**Over-representation of Genes With Respect to Function is a Very Weak Sign of Synergy**

*Luis Pedro Coelho*

*Instituto de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisboa, Portugal*

**Summary Sentence**: The starting assumption was that two drugs will act in a synergistic fashion if they disturb different genes with similar functionality.

**Background/Introduction**

We assumed that synergy would come from multiple drugs affecting different genes with similar function as these would then complement each other rather than be independent perturbations (Léhar et al., 2007; Yeh et al., 2009).

In order to detect this class of situations, we first detected a set of perturbed genes. We then mapped these to GO terms to obtain a set of perturbed GO terms. For any pair of drugs, we can measure the similarity between their perturbed GO term sets and predict the similarity in genes sets. If the actual similarity in genes is lower than predicted, then this pair is assumed to be synergistic.

**Methods**

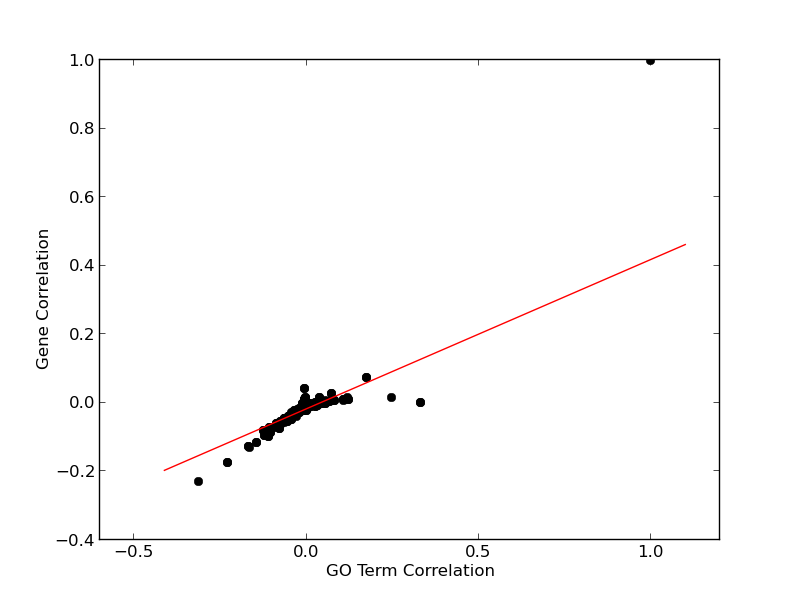
As a first step, the media and DMSO data was used to determine a baseline variation and genes were considered disturbed if they moved more than 1.5 standard deviations away from the mean at any time point. This is a noisy measurement, but it will only be used very indirectly. We had tried stricter measures (e.g., 1.5 std. deviations on multiple time-points) but those led to a very small number of perturbed genes on some drugs (including no perturbation).

We applied the same procedure after grouping genes by GO terms. This grouping was performed by assigning to each GO term and condition the sum of all the expression value of all genes that are associated with this term. After filtering as above, we obtained a set of GO terms. Only the molecular function vocabulary was used for submission.

Each drug was thus characterized by two signatures:

1. A set of disturbed genes, represented as a binary vector,.
2. A set of disturbed molecular function GO terms, equally represented as a binary vector, .

For each pair of drugs, we computed the Pearson correlation of the gene perturbation vectors, , and its GO terms perturbation vector, . There is a roughly linear correlation between these two values, ,as can be seen in Figure 1.



*Figure 1: Gene correlations as a function of GO term correlations. Each circle represents a drug pair, the straight line is the best least-squares fit.*

However, the is not perfect. the distance to the regression line is our measure of synergy. Drugs pairs for which the genes are less correlated than predicted by the correlation at the GO term were predicted to be synergistic.

**Conclusion/Discussion**

This method was very simple and it obtained only mediocre results. With the benefit of the testing data, which was not available at the time of the competition, we can test a few variations and measure whether they would have been better than the submission. In particular, the restriction to the Molecular Function vocabulary was somewhat arbitrary and we can test other GO vocabularies. In total, there are 6 possible combinations of vocabularies. The best result is the combination of the molecular function and the biological process vocabularies, but the p-value is 16%. The best single vocabulary is not molecular function, but biological process. In retrospect, this may have been a better embodiment of the idea that was underlying this method, but the result is still of limited value, as the resulting p-value is 21%.

Another variation results from reversing the prediction and predict GO term correlation based on gene correlation, with synergy being again measured as the deviation from prediction. The best results are obtained with the molecular function, or the biological process vocabularies, or both combined. These three combinations all result in p-values of 16% or 17%.

We thus conclude that fundamentally, this model was of limited value and a richer model would have been necessary for adequate prediction.

**References**

Cokol, Murat, Hon Nian Chua, Murat Tasan, Beste Mutlu, Zohar B. Weinstein, Yo Suzuki, Mehmet E. Nergiz et al. "Systematic exploration of synergistic drug pairs." *Molecular systems biology* 7, no. 1 (2011).

Yeh PJ, Hegreness MJ, Aiden AP, Kishony R (2009) Drug interactions and the evolution of antibiotic resistance. *Nat Rev Microbiol* **7**: 460–466

Lehár, Joseph, et al. "Chemical combination effects predict connectivity in biological systems." *Molecular systems biology* 3.1 (2007).

**Authors Statement**

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