



**Supplementary Figure:** Results shown are for a two group simulation with  $n=4$  samples for group 1 and  $n=3$  for group 2 with 1 more variable sample per group. The more variable samples were simulated to have a library size that was  $X\%$  as large as the regular samples (as indicated in the plot title).

Top row: Each panel shows results from simulations with different true positive fold-changes (FCs): 1.5-fold, 2-fold and 4-fold (as indicated in the plot title). The y-axis shows the cumulative number of false positives amongst the top 200 genes from each analysis and the x-axis indicates the simulated sample variability of the more variable samples. Middle row: Boxplots of the number of genes with false discovery rate ( $\text{FDR} < 0.1$ ) (y-axis) for different sample variability settings (x-axis) and true positive FCs (indicated in plot title). Bottom row: Boxplots of empirical FDR at a  $\text{FDR}$  cut-off of 0.1 (y-axis) for each method for various sample variabilities of the more variable samples (x-axis) and different true positive FCs (indicated in plot title). Methods have been colour coded according to the legend in panel 4, except for the top 3 panels where the 'Sample Removal' method is black rather than grey. Results from 100 independent simulations under each setting are shown.









