

- ・ AI : 新しい発見
- ・ 自動化 : 研究の効率化
- ・ 視覚化 : 直感的な意思決定



# LSKB Ver.7 概要

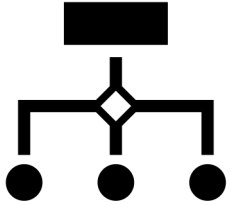
2021/12/20

緑川 淳

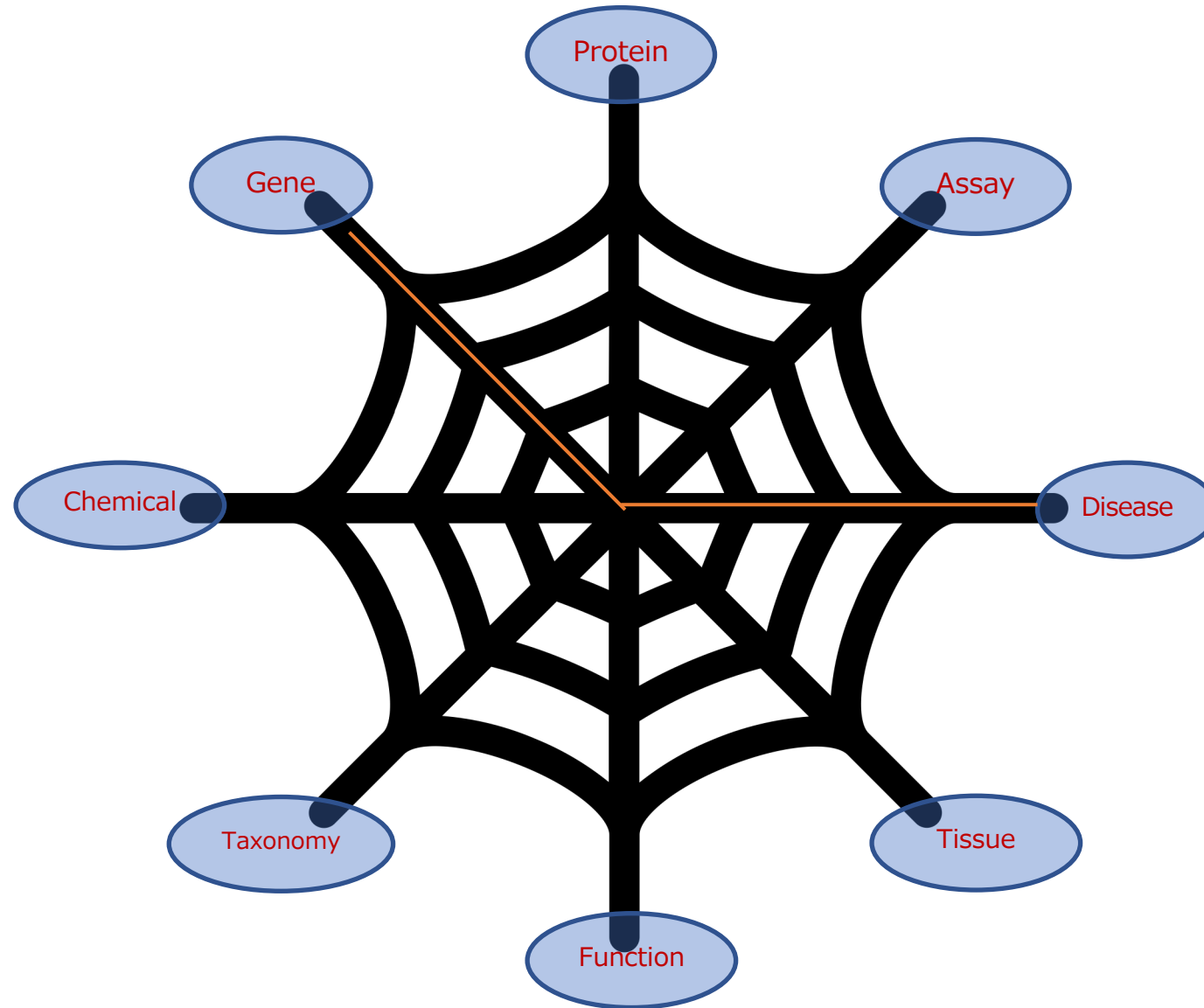
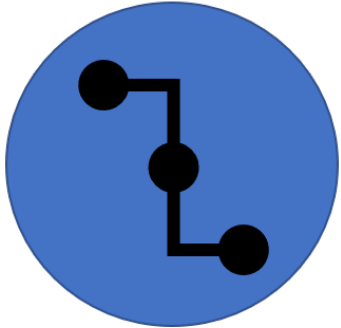
株式会社ワールドフュージョン

# LSKB is scientific web

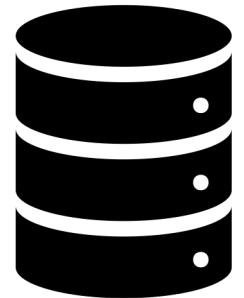
Ontology



Interaction

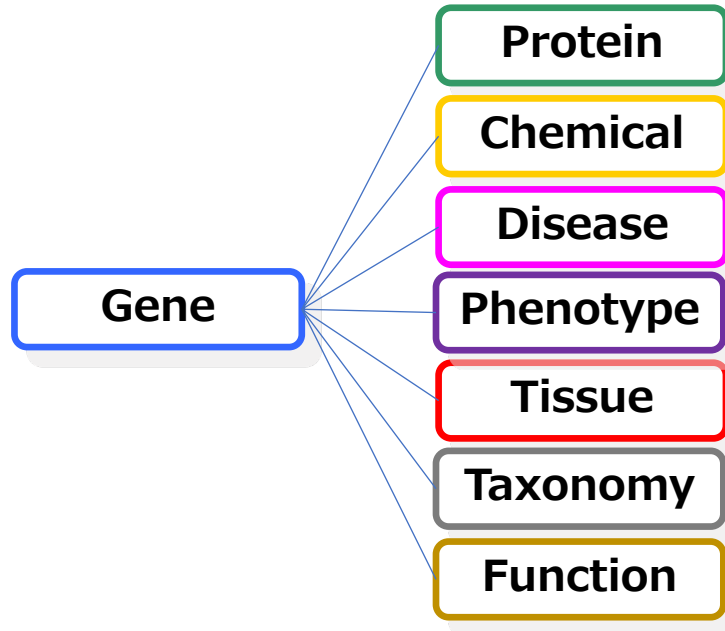
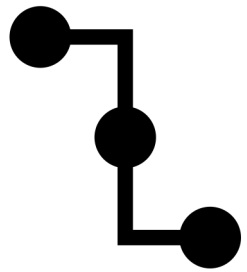


Contents



# LSKB Interaction

Interaction



## 1) Text-Mining

- 20 years of PubMed literature (3 levels)
- Clinical Trial
- Assay Description

## 2) Assay Data

- Target Gene/Protein - Chemical
- Activities Endpoint
- Mode : Inhibition agonism/antagonism
- Expression: up/down regulation
- GWAS

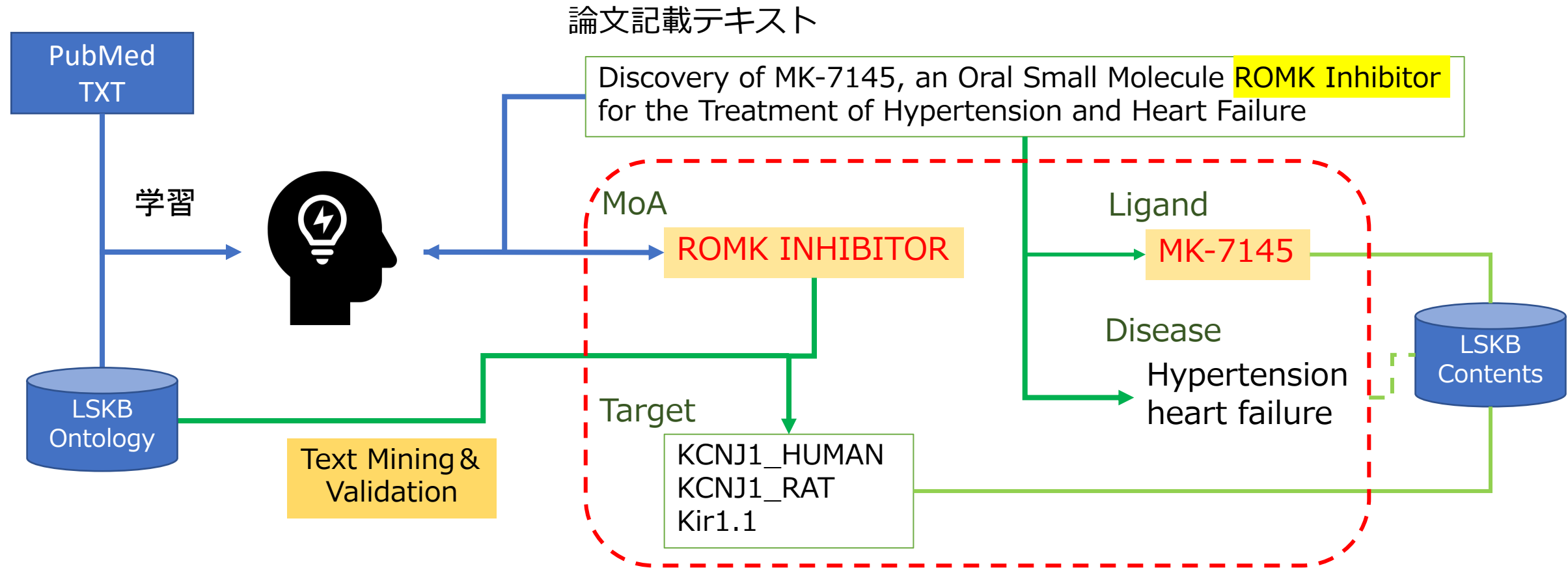
## 3) Curated Annotation

- Disease Target
- Gene Ontology
- Pathway

## 4) AI Curated Annotation

- Mechanism of Action

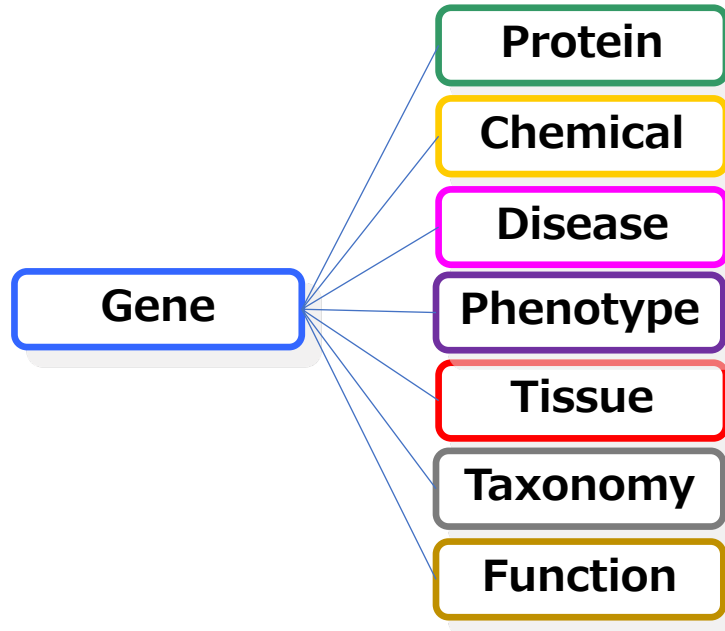
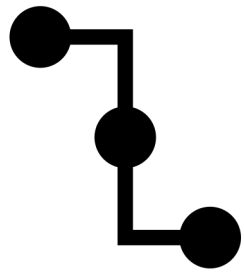
# 新規MoA 抽出と後処理プロセス概要



ラベル	Recall(再現率)	Precision(精度)	F1 Score
MoA	98.30%	95.30%	96.80%

# LSKB Interaction

Interaction



## 1) Text-Mining

- 20 years of PubMed literature (3 levels)
- Clinical Trial
- Assay Description

## 2) Assay Data

- Target Gene/Protein - Chemical
- Activities Endpoint
- Mode : Inhibition agonism/antagonism
- Expression: up/down regulation
- GWAS

## 3) Curated Annotation

- Disease Target
- Gene Ontology
- Pathway

## 4) AI Curated Annotation

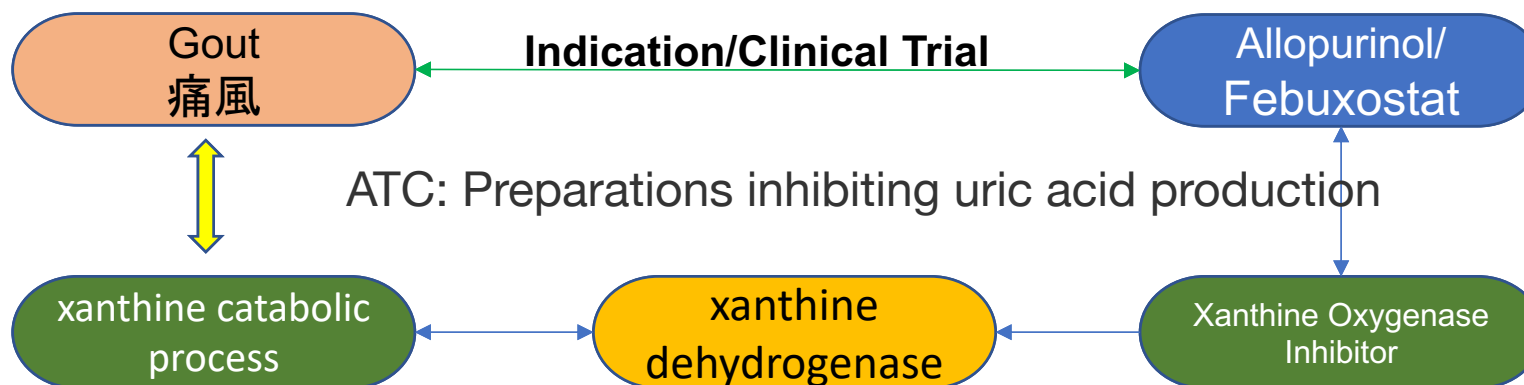
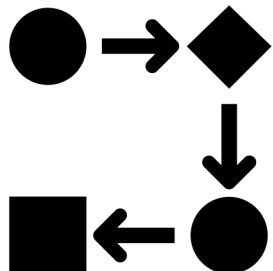
- Mechanism of Action

## 5) Misc

- **GO-MoA database**

**Disease related biological function database constructed by relationship of mechanism of action and the disease.**

## GO-MoA database

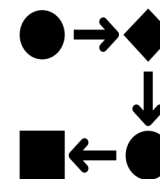


- **12,973 diseases related biological function**
- **99,610 combination of GO-MoA**
- **12,731 drugs and 2,164 targets are included.**
- **2,042 considerable human targets**
  - **functionally similar but unknown target to the disease**
- **4,019 considerable drugs**
  - **for drug repurposing**
- **750k active compounds to the targets**  
In case of p-value <0.0001



# Amyotrophic Lateral Sclerosis (ALS)

GO-MoA database



Disease

Amyotrophic Lateral Sclerosis

Rare

T047:Disease or Syndrome

BioAssayOntology ExperimentalFactorOntology Guide to Pharmacology Medgen Orphanet UMLS

RNA

GO Top Layer :

Border of pValue :

Sort by : GO

3rd 4th

1

Drug

<<

GO (Biological Process) Drug

0 9 18 27 36



## Sub-categories of regulation of RNA metabolic process

Gene Drug

GO (Biological Process)	# of Hit Drug	# of Hit Considerable Drug	p-Value	References	# of MOA
positive regulation of transcription from RNA polymerase II promoter	28	1657	1.68E-7	0	18
negative					

## Number of common drugs

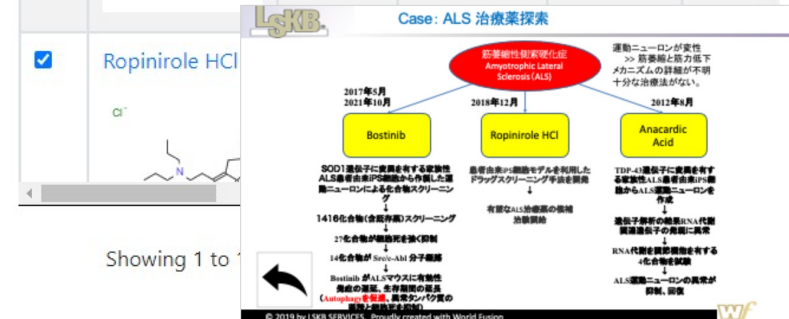
	2	3	4	5	6	7	8	9
Drug	13	0	5	3	0	1	0	2
Considerable Drug	660	137	150	91	4	57	58	113

## positive regulation of transcription from RNA polymerase II promoter related drugs

Send Chemical ID Filter Data Apply Filter(s) Download SD File Download TXT File

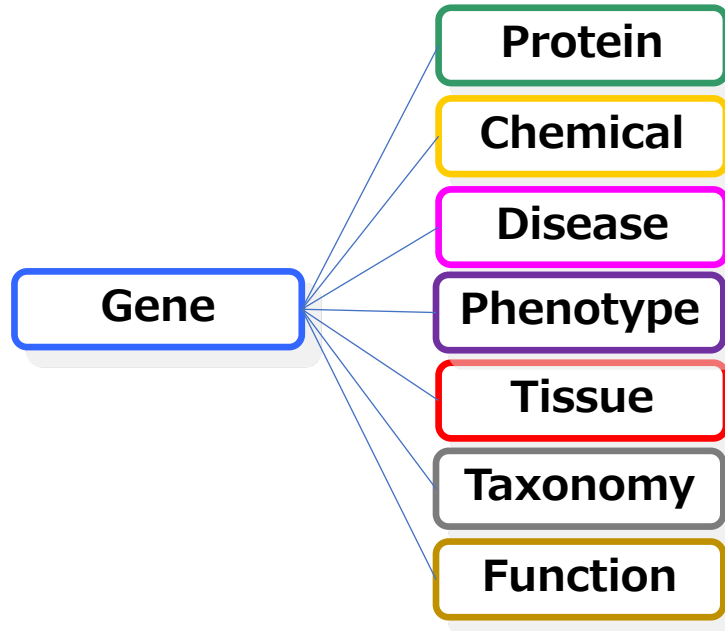
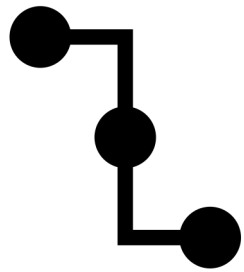
Show 25 entries

<input checked="" type="checkbox"/>	Chemical	Molecule Type	Phase	Withdrawn
<input checked="" type="checkbox"/>	ropinirole	Small molecule	0	<ul style="list-style-type: none"><li>N04</li><li>N04 Dop</li></ul>



# LSKB Interaction

Interaction



## 1) Text-Mining

- 20 years of PubMed literature (3 levels)
- Clinical Trial
- Assay Description

## 2) Assay Data

- Target Gene/Protein - Chemical
- Activities Endpoint
- Mode : Inhibition agonism/antagonism
- Expression: up/down regulation
- GWAS

## 3) Curated Annotation

- Disease Target
- Gene Ontology
- Pathway

## 4) AI Curated Annotation

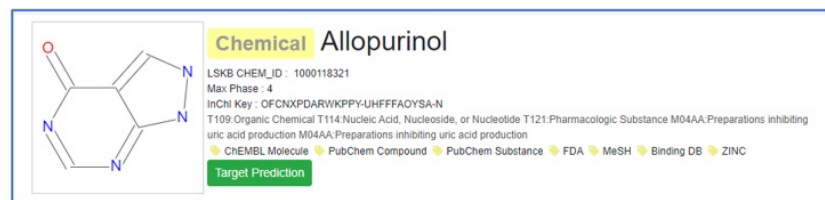
- Mechanism of Action
- **Gene RIF**

## 5) Misc

- Protein Pocket Similarity



# Similar Proteinと 適応症



**Function** Xanthine dehydrogenase inhibitor

XDH\_HUMAN  
Xanthine  
dehydrogenase/oxidase



Known Pocket Similar

HMDH\_HUMAN  
3-hydroxy-3-methylglutaryl-  
coenzyme A reductase

**Function** Hydroxymethylglutaryl-CoA Reductase Inhibitors

MeSH NDFRT ChEMBL



**Disease** Gout

T047:Disease or Syndrome

BioAssayOntology ExperimentalFactorOntology Guide to Pharmacology Medgen UMLS

**Disease** Renal Insufficiency

T047:Disease or Syndrome

Medgen UMLS

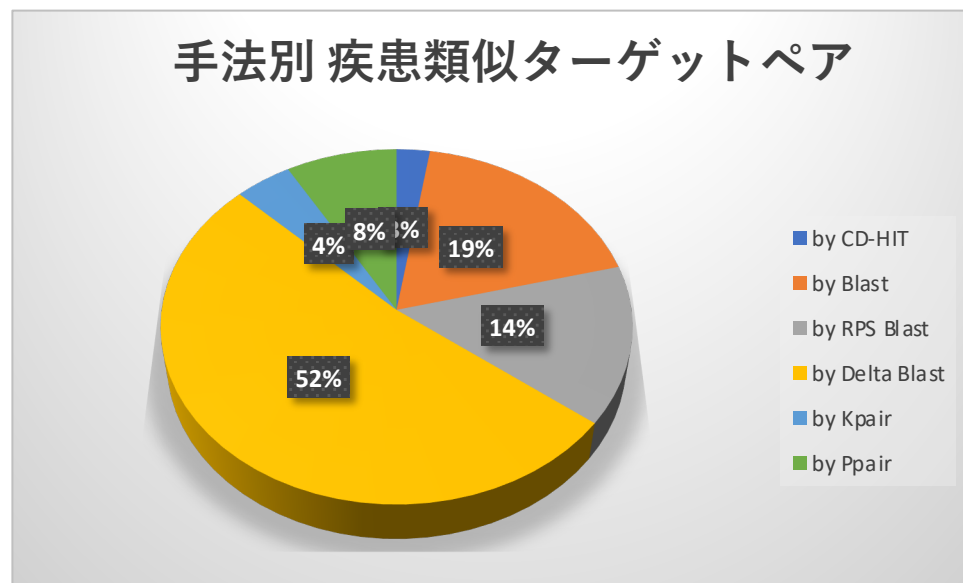
**Disease** Hypercholesterolemia

T047:Disease or Syndrome

Guide to Pharmacology Medgen UMLS

# 同一疾患における類似ターゲットタンパク質

- Phase4/3の疾患-Drugのターゲットタンパク質において、類似タンパク質ペアは611,731件
- 既知ポケットの形状の類似性 (Kpair: PoSSuM) では45,181件

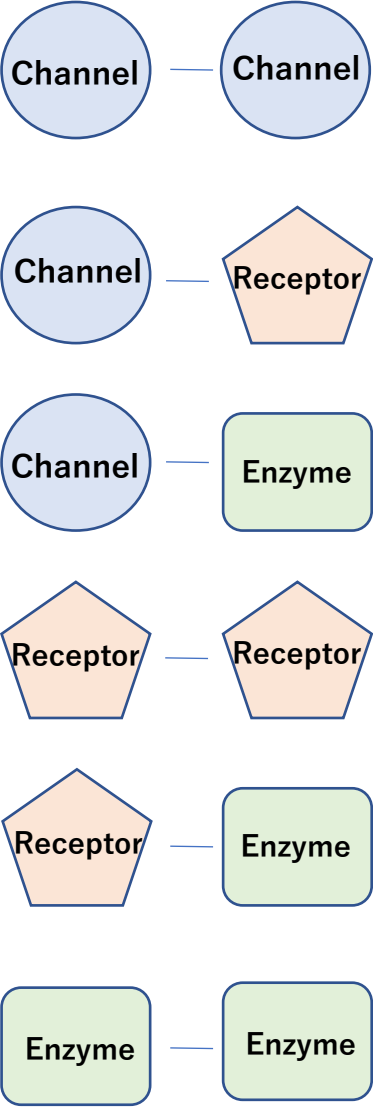


PoSSuM: <http://possum.cbrc.jp/PoSSuM/>

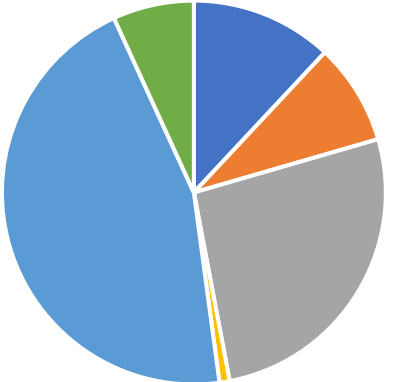
MeSH Classification	Ratio
Nervous System Diseases	13.1%
Neoplasms	11.2%
Pathological Conditions, Signs and Symptoms	6.8%
Cardiovascular Diseases	5.8%
Male Urogenital Diseases	5.3%
Congenital, Hereditary, and Neonatal Diseases and Abnormalities	5.2%
Female Urogenital Diseases and Pregnancy Complications	4.8%
Mental Disorders	4.8%
Digestive System Diseases	4.5%
Bacterial Infections and Mycoses	3.9%
Skin and Connective Tissue Diseases	3.8%
Immune System Diseases	3.7%
Hemic and Lymphatic Diseases	3.5%
Musculoskeletal Diseases	3.1%
Respiratory Tract Diseases	3.0%
Nutritional and Metabolic Diseases	2.8%
Endocrine System Diseases	2.7%
Eye Diseases	2.5%

# Nervous System Disease クラスにおけるターゲット間の ポケット形状の類似タンパク質ペア

UNIPROT_ENTRY_NAME	UNIPROT_ENTRY_NAME
KCNK2_HUMAN	PCSK9_HUMAN
KCNKA_HUMAN	PCSK9_HUMAN
GCR_HUMAN	SCN9A_HUMAN
DACB_ECOLI	PBPB_ECOLI
KCNKA_HUMAN	PPARG_HUMAN
P2Y12_HUMAN	PDE5A_HUMAN
DRD3_HUMAN	PPARG_HUMAN
MCR_HUMAN	PPARG_HUMAN
KCNKA_HUMAN	PDE4D_HUMAN
KCNKA_HUMAN	PDE9A_HUMAN
PDE5A_HUMAN	PPARG_HUMAN
PDE4D_HUMAN	SCN9A_HUMAN
GCR_HUMAN	VDR_HUMAN
KCNKA_HUMAN	PDE6C_HUMAN
KCNK2_HUMAN	PDE4B_HUMAN
ANT3_HUMAN	GCR_HUMAN
PDE4B_HUMAN	PPARG_HUMAN
KCNKA_HUMAN	SCN9A_HUMAN
KCNKA_HUMAN	PDE5A_HUMAN
MCR_HUMAN	PAR1_HUMAN
ANT3_HUMAN	THRB_HUMAN
HMDH_HUMAN	PDE4B_HUMAN
PDE4A_HUMAN	PPARG_HUMAN



ターゲットタンパク質クラス間のペア

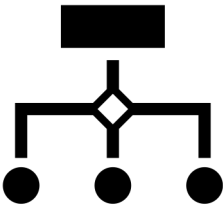


■ C-E ■ C-R ■ R-E ■ C-C ■ R-R ■ E-E

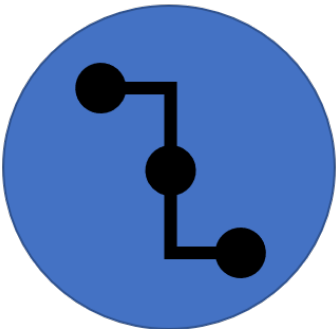
Pair-Class	Count
C-E	14
C-R	10
R-E	31
C-C	1
R-R	53
E-E	8

# LSKB is scientific web

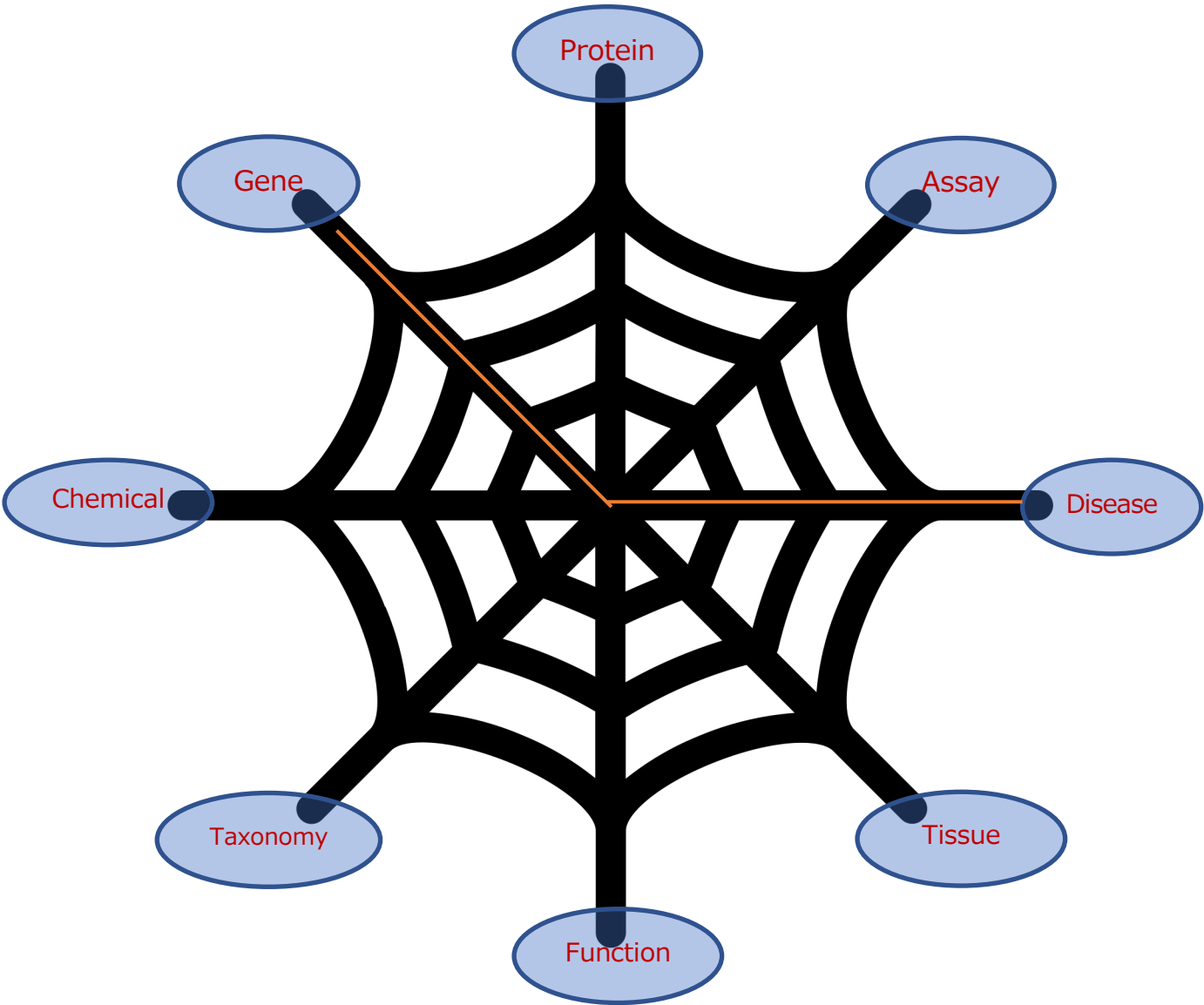
Ontology



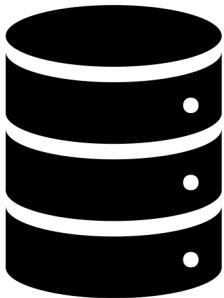
Interaction



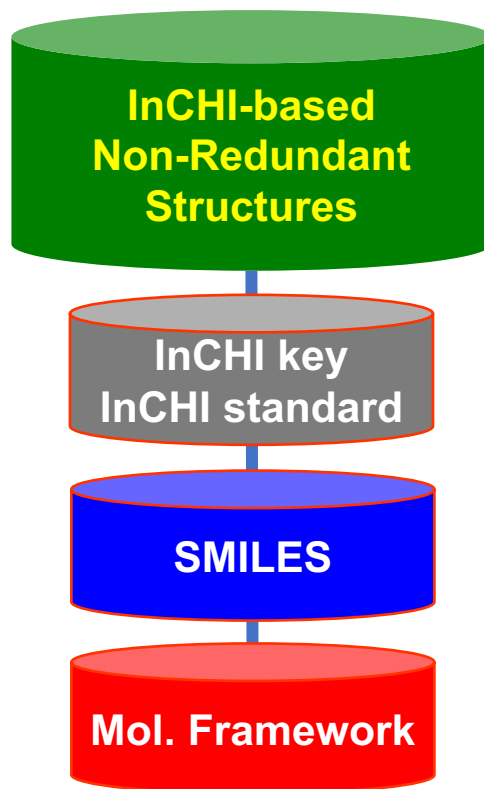
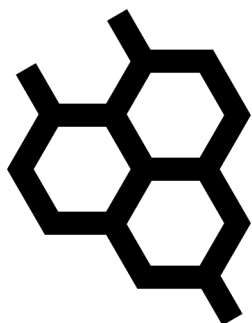
Structure



Contents



## Structure



### 1) Huge Structure Collection

- InCHI-based Non-redundant
- 100Million
- Molecular Framework

### 2) Structure Search

- Exact Match
- Similarity Search
- Substructure Search
- Molecular Framework search

### 3) Prediction/Analysis

- Target Prediction
- Target Confirmation
- Mechanism Explorer

### 4) Misc

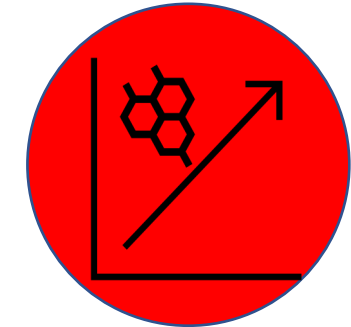
- Multiple structures search
- Automatic saving of search result

# LSKB is scientific web

Bioinformatics



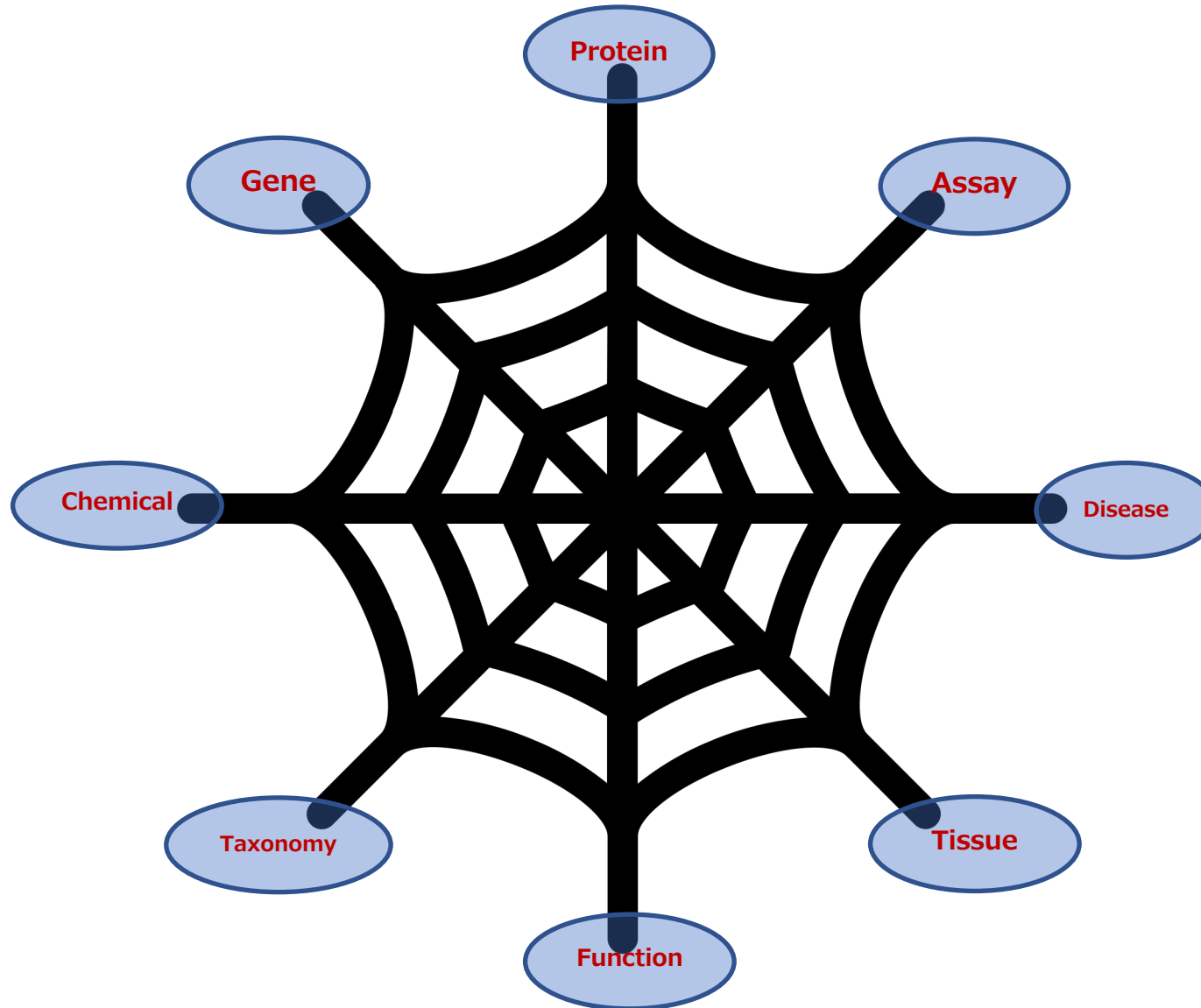
Cheminformatics



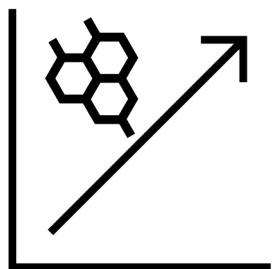
Exploration



Analysis



## Cheminformatics



Target Prediction

Target Confirmation

Chemical Properties

Molecular Framework

Go To MOE

Workflow





### LSKB の起動

読み込んだPDBファイルの情報を  
LSKBで確認

QFSS による LSKB搭載化合物の高速  
検索

デザインした化合物のTarget  
Prediction

抽出した骨格/部分構造のターゲットを  
探す Target Confirmation

MOESAIC でのLSKB構造検索

... etc

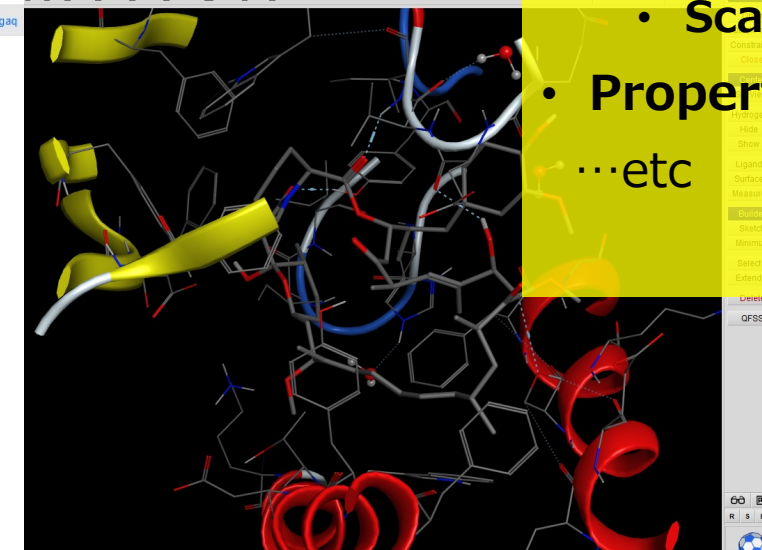
### MTOR\_HUMAN

> UniProt

Protein Names	Serine/threonine-protein kinase mTOR (2.7.11.1) (FK506-binding protein 12-rapamycin) (Mammalian target of rapamycin) (Mechanistic target of rapamycin) (Rapa
Taxonomy	<a href="#">Homo sapiens</a> [ Taxonomy ID : 9606 ] > <a href="#">NCBI</a>

OFF-X

Check All	Clear	-----	Go
PDB ID	Title	Ligand	Source / Resolution
<input type="checkbox"/> 1aue	FKBP-RAPAMYCIN BINDING DOMAIN (FRB) OF THE FKBP-RAPAMYCIN ASSOCIATED PROTEIN		Xray 2.33
<input type="checkbox"/> 1f6	MUNOPHILIN-P12-RAPAMYCIN COMPLEX FRAP	RAP	Xray 2.70
<input type="checkbox"/> 1n1	MUNOPHILIN-P12-RAPAMYCIN COMPLEX FRAP	RAD	Xray 2.20
<input type="checkbox"/> 2fap	THE STRUCTURE OF THE IMMUNOPHILIN-	RAD	Xray 2.20
<input type="checkbox"/> 2gaq			



- PDBファイルをMOEで開く
  - **Docking Study** など
- 活性情報を含むSDF ファイルからMOEでの解析
  - **QSAR**
  - **Scaffold 抽出**
- **Property**計算
- ...etc





## Exploration



From Disease list

Target Explorer

GO Explorer  
-Biological Process-

From Gene list

Disease Explorer

Expression Explorer

From Chemical structure & Gene list

Mechanism Explorer

Drug Repurposing  
Procedure

LSKB



World Fusion

# Target Explorer



MeS

Ne

Eye

Glaucoma  
Glaucoma, Open-Angle  
Angle Closure Glaucoma  
Low Tension Glaucoma  
Hydrophthalmos

Disease Group vs Gene   Disease vs Gene   Matrix Disease vs Gene

Download TXT File

Show 25 entries

Rank	Target Gene	Disease Num	OMIM	MedGen	Human Phenotype Ontology	Array Experiment	Gene RIF	SNP (Contribution)	References	References LV1	References LV2	Pathway	Therapeutic Target	Max Phase Drug/Exp	Assay Target	Confidence Score Mean	Confidence Score Max	SUM of Confidence Score
1	CA2	3	0	0	0	1	0	0.09	48	4	4	1	Yes	4	CHEMBL205	0.38	0.38	1.07
2	PTGFR	2	0	0	0	0	3	0.19	47	16	16	0	Yes	4	CHEMBL1987	0.46	0.47	0.92
3	MYOC	5	0	0	0	1	100	5.27	537	427	413	0	Yes	0		0.18	0.26	0.89
4	CYP1B1	5	1	0	0	0	84	3.42	225	186	184	0	Yes	0	CHEMBL4878	0.18	0.44	0.88
5	ADRB1	2	0	0	0	0	1	0.0	0	0	0	0	Yes	4	CHEMBL213	0.42	0.42	0.83
6	CA4	2	0	0	0	0	0	0.14	10	2	2	1	Yes	4	CHEMBL3729	0.4	0.46	0.79
7	CA12	2	0	0	0	1	1	0.0	1	1	1	1	Yes	4	CHEMBL3242	0.4	0.46	0.79
8	ADRB2	2	0	0	0	0	2	0.05	2	2	2	0	Yes	4	CHEMBL210	0.39	0.45	0.78
9	VEGFA	4	0	0	0	2	3	0.04	318	38	37	0	Yes	3	CHEMBL1783	0.18	0.4	0.73
10	OPTN	4	0	0	0	1	31	1.76	155	123	122	0		0		0.18	0.26	0.73
11	WDR36	5	0	0	0	0	17	4.16	17	9	9	0		0		0.14	0.21	0.72
12	CA1	2	0	0	0	0	0	0.0	23	0	0	1	Yes	4	CHEMBL261	0.34	0.35	0.68
13	LTBP2	3	0	0	0	1	7	1.77	19	13	13	0		0		0.19	0.23	0.56
14	FOXC1	3	0	0	0	0	11	0.77	48	29	29	0		0		0.18	0.22	0.54
15	LOXL1	5	0	0	0	1	42	3.0	70	47	47	1		0		0.1	0.15	0.52
16	OPA1	4	0	0	0	0	9	0.26	257	243	228	0		0		0.13	0.17	0.51
17	MTHFR	4	0	0	0	2	16	0.86	29	20	20	0		0		0.13	0.14	0.51
18	NOS3	4	0	0	0	0	10	1.2	25	20	20	0		0	CHEMBL4803	0.12	0.13	0.49
19	ASB10	3	0	0	0	0	3	5.41	5	5	5	0		0		0.16	0.2	0.48
20	ATOH7	4	0	0	0	0	5	0.22	2	2	2	0		0		0.12	0.2	0.48
21	CAV1	4	0	0	0	2	12	1.59	24	13	13	0		0	CHEMBL3808270	0.12	0.14	0.48
22	VAV2	5	0	0	0	0	3	0.21	2	2	2	0		0		0.1	0.12	0.48

Showing 1 to 25 of 4,475 entries

Previous 1 2 3 4 5 ... 179 Next





## Exploration



From Disease list

Target Explorer

GO Explorer  
-Biological Process-

From Gene list

Disease Explorer

Expression Explorer

From Chemical structure & Gene list

Mechanism Explorer

Drug Repurposing  
Procedure

LSKB



# Disease Explorer



## Interest Genes (Predicted targets)

[MS4A1](#)  
[IL1B](#)  
[IL6R](#)  
[IL17A](#)  
[OPRK1](#)  
[TNF](#)

Literature mining

Database

Gene Expression

SNPs Information

Indication

Clinical Trial

Pathway

- MS4A1
- MS4A2
- IL1B
- IL6R
- IL17A
- OPRK1

Send Disease ID : 3823 / 3823 [Download TXT File](#)

<input checked="" type="checkbox"/>	Rank	Disease	Gene Num	OMIM	MedGen	Human Phenotype Ontology	GEO	Gene RIF	SNP (rsID)	References	References LV1	References LV2	Pathway	Therapeutic Target	Max Phase Drug/Exp	Confidence Score Mean	Confidence Score Max	SUM of Confidence Score
<input checked="" type="checkbox"/>	1	Rheumatoid Arthritis	6	0	0	0	14	163	24	5873	1768	1757	22	Yes	4	0.411	0.592	2.467
<input checked="" type="checkbox"/>	2	Immune System Diseases	7	0	0	0	0	1	7	243	21	21	12	Yes	4	0.3	0.5	2.1
<input checked="" type="checkbox"/>	3	Diabetes mellitus type 1	6	0	0	0	3	36	8	385	83	82	23	Yes	3	0.321	0.483	1.925
<input checked="" type="checkbox"/>	4	Behcet Syndrome	6	0	0	0	0	26	2	227	22	22	0	Yes	3	0.279	0.383	1.675
<input checked="" type="checkbox"/>	5	Spondyloarthritis 1	6	0	0	0	3	39	13	1906	517	511	2	Yes	4	0.258	0.525	1.55
<input checked="" type="checkbox"/>	6	Osteoarthritis	6	0	0	0	5	56	7	804	122	119	0	Yes	3	0.254	0.392	1.525
<input checked="" type="checkbox"/>	7	Leukemia	7	0	0	0	40	50	3	1438	327	323	34	Yes	3	0.211	0.417	1.475
<input checked="" type="checkbox"/>	8	Ankylosing Spondylitis	5	0	0	0	2	24	5	1154	302	298	0	Yes	4	0.293	0.508	1.467
<input checked="" type="checkbox"/>	9	Crohn disease-associated growth failure	6	0	0	0	14	41	16	1594	218	215	9	Yes	4	0.243	0.55	1.458
<input checked="" type="checkbox"/>	10	Psoriasis	7	0	0	0	14	66	10	2035	669	664	9	Yes	4	0.207	0.533	1.45
<input checked="" type="checkbox"/>	11	Psoriasis 1	7	0	0	0	14	65	10	2035	669	664	9	Yes	4	0.207	0.533	1.45
<input checked="" type="checkbox"/>	12	Psoriasis 7	7	0	0	0	14	65	10	2035	669	664	9	Yes	4	0.207	0.533	1.45
<input checked="" type="checkbox"/>	13	Psoriasis 11	7	0	0	0	14	65	10	2035	669	664	9	Yes	4	0.207	0.533	1.45
<input checked="" type="checkbox"/>	14	Psoriasis 13	7	0	0	0	14	65	10	2035	669	664	9	Yes	4	0.207	0.533	1.45
<input checked="" type="checkbox"/>	15	Psoriasis 2	7	0	0	0	14	65	10	2035	669	664	9	Yes	4	0.207	0.533	1.45





## Exploration



From Disease list

Target Explorer

GO Explorer  
-Biological Process-

From Gene list

Disease Explorer

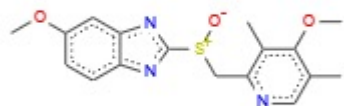
Expression Explorer

From Chemical structure & Gene list

Mechanism Explorer

Drug Repurposing  
Procedure

# Mechanism Explorer



Target  
Prediction

Expand Similar protein

化合物により、起こった事  
象の原因を探る

Expression Experiment file  
With Gene ID

24159  
24309  
24330  
24385  
24411  
24413  
...

Summary

Predicted targets

Similar proteins

GO summary

Pathway summary

OMIM summary

Expressions with annotation

Expressions vs prediction

Additional file

FILTERS

See also other tables, to check filter condition...

EXPRESSION

not

(NULL)

Delete

Welch\_t\_test-FDR\_Welch\_t\_test-FDR

≤

0.05

Delete

ENTRY\_NAME

contains

Any string

Add Filter

Clear

Filter operation mode

AND

OR

<

<

>

>

📄

📄

📄

UniProt Entry Name

Action

---

Page 1 of 3

Total 1226 records

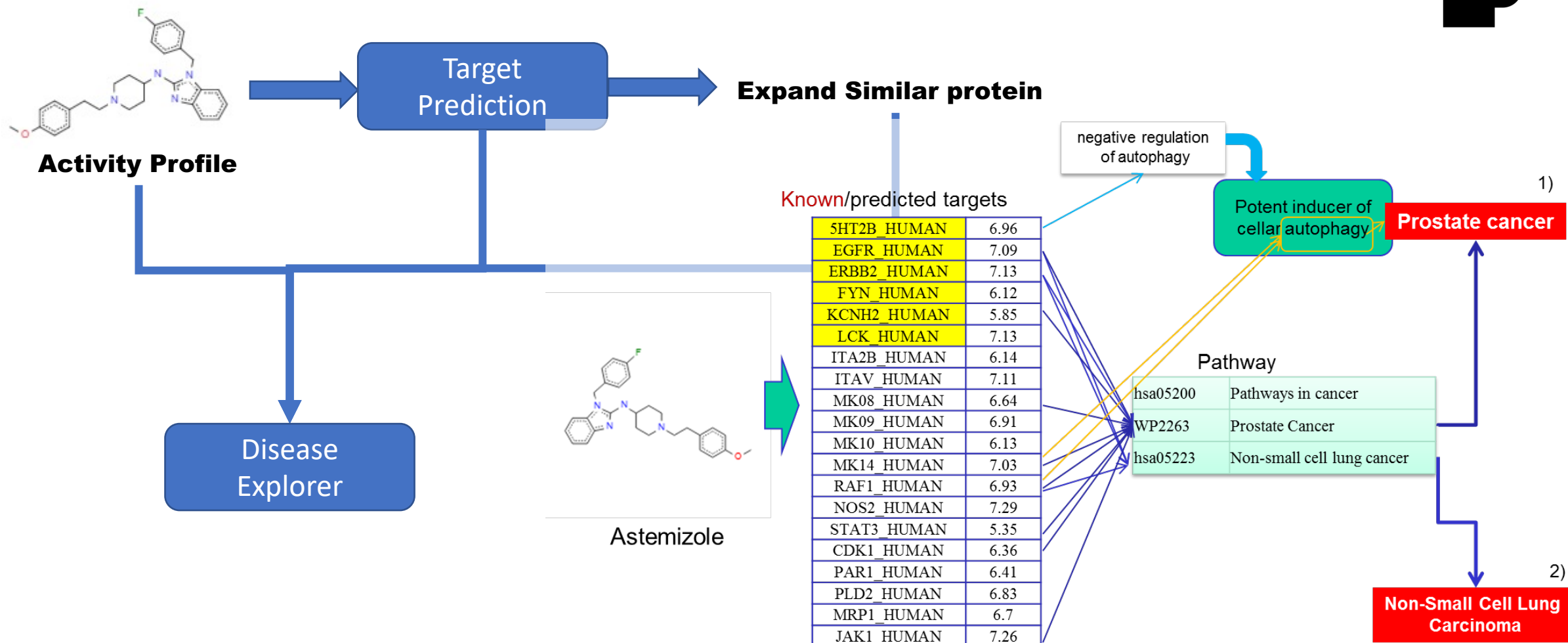
58 records, filter passed

	ENTRY NAME	NAME	PREDICTED	SIMILAR	EXPRESSION	Log2Ratio(S2/S1)	Log2Ratio(S2/S1)	Welch t test-p	Welch t test-p	Welch t test-FDR	Welch t test-FDR
<input type="checkbox"/>	<a href="#">MRP1_HUMAN</a>	Multidrug resistance-associated protein 1		YES	YES	0.133595898		9.35E-04		0.008402893	
<input type="checkbox"/>	<a href="#">MMP11_HUMAN</a>	Stromelysin-3		YES	YES	-0.404556852		0.002175704		0.015769683	
<input type="checkbox"/>	<a href="#">CP3A7_HUMAN</a>	Cytochrome P450 3A7		YES	YES	0.141544364		0.010024086		0.046117305	
<input type="checkbox"/>	<a href="#">NR1H4_HUMAN</a>	Bile acid receptor		YES	YES	-0.148589815		0.006954139		0.036109377	
<input type="checkbox"/>	<a href="#">PHOP1_HUMAN</a>	Phosphoethanolamine/phosphocholine phosphatase	YES		YES	0.449497769		2.65E-12		4.70E-10	
<input type="checkbox"/>	<a href="#">CATL2_HUMAN</a>	Cathepsin L2		YES	YES	0.225004207		0.00102569		0.009093238	
<input type="checkbox"/>	<a href="#">MRP1_RAT</a>	Multidrug resistance-associated protein 1		YES	YES	0.133595898		9.35E-04		0.008402893	
<input type="checkbox"/>	<a href="#">CATL1_RAT</a>	Cathepsin L1		YES	YES	0.225004207		0.00102569		0.009093238	
<input type="checkbox"/>	<a href="#">STAT1_HUMAN</a>	Signal transducer and activator of transcription 1-alpha/beta		YES	YES	0.07416003		0.00850474		0.041271291	
<input type="checkbox"/>	<a href="#">GLP1R_RAT</a>	Glucagon-like peptide 1 receptor		YES	YES	-0.123849715		1.36E-04		0.00187269	
<input type="checkbox"/>	<a href="#">CAH14_HUMAN</a>	Carbonic anhydrase 14		YES	YES	-0.51539403		0.007698595		0.038635491	
<input type="checkbox"/>	<a href="#">FABP1_HUMAN</a>	Fatty acid-binding protein, intestinal		YES	YES	-0.20605397		3.22E-04		0.003762708	
<input type="checkbox"/>	<a href="#">CAH3_HUMAN</a>	Carbonic anhydrase 3		YES	YES	-0.556699937		7.78E-05		0.001187427	
<input type="checkbox"/>	<a href="#">ADRB2_HUMAN</a>	Beta-2 adrenergic receptor		YES	YES	-0.320416739		0.001998162		0.014811138	
<input type="checkbox"/>	<a href="#">DHB7_HUMAN</a>	3-keto-steroid reductase		YES	YES	0.254163992		7.11E-06		1.57E-04	
<input type="checkbox"/>	<a href="#">MRP4_HUMAN</a>	Multidrug resistance-associated protein 4		YES	YES	0.204928218		3.92E-04		0.004411943	
<input type="checkbox"/>	<a href="#">B2CL1_RAT</a>	Bcl-2-like protein 1		YES	YES	0.37580736		5.16E-06		1.23E-04	
<input type="checkbox"/>	<a href="#">AA2BR_HUMAN</a>	Adenosine receptor A2b		YES	YES	-0.240745765		0.00315859		0.020849636	

Top



# Drug Repurposing Procedure

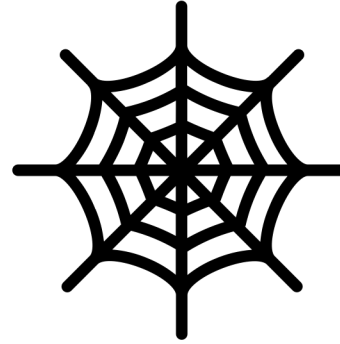


Target Prediction & map the targets to pathway by LSKB.

- 1) Drug repurposing from an academic perspective  
DOI: 10.1016/j.ddstr.2011.10.002
- 2) Repurposing Cationic Amphiphilic Antihistamines for Cancer Treatment  
<http://dx.doi.org/10.1016/j.ebiom.2016.06.013>



# LSKB Version 7



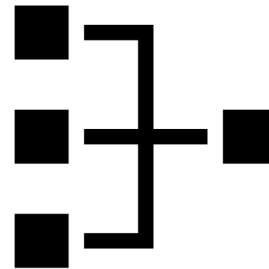
**Content**



**Analysis**

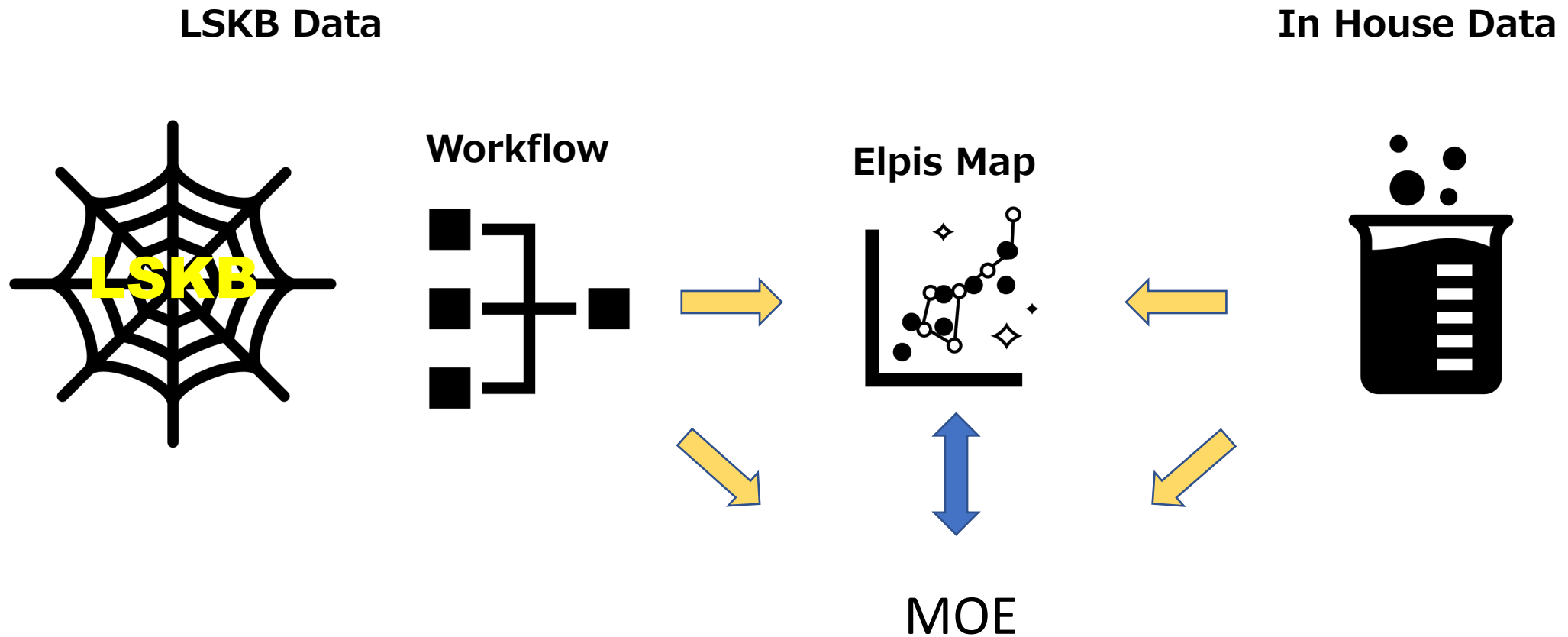


**Workflow**





# Elpis Map : Decision Support Tool





# LSKB Ver.7

1. GUI を刷新
  - Top Menu の再構成
  - 新しいテーブルの導入
2. Deep Learning により 抽出した MoA の追加
  - ターゲット/Ligand/疾患との関連性の拡大
3. GO-MoAの Update
  - 新規 MoAの追加と 不要なアノテーションの整理
  - 疾患から 新しい ターゲット、ターゲット遺伝子から新たな適応症、化合物から 新たな適応症の探索
4. Workflowの Update
  - 目的とするデータ収集を可能に
  - 処理の自動化
  - LSKBのデータベースを最大限に活用できる探索ツール
5. Elpis Mapの搭載
  - LSKB 搭載データと 自社データの比較が容易
  - 意思決定を強力にサポート

etc



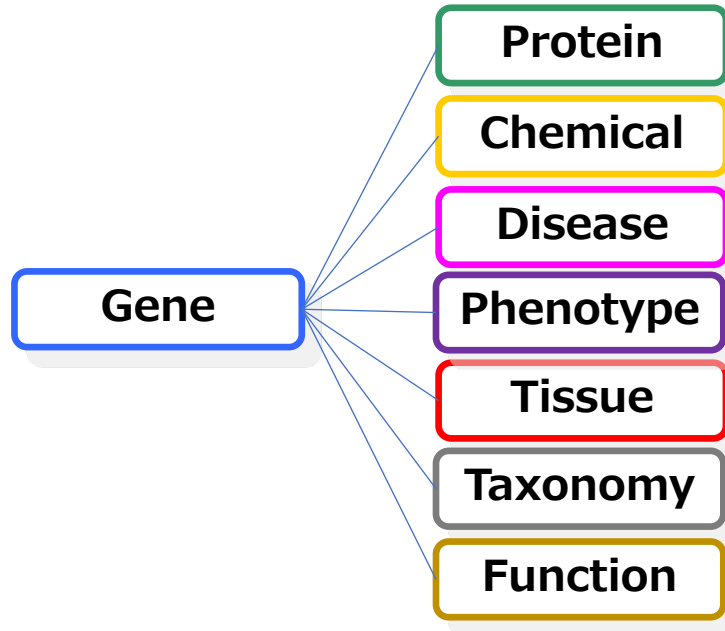
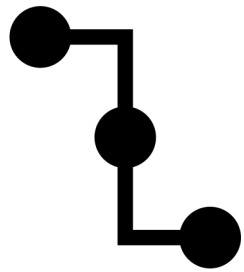
ご清聴ありがとうございました。

ご質問は  
[support@lskb.jp](mailto:support@lskb.jp)  
までお願いいたします。

<https://www.lskb.jp/>

# LSKB Interaction

Interaction



## 1) Text-Mining

- 20 years of PubMed literature (3 levels)
- Clinical Trial
- Assay Description

## 2) Assay Data

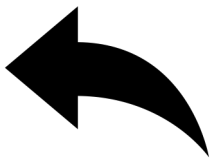
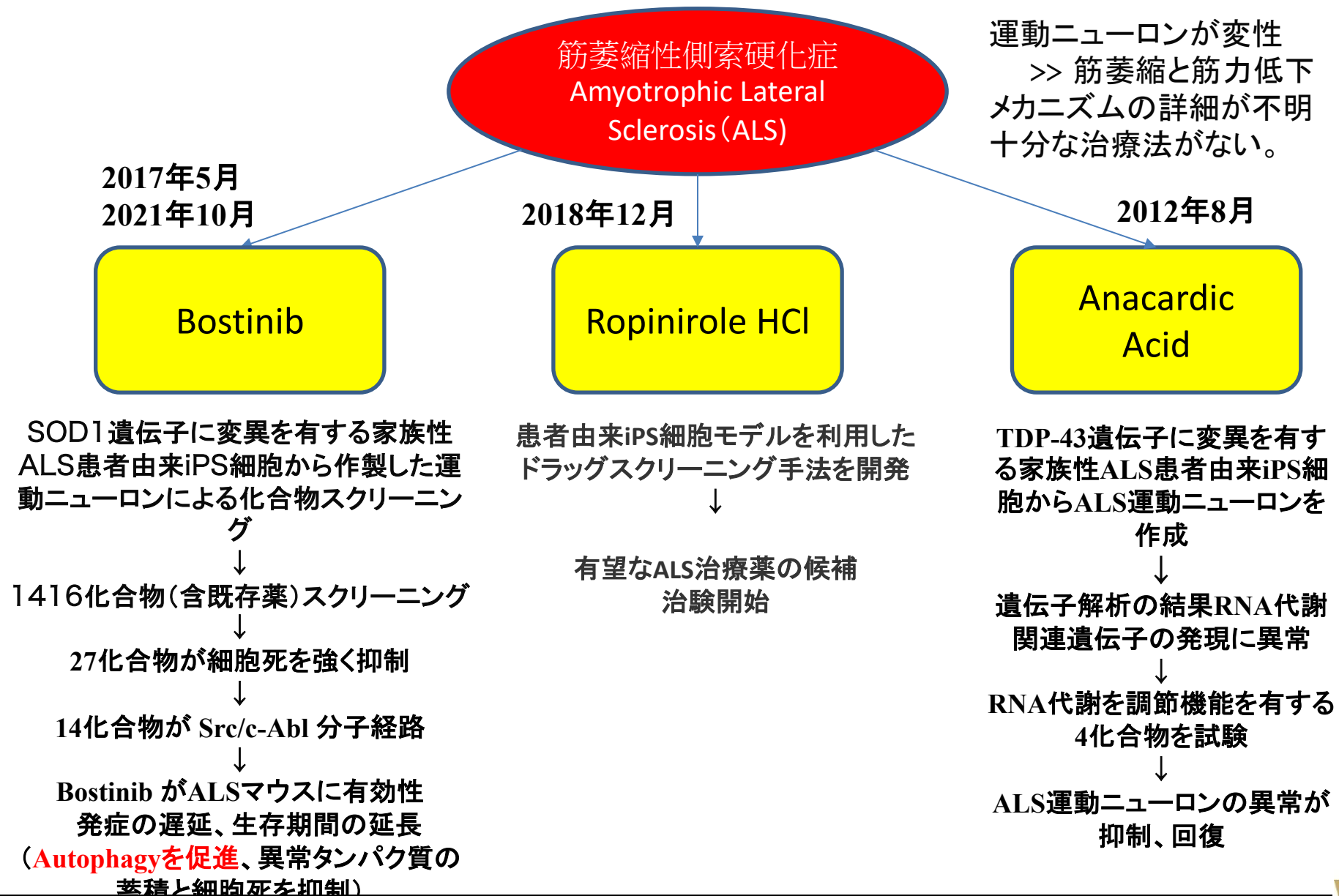
- Target Gene/Protein - Chemical
- Activities Endpoint
- Mode : Inhibition agonism/antagonism
- Expression: up/down regulation
- GWAS

## 3) Curated Annotation

- Disease Target
- Gene Ontology
- Pathway

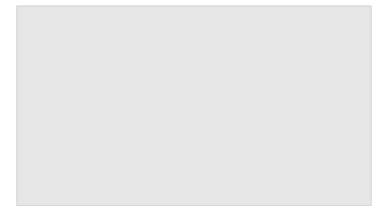
## 4) AI Curated Annotation

- Mechanism of Action
- Gene RIF



# 結果

- GeneRIFの搭載文章 1,486,580件のうち、162,988件（Gene 19190件）をBERTのMNLI タスクを実行した。
  - 遺伝子の生物種の内訳は Human (10523)、Mouse (5323)、Rat (2671)、Zebrafish (273) であった。
- 9 種類の仮説を自動生成（計 1,499,079 件）した。その内訳は
  - “遺伝子と疾患との関連” 仮説で取得された 疾患は 10502件、遺伝子は18135件であった。
  - “遺伝子発現の負の変動と疾患との関連性” 仮説で取得された 疾患は 3714件、遺伝子は 5673件であった。
- 各仮説の判定結果（Entailment）をランダムに20-25件 前提文章と比較検証したところ、精度は 80-100%であった。
- 実際の利用場面（Entailment）の MNLIのタスクの精度は F1-Score : 0.832であった。



# 正しく判断できた事例

仮説：「遺伝子」の負の発現変動が「疾患」に起因する。

GENE	DISEASE	GENERIF
slc41a1	nephronophthisis	<p><b>Knockdown</b> of slc41a1 expression in zebrafish resulted in ventral body curvature, hydrocephalus, and cystic kidneys, similar to the effects of knocking down other nephronophthisis genes.</p> <p>ゼブラフィッシュにおけるslc41a1発現のノックダウンは、他の<b>髄質性嚢胞腎</b> 遺伝子をノックダウンする効果と同様に、腹側体の湾曲、水頭症、および嚢胞性腎をもたらしました。</p>
ALDH1	carcinogenesis	<p><b>loss of</b> ALDH1A1 expression is suggested to promote carcinogenesis especially in the smoking-related lung adenocarcinomas</p> <p>ALDH1A1発現の喪失は、特に喫煙関連の肺腺癌において<b>発癌</b>を促進することが示唆されています。</p>
ALDH1	metastatic renal cell carcinoma	<p><b>Low</b> ALDH1 expression is associated with high stage and metastatic renal cell carcinoma.</p> <p>低ALDH1発現は、ハイステージおよび<b>転移性腎細胞癌</b>と関連しています。</p>
BIN1	cutaneous T-cell lymphoma	<p><b>Reduced</b> BIN1 expression is associated with cutaneous T-cell lymphoma.</p> <p>BIN1発現の低下は、<b>皮膚T細胞リンパ腫</b>に関連しています。</p>
Bcl-2	severe congenital neutropenia	<p>a <b>selective decrease</b> in Bcl-2 expression was seen in myeloid progenitor cells of pateients with Kostmann syndrome: severe congenital neutropenia.</p> <p>コストマン症候群患者：<b>重症先天性好中球減少症</b>の骨髓前駆細胞でBcl-2発現の選択的減少が見られました。</p>