Exploratory Data Analysis: Methods and Results

Romualdo Benigni

Istituto Superiore di Sanita'
Rome Italy

rbenigni@iss.it

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

  Somatic Mutation = Cancer

- Variations: epigenetic, chromosomal, and cancer stem-cell theories

- Cancer originates at the cellular level of biologic organization, and carcinogens directly alter the DNA structure or function in cells in the tissue from which cancer arises

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

- Later on, recognition of **nongenotoxic / epigenetic carcinogens**
- Do not bind covalently to DNA, do not directly cause DNA damage, and are usually negative in the standard mutagenicity assays
- Risk they pose is of the greatest concern:
- A remarkable proportion of recognized **human carcinogens** act by nongenotoxic mechanisms (e.g., 2,3,7,8-tetrachlorodibenzo-p-dioxin, 1,4-dichlorobenzene, polychlorinated biphenyls, etc…)

Human Carcinogens *(IARC Group 1)* negative in genotox assays

Dimethylarsinic acid
Monosodium methane arsenate
Beryllium
Beryllium sulfate tetrahydrate
Chromium carbonyl
Cyclosporin
Estrogens, non-steroidal
Chlorotrianise
Estrogen/progesterone therapy
Estradiol
Estrogens, steroidal
Ethinyl estradiol
Ethanol
Gallium arsenide
Nickel sulfate hexahydrate
Nickel (II) oxide
Nickelocene
2,3,7,8-Tetrachlorodibenzo-para-dioxin

Hernandez et al., 2009, Mutat.Res., 682: 94-109.

Evolution of theories on the early stages of carcinogenesis:

- **Somatic mutation and its variations (**epigenetic, chromosomal, stem-cell):

Cancer originates at the cellular level of biological organization

Carcinogens alter DNA structure or function in cells in tissues from which cancer arises

- **Tissue organization field:**

Cancer arises from disruption of tissue microachitecture

Mutations and genetic instability as consequence of disruption of the morphostat gradient

Baker et al., 2010, J.Clin.Oncol., 28: 3215-3218

Looking for further mutagenicity Short-Term Tests (STT) to predict carcinogenicity

- Hypothesis on **complementarity**

  to cover the spectrum of cancer-relevant factors:

  different **genetic endpoints** (gene mutation, chromosomal damage)

  different **cells** (bacterial, mammalian)

  animals (ADME)

- Development of **> 100 STTs**

- Implemented into **regulations**:

In vitro {bacteria + mammalian cells} {gene mutations + chrom. damage}

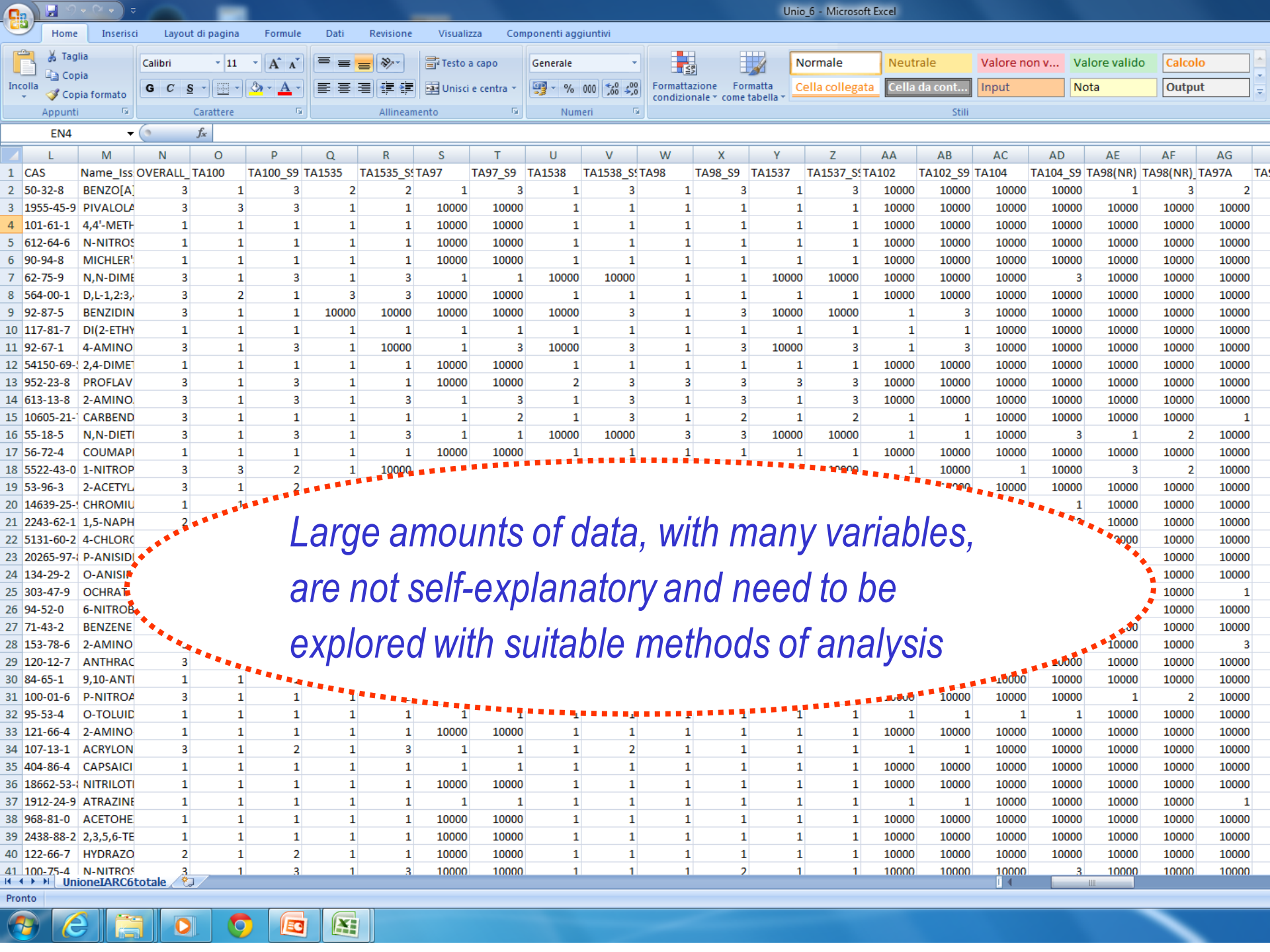
In vivo {filter *in vitro* false positives}

Biology is an experimental science

What large **databases** point to

when interrogated with adequate

Multivariate Data Analysis methods ?


Unio_6 - Microsoft Excel
Home
Inserisci
Layout di pagina
Formule
Dati
Revisione
Visualizza
Componenti aggiuntivi
Incolla
Taglia
Copia
Copia formato
Calibri
11
Testo a capo
Unisci e centra
Generale
Formattazione condizionale
Formatta come tabella
Normale
Neutrale
Valore non v...
Valore valido
Calcolo
Cella collegata
Cella da cont...
Input
Nota
Output
Appunti
Carattere
Allineamento
Numeri
Stili
EN4
Large amounts of data, with many variables, are not self-explanatory and need to be explored with suitable methods of analysis
UnioneIARC6totale
Pronto

Factor or Principal Component Analysis:

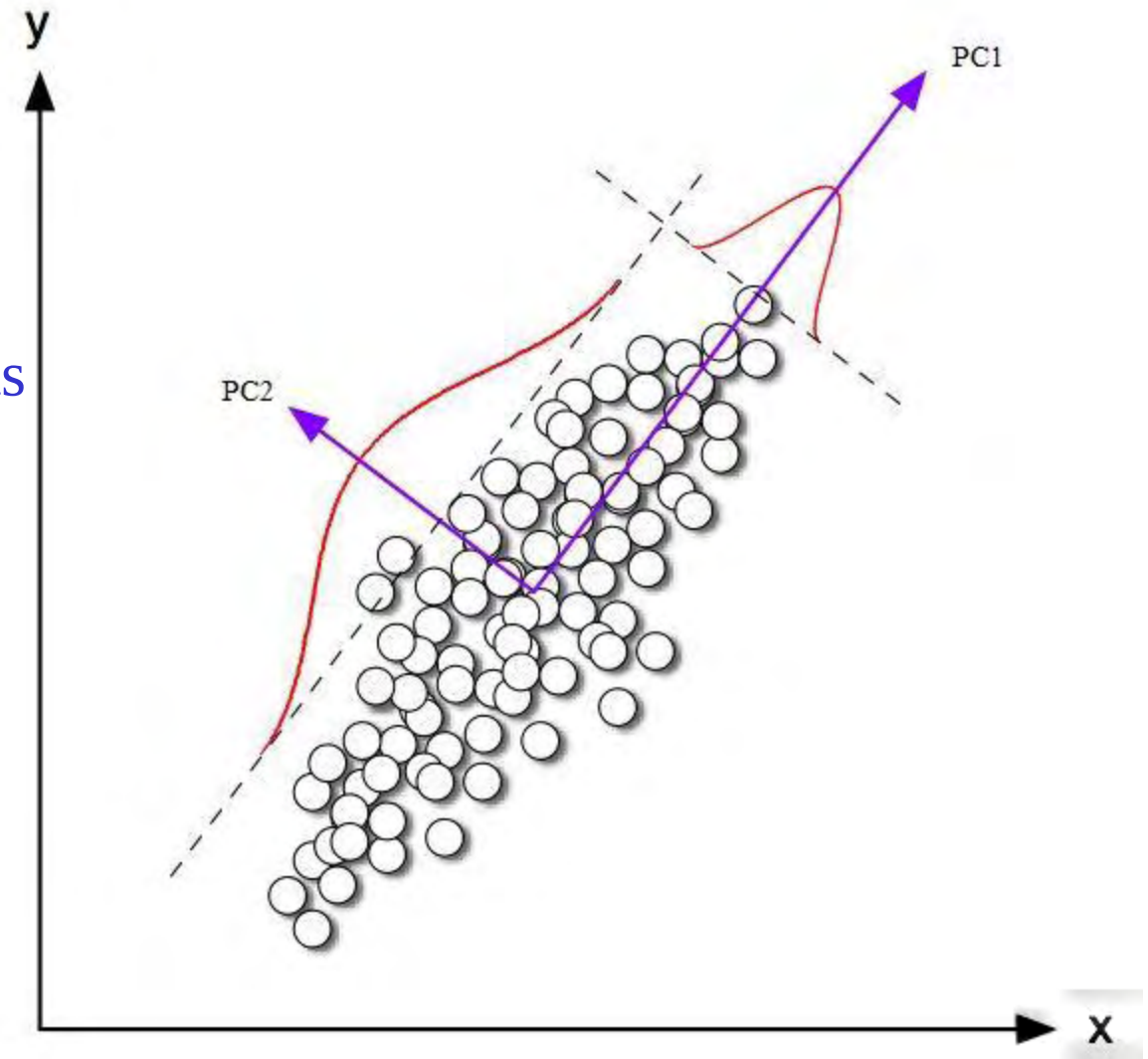
Cleans data from intercorrelations

Identifies main trends

Constructs global indices

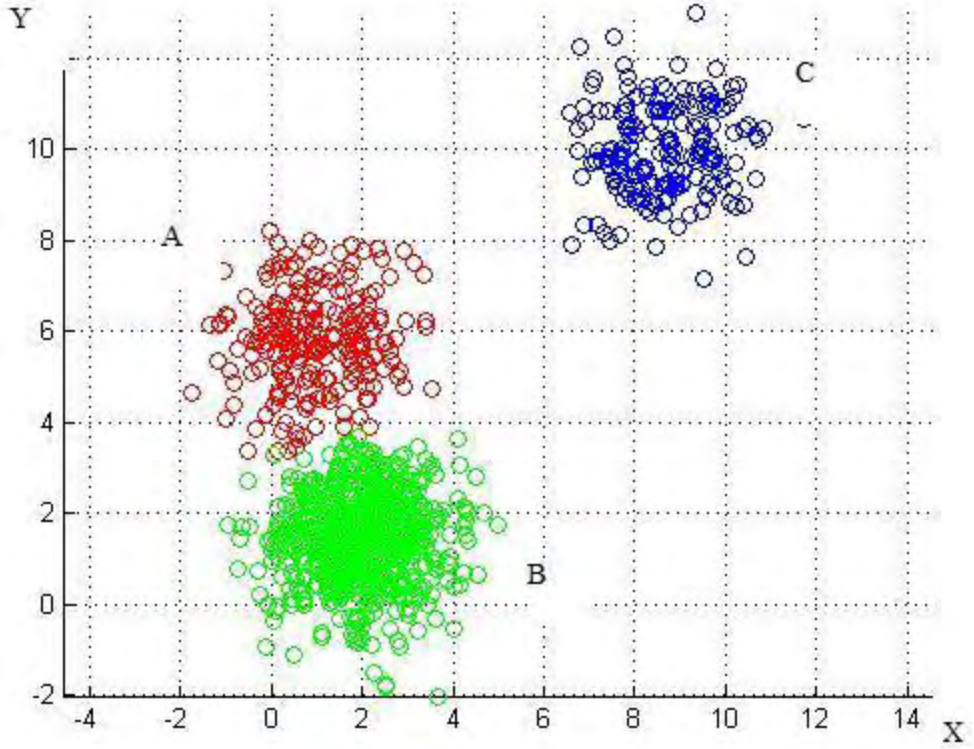
Highlights orthogonal effects

Permits easy visualization



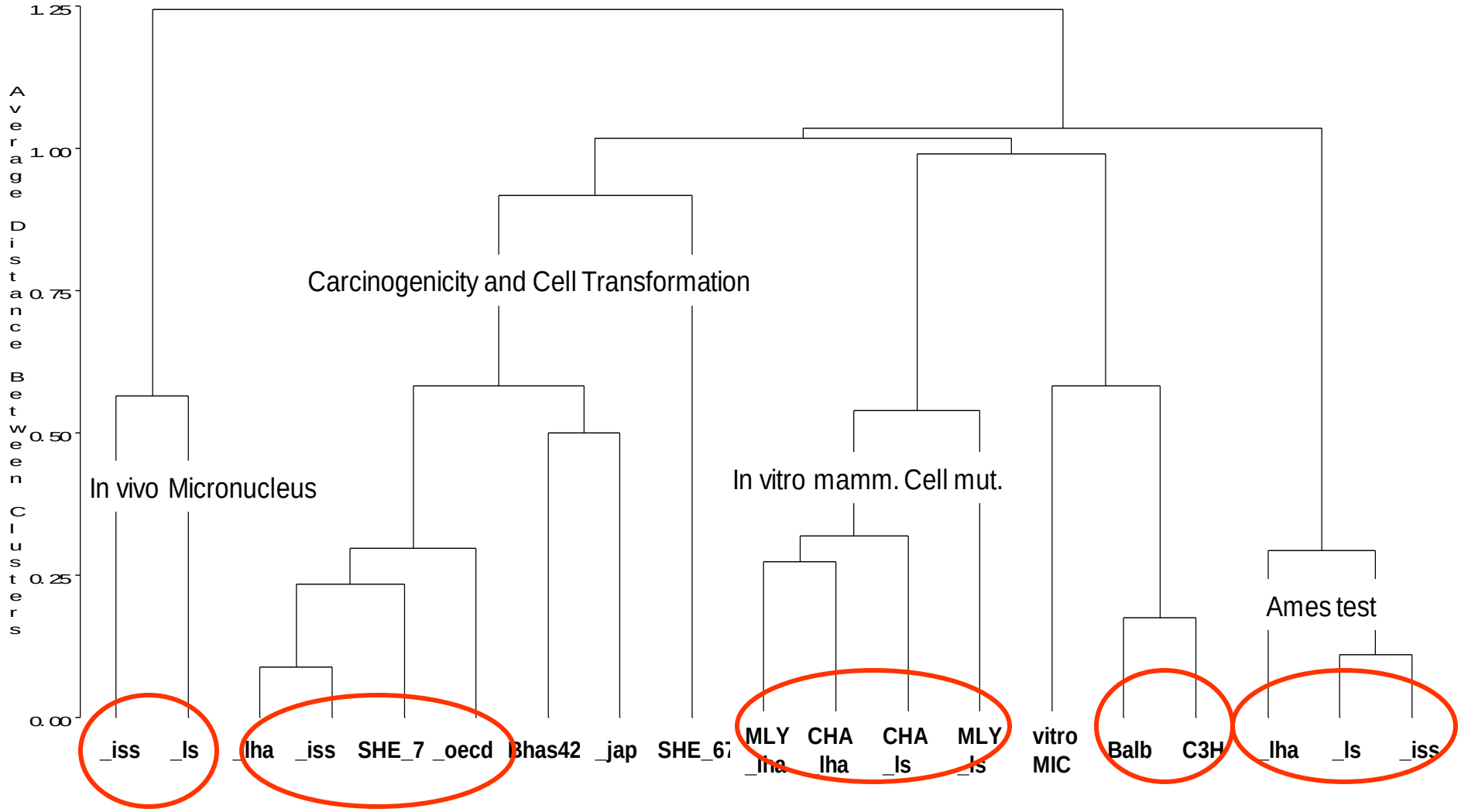
Cluster Analysis:

Identifies groups with homogeneous characteristics
Complementary to Principal Component Analysis



Cluster Analysis:

Identifies groups with homogeneous characteristics
Complementary to Principal Component Analysis



An old comparative carcinogenicity prediction exercise:

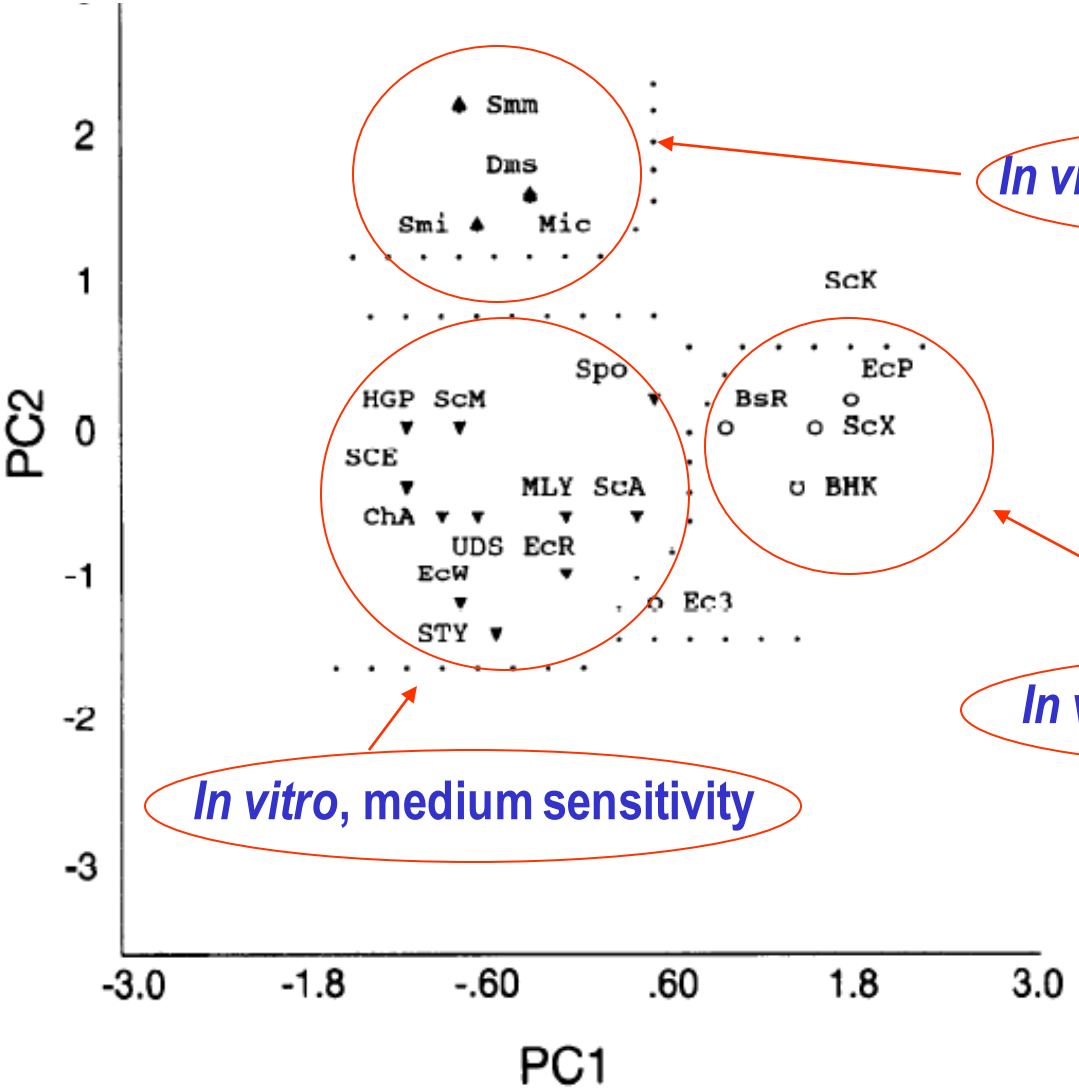
42 chemicals *systematically tested in* **22 mutagenicity STTs**

De Serres and Ashby, 1981, Evaluation of Short-Term Tests for Carcinogens,
Progr. Mutat.Res. Vol 1

Conclusion(s):

- Central role for ***Salmonella***
- Continue with exploration of different **genetic endpoints** and different **cells**

Similarity map of responses of **22 mut. STTs** to **42 chemicals**



In vivo, low sensitivity

In vitro, high sensitivity

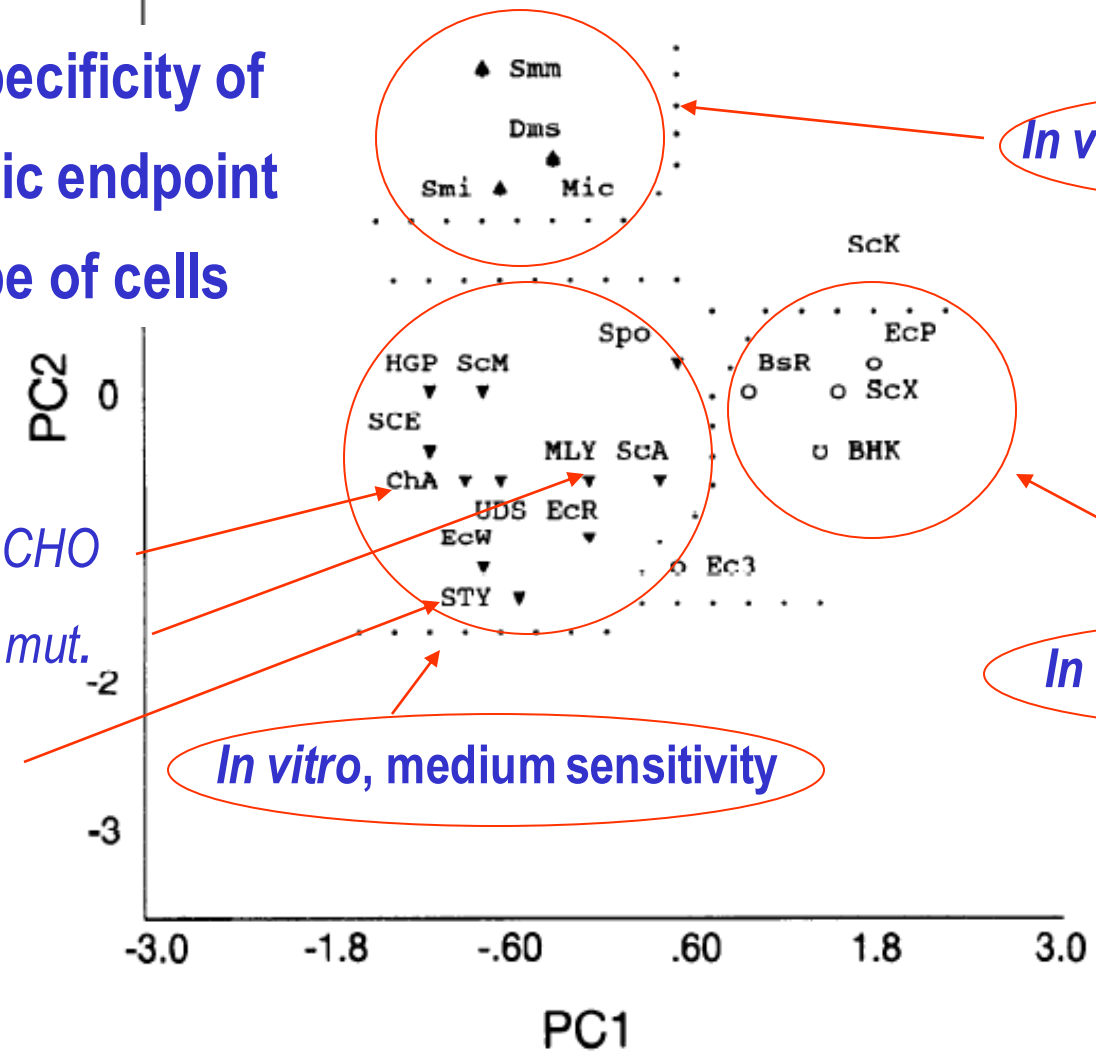
In vitro, medium sensitivity

Analysis: Benigni and Giuliani, 1985, Mutat.Res., 147: 139 - 151

Data from: De Serres and Ashby, 1981, Evaluation of Short-Term Tests for Carcinogens, Progr. Mutat.Res. Vol 1

Similarity map of responses of **22 mut. STTs** to **42 chemicals**

No specificity of genetic endpoint or type of cells



In vivo, low sensitivity

In vitro, high sensitivity

In vitro, medium sensitivity

Chrom. Aberr. CHO

Mouse lymph. mut.

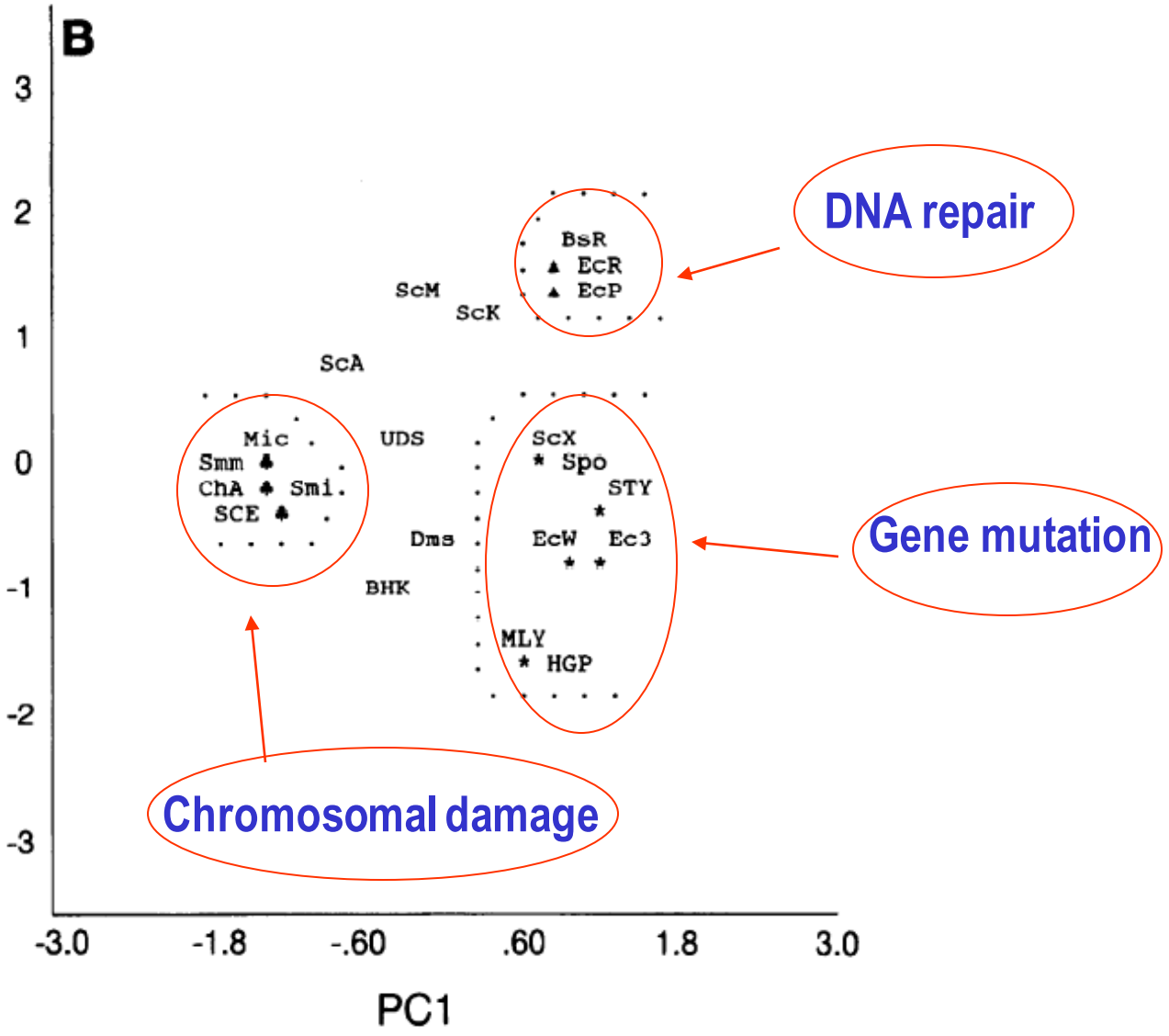
Salmonella

Analysis: Benigni and Giuliani, 1985, Mutat.Res., 147: 139 - 151

Data from: De Serres and Ashby, 1981, Evaluation of Short-Term Tests for Carcinogens, Progr. Mutat.Res. Vol 1

A questionnaire for biologists:

how 22 mutagenicity STTs will respond to 42 chemicals ?



Benigni and Giuliani, 1988, Mutat. Res. 205: 227; Benigni and Giuliani, 1994, Amer.J.Physiol., 266: R1697

Looking for further mutagenicity Short-Term Tests to predict carcinogenicity

- Looking for complements to *Salmonella (Ames)*:
  National Toxicology Progran evaluation of four *in vitro* STTs
  (n. chemicals = 114)

- Different **genetic endpoints** (gene mutation, chromosomal damage), and different **cells** (bacterial, mammalian)

  Salmonella typhimurium (Ames)
  Chromosomal Aberrations in CHO cells
  SCEs in CHO cells
  Mouse Lymphoma mutation

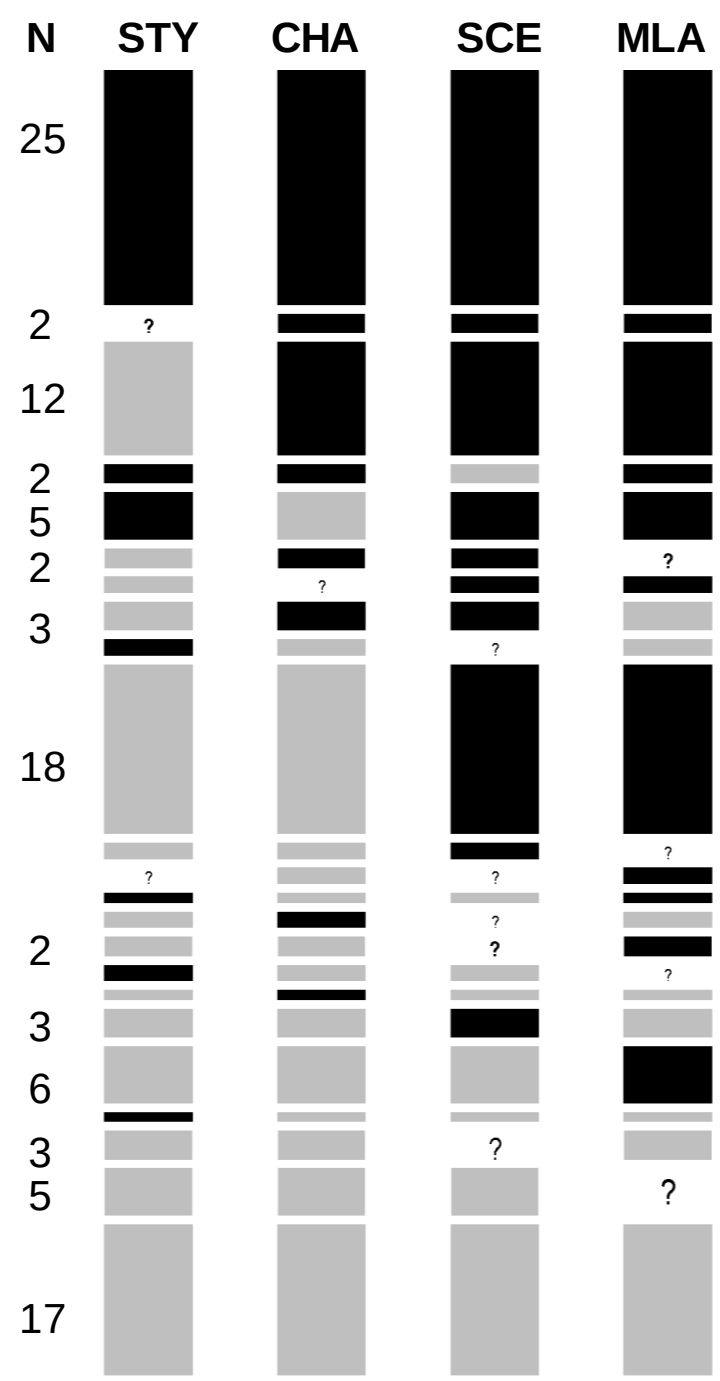
Tennant et al., 1987 Science 236: 933-941
Zeiger et al., 1990 Environ. Mut. Mutagen. 16(18): 1-14

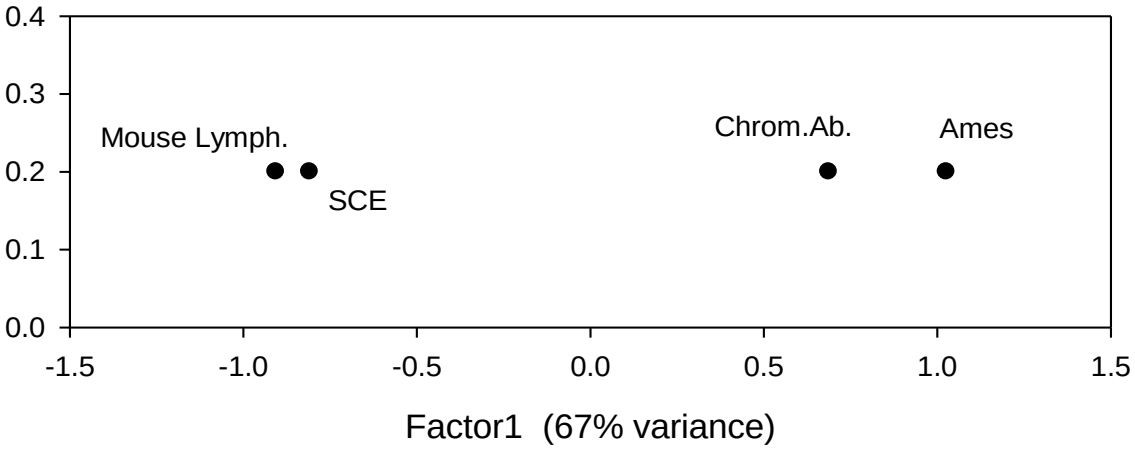
Relevance of STTs to rodent carcinogenicity

	Chi-square (p)
Ames	< 0.0001
Chrom. Ab (CHO)	0.011
Sce (CHO)	0.277
Mouse Lymphoma mut.	0.305
Batteries	
Ames + Chrom.Ab.	0.0004
Ames + Mouse Lymphoma	0.120

(our elaborations)

Genetic Toxicity Profiles: 114 NTP Chemicals



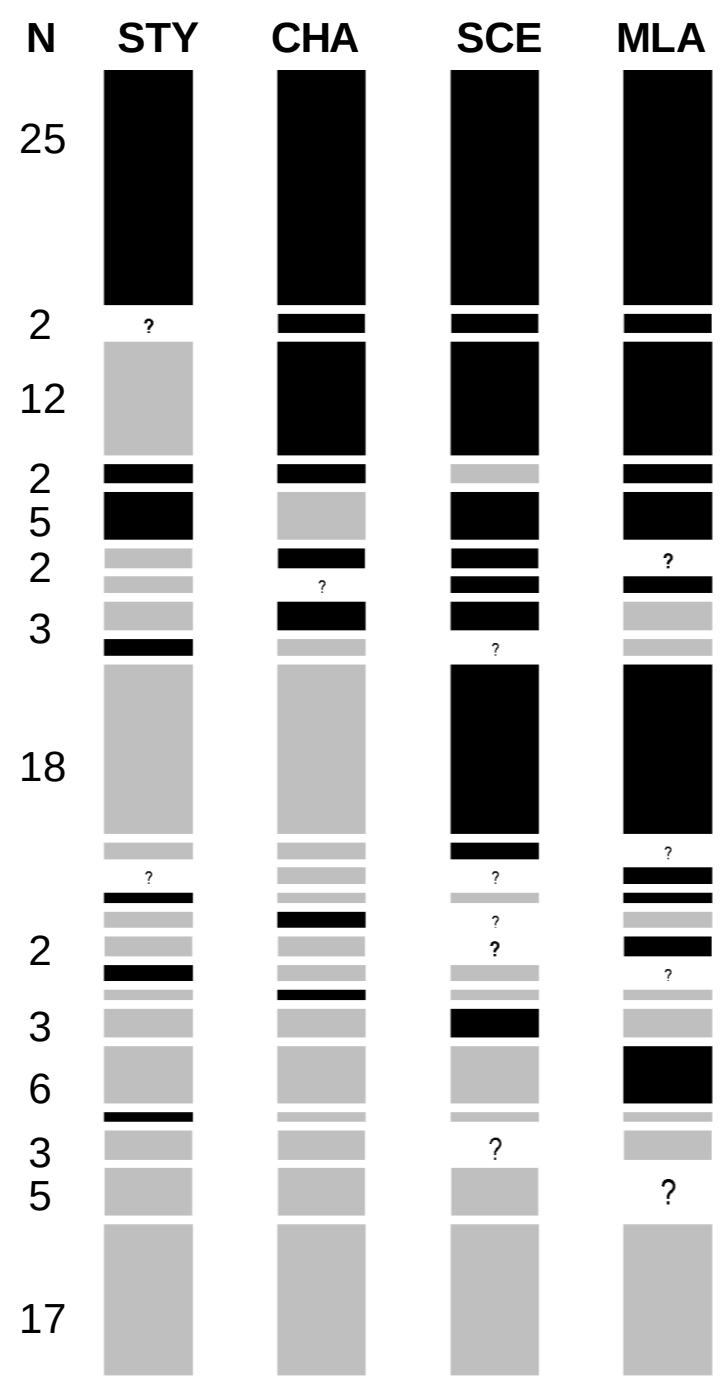


positive

negative

? equivocal

Benigni 1992 Mutagenesis 7: 335-341



Genetic Toxicity Profiles: 114 NTP Chemicals

Positive in Salmonella (and other tests): high correlation with carcinogenicity

Negative in Salmonella and positive in other tests: no correlation with carcinogenicity

positive

negative

? equivocal

Benigni 1992 Mutagenesis 7: 335-341

STTs to predict carcinogenicity: state-of-the-art

- **Mutagenicity = Carcinogenicity ?**

  Only within a limited area of the chemical space, i.e., **DNA-reactive** chemicals

- DNA-reactive chemicals induce **cancer, and a wide spectrum of mutations**

- Most predictive mutagenicity-based assay: **Ames test**

- **mammalian *in vitro* assays** when Ames-negative :
  no correlation with carcinogenicity *(too many false positives)*

- No reliable ***in vivo* STTs** (e.g., micronucleus) *(too many false negatives)*

Benigni R. et al., Exp.Opinion Drug Metab.Toxicol., 2010, 6: 1-11.
Zeiger E Regulat.Pharmacol.Toxicol. 1998;28:85-95.

Ames test *versus* rodent carcinogenicity

Carcinogenicity	Ames test: neg	Ames test: pos
Non carcinogens	233	76
Carcinogens: Non DNA-reactive	136	34
Carcinogens: DNA-reactive	79	277

Ames identifies DNA-reactive carcinogens

Results from 835 chemicals in ISSCAN v3a

http://www.iss.it/ampp/dati/cont.php?id=233&lang=1&tipo=7

Ames test *versus* rodent carcinogenicity

Carcinogenicity	Ames test neg	Ames test pos
Non carcinogens	233	76
Carcinogens: Non DNA-reactive	136	34
Carcinogens: DNA-reactive	79	277

Ames mutagen: 80% probability of being a carcinogen

Results from 835 chemicals in ISSCAN v3a
http://www.iss.it/ampp/dati/cont.php?id=233&lang=1&tipo=7

Increasing recognition of limitations

attempts and debates going on (EFSA, ECVAM, US EPA, etc..)

- **mammalian *in vitro* too sensitive**:
  - Lower doses
  - Manipulate genes
  - Modify positivity criteria
- ***in vivo* too insensitive**:
  - Develop further tests (e.g., Comet assay)

- **omics *in vitro* assays**
  - to trace molecular perturbations related to specific biochemical pathways (e.g., Toxcast project)

All this for the future, to be validated

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

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)

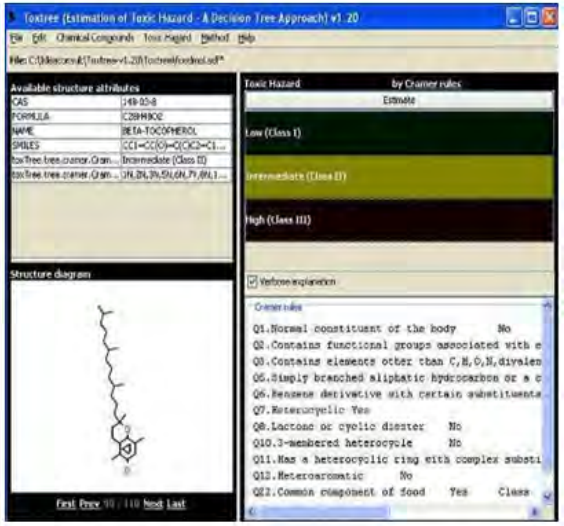
uropean Commission > JRC > IHCP > Our Laboratories > Computational Toxicology and Modelling > QSAR Tools > Toxtree

- 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)*

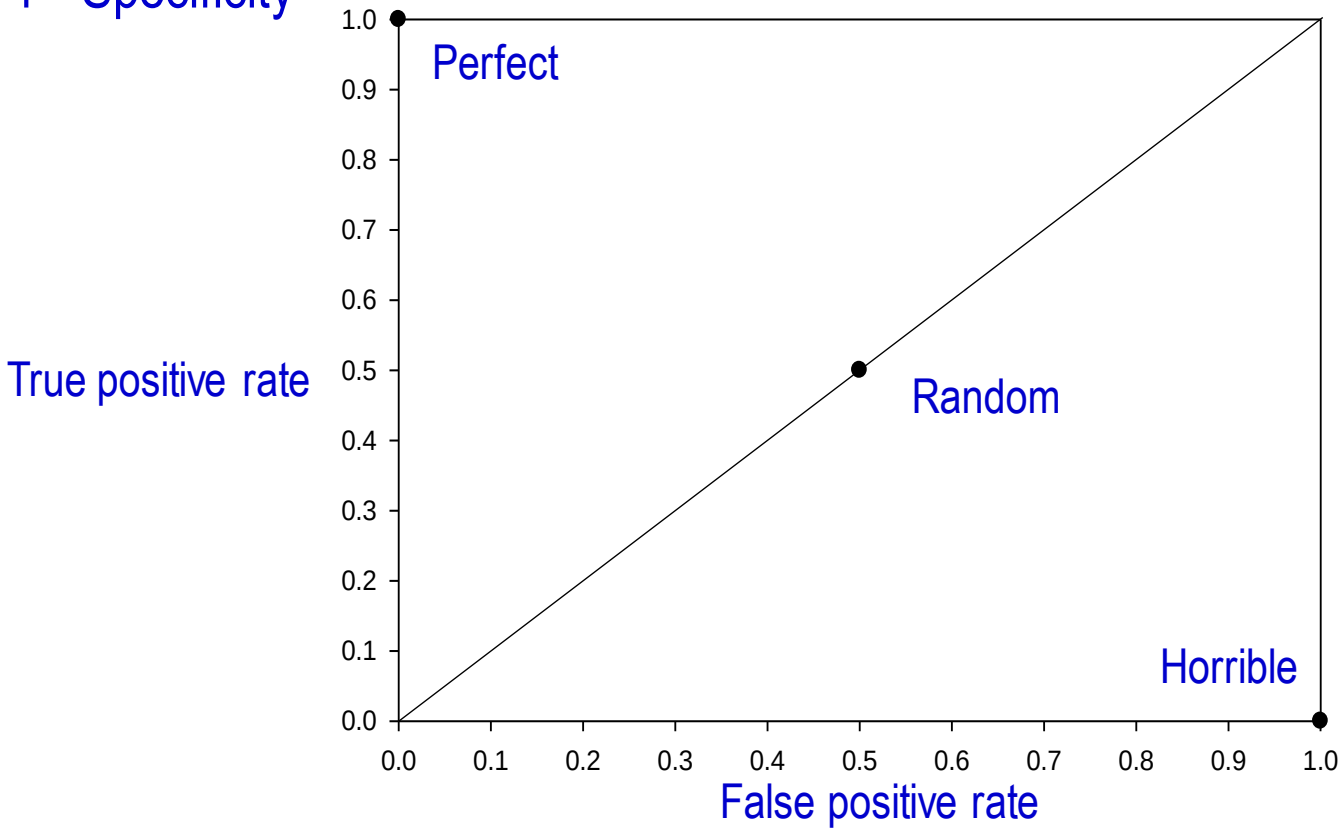
ROC graph: A simple, graphical way of comparing predictions with results

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

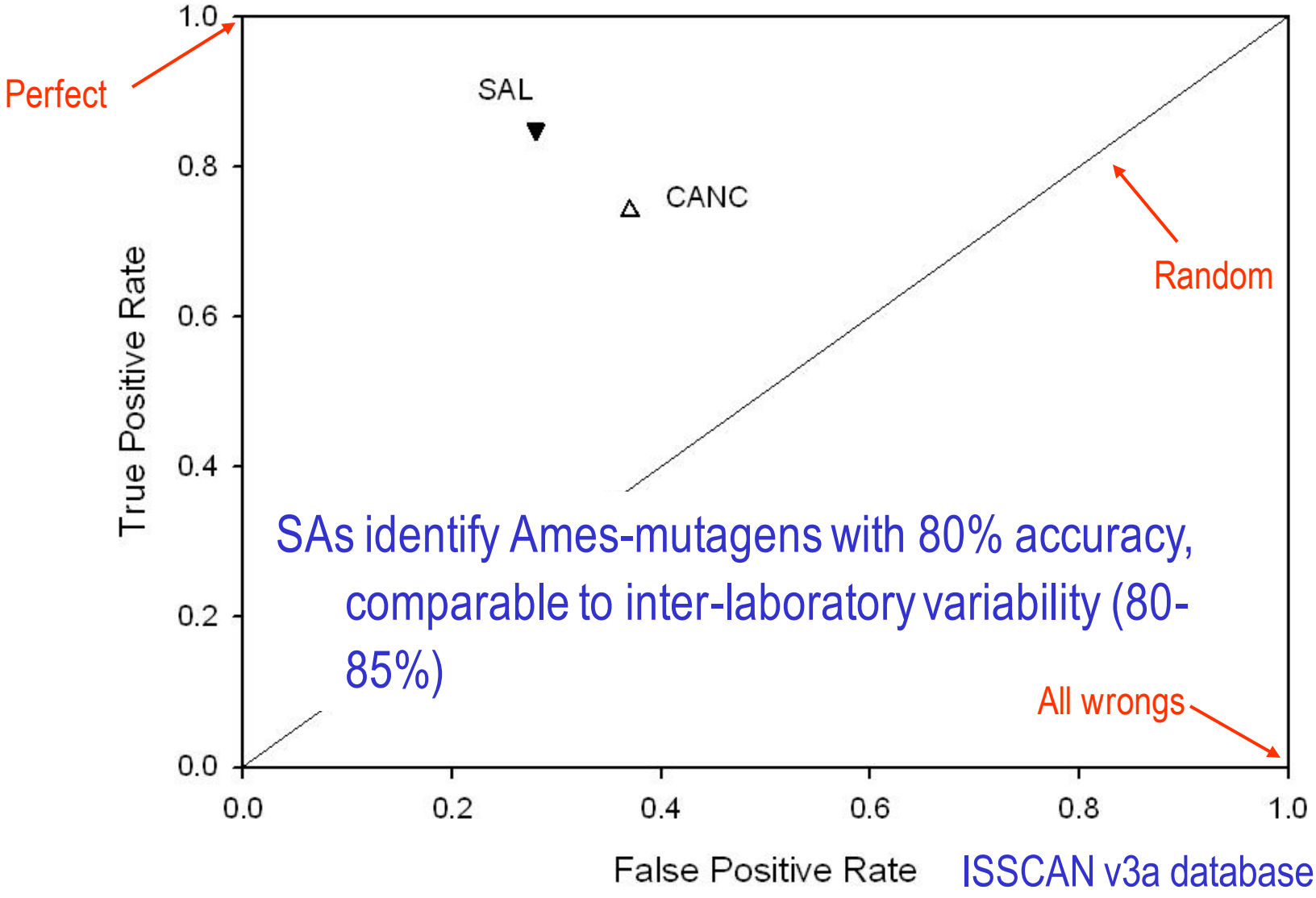
= Sensitivity

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

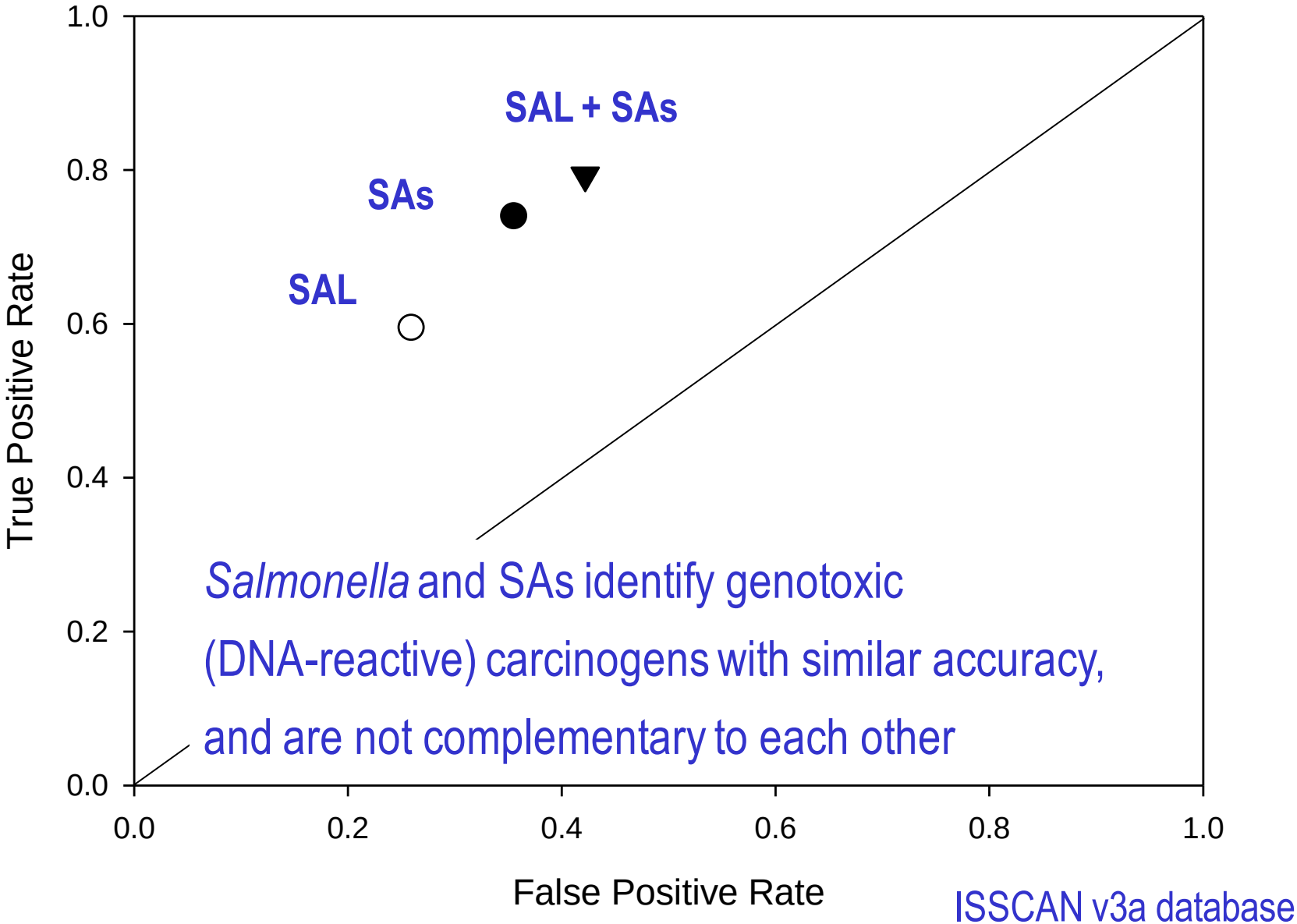
= 1 - Specificity



Canc / Mut SAs: agreement with Carcinogenicity and *Salmonella* (Ames)



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



ISSCAN v3a database

Non-mutagenicity *in vitro* assays for non-DNA-reactive carcinogens ?

ISSCTA: a database of Cell Transformation Assays results

included into the **OECD (Q)SAR Toolbox** and the **ISSTOX** database

n = 370 (including inorganics and organics)

Systems

primary normal diploid cells: Syrian Hamster Embryo cells assay (pH 7)
Syrian Hamster Embryo cells assay (pH 6.7)

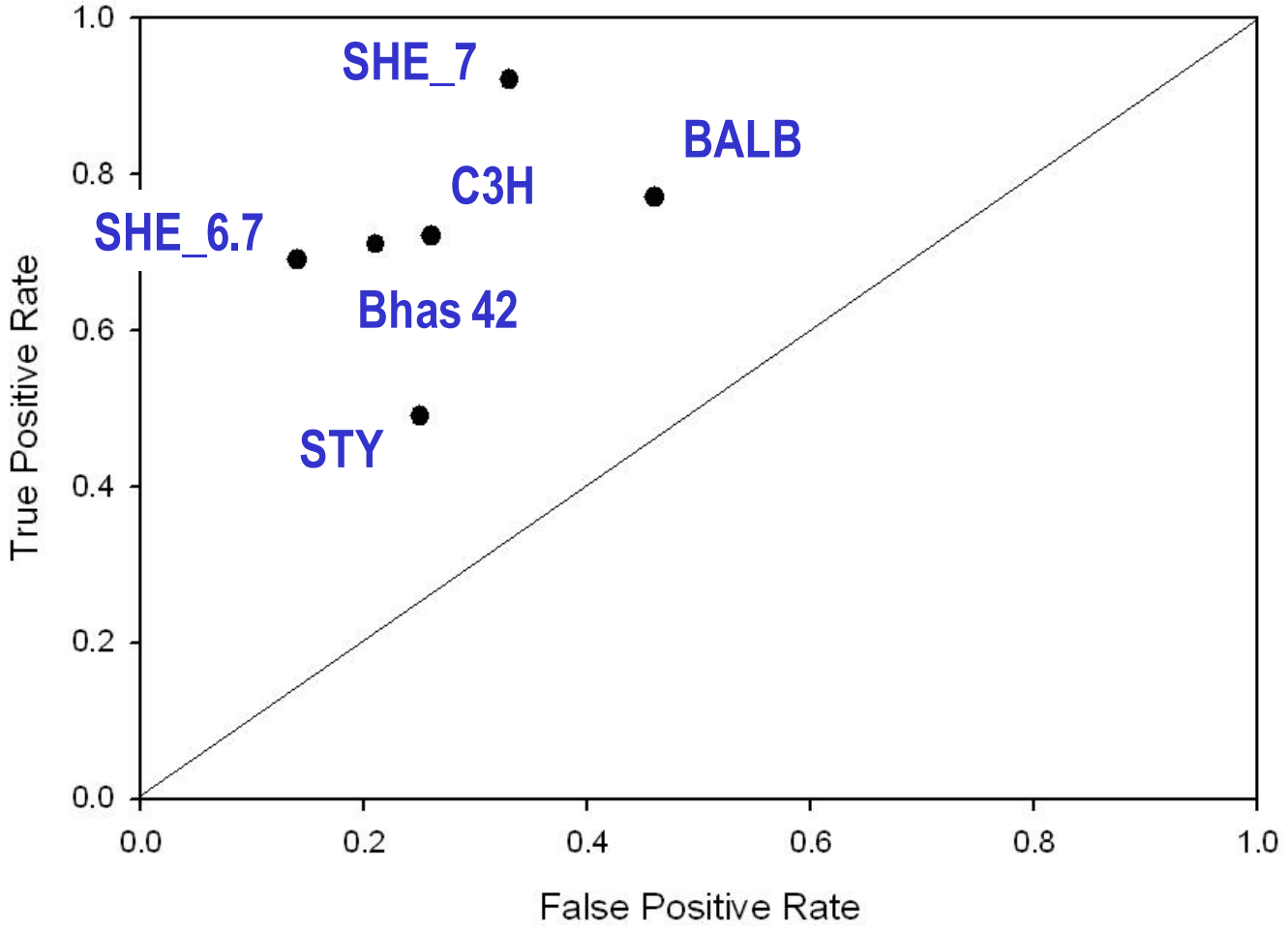
immortalized aneuploid mouse cells: BALB/c 3T3
C3H10T1/2
Bhas 42

Data: *OECD vol.31, 2007*

Sakai et al., 2010 Mutat.Res. 702:100-122

Cell Transformation Assays: carcinogenicity prediction

A new analysis on ISSCTA, including inorganics and Bhas 42



SHE pH≥7 Cell transformation *versus* rodent carcinogenicity

Carcinogenicity	SHE neg	SHE pos
Negatives	36	18
Non DNA-reactive	6	70
DNA-reactive	5	59

Carcinogens

Sensitive to DNA-reactive and non-DNA-reactive carcinogens

Chemicals n = 194, from OECD vol. 31

Improving testing strategy with available tools ?

***in vitro* assays** {*Salmonella* and Cell transformation}
sensitive to both DNA-reactive and non DNA-reactive carcinogens

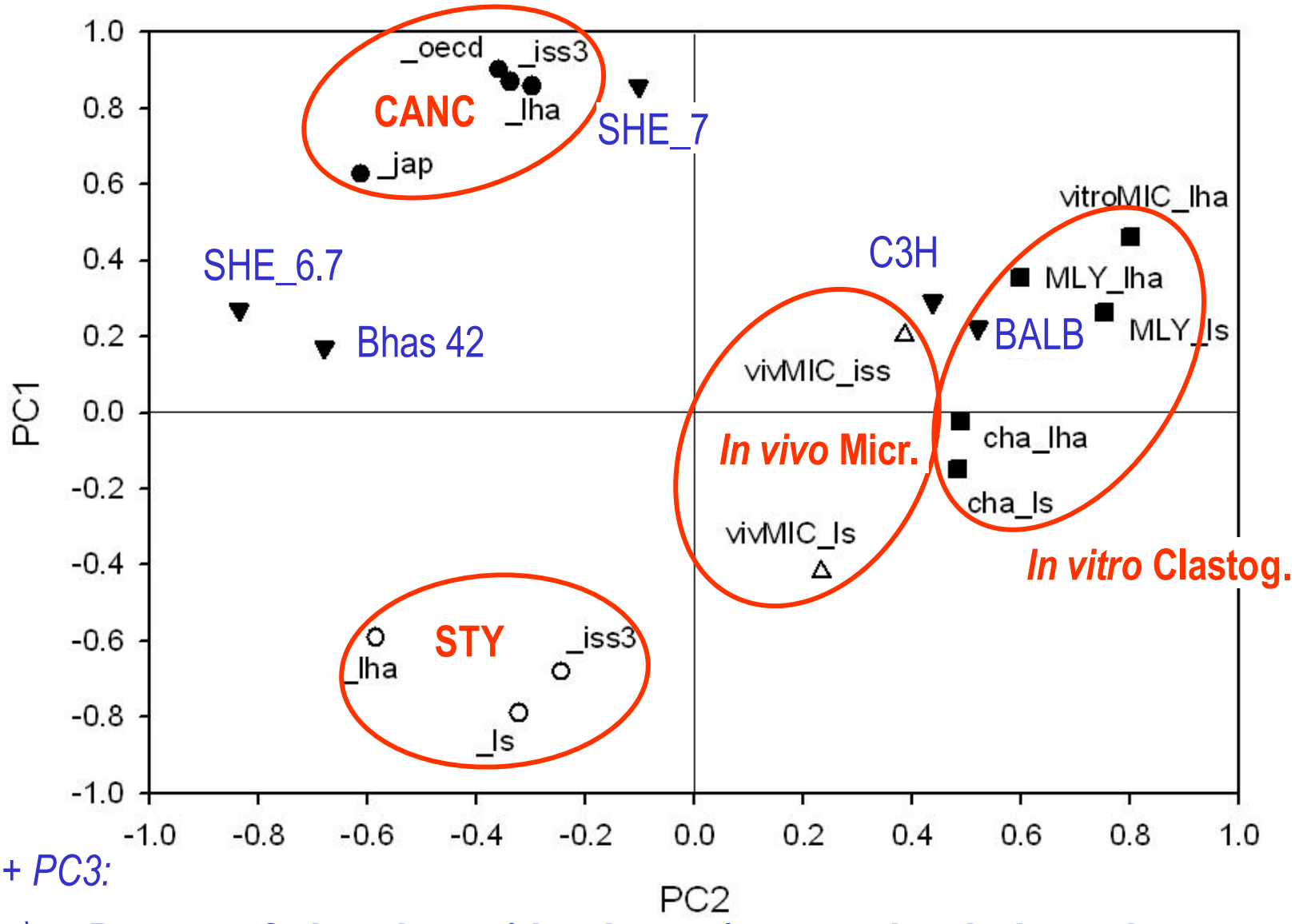
Structural Alerts to predict / rationalize experimental results

A tiered approach to carcinogens identification

only ~ 5-10% undetected carcinogens

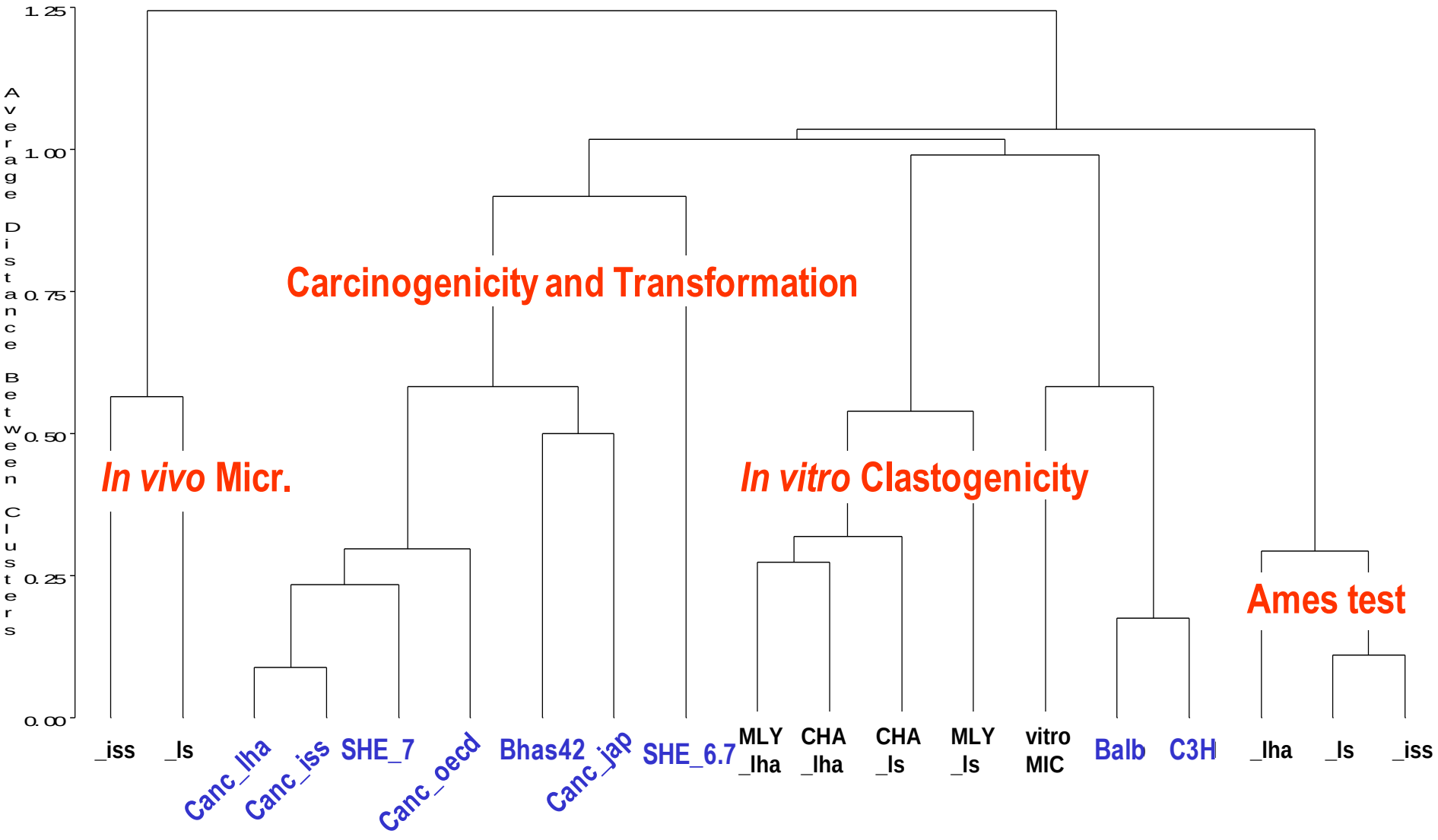
Tiered Approach A			Tiered Approach B		
	Noncarc	Carc		Noncarc	Carc
Initial Sample	36	86	Initial Sample	52	130
	After Tier 1			After Tier 1a	
STY Negative	27	41	SAgeno Negative	32	66
				After Tier 1b	
			SA nongeno Negative	27	43
	After Tier 2			After Tier 2	
SHE Negative	17	8	SHE Negative	17	5
% initial sample	47 %	9%		33%	4%

Meta-"omics" of toxicological endpoints



(PC1 + PC2 + PC3: 0.74 Exp. Var.)

Panels of chemicals (databases) as mechanistic probes


Average Distance Between Clusters
1.25
1.00
0.75
0.50
0.25
0.00
Carcinogenicity and Transformation
In vivo Micr.
In vitro Clastogenicity
Ames test
_iss
_ls
Canc_lha
Canc_iss
SHE_7
Canc_oecd
Bhas42
Canc_jap
SHE_6.7
MLY _lha
CHA _lha
CHA _ls
MLY _ls
vitro MIC
Balb
C3H
_lha
_ls
_iss

Tiered Approach to the identification of IARC carcinogens

Group 1, human carcinogens: **69 / 70**

(except ethanol alcohol in beverages)

Group 2a, probably carcinogenic to humans: **47 / 47**

(only on the basis of STY and SAs)

Group 2b, possibly carcinogenic to humans: **210 / 212**

(exceptions areTitanium dioxide and Nitrilotriacetic acid. Titanium dioxide is classified by IARC as a group together with its salts, so it is not an individual chemical species suited for experimental tests. Nitrilotriacetic acid is a chelating agent whose carcinogenic effects are likely to be secondary ones, thus being difficult to be deduced from its very structure)

Conclusions and Perspectives

- Simpler ad more efficient strategies for *in vitro* pre-screening of carcinogenicity already available

- SAs play a key role

- Expand SAs for nongenotoxic carcinogens to decrease reliance on experiments (mechanistic knowledge plus data mining)

- Need to improve overall specificity

- **Freely-available tools generated at the ISS**

Structural Alerts for - canc / *in vitro* mut
- *in vivo* micronucleus

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

OECD (Q)SAR Toolbox
http://www.oecd.org/document/54/0,3746,en_2649_37465_42923638_1_1_1_37465,00.html

ISSTOX curated toxicological databases
http://www.iss.it/meca/index.php?lang=1&anno=2013&tipo=25

- **Acknowledgements**

Alessandro Giuliani

Cecilia Bossa

Olga Tcheremenskaia

Chiara Battistelli

Mauro Colafranceschi

References

- Lebart R., Morineau A., and Warwick K.M. . (1984): Multivariate descriptive statistical analysis, New York: Wiley.

- Benigni R. and A. Giuliani (1994) Quantitative modeling and biology: the multivariate approach, *Amer. J. Physiol., 266 (Regulatory Integrative Comp. Physiol.* 35): R1697-R1704.

- Benigni R. (2012) Alternatives to the carcinogenicity bioassay for toxicity prediction: are we there yet?, *Expert Opin.Drug Metab.Toxicol.*, 8: 1-11.

- R. Benigni, C. Bossa, C. Battistelli, and O. Tcheremenskaia (2013) IARC Class 1 and 2 carcinogens are successfully identified by an alternative strategy that detects DNA-reactivity and Cell Transformation ability of chemicals. *Mutat.Res.,* accepted.