UNC ESHELMAN SCHOOL OF PHARMACY

MML UNC.EDU

A Combined Use of *In Vitro* Screening and Cheminformatics Approaches Improves the Accuracy of *In Vivo* Toxicity Models

Alexander Tropsha, Ivan Rusyn, Hao Zhu, Denis Fourches, Lin Ye, Ann Richard, Todd Martin

UNC-Chapel Hill and US EPA

Carolina Center for Computational Toxicology

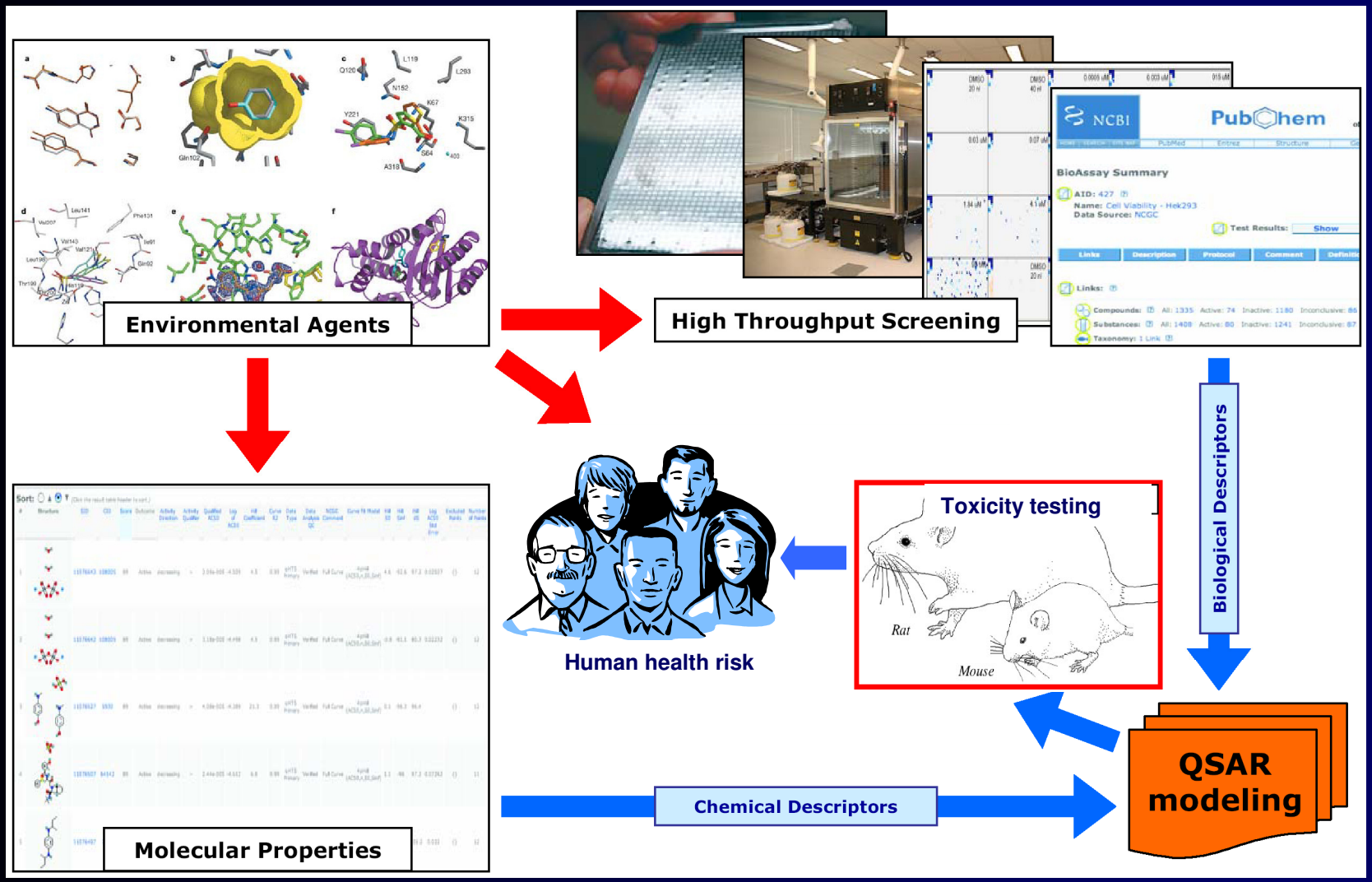
and

Carolina Center for Environmental Bioinformatics

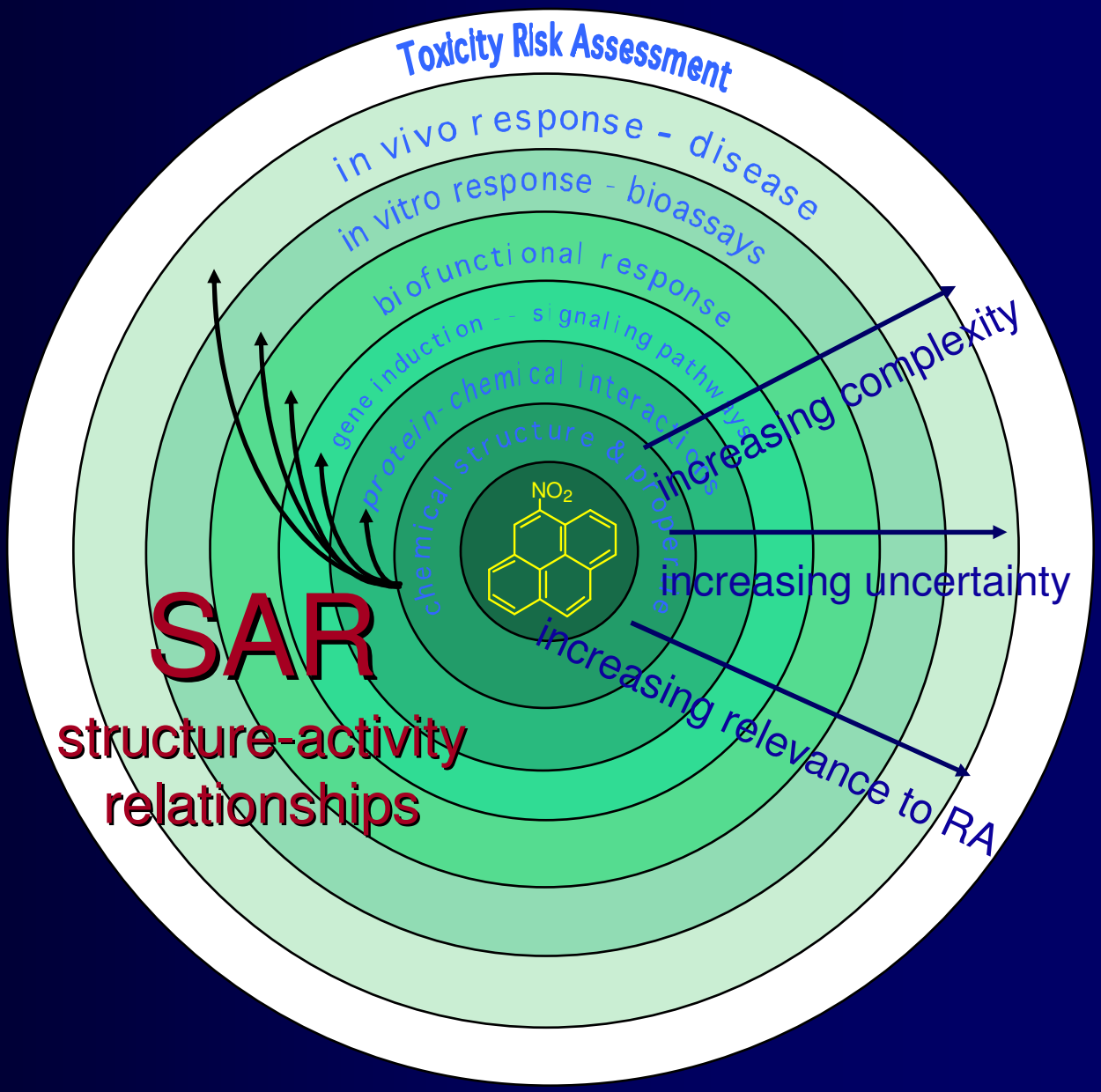
OUTLINE

- Introduction: the chemical structure – toxicity modeling continuum
- A little bit of methodology: predictive QSTR modeling workflow
- Applications
  - Prediction of chemical carcinogenicty in rodents using hybrid chemical and biological descriptors
  - Consensus QSAR modeling of aquatic toxicity
  - Structure – In vitro – In vivo Correlations: Biological Data Partitioning and Hierarchical Modeling of Rodent Chemical Toxicity
  - Concordance between animal and human DILI data
- Conclusions: Toxico-cheminformatics is a decision support tool

Chemical Structure – Toxicity Data Continuum.

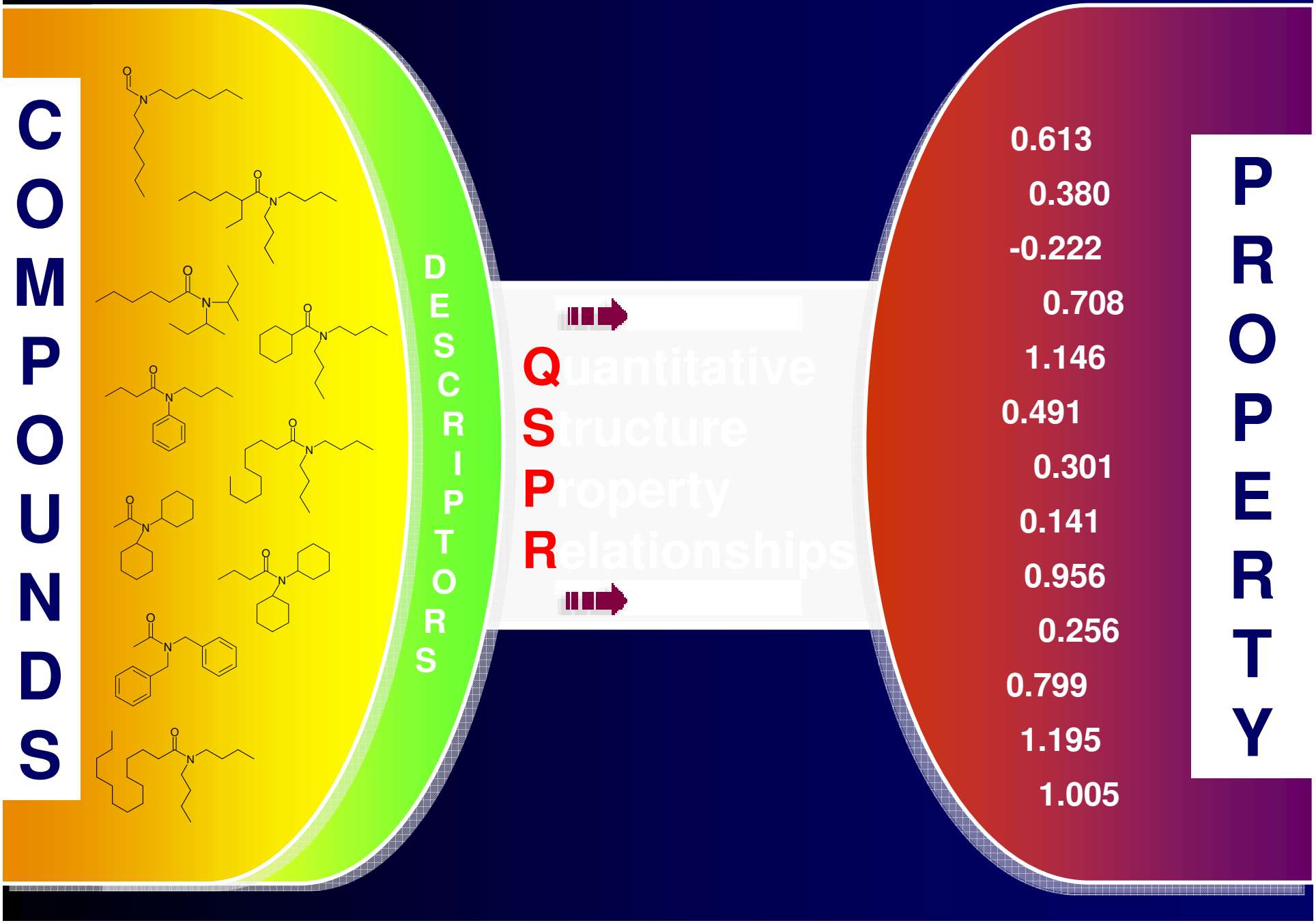


Chemocentric view of biological data



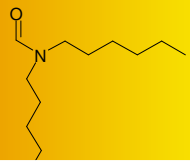
Slide courtesy of Dr. Ann Richard (EPA)

Principles of QSAR/QSPR modeling

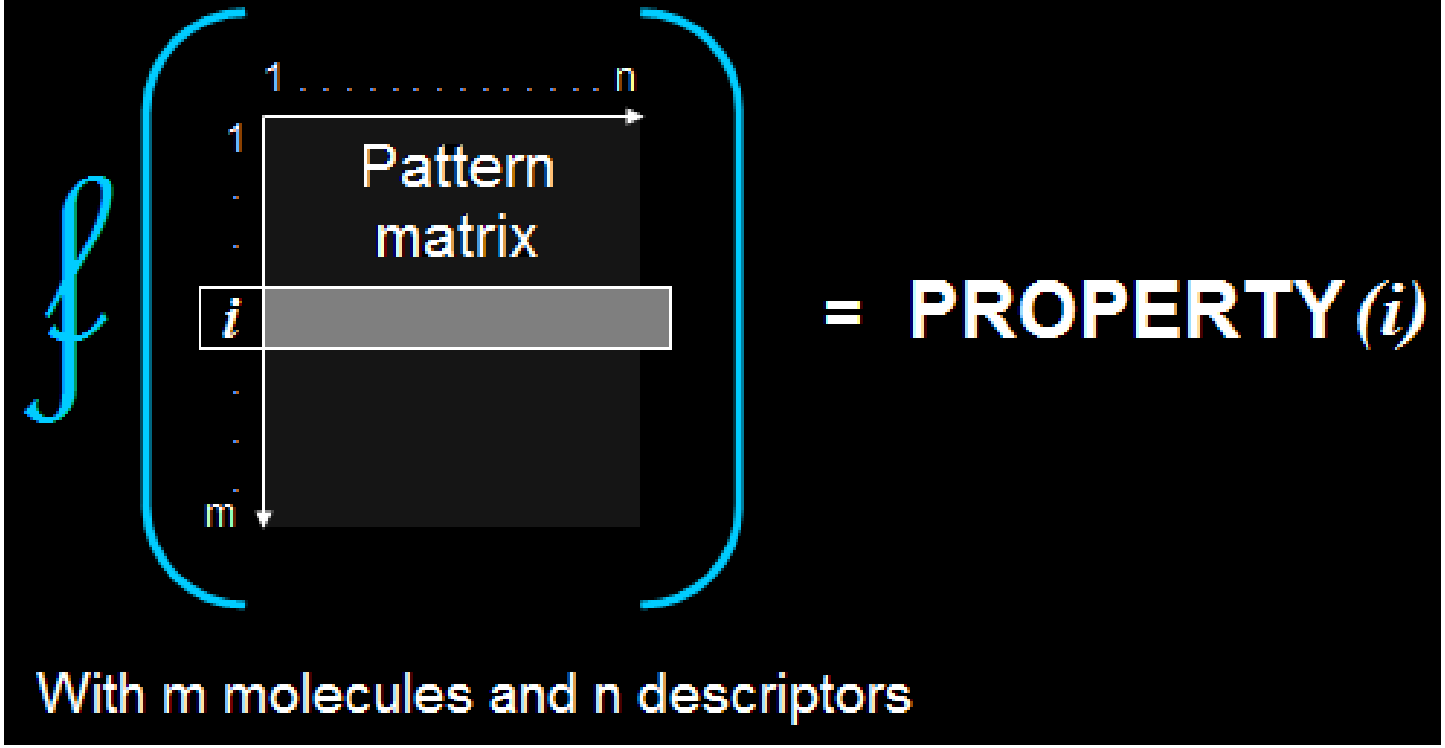


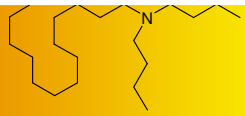
Principles of QSAR/QSPR modeling

COMPOUNDS



0.613

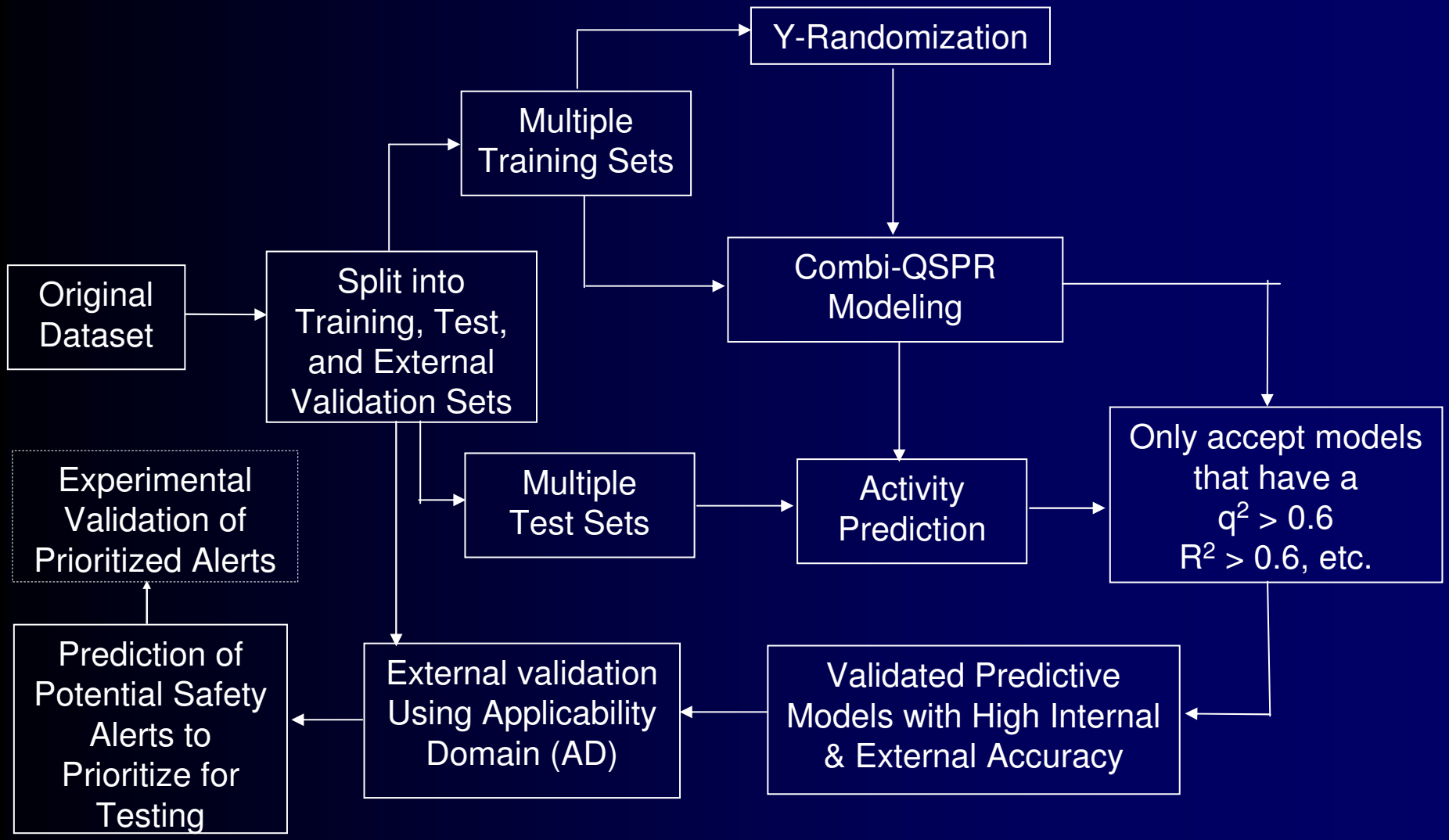




1.005

PROPERTY

Predictive QSAR Workflow*



Tropsha, A., Golbraikh, A. Predictive QSAR Modeling Workflow, Model Applicability Domains, and Virtual Screening. Curr. Pharm. Des., 2007, 13, 3494-3504.

Experimental Study I: The Use of High Throughput Screening Data as Additional Biological Descriptors Improves the Prediction Accuracy of Conventional QSAR Models of Chemical Carcinogenicity*

Zhu et al, EHP, 2008, (116): 506-513

NTP-HTS Content Summary of 1408 Compounds

- Chemical Types:
  - Organic: 1,348
  - Inorganic: 27
  - Organometallic: 19
  - No structure: 14
- 1348 Organic compounds contain:
  - Normal: 1,279
  - Complex: 51
  - Salts: 20
  - Duplicates: 53
- Finally, 1,289 unique organic compounds identified

Characteristics of the Experimental Activities of 1,289 Compounds

	BJ	Jurkat	Hek293	HepG2	MRC5	SK-N-SH	General
Actives	42	121	63	41	37	74	140
Inconclusives	44	89	79	47	44	54	90
Inactives	1,203	1,079	1,147	1,201	1,208	1,161	1,059

Additional biological data on 1,289 Compounds

NTP-HTS	NTPBSI	NTPGTZ	HPVCSI	**CPDB**	IRISSI
1,289	1,153	1,053	423	**383**	181

NTPBSI: National Toxicology Program Chemical Structure Index file
NTPGTZ: National Toxicology Program genotoxicity
HPVCSI: High Production Volume Chemicals
CPDB: Carcinogenic Potency Data Base All Species
IRISSI: EPA Integrated Risk Information System

The table is based on the DSSTox project of Dr. Ann Richard at EPA.

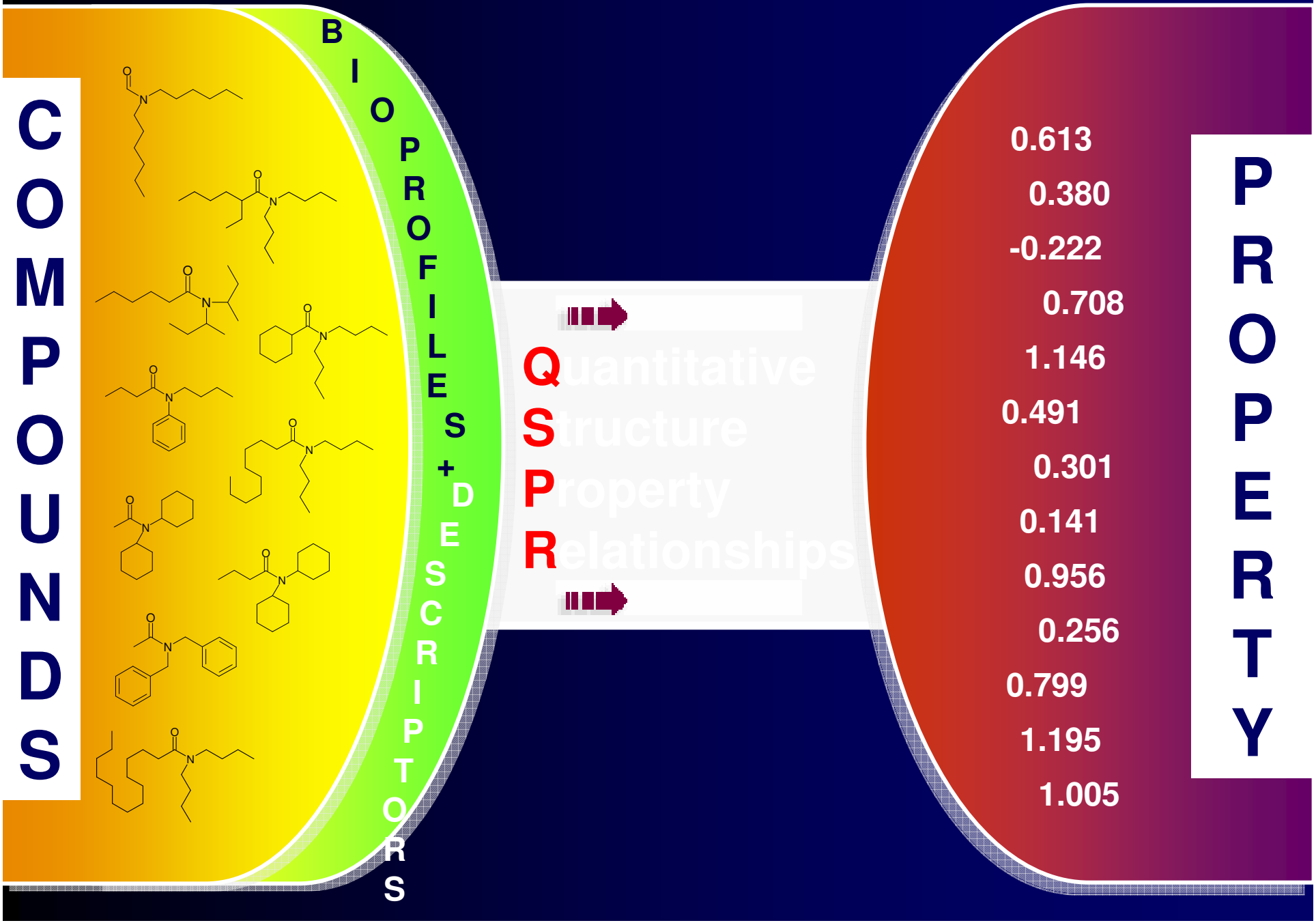
Division of the dataset into modeling and external Sets

- 314 out of 383 CPDB compounds after removing 69 compounds with inconclusive carcinogenicity results.
- Randomly excluded 50 compounds as external test set.
- Using sphere exclusion approach to split the remaining 264 compounds into multiple training/test set pairs.

The Relationship between HTS Results and Rodent Carcinogenicity

	HTS actives	**HTS inconclusives**	**HTS inactives**
CPDB actives	30	12	136
CPDB Inactives	9	13	114
Correlation	77%	-	46%

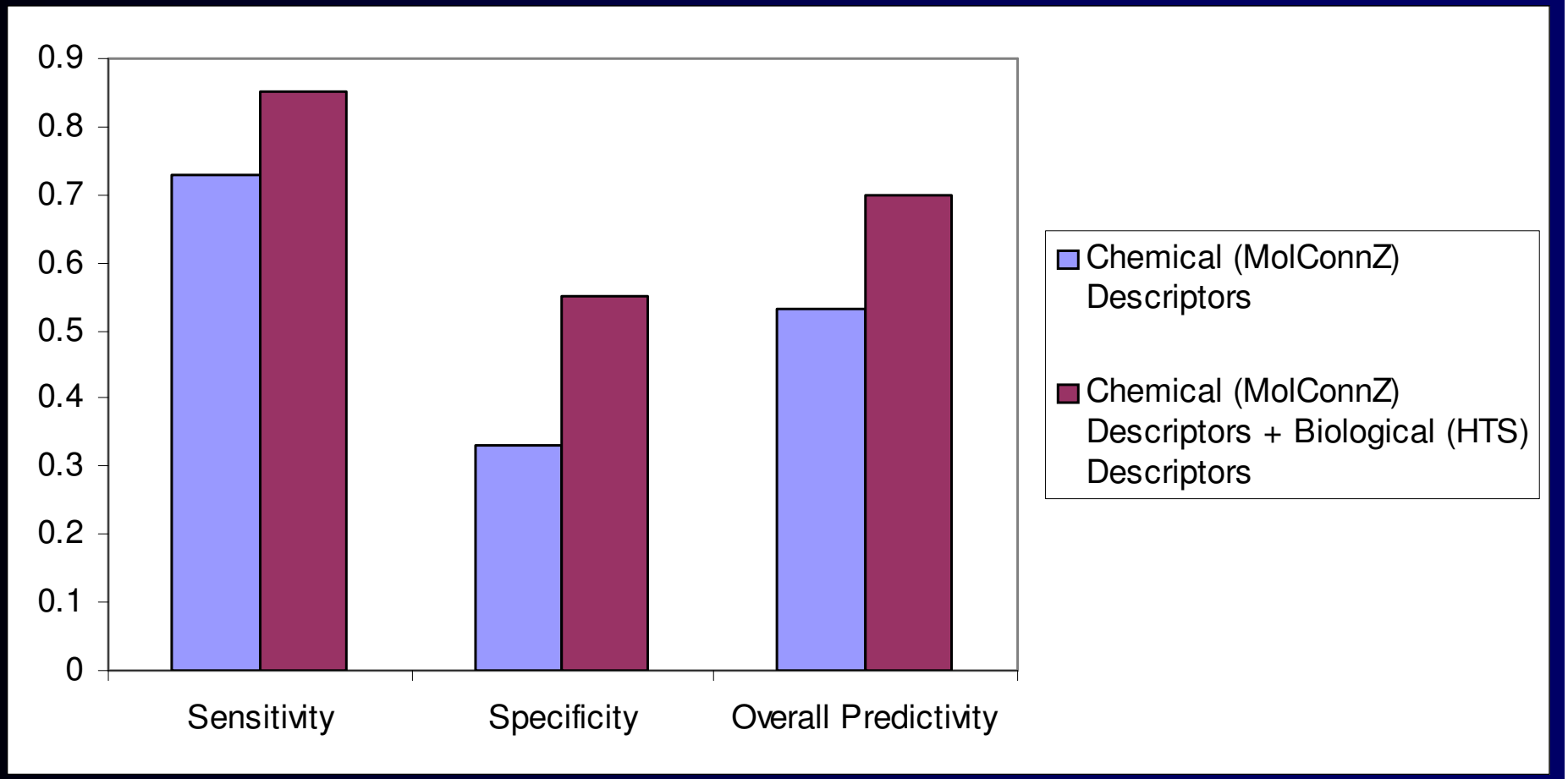
Principle of QSAR/QSPR modeling



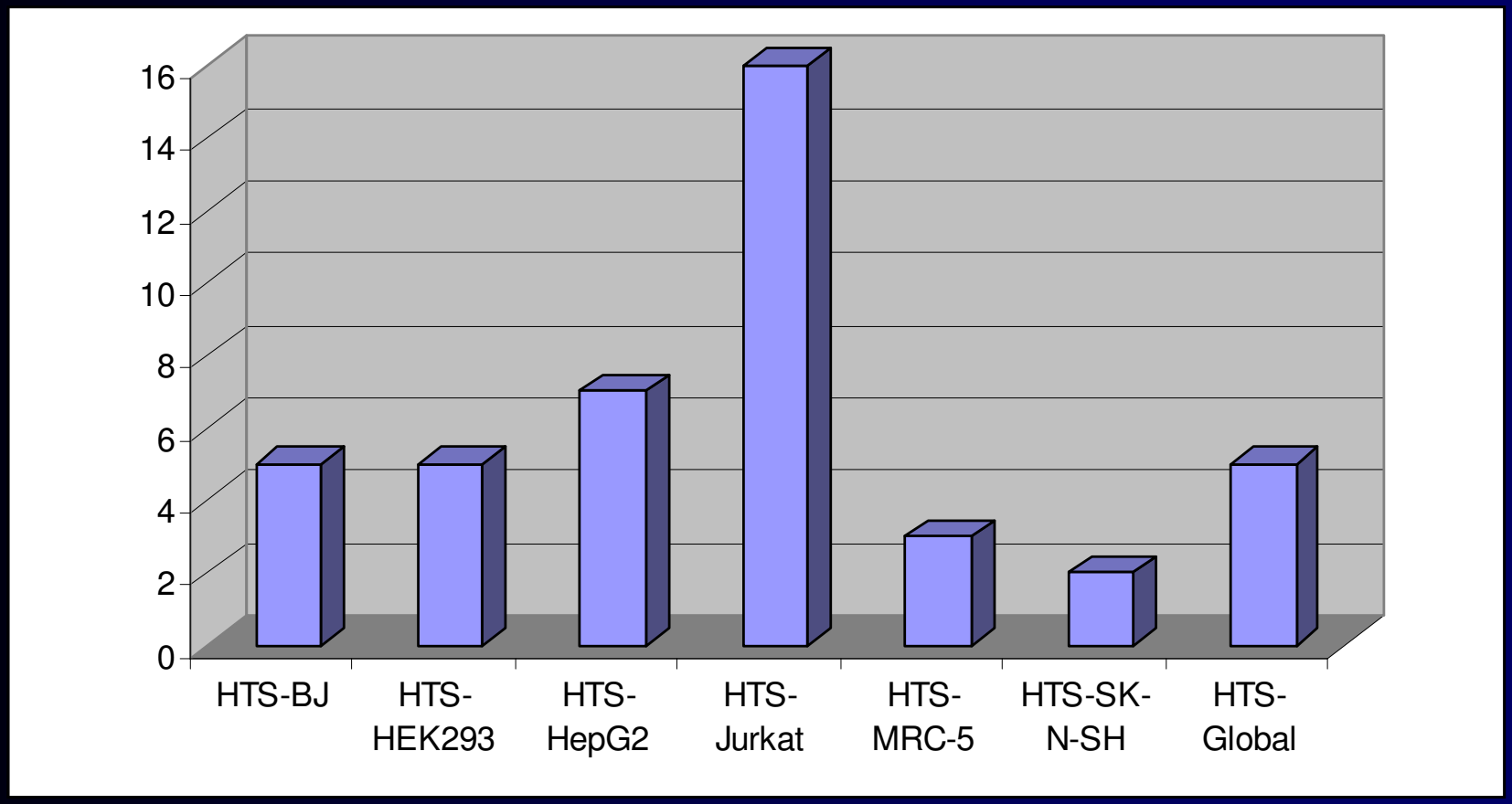
Prediction accuracy for the external dataset of 50 Compounds

	Chemical descriptors only		Combined descriptors	
	Exp. Actives	Exp. Inactives	Exp. Actives	Exp. Inactives
Pred. actives	18	8	22	6
Pred. inactives	8	10	6	12
Predictivity	69.2%	55.5%	78.6%	66.7%
Overall Predictivity	**62.3%**		**72.7%**	
Coverage	88%		92%	

Comparison between Predictive Power of QSAR Models using Conventional vs. Hybrid Descriptors.



Relative contributions of HTS descriptors to 34 acceptable models



Conclusions of the Study I

- 1. NTP-HTS screening data have limited predictive power for rodent carcinogenicity.
- 2. Using the NTP-HTS data as biological fingerprint descriptors significantly improved the overall QSAR-based prediction accuracy of rodent carcinogenicity.
- 3. With sufficient improvements in resulting model predictive performance, *in vitro* HTS bioassays, coupled with traditional chemical structure-based descriptors, may be ultimately helpful in prioritizing or partially replacing *in vivo* toxicity testing

Experimental Study II: Combinatorial QSAR Modeling of Chemical Toxicants Tested against *Tetrahymena pyriformis**

*Zhu et al, JCIM, 2008, (48): 766-784 ; Tetko et al, JCIM, 2008, ASAP

International Virtual Collaboratory* of Computational Chemical Toxicology

- USA: UNC-Chapel Hill (UNC) - H. Zhu and A. Tropsha
- France: University of Louis Pasteur (ULP) – D. FOURCHES and A. VARNEK
- Italy: University of Insubria (UI) – E. PAPA and P. GRAMATICA
- Sweden: University of Kalmar (UK) – T. ÖBERG
- Germany: Munich Information Center for Protein Sequences/Virtual Computational Chemistry Laboratory (VCCLAB)– I. TETKO
- Canada: University of British Columbia (UBC) – A. CHERKASOV

*a new networked organizational form that also includes social processes; collaboration techniques; formal and informal communication; and agreement on norms, principles, values, and rules

The *T. pyriformis* toxicity dataset

- Compiled from several publications of T. Schultz's group (2001-2005) and the Tetratox website (http://www.vet.utk.edu/TETRATOX/)
- Corrected over 100 errors (chemical structures, chemical name and CAS ids).
- 983 unique compounds: 644 compound in modeling set; 339 compound in the external validation set I.
- 110 new compounds from a recent publication (Schultz et al, 2007) and used as the external validation set II.

Different countries, different groups, different tools – shared basic principles

- Explore and combine various QSAR approaches
- Use extensive model validation and applicability domains
- Consider external prediction accuracy as the ultimate criteria of model quality

$$Q^2_{abs} = 1 - \sum_Y (Y_{\exp} - Y_{LOO})^2 \Big/ \sum_Y (Y_{\exp} - \langle Y \rangle_{\exp})^2 \qquad (1)$$

$$R^2_{abs} = 1 - \sum_Y (Y_{\exp} - Y_{pred})^2 \Big/ \sum_Y (Y_{\exp} - \langle Y \rangle_{\exp})^2 \qquad (2)$$

$$MAE = \sum_Y \left| Y - Y_{pred} \right| \Big/ n \qquad (3)$$

Overview of the Approaches (15 methodologies total)

Group ID	Modeling Techniques	Descriptor Type	Applicability Domain
UNC	*k*NN, SVM	MolConnZ, Dragon	Euclidean distance threshold between a test compound and compounds in the modeling set
ULP	MLR, *k*NN, SVM	Fragments	Euclidean distance threshold between a compound and compounds in the modeling set; bounding box
UI	OLS	Dragon	Leverage approach
UK	PLS	Dragon	Residual standard deviation and leverage within the PLSR model
MIPS	ASNN	E-state	Maximal correlation coefficient of the test molecule to the training set molecules in the space of models
UBC	MLR, ANN, SVM, PLS	IND_I	Descriptor variability

The Prediction of the Two Evaluation Sets by Consensus Models

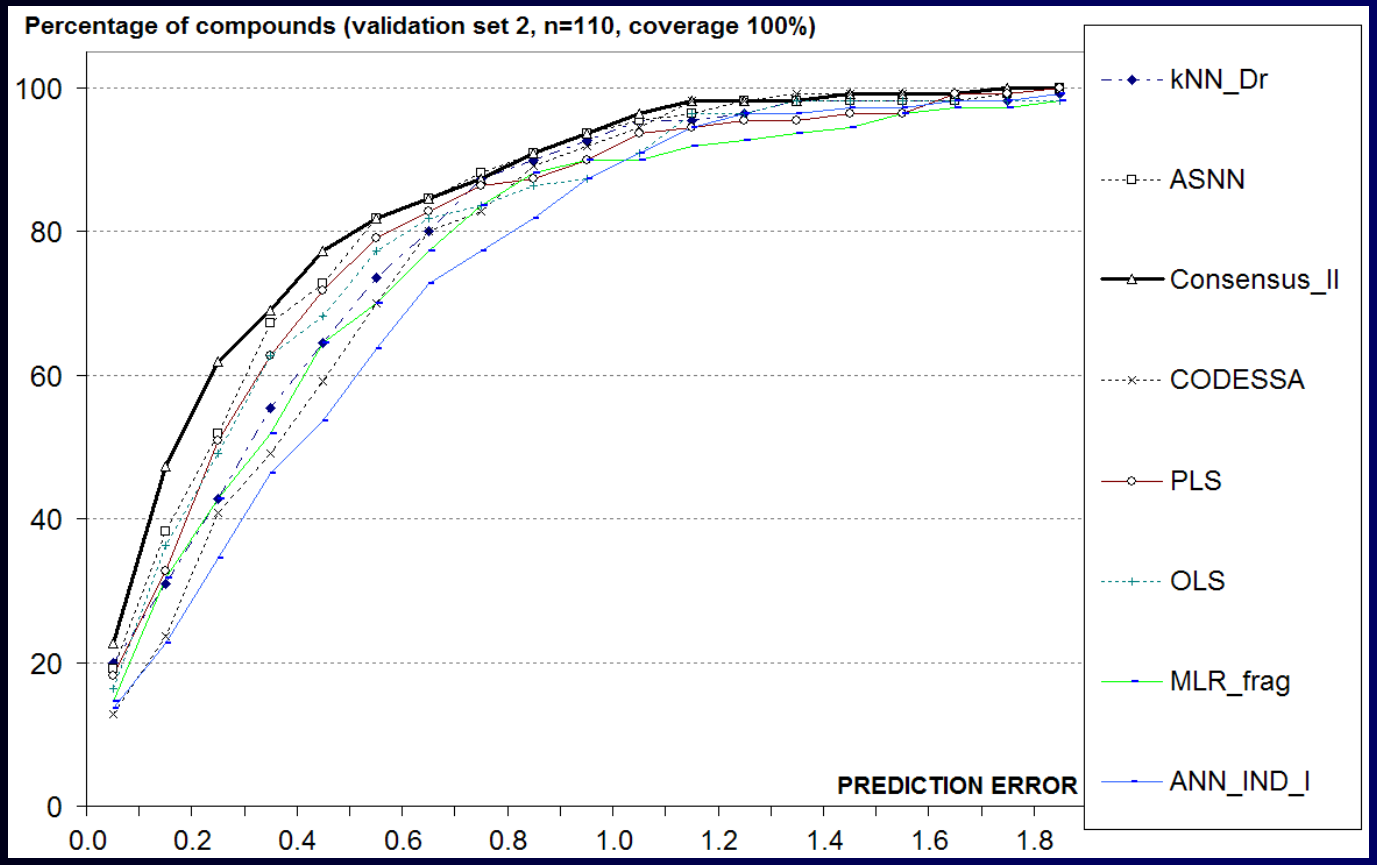
Model	Group ID	1st Evaluation Set (n=339)			2nd Evaluation Set (n=110)		
		R_I^2	SE_I	Coverage	R_{II}^2	SE_{II}	Coverage
*k*NN-Dragon	UNC	0.87	0.30	80.2%	0.77	0.29	52.7%
*k*NN-MolconnZ	UNC	0.86	0.31	84.3%	0.50	0.36	53.6%
SVM-Dragon	UNC	0.82	0.39	80.2%	0.83	0.31	52.7%
SVM-MolconnZ	UNC	0.84	0.37	84.3%	0.59	0.41	53.6%
*k*NN-Fragmental	ULP	0.71	0.47	100%	0.41	0.53	100%
SVM-Fragmental	ULP	0.78	0.49	100%	0.46	0.62	100%
MLR	ULP	0.82	0.43	97.3%	0.48	0.62	95.5%
MLR-CODESSA	ULP	0.72	0.47	100%	0.59	0.44	100%
OLS	UI	0.77	0.43	98.5%	0.59	0.49	98.2%
PLS	UK	0.81	0.40	96.1%	0.60	0.49	95.5%
ASNN	MISP	0.88	0.33	87.4%	0.76	0.40	71.8%
PLS-IND_I	UBC	0.74	0.39	99.7%	0.45	0.54	100%
MLR-IND_I	UBC	0.75	0.40	99.7%	0.46	0.53	100%
ANN-IND_I	UBC	0.76	0.39	99.7%	0.46	0.53	100%
SVM-IND_I	UBC	0.79	0.35	99.7%	0.53	0.46	100%
Consensus Model	-	0.87	0.27	100%	0.70	0.34	100%

Which model is best?

- Observation: Models that afford most accurate predictions for the validation sets are not necessarily ranked as top models for the modeling set.
- Back to choices and practices: So how do we choose "the best" models?

  Should we choose?
- Consensus Prediction
  - Only predict compounds within the applicability domain of most models
  - For each compound, exclude predictions that have high deviations from the mean value
  - Final predicted value is the average all predictions.

Consensus Model gives the lowest MAE of prediction (Validation Set II)



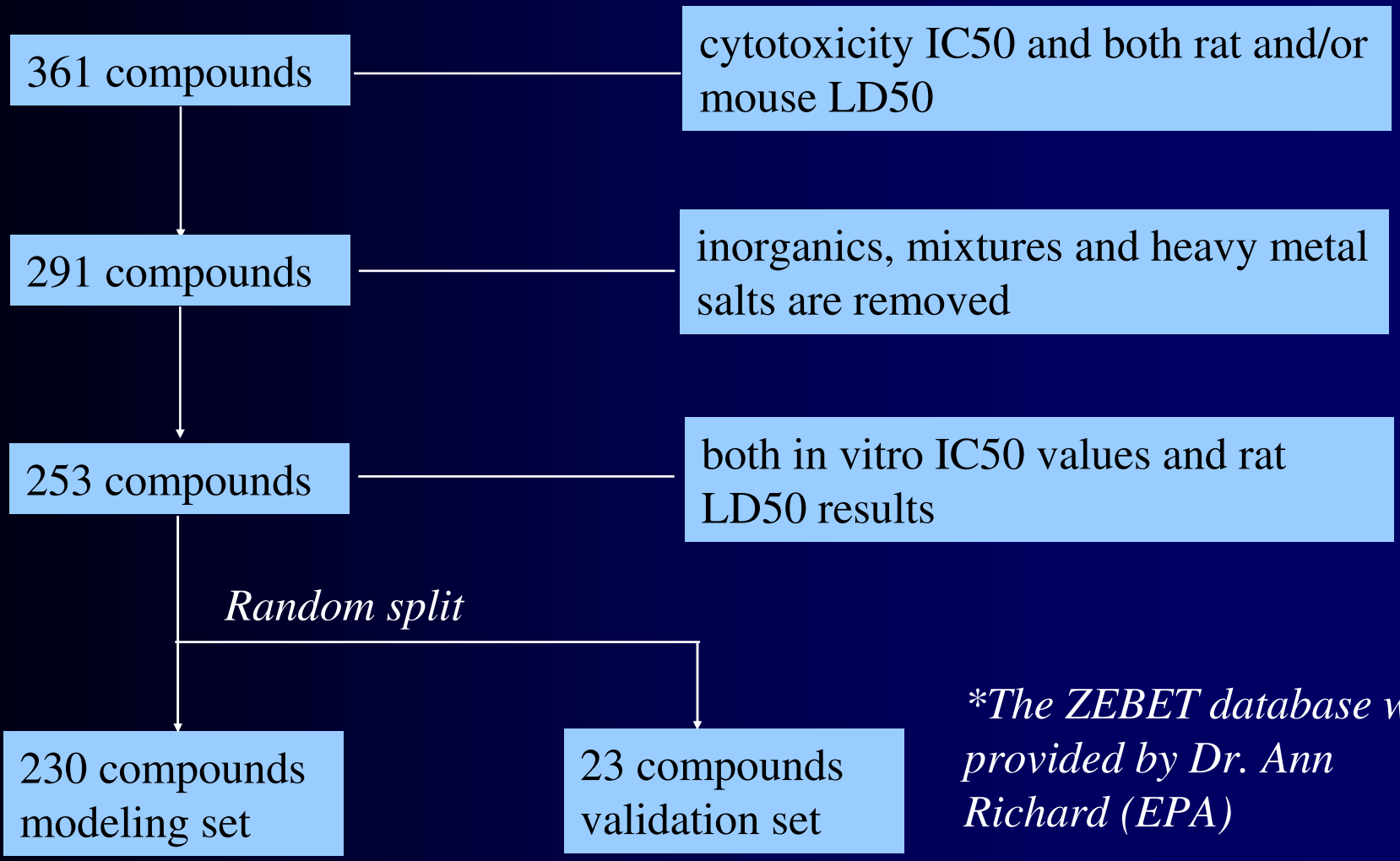
Conclusions of the aquatic toxicity modeling

- Training set modeling is insufficient to guarantee externally predictive models
- The use of AD is critical to achieve respectable external predictivity of individual models BUT one should keep in mind the balance between predictivity and space coverage
- Consensus prediction
  - affords the high predictive power
  - lowest MAE
  - stable against relatively inefficient individual models
  - Avoids the problem of choice!!!

Experimental Study III: Structure – In vitro – In vivo Correlations: Biological Data Partitioning and Hierarchical Modeling of Rodent Chemical Toxicity*

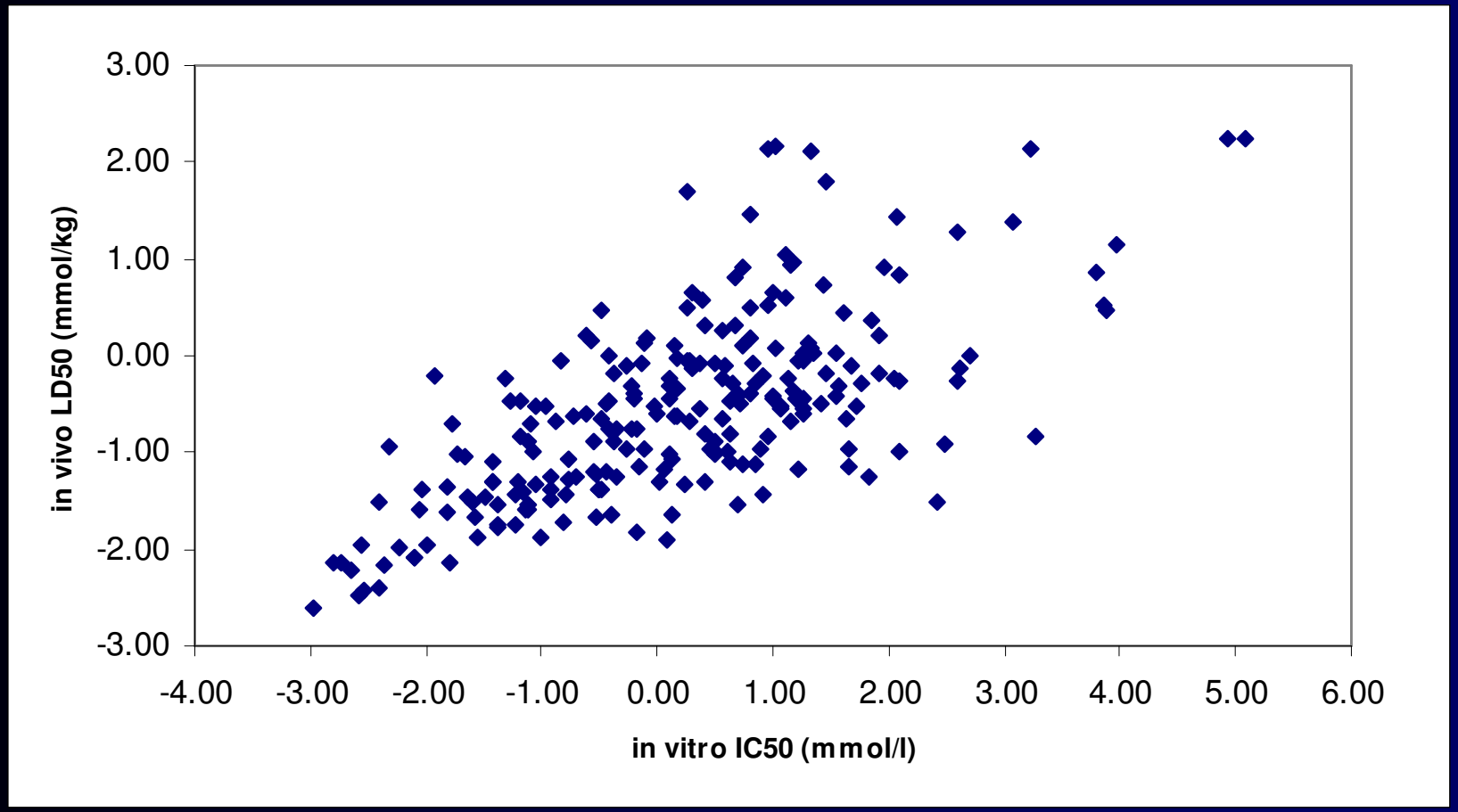
*Zhu et al, in preparation

ZEBET Database* and Data Preparation



**The ZEBET database was provided by Dr. Ann Richard (EPA)*

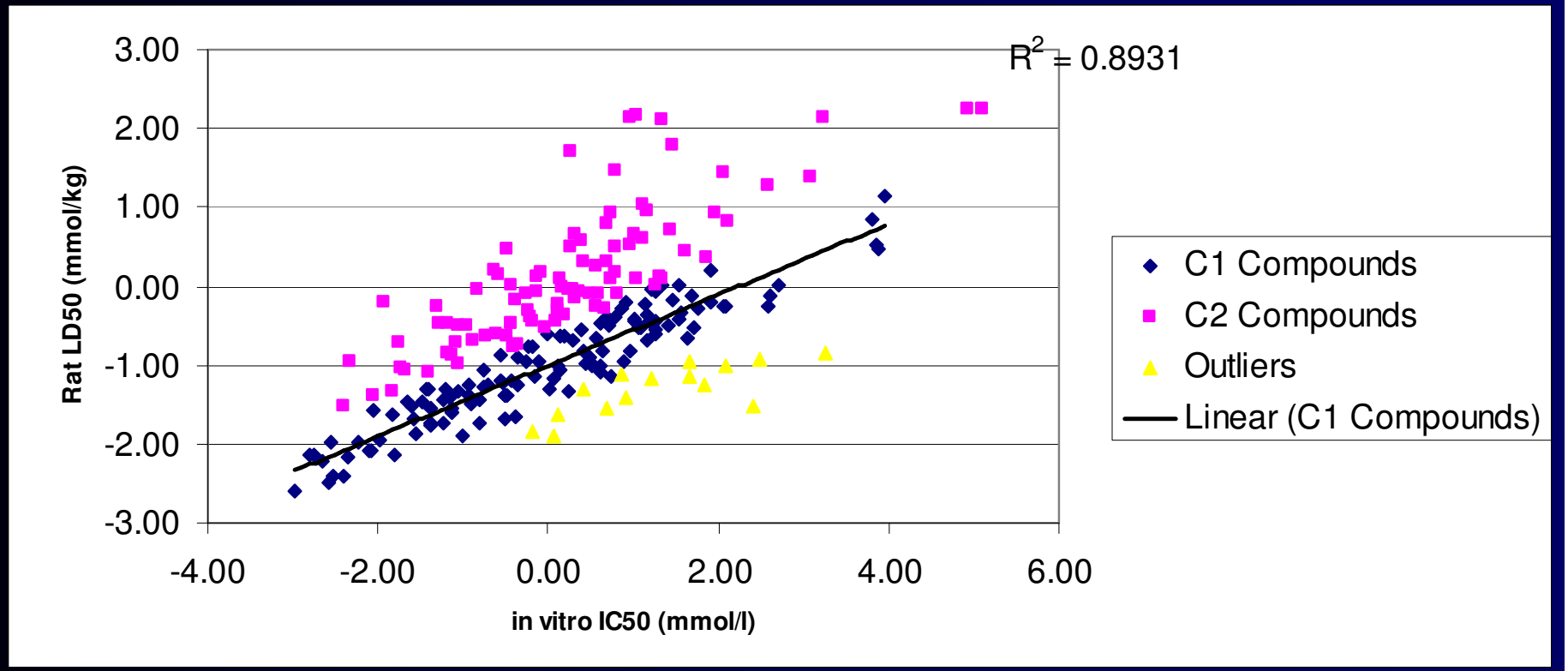
Poor in vitro-in vivo Correlation Between IC50 and Rat LD50 Values



R^2=0.46

A New Method to Use *in vitro* Toxicity Results to Assist the QSAR Modeling of *in vivo* Toxicity Endpoint

- IC50 vs. rat LD50 values



Moving Regression for Data Partitioning

$$\eta(x_i, y_i) = \begin{cases} 1, \text{if } y_i \in [ax_i + b - d_1, ax_i + b + d_2] \\ 0, \text{otherwise} \end{cases}$$

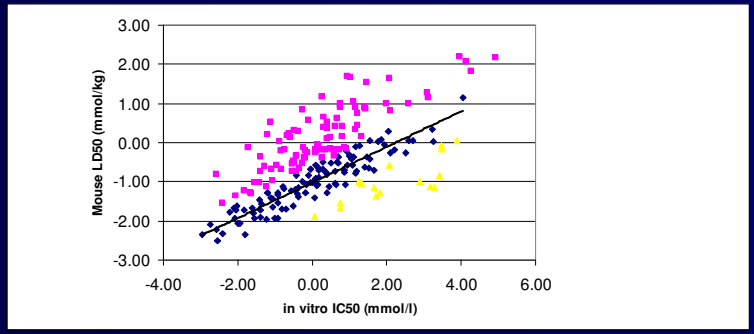
$$F(a,b) = \sum_{i=1}^{n} \eta(x_i, y_i)(y_i - ax_i - b)^2$$

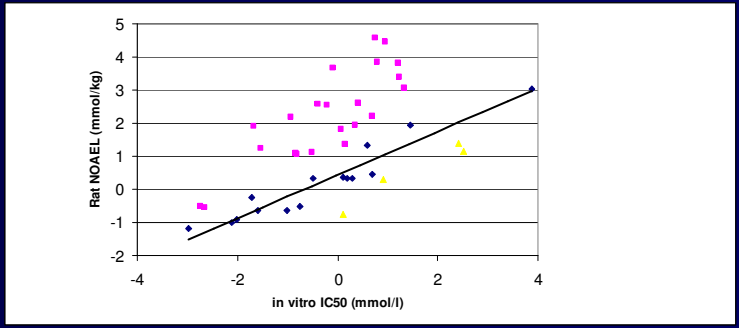
$$\eta(x_i, y_i) \sim \frac{1}{2}\left\{\frac{1}{1+\exp[-P_1(y_i - ax_i - b + d_1)]} + \frac{1}{1+\exp[P_2(y_i - ax_i - b - d_2)]}\right\}$$

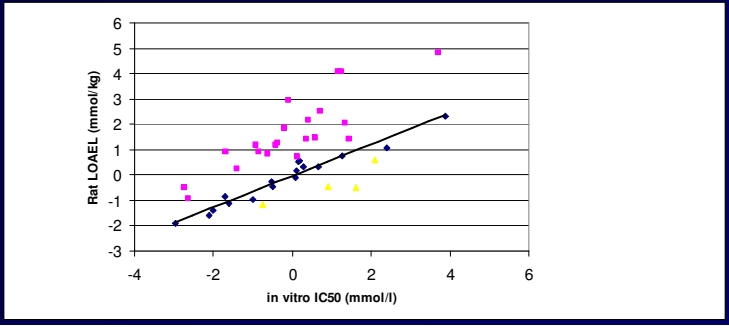
$$F(a,b) = \sum_{i=1}^{n} \frac{1}{2}\left\{\frac{1}{1+\exp[-P_1(y_i - ax_i - b + d_1)]} + \frac{1}{1+\exp[P_2(y_i - ax_i - b - d_2)]}\right\}(y_i - ax_i - b)^2$$

Cytotoxicity IC50 Values vs. Other in vivo Toxicity Results

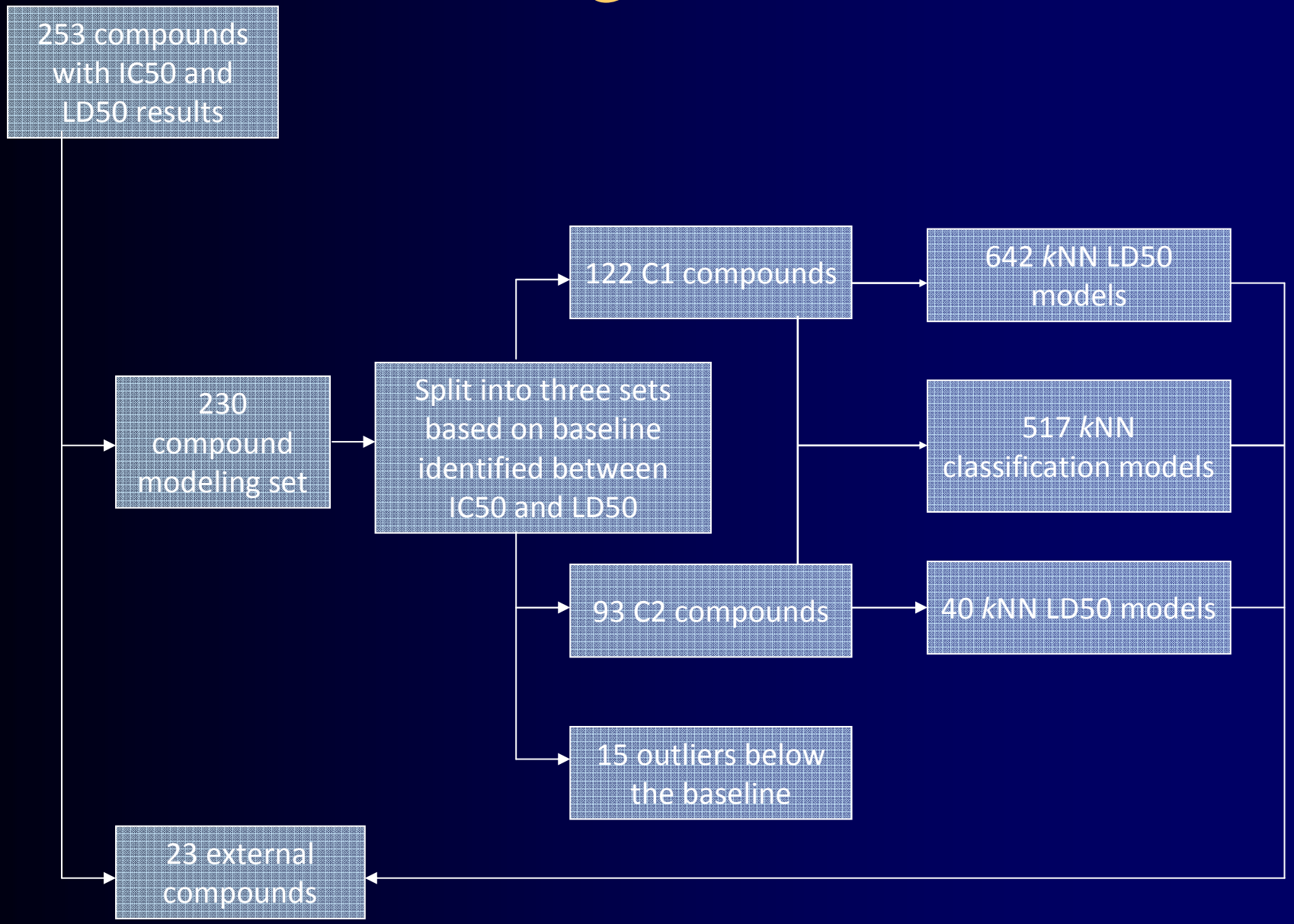
- IC50 vs. mouse LD50 values
- IC50 vs. rat NOAEL values
- IC50 vs. rat LOAEL values



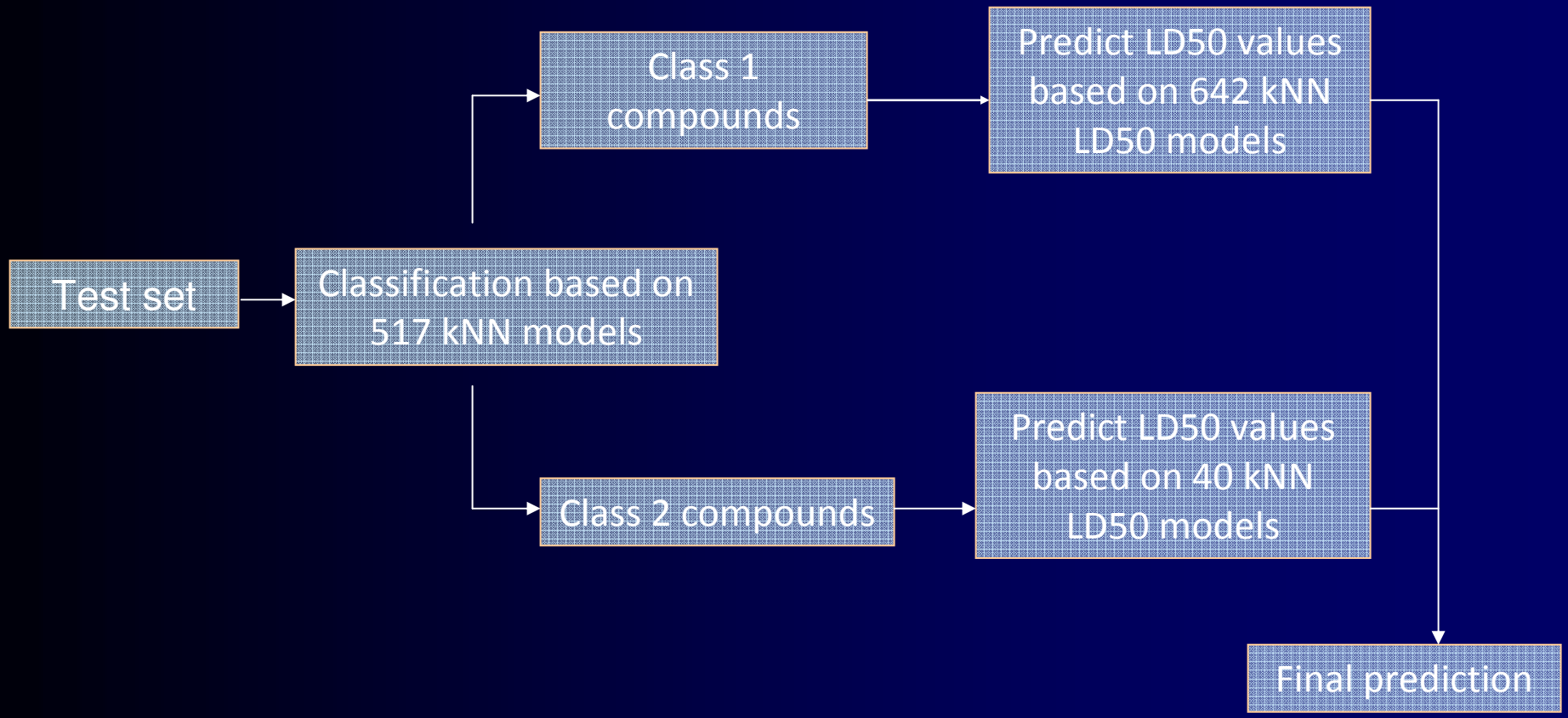




Modeling Workflow



Prediction Workflow



Classification of the Rat LD50 Values for the External Set of 23 Compounds

No AD:
Classification rate = 62%

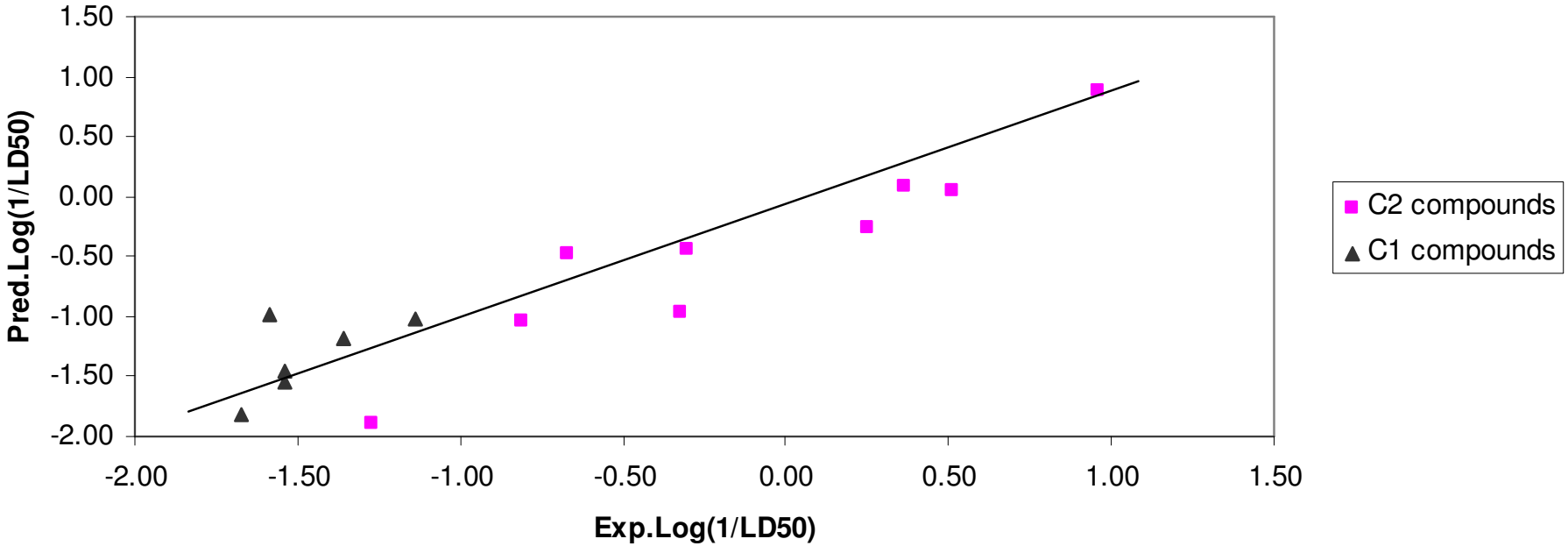
	Pred. C1	Pred. C2
Exp. C1	7	2
Exp. C2	6	5

With AD:
Classification rate = 78%

	Pred. C1	Pred. C2
Exp. C1	6	0
Exp. C2	4	5

Prediction of the Rat LD50 Values of the External 23 Compounds

- R^2=0.79, *MAE*=0.37, Coverage=74% (17 out of 23)



Prediction of New ZEBET Compounds

- Additional 115 ZEBET compounds with rat LD50 testing results obtained from Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM).

- R^2, *MAE* and prediction coverage of 0.60, 0.46, and 62%

Comparison Between Our Model and Toxicity Prediction by Komputer Assisted Technology (TOPKAT) LD50 Predictor

- 27 out of the 115 new ZEBET compounds that do not exist in the TOPKAT LD50 training set (version 6.1).
- Prediction of 27 new ZEBET compounds

	Our model		TOPKAT	
	No AD	With AD	No AD	With AD
R^2	0.69	0.73	0.16	0.50
MAE	0.42	0.34	0.78	0.46
Coverage	100%	70%	100%	70%

RTECS Rat Oral LD50 Dataset Overview

- 7,385 unique compounds total after removing inorganic compounds and mixtures.
  - **provided by Dr. Todd Martin (EPA)**

- Split the whole dataset into two parts: 3,472 compound modeling set and 3,913 compound validation set. All the compounds in the validation set were not included in the TOPKAT LD50 Predictor training set (version 6.1).

QSAR Methods

- UNC: Random Forrest (RF) and kNN
- EPA: Hierarchical Modeling, Nearest Neighbor (NN), FDA QSAR
- Descriptors: Dragon descriptors, fragmental descriptors
- Various types of Applicability Domain (AD)

7 Individual QSAR Models

- UNC: RF (two models) and kNN*
- EPA: Hierarchical with fragment constraint, Hierarchical no fragment constraint, NN, FDA QSAR

*UNC group used two modeling set: the original 3,472 compound modeling set and a reduced modeling set (2,475) after removing 997 structural outliers. RF model were developed for both sets and kNN models were only developed for reduced set.

External Validation Results for 7 Individual Models

	RF_full set	RF_red set	kNN_red set	Hierarchical with fragment constraint	Hierarchical no fragment constraint	Nearest neighbor	FDA
R^2	0.57	0.7	0.66	0.36	0.27	0.24	0.29
MAE	0.46	0.41	0.44	0.58	0.60	0.61	0.60
Coverage	50%	20%	20%	66%	93%	97%	95%

The Comparison Between Combi-QSAR and TOPKAT Results

	Consensus (at least 1 model)	TOPKAT	Consensus (70% of models.)	TOPKAT	Consensus (All models)	TOPKAT
R^2	0.41	0.19	0.62	0.39	0.76	0.60
MAE	0.54	0.77	0.44	0.60	0.38	0.50
Coverage	100%	100%	41%	41%	16%	16%

Experimental Study IV: Concordance between animal and human DILI data *

*Zhu et al, in preparation

Introduction

Hepatotoxicity is a major safety concern for drug development, as being a leading cause of candidate attrition.

Sources : M. Fung et al. Drug Information Journal, 2001.
MDS Pharma Services Issue, 2008, 7, 1-13.

Recently, the Safety Intelligence Program (**SIP**) group members performed a data analysis in order "*to assess the degree of concordance across species for drug-induced liver effects*", and thus, to complete the "*Non-Clinical Guideline On Drug-Induced Hepatoxicity*" published by the European Medicines Agency (EMEA).

One of the SIP goals is to contribute to the challenging quest for accurate tools to predict the drug-induced liver injury (DILI) potential associated with drug candidates approaching clinical use.

biowisdom

Intelligence in healthcare

Prepublication Memorandum

From the Safety Intelligence Program Board

Authors:

Steven S
David Co
Julie Ban
Jack Rey

Contact I

sip™

biowisdom

27th June

© BioWisc

Introduction

The Safety Intelligence Program (SIP) Board welcomes the opportunity to comment on the CHMP Draft Non-clinical Guidelines on Drug-induced Hepatotoxicity. SIP is an industry led initiative that harnesses the expertise of its pharmaceutical members, BioWisdom and other key stakeholders to build a comprehensive and high quality intelligence resource for use in the practice of drug safety assessment. SIP strives to ensure that the benefit/risk decisions made for every compound in the development pipeline or drug on the market is based on having visibility to the best information possible.

The 2008 priority for SIP is to focus on hepatotoxicity, in recognition of the challenge in being able to predict, monitor and manage the hepatotoxicity risk associated with new chemical entities approaching approved clinical use. SIP leverages the huge amount of publicly available information to generate an intelligence resource for the safety science communities working in drug development. This intelligence resource is created using BioWisdom's established technology platform (Sofia™) that enables the systematic generation of semantically consistent assertional meta-data. Assertional meta-data comprise relationships between distinct entities, for example, 'Acetaminophen INDUCES Hepatic Necrosis' or 'Bosentan INHIBITS ATP Binding Cassette, Subfamily B, Member 11'. With the capability to reference the original citation, the assertional meta-data can be analysed systematically to reveal new insights related to specific topics.

Here we present a prepublication report that highlights the power of being able to perform systematic and comprehensive analyses on assertional meta-data that captures the current status of knowledge pertaining to a particular area. As an example here, we use an analysis of the degree of concordance of compound-induced effects in the liver between preclinical species and human, referencing specifically, the following statement made in the draft guidelines (section 5, point iii):

"With respect to the animal species used in standard non-clinical studies, a general assumption is that the higher the species (rodent, non-rodent, non-human primate) that demonstrates signs of liver toxicity or histopathological adverse responses, the greater the relevance of clarifying mechanisms responsible for liver toxicity."

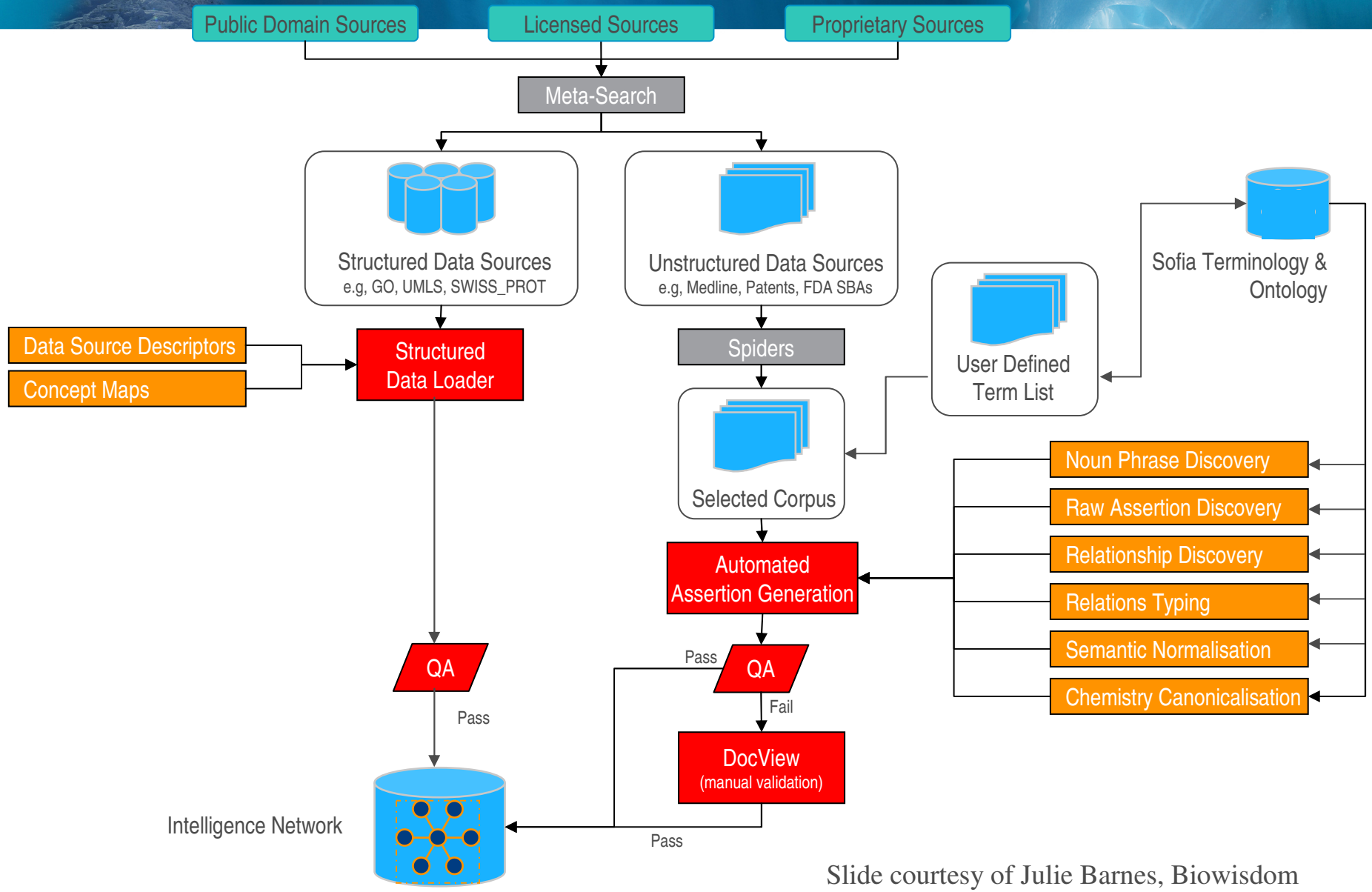
We submit this analysis to the EMEA because we believe it forms part of the necessary evaluation of historic knowledge that will advance our collective understanding of drug-induced hepatotoxicity and will ultimately lead to an improved capability to assess risk of new chemical entities for liver injury.

Brief Methodology

BioWisdom's Sofia platform was used to generate assertions that describe the effects of known chemicals in the liver. Vocabularies/thesauri describing >150,000 distinct chemical names and >6000 liver pathologies, physiological processes and clinical chemistry liver biomarkers were used to generate putative assertions, from publicly accessible information. Specifically, here we used Medline abstracts and European Public Assessment Reports (EPARs) published by the EMEA. The assertions were passed through a QC process to ensure they accurately reflected (to >97%) the statements made by the authors in the documents. Each assertion was supported by one or more pieces of evidence. Extracted assertions were "semantically normalised" to deal with the inconsistencies inherent in the way authors describe their observations. This process yields a

© BioWisdom Ltd, 2008 Non-Confidential 3/7

Intelligence Network Build Process



Slide courtesy of Julie Barnes, Biowisdom

Data transformation for the Venn diagram

- Species profile for each compound (951) was retrieved from the original data. This step was done automatically with a program written in Delphi.

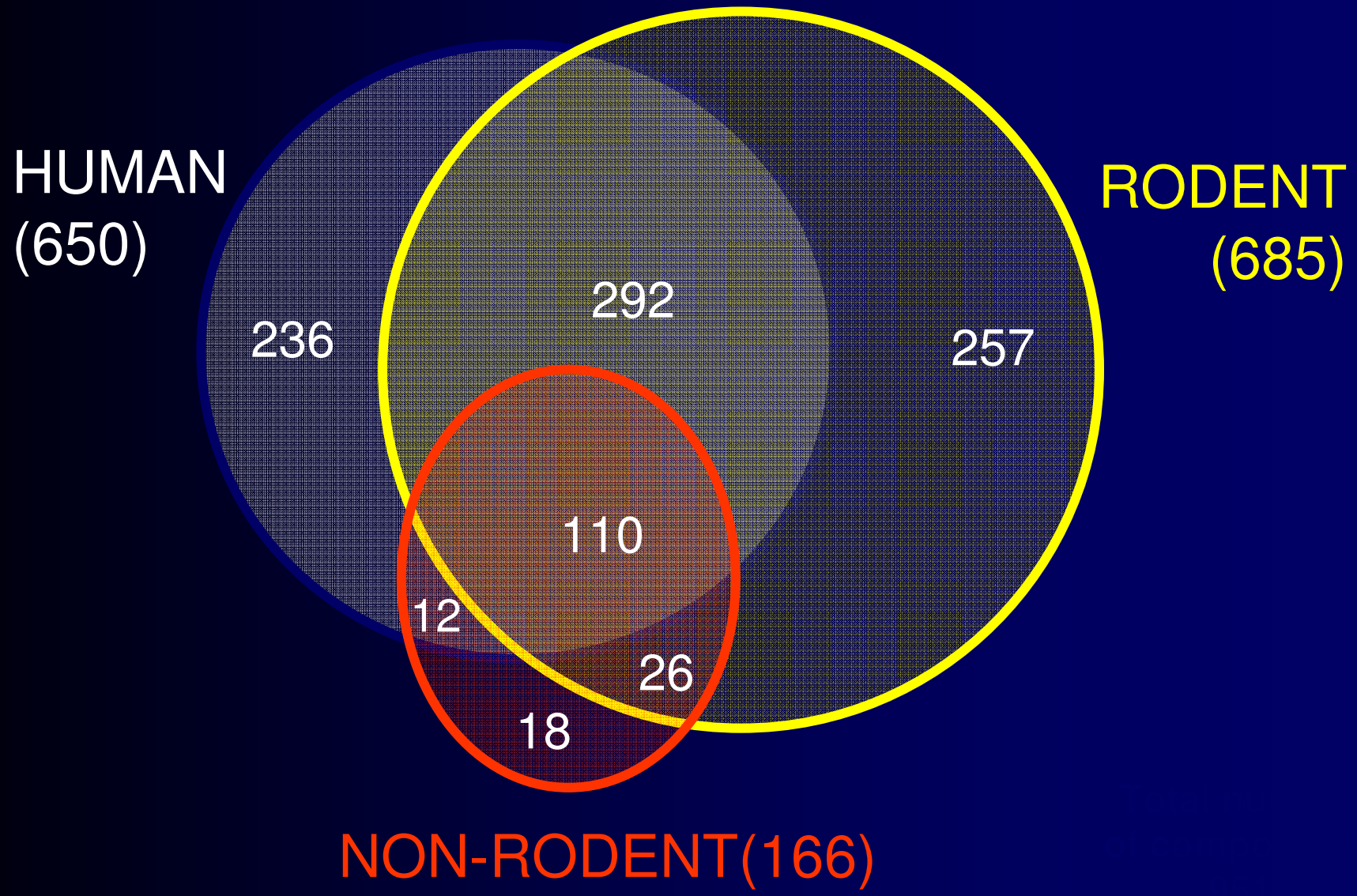
- Then, the table required to draw the Venn diagram was calculated:

		A	B	C		A	B	C	AB	AC	BC	ABC
ID	Name	HUMAN	RODENT	NON-RODENT		only	only	only				
1	(R)-Roscovitine	0	1	0		0	1	0	0	0	0	0
2	17-Methyltestosterone	0	1	0								0
3	1-alpha-Hydroxycholecalciferol	1	0	0								0
4	2,3-Dimercaptosuccinic acid	1	1	0								0
5	2,4,6-Trinitrotoluene	1	0	0								0
6	2-Deoxy-D-glucose	1	1	0								0
7	2'-fluoro-5-methylarabinosyluracil	1	0	0								0
8	2-Methoxyestradiol	1	1	0								0
9	4-aminobenzoic acid	0	1	0								0
10	4-Hydroxytamoxifen	1	1	0								0
11	5 fluorouracil	1	1	1								1
12	5-Azacitidine	1	1	0								0
13	5-Bromouracil	0	1	0								0
14	5-fluoro-2'-deoxyuridine	1	1	0								0
15	6-Mercaptopurine	1	1	0								0
16	Acadesine	0	1	0								0
17	Acarbose	1	1	0								0
18	Acebutolol	1	1	0								0
19	Acenocoumarol	1	0	0								0
20	Acetamide	0	1	0								0
21	Acetaminophen	1	1	1								1
22	Acetazolamide	1	1	1								1
23	Acetic acid	1	1	1								1
24	Acetohexamide	1	0	0								0
25	Acetohydroxamic acid	0	1	0		0	1	0	0	0	0	0

It should be emphasized that we assume that <u>each compound has been tested in all species, i.e., humans, rodents and non-rodents.</u>

"1" = toxic
"0" = non-toxic

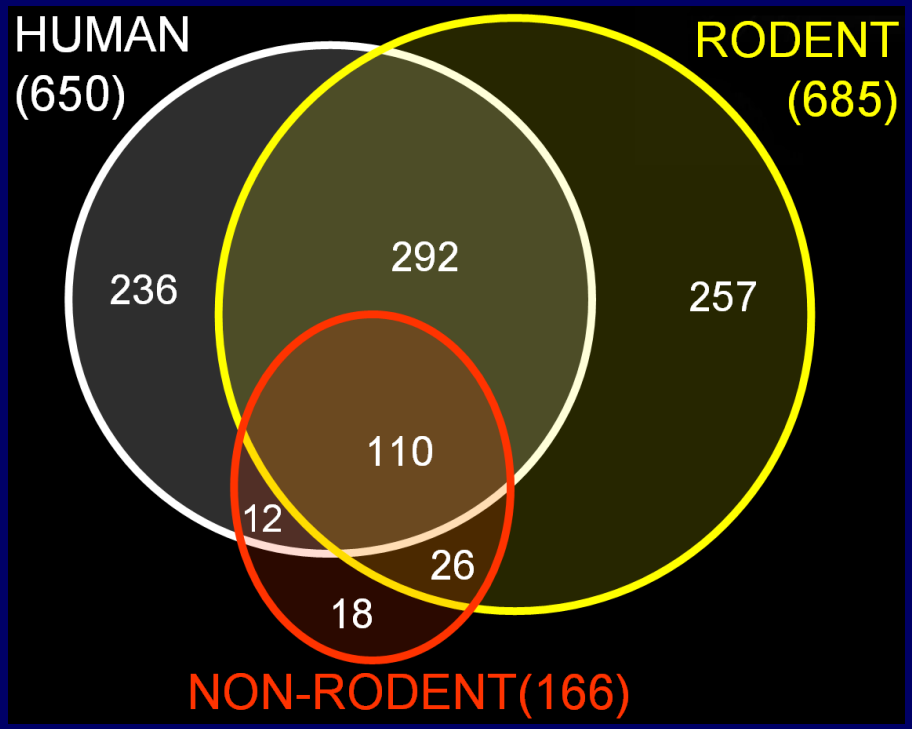
Results : the Venn Diagram of Curated Biowisdom data



Results : the Venn Diagram of Curated Biowisdom data

→ Concordance between humans and rodents ?

In our opinion it should reflect data on BOTH toxic and non-toxic compounds!



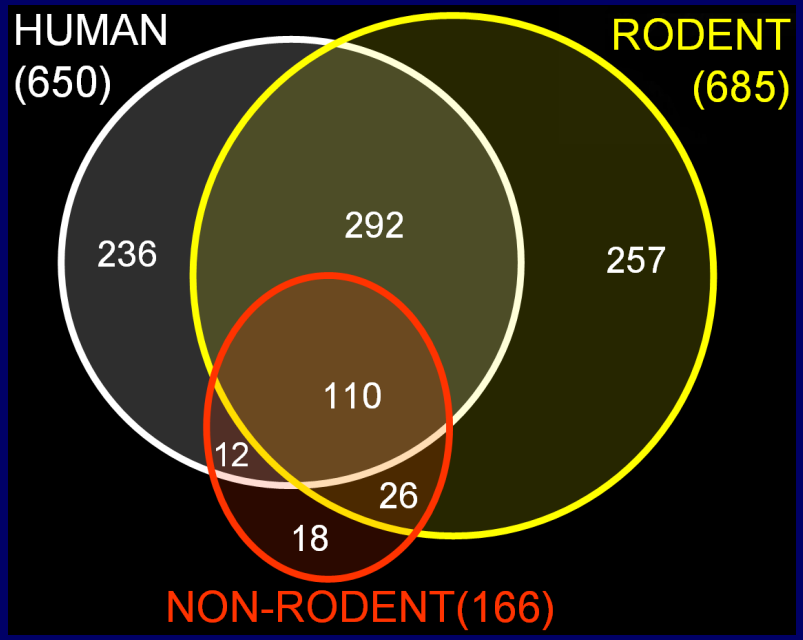
$$\text{Concordance} = \frac{\#\text{ toxicants BOTH for humans/rodents} \;+\; \#\text{ NON-toxicants BOTH for humans/rodents}}{\text{Total of tested chemicals}}$$

$$\text{Concordance} = \frac{(292 + 110) \;+\; 18}{951} \simeq \mathbf{44.2\ \%}$$

(Using Biowisdom initial data – 1061 compounds, we found concordance ≃ 42.4 %)

Conclusions about concordance across species

	H	R	NR
H		44.2%	39.9%
R	44.2%		39.1%
NR	39.9%	39.1%	



- Calculated concordance values between species are very close.

- Surprisingly, there is no large gap between concordance values between concordances H/R and H/NR is less than 5%) as one could suppose.

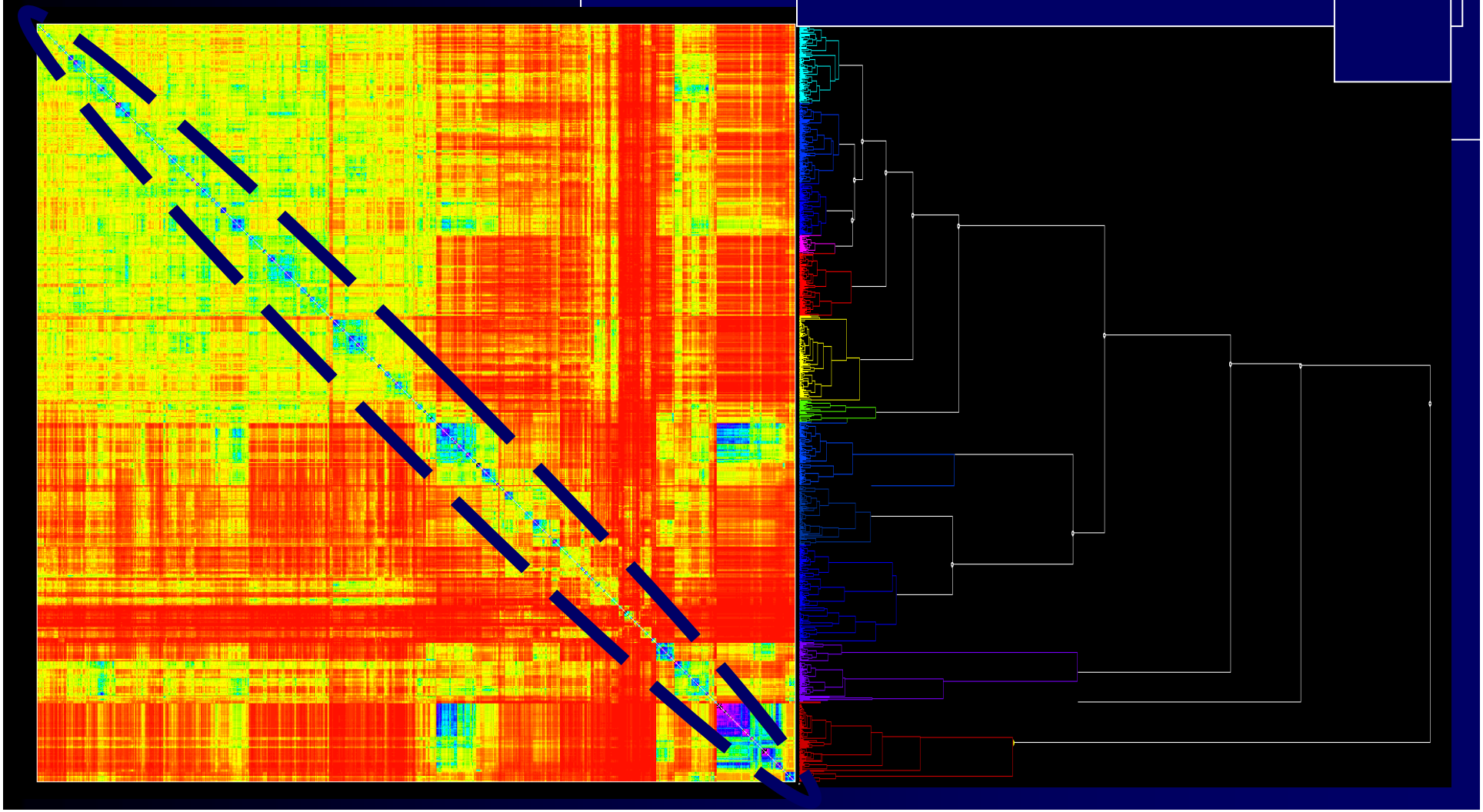
- These results are valid if and ONLY IF the following assumption is correct: each compound has been effectively tested for each category H, R and NR, and in each case, found either toxic or non-toxic.

For example: we assume that 18 compounds that have been found to be only toxic for non-rodents have been tested on both humans and rodents and found to be non-toxic.

1. Clustering of 951 compounds in chemical space

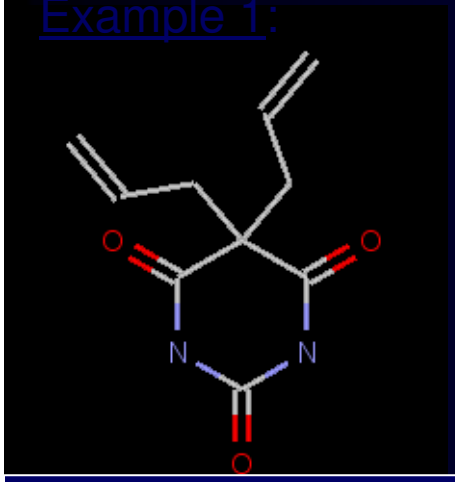
The cluster analysis has been performed using fragment descriptors, hierarchical algorithm, Euclidean metrics between compounds, and a complete linkage between clusters.

Small clusters have then been identified with pretty high levels of similarity between compounds.

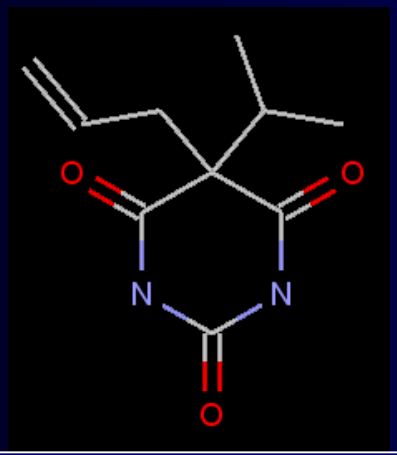


1. Clustering of compounds in chemical space

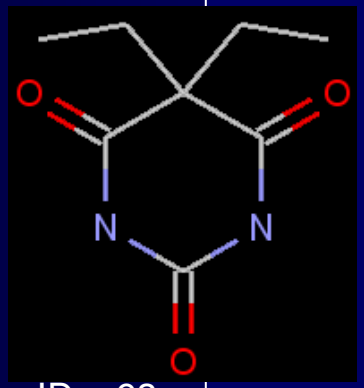
Example 1:



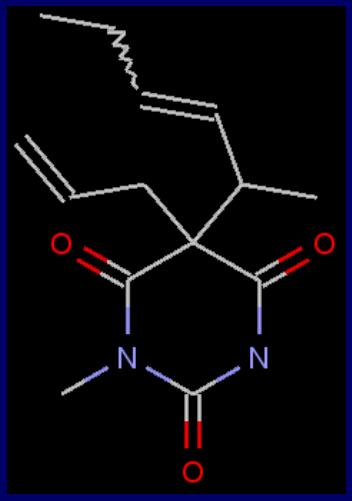
ID = 45
HUMAN = **0**
RODENT = **1**
NON-RODENT = **0**



ID = 76
HUMAN = **0**
RODENT = **1**
NON-RODENT = **0**

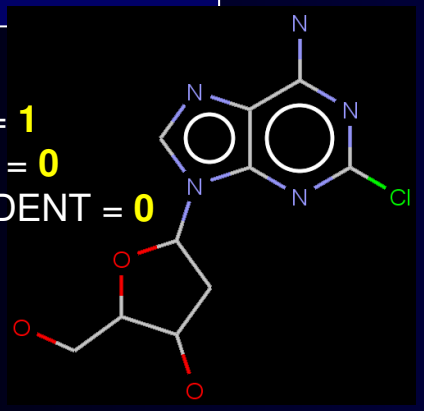


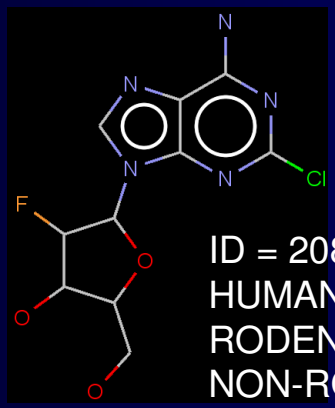
ID = 93
HUMAN = **0**
RODENT = **1**
NON-RODENT = **0**



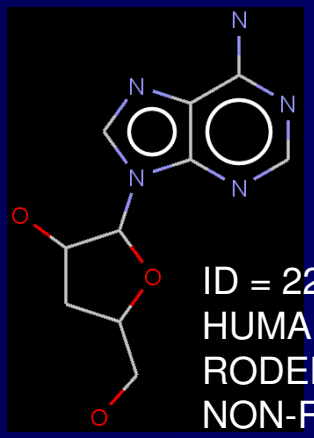
HUMAN = **0**
RODENT = **1**
NON-RODENT = **0**

ID = 201
HUMAN = **1**
RODENT = **0**
NON-RODENT = **0**



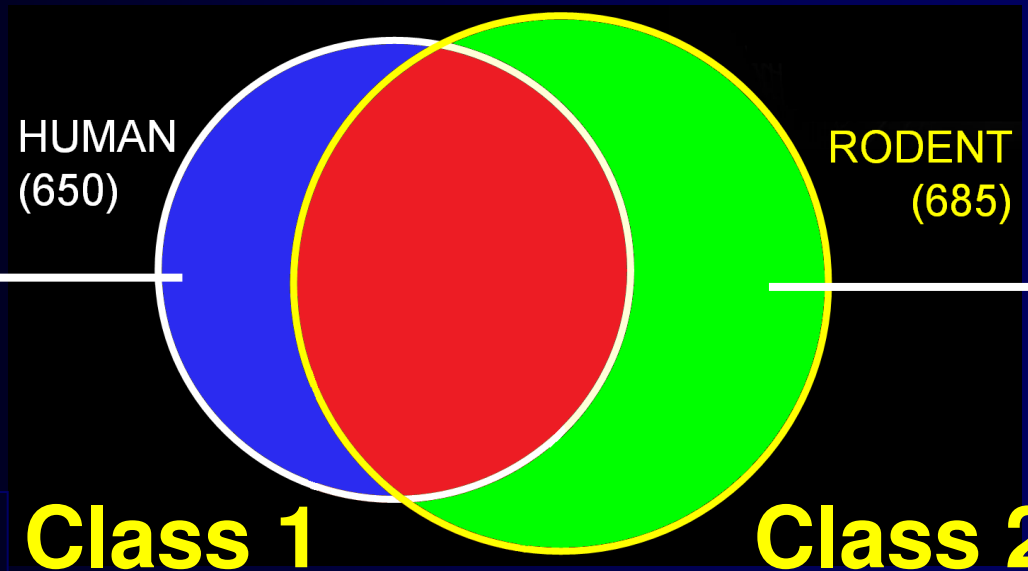


ID = 208
HUMAN = **1**
RODENT = **0**
NON-RODENT = **0**



ID = 223
HUMAN = **0 (???)**
RODENT = **1 (???)**
NON-RODENT = **0**

3. QSAR based classification



HUMAN ONLY

Class 1
(248)

Class 2
(283)

RODENT ONLY

Compounds found to be <u>toxic for humans</u> only

Compounds found to be <u>non-toxic for humans</u>

Could we predict the class of a compound from its structure only ?

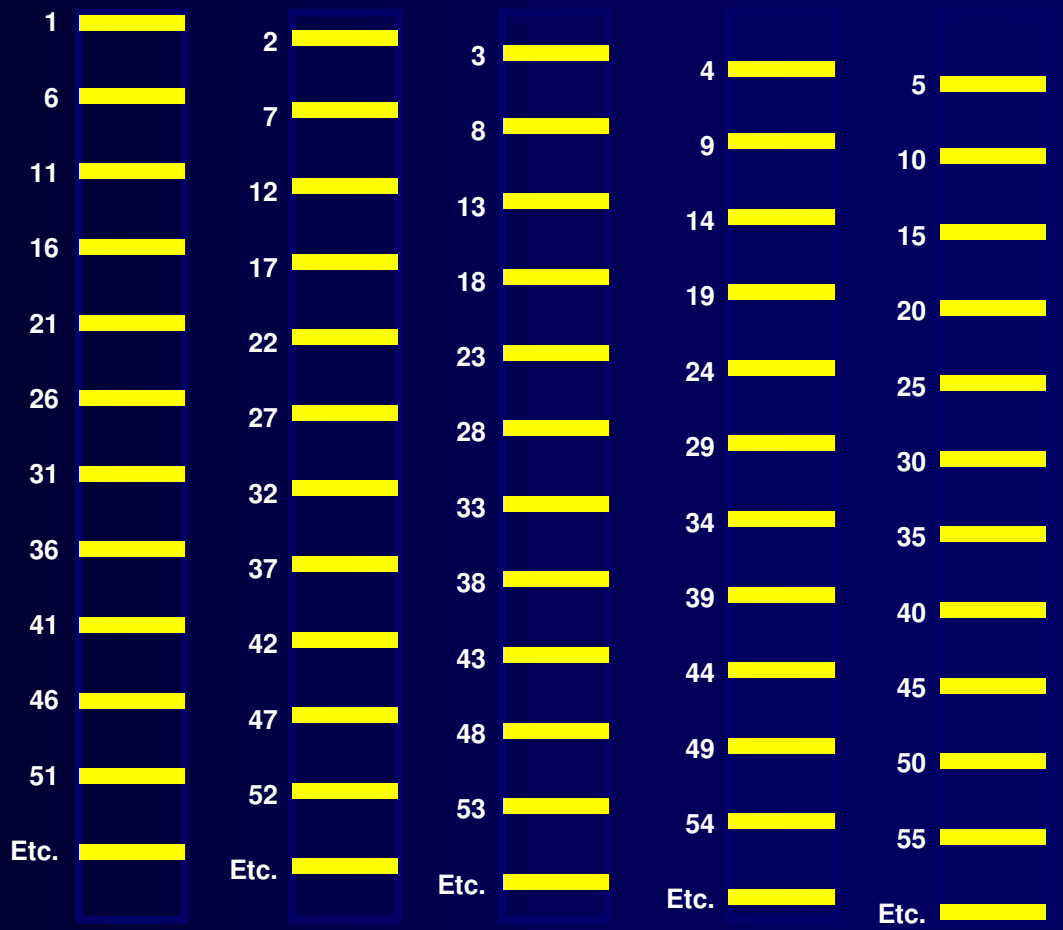
QSAR based classification

HUMANS

RODENS

External 5 fold

Cross validation



20% of compounds → EXTERNAL SET

80% of compounds → MODELING SET

Models are built using the modeling set ONLY.

QSAR based classification

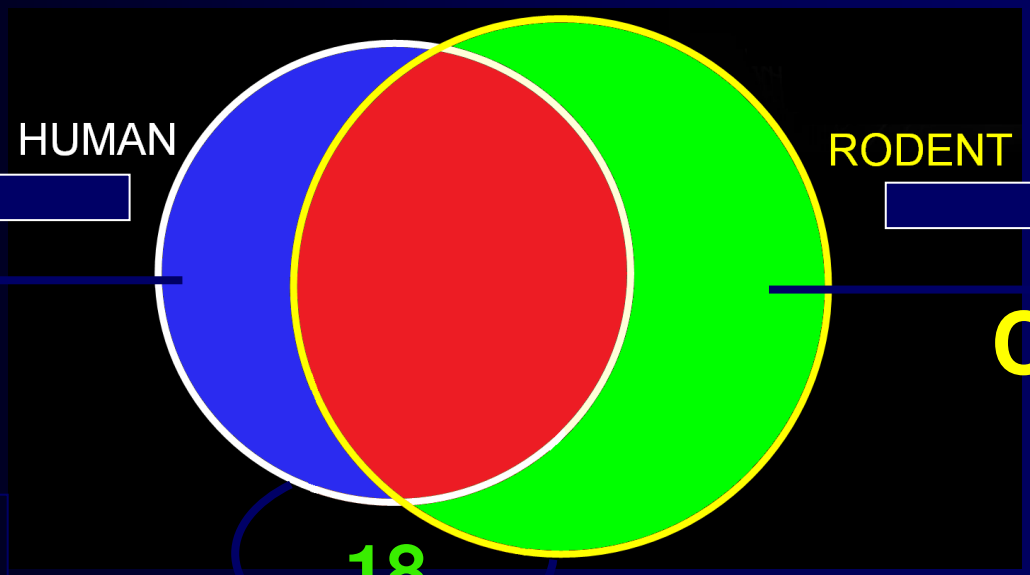
Using SUPPORT VECTOR MACHINES (SVM)

Accuracy (%) = (number of compounds correctly predicted)/(total number of compounds)

Fold	Modeling set 5 fold CV	Modeling set Accuracy	External set Accuracy	*Model ID*	
1	62.3%	88.2%	71.0%	*217*	
	62.9%	77.6%	67.3%	*162*	*Dragon*
2	64.9%	81.2%	64.2%	*112*	
	67.5%	81.2%	55.7%	*197*	*Dragon*
3	62.4%	91.3%	64.2%	*194*	
	65.2%	91.1%	61.3%	*198*	*Dragon*
4	64.9%	99.3%	72.6%	*208*	
	62.1%	84.9%	68.9%	*151*	*Dragon*
5	63.3%	82.6%	68.9%	*205*	
	61.9%	94.4%	70.8%	*175*	*Dragon*

NB: The results are preliminary, could be improved.

3. QSAR based classification



Class 1
(248)

HUMAN
ONLY

Class 2
(283)

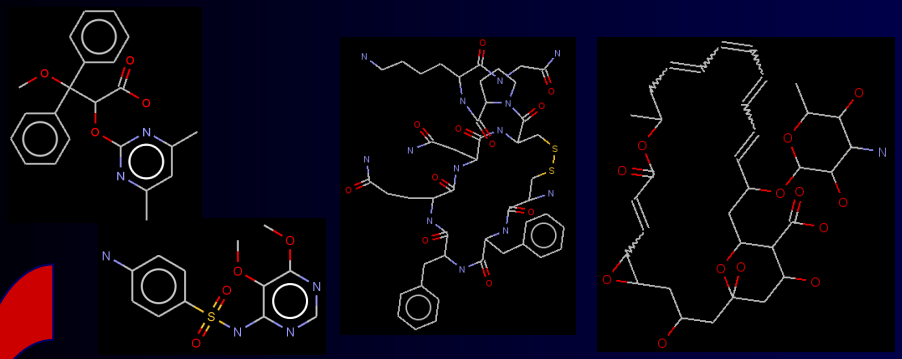
RODENT
ONLY

EXTERNAL SET
(18 compounds being non-toxic
in humans)

QSAR MODELS

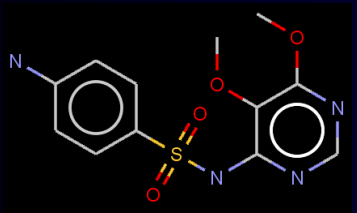
3. QSAR based classification

Compounds	Modeling set 5 fold CV	Modeling set Accuracy	External set Accuracy	*Model ID*	*Descriptors*
18	62.9%	92.5%	77.8%	*206*	*Fragments*
	64.0%	97.9%	66.7%	*141*	*Dragon*



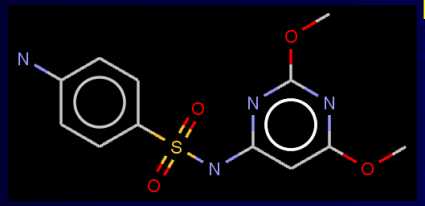
14 of 18 compounds are predicted as non-toxic for humans.

Hepatotoxicity induced by 4 compounds is not well predicted. BUT:



Sulfadoxine (ID=820)
Human = 0
Rodent = 0
Non-Rodent = 1

Missing/incorrect data ???



IN THE TRAINING SET:

Sulfadimethoxine (ID=1)
Human = 1
Rodent = 0
Non-Rodent = 1

Conclusions of Study IV

- We focused on the concordance analysis across species for hepatotoxicity induced by drugs. Results showed close concordance values (~40%) for Human/Rodents and Human/Non-Rodents **However, this conclusion is valid if and only if we assume that each compound of the set has been effectively tested for each category (human, rodent and non-rodent), and in each case, found either toxic or non-toxic.**

- Cluster analysis allowed us to identify multiple clusters, in which compounds belong to congeneric series. Similar toxicity profiles are observed for certain clusters **Similarity in chemical space could help to double check the toxicity reported in the literature for different species.**

- QSAR models have been generated to predict the toxicity of compounds for humans. Despite the apparent diversity of data, models show fairly good prediction power assessed by five-fold cross validation procedures, and confirmed by the application of models to an external set of 18 compounds (under the same assumption of the completeness of toxicity testing across all compounds and species).

Final Thoughts

Nothing that is worth knowing can be taught.

Oscar Wilde

- Focus on accurate prediction of external datasets is much more critical than accurate fitting of existing data
  - validation!!!
  - applicability domain
  - consensus prediction using all acceptable models
  - Ideally, experimental validation of a small number of computational hits
- Predictive QSAR workflow with extensive validation affords statistically significant models
  - reliable property predictors
  - decision support tools in selecting experimental screening sets
- HTS and –omics data may be insufficient to achieve the desired accuracy of the end point property prediction BUT should be explored as biodescriptors in combination with chemical descriptors

ACKNOWLEDGMENTS

UNC ASSOCIATES

Former:

-Stephen CAMMER
-Sung Jin CHO
-Weifan ZHENG
- Min SHEN
-Bala KRISHNAMOORTHY
-Shuxing ZHANG
–Peter ITSKOWITZ
–Scott OLOFF
-Shuquan ZONG
-Raed KHASHAN
– Jun FENG
– Yun-De XIAO
–Yuanyuan QIAO
–Ruchir SHAH
–Patricia LIMA
–Assia KOVACHEVA
–Julia GRACE
–Hao HU

Current

•Structural bioinformatics group:

– Yetian CHEN
– Tanarat KIETSAKORN
–Theo Walker
– Berk ZAFER
– Denis FOURCHES
– Georgii ABRAMOCHKIN

Cheminformatics group:

- Kun WANG
– Sasha GOLBRAIKH
– Simon WANG
– Chris GRULKE
- Hao TANG
- Hao ZHU
- Tong-Ying Wu
- Achintya SAHA
- Rima HAJJO
M. KARTHIKEYAN
- Lin YE
- Lying ZHANG
- Mihir SHAH
- Jui-Hua HSIEH
- Aleks SEDYKH

- Collaborators
  - Ann Richard (EPA)
  - Todd Martin (EPA)
  - Ivan Rusyn (UNC)
  - Fred Wright (UNC)
  - Julie Barnes (Biowisdom)

- Funding
  - NIH
    - P20-HG003898 (RoadMap)
    - R21GM076059 (RoadMap)
    - R01-GM66940
    - GM068665
  - EPA (RD 83382501 and R832720)