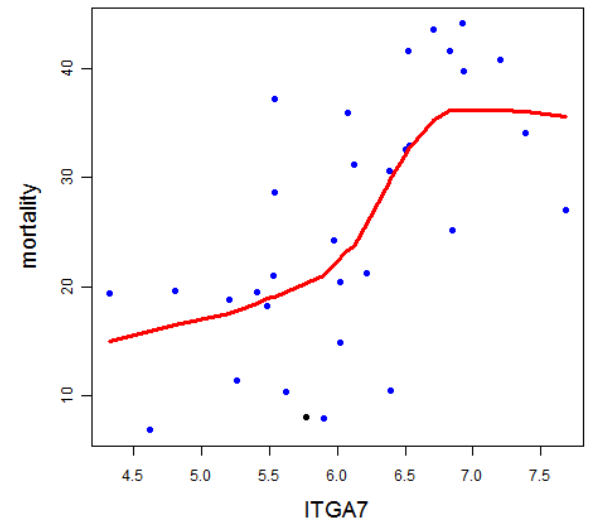
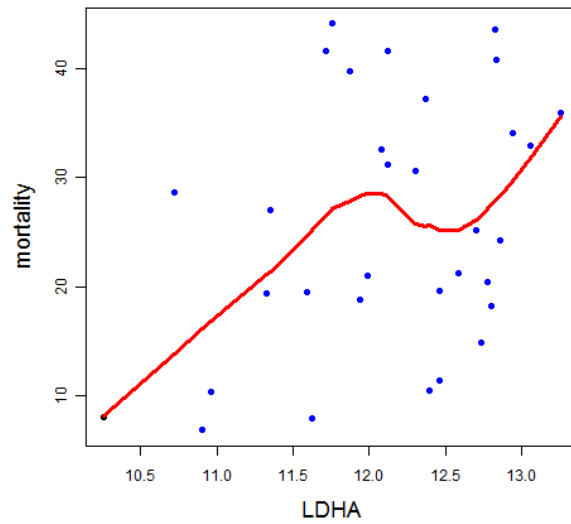
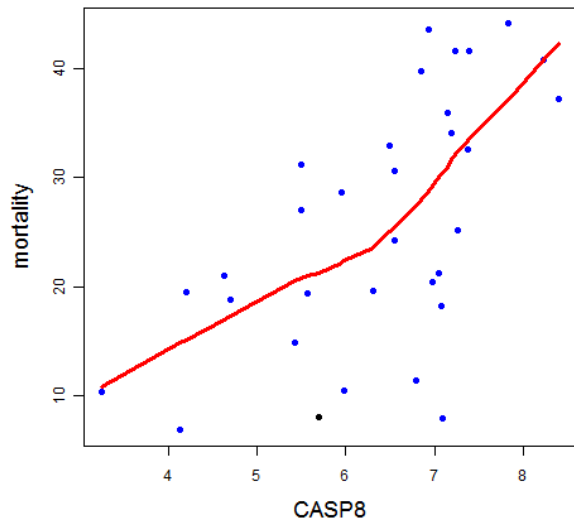
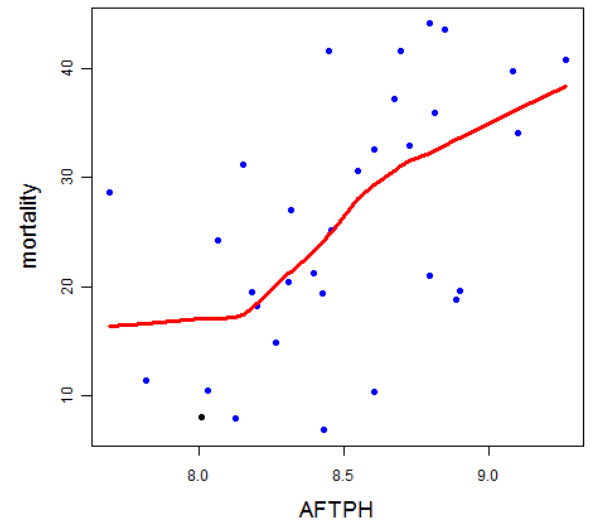
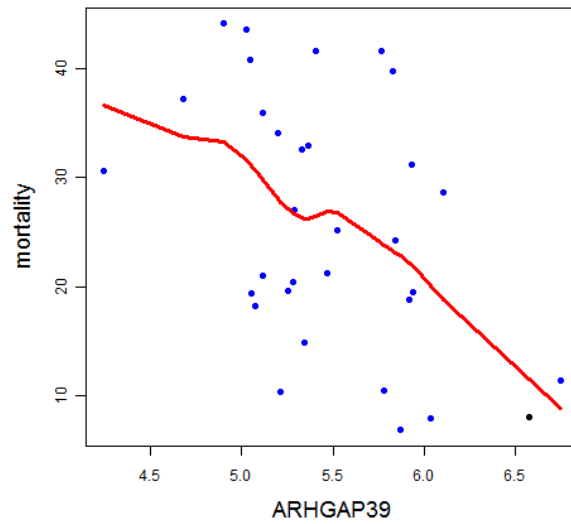
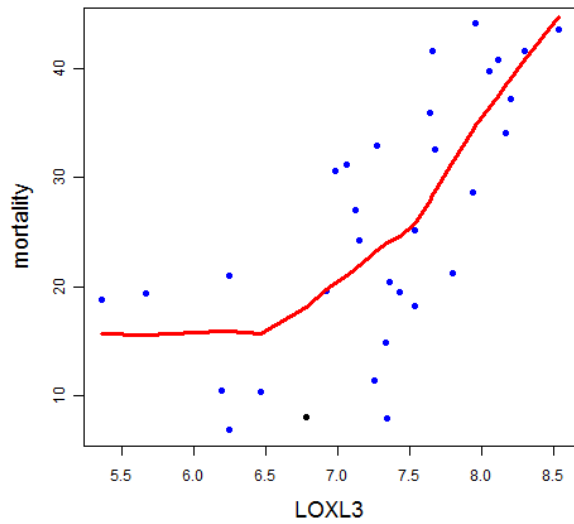
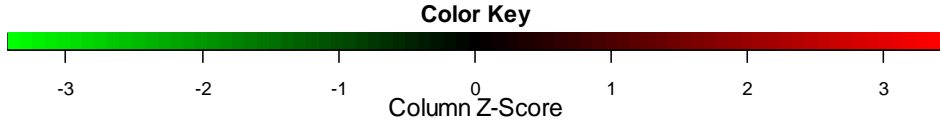
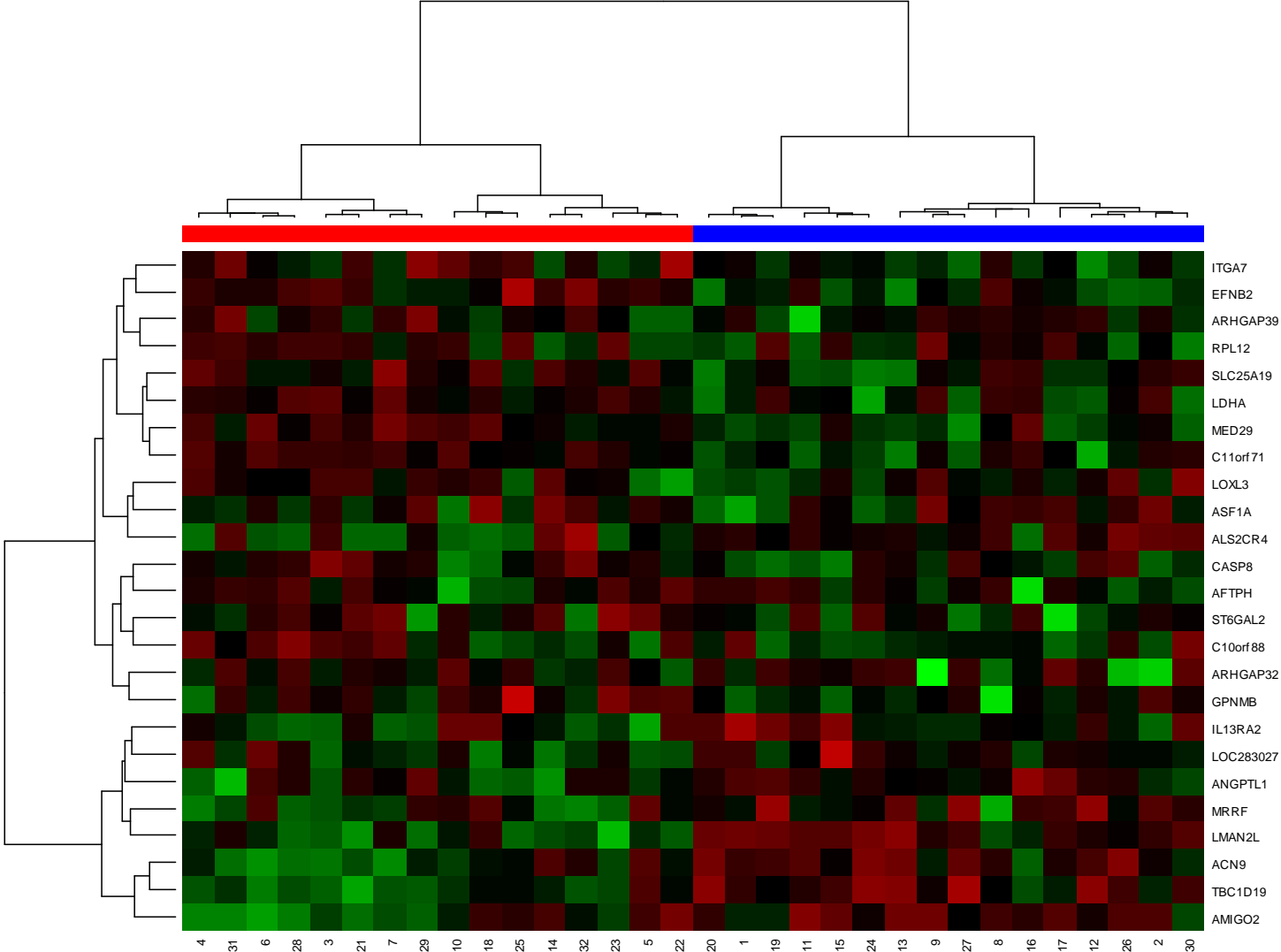


# Supplementary Fig. S 1

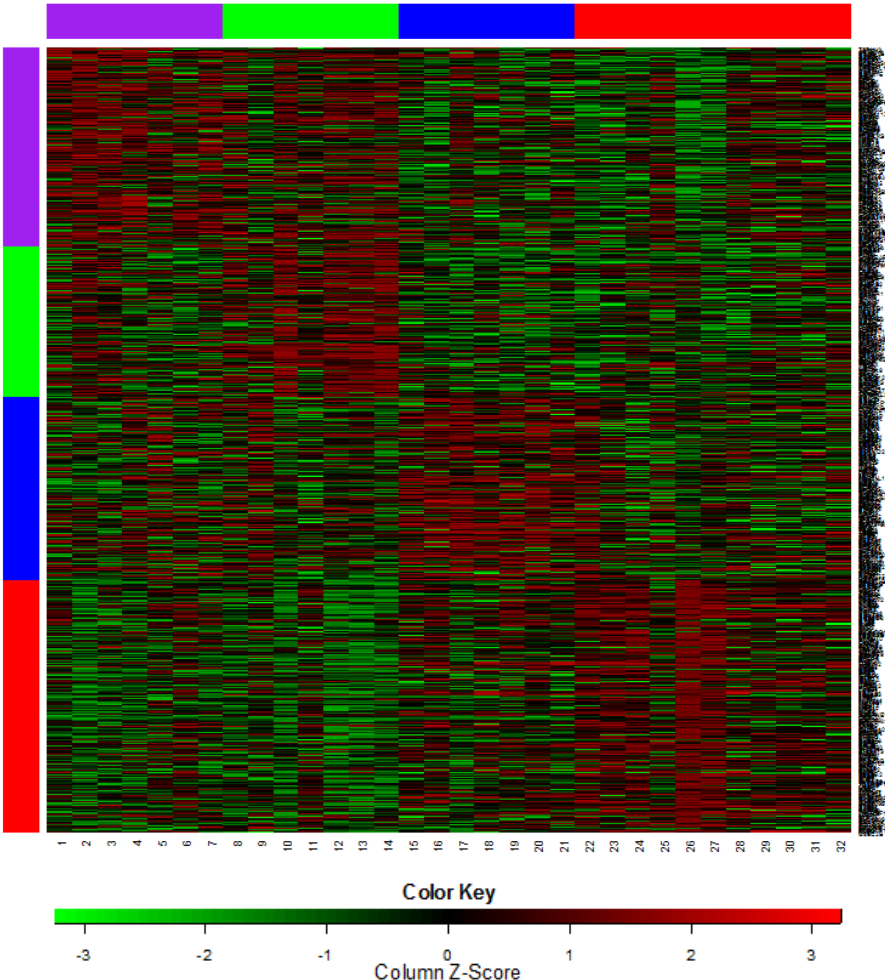


# Supplementary Fig.S2

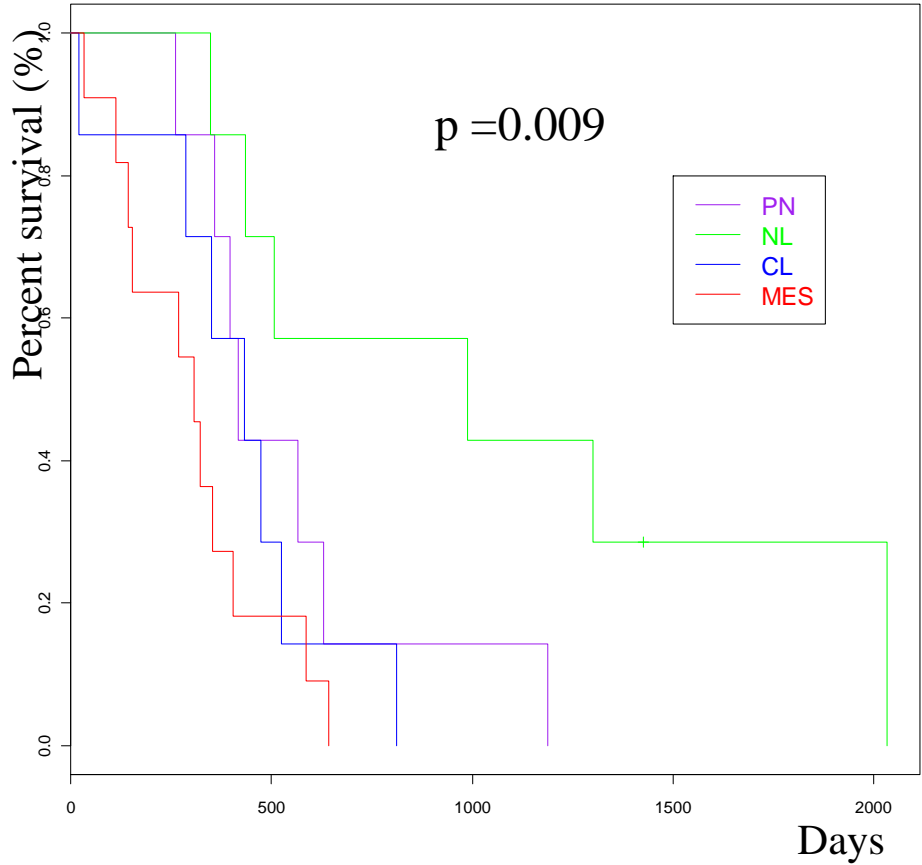


# Supplementary Fig.S3

(A)



(B)



**Supplementary Figure S1: Marginal plots for the six genes.** Values on the vertical axis represent the expected number of deaths for a given gene expression; the horizontal axis represents the gene expression value. Each point corresponds to an individual; patients with higher values on the vertical axis have higher risks. Lines represent a LOWESS (locally weighted scatter plot), with smoothing to show relationships between predicted values by the random survival forests model and gene expressions based on the scatter plot.

**Supplementary Figure S2: Heat map for selected genes.** Blue and red bars (top) indicate two groups classified by the fitted random survival forests model. The heat map shows patients sorted by estimated mortalities (columns), and genes sorted by expression (rows). For patients with poor survival (blue bar, top), upper genes are underexpressed; lower genes are overexpressed. However, expression patterns for long-surviving patients (right cluster) might be reversed for poor survivors. Thus, in the heat map, the indicated genes distinguish between poor and good survivors.

**Supplementary Figure S3: Classification by cell-of-origin associated with survival.** Clustering analysis for classification into Proneural (purple), Neural (green), Classical (blue), and Mesenchymal (red) subgroups using the gene expression-based method of Verhaak et al. (B) Kaplan–Meier curves comparing these subgroups.