	QMRF identifier (JRC Inventory): **To be entered by JRC**
	QMRF Title: **Mutagenicity (Ames test) CONSENSUS model - v. 1.0.4**
	Printing Date: **14-04-2022**

1.QSAR identifier

1.1.QSAR identifier (title):

Mutagenicity (Ames test) CONSENSUS model (version 1.0.4)

1.2.Other related models:

The model provides a qualitative prediction of mutagenicity on *Salmonella typhimurium* (Ames test), applying a consensus approach based on the four QSAR models currently available in VEGA. It is implemented inside the VEGA online platform, accessible at: www.vegahub.eu. See QMRF of the 4 individual models as further reference.

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:

April 2022

2.2.QMRF author(s) and contact details:

[1] Alessio Gamba Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy alessio.gamba@marionegri.it https://www.marionegri.it/

[2] Giuseppa Raitano Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy giuseppa.raitano@marionegri.it https://www.marionegri.it/

2.3.Date of QMRF update(s):

2.4.QMRF update(s):

Version 1.0.1: changed weights for Applicability Domain conversion.

Version 1.0.2: changed the main algorithm when experimental values are found.

2.5.Model developer(s) and contact details:

[1] Alberto Manganaro Istituto di Ricerche Farmacologiche Mario Negri - IRCSS Via Mario Negri 2, 20156 Milano, Italy alberto.manganaro@marionegri.it https://www.marionegri.it/

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

2.6.Date of model development and/or publication:

2014

2.7.Reference(s) to main scientific papers and/or software package:

[1] Evaluation of QSAR models for the prediction of Ames genotoxicity: a retrospective exercise on the chemical substances registered under the EU REACH regulation. Antonio Cassano, Giuseppa Raitano, Enrico Mombelli, Alberto Fernández, Josep Cester, Alessandra Roncaglioni, Emilio Benfenati. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2014; 32(3): 273–298. doi: 10.1080/10590501.2014.938955

2.8.Availability of information about the model:

The model is non-proprietary and the training set is available.

2.9.Availability of another QMRF for exactly the same model:

Another QMRF is not available.

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Histidine-dependent strains of *Salmonella typhimurium* (Ames test)

3.2.Endpoint:

TOX 7.6.1. Genetic toxicity in vitro

3.3.Comment on endpoint:

Mutagenic toxicity is the capacity of a substance to cause genetic mutations. This property is of high public concern because it has a close relationship with carcinogenicity and eventually reproductive toxicity: most of the mutagenic substances are suspected carcinogenic substance in case a genotoxic mechanism is considered. The Ames test is the basic in vitro assay to detect mutagens. The relevant test guideline covering this endpoint is OECD TG 471. The training set is based on test results from either the original version of the test guideline from 1983 or a newer version from 1997. The endpoint covers the DNA base-pair substitution and frameshift mutagenic mechanisms that are covered by the Ames tester strains: TA 1535, TA100, TA 98, and TA 1537 or TA97 or TA 97a. A part of the training set data additionally covers cross-linking mutagenic events measured by the inclusion of the E.coli WP2 or E.coli WP2 (pKM101) or TA 102 test strains. The test strains for DNA cross-links were included in the 1997 guideline update. As the training set does not systematically cover DNA cross-links, mutagenic substances acting by this mechanism may be under-predicted.

The endpoint is measured on the parent compound and the metabolites generated in vitro by the employed S9 mix of enzyme-induced rodent liver homogenates. In a few cases, liver homogenates from hamsters may have been used.

3.4.Endpoint units:

Adimensional

3.5.Dependent variable:

Binary classification as: Mutagenic / Non-Mutagenic

3.6.Experimental protocol:

Based on the OECD 471 test guideline. Ames test is an in vitro model of chemical mutagenicity and consists of a range of bacterial strains that together are sensitive to a large array of DNA-damaging agents.

3.7.Endpoint data quality and variability:

The estimated inter-laboratory reproducibility rate of S. typhimurium test data is 85% [ref.3, sect.9.2]

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

The model performs a consensus assessment based on the predictions of the available VEGA mutagenicity models (CAESAR, SARpy, ISS and KNN). See the individual QMRFs [3].

The development of the original version of the model is described in the article: Antonio Cassano, Giuseppa Raitano, Enrico Mombelli, Alberto Fernández, Josep Cester, Alessandra Roncaglioni, Emilio Benfenati - Evaluation of QSAR models for the prediction of Ames genotoxicity: a retrospective exercise on the chemical substances registered under the EU REACH regulation, J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2014; 32(3): 273–298. doi: 10.1080/10590501.2014.938955. In the original version of the model, the consensus model was elaborated, and its performance evaluated, by integrating the output of the three models (not four as in the last version) implemented within the VEGA platform (i.e., CAESAR, SARPY, and ISS) through the ADI provided by the software.

The equation for the output of the consensus model was the following:

CONSENSUS = ((±1)*ADcaesar + (±1) *ADsarpy + (±1) *ADiss)/(ADcaesar + ADsarpy + ADiss)

In the equation, the variables ADcaesar, ADsarpy, and ADiss represent the ADI of the CAESAR, SARPY, and ISS models respectively. In the numerator of the equation, each ADI is multiplied by +1 if the prediction is positive whereas it is multiplied by -1 if the prediction is negative. If all the ADI values of the models are zero, the more frequent value among the outputs of the models is taken as the consensus output. Finally, a threshold of -0.1 has been selected to obtain, from Equation 5, a binary output.

4.2. Explicit algorithm:

The consensus algorithm uses the Applicability Domain assessment of each single model's prediction as its weight, so that the final assessment will be more influenced by the single models that produced more reliable predictions.

The Applicability Domain assessment of each model is converted to a numerical value in the range [from 0 to 1] with the following scheme:

AD Assessment	*Value / Weight/Index Reliability*
Experimental value	1.0
High reliability	0.9
Moderate reliability	0.6
Low reliability	0.2

For each prediction class (i.e. mutagenic or non-mutagenic), a score is calculated as the sum of the weights for each model that produced that prediction. The calculated score is normalized over the number of used models, so that it has a theoretical [0..1] range.

So, the consensus score (CS) is calculated separately for the two outcomes (Mutagenic consensus score CS_M and non-mutagenic consensus score CS_{nM}) as follow:

$$\begin{cases} CS_M = \dfrac{\sum IR_M}{nr.tot\ M\ models} \\ CS_{nM} = \dfrac{\sum IR_{nM}}{nr.tot\ nM\ models} \end{cases}$$

where IR is the index reliability of each model (see table above) related to positive prediction IR_M or negative prediction IR_{nM}. Both consensus scores are normalized respectively for the total number of model that predict mutagenic ($nr.tot\ M\ models$) and non-mutagenic ($nr.tot\ nM\ models$)

The final class assignment is done according to the value of the CS: the compound is assigned to the positive class if $CS_M \geq CS_{nM}$.

If at least one experimental value is available, the CS is calculated as 1 if all values are concordant or as a ratio depending on the prevalence of the experimental responses. In this case, the reported number of models used refers only to the number of models having an experimental value.

For the purpose of this consensus approach, the “suspect mutagenic / non-mutagenic” predictions are considered simply as “mutagenic / non-mutagenic” predictions.

In general, in the calculation of the consensus score all the predictions of the individual models are taken into consideration: no threshold is applied to the ADI value of the single predictions just as there is no threshold for the consensus itself.

With this approach, the score of the final prediction can be used as a measure of the reliability of the produced consensus assessment. Indeed, the score would achieve its maximum value (1) only if one or more models found experimental values and these values are in agreement. In all other cases, the score will result in lower values.

4.3.Descriptors in the model:

NA

4.4.Descriptor selection:

The input for this model is the predictions of the other models so no other types of descriptors are directly used by the consensus.

4.5.Algorithm and descriptor generation:

NA

4.6.Software name and version for descriptor generation:

NA

4.7.Chemicals/Descriptors ratio:

NA

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

With this approach, the score of the final prediction can be used as a measure of the reliability of the produced consensus assessment. Indeed, the score would achieve its maximum value (1) only if one or more models found experimental values and these values are in agreement. In all other cases, the score will result in lower values.

5.2.Method used to assess the applicability domain:

See 4.2

5.3.Software name and version for applicability domain assessment:

VEGA (www.vegahub.eu)

5.4.Limits of applicability:

The model is not applicable to inorganic chemicals and substances containing unusual elements (i.e., different from C, O, N, S, P, Cl, Br, F, I). Salts can be predicted only if converted to the neutralized form.

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

The model performs a consensus assessment based on the predictions of the available VEGA mutagenicity models (CAESAR, SARpy, ISS and KNN) and their training sets are available.

6.2.Available information for the training set:

CAS RN: No

Chemical Name: No

Smiles: No

Formula: No

INChI: No

MOL file: No

NanoMaterial: No

6.3.Data for each descriptor variable for the training set:

The datasets are available for the 4 individual models.

6.4.Data for the dependent variable for the training set:

The datasets are available for the 4 individual models.

6.5.Other information about the training set:

NA

6.6.Pre-processing of data before modelling:

NA

6.7.Statistics for goodness-of-fit:

No statistics are provided since in presence of experimental data the outcome of the consensus will be calculated according to experimental data only. The statistic of each individual model is provided within the relative QMRF

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

NA

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

NA

6.10.Robustness - Statistics obtained by Y-scrambling:

NA

6.11.Robustness - Statistics obtained by bootstrap:

NA

6.12.Robustness - Statistics obtained by other methods:

NA

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

The model performs a consensus assessment based on the predictions of the available VEGA mutagenicity models (CAESAR, SARpy, ISS and KNN) and their individual external validation sets are available as in the relative QMRF. Those external validation sets are composed of sets of data not in common with the training and the test sets of the single models. Those data were selected from a big dataset comprising public and proprietary data [4] [5]. In this case the external validation set is composed of the set of data not in common with all the training and the test sets of the single models.

7.2.Available information for the external validation set:

NA

7.3.Data for each descriptor variable for the external validation set:

NA

7.4.Data for the dependent variable for the external validation set:

NA

7.5.Other information about the external validation set:

The external validation set is composed of 12240 substances, 1619 experimentally positive and 10621 experimentally negative on Ames test.

7.6.Experimental design of test set:

NA

7.7.Predictivity - Statistics obtained by external validation:

Four compounds were not predicted by all the single models, then the available predictions for the statistical assessment were 12236.

	no threshold applied	score <0,3	score >=0,3	score <0,5	>=0,5
TP	1014	340	674	614	400
FN	605	190	415	412	193
FP	3156	1593	1563	2614	542
TN	7461	1678	5783	3792	3669
Tot predicted	12236	3801	8435	7432	4804

accuracy	0,69	0,53	0,77	0,59	0,85
sensitivity	0,63	0,64	0,62	0,60	0,67
specificity	0,70	0,51	0,79	0,59	0,87
MCC	0,24	0,11	0,31	0,13	0,45

7.8.Predictivity - Assessment of the external validation set:

NA

7.9.Comments on the external validation of the model:

The distribution of the external validation dataset is unbalanced: the 87% of the compounds is non mutagenic experimentally.

8.Providing a mechanistic interpretation - OECD Principle 5

8.1.Mechanistic basis of the model:

The model provides a qualitative prediction of mutagenicity on Salmonella typhimurium (Ames test), applying a consensus approach based on the four QSAR models currently available in VEGA. Two models of them (ISS and SARpy) are structural alerts based. In particular, ISS structural alerts are expert based meanwhile SARpy includes statistical alerts to identify both toxic and non-toxic compounds. Thus, the mechanisms associated to the effect can be explored in this way.

8.2.A priori or a posteriori mechanistic interpretation:

The ISS model is based on a prioiri knowledge. The SARpy model is based on a posteriori interpretation.

8.3.Other information about the mechanistic interpretation:

NA

9.Miscellaneous information

9.1.Comments:

NA

9.2.Bibliography:

[1] Evaluation of QSAR models for the prediction of ames genotoxicity: a retrospective exercise on the chemical substances registered under the EU REACH regulation. Antonio Cassano, Giuseppa Raitano, Enrico Mombelli, Alberto Fernández, Josep Cester, Alessandra Roncaglioni, Emilio Benfenati. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2014; 32(3): 273–298. doi: 10.1080/10590501.2014.938955

[2] T. Ferrari, D. Cattaneo, G. Gini, N. Golbamaki Bakhtyari, A. Manganaro, E. Benfenati, “Automatic knowledge extraction from chemical structures: the case of mutagenicity prediction”, SAR and QSAR in Environmental Research (2013), vol. 24 issue 5, 365-83. http://dx.doi.org/10.1080/1062936X.2013.773376

[3] Guide to Mutagenicity Consensus version 1.0.3. Available as software help guide to the models when installing it.

[4] Honma M, Kitazawa A, Cayley A, Williams RV, Barber C, Hanser T, Saiakhov R, Chakravarti S, Myatt GJ, Cross KP, Benfenati E, Raitano G, Mekenyan O, Petkov P, Bossa C, Benigni R, Battistelli CL, Giuliani A, Tcheremenskaia O, DeMeo C, Norinder U, Koga H, Jose C, Jeliazkova N, Kochev N, Paskaleva V, Yang C, Daga PR, Clark RD, Rathman J. Improvement of quantitative structure-activity relationship (QSAR) tools for predicting Ames mutagenicity: outcomes of the Ames/QSAR International Challenge Project. Mutagenesis. 2019 Mar 6;34(1):3-16. doi: 10.1093/mutage/gey031. PMID: 30357358; PMCID: PMC6402315.

[5] Cassano, A.; Raitano, G.; Mombelli, E.; Fernández, A.; Cester, J.; Roncaglioni, A.; Benfenati, E. Evaluation of QSAR Models for the Prediction of Ames Genotoxicity: A Retrospective Exercise on the Chemical Substances Registered Under the EU REACH Regulation. J. Environ. Sci. Health C Environ. Carcinog. Ecotoxicol. Rev. 2014, 32, 273–298. DOI: 10.1080/10590501.2014.938955.

9.3.Supporting information:

All available datasets are present in the model inside the VEGA software.

10.Summary (JRC QSAR Model Database)

10.1.QMRF number:

To be entered by JRC

10.2.Publication date:

To be entered by JRC

10.3.Keywords:

To be entered by JRC

10.4.Comments:

To be entered by JRC