

# Drug Repositioning by Merging Active Subnetworks Validated in Cancer and COVID-19

M. Lucchetta\*      M. Pellegrini†

May 11, 2021

## Abstract

Computational Drug Repositioning aims at ranking and selecting existing drugs for use in novel diseases or existing diseases for which these drugs were not originally designed. Using vast amounts of available omic data in digital form within an *in silico* screening has the potential for speeding up considerably the shortlisting of promising candidates in response to outbreaks of diseases such as COVID-19 for which no satisfactory cure has yet been found. We describe DrugMerge as a methodology for preclinical computational drug repositioning based on merging multiple drug rankings obtained with an ensemble of Disease Active Subnetwork construction algorithms. DrugMerge uses differential transcriptomic data from cell lines/tissues of patients affected by the disease and differential transcriptomic data from drug perturbation assays, in the context of a large gene co-expression network. Experiments with four benchmark diseases (Asthma, Rheumatoid Arthritis, Prostate Cancer, and Colorectal Cancer) demonstrate that our method detects in first position drugs in clinical use for the specified disease, in all four cases. Our method is competitive with the state-of-the-art tools such as CMAP (Connectivity Map). Application of DrugMerge to COVID-19 data found rankings with many drugs currently in clinical trials for COVID-19 in top positions, thus showing that DrugMerge is able to mimic human expert judgment.

## 1 Introduction

Among the strategies pharmacological science adopts for tackling diseases (either new or old), there are: the development of new drugs, the development of vaccines, and the repurposing of existing drugs. While the development of vaccines and novel drugs are often effective, they are also expensive and time-consuming. Thus repurposing of existing drugs is a cheaper and faster route to explore when one wishes to hedge one's bets [Ashburn and Thor, 2004],[Nosengo, 2016].

As data about drugs and diseases accumulate in databases comprising many aspects: genomic, transcriptomic, phenotypic, clinical, and epidemiological, it becomes feasible and desirable to employ this body of accumulated knowledge within a computerized system so as to have a pre-screen *in silico* of the candidate drugs to be repurposed for a specific disease [Sam and Athri, 2019]. These shortlisted drugs are in reduced number with respect to the initial pool of drugs and thus are more amenable to the subsequent steps of the full pipeline (including *in vitro* and *in vivo* experiments, clinical trials, and eventual deployment in clinical practice). For existing drugs often their safety profile for humans is already known, therefore once the initial *in vitro* and *in vivo* assays are positive, they can be moved directly to clinical trials of phase II or phase III (thus skipping phases 0 and I), and accelerating the drug approval lifecycle.

---

\*IIT-CNR, Pisa, Italy, and Dipartimento di Biotecnologie, Chimica e Farmacia, Università degli Studi di Siena, Siena, Italy

†IIT-CNR, Pisa, Italy. Contact author [marco.pellegrini@iit.cnr.it](mailto:marco.pellegrini@iit.cnr.it)

The recent pandemic sparked by the SARS-CoV2 virus has placed drug repurposing in the spotlight because of the pressure on the public health systems to find cures for acute cases of patients infected by SARS-CoV2. At the moment of writing, remdesivir has been approved by regulatory bodies to treat acute COVID-19 cases needing hospitalization [Rubin et al., 2020]. No other drug or treatment has passed all phases of clinical trials leading to full approval by the Federal Drug Administration (FDA) or the European Medicines Agency (EMA).

Initial drug investigations on COVID-19 were based on the accumulated experience of the efficacy of drugs *in vitro* and *in vivo* experiments for SARS-Cov, MERS, and other coronaviruses. Computerized drug repurposing however aims at a more systematic approach to candidate drug ranking and selection. Moreover, while having accrued data from other viruses belonging to the same family is a bonus, we should not always rely on having such data handy.

In a longer perspective, both the possible variations in the behavior of SARS-CoV2, due to genetic mutations, or the emergence of new human viral diseases due to cross-species transmission, imply that an easy to use, accurate, and fast pre-screening and ranking of repurposable drugs aimed at countering any new viral threats is needed. Our study is a contribution in this direction.

Network-based drug repurposing is a recent new approach to drug repurposing that may increase the effectiveness of drug repurposing pipelines. In particular network-based drug repurposing can merge many 'omics' data sets in a unified framework, thus such approach may be more robust and capable of including systemic biological effects [Barabasi and Oltvai, 2004], [Barabási et al., 2011], [Cheng et al., 2019], and [Gysi et al., 2020].

In [Lucchetta, 2020] we established Core&Peel as a competitive method for building disease active subnetworks compared to several state-of-the-art algorithms in several known benchmark diseases. Moreover, we applied Core&Peel and the other algorithms to build a collection of active subnetworks from gene expression data of human tissues infected by COVID-19. Subsequent pathway enrichment analysis revealed that in a COVID-19 infection are activated pathways similar to those activated in other diseases (mostly viral infections). These observations had implications for the choice of existing drugs likely to have an impact on COVID-19 to be short-listed for *in vitro* experiments and eventually go on to clinical trials. Although often the connection between a disease-related pathway and a repurposable drug was obvious in the specific context, sometimes the connection was hard to establish since, for example, for a relevant disease-related pathway too many drugs are known to be effective. Moreover, there was no quantitative way to prioritize the potentially repurposable drugs obtained from this type of qualitative analysis.

In this paper, building upon [Lucchetta, 2020], we describe a systematic way to rank repurposable drugs for a specific disease, so as to take advantage of multiple active subnetwork algorithms and multiple gene expression data sets. We use known clinically approved drugs for benchmark diseases, and trial drugs for COVID-19 as golden standard. We show that our method has better measurable performance than C-map [Lamb et al., 2006], a widely used drug ranking portal, as well as COVID-19 specific drug rankings in [Taguchi and Turki, 2020] and [Mousavi et al., 2020]. The quality of our ranking results on COVID-19 is arguably comparable to those in [Gysi et al., 2020] and [Zhou et al., 2020b], which use however quite different strategies. The whole workflow of our method is shown in Figure 1.

Ruiz et al. [Ruiz et al., 2020] observe that "...a drug's effectiveness can often be attributed to targeting genes that are distinct from disease-associated genes but that affect the same functional pathways.", in contrast to "... existing approaches assume that, for a drug to treat a disease, the proteins targeted by the drug need to be close to or even need to coincide with the disease-perturbed proteins." Ruiz et al. thus build a multi-layered interaction graph comprising proteins, drugs, diseases, and functional pathways (extracted from the GO repository), and devise a random-walk diffusion algorithm to connect the drug nodes and the disease nodes of the model through paths that include the functional nodes of the graph.

Our approach is intermediate between the functional-based method advocated by Ruiz et al. [Ruiz et al., 2020] and protein proximity based methods (e.g. [Gysi et al., 2020]).

We do not represent GO pathways as functional nodes in the model, instead, we rely on a large and comprehensive gene co-expression network to provide the functional association in an implicit way. Dense subgraphs of the co-expression network will be formed by genes that have a coherent pattern of being up/down expressed, and thus are likely to be involved in specific cellular processes, either directly, by expressing binding proteins, or indirectly through cascading regulatory action chains, possibly involving also ncRNA and other regulatory elements not explicitly modelled. All algorithms used in this paper to build Active Network will take advantage of the dense neighborhood of the co-expression network, either directly (e.g. Core&Peel, ModuleDiscoverer) or indirectly via the augmented local graph connectivity (KPM, Degas, ClustEx). In contrast with the proximity-based methods, we do not define any ad hoc explicit distance function over a network, since the Active Network algorithms already provide a clear-cut demarcation for the part of the global gene network onto which a perturbation should act, thus we can use directly indices for measuring module enrichment or incidence of the drug's DEG onto the disease's Active Network.

We follow a network-based approach to drug repositioning and we focus mainly on the human transcriptional response to the disease, rather than the interaction between the disease agent (e.g. a virus) and the transcriptional response to it. In the case of COVID-19, also it is known that an over-response by the immune system is often the main cause of death in acute cases [Moore and June, 2020], [Laing et al., 2020], [Li et al., 2020] [de la Rica et al., 2020], [Tang et al., 2020], [Cron, 2020], [Hojyo et al., 2020], [Dorward et al., 2020], [Merad and Martin, 2020], and [Fajgenbaum and June, 2020].

Our strategy for COVID-19 is thus to prioritize drugs that act on a disease active subnetwork embedded into a global human gene co-expression network, to counter the virus' global effects. This approach is complementary to a different one that aims at drugs interfering with the interactions of the host-proteins with the viral-protein. While the former approach aims at controlling the global transcriptional response of human tissues to the virus, the latter favors drugs blocking key aspects of the virus actions in entering the host cells, and activating the cells' biochemical machinery for reproduction.

This paper is organized as follows. Section 2 reports the main results of our study on ranking for drug repositioning in four benchmark diseases and COVID-19. Section 3 describes in detail the algorithmic pipeline for the proposed DrugMerge method. Section 4 covers several strong points, current limitations, and perspective improvements of DrugMerge. Section 5 briefly gives an overview of the literature on computational drug repositioning, with more in-depth descriptions of previous methods more relevant for the development of DrugMerge. In Section 6 we make final comments on the potential of computational drug repositioning for addressing future health needs. In appendix A we give a brief synopsis of the supplementary tables and files.

## 2 Results

We compute as main quality measures the reciprocal hit ranking (RHR) and the precision at 20 (precision@20) of a ranking with respect to a golden standard (the TTD records for four benchmark diseases, and the trials.gov records for COVID-19). Each measure is then normalized via a z-score, and corresponding p-value, referred to the distribution of quality measures given by taking repeatedly random permutations of the input drug data set.

The results are shown in Figures from 2 to 9 and Supplementary Tables from S1 to S8. For DrugMerge we report only significant results at p-value  $\leq 0.05$  for at least one of the two quality measures. The DrugMerge method could attain significant results for all four benchmark diseases (asthma, rheumatoid arthritis, colorectal cancer, and prostate cancer) on at least one Drug Database (GEO, DrugMatrix, CMAP, L1000), and often on more than one. Moreover, as reported in the actual rankings in Supplementary Table S7, for the four benchmark data sets the reported drug ranking always has a drug in clinical use at the top of the list for the chosen Drug data set (attaining the

best p-value). In Figure 6 we compare, using the p-value, the performance of DrugMerge versus the CMAP algorithm [Lamb et al., 2006] on the four benchmark diseases. DrugMerge has always a better performance in terms of p-value, even when the same CMAP drug database is used also by DrugMerge. Next we discuss in detail results for single diseases.

## 2.1 DrugMerge results on Asthma

Figure 2 and Supplementary Table S1 show that DrugMerge on the GEO Drug dataset finds one clinically relevant drug (prednisolone) in first position both when all drugs are used and when only FDA approved drugs are used. The precision@20 is not statistically significant, but the RHR of this single hit is very significant. A closer look at the GEO data and at the TTD records reveals that prednisolone is the only drug from TTD present in the GEO dataset. The initial list of differentially expressed genes is measured from a cohort of patients with acute asthma with respect to healthy patients, and interestingly the hit drug prednisolone is used specifically for acute asthma<sup>1</sup>. The DrugMerge algorithm does not find any hit in the L1000 drug dataset. An examination of the L1000 data shows that only 5 asthma-related drugs as reported by TTD are in the L1000 data (budesonide, fluticasone propionate, mometasone furoate, beclomethasone, and colforsin) which are used to treat chronic asthma or mild/moderate persistent asthma. Similar considerations hold for the TTD drugs present in CMAP Drug data. Thus DrugMerge appears to be able to differentiate the acute vs chronic forms of the asthma disease and rank higher drugs that are fit for the subclass of asthma patients assessed in the data set (acute asthma patients).

Among the top positions in the ranking reported in Supplementary File S7, we find drugs with known effects in asthma patients or animal models of asthma, and one of these went into clinical trial stage: deferasirox [Sompornrattanaphan et al., 2020], rosiglitazone [Lee et al., 2016], valproic acid<sup>2</sup>, dexamethasone [Sarkari, 2007], celecoxib [Martin-Garcia et al., 2003], and tamoxifen [Mansilla et al., 2020].

Rosiglitazone has been tested on a murine model of chronic asthma [Lee et al., 2016], suggesting that its intranasal administration can prevent air way inflammation.

A comparison study between dexamethasone versus hydrocortisone in severe acute pediatric asthma [Sarkari, 2007] showed that the mean length of hospitalization in children receiving dexamethasone was significantly shorter than those receiving hydrocortisone. Celecoxib is a COX-2 inhibitor and nonsteroidal anti-inflammatory drug.

A study of 33 asthma patients [Martin-Garcia et al., 2003] demonstrated that celecoxib is a suitable drug in aspirin-induced and/or nonsteroidal anti-inflammatory drug-induced asthma patients.

Tamoxifen is an estrogen receptor modulator mainly used for the treatment of breast cancer, but it has also been reported to have anti-inflammatory activity. A recent study [Mansilla et al., 2020] has tested tamoxifen on three different mouse models, showing that tamoxifen can reduce inflammatory infiltration of neutrophils in the airways.

For the remaining drugs in the top positions, we focus on their mechanism of action and we search in literature if this can have relevance in asthma. Using this criterion, we find captopril, letrozole, decitabine, imatinib, and probucol as relevant drugs.

Captopril belongs to the class of drugs known as angiotensin-converting enzyme (ACE) inhibitors and is used primarily to lower high blood pressure (hypertension). It has been proved that asthmatic subjects with comorbid hypertension display evidence of enhanced asthma morbidity [Christiansen et al., 2016].

Letrozole inhibits aromatase, an enzyme that catalyzes the synthesis of estrogen. Asthma prevalence and severity are greater in women than in men, suggesting this is in part related to female steroid sex hormones, such as estrogen. Estrogen receptors are found on numerous immunoregulatory cells and estrogen's actions skew immune responses toward allergy [Bonds and Midoro-Horiuti, 2013].

<sup>1</sup><https://www.mayoclinic.org/diseases-conditions/asthma/in-depth/asthma-medications/art-20045557>

<sup>2</sup>In clinical trial <https://clinicaltrials.gov/ct2/show/NCT00153270>

Decitabine is a hypomethylating agent and is used to treat myelodysplastic syndromes and acute myeloid leukemia. Yang et al. [Yang et al., 2015] demonstrated that DNA methylation in specific gene loci are associated with asthma and suggest that epigenetic changes might play a role in establishing the immune phenotype associated with childhood asthma.

Imatinib is a tyrosine kinase inhibitor. Anti-inflammatory effects of tyrosine kinase inhibitors have been reported in animal models of allergic asthma, suggesting that this kind of drug can be a very attractive strategy for the treatment of asthma [Wong and Leong, 2004].

Finally, probucol lowers the level of cholesterol through the inhibition of cholesterol synthesis and deletion of cholesterol absorption. Ramaraju et al. [Ramaraju et al., 2013] found a modest but significant association between higher levels of serum cholesterol and asthma.

## 2.2 DrugMerge results on Rheumatoid Arthritis (RA)

Figures 3 and Supplementary Table S2 show that DrugMerge finds significant rankings in several drug databases (L1000, DrugMatrix, GEO), with GEO attaining the best p-value, using several algorithms, both when all drugs are considered or when only FDA-approved drugs are considered. For the GEO dataset Supplementary Table S7, we find in first position the clinically relevant drug dexamethasone. Other drugs in use according to the TTD records that appear in the top twenty positions are: imatinib, methylprednisolone, and prednisolone.

Among the top positions in the ranking, we find drugs with known effects in rheumatoid arthritis patients or animal models of rheumatoid arthritis, and one of these went into clinical trial stage: rosiglitazone<sup>3</sup>, pioglitazone [Suke et al., 2013], estradiol [Romo-García et al., 2020], captopril [Martin et al., 1984], vitamin c [Carr et al., 2015], baclofen [Huang et al., 2015], deferasirox [Polson et al., 1986], decitabine [Petralia et al., 2019], sirolimus [Wen et al., 2019], hydrocortisone [Boland, 1952], paclitaxel [Kurose et al., 2001], and ethinylestradiol [Subramanian et al., 2005].

Suke et al. [Suke et al., 2013] studied the effects of combined pioglitazone and prednisolone on adjuvant-induced arthritis in rats. This study suggested that the combination of these two drugs was effective in modulating the inflammatory response and suppress arthritis progression.

Estradiol is an estrogen steroid female hormone, and estrogens have a direct action upon the immune system. The role of the estrogens in rheumatoid arthritis has been studied here [Romo-García et al., 2020].

Captopril is prescribed for hypertension but it has immunosuppressant activity, as well. Therefore, captopril was considered a potential slow-acting drug for treating rheumatoid arthritis as demonstrated in [Martin et al., 1984].

Vitamin C is a vitamin found in various foods and it is important for immune system function. It has also been studied the role of the vitamin c in treating pain [Carr and McCall, 2017], in particular, an administration of high-dose vitamin C in patients with rheumatoid arthritis showed a complete decrease in pain [Carr et al., 2015].

Baclofen is a medication used to treat muscle spasticity. Huang et al. [Huang et al., 2015] investigated the effects of baclofen in murine collagen-induced arthritis, proving that baclofen alleviated the clinical development of arthritis.

Decitabine has already been found in asthma analysis and as mentioned before, it inhibits DNA methylation. Petralia et al. [Petralia et al., 2019] have studied the effect of decitabine in a murine model of rheumatoid arthritis and have demonstrated that decitabine administration was associated with a significant improvement of the clinical condition.

Sirolimus, also known as rapamycin, has immunosuppressant functions and is used to prevent rejection in organ transplants. Wen et al. [Wen et al., 2019] studied the safety, tolerance, and efficacy of sirolimus in patients with active RA treated with low-dose sirolimus combined with original therapy. They showed that this therapy alleviates clinical symptoms and decreases the immunosuppressive applications in patients with active RA.

<sup>3</sup>In clinical trial <https://clinicaltrials.gov/ct2/show/NCT00379600>

Hydrocortisone is a treatment for acute episodes of rheumatic disorders, including rheumatoid arthritis <sup>4</sup>.

Paclitaxel is an anticancer agent and is classified as a microtubule-stabilizing agent. Kurose et al. [Kurose et al., 2001] studied the effects of paclitaxel on cultured synovial cells from patients with rheumatoid arthritis. The data suggest paclitaxel as a possible therapy for RA.

Ethinylestradiol is an active estrogen and component of birth control pills. Subramaniam et al. [Subramaniam et al., 2005] studied the effectiveness of ethinylestradiol in treating collagen-induced arthritis mice, noticing a decreased proliferation and secretion of pro-inflammatory factors.

Focusing on the action mechanism of the remaining drugs, we find some links with rheumatoid arthritis, such as amoxicillin [Ogrendik, 2014], and niacin [Peclat et al., 2020].

Amoxicillin is an antibiotic used to treat several bacterial infections. Since the 1930s, RA has been treated with antibiotics and there have been several reports in the literature indicating that periodontal pathogens are a possible cause of the disease [Ogrendik, 2014].

Niacin, more commonly known as vitamin B3, is a precursor of the coenzymes nicotinamide-adenine dinucleotide (NAD<sup>+</sup>). Recent studies have identified potential therapeutic approaches for boosting NAD<sup>+</sup> to treat rheumatologic diseases, including rheumatoid arthritis. In particular, they focused on the enzymatic activity of CD38, one of the main enzymes in NAD<sup>+</sup> catabolism [Peclat et al., 2020].

### 2.3 DrugMerge results on Colorectal Cancer

Figures 4 and Supplementary Table S3 show that DrugMerge finds significant rankings in several drug databases (L1000, CMAP, GEO), using several algorithms, both when all drugs are considered or when only FDA-approved drugs are considered. For the L1000 dataset Supplementary Table S7, we find in the first position the clinically relevant drug erlotinib. Other drugs in clinical use according to the TTD records that appear in the the top twenty positions are: lapatinib and sunitinib. Among the top positions in the ranking, we find drugs with known effects in colorectal cancer patients, related animal models, or cell lines, and one of these went into clinical trial stage: mitoxantrone [Cornelia et al., 1991], sorafenib [Kacan et al., 2016], sutent (sunitinib) <sup>5</sup>, dasatinib [Scott et al., 2017], sirolimus (rapamycin) [Wagner et al., 2009], azacitidine [Azad et al., 2017], imatinib [Samei et al., 2016], tamoxifen [Kuruppu et al., 1998], doxorubicin [Lee et al., 2017], naltrexone [Ma et al., 2020], and digoxin [Hou et al., 2020].

Mitoxantrone has been tested in 13 patients with advanced or metastatic colon carcinoma and the authors observed that mitoxantrone was moderately effective [Cornelia et al., 1991].

Sorafenib is a kinase inhibitor drug approved for the treatment of some cancers. In [Kacan et al., 2016], the authors showed that sorafenib has a strong antitumor and antiangiogenic effect on the colorectal cancer cell line.

Scott et al. [Scott et al., 2017] demonstrated that dasatinib had significant anti-proliferative activity in a subset of a colorectal cancer cell line, even if dasatinib is currently being studied in combination with chemotherapy, as its use as a single agent appears limited.

Wagner et al. [Wagner et al., 2009] studied the effects of rapamycin in mice and showed that it was able to suppress advanced stage colorectal cancer.

Azacitidine has been tested in combination with entinostat in metastatic colorectal cancer patients, resulting in a tolerable therapy [Azad et al., 2017].

Imatinib is an oral chemotherapy medication used to treat cancer and it has been proved that it inhibits proliferation of colon cancer cells [Samei et al., 2016].

Kuruppu et al [Kuruppu et al., 1998] demonstrated that tamoxifen has a potent inhibitory action on colorectal cancer metastases in a murine model.

<sup>4</sup><https://www.drugs.com/monograph/hydrocortisone-systemic.html#ra>

<sup>5</sup>In clinical trial <https://clinicaltrials.gov/ct2/show/NCT00961571>

Doxorubicin is a chemotherapy medication used to treat cancer. In [Lee et al., 2017] the authors propose doxorubicin-loaded oligonucleotides attached to gold nanoparticles as a drug delivery system for cancer chemotherapy. They tested this approach in colon cancer cell lines, demonstrating the drug's efficacies such as *in vitro* cytotoxicity.

Naltrexone is primarily used to manage alcohol or opioid dependence. Ma et al. [Ma et al., 2020] explored the inhibitory effect of low-dose naltrexone on colorectal cancer progression *in vivo* and *in vitro*.

Digoxin is used to treat various heart conditions such as atrial fibrillation, and heart failure. A recent study [Hou et al., 2020] has examined the effects of digoxin on two kinds of colorectal cancer cells. Their findings suggest that digoxin has the potential to be applied as an antitumor drug via inhibiting proliferation and metastasis.

For the remaining drugs in the top twenty positions, we focus on their action mechanism, and we find some links with colorectal cancer: epirubicin [Coss et al., 2009], everolimus [Thiem et al., 2013], itavastatin [Li et al., 2021], and dactinomycin [Xu et al., 2020].

Epirubicin is a chemotherapy agent, intercalating into DNA and inhibiting topoisomerase II, thereby inhibiting DNA replication. In [Coss et al., 2009] the authors demonstrate topoisomerase IIa gene and protein alterations in colorectal cancer and show that increased expression is associated with pathologically advanced disease.

Everolimus is a medication used as an immunosuppressant and in the treatment of several tumors. Its main mechanism of action is the mTOR inhibition. Thiem et al. [Thiem et al., 2013] demonstrated that the mTORC1 inhibition reduces the inflammation associated with the gastrointestinal tumor. Itavastatin (or pitavastatin) belongs to the class of statins, which are inhibitors of HMG-CoA reductase, the enzyme that catalyzes the first step of cholesterol synthesis. A meta-analysis work [Li et al., 2021] showed that statin use is a protective factor for colorectal cancer prognosis, and it might be significantly associated with lower overall mortality.

Finally, dactinomycin is a chemotherapy medication and can inhibit transcription. Recent studies have revealed crucial roles of transcription regulation in colorectal cancer development. In particular, deregulation of transcription factors is pretty frequent, and drastic changes in gene expression profiles play fundamental roles in the multistep process of tumorigenesis [Xu et al., 2020].

## 2.4 DrugMerge results on Prostate Cancer

Figures 5 and Supplementary Table S4 show that DrugMerge finds significant rankings in several drug databases (L1000, GEO) for prostate cancer, using several algorithms, both when all drugs are considered or when only FDA approved drugs are considered. For the L1000 data set Supplementary Table S7, we find in the first position the clinically relevant drug mitoxantrone. Other drugs in clinical use according to the TTD records that appear in the top twenty positions are: dasatinib, decitabine, menadione, and doxorubicin hydrochloride.

Among the top twenty positions in the ranking we find drugs with known effects in prostate cancer patients, related animal models, or cell lines, and one of these went into clinical trial stage: erlotinib [Gravis et al., 2008], sorafenib [Zaborowska et al., 2012], lapatinib [Whang et al., 2013], sirolimus [Nandi et al., 2021], sunitinib [Michaelson et al., 2009], azacitidine [Singal et al., 2015], everolimus [Morgan et al., 2008], gemcitabine<sup>6</sup>, ciclosporine [Kawahara et al., 2015], and auranofin [Liu et al., 2019].

A study of 30 patients with advanced or metastatic prostate cancer has been conducted to test the effect of erlotinib [Gravis et al., 2008], and it demonstrated an improvement in clinical benefit. Sorafenib has been tested in patients with hormone-refractory prostate cancer [Zaborowska et al., 2012], indicating that the treatment may be associated with good outcomes in terms of overall survival.

<sup>6</sup>In clinical trial <https://clinicaltrials.gov/ct2/show/NCT00276549>

Lapatinib is a dual tyrosine kinase inhibitor of the epidermal growth factor 1 (EGFR) and 2 (HER2). In a study of twenty-nine castration-resistant prostate cancer patients [Whang et al., 2013], lapatinib showed single-agent activity in a small subset of unselected patients.

Nandi et al. [Nandi et al., 2021] investigated the effect of sirolimus encapsulated liposomes in two different prostate cancer cell lines, and noticed an antiproliferative effect.

Sunitinib, an inhibitor of tyrosine kinase receptors, has been studied in men with castration-resistant prostate cancer [Michaelson et al., 2009], indicating that it is well tolerated and may have modest benefit in this patient population.

Azacitidine, a demethylating agent, has been tested in combination with docetaxel, and prednisone in patients with metastatic castration-resistant prostate cancer and showed to be active in this subset of population [Singal et al., 2015].

Everolimus, an mTOR inhibitor, has been assessed in mice, resulting in an inhibition of prostate cancer growth [Morgan et al., 2008].

For the remaining drugs, we focus on their action mechanism, and we find some links with prostate cancer for imatinib [Vicentini et al., 2003], epirubicin [Patra et al., 2011], and mitomycin C [Gurumurthy et al., 2001].

Imatinib is an oral chemotherapy medication and functions as a tyrosine kinase inhibitor. Vicentini et al. [Vicentini et al., 2003] investigated the effects of the epidermal growth factor receptor tyrosine kinase inhibitor on human prostatic cancer cell lines, suggesting that it may have the potential in blocking tumor growth and progression.

Epirubicin has been already found in colorectal cancer and as mentioned before, it is a topoisomerase II inhibitor. An *in vitro* cell-based assay [Patra et al., 2011] demonstrated that MHY336, a topoisomerase II inhibitor, significantly inhibited the proliferation of three prostate cancer cell lines. Finally, mitomycin C is used as a chemotherapeutic agent and it is a potent DNA crosslinker. DNA crosslinking has useful merit in chemotherapy and targeting cancerous cells for apoptosis [Deans and West, 2011]. It is known that evading cell death is one hallmark of cancer, in particular, prostate cancer cells develop multiple apoptosis blocking strategies during the various stages of tumor progression [Gurumurthy et al., 2001].

## 2.5 DrugMerge results on COVID-19

Figures 7 and 8 and Supplementary Table S5 show that DrugMerge finds significant rankings in all drug databases (L1000, CMAP, GEO, DrugMatrix), using a variety of algorithms and their combinations, both when all drugs are considered and when only FDA approved drugs are considered. For COVID-19 we have a choice of three disease data sets (cells, BALF, PBMC), four Drug datasets (L1000, CMAP, GEO, and DrugMatrix), and several algorithms and weighting functions. We proceed by first running a single algorithm on a single drug data set, collecting the configurations that give the best significant performance measures. Next, we seek improvements by merging the rankings for several algorithms with significant performance on a single disease data set. And finally, we combine the improved significant results by merging the rankings obtained by several significant algorithms on several COVID-19 disease data sets. Figures 7 and 8 and Supplementary Table S5 report the best results in each phase, while Table S5.1 report the full algorithmic settings.

We focus now on two such rankings in Supplementary file S7 attaining the best p-values on RHR for all drugs or FDA-approved drugs only.

The highest p-value on RHR when only FDA-approved drugs are considered is attained on the CMAP Drug dataset using all three disease data sets for COVID-19 and three algorithms (ClustEx, Core&Peel, Degas), (Alg 01 in Figure 7, see Supplementary Table S5.1). The top two positions of the ranking report two drugs listed in clinical trial on [www.trials.gov](http://www.trials.gov) for COVID-19: etoposide and mefloquine. Within the first twenty positions, further 5 drugs are in clinical trials for COVID-19: colchicine, chlorpromazine, propofol, tretinoin, and ivermectin. Among the top ten positions in the ranking we find drugs with known effects animal models or cell lines or proposed for further scrutiny



*in silico* screenings: thioridazine [Otręba et al., 2020], prochlorperazine [Kow and Hasan, 2020], primaquine [Sachdeva et al., 2020], fluphenazine [Nazeam et al., 2020].

The highest p-value on RHR when all drugs are considered is attained by KPM on the L1000 drug data set with the BALF COVID-19 disease data set (Alg 17 in Figure 8, see Suppl. Table S5.1). The top position of the ranking reports a drug listed in clinical trial on [www.trials.gov](http://www.trials.gov) for COVID-19: sirolimus. Within the first twenty positions further 4 drugs are in clinical trials for covid19: decitabine, tamoxifen, simvastatin, and imatinib. Among the top ten positions in the ranking we find drugs with known effects in animal models or cell lines or proposed for further scrutiny in *in silico* screenings: lapatinib [Raymonda et al., 2020], dasatinib [Weisberg et al., 2020], mitoxantrone [Lokhande et al., 2020], everolimus [Terrazzano et al., 2020], sunitinib [Wang et al., 2020b].

For the remaining drugs in the top ten positions in both lists, we determine the main mechanism of action (i.e we determine the main pathways affected by the drug having the desired pharmacological effect) and we then look for hints in the literature of relevance of these pathways for COVID-19.

In the first list, we find methylergometrine, which is a partial agonist and antagonist of serotonin, dopamine, and  $\alpha$ -adrenergic receptors. Costa et al. [Costa et al., 2020] conjectured that serotonin inhibitors could have a neuroprotective effect in patients affected by COVID-19.

In the second list, we find erlotinib, stuent/sunitinib, azacitidine, and sorafenib. Erlotinib is an epidermal growth factor receptor (EGFR) inhibitor in use for the treatment of non-small cell lung cancer (NSCLC) and pancreatic cancer. Epidermal growth factor receptor (EGFR) signaling is known to be involved in the progress of viral infections [Hondermarck et al., 2020], including those caused by SARS [Venkataraman and Frieman, 2017] and SARS-Cov2 [Klann et al., 2020], thus it is a plausible target pathway to consider.

Sunitinib (Sutent) multi-targeted receptor tyrosine kinase (RTK) inhibitor used in the treatment of renal cell carcinoma (RCC) and imatinib-resistant gastrointestinal stromal tumor (GIST). Weidberg et al. [Weisberg et al., 2020] recently advocated the repurposing of kinase inhibitors for the treatment of COVID-19.

Azacitidine has been shown effective against human immunodeficiency virus (HIV) *in vitro* [Dapp et al., 2009] and human T-lymphotropic virus (HTLV) [Diamantopoulos et al., 2012]. Its action is mainly at a epigenetic level, inhibiting of DNA methyltransferase, causing hypomethylation of DNA. The role of epigenetics in COVID-19 is a growing area of research with large potential for advances in drug target identification [Pruimboom, 2020], [Schäfer and Baric, 2017], and [Chlamydas et al., 2021].

Sorafenib is a kinase inhibitor drug approved for the treatment of advanced primary renal cell carcinoma, advanced primary liver cancer, and forms of resistant advanced thyroid carcinoma. Sorafenib is a protein kinase inhibitor with activity against the Raf/Mek/Erk pathway [Liu et al., 2006]. Several authors have explored the role of MAP Kinase signalling pathway in the coronavirus family at large [Cai et al., 2007] and for SARS-cov2 in particular [Hemmat et al., 2021],[Ghasemnejad-Berenji and Pashapour, 2021], and [Zhou et al., 2020a].

In Supplementary File S8 and figure 9 we report performance measures for the proposed COVID-19 drug repurposing rankings from Taguchi et al 2020 [Taguchi and Turki, 2020], Mousavi et al 2020 [Mousavi et al., 2020], Gysi et al 2020 [Gysi et al., 2020], and Zhou et al 2020 [Zhou et al., 2020b].

For Taguchi et al 2020, and Mousavi et al 2020 the drug data sets used are overlapping with ours thus it is possible to compute the significance of the performance measures on a common basis. Interestingly as shown in Supplementary File S8 no such rankings attain the significance threshold p-value  $\leq 0.05$  in any configuration.

For Gysi et al. 2020 and Zhou et al. 2020, the drug data sets used are different from the ones we use, and the two methods are not based on gene expression data, and it is not possible to determine with confidence the initial pool of drug candidates. For this reason, we do not compute the normalized significance, and we report directly the absolute performance measures. The results of Gysi et al 2020 are in absolute value superiors to those of Zhou et al 2020 and qualitatively comparable to those obtained by DrugMerge on GEO data for precision@20, and on all four Drug

datasets for RHR. In particular, Gysi et al 2020 [Gysi et al., 2020] use data on interactions between viral and human proteins in [Gordon et al., 2020] to seed their active subnetwork and finish their listing with the help of expert advice, which we do not use. Gysi et al. 2020 also have a notion of 'positive' drug action akin to the one used in our paper.

## 3 Methods

### 3.1 Drug Ranking based on p-value

We use the Core&Peel technique described in [Lucchetta, 2020], [Pellegrini et al., 2016] and 4 other methods (ClustEx [Gu et al., 2010], ModuleDiscoverer [Vlaic et al., 2018], Degas [Ulitsky et al., 2010], and KeyPathwayMiner [Alcaraz et al., 2011]) to produce active subnetworks for a given disease  $D$ , starting from a list of differentially expressed genes in samples affected by the disease w.r.t. healthy controls  $DEG(D)$  and a gene co-expression network. In [Lucchetta, 2020], we employed the gene co-expression network made available by the DREAM challenge [Choobdar et al., 2019], and DEG for seven different datasets: two inflammatory disease microarray experiments (asthma and rheumatoid arthritis), two cancer RNA-Seq studies (prostate and colorectal cancer), and three COVID-19 datasets called BALF (bronchoalveolar lavage fluid RNA-Seq), PBMC (infected patient peripheral blood mononuclear cells) and COVID19 cells (human adenocarcinomic alveolar basal epithelial A549 cells). In [Lucchetta, 2020], we compared the performance of Core&Peel and the other four algorithms in detecting active subnetworks, resulting in that Core&Peel is a competitive method in that purpose. In this work, we use the five active subnetworks detected by all five methods as inputs of DrugMerge. Let  $AN$  be any such active network. For a given Drug perturbation datasets (for us: GEO, DrugMatrix, CMAP, and L1000) using the *enrichR* R/Bioconductor package [Kuleshov et al., 2016], we obtain a listing of drugs enriched in  $AN$  ranked by their adjusted p-value in reverse order (from the smallest to the largest). We cut this list at the adjusted p-value of 0.05, retaining only the significantly enriched drugs.

### 3.2 Drug perturbation databases

To predict the potentially repurposable drugs in each active subnetwork (one for each disease and each method), we used four drug perturbation databases available in Enrichr (<https://maayanlab.cloud/Enrichr/#stats>). These databases are called 'Drug\_Perturbations\_from\_GEO\_2014', 'DrugMatrix', 'LINCS\_L1000\_Chem\_Pert\_down', 'LINCS\_L1000\_Chem\_Pert\_up', 'Old\_CMAP\_down', and 'Old\_CMAP\_up'. The GEO- and DrugMatrix-related analysis has been performed in June 2020, that one for L1000 and CMAP has been performed in January 2021. Each database includes several drugs and the corresponding genes perturbed by the drug, distinguishing if these genes are up- or down-regulated. The Drug Perturbations database from GEO was compiled using experiments from GEO [Barrett et al., 2012] where gene expression levels were measured before and after the administration of a drug. This library has drug signatures for several tissues, and species including the *homo sapiens*, *mus musculus* and *rattus norvegicus*. In total, there are 701 terms and 132 unique drugs (where we do not consider the species, the tissues, and the modularity of the genes). DrugMatrix (DM) [Svoboda et al., 2019], [Ganter et al., 2005] is one of the largest rat toxicogenomics databases and has drug signatures available on liver, heart, kidney, thigh muscle, and primary hepatocytes. In total, it includes 7876 terms and 656 unique drugs. The Connectivity Map (also called CMAP) [Lamb et al., 2006], [Lamb, 2007] is a collection of gene expression data from five different human cells perturbed with many chemicals and genetic reagents. In total, CMAP applied 1309 compounds yielding 6100 profiles. The Connectivity Map project entered a most recent version of CMAP, as part of NIH's Library of Integrated Network-Based Cellular Signatures (LINCS) program, called LINCS-L1000. It comprises  $\sim 5000$  genetic perturbations and  $\sim 15000$  perturbations induced by

chemical compounds [Musa et al., 2018] across 98 different cell lines. The two CMAP projects differ also in sequencing platforms. The LINCS-L1000 project replaced the Affymetrix GeneChips used by the original CMAP with Luminex bead arrays [Lim and Pavlidis, 2019], which has been developed to facilitate rapid, flexible, and high-throughput gene expression profiling at lower costs. Using the *enrichR* package, in the L1000 database, we obtained 33132 terms and 4117 compounds.

### 3.3 Filtering of drug ranking

We have noticed that a straightforward application of the p-value based drug ranking includes in top ranking positions some molecules and drugs that are known to be either generally toxic or to induce diseases in animal models. For this reason, we introduce the option of applying two filters to a ranked list of drugs. The first filter is used to retain only drugs that are approved by the FDA (we use a list of drugs downloaded from <https://www.accessdata.fda.gov> on September 9th 2020, with 6315 entries). For FDA-approved drugs usually adverse effects and toxicity levels are known and probably acceptable in most cases. The second filter retains drugs that by the directionality analysis appear to have a 'positive' action on the disease, i.e. genes are up/down regulated antagonistic to the disease effect. This list of positive drugs can be obtained with the method described in the next section.

### 3.4 Drug ranking based on drug positivity scores

For this analysis, we exploit the fact that two of the methods we employ to build active subnetworks produce the active sub-network as a union of dense sub-graphs (ego-networks) of a large co-expression network (in detail this co-expression network has been produced for the DREAM challenge on disease module detection [Choobdar et al., 2019]). Since such network is built on the correlation of the expression-vectors over thousands of conditions, a high correlation vector implies that a pair of genes are generally both up or both down regulated at any given time. In a dense ego-network, such tendency is coherently common to all genes in the ego-network. We exploit this phenomenon as follows. For an ego-network that is significantly enriched in genes down-regulated by the disease, we subtract the number of up-regulated genes by the drug from the number of down-regulated genes by the drug. If this number is positive we say that the drug has a *positive* effect of the disease (for this ego-network). We do a symmetric calculation of the case of ego-networks that are significantly enriched in genes up-regulated by the disease. Summing these effects (with their sign) over all the ego-networks comprising the active subnetwork, we can compute the directionality of action of the drug and a magnitude (directionality score). We retain the drugs that have a positive directionality score and we rank them by the magnitude of the score. Note that this scheme relies just on counting the number of up/down regulated genes, not on the magnitude of the expression, as long as it is significant. In a variant of this scheme we can use the size of the ego-networks as a weighting factor. This method is justified by an extension to the ego-networks of the phenomenon Chen et al. [Chen et al., 2017] have noticed that in preclinical models of breast, liver, and colon cancers, reversal of gene expression between drug and disease correlates well with drug efficacy. Furthermore, Chen et al. validated the efficacy of drugs in xenograft models of the diseases.

### 3.5 Merging of Drug lists

Given two or more ranked lists of drugs, we wish to aggregate them into a single list that takes into account the ranking of any drug in each of the lists where it occurs. Following Gysi et al. [Gysi et al., 2020] we use the C-rank routine in [Zitnik et al., 2018b] to perform rank aggregation. C-rank is

defined in [Zitnik et al., 2018b] to work with any custom-defined ranking of items<sup>7</sup>. Interestingly this rank aggregation method is oblivious of the origin of the ranking and we can thus easily aggregate lists obtained with different active subnetwork algorithms on different disease-DEG sets. The only caveat is that the lists supplied to c-rank must be of equal length, thus the case of missing drugs must be handled, by fitting the drugs missing from a list at its bottom with a scaled mean score value obtained from the other lists.

### 3.6 Performance evaluation

The performance evaluation aims at assessing how well a predicted ranked list gives priorities in accordance to human expert judgment as codified by disease-treatment pairs extracted from the *Therapeutic Target Database* (TTD) [Chen et al., 2002] for the four benchmark diseases. Since for COVID-19 currently only one drug has gained acceptance as an approved clinical treatment, we resort to the list of drugs for the treatment of COVID-19 undergoing clinical trial as listed in the portal <https://www.clinicaltrials.gov/> (at Sept 5th, 2020).

We use two main measures. The first is precision at 20 (precision@20), that is the number of drugs in the first 20 positions of the proposed list of repurposable drugs that are hits in the TTD-derived listing, or, for COVID-19, in the clinicaltrials.gov listing.

The second is the Reciprocal Hit-Rank (RHR) [Deshpande and Karypis, 2004]. For a list  $L$  of predictions length  $N$ , let the hits be in positions  $p_1, p_2, \dots, p_h$ . We define :

$$RHR(L) = \sum_{i=1}^h \frac{1}{p_i}.$$

Note that this formula does not imply a fixed length for the lists being compared, and it is fair in comparing both long and short lists. A longer list may have more hits by chance, however it will have positive but diminishing returns for hits in its tail. We prefer not to use ROC-related estimates that may be inaccurate when the hits and miss classes are very unbalanced [Hanczar et al., 2010] [Brown and Patel, 2018].

### 3.7 Statistical significance

Let  $D$  be a set of drugs and  $H \subset D$  a golden standard for a disease (e.g. the drugs in  $D$  that are in clinical use for the disease). We generate  $n$  random permutations uniformly at random  $r_i(D)$ , for  $i = 1, \dots, n$  and we take, for a performance measure  $\mu_H()$  depending on  $H$ , the mean  $m$  and standard deviation  $sd$  of the sequence  $M = [\mu_H(r_i(D)) \text{ for } i = 1, \dots, n]$ . For a given drug ranking  $R(D)$ , its *zscore* for  $\mu_H()$  is its deviation from the mean of the above described random distribution normalized by the standard deviation of the random distribution:

$$zscore(R(D)) = \frac{\mu_H(R(D)) - m(M)}{sd(M)}$$

Assuming that the random distribution of the values of  $\mu_H(r_i(D))$  is normal, we obtain the (two sided) p-value by plugging the z-score value into the complementary cumulative distribution function (ccdf) of the normal distribution from the package `scipy.stats` of SciPy.org.

The number  $n$  of random permutations is determined as follows. Starting with  $n = 100,000$ , we increment  $n$  in steps of 10,000, stopping when the value of both  $m$  and  $sd$  are stable, with a displacement from the previous iteration less than 0.01.

---

<sup>7</sup>Thus we do not make use of the graph-based community ranking scores for which c-rank was originally designed, instead we use it here as a generic rank aggregation method

### 3.8 Drug name and aliases

We handle drug aliases by using the MESH<sup>8</sup> database. Moreover, we identified commercial drug names with their principal active ingredient. Ambiguous cases were checked by hand against the pharmacological literature and DrugBank records (<https://go.drugbank.com/>).

### 3.9 CMAP algorithm

To compare the DrugMerge performance in the four benchmark diseases (asthma, rheumatoid arthritis, prostate cancer, and colorectal cancer), we used the *PharmacoGx* [Smirnov et al., 2016] R package, which performs the Connectivity Map (CMAP) analysis [Lamb et al., 2006]. For each disease, we used the differentially expressed genes (DEGs) as gene input to identify drugs with therapeutic potential in each disease. In particular, we used the *downloadPSet* function to download the CMAP database, and the *drugPerturbationSig* function to identify differential gene expressions induced by drug treatment. Finally, the *connectivityScore* function compares drug signatures against disease signatures (DEGs), by assigning connectivity score and p-value for each drug. The connectivity score determines the correlation between the drug and disease signatures and it ranges from  $-1$  to  $1$ . Since we are looking for disease treatments or drug repurposing, the CMAP drugs should be anti-correlated with disease signatures. For this reason, we ordered the final CMAP drugs according to the increasing order of the connectivity score (from negative to positive values). We also filtered the CMAP drugs according to  $p\text{-value} < 0.05$  and we used both lists with or without p-value filtering. Starting from these lists, we calculated the RHR and precision@20 with z-score and p-values associated. After that, we selected those results with the best values of RHR and compare them with values detected by DrugMerge as reported in Figure 6 and Table S6.

## 4 Discussion

We can identify three main strategies in Drug Repositioning (DR): the first is disease-oriented DR (i.e. find repurposable drugs for a specified disease), the second is drug-oriented DR (finding a disease for a specified drug), the third is generic DR (find any novel drug-disease pair). The approaches may vary significantly across the three types, and also the validation methodology is strongly influenced by the objective of the strategy. In this study we tackle disease-oriented DR, however, we should point out that the base methodology can in principle be adapted easily to the other two strategies, provided sufficient high-quality data is available.

DrugMerge is a proof-of-concept exploring a novel use of disease active subnetworks in the context of drug repositioning aimed at a specific disease. The tests involving four benchmark diseases with drugs in clinical use did confirm the superior performance of the proposed method over the current state of the art (e.g. CMAP) in placing such clinically approved drugs in top ranking positions. In this case, the evaluation criterion is very strict since we were able to guess correctly the efficacy and safety of drugs administered to patients basing our analysis only on transcriptomic data from *in vitro* cell lines and tissues. Interestingly, while efficacy is in principle detectable by examining specific types of cells or tissues primarily affected by a disease, safety is a much broader concept involving all human organs and tissues, thus harder to pinpoint with localized experiments either *in vitro* or *in silico*.

Tests on the data from the COVID-19 disease also show, according to our validation criterion, a performance superior to some of the approaches proposed in the literature, while for some others we notice a broadly equivalent outcome, even when using different approaches and different input data. Results on COVID-19 should be taken with caution, however, as our validation methodology implies that the proposed ranking is able to mimic the best judgment of human experts, which may

---

<sup>8</sup><https://www.nlm.nih.gov/mesh/>

be based on their prior knowledge and expertise, data available to them, and their assessment of the odds, leading to a shortlisting of such drugs in clinical trials. Thus our validation methodology might indeed be failing us in cases where the disease to be tackled has key features that escape our present general understanding. What can be gained in our exercise is a way to make drug shortlisting for *in vitro* further assessment (and eventual accession to clinical trial) more systematic, while also using relatively cost effective data sources. Besides the global gene co-expression network, which we can assume as background knowledge, our method relies on rather standard and relatively inexpensive RNA-Seq and microarray transcriptomic measurements of differentially expressed genes in infected cells/tissues. Systematization and finding alternatives to pure serendipity (discovery by chance) are key aspects of all *in silico* based drug repositioning efforts.

One bonus of this parsimonious data requirement is that it becomes easier to track the evolution of viral Mechanism of Action and disease etiology as the viral agent mutates, thus in principle the drug ranking, and thus follow up decisions, may change rapidly when transcriptomic data on new viral strains of COVID-19 become available.

Sars-Cov2 is a member of the coronavirus family, responsible also for the Sars-CoV and MERS epidemics. Many studies, in particular at the end of 2019 and beginning of 2020, make up for the paucity of data on Sars-Cov2 by making use of genomic, transcriptomic, proteomic, and pharmacological data relative to Sars-CoV and MERS and other members of the coronavirus family, including those known to be hosted by other animal species.

In our study, we chose not to use data from viruses different from Sars-Cov2 as primary data. This choice comes from two orders of considerations. The first is that data from close relatives of the pathogen may not always be available, and there is thus a general long term advantage attained by methods that rely just on relatively small and focused data. The second issue is that we focus not on the pathogen in itself but on the human transcriptional response to it, which is known to be rather different across Sars-Cov2, Sars-CoV, and MERS epidemics. In this line of reasoning, it makes more sense to highlight what makes human transcriptional response to Sars-Cov2 different from that induced by other coronaviruses, even when the causing pathogens are phylogenetically very close.

We use the list merging procedure to merge lists obtained by several algorithms and several DE gene data sets from different cell lines/tissues but always referred to a single drug data sets, and a single disease. Different drug data sets are often obtained using different initial candidate drugs and use different technologies. Moreover sometimes, due to the different technologies and other experimental settings, even when we restrict the analysis to common drugs, replicability across drug data sets may be low (see e.g. [Lim and Pavlidis, 2019] [Lin et al., 2020] reporting low replicability across CMAP and L1000 data using the CMAP algorithm). Moreover the performance measures we use measure the quality of the whole ranking, and the ranking itself gives a way to compare drugs listed by position. However, based on these measures we refrain from comparing (and thus give a relative order) to two drugs in different rankings obtained with different perturbation drug data sets. Such a global ranking would need further insight and technological bias correction, which we leave for future research.

Drug DE gene data report often measurements for a range of different dosages taken at a range of different time points. For the p-valued ranking, we use the dosage/time point giving the lowest p-value. For rankings based on positivity score we take the average positivity score. These choices are justified for an initial assessment of drug potential to be effective against a disease. Dosage and time points could be taken into account together with available dose-dependent toxicity data in a second level of analysis when additional data is collected in *in vitro* experiments.

For this study, we use all available transcriptomic drug and disease data regardless of the cell lines/tissues/species for which the transcriptomic measurements are taken. This choice increases the robustness of the results in general, however, it might miss some more subtle tissue-specific phenomena [Newman et al., 2020].

On the one hand, it is known that the SARS-Cov2 virus infects several human organs includ-

ing lungs, liver, kidneys, and gut [Mallapaty, 2020], and the cells transcriptional response may in principle be diverse in each tissue. However, Dudley et al. [Dudley et al., 2009] analyzing a large repository of disease-related gene differential expression data conclude that "molecular signature of disease across tissues is overall more prominent than the signature of tissue expression across diseases". Our analysis of active subnetworks extracted from three COVID-19 data sets (BALF, PBMC, cells) also supports for COVID-19 a general high similarity of gene expression across the sample types. Differences in the extracted active networks are mainly due to the algorithms used.

Also, drugs may act differently in different cell types. For example, Chloroquine has been found to be ineffective in infected human lung cells [Hoffmann et al., 2020], while it was found effective in infected Vero cells (a cell line isolated from kidney epithelial cells extracted from an African green monkey) [Wang et al., 2020a]. Similar tissue-specific differences in drug sensitivity are noticed by Dittmar et al. [Dittmar et al., 2020]. In principle, if a variety of options are available, one may opt for including only the data drawn from biological samples more similar to the actual tissues affected by the disease in patients. In the context of Cancer drug repositioning, a tool like CELLector [Najgebauer et al., 2020] may be beneficial for a systematic cell/tissue pre-selection.

In our study we use a large variety of cellular types for drug perturbation analysis. Drug matrix [Ganter et al., 2005] has drug perturbation signatures available on liver, heart, kidney, thigh muscle and primary hepatocytes (from *rattus norvegicus*). GEO [Barrett et al., 2012] has drug perturbation signatures for several tissues and species including *homo sapiens*, *mus musculus* and *rattus norvegicus*. CMAP measures drug perturbations in breast cancer epithelial cell line MCF7 (MCF7/ssMCF7), in prostate cancer epithelial cell line PC3, and the nonepithelial lines HL60 (leukemia) and SKMEL5 (melanoma). L1000 collects drug perturbation data from 98 different cellular contexts including human primary cell lines and human cancer cell lines [Lim and Pavlidis, 2019].

For COVID19 we report two drug rankings attaining top performance, one is obtained by merging data from all three tissues used to measure DE genes, therefore it indicates evidence for drugs that may act on a general systemic level across different human organs. This is complemented by the second ranking which is attained specifically for BALF measurements (bronchoalveolar lavage fluid) and thus it may indicate drugs with a specific action in pulmonary tissues.

One advantage of our approach, in contrast with many others, is that we do not have an initial explicit guess for a putative target protein for the drug or a putative target pathway (or mechanism of action). However, once relevant drugs are shortlisted we can analyze their likely mechanism of action, either through the literature or through the use of specialized databases (e.g. DrugPath [Jaundoo and Craddock, 2020]).

In this study, we have used four public Drug related data sets (GEO, DrugMatrix, CMAP, and L1000) representing a sufficient drug basis for our investigation and the validation of the proposed methodology. However, a much larger amount of useful data on drugs is proprietary therefore it is important to devise incentives for sharing these repositories of useful data in a global drug repositioning strategy. As a consequence of the non-complete drug coverage of our data sets we could not assess some drugs whose use in COVID-19 cases has been advocated, such as remdesivir (approved by FDA in October 2020 [Rubin et al., 2020]) or lopinavir (not approved by FDA). Cao et al. [Cao et al., 2020] recently reported the results of a randomized clinical trial involving 199 hospitalized adult patients with severe COVID-19: no benefit was observed with lopinavir-ritonavir treatment over the standard care cohort. Interestingly, we have data from DrugMatrix on ritonavir (not approved by FDA), which fails the positivity score filter, thus our results on ritonavir are consistent with the findings in [Cao et al., 2020].

Comorbidities in patients affected by COVID-19 are important clinical aspects, as it is known that comorbidities are one of the strongest risk factors for such patients, strongly affecting mortality rates [Sanyaolu et al., 2020]. Comorbidities may alter drastically the effect of drugs. The disease gene expression data (for all five diseases) we use in this study do not come from patients with known comorbidities, therefore the results of our study should be interpreted in such a setting. Some

authors ([Fiscon et al., 2020]) try to exploit known comorbidities as a positive tool in modeling the relevant features of COVID-19, while other authors, based on protein target analysis, have a more skeptical view ([Gysi et al., 2020]): "In summary, we find that the SARS-CoV2 targets do not overlap with disease genes associated with any major diseases, indicating that a potential COVID-19 treatment can not be derived from the arsenal of therapies approved for specific diseases." The effect of concomitant comorbidities on repurposed drugs might be better approached when transcriptomic data from these sub-populations of COVID-19 patients become available.

In this study, we do not treat the issue of combinations of drugs. This is a novel area of research in which network-based approaches have been shown to give novel insight and may help in coping with the combinatorial explosion of drug combinations. However, validated data on drug combinations is limited thus making the task of performance evaluation and method validation more complex w.r.t the analysis for single drugs. We plan to incorporate models of multi-drug/disease interactions in our setting as future research [Cheng et al., 2019].

Drugs can be toxic or cause adverse events. Toxicity can be detected via ad-hoc *in vitro* and *in vivo* experiments or it can be predicted via *in silico* models [Zhang et al., 2018],[Ganter et al., 2005],[Igarashi et al., 2015],[Alexander-Dann et al., 2018],[Poleksic and Xie, 2019]. Any of these techniques should be applied downstream to the use of DrugMerge for extra safety. DrugMerge uses two forms of filtering to handle adverse effects and toxicity. The first mechanism is the inclusion of drugs having a contrary effect to the disease as measured by the positivity score (which in turn is similar to a mechanism also used by [Gysi et al., 2020]). The second filtering is via the possibility to limit the analysis to FDA-approved drugs, for which toxicity and adverse effects are known and have been considered to be outweighed by the benefits in certain conditions. This mild form of filtering is sufficient to increase the performance in our tests and thus mimic human expert judgment, which includes consideration of toxicity and adverse effects. However, toxicity should be always be reconsidered ex-post to finalize any shortlisting decision.

Since we base our approach on the transcriptomic responses of human cells/tissues to drugs and diseases, the cause of the disease (viral, bacterial, genetic, metabolic, etc..) is not critical for the application of the methodology. Indeed in our study, we applied it to a variety of cases. Asthma is an inflammatory disease of the airways of the lungs having several triggering causes, including exposure to air pollution and allergens. Rheumatoid arthritis (RA) is a long-term autoimmune disorder that primarily affects joints. Cancer is generally associated with alterations of the genetic code in somatic human cells. Finally, COVID-19 is of viral origin. The advantage of generality is balanced by the possibility that a high ranking drug is effective against symptoms or side-effects of the diseases, rather than acting on its main mechanism of action. In one of the drug rankings for COVID-19, we find within the top 20 positions prochlorperazine which has both antiviral activity [Otręba et al., 2020] and is also a broad-spectrum antiemetic contrasting covid19 symptoms like nausea and vomiting [Kow and Hasan, 2020], [Zhang et al., 2020a]. At this level of analysis, DrugMerge does not yet distinguish the different actions a drug might have thus further investigations are needed to untangle complex drug actions.

## 5 Related work

Computational drug repositioning is a burgeoning field of research (see. e.g. [Li et al., 2016], [Levin et al., 2020] [Galindez et al., 2021], [Dotolo et al., 2020], [Pushpakom et al., 2019], [Alaimo and Pulvirenti, 2019], [Lotfi Shahreza et al., 2018], [Xue et al., 2018], [Jarada et al., 2020], [Low et al., 2020]) receiving much attention for its high potential impact on coping with present and future pandemic events by novel pathogens [Zhou et al., 2020c] [Nabirovichkin et al., 2020].

For the four benchmark diseases considered in our study, studies on repurposed drugs for tumors are reported in [Nowak-Sliwinska et al., 2019], [Zhang et al., 2020b], [Turanli et al., 2018], for asthma in [Kruse and Vanijcharoenkarn, 2018], and for rheumatoid arthritis in [Hu et al., 2019], and [Kim



et al., 2018].

A survey focusing on repurposing anti-cancer drugs for covid19 is in [Ciliberto et al., 2020]. Computational drug repositioning is gaining much attention during the current covid19 pandemic as reported in [Singh et al., 2020].

Network-based drug repositioning [Barabasi and Oltvai, 2004] is an approach that leverages on building and analyzing various types of biological networks integrating several layers of 'omic' data [Re and Valentini, 2013]. Also, drug perturbation data bases play a central role in computational drug profiling [Duan et al., 2014], [Lamb et al., 2006], [Subramanian et al., 2017], [Ganter et al., 2005]. Next, we focus on some results in the lines of research that are most relevant for our study.

Taguchi et al. [Taguchi and Turki, 2020] apply an unsupervised method based on tensor decomposition to perform feature extraction of gene expression profiles in multiple lung cancer cell lines infected with severe acute respiratory syndrome coronavirus 2 [Blanco-Melo et al., 2020]. They thus identified, using Enrichr [Kuleshov et al., 2016], drug candidates that significantly altered the expression of the 163 genes selected in the previous phase. Drug perturbation data is collected from GEO, DrugMatrix, L1000, and other repositories.

Ruiz et al. [Ruiz et al., 2020] build a multi-layer disease-gene-drug-pathway network and develop a random-walk-based score that captures indirect effects of drugs on diseases via commonly affected pathways. They validate their approach in a leave-one-out validation on a golden standard of 6000 drug-disease pairs commonly used in clinical practice. The paper, however, does not report on the specific application of this approach to COVID-19.

Fiscon et al. [Fiscon et al., 2020], [Fiscon et al., 2021] use an approach based on computing distances between regions of a large human interactome network affected by a disease (disease targets) and those affected by drugs (drug perturbations). This work relies on disease target similarities between a group of diseases including covid19, SARS-coV, MERS, and others. The final assessment of drug candidates for COVID19 is done with the C-map database.

Gysi et al. [Gysi et al., 2020] define several proximity-based distance functions between covid19 human target proteins (as listed in [Gordon et al., 2020]) and drug protein targets as listed in DrugBank [Wishart et al., 2018], in order to prioritize repurposable drugs. Several refinement taking into account tissue specificity, drug action on gene expression levels (using data from [Blanco-Melo et al., 2020]), comorbidities, and drug toxicity lead to a final list of 81 repurposable drugs for covid19. The main measure of performance chosen is the AUC of the predicted list versus the list of drugs currently employed in clinical trials for the treatment of COVID-19 as listed in <https://www.clinicaltrials.gov/>.

Our approach to performance evaluation is similar to the one in [Gysi et al., 2020] since we also use drugs in clinical trials as the golden standard for covid19. Thus we are both measuring how well our automated ranking systems come close to the collective wisdom of human experts that shortlisted existing drugs for repurposing on COVID-19 during 2020 based on a large variety of considerations such as available clinical and preclinical data, pharmacological background, guesses mechanism of action. However, we decided to use precision@20 and the Reciprocal Hit Ranking as quantitative measures since they are quite intuitive, more suitable for a ranking problem, and they handle more uniformly lists of candidate drugs that can span from a few dozen drugs to a few thousand.

Zhou et al. [Zhou et al., 2020b] also use proximity-based measures in biological networks to find a list of 16 repurposable drugs. It should be noticed that Zhou et al. use data collected from the family of human coronaviruses (HCoVs) to build the network, thus relying heavily on evolutionary conservation of relevant coding parts of the viral genome across the species in this family and on phylogenetic considerations.

Sadegh et al. [Sadegh et al., 2020] use the algorithm KeyPathwayMiner [Alcaraz et al., 2011] to define the active subnetwork in an integrated COVID-19 resource network from CoronaVirus Explorer CoVex resource (<https://exbio.wzw.tum.de/covex/>). The main contribution of [Alcaraz et al., 2011] is a resource (CoVex) that can be used interactively in many different scenarios and

modalities to explore disease-gene-drug relationships for COVID-19.

Mall et al. [Mall et al., 2020] use deep learning induced vector embedding of drugs and viral proteins to predict drug-viral protein activity and propose a short list of 15 drugs potentially useful against COVID-19.

A second very popular approach uses knowledge of the 3-dimensional molecular configuration of proteins and drugs to simulate, via docking or molecular dynamics, the most promising drug-target bindings. Most research in docking-based drug repositioning starts with an assessment of one (or a few) protein acting as potential drug targets [Chellapandi and Saranya, 2020].

Seo et al. [Seo et al., 2020] take three dimensional structure of the main protease (Mpro) of SARS-CoV-2 as a target, and use docking simulations on a supercomputer to evaluate binding affinity between Mpro and drug candidates listed in the SWEETLEAD library and the ChEMBL database (19,168 molecules), thus shortlisting 43 drugs, which, after molecular dynamics simulations, are then reduced to 8. A similar docking-based screening is also used in [Gimeno et al., 2020], [Shah et al., 2020], [Elmezayen et al., 2020], [Wang, 2020], [Trezza et al., 2020], and [Mahapatra et al., 2020].

Besides network-based and docking-based approaches there is a variety of other principles and approaches that have been used both *in silico* and *in vitro*: Luminescent cell viability assay and cell imaging [Riva et al., 2020a] [Riva et al., 2020b], morphological cell profiling [Mirabelli et al., 2020], co-morbidity [Luna et al., 2020], deep learning [Zeng et al., 2020], network proximity based [Stolfi et al., 2020], quantum mechanical scoring [Cavasotto and Filippo, 2020], multi-objective optimization [Chen et al., 2020], bipartite network projections [Re and Valentini, 2013], matrix factorization [Huang et al., 2020], machine learning [Loucera et al., 2020], network node scoring [Nam et al., 2020] [Ge et al., 2020], semi-supervised learning [Iorio et al., 2015], variants of gene set enrichment analysis (GSEA) [Fang et al., 2021], multi-omic integration [Tomazou et al., 2021].

A more direct experimental approach has been described in [Gordon et al., 2020] and [Tutuncoglu et al., 2020]. Gordon et al. [Gordon et al., 2020] performed a direct proteomic assay to uncover 332 human proteins potentially interacting with 26 SARS-Cov2 proteins. Sixty-nine drugs targeting these proteins are then shortlisted as repurposable drugs for covid19, using a mix of chemoinformatic searches and expert advice. Note that in this setting a priority ranking is not provided.

A main theme along with drug ranking is that of drug combinations [Muratov and Zakharov, 2020]. Often drug combinations offer lower toxicity and more effective disease treatment. Most drug combinations currently under consideration result from ad-hoc considerations, as screening drug combinations automatically incurs easily in a combinatorial explosion of cases to be considered. A few network-based principled approaches have been proposed and may be integrated into DrugMerge [Artigas et al., 2020], [Zhou et al., 2020b], [Bobrowski et al., 2020], [Bobrowski et al., 2021], [Jin et al., 2020], [Cheng et al., 2019], and [Zitnik et al., 2018a].

While the final hallmark of success is the identification of a repurposed drug that can pass successfully all stages of drug approval for clinical use (see e.g. [Edwards, 2020]), it is important to be able to perform *in silico* validation on intermediate results. A recent study by Brown and Patel [Brown and Patel, 2018] gives a critical assessment of the options for *in silico* drug validations, assessing the weak and strong points of each strategy.

## 6 Conclusions

In this study, we describe DrugMerge, a methodology for ranking repurposable drugs where we rank drugs by their ability to affect collections of disease active subnetworks (computed by an ensemble of state-of-the-art algorithms) and contrast the perturbation induced by the disease in such sub-networks. The rationale behind our approach is a model in which drugs and diseases interact indirectly by their mutual action on the sub-networks rather than directly by affecting the same target proteins. Moreover, we are able to leverage on an ensemble of different algorithms for

computing active subnetworks, and then merge the resulting rankings, thus increasing the robustness of the methodology.

Results of four benchmark diseases are encouraging since we could find in all four cases drugs in clinical use (thus both safe and effective) in first ranking positions, and several other drugs in clinical use in high ranking positions. These results imply that DrugMerge when provided with input data from transcriptional assays, has the potential for coming close to expert human judgment, and guess drugs in actual clinical usage. For COVID-19 we measure how close we come to guessing drugs that have been shortlisted for clinical trials according to the FDA records. For most drugs in high ranking positions but not on clinical trials we have found several supporting evidence in the literature or by direct assessment of their likely Mechanism of Action, that makes their indication by DrugMerge biologically plausible.

DrugMerge should be considered as a proof-of-concept and its predictions should be considered as indications for shortlisting drugs for further *in vitro* analysis rather than absolute predictions. In particular, aspects like dosage, tissue specificity, and a more refined analysis of possible adverse side effects will need to be tackled further in developing DrugMerge towards a full-fledged product.

Prioritizing drugs for *in vitro* testing and eventual accession to clinical trials is valuable for decision makers in a pandemic situation. Lack of prioritization in drug repurposing efforts has led to some drugs being over-researched, early in the pandemic, even based on dubious evidence (e.g. hydroxychloroquine in COVID-19). Moreover having many small clinical trials has the effect of reducing the pool of eligible patients available for larger and more promising trials<sup>9</sup>.

A tool such as DrugMerge aiming at drug ranking might be used in conjunction with an activity of data collection that targets specific sub-populations of patients. For example, most trials to date for repurposed drugs in COVID-19 have focused on hospitalized patients, even though prevention might be more valuable and effective if initiated earlier in disease progression. Also transcriptomic data from patients affected by frequent comorbidities might help towards the study of drugs specifically selected for sub-populations at higher risk of mortality [Stephenson et al., 2021].

This paper focuses on the preclinical considerations that may indicate a drug to be shortlisted for further preclinical *in vitro* and *in vivo* testing, leading to eventual clinical trial. Naturally, all evidences, which may conceivably be noisy and often contradictory, should be gathered and considered before pushing a drug further down the pipeline.

## 7 Competing interests

The authors declare no conflict of interest for this article.

## 8 Data Availability

Primary data used in this study is available in public repositories. Intermediate results, processing software and scripts will be made available on GitHub.

## 9 Funding

M.L.'s doctoral fellowship is funded by CNR, Institute of Informatics and Telematics. The research exposed in this article has been conducted as curiosity-driven free research by the authors.

---

<sup>9</sup>Generic Drug Repurposing for COVID-19 and Beyond. Susan Athey, Rena Conti, Richard Frank, and Jonathan Gruber Institute for Health System Innovation & Policy. July2020. <https://www.gsb.stanford.edu/faculty-research/publications/generic-drug-repurposing-covid-19-beyond>

## 10 Author's contributions

M.P. designed the study. M.L. performed the analyses. M.P. and M.L. interpreted the results and wrote the manuscript. All authors reviewed the manuscript.

## References

- [Alaimo and Pulvirenti, 2019] Alaimo, S. and Pulvirenti, A. (2019). Network-based drug repositioning: Approaches, resources, and research directions. In *Computational Methods for Drug Repurposing*, pages 97–113. Springer.
- [Alcaraz et al., 2011] Alcaraz, N., Küçük, H., Weile, J., Wipat, A., and Baumbach, J. (2011). Key-pathwayminer: detecting case-specific biological pathways using expression data. *Internet Mathematics*, 7(4):299–313.
- [Alexander-Dann et al., 2018] Alexander-Dann, B., Pruteanu, L. L., Oerton, E., Sharma, N., Berindan-Neagoe, I., Módos, D., and Bender, A. (2018). Developments in toxicogenomics: understanding and predicting compound-induced toxicity from gene expression data. *Molecular omics*, 14(4):218–236.
- [Artigas et al., 2020] Artigas, L., Coma, M., Matos-Filipe, P., Aguirre-Plans, J., Farrés, J., Valls, R., Fernandez-Fuentes, N., de la Haba-Rodriguez, J., Olvera, A., Barbera, J., et al. (2020). In-silico drug repurposing study predicts the combination of pirlfenidone and melatonin as a promising candidate therapy to reduce sars-cov-2 infection progression and respiratory distress caused by cytokine storm. *PloS one*, 15(10):e0240149.
- [Ashburn and Thor, 2004] Ashburn, T. T. and Thor, K. B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nature reviews Drug discovery*, 3(8):673–683.
- [Azad et al., 2017] Azad, N. S., El-Khoueiry, A., Yin, J., Oberg, A. L., Flynn, P., Adkins, D., Sharma, A., Weisenberger, D. J., Brown, T., Medvari, P., et al. (2017). Combination epigenetic therapy in metastatic colorectal cancer (mcr) with subcutaneous 5-azacitidine and entinostat: a phase 2 consortium/stand up 2 cancer study. *Oncotarget*, 8(21):35326.
- [Barabási et al., 2011] Barabási, A.-L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nature reviews genetics*, 12(1):56–68.
- [Barabasi and Oltvai, 2004] Barabasi, A.-L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature reviews genetics*, 5(2):101–113.
- [Barrett et al., 2012] Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., et al. (2012). Ncbi geo: archive for functional genomics data sets—update. *Nucleic acids research*, 41(D1):D991–D995.
- [Blanco-Melo et al., 2020] Blanco-Melo, D., Nilsson-Payant, B., Liu, W.-C., Møller, R., Panis, M., Sachs, D., Albrecht, R., et al. (2020). Sars-cov-2 launches a unique transcriptional signature from in vitro, ex vivo, and in vivo systems. *BioRxiv*.
- [Bobrowski et al., 2020] Bobrowski, T., Chen, L., Eastman, R. T., Itkin, Z., Shinn, P., Chen, C. Z., Guo, H., Zheng, W., Michael, S., Simeonov, A., et al. (2020). Discovery of synergistic and antagonistic drug combinations against sars-cov-2 in vitro. *bioRxiv*. doi:10.1101/2020.06.29.178889.

- [Bobrowski et al., 2021] Bobrowski, T., Chen, L., Eastman, R. T., Itkin, Z., Shinn, P., Chen, C. Z., Guo, H., Zheng, W., Michael, S., Simeonov, A., et al. (2021). Synergistic and antagonistic drug combinations against sars-cov-2. *Molecular Therapy*, 29(2):873–885.
- [Boland, 1952] Boland, E. W. (1952). Rheumatoid arthritis—experiences with hydrocortisone (free alcohol) and hydrocortisone acetate. *California medicine*, 77(1):1.
- [Bonds and Midoro-Horiuti, 2013] Bonds, R. S. and Midoro-Horiuti, T. (2013). Estrogen effects in allergy and asthma. *Current opinion in allergy and clinical immunology*, 13(1):92.
- [Brown and Patel, 2018] Brown, A. S. and Patel, C. J. (2018). A review of validation strategies for computational drug repositioning. *Briefings in bioinformatics*, 19(1):174–177.
- [Cai et al., 2007] Cai, Y., Liu, Y., and Zhang, X. (2007). Suppression of coronavirus replication by inhibition of the mek signaling pathway. *Journal of virology*, 81(2):446–456.
- [Cao et al., 2020] Cao, B., Wang, Y., Wen, D., Liu, W., Wang, J., Fan, G., Ruan, L., Song, B., Cai, Y., Wei, M., et al. (2020). A trial of lopinavir–ritonavir in adults hospitalized with severe covid-19. *New England Journal of Medicine*, 382:1787–1799.
- [Carr et al., 2015] Carr, A., Vissers, M., and Cook, J. (2015). Parenteral vitamin c relieves chronic fatigue and pain in a patient presenting with rheumatoid arthritis and mononeuritis multiplex secondary to cns vasculitis. *Case Rep Clin Path*, 2(2):57–61.
- [Carr and McCall, 2017] Carr, A. C. and McCall, C. (2017). The role of vitamin c in the treatment of pain: new insights. *Journal of translational medicine*, 15(1):1–14.
- [Cavasotto and Filippo, 2020] Cavasotto, C. and Filippo, J. D. (2020). In silico drug repurposing for covid-19: Targeting sars-cov-2 proteins through docking and quantum mechanical scoring. *ChemRxiv*.
- [Chellapandi and Saranya, 2020] Chellapandi, P. and Saranya, S. (2020). Genomics insights of sars-cov-2 (covid-19) into target-based drug discovery. *Medicinal Chemistry Research*, pages 1–15.
- [Chen et al., 2017] 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. *Nature communications*, 8(1):1–12.
- [Chen et al., 2020] Chen, H., Cheng, F., and Li, J. (2020). idrug: Integration of drug repositioning and drug-target prediction via cross-network embedding. *PLoS computational biology*, 16(7):e1008040.
- [Chen et al., 2002] Chen, X., Ji, Z. L., and Chen, Y. Z. (2002). Ttd: therapeutic target database. *Nucleic acids research*, 30(1):412–415.
- [Cheng et al., 2019] Cheng, F., Kovács, I. A., and Barabási, A.-L. (2019). Network-based prediction of drug combinations. *Nature communications*, 10(1):1–11.
- [Chlamydas et al., 2021] Chlamydas, S., Papavassiliou, A. G., and Piperi, C. (2021). Epigenetic mechanisms regulating covid-19 infection. *Epigenetics*, 16(3):263–270.
- [Choobdar et al., 2019] Choobdar, S., Ahsen, M. E., Crawford, J., Tomasoni, M., Fang, T., Lamparter, D., Lin, J., Hescott, B., Hu, X., Mercer, J., et al. (2019). Assessment of network module identification across complex diseases. *Nature methods*, 16(9):843–852.

- [Christiansen et al., 2016] Christiansen, S. C., Schatz, M., Yang, S.-J., Ngor, E., Chen, W., and Zuraw, B. L. (2016). Hypertension and asthma: a comorbid relationship. *The Journal of Allergy and Clinical Immunology: In Practice*, 4(1):76–81.
- [Ciliberto et al., 2020] Ciliberto, G., Mancini, R., and Paggi, M. G. (2020). Drug repurposing against covid-19: focus on anticancer agents. *Journal of Experimental & Clinical Cancer Research*, 39:1–9.
- [Cornelia et al., 1991] Cornelia, G., Casaretti, R., Cornelia, P., Daponte, A., Parziale, A., Iervolino, V., Santillo, G., and Zarrilli, D. (1991). Treatment of advanced colorectal cancer with mitoxantrone, high dose folinic acid and fluorouracil. *Tumori Journal*, 77(5):445–446.
- [Coss et al., 2009] Coss, A., Tosetto, M., Fox, E. J., Sapetto-Rebow, B., Gorman, S., Kennedy, B. N., Lloyd, A. T., Hyland, J. M., O’Donoghue, D. P., Sheahan, K., et al. (2009). Increased topoisomerase  $\alpha$  expression in colorectal cancer is associated with advanced disease and chemotherapeutic resistance via inhibition of apoptosis. *Cancer letters*, 276(2):228–238.
- [Costa et al., 2020] Costa, L. H., Santos, B. M., and Branco, L. G. (2020). Can selective serotonin reuptake inhibitors have a neuroprotective effect during covid-19? *European journal of pharmacology*, 889:173629.
- [Cron, 2020] Cron, R. Q. (2020). Coronavirus is the trigger, but the immune response is deadly. *The Lancet Rheumatology*.
- [Dapp et al., 2009] Dapp, M. J., Clouser, C. L., Patterson, S., and Mansky, L. M. (2009). 5-azacytidine can induce lethal mutagenesis in human immunodeficiency virus type 1. *Journal of virology*, 83(22):11950–11958.
- [de la Rica et al., 2020] de la Rica, R., Borges, M., and Gonzalez-Freire, M. (2020). Covid-19: In the eye of the cytokine storm. *Frontiers in Immunology*, 11:2313.
- [Deans and West, 2011] Deans, A. J. and West, S. C. (2011). Dna interstrand crosslink repair and cancer. *Nature reviews cancer*, 11(7):467–480.
- [Deshpande and Karypis, 2004] Deshpande, M. and Karypis, G. (2004). Item-based top-n recommendation algorithms. *ACM Transactions on Information Systems (TOIS)*, 22(1):143–177.
- [Diamantopoulos et al., 2012] Diamantopoulos, P. T., Michael, M., Benopoulou, O., Bazanis, E., Tzeletas, G., Meletis, J., Vayopoulos, G., and Viniou, N.-A. (2012). Antiretroviral activity of 5-azacytidine during treatment of a htlv-1 positive myelodysplastic syndrome with autoimmune manifestations. *Virology journal*, 9(1):1–5.
- [Dittmar et al., 2020] Dittmar, M., Lee, J. S., Whig, K., Segrist, E., Li, M., Jurado, K., Samby, K., Ramage, H., Schultz, D., and Cherry, S. (2020). Drug repurposing screens reveal fda approved drugs active against sars-cov-2. *bioRxiv*.
- [Dorward et al., 2020] Dorward, D. A., Russell, C. D., Um, I. H., Elshani, M., Armstrong, S. D., Penrice-Randal, R., Millar, T., Lerpiniere, C. E., Tagliavini, G., Hartley, C. S., et al. (2020). Tissue-specific immunopathology in fatal covid-19. *American Journal of Respiratory and Critical Care Medicine*, 203(2):192–201.
- [Dotolo et al., 2020] Dotolo, S., Marabotti, A., Facchiano, A., and Tagliaferri, R. (2020). A review on drug repurposing applicable to covid-19. *Briefings in bioinformatics*.
- [Duan et al., 2014] Duan, Q., Flynn, C., Niepel, M., Hafner, M., Muhlich, J. L., Fernandez, N. F., Rouillard, A. D., Tan, C. M., Chen, E. Y., Golub, T. R., et al. (2014). Lincs canvas browser: interactive web app to query, browse and interrogate lincs l1000 gene expression signatures. *Nucleic acids research*, 42(W1):W449–W460.

- [Dudley et al., 2009] Dudley, J. T., Tibshirani, R., Deshpande, T., and Butte, A. J. (2009). Disease signatures are robust across tissues and experiments. *Molecular systems biology*, 5(1):307.
- [Edwards, 2020] Edwards, A. (2020). What are the odds of finding a covid-19 drug from a lab repurposing screen? *Journal of Chemical Information and Modeling*.
- [Elmezayen et al., 2020] Elmezayen, A. D., Al-Obaidi, A., Şahin, A. T., and Yelekçi, K. (2020). Drug repurposing for coronavirus (covid-19): in silico screening of known drugs against coronavirus 3cl hydrolase and protease enzymes. *Journal of Biomolecular Structure and Dynamics*, pages 1–13.
- [Fajgenbaum and June, 2020] Fajgenbaum, D. C. and June, C. H. (2020). Cytokine storm. *New England Journal of Medicine*, 383(23):2255–2273.
- [Fang et al., 2021] Fang, M., Richardson, B., Cameron, C. M., Dazard, J.-E., and Cameron, M. J. (2021). Drug perturbation gene set enrichment analysis (dpgsea): a new transcriptomic drug screening approach. *BMC bioinformatics*, 22(1):1–14.
- [Fiscon et al., 2020] Fiscon, G., Conte, F., D’Offizi, G., Farina, L., and Paci, P. (2020). Saverunner: a network-based algorithm for drug repurposing and its application to covid-19. *arXiv preprint arXiv:2006.03110*.
- [Fiscon et al., 2021] Fiscon, G., Conte, F., Farina, L., and Paci, P. (2021). Saverunner: A network-based algorithm for drug repurposing and its application to covid-19. *PLOS Computational Biology*, 17(2):e1008686.
- [Galindez et al., 2021] Galindez, G., Matschinske, J., Rose, T. D., Sadegh, S., Salgado-Albarrán, M., Späth, J., Baumbach, J., and Pauling, J. K. (2021). Lessons from the covid-19 pandemic for advancing computational drug repurposing strategies. *Nature Computational Science*, 1(1):33–41.
- [Ganter et al., 2005] Ganter, B., Tugendreich, S., Pearson, C. I., Ayanoglu, E., Baumhueter, S., Bostian, K. A., Brady, L., Browne, L. J., Calvin, J. T., Day, G.-J., et al. (2005). Development of a large-scale chemogenomics database to improve drug candidate selection and to understand mechanisms of chemical toxicity and action. *Journal of biotechnology*, 119(3):219–244.
- [Ge et al., 2020] Ge, Y., Tian, T., Huang, S., Wan, F., Li, J., Li, S., Yang, H., Hong, L., Wu, N., Yuan, E., et al. (2020). A data-driven drug repositioning framework discovered a potential therapeutic agent targeting covid-19. *BioRxiv*.
- [Ghasemnejad-Berenji and Pashapour, 2021] Ghasemnejad-Berenji, M. and Pashapour, S. (2021). Sars-cov-2 and the possible role of raf/mek/erk pathway in viral survival: is this a potential therapeutic strategy for covid-19? *Pharmacology*, 106(1-2):119–122.
- [Gimeno et al., 2020] Gimeno, A., Mestres-Truyol, J., Ojeda-Montes, M. J., Macip, G., Saldivar-Espinoza, B., Cereto-Massagué, A., Pujadas, G., and Garcia-Vallvé, S. (2020). Prediction of novel inhibitors of the main protease (m-pro) of sars-cov-2 through consensus docking and drug reposition. *International Journal of Molecular Sciences*, 21(11):3793.
- [Gordon et al., 2020] Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O’Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., et al. (2020). A sars-cov-2 protein interaction map reveals targets for drug repurposing. *Nature*, pages 1–13.
- [Gravis et al., 2008] Gravis, G., Bladou, F., Salem, N., Goncalves, A., Esterni, B., Walz, J., Bagatini, S., Marcy, M., Brunelle, S., and Viens, P. (2008). Results from a monocentric phase ii trial of erlotinib in patients with metastatic prostate cancer. *Annals of oncology*, 19(9):1624–1628.

- [Gu et al., 2010] Gu, J., Chen, Y., Li, S., and Li, Y. (2010). Identification of responsive gene modules by network-based gene clustering and extending: application to inflammation and angiogenesis. *BMC systems biology*, 4(1):47.
- [Gurumurthy et al., 2001] Gurumurthy, S., Vasudevan, K. M., and Rangnekar, V. M. (2001). Regulation of apoptosis in prostate cancer. *Cancer and Metastasis Reviews*, 20(3):225–243.
- [Gysi et al., 2020] Gysi, D. M., Ítalo Do Valle, Zitnik, M., Ameli, A., Gan, X., Varol, O., Sanchez, H., Baron, R. M., Ghiassian, D., Loscalzo, J., and Barabási, A.-L. (2020). Network medicine framework for identifying drug repurposing opportunities for covid-19.
- [Hanczar et al., 2010] Hanczar, B., Hua, J., Sima, C., Weinstein, J., Bittner, M., and Dougherty, E. R. (2010). Small-sample precision of roc-related estimates. *Bioinformatics*, 26(6):822–830.
- [Hemmat et al., 2021] Hemmat, N., Asadzadeh, Z., Ahangar, N. K., Alemohammad, H., Najafzadeh, B., Derakhshani, A., Baghbanzadeh, A., Baghi, H. B., Javadrashid, D., Najafi, S., et al. (2021). The roles of signaling pathways in sars-cov-2 infection; lessons learned from sars-cov and mers-cov. *Archives of Virology*, pages 1–22.
- [Hoffmann et al., 2020] Hoffmann, M., Mösbauer, K., Hofmann-Winkler, H., Kaul, A., Kleine-Weber, H., Krüger, N., Gassen, N. C., Müller, M. A., Drosten, C., and Pöhlmann, S. (2020). Chloroquine does not inhibit infection of human lung cells with sars-cov-2. *Nature*, pages 1–5.
- [Hojyo et al., 2020] Hojyo, S., Uchida, M., Tanaka, K., Hasebe, R., Tanaka, Y., Murakami, M., and Hirano, T. (2020). How covid-19 induces cytokine storm with high mortality. *Inflammation and Regeneration*, 40(1):1–7.
- [Hondermarck et al., 2020] Hondermarck, H., Bartlett, N. W., and Nurcombe, V. (2020). The role of growth factor receptors in viral infections: An opportunity for drug repurposing against emerging viral diseases such as covid-19? *FASEB BioAdvances*, 2(5):296–303.
- [Hou et al., 2020] Hou, Y.-Q., Wang, Y.-Y., Wang, X.-C., Liu, Y., Zhang, C.-Z., Chen, Z.-S., Zhang, Z., Wang, W., and Kong, D.-X. (2020). Multifaceted anti-colorectal tumor effect of digoxin on hct8 and sw620 cells in vitro. *Gastroenterology Report*, 8(6):465–475.
- [Hu et al., 2019] Hu, R.-Y., Tian, X.-B., Li, B., Luo, R., Zhang, B., and Zhao, J.-M. (2019). Individualized drug repositioning for rheumatoid arthritis using weighted kolmogorov–smirnov algorithm. *Pharmacogenomics and Personalized Medicine*, 12:369.
- [Huang et al., 2020] Huang, J., Chen, J., Zhang, B., Zhu, L., and Cai, H. (2020). Evaluation of gene–drug common module identification methods using pharmacogenomics data. *Briefings in Bioinformatics*.
- [Huang et al., 2015] Huang, S., Mao, J., Wei, B., and Pei, G. (2015). The anti-spasticity drug baclofen alleviates collagen-induced arthritis and regulates dendritic cells. *Journal of cellular physiology*, 230(7):1438–1447.
- [Igarashi et al., 2015] Igarashi, Y., Nakatsu, N., Yamashita, T., Ono, A., Ohno, Y., Urushidani, T., and Yamada, H. (2015). Open tg-gates: a large-scale toxicogenomics database. *Nucleic acids research*, 43(D1):D921–D927.
- [Iorio et al., 2015] Iorio, F., Shrestha, R. L., Levin, N., Boilot, V., Garnett, M. J., Saez-Rodriguez, J., and Draviam, V. M. (2015). A semi-supervised approach for refining transcriptional signatures of drug response and repositioning predictions. *PloS one*, 10(10):e0139446.



- [Jarada et al., 2020] Jarada, T. N., Rokne, J. G., and Alhajj, R. (2020). A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. *Journal of Cheminformatics*, 12(1):1–23.
- [Jaundoo and Craddock, 2020] Jaundoo, R. and Craddock, T. J. (2020). Drugpath: The drug gene pathway meta-database. *International journal of molecular sciences*, 21(9):3171.
- [Jin et al., 2020] Jin, W., Barzilay, R., and Jaakkola, T. (2020). Modeling drug combinations based on molecular structures and biological targets. *arXiv preprint arXiv:2011.04651*.
- [Kacan et al., 2016] Kacan, T., Nayir, E., Altun, A., Kilickap, S., Babacan, N. A., Ataseven, H., and Kaya, T. (2016). Antitumor activity of sorafenib on colorectal cancer. *Journal of Oncological Sciences*, 2(2-3):53–57.
- [Kawahara et al., 2015] Kawahara, T., Kashiwagi, E., Ide, H., Li, Y., Zheng, Y., Ishiguro, H., and Miyamoto, H. (2015). The role of nfatc1 in prostate cancer progression: cyclosporine a and tacrolimus inhibit cell proliferation, migration, and invasion. *The Prostate*, 75(6):573–584.
- [Kim et al., 2018] Kim, K.-J., Rai, N., Kim, M., and Tagkopoulos, I. (2018). A network-based model for drug repurposing in rheumatoid arthritis. *BioRxiv*, page 335679.
- [Klann et al., 2020] Klann, K., Bojkova, D., Tascher, G., Ciesek, S., Münch, C., and Cinatl, J. (2020). Growth factor receptor signaling inhibition prevents sars-cov-2 replication. *Molecular cell*, 80(1):164–174.
- [Kow and Hasan, 2020] Kow, C. and Hasan, S. (2020). Prochlorperazine for nausea and vomiting accompanied covid-19. *Journal of Gastroenterology and Hepatology (Australia)*.
- [Kruse and Vanijcharoenkarn, 2018] Kruse, R. and Vanijcharoenkarn, K. (2018). Drug repurposing to treat asthma and allergic disorders: Progress and prospects. *Allergy*, 73(2):313–322.
- [Kuleshov et al., 2016] Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research*, 44(W1):W90–W97.
- [Kurose et al., 2001] Kurose, A., Yoshida, W., Yoshida, M., and Sawai, T. (2001). Effects of paclitaxel on cultured synovial cells from patients with rheumatoid arthritis. *Cytometry: The Journal of the International Society for Analytical Cytology*, 44(4):349–354.
- [Kuruppu et al., 1998] Kuruppu, D., Christophi, C., Bertram, J. F., and O'BRIEN, P. E. (1998). Tamoxifen inhibits colorectal cancer metastases in the liver: a study in a murine model. *Journal of gastroenterology and hepatology*, 13(5):521–527.
- [Laing et al., 2020] Laing, A. G., Lorenc, A., Del Barrio, I. D. M., Das, A., Fish, M., Monin, L., Muñoz-Ruiz, M., McKenzie, D. R., Hayday, T. S., Francos-Quijorna, I., et al. (2020). A dynamic covid-19 immune signature includes associations with poor prognosis. *Nature medicine*, 26(10):1623–1635.
- [Lamb, 2007] Lamb, J. (2007). The connectivity map: a new tool for biomedical research. *Nature reviews cancer*, 7(1):54–60.
- [Lamb et al., 2006] 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(5795):1929–1935.

- [Lee et al., 2017] Lee, C.-S., Kim, H., Yu, J., Yu, S. H., Ban, S., Oh, S., Jeong, D., Im, J., Baek, M. J., and Kim, T. H. (2017). Doxorubicin-loaded oligonucleotide conjugated gold nanoparticles: A promising in vivo drug delivery system for colorectal cancer therapy. *European journal of medicinal chemistry*, 142:416–423.
- [Lee et al., 2016] Lee, H. Y., Rhee, C. K., Kang, J. Y., Park, C. K., Lee, S. Y., Kwon, S. S., Kim, Y. K., and Yoon, H. K. (2016). Effect of intranasal rosiglitazone on airway inflammation and remodeling in a murine model of chronic asthma. *The Korean journal of internal medicine*, 31(1):89.
- [Levin et al., 2020] Levin, J. M., Oprea, T. I., Davidovich, S., Clozel, T., Overington, J. P., Vanhaelen, Q., Cantor, C. R., Bischof, E., and Zhavoronkov, A. (2020). Artificial intelligence, drug repurposing and peer review. *Nature Biotechnology*, 38(10):1127–1131.
- [Li et al., 2016] Li, J., Zheng, S., Chen, B., Butte, A. J., Swamidass, S. J., and Lu, Z. (2016). A survey of current trends in computational drug repositioning. *Briefings in bioinformatics*, 17(1):2–12.
- [Li et al., 2021] Li, L., Cui, N., Hao, T., Zou, J., Jiao, W., Yi, K., and Yu, W. (2021). Statins use and the prognosis of colorectal cancer: a meta-analysis. *Clinics and Research in Hepatology and Gastroenterology*, 45(5):101588.
- [Li et al., 2020] Li, S., Zhang, Y., Guan, Z., Li, H., Ye, M., Chen, X., Shen, J., Zhou, Y., Shi, Z.-L., Zhou, P., et al. (2020). Sars-cov-2 triggers inflammatory responses and cell death through caspase-8 activation. *Signal transduction and targeted therapy*, 5(1):1–10.
- [Lim and Pavlidis, 2019] Lim, N. and Pavlidis, P. (2019). Evaluation of connectivity map shows limited reproducibility in drug repositioning. *BioRxiv*, page 845693.
- [Lin et al., 2020] Lin, K., Li, L., Dai, Y., Wang, H., Teng, S., Bao, X., Lu, Z. J., and Wang, D. (2020). A comprehensive evaluation of connectivity methods for 11000 data. *Briefings in bioinformatics*, 21(6):2194–2205.
- [Liu et al., 2006] Liu, L., Cao, Y., Chen, C., Zhang, X., McNabola, A., Wilkie, D., Wilhelm, S., Lynch, M., and Carter, C. (2006). Sorafenib blocks the raf/mek/erk pathway, inhibits tumor angiogenesis, and induces tumor cell apoptosis in hepatocellular carcinoma model plc/prf/5. *Cancer research*, 66(24):11851–11858.
- [Liu et al., 2019] Liu, N., Guo, Z., Xia, X., Liao, Y., Zhang, F., Huang, C., Liu, Y., Deng, X., Jiang, L., Wang, X., et al. (2019). Auranofin lethality to prostate cancer includes inhibition of proteasomal deubiquitinases and disrupted androgen receptor signaling. *European journal of pharmacology*, 846:1–11.
- [Lokhande et al., 2020] Lokhande, K. B., Doiphode, S., Vyas, R., and Swamy, K. V. (2020). Molecular docking and simulation studies on sars-cov-2 mpro reveals mitoxantrone, leucovorin, birinapant, and dynasore as potent drugs against covid-19. *Journal of Biomolecular Structure and Dynamics*, pages 1–12.
- [Lotfi Shahreza et al., 2018] Lotfi Shahreza, M., Ghadiri, N., Mousavi, S. R., Varshosaz, J., and Green, J. R. (2018). A review of network-based approaches to drug repositioning. *Briefings in bioinformatics*, 19(5):878–892.
- [Loucera et al., 2020] Loucera, C., Esteban-Medina, M., Rian, K., Falco, M. M., Dopazo, J., and Peña-Chilet, M. (2020). Drug repurposing for covid-19 using machine learning and mechanistic models of signal transduction circuits related to sars-cov-2 infection. *Signal transduction and targeted therapy*, 5(1):1–3.

- [Low et al., 2020] Low, Z. Y., Farouk, I. A., and Lal, S. K. (2020). Drug repositioning: New approaches and future prospects for life-debilitating diseases and the covid-19 pandemic outbreak. *Viruses*, 12(9):1058.
- [Lucchetta, 2020] Lucchetta, M. and Pellegrini, M. (2020). Finding disease modules for cancer and covid-19 in gene co-expression networks with the core&peel method. *Sci. Rep.*, 10:17628. <https://doi.org/10.1038/s41598-020-74705-6>.
- [Luna et al., 2020] Luna, B., Ramírez, M., and Galán, E. (2020). Network analysis and disease subnets for the sars-cov-2/human interactome.
- [Ma et al., 2020] Ma, M., Wang, X., Liu, N., Shan, F., and Feng, Y. (2020). Low-dose naltrexone inhibits colorectal cancer progression and promotes apoptosis by increasing m1-type macrophages and activating the bax/bcl-2/caspase-3/parp pathway. *International immunopharmacology*, 83:106388.
- [Mahapatra et al., 2020] Mahapatra, S., Nath, P., Chatterjee, M., Das, N., Kalita, D., Roy, P., and Satapathi, S. (2020). Repurposing therapeutics for covid-19: Rapid prediction of commercially available drugs through machine learning and docking. *medRxiv*.
- [Mall et al., 2020] Mall, R., Elbasir, A., Al Meer, H., Chawla, S., and Ullah, E. (2020). Data-driven drug repurposing for covid-19. *ChemRxiv chemrxiv*, 12661103:v1.
- [Mallapaty, 2020] Mallapaty, S. (2020). Mini organs reveal how the coronavirus ravages the body. *Nature*, 583(7814):15–16.
- [Mansilla et al., 2020] Mansilla, A., Soto, J., Henriquez, C., Philp, A. R., Mall, M. A., Sarmiento, J., and Flores, C. A. (2020). Tamoxifen reduces inflammatory infiltration of neutrophils in the airways. *bioRxiv*.
- [Martin et al., 1984] Martin, M., McKenna, F., Bird, H., Surrall, K., Dixon, J., and Wright, V. (1984). Captopril: a new treatment for rheumatoid arthritis? *The Lancet*, 323(8390):1325–1328.
- [Martin-Garcia et al., 2003] Martin-Garcia, C., Hinojosa, M., Berges, P., Camacho, E., Garcia-Rodriguez, R., and Alfaya, T. (2003). Celecoxib, a highly selective cox-2 inhibitor, is safe in aspirin-induced asthma patients. *Journal of investigational allergology & clinical immunology*, 13(1):20–25.
- [Merad and Martin, 2020] Merad, M. and Martin, J. C. (2020). Pathological inflammation in patients with covid-19: a key role for monocytes and macrophages. *Nature Reviews Immunology*, 20(6):355–362.
- [Michaelson et al., 2009] Michaelson, M. D., Regan, M., Oh, W., Kaufman, D., Olivier, K., Michaelson, S., Spicer, B., Gurski, C., Kantoff, P., and Smith, M. (2009). Phase ii study of sunitinib in men with advanced prostate cancer. *Annals of Oncology*, 20(5):913–920.
- [Mirabelli et al., 2020] Mirabelli, C., Wotring, J. W., Zhang, C. J., McCarty, S. M., Fursmidt, R., Kadambi, N. S., Amin, A. T., O’Meara, T. R., Pretto-Kernahan, C. D., Spence, J. R., et al. (2020). Morphological cell profiling of sars-cov-2 infection identifies drug repurposing candidates for covid-19. *bioRxiv*.
- [Moore and June, 2020] Moore, J. B. and June, C. H. (2020). Cytokine release syndrome in severe covid-19. *Science*, 368(6490):473–474.

- [Morgan et al., 2008] Morgan, T. M., Pitts, T. E., Gross, T. S., Poliachik, S. L., Vessella, R. L., and Corey, E. (2008). Rad001 (everolimus) inhibits growth of prostate cancer in the bone and the inhibitory effects are increased by combination with docetaxel and zoledronic acid. *The Prostate*, 68(8):861–871.
- [Mousavi et al., 2020] Mousavi, S. Z., Rahmanian, M., and Sami, A. (2020). A connectivity map-based drug repurposing study and integrative analysis of transcriptomic profiling of sars-cov-2 infection. *Infection, Genetics and Evolution*, 86:104610.
- [Muratov and Zakharov, 2020] Muratov, E. and Zakharov, A. (2020). Viribus unitis: Drug combinations as a treatment against covid-19. *ChemRxiv*.
- [Musa et al., 2018] Musa, A., Ghoraie, L. S., Zhang, S.-D., Glazko, G., Yli-Harja, O., Dehmer, M., Haibe-Kains, B., and Emmert-Streib, F. (2018). A review of connectivity map and computational approaches in pharmacogenomics. *Briefings in bioinformatics*, 19(3):506–523.
- [Nabirovichkin et al., 2020] Nabirovichkin, S., Peluffo, A. E., Rinaudo, P., Yu, J., Hajj, R., and Cohen, D. (2020). Next-generation drug repurposing using human genetics and network biology. *Current Opinion in Pharmacology*.
- [Najgebauer et al., 2020] Najgebauer, H., Yang, M., Francies, H. E., Pacini, C., Stronach, E. A., Garnett, M. J., Saez-Rodriguez, J., and Iorio, F. (2020). Collector: genomics-guided selection of cancer in vitro models. *Cell Systems*, 10(5):424–432.
- [Nam et al., 2020] Nam, Y., Yun, J.-S., Lee, S. M., Park, J. W., Chen, Z., Lee, B., Verma, A., Ning, X., Shen, L., and Kim, D. (2020). Network reinforcement driven drug repurposing for covid-19 by exploiting disease-gene-drug associations. *arXiv preprint arXiv:2008.05377*.
- [Nandi et al., 2021] Nandi, U., Onyesom, I., and Douroumis, D. (2021). Anti-cancer activity of sirolimus loaded liposomes in prostate cancer cell lines. *Journal of Drug Delivery Science and Technology*, 61:102200.
- [Nazeam et al., 2020] Nazeam, J., Mohammed, E. Z., Raafat, M., Houssein, M., Elkafoury, A., Hamdy, D., and Jamil, L. (2020). Based on principles and insights of covid-19 epidemiology, genome sequencing, and pathogenesis: Retrospective analysis of sinigrin and prolixinrx (fluphenazine) provides off-label drug candidates. *SLAS DISCOVERY: Advancing the Science of Drug Discovery*, 25(10):1123–1140.
- [Newman et al., 2020] Newman, R., Pietras, C. M., Qu, F., Sapashnik, D., Kofman, L., Butze, S., Ocitti, F., Fried, I., and Slonim, D. (2020). Improving cell-specific drug connectivity mapping with collaborative filtering. *bioRxiv*.
- [Nosengo, 2016] Nosengo, N. (2016). Can you teach old drugs new tricks? *Nature News*, 534(7607):314.
- [Nowak-Sliwinska et al., 2019] Nowak-Sliwinska, P., Scapozza, L., and Altaba, A. R. (2019). Drug repurposing in oncology: Compounds, pathways, phenotypes and computational approaches for colorectal cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1871(2):434–454.
- [Ogrendik, 2014] Ogrendik, M. (2014). Antibiotics for the treatment of rheumatoid arthritis. *International journal of general medicine*, 7:43.
- [Otręba et al., 2020] Otręba, M., Kośmider, L., and Rzepecka-Stojko, A. (2020). Antiviral activity of chlorpromazine, fluphenazine, perphenazine, prochlorperazine, and thioridazine towards rna-viruses. a review. *European journal of pharmacology*, 887:173553.

- [Patra et al., 2011] Patra, N., De, U., Kang, J.-A., Kim, J. M., Ahn, M. Y., Lee, J., Jung, J. H., Chung, H. Y., Moon, H. R., and Kim, H. S. (2011). A novel epoxypropoxy flavonoid derivative and topoisomerase ii inhibitor, mhy336, induces apoptosis in prostate cancer cells. *European journal of pharmacology*, 658(2-3):98–107.
- [Peclat et al., 2020] Peclat, T. R., Shi, B., Varga, J., and Chini, E. N. (2020). The nadase enzyme cd38: an emerging pharmacological target for systemic sclerosis, systemic lupus erythematosus and rheumatoid arthritis. *Current opinion in rheumatology*, 32(6):488–496.
- [Pellegrini et al., 2016] Pellegrini, M., Baglioni, M., and Geraci, F. (2016). Protein complex prediction for large protein protein interaction networks with the core&peel method. *BMC bioinformatics*, 17(12):372.
- [Petralia et al., 2019] Petralia, M. C., Mazzon, E., Basile, M. S., Cutuli, M., Di Marco, R., Scandurra, F., Saraceno, A., Fagone, P., Nicoletti, F., and Mangano, K. (2019). Effects of treatment with the hypomethylating agent 5-aza-2'-deoxycytidine in murine type ii collagen-induced arthritis. *Pharmaceuticals*, 12(4):174.
- [Poleksic and Xie, 2019] Poleksic, A. and Xie, L. (2019). Database of adverse events associated with drugs and drug combinations. *Scientific reports*, 9(1):1–9.
- [Polson et al., 1986] Polson, R. J., Jawad, A. S., Bomford, A., Berry, H., and Williams, R. (1986). Treatment of rheumatoid arthritis with desferrioxamine. *QJM: An International Journal of Medicine*, 61(3):1153–1158.
- [Pruimboom, 2020] Pruumboom, L. (2020). Methylation pathways and sars-cov-2 lung infiltration and cell membrane-virus fusion are both subject to epigenetics. *Frontiers in cellular and infection microbiology*, 10:290.
- [Pushpakom et al., 2019] 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. *Nature reviews Drug discovery*, 18(1):41–58.
- [Ramaraju et al., 2013] Ramaraju, K., Krishnamurthy, S., Maamidi, S., Kaza, A. M., and Balasubramaniam, N. (2013). Is serum cholesterol a risk factor for asthma? *Lung India: Official Organ of Indian Chest Society*, 30(4):295.
- [Raymonda et al., 2020] Raymonda, M. H., Ciesla, J. H., Monaghan, M., Leach, J., Asantewaa, G., Smorodintsev-Schiller, L. A., Lutz, M. M., Schafer, X. L., Takimoto, T., Dewhurst, S., et al. (2020). Pharmacologic profiling reveals lapatinib as a novel antiviral against sars-cov-2 in vitro. *bioRxiv*.
- [Re and Valentini, 2013] Re, M. and Valentini, G. (2013). Network-based drug ranking and repositioning with respect to drugbank therapeutic categories. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 10(6):1359–1371.
- [Riva et al., 2020a] Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Burgstaller, S., Pache, L., De Jesus, P., Hull, M. V., Chang, M., et al. (2020a). A large-scale drug repositioning survey for sars-cov-2 antivirals. *bioRxiv*.
- [Riva et al., 2020b] Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Pache, L., Burgstaller-Muehlbacher, S., De Jesus, P. D., Teriete, P., Hull, M. V., et al. (2020b). Discovery of sars-cov-2 antiviral drugs through large-scale compound repurposing. *Nature*, 586(7827):113–119.

- [Romo-García et al., 2020] Romo-García, M. F., Zapata-Zuñiga, M., Enciso-Moreno, J. A., and Castañeda-Delgado, J. E. (2020). The role of estrogens in rheumatoid arthritis physiopathology. In *Rheumatoid Arthritis-Other Perspectives towards a Better Practice*. IntechOpen.
- [Rubin et al., 2020] Rubin, D., Chan-Tack, K., Farley, J., and Sherwat, A. (2020). Fda approval of remdesivir—a step in the right direction. *New England Journal of Medicine*, 383(27):2598–2600.
- [Ruiz et al., 2020] Ruiz, C., Zitnik, M., and Leskovec, J. (2020). Discovery of disease treatment mechanisms through the multiscale interactome. *bioRxiv*.
- [Sachdeva et al., 2020] Sachdeva, C., Wadhwa, A., Kumari, A., Hussain, F., Jha, P., and Kaushik, N. K. (2020). In silico potential of approved antimalarial drugs for repurposing against covid-19. *OMICS: A Journal of Integrative Biology*, 24(10):568–580.
- [Sadegh et al., 2020] Sadegh, S., Matschinske, J., Blumenthal, D. B., Galindez, G., Kacprowski, T., List, M., Nasirigerdeh, R., Oubounyt, M., Pichlmair, A., Rose, T. D., and et al. (2020). Exploring the sars-cov-2 virus-host-drug interactome for drug repurposing. *Nature Communications*, 11(1).
- [Sam and Athri, 2019] Sam, E. and Athri, P. (2019). Web-based drug repurposing tools: a survey. *Briefings in bioinformatics*, 20(1):299–316.
- [Samei et al., 2016] Samei, L., Yaling, P., Lihua, Y., Yan, Z., and Shuyan, J. (2016). Effects and mechanism of imatinib in inhibiting colon cancer cell proliferation. *Medical science monitor: international medical journal of experimental and clinical research*, 22:4126.
- [Sanyaolu et al., 2020] Sanyaolu, A., Okorie, C., Marinkovic, A., Patidar, R., Younis, K., Desai, P., Hosein, Z., Padda, I., Mangat, J., and Altaf, M. (2020). Comorbidity and its impact on patients with covid-19. *SN comprehensive clinical medicine*, pages 1–8.
- [Sarkari, 2007] Sarkari, S. E. B. (2007). Comparative efficacy of dexamethasone versus hydrocortisone in severe acute pediatric asthma. *Iranian Journal of Allergy, Asthma and Immunology*, pages 159–160.
- [Schäfer and Baric, 2017] Schäfer, A. and Baric, R. S. (2017). Epigenetic landscape during coronavirus infection. *Pathogens*, 6(1):8.
- [Scott et al., 2017] Scott, A. J., Song, E.-K., Bagby, S., Purkey, A., McCarter, M., Gajdos, C., Quackenbush, K. S., Cross, B., Pitts, T. M., Tan, A. C., et al. (2017). Evaluation of the efficacy of dasatinib, a src/abl inhibitor, in colorectal cancer cell lines and explant mouse model. *PLoS One*, 12(11):e0187173.
- [Seo et al., 2020] Seo, S., Park, J. W., An, D., Yoon, J., Paik, H., and Hwang, S. (2020). Supercomputer-aided Drug Repositioning at Scale: Virtual Screening for SARS-CoV-2 Protease Inhibitor. *ChemRxiv*.
- [Shah et al., 2020] Shah, B., Modi, P., and Sagar, S. R. (2020). In silico studies on therapeutic agents for covid-19: Drug repurposing approach. *Life Sciences*, page 117652.
- [Singal et al., 2015] Singal, R., Ramachandran, K., Gordian, E., Quintero, C., Zhao, W., and Reis, I. M. (2015). Phase i/ii study of azacitidine, docetaxel, and prednisone in patients with metastatic castration-resistant prostate cancer previously treated with docetaxel-based therapy. *Clinical genitourinary cancer*, 13(1):22–31.
- [Singh et al., 2020] Singh, T. U., Parida, S., Lingaraju, M. C., Kesavan, M., Kumar, D., and Singh, R. K. (2020). Drug repurposing approach to fight covid-19. *Pharmacological Reports*, pages 1–30.

- [Smirnov et al., 2016] Smirnov, P., Safikhani, Z., El-Hachem, N., Wang, D., She, A., Olsen, C., Freeman, M., Selby, H., Gendoo, D. M., Grossmann, P., et al. (2016). Pharmacogx: an r package for analysis of large pharmacogenomic datasets. *Bioinformatics*, 32(8):1244–1246.
- [Sompornrattanaphan et al., 2020] Sompornrattanaphan, M., Krikeerati, T., Wongsas, C., Thongn-garm, T., and Yampayon, K. (2020). Successful deferasirox rechallenge and treating through reaction in a patient with challenge-proven mild immediate reaction: A case report. *Journal of Asthma and Allergy*, 13:557.
- [Stephenson et al., 2021] Stephenson, E., Reynolds, G., Botting, R. A., Calero-Nieto, F. J., Morgan, M. D., Tuong, Z. K., Bach, K., Sungnak, W., Worlock, K. B., Yoshida, M., et al. (2021). Single-cell multi-omics analysis of the immune response in covid-19. *Nature Medicine*, pages 1–13.
- [Stolfi et al., 2020] Stolfi, P., Manni, L., Soligo, M., Vergni, D., and Tieri, P. (2020). Designing a network proximity-based drug repurposing strategy for covid-19. *Frontiers in cell and developmental biology*, 8.
- [Subramanian et al., 2017] 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(6):1437–1452.
- [Subramanian et al., 2005] Subramanian, S., Tovey, M., Afentoulis, M., Krogstad, A., Vandembark, A. A., and Offner, H. (2005). Ethinyl estradiol treats collagen-induced arthritis in dba/11acj mice by inhibiting the production of  $\text{tnf-}\alpha$  and  $\text{il-1}\beta$ . *Clinical Immunology*, 115(2):162–172.
- [Suke et al., 2013] Suke, S. G., Negi, H., Mediratta, P., Banerjee, B., and Sharma, K. (2013). Anti-arthritis and anti-inflammatory activity of combined pioglitazone and prednisolone on adjuvant-induced arthritis. *European journal of pharmacology*, 718(1-3):57–62.
- [Svoboda et al., 2019] Svoboda, D. L., Saddler, T., and Auerbach, S. S. (2019). An overview of national toxicology program’s toxicogenomic applications: Drugmatrix and toxfx. *Advances in Computational Toxicology*, pages 141–157.
- [Taguchi and Turki, 2020] Taguchi, Y. and Turki, T. (2020). A new advanced in silico drug discovery method for novel coronavirus (sars-cov-2) with tensor decomposition-based unsupervised feature extraction. *Preprints*.
- [Tang et al., 2020] Tang, Y., Liu, J., Zhang, D., Xu, Z., Ji, J., and Wen, C. (2020). Cytokine storm in covid-19: the current evidence and treatment strategies. *Frontiers in immunology*, 11:1708.
- [Terrazzano et al., 2020] Terrazzano, G., Rubino, V., Palatucci, A. T., Giovazzino, A., Carriero, F., and Ruggiero, G. (2020). An open question: is it rational to inhibit the mtor-dependent pathway as covid-19 therapy? *Frontiers in pharmacology*, 11:856.
- [Thiem et al., 2013] Thiem, S., Pierce, T. P., Palmieri, M., Putoczki, T. L., Buchert, M., Preaudet, A., Farid, R. O., Love, C., Catimel, B., Lei, Z., et al. (2013). mtorc1 inhibition restricts inflammation-associated gastrointestinal tumorigenesis in mice. *The Journal of clinical investigation*, 123(2).
- [Tomazou et al., 2021] Tomazou, M., Bourdakou, M. M., Minadakis, G., Zachariou, M., Oulas, A., Karatzas, E., Loizidou, E., Kakouri, A., Christodoulou, C., Savva, K., Zanti, M., Onisiforou, A., Afxenti, S., Richter, J., Christodoulou, C. G., Kyprianou, T., Kolios, G., Dietis, N., and Spyrou, G. M. (2021). Multi-omics data integration and network-based analysis drives a multiplex drug repurposing approach to a shortlist of candidate drugs against COVID-19. *Briefings in Bioinformatics*. bbab114.

- [Trezza et al., 2020] Trezza, A., Iovinelli, D., Santucci, A., Prischi, F., and Spiga, O. (2020). An integrated drug repurposing strategy for the rapid identification of potential sars-cov-2 viral inhibitors. *Scientific reports*, 10(1):1–8.
- [Turanli et al., 2018] Turanli, B., Grötli, M., Boren, J., Nielsen, J., Uhlen, M., Arga, K. Y., and Mardinoglu, A. (2018). Drug repositioning for effective prostate cancer treatment. *Frontiers in Physiology*, 9:500.
- [Tutuncuoglu et al., 2020] Tutuncuoglu, B., Cakir, M., Batra, J., Bouhaddou, M., Eckhardt, M., Gordon, D. E., and Krogan, N. J. (2020). The landscape of human cancer proteins targeted by sars-cov-2. *Cancer discovery*, 10(7):916–921.
- [Ulitsky et al., 2010] Ulitsky, I., Krishnamurthy, A., Karp, R. M., and Shamir, R. (2010). Degas: de novo discovery of dysregulated pathways in human diseases. *PloS one*, 5(10).
- [Venkataraman and Frieman, 2017] Venkataraman, T. and Frieman, M. B. (2017). The role of epidermal growth factor receptor (egfr) signaling in sars coronavirus-induced pulmonary fibrosis. *Antiviral research*, 143:142–150.
- [Vicentini et al., 2003] Vicentini, C., Festuccia, C., Gravina, G. L., Angelucci, A., Marronaro, A., and Bologna, M. (2003). Prostate cancer cell proliferation is strongly reduced by the epidermal growth factor receptor tyrosine kinase inhibitor zd1839 in vitro on human cell lines and primary cultures. *Journal of cancer research and clinical oncology*, 129(3):165–174.
- [Vlaic et al., 2018] Vlaic, S., Conrad, T., Tokarski-Schnelle, C., Gustafsson, M., Dahmen, U., Guthke, R., and Schuster, S. (2018). Modulediscoverer: identification of regulatory modules in protein-protein interaction networks. *Scientific reports*, 8(1):1–11.
- [Wagner et al., 2009] Wagner, M., Roh, V., Strehlen, M., Laemmle, A., Stroka, D., Egger, B., Trochsler, M., Hunt, K. K., Candinas, D., and Vorburger, S. A. (2009). Effective treatment of advanced colorectal cancer by rapamycin and 5-fu/oxaliplatin monitored by timp-1. *Journal of gastrointestinal surgery*, 13(10):1781–1790.
- [Wang, 2020] Wang, J. (2020). Fast identification of possible drug treatment of coronavirus disease-19 (covid-19) through computational drug repurposing study. *Journal of Chemical Information and Modeling*.
- [Wang et al., 2020a] Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., and Xiao, G. (2020a). Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-ncov) in vitro. *Cell research*, 30(3):269–271.
- [Wang et al., 2020b] Wang, P.-G., Tang, D.-J., Hua, Z., Wang, Z., and An, J. (2020b). Sunitinib reduces the infection of sars-cov, mers-cov and sars-cov-2 partially by inhibiting ap2m1 phosphorylation. *Cell discovery*, 6(1):1–5.
- [Weisberg et al., 2020] Weisberg, E., Parent, A., Yang, P. L., Sattler, M., Liu, Q., Liu, Q., Wang, J., Meng, C., Buhrlage, S. J., Gray, N., et al. (2020). Repurposing of kinase inhibitors for treatment of covid-19. *Pharmaceutical research*, 37(9):1–29.
- [Wen et al., 2019] Wen, H.-Y., Wang, J., Zhang, S.-X., Luo, J., Zhao, X.-C., Zhang, C., Wang, C.-H., Hu, F.-Y., Zheng, X.-J., Cheng, T., et al. (2019). Low-dose sirolimus immunoregulation therapy in patients with active rheumatoid arthritis: A 24-week follow-up of the randomized, open-label, parallel-controlled trial. *Journal of immunology research*, 2019.



- [Whang et al., 2013] Whang, Y. E., Armstrong, A. J., Rathmell, W. K., Godley, P. A., Kim, W. Y., Pruthi, R. S., Wallen, E. M., Crane, J. M., Moore, D. T., Grigson, G., et al. (2013). A phase ii study of lapatinib, a dual egfr and her-2 tyrosine kinase inhibitor, in patients with castration-resistant prostate cancer. *Urologic Oncology: Seminars and Original Investigations*, 31(1):82–86.
- [Wickham et al., 2016] Wickham, H., Chang, W., and Wickham, M. H. (2016). Package ‘ggplot2’. *Create Elegant Data Visualisations Using the Grammar of Graphics. Version*, 2(1):1–189.
- [Wishart et al., 2018] 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 research*, 46(D1):D1074–D1082.
- [Wong and Leong, 2004] Wong, W. F. and Leong, K. P. (2004). Tyrosine kinase inhibitors: a new approach for asthma. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*, 1697(1-2):53–69.
- [Xu et al., 2020] Xu, H., Liu, L., Li, W., Zou, D., Yu, J., Wang, L., and Wong, C. C. (2020). Transcription factors in colorectal cancer: molecular mechanism and therapeutic implications. *Oncogene*, pages 1–15.
- [Xue et al., 2018] Xue, H., Li, J., Xie, H., and Wang, Y. (2018). Review of drug repositioning approaches and resources. *International journal of biological sciences*, 14(10):1232.
- [Yang et al., 2015] Yang, I. V., Pedersen, B. S., Liu, A., O’Connor, G. T., Teach, S. J., Kattan, M., Misiak, R. T., Gruchalla, R., Steinbach, S. F., Szeffler, S. J., et al. (2015). Dna methylation and childhood asthma in the inner city. *Journal of Allergy and Clinical Immunology*, 136(1):69–80.
- [Zaborowska et al., 2012] Zaborowska, M., Szmit, S., and Szczylik, C. (2012). Sorafenib in progressive castrate-resistant prostate cancer. can we talk about a new therapeutic option? *Archives of medical science: AMS*, 8(3):528.
- [Zeng et al., 2020] Zeng, X., Song, X., Ma, T., Pan, X., Zhou, Y., Hou, Y., Zhang, Z., Li, K., Karypis, G., and Cheng, F. (2020). Repurpose open data to discover therapeutics for covid-19 using deep learning. *Journal of proteome research*, 19(11):4624–4636.
- [Zhang et al., 2018] Zhang, L., Zhang, H., Ai, H., Hu, H., Li, S., Zhao, J., and Liu, H. (2018). Applications of machine learning methods in drug toxicity prediction. *Current topics in medicinal chemistry*, 18(12):987–997.
- [Zhang et al., 2020a] Zhang, T., Liu, D., Tian, D., and Xia, L. (2020a). The roles of nausea and vomiting in covid-19: did we miss something? *Journal of Microbiology, Immunology and Infection*.
- [Zhang et al., 2020b] Zhang, Z., Zhou, L., Xie, N., Nice, E. C., Zhang, T., Cui, Y., and Huang, C. (2020b). Overcoming cancer therapeutic bottleneck by drug repurposing. *Signal transduction and targeted therapy*, 5(1):1–25.
- [Zhou et al., 2020a] Zhou, L., Huntington, K., Zhang, S., Carlsen, L., So, E.-Y., Parker, C., Sahin, I., Safran, H., Kamle, S., Lee, C.-M., et al. (2020a). Mek inhibitors reduce cellular expression of ace2, perk, prb while stimulating nk-mediated cytotoxicity and attenuating inflammatory cytokines relevant to sars-cov-2 infection. *Oncotarget*, 11(46):4201.
- [Zhou et al., 2020b] Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., and Cheng, F. (2020b). Network-based drug repurposing for novel coronavirus 2019-ncov/sars-cov-2. *Cell discovery*, 6(1):1–18.

- [Zhou et al., 2020c] Zhou, Y., Wang, F., Tang, J., Nussinov, R., and Cheng, F. (2020c). Artificial intelligence in covid-19 drug repurposing. *The Lancet Digital Health*.
- [Zitnik et al., 2018a] Zitnik, M., Agrawal, M., and Leskovec, J. (2018a). Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*, 34(13):i457–i466.
- [Zitnik et al., 2018b] Zitnik, M., Leskovec, J., et al. (2018b). Prioritizing network communities. *Nature communications*, 9(1):1–9.

## A Supplementary files

- S1 **DrugMerge-asthma-table1.csv**. The table reports the ranking function, the filter, drug perturbation data set, algorithmic configuration attaining best performance on the asthma DEG dataset. It is reported the reverse hits ranking and precision at 20, with corresponding z-scores and p-values. The csv file uses ';' as column separator.
- S2 **DrugMerge-arthritis-table1.csv** Table similar to S1 for Rheumatoid Arthritis DEG dataset.
- S3 **DrugMerge-colorectal-cancer-table1.csv** Table similar to S1 for Colorectal Cancer DEG dataset.
- S4 **DrugMerge-prostate-cancer-table1.csv** Table similar to S1 for Prostate Cancer DEG dataset.
- S5 **DrugMerge-covid19-table4.csv** Table similar to S1 for Covid19 DEG datasets.
- S5.1 **Table-algorithm-alias.csv** Table mapping synthetic algorithms identifiers in Figures 7 and 8 with algorithmic configurations in Table S5.
- S6 **Summary-four-diseases-CMAPalg-vs-DrugMerge.xlsx** Table reporting comparative performance of CMAP and DrugMerge on four benchmark data sets (Asthma, Rheumatoid Arthritis, Prostate cancer, and Colorectal cancer). It is reported the reverse hits ranking and precision at 20, with corresponding z-scores and p-values.
- S7 **Drug-lists-final.xlsx** Drug rankings with best rhr pvalue performance for DrugMerge on four benchmark data sets (Asthma, Rheumatoid Arthritis, Prostate cancer, and Colorectal cancer), and covid19 (both with FDA approved only drugs, or all drugs).
- S8 **Summary-table-comparative-results.xlsx** Table reporting comparative performance of several published drug rankings on covid19. It is reported the reverse hits ranking and precision at 20, with corresponding z-scores and p-values.

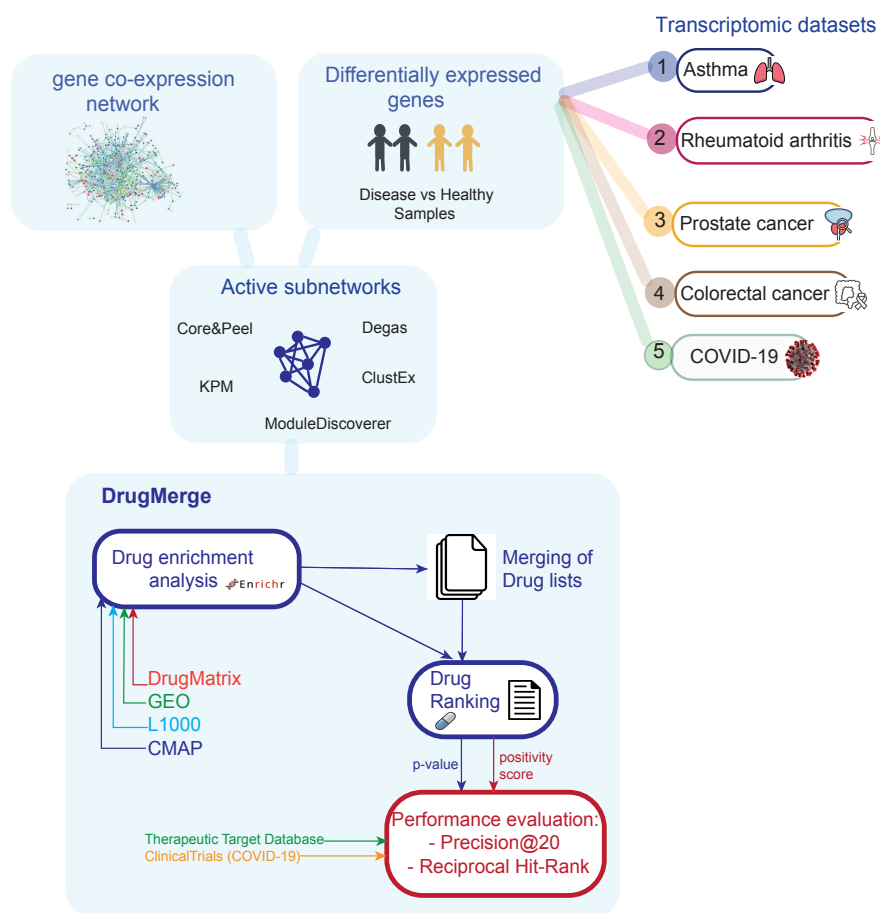


Figure 1: Overview of the workflow. The first part of the figure shows how we generated the data inputs of DrugMerge. These data have been created in [Lucchetta, 2020], and we reported them here for sake of completeness. We used several transcriptomic data to identify differentially expressed genes and the gene co-expression network to find one active subnetwork for each method used (Core&Peel, ModuleDiscoverer, Degas, ClustEx, and KeyPathwayMiner). For each of them, we performed the drug enrichment analysis on four drug perturbation databases: DrugMatrix, GEO, L1000, and CMAP. After that, we ranked the drugs according to the p-value or the positivity score. We also merged more drug lists, which come from different algorithms for the same disease and we applied the same ranking strategy. Finally, we evaluated the DrugMerge performance through the Precision@20 and the Reciprocal Hit-Rank using the golden-standard drugs, collected in the Therapeutic Target Database, and ClinicalTrials for COVID-19.

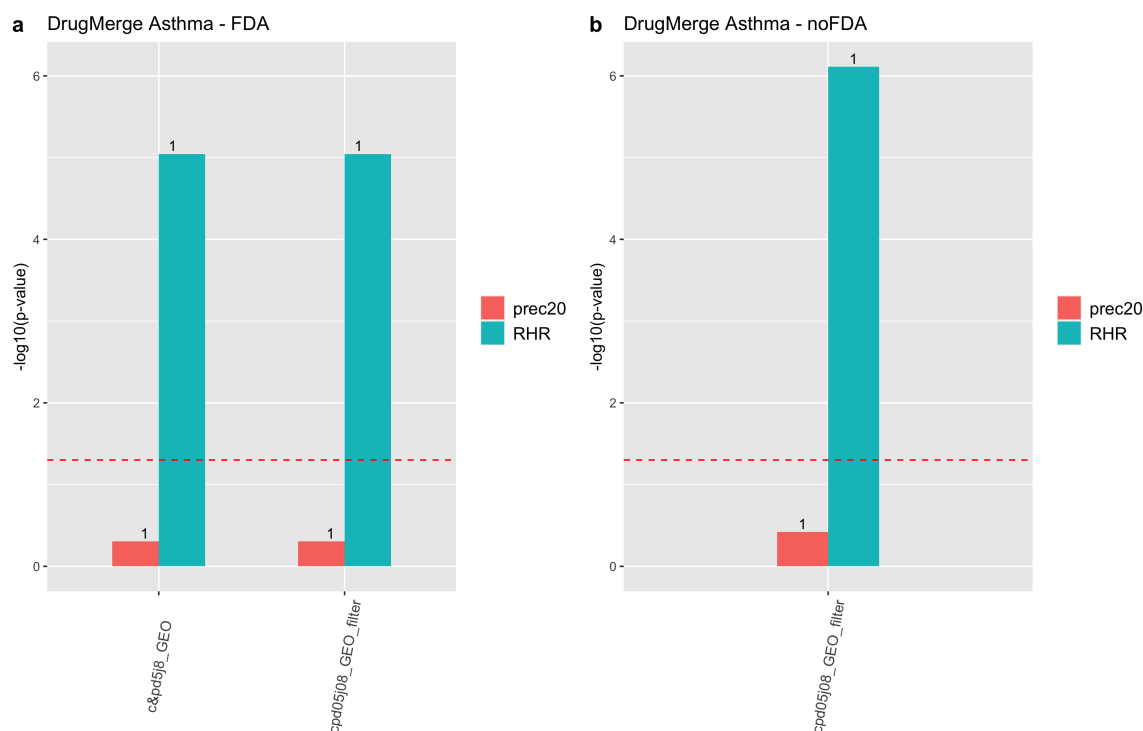


Figure 2: DrugMerge performance on Asthma data. The bars represent the  $-\log_{10}(\text{pvalue})$  with respect to the precision@20 (red), and the RHR (light blu). On the x-axis, different algorithms or combinations of them are shown. The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the dotted line show a significant p-value. The a) panel represents the DrugMerge performance when only FDA-approved drugs are considered; b) when all drugs (without any FDA filtering) are considered. The plot has been generated by the *ggplot2* [Wickham et al., 2016] R package.

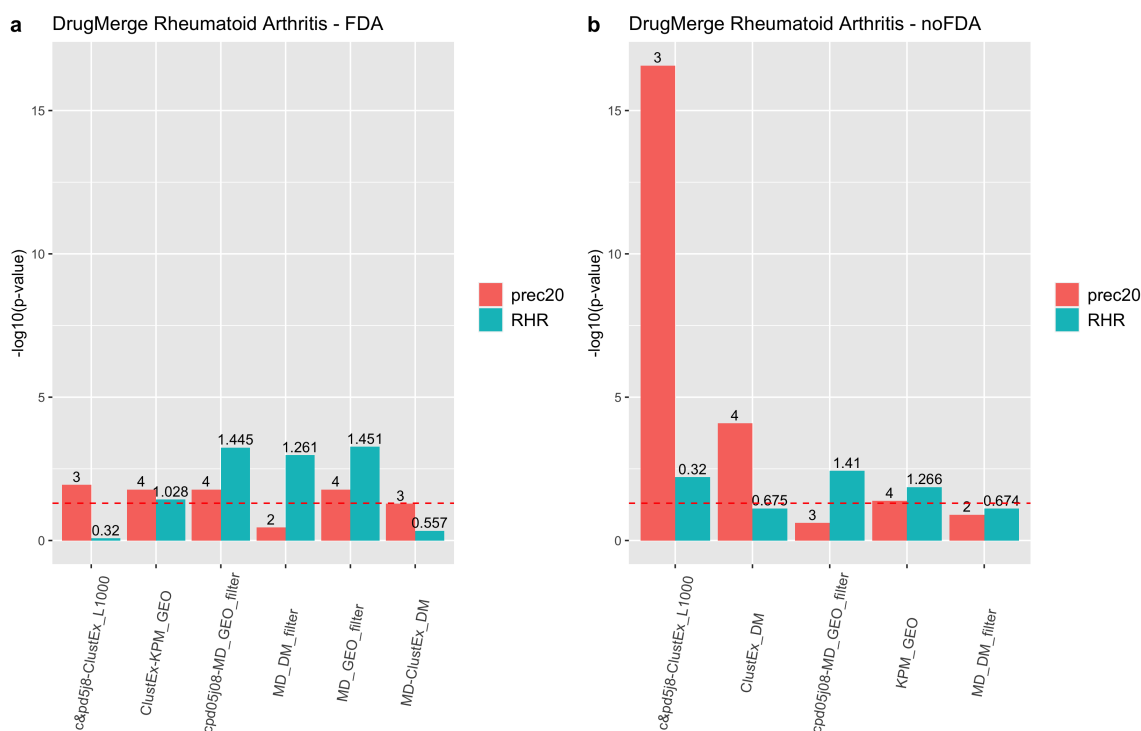


Figure 3: DrugMerge performance on Rheumatoid Arthritis data. The bars represent the  $-\log_{10}(pvalue)$  with respect to the precision@20 (red), and the RHR (light blu). On the x-axis, different algorithms or combinations of them are shown. The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the dotted line show a significant p-value. The a) panel represents the DrugMerge performance when only FDA-approved drugs are considered; b) when all drugs (without any FDA filtering) are considered. The plot has been generated by the *ggplot2* [Wickham et al., 2016] R package.

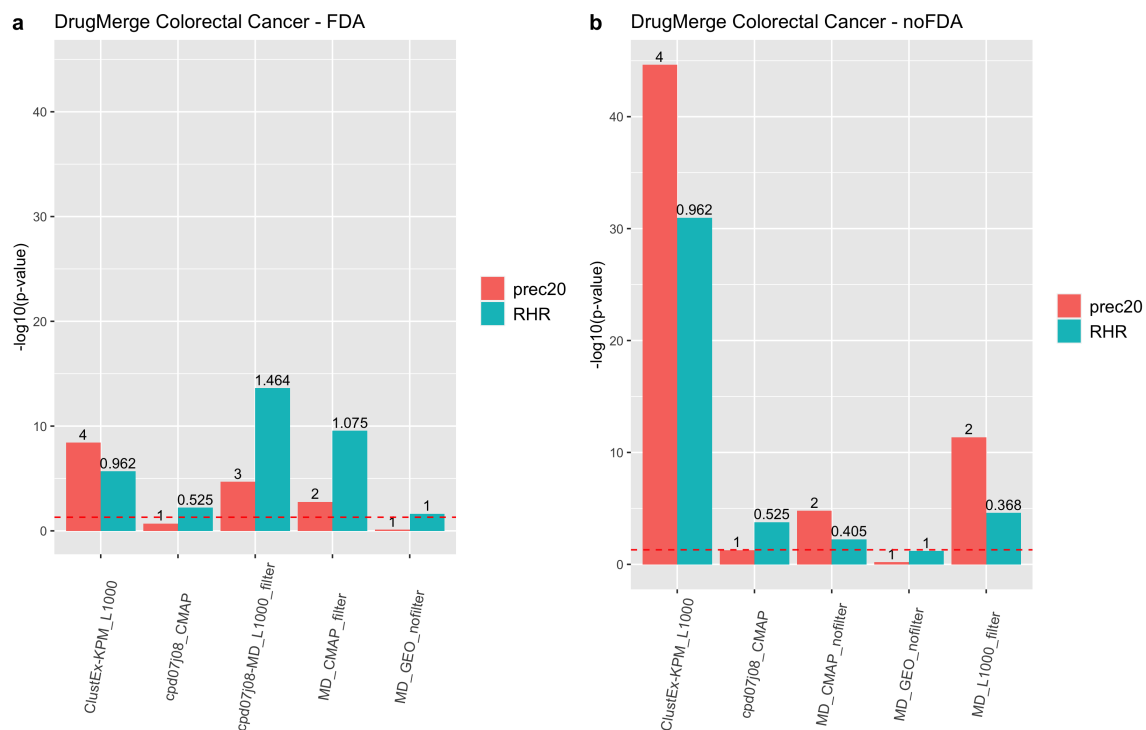


Figure 4: DrugMerge performance on Colorectal Cancer data. The bars represent the  $-\log_{10}(\text{pvalue})$  with respect to the precision@20 (red), and the RHR (light blue). On the x-axis, different algorithms or combinations of them are shown. The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the line show a significant p-value. The a) panel represents the DrugMerge performance when only FDA-approved drugs are considered; b) when all drugs (without any FDA filtering) are considered. The plot has been generated by the *ggplot2* [Wickham et al., 2016] R package.

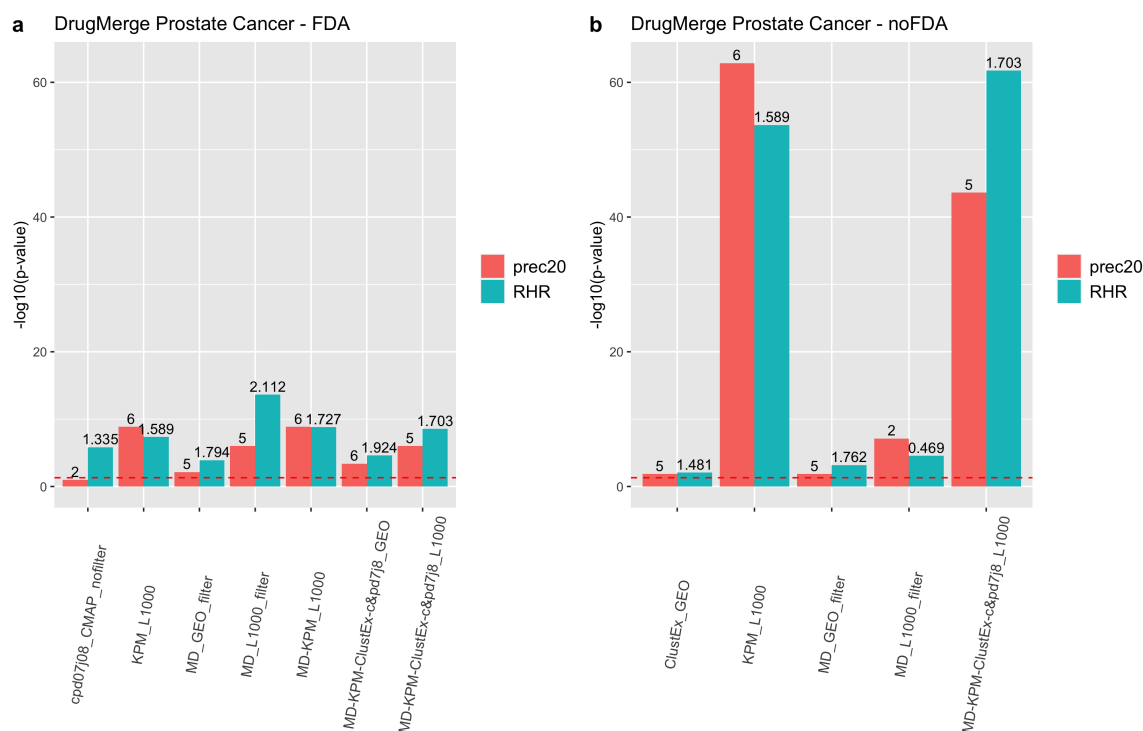


Figure 5: DrugMerge performance on Prostate Cancer data. The bars represent the  $-\log_{10}(pvalue)$  with respect to the precision@20 (red), and the RHR (light blu). On the x-axis, different algorithms or combinations of them are shown. The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the dotted line show a significant p-value. The a) panel represents the DrugMerge performance when only FDA-approved drugs are considered; b) when all drugs (without any FDA filtering) are considered. The plot has been generated by the *ggplot2* [Wickham et al., 2016] R package.

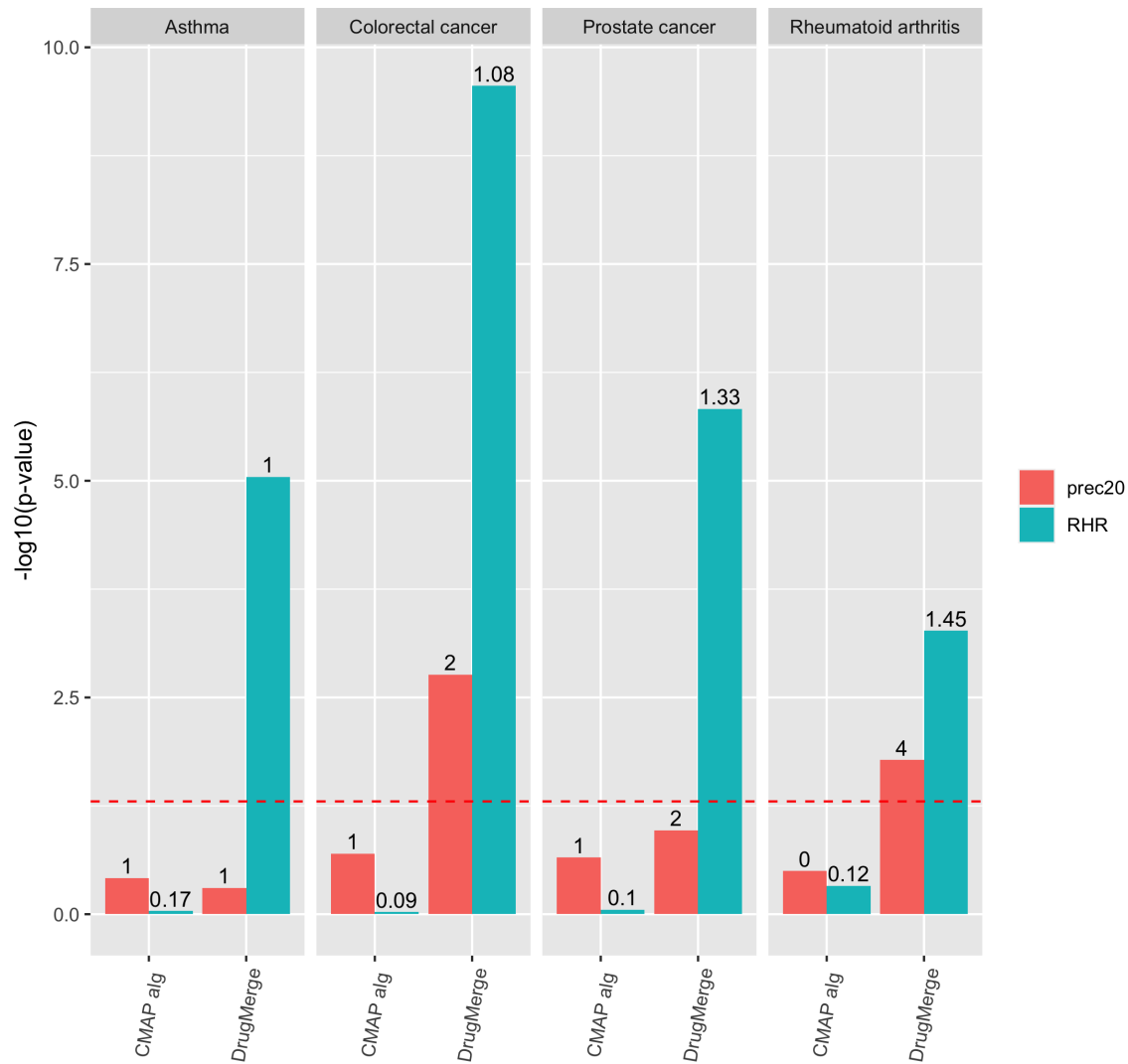


Figure 6: Comparison of performance of DrugMerge and the CMAP algorithm across the four benchmark diseases. For Colorectal Cancer and Prostate Cancer both methods use CMAP data. For Asthma and Rheumatoid Arthritis DrugMerge uses GEO data. The bars represent the  $-\log_{10}(pvalue)$  with respect to the precision@20 (red), and the RHR (light blu). The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the dotted line show a significant p-value. The plot has been generated by the *ggplot2* [Wickham et al., 2016] R package.



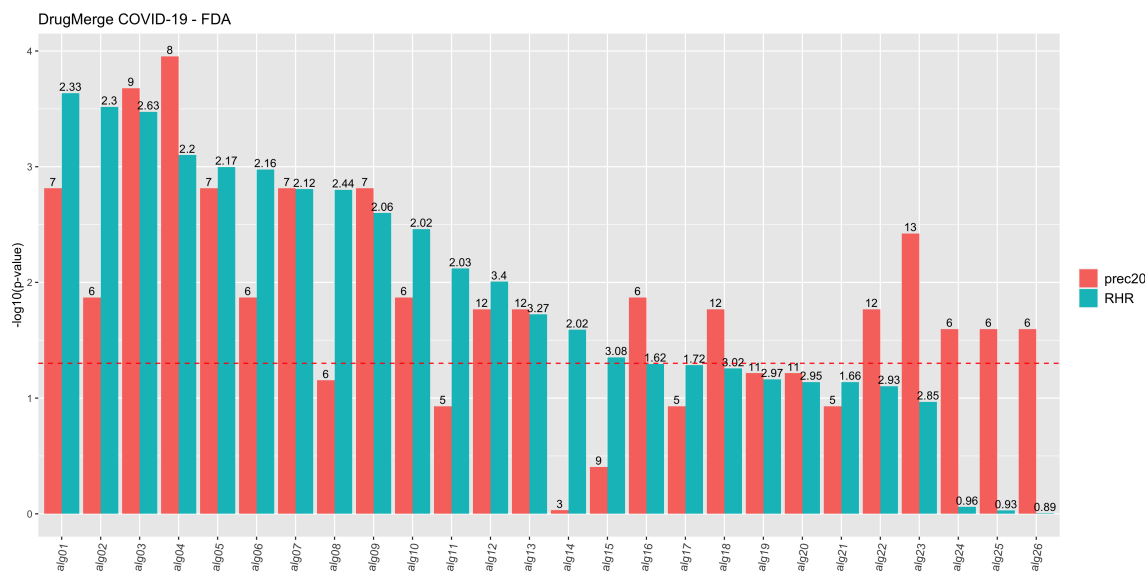


Figure 7: DrugMerge performance on COVID-19 data restricted to drugs with FDA approval. The bars represent the  $-\log_{10}(pvalue)$  with respect to the precision 20 (red), and the RHR (light blu). On the x-axis, different algorithms or combinations of them are indexed with aliases, since the complete algorithm names are too long. The mapping between them is on Table S5.1. The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the dotted line show a significant p-value.

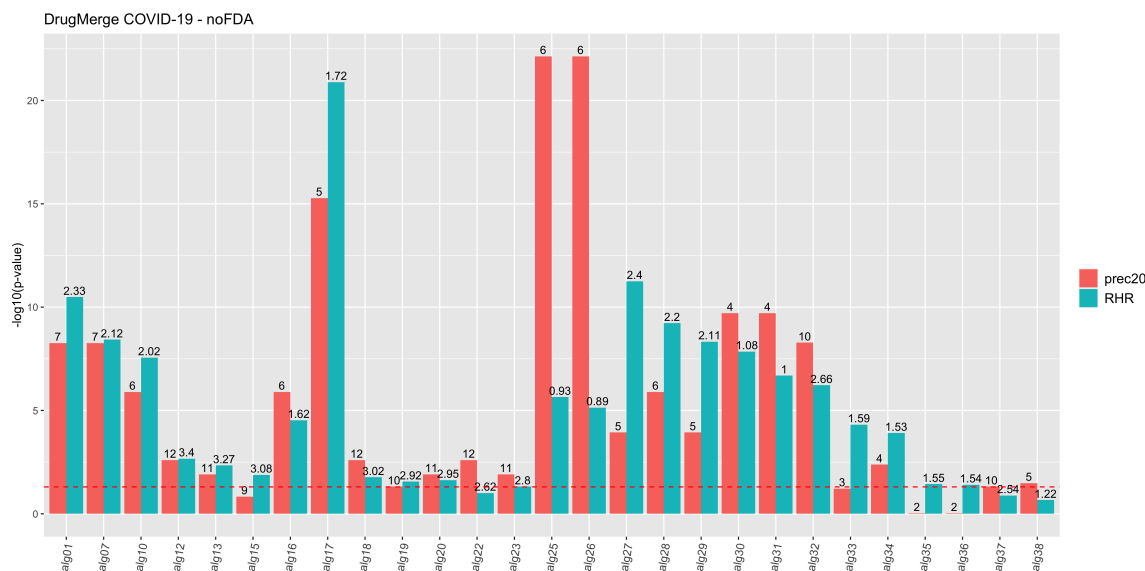


Figure 8: DrugMerge performance on COVID-19 data without FDA approval restriction. The bars represent the  $-\log_{10}(pvalue)$  with respect to the precision@20 (red), and the RHR (light blue). On the x-axis, different algorithms or combinations of them are indexed with aliases, since the complete algorithm names are too long. The mapping between them is on Table S5.1. The numbers on the top of the bars show the absolute values of precision@20 or RHR. The dotted red line represents the limit of significance ( $-\log_{10}(0.05)$ ). All the bars above the dotted line show a significant p-value.

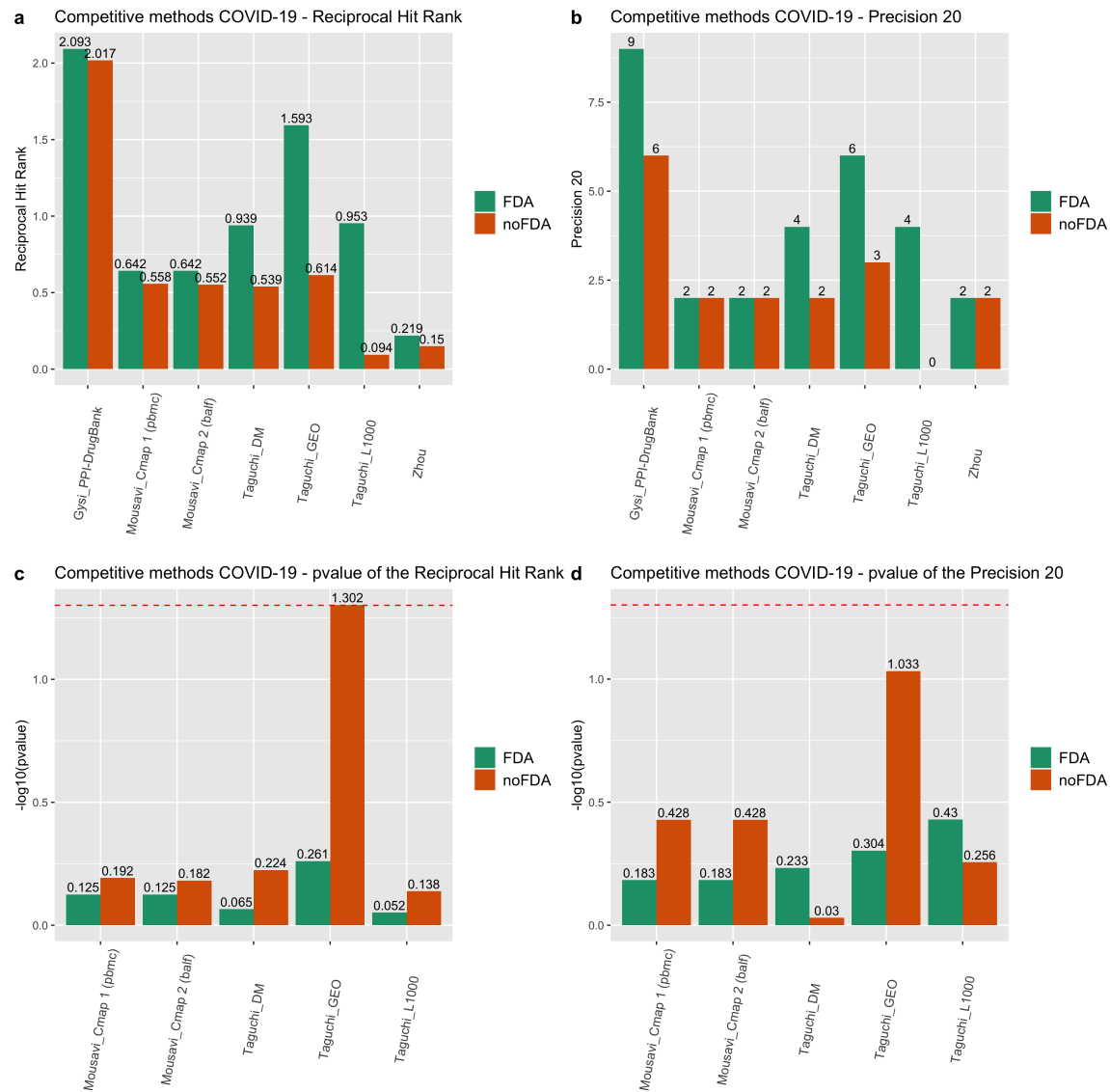


Figure 9: Performance of drug rankings reported in literature for COVID-19 drug repurposing. The dotted red line indicates the 0.05 significance threshold.