Supplementary Information for *"Improving target-disease association prediction through a graph neural network with credibility information"*

Chang Liu [1,†], Cuinan Yu [2,†], Yipin Lei [1,†],
Kangbo Lyu [1], Tingzhong Tian [1], Qianhao Li [3], Dan Zhao [1,*], Fengfeng Zhou [2,*], and Jianyang Zeng [1,*]

[1]Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China.
[2]Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, College of Computer Science and Technology, Jilin University, Changchun, Jilin, China.
[3]Machine Learning Department, Silexon AI Technology Co., Ltd., Nanjing, Jiangsu Province, China.

1 Baseline methods

In this section, we described the baseline methods used in comparison with CreaTDA.

1.1 GTN

Graph transformer networks (GTN) [5] is a state-of-the-art Graph Neural Network (GNN) that has been shown to achieve excellent performance on three benchmark node-classification tasks on heterogeneous networks (HNs). GTN utilizes "meta-paths", i.e., paths on an HN connected with heterogeneous edges, based on which new adjacency matrices can be generated by multiplying the individual adjacency matrices along the segments of the meta-path. Instead of relying on the pre-defined meta-paths that require domain knowledge or manual selection, GTN learns meta-paths by softly selecting adjacency matrices based on the attention mechanism.

Though GTN is a node classification model, we can easily reform it to perform TDA prediction by appending a network reconstruction module as in **Definition 6** in the main text (but without encoding credibility

information):

$$\min \sum_{r \in R} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} (m(e) - f^1(u)^T G_r H_r^T f^1(v))^2, \tag{1}$$

in which $f^1(u)$ and $f^1(v)$ come from Z (see Eq. 5 in [5]), a matrix representing all of the learned node embeddings.

1.2 DTINet

DTINet [3] is a machine learning framework for drug-target interaction (DTI) prediction. DTINet is based on random walk with restart (RWR) and diffusion component analysis (DCA), together with a final projection step similar to Eq. 1. To reform the DTI prediction task for TDA prediction, we only need to modify the RWR step, which generates diffusion states, i.e., latent representations, of relevant nodes. More Specifically, DTINet integrates all individual networks containing drugs and proteins (except for the drug-protein interaction network) to compute diffusion states of drug and protein nodes, respectively. We modified this by integrating all networks containing diseases and proteins (except for the TDA network) to compute the diffusion states of disease and protein nodes separately.

1.3 RGCN

Relational graph convolutional networks (RGCN) [4] is a simple GNN designed for addressing the prediction tasks related to the highly multi-relational graphs. RGCN employs the linear message transformations as simple graph convolutional layers applied on the hidden representations of both nodes and relations. To perform TDA prediction, we leveraged the same learning module as defined in Eq. 1, after obtaining the node embeddings $f^1(\cdot)$ through the forward propagation module of RGCN.

1.4 HGT

Heterogeneous Graph Transformer (HGT) [1] is one of the state-of-the-art GNNs tackling the large-scale heterogeneous graphs with millions of nodes. HGT adopts Heterogeneous Mutual Attention [1] to aggregate information from neighbors. Briefly, in each iteration, query vectors are generated from the source node, while the key and value vectors are generated from the neighborhood nodes. With the generated query, key, and value vectors, the attention mechanism is then employed. HGT additionally initiates a balanced neighborhood sampler, namely HGSampling [1], for large-scale training, which is not included in our imple-

mentation as only a relatively small amount of nodes are in the HN used (see Section 2.1 in the main text). HGT learns the contextualized node representations for the downstream tasks such as link prediction, and we use the same module as defined in Eq. 1 to obtain TDA predictions in our baseline comparison tests.

2 Ablation studies in the cluster-wise cross-validation test

2.1 Control models

As mentioned in the main text, we developed four models, namely CreaTDA_og (no credibility encoded), CreaTDA_rl (random soft labels), CreaTDA_rw (random penalty weights), and CreaTDA_rlrw (both random soft labels and random penalty weights) as control.

We first introduce the mathematical terms used in defining these control models, inheriting the notations defined in the main text:

Definition 1 (Random soft label) For an edge $e = (i, j, r)$ of edge-type $r \in R_c$ between entities i and j, its random soft label is defined as:

$$l'(e) = \begin{cases} \mathit{uniform}(\sigma(\alpha), 1), & m(e) = 1, \\ 0, & m(e) = 0, \end{cases} \tag{2}$$

where both $m(e)$ and α are the same as in Eq. 3 in the main text, $\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function, and $\mathit{uniform}(a, b)$ stands for a uniform sampler from $[a, b]$, whose values are computed offline before training.

Definition 2 (Random penalty weight) For an edge $e = (i, j, r)$ of edge-type $r \in R_c$ between entities i and j, the random penalty weight of the reconstruction error on e is defined as:

$$w'(e) = \begin{cases} \mathit{uniform}(\sigma(\beta), 1), & m(e) = 1, \\ 1, & m(e) = 0, \end{cases} \tag{3}$$

where both $m(e)$ and β are the same as in Eq. 4 in the main text, $\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid function, and $\mathit{uniform}(a, b)$ is the same as in Eq. 2 above, albeit performed independently.

We then define the optimization objectives of the four models as follows:

Definition 3 (CreaTDA_og) For the same parameter set Θ as defined in Eq. 5 in the main text, the

optimization objective of CreaTDA_og is:

$$\min_{\Theta} \sum_{r \in R} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} (m(e) - f^1(u)^T G_r H_r^T f^1(v))^2. \tag{4}$$

Definition 4 (CreaTDA_rl) For the same parameter set Θ as defined in Eq. 5 in the main text, the optimization objective of CreaTDA_rl is:

$$\begin{aligned} &\min_{\Theta} \sum_{r \in R \setminus R_c} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} (m(e) - f^1(u)^T G_r H_r^T f^1(v))^2 \\ &+ \sum_{r \in R_c} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} w(e)(l'(e) - f^1(u)^T G_r H_r^T f^1(v))^2. \end{aligned} \tag{5}$$

Definition 5 (CreaTDA_rw) For the same parameter set Θ as defined in Eq. 5 in the main text, the optimization objective of CreaTDA_rw is:

$$\begin{aligned} &\min_{\Theta} \sum_{r \in R \setminus R_c} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} (m(e) - f^1(u)^T G_r H_r^T f^1(v))^2 \\ &+ \sum_{r \in R_c} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} w'(e)(l(e) - f^1(u)^T G_r H_r^T f^1(v))^2. \end{aligned} \tag{6}$$

Definition 6 (CreaTDA_rlrw) For the same parameter set Θ as defined in Eq. 5 in the main text, the optimization objective of CreaTDA_rlrw is:

$$\begin{aligned} &\min_{\Theta} \sum_{r \in R \setminus R_c} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} (m(e) - f^1(u)^T G_r H_r^T f^1(v))^2 \\ &+ \sum_{r \in R_c} \sum_{\substack{u,v \in V \\ e=(u,v,r) \in E}} w'(e)(l'(e) - f^1(u)^T G_r H_r^T f^1(v))^2. \end{aligned} \tag{7}$$

2.2 Ablation study results

	CreaTDA_og	CreaTDA_rlrw	CreaTDA_rl	CreaTDA_rw	CreaTDA
AUROC	0.810 ± 0.008	0.806 ± 0.007	0.808 ± 0.009	0.808 ± 0.007	$\mathbf{0.814 \pm 0.007}$
AUPR	0.501 ± 0.017	0.490 ± 0.013	0.502 ± 0.022	0.501 ± 0.014	$\mathbf{0.516 \pm 0.016}$

Table S1: Results of ablation studies in the cluster-wise cross validation test (mean ± standard deviation). The results where CreaTDA outperformed all control methods are presented in boldface.

We can see from Table S1 that encoding credibility information improved the performance of CreaTDA on the cluster-wise cross validation test.

3 P-values of the Spearman correlations between the output values of top-k predictions and corresponding C_r values

k	GTN	DTINet	RGCN	HGT	CreaTDA_og	CreaTDA_rlrw	CreaTDA_rl	CreaTDA_rw	CreaTDA
200	0.243	0.964	0.368	0.227	0.669	0.892	0.561	0.467	0.024
500	0.757	0.666	0.669	0.002	0.411	0.582	0.079	0.022	0.012
1000	0.983	0.247	0.805	0.657	0.023	0.604	0.292	0.002	0.0
1500	0.515	0.829	0.321	0.655	0.006	0.54	0.007	0.0	0.0
2000	0.314	0.033	0.406	0.382	0.002	0.692	0.0	0.0	0.0
2500	0.8	0.027	0.444	0.218	0.253	0.439	0.0	0.0	0.0
3000	0.408	0.005	0.109	0.277	0.392	0.217	0.0	0.0	0.0

Table S2: The P-values of the Spearman correlations between the output values of the top-k ($k = 200, 500, 1000, 1500, 2000, 2500, 3000$) novel predictions produced by all models (i.e., CreaTDA and all of the baseline and control models), and their corresponding C_r values.

4 The Trie hashing algorithm for computing C_r values

To search for the co-occurrence between thousands of entities in millions of papers efficiently, we adopted the Trie hashing algorithm [2] for substring matching.

Algorithm 1 Substring Matching with Trie Hashing

Goal: Record co-occurrence $co[i, j]$ of M entities in N papers.
Collect all M names with at most k characters of required entities as $entity\_list[M][k]$.
Build Trie Tree:
for all $entity\_name \in entity\_list$ **do**
  $current\_node \leftarrow tree\_root$
  for all $character \in entity\_name$ **do**
    if $current\_node.has\_child(character)$ **then**
      $current\_node \leftarrow current\_node.child(character)$
    else
      $current\_node \leftarrow current\_node.add\_child(character)$
    end if
  end for
  $current\_node.color \leftarrow entity\_index$
end for
Search in N articles:
for all $article \in article\_list$ **do**
  for all $phrase \in article$ **do**
    $current\_node \leftarrow tree\_root$
    for all $character \in phrase$ **do**
      if $current\_node.has\_child(character)$ **then**
        $current\_node \leftarrow current\_node.child(character)$
      else
        $current\_node \leftarrow NULL$
        Break to match the next phrase.
      end if
    end for
    if $current\_node \neq NULL$ **then**
      Record $i \leftarrow current\_node.color$.
      for all recorded entity indices j in this article **do**
        $co[i, j] \leftarrow co[i, j] + 1$
      end for
    end if
  end for
end for

5 Top-200 predictions of CreaTDA trained on the whole HN

protein_idx	disease_idx	disease	protein_number	protein_name	output
501	1729	calculi	P08069	IGF1R	0.626
247	2566	hyperoxia	P31749	AKT1	0.598
2	2723	fasciitis	P35228	NOS2	0.608
636	2493	sneezing	P05412	JUN	0.642
94	2163	warts	P15692	VEGFA	0.848
553	1312	fibroma	P02778	CXCL10	0.757
553	1881	tachypnea	P02778	CXCL10	0.575
827	1729	calculi	P09874	PARP1	0.759
2	2572	retinal detachment	P35228	NOS2	0.814
654	1554	starvation	P05164	MPO	0.751
247	2215	connective tissue diseases	P31749	AKT1	0.559
690	1586	tuberous sclerosis	P01100	FOS	0.619
247	1763	hepatorenal syndrome	P31749	AKT1	0.574
459	1838	fragile x syndrome	P04150	NR3C1	0.668
440	2021	bronchiolitis	P00533	EGFR	0.578
664	1312	fibroma	P11413	G6PD	0.791
94	1600	virus diseases	P15692	VEGFA	0.632
542	2296	corneal edema	P05231	IL6	0.621
542	2322	esophageal motility disorders	P05231	IL6	0.585
698	1313	focal nodular hyperplasia	P00450	CP	0.583
1005	2102	mild cognitive impairment	Q16665	HIF1A	0.631
690	1310	fat necrosis	P01100	FOS	0.58
459	1839	lymphadenitis	P04150	NR3C1	0.715
690	2021	bronchiolitis	P01100	FOS	0.684
690	2767	thalassemia	P01100	FOS	0.59
660	2414	mycoplasma infections	P05362	ICAM1	0.694
654	2021	bronchiolitis	P05164	MPO	0.608
73	2329	failure to thrive	P10275	AR	0.564
580	2029	chromosomal instability	P17936	IGFBP3	0.653
624	2163	warts	P24385	CCND1	0.71
636	2242	acromegaly	P05412	JUN	0.678

660	2661	sleep apnea syndromes	P05362	ICAM1	0.626
440	2333	fetal hypoxia	P00533	EGFR	0.612
654	2566	hyperoxia	P05164	MPO	0.578
661	1542	scurvy	P19838	NFKB1	0.578
672	2029	chromosomal instability	P11021	HSPA5	0.563
690	2395	macroglossia	P01100	FOS	0.56
715	1307	facies	P42574	CASP3	0.555
560	1990	pressure ulcer	P01579	IFNG	0.681
654	1855	short bowel syndrome	P05164	MPO	0.618
614	1881	tachypnea	P09038	FGF2	0.612
610	1109	monosomy	P99999	CYCS	0.601
593	2482	seminoma	P01137	TGFB1	0.574
654	1854	retinopathy of prematurity	P05164	MPO	0.74
15	1554	starvation	P02452	COL1A1	0.723
94	1109	monosomy	P15692	VEGFA	0.718
542	2477	sarcoma, clear cell	P05231	IL6	0.7
593	2021	bronchiolitis	P01137	TGFB1	0.675
542	2164	anovulation	P05231	IL6	0.653
636	2215	connective tissue diseases	P05412	JUN	0.627
560	2395	macroglossia	P01579	IFNG	0.619
614	1766	irritable bowel syndrome	P09038	FGF2	0.591
151	1612	chlamydia infections	P29474	NOS3	0.589
150	2102	mild cognitive impairment	P35354	PTGS2	0.577
94	1546	shwartzman phenomenon	P15692	VEGFA	0.795
624	1834	marijuana abuse	P24385	CCND1	0.776
138	2378	leg ulcer	P10415	BCL2	0.739
827	1844	trismus	P09874	PARP1	0.72
668	1933	complex regional pain syndromes	P01138	NGF	0.686
924	1422	malnutrition	P21397	MAOA	0.685
837	1554	starvation	P55210	CASP7	0.659
150	2118	primary sclerosing cholangitis	P35354	PTGS2	0.648
636	1600	virus diseases	P05412	JUN	0.636
440	1313	focal nodular hyperplasia	P00533	EGFR	0.623

138	2767	thalassemia	P10415	BCL2	0.611
1005	1313	focal nodular hyperplasia	Q16665	HIF1A	0.606
624	2559	immune complex diseases	P24385	CCND1	0.584
440	2344	halitosis	P00533	EGFR	0.58
420	1773	mucocele	P09601	HMOX1	0.579
654	1846	ataxia with vitamin e deficiency	P05164	MPO	0.577
247	2004	osteochondrodysplasias	P31749	AKT1	0.572
690	2079	hypolipoproteinemias	P01100	FOS	0.569
636	2650	pulmonary atresia	P05412	JUN	0.567
247	1885	foot diseases	P31749	AKT1	0.566
732	1422	malnutrition	P15559	NQO1	0.563
119	2661	sleep apnea syndromes	P37231	PPARG	0.56
590	1387	jaw diseases	P29460	IL12B	0.56
715	2322	esophageal motility disorders	P42574	CASP3	0.553
636	1310	fat necrosis	P05412	JUN	0.849
636	1586	tuberous sclerosis	P05412	JUN	0.827
440	1859	brugada syndrome	P00533	EGFR	0.812
138	1628	fetal weight	P10415	BCL2	0.8
593	1837	delirium, dementia, amnestic, cognitive disorders	P01137	TGFB1	0.798
1393	1512	prion diseases	P04179	SOD2	0.784
138	2102	mild cognitive impairment	P10415	BCL2	0.784
1210	1668	megacolon	P20248	CCNA2	0.784
1202	1625	vascular neoplasms	P35869	AHR	0.775
629	1693	adrenal cortex neoplasms	P08183	ABCB1	0.761
690	1997	cattle diseases	P01100	FOS	0.716
594	1225	bronchitis	P35222	CTNNB1	0.711
711	2029	chromosomal instability	P98170	XIAP	0.706
711	1554	starvation	P98170	XIAP	0.7
1054	2368	intestinal polyposis	P08684	CYP3A4	0.698
683	2029	chromosomal instability	P08254	MMP3	0.697
1038	2029	chromosomal instability	P31350	RRM2	0.696
827	1837	delirium, dementia, amnestic, cognitive disorders	P09874	PARP1	0.691
636	1793	proctocolitis	P05412	JUN	0.684

1033	1837	delirium, dementia, amnestic, cognitive disorders	P05177	CYP1A2	0.684
501	1625	vascular neoplasms	P08069	IGF1R	0.681
636	1275	delayed graft function	P05412	JUN	0.677
884	1554	starvation	P06400	RB1	0.676
363	2029	chromosomal instability	P08253	MMP2	0.676
1118	1554	starvation	P05067	APP	0.67
654	1387	jaw diseases	P05164	MPO	0.668
1005	2126	rhinitis, allergic, perennial	Q16665	HIF1A	0.665
138	2766	nasopharyngitis	P10415	BCL2	0.662
636	2197	acquired hyperostosis syndrome	P05412	JUN	0.659
690	1793	proctocolitis	P01100	FOS	0.654
697	1600	virus diseases	P19438	TNFRSF1A	0.652
624	1247	carotid stenosis	P24385	CCND1	0.65
73	1773	mucocele	P10275	AR	0.649
642	2280	choledocholithiasis	P01584	IL1B	0.645
1116	1766	irritable bowel syndrome	P04637	TP53	0.643
1116	1191	alopecia areata	P04637	TP53	0.64
1036	780	mercury poisoning, nervous system	O00206	TLR4	0.633
1135	1987	myxoma	Q07817	BCL2L1	0.633
1116	2661	sleep apnea syndromes	P04637	TP53	0.631
560	1885	foot diseases	P01579	IFNG	0.628
400	2056	gastroparesis	P01375	TNF	0.627
1054	2159	dermatitis, occupational	P08684	CYP3A4	0.626
636	2336	flatulence	P05412	JUN	0.622
1033	2164	anovulation	P05177	CYP1A2	0.621
420	1837	delirium, dementia, amnestic, cognitive disorders	P09601	HMOX1	0.62
626	2506	systemic inflammatory response syndrome	P60568	IL2	0.62
15	2004	osteochondrodysplasias	P02452	COL1A1	0.617
501	2164	anovulation	P08069	IGF1R	0.616
1393	1763	hepatorenal syndrome	P04179	SOD2	0.616
658	1859	brugada syndrome	P06493	CDK1	0.615
138	2716	hip fractures	P10415	BCL2	0.615
642	2199	femoral fractures	P01584	IL1B	0.613

400	2234	pelvic inflammatory disease	P01375	TNF	0.609
636	1356	histiocytosis	P05412	JUN	0.606
642	1505	polyomavirus infections	P01584	IL1B	0.605
924	1837	delirium, dementia, amnestic, cognitive disorders	P21397	MAOA	0.605
636	1923	ascorbic acid deficiency	P05412	JUN	0.602
1054	2586	critical illness	P08684	CYP3A4	0.598
654	1612	chlamydia infections	P05164	MPO	0.596
697	1181	adenocarcinoma, bronchiolo-alveolar	P19438	TNFRSF1A	0.594
1116	1384	iridocyclitis	P04637	TP53	0.594
420	2197	acquired hyperostosis syndrome	P09601	HMOX1	0.59
1393	402	dystocia	P04179	SOD2	0.589
1033	2401	marfan syndrome	P05177	CYP1A2	0.588
629	2548	insulinoma	P08183	ABCB1	0.587
690	1886	hyperkeratosis, epidermolytic	P01100	FOS	0.587
642	1306	facial dermatoses	P01584	IL1B	0.587
1116	1584	tuberculosis, pulmonary	P04637	TP53	0.585
138	2308	diverticulitis, colonic	P10415	BCL2	0.584
690	2474	rhinitis, allergic, seasonal	P01100	FOS	0.584
1135	2566	hyperoxia	Q07817	BCL2L1	0.583
711	1624	splenic neoplasms	P98170	XIAP	0.582
711	2102	mild cognitive impairment	P98170	XIAP	0.581
1116	2646	obstetric labor complications	P04637	TP53	0.581
636	2401	marfan syndrome	P05412	JUN	0.581
629	1624	splenic neoplasms	P08183	ABCB1	0.58
642	2300	cytomegalovirus infections	P01584	IL1B	0.58
636	1416	lymphedema	P05412	JUN	0.58
138	2427	oligomenorrhea	P10415	BCL2	0.579
560	1574	tracheal diseases	P01579	IFNG	0.577
642	2396	malabsorption syndromes	P01584	IL1B	0.576
138	2748	aggressive periodontitis	P10415	BCL2	0.576
624	1353	hidradenitis suppurativa	P24385	CCND1	0.575
1116	1971	spondylitis	P04637	TP53	0.575
94	2381	leukemia, myelomonocytic, acute	P15692	VEGFA	0.575

1054	2293	colorectal neoplasms, hereditary nonpolyposis	P08684	CYP3A4	0.573
1132	1766	irritable bowel syndrome	P01308	INS	0.573
1054	1387	jaw diseases	P08684	CYP3A4	0.573
636	1885	foot diseases	P05412	JUN	0.573
1116	2274	cerebral arterial diseases	P04637	TP53	0.572
1054	2333	fetal hypoxia	P08684	CYP3A4	0.572
641	2414	mycoplasma infections	P13500	CCL2	0.57
1118	1834	marijuana abuse	P05067	APP	0.569
933	2159	dermatitis, occupational	P11802	CDK4	0.569
697	1353	hidradenitis suppurativa	P19438	TNFRSF1A	0.567
1135	2567	neoplastic processes	Q07817	BCL2L1	0.567
1393	2118	primary sclerosing cholangitis	P04179	SOD2	0.566
636	1439	microvascular angina	P05412	JUN	0.565
1116	2249	airway remodeling	P04637	TP53	0.563
636	1417	lymphocele	P05412	JUN	0.563
367	1612	chlamydia infections	P00390	GSR	0.562
1118	1318	giant cell arteritis	P05067	APP	0.562
363	2215	connective tissue diseases	P08253	MMP2	0.562
999	1628	fetal weight	Q13547	HDAC1	0.561
65	2548	insulinoma	O95477	ABCA1	0.561
367	1610	arrest of spermatogenesis	P00390	GSR	0.561
2	1122	polyploidy	P35228	NOS2	0.56
1393	1109	monosomy	P04179	SOD2	0.56
367	1628	fetal weight	P00390	GSR	0.557
891	1422	malnutrition	P02461	COL3A1	0.557
1116	2650	pulmonary atresia	P04637	TP53	0.556
440	1445	multiple sclerosis	P00533	EGFR	0.556
646	2129	silicosis	P27361	MAPK3	0.556
1135	35	lead poisoning, nervous system	Q07817	BCL2L1	0.555
1054	1542	scurvy	P08684	CYP3A4	0.555
544	2401	marfan syndrome	P03956	MMP1	0.555
624	2354	hernia, abdominal	P24385	CCND1	0.555
1013	2102	mild cognitive impairment	P00441	SOD1	0.555

560	2368	intestinal polyposis	P01579	IFNG	0.555
918	1422	malnutrition	P09210	GSTA2	0.553
654	35	lead poisoning, nervous system	P05164	MPO	0.553
138	1802	schizotypal personality disorder	P10415	BCL2	0.553

Table S3: The top-200 novel predictions of CreaTDA trained on the whole HN. "protein_idx" and "disease_idx" represent the row and column indices in our constructed TDA network, repsectively. "protein_number" represents the uniprot ID of a given protein. "output" represents the output value of a predicted TDA.

References

[1] Z. Hu, Y. Dong, K. Wang, and Y. Sun. Heterogeneous graph transformer. In *Proceedings of The Web Conference 2020*, pages 2704–2710, 2020.

[2] W. Litwin. Trie hashing. In *Proceedings of the 1981 ACM SIGMOD international conference on Management of data*, pages 19–29, 1981.

[3] Y. Luo, X. Zhao, J. Zhou, J. Yang, Y. Zhang, W. Kuang, J. Peng, L. Chen, and J. Zeng. A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information. *Nature communications*, 8(1):1–13, 2017.

[4] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. v. d. Berg, I. Titov, and M. Welling. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pages 593–607. Springer, 2018.

[5] S. Yun, M. Jeong, R. Kim, J. Kang, and H. J. Kim. Graph transformer networks. *Advances in neural information processing systems*, 32, 2019.