Explore Small Molecule-induced Genome-wide Transcriptional Profiles for Novel Inflammatory Bowel Disease Drug

Xiaoshu Cai, BS1, Yang Chen, PhD2, Zhen Gao, PhD2, Rong Xu, PhD2

1Department of Electrical Engineering and Computer Science, School of Engineering, Case Western Reserve University, Cleveland, Ohio, USA; 2Department of Epidemiology & Biostatistics, School of Medicine, Case Western Reserve University, Cleveland, Ohio, USA

Abstract

Inflammatory Bowel Disease (IBD) is a chronic and relapsing disorder, which affects millions people worldwide. Current drug options cannot cure the disease and may cause severe side effects. We developed a systematic framework to identify novel IBD drugs exploiting millions of genomic signatures for chemical compounds. Specifically, we searched all FDA-approved drugs for candidates that share similar genomic profiles with IBD. In the evaluation experiments, our approach ranked approved IBD drugs averagely within top 26% among 858 candidates, significantly outperforming a state-of-art genomics-based drug repositioning method (p-value < e-8). Our approach also achieved significantly higher average precision than the state-of-art approach in predicting potential IBD drugs from clinical trials (0.072 vs. 0.043, p<0.1) and off-label IBD drugs (0.198 vs. 0.138, p<0.1). Furthermore, we found evidences supporting the therapeutic potential of the top-ranked drugs, such as Naloxone, in literature and through analyzing target genes and pathways.

Introduction

Inflammatory bowel disease (IBD) is a group of chronic and relapsing inflammatory conditions in the intestinal tract. It affects 3.6 million people in the USA and Europe, and has an increasing incidence around world in recent years. IBD is currently not medically curable. Available drug options, such as corticosteroids, aminosalicylates and targeted drugs, focus on remitting the inflammation and are only effective for a subset of the patients. In addition, these drugs may cause severe side effects, such as an increased rate of malignancies or infectious diseases. Traditional drug development experiments have identified novel targeted drugs, such as TNF-alpha blockers, which showed a promising efficacy in treating IBD. However, recent clinical researches show that a large number of patients do not respond to these drugs. Therefore, discovering new drug treatment for IBD is of compelling need.

Computational drug repositioning approaches predict new indications for existing drugs, which have passed a number of toxicity and other tests, thus may lead to rapid and cost-effective drug discovery. Existing drug repositioning strategies can be classified into three categories: drug-based, disease-based and profile-based (Figure 1). Drug-based approaches (Figure 1 (a)) assume that similar drugs can be used to treat the same disease. They leverage chemical structure, side-effects, and other drug characteristics to determine the drug similarity. Disease-based approaches (Figure 1 (b)) predict new drug indications exploiting disease-disease similarity based on disease phenotypes, genetic and genomic signatures. Recently, profile-based drug repositioning approaches (Figure 1(c)) have been developed to identify drug candidates for small-cell lung cancer, Parkinson’ disease, as well as IBD. Unlike the drug-based and disease-based strategy, profile-based approaches do not depend on existing drug-treatment knowledge and may have increased ability to discover new drug-disease associations. Early profile-based approaches match the drug target genes with the disease-associated genes obtained from GWAS, OMIM and network-based disease-gene predictions. Several recent studies also investigate the similarity in phenotype profiles between diseases and drugs.

Most current profile-based drug repositioning approaches exploit the well-studied drug targets to match genetic and phenotypic profiles of the drugs and diseases. However, the interactions between chemical compounds and human proteins are complex, and the off-targets of existing drugs may provide a unique path to link diseases with new potential treatments. For example, Connectivity Map (CMap) is the first database that provides the genomic signatures associated with perturbations of small molecular compounds, and has facilitated the drug discovery for IBD. Recently, The Library of Integrated Network-based Cellular Signatures (LINCS) offers more comprehensive drug-gene interaction data through significantly extending the experiments in the first version of CMap on much more cell lines (from 4 to 18) with more perturbagens (from 164 to 20413 chemical perturbagens) using more signatures (from 564 to millions) through
calculating drug-drug similarities based on gene expression profiles. In this study, we developed methods to mine genomic profiles for approved drugs from vast amounts of data in LINCS and matched profiles between IBD and the drugs. We demonstrated the improved performance of IBD drug prediction comparing with the state-of-art approach based on CMap\textsuperscript{14}, and show several valuable new drug candidates. Our framework can be easily extended to predicting drugs for other diseases with the LINCS data.

![Diagram](Image)

Figure 1. Computational drug repositioning strategies: (a) disease-based methods, (b) drug-based methods, and (c) profile-based methods

![Diagram](Image)

Figure 2. Workflow of our drug repositioning approach based on LINCS L1000 data

Methods

Our drug repositioning framework contains three steps (Figure 2): (i) generating genomic profiles for approved drugs from the LINCS data; (ii) constructing gene expression profile for IBD based on GEO; (iii) ranking all the candidate drugs based on the genomic similarity to IBD. The following sections describe the three steps as well as the evaluation method in detail.

A. Identify genomic profiles for approved drugs from the LINCS data

The Library of Integrated Network-based Cellular Signatures (LINCS) dataset is a catalog of gene-expression data collected from human cells treated with chemical compounds and genetic reagents. The LINCS L1000 dataset contains gene expression-profiling assays for 20413 small-molecular compound perturbagens and 22268 genes\textsuperscript{14}. The expression profiles are generated for 1171 landmark transcripts while the rest of the transcriptome,
approximately 22,000 genes, are estimated by a model built from computational processing of thousands of gene expression datasets from Gene Expression Omnibus (GEO). Several recent studies computed transcriptional similarity based on landmark genes. Here, we explored the methods to integrate the complete data in LINCS L1000 and gain more prediction power in drug repositioning.

Gene expression data are available at different levels in the LINCS L1000 database (http://www.lincscloud.org/). We downloaded the differential expression level data, which consist of Z-scores representing the significance of the associations between genes and chemical perturbations. The data are stored in a 110GB GCTX file, which is a binary format file that allows for high-performance access. We extracted a large gene expression matrix from the file. The 22,268 rows of the matrix represent probe sets (corresponding to genes) and the columns are millions of signatures, which represent replicated experiments for each chemical perturbation on all the genes. We obtained a list of 1,761 FDA-approved drugs from DrugBank and mapped them to 858 identical drugs in the L1000 database. Then we extracted the sub-matrix for the 858 approved drugs through the programmatic access to the dataset. The sub-matrix consists of 22,268 rows and 100,041 columns (each column represents one signature, and one drug corresponds to multiple signatures), and the Z-score stored in it was used to identify genomic profiles for the approved drugs.

In the sub-matrix of gene expression data for the 858 approved drugs, each gene may correspond to multiple probe sets (rows) and each drug is mapped to multiple signatures (columns). As a result, a drug-gene link is represented by a multiple-to-multiple relationship in the matrix. To represent the strength of a drug-gene association, we counted the number of genes that are over and under-expressed and constructed a sign score vector for each drug. Specifically, we determined the genes as over-expressed if the Z-score was greater than 2 and under-expressed if the Z-score was less than -2 following the previous study (Figure 3(a)). For example, in figure 3, we marked the over-expressed genes as red and the under-expressed genes as blue in the Z-score matrix. Then we represented the gene-drug relationship by extracting corresponding probe sets (rows) and relative signatures (columns) into a small matrix (Figure 3(c)). Last, we computed a sign score \( S_{\text{sign}} \) for the drug-gene pair:

\[
S_{\text{sign}} = \frac{\text{Number of over and under-expressed values}}{\text{Number of rows} \times \text{Number of columns}}
\]

Given a drug, a sign score was calculated for each of the 12,716 genes, and vector consisting of 12,716 sign scores was constructed to represent the genomic profile for the drug:

\[
drug_i = < S_{\text{sign}_1}, S_{\text{sign}_2}, ..., S_{\text{sign}_i}, ..., S_{\text{sign}_n} >
\]

Figure 3. The schematic representation of workflow to construct drug-gene profile: (a) Representation of determining over- and under-expression, (b) Z-score matrix of gene expression level profile in LINCS can be painted using our schema, (c) Relationship between gene and drug in LINCS
B. Identify the genomic profile for IBD from GEO

We downloaded the study GSE6731 and identified the significant differentially expressed genes for IBD. A total of 36 samples assembled as gene expression patterns of active and inactive areas of IBD, infectious colitis and healthy controls from human colonoscopy biopsy. For a fair comparison, we selected the same study as the CMap-based method used. We used the GEO2R software\textsuperscript{29} to detect the significant genes. Then, we converted P-values generated from GEO2R to Z-scores and set the threshold of 2/-2 (the same one in Method A) to identify significant genes for disease genomic profiles.

C. Calculate genomic similarity between IBD and drugs

We computed the cosine similarity between drug and disease genomic profiles to estimate the potential of each drug in treating IBD. The similarity is defined as:

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

in which A and B are sign score vectors (represented as in equation(2)) of drug and disease, respectively. Then we ranked the candidate drugs based on this genomic similarity.

D. Compare with an existing baseline method

We compared our approach with a state-of-art drug repositioning method for IBD, which exploited the CMap data to identify the genomic profiles for drugs\textsuperscript{14}. Our method and the CMap-based approach used the same GEO dataset to identify significant genes for IBD. They provided a ranked list of 79 drugs for IBD, and considered the rest of the 164 drugs irrelevant to IBD. We simulated the full rank of CMap-based approach by randomizing the irrelevant drugs for 100 times for the purpose of fair comparison. Then we constructed evaluation sets containing approved IBD drugs and novel IBD drugs respectively, and compared the two methods in prioritizing the evaluation drugs.

In our recent studies, a large-scale drug-disease treatment knowledge base was constructed based on multiple resources\textsuperscript{30,31,32}. The databases include 9,216 drug-disease treatment pairs extracted from FDA drug labels, 34,306 pairs extracted from 22 million published biomedical literature abstracts, 69,724 disease-treatment pairs extracted from clinical trials and 111862 pairs from the FDA post marking surveillance system.

a. Evaluation on FDA-approved IBD drugs

We extracted 25 FDA-approved IBD drugs from FDA drug labels in this knowledge base, and found 10 of them appear in the LINCS dataset. The intersected approved drug set consists of 8 corticosteroids and 2 aminosalicylates. We plotted the number of approved IBD drugs that fell into 10 ranking ranges to estimate their distribution. Mean and median of the ranks for the approved IBD drugs were computed and compared with the CMap-based method. We also evaluated whether the two methods generated significantly different ranks for the approved IBD drugs.

b. Evaluation on novel IBD drugs

Based on this knowledge base, we constructed two evaluation sets as the proxy of novel IBD drugs. The first sets contain 58 potential IBD drugs that have been tested in clinical trials. The second one contains 238 off-label IBD drugs collected from the FDA post marking surveillance system. FDA-approved IBD drugs are removed from both sets. We used the histogram plot to estimate the distribution of drug ranks in two evaluation sets, respectively. We also calculated the precision-recall curve and the mean average precision to evaluate the ranking of the two methods.

Results

A. Our system prioritized the FDA-approved IBD drugs and outperformed the CMap-based method

Figure 4 show that our approach prioritized most approved IBD drugs. In particular, our method ranked 5 out of 10 approved IBD drugs within top 10% among 858 drugs. The average rank of 10 FDA-approved IBD drugs among 858 drugs is 26% (significantly higher than the random rankings, p-value < e-8, T test). The median rank is 19.5%, which is also significantly lower than random case (p-value < e-8, Wilcoxon signed sum test). Together, the results show that our ranking prioritized most of the approved IBD drugs and performed significantly better than random ranking. These results also support that our methods of processing the LINCS data and calculating the disease-drug profile similarity are reasonable. We noticed that two approved drugs, Methylprednisolone and Balsalazide, were ranked a bit lower than 50%. This might be due to small sample size for the two drugs: most of the chemical
compounds in LINCS have several hundreds of replicated signatures, while the sample size of Methylprednisolone and Balsalazide is only 10% comparing to the average.

We also compared with the CMap-based method in predicting approved IBD drugs. Because of their smaller search scope (only three approved IBD drugs appear in the data), the CMap-based system only identified one approved IBD drug as the significant candidates (within top 79). The average rank of all the approved IBD drugs in the simulated ranks generated by the CMap-based method is 34%, which is significantly lower than ours average rank (p-value < e-8, T test).

![Figure 4. Distribution of FDA-approved IBD drugs](image)

B. Our system identified novel IBD drugs and performed significantly better than the CMap-based method

In the two evaluation sets of novel IBD drugs, 24 drugs from the clinical trials and 141 from the post marketing surveillance system appear in our drug repositioning system. We ranked 7 drugs from the clinical trials within top 10%, and 13 within 30%. In addition, 23 off-label IBD drugs from the post marketing surveillance system were ranked within top 10% and 88 of them (58%) were ranked above 50%. We plot the distribution of the evaluation drugs among all ranked list in figure 5(a)-(b), and found an increasing trend for the number of positive drugs when the rank moves from the bottom to the top. In contrast, the off-label drugs spread evenly over different ranks in the result of the CMap-based approach (figure 5(d)). The evaluation drugs from the clinical trials are even likely to be ranked in the bottom by the CMap-based approach (figure 5(c)). The comparison of precision-recall curves between the two methods (figure 6) suggests the same result: our approach achieved a mean average precision of 0.072 and 0.198 for the two evaluation sets respectively, which are significantly better than 0.043 (p<0.1) and 0.138 (p<0.1) obtained by the CMap-based approach.

We noticed that in predicting drugs from clinical trials, the precisions at different cutoffs of recall are relatively low for both two methods. This mainly depends on the size of gold standard set, as CMap-based approach intersected only 7 drugs from clinical trials and 24 potential drugs from clinical trials appeared in our approach. This precision recall curve and corresponding mean average precision value do not represent the true predictive power of two methods. We evaluated precision recall here only for comparison of rankings of two methods.
Figure 5. Our approach: (a) Distribution of count of drugs from clinical trials. (b) Distribution of count of drugs from FDA post marking surveillance system. CMap-based approach: (c) Distribution of count of drugs from clinical trials. (d) Distribution of count of drugs from FDA post marking surveillance system.

Figure 6. Precision-recall curves in ranking novel drugs from clinical trials (left); Precision-recall curve in ranking novel drugs from FDA post marking surveillance system (right)
C. Top-ranked predicted drugs show promise in treating IBD

Among the top-ranked drugs, we found Naloxone, which was ranked in top one, and Naltrexone in top seven have the potential in treating IBD. Both of them are opioid antagonists and have similar effects\textsuperscript{33, 34}. Low-dose naltrexone (LDN) has been demonstrated as a novel anti-inflammatory treatment for chronic pain conditions, such as fibromyalgia, Crohn’s disease, multiple sclerosis, and complex regional pain syndrome\textsuperscript{33}. Naloxone is traditionally approved for complete or partial reversal of narcotic depression. We further extracted the significantly differentially expressed genes induced by Naloxone from LINCS, and found many common drug-interacting genes are closely associated with IBD pathways (Table 1). For example, the target gene CXCL1 is over-expressed in Naloxone and influences the pathway of Cytokine-cytokine receptor interaction, which regulates the innate as well as adaptive inflammatory host defenses. IL18 and STAT1 are over-expressed in Naloxone, and contained in the pathway directly associated with IBD. We ranked the pathways related to Naloxone and IBD respectively based on the gene-pathway associations and found that top six pathways are consistent between the drug and the disease (Table 2). The pathway analysis results support the therapeutic potential of Naloxone for IBD.

In addition, previous researches show that severity of IBD varies during pregnancy, menopause, or oral contraceptives use, and gonadal sex hormones may be involved in the IBD pathogenesis\textsuperscript{35}. Interestingly, Estradiol and Estriol, which are ranked in the second and fourth by our approach, are two female sex hormones. An earlier Estrogen study shows that female sex hormones may result in less severed IBD and this correlates with decreased pro-inflammatory mediators levels\textsuperscript{35}. Scopolamine, a secondary metabolite from Solanaceae family, is atropine alkaloid drug with muscarinic antagonist effects\textsuperscript{36} and has been used to control gastric peristalsis in bowel diseases\textsuperscript{37}. The 6th in the rank list is Doxepin, which is quite interesting as it’s depression drug while shown Anti-ulcer effect\textsuperscript{38}. Anti-ulcer effect shows relevance for treating IBD\textsuperscript{39}. In the 10th of the list is Nafcillin, which is a narrow-spectrum antibiotics of the penicillin class, and it is interestingly found to enhance innate immune-mediated killing of *Staphylococcus aureus*\textsuperscript{40}.

| Gene symbol | Pathway |
|-------------|---------|
| HLA-DMA     | Reactome immune system |
| RELB        | Pid IL12 _2pathway |
| CXCL1       | Cytokine-cytokine receptor interaction |
| IL18, STAT1 | Inflammatory bowel disease |
| IL6         | Jak-STAT signaling pathway |

**Table 1.** Common target genes that are significantly differentially expressed and related pathway of top-ranked drug Naloxone and IBD

| Rank | Naloxone Pathway | IBD Pathway |
|------|------------------|-------------|
| 1    | REACTOME_GPCR_LIGAND_BINDING | REACTOME_GPCR_LIGAND_BINDING |
| 2    | REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_SYSTEM | REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_SYSTEM |
| 3    | REACTOME_PLATELET_ACTIVATION_SIGNALING_AND_AGGREGATION | REACTOME_PLATELET_ACTIVATION_SIGNALING_AND_AGGREGATION |
| 4    | REACTOME_METABOLISM_OF_LIPIDS_AND_LIPOPROTEINS | REACTOME_METABOLISM_OF_LIPIDS_AND_LIPOPROTEINS |
| 5    | REACTOME_ADAPTIVE_IMMUNE_SYSTEM | REACTOME_ADAPTIVE_IMMUNE_SYSTEM |
| 6    | REACTOME_HEMOSTASIS | REACTOME_HEMOSTASIS |
| 7    | REACTOME_METABOLISM_OF_CARBOHYDRATES | REACTOME_ANTIGEN_PROCESSING_UBIQUITINATION_PROTEASOME_DEGRADATION |
| 8    | REACTOME_NONSENSE_DECAY_ENHANCED_BY_THE_EXON_JUNCTION_COMPLEX | REACTOME_PI3K CASCADE |
| 9    | REACTOME_ANTIGEN_PROCESSING_UBIQUITINATION_PROTEASOME_DEGRADATION | REACTOME_SIGNALING_BY_FGFR |
| 10   | REACTOME_METABOLISM_OF_PROTEINS | REACTOME_METABOLISM_OF_CARBOHYDRATES |
Discussion

In this study, we developed a drug repositioning framework based on genomic profile matching between drugs and diseases. We leveraged the data of differentially expressed genes induced by drugs from the most recent pharmacological database called LINCS. Our work has several limitations that can be improved in the future studies. First, the evaluation of FDA-approved drugs shows that the prediction power is affected by the sample size of drug expression signature. When the sample size of signatures of a drug is much smaller than the average size of replicates of signatures, the rank of it is relatively low as a result. In the future, we plan to experiment on adding weights to normalize the sample size in the sign score for drug-gene pairs and balance the low rank of minority caused by small sample size.

In addition, we currently focus on investigating the genomic profile similarity between drugs and diseases, and treat the over- and under-expression equally in calculating the similarity. In the future, we will consider the direction of the gene expression level and only match the significant genes with opposite directions between drugs and diseases. We also plan to incorporate higher-level phenotype profiles in evaluating the relationships between drugs and diseases. Finally, we currently detected significantly expressed genes for IBD from one GEO study, primarily in order to make a fair comparison with the CMap-based approach. Since the gene expression studies in GEO are highly noisy, we will integrate the results based on multiple studies later when applying our system in predicting new drugs for IBD and other disorders.

Conclusion

We have developed a novel drug prediction framework leveraging drug-induced gene expression profiles in LINCS. We demonstrated that our approach identifies approved and novel IBD drugs and achieved significantly better performance comparing with a state-of-art drug repositioning approach based on an early pharmacological database called CMap. Literature-review and pathway analysis show that the top-ranked predicted drugs have the potential in treating IBD.

Funding

The project was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under the NIH Director’s New Innovator Award number DP2HD084068.

References

1. Baumgart, Daniel C., and William J. Sandborn. Inflammatory bowel disease: clinical aspects and established and evolving therapies. The Lancet 369.9573 (2007): 1641-1657.
2. Martínez-Montiel, M. P., et al. Pharmacologic therapy for inflammatory bowel disease refractory to steroids. Clinical and experimental gastroenterology 8 (2015): 257.
3. Thomas, Arul, and Nilesh Lodhia. Advanced therapy for inflammatory bowel disease: a guide for the primary care physician. The Journal of the American Board of Family Medicine 27.3 (2014): 411-420.
4. de Mattos, Bruno Rafael Ramos, et al. Inflammatory Bowel Disease: An Overview of Immune Mechanisms and Biological Treatments. Mediators of Inflammation 501 (2015): 493012.
5. Loftus, Edward V. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. Gastroenterology 126.6 (2004): 1504-1517.
6. Engel, Matthias A., and Markus F. Neurath. New pathophysiological insights and modern treatment of IBD. Journal of gastroenterology 45.6 (2010): 571-583.
7. Keiser, MJ, Setola V, Irwin, JJ, et al. Predicting new molecular targets for known drugs. Nature, 2009, 462(7270): 175-181.
8. Dudley, JT, Deshpande, T, Butte, AJ. Exploiting drug–disease relationships for computational drug repositioning. Brief Bioinform, 2011, bbr013.
9. Campillos, M., Kuhn, M., Gavin, A. C., Jensen, L. J., Bork, P. (2008). Drug target identification using side-effect similarity. Science, 321(5886), 263-266.
10. Gottlieb, A., Stein, G. Y., Ruppin, E., Sharan, R. (2011). PREDICT: a method for inferring novel drug indications with application to personalized medicine. Mol syst biol, 7
11. Suthram, S, Dudley, JT, Chiang, AP, Chen, R, Hastie, TJ, Butte, AJ. Network-based elucidation of human
disease similarities reveals common functional modules enriched for pluripotent drug targets. PLoS Comput Biol, 2010, 6(2): e1000662.

12. Awaad, Amani S., Reham M. El-Meligy, and Gamal A. Soliman. Natural products in treatment of ulcerative colitis and peptic ulcer. Journal of Saudi chemical society 17.1 (2013): 101-124.

13. Chen Y, Cai X, Xu R (2015) Combing Human Disease Genetics and Mouse Model Phenotypes Towards Drug Repositioning for Parkinson’s Disease. 2015 Annual American Medical Informatics Association Symposium, CA (accepted).

14. Dudley, Joel T., et al. Computational repositioning of the anticonvulsant topiramate for inflammatory bowel disease. Science translational medicine 3.96 (2011): 96ra76-96ra76.

15. Sanseau, P, Agarwal, P, Barnes, MR, Pastinen, T, Richards, JB, Cardon, LR, Mooser, V. Use of genome-wide association studies for drug repositioning. Nat biotechnol, 2012, 30(4): 317-320.

16. Okada, Y, Wu, D, Trynka, G, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature, 2014, 506(7488): 376-381.

17. Wang, ZY, Zhang, HY. Rational drug repositioning by medical genetics. Nat Biotech, 2013, 31(12): 1080-1082.

18. Chen Y, Xu R (2015) Network-based Gene Prediction for Plasmodium falciparum Malaria Towards Genomics-based Drug Discovery. BMC Genomics, 16(Suppl 7):S9

19. Chen Y, Li L, Zhang GQ, Xu R (2015) Phenome-driven Disease Genetics Prediction Towards Drug Discovery. Bioinformatics, 31 (12): i276-i283 doi: 10.1093/bioinformatics/btv245

20. Chen Y, Zhang X, Zhang GQ, Xu R (2014) Comparative Analysis of A Novel Disease Phenotype Network Based on Clinical Manifestations, Journal of Biomedical Informatics, 53:113-20

21. Chen, Y., Li, L., & Xu, R. (2015). Disease Comorbidity Network Guides the Detection of Molecular Evidence for the Link Between Colorectal Cancer and Obesity. AMIA Summits on Translational Science Proceedings, 2015, 201.

22. Chen Y, Gao Z, Wang B, Xu R (2015) Towards precision medicine-based therapies for glioblastoma: interrogating human disease genomics and mouse phenotypes. ICIBM 2015 (accepted).

23. Lamb J, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. Sep 29.2006 313:1929.

24. http://support.lincscloud.org/

25. Vidović, Dušica, Amar Koleti, and Stephan C. Schürer. Large-scale integration of small molecule-induced genome-wide transcriptional responses, Kinome-wide binding affinities and cell-growth inhibition profiles reveal global trends characterizing systems-level drug action. Frontiers in genetics 5 (2014).

26. Ma’ayan, Avi, et al. Lean Big Data integration in systems biology and systems pharmacology. Trends in pharmacological sciences 35.9 (2014): 450-460.

27. Duan, Qiaonan, et al. LINCS Canvas Browser: interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures. Nucleic acids research (2014): gku476.

28. Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., et al. (2014). DrugBank 4.0: shedding new light on drug metabolism. Nucleic acids research, 42(D1), D1091-D1097.

29. Barrett, T, Willhite, SE, Ledoux, P, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res, 2013, 41(D1): D991-D995.

30. Xu R, Wang Q. Large-scale extraction of drug-disease treatment pairs from biomedical literature for drug repurposing. BMC Bioinformatics, 2013, 14(1): 181.

31. Xu R, Li L, Wang Q. Towards building a disease-phenotype relationship knowledge base: large scale extraction of disease-manifestation relationship from literature. Bioinformatics, 2013, doi: 10.1093/bioinformatics/btt359.

32. Xu R, Wang Q. Automatic signal prioritizing and filtering approaches in detecting post-marketing cardiovascular events associated with targeted cancer drugs from the FDA Adverse Event Reporting System (FAERS). J Biomed Inform, 2014, 171-177.

33. Younger, Jarred, Luke Parkitny, and David McLain. The use of low-dose naltrexone (LDN) as a novel anti-inflammatory treatment for chronic pain. Clinical rheumatology 33.4 (2014): 451-459.

34. Syed, Yahiya Y., and Gillian M. Keating. Extended-release intramuscular naltrexone (VIVITROL®): a review of its use in the prevention of relapse to opioid dependence in detoxified patients. CNS drugs 27.10 (2013): 851-861.

35. Cook, Lydia C., et al. The role of estrogen signaling in a mouse model of inflammatory bowel disease: a Helicobacter hepaticus model. PloS one 9.4 (2014): e94209.

36. Ziegler, Jörg, and Peter J. Facchini. Alkaloid biosynthesis: metabolism and trafficking. Annu. Rev. Plant Biol. 59 (2008): 735-769.

37. Katoh, Kansei, et al. Comparison of gastric peristalsis inhibition by scopolamine butylibromide and glucagon:
evaluation by electrogastrography and analysis of heart rate variability. Journal of gastroenterology 38.7 (2003): 629-635.

38. Leitold, M., W. Fleissig, and A. Merk. Anti-ulcer and secretion-inhibitory properties of the tricyclic derivative doxepin in rats and dogs. Arzneimittel-Forschung 34.4 (1983): 468-473.

39. Kishimoto, Shinya, et al. Rebamipide, an antiulcer drug, prevents DSS-induced colitis formation in rats. Digestive diseases and sciences 45.8 (2000): 1608-1616.

40. Sakoulas, George, et al. Nafcillin enhances innate immune-mediated killing of methicillin-resistant Staphylococcus aureus. Journal of Molecular Medicine 92.2 (2014): 139-149.