

生命と環境のフェノーム統合データベース

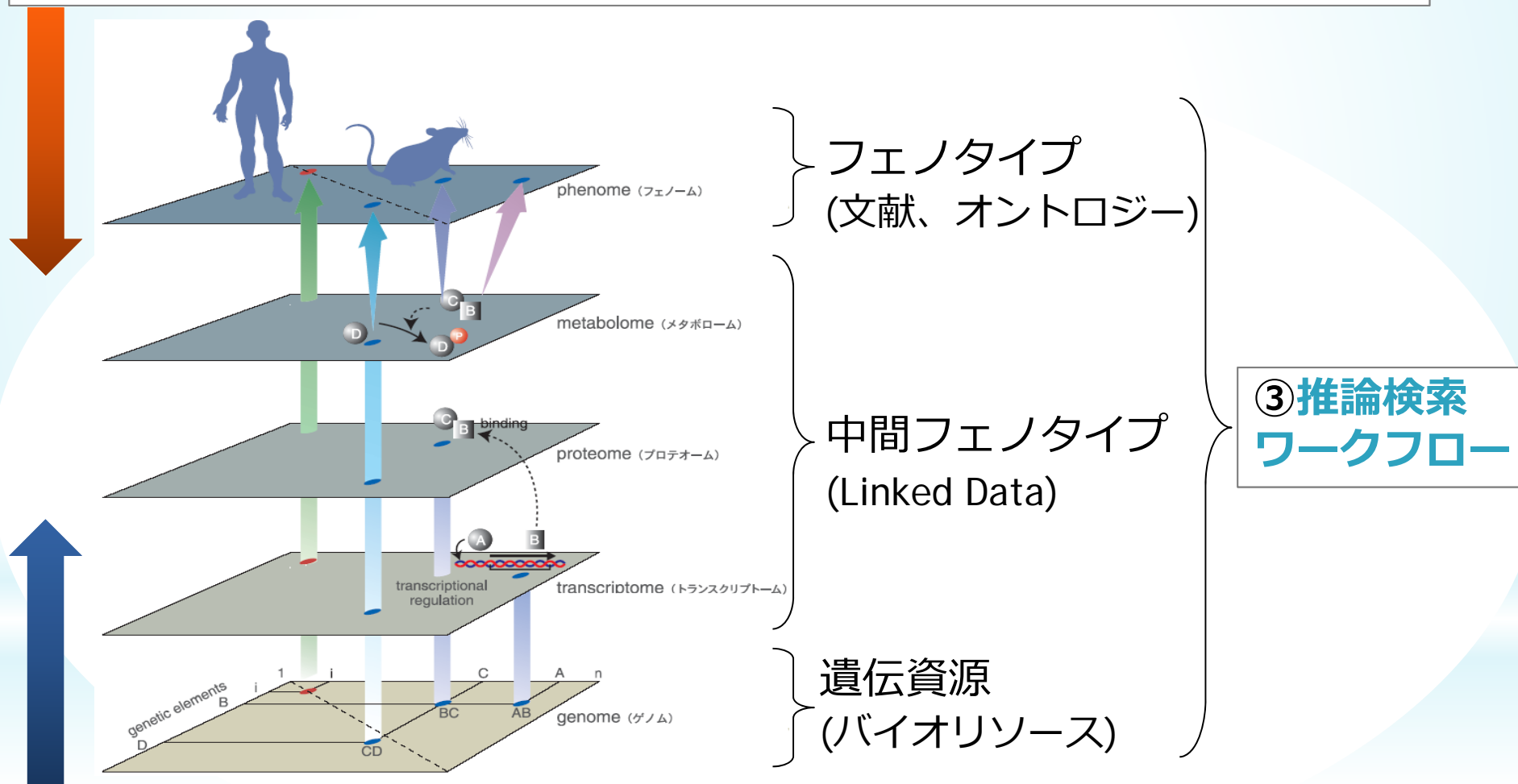
統合化推進プログラム 平成24年度 進捗報告会

2013年1月21日(月)

(独) 理化学研究所 豊田哲郎、柘屋啓志

フェノーム統合データベース

① **フェノタイプ情報**から**バイオリソース**を検索できるようにする



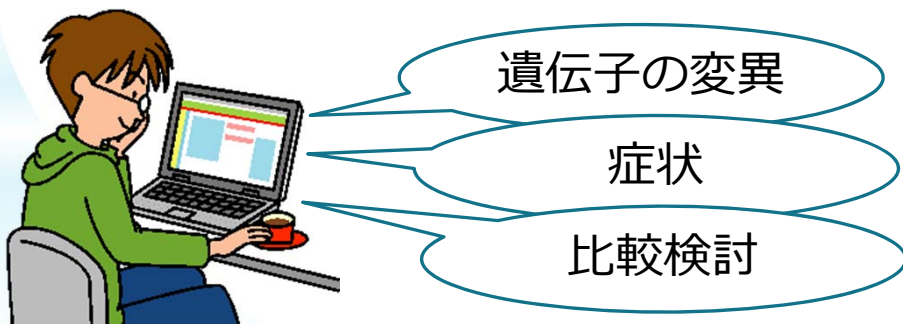
② **バイオリソース**から**フェノタイプ情報**を検索できるようにする

フェノーム統合DBは、バイオリソースの利用を促進する

従来： リソース名を知らないと検索できなかった

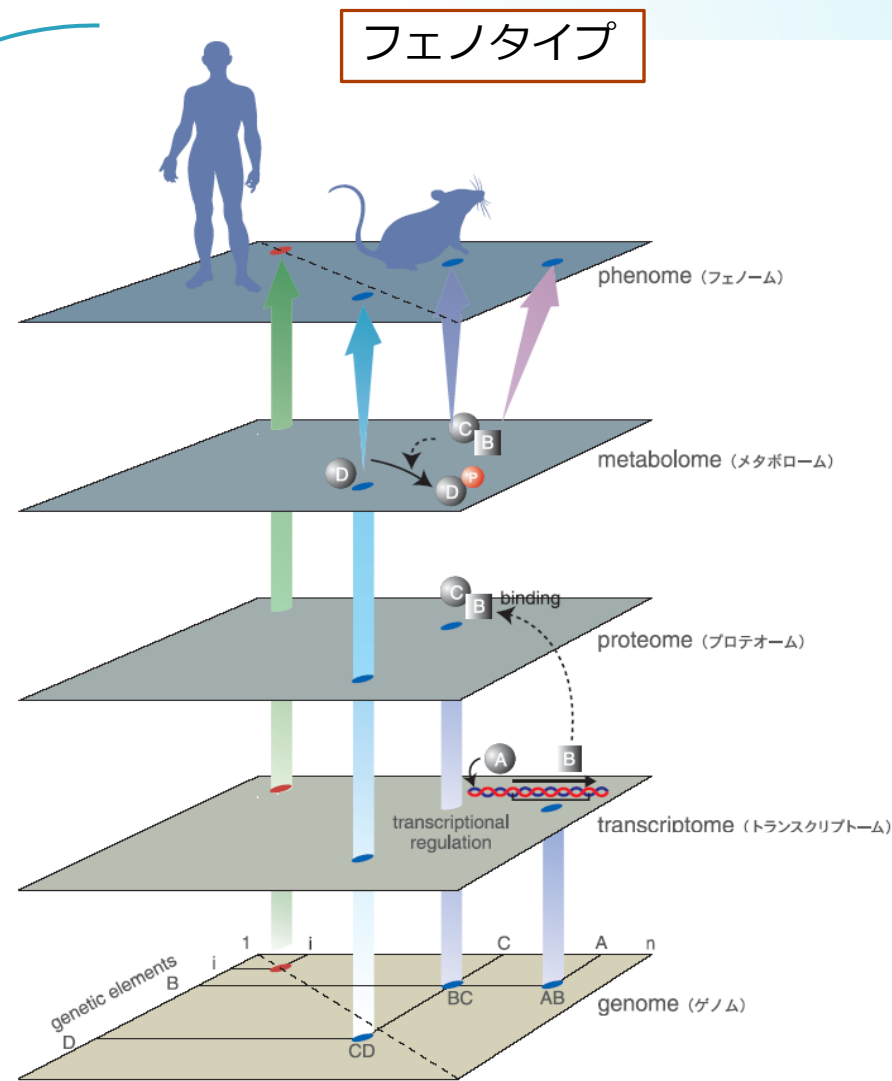


本データベース： フェノタイプ経由で研究にふさわしいリソースを探索、比較検討



バイオリソースの利用が促進される効果

バイオリソースの利用価値が高まる！



遺伝資源 (バイオリソース)

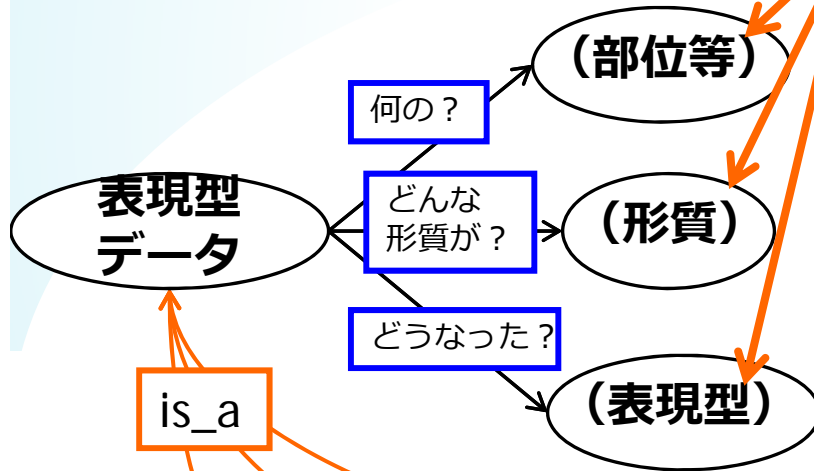
フェノーム情報の整備のために RDF技術を活用

表現型データ汎用RDFスキーム (抜粋)

➢ OBOコンソーシアム提案の改良型

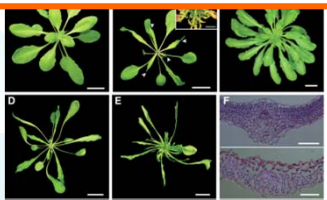
DB値/オントロ
ジーで代入

何の?
どんな
形質が?
どうなった?

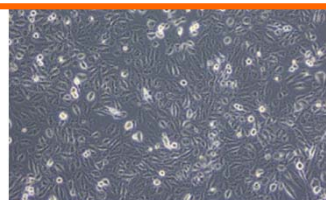


表現型アノテーション	系統	測定部位	測定対象の形質	測定結果	何の?/形質
phenotype of M100919	M100919	赤血球細胞数の増加	-	has number of	has extra parts of type erythrocyte
phenotype of M100451	M100451	短足症	foot	size	decreased size -
phenotype of M101156	M101156	光受容体外節の短縮	photoreceptor outer segment	length	decreased length -
phenotype of M101152	M101152	聴覚障害	sensory perception of sound	rate	decreased rate -
phenotype of M100646	M100646	血中インスリンレベル上昇	blood	concentration of	increased concentration insulin
phenotype of M100856	M100856	短指	limb digit	length	decreased length -
phenotype of M100702	M100702	インシュリンの血中濃度減少	blood	concentration of	decreased concentration insulin
phenotype of M100392	M100392	高血糖	blood	concentration of	increased concentration glucose
phenotype of M100210	M100210	高血糖	blood	concentration of	increased concentration glucose

植物表現型特性情報



培養細胞特性情報



KOマウス表現型情報



(マウス表現型の例)

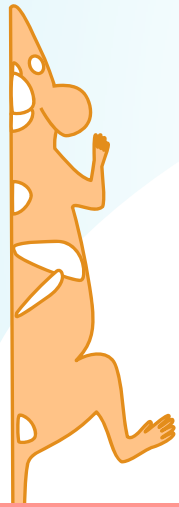
生物横断的に特性情報を格納できる汎用スキーム (各生物の特性情報 is_a 表現型データ)

開発したスキームの有用性

部位と特性で、リソース情報を整理

異常部位の内訳 (MAオントロジー分類)

神経系の異常：10系統
骨格系：36系統



長骨：16
長骨以外：5
中軸系：12
関節：3

表皮系：87系統
結合組織系：17系統
血液、リンパ系：102系統

長骨異常の内訳 (PATOオントロジー分類)

形態：35

サイズ：20

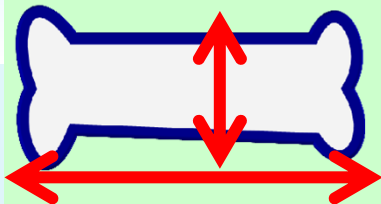
長さ短縮：13

径の短縮：6

形成不全：1

形態以外：1

形状：15



長骨の長軸方向成長
コントロール変異：
13系統、太さコン
ロール変異：6系統

柔軟なリソース検索の実現

「骨」に異常のあるマウスは？

異常部位でフィル
ター

表現型アノテーション	系統	mammalian phenotype ontology	測定部位	測定対象の形質	測定結果
phenotype of M100451	M100451	長骨骨端の形態異常	long bone epiphysis	morphology	abnormal
phenotype of M100856	M100856	長骨骨端軟骨板増殖帯の異常	long bone epiphyseal plate proliferative zone	physical object quality	abnormal
phenotype of B6.129S-Osmr<tm1Mtan>	B6.129S-Osmr<tm1Mtan>	骨髓球系前駆細胞の形態異常	common myeloid progenitor	morphology	abnormal
phenotype of M100856	M100856	下肢帯骨の形態異常	pelvic girdle bone	morphology	abnormal
phenotype of M101679	M101679	頭部の短縮	cranium	decreased diameter	decreased anterior-posterior diameter
phenotype of M100906	M100906	頭部の短縮	cranium	decreased diameter	decreased anterior-posterior diameter
phenotype of M101532	M101532	頭部の短縮	cranium	decreased diameter	-
phenotype of PIC/Nem-Pdn/+	PIC/Nem-Pdn/+	骨癒合症	bone	attachment quality	attached to
phenotype of M100576	M100576	頭蓋骨の短縮	cranium	height	decreased height
phenotype of M100117	M100117	半球状の頭蓋	cranium	curvature	domed
phenotype of M101682	M101682	頭部の短縮	cranium	decreased diameter	decreased anterior-posterior diameter
phenotype of M100566	M100566	頭部の短縮	cranium	decreased diameter	decreased anterior-

マウス系統のリスト

RDFの長所を活かして バイオリソースを「分かりやすく」提示

[See the Original Site](#)

マウス系統 (マウス系統 (理研バイオリソースセンター))

C57BL/6-KitW-v/+ [オリジナルサイト](#) [\[ページ一覧を見る\]](#)

BRC ID	RBRC00142 関連情報を検索
正式系統名	B6.Cg-KitW-v/Rbrc 関連情報を検索
由来生物の分類群	Mus musculus 関連情報を見る 関連情報を検索
系統カテゴリー	Spontaneous Mutation 関連情報を見る 関連情報を検索
アレル	Kit<W-v> 関連情報を見る 関連情報を検索
提供者	Yukihiko KITAMURA 関連情報を検索
提供機関	阪大 関連情報を見る 関連情報を検索
系統の保存状態	Frozen embryos 関連情報を見る 関連情報を検索
提供までの期間	凍結卵より作出した子孫を2~4ヶ月以内に提供可能、 凍結胚を1ヶ月以内に提供可能 関連情報を見る
リソースに関する文献	PMID:1559233, PMID:11578878 関連情報を見る
説明(英語)	B6.Cg-Kit<W-v>. Viable dominant spotting. Spontaneous mutation arose in C57BL. The Kit<W-v> mutation arose spontaneously in the C57BL strain. Heterozygote for Kit<W-v> are gray with light belly, white spot, light tail and feet. Homozygous mice are white with black eyes. The c-Kit cDNA from Kit<W-v>/Kit<W-v> mice showed an AC to T point mutation at nucleotide 2007 that resulted in the change from threonine to methionine at position 660. 関連情報を検索
供与元	発注する
"系統" が C57BL/6-KitW-v/+ の表現型アノテーション	31件 関連情報を検索

メタ情報

ID	URN:RIKEN:cria315s1ria315u10000000142i
URI	http://scinets.org/item/cria315s1ria315u10000000142i
外部ID	RBRC00142
所属ページ	http://www2.brc.riken.jp/lab/animal/detail.php?brc_no=RBRC00142
最終更新日	2012-12-07

画像 

おすすめ (関連マウス系統) [\[グラフを表示\]](#)

同じ遺伝子に変異を持つマウス系統 (11)

1 個のMouse MGI Geneが一致

C;B6-a/a · B/B-WSH/+	C;B6-a/a · b/b-WSH/+	B6C3Fe a/a-Kit<W-Rb...>	C3H/HeN-WR1
			
共通のMouse MGI Gene • Kit	共通のMouse MGI Gene • Kit	共通のMouse MGI Gene • Kit	共通のMouse MGI Gene • Kit

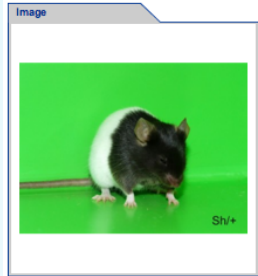
RDFのリンクは、
関連情報に延々と
繋がっている

最初に最低限の情報
へのリンクを見せて
おき、 **▼ 関連情報を見る**
をクリックすること
で、関連情報へと誘
導する

全てのリソースは
オープンに利用可能

“Amazon”型の
「おすすめマウス」
へのリンク (後述)

▼ 関連情報を見る をクリックすることで、 関連情報を閲覧



Announcement
Reference paper of this database site
The RIKEN integrated database of mammals
nar.oxfordjournals.org
The RIKEN integrated database of mammals (http://icdri.riken.jp/db/mammal) is the official undertaking to integrate its mammalian databases produced from multiple large-scale programs that have been promoted by the institute. The database integrates not only RIKEN's original data...

Directory
Basic Formal Ontology 1.1.1
CHEBI Chemical entities of biological interest
Genetics Ontology Lite
The RIKEN integrated database of mammals
RIKENBASE
Mouse strain in RIKEN BioResource Center
Mouse strain
RIKEN BioResource
Test results of genetic background
Genotyping PCR protocol
Phenotype annotation
RIKEN disease model mice
The other disease model mice
RIKEN disease model mice
Common disease model mice
RIKEN disease model mice
Skeletal dysplasia model mice
RIKEN disease model mice
Sensory organ disease model mice
RIKEN disease model mice
Haematological disorder model mice
RIKEN disease model mice
RIKEN ENI induced mouse line (in BRC)

Kit<W-sh> アレルによってもたらされる表現型を列挙

See the Original Site

Mouse Strain (Mouse strain in RIKEN BioResource Center)

C57BL/6-WshWsh Original Site [Browse list of all records]

BRC ID	RBRC01888
Nomenclature of Mouse Strains	B6(C)-KitW-sh/Rbrc
taxon	Mus musculus
parent node	Mus musculus
subspecies	Mus musculus
Database Registry	36 subjects
Cell lines	1755 subjects
Mouse Strain	5037 subjects
Biodata repository	RIKEN RefDB/READ: RIKEN BRC
PDB (Protein Data Bank)	3051 subjects
Swiss-Prot Rodents	16401 subjects
RefSeq Mouse_mRNA	36441 subjects
Mouse Ensembl Gene Annotation Release 59	36229 subjects
Mouse Ensembl Transcript Annotation Release 59	87862 subjects
Mouse MGI Gene	84288 subjects
Genome version	NCBI mouse NCBI mouse NCBI mouse
Transcriptome data	Mouse 430 Contains RE Contains RE validation... Contains tip analysis.
Protein related data	Mouse 2DE Mbyte). Mouse 23-Plus Kbyte)
BRC Strain Type	Spontaneous Mutation
has_allele	Kit<W-sh> allele type MGI marker
Other database links(Ensembl)	Spontaneous Kit Other database links(Ensembl)
genotype	Kit
species	Kit
physical description	Kit
synonym	Kit
EntrezGene Mouse	Kit
MGI Allele	104 subjects
CDT-DB	CD15215, CD11943

実験用マウス
"Mus musculus"

Mus musculusにおける他のバイオリソースへと誘導

このマウスは、
"Kit<W-sh>"
アレルを持つ

Kit<W-sh> アレルとは、
Kit遺伝子のバリエーション

Strain	C57BL/6	Hide Detail
Strain Type		
Holder Site	http://www.mgi.jp	
Allele	973 sub	
Mouse Strain analyzed in FANTOM 4	C57BL/6	
Mouse Strain	221 sub	
Keisuke OBOKI		
Institute	RIKEN	Hide Detail
alias	理化学研究所 (独)理化学研究所	
kakeiInstituteCode	7705	Hide Detail
Institute	199	
Institute	46 subjects	
Mouse Strain	46 subjects	
Live mice, Frozen embryos		
BRC Mouse strain availability		
Live mouse will be provided 1 to 3 months		
Cryopreserved embryos will be provided within 1 m.		
Japanese description	マスト細胞、メラノサイトの欠損マウス。の機能解析。変異の位置や種類は未解析。	
reference of the resource	reference of the resource	
Journal	Blood, 83, 35	
PMID: 7515715		
PMID: 16127161		
PMID: 10515869	Blood, 94, 2658-2666 (1999).	1999
PMID: 1381623	Blood, 15, 1448-1453 (1992).	1992
PMID: 7521	Development, 118, 705-717 (1993).	1993
PMID: 71	Genet. Res., 39, 315-322 (1982).	1982
Description	The Kit gene arose spontaneously in the (C3H/HeN x 101/HJF1) at the MRC (Lyon et al., 1982). white coat with black eyes. Heterozygotes are black with a white sash. This strain is useful for and neural crest migration.	
Genotyping protocol		
Supplier	http://www2.brc.riken.jp/lab/animal/detail.php?brc_no=RBRC01888	

このシステムは
"C57BL" 系統に由来する。

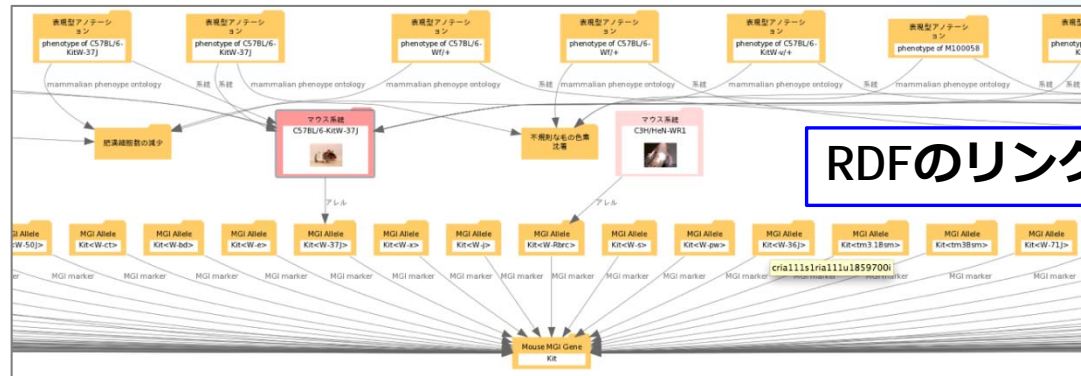
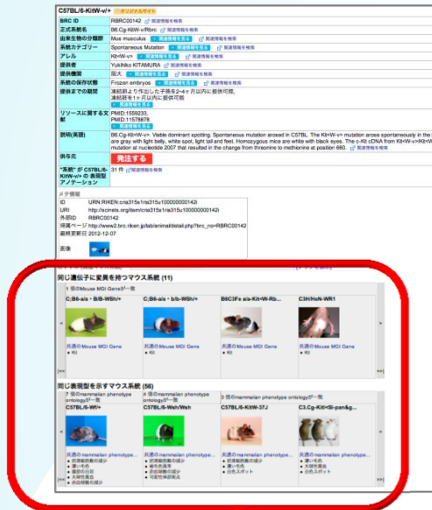
他に"C57BL"由来系統は何があるか

本システムの寄託者は
理研

リファレンス情報

RDFの繋がりを機械的に探索し、「お勧めマウス」を自動提示

RDFのリンク辿る事で、類似の表現型を示すマウスを自動収集。
 “Amazon” 風のインターフェースで、より目的に合ったリソースに誘導。



RDFのリンク (抜粋)

同じ表現型を示すマウス系統 (56) (各系統の示す表現型の一致度が高い順番に提示)

<p>7 個のmammalian phenotype ontologyが一致 C57BL/6-Wf/+</p>  <p>共通の mammalian phenotype...</p> <ul style="list-style-type: none"> ● 肥満細胞数の減少 ● 薄い毛色 ● 腹部の白斑 ● 大球性貧血 ● 赤血球数の減少 	<p>4 個のmammalian phenotype ontologyが一致 C57BL/6-Wsh/Wsh</p>  <p>共通の mammalian phenotype...</p> <ul style="list-style-type: none"> ● 肥満細胞数の減少 ● 被毛色異常 ● 赤血球数の減少 ● 可変性体部斑点 	<p>3 個のmammalian phenotype ontologyが一致 C57BL/6-KitW-37J</p>  <p>共通の mammalian phenotype...</p> <ul style="list-style-type: none"> ● 肥満細胞数の減少 ● 薄い毛色 ● 白色スポット 	<p>3個のmammalian phenotype ontologyが一致 C3.Cg-Kitl<Sl-pan&g...</p>  <p>共通の mammalian phenotype...</p> <ul style="list-style-type: none"> ● 薄い毛色 ● 大球性貧血 ● 白色スポット
--	---	---	--

類似度高

類似度低

フェノーム統合の高度化に向けて

- H23 リソースを集めて繋ぐ・公開する
- H24 分かりやすく見せる・探せる
- H25 ユーザのデータ解析に役立つ



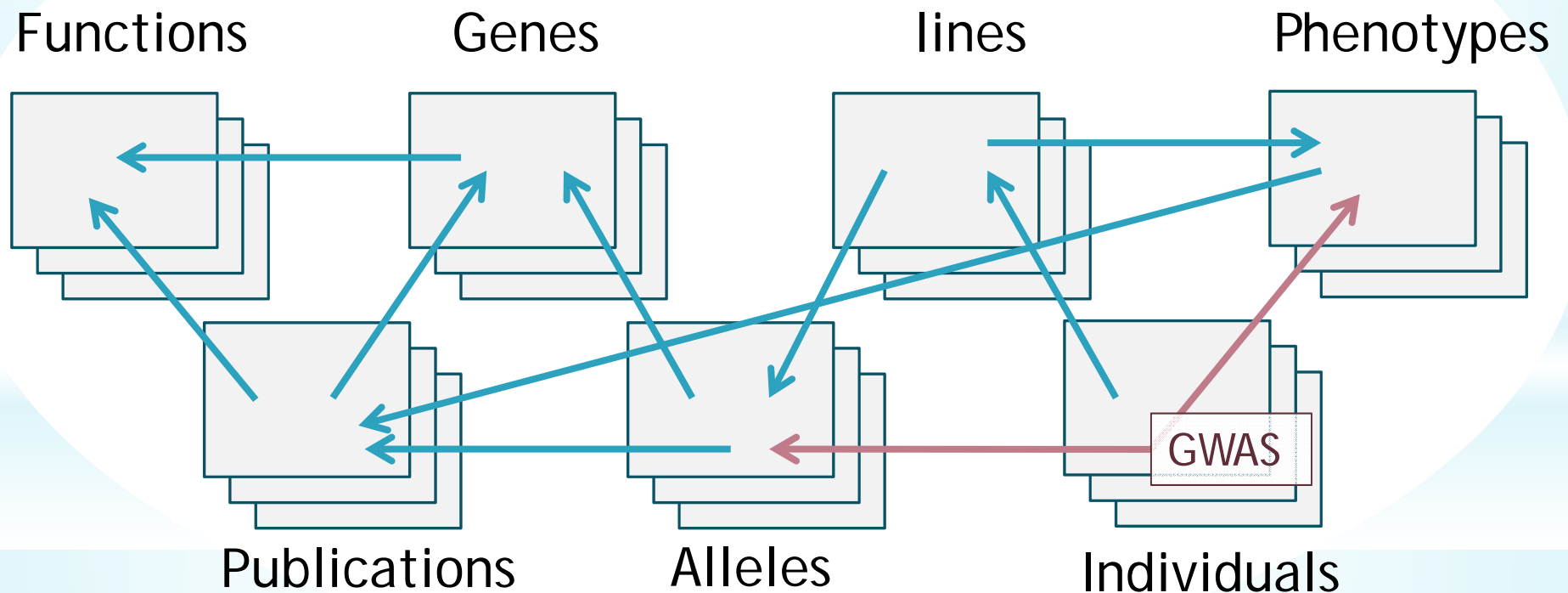
SWAS解析

(Semantic-Web Association Study)

Semantic Webによる相関解析の拡張

Kobayashi & Toyoda, *Bioinformatics*, 2008

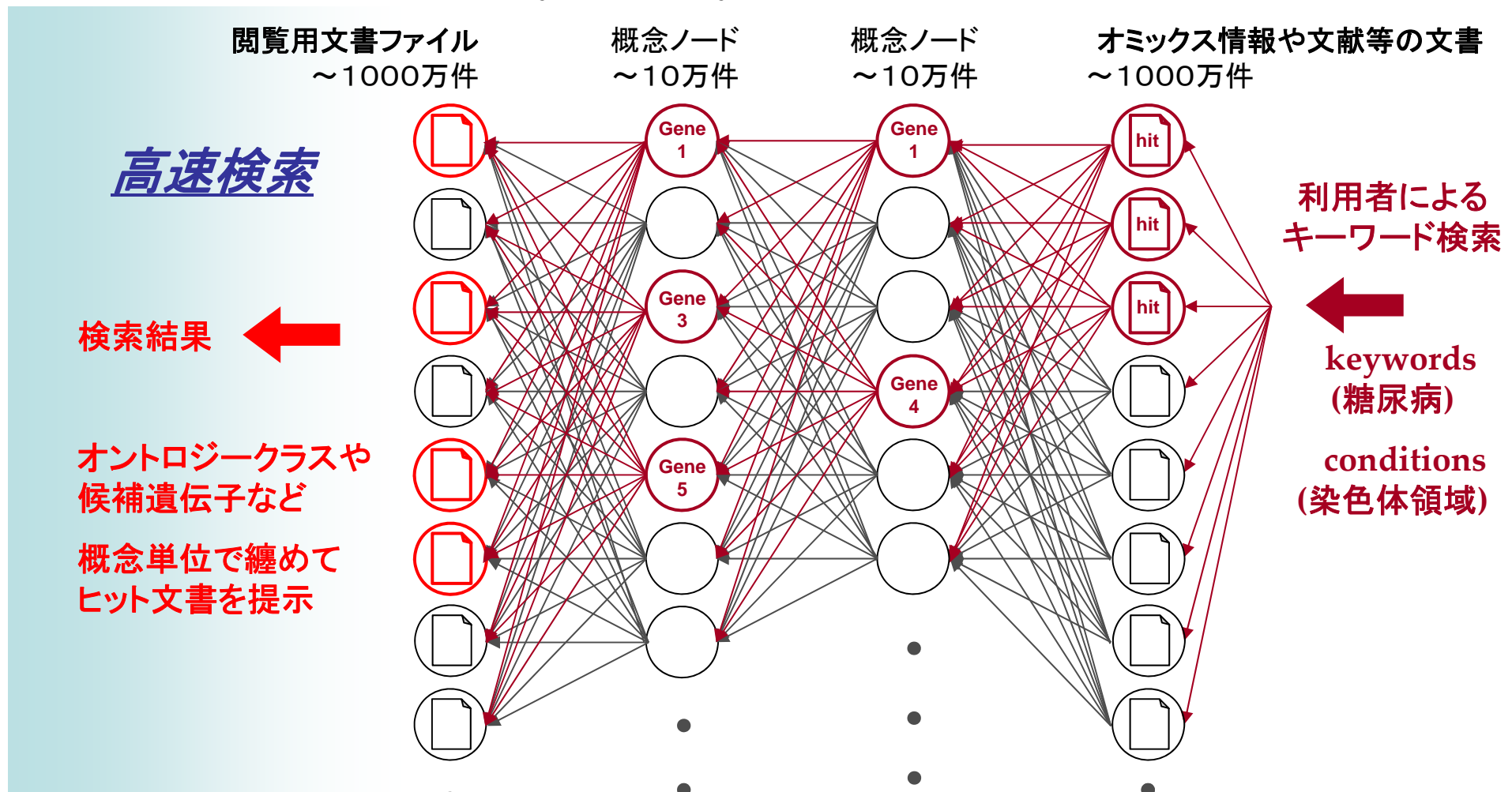
- GWAS: Genome-Wide Association Study
- SWAS: Semantic-Web Association Study



SWAS検索エンジン(PosMed)の適用

統計的なセマンティックウェブ検索システム

Kobayashi & Toyoda *Bioinformatics* (2008)



すべての計算が1秒程度で終わる → 高速ウェブ検索の実現

数万個の集計表を高速に計算する特許技術を適用

すべての Gene (X=1,2,,,数万) に対して、
三次元集計表を瞬時に作成する技術

	検索ワードを含む	検索ワードを含まない
Gene Xを含む	A個のドキュメント	B個のドキュメント
Gene Xを含まない	C個のドキュメント	D個のドキュメント

プロパティ(ドキュメントの種類)による
分類(Medline, OMIM, PPI,
Bio resources 等)

この表に基づいて、統計量の算出や
フィッシャーの正確検定などの
統計的推論を行うことができる

遺伝子名や
概念ノード等
(~10万件)

文献等の文書ファイル
(~千万件)

利用者による
フェノタイプ検索
(検索ワード)

キュレーターが対応づけた
遺伝子と文書の対応関係

文献と遺伝子の正確な対応付けルールを構築 → 正確なテキストマイニング

SWAS検索でバイオリソースの フェノタイプを推論

search bioresource condition bioresource name species mouse

keyword diabetes GWAS bioresource name search clear recent 4 years 4 show 20

Hit resources

download rank list
set as target

check all

- C57BL/6J
Myd88
- IGIF/IL-18 KO
Il18
- C3H/HeJ
Cntnap1
- C3HeB/FeJ
Asgr1
- C57BL/6JStfdAnu-
Cd4^{thb}/A...
Cd4
- B6.129-Tnfrsf1a^{tm1Mak}
Tnfrsf1a
- C.CBA-Tg(CD2-
IL5)5C2...
Cd2
- B6.CBACa-Tg(CD2-
IL5)...
Cd2
- CBA/Ca-Tg(CD2-
IL5)5C...
Cd2
- CBA/Ca-Tg(CD2-
IL5)5C...
Cd2
- C3.SW-H2/Sn...
Cntnap1
- C.Cg-Myd88^{tm1Aki}
Tg(CD2-IL...
Myd88
- NOD.Cg-Tnfrsf1a^{tm1BIt}
Fas^{DF}
Tnfrsf1a
- C57BR/cdJ
Cd4
- SHM/Nem-Re/+shm/+
Cntnap1
- C57BL/6JAnu-
Cd4^{thb}/App
Cd4
- RKO-1

Links from the resource to the keyword

BALB/cByJ ← Acads ← diabetes GWAS

Mouse bioresource ← Provenance ← Mouse Locus ← keyword

About the resource

Acads 182 2 show graph

Symbol	Acads
Name	acyl-Coenzyme A dehydrogenase, short chain
Other aliases	SCAD, Bcd-1, Bcd1, Hdlq8

Documents about the resource

all (2/182)	Mouse bioresource (0/3)	Gene ontology (0/16)	MEDLINE (1/125)	Mammalian phenotype ontology (0/13)	Mouse protein-protein interaction (0/0)	Mouse gene (0/1)	OMIM (1/13)	Pathway information (0/11)	other databases
-------------	-------------------------	----------------------	-----------------	-------------------------------------	---	------------------	-------------	----------------------------	-----------------

show related entities

keyword diabetes GWAS year all Find

show digest order by fact download

1-20 of 182 Page 1 next

Documents related to Acads

2: The minor C-allele of rs2014355 in **ACADS** is associated with reduced insulin release following an oral **glucose** load.
PMID:21211036
BMC medical genetics 2011
AB-

A genome-wide association study (GWAS) using metabolite concentrations as proxies for enzymatic activity, suggested that two variants: rs2014355 in the gene encoding short-chain acyl-coenzyme A dehydrogenase (**ACADS**) and rs11161510 in the gene encoding medium-chain acyl-coenzyme A dehydrogenase (**ACADM**) impair fatty acid β -oxidation.

Chronic exposure to fatty acids due to an impaired β -oxidation may down-regulate the glucose-stimulated insulin release and result in an increased risk of type 2 diabetes (T2D). ... The C-allele was not associated with T2D in the case-control analysis (OR 1.07, 95% CI 0.96-1.18, P = 0.21). rs11161510 of **ACADM** did not associate with any indices of glucose-stimulated insulin

キーワード
「糖尿, GWAS」で
バイオリソースを検索



GWASで糖尿原因遺
伝子とされている
Acads遺伝子と文献
共起のある、
マウスリソースが提示
された。

SWAS検索でシロイヌナズナ変異体の 表現型を推論検索

search bioresource condition bioresource name species arabidopsis

keyword **flowering** bioresource name search clear recent 4 years 4 show 20

Hit resources

download rank list

set as target

check all

1. F24649, Mutant of FO... >> AT1G65480
2. Z032806, Mutant of A... AT5G65060
3. Z031044, Mutant of A... AT3G04610
4. F23728, Mutant of FO... AT1G65480
5. F12202, Mutant of FO... AT1G65480
6. Z100145, Mutant of A... AT3G19140
7. 11-2516-1, Mutant of... AT1G04400
8. F09944, Mutant of FO... AT5G35840
9. 16-2733-1, Mutant of... AT1G24260
10. F10518, Mutant of FO... AT2G46790
11. Z089550, Mutant of A... AT5G60100
12. Z050250, Mutant of A... AT5G17690
13. Z077144, Mutant of A... AT1G68050
14. F13347, Mutant of FO...

Links from the resource to the keyword

F24649, Mutant of FO... Arabidopsis bioresource ← Provenance ← AT1G65480 Arabidopsis Locus ← flowering keyword

About the resource

AT1G65480 241 234 show graph

Symbol	AT1G65480
Name	PEBP (phosphatidylethanolamine-binding protein) family protein
Other aliases	FT, FLOWERING LOCUS T
ID	AT1G65480

Documents about the resource

all (234/241)

Arabidopsis gene (1/1)	At co-expression (0/2)	Arabidopsis protein-protein interaction (0/5)	At phenotype (2/2)	At protein (0/0)	MEDLINE (230/230)	Plant phenotype (1/1)
------------------------	------------------------	---	--------------------	------------------	-------------------	-----------------------

show related entities

keyword **flowering** year all Find

documents

publication year

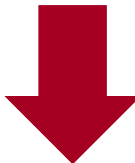
show digest order by fact download

1-20 of 241 Page 1 next

Documents related to AT1G65480

- 1: Isolation and characterization of a **FLOWERING LOCUS T** homolog from pineapple (*Ananas comosus* (L.) Merr).
PMID:22710136
Gene 2012 Sep 1
AB- **FLOWERING LOCUS T** (FT)-like genes are crucial regulators of **flowering** in angiosperms. ...
- 2: Expression of vernalization responsive genes in wheat is associated with histone H3 trimethylation.
PMID:22684814
Molecular genetics and genomics - MGG 2012 Jul

検索キーワード
「開花」で検索



既存のデータベースには変異体F24649の開花に関する表現型情報は含まれていないが、推論検索により関係性が示された。

疾患関連遺伝子の候補群の中から機能解釈可能なものを SWAS検索

Bioinformatics, 2008
Nucleic Acids Res., 2009
Plant Cell & Phys., 2009

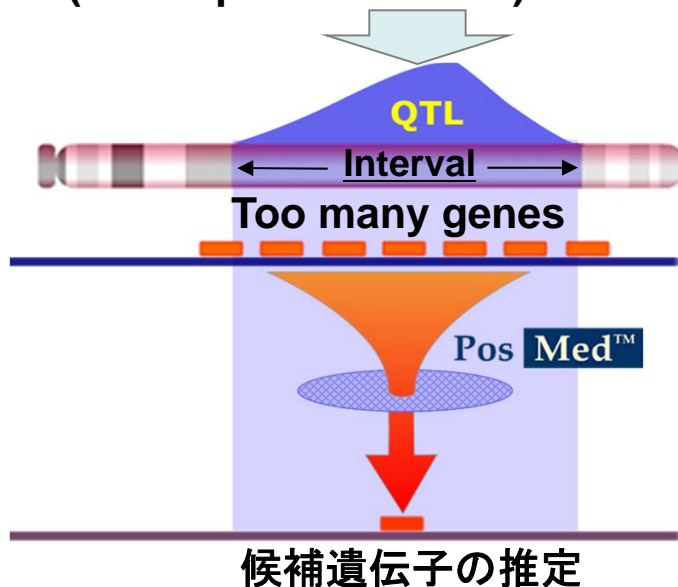


複合検索条件

キーワード: type 2 diabetes
ゲノム条件: 6番染色体の遺伝子

ポジショナルクローニングでの
候補遺伝子選び

交配実験によるラフマッピング
(~10Mbp wide interval)



原因遺伝子変異の同定

候補遺伝子の
ランキング結果

Rank	Interval	Mouse Locus	Co-cited	Drug	Metabolite	Keyword
1	P value: 4.44E-5044	Symbol: Pparg Name: peroxisome proliferator activated receptor gamma Other aliases: Nrlc3,PPAR-gamma,PPar-gamma2 ID: MGI:97747 Document: with keyword Link to: MGI Position: Mm.6.116786416-116826626	P value: 2.48E-782 Documents: 1190hits, with keyword	Symbol: Insulin ID: DRG:00491		P value: 6.48E-8740 Document: 9257 hits with keyword Keyword Document: 12584 hits Keyword: "type 2 diabetes"
2	P value: 1.41E-1819	Symbol: Iapp Name: islet amyloid polypeptide Other aliases: amylin ID: MGI:96382 Document: with keyword Link to: MGI Position: Mm.6.143141688-143147087	P value: 1.16E-770 Documents: 712hits, with keyword	Symbol: Insulin ID: DRG:00491		P value: 6.48E-8740 Document: 9257 hits with keyword Keyword Document: 12584 hits Keyword: "type 2 diabetes"
3	P value: 1.41E-404	Symbol: Abcc9 Name: ATP-binding cassette, sub-family C (CFTR/MRP), member 9 Other aliases: SUR2A,SUR2B,Sur2 ID: MGI:1352630 Document: with keyword Link to: MGI Position: Mm.6.143432404-143545480	P value: 7.48E-193 Documents: 87hits, with keyword		Symbol: C23H2CIN305S Name: Glyburide Other aliases: Glibenclamide ID: CPD:C07022	P value: 2.66E-277 Document: 297 hits with keyword Keyword Document: 12584 hits Keyword: "type 2 diabetes"
4	P value: 4.74E-231	Symbol: Slc2a3 Name: solute carrier family 2 (facilitated glucose transporter), member 3 Other aliases: Glut-3,Glut3 ID: MGI:95757 Document: with keyword Link to: MGI Position: Mm.6.123394790-123409492	P value: 9.87E-284 Documents: 115hits, with keyword		Symbol: Slc2a4 Name: solute carrier family 2 (facilitated glucose transporter), member 4 Other aliases: Glut-4,Glut4 ID: MGI:95750 Link to: MGI	P value: 1.82E-131 Document: 140 hits with keyword Keyword Document: 12584 hits Keyword: "type 2 diabetes"
5	P value: 4.03E-281	Symbol: A2m Name: alpha-2-macroglobulin Other aliases: A2mp ID: MGI:2449119 Document: with keyword	P value: 6.40E-242 Documents: 300hits, with keyword		Symbol: Plg Name: plasminogen Other aliases: Pg ID: MGI:97520 Link to: MGI	P value: 1.17E-47 Document: 209 hits with keyword

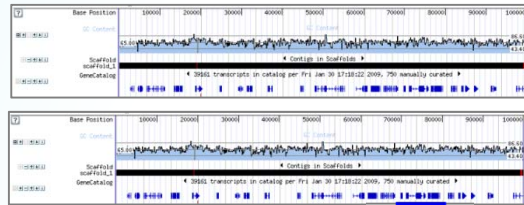
ENUマウス mutaジェネシスでは、約65のマウス系統の原因遺伝子の探索成功に貢献
→ 今後は、NGSで見つかるヒト細胞の変異部位の機能解釈に応用していく

用途例：がん細胞の原因変異遺伝子の *in silico* 絞り込み

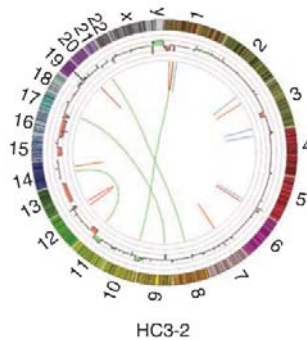
① NGSデータの比較による変異部位の同定

正常細胞

がん細胞

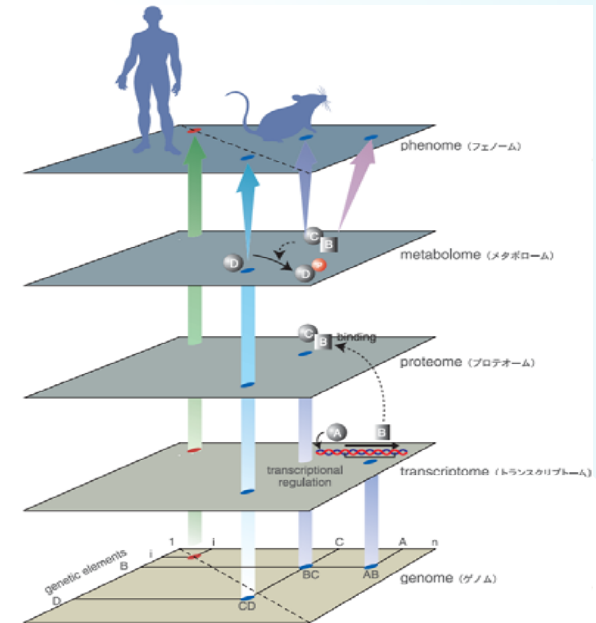


② 変異のある遺伝子のリストの作成 (数百~数千)



Gene1
Gene2
.
.
.

変異遺伝子リスト



フェノーム統合DB

③ 変異遺伝子とフェノタイプ情報を推論的に結び付けて提示 (原因遺伝子変異の絞り込み)

Gene1

4. **IL1A . interleukin 1, alpha**

<p>Interval</p> <p>Human Locus</p> <p>Symbol: IL1A Name: interleukin 1, alpha Other aliases: IL1F1, IL-1A, IL1-ALPHA, IL1 P value: 9.36E-2404 ID: HGNC:5991 set as target 2 Link to: RIKEN SciNetS HGNC Position: Hs:2:113531492-113542187</p>	<p>Mouse Locus</p> <p>Mouse Locus</p> <p>Symbol: Il1a Name: interleukin 1 alpha Other aliases: Il-1a ID: MGI:96542 7490 Link to: RIKEN SciNetS MGI Position: Mm:2:128125346-129135708</p>	<p>Co-citation</p> <p>P value: 9.36E-2404 4510</p>	<p>Mouse Locus</p> <p>Mouse Locus</p> <p>Symbol: Tnf Name: tumor necrosis factor Other aliases: tumor necrosis factor-alpha, TNF-alpha, TNFalpha, TNF alpha, Tnfa, DIF, Tnfsf1a ID: MGI:104798 123110 Link to: RIKEN SciNetS MGI Position: Mm:17:35338335-35338941</p>	<p>Keyword</p> <p>P value: 2.47E-3184 Keyword: cancer</p>
--	---	--	--	---

Gene2

5. **TGFA . transforming growth factor, alpha**

<p>Interval</p> <p>Human Locus</p> <p>Symbol: TGFA Name: transforming growth factor, alpha P value: 1.73E-2248 ID: HGNC:11765 set as target 2 Link to: RIKEN SciNetS HGNC Position: Hs:2:70874412-70781925</p>	<p>Mouse Locus</p> <p>Mouse Locus</p> <p>Symbol: Tgfa Name: transforming growth factor alpha ID: MGI:98724 5677 Link to: RIKEN SciNetS MGI Position: Mm:6:86145415-86225151</p>	<p>Co-citation</p> <p>P value: 5.35E-4613 4209</p>	<p>Mouse Locus</p> <p>Mouse Locus</p> <p>Symbol: Egf Name: epidermal growth factor ID: MGI:95290 45703 Link to: RIKEN SciNetS MGI Position: Mm:3:129380496-129458234</p>	<p>Keyword</p> <p>P value: 1.73E-2248 Keyword: cancer</p>
--	---	--	--	---

フェノタイプ：
Cancer

<http://biolod.org/PosMed> で公開中

The screenshot shows the PosMed website interface. At the top, there is a search bar with the text "keyword" and a "search" button. Below the search bar, there is a "Search Categories" sidebar with options: Gene, Bioresource, Phenotype, Disease, Metabolite, Drug, Researcher, Document set, and Database. The main content area features a diagram illustrating the concept of "genetic intervals(s)" and "filtering" based on a "Phenotypic Keyword". The diagram shows a chromosome with a highlighted region, a funnel labeled "Pos Med™" filtering information, and a red arrow pointing to a specific gene. Text on the page includes: "Tutorial in 2013 [English] Sample trial (search 'diabetes' related mouse bioresource of RIKEN Bio Resource Center)", "PosMed: 'Positional MEDLINE' assists your positional-cloning studies", "Since whole genome sequences were first elucidated, knowledge-based ranking of candidate genes has become one of the most important bioinformatics tasks in the forward-genetics and positional-cloning approaches to identify phenotype-responsible gene mutations.", "This task requires creating a form of artificial intelligence that can solve a genetic researcher's problems by learning computationally a vast amount of information accumulated in documents and published data.", "We have developed a system named 'PosMed,' an artificial intelligence that guides you to the key information waiting to be discovered in the sea of data.", "Please check our tutorial for details.", "PosMed: inferential search of disease-related genes and bioresources", and "Bioresource search provides not only normal direct keyword search but also inference searches using various biomedical databases, including OMIM, MEDLINE, ICD-10. These indirect search functions are unique and not provided at any other bioresource databases." The RIKEN Mouse Bioresource Search logo is also visible.

月間ユニークユーザー 16,000 人 (昨年12月)

Biophenome Linked Open Databases

<http://biolod.org>

The screenshot shows the BioLOD website homepage. At the top, there is a navigation menu with links for Home, BioLODについて, データベースリスト, チュートリアル, ユースケース, リンク, and ニュース. Below the navigation is a search bar with the placeholder text "Enter keyword..." and a "Search" button. The main content area is divided into several sections:

- BioLOD (Biophenome Linked Open Databases)**: A section describing the database's purpose and capabilities, including a search bar and a "Search" button.
- Reference papers of this database:** A list of three scientific papers related to the database.
- BioLOD data statistics**: A table showing the number of Databases, Classes, Instances, and Literature for Open Data, Not Open Data, and Total (Browsable).
- Integrated databases in various field**: A list of four integrated databases with their respective instance counts.
- Facebookもチェック**: A section for social media links, including a Facebook link.

BioLOD data statistics

	Databases	Classes	Instances	Literature
Open Data	139	715	9,366,352	272,163
Not Open Data	66	228	286,429	0
Total (Browsable)	205	943	9,652,781	272,163

as of January 2013.

Integrated databases in various field

- マウス系統 - マウス系統 (理研バイオリソースセンター) (5115 Instances)
- 細胞株 - 細胞株 (理研バイオリソースセンター) (3611 Instances)
- シロイヌナズナ変異株観察情報 - シロイヌナズナ変異体フェノーム情報 (14629 Instances)
- 表現型の変異 - 植物フェノーム (487 Instances)

今年度の成果

■ リソース基本情報の整備

- マウス: 理研内(昨年度4929件)+新規186件、理研外 更新12911+新規6287件
- 培養細胞: 新規3611件
- 植物: 新規44310件追加中

■ プロパティ標準化作業とフェノーム記述の識別子体系化と評価

- 一般的に使われるオントロジーを用いて、表現型データ記述の汎用RDFスキーマを作成: 34データベースが対応済

■ フェノーム情報の整備

- マウス: 新規 874件
- 培養細胞: 新規 5000件
- シロイヌナズナ: 以下のデータをPosMedによる推論検索対象に追加
 - 変異株3763種類の表現型情報: 14631件
 - 文献キュレーションによるシロイヌナズナ表現型情報: 824件

■ リソース提示のインターフェース開発(フェノーム利用ワークフロー開発)

- RDF形式データを用いたバイオリソース提示インターフェース
- 疾患モデル推論システム開発(連携支援システム)

実施計画

H24年度	<ul style="list-style-type: none">■ プロパティ標準化作業と識別子体系化作業を実施し、その成果を公開■ マウス系統と細胞株と植物株の統合■ それらのフェノーム情報の統合■ フェノーム情報を活用した連携研究を支援するシステム構築（おすすめ機能）
H25年度	<ul style="list-style-type: none">■ プロパティ標準や識別子体系の評価と向上（アップロード機能の実装とデモ）■ 細胞株培養条件および微生物その他の有用性情報の統合化作業■ バイオリソースを対象とした次世代シーケンサーデータの統合化■ ワークフローとしてのPosMed推論検索の高度化と普及化■ 構築したシステムやデータのNBDCへ受け渡し



研究ニーズを捉えるインターフェースの追求

- ✓ GWAS, コホート研究からモデル生物での検証へ
- ✓ フェノームデータでバイオリソースの価値を高める