

Interações entre Droga e Doença por meio de Genes

Mateus Siqueira Batista Nicolas Bissoli Nattis

MC536 - Instituto de Computação, UNICAMP

2020

Proposta

Obter dados de interações entre genes e drogas, e entre genes e doenças.

Proposta

Obter dados de interações entre genes e drogas, e entre genes e doenças.

Através disso, podemos relacionar a interação entre estas drogas e as doenças.

Proposta

Obter dados de interações entre genes e drogas, e entre genes e doenças.

Através disso, podemos relacionar a interação entre estas drogas e as doenças.

Exemplo:

- Droga A ativa o gene X.
- Gene X tem relação de causa com as doenças α , γ .
- Portanto, a droga A tem relação de causa com as doenças α , γ .



- Dados sobre interações droga-gene e o genoma drogável.
- Extraído de mais de trinta fontes confiáveis.
- Dados extraídos via arquivo .tsv disponibilizados.



- Plataforma contendo uma das maiores coleções publicamente disponíveis de genes e variantes associados a doenças humanas.
- Dados extraídos via banco de dados SQLite disponibilizado.

A relação de droga-doença ocorre de modo intuitivo a partir das relações de doença e droga com o gene.

		Droga-Gene	
		ativação	inibição
Gene-Doença	ativação	ativação	inibição
	inibição	inibição	ativação

Tipos de Interação

Activator

Activator

Inducer

Agonist

Binder

Chaperone

Partial Agonist

Potentiator

Stimulator

Substrate

Inhibitor

Antibody

Antagonist

Inhibitor

Antisense
Oligonucleotide

Blocker

Inhibitory
Allosteric
Modulator

Inverse Agonist

Partial Antagonist

Suppressor

DgiDb

Susceptibility
Mutation

Chromosomal
Rearrangement

Somatic Causal
Mutation

Fusion Gene

Germline Causal
Mutation

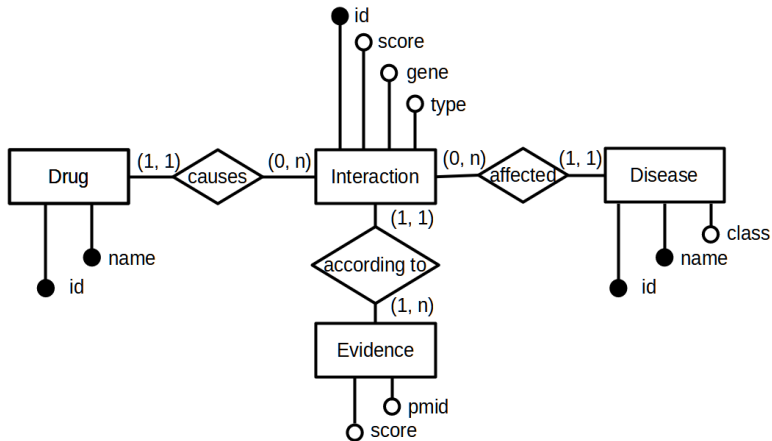
GeneticVariation

Biomarker

Therapeutic

DisgeNet

Modelo Conceitual



Modelo Lógico

Drug(DrugId, Name)

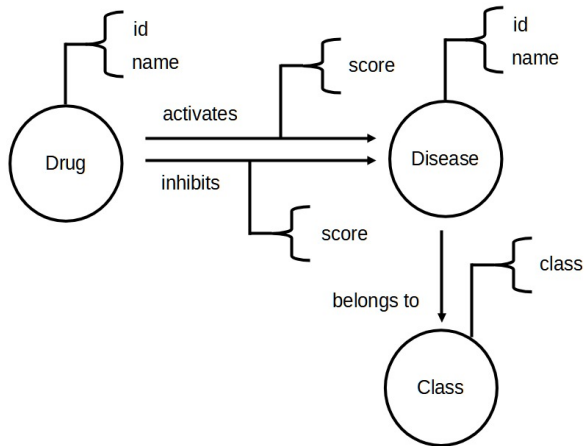
Disease(DiseaseId, Name, Class)

Modelo Lógico

```
Interaction(InteractionId,  
           DrugId,  
           DiseaseId,  
           Score,  
           Gene,  
           Type)
```

```
Evidence(InteractionId,  
         Pmid,  
         Score)
```

Modelo Lógico de Redes Complexas



Evolução do projeto

- Removemos classes de drogas e aliases (fora do escopo).
- Dividimos as interações do Neo4j em duas.
- Usamos SQLite e TSV ao invés de requisições HTTP (muito lento).

Transformação

- Transformar DISGENET SQLite em TSV.
- Ler TSV das fontes.
- Produzir TSV final.
- Transformar TSV final em SQLite.
- Transformar TSV final em CSV para Neo4j.

Tratamento de Dados

■ Confibialidade

- Os dados coletados possuem níveis de confiabilidade variáveis, que foram levados em conta e tranformados em um Score.
- Há um Score atrelado ao **Dgldb** e outro ao **DisGeNET**. Assim, obtivemos um Score global da interação multiplicando um pelo outro.

Perguntas de Análise

Selecione o nome das drogas e doenças das N interações com maior pontuação.

```
SELECT Dr.Name, Di.Name, I.Type, I.Score  
FROM Interaction as I, Drug as Dr, Disease as Di  
WHERE Dr.Id = I.DrugId AND Di.Id = I.DiseaseId  
ORDER BY I.Score DESC LIMIT 10;
```


Perguntas de Análise

Selecione o nome das drogas e doenças das N interações com maior pontuação.

Dr.Name	Di.Name	I.Type	I.Score
NITISINONE	Tyrosinemia, Type III	0	0.861673561599624
BUROSUMAB	Autosomal dominant hypophosphatemic rickets	0	0.33348003403456
NITISINONE	Hawkinsinuria	0	0.669172021242262
BUROSUMAB	Hypophosphatemic Rickets	0	0.416674001701728
BUROSUMAB	TUMORAL CALCINOSIS, HYPERPHOSPHATEMIC, FAMILIAL	0	0.16674001701728
NITISINONE	Tyrosinemias	0	0.40333656074876
GOLODIRSEN	Muscular Dystrophy, Duchenne	1	0.388903558959012
GLEMBATUMUMAB VEDOTIN	Malignant neoplasm of breast	1	0.386666666666667
UROFOLLITROPIN	Ovarian Hyperstimulation Syndrome	1	0.346659625033008
THIAMINE	Dystonia	1	0.333339201361383

Perguntas de Análise

Quais drogas tem relação com a doença (Acute lymphocytic leukemia) C0023449?

```
SELECT Dr.Id, Dr.Name, I.Type, I.Score  
FROM INTERACTION as I, DRUG as Dr, Disease as Di  
WHERE Di.id ='C0023449' AND I.DiseaseId = Di.id  
ORDER BY I.Score DESC LIMIT 10;
```

Perguntas de Análise

Quais drogas tem relação com a doença (Acute lymphocytic leukemia) C0023449?

Dr.Id	Dr.Name	I.Type	I.Score
chembl:CHEMBL398707	HYDROMORPHONE	0	0.0388903558959012
chembl:CHEMBL3545253	FLORTAUCIPIR F 18	0	0.0388903558959012
chembl:CHEMBL2	PRAZOSIN	0	0.0388903558959012
chembl:CHEMBL1621597	IPRATROPIUM	0	0.0388903558959012
chembl:CHEMBL157101	KETOCONAZOLE	0	0.0388903558959012
chembl:CHEMBL1670	MITOTANE	0	0.0388903558959012
chembl:CHEMBL723	CARVEDIOL	0	0.0388903558959012
chembl:CHEMBL157138	LISURIDE	0	0.0388903558959012
chembl:CHEMBL2103830	FOSTAMATINIB	0	0.0388903558959012
chembl:CHEMBL1201250	BENZQUINAMIDE	0	0.0388903558959012

Perguntas de Análise

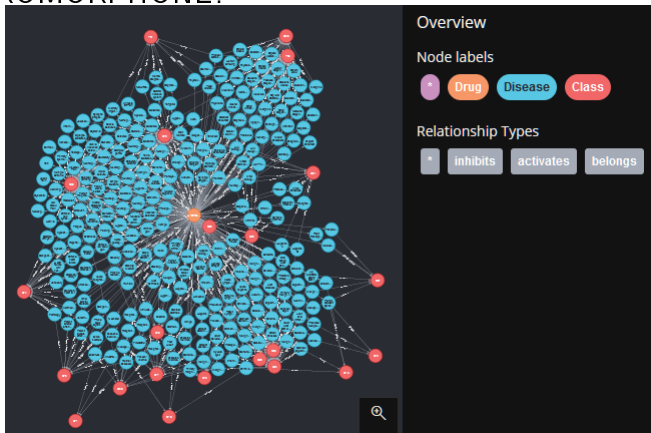
Quais classes de doenças são mais ativadas pela HYDROMORPHONE?

- Uma análise aprofundada é necessária, mas podemos partir da seguinte.

```
MATCH (dr:Drug)-[activates]->(di:Disease)
-[belongs]->(c:Class)
WHERE dr.name='HYDROMORPHONE'
RETURN dr, activates, di, belongs, c
```

Perguntas de Análise

Quais classes de doenças são mais ativadas pela HYDROMORPHONE?



Obrigado pela atenção!