

日本化学工業協会・新LRI第5期研究報告会シンポジウム
～毒性予測の今後～

東京証券会館 8階ホール

2017/8/25

毒性発現経路および 化学構造情報を指標とした ヒト有害反応の予測法の開発

明治薬科大学・臨床薬剤学研究室

植沢 芳広

日本化学工業協会・新LRI第5期研究報告会シンポジウム
～毒性予測の今後～

東京証券会館 8階ホール

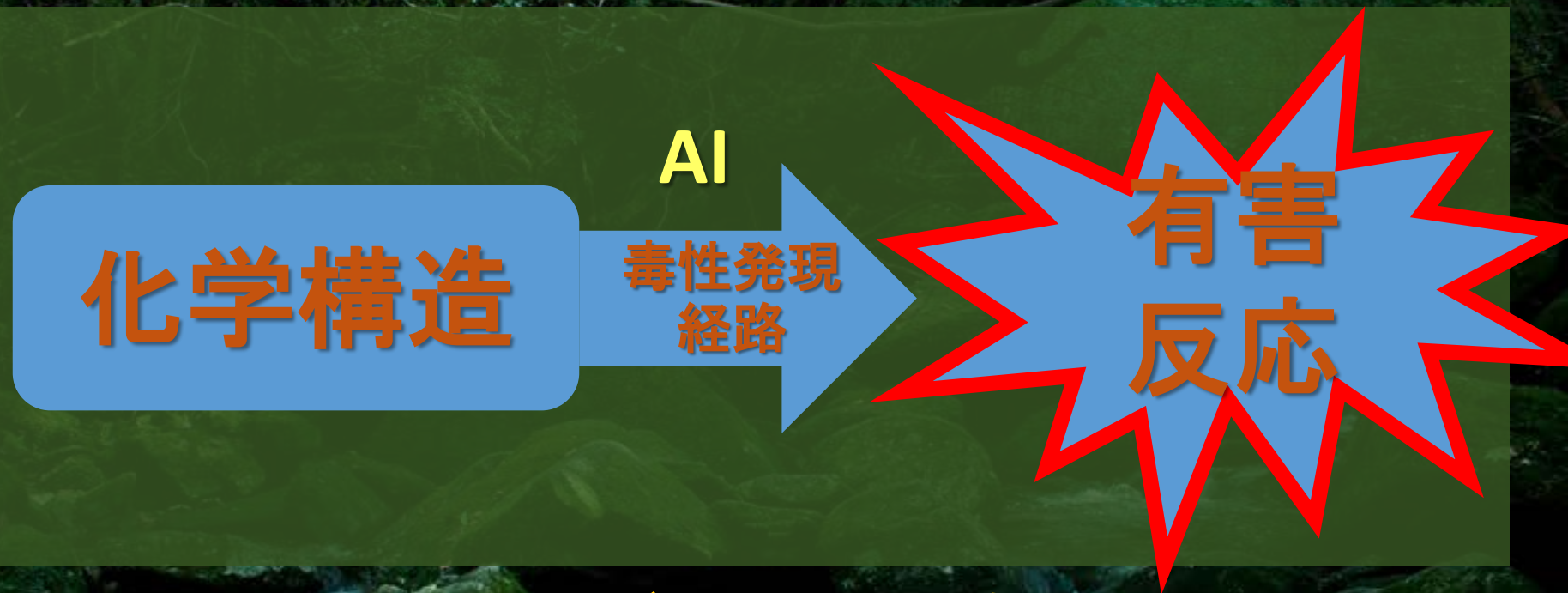
2017/8/25

毒性発現経路および 化学構造情報を指標とした ヒト有害反応の予測法の開発

明治薬科大学 臨床薬剤学研究室

植沢 芳広

毒性発現経路および化学構造情報を指標とした ヒト有害反応の予測法の開発



予測モデルの構築

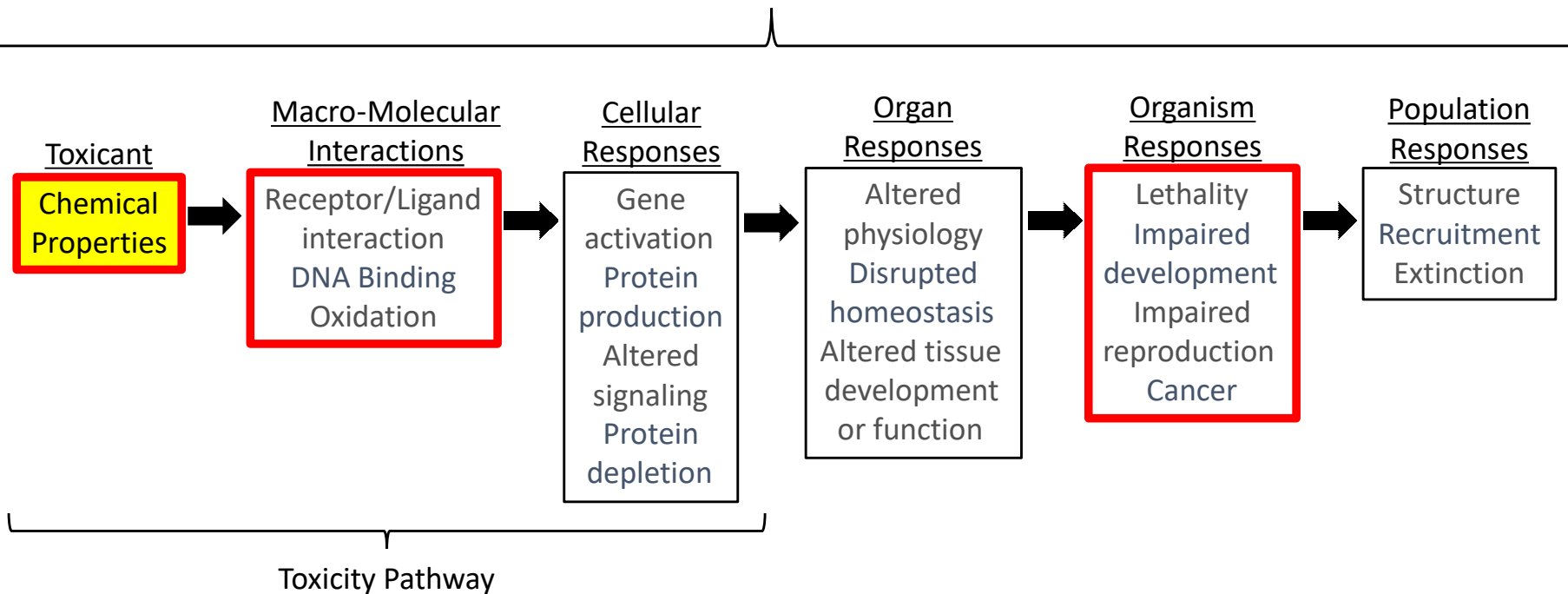
- 創薬初期のフィルター
- 新薬に対する注意喚起
- 化学品の管理

毒性発現経路 (AOP)

- Adverse Outcome Pathway: 有害性発現経路
- OECDでは有害性発現メカニズムをAOPに基づいて整理することを推奨
- ある毒性の原因となる分子レベルの反応から、細胞レベル、臓器レベル、生体レベルなどを経て、最終的な有害性発現に至るまでの、各レベルにおける毒性メカニズムの知見を整理し、それらの因果関係を経路として表したものの。
- 構造毒性相関を見出すことが困難な毒性評価が可能になると期待される。

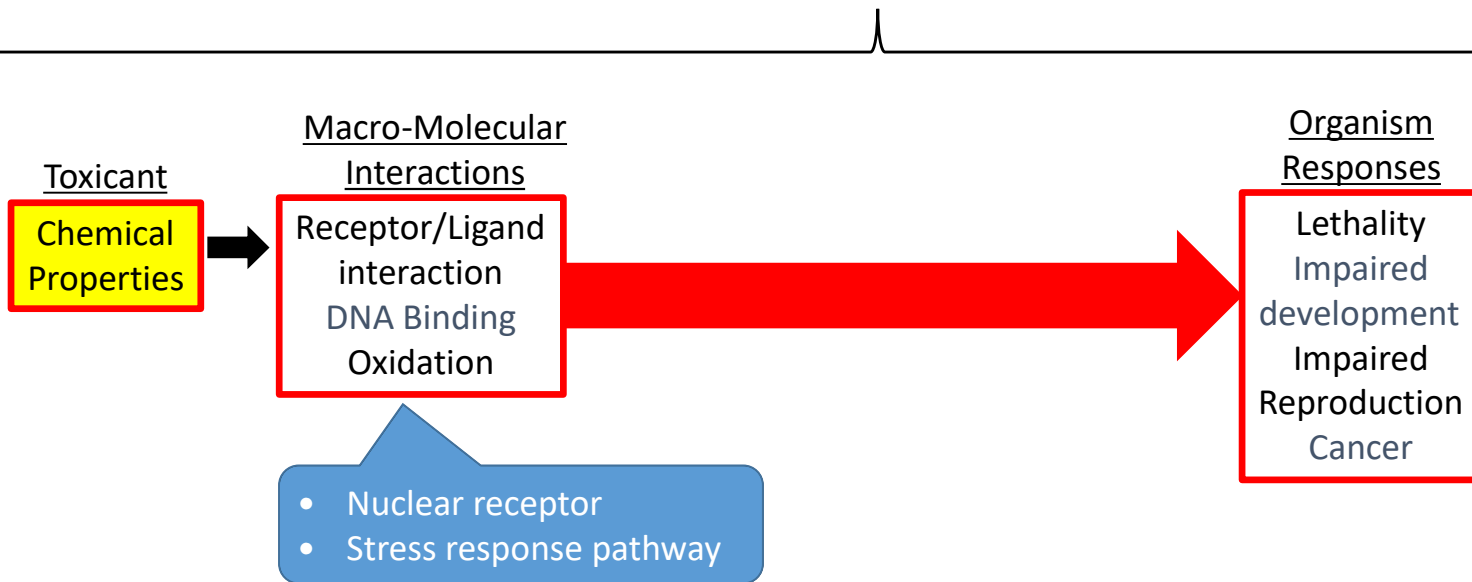
AOP

Adverse Outcome Pathway



AOP

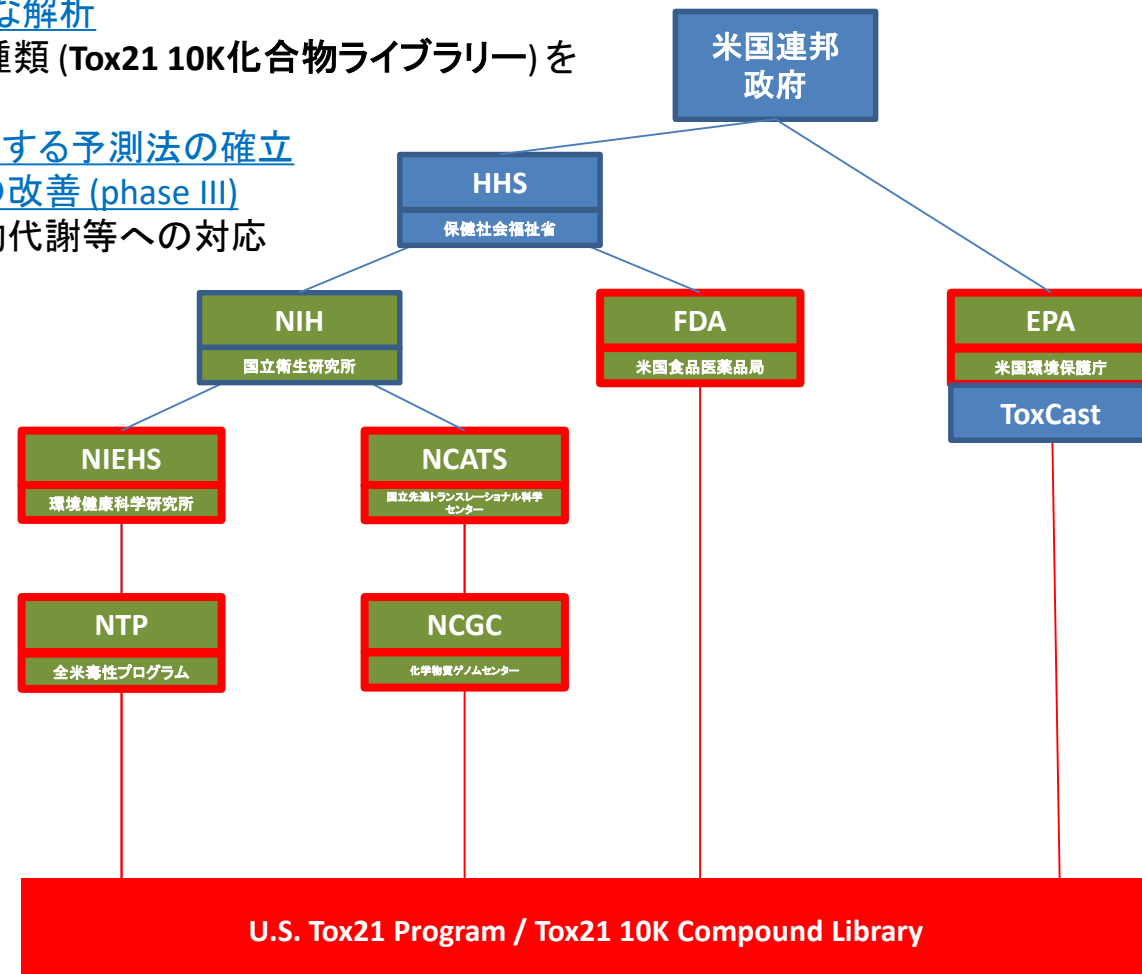
Adverse Outcome Pathway



Toxicology in the 21st Century (Tox21)

米国NIH、EPA、FDAによる施設間共同研究(2008～)

- ① 毒性発現機序の理解
- ② 毒性物質の評価法の確立
 - ロボットを投入した定量的超ハイスループットスクリーニングシステムの構築
 - 重要な核内受容体・ストレス応答パスウェイを調査
- ③ 既知物質に関する網羅的・実験的な解析
 - 代表的な化合物群約10,000種類 (Tox21 10K化合物ライブラリー) を選定
- ④ 有害性が未確認である化合物に対する予測法の確立
- ⑤ 生物学的適用範囲および妥当性の改善 (phase III)
 - ES細胞、iPS細胞を用いた異物代謝等への対応



Tox21-AOPs

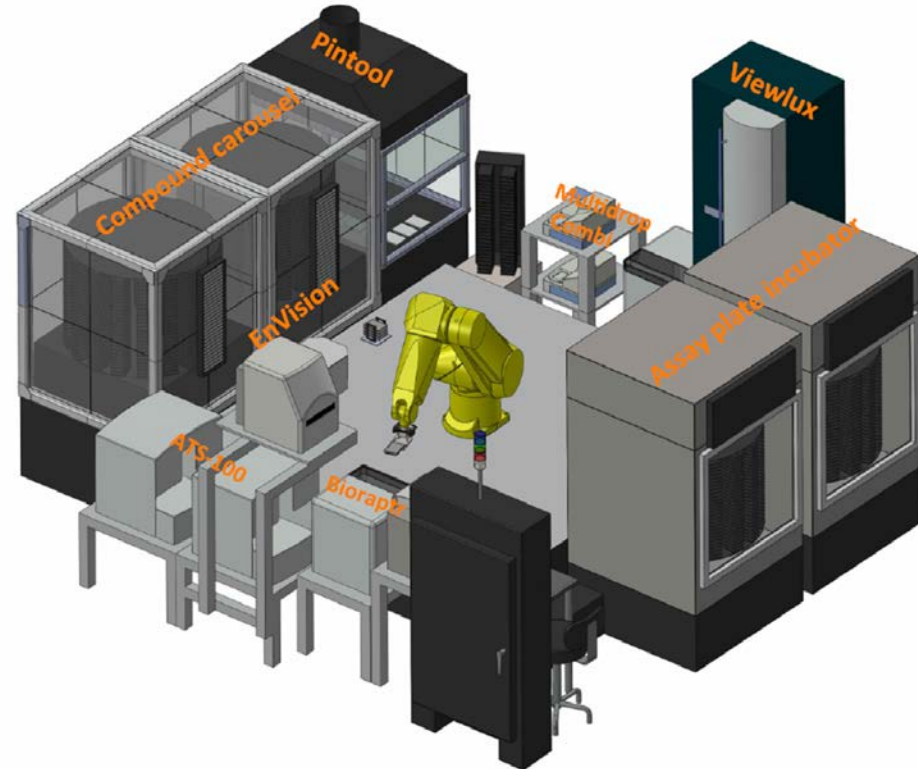
重要な核内受容体・ストレス応答パスウェイを調査



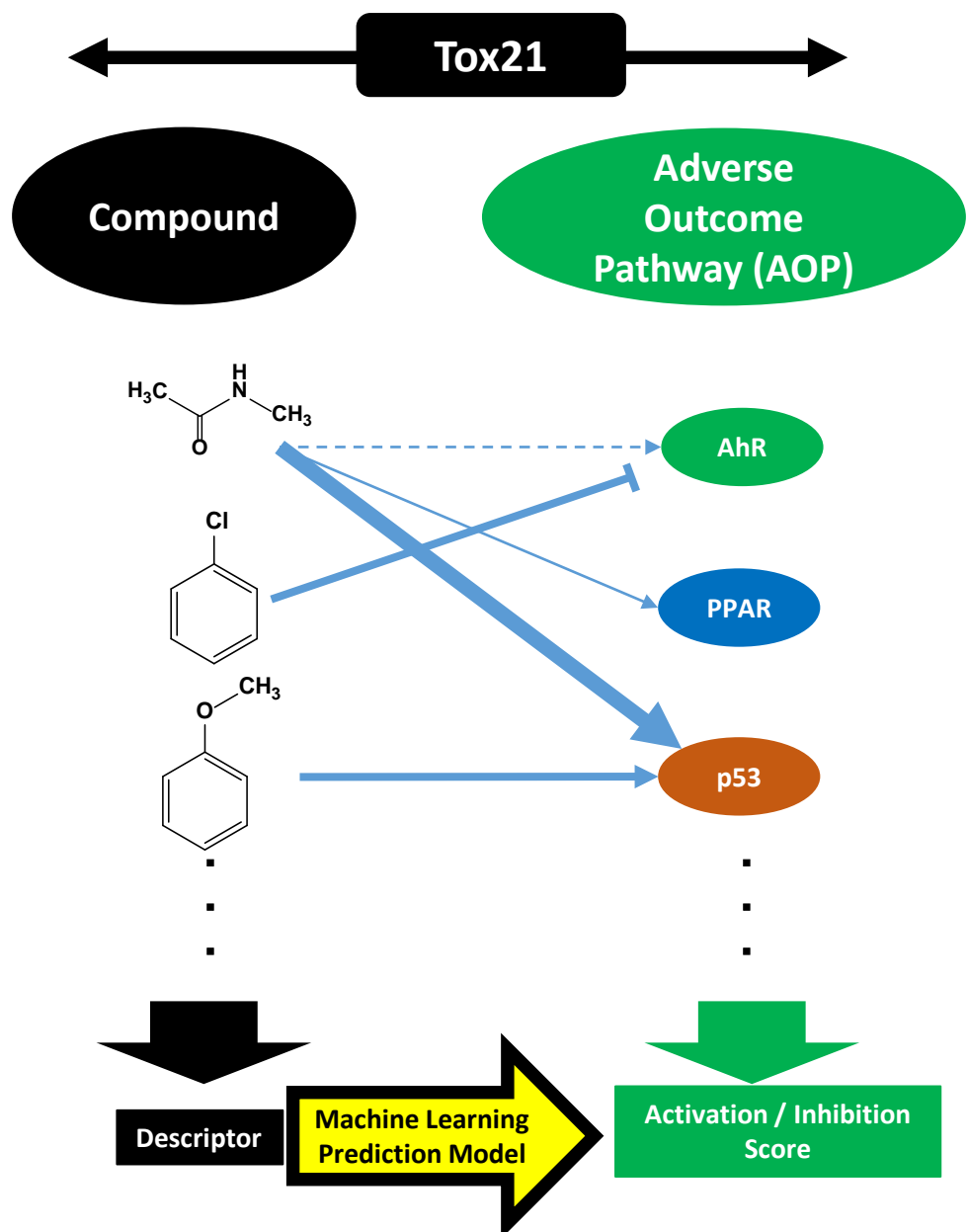
NCBI-PubChemに公開

AhR	Aryl hydrocarbon receptor
AP-1	Activating protein-1
AR	Androgen receptor
ARE	Antioxidant response element
Aromatase	Aromatase
ATAD5	ATPase family AAA domain-containing protein 5
CAR	Constitutive androstane receptor
ERa	Estrogen receptor alpha
ERR	Estrogen related receptor
ERstress	Endoplasmic reticulum stress
FXR	Farnesoid X receptor
GR	Glucocorticoid receptor
H2AX	Histone H2A.x
HIF-1	Hypoxia Inducible Factor-1
HSR	Heat shock response element
MMP	Mitochondrial membrane potential
NFkB	Nuclear factor-kappa B
p53	p53
PGC	Pleiotropic PPARgamma coactivator
PPARd	Peroxisome proliferator-activated receptor delta
PPARg	Peroxisome proliferator-activated receptor gamma
RAR	Retinoic acid receptor
RORg	Retinoic acid receptor-related orphan receptor
RXR	Retinoid X receptor
TR	Thyroid hormone receptor
TSHR	Thyroid stimulating hormone receptor
VDR	Vitamin D receptor
dsbDNA damage	Stalled replication forks
srf DNA damage	DNA double-strand breaks

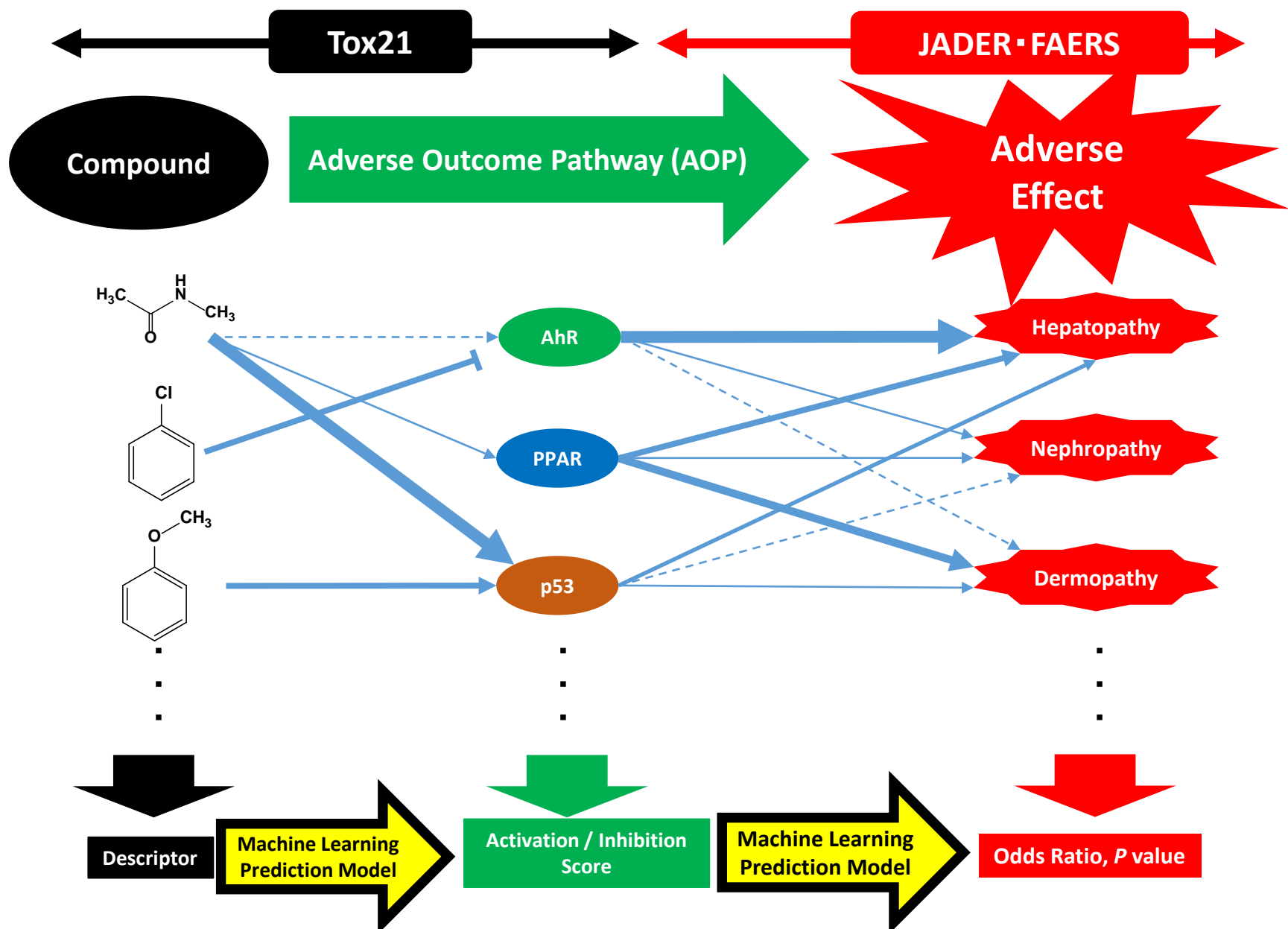
ウルトラハイスループットスクリーニングによる10,000化合物の網羅的レポーター遺伝子アッセイ



Relationship among Compounds, AOPs and Adverse Effects



Relationship among Compounds, AOPs and Adverse Effects



副作用データベース JADER

我が国における
大規模医薬品副作用データベース
Japanese Adverse Drug Event Report
database

2017年5月現在 45万件

290万件(7000医薬品)

- 独立行政法人医薬品医療機器総合機構 (PMDA)のHPにおいて公開
- 2004年4月以降の日本における医薬品による副作用報告
- 40万件を超える症例と、290万件におよぶ医薬品データが蓄積

症例一覧テーブル
demoCCYMM.csv

識別番号
報告回数
性別
年齢
体重
身長
報告年度・四半期
状況
報告の種類
報告者の資格

医薬品情報テーブル
drugCCYMM.csv

識別番号
報告回数
医薬品連番
医薬品の関与
医薬品(一般名)
医薬品(販売名)
経路
投与開始日
投与終了日
投与量
投与単位
分割投与回数
使用理由
医薬品の処置
再投与による再発の有無

2000種類超の
化学構造

副作用テーブル
reacCCYMM.csv

識別番号
報告回数
有害事象連番
有害事象
転帰
有害事象の発現日

71万件(8000副作用)

原疾患テーブル
histCCYMM.csv

識別番号
報告回数
原疾患等連番
原疾患等

90万件(8300疾患)

The screenshot shows the PMDA website interface. At the top, it says '独立行政法人 医薬品医療機器総合機構' and 'Pharmaceuticals and Medical Devices Agency'. There are navigation tabs for 'PMDAについて', '安全性情報', '添付文書等', '医療用医薬品', and '医療機器'. Below that, there are buttons for '一般の方向け', '医療従事者向け', 'アカデミア向け', and '企業向け'. A search bar is visible with the text 'サイト内検索' and a search button. At the bottom, there is a section titled '医薬品副作用データベース 利用規約' (JADER Database Terms of Use).

JADERによる副作用と薬物の解析

- ・解析用データテーブルから**2×2分割表**を網羅的に作成
- ・**報告オッズ比(ROR)**とFisherの正確検定の**P値**を算出

	各副作用	その他の副作用
各被疑薬	a	c
その他の医薬品	b	d

$$\text{ROR} = \frac{a/c}{b/d}$$

JADERによる副作用と薬物の解析

- ・解析用データテーブルから**2×2分割表**を網羅的に作成
- ・**報告オッズ比(ROR)**とFisherの正確検定の**P値**を算出

	出血性胃潰瘍	その他の副作用
NSAIDs	360	1,172
その他の医薬品	20,484	886,306

$$\text{ROR} = \frac{a/c}{b/d} = 13.3$$

JADERによる副作用と薬物の解析

- ・解析用データテーブルから2×2分割表を網羅的に作成
- ・報告オッズ比(ROR)とFisherの正確検定のP値を算出

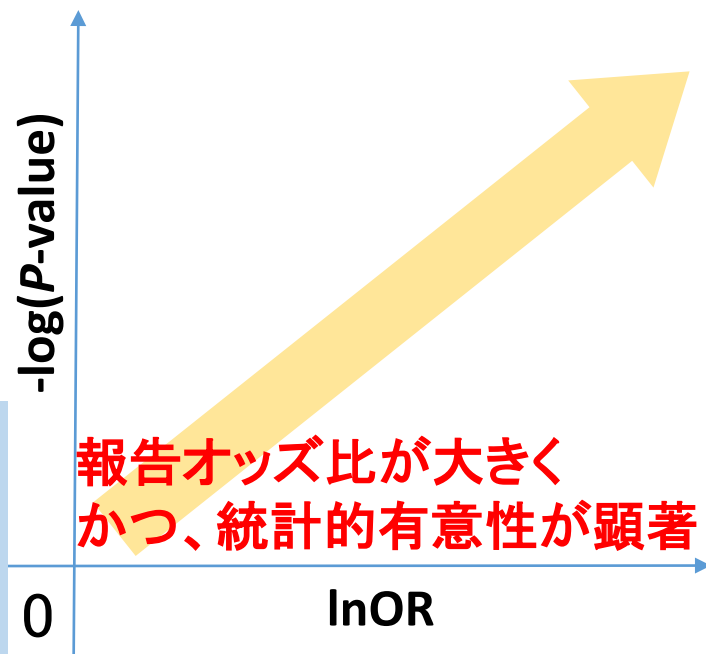
	各副作用	その他の副作用
各被疑薬	a	c
その他の医薬品	b	d

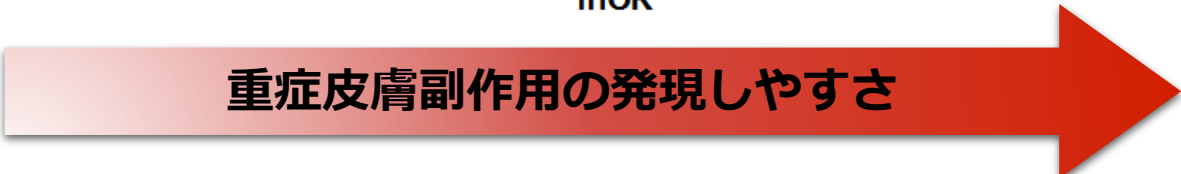
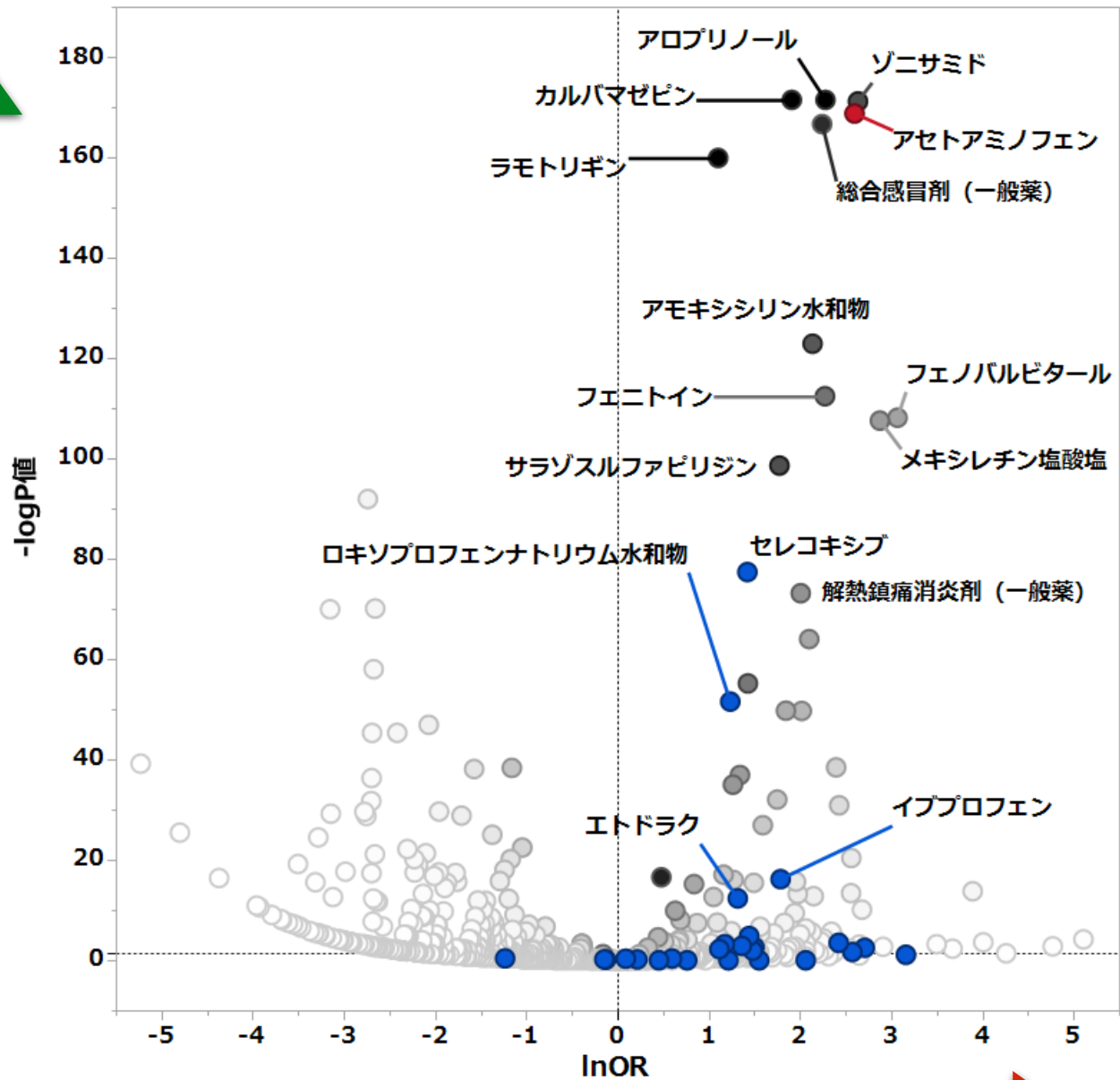
$$\text{ROR} = \frac{a/c}{b/d}$$

Volcano plot の作成

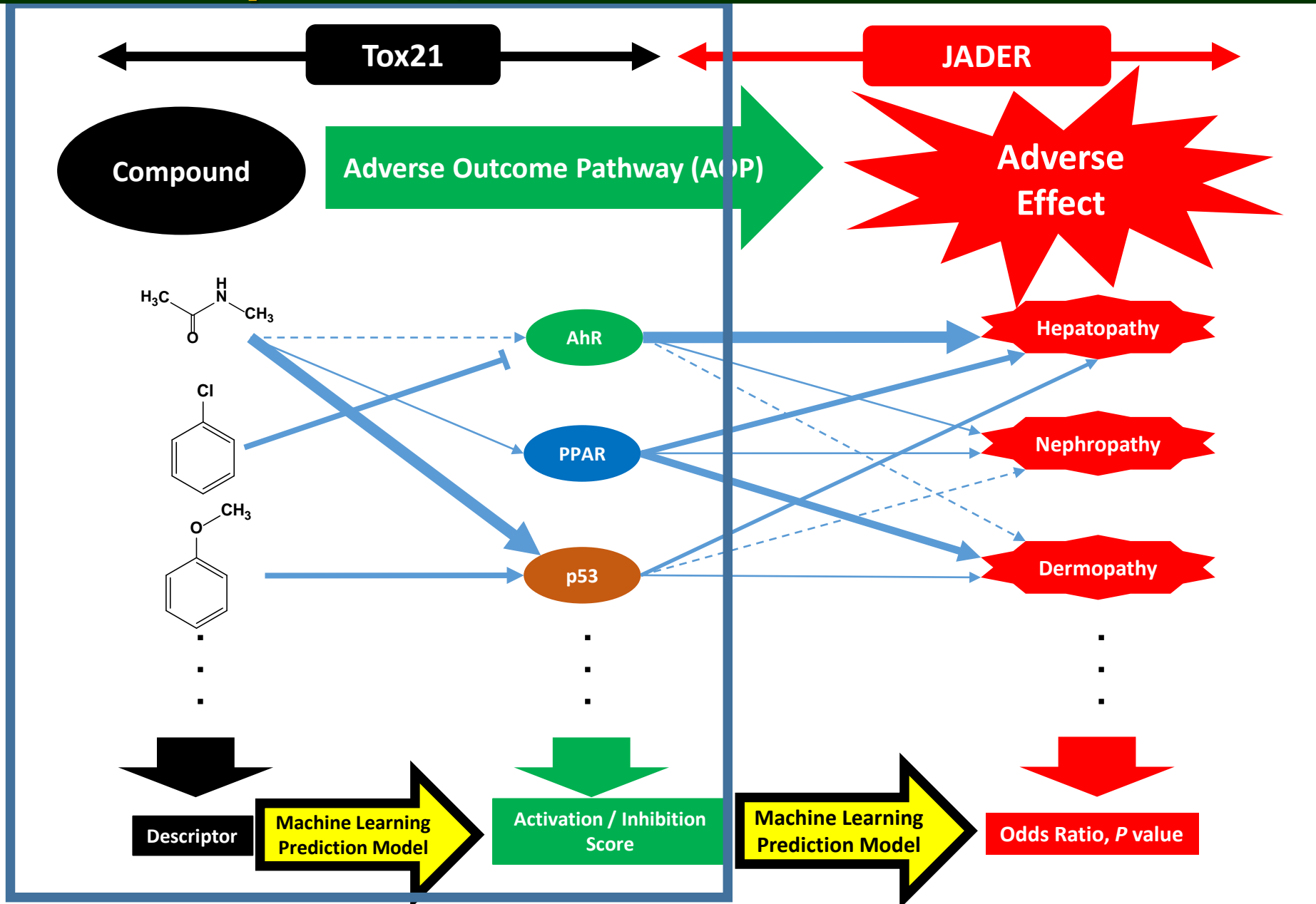
- x軸 : 報告オッズ比の自然対数 (lnOR)
- y軸 : P値の逆数の常用対数 (-log P値)

- ・副作用と関連の深い医薬品の観察法
- ・データベース内の副作用・医薬品間の関係を二次元表面上に圧縮表示





Relationship among Compounds, AOPs and Adverse Effects



NIH-Tox21 DATA Challenge 2014 (有害性発現経路予測)



Tox21 Data Challenge 2014

Login

» Home

About

Registration

Data/Resources

Submissions

Discussion

Leaderboard **NEW**

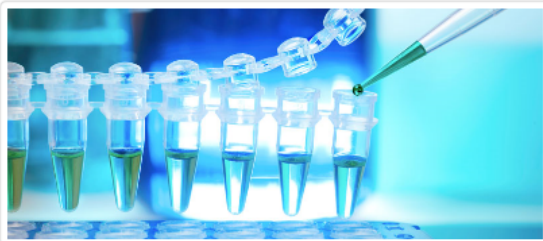
Contact Us



Open call ends: November 14, 2014



About the Data 3



NCATS will provide assay activity data and chemical structures on the Tox21 collection of ~10,000 compounds (Tox21 10K). A collection of compounds

The Challenge

The 2014 *Tox21* data challenge is designed to help scientists understand the potential of the chemicals and compounds being tested through the *Toxicology in the 21st Century* initiative to disrupt biological pathways in ways that may result in toxic effects.

The goal of the challenge is to "crowdsource" data analysis by independent researchers to reveal how well they can predict compounds' interference in biochemical pathways using only chemical structure data. The computational models produced from the Challenge could become decision-making tools for government agencies in determining which environmental chemicals



All challenge winners will receive the opportunity to submit a paper for publication in a special thematic issue of *Frontiers in Environmental Science* and recognition on the NCATS website and via social media.

Key Dates 0

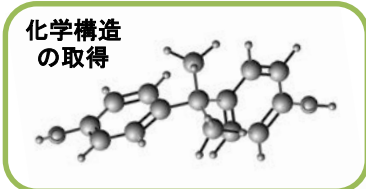
August 18, 2014
NCATS begins accepting submissions

Tox21 DATA Challenge 2014の 概要

- **Tox21の目標:**
 - ◆ 計算毒性学「化学構造からの毒性関連活性の予測」
- 現時点でどの程度の予測が可能であるか？
- Tox21 10K化合物ライブラリーを使用した毒性予測コンペティションを開催
- NIH国立先進トランスレーショナル科学センター (NCATS)が担当

主催者側

Tox21 10K ライブラリー
多様な化学物質
10,000種類



ウルトラHTS
ロボットを用いた
超高速アッセイによる
核内受容体・ストレス
応答パスウェイに対
する活性化能の評価

数値化
計算

構造記述子
分子量、脂溶性、
原子/結合数、その他

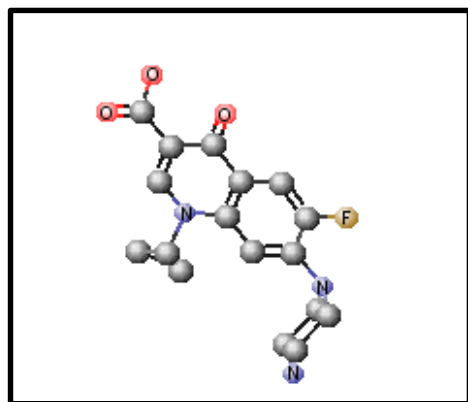
多変量解析
機械学習

参加者側

予測モデル
多数の化合物の構造と生理活性の間
に成立するようなルールを導き出す

定量的構造活性相關

Quantitative Structure–Activity Relationship (QSAR)



Experiment

Biological Activity

- Inhibitory effect
- Binding affinity
- EC50
- IC50
- etc...

Computational
calculation

Prediction

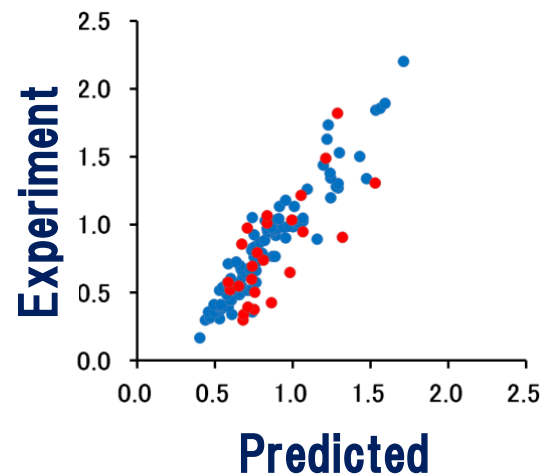
Descriptor
記述子

Structural/physicochemical properties

- Molecular weight
- Lipophilicity
- Fragment
- etc...

- Statistical analysis
- Machine learning

Prediction model



Tox21 Data Challenge 2014における 予測ターゲット

Tox21では、広範な毒性評価のためのターゲットとして種々の毒性発現経路(AOP)を採用した。


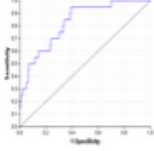
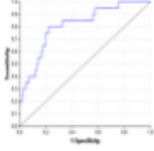
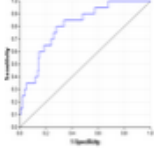
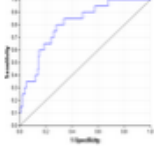
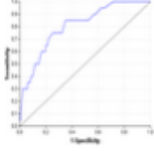
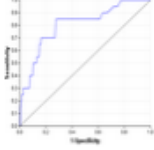

核内受容体関連パスウェイ

- ① エストロゲン受容体 α リガンド結合部位 (ER,LBD)
- ② エストロゲン受容体 α 前配列 (ER,full)
- ③ アロマターゼ (aromatase)
- ④ 芳香族炭化水素受容体 (AhR)
- ⑤ アンドロゲン受容体全配列 (AR,full)
- ⑥ アンドロゲン受容体リガンド結合部位 (AR,LBD)
- ⑦ ペルオキシソーム増殖因子活性化受容体 γ (PPAR-gamma)

ストレス応答パスウェイ

- ① 転写因子Nrf2/アンチオキシダント応答配列 (Nrf2/ARE)
- ② 熱ショック因子応答配列 (HSE)
- ③ ATPaseファミリー、AAADメイン含有5 (ATAD5)
- ④ ミトコンドリア膜電位 (MMP)
- ⑤ p53

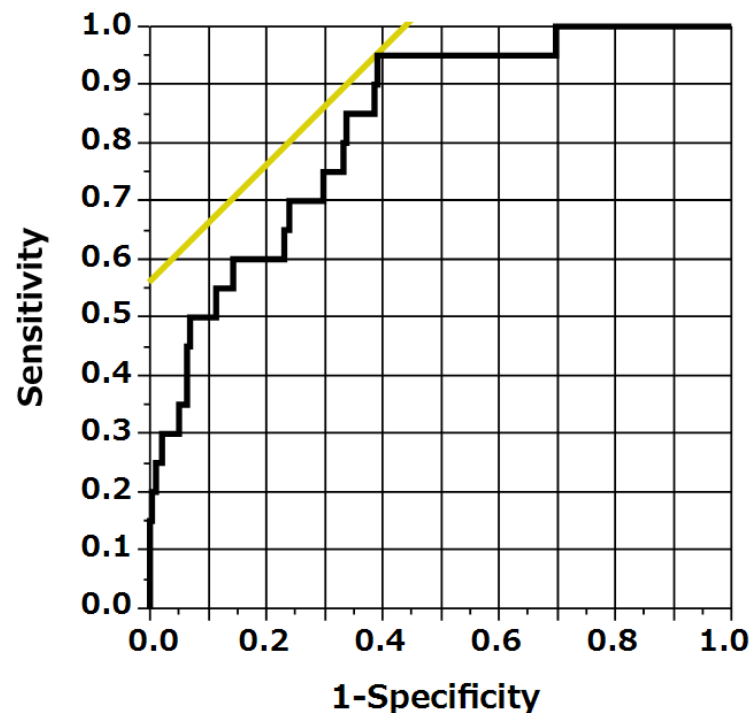
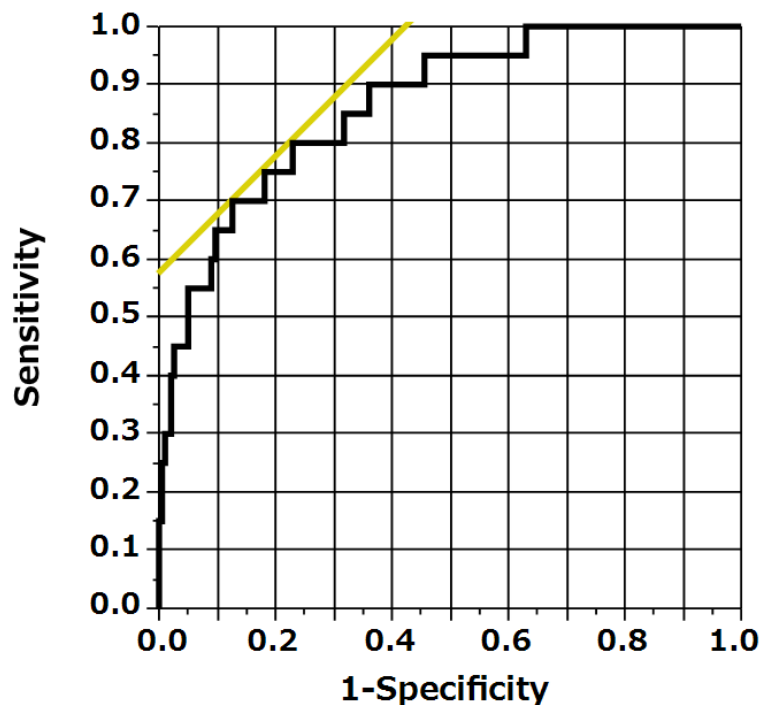


Team Name	Rank	AUC-ROC	Curve
 microsomes	1	0.82732759	
Bioinf@JKU	2	0.81439655	
Bioinf@JKU-ensemble4	3	0.80737069	
Bioinf@JKU-ensemble3	4	0.80737069	
AMAZIZ	5	0.8062931	
T	6	0.80508621	
Bioinf@JKU-ensemble2	7	0.80107759	

新規ER活性化物質予測モデルの精度

New model in the present study

Best model in Tox21 Challenge



$\text{ROC_AUC}_{\text{Evaluation}} = 0.866$

$\text{ROC_AUC}_{\text{Evaluation}} = 0.827$

Y Asako, Y Uesawa, High-Performance Prediction of Agonists on Human Estrogen Receptor Based on Chemical Structures, *Molecules* **22**(4) 1-10 (2017)

Y Uesawa, Rigorous selection of random forest models for identifying compounds that activate toxicity-related pathways, *Front Environ Sci* **4**(9) 1-6 (2016)

Winners

Organization	Team Name	Team Member	Challenge Assay
Johannes Kepler University Linz, Austria	Bioinf@JKU	G. Klambauer, et al.	AhR, Nrf2/ARE
	Bioinf@JKU-ensemble1	G. Klambauer, et al.	ER(full), HSE
	Bioinf@JKU-ensemble3	G. Klambauer, et al.	AR(LBD)
	Bioinf@JKU-ensemble4	G. Klambauer, et al.	PPAR-gamma
Technical University of Munich, Germany	AMAZIZ	A. Abdelaziz	ATAD5, MMP
Budapest University of Technology and Economics, Hungary	Dmlab	G. Barta	AR(full), Aromatase, p53
Meiji Pharmaceutical University, Japan	Microsomes	Y. Uesawa	ER(LBD)

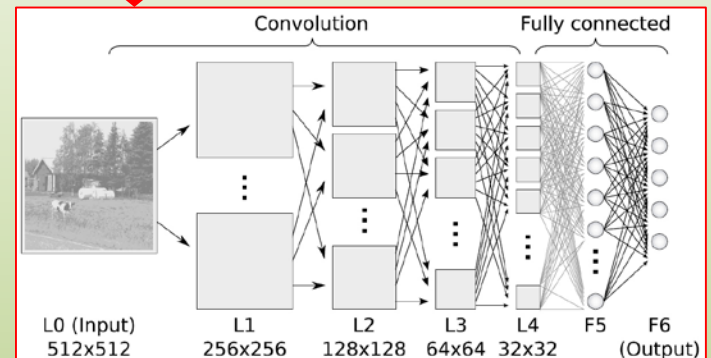
・リンツ・ヨハネス・ケプラー大学(オーストリア、総合優勝)

・ミュンヘン工科大学(ドイツ)

・ブダペスト工科経済大学(ハンガリー)

・明治薬科大学(日本)

深層学習 (Deep Learning) を使用



* Winnersにより構築された予測モデルの詳細は、Frontiers in Environmental Science誌の特集号参照

ディープラーニング（深層学習）

- 人工ニューラルネットワークの一種である。
- パターン認識で用いられてきた従来のニューラルネットワークは、入力層，中間層，出力層の3層程度から成る構成であった。
- 近年登場したディープラーニングは，中間層を重層化することによって学習を繰り返す結果，画像や音声認識の分野において極めて高い精度を達成できることが分かり話題となっている。
- 創薬分野においても，2012年にMerck社が行ったMerck Molecular Activity Challengeにおいてディープラーニングが最高精度を示したことで注目された。
- 欠点として，高度な計算環境を要求すること，および開発や調整が困難であることが挙げられる。

（薬学会・薬学用語解説2015.10.FYI）

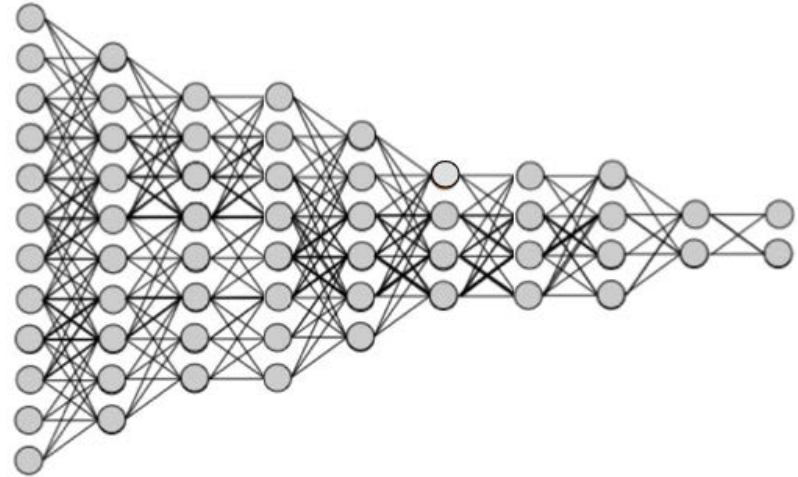
ディープラーニングの利点と欠点

- 利点

- ✓ 高い汎化性能
- ✓ 特徴表現学習

- 欠点

- ✓ 比較的高度な計算環境を要求
- ✓ ハイパーパラメータ等の調整が困難
- ✓ そのパフォーマンスを発揮するには大規模なサンプルが要求される



ディープラーニングの特徴：特徴表現学習

RF, SVM, NN

大規模データ

特徴量の抽出
記述子計算(人間)

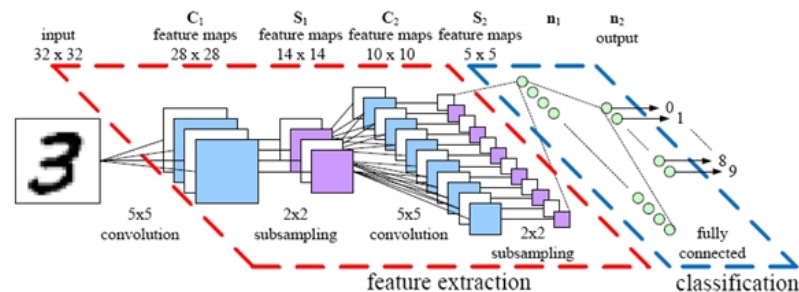
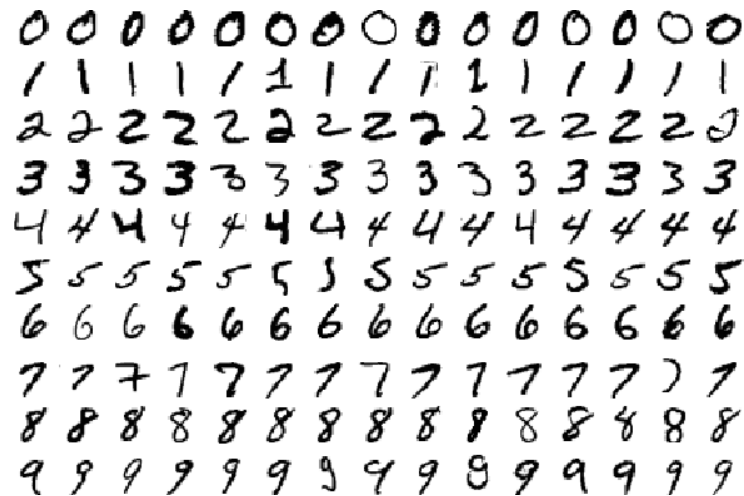
予測モデル

Deep Learning

大規模データ

特徴量の抽出
特徴表現学習

予測モデル



ディープラーニングによる名人芸のモデル化

熟練の研究者

大規模な経験

特徴量の抽出
化合物の雰囲気
(人間)

予測



Deep Learning

大規模データ

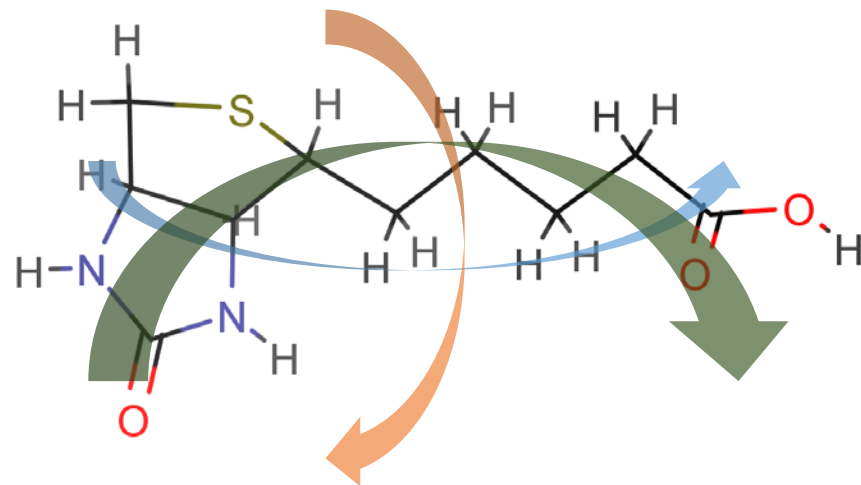
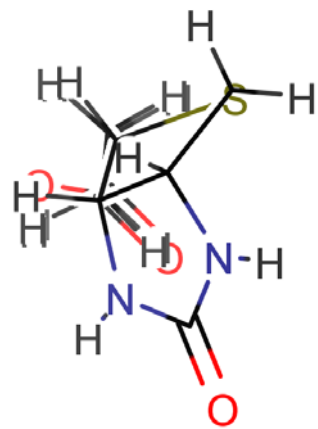
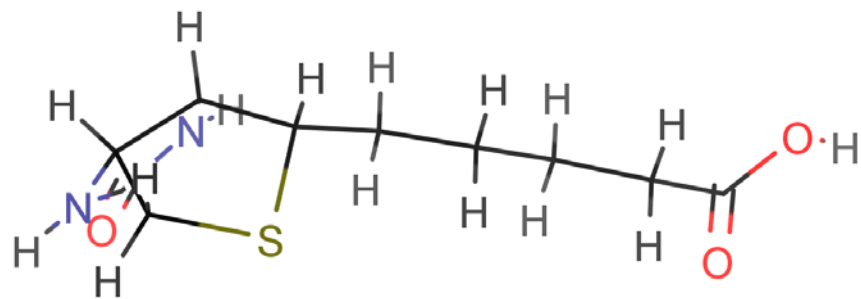
特徴量の抽出
特徴表現学習

予測モデル



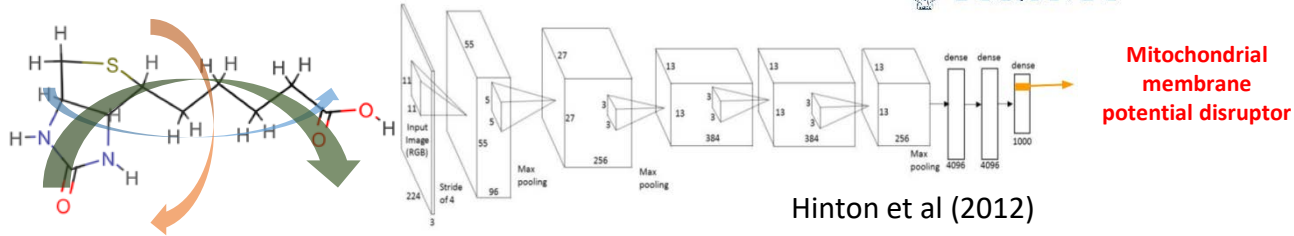
Deep
Learning

DeepSnapにおける 構造情報の入力

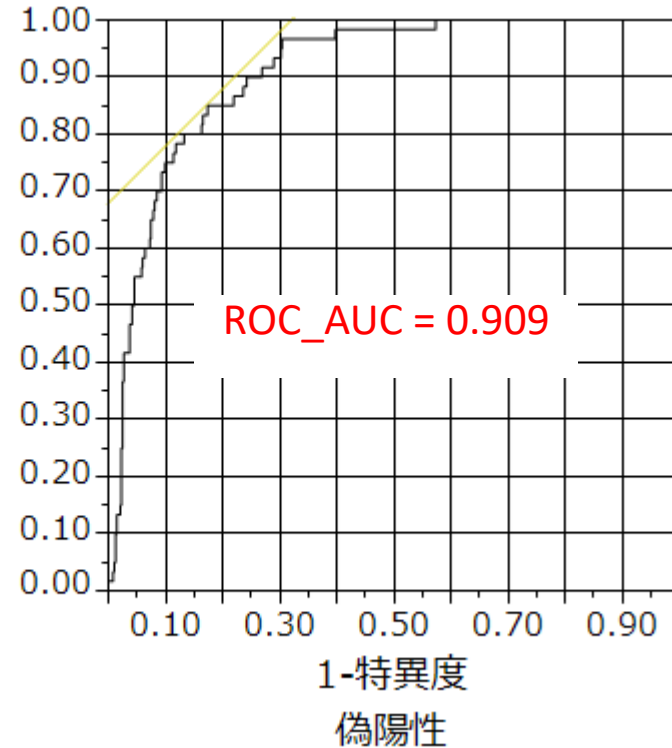
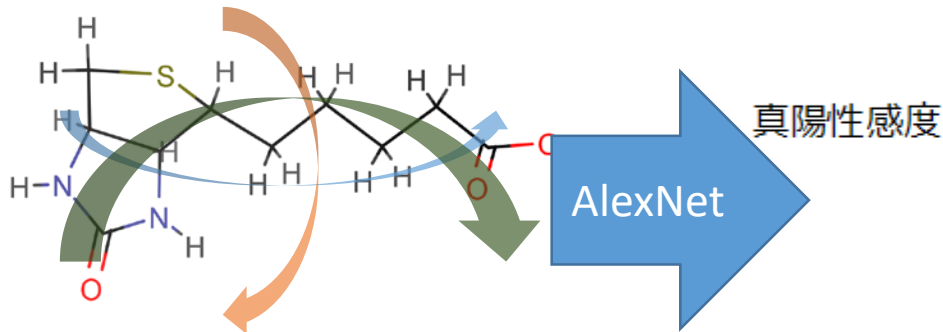


Deep Learningを用いた ミトコンドリア膜電位攪乱物質の予測

AlexNet



外部検証セット化合物のROC曲線



1万化合物を解析

Acknowledgments



This work was supported by

- The Long-Range Research Initiative (LRI) research program from the Japan Chemical Industry Association (JCIA).

A photograph of a stream flowing through a dense forest. The water is clear and white with foam as it cascades over large, moss-covered rocks. The surrounding trees and foliage are a vibrant green, creating a serene and natural atmosphere. The text "Thank you for your attention" is overlaid in a white, italicized serif font across the middle of the image.

Thank you for your attention