UNC
ESHELMAN
SCHOOL OF PHARMACY

MML
UNC.EDU

(Challenges of) Computer-Aided Nanotoxicology

Alexander Tropsha and Denis Fourches
University of North Carolina at Chapel Hill, USA

Outline

MML UNC.EDU

1. Availability of data on structure and biological activity of nanomaterials

2. Application of the Quantitative Structure Activity Relationships (QSAR) approach to nanostructure-activity relationship (QNTR) modeling

3. Preliminary studies

CS1- QNTR of whole NPs: Modeling of cellular effects induced by diverse NP [1]

CS2- QNTR of surface modifiers: Modeling of the NPs uptake in PaCa2 cancer cells [2]

4. Future Studies

[1] Shaw et al. PNAS, 2008, 105, 7387-7392

[2] Weissleder et al. *Nat. Biotechnol.*, 2005, 23 (11), 1418-1423

Introduction

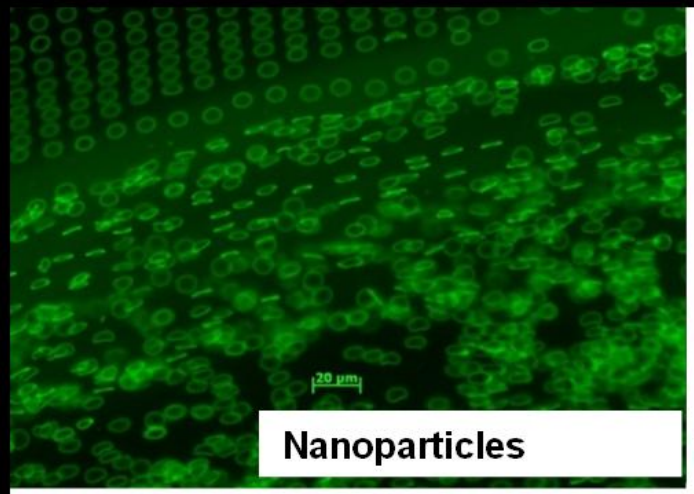
- There are ~**600** manufacturer-identified nanotechnology-based consumer products currently on the market (Woodrow Wilson International Center for Scholars, 2008).

- Acute or repeated exposure to MNPs present in commercial products may cause systemic, cellular, and/or genomic toxicities. Methods should be developed that are “capable of studying the relationship between deposited particles…to determine which aspects…are best predictors of adverse health effects”*

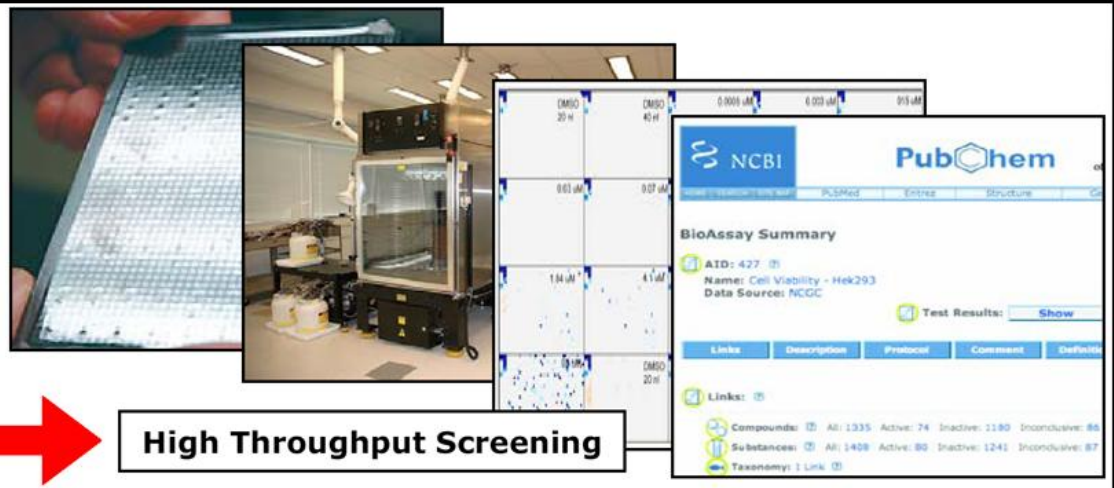
Development of « *in silico* » approaches

* Luther, W. editor, Industrial Application of Nanomaterials - Chances and Risks, Technological Analysis http://www.zukuenftigetechnologien.de/11.pdf

QNTR modeling workflow



Nanoparticles

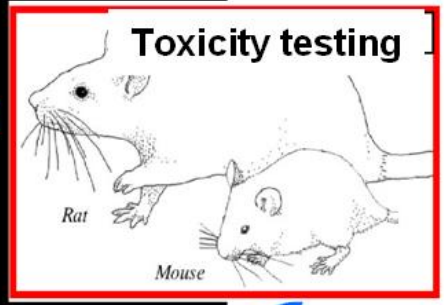


High Throughput Screening

Molecular weight, compositions and geometrical parameters, physico-chemical properties (acidic, basic, neutral, amphi- or lipophilic etc.)

Molecular properties

Human health risk

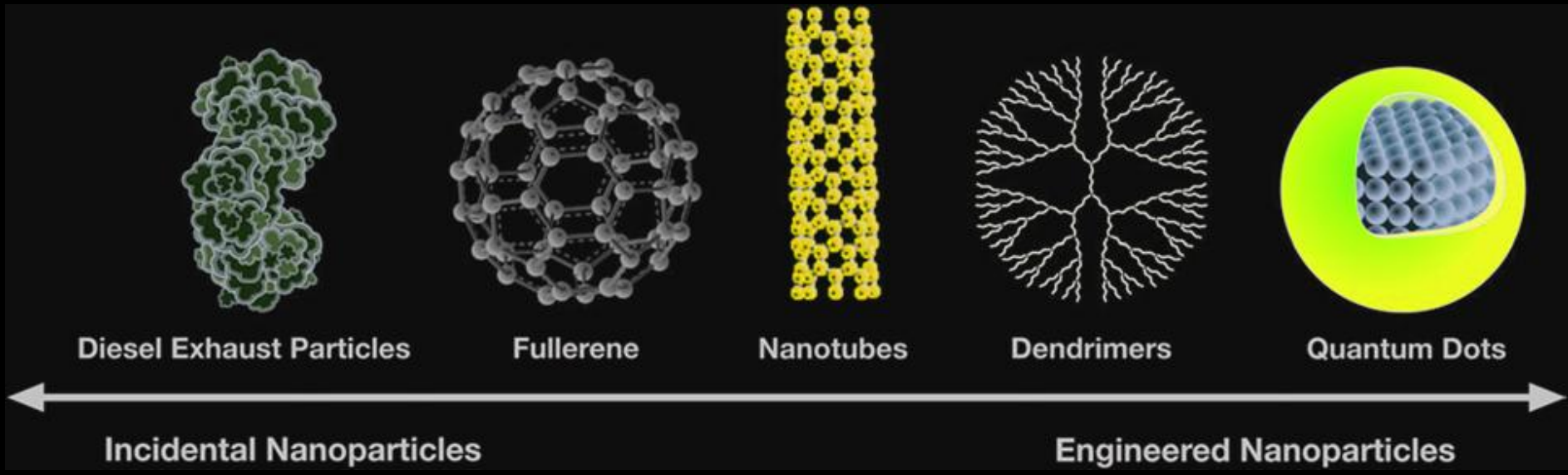


Biological Descriptors

Chemical Descriptors

QNTR modeling

Difficulties in Modeling of Nanoparticles

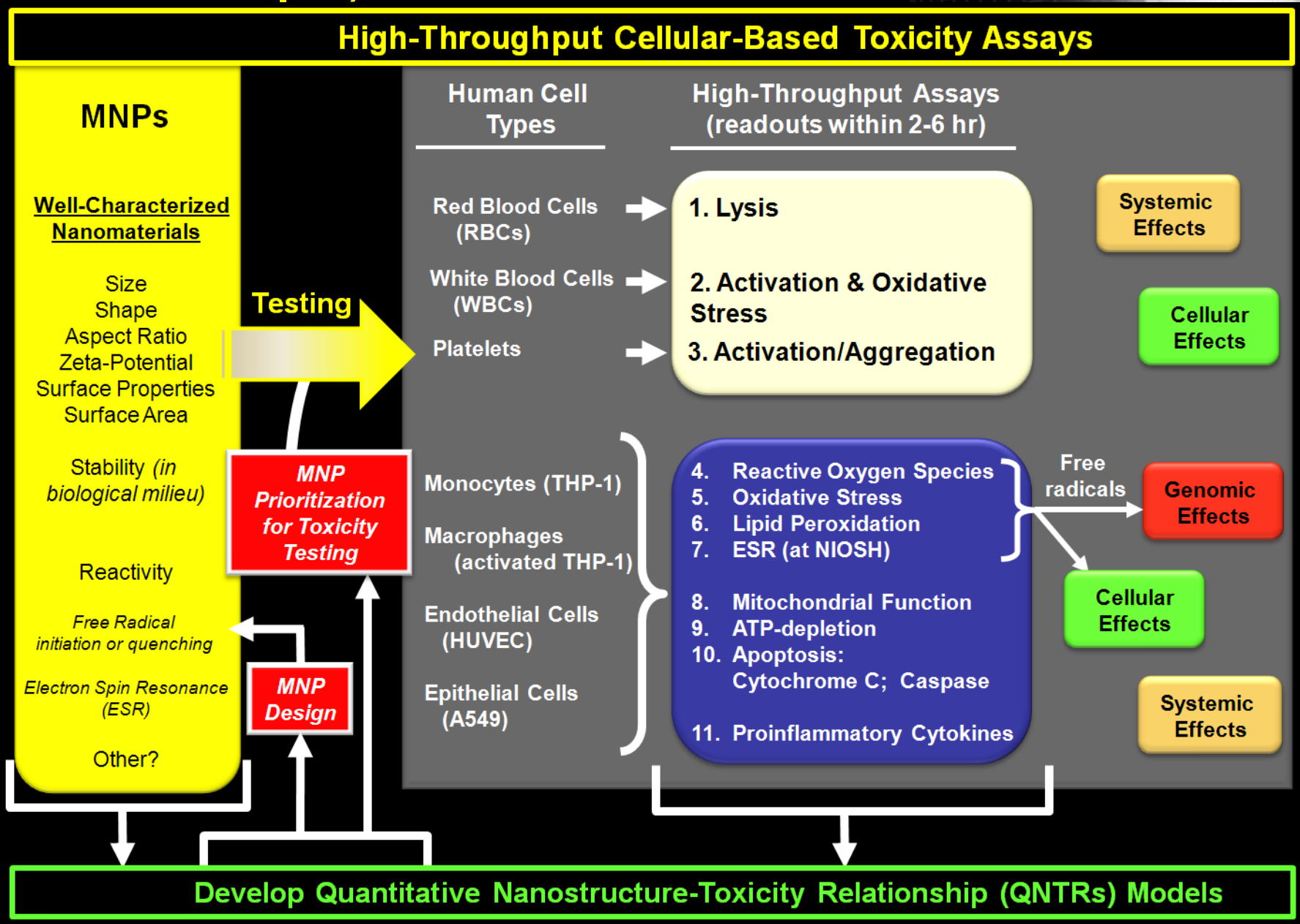


S. Stern and S. McNeil, Toxicological Sciences, 101(1), 4-21, 2008.

- NP structures are very diverse → a real challenge to develop quantitative parameters (descriptors) of MNPs.
- Systematic physico-chemical, geometrical, structural and biological studies of NPs are nearly absent.
- Computational modeling of nanoparticles is only beginning to emerge; best if done in collaboration with experimental scientists.

Emerging collaborative project with Dr. Russ Mumper, UNC

MML UNC.EDU



Research hypotheses

1. The effects of NPs on different types of human cells depend on the physical/chemical/geometrical properties of the NPs.

CHARACTERIZATION OF PARTICLES

Composition, size, shape, aspect ratio, surface area, chemistry/morphology, zeta potential, chemical reactivity, structural descriptors.

2. High-throughput cellular-based assays with endpoints within 2-6 hr provide useful and predictive information about long-term biological properties of NPs.

Nano-bio interactions with human cells occur relatively rapidly, but the effects of these interactions are manifested over much longer time periods.

Research hypotheses

1. The effects of NPs on different types of human cells depend on the physical/chemical/geometrical properties of the NPs.

2. High-throughput cellular-based assays with endpoints within 2-6 hr provide useful and predictive information about long-term biological properties of NPs.

3. Toxicological data obtained from *in-vitro* cellular-based toxicity assays will correlate reasonably with *in-vivo* findings.

4. Using physical/chemical characterization and toxicological screens for an ensemble of MNPs, it will be possible to develop **predictive Quantitative Nanostructure – Toxicity (QNTR) models.**

- **Fundamental, comprehensive and predictive knowledge** of the nanotoxicology of MNPs;
- Improvements of experimental design and prioritized toxicity testing, to obtain safer and more efficient nanoparticles.

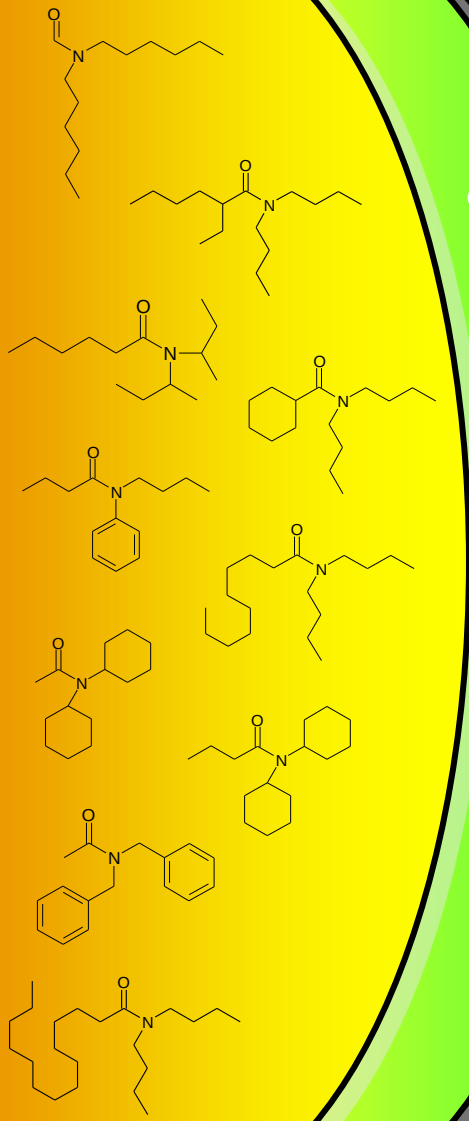
Outline

1. Availability of data on structure and biological activity of nanomaterials

2. Application of the Quantitative Structure Activity Relationships (QSAR) approach to nanostructure-activity relationship (QNTR) modeling

Principles of QSAR modeling

COMPOUNDS



DESCRIPTORS

Thousands of molecular descriptors are available for organic compounds

constitutional, topological, structural, quantum mechanics based, fragmental, steric, pharmacophoric, geometrical, thermodynamical conformational, etc.

Quantitative Structure Activity Relationships

- **Building of models** using machine learning methods (NN, SVM etc.);

- **Validation of models** according to numerous statistical procedures, and their **applicability domains**.

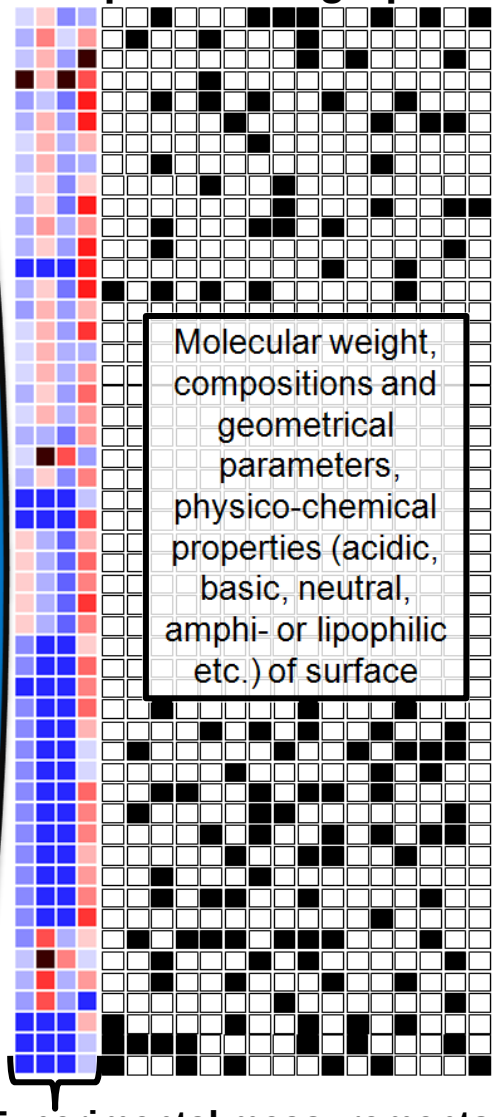
0.613
0.380
-0.222
0.708
1.146
0.491
0.301
0.141
0.956
0.256
0.799
1.195
1.005

ACTIVITY

Introducing QNTR modeling

NANOMATERIALS

Nanoparticle fingerprints



Molecular weight, compositions and geometrical parameters, physico-chemical properties (acidic, basic, neutral, amphi- or lipophilic etc.) of surface

DESCRIPTORS

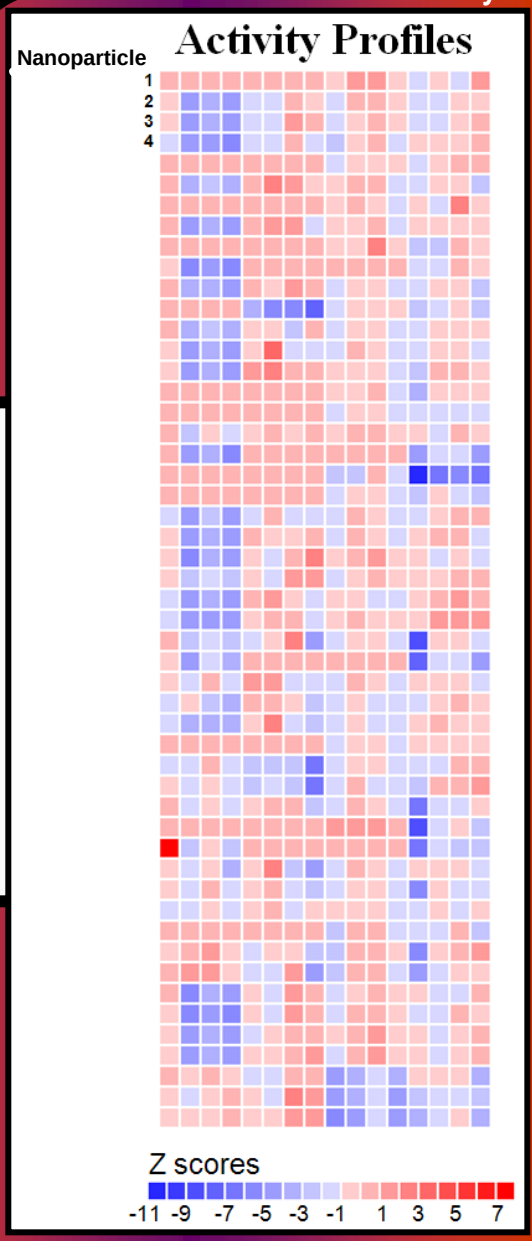
Experimental measurements
(size, relaxivities, zeta potential etc.)

Quantitative Nanostructure Toxicity Relationships

- **Building of models** using machine learning methods (NN, SVM etc.);

- **Validation of models** according to numerous statistical procedures, and their applicability domains.

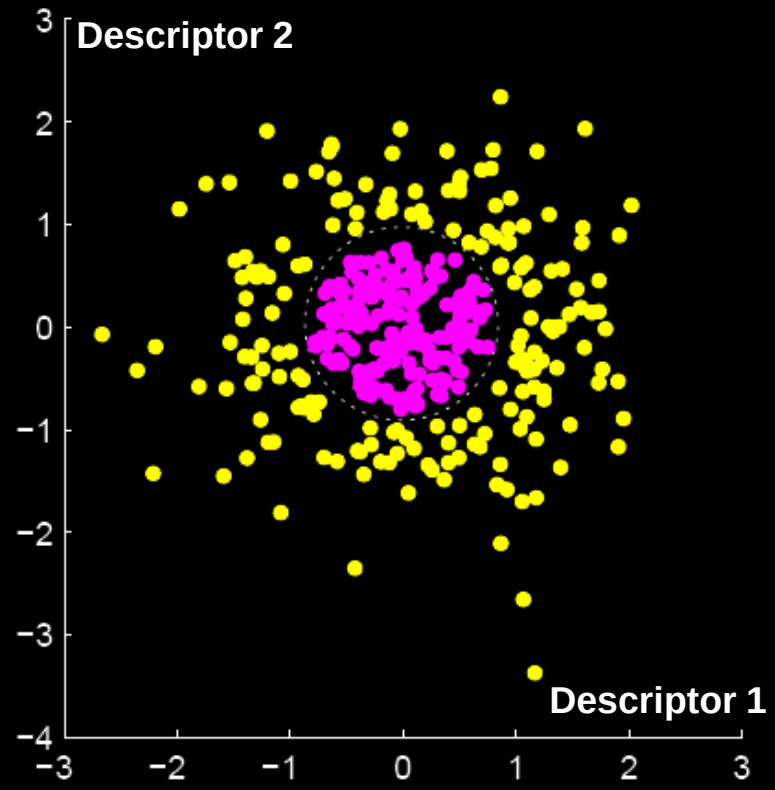
High-throughput cellular-based assays



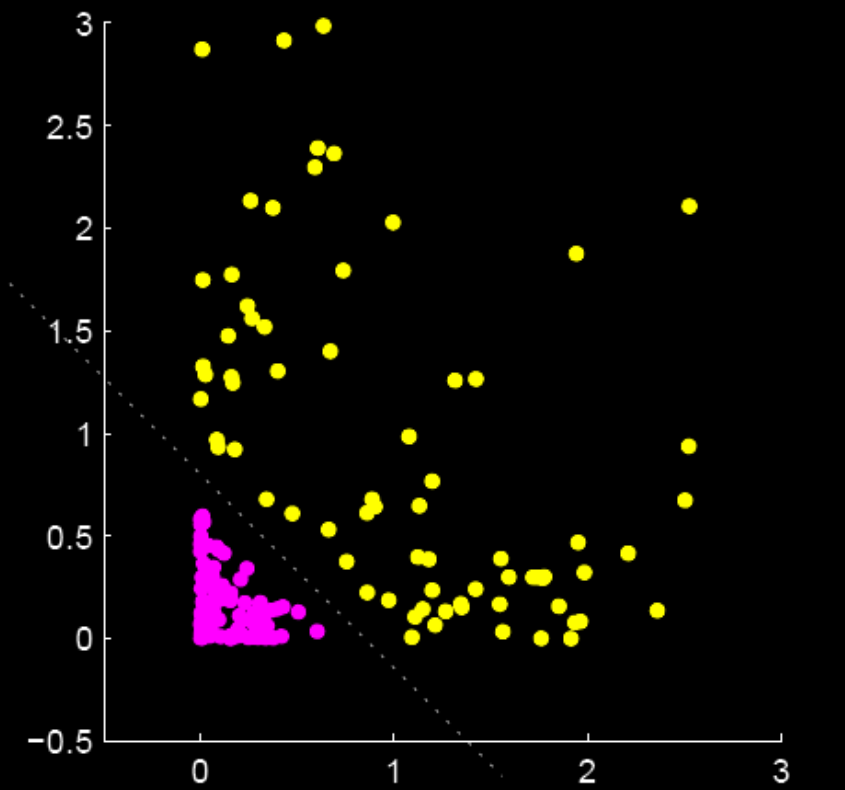
Support Vector Machine (SVM)

Introduced by Vapnik (1995), the SVM approach identifies the best linear separation between two classes of data. In a multidimensional descriptor space, such separation is realized by a hyperplane leading to the best linear segregation between data in the feature space.

What is the feature space ?



Kernel Function
(Ex: x^2, y^2)



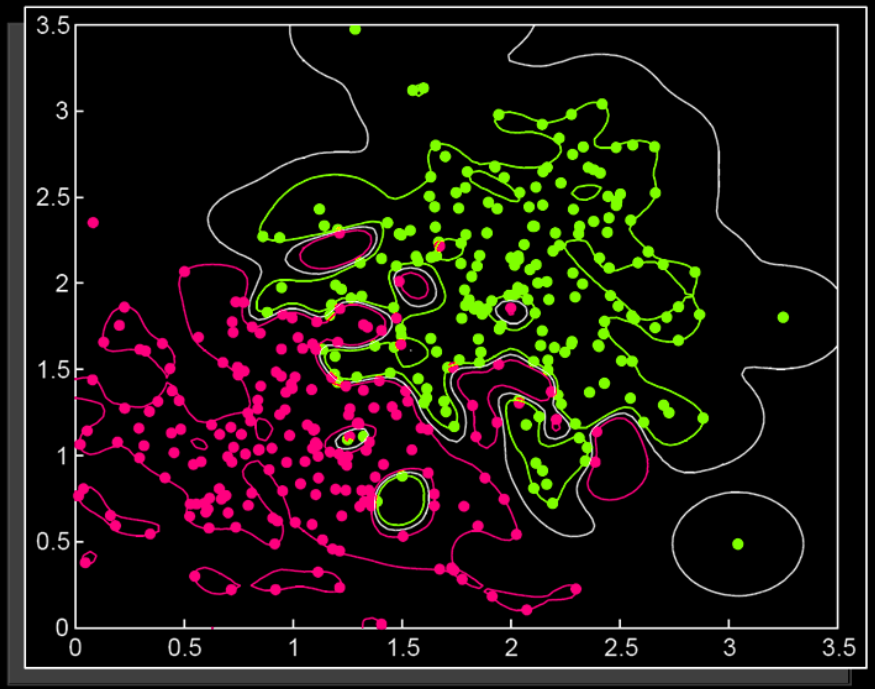
Descriptor space : no linear separation between data points.

Feature space : linear separation exists.

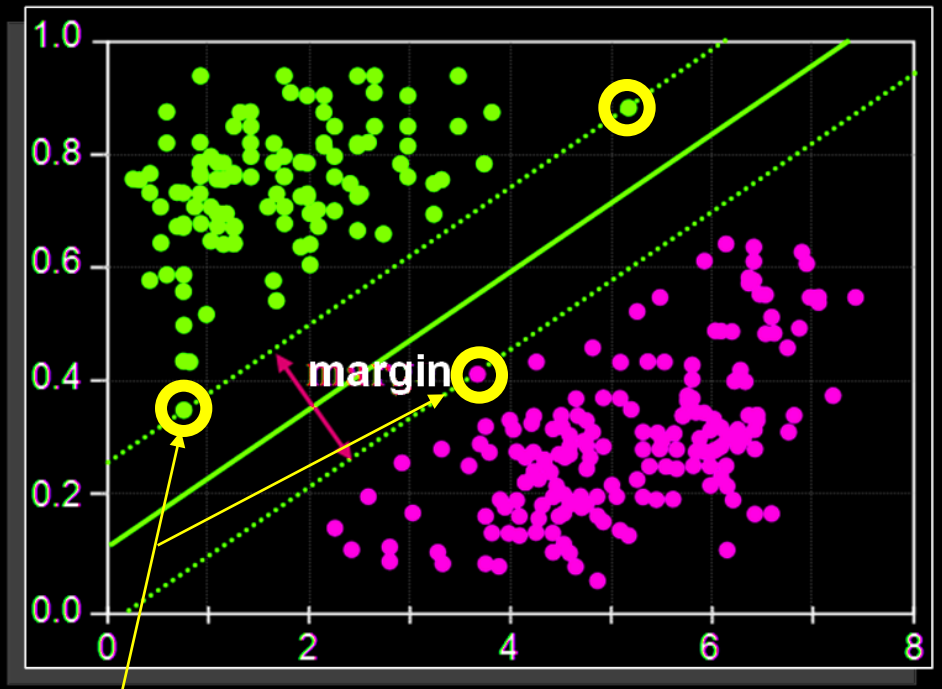
Support Vector Machine (SVM)

The SVM algorithm tends to maximize the margin around the hyperplane separating the two class of compounds. Different kernel functions and parameters have to be optimized (grid search) in order to identify the best models.

Descriptor space



Feature space



Support vectors

Outline

MML
UNC.EDU

1. Availability of data on structure and biological activity of nanomaterials

2. Application of the Quantitative Structure Activity Relationships (QSAR) approach to nanostructure-activity relationship (QNTR) modeling

3. Preliminary studies

CS1- QNTR of whole NPs: Modeling of cellular effects induced by diverse NP [1]

CS2- QNTR of surface modifiers: Modeling of the NPs uptake in PaCa2 cancer cells [2]

[1] Shaw et al. PNAS, 2008, 105, 7387-7392

[2] Weissleder et al. *Nat. Biotechnol.*, 2005, 23 (11), 1418-1423

Case Study 1

Recently[1], 51 diverse NPs were tested *in-vitro* against 4 cell lines in 4 different assays at 4 different concentrations (→ 51x64 data matrix).

NANOPARTICLES

- → cross-linked iron oxide (**CLIO**)-based (23 NPs)
- → pseudocaged nanoparticle (**PNP**)-based (19 NPs)
- → monocrystalline iron oxide nanoparticle (**MION**)-based (4 NPs)
- → **quantum dot**-based with a CdSe core, a ZnS shell, and a polymer coating (3 NPs)
- → two other iron-based MNPs: Feridex IV (approved for in vivo imaging) and Ferrum Hausmann (approved for iron supplementation)

[1] Shaw et al. Perturbational profiling of nanomaterial biologic activity. PNAS, 2008, 105, 7387-7392

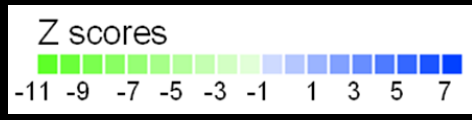
Case Study 1

Recently[1], 51 diverse NPs were tested *in-vitro* against 4 cell lines in 4 different assays at 4 different concentrations (→ 51x64 data matrix).

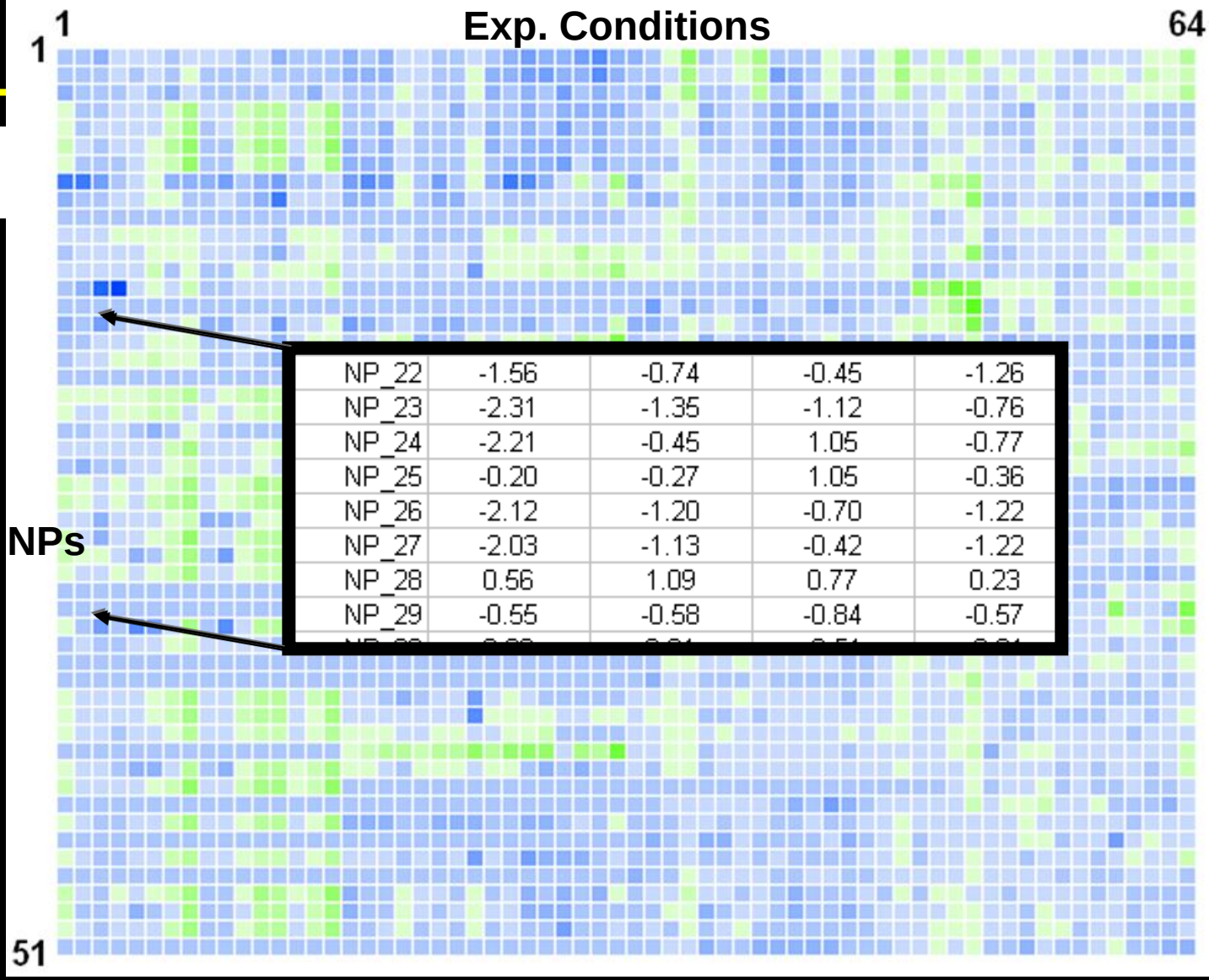
Cell lines	x	Assays	x	Concentrations
→ Vascular cells (endothelial) → Vascular cells (smooth muscle cells) → Monocytes → Hepatocytes		→ ATP content → Reducing equivalents → Caspase-mediated apoptosis → Mitochondrial membrane potential		→ 0.01 → 0.03 → 0.1 → 0.3 mg/ml Fe for iron-based nanoparticles

[1] Shaw et al. Perturbational profiling of nanomaterial biologic activity. PNAS, 2008, 105, 7387-7392

Case Study 1



Initial activity matrix



NP_22	-1.56	-0.74	-0.45	-1.26
NP_23	-2.31	-1.35	-1.12	-0.76
NP_24	-2.21	-0.45	1.05	-0.77
NP_25	-0.20	-0.27	1.05	-0.36
NP_26	-2.12	-1.20	-0.70	-1.22
NP_27	-2.03	-1.13	-0.42	-1.22
NP_28	0.56	1.09	0.77	0.23
NP_29	-0.55	-0.58	-0.84	-0.57

Z scores: assay values were expressed in units of standard deviations of the distribution obtained when cells are treated with PBS (Phosphate Buffered Saline) alone.

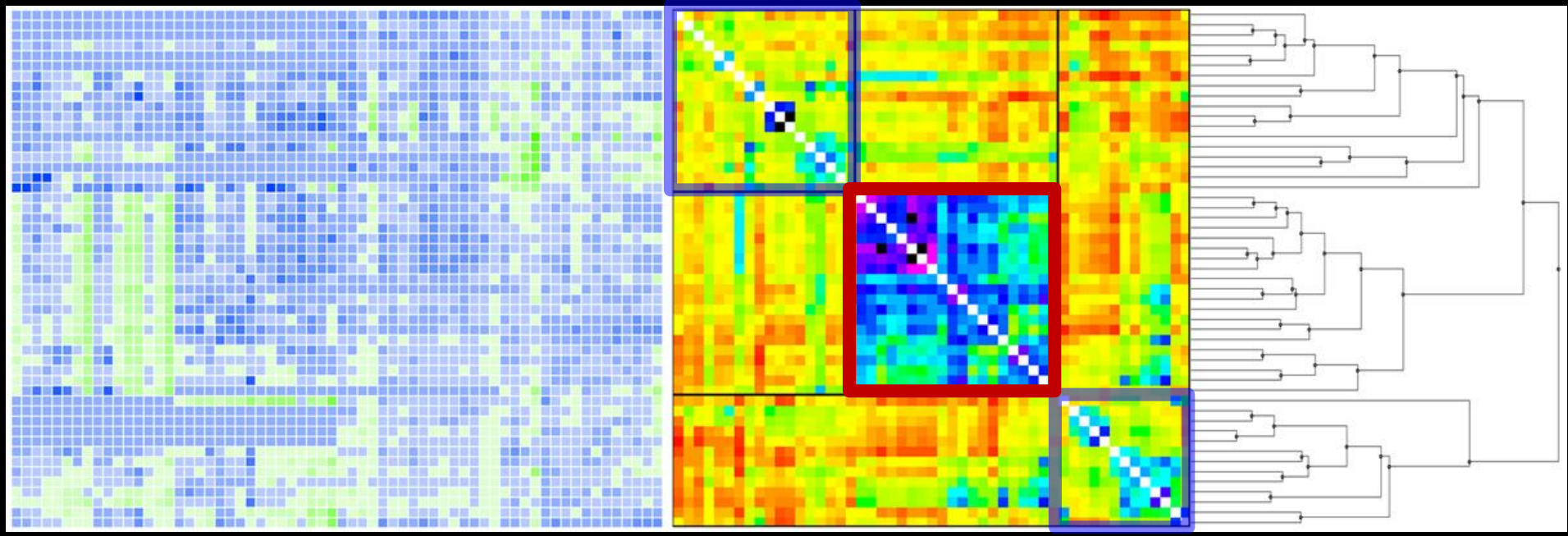
$$Z_{NP} = (\mu_{NP} - \mu_{PBS}) / \sigma_{PBS}$$

μ_{NP} : mean of control tests with PBS

σ_{NP} : standard deviation of control tests with tests

CS1. Hierarchical clustering of the activity matrix

After the normalization of data, ISIDA/Cluster program* was used to cluster the activity matrix (51 * 64), using Johnson's hierarchical method, Euclidean metrics and complete linkage.



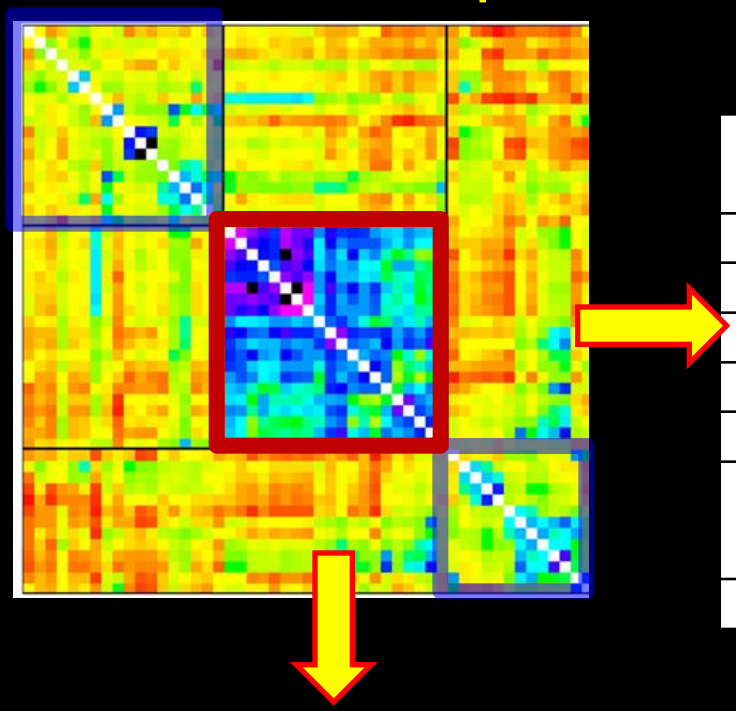
Clustered Activity Matrix

Clustered Distance Matrix

Dendogram

* http://infochim.u-strasbg.fr

Critical information concerning structure-activity relationships can be extracted from the analysis of clusters



NP type	CLUSTER 1	CLUSTER 2	CLUSTER 3	Total
CLIO	7	13	3	23
PNP	7	2	10	19
MION	0	4	0	4
Qt-dot	3	0	0	3
Feridex	0	1	0	1
Ferrum Haussmann	1	0	0	1
Total	18	20	13	

NP Core	CLUST 1	CLUST 2	CLUST 3	Total
Fe_2O_3	5	0	9	14
Fe_3O_4	9	20	4	33
Cd-Se	3	0	0	3
Fe(III)	1	0	0	1
Total	18	20	13	

A given metal core (i.e, Fe_3O_4) or NP category (i.e, Qt-dot), will induce similar biological effects in most cases, independent of the surface modifications.

CS1. QNTR matrix

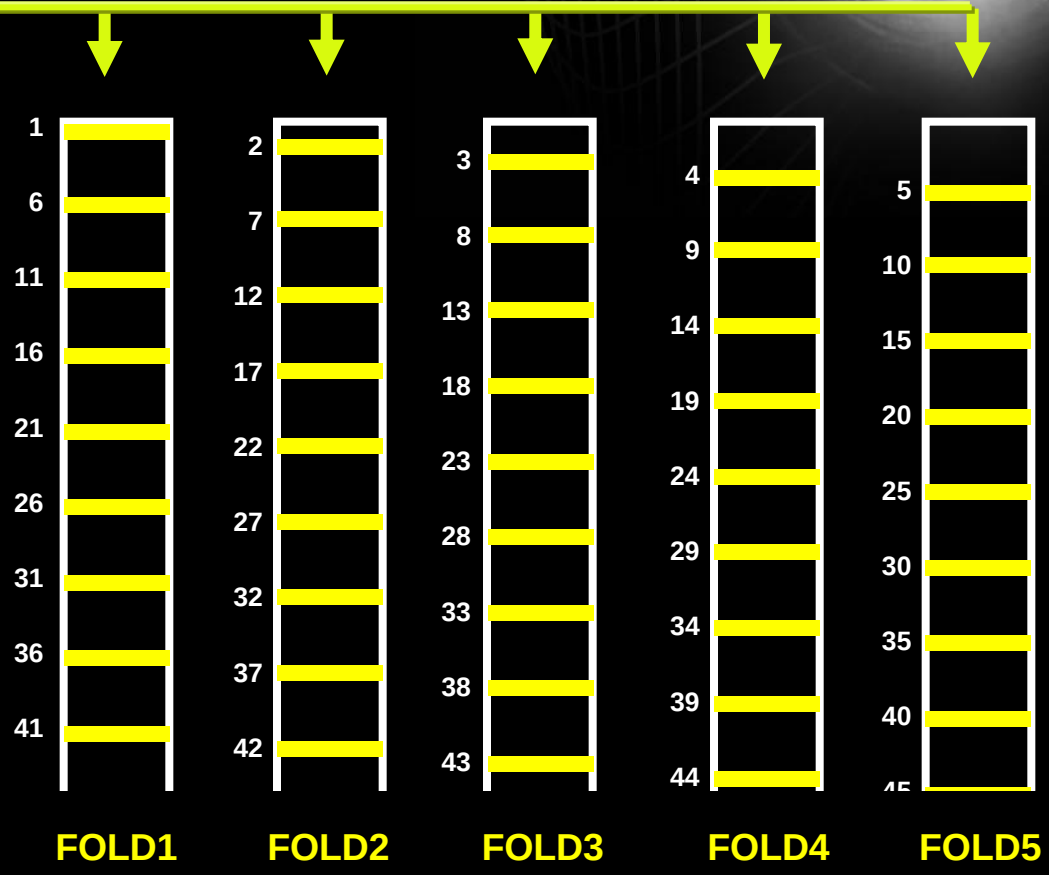
	Effect	Size		Zeta pot.		Relaxitivities	
NP-01	High	0.4865	0.5278	0.2941	0.3986		
NP-02	Low	0.4054	0.7222	0.4837	0.6476		
NP-03	High	0.4324	0.5833	0.3529	1.0000		
NP-04	Low	1.0000	0.5833	1.0000	0.7991		
NP-05	High	0.3649	0.4722	0.2353	0.9403		
NP-06	High	0.3919	0.6111	0.3333	0.9079		
NP-07	High	0.5135	0.5833	0.4052	0.6270		

For 44 NPs, size, zeta potential and relaxitivities were available, and then normalized between 0 and 1, to form the QNTR matrix.

Is it possible to predict whether a given particle will induce low or high biological effects using QNTR models?

CS1. QNTR modeling of 44 diverse NPs

- 44 diverse NPs
- MML-WinSVM program for Windows
- 5 fold external cross-validation procedure



20% of compounds → EXTERNAL SET

80% of compounds → MODELING SET

Models are built using the modeling set ONLY.

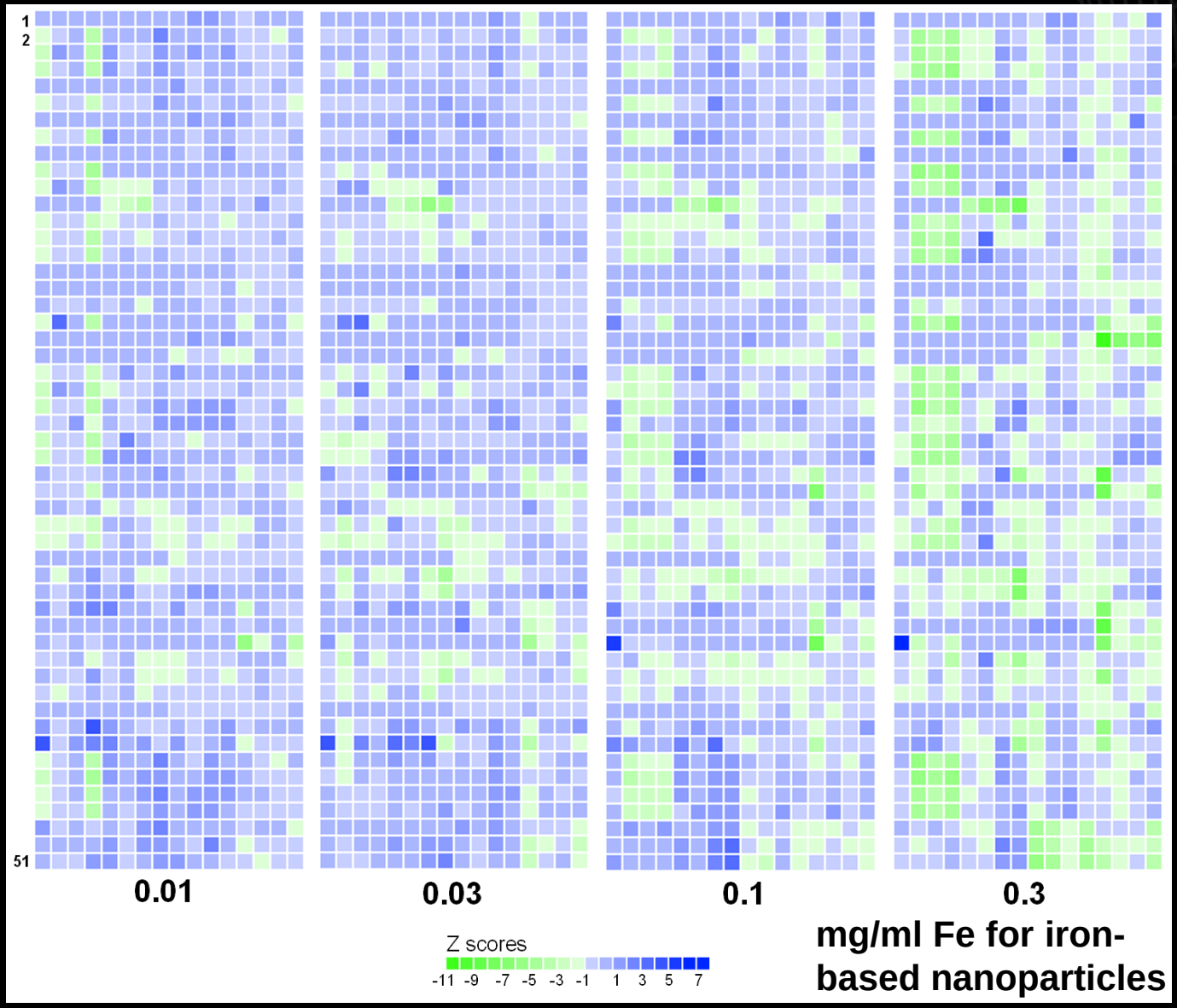
CS1. QNTR modeling results of 44 diverse NPs using MML-WinSVM and a 5 fold external cross-validation

Fold	MODELING SETS				EXTERNAL SETS				
	n	# models	% accuracy internal 5-fold CV	% accuracy	*n*	% accuracy	% CCR[a]	% Sensitivity	% Specificity
1	35	11	51.4 – 60.0	71.4 – 82.9	9	**78**	83	67	100
2	35	13	51.4 – 60.0	71.4 – 77.1	9	**78**	75	50	100
3	35	16	57.1 – 62.9	74.3 – 82.9	9	**78**	78	80	75
4	35	11	60.0 – 62.9	77.1 – 88.6	9	**56**	55	50	60
5	36	4	66.7	83.3 – 86.1	8	**75**	67	33	100
					44	**73**	**73**	**60**	**86**

[a]CCR – Correct Classification Rate.

Prediction performances are surprisingly good : the overall prediction accuracy for those 44 NPs is equal to 73 %

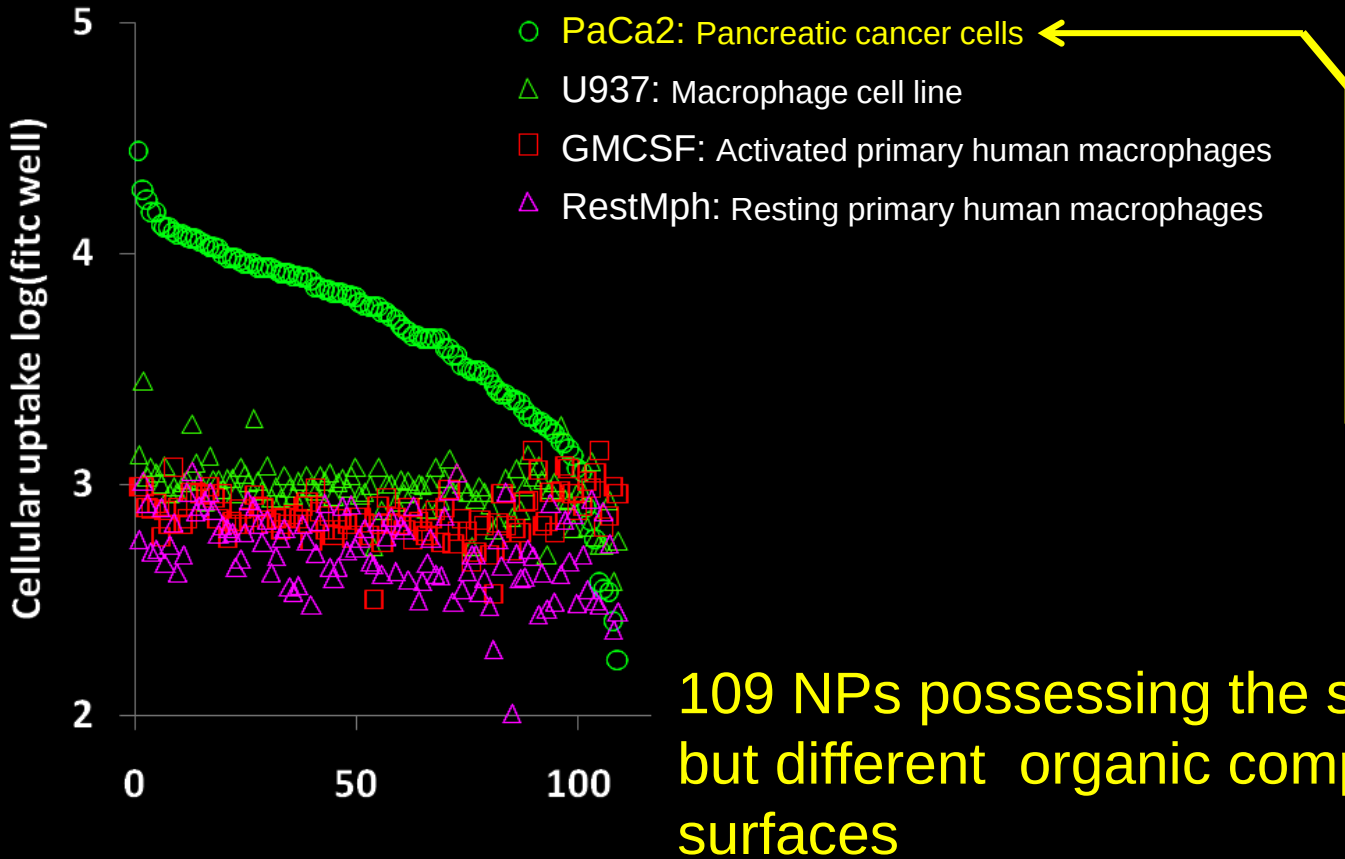
CS1. Dose-dependency of NP effects



From the analysis of activity matrix, we also show that **biological effects induced by NPs are dose-dependent.**

CS2. QNTR of surface modifiers: modeling of the NPs uptake in PaCa2 cancer cells*

Nanomaterials with precise biological functions have high potential for use in biomedical applications. Recently, Weissleder et al.* investigated whether the multivalent attachment of small organic molecules on a same NP can modify its binding affinity to certain cells.



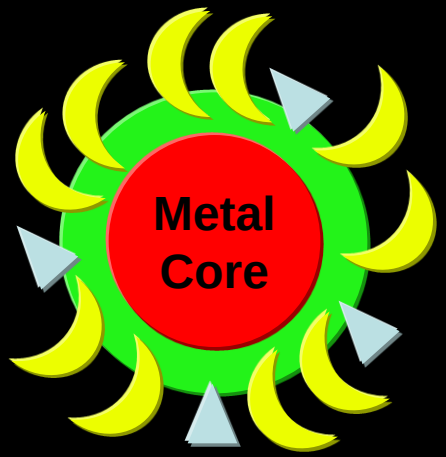
Unlike the other cell lines, the PaCa2 pancreatic cancer cells showed diverse cellular uptakes for different NPs.

109 NPs possessing the same core (CLIO) but different organic compounds on their surfaces

* Weissleder et al. *Nat. Biotechnol.*, 2005, 23 (11), 1418-1423

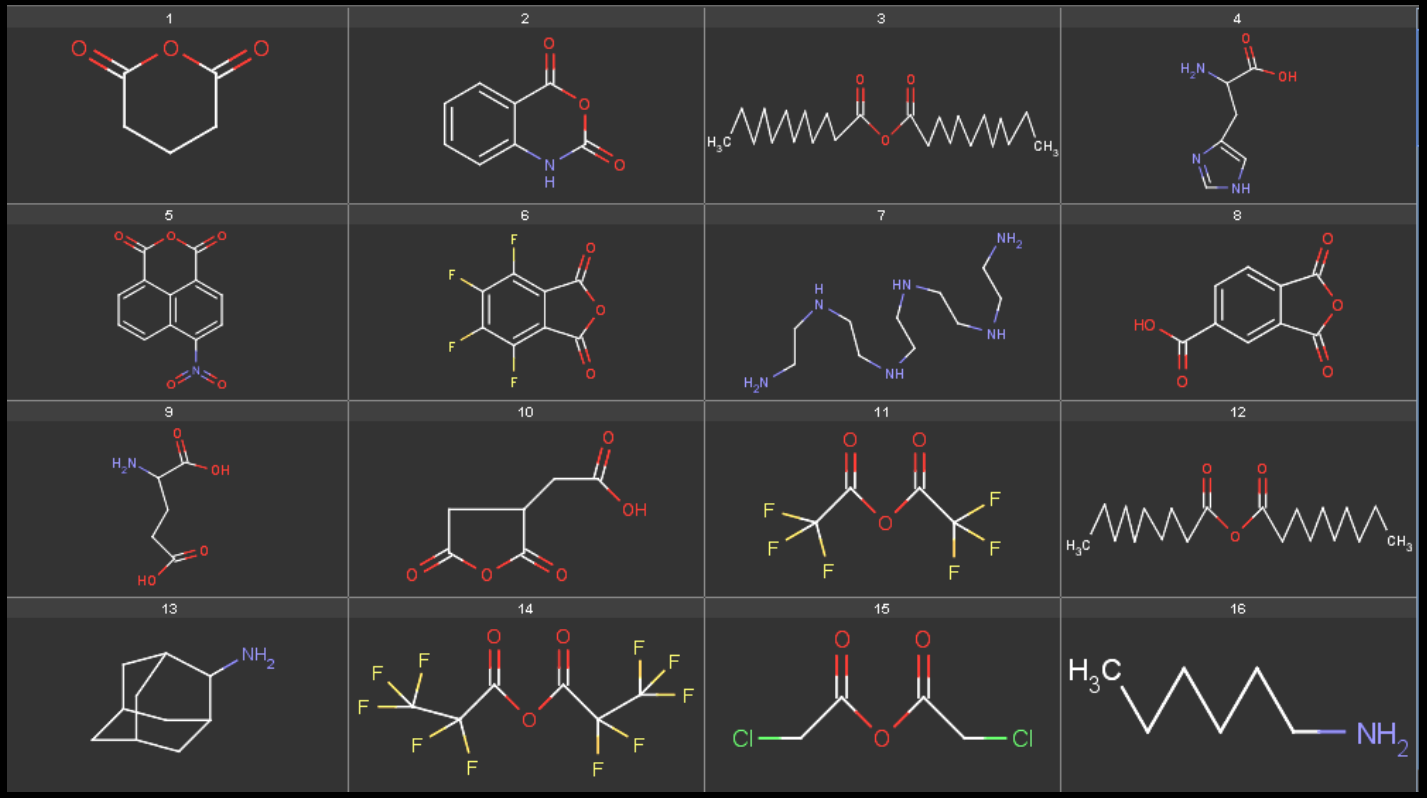
CS2. Dataset representation

109 NPs possessing the same core (CLIO) but different organic compounds on their surfaces



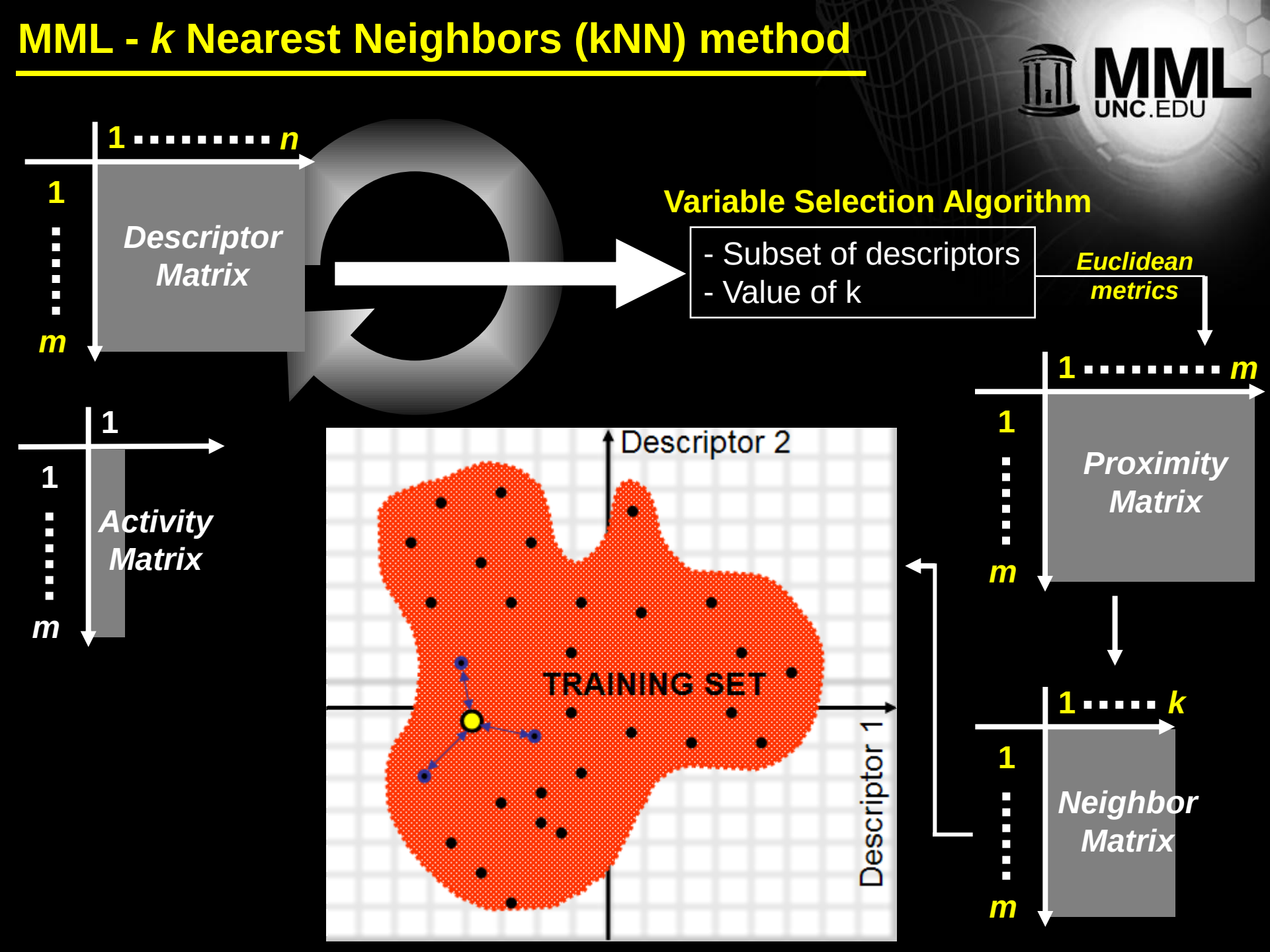
FITC (fluorescein isothiocyanate)

Small organic compounds



Each NP was represented by one single organic compound of its surface. Classical Dragon molecular descriptors, used for standard QSAR studies, have been calculated for those small organic molecules only.

MML - *k* Nearest Neighbors (kNN) method



CS2. QNTR modeling results

- 109 surface modifiers and their corresponding PaCa2 uptake
- MML-kNN program using 2D Dragon molecular descriptors
- 5 fold external cross-validation procedure

PREDICTION PERFORMANCES OF MODELS ARE GOOD

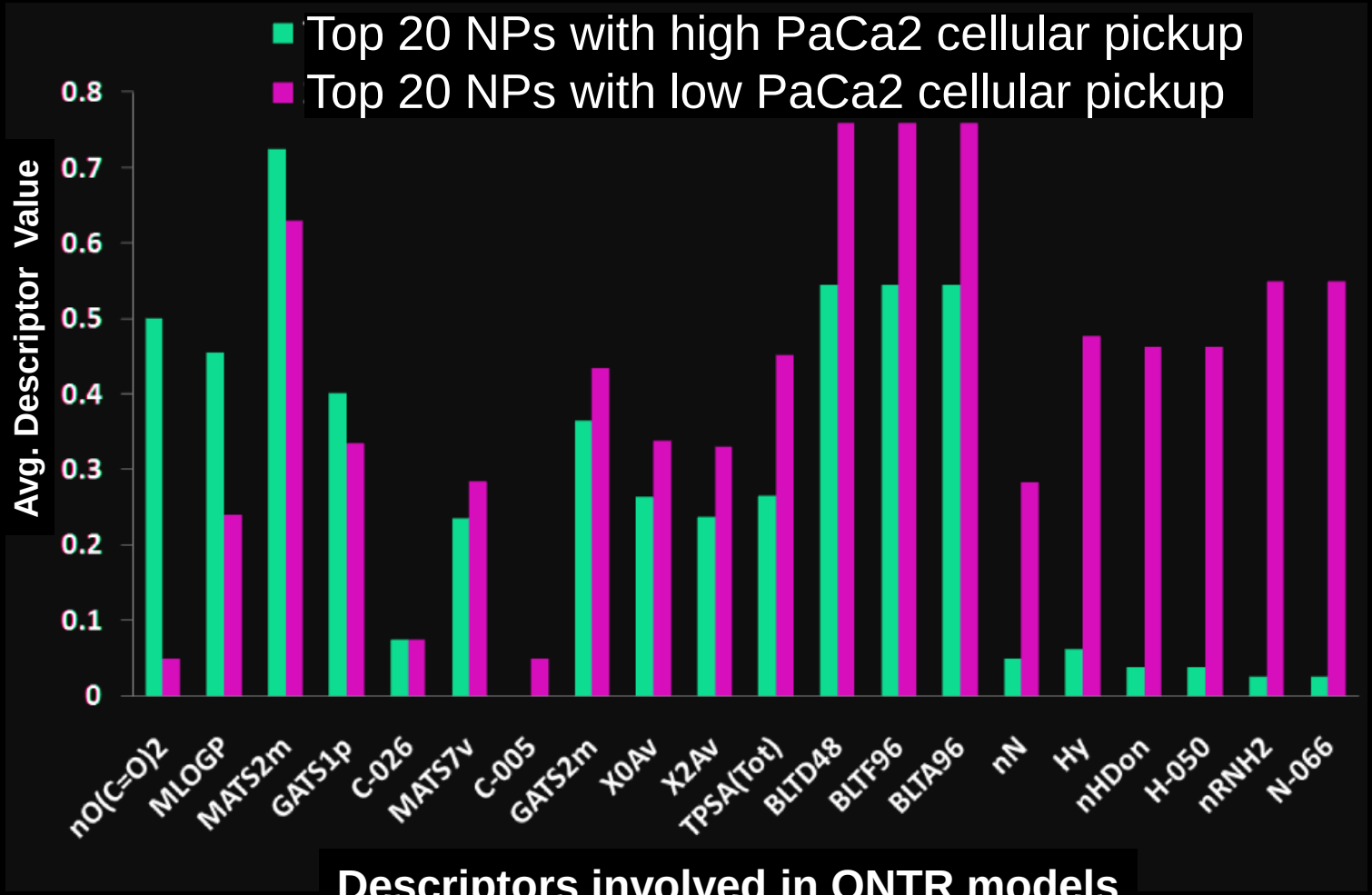
Fold	$Q^2/R^2/R^2_{ext}$	MAE
1	0.78/0.85/0.87	0.15
2	0.80/0.86/0.79	0.15
3	0.83/0.90/0.73	0.20
4	0.84/0.91/0.61	0.17
5	0.84/0.88/0.49	0.19

Total R^2_{ext} = 0.72, MAE = 0.18

Using our models, experimentalists could avoid some expensive and time consuming compound synthesis and cell based assays as well.

CS2. QNTR modeling results

The QNTR approach also gives critical information concerning the contributions of each descriptor to NP cellular uptake:



Conclusions

- Preliminary modeling results demonstrate that QNTR models can successfully predict the biological effects of NPs from their descriptors either experimentally measured (e.g., first case study) , or calculated (second case study).

- To increase the accuracy and impact of models on the experiments, we need more systematic experimental data (structural and biological).

- QNTR approach may allow rational design or prioritization of novel NPs with desired target (physical and biological) properties.

ACKNOWLEDGEMENTS

-Dr. Russell Mumper, UNC
-Dr. Lin Ye, UNC

- We also thank Drs. Shaw and Weissleder for providing us with the NP datasets.

The support from **NIH** (grants P20-HG003898 and R01-GM66940) is gratefully acknowledged.

UNC ESHELMAN SCHOOL OF PHARMACY

Molecular Modeling Laboratory

Principal Investigator
Alexander Tropsha

Research Professors
Clark Jeffries, Alexander Golbraikh, Hao Zhu, Simon Wang, M. Karthikeyan

Postdoctoral Fellows
Denis Fourches, Eugene Muratov, Georgiy Abramochkin, Lin Ye

MML
UNC.EDU

Visiting Research Scientist
Aleks Sedykh

Graduate Research Students
Christopher Grulke, Kun Wang, Hao Tang, Jui-Hua Hsieh, Rima Hajjo, Tanarat Kietsakorn, Tony Wu, Liying Zhang, Melody Luo, Guiyu Zhao, Andrew Fant

Research Programmers
Theo Walker

System Administrator
Mihir Shah

Adjunct Members
Weifan Zheng, Shubin Liu

Administrative Assistant
Paula Press