



**Figure S3: CladePP increases sensitivity.** The Bars in each column indicate the maximum ratio score (relative to the 79 HRR gold standard) for each gene when the NPP matrix is clustered across a different clade. Simply, it is the ratio of HRR genes from the total number of genes that cluster with a gene of interest in a specific clade. The genes are listed according to their order in the hierarchical clustering dendrogram. Peaks in the colored bars correspond to clusters with a high fraction of HRR genes. List of genes (HRR gold standard genes in bold) with their profiles and FDR adjusted p-values (calculated using hypergeometric test) for the enrichment of HRR genes are given for selected clusters. Note that some genes cluster together in several different clades. The CladePP column is the combination of the highest gene scores from the different clades.