

生態毒性QSARシステムの解説 (生態毒性予測システムKATE)

(独)国立環境研究所 環境リスク研究センター

蓮沼 和夫

大分大学教育福祉科学部

吉岡 義正

2011年3月16日(水)東京会場

2011年3月17日(木)大阪会場

はじめに

QSARで得られた予測結果は、化審法の届出に必要な生態毒性試験結果として利用することは出来ません

QSARとは

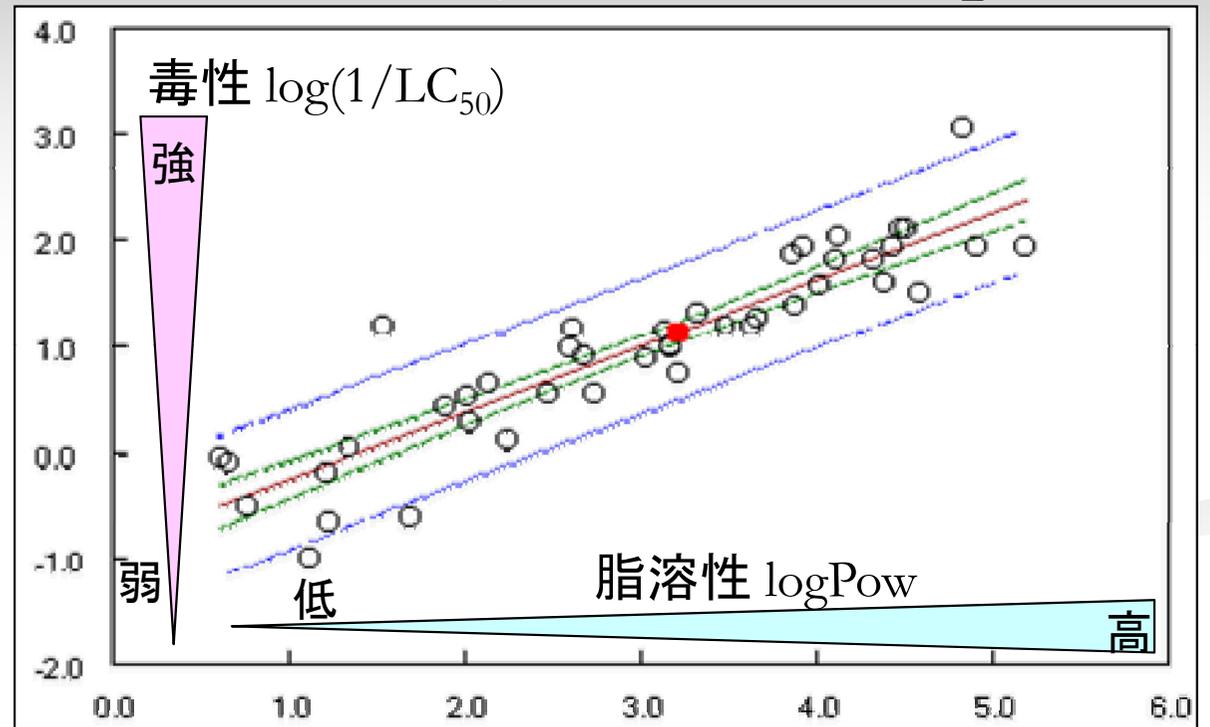
■ Quantitative Structure-Activity Relationship

■ 毒性

(生物学的活性)

魚類 半数致死量 LC_{50} : 7.8 mg/L

相関関係

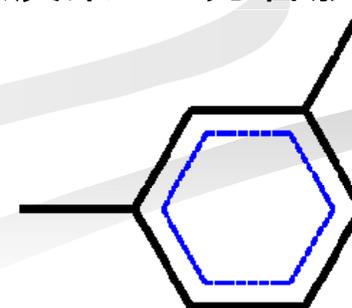


■ 化学物質の構造上の特徴

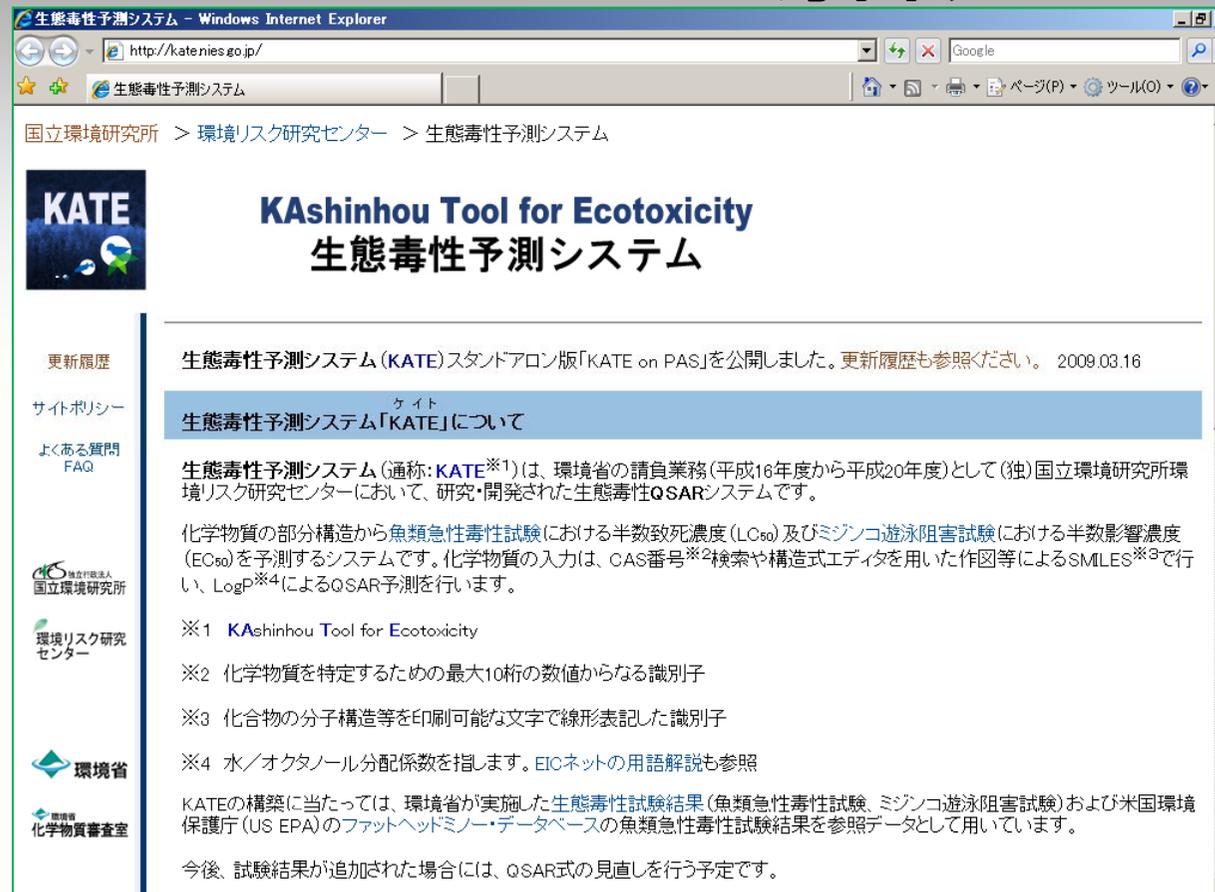
脂肪族炭素: 2 芳香族炭素: 6

■ 物理化学的パラメータ

脂溶解性の指標 $\log P_{ow}$ (オクタノール/水分配係数): 3.2



KATEの解説



- KATE [KAshinhou Tool for Ecotoxicity] とは
 - 化学物質の部分構造から毒性値を予測：
 - 魚類急性毒性試験における半数致死濃度 (LC_{50})
 - ミジンコ遊泳阻害試験における半数影響濃度 (EC_{50})

KATE開発の履歴

2011年3月

- スタンドアロン版「KATE on PAS 2011」とインターネット版「KATE on NET 2011」公開予定

2009年3月

インターネット版「KATE on NET」に加え
スタンドアロン版「KATE on PAS」を公開¹

2008年1月

試用版 (KATE Ver0.1) 公開

2007年7月

三省合同審議会※に対し、
予測結果の提供を開始

注)※: 薬事・食品衛生審議会
薬事分科会化学物質安全対策
部会化学物質調査会、化学物
質審議会審査部会、中央環境
審議会環境保健部会化学物質
審査小委員会

2004年

環境省の請負業務として研究・開発
開始(2004年度～2010年度)

KATEの予測方法

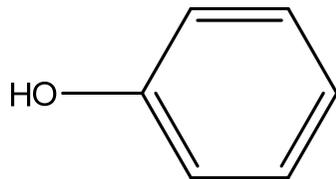
ユーザーの操作

化学物質の構造
(SMILES)の入力

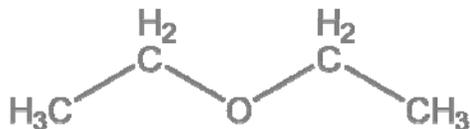
- SMILES: 化合物の分子構造を線形表記した識別子

- フェノール:

c1ccccc1O



- ジエチルエーテル: CCOCC



KATE内部での動作

クラス分類

QSAR 式を選択

参照物質データ

- ・環境省生態毒性試験結果
- ・U.S. EPA アットヘッドミラー データベース

logPow代入

- 実測値が望ましいが、無くても毒性予測可能
- ・スタンドアロン版: U.S. EPA KOWWIN
 - ・インターネット版: BioByte社 ClogP

予測毒性値を表示

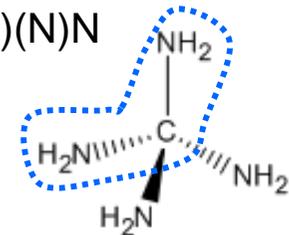
- 化学物質の構造に基づくクラス分類
- logPowの値から毒性値を予測

PASとは

- PAS (Platform for Assessment from Structure)^注は・・・
 - 構造分類に基づく物性や毒性を予測するための独自のシステム
 - 部分構造の取得プログラム (FITS; Fragment Identification by Tree Structure)、構造図の表示・入力プログラムなどからなる統合システム
- FITSは部分構造の規定に独自のルールを使用
 - 主体部分は、1次元構造を基本としたFITS記述です。
F/01211/**C=CNC=C**/1JnC=O,3V3,3B3,2Cy,3Cy,4Cy,2Rs4,/ |

例: NC(N)(N)Nの構造でNCNの構造の数を、目的に応じて1-6個まで定義できます。

SMILES: NC(N)(N)N

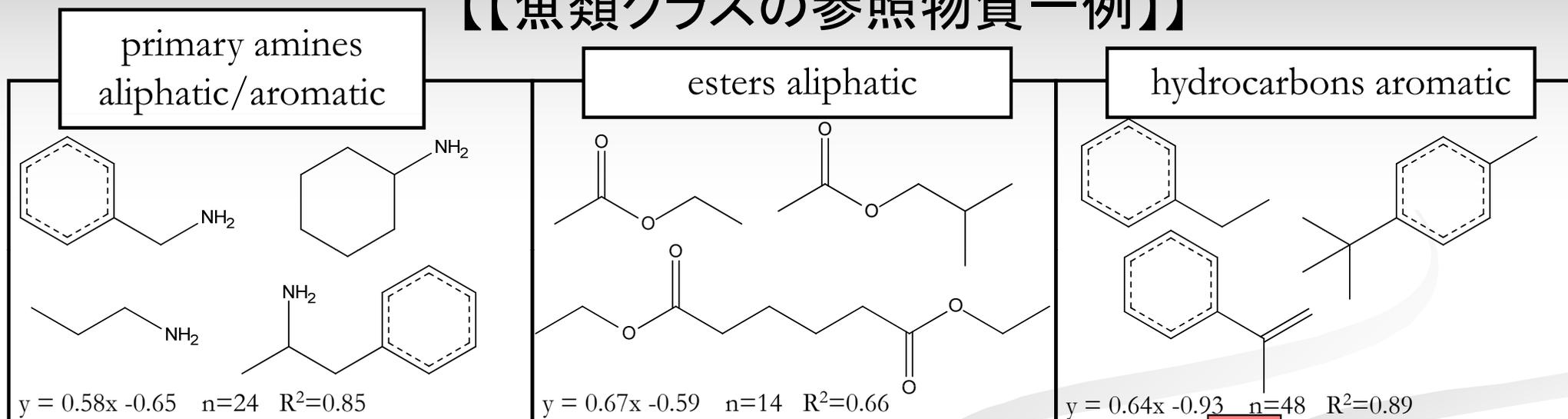


注: PASの開発は、2000~2002年度(H12~14)環境省環境研究総合促進費「環境中の複合化学物質による次世代影響リスクの評価とリスク支援に関する研究」の一環として大分大学で実施。また、「環境データの解析と環境中生物影響評価に関する研究」として、2005~2008年度(H17~20)には(独)国立環境研究所と大分大学との委託・共同研究として実施。

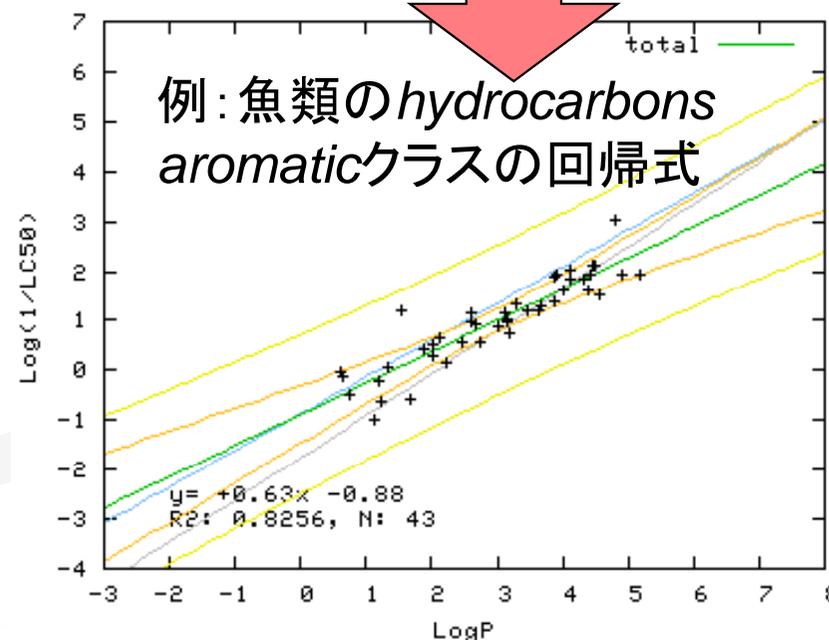
KATEのクラスについて

- 魚類・甲殻類合わせ約80種類のQSAR式・クラスが存在

【魚類クラスの参照物質一例】



	魚類	甲殻類
式の数(参照データ 2以上)	41	40
データ数が十分な式の数 (参照データ 5以上)	36	22
信頼できる式の数 (参照データ 5以上 かつR ² が0.7以上)	20	10



*Neutral Organics*クラスについて

- 脂肪族炭化水素、脂肪族・芳香族エーテル、脂肪族・芳香族ケトン、アルコールといった単純な麻酔作用のみで毒性が説明できると考えられるクラスの物質は、*Neutral Organics*というクラスとして再定義
- *Neutral Organics*に分類されるクラス：
 - alcohols aliphatic
 - ethers aliphatic
 - ethers with aromatic
 - hydrocabons aliphatic
 - nitriles aliphatic
 - phosphates

スタンドアロン版KATE on PAS 入力画面

KATE on PAS

ファイル (F) Help ユーザDB管理 (M)

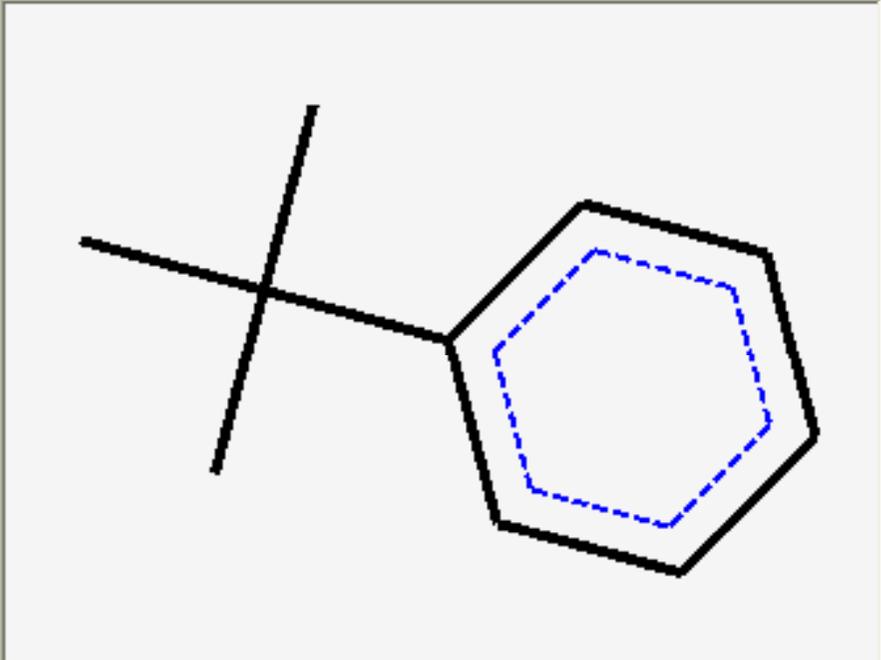
ID番号 CAS番号

SMILES ←SMILES

名称

logPow 溶解度 (mg/L)

logPow実測値(EPA)=4.11 (使用)
logPow推定値(EPA)=3.90



The image shows a screenshot of the KATE on PAS software interface. The window title is "KATE on PAS". The menu bar includes "ファイル (F)", "Help", and "ユーザDB管理 (M)". Below the menu bar, there are input fields for "ID番号" and "CAS番号", followed by buttons for "検索開始", "図表示", "拡大図", and "クリア". The "SMILES" field contains the string "CC(C)(C)c1ccccc1" with a red arrow pointing to it and the text "←SMILES". Below this is an empty "名称" field. Further down, there are buttons for "QSAR実行", "logPow", and "溶解度 (mg/L)", along with another "ユーザDBに登録" button and a "図のコピー" button. A yellow text area on the left displays "logPow実測値(EPA)=4.11 (使用)" and "logPow推定値(EPA)=3.90". On the right, a chemical structure of 1,1,1-trimethylbenzene is shown with a dashed blue outline.

KATE on PAS出力(概要)画面

QSAR予測結果

戻る (X) 拡大図 (B) 参照物質 (R)

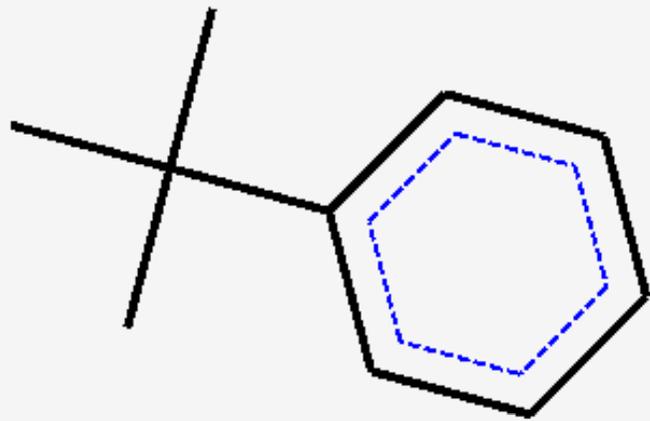
概要 詳細

分子量 134.22 使用logPow 4.11

溶解度

logPow実測値(EPA)=4.11 (使用)
logPow推定値(EPA)=3.90

KATEでは予め定義された部分構造の有無で、クラスの適用範囲を判断します。予測する化学物質のもつ部分構造のすべてが、[そのクラス]や[Neutral Organicクラス]の参照物質に含まれるかで、○、△、×に評価します。○、△の場合は[そのクラス]の適用範囲内と判断し



```

*****
----- KATE on PAS Estimation -----
SMILES      : CC(C)(C)c1ccccc1
名称等      :
組成式      : C10H14   分子量 :134.22
logP        : 4.110 (EPA 実測値)
----- KATE 生態毒性推定値 -----
推定値      logP範囲      判定
-----
クラス      : hydrocarbons aromatic
試験生物    : Fish
エンドポイント: 96-hr LC50    2.6 mg/L    0.600<***<5.170    ○
-----
クラス      : hydrocarbons aromatic
試験生物    : Daphnid
エンドポイント: 48-hr EC50    1.1 mg/L    0.650<***<5.170    ○
-----
=== 判定 ===
○ ; [このクラス]の参照物質にも含まれ、適用範囲内と判断されます。

=== 検出部分構造と数 ===
4909 脂肪族 C          4
4910 芳香族 原子      6

=== カテゴリー 部分構造 ===
5056 脂肪族C-CH3      3
5063 芳香族t-Bu       1
5201 benzene          1
5236 Aromatics        1
    
```

予測毒性値

判定

KATE on PAS出力(詳細)画面

参照物質(QSAR式作成の根拠となった
実測毒性値が入手可能な物質)一覧

QSAR予測結果

戻る (X) 拡大図 (B) 参照物質 (R)

概要 詳細

Clear Selection Fish 96-hr LC50 hydro

100414 : Benzene, ethyl- :1000628
 100696 : Pyridine, 2-ethyl- :1000135
 100710 : Pyridine, 2-ethyl- :1000235
 104518 : Benzene, butyl- :1000341
 106423 : Benzene, 1,4-dimethyl- :1000612
 108383 : Benzene, 1,3-dimethyl- :1000490
 108883 : Benzene, methyl- :1000385
 108894 : Pyridine, 4-methyl- :1000239
 108996 : Pyridine, 3-methyl- :1000552
 109068 : Pyridine, 2-methyl- :1000020
 109977 : 1H-Pyrrole :1000417
 110009 : Furan :1000120
 110021 : Thiophene :1000916

SMILES : CC(C)(C)c1ccccc1
 名称等 :
 組成式 : C10H14 分子量 :134.22
 logP : 4.110 (EPA 実測値)

----- KATE on PAS Estimation -----

	推定値	logP範囲	判定
クラス : hydrocarbons aromatic			
試験生物 : Fish			
エンドポイント: 96-hr LC50	2.6 mg/L	0.600<***<5.170	○

回帰式: $\log(1/C, \text{mmol}) = -0.883 + 0.630 * \text{LogP}$ n=43 R2=0.826

=== 判定 ===
 ○ ; [このクラス]の参照物質にも含まれ、適用範囲内と判断されます。

=== クラス選択理由 ===

設定	検出
芳香族 原子	>0, 6
脂肪族 N NO	=0, 0
脂肪族 O	=0, 0
脂肪族 S	=0, 0
脂肪族 P	=0, 0
脂肪族 carbon-halogen	=0, 0
脂肪族 NO2	=0, 0
芳香族 n	=0, 0
芳香族 o	=0, 0
芳香族 s	=0, 0
芳香族 carbon-halogen	=0, 0

=== 検出部分構造と数 ===

4909 脂肪族 C	4
4910 芳香族 原子	6

=== カテゴリー 部分構造 ===

5056 脂肪族C-CH3	3
5063 芳香族c-t-Bu	1
5201 benzene	1
5236 Aromatics	1

Toxicity Determined $\log(1/[\text{mmol/L}])$

logPow

緑線=回帰式の信頼限界(90%両側), 青線=予測値の信頼限界(90%両側)

予測結果の適用範囲について(判定)

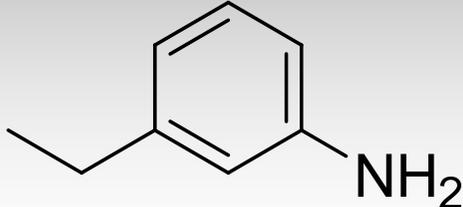
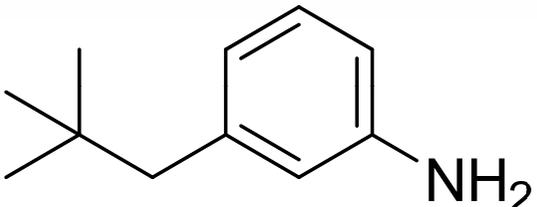
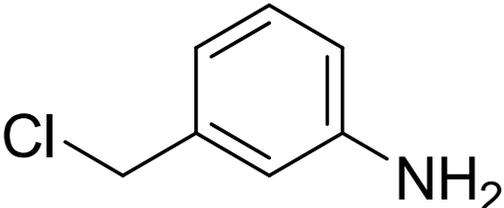


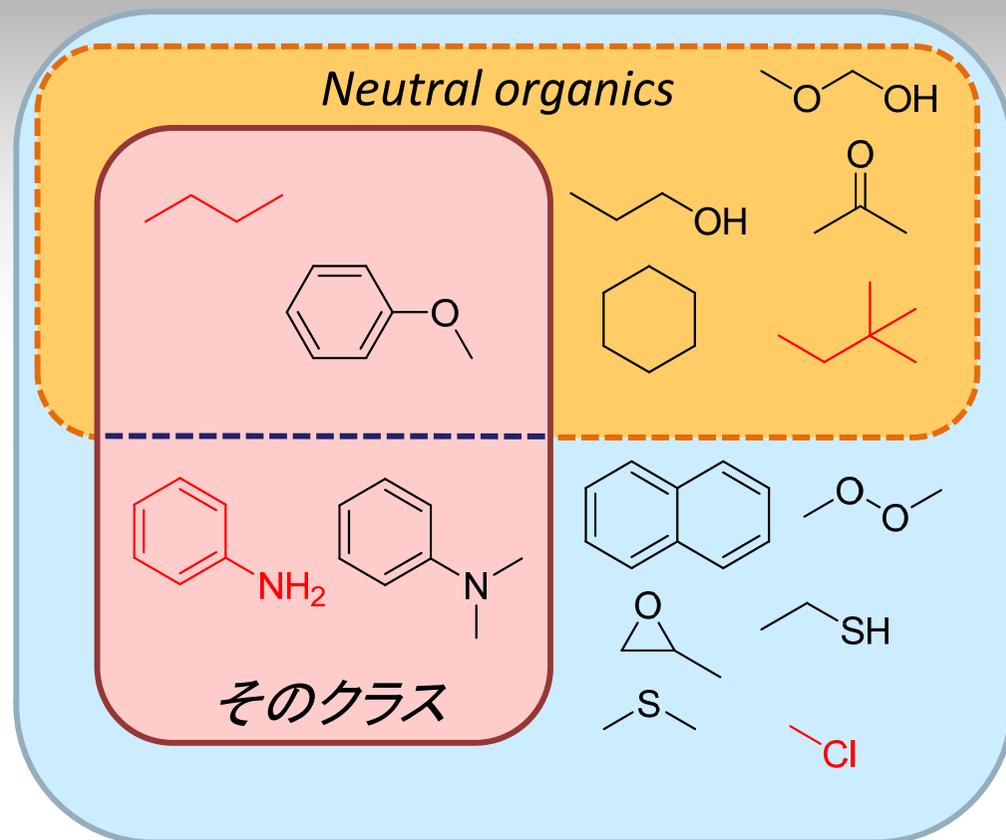
- 構造C判定: 予測する化学物質のもつ部分構造すべてが、
 - ○: [そのクラス]の参照物質にも含まれる。
 - △: [そのクラス]または[Neutral Organicクラス]の参照物質にも含まれる。
 - ×: [そのクラス]や[Neutral Organicクラス]の参照物質には含まれない部分構造がある。

として評価される。

- logP判定: 予測した物質のlogPowがQSAR回帰式の有効範囲内に入っているか(内挿であるか)で評価される。
 - スタンドアロン版: 有効範囲外の場合、『>P』又は『<P』と評価
 - インターネット版: 有効範囲内の場合『○』、範囲外は『×』と評価

予測結果の適用範囲について(判定)

	C(1)	C(2)
	YES 構造C判定 : ○	YES
	NO 構造C判定 : △	YES
	NO 構造C判定 : ×	NO



構造C 判定

C(1): 予測した化学物質の全ての部分構造が、「そのクラス」の参照物質に含まれるか。

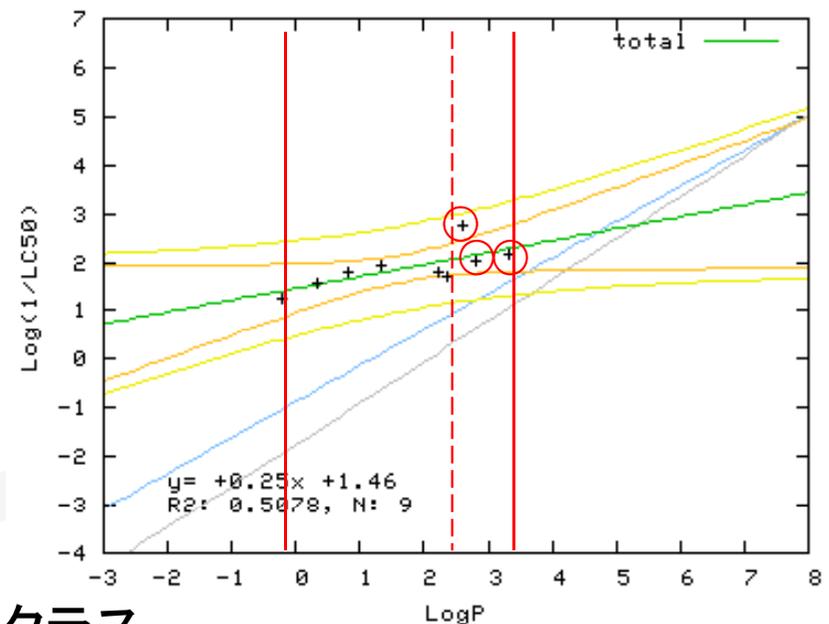
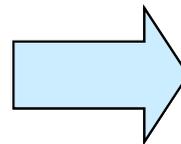
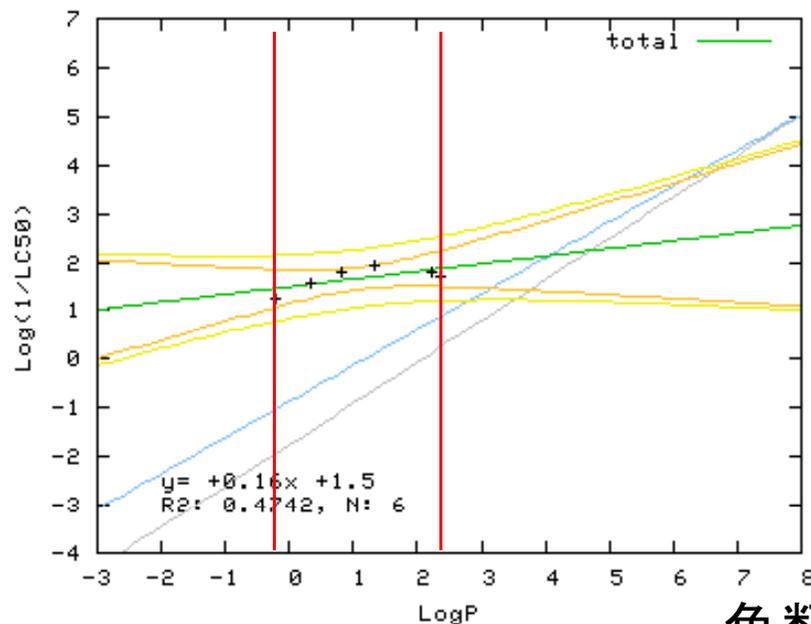
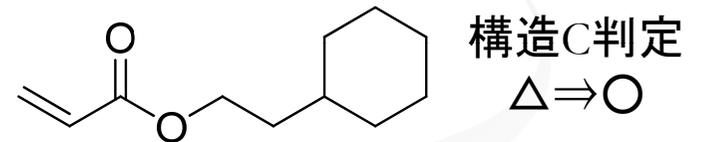
C(2): 予測した化学物質の全ての部分構造が、「そのクラス」又は *Neutral Organics* の参照物質に含まれるか。

2011年3月版での改良点

■ 参照物質の追加①

	参照物質数	
	魚類	甲殻類
2009年3月版	539	260
2011年3月版	599	321

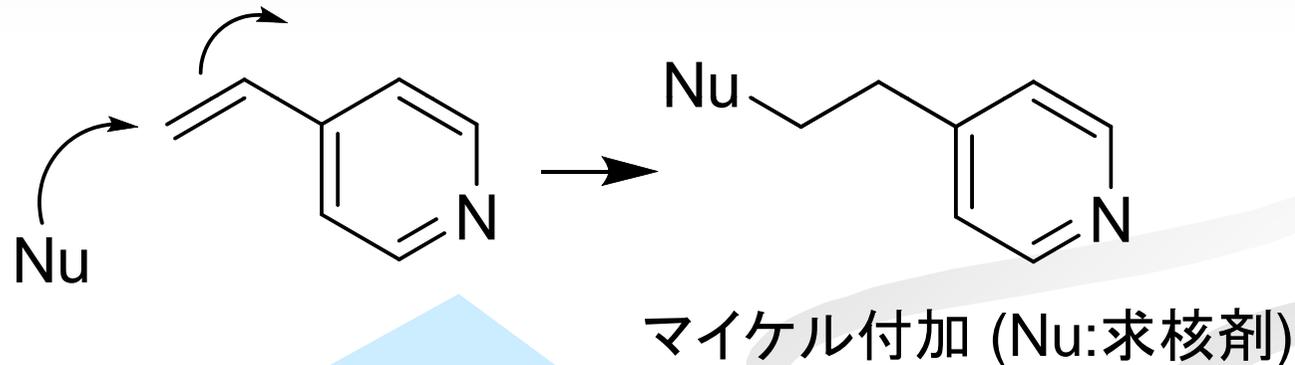
- 予測精度の向上
- logPow 範囲の拡大
- 予測可能な構造の拡大



魚類 *acrylates*クラス

2011年3月版での改良点

- フラグメント・分類ルールの変更②
- 構造C判定の変更③
 - 皮膚感作性の反応性²に関する部分構造を追加³



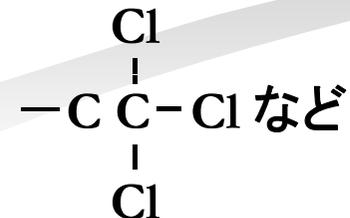
4 - ビニルピリジン

魚類LC₅₀: 1 mg/L(実測値)

120mg/L(予測値、*hydrocarbons aromatic*クラス)

構造C判定: ○⇒× (2009年3月版⇒2011年3月版)

- 精度向上のための部分構造を追加



ご静聴ありがとうございました

参考文献

- 1 A. Furuhamma, T. Toida, N. Nishikawa, Y. Aoki, Y. Yoshioka, H. Shiraishi, Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system, March 2009 version, SAR QSAR Environ. Res., 21 (2010), pp. 403-413.
- 2 S.J. Enoch, J.C. Madden, and M.T.D. Cronin, Identification of mechanisms of toxic action for skin sensitisation using a SMARTS pattern based approach, SAR QSAR Environ. Res., 19 (2008), pp. 555-578.
- 3 A. Furuhamma, K. Hasunuma, Y. Aoki, Y. Yoshioka and H. Shiraishi, Application of chemical reaction mechanistic domains to an ecotoxicity QSAR model, KAshinhou Tool for Ecotoxicity (KATE), SAR QSAR Environ. Res., (2011), *in press*.