KNApSAcKを用いた植物の 効能メカニズム解明のための基盤構築

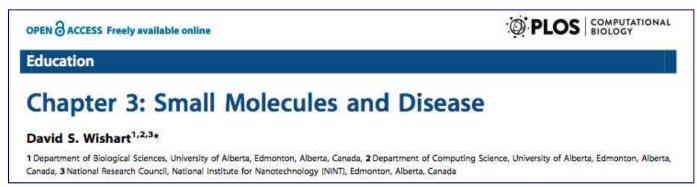
京都大学 化学研究所 西村陽介

H26.03.02

統合データ解析トライアル・終了報告会

KNApSAcK

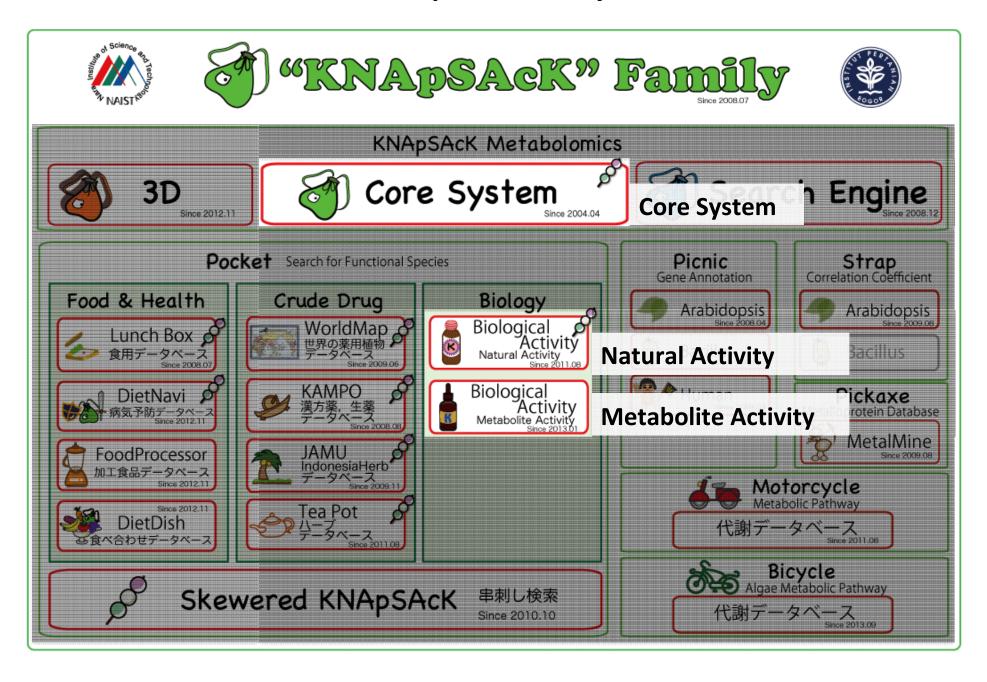
- 生物種と代謝産物の関係を体系化することを目的とした、二次代謝産物データベース
- ・ 収載数が多い
 - 約2万生物種、約5万代謝産物
 - 約10万生物種-代謝産物ペア
- 植物代謝産物のデータベースとして総説で紹介されている。



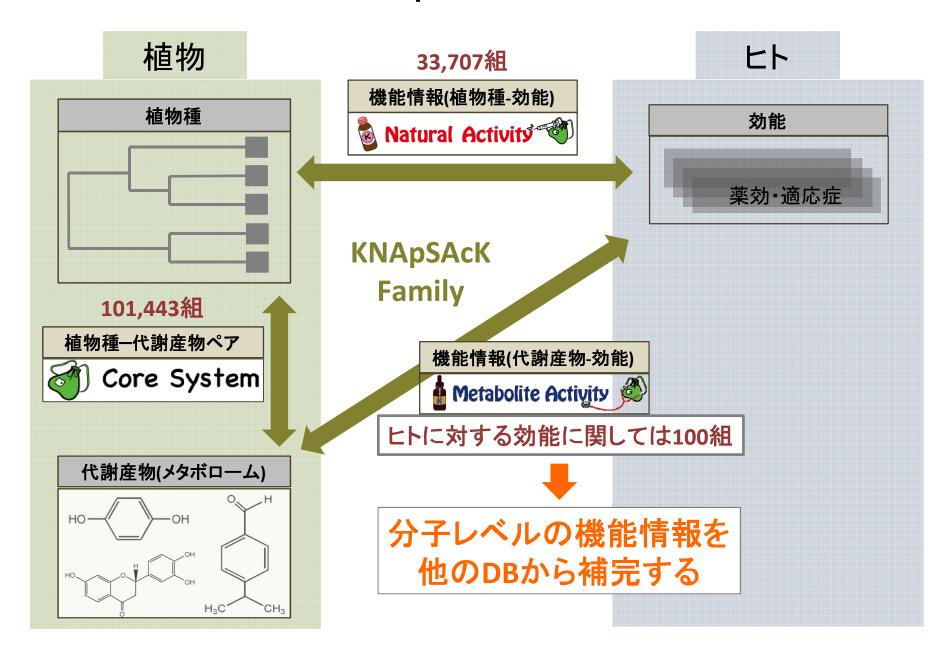
植物のヒトに対する効能メカニズムの解明

- 植物は古くから解毒、鎮痛、病気の治癒等の為に利用され、日々 の食事においても人々は多様な植物を摂取している。
- 植物代謝産物から数多くの医薬品の有効成分が発見されている
 - モルヒネ(鎮痛剤)
 - パクリタキセル(抗がん剤)
 - ジギトキシン(強心剤)
- 近年の大規模解析技術の発達に伴って、ヒトタンパク質と代謝産物の相互作用情報が蓄積されている(分子レベルの機能情報)。
- 分子レベルの機能情報を利用して、植物のヒトに対する効能メカニズムを解明することにより、植物の摂取法や、生薬としての利用についての理解が深まる。

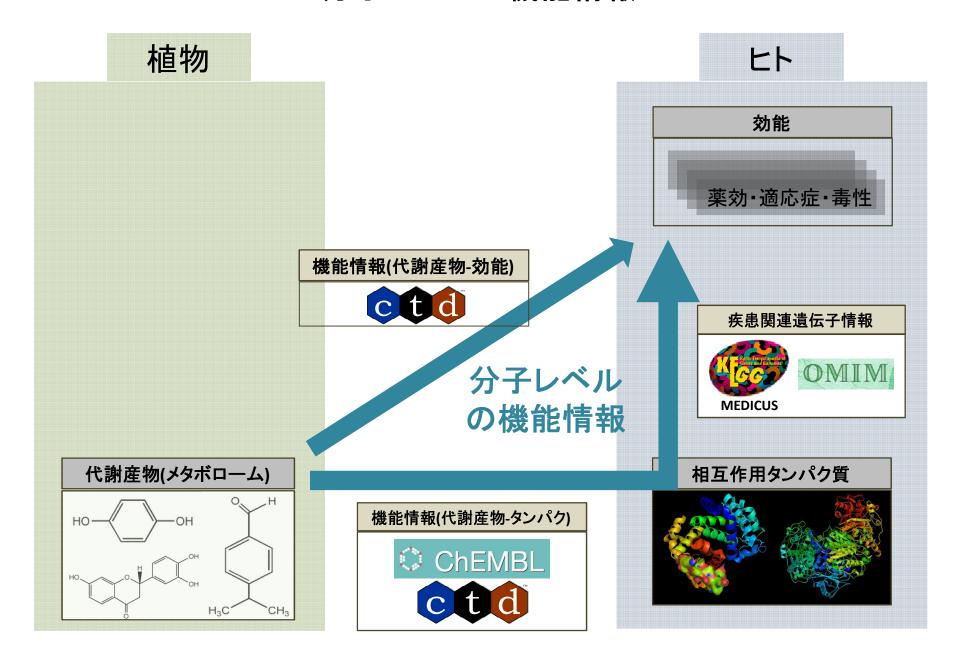
KNApSAcK Family



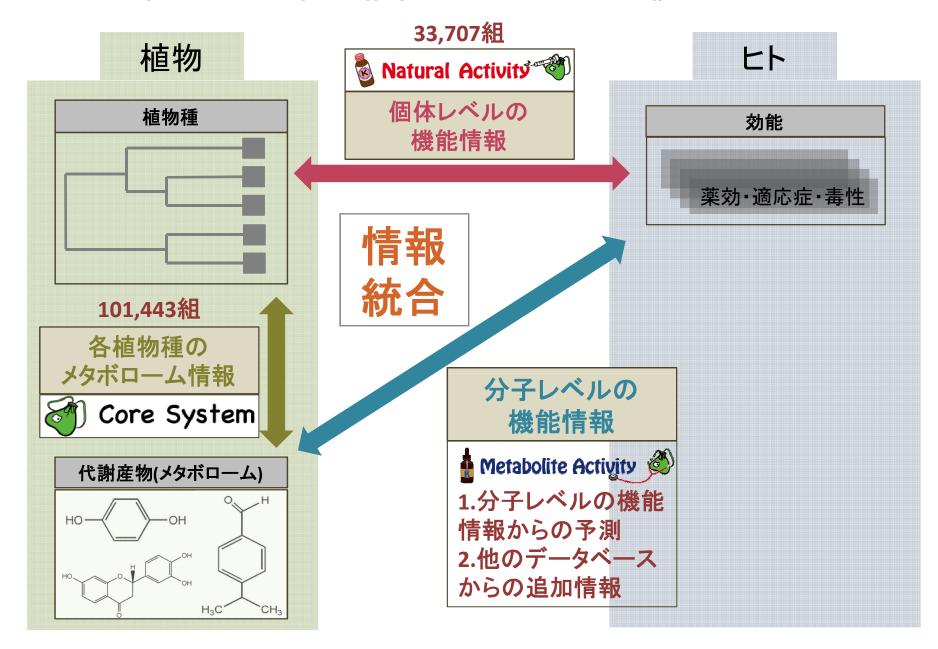
KNApSAcKの情報



分子レベルの機能情報



個体レベルの機能情報の予測のための統合スキーム



PCIDBの作成

PCIDB

Metabolites -

Species *

Metabolite Clusters -

PCIDB - PhytoChemical Interactions DB

植物は解毒、鎮痛、病気の治癒等のために各地で利用されている。近年大規模解析技術の発達に伴って、ヒトタンパク質と代謝産物の相互作用情報が蓄積されており、この情報を活用することにより、様々な病気に対する治療への応用が期待される。

PCIDBは、植物を中心とする代謝産物とその由来植物種の関係が10万レコード以上収載されているKNApSAcK Databaseと、分子レベルの機能情報が取得出来る ChEMBLや、CTD等のデータベースを用いて、これらの情報を統合することにより、植物代謝産物の相互作用タンパク質情報や、ヒトに対する効能情報を取得するためのデータベースである。

Search KNApSAcK Metabolite

Search KNApSAcK Species

by ID

e.g.) C000003

Search

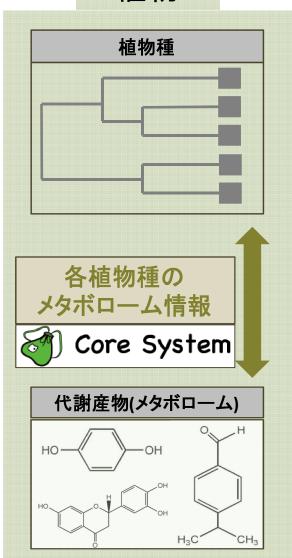
Search By Name

e.g.) Vitis

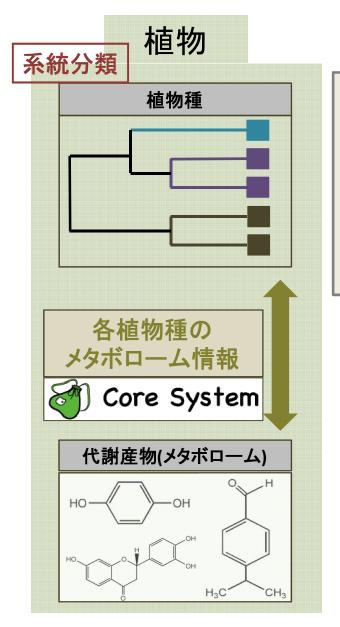
Search

KNApSAcK coreのデータ



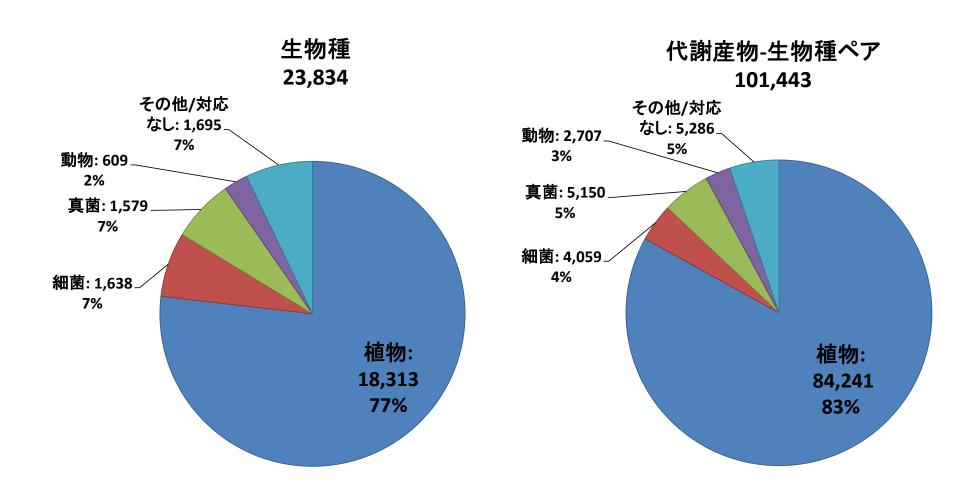


KNApSAcK生物種の系統分類



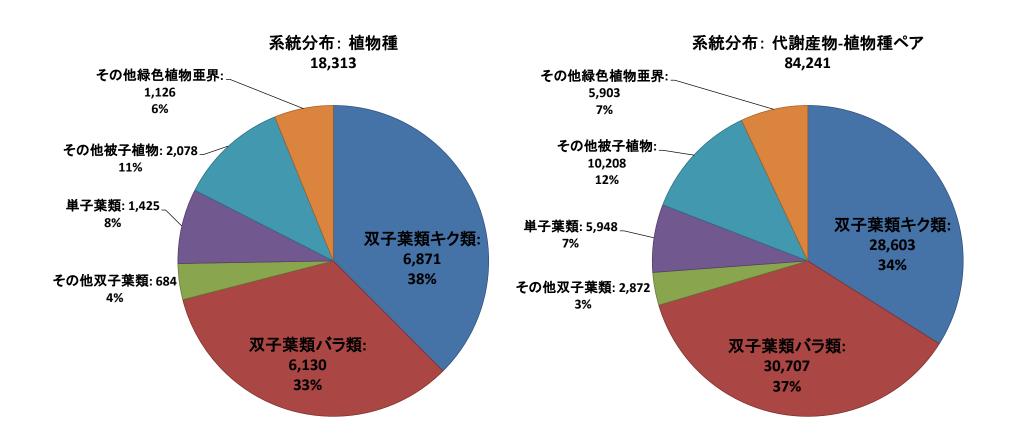
- NCBI taxonomyへの対応
- 生物種系統分類
 - (super)kingdom
 - plant taxonomic class
 - plant family

系統分類結果: (super)kingdom



KNApSAcKは大部分が植物に関するデータであることを確認

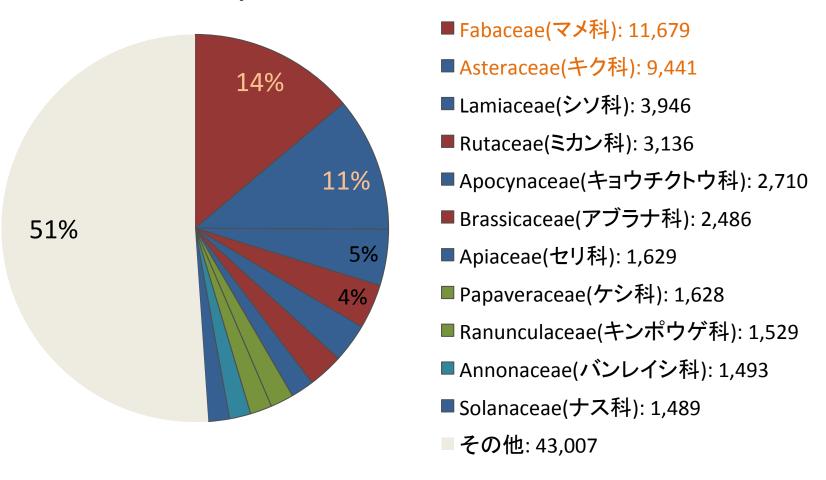
系統分類結果: plant taxonomic class



「双子葉類キク類」「双子葉類バラ類」「その他」に3分される

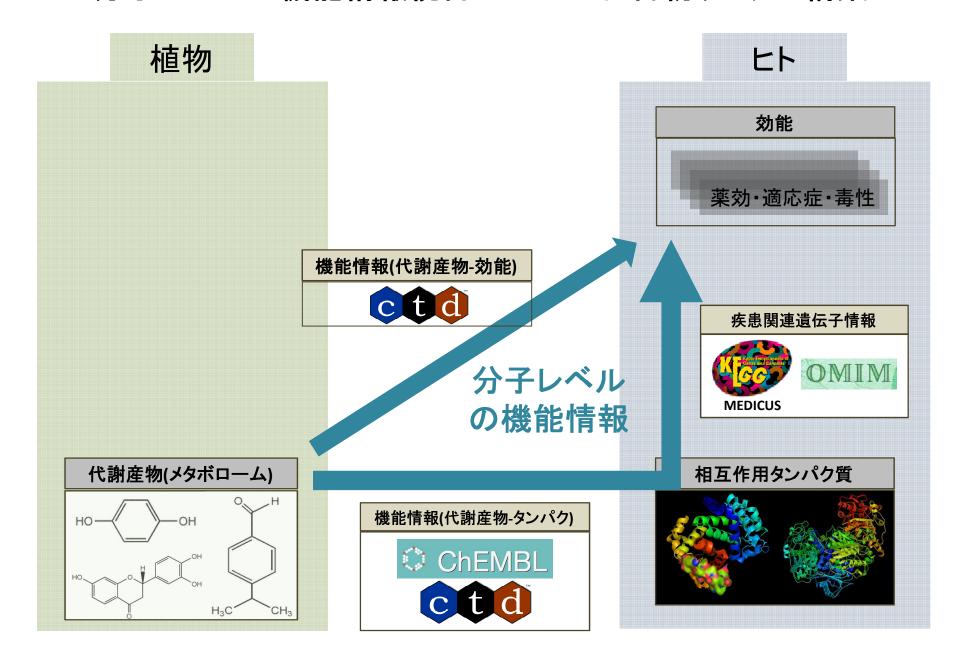
系統分類結果: plant family

Top11/430: 代謝産物-植物種ペア



大きいfamilyは「マメ科」「キク科」など

分子レベルの機能情報統合のための化合物リンクの構築



KNApSAcKとChEMBL, CTDとの化合物リンク構築

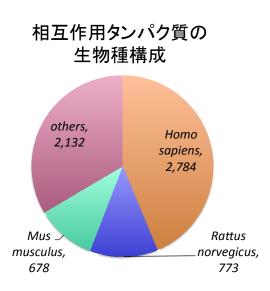
- Standard InChI (ChEMBL)
 - IUPACによる、標準的な分子構造記述法
 - 立体異性体を考慮しない化合物リンクも構築した (Main Layer)

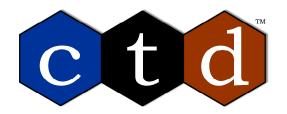
- CAS登録番号 (CTD)
 - 化学物質IDのデファクトスタンダード



Release 17 (Aug. 2013)

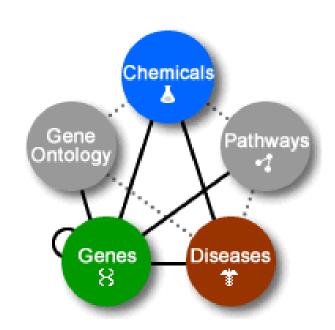
- 医薬品及び開発化合物のデータベース
- ・ ターゲット(タンパク質等)に関するアッセイ情報
 - 文献からマニュアル抽出 + PubChem からの情報
 - 専門家によるキュレーション
- >73万アッセイ、>1200万活性情報、>130万化合物
- 6,367 相互作用タンパク質
 - ヒトでは2,784個



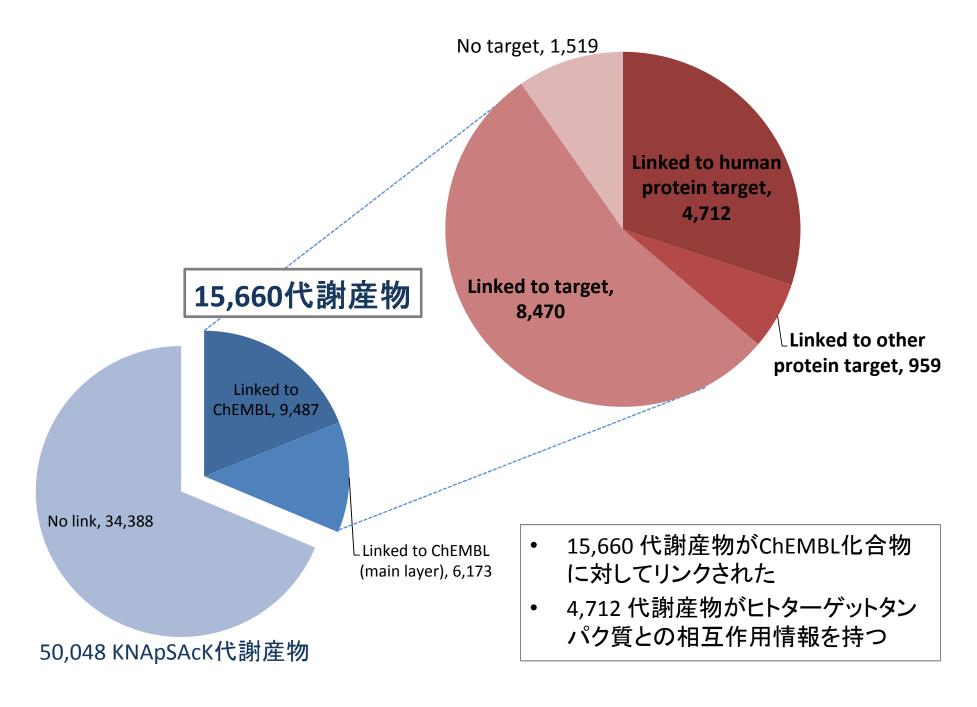


Ver. Oct. 2013

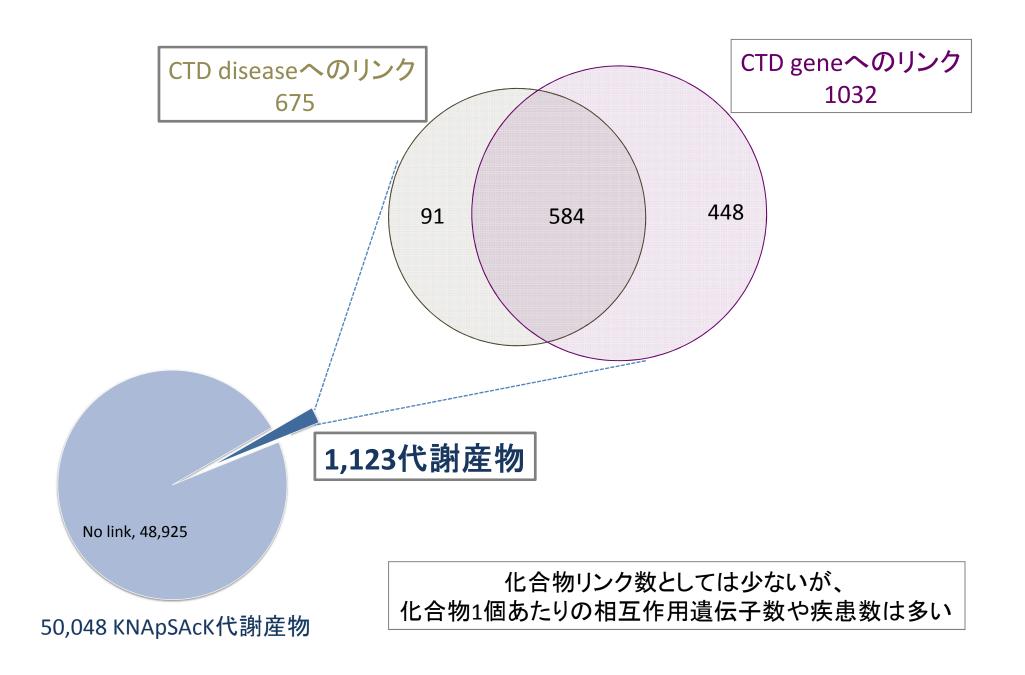
- 環境物質が人体に与える影響を文献から抽出
- C(化合物)-G(遺伝子)相互作用情報
 - 直接の相互作用及び、結果としてのリン酸化などの間接的な作用を含む
 - 53個の相互作用述語表現
 - C (MeSH Term) G (Entrez gene id)
 - >88万レコード (curated)
- C(化合物) D(疾患)相互作用情報
 - therapeutic か marker/mechanism に分類される
 - C (MeSH Term) D (MeSH Term)
 - >18万レコード (curated)



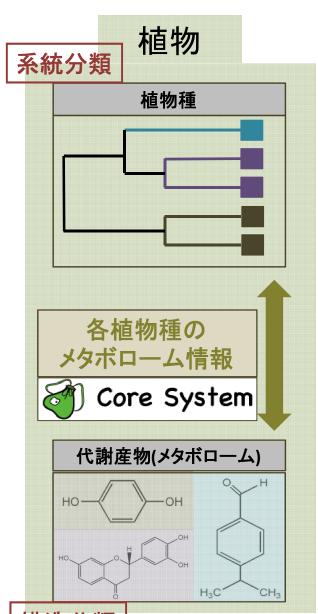
KNApSAcKとChEMBLの化合物リンク (Standard InChl)



KNApSAcKとCTDの化合物リンク(CAS登録番号)



代謝産物の構造分類



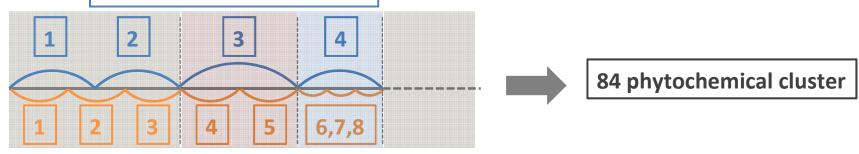
- 母核を用いた化合物分類
 - KEGG BRITEの分類を利用
 - br08003, Phytochemical Compounds
- 部分構造を利用した分類 (KCF-S cluster)
 - KCF-S (GIW2013, Kotera et al.,)
 - ・ 化合物の部分構造ディスクリプタ
- ・ これらの共通要素分類 (phytochemical cluster)
 - 母核分類と部分構造分類の共通要素を取り出す

構造分類

phytochemical clusters

- KEGG BRITE 母核分類 (br08003, Phytochemical Compounds)
 - A, B, C, Dの4階層
 - C階層で分類: 163個
- KCF-S cluster
 - 8,955 clusters (3,245 singletons)
 - 10,000 descriptors
 - similarity / clustering cutoff value = 0.7 / 0.5
- phytochemical cluster (これらの共通要素分類)
 - 84 clusters
 - 161母核分類と994 KCF-S clustersから構成
 - 植物代謝産物のうち、51%が含まれる
 - ChEMBLの相互作用情報については、73%が含まれる

KEGG BRITE 母核分類: 161



部分構造分類:994

phytochemical clusters

PCIDB

Metabolites ▼

Species *

Metabolite Clusters ▼

help

84 phytochemical clusters

No	0.	# of metabolites		# of KCF-S clusters	Classes of Phytochemical Compounds in KEGG BRITE		
1		176		63	Pyrrolidine alkaloids Tropane alkaloids Piperidine alkaloids Pyridine alkaloids Acetate derived alkaloids		
2	131			20	Pyrrolizidine alkaloids Indolizidine alkaloids		
3	3 175		16	Quinolizidine alkaloids			
4		1605]	95	Isoquinoline alkaloids Indole alkaloids Pyrroloindole alkaloids		
	hemica	I			Others		
	ster 選択	2		2	Thiazole alkaloids		
6		186		22	Tyramine derivatives Caffeate derivatives Coniferyl alcohol derivatives Paracoumaryl alcohol derivatives Sinapate derivatives Others		
7		243		43	Acridone alkaloids Quinazoline alkaloids Quinoline alkaloids		
8		2		2	Imidazole alkaloids		
9		9		2	Phenylalanine derived alkaloids		
10)	394		20	Terpenoid alkaloids		
i				T. Control of the con			

各phytochemical clusterの情報

PCIDB Metabolites - Species - Metabolite Clusters - help

phytochemical cluster No. 3 (168 metabolites)

Index
Plant Species
BR08003 category (1)
BR08003 list (34)
KCF-S cluster (16)
Metabolite list (168)
Human Protein / Gene in interactions
Related Diseases

各項目の リストとリンク

KEGG BRITE母核分類 (Quinolizidine alkaloids)

Plant Species

Cumulative plant class count in this cluster

植物クラス分類 (ほぼバラ類)

class name	count	
rosids	1016	
Magnoliophyta	29	
asterids	23	
Embryophyta	12	
core eudicotyledons	9	
Spermatophyta	4	

Cumulative plant family count in this cluster (Top 10)

ファミリー分類 (ほぼマメ科)

class name	count
Fabaceae	1012
Berberidaceae	25
Orobanchaceae	16
Lycopodiaceae	12
Cupressaceae	4
Daphniphyllaceae	4
Convolvulaceae	4
Nymphaeaceae	3
Viscaceae	3
Styracaceae	3

BR08003 category (1)

BR08003 Category		# of metabolite		
	Quinolizidine alkaloids	31		

BR08003 list (34)

BR08003 Category	KEGG ID	KNApSAcK ID
Quinolizidine alkaloids	C09856	C00001934
Quinolizidine alkaloids	C09860	C00001935
Quinolizidine alkaloids	C09874	C00001943
Quinolizidine alkaloids	C09945	C00001971
Quinolizidine alkaloids	C10747	C00002204
Quinolizidine alkaloids	C10750	C00002207
Quinolizidine alkaloids	C10751	C00002208
Quinolizidine alkaloids	C10752	C00002209
Quinolizidine alkaloids	C10753	C00002210
Quinolizidine alkaloids	C10754	C00002211
Quinolizidine alkaloids	C10755	C00002212
Quinolizidine alkaloids	C10757	C00002214
Quinolizidine alkaloids	C10758	C00002215
Quinolizidine alkaloids	C10760	C00002216
Quinolizidine alkaloids	C10762	C00002217
Quinolizidine alkaloids	C10763	C00002218
Quinolizidine alkaloids	C10764	C00002219
Quinolizidine alkaloids	C10767	C00002220

KCF-S cluster (16)

KCF-S ID	# of metabolite	
No. 85	49	
No. 384	19	
No. 424	18	
No. 376	17	
No. 555	15	
No. 1023	9	
No. 1211	8	
No. 1838	6	
No. 1636	6	
No. 2032	5	
No. 1268	5	
No. 2957	4	
No. 3172	3	
No. 5313	2	
No. 6163	1	
No. 6875	1	

関連するKCF-S cluster 情報

各phytochemical clusterの情報(代謝産物リスト)

	te list (168)	化合物リ		関連疾患	CF-S cluster		
代謝) KNApSAcK ID	在物情報 name	ChEMBL link	CTD link	# of proteins in ChEMBL interaction / related OMIM / related KEGG DISEASE	# of genes in CTD interaction / related diseases	KCF-S cluster	figure
C00036283 d	Dimethamine					No. 85	H H H H H
C00032558 &	11,12-Seco-12,13- didehydromultiflorine		C058266			No. 85	
C00007720 @	N-Acetylcytisine	CHEMBL1513538 CHEMBL1553838		2/3/7		No. 85	HO,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
C00007713 @	(-)-N-(3-Oxobutyl)cytisine	CHEMBL190851		4/3/1		No. 85	H N CH ₅
C00026280 d	Oxymatrine / Ammothamnine / Matrine N-oxide / (+)-Matrine N-oxide / Matrine 1beta-oxide	CHEMBL458337 CHEMBL1358502	C037573	3/11/6	0/4	No. 85	N HIMM H
C00007712 &	Rhombifoline / (-)-Rhombifoline	CHEMBL1397610	C119591			No. 85	N CH ₃
C00007704 d	Cytisine / (-)-Cytisine	CHEMBL47039 CHEMBL497939 CHEMBL1628606	C004712	27 / 36 / 33	6/1	No. 85	H N OH

各phytochemical clusterの情報

PCIDB Metabolites - Species - Metabolite Clusters - help

phytochemical cluster No. 3 (168 metabolites)

Index
Plant Species
BR08003 category (1)
BR08003 list (34)
KCF-S cluster (16)
Metabolite list (168)
Human Protein / Gene in interactions
Related Diseases

各項目の リストとリンク

KEGG BRITE母核分類 (Quinolizidine alkaloids)

Plant Species

Cumulative plant class count in this cluster

植物クラス分類 (ほぼバラ類)

class name	count
rosids	1016
Magnoliophyta	29
asterids	23
Embryophyta	12
core eudicotyledons	9
Spermatophyta	4

Cumulative plant family count in this cluster (Top 10)

ファミリー分類 (ほぼマメ科)

class name	count
Fabaceae	1012
Berberidaceae	25
Orobanchaceae	16
Lycopodiaceae	12
Cupressaceae	4
Daphniphyllaceae	4
Convolvulaceae	4
Nymphaeaceae	3
Viscaceae	3
Styracaceae	3

BR08003 category (1)

BR08003 Category	# of metabolite		
Quinolizidine alkaloids	31		

BR08003 list (34)

BR08003 Category	KEGG ID	KNApSAcK ID
Quinolizidine alkaloids	C09856	C00001934
Quinolizidine alkaloids	C09860	C00001935
Quinolizidine alkaloids	C09874	C00001943
Quinolizidine alkaloids	C09945	C00001971
Quinolizidine alkaloids	C10747	C00002204
Quinolizidine alkaloids	C10750	C00002207
Quinolizidine alkaloids	C10751	C00002208
Quinolizidine alkaloids	C10752	C00002209
Quinolizidine alkaloids	C10753	C00002210
Quinolizidine alkaloids	C10754	C00002211
Quinolizidine alkaloids	C10755	C00002212
Quinolizidine alkaloids	C10757	C00002214
Quinolizidine alkaloids	C10758	C00002215
Quinolizidine alkaloids	C10760	C00002216
Quinolizidine alkaloids	C10762	C00002217
Quinolizidine alkaloids	C10763	C00002218
Quinolizidine alkaloids	C10764	C00002219
Quinolizidine alkaloids	C10767	C00002220

KCF-S cluster (16)

KCF-S ID	# of metabolite	
No. 85	49	
No. 384	19	
No. 424	18	
No. 376	17	
No. 555	15	
No. 1023	9	
No. 1211	8	
No. 1838	6	
No. 1636	6	
No. 2032	5	
No. 1268	5	
No. 2957	4	
No. 3172	3	
No. 5313	2	
No. 6163	1	
No. 6875	1	



KCF-S cluster へのリンクを選択

各KCF-S clusterの情報

PCIDB Metabolites - Species - Metabolite Clusters -

KCF-S cluster No. 85

Relevant Phytochemical cluster No. 3

Index
Plant Species
BR08003 category (1)
BR08003 list (6)
Metabolite list (49)
Human Protein / Gene in interactions
Related Diseases

各項目の リストとリンク

Plant Species

Cumulative plant class count in this cluster

植物クラス分類 (ほぼバラ類)

class name	count	
rosids	299	
Magnoliophyta	12	
asterids	4	
core eudicotyledons	2	
Spermatophyta	2	

Cumulative plant family count in this cluster

ファミリー分類 (ほぼマメ科)

class name	count
Fabaceae	298
Berberidaceae	12
Cupressaceae	2
Styracaceae	2
Amaranthaceae	1
Daphniphyllaceae	1
Convolvulaceae	1
Orobanchaceae	1
Euphorbiaceae	1

BR08003 category (1)

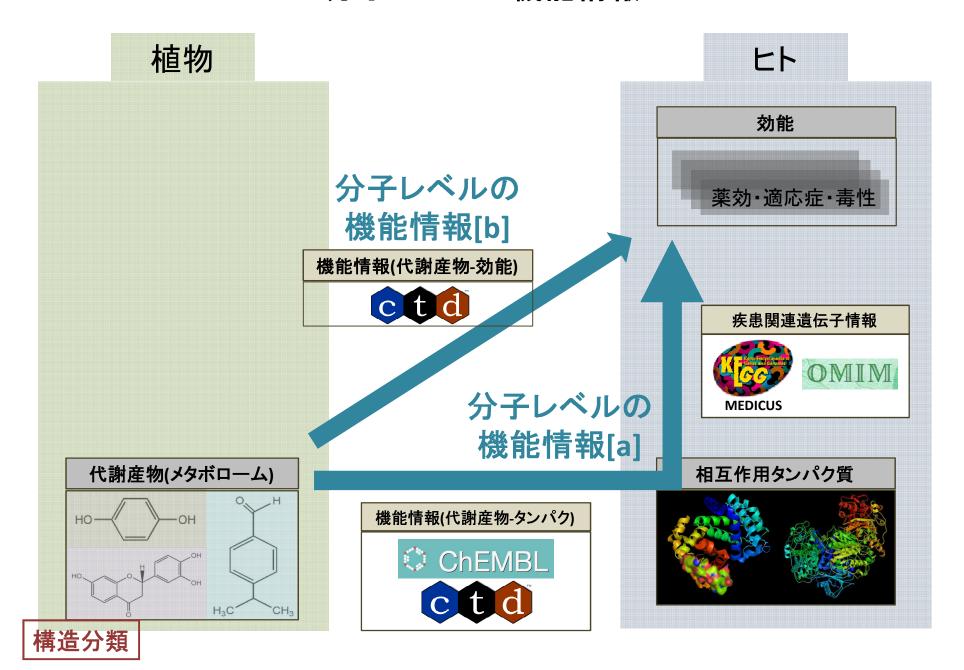
BR08003 Category	KEGG COMPOUND count
Quinolizidine alkaloids	5

BR08003 list (6)

BR08003 Category	KEGG ID	KNApSAcK ID
Quinolizidine alkaloids	C10752	C00002209
Quinolizidine alkaloids	C02621	C00002221
Quinolizidine alkaloids	C10772	C00002225
Quinolizidine alkaloids	C10774	C00002227
Quinolizidine alkaloids	C10776	C00002229
Quinolizidine alkaloids	C10752	C00007692

KEGG BRITE母核分類 (Quinolizidine alkaloids)

分子レベルの機能情報



各phytochemical clusterの情報(分子レベルの機能情報[a])

ChEMBL相互作用タンパク質の情報

Human Protein / Gene in interactions

相互作用の対象となる KNApSAcK代謝産物 各相互作用タンパク質 に関連する疾患の数(OMIM / KEGG)

117 ChEMBL Protein in interactions

accession	description	class description	KNApSAcK metabolite in interactions	# of diseases (OMIM / KEGG)
Q03164	Histone-lysine N-methyltransferase 2A	Enzyme	C00002207 C00002218 C00002224 C00002227 C00002236 C00002238 C00007681 C00007682 C00007735 C00007736 C00007757 C00023313 C00026289 C00026301 C00026325 C00026326 C00026329	1/2
O00255	Menin	Unclassified protein	C00002207 C00002218 C00002224 C00002227 C00002236 C00002238 C00007681 C00007682 C00007735 C00007736 C00007757 C00023313 C00026289 C00026301 C00026325 C00026326 C00026329	2/5
P28482	Mitogen-activated protein kinase 1	Erk	C00002227 C00002236 C00007681 C00007682 C00007757 C00023313 C00026289 C00026301 C00026329	0/0
P10635	Cytochrome P450 2D6	Cytochrome P450 2D6	C00002207 C00002218 C00002236 C00002238 C00007681 C00007682 C00026325 C00026326 C00026329	1/0
Q92830	Histone acetyltransferase KAT2A	Enzyme	C00002227 C00002236 C00007681 C00007682 C00007757 C00023313 C00026289 C00026301 C00026329	0/0
P33261	Cytochrome P450 2C19	Cytochrome P450 2C19	C00002207 C00002218 C00002236 C00002238 C00007681 C00007682 C00026325 C00026326 C00026329	1/1
P11712	Cytochrome P450 2C9	Cytochrome P450 2C9	C00002218 C00002236 C00007681 C00007682 C00007735 C00026325 C00026326 C00026329	0/1
P05177	Cytochrome P450 1A2	Cytochrome P450 1A2	C00002218 C00002236 C00007681 C00007682 C00026325 C00026326 C00026329	0/0
P08684	Cytochrome P450 3A4	Cytochrome P450 3A4	C00002218 C00002236 C00007681 C00007682 C00026325 C00026326 C00026329	0/1
P02545	Prelamin-A/C	Unclassified protein	C00002218 C00002236 C00007681 C00007682 C00026325 C00026326 C00026329	11 / 10
P83916	Chromobox protein homolog 1	Unclassified protein	C00002218 C00007758 C00026281 C00026325 C00026326	0/0
P11387	DNA topoisomerase 1	Isomerase	C00002227 C00007757 C00023313 C00026289 C00026301	0/0







相互作用タンパク質のUniProt ID、名前、分類

各phytochemical clusterの情報(分子レベルの機能情報[a])

CTD相互作用タンパク質の情報

相互作用の対象となる KNApSAcK代謝産物

11 Gene in CTD interactions

	lacktriangle				
gene	gene name	gene description	KNApSAcK meta	bolite in	interactions
1135	CHRNA2	cholinergic receptor, nicotinic, alpha 2 (neuronal)	C00002218		
1136	CHRNA3, LNCR2, NACHRA3, PAOD2	cholinergic receptor, nicotinic, alpha 3 (neuronal)	C00002218		
1137	CHRNA4, BFNC, EBN, EBN1, NACHR, NACHRA4, NACRA4	cholinergic receptor, nicotinic, alpha 4 (neuronal)	C00002218		
1139	CHRNA7, CHRNA7-2, NACHRA7	cholinergic receptor, nicotinic, alpha 7 (neuronal)	C00002218		
1141	CHRNB2, EFNL3, nAChRB2	cholinergic receptor, nicotinic, beta 2 (neuronal)	C00002218		
1143	CHRNB4	cholinergic receptor, nicotinic, beta 4 (neuronal)	C00002218	_	
4609	MYC, MRTL, MYCC, bHLHe39, c-Myc	v-myc avian myelocytomatosis viral oncogene homolog	C00002227		
4893	NRAS, ALPS4, N-ras, NRAS1, NS6	neuroblastoma RAS viral (v-ras) oncogene homolog	C00002227	4_	C00002227
7150	TOP1, TOPI	topoisomerase (DNA) I (EC:5.99.1.2)	C00002227		へのリンクを選択
7157	TP53, BCC7, LFS1, P53, TRP53	tumor protein p53	C00002227	,	
1565	CYP2D6, CPD6, CYP2D, CYP2D7AP, CYP2D7BP, CYP2D7P2, CYP2D8P2, CYP2DL1, CYPIID6, P450-DB1, P450C2D, P450DB1	cytochrome P450, family 2, subfamily D, polypeptide 6 (EC:1.14.14.1)	C00002236		





1

相互作用タンパク質のEntrez gene ID、名前、説明

各代謝産物の情報(分子レベルの機能情報[a])

ChEMBL/CTD相互作用タンパク質の情報

Human Protein / Gene in interaction

ChEMBL Assay / Activity の詳細情報

5 ChEMBL Protein in interactions

accession	description	class description	compound	assay ID (# of activities)	# of diseases (OMIM / KEGG)
P11387	DNA topoisomerase 1	Isomerase	CHEMBL1824581	CHEMBL1826307 (1)	0/0
Q92830	Histone acetyltransferase KAT2A	Enzyme	CHEMBL1733145	CHEMBL1738606 (1)	0/0
P28482	Mitogen-activated protein kinase 1	Erk	CHEMBL1396816	CHEMBL1613808 (1)	0/0
O00255	Menin	Unclassified protein	CHEMBL1396816	CHEMBL1614531 (1)	2/5
Q03164	Histone-lysine N-methyltransferase 2A	Enzyme	CHEMBL1396816	CHEMBL1614531 (1)	1/3

CTD interaction (4)

compound	gene	gene name	gene description	interaction	interaction type	form	reference pmid
C034244	4609	MYC MRTL MYCC bHLHe39 c-Myc	v-myc avian myelocytomatosis viral oncogene homolog	matrine results in decreased expression of MYC mRNA	decreases expression	mRNA	11489473
C034244	4893	NRAS ALPS4 N-ras NRAS1 NS6	neuroblastoma RAS viral (v-ras) oncogene homolog	matrine results in increased expression of NRAS mRNA	increases expression	mRNA	11489473
C034244	7150	TOP1 TOPI	topoisomerase (DNA) I (EC:5.99.1.2)	matrine analog results in decreased activity of TOP1 protein	decreases activity	protein	21194615
C034244	7157	TP53 BCC7 LFS1 P53 TRP53	tumor protein p53	matrine results in increased expression of TP53 mRNA	increases expression	mRNA	11489473

相互作用の説明、分類、作用対象

引用元文献

各phytochemical clusterの情報(分子レベルの機能情報[a])

ChEMBL相互作用タンパク質に関連する疾患情報

Related Diseases

Diseases related to proteins in ChEMBL interactions

OMIM (81)

ОМІМ	preferred title	UniProt
#300438	17-beta-hydroxysteroid dehydrogenase x deficiency	Q99714
#100100	Abdominal muscles, absence of, with urinary tract abnormality and cryptorchidism	P20309
#219080	Acth-independent macronodular adrenal hyperplasia; aimah	P63092
#103780	Alcohol dependence	P08172 P14416 P31645
#614373	Amyotrophic lateral sclerosis 16, juvenile; als16	Q99720
#602025	Body mass index quantitative trait locus 9; bmiq9	P41968
#300615	Brunner syndrome	P21397
#115200	Cardiomyopathy, dilated, 1a; cmd1a	P02545
#212112	Cardiomyopathy, dilated, with hypergonadotropic hypogonadism	P02545
#605588	Charcot-marie-tooth disease, axonal, type 2b1; cmt2b1	P02545
#162800	Cyclic neutropenia	P08246
#612522	Diabetes mellitus, insulin-dependent, 22; iddm22	P51681
#609535	Drug metabolism, poor, cyp2c19-related	P33261
#608902	Drug metabolism, poor, cyp2d6-related	P10635
#181350	Emery-dreifuss muscular dystrophy 2, autosomal dominant; edmd2	P02545

1	_	
OMIM	phenotype II	

疾患に関連するタンパク質ID

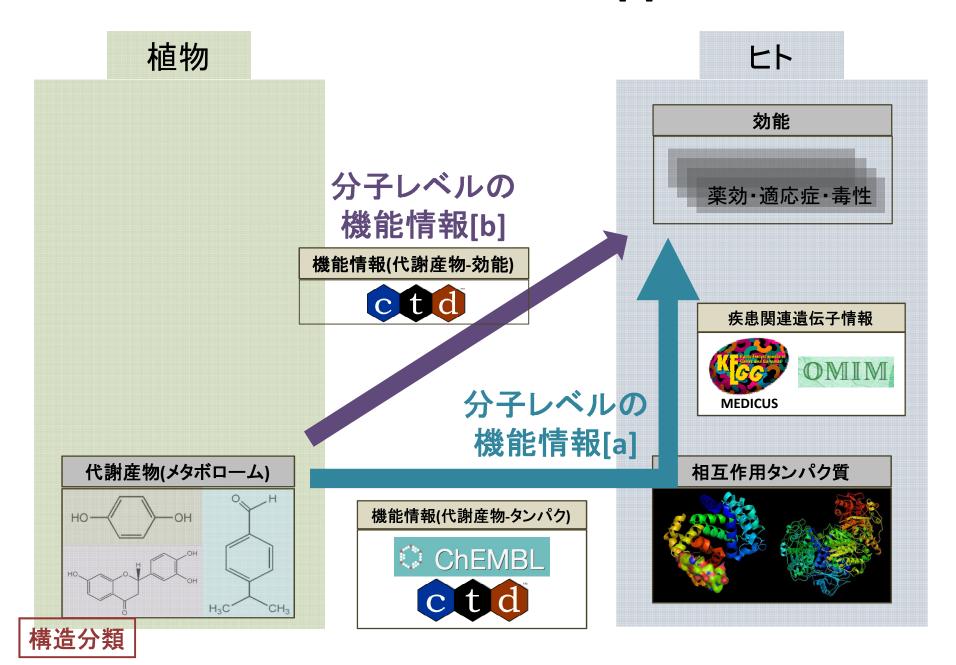
KEGG DISEASE (68)

name	UniProt
Adrenal carcinoma	O00255 (related)
Carcinoid	O00255 (related)
Malignant islet cell carcinoma	O00255 (related)
Primary hyperparathyroidism	O00255 (related)
Pituitary adenomas	O00255 (related)
Oral cancer	P00533 (related) P00533 (marker)
Esophageal cancer	P00533 (related) P35354 (related)
Gastric cancer	P00533 (related) P04626 (related)
Bladder cancer	P00533 (related) P04626 (related)
Choriocarcinoma	P00533 (related) P03956 (related) P04626 (related)
Cervical cancer	P00533 (related) P04626 (related)
Glioma	P00533 (related) P00533 (marker)
Laryngeal cancer	P00533 (related) P00533 (marker)
	Adrenal carcinoma Carcinoid Malignant islet cell carcinoma Primary hyperparathyroidism Pituitary adenomas Oral cancer Esophageal cancer Gastric cancer Bladder cancer Choriocarcinoma Cervical cancer Glioma



疾患に関連するタンパク質ID

分子レベルの機能情報[b]



各phytochemical clusterの情報(分子レベルの機能情報[b])

CTDの化合物と直接関連する疾患情報

Diseases related to CTD interactions

14 disease in interactions with metabolites

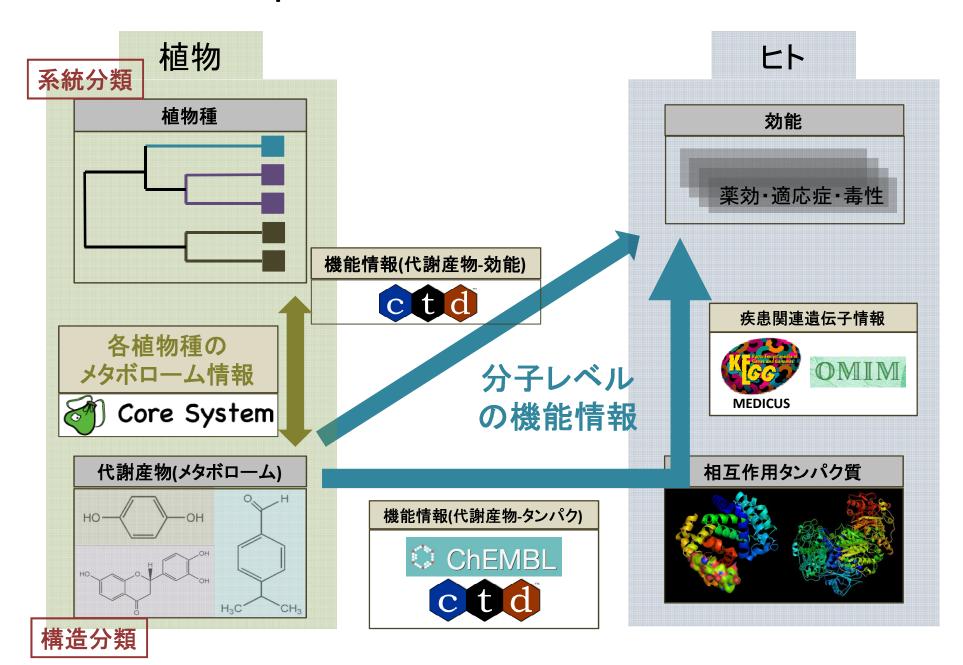
MESH or OMIM	name	KNApSAcK metabolite
D001145	Arrhythmias, Cardiac	C00002214 C00002236
D006331	Heart Diseases	C00002227
D064420	Drug-Related Side Effects and Adverse Reactions	C00002214
D014693	Ventricular Fibrillation	C00002214
D002289	Carcinoma, Non-Small-Cell Lung	C00002227
D002779	Cholestasis	C00002227
D056487	Drug-Induced Liver Injury, Chronic	C00002227
D012640	Seizures	C00002218
D008113	Liver Neoplasms	C00002227
D003092	Colitis	C00007758
D019694	Hepatitis B, Chronic	C00007758
D007249	Inflammation	C00007758
D008106	Liver Cirrhosis, Experimental	C00007758
D054537	Atrioventricular Block	C00002236



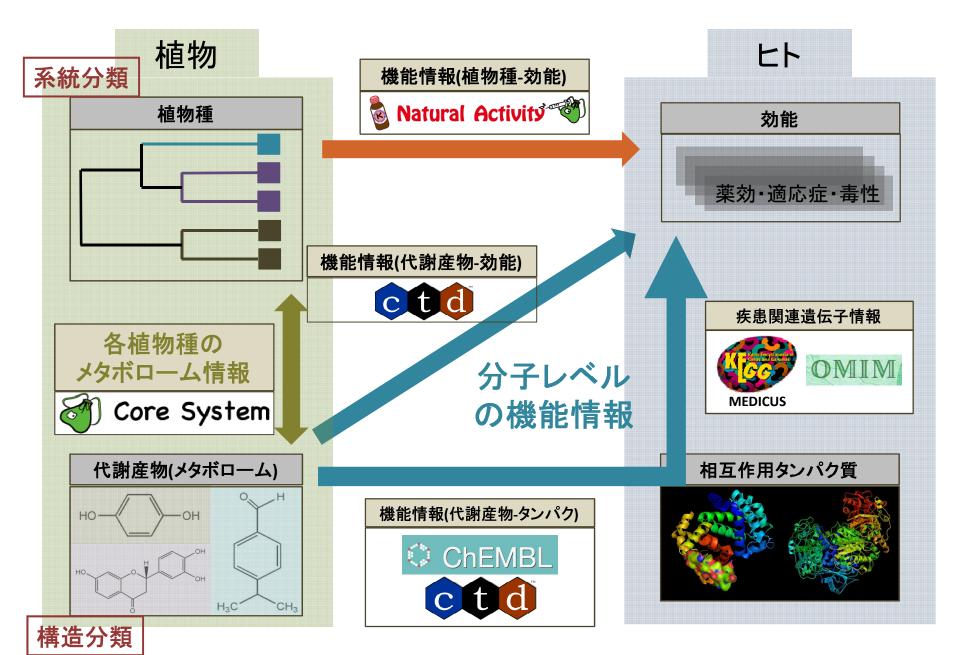


疾患に関連する KNApSAcK代謝産物ID

KNApSAcK Coreと分子レベルの機能情報



Natural Activityの検証



各Speciesのページ

Natural Activityの検証例

PCIDB

/letabolites

Species *

Metabolite Clusters *

Acnistus arborescens

Index
Plant Species
Natural Activity
Metabolite list (3)
Human Protein
/ Gene in interactions
Related Diseases

Species

KNApSAcK エントリー情報

KNApSAcK Entry

Organism name	Acnistus arborescens
Genus	Acnistus
Family	Solanaceae
Kingdom	Plantae

NCBI taxonomy

NCBI taxonomy情報

Entry

Linked NCBI taxonomy name	Acnistus arborescens	
Linked NCBI taxonomy ID	362341	
Linked level	species	

Family

マメ科

Family in NCBI taxonomy	Solanaceae
ID	4070

Kingdom (Superkingdom)

Kingdom (Superkingdom) in NCBI taxonomy	Viridiplantae
ID	33090

Plant class

Plant class	asterids
ID	71274

Natural Activity

List (8)

Natural Activityのリスト

Species	Activity		
Acnistus arborescens (L.) Schltdl.	Anticancer	坮億/∧∽	iconcor)
Acnistus arborescens (L.) Schltdl.	Antileukemic	抗癌(Ant	
Acnistus arborescens (L.) Schltdl.	Antimalarial	抗白血病	(Antileukemic)
Acnistus arborescens (L.) Schltdl.	Antiproliferant		
Acnistus arborescens (L.) Schltdl.	Cytotoxic		
Acnistus arborescens (L.) Schltdl.	Diuretic		
Acnistus arborescens (L.) Schltdl.	Emollient		
Acnistus arborescens (L.) Schltdl.	Piscicide		

Natural Activityの検証例

Human Protein / Gene in interactions

27 ChEMBL Protein in interactions

accession	description	class description	KNApSAcK me	etabolite in interactions	# of diseases (OMIM / KEGG)
P17861	X-box-binding protein 1	Unclassified protein	C00003676		1/0
Q99700	Ataxin-2	Unclassified protein	C00003676		1/1
P49798	Regulator of G-protein signaling 4	Unclassified protein	C00003676		2/0
P06746	DNA polymerase beta	Enzyme	C00003676		0/0
P42574	Caspase-3	C14	C00003676		0/0
P28482	Mitogen-activated protein kinase 1	Erk	C00003676		0/0
P27361	Mitogen-activated protein kinase 3	Erk	C00003676		0/0
P11473	Vitamin D3 receptor	NR1I1	C00003676		2/3
P42858	Huntingtin	Unclassified protein	C00003676		1/1
P84022	Mothers against decapentaplegic homolog 3	Unclassified protein	C00003676		2/0
O75496	Geminin	Unclassified protein	C00003676		0/0
P51151	Ras-related protein Rab-9A	Unclassified protein	C00003676		0/0
P32245	Melanocortin receptor 4	Melanocortin receptor	C00003676		1/0
Q9Y253	DNA polymerase eta	Enzyme	C00003676		1/1
P07355	Annexin A2	Unclassified protein	C00003676	Acnistus arborescensが持つ	0/0
O15118	Niemann-Pick C1 protein	Unclassified protein	C00003676	代謝産物 C00003676と	1/1
Q96QE3	ATPase family AAA domain-containing protein 5	Unclassified protein	C00003676	P31749 (Akt)の相互作用	0/0
P31749	RAC-alpha serine/threonine-protein kinase	Akt	C00003676		4 / 1
Q9UNA4	DNA polymerase iota	Enzyme	C00003676		0/0
Q6W5P4	Neuropeptide S receptor	Neuropeptide receptor	C00003676		1/0
P27695	DNA-(apurinic or apyrimidinic site) lyase	Enzyme	C00003676		0/0
P10636	Microtubule-associated protein tau	Unclassified protein	C00003676		4/3

各Speciesのページ

Natural Activityの検証例

Related Diseases

Diseases related to proteins in ChEMBL interactions

OMIM (22)

OMIM	preferred title	UniProt
#608584	Asthma-related traits, susceptibility to, 2	Q6W5P4
#114480	Breast cancer	P31749
#114500	Colorectal cancer; crc	P31749 P84022 Q14191
#615109	Cowden syndrome 6; cws6	P31749
#600274	Frontotemporal dementia; ftd	P10636
#137800	Glioma susceptibility 1; glm1	O75874
#143100	Huntington disease; hd	P42858
#613795	Loeys-dietz syndrome, type 3; lds3	P84022
#612371	Major affective disorder 7; mafd7	P17861
#607948	Mycobacterium tuberculosis, susceptibility to	P11473
#257220	Niemann-pick disease, type c1; npc1	O15118
#601665	Obesity	P32245
#260540	Parkinson-dementia syndrome	P10636
#172700	Pick disease of brain	P10636
#176920	Proteus syndrome	P31749
#604906	Schizophrenia 9; sczd9	P49798
#181500	Schizophrenia; sczd	P49798
#183090	Spinocerebellar ataxia 2; sca2	Q99700
#601104	Supranuclear palsy, progressive, 1; psnp1	P10636
#277440	Vitamin d-dependent rickets, type 2a; vddr2a	P11473
#277700	Werner syndrome; wrn	Q14191
#278750	Xeroderma pigmentosum, variant type; xpv	Q9Y253

KEGG DISEASE (16)

KEGG	name	UniProt
H00136	Niemann-Pick disease type C (NPC)	O15118 (related)
H00081	Hashimoto's thyroiditis	P01215 (marker)
H00082	Graves' disease	P01215 (marker)
H00250	Congenital nongoitrous hypothyroidism (CHNG)	P01215 (marker)
H00058	Amyotrophic lateral sclerosis (ALS)	P10636 (related)
H00077	Progressive supranuclear palsy (PSP)	P10636 (related)
H00078	Frontotemporal lobar degeneration (FTLD)	P10636 (related)
H00342	Tuberculosis	P11473 (related)
H00784	Localized autosomal recessive hypotrichosis	P11473 (related)
H01143	Vitamin D-dependent rickets	P11473 (related)
H00539	PTEN hamartoma tumor syndrome (PHTS)	P31749 (related)
H00059	Huntington's disease (HD)	P42858 (related)
H00296	Defects in RecQ helicases	Q14191 (related)
H00063	Spinocerebellar ataxia (SCA)	Q99700 (related)
H00403	Disorders of nucleotide excision repair	Q9Y253 (related)

P31749 (Akt)は癌や腫瘍に関連がある

Diseases related to CTD interactions

2 disease in interactions with metabolites

MESH or OMIM	name	KNApSAcK metabolite
D007249	Inflammation	C00003676
D009369	Neoplasms	C00003676

CTDに関しても、代謝産物C00003676の 腫瘍形成に対する関連情報がある

研究開発成果

- KNApSAcK生物種と系統分類の標準化
 - NCBI Taxonomyへの生物種名のマッピングと、それを用いた生物種分類
- KNApSAcKと、ChEMBL, CTDとの化合物リンクの作成
- KNApSAcK代謝産物の構造類似性に基づく分類及び母核分類を利用した分類の作成
- ・ 分子レベルの機能情報の統合
- ・ PCIDB(PhytoChemical Interactions DB)の作成と公開

今後の発展性

- ・ 統一的な疾患分類体系の構築
 - MetaMap, UMLS
- ChEMBLやCTD以外の分子レベルの機能情報の統合
 - 例) 医薬品: KEGG DRUG、生薬: KEGG ENVIRON、副作用: SIDER
- ChEMBLアッセイの活性詳細情報や、ヒト以外の相互作用タンパク質情報の利用
 - 感染症(抗生物質、抗真菌薬、抗ウイルス薬等への応用)
- KNApSAcKに登録されていない代謝産物情報の利用
 - 新規代謝産物に関して、KCF-S clusterやphytochemical clusterを 利用した、分子レベルの相互作用情報予測