KGML-xDTD: A Knowledge Graph-based Machine Learning Framework for Drug Treatment Prediction and Mechanism Description

A PREPRINT

Chunyu Ma
Huck Institutes of the Life Sciences
Pennsylvania State University
State College, PA 16802, USA
cqm5886@psu.edu

Zhihan Zhou
Department of Computer Science
Northwestern University
Evanston, IL 60208, USA
zhihanzhou@u.northwestern.edu

Han Liu
Department of Computer Science
Northwestern University
Evanston, IL 60208, USA
hanliu@northwestern.edu

David Koslicki
Department of Computer Science
Department of Biology
Huck Institutes of the Life Sciences
Pennsylvania State University
State College, PA 16802, USA
dmk333@psu.edu

ABSTRACT

Background: Computational drug repurposing is a cost- and time-efficient approach that aims to identify new therapeutic targets or diseases (indications) of existing drugs/compounds. It is especially critical for emerging and/or orphan diseases due to its cheaper investment and shorter research cycle compared with traditional wet-lab drug discovery approaches. However, the underlying mechanisms of action (MOAs) between repurposed drugs and their target diseases remain largely unknown, which is still a main obstacle for computational drug repurposing methods to be widely adopted in clinical settings.
Results: In this work, we propose `KGML-xDTD`: a **K**nowledge **G**raph-based **M**achine **L**earning framework for e**x**plainably predicting **D**rugs **T**reating **D**iseases. It is a two-module framework that not only predicts the treatment probabilities between drugs/compounds and diseases but also biologically explains them via knowledge graph (KG) path-based, testable mechanisms of action (MOAs). We leverage knowledge-and-publication based information to extract biologically meaningful "demonstration paths" as the intermediate guidance in the Graph-based Reinforcement Learning (GRL) path-finding process. Comprehensive experiments and case study analyses show that the proposed framework can achieve state-of-the-art performance in both predictions of drug repurposing and recapitulation of human-curated drug MOA paths.
Conclusions: `KGML-xDTD` is the first model framework that can offer KG-path explanations for drug repurposing predictions by leveraging the combination of prediction outcomes and existing biological knowledge and publications. We believe it can effectively reduce "black-box" concerns and increase prediction confidence for drug repurposing based on predicted path-based explanations, and further accelerate the process of drug discovery for emerging diseases.

Keywords Drug Repurposing · Reinforcement Learning · Biomedical Knowledge Graph

1 Introduction

Traditional drug development is a time-consuming process (from initial chemical identification to clinical trials and finally to FDA approval) that takes around 10-15 years and also comes along with billions-of-dollars investments and high failure rates [Berdigaliyev and Aljofan, 2020]. Considering the rapid pace of novel disease evolution, it is urgent to find a more efficient and economical drug discovery method. Fortunately, it has been observed that a single drug

can often be effective in treating multiple diseases. For example, thalidomide was originally used as an anti-anxiety medication [Miller, 1991], and was later found to have the anti-cancer potential for the treatment of cancers [Verheul et al., 1999, Singhal et al., 1999]. Hence, drug repurposing, also known as the identification of new uses for existing drugs/compounds, might bring us hope to address this urgent need with the advantage of a shorter research cycle, lower development cost, and more preexisting safety tests.

Existing drug repurposing approaches can roughly be categorized into experimental-based approaches (e.g., binding affinity assays [Kairys et al., 2019], phenotypic screening [Aulner et al., 2019]), clinical-based approaches (e.g., off-label drug use analysis [Rusz et al., 2021]), and computational-based approaches (e.g., chemical-structure-based [Swamidass, 2011], and GWAS-based approaches [Sanseau et al., 2012]). Compared with the former two approaches, the computational approaches are more cost- and time-efficient, particularly when the goal is to prioritize a large number of target drugs/compounds for follow-up experimental investigation. Among all computational drug repurposing methods, the integration of multiple biomedical data sources into a so-called **biomedical knowledge graph** (BKG) for drug discovery has become popular in recent year [Bonner et al., 2022] due to the increasing availability of curated biomedical databases such as DrugBank [Wishart et al., 2017a], ChEMBL [Gaulton et al., 2012], HMDB [Wishart et al., 2017b] and the advancement of semantic web techniques [Kanza and Graham Frey, 2021]. There are three types of existing BKGs: database-based BKGs, literature-based BKGs, and mixed BKGs. The database-based BKGs (e.g., *Hetionet* [Himmelstein et al., 2017], *BioKG* [Walsh et al., 2020], *CBKH* [Su et al., 2021]) are constructed by integrating biomedical data and their relations stored in existing biological databases. The literature-based BKGs (e.g., *GNBR* [Percha and Altman, 2018]) are built by leveraging Natural Language Processing (NLP) techniques to extract semantic information from a large amount of available biomedical literature and electronic health record (EHR) data, which are mostly disease-specific [Zhang et al., 2021, Wang et al., 2021a, Li et al., 2020]. The mixed BKGs (e.g., *CKG* [Santos et al., 2022], *RTX-KG2* [Wood et al., 2022]) are generated by combining the knowledge sources from the above two methods.

Based on these BKGs, several machine learning methods have been proposed or implemented for drug repurposing prediction by treating it as a link prediction task in the BKGs. For example, Himmelstein et al. [2017] used the so-called degree-weighted path count (DWPC) to assess the prevalence of 1,206 metapaths and then classified drug-disease treatment relations by fitting these DWPC features to a logistic regression model. Ioannidis et al. [2020] proposed a novel graph neural network model I-RGCN to learn the node and relation embeddings for the Covid-19 drug repurposing task. Zhang et al. [2021] recently predicted the possible drugs for Covid-19 with five existing popular knowledge graph completion methods (e.g. TransE [Bordes et al., 2013], RotatE [Sun et al., 2019], DistMult [Yang et al., 2014], ComplEx [Trouillon et al., 2016], and STELP [Wang et al., 2021b]). Although some of these models have shown good performance in drug repurposing prediction on the small-scale BKGs, none have been scaled to massive BKGs with more than millions of nodes and edges and make a comprehensive comparison. More importantly, most of them lack the biological explanatory ability for their predictions, which limits their applicability in clinical research.

Currently, there are few computational models designed for drug repurposing *explanations*. A common and intuitive explanation based on a biomedical knowledge graph for drug repurposing leverages the semantic BKG-based paths between given drug-disease pairs. Sosa et al. [2020] applied a graph embedding model UKGE [Chen et al., 2019], which utilizes the weighted (the frequency of relation appeared in literature) relation edges in a literature-based KG *GNBR*, to identify new indications of drugs for rare diseases and then explain the results via the highest-ranking paths based on confidence scores. However, this method is only applicable in the literature-based BKGs with the weighted edge information. Most BKGs using database-based knowledge don't contain such information. Sang et al. [2019] proposed GrEDeL that combines the TransE embedding method with a Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) model to predict drug-disease relation. By using the embeddings of BKG-paths as model input for predictions, they can provide path-based explanations. However, they claimed that the effectiveness of the approach relies heavily on the NLP tool SemRep, which is reported to have high false positives in named entity recognition [Kilicoglu et al., 2020]. Also, they didn't fully evaluate how biologically reasonable their predicted path-based mechanisms of action (MOAs) are.

Besides the existing methods above, we view reinforcement learning (RL) as a promising solution for drug repurposing explanation. RL models solve the decision-making problem, in which an agent learns how to take appropriate actions to maximize cumulative rewards through interactions with the environment. RL has achieved widespread success in various domains, including games, recommendation systems, healthcare, transportation, etc [Li, 2019]. Graph Reinforcement Learning (GRL), first proposed around in 2017, aims to solve graph mining tasks such as link prediction [Chen et al., 2022], adversarial attacks [Sun et al., 2020], and relational reasoning [Zhou et al., 2021]. Unlike its applications in other domains, one of the biggest challenges in GRL is finding an appropriate reward to guide the path searching in specific domains. To address the issue of finding biologically reasonable BKG-based paths for drug repurposing, it is crucial to incorporate biomedical domain knowledge to guide the path-finding process. Liu et al. [2021] developed an RL-based model "PoLo" that utilizes the biological meta-paths identified in Himmelstein et al.

[2017] via the "DWPC" method to supervise path searching for drug repurposing. However, the "PoLo" model does not scale to a massive and complex BKG (e.g., CKG and RTX-KG2) due to its dependence on the "DWPC" method that is reported to be computationally inefficient [Womack et al., 2019].

In this article, we describe `KGML-xDTD`: a **K**nowledge **G**raph-based **M**achine **L**earning framework for e**x**plainably predicting **D**rugs **T**reating **D**iseases, which contains two modules for both drug repurposing prediction and explanation. We propose to amplify the ability of RL model in biologically meaningful path searching by utilizing the biologically meaningful "demonstration paths" and pre-trained drug-repurposing model probability as rewards. We incorporate this idea into the appropriate models (e.g., GraphSAGE [Hamilton et al., 2017], Random Forest, and ADAC RL [Zhao et al., 2020] models) and then make them applicable to the explainable drug repurposing problem at massive data scale and complexity. By comparing with the existing popular drug repurposing models and evaluating the predicted paths with an expert-curated path-based drug MOA database *DrugMechDB* [Mayers et al., 2020], we show that the proposed model framework can achieve state-of-the-art performance in both predictions of drug repurposing and recapitulation of human-curated drug MOA paths provided by DrugMechDB. In further case studies, by comparing the model predictions with the real regulatory networks, we show that the proposed framework effectively identifies biologically reasonable BKG-based paths for real-world applications.

2 Materials and Methods

2.1 Datasets

2.1.1 Customized Biomedical Knowledge Graph

To accommodate biomedical-reasonable predictions of drugs' indications and their mechanisms of action, the ideal biomedical knowledge graph should integrate biomedical knowledge from comprehensive and diverse databases and publications, as well as accurately identify and merge different identifiers representing the same biological entity into one (For example, "CHEBI:2367" and "CHEMBL455626" are two distinct identifiers separately presented in ChEBI database [Degtyarenko et al., 2008] and ChEMBL database [Gaulton et al., 2012] but represent the same compound "abyssinone I"). Thus, we utilize the canonicalized version of the Reasoning Tool X Knowledge Graph 2 (*RTX-KG2c*) [Wood et al., 2022], one of the largest open-source biomedical knowledge graph (BKG) that integrates knowledge from extensive human-curated and Publication-based databases, and has been widely used in the Biomedical Data Translator Project [Consortium, 2019, Translator Consortium, 2019]. Compared to other commonly used open-source BKGs mentioned above, *RTX-KG2c* is a biolink-model-based [1] standardized [Unni et al., 2022] and regularly-updated BKG that efficiently merges biologically and semantically equivalent nodes and edges via multiple curation steps. The version 2.7.3 of *RTX-KG2c* that we use contains around 6.4M nodes and 39.3M edges with knowledge from 70 public biomedical sources, where all biological concepts (e.g., "ibuprofen") are represented as vertices and all concept-predicate-concept (e.g., "ibuprofen - increases activity of - GP1BA gene") are presented as edges. For drug repurposing purposes, we customized *RTX-KG2c* with four principles (Please see more details in Appx. A): 1). excluding the nodes whose categories are irrelevant to drug repurposing explanation (e.g., "GeographicLocation" and "Device"); 2). filtering out the low-quality edges based on our criteria; 3). removing the hierarchically redundant edges; 4). excluding all drug-disease edges. After these processing steps, 3,659,165 nodes with 33 distinct categories (Figure 1 a) and 18,291,237 edges with 74 distinct types (Figure 1 b) are left in our customized biomedical knowledge graph, which is used for downstream model training.

2.1.2 Data Sources for Model Training

To train the `KGML-xDTD` framework for drug repurposing prediction and its explanation, we utilize four human-curated and NLP-derived training datasets:

- **MyChem Data** [Xin et al., 2018] is provided by the BioThings API collection [Xin et al., 2016], which contains up-to-date human-curated annotations regarding indication and contraindication for drugs. We use drug-disease pairs with the relation "indication" as true positives while those with "contraindication" as true negatives.
- **SemMedDB Data** [Kilicoglu et al., 2012] is provided by the Semantic MEDLINE Database (SemMedDB), which leverages natural language processing (NLP) techniques to extract semantic triples with "treats" and "negatively treats" relations from PubMed abstracts. We use drug-disease pairs with the relation "treats" as true positives while those with "negatively treats" as true negatives.
- **NDF-RT Data** [Brown et al., 2004] is provided by National Drug File – Reference Terminology from Veterans Health Administration (VHA) which contains human-curated information on drug interaction, indications, and

[1] The biolink model [Unni et al., 2022] is a universal and standardized BKG ontology framework

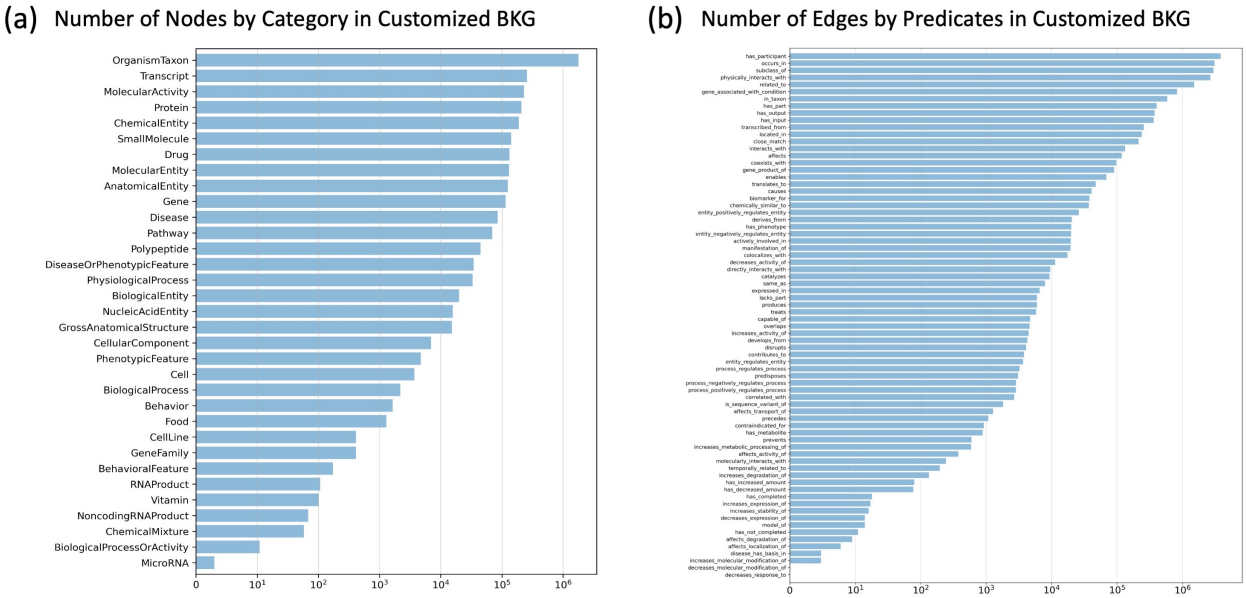

Figure 1: Number of nodes by category (a) and number of edges by predicate (b) in customized Biomedical Knowledge Graph (BKG).

contraindications. We use drug-disease with therapeutics label "indications" as true positives while those with "contraindications" as true negatives.

- **RepoDB Data** [Brown and Patel, 2017] is a standard set of successful and failed drug-disease pairs in clinical trials collected by the Blavatnik Institute at Harvard Medical School. We use drug-disease with the status "approved" as true positives while those with "withdrawn" as true negatives.

We further filter drug-disease pairs from SemMedDB Data due to publication bias and possible NLP mistakes by using both the co-occurrence frequency and the PubMed-publication-based Normalized Google Distance (NGD) [Cilibrasi and Vitanyi, 2007] defined below:

$$NGD(c1, c2) = \frac{max\{log\mathcal{N}(c1), log\mathcal{N}(c2)\} - log\mathcal{N}(c1, c2)}{logN - min\{log\mathcal{N}(c1), log\mathcal{N}(c2)\}} \tag{1}$$

where $c1$ and $c2$ are two biological concepts used in the customized BKG; $\mathcal{N}(c1)$ and $\mathcal{N}(c2)$ respectively represent the total number of unique PubMed IDs associated with $c1$ and $c2$; $\mathcal{N}(c1, c2)$ is the total number of unique PubMed IDs shared between $c1$ and $c2$; N is the total number of pairs of Medical Subject Heading (MeSH) terms annotations in PubMed database. Only the SemMedDB drug-disease pairs with at least 10 supporting publications and an NGD score of 0.6 or lower are left for the downstream model training.

These datasets are pooled together and then processed by 1). mapping the raw identifiers of drugs and diseases to the identifiers used in the customized BKG; 2) removing duplicate drug-disease pairs in both the true positive set and the true negative set. Table 1 shows the drug-disease pair count from each data source after data pre-processing.

Table 1: Pair count of true positive (indications) and true negative (contraindications or no effect) data from four data sources after data pre-processing.

Source	**True Positive (Treats)**	**True Negative (Not Treat)**
MyChem	3,663	26,795
SemMedDB	8,255	11
NDF-RT	3,421	5,119
RepoDB	2,127	738
Shared	3,971	526
Total	21,437	33,189

Note that 'shared' means those pairs are from two or more data sources.

2.1.3 DrugMechDB

DrugMechDB[2] [Mayers et al., 2022], to our best knowledge, is the first human-curated path-based database for explaining the mechanism of action (MOA) from a drug to a disease in an indication, with 3,593 MOA paths for 3,327 unique drug-disease pairs. These paths are extracted from free-text descriptions from DrugBank, Wikipedia, and other literature sources, and then have been curated by subject matter experts and also follow the schema of Biolink model. Hence, we can match them to nodes and edges used in the RTX-KG2 BKG via the *Node Synonymizer* function [Wood et al., 2022]. Since the length of predicted MOA paths generated by the `KGML-xDTD` framework is fixed to 3 in this study, we consider those 3-hop BKG-based paths of which all four nodes show up in the complete DrugMechDB MOA paths are the correct matched paths. Thus, we find 472 unique drug-disease pairs, of which each has at least one such correct matched path in all possible 3-hop paths between drug and disease in the customized BKG. We use these matched paths as true positive biologically meaningful paths for the evaluation of the model-predicted paths.

2.2 Model Framework

The model framework of `KGML-xDTD` consists of two modules: a drug repurposing prediction module that combines the advantages of GraphSAGE [Hamilton et al., 2017] and a Random Forest model, and an MOA prediction module that utilizes an adversarial actor-critic reinforcement learning (RL) model. We show the overview of the entire model framework in Figure 2. The implementation details of each module in `KGML-xDTD` framework are presented in Appx B.

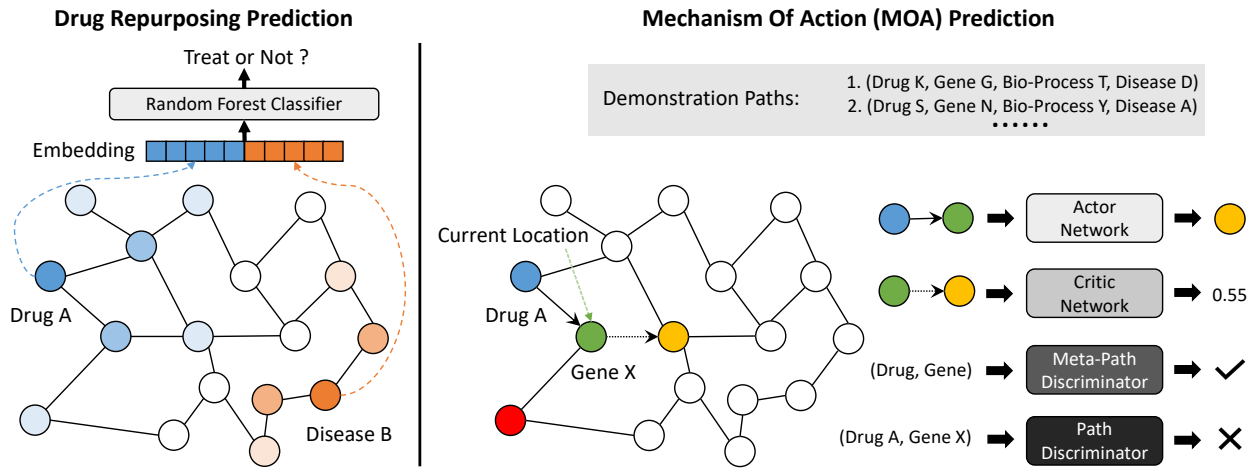

Figure 2: Illustration of entire `KGML-xDTD` model framework: drug repurposing prediction module (left) and mechanism of action (MOA) prediction module (right).

2.2.1 Notations

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a directed biomedical knowledge graph, where each node $v \in \mathcal{V}$ represents a biological entity (e.g., a specific drug, disease, gene, or pathway, etc.) and each edge $e \in \mathcal{E}$ represents a biomedical relationship (e.g., *interacts-with, see more in Figure 1 b*). We use $\mathcal{V}^{\text{drug}}$ to represent all the drug nodes (the nodes with the categories of "Drug" and "Small Molecule" in the customized BKG) and $\mathcal{V}^{\text{disease}}$ to represent all the disease nodes (the nodes with the categories of "Disease", "PhenotypicFeature", "BehavioralFeature" and "DiseaseOrPhenotypicFeature" in the customized BKG). For each notation, we use bold formatting to represent its embedding (e.g., $\boldsymbol{v}$ represents the embedding of v).

2.2.2 Drug Repurposing Prediction

Drug repurposing aims to identify new indications of existing drugs/compounds. We solve it as a link prediction problem on the graph $\mathcal{G}$. Specifically, given any drug-disease pair (v_i, v_j) where $v_i \in \mathcal{V}^{\text{drug}}$ and $v_j \in \mathcal{V}^{\text{disease}}$, we predict the probability that drug i can be used to treat disease j. We first use GraphSAGE to calculate the embedding for each node. Ideally, the node embeddings should contain two kinds of information: node attributes and node neighborhoods.

To capture the neighborhood information, we optimize GraphSAGE to encourage neighbor nodes to have similar embeddings and non-neighbor nodes to have distinct embeddings. Specifically, we perform random walks for each node to collect its neighborhood information and train the model to maximize a node's similarity with its neighbor

[2] `https://sulab.github.io/DrugMechDB`

nodes. For a node u, the loss is calculated as:

$$L_{\mathcal{G}}(\boldsymbol{z}_u) = -\log(\sigma(\boldsymbol{z}_u^\top \boldsymbol{z}_v)) - k \cdot \mathbb{E}_{v_n \sim P_n(v)} \log(\sigma(\boldsymbol{z}_u^\top \boldsymbol{z}_{v_n})) \tag{2}$$

where $\boldsymbol{z}_u, \boldsymbol{z}_v$ are respectively the embeddings of nodes u, v, σ is the sigmoid function, v is a node that co-occurs with u in fixed-length random walks, P_n represents negative sampling distribution, and k indicates the number of negative samples (nodes not in u's fixed-length neighborhood).

To capture the node attributes information, we utilize the PubMedBERT model [Gu et al., 2022], a pre-trained language model designed for biomedical texts, to generate a node attribute embedding for each node based on the concatenation of the node's name and category. We further compress the embeddings to 100 dimensions with Principal Components Analysis (PCA) to reduce memory usage and use them as the initial node feature for GraphSAGE. In this way, the final GraphSAGE embedding of each node should contain the information regarding both graph topology and node attributes. We concatenate the GraphSAGE embeddings of drug-disease pairs and use them as input of a Random Forest model to classify each drug-disease pair into one of the "not treat", "treat", and "unknown" classes. We obtain "treat" and "not treat" drug-disease pairs from four data sources (described in Sec. 2.1.2). We generate "unknown" drug-disease pairs through negative sampling [Mikolov et al., 2013], that is, replacing the drug or disease identifier in each "treat" drug-disease pair with a random drug or disease identifier to generate a new pair that does not appear in both the "treat" and "not treat" classes. Specifically, for each unique "treat" drug-disease pair, we respectively replace its drug identifier with 30 other random drug identifiers as well as replace its disease identifier with 30 other random disease identifiers to make 60 new drug-disease pairs for the "unknown" class.

2.2.3 Mechanism of Action (MOA) Prediction

When potential indications of a given drug are identified by the drug repurposing prediction module, a natural yet essential question is: can we biologically explain the predictions? We solve this by employing a reinforcement learning (RL) model to predict the BKG-based MOA paths, which are the paths on the knowledge graph from drug nodes to disease nodes. These BKG-based MOA paths can semantically describe an abstract biological process of how a drug treats a disease.

Demonstration paths To encourage the RL agent to terminate the path searching at the expected diseases through a biologically reasonable path, we leverage so-called "demonstration paths", a set of biologically likely paths (e.g., drug1-gene1-protein3-disease1), that explains the underlying reasons for why a drug can treat a disease. We extract 396,705 demonstration paths from the customized BKG using the known drug-target interactions collected from two curated biomedical data sources: DrugBank (v5.1) and Molecular Data Provider (v1.2)[3], as well as the PubMed-publication-based Normalized Google Distance (NGD) (see Equation 1). We show more details regarding demonstration path extraction in Appx. C.

Adversarial Actor-critic Reinforcement Learning We formulate the MOA prediction as a path-finding problem and adapt the Adversarial Actor-Critic Reinforcement Learning model [Zhao et al., 2020] to solve it. Reinforcement learning is defined as a Markov Decision Process (MDP) which contains:

States: Each state s_t at time t is defined as $s_t = \big(v_{drug}, v_t, (v_{t-1}, e_t), \ldots, (v_{t-K}, e_{t-(K-1)})\big)$ where $v_{drug} \in \mathcal{V}^{\text{drug}}$ is a given starting drug node; $v_t \in \mathcal{V}$ represents the node where the agent locates at time t; the tuple $\big(v_{t-K}, e_{t-(K-1)}\big)$ represents the previous Kth node and $(K-1)$th predicate. For the initial state s_0, the previous nodes and predicates are substituted by a special dummy node and predicate. We concatenate the embedding of all nodes and predicates of s_t to get the state embedding $\boldsymbol{s}_t$, where the node embeddings are node attribute embeddings generated with the PubMedBERT model (see Sec. 2.2.2) and the predicate embeddings employ one-hot vectors.

Actions: The action space A_t of each node v_t includes a self-loop action a_{self} and the actions to reach its outgoing neighbors in the graph $\mathcal{G}$. Due to memory limitation and extremely large outdegree of certain nodes in the knowledge graph, we prune the neighbor actions based on the PageRank scores if a node has more than 3,000 neighbors. Specifically, we let $A_t = (a_{self}, a_1, \ldots, a_k, \ldots, a_{n_{v_t}})$ where n_{v_t} is out-degree of node $v_t \in \mathcal{V}$. For each action $a_t = (e_t, v_{t+1}) \in A_t$ taken at time t, we concatenate its node and predicate embeddings to obtain action embedding $\boldsymbol{a}_t$. We learn two embedding matrices $E^{N_{\text{n}} \times \text{d}}$ and $E^{N_{\text{p}} \times \text{d}}$ respectively for nodes and predicates[4], where d represents the embedding dimension, N_{n} represents the number of nodes in graph, and N_{p} represents the number of predicate categories in graph.

Rewards: During the path searching process, the agent only receives a terminal reward $R_{e,T}$ from environment (that is, there is no intermediate reward from environment: $R_{e,t} = 0, \forall t < T$). Let v_T be the last node of the path, and $\mathcal{N}_{drug}$

[3] `https://github.com/NCATSTranslator/Translator-All/wiki/Molecular-Data-Provider`

[4] Each sub-network uses separate embedding matrices.

be the known diseases that drug v_{drug} can treat. The terminal reward $R_{e,T}$ from environment is calculated with the drug repurposing model via:

$$R_{e,T} = \begin{cases} 1, & \text{if } v_T \in \mathcal{N}_{drug}. \\ p_{treat}, & \text{if } v_T \notin \mathcal{N}_{drug}; v_T \in \mathcal{V}^{\text{disease}} \text{ and } f(v_{drug}, v_T) \text{ is predicted as "treat".} \\ 0, & \text{if } v_T \notin \mathcal{N}_{drug}; v_T \in \mathcal{V}^{\text{disease}} \text{ and } f(v_{drug}, v_T) \text{ is not predicted as "treat".} \\ -1, & \text{if } v_T \notin \mathcal{V}^{\text{disease}}. \end{cases}$$

where p_{treat} is the "treat" class probability predicted by the drug repurposing model f.

The Adversarial Actor-critic RL model consists of four sub-networks that share the same model architecture MLP^i (note that i represents the id of each sub-network described later, such as a for *actor network*, c for *critic network*, etc.) but with different parameters:

$$\text{MLP}^i(X) = BA(BA(XW_1^i + b_1^i)W_2^i + b_2^i)W_3^i + b_3^i \tag{3}$$

where $\{W_1^i, W_2^i, W_3^i, b_1^i, b_2^i, b_3^i\}$ are the parameters and biases of linear transformations, BA represents a batch normalization layer followed by an ELU activation function.

Actor network: The actor network learns a path-finding policy π_θ (note that θ represents all parameters of actor network) to guide the agent to choose an action a_t from the action space A_t based on the current state s_t:

$$\pi_\theta(a_t|s_t, A_t) = \text{softmax}(\boldsymbol{A}_t \odot \text{MLP}^a(\boldsymbol{s}_t)) \tag{4}$$

where $\boldsymbol{A}_t$ is the embedding matrix of the action space A_t; $\odot$ represents the dot product. Here, $\pi_\theta(a_t|s_t, A_t)$ represents the probability of choosing action a_t at time t from the action space A_t given the state s_t.

Critic network: The critic network [Lillicrap et al., 2015] estimates the expected reward $Q_\phi(s_t, a_t)$ (note that ϕ represents all parameters of the critic network) if the agent takes the action a_t at the state s_t by:

$$Q_\phi(s_t, a_t) = \text{MLP}^c(\boldsymbol{s}_t) \odot \boldsymbol{a}_t \tag{5}$$

Path discriminator network: Since the RL agent only receives a terminal reward $R_{e,T}$ from environment indicating whether it reaches an expected target, to encourage the agent to find biologically reasonable paths and provide intermediate rewards, we further guide it with demonstration paths. This network is essentially a binary classifier that distinguishes whether a path segment (s_t, a_t) is from demonstration paths or generated by the actor network. We treat all the known demonstration path segments (s_t^D, a_t^D) as positive samples and all actor-generated non-demonstration path segments (s_t^{ND}, a_t^{ND}) as negative samples. The path discriminator network $D_p(s, a) = \text{sigmoid}(\text{MLP}^p(\boldsymbol{s} \oplus \boldsymbol{a}))$, where $\boldsymbol{s}$ and $\boldsymbol{a}$ are respectively the embeddings of the state s and the action a; $\oplus$ represents the concatenation operator, is optimized with:

$$L_p = -\mathbb{E}_{(s,a)\sim P_D}[\log(D_p(s,a))] - \mathbb{E}_{(s,a)\sim P_A}[\log(1 - D_p(s,a))] \tag{6}$$

where P_D and P_A respectively represent the demonstration path segment distribution and the actor-generated non-demonstration path segment distribution. Based on the probability $D_p(s_t, a_t)$, the path-discriminator-based intermediate reward $R_{p,t}$ is calculated as:

$$R_{p,t} = \log(D_p(s_t, a_t)) - \log(1 - D_p(s_t, a_t)). \tag{7}$$

Meta-Path discriminator network: Similar to the path discriminator, this network aims to judge whether the meta-path of the actor-generated paths is similar to that of demonstration paths. The meta-path is the path of node categories (e.g., ["Drug"→"Gene"→"BiologicalProcess"→"Disease"]). Similarly, the meta-path discriminator $D_m(M) = \text{sigmoid}(\text{MLP}^m(\boldsymbol{M}))$, where $\boldsymbol{M}$ is the embedding of the meta-path M defined as the concatenation of learned category embeddings of all nodes that appear in the path, is also a binary classifier where the meta-paths of demonstration paths are treated as positive samples while others are negative samples. We optimize it with the following loss:

$$L_m = -\mathbb{E}_{M\sim P_D^M}[\log(D_m(M))] - \mathbb{E}_{M\sim P_A^M}[\log(1 - D_m(M))] \tag{8}$$

where P_D^M and P_A^M respectively represent the demonstration meta-path distribution and the actor-generated non-demonstration meta-path distribution. The intermediate reward $R_{m,t}$ generated by the meta-path discriminator is calculated by:

$$R_{m,t} = \log(D_m(M)) - \log(1 - D_m(M)). \tag{9}$$

The integrated intermediate reward R_t at time t is then calculated as:

$$R_t = \alpha_p R_{p,t} + \alpha_m R_{m,t} + (1 - \alpha_p - \alpha_m)\gamma^{T-t} R_{e,T} \tag{10}$$

where $\alpha_p \in [0, 1]$ and $\alpha_m \in [0, 1 - \alpha_p]$ are hyperparameters, γ is the decay coefficient, and $R_{e,T}$ is defined in the "Rewards" section above.

To optimize the critic network, we minimize the Temporal Difference (TD) error [Sutton, 1988] with loss:

$$L_c = \mathrm{TD}^2 = [(R_t + Q_\phi(s_{t+1}, a_{t+1})) - Q_\phi(s_t, a_t)]^2. \tag{11}$$

Since the goal of the actor network is to achieve the largest expected reward by learning an optimal actor policy, we optimize the actor network by maximizing $J(\theta) = \mathbb{E}_{a\sim\pi_\theta}[Q_\phi(s_t, a)]$. We use the REINFORCE algorithm [Williams, 1992] to optimize the parameters. To encourage more diverse exploration in finding paths, we use the entropy of π_θ as a regularization term and optimize the actor network with the following stochastic gradient of the loss function L_a:

$$\nabla_\theta L_a = -\nabla_\theta J(\theta) = -\mathbb{E}_{\pi_\theta}[\nabla_\theta \mathrm{TD} \log \pi_\theta(a_t|s_t)] - \alpha \nabla_\theta \mathrm{entropy}(\pi_\theta) \tag{12}$$

where π_θ is the action probability distribution based on the actor policy, and α is the entropy weight.

We follow Zhao et al. [2020] to train the Adversarial Actor-critic RL model in a multi-stage way. First, we initialized the actor network using the behavior cloning method [PomerleauDean, 1991] in which the training set of demonstration paths is used to guide the sampling of the agent with Mean Square Error (MSE) loss. Then, in the first z epochs, we freeze the parameters of the actor network and the critic network and respectively train the path discriminator network and meta-path discriminator network by minimizing L_p and L_m. After z epochs, we unfreeze the actor network and the critic network and optimize them together by minimizing a joint loss $L_{joint} = L_a + L_c$.

3 Results

3.1 Evaluation Settings

3.1.1 Data Split

The post-processed drug-disease pairs (described in Sec. 2.1.2) are split into training, validation, and test sets where the drug-disease pairs of each unique drug are randomly split according to a ratio of 8/1/1. For example, let's say drugA has 10 known diseases that it treats (e.g., drugA-disease1, . . . , drugA-disease10), 8 pairs are randomly split into the training set, 1 pair is to the validation set, 1 pair to the test set. With this data split method, the model can be exposed to every drug in the training set, which complies with our goal of predicting new indications of known drugs and their potential mechanisms of action (MOAs) based on the MOA of known target diseases.

3.1.2 Evaluation Metrics

The proposed framework `KGML-xDTD` is evaluated on two types of tasks: predicting drug-disease "treat" probability (i.e., drug repurposing prediction) as well as identifying biologically reasonable BKG-based MOA paths from all candidates (i.e., MOA prediction). These two tasks are evaluated based on classification accuracy-based metrics (e.g., accuracy, macro f1 score) and ranking-based metrics (e.g., mean percentile rank, mean reciprocal rank, and proportion of ranks smaller than K) defined as follows:

Accuracy (ACC) is the fraction of the model classification is correct, computed as:

$$\text{ACC} = \frac{\text{Number of correct classifications}}{\text{Total number of drug-disease pair classifications}} \tag{13}$$

Macro F1 score (Macro-F1) is the unweighted mean of all the per-class F1 scores:

$$F1^c = 2 * \frac{precision^c \times recall^c}{precision^c + recall^c} \qquad \text{Macro-F1} = \frac{1}{|C|}\sum_{c\in C} F1^c \tag{14}$$

where C presents classification classes (e.g.,"treat," "not treat," and "unknown").

Mean Percentile Rank (MPR) is the average percentile rank of the 3-hop correct DrugMeshDB-based matched path (described in Sec. 2.1.3) of true positive drug-disease pairs:

$$\text{MPR} = \frac{1}{|PR|}\sum_{pr\in PR} pr \tag{15}$$

where PR is a list of percentile ranks of correct matched paths of true positive drug-disease pairs ("treat" category).

Mean Reciprocal Rank (MRR) is the average inverse rank of true positive drug-disease pairs ("treat" category) or their 3-hop correct matched paths:

$$\text{MRR} = \frac{1}{|R|}\sum_{r\in R} \frac{1}{r} \tag{16}$$

where R is a list of ranks of true positive drug-disease pairs or their correct matched paths.

Hit@K is the proportion of ranks smaller thank K for true positive drug-disease pairs ("treat" category) or their 3-hop correct matched paths:

$$\text{Hit@K} = \frac{1}{|R|} \sum_{r \in R} |r \leq k| \tag{17}$$

where R is a list of ranks of true positive drug-disease pairs or their correct matched paths.

3.1.3 Drug Repurposing Prediction Evaluation Method

We utilize the metrics *ACC* and *Macro-F1* to measure the accuracy of drug repurposing prediction of our `KGML-xDTD` framework while using ranking-based metrics *MRR* and *Hit@K* to show its capability in reducing false positive (i.e., the false drug-disease pairs ranking higher among possible drug-disease candidates). We use the following three methods to generate non-true-positive drug-disease candidates for each true positive drug-disease pair for the *MRR* and *Hit@K* calculation:

- **Drug-rank-based replacement**: For each true positive drug-disease pair, the drug-rank-based replacement pairs are generated by replacing the drug entity with each of all 274,676 other drugs in the customized BKG while excluding all known true positive drug-disease pairs.
- **Disease-rank-based replacement**: For each true positive drug-disease pair, the disease-rank-based replacement pairs are generated by replacing the disease entity with each of all 124,638 other diseases in the BKG while excluding all known true positive drug-disease pairs.
- **Combined Replacement**: For each true positive drug-disease pair, the combined replacement pairs are the combination of all replacement pairs of the above two methods. All known true positive drug-disease pairs are excluded from these replacement pairs.

Due to the massive size of possible drug-disease candidates, some baseline models (e.g., GAT and GraphSAGE+SVM) are not applicable in this setting within a reasonable time (e.g., a week). Thus, we also use a small subset of drug-disease replacement (we use 1000 random drug-disease pairs from the combined replacement set above: 500 with drug id replacement and 500 with disease id replacement) to calculate *MRR* and *Hit@K* in order to compare `KGML-xDTD` with all baselines. In addition, since the drug repurposing prediction module of `KGML-xDTD` framework does 3-class classification while other baselines do 2-class classification, for a fair comparison, we re-calculate *ACC* and *Macro-F1* for `KGML-xDTD` by excluding the "unknown" class.

3.1.4 MOA Prediction Evaluation Method

For the evaluation of MOA prediction, we use the DrugMechDB [Mayers et al., 2020] to obtain the expert-verified MOA paths as ground-truth data and match each biological concept in these verified MOA paths to the biological entities used in the customized BKG, and then generate the correct matched paths (described in Sec. 2.1.3). we first calculate the path scores for all 3-hop KG paths between drug and disease with the path-finding policy learned from Adversarial Actor-critic Reinforcement Learning (RL) model using equation:

$$\text{path score} = \sum_{i=1}^{k} \delta^{i-1} \times \log(P_i \times N_i) \tag{18}$$

where k is the number of hops in this path; δ is a decay coefficient (we set it to 0.9 in this study); P_i represents the probability of choosing action a_i in the ith hop following this path based on the trained RL model; N_i is the number of possible actions in the ith hop.

With these path scores, we obtain the ranks of the matched BKG-based MOA paths and calculate the ranking-based metrics *MPR*, *MRR*, and *Hit@K*. For those drug-disease pairs with multiple BKG-based MOA paths, we use the highest ranks of their paths. We compare `KGML-xDTD` with the baseline models based on these metrics to show the capability of MOA prediction module of `KGML-xDTD` in identifying biologically reasonable BKG-based MOA paths from a massive and complex BKG with comparably low false positive. In addition, we further perform two case studies to evaluate the effectiveness of `KGML-xDTD` in identifying the biologically reasonable BKG-based paths.

3.2 Drug Repurposing Prediction Evaluation

For drug repurposing prediction evaluation, we compare the `KGML-xDTD` model framework against several state-of-the-art (SOTA) KG-based models and variants of `KGML-xDTD` for drug repurposing prediction based on the method described in Sec. "Drug Repurposing Prediction Evaluation Method" on pag 3.1.4.

We use eight different SOTA KG-based models as baseline models that are commonly used for BKG-based drug repurposing [Zhang et al., 2021, Hsieh et al., 2020]. `TransE` [Bordes et al., 2013], `TransR` [Lin et al., 2015], `RotatE` [Sun et al., 2019] are the translation-distance-based models that regard a relation (e.g., "treats") as a "translation"/"rotation" (e.g., a kind of spatial transformation) from a head entity (e.g., a drug node) to a tail entity (e.g., a disease node). `DistMult` [Yang et al., 2014] is a bilinear model that measures the latent semantic similarity of a knowledge-graph triple (head entity, relation/predicate, tail entity) with a trilinear dot product. `ComplEx` [Trouillon et al., 2016] and `ANALOGY` [Liu et al., 2017] are the extensions of `DistMult` that consider more complex relations (e.g., asymmetric relations). `SimpLE` [Kazemi and Poole, 2018] is a tensor-factorization-based model to learn the semantic relation of a knowledge-graph triple. `GAT` [Veličković et al., 2018] is a popular graph neural model that leverages the important graph topology structure based on self-attention mechanism for graph-associated tasks (e.g., link prediction). Implementation details of these baselines are presented in Appx. D.

Besides these SOTA baseline models, we also compare the drug repurposing prediction module in `KGML-xDTD` with its several variants to show the effectiveness of model components. For example, to show efficacy of the combination of GraphSage and Random Forest (RF), we use a pure GraphSAGE for link prediction (`GraphSAGE-link`), the combination of GraphSage and logistic model (`GraphSAGE-logistic`), and the combination of GraphSage and Support Vector Machine (SVM) model (`GraphSAGE-SVM`). To demonstrate the effectiveness of node attribute embeddings (described in Sec. 2.2.2) in improving repurposing prediction, we conduct an ablation experiment that replaces node attribute embeddings (NAEs) with random embeddings (initialized with the Xavier method [Glorot and Bengio, 2010]) as GraphSage initialized embeddings (`KGML-xDTD w/o NAE`); to support rationality of setting "unknown" class through negative sampling (described in Sec. 2.2.2), we modify the drug repurposing prediction module for 2-class classification[5] (`2-class KGML-xDTD`) as a baseline comparison model.

Table 2: The performance comparison of Drug Repurposing Prediction (DRP) between `KGML-xDTD` and different baseline models based on test dataset (described in Sec. 3.1.1). The top panel shows the performance of state-of-the-art (SOTA) baseline models; the middle panel shows the performance of variants of `KGML-xDTD` model framework; the bottom panel shows the performance of `KGML-xDTD` model framework.

Model	**Accuracy**	**Macro F1 score**	**MRR**	**Hit@1**	**Hit@3**	**Hit@5**
TransE	0.708	0.708	0.280	0.117	0.301	0.449
TransR	0.858	0.855	0.301	0.123	0.344	0.518
RotatE	0.704	0.704	0.253	0.075	0.283	0.450
DistMult	0.555	0.495	0.173	0.041	0.143	0.259
ComplEx	0.624	0.460	0.132	0.020	0.106	0.194
ANALOGY	0.594	0.465	0.180	0.045	0.147	0.274
SimplE	0.599	0.472	0.163	0.038	0.137	0.244
GAT	**0.936**	**0.934**	0.003	0.001	0.001	0.001
GraphSAGE-link	0.919	0.915	0.002	0	0	0
GraphSAGE+logistic	0.791	0.784	0.002	0	0	0
GraphSAGE+SVM	0.807	0.793	0.002	0	0	0
KGML-xDTD w/o NAEs	0.909 (0.898*)	0.891 (0.892*)	0.150	0.029	0.138	0.243
2-class KGML-xDTD	0.929	0.925	0.271	0.177	0.310	0.381
KGML-xDTD (ours)	0.935 (0.930*)	0.923 (0.926*)	**0.356**	**0.206**	**0.407**	**0.522**

1. The values with * inside the parenthesis are the adjusted results by excluding the "unknown" category for a fair comparison.
2. The ranking metrics (e.g., "MRR" and "Hit@K") are calculated with the random drug-disease replacement method (i.e., we use 1,000 random drug-disease pairs to calculate the ranks of true positive drug-disease pairs in test dataset).
3. "NAEs" represents node attribute embeddings.

Table 2 shows the performance of `KGML-xDTD` model framework and other baseline models in the task of drug repurposing prediction based on the metrics described in Sec. 3.1.2 and test dataset along with 1,000 random drug-disease replacement pairs. As shown in the table, on the one hand, the `KGML-xDTD` outperforms most of the baseline models and achieves comparable performance as `GAT` in classification-based metrics (e.g., accuracy, macro f1 score), indicating its effectiveness in classifying known "treat" and "not treat" drug-disease pairs with both attribute and neighborhood information on the knowledge graph. On the other hand, `KGML-xDTD`'s exceptional performance in ranking-based metrics shows its superiority over baselines in identifying new indications of existing drugs out of a large number of possible drug-disease pairs with relatively low false positives (Figure 3 further supports this conclusion with

[5] only considers true positive and true negative.

three different "complete" replacement methods), which is of great importance for guiding clinical research. Besides, by comparing `2-class KGML-xDTD` with the vanilla GraphSAGE model (e.g., `GraphSAGE-link`), we demonstrate the effectiveness of the Random Forest model over a neural network classifier in this task. The comparison between `KGML-xDTD w/o NAE` and `KGML-xDTD` shows that the `KGML-xDTD` benefits from the use of node attribute embeddings for drug repurposing prediction while the comparison with `2-class KGML-xDTD` indicates the effectiveness of using negative sampling to generate "unknown" drug-disease pairs for model training. With the "unknown" drug-disease pairs, the `KGML-xDTD` model achieves significant improvement in ranking-based metrics, which is essential when applying to real-world drug repurposing because it can reduce the false positives.

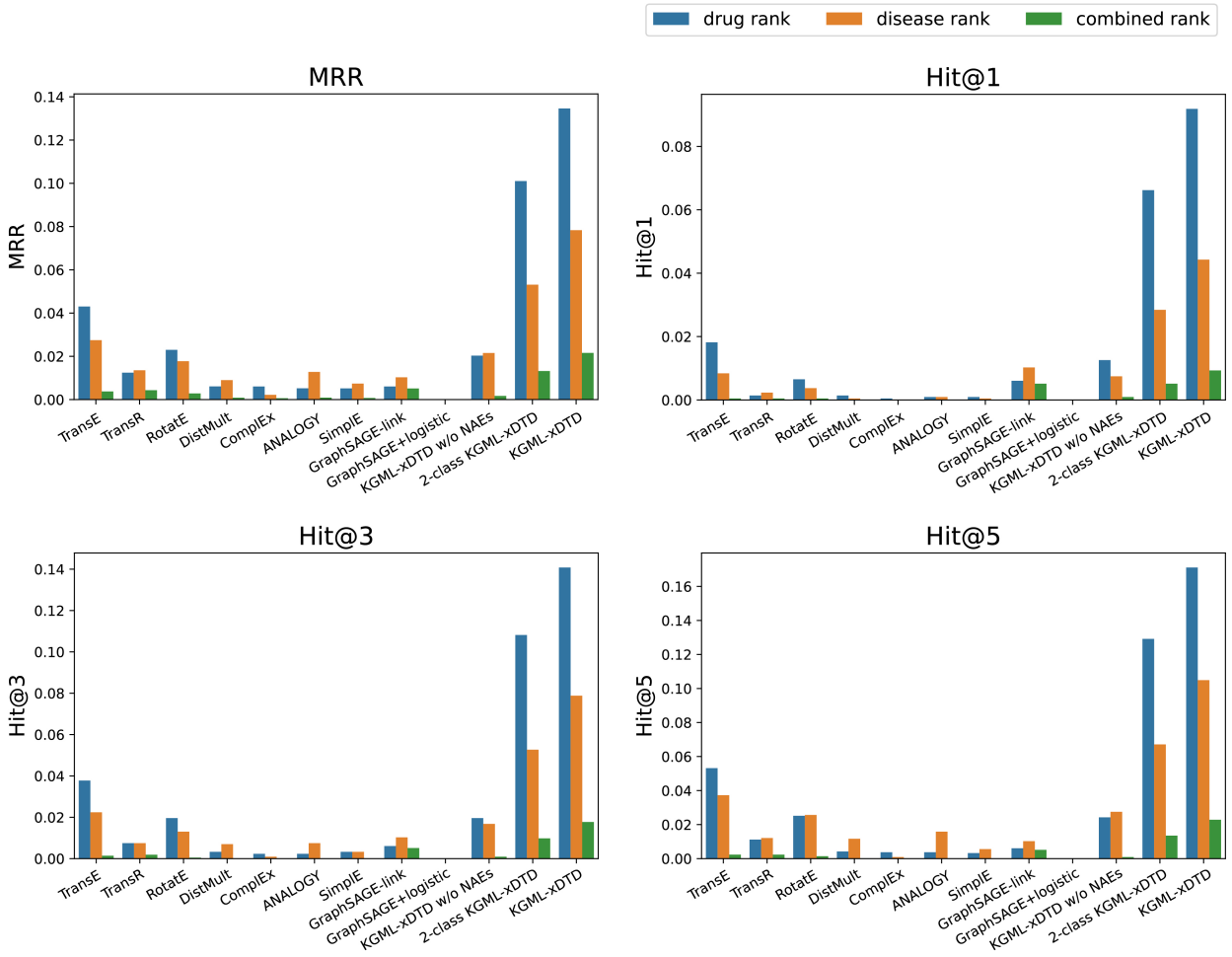

Figure 3: The performance comparison of Drug Repurposing Prediction (DRP) between `KGML-xDTD` and different baseline models (`GAT` and `GraphSAGE+SVM` are excluded due to computation time constraints) based on test dataset using three "complete" replacement methods (i.e., Drug-rank-based replacement", "Disease-rank-based replacement", and "Combined replacement") to generate non-true-positive drug-disease candidates for each true positive drug-disease pair for MRR and Hit@K calculation (see more details in Sec. 3.1.3). The legend "drug rank", "disease rank" and "combined rank" respectively correspond to the methods of "Drug-rank-based replacement", "Disease-rank-based replacement", and "Combined replacement" described in Sec. 3.1.3.

3.3 MOA prediction evaluation

For MOA prediction, we evaluate how well the `KGML-xDTD` can identify the DrugMechDB-matched BKG-based MOA paths (described in Sec. 3.1.4) from a large number of possible paths in the customized BKG by utilizing ranking-based metrics (e.g., MPR, MRR, and Hit@K) and two specific case studies.

There are few machine learning models designed for the task of identifying biologically meaningful paths from biomedical knowledge graphs for explaining drug repurposing. To our best knowledge, although three models (e.g.,

UKGE [Sosa et al., 2020], GrEDeL [Sang et al., 2019], Polo [Liu et al., 2021], all mentioned in Sec. 1) were proposed and can be used for this goal, they all have certain constraints and cannot be used as baseline models for comparison. The UKGE model cannot be applied to BKGs without weighted edge information (e.g., frequency of relation appeared in literature). The authors of the GrEDeL model don't provide the code to implement this model. The Polo model cannot be trained within a reasonable time (e.g., within two weeks) on a massive and complex BKG (e.g., RTX-KG2) due to its dependence on a computationally inefficient method "DWPC" [Himmelstein et al., 2017]. Therefore, we choose the `MultiHop` reinforcement learning model [Lin et al., 2018] as a baseline model since it uses the LSTM model framework as the GrEDel model and allows using a self-defined reward shaping strategy in its reward function as what we do in the `KGML-xDTD` model (i.e., we can use the same reward strategy described in Sec. 2.2.3). Furthermore, we also compare with an ablated version of `KGML-xDTD` (i.e., `KGML-xDTD w/o DP` which does not take advantage of the demonstration paths by setting α_p and α_m in Function 9 as 0) as another baseline model to show the importance of demonstration paths.

Table 3: The performance comparison of Mechanism of Action (MOA) Prediction between `KGML-xDTD` and different baseline models (e.g., `MultiHop` and `KGML-xDTD w/o DP`) based on test dataset (described in Sec. 3.1.1). The metrics in this table are calculated using path scores and all non-DrugMechDB-matched 3-hop paths between drug and disease as "negative" paths for each true positive drug-disease pair (see more details in Sec. 3.1.4).

Model	**MPR**	**MRR**	**Hit@1**	**Hit@10**	**Hit@50**	**Hit@100**	**Hit@500**
MultiHop	61.400%	0.027	0.017	0.042	0.067	0.118	0.345
KGML-xDTD w/o DP	72.965%	0.015	0.008	0.017	0.067	0.160	0.403
KGML-xDTD (ours)	**94.696%**	**0.109**	**0.059**	**0.193**	**0.496**	**0.613**	**0.849**

We show the comparison results between the `KGML-xDTD` model framework and different baseline models for MOA prediction evaluation in Table 3. Although all the models receive the same terminal rewards from the environment, that is given by the drug repurposing prediction module of `KGML-xDTD`, the MOA prediction module of `KGML-xDTD` achieves significantly better performance in identifying DrugMechDB-matched BKG-based MOA paths than the other two baselines across all ranking-based metrics. Comparison between the `KGML-xDTD` with and without demonstration paths (i.e., `KGML-xDTD w/o DP`) further illustrates the great effectiveness of using demonstration paths to guide the path-finding process. Due to the massive searching space and sparse rewards, the RL agent often fails to find reasonable paths out of many possible choices, while our model `KGML-xDTD`, with the intermediate guidance provided by the demonstration path, is able to identify the biologically reasonable choices with a much higher probability. Moreover, regardless of the `KGML-xDTD` with and without demonstration paths, we can see that they both are superior to the `MultiHop`, which illustrates that the actor-critic model structure might be more effective than LSTM for this task.

To further evaluate the performance of `KGML-xDTD` model framework in identifying biologically relevant MOA paths for drug repurposing, we present two different case studies to explore the potential repurposed drugs and their potential mechanism for two rare genetic diseases: Hemophilia B and Huntington's disease.

3.3.1 Case 1: Hemophilia B

Hemophilia B, also known as factor IX deficiency or Christmas disease, is a rare genetic disorder that results in prolonged bleeding in patients. It is caused by mutations in the factor IX (F9) gene, which is located on the X chromosome. Table 4 displays the top 10 drugs/treatments predicted by the `KGML-xDTD` model framework, including both those that

Table 4: Top 10 predicted drugs/treatments for hemophilia B (note that the drugs highlighted in red color are used in the training set).

Drug/Treatment	**Prob.**	**Publications**
Eptacog Alfa (rFVIIa)	0.833	[Croom and McCormack, 2008, Minno, 2015]
Nonacog Alfa (rFIX)	0.803	[Rendo et al., 2015]
Viral Vector	0.780	[Driessche et al., 2001]
Factor VIIa	0.748	[Roberts et al., 2004, Croom and McCormack, 2008]
Recombinant FVIIa (rFVIIa)	0.724	[Roberts et al., 2004, Croom and McCormack, 2008]
Thrombin	0.709	[Negrier et al., 2019]
Factor IX	0.708	[Goodeve, 2015]
Epicriptine	0.702	
Hyperbaric Oxygen	0.660	
Triamcinolone	0.649	

are used in the training set (highlighted in red) and those that are not. Besides those known drugs/treatments used in the training set, the majority of the remaining seven drugs/treatments on the list are supported by published research and have the potential to treat hemophilia B. For example, the activated human-derived coagulation factor VII (i.e., Factor VIIa) or the recombinant activated factor VII (i.e., rFVIIa) is one of the proteins that can cause blood clots as an important part of the blood coagulation regulatory network (as shown in the Figure 4). This protein is used as an effective inhibitor in the treatment of patients with hemophilia B [Roberts et al., 2004, Croom and McCormack, 2008]. Thrombin is a key enzyme in the maintenance of normal hemostatic function. It has been reported that using thrombin as a therapeutic strategy can help prevent bleeding in patients with hemophilia [Negrier et al., 2019]. The use of recombinant factor IX therapy is a recommended treatment option for individuals with hemophilia B [Goodeve, 2015]. Some examples of recombinant factor IX products include BeneFIX, Rixubis, Ixinity, Alprolix Idelvion, and Rebinyn.

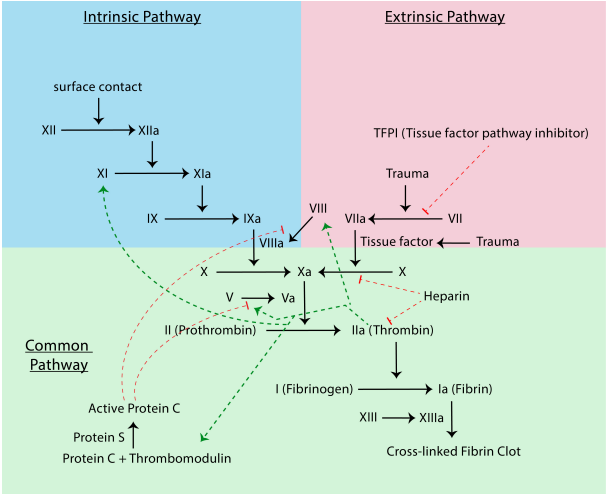

Figure 4: Blood Coagulation Regulatory Network with arrows for molecular reactions (black), positive feedback (green), and negative feedback (red).

To further assess the biological explanations of the predicted 3-hop BKG-based MOA paths for the treatment of hemophilia B, we have used the curated DrugMechDB-based MOA paths, which are not used in the model training process. DrugMechDB contains relevant MOA paths of hemophilia B treatment only for Eptacog Alfa and Nonacog Alfa. We show the comparisons between the subgraphs with the top 10 predicted 3-hop BKG-based paths and the curated DugMechDB-based MOA paths for Eptacog Alfa and Nonacog Alfa in Figure 5. The corresponding biological entities between the predicted paths and the curated DrugMechDB-based paths are highlighted in red color. Although the predicted paths can't exactly match the DrugMechDB-based MOA paths due to the limited path length and some missing semantic relationships in the customized biomedical knowledge graph, key biological entities (such as Coagulation Factor VII, Coagulation Factor X, and Coagulation Factor IX) that are important for the treatment of hemophilia B are present in the subgraphs of the top 10 predicted paths. As shown in Figure 4, the treatment of hemophilia B involves a complex molecular network of blood coagulation, and many of the coagulation factors (such as factor VII, factor VIIa, factor III, factor II, factor VIII, factor IX, and factor X) present in the subgraphs of the top 10 predicted paths are also part of this molecular network. In Appx. D, we also utilize the `KGML-xDTD` model framework to predict the top 10 3-hop BKG-based paths, which can serve as biological explanations of the predicted "treats" relationship between Factor VIIa and Hemophilia B (shown in Table 4). This particular drug/treatment - disease pair is not included in the training set and thus can be used to indicate how `KGML-xDTD`'s MOA path predictions can contribute to the explanation of the predicted drug repurposing results. The predicted paths show molecular details akin to those in Figure 4 for treating hemophilia B. As a result, the predicted paths by `KGML-xDTD` model framework can help identify key molecules in the real drug action regulatory network, thereby aiding in explaining drug repurposing to some extent.

3.3.2 Case 2: Huntington's disease

Huntington's disease (HD) is a rare neurogenetic disorder that typically occurs in midlife with symptoms of depression, uncontrolled movements, and cognitive decline. While there is currently no drug/treatment that can alter the course of HD, some drugs/treatments can be useful for the treatment of its symptoms in abnormal movements (e.g., chorea) and psychiatric phenotypes. We show ten drugs/treatments with the highest predicted probability by the `KGML-xDTD` model framework after manual processing in Table 5. This processing involves excluding the chemotherapeutic drugs

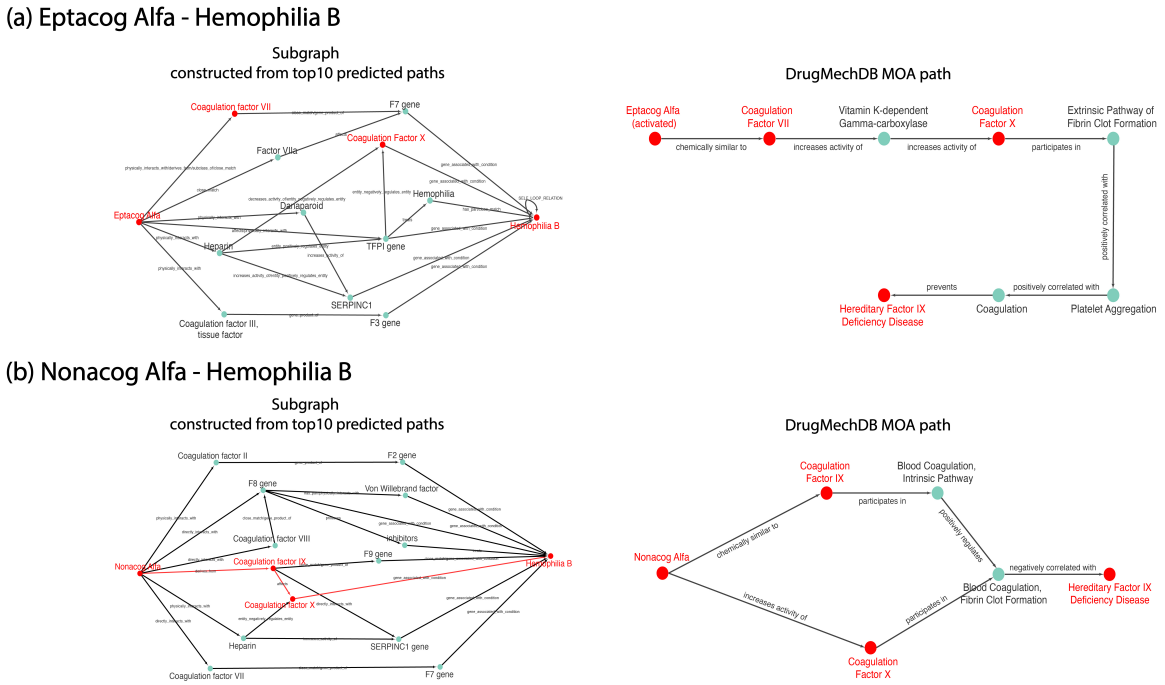

Figure 5: Comparison between the top 10 predicted 3-hop paths (integrated into a subgraph for better visualization) and the curated DugMechDB-based MOA paths for Eptacog Alfa (a) and Nonacog Alfa (b). On the left is a network representation of the top 10 `KGML-xDTD` predicted paths. On the right are network representations of the human-curated DrugMechDB mechanism of action paths. In both, the corresponding biological entities/vertices between the top 10 predicted 3-hop paths and MOA paths are highlighted in red color. The red edges indicate a top-10-`KGML-xDTD` predicted path where all entities show up in the DugMechDB-based MOA paths.

from the predicted drug candidate list due to their potential risk of cytotoxicity to normal cells (which could lead to false positives for drug repurposing of non-cancer diseases [Sourimant et al., 2021, Gysi et al., 2021]), and only presenting the top 5 results in the training set, and top 5 from the test or validation set. From this table, it can be observed that many of the top-ranked predicted drugs have been supported by publications as potential treatments for the symptoms of HD. Since there is currently no effective treatment for HD, DrugMechDB does not have a corresponding MOA path for comparison. To analyze the predicted paths by the `KGML-xDTD` model framework for the predicted non-chemotherapeutic drugs/treatments that are not included in the training set (shown in black in Table 5), we present their top 10 predicted paths (integrated into different subgraphs) in Figure 6. From these predicted paths, we can see that most of them are biologically relevant. For example, the subfigure (a) of Figure 6 shows that Risperidone is predicted to be useful for the treatment of HD by decreasing the activity of the genes associated with the 5-Hydroxytryptamine receptor (e.g., HTR1A, HTR2A, HTR2C, HTR7) and dopamine receptor (e.g., DRD1, DRD2, DRD3) which have been proven to be involved in the pathogenesis of depressive disorders [Yohn et al., 2017, Delva and Stanwood, 2021]. The presence of depressive symptoms is a significant characteristic of HD [Coppen and Roos, 2017]. Entinostat is predicted to have the potential to alleviate the symptoms of HD by inhibiting the functions of histone deacetylase genes such as HDAC1, HDAC2, HDAC6 (see subfigure (b) of Figure 6), and one of the predicted 3-hop MOA paths ("Entinostat" $\rightarrow$ "decreases activity of" $\rightarrow$ "HDAC1 gene" $\rightarrow$ "interacts with" $\rightarrow$ "Histone H4" $\rightarrow$ "gene associated with condition" $\rightarrow$ "Huntington's disease") is supported by the previous research [Shukla and Tekwani, 2020, Yu et al., 2009]. Primaquine is predicted to act on the NQ02 gene and the IKBKG gene to potentially play a therapeutic role in neurodegenerative disease, reported in [Voronin et al., 2021, Singh and Singh, 2020]. According to the predicted MOA paths, Isradipine may have a potential therapeutic effect for HD by mainly regulating the genes of the Calcium Voltage-Gated Channel, including CACNA1S, CACNA1D, CACNA1C, CACNB2, CACNA2D2. These genes may be associated with the symptoms of HD, such as chorea, depression, and dementia [Yagami et al., 2012]. Lastly, Amifampridine is predicted to regulate the genes of the Potassium Voltage-Gated Channel which are potentially associated with HD [Noh et al., 2019]. All these examples indicate that the predicted BKG-based MOA paths can explain the mechanism of repurposed drugs to some extent.

Table 5: Top 5 predicted drugs/treatments used in the training set (highlighted in red color) and the top 5 non-chemotherapeutic predicted drugs/treatments that are not in the training set for Huntington's Disease.

Drug/Treatment	Prob.	Publications
Pimozide	0.939	[Arena et al., 1980, Videnovic, 2013]
Therapeutic Agent	0.939	
Olanzapine	0.938	[Paleacu et al., 2002, Squitieri et al., 2001]
Riluzole	0.935	[Group, 2003]
Antipsychotic Agent	0.932	[Unti et al., 2017]
Risperidone	0.893	[Coppen and Roos, 2017, Duff et al., 2008]
Entinostat	0.888	[Shukla and Tekwani, 2020]
Primaquine	0.887	
Isradipine	0.884	[Miranda et al., 2019]
Amifampridine	0.882	

4 Discussion

In this work, we propose `KGML-xDTD`, a two-module, knowledge graph-based machine learning framework that not only predicts the treatment probabilities between drugs/compounds and diseases but also provides biological explanations for these predictions through the predicted paths in a massive biomedical knowledge graph with comprehensive biomedical data sources as potential mechanisms of action. This framework can assist medical researchers in quickly identifying the potential drug/compound-disease pairs that might have a treatment relationship, which can accelerate the process of drug discovery for emerging diseases. Additionally, by leveraging the KG-based paths predicted by the framework, medical professionals (e.g., doctors and licensed medical practitioners) can straightforwardly assess the accuracy of the predictions via the predicted MOAs, which can help to reduce false positives that may be produced by the "black-box" operation of traditional machine learning models.

Although previous research [Himmelstein et al., 2017, Zhang et al., 2021, Ioannidis et al., 2020] has applied a variety of models to the task of drug repurposing using biomedical knowledge graphs (BKGs), these approaches are implemented in the small-scale BKGs and many do not scale to larger graphs. As biotechnology advances and the volume of data in biomedical databases increases, BKGs are becoming larger and more complex. In our comparison with state-of-the-art KG-based models for drug repurposing, we find that the `KGML-xDTD` model had higher accuracy with lower false positives when applied to a massive and complex biomedical knowledge graph *RTX-KG2c*. By evaluating the predicted paths with DrugMechDB and two case studies, we show that the model can capture some key biological entities involved in real drug action regulatory networks.

It is widely acknowledged that drug repurposing is one of the most challenging problems in biomedicine, and current AI techniques are still in the early stages of addressing it. Many other AI models, such as those based on chemical structure, drug-target interactions, and drug perturbations of gene expression, are developed for solving this goal. They may offer more accurate predictions but also have limitations in terms of cost and the availability of samples for specific diseases. Biomedical knowledge graph (BKG)-based machine learning models, such as the `KGML-xDTD` model, offer a cost- and time-efficient alternative due to the large volume of biomedical knowledge stored in public databases and publications. The `KGML-xDTD` model framework is not intended to replace or beat these models, but rather provides a complementary approach that leverages emerging knowledge graphs for drug repurposing.

Future work to further enhance the `KGML-xDTD` model framework might include extending the predicted paths for more specific explanations, and considering the negative drug-disease pairs so that the model can explain why certain drugs are harmful to diseases.

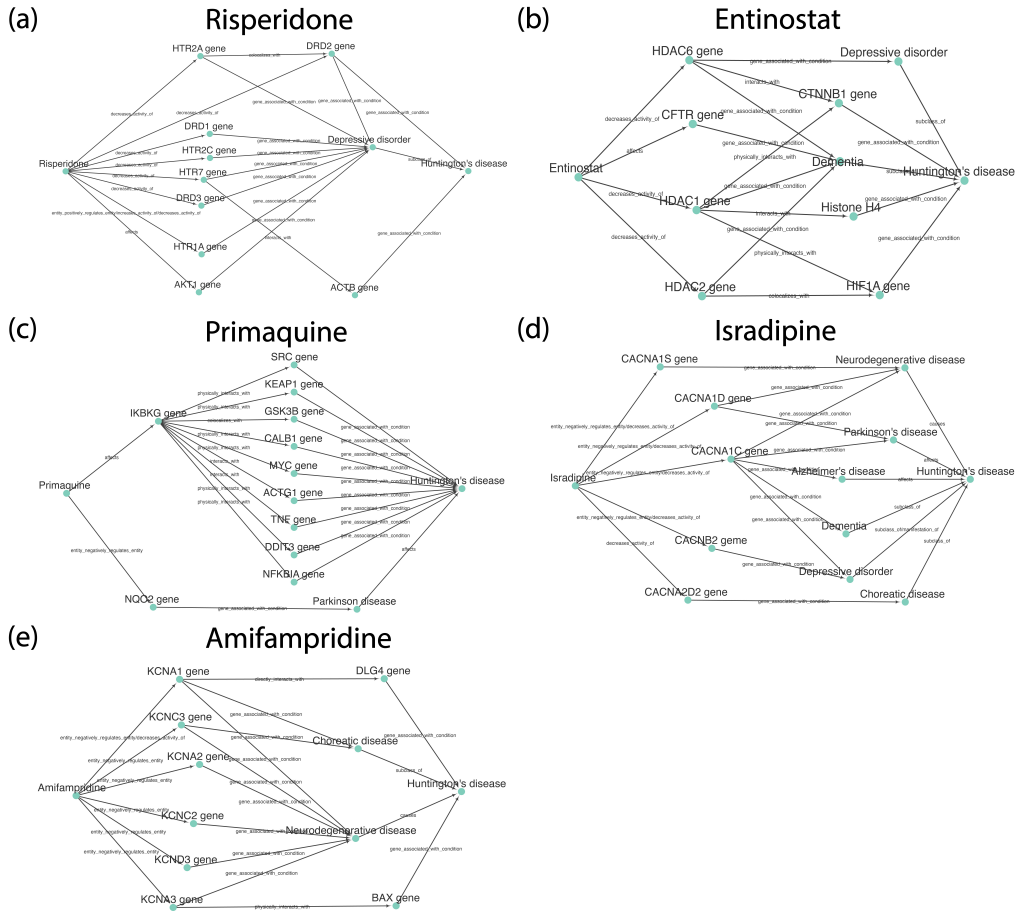

Figure 6: Top 10 predicted 3-hop BKG-based MOA paths (integrated into subgraphs) for top 5 non-chemotherapeutic predicted drugs/treatments that are not included in the training set for Huntington's disease.

References

Nurken Berdigaliyev and Mohamad Aljofan. An overview of drug discovery and development. *Future Medicinal Chemistry*, 12(10):939–947, 2020. ISSN 1756-8919. doi:10.4155/fmc-2019-0307.

M T Miller. Thalidomide embryopathy: a model for the study of congenital incomitant horizontal strabismus. *Transactions of the American Ophthalmological Society*, 89:623–74, 1991. ISSN 0065-9533.

H M W Verheul, D Panigrahy, J Yuan, and R J D'Amato. Combination oral antiangiogenic therapy with thalidomide and sulindac inhibits tumour growth in rabbits. *British Journal of Cancer*, 79(1):114–118, 1999. ISSN 0007-0920. doi:10.1038/sj.bjc.6690020.

Seema Singhal, Jayesh Mehta, Raman Desikan, Dan Ayers, Paula Roberson, Paul Eddlemon, Nikhil Munshi, Elias Anaissie, Carla Wilson, Madhav Dhodapkar, Jerome Zeldis, David Siegel, John Crowley, and Bart Barlogie. Antitumor activity of thalidomide in refractory multiple myeloma. *The New England Journal of Medicine*, 341(21): 1565–1571, 1999. ISSN 0028-4793. doi:10.1056/nejm199911183412102.

Visvaldas Kairys, Lina Baranauskiene, Migle Kazlauskiene, Daumantas Matulis, and Egidijus Kazlauskas. Binding affinity in drug design: experimental and computational techniques. *Expert Opinion on Drug Discovery*, 14(8): 755–768, 2019. ISSN 1746-0441. doi:10.1080/17460441.2019.1623202.

Nathalie Aulner, Anne Danckaert, JongEun Ihm, David Shum, and Spencer L. Shorte. Next-Generation Phenotypic Screening in Early Drug Discovery for Infectious Diseases. *Trends in Parasitology*, 35(7):559–570, 2019. ISSN 1471-4922. doi:10.1016/j.pt.2019.05.004.

Carmen-Maria Rusz, Bianca-Eugenia Ősz, George Jîtcă, Amalia Miklos, Mădălina-Georgiana Bătrînu, and Silvia Imre. Off-Label Medication: From a Simple Concept to Complex Practical Aspects. *International Journal of Environmental Research and Public Health*, 18(19):10447, 2021. ISSN 1661-7827. doi:10.3390/ijerph181910447.

S. Joshua Swamidass. Mining small-molecule screens to repurpose drugs. *Briefings in Bioinformatics*, 12(4):327–335, 2011. ISSN 1467-5463. doi:10.1093/bib/bbr028.

Philippe Sanseau, Pankaj Agarwal, Michael R Barnes, Tomi Pastinen, J Brent Richards, Lon R Cardon, and Vincent Mooser. Use of genome-wide association studies for drug repositioning. *Nature Biotechnology*, 30(4):317–320, 2012. ISSN 1087-0156. doi:10.1038/nbt.2151.

Stephen Bonner, Ian P Barrett, Cheng Ye, Rowan Swiers, Ola Engkvist, Andreas Bender, Charles Tapley Hoyt, and William L Hamilton. A review of biomedical datasets relating to drug discovery: a knowledge graph perspective. *Briefings in Bioinformatics*, sep 2022. doi:10.1093/bib/bbac404.

David S. Wishart, Yannick D. Feunang, An C. Guo, Elvis J. Lo, Ana Marcu, Jason R. Grant, Tanvir Sajed, Daniel Johnson, Carin Li, Zinat Sayeeda, Nazanin Assempour, Ithayavani Iynkkaran, Yifeng Liu, Adam Maciejewski, Nicola Gale, Alex Wilson, Lucy Chin, Ryan Cummings, Diana Le, Allison Pon, Craig Knox, and Michael Wilson. Drugbank 5.0: a major update to the drugbank database for 2018. *Nucleic Acids Research*, 46(D1):gkx1037–, 2017a. ISSN 1362-4962. doi:10.1093/nar/gkx1037.

Anna Gaulton, Louisa J. Bellis, A. Patricia Bento, Jon Chambers, Mark Davies, Anne Hersey, Yvonne Light, Shaun McGlinchey, David Michalovich, Bissan Al-Lazikani, and John P. Overington. Chembl: a large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40(D1):D1100–D1107, 2012. ISSN 0305-1048. doi:10.1093/nar/gkr777.

David S. Wishart, Yannick Djoumbou Feunang, Ana Marcu, An Chi Guo, Kevin Liang, Rosa Vázquez-Fresno, Tanvir Sajed, Daniel Johnson, Carin Li, Naama Karu, Zinat Sayeeda, Elvis Lo, Nazanin Assempour, Mark Berjanskii, Sandeep Singhal, David Arndt, Yonjie Liang, Hasan Badran, Jason Grant, Arnau Serra-Cayuela, Yifeng Liu, Rupa Mandal, Vanessa Neveu, Allison Pon, Craig Knox, Michael Wilson, Claudine Manach, and Augustin Scalbert. Hmdb 4.0: the human metabolome database for 2018. *Nucleic Acids Research*, 46(D1):gkx1089–, 2017b. ISSN 1362-4962. doi:10.1093/nar/gkx1089.

Samantha Kanza and Jeremy Graham Frey. Semantic technologies in drug discovery. In Olaf Wolkenhauer, editor, *Systems Medicine*, pages 129–144. Academic Press, Oxford, 2021. ISBN 978-0-12-816078-7. doi:https://doi.org/10.1016/B978-0-12-801238-3.11520-X.

Daniel Scott Himmelstein, Antoine Lizee, Christine Hessler, Leo Brueggeman, Sabrina L Chen, Dexter Hadley, Ari Green, Pouya Khankhanian, and Sergio E Baranzini. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife*, 6:e26726, 2017. doi:10.7554/elife.26726.

Brian Walsh, Sameh K Mohamed, and Vít Nováček. BioKG: A Knowledge Graph for Relational Learning On Biological Data. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 3173–3180, 2020. doi:10.1145/3340531.3412776.

Chang Su, Yu Hou, Winston Guo, Fayzan Chaudhry, Gregory Ghahramani, Haotan Zhang, and Fei Wang. Cbkh: The cornell biomedical knowledge hub. *medRxiv*, 2021. doi:10.1101/2021.03.12.21253461.

Bethany Percha and Russ B Altman. A global network of biomedical relationships derived from text. *Bioinformatics*, 34(15):2614–2624, 2018. ISSN 1367-4803. doi:10.1093/bioinformatics/bty114.

Rui Zhang, Dimitar Hristovski, Dalton Schutte, Andrej Kastrin, Marcelo Fiszman, and Halil Kilicoglu. Drug repurposing for covid-19 via knowledge graph completion. *Journal of Biomedical Informatics*, 115:103696, 2021. ISSN 1532-0464. doi:10.1016/j.jbi.2021.103696.

Qingyun Wang, Manling Li, Xuan Wang, Nikolaus Parulian, Guangxing Han, Jiawei Ma, Jingxuan Tu, Ying Lin, Ranran Haoran Zhang, Weili Liu, Aabhas Chauhan, Yingjun Guan, Bangzheng Li, Ruisong Li, Xiangchen Song, Yi Fung, Heng Ji, Jiawei Han, Shih-Fu Chang, James Pustejovsky, Jasmine Rah, David Liem, Ahmed ELsayed, Martha Palmer, Clare Voss, Cynthia Schneider, and Boyan Onyshkevych. COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Demonstrations*, pages 66–77, 2021a. doi:10.18653/v1/2021.naacl-demos.8.

Nan Li, Zhihao Yang, Ling Luo, Lei Wang, Yin Zhang, Hongfei Lin, and Jian Wang. KGHC: a knowledge graph for hepatocellular carcinoma. *BMC Medical Informatics and Decision Making*, 20(Suppl 3):135, 2020. doi:10.1186/s12911-020-1112-5.

Alberto Santos, Ana R. Colaço, Annelaura B. Nielsen, Lili Niu, Maximilian Strauss, Philipp E. Geyer, Fabian Coscia, Nicolai J. Wewer Albrechtsen, Filip Mundt, Lars Juhl Jensen, and Matthias Mann. A knowledge graph to interpret clinical proteomics data. *Nature Biotechnology*, 40(5):692–702, 2022. ISSN 1087-0156. doi:10.1038/s41587-021-01145-6.

E. C. Wood, Amy K. Glen, Lindsey G. Kvarfordt, Finn Womack, Liliana Acevedo, Timothy S. Yoon, Chunyu Ma, Veronica Flores, Meghamala Sinha, Yodsawalai Chodpathumwan, Arash Termehchy, Jared C. Roach, Luis Mendoza, Andrew S. Hoffman, Eric W. Deutsch, David Koslicki, and Stephen A. Ramsey. RTX-KG2: a system for building a semantically standardized knowledge graph for translational biomedicine. *BMC Bioinformatics*, 23(1):400, 2022. doi:10.1186/s12859-022-04932-3.

Vassilis N. Ioannidis, Da Zheng, and George Karypis. Few-shot link prediction via graph neural networks for covid-19 drug-repurposing. *CoRR*, abs/2007.10261, 2020.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. Translating Embeddings for Modeling Multi-relational Data. In *Neural Information Processing Systems (NIPS)*, pages 1–9, South Lake Tahoe, United States, December 2013.

Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang. RotatE: Knowledge Graph Embedding by Relational Rotation in Complex Space. *arXiv*, 2019.

Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding Entities and Relations for Learning and Inference in Knowledge Bases. *arXiv*, 2014.

Tho Trouillon, Johannes Welbl, Sebastian Riedel, ric Gaussier, and Guillaume Bouchard. Complex Embeddings for Simple Link Prediction. *arXiv*, 2016. doi:10.48550/arXiv.1606.06357.

Bo Wang, Tao Shen, Guodong Long, Tianyi Zhou, Ying Wang, and Yi Chang. Structure-augmented text representation learning for efficient knowledge graph completion. *Proceedings of the Web Conference 2021*, pages 1737–1748, 2021b. doi:10.1145/3442381.3450043.

Daniel N Sosa, Alexander Derry, Margaret Guo, Eric Wei, Connor Brinton, and Russ B Altman. A Literature-Based Knowledge Graph Embedding Method for Identifying Drug Repurposing Opportunities in Rare Diseases. *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing*, 25:463–474, 2020.

Xuelu Chen, Muhao Chen, Weijia Shi, Yizhou Sun, and Carlo Zaniolo. Embedding uncertain knowledge graphs. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33:3363–3370, 2019. ISSN 2159-5399. doi:10.1609/aaai.v33i01.33013363.

Shengtian Sang, Zhihao Yang, Xiaoxia Liu, Lei Wang, Hongfei Lin, Jian Wang, and Michel Dumontier. GrEDeL: A Knowledge Graph Embedding Based Method for Drug Discovery From Biomedical Literatures. *IEEE Access*, 7: 8404–8415, 2019. ISSN 2169-3536. doi:10.1109/access.2018.2886311.

Halil Kilicoglu, Graciela Rosemblat, Marcelo Fiszman, and Dongwook Shin. Broad-coverage biomedical relation extraction with SemRep. *BMC Bioinformatics*, 21(1), 2020. doi:10.1186/s12859-020-3517-7.

Yuxi Li. Reinforcement Learning Applications. *arXiv*, 2019. doi:10.48550/arxiv.1908.06973.

Ling Chen, Jun Cui, Xing Tang, Yuntao Qian, Yansheng Li, and Yongjun Zhang. RLPath: a knowledge graph link prediction method using reinforcement learning based attentive relation path searching and representation learning. *Applied Intelligence*, 52(4):4715–4726, 2022. ISSN 0924-669X. doi:10.1007/s10489-021-02672-0.

Yiwei Sun, Suhang Wang, Xianfeng Tang, Tsung-Yu Hsieh, and Vasant Honavar. Adversarial Attacks on Graph Neural Networks via Node Injections: A Hierarchical Reinforcement Learning Approach. *Proceedings of The Web Conference 2020*, pages 673–683, 2020. doi:10.1145/3366423.3380149.

Xingchen Zhou, Peng Wang, Qiqing Luo, and Zhe Pan. Multi-hop Knowledge Graph Reasoning Based on Hyperbolic Knowledge Graph Embedding and Reinforcement Learning. *The 10th International Joint Conference on Knowledge Graphs*, pages 1–9, 2021. doi:10.1145/3502223.3502224.

Yushan Liu, Marcel Hildebrandt, Mitchell Joblin, Martin Ringsquandl, Rime Raissouni, and Volker Tresp. Neural multi-hop reasoning with logical rules on biomedical knowledge graphs. In Ruben Verborgh, Katja Hose, Heiko Paulheim, Pierre-Antoine Champin, Maria Maleshkova, Oscar Corcho, Petar Ristoski, and Mehwish Alam, editors, *The Semantic Web*, pages 375–391, Cham, 2021. Springer International Publishing. ISBN 978-3-030-77385-4.

Finn Womack, Jason McClelland, and David Koslicki. Leveraging distributed biomedical knowledge sources to discover novel uses for known drugs. *bioRxiv*, page 765305, 2019. doi:10.1101/765305.

William L Hamilton, Rex Ying, and Jure Leskovec. Inductive representation learning on large graphs. *arXiv*, 2017.

Kangzhi Zhao, Xiting Wang, Yuren Zhang, Li Zhao, Zheng Liu, Chunxiao Xing, and Xing Xie. Leveraging Demonstrations for Reinforcement Recommendation Reasoning over Knowledge Graphs. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 239–248, 2020. doi:10.1145/3397271.3401171.

Michael Mayers, Dylan Steinecke, and Andrew I. Su. Database of mechanism of action paths for selected drug-disease indications, March 2020. URL `https://doi.org/10.5281/zenodo.3708278`.

Kirill Degtyarenko, Paula de Matos, Marcus Ennis, Janna Hastings, Martin Zbinden, Alan McNaught, Rafael Alcántara, Michael Darsow, Mickaël Guedj, and Michael Ashburner. ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Research*, 36(Database issue):D344–D350, 2008. ISSN 0305-1048. doi:10.1093/nar/gkm791.

The Biomedical Data Translator Consortium. Toward A Universal Biomedical Data Translator. *Clinical and Translational Science*, 12(2):86–90, 2019. ISSN 1752-8054. doi:10.1111/cts.12591.

Translator Consortium. The biomedical data translator program: Conception, culture, and community. *Clinical and Translational Science*, 12(2):91–94, March 2019. doi:10.1111/cts.12592.

Deepak R Unni, Sierra AT Moxon, Michael Bada, Matthew Brush, Richard Bruskiewich, J Harry Caufield, Paul A Clemons, Vlado Dancik, Michel Dumontier, Karamarie Fecho, et al. Biolink model: A universal schema for knowledge graphs in clinical, biomedical, and translational science. *Clinical and Translational Science*, 2022. doi:10.1111/cts.13302.

Jiwen Xin, Cyrus Afrasiabi, Sebastien Lelong, Julee Adesara, Ginger Tsueng, Andrew I. Su, and Chunlei Wu. Cross-linking biothings apis through json-ld to facilitate knowledge exploration. *BMC Bioinformatics*, 19(1):30, 2018. doi:10.1186/s12859-018-2041-5.

Jiwen Xin, Adam Mark, Cyrus Afrasiabi, Ginger Tsueng, Moritz Juchler, Nikhil Gopal, Gregory S. Stupp, Timothy E. Putman, Benjamin J. Ainscough, Obi L. Griffith, Ali Torkamani, Patricia L. Whetzel, Christopher J. Mungall, Sean D. Mooney, Andrew I. Su, and Chunlei Wu. High-performance web services for querying gene and variant annotation. *Genome Biology*, 17(1):91, 2016. doi:10.1186/s13059-016-0953-9.

Halil Kilicoglu, Dongwook Shin, Marcelo Fiszman, Graciela Rosemblat, and Thomas C. Rindflesch. Semmeddb: a pubmed-scale repository of biomedical semantic predications. *Bioinformatics*, 28(23):3158–3160, 2012. ISSN 1367-4803. doi:10.1093/bioinformatics/bts591.

Steven H Brown, Peter L Elkin, S Trent Rosenbloom, Casey Husser, Brent A Bauer, Michael J Lincoln, John Carter, Mark Erlbaum, and Mark S Tuttle. Va national drug file reference terminology: a cross-institutional content coverage study. *Studies in health technology and informatics*, 107(Pt 1):477–81, 2004. ISSN 0926-9630.

Adam S. Brown and Chirag J. Patel. A standard database for drug repositioning. *Scientific Data*, 4(1):170029, 2017. doi:10.1038/sdata.2017.29.

R.L. Cilibrasi and P.M.B. Vitanyi. The Google Similarity Distance. *IEEE Transactions on Knowledge and Data Engineering*, 19(3):370–383, 2007. ISSN 1041-4347. doi:10.1109/tkde.2007.48.

Michael Mayers, Roger Tu, Dylan Steinecke, Tong Shu Li, Núria Queralt-Rosinach, and Andrew I Su. Design and application of a knowledge network for automatic prioritization of drug mechanisms. *Bioinformatics*, 38(10):btac205, 2022. ISSN 1367-4803. doi:10.1093/bioinformatics/btac205.

Yu Gu, Robert Tinn, Hao Cheng, Michael Lucas, Naoto Usuyama, Xiaodong Liu, Tristan Naumann, Jianfeng Gao, and Hoifung Poon. Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing. *ACM Transactions on Computing for Healthcare*, 3(1):1–23, 2022. ISSN 2691-1957. doi:10.1145/3458754.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality, 2013.

Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv*, 2015.

Richard S. Sutton. Learning to predict by the methods of temporal differences. *Machine Learning*, 3(1):9–44, 1988. ISSN 0885-6125. doi:10.1007/bf00115009.

Ronald J. Williams. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3-4):229–256, 1992. ISSN 0885-6125. doi:10.1023/a:1022672621406.

A PomerleauDean. Efficient training of artificial neural networks for autonomous navigation. *Neural Computation*, 1991.

Kanglin Hsieh, Yinyin Wang, Luyao Chen, Zhongming Zhao, Sean Savitz, Xiaoqian Jiang, Jing Tang, and Yejin Kim. Drug repurposing for COVID-19 using graph neural network and harmonizing multiple evidence. *arXiv*, 2020. doi:10.48550/arxiv.2009.10931.

Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. Learning entity and relation embeddings for knowledge graph completion. In *In Proceedings of AAAI'15*, 2015.

Hanxiao Liu, Yuexin Wu, and Yiming Yang. Analogical inference for multi-relational embeddings, 2017.

Seyed Mehran Kazemi and David Poole. Simple embedding for link prediction in knowledge graphs, 2018.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph Attention Networks. *International Conference on Learning Representations*, 2018. accepted as poster.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In Yee Whye Teh and D. Mike Titterington, editors, *AISTATS*, volume 9 of *JMLR Proceedings*, pages 249–256. JMLR.org, 2010.

Xi Victoria Lin, Richard Socher, and Caiming Xiong. Multi-Hop Knowledge Graph Reasoning with Reward Shaping. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3243–3253, 2018. doi:10.18653/v1/d18-1362.

Katherine F Croom and Paul L McCormack. Recombinant factor VIIa (eptacog alfa): a review of its use in congenital hemophilia with inhibitors, acquired hemophilia, and other congenital bleeding disorders. *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy*, 22(2):121–36, 2008. ISSN 1173-8804. doi:10.2165/00063030-200822020-00005.

Giovanni Di Minno. Eptacog alfa activated: a recombinant product to treat rare congenital bleeding disorders. *Blood Reviews*, 29:S26–S33, 2015. ISSN 0268-960X. doi:10.1016/s0268-960x(15)30005-9.

Pablo Rendo, Lynne Smith, Hsiao-Yu Lee, and Frank Shafer. Nonacog alfa: an analysis of safety data from six prospective clinical studies in different patient populations with haemophilia B treated with different therapeutic modalities. *Blood coagulation & fibrinolysis : an international journal in haemostasis and thrombosis*, 26(8):912–8, 2015. ISSN 0957-5235. doi:10.1097/mbc.0000000000000359.

Thierry Driessche, Desire Collen, and Marinee Chuah. Viral Vector-Mediated Gene Therapy for Hemophilia. *Current Gene Therapy*, 1(3):301–315, 2001. ISSN 1566-5232. doi:10.2174/1566523013348508.

Harold R. Roberts, Dougald M. Monroe, and Gilbert C. White. The use of recombinant factor VIIa in the treatment of bleeding disorders. *Blood*, 104(13):3858–3864, 2004. ISSN 0006-4971. doi:10.1182/blood-2004-06-2223.

Claude Negrier, Midori Shima, and Maureane Hoffman. The central role of thrombin in bleeding disorders. *Blood Reviews*, 38:100582, 2019. ISSN 0268-960X. doi:https://doi.org/10.1016/j.blre.2019.05.006.

A. C. Goodeve. Hemophilia B: molecular pathogenesis and mutation analysis. *Journal of Thrombosis and Haemostasis*, 13(7):1184–1195, 2015. ISSN 1538-7933. doi:10.1111/jth.12958.

Julien Sourimant, Megha Aggarwal, and Richard K Plemper. Progress and pitfalls of a year of drug repurposing screens against covid-19. *Current Opinion in Virology*, 49:183–193, 2021. ISSN 1879-6257. doi:https://doi.org/10.1016/j.coviro.2021.06.004.

Deisy Morselli Gysi, Í talo do Valle, Marinka Zitnik, Asher Ameli, Xiao Gan, Onur Varol, Susan Dina Ghiassian, J. J. Patten, Robert A. Davey, Joseph Loscalzo, and Albert-László Barabási. Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proceedings of the National Academy of Sciences of the United States of America*, 118(19):e2025581118, 2021. ISSN 0027-8424. doi:10.1073/pnas.2025581118.

Christine N. Yohn, Mark M. Gergues, and Benjamin Adam Samuels. The role of 5-HT receptors in depression. *Molecular Brain*, 10(1):28, 2017. doi:10.1186/s13041-017-0306-y.

Nella C Delva and Gregg D Stanwood. Dysregulation of brain dopamine systems in major depressive disorder. *Experimental Biology and Medicine*, 246(9):1084–1093, 2021. ISSN 1535-3702. doi:10.1177/1535370221991830.

Emma M. Coppen and Raymund A. C. Roos. Current Pharmacological Approaches to Reduce Chorea in Huntington's Disease. *Drugs*, 77(1):29–46, 2017. ISSN 0012-6667. doi:10.1007/s40265-016-0670-4.

Surabhi Shukla and Babu L. Tekwani. Histone Deacetylases Inhibitors in Neurodegenerative Diseases, Neuroprotection and Neuronal Differentiation. *Frontiers in Pharmacology*, 11:537, 2020. ISSN 1663-9812. doi:10.3389/fphar.2020.00537.

In Tag Yu, Jin-Yong Park, Sung Hyun Kim, Jeong-sik Lee, Yong-Seok Kim, and Hyeon Son. Valproic acid promotes neuronal differentiation by induction of proneural factors in association with H4 acetylation. *Neuropharmacology*, 56(2):473–480, 2009. ISSN 0028-3908. doi:10.1016/j.neuropharm.2008.09.019.

Mikhail V. Voronin, Ilya A. Kadnikov, Liana F. Zainullina, Ilya O. Logvinov, Ekaterina R. Verbovaya, Tatyana A. Antipova, Yulia V. Vakhitova, and Sergei B. Seredenin. Neuroprotective properties of quinone reductase 2 inhibitor m-11, a 2-mercaptobenzimidazole derivative. *International Journal of Molecular Sciences*, 22(23), 2021. ISSN 1422-0067. doi:10.3390/ijms222313061.

Shareen Singh and Thakur Gurjeet Singh. Role of Nuclear Factor Kappa B (NF-KB) Signalling in Neurodegenerative Diseases: An Mechanistic Approach. *Current Neuropharmacology*, 18(10):918–935, 2020. ISSN 1570-159X. doi:10.2174/1570159x18666200207120949.

T Yagami, H Kohma, and Y Yamamoto. L-type voltage-dependent calcium channels as therapeutic targets for neurodegenerative diseases. *Current medicinal chemistry*, 19(28):4816–27, 2012. ISSN 0929-8673. doi:10.2174/092986712803341430.

Wonjun Noh, Sojeong Pak, Geunho Choi, Sungchil Yang, and Sunggu Yang. Transient Potassium Channels: Therapeutic Targets for Brain Disorders. *Frontiers in Cellular Neuroscience*, 13:265, 2019. ISSN 1662-5102. doi:10.3389/fncel.2019.00265.

R Arena, A Iudice, P Virgili, P Moretti, and G Menchetti. Huntington's disease: clinical effects of a short-term treatment with pimozide. *Advances in biochemical psychopharmacology*, 24:573–5, 1980. ISSN 0065-2229.

Aleksandar Videnovic. Treatment of huntington disease. *Current treatment options in neurology*, 15(4):424–38, 2013. ISSN 1092-8480. doi:10.1007/s11940-013-0219-8.

D Paleacu, M Anca, and N Giladi. Olanzapine in Huntington's disease: Olanzapine in Huntington's disease. *Acta Neurologica Scandinavica*, 105(6):441–444, 2002. ISSN 0001-6314. doi:10.1034/j.1600-0404.2002.01197.x.

F Squitieri, M Cannella, A Porcellini, L Brusa, M Simonelli, and S Ruggieri. Short-term effects of olanzapine in Huntington disease. *Neuropsychiatry, neuropsychology, and behavioral neurology*, 14(1):69–72, 2001. ISSN 0894-878X.

Huntington Study Group. Dosage effects of riluzole in Huntington's disease: a multicenter placebo-controlled study. *Neurology*, 61(11):1551–6, 2003. ISSN 0028-3878. doi:10.1212/01.wnl.0000096019.71649.2b.

E. Unti, S. Mazzucchi, G. Palermo, U. Bonuccelli, and R. Ceravolo. Antipsychotic drugs in Huntington's disease. *Expert Review of Neurotherapeutics*, 17(3):227–237, 2017. ISSN 1473-7175. doi:10.1080/14737175.2016.1226134.

Kevin Duff, Leigh J. Beglinger, Margaret E. O'Rourke, Peg Nopoulos, Henry L. Paulson, and Jane S. Paulsen. Risperidone and the Treatment of Psychiatric, Motor, and Cognitive Symptoms in Huntington's Disease. *Annals of Clinical Psychiatry*, 2008. doi:10.1080/10401230701844802.

Artur S. Miranda, Pablo Leal Cardozo, Flavia R. Silva, Jessica M. de Souza, Isabella G. Olmo, Jader S. Cruz, Marcus Vinícius Gomez, Fabiola M. Ribeiro, and Luciene B. Vieira. Alterations of Calcium Channels in a Mouse Model of Huntington's Disease and Neuroprotection by Blockage of CaV1 Channels. *ASN NEURO*, 11: 1759091419856811, 2019. doi:10.1177/1759091419856811.

Qing Cong, Zhiyong Feng, Fang Li, Li Zhang, Guozheng Rao, and Cui Tao. Constructing biomedical knowledge graph based on semmeddb and linked open data. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1628–1631, 2018. doi:10.1109/BIBM.2018.8621568.

Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. URL `http://ilpubs.stanford.edu:8090/422/`. Previous number = SIDL-WP-1999-0120.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2015.

A Section S1. Biomedical Knowledge Graph RTX-KG2c Pre-processing

The dataset of *RTX-KG2c* (v2.7.3) [Wood et al., 2022] is accessed via https://github.com/RTXteam/RTX-KG2. We pre-process the raw data of *RTX-KG2c* by the following four principles:

1. Since we are mainly interested in the categories relevant to drug mechanisms of action (MOAs), we exclude the nodes with categories that are not expected to be useful for drug repurposing explanation (e.g., "GeographicaLocation", "Device", "InformationResource").
2. One of data sources used in the *RTX-KG2c* is the Semantic MEDLINE Database (SemMedDB) [Kilicoglu et al., 2012], one of the most widely used NLP-derived biomedical knowledge sources, that has been found recently [Cong et al., 2018] to contain some inconsistent relations due to the immature NLP techniques even though it contains many latest-found relations (e.g., the relations with Covid-19). To improve the quality of SemMedDB-based edges, we filter out some edges based on the following criteria:
   - Each remaining SemMedDB-based edge must be supported by at least 10 publications.
   - The PubMed-publication-based NGD score [Cilibrasi and Vitanyi, 2007] (defined in Equation 1 in the main text) of the two end nodes should be higher than 0.6.
3. The *RTX-KG2c* is a multigraph that allows multiple edges to connect between two nodes. These edges present different relations (a.k.a. predicates) between the nodes, and follow the predicate hierarchy [6] used in the Biolink model [Unni et al., 2022]. As a result, the *RTX-KG2c* contains some hierarchically redundant edges between two nodes. To simplify the process of path finding for the downstream MOA predictions, we only reserve the "leaf" predicates (i.e., the most specific predicates) in the Biolink semantic relation hierarchy if there are hierarchically associated edges between two nodes. For example, if there are edges between two nodes representing the predicates "affected by", "entity regulated by entity" and "entity positively regulated by entity", we would remove the "affected by" and "entity regulated by entity" edges because they are more general (i.e., "ancestor") predicates in the hierarchy compared to "entity positively regulated by entity". Removing those more general predicates does not affect the interpretability of the paths because the "leaf" predicates contain more precise semantic information.
4. To avoid the training information leakage, we exclude all existing edges that directly connect the potential drug nodes (nodes with categories "Drug" or "SmallMolecule") to potential disease nodes (nodes with categories "Disease", "PhenotypicFeature", "BehavioralFeature" or "DiseaseOrPhenotypicFeature") in the *RTX-KG2c*.

After applying the pre-processing steps to *RTX-KG2c* (v2.7.3) described above, the post-processed/customized biomedical knowledge graph (BKG) consists of 3,659,165 nodes with 33 distinct categories, and 18,291,237 edges with 74 distinct types.

B Section S2. Implementation Details of `KGML-xDTD` Model Framework

B.1 Drug Repurposing Prediction Module

In the drug repurposing prediction module, we utilize the source code [7] provided by Hamilton et al. [2017] to train unsupervised GraphSAGE embeddings with its "big" mean-based aggregator and two hidden layers of dimensions [256, 256]. For the random walk setting, we performed 10 walks each with length of 100. As for other parameters, the number of epochs is set to 10, the neighbor sampling size of each layer is 96, the learning rate is 0.001, the batch size is 256, and the maximum number of iterations per epoch is 10,000. Instead of using the default identity embeddings as the initial features, we use node attribute embeddings generated by the pre-trained PubMedBert model [Gu et al., 2022] with concatenation of the node's name and category, and then reduce their dimensions to 100 using Principal Component Analysis (PCA). The final output output embedding vector for each node has a dimension of 512. With these GraphSAGE embedding vectors, we concatenate the embedding vectors of each drug-disease pair in the training set as input features and use the *RandomForestClassifier* function of scikit-learn (v1.0) python package to train a Random Forest model. We run a grid search using the *GridSearchCV* function to determine the optimal parameter set for the Random Forest model from a range of depths {5, 10, 15, 20, 25, 30, 35} and number of trees {500, 1000, 1500, 2000}. The best parameter set for the Random Forest model uses the maximum depth $max_depth = 35$ and the number of trees $n_estimators = 2000$.

[6] Visualization of hierarchical predicates in the Biolink Model v2.1.0 `http://tree-viz-biolink.herokuapp.com/predicates/2.1.0`

[7] `https://github.com/williamleif/GraphSAGE`

B.2 Mechanism of Action (MOA) Prediction Module

In the mechanism of action prediction module, we design a reinforcement learning model following the model framework of Zhao et al. [2020] for drug repurposing purpose (see Sec. "Mechanism of Action (MOA) Prediction" in the main text). We set the state history length $K = 2$ and the maximum length of path $T = 3$. In order to make the customized biomedical knowledge graph (described in Sec. A above) enable training on a 48GB Quadro RTX 8000 GPU, we prune the action space of each node to a maximum size of 3,000 based on the PageRank score [Page et al., 1999] (calculated by the *pagerank* function of NetworkX (v2.7.1) python package). We set the dimensions of all lookup matrices used within actor, critic, meta-path discriminator, and path discriminator networks to 100. The dimensions of hidden layers of the actor network and critic network are both set to 512. We set the dimensions of hidden layers of the path discriminator with [512, 512] and used the dimension set of [512, 256] for the hidden layers of the meta-path discriminator. We use Xavier initialization [Glorot and Bengio, 2010] for the embeddings of all lookup matrices and the network layers. The weight of the path discriminator reward α_p is set to 0.006 while the meta-factor of the path discriminator reward $a_m = 0.012$. We respectively assign 0.99 and 0.005 to the decaying coefficient γ of $R_{e,T}$ and the weight α of entropy term. We optimize all networks using the Adam optimization algorithm [Kingma and Ba, 2015] with a learning rate of 0.0005. The mini-batch size is set to 32 with a path rollout of 35. The dropout rates of all subnetworks are set to 0.3 and the action dropout rate is set to 0.5.

C Section S3. Implementation Details of Demonstration Path Extraction

The demonstration paths are a set of multi-hop BKG-based paths that can be used to guide the agent in the reinforcement learning model of `KGML-xDTD` model framework to find biologically reasonable BKG-based MOA paths. It can be formulated as $P^k = \{p^k_{s,t} | v_s \in \mathcal{V}^{\text{drug}}; v_t \in \mathcal{V}^{\text{disease}}\}$ where $p^k_{s,t}$ is a multi-hop demonstration path with maximum path length k starting from a drug node v_s and ending at a disease node v_t. Given a potential drug node and a potential disease node (defined in Sec. A above), the number of paths in the customized biomedical knowledge graph between them grows exponentially as the value of k increases. Therefore, we set $k = 3$ to guarantee that the agent can find biologically meaningful predicted MOA paths in a reasonable amount of time. To extract reasonable demonstration paths from the customized biomedical knowledge graph, we use the known drug-target interactions collected from two curated biomedical data sources (e.g., DrugBank (v5.1) and Molecular Data Provider (v1.2)[8]), as well as the PubMed-publication-based Normalized Google Distance (NGD) (see Equation 1 in the main text). A demonstration path extracted from the customized biomedical knowledge graph must satisfy the following two requirements:

1. The edge connecting the drug node to the first intermediate node in a demonstration path must be supported by DrugBank or Molecular Data Provider as a known drug-target interaction, and must also have a Normalized Google Distance (NGD) score of 0.6 or lower.
2. The edge connecting the second intermediate node to the disease node in a demonstration path should have a Normalized Google Distance (NGD) score of 0.6 or lower.

Only the true positive pairs (see Table 1 in the main text) collected from four human-curated and NLP-derived training datasets (described in Sec. "Data Sources for Model Training" in the main text) are used to extract demonstration paths. We also filter out any true positive pairs that are not reachable from a drug node to its corresponding disease node within maximum of 3 hops in the customized biomedical knowledge graph. Out of these 21,437 true positive pairs, 8,495 are able to find at least one demonstration path that meet the requirement. We finally find 396,705 demonstration paths for 8,495 true positive drug-disease pairs which are used for reinforcement learning training.

D Section S4. Implementation Details of Baseline Models

We use the OpenKE library [9] to implement the models `TransE`, `TransR`, `RotatE`, `DistMult`, `ComplEx`, `ANALOGY`, and `SimpLE` with their default parameter settings (we adjusted the hyperparameters for some models to ensure we can finish the training on GPUs for a reasonable time). Table 6 shows the detailed hyperparameter setting of these baseline model. For other baseline models, we use the PyTorch Geometric [10] framework to implement the `GAT` and `GraphSAGE-link` model and use the same GraphSAGE embeddings (mentioned above) with scikit-learn (v1.0) python package to implement the `GraphSAGE+logistic`, `GraphSAGE+SVM` and `2-class GraphSAGE+RF` models with the grid-search-based optimal parameter settings.

[8] `https://github.com/NCATSTranslator/Translator-All/wiki/Molecular-Data-Provider`

[9] `https://github.com/thunlp/OpenKE`

[10] `https://github.com/pyg-team/pytorch_geometric`

Table 6: Hyperparameters used for Baseline Models.

Model	Hidden Dim.	Num. Epochs	Batch Size	Learning Rate	Optimizer
TransE	100	10000	1000	1	SGD
TransR	50	2000	1000	1	SGD
RotatE	30	2000	1000	2e-5	Adam
DistMult	100	10000	1000	0.5	Adagrad
ComplEx	50	2000	500	0.5	Adagrad
ANALOGY	20	2000	500	0.5	Adagrad
SimpLE	100	2000	500	0.5	Adagrad

For the implementation of the `MultiHop` model, we use the source code[11] provided by Lin et al. [2018] and modify its reward function by using our defined reward shaping strategy (described in Sec. "Adversarial Actor-critic Reinforcement Learning" in the main text). We set all its parameters the same as the reinforcement learning model in the `KGML-xDTD` model framework if they are available otherwise we use the default parameters.

Section S5. A subgraph with the top 10 predicted 3-hop MOA paths for Factor VIIa – Hemophilia B pair

We utilize the `KGML-xDTD` model framework to predict the top 10 3-hop BKG-based paths, serving as biological explanations of the predicted "treats" relationship between Factor VIIa and Hemophilia B (shown in Table 4 for case study 1 in the main text). This particular drug/treatment - disease pair is not used in the training set and thus can be used to indicate how KGML-xDTD's MOA path predictions can contribute to the explanation of the predicted drug repurposing results. The predicted paths show similar molecular details as those in Figure 4 in the main text for treating hemophilia B. The path that we highlighted in red is the one in which all nodes can align with the key molecules in the real drug action regulatory network shown in Figure 4.

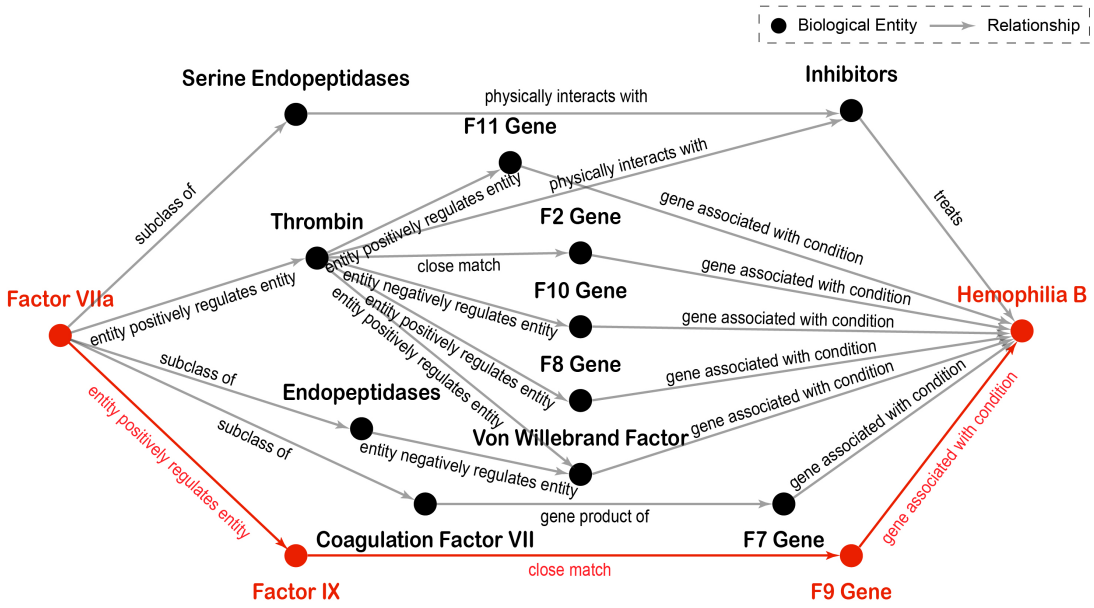

Figure 7: Top 10 predicted 3-hop paths (integrated into a graph for better visualization) generated by the `KGML-xDTD` model framework serve as biological explanations of the predicted "treats" relationship between Factor VIIa and Hemophilia B. The path highlighted with red is the one in which all nodes can align with the key molecules in the real drug action regulatory network shown in Figure 4 in the main text.

[11] `https://github.com/salesforce/MultiHopKG`