

Article

Drug Repurposing Using Modularity Clustering in Drug-Drug Similarity Networks Based on Drug-Gene Interactions

Vlad Groza ¹, Mihai Udrescu ^{1,*} , Alexandru Bozdog ¹ and Lucreția Udrescu ² 

¹ Department of Computer and Information Technology, University Politehnica of Timișoara, 300223 Timișoara, Romania; vlad.groza@student.upt.ro (V.G.); alexandru.bozdog@cs.upt.ro (A.B.)

² Department I—Drug Analysis, “Victor Babeș” University of Medicine and Pharmacy Timișoara, 300041 Timișoara, Romania; udrescu.lucretia@umft.ro

* Correspondence: mihai.udrescu@cs.upt.ro

Abstract: Drug repurposing is a valuable alternative to traditional drug design based on the assumption that medicines have multiple functions. Computer-based techniques use ever-growing drug databases to uncover new drug repurposing hints, which require further validation with in vitro and in vivo experiments. Indeed, such a scientific undertaking can be particularly effective in the case of rare diseases (resources for developing new drugs are scarce) and new diseases such as COVID-19 (designing new drugs require too much time). This paper introduces a new, completely automated computational drug repurposing pipeline based on drug-gene interaction data. We obtained drug-gene interaction data from an earlier version of DrugBank, built a drug-gene interaction network, and projected it as a drug-drug similarity network (DDSN). We then clustered DDSN by optimizing modularity resolution, used the ATC codes distribution within each cluster to identify potential drug repurposing candidates, and verified repurposing hints with the latest DrugBank ATC codes. Finally, using the best modularity resolution found with our method, we applied our pipeline to the latest DrugBank drug-gene interaction data to generate a comprehensive drug repurposing hint list.

Keywords: bioinformatics; drug repurposing; complex network analysis; modularity clustering; ATC code



Citation: Groza, V.; Udrescu, M.; Bozdog, A.; Udrescu, L. Drug Repurposing Using Modularity Clustering in Drug-Drug Similarity Networks Based on Drug-Gene Interactions. *Pharmaceutics* **2021**, *13*, 2117. <https://doi.org/10.3390/pharmaceutics13122117>

Academic Editor: David Barlow

Received: 24 October 2021

Accepted: 2 December 2021

Published: 8 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The growth in the number of newly approved pharmaceutical substances has stagnated despite the ever-growing resources that the industry allocates [1–4]. Designing, developing, and testing new medicines is an expensive, long, and cumbersome process [5], which becomes explicitly bothersome for new rare diseases—because funds are limited—and new pathogen epidemics—stopping the disease spread requires a rapid therapeutic solution [6,7]. One convenient alternative to the pharmaceutical industry's productivity challenges is drug repurposing, underpinned by the R&D in the pharmaceutical industry, as well as the observations and long-time experience indicating the favorable polypharmacological profile of drugs (in other words, most pharmaceutical substances tend to have multiple functions) [8–10]. The trend that calls for drug repurposing techniques is in sync with the recent expansion of Big Data and machine learning in genetics, biology, and medicine; therefore, we witnessed the development of a wide array of computer-based methodologies to uncover new drug repurposing [11–13].

A significant area in computational repurposing (or repositioning) relies on the complex network representations of various drug interaction/relationship types, e.g., drug-drug [14], drug-target [15–17], drug-side effect [18], drug-gene. The networks consist of nodes/edges—representing drugs, targets, genes, or side effects—and links/edges—representing interactions or other types of relationships [19]. The network of specific drug interactions allows for the characterization of a complex biological system under therapy; therefore, researchers can use computational techniques and network science principles to

explore the interplay between microscale interactions and macroscale behavior [14]. An important area in network science is community/cluster detection and analysis [20,21]. The assumption is that nodes from a distinct cluster have similar topological properties and, thus, share a common feature; this results in drug repurposing opportunities [6]. (If most drugs in a cluster have a particular therapeutic function, then it is reasonable to assume that the function also exists at least in some of the other drugs in the cluster). Many network-based computational drug repurposing methods use topological network features, such as centralities (topological indicators/measures of a node's importance in the network) and modularity, to identify potential repositioning [22,23].

All computational drug repositioning methods produce lists of hints or predictions that require testing or confirmation in silico (e.g., molecular docking) [24], in vitro, and in vivo [25]. One can also indirectly prove the effectiveness of the computational technique by applying it on an earlier database version and testing the predictions on the latest data [14,22]. The existing computational pipelines predicted several important drug repurposings. Moreover, the crisis generated by the COVID-19 pandemic called for drug repurposing solutions to counter SARS-CoV-2 infections.

In our prior study, we also approached the problem of drug repositioning by building a drug–drug interaction network [14] and a drug–drug similarity network based on drug–target interactions [22]; we used the corresponding drug–drug and drug–target interaction data from DrugBank 4.1 and 4.2, respectively. In [14], we used community detection with energy-based layouts and fixed modularity; in [22], we also used energy-based layouts and fixed modularity, as well as ranking nodes by network centralities; in both previous approaches, we labeled the clusters and confirmed predictions with expert analysis.

In this paper, we also use a method based on network community detection and analysis. To this end, we build a drug–drug similarity network, because similarity networks are better suited for community detection: Nodes in the same community are more likely to be similar. Indeed, many other computational drug repurposing methods operate on similarity networks [26,27], with similarity defined on various criteria—from drug–target interactions [22] to adverse effects [18]. We find inspiration in the diseasome project [28,29] based on processing a disease–gene bipartite network (i.e., with two types of nodes, namely, genes and diseases); the processing of the disease–gene network projects it as either a gene–gene similarity or a disease–disease similarity network. In the gene–gene network, a link between two genes exists if there is at least one common disease with which they interact; in the disease–disease network, a link between two diseases exists if at least one gene is responsible for both diseases.

Our method builds a drug–gene interaction network with drug–gene interaction data from the earlier DrugBank 5.0.9 version, then projects it as a drug–drug similarity network; this is the first drug repurposing method derived from a gene-based drug–drug similarity network to the best of our knowledge. Our drug–drug similarity network is weighted—the weight of the link between two nodes/drugs represents the number of genes with which the two drugs interact in the same manner. We then use modularity-based network clustering to identify drug communities/clusters. We adopt the same assumption as in the case of the diseasome analysis in [30] that nodes inside the same community most probably share a common function or property. In this manner, if a drug inside one community does not have the ATC code level 1 of the majority, then we hypothesize that the drug can be repurposed accordingly. Nonetheless, we improve the efficiency of the approach by providing an automated procedure for tuning modularity resolution [31] by comparing the ATC code level 1 predicted with our method applied to DrugBank 5.0.9 [32] with the level 1 ATC codes of the drug in the latest DrugBank version 5.1.8 [33]. Finally, we apply our pipeline—with the optimized modularity resolution—to the latest DrugBank data to generate a new list of repurposing hints, which we support by existing literature findings. Refer to the overview of our proposed methodology in Figure 1. We only considered drugs listed as *approved* in DrugBank.

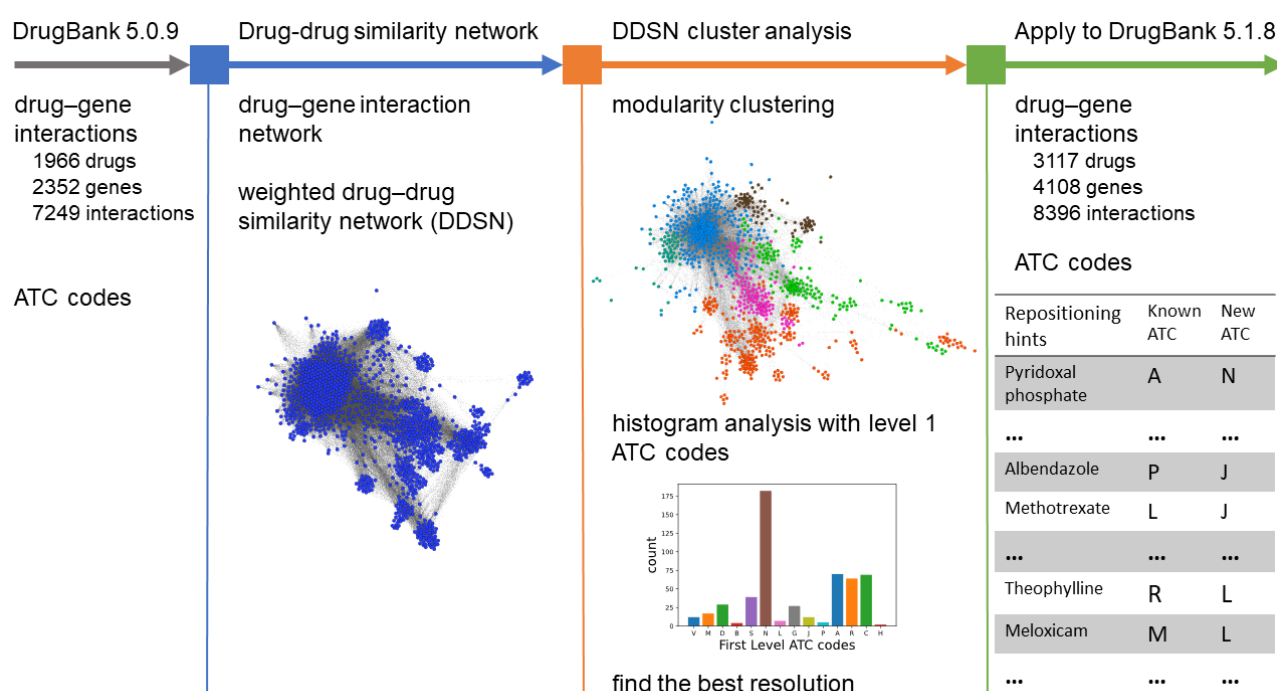


Figure 1. The overview of our proposed computational drug repurposing pipeline. In the first step, we use drug–gene interaction information from DrugBank 5.0.9 to build the (bipartite) drug–gene interaction network, which we then projected as a drug–drug similarity network (DDSN). In the second step, we used modularity class network clustering to identify drug communities with shared properties, analyzed the DrugBank 5.0.9 first-level ATC code histograms in each community to predict new drug properties, and checked these predictions against the latest DrugBank 5.1.8 level 1 ATC codes. The procedure in the second step allows maximizing the number of confirmed repositionings by adjusting modularity resolution. The third step uses our method with the optimized resolution value determined in the second step to generate a repurposing hints list according to DrugBank 5.1.8.

Three arguments support the novelty of the research presented in this paper. First, this manuscript is—to the best of our knowledge—the first to build and process a DDSN based on drug–gene interaction data. Second, we present a novel method (based on level 1 ATC codes) that labels clusters and generates repositioning hints automatically. Third, we tuned modularity resolution algorithmically and automatically confirmed repositioning hints by comparing two chronologically distinct DrugBank versions.

From a pharmacological perspective, our overarching contribution is to develop, for the first time, and promote the drug–gene interaction networks as a valuable analytical, screening, and visualization tool in drug repositioning. Our method can complement existing computational repositioning pipelines; therefore, it can be integrated into more sophisticated ensemble methods.

2. Materials and Methods

In this section, we present the conceptual description of our algorithmic drug repositioning method from Figure 1. The thorough technical implementation and description are provided on our GitHub page <https://github.com/GrozaVlad/Drug-repurposing-using-DDSNs-and-modularity-clustering> (last commit on 21 October 2021). We used *Nodejs* with packets *xml-js* (for parsing the DrugBank xml files) and *pg* (for interacting with the PostgreSQL database), and *Docker* and *Docker-compose* for containerized databases [34]. For building and clustering DDSN, we used the *Python* packages *Psycopg2*, *Pandas* [35], *NetworkX* [36], and *Cdlib* [37]; for visualizing the networks, we used *Gephi* [38]. The hardware platform for running this project was a MacBook Pro, Intel Core i9—2400 MHz with 16 GB RAM, GPU Radeon Pro 560× 4 GB.

2.1. Databases

In order to facilitate an automated procedure of validating our drug repurposing pipeline, we used the earlier DrugBank version 5.0.9 to generate repurposing predictions in one of the anatomical or pharmacological groups described by the first-level ATC codes, then we validated the predictions with the ATC codes with the latest DrugBank version 5.1.8 (last accessed on 30 September 2021).

In DrugBank version 5.0.9, there are 1966 drugs, 2352 genes, and 7249 drug–gene interactions; the interaction types are part of the set $I_e = \{\text{inhibitor, agonist, antagonist, other/unknown, ligand, partial agonist, inducer, other, suppressor, binder, antibody, modulator, allosteric modulator, potentiator, neutralizer, stimulator, activator, component of, substrate, inactivator, blocker, antisense oligonucleotide}\}$. In the latest DrugBank version 5.1.8, there are 3117 drugs, 4108 genes, and 8396 drug–gene interactions with interaction types part of the set $I_l = \{\text{inhibitor, agonist, antagonist, other/unknown, antibody, substrate, ligand, partial agonist, inducer, other, suppressor, binder, potentiator, modulator, activator, cofactor, degradation, positive allosteric modulator, incorporation into and destabilization, allosteric modulator, neutralizer, stimulator, binding, inactivator, inverse agonist, blocker, chaperone, inhibition of synthesis, antisense oligonucleotide, gene replacement, regulator}\}$. Refer to Section 4.1 for explanations.

We chose DrugBank [33] because it is a comprehensive, versioned, and scientifically curated (i.e., robust) database with consistent support for in silico drug design and repositioning space exploration [32].

2.2. Building the Drug–Drug Similarity Network

The bipartite drug–gene interaction network is a graph $\mathcal{G} = (V, E)$, where V is the set of vertices or nodes, and E is the set of edges. The network \mathcal{G} is bipartite because $V = V_D \cup V_G$, where V_D is the set of drugs and V_G is the set of genes. The edges $e_{ij} \in E$ represent interactions between a drug $D_i \in V_D$ and a gene $G_j \in V_G$ (the interaction is of the type $T_k \in I$, with I defined in Section 2.1). An example of such a drug–gene bipartite graph is presented in Figure 2a, with 4 drugs, 3 genes, and 3 types of drug–gene interactions.

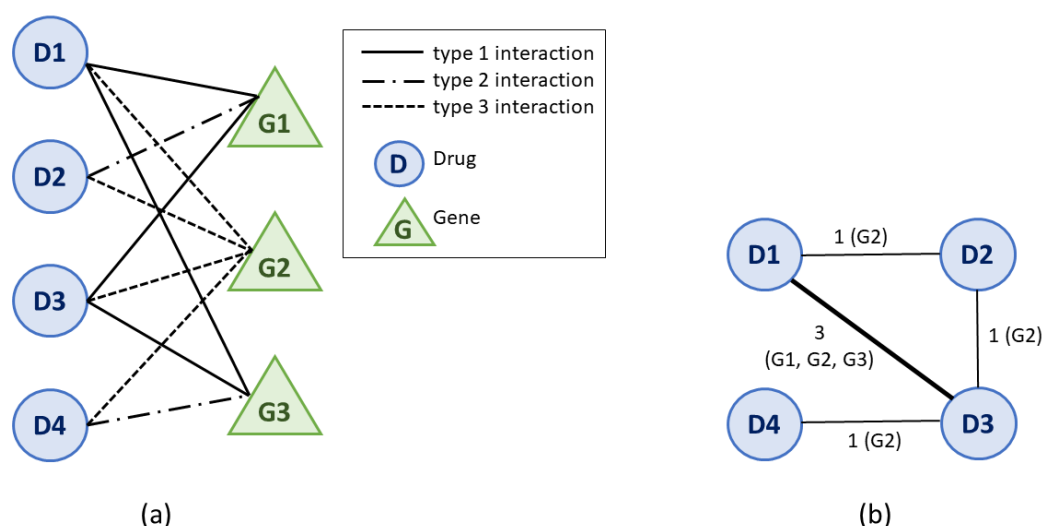


Figure 2. An illustrative example of projecting the bipartite drug–gene interaction graph \mathcal{G} (a) into a weighted drug–drug similarity network \mathcal{W} (b). In our example, \mathcal{G} has 4 drugs (D_1, D_2, D_3 , and D_4), 3 genes (G_1, G_2 , and G_3), and 3 types of drug–gene interactions. In the drug–drug similarity network from panel (b), nodes are drugs, and links between two drugs represent the number of genes with which the drugs interact in the same manner. For instance, as shown, the link $w_{1,3}$ between nodes/drugs D_1 and D_3 has a weight of 3 because D_1 and D_3 have the same type of interaction with genes G_1, G_2 , and G_3 .

From the drug–gene bipartite network \mathcal{G} , we generated the weighted drug–drug similarity network $\mathcal{W} = (V_D, W)$ using network projection [39]. In the DDSN, the nodes represent drugs, and a link between two nodes exists if there is at least one gene with which the two drugs interact in the same manner (i.e., the interactions are of the same type $T_k \in I$). In Figure 2b, we present the DDSN projection of the drug–gene example network in Figure 2. The network is weighted because two drugs D_i and D_j can have the same type of interactions with m genes; therefore, the weight of edge $w_{ij} \in W$ is m .

2.3. Network Clustering Analysis

The clustering of network $\mathcal{G} = (V, E)$ is the process of classifying all nodes $v_i \in V$ in one of the n (disjoint) subsets C_j , with $V = \bigcup_{j=1}^n C_j$, according to their topological properties. In this paper, we use modularity-based clustering because of its proven effectiveness in drug network analysis [14,22,23]. As defined in [40], the modularity of a clustering \mathcal{C} in a weighted network such as our DDSN—represented as \mathcal{W} —is defined as follows.

$$M = \frac{1}{2a} \sum_{ij} \left(w_{ij} - \frac{k_i k_j}{2a} \right) p(C_i, C_j). \quad (1)$$

In Equation (1), $a = \frac{1}{2} \sum_{ij} w_{ij}$; i and j are the indexes of nodes $v_i, v_j \in V_D$; k_i and k_j are the node degrees (i.e., the sums of weights of incident edges) for nodes $v_i, v_j \in V_D$; w_{ij} is the adjacency matrix of nodes in \mathcal{W} ; C_i and C_j are the communities that include nodes $v_i, v_j \in V_D$, respectively; and p is a function $p(x, y)$ that returns 1 if $x = y$ and 0 otherwise. (In our DDSN, nodes v_i and v_j are drugs D_i and D_j , respectively).

The modularity of clustering \mathcal{C} is a value $M_{\mathcal{C}} \in [-1, 1]$, representing the edge density within the clusters with respect to the edge density between clusters. The clustering algorithms are based on modularity search for the best partitioning \mathcal{C} of the node-set such that the value of M is maximized. The problem is that an exhaustive search for the best modularity entails large computational burden. Consequently, in practice, heuristic algorithms approximate optimal modularity clustering. However, if the network is very large, such approximations cannot identify small-size clusters—even if the density of internal edges is high and the density of edges between these small clusters and the rest of the network is low.

In this paper, we use the modularity-based clustering algorithm from [41], which controls the resolution of the clustering using a recursive procedure that starts with each node being a cluster and then moving nodes v_i (i.e., D_i in our DDSN) to a different cluster C_j if this generates a positive modularity gain expressed as follows.

$$\Delta M = \left[\frac{K_{C_j}^* + K_i^{C_j}}{2a} - \left(\frac{K_{C_j} + K_i}{2a} \right)^2 \right] - \left[\frac{K_{C_j}^*}{2a} - \left(\frac{K_{C_j}}{2a} \right)^2 - \left(\frac{K_i}{2a} \right)^2 \right]. \quad (2)$$

In Equation (2), $K_{C_j}^*$ is the sum of the weights of all edges within cluster C_j ; K_{C_j} is the sum of the weights of all edges incident to nodes in cluster C_j ; K_i is the sum of the weights of all edges incident to node v_i (D_i in DDSN); and $K_i^{C_j}$ is the sum of the weights of links from v_i to all nodes in cluster C_j . The algorithm controls the clustering resolution using the value of $\lambda = \Delta M$ —a lower λ determines a higher number of clusters.

2.4. Tuning Resolution λ

Using Algorithm 1, we tune the modularity resolution to achieve efficiency in predicting new drug properties. To this end, we try λ values in the $[0.1, 5]$ interval, with a step of 0.1, generate the modularity clustering \mathcal{C} for each resolution value ($\text{Clustering}(\mathcal{G}, \lambda)$), and determine the dominant property \mathcal{P}_i in each cluster $C_i \in \mathcal{C}$. The dominant property \mathcal{P}_i corresponds to the level 1 ATC code of the majority of drugs in cluster i , $D_j \in C_j$, as resulting from the level 1 ATC code histogram of C_i , and denoted $\mathcal{A}^1(C_i)$. Then, for each

drug D_j in each cluster C_i , we checked the list of first level ATC codes for drug D_j (denoted $A^1(D_j)$) against the drug's cluster dominant property \mathcal{P}_i . If \mathcal{P}_i is not in the list of DrugBank 5.0.9 level 1 ATC codes for D_j (i.e., $A^1(D_j)$), but it is present in the list of DrugBank 5.1.8 level 1 ATC codes (i.e., $A_c^1(D_j)$), then we consider this as a confirmed repositioning of D_j to property \mathcal{P}_i . As such, we will add drug D_j to the list of repositionings confirmed with DrugBank 5.1.8 level 1 ATC codes, \mathcal{R}^c . Value λ_{max} corresponds to \mathcal{R}^c with the biggest number of elements, namely $\max\{|\mathcal{R}^c|\}$.

Algorithm 1 Find the parameter λ , such that the clustering \mathcal{C} of nodes/drugs D_i in \mathcal{G} with modularity resolution λ (i.e., $\text{Clustering}(\mathcal{G}, \lambda)$) produces the biggest number of repositionings confirmed with the level 1 ATC codes in DrugBank 5.1.8.

Input: Drug-drug similarity network $\mathcal{G} = (V_D, E)$ based on drug-gene interaction data from DrugBank 5.0.9., ATC codes for drugs in DrugBank versions 5.0.9 and 5.1.8

Output: The λ value that generates the highest number of confirmed repositionings.

```

1: for  $\lambda$  in range (0.1 to 5), with 0.1 steps do
2:    $\mathcal{C} \leftarrow \text{Clustering}(\mathcal{G}, \lambda)$ 
3:   for all  $C_i \in \mathcal{C}$  do
4:      $\mathcal{P}_i \leftarrow \mathcal{A}^1(C_i)$ 
5:      $\mathcal{R}_i^c \leftarrow \emptyset$ 
6:     for all  $D_j \in C_i$  do
7:       if then  $\mathcal{P}_i \notin A^1(D_j)$  &  $\mathcal{P}_i \in A_c^1(D_j)$ 
8:          $\mathcal{R}_i^c \leftarrow \mathcal{R}_i^c \cup \{D_j\}$ 
9:       end if
10:    end for
11:  end for
12:   $\mathcal{R}^c \leftarrow \bigcup_i \mathcal{R}_i^c$ 
13: end for
14: Return the value of  $\lambda_{max}$  corresponding to  $\max\{|\mathcal{R}^c|\}$ 

```

2.5. Generating New Repurposing Hints

We generated a list of new repositioning hints using the modularity clustering with the resolution value determined by Algorithm 1 in Section 2.4. Algorithm 2 presents the method we follow: Cluster the DDSN built with drug–gene interaction information from DrugBank 5.1.8 using the tuned resolution λ_{max} ($\mathcal{C} = \text{Clustering}(\mathcal{G}, \lambda_{max})$); determine the dominant property \mathcal{P}_i of each cluster $C_i \in \mathcal{C}$ as resulted from C_i 's level 1 ATC code histogram (denoted $\mathcal{A}^1(C_i)$); and check for each drug D_j in each cluster C_i the list of first level ATC codes of D_j (denoted $A^1(D_j)$) against its cluster's dominant property \mathcal{P}_i . If the cluster's dominant property \mathcal{P}_i is not in $A^1(D_j)$ (the list of D_j level 1 ATC codes), we hint that D_j can be repositioned to \mathcal{P}_i . Consequently, we add these repositioning cases as drug–predicted property pairs (D_j, \mathcal{P}_i) to the repositioning hints list \mathcal{N} .

Algorithm 2 Generate the list of drug repurposing hints by clustering the DDSN \mathcal{G} with the tuned modularity resolution.

Input: Drug–drug similarity network $\mathcal{G} = (V_D, E)$ based on drug–gene interaction data from DrugBank 5.1.8, λ_{max} , and the ATC codes for drugs in DrugBank 5.1.8.

Output: The repositioning hints \mathcal{N} as a list of drug–predicted property pairs, $(D_j, \mathcal{A}^1(C_i))$.

```

1:  $\mathcal{C} \leftarrow \text{Clustering}(\mathcal{G}, \lambda_{max})$ 
2:  $\mathcal{N} \leftarrow \emptyset$ 
3: for all  $C_i \in \mathcal{C}$  do
4:    $\mathcal{P}_i \leftarrow \mathcal{A}^1(C_i)$ 
5:   for all  $D_j \in C_i$  do
6:     if  $\mathcal{P}_i \notin \mathcal{A}^1(D_j)$  then
7:        $\mathcal{N} \leftarrow \mathcal{N} \cup \{(D_j, \mathcal{P}_i)\}$ 
8:     end if
9:   end for
10: end for
11: Return the list of drug repositionings  $\mathcal{N}$  as drug–predicted property pairs

```

3. Results

3.1. DDSN Using Drug–Gene Interactions from DrugBank 5.0.9

Following the algorithmic approach presented in Figure 1, according to the methods described in Sections 2.2–2.5, we employ cluster-based network analysis on the drug–drug similarity network (DDSN) built with drug–gene interaction information from DrugBank 5.0.9 to search for the most effective modularity resolution λ_{max} —in other words, the modularity resolution that produces the highest number of drug repositionings confirmed with level 1 ATC codes from DrugBank 5.1.8. Figure 3 presents the result of running Algorithm 1 from Section 2.4; the best results correspond to resolutions 1.9 and 2.0 (the same nine confirmed repositionings in both cases). Henceforth, we will consider $\lambda_{max} = 2.0$.

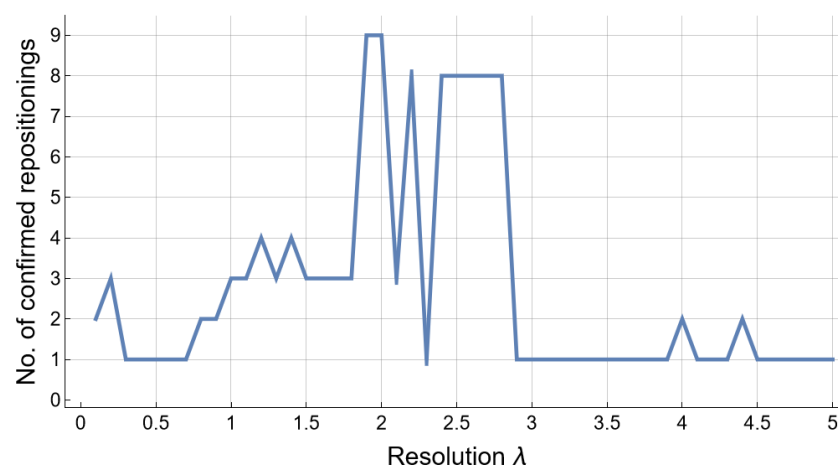


Figure 3. The number of confirmed repositionings \mathcal{R}^c for resolution λ values in the $[0.1, 5]$ interval, with a step of 0.1, after running Algorithm 1 on the DDSN \mathcal{G} built with drug–gene interaction information from DrugBank 5.0.9. The highest number of repositionings confirmed with level 1 ATC codes from DrugBank 5.1.8 (i.e., 9) corresponds to resolutions 1.9 and 2.0.

Figure 4 presents the largest connected component of the DDSN, constructed with drug–gene interaction data from DrugBank 5.0.9 and clustered with modularity resolution $\lambda_{max} = 2.0$; the text indicates the topological coordinates of repositionings confirmed with DrugBank 5.1.8 data.

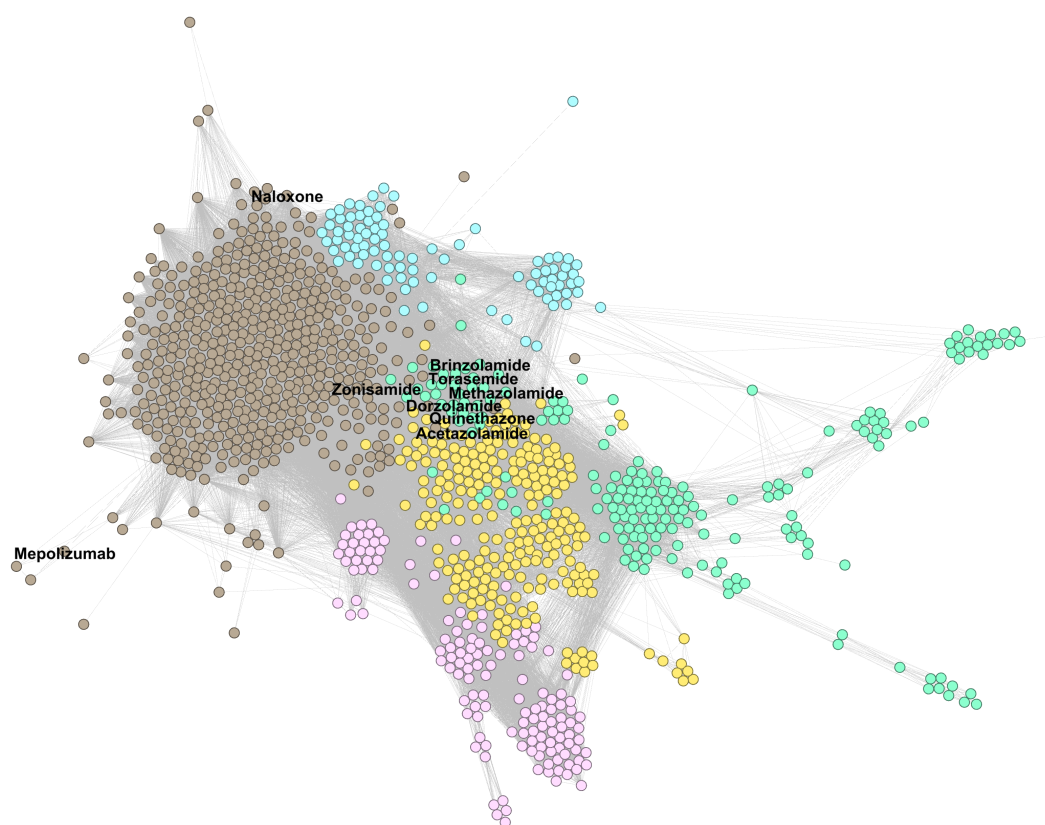


Figure 4. Drug–drug similarity network (DDSN) built with drug–gene interaction data from DrugBank 5.0.9, clustered using modularity classes for resolution $\lambda_{max} = 2.0$. We indicate the position of drugs repositioned and confirmed (with level 1 ATC codes from DrugBank 5.1.8) them by labeling the corresponding nodes with their names. The brown nodes represent drugs in cluster C_0 (512 drugs), yellow nodes represent drugs in cluster C_1 (238 drugs), green nodes represent drugs in cluster C_2 (197 drugs), pink nodes represent drugs in cluster C_3 (143 drugs), and light blue nodes represent drugs in cluster C_4 (88 drugs).

In Figure 4, nodes represent drugs, and links represent similarity relationships based on drug–gene interactions, as described in Section 2.2; node colors correspond to specific clusters, as determined by the modularity class, and all links are represented with grey lines.

In Appendix A.1, Figures A1–A3, present zoomed details of DDSN from Figure 4 in the vicinity of nine confirmed repositionings corresponding to $\lambda_{max} = 2.0$. The repositionings come from cluster C_0 —brown and cluster C_2 —green nodes. We indicated the drug repositionings confirmed with DrugBank 5.1.8 data with red arrows (\rightarrow) in Figures A1 and A2; in Figure A3, we have many confirmed repurposed drugs and a high density of nodes; hence, red diamonds (\diamond) were used instead of arrows.

The zoomed details provided by Figures A1 and A2 show that mepolizumab and naloxone are within cluster C_0 (brown nodes), where the dominant property is given by the level 1 ATC code N—*Nervous system*, followed by code R—*Respiratory system*. As such, our method automatically predicts that mepolizumab (listed as L—*Antineoplastic and immunomodulatory drugs* in DrugBank 5.0.9) acts as a drug with level 1 ATC code R. (In Appendix A.2, Figure A4 shows that in cluster C_0 —in addition to the dominant level 1 ATC codes N—we also have many subcluster drugs with level 1 ATC codes A—*Alimentary tract and metabolism*; R—*Respiratory system*; and C—*Cardiovascular system*). Our method predicts that naloxone (an opioid overdose antidote in DrugBank 5.0.9) also acts on the nervous system (first level ATC N). The more recent DrugBank 5.1.8 confirms the predictions, listing mepolizumab with first level ATC code R and naloxone with N (see more details in Section 3.3.1).

In Appendix A.1, Figure A3, we zoom in to the region in DrugBank 5.0.9 DDSN with the confirmed repositionings in cluster C_2 (green nodes), with the dominant level

1 ATC code G—*Genitourinary system and sex hormones* (see the histogram in Appendix A.2 Figure A4). The confirmed repositionings in cluster C_2 are torasemide (ATC level 1 code C, cardiovascular system), quinetazone (C), methazolamide (S, sensory organs), acetazolamide (S), dorzolamide (S), and brinzolamide (S). Zonisamide (N, nervous system) is a brown node (cluster C_0) but in the close vicinity of cluster C_2 ; therefore, one can expect functional overlappings [14]. Our method automatically predicts that all these drugs have genitourinary system properties, and DrugBank 5.1.8 confirms the predictions (see the detailed description in in Section 3.3.1).

Using ATC codes as references for drug repurposing is already used in the state-of-the-art contexts, although confirmations based on ATC codes are very conservative (i.e., the World Health Organization assigns new ATCs after a long and thorough process) [25,42]. Confirming the predicted drug repositionings by performing a research literature review will reveal many more confirmations [25,43]. By this logic, our analysis of DrugBank 5.0.9 does not reveal many confirmed repurposings, yet it helps tune the modularity resolution λ .

3.2. DDSN Using Drug–Gene Interactions from DrugBank 5.1.8

According to the algorithmic approach presented in Figure 1, we generated the DDSN based on the drug–gene interactions reported in DrugBank 5.1.8 and clustered DDSN using the modularity classes obtained for resolution λ_{max} (by employing Algorithm 1 with the results presented in Section 3.1). We display the largest connected component of the DrugBank 5.1.8 DDSN in Figure 5, with cluster C_0 (brown nodes) having the dominant level 1 ATC code N—*Nervous system*; clusters C_1 and C_2 (green and orange nodes) J—*Anti-infectives for systemic use*; cluster C_3 (light blue nodes) L—*Antineoplastic and immunomodulating agents*; and cluster C_4 (pink nodes) A—*Alimentary tract and metabolism*.

By running Algorithm 2 on the DDSN built with DrugBank 5.1.8 data and clustered with modularity classes at resolution λ_{max} , we generated lists of drug repurposing hints for each drug cluster. In the Supplementary Materials Table S1 file *DDSN-results.xls*, tab *DB 5.1.8 resolution 2.0*, we present the first 10 drug clusters and the entire list of drug repurposing candidates generated with Algorithm 2 (759 candidates).

Generating a list of 759 drug repurposing candidates with the latest DrugBank data and experimental confirmation is beyond the focus of our paper, and we select the first 10 drugs in each cluster in terms of betweenness/degree centrality (the methodology used in [22]) and checked them with the state-of-the-art scientific literature. For checking repositioning hints, we searched for articles in PubMed. The terms we used to search the literature were the name of the drug and the words/pharmacological terms that form level 1 of the ATC code. For example, our methodology predicted for methotrexate ATC code with level 1 J—*Anti infectives for systemic use*; we searched for the confirmation of this prediction by using keywords *methotrexate anti-infective*, as well as keywords representing therapeutic groups included in class J (i.e., *methotrexate antiviral*, *methotrexate antibacterial*, or *methotrexate antimycotic*). The confirmation results of our extensive literature check are presented in Table 1, showing the drug name, cluster number, current level 1 ATC code, predicted level 1 ATC code, and confirmation references. We also added a detailed discussion of the repurposing hints from Table 1 in Section 3.2.

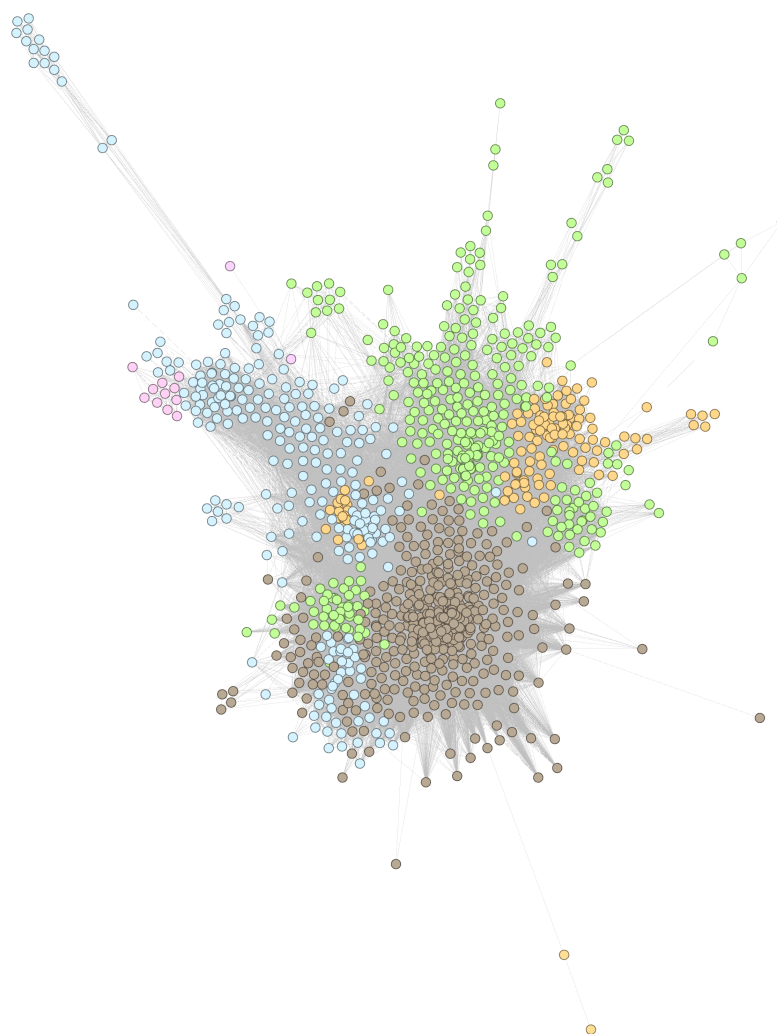


Figure 5. Drug–drug similarity network (DDSN) built with drug–gene interaction data from DrugBank 5.1.8, clustered using modularity classes for resolution $\lambda_{max} = 2.0$. The brown nodes represent drugs in cluster C_0 (479 drugs), green nodes represent drugs in cluster C_1 (346 drugs), light blue nodes represent drugs in cluster C_2 (270 drugs), orange nodes represent drugs in cluster C_3 (129 drugs), and pink nodes represent drugs in cluster C_4 (12 nodes).

We present the topological DDSN placement of Pyridoxal phosphate—predicted repositioning from cluster C_0 —in Figure 6, where a red diamond (\diamond) marks the exact position.

In Figure 7, we illustrate the position of albendazole and methotrexate in the DDSN built with DrugBank 5.0.8 data as predicted drug repositionings from cluster C_1 . Other drug repurposing candidates from cluster C_1 (presented in Table 1) are shown in Appendix B.1 and Figure A5: simvastatin, fluvastatin, lovastatin, and atorvastatin.

Figure 8 displays the DrugBank 5.0.8 DDSN placement of cholecalciferol, ergocalciferol, and calcifediol—drug repurposing candidates from cluster C_2 . In Appendix B.1, Figures A6–A8, we identify the topological positions of the other drug repurposing candidates in cluster C_2 (Table 1): meloxicam, theophylline, and chloroquine.

Table 1. The list of drug repurposing candidates generated with our methodology in Figure 1 on data from DrugBank 5.1.8, and confirmed with scientific literature. The rows correspond to drugs or drug classes (for example, simvastatin, fluvastatin, lovastatin, and atorvastatin are statins). The columns indicate—from left to right—the name, the cluster, the current level 1 ATC code in DrugBank 5.1.8, the predicted level 1 ATC code, and the confirmation references for the drug (or drug class) in each row.

Drug	Cluster	Current Level 1 ATC	Predicted Level 1 ATC	References
Pyridoxal phosphate	C ₀	A	H	[44,45]
Albendazole	C ₁	P	J	[46,47]
Methotrexate	C ₁	L	J	[48–50]
[Simvastatin	C ₁	C	J	[51,52]
Fluvastatin				
Lovastatin				
Atorvastatin				
Theophylline	C ₂	R	L	[14,53]
Meloxicam	C ₂	M	L	[54–56]
[Cholecalciferol	C ₂	M, A	L	[57,58]
Ergocalciferol				
Calcifediol				
Chloroquine	C ₂	P	L	[59–63]
[Mecasermin	C ₄	H	A	[64–66]
Mecasermin rinfabate				
Ornithine	C ₂₅	A	N	[67]

We also show the placement of drug repurposing candidates mecasermin and mecasermin rinfabate (in Figure 9, in cluster C_4 , with red diamonds \diamond) and ornithine (in Figure 10, in cluster C_{25} , with a red arrow \rightarrow).

The histograms showing the dominant properties (as level 1 ATC codes) in clusters C_0 , C_1 , C_2 , and C_4 are presented in Appendix B.2, Figure A9.

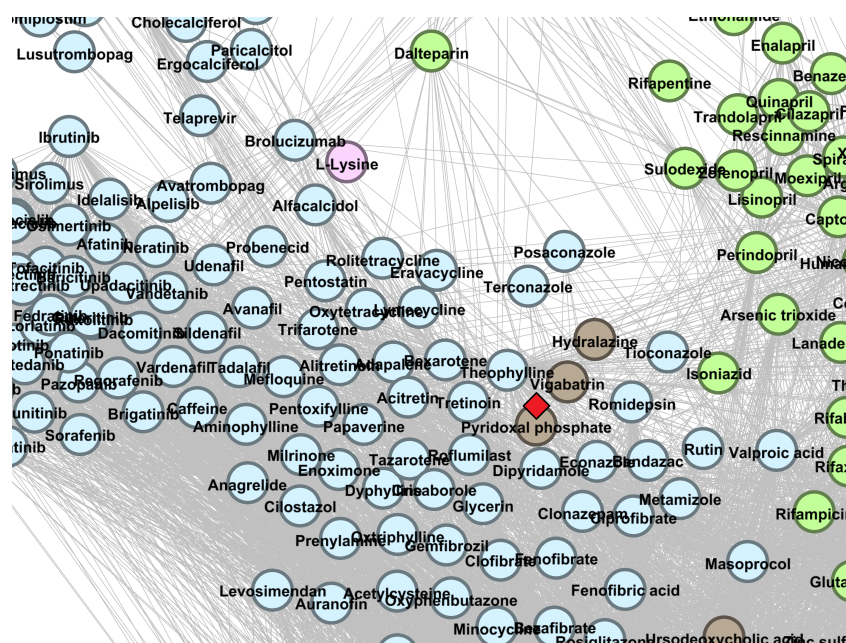


Figure 6. The DrugBank 5.1.8 DDSN network’s zoomed detail shows the repositioning within cluster C_0 (brown nodes) with a red diamond (\diamond). Our repositioning pipeline predicts that pyridoxal phosphate (currently at ATC level 1 code A—*Alimentary tract and metabolism*) has properties described by the level 1 ATC code N—*Nervous system*.

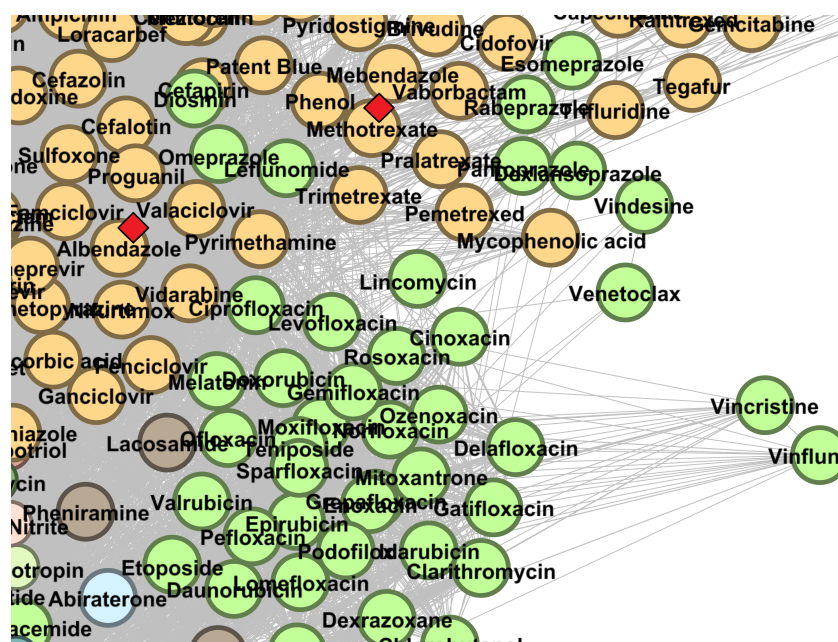


Figure 7. The DrugBank 5.1.8 DDSN network's zoomed detail shows two repositionings within cluster C_1 (green nodes) with a red diamond (◊). Our repositioning pipeline predicts that albendazole and methotrexate (currently at ATC level 1 codes P–Antiparasitic products, insecticides, and repellents and L–Antineoplastic and immunomodulating agents, respectively) have properties described by the level 1 ATC code J–Anti infectives for systemic use.

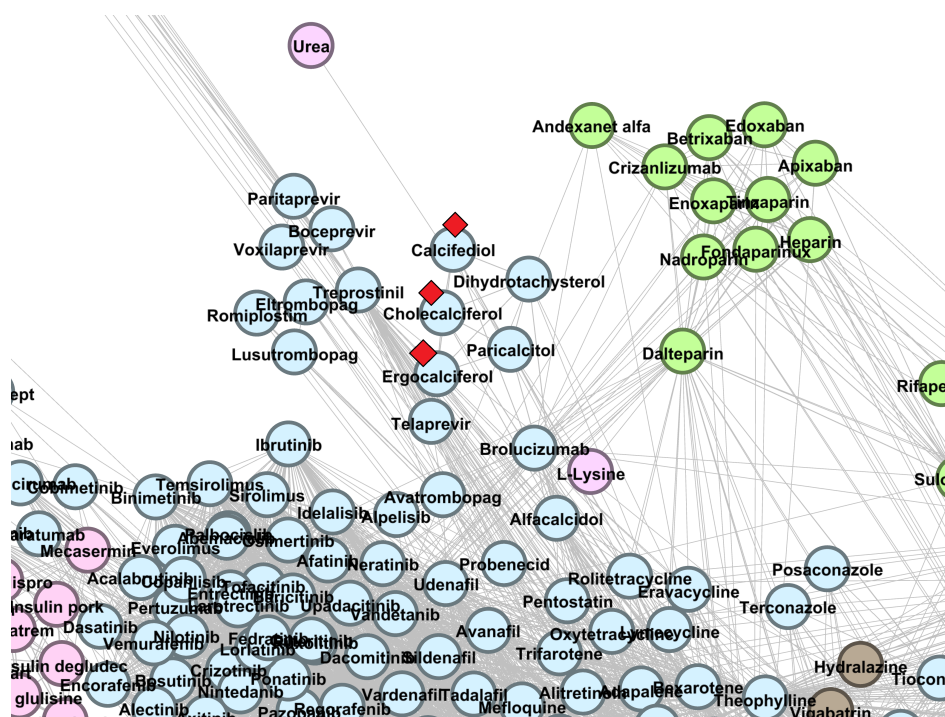


Figure 8. The DrugBank 5.1.8 DDSN network's zoomed detail shows three repositionings (vitamin D derivatives) within cluster C_2 (light blue nodes) with a red diamond (◊). Our repositioning pipeline predicts that cholecalciferol, ergocalciferol, and calcifediol (currently at ATC level 1 codes A–Alimentary tract and metabolism and M–Musculo-skeletal system) have properties described by the level 1 ATC code L–Antineoplastic and immunomodulating agents.

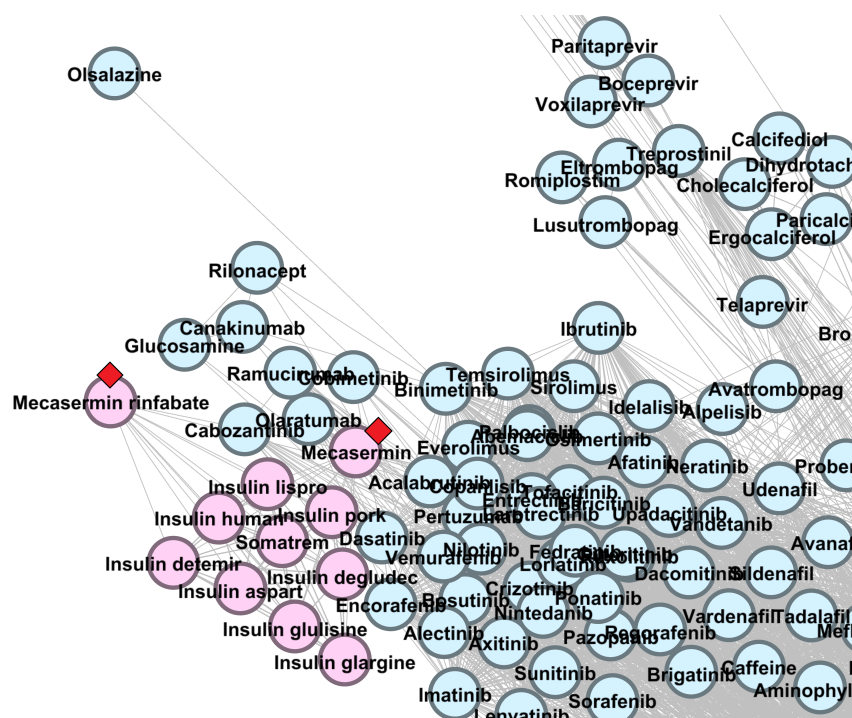


Figure 9. The DrugBank 5.1.8 DDSN network’s zoomed detail shows two repositionings within cluster C_4 (pink nodes) with a red diamond (\diamond). Our repositioning pipeline predicts that mecasermin and mecasermin rinfabate (currently at ATC level 1 codes H-*Systemic hormonal preparations, excluding sex hormones and insulins*) have properties described by the level 1 ATC code A-*Alimentary tract and metabolism*.

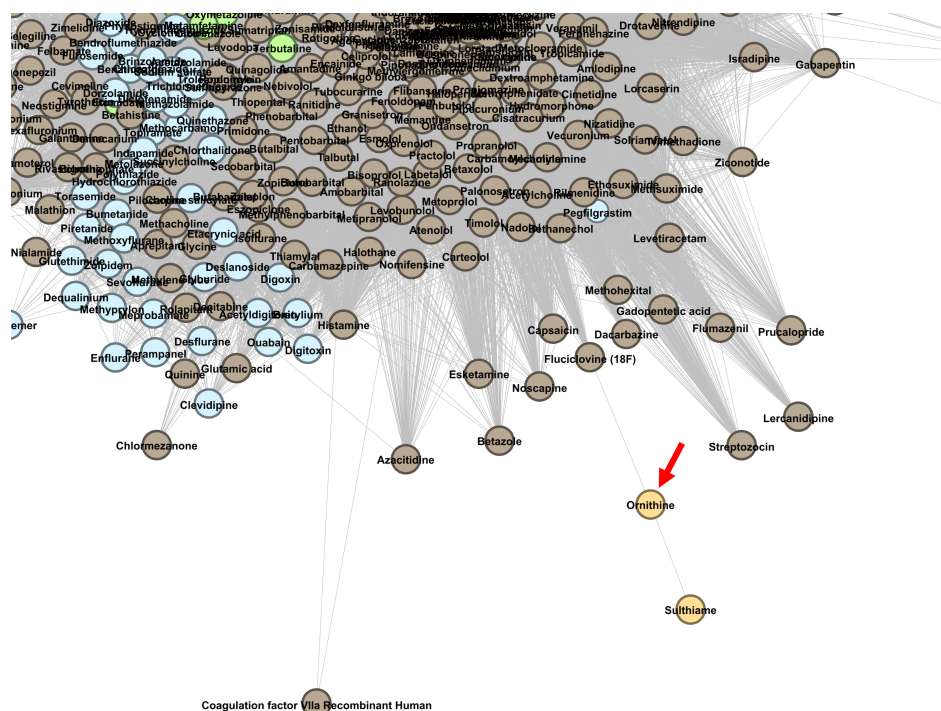


Figure 10. The DrugBank 5.1.8 DDSN network’s zoomed detail shows a repositioning within cluster C₂₅ (light orange) with a red arrow (→). Our method predicts that ornithine (currently at ATC level 1 code A–*Alimentary tract and metabolism*) has properties described by the level 1 ATC code N–*Nervous system*.

3.3. Repositioning Confirmations

3.3.1. Confirmed Drug Repositionings in DrugBank 5.0.9

This section discusses the drug repositioning hits generated with our methodology in DrugBank 5.0.9 and confirmed with the level 1 ATC codes in DrugBank 5.1.8. Our procedure confirmed the predicted hints in modularity classes 0 and 2.

Modularity Cluster C_0

In modularity cluster C_0 , DrugBank 5.1.8 confirms mepolizumab and naloxone (see Figures A1 and A2). Naloxone (ATC code V03AB15) is a μ -opioid receptor antagonist indicated in the treatment of opioid overdose. In DrugBank 5.0.9, naloxone's first level ATC is V-*Various*; its level 4 (V03AB) means naloxone is in the *Antidotes* category.

Our methodology predicts naloxone's level 1 ATC as N-*Nervous system*; the latest DrugBank 5.1.8 adds two N level 1 ATC codes to naloxone (level 4 ATC category *Natural opium alkaloids* for the combinations with hydromorphone and oxycodone), thus confirming our prediction.

Mepolizumab (ATC code L04AC06) is a monoclonal antibody acting as an antagonist of interleukin-5, included in the L-*Antineoplastic and immunomodulating agents* level 1 ATC category by DrugBank 5.0.9.

DrugBank 5.1.8 does not list the L04AC06 code anymore for mepolizumab; instead, it uses the level 1 ATC code R-*Respiratory system* (the level 4 ATC is R03DX, which includes *other systemic drugs for obstructive airways diseases*, as mepolizumab is indicated in severe eosinophilic asthma).

Modularity Cluster C_2

In modularity cluster C_2 , DrugBank 5.1.8 confirms torasemide, methazolamide, acetazolamide, dorzolamide, brinzolamide, zonisamide, and quinetazone (see Figure A3).

Torasemide, quinetazone, methazolamide, acetazolamide, dorzolamide, and zonisamide, brinzolamide (ATC codes: C03CA04, C03BA02/C03BB02, S01EC05, S01EC01, S01EC03, N03AX15, S01EC04/S01EC54) are sulfonamide compounds with various pharmacodynamic effects. According to DrugBank 5.0.9, torasemide and quinetazone are diuretics used as antihypertensive drugs, included in the C-*Cardiovascular system* level 1 ATC category. Zonisamide is an antiepileptic drug (level 1 ATC N-*Nervous system*). Methazolamide, acetazolamide, dorzolamide, and brinzolamide are carbonic anhydrase inhibitors used in glaucoma (level 1 ATC S-*Sensory organs*).

Our methodology predicts G-*Genito urinary system and sex hormones* as the level 1 ATC code for torasemide, quinetazone, methazolamide, acetazolamide, dorzolamide, zonisamide, and brinzolamide. Indeed, the latest DrugBank 5.1.8 version includes all these drugs in the G level 1 ATC category—more precisely, in the G01AE level 4 ATC category of *Anti-infective and antiseptics* having a sulfonamide-based chemical structure.

3.3.2. Drug Repositioning Hints in DrugBank 5.1.8

This section discusses the validity of some drug repositioning hints generated with our methodology in DrugBank 5.1.8; as this is the latest database version, we cannot use the same confirmation procedure based on ATC codes. Consequently, we provide evidence found in the state-of-the-art literature as confirmation clues. However, as both the number of clusters and their size prohibit an exhaustive literature search, we focus on the clusters with confirmed drug repurposing candidates—clusters C_0 , C_1 , C_2 , C_4 , and C_{25} .

Pyridoxal phosphate (cluster C_0 , ATC code A11HA06) is the active form of vitamin B6 and belongs to the A-*Alimentary tract and metabolism* level 1 ATC category, along with the rest of water-soluble and fat-soluble vitamins. Our method predicts pyridoxal phosphate as level 1 ATC code N-*Nervous system* (see Figure 6); H-S Wang et al. reported that pyridoxal phosphate controls idiopathic intractable epilepsy in children [44]. P.B. Mills and team identified two groups of patients with neonatal epileptic encephalopathy (determined by PNPO mutations) that respond to pyridoxal phosphate [45].

Albendazole (cluster C_1 , ATC code P02CA03) is an antiparasitic drug (first level ATC *P–Antiparasitic products, insecticides and repellents*) efficient in various helminthic infections. Our methodology predicts J as level 1 ATC code, suggesting potential systemic anti-infective effects (see Figure 7). Of note, ATC lists drug classes such as antivirals, antibacterials, antimycotics, and vaccines in the J–*Anti-infectives for systemic use* category. In vitro results show that albendazole exerts antifungal activity against *Aspergillus* spp. [46]; moreover, experiments on mice revealed antifungal effects against *Pneumocystis carinii* [47], confirming the new potential antifungal medical use of albendazole.

Methotrexate (cluster C_1 , ATC codes L04AX03, L01BA01) is an anticancer and immunosuppressant agent; therefore, the level 1 ATC is L–*Antineoplastic and immunomodulating agents*. We predict the first level J–*Anti-infectives for systemic use* (see Figure 7). The literature survey reveals several papers reporting in vitro antiviral effects of methotrexate in a dose-dependent manner on SARS-CoV-2 [48] and Zika virus replication [49]; methotrexate also prevents the replication of human cytomegalovirus and inhibits viral DNA synthesis [50].

Simvastatin, fluvastatin, lovastatin, and atorvastatin (cluster C_1 , ATC codes A10BH51/C10AA01/C10BX04/C10BA02/C10BX01/C10BA04, C10AA04, C10AA02/C10BA01, and C10BX15/C10AA05/C10BX03/C10BA05/C10BX11/C10BX08/C10BX06/C10BX12) are HMG-CoA reductase inhibitors (also called statins) that lower serum lipid levels, reducing the risk of cardiovascular events caused by hyperlipidemia; they are in the level 1 ATC C–*Cardiovascular system* class. The first level of their ATC code, as predicted by our method, is J–*Anti-infectives for systemic use* (see Figure A5), confirmed by literature; as such, simvastatin exhibits in vitro antimicrobial effect on methicillin-susceptible *Staphylococcus aureus* [51]. S.P. Parihar et al. [52] review the literature reporting preclinical and clinical evidence of statins effects in viral, parasitic, fungal, and bacterial infections, pointing out the factors that influence the response to statins, such as human polymorphism, metabolism, and drug interactions; this review includes data on all mentioned statins. Our algorithm predicts that all statins in cluster C_1 are potential anti-infective agents. As shown, for the statins we highlighted in Figure A5, we found literature confirming our prediction; for the other statins, new experiments and studies may provide confirmation.

Theophylline (cluster C_2 , ATC codes R03DA54, R03DA74, R03DA20, R03DA04, and R03DB04) is a methylxanthine derivative used to treat obstructive respiratory conditions, such as asthma and COPD, hence having R–*Respiratory system* as first level ATC code. Our methodology indicates theophylline's *Anticancer and immunomodulating properties*, as reflected by the predicted ATC first level L (see Figure A7), thus further confirming the repositioning proposed by our previous research [14]. Indeed, recent literature demonstrates the anticancer properties of theophylline in breast and cervical cell lines [53].

Meloxicam (cluster C_2 , ATC codes M01AC56 and M01AC06) is an oxicam derivative with anti-inflammatory and antirheumatic properties of the M–*Musculo-skeletal system* ATC category. Our network-based methodology predicts L as the first level of the ATC code (see Figure A6). The literature confirms our prediction of the anticancer properties of meloxicam: Meloxicam inhibits tumor growth in COX-2 positive colorectal cancer [54]. Tsubouchi et al. report that COX-2 plays a significant role in the pathogenesis and progression of non-small cell lung cancer (NSCLC), demonstrating the inhibitory effect of meloxicam on the NSCLC growth by preferentially inhibiting COX-2 [55]. Reference [56] shows that meloxicam is efficient in osteosarcoma in both COX-2-dependent and independent inhibitory manners.

Cholecalciferol, ergocalciferol, and calcifediol (cluster C_2 , ATC codes M05BB09/M05BX53/M05BB07/M05BB08/A11CC55/M05BB05/A11CC05/M05BB03/M05BB04, A11CC01, and A11CC06) are vitamin D analogs. Cholecalciferol (vitamin D3) is a fat-soluble vitamin (ATC level 1 A–*Alimentary tract and metabolism*, a category which includes hydro-soluble and lipo-soluble vitamins) with a well-established role in bone mineralization (ATC second level M05–*Musculo-skeletal system, drugs for treatment of bone diseases*). Ergocalciferol and calcifediol are also grouped in A–*Alimentary tract and metabolism* level 1 ATC. We predict these drugs as targeting diseases at level 1 ATC code L–*Antineoplastic and immunomodulating agents*.

lating agents (see Figure 8). There is extensive literature reporting the beneficial effects of vitamin D analogs in different cancers and highlighting the epidemiological, preclinical, and clinical results; all these back up their evolution as prophylactic and curative anticancer drugs [57,58].

Chloroquine (cluster C_2 , ATC code P01BA01) is an antimalarial drug; consequently, it belongs to the P–*Antiparasitic products, insecticides and repellents* level 1 ATC category. According to our results, the predicted first-level ATC is L–*Antineoplastic and immunomodulating agents* for chloroquine (dominant in cluster C_1 , see Figure A8). Multiple research reviews report in vitro, in vivo, and clinical trials testing chloroquine’s anticancer effect in glioblastoma [59] and other types of cancers [60–63], hence supporting the potential repositioning of chloroquine as an anticancer drug, as uncovered by our methodology.

Mecasermin and mecasermin rinfabate (cluster C_4 , ATC codes H01AC03, H01AC05) are recombinant insulin-like growth factor-1 drugs indicated in growth failure in children with primary IGF-1 deficiency and, hence, are included in the H–*Systemic hormonal preparations, excluding sex hormones and insulins*. Literature and medicine regulatory authorities reports present the secondary pharmacologic actions of mecasermin and mecasermin rinfabate, including the anabolic and insulin-like effects (i.e., hypoglycemia) [64–66]; these pharmacologic effects could place the drugs in the A–*Alimentary tract and metabolism* level 1 ATC, as predicted by our methodology (see Figure 9).

Ornithine (cluster C_{25} , ATC code A05BA06) is a non-essential amino acid indicated as nutritional supplementation and for a good liver function and included in the A–*Alimentary tract and metabolism* level 1 ATC. M. Miyake et al. suggest that L-ornithine may interfere with the Central Nervous System, following a randomized, double-blind controlled trial that demonstrated that L-ornithine relieved stress and improved sleep quality in humans compared to the placebo group [67]. Indeed, we predicted ornithine at level 1 ATC N–*Nervous system* (see Figure 10).

4. Discussion

In this section, we discuss the particularities of our method, namely the data we use, the limitations of our method and its validation with ATC codes, and the way to integrate it into an ensemble drug repositioning framework.

4.1. Drug–Gene Interactions

The method we propose in this paper uses drug–gene interaction data from DrugBank versions 5.0.9 and 5.1.8. Table 2 presents examples of drug–gene interactions and their corresponding types, as defined by DrugBank 5.1.8 (see a detailed list of drug–gene interaction types in the Supplementary Materials Table S1 file *DDSN-results.xls* and how to retrieve such drug–gene interactions from DrugBank in the GitHub page <https://github.com/GrozaVlad/Drug-repurposing-using-DDSNs-and-modularity-clustering> (last commit on 21 October 2021)).

4.2. Method Limitations

The mechanisms that influence the polypharmacological profile of drugs are highly complex. Indeed, the medicinal compound interacts with a complex system represented by the human organism. Complex systems are context-dependent; in other words, any detail at the micro-scale influences the macroscale behavior. As such, many factors can be considered when analyzing the functions of any pharmaceutical substance: from the chemical structure to various types of relationships and interactions, as well as pharmacokinetics and pharmacodynamics. By this logic, our approach is limited to considering a narrow informational angle, namely drug–gene interactions. Nonetheless, considering many mechanisms and types of data simultaneously within the same model would be prohibitively complex, and the networks would become much too dense for any centrality of community analysis. Even considering one type of information has become significantly complex; for instance, the drug–drug interaction networks in DrugBank 3.0 had an average degree

of ~20, and in DrugBank 5.1.8 the average DDI network degree is ~600). Recent literature [68–70] advances the so-called ensemble methods to address this new situation of being confronted with an overabundance rather than scarcity of data (see Section 4.4).

Table 2. Examples of drug–gene interactions listed in DrugBank.

Drug Name	Gene Name	Interaction Type
Alteplase	PLG	activator
Hydromorphone	OPRK1	agonist
Varenicline	CHRNA2	partial agonist
Prazosin	ADRA1B	antagonist
Ascorbic acid	EGLN1	chaperone
Pyridoxal phosphate	GAD1	cofactor
Vardenafil	PDE6G	allosteric modulator
Trastuzumab	ERBB2	antibody
Nusinersen	SMN2	antisense oligonucleotide
Methysergide	HTR1F	binder
Tiapropride	DRD2	blocker
Carvedilol	KCNJ4	inhibitor
Clobetasol propionate	ANXA1	inducer
Clofazimine	PPARG	modulator
Cerliponase alfa	IGF2R	ligand
Filgrastim	CSF3R	stimulator
Dalteparin	SERPINC1	potentiator
Vitamin A	RDH13	substrate
Nedocromil	CYSLTR1	suppressor
Belimumab	TNFSF13B	neutralizer
Esmirtazapine	HRH1	inverse agonist
Procainamide	DNMT1	other
Haloperidol	HTR2A	other/unknown

4.3. Labeling and Validation with ATC Codes

Employing computational methods (i.e., data mining and machine learning) in drug repositioning is generally hampered because we do not have a robust ground truth. Indeed, databases such as DrugBank record positive information about the drugs' known properties and functions, yet the absence of evidence is not evidence of absence (some drug properties may be hidden, and only future experiments can fully reveal them). That is why performance evaluation and validation of computational drug repositioning models are still an open issue; therefore, researchers adopt ad hoc, particular strategies, which are hard to compare [71]. Consequently, we resorted to making predictions with an older database version and then validating them with the latest version. However, even the latest database still cannot contain exhaustive information about drug functions. Furthermore, the negative information on drug functions/effects (stating what properties a drug does not have) will help prune the vast search space in drug repositioning. Unfortunately, negative information is scarce and scattered throughout the literature; to the best of our knowledge, no comprehensive dataset contains such data based on experimental results. As such, the existing negative information cannot be used algorithmically/automatically. As explained, one feasible method for filtering the noise and navigating the search space affected by uncertainty—an approach supported by recent research—is to integrate tools (such as the one we propose here) in ensemble methods.

Many computational drug repositioning methods based on complex networks rely on community detection and community labeling. However, labeling can be cumbersome and subjective; thus, we decided to use ATC codes, since this system is the standard for classifying medicines accepted by the WHO. Furthermore, the automated approach is fostered because the ATC code aggregates all information about a drug in a combination of

letters and numbers, which are easier to process algorithmically. The ATC code classifies drugs on five levels considering three criteria simultaneously: anatomical (A)—the first level; therapeutic (T)—levels 2 and 3; and chemical (C)—levels 4 and 5. The anatomical criterion indicates the anatomical level or the physiological organ systems on which a specific drug acts. Each anatomical level is indicated in the ATC code by a letter (e.g., A—*Alimentary tract and metabolism*, C—*Cardiovascular system*, M—*Musculoskeletal system*, or R—*Respiratory system*); the ATC system contains 14 anatomical groups. Level 2 represents the therapeutic classification criterion and is encoded by two digits. Level 3 (encoded by a letter) indicates the particular pharmacological group of the drug. Level 4 (encoded by a letter) indicates the chemical class of the drug. Level 5 is encoded by two digits the chemical structure of the drug. This paper only used the first-level ATC codes for labeling and validation of prediction, although drug function is more precisely expressed by levels 1–3; we opted perform this because the sophisticated hierarchical clustering algorithms entailed by such an approach would have unnecessarily intensified the computational character of our study.

4.4. Method Application

When the problem at hand is too complex to solve by employing a single model, machine learning uses an ensemble strategy [72], which trains several models on the same set of data to operate collectively for solving the problem. This strategy is already used in bioinformatics to approach complex problems such as motif discovery in ChIP-Seq data [73]. The problem of drug repositioning is also very complex; however, prediction accuracy is not the primary indicator of success (the benefit of correctly predicting even a few drug repositionings is more significant than the cost of experiments entailed by testing the wrong predictions [74].) As such, very recent literature advances the idea of using ensemble methods for drug repositioning [69,70].

In this context, considering that—as explained in Section 4.2—our method uses drug–gene interaction data that partially describes the behavior of drugs, we indicate the ensemble strategy as a method to use our method. As shown in Figure 11, drug repositioning prediction based on drug–gene interaction data may be Method i from the group of machine learning methods based on distinct models {Method1, Method2, . . . Method m }. The repositioning hints list i is aggregated (i.e., via voting, averaging, or other procedures) to produce a final drug repositioning hints list. The aggregation process may use pharmacological expertise, e.g., to adjust the weights of a weighted average. However, implementing the ensemble strategy is beyond the scope of this paper, which aims to analyze and promote—for the first time—the beneficial role of drug–gene interaction networks for computational drug repositioning.

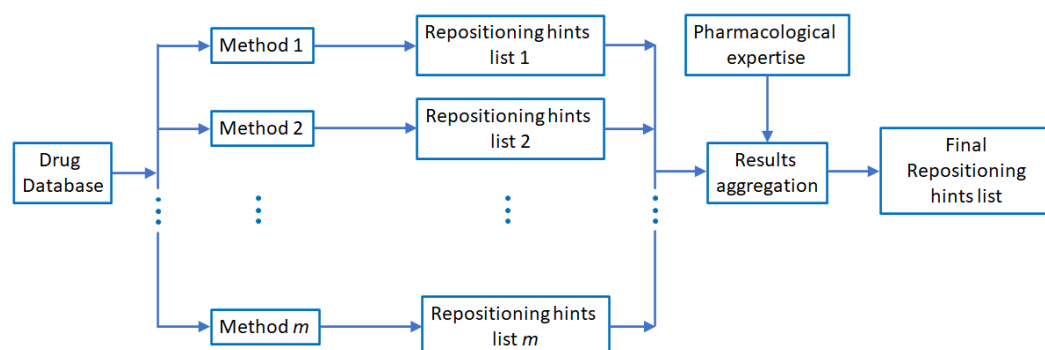


Figure 11. Overview of the ensemble strategy in drug repositioning. A group of machine learning and data mining methods {Method1, Method2, . . . Method m }, implementing various models and using distinct features (e.g., drug–drug interactions, drug–target interactions, drug–gene interactions, drug–adverse reactions relationships, pharmacokinetic properties) from the same comprehensive dataset and predicting a list of drug repositioning hints. Each method Method i generates its repositioning hints list, and an aggregation process assembles all lists in the final repositioning hints list.

5. Conclusions

In this paper, we propose a new drug repurposing methodology based on algorithmic complex network analysis. To this end, we introduce an original method of building the Drug–Drug Similarity Network (DDSN) using drug–gene interactions from DrugBank, clustering DDSN with modularity classes, and labeling each cluster with the dominant first level ATC code of drugs within the cluster. The assumption that results in drug repurposing hints is that drugs in a cluster share the dominant property of the cluster. We use an automated procedure to tune modularity resolution, to apply our methodology on a DDSN built with data from DrugBank 5.0.9, to generate the list of drug repurposing hints (i.e., drugs for which the first level ATC does not match the dominant cluster label), and to check it against ATC codes in DrugBank 5.1.8.

By running our method on the DrugBank 5.1.8 DDSN, we generated a consistent list of drug repositioning candidates; we select the top betweenness/degree drugs in each cluster and perform a preliminary validation with state-of-the-art experimental results reported in the literature. Due to the fact that we collected many literature confirmations of our method's predictions, we argue that our fully automated pipeline, based on Big Data and unsupervised machine learning, is a practical tool that can substantially narrow the enormous search space in drug repositioning.

To summarize, the overarching methodological contributions of our paper are listed as follows:

- (i) A new method to build weighted drug–drug similarity networks based on drug–gene interactions;
- (ii) An automated procedure to optimize the modularity resolution such that network clustering maximizes the number of identified drug repurposings. A known/confirmed drug repurposing is a drug with more level 1 ATC codes in the latest drug database, compared with the earlier database—used to generate the drug–drug similarity network;
- (iii) A new drug repurposing list was generated with our pipeline from the latest DrugBank 5.1.8 by analyzing the three most representative clusters.

In the present context, affected by the COVID-19 pandemic, we believe that the most promising findings/results presented in our paper are the anti-infective effects of statins, especially their potential antiviral effects. Indeed, the very recent comprehensive study [6] also finds, following in vitro screening, that fluvastatin presents what the authors call “strong effect” against SARS-CoV-2.

Considering all aspects presented in Section 4.2, we will extend our research on drug–gene interaction networks by implementing hierarchical clustering to predict ATC codes on levels 1–3, developing a dedicated cluster overlapping algorithm as a drug repositioning prediction strategy (i.e., one would reasonably expect that drugs in the overlapping zone would inherit the dominant properties of the respective clusters) and integrating the drug–gene network method into an ensemble strategy. These future objectives require substantial reliance on developing bioinformatic tools, entailing algorithm design, machine learning, and Big Data analytics.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/pharmaceutics13122117/s1>, Table S1: DDSN-results.

Author Contributions: Conceptualization, M.U. and L.U.; methodology, V.G., M.U., A.B. and L.U.; software, V.G. and A.B.; validation, V.G., A.B. and L.U.; formal analysis, M.U.; investigation, V.G. and A.B.; resources, M.U. and L.U.; data curation, V.G. and L.U.; writing—original draft preparation, M.U. and L.U.; writing—review and editing, V.G., A.B. and L.U.; visualization, V.G., M.U. and L.U.; supervision, M.U. and L.U.; project administration, M.U. and L.U.; funding acquisition, L.U. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by a grant of the Romanian Ministry of Education and Research, CCCDI-UEFISCDI, project number PN-III-P2-2.1-PED-2019-2842, within PNCDI III.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study uses only public database data.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ATC	Anatomical Therapeutic Chemical;
COPD	Chronic Obstructive Pulmonary Disease;
COX-2	Cyclooxygenase-2;
DDSN	Drug–Drug Similarity Network;
NSCLC	Non-Small Cell Lung Cancer.

Appendix A. Repositionings and Statistics for DrugBank 5.0.9 DDSN

Appendix A.1. DDSN Zoomed Details

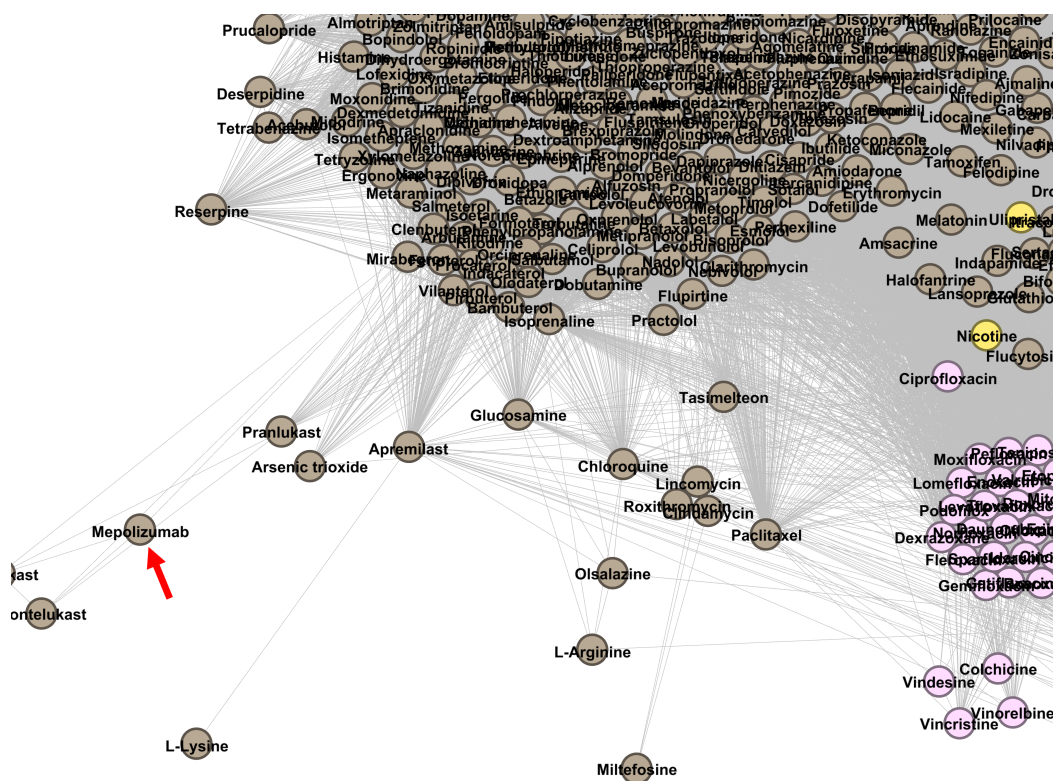


Figure A1. The zoomed detail of the DDSN network built with drug–gene interaction data from DrugBank 5.0.9, which shows the relative position of mepolizumab within cluster C_0 (brown nodes) with a red arrow (\rightarrow). Our repositioning pipeline predicts that mepolizumab—listed as antineoplastic in DrugBank 5.0.9—also acts as a drug with level 1 ATC code R (*Respiratory system*), confirmed by the more recent DrugBank version 5.1.8.

Appendix A.2. DDSN Cluster Histograms

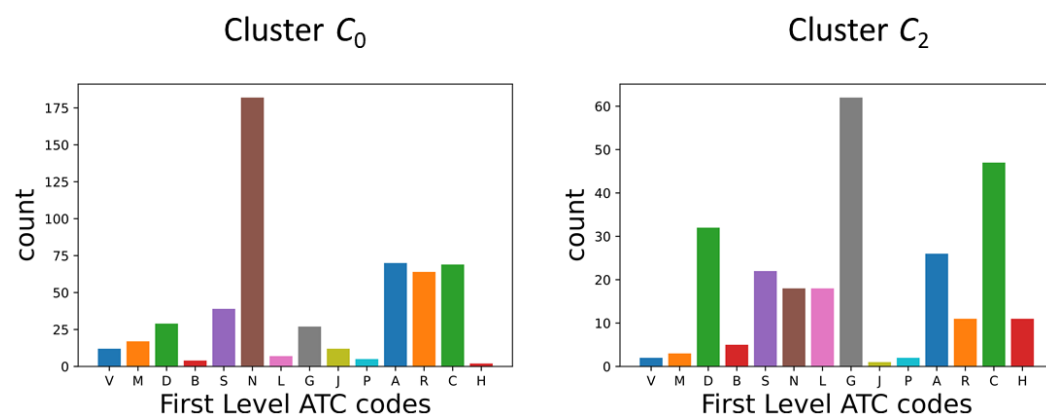


Figure A4. Histograms of level 1 ATC codes in the DrugBank 5.0.9 DDSN clusters holding drug repositionings confirmed by DrugBank 5.1.8: cluster C₀ (brown nodes) in the left panel and cluster C₂ (green nodes) in the right panel. The dominant property in cluster C₀ is N–Nervous system, with many subcluster drugs with level 1 ATC codes A, R, and C (Alimentary tract and metabolism, Respiratory system, and Cardiovascular system, respectively). The dominant properties in cluster C₂ are G, C, and D (Genito urinary system and sex hormones, Cardiovascular system, and Dermatologicals, respectively).

Appendix B. Repositionings and Statistics for DrugBank 5.1.8 DDSN

Appendix B.1. DDSN Zoomed Details

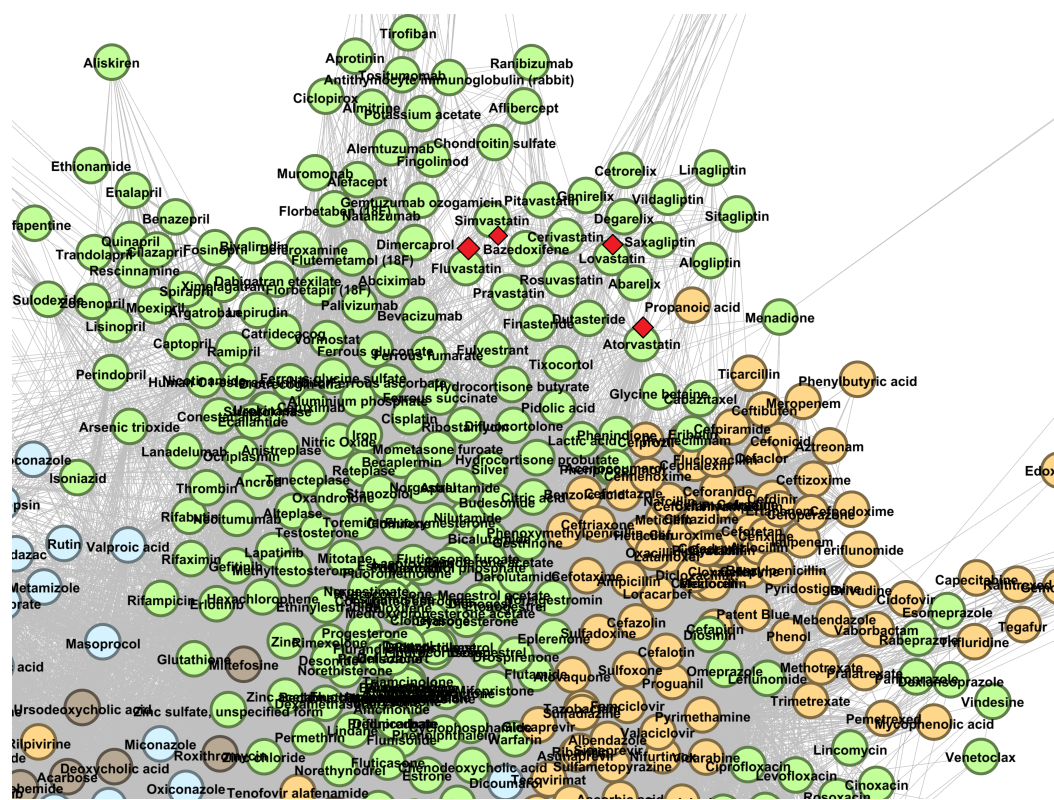


Figure A5. The DrugBank 5.1.8 DDSN network's zoomed detail shows four repositionings within cluster C₁ (green nodes) with a red diamond (◊). Our repositioning pipeline predicts that simvastatin, fluvastatin, lovastatin, and atorvastatin (currently at ATC level 1 codes C–Cardiovascular system) have properties described by the level 1 ATC code J–Anti infectives for systemic use.

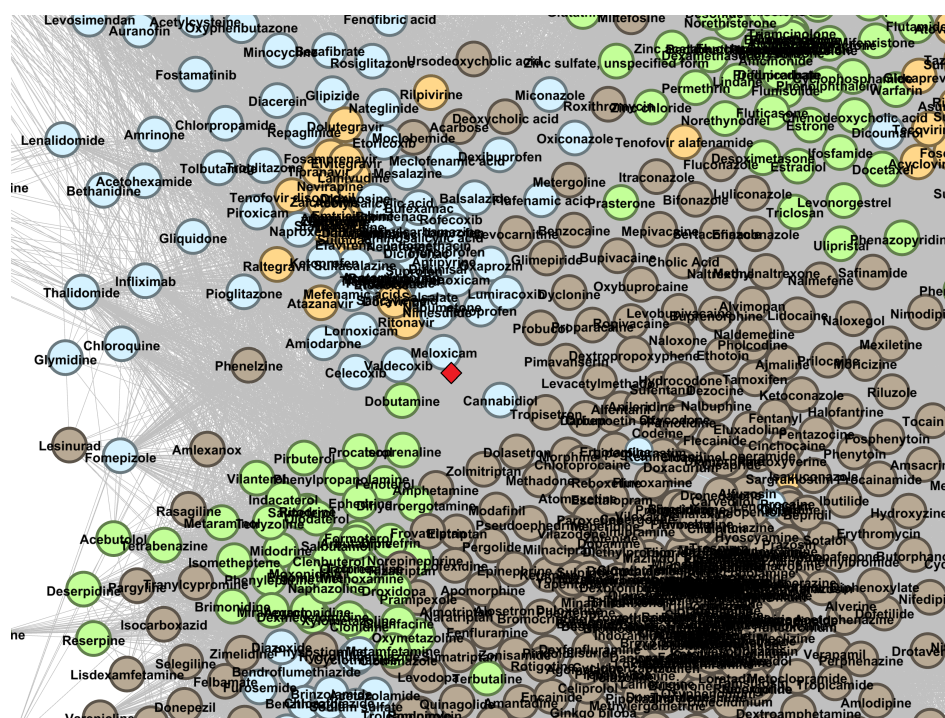


Figure A6. The DrugBank 5.1.8 DDSN network's zoomed detail shows a repositionings within cluster C_2 (light blue nodes) with a red diamond (◆). Our repositioning pipeline predicts that meloxicam (currently at ATC level 1 code M–Musculo-skeletal system) has properties described by the level 1 ATC code L–Antineoplastic and immunomodulating agents.

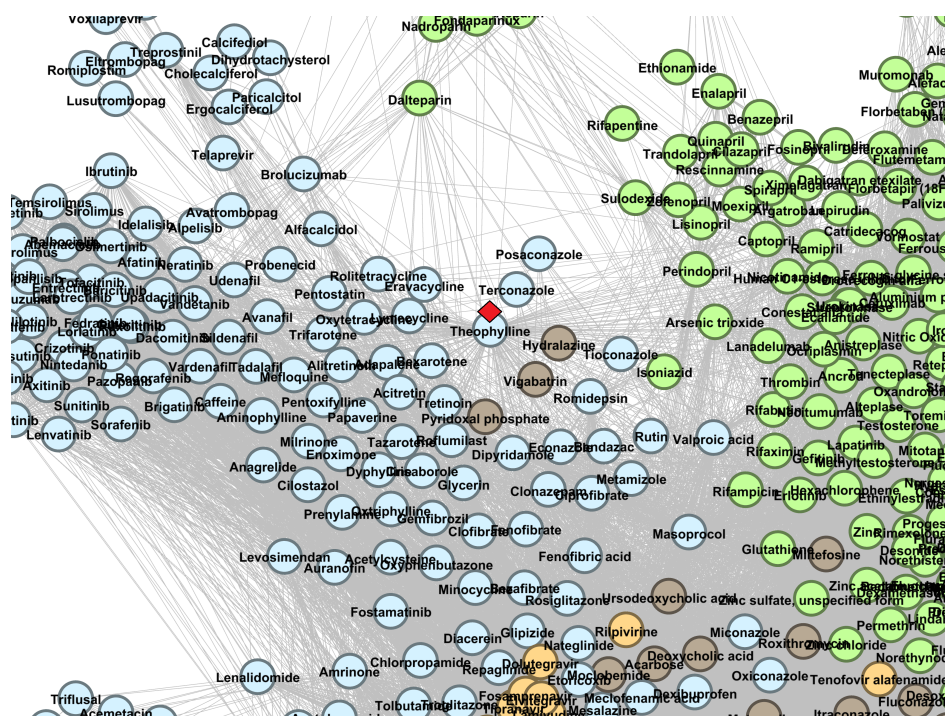


Figure A7. The DrugBank 5.1.8 DDSN network's zoomed detail shows a repositioning within cluster C_2 (light blue nodes) with a red diamond (◆). Our repositioning pipeline predicts that theophylline (currently at ATC level 1 code R–Respiratory system) has properties described by the level 1 ATC code L–Antineoplastic and immunomodulating agents.

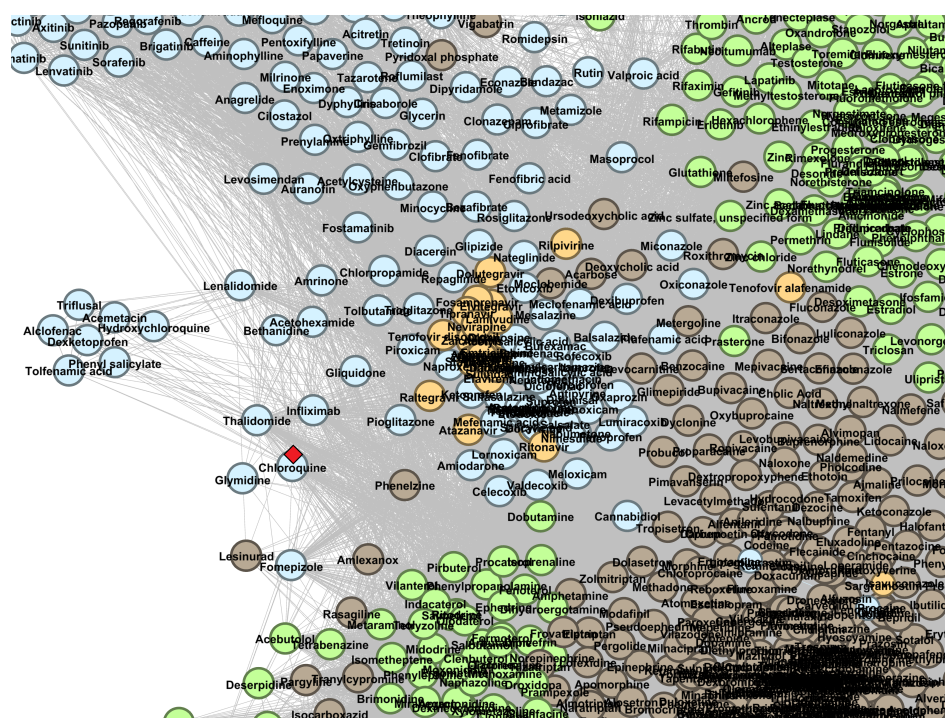


Figure A8. The DrugBank 5.1.8 DDSN network's zoomed detail shows repositioning within cluster C_2 (light blue nodes) with a red diamond (◊). Our repositioning pipeline predicts that chloroquine (currently at ATC level 1 code P–Antiparasitic products, insecticides and repellents) has properties described by the level 1 ATC code L–Antineoplastic and immunomodulating agents.

Appendix B.2. DDSN Cluster Histograms

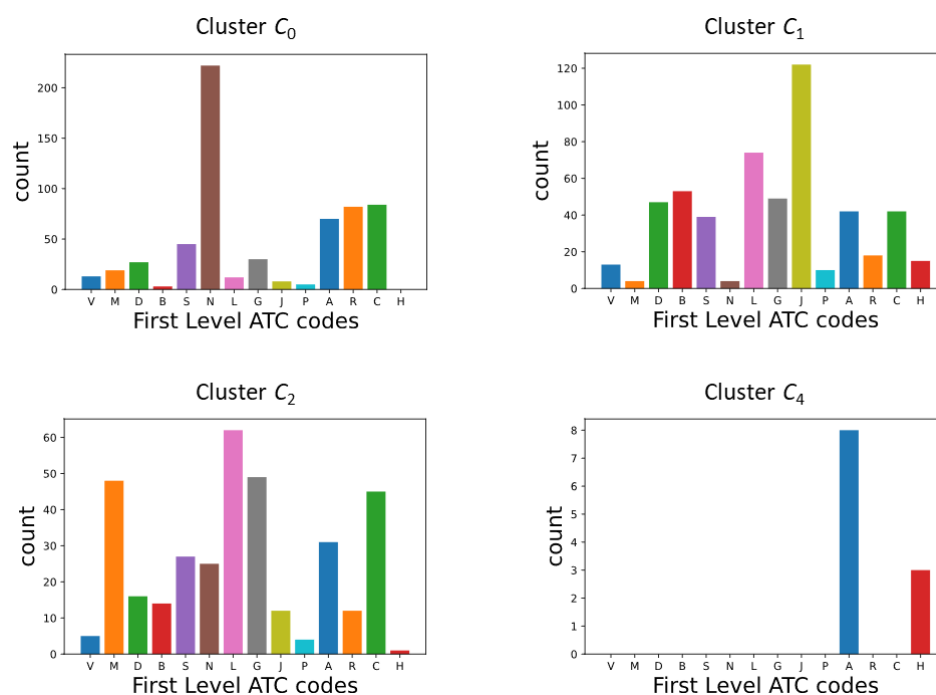


Figure A9. Histograms of level 1 ATC codes in the DrugBank 5.1.8 DDSN clusters holding drug repositionings confirmed by literature review: cluster C_0 (brown nodes), cluster C_1 (green nodes), cluster C_2 (light blue nodes), and cluster C_4 (pink nodes). The dominant property in cluster C_0 is N–Nervous System, J–Anti-infectives for systemic use in cluster C_1 , L–Antineoplastic and immunomodulating agents in cluster C_2 , and A–Alimentary Tract and Metabolism in cluster C_4 .

References

- Munos, B. Lessons from 60 years of pharmaceutical innovation. *Nat. Rev. Drug Discov.* **2009**, *8*, 959–968.
- Dickson, M.; Gagnon, J.P. The cost of new drug discovery and development. *Discov. Med.* **2009**, *4*, 172–179.
- Chen, X.Q.; Antman, M.D.; Gesenberg, C.; Gudmundsson, O.S. Discovery pharmaceuticals—Challenges and opportunities. *AAPS J.* **2006**, *8*, E402–E408.
- Pammolli, F.; Magazzini, L.; Riccaboni, M. The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug Discov.* **2011**, *10*, 428–438.
- Lombardino, J.G.; Lowe, J.A. The role of the medicinal chemist in drug discovery—Then and now. *Nat. Rev. Drug Discov.* **2004**, *3*, 853–862.
- Gysi, D.M.; Do Valle, Í.; Zitnik, M.; Ameli, A.; Gan, X.; Varol, O.; Ghiassian, S.D.; Patten, J.; Davey, R.A.; Loscalzo, J.; et al. Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2025581118.
- Meganck, R.M.; Baric, R.S. Developing therapeutic approaches for twenty-first-century emerging infectious viral diseases. *Nat. Med.* **2021**, *27*, 401–410.
- Ashburn, T.T.; Thor, K.B. Drug repositioning: Identifying and developing new uses for existing drugs. *Nat. Rev. Drug Discov.* **2004**, *3*, 673–683.
- Reddy, A.S.; Zhang, S. Polypharmacology: Drug discovery for the future. *Expert Rev. Clin. Pharmacol.* **2013**, *6*, 41–47.
- Pinzi, L.; Tinivella, A.; Caporusio, F.; Rastelli, G. Drug repurposing and polypharmacology to fight SARS-CoV-2 through inhibition of the main protease. *Front. Pharmacol.* **2021**, *12*, 84.
- Aliper, A.; Plis, S.; Artemov, A.; Ulloa, A.; Mamoshina, P.; Zhavoronkov, A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol. Pharm.* **2016**, *13*, 2524–2530.
- Zhou, Y.; Wang, F.; Tang, J.; Nussinov, R.; Cheng, F. Artificial intelligence in COVID-19 drug repurposing. *Lancet Dig. Health* **2020**, *2*, e667–e676.
- Pushpakom, S.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.; Guilleams, T.; Latimer, J.; McNamee, C.; et al. Drug repurposing: Progress, challenges and recommendations. *Nat. Rev. Drug Discov.* **2019**, *18*, 41–58.
- Udrescu, L.; Sbârcea, L.; Topîrceanu, A.; Iovanovici, A.; Kurunczi, L.; Bogdan, P.; Udrescu, M. Clustering drug-drug interaction networks with energy model layouts: community analysis and drug repurposing. *Sci. Rep.* **2016**, *6*, 1–10.
- Cheng, F.; Desai, R.J.; Handy, D.E.; Wang, R.; Schneeweiss, S.; Barabasi, A.L.; Loscalzo, J. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat. Commun.* **2018**, *9*, 1–12.
- Luo, Y.; Zhao, X.; Zhou, J.; Yang, J.; Zhang, Y.; Kuang, W.; Peng, J.; Chen, L.; Zeng, J. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nat. Commun.* **2017**, *8*, 1–13.
- AY, M.; Goh, K.I.; Cusick, M.E.; Barabasi, A.L.; Vidal, M. Drug–target network. *Nat. Biotechnol.* **2007**, *25*, 1119–1127.
- Ye, H.; Liu, Q.; Wei, J. Construction of drug network based on side effects and its application for drug repositioning. *PLoS ONE* **2014**, *9*, e87864.
- Lotfi Shahreza, M.; Ghadiri, N.; Mousavi, S.R.; Varshosaz, J.; Green, J.R. A review of network-based approaches to drug repositioning. *Brief. Bioinf.* **2018**, *19*, 878–892.
- de Oliveira, T.B.; Zhao, L.; Faceli, K.; de Carvalho, A.C. Data clustering based on complex network community detection. In Proceedings of the 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–6 June 2008; pp. 2121–2126.
- Yang, Z.; Algesheimer, R.; Tessone, C.J. A comparative analysis of community detection algorithms on artificial networks. *Sci. Rep.* **2016**, *6*, 1–18.
- Udrescu, L.; Bogdan, P.; Chiş, A.; Sîrbu, I.O.; Topîrceanu, A.; Văruţ, R.M.; Udrescu, M. Uncovering New Drug Properties in Target-Based Drug-Drug Similarity Networks. *Pharmaceutics* **2020**, *12*, 879.
- Badkas, A.; De Landtsheer, S.; Sauter, T. Topological network measures for drug repositioning. *Brief. Bioinf.* **2021**, *22*, bbaa357.
- Pérez-Moraga, R.; Forés-Martos, J.; Suay-García, B.; Duval, J.L.; Falcó, A.; Climent, J. A COVID-19 Drug Repurposing Strategy through Quantitative Homological Similarities Using a Topological Data Analysis-Based Framework. *Pharmaceutics* **2021**, *13*, 488.
- Iorio, F.; Bosotti, R.; Scacheri, E.; Belcastro, V.; Mithbaokar, P.; Ferriero, R.; Murino, L.; Tagliaferri, R.; Brunetti-Pierri, N.; Isacchi, A.; et al. Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14621–14626.
- Cheng, F.; Li, W.; Wu, Z.; Wang, X.; Zhang, C.; Li, J.; Liu, G.; Tang, Y. Prediction of polypharmacological profiles of drugs by the integration of chemical, side effect, and therapeutic space. *J. Chem. Inf. Model.* **2013**, *53*, 753–762.
- Gottlieb, A.; Stein, G.Y.; Rupp, E.; Sharan, R. PREDICT: A method for inferring novel drug indications with application to personalized medicine. *Mol. Syst. Biol.* **2011**, *7*, 496.
- Langhauser, F.; Casas, A.I.; Guney, E.; Menche, J.; Geuss, E.; Kleikers, P.W.; López, M.G.; Barabási, A.L.; Kleinschmitt, C.; Schmidt, H.H.; et al. A disease cluster-based drug repurposing of soluble guanylate cyclase activators from smooth muscle relaxation to direct neuroprotection. *NPJ Syst. Biol. Appl.* **2018**, *4*, 1–13.
- Goh, K.I.; Choi, I.G. Exploring the human diseaseome: The human disease network. *Brief. Funct. Gen.* **2012**, *11*, 533–542.

30. Goh, K.I.; Cusick, M.E.; Valle, D.; Childs, B.; Vidal, M.; Barabási, A.L. The human disease network. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8685–8690.
31. Lancichinetti, A.; Fortunato, S. Limits of modularity maximization in community detection. *Phys. Rev. E* **2011**, *84*, 066122.
32. Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucl. Acids Res.* **2006**, *34*, D668–D672.
33. 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. DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucl. Acids Res.* **2018**, *46*, D1074–D1082.
34. Merkel, D. Docker: Lightweight linux containers for consistent development and deployment. *Linux J.* **2014**, *2014*, 2.
35. McKinney, W. Pandas: A foundational Python library for data analysis and statistics. *Python High Perform. Sci. Comput.* **2011**, *14*, 1–9.
36. Hagberg, A.; Swart, P.; Chult, D.S. *Exploring Network Structure, Dynamics, and Function Using NetworkX*; Technical report; Los Alamos National Lab. (LANL): Los Alamos, NM, USA, 2008.
37. Rossetti, G.; Milli, L.; Cazabet, R. CDLIB: A python library to extract, compare and evaluate communities from complex networks. *Appl. Netw. Sci.* **2019**, *4*, 1–26.
38. Bastian, M.; Heymann, S.; Jacomy, M. Gephi: An open source software for exploring and manipulating networks. In Proceedings of the Third international AAAI Conference on Weblogs and Social Media, San Jose, CA, USA, 17–20 May 2009.
39. Zhou, T.; Ren, J.; Medo, M.; Zhang, Y.C. Bipartite network projection and personal recommendation. *Phys. Rev. E* **2007**, *76*, 046115.
40. Girvan, M.; Newman, M.E. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 7821–7826.
41. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, P10008.
42. Peng, Y.; Wang, M.; Xu, Y.; Wu, Z.; Wang, J.; Zhang, C.; Liu, G.; Li, W.; Li, J.; Tang, Y. Drug repositioning by prediction of drug's anatomical therapeutic chemical code via network-based inference approaches. *Brief. Bioinf.* **2021**, *22*, 2058–2072.
43. Tan, F.; Yang, R.; Xu, X.; Chen, X.; Wang, Y.; Ma, H.; Liu, X.; Wu, X.; Chen, Y.; Liu, L.; et al. Drug repositioning by applying 'expression profiles' generated by integrating chemical structure similarity and gene semantic similarity. *Mol. BioSyst.* **2014**, *10*, 1126–1138.
44. Wang, H.; Kuo, M.; Chou, M.; Hung, P.; Lin, K.; Hsieh, M.; Chang, M. Pyridoxal phosphate is better than pyridoxine for controlling idiopathic intractable epilepsy. *Arch. Dis. Child.* **2005**, *90*, 512–515.
45. Mills, P.B.; Camuzeaux, S.S.; Footitt, E.J.; Mills, K.A.; Gissen, P.; Fisher, L.; Das, K.B.; Varadkar, S.M.; Zuberi, S.; McWilliam, R.; et al. Epilepsy due to PNPO mutations: genotype, environment and treatment affect presentation and outcome. *Brain* **2014**, *137*, 1350–1360.
46. Berthet, N.; Faure, O.; Bakri, A.; Ambroise-Thomas, P.; Grillot, R.; Brugere, J.F. In vitro susceptibility of *Aspergillus* spp. clinical isolates to albendazole. *J. Antimicrob. Chemother.* **2003**, *51*, 1419–1422.
47. Bartlett, M.S.; Edlind, T.D.; Lee, C.H.; Dean, R.; Queener, S.F.; Shaw, M.M.; Smith, J.W. Albendazole inhibits *Pneumocystis carinii* proliferation in inoculated immunosuppressed mice. *Antimicrob. Agents Chemother.* **1994**, *38*, 1834–1837.
48. Caruso, A.; Caccuri, F.; Bugatti, A.; Zani, A.; Vanoni, M.; Bonfanti, P.; Cazzaniga, M.E.; Perno, C.F.; Messa, C.; Alberghina, L. Methotrexate inhibits SARS-CoV-2 virus replication “in vitro”. *J. Med. Virol.* **2021**, *93*, 1780–1785.
49. Beck, S.; Zhu, Z.; Oliveira, M.F.; Smith, D.M.; Rich, J.N.; Bernatchez, J.A.; Siqueira-Neto, J.L. Mechanism of action of methotrexate against Zika virus. *Viruses* **2019**, *11*, 338.
50. Lembo, D.; Griboudo, G.; Cavallo, R.; Riera, L.; Angeretti, A.; Hertel, L.; Landolfo, S. Human cytomegalovirus stimulates cellular dihydrofolate reductase activity in quiescent cells. *Intervirology* **1999**, *42*, 30–36.
51. Jerwood, S.; Cohen, J. Unexpected antimicrobial effect of statins. *J. Antimicrob. Chemother.* **2008**, *61*, 362–364.
52. Parihar, S.P.; Guler, R.; Brombacher, F. Statins: A viable candidate for host-directed therapy against infectious diseases. *Nat. Rev. Immunol.* **2019**, *19*, 104–117.
53. Chang, Y.L.; Hsu, Y.J.; Chen, Y.; Wang, Y.W.; Huang, S.M. Theophylline exhibits anti-cancer activity via suppressing SRSF3 in cervical and breast cancer cell lines. *Oncotarget* **2017**, *8*, 101461.
54. Goldman, A.P.; Williams, C.S.; Sheng, H.; Lamps, L.W.; Williams, V.P.; Pairet, M.; Morrow, J.D.; DuBois, R.N. Meloxicam inhibits the growth of colorectal cancer cells. *Carcinogenesis* **1998**, *19*, 2195–2199.
55. Tsubouchi, Y.; Mukai, S.; Kawahito, Y.; Yamada, R.; Kohno, M.; Inoue, K.; Sano, H. Meloxicam inhibits the growth of non-small cell lung cancer. *Anticancer Res.* **2000**, *20*, 2867–2872.
56. Naruse, T.; Nishida, Y.; Hosono, K.; Ishiguro, N. Meloxicam inhibits osteosarcoma growth, invasiveness and metastasis by COX-2-dependent and independent routes. *Carcinogenesis* **2006**, *27*, 584–592.
57. Deeb, K.K.; Trump, D.L.; Johnson, C.S. Vitamin D signalling pathways in cancer: potential for anticancer therapeutics. *Nat. Rev. Cancer* **2007**, *7*, 684–700.
58. Chiang, K.C.; Chen, T.C. The anti-cancer actions of vitamin D. In *Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry—Anti-Cancer Agents)*; Bentham Science Publishers: Sharjah, United Arab Emirates, 2013; Volume 13, pp. 126–139.
59. Weyerhäuser, P.; Kantelhardt, S.R.; Kim, E.L. Re-purposing chloroquine for glioblastoma: potential merits and confounding variables. *Front. Oncol.* **2018**, *8*, 335.

60. Verbaanderd, C.; Maes, H.; Schaaf, M.B.; Sukhatme, V.P.; Pantziarka, P.; Sukhatme, V.; Agostinis, P.; Bouche, G. Repurposing Drugs in Oncology (ReDO)—Chloroquine and hydroxychloroquine as anti-cancer agents. *ecancermedicalscience* **2017**, *11*, 781.
61. Dolgin, E. Anticancer autophagy inhibitors attract ‘resurgent’ interest. *Nat. Rev. Drug Discov.* **2019**, *18*, 408–410.
62. Varisli, L.; Cen, O.; Vlahopoulos, S. Dissecting pharmacological effects of chloroquine in cancer treatment: Interference with inflammatory signaling pathways. *Immunology* **2020**, *159*, 257–278.
63. Zhou, W.; Wang, H.; Yang, Y.; Chen, Z.S.; Zou, C.; Zhang, J. Chloroquine against malaria, cancers and viral diseases. *Drug Discov. Today* **2020**, 2012–2022.
64. Kemp, S.F.; Thraillkill, K.M. Mecasermin rinfabate for severe insulin-like growth factor-I deficiency. *Clin. Pract.* **2007**, *4*, 133.
65. IPLEXTM (Mecasermin Rinfabate [rDNA Origin] Injection). Available online: https://www.accessdata.fda.gov/drugsatfda_docs/label/2007/021884s001lbl.pdf (accessed on 21 October 2021).
66. INCRELEX, INN: Mecasermin. Scientific Discussion. Available online: https://www.ema.europa.eu/en/documents/scientific-discussion/increlex-epar-scientific-discussion_en.pdf (accessed on 21 October 2021).
67. Miyake, M.; Kirisako, T.; Kokubo, T.; Miura, Y.; Morishita, K.; Okamura, H.; Tsuda, A. Randomised controlled trial of the effects of L-ornithine on stress markers and sleep quality in healthy workers. *Nutr. J.* **2014**, *13*, 1–8.
68. Zhou, X.; Wang, M.; Katsyv, I.; Irie, H.; Zhang, B. EMUDRA: Ensemble of multiple drug repositioning approaches to improve prediction accuracy. *Bioinformatics* **2018**, *34*, 3151–3159.
69. Wang, J.; Wang, W.; Yan, C.; Luo, J.; Zhang, G. Predicting Drug-Disease Association Based on Ensemble Strategy. *Front. Genet.* **2021**, *12*, 548.
70. Ghorbanali, Z.; Zare-Mirakabad, F.; Mohammadpour, B. DRP-VEM: Drug repositioning prediction using voting ensemble. *arXiv* **2021**, arXiv:2110.01403.
71. Jarada, T.N.; Rokne, J.G.; Alhadj, R. A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. *J. Cheminf.* **2020**, *12*, 1–23.
72. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, e1249.
73. Lihu, A.; Holban, Ş. A review of ensemble methods for de novo motif discovery in ChIP-Seq data. *Brief. Bioinf.* **2015**, *16*, 964–973.
74. Xue, H.; Li, J.; Xie, H.; Wang, Y. Review of drug repositioning approaches and resources. *Int. J. Biol. Sci.* **2018**, *14*, 1232.