Integrative Prediction Approaches in Genotoxicity / Carcinogenicity:

Structure-Activity Relationships. Part I

Romualdo Benigni

Istituto Superiore di Sanita'
Rome Italy

rbenigni@iss.it

Traditional toxicology has been the major source of information

Now, opportunities to accept “alternative” approaches

Three **Rs**:

Replacement, **R**eduction, and **R**efinement of animal testing

with the aims of shortening times of toxicity testing, protecting animal health and welfare, and saving money

Opportunities to accept “alternative” approaches in EU

REACH: EU new regulation of chemicals

Registration, **E**valuation and **A**ssessment of **C**hemicals

- **non-testing methods**, including **(Q)SARs, read-across and chemical category** approaches will be used more extensively and more systematically than under previous EU legislation….

(Q)SAR:
(Quantitative) Structure-Activity Relationships

Read-Across:
data gaps filling approach; information for one or more *source* chemicals is used to make a prediction for a *target* chemical, considered to be similar in some way.

Chemical category:
group of chemicals whose physicochemical and toxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity

Organization for Economic Cooperation and Development (OECD) Principles

*To facilitate the consideration of a (Q)SAR model for **regulatory purposes**, it should be associated with the following information:*

- ***1) a defined endpoint***
- ***2) an unambiguous algorithm***
- ***3) a defined domain of applicability***
- ***4) appropriate measures of goodness-of–fit, robustness and predictivity***
- ***5) a mechanistic interpretation, if possible***

Identification of chemical mutagens and carcinogens:

background information

Mechanistic findings at the basis of the science and regulation of mutagens and carcinogens

- **Miller**s' electrophilic (DNA-) reactivity theory of carcinogenesis (not including nongenotoxic carcinogens)

- **Chemical mutagenicity:**
  Malling's *in vitro* metabolic activation (**S30, S9**);
  ***Salmonella*, or Ames**' test for DNA-reactive chemicals

- Structure-Activity (carcinogenicity) Relationships (**Ashby**'s **Structural Alerts**)

Mechanistic findings at the basis of the science and regulation of mutagens and carcinogens

- Because of the success **of Miller**s' electrophilic reactivity theory of carcinogenesis, and of **Ames**' test:

  major research efforts on the hypothesis **Mutation = Cancer**

  development of **> 100 STTs**

- Later on, recognition of **nongenotoxic carcinogens**

Backing up the mutagenicity STTs with **Structure-Activity** concepts

- To model / predict the mutagenicity or carcinogencity test results
- To complement the mutagenicity tests

Structure-activity relationship concepts:

application to different issues, through different approaches

Coarse-grain **Structure Alerts**

Fine-tuned Quantitative Structure-Activity Relationships (**QSAR**)

Structure Alerts (SA)

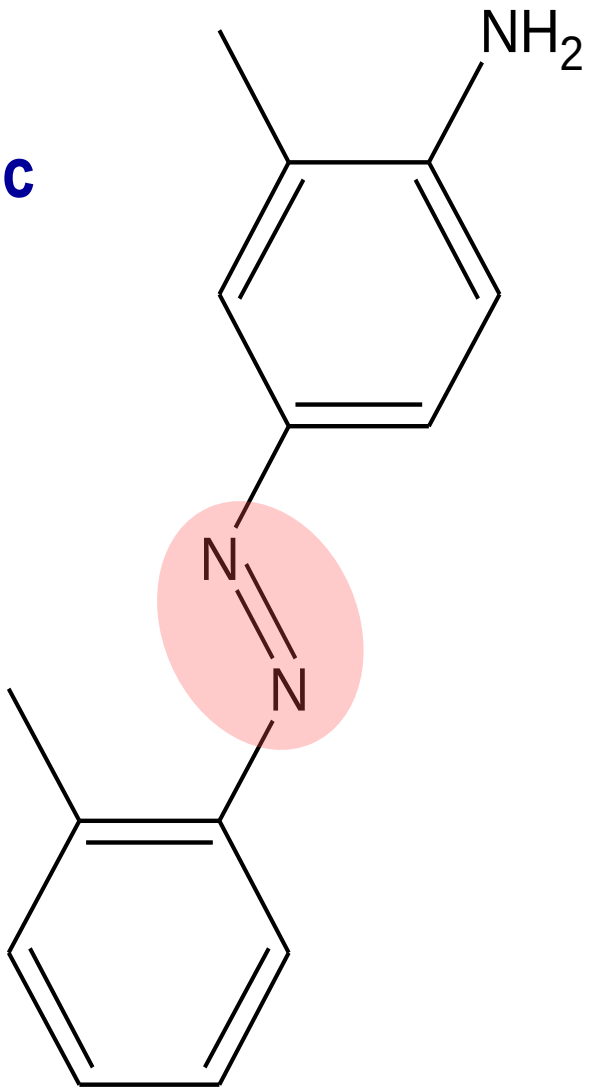
Functional groups or **Substructures**

linked to

toxic (carcinogenic / mutagenic) effects of chemicals

Several *Azo-dyes* are carcinogenic

(*generation of aromatic amines*)



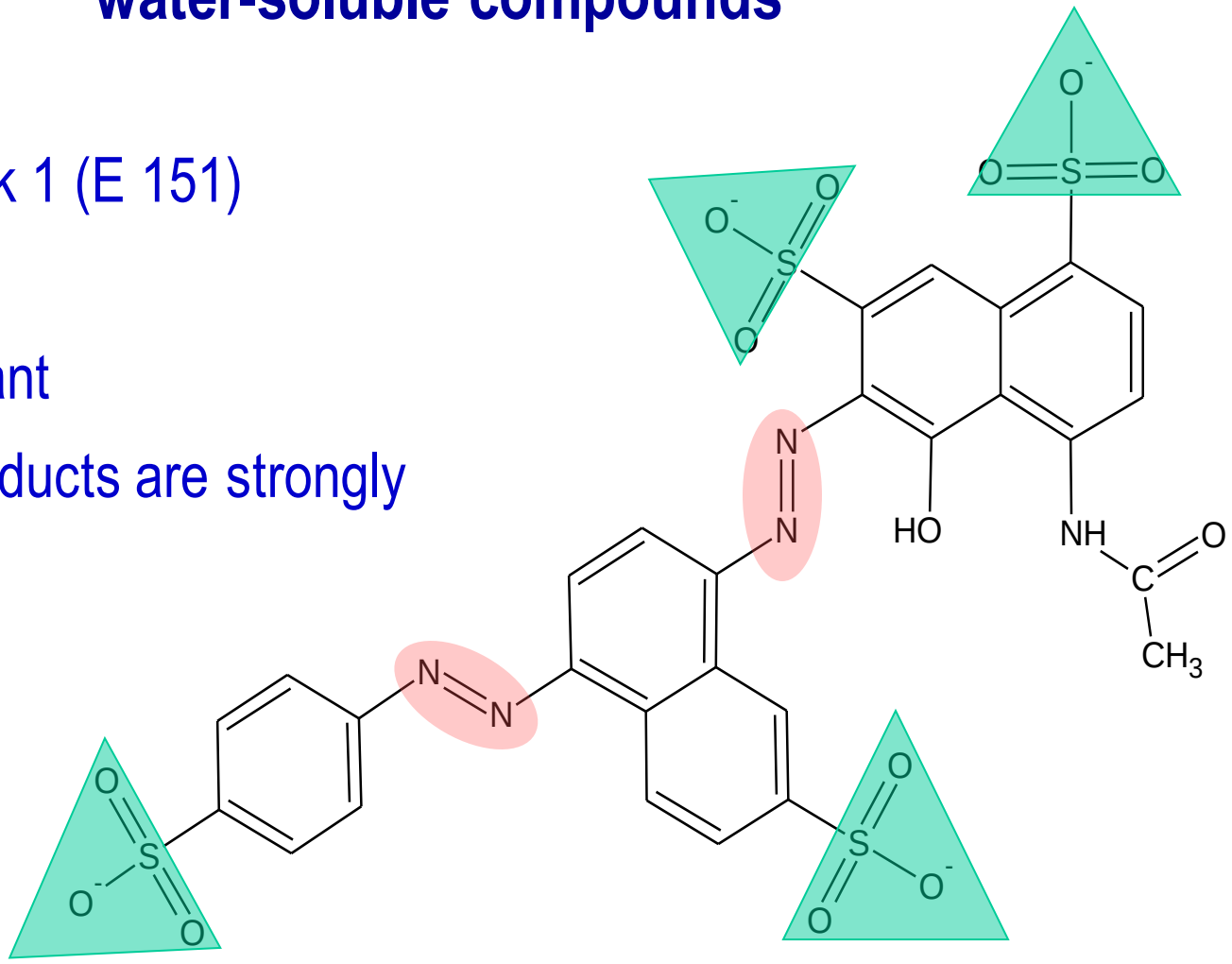
Butter yellow

- Food colorant
- Carcinogenic

Hydrophilic sulfonic acid groups generate water-soluble compounds

C.I. Food Black 1 (E 151)

- Safe colorant
- All split products are strongly hydrophilic

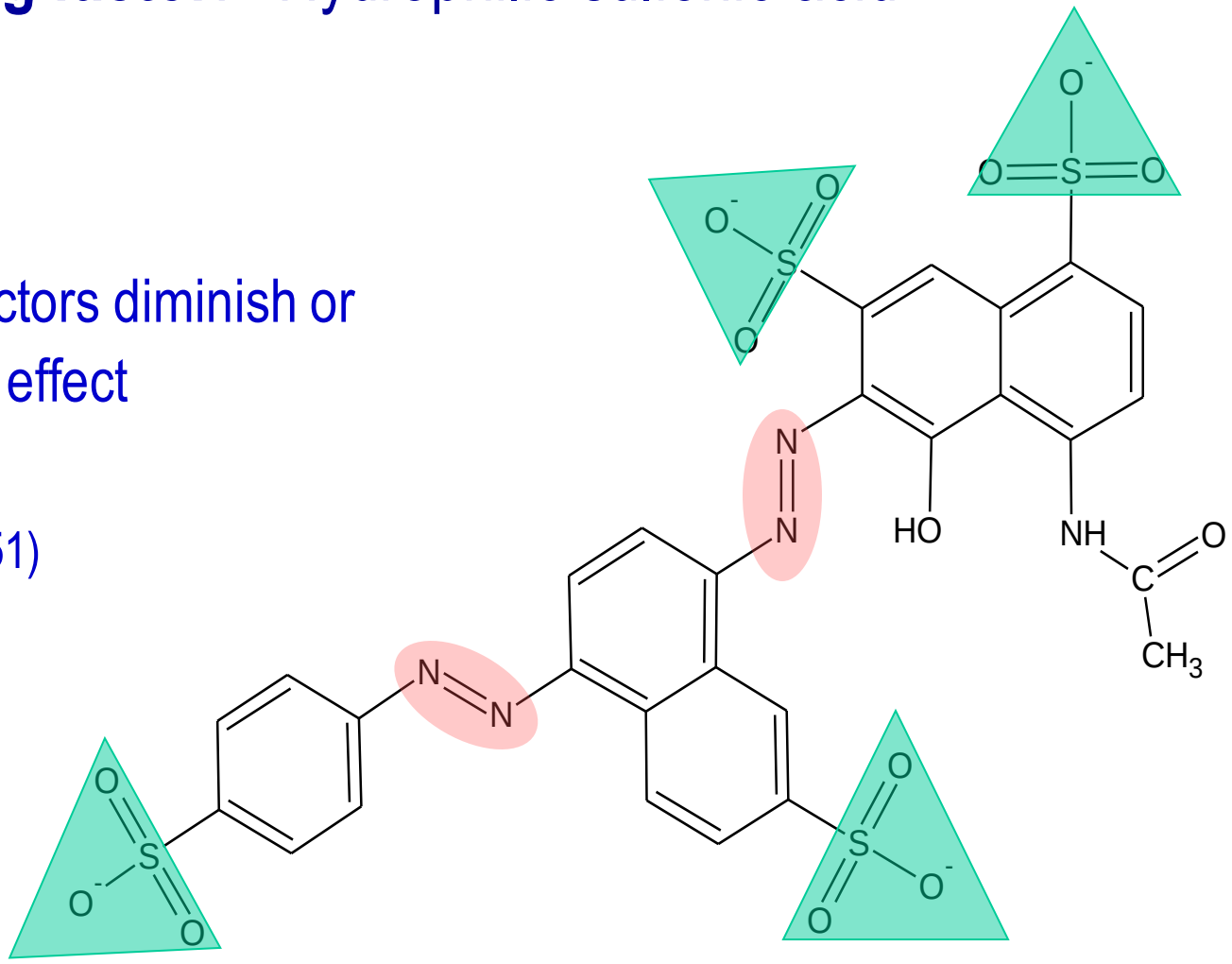


SA: Aromatic Diazo

Modulating factor: Hydrophilic sulfonic acid

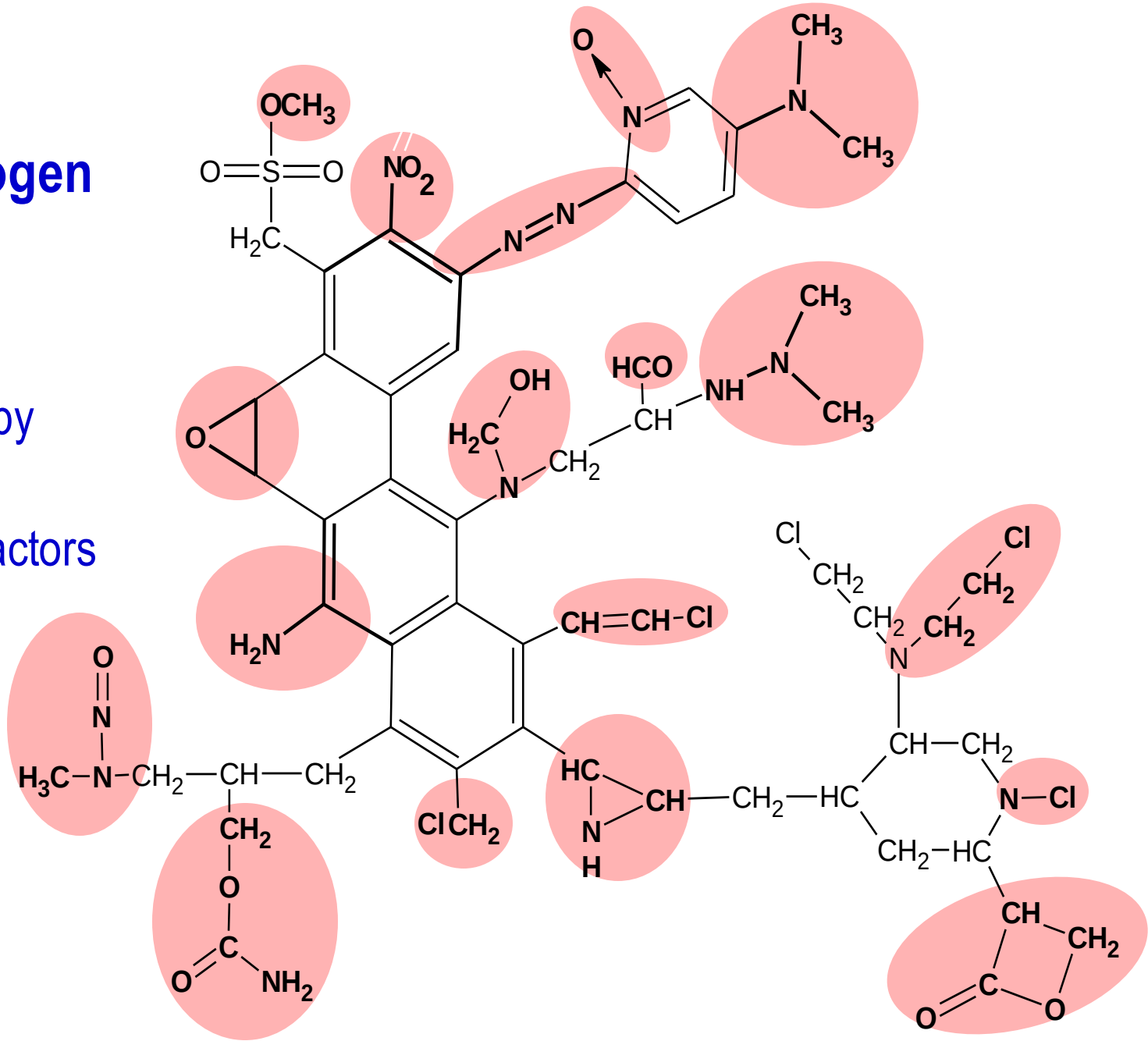
The modulating factors diminish or abolish the SA effect

C.I. Food Black 1 (E 151)



Ashby's Poly-carcinogen

Some alerts accompanied by detoxifying (modulating) factors



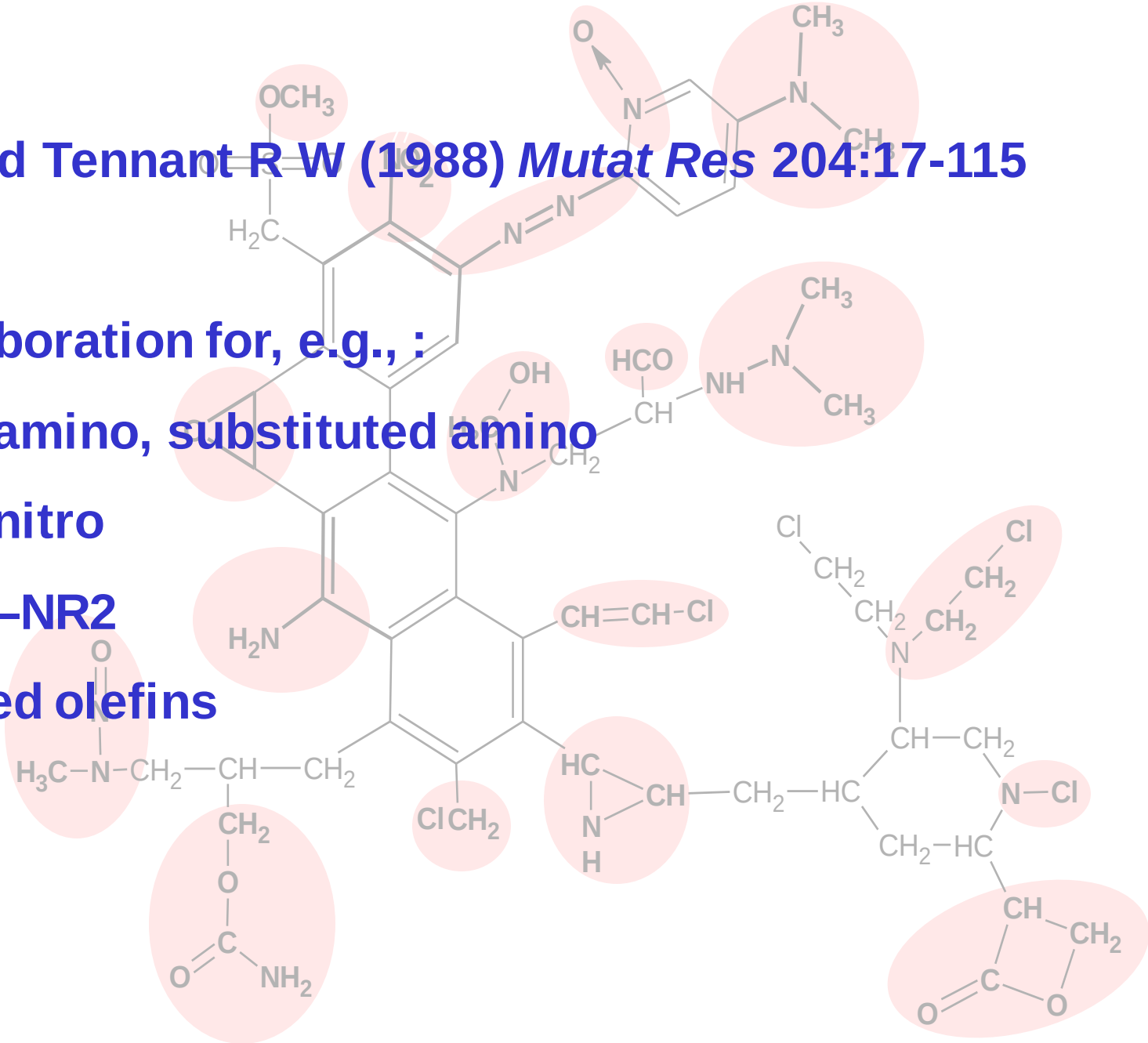
Ashby (1995) Environ.Mutag. 7: 919-921

J Ashby and Tennant R W (1988) *Mutat Res* 204:17-115

Further elaboration for, e.g., :

- **Aromatic amino, substituted amino**
- **Aromatic nitro**
- **Aromatic –NR2**
- **Chlorinated olefins**
- **PAH**

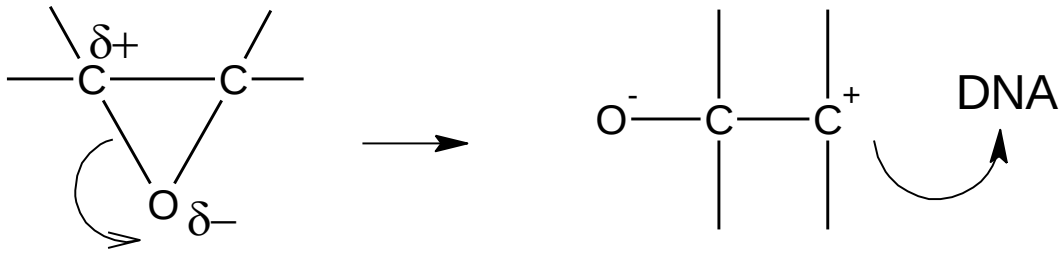
etc…



Ashby J (1978) *Cancer* 37:904-923

SA: chemical class that provokes toxic effects through one or few shared mechanisms of action

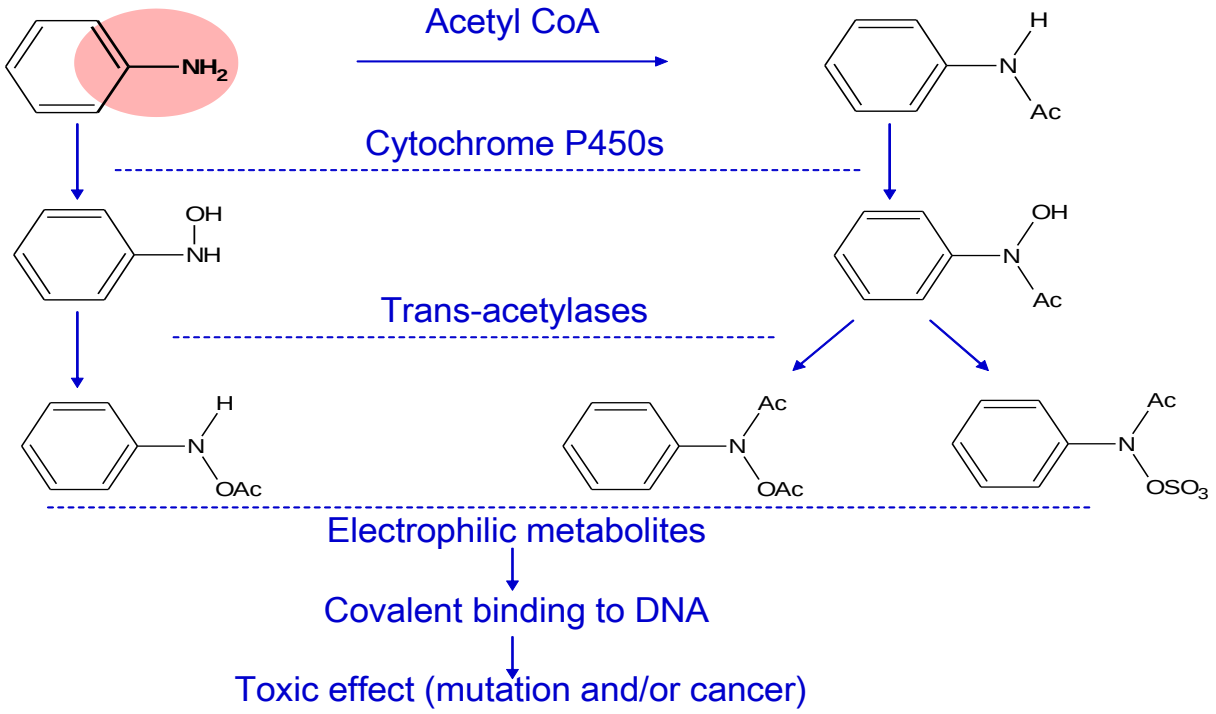
direct-acting carcinogens: e.g., ***epoxides***, aziridines, sulfur and nitrogen mustards, α-haloethers, and lactones



Strained ring Carbonium ion Alkylation

SA: chemical class that provokes toxic effects through one or few shared mechanisms of action

Metabolically activated carcinogens: e.g., ***aromatic amines***



SA: chemical class that provokes toxic effects through one or few shared mechanisms of action

complex carcinogens: e.g., ***aliphatic halogens***

From genotoxic to epigenetic, with increasing degree of halogenation and depending on the carbon skeleton

Short-chain monohalogenated alkanes (and alkenes) direct-acting alkylating agents;

Dihalogenated alkanes: alkylating or cross-linking agents.

Polyhaloalkanes: by free radical or nongenotoxic mechanisms, or reductive dehalogenation to yield haloalkenes.

Halogenated cycloalkanes (and cycloalkenes): possibly epigenetic or direct alkylation after metabolic transformation

Toxtree: Rulebases for mutagens / carcinogens

Structure-based approach consisting of:

-New compilation of ***Structure Alerts*** (32 genotox / DNA-reactive; 23 non-genotox)

-Three mechanistically-based ***QSARs*** for congeneric classes (aromatic amines, aldehydes)

Expert system **Toxtree:** Open-source, freely available

JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection (IHCP)

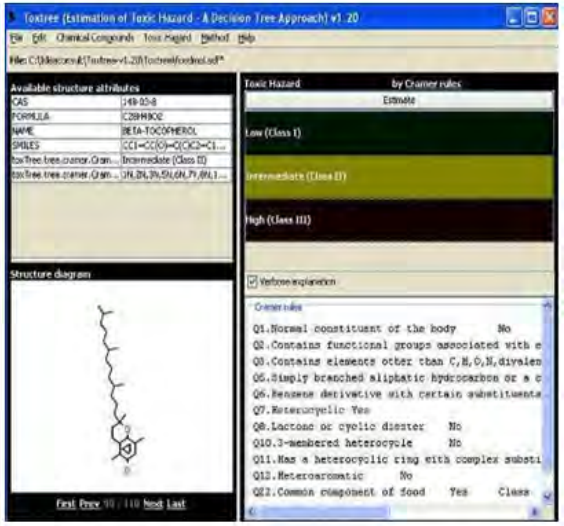
uropean Commission > JRC > IHCP > Our Laboratories > Computational Toxicology and Modelling > QSAR Tools > Toxtree

- Computational Toxicology and Modelling
- Background
- Information Sources
- Publications
- QSAR Tools
  - Stat4tox - Software for the Statistical Evaluation of In Vitro Assays
  - Danish (Q)SAR Database
  - Toxtree

Toxtree

Toxtree is a flexible and user-friendly open-source application that places chemicals into categories and predicts various kinds of toxic effect by applying decision tree appro
implemented:

- the Cramer classification scheme
- an Extended Cramer scheme
- the Kroes TTC decision tree
- the Verhaar scheme for aquatic modes of action
- rulebases for skin and eye irritation and corrosion
- the Benigni-Bossa rulebase for mutagenicity and carcinogenicity
- the ToxMic rulebase for the in vivo micronucleus assay



is made freely available as a service to scientific researchers ar

Toxtree v. 2.5

http://ihcp.jrc.ec.europa.eu/our_labs/predictive_toxicology/qsar_tools/toxtree

Toxtree v. 2.6.0

http://toxtree.sourceforge.net/download.html#Toxtree_2.6.0

http://toxtree.sourceforge.net/predict/ *(web version)*

European Commission

JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection (IHCP)

Computational Toxicology and Modelling
Background
Information Sources

Toxtree

Toxtree is a flexible and user-friendly open-source application that places chemicals into categories and predicts various kinds of toxic effect by applying decision tree appro
implemented:

- the Cramer classification scheme

eme

ree

aquatic modes of

e irritation and

se for
genicity

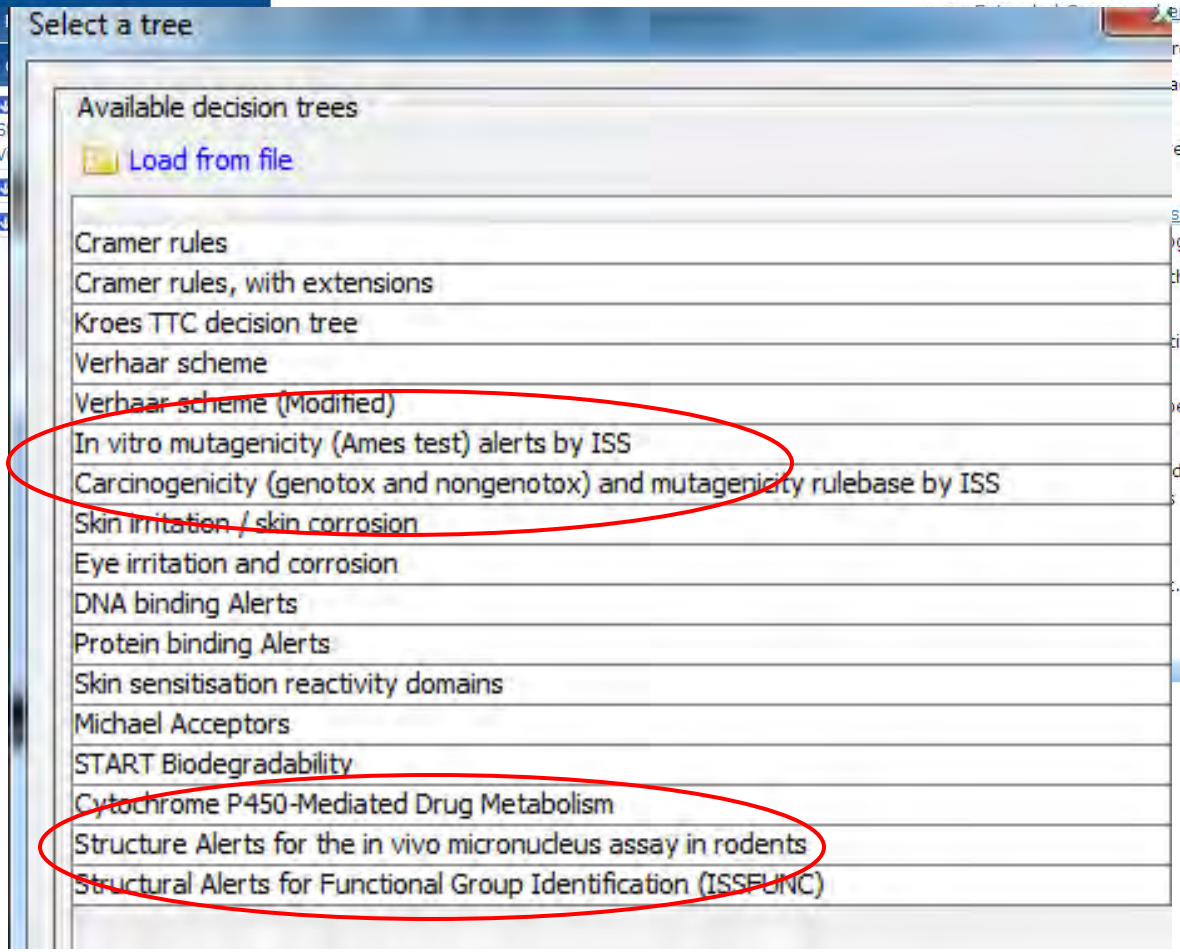
the in vivo

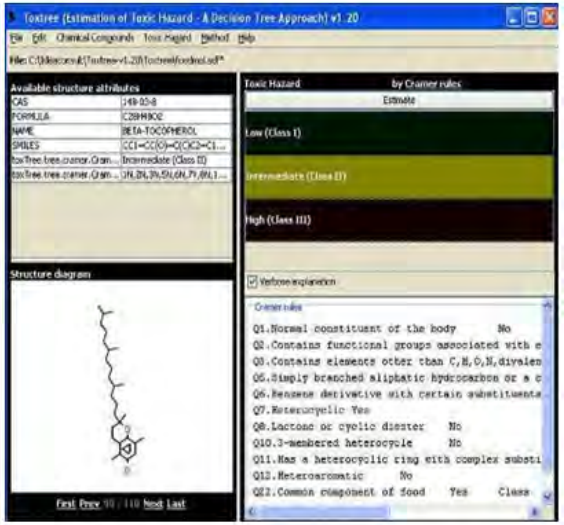
tification of

ersistance /

dentification of
(ISSFUNC

. The software is made freely available as a service to scientific researchers ar





Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v1.40

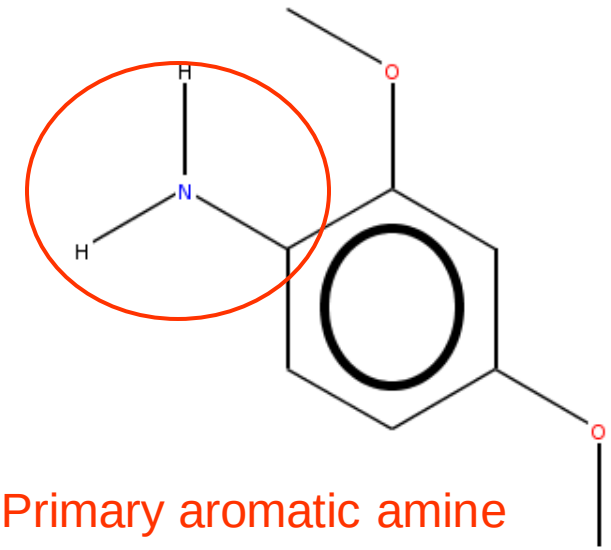
File Edit Chemical Compounds Toxic Hazard Method Help

<< >> Enter SMILES: CCC Go!

Available structure attributes

B5STM1	1,0000
Benigni / Bossa rulebase (for mutagenicity a...	,SA1N,SA2N,SA3N,SA4N,SA5N,SA6N,SA7N,...
EHOMO	-8,1427
ELUMO	0,3737
For a better assessment a QSAR calculation ...	NO
I(An)	true
I(BiBr)	false
I(NO2)	false
Idist	0
LSTM1	2,0600
MR2	0,7900

Structure diagram



Primary aromatic amine

First Prev 1 / 1 Next Last

Toxic Hazard by Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)

Estimate

Structural Alert for genotoxic carcinogenicity

Structural Alert for nongenotoxic carcinogenicity

No alerts for carcinogenic activity

Potential S. typhimurium TA100 mutagen based on QSAR

Unlikely to be a S. typhimurium TA100 mutagen based on QSAR

Potential carcinogen based on QSAR

Unlikely to be a carcinogen based on QSAR

For a better assessment a QSAR calculation could be applied.

Verbose explanation

Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)

```
QSA1.Acyl halides          No
QSA2.Alkyl (C<5) or benzyl ester of sulphonic or phosphonic a
QSA3.N-methylol derivatives      No
QSA4.Monohaloalkene      No
QSA5.S or N mustard      No
QSA6.Propiolactones and propiosultones  No
```

Completed.

Toxtree (Estimation of Toxic Hazard - A Decision Tree Approach) v2.1.0

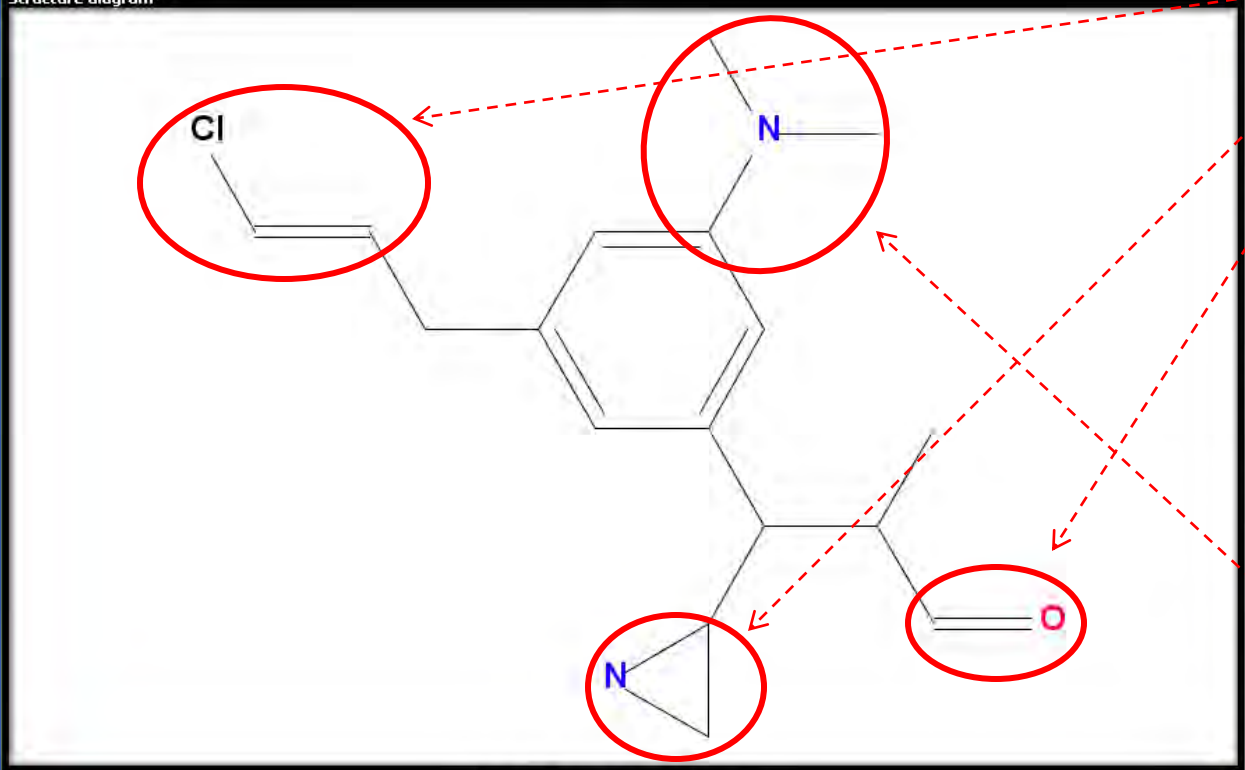
File Edit Chemical Compounds Toxic Hazard Method Help

<< >> Enter SMILES: Go!

Available structure attributes

B5STM1	2,0400
EHOMO	-8,7053
ELUMO	-0,0108
Error when applying the decision tree	YES
For a better assessment a QSAR calculation could be applied.	NO
LSTM1	2,8700
MR2	0,1000
MR3	NaN
MR5	NaN
MR6	0,1000
Negative for genotoxic carcinogenicity	NO
Negative for nongenotoxic carcinogenicity	YES
Potential S. typhimurium TA100 mutagen based on QSAR	NO
Potential carcinogen based on QSAR	NO
Proceed with QSAR6 and QSAR8?	YES
QSAR6,8 applicable?	YES
SA1	NO
SA10	NO
SA11	YES
SA12	NO
SA13	NO

Structure diagram



First Prev 1 / 1 Next Last

Toxic Hazard by Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)

Estimate

Structural Alert for genotoxic carcinogenicity

Structural Alert for nongenotoxic carcinogenicity

Potential S. typhimurium TA100 mutagen based on QSAR

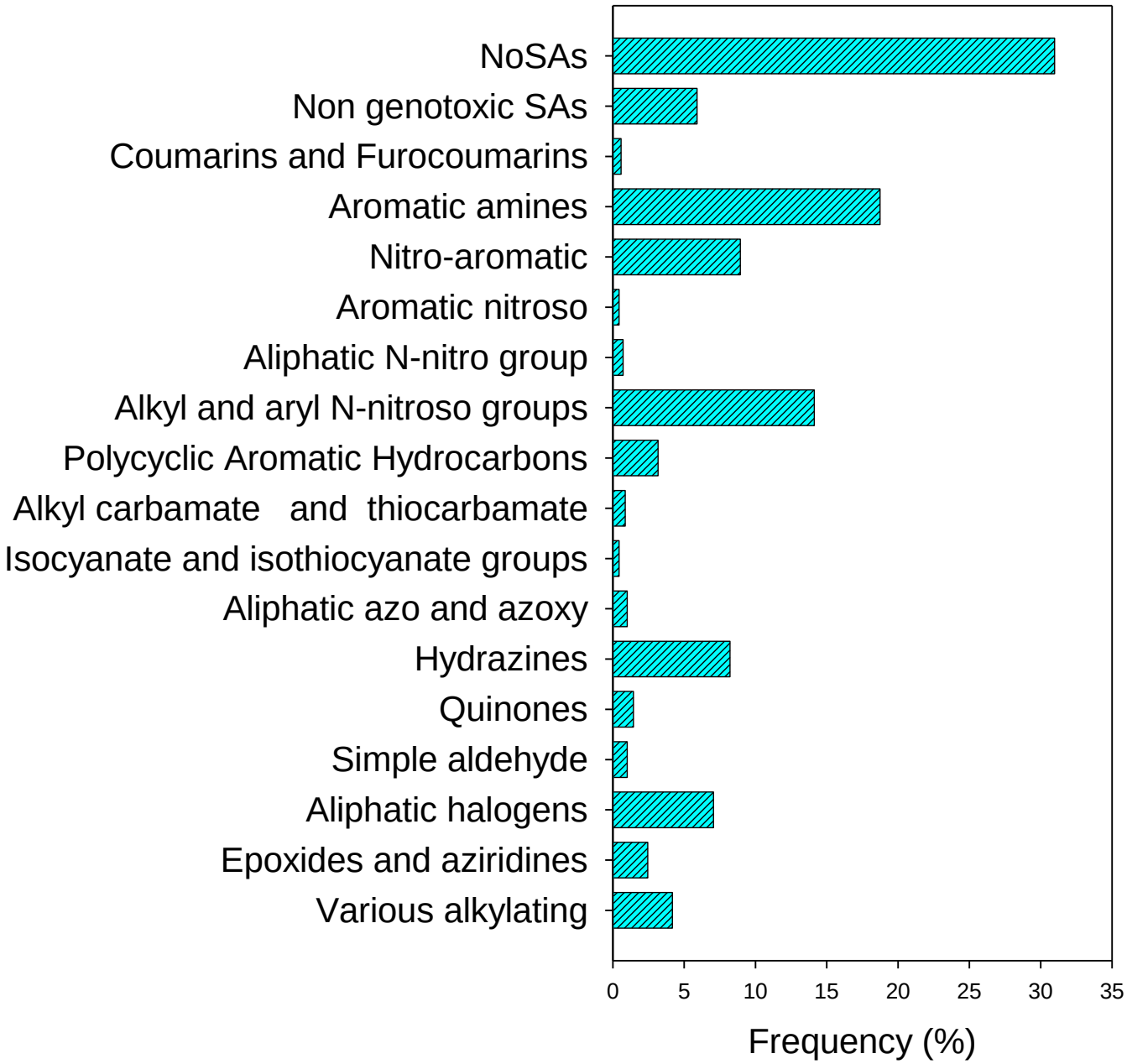
Unlikely to be a S. typhimurium TA100 mutagen based on QSAR

Verbose explanation

Benigni / Bossa rulebase (for mutagenicity and carcinogenicity)

- QSA1.Acyl halides No
- QSA2.Alkyl (C5) or benzyl ester of sulphonic or phosphonic acid No
- QSA3.N-methylol derivatives No
- QSA4.Monohaloalkene Yes
- QSA5.S or N mustard No
- QSA6.Propiolactones and propiosultones No
- QSA7.Epoxides and aziridines Yes
- QSA8.Aliphatic halogens No
- QSA9.Alkyl nitrite No
- QSA11.Simple aldehyde Yes
- QSA12.Quinones No
- QSA13.Hydrazine No
- QSA14.Aliphatic azo and azoxy No
- QSA15.Isocyanate and isothiocyanate groups No
- QSA16.Alkyl carbamate and thiocarbamate No
- QSA18.Polycyclic Aromatic Hydrocarbons No
- QSA19.Heterocyclic Polycyclic Aromatic Hydrocarbons No
- QSA21.Alkyl and aryl N-nitroso groups No
- QSA22.Azide and triazene groups No
- QSA23.Aliphatic N-nitro No
- QSA24.α,β unsaturated alkoxy No
- QSA25.Aromatic nitroso group No
- QSA26.Aromatic ring N-oxide No
- QSA27.Nitro aromatic No
- QSA28.Primary aromatic amine, hydroxyl amine and its derived esters (with restrictions) No
- QSA28bis.Aromatic mono- and dialkylamine Yes
- QSA28ter.Aromatic N-acyl amine No
- QSA29.Aromatic diazo No
- QSA30.Coumarins and Furocoumarins No
- QGenotoxic alert?.At least one alert for genotoxic carcinogenicity fired? Yes Class Structural Alert for genotoxic carcinogenicity
- QSA17.Thiocarbonyl (Nongenotoxic carcinogens) No

Profile of the Kirkland's database on carcinogens



Many **Toxtree** rulebases included in the **OECD (Q)SAR Toolbox**

Download and documentation:
www.qsartoolbox.org

OncoLogic™ - A Computer System to Evaluate the Carcinogenic Potential of Chemicals

You will need Adobe Reader to view some of the files on this page. See EPA's PDF page to learn more.

New OncoLogic™ Version 8.0 - Released September 2013

OncoLogic™ is a desktop computer program that evaluates the likelihood that a chemical may cause cancer. OncoLogic™ has been peer reviewed, runs on a Windows® PC, and is bein evaluate cancer potential of chemicals. The OncoLogic™ installer is posted at the bottom of this page.

NEW FOR VERSION 8.0 RELEASE: Updates to the system include a new CAS/Name look-up feature under the "Organics SAR" module for approximately 1500 chemicals for which av need to draw the chemical structure for these substances. Additionally, four new "simplified" chemical classes have been added to the "Organics SAR" module including aliphatic alcoho existing logic previously available for the other classes in OncoLogic remain unchanged.

New and existing users of OncoLogic 8.0 should report any model logic issues, inconsistencies with previous version outputs, or general software problems to the technical points of co

How does OncoLogic™ predict cancer potential of a chemical?

OncoLogic™ predicts cancer-causing potential by:

1. applying the rules of structure activity relationship (SAR) analysis,
2. mimicking the decision logic of human experts, and
3. incorporating knowledge of how chemicals cause cancer in animals and humans.

What is a Structure Activity Relationship (SAR)?

SAR is a technique used by chemists, biologists, and other scientists to correlate the biological activity of a chemical to its structure. When performing SAR analysis, a scientist will gro arrangement and distribution of functional groups, and analyzes the contribution of each factor to biological activities. Other considerations include the conditions under which humans

What is an expert system?

An expert system is a computer program that mimics the judgment of experts by following sets of knowledge rules that are based on studies of how chemicals cause cancer in animals from the user and following the knowledge rules incorporated into the system, uses the responses to construct an estimation of the most likely results.

What are the chemical classes evaluated by OncoLogic™?

Currently OncoLogic™ has subsystems that can evaluate fibers, metals, polymers, and more than 48 classes of organic chemicals. If the chemical of interest can not be placed into one

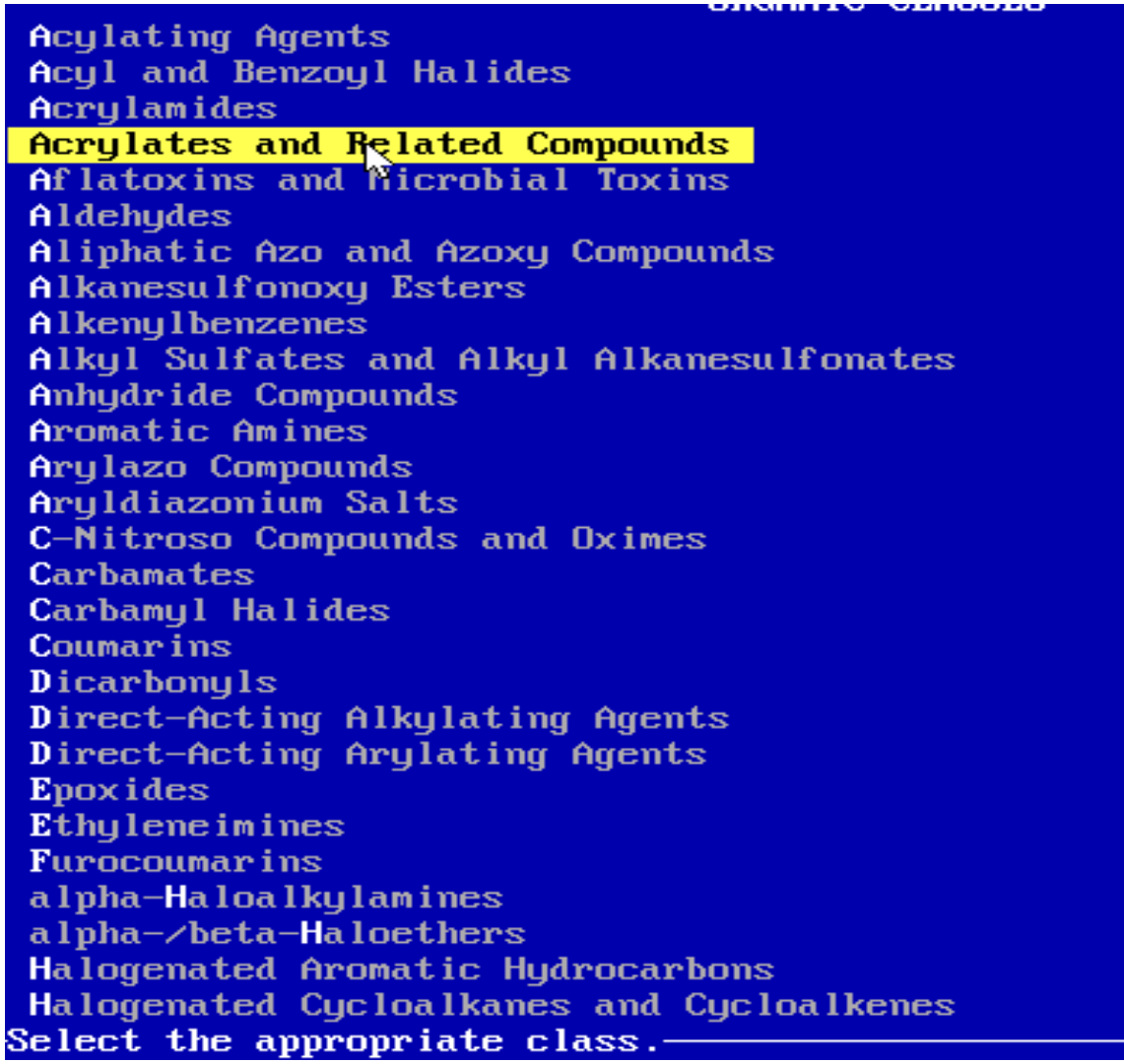
Free download and documentation:

http://www.epa.gov/oppt/sf/pubs/oncologic.htm

Running OncoLogic®:
Organics Module

- Select chemical class
  - 48 total
  - Description in Manual
  - Hit "F1" to view sample structures

- Absence of structure in OncoLogic provides suggestive, *but not definitive*, evidence of lack of major cancer concern. Functional arm should be used if possible.



Checking the agreement between

SAs and experimental results

A chemical with a SA (with no modulating factors)

is predicted as potentially toxic

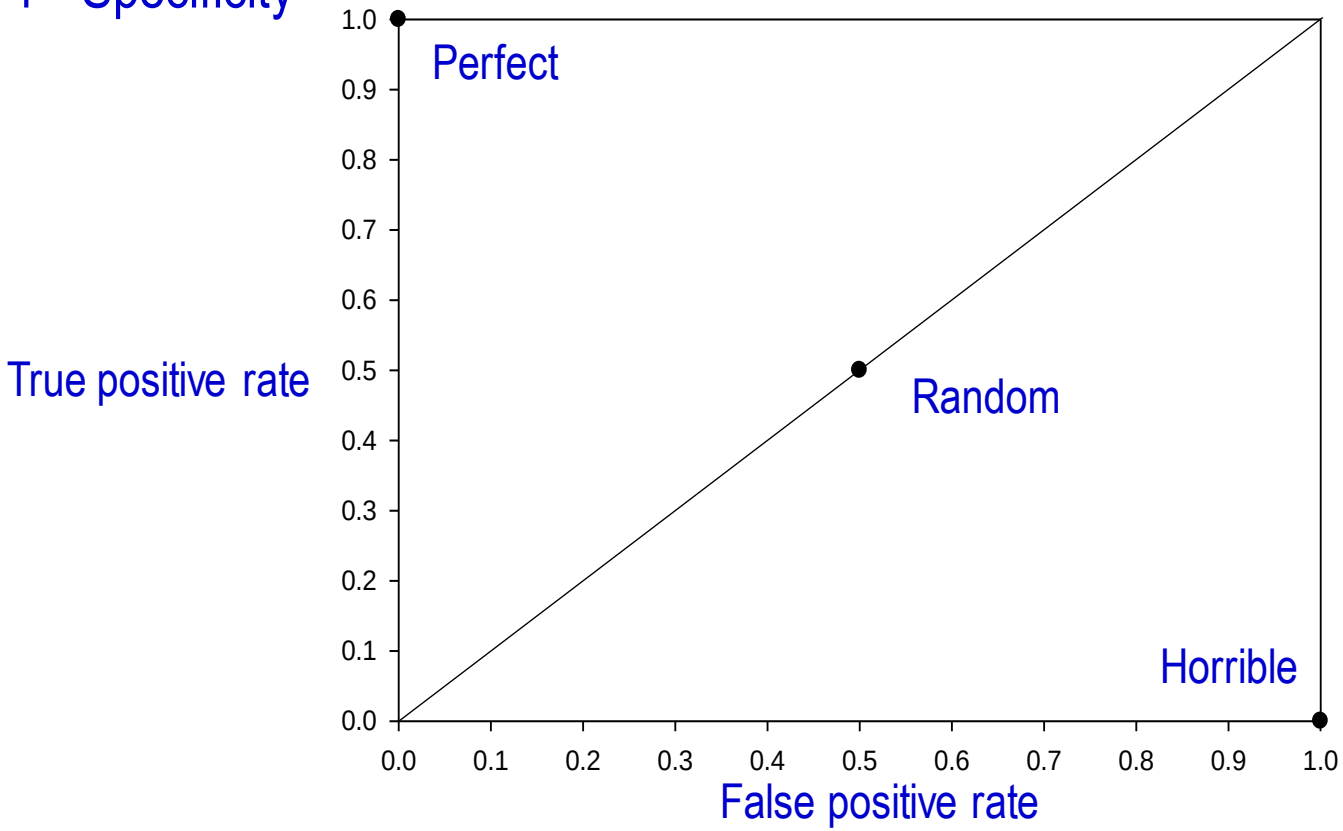
ROC graph: A simple, graphical way of comparing predictions with results

True positive rate = (Positives predicted as positive) / (Real positives)

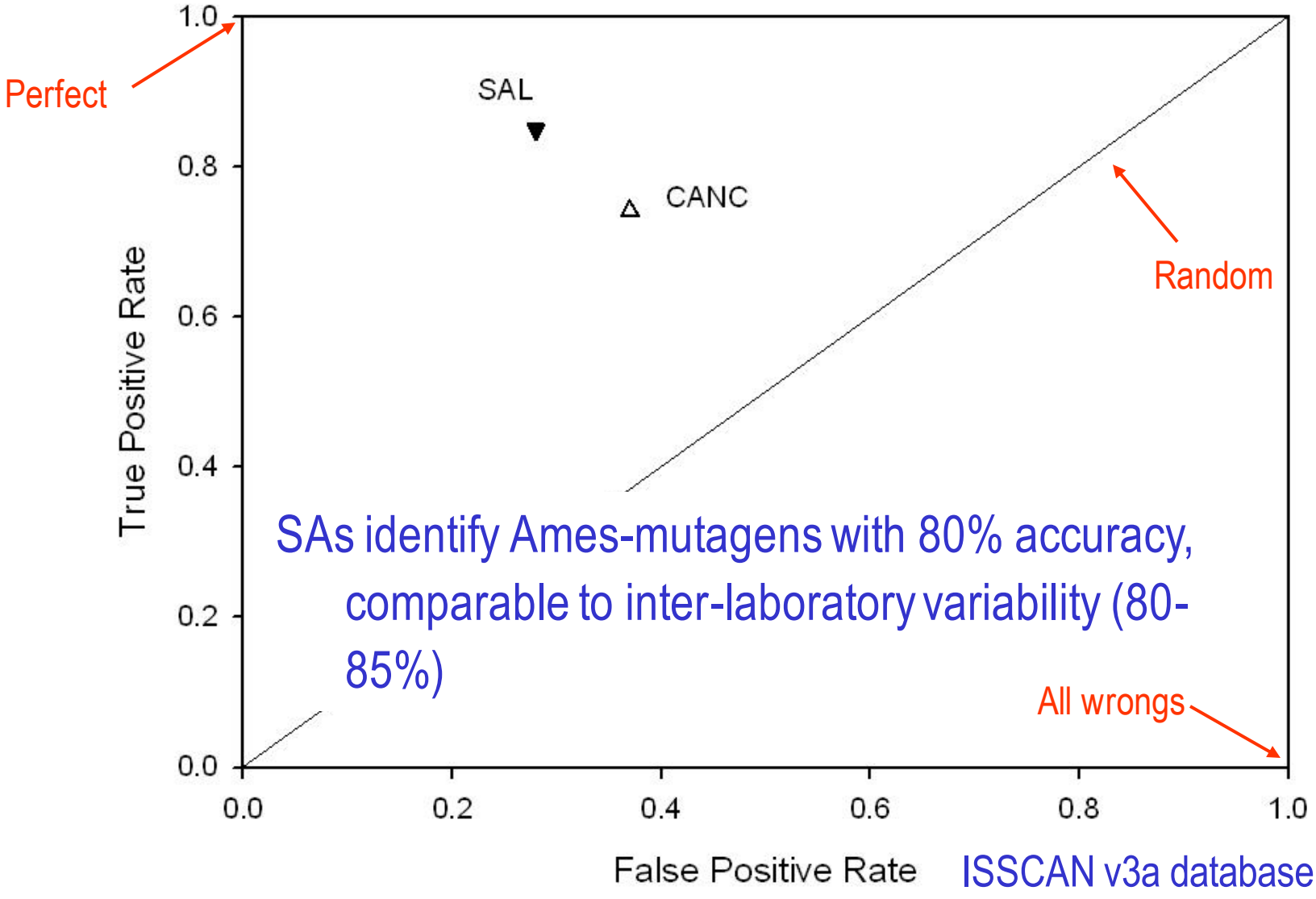
= Sensitivity

False Positive Rate = (Negatives predicted as positive) / (Real negatives)

= 1 - Specificity



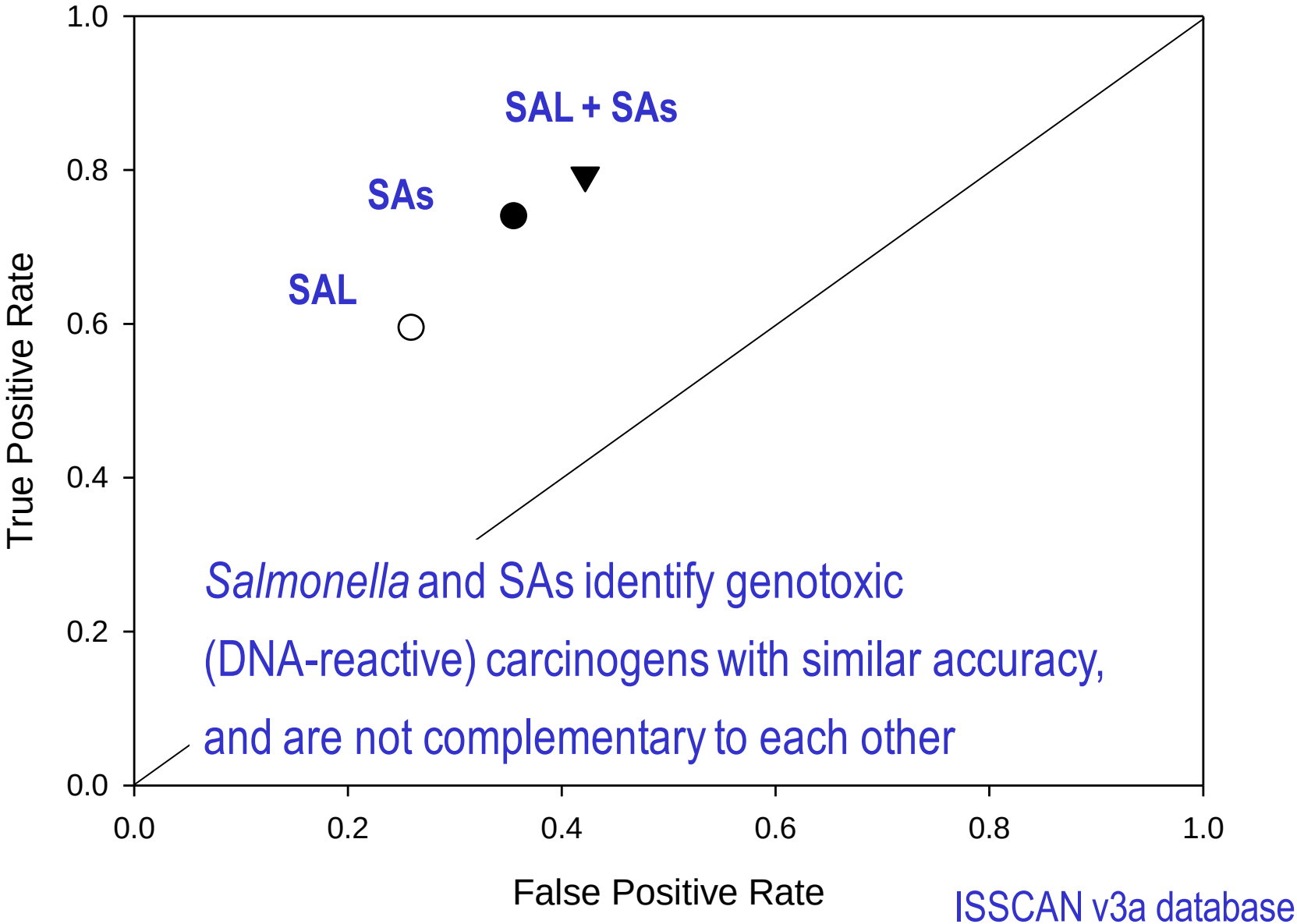
Toxtree SAs: agreement with Carcinogenicity and *Salmonella* (Ames)



SAs identify Ames-mutagens with 80% accuracy, comparable to inter-laboratory variability (80-85%)

ISSCAN v3a database

Carcinogenicity prediction: *Salmonella* (Ames) *versus* SAs



ISSCAN v3a database

Which use for the Structure Alerts ?

Success story: priority setting by human experts

Out of 400 chemicals tested by NCI/NTP:

- 2/3 selected as suspect carcinogens (n=267)
  68% carcinogenic (n=187)

- 1/3 selected on production/exposure considerations (n=133)
  20% carcinogenic (n=26); 6.8% positive in two species (n=9)

Fung, Barrett, and Huff (1995) Environ.Health Perspect. 103: 680-683

Knowledge on DNA-reactivity (coded in SAs):

- Reliable enough to predict *Salmonella* results, and identify many carcinogens
- Identify human carcinogens
- Basis for successful priority setting in NTP bioassays (70% carcinogens among structurally suspect chemicals, only 10% among high exposure chemicals)
- Contribution to reduce DNA-reactive carcinogens among synthetic chemicals (pesticides, pharmaceuticals)

Which use for the Structure Alerts ?

Great tool for **coarse-grain** characterization of the chemicals:

- Description of sets of chemicals
- Preliminary hazard characterization
- Category formation (e.g., regulation, fine-tuned QSAR, etc…)
- Priority setting (enriching the target)

General References on QSAR

- Hansch,C., Hoekman,D., Leo,A., Weininger,D., and Selassie,C.D. (2002): Chem-bioinformatics: comparative QSAR at the interface between chemistry and biology. *Chem.Revs.*, 102:783-812.

- Hansch,C. and Leo,A. (1995): *Exploring QSAR. 1. Fundamentals and applications in chemistry and biology.* American Chemical Society, Washington, D.C.

- Franke,R. and Gruska,A. (2003): General introduction to QSAR. In: *Quantitative Structure-Activity Relationhsip (QSAR) models of mutagens and carcinogens.*, edited by R.Benigni, pp. 1-40. CRC Press, Boca Raton.

References on Structural Alerts

- Ashby,J. (1985) Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. Environ.Mutagen., 7, 919-921.

- Benigni,R. and Bossa,C. (2008) Structure Alerts for carcinogenicity, and the Salmonella assay system: a novel insight through the chemical relational databases technology. Mutat.Res.Rev., 659, 248-261.

- Enoch,S.J. and Cronin,M.T. (2012) Development of new structural alerts suitable for chemical category formation for assigning covalent and non-covalent mechanisms relevant to DNA binding. Mutat.Res., 743, 10-19..

- Kalgutkar,A.S., Gardner,I., Obach,R.S., Shaffer,C.L., Callegari,E., Henne,K.R., Mutlib,A.E., Dalvie,D.K., Lee,J.S., Nakai,Y., O'Donnell,J.P., Boer,J., and Harriman,S.P. (2005) A comprehensive listing of bioactivation pathways of organic functional groups. Curr.Drug Metab., 6, 161-225.

References on Carcinogenicity Mechanisms and Oncologic analysis

- Lai,D.Y. and Woo,Y.T. (2012) Reducing Carcinogenicity and Mutagenicity Through Mechanism-based Molecular Design of Chemicals. In Boethling,R.S. and Voutchkova,A.M. (eds.) Green Processes. Volume 9: Designing Safer Chemicals. Wiley-VCH, Heidelberg.

- Woo,Y.T. and Lai,D.Y. (2010) Mechanism-based Structure-Activity Relationship Analysis of Chemical Carcinogens. In Hsu,G. and Stedeford,T. (eds.) Cancer Risk Assessment: Chemical Carcinogenesis, Hazard Evaluation, and Risk Quantification. Wiley, New York, pp 517-56.

- Woo,Y.T. and Lai,D.Y. (2005) Oncologic: a mechanism based expert system for predicting the carcinogenic potential of chemicals. In Helma,C. (ed.) Predictive toxicology. Taylor and Francis, Boca Raton, pp 385-413.