MetaboBank
サイトビジット 2019 0712

有田正規　　遺伝研
金谷重彦　　NAIST
櫻井望　　遺伝研
平川英樹　　かずさDNA
（福島敦史　　理研CSRS）

メンバー紹介

遺伝研： **有田 (MassBank)、櫻井 (メタデータ、サーバ)**、川島、多田
直接雇用：時松敏明 （メタデータ項目、BioSample摺合）
佐藤充治 （代謝マップエディタ）

奈良先： **金谷（代謝マップ）**、小野、他ラボメンバー
直接雇用：森田晶（システム）、大橋美名子（補佐）、ほか学生

かずさ： **平川（統括）、櫻井（食品データ）**
直接雇用：長崎英樹（MassBase整理）、大澤祥子（補佐）

理研CSRS： 福島（PMM作成）、津川、小林
直接雇用：高橋みき子（PRIMe整理）、ほか学生

発表は有田（10）、金谷（10）、櫻井（10）、有田（10）

開始時の問題点と目標

データ再利用の難しさ

- 生データの取得法、フォーマットが多様
  →国際標準化（Lipidomics）、ABFフォーマットの採用
- データ統合が困難
  →CSRSかずさは統合、EBIと連携、BioSample等の利用

最終目標

- 論文投稿時に登録するリポジトリ = 恒久的なID付与
- 統合解析のできるデータサーバ
  - 生データに解析結果や測定法がぶら下がる構造　= 他DBとの連携
    （例：Metabolonoteの分類法）
  - 解析結果（例えばスペクトル、化合物リスト）の検索 = MassBank
  - データ登録者へのリワード

2020年までのマイルストーン

いずれも国際標準が前提

遺伝研： **恒久的なデータ・リポジトリの開設** （遺伝研アドレス）

初期データ： 理研CSRSおよびかずさDNA研のデータ

（MassBase, DropMet, Komics）

投稿データ： 共同研究先を優先

プロジェクトデータ： AMED微生物叢領域との連携

かずさ： 保有するメタボロームデータの再解析

PowerGet等、ソフトウェアのアップデート　遺伝研でリリース

理研　： 保有するメタボロームデータの再解析　　理研でリリース

奈良先： 生体活性オントロジー構築、メタ代謝マップ構築

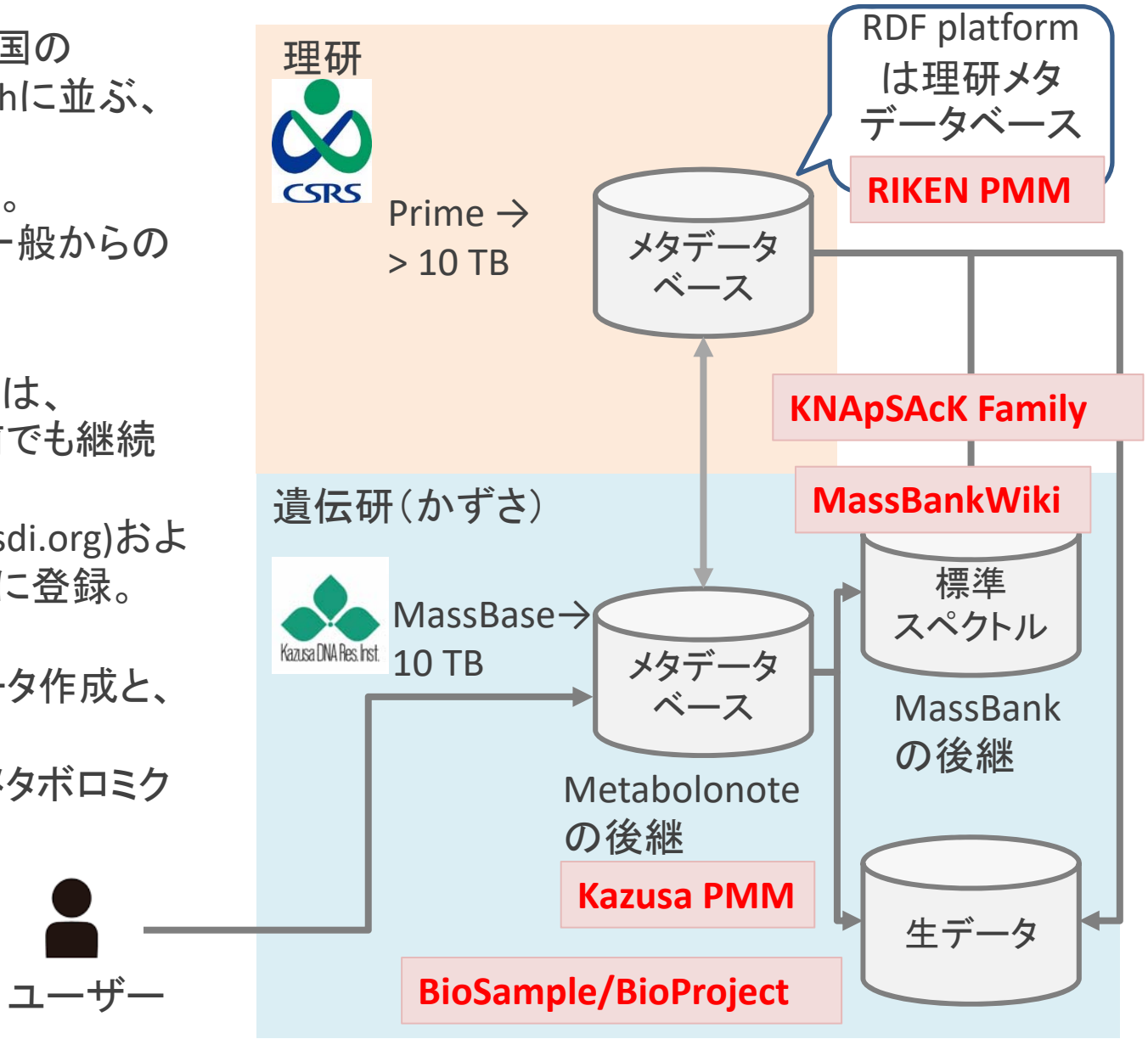
MetaboBank 全体構想

欧州のMetaboLights, 米国のMetabolomics Workbenchに並ぶ、生データのリポジトリ。
DDBJサービスとして開設。
平成31～2年度には、一般からの投稿を受付。

標準スペクトルのサーバは、MassBank wikiという名前でも継続

データはOmics DI (omicsdi.org)および metabolomeXchange に登録。

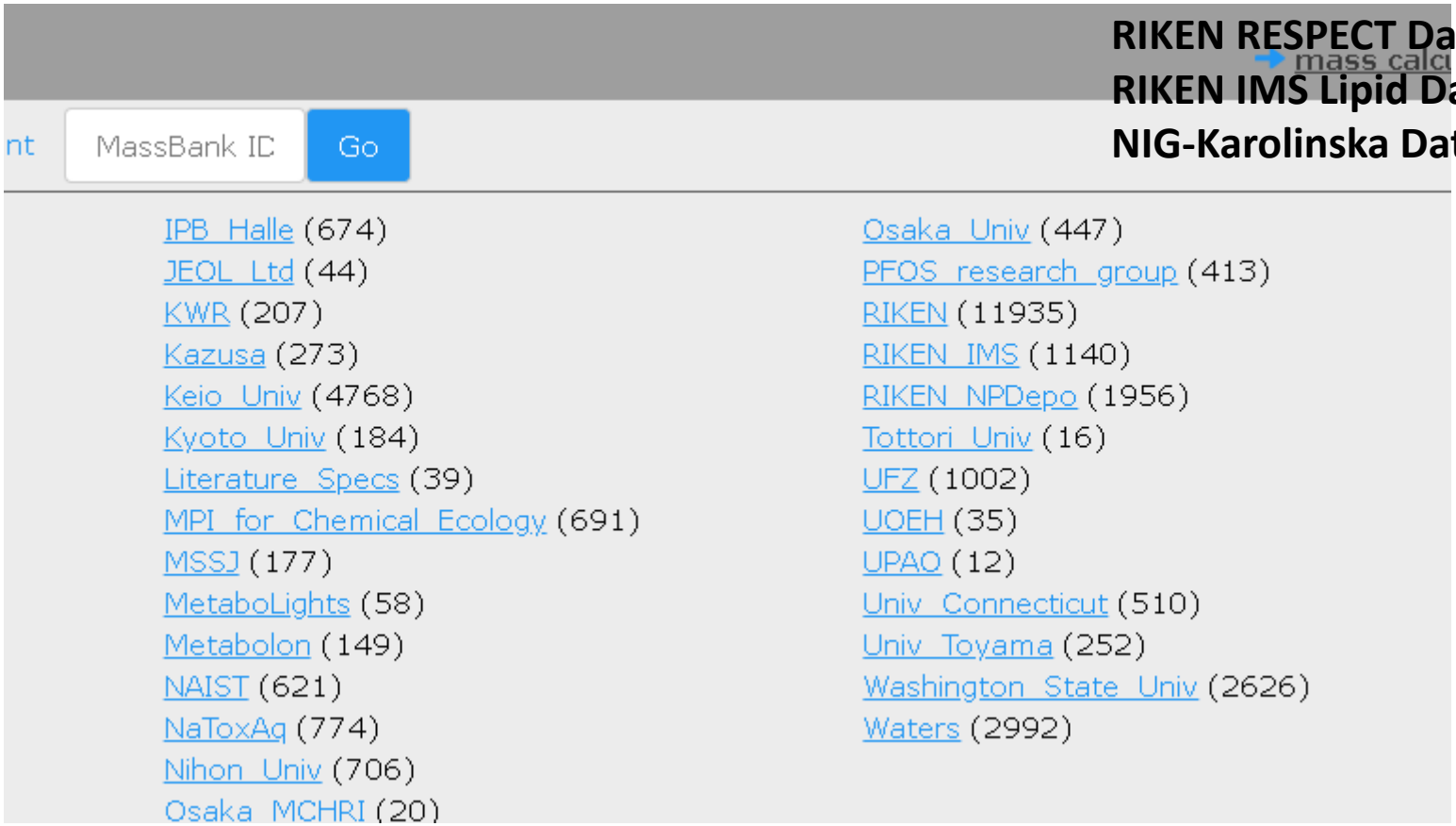
今年度は植物のメタデータ作成と、フォーマット決定。
来年度以降、医薬学のメタボロミクスへ拡張



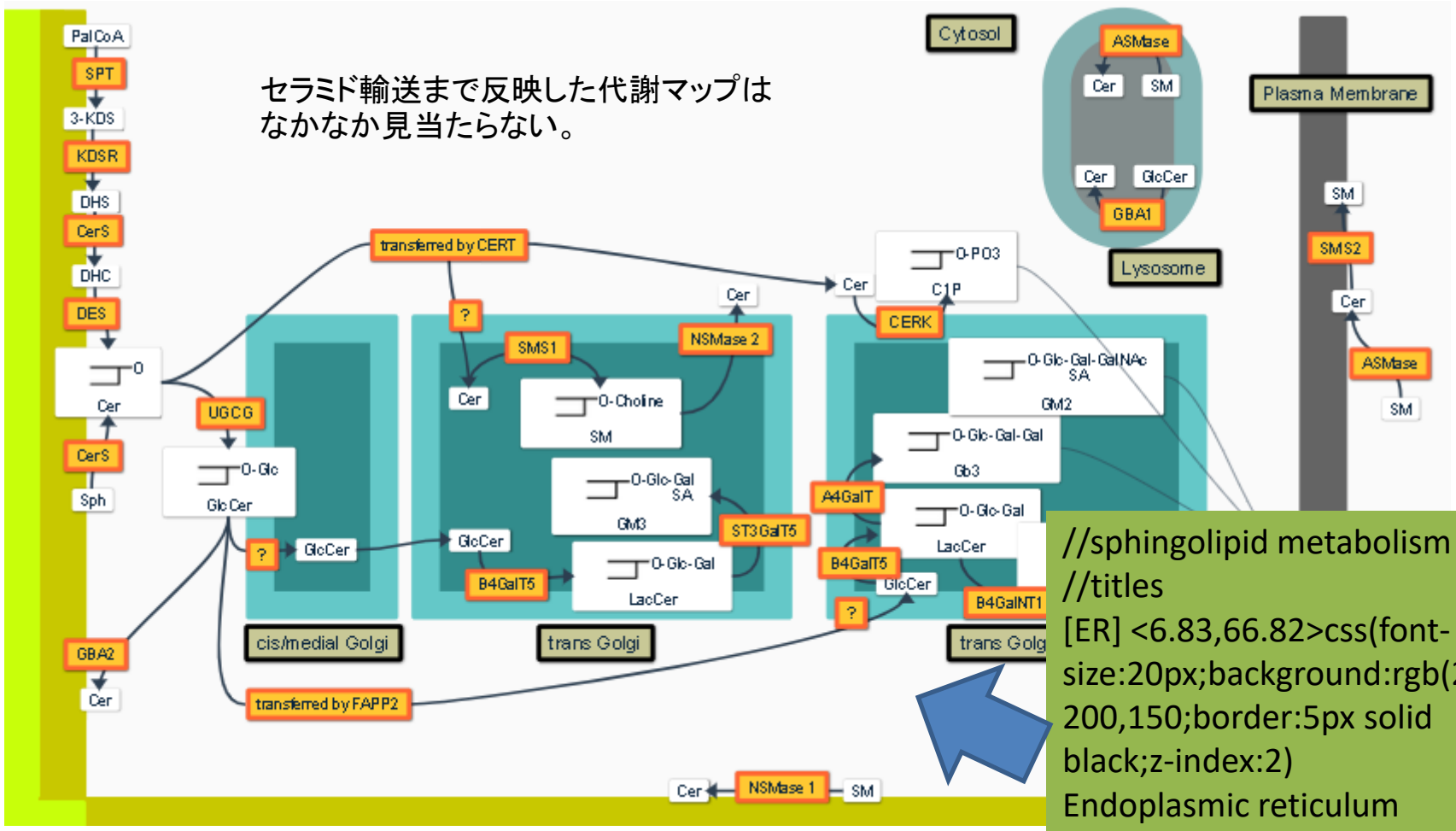
MassBank GitHub Record

MassBank.eu側がまず登録

MSSJデータ 177 にも対応
RIKEN PLASMA Data (5000程度)
RIKEN RESPECT Data (4000程度)
RIKEN IMS Lipid Data (1000程度)
NIG-Karolinska Data (800程度)



代謝動態データベース及びマップの構築



テキスト情報と連動したオープンソースの代謝マップエディタ作成
セラミド生合成マップを記載して、各種DBとリンク。
LipidBank, MassBank wiki　データベース内に設置予定

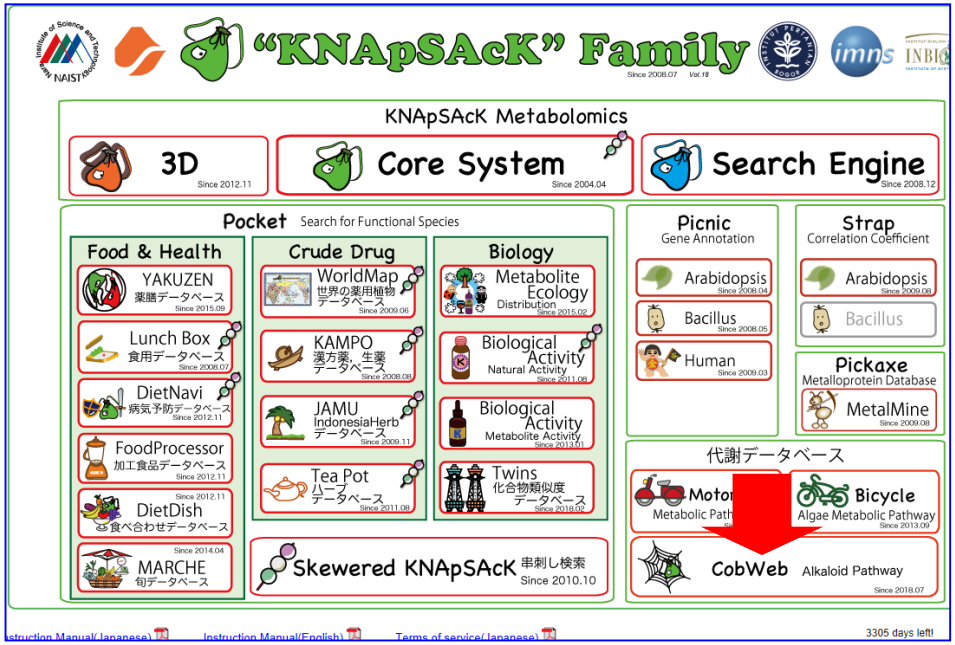
```
//sphingolipid metabolism map
//titles
[ER] <6.83,66.82>css(font-size:20px;background:rgb(200,200,150;border:5px solid black;z-index:2)
Endoplasmic reticulum

[Cytosol] <70,1>css(font-size:20px;background:rgb(200,200,150;border:5px solid black;z-index:2)
Cytosol
```

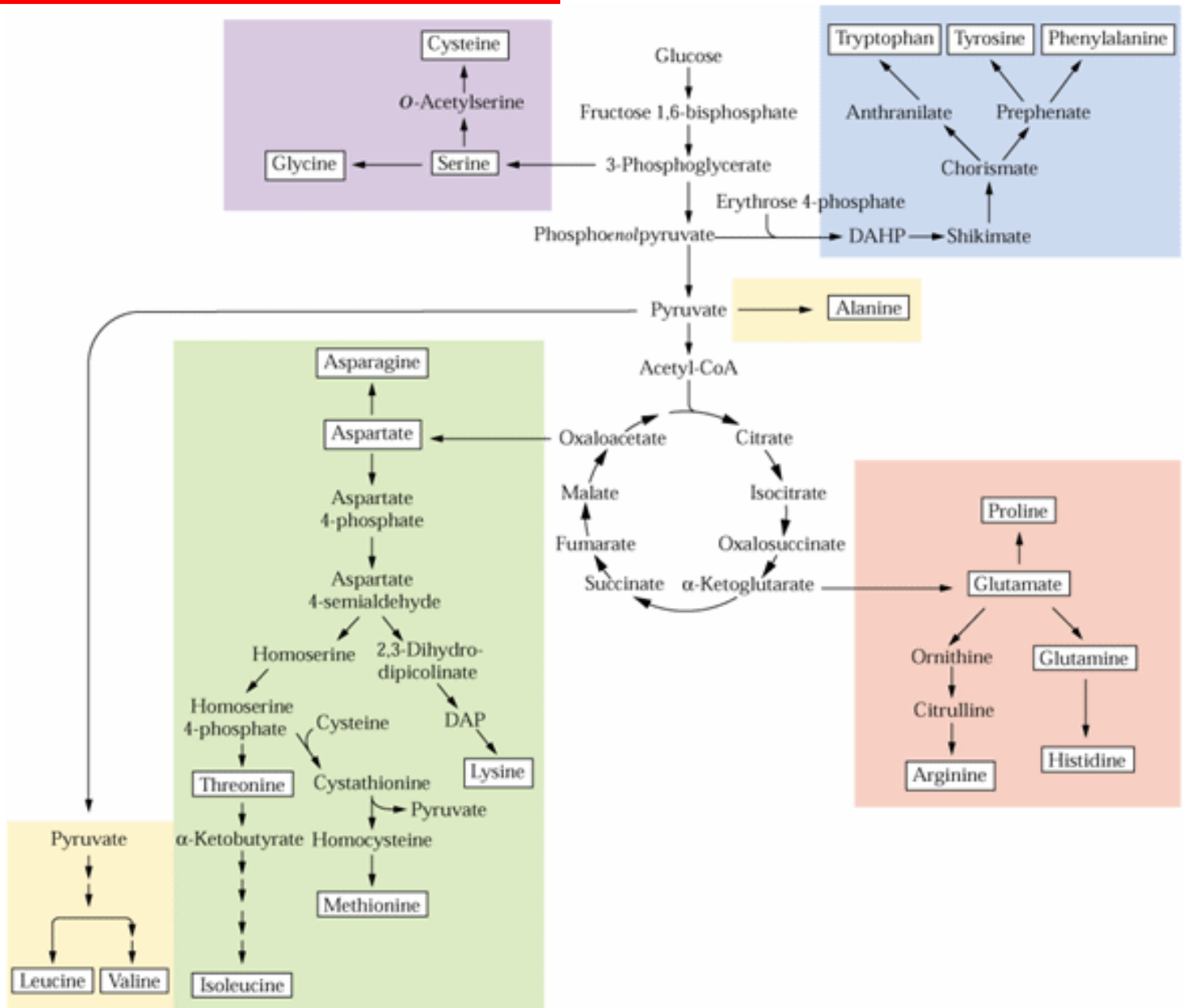
二次代謝生合成経路の悉皆的解析を目指したデータサイエンス : KNApSAcK family DBの展開

Shigehiko Kanaya
Nara Institute of Science and Technology (NAIST)
Graduate School of Science and Technology
skanaya@gtc.naist.jp

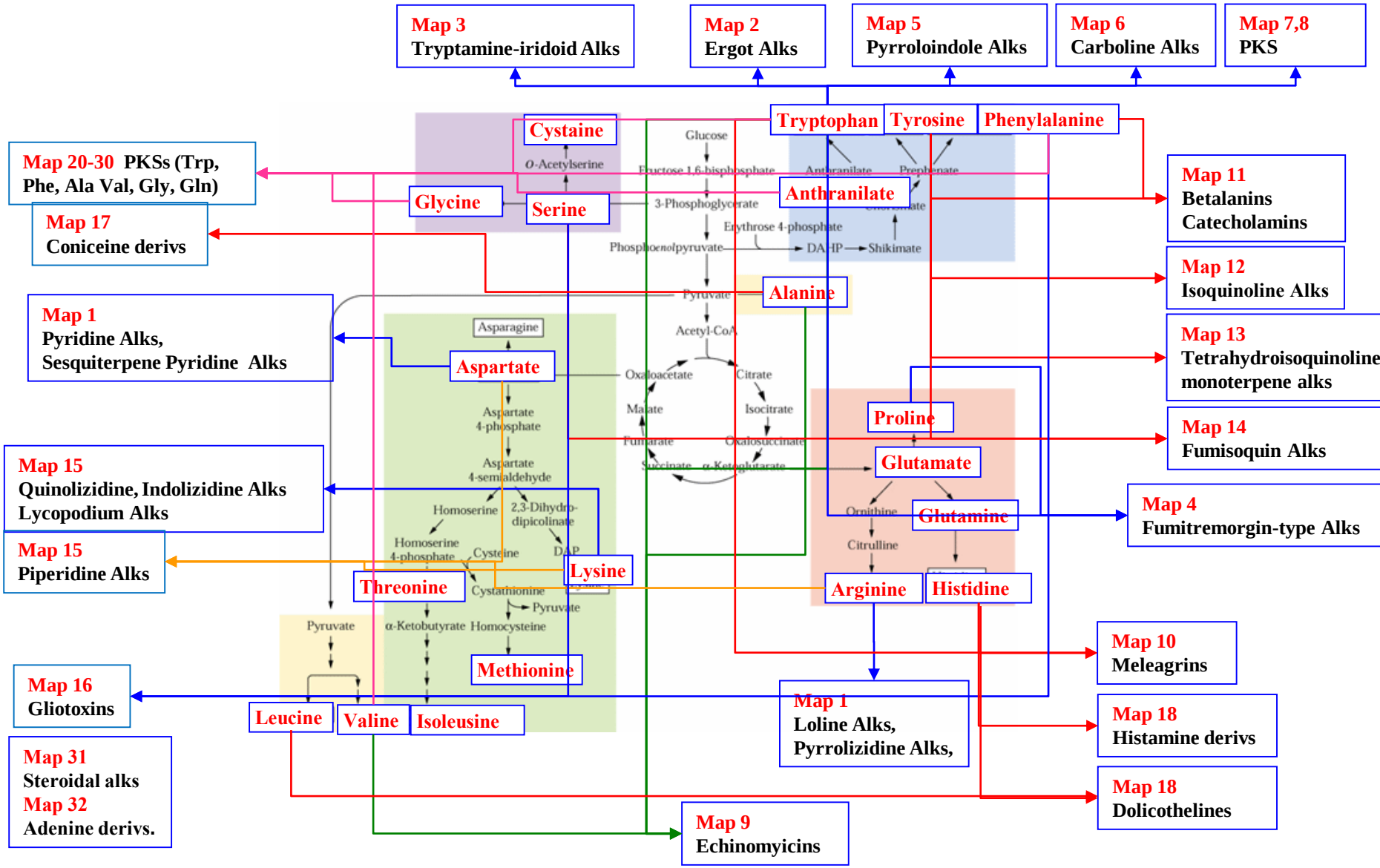
Last updata 2019/4/1
51,179 unique metabolites
116,315 species-metabolite pairs
Plant Cell Physiol.53, e1(1-12) (2012)



Primary Metabolic pathway: directly involved in normal growth, development, and reproduction



Secondary Metabolism: are not absolutely required for the survival of the organism. Alkaloids are biosynthesized by primary metabolites such as amino acids nucleotide, steroids, and secondary metabolites such as terpenoids..



“CobWeb”

Please input MetaboliteName. Multiple searches are possible for MetaboliteName. The delimiter is : (colon).

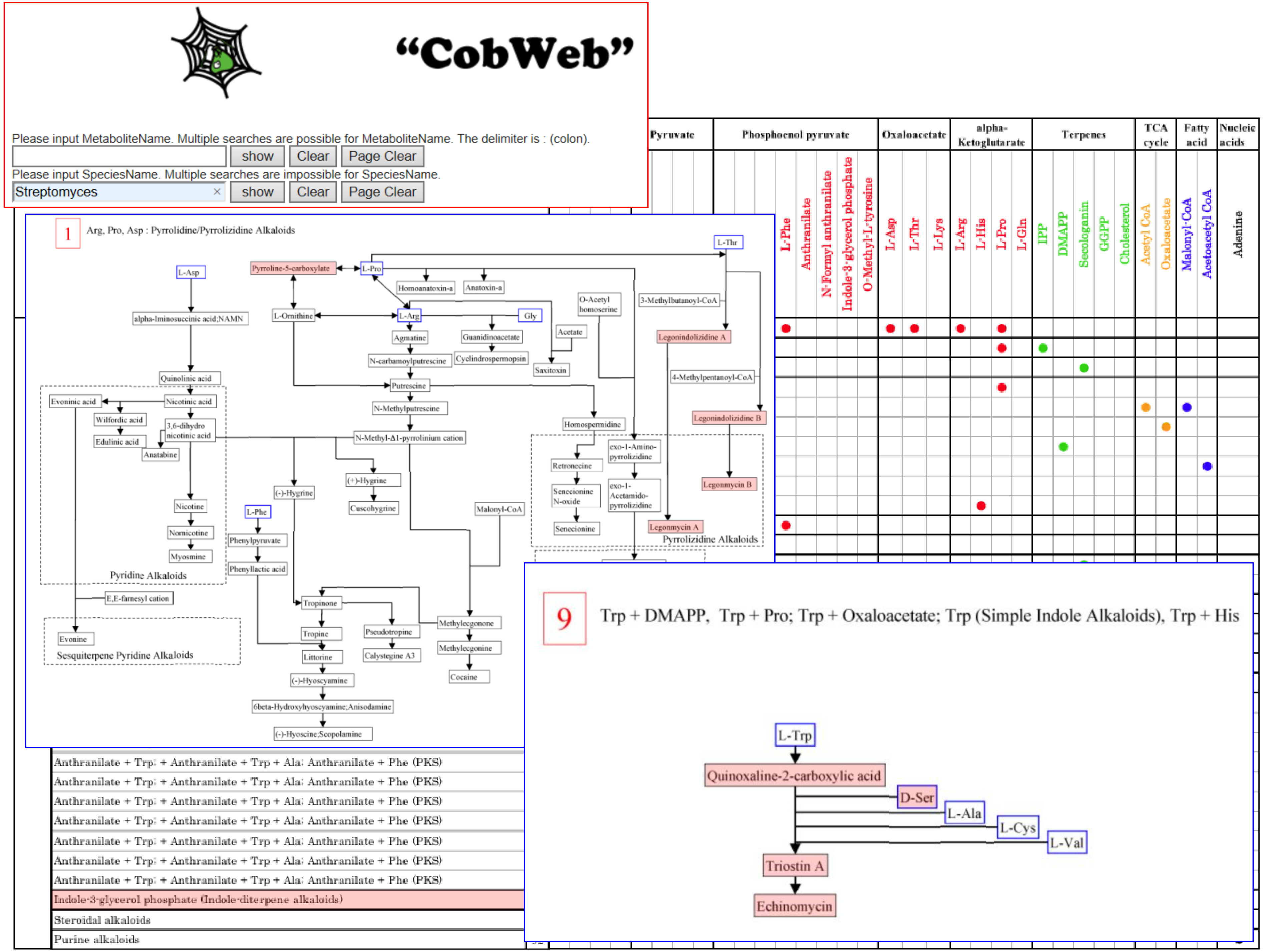
show | Clear | Page Clear

Please input SpeciesName. Multiple searches are impossible for SpeciesName.

Streptomyces × show | Clear | Page Clear

Starting substance (SS)

Pathway Map ID			Pyruvate				Phosphoenol pyruvate								Oxaloacetate			alpha-Ketoglutarate				Terpenes					TCA cycle		Fatty acid		Nucleic acids
			Gly	L-Ser	D-Ser	L-Cys	L-Ala	D-Ala	L-Leu	L-Val	L-Trp	D-Trp	L-Tyr	L-Phe	Anthranilate	N-Formyl anthranilate	Indole-3-glycerol phosphate	O-Methyl-L-tyrosine	L-Asp	L-Thr	L-Lys	L-Arg	L-His	L-Pro	L-Gln	IPP	DMAPP	Secologanin	GGPP	Cholesterol	Acetyl CoA	Oxaloacetate	Malonyl-CoA	Acetoacetyl CoA	Adenine
	Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids	1	●											●					●	●		●		●										
	Trp: Ergot Alkaloids	2					●			●	●													●		●									
	Trp: Monoterpenoid Indole Alkaloids	3									●																	●							
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	4									●													●											
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	5									●																			●		●			
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	6									●																				●				
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	7									●																●								
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	8									●																						●		
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	9			●	●	●			●	●																								
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	10									●												●												
	Tyr: Betanin group; Tyr + Phe, Colchicine	11											●	●																					
	Tyr: Isoquinoline Alkaloids (Benzylisoquinoline Alkaloids)	12											●																						
	Tyr + Secologanin (Tetrahydroisoquinoline monoterpene alkaloids), Tyr + Ser	13											●															●							
	Tyr + Secologanin (Tetrahydroisoquinoline monoterpene alkaloids), Tyr + Ser	14		●									●																						
	Lys: Quinolizidine, Piperidine Alkaloids, : Indolizidine Alkaloids Lycopodium Alkaloids	15																	●		●	●													
	Phe + Ser	16		●										●																						
	Ala + Acetyl-CoA + Malonyl-CoA	17					●																								●		●		
	His: Imidazole alkaloid(histidine derivatives)	18							●														●												
	Anthraniate(Acridine Alkaloids)	19													●			●	●																
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	20									●				●																				
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	21												●	●																				
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	22					●	●			●				●																				
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	23					●			●		●			●																				
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	24					●					●				●																			
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	25	●								●				●																				
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	26	●												●																				
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	27													●									●											
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	28													●																		●		
	Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)	29													●										●										
	Indole-3-glycerol phosphate (Indole-diterpene alkaloids)	30															●											●							
	Steroidal alkaloids	31																											●						
	Purine alkaloids	32																																	●


"CobWeb"
Please input MetaboliteName. Multiple searches are possible for MetaboliteName. The delimiter is : (colon).
show
Clear
Page Clear
Please input SpeciesName. Multiple searches are impossible for SpeciesName.
Streptomyces
show
Clear
Page Clear
Pyruvate
Phosphoenol pyruvate
Oxaloacetate
alpha-Ketoglutarate
Terpenes
TCA cycle
Fatty acid
Nucleic acids
L-Phe
Anthranilate
N-Formyl anthranilate
Indole-3-glycerol phosphate
O-Methyl-L-tyrosine
L-Asp
L-Thr
L-Lys
L-Arg
L-His
L-Pro
L-Gln
IPP
DMAPP
Secologanin
GGPP
Cholesterol
Acetyl CoA
Oxaloacetate
Malonyl-CoA
Acetoacetyl CoA
Adenine
1
Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids
L-Thr
L-Asp
Pyrroline-5-carboxylate
L-Pro
Homoanatoxin-a
Anatoxin-a
O-Acetyl homoserine
3-Methylbutanoyl-CoA
alpha-Iminosuccinic acid;NAMN
L-Ornithine
L-Arg
Gly
Agmatine
Guanidinoacetate
Acetate
Legonindolizidine A
N-carbamoylputrescine
Cyclindrospermopsin
Saxitoxin
4-Methylpentanoyl-CoA
Quinolinic acid
Putrescine
Evoninic acid
Nicotinic acid
N-Methylputrescine
Legonindolizidine B
Wilfordic acid
3,6-dihydro nicotinic acid
Homospermidine
Edulinic acid
N-Methyl-Δ1-pyrrolinium cation
Anatabine
exo-1-Amino-pyrrolizidine
Retronecine
(+)-Hygrine
Legonmycin B
(-)-Hygrine
Senecionine N-oxide
exo-1-Acetamido-pyrrolizidine
Nicotine
Cuscohygrine
Malonyl-CoA
L-Phe
Legonmycin A
Senecionine
Pyrrolizidine Alkaloids
Nornicotine
Phenylpyruvate
Myosmine
Phenyllactic acid
Pyridine Alkaloids
E,E-farnesyl cation
Tropinone
Methylecgonone
Evonine
Tropine
Pseudotropine
Methylecgonine
Sesquiterpene Pyridine Alkaloids
Littorine
Calystegine A3
Cocaine
(-)-Hyoscyamine
6beta-Hydroxyhyoscyamine;Anisodamine
(-)-Hyoscine;Scopolamine
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Anthranilate + Trp; + Anthranilate + Trp + Ala; Anthranilate + Phe (PKS)
Indole-3-glycerol phosphate (Indole-diterpene alkaloids)
Steroidal alkaloids
Purine alkaloids
9
Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His
L-Trp
Quinoxaline-2-carboxylic acid
D-Ser
L-Ala
L-Cys
L-Val
Triostin A
Echinomycin

"CobWeb" from KNApSAcK

Please input MetaboliteName. Multiple searches are possible for MetaboliteName. The delimiter is : (colon).

show Clear Page Clear

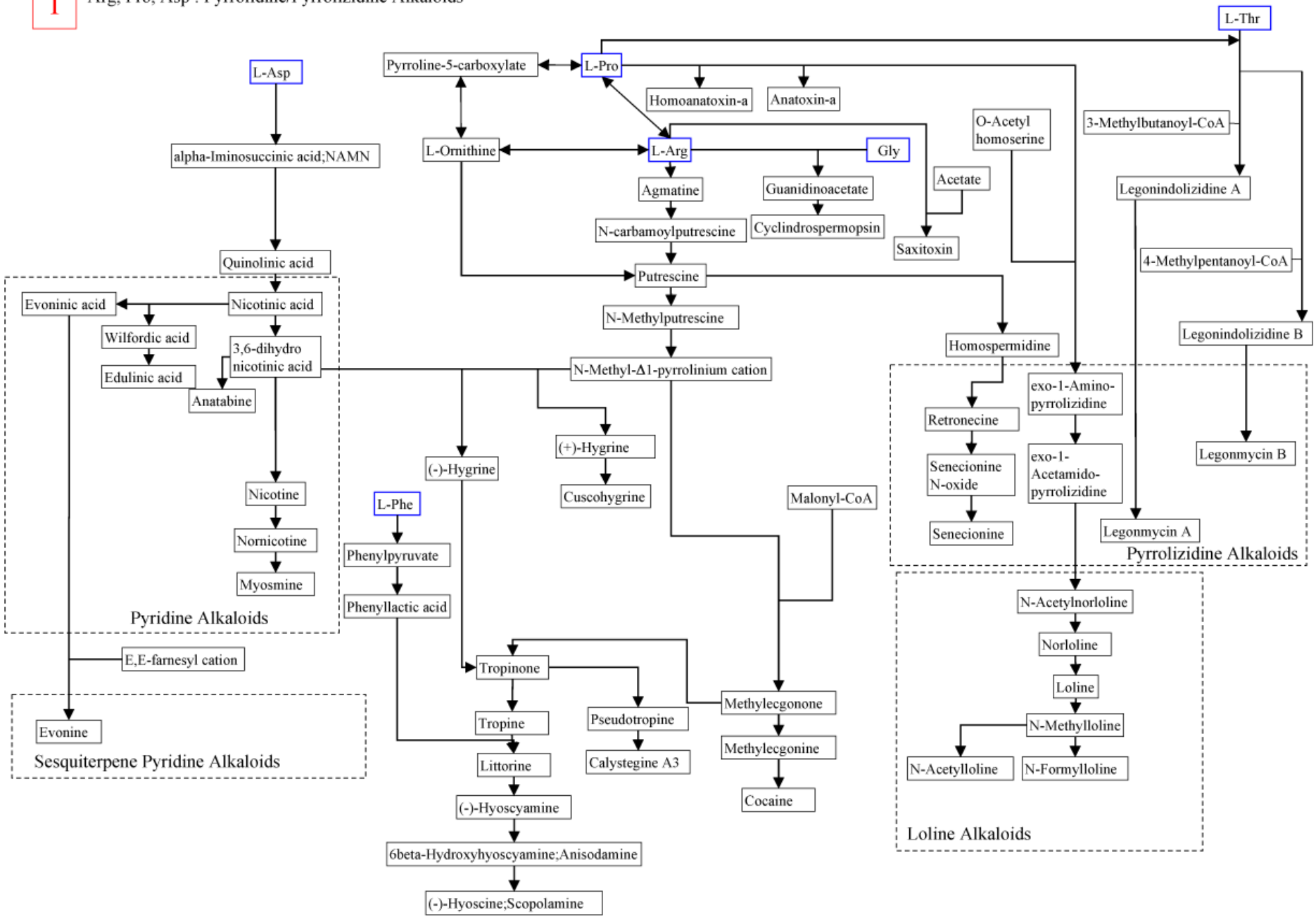
Please input SpeciesName. Multiple searches are impossible for SpeciesName.

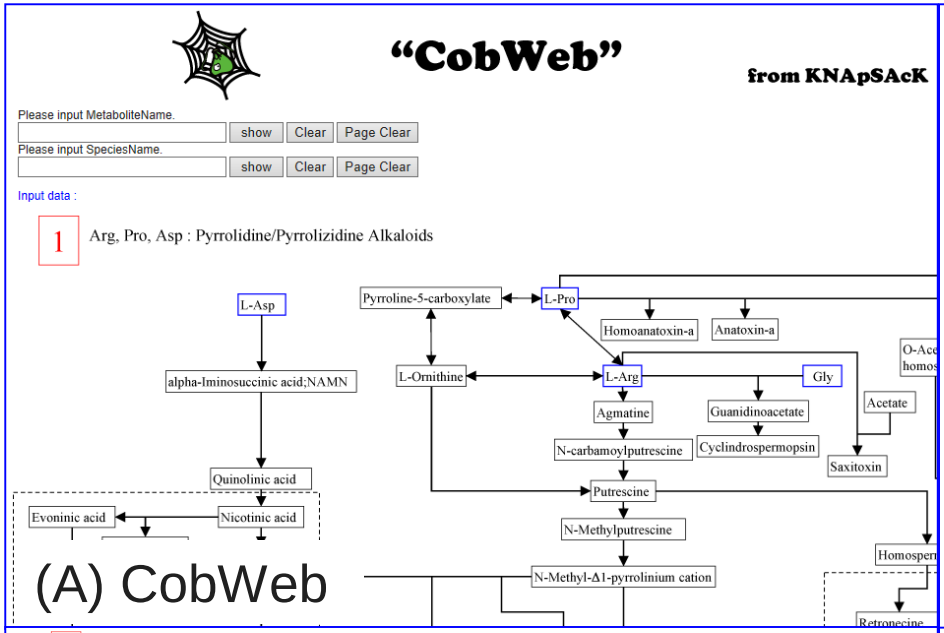
Nicotiana

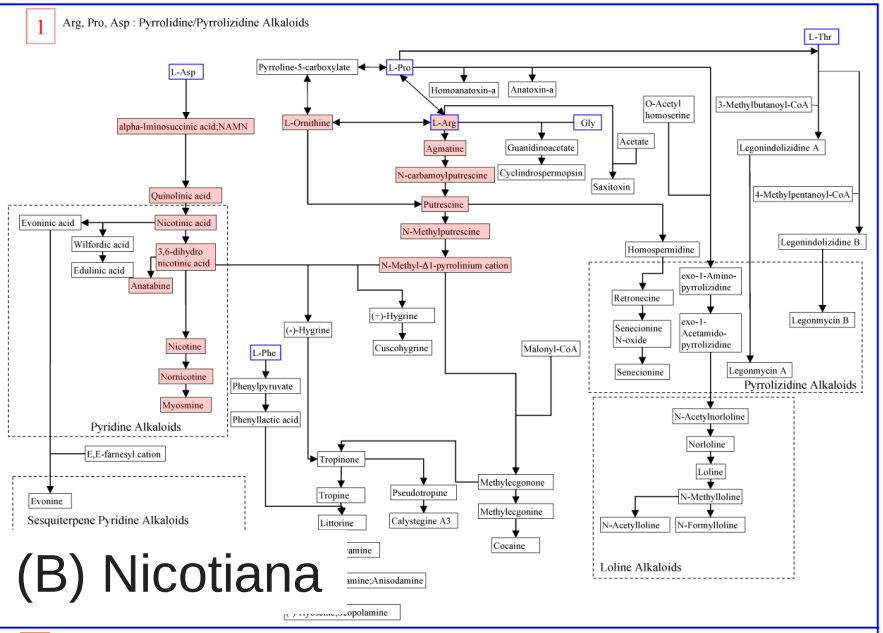
show Clear Page Clear

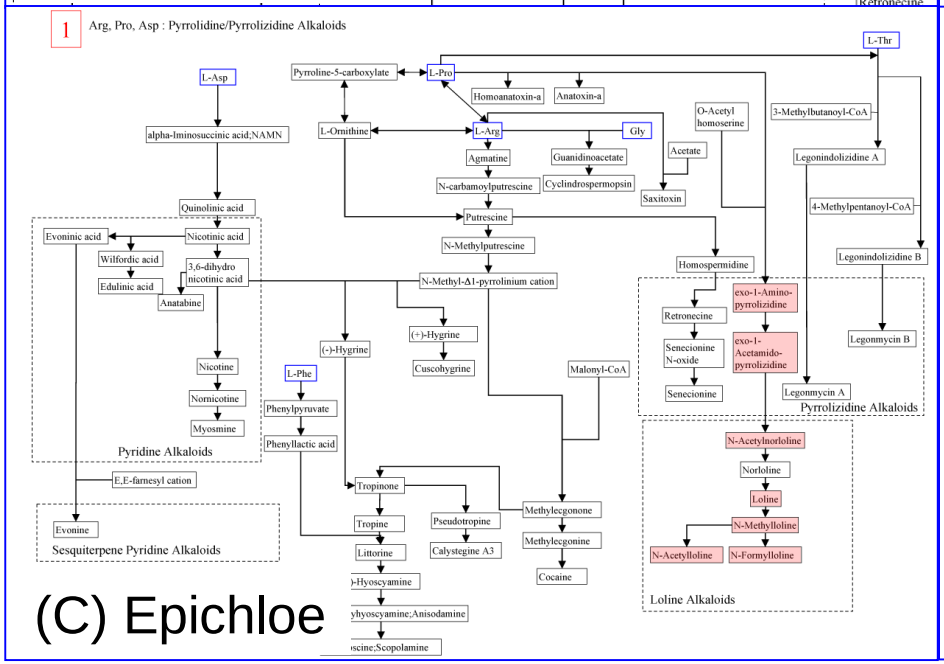
Group of SS			3-phospho glycerate				Pyruvate				Phosphoenol pyruvate								Oxaloacetate			alpha-Ketoglutarate				Terpenes					TCA cycle		Fatty acid		Nucleic acids
Starting substance (SS)			Gly	L-Ser	D-Ser	L-Cys	L-Ala	D-Ala	L-Leu	L-Val	L-Trp	D-Trp	L-Tyr	L-Phe	Anthranilate	N-Formyl anthranilate	Indole-3-glycerol phosphate	O-Methyl-L-tyrosine	L-Asp	L-Thr	L-Lys	L-Arg	L-His	L-Pro	L-Gln	IPP	DMAPP	Secologanin	GGPP	Cholesterol	Acetyl CoA	Oxaloacetate	Malonyl-CoA	Acetoacetyl CoA	Adenine
Pathway Map ID	Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids	1	●											●					●	●		●		●											
	Trp: Ergot Alkaloids	2					●			●	●													●		●									
	Trp: Monoterpenoid Indole Alkaloids	3									●																	●							
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	4									●													●											
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	5									●																				●		●		
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	6									●																					●			
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	7									●																●								
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	8									●																							●	
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	9			●	●	●			●	●																								
	Trp + DMAPP, Trp + Pro: Trp + Oxaloacetate: Trp (Simple Indole Alkaloids), Trp + His	10									●												●												
	Tyr: Betanin group: Tyr + Phe, Colchicine	11											●	●																					
	Tyr: Isoquinoline Alkaloids (Benzylisoquinoline Alkaloids)	12											●																						
	Tyr + Secologanin (Tetrahydroisoquinoline monoterpene alkaloids), Tyr + Ser	13											●															●							
	Tyr + Secologanin (Tetrahydroisoquinoline monoterpene alkaloids), Tyr + Ser	14		●									●																						
	Lys: Quinolizidine, Piperidine Alkaloids, : Indolizidine Alkaloids Lycopodium Alkaloids	15																	●		●	●													
	Phe + Ser	16		●										●																						
	Ala + Acetyl-CoA + Malonyl-CoA	17					●																									●		●		
	His: Imidazole alkaloid(histidine derivatives)	18							●														●													
	Anthraniate(Acridine Alkaloids)	19													●			●	●																	
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	20									●				●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	21												●	●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	22					●	●			●				●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	23					●			●		●			●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	24					●					●				●																			
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	25	●								●				●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	26	●												●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	27													●									●											
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	28													●																		●		
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	29													●										●										
	Indole-3-glycerol phosphate (Indole-diterpene alkaloids)	30															●												●						
	Steroidal alkaloids	31																												●					
	Purine alkaloids	32																																	●

1 Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids




"CobWeb"
from KNApSAcK
Please input MetaboliteName.
show
Clear
Page Clear
Please input SpeciesName.
show
Clear
Page Clear
Input data :
1
Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids
L-Asp
alpha-Iminosuccinic acid;NAMN
Quinolinic acid
Evoninic acid
Nicotinic acid
Pyrroline-5-carboxylate
L-Pro
Homoanatoxin-a
Anatoxin-a
L-Ornithine
L-Arg
Gly
Agmatine
Guanidinoacetate
Acetate
N-carbamoylputrescine
Cyclindrospermopsin
Saxitoxin
Putrescine
N-Methylputrescine
N-Methyl-Δ1-pyrrolinium cation
(A) CobWeb


1
Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids
L-Thr
L-Asp
Pyrroline-5-carboxylate
L-Pro
Homoanatoxin-a
Anatoxin-a
O-Acetyl homoserine
3-Methylbutanoyl-CoA
alpha-Iminosuccinic acid;NAMN
L-Ornithine
L-Arg
Gly
Legonindolizidine A
Agmatine
Guanidinoacetate
Acetate
N-carbamoylputrescine
Cyclindrospermopsin
Saxitoxin
4-Methylpentanoyl-CoA
Quinolinic acid
Putrescine
N-Methylputrescine
Evoninic acid
Nicotinic acid
Wilfordic acid
3,6-dihydro nicotinic acid
Edulinic acid
Anatabine
N-Methyl-Δ1-pyrrolinium cation
Homospermidine
Legonindolizidine B
exo-1-Amino-pyrrolizidine
Retronecine
Senecionine N-oxide
exo-1-Acetamido-pyrrolizidine
Legonmycin B
Senecionine
Legonmycin A
Pyrrolizidine Alkaloids
Nicotine
Nornicotine
Myosmine
(-)-Hygrine
(+)-Hygrine
Cuscohygrine
L-Phe
Phenylpyruvate
Phenyllactic acid
Malonyl-CoA
Tropinone
Tropine
Pseudotropine
Methylecgonone
Calystegine A3
Methylecgonine
Littorine
Cocaine
N-Acetylnorloline
Norloline
Loline
N-Methylloline
N-Acetylloline
N-Formylloline
Loline Alkaloids
Pyridine Alkaloids
E,E-farnesyl cation
Evonine
Sesquiterpene Pyridine Alkaloids
(B) Nicotiana


1
Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids
L-Thr
L-Asp
Pyrroline-5-carboxylate
L-Pro
Homoanatoxin-a
Anatoxin-a
O-Acetyl homoserine
3-Methylbutanoyl-CoA
alpha-Iminosuccinic acid;NAMN
L-Ornithine
L-Arg
Gly
Legonindolizidine A
Agmatine
Guanidinoacetate
Acetate
N-carbamoylputrescine
Cyclindrospermopsin
Saxitoxin
4-Methylpentanoyl-CoA
Quinolinic acid
Putrescine
N-Methylputrescine
Evoninic acid
Nicotinic acid
Wilfordic acid
3,6-dihydro nicotinic acid
Edulinic acid
Anatabine
N-Methyl-Δ1-pyrrolinium cation
Homospermidine
Legonindolizidine B
exo-1-Amino-pyrrolizidine
Retronecine
Senecionine N-oxide
exo-1-Acetamido-pyrrolizidine
Legonmycin B
Senecionine
Legonmycin A
Pyrrolizidine Alkaloids
Nicotine
Nornicotine
Myosmine
(-)-Hygrine
(+)-Hygrine
Cuscohygrine
L-Phe
Phenylpyruvate
Phenyllactic acid
Malonyl-CoA
Tropinone
Tropine
Pseudotropine
Methylecgonone
Calystegine A3
Methylecgonine
Littorine
Cocaine
N-Acetylnorloline
Norloline
Loline
N-Methylloline
N-Acetylloline
N-Formylloline
Loline Alkaloids
Pyridine Alkaloids
E,E-farnesyl cation
Evonine
Sesquiterpene Pyridine Alkaloids
(C) Epichloe

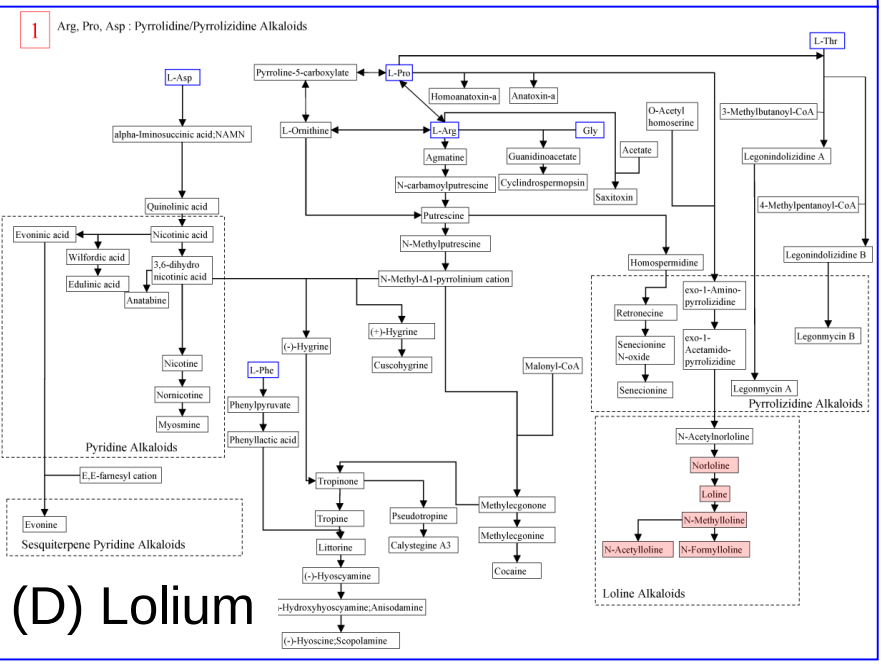

1
Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids
L-Thr
L-Asp
Pyrroline-5-carboxylate
L-Pro
Homoanatoxin-a
Anatoxin-a
O-Acetyl homoserine
3-Methylbutanoyl-CoA
alpha-Iminosuccinic acid;NAMN
L-Ornithine
L-Arg
Gly
Legonindolizidine A
Agmatine
Guanidinoacetate
Acetate
N-carbamoylputrescine
Cyclindrospermopsin
Saxitoxin
4-Methylpentanoyl-CoA
Quinolinic acid
Putrescine
N-Methylputrescine
Evoninic acid
Nicotinic acid
Wilfordic acid
3,6-dihydro nicotinic acid
Edulinic acid
Anatabine
N-Methyl-Δ1-pyrrolinium cation
Homospermidine
Legonindolizidine B
exo-1-Amino-pyrrolizidine
Retronecine
Senecionine N-oxide
exo-1-Acetamido-pyrrolizidine
Legonmycin B
Senecionine
Legonmycin A
Pyrrolizidine Alkaloids
Nicotine
Nornicotine
Myosmine
(-)-Hygrine
(+)-Hygrine
Cuscohygrine
L-Phe
Phenylpyruvate
Phenyllactic acid
Malonyl-CoA
Tropinone
Tropine
Pseudotropine
Methylecgonone
Calystegine A3
Methylecgonine
Littorine
Cocaine
(-)-Hyoscyamine
(-)-Hydroxyhoscyamine;Anisodamine
(-)-Hyoscine;Scopolamine
N-Acetylnorloline
Norloline
Loline
N-Methylloline
N-Acetylloline
N-Formylloline
Loline Alkaloids
Pyridine Alkaloids
E,E-farnesyl cation
Evonine
Sesquiterpene Pyridine Alkaloids
(D) Lolium

Fig.5

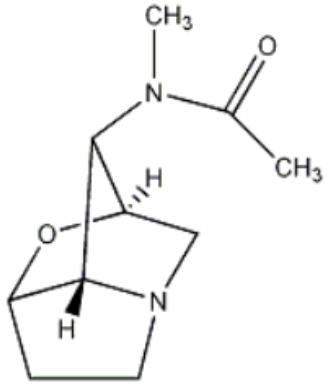
Metabolite Information

Name	N-Acetylloline NAL
Formula	C10H16N2O2
Mw	196.12117777
CAS RN	4914-36-7
C_ID	C00023611 3D , Twins (%)
InChIKey	YIZSKLHCDNIMHK-YEXJWTGLNA-N
InChICode	InChI=1S/C10H16N2O2/c1-6(13)11(2)9-8-5-12-4-3-7(14-8)10(9)12/h7-10H,3-5H2,1-2H3/t7-,8+,9-,10+/m0/s1
SMILES	O1[C@@H]2[C@H]3N(C[C@@H]1[C@@H]3N(C)C(=O)C)CC2
Start Substs in Alk. Biosynthesis (Prediction)	

Organism

Kingdom	Family	Species	Reference
Fungi	Clavicipitaceae	Epichloe coenophiala e19	Ref.
Fungi	Clavicipitaceae	Epichloe coenophiala e4163	Ref.
Fungi	Clavicipitaceae	Epichloe festucae E2368	Ref.
Fungi	Clavicipitaceae	Epichloe uncinata e167	Ref.
Fungi	Hypocreaceae	Acremonium coenophialum	Ref.
Plantae	Poaceae	Festuca arundinacea W	Ref.
Plantae	Poaceae	Lolium arundinaceum	Ref.
Plantae	Poaceae	Lolium cuneatum	Ref.
Plantae	Poaceae	Lolium pratense	Ref.
-	-	Neotyphodium pratense	Ref.
-	-	Neotyphodium siegelii	Ref.
-	-	Neotyphodium uncinatum	Ref.
-	-	Neotyphodium coenophialum	Ref.
-	-	Neotyphodium occultans	Ref.

Structural formula

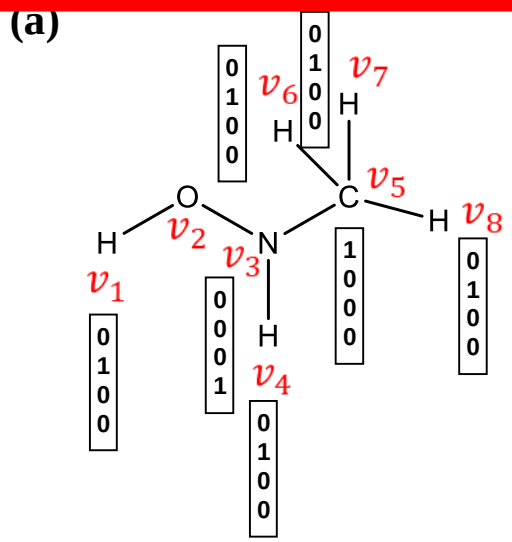


zoom in

Fig.6

578 alkaloids with known starters, but 11,000 structure-known alkaloids with unknown starters.

Prediction of starting substances of alkaloids based on Graph Convolutional Neural Networks (GCNN



BMC Bioinf., (2019) in press

Classification of alkaloids according to the starting substances of their biosynthetic pathways using graph convolutional neural networks

Ryohei Eguchi[1,2], Naoaki Ono[1,2*], Aki Hirai (Morita)[1], Tetsuo Katsuragi[3], Satoshi Nakamura[1,2], Ming Huang[1], Md. Altaf-Ul-Amin[1] and Shigehiko Kanaya[1,2]

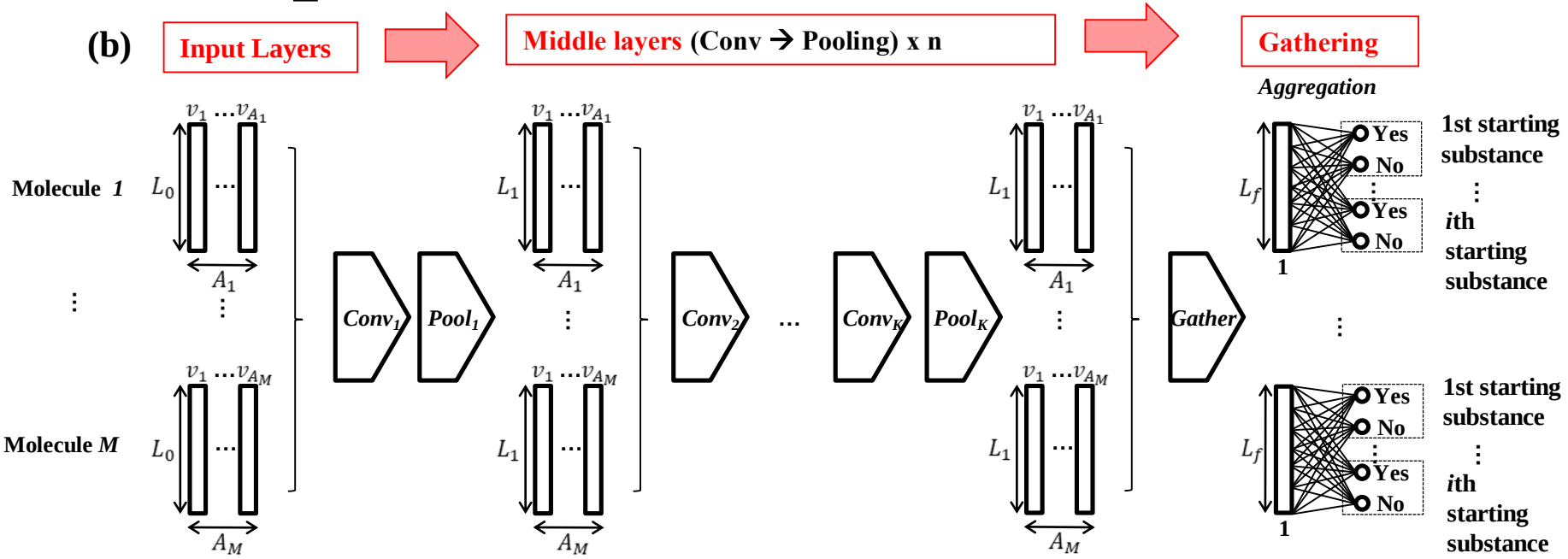
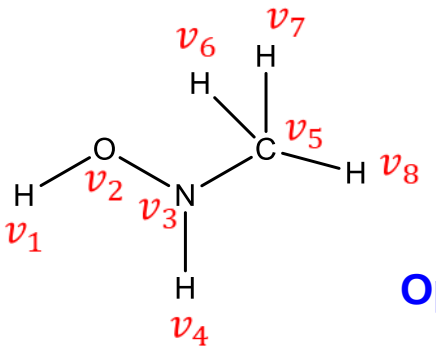


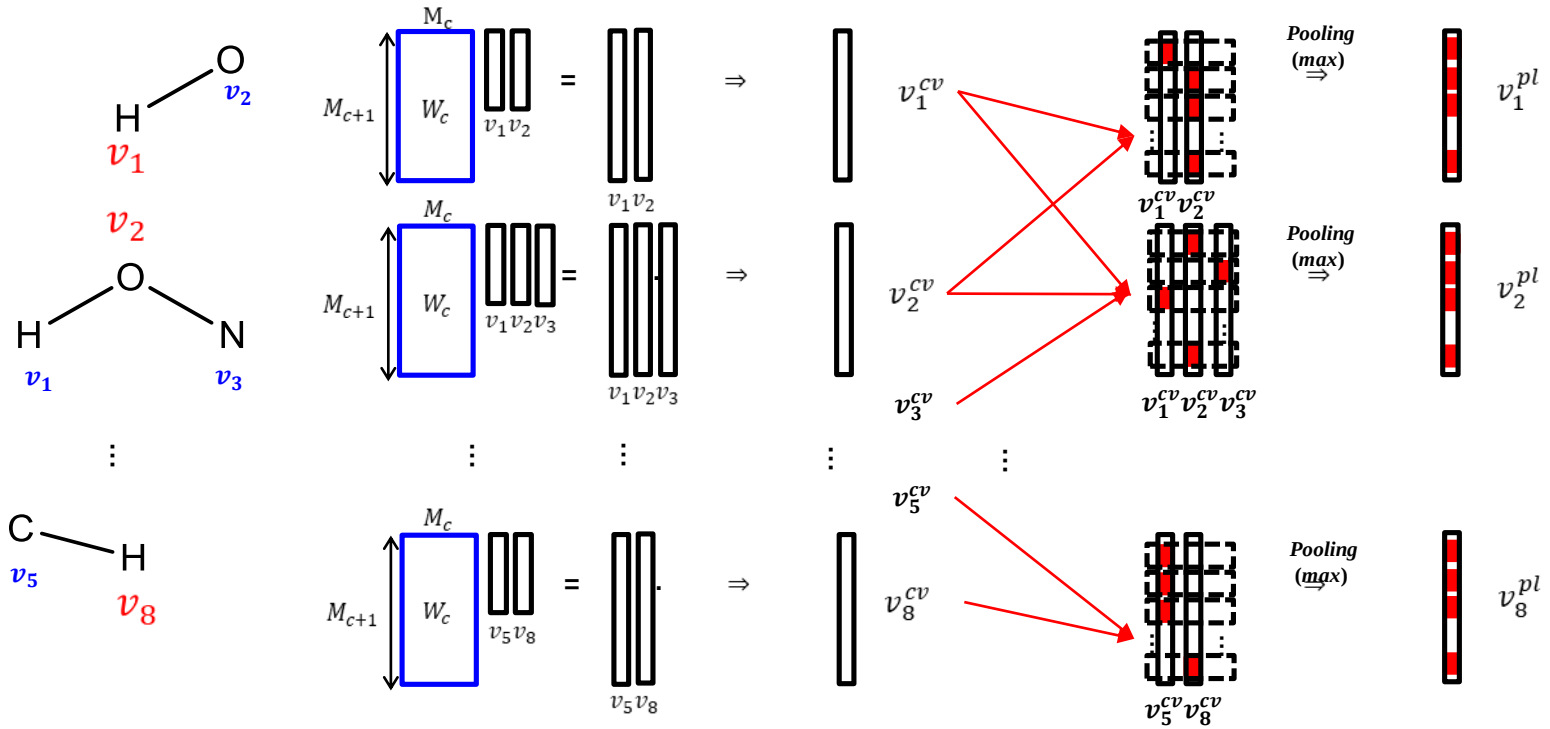
Figure 1 schematic diagram of MGCNN. (a) Explanation of one hot vectors for a molecule, (b) Schematic diagram of MGNN (details are explained in the text). In the one hot vectors for *i*th molecule visualized in (a), A_i in (b) is set as 8.

Graph Convolutional Neural Networks



Optimization of Weight W

Selection of targeted and neighboring

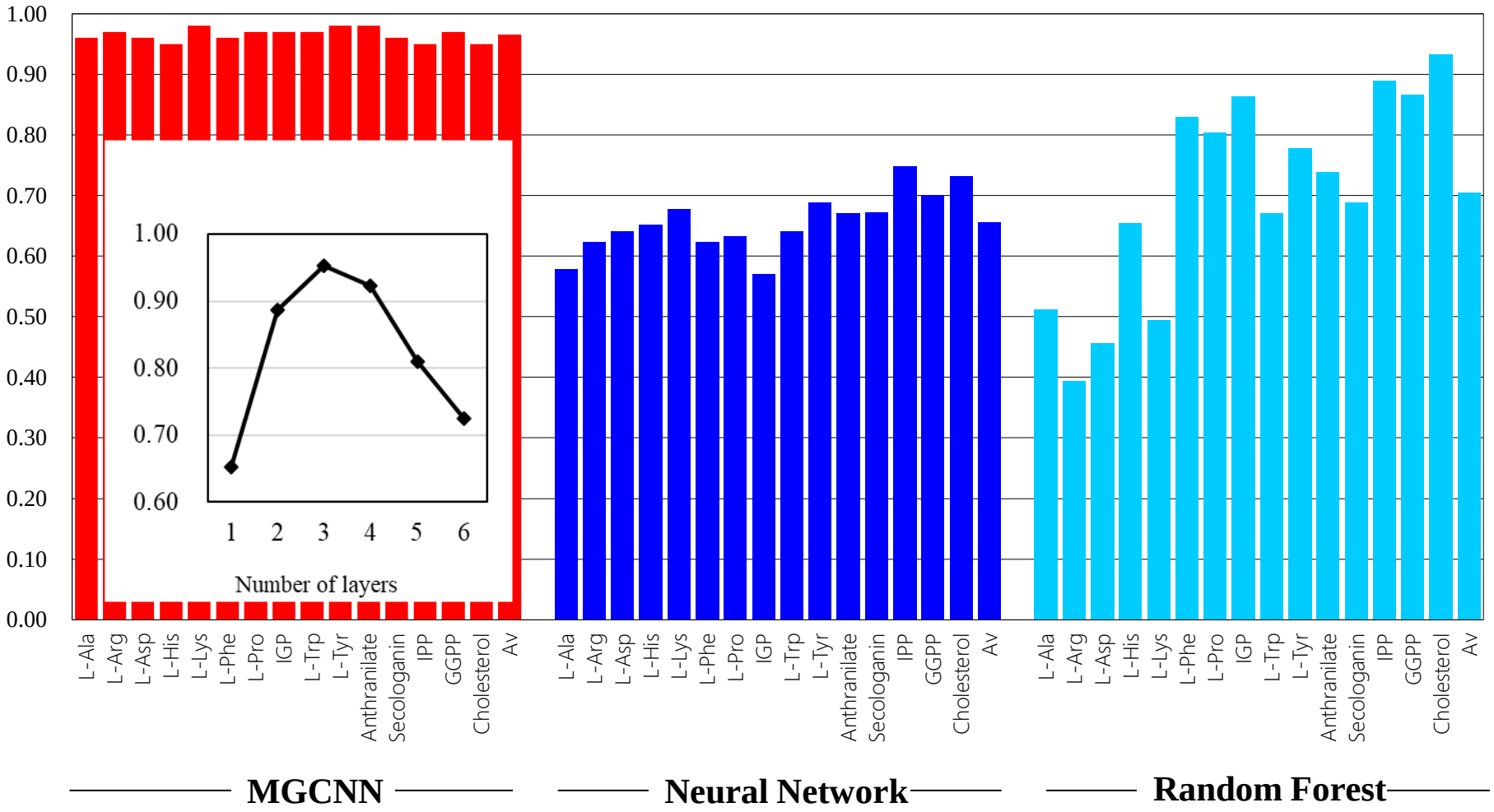


Optimization of matrix Wc

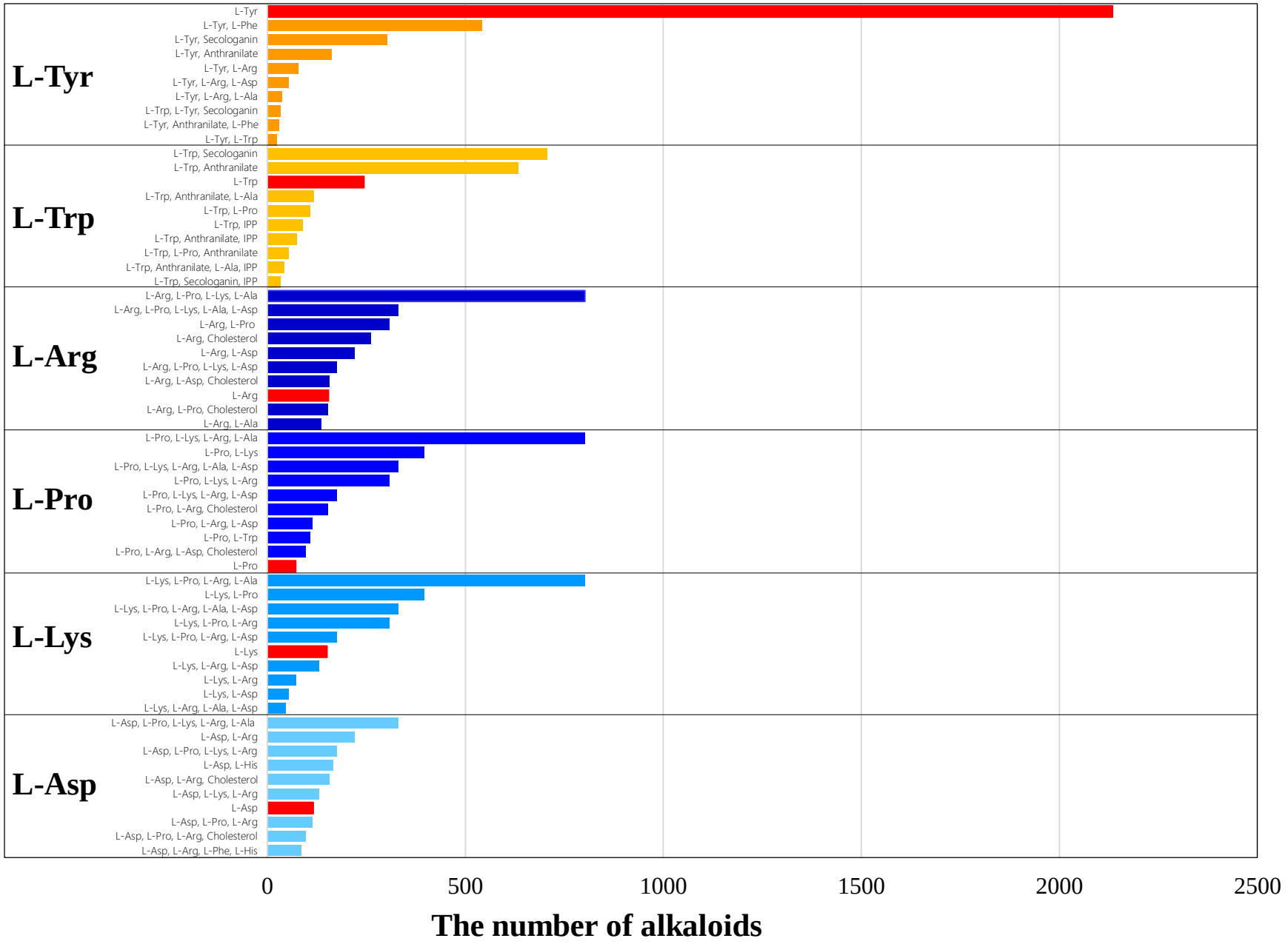
(a) Convolution **(b) Pooling**

Figure2 (a) Convolution and (b) pooling layers

Accuracy for GCNN, Neural Network, and Random Forest



The best 10 combination for six major starting substances in KNApSAcK alkaloids



Biopathways involved in L-Tyrosine

Group of SS		3-phospho glycerate				Pyruvate				Phosphoenol pyruvate							Oxaloacetate			alpha-Ketoglutarate				Terpenes					TCA cycle		Fatty acid		Nucleic acids
Starting substance (SS)		Gly	L-Ser	D-Ser	L-Cys	L-Ala	D-Ala	L-Leu	L-Val	L-Trp	D-Trp	L-Tyr	L-Phe	Anthranilate	N-Formyl anthranilate	Indole-3-glycerol phosphate	O-Methyl-L-tyrosine	L-Asp	L-Thr	L-Lys	L-Arg	L-His	L-Pro	L-Gln	IPP	DMAPP	Secologanin	GGPP	Cholesterol	Acetyl CoA	Oxaloacetate	Malonyl-CoA	Acetoacetyl CoA	Adenine
Pathway Map ID	Arg, Pro, Asp : Pyrrolidine/Pyrrolizidine Alkaloids	1	●										●					●	●		●		●											
	Trp: Ergot Alkaloids	2				●			●	●													●		●									
	Trp: Monoterpenoid Indole Alkaloids	3								●																	●							
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	4								●													●											
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	5								●																			●		●			
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	6								●																				●				
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	7								●																●								
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	8								●																						●		
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	9			●	●	●			●	●																							
	Trp + DMAPP, Trp + Pro; Trp + Oxaloacetate; Trp (Simple Indole Alkaloids), Trp + His	10								●												●												
	Tyr: Betanin group: Tyr + Phe, Colchicine	11										●	●																					
	Tyr: Isoquinoline Alkaloids (Benzylisoquinoline Alkaloids)	12										●																						
	Tyr + Secologanin (Tetrahydroisoquinoline monoterpene alkaloids), Tyr + Ser	13										●															●							
	Tyr + Secologanin (Tetrahydroisoquinoline monoterpene alkaloids), Tyr + Ser	14		●								●																						
	Lys: Quinolizidine, Piperidine Alkaloids, : Indolizidine Alkaloids Lycopodium Alkaloids	15																●		●	●													
	Phe + Ser	16		●									●																					
	Ala + Acetyl-CoA + Malonyl-CoA	17				●																									●			
	His: Imidazole alkaloid(histidine derivatives)	18						●														●												
	Anthraniate(Acridine Alkaloids)	19												●			●	●																
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	20								●				●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	21											●	●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	22				●	●			●				●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	23				●			●		●			●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	24				●					●					●																		
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	25	●							●				●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	26	●											●																				
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	27												●									●											
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	28												●																		●		
	Anthranilate + Trp: + Anthranilate + Trp + Ala: Anthranilate + Phe (PKS)	29												●										●										
	Indole-3-glycerol phosphate (Indole-diterpene alkaloids)	30														●												●						
	Steroidal alkaloids	31																											●					
	Purine alkaloids	32																																●

Predicted Starting substances are reflected in KNApSAcK Core

input word = C00026421

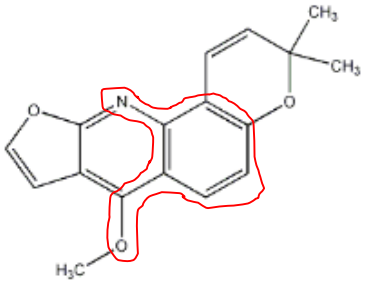
Metabolite Information

Name	Dutadrupine
Formula	C17H15NO3
Mw	281.10519335
CAS RN	80151-78-6
C_ID	C00026421 3D , Twins (%)
InChIKey	BNUBFEHGKQEPIQ-UHFFFAOYSA-N
InChICode	InChI=1S/C17H15NO3/c1-17(2)8-6-10-13(21-17)5-4-11-14(10)18-16-12(7-9-20-16)15(11)19-3/h4-9H,1-3H3
SMILES	c12c(ccc3c1nc1c(c3OC)cco1)OC(C=C2)(C)C
Start Substs in Alk. Biosynthesis (Prediction)	Anthranilate

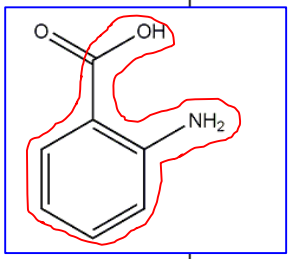
Organism

Kingdom	Family	Species	Reference
Plantae	Rutaceae	Dutaillyea drupacea	Ref.
Plantae	Rutaceae	Melicope erromangensis	Ref.
Plantae	Rutaceae	Melicope semecarpifolia W	Ref.
-	-	Almeidia coerulea	Ref.

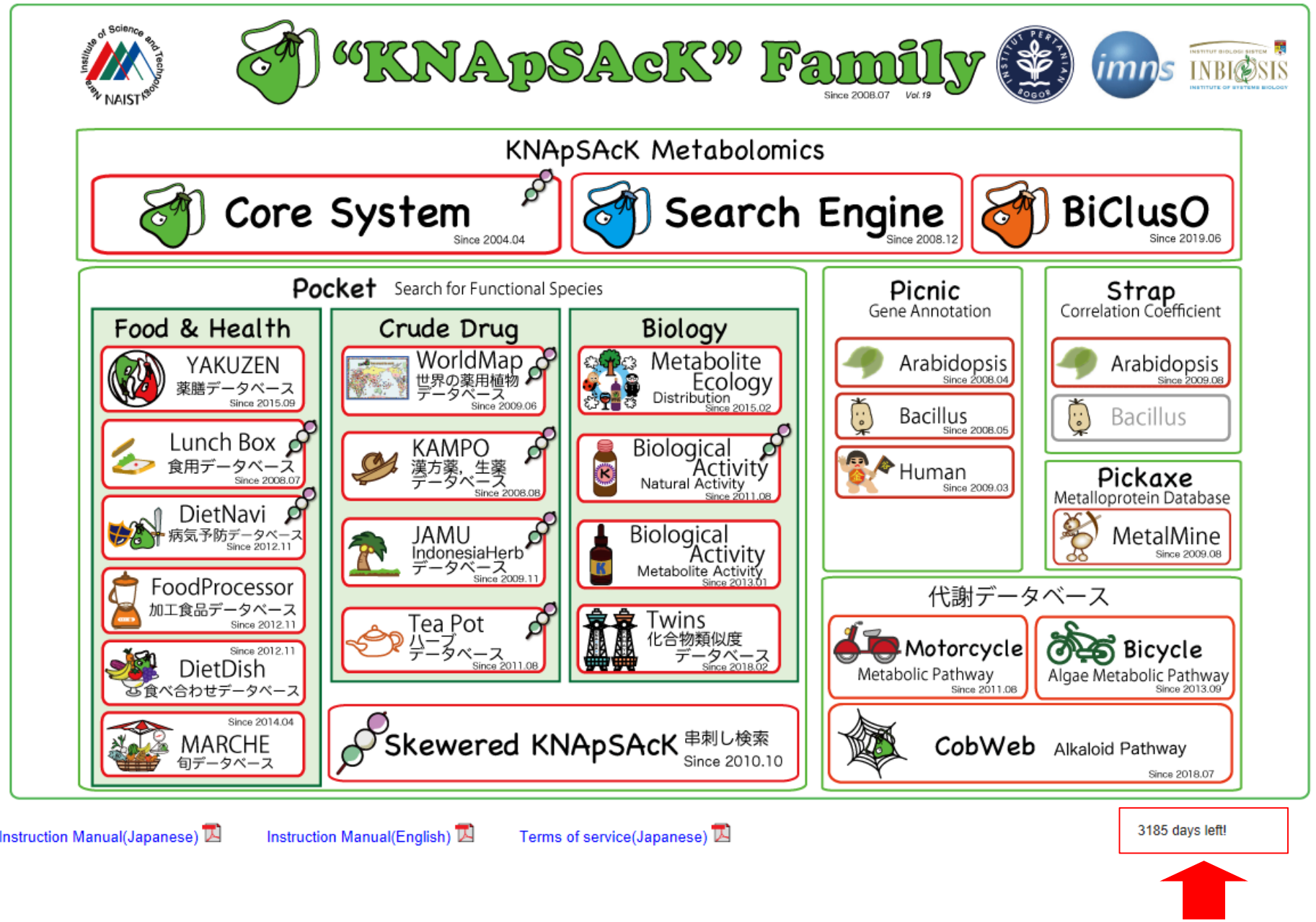
Structural formula



zoom in



Anthranilate



Instruction Manual(Japanese) Instruction Manual(English) Terms of service(Japanese)

3185 days left!

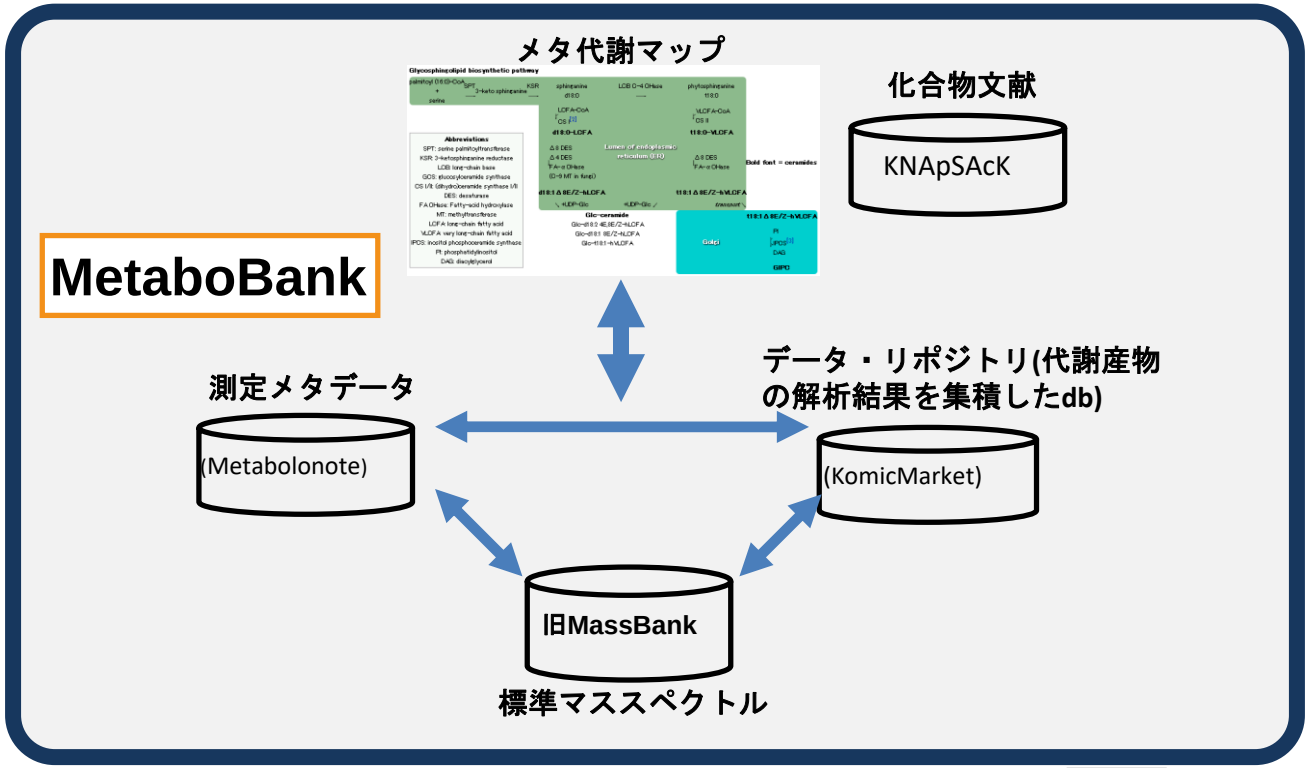
MetaboBankの開発
かずさDNA研のデータ整備

2019年7月12日　統合DBプロジェクト 2019年度サイトビジット（遺伝研）

櫻井 望
国立遺伝学研究所

長崎英樹、平川英樹
かずさDNA研究所

開発の概要

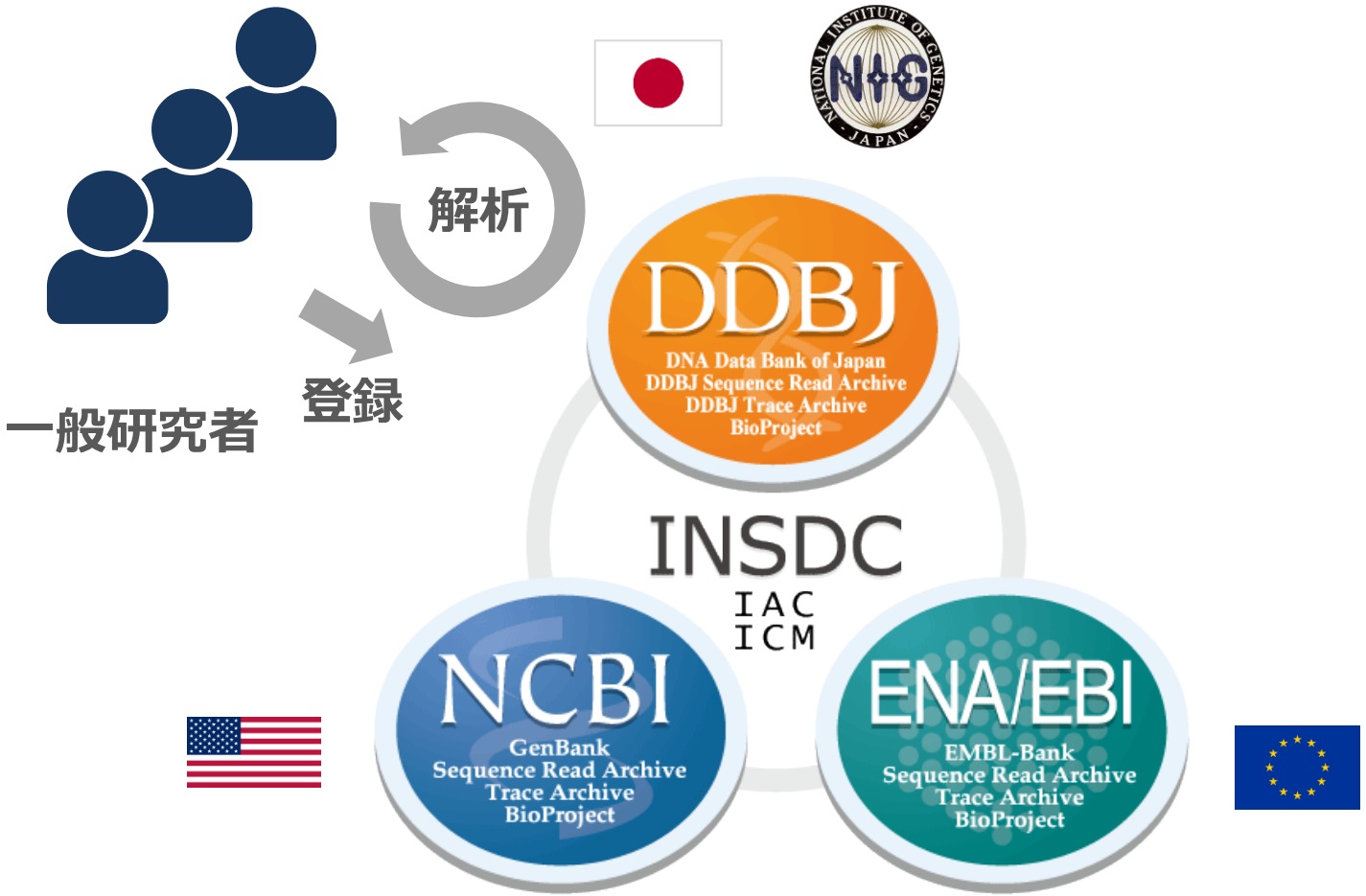


- 実験データ
- メタデータ

かずさDNA研究所

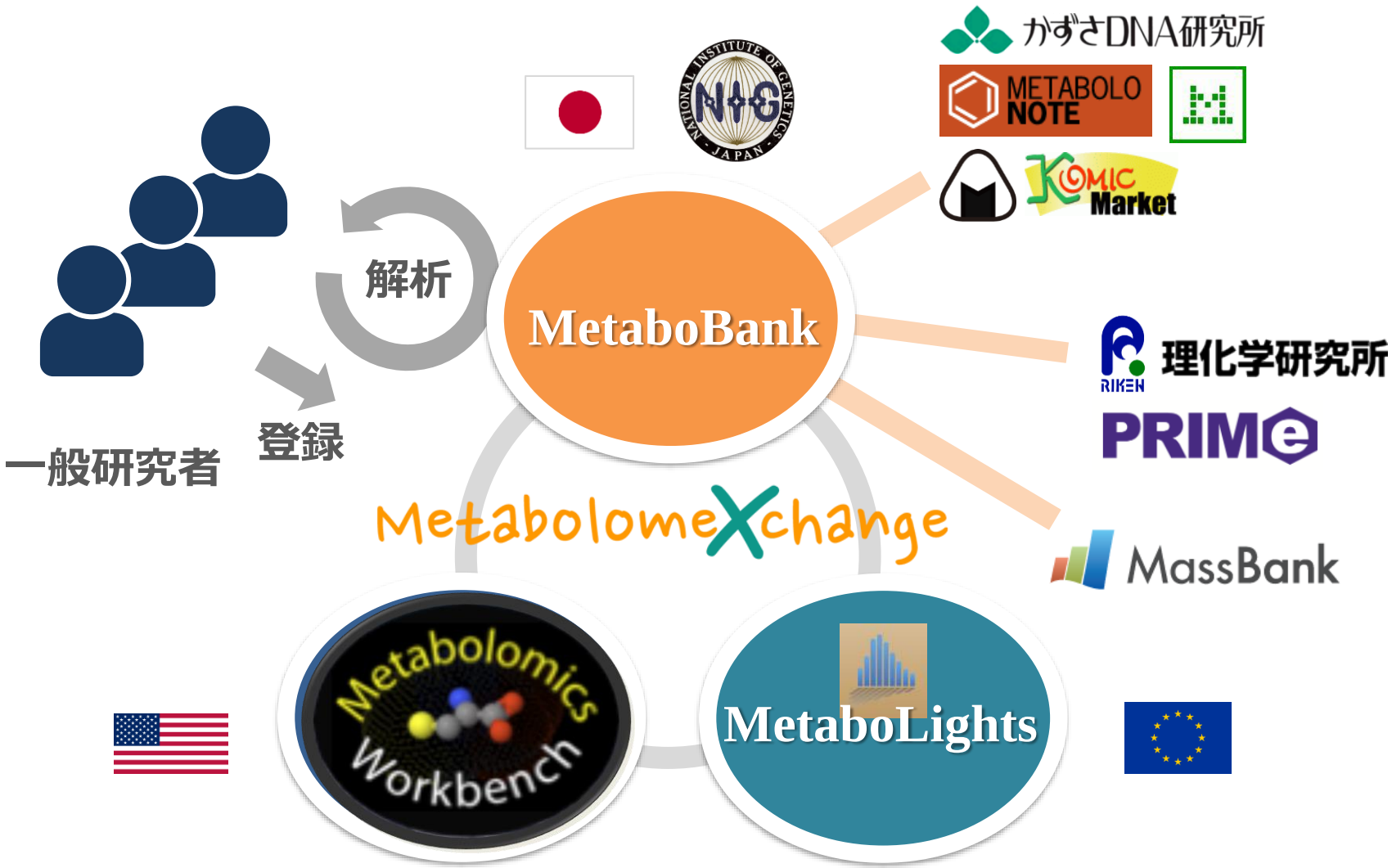
RIKEN
理化学研究所

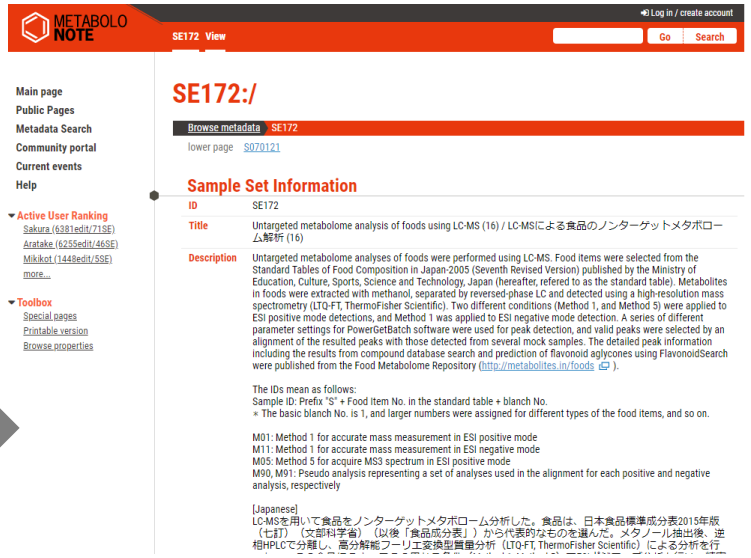
DDBJ：　DNA配列登録の拠点の一つ



INSDC: International Sequence Database Collaboration

メタボロームデータ登録のアジア拠点の構築



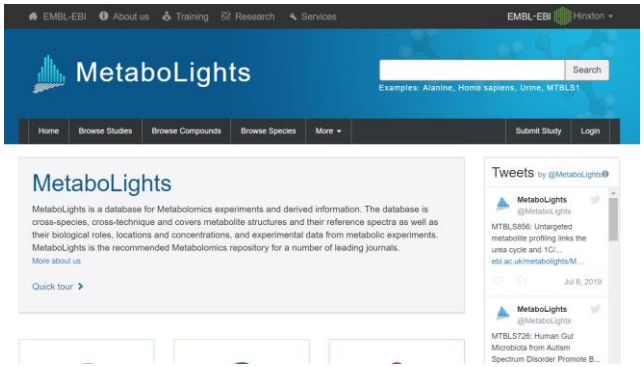


Metabolonote

H29年度 調査研究予算で連携達成

2019年5月14日

EBIミーティング



2019年6月26日

EBI, NCBIミーティング@Metabolomics2019

本日の内容

- MetaboBank インターフェースの開発
- かずさDNA研究所のデータ整備

進捗：　RDF検索（実装テスト）

MetaboBank

Home　Repository　Database　APIs　Statistics　Help

登録・データ管理

Simple Search　Advanced Search

シンプルな全文サーチ

arabidopsis × root ×

Search　AND OR

Search Results

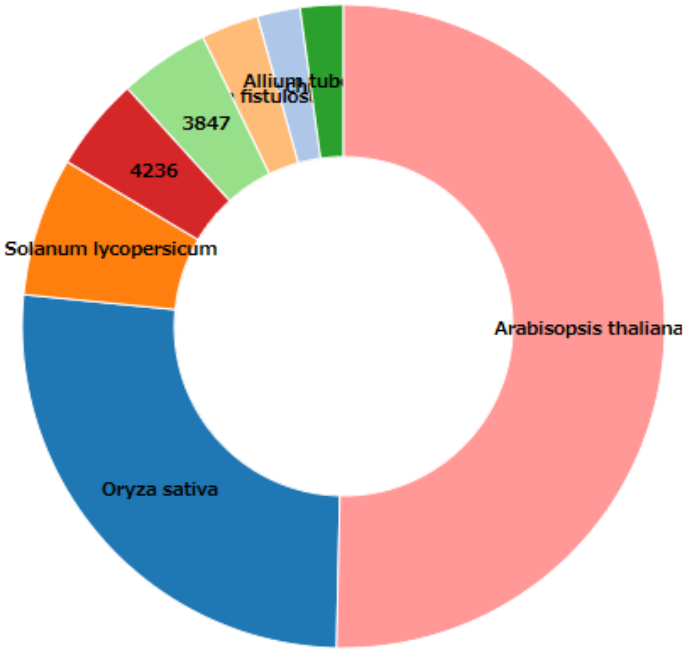
Show 10 entries

Project ID	Sample	Experiments	#Analysis
RPMM0006	Metabolomic correlation-network modules in Arabidopsis based on a graph-clustering approach	[GC-TOF/MS analysis]	54
RPMM0019	Physiological Roles of the beta-Substituted Alanine Synthase Gene Family in Arabidopsis	[Metabolites analyzed by GC-TOF/MS, Metabolites analyzed by CE-TOF/MS]	36

Showing 1 to 2 of 2 entries

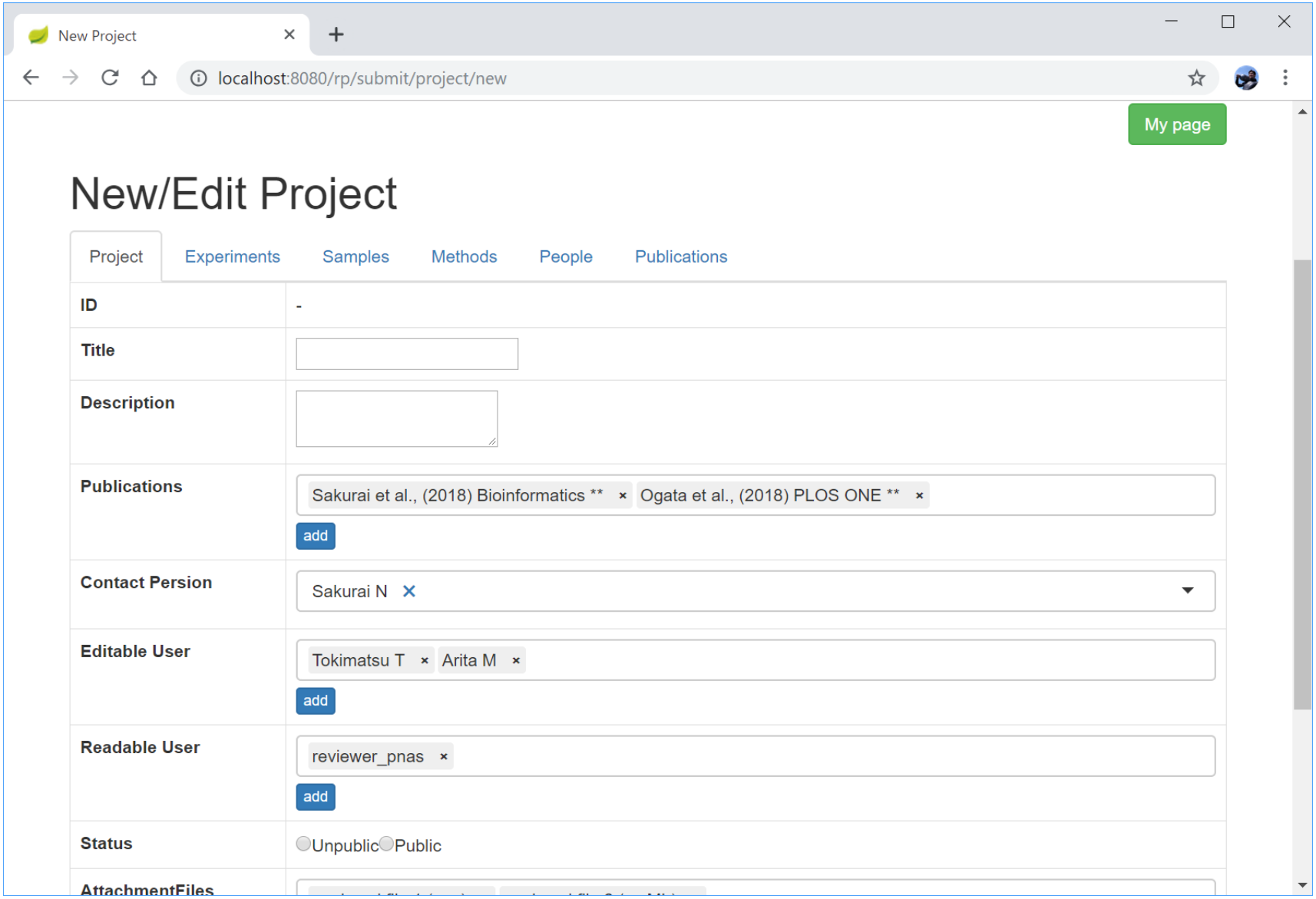
Statistics

Statistical information about this website.



RDFによる柔軟な絞り込み検索

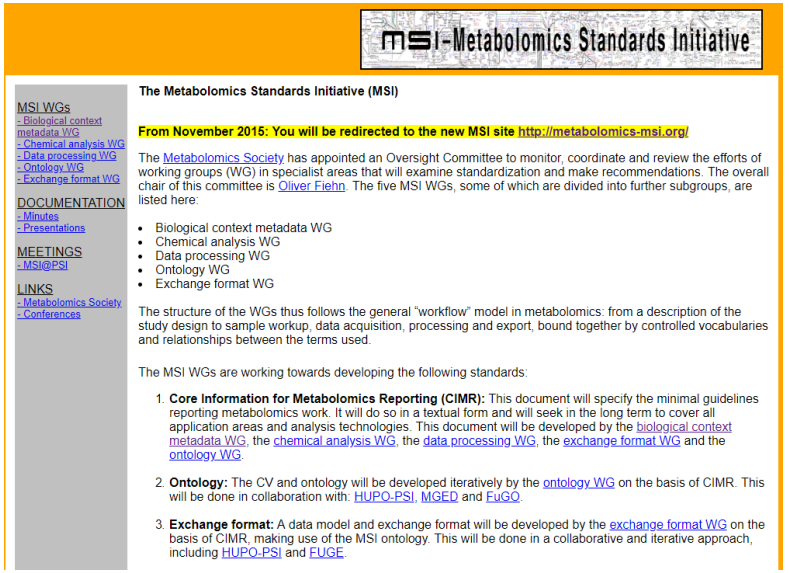
進捗：　データ入力UIの検討（Mockup）



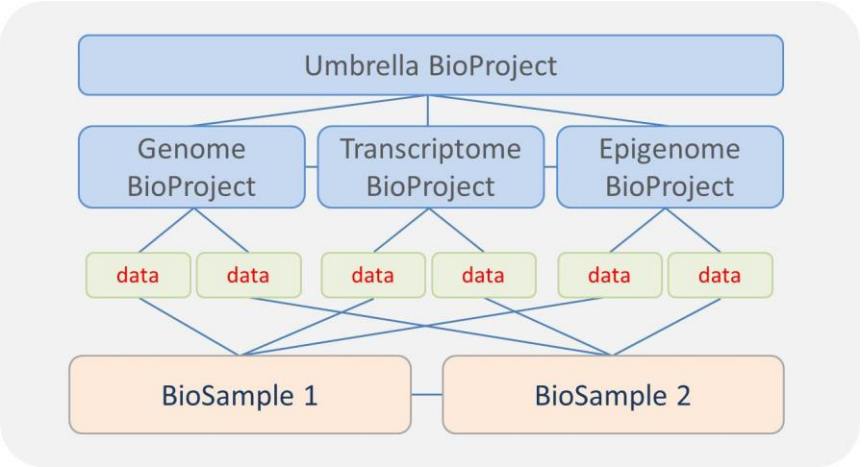
データ登録の課題

- Metabolomics Standards Initiative (MSI)準拠

2007~



- BioSample/BioProject

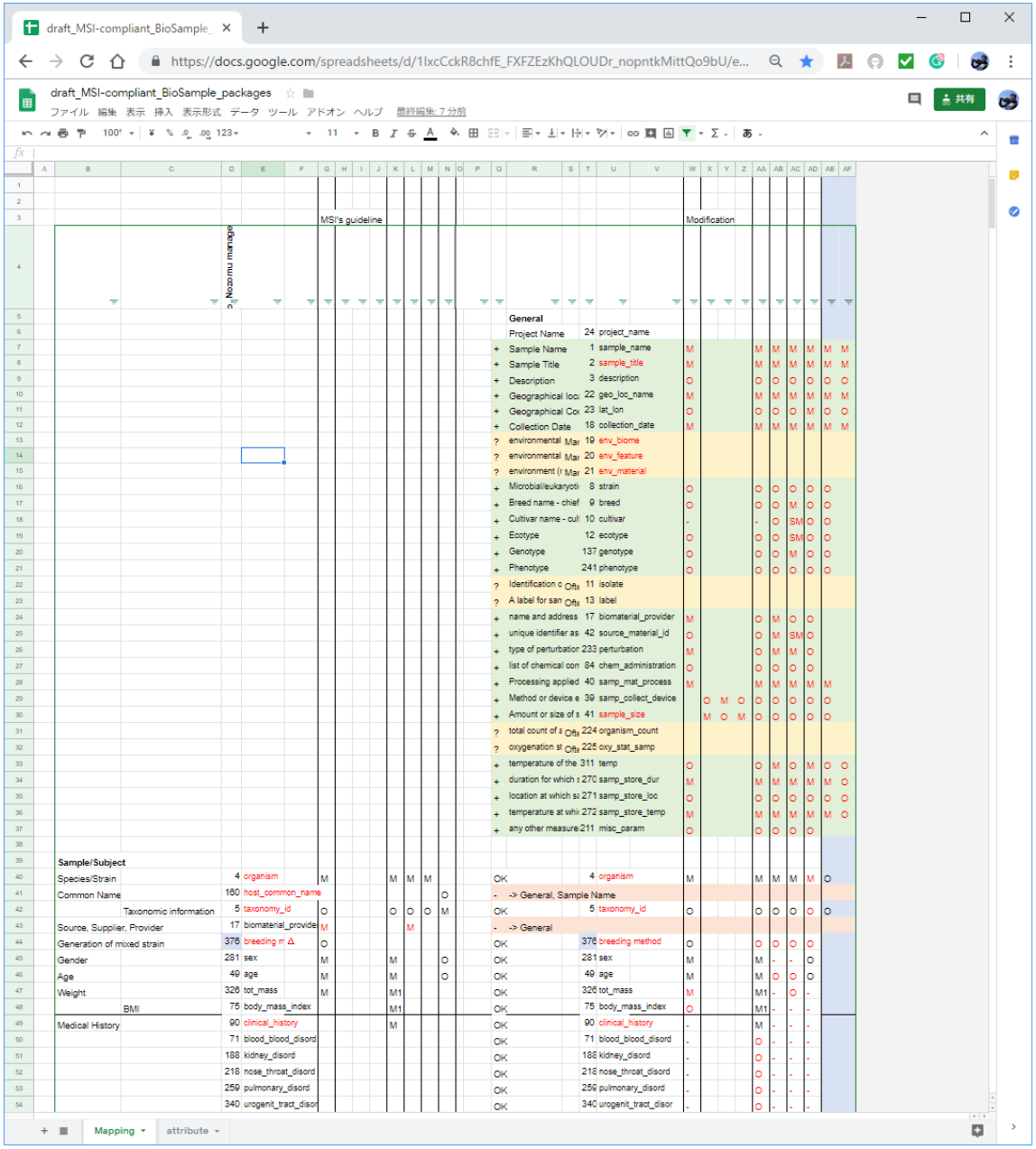


- RDF化

Comparison of the repositories (1)

	MetaboLights	Metabolomics Workbench	MetaboBank (our plan)
BioSample support	no	no	yes
MSI's reporting guideline	recommended	recommended	compliant
Metadata Input Interfaces	File-based (ISA creator), Web GUI (beta)	Web GUI, File-based (flat-file)	File-based (excel), Web GUI
File transfer	FTP client, Aspera client (project-based management)	FTP client, Aspera client	SFTP, Aspera, HTTPS, (Under investigation)
Group editing	no	no	yes
Private access for peer review	yes (on another server)	(probably no)	yes
Embargo	yes (up to 60 month)	yes (no limitation)	yes
Local installation	yes	(probably no)	yes (for duplication of the system)

MSIに準拠したBioSampleパッケージの作成



BioSample：
必須項目は少ないが、400近い項目が定義され、ゲノム情報等に使われている

MSI:
細かい必須項目が存在

Comparison of the repositories (1)

	MetaboLights	Metabolomics Workbench	MetaboBank (our plan)
BioSample support	no	no	yes
MSI's reporting guideline	recommended	recommended	compliant
Metadata Input Interfaces	File-based (ISA creator), Web GUI (beta)	Web GUI, File-based (flat-file)	File-based (excel), Web GUI
File transfer	FTP client, Aspera client (project-based management)	FTP client, Aspera client	SFTP, Aspera, HTTPS, (Under investigation)
Group editing	no	no	yes
Private access for peer review	yes (on another server)	(probably no)	yes
Embargo	yes (up to 60 month)	yes (no limitation)	yes
Local installation	yes	(probably no)	yes (for duplication of the system)

Comparison of the repositories (2)

	MetaboLights	Metabolomics Workbench	MetaboBank (our plan)
API	(probably no)	yes (RESTful)	yes (RESTful)
Search study by	technology (MS/NMR), organism, organism part, status (public/submitted)	subject type, species, project/study title, institution, submitted year, analysis type (MS/NMR), free text	RDF-based. technology, organism, organism part, free text,
Search compound by	organism, organism part, pathways, reactions, technology (MS/NMR), status (public/submitted), compound name	species, compound name, MS ion mode (positive/negative), compound substructure, compound class, m/z (adduct, compound classes), retention time,	organism, organism part, compound name, m/z, retention time, MSn
Data analysis	no	normalization, clustering, univariate- and multivariate analysis, mapping to the pathways…	yes (not planed in detail yet)

Comparison of the repositories (3)

	MetaboLights	Metabolomics Workbench	MetaboBank (our plan)
Data Structure	Study Assay	Project Study (per Analysis) Sample	Project Experiment Sample Measurement
Definition of the experimental run	Sample name + protocols	Sample x Analysis	Sample x Measurement
multilingual	no	no	yes
ID	MTBLS (study) MTBLC (compound)	PR (project), ST (study), SA (sample), TR (treatment), SP (sample prep), CO (collection), AN (analysis), NM (NMR), MS (MS)	Measurement, BioSample, BioProject

Metadata input framework

Definition of the metadata attributes

JSON Schema

```
{
  "title": "Sample",
  "type": "object",
  "required": [
    "sample_name",
    "sample_title",
    "collection_date",
    "organism"
  ],
  "properties": {
    "sample_name": {
      "title" : "Sample Name",
      "type": "string",
      "minLength": 1,
```

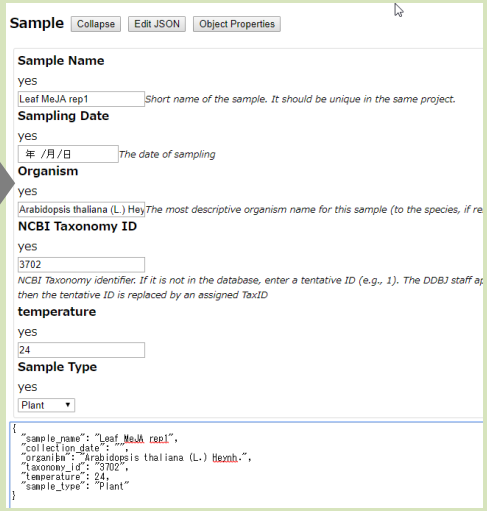
Definition of required attributes

JSON

```
[
    {
        "condition": {
            "property_name": "sample_type",
            "property_value": "Plant",
            "condition": "equals"
        },
        "required": [
            "temperature",
            "sample_name",
            "taxonomy_id",
            "collection_date",
            "organism"
        ],
        "required-either-of": [
        ],
        "optional": [
```

Automatic generation of data input forms

JavaScript



ANGULAR

AJV Json Schema Validator

Metadata JSON

```
[
    {
        "condition": {
            "property_name": "sample_type
            "property_value": "Plant",
            "condition": "equals"
        },
        "required": [
            "temperature",
            "sample_name",
            "taxonomy_id",
            "collection_date",
            "organism"
        ],
        "required-either-of": [
        ],
        "optional": [
```

Definition of ontologies

JSON or OWL

RDF

XML

NoSQL db

elastic

DDBJ BioSample/ BioProject registration system

今後の予定

９月

MetaboBank（データレポジトリ部）プロトタイプ（内部テスト可）の完成

12月

かずさ・理研データの搭載、負荷試験、セキュリティーテスト

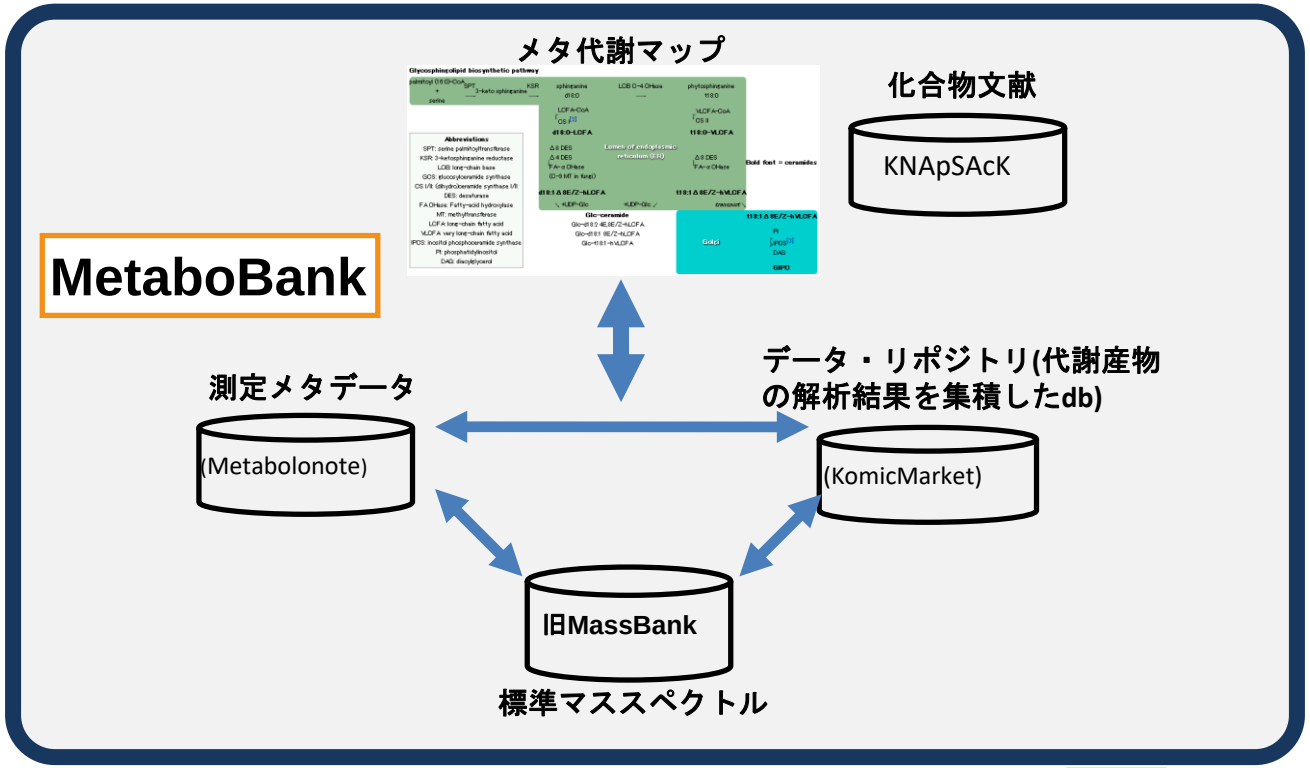
3月

一般データの受け入れ開始

本日の内容

- MetaboBank インターフェースの開発
- かずさDNA研究所のデータ整備

かずさ研の役割



- 実験データ
- メタデータ

かずさDNA研究所

RIKEN
理化学研究所

Metabolonoteメタデータ移植作業

http://metabolonote.kazusa.or.jp/

Ara *et al*. Front Bioeng Biotechnol. 2015 Apr 7;3:38.

- 植物、微生物等：82件
- 食品：16件
- その他：6件

計104件

- 理研関連：82件（理研で整備）

RDFへの変換

- RDF (Resource Description Framwork)

- Metabolonoteメタデータからのデータ構築

Analytical Method Details Information

ID	MS01
Title	LC-FT-MS, ESI, Positive (method 1)
Instrument	Agilent1100 HPLC (Agilent), LTQ-FT (Thermo Fisher Scientific)
Instrument Type	LC-FTICR-MS
Ionization	ESI
Ion Mode	Positive
Description	Frozen powder of material was extracted with three times volumes of methanol containing 25 µM of 7-hydroxy-5-methylflavone as an internal standard. After homogenization using a Mixer Mill MM 300 (Quiagen) at 25 Hz for 2 min twice, homogenates were centrifuged (17,400 g, 5 min, 4 °C). The supernatant was filtered through 0.2 µm PTFE membrane (Millipore). Hydrophobic compounds in the filtrate were removed by absorbing to C18 silica column (MonoSpin C18, GL Science, Tokyo, Japan). Mock sample was prepared with the same procedure without adding the plant material. 20 µl of methanol solution was applied to a TSK-gel column ODS-100V (4.6 x 250 mm, 5 µm; TOSOH). Water (HPLC grade; solvent A) and acetonitrile (HPLC grade; solvent B) were used as the mobile phase with 0.1% v/v formic acid added to both solvents. The gradient program was as follows: 3% B (0 min), 97% B (90 min), 97% B (100 min), 3% B (100.1 min), and 3% B (107 min). The flow rate was set to 0.25 ml/min (0-100 min) and 0.5 ml/min (100.1-107 min). The flow rate was set to 0.5 ml/min, and the column oven temperature was set at 40 °C. Compounds are detected in ESI-positive mode in the range m/z 100-1500. Multistage MSn analyses were carried out using collision induced dissociation (CID) in a linear ion trap detector at a normalized collision energy of 35.0% and an isolation width of 4.0 (m/z), and monitored by both ion trap detector and FT-ICR detector at 25,000 (at m/z 400) mass resolution. The ESI setting was as follows: spray voltage 4.0 kV and capillary temperature 300 °C. Nitrogen sheath gas and auxiliary gas were set at 40 and 15 arbitrary units, respectively. To monitor HPLC elution, a photodiode array detector was used in the wavelength range 200−650 nm. Mass scan events were set as follows (referred as method 1): full mass scan with FT-ICR in resolution 100,000 and MS2 scans for the most intense 5 ions of full mass scan with ion trap (IT). A dynamic exclusion setting was applied as follows: repeat count, 3; repeat duration, 30 sec; exclusion list size, 500; margin, 10 ppm, and exclusion duration, 20. The data were acquired by Xcalibur software version 2.07 (Thermo Fisher Scientific).

サンプル調整

HPLC

マス

共通の項目(クラス)でリンクされた50シートのエクセルファイルからRDF化
合わせて正誤確認(記述ミス、著者所属確認等)

Chromatography クロマトグラフィ Chromatography	comment コメント rdfs:comment rdf:langString	description 説明 dcterms:description rdf:langString	temperature gradient 温度勾配 temperatureGradient xsd:string	column type カラムの種類 columnType xsd:string	column temperature カラム温度 columnTemperature Temperature	column pressure カラム圧力 columnPressure Pressure	column name カラム名 columnName xsd:string
pm_chromato:SE112_MS01	"Water (HPLC grade; solvent A) and acetonitrile (HPLC grade; solvent B) were used as the mobile phase with 0.1%	[temp:40C]		"TSK-gel column ODS-100V (4.			
pm_chromato:SE112_MS05	"Water (HPLC grade; solvent A) and acetonitrile (HPLC grade; solvent B) were used as the mobile phase with 0.1%	[temp:40C]		"TSK-gel column ODS-100V (4.			
pm_chromato:SE112_MS11	"Water (HPLC grade; solvent A) and acetonitrile (HPLC grade; solvent B) were used as the mobile phase with 0.1%	[temp:40C]		"TSK-gel column ODS-100V (4.			

MetabolonoteのXMLファイルからの変換処理

```
<Page ID="" Title="/S01/M01">
<Template Name="M"><Field Name="M_ID">M01</Field><Field Name="M_Title">LC-FTICR-MS, ESI Positive analysis</Field><Field
Name="M_Method Set ID">MS1</Field><Field Name="M_Sample Amount">6.7 mg</Field><Field Name="M_Comment">[MassBase ID] MDLC1_05711</
Field></Template>
<Free_Text id="1">{{LinkTo_MassBase|MDLC1_05711}}</Free_Text>
</Page>
<Page ID="" Title="/S01/M01/D01">
<Template Name="D"><Field Name="D_ID">D01</Field><Field Name="D_Title">PowerGet data analysis for Bio-MassBank</Field><Field
Name="D_Data Analysis Set ID">DS1</Field><Field Name="D_Recommended decimal places of m/z">6|ITMS 2</Field></Template>
<Free_Text id="1">{{LinkTo_BioMassBank|keyword=SE1_S01_M01_D01}}
&amp;lt;!-- {{LinkTo_KomicMarketTmp}} --&amp;gt;
{{LinkTo_KM2|SE1}}</Free_Text>
</Page>
<Page ID="" Title="/S01/M01/D02">
<Template Name="D"><Field Name="D_ID">D02</Field><Field Name="D_Title">PowerGet data analysis for KomicMarket2</Field><Field
Name="D_Data Analysis Set ID">DS2</Field><Field Name="D_Recommended decimal places of m/z">6|ITMS 2</Field></Template>
<Free_Text id="1">&amp;lt;!-- {{LinkTo_KomicMarketTmp}} --&amp;gt;
{{LinkTo_KM2|SE1}}</Free_Text>
```

```
#S01/M01
M_ID:   M01
M_Title:    LC-FTICR-MS, ESI Positive analysis
M_Method Set ID:    MS1
M_Sample Amount:    6.7 mg
M_Comment:  [MassBase ID] MDLC1_05711


#S01/M01/D01
D_ID:   D01
D_Title:    PowerGet data analysis for Bio-MassBank
D_Data Analysis Set ID: DS1
D_Recommended decimal places of m/z:    6|ITMS 2


#S01/M01/D02
D_ID:   D02
D_Title:    PowerGet data analysis for KomicMarket2
D_Data Analysis Set ID: DS2
D_Recommended decimal places of m/z:    6|ITMS 2
```

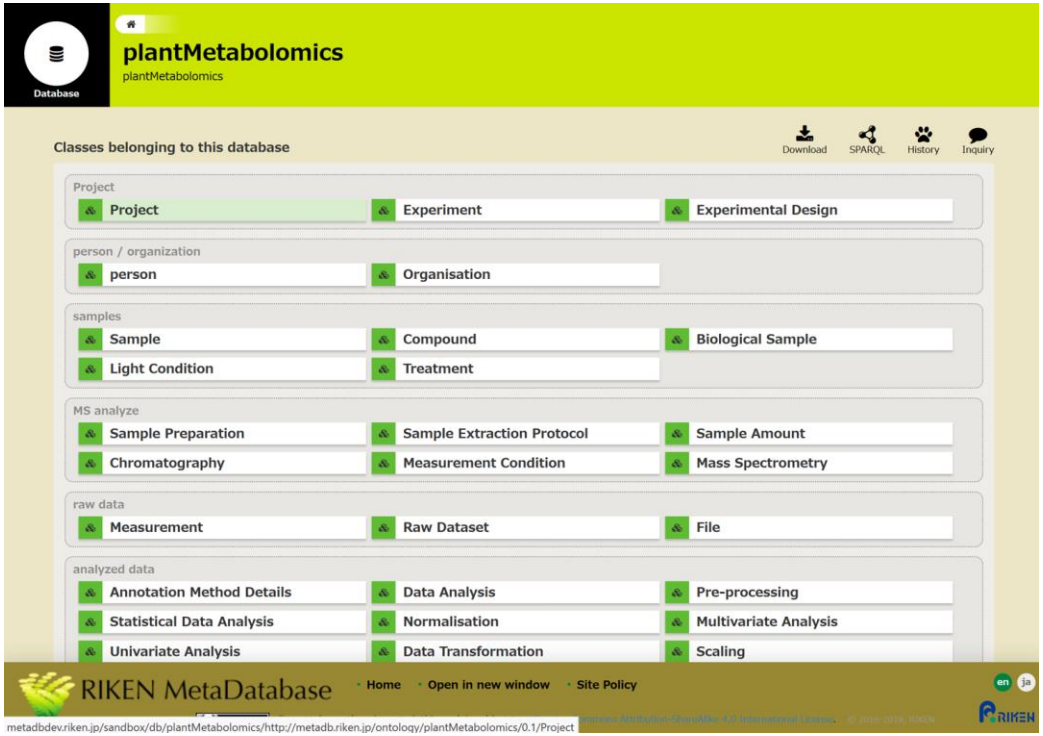
PGDBjやMassBaseなど外部へのリンクもまとめて出力

でも基本データ手入力

進捗：

- 104件完了
- RIKENメタデータデータベースに登録

➡ 検索テスト･入力フォーマット(6/3版)の調整



RPMM移植試験で使用したメタデータ
SE1　基本的
SE6　LTQ-Orbitrap使用
SE18　Mouse liver 解析書き込み多い
SE40 サンプル多い
SE61 サンプル、解析種類多い
SE62 サンプル、解析書き込み多い
SE88 Micro-Tom 解析種類多い
SE97 Lotus japonicus 解析種類多い
SE112 食品レポジトリ
SE114 食品レポジトリ
SE116 食品レポジトリ
SE189 解析書き込み多い

今後の予定

- **実験データの再解析**

  化合物データベースの充実化によって検索でヒットする化合物が増えることが見込まれる。

- **植物ゲノム統合DBへのメタボローム データの統合**

  ゲノム関連データとの横断的な検索を可能にする。

RIKEN Plant Metabolome MetaDatabase: an integrated plant metabolome data repository based on the semantic web

RIKEN
Plant Metabolome MetaDatabase

http://metabobank.riken.jp/

Atsushi Fukushima
RIKEN Center for Sustainable Resource Science

2019 07 12
Togo DB Site Visit

RIKEN CSRS

Aim of this project

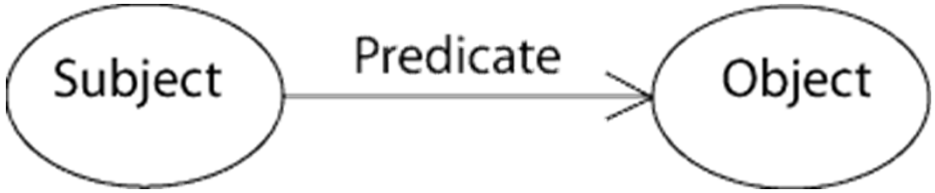
- Development of the RIKEN Plant Metabolome MetaDatabase (RIKEN PMM)
  - Semantic web based
  - Construction of the Plant Metabolomics Ontology
- Reanalysis/reannotation of GC-MS data in plants
  - Versioning of GC-MS spectra
  - Improvement of metabolite annotation

RDF/SPARQL

A standard model for data interchange on the Web

RDF (Resource Description Framework)

- has features that facilitate data merging even if the underlying schemas differ
- specifically supports the evolution of schemas over time without requiring all the data consumers to be changed



subject–predicate–object, known as triples

SPARQL - SPARQL Protocol and RDF Query Language

- an RDF query language
- can retrieve and manipulate data stored in RDF format

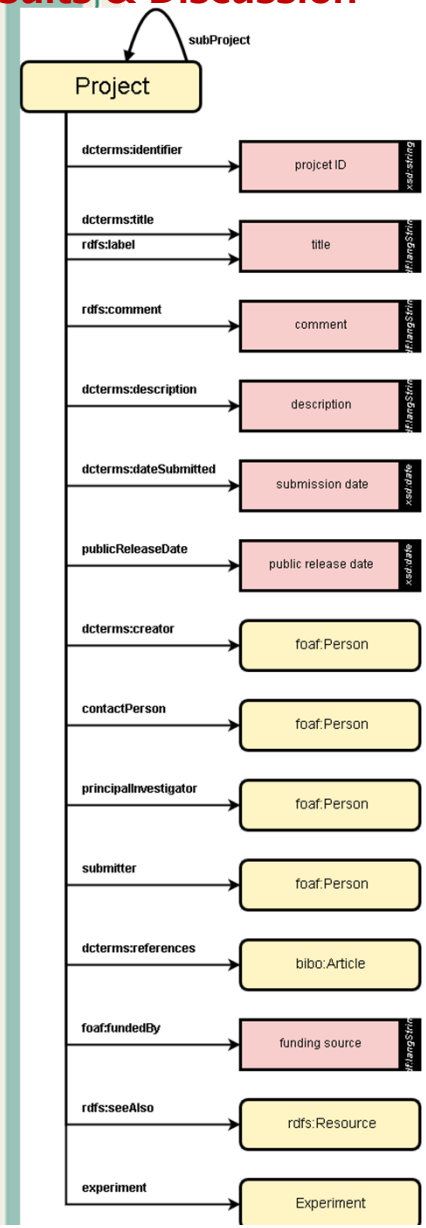
https://ja.wikipedia.org/wiki/Resource_Description_Framework

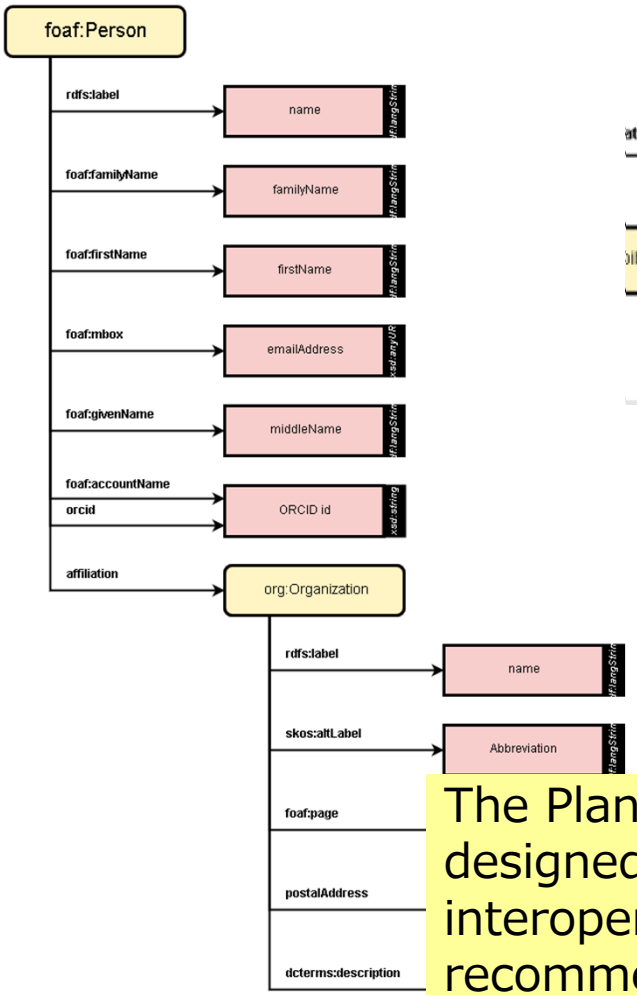
RIKEN MetaDatabase

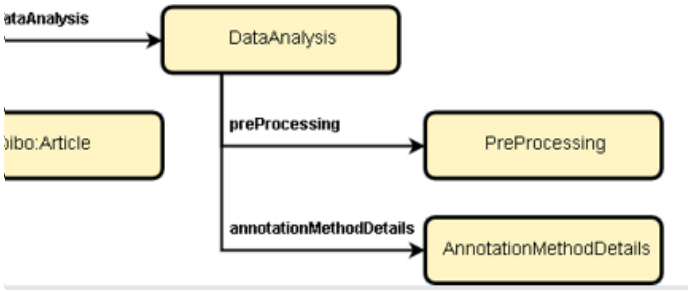
RIKEN MetaDatabase
http://metadb.riken.jp/

Kobayashi, N et al. LNCS 10055, 99-115 (2016)

RDF Schema in RPMM







The Plant Metabolomics Ontology is designed to realise metadata interoperability according to the recommendation of the Metabolomics Standards Initiative (MSI, http://www.metabolomics-msi.org/)

Metadata registration

Using a spreadsheet (e.g., MS Excel)

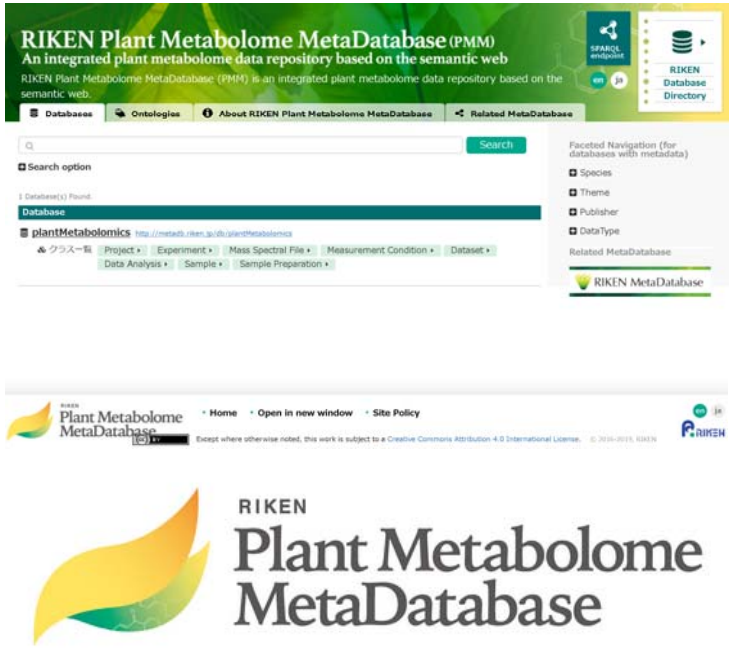
	A	B	C	D	E	G	H	I	J	
1	EnglishAttributior	Biological Sample	biological sample	name	comment	description	taxonomy	cultivar	genotype	plant organ
2	日本語属性	生物サンプル	生物サンプルID	名称	コメント	説明	分類群	品種	遺伝型	植物器官
3	PropertyURI		dcterms:identifie	rdfs:label	rdfs:comment	dcterms:descr	obo:RO_0002162	cultivar	genotype	plantOrgan
4	Datatype	BiologicalSample	xsd:string	rdf:langStrin	rdf:langString	rdf:langStrin	owl:Class	xsd:strir	xsd:string	owl:Class
5		pm_biosamp:test699_F_037_1	"F_037_1"		"This is test data"@en		http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
6		pm_biosamp:test699_F_037_old_1	"F_037_old_1"		"This is test data"@en		http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
7		pm_biosamp:test699_F_040_1	"F_040_1"		"This is test data"@en		http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
8		pm_biosamp:test699_F_040_old_1	"F_040_old_1"		"This is test data"@en		http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
9		pm_biosamp:test699_F_043_1	"F_043_1"		"This is test data"@en		http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo

An example of spreadsheet describing RIKEN PMM's "BiologicalSample" class.

G	H	I	J	
description	taxonomy	cultivar	genotype	plant organ
説明	分類群	品種	遺伝型	植物器官
dcterms:descr	obo:RO_0002162	cultivar	genotype	plantOrgan
rdf:langStrin	owl:Class	xsd:strir	xsd:string	owl:Class
	http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
	http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
	http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
	http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo
	http://purl.bioontology.org/ontology/NCBITAXON/	"Reiyo"		http://purl.obo

Generation of PMM's RDF

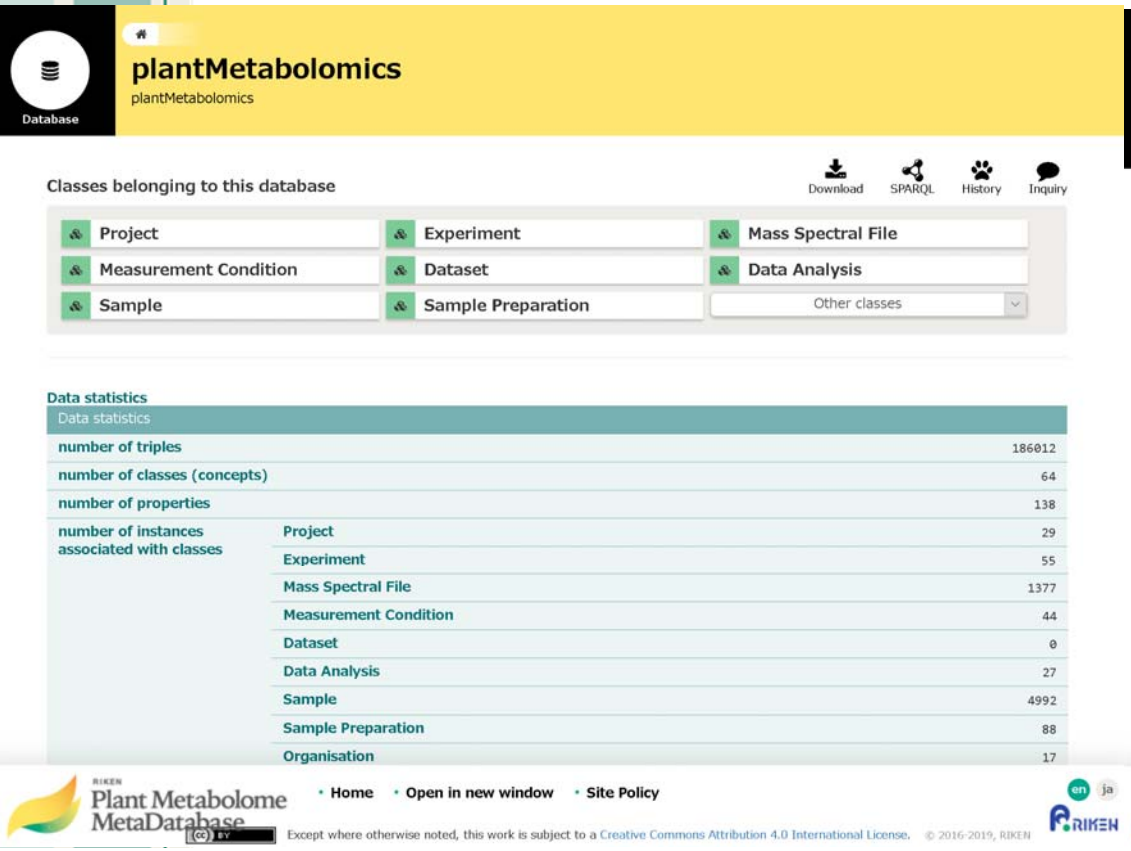
RIKEN PMM



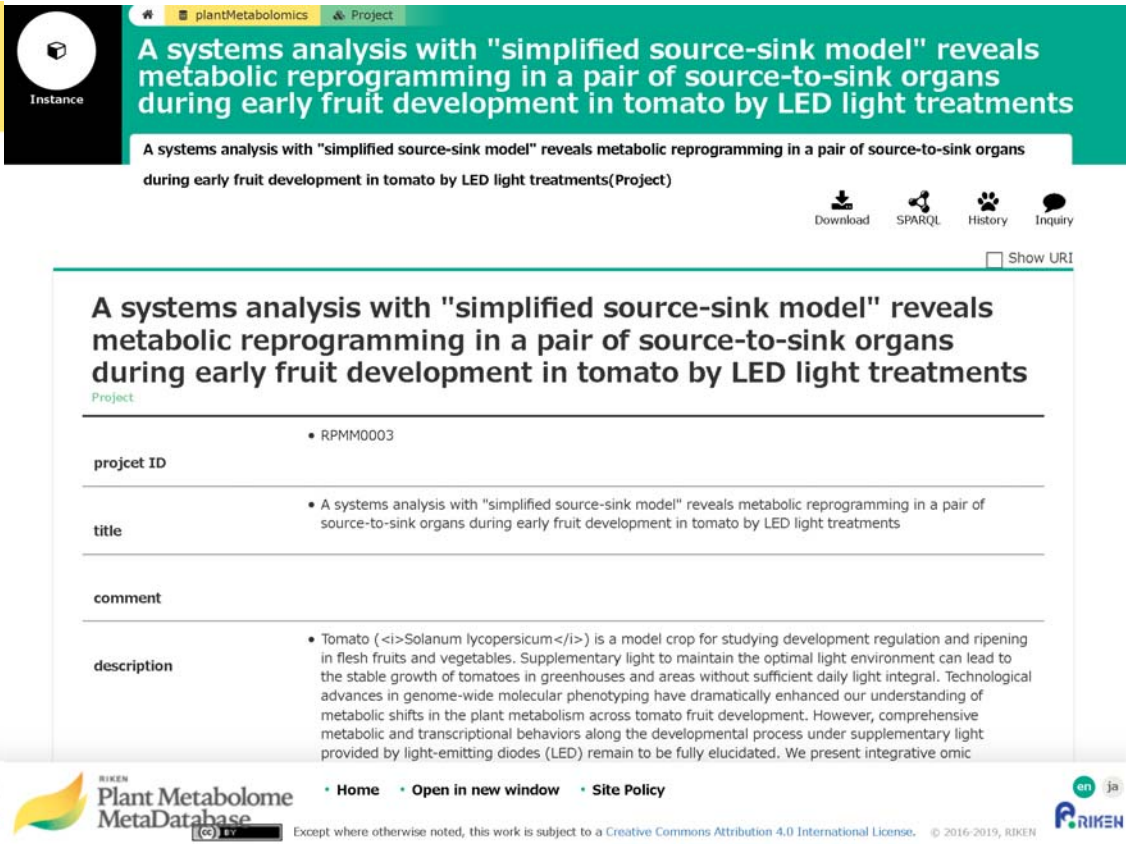
Results & Discussion

RIKEN PMM's graphical user interface (development ver)

Summary page



Project page



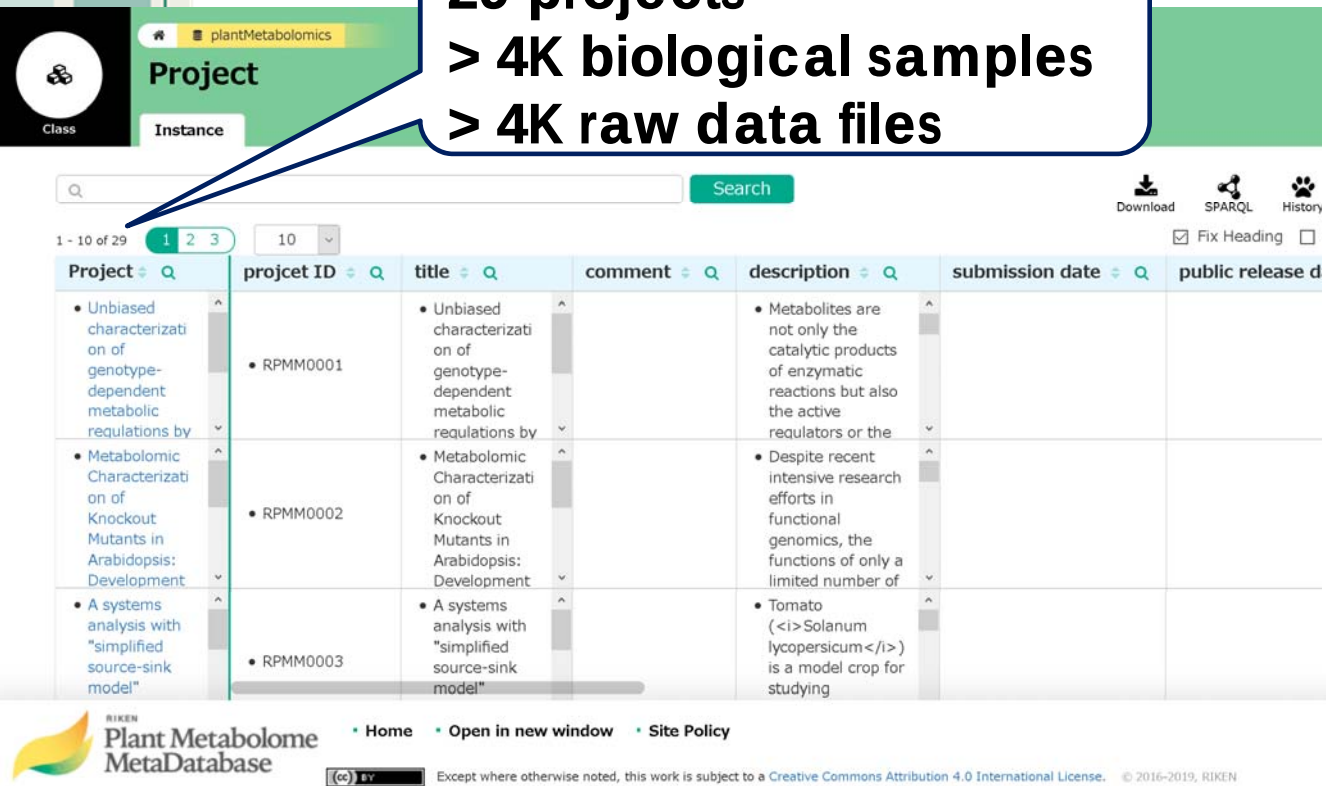
Results & Discussion

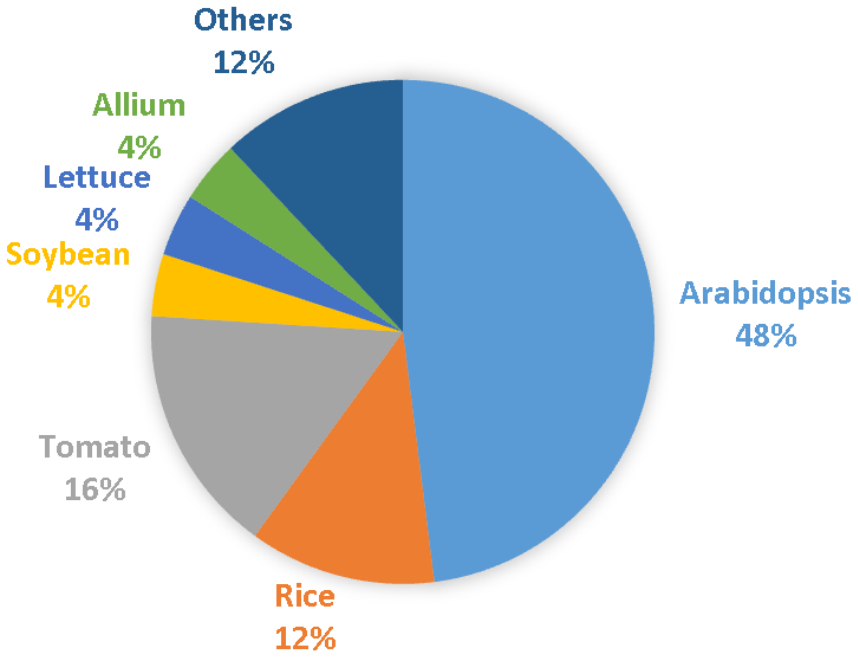
The current status of RPMM (development ver)

Summary page

Species stat

29 projects
> 4K biological samples
> 4K raw data files





Summary

- developed a semantic web-based repository database known as the RIKEN PMM using native RDF technologies

➔29 projects with high-quality metadata

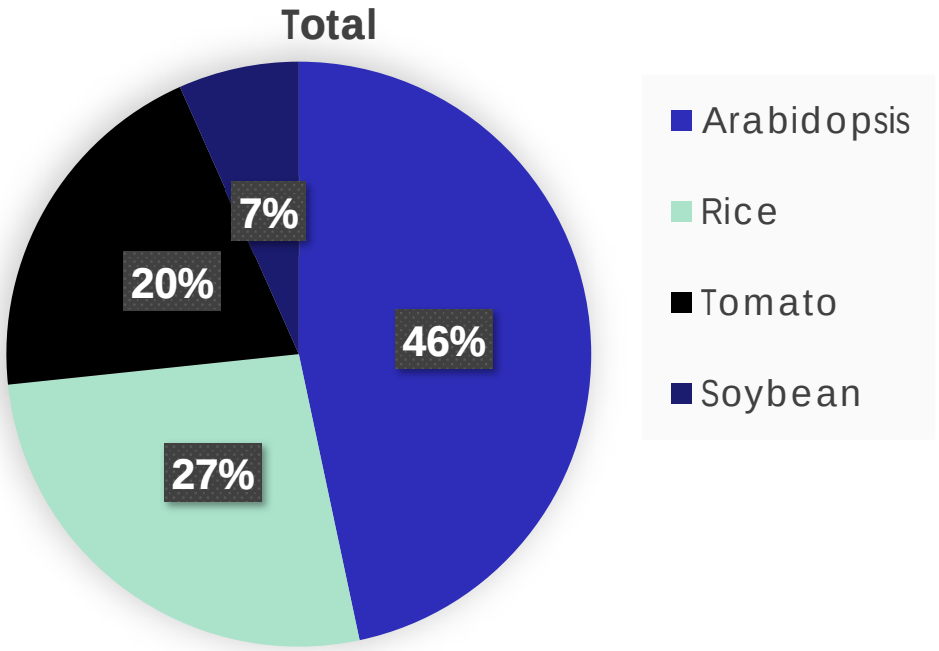
➔ efficient data sharing, spreading and retrieving in plant metabolomics

RIKEN Plant Metabolome MetaDatabase

http://metabobank.riken.jp/

Re-analysis of GC/MS data

15 Re-analyzed datasets



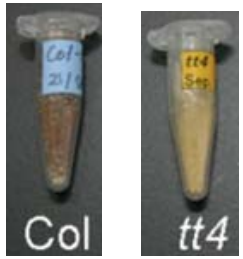
An example

Metabolomic correlation-network modules in Arabidopsis based on a graph-clustering approach

Fukushima et al. (2011); Kusano et al. (2007)

RPMM0006, MTBLS45, SE47

- *Arabidopsis thaliana*
- 3 genotypes (Col-0, *mto1*, and *tt4*)
- Root samples (biological replicates, *n* >= 15/genotype)
  53 samples
- Related to **RPMM0001** (MTBLS40 and SE46) (aerial part)
- GC-MS



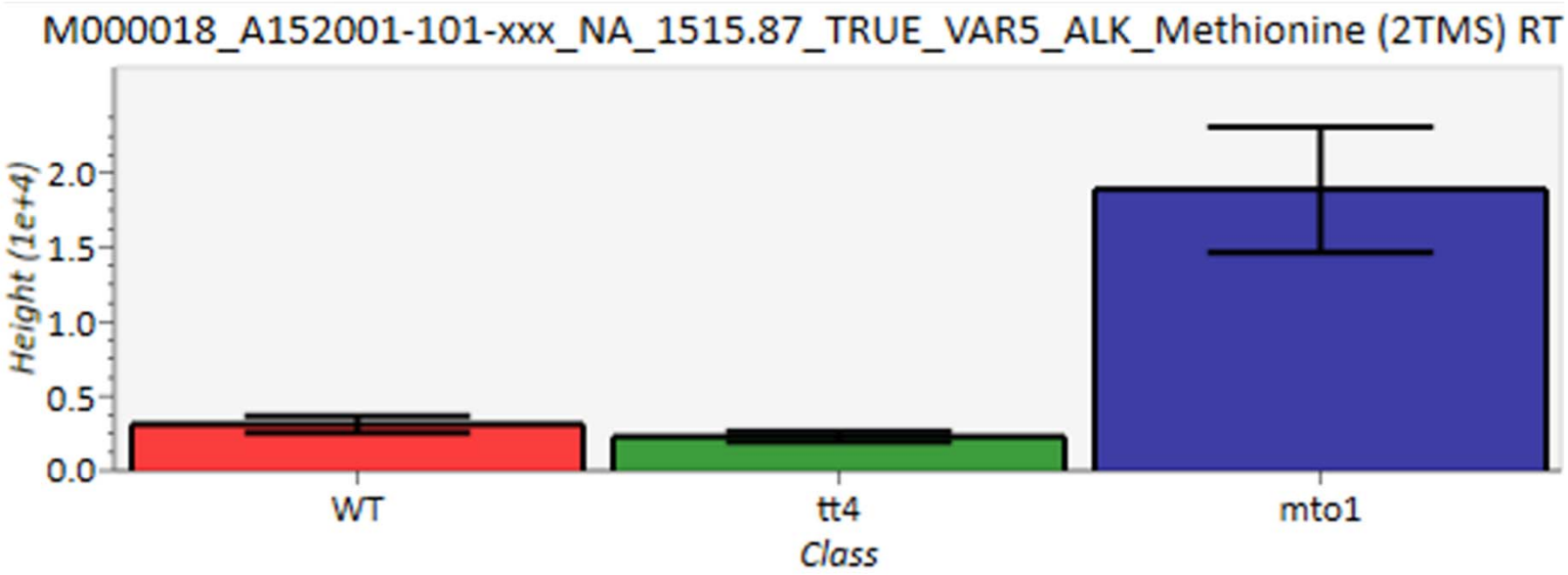
Re-analysis method

- MS-DIAL (Tsugawa et al.)
- Comparison with original annotation (Kusano et al.)

Results

	Extracted peaks	Annotated peaks	Known peaks
Original annotation	518	166	84
Re-analyzed annotation	444	253	204
		RI (error) < 10	58

Match rate: 58/84 = **69%**


M000018_A152001-101-xxx_NA_1515.87_TRUE_VAR5_ALK_Methionine (2TMS) RT
Height (1e+4)
2.0
1.5
1.0
0.5
0.0
WT
tt4
mto1
Class