**Table S1**. The drug vector space model is in comparison with the correlations of gene expression signature and structure similarity between two drugs targeting the same disease

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **1st drug** | **2nd drug** | **Drug vector space model** | | | **Gene expression signature[a]** | | **Structure similarity** |
| **Jaccard Index (*JI*)** | | **Number of shared genes** | **Pearson correlation** | ***P* value** | **Tanimoto coefficient (*TC*)** |
| **Breast cancer (MCF-7[b])** | | |
| Doxorubicin | Estradiol | **0.5397\*** | | 90 | **0.1849** | **0.0135** | **0.5441** |
| Doxorubicin | Progesterone | **0.5189\*** | | 77 | **0.0420** | **0.5954** | **0.4234** |
| Estradiol | Progesterone | **0.6683\*** | | 81 | **0.6174** | **<2.2E-16** | **0.5670** |
| **Prostate cancer (PC-3[b])** | | |
| Letrozole | Levonorgestrel | 0.3125 | | 9 | 0.0112 | 0.9559 | 0.0661 |
| Letrozole | Raloxifene | 0.3200 | | 12 | 0.1190 | 0.4766 | 0.5994 |
| Levonorgestrel | Raloxifene | 0.2784 | | 16 | 0.1692 | 0.2901 | 0.1781 |
| Metformin | Paclitaxel | **0.5100\*** | | 38 | **0.5660** | **4.11E-09** | **0.0818** |
| Metformin | Tamoxifen | **0.5485\*** | | 41 | **0.2242** | **0.0347** | **0.0447** |
| Paclitaxel | Tamoxifen | **0.6452\*** | | 46 | **0.4203** | **1.48E-05** | **0.5224** |
| **Leukemia (HL-60[b])** | | |
| Ganciclovir | Ribavirin | 0.2761**#** | | 13 | 0.0020 | 0.9917 | 0.5897 |
| Ganciclovir | Thalidomide | 0.3118 | | 19 | -0.0204 | 0.9132 | 0.3441 |
| Ribavirin | Thalidomide | 0.3667 | | 18 | 0.2817 | 0.1388 | 0.2908 |

**[a]**The drugs were chosen according to the higher level of repurposed drug score (***RDS***) and the availability in the database of Connectivity Map (build 02, <http://www.broadinstitute.org/cmap/>). For example, the most three common dugs (the top 3 *RDS* drugs) in prostate cancer are Ethinyl estradiol, Diethylstilbestrol and Levonorgestrel. However, there is no experimental data about Ethinyl estradiol and Diethylstilbestrol, we used Letrozole, Ralozifene, Metformin, Paclitaxel and Tamoxifen instead.

**[b]**Only six human cancer cell lines were treated with different drugs in the study of Connectivity Map. In this data set, three of these cell lines (MCF-7, PC-3 and HL-60) are treated more than one thousand times of drug treatment, and thus selected for testing our hypothesis regarding drug-drug relations.

**\***The ***JI*** score of drug vector space model, which is higher than 0.5, is significantly supported by either or both (labeled with underline) of gene expression signature only (***p*** < 0.05, and labeled in red) and Tanimoto coefficient only (***TC*** > 0.4, and labeled in green).

**#**The drug pair is highlighted for the condition that drug vector space model (***JI*** < 0.5) and gene expression signature (***p*** > 0.05) are not significant but structure similarity (***TC*** > 0.4).