Journal of Genetics and Genomics xxx (2018) 1–9

Contents lists available at ScienceDirect

Journal of Genetics and Genomics

Journal homepage: www.journals.elsevier.com/journal-of-genetics-and-genomics/

Original research

The DrugPattern tool for drug set enrichment analysis and its prediction for beneficial effects of oxLDL on type 2 diabetes

Chuanbo Huang [a, b, 1], Weili Yang [a, 1], Junpei Wang [a, 1], Yuan Zhou [a], Bin Geng [c], Georgios Kararigas [d], Jichun Yang [a, **], Qinghua Cui [a, e, *]

[a] Department of Biomedical Informatics, Department of Physiology and Pathophysiology, MOE Key Lab of Cardiovascular Sciences, School of Basic Medical Sciences, Center for Non-Coding RNA Medicine, Peking University, Beijing 100191, China
[b] School of Mathematics Sciences, Huaqiao University, Quanzhou 362021, China
[c] Hypertension Center, Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, State Key Laboratory of Cardiovascular Disease, National Center for Cardiovascular Diseases, Beijing 100037, China
[d] Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health, Institute of Gender in Medicine and Center for Cardiovascular Research, DZHK (German Centre for Cardiovascular Research), 10115, Berlin, Germany
[e] Center of Bioinformatics, Key Laboratory for Neuro-Information of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Article history:
Received 8 January 2018
Received in revised form
18 May 2018
Accepted 4 July 2018
Available online xxx

Keywords:
Enrichment analysis
Drug
Data mining
oxLDL
Type 2 diabetes

ABSTRACT

Enrichment analysis methods, e.g., gene set enrichment analysis, represent one class of important bioinformatical resources for mining patterns in biomedical datasets. However, tools for inferring patterns and rules of a list of drugs are limited. In this study, we developed a web-based tool, DrugPattern, for drug set enrichment analysis. We first collected and curated 7019 drug sets, including indications, adverse reactions, targets, pathways, etc. from public databases. For a list of interested drugs, DrugPattern then evaluates the significance of the enrichment of these drugs in each of the 7019 drug sets. To validate DrugPattern, we employed it for the prediction of the effects of oxidized low-density lipoprotein (oxLDL), a factor expected to be deleterious. We predicted that oxLDL has beneficial effects on some diseases, most of which were supported by evidence in the literature. Because DrugPattern predicted the potential beneficial effects of oxLDL in type 2 diabetes (T2D), animal experiments were then performed to further verify this prediction. As a result, the experimental evidences validated the DrugPattern prediction that oxLDL indeed has beneficial effects on T2D in the case of energy restriction. These data confirmed the prediction accuracy of our approach and revealed unexpected protective roles for oxLDL in various diseases. This study provides a tool to infer patterns and rules in biomedical datasets based on drug set enrichment analysis. DrugPattern is available at http://www.cuilab.cn/drugpattern.

1. Introduction

Recently, the performances of techniques in data collection, storage, processing and analysis have changed dramatically (Mayer-Schonberger, 2016). As a result, the amount of data in both biology and medicine has increased remarkably (Rumsfeld et al., 2016). So far, the open-access big data provide us with unprecedented opportunities for scientific discovery in health and disease (Rumsfeld et al., 2016). Big data analysis (BDA) has recently been used to discover and predict novel biomarkers of cardiovascular diseases (Gerstein et al., 2015), trends of flu (Ginsberg et al., 2009), cancer targets (Jiang and Liu, 2015), and novel subtypes of type 2 diabetes (Li et al., 2015). These studies reveal that BDA is exhibiting great potential in knowledge inference (Chaussabel and Pulendran, 2015) and information prediction in precision medicine (de Lemos et al., 2015). Among the biomedical BDA tools, set enrichment analysis-based methods play important roles in mining patterns, including gene set enrichment analysis tools, e.g., DAVID

* Corresponding author. Department of Biomedical Informatics, Department of Physiology and Pathophysiology, MOE Key Lab of Cardiovascular Sciences, School of Basic Medical Sciences, Peking University, Beijing 100191, China.
** Corresponding author.
E-mail addresses: yangj@bjmu.edu.cn (J. Yang), cuiqinghua@bjmu.edu.cn (Q. Cui).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jgg.2018.07.002

Bioinformatics (Huang da et al., 2009) and GSEA (Subramanian et al., 2005), miRNA set enrichment analysis tools, e.g., TAM (Lu et al., 2010), metabolite set enrichment analysis, e.g. MSEA (Xia and Wishart, 2010) and MPEA (Kankainen et al., 2011), and phenotype set enrichment analysis, e.g., PSEA (Ried et al., 2012). The fast development of drug-related techniques, such as drug-repurposing (Sleire et al., 2017) and network pharmacology (Boezio et al., 2017), leads to the generation of large lists of drugs. For a list of drugs, people always want to know whether there are some patterns and rules behind them, which could be useful for scientists to obtain valuable clues for the potential targets they are interested in. However, tools for discovering patterns and rules in such lists of drugs are limited. Recently, Napolitano et al. (2016) developed a tool named drug-set enrichment analysis (DSEA) to investigate the mode of action of drugs. However, DSEA cannot perform enrichment analysis for drug sets but for gene sets and gene pathways using the list of genes whose expression profiles can be regulated by drugs. Therefore, a tool for enrichment analysis of drug sets is greatly needed. Given the multiple effects of chemicals, particularly the therapeutic and adverse effects, a tool for enrichment analysis of drug sets could be used to evaluate the "good" and "bad" aspects of the factors of potential interest.

Accumulating datasets in public databases provide us opportunities to collect and curate a variety of drug sets. We define a drug set as a group of drugs with the same biomedical item, for example, the drugs with the same targets could be a drug set. In this study, we first collected and curated 7019 drug sets from public databases. This makes it possible to infer enriched patterns in a new drug list. For this purpose, we developed a drug set enrichment analysis tool, DrugPattern. Any factor can be analyzed using DrugPattern if it can be associated with a list of drugs. To confirm the accuracy of DrugPattern, we applied it to predict the effects of oxidized low-density lipoprotein (oxLDL). LDL, especially the oxidized form, is regarded as the bad cholesterol. Lowering oxLDL level has been an important strategy for the protection and prevention of coronary heart disease. However, unexpected findings have been reported that the beneficial effects of LDL-lowering drugs vary in different individuals (Goldstein and Brown, 2015). Our predictions and corroborating validation with animal experimentation also reveal unexpected results demonstrated by using DrugPattern.

2. Results

2.1. Generation of DrugPattern and its application for predicting oxLDL effects

We collected and curated 7019 drug sets, which cover seven big classes, including adverse drug reaction (3632 drug sets), target (1970 drug sets), pathway (603 drug sets), therapeutic or general categories (329 drug sets), disease (240 drug sets), chemical structure classification (159 drug sets), and anatomical therapeutic chemical classification (ATC, 86 drug sets). We then developed a web-based tool, DrugPattern (http://www.cuilab.cn/drugpattern), to discover regular patterns and rules in a given drug list. For a given list of interested drugs, DrugPattern calculates the enrichment of the given list of drugs in each of the 7019 drug sets. To validate the accuracy of DrugPattern, we focused on oxLDL. As shown in Fig. 1A, we first dissected the gene-expression profiles induced by oxLDL treatment. Next, by searching the gene-expression profiles induced by more than 1000 drugs in the Connectivity Map (CMap) database, which was originally designed for drug-repurposing, we obtained the drugs that show high similarity in gene-expression profiles with oxLDL. As a result, we got 66 oxLDL-similar drugs in gene-expression profiles (Table S1). Then, we entered these drugs into DrugPattern for enrichment analysis of these drugs in the 7019 drug sets. Finally, we focused on the predicted enriched disease of these drugs. The prediction suggested that oxLDL has potential to defend some certain diseases. This means that oxLDL could have beneficial effects on these diseases (Fig. 1B).

The link of LDL to coronary heart disease as a major risk factor represents a most important biomedical finding in the 20th century (Goldstein and Brown, 2015). Moreover, LDL-lowering therapy has been an important strategy for the protection and prevention of coronary heart disease (Goldstein and Brown, 2015). However, in the new century, accumulating evidences have revealed unexpected results. For example, the association between serum total cholesterol level and all-cause mortality is negative but not positive (Hamazaki et al., 2015; Ravnskov et al., 2016). Therefore, it is becoming timely necessary to re-evaluate the roles of LDL, the cholesterol carrier, in various diseases. In the era of precision medicine, this is critical to identify the patients who should receive LDL-lowering therapy and those who should not. In this study, we used DrugPattern to infer the effects of oxLDL, which has always been considered to be more deleterious than LDL. As a result, DrugPattern predicted that oxLDL has beneficial effects on a number of diseases including bacterial infection, type 2 diabetes, leprostatic, influenza virus, cancer, glaucoma and depression (Fig. 1B). Literature mining further confirmed that high cholesterol level may indeed protect against infection (Ravnskov, 2003), cancer (Ravnskov et al., 2016) and depression (Persons and Fiedorowicz, 2016), suggesting that DrugPattern has a good prediction accuracy. In the case of type 2 diabetes, increased oxLDL levels have been observed in type 2 diabetic patients (Marin et al., 2015), leading to the hypothesis that oxLDL may contribute to T2D (Marin et al., 2015). However, due to conflicting published reports, the relationship between oxLDL and type 2 diabetes still remains unclear.

2.2. Animal experiments confirmed the beneficial effects of oxLDL on type 2 diabetes

Given the potential beneficial effects of oxLDL on type 2 diabetes predicted by DrugPattern, animal experiments were then performed to verify this prediction. For doing so, mice fed on high fat diet (HFD) for 4 months were treated with human oxLDL for 3 weeks. In clinical practice, obese type 2 diabetic patients are always suggested to control their body weight and energy intake. To mimic these clinical strategies, HFD was replaced with normal diet (ND) once the mice began receiving oxLDL treatment. During the period of oxLDL treatment, the body weight of both the oxLDL-treated and control groups of mice decreased gradually, but there was no obvious difference between them at each time point (Fig. 2A). oxLDL treatment also had no significant effect on the weight of liver, white adipose, heart and kidney (data not shown). At 7, 14 and 21 days post oxLDL treatment, the fasting blood glucose levels of oxLDL-treated mice were significantly lower than that of control mice (Fig. 2B). Oral glucose tolerance tests (OGTTs) revealed that the overall glucose intolerance was also significantly improved at 14 days post oxLDL treatment onset (Fig. 2C–F). Insulin tolerance tests (ITTs) indicated that the global insulin sensitivity was also increased at 18 days post oxLDL treatment (Fig. 3A and B). In support, serum insulin level has a tendency to decrease after oxLDL treatment ($P = 0.057$, Fig. 3C). Both serum triglyceride and cholesterol levels were also reduced after oxLDL treatment (Fig. 3D and E). oxLDL treatment slightly increased serum aspartate aminotransferase (AST) and alanine transaminase (ALT) activities, but no statistical difference was found when compared with control mice (Fig. 4A and B). Both the mRNA and protein levels of gluconeogenic factors PEPCK and G6Pase were reduced, whereas mRNA levels of

Please cite this article in press as: Huang, C., et al., The DrugPattern tool for drug set enrichment analysis and its prediction for beneficial effects of oxLDL on type 2 diabetes, Journal of Genetics and Genomics (2018), https://doi.org/10.1016/j.jgg.2018.07.002