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], Lang Wu[4], Ruiyuan Pan[5], Ravi V. Shah[6], Eric R. Gamazon[7,8,*], Dan Zhou[1,2,*]

[1] The Second Affiliated Hospital and School of Public Health, 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] Cancer Epidemiology Division, Population Sciences in the Pacific Program, University of Hawaii Cancer Center, University of Hawaii at Manoa, Honolulu, HI, USA

[5] Key Laboratory of Mental Health of the Ministry of Education, Guangdong-Hong Kong-Macao Greater Bay Area Center for Brain Science and Brain-Inspired Intelligence, Guangdong-Hong Kong Joint Laboratory for Psychiatric Disorders, Guangdong Province Key Laboratory of Psychiatric Disorders, Guangdong Basic Research Center of Excellence for Integrated Traditional and Western Medicine for Qingzhi Diseases, Department of Neurobiology, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China

[6] Division of Cardiology, Vanderbilt University Medical Center, Nashville, TN, USA;

[7] Division of Genetic Medicine, Vanderbilt University Medical Center, Nashville, TN, USA;

[8] Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA;

[#,*]equally contributed

[*]Corresponding to:

Dan Zhou: danzhou@zju.edu.cn

Eric R. Gamazon: egamazon@gmail.com

Abstract

Drug repositioning presents a streamlined and cost-efficient way to expand the range of therapeutic possibilities. Drugs with human genetic evidence are more likely to advance successfully through clinical trials towards FDA approval. Single gene-based drug repositioning methods have been implemented, but approaches leveraging a broad spectrum of molecular signatures remain underexplored. We propose a framework called "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 response profiles in a high-dimensional normed space, quantifying the reversal potential of candidate drugs in a disease-related cell-based screening. Here, we implemented this framework to identify potential therapeutics relevant to COVID-19 and type 2 diabetes (T2D). For COVID-19, we identified 36 drugs showing potential reversal roles. Notably, nearly 70% (25/36) of the drugs have been linked to COVID-19 from other studies, with seven drugs supported by ongoing/completed clinical trials. For T2D, we observed reversal signals for 16 compounds on multiple disease signatures. Five drugs are supported by published literature, affirming potential therapeutic value. In summary, we propose a comprehensive genetics-anchored framework integrating population-level signatures and cell-based

screening that has the potential to accelerate the search for new therapeutic strategies.

Key words

Drug repositioning; signature matching; TWAS; reversal distance; COVID-19; type 2 diabetes

Introduction

Conventional drug development is expensive in cost and time, spanning over a decade with substantial financial investment to go from target identification to therapeutics(DiMasi et al., 2003); Wouters et al. (2020). Drug repositioning represents a cost-effective strategy to extend the utility of existing drugs, leveraging previous studies on safety in preclinical and clinical studies to accelerate drug application.

Genetic disease associations may play a significant role in repositioning, since empirical data suggests that drug targets with genetic support are 2.6 times greater to be successful(Minikel et al., 2024). Genome-wide, transcriptome-wide, and proteome-wide association studies (GWAS, TWAS, PWAS) are powerful genetic tools to predict drug repurposing candidates: for example, the transcriptome-wide significant association between *TYK2* and critical illness in COVID-19(Pairo-Castineira et al., 2021), facilitated the drug repurposing of baricitinib, which has demonstrated therapeutic benefit in subsequent clinical trials(Rubin, 2022).

However, drug repositioning focused on a single gene target may not be sufficient to elicit a meaningful therapeutic response as complex diseases involve intricate networks of pathways and interactions(Pushpakom et al., 2019). The growing evidence for polypharmacology calls for the

adoption of experimental-computational multi-target approaches(Paolini et al., 2006). Most existing approaches rely on differential gene expression (DGE) to generate a list of prioritized 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(Porcu et al., 2021). Moreover, 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, a method that performs an inverse pattern-matching alignment of disease signatures and drug response profiles(Chen et al., 2017) suggests that, for a given disease signature comprising genes with altered up/downregulation, a drug that reverses the disease-associated gene expression changes may hold therapeutic potential for the disease. Genetically-anchored TWAS minimizes the problem of reverse causality. Indeed, Wu et al. applied PrediXcan models to find drug repurposing candidates using the Integrative Library of Integrated Network-based Cellular Signatures (iLINCS) platform(Wu et al., 2022). 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)(Subramanian et al., 2017). Methodologically, previous approaches (e.g., a widely adopted nonparametric pattern-matching method based on the Kolmogorov−Smirnov (KS) statistics(Lamb et al., 2006)) do not fully consider the pattern of alignment between the (transcript)omic features of drug induction and disease susceptibility, including, explicitly, genetically-determined risk, within a multitarget or polygenic framework. In summary, a genetics-informed omics-based repurposing framework that addresses critical methodological gaps is needed.

To address these gaps, here we developed Transcriptome-informed Reversal Distance (TReD) to integrate cell-based drug response 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 and T2D.

Results

Study overview

Genetics-informed single-gene matching and omics profile-based signature matching are two promising approaches used in drug repurposing, with diverse types of omics data providing a valuable opportunity. 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 (Fig 1A, Supplemental Fig. 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 drug response profiles (CMAP LINCS 2020) (Fig. 1B, Supplemental Fig. S1). In this study, we focused primarily on the drug response profile data arising from small-molecule perturbations in immune-related cells and blood glucose metabolism-related tissues for the specific TReD applications to COVID-19 and T2D, respectively. TReD is based on transcriptome signature matching, which relies on the signature reversion principle (Wagner et al., 2015). In contrast to previous transcriptome signature pattern-matching studies, we proposed a

"reversal distance" which quantifies the drug's potential in reversing the disease-associated gene expression, using a metric that is scale-invariant and directionality-focused. To conduct hypothesis testing and evaluate the statistical significance, we employed a permutation approach.

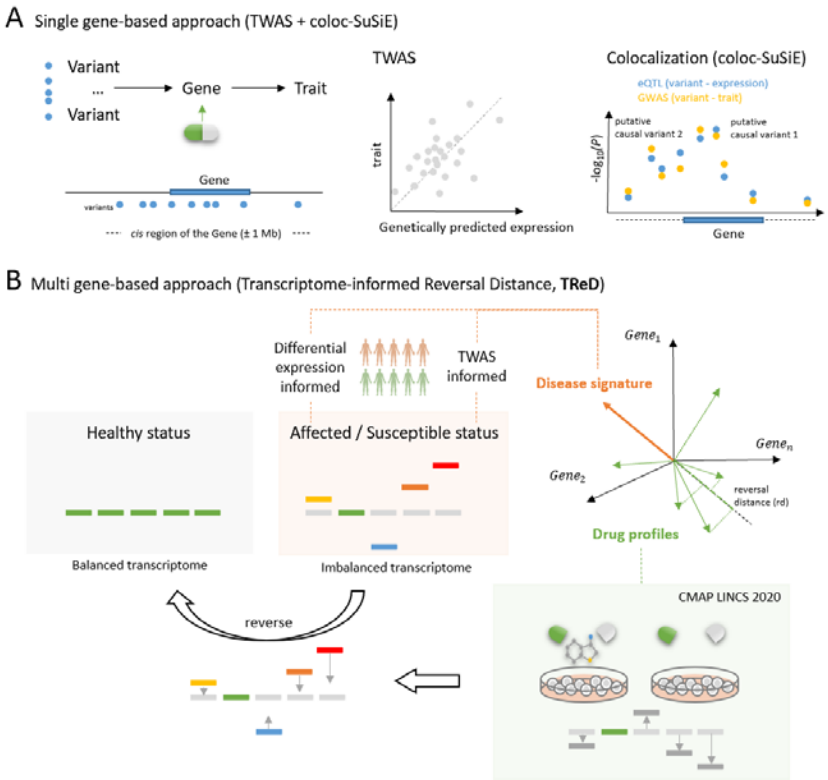

Figure 1. The framework overview. (A) Single gene-based approach integrates TWAS (e.g., PrediXcan) and colocalization analysis (e.g., coloc-SuSiE). (B) Multi-gene-based approach TReD (Transcriptome-informed Reversal Distance) utilizes DGE- and TWAS-derived population-level gene expression profiles as disease signatures, reflecting affection status and germline susceptibility, respectively. By embedding the population-level disease signatures (orange vector) and the cell-based drug response 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 (**Methods**).

Single-target drug repositioning using TWAS and coloc-SuSiE

We performed TWAS to identify association between gene expression and a given phenotype. Models to predict gene expression were fit using a reference panel, subsequently applying gene expression models to GWAS of COVID-19 severity and T2D. We used data from Genotype-Tissue Expression (GTEx) v8, focusing on select tissues potentially relevant to COVID-19 severity (whole blood, lung, lymphocytes, and spleen samples(Vanderbeke et al., 2021)) or T2D (visceral adipose, liver, pancreas, skeletal muscle). Genes that survived Bonferroni correction for each tissue were considered significant (Fig. 2, Supplemental Table S1). For TWAS results significant after multiple testing, we performed coloc/coloc-SuSiE. eQTL and GWAS signals were determined to "colocalize" if maximum posterior probability of colocalization (i.e., PP.H4) was greater than or equal to 0.85 for coloc or coloc-SuSiE. The results of coloc are in Supplemental Table S2 and those of coloc-SuSiE are in Supplemental Table S3. We further examined which prioritized disease-related genes might be drug targets. Drug target genes were defined using public databases, including DrugBank(Wishart et al., 2018), the Drug Gene Interaction Database (DGIdb)(Freshour et al., 2021), and the Therapeutic Target Database (TTD)(Zhou et al., 2022).

For COVID-19 severity, we identified 34, 45, 27, and 43 significant genes in blood, lung, lymphocytes, and spleen, respectively in TWAS. Nine genes showed significance across four tissues. Notably, *tyrosine kinase 2* (*TYK2)* was significant in all tissues with a concordant positive effect, signifying that an increased expression of *TYK2* would be 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(Hu et al., 2021). High expression of *TYK2* and host-driven inflammatory lung injury are related to life-threatening COVID-19(Pairo-Castineira et al., 2021). *RAB2A* also showed a positive effect in TWAS in the four tissues. *RAB2A* is required for SARS-CoV-2 replication(Li et al., 2020), and higher expression of *RAB2A* is associated with poorer COVID-19 outcome(Pairo-Castineira et al., 2023), consistent with directionality we observed.

In the four analyzed tissues, we identified a total of 28 unique genes post coloc/coloc-SuSiE (Supplemental Fig. S2). Among these targets, 50% (14/28) have been reported to be associated with COVID-19 (*DPP9*(Wang et al., 2021a), *WNT3*(Wu et al., 2021), *ABO*(Ellinghaus et al., 2020; Ray et al., 2021), *NAPSA*(Krishnamoorthy et al., 2023), *GSDMB*(Chang et al., 2022), *RAB2A*(Pairo-Castineira et al., 2023), *OAS1*(Banday et al., 2022), *OAS3*(Pairo-Castineira et al., 2021), *LRRC37A2*(Zhu et al., 2023), *MAPT*(Cao et al., 2023), *FUT2*(Andreakos et al., 2022; Kousathanas et al., 2022), *IKZF3*(Yao et al., 2023), *ICAM5*(Kousathanas et al., 2022), *CCR9*(Ellinghaus et al., 2020; Yao et al., 2021)). *CCR9* and *MAPT* are druggable genes: drugs such as docetaxel and astemizole, which target *MAPT,* had COVID-19-related support. Docetaxel can inhibit the active site of SARS-CoV-2 RdRP(Bagabir, 2023). Astemizole (antihistamine) can interfere with SARS-COV-2 entry (Wang et al., 2021b).

For T2D, 86 genes were identified by both TWAS and coloc/coloc-SuSiE in tissues examined (Fig. 3). We further investigated their associations with T2D in the Type 2 Diabetes Knowledge Portal(Costanzo et al., 2023). Among the 86 genes, ten have been classified as potential T2D effectors, including *Insulin receptor substrate 1* (*IRS1*), a gene central in insulin signal transmission(Sun et al., 1992). We also identified T2D-associated genes *PTH1R* and *NOTCH2*, both druggable. *PTH1R* showed a negative effect

on T2D in visceral adipose tissue. Teriparatide (TPTD) and abaloparatide agonize *PTH1R*, and TPTD may improve glucose control in select populations (Mazziotti et al., 2014). Similarly, *NOTCH2* showed a negative effect on T2D in both liver and visceral adipose tissue (though drugs targeting NOTCH2 like nirogacestat and rg-4733 are inhibitory).

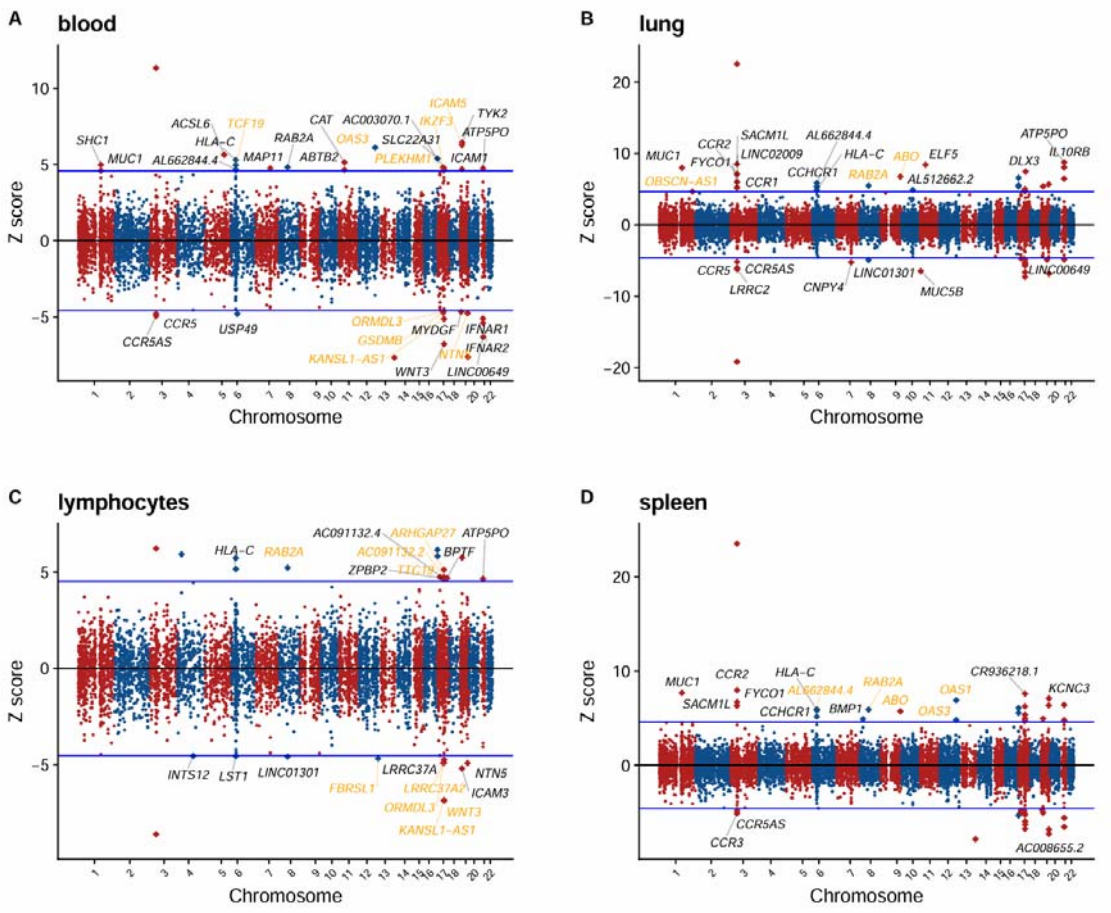

Figure 2. TWAS and colocalization analysis for COVID-19. Manhattan plots of TWAS results for COVID-19 severity in whole blood (A), lung (B), lymphocytes (C), and spleen (D), respectively. The *Z*-scores for TWAS are shown. *Z*-score > 0 indicates a gene for which higher genetically-determined expression is associated with COVID-19 severity. The labelled genes survived multiple testing corrections (over or below the blue lines) given the number of genes tested. Genes that further pass the significance threshold from the colocalization analysis are colored gold.

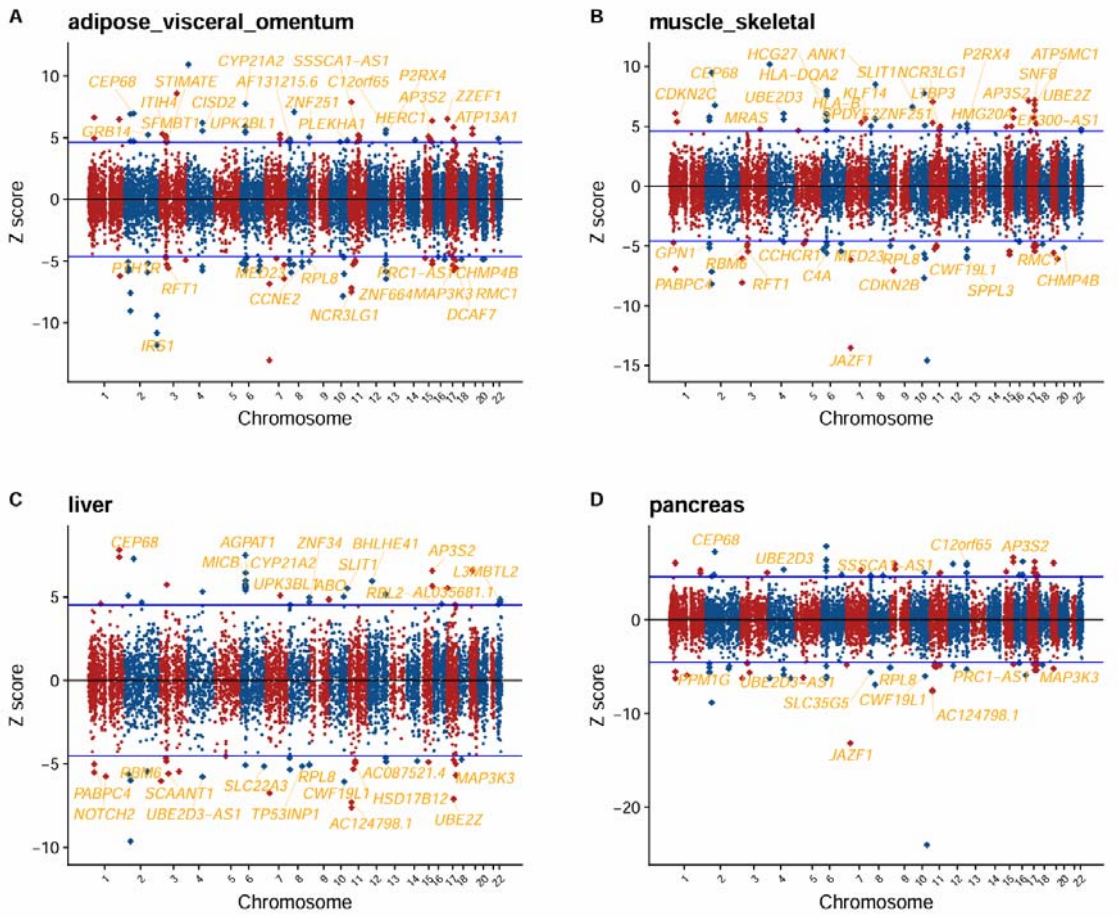

Figure 3. TWAS and colocalization analysis for T2D. Manhattan plots of TWAS results for T2D in visceral adipose (A), skeletal muscle (B), liver (C), and pancreas (D), respectively. The *Z*-scores for TWAS are shown. *Z*-score > 0 indicates a gene for which higher genetically-determined expression is associated with T2D. The labelled genes survived multiple testing corrections (over or below the blue lines) given 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

Next, we applied TReD to prioritize potential drug repurposing candidates for COVID-19 and T2D.

We sought to identify significant reversal effects from the cellular drug response profiles on the population-level disease signatures at transcriptome level (**Methods**). In addition to four TWAS-derived disease signatures, we included three DGE-based disease signatures (for COVID-19: ALV (alveolar), EXP (expansion), BALF (bronchoalveolar lavage fluid); for T2D: islets, myoblasts, myotubes; Table S4 and Methods). Rules for prioritizing drugs are in **Methods**. Briefly, we defined a metric called reversal distance (*rd*) as a metric of how a drug may counteract a disease signature (defined by cellular transcriptomic response to the drug). For a drug to be considered a potential candidate, its reversal distance must be two or more standard deviations away from the mean *rd* of the null distribution and show nominal significance in a permutation test. To ensure robustness, we further required consistent results across technical LINCS replicates, and at least one drug response profile analysis (permutation p-value) for the given drug must pass Bonferroni correction.

For COVID-19, we identified 707 drugs showing nominal significance in any of the seven disease signatures. For perturbations without repeated LINCS experiments, a consistent result in a confirmatory cell line is required to support the candidate drug response profile. Additionally, we highlighted drugs that exhibited a reverse effect in at least four disease signatures. Lastly, we prioritized 36 drugs for COVID-19 that show significance ($P_{adj} < 0.05$) in at least one drug response profile (Figure 4). Table 1 provides an overview of the 36 drugs, including the reversal distance across the different disease signatures and whether there is any prior evidence linking to COVID-19 and/or clinical studies for the treatment of COVID-19. Notably, nearly 70% of the drugs (25/36) are reported related to COVID-19, and seven drugs have ongoing/completed clinical trials (Table 1).

Of note, over 40% of candidates are immunomodulatory drugs, especially JAK inhibitors. Baricitinib, significantly reversing five disease signatures (Fig. 4), is a JAK1/2 and TYK2 inhibitor

approved for COVID-19 therapy in 2022. Ruxolitinib significantly reversed six disease signatures (Fig.4), with a variety of reports supporting divergent clinical efficacy (Koschmieder et al., 2020; Vannucchi et al., 2021) (Han et al., 2022) (Rein et al., 2022). Analogously, enzastaurin—investigational for treatment of glioblastoma multiforme via selectively inhibition of PKC-beta and other kinases involved in cell proliferation and survival—appeared effective in reversing six disease signatures (**Figure 4)**. Of note, SARS-CoV can activate the PKC members(Ji et al., 2009), and inhibitors of protein kinase C family members, including enzastaurin, may decrease SARS-CoV-2 replication *in vitro* (Huang et al., 2022)

For T2D, we identified 78 drugs/compounds as potential candidates for at least two T2D disease signatures (one TWAS and one DGE). We further filtered out 62 drugs/compounds without any drug response profile passing multiple comparison. This resulted in a total of 16 drugs/compounds (Fig. 5). Five of the drugs/compounds have been previously reported as potential options for the prevention or treatment of T2D. For example, imatinib (notably elevated reversal distance) has been shown to improve insulin sensitivity and maintain glucose homeostasis(Veneri et al., 2005) potentially via effects on endoplasmic reticula stress (Han et al., 2009). Pimozide also showed a significant reversal effect and may alleviate hyperglycemia in diet-induced obesity through inhibition of skeletal muscle ketone oxidation(Al Batran et al., 2020). Batimastat demonstrated a significant reversal effect, with potential roles in T2D-related complications as a matrix metalloproteinase inhibitor. (Ramnath et al., 2014). Collectively, these results suggest pathophysiologic plausibility of TReD-defined targets in T2D.

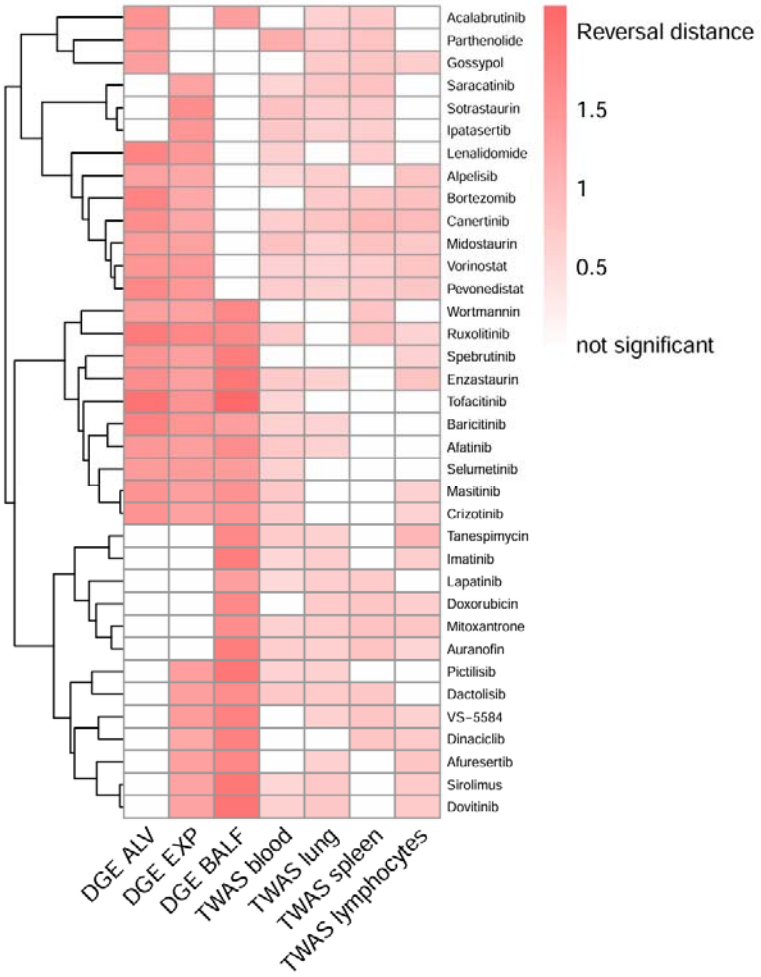

Figure 4. Drug repositioning candidates with potential reversal effect on COVID-19 disease signatures from DGE and TWAS. In total, 36 drugs were found by our Transcriptome-informed Reversal Distance (TReD) framework (**Methods**) to have a potential reversal effect on COVID-19 as indicated by at least 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 (the information on significance can be found in Supplemental Table S5). 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.

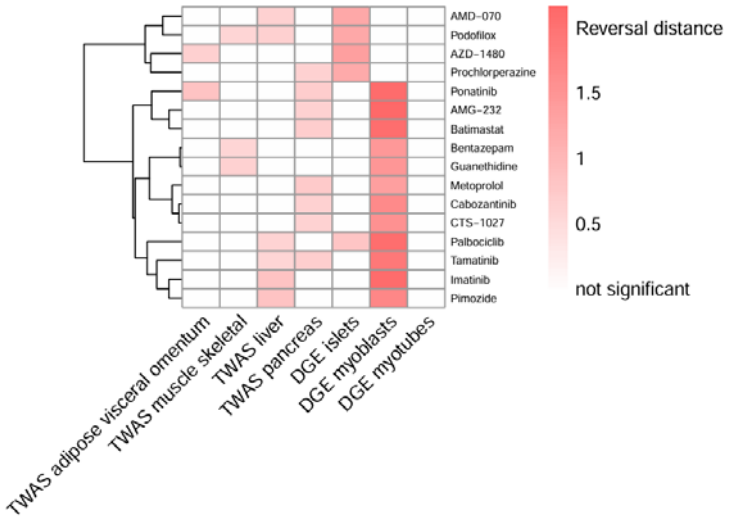

Figure 5. Drug repositioning candidates with potential reversal effect on T2D disease signatures from DGE and TWAS. In total, 16 drugs were found by our Transcriptome-informed Reversal Distance (TReD) framework (**Methods**) to have a potential reversal effect on T2D as indicated by two of the seven 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 (the information on significance can be found in Supplemental Table S6). 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, are described in Methods.

Table 1. DrugBank annotations for drugs with significant reversal effects on 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
		DGE ALV	DGE EXP	DGE BALF	TWAS blood	TWAS lung	TWAS spleen	TWAS lymphocytes		

Wortmannin	PI3Ks inhibitor	1.350	1.311	1.699	n.s.	n.s.	0.829	n.s.	Yes(Jiang et al., 2022)	No
VS-5584	Antineoplastic	n.s.	1.426	1.784	n.s.	0.644	0.794	0.640	Yes(Bharadwaj et al., 2022)	No
Vorinostat	Antineoplastic	1.465	1.458	n.s.	0.632	0.597	0.700	0.813	Yes(Sinha et al., 2020)	No
Tofacitinib	Antineoplastic and Immunomodulating	1.991	1.518	2.159	0.583	n.s.	n.s.	n.s.	Yes(Guimarães et al., 2021)	Yes
Tanespimycin	Antitumor antibiotics	n.s.	n.s.	1.665	0.740	0.653	n.s.	1.026	No	No
Spebrutinib	Immunomodulatory	1.507	1.333	1.828	n.s.	n.s.	n.s.	0.634	Yes(Kaliamurthi et al., 2021)	No
Sotrastaurin	Immunosuppressant	n.s.	1.625	n.s.	0.855	0.720	0.727	n.s.	Yes(Huang et al., 2022)	No
Sirolimus	Immunosuppressant	1.323	1.299	1.938	0.555	0.743	n.s.	0.733	Yes(Bischof et al., 2021)	Yes
Selumetinib	Antineoplastic	1.426	1.406	1.411	0.667	n.s.	n.s.	n.s.	Yes(Xie et al., 2022)	No
Ruxolitinib	Antineoplastic	1.871	1.685	1.666	0.722	n.s.	0.878	0.592	Yes(Lamontagne et al., 2020)	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(Bahrami et al., 2020)	No
Mitoxantrone	Chemotherapy drugs	n.s.	1.242	1.599	0.648	0.729	0.847	0.821	Yes(Preet et al., 2022)	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(Drayman et al., 2021)	Yes
Lenalidomide	Antineoplastic and Immunomodulating	1.750	1.441	n.s.	0.646	n.s.	0.714	n.s.	Yes(von Metzler et al., 2022)	No
Lapatinib	Antineoplastic	n.s.	n.s.	1.357	0.541	0.709	0.733	n.s.	Yes(Saul et al., 2023)	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(van de Veerdonk et al., 2022)	Yes
Gossypol	Polyphenolic compounds	1.374	n.s.	n.s.	n.s.	0.738	0.824	0.711	Yes(Wang et al., 2022)	No

Enzastaurin	Antineoplastic	1.616	1.346	1.957	0.738	0.642	n.s.	0.840	Yes(Huang et al., 2022)	No
Doxorubicin	Chemotherapy drugs	n.s.	n.s.	1.660	n.s.	0.730	0.798	0.685	Yes(Sajid Jamal et al., 2022)	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(Ahmad et al., 2023)	No
Dactolisib	Antineoplastic	n.s.	1.363	1.629	0.761	0.732	0.772	n.s.	Yes(Tomazou et al., 2021)	No
Crizotinib	Antineoplastic	1.513	1.318	1.459	0.720	n.s.	n.s.	0.625	Yes(Chen et al., 2021)	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	0.819	0.883	Yes(Li et al., 2022)	No
Baricitinib	Immunosuppressant	1.767	1.493	1.393	0.633	0.594	n.s.	n.s.	Yes(2022)	Yes
Auranofin	Antiinflammatory and Antirheumatic	n.s.	n.s.	1.815	0.715	0.664	0.844	0.588	Yes(Biji et al., 2021)	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(2020)	Yes
Saracatinib	Antineoplastic	n.s.	1.279	n.s.	0.576	0.771	0.859	n.s.	No	NO

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.

Table 2. DrugBank annotations for drugs with significant reversal effects on two disease signatures for T2D.

		Reversal distance for each of the disease signatures							Evidence for T2D	Evidence for clinical trials
Drug Hit	Description	DGE islets	DGE myoblasts	DGE myotubes	TWAS adipose	TWAS muscle	TWAS liver	TWAS pancreas		
AMG-232	Antineoplastic	n.s	2.452	n.s	n.s	n.s	n.s	0.643	No	No
Batimastat	Antineoplastic	n.s.	2.063	n.s	n.s.	n.s	n.s	0.715	Yes(Nose et al., 2003)	No
Imatinib	Tk inhibitor	n.s	2.094	n.s.	n.s.	n.s.	0.666.	n.s.	Yes(Veneri et al., 2005)	No
Pimozide	Antipsychotic	n.s.	1.692	n.s	n.s	n.s	0.830	n.s	Yes(Al	No

									Batran et al., 2020)	
Bentazepam	Antipsychotic	n.s	1.448	n.s	n.s	0.586	n.s	n.s	No	No
Guanethidine	Antihypertensive drug	n.s.	1.463	n.s.	n.s	0.652	n.s	n.s.	No	No
Cabozantinib	Antineoplastic	n.s	1.657	n.s	n.s	n.s	n.s	0.634	Yes(Ye et al., 2022)	No
Metoprolol	Antiarrhythmic	n.s	1.371	n.s	n.s	n.s.	n.s.	0.735	Yes(Sharma et al., 2008)	No
AMD-070	CXCR4 antagonist	1.231	n.s	n.s	n.s	n.s	0.645	n.s.	No	No
AZD-1480	JAK inhibitor	1.353	n.s	n.s	0.666	n.s	n.s	n.s	No	No
Prochlorperazine	Antipsychotic	1.202	n.s.	n.s.	n.s.	n.s.	n.s.	0.652	No	No
CTS-1027	Amines	n.s.	1.556	n.s.	n.s.	n.s.	n.s.	0.631	No	No
Podofilox	Antimitotic	1.231	n.s.	n.s.	n.s.	0.588	0.657	n.s.	No	No
Palbociclib	Antineoplastic	0.789	2.057	n.s.	n.s.	n.s.	0.598	n.s.	No	No
Ponatinib	Antineoplastic	n.s.	2.103	n.s.	0.848	n.s.	n.s.	0.699	No	No
Tamatinib	Syk inhibitor	n.s.	1.922	n.s.	n.s.	n.s.	0.584	0.688	No	No

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 (islets, myoblasts, and myotubes), are described in Methods.

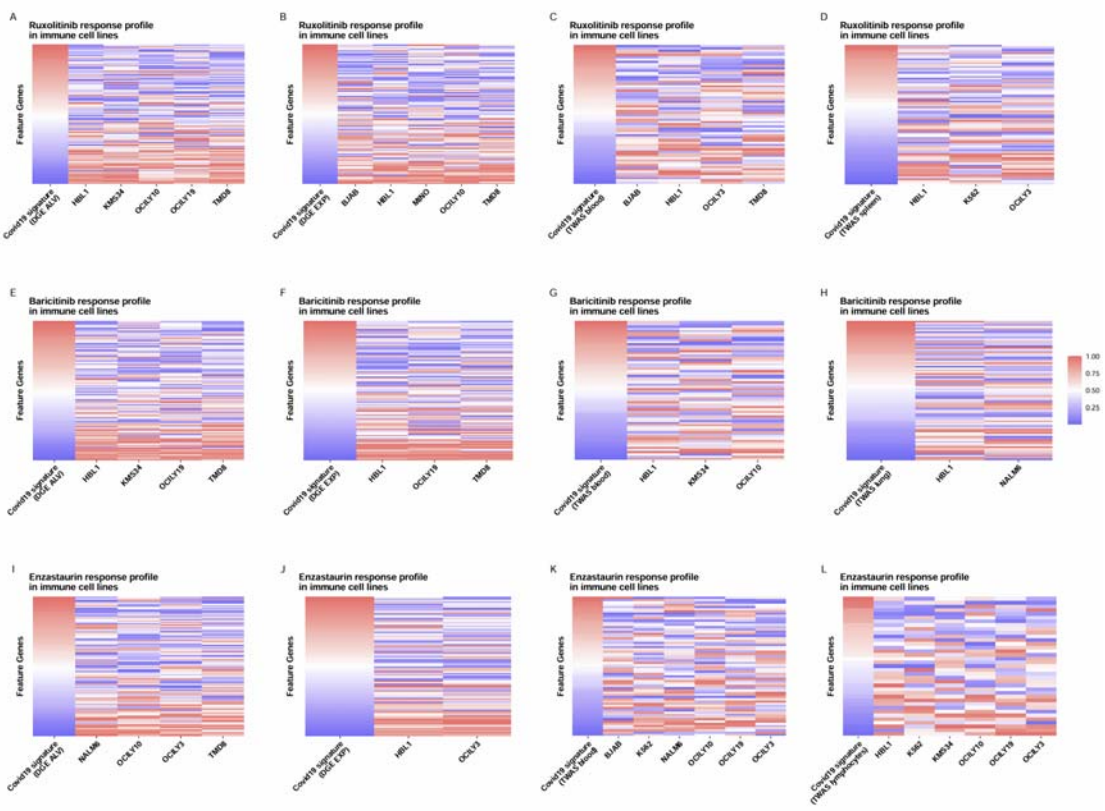

Figure 6. TReD-identified relationships between COVID-19 signatures and drug response profiles. The potential reversal effects of Ruxolitinib (A-D), Baricitinib (E-H), and Enzastaurin (I-L), which were identified against at least four disease signatures of COVID-19, are visualized using a heatmap. Taking (a) as an example, the first column in the heatmap represents differential gene expression (DGE) from the ALV 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 cells. For drugs with multiple significant profiles, the drug response profiles (e.g., different duration 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.

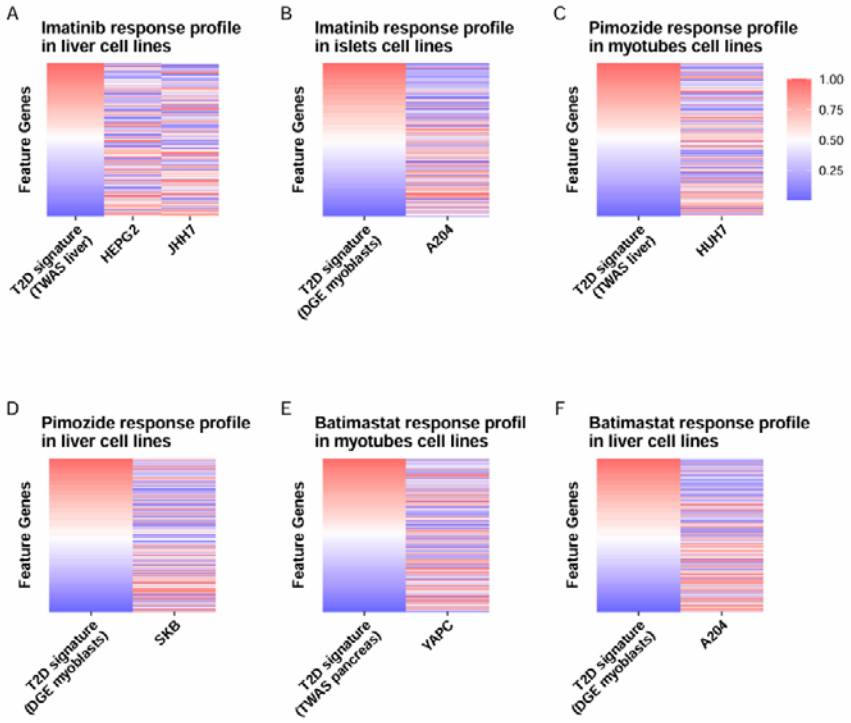

Figure 7. TReD-identified relationships between T2D signatures and drug response profiles. The potential reversal effects of Imatinib (A-B), Pimozide (C-D), and Batimastat (E-F), which were identified against two disease signatures of T2D, are visualized using a heatmap. Taking (a) as an example, the first column in the heatmap represents the transcriptome-wide association analysis (TWAS) derived disease signature in liver ranked by fold change. Red and blue denote higher and lower expressed in affected samples, respectively. The remaining columns represent gene expression changes after drug treatment in liver cells. For drugs with multiple significant profiles (e.g., different duration after treatment), the drug response profiles with the largest reversal distance are presented. The source data for the disease signatures, including TWAS (four tissue/cell types) and DGE, are described in **Methods**.

Discussion

Drug development pipelines have traditionally relied on high-throughput rational or biologic synthesis and screening approaches. Recent high-dimensional genomic and transcriptional approaches have yielded a path to prioritize potential targets for directed study(Sass et al., 2024). However powerful, traditional genomic approaches (e.g., TWAS) alone only identify potential pathways to target, without taking advantage of the large bank of information on druggability and cellular responses to drugs already in existence. Here, we developed a framework to identify potential drug repurposing candidates by combining population-level disease signatures via human genetics (TWAS/DGE) and cellular responses to drugs at the transcriptional level (TReD). These approaches were successfully applied to COVID-19 and T2D, identifying known (e.g., baricitinib/TYK2) and novel drug-disease combinations.

Using TReD, we found a significant correlation between potential drug candidates and associated disease, specifically within pathophysiologies central to the two conditions studied. For example, in COVID-19, nearly half (16/36) of identified drugs were immunomodulatory (critical to COVID-19 pathogenesis), with prior study in COVID-19. Indeed, one target identified by TReD has in-human approval as a COVID-19 therapeutic (baricitinib in severe COVID-19), with others recommended as an alternative (tofacitinib (2021)). Our results in T2D were similar in scope. Among the prioritized drugs, we have highlighted pimozide and imatinib as having strong associations with T2D. Collectively, these results demonstrate the ability of this novel TReD framework to identify drug repositioning candidates.

We also investigated the drugs that are *positively* correlated with the disease signatures of COVID-19. A few such top-ranked drugs (including testosterone(Yassin et al., 2023) and

otamixaban(Hempel et al., 2021)) were also supported by published studies as potential treatment options. This scenario may arise, for example, when a disease signature targeted by such a drug lies in a co-expression network with another disease signature whose net reversal is a potential therapeutic option. This result indicates that, in reality, the biological system operates in a much more complex fashion.

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 response 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. 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 and T2D, our framework is highly generalizable and can be implemented to identify drug repositioning candidates for other diseases.

While encouraging as a first attempt to integrate TWAS and cellular response profiles, TReD has a key limitation: this approach is dependent on comprehensive curation of drug response profiles over a wide range of drug targets, cell types, and disease conditions. Indeed, while we were broad in scope for inclusion of curated data for these elements, the cellular response profiles were obtained primarily in cancer cell lines, which may not necessarily phenocopy the cell types and biological context of interest for the condition under study (e.g., COVID-19). Although we requested a confirmatory cell line with concordant response to support the findings, the treatment dose and duration in the confirmatory cell line are not always consistent. In addition, we recognize that most medications here are necessarily anti-cancer therapeutics, which may not be clinically translatable to other broad conditions (e.g., T2D) due to off-target (side) effects. Finally, other "omes" beyond transcriptional responses are likely relevant (e.g., protein expression, non-coding RNA) that may encode trans-organ signaling pathways and a broad cell atlas relevant to multi-system disorders (e.g., T2D). Our study does not consider the complex issues of drug-target interactions and drug-drug interaction, and consideration of in vivo pharmacokinetics, pharmacodynamics, and substrate processing are critical. The solution to these limitations is either a strength of older drug discovery approaches—broad, rapid screening functional assays across cellular phenotypes and broad synthetic or derivatized chemical libraries or biologics for screening (e.g., rapid lentiviral CRISPR screening(Chaffin et al., 2022))—or a strength of ongoing drug development pipelines (Phase 0 and animal system studies). The role for

TReD in this pipeline is to hone the potential space of therapeutics and provide a multi-dimensional platform to integrate across molecular genetic information to improve drug discovery efficiency.

In summary, integrating transcriptome-level profiles from cell-based drug response profiles and population-based disease signature studies, we developed a framework, TReD, to identify drug repurposing candidates. Deploying TReD in COVID-19 and T2D successfully identified both known and novel candidates that capture underlying pathobiological mechanisms important in each disease. These results suggest that computational approaches that unite molecular genetic information with cellular response profiles may be effective in drug repositioning across a wide array of disorders. Future work collating cell types, pharmacotherapies, and disease states is needed to fully realize the potential of this and analogous approaches.

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 refers to hospitalization with a laboratory confirmation of SARS-CoV2 infection. This dataset contains 44,989 cases and 2,356,386 controls. The summary statistics of GWAS of T2D were obtained from Xue et al's study. This dataset contains 62,892 T2D cases and 596,424 controls of European ancestry(Xue et al., 2018). (http://cnsgenomics.com/data/t2d/Xue_et_al_T2D_META_Nat_Commun_2018.gz). Additional descriptions of the included studies can be found in Supplemental Table S7.

TWAS and colocalization for COVID-19 severity and T2D

To identify disease associated genes, we applied TWAS with the best prediction models among PrediXcan(Gamazon et al., 2015), UTMOST(Hu et al., 2019), and JTI(Zhou et al., 2020) according to the prediction accuracy estimated by 5-fold cross-validation. We used the residuals of the normalized gene expression after regressing out relevant covariates. SNPs in the cis region (i.e.,1 Mb upstream and downstream) of the gene were considered candidate features for the gene's prediction models. Applying the prediction model to GWAS dataset, we would estimate the genetically predicted expression level for each individual. The significance of the effect of the predicted expression of a gene on a trait was determined by linear regression. In practice, summary statistics-based association test was performed. Summary-statistics based association test utilizes the weights from the prediction model of the gene under consideration, the covariances of the SNPs included in the prediction model for the gene, and the GWAS effect size coefficient for each SNP. Notably, the summary statistics-based approach results in nearly perfect concordance with the individual-level based approach(Barbeira et al., 2018).

SARS-CoV-2 directly attacks pulmonary alveolar cells and causes lung injury. Pulmonary infections leading to COVID-19 induce adaptive immune responses while effector immune cells are generated and recruited in the lymphocytes, spleen, and blood. Therefore, blood, lung, lymphocytes, and spleen were considered as relevant tissues for COVID-19. Visceral adipose, liver, pancreas, and skeletal muscle were considered relevant tissues for T2D. For TWAS results that passed Bonferroni multiple testing correction, we performed colocalization to provide additional support under different sets of assumptions. To perform colocalization analysis in the presence of multiple potentially causal variants, we applied coloc(Giambartolomei et al., 2014) to the GWAS hits decomposed by SuSiE(Wallace, 2021) in each GWAS locus, which spans the SNPs located within 200 kb (in each

direction) of the lead SNP. Notably, *coloc-SuSiE* is more generalizable since it relaxes the "one causal variant" assumption of regular *coloc*. We used European individuals in the 1000 Genomes Project as the LD panel. If SuSiE failed to converge after 1000 iterations for either eQTL or GWAS summary statistics, we instead used the regular *coloc*. eQTL and GWAS signals were determined to "colocalize" if the maximum posterior probability of colocalization (i.e., PP.H4) was greater than or equal to 0.85 using either *coloc* or *coloc-SuSiE*. 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 an online resource on drug-gene interactions and the druggable genome(Cannon et al., 2024).

DGE for COVID19 and T2D

In COVID-19, the DGE data—consisting of independent gene expression signatures labelled 'ALV', 'EXP', 'BALF'—were derived from the study of Le et al(Le et al., 2021). The ALV signature was generated from RNA-seq data in human adenocarcinomic alveolar basal epithelial cells (GEO Dataset: GSE147507; n=67) (Blanco-Melo et al., 2020) (Supplemental Table S8). After Benjamini-Hochberg adjustment, the study identified 120 differentially expressed genes—100 upregulated and 20 downregulated. The EXP signature was obtained by RNA-seq on organoid samples (Lamers et al., 2020) (Supplemental Table S9) (GSE149312; n=22). After Benjamini-Hochberg adjustment and applying a fold-change cutoff (|log2 FC| > 2), 125 genes remained significant. The BALF signature was estimated by RNA-seq analysis of the bronchoalveolar lavage fluid samples from two COVID-19 patients and three controls (Xiong et al., 2020). After

DESeq2 processing, 1349 genes were used to define the BALF signature (Table S10).

Three DGE datasets (labelled as 'islets', 'myoblasts', and 'myotubes') were included in the T2D analysis. The islet signature was generated from microarray data on pancreatic islets derived from T2D cases and normal glucose-tolerant controls (Gunton et al., 2005), consisting of 370 differentially expressed genes at $p < 0.01$ (Supplemental Table S11). The other two signatures were obtained from the Davegårdh C et al study (Davegårdh et al., 2021). Based on a false discovery rate less than 5%, the study identified 577 differentially expressed genes in myoblasts and 42 differentially expressed genes in myotubes (Supplemental Table S12 and Table S13)

The CMAP LINCS 2020 dataset

The Library of Integrated Network-Based Cellular Signatures (LINCS) is a NIH Common Fund program. Its primary goal is to generate a comprehensive atlas of perturbation-response signatures. The CMAP LINCS 2020 data contains drug response profiles (defined by drug-induced gene expression profiles in response to perturbations). Here, we mainly focused on the transcriptomic response to small-molecule (drugs/compounds) perturbations in immune-related cells and blood glucose metabolism-related tissues for the application of the TReD to COVID-19 and T2D, respectively. We downloaded the level 5 beta data consisting of drug response profiles arising from small-molecule perturbations from CMAP LINCS 2020 (https://s3.amazonaws.com/macchiato.clue.io/builds/LINCS2020/level5/level5_beta_trt_cp_n720216x12328.gctx),

Candidate drug prioritization

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 signature as a vector in a

high-dimensional space (whose dimension n is the total number of genes) with each nonzero element of the vector given by a disease-associated gene. The disease signature can be derived from either TWAS or DGE. The vector for each candidate drug is embedded in the same n-dimensional space. For each disease signature vector and drug response profile pair, we compute the reversal distance. Reversal distance measures a drug's ability to counteract disease-related gene expression. A positive value suggests potential preventive or therapeutic effects, while a negative value indicates the drug may worsen the disease. A higher positive reversal distance suggests a stronger counteracting effect. 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:

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

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

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

$rd(\overrightarrow{G}, \overrightarrow{D})$. Equation 2 defines a function:

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

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 $\overrightarrow{G}$ determines which genes are the "relevant" elements of the drug response profile vector $\overrightarrow{D}$.

Of interest to us is the sampling distribution of the cosine similarity ω. Consider the function $h(s, t, u) = \frac{s}{\sqrt{tu}}$. Note that ω is equal to $h\left(\vec{G} \,.\, \vec{D},\ \vec{G} \,.\, \vec{G},\ \vec{D} \,.\, \vec{D}\right)$. Furthermore, $\vec{G} \,.\, \vec{G}$ and $\vec{D} \,.\, \vec{D}$ are $\chi^2(n)$ distributed. We leverage the so-called Delta Method. Using the first two terms of the Taylor series expansion and setting $\Gamma = (s, t, u)$ with expected value $\gamma = \mathrm{E}[\Gamma]$, we obtain approximate expressions for the mean and variance of $h(\Gamma)$:

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

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

where ∇h is the gradient.

In this framework, drug repurposing is a search problem: the vector $\overrightarrow{G}$ is evaluated against all possible $\overrightarrow{D}$, 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, $rd(\overrightarrow{G}, \overrightarrow{D_i})$, we get a probability distribution consisting of the K probabilities p_i conditional on the disease signature information:

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

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

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

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

We converted the gene-disease association signals 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 was rank-transformed.

For each drug response profile, the p-value of the reverse distance (*rd*) is estimated by a permutation test. To generate the null distribution, we randomly shuffle the labels 1000 times to create numerous permuted samples. These samples are generated under the assumption of the null hypothesis, meaning there is no effect or association. In practice, the rank-ratio of disease signature is shuffled randomly. A null reversal distance is calculated on the basis of the shuffled disease signature. Then, we calculate the real reversal distance and estimate the proportion of permuted reversal distance that are as extreme or more extreme than the reversal distance obtained from the real data. This proportion represents the p-value. For each of the disease signatures, a significant drug response profile is defined as showing nominal significance in the permutation test ($P < 0.05$) and large rd (over 2 standard deviations from the mean across all the drug response profiles). In addition, we further adjust the experiment-wise permutation p-value (pooling different perturbations) across the tested cell lines using Bonferroni correction. For each disease, drugs with at least one significant ($P_{adj} < 0.05$) drug response profile were prioritized.

We also implemented a degree-preserving permutation approach. Let $G = (V, E)$ be a graph,

where V is the set of nodes (genes) and E is the set of edges (defined by a co-expression network in a given tissue). To generate a permutation which preserves the degree distribution of the network, we randomly select two edges (say, $a \leftrightarrow b$ and $c \leftrightarrow d$) with non-overlapping nodes, remove the edges between the nodes, and create new edges ($a \leftrightarrow c$ and $b \leftrightarrow d$) by swapping the nodes until a certain proportion (80%) of the edges have been shuffled. The null reversal distance is computed for each such permutation.

Application of TReD to COVID-19 severity and T2D

Leveraging the TWAS/DGE results and the cell-based gene expression data from CMAP LINCS 2020, we implemented the TReD framework to prioritize drug repositioning candidates for COVID-19 and T2D. Immune-related and T2D-related cell lines (matching the T2D-related tissues) have been annotated by multiple resources including the CMAP LINCS 2020 and the CCLE (cancer cell lines which have been used to study COVID-19 and other diseases(Bakowski et al., 2021)). Information on drugs and drug targets is curated in the bioinformatics and chemoinformatics resource DrugBank (DrugBank Release Version 5.1.10 | DrugBank Online) A total of 1112 drugs and 12 immune-related cell types were included for COVID-19, while 1980 drugs and 10 related cell types were included for T2D.

We leveraged the TWAS results of multiple tissues and three DGE profiles as disease signatures (as defined in the previous section). For perturbations with repeated LINCS experiments, consistent direction of *rd* is required. i.e., drug response profile with repeated experiments showing inconsistent direction of *rd* will be removed. Considering the multiple testing, we prioritize drugs which show at least one significant drug response profile ($P_{adj} < 0.05$) for each disease.

Software Availability

The source codes used in this study are available at GitHub(https://github.com/zdangm/TReD). Source codes are also provided as Supplemental Code.

Acknowledgments

This research is supported by the National Natural Sciences Foundation of China 82204118 (D.Z.), the Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province 2020E10004 (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.). R.V.S. is supported by grants from the National Institutes of Health.

Author contributions: Conceptualization, Dan Zhou and Eric R. Gamazon; formal analysis, Chunfeng He, Yue Xu, Yuan Zhou, and Ran Meng; writing--original draft, Chunfeng He, and Yue Xu; writing--review & editing, Dan Zhou, Eric R. Gamazon, Yuan Zhou, Jiayao Fan, Chunxiao Cheng, Lang Wu and Ravi V. Shah.

Competing interest statement：Dr. Shah has equity ownership in Thryv Therapeutics. Dr. Shah is a co-inventor on a patent for ex-RNAs signatures of cardiac remodeling and a pending patent on proteomic signatures of disease. The remaining authors read and approved the final manuscript.

References

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

(2022). 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.

Ahmad, S., Pasha Km, M., Raza, K., Rafeeq, M.M., Habib, A.H., Eswaran, M., and Yadav, M.K. (2023). Reporting dinaciclib and theodrenaline as a multitargeted inhibitor against SARS-CoV-2: an in-silico study. J Biomol Struct Dyn *41*, 4013-4023.

Al Batran, R., Gopal, K., Capozzi, M.E., Chahade, J.J., Saleme, B., Tabatabaei-Dakhili, S.A., Greenwell, A.A., Niu, J., Almutairi, M., Byrne, N.J., et al. (2020). Pimozide Alleviates Hyperglycemia in Diet-Induced Obesity by Inhibiting Skeletal Muscle Ketone Oxidation. Cell Metab *31*, 909-919.e908.

Andreakos, E., Abel, L., Vinh, D.C., Kaja, E., Drolet, B.A., Zhang, Q., O'Farrelly, C., Novelli, G., Rodríguez-Gallego, C., Haerynck, F., et al. (2022). A global effort to dissect the human genetic basis of resistance to SARS-CoV-2 infection. Nat Immunol *23*, 159-164.

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

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

Bakowski, M.A., Beutler, N., Wolff, K.C., Kirkpatrick, M.G., Chen, E., Nguyen, T.H., Riva, L., Shaabani, N., Parren, M., Ricketts, J., et al. (2021). Drug repurposing screens identify chemical entities for the development of COVID-19 interventions. Nat Commun *12*, 3309.

Banday, A.R., Stanifer, M.L., Florez-Vargas, O., Onabajo, O.O., Papenberg, B.W., Zahoor, M.A., Mirabello, L., Ring, T.J., Lee, C.H., Albert, P.S., et al. (2022). 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.

Barbeira, A.N., Dickinson, S.P., Bonazzola, R., Zheng, J., Wheeler, H.E., Torres, J.M., Torstenson, E.S., Shah, K.P., Garcia, T., Edwards, T.L., et al. (2018). Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun *9*, 1825.

Bharadwaj, S., Azhar, E.I., Kamal, M.A., Bajrai, L.H., Dubey, A., Jha, K., Yadava, U., Kang, S.G., and Dwivedi, V.D. (2022). 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.

Biji, A., Khatun, O., Swaraj, S., Narayan, R., Rajmani, R.S., Sardar, R., Satish, D., Mehta, S., Bindhu, H., Jeevan, M., et al. (2021). Identification of COVID-19 prognostic markers and therapeutic targets through meta-analysis and validation of Omics data from nasopharyngeal samples. EBioMedicine *70*, 103525.

Bischof, E., Siow, R.C., Zhavoronkov, A., and Kaeberlein, M. (2021). 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.

Cannon, M., Stevenson, J., Stahl, K., Basu, R., Coffman, A., Kiwala, S., McMichael, J.F., Kuzma, K., Morrissey, D., Cotto, K., et al. (2024). DGIdb 5.0: rebuilding the drug-gene interaction database for precision medicine and drug discovery platforms. Nucleic Acids Res *52*, D1227-d1235.

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

Chaffin, M., Papangeli, I., Simonson, B., Akkad, A.D., Hill, M.C., Arduini, A., Fleming, S.J., Melanson, M., Hayat, S., Kost-Alimova, M., et al. (2022). Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. Nature *608*, 174-180.

Chang, J.J., Gleeson, J., Rawlinson, D., De Paoli-Iseppi, R., Zhou, C., Mordant, F.L., Londrigan, S.L., Clark, M.B., Subbarao, K., Stinear, T.P., et al. (2022). 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.

Chen, B., Ma, L., Paik, H., Sirota, M., Wei, W., Chua, M.S., So, S., and Butte, A.J. (2017). Reversal of cancer gene expression correlates with drug efficacy and reveals therapeutic targets. Nat Commun *8*, 16022.

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

Costanzo, M.C., von Grotthuss, M., Massung, J., Jang, D., Caulkins, L., Koesterer, R., Gilbert, C., Welch, R.P., Kudtarkar, P., Hoang, Q., et al. (2023). The Type 2 Diabetes Knowledge Portal: An open access genetic resource

dedicated to type 2 diabetes and related traits. Cell Metab *35*, 695-710.e696.
DiMasi, J.A., Hansen, R.W., and Grabowski, H.G. (2003). The price of innovation: new estimates of drug development costs. J Health Econ *22*, 151-185.
Drayman, N., DeMarco, J.K., Jones, K.A., Azizi, S.A., Froggatt, H.M., Tan, K., Maltseva, N.I., Chen, S., Nicolaescu, V., Dvorkin, S., et al. (2021). Masitinib is a broad coronavirus 3CL inhibitor that blocks replication of SARS-CoV-2. Science *373*, 931-936.
Ellinghaus, D., Degenhardt, F., Bujanda, L., Buti, M., Albillos, A., Invernizzi, P., Fernández, J., Prati, D., Baselli, G., Asselta, R., et al. (2020). Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med *383*, 1522-1534.
Freshour, S.L., Kiwala, S., Cotto, K.C., Coffman, A.C., McMichael, J.F., Song, J.J., Griffith, M., Griffith, O.L., and Wagner, A.H. (2021). Integration of the Drug-Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. Nucleic Acids Res *49*, D1144-d1151.
Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., Nicolae, D.L., Cox, N.J., et al. (2015). A gene-based association method for mapping traits using reference transcriptome data. Nature Genetics *47*, 1091-1098.
Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., and Plagnol, V. (2014). Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet *10*, e1004383.
Guimarães, P.O., Quirk, D., Furtado, R.H., Maia, L.N., Saraiva, J.F., Antunes, M.O., Kalil Filho, R., Junior, V.M., Soeiro, A.M., Tognon, A.P., et al. (2021). Tofacitinib in Patients Hospitalized with Covid-19 Pneumonia. N Engl J Med *385*, 406-415.
Han, M.K., Antila, M., Ficker, J.H., Gordeev, I., Guerreros, A., Bernus, A.L., Roquilly, A., Sifuentes-Osornio, J., Tabak, F., Teijeiro, R., et al. (2022). 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.
Han, M.S., Chung, K.W., Cheon, H.G., Rhee, S.D., Yoon, C.H., Lee, M.K., Kim, K.W., and Lee, M.S. (2009). Imatinib mesylate reduces endoplasmic reticulum stress and induces remission of diabetes in db/db mice. Diabetes *58*, 329-336.
Hempel, T., Elez, K., Krüger, N., Raich, L., Shrimp, J.H., Danov, O., Jonigk, D., Braun, A., Shen, M., Hall, M.D., et al. (2021). Synergistic inhibition of SARS-CoV-2 cell entry by otamixaban and covalent protease inhibitors: pre-clinical assessment of pharmacological and molecular properties. Chem Sci *12*, 12600-12609.
Hu, X., Li, J., Fu, M., Zhao, X., and Wang, W. (2021). The JAK/STAT signaling pathway: from bench to clinic. Signal Transduct Target Ther *6*, 402.
Hu, Y., Li, M., Lu, Q., Weng, H., Wang, J., Zekavat, S.M., Yu, Z., Li, B., Gu, J., Muchnik, S., et al. (2019). A statistical framework for cross-tissue transcriptome-wide association analysis. Nature Genetics *51*, 568-576.
Huang, C., Feng, F., Shi, Y., Li, W., Wang, Z., Zhu, Y., Yuan, S., Hu, D., Dai, J., Jiang, Q., et al. (2022). Protein Kinase C Inhibitors Reduce SARS-CoV-2 Replication in Cultured Cells. Microbiol Spectr *10*, e0105622.
Ji, H.L., Song, W., Gao, Z., Su, X.F., Nie, H.G., Jiang, Y., Peng, J.B., He, Y.X., Liao, Y., Zhou, Y.J., et al. (2009). SARS-CoV proteins decrease levels and activity of human ENaC via activation of distinct PKC isoforms. Am J Physiol Lung Cell Mol Physiol *296*, L372-383.
Jiang, Y., Yan, Q., Liu, C.X., Peng, C.W., Zheng, W.J., Zhuang, H.F., Huang, H.T., Liu, Q., Liao, H.L., Zhan, S.F., et al. (2022). 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.
Kaliamurthi, S., Selvaraj, G., Selvaraj, C., Singh, S.K., Wei, D.Q., and Peslherbe, G.H. (2021). Structure-Based

Virtual Screening Reveals Ibrutinib and Zanubrutinib as Potential Repurposed Drugs against COVID-19. Int J Mol Sci *22*.

Koschmieder, S., Jost, E., Cornelissen, C., Müller, T., Schulze-Hagen, M., Bickenbach, J., Marx, G., Kleines, M., Marx, N., Brümmendorf, T.H., et al. (2020). Favorable COVID-19 course despite significant comorbidities in a ruxolitinib-treated patient with primary myelofibrosis. Eur J Haematol *105*, 655-658.

Kousathanas, A., Pairo-Castineira, E., Rawlik, K., Stuckey, A., Odhams, C.A., Walker, S., Russell, C.D., Malinauskas, T., Wu, Y., Millar, J., et al. (2022). Whole-genome sequencing reveals host factors underlying critical COVID-19. Nature *607*, 97-103.

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

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science *313*, 1929-1935.

Lamontagne, F., Agarwal, A., Rochwerg, B., Siemieniuk, R.A., Agoritsas, T., Askie, L., Lytvyn, L., Leo, Y.S., Macdonald, H., Zeng, L., et al. (2020). A living WHO guideline on drugs for covid-19. Bmj *370*, m3379.

Li, B., Clohisey, S.M., Chia, B.S., Wang, B., Cui, A., Eisenhaure, T., Schweitzer, L.D., Hoover, P., Parkinson, N.J., Nachshon, A., et al. (2020). Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. Nat Commun *11*, 164.

Li, T.W., Kenney, A.D., Park, J.G., Fiches, G.N., Liu, H., Zhou, D., Biswas, A., Zhao, W., Que, J., Santoso, N., et al. (2022). SARS-CoV-2 Nsp14 protein associates with IMPDH2 and activates NF-κB signaling. Front Immunol *13*, 1007089.

Mazziotti, G., Maffezzoni, F., Doga, M., Hofbauer, L.C., Adler, R.A., and Giustina, A. (2014). Outcome of glucose homeostasis in patients with glucocorticoid-induced osteoporosis undergoing treatment with bone active-drugs. Bone *67*, 175-180.

Minikel, E.V., Painter, J.L., Dong, C.C., and Nelson, M.R. (2024). Refining the impact of genetic evidence on clinical success. Nature *629*, 624-629.

Nose, A., Mori, Y., Uchiyama-Tanaka, Y., Kishimoto, N., Maruyama, K., Matsubara, H., and Iwasaka, T. (2003). Regulation of glucose transporter (GLUT1) gene expression by angiotensin II in mesangial cells: involvement of HB-EGF and EGF receptor transactivation. Hypertens Res *26*, 67-73.

Pairo-Castineira, E., Clohisey, S., Klaric, L., Bretherick, A.D., Rawlik, K., Pasko, D., Walker, S., Parkinson, N., Fourman, M.H., Russell, C.D., et al. (2021). Genetic mechanisms of critical illness in COVID-19. Nature *591*, 92-98.

Pairo-Castineira, E., Rawlik, K., Bretherick, A.D., Qi, T., Wu, Y., Nassiri, I., McConkey, G.A., Zechner, M., Klaric, L., Griffiths, F., et al. (2023). GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. Nature *617*, 764-768.

Paolini, G.V., Shapland, R.H., van Hoorn, W.P., Mason, J.S., and Hopkins, A.L. (2006). Global mapping of pharmacological space. Nat Biotechnol *24*, 805-815.

Porcu, E., Sadler, M.C., Lepik, K., Auwerx, C., Wood, A.R., Weihs, A., Sleiman, M.S.B., Ribeiro, D.M., Bandinelli, S., Tanaka, T., et al. (2021). Differentially expressed genes reflect disease-induced rather than disease-causing changes in the transcriptome. Nat Commun *12*, 5647.

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

Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J.,

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

Ramnath, R., Foster, R.R., Qiu, Y., Cope, G., Butler, M.J., Salmon, A.H., Mathieson, P.W., Coward, R.J., Welsh, G.I., and Satchell, S.C. (2014). Matrix metalloproteinase 9-mediated shedding of syndecan 4 in response to tumor necrosis factor α: a contributor to endothelial cell glycocalyx dysfunction. Faseb j *28*, 4686-4699.

Ray, J.G., Schull, M.J., Vermeulen, M.J., and Park, A.L. (2021). 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.

Rein, L., Calero, K., Shah, R., Ojielo, C., Hudock, K.M., Lodhi, S., Sadaka, F., Bellam, S., Palma, C., Hager, D.N., et al. (2022). Randomized Phase 3 Trial of Ruxolitinib for COVID-19-Associated Acute Respiratory Distress Syndrome. Crit Care Med *50*, 1701-1713.

Rubin, R. (2022). Baricitinib Is First Approved COVID-19 Immunomodulatory Treatment. Jama *327*, 2281.

Sajid Jamal, Q.M., Alharbi, A.H., and Ahmad, V. (2022). 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.

Sass, F., Ma, T., Ekberg, J.H., Kirigiti, M., Ureña, M.G., Dollet, L., Brown, J.M., Basse, A.L., Yacawych, W.T., Burm, H.B., et al. (2024). NK2R control of energy expenditure and feeding to treat metabolic diseases. Nature *635*, 987-1000.

Saul, S., Karim, M., Ghita, L., Huang, P.T., Chiu, W., Durán, V., Lo, C.W., Kumar, S., Bhalla, N., Leyssen, P., et al. (2023). Anticancer pan-ErbB inhibitors reduce inflammation and tissue injury and exert broad-spectrum antiviral effects. J Clin Invest.

Sharma, V., Dhillon, P., Wambolt, R., Parsons, H., Brownsey, R., Allard, M.F., and McNeill, J.H. (2008). Metoprolol improves cardiac function and modulates cardiac metabolism in the streptozotocin-diabetic rat. Am J Physiol Heart Circ Physiol *294*, H1609-1620.

Sinha, S., Cheng, K., Schäffer, A.A., Aldape, K., Schiff, E., and Ruppin, E. (2020). In vitro and in vivo identification of clinically approved drugs that modify ACE2 expression. Mol Syst Biol *16*, e9628.

Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell *171*, 1437-1452.e1417.

Sun, X.J., Miralpeix, M., Myers, M.G., Jr., Glasheen, E.M., Backer, J.M., Kahn, C.R., and White, M.F. (1992). Expression and function of IRS-1 in insulin signal transmission. J Biol Chem *267*, 22662-22672.

Tomazou, M., Bourdakou, M.M., Minadakis, G., Zachariou, M., Oulas, A., Karatzas, E., Loizidou, E.M., Kakouri, A.C., Christodoulou, C.C., Savva, K., et al. (2021). 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*.

van de Veerdonk, F.L., Giamarellos-Bourboulis, E., Pickkers, P., Derde, L., Leavis, H., van Crevel, R., Engel, J.J., Wiersinga, W.J., Vlaar, A.P.J., Shankar-Hari, M., et al. (2022). A guide to immunotherapy for COVID-19. Nat Med *28*, 39-50.

Vanderbeke, L., Van Mol, P., Van Herck, Y., De Smet, F., Humblet-Baron, S., Martinod, K., Antoranz, A., Arijs, I., Boeckx, B., Bosisio, F.M., et al. (2021). Monocyte-driven atypical cytokine storm and aberrant neutrophil activation as key mediators of COVID-19 disease severity. Nat Commun *12*, 4117.

Vannucchi, A.M., Sordi, B., Morettini, A., Nozzoli, C., Poggesi, L., Pieralli, F., Bartoloni, A., Atanasio, A., Miselli, F., Paoli, C., et al. (2021). Compassionate use of JAK1/2 inhibitor ruxolitinib for severe COVID-19: a prospective observational study. Leukemia *35*, 1121-1133.

Veneri, D., Franchini, M., and Bonora, E. (2005). Imatinib and regression of type 2 diabetes. N Engl J Med *352*, 1049-1050.

von Metzler, I., Campe, J., Huenecke, S., Raab, M.S., Goldschmidt, H., Schubert, R., Rabenau, H.F., Ciesek, S., Serve, H., and Ullrich, E. (2022). 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.

Wagner, A., Cohen, N., Kelder, T., Amit, U., Liebman, E., Steinberg, D.M., Radonjic, M., and Ruppin, E. (2015). Drugs that reverse disease transcriptomic signatures are more effective in a mouse model of dyslipidemia. Mol Syst Biol *11*, 791.

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

Wang, L., Balmat, T.J., Antonia, A.L., Constantine, F.J., Henao, R., Burke, T.W., Ingham, A., McClain, M.T., Tsalik, E.L., Ko, E.R., et al. (2021a). An atlas connecting shared genetic architecture of human diseases and molecular phenotypes provides insight into COVID-19 susceptibility. Genome Med *13*, 83.

Wang, W., Li, W., Wen, Z., Wang, C., Liu, W., Zhang, Y., Liu, J., Ding, T., Shuai, L., Zhong, G., et al. (2022). Gossypol Broadly Inhibits Coronaviruses by Targeting RNA-Dependent RNA Polymerases. Adv Sci (Weinh) *9*, e2203499.

Wang, X., Lu, J., Ge, S., Hou, Y., Hu, T., Lv, Y., Wang, C., and He, H. (2021b). Astemizole as a drug to inhibit the effect of SARS-COV-2 in vitro. Microb Pathog *156*, 104929.

Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., et al. (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res *46*, D1074-d1082.

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

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

Wu, P., Feng, Q., Kerchberger, V.E., Nelson, S.D., Chen, Q., Li, B., Edwards, T.L., Cox, N.J., Phillips, E.J., Stein, C.M., et al. (2022). Integrating gene expression and clinical data to identify drug repurposing candidates for hyperlipidemia and hypertension. Nat Commun *13*, 46.

Xie, J., Klemsz, M.J., Kacena, M.A., Sandusky, G., Zhang, X., and Kaplan, M.H. (2022). Inhibition of MEK signaling prevents SARS-CoV2-induced lung damage and improves the survival of infected mice. J Med Virol *94*, 6097-6102.

Xue, A., Wu, Y., Zhu, Z., Zhang, F., Kemper, K.E., Zheng, Z., Yengo, L., Lloyd-Jones, L.R., Sidorenko, J., Wu, Y., et al. (2018). Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun *9*, 2941.

Yao, M., Huang, X., Guo, Y., Zhao, J.V., and Liu, Z. (2023). 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.

Yao, Y., Ye, F., Li, K., Xu, P., Tan, W., Feng, Q., and Rao, S. (2021). 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.

Yassin, A., Sabsigh, R., Al-Zoubi, R.M., Aboumarzouk, O.M., Alwani, M., Nettleship, J., and Kelly, D. (2023). Testosterone and Covid-19: An update. Rev Med Virol *33*, e2395.

Ye, H., Yan, J., Wang, Q., Tian, H., and Zhou, L. (2022). The protective effects of cabozantinib against high glucose-induced damages in in vitro renal glomerular endothelial cells model via inhibition of early growth

response-1 (Egr-1). Bioengineered *13*, 10605-10616.
Zhou, D., Jiang, Y., Zhong, X., Cox, N.J., Liu, C., and Gamazon, E.R. (2020). A unified framework for joint-tissue transcriptome-wide association and Mendelian randomization analysis. Nat Genet *52*, 1239-1246.
Zhou, Y., Zhang, Y., Lian, X., Li, F., Wang, C., Zhu, F., Qiu, Y., and Chen, Y. (2022). Therapeutic target database update 2022: facilitating drug discovery with enriched comparative data of targeted agents. Nucleic Acids Res *50*, D1398-d1407.
Zhu, Z., Chen, X., Wang, C., Zhang, S., Yu, R., Xie, Y., Yuan, S., Cheng, L., Shi, L., and Zhang, X. (2023). An integrated strategy to identify COVID-19 causal genes and characteristics represented by LRRC37A2. J Med Virol *95*, e28585.