

# Dr AFC: drug repositioning through anti-fibrosis characteristic

Dingfeng Wu<sup>†</sup>, Wenxing Gao<sup>†</sup>, Xiaoyi Li<sup>†</sup>, Chuan Tian, Na Jiao, Sa Fang, Jing Xiao, Zhifeng Xu, Lixin Zhu, Guoqing Zhang and Ruixin Zhu

Corresponding authors: Ruixin Zhu, Putuo People's Hospital, Department of Bioinformatics, Tongji University, 1239 Siping Road, Shanghai 200092, P. R. China. Tel: 86-21-6598-1041. E-mail: rxzhu@tongji.edu.cn; Guoqing Zhang, Bio-Med Big Data Center, Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai 200031, P. R. China. Tel: 86-21-5492-0465. E-mail: gqzhang@picb.ac.cn; Lixin Zhu, Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, Department of Colorectal Surgery, the Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510655, P. R. China. Tel: 86-199-46256235. E-mail: zhulx6@mail.sysu.edu.cn

<sup>†</sup>These authors contributed equally to this work.

## Abstract

Fibrosis is a key component in the pathogenic mechanism of a variety of diseases. These diseases involving fibrosis may share common mechanisms and therapeutic targets, and therefore common intervention strategies and medicines may be applicable for these diseases. For this reason, deliberately introducing anti-fibrosis characteristics into predictive modeling may lead to more success in drug repositioning. In this study, anti-fibrosis knowledge base was first built by collecting data from multiple resources. Both structural and biological profiles were then derived from the knowledge base and used for constructing machine learning models including Structural Profile Prediction Model (SPPM) and Biological Profile Prediction Model (BPPM). Three external public data sets were employed for validation purpose and further exploration of potential repositioning drugs in wider chemical space. The resulting SPPM and BPPM models achieve area under the receiver operating characteristic curve (area under the curve) of 0.879 and 0.972 in the training set, and 0.814 and 0.874 in the testing set. Additionally, our results also demonstrate that substantial amount of multi-targeting natural products possess notable anti-fibrosis characteristics and might serve as encouraging candidates in fibrosis treatment and drug repositioning. To leverage our methodology and findings, we developed repositioning prediction platform, drug repositioning based on anti-fibrosis characteristic that is freely accessible via <https://www.biosino.org/drafc>.

**Key words:** fibrosis; drug repositioning; web server; knowledge base; natural products

**Dingfeng Wu** is a postdoc fellow at Tongji University, Shanghai, China. His current research interests include machine learning and drug repositioning. **Wenxing Gao, Xiaoyi Li, Sa Fang and Jing Xiao** are Master's Degree students at Tongji University, Shanghai, China. They are working on bioinformatics and multi-omics analysis.

**Chuan Tian** is a scientist at Relay Therapeutics, Cambridge, United States. His current research interests include machine learning and drug repositioning. **Na Jiao** is a postdoc fellow at Sun Yat-sen University, Guangzhou, China. Her current research interests include drug repositioning and gut microbiome.

**Zhifeng Xu** is a junior student at Tongji University, Shanghai, China. He is working on bioinformatics.

**Lixin Zhu** is an associate professor at Sun Yat-sen University, Guangzhou, China. His current research interests include drug repositioning and gut microbiome.

**Guoqing Zhang** is a professor at Shanghai Institute of Nutrition and Health, Chinese Academy of Sciences, Shanghai, China. His current research interests include machine learning and natural language processing.

**Ruixin Zhu** is a professor at Tongji University, Shanghai, China. His current research interests include drug repositioning, microbiome and bioinformatics.

Submitted: 30 March 2020; Received (in revised form): 13 May 2020

© The Author(s) 2020. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Introduction

Fibrosis is defined as the process of excessive accumulation of fibrous connective tissue in most tissues or organs, where normal cell injury, transdifferentiation and abnormal proliferation, and deposition of extracellular matrix can disrupt tissue function. In the new era of 21st century, the morbidity and mortality rates of various fibrotic diseases have increased progressively, bringing an enormous global health burden. In developed countries, fibroproliferative diseases are responsible for nearly 45% of deaths [1]. One of the well-known fibrotic diseases, idiopathic pulmonary fibrosis (IPF), has a poor prognosis within the 5-year survival rate less than 30% and median survival ranging from 3 to 5 years [2]. The situation of IPF patients is even worse than those with various types of cancers [3]. As data obtained by Clinical Practice Research Datalink revealed, the prevalence of IPF patients in board case definitions has doubled from 19.94 per 100 000 patients in 2000 to 38.82 per 100 000 patients in 2012, and a 80% increase in incidence was observed [4]. Another life-threatening fibrotic disease, cardiac fibrosis, is one of the leading factors causing heart failure [5]. A research from 2008 to 2014 revealed that in 318 patients with systolic dysfunction, 78% had one type of myocardial fibrosis, whereas 25% had at least two types [6].

The polypharmacology of most anti-fibrosis drugs could improve therapeutic efficacy. Recent studies have found that, firstly, fibrosis is the common pathogenic process in most diseases. For example, there are multiple common cellular processes between lung cancer and IPF, including inflammation, cell apoptosis and tissue infiltration [7]. Secondly, fibrosis-related processes have common mechanisms, targets and drugs [8, 9]. A multi-organ fibrosis research discovered 90 common differentially expressed genes across lung, heart, liver and kidney. In the two most active gene networks generated by ingenuity pathway analysis, these genes play a key role in connective tissue disorders and genetic, skeletal and muscular disorders [10]. Similarly, another multi-organ fibrosis research also obtained 11 metzincin-related differentially expressed genes across heart, lung, liver, kidney and pancreas including *THBS2*, *TIMP1*, *COL1A2*, *COL3A1*, *HYOU1*, *MMP2* and *MMP7* [11]. Thirdly, fibrosis is a complicated pathological process involving multiple pathways, thus multi-target drugs are more appropriate for fibrosis-related diseases [9]. Different pathways interact and counter-interact with each other to establish a 'check-and-balance' system, for instance, the core regulators, transforming growth factor- $\beta$  and connective tissue growth factor signaling pathways could collaborate to elicit pulmonary and renal fibrosis [12, 13]. In summary, these evidences indicate that anti-fibrosis intervention strategies and medicines may be applicable for diseases that were not originally considered and targeted through targeting their common fibrosis-related mechanisms. Therefore, compounds that more specifically target anti-fibrosis could have greater potential of repositioning and are more applicable for drug repositioning research.

Drug repositioning or repurposing refers to the 'reuse of old drugs', recycling existing drugs for new medical indications. Compared with *de novo* drug discovery, drug repositioning has obvious advantages such that it could significantly shorten drug development periods, reduce laboratory cost and minimize potential safety risk. Nowadays, drug repositioning is one of the most productive strategies in drug development [14]. With the advancement of high-throughput sequencing technology and deep learning, various data-driven computational prediction and analytic models stand out [15, 16], including similarity

ensemble approach (SEA) [17] and connectivity map (CMap) [18]. SEA clusters ligands into sets and calculates the similarity scores between ligand sets from ligand topology [17]. CMap computes the similarity to 'signatures' deduced from compound-induced gene profiles to quantify the biological functional relationships between compounds. Moreover, the relationship between compounds and diseases could also be quantified in reversed manner [18]. However, with countless repositioning methods and algorithms developed [19–21], there are still no attempts hitherto in introducing anti-fibrosis characteristic into drug repositioning strategy.

For the first time, we built the anti-fibrosis knowledge base upon anti-fibrosis targeted research. Based on the knowledge base, two repositioning models, Structural Profile Prediction Model (SPPM) and Biological Profile Prediction Model (BPPM) were constructed to achieve high prediction accuracy. Centered on these two models, we then developed a repositioning computing platform, drug repositioning based on anti-fibrosis characteristic (Dr AFC), to accelerate the exploratory process of repositioning drugs and contribute to the cutting-edge study of its underlying mechanisms.

## Materials and methods

### Data sets

#### Anti-fibrosis knowledge base

Anti-fibrosis-related literatures were collected through key word query 'fibrosis AND target' from PubMed from January 1 2000 to October 31 2019. The compound–target interaction information on 'fibrosis' were collected from Comparative Toxicogenomics Database (CTD) [22] from January 1 2000 to October 31 2019. Anti-fibrosis trials were collected from [ClinicalTrials.gov](http://ClinicalTrials.gov) [23] from January 1 2000 to October 31 2019. Meanwhile, approved anti-fibrosis drugs from DrugBank (Version 5.1.3) were collected. Finally, anti-fibrosis treatments, targets and compound–target interactions were extracted and aggregated into the knowledge base.

#### Model construction

Structural and biological profiles of compounds were collected from DrugBank [24] and CMap (build02), respectively, and used for model construction. A total of 2408 approved drugs in DrugBank and 1223 compounds in the anti-fibrosis knowledge base served as the raw data for SPPM construction. In total, 6100 biological profiles (gene expression) of 1309 small molecules in CMap served as the raw data for BPPM.

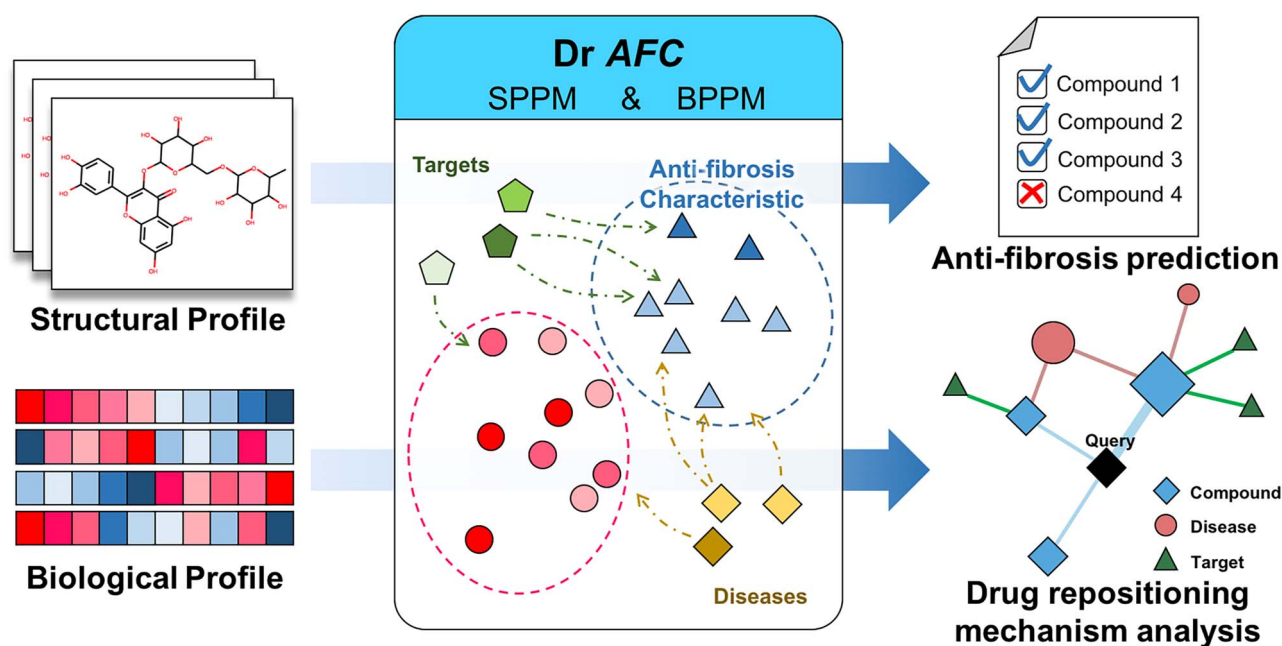
#### Case studies

A total of 20 263 natural products from Traditional Chinese Medicine Integrated Database (TCMID) [25], 5968 DrugBank experimental drugs [24] and 5000 random compounds from ChEMBL [26] were collected as external validations and used for SPPM case studies. And external biological profiles from Gene Expression Omnibus database (GSE85871) that contains transcriptomics perturbation profiles of 105 natural products in Michigan Cancer Foundation-7 cell line were used for BPPM case studies.

## Methods

### Pre-processing of modeling data

In raw chemical structures (from DrugBank approved drugs and the anti-fibrosis knowledge base) and biological profiles (from



**Figure 1.** The schematic of Dr AFC construction. Based on SPPM and BPPM, Dr AFC was constructed, which allows users to upload compound structures or compound-induced biological profiles for repositioning potential prediction.

CMap) data, compounds that exist in the anti-fibrosis knowledge base were labeled as positive candidates, whereas the rest were labeled as negative candidates. Then, chemical structures were converted into chemical fingerprints (166-bits MACCS keys) for processing chemical information in a fast and convenient way using RDKit [27]. As to biological profiles, Quantile Transformer was used to transform biological profiles into ranking orders to improve model generalizability and also made data sets from different batches and platforms more comparable.

One-class support vector machine ( $\nu=0.3$ ) was performed to estimate sample quality, remove outliers and confirm final positive and negative samples. In total, 70% of final samples were used as training set for model selection and super-parameter determination, whereas the remainder as testing set for model validation.

#### Anti-fibrosis model construction and validation

Four different machine learning algorithms were selected for modeling on training set, including logistic regression with l1 and l2 penalties, decision tree, random forest and gradient boosting. Among them, method with highest precision and area under the curve (AUC) calculated by 5-fold cross-validation was selected for subsequent analysis. Iterative feature elimination (IFE) algorithm was performed to select optimal feature set through one-by-one feature deletion. Finally, SPPM and BPPM were constructed based on optimal modeling algorithm and feature set, and further validated by testing set. The details of model construction are shown in [Supplementary Figure S1](#).

#### Drug repositioning mechanism analysis

Network-based inference approaches are widely applied in the realm of drug repositioning [20, 21]. Here, we infer the potential drug repositioning mechanism through compound-target-disease network. Firstly, based on SPPM and BPPM, the repositioning characteristics of compounds were predicted through

their structural or biological profiles, in which compounds with reposition score  $> 0.5$  were considered as anti-fibrosis and having repositioning potential. Next, the anti-fibrosis characteristic and potential repositioning mechanisms of these candidates were explored on the basis of compound-target-disease corresponding information in the anti-fibrosis knowledge base. Similar compounds that may interact with same targets and diseases were calculated through Tanimoto similarity of chemical structural fingerprints or Spearman's rank correlation coefficient of biological profiles. Targets and disease information of compounds reported in previous researches were refined from the anti-fibrosis knowledge base to explore anti-fibrosis mechanism of compounds. Finally, the potential mechanisms among compounds in compound-target-disease network displayed in drug repositioning analysis were used to help propose feasible drug repositioning solutions.

#### Webserver construction of Dr AFC

Dr AFC was constructed through PostgreSQL database and Django framework. This platform serves as a practical tool for prediction of drug repositioning potential based on compound structures (via SPPM) and biological profiles (via BPPM) as well as displaying compound-target-disease network of drug repositioning mechanisms. Meanwhile, Dr AFC also integrated toolkits such as quantitative estimate of drug-likeness from Silicos-it [28], and similarity calculation and structure matching borrowed from RDkit to provide convenient web-based calculations for users.

The overall process is shown in [Figure 1](#).

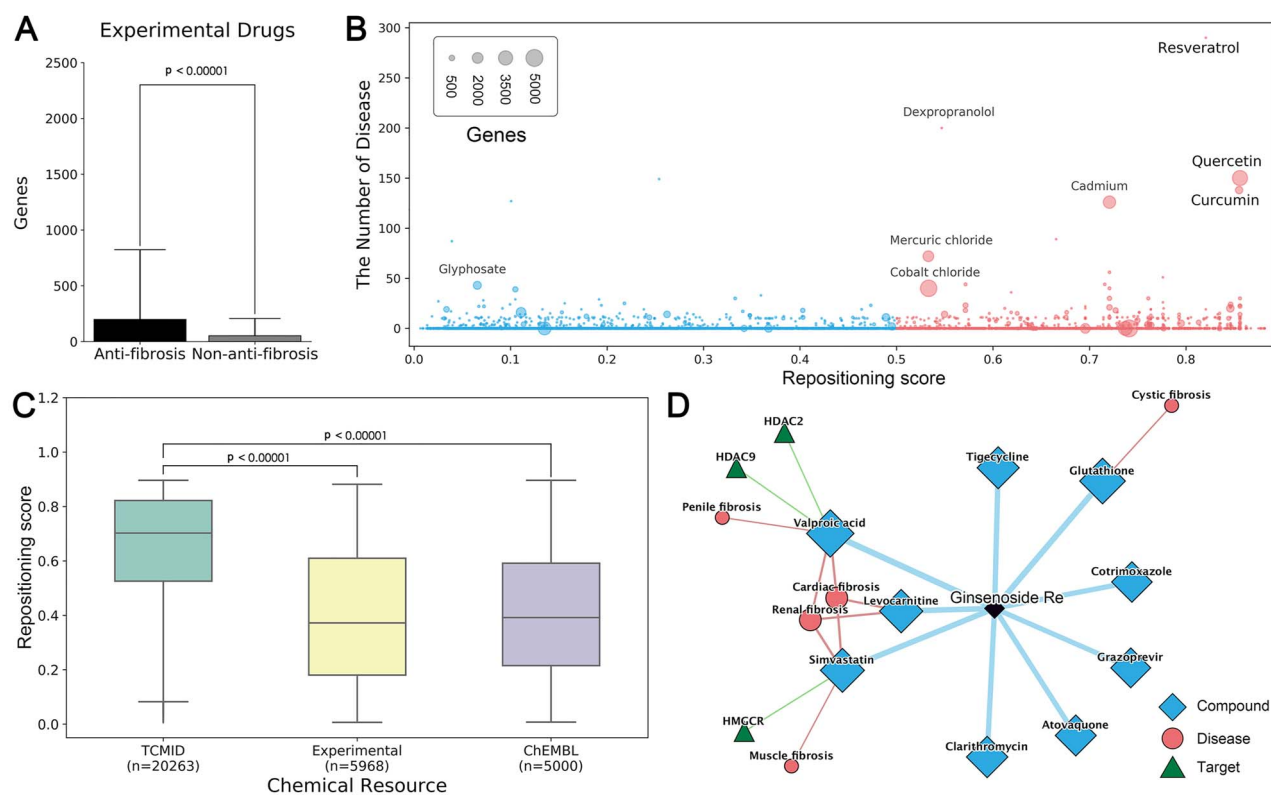
## Results

### SPPM and BPPM show high performances for anti-fibrosis prediction

To construct the anti-fibrosis knowledge base, 7058 fibrosis-related references from PubMed, 302 from CTD [22] and







**Figure 3.** Case studies of Dr AFC. (A) Comparison of the number of genes interacting with compounds predicted as anti-fibrosis and non-anti-fibrosis. (B) The distribution of related genes, diseases and repositioning score for Drugbank experimental drugs. Compounds with repositioning score > 0.5 were considered as anti-fibrosis and had repositioning potential. (C) The distribution of repositioning scores in different data sets. (D) Drug repositioning mechanism analysis of ginsenoside Re by Dr AFC. (P-value was calculated by two-sided Wilcoxon rank sum test).

presumably more genetic effects thus greater repositioning potential (Figure 3A).

In Drugbank experimental drugs, multiple drugs with great repositioning potential (related genes > 500 and diseases > 20, Figure 3B, Supplementary Table S3) were already developed for fibrotic diseases and other diseases. Quercetin was discovered to ameliorate liver fibrosis through regulating macrophage infiltration and polarization, and it could alleviate IPF through fibroblasts apoptosis [29, 30]. Based on our results, we confirm that quercetin interacts with numerous genes and is strongly linked to multiple diseases (repositioning score=0.856, related genes=3938 and diseases=150, Supplementary Table S3). Another natural compound from turmeric (extracted from turmeric plant), curcumin (repositioning score=0.855, related genes=903 and diseases=138, Supplementary Table S3), is also believed for treating multiple fibrotic diseases. It could inhibit fibroblast proliferation and myofibroblast differentiation in IPF [31], whereas inhibit oxidative stress and exhibit anti-inflammatory effect in liver fibrosis [32]. Apart from fibrosis, curcumin has also been applied for osteoarthritis and rheumatoid arthritis treatment [33, 34]. Moreover, other drugs, such as resveratrol also has great repositioning potential (repositioning score=0.821, Figure 3B).

#### Natural compounds are the better repositories for drug repositioning

To expand the resources of potential repositioning drugs and further explore the chemical space, we introduced two

external molecule sets, natural products from TCMID [25] and random compounds from ChEMBL [26]. SPPM was used to predict the repositioning potential of compounds from both external molecule sets in a structural perspective. The results show that 35.42, 77.26 and 37.04% of compounds could be potentially repositioned from DrugBank experimental drugs, TCMID and ChEMBL, respectively. The reserves in natural products from TCMID are significantly higher than others, indicating that natural products are promising repositioning repositories and worth further investigations (Figure 3C).

From a genetic perspective, BPPM was used to predict the repositioning potential of 105 natural products (GSE85871) using their gene profiles. The results show that a total of 66 natural products have anti-fibrosis characteristic and repositioning potential, including ginsenoside Re (repositioning score=0.979), muscone (repositioning score=0.974) and cinnamic acid (repositioning score=0.948) (Supplementary Table S4). Among them, ginsenoside Re is potentially impacting HDAC2, HDAC9 and HMGCR and playing anti-fibrosis roles via 'inflammation', 'preventing collagen deposition' and 'targeting myeloperoxidase', as reported by drug repositioning mechanism analysis tools from Dr AFC (Figure 3D). Ginsenoside Re is the extract of *Panax ginseng*, which exhibit protective effects in neural and systematic inflammations through inhibiting the interaction between LPS and TLR4 in macrophages [35]. It was reported to exert anti-fibrosis effect on cardiac fibrosis through down-regulating the expression of p-Smad3, collagen I and reducing the augmentation of collagen fibers [36]. Apart from fibrosis, ginsenoside Re could alleviate inflammation through inhibiting



matching tools. Users could upload their compounds in SMILES and perform these ligand-based calculations. Drug-likeness estimation could evaluate and score the compound drug-likeness, ranging from 0 to 1 with higher score indicating higher potential for lead compound. Chemical similarity calculation and structure matching provide convenient ways for users to search compound with similar structures, same structures or substructures, supporting single compound calculation and simultaneous calculation for multiple compounds.

## Discussion

Fibrosis is the common mechanism underlying diseases that has attracted global attention. Anti-fibrosis characteristic of a compound could essentially infer its great repositioning potential. However, the anti-fibrosis characteristic has not been thoroughly considered and extensively applied into the realm of drug discovery till now. In this study, we first bridge the gap by developing a platform that can provide intensive information conveniently on Dr AFC data (<https://www.biosino.org/drafc>). This *in silico* platform also provides a highly accurate way to generate data for rational drug design via combining the advanced machine-learning algorithm.

Dr AFC was built upon the anti-fibrosis knowledge base, which pioneered the curation and deployment of fibrosis-related studies over the past decade. Structural profile (SPPM) and biological profile (BPPM) that show extraordinary capabilities in drug repositioning prediction (with AUC 0.814 and 0.874, respectively) were integrated into Dr AFC. BPPM shows slightly higher performance than SPPM based on AUC metric. The possible reason could be that biological profile is more tolerant and could contain information reflecting an overall effect of compound in the body. Biological profile has shown its advantage in multiple repositioning algorithms previously, such as CMap [18], L1000CDS<sup>2</sup> [39] and MANTRA [40]. Besides, certain therapies without available structural profile such as biotech drugs or cocktail therapies could also be studied in repositioning research based on their biological profiles.

In BPPM, 47 biological markers exhibited strong prediction capabilities. These genes are directly or indirectly linked to various fibrotic diseases. Notably, ribosomal proteins including RPL30, MRPL15, RPL32, RPS3A, RPLP0, RPL7, RPL23A and RPL13A are the main part of these biological markers. Ribosomes serve as significant regulators in immune signaling pathways, tumorigenesis pathways, and cardiovascular and metabolic diseases [41, 42]. For example, the expression of RPL30 is negatively correlated with carcinogenesis process in medulloblastoma that is usually accompanied by desmoplasia and could thus serve as a prognosis biomarker [43]. Besides, the over-activation of ribonucleic acid (RNA) polymerase in the biogenesis of ribosomes could enhance protein synthesis and decrease translation accuracy, in turn triggering cancers or exacerbating cancer processes [44]. Furthermore, some biological markers are associated with the spliceosome formation including RBM8A, HNRNPA3, SNRPG and DHX15. Spliceosome is a large molecular machine composed of five small nuclear ribonucleic acids and many other proteins, and serves as the catalyzer of pre-RNA introns that are crucial for protein expression and function. It has been reported to be closely associated with multiple diseases including cystic fibrosis and pulmonary fibrosis [45, 46].

Based on external molecule sets, natural products are validated to have the most appealing anti-fibrosis characteristics and repositioning potential among chemicals from different

sources. Natural products provide a wealth of valuable natural resources for modern medicine and are seen as promising and popular candidates for drug repositioning studies [47]. Their natural scaffold novelty, structural complexity, abundant stereochemistry and 'metabolite-likeness' mainly account for their broad-spectrum of biological activities [48, 49]. The multi-targeting and synergistic effects of natural products exhibit great advantages in treating diseases undergoing sophisticated mechanisms, such as fibrosis [50]. Our studies show that natural products such as ginsenoside have great anti-fibrosis characteristic and repositioning potential and should give top priority to consider repositioned drug discovery. Additionally, the natural products in Drugbank experimental drugs, such as quercetin, curcumin and resveratrol, also highlight their strong repositioning capabilities. Therefore, natural products are a fruitful and promising source for future drug development studies.

## Conclusion

In summary, based on anti-fibrosis characteristics, we constructed two predictive repositioning models, SPPM and BPPM, which predict the anti-fibrosis characteristics and repositioning potential from compound structures and/or compound-induced biological profiles. SPPM and BPPM take the advantage of therapeutic commonality and universality of fibrotic diseases, and dramatically increase the success rate of drug repositioning predications. This study not only established a highly efficient strategy of predicting repositioning, but also developed a convenient and user-friendly computing platform, Dr AFC (<https://www.biosino.org/drafc>), for studying fibrosis mechanisms and drug repositioning.

### Key Points

- Fibrosis is the common mechanism of diseases, which could be applied in drug repositioning.
- We developed a convenient and user-friendly computing platform, Dr AFC, for studying fibrosis mechanisms and drug repositioning.
- Dr AFC shows high performance on both cross-validation and external validation, which demonstrates its potential applications in drug discovery.
- Natural compounds proved to be the better repositories for drug repositioning.

## Supplementary Data

Supplementary data are available online at <https://academic.oup.com/bib>.

## Funding

National Key R&D Program of China (2017YFC0907505, 2016YFC0901904, 2017YFC0908404 to G.Z.); National Natural Science Foundation of China (81774152 to R.Z., 81770571 to L.Z.); National Postdoctoral Program for Innovative Talents of China (BX20190393 to N.J.); China Postdoctoral Science Foundation (2019M651568 to D.W., 2019M663252 to N.J.); National Science Foundation of Shanghai (16ZR1449800 to R.Z.); Science and Technology Service Network Initiative of Chinese Academy of Sciences (Y919C11011 to G.Z.);

University at Buffalo Community of Excellence in Genome, Environment and Microbiome (to L.Z.).

## References

- Wynn TA. Cellular and molecular mechanisms of fibrosis. *J Pathol* 2008;**214**:199–210.
- Vancheri C, du Bois RM. A progression-free end-point for idiopathic pulmonary fibrosis trials: lessons from cancer. *Eur Respir J* 2013;**41**:262–9.
- Homma S, Bando M, Azuma A, et al. Japanese guideline for the treatment of idiopathic pulmonary fibrosis. *Respir Investig* 2018;**56**:268–91.
- Strongman H, Kausar I, Maher TM. Incidence, prevalence, and survival of patients with idiopathic pulmonary fibrosis in the UK. *Adv Ther* 2018;**35**:724–36.
- Horn MA, Trafford AW. Aging and the cardiac collagen matrix: novel mediators of fibrotic remodelling. *J Mol Cell Cardiol* 2016;**93**:175–85.
- Almehmadi F, Joncas SX, Nevis I, et al. Prevalence of myocardial fibrosis patterns in patients with systolic dysfunction: prognostic significance for the prediction of sudden cardiac arrest or appropriate implantable cardiac defibrillator therapy. *Circ Cardiovasc Imaging* 2014;**7**:593–600.
- Ballester B, Milara J, Cortijo J. Idiopathic pulmonary fibrosis and lung cancer: mechanisms and molecular targets. *Int J Mol Sci* 2019;**20**:593.
- Li X, Zhu L, Wang B, et al. Drugs and targets in fibrosis. *Front Pharmacol* 2017;**8**:855.
- Liu J, Sun D, Liu J, et al. FibroAtlas: a database for the exploration of fibrotic diseases and their genes. *Cardiol Res Pract* 2019;**2019**:4237285.
- Wenzke KE, Cantemir-Stone C, Zhang J, et al. Identifying common genes and networks in multi-organ fibrosis. *AMIA Jt Summits Transl Sci Proc* 2012;**2012**:106–15.
- Rodder S, Scherer A, Korner M, et al. A subset of metzincins and related genes constitutes a marker of human solid organ fibrosis. *Virchows Arch* 2011;**458**:487–96.
- Meng XM, Nikolic-Paterson DJ, Lan HY. TGF-beta: the master regulator of fibrosis. *Nat Rev Nephrol* 2016;**12**:325–38.
- Wang Q, Usinger W, Nichols B, et al. Cooperative interaction of CTGF and TGF-beta in animal models of fibrotic disease. *Fibrogenesis Tissue Repair* 2011;**4**:4.
- Xue H, Li J, Xie H, et al. Review of drug repositioning approaches and resources. *Int J Biol Sci* 2018;**14**:1232.
- Qiu TY, Qiu JX, Feng J, et al. The recent progress in proteochemometric modelling: focusing on target descriptors, cross-term descriptors and application scope. *Brief Bioinform* 2017;**18**:125–36.
- Liu L, Tsompana M, Wang Y, et al. Connection map for compounds (CMC): a server for combinatorial drug toxicity and efficacy analysis. *J Chem Inf Model* 2016;**56**:1615–21.
- Keiser MJ, Roth BL, Armbruster BN, et al. Relating protein pharmacology by ligand chemistry. *Nat Biotechnol* 2007;**25**:197–206.
- Lamb J, Crawford ED, Peck D, et al. The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006;**313**:1929–35.
- Sam E, Athri P. Web-based drug repurposing tools: a survey. *Brief Bioinform* 2019;**20**:299–316.
- Wu Z, Cheng F, Li J, et al. SDTNBI: an integrated network and chemoinformatics tool for systematic prediction of drug-target interactions and drug repositioning. *Brief Bioinform* 2017;**18**:333–47.
- Peng Y, Wang M, Xu Y, et al. Drug repositioning by prediction of drug's anatomical therapeutic chemical code via network-based inference approaches. *Brief Bioinform* 2020;**1**–15. doi: 10.1093/bib/bbaa027.
- Davis AP, Grondin CJ, Johnson RJ, et al. The Comparative Toxicogenomics Database: update 2019. *Nucleic Acids Res* 2019;**47**:D948–54.
- Zarin DA, Fain KM, Dobbins HD, et al. 10-year update on study results submitted to clinicaltrials.gov. *N Engl J Med* 2019;**381**:1966–74.
- Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* 2018;**46**:D1074–82.
- Huang L, Xie D, Yu Y, et al. TCMID 2.0: a comprehensive resource for TCM. *Nucleic Acids Res* 2018;**46**:D1117–20.
- Mendez D, Gaulton A, Bento AP, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res* 2019;**47**:D930–40.
- Lovrić M, Molero JM, Kern R. PySpark and RDKit: Moving towards Big Data in Cheminformatics. *Mol Inform* 2019;**38**:e1800082.
- Sanders MP, Barbosa AJ, Zarzycka B, et al. Comparative analysis of pharmacophore screening tools. *J Chem Inf Model* 2012;**52**:1607–20.
- Li X, Jin Q, Yao Q, et al. The flavonoid quercetin ameliorates liver inflammation and fibrosis by regulating hepatic macrophages activation and polarization in mice. *Front Pharmacol* 2018;**9**:72.
- Hohmann MS, Habel DM, Coelho AL, et al. Quercetin enhances ligand-induced apoptosis in senescent idiopathic pulmonary fibrosis fibroblasts and reduces lung fibrosis in vivo. *Am J Respir Cell Mol Biol* 2019;**60**:28–40.
- Smith MR, Gangireddy SR, Narala VR, et al. Curcumin inhibits fibrosis-related effects in IPF fibroblasts and in mice following bleomycin-induced lung injury. *Am J Physiol Lung Cell Mol Physiol* 2010;**298**:L616–25.
- Bruck R, Ashkenazi M, Weiss S, et al. Prevention of liver cirrhosis in rats by curcumin. *Liver Int* 2007;**27**:373–83.
- Henrotin Y, Priem F, Mobasher A. Curcumin: a new paradigm and therapeutic opportunity for the treatment of osteoarthritis: curcumin for osteoarthritis management. *Springerplus* 2013;**2**:56.
- Chandran B, Goel A. A randomized, pilot study to assess the efficacy and safety of curcumin in patients with active rheumatoid arthritis. *Phytother Res* 2012;**26**:1719–25.
- Kim JH, Yi YS, Kim MY, et al. Role of ginsenosides, the main active components of *Panax ginseng*, in inflammatory responses and diseases. *J Ginseng Res* 2017;**41**:435–43.
- Wang QW, Yu XF, Xu HL, et al. Ginsenoside re improves isoproterenol-induced myocardial fibrosis and heart failure in rats. *Evid Based Complement Alternat Med* 2019;**2019**:3714508.
- Lee IA, Hyam SR, Jang SE, et al. Ginsenoside re ameliorates inflammation by inhibiting the binding of lipopolysaccharide to TLR4 on macrophages. *J Agric Food Chem* 2012;**60**:9595–602.
- Lee S, Lee MS, Kim CT, et al. Ginsenoside Rg3 reduces lipid accumulation with AMP-activated protein kinase (AMPK) activation in HepG2 cells. *Int J Mol Sci* 2012;**13**:5729–39.
- Duan Q, Reid SP, Clark NR, et al. L1000CDS(2): LINCS L1000 characteristic direction signatures search engine. *NPJ Syst Biol Appl* 2016;**2**:16015.
- Carrella D, Napolitano F, Rispoli R, et al. Mantra 2.0: an online collaborative resource for drug mode of action and



- repurposing by network analysis. *Bioinformatics* 2014;**30**:1787–8.
41. Zhou X, Liao WJ, Liao JM, et al. Ribosomal proteins: functions beyond the ribosome. *J Mol Cell Biol* 2015;**7**:92–104.
  42. Wang W, Nag S, Zhang X, et al. Ribosomal proteins and human diseases: pathogenesis, molecular mechanisms, and therapeutic implications. *Med Res Rev* 2015;**35**:225–85.
  43. De Bortoli M, Castellino RC, Lu XY, et al. Medulloblastoma outcome is adversely associated with overexpression of EEF1D, RPL30, and RPS20 on the long arm of chromosome 8. *BMC Cancer* 2006;**6**:223.
  44. Pelletier J, Thomas G, Volarevic S. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat Rev Cancer* 2018;**18**:51–63.
  45. Box JA, Bunch JT, Tang W, et al. Spliceosomal cleavage generates the 3' end of telomerase RNA. *Nature* 2008;**456**:910–4.
  46. Liu X, Jiang Q, Mansfield SG, et al. Partial correction of endogenous DeltaF508 CFTR in human cystic fibrosis airway epithelia by spliceosome-mediated RNA trans-splicing. *Nat Biotechnol* 2002;**20**:47–52.
  47. Shen B. A new golden age of natural products drug discovery. *Cell* 2015;**163**:1297–300.
  48. Harvey AL, Edrada-Ebel R, Quinn RJ. The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov* 2015;**14**:111–29.
  49. Zhao C, Ye Z, Ma ZX, et al. A general strategy for diversifying complex natural products to polycyclic scaffolds with medium-sized rings. *Nat Commun* 2019;**10**:4015.
  50. Yan H, Huang Z, Bai Q, et al. Natural product andrographolide alleviated APAP-induced liver fibrosis by activating Nrf2 antioxidant pathway. *Toxicology* 2018;**396–397**:1–12.