

# Integrated Gene Set Analysis for microRNA Studies

Estimation of the type I error rate

2016-02-04

## Methods

The estimation of the type I error rate have been computed from a permutation strategy of the data.

We have created a table of all miRNA - gene target pairs, and then randomly permuted the gene column. This would have the effect of preserving the number of genes each miRNA targets, and the genes and the number of miRNAs each gene is associated with, but would remove all biological association within and between miRNAs.

This permutation strategy was performed 100 times. Median and percentiles were used to describe the percentage of significant results.

All scripts are available in <https://github.com/dmontaner-papers/gsa4mirna> (folder: scripts\_permutation).

## Results

For each cancer type, this table shows the percentage of significant results.

The amount of False Positives (FP) found in the permuted analysis in **paired studies** remains under the expected thresholds according to the p-value cutoff selected.

Table 1: % FP in paired studies

Cancer	Median	Percentile 5	Percentile 95
BLCA	0	0	0.037
BRCA	0	0	0.055
CESC	0.037	0	0.405
ESCA	0.018	0	0.018
HNSC	0.11	0.037	0.883
KICH	0.258	0.055	0.662
KIRC	0	0	0.055
KIRP	1.031	0.368	2.024
LIHC	0	0	0.147
LUAD	1.104	0.276	1.987
LUSC	1.288	0.57	2.319
PAAD	0	0	0.018
PCPG	0	0	0.055
PRAD	0.147	0	0.865
STAD	0.092	0	0.736
THCA	0	0	0
UCEC	0.129	0	0.645

The amount of False Positives found in the permuted analysis in **unpaired studies** remains under the expected thresholds according to the p-value cutoff selected except for COAD which shows a median next to this threshold.

Table 2: % FP in unpaired studies

Cancer	Median	Percentile 5	Percentile 95
BLCA	0	0	0.037
BRCA	0	0	0.018
CESC	0.184	0	0.755
COAD	5.19	3.466	6.028
ESCA	0.791	0.276	1.463
HNSC	0.865	0.055	1.987
KICH	0	0	0.04
KIRC	0.313	0.11	1.141
KIRP	0.994	0.386	1.664
LIHC	4.049	2.462	6.128
LUAD	0	0	0.092
LUSC	1.224	0.54	1.844
PAAD	0.055	0	0.092
PCPG	0	0	0.11
PRAD	0	0	0.184
READ	0.313	0.055	2.337
SKCM	1.601	0.791	2.882
STAD	0.083	0	0.242
THCA	0	0	0.018
UCEC	2.797	1.692	4.529