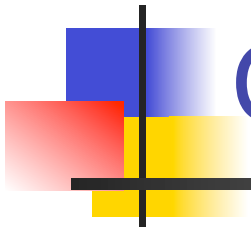
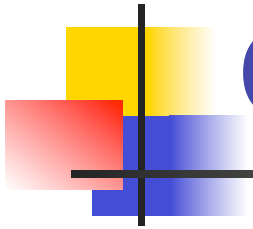


Ranking Semantic Associations in Chemical System Biology Space

Xiao Dong
School of Informatics and Computing
Indiana University at Bloomington

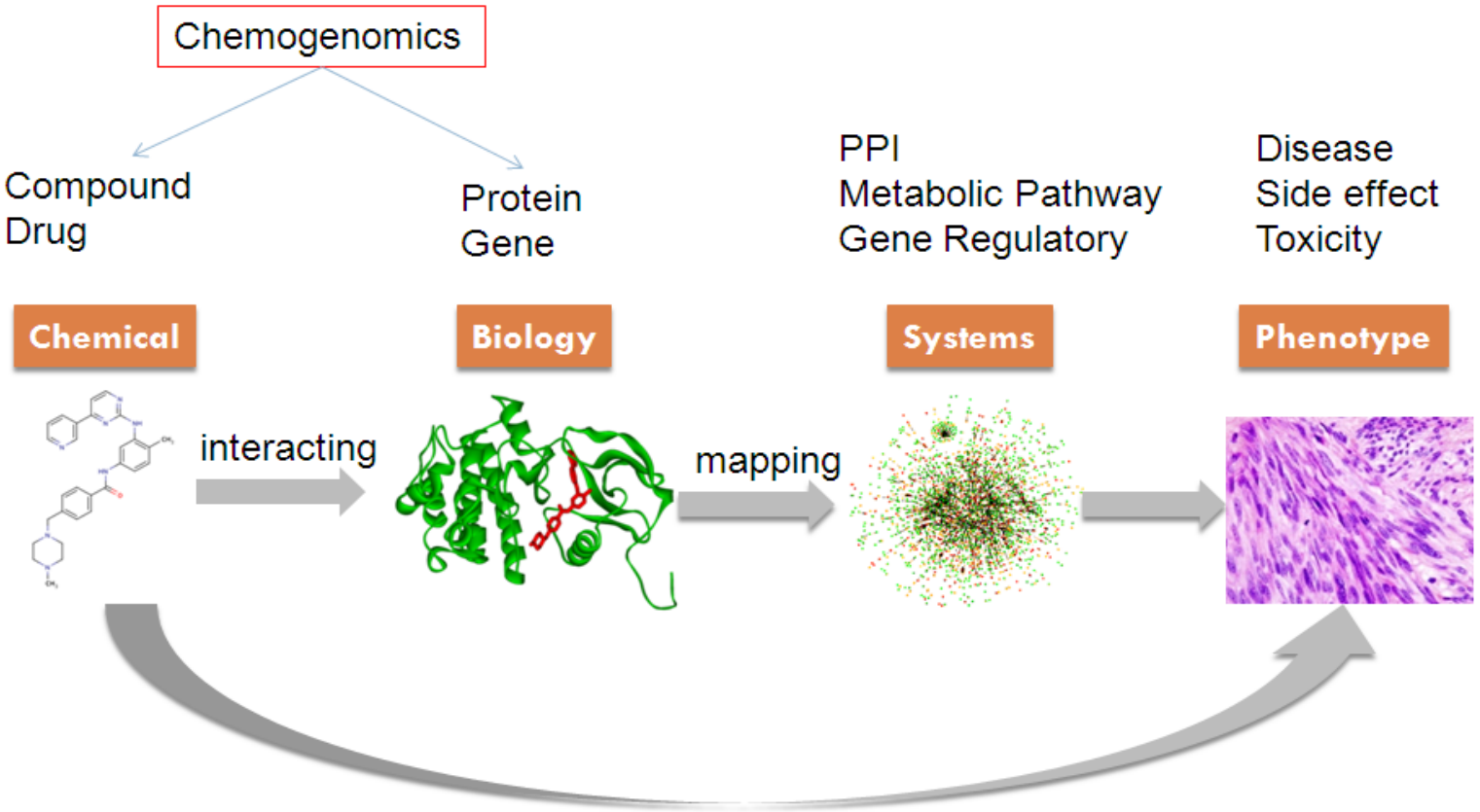
Presented at FWCS2010, April 26th
WWW2010, Raleigh, NC



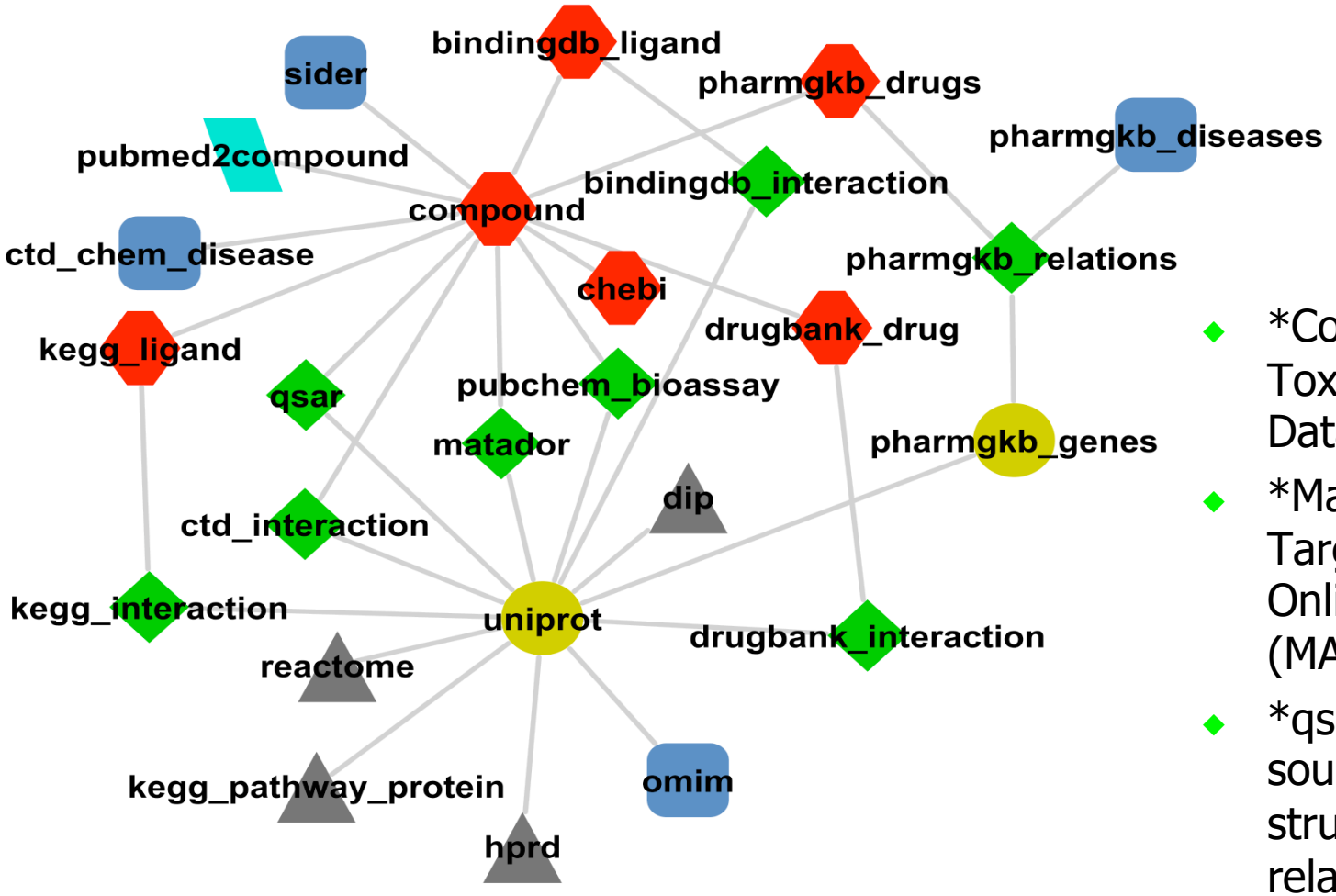
Chem2Bio2RDF

- www.chem2bio2rdf.org
- Initiated by Prof. Ding (NIH VIVO investigator) at Fall 2009.
- Work done at chemical informatics group (Prof. Wild) at IU.
- Primary goal is to bring chemogenomics data source onto linked data cloud.

Chemogenomics and Chemical Systems Biology

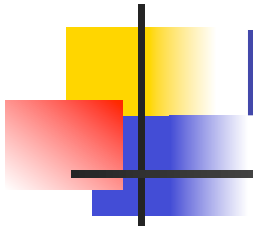


As of Today



- *Comparative Toxicogenomics Database (CTD)
- *Manually Annotated Targets and Drugs Online Resource (MATADOR)
- *qsar is a small online source for quantitative structure activity relationship

www.chem2bio2rdf.org

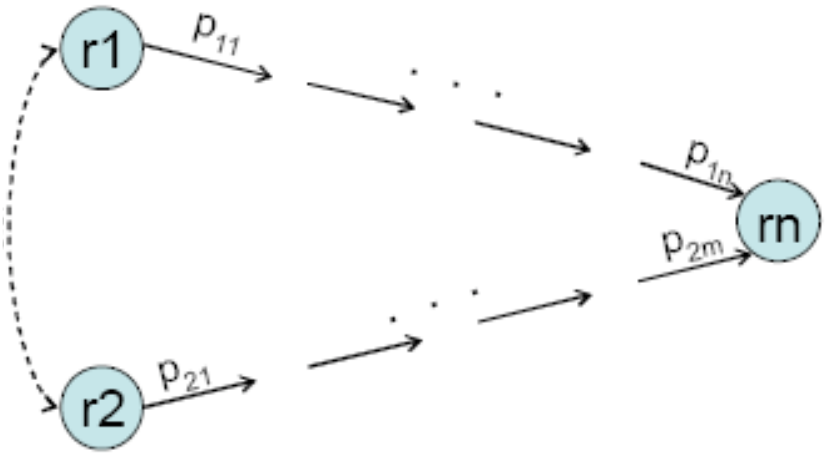


Motivation

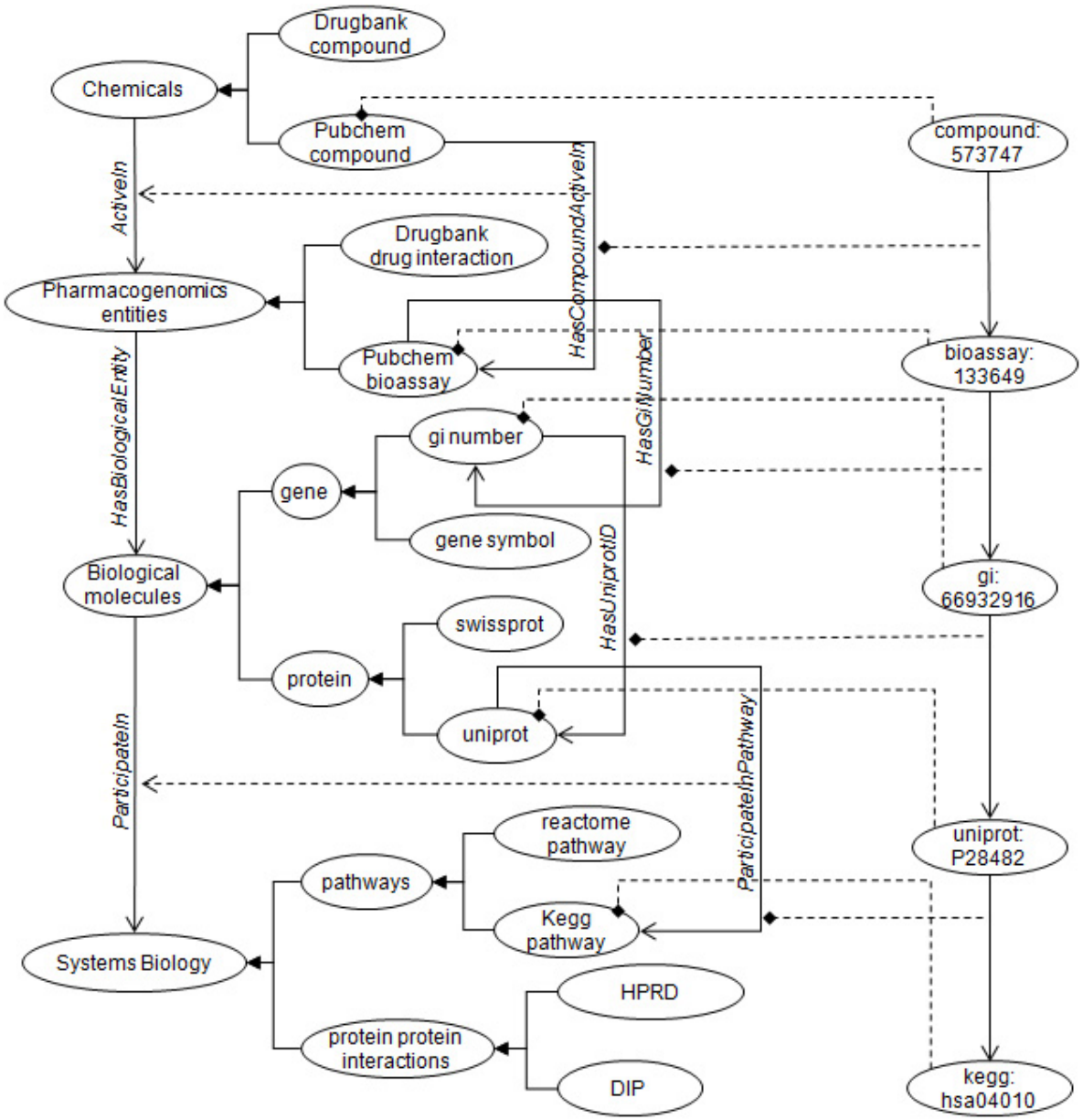
- Many important research questions can be boiled down to indentifying crucial associations among biological entities
  - find gene targets which the compound is active against
  - find compounds associated to a certain disease (Alzheimer)
  - find compounds cause a given adverse drug effect (Hypertension)
  - find all the active compounds in PubChem sharing at least two common targets with a FDA approved drug
  - find all the compounds in PubChem active towards at least two targets that are in the same pathway
  - find KEGG pathways containing at least two of the targets associated with a given side effect

Semantic Associations Formally

- Semantic association refers to a sequence of {subject property object} triples that connect two instances
- ρ-path association
- ρ-join association
- ρ-iso association

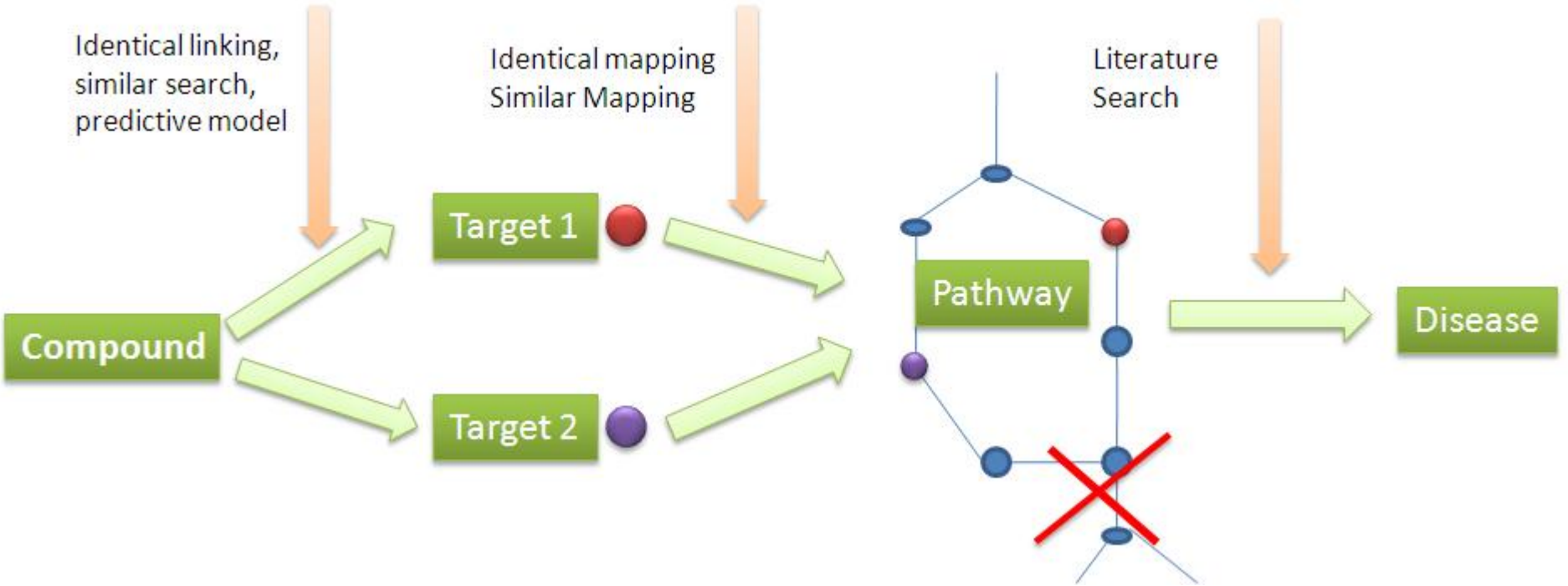


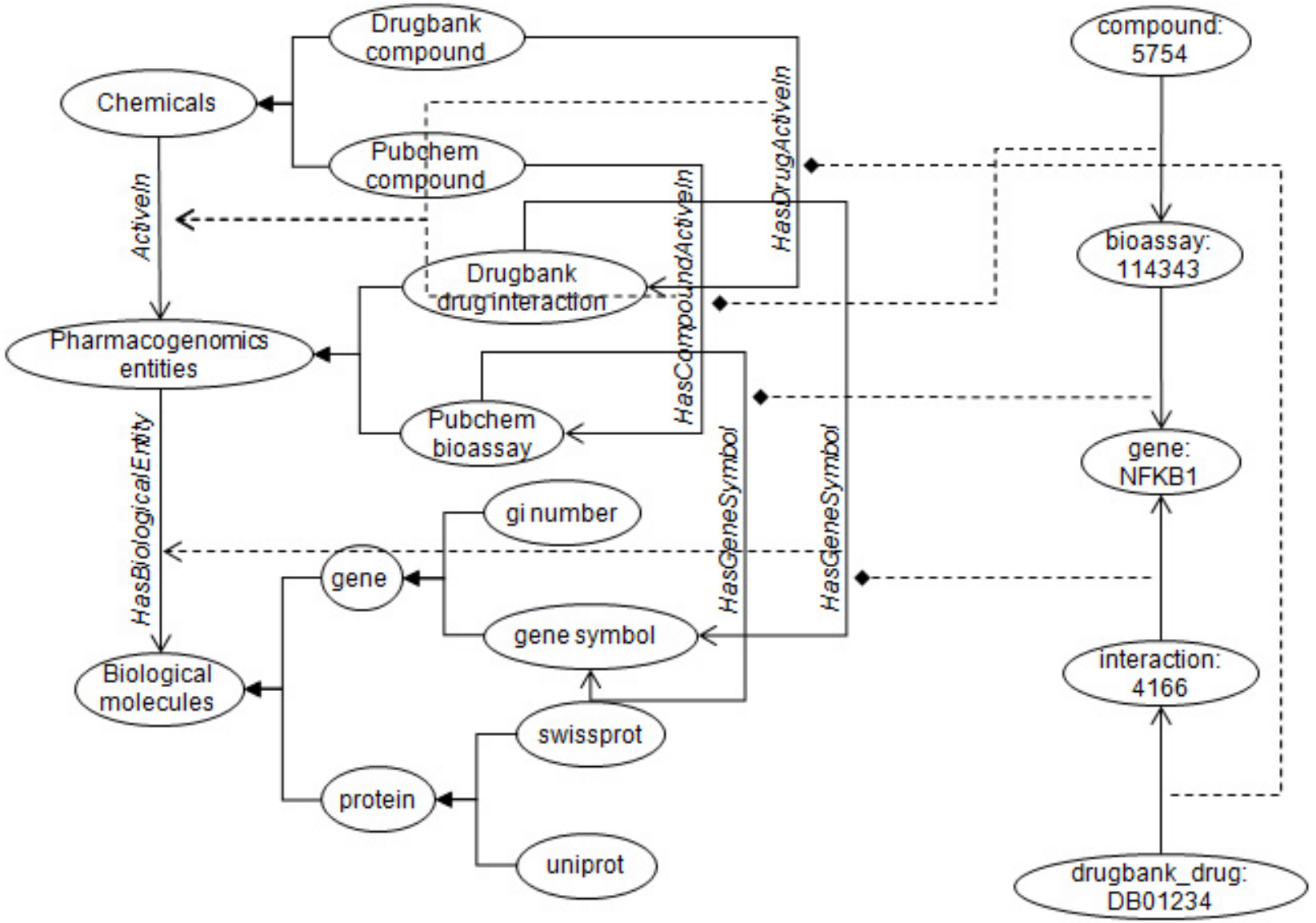
SemRank by *Anyanwu, K. et al.* WWW 2005



ρ -path in chemical systems biology

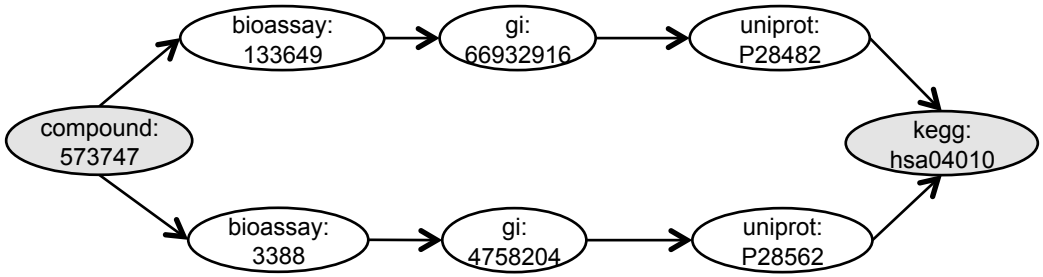
Multiple Pathway Inhibitor



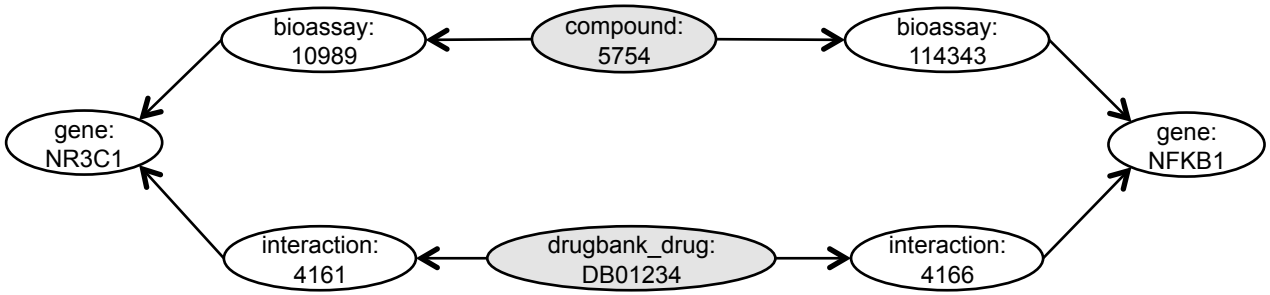


ρ -join in chemical systems biology

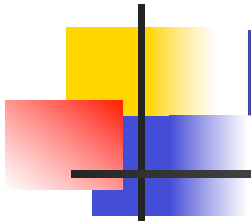
Two Intriguing Cases



(b) Double ρ Path Association, Identify multiple pathway inhibitor

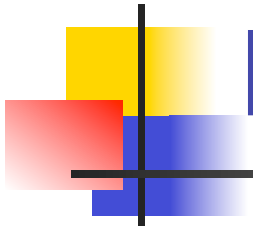


(a) Double ρ Join Association, polypharmacology



Polypharmacology

- Conventional drug embraces the dogma "one gene, one drug, one disease"
- Polypharmacology focuses on multi-target drugs
  - Identify target leads to unwanted side effect
  - Enhance therapeutic potency
- Identify compounds sharing targets with drugs of known polypharmacology



Ranking

- Intrinsic metrics
  - Multiple Path Aggregator is applied to promote the special case associated pairs conform to special interests (i.e. multiple pathway inhibitor or polypharmacology)
- Extrinsic metrics
  - Quality
  - Specificity
  - Rarity/Popularity
- Weighted sum taken as the total ranking score
- Evaluation using pubmed literature abstract co-occurrence score

Association Pairs (Compound/ Dexamethasone)

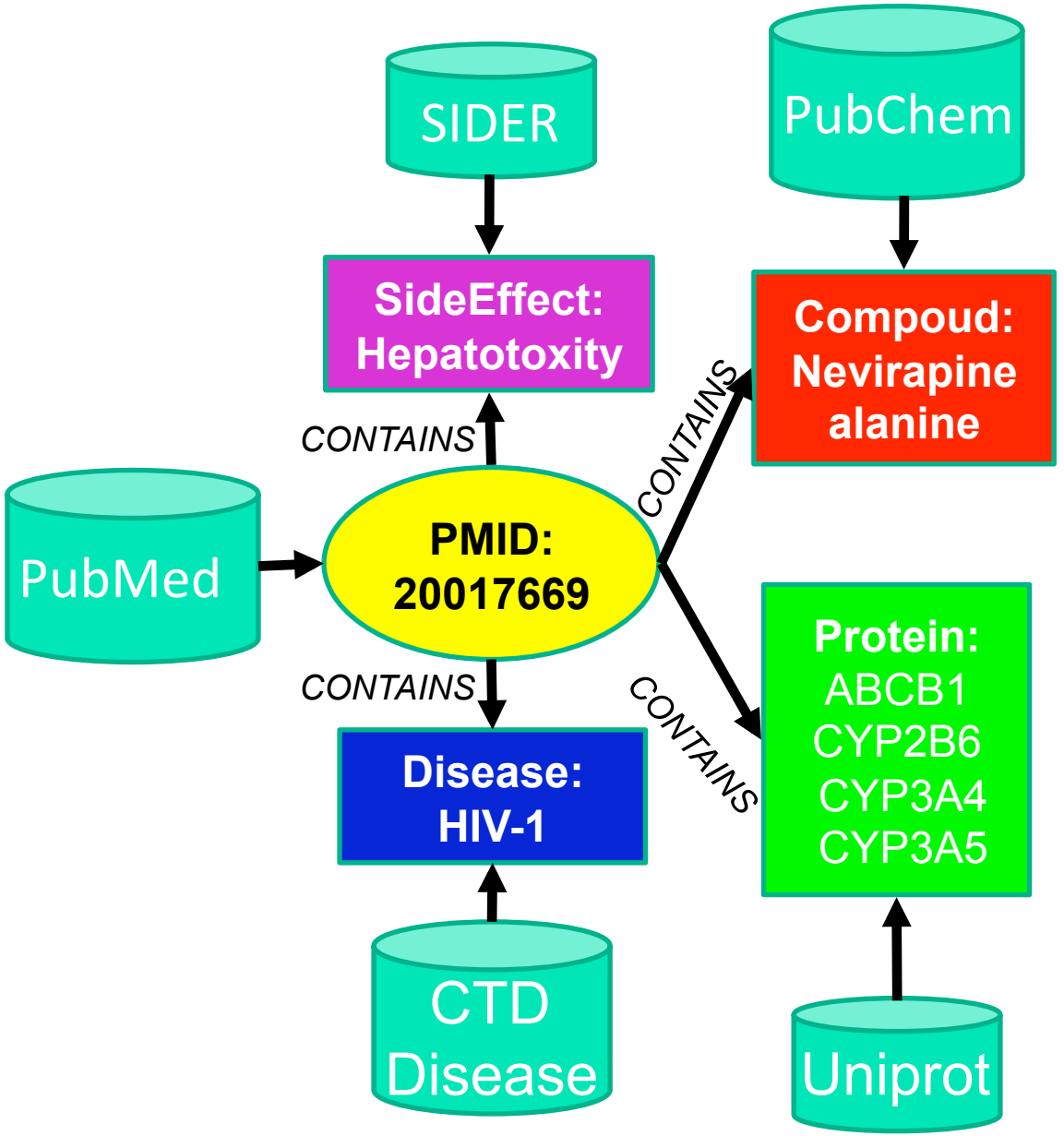
data source	number of associations found	polypharmacology cases found
PubChem Bioassay	1123	2
CTD	318	21
BindDB	117	0
TTD	30	0
MATADOR	0	N/A
QSAR	0	N/A
PharmGKB	0	N/A

Chem/Bio Entity Extraction from PubMed

PMID: 20017669

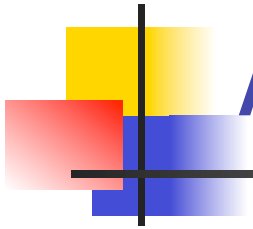
Title: Nevirapine-induced hepatotoxicity and pharmacogenetics: a retrospective study in a population from Mozambique.

Abstract: Aims: Nevirapine is widely used to treat HIV-1 infection to prevent mother-to-child transmission; unfortunately adverse drug reactions have been reported. Our aim was to identify genes/variants involved in nevirapine-induced hepatotoxicity. MATERIALS & METHODS: Patients from Mozambique, 78 with nevirapine-induced hepatotoxicity and 78 without adverse events, were genotyped for ABCB1, CYP2B6, CYP3A4 and CYP3A5 gene variants. We conducted a case-control association study and a genotype/phenotype correlation analysis. RESULTS: The ABCB1 c.3435C>T SNP was associated with hepatotoxicity ($p = 0.038$), with the variant T allele showing a protective effect (odds ratio: 0.42). Moreover, four SNPs in the CYP2B6 and CYP3A5 genes resulted significantly correlated with transaminase values. In particular, for the CYP2B6 c.983T>C SNP, the difference in the alanine aminotransferase mean values were highly significant between TT and TC genotypes ($p < 0.001$). CONCLUSION: Our preliminary results confirm the contribution of the ABCB1 c.3435C>T SNP in nevirapine-induced hepatotoxicity risk and, at the same time, suggest the necessity for further studies.



Validation with Literature Co-occurrence

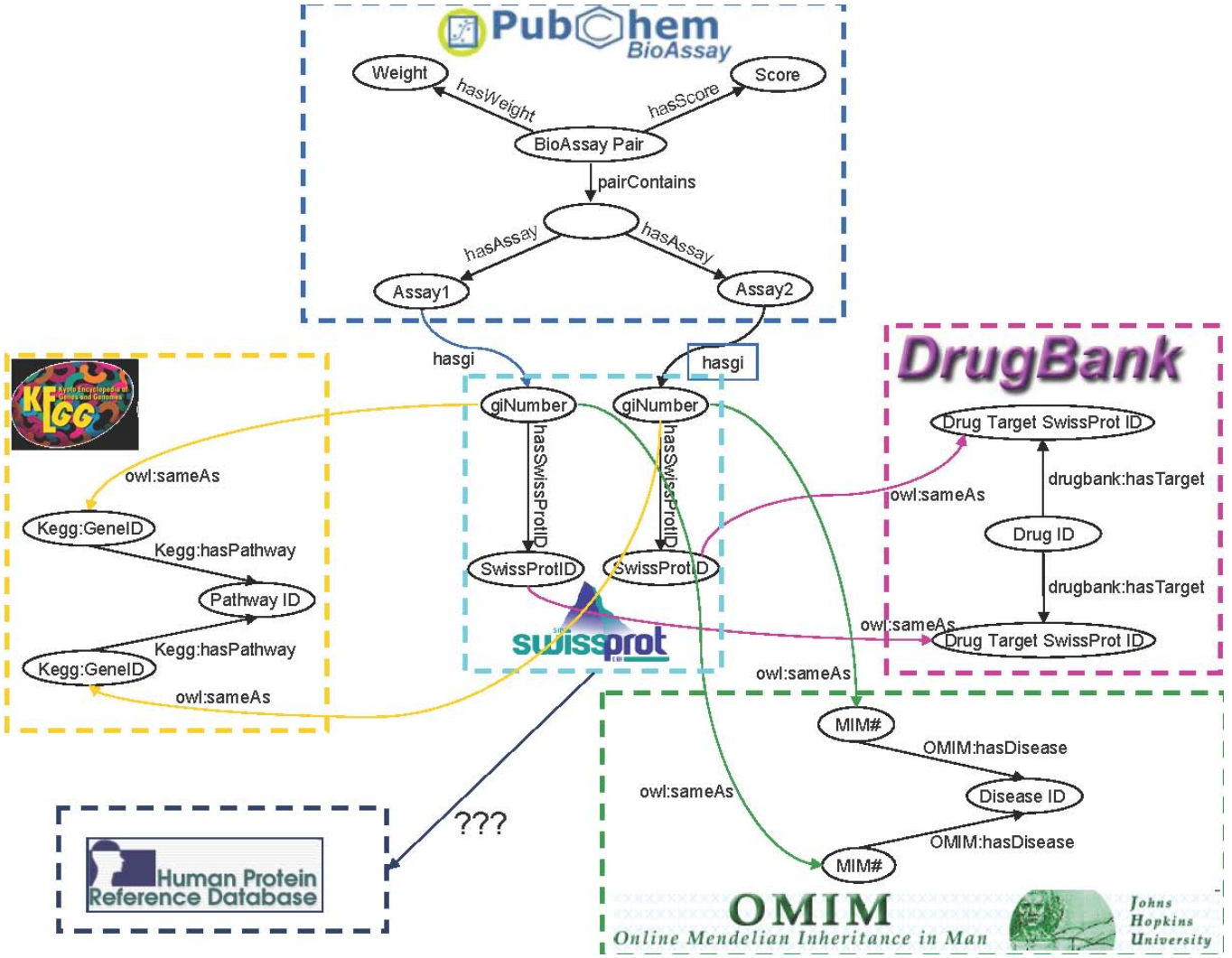
PubChem ID	quality	distinctiveness	specificity	ranking score	
443495	4	3	0	7	✓
55245	4	3	0	7	✓
969516	3	2.01	0	5.01	
74990	3	2.01	0	5.01	✓
5743	2	1.02	1	4.02	✓



Acknowledgement

- NIH VIVO project
- Eli Lilly and Company

As of Oct. 2009



Extends to LODD and Bio2RDF