Regression on GO Terms

Luis Pedro Coelho

Instituto de Medicina Molecular

**Summary Sentence**: Two drugs will act in a synergistic fashion if they disturb different genes with related functionality.

**Background/Introduction**

**Methods**

For the implementation of this idea, I relied, on the GO terms (molecular function and biological process vocabularies only).

First, the media and DMSO data was used to determine a baseline variation and genes were considered disturbed if they moved more than 1.5 standard deviations away from the mean at any time point. This is a noisy measurement, but it will only be used very indirectly. As above, genes were also grouped together by GO terms.

Each drug was thus characterized by two signatures:

1. A set of disturbed genes
2. A set of disturbed GO terms

For each pair of drugs, I computed the correlation between its disturbed genes and its disturbed GO terms.

There is a roughly linear correlation between these two values, but the distance to the regression line is our measure of synergy. Drugs pairs for which the GO terms are more correlated than predicted by the correlation at the gene level were predicted to be synergistic.

**Conclusion/Discussion**

**References**

**Authors Statement**

LPC developed the methodology, implemented it, and wrote the report.