Electronic Supplementary Material (ESI) for Molecular BioSystems.

Supplementary materials

Performance evaluation for different methods

In the following, we shall to evaluate effects for different 3D structure similarity thresholds, correlation coefficients and clustering algorithms. Since the count of modules generated by different methods and drug numbers in different modules are different, we should choose a standard to reduce this effect. In addition, the predication of our method are mainly based on modules and drug's ATC code, we choose the average percentage (AP) of dominant attribute (ATC) in all modules as standard. AP-Score is defined as follow:

$$AP(M) = \frac{\sum_{i=1}^{m} n_i}{\sum_{j=1}^{m} N_j}$$

Where m is the count of modules generated by method M. N_j is the count of drugs with ATC code in module j and n_i is the count of drugs with dominant attribute (ATC) in module i. For example, if there are 12 drugs in module i, 10 drugs with ATC code (2 drugs without ATC), and 6 of 10 drugs with dominant attribute (N represents nervous system), then N_i is 12 and n_i is 6.

The selection of threshold for 3D structure similarity score

Our method's main approach is constructing 'expression profile' in which each sore is a consensus response score (CRS) for every protein to each drug. CRS quantifies the degree of consistence for drug's chemical structure and a protein's function. For a given drug, we choose some drugs as benchmarks which should include drugs with high, medium and low similarities to this drugs. At first, we computed the average 3D structure similarity score for drug pairs among our dataset (C^2_{965} drug pairs), the score is 0.5580 (the 3D scores in our manuscript ranging from 0 to 2). Hence, the threshold should under 0.5580. In addition, we counted the number of drug pairs exceeding some thresholds (shown in Figure suppl. 1.). We can see that the number of drug pairs decreasing significantly when thresholds larger than 0.4. So we used the method described in our manuscript to construct 'expression profile', drug similarity network and mine modules using different thresholds (shown in table 1). The resulting AP-score is shown in table suppl. 1 (we just used Pearson correlation coefficient in our method for assessing different threshold). We can see that the AP-score corresponding 0.4 is the largest. Hence, we choose 0.4 as the threshold for 3D structure similarity score.

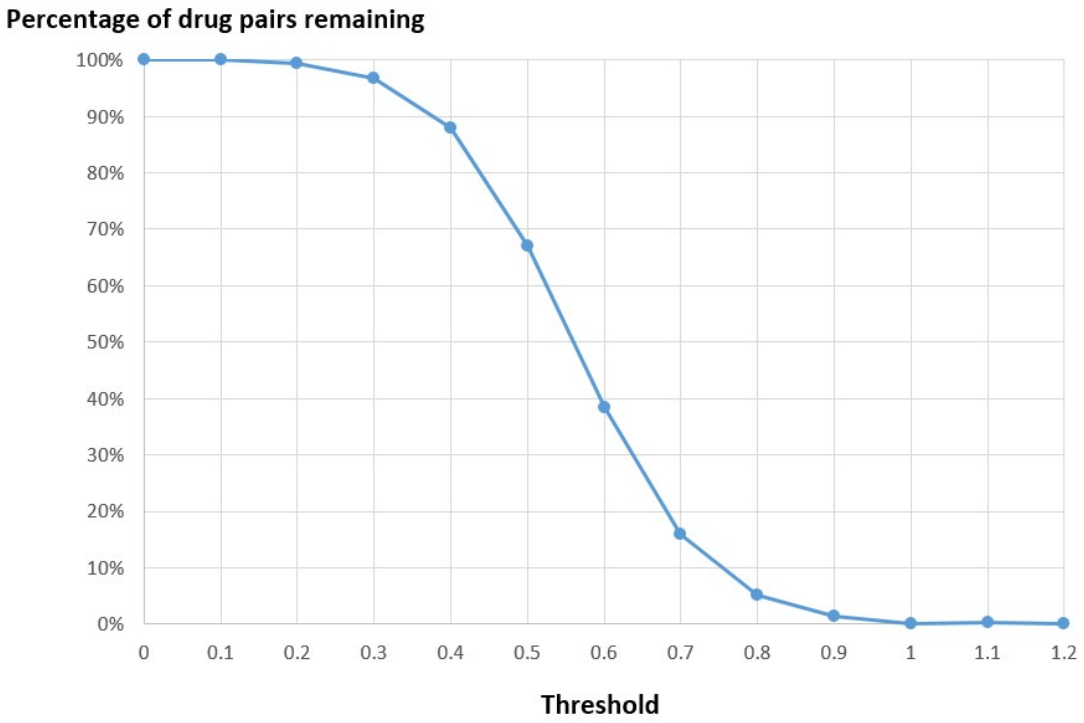

Figure suppl. 1. The percentage of drug pairs exceeding different threshold.

Table suppl. 1. The AP-score for different threshold.

Threshold	0	0.1	0.2	0.3	**0.4**	0.5
AP-score	0.4009	0.3578	0.3709	0.3865	**0.4488**	0.4243

The effect of different correlation coefficients

In this section, we assess the effects for different correlation coefficients (Pearson correlation coefficient, Spearman's rank correlation coefficient, Kendall rank correlation coefficient). In our method, we used correlation coefficient twice in consensus report score and drug similarity. In particular, we choose 0.4 as the threshold for 3D structure similarity score for each of three correlation coefficients. The resulting AP score is shown in table suppl. 2. We can see that the AP-score corresponding Pearson correlation coefficient is the largest. Hence, we choose Pearson correlation coefficient in our method.

Table suppl. 2. The AP-score for different correlation coefficient.

Threshold	**Pearson**	Spearman	Kendall
AP-score	**0.4488**	0.3763	0.4041

The effect of different clustering algorithms

In this section, we study how different clustering algorithms lead to different results in our method. We chose three popular clustering algorithms (MCODE[1], MINE[2], NeMo[3]). MCODE is a very popular clustering algorithm which utilizes vertex weighting to grow clusters from a starting vertex of high local weight by iteratively adding neighboring vertices with similar weights.

MINE is an agglomerative clustering algorithm using a modified vertex weighting strategy. It has strengths for the identification of modules in dense, highly interconnected networks. NeMo is based on SPLAT[4]and CODENSE[5], identifying frequent dense subgraphs in input networks. We employed these three clustering algorithms to identified modules, then we calculated the AP score for rach of them (shown in table suppl. 3). We can see that MINE gets the the largest AP-score , but it just identified 7 modules and one module including 862 drugs. Hence, we used MCODE to identified modules in our manuscript.

Table suppl. 3. The AP-score for different clustering algorithms.

Threshold	**MCODE**	MINE	NeMo
AP-score	0. 4488	0. 6111	0. 4272

1. Bader GD, Hogue CW: **An automated method for finding molecular complexes in large protein interaction networks**. *BMC bioinformatics* 2003, **4**:2.
2. Rhrissorrakrai K, Gunsalus KC: **MINE: Module Identification in Networks**. *BMC bioinformatics* 2011, **12**:192.
3. Yan X, Mehan MR, Huang Y, Waterman MS, Yu PS, Zhou XJ: **A graph-based approach to systematically reconstruct human transcriptional regulatory modules**. *Bioinformatics* 2007, **23**(13):i577-586.
4. Yan X: **Mining closed relational graphs with connectivity constraints.** *Proceedings of the Int Conf on Konwledge Discovery and Data Mining* 2005:324-333.
5. Hu H, Yan X, Huang Y, Han J, Zhou XJ: **Mining coherent dense subgraphs across massive biological networks for functional discovery**. *Bioinformatics* 2005, **21 Suppl 1**:i213-221.