KATE on NET 2011
生態毒性予測システム「KATE2011」インターネット版
操作マニュアル 2011.03.31

※　KATE は、化学物質の生態毒性（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}））を予測するシステムです。

※　本システム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。

＊ご質問等がございましたら、下記までお問い合わせ下さい＊
国立研究開発法人　国立環境研究所　環境リスク・健康研究センター　kate@nies.go.jp

Copyright(C) 2008-2011 Ministry of the Environment, Government of Japan.
All Rights Reserved

マニュアル改訂履歴

バージョン	発行日	改訂履歴
第 1 版	2011 年 3 月 31 日	初版発行
第 1.1 版	2016 年 2 月 2 日	リンク URL の変更、予測区間に関する説明と図の追加・変更
第 1.2 版	2016 年 4 月 18 日	名称変更「環境リスク研究センター」→「環境リスク・健康研究センター」

はじめに

生態毒性予測システム「KATE」：Kashinho Tool for Ecotoxicity とは

環境省の請負業務（平成 16 年度から平成 22 年度）として（独）国立環境研究所環境リスク研究センターにおいて、研究・開発された生態毒性 QSAR システムです。2008 年 1 月に試用版、2009 年 3 月に KATE2009 が公開されました。*1

化学物質の部分構造から魚類急性毒性試験における半数致死濃度（LC50）*2 及びミジンコ遊泳阻害試験における半数影響濃度（EC50）*2 を予測するシステムです。化学物質の入力は、CAS 番号検索や構造式エディタを用いた作図等による SMILES*3 入力行い、LogP*4 による QSAR 予測を行います。

KATE の構築に当たっては、環境省が実施した生態毒性試験結果（魚類急性毒性試験、ミジンコ遊泳阻害試験）*5 及び米国環境保護庁（US EPA）のファットヘッドミノー・データベースの魚類急性毒性試験結果*6 を参照データとして用いています。今後、試験結果が追加された場合には QSAR 式の見直しを行う予定です。

「KATE」の公開形体について

KATE にはパソコンにインストールして使用するスタンドアロン版「KATE on PAS」<http://kate.nies.go.jp/onpas.html>とインターネット上のブラウザ画面で操作するインターネット版「KATE on NET」<http://kate.nies.go.jp/onnet.html>があります。なお、本システムの一部は大分大学との共同研究の成果および Daylight 社製のシステム*7 とその結果を使用しております。

「KATE on NET」の留意事項

KATE on NET（インターネット版）は、入力が同じであれば KATE on PAS（スタンドアロン版）と同じ予測結果を出力します。ユーザによる LogP 入力を省略すると、予測結果の相違が生じます。

「KATE2009 以降」の更新内容

■　実測 LogP 値による毒性予測を基本とします。化学物質の入力画面に LogP 入力の欄も追加されました。ユーザーが LogP 値を指定しない場合は、Daylight 社*7 から提供される CLogP プログラムを使用して、BioByte 社*8 のアルゴリズムにより算出した CLogP 結果を使用します。なお、ライセンスの契約上、ClogP 値は提供できません。（注意：「KATE on PAS」では US EPA で開発された KOWWIN による計算 LogP 値*9 を用いています。）

■　構造分類のプログラムを変更し、KATE on PAS と同じ部分構造の取得プログラム「FITS；Fragment Identification by Tree Structure」を導入しました。

■　環境省の生態毒性試験結果を追加し、QSAR モデルの見直しを行いました。

■　その他 KATE のホームページも参照ください。

FITS を用いたシステムでの SMILES の入出力について

- FITS には SMILES の正規化のための独自ルーチンが組み込まれており、入力された SMILES は正規化処理されますが、表示には入力した SMILES が出力されます。
- ナトリウム Na、カリウム K は水素 H に置換して予測しないと動作しません。
- イオンは窒素イオン N^+、リンイオン P^+以外は対応していません。
- 原子番号 55 以上の原子は対応していません。
- 異常原子価の構造はうけつけません。共鳴構造を考慮にいれて SMILES を修正ください。
  →Daylight 社の SMILES ツールキットで内部変換された SMILES を用いるとエラーになることがあります。具体的な SMILES の例は 7.Appendix(2)（p.21）を参照ください。

(Q)SAR： (Quantitative) Structure-Activity Relationship、（定量的）構造活性相関とは

化学物質の構造上の特徴又は物理化学定数と生物学的活性（毒性等）の相関関係を構造活性相関（SAR）といい、定量的なものを定量的構造活性相関（QSAR）といいます。両者を併せて(Q)SAR と記載することもあります。構造活性相関は、例えば、特定の官能基の有無から物質の有害性の多寡を推測することを指し、構造を手掛かりに毒性等を定量的に算出する仕組みをいわゆる QSAR モデルと呼びます。

海外で開発されている生態毒性 QSAR モデルとしては、米国環境保護庁（US EPA）の ECOSAR[*10] などがあります。

KATE 2011 の更新履歴

（変更１）バージョン管理

KATE の更新は年号によるバージョン管理を実施します。

（変更２）データの追加

2009 年 3 月版公開以降に報告された参照物質および実測 log P データを追加しました。また、ClogP を用いた QSAR 式の種類の更新も同時に行いました。

（変更３）ＱＳＡＲモデルの更新

化学物質の部分構造の分類ルールおよび部分構造の切り出しの定義を更新しました。更に、構造

C 判定に皮膚感作性の反応性に関する部分構造の一部修正と精度向上のための部分構造を追加しました。

皮膚感作性の部分構造の引用文献：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.

生態毒性への適用の評価に関する文献： A. Furuhama, 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*.

その他

本システムで得られた予測結果は、十分な予測精度を保証するものではありません。化学物質の生態毒性影響の程度について、参考値を得るためのツールの一つとしてご利用ください。また、「化学物質の審査及び製造等の規制に関する法律（化審法）」に基づく届出*11に必要な生態毒性試験結果として利用することはできません。 著作権、リンク等については、サイトポリシーのページをご覧ください。

参考 URLs

*1 http://kate.nies.go.jp/

論文情報 A. Furuhama, T. Toida, N. Nishikawa, Y. Aoki, Y. Yoshioka, and H. Shiraishi : Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system, March 2009 version : SAR QSAR Environ. Res., Volume 21 Issue 5, 403 (2010).

*2 http://www.env.go.jp/chemi/kagaku/hourei/tuuchi/mat12.pdf

*3 化合物の分子構造等を印刷可能な文字で線形表記した識別子。Daylight 社のホームページも参照ください。http://www.daylight.com/smiles/index.html

*4 水／オクタノール分配係数を指します。EIC ネットの用語解説も参照ください。
http://www.eic.or.jp/ecoterm/?act=view&serial=295

*5 http://www.env.go.jp/chemi/sesaku/seitai.html

*6 http://archive.epa.gov/med/med_archive_03/web/html/fathead_minnow.html

*7 http://www.daylight.com/

*8 http://www.biobyte.com/index.html

*9 http://www.epa.gov/tsca-screening-tools/epi-suitetm-estimation-program-interface

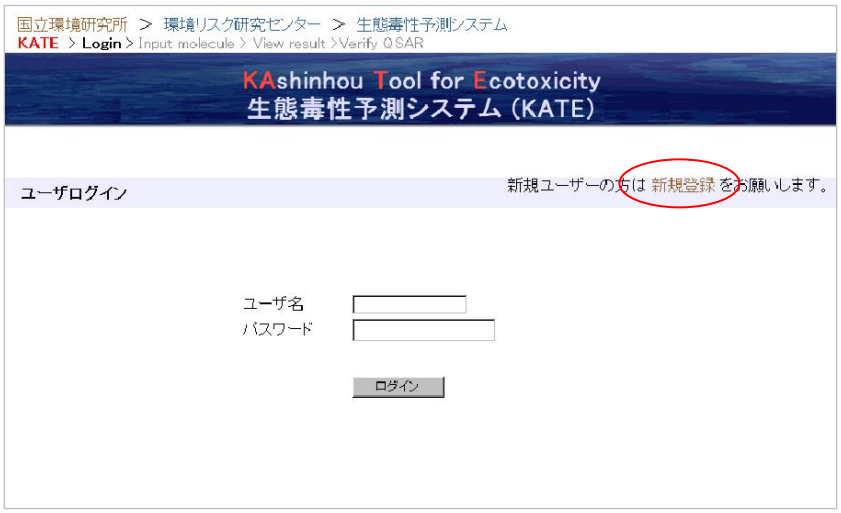
*10
http://www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-ecosar-predictive-model

*11 http://www.env.go.jp/chemi/info/guide01.html

1.　新規ユーザ登録　https://kate.nies.go.jp/nies/user_regist_input.php

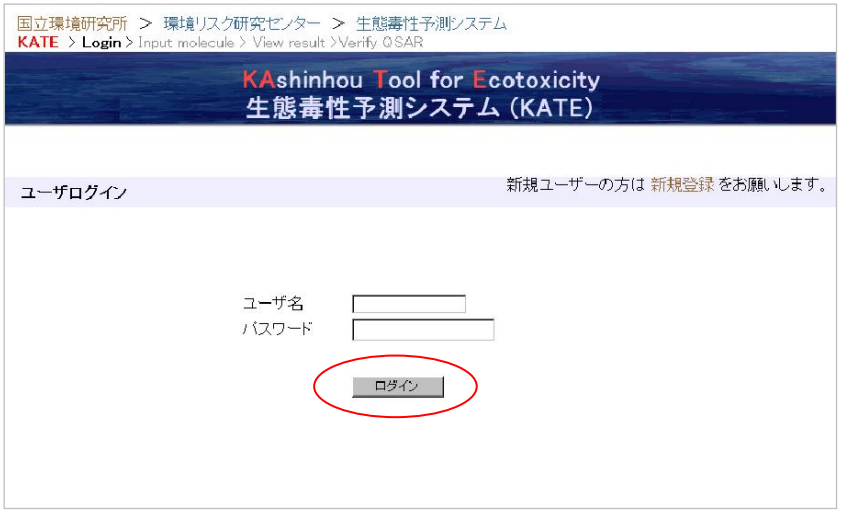
初めて使用する場合は、ユーザ名及びパスワードの登録を行ってください。ユーザ名は 5 文字以上必要です。また、既に登録されているユーザ名は登録することはできません。

なお、ユーザ名またはパスワードを忘れた場合も新規登録を行って下さい。長期間使用されていないユーザ名は定期的に削除されます。



2.　ログイン　https://kate.nies.go.jp/nies/index.php

1.で登録したユーザ名及びパスワードを入力して、ログインしてください。



＊パスワードの変更も可能です。

3.　化学物質の入力

(1) 1 つの化合物の予測

化学物質は、SMILES で入力してください。SMILES とは、構造式を文字列で表現するフォーマットです。Daylight 社のウェブページに詳しい解説があります。

< http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>

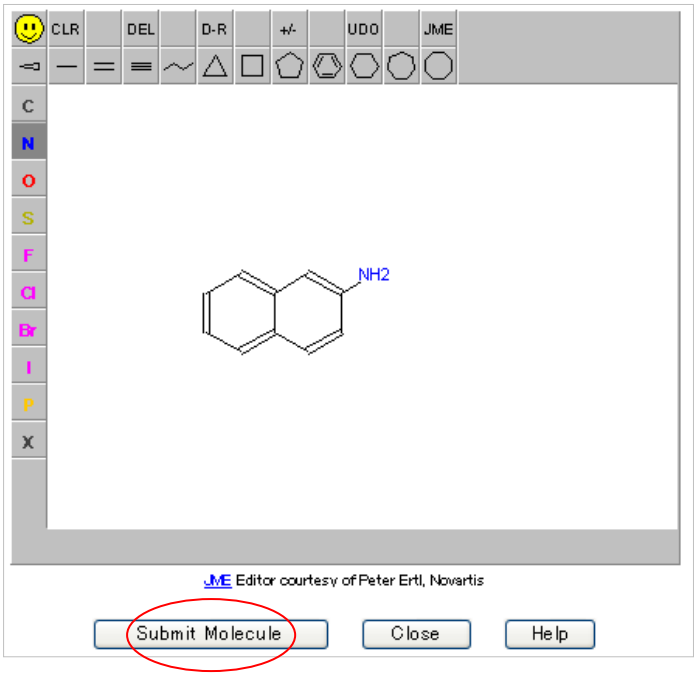
1 構造式エディタを用いて構造式を描くと SMILES に変換が出来ます。また、2 化学物質の CAS 番号、3 WebKis-Plus（化学物質 DB）を使用して、物質を選択すると、名称、CAS 番号の入力及び SMILES への変換をすることが出来ます。

なお、米国環境保護庁（US EPA）が無償で提供する EPI Suite を使用して、CAS 番号から SMILES へ変換することも出来ます。EPI Suite の結果を用いる場合には、お手数ですが、EPI Suite で表示された SMILES をコピーして KATE の SMILES 入力の欄に貼り付けてください。

補足情報は入力すると予測結果のページに表示されます。必須項目ではありません。

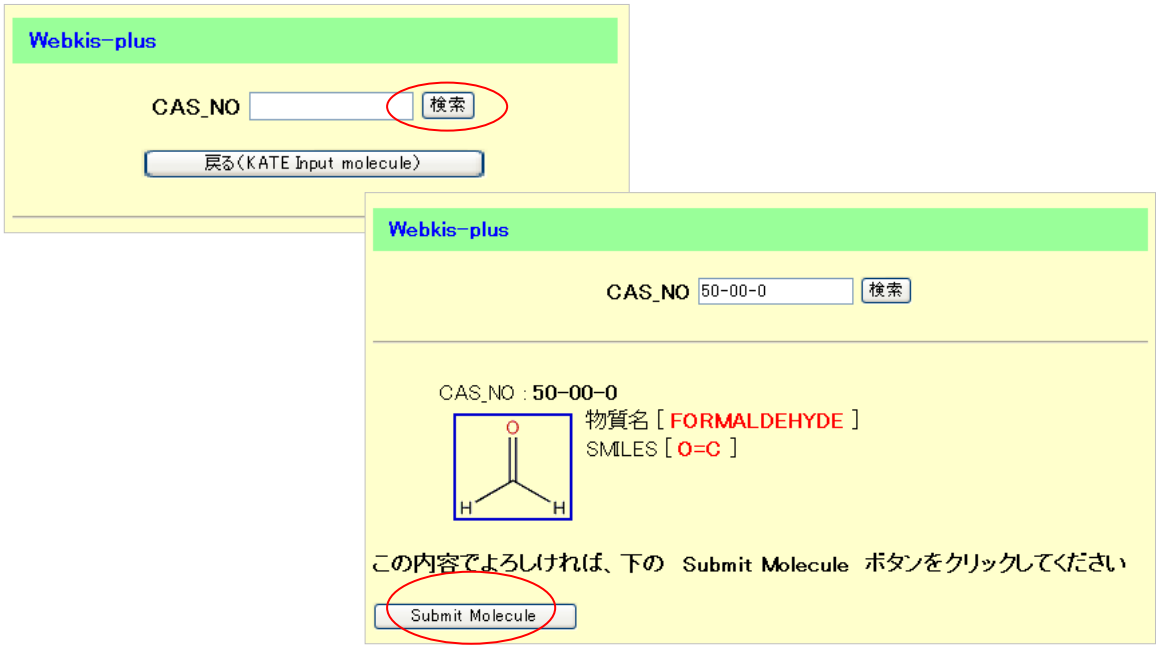
① 構造式エディタ　※構造式エディタには Java Runtime Plug-in が必要です。

構造式を描き、「Submit Molecule」をクリックします。



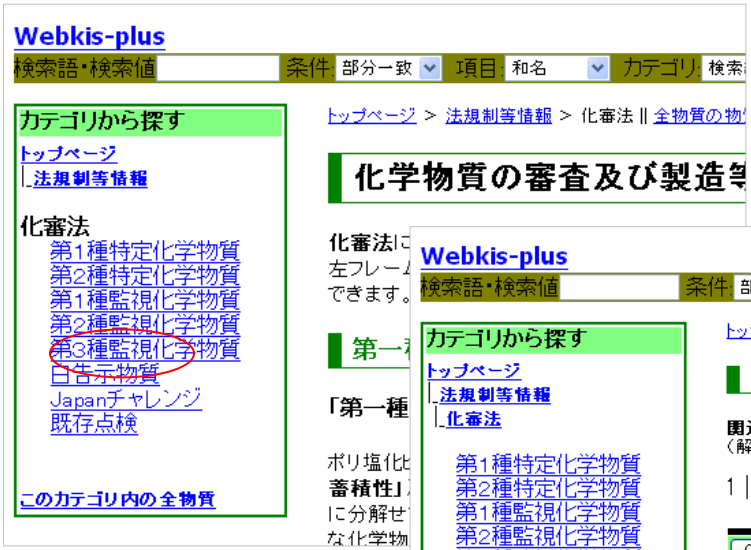
② CAS 番号検索

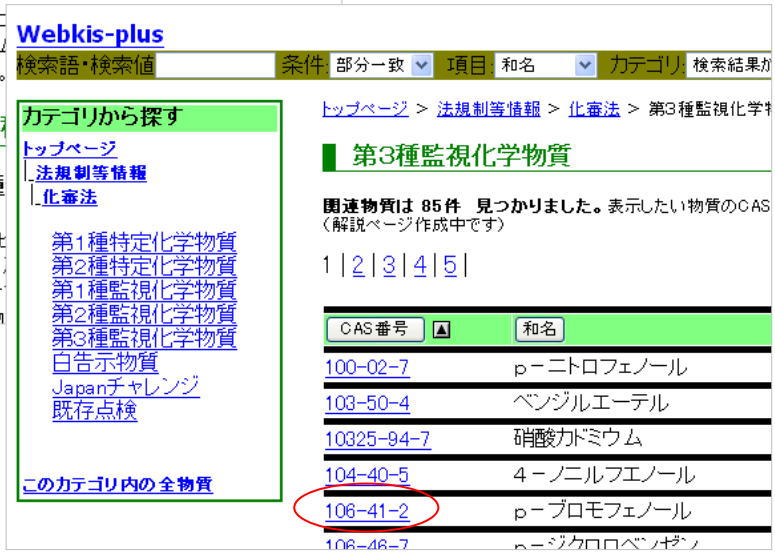
CAS 番号を入力し、「検索」をクリックすると物質が表示されます。構造式等を確認し、「Submit Molecule」をクリックします。

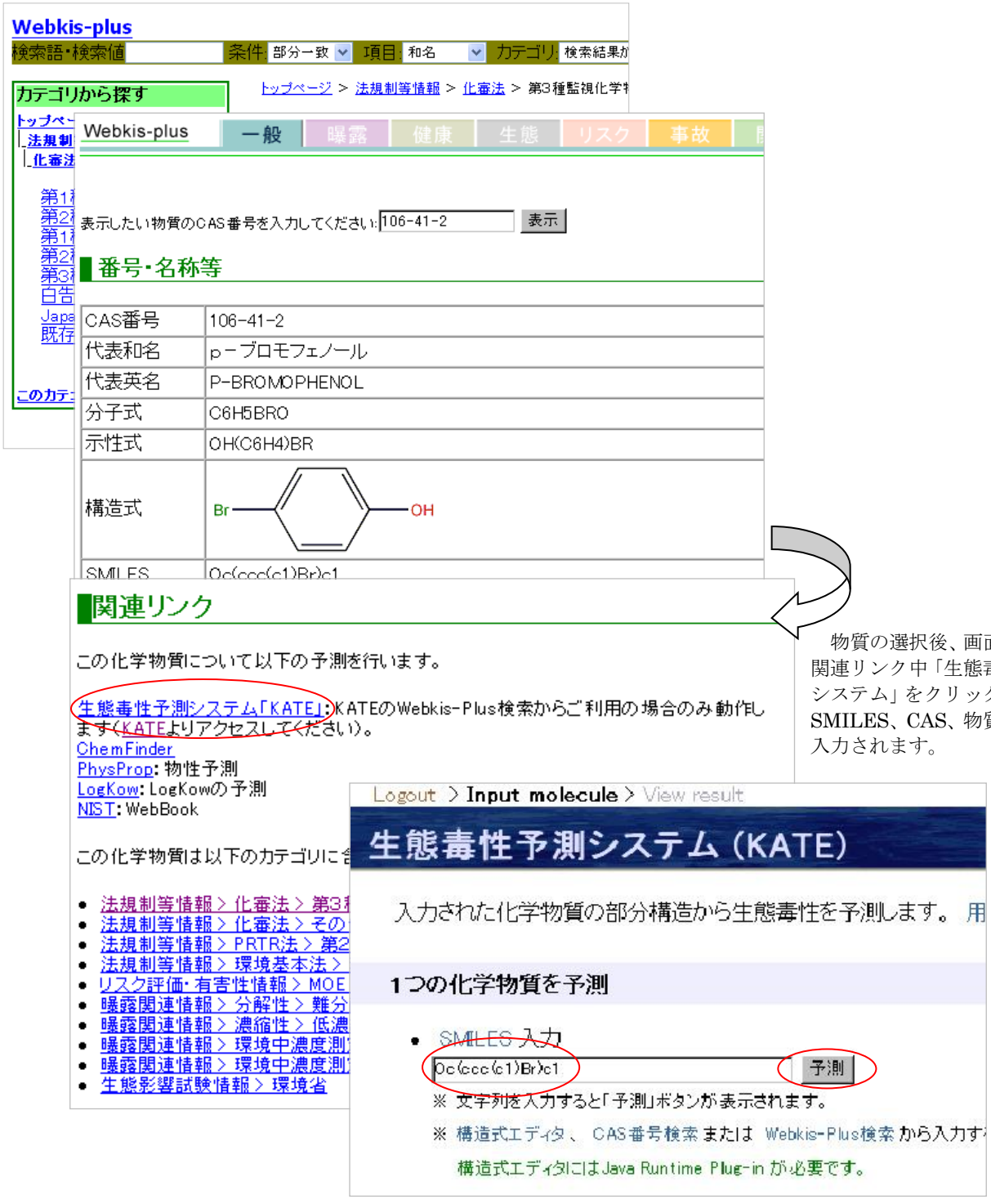


③　Webkis-plus

まず、物質を選択します。







物質の選択後、画面下部の関連リンク中「生態毒性予測システム」をクリックするとSMILES、CAS、物質名称が入力されます。

LogP

水/オクタノール分配係数(Octanol Water Partition Coefficient)を LogP と表示します。現バージョンの「KATE」では LogP による単回帰での毒性予測を行い、実測 LogP 値の入力を推奨します。インターネット版「KATE on NET」では LogP 値が入力されない場合は Daylight 社のプログラムを用いて、BioByte 社の提供するデータベースの logP 実測値(LogPstar)が存在すればその値を優先し、実測値が登録されていない場合は入力された SMILES に基づき BioByte 社のアルゴリズムで計算された ClogP 値での予測を行います。なお、ライセンスの契約上、ClogP 値は提供できません。

(2)複数の化学物質を予測

SMILES リスト

KATE では複数の SMILES を連続して処理するために一行に 1 物質のデータの連続からなる独自のファイルフォーマット「SMILES リスト」を利用します。

SMILES リスト例 (1)

10047-28-6　　CCCCOC(=O)CS
5954-68-7　　CC(=C)CS
28467-71-2　　CC1(CC2(C)CC3(Br)C1)CC(Br)(C2)C3
693-58-3　　CCCCCCCCCBr
21912-23-2　　CCCCCCCCC(Br)CBr
112-29-8　　CCCCCCCCCCBr

※順序は CAS 番号 *Tab* SMILES *改行*の順番で指定されます。

SMILES リスト例 (2)

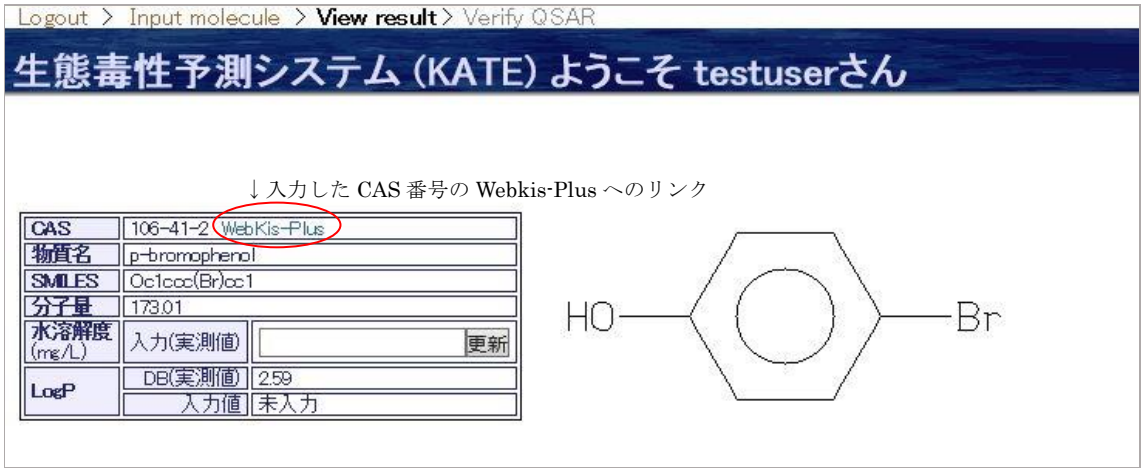
chemical1　　-0.8　　CCCCOC(=O)CS
chemical2　　1.2　　CC(=C)CS
chemical3　　1.3　　CC1(CC2(C)CC3(Br)C1)CC(Br)(C2)C3
chemical4　　1.4　　CCCCCCCCCBr
chemical5　　4.1　　CCCCCCCCC(Br)CBr
chemical6　　4.25　　CCCCCCCCCCBr

※順序は名称 *Tab* LogP 値 *Tab* SMILES *改行*の順番で指定されます。

※KATE の入力に関する用語について　https://kate.nies.go.jp/nies/glossary.php も参照ください。

4.　結果の表示

ページの上部には入力した CAS 番号、物質名、SMILES が表示されます。ルールに従った CAS 番号であれば WebKis-Plus へのリンクが表示されます。分子量は入力した情報から計算した値が表示されます。入力した LogP と DB の LogP 値は表示されますが、KATE 内部で計算した ClogP による計算値はライセンスの契約上、提供できません。



中段には、化学物質のクラス分類、Fish（魚類）、Daphnid（ミジンコ類）の予測結果等が表示されます。参照物質数が十分でなく、QSAR 式が存在しない場合は予測結果欄に「Unavailable」と表示されます。

予測結果欄の右には、信頼水準 95%の予測区間（回帰式の信頼区間）の欄が表示されます。

なお、LogP の欄にはユーザが入力した LogP 値による予測では「LogP 入力値」、ユーザが LogP 値の指定をせず、BioByte 社の提供するデータベースの logP 実測値(LogPstar)が存在した場合は「logP 実測値(LogPstar)」、ClogP 値を用いた場合は、「ClogP 予測値」と表示されます。

QSAR結果

クラス※1	エンドポイント			LogP	予測結果(mg/L)	予測区間(mg/L)(信頼水準 95%)	判定		
							構造C※2	LogP※3	溶解度※4
ethers aliphatic	Fish	96hr	LC50	CLogP予測値	0.57	[0.18, 1.8]	○	×	未入力
Neutral Organics	Fish	96hr	LC50	CLogP予測値	0.11	[0.018, 0.71]	○	○	未入力
ethers aliphatic	Daphnid	48hr	EC50	CLogP予測値	Unavailable	[-, -]	○	×	未入力
Neutral Organics	Daphnid	48hr	EC50	CLogP予測値	0.14	[0.016, 1.2]	○	×	未入力

※1 化学物質が複数のクラスに分類される場合があります。判定の項目が○のクラスを優先してください。
すべて○の場合、Neutral Organics でないクラスを採用するか、Verify QSARの画面で参照物質やQSAR式の決定係数などを参考に判断して下さい。

複数のクラスに分類される場合には、全てのクラスにおける結果が表示されます。判定の項目（構造 C、LogP）に表示される○の数が多いクラスを優先し、△よりも○を優先してください。但し、Neurtal Organics と Neurtal Organics 以外のクラスで○の数が同じ場合には、Neurtal Organics 以外のクラスを採用するか、Verify QSAR の画面で参照物質や QSAR 式の決定係数などを参考に判断してください。

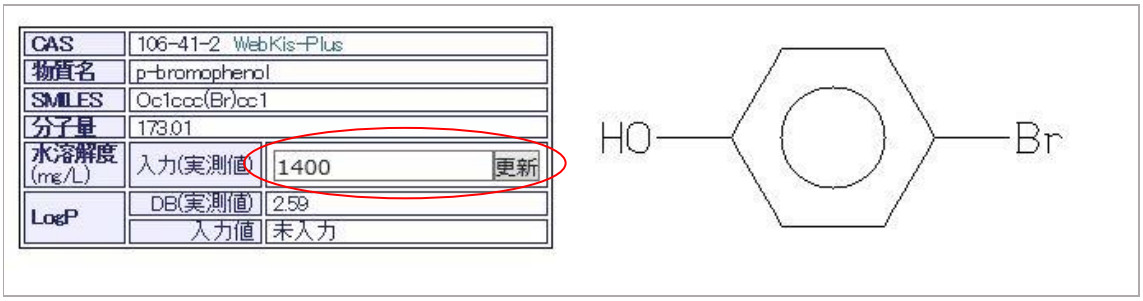
QSAR結果

クラス※1	エンドポイント			LogP	予測結果(mg/L)	予測区間(mg/L)(信頼水準 95%)	判定		
							構造C※2	LogP※3	溶解度※4
ethers aliphatic	Fish	96hr	LC50	CLogP予測値	0.57	[0.18, 1.8]	○	×	未入力
Neutral Organics	Fish	96hr	LC50	CLogP予測値	0.11	[0.018, 0.71]	○	○	未入力
ethers aliphatic	Daphnid	48hr	EC50	CLogP予測値	Unavailable	[-, -]	○	×	未入力
Neutral Organics	Daphnid	48hr	EC50	CLogP予測値	0.14	[0.016, 1.2]	○	×	未入力

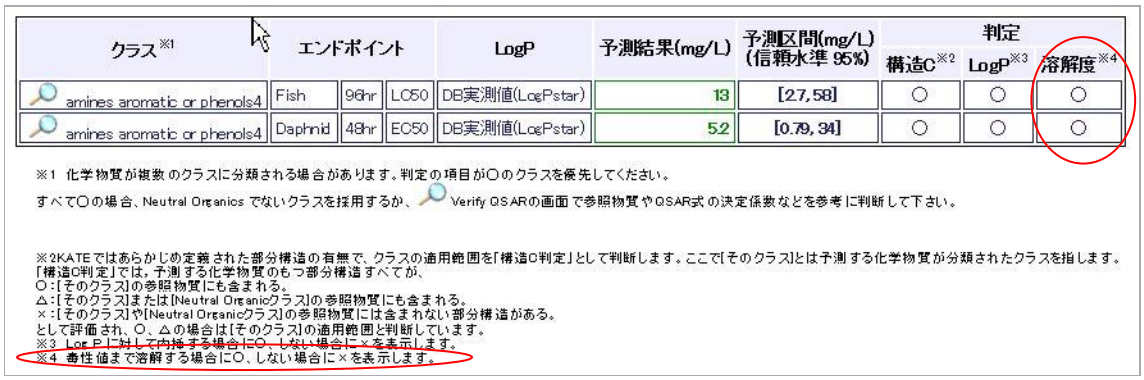
※1 化学物質が複数のクラスに分類される場合があります。判定の項目が○のクラスを優先してください。
すべて○の場合、Neutral Organics でないクラスを採用するか、Verify QSARの画面で参照物質やQSAR式の決定係数などを参考に判断して下さい。

↑の○が多いほど信頼性が高い予測結果になります。

14 ページの予測結果画面で水溶解度を入力して更新ボタンをクリックすると、予測された毒性値と水溶解度との関係が○、×で表示されます。水溶解度は半角で入力して下さい。



CAS	106-41-2 WebKis-Plus	
物質名	p-bromophenol	
SMILES	Oc1ccc(Br)cc1	
分子量	173.01	
水溶解度 (mg/L)	入力(実測値)	1400 更新
LogP	DB(実測値)	2.59
	入力値	未入力



クラス※1	エンドポイント			LogP	予測結果(mg/L)	予測区間(mg/L)(信頼水準 95%)	判定		
							構造C※2	LogP※3	溶解度※4
amines aromatic or phenols4	Fish	96hr	LC50	DB実測値(LogPstar)	13	[2.7, 58]	○	○	○
amines aromatic or phenols4	Daphnid	48hr	EC50	DB実測値(LogPstar)	5.2	[0.79, 34]	○	○	○

※1 化学物質が複数のクラスに分類される場合があります。判定の項目が○のクラスを優先してください。
すべて○の場合、Neutral Organics でないクラスを採用するか、Verify QSARの画面で参照物質やQSAR式の決定係数などを参考に判断して下さい。

※2KATEではあらかじめ定義された部分構造の有無で、クラスの適用範囲を「構造C判定」として判断します。ここで[そのクラス]とは予測する化学物質が分類されたクラスを指します。
「構造C判定」では，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、○、△の場合は[そのクラス]の適用範囲と判断しています。
※3 Log P に対して内挿する場合に○、しない場合に×を表示します。
※4 毒性値まで溶解する場合に○、しない場合に×を表示します。

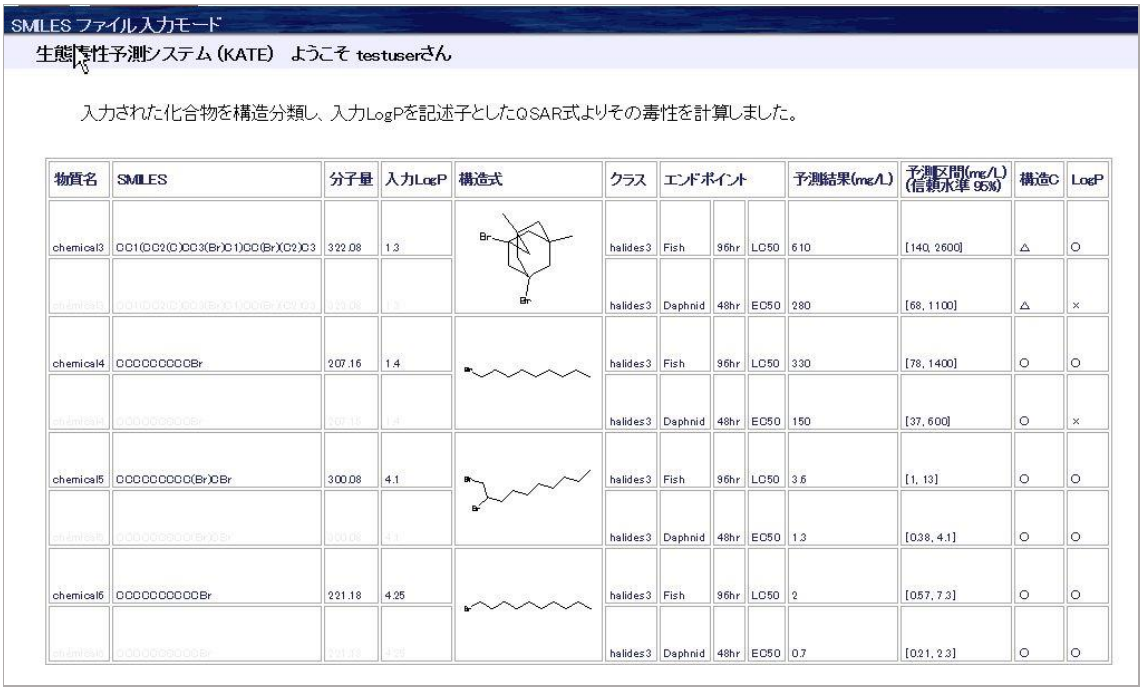
なお、LogP や水溶解度を予測するソフトウエアには米国環境保護庁（US EPA）が無償で提供する EPI Suite などがあります。

下段には、化学物質に含まれる部分構造の一覧が表示されます。

部分構造一覧

ID	名称	個数
5012	芳香族c-OH	1
5087	芳香族c-Br	1
5201	benzene	1
5236	Aromatics	1
5273	Phenols Single	1
4130	Aromatic Carbon	6
4285	-OH [hydroxy, aromatic attach]	1
4146	-Br [bromine, aromatic attach]	1
4902	脂肪族 O	1
4910	芳香族 原子	6
4915	芳香族 carbon-halogen	1
4917	フェノール	1
6040	フェノール	1

また、複数の化学物質の予測を行った場合には、結果は以下のように表示されます。



SMILES ファイル入力モード

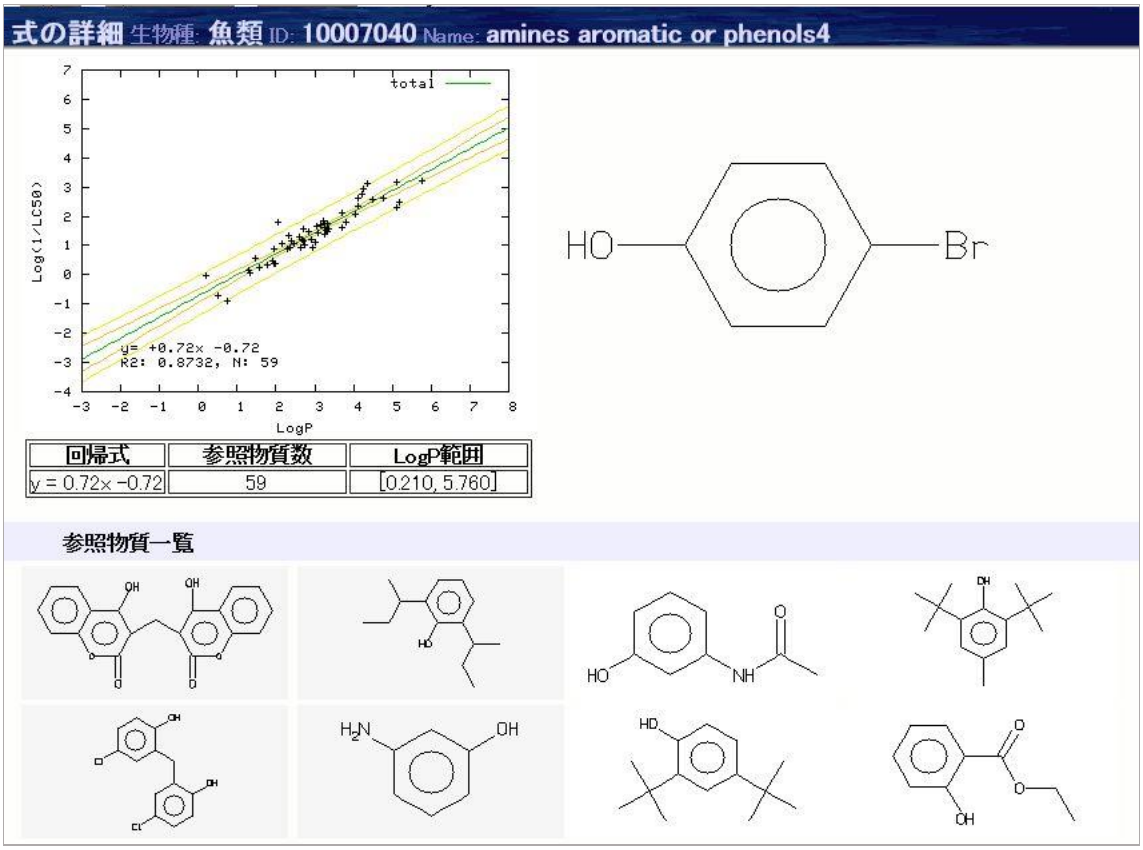
生態毒性予測システム（KATE） ようこそ testuserさん

入力された化合物を構造分類し、入力LogPを記述子としたQSAR式よりその毒性を計算しました。

物質名	SMILES	分子量	入力LogP	構造式	クラス	エンドポイント			予測結果(mg/L)	予測区間(mg/L)(信頼水準 95%)	構造C	LogP
chemical3	CC1(CC2(C)CC3(Br)C1)CC(Br)(C2)C3	322.08	1.3		halides3	Fish	96hr	LC50	610	[140, 2600]	△	○
					halides3	Daphnid	48hr	EC50	280	[68, 1100]	△	×
chemical4	CCCCCCCCBr	207.16	1.4		halides3	Fish	96hr	LC50	330	[78, 1400]	○	○
					halides3	Daphnid	48hr	EC50	150	[37, 600]	○	×
chemical5	CCCCCCCC(Br)CBr	300.08	4.1		halides3	Fish	96hr	LC50	3.6	[1, 13]	○	○
					halides3	Daphnid	48hr	EC50	1.3	[0.38, 4.1]	○	○
chemical6	CCCCCCCCCBr	221.18	4.25		halides3	Fish	96hr	LC50	2	[0.57, 7.3]	○	○
					halides3	Daphnid	48hr	EC50	0.7	[0.21, 2.3]	○	○

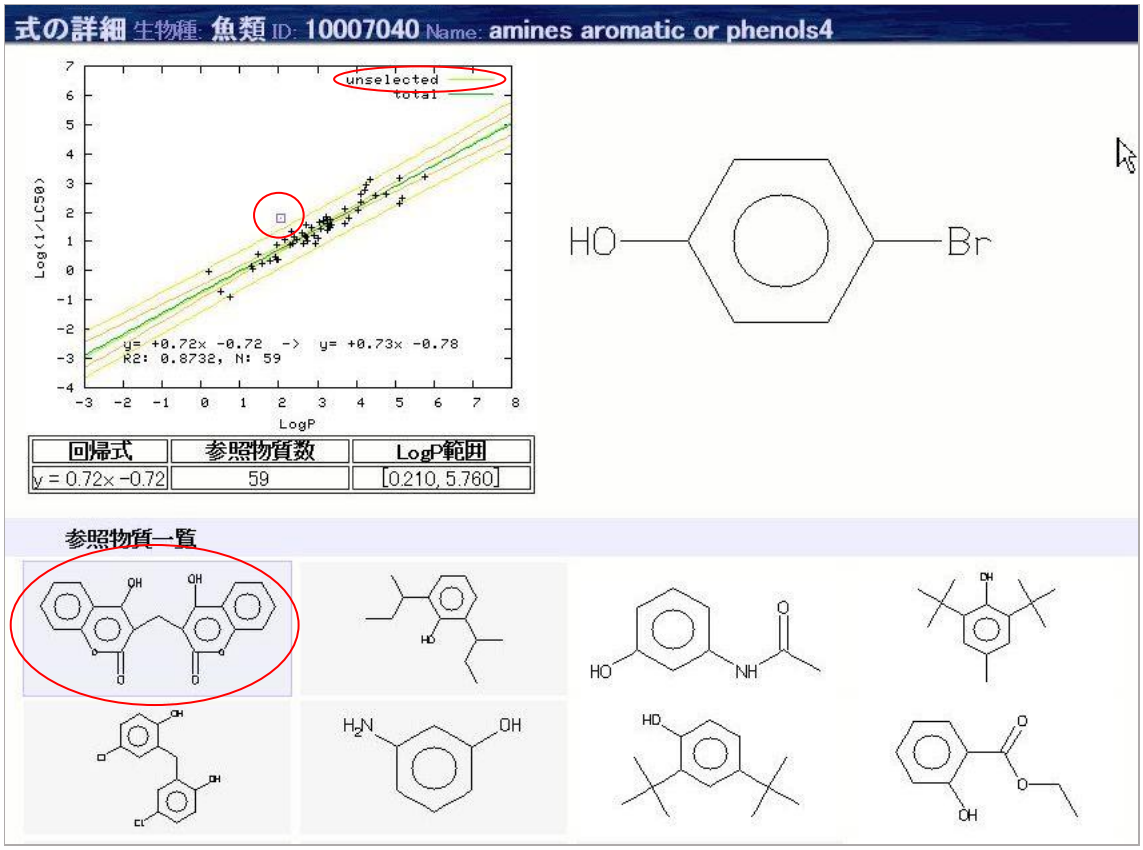
5.　参照物質の閲覧と比較

14 ページの予測結果画面で、分類されたクラスの をクリックすると、そのクラスの回帰直線及び参照物質一覧が表示されます。

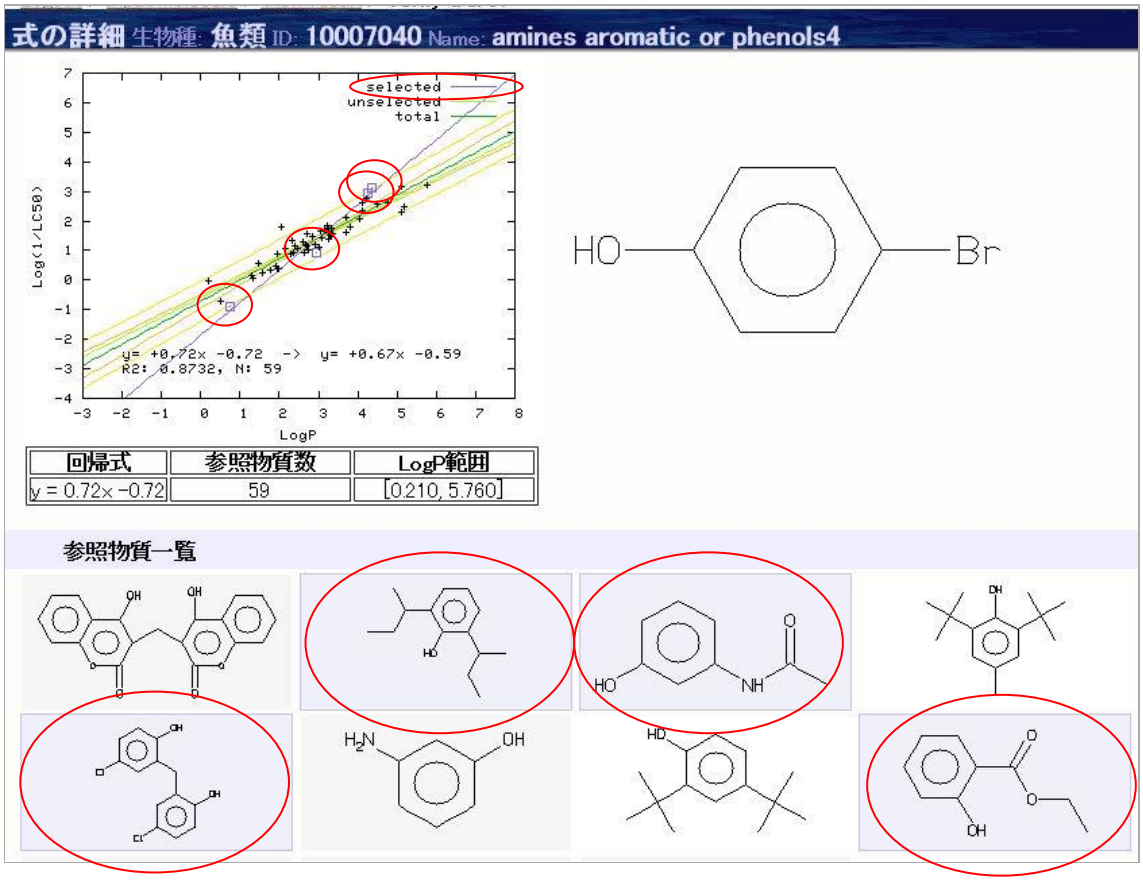


グラフ中の緑色の線は「回帰直線」、オレンジの線は「回帰式の信頼水準 95%の信頼限界」、黄色の線は「予測値の信頼水準 95%の信頼限界」を示します。また、回帰式の上側にある物質は構造式の背景が灰色、下側にある物質は背景が白色になります。

下段の参照物質の構造式またはグラフ上のポイントをクリックして選択すると、その物質を除いた回帰直線（unselected、薄黄緑色）を表示させることが出来ます。グラフ上では、選択された物質は□で示されます。もう一度物質をクリックすると選択が解除されます。



さらに、４つ以上の物質を選択すると、それらを参照物質とした回帰直線（selected、紫色）がグラフ上に表示されます。



6.　ログアウト

作業が終了したら、ログアウトをしてください。ログアウトをせずに終了すると、作業されていた最後の物質の予測結果がサーバーに残りますので、ご注意ください。なお、ログアウトを忘れた場合でも、同じユーザ名で再度ログイン、ログアウトをすれば情報は削除されます。

おわり

7. Appendix（付録）

（1）Appendix： 結果画面の「判定」で示される内容について

構造 C 判定

KATE ではあらかじめ定義された部分構造の有無で、クラスの適用範囲を「構造 C 判定」として判断します。ここで[そのクラス]とは予測する化学物質が分類されたクラスを指します。

「構造 C 判定」では，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、○、△の場合は[そのクラス]の適用範囲と判断しています。

LogP 判定

スタンドアロン版 KATE on PAS では logP 値が回帰式の有効範囲未満の場合は<P、logP 値が回帰式の有効範囲を超えている場合は>P がそれぞれ表示されます。

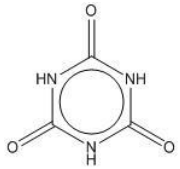
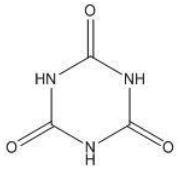
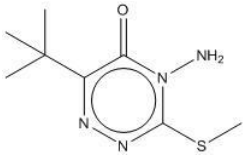
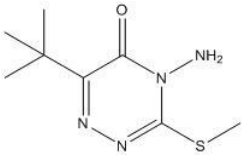
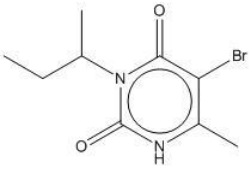
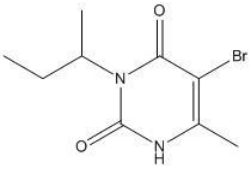
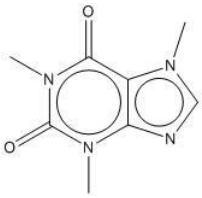
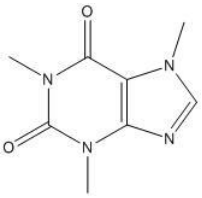
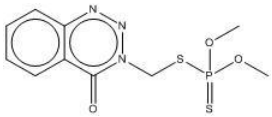
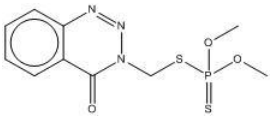
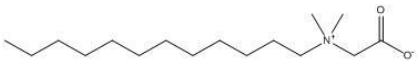
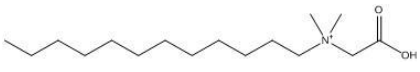
インターネット版 KATE on NET では Log P に対して内挿する場合に○、しない場合に×が「LogP」の欄に表示されます。

溶解度判定

スタンドアロン版 KATE on PAS では毒性レベルが溶解度を上回る場合に>S と表示されます。

インターネット版 KATE on NET では毒性値まで溶解する場合に○、しない場合に×が「溶解度」の欄に表示されます。

（2）KATE での入力時に SMILES の修正が必要な例

不可の SMILES と構造式	修正後の SMILES と構造式
CAS 番号：108-80-5	
 O=c1[nH]c(=O)[nH]c(=O)[nH]1	 O=C1NC(=O)NC(=O)N1
CAS 番号：21087-64-9	
 CSc1nnc(c(=O)n1N)C(C)(C)C	 CSC1=NN=C(C(=O)N1N)C(C)(C)C
CAS 番号：314-40-9	
 CCC(C)n1c(=O)[nH]c(C)c(Br)c1=O	 CCC(C)N1C(=O)NC(C)=C(Br)C1=O
CAS 番号：58-08-2	
 Cn1cnc2n(C)c(=O)n(C)c(=O)c12	 CN1C=NC2N(C)C(=O)N(C)C(=O)C1=2
CAS 番号：86-50-0	
 COP(=S)(OC)SCn1nnc2ccccc2c1=O	 COP(=S)(OC)SCN1N=Nc2ccccc2C1=O
CAS 番号：683-10-3	
 CCCCCCCCCCCC[N+](C)(C)CC(=O)[O-]	 CCCCCCCCCCCC[N+](C)(C)CC(=O)O