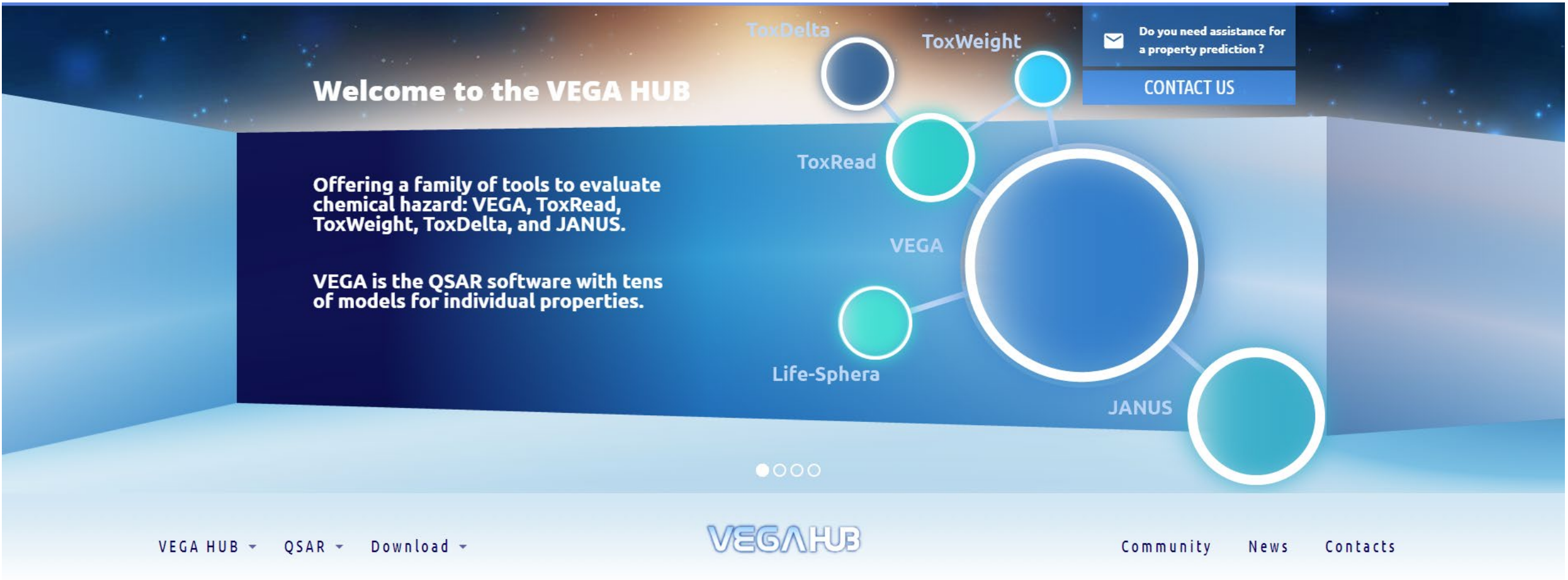
www.vegahub.eu



VEGA, introduction and novelty generated in this project

Alberto Manganaro, KODE s.r.l., Italy

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461

VEGAHUB

- Pre-built tools
- Links to SW to develop new models

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461

VEGAHUB

- Pre-built tools

VEGA
ToxRead
JANUS
Sphera
(toDIVINE)

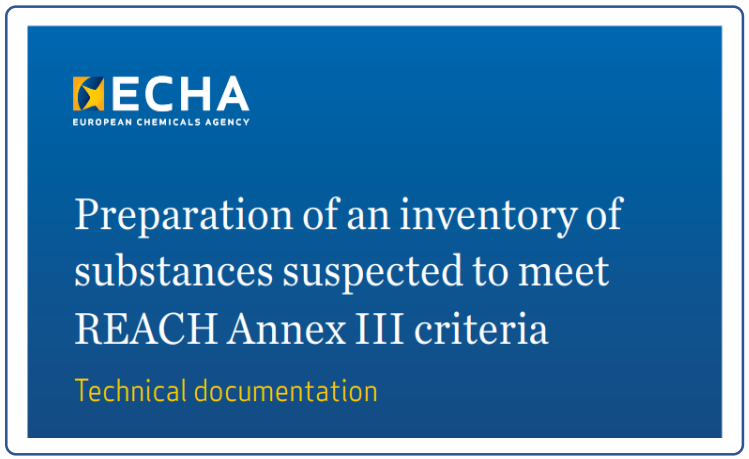
CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH
LIFE17 GIE/IT/000461

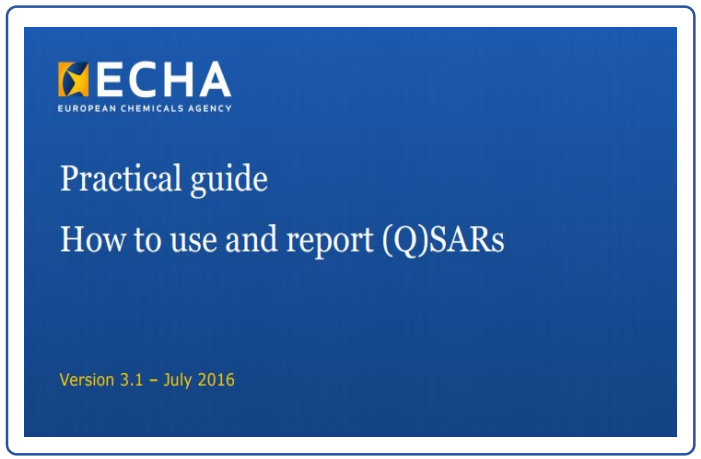
VEGA

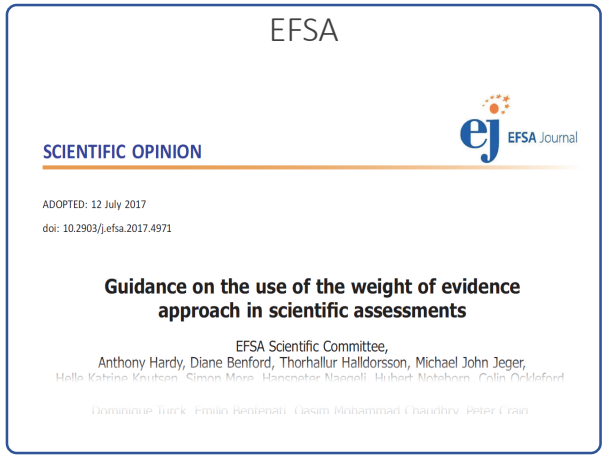
Regulatory use

- **OECD**
- **ECHA**
- **EFSA**

- Docked to the OECD QSAR Toolbox
- Projects from Italian Health and Environmental Ministries
- Used by German UBA for prioritization and substance evaluation







CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH
LIFE17 GIE/IT/000461

VEGA

Sources of the models

- EC / UBA / EFSA Projects
- Arnot-Gobas model for BCF (US EPA – EPI Suite software)
- Meylan model for BCF (US EPA – EPI Suite software)
- MOA aquatic (US EPA – TEST)
- Koc model (NIH – OPERA software)
- Koa model (NIH – OPERA software)
- Cramer class model (JRC – Toxtree software)
- Verhaar class model (JRC – Toxtree software)
- Mutagenicity ISS model (JRC – Toxtree software)

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461

Documentation

- ❑ QRMF
- ❑ Datasets
- ❑ Guides

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461

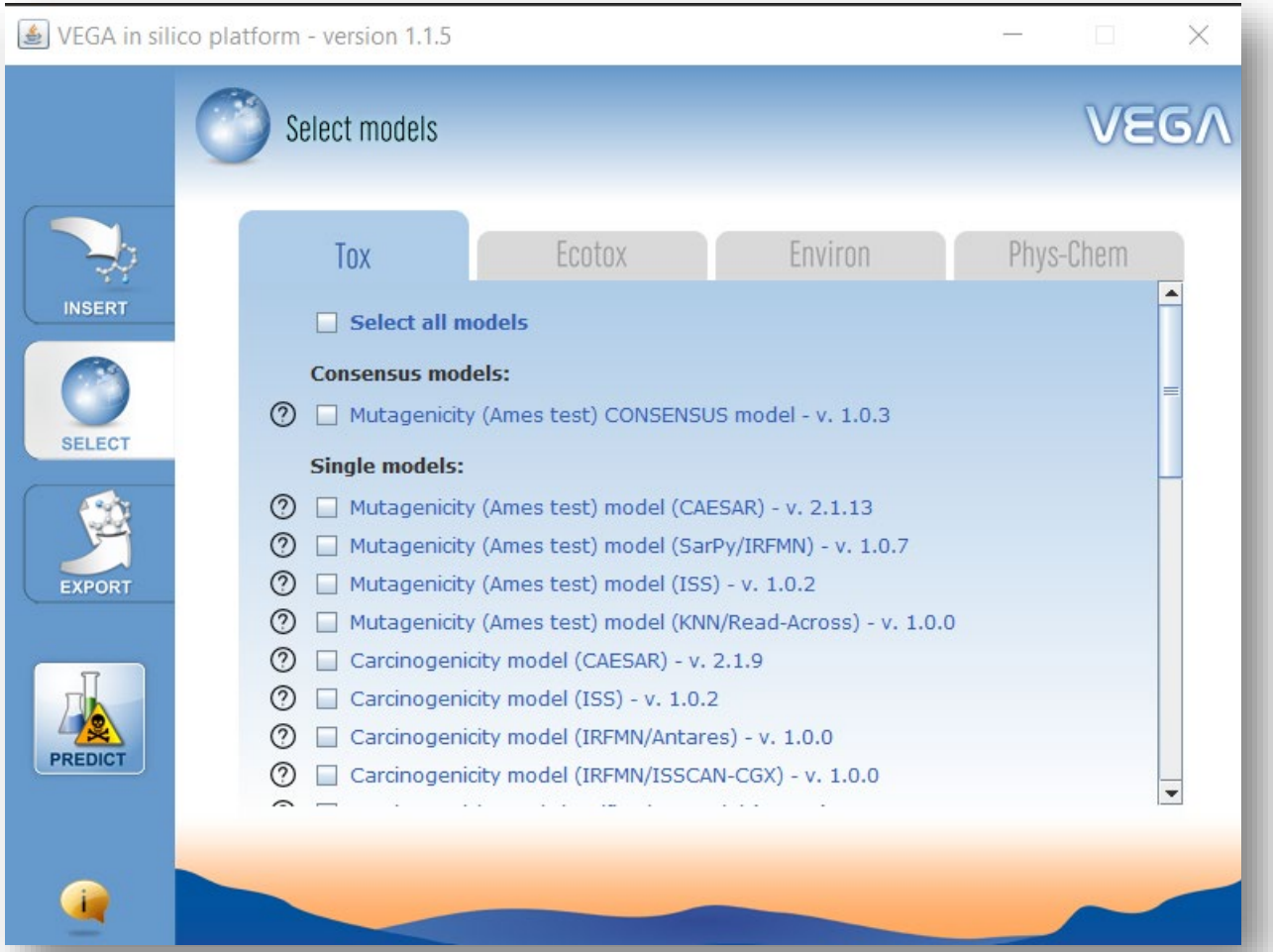
The conceptual architecture

- Individual models
- Tool for reliability, to be used also for read-across

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH
LIFE17 GIE/IT/000461

VEGA

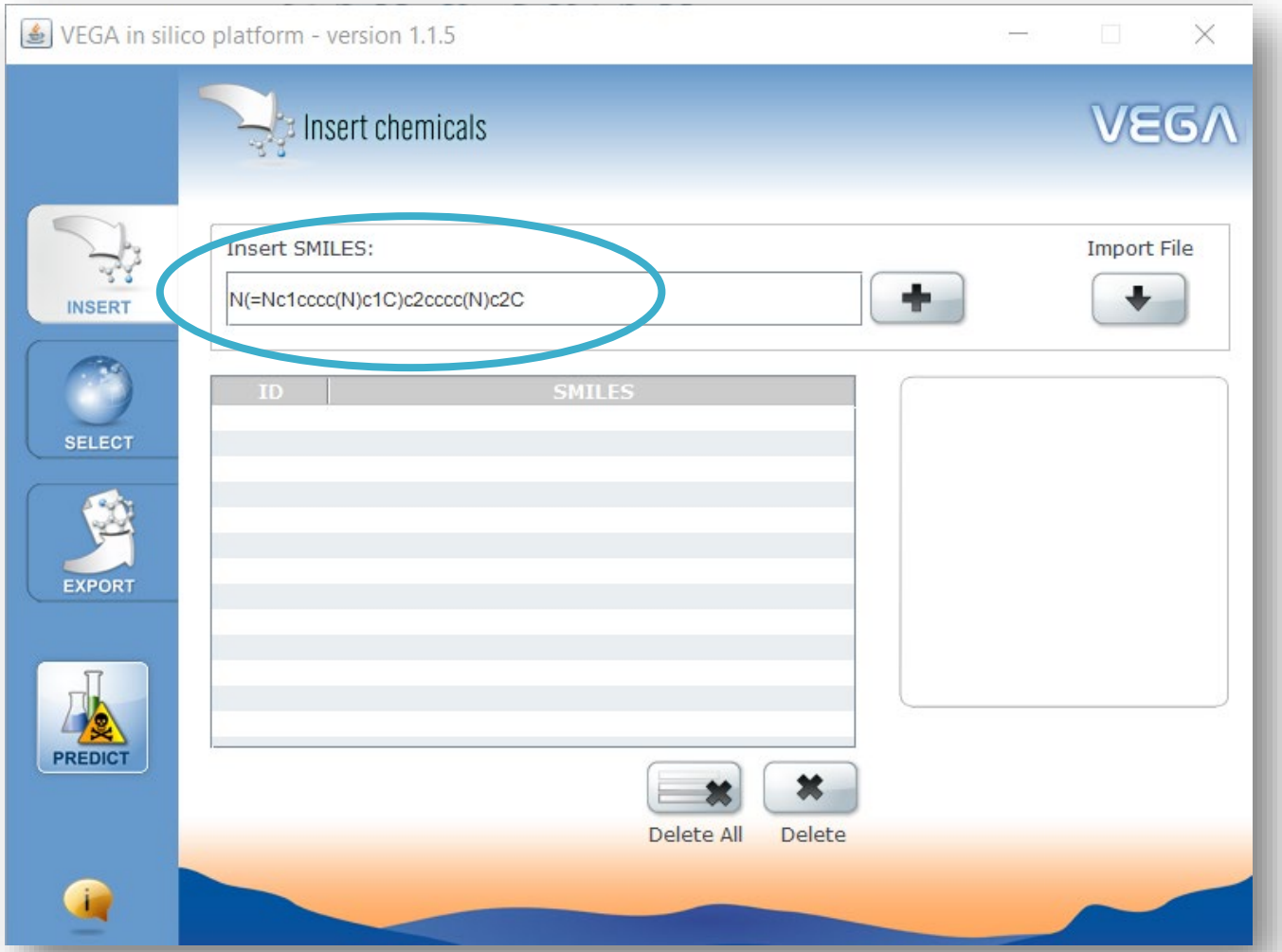
Predictive models



80 models regarding properties on

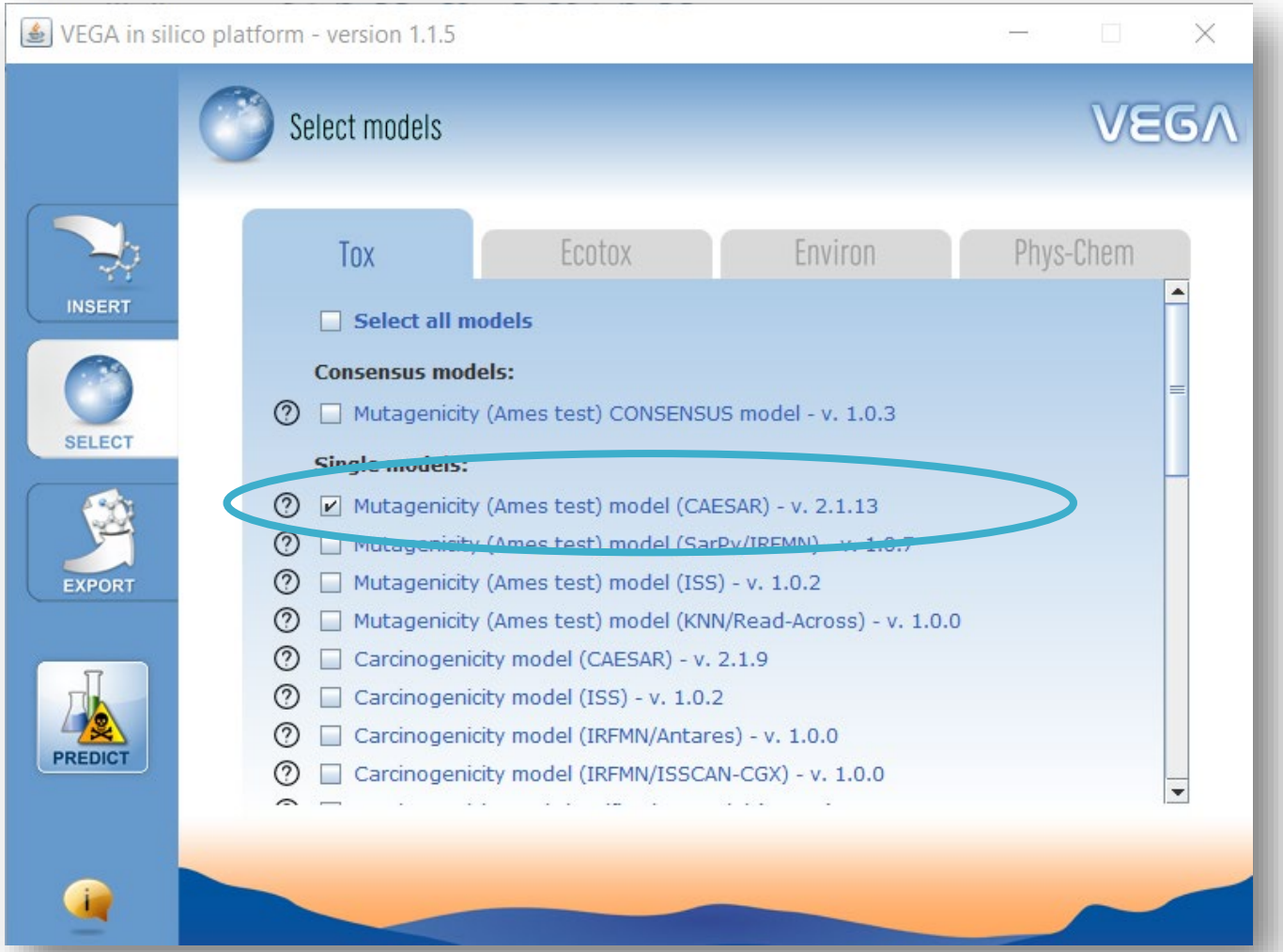
- Human toxicity
- Eco-toxicity
- Environmental
- Physico-chemical
- Toxicokinetics

Predictive models – Examples



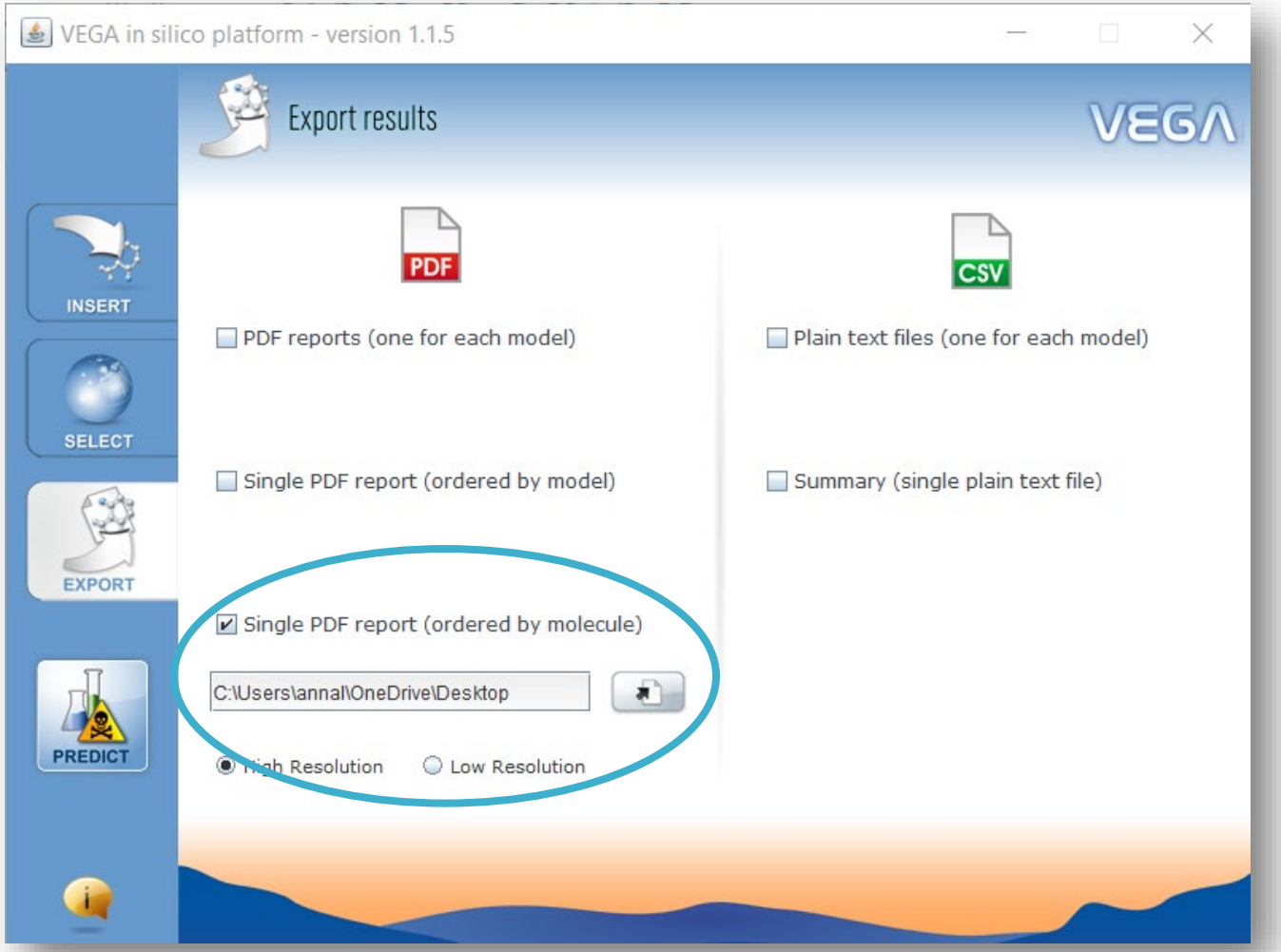
Insert page

Predictive models – Examples



Select page

Predictive models – Examples



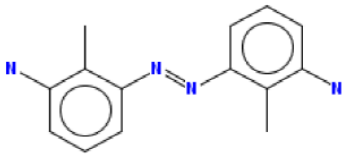
Export page

Predictive models – Examples

VEGA Mutagenicity (Ames test) model (CAESAR) 2.1.13 page 1

1. Prediction Summary

Prediction for compound Molecule 0



Prediction: Reliability:

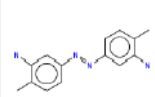
Prediction is Mutagenic, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.
The following relevant fragments have been found: SA29 Aromatic diazo

Compound: Molecule 0
Compound SMILES: N(=Nc1cccc(N)c1C)c2cccc(N)c2C
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural alerts: SA29 Aromatic diazo
Reliability: the predicted compound is into the Applicability Domain of the model
Remarks:
none

Output
Summary page

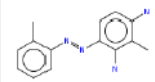
Predictive models – Output other pages

3.1 Applicability Domain:
Similar Compounds, with Pre



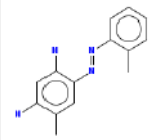
CAS: 119631-00-4
Dataset id: 3482 (Trainir
SMILES: N(=Nc1ccc(c(N
Similarity: 0.977

Experimental value: Muta
Predicted value: Mutager



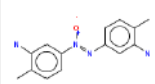
CAS: 84434-45-7
Dataset id: 2696 (Test s
SMILES: N(=Nc1ccccc1(
Similarity: 0.951

Experimental value: Muta
Predicted value: Mutager



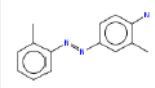
CAS: 7467-29-0
Dataset id: 3021 (Trainir
SMILES: N(=Nc1cc(c(N)(
Similarity: 0.949

Experimental value: Muta
Predicted value: Mutager



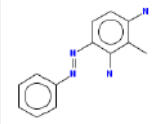
CAS: 5857-92-1
Dataset id: 2566 (Trainir
SMILES: [O-][N+](=Nc1c
Similarity: 0.93

Experimental value: Muta
Predicted value: Mutager



CAS: 97-56-3
Dataset id: 594 (Training
SMILES: N(=Nc1ccccc1(
Similarity: 0.925

Experimental value: Muta
Predicted value: Mutager



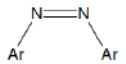
CAS: 84434-42-4
Dataset id: 1799 (Trainir
SMILES: N(=Nc1ccc(N)c
Similarity: 0.924

Experimental value: Muta
Predicted value: Mutager

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts:

Fragment found: SA29 Aromatic diazo

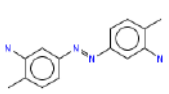


Ar = Any aromatic/heteroaromatic ring

- Chemicals with a sulfonic acid group (-SO3H) on both rings linked to the diazo group are excluded.

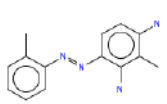
Aromatic diazo. If a sulfonic acid group (-SO3H) is present on each of the rings that contain the diazo group, the substance should be not classified.

Following, the most similar compounds from the model's dataset having the same fragment.



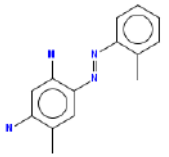
CAS: 119631-00-4
Dataset id: 3482 (Training set)
SMILES: N(=Nc1ccc(c(N)c1)C)c2ccc(c(N)c2)C
Similarity: 0.977

Experimental value: Mutagenic
Predicted value: Mutagenic



CAS: 84434-45-7
Dataset id: 2696 (Test set)
SMILES: N(=Nc1ccccc1C)c2ccc(N)c(c2(N))C
Similarity: 0.951

Experimental value: Mutagenic
Predicted value: Mutagenic



CAS: 7467-29-0
Dataset id: 3021 (Training set)
SMILES: N(=Nc1cc(c(N)cc1(N))C)c2ccccc2C
Similarity: 0.949

Experimental value: Mutagenic
Predicted value: Mutagenic

Domain of the model.

nental value in the training set have been found.

nd in the training set is good.

experimental values that agree with the predicted

le the descriptor range of the compounds of the

ve been found in the compounds of the training

ng this aspect.

d be reviewed by an expert.

rding this aspect.

Predictive models – ADI parameters

- **Visualization of similar substances**
- **Similarity index** *(chemical; sub-indices)*
- **Chemometric check** *(descriptor space)*
- **Atom centered-fragment** *(chemical)*

Chemical input space

- **Check of the descriptor sensitivity** *(algorithm)*
- **Uncertainty** *(algorithm)*

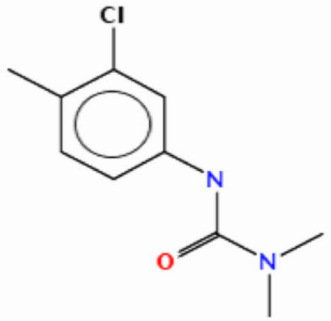
Characteristics of the algorithm

- **Fragments for outliers** *(output space)*
- **Prediction Accuracy** *(output space)*
- **Prediction Concordance** *(tox) exploration)*

Input and output toxicological space

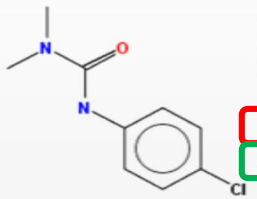
Predictive models – ADI accuracy

Prediction for the compound no. 1: CN(C)C(=O)NC1=CC(Cl)=C(C)C=C1



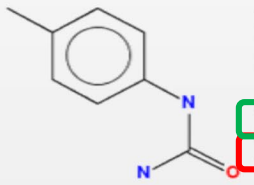
Carcinogenic: Non-Positive
Class indices: Positive=0.079, Non-Positive=0.921
Remarks for the prediction:

The following chemicals similar to the query compound have been identified in the CAESAR database:



Dataset id: 162
SMILES: Clc1ccc(cc1)NC(=O)N(C)C
Similarity: 0.884

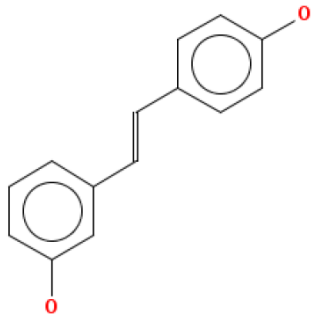
Experimental class: Positive
Predicted class: Non-Positive



Dataset id: 754
SMILES: O=C(Nc1ccc(cc1)C)N
Similarity: 0.857

Experimental class: Non-Positive
Predicted class: Positive

Predictive models – ADI concordance



Prediction: Reliability:

Prediction is Possible NON-Mutagenic, but the result shows some critical aspects, which require to be checked:
- similar molecules found in the training set have experimental values that disagree with the predicted value

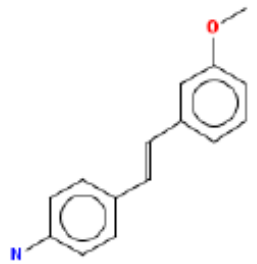
Global AD Index
AD index = 0.719
Explanation: the predicted compound could be out of the Applicability Domain of the model.

Similar molecules with known experimental value
Similarity index = 0.901
Explanation: strongly similar compounds with known experimental value in the training set have been found.

Accuracy of prediction for similar molecules
Accuracy index = 1
Explanation: accuracy of prediction for similar molecules found in the training set is good.

Concordance for similar molecules
Concordance index = 0.33
Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value.

Atom Centered Fragments similarity check
ACF index = 1
Explanation: all atom centered fragment of the compound have been found in the compounds of the training set.

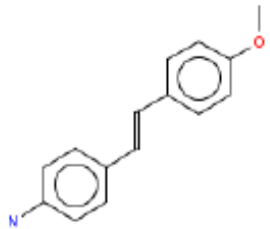


Compound #1

CAS: 154028-32-7
Dataset id: 2989 (Training set)
SMILES: O(c2cccc(C=Cc1ccc(N)cc1)c2)C
Similarity: 0.907

Experimental value: Mutagenic
Predicted value: Mutagenic

Alerts (not found in the target): SM44; SM104

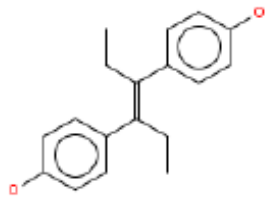


Compound #2

CAS: 7570-37-8
Dataset id: 1345 (Training set)
SMILES: O(c1ccc(cc1)C=Cc2ccc(N)cc2)C
Similarity: 0.905

Experimental value: Mutagenic
Predicted value: Mutagenic

Alerts (not found in the target): SM44; SM104



Compound #3

CAS: 56-53-1
Dataset id: 2731 (Test set)
SMILES: Oc1ccc(cc1)C(=C(c2ccc(O)cc2)CC)CC
Similarity: 0.893

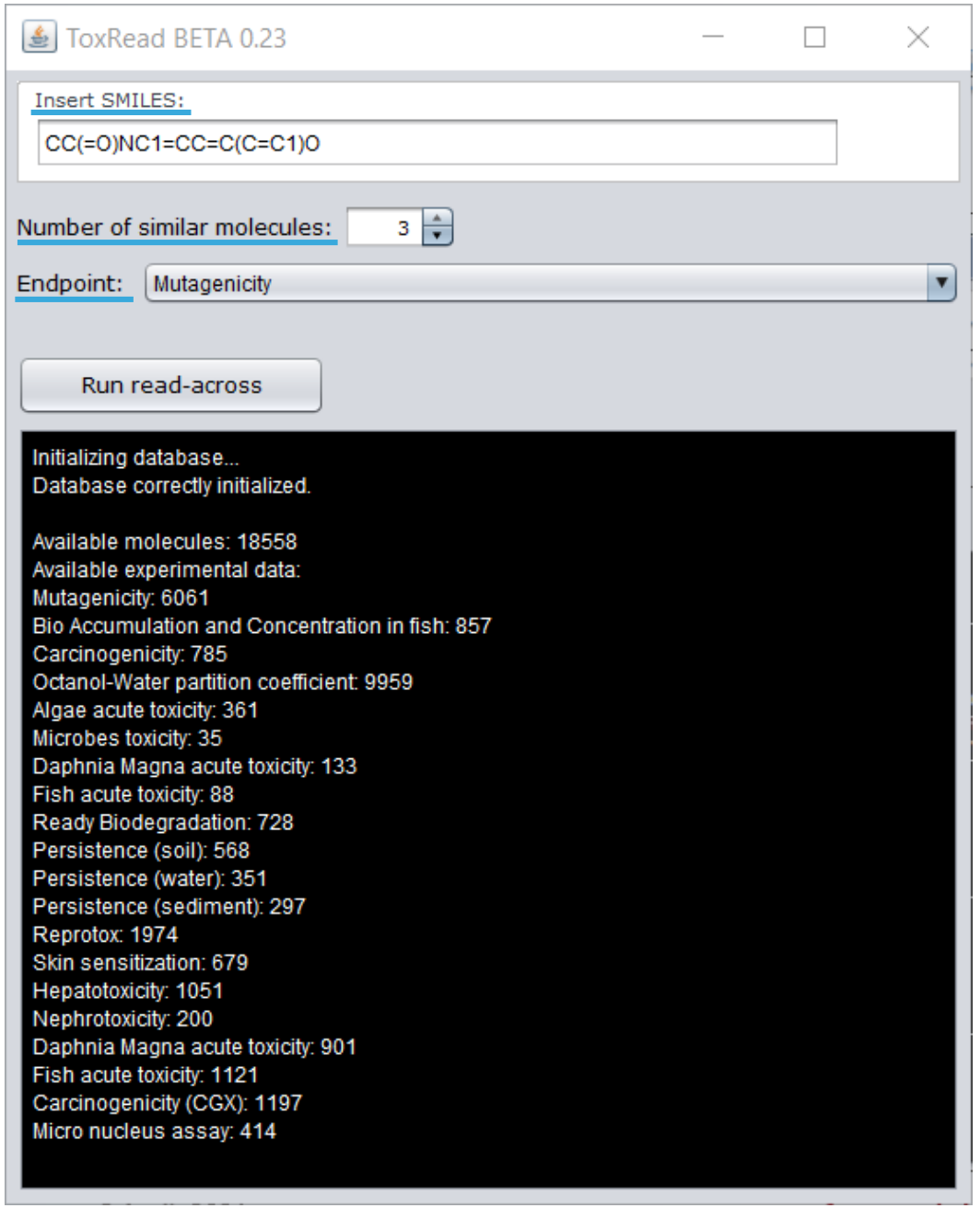
Experimental value: NON-Mutagenic

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH
LIFE17 GIE/IT/000461

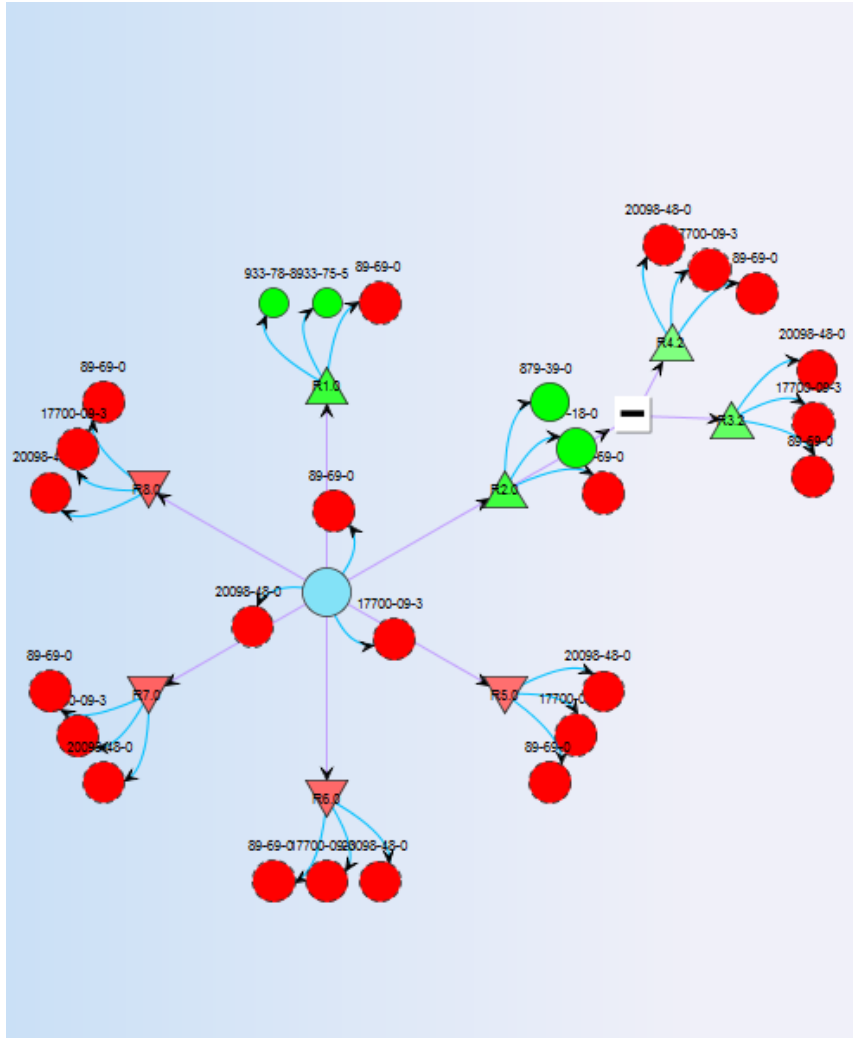
Read across tool – Examples

toxRead

- More than 18500 molecules
- 20 endpoints
- Expert based rules/statistic based rules

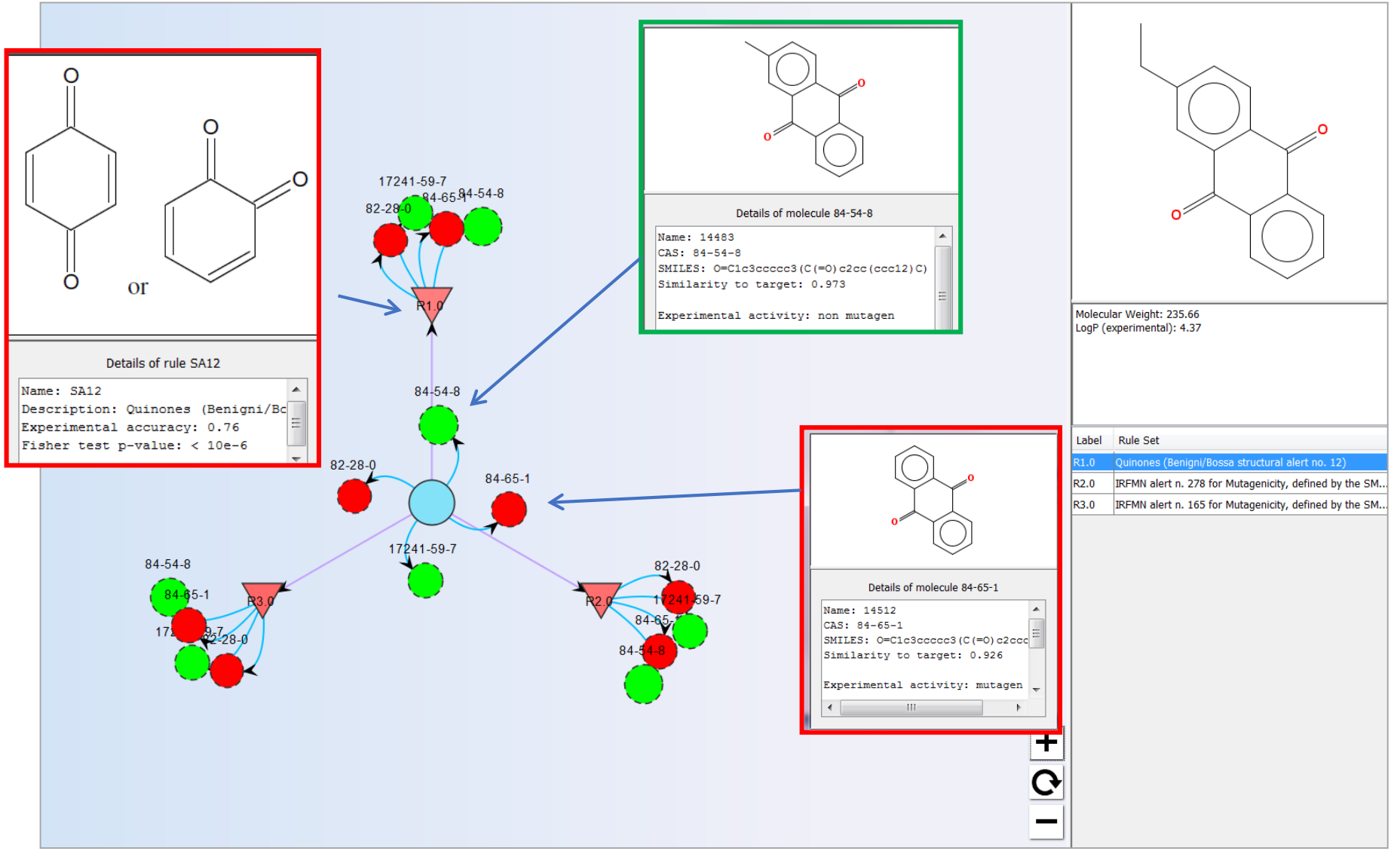


Read across tool – User interface

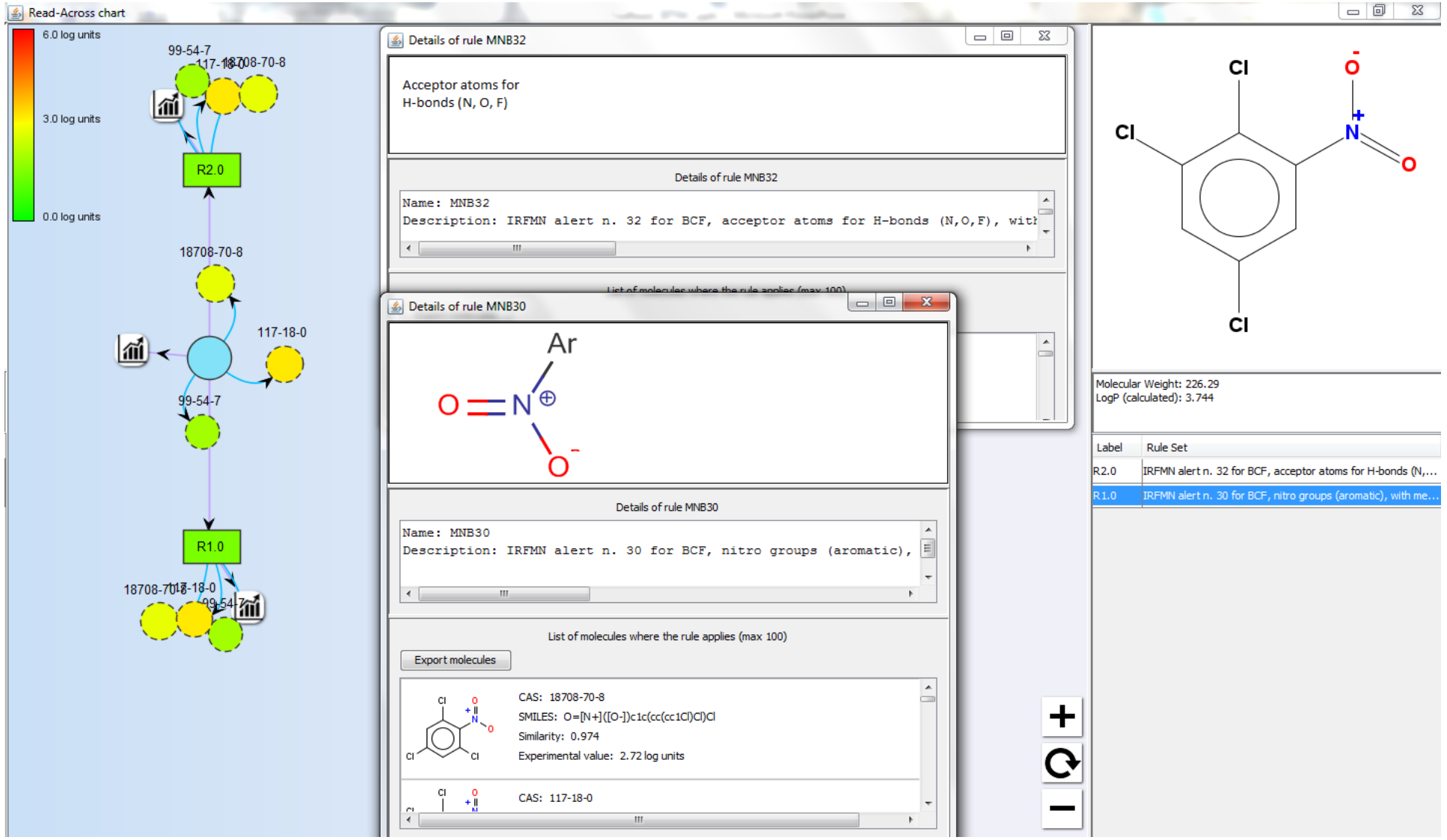


- Target molecule in the center of the visualization panel
- Target compound directly connected to most similar molecules (in inner circle)
- Structural alerts in the triangles
- Paths connect molecules sharing the same structural alert
- Shape: circles are molecules, triangles are structural alerts
- Circle dimension: related to similarity
- Color: red or green with different saturation indicates active or non active at different levels
- Clicking on nodes shows structure, explanation, etc.

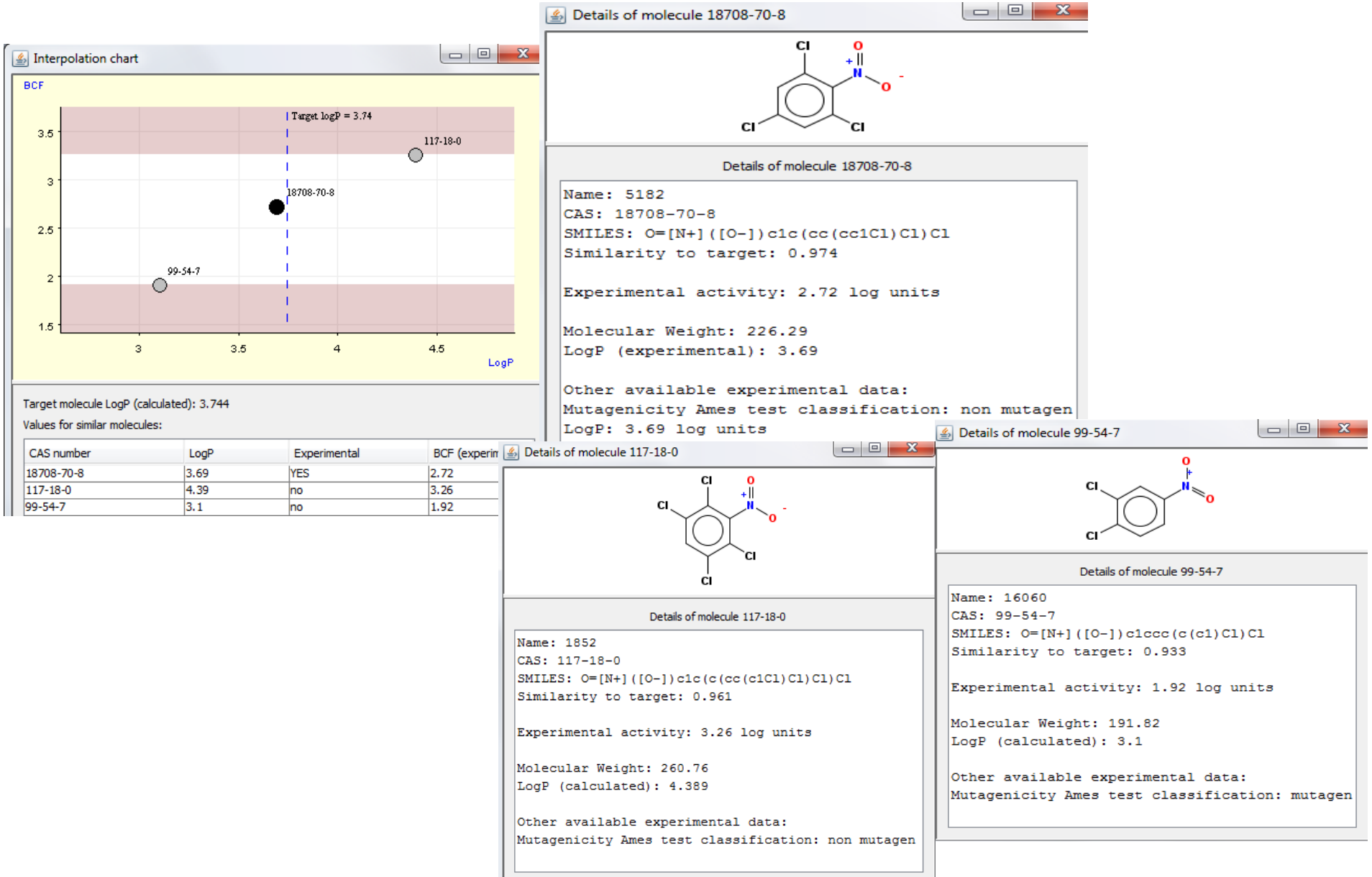
Read across tool – A closer look



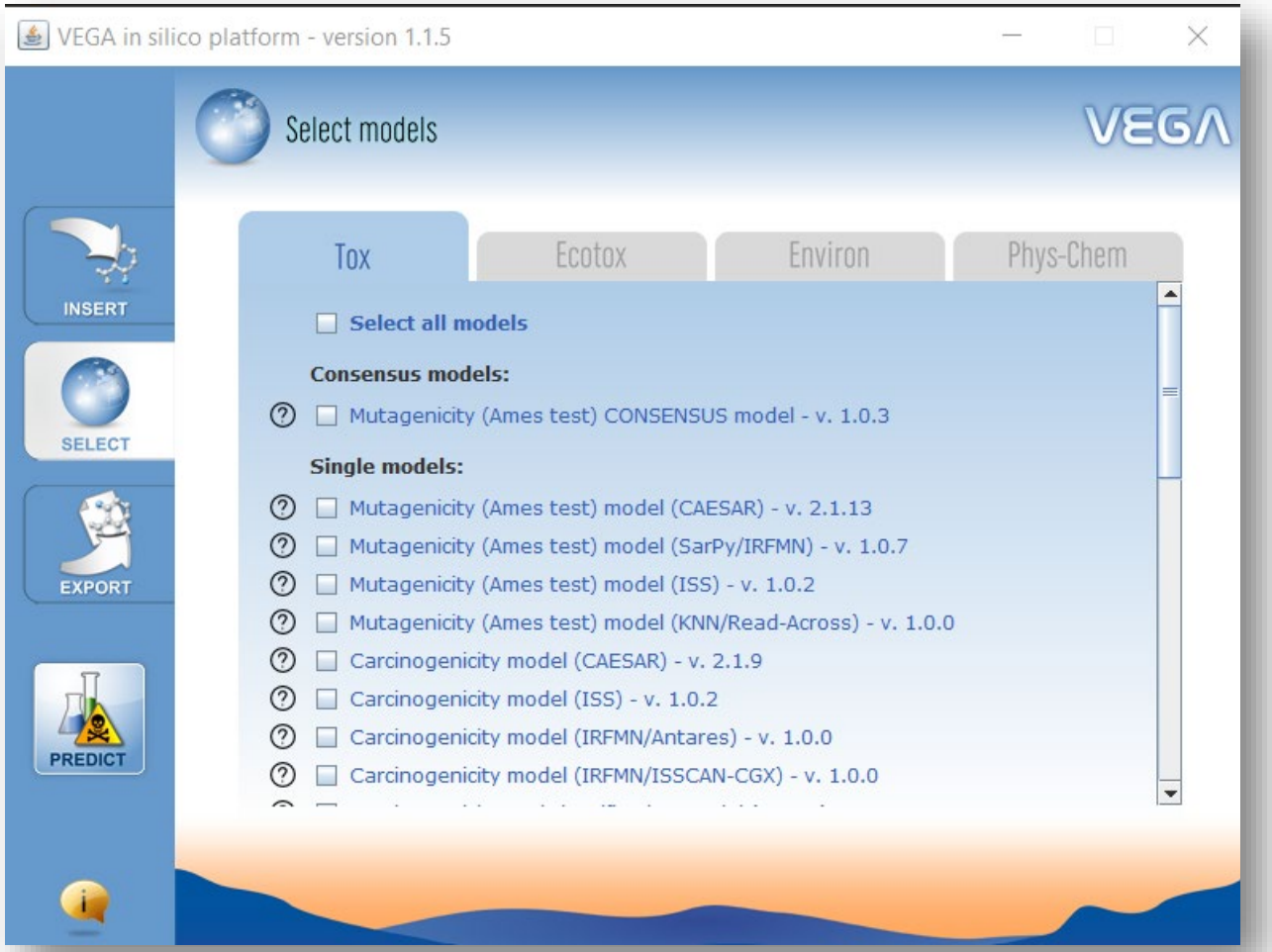
Read across tool – Quantitative endpoints: graph



Read across tool – Quantitative endpoints (e.g., BCF)



Predictive models – Novelties. New models

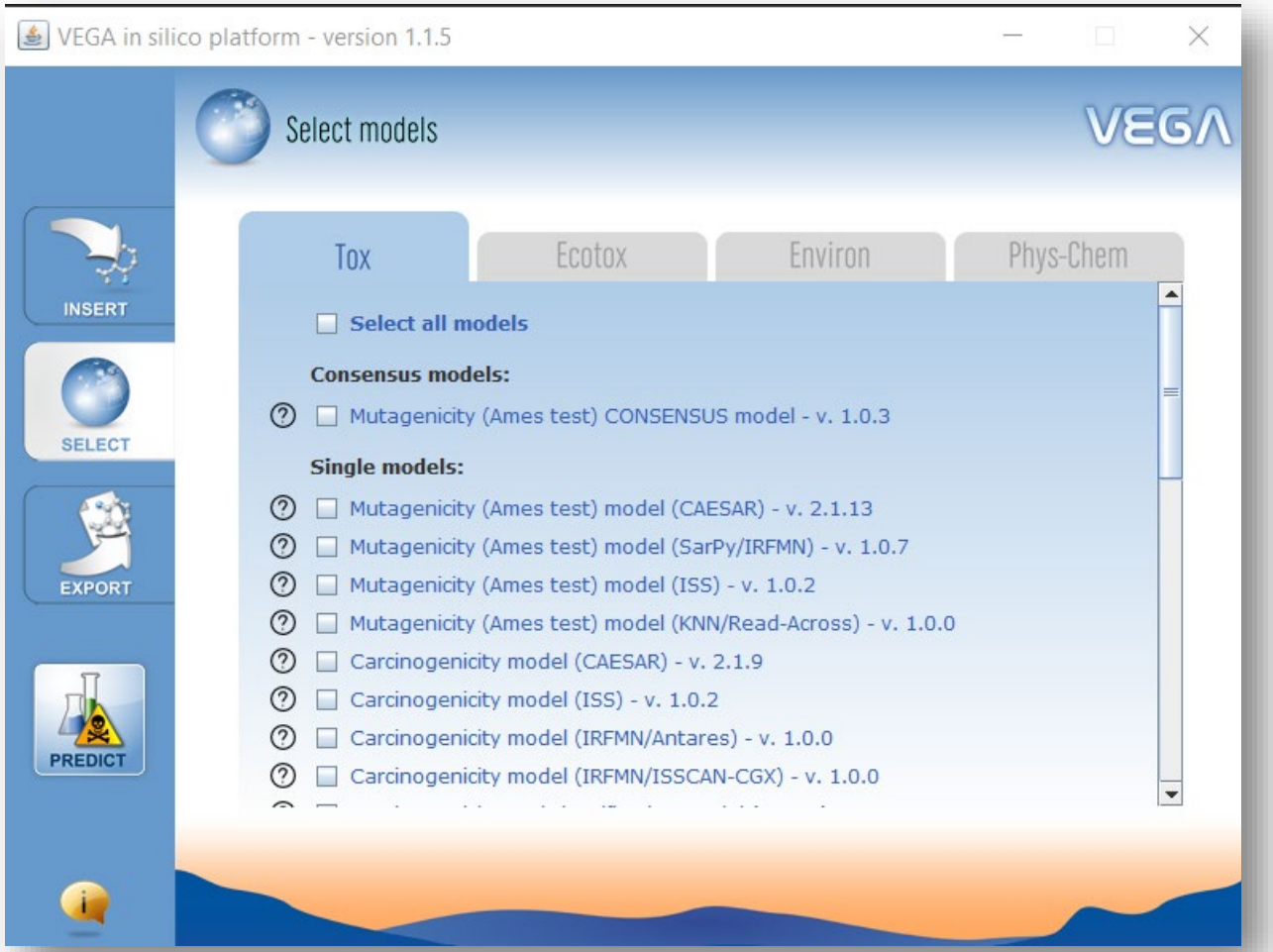


- **18** new models implemented in Vega

- **24** new models already developed – Implementation ongoing

- New concept of similarity

Predictive models – Novelties. QMRF-QPRF



- **18** new models implemented in Vega

- Improvements of the output to facilitate QPRF
- Revision of QMRFs
- integration of the models for the same endpoint

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH
LIFE17 GIE/IT/000461

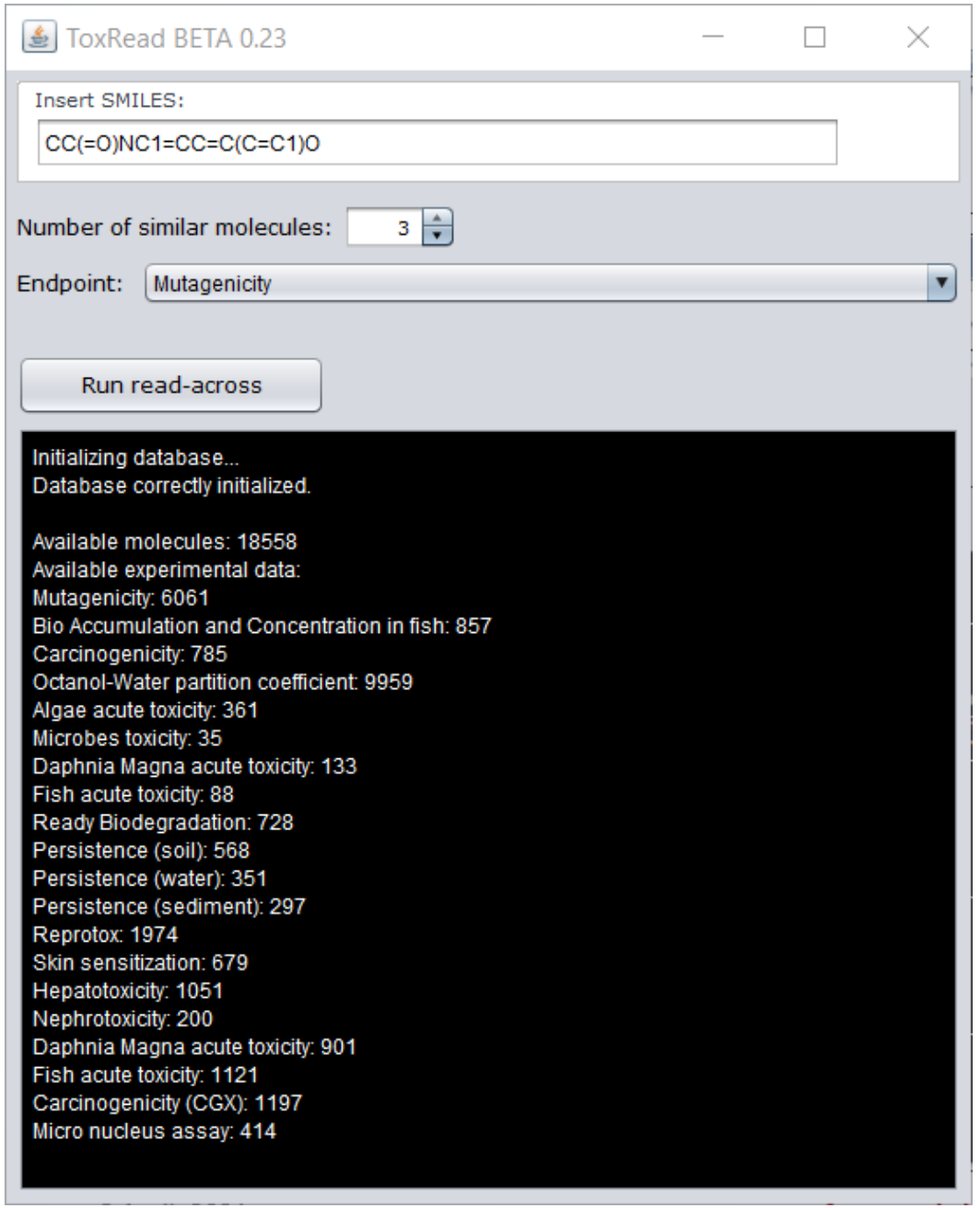
Read across tool – Novelties

❖ New sets of rules available for implementation:

- in vivo Miconucleus assay
- Skin Irritation
- Eye Irritation
- LD50, rat
- Androgen receptor

❖ New sets of rules under evaluation:

- Respiratory sensitization in human
- Algae acute
- Terrestrial toxicity – eartworm
- Skin sens



CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461

In the future, a closer link

toxRead VS VEGA

No report

No batch mode

For exploration/analysis

More rules

More substances

All in the same page

Thank you!