2015/10/21 計算生命科学の基礎Ⅱ

「生物システムの設計：システム生物学から合成生物学へ」

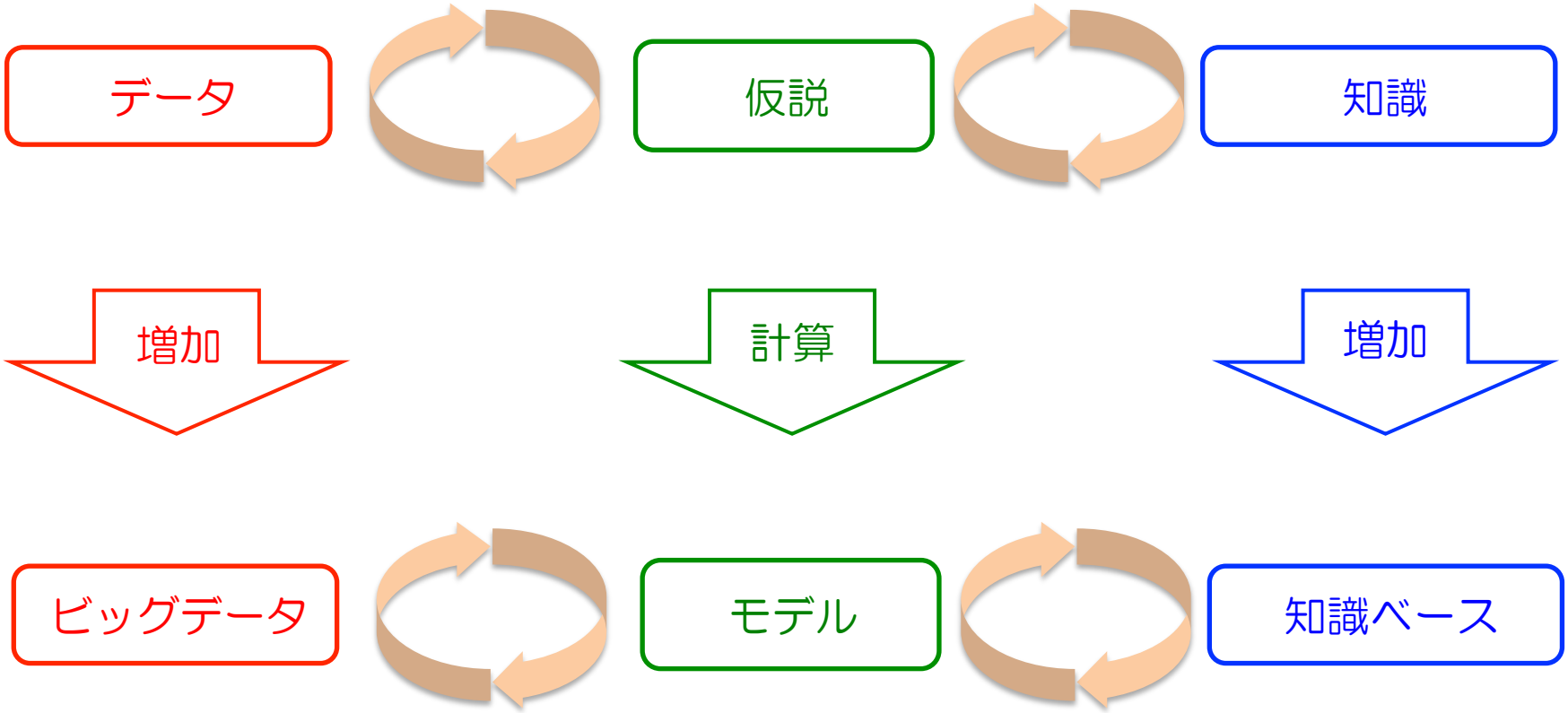
荒木 通啓　（Michihiro Araki）
araki@port.kobe-u.ac.jp

神戸大学　自然科学系先端融合研究環
Organization of Advanced Science and Technology
Kobe University

内容

- はじめに
- 遺伝子回路設計
- 代謝経路設計

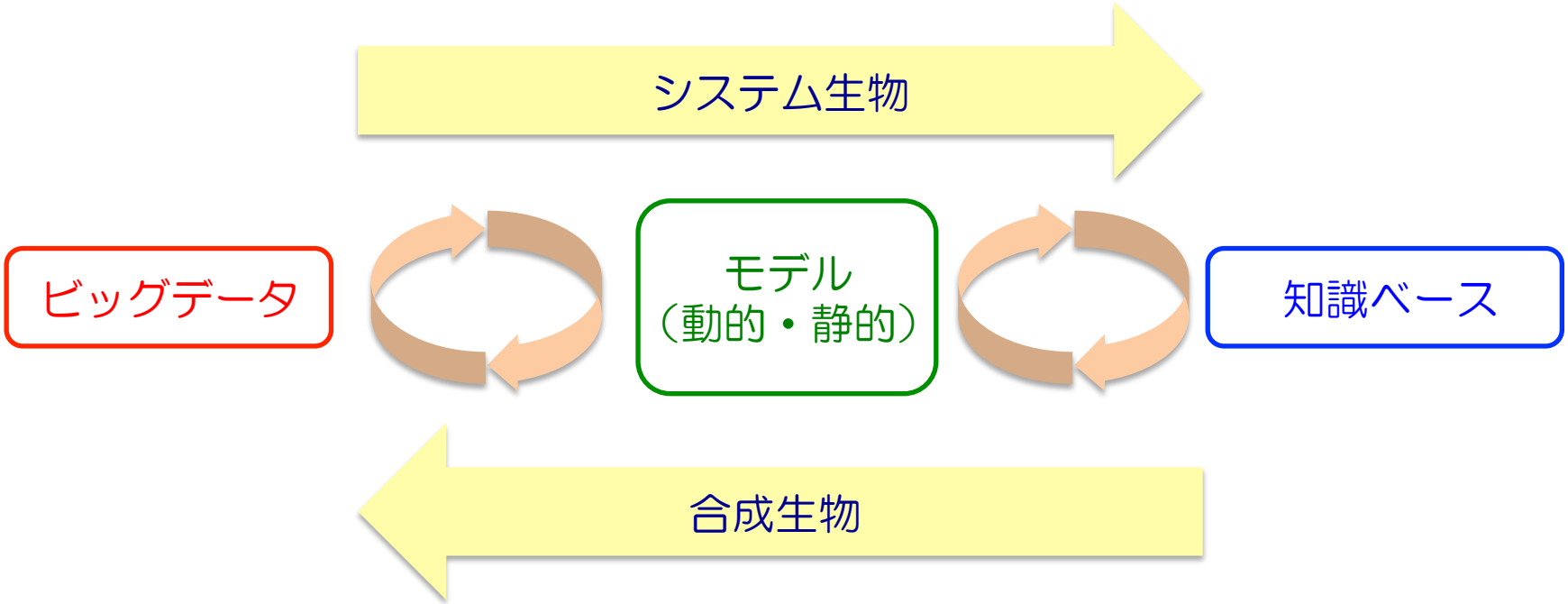
はじめに



計算・情報科学の必要性

はじめに

キーワード）
要素同定、アノテーション、データベース、ネットワーク解析、システム解析、パラメータ同定、シミュレーション、階層性、モチーフなど

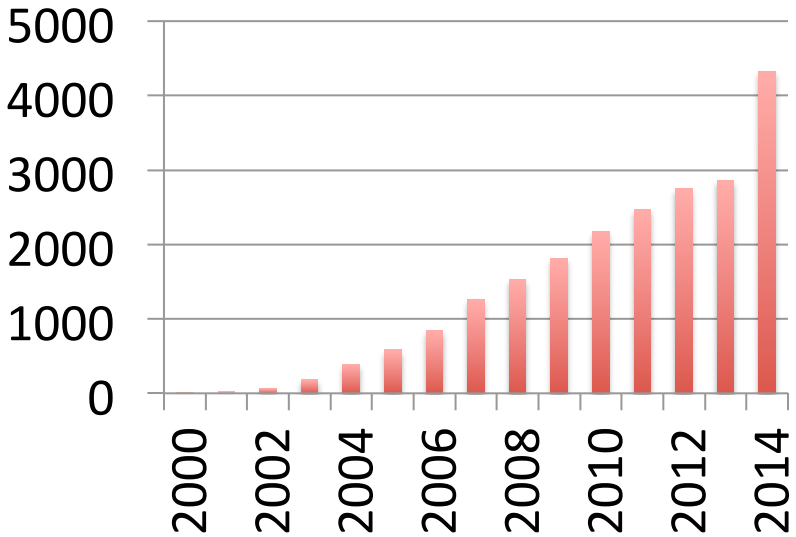


キーワード）
ゲノム細胞工学（ミニマムゲノム、ゲノム合成）、非天然塩基・アミノ酸、ゲノム編集、BioBrick、人工細胞、遺伝子回路設計、代謝経路設計など

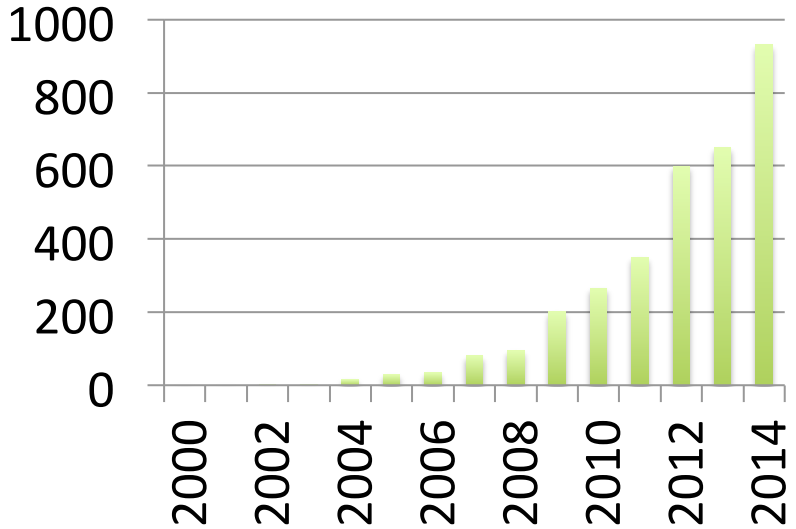
はじめに

PubMed検索の結果

システム生物学
Systems Biology



合成生物学
Synthetic Biology



生物システム

分子情報

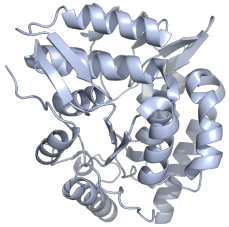
- 塩基配列
  - DNA、RNA

atgtccactgcggtcctggaaaacccaggcttgggc
aggaaactctctgactttggacag

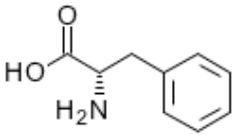
- アミノ酸配列
  - タンパク質

MSTAVLENPGLGRKLSDFGQETSYIED
NCNQNGAISLIFSLKEEVGALAKVLRLF

- 立体構造
  - タンパク質



- 化学構造
  - 化合物

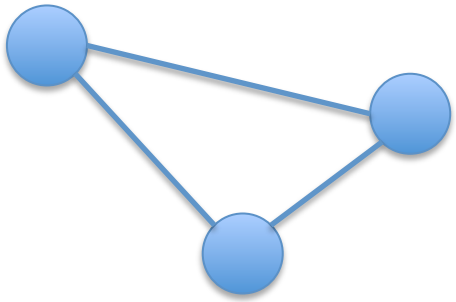


C00079

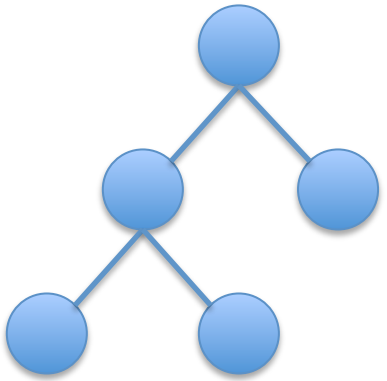
生物システム

ネットワーク情報

- グラフ
  - ペア情報
  - パスウェイ



- ツリー
  - 分類情報
  - オントロジー



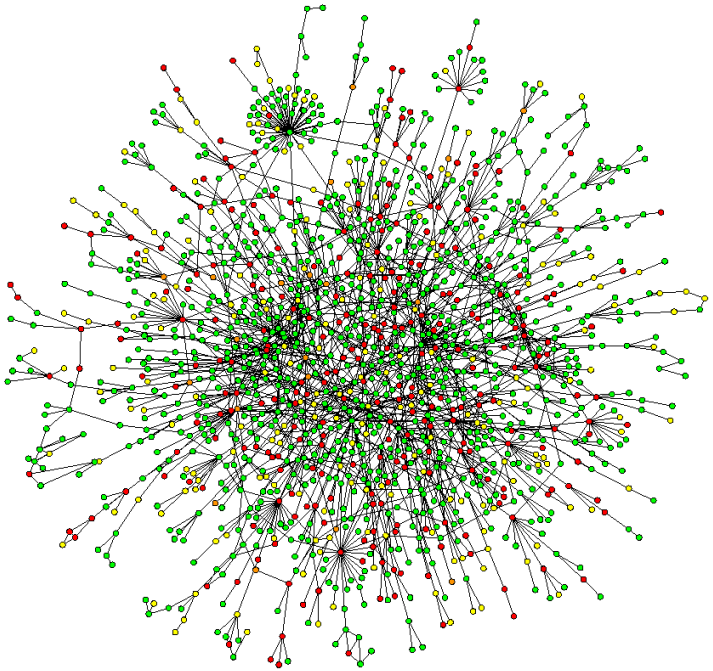
生物システム

表現方法：データベース

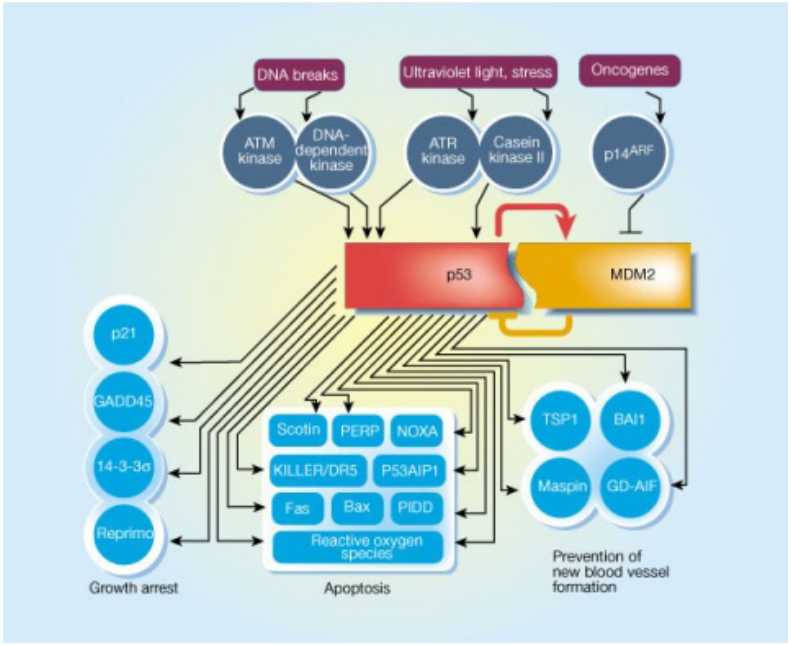
- 要素
  - 遺伝子、タンパク質、化合物
- ペア
  - 相互作用、制御関係、反応
- グラフ
  - パスウェイ
- 階層リスト
  - オントロジー

生物システム

遺伝子回路



Barabasi A-L *et al.*



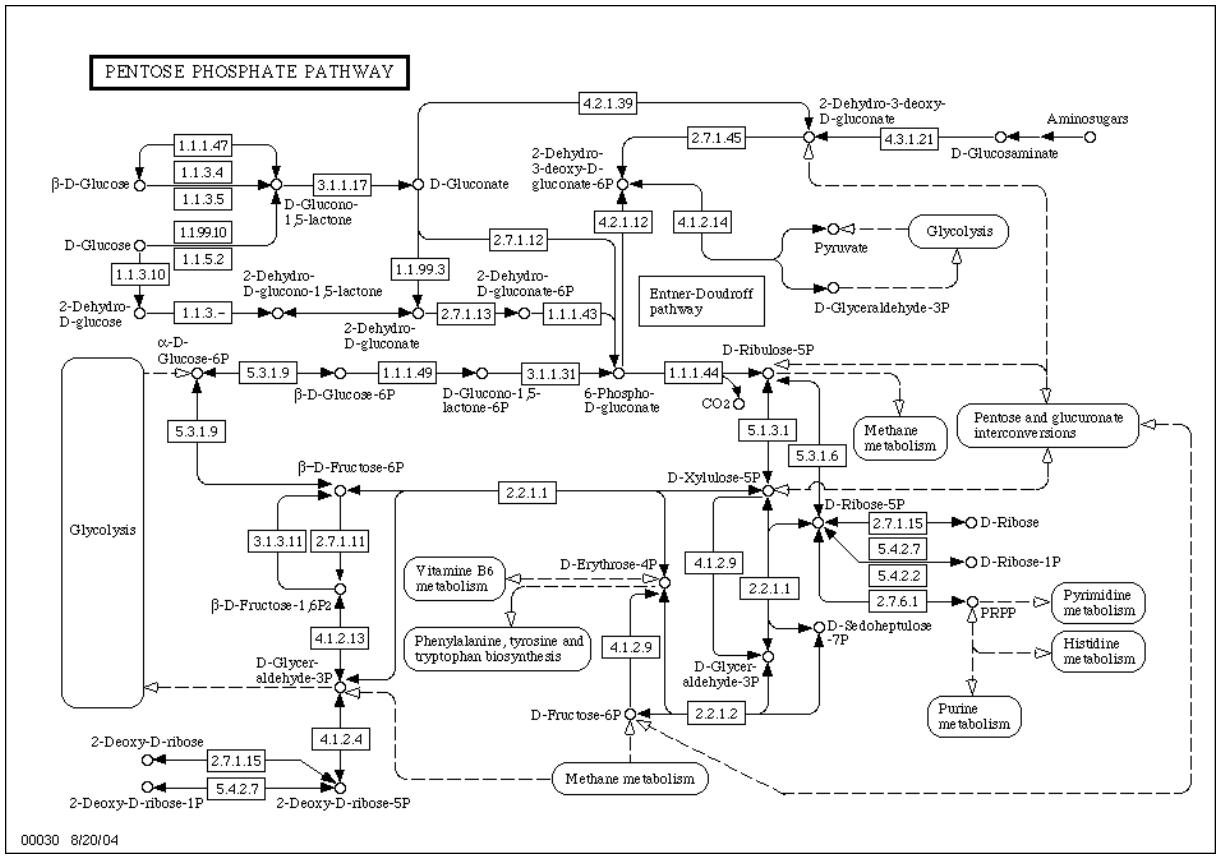
Vogelstein B *et al.*

ランダムな進化　→　一般的な法則・モデル？　→　遺伝子回路設計？

生物システム

代謝経路

KEGG PATHWAY



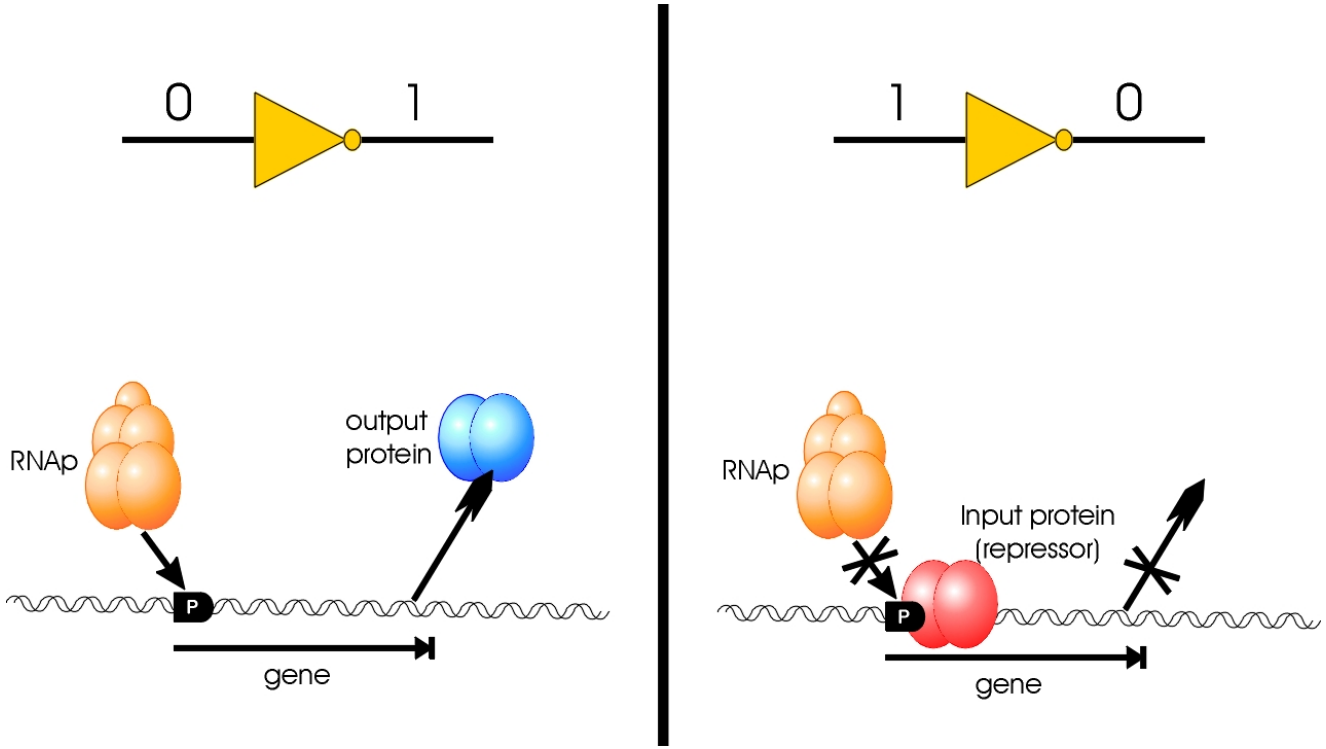
代謝経路制御・拡張　→　一般的な解析モデル？　→　代謝経路設計？

内容

- はじめに
- 遺伝子回路設計
- 代謝経路設計

遺伝子回路：転写制御

例）リプレッサー



転写制御モデル

- 入力関数：転写因子の効果の強さ

$$\text{Y の産生速度} = f(X*)$$

- アクチベーターのヒル関数

$$f(X^*) = \frac{\beta X^{*n}}{K^n + X^{*n}}$$

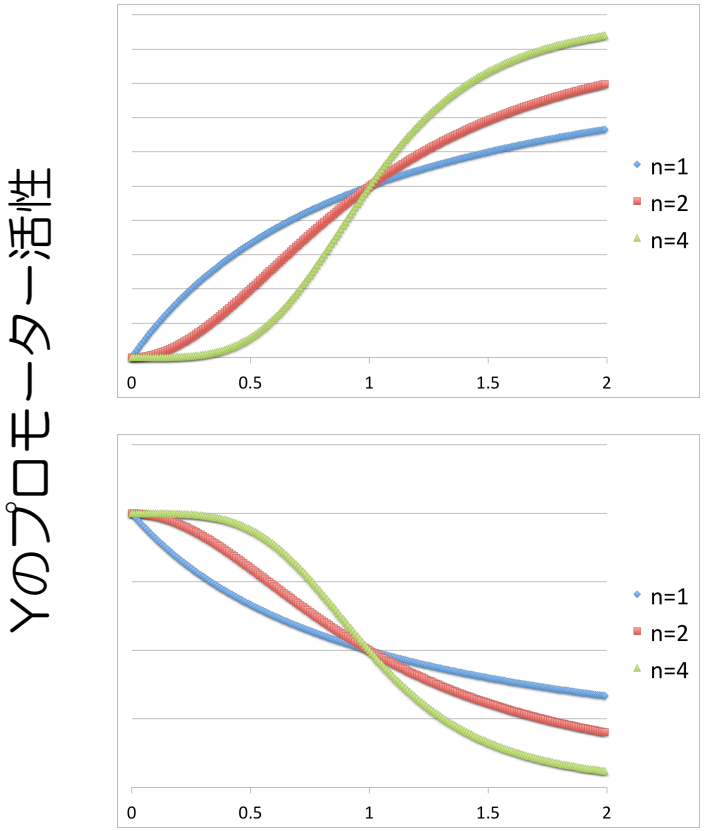
$$f(X^*) = \beta\theta(X^* > K)$$

- リプレッサーのヒル関数

$$f(X^*) = \frac{\beta}{1 + (\frac{X^*}{K})^n}$$

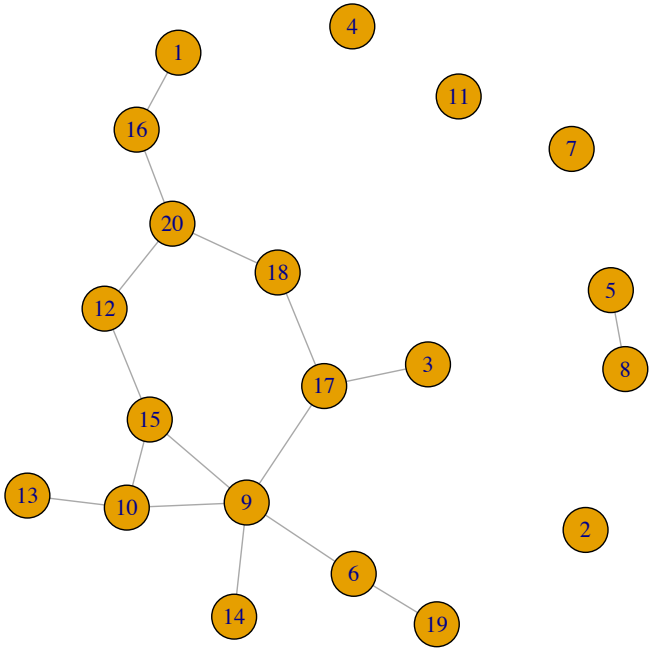
$$f(X^*) = \beta\theta(X^* < K)$$

β：最大発現レベル
K：活性化係数
n：ヒル係数
Θ：ステップ関数



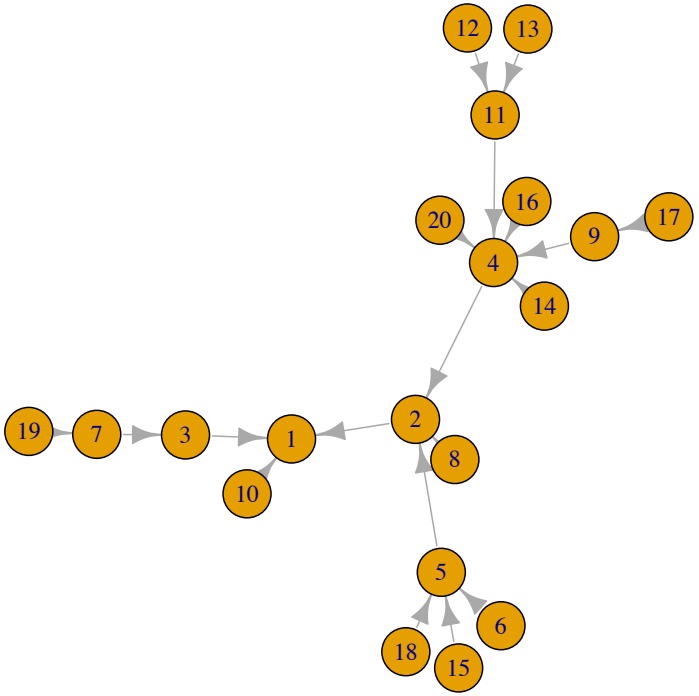
ネットワーク

ランダムネットワーク



G(n,p) or G(n,m) graph:
- n個のvertices
- 任意の２vertices間にedgeがある確率p
- 任意の２vertices間にm個のedgeがある

実際のネットワーク

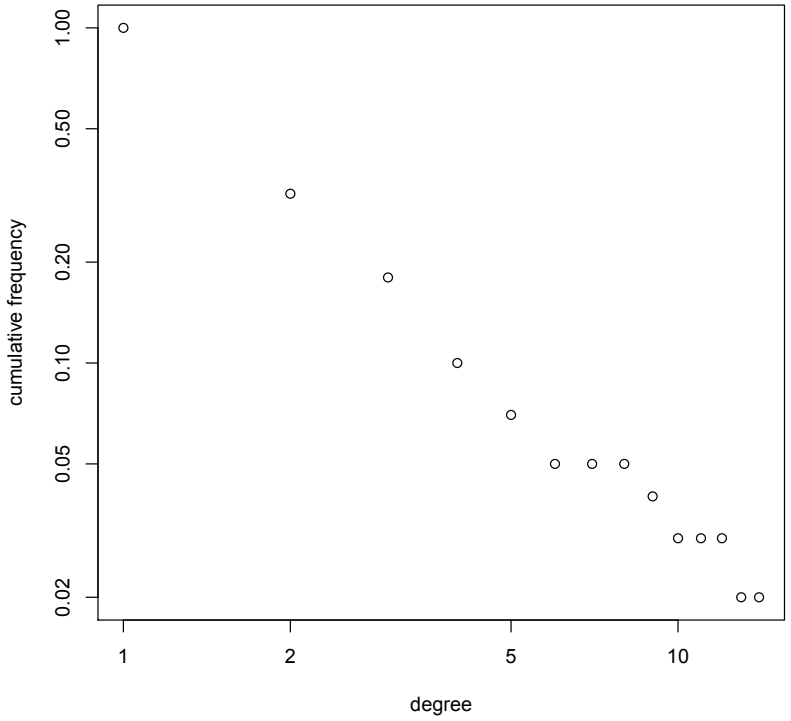


一つのvertexからスタートしてvertexを追加
vertexが選ばれる確率：
- P[i] ~ k[i]^alpha + a

ネットワーク

スケールフリー性

- エッジの数をノード毎に数えてプロット

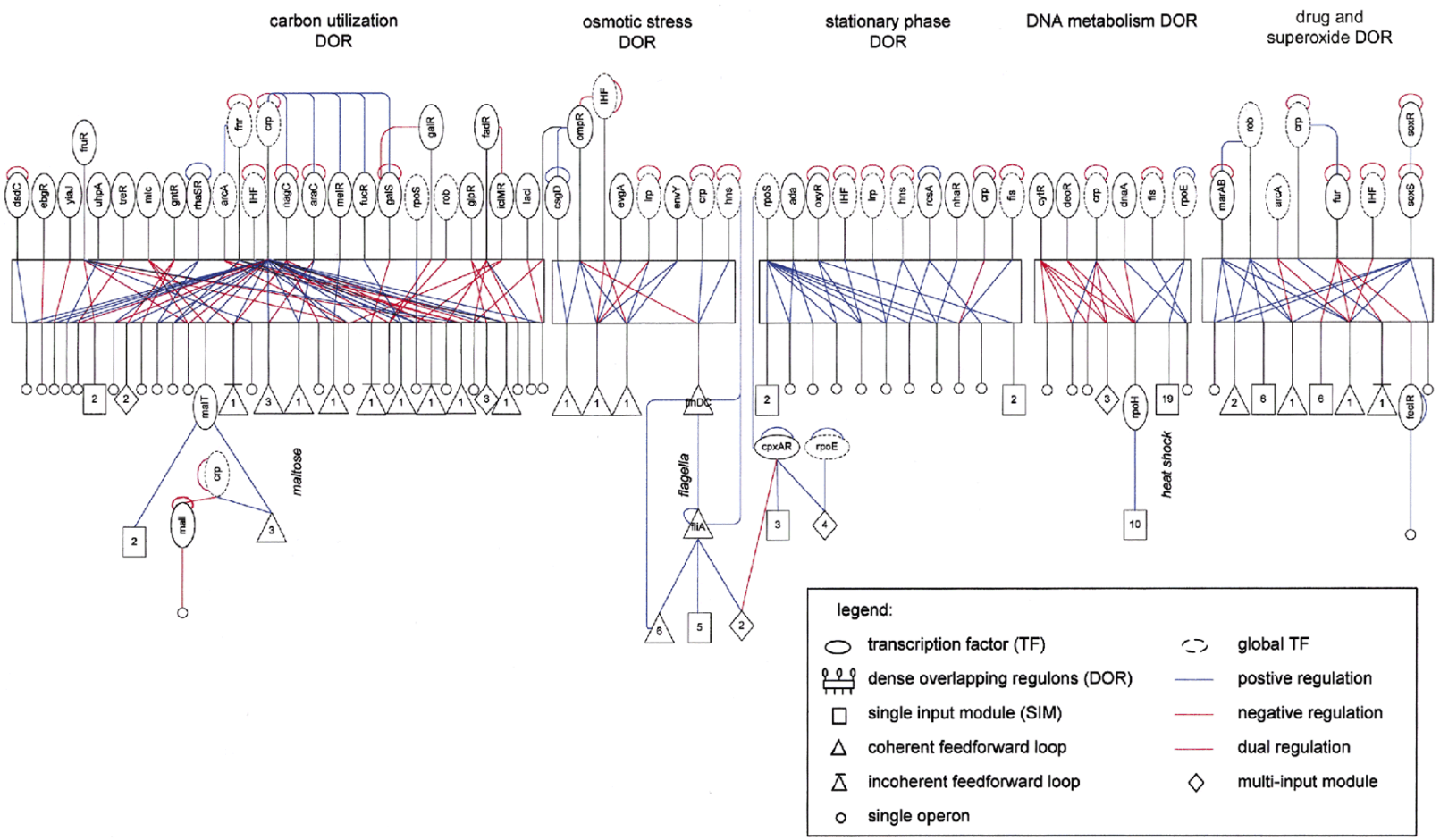


例）　WWW
Social network
Biological network

特徴）　Modular structure
ハブの存在
短い経路
冗長性

転写制御ネットワーク

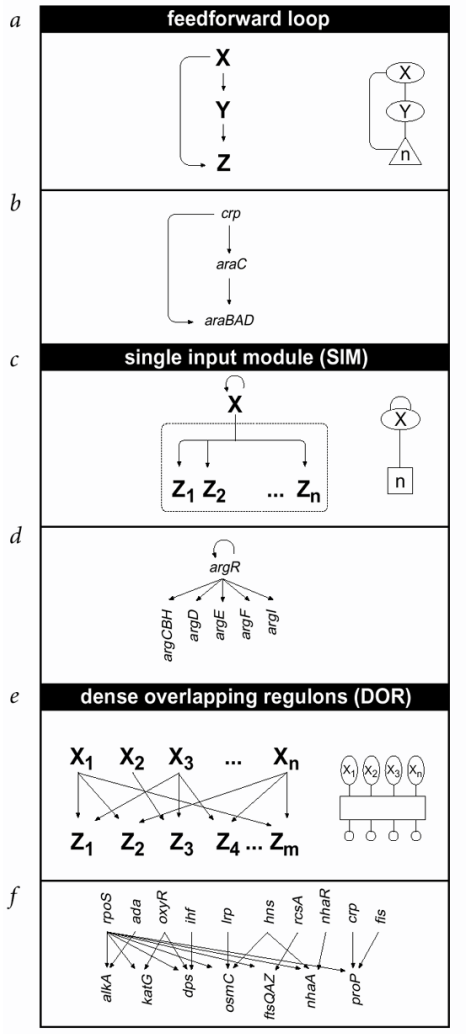
例）大腸菌：Regulon DB



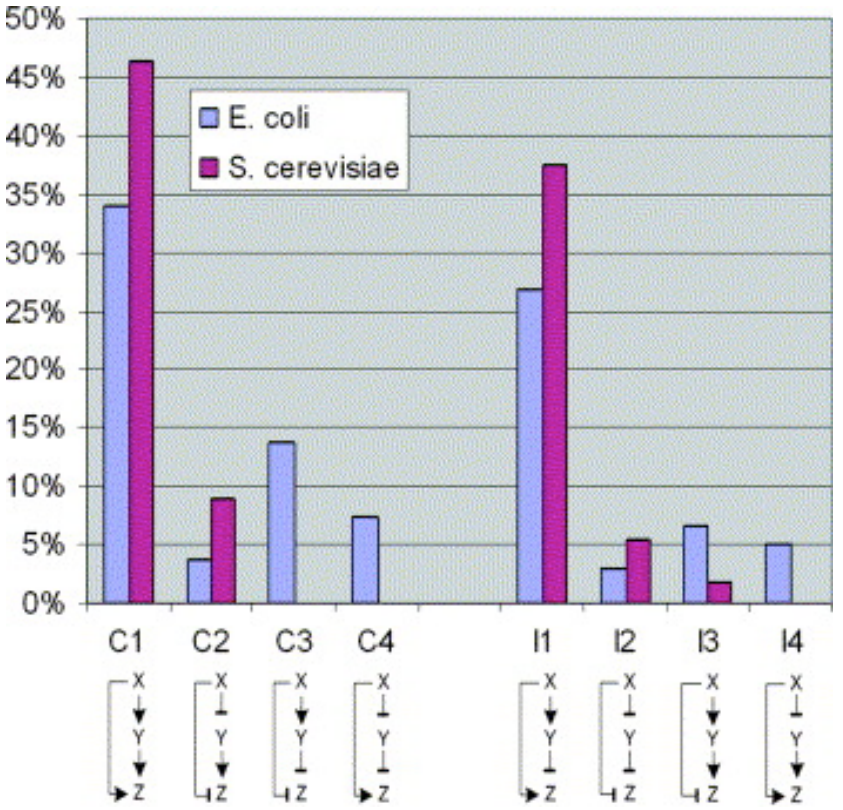
Shen-Orr *et al.* Nat Genet 31, 64-68 (2002)

ネットワークモチーフ

モチーフ

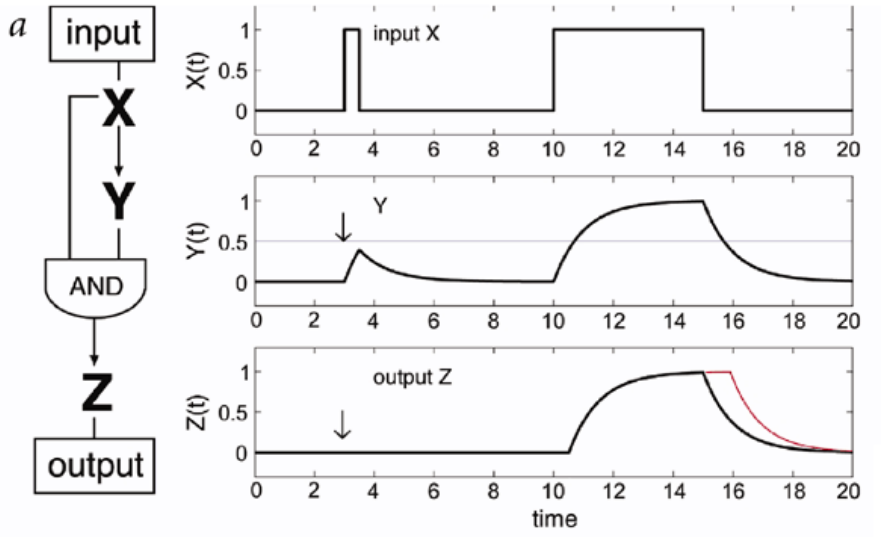


フィードフォワード



Shen-Orr *et al.* Nat Genet 31, 64-68 (2002)
Mangan and Alon JMB 356, 1073-1081 (2006)

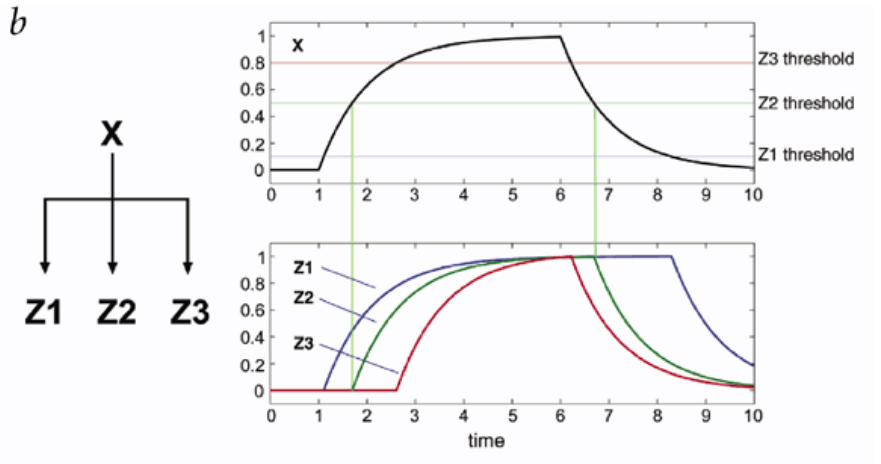
ネットワークモチーフ



フィードフォワード（FFL）

$$\frac{dZ}{dt} = \beta_z \theta(X^* > K_{xz}) \theta(Y^* > K_{yz}) - \alpha_z Z$$

オンオフ感知性の遅延



単入力モジュール（SIM）

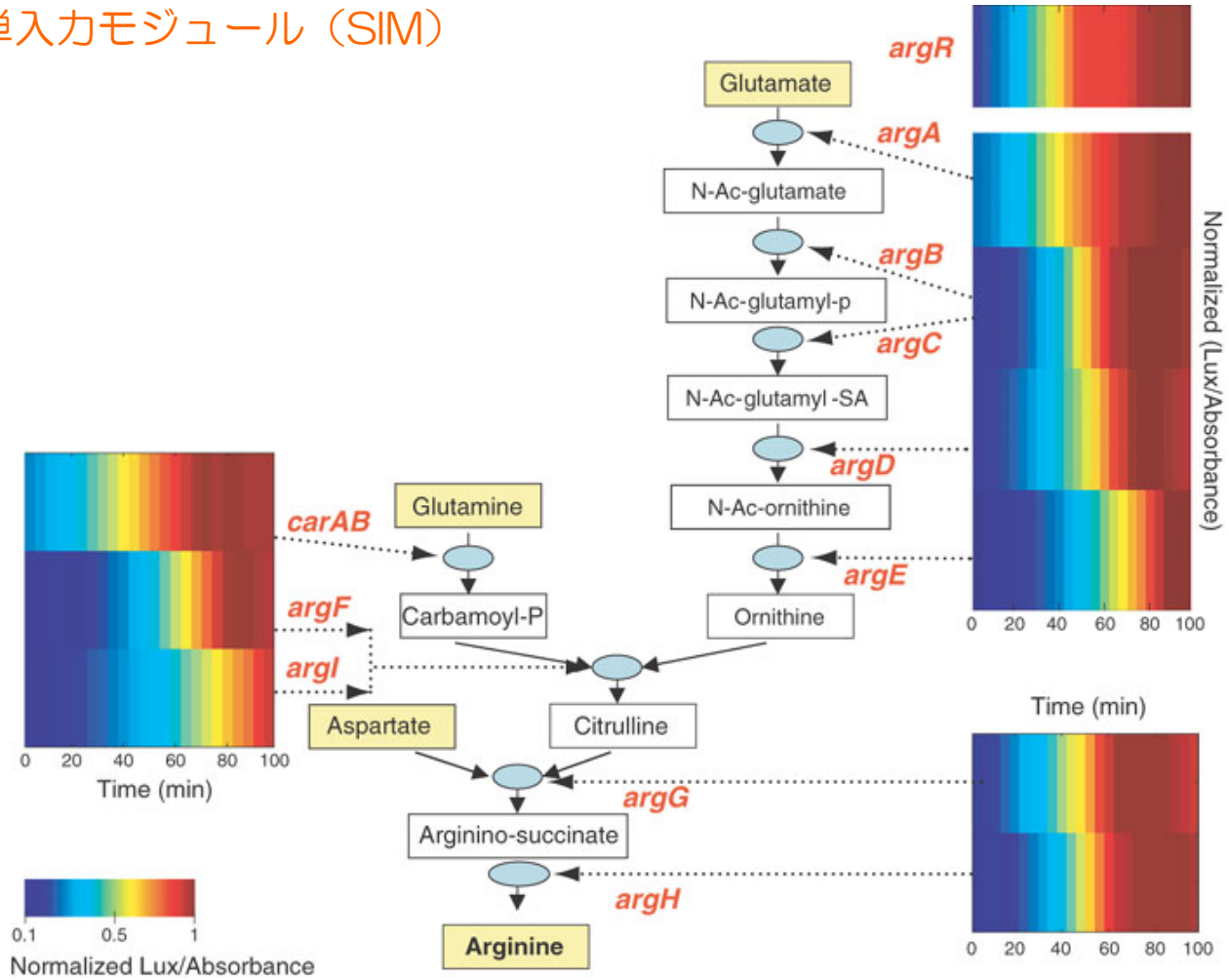
$$\frac{dZ_i}{dt} = \beta_{Zi} \theta(X^* > K_{xz_i}) - \alpha_{z_i} Z_i$$

遺伝子発現の時間プログラム

Shen-Orr *et al.* Nat Genet 31, 64-68 (2002)

ネットワークモチーフ

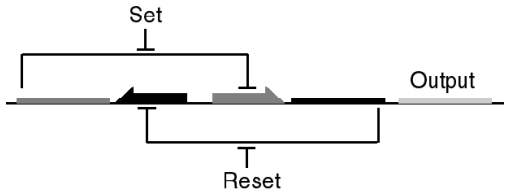
事例）単入力モジュール（SIM）

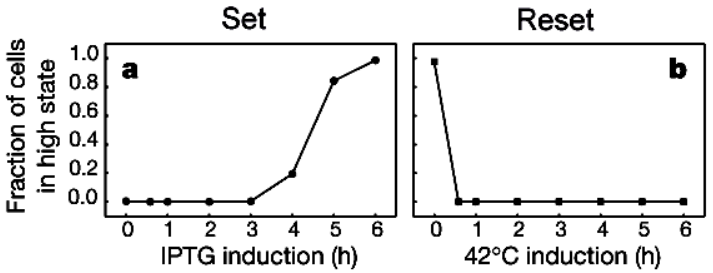


Zaslaver A *et al.* Nat Genet 36, 486-491 (2004)

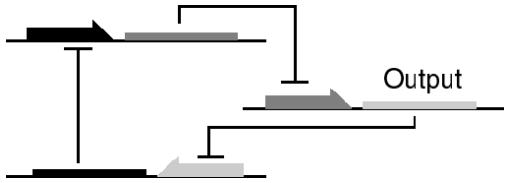
遺伝子回路設計：Toy Model

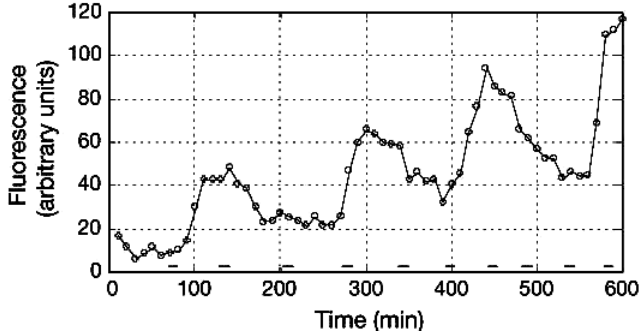
Memory (Gardner, *et al.*)



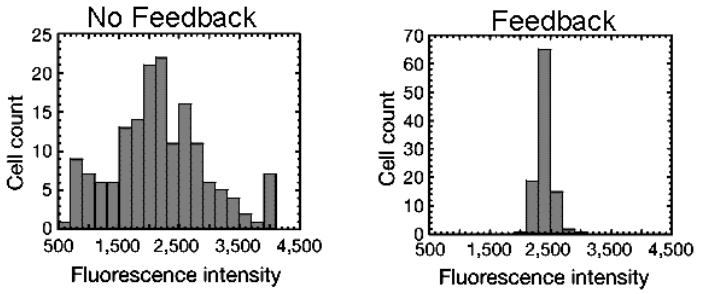


Clock (Elowitz, *et al.*)



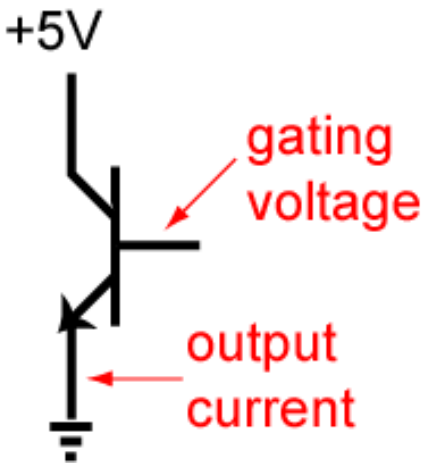


Noise Reduction (Becskei, *et al.*)

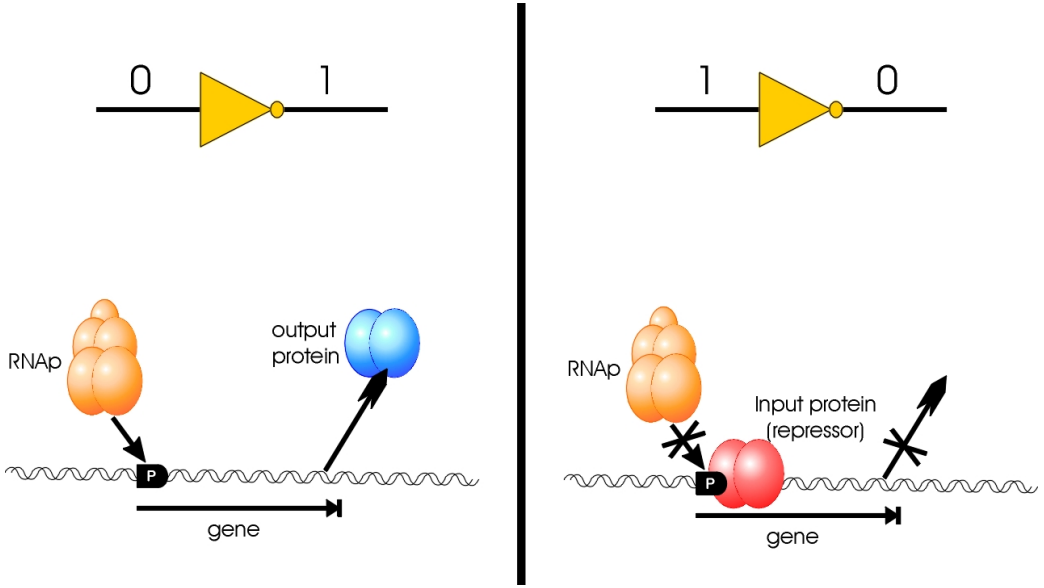


工学スイッチと遺伝子スイッチ

Electronic switch



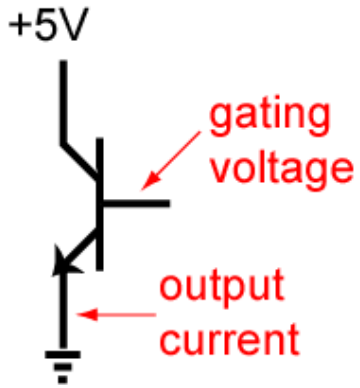
Genetic switch



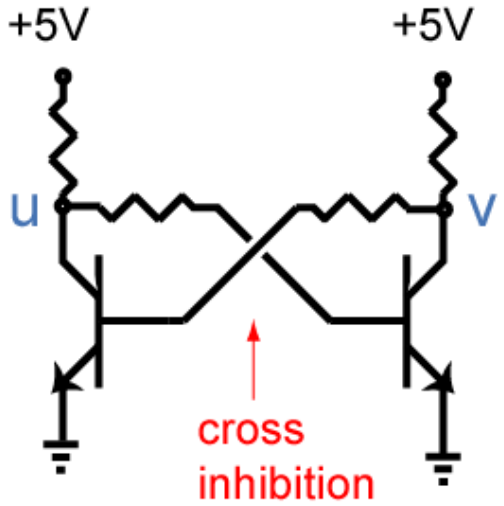
遺伝子トグルスイッチ

Electronic memory

Transistor

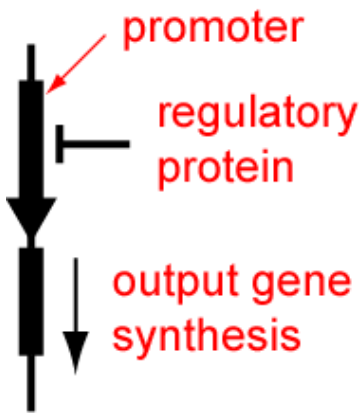


RS latch (bistable)

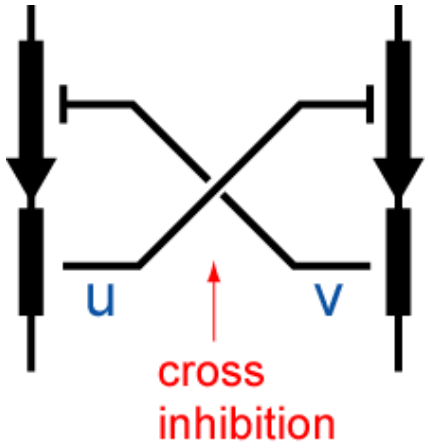


Biochemical memory

Repressible promoter

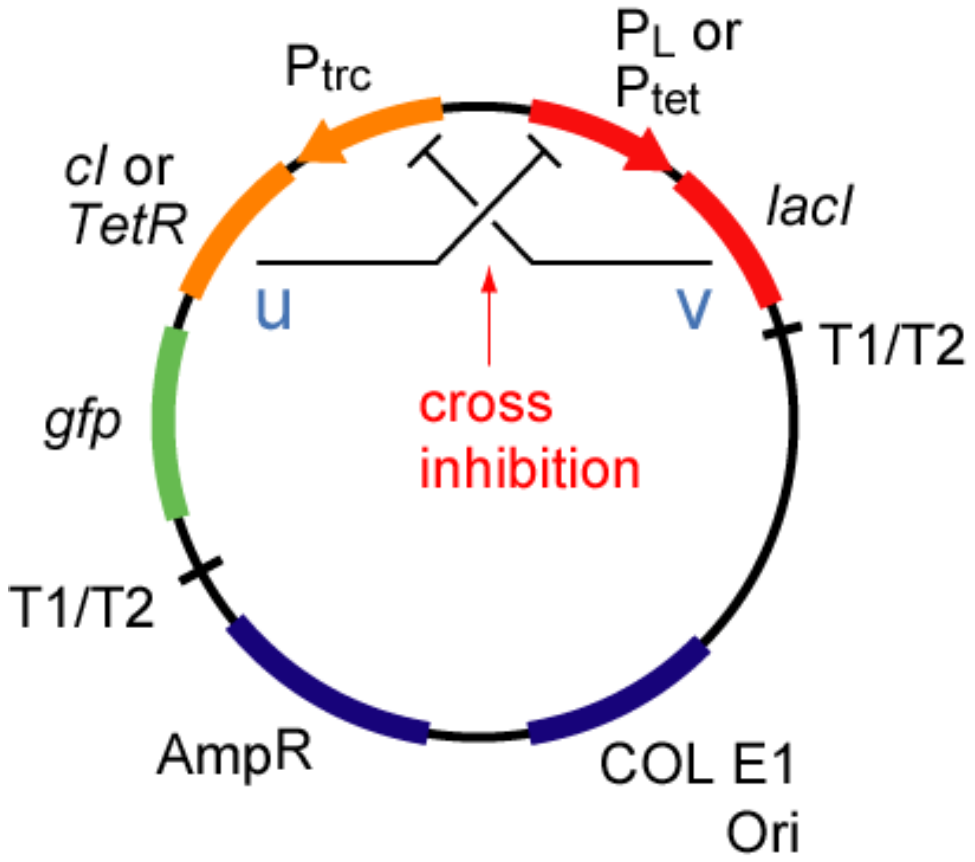


Genetic toggle switch (bistable)



遺伝子トグルスイッチ

Development of a genetic toggle switch in E. coli (biochemical memory)



Gardner TS, Cantor CR, Collins JJ
Nature 403, 339-342 (2000)

トグルモデル

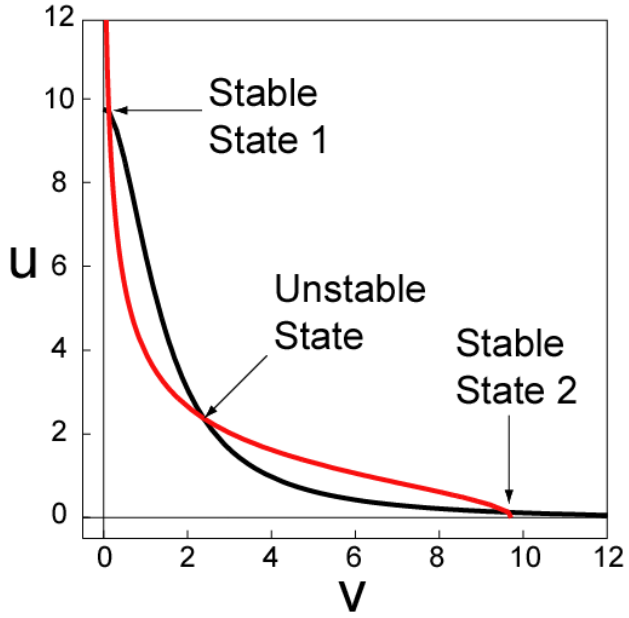
Nonlinear ODE model: reduced rate equations for transcription and translation

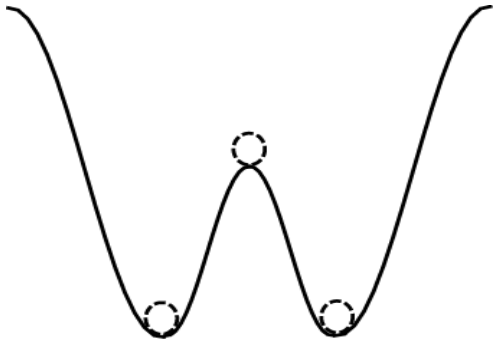
$$0 \cancel{\dot{u}} = \frac{\alpha_1}{1 + v^{\beta}} - u$$

$$0 \cancel{\dot{v}} = \frac{\alpha_2}{1 + u^{\gamma}} - v$$

$$u = \frac{\alpha_1}{1 + v^{\beta}}$$

$$v = \frac{\alpha_2}{1 + u^{\gamma}}$$

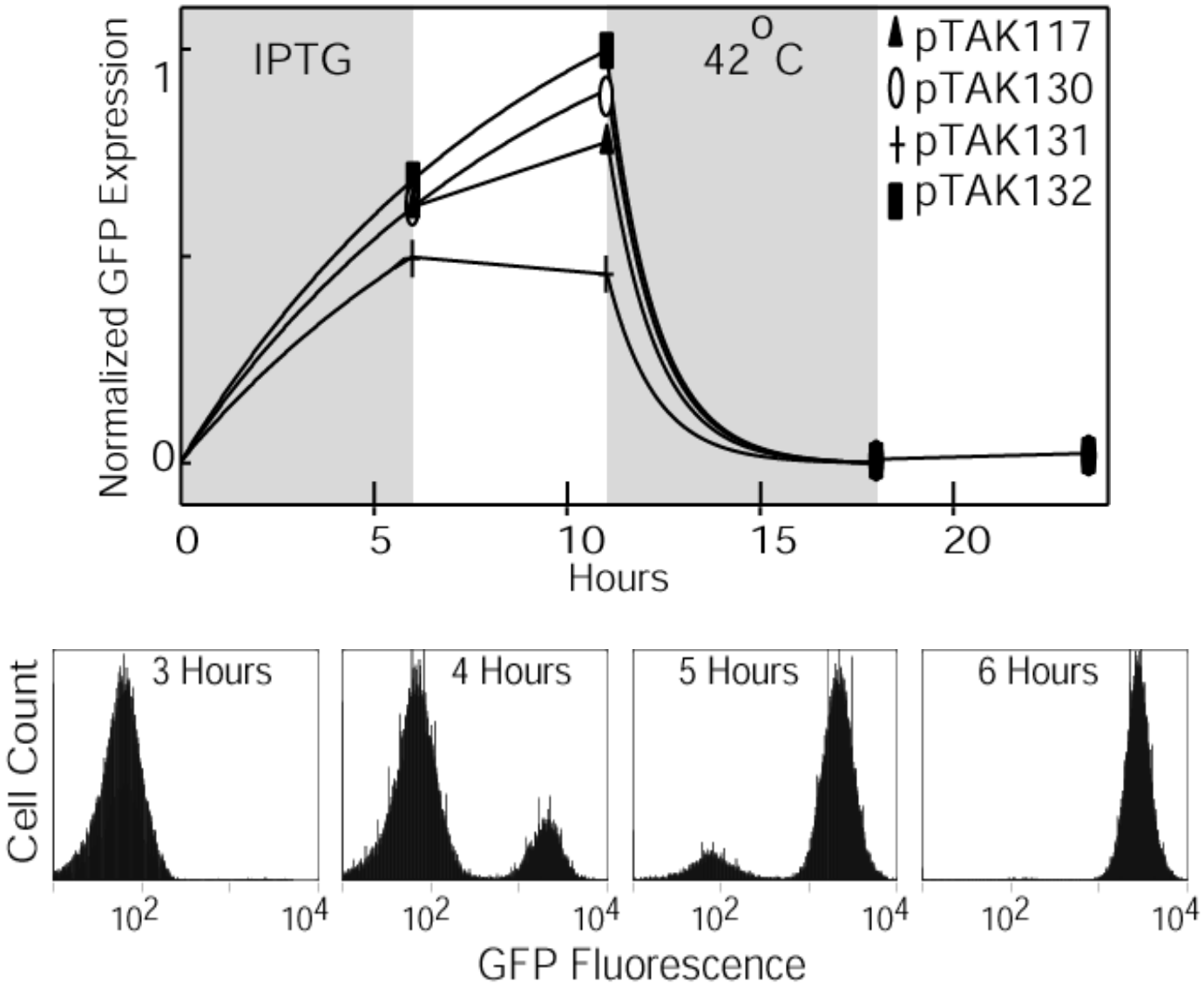




双安定状態（Bistability）

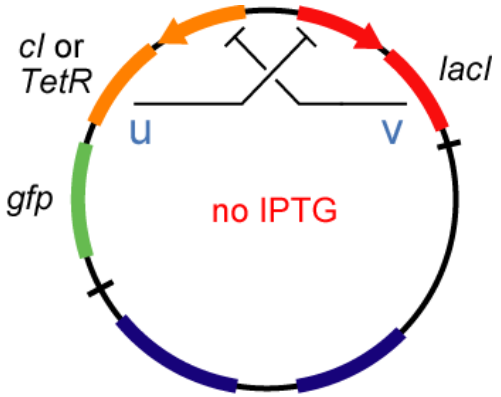
Bistable Switch

LacI/CI*ts* Toggle

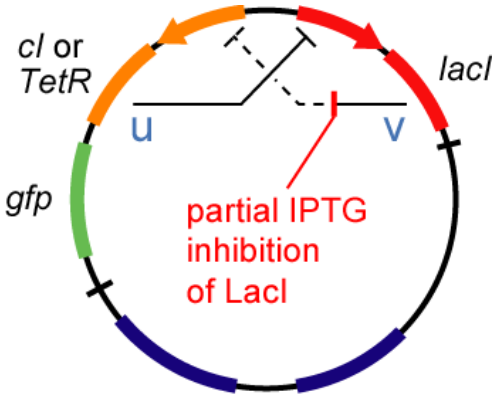


Control of Bistability

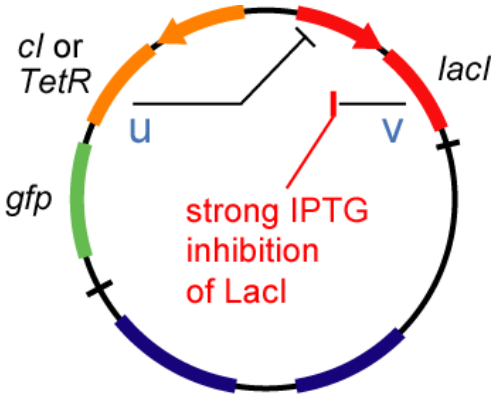
Bistablilty

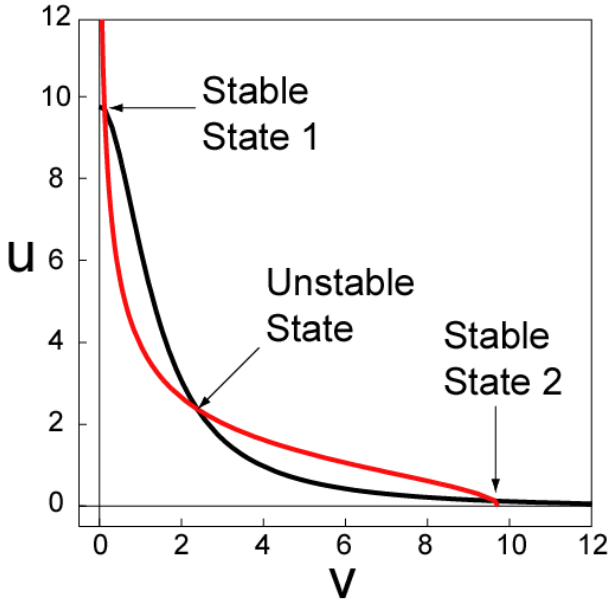


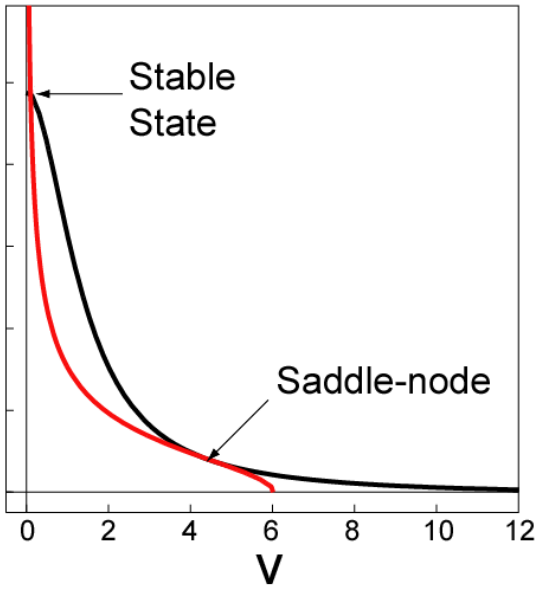
Bifurcation

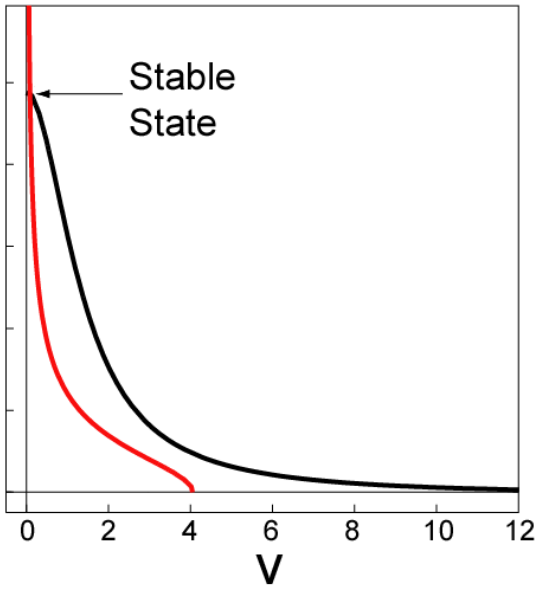


Monostability

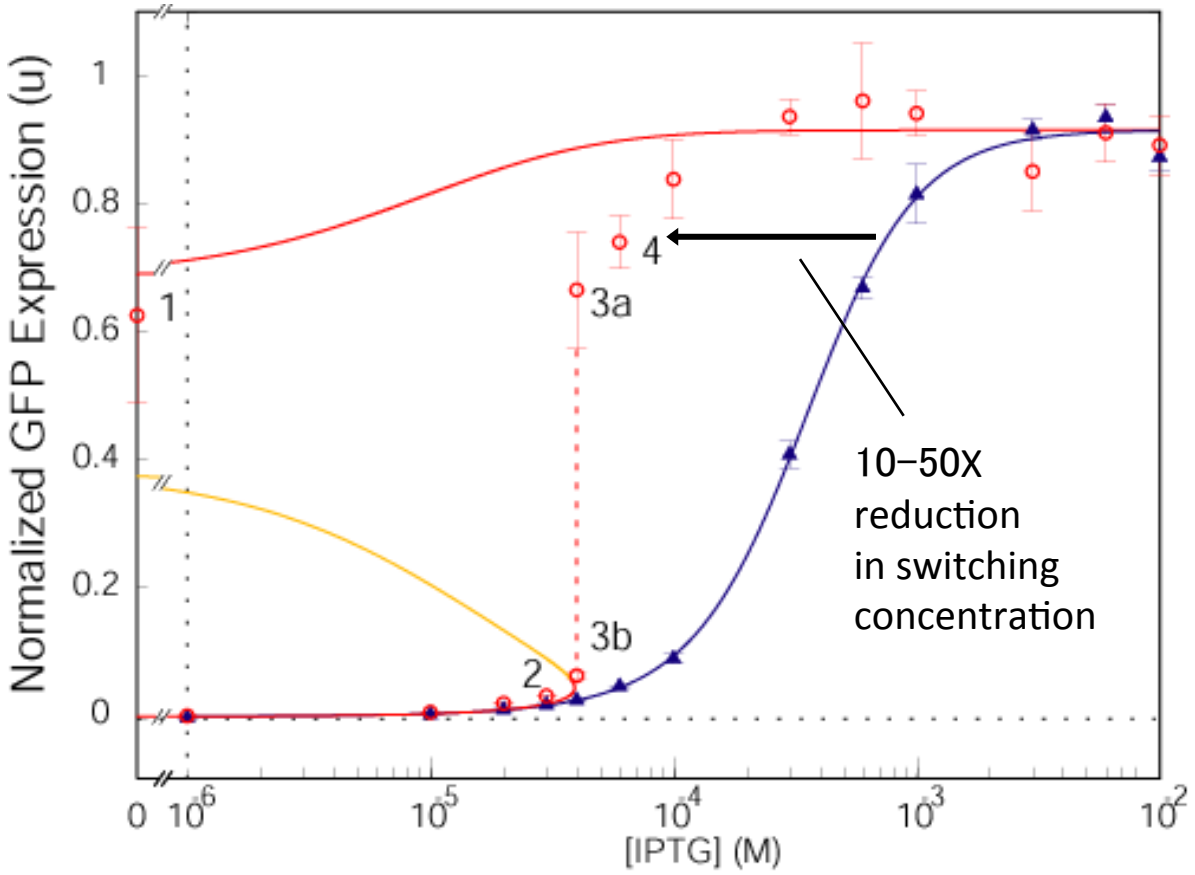




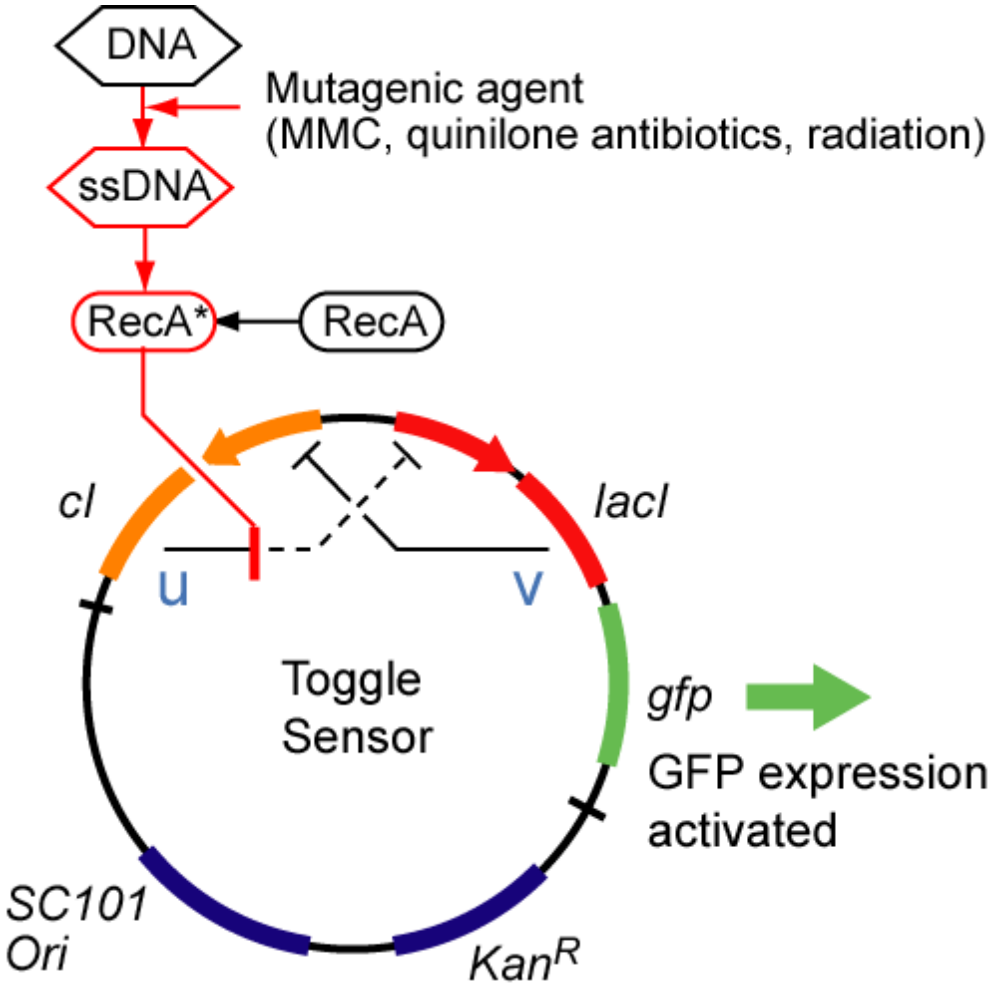


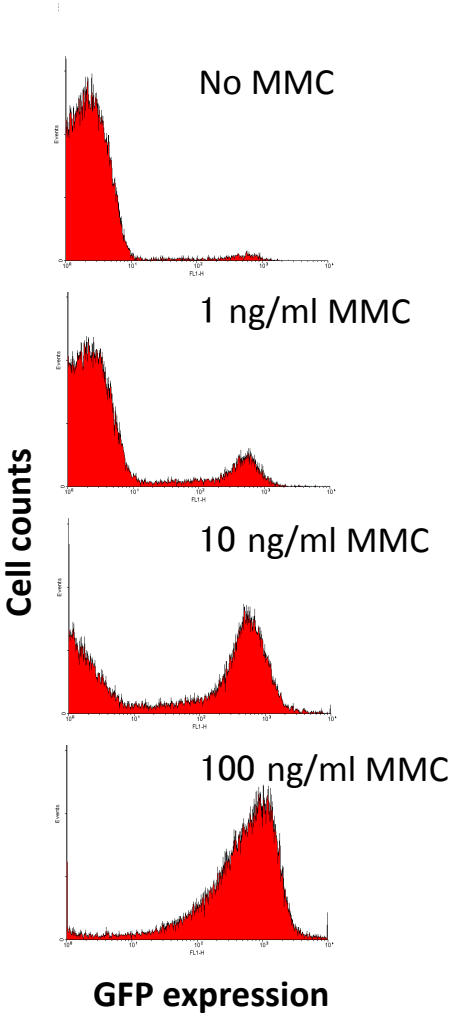


All-or-none switching



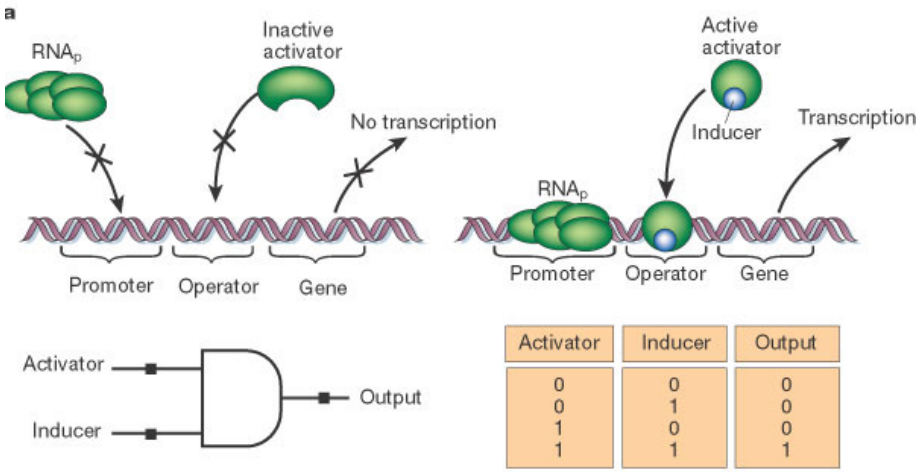
Programmable Cells

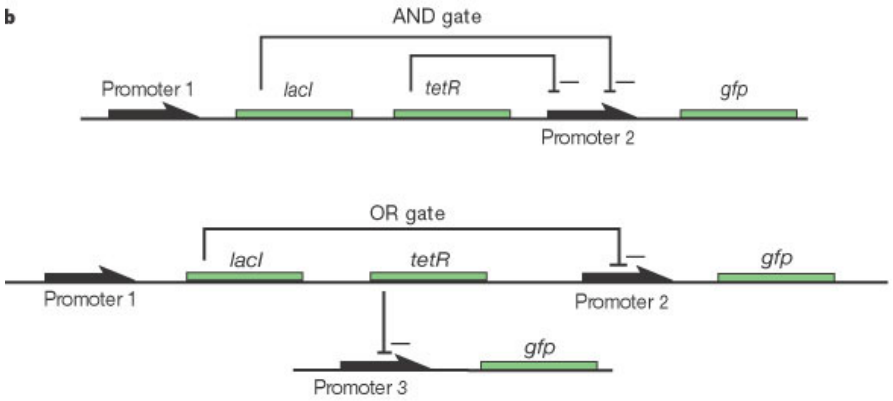


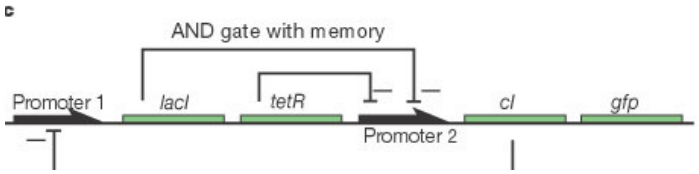


Kobayashi H, Kaern M, Araki M et al.
PNAS 101(22):8414-8419 (2004)

AND/OR GATE

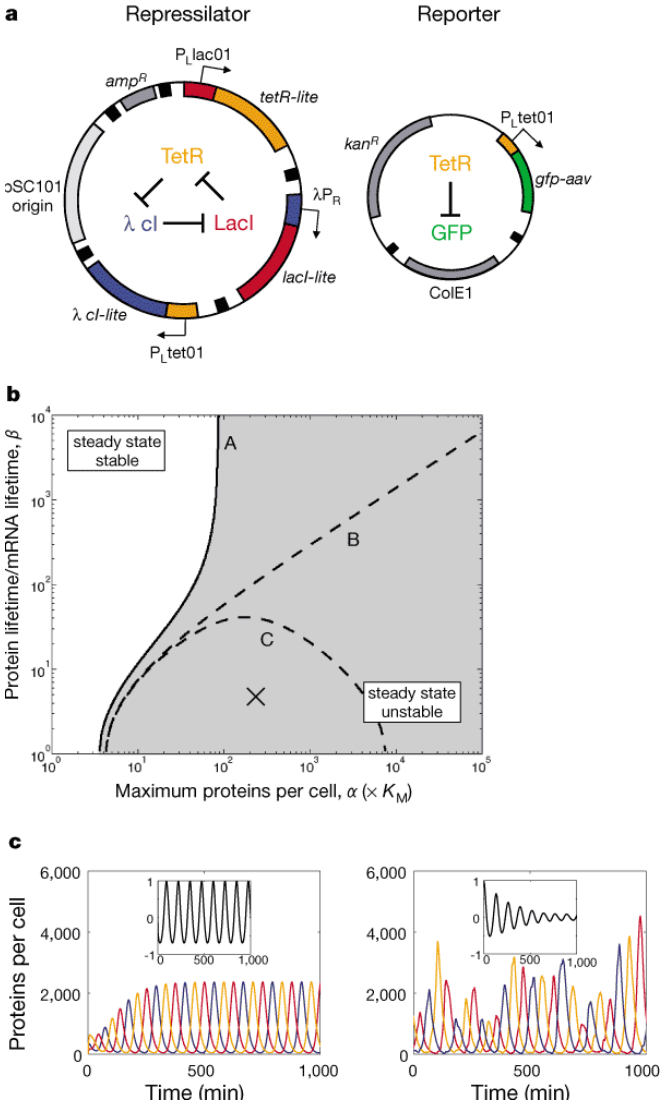






Hasty J, McMillen D, Collins JJ
Nature 420, 224-230 (2002)

Genetic Clock



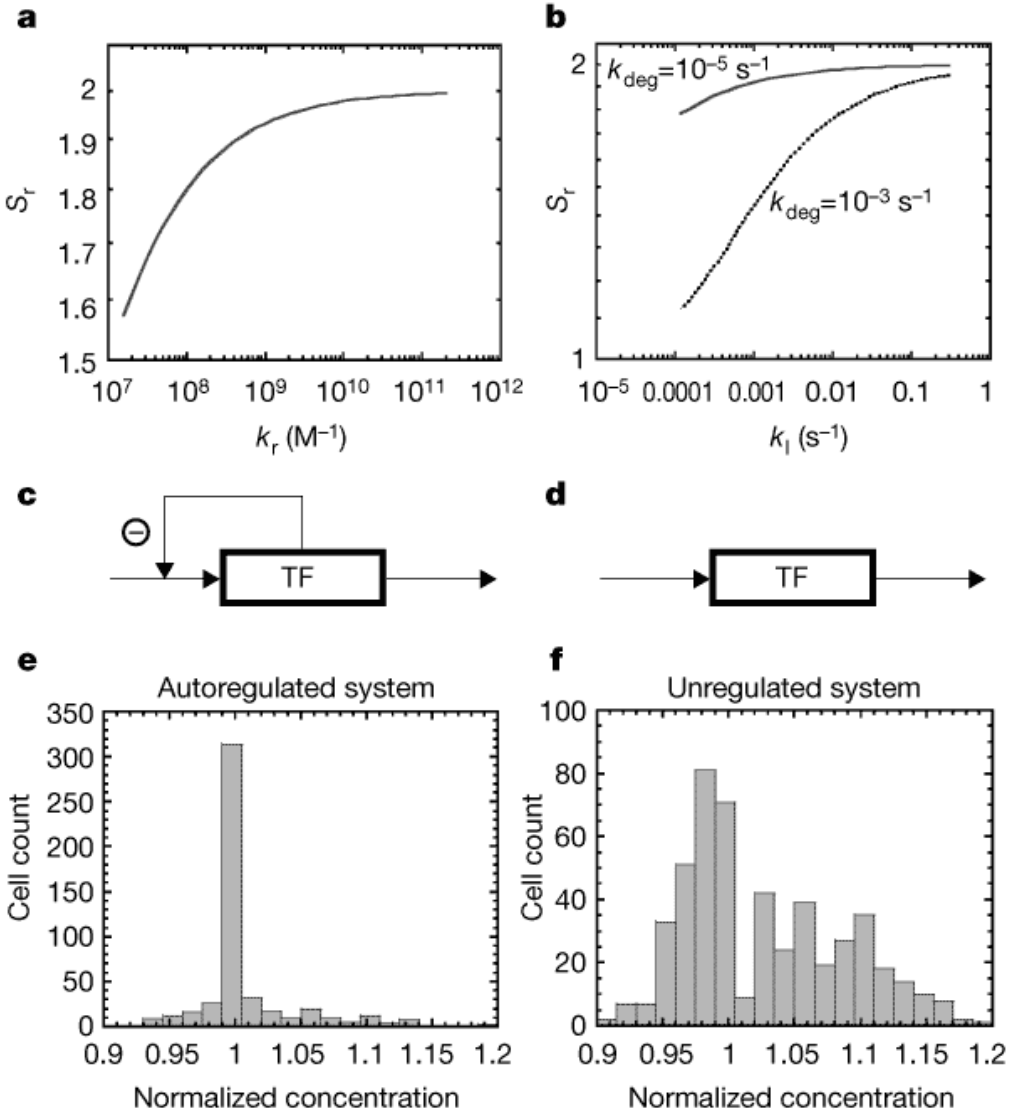
$$\dot{m}_i = -m_i + \frac{\alpha}{(1+p_j^n)} + \alpha_0$$

$$\dot{p}_i = -\beta(p_i - m_i)$$

Repressor-protein concentrations: pi
mRNA concentrations: mi
(where i is lacI, tetR or cI)

Elowitz M, Leibler S
Nature 403, 335-338 (2000)

Noise Reduction



$$S_{unreg} = f'_{unreg}(R^*) = -k_{deg}$$

$$S_{auto} = f'_{auto}(R^*) = -\frac{nk_pPk_Iak_r}{(1+k_pP+k_rR^*)^2} - k_{deg}$$

$$S_r = \frac{S_{auto}}{S_{unreg}}$$

Repressor-protein concentrations: R
RNA polymerase concentrations: P
Binding constant of polymerase: kp
Binding constant of repressor: kr
Promoter isomerization rate: ki

Becskei A, Serrano L
Nature 405, 590-593 (2000)

内容

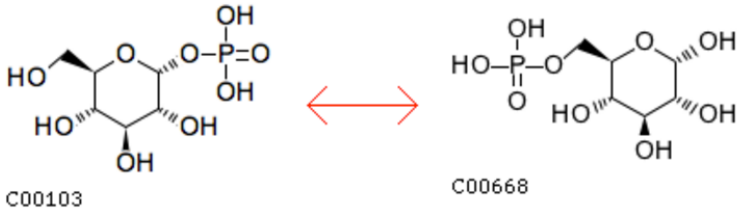
- はじめに
- 遺伝子回路設計
- 代謝経路設計

代謝反応

KEGG REACTION

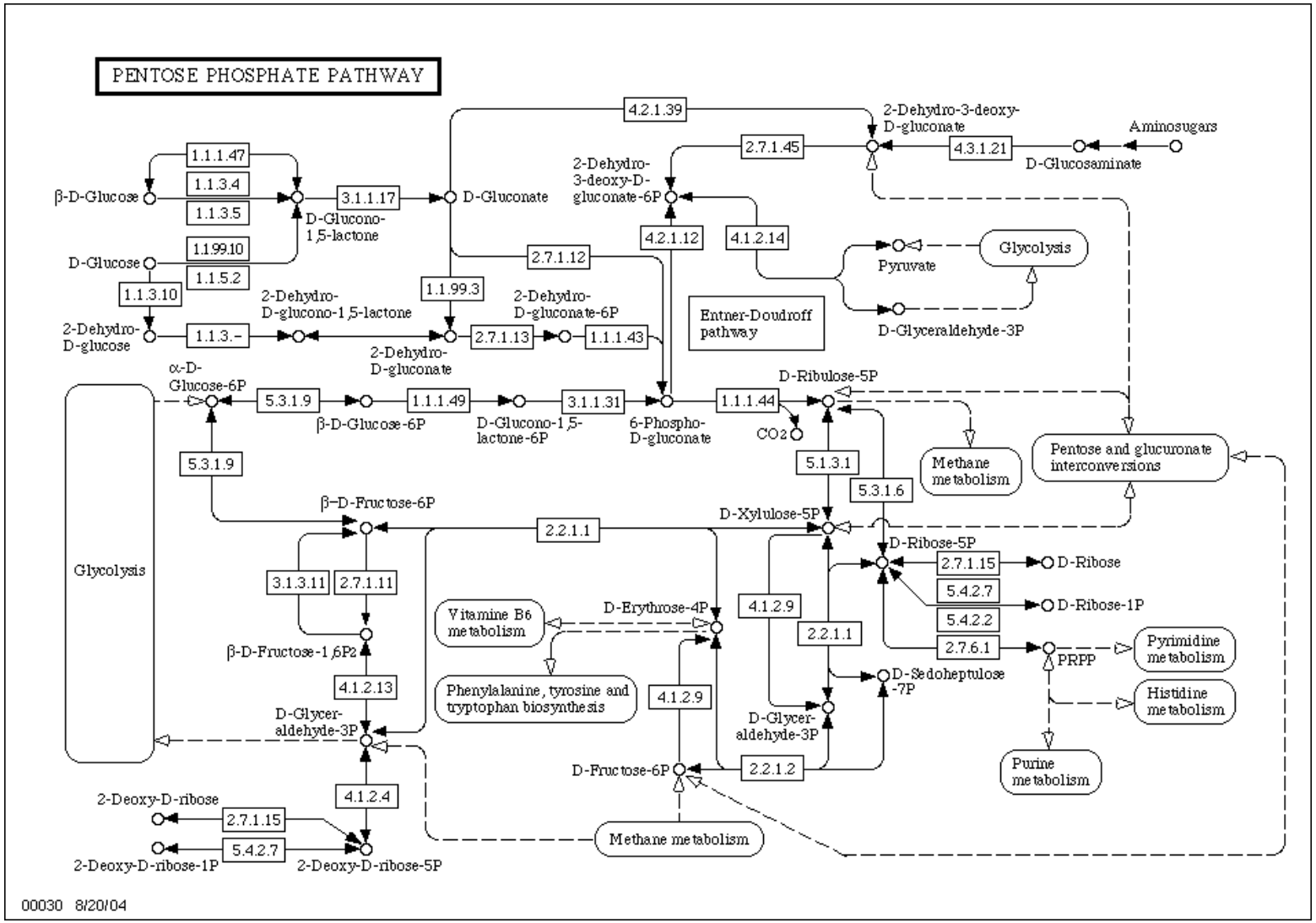
REACTION: R00959

Help

Entry	R00959 Reaction
Name	alpha-D-Glucose 1-phosphate 1,6-phosphomutase
Definition	D-Glucose 1-phosphate <=> alpha-D-Glucose 6-phosphate
Equation	C00103 <=> C00668
	 C00103 C00668
RPair	RP01197 C00103_C00668 main [RC:RC00408]
Enzyme	5.4.2.2 5.4.2.5
Pathway	rn00010 Glycolysis / Gluconeogenesis rn00052 Galactose metabolism rn00500 Starch and sucrose metabolism rn00520 Amino sugar and nucleotide sugar metabolism rn01100 Metabolic pathways rn01110 Biosynthesis of secondary metabolites rn01120 Microbial metabolism in diverse environments rn01130 Biosynthesis of antibiotics
Module	M00549 Nucleotide sugar biosynthesis, glucose => UDP-glucose
Orthology	K01835 phosphoglucomutase [EC:5.4.2.2] K15778 phosphomannomutase / phosphoglucomutase [EC:5.4.2.8 5.4.2.2] K15779 phosphoglucomutase / phosphopentomutase [EC:5.4.2.2 5.4.2.7]

代謝パスウェイ

KEGG PATHWAY

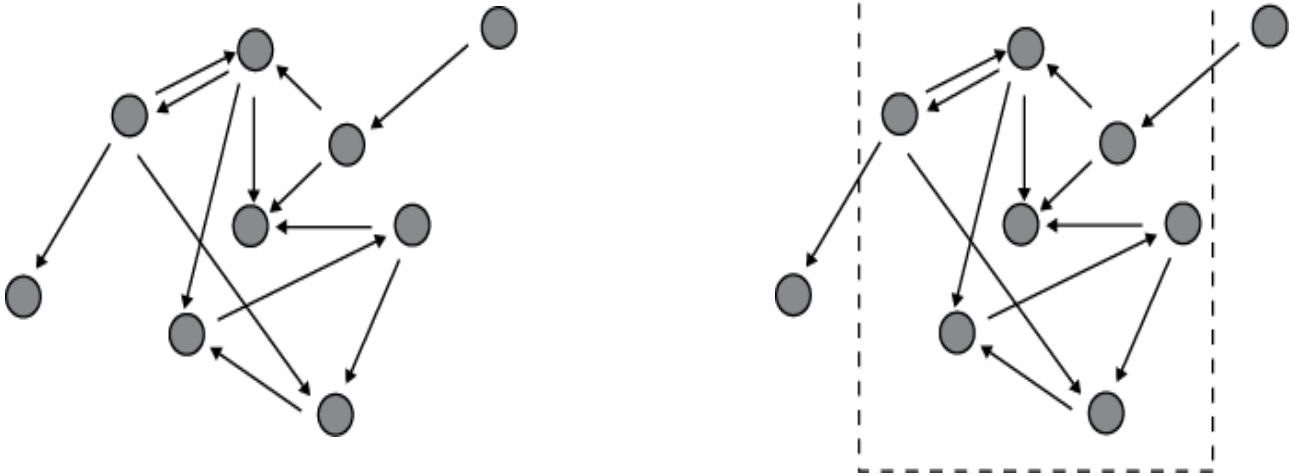


生化学反応解析

- 生化学反応系のモデル化
- 数学モデルの作成：仮定と近似
- パラメーター推定、感度解析
- 実験値と計算値の比較
- モデルの適否判定
- モデルの最適化

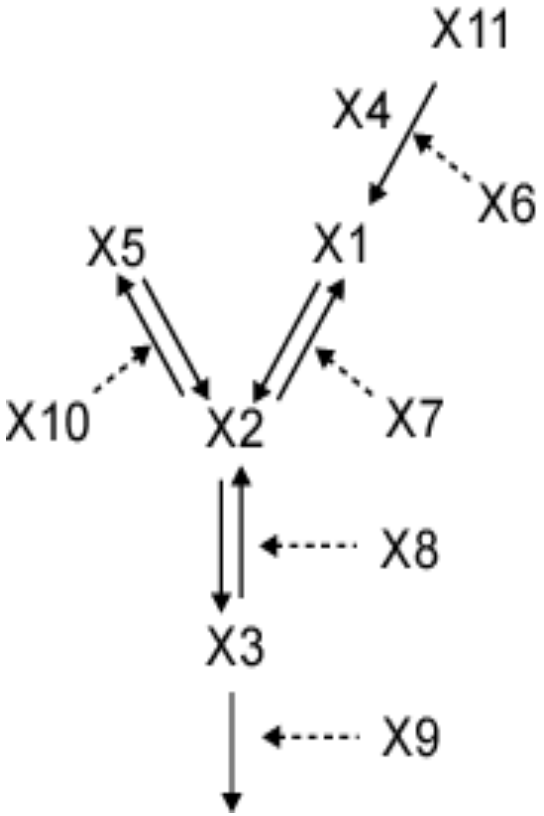
生化学反応系

- 構成物質の決定
- 相互作用の決定
- 記述方法の決定



生化学反応系

Example of proper map: Glycolysis



Dependent Variables:
X1 = Glucose-1-phosphate
X2 = Glucose-6-phosphate
X3 = Fructose-1-phosphate
X4 = Pi

Independent Variables:
X5 = Glucose
X6 = Phosphorylase
X7 = Phosphoglucomutase
X8 = Phosphoglucose isomerase
X9 = Phosphofructokinase
X10 = Glucokinase
X11 = Glycogen

ODE

数値解法

$$dy/dt = f(y,t)$$
$$y(t0) = y0$$

- 離散的な点列 $\{ti\}$ 上での正しい解 $\{y(ti)\}$ の近似解 $\{yi\}$

- 代表的な方法
  - オイラー法
  - ２次のルンゲークッタ法
  - ４次のルンゲークッタ法

  など

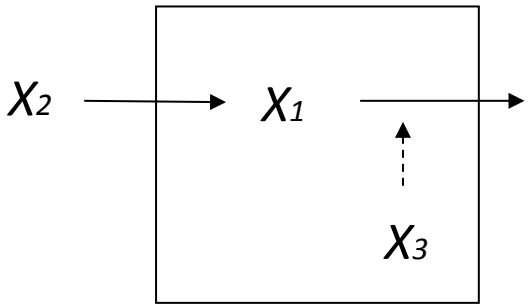
- 数値解法は様々なソフトに組み込まれている

S-system

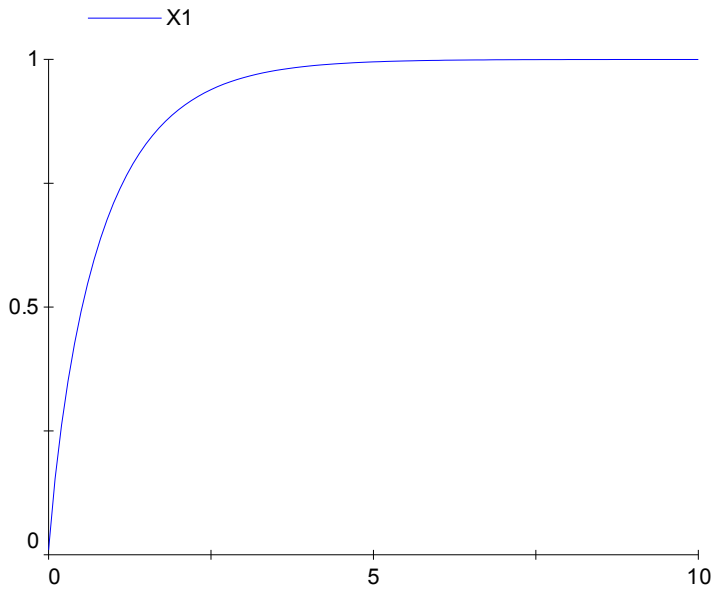
General equation

$$dX_i / dt = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n+m} X_j^{h_{ij}} \qquad for\ i = 1, 2, \ldots, n$$

Example

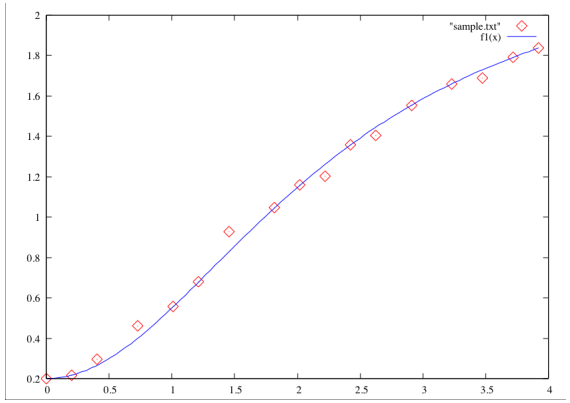


$$dX_1 / dt = \alpha_1 X_2 - \beta_1 X_1^{0.5} X_3^{-1}$$
$$dX_2 / dt = 2$$
$$dX_3 / dt = 0.5$$



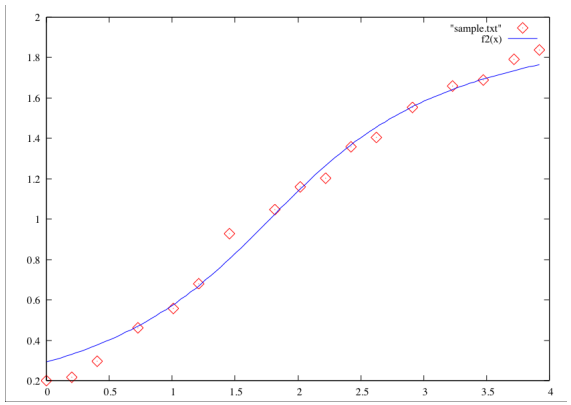
モデル選択

Hill function



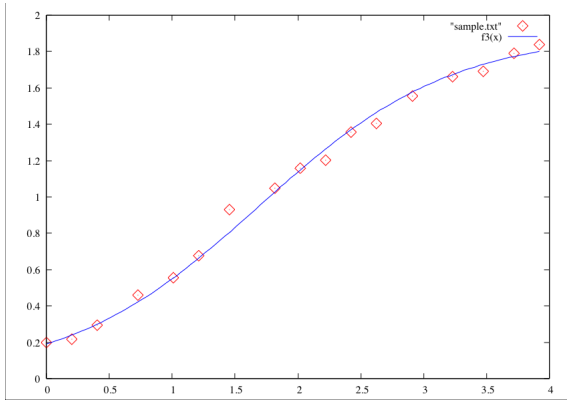
f1(x) = 2.2x^2/(2.3^2+x^2)+0.2

Arctangent function



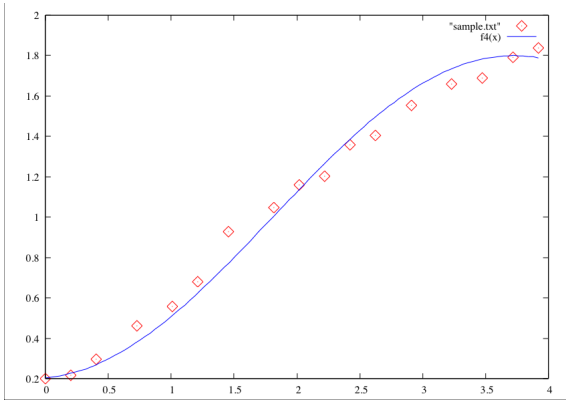
f2(x) = 0.7*arctg(0.9x-1.6)+1

Logistic function



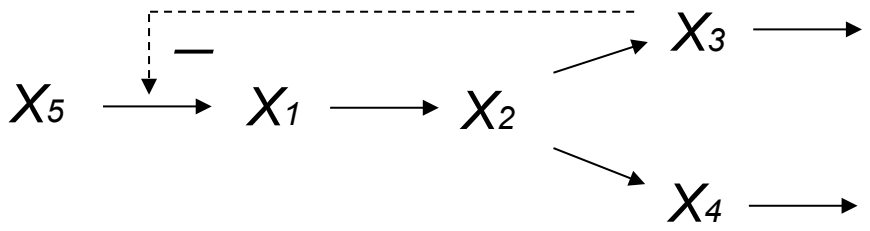
f3(x) =1.9/(1+exp(2.2-1.3x))

Sine function



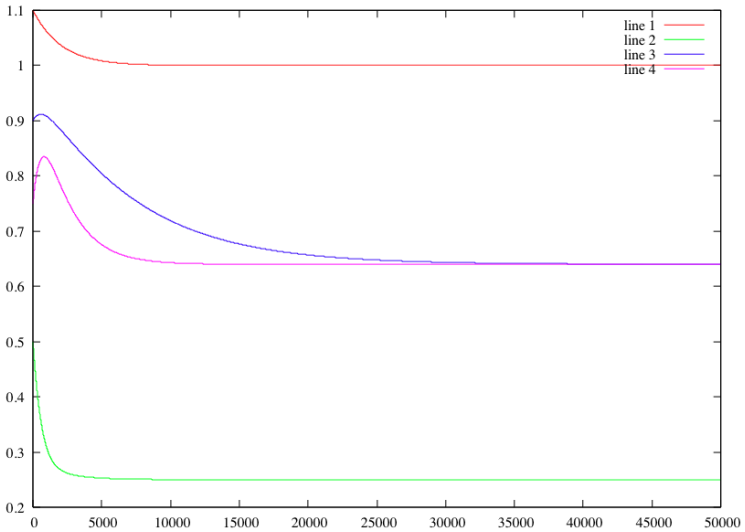
f4(x) = 0.8sin(0.82x+4.8)+1

パラメータ依存性

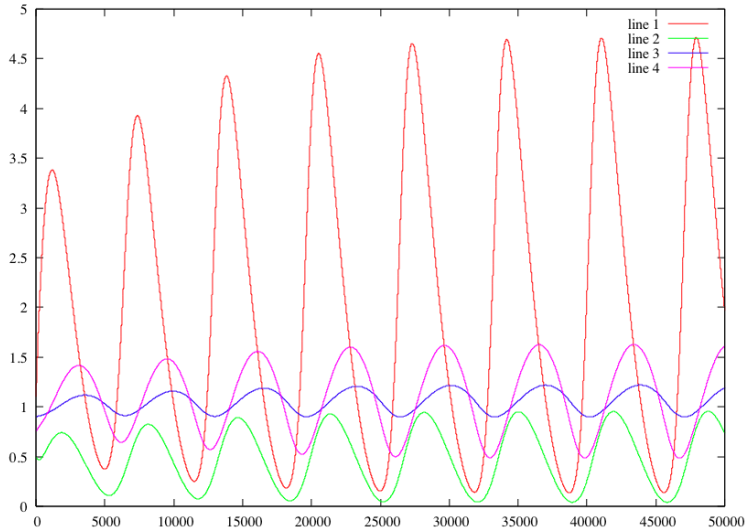


$$dX_1/dt = 10X_3^{g_{13}}X_5 - 5X_1^{0.5}$$
$$dX_2/dt = 5X_1^{0.5} - 10X_2^{0.5}$$
$$dX_3/dt = 2X_2^{0.5} - 1.25X_3^{0.5}$$
$$dX_4/dt = 8X_2^{0.5} - 5X_4^{0.5}$$
$$dX_5/dt = 0.5$$

$g_{13} = 0$

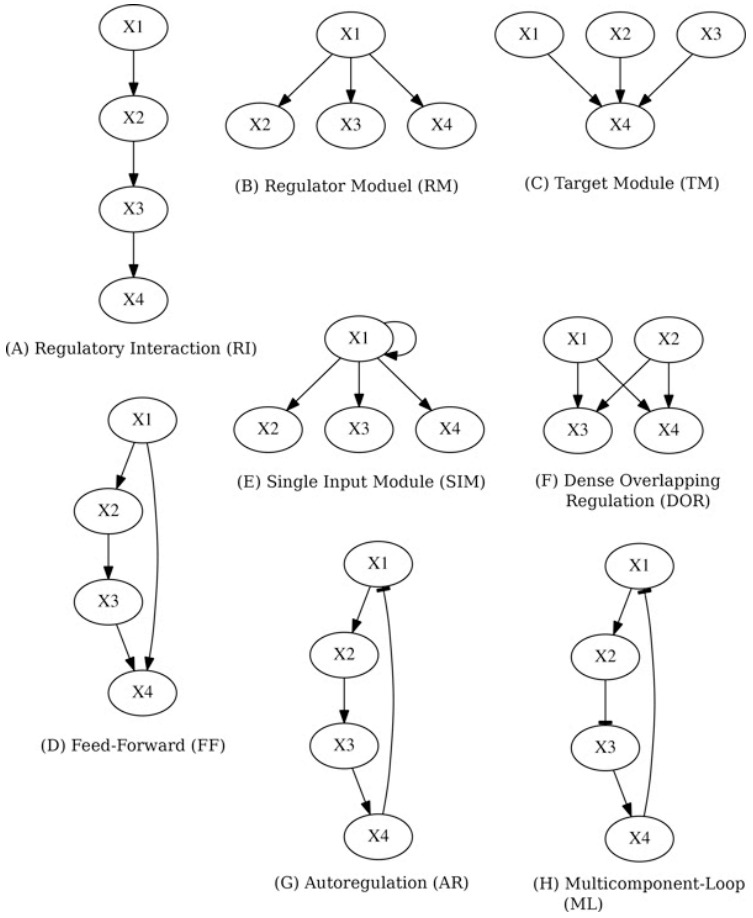


$g_{13} = -16$



パラメータ推定

S-system解析：時系列データからのパラメータ最適化・ネットワーク推定



(A) Regulatory Interaction (RI)

(B) Regulator Moduel (RM)

(C) Target Module (TM)

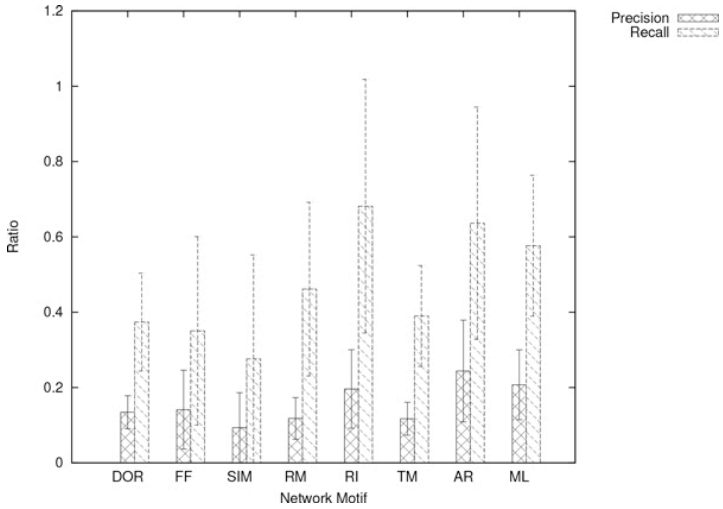
(D) Feed-Forward (FF)

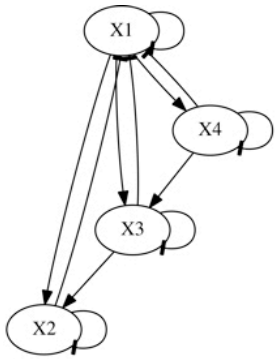
(E) Single Input Module (SIM)

(F) Dense Overlapping Regulation (DOR)

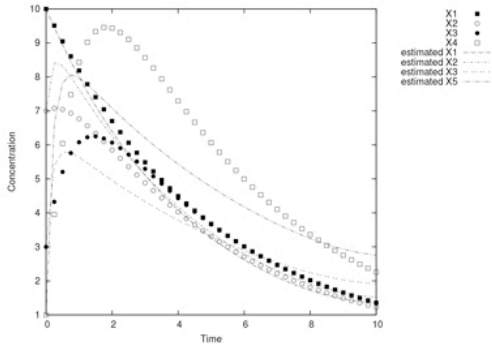
(G) Autoregulation (AR)

(H) Multicomponent-Loop (ML)





(A) Example of identified network for SIM model



(B) Example of time-series calculated from identified network.

(C) Example of estimated S-system parameter values.

7.50	-2.58	-0.87	-1.28	-1.49	0.22	1.00	0.00	0.00	0.00
8.98	1.96	-1.68	0.75	0.00	8.60	0.00	1.00	0.00	0.00
7.93	1.59	0.00	-2.90	1.28	3.74	0.00	0.00	1.00	0.00
9.59	0.94	0.00	0.00	-0.60	2.56	0.00	0.00	0.00	1.00

Nakatsui M, Araki M, Kondo A
BMC Systems Biology 7(Suppl 6):S4

反応記述言語

Systems Biology Markup Language (SBML)

XML形式で生化学反応を記述

Compartment
Species
Reaction
Parameter
Unit Definition
Rule

$$k = \frac{k_3}{k_2}, \quad s_2 = \frac{k\,x}{1 + k_2}, \quad A = 0.10\,x$$

```
<model>
    ...
    <listOfRules>
        <assignmentRule variable="k">
            <notes>
                <xhtml:p>
                    k = k3/k2
                </xhtml:p>
            </notes>
            <math xmlns="http://www.w3.org/1998/Math/MathML">
                <apply>
                    <divide/>
                    <ci> k3 </ci>
                    <ci> k2 </ci>
                </apply>
            </math>
        </assignmentRule>
        <assignmentRule variable="s2">
            <notes>
                <xhtml:p>
                    s2 = (k * x)/(1 + k2)
                </xhtml:p>
            </notes>
            <math xmlns="http://www.w3.org/1998/Math/MathML">
                <apply>
                    <divide/>
                    <apply>
                        <times/>
                        <ci> k </ci>
                        <ci> x </ci>
                    </apply>
                    <apply>
                        <plus/>
                        <cn> 1 </cn>
                        <ci> k2 </ci>
                    </apply>
                </apply>
```

http://sbml.org

シミュレーション方法

Process	Phenomena	Models
Metabolism	Enzymatic reaction	ODE, S-systems, FBA
Signal transduction	Molecular binding Enzymatic reaction	DAE Stochastic algorithms
Gene expression	Molecular binding Polymerization Degradation	DAE, S-systems Boolean networks Stochastic algorithms

シミュレータ

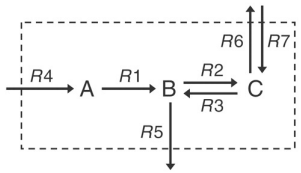
	Capabilities					Frameworks							API	Dep.	Platforms	SBML		Availabil.		
	Creation	Simulation	Analysis	Database	Utility	ODE	DAE	PDE	Stochastic	Events	Logical	Other				Import	Export	Open source	Academic use	Commercial use
acslXtreme	•														W	•			$	$
ALC	•					•	•		•			•			L, W, M, B		•	•	F	F
Asmparts	•				•	•									L,W	•	•	•	F	F
Antimony	•				•								C, C++		L, W, M	•	•	•	F	F
AutoSBW			•			•							SBW	SBW	L, W, M	•	•	•	F	F
AVIS												•		various	L	•		•	F	F
BALSA	•													Sigtran						
BASIS	•	•		•					•	•			WS		B	•	•	•	F	F
BetaWB	•	•	•						•	•					L,W,M		•		F	F
BiGG				•											B		•		F	
BiNoM	•		•		•							•			L, W, M	•	•	•	F	F
BiNoM Cytoscape Plugin	•		•		•							•		Cytoscape	L, W, M	•	•	•	F	F
BIOCHAM		•			•	•									L,W,M		•	•	F	F
BioCharon	•	•	•		•	•								CHARON						
Biological Networks	•		•		•										L,W,M	•	•		F	$
BioCyc				•													•		F	$
BioGrid																				

http://sbml.org

化学量論的解析

フラックスバランス解析（FBA）

I. Reaction network formalism



Chemical reactions

Internal	Exchange
*R*1: −1 A → 1 B	*R*4: 1 A
*R*2: −1 B → 1 C	*R*5: −1 B
*R*3: −1 C → 1 B	*R*6: −1 C
	*R*7: 1 C

S =

	*R*1	*R*2	*R*3	*R*4	*R*5	*R*6	*R*7
A	−1	0	0	1	0	0	0
B	1	−1	1	0	−1	0	0
C	0	1	−1	0	0	−1	1

II. FBA formulation

Dynamic mass balance

$$\frac{dC}{dt} = \mathbf{S}v$$

C : Concentration
t : Time
$\mathbf{S}$: Stoichiometric matrix
v : Flux vector

Steady-state assumption

$\mathbf{S}v = 0$

LP formulation

Objective: max $Z = v_5$

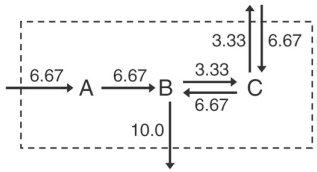
Constraints:

$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{bmatrix} R1 & R2 & R3 & R4 & R5 & R6 & R7 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ \vdots \\ v_7 \end{bmatrix} = \mathbf{0} \qquad 0 \le v_1,\ldots,v_7 \le 10$$

III. Hypothetical flux distribution at steady-state

$Z = 10$

$v = [6.67\ 3.33\ 6.67\ 6.6710.0\ 3.33\ 6.67]^T$



反応マトリックス記述

定常状態仮定
定式化

線形計画法

Lee JM, Gianchandani EP, Papin JA
Brief Bioinformatics 7, 140-150 (2006)

ゲノムスケールモデル

SBMLによる記述

```
<reaction id="R_ACKr" name="acetate kinase">
 <notes>
  <html xmlns="http://www.w3.org/1999/xhtml">
  <p>GENE_ASSOCIATION: (b3115 or b2296 or b1849)</p>
  <p>GENE_LIST: b1849 b2296 b3115</p>
   <p>SUBSYSTEM: Pyruvate Metabolism</p></html>
 </notes>
 <listOfReactants>
  <speciesReference species="M_ac_c"/>
  <speciesReference species="M_atp_c"/>
 </listOfReactants>
 <listOfProducts>
  <speciesReference species="M_actp_c"/>
  <speciesReference species="M_adp_c"/>
 </listOfProducts>
 <kineticLaw>
  <math xmlns="http://www.w3.org/1998/Math/MathML">
   <ci> FLUX_VALUE </ci>
  </math>
  <listOfParameters>
   <parameter id="LOWER_BOUND" value="-1000" units="mmol_per_gDW_per_hr"/>
   <parameter id="UPPER_BOUND" value="1000" units="mmol_per_gDW_per_hr"/>
   <parameter id="OBJECTIVE_COEFFICIENT" value="0"/>
   <parameter id="FLUX_VALUE" value="0" units="mmol_per_gDW_per_hr"/>
  </listOfParameters>
 </kineticLaw>
</reaction>
```

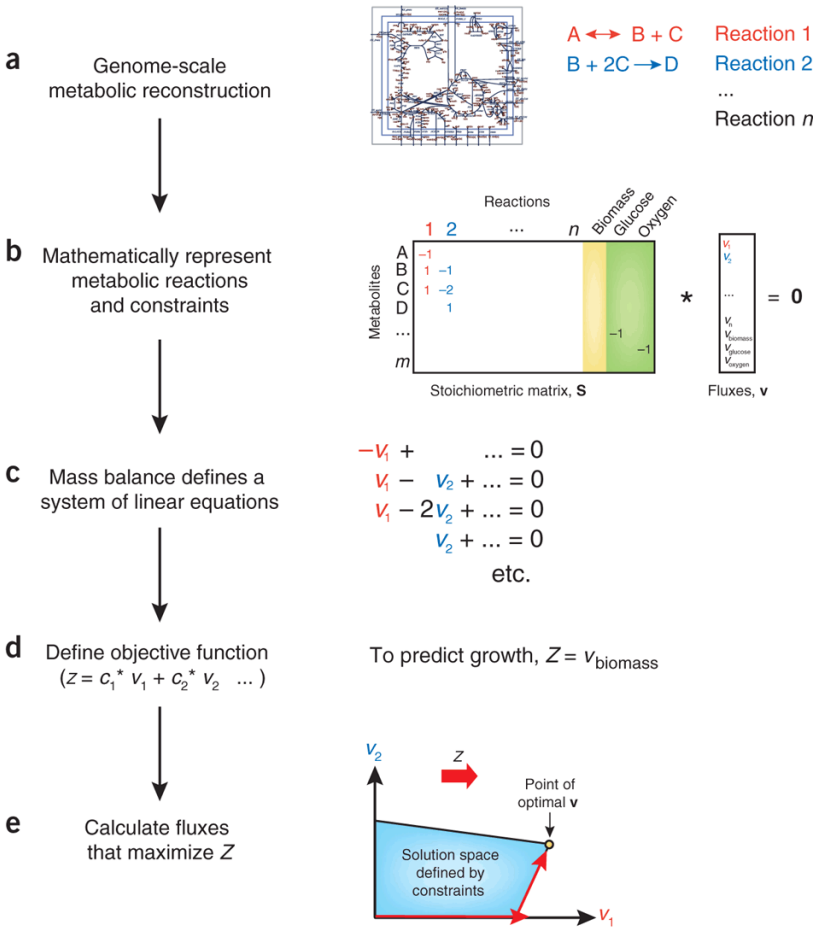
反応（酵素名称）

遺伝子ID

反応種（化合物）

計算パラメータ

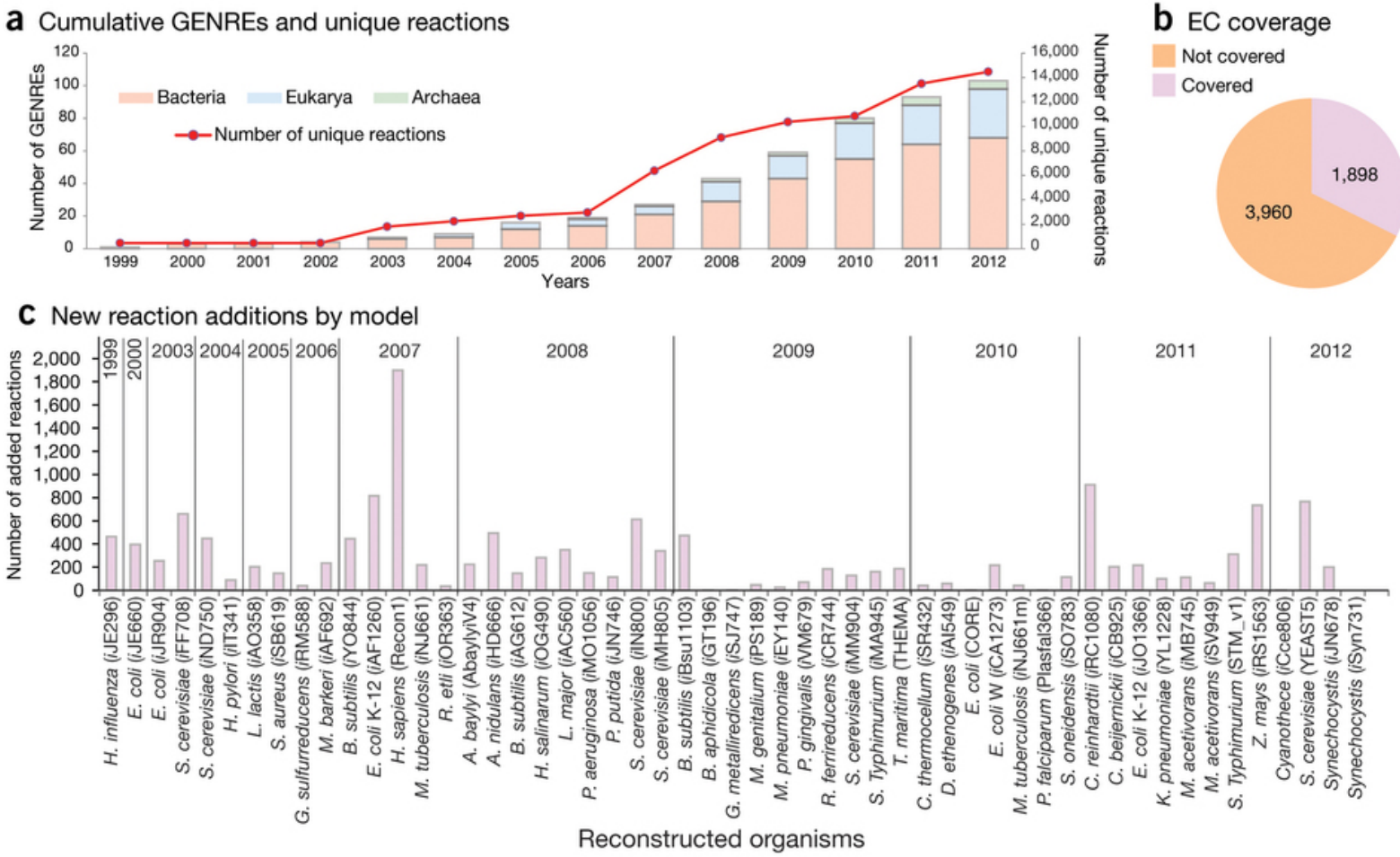
ゲノムスケールモデル



Oath JD, Thiele I, Palsson BO
Nature Biotechnology 28, 245-248 (2010)

ゲノムスケールモデル

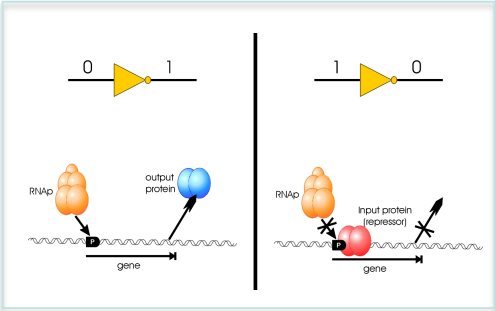
Genome-scale network reconstruction (GENRE)



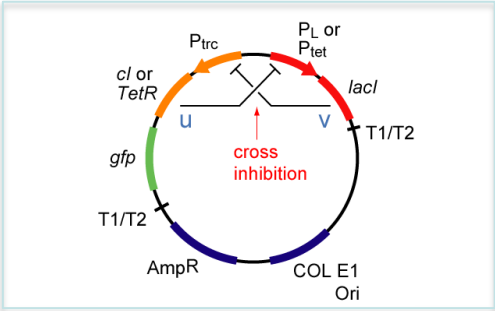
Monk J, Nogales J, Palsson BO
Nature Biotechnology 32, 447-452(2014)

代謝経路設計：化学情報解析

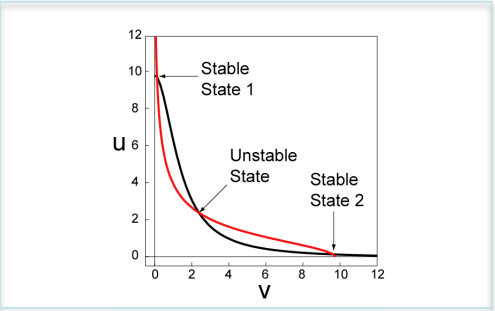
単位/部品　　組合せ/再構築　　表現型/構造



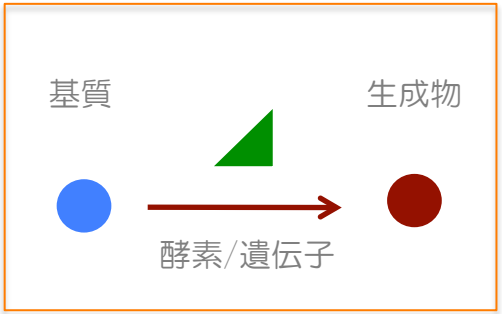
遺伝子/プロモータ



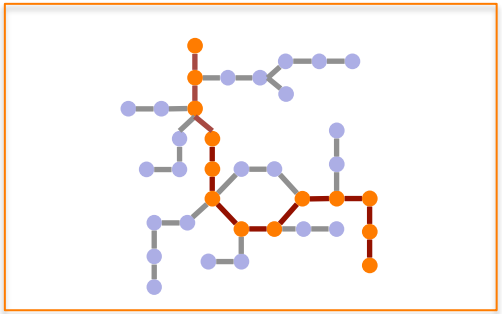
遺伝子回路



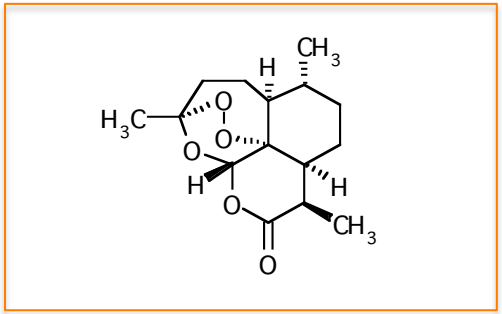
スイッチ



反応/酵素

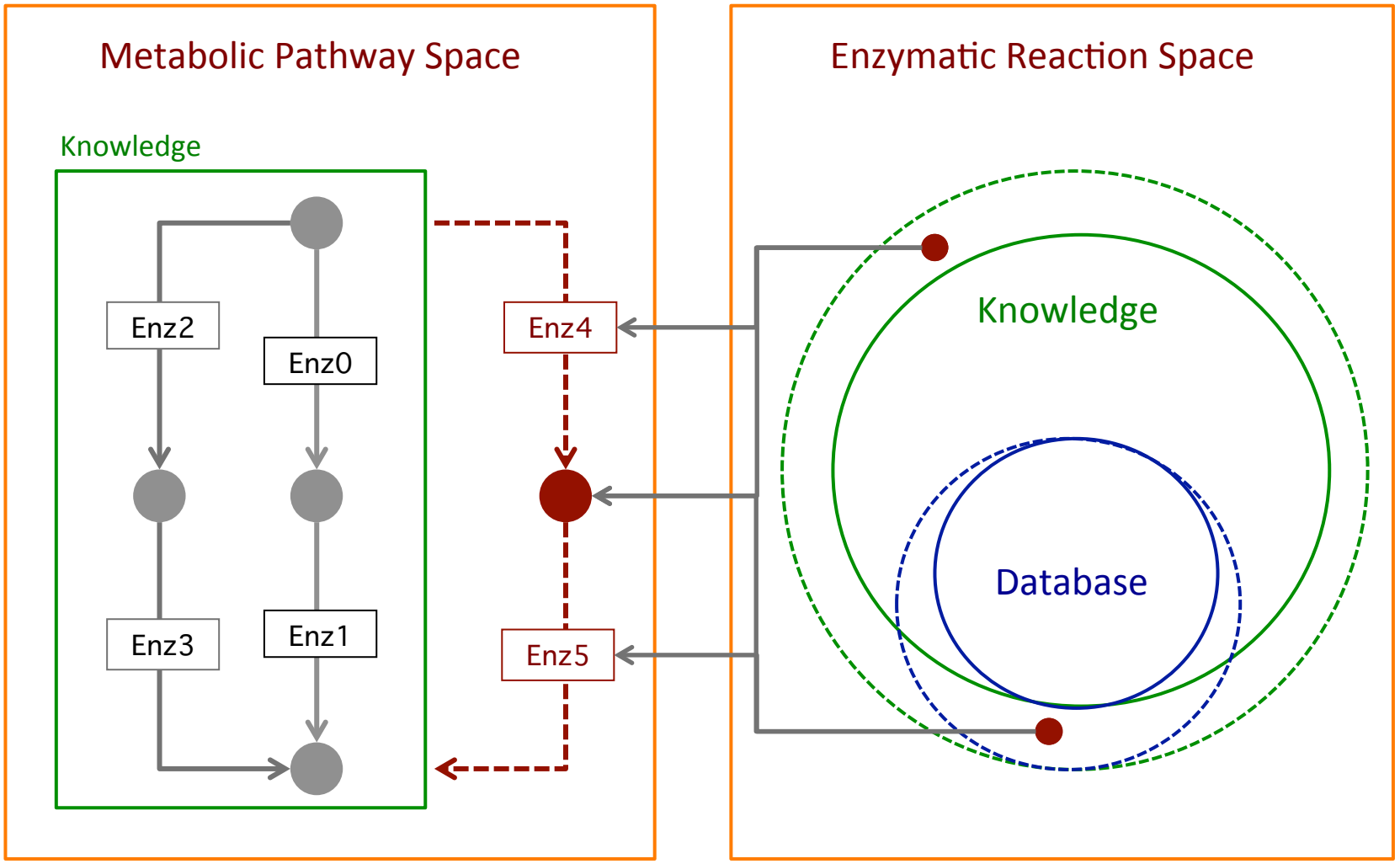


代謝経路



代謝化合物

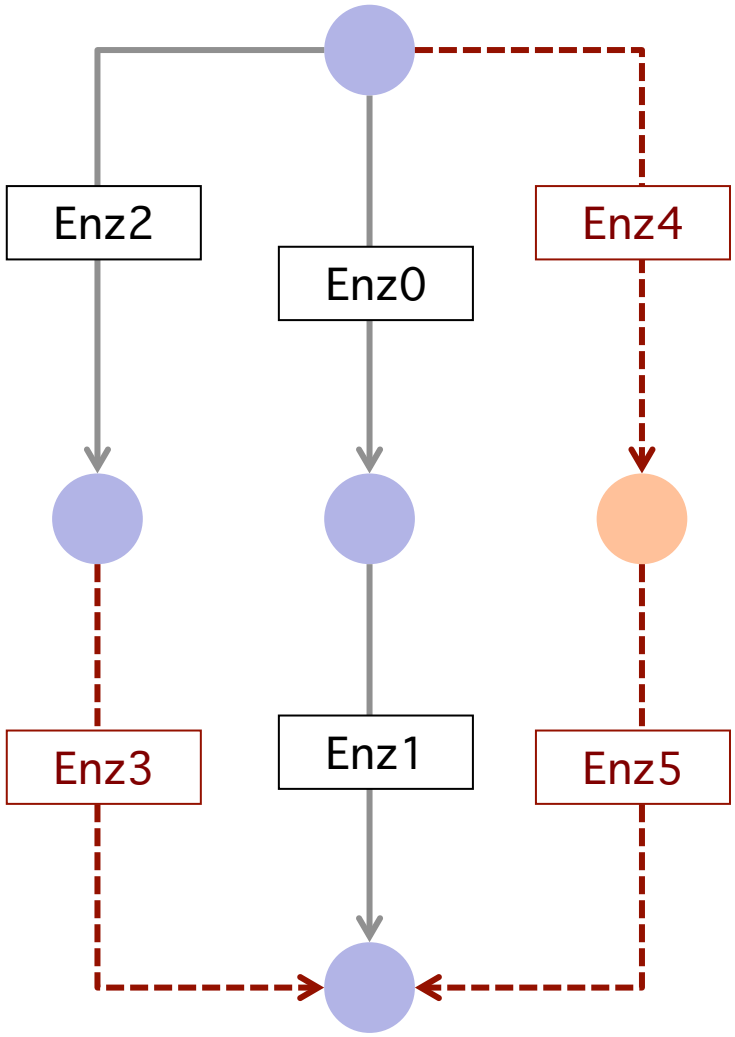
代謝パスウェイの知識拡張



代謝反応データベース

Database	Entries	Description
KEGG COMPOUND	17,091	Compounds
PubChem	~50,000,000	Compounds
KEGG REACTION	9,400	Reactions
KEGG RPAIR	14,218	Reactant Pairs
KEGG ENZYME	6,118	EC Numbers
BRENDA	6,130	Different Enzymes

代謝経路設計



Known:

Unknown:

PATH_0 (Enz0_Enz1): Known

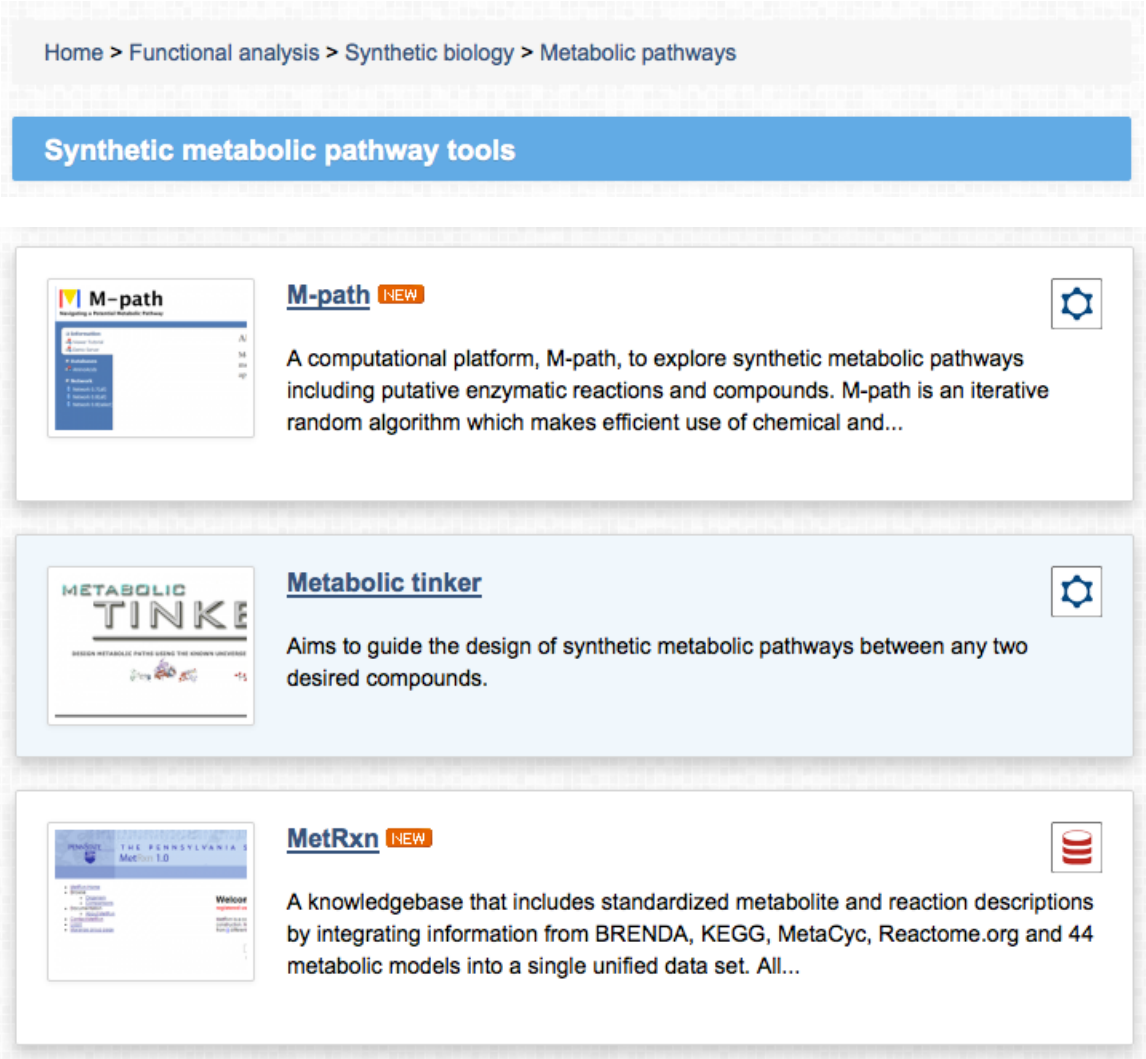
PATH_1 (Enz2_Enz3): Unknown Ez

PATH_2 (Enz4_Enz5): Unknown Ez, Co

設計方法

omictools.com/

設計方法


Home > Functional analysis > Synthetic biology > Metabolic pathways
Synthetic metabolic pathway tools
M-path
M-path NEW
A computational platform, M-path, to explore synthetic metabolic pathways including putative enzymatic reactions and compounds. M-path is an iterative random algorithm which makes efficient use of chemical and...
METABOLIC TINKE
Metabolic tinker
Aims to guide the design of synthetic metabolic pathways between any two desired compounds.
MetRxn NEW
A knowledgebase that includes standardized metabolite and reaction descriptions by integrating information from BRENDA, KEGG, MetaCyc, Reactome.org and 44 metabolic models into a single unified data set. All...

設計方法

Related sites M-path

PathPred
A web-based server to predict plausible enzyme-catalyzed reaction pathways from a query compound...

Desharky
A Monte Carlo algorithm that finds a metabolic pathway from a target compound by exploring a...

FMM
It can reconstruct metabolic pathways form one metabolite to another metabolite among different...

MetRxn
A knowledgebase that includes standardized metabolite and reaction descriptions by integrating...

MRSD
A tool to search and design routes based on the weighted compound transform diagraph.

SobolHDMR
A general purpose metamodeling software.

FindPath
An unified system predicting and ranking the possible pathways according to their metabolic...

XTMS
A web-based pathway analysis platform which provides full access to the set of pathways that can...

Metabolic tinker
Aims to guide the design of synthetic metabolic pathways between any two desired compounds.

設計方法

Network (knowledge) based approaches

- Handorf *et al.*, 2005; Noor *et al.*, 2010; Chou *et al.*, 2009; McClymont and Soyer, 2013; Kumar *et al.*, 2012; Xia *et al.*, 2011; Rodrigo *et al.*, 2008

The methods are effective in finding known heterologous enzymatic reactions, but it ignores any pathway with an unknown reaction step.

Reaction (chemical) rule based approaches

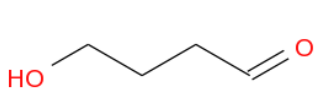
- Yim *et al.*, 2011; Hatzimanikatis *et al.*, 2005; Carbonell *et al.*, 2012, 2014; Henry *et al.*, 2010; Cho *et al.*, 2010
- Araki *et al.*, 2015

These methods have been limited to a small number of either reactions or compounds to avoid combinatorial explosion in reconstructing synthetic metabolic pathways.

問題設定

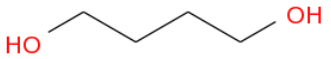
Reaction

4-Hydoroxybutanal



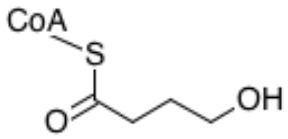
???

1, 4-BDO



Compound

4HB-CoA

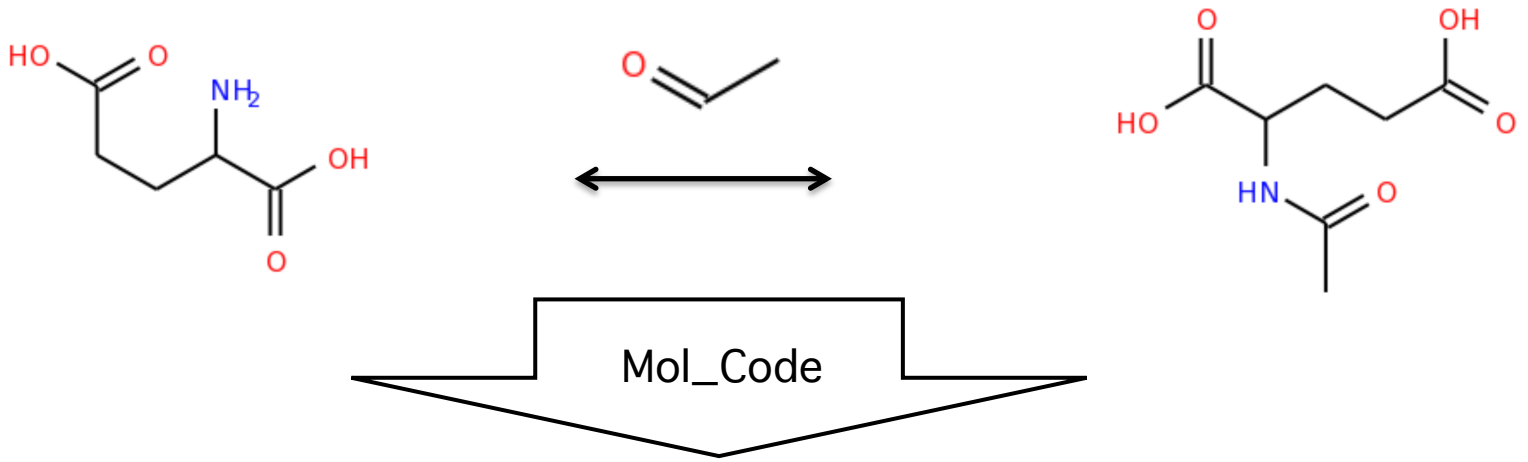


- CoA

???

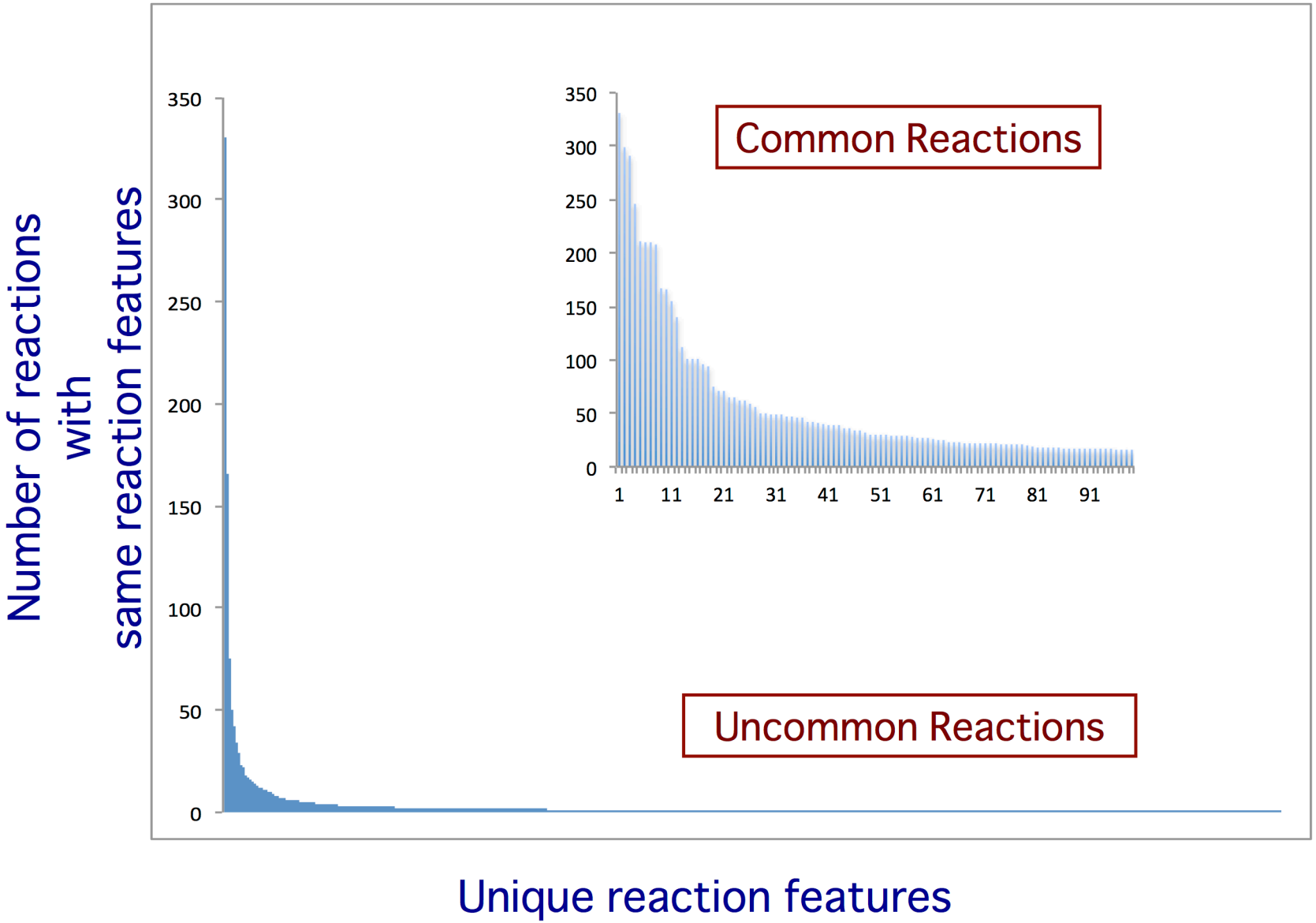
化学・反応表現

① Glutamate　② Acetyl group　③ N-acetyl glutamate

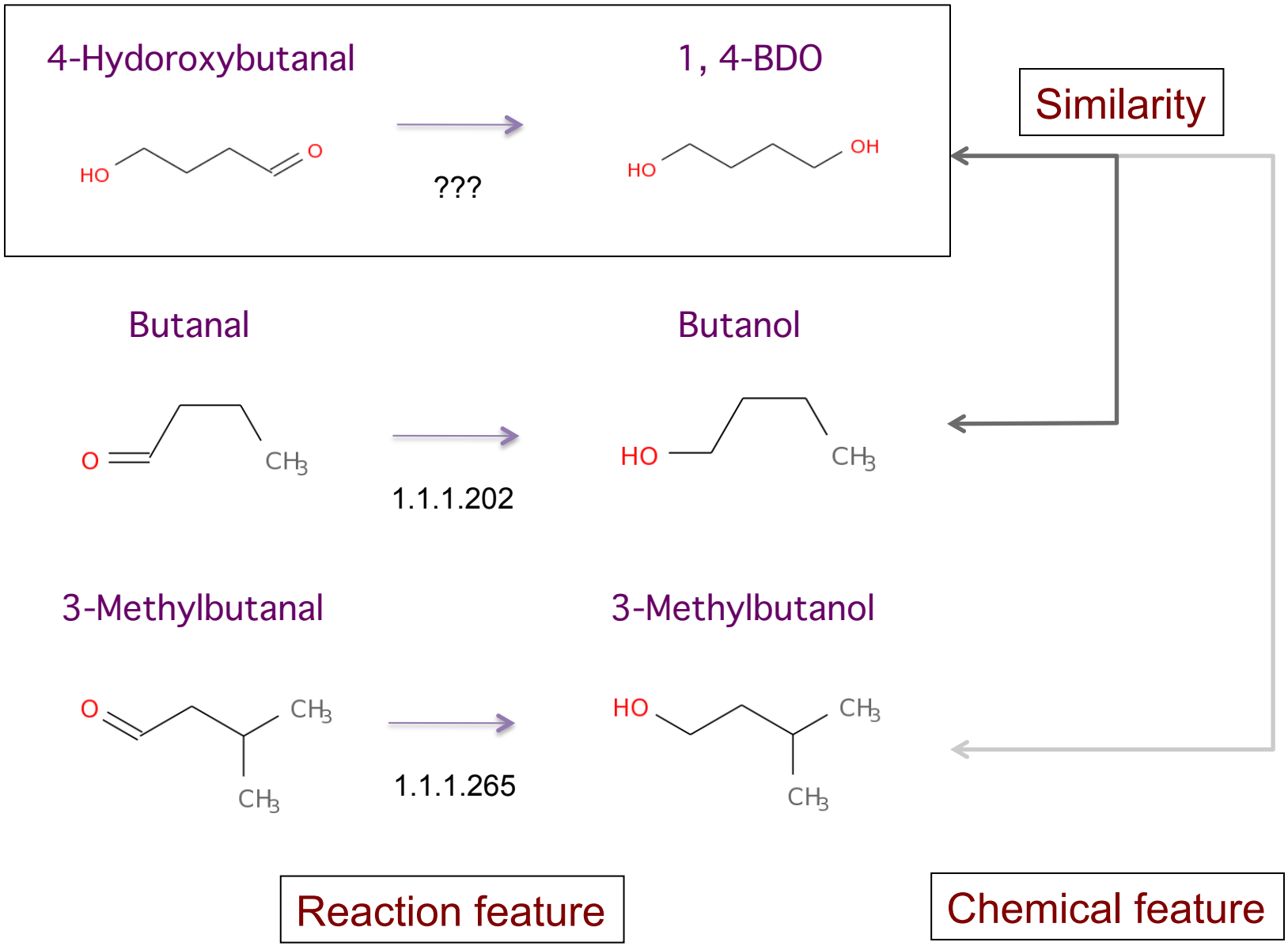


Glutamate	Acetylation	N-acetyl glutamate
(2 3 0 0 ・・・・・ 1 2 0)	(1 1 0 0 ・・・・・0 0 1) ⟷ (-1 -1 0 0 ・・・・・0 0 -1)	(3 4 0 0 ・・・・・1 2 1)
Feature vector (Chemical feature)	Deacetylation Feature vector differences (Reaction feature)	Feature vector (Chemical feature)

反応バリエーション

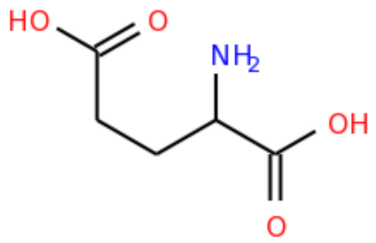

Number of reactions with same reaction features
350
300
250
200
150
100
50
0
Common Reactions
350
300
250
200
150
100
50
0
1
11
21
31
41
51
61
71
81
91
Uncommon Reactions
Unique reaction features

反応推定


4-Hydoroxybutanal
1, 4-BDO
HO
O
???
HO
OH
Similarity
Butanal
Butanol
O
CH3
1.1.1.202
HO
CH3
3-Methylbutanal
3-Methylbutanol
O
CH3
CH3
1.1.1.265
HO
CH3
CH3
Reaction feature
Chemical feature

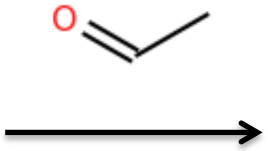
化合物推定

① Glutamate



(2 3 0 0 ・・・・・・ 1 2 0)

② Acetyl group

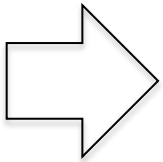


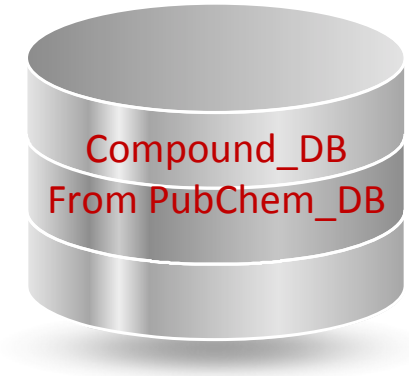
(1 1 0 0 ・・・・・・0 0 1)

???

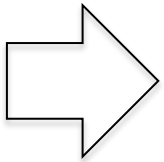
(3 4 0 0 ・・・・・・1 2 1)

Structure Search

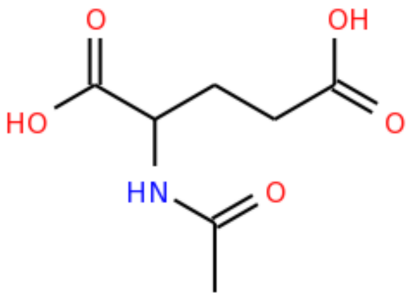




Assignment

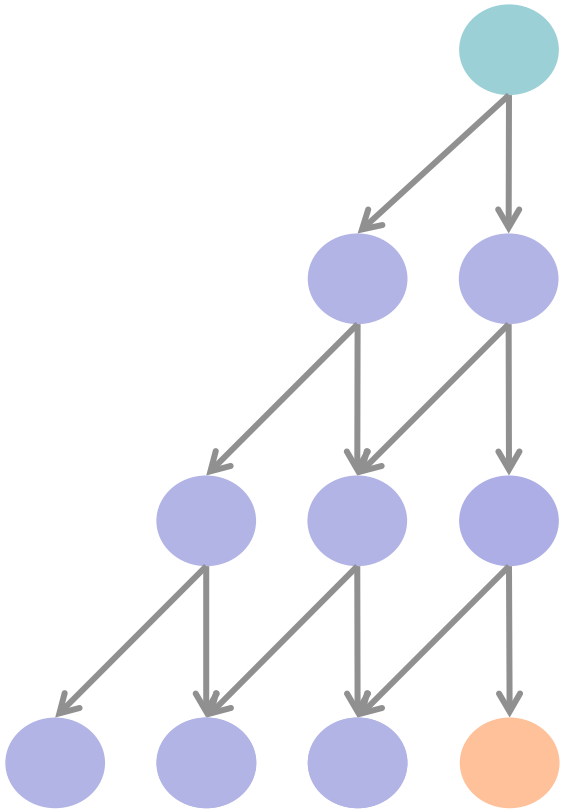


③ N-acetyl glutamate



設計方法

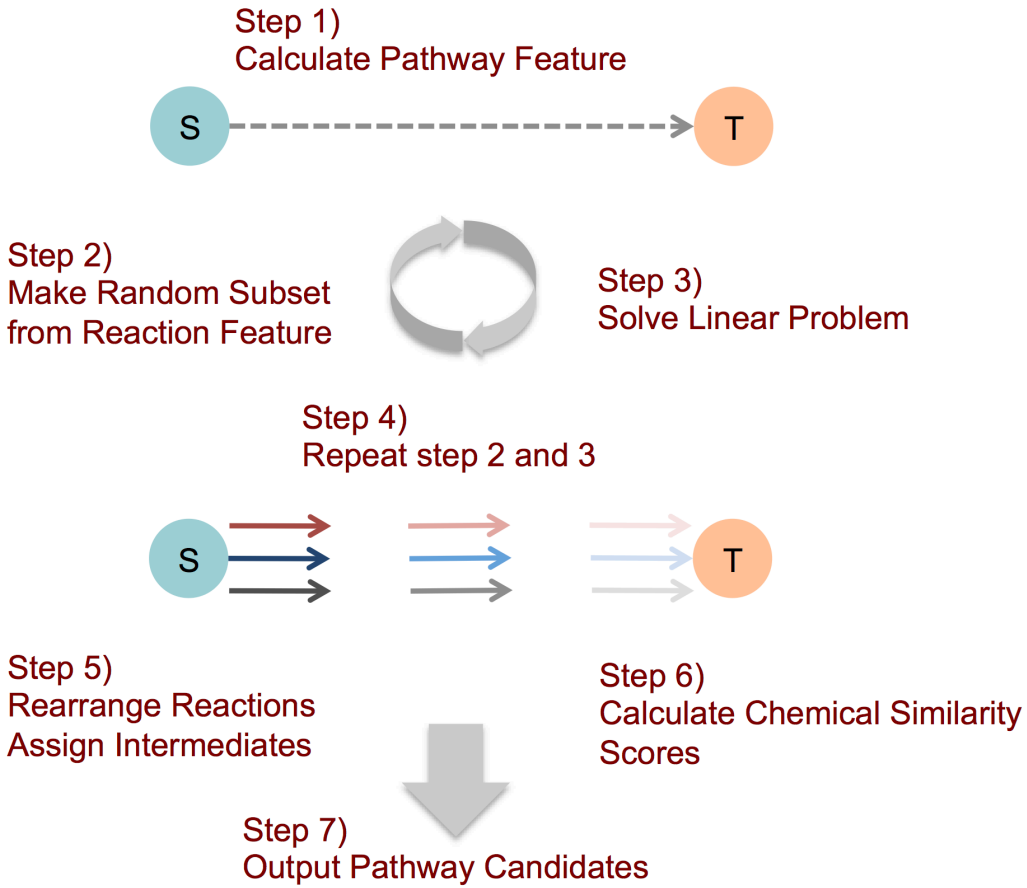
Brute-force method



Incorrect	Correct

Reaction	Compound
1,000	
	1,000
1,000	
	1,000,000
1,000	
	1,000,000,000

設計方法



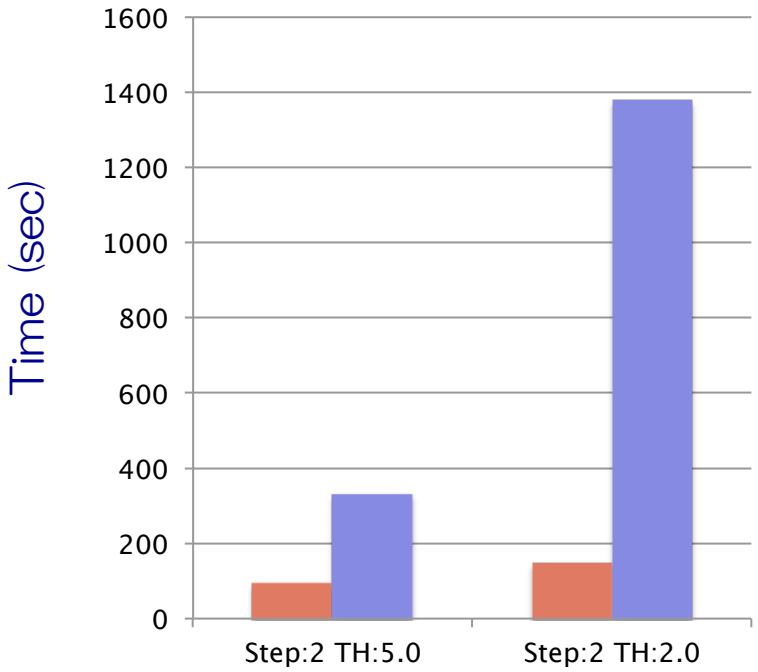
Araki M *et al.* Bioinformatics 31, 905-911 (2015)

計算時間

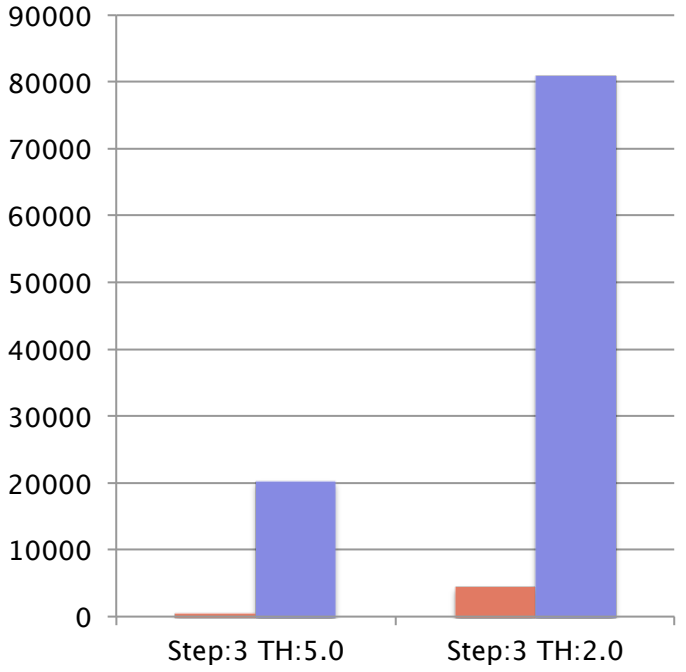
Example) Glucose ~ Succinate

Proposed（■） vs Sequential（■）

Reaction Data
- 1553 (TH: 2)
- 599 (TH: 5)



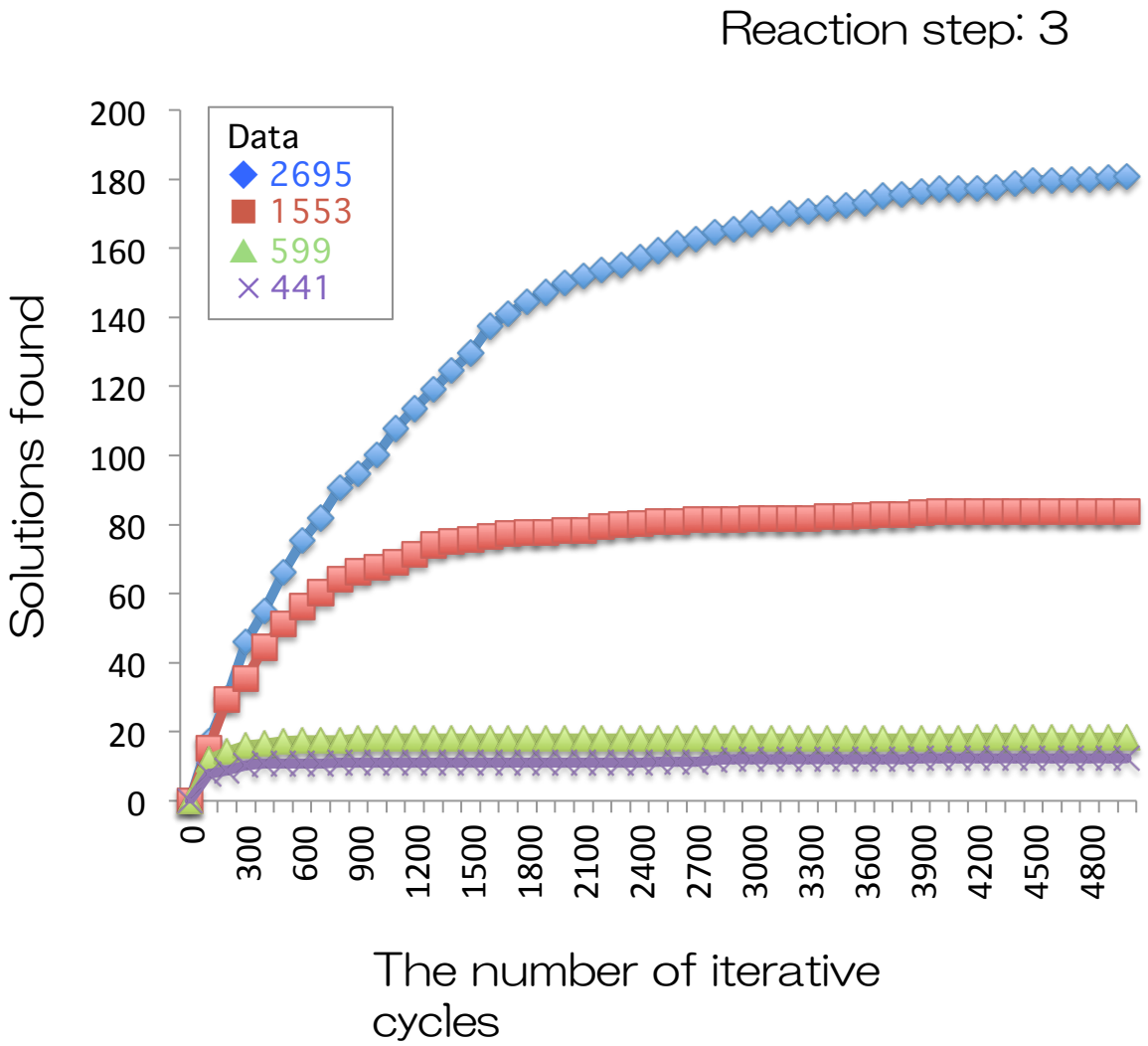
Reaction Steps：2



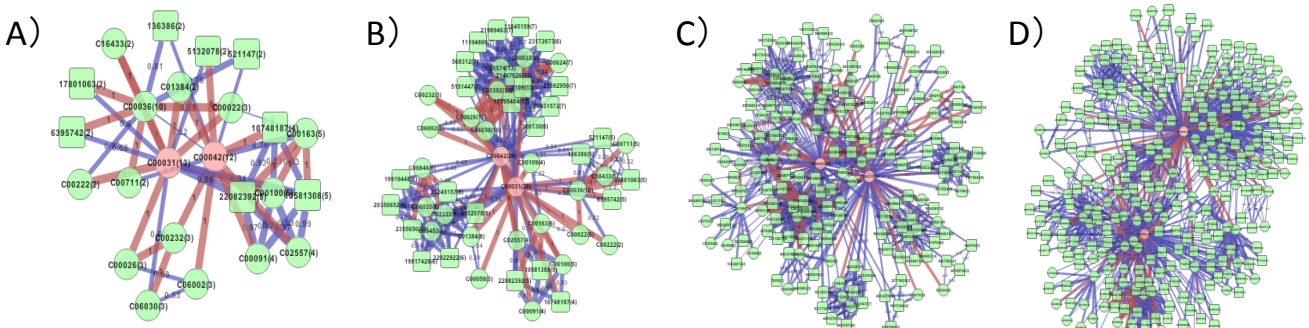
Reaction Steps：3

Iterative Enumeration

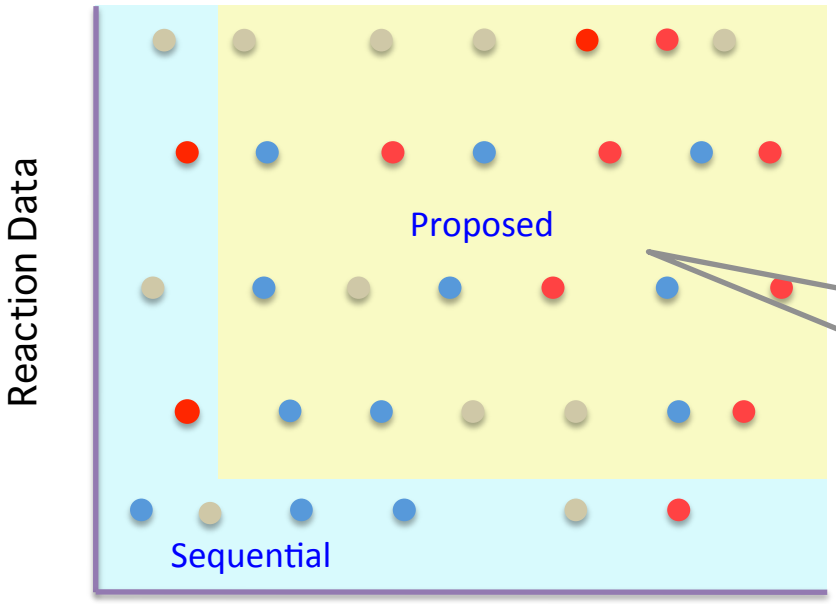
Example) Glucose ~ Succinate



Advantage



Reaction Data: A < B < C < D



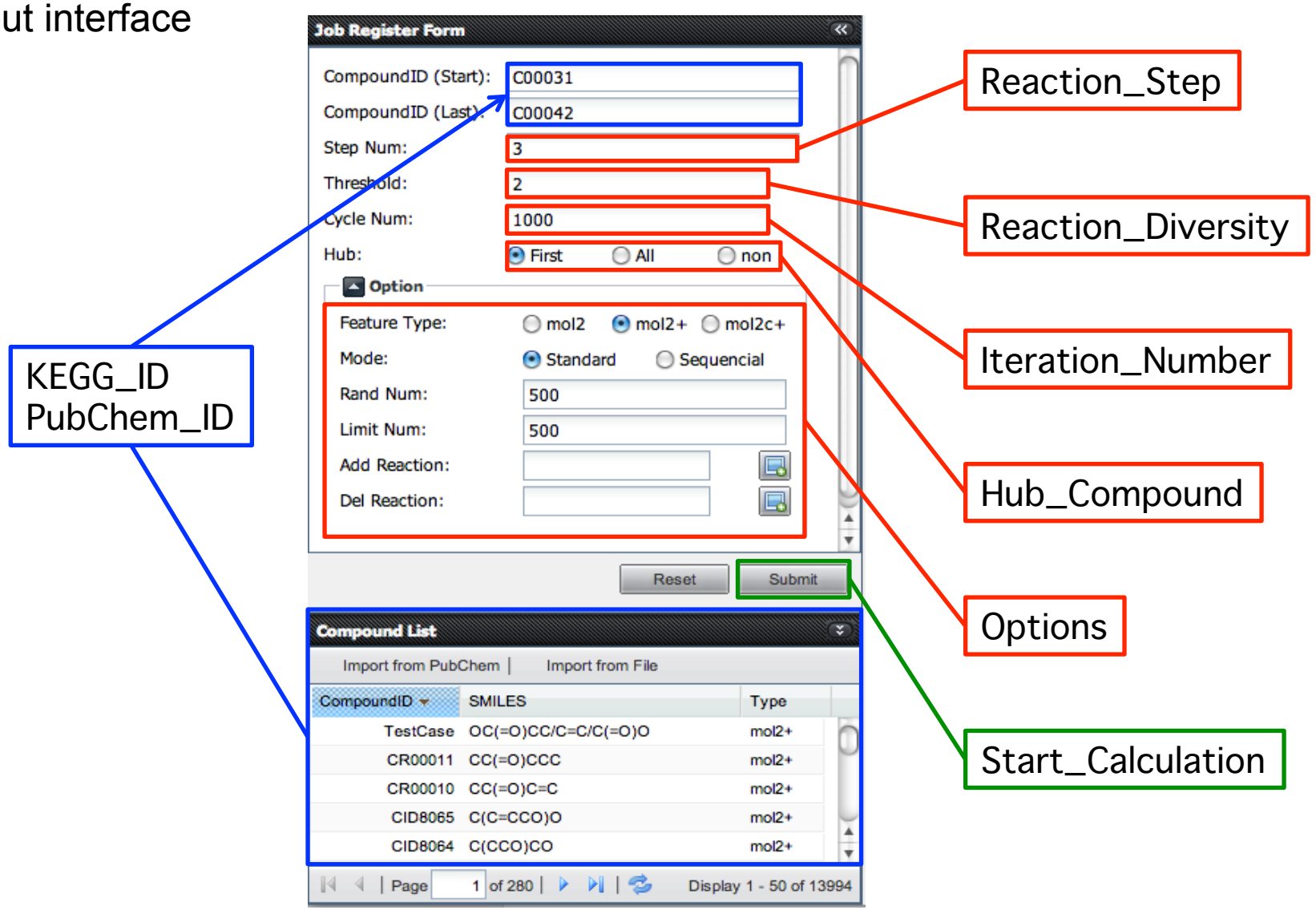
Increasing network diversity with putative compounds and reactions

Expanding the diversity of chemical and reaction information available for finding novel metabolic pathways

M-path

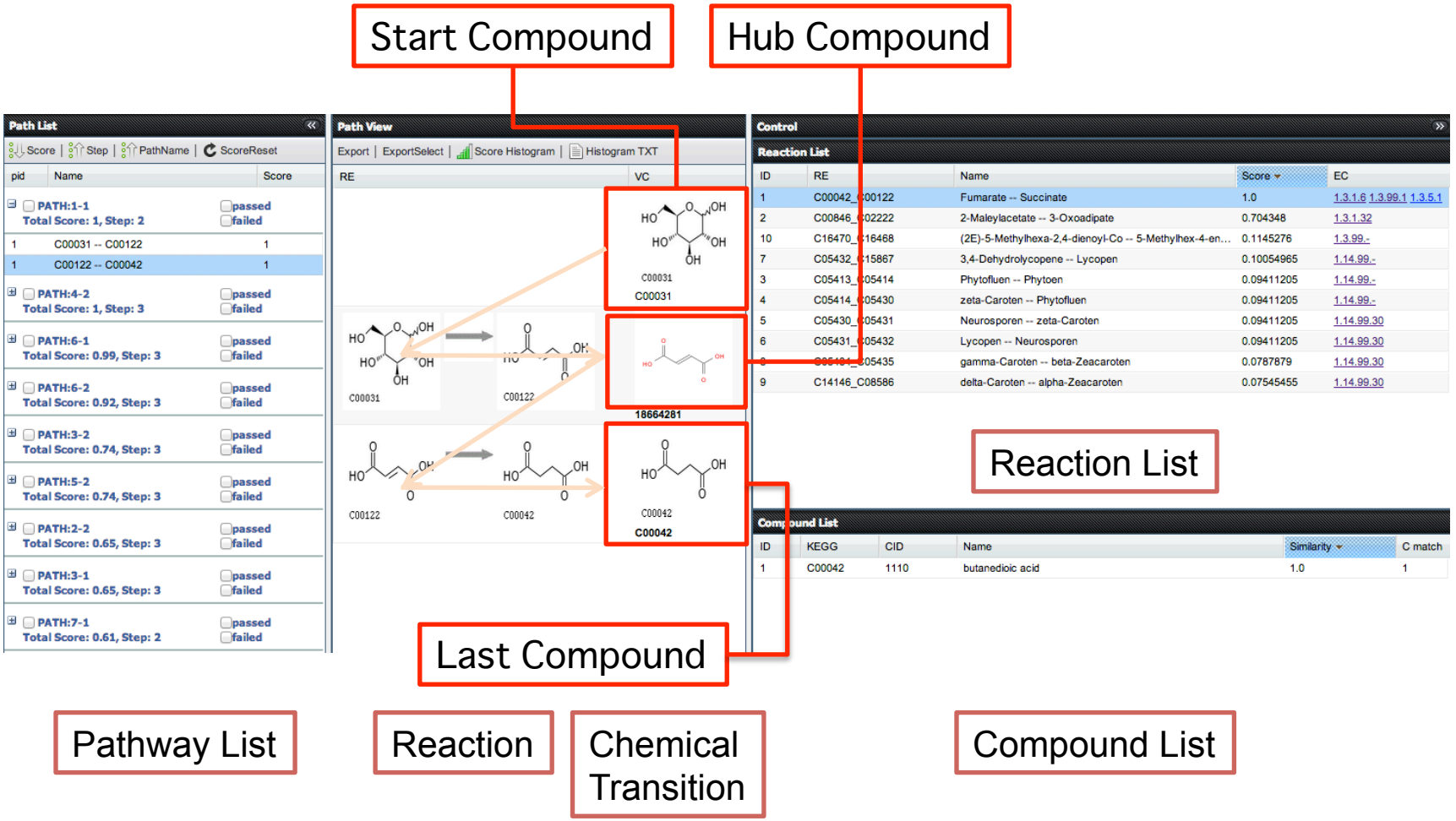
http://bp.scitec.kobe-u.ac.jp/m-path/server/

Input interface



M-path

Example) Glucose ~ Succinate



M-path

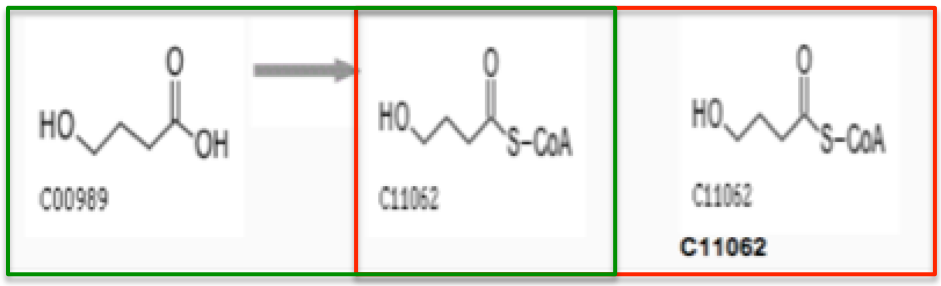
Enzymatic Reactions

EC_number

Reaction

ID	RE	Name	Score	EC
40	C11062_C00989	4-Hydroxybutanoic acid -- 4-Hydroxybutyryl-Co	1.0	3.1.2 6.2.1 2.8.3 u...
42	C05668_C01013	3-Hydroxypropanoate -- 3-Hydroxypropionyl-...	1.0	3.1.2 6.2.1 2.8.3 u...
44	C04047_C01188	3-Hydroxy-2-methylpropanoate -- 3-Hydroxy-2..	1.0	3.1.2 6.2.1 2.8.3 u...
69	C06000_C06001	(S)-3-Hydroxyisobutyrat -- (S)-3-Hydroxyisobu..	1.0	3.1.2 6.2.1 2.8.3 u...
9	C00136_C00246	Butanoic acid -- Butanoyl-CoA	0.990338	3.1.2 6.2.1 2.8.3 u...
95	C16268_C16267	Cyclopropanecarboxylat -- Cyclopropanecarb...	0.990338	3.1.2 6.2.1 2.8.3 u...
2	C00091_C00042	Succinate -- Succinyl-CoA	0.985714	3.1.2 6.2.1 2.8.3 u...
4	C00083_C00383	Malonate -- Malonyl-CoA	0.985646	3.1.2 6.2.1 2.8.3 u...
16	C00827_C00186	(S)-Lactate -- Lactoyl-Co	0.985646	3.1.2 6.2.1 2.8.3 u...
34	C00683_C02170	Methylmalonate -- (S)-Methylmalonyl-CoA	0.985646	3.1.2 6.2.1 2.8.3 u...
45	C01213_C02170	Methylmalonate -- (R)-Methylmalonyl-CoA	0.985646	3.1.2 6.2.1 2.8.3 u...
60	C03188_C04025	alpha,omega-Dicarboxylic aci -- omega-Carb...	0.985646	3.1.2 6.2.1 2.8.3 u...
6	C00100_C00163	Propanoate -- Propanoyl-CoA	0.980583	3.1.2 6.2.1 2.8.3 u...
10	C04348_C00149	(S)-Malate -- (3S)-3-Carboxy-3-hydroxypropan..	0.976636	3.1.2 6.2.1 2.8.3 u...
36	C00904_C00815	Citramalate -- Citramalyl-Co	0.976636	3.1.2 6.2.1 2.8.3 u...
37	C01011_C00815	Citramalate -- (3S)-Citramalyl-CoA	0.976636	3.1.2 6.2.1 2.8.3 u...
41	C01011_C02614	(S)-2-Methylmalate -- (3S)-Citramalyl-CoA	0.976636	3.1.2 6.2.1 2.8.3 u...
14	C00332_C00164	Acetoacetate -- Acetoacetyl-CoA	0.97619	3.1.2 6.2.1 2.8.3 u...
22	C00356_C03761	3-Hydroxy-3-methylglutarate -- (S)-3-Hydroxy-...	0.972093	3.1.2 6.2.1 2.8.3 u...
55	C03058_C02630	2-Hydroxyglutarat -- 2-Hydroxyglutaryl-Co	0.972093	3.1.2 6.2.1 2.8.3 u...

Chemical Similarity Scores



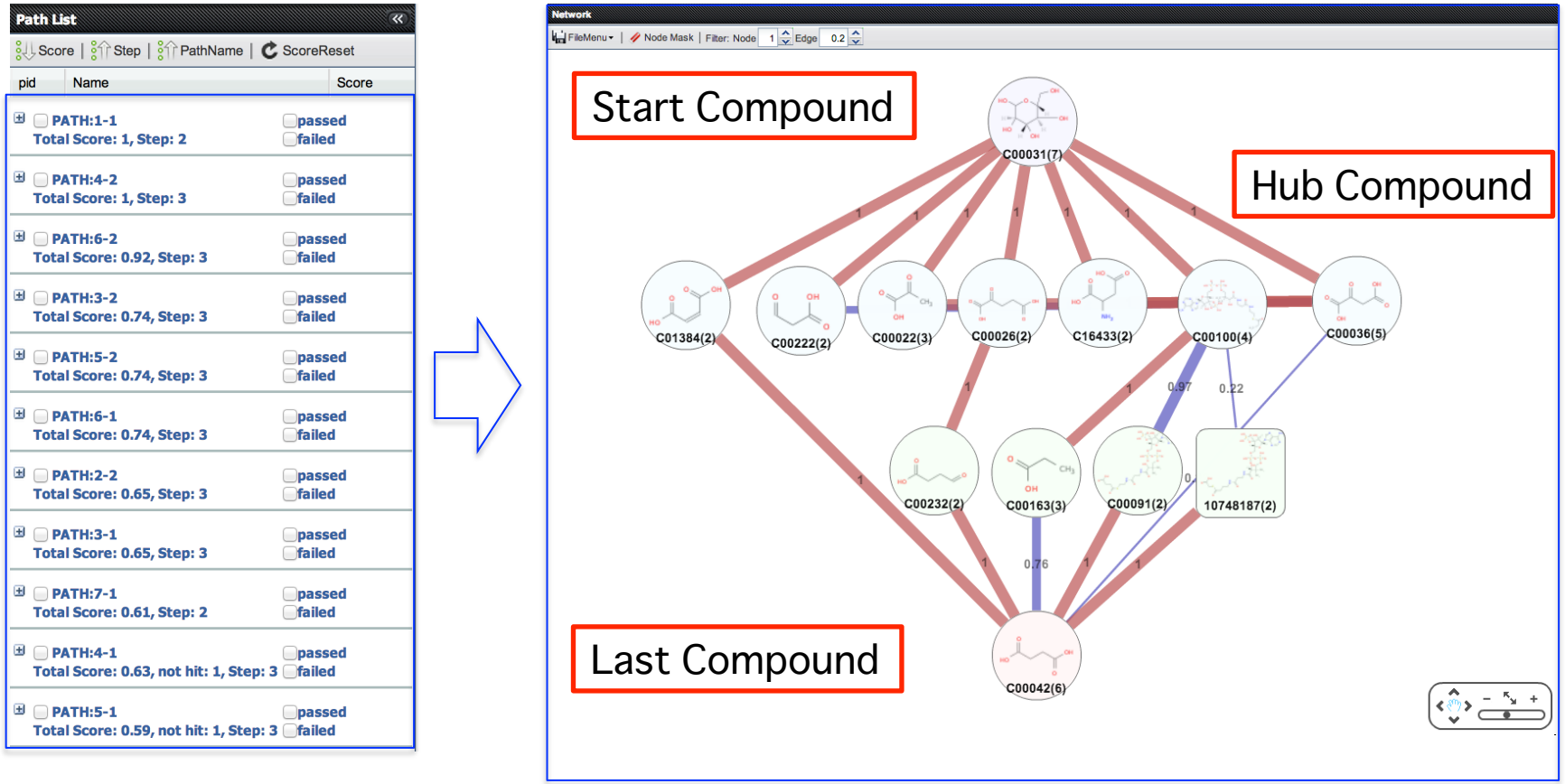
Multiple Alignment
Find Motifs

Genome mining

Enzyme/Gene Selection

M-path

Example) Glucose ~ Succinate



展開

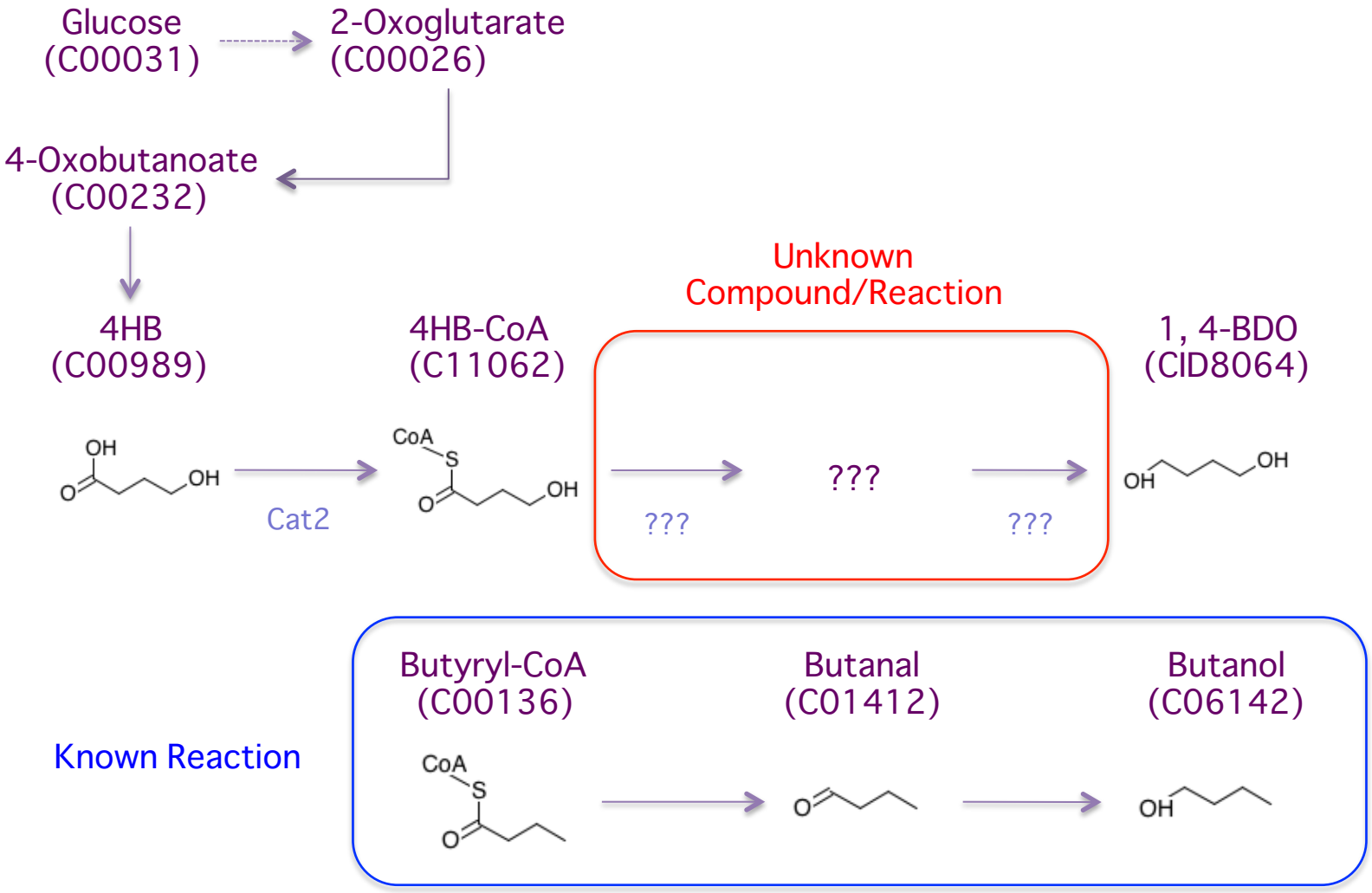
1) Design metabolic pathways for specified target compounds

2) Enumerate reachable compounds from specified start compounds

3) Expand the scope of metabolic pathways

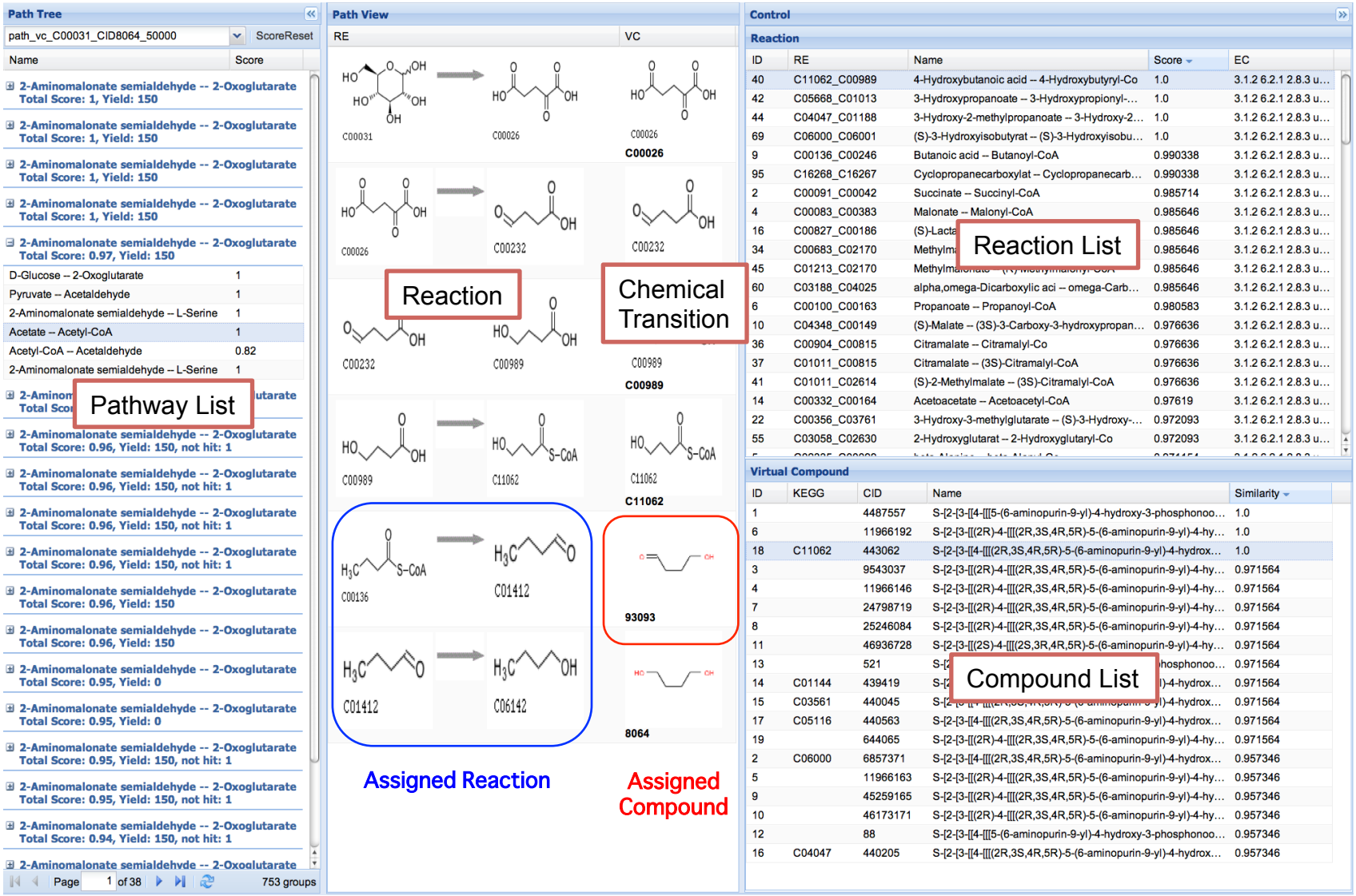
4) Find genes/metabolites using omics data

事例）1,4-Butanediol



Genomatica（Nat. Chem. Biol. 2011）

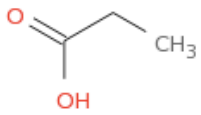
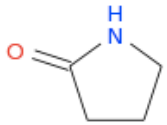
事例）1,4-Butanediol



Reachable compounds

Calculate N steps from start compounds (eg. glutamate)

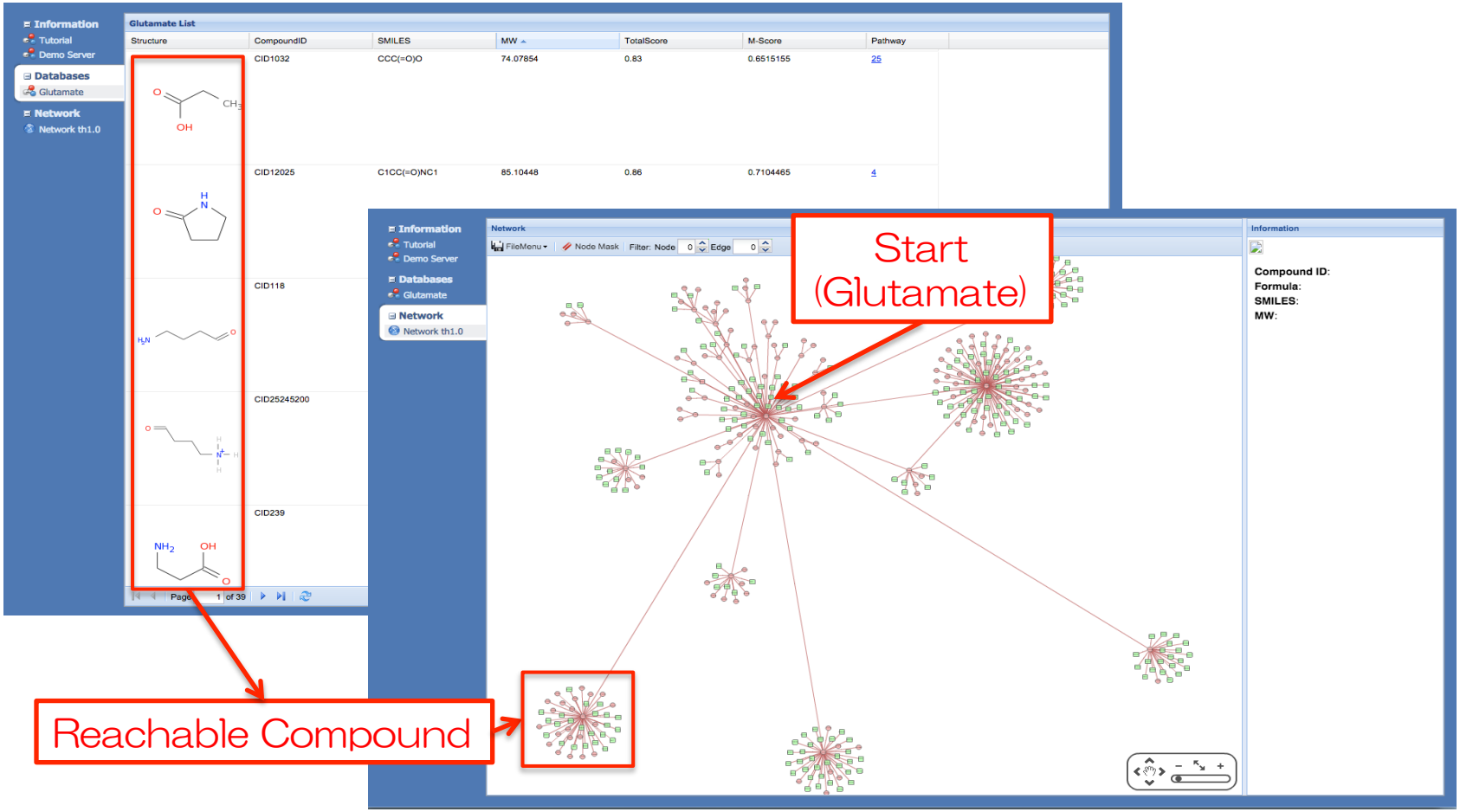
Glutamate List

Structure	CompoundID	SMILES	MW	TotalScore	M-Score	Pathway
	CID1032	CCC(=O)O	74.07854	0.83	0.6515155	25
	CID12025	C1CC(=O)NC1	85.10448	0.86	0.7104465	4
	CID118	C(CC=O)CN	87.12036	1.0	1.0	8

Page 1 of 39 | Displaying Compounds 1 - 50 of 1928

Reachable compounds

Integrate all pathways from glutamate



Expanding Metabolic Pathway

PubChem Database（～50M）

データ抽出

Any Chemicals
(ex:α-Amino Acids～7k)

MW = 99.1　MW = 100

MW = 99.1　MW = 101

網羅的
パスウェイ計算

M-path

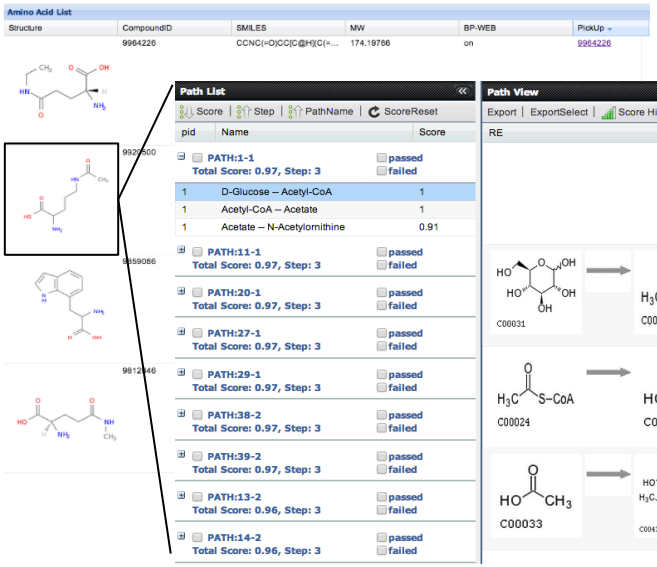
計算条件検討
スコア比較

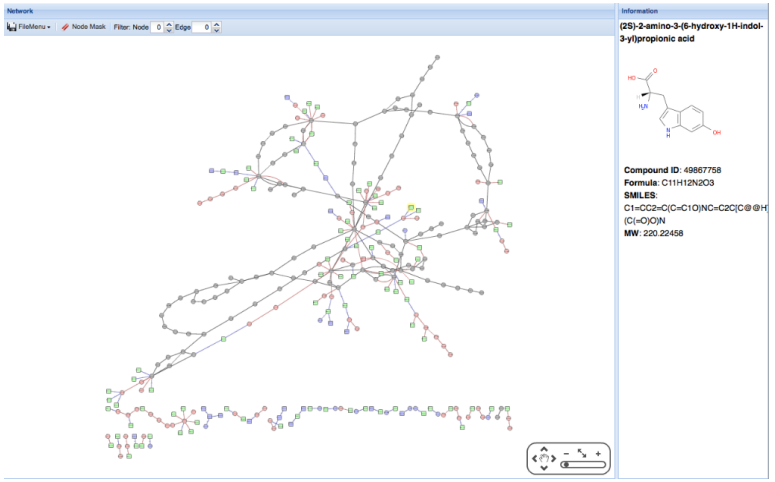
独自の
化合物DB構築

統合

Comprehensive Chemical Pathway Database

パスウェイ出力・選択
ネットワーク解析

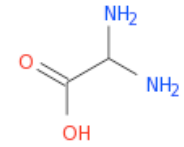
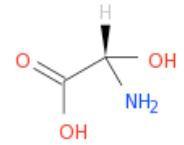
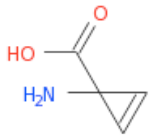
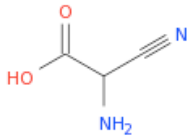




Expanding Metabolic Pathway

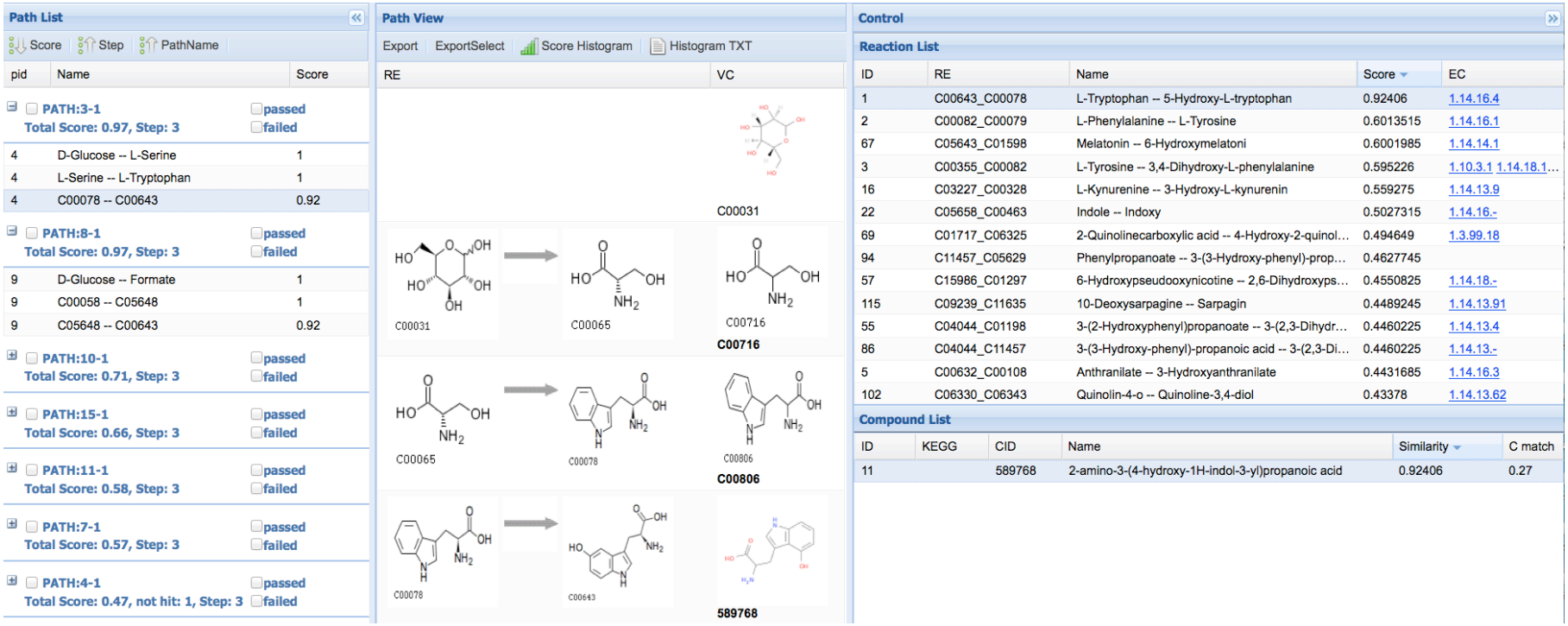
http://bp.scitec.kobe-u.ac.jp/m-path/db/

Amino Acid List

Structure	CompoundID	SMILES	MW	TotalScore	M-Score	Pathway
	CID18976960	C(C(=O)O)(N)N	90.08124		0.430622	yes
	CID16744473	[C@H](C(=O)O)(N)O	91.066	0.88		yes
	CID10996968	C1=CC1(C(=O)O)N	99.088	0.77	0.3058555	yes
	CID533486	C(#N)C(C(=O)O)N	100.07606	0.86	0.6555185	yes

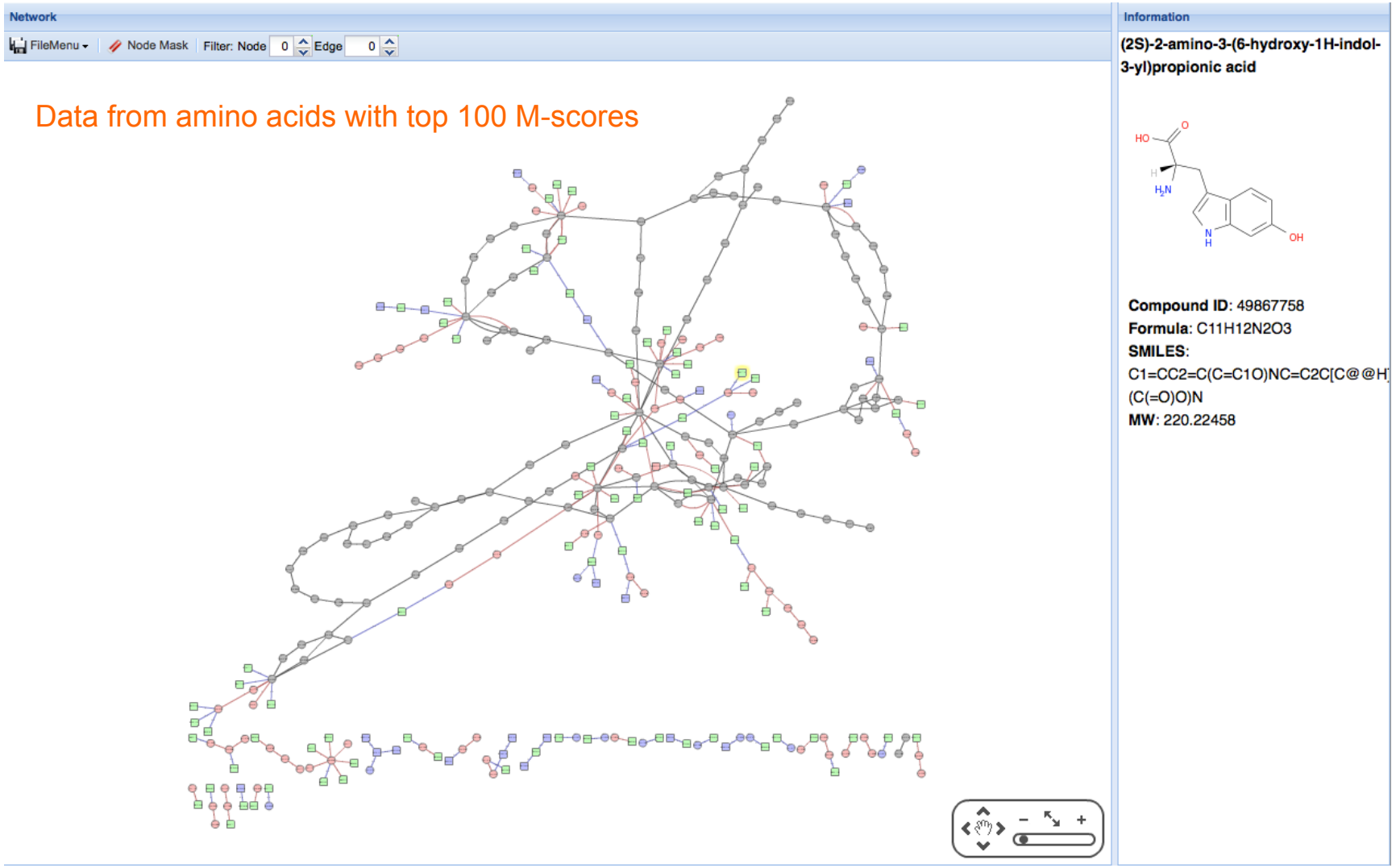
Expanding Metabolic Pathway

http://bp.scitec.kobe-u.ac.jp/m-path/db/



Expanding Metabolic Pathway

http://bp.scitec.kobe-u.ac.jp/m-path/aa/



ありがとうございました

荒木 通啓　（Michihiro Araki）
araki@port.kobe-u.ac.jp

神戸大学　自然科学系先端融合研究環
Organization of Advanced Science and Technology
Kobe University