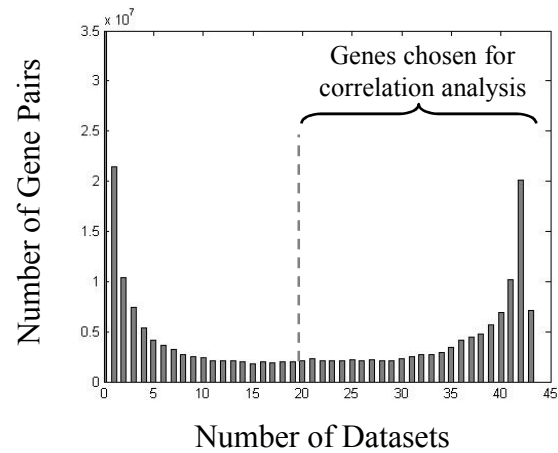


Additional data file 1



Distribution of the number of datasets in which each gene pair appears

For each gene pair, the number of datasets in which they appear simultaneously was calculated. The data is plotted as a histogram showing the number of gene pairs exist per a specific number of datasets. Only gene pairs that appear in at least 20 datasets (dashed line) were used for further analysis.