I modelli QSAR in VEGA per quali endpoints ed il loro utilizzo.

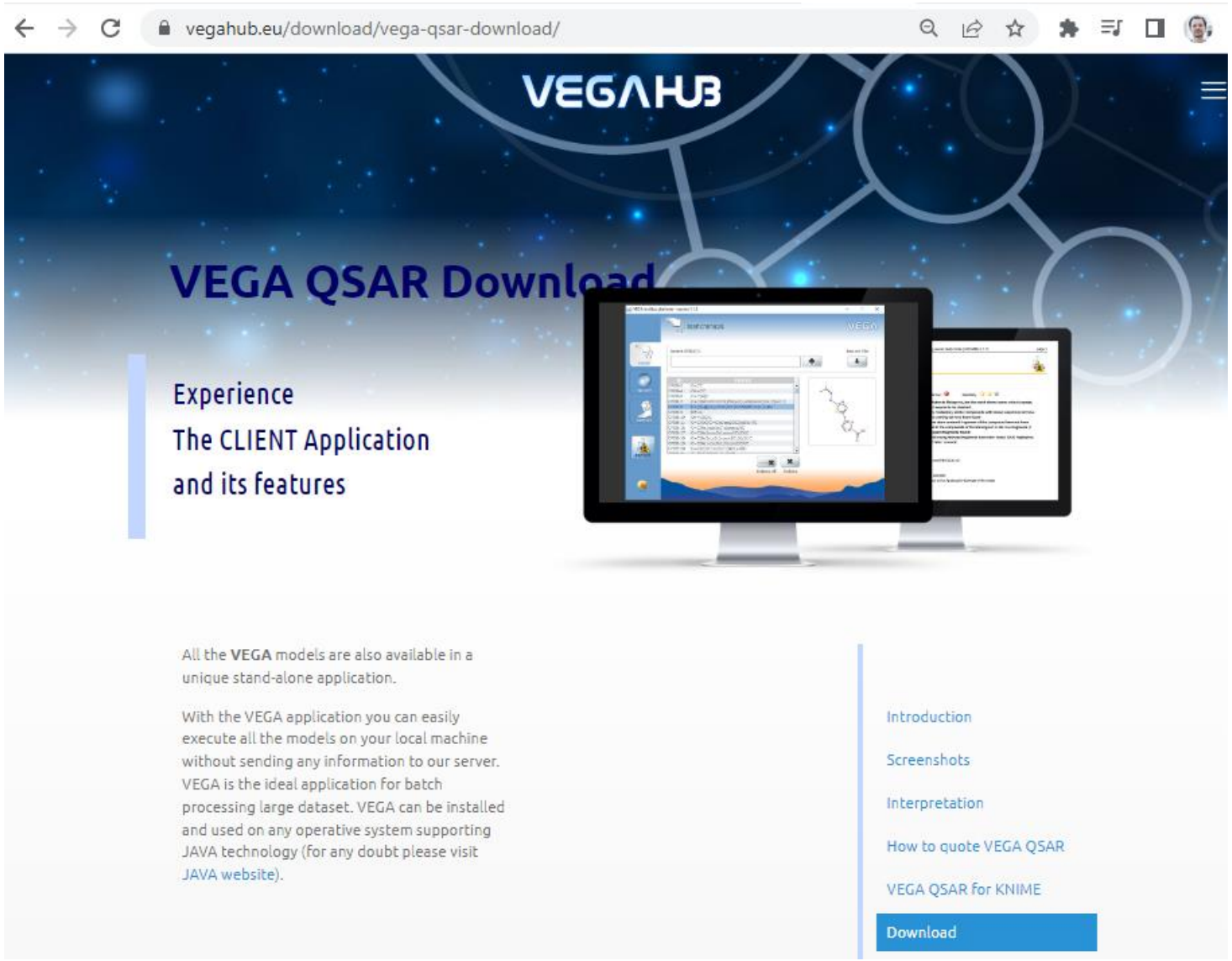
CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461

VEGA

- Applicazione stand-alone che contiene numerosi **modelli QSAR** per predire proprietà ed attività biologiche a partire dalla struttura di un composto
- Genera dei **report completi** (file PDF) e riassunti dei valori in formato di testo
- Open source e **gratuita**, si scarica dal sito: hiips://www.vegahub.eu/
- Esecuzione in locale (non invia dati a server esterni)
- Ultima versione: **112 modelli disponibili**

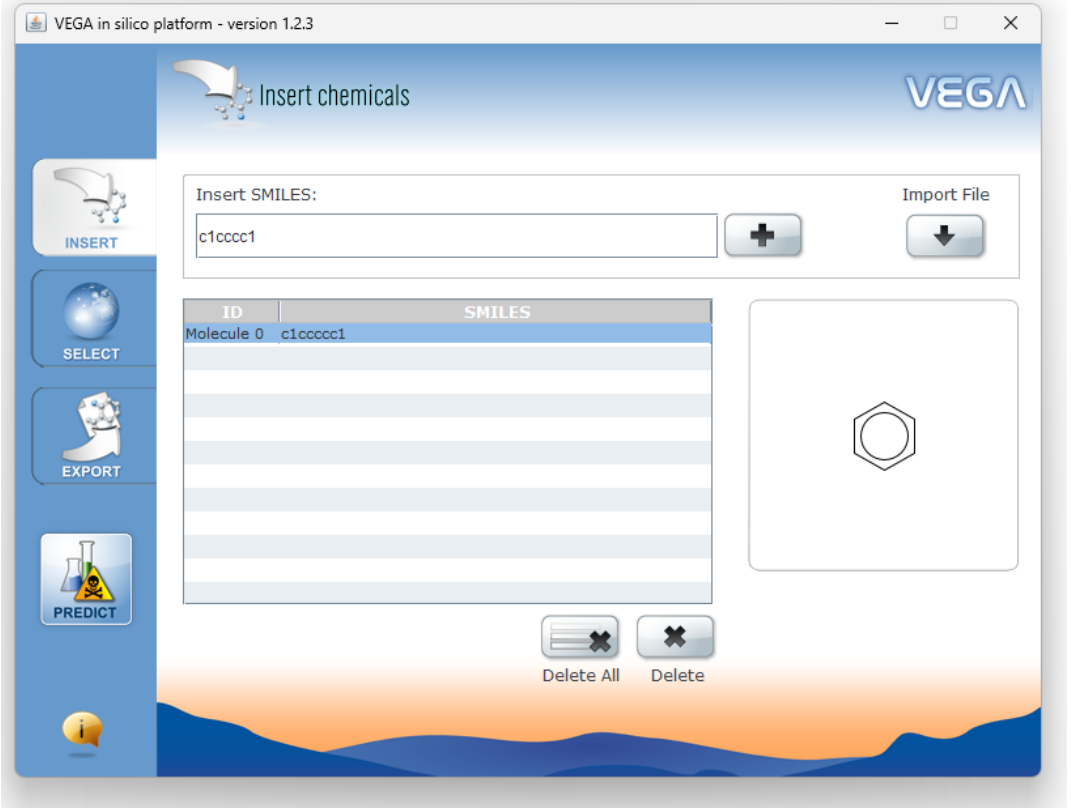
Il sito VegaHub



Funzionamento

Facile utilizzo, l'utente deve inserire le strutture delle molecole di interesse (ad esempio in semplice formato SMILES), selezionare i modelli desiderati e lanciare i calcoli.

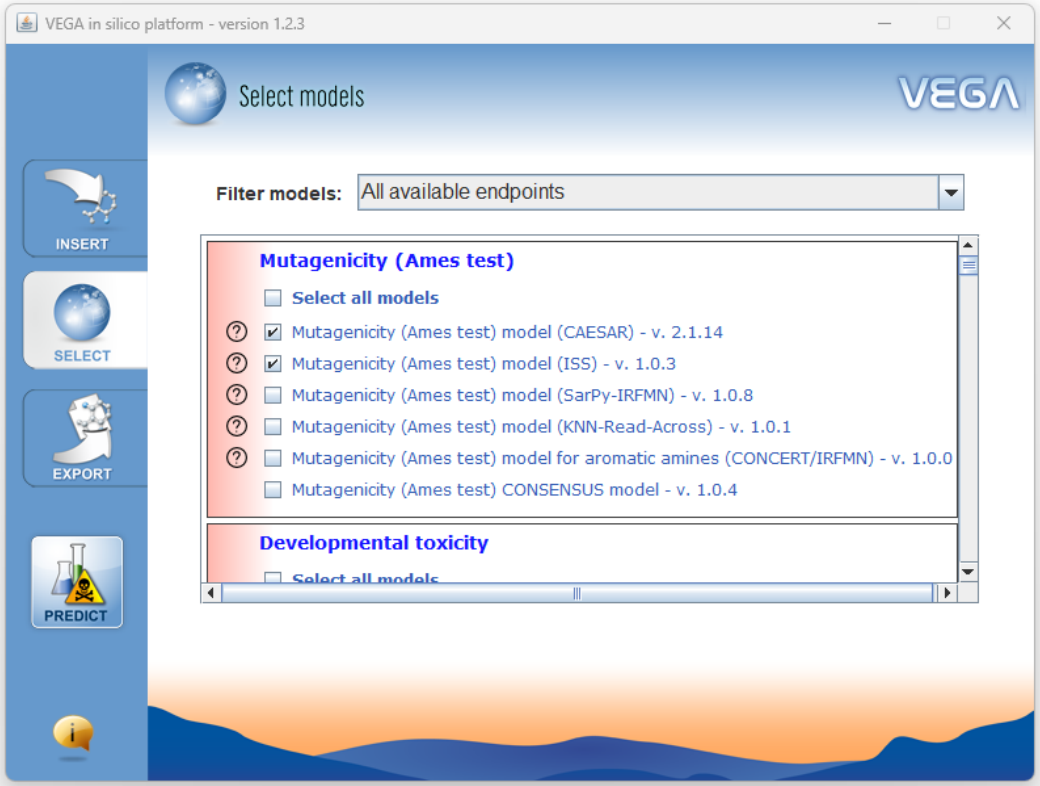
Il software salva nella directory desiderata i risultati nel formato selezionato (report completo PDF e/o sommario dei risultati in formato di testo)



I modelli

Modelli suddivisi in macro-sezioni:

- **Human toxicity**
- **Ecotoxicity**
- **Fate & Distribution**
- **Physical-chemical properties**
- **Human & ecological PBPK**

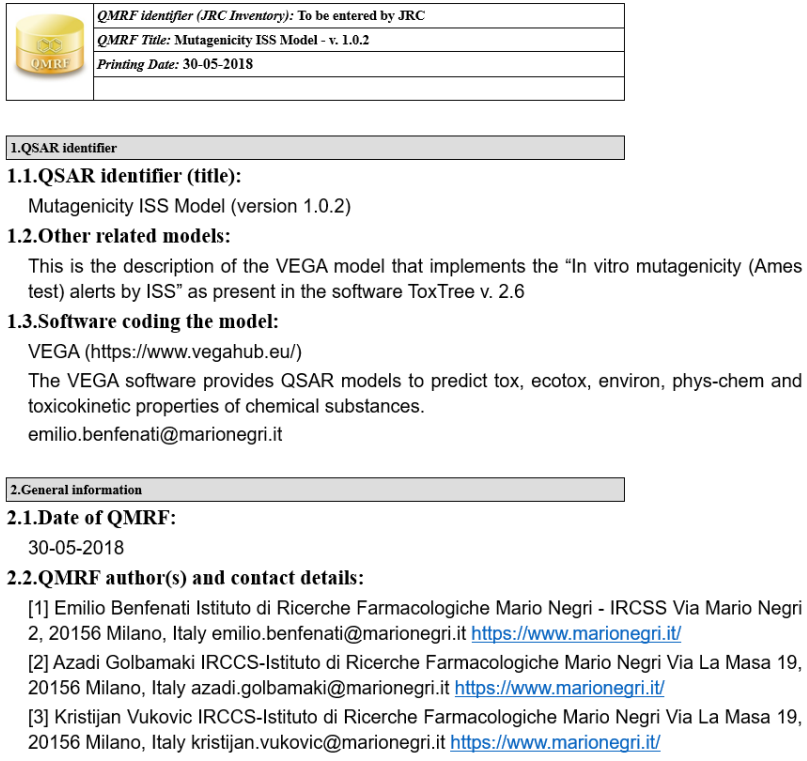


Per molti endpoint sono disponibili più modelli QSAR (sviluppati con approcci differenti)

I modelli

I modelli sono accompagnati dai **QMRF** (QSAR Model Reporting Format), che riportano tutti i dettagli scientifici riguardanti la definizione del modello - importanti per **l'accettazione** del modello ed il suo **utilizzo a fini normativi** - quali:

- Definizioni degli algoritmi del modello
- Origine dei dati sperimentali usati nel modello
- Definizione del dominio di applicabilità
- Riferimenti scientifici (pubblicazioni)



QMRF	*QMRF identifier (JRC Inventory):* **To be entered by JRC**
	QMRF Title: **Mutagenicity ISS Model - v. 1.0.2**
	Printing Date: **30-05-2018**

1.QSAR identifier

1.1.QSAR identifier (title):

Mutagenicity ISS Model (version 1.0.2)

1.2.Other related models:

This is the description of the VEGA model that implements the "In vitro mutagenicity (Ames test) alerts by ISS" as present in the software ToxTree v. 2.6

1.3.Software coding the model:

VEGA (https://www.vegahub.eu/)

The VEGA software provides QSAR models to predict tox, ecotox, environ, phys-chem and toxicokinetic properties of chemical substances.

emilio.benfenati@marionegri.it

2.General information

2.1.Date of QMRF:

30-05-2018

2.2.QMRF author(s) and contact details:

[1] Emilio Benfenati Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy emilio.benfenati@marionegri.it https://www.marionegri.it/

[2] Azadi Golbamaki IRCCS-Istituto di Ricerche Farmacologiche Mario Negri Via La Masa 19, 20156 Milano, Italy azadi.golbamaki@marionegri.it https://www.marionegri.it/

[3] Kristijan Vukovic IRCCS-Istituto di Ricerche Farmacologiche Mario Negri Via La Masa 19, 20156 Milano, Italy kristijan.vukovic@marionegri.it https://www.marionegri.it/

Endpoint disponibili

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

Human Toxicity

Mutagenicity (Ames test)	Estrogen and Androgen Receptor-mediated effect
Developmental/Reproductive Toxicity	Androgen Receptor-mediated effect
Carcinogenicity	Thyroid Receptor
Acute Toxicity (LD50)	Glucocorticoid Receptor
Skin Sensitization	Thyroperoxidase Inhibitory Activity
Skin Irritation model	Endocrine Disruptor activity
Eye Irritation	NOAEL and LOAEL (generale e fegato)
Chromosomal aberration	Cramer classification
In vitro / in vivo Micronucleus	Hepatotoxicity

Endpoint disponibili

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

Ecotoxicity

BCF	Algae Chronic Toxicity
Fish Acute Toxicity	Verhaar classification
Fathead Minnow Acute Toxicity	MOA fish toxicity
Daphnia Magna Acute Toxicity	Bee acute toxicity
Guppy Acute Toxicity	Earthworm Toxicity
Algae Acute Toxicity	Sludge (EC50) Toxicity
Fish Chronic Toxicity	Zebrafish embryo AC50
Daphnia Magna Chronic Toxicity	

Endpoint disponibili

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

Fate & Distribution

Ready Biodegradability
Persistence (sediment) qualitative
Persistence (soil) qualitative
Persistence (water) qualitative
Persistence (sediment) quantitative half-life
Persistence (soil) quantitative half-life
Persistence (water) quantitative half-life
Air Half-Life

Endpoint disponibili

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

Physical-Chemical properties and PBPK

LogP	Plasma Protein Binding
Water solubility	Aromatase activity
Vapour Pressure	P-Glycoprotein activity
Melting Point	Hepatic Steatosis MIE assays
Hydrolysis	Skin Permeation
Henry's Law	Adipose tissue-blood partition
KOA	Total body elimination half-life
KOC	kM/Half-Life

L’output di VEGA

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

- I modelli generano un **report completo in PDF**, contenente numerose informazioni importanti per valutare la predizione stessa, e da usare come materiale di supporto per utilizzi normativi
- Le predizioni QSAR non devono essere usate con un approccio “**black box**”, ma analizzate alla luce di tutte le informazioni di accompagnamento - anche al fine di considerare più modelli per la stessa proprietà, ad esempio usando modelli meccanicistici (frammenti/regole) insieme a modelli puramente statistici

L’output di VEGA

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

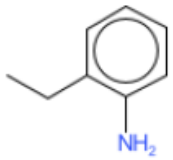
I report di VEGA contengono molte informazioni fondamentali per interpretare ed usare i risultati, come:

- Valore finale dell’**Applicability Domain**, scomposto in diversi sotto elementi in modo da comprendere con precisioni eventuali motivi di predizioni poco affidabili
- **Molecole più simili** a quella predetta, disponibili nel dataset del modello e quindi con valore sperimentale disponibile (anche per eventuale approccio read-across)
- **Spiegazione di frammenti/regole** identificate nella molecola predetta

Report PDF

1. Prediction Summary

Prediction for compound Molecule 1 -



Prediction: Reliability:

Prediction is Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value

The following alerts have been found: SA28 Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions)

Compound: Molecule 1
Compound SMILES: Nc1ccccc1CC
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural Alerts: SA28 Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions)
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

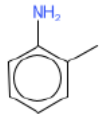
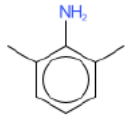
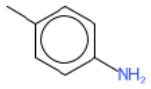
CONCERTREACH CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH
LIFE17 GIE/IT/000461

Report PDF

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

3.1 Applicability Domain:

Similar Compounds, with Predicted and Experimental Values

	Compound #1 CAS: 636-21-5 Dataset id:32 (Training Set) SMILES: Nc1ccccc1C Similarity: 0.948 Experimental value : NON-Mutagenic Predicted value : Mutagenic Alerts (found also in the target): SA28 Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions)
	Compound #2 CAS: 87-62-7 Dataset id:70 (Training Set) SMILES: Nc1c(cccc1C)C Similarity: 0.945 Experimental value : NON-Mutagenic Predicted value : NON-Mutagenic
	Compound #3 CAS: 540-23-8 Dataset id:632 (Training Set) SMILES: Nc1ccc(cc1)C Similarity: 0.936 Experimental value : Mutagenic Predicted value : Mutagenic Alerts (found also in the target): SA28 Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions)

Report PDF

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

3.2 Applicability Domain:

Measured Applicability Domain Scores

✗	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.946 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.499 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✗	Concordance for similar molecules Concordance index = 0 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..

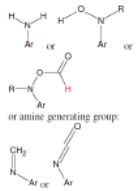
Report PDF

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

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 1) Reasoning on fragments/structural alerts :.

Fragment found: SA28 Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions)

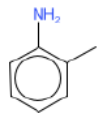


Ar = Any aromatic/heteroaromatic ring
R= Any atom/group

- Chemicals with ortho-disubstitution, or with an ortho carboxylic acid substituent are excluded.
- Chemicals with a sulfonic acid group (-SO3H) on the same ring of the amino group are excluded .

Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions). However: Aromatic amino groups with ortho disubstitutions or with a carboxylic acid substituent in ortho position are excluded. If a sulfonic acid group (-SO3H) is present on the ring that contains also the amino group, the substance should be excluded from the alert. The following structures should also be included: O=C=NC1=CC=CC=C1 and C([H])([H])=NC1=CC=CC=C1. The possibility that the Nitrogen atom of hydroxyl amine is part of a cycle, should be excluded.

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



CAS: 636-21-5
Dataset id:32 (Training Set)
SMILES: Nc1ccccc1C
Similarity: 0.948

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA28 Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions)

GRAZIE

CONCERTREACH
CONCERTING EXPERIMENTAL DATA AND IN SILICO MODELS FOR REACH

LIFE17 GIE/IT/000461