Integrating population-level and cell-based signatures for drug repositioning

Chunfeng He[1,2,#], Yue Xu[1,2,#], Yuan Zhou[3], Jiayao Fan[1,2], Chunxiao Cheng[1], Ran Meng[1], Eric R. Gamazon[4,5], Dan Zhou[1,2]

[1]School of Public Health and the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China;
[2]The Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province, Hangzhou, Zhejiang, China;
[3]Department of Biostatistics and Center for Quantitative Sciences, Vanderbilt University Medical Center, Nashville, TN, USA;
[4]Vanderbit Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA;
[5]Data Science Institute, Vanderbilt University Medical Center, Nashville, TN, USA;

[#]equally contributed

[*]Corresponding to:
Dan Zhou: zdangm@gmail.com
Eric Gamazon; Email: egamazon@gmail.com

Abstract

Drug repositioning presents a streamlined and cost-efficient way to expand the range of therapeutic possibilities. Furthermore, drugs with genetic evidence are more likely to progress successfully through clinical trials towards FDA approval. Exploiting these developments, single gene-based drug repositioning methods have been implemented, but approaches leveraging the entire spectrum of molecular signatures are critically underexplored. Most multi-gene-based approaches rely on differential gene expression (DGE) analysis, which is prone to identify the molecular consequence of disease and renders causal inference challenging. We propose a framework TReD (Transcriptome-informed Reversal Distance) that integrates population-level disease signatures robust to reverse causality and cell-based drug-induced transcriptome response profiles. TReD embeds the disease signature and drug profile in a high-dimensional normed space, quantifying the reversal potential of candidate drugs in a disease-related cell screen assay. The robustness is ensured by evaluation in additional cell screens. For an application, we implement the framework to identify potential drugs against COVID-19. Taking transcriptome-wide association study (TWAS) results from four relevant tissues and three DGE results as disease features, we identify 37 drugs showing potential reversal roles in at least four of the seven disease signatures. Notably, over 70% (27/37) of the drugs have been linked to COVID-19 from other studies, and among them, eight drugs are supported by ongoing/completed clinical trials. For example, TReD identifies the well-studied JAK1/JAK2 inhibitor baricitinib, the first FDA-approved immunomodulatory treatment for COVID-19. Novel potential candidates, including enzastaurin, a selective inhibitor of PKC-beta which can be activated by SARS-CoV-2, are also identified. In summary, we propose a comprehensive genetics-anchored framework integrating population-level signatures and cell-based screens that can accelerate the search for new therapeutic strategies.

Key words

drug repositioning; signature matching; TWAS; reversal distance; COVID-19

Introduction

The conventional model of drug development is commonly perceived as a time-consuming and costly endeavor, with recent statistics revealing that it typically spans 10-15 years, and necessitates a financial investment exceeding $1.3 billion to introduce a new drug to the market[1,2]. In contrast, drug repositioning is a cost-effective strategy to extend the utility of existing drugs and offer shorter pathways to approval for new therapeutic indications, since the safety has already been confirmed in preclinical and the first two stages of clinical trials.

Leveraging genetic evidence of disease association helps to address this challenge since empirical data have shown that drug targets with genetic support are more likely to be successful[3,4]. Genome-wide association analysis (GWAS) and Transcriptome-wide association analysis (TWAS)[5]/ Proteome-wide association study (PWAS)[6] are powerful genetic tools to predict drug repurposing candidates. For example, the genome-wide significant association between *TYK2* and critical illness in COVID-19[7], directly led to the drug repurposing of baricitinib, which has demonstrated clear therapeutic benefit in subsequent clinical trials[8-10].

However, drug repositioning focused on a single gene target may be blind to molecular processes that are inevitably connected, through complex networks, with multiple levels of the organizational hierarchy of biological systems[11]. The growing evidence for polypharmacology entails the adoption of experimental and computational multitarget approaches[12,13]. Utilizing the transcriptome at the cellular level, a method that performs an inverse pattern-matching alignment of disease signatures and drug profiles[14,15] posits that, for a given disease signature comprising a set of genes with altered up and down-regulation, a drug that reverses the disease-associated gene expression changes holds therapeutic potential for the disease. Most existing approaches rely on differential gene expression (DGE) to generate a list of prioritzed genes. However, the transcriptome perturbations identified by DGE are challenging to interpret as molecular consequences or causes of disease, with one recent study claiming a higher likelihood of implicated genes being of the former class[16]. Consequently, methodologies grounded in DGE analysis inherently struggle to determine the direction of pharmacological effect (therapeutic or harmful) on disease when lacking the context of a known disease pathogenesis. In contrast, genetically-anchored TWAS minimizes

the problem of reverse causality. Patrick Wu et al. applied the PrediXcan[17] models to find drug repurposing candidates using the Integrative Library of Integrated Network-based Cellular Signatures (iLINCS) platform[18] . The latest release (CMAP LINCS 2020) of this platform contains over 3 million gene expression profiles with > 80,000 perturbations and > 200 cell lines. The resource is roughly a 3-fold expansion relative to the previous version (2017)[19]. Methodologically, previous approaches (for example, a widely adopted nonparametric pattern-matching strategy based on the Kolmogorov−Smirnov (KS) statistic[20,21]) do not fully consider the pattern of alignment between the (transcript)omic features of drug induction and disease susceptibility[14,22] within a multitarget or polygenic framework. In summary, a genetics-informed omics-based repurposing framework that addresses these methodological gaps is needed.

We developed Transcriptome-informed Reversal Distance (TReD) to integrate cell-based drug profiles (CMAP LINCS 2020) and population-level disease signatures (TWAS and DGE) to identify drug repurposing candidates. For illustration, we applied the framework to search for drug repurposing candidates for COVID-19.

Results

An overview of the study

Genetics-informed single-gene matching and omics profile-based signature matching are two promising approaches used in drug repurposing[23], with diverse types of omics data providing a valuable opportunity[24]. In this study, we integrated these two approaches, leveraging transcriptome data to identify potential drug repurposing candidates (Figure 1). A single gene-based approach that combines TWAS and colocalization analysis (coloc-SuSiE) detects genes associated with disease (Figure 1a, Figure S1). Subsequently, drugs targeting these genes are prioritized as candidate therapeutics. For the multi gene-based approach, we developed TReD, integrating TWAS- and DGE- derived disease signatures and cell-based transcriptome-wide drug profiles (CMAP LINCS 2020) (Figure 1b, Figure S1). TReD is based on transcriptome signature matching, which relies on the signature reversion principle[25]. Different from previous transcriptome signature pattern-matching studies[26], we calculated the

projection distance from the cosine similarity with the drug signature, incorporating the transcriptome profile of all featured genes (**Methods**). The reversal distance represents the drug's effect in the opposite direction of the disease-associated gene expression. To conduct hypothesis testing and evaluate the statistical significance, we employed a permutation approach. Drugs showing a large reversal distance and a significant p-value in at least two disease-related cell lines are considered as potential candidates. We applied this framework to identify drug repurposing candidates for COVID-19.

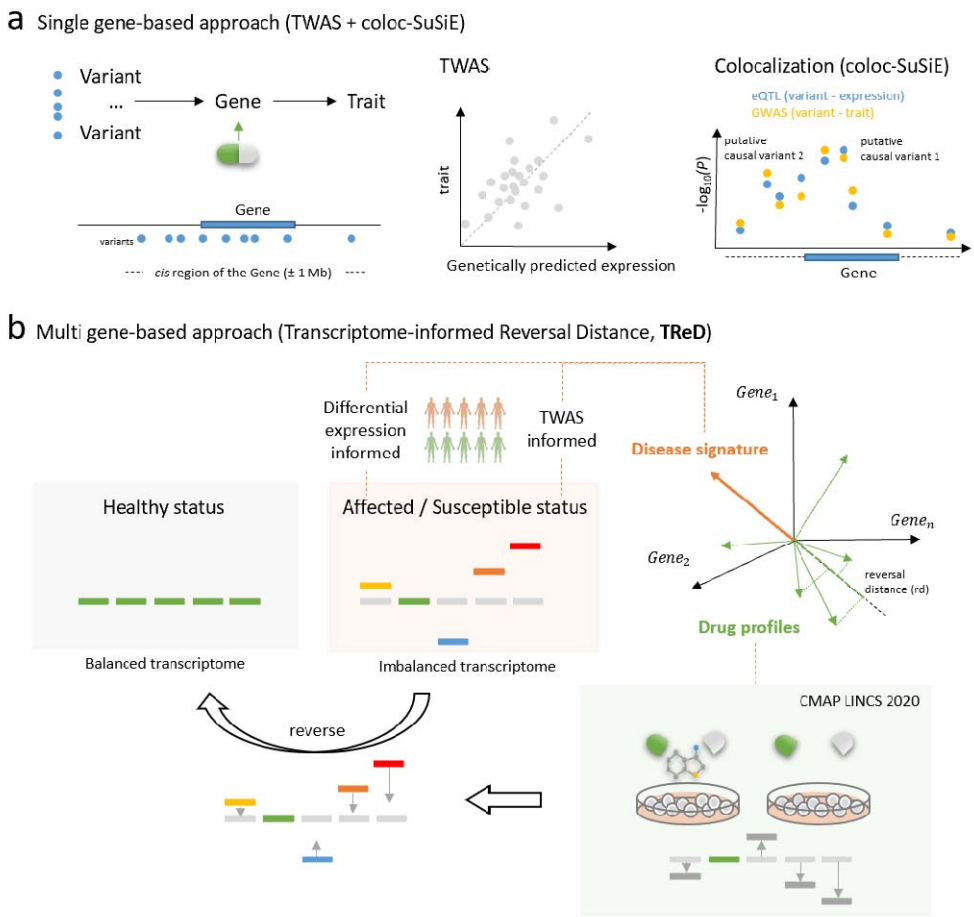
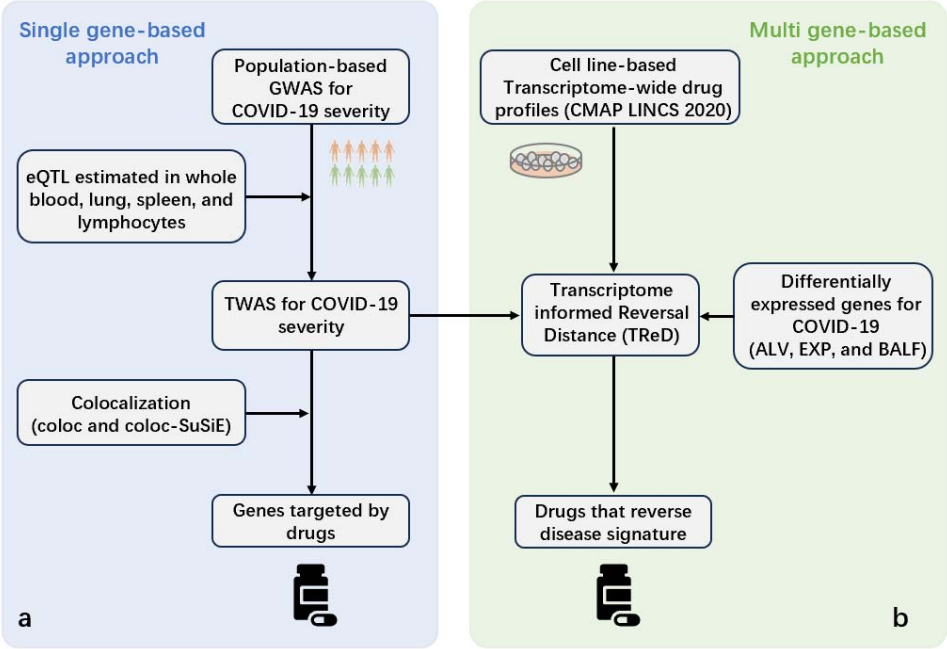


Figure 1. The framework overview. (a) Single gene-based approach integrates TWAS (e.g., PrediXcan) and colocalization analysis (e.g., coloc-SuSiE). (b) Multi-gene-based approach TReD (Transcriptome-informed Reversal Distance) utilizes DGE- and TWAS-derived population-level transcriptome profiles as disease signatures, reflecting affection status and germline susceptibility, respectively. By embedding the population-level disease signatures (orange vector) and the cell-based drug profiles (green vectors) in the same high-dimensional space, the reversal distance (*rd*) is estimated for each drug. A permutation test is used to estimate the statistical significance of the reversal distance.



Supplementary Figure 1. Application of TReD to identify drug repurposing candidates for COVID-19 (a) GWAS summary statistics of COVID-19 severity and eQTL data in four immune-related tissues were combined by PrediXcan and coloc/coloc-SuSiE, respectively. Genes that were significant in TWAS after Bonferroni correction and, in addition, exhibited positive colocalization (PP.H4 for coloc and coloc-SuSiE ≥ 0.85) were prioritized as potential drug targets. (b) Four TWAS results ($P_{FDR} < 0.05$), three DGE results for COVID-19 severity and cell-based transcriptome drug profiles from CMAP LINCS 2020 were integrated by TReD to obtain drugs that reverse the disease signatures.

Single-target drug repositioning using TWAS and coloc-SuSiE

We performed a TWAS to identify the associations between gene expression and COVID-19 severity. This was achieved by training gene expression prediction models using a reference panel and applying in the models to a GWAS of COVID-19 severity. For this analysis, we used data from Genotype-Tissue Expression (GTEx) v8, focusing on four tissues relevant to COVID-19, namely whole blood, lung, lymphocytes, and spleen samples. Genes that survived Bonferroni correction for each tissue were considered significant (Figure 2, Table S1).

We identified 34, 45, 27, and 43 significant genes in blood, lung, lymphocytes, and spleen, respectively. Nine genes showed significance across four tissues. Notably, *tyrosine kinase 2*

(*TYK2)* was significant in all tissues with a concordant positive effect, which means that an increased expression of *TYK2* is associated with an increased risk of severe COVID-19. *TYK2* is an intracellular kinase belonging to the Janus kinase (JAK) family, which plays a pivotal role in the Janus Kinase–Signal Transducer and Activator of Transcription (JAK-STAT) pathway. Previous studies have also revealed that high expression of *TYK2* and host-driven inflammatory lung injury are related to life-threatening COVID-19[7,27]. *RAB2A* also shows a positive effect in TWAS results of the four tissues. The gene is required for SARS-CoV-2 replication[28] and it has been reported that higher expression of this gene is associated with worse COVID-19 outcome[29], aligning with our findings.

For TWAS results that passed our significance threshold after accounting for multiple testing, we performed coloc/coloc-SuSiE, positing different assumptions for causal gene search. Both coloc and coloc-SuSiE were performed using their default parameters. We considered that the signals of eQTL and GWAS were colocalized if the maximum posterior probability of colocalization (i.e., PP.H4 for coloc or coloc-SuSiE) was greater than or equal to 0.85. The results of coloc can be found in Table S2 and coloc-SuSiE in Table S3. In the four analyzed tissues, we identified a total of 28 unique genes (Figure S2), among these genes, 50% (14/28) have been reported to be associated with COVID-19 (*DPP9*[30,31], *WNT3*[32,33], *ABO*[34,35], *NAPSA*,[36] *GSDMB*[37,38], *RAB2A*[29], *OAS1*[39], *OAS3*[7], *LRRC37A2*[40], *MAPT*[32], *FUT2*[41,42], *IKZF3*[43], *ICAM5*[41,44], *CCR9*[34,45]).

Next, we examined which of the prioritized disease-related genes are possible drug targets, thereby identifying potential therapeutic drugs for the disease. Drug target genes were defined using public databases, including DrugBank[46], The Drug Gene Interaction Database (DGIdb)[47] and the Therapeutic Target Database (TTD)[48]. *CCR9* and *MAPT* are druggable genes. Drugs such as Docetaxel and Astemizole, which target *MAPT,* had COVID-19-related support. *MAPT* encodes the microtubule-associated protein tau (MAPT). Tau protein binds to microtubules and helps in maintaining their stability. Docetaxel belongs to a class of drugs known as taxanes and is a chemotherapy medication used to treat various types of cancer. Interestingly, Docetaxel was noted to have the potential to bind the active site residues of SARS-CoV-2 RdRP, and exert antiviral effects by inhibiting the activity of SARS-CoV-2 RdRP[49]. Astemizole is an antihistamine medication that is used to relieve allergy symptoms.

Wang X, et al. found that astemizole can antagonize ACE2 and inhibit the entry of SARS-COV-2 spike pseudovirus into ACE2-expressed HEK293T cells (ACE2hi cells)[50]. This highlights astemizole as a potential drug candidate that can be repurposed in anti-coronavirus therapies.

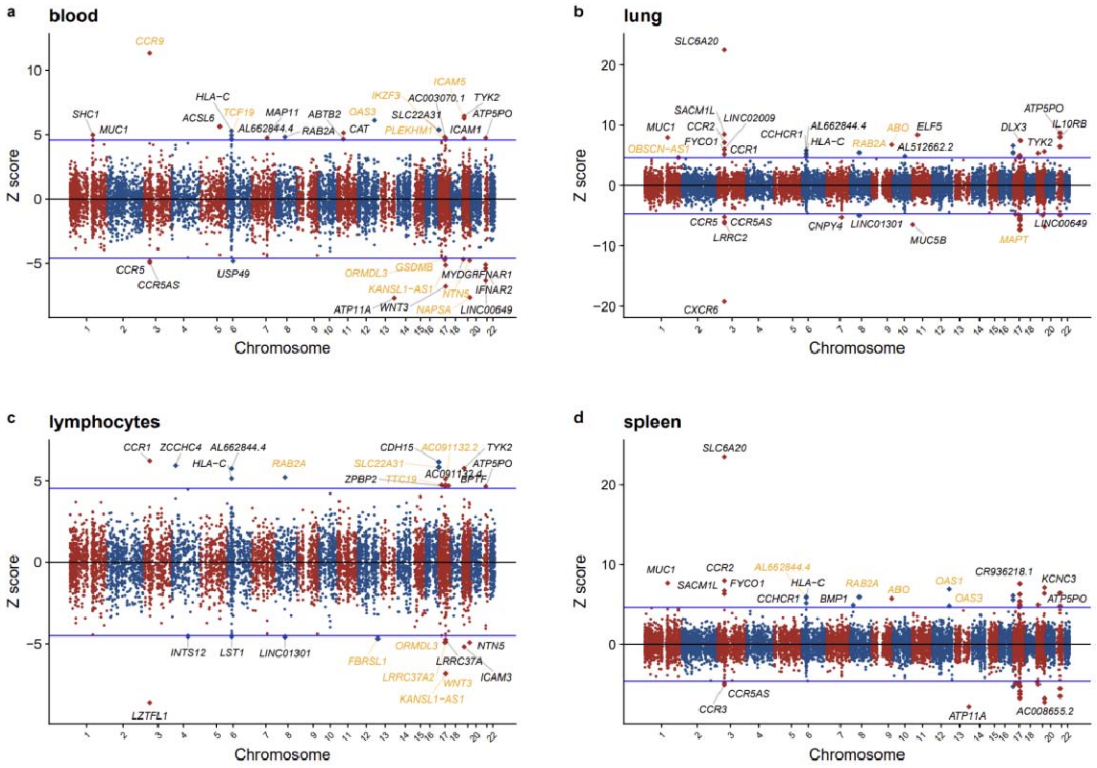


Figure 2. TWAS and colocalization results. Manhattan plots of TWAS results for COVID-19 severity in whole blood (a), lung (b), lymphocytes (c), and spleen (d), respectively. The Z-scores for TWAS are shown. Z-score > 0 denotes a gene for which increased predicted expression is associated with COVID-19 severity. The highlighted genes survived multiple testing corrections (over or below the blue lines) for the number of genes tested. Genes that further pass the significance threshold from the colocalization analysis are colored gold.

Multi-target drug repositioning using TReD

We applied the TReD framework to identify potential drug repurposing candidates with a potentially therapeutic benefit for COVID-19. This was achieved by identifying drugs inducing significant reversal effects of the cell-based transcriptome drug profiles on the population-level transcriptome disease signatures (**Methods**). In addition to the four TWAS-derived disease signatures, we included three DGE-based disease signatures. The

detailed results of TReD can be found in Table S4. A significant drug was defined by 2sd outlier and $P < 0.05$ in permutation test (**Methods**). For each disease signature (from either TWAS or DGE), drugs that met both the reversal distance threshold and the significance in at least two cell lines were considered as potential candidates. We obtained 263 drug profiles meeting the criteria for any of the seven disease signatures (Table S5). Here, we focused on the 37 drugs that showed significant reversal effects against at least four disease signatures (Figure 3). Table 1 provides an overview of the 37 drugs, including the reversal distance distribution across the different disease signatures and whether there is any prior evidence linking them to COVID-19 and/or clinical studies for the treatment of COVID-19. Notably, more than 70% of the drugs (27/37) have been reported to be related to COVID-19, and eight drugs have ongoing/completed clinical trials.

It is worth noting that over 40% of the candidates are immunomodulatory drugs, especially JAK inhibitors. Baricitinib, significantly reversing four disease signatures (Figure 4b), is a JAK1/2 and TYK2 inhibitor, and has been approved for the treatment of COVID-19 by FDA in 2022. Baricitinib, in a large clinical trial for hospitalized COVID-19, exhibited a clear therapeutic benefit[10]. Ruxolitinib significantly reversed all seven disease signatures (Figure 4a). As a JAK1 and JAK2 inhibitor, ruxolitinib has been used for around 10 years in the clinic for the treatment of myelofibrosis and polycythemia vera. Case reports or small cases series have reported that Ruxolitinib might be helpful in combating COVID-19[51-53]. The RUXCOVID[54] and phase 3 RUXCOVID-DEVENT[55] study showed no statistically significant improvement in mortality through day 29 compared with placebo. However, in the RUXCOVID-DEVENT study, when US study participants were analyzed separately, significant benefits were observed[55], indicating effectiveness.

Enzastaurin was found to be effective in reversing six disease signatures (Figure 4). Enzastaurin is an investigational agent for the treatment of relapsed glioblastoma multiforme (GBM). Enzastaurin was designed to selectively inhibit PKC-beta and other kinases involved in cell proliferation and survival. SARS-CoV can activate the PKC members[56], suggesting that these might be involved in SARS-CoV-2 infection. An *in vitro* study revealed that bisindolylmaleimide IX, a PKC inhibitor, targets the viral protease 3CLpro[57]. Another study found that several inhibitors of protein kinases C (PKCs) including enzastaurin decrease the

replication of SARS-CoV-2 in cultured cells[58]. The evidence above indicates that enzastaurin is a potential therapeutic agent against SARS-CoV-2.

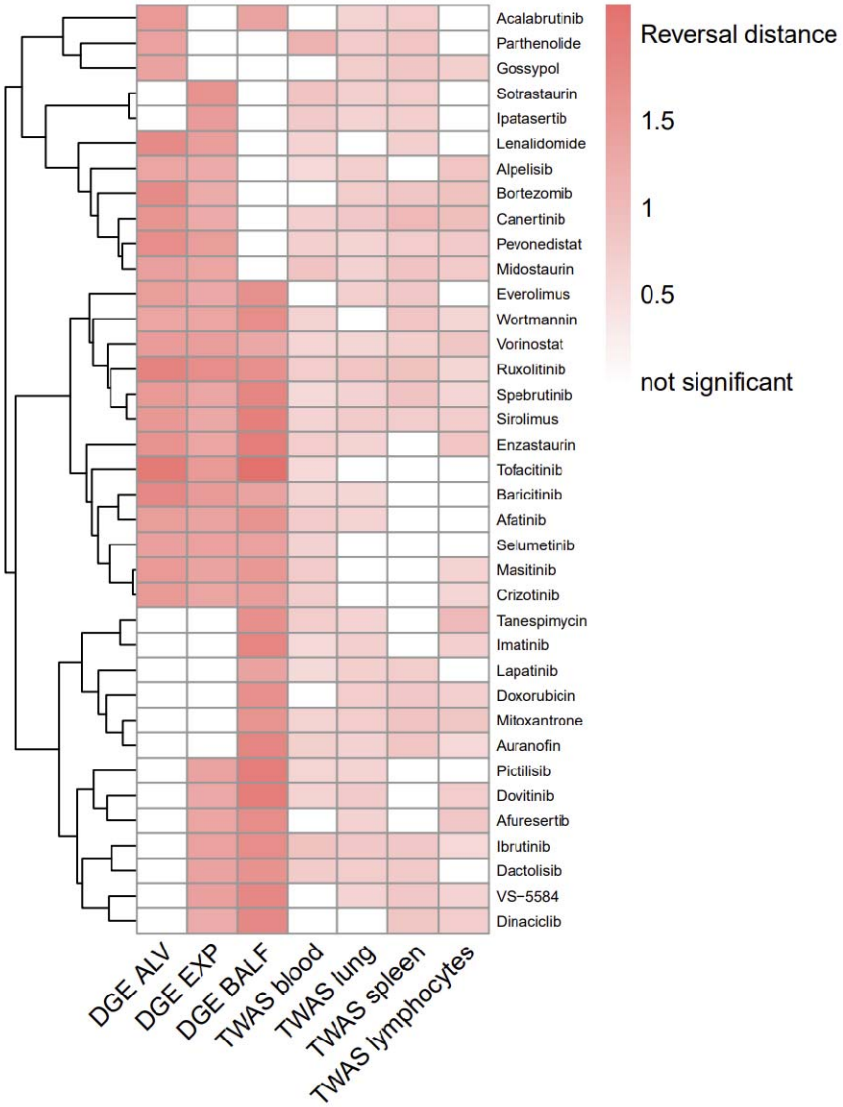


Figure 3. Drug repositioning candidates with potential reversal effect on the disease signatures from DGE and TWAS. In total, 37 drugs were found by our Transcriptome-informed Reversal Distance (TReD) framework (**Methods**) to have a potential reversal effect on COVID-19 as indicated by at least four of the seven tested transcriptome-based disease signatures (x-axis). The heatmap shows the reversal distance for each drug response profile / disease signature pair. Blank cells denote "not significant" in the permutation test. Hierarchical clustering was performed on the drug response profile (y-axis). The source data for the disease signatures, including TWAS (four tissue/cell types) and DGE (ALV, EXP, and BALF), are described in Methods.

Table 1. DrugBank annotations for drugs with significant reversal effects in at least four disease signatures for COVID-19.

Drug Hit	Description	Reversal distance for each of the disease signatures							Evidence for COVID-19	Evidence for clinical trials
		ALV	EXP	BALF	blood	lung	spleen	lymphocytes		
Wortmannin	PI3Ks inhibitor	1.350	1.347	1.699	0.653	n.s.	0.829	0.603	Yes[59]	No
VS-5584	Antineoplastic	n.s.	1.426	1.784	n.s.	0.644	0.794	0.640	Yes[60]	No
Vorinostat	Antineoplastic	1.465	1.458	1.305	0.632	0.597	0.700	0.813	Yes[61]	No
Tofacitinib	Antineoplastic and Immunomodulating	1.991	1.518	2.159	0.583	n.s.	n.s.	n.s.	Yes[62]	Yes
Tanespimycin	Antitumor antibiotics	n.s.	n.s.	1.665	0.740	0.653	n.s.	1.026	No	No
Spebrutinib	Immunomodulator	1.507	1.339	1.828	0.554	0.674	0.860	0.624	Yes[63]	No
Sotrastaurin	Immunosuppressant	n.s.	1.625	n.s.	0.855	0.720	0.727	n.s.	Yes[58]	No
Sirolimus	Immunosuppressant	1.531	1.299	1.939	0.653	0.743	0.731	0.733	Yes[64]	Yes
Selumetinib	Antineoplastic	1.426	1.406	1.411	0.667	n.s.	n.s.	n.s.	Yes[65]	No
Ruxolitinib	Antineoplastic and Immunomodulating	1.871	1.685	1.666	0.722	0.845	0.878	0.592	Yes[66]	Yes
Pictilisib	PI3K inhibitors	n.s.	1.377	1.944	0.615	0.636	n.s.	n.s.	No	No
Pevonedistat	Antineoplastic	1.686	1.450	n.s.	0.709	0.650	0.732	0.785	No	No
Parthenolide	Anti-Inflammatory	1.403	n.s.	n.s.	1.138	0.753	0.841	n.s.	Yes[67]	No
Mitoxantrone	Chemotherapy drugs	n.s.	n.s.	1.599	0.648	0.729	0.852	0.821	Yes[68]	No
Midostaurin	Antineoplastic	1.438	1.325	n.s.	0.863	0.672	0.853	0.748	No	No
Masitinib	Antineoplastic and Immunomodulating	1.511	1.372	1.529	0.750	n.s.	n.s.	0.639	Yes[69]	Yes
Lenalidomide	Antineoplastic and Immunomodulating	1.750	1.441	n.s.	0.646	n.s.	0.714	n.s.	Yes[70]	No
Lapatinib	Antineoplastic	n.s.	n.s.	1.357	0.541	0.709	0.733	n.s.	Yes[71]	No
Ipatasertib	Antineoplastic	n.s.	1.478	n.s.	0.768	0.654	0.708	n.s.	No	No
Imatinib	Antineoplastic	n.s.	n.s.	1.814	0.552	0.700	n.s.	0.719	Yes[72]	Yes
Ibrutinib	Antineoplastic	n.s.	1.413	1.685	0.876	0.793	0.789	0.581	Yes[63]	Yes
Gossypol	Polyphenolic compounds	1.374	n.s.	n.s.	n.s.	0.738	0.824	0.711	Yes[73]	No
Everolimus	Antineoplastic	1.457	1.288	1.648	n.s.	0.698	0.797	n.s.	Yes[74]	No
Enzastaurin	Antineoplastic	1.616	1.346	1.957	0.738	0.642	n.s.	0.840	Yes[58]	No
Doxorubicin	Chemotherapy drugs	n.s.	n.s.	1.660	n.s.	0.730	0.798	0.685	Yes[75]	No
Dovitinib	Multi-targeted kinase inhibitor	n.s.	1.304	1.964	0.652	0.782	n.s.	0.737	No	No
Dinaciclib	Antineoplastic	n.s.	1.236	1.800	n.s.	n.s.	0.816	0.738	Yes[76]	No
Dactolisib	Antineoplastic	n.s.	1.363	1.629	0.761	0.732	0.772	n.s.	Yes[77]	No
Crizotinib	Antineoplastic	1.513	1.318	1.459	0.720	n.s.	n.s.	0.625	Yes[78]	No
Canertinib	Antineoplastic	1.613	1.253	n.s.	0.687	0.801	1.029	0.957	No	No
Bortezomib	Antineoplastic	1.766	1.241	n.s.	n.s.	0.734	0.826	0.883	Yes[79]	No
Baricitinib	Immunosuppressant	1.767	1.493	1.393	0.633	0.594	n.s.	n.s.	Yes[10]	Yes
Auranofin	Antiinflammatory and Antirheumatic	n.s.	n.s.	1.815	0.715	0.664	0.844	0.588	Yes[80]	No
Alpelisib	Antineoplastic	1.319	1.249	n.s.	0.582	0.711	n.s.	0.840	No	No
Afuresertib	Akt inhibitor	n.s.	1.334	1.680	n.s.	0.669	n.s.	0.794	No	No
Afatinib	Antineoplastic	1.442	1.369	1.628	0.744	0.641	n.s.	n.s.	No	No

Acalabrutinib	Antineoplastic	1.526	n.s.	1.355	n.s.	0.650	0.722	n.s.	Yes[81]	Yes

The estimated reversal distance is shown. n.s. denotes not significant in the permutation test. The source data for the disease signatures, including TWAS (four tissue/cell types) and DGE (ALV, EXP, and BALF), are described in Methods.

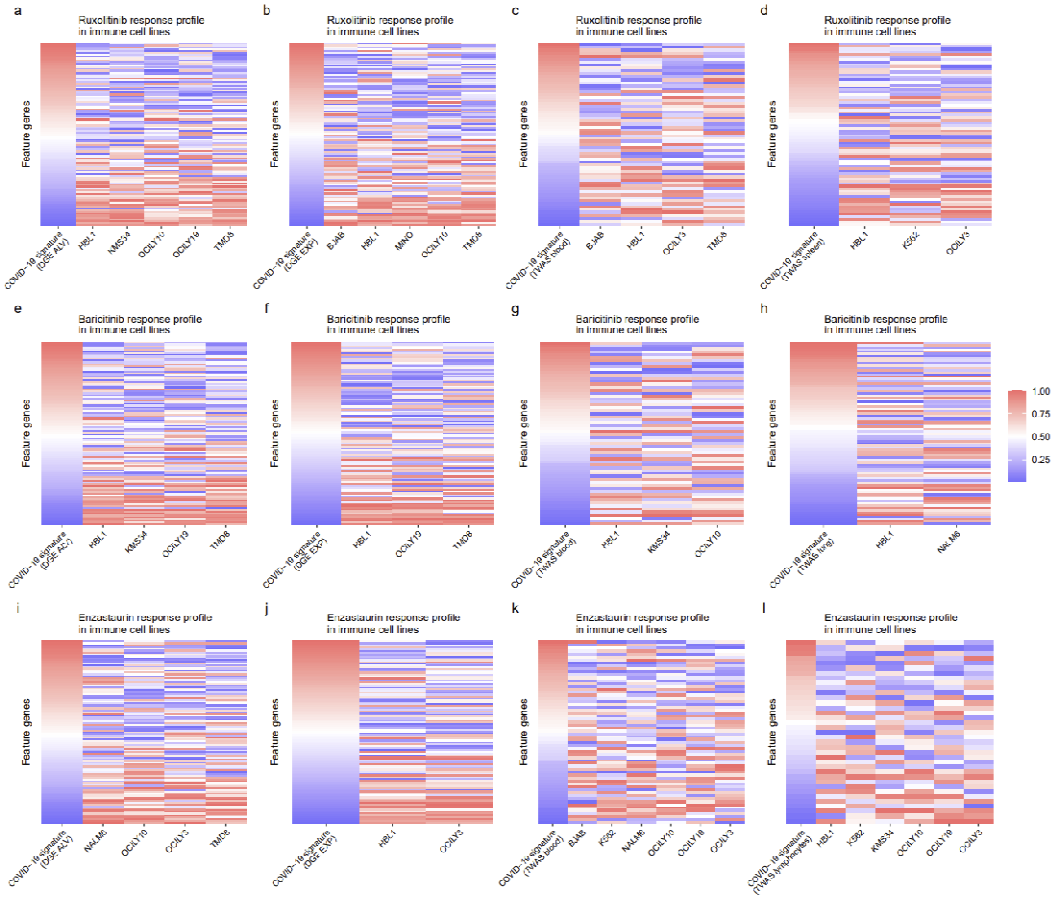


Figure 4. TReD-identified relationships between COVID-19 signatures and drug profiles. The potential reversal effects of Ruxolitinib (a-d), Baricitinib (e-h), and Enzastaurin (i-l), which were identified against at least four disease signatures of COVID-19, are visualized using a heatmap. Taking (a) as an example, the first column in the heatmap represents differential gene expression (DGE) from the ALV dataset gene expression disease signature ranked by fold change. Red and blue denote higher and lower expressed in affected samples, respectively. The remaining columns represent gene expression changes after Ruxolitinib treatment in immune cell lines. For drugs with multiple significant profiles, the drug profiles (e.g., different time points after treatment) with the largest reversal distance are presented. The source data for the disease signatures, including TWAS (four tissue/cell types) and DGE (ALV and EXP), are described in Methods.

Discussion

We developed a framework to identify potential drug repurposing candidates, incorporating population-level disease signatures using TWAS/DGE and cell-based drug-induced transcriptome responses. To demonstrate the effectiveness of this framework, we applied it to the up-to-date GWAS for COVID-19. In our quest to uncover the transcriptome signatures relevant to COVID-19 using TWAS, we identified the *TYK2* gene, which exhibited a clear and substantial association with COVID-19 severity.

Using the proposed multi-target approach TReD, nearly half (16/37) of the identified drugs were categorized as immunotherapeutic agents. Notably, all sixteen drugs have been studied in the context of COVID-19. SARS-CoV-2 infection can induce dysfunctional immune responses and aggressive inflammatory responses that are often associated with severe manifestations[82,83]. Therefore, hospitalized patients with severe or critical COVID-19 may benefit from treatment with immunomodulatory drugs. Cytokines play pivotal roles as key regulators of the immune response, acting through specific cytokine receptors. The human genome encodes over 40 cytokine receptors, all signaling *via* JAKs, that are appended non-covalently to their cytosolic tails. Four JAKs are coded by the human genome, namely *JAK1*, *JAK2*, *JAK3*, and *TYK2*. Activated receptors induce the activation of Signal Transducers and Activators of Transcription (STATs) *via* JAKs. This pathway is crucial for the intracellular signaling of numerous cellular processes, playing an essential role in both normal and pathological states of immune-mediated inflammatory diseases. Furthermore, JAK-STAT activation promotes senescence of SARS-CoV-2 infected cells, which amplifies inflammation[84]. JAK inhibition prevents downstream phosphorylation and hence activation of STAT. Therefore, JAK inhibitors have been suggested as potential therapies against systemic inflammation in COVID-19[85,86]. Notably, several meta-analysis studies have reported the safety and efficacy of JAK inhibitors, particularly baricitinib and ruxolitinib in COVID-19[87,88]. In May 2022, the FDA approved a new indication for baricitinib in hospitalized adults with COVID-19 who require oxygen support. Notably, baricitinib exhibited significance against four disease signatures, while ruxolitinib showed significance across all seven disease signatures in our study. Tofacitinib, also identified by TReD, is recommended as an

alternative option by the NIH in cases where baricitinib is not available[89]. Over half of the drugs in our results are substantiated by existing studies, supporting their relevance to COVID-19. Collectively, these results demonstrate the ability of the TReD framework to identify drug repositioning candidates.

The strengths of our framework stem from several key aspects. Firstly, the amalgamation of single-target and multiple-target approaches enables a more comprehensive therapeutic search, and this is particularly valuable for addressing complex diseases involving multiple biomolecules or mechanisms. Secondly, the framework incorporates two complementary methods (DGE and TWAS) as population-level sources for disease signatures. Although TWAS may be underpowered with a limited sample size, a critical feature of such a genetics-anchored method ensures that it identifies *potentially* regulatory or candidate causal genes for the disease, rather than genes altered by the disease. In contrast, DGE is prone to identify the transcriptomic consequences of the disease; nevertheless, it remains well-powered even with limited sample sizes. The combination of these two methods should enrich for upstream mediators and molecular consequences, facilitating the discovery of new drug indications or the repositioning of existing drugs. Thirdly, the use of the reversal distance measurement in our study capitalizes on available (and growing) information from transcriptome-level disease-associated gene features and drug profiles, with biologically relevant interpretation of an effective therapeutic effect on disease. This differs from previous and, we would argue, less interpretable methods, such as those that have relied solely on a non-parametric KS statistic, focusing on the maximum difference in the relative ranks of the upregulated and downregulated genes[90]. Fourthly, we leveraged the latest CMAP LINCS 2020 data which contains over three million gene expression profiles with > 80,000 perturbations and > 200 cell lines affecting over 12,328 genes, providing a comprehensive candidate pool for drug repositioning. While our study focused on COVID-19, our framework is highly generalizable and can be implemented to identify drug repositioning candidates for other diseases.

The limitations of the framework should also be acknowledged. Firstly, limited by data availability, only the transcriptome was considered. Comprehensive data from other omics disciplines for COVID-19 are currently lacking or, at best, very limited. Secondly, our study

does not consider the highly complex issues of drug-target interactions and drug-drug interactions. Thirdly, i*n vivo* and *in vitro* drug responses may be significantly different, and indeed the actual therapeutic efficacy is likely to be far more multifaceted than the simple alignment of transcriptome profiles. *In vivo* drug responses are influenced by complex factors like pharmacokinetics and pharmacodynamics, which may not be easily aligned with *in vitro* observations. Despite these limitations, it should be noted that the TReD *statistical* methodology is sufficiently flexible to integrate related methodological improvements, involving other types of omics data, more complex therapeutic profiles, and more realistic *in vivo* pharmacologic models.

In summary, integrating transcriptome-level profiles from cell-based response profiles and population-based disease signature studies, we developed a framework, TReD, to identify drug repurposing candidates. In practice, we successfully identified well-studied candidates and several other promising candidate drugs, such as Enzastaurin, for COVID-19 treatment and prevention. The framework demonstrates effectiveness in drug repositioning and is highly applicable across various diseases.

Materials and Methods

GWAS for COVID-19 severity

The GWAS summary statistics for COVID-19 severity were downloaded from COVID-19 Host Genetics Initiative (https://www.covid19hg.org/), the latest version, release 7 of the hospitalized versus population data. The hospitalized phenotype includes patients who were hospitalized with a laboratory-confirmed SARS-Cov2 infection. This dataset contains 44,989 cases and 2,356,386 controls. Additional descriptions of the included studies can be found in Table S6.

TWAS and colocalization for COVID-19 severity

To identify COVID-19 associated genes, we applied TWAS with the best prediction models among PrediXcan[91], UTMOST[92], and JTI[93] according to the prediction accuracy estimated by 5-fold cross-validation. Briefly, the residuals of the normalized gene expression levels were used after regressing out covariates. SNPs within 1 Mb upstream and downstream

of the gene body were considered as predictor variables in the models. TWAS summary statistics-based association test was performed. Blood, lung, lymphocytes, and spleen were considered as relevant tissues for COVID-19. For TWAS results that passed Bonferroni multiple testing correction, we performed colocalization to provide additional support under different sets of assumptions. To test colocalization in the presence of multiple causal variants, we applied coloc[94] to the signals decomposed by SuSiE[95] for each locus, including the variants located within ± 200 kb of the lead variant (*coloc-SuSiE*). Notably, *coloc-SuSiE* is more generalizable since it relaxes the "one causal variant" assumption of regular *coloc*. We used European participants in the 1000 genomes project as the LD reference. If SuSiE did not converge after 1000 iterations for either eQTL or GWAS summary statistics, we instead used the regular *coloc*. We considered that the signals of eQTL and GWAS were colocalized if the maximum posterior probability of colocalization (i.e., PP.H4 for *coloc* and *coloc-SuSiE*) was equal or larger than 0.85. TWAS significant genes showing positive colocalization evidence were then used to search for drugs targeting the genes, via DrugBank, TTD, DGIdb. DrugBank is a comprehensive and up-to-date resource for searching over 500,000 drugs and drug products, their targets, pathways, indications, and pharmacology. TTD is a comprehensive database of therapeutic targets, drugs, biomarkers and scaffolds for various diseases and ICD identifiers. DGIdb is a web resource that provides information on drug-gene interactions and druggable genes from publications, databases, and other web-based sources.

DGE for COVID19

The DGE data (three independent gene expression signatures labeled 'ALV', 'EXP', 'BALF') are directly derived from the study of Brian L. Le et al[26]. The ALV signature was estimated in human adenocarcinomic alveolar basal epithelial cells by comparing SARS-CoV-2 infection to mock-infection conditions[96] (Table S7). The EXP signature was obtained by comparing the difference in gene expression profile between organoid samples infected with SARS-CoV-2 and controls[97] (Table S8). The BALF signature was estimated by comparing the transcriptome between the bronchoalveolar lavage fluid samples from two COVID-19 patients (two samples per patient) and three healthy controls (Table S9).

The CMAP LINCS 2020 dataset

The Library of Integrated Network-Based Cellular Signatures (LINCS) is a program

funded by the NIH Common Fund. Its primary goal is to generate a large-scale and comprehensive catalog of perturbation-response signatures by utilizing a diverse collection of perturbations across many model systems and assay types. The CMAP LINCS 2020 data contains ~3 million gene expression profiles of instances which are defined as changes in the assay readout of a perturbed sample when compared to a control sample involving the perturbation reagents (e.g., small molecules, sgRNAs, shRNAs, antibodies and disease states) and model system reagents (cell lines, iPSCs, differentiated cells, and primary cells). Here, we mainly focused on the transcriptomic response in immune-related cell lines due to small-molecule perturbations in the application of the TReD methodology to COVID-19. We downloaded the level 5 beta data consisting of gene expression profiles due to small-molecule perturbations from CMAP LINCS 2020 (https://s3.amazonaws.com/macchiato.clue.io/builds/LINCS2020/level5/level5_beta_trt_cp_n720216x12328.gctx),

Evaluation of the reversal potential of candidate drugs

To quantify the extent of reversal on disease of a specific drug, we used the cosine similarity metric. For a given disease, we first represent the set of disease signatures as a vector in a high-dimensional space (whose dimension is the total number of genes) with each nonzero element of the vector given by a disease-associated gene. The disease signatures can be derived from either TWAS or DGE. The vector for each candidate drug is embedded in the same space. For each disease signature vector and drug response pair, we compute the reversal distance. Let $\overrightarrow{G}$ denote a disease signature vector, $\overrightarrow{D}$ a drug profile vector, and rd the reversal distance. The rd quantifies the potential effect of the drug in reversing the disease signatures:

$$\omega := \cos\theta = \frac{\overrightarrow{G} \,.\, \overrightarrow{D}}{\left|\left|\overrightarrow{G}\right|\right| \left|\left|\overrightarrow{D}\right|\right|} \qquad (1)$$

$$rd = \left|\left|\overrightarrow{G}\right|\right| \cos(\pi - \theta) \qquad (2)$$

Here, $||\vec{a}|| = \sqrt{\vec{a} \,.\, \vec{a}}$ denotes the Euclidean norm of the vector $\vec{a}$. Note that the cosine similarity ω (equation 1) always lies in $[-1,1]$, with -1 indicating exactly opposite orientation of the disease signature and the drug response profile and +1 indicating parallel

orientation. The change in ω in response to a perturbation in the i-th element of the drug response profile $\vec{D}$, d_i, or the i-th element of the gene signature $\vec{G}$, g_i, is given by:

$$\frac{\partial \omega}{\partial d_i} = \frac{g_i}{\left\|\vec{G}\right\|\left\|\vec{D}\right\|} - \frac{d_i(\vec{G}\,.\,\vec{D})}{\left\|\vec{G}\right\|\left\|\vec{D}\right\|^3}$$

$$\frac{\partial \omega}{\partial g_i} = \frac{d_i}{\left\|\vec{G}\right\|\left\|\vec{D}\right\|} - \frac{g_i(\vec{G}\,.\,\vec{D})}{\left\|\vec{D}\right\|\left\|\vec{G}\right\|^3}$$

Each quantity measures the sensitivity of the metric. In equation 2, θ, and thus the reversal distance rd, is a function of $\vec{G}$ and $\vec{D}$. When we wish to emphasize this dependence, we will write rd as $rd(\vec{G},\vec{D})$. Equation 2 defines a function:

$$d\left(\vec{G},\vec{D}\right) = \frac{2\arccos\left(rd\left(\vec{G},\vec{D}\right)/\left\|\vec{G}\right\|\right)}{\pi},$$

which satisfies the conditions for a distance function, namely, non-negativity: $d\left(\vec{a},\vec{b}\right) \geq 0$; symmetry: $d\left(\vec{a},\vec{b}\right) = d\left(\vec{b},\vec{a}\right)$; the triangle inequality; and the coincidence axiom: $d\left(\vec{a},\vec{b}\right) = 0$ if and only if $\vec{a} = \vec{b}$. Despite the symmetry of the distance function, there is clearly an underlying asymmetry, as the disease signature vector $\vec{G}$ determines which genes are the "relevant" elements of the drug response profile vector $\vec{D}$.

In this framework, drug repurposing is a search problem: the vector $\vec{G}$ is evaluated against all possible $\vec{D}$, which are prioritized using rd. A larger rd indicates a stronger reversal effect. Assuming K total number of drugs tested and the reversal distance for the i-th drug, $rd(\vec{G},\vec{D_i})$, we get a probability distribution consisting of the K probabilities p_i conditional on the disease signature information:

$$p_i = P\left(\vec{D_i}\,\middle|\,\vec{G}\right) = \frac{\exp\left(rd(\vec{G},\vec{D_i})\right)}{\sum_{j=1}^{K}\exp\left(rd(\vec{G},\vec{D_j})\right)}$$

by applying the softmax function. One drug repurposing approach is to identify the drug that maximizes the reversal distance:

$$\mathrm{D}^* = \underset{\vec{D}}{\operatorname{argmax}}\left(\ln\left(\left\|\vec{G}\right\|\right) + \ln\left(\cos\left(\pi - \theta(\vec{G},\vec{D})\right)\right)\right)$$

Here, we instead prioritized a potentially larger set of drugs, i.e., the top candidates that meet a significance threshold.

Of interest to us is the sampling distribution of the cosine similarity ω. Consider the function $h(s,t,u) = \frac{s}{\sqrt{tu}}$, where s, t, and u are asymptotically normal distributed. Note that ω is equal to $h\left(\vec{G}\,.\,\vec{D},\ \vec{G}\,.\,\vec{G},\ \vec{D}\,.\,\vec{D}\right)$. We leverage the so-called Delta Method. Using the first two terms of the Taylor series expansion and setting $\Gamma = (s,t,u)$ and its expected value $\gamma = \mathrm{E}[\Gamma]$, we obtain approximate expressions for the mean and variance of $h(\Gamma)$:

$$\mathrm{E}[h(\Gamma)] \approx h(\gamma) + \nabla h(\gamma)^T\,.\,\mathrm{E}[\Gamma - \gamma]$$

$$\mathrm{Var}\big(h(\Gamma)\big) \approx \nabla h(\gamma)^T\,.\,\mathrm{Cov}(\Gamma)\,.\,\nabla h(\gamma)$$

In our implementation, the gene-disease association signals are converted into rank ratios, which, for each gene, is defined as the rank of the gene divided by the total number of genes. Similarly, each element of the drug response profile is given by the rank-transformed cell-based drug-induced gene expression response.

We performed a test of significance of each candidate drug. For each drug profile, a p-value for the reversal distance is estimated by permutation. For each permutation, the rank ratios are randomly shuffled, and the corresponding null reversal distance is computed. This process constructs an empirical null distribution of reversal distance for each drug. Subsequently, for each drug, we calculated the probability that the empirical reversal distance could have originated from the null distribution of the reversal distance. This involved counting the number of null reversal distance values greater than the real reversal distance and dividing this count by the total number of permutations to generate an empirical p-value. Drugs showing significance in the permutation test ($P < 0.05$) and large rd (over 2 standard deviations from the mean across the tested drugs) are defined as strong candidates.

Application of TReD to COVID-19 severity

Leveraging the TWAS/DGE results and the cell-based gene expression data from CMAP LINCS 2020, we implemented the TReD framework to identify drug repositioning candidates for COVID-19. Immune-related cell lines have been annotated by the Cancer Cell Line Encyclopedia (CCLE) (https://sites.broadinstitute.org/ccle/). Information on drugs and drug targets is curated in the bioinformatics and chemoinformatics resource DrugBank (DrugBank Release Version 5.1.10 | DrugBank Online). In this study, 2,413 drugs and 12 immune-related cell types were included.

We leveraged the TWAS results of four tissues (blood, lung, lymphocytes, and spleen) and three DGE profiles as disease signatures (as defined in the previous section). Drug candidates were identified by permutation (1000 times sampling) and whether the reversal distance was greater than mean + 2 standard deviation. To ensure the robustness of the reversal relationship, we considered only those drugs whose reversal potential could be verified in additional cell line(s).

Data availability

Links for databases and datasets used in this study: COVID-19 severity GWAS summary statistics: COVID19-hg GWAS meta-analyses round 7 (covid19hg.org); CMAP LINCS 2020: https://s3.amazonaws.com/macchiato.clue.io/builds/LINCS2020/level5/level5_beta_trt_cp_n720216x12328.gctx; CCLE: Cancer Cell Line Encyclopedia (CCLE) (broadinstitute.org); DrugBank Release Version 5.1.10: DrugBank Release Version 5.1.10 | DrugBank Online;

Code availability

The used scripts are available at https://github.com/zdangm/TReD.

Funding

This research is supported by the Fundamental Research Funds from the Central Universities of Zhejiang University (D.Z.), the National Natural Sciences Foundation of China 82204118 (D.Z.), the National Institutes of Health (NIH) Genomic Innovator Award R35HG010718 (E.R.G.), NIH/NHGRI R01HG011138 (E.R.G.), NIH/NIA R56AG068026 (E.R.G.), and NIH/NIGMS R01GM140287 (E.R.G.).

References

1. Wouters, O.J., McKee, M. & Luyten, J. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. *Jama* **323**, 844-853 (2020).

2. DiMasi, J.A., Hansen, R.W. & Grabowski, H.G. The price of innovation: new estimates of drug development costs. *J Health Econ* **22**, 151-85 (2003).

3. King, E.A., Davis, J.W. & Degner, J.F. Are drug targets with genetic support twice as likely to be approved? Revised estimates of the impact of genetic support for drug mechanisms on the probability of drug approval. *PLoS Genet* **15**, e1008489 (2019).

4. Nelson, M.R. *et al.* The support of human genetic evidence for approved drug indications. *Nat Genet* **47**, 856-60 (2015).

5. Gamazon, E.R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nat Genet* **47**, 1091-8 (2015).

6. Zheng, J. *et al.* Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet* **52**, 1122-1131 (2020).

7. Pairo-Castineira, E. *et al.* Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92-98 (2021).

8. Marconi, V.C. *et al.* Efficacy and safety of baricitinib for the treatment of hospitalised adults with COVID-19 (COV-BARRIER): a randomised, double-blind, parallel-group, placebo-controlled phase 3 trial. *Lancet Respir Med* **9**, 1407-1418 (2021).

9. Kalil, A.C. *et al.* Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19. *N Engl J Med* **384**, 795-807 (2021).

10. Baricitinib in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial and updated meta-analysis. *Lancet* **400**, 359-368 (2022).

11. Houghten, R.A. *et al.* Mixture-based synthetic combinatorial libraries. *J Med Chem* **42**,

3743-78 (1999).

12. Paolini, G.V., Shapland, R.H., van Hoorn, W.P., Mason, J.S. & Hopkins, A.L. Global mapping of pharmacological space. *Nat Biotechnol* **24**, 805-15 (2006).

13. Hopkins, A.L. Network pharmacology: the next paradigm in drug discovery. *Nat Chem Biol* **4**, 682-90 (2008).

14. Chen, B. *et al.* Computational Discovery of Niclosamide Ethanolamine, a Repurposed Drug Candidate That Reduces Growth of Hepatocellular Carcinoma Cells In Vitro and in Mice by Inhibiting Cell Division Cycle 37 Signaling. *Gastroenterology* **152**, 2022-2036 (2017).

15. Chen, B. *et al.* Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets. *Nat Commun* **8**, 16022 (2017).

16. Porcu, E. *et al.* Differentially expressed genes reflect disease-induced rather than disease-causing changes in the transcriptome. *Nat Commun* **12**, 5647 (2021).

17. Barbeira, A.N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* **9**, 1825 (2018).

18. Pilarczyk, M., Najafabadi, M.F., Kouril, M., Vasiliauskas, J. & Medvedovic, M. Connecting omics signatures of diseases, drugs, and mechanisms of actions with iLINCS. (2019).

19. Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* **171**, 1437-1452.e17 (2017).

20. Lamb, J. *et al.* The Connectivity Map: using gene-expression signatures to connect

small molecules, genes, and disease. *Science* **313**, 1929-35 (2006).

21. Hu, R.Y. *et al.* Individualized Drug Repositioning For Rheumatoid Arthritis Using Weighted Kolmogorov-Smirnov Algorithm. *Pharmgenomics Pers Med* **12**, 369-375 (2019).

22. Le, B.L., Iwatani, S., Wong, R.J., Stevenson, D.K. & Sirota, M. Computational discovery of therapeutic candidates for preventing preterm birth. *JCI Insight* **5**(2020).

23. Pushpakom, S. *et al.* Drug repurposing: progress, challenges and recommendations. *Nat Rev Drug Discov* **18**, 41-58 (2019).

24. Reay, W.R. & Cairns, M.J. Advancing the use of genome-wide association studies for drug repurposing. *Nat Rev Genet* **22**, 658-671 (2021).

25. Wagner, A. *et al.* Drugs that reverse disease transcriptomic signatures are more effective in a mouse model of dyslipidemia. *Mol Syst Biol* **11**, 791 (2015).

26. Le, B.L. *et al.* Transcriptomics-based drug repositioning pipeline identifies therapeutic candidates for COVID-19. *Sci Rep* **11**, 12310 (2021).

27. Horby, P. *et al.* Dexamethasone in Hospitalized Patients with Covid-19. *N Engl J Med* **384**, 693-704 (2021).

28. Li, B. *et al.* Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nat Commun* **11**, 164 (2020).

29. Pairo-Castineira, E. *et al.* GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. *Nature* **617**, 764-768 (2023).

30. Muhammad, J.S., Siddiqui, R. & Khan, N.A. COVID-19: Is There a Link between Alcohol Abuse and SARS-CoV-2-Induced Severe Neurological Manifestations? *ACS*

Pharmacol Transl Sci **4**, 1024-1025 (2021).

31. Wang, L. *et al.* An atlas connecting shared genetic architecture of human diseases and molecular phenotypes provides insight into COVID-19 susceptibility. *Genome Med* **13**, 83 (2021).

32. Cao, H., Baranova, A., Song, Y., Chen, J.H. & Zhang, F. Causal associations and genetic overlap between COVID-19 and intelligence. *Qjm* (2023).

33. Wu, L., Zhu, J., Liu, D., Sun, Y. & Wu, C. An integrative multiomics analysis identifies putative causal genes for COVID-19 severity. *Genet Med* **23**, 2076-2086 (2021).

34. Ellinghaus, D. *et al.* Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N Engl J Med* **383**, 1522-1534 (2020).

35. Ray, J.G., Schull, M.J., Vermeulen, M.J. & Park, A.L. Association Between ABO and Rh Blood Groups and SARS-CoV-2 Infection or Severe COVID-19 Illness : A Population-Based Cohort Study. *Ann Intern Med* **174**, 308-315 (2021).

36. Krishnamoorthy, S., Li, G.H. & Cheung, C.L. Transcriptome-wide summary data-based Mendelian randomization analysis reveals 38 novel genes associated with severe COVID-19. *J Med Virol* **95**, e28162 (2023).

37. Cui, H. & Zhang, L. Key Components of Inflammasome and Pyroptosis Pathways Are Deficient in Canines and Felines, Possibly Affecting Their Response to SARS-CoV-2 Infection. *Front Immunol* **11**, 592622 (2020).

38. Chang, J.J. *et al.* Long-Read RNA Sequencing Identifies Polyadenylation Elongation and Differential Transcript Usage of Host Transcripts During SARS-CoV-2 In Vitro Infection. *Front Immunol* **13**, 832223 (2022).

39. Banday, A.R. *et al.* Genetic regulation of OAS1 nonsense-mediated decay underlies association with COVID-19 hospitalization in patients of European and African ancestries. *Nat Genet* **54**, 1103-1116 (2022).

40. Zhu, Z. *et al.* An integrated strategy to identify COVID-19 causal genes and characteristics represented by LRRC37A2. *J Med Virol* **95**, e28585 (2023).

41. Kousathanas, A. *et al.* Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* **607**, 97-103 (2022).

42. Andreakos, E. *et al.* A global effort to dissect the human genetic basis of resistance to SARS-CoV-2 infection. *Nat Immunol* **23**, 159-164 (2022).

43. Yao, M., Huang, X., Guo, Y., Zhao, J.V. & Liu, Z. Disentangling the common genetic architecture and causality of rheumatoid arthritis and systemic lupus erythematosus with COVID-19 outcomes: Genome-wide cross trait analysis and bidirectional Mendelian randomization study. *J Med Virol* **95**, e28570 (2023).

44. Zheng, J. *et al.* Multi-ancestry Mendelian randomization of omics traits revealing drug targets of COVID-19 severity. *EBioMedicine* **81**, 104112 (2022).

45. Yao, Y. *et al.* Genome and epigenome editing identify CCR9 and SLC6A20 as target genes at the 3p21.31 locus associated with severe COVID-19. *Signal Transduct Target Ther* **6**, 85 (2021).

46. Wishart, D.S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res* **46**, D1074-d1082 (2018).

47. Freshour, S.L. *et al.* Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res* **49**, D1144-d1151 (2021).

48. Zhou, Y. *et al.* Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. *Nucleic Acids Res* **50**, D1398-d1407 (2022).

49. Bagabir, S.A. Investigating the potential of natural compounds as novel inhibitors of SARS-CoV-2 RdRP using computational approaches. *Biotechnol Genet Eng Rev*, 1-21 (2023).

50. Wang, X. *et al.* Astemizole as a drug to inhibit the effect of SARS-COV-2 in vitro. *Microb Pathog* **156**, 104929 (2021).

51. D'Alessio, A. *et al.* Low-dose ruxolitinib plus steroid in severe SARS-CoV-2 pneumonia. *Leukemia* **35**, 635-638 (2021).

52. Koschmieder, S. *et al.* Favorable COVID-19 course despite significant comorbidities in a ruxolitinib-treated patient with primary myelofibrosis. *Eur J Haematol* **105**, 655-658 (2020).

53. Vannucchi, A.M. *et al.* Compassionate use of JAK1/2 inhibitor ruxolitinib for severe COVID-19: a prospective observational study. *Leukemia* **35**, 1121-1133 (2021).

54. Han, M.K. *et al.* Ruxolitinib in addition to standard of care for the treatment of patients admitted to hospital with COVID-19 (RUXCOVID): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Rheumatol* **4**, e351-e361 (2022).

55. Rein, L. *et al.* Randomized Phase 3 Trial of Ruxolitinib for COVID-19-Associated Acute Respiratory Distress Syndrome. *Crit Care Med* **50**, 1701-1713 (2022).

56. Ji, H.L. *et al.* SARS-CoV proteins decrease levels and activity of human ENaC via activation of distinct PKC isoforms. *Am J Physiol Lung Cell Mol Physiol* **296**, L372-83

(2009).

57. Gupta, Y. *et al.* Bisindolylmaleimide IX: A novel anti-SARS-CoV2 agent targeting viral main protease 3CLpro demonstrated by virtual screening pipeline and in-vitro validation assays. *Methods* **195**, 57-71 (2021).

58. Huang, C. *et al.* Protein Kinase C Inhibitors Reduce SARS-CoV-2 Replication in Cultured Cells. *Microbiol Spectr* **10**, e0105622 (2022).

59. Jiang, Y. *et al.* Insights into potential mechanisms of asthma patients with COVID-19: A study based on the gene expression profiling of bronchoalveolar lavage fluid. *Comput Biol Med* **146**, 105601 (2022).

60. Bharadwaj, S. *et al.* SARS-CoV-2 M(pro) inhibitors: identification of anti-SARS-CoV-2 M(pro) compounds from FDA approved drugs. *J Biomol Struct Dyn* **40**, 2769-2784 (2022).

61. Sinha, S. *et al.* In vitro and in vivo identification of clinically approved drugs that modify ACE2 expression. *Mol Syst Biol* **16**, e9628 (2020).

62. Guimarães, P.O. *et al.* Tofacitinib in Patients Hospitalized with Covid-19 Pneumonia. *N Engl J Med* **385**, 406-415 (2021).

63. Kaliamurthi, S. *et al.* Structure-Based Virtual Screening Reveals Ibrutinib and Zanubrutinib as Potential Repurposed Drugs against COVID-19. *Int J Mol Sci* **22**(2021).

64. Bischof, E., Siow, R.C., Zhavoronkov, A. & Kaeberlein, M. The potential of rapalogs to enhance resilience against SARS-CoV-2 infection and reduce the severity of COVID-19. *Lancet Healthy Longev* **2**, e105-e111 (2021).

65. Xie, J. *et al.* Inhibition of MEK signaling prevents SARS-CoV2-induced lung damage and improves the survival of infected mice. *J Med Virol* **94**, 6097-6102 (2022).

66. Lamontagne, F. *et al.* A living WHO guideline on drugs for covid-19. *Bmj* **370**, m3379 (2020).

67. Bahrami, M., Kamalinejad, M., Latifi, S.A., Seif, F. & Dadmehr, M. Cytokine storm in COVID-19 and parthenolide: Preclinical evidence. *Phytother Res* **34**, 2429-2430 (2020).

68. Preet, G., Oluwabusola, E.T., Milne, B.F., Ebel, R. & Jaspars, M. Computational Repurposing of Mitoxantrone-Related Structures against Monkeypox Virus: A Molecular Docking and 3D Pharmacophore Study. *Int J Mol Sci* **23**(2022).

69. Drayman, N. *et al.* Masitinib is a broad coronavirus 3CL inhibitor that blocks replication of SARS-CoV-2. *Science* **373**, 931-936 (2021).

70. von Metzler, I. *et al.* COVID-19 in multiple-myeloma patients: cellular and humoral immunity against SARS-CoV-2 in a short- and long-term view. *J Mol Med (Berl)* **100**, 463-470 (2022).

71. Saul, S. *et al.* Anticancer pan-ErbB inhibitors reduce inflammation and tissue injury and exert broad-spectrum antiviral effects. *J Clin Invest* (2023).

72. van de Veerdonk, F.L. *et al.* A guide to immunotherapy for COVID-19. *Nat Med* **28**, 39-50 (2022).

73. Wang, W. *et al.* Gossypol Broadly Inhibits Coronaviruses by Targeting RNA-Dependent RNA Polymerases. *Adv Sci (Weinh)* **9**, e2203499 (2022).

74. Colmenero, J. *et al.* Epidemiological pattern, incidence, and outcomes of COVID-19 in

liver transplant patients. *J Hepatol* **74**, 148-155 (2021).

75. Sajid Jamal, Q.M., Alharbi, A.H. & Ahmad, V. Identification of doxorubicin as a potential therapeutic against SARS-CoV-2 (COVID-19) protease: a molecular docking and dynamics simulation studies. *J Biomol Struct Dyn* **40**, 7960-7974 (2022).

76. Ahmad, S. *et al.* Reporting dinaciclib and theodrenaline as a multitargeted inhibitor against SARS-CoV-2: an in-silico study. *J Biomol Struct Dyn* **41**, 4013-4023 (2023).

77. Tomazou, M. *et al.* Multi-omics data integration and network-based analysis drives a multiplex drug repurposing approach to a shortlist of candidate drugs against COVID-19. *Brief Bioinform* **22**(2021).

78. Chen, T., Philip, M., KA, L.C. & Tyagi, S. A multi-modal data harmonisation approach for discovery of COVID-19 drug targets. *Brief Bioinform* **22**(2021).

79. Li, T.W. *et al.* SARS-CoV-2 Nsp14 protein associates with IMPDH2 and activates NF-κB signaling. *Front Immunol* **13**, 1007089 (2022).

80. Biji, A. *et al.* Identification of COVID-19 prognostic markers and therapeutic targets through meta-analysis and validation of Omics data from nasopharyngeal samples. *EBioMedicine* **70**, 103525 (2021).

81. Repurposing Cancer Drugs for COVID-19. *Cancer Discov* **10**, Of2 (2020).

82. Tay, M.Z., Poh, C.M., Rénia, L., MacAry, P.A. & Ng, L.F.P. The trinity of COVID-19: immunity, inflammation and intervention. *Nat Rev Immunol* **20**, 363-374 (2020).

83. Merad, M., Blish, C.A., Sallusto, F. & Iwasaki, A. The immunology and immunopathology of COVID-19. *Science* **375**, 1122-1127 (2022).

84. Lee, S. *et al.* Virus-induced senescence is a driver and therapeutic target in COVID-19.

Nature **599**, 283-289 (2021).

85. Heidel, F. & Hochhaus, A. Holding CoVID in check through JAK? The MPN-approved compound ruxolitinib as a potential strategy to treat SARS-CoV-2 induced systemic hyperinflammation. *Leukemia* **34**, 1723-1725 (2020).

86. Richardson, P. *et al.* Baricitinib as potential treatment for 2019-nCoV acute respiratory disease. *Lancet* **395**, e30-e31 (2020).

87. Wijaya, I. *et al.* The use of Janus Kinase inhibitors in hospitalized patients with COVID-19: Systematic review and meta-analysis. *Clin Epidemiol Glob Health* **11**, 100755 (2021).

88. Patoulias, D., Doumas, M., Papadopoulos, C. & Karagiannis, A. Janus kinase inhibitors and major COVID-19 outcomes: time to forget the two faces of Janus! A meta-analysis of randomized controlled trials. *Clin Rheumatol* **40**, 4671-4674 (2021).

89. in *Coronavirus Disease 2019 (COVID-19) Treatment Guidelines* (National Institutes of Health (US), Bethesda (MD), 2021).

90. Sirota, M. *et al.* Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Sci Transl Med* **3**, 96ra77 (2011).

91. Gamazon, E.R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics* **47**, 1091-1098 (2015).

92. Hu, Y. *et al.* A statistical framework for cross-tissue transcriptome-wide association analysis. *Nature Genetics* **51**, 568-576 (2019).

93. Zhou, D. *et al.* A unified framework for joint-tissue transcriptome-wide association and Mendelian randomization analysis. *Nat Genet* **52**, 1239-1246 (2020).

94. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* **10**, e1004383 (2014).

95. Wallace, C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet* **17**, e1009440 (2021).

96. Blanco-Melo, D. *et al.* Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* **181**, 1036-1045.e9 (2020).

97. Lamers, M.M. *et al.* SARS-CoV-2 productively infects human gut enterocytes. *Science* **369**, 50-54 (2020).