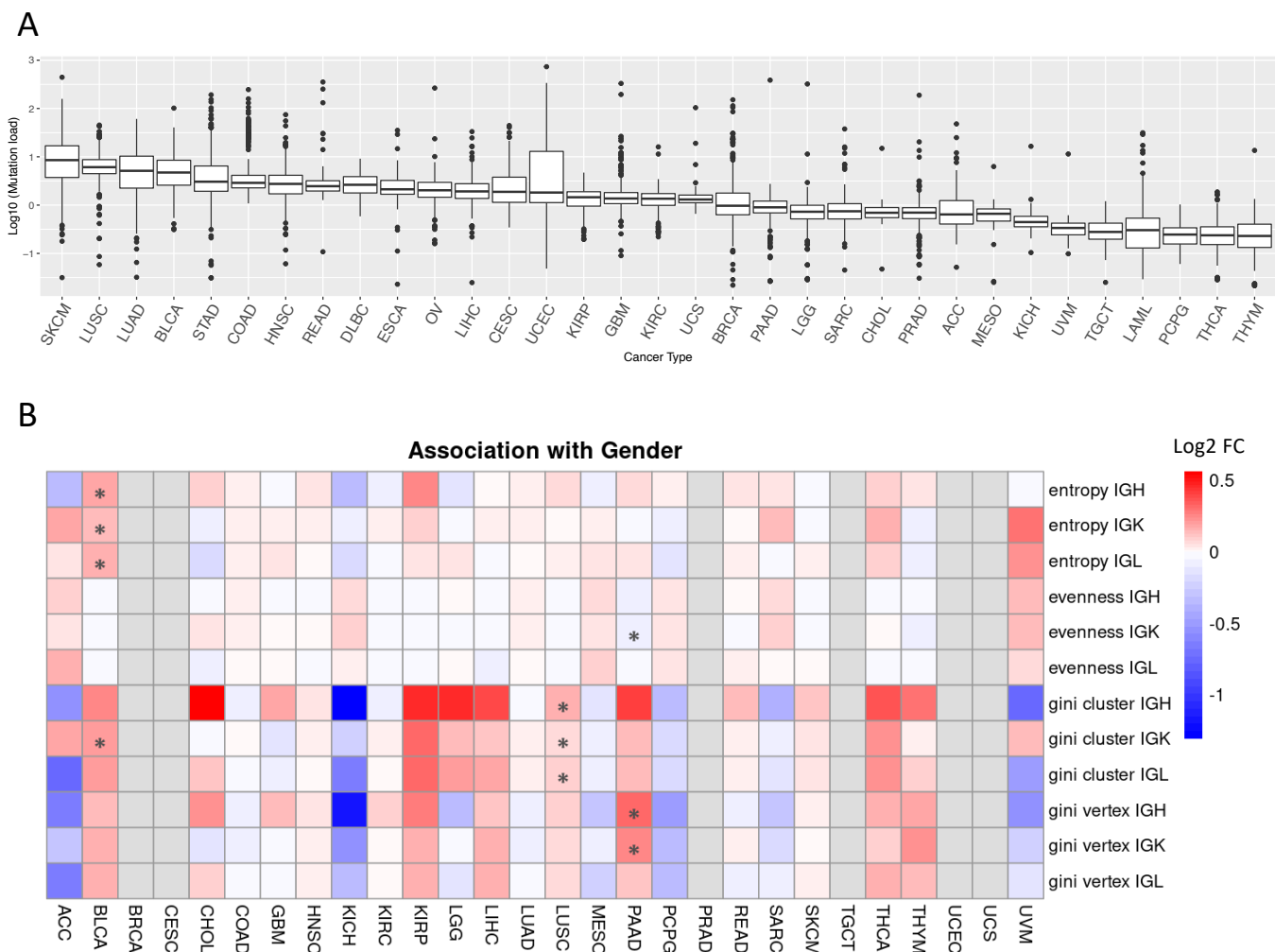
PRAD





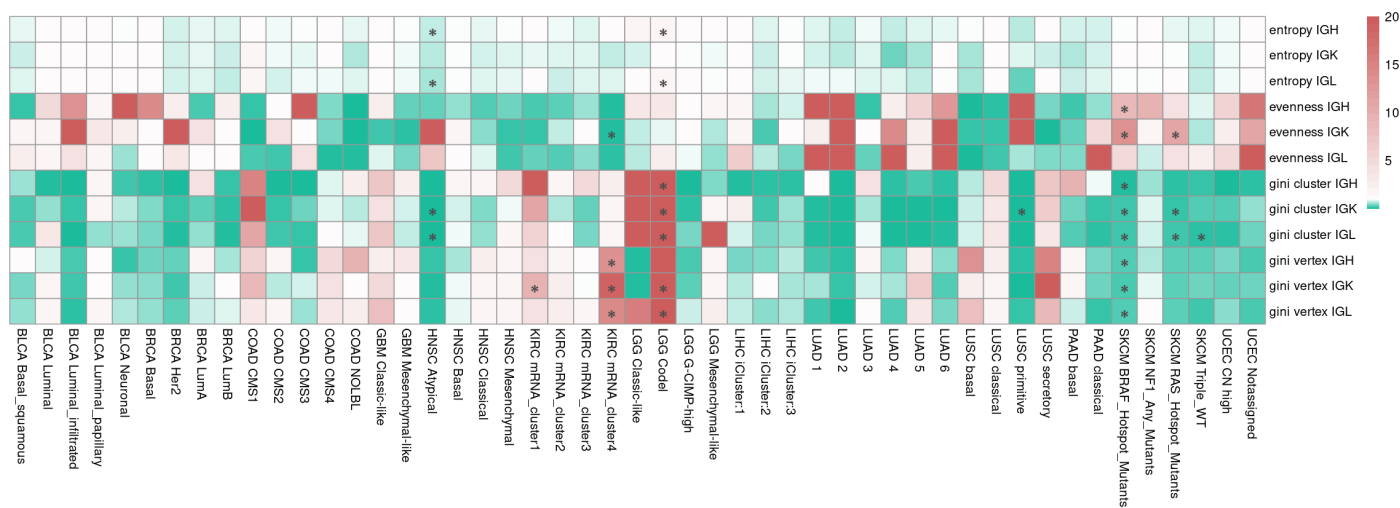
(a)-(r) Boxplots comparing the B cell repertoire features between tumor subtypes.

Supplementary Figure 7



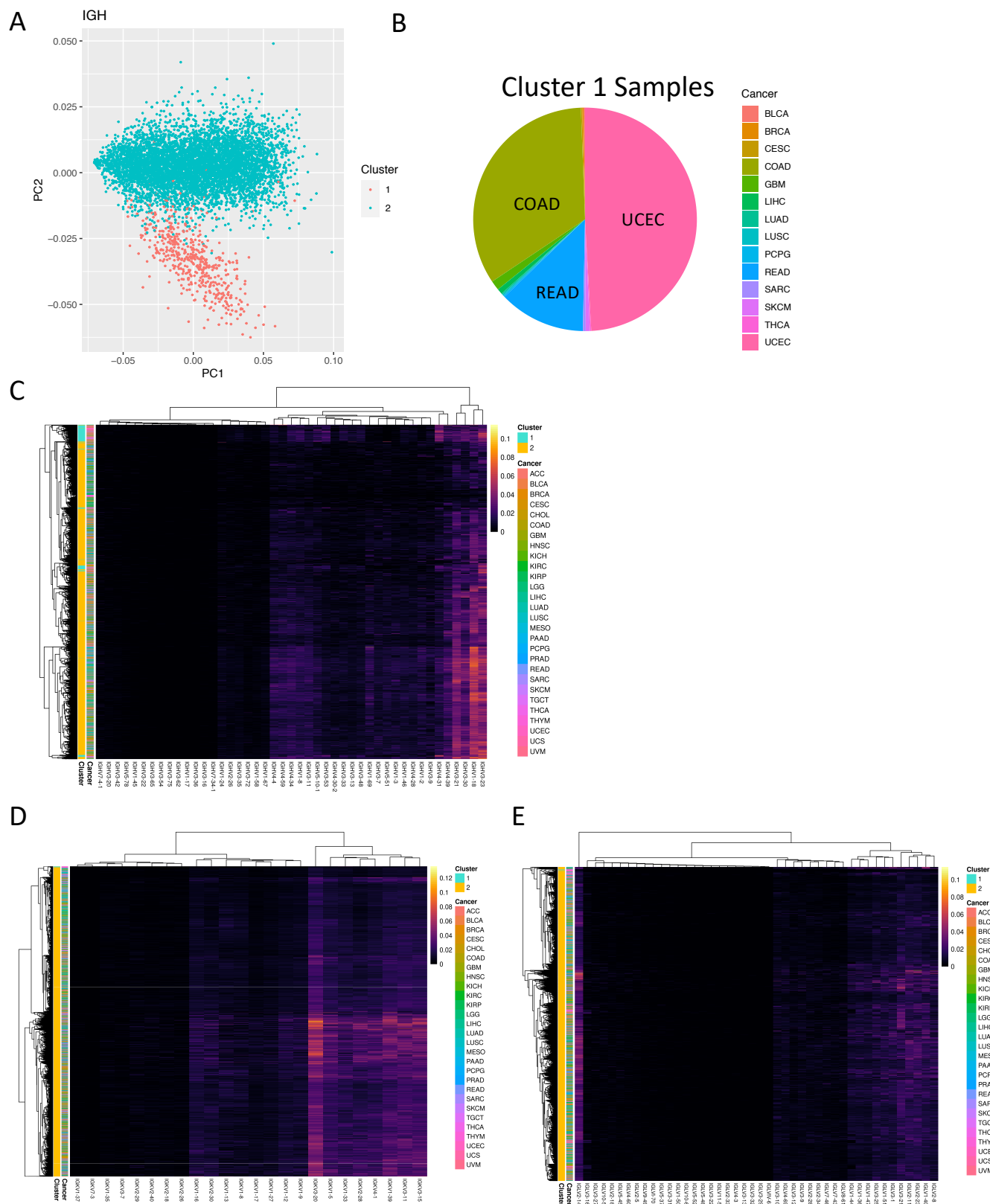
(a) Boxplots of mutation load across tumor types. (b) Heatmap showing the log2 fold ratio between the mean value in females and the mean value in males. The Wilcoxon rank-sum test was used to calculate significance and significant comparisons with FDR < 0.05 are marked by an asterisk.

Supplementary Figure 8



Heatmap showing the hazard ratio from a Cox proportional hazards model for each B cell repertoire feature. Columns are individual tumor subtypes. Red indicates a hazard ratio greater than 1 and green indicates a hazard ratio less than 1. Significant associations (FDR < 0.05) are marked by an asterisk.

Supplementary Figure 9



(a) PCA plot of IGH V gene usage. Samples are colored by the clusters identified by k-means clustering (k=2). (b) Pie chart showing the proportion of each tumor type in cluster 1. (c) Heatmap of IGH V genes associated with cluster 1. (d) Heatmap of IGK V genes associated with THYM samples. (e) Heatmap of IGL V genes associated with THYM samples.