# QSAR Toolbox @ U.S. EPA

## Connection with OncoLogic™

Dr. Brianne Raccor

Dr. Martin Phillips

U.S. EPA

Office of Pollution Prevention and Toxics

Data Gathering & Analysis Division

# Biochemical Basis of Cancer Development

Carcinogenesis is a multi-step, multi-factorial process that requires both genetic and/or epigenetic cellular changes to progress to invasive cancer.

- 3 Basic Stages of Cancer Formation
  - Initiation
    - Gene mutations
    - Epigenetic changes
  - Promotion
    - Clonal expansion of cancer phenotype (e.g., defect in terminal differentiation, growth control, or decreased sensitivity to apoptosis)
  - Progression
    - Activation of proto-oncogenes or inactivation tumor suppression genes and antimetastatic genes

# OncoLogic™: A mechanism-based expert system for predicting carcinogenic potential

- Developed by domain experts in collaboration with expert system developer
- Knowledge from SAR on >10K chemicals
- Class-specific approach to optimize predictive capability
- Considers all relevant factors including biological input when possible
- Predictions with scientific rationale and ordinal ranking
- In January 2021, OncoLogic 8™ was migrated to a newer standalone platform - OncoLogic 9™
- Available for free download from https://tinyurl.com/oncologic

# OncoLogic™ - Expert System

## HOW IT WORKS

- Mimic the thinking and reasoning of human experts using knowledge-based rules for chemical classes to predict cancer concern
  - Assigns a baseline concern level ranging from low to high for each structure class
  - Evaluates how substituents on the chemical may affect carcinogenicity
    - Concern level changes accordingly

## BENEFITS

- Expedites the decision-making process
- Allows sharing of knowledge
- Reduces error and inconsistency
- Formalizes knowledge rules for cancer hazard identification
- Allows non-experts to reach scientifically supportable conclusions

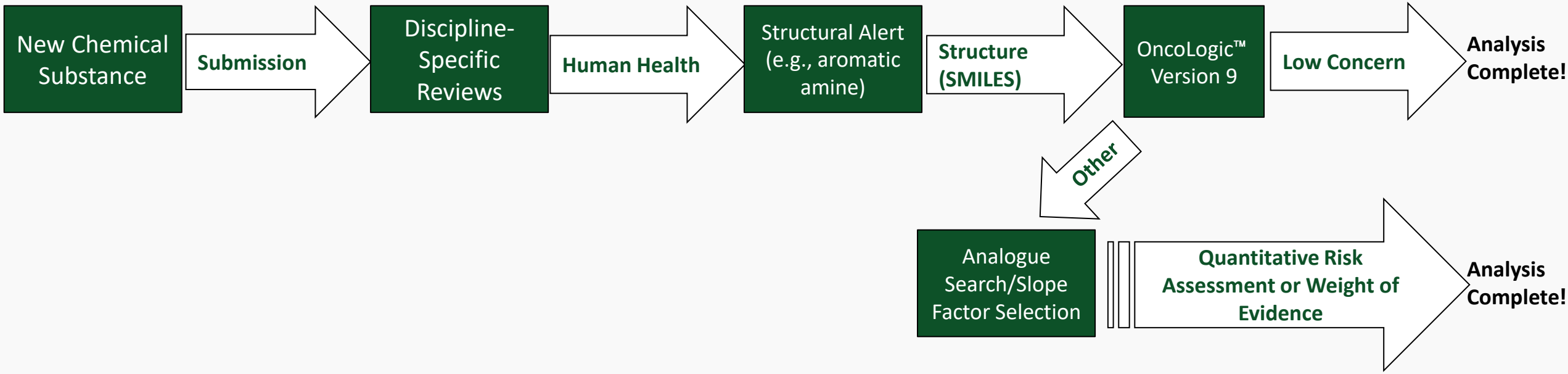
# OncoLogic™ - Concern Levels

| OncoLogic™ Concern | Definition |
|---|---|
| Low | Unlikely to be carcinogenic |
| Marginal | Likely to have equivocal carcinogenic activity |
| Low – Moderate | Likely to be weakly carcinogenic |
| Moderate | Likely to be a moderately active carcinogen |
| Moderate – High | Highly likely to be a moderately active carcinogen |
| High | Highly likely to be a potent carcinogen |

# EPA TSCA New Chemicals Workflow

| OncoLogic™ Concern |
|---|
| Low |
| Marginal |
| Low – Moderate |
| Moderate |
| Moderate – High |
| High |

- Refines class-based structural alerts to better predict cancer hazards



# Integration of OncoLogic 9™ with QSAR Toolbox

- For analysis of some chemical classes (e.g., organophosphates), OncoLogic 9™ requires the user to answer a series of questions for the prediction of carcinogenic potential
- The aim of the current project is to reduce the need for expert knowledge through the integration of QSAR Toolbox with OncoLogic 9™
- The data and knowledge that exists in QSAR Toolbox would provide automatic answers to the questions asked by  OncoLogic 9™
- Currently, QSAR Toolbox can automatically answer questions related to predicting concern for organophosphorus compounds

# Integration of OncoLogic 9™ with QSAR Toolbox

Questions that may be asked by OncoLogic 9™ for organophosphorus compounds, which have been integrated with QSAR Toolbox:

1. Select the results of the *IN VIVO* Genotoxicity Testing
    - Answers: Positive, Negative, Unknown
2. Select the results of the *IN VITRO* Genotoxicity Testing
    - Answers: Positive, Negative, Unknown
3. Is the chemical highly toxic?
    - Answers: Yes, No, Unknown
4. Is rapid hydrolysis or detoxification expected?
    - Answers: Yes, No, Unknown

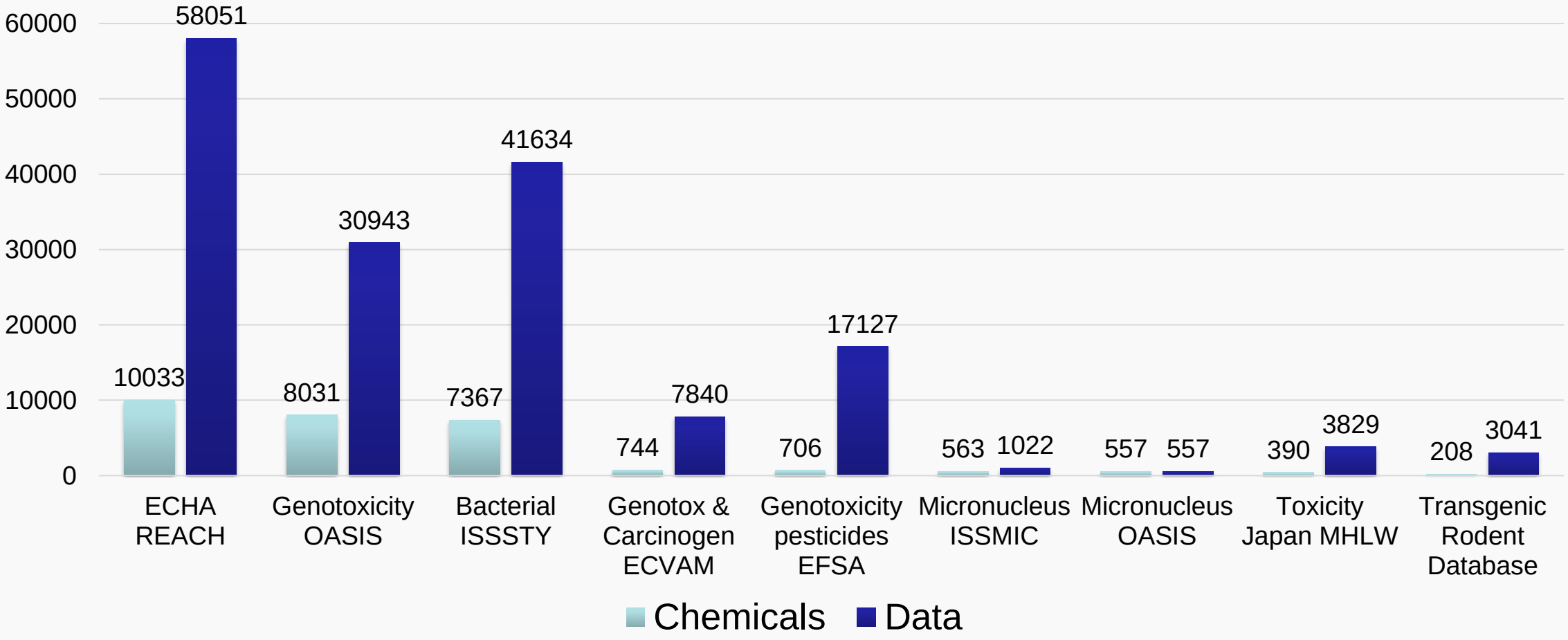
# Integrating QSAR Toolbox with OncoLogic 9™ for Classification of Genotoxicity

- QSAR Toolbox databases provide *in vivo* and *in vitro* genotoxicity data
- The system provides an automatic answer to the requested information about genotoxicity and extracts experimental data from QSAR Toolbox databases
- Answers are based on:
  - Positive – worst-case scenario considered (at least one positive experimental data point was identified)
  - Negative – only negative data was found
  - Unknown – i.e., equivocal (technically compromised, inconclusive, etc.) or the chemical has no data
- The answer is explained in the final report

# Databases Available from QSAR Toolbox

- Databases including *in vivo/ in vitro* genotoxicity data are:
  - ECHA REACH
  - Genotoxicity & Carcinogenicity ECVAM
  - Genotoxicity OASIS
  - Genotoxicity pesticides EFSA
  - Bacterial ISSSTY
  - Micronucleus ISSMIC
  - Micronucleus OASIS
  - Toxicity Japan MHLW
  - Transgenic Rodent Database

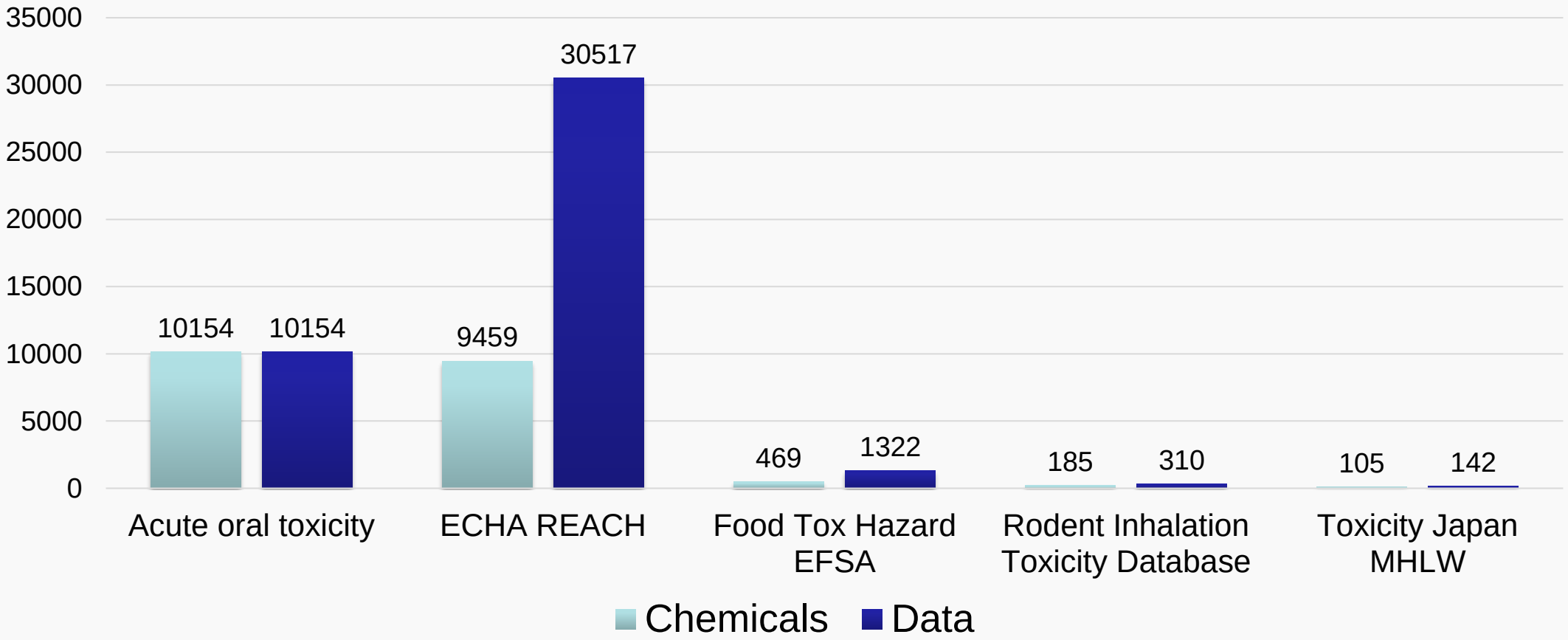
# Distribution of *In Vitro/In Vivo* Genotoxicity Data Across QSAR Toolbox Databases



# Integrating QSAR Toolbox with OncoLogic 9™ for Classification of Toxicity

- OncoLogic 9™ incorporates toxicity data into overall classification for level of concern
- Potential answers: Yes, No, Unknown
- QSAR Toolbox can extract documented **acute toxicity** data from the following Toolbox databases:
  - Acute Oral Toxicity Database
  - ECHA REACH
  - Food Tox Hazard EFSA
  - Rodent Inhalation Toxicity Database
  - Toxicity Japan MHLW

# Distribution of Acute Toxicity Data Across QSAR Toolbox Databases



# Classification of Chemicals as Highly Toxic Based on $LD_{50}$ or $LC_{50}$ Extracted from QSAR Toolbox

| Answers | Oral (gavage, capsule, drinking water, feed, unspecified) | Dermal | Inhalation (aerosol, gas) | Inhalation (vapor) | Inhalation (dust, mist) |
|---|---|---|---|---|---|
| | Mass fraction (mg/ kg) | Mass fraction (mg/kg) | Volume concentration (ppm) | Mass concentration (mg/L) | Mass concentration(mg/L) |
| Yes | ≤ 5 | ≤ 50 | ≤ 100 | ≤ 0.5 | ≤ 0.05 |
| No | > 5 | > 50 | > 100 | > 0.5 | > 0.05 |
| Unknown | Not “yes” or “no” category, no data available | | | | |

# Classification of Hydrolysis Based on Calculation of Half-Life

- Question: "Is rapid hydrolysis or detoxification expected?“
- Potential Answers: Yes, No, Unknown
- QSAR Toolbox can model hydrolysis half-life values or search for experimental kinetic hydrolysis half-life data, classify chemicals into the following categories, and use the category to automatically answer the question in OncoLogic 9™:
  - very slow (>100 days)
  - slow (30 - 100 days),
  - moderate (10-30 days)
  - fast (5-10 days)

  } No

  - very fast (1-5 days)
  - extremely fast (< 1 day)

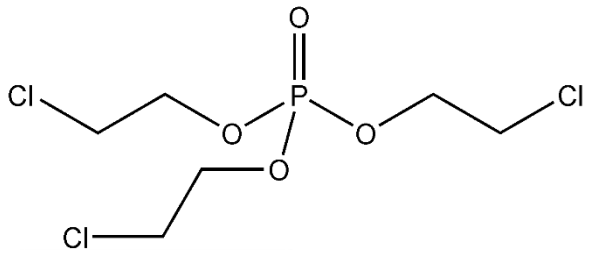
  } Yes

# Example for the Classification of *In Vivo* Genotoxicity

OncoLogic™ 9 report

| Chemical class | Level of concern |
|---|---|
| **Organophosphorus Compounds** | |
| Trialkyl (thio)phosphates | Moderate |

**OncoLogic Justification Report**




The concern based on structure-activity relationship consideration for this trialkyl phosphate-type compound is LOW-MODERATE

The concern based on the functional properties of this trialkyl phosphate-type compound is HIGH-MODERATE

Considering both the SAR analysis and the functional properties of the trialkyl phosphate-type compound, the average level of concern is used. The final level of concern is MODERATE

The effect of any highlighted substituents is uncertain.

JUSTIFICATION

Organophosphorus (OP) compounds are organic compounds containing phosphorus (P) as a constituent.The principal carcinogenic OP compounds are:

i) alkyl phosphates, phosphonates, phosphoroamidates and their thio derivatives (phosphorothioates,phosphorothiolates, thiophosphonates, thiophosphoroamidates.
ii) phosphoramides, and
iii) phosphorus-containing nitrogen mustards

The structural features that may affect carcinogenicity of alkyl phosphates, phosphonates, phophoroamidates, and their related thio derivates include:

i) nature of alkyl groups (size, degree of branching, halogenation) which may serve as an alkylating moiety.
ii) presence or absence of electron-withdrawing group which may enhance alkylating activity.
iii) ease or resistance to detoxifying hydrolysis which may abolish alkylating activity.
iv) potentially carcinogenic hydrolysis product(s).

Phosphoramides and certain phosphorus-containing nitrogen-mustards (e.g. cylcophosphamides) may require metabolic activation to alkylating intermediates.

The baseline level of concern for this unsubstituted trialkyl phosphate-type compound, where R1 is ethyl, R2 is ethyl, R3 is ethyl, is LOW-MODERATE.

There are no or inadequate carcinogenicity data on this compound. The concern level derived for this compound is based on structure-activity relationship analysis indicative of potential alkylating activity.

The single chloro, bromo, or iodo substituent on the beta position of R1 is expected to increase the level of concern.

The single chloro, bromo, or iodo substituent on the beta position of R2 is expected to increase the level of concern.

The single chloro, bromo, or iodo substituent on the beta position of R3 is expected to increase the level of concern.

Therefore the level of concern is raised to UNKNOWN.

Because the compound is genotoxic IN VIVO, and its overall toxicity is not high, the carcinogenicity concern based on its functional properties is HIGH-MODERATE.

The concern level based on SAR analysis differs from the level based on the functional properties of the compound. The concern level assigned for the compound is the average of the levels derived from SAR analysis and consideration of the functional properties.

The final level of concern for this trialkyl phosphate-type compound is MODERATE.

The following questions have been automatically answered by QSAR Toolbox:

Question: "Select the results of the IN VIVO Genotoxicity Testing."
Answer: "Positive"

Question: "Is the compound highly toxic?"
Answer: "No"

Explain Answers

# Explanation of QSAR Toolbox Classification of *In Vivo* Genotoxicity

for the compound is the average of the levels derived from SAR analysis and consideration of the functional properties.

The final level of concern for this trialkyl phosphate-type compound is MODERATE.

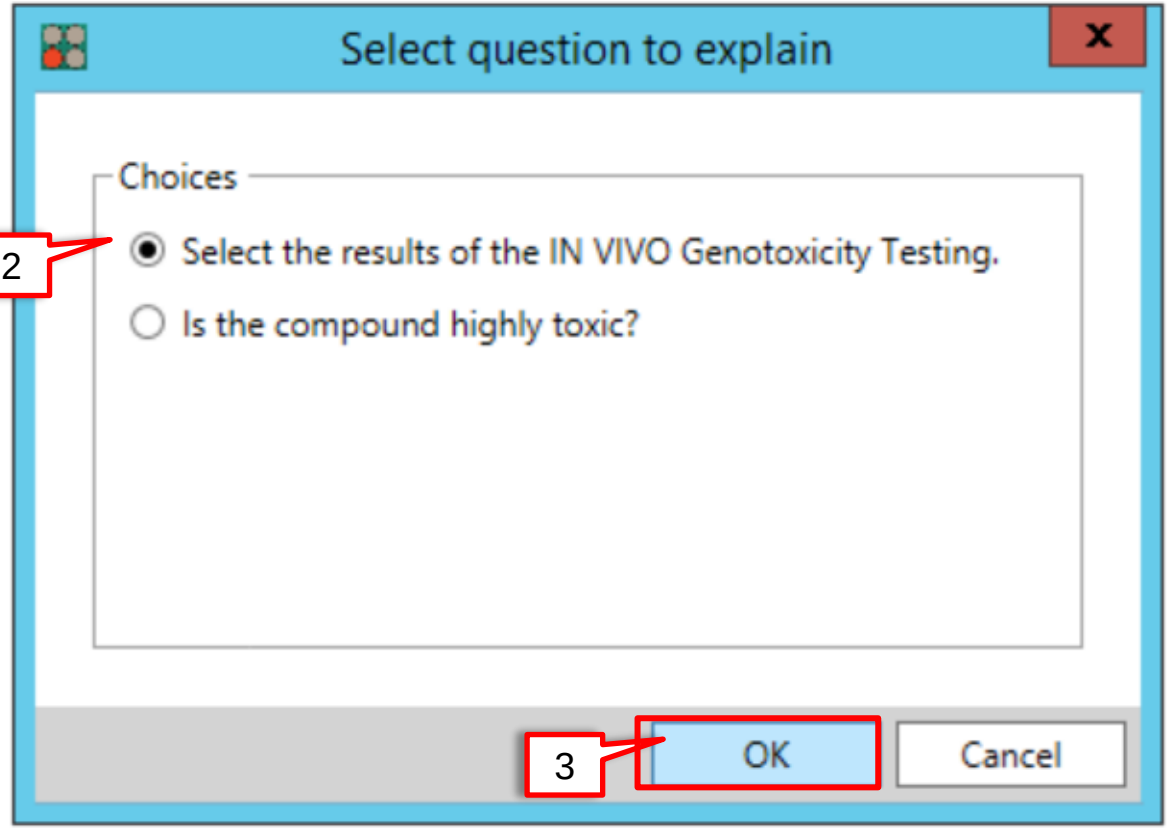
The following questions have been automatically answered by QSAR Toolbox:

Question: "Select the results of the IN VIVO Genotoxicity Testing."
Answer: "Positive"

Question: "Is the compound highly toxic?"
Answer: "No"

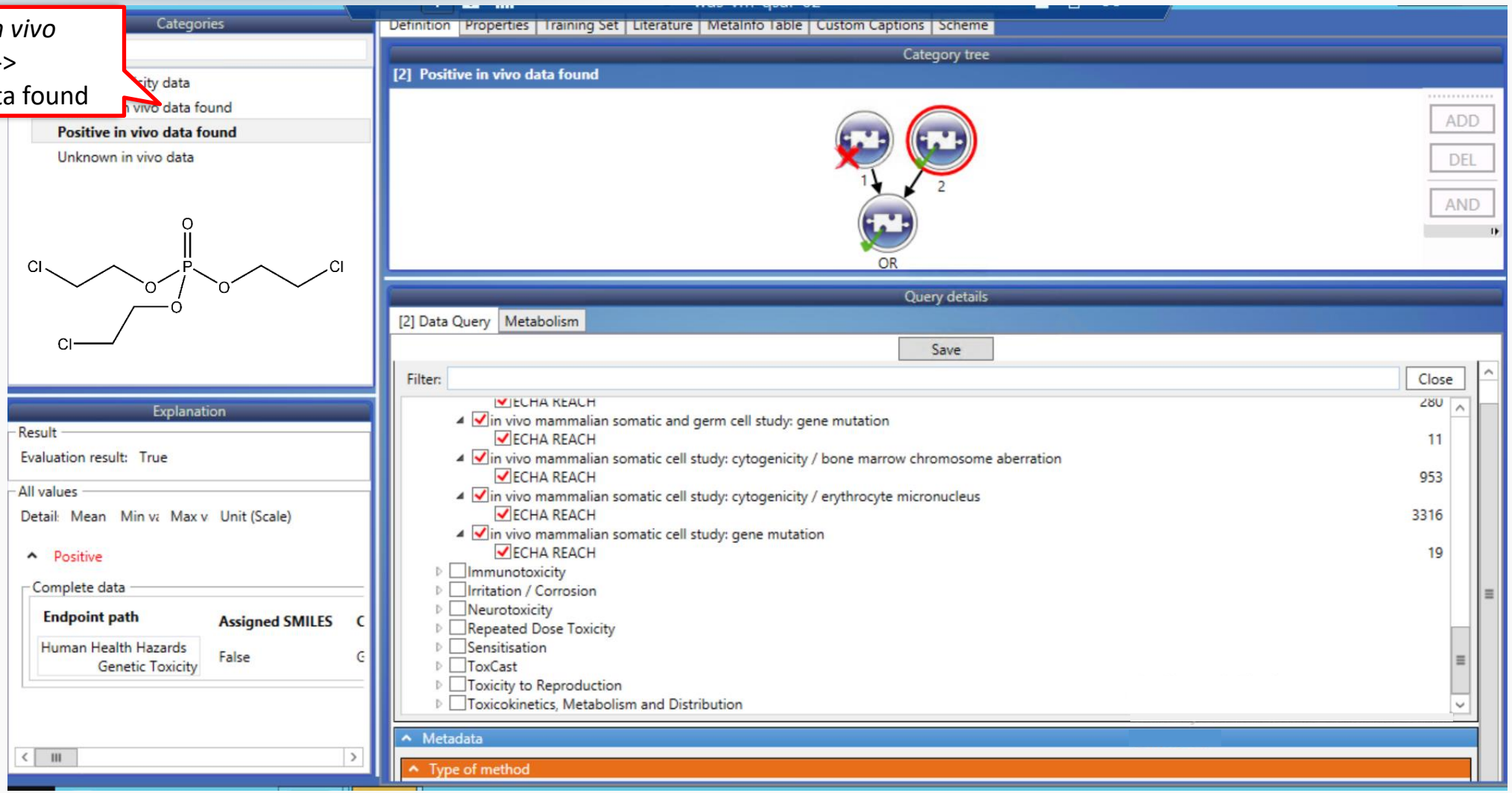
Explain Answers

1



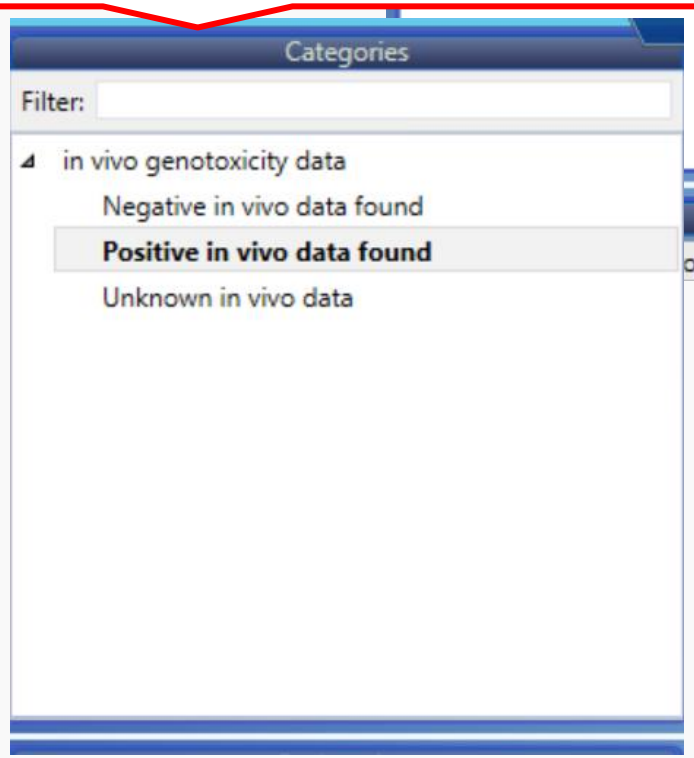
# QSAR Toolbox Explanation

Explanation for: *in vivo* genotoxicity data -> Positive *in vivo* data found

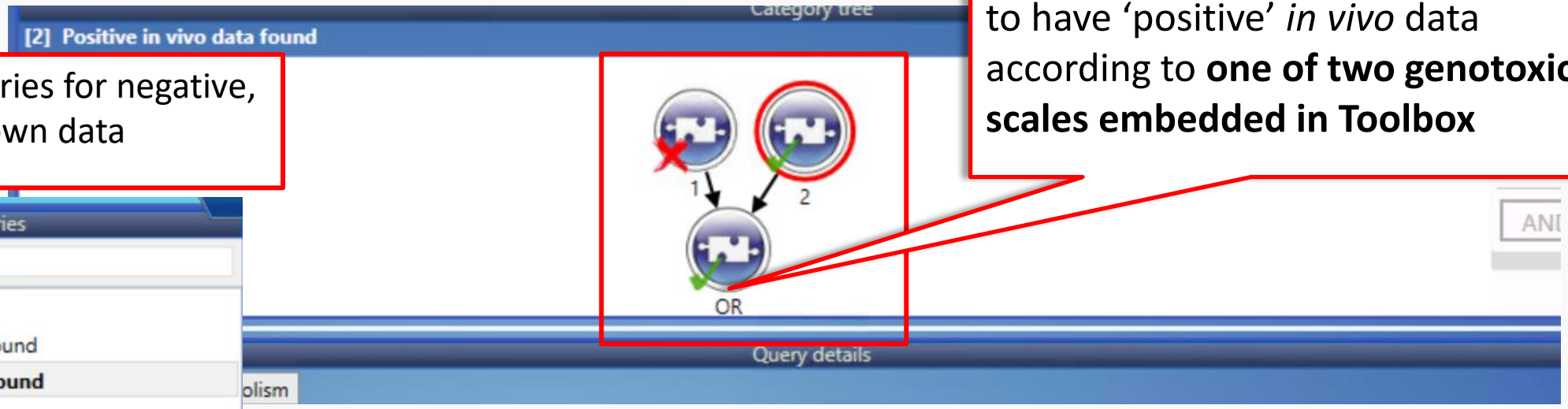


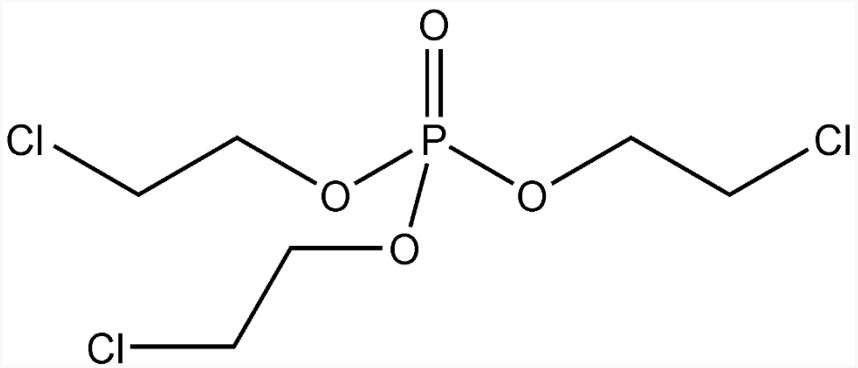
# QSAR Toolbox Categories

The defined categories for negative, positive, and unknown data

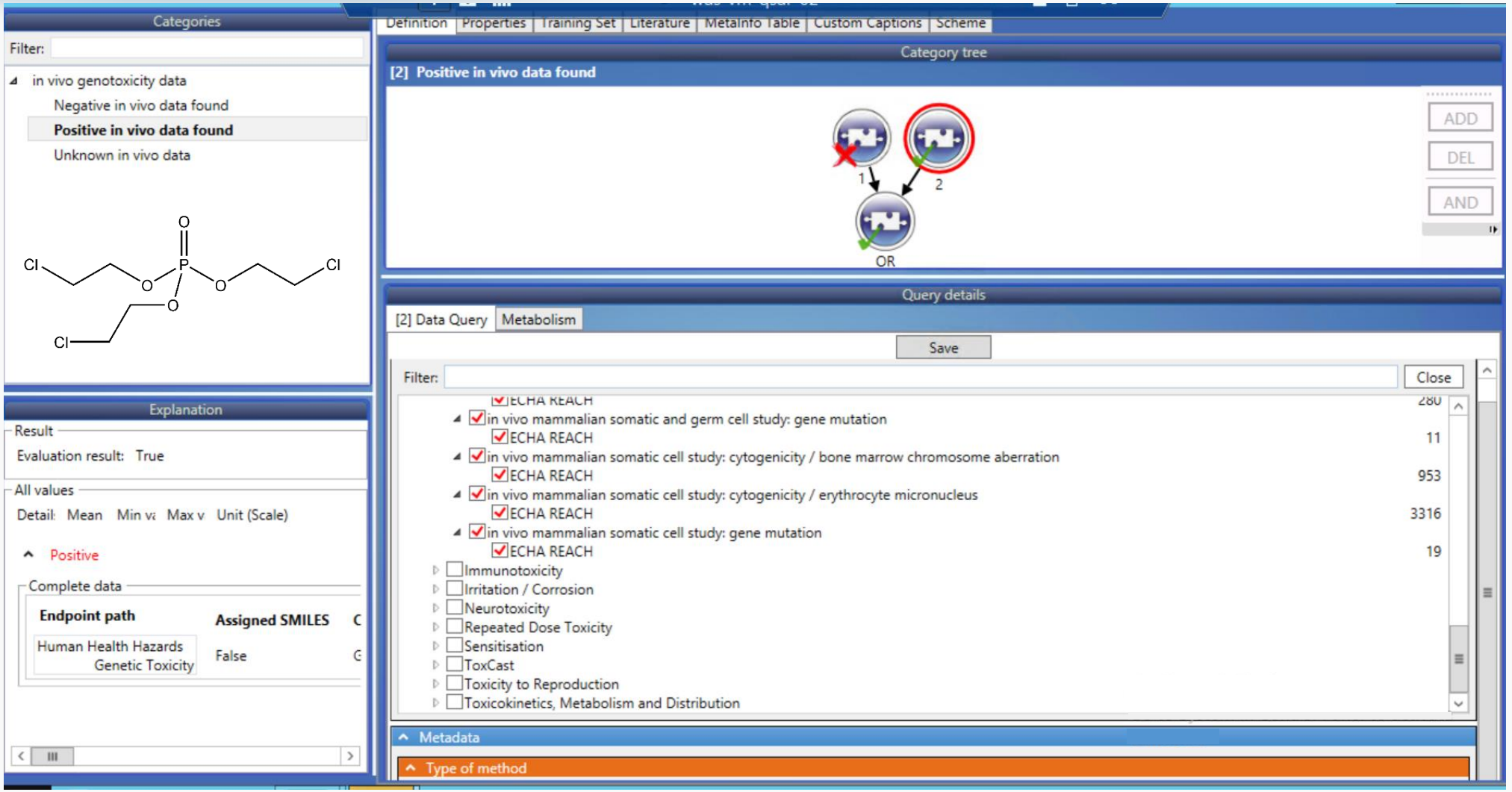



Data boundaries requiring chemicals to have 'positive' *in vivo* data according to **one of two genotoxicity scales embedded in Toolbox**

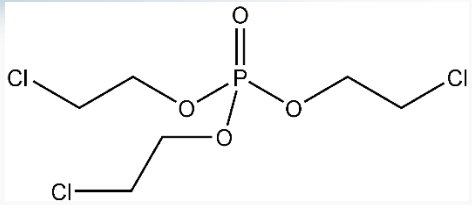


# QSAR Toolbox Explanation



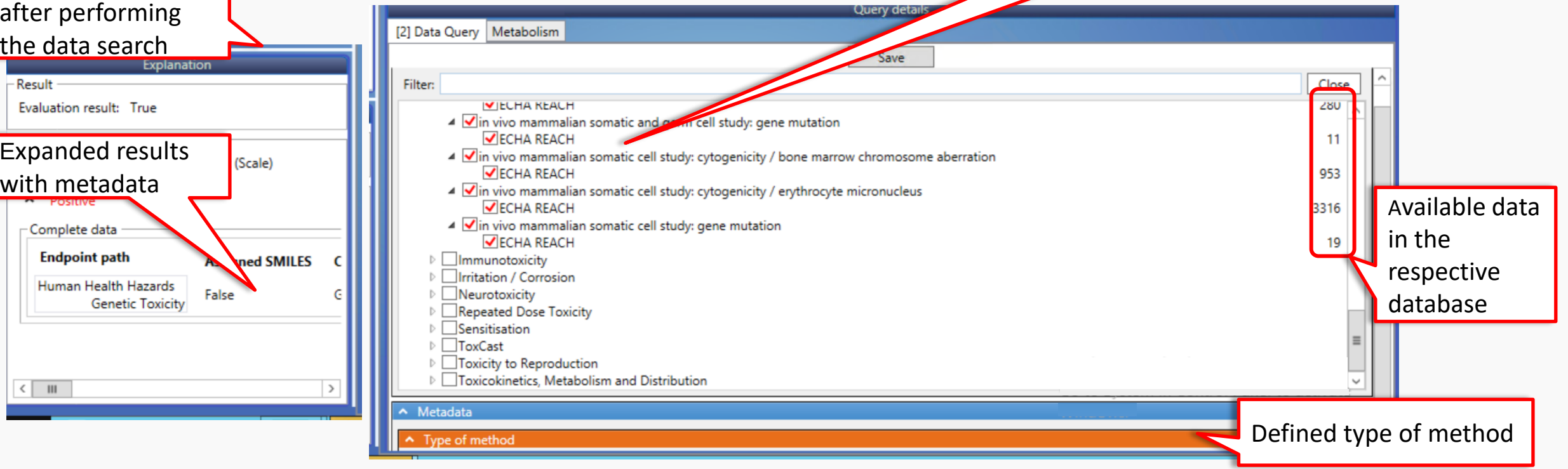
# Databases Searched by QSAR Toolbox




Selected databases with genotoxicity data used for searching experimental data

Results derived after performing the data search

Expanded results with metadata




Available data in the respective database

Defined type of method

# QSAR Toolbox Explanation Summary

The defined categories for negative, positive, and unknown data

Data boundaries requiring chemicals to have 'positive' *in vivo* data according to **one of two genotoxicity scales embedded in Toolbox**

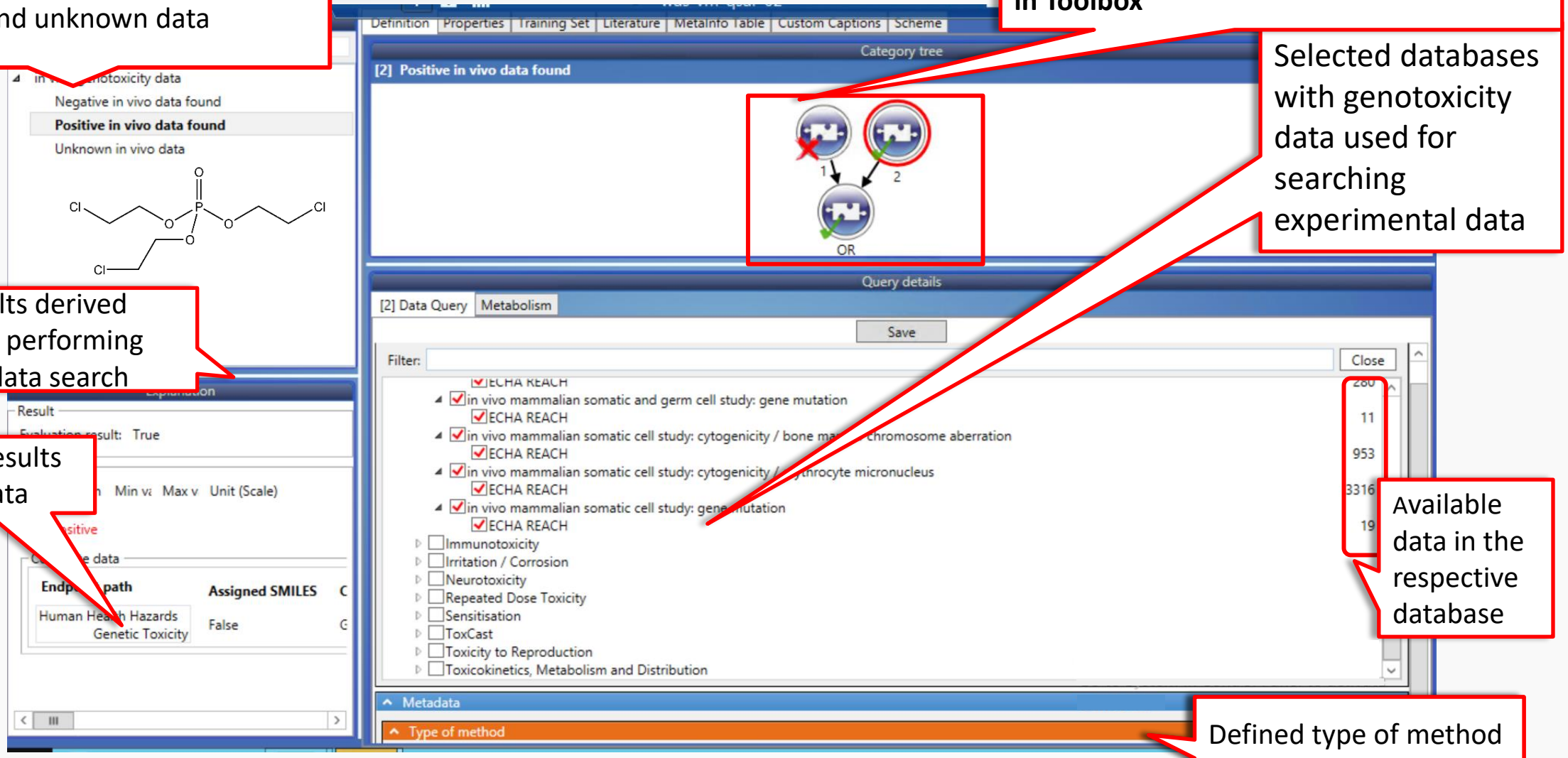
Selected databases with genotoxicity data used for searching experimental data

Results derived after performing the data search

Expanded results with metadata

Available data in the respective database

Defined type of method



# Recent Updates

1. **New version of OncoLogic™ 9.2**
   - Automatic prediction of alkylating ability for an organophosphorus compound
   - Automatic prediction of chelating ability for an organophosphorus compound
2. **OncoLogic™ 9 Add-In for QSAR Toolbox**
   - Evaluation of the carcinogenic potential of multiple compounds using batch mode
   - Add-In is used to extract information from QSAR Toolbox to automatically answer OncoLogic™ questions
3. **QSAR Toolbox update (QSAR Toolbox 4.5 Service Pack 1)**

# Acknowledgements


- US EPA
  - Dr. Tala Henry
  - Dr. Yin-Tak Woo (retired)
- OECD
  - Dr. Patience Brown
  - Tomoko Aoyagi
- ECHA
  - Dr. Andrea Gissi
- LMC
  - Dr. Ovanes Mekenyan
  - Valentin Marinov

# THANK YOU!

Questions?

**Brianne Raccor, Ph.D.**
raccor.brianne@epa.gov

**Martin B. Phillips, Ph.D., DABT**
phillips.martinb@epa.gov