

◇ **第四部 インシリコ創薬関連**
Part4. Insilico drug design

株式会社 インシリコデータ
湯田 浩太郎

Contents:

挨拶: Greetings:

株式会社 インシリコデータ (In Silico Data, Ltd.)

湯田 浩太郎 (Kohtaro Yuta)

◆導入 計算毒性学と「化学データサイエンス」

Introduction: Computational Toxicology and “Chemical Data Science”

◇第一部 計算機化学 (Computer Chemistry) 関連

Part1. Computer Chemistry

◇第二部 化学多変量解析／パターン認識 (ケモメトリックス (Chemometrics)) 関連

Part2. Chemical multivariate analysis / pattern recognition (Chemometrics)

◇第三部 人工知能 (Artificial Intelligence) 関連

Part3. Artificial Intelligence

◇第四部 インシリコ創薬関連

Part4. Insilico drug design

□創薬でのデータサイエンス展開例

Examples of data science development in drug discovery

◇創薬時における化合物と種々特性の関係

Relationships between compounds and various properties during drug discovery

創薬への**インテグレートド概念**の提案

Proposed integrated concept for drug discovery

◇創薬時における「**並列創薬**」の提案

Proposal of “Parallel Drug Discovery” at the time of drug discovery

上記**インテグレートド概念**を実施するインシリコ手法
として「**並列創薬**」を提案し、システム上で実施

Proposed “Parallel Drug Discovery” as an in silico
method to implement the above integrated concept and
implemented it on the system

◇化合物と種々特性の関係：インテグレートド概念

Relationship between compounds and various properties:

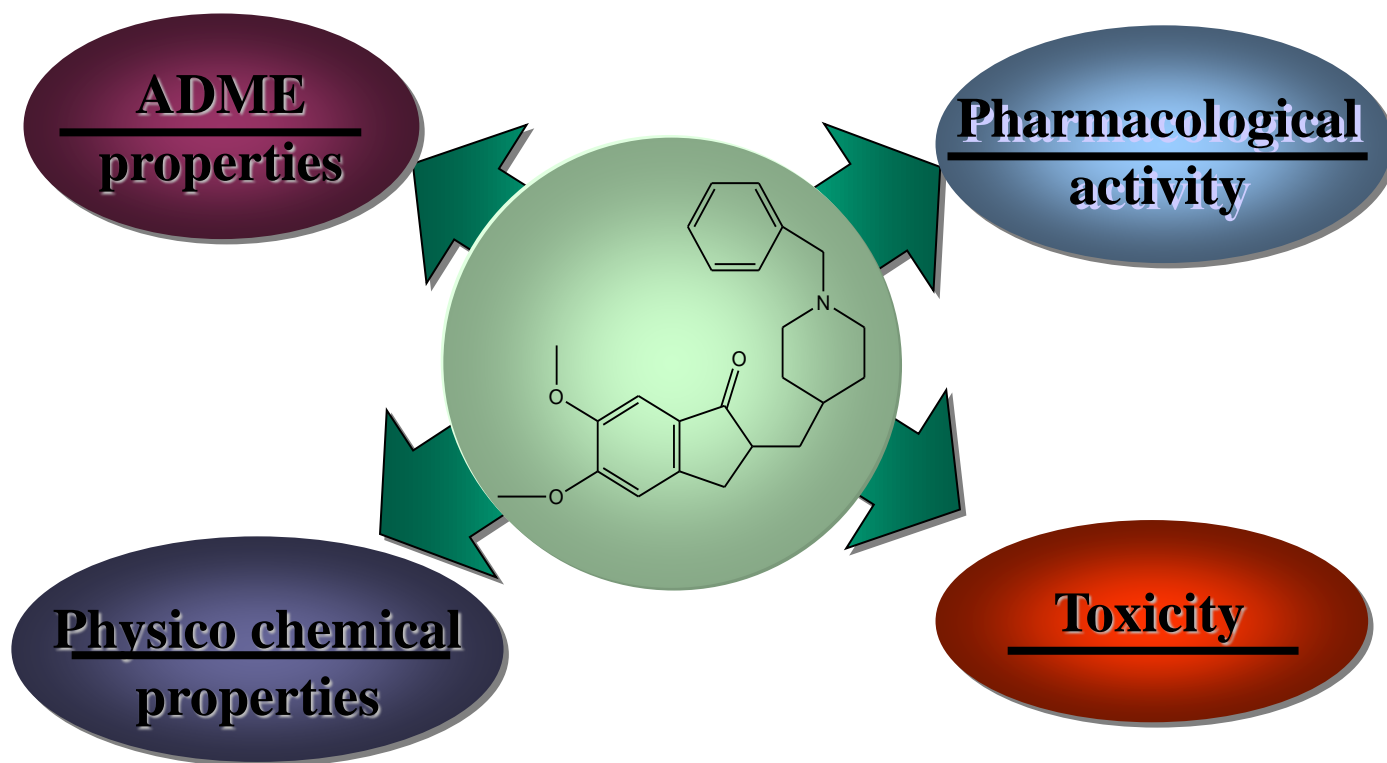
Integrated concept

- 化合物関連の様々な研究や技術は何らかの基本概念を基に開発／展開される
- 新たに基本概念を設定
- 新たな基本概念に従った化合物関連の様々な手法が展開可能となる
- この新たな基本概念として「**インテグレートド概念**」を提案した
 - Various research and technologies related to compounds are developed / deployed based on some basic concept
 - New basic concept set
 - Various compound-related methods based on new basic concepts can be developed
 - Proposed "Integrated concept" as this new basic concept

化合物は
構造式が決まれば
総ての特性が決定する

All properties of a compound are determined once the structural formula is determined

□インテグレートド概念図
Integrated conceptual diagram



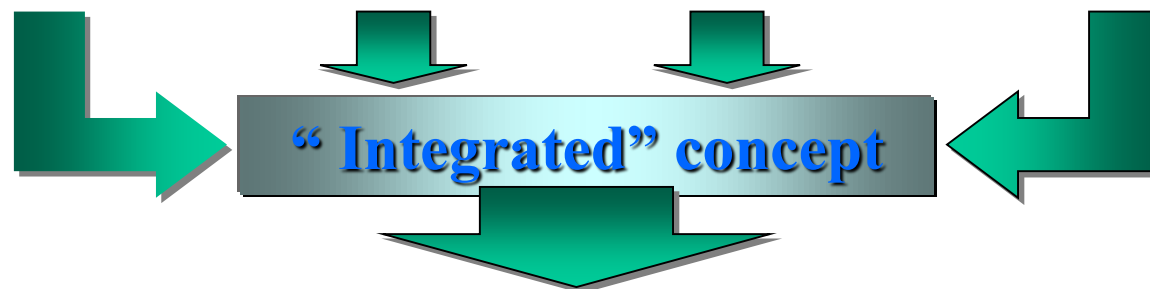
Relations between compound structure and various properties.

□インテグレートド概念図
Integrated conceptual diagram

‘Integrated’ concept for drug development

Activity + ADME + Toxicity + Property

All drug properties shall be considered at the same time



“Integrated” in silico screening & drug design

The basic structure of ‘Integrated’ conception.

□インテグレートッド実施例 Integrated implementation example

- ◇複数の特性を同時に評価しながら化合物をスクリーニングする
予め複数の予測モデルを用意し、これらの複数の要因について同時に予測／評価する

“インテグレートッド”インシリコスクリーニング “Integrated” in silico screening

1次元スクリーニング

1薬理活性／化合物

A	B	C	D
Order Col	Structure	R1 Label	Result (alpha 1)
1	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	1671.08
2	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	1722.88
3	<chem>c1ccc(cc1)C=C</chem>	Benzoylvinyl	2536.82
4	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.64
5	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
6	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
7	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
8	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
9	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
10	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
11	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
12	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
13	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
14	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
15	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15
16	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15

従来型のアプローチ

2次元スクリーニング

複数の薬理活性
複数のADME特性
複数の毒性
複数の物性

化合物

A	B	C	D	E	F	G	H	I	J	K	L
Order Col	Structure	R1 Label	Result (alpha 1)	Result (alpha 1b)	BSA	CYP1A9	CYP3A	Carcinogenicity	AMES	logP	logD
1	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	1671.08	1839.2	1647.09	1	0	1	0	2.90812	-0.341246
2	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	1722.88	1879.38	1864.05	1	0	1	0	2.02751	-0.321069
3	<chem>c1ccc(cc1)C=C</chem>	Benzoylvinyl	2536.82	2555.85	2931.72	0	0	0	1	2.42194	-1.42444
4	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.64	2768.54	2674.39	1	1	1	0	2.39057	2.59433
5	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15	2768.54	2674.39	1	1	1	0	2.39057	2.59433
6	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15	2768.54	2674.39	1	1	1	0	2.6236	2.19323
7	<chem>c1ccc(cc1)C(=O)O</chem>	Phenylacetyl	2555.15	2768.54	2674.39	1	1	1	0	2.59095	1.61159
8	<chem>c1ccc(cc1)C(=O)O</chem>	Benzamido	3730.95	4064.31	3754.68	1	0	1	1	1.15520	1.57961
9	<chem>c1ccc(cc1)C(=O)O</chem>	Benzamido	3730.95	4064.31	3754.68	1	0	1	1	1.15520	1.57961
10	<chem>c1ccc(cc1)C(=O)O</chem>	Carbamoyl	3775.32	4369.63	3285.43	0	1	0	0	-1.19859	3.06413
11	<chem>c1ccc(cc1)C(=O)O</chem>	Bromo	3777.35	3799.94	4109.98	1	0	0	1	1.38841	-0.533316
12	<chem>c1ccc(cc1)C(=O)O</chem>	Benzylidene	3783.45	4182.23	4106.3	0	0	1	0	1.05442	0.384083
13	<chem>c1ccc(cc1)C(=O)O</chem>	Benzyl	3786.09	4006.94	3935.08	1	0	0	1	2.11857	0.24521
14	<chem>c1ccc(cc1)C(=O)O</chem>	Carbamoyl	3790.39	4247.71	3922.87	1	0	0	0	-0.65745	0.893304
15	<chem>c1ccc(cc1)C(=O)O</chem>	Phenoxy	3798.74	4230.52	3270.84	0	1	0	0	2.46775	1.10953
16	<chem>c1ccc(cc1)C(=O)O</chem>	Carbamoyl	3805.29	4080.55	3223.85	0	0	1	0	-1.29534	2.48536

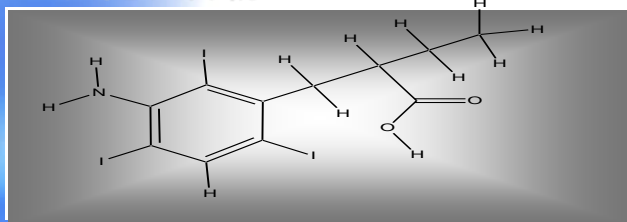
次世代型アプローチ

□インテグレートッド実施例

“インタラクティブ(リアルタイム)”ドラッグデザイン

* 構造修正による薬理活性/ADME/毒性/物性変化の
即時チェックによるドラッグデザインの実施

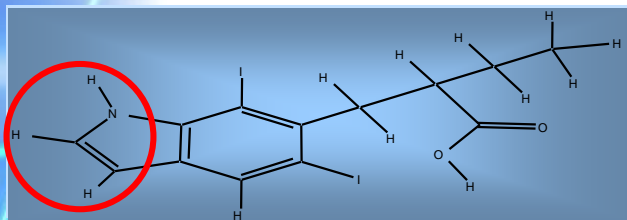
出発化合物



構造修正



薬理活性
ADME
毒性
物性



修正後の化合物

antibacterial	Antiinflammatory	anticancer	pesticide
carcinogenicity	Ames test	LD50	others

ADME and properties

Caco-2	BBB	CYP
LogP	pKa	LogD _{7.4}

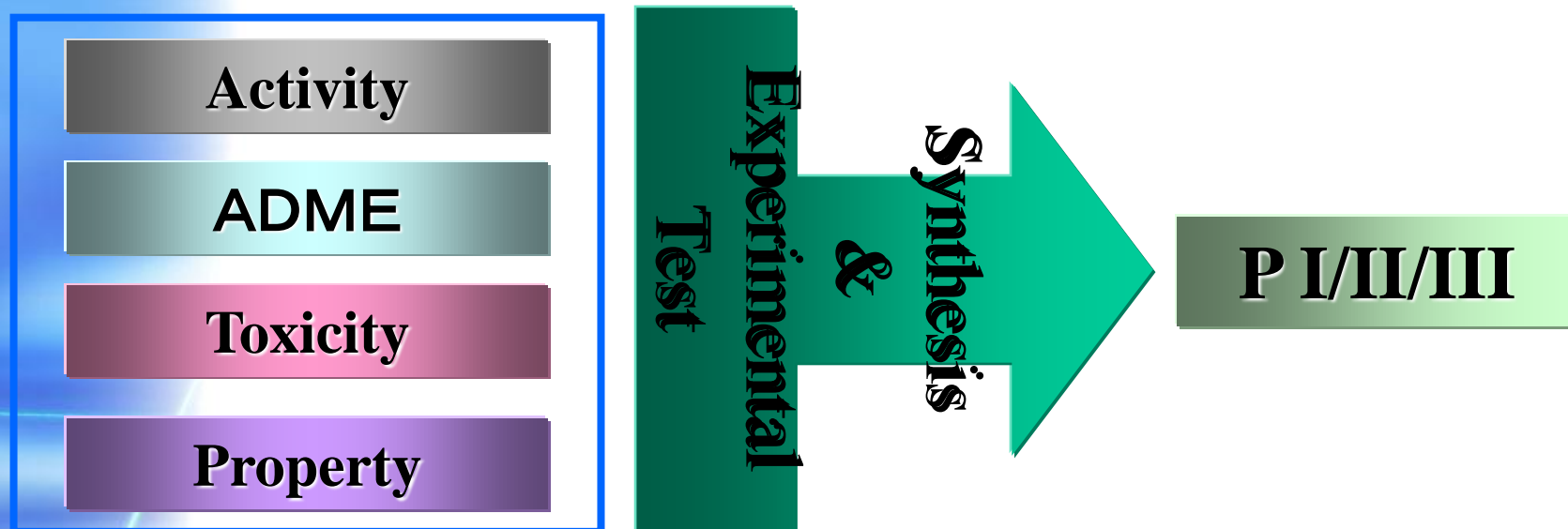
antibacterial	Antiinflammatory	anticancer	pesticide
carcinogenicity	Ames test	LD50	others

場所や時間を問わず、思いついた時に
構造式を書きただけに必要な情報が得られる

複数の予測モデルを用意し、構造式を変更するとリアルタイムで予測値が変わる。
この変化を見ながら、化合物構造式を対話型でデザインする

□ 並列創薬概念図

‘Parallel Drug Design’



In Silico Screening

Basic flow of the ‘Parallel Drug Design’.

並列創薬:

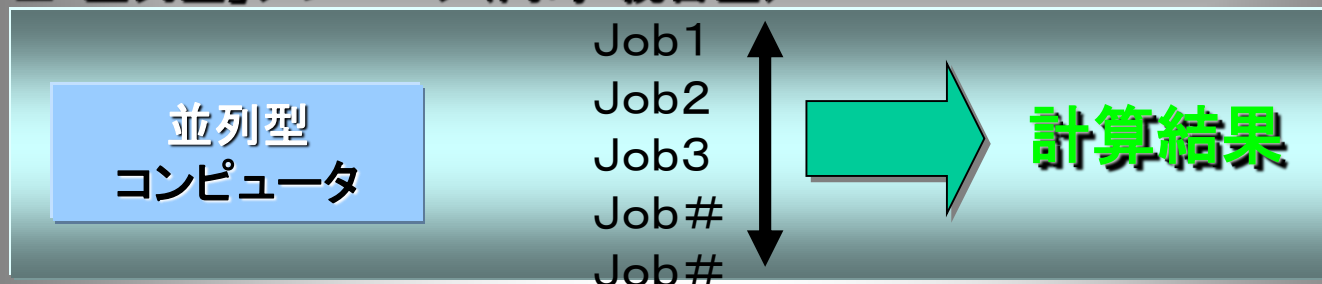
薬理活性／ADME／毒性／物性等の諸特性を同時評価しつつ化合物デザイン実施

□ 並列創薬概念図

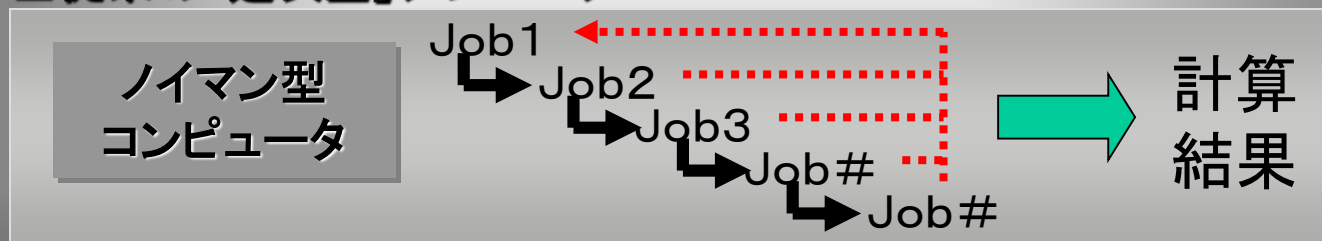
コンピュータ技術に例えた比較

「並列型」アプローチと「逐次型」アプローチ

□ 「並列型」アプローチ (同時・統合型)



□ 従来の「逐次型」アプローチ



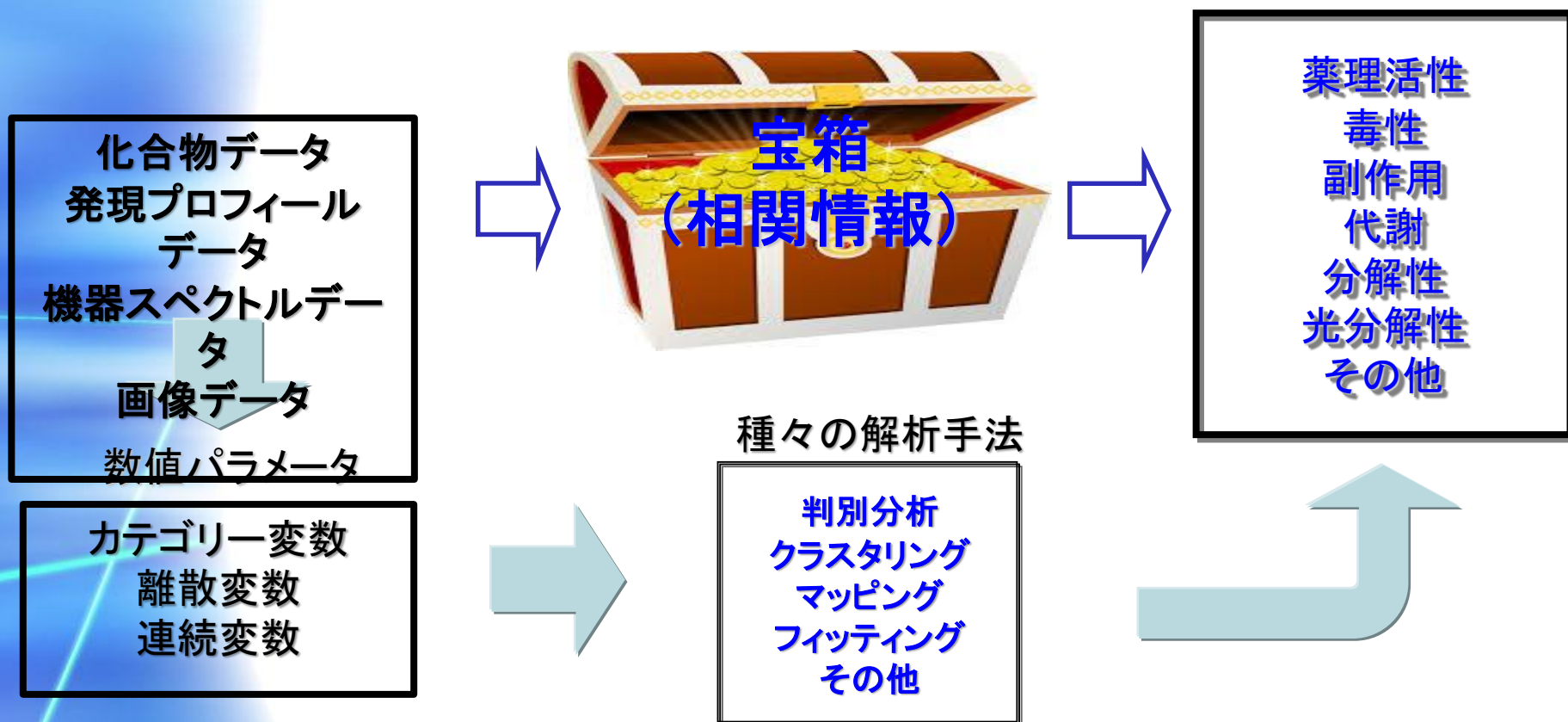
並列型アプローチ:

インテグレートド概念で提案する様々な化合物デザイン手法次世代型アプローチ

逐次型アプローチ:

従来型で、特性単位で値を決定し、残る特性を順番に決定してゆく

◇構造－活性/毒性/ADME/物性相関、メタボロミクス、機器スペクトル解析



◇構造－活性相関への展開

ケモトリックス解析による構造－活性相関は関連技術の発展や開発により様々な適用パターンで急速に展開されている

①構造－活性相関(Hansch-Fujita法)

データ解析手法として重回帰、パラメーターはHansch-Fujitaパラメータ

②ケモトリックスによる構造－活性/毒性/ADME/物性相関

データ解析手法はケモトリックス、パラメーターは多種利用

③バーチャルスクリーニング

化合物は化合物データベースより、パラメーターは種々

④インシリコHTSスクリーニング

HTS結果の化合物、パラメーターは種々

⑤ドラグリポジショニング

種々薬理活性予測モデル、パラメーターは種々

◇バイオ関連解析への展開

バイオ関連研究分野でも多変量解析／パターン認識の展開による
様々な解析研究が展開されている

①遺伝子(ゲノム)解析

遺伝子配列をターゲットとしてホモロジー(相同性)検索、モチーフ検索

②発現プロフィール解析

マイクロアレイやシーケンシングにより、ある個体の器官、組織、
細胞ごとに調べられた遺伝子発現の全体的な様子で、理想的には
全遺伝子の発現量として表現される。

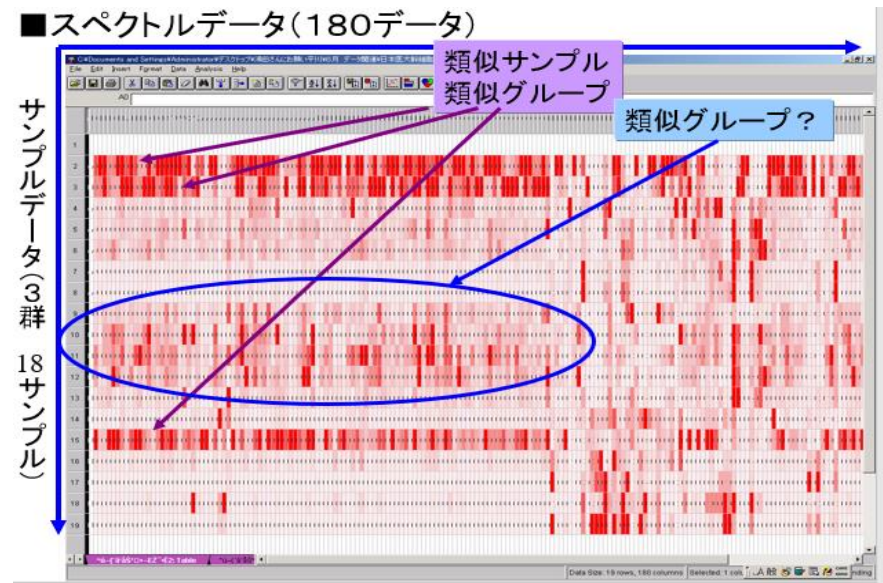
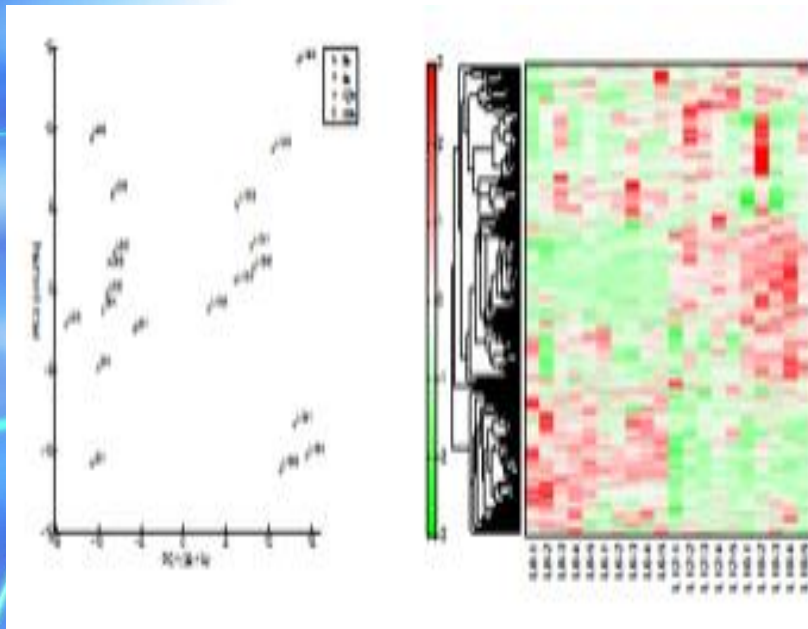
<https://www.weblio.jp/content/発現プロフィール>

③SNPs解析

SNP(1塩基多型)を遺伝子中より発見するアプローチ。
この解析には統計手法が適用される

◇メタボロミクス解析への展開

メタボロミクス (Metabolomics) あるいはメタボローム解析 (Metabolomic analysis) とは、細胞の活動によって生じる特異的な分子を網羅的に解析することである[1]。メタボロームという語は、ある生物の持つ全ての代謝産物(メタボライト)を表す。伝令RNAの発現データやプロテオームの解析だけでは細胞で何が起きているのか分からないが、メタボロームのプロファイルは細胞のある瞬間の生理を明らかにすることができる。



幹細胞のNMRデータを用いた解析

4. ルールベース型人工知能事例

・EMILシステムとは

EMIL とは:(**E**xample-**M**ediated-**I**nnovation-for-**L**ead-**E**volution)

・京都大学教授(現名誉教授)発案による、ルールベース型人工知能システム

実施目的:

・メディシナルケミストが有する化合物変換ノウハウを 用いて
目的薬理活性向上を効率的に実現する化合物構造式変換を行う
人工知能システム

* 変更された化合物は「バイオアナログウス(生物学的等価)」化合物と
命名された

変換ルールの取り出し:

・製薬および農薬関連企業における最先端のメディシナルケミストが召集され、
変換ルール取り出しチームを結成して実施

開発言語(当初):ハイブリッド形態

化合物関連操作等 ⇒ FORTRAN(科学関連分野言語)

ルール稼働用言語 ⇒ LISP(人工知能関連分野言語)

4. ルールベース型人工知能事例

・EMILシステムとは

EMIL

開発目的: (Example-Mediated-Innovation-for-Lead-Evolution)
創薬研究者が有する化合物構造式の変換ノウハウ活用

- ・創薬研究者(特にメディシナルケミスト)は薬理活性の最適化を目指して化合物構造式のデザインが必要
- ・化合物構造式のデザイン可能性はほぼ無限にある
- ・開発目的に従った結果となるようにデザインしたい
- ・薬理活性のみならず、毒性や副作用回避も必要
- ・QSARやドッキングが適用困難な場合も多い

□プロの創薬研究者は、過去の豊富なデザイン経験より、効率的なデザイン手法についての様々な知識を持つ

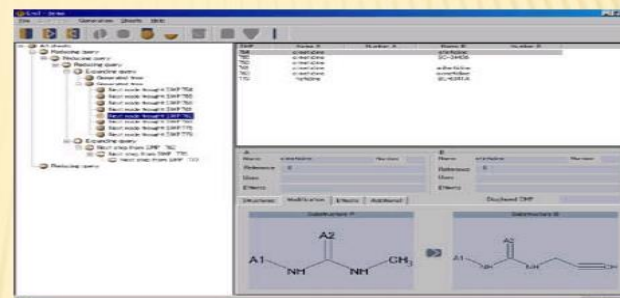
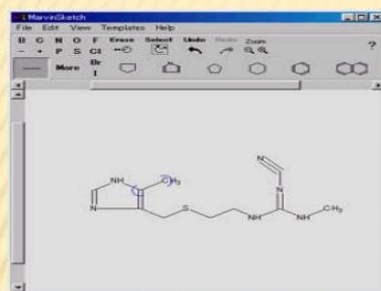
貴重なノウハウ:活用したい

4. ルールベース型人工知能事例

・EMILシステムとは

EMILシステムの流れ図

In Silico Data
Miracles by the KY-methods



過去の構造変換ルールの適用
医薬品／農薬

バイオアナログウス
化合物群

化合物1

化合物2

化合物3

化合物4

化合物1-1

化合物1-2

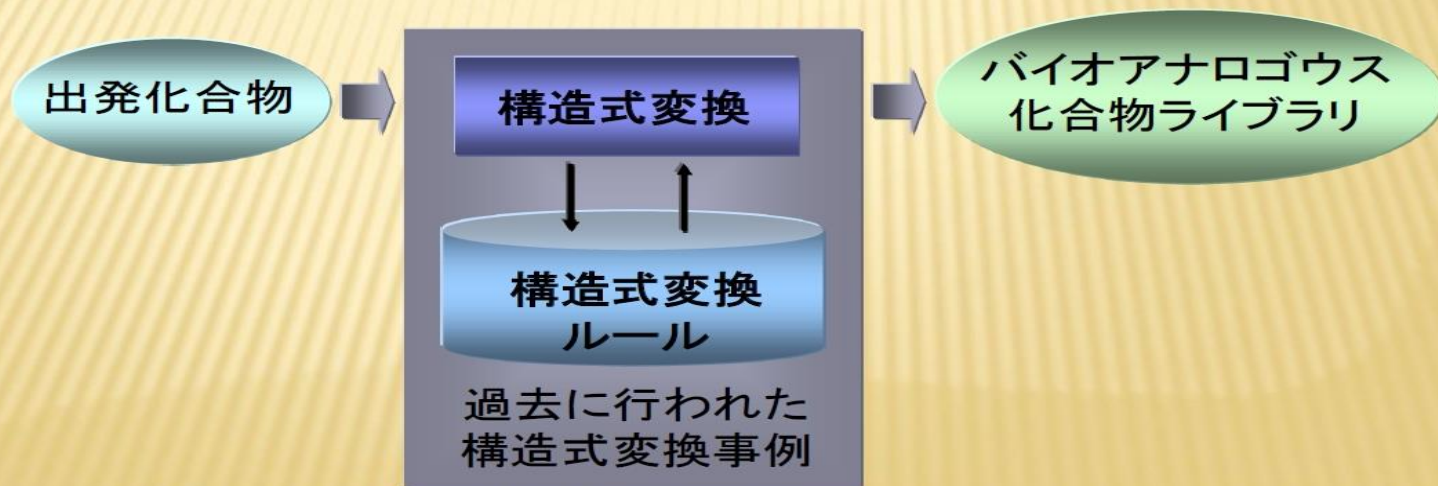
4. ルールベース型人工知能事例

・EMILシステムとは

In Silico Data
Miracles by the KY-methods

EMILシステム(人工知能システム)の機能と流れ

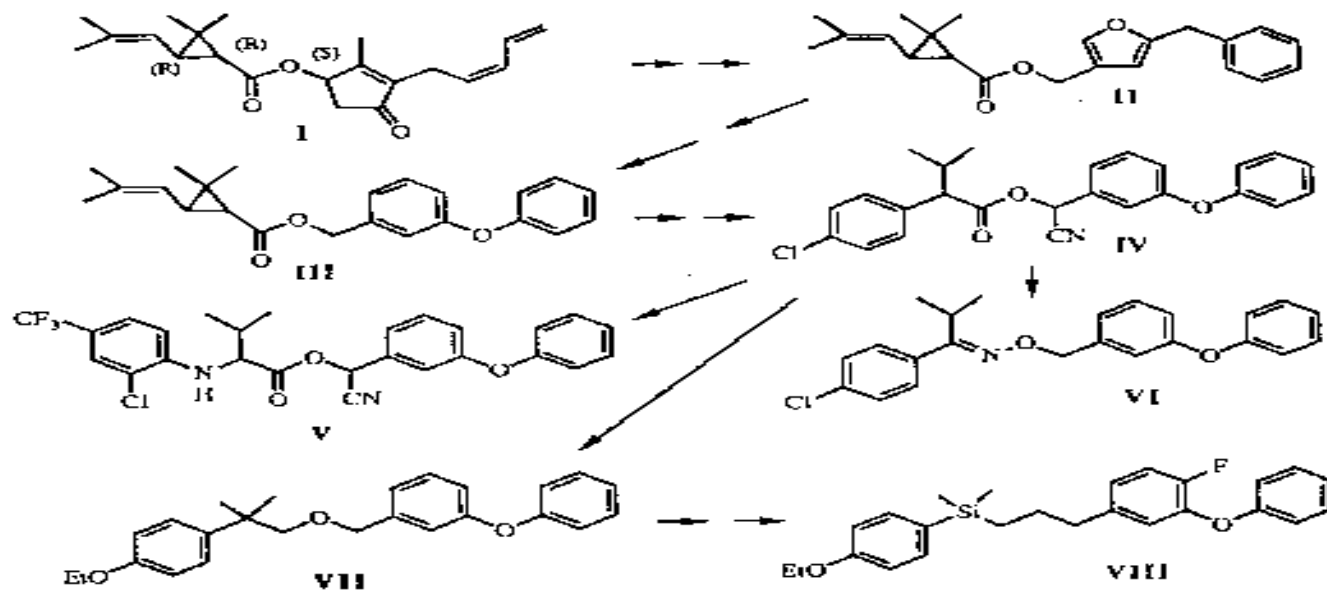
- 過去に行われた医薬品や農薬等の構造変換事例に従って、新たにバイオアナログウス化合物群を創出するプログラム



4. ルールベース型人工知能事例

・EMILシステムとは

化合物構造のバイオアナログウス変換事例



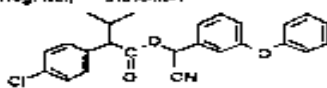
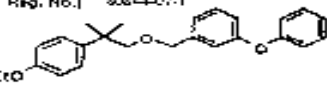
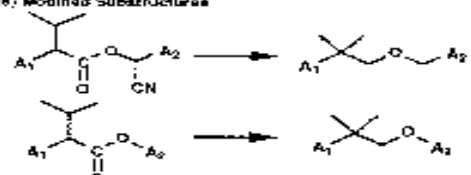

http://cicsj.chemistry.or.jp/14_1/14_1emil.htmlより引用

4. ルールベース型人工知能事例

・研究者のノウハウ収集とプログラム化

化合物構造変換シート (創薬用と農薬用の二種類あり)

EMIL DATA SHEET

Structure A	Med. Class.	Use	Insecticide	Date	Ser. No.	Reg. No.	Prepared by	Structure B	Med. Class.	Use	Insecticide																																																												
AGRO. Effects	Structural Type	Structural Type	Structural Type					AGRO. Effects	Structural Type	Structural Type	Structural Type																																																												
<p>Name (General Name) Fenoxazole (Code No.) S-5602 (CA Reg. No.) 51630-56-1</p>  <p>(Institution) Sumitomo Chemical</p> <p>(1) Mode of Action Prolongation of the Open State of Sodium Channels Block of the Nerve Conduction.</p> <p>(2) Structural Requirements The Acyl Moiety: S-S-R Form. C1-C6 can be replaced by E and F, and p-C by E, R, and O (except nitro) at various positions.</p> <p>(3) Activity (Potency and Features) LUS Suspendant form 0.21 mg/kg Branched form 0.13 mg/head Aqueous emulsion 0.15 mg/head Lethal dose 441 mg/kg LD₅₀ (rat): 0.25 mg/kg i.p.</p> <p>(4) Physicochemical Data Vapor pressure: 2.4 X 10⁻⁷ mmHg LogP (octanol): 6.12 at 25°</p> <p>(5) Others Commercialized for active pharmaceuticals.</p>			Pyrethroids	9/01/2003			Harosuke Yoshitaka	<p>Name (General Name) Dofenpropr (Code No.) MET130C (CA Reg. No.) 80844-07-1</p>  <p>(Institution) Mitsui-Toatsu Chemical</p> <p>(1) Mode of Action Prolongation of the Open State of Sodium Channels Repetitive Excitation of the Nervous System</p> <p>(2) Structural Requirements Original Me group and the connecting ClS are essential. Introduction of CN at the p-positions in the phenyl group is unfavorable.</p> <p>(3) Activity (Potency and Features) LUS Suspendant form 0.15 mg/kg Branched form 0.21 mg/head Aqueous emulsion 0.09 mg/head LD₅₀ (rat, oral): 50,000 mg/kg LD₅₀ (rat, i.p.): 5 ppm (24 h)</p> <p>(4) Physicochemical Data mp: 64-37.5° Vapor pressure: 1.4 X 10⁻⁷ mmHg (10°) logP (octanol): 6.87</p> <p>(5) Others Formerly marketed as a paddy field insecticide.</p>																																																															
References Satohei Niimata, "Design of Bioactive Molecules", H. Yoshioka and K. Shuzhi, Eds., Soft Science Publications, Tokyo, 1986, p.338.		Structural Evolution																																																																					
				(b) Modified Substructures																																																																			
																																																																							
				(c) Effects of the Modification																																																																			
				<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th>Unrecognized Acceptor</th> <th>Softness</th> <th>Hardness</th> <th>Electrostatic</th> <th>Hydrophobicity</th> <th>Hydrophilicity</th> <th>Hydrophobicity</th> <th>Hydrophilicity</th> <th>Hydrophobicity</th> <th>Hydrophilicity</th> <th>Hydrophobicity</th> <th>Hydrophilicity</th> </tr> </thead> <tbody> <tr> <td>Electrostatic</td> <td>+</td> <td>-</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> </tr> <tr> <td>Hydrophobicity</td> <td>+</td> <td>-</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> </tr> <tr> <td>Hydrophilicity</td> <td>+</td> <td>-</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> </tr> <tr> <td>Aq. Solubility</td> <td>+</td> <td>-</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> <td>+</td> </tr> </tbody> </table>								Unrecognized Acceptor	Softness	Hardness	Electrostatic	Hydrophobicity	Hydrophilicity	Hydrophobicity	Hydrophilicity	Hydrophobicity	Hydrophilicity	Hydrophobicity	Hydrophilicity	Electrostatic	+	-	+	+	+	+	+	+	+	+	+	Hydrophobicity	+	-	+	+	+	+	+	+	+	+	+	Hydrophilicity	+	-	+	+	+	+	+	+	+	+	+	Aq. Solubility	+	-	+	+	+	+	+	+	+	+	+
Unrecognized Acceptor	Softness	Hardness	Electrostatic	Hydrophobicity	Hydrophilicity	Hydrophobicity	Hydrophilicity	Hydrophobicity	Hydrophilicity	Hydrophobicity	Hydrophilicity																																																												
Electrostatic	+	-	+	+	+	+	+	+	+	+	+																																																												
Hydrophobicity	+	-	+	+	+	+	+	+	+	+	+																																																												
Hydrophilicity	+	-	+	+	+	+	+	+	+	+	+																																																												
Aq. Solubility	+	-	+	+	+	+	+	+	+	+	+																																																												
				(d) Pro- and/or Epi-Isomers for the Modification																																																																			
				 <p>Instead of the diast. Epi-isomer between C1 and C1 for fenvalerate, that between E1 and C1, optimization of the C number and conversion of ester to ether bonding types were tried.</p>																																																																			

http://cicsj.chemistry.or.jp/14_1/14_1emil.htmlより引用

4. ルールベース型人工知能事例 ・ EMILシステムを用いたドラグリポジショニング

□ EMILシステムを用いたドラグリポジショニング

実施目的:

既存薬物(Tiaramide)誘導体の新規薬理活性の探索

手順と基本技術:

1. 探索する新規薬理活性を5種類選択する
2. 薬理活性ごとに判別関数作成(高速スクリーニング用)
3. Tiaramideの誘導体をEMILを用いて生成
4. 生成された誘導体化合物について薬理活性チェック

4. ルールベース型人工知能事例

・ EMILシステムを用いたドラグリポジショニング

インシリコ創薬への人工知能の部分適用事例  In Silico Data
Miracles by the KY-methods

マルチカテゴリースクリーニング実施事例

■ 5種類の薬理活性に関する同時スクリーニング

1. CCK-A (Cholecystokinin-A) 受容体拮抗剤: Malonamide誘導体、急性膵炎治療薬
2. Ca^{2+} 遊離抑制作用: 1,5-benzothiazepine誘導体、降圧薬、抗狭心症薬
3. 抗潰瘍作用: 4-Bis[(aminoalkyl)amino]-9,10-anthracenediones誘導体
4. 抗菌作用: Quinolone誘導体
5. 抗アレルギー作用: Imidazo誘導体

■ 薬理活性スクリーニング対象化合物群 (化合物ライブラリー)

人工知能システムEMILを用いて創出された、バイオアナログウス化合物群

■ 薬理活性スクリーニング (多変量解析/パターン認識手法)

データ解析支援システムADAPTを用いて5種類の活性スクリーニングを行った

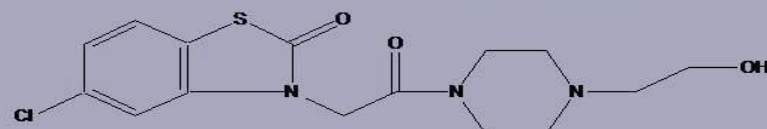
4. ルールベース型人工知能事例

・ EMILシステムを用いたドラグリポジショニング

EMIL実行上での初期化合物と創出化合物群

■ 出発化合物構造式

Tiaramide
(Antiasthmatic, Anti-inflammatory)



■ 適用した構造式変換ルール

医薬品関連変換ルール

■ 変換結果

第一次変換出力構造式数 78化合物

第二次変換出力構造式数 76化合物

(第一次変換化合物中の1化合物を出発化合物とした)

4. ルールベース型人工知能事例

・ EMILシステムを用いたドラグリポジショニング

グ

化合物活性スクリーニング結果 (1)



- 1. CCK-A (Cholecystokinin-A) 受容体拮抗剤; Malonamide誘導体、急性膵炎治療薬
- 2. Ca²⁺遊離抑制作用; 1,5-benzothiazepine誘導体、降圧薬、抗狭心症薬
- 3. 抗潰瘍作用; 4-Bis[(aminoalkyl)amino]-9,10-anthracenediones誘導体
- 4. 抗菌作用; Quinolone誘導体
- 5. 抗アレルギー作用; Imidazo誘導体

化合物ID	活性1	活性2	活性3	活性4	活性5	活性総数
1	x	○	○	○	○	4
2	x	○	○	○	○	4
3	○	x	x	x	x	1
4	x	x	○	○	○	3
5	x	x	○	○	○	3
6	x	x	○	○	○	3
7	x	x	x	○	○	1
8	○	x	○	○	x	3
9	○	x	○	○	○	4
10	x	x	○	○	○	3
⋮						⋮

4. ルールベース型人工知能事例

・ EMILシステムを用いたドラグリポジショニング

□ EMILシステムを用いたドラグリポジショニング

実験結果:

新規薬理活性可能性ありとされる化合物を効率よく探索できた

結果の考察:

1. 複数薬理活性候補化合物が頻度高く発見できた
EMILにより創出された段階で薬理活性候補化合物が多数含まれた
→EMILの構造変換ルールの効果を実証
2. 化学多変量解析／パターン認識(ケモメトリクス)により、
5種類の薬理活性について効率よくチェック出来た
3. 全てバーチャル化合物を用いて短時間で実施された