Helsinki

EPA's ToxCast™ Program for Predicting Hazard and Prioritizing Toxicity Testing of Environmental Chemicals

Robert Kavlock

October, 2007

Goals of Computational Toxicology at the EPA

- Screening and Prioritization of Environmental Chemicals
  - –Many widely used (>10,000)
  - –Few completely characterized (1,000-2,000)
  - –Reduce use of animal testing

  Model / characterize known toxic chemicals
  - –Hazard ID
  - –Mode of action
  - –Mechanism of action
  - –Dose response
- Working hypothesis
  - –A small number of commonly used chemicals account for a significant amount of human disease / ecological damage

ToxCast Background

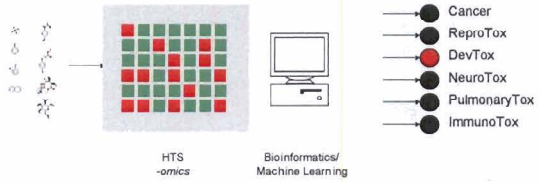
- Research program developed by EPA's National Center for Computational Toxicology
- Formulated to address chemical screening and prioritization needs
- Based on experience of the pharmaceutical industry
- Comprehensive use of current technology
- Phased approach to evaluate utility
- Committed to stakeholder involvement and release of data to public domain

Current Approach for Toxicity Testing

in vivo testing

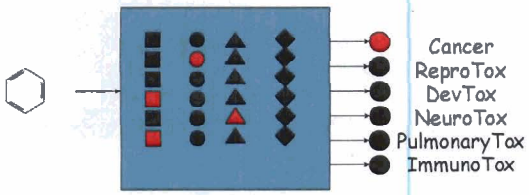
Future of Toxicity Testing

in vitro testing *in silico* analysis



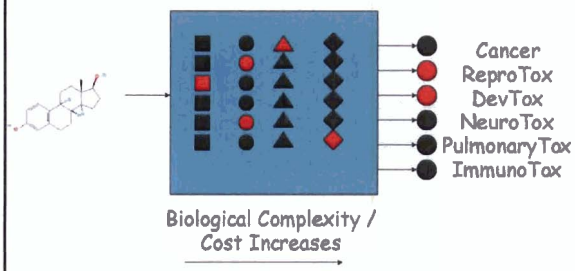
The ToxCast™ Project

Toxicity Prediction Tomorrow: Find Pattern of Assays that Predicts Tox

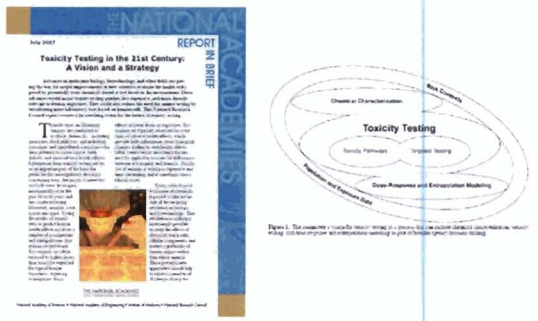


Biological Complexity / Cost Increases

Toxicity Prediction Tomorrow: Find Pattern of Assays that Predicts Tox



Biological Complexity / Cost Increases

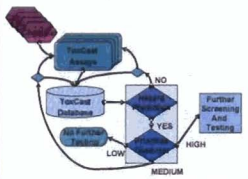


ToxCast "Target" Chemicals

- Environmental Chemicals
  - -Used in large quantities
  - -Found in drinking water/air/soil
- Source Lists
  - -HPV
  - -MPV (IUR)
  - -Pesticides (Actives & Inerts)
  - -TRI
  - -DWC
  - -Endocrine Disruptors (EDC73)
- Total of ~11,000 unique chemicals

Phased Development

Phase	Number of Chemicals	Chemical Criteria	Purpose	Est. Cost per Chemical	Target Date
I	>300	Data Rich (pesticides)	Signature Development	$20k	FY07-08
II	>1000	Expanded Structure and Use Diversity	Evaluation and Extension	$12-15k	FY08-09
III	Thousands	Data poor	Prediction and Prioritization	$6-10k	FY10-12



- Deliver an affordable, science-based system for categorizing chemicals
- Increasing confidence as database grows
- Identify potential mechanisms of action
- Refine and reduce use of animals in hazard identification and risk assessment

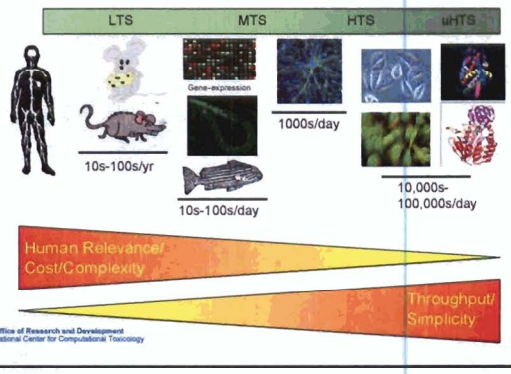
Example of Chemical Classes to be Investigated in Phase I

299 with complete OPP toxicity datasets
30 Carbamates (plus one metabolite)
33 Organophosphates (plus several metabolites)
12 Pyrethroids
12 Triazines (plus one metabolite)
17 Azole Fungicides (plus one metabolite)
13 Organochlorines
7 Phthalates (and several metabolites)
16 HPVs, 11 HPV challenge

Classification based on OPPIN

High-Throughput Screening Assays

batch testing of chemicals for pharmacological/toxicological end points using automated liquid handling, detectors, and data acquisitio n



EPA

BioFocus DPI

NOVASCREEN

PHYLONIX

ToxCast™

BioSeek

attagene

Nine contracts and one IAG providing chemical procurement, biochemical assays, cellular reporter assays and genomics, complex human cell responses, and model organisms; capacity to screen up to 10,000 chemicals in over 400 assays by 2012

CELLUMEN

NCGC

ACEA

EXPRESSION ANALYSIS

Office of Research and Development
National Center for Computational Toxicology

12

EPA

NOVASCREEN
a Caliper Life Sciences company

Biochemical and Cellular High Throughput Screening (HTS)

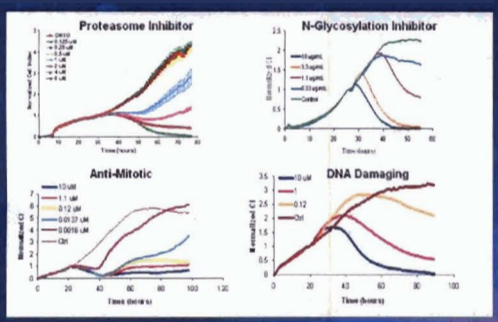
Number of Assays

- 30 Cytochrome P450s
- 81 GPCRs
- 22 Ion Channels
- 28 Kinases
- 24 Nuclear Receptors
- 19 Phosphatases
- 9 Transporters
- 27 Other Enzymes

Office of Research and Development
National Center for Computational Toxicology

13

ACEA RT-CES™ Impedance-based Biomonitoring of Cellular Cytotoxicity



ACEA

EPA

High Content Cytotoxicity Assessment

CELLUMEN

Stress Pathway Activation

Organelle Functions

Oxidative Stress

DNA Damage

Cell Cycle

CSK Integrity

- HepG2
- Acute, early & chronic exposure
- Multiple mechanisms of toxicity
- 384-well capacity

Office of Research and Development
National Center for Computational Toxicology

15

EPA

BioMAP Systems
Inflammation / Autoimmune Disease / Respiratory / Cardiovascular

BioSeek

Primary Human Cell Types	Disease Relevance	BioMAP® System
Endothelial cells (EC)	Th1/ Th2 inflammation, allergy, asthma, dermatitis, angiogenesis, wound healing, restenosis, atherosclerosis (coronary art.)	3C, 4H, CASM
EC + Peripheral Blood Mononuclear Cells	Th1 inflammation, psoriasis, COPD fibrosis, monocyte and T cell responses	LPS, SAg, MFNo
EC + Macrophages	Macrophage responses, arthritis, COPD, fibrosis	
EC + Mast cells	Asthma, allergy, dermatitis, fibrosis	MyBg, MCIgE
EC + Smooth Muscle Cells	Vascular biology, restenosis, atherosclerosis	HSM3C, HTH2
EC + Th2 blasts	Allergy, asthma	
Fibroblasts	Arthritis, asthma, dermatitis, fibrosis, psoriasis, wound healing	HDFT, HDF3CT, HDF3C, HDFNo, HDF3CGF
Myofibroblasts	Fibrosis, COPD, wound healing	
Keratinocytes	Psoriasis, dermatitis, wound healing	KJ3CT, MyoF
Keratinocytes + Fibroblasts	Psoriasis, dermatitis, wound healing	KF3CT, KFNo
Bronchial Epithelial Cells	Th1 and Th2 inflammation Allergy, asthma, fibrosis, COPD	BE3C, BE4T
Bronchial Epithelial Cells + Fibroblasts	Asthma, allergy, fibrosis, COPD	BF4T, BMPC, CASM3C
Smooth Muscle Cells	Vascular inflammation, asthma, COPD, fibrosis (coronary artery SMC)	

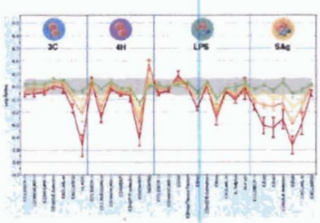
Office of Research and Development
National Center for Computational Toxicology

EPA

Biomarkers of Cell Function in Complex Human In Vitro Culture Systems

BioSeek

System	Environment	Cell Types	Readouts
3C			
4H			
LPS			
SAg			



Berg, et al JPET M 53:67-74 (2006)

Office of Research and Development
National Center for Computational Toxicology

17

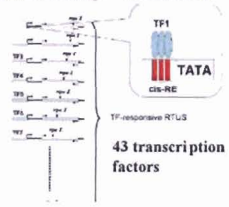
Mixed Cultures of Primary Cells

Integrated Discrete Multiple Organ Cell Culture
IDMOC
(U. S. Patent Allowed November 2006; published March, 2007)

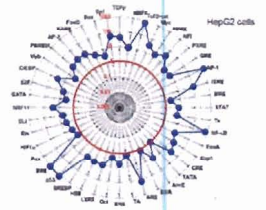
Transcription Factor Activity Profiling

Cis-Factorial™ Biosensors

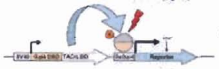
attagene



43 transcription factors



Trans-Factorial™ Biosensors



24 nuclear receptors

Nuclear Receptor Activation

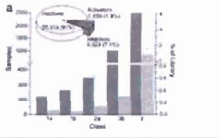
PNAS

Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries

James Inglese*, Douglas S. Auld, Ajit Jadhav, Ronald L. Johnson, Anton Simeonov, Adam Yasgar, Wei Zheng, and Christopher P. Austin

NIH Chemical Genomics Center, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892-3370

Communicated by Francis S. Collins, National Institutes of Health, Bethesda, MD, May 31, 2006 (received for review April 12, 2006)



PNAS August 2006 vol.103 no.31 11473-11478

Reporter gene assays

From Invitrogen:
- AR
- ERa
- FXR
- GR
- LXRb
- PPARd
- RXR
- TRb
- VDR
- RORa

Considering from other sources:
- PPARa
- PPARb/d
- PPARg
- PXR
- CAR
- FXR
- RXRa
- ERRa
- TRa
- PR
- ERa
- ERb
- AR
- GR
- MR
- LXRa
- LXRb
- VDR
- RORa

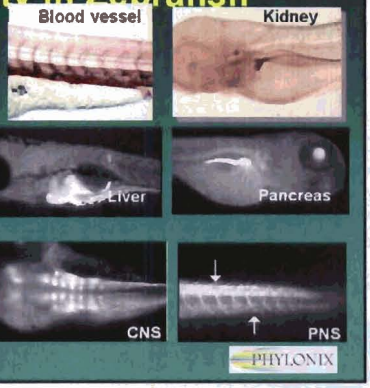
Gene Expression Profiling by Microarray and PCR

EXPRESSION ANALYSIS

Human BeadChips assay up to 48,000 transcripts.
Mouse BeadChips assay up to 47,000 transcripts.
RatRef BeadChips assay up to 22,000 transcripts.
Customized chips up to 1400 genes in 96well format.
Individual or multiplexed PCR (≤ 48 transcripts in parallel).

Phylonix Assessment of Toxicity in Zebrafish

- Vessels
- Kidney
- Liver
- Pancreas
- Cartilage
- CNS
- Motor Neurons

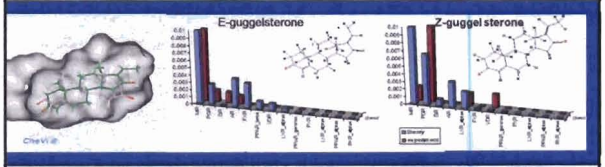


4/18/2007

PHYLONIX

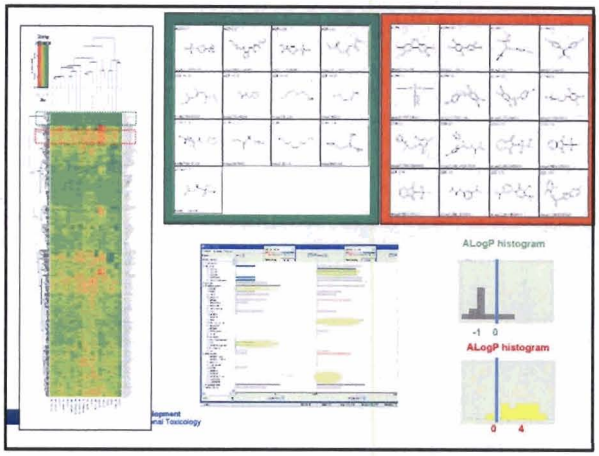
In Silico Screening

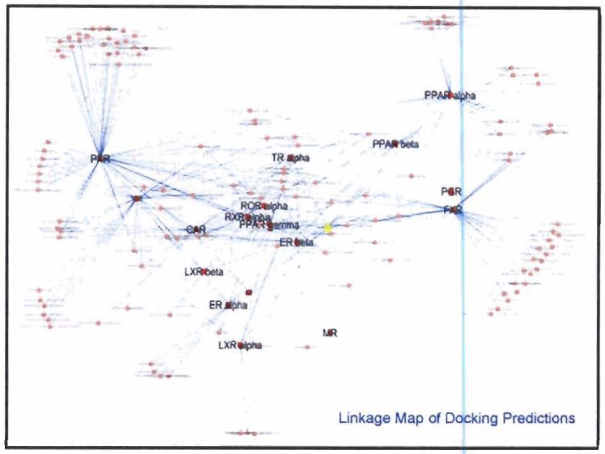
Molecular Docking: Guggelsterone Target Profiling

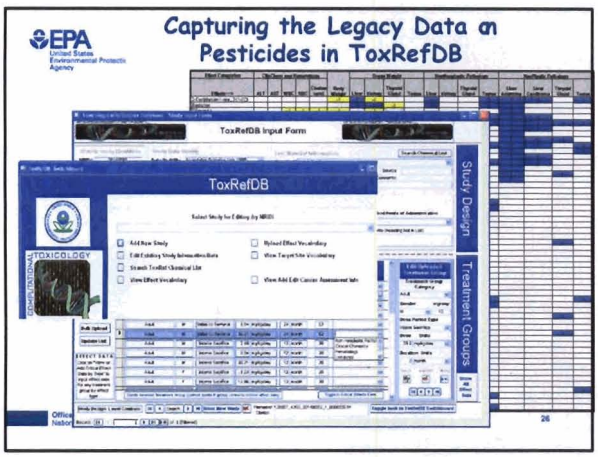


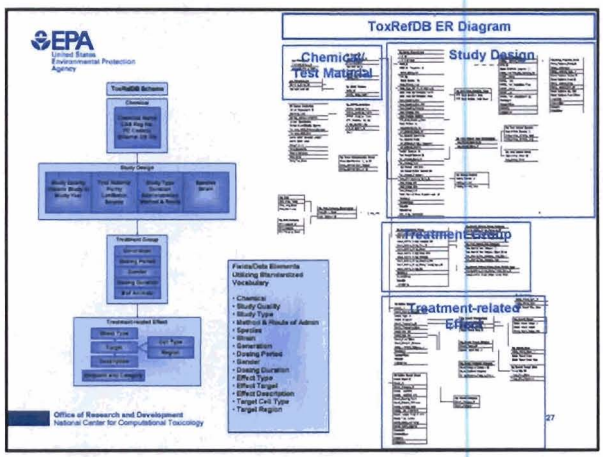
Docking of both E/Z-Guggelsterone geometric isomers against multiple crystal-structure derived human NR targets in their agonist-associated (active) conformation (from www.pdb.org) and MMFFx optimized ligand set geometries with AM1-BCC charges assigned from **MOE (CCG Canada)** as found in ***KiBank*** **(Aizawa 2004)**, curated from the original publication on guggelsterone polypharmacology **(Burris, 2005)**. In the computational toxicology framework we may also pose this question in terms of polytoxicology or *pan-agonism* associated with an adverse rather than therapeutic effect. All performed in **eHiTS** on "fast" screening mode (fewer match-poses generated) (Zsoldos et al 2006) against the diverse set of targets The docked structure of E/Z isomers are shown docked within the binding pocket from MR (mineralcorticoid receptor) one of the top hits for both isomers. The structural formula is shown overlayed on the experiment/theory rank ordered bar graphs (magnitude – normalized binding affinity (K_d) to largest value, so large bars = high affinity).

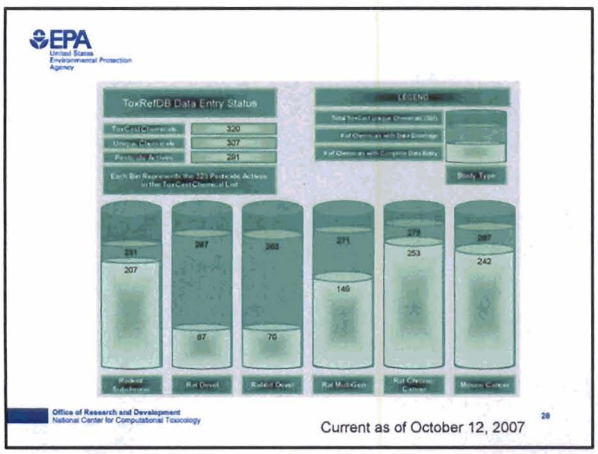
Michael-Rock Goldsmith,
Molecular Modeling Group, NCCT/USEPA

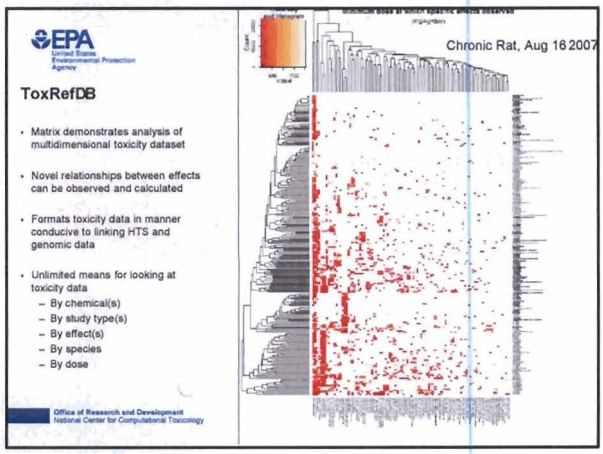

ALogP histogram
-1 0
ALogP histogram
0 4

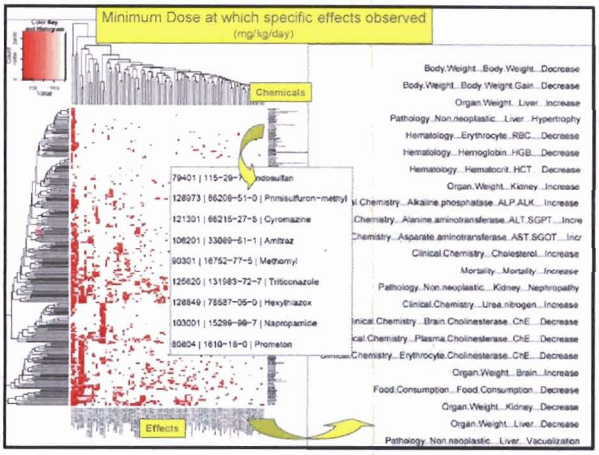

PPARalpha
PPARbeta
TRalpha
PXR
RORalpha
RXRalpha
PPARgamma
CAR
ERbeta
LXRbeta
ERalpha
MR
LXRalpha
Linkage Map of Docking Predictions


EPA
United States Environmental Protection Agency
Capturing the Legacy Data on Pesticides in ToxRefDB
ToxRefDB Input Form
ToxRefDB
Study Design
Treatment Groups
26


EPA
United States Environmental Protection Agency
ToxRefDB ER Diagram
Chemical/ Test Material
Study Design
Treatment Group
Treatment-related Effect
Office of Research and Development
National Center for Computational Toxicology
27


EPA
United States Environmental Protection Agency
ToxRefDB Data Entry Status
320
307
291
Study Type
231
207
87
287
87
303
70
271
146
279
253
287
242
Office of Research and Development
National Center for Computational Toxicology
Current as of October 12, 2007
28


EPA
United States Environmental Protection Agency
ToxRefDB
Matrix demonstrates analysis of multidimensional toxicity dataset
Novel relationships between effects can be observed and calculated
Formats toxicity data in manner conducive to linking HTS and genomic data
Unlimited means for looking at toxicity data
By chemical(s)
By study type(s)
By effect(s)
By species
By dose
Chronic Rat, Aug 16 2007
Office of Research and Development
National Center for Computational Toxicology

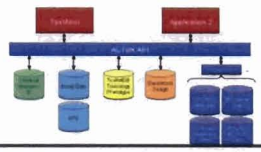


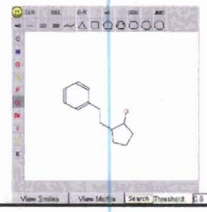
ACToR

Aggregated Computational Toxicology Resource

- Aggregating the world's chemical structure, bioassay and toxicology data for environmental chemicals
  - –130 sources of data
  - –Over 10M chemicals
- Searchable by: Data source, name, CAS, chemical structure
- Will manage all ToxCast Data
- Public release planned for 2008

Structure Search





Tracking Multiple Sources forOverlaps

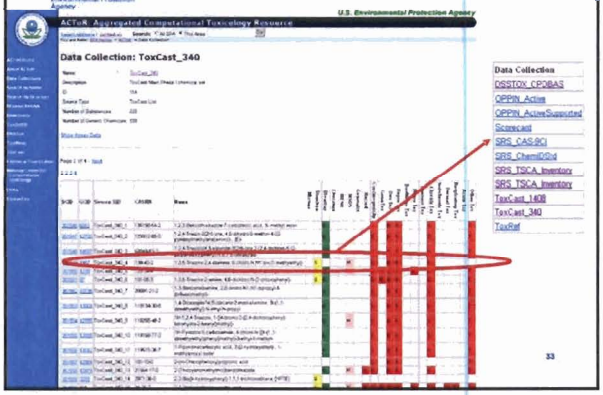
Overlap Matrix

57 Chemicals common to ToxCast I and EDSP 1

328 Chemicals common to Pesticide Actives and ToxCast i

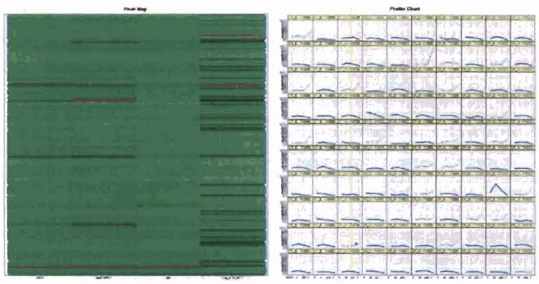
32

Aggregating Data from Global Sources



33

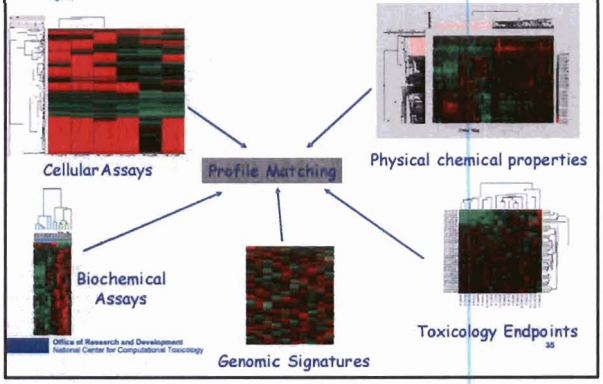
Initial ToxCast Screening Results



Office of Research and Development
National Center for Computational Toxicology

34

Correlating HTS to Toxicity



Office of Research and Development
National Center for Computational Toxicology

35

ToxCast Proof-of-Concept: Signatures of Toxicity for ~300 Pesticides

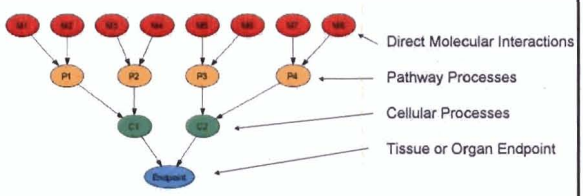
	Chemical	HTS			HCS			Genomics			Toxicity		
Chemical	Physico-Chemical Properties			In-vitro / Biochemical Assays			Cellular Assays			Gene Expression Signatures			Toxicity Endpoints		
	P1	...	PN	A1	...	AN	C1	...	CN	S1	...	SN	T1	...	TN
C1															
C2															
C3															
...															
CN															

Data will be both quantitative and categorical

Finding Tests For Toxicity - Classification

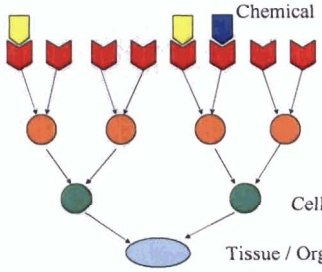
- Goal: Find "classifiers" that accurately predict endpoints
  - –Evaluate standard methods (NN, KNN, SVM, SLR, GA, CART)
- Use all available data
  - –ToxCast & public sources
  - –HTS, HCS, genomics, physicochemical properties, calculated properties
- Properties of an ideal classifier
  - –Accurate (low false positive & false negative rates)
  - –Inexpensive and easy to measure for new chemicals
  - –Easy to interpret – provides biological insight

ToxCast Data Analysis Simulation Model

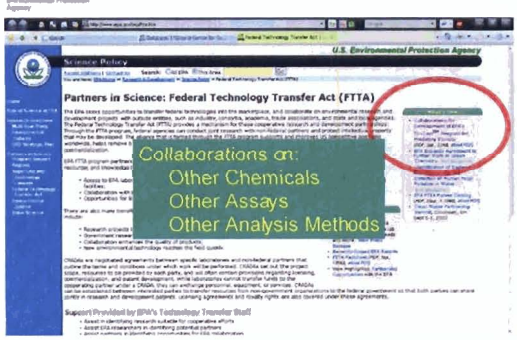


Evaluate algorithms as f(sample size, model complexity, noise)

Thinking Mechanistically

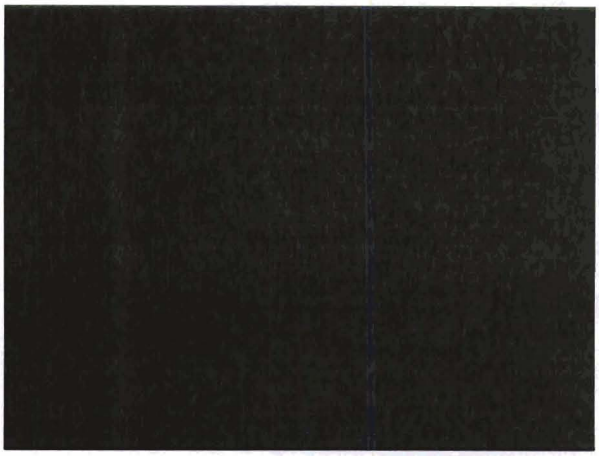


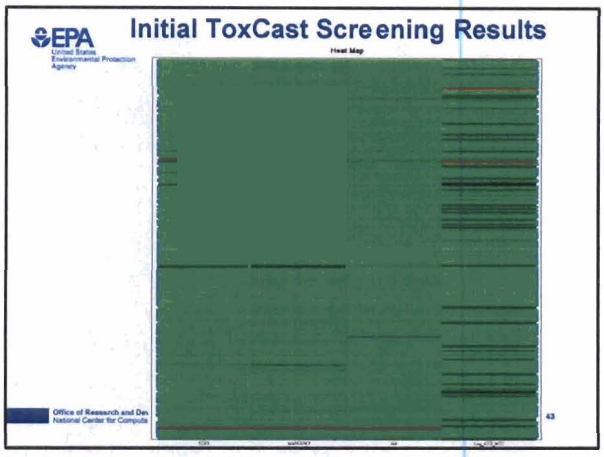
Seeking Collaborative Partners

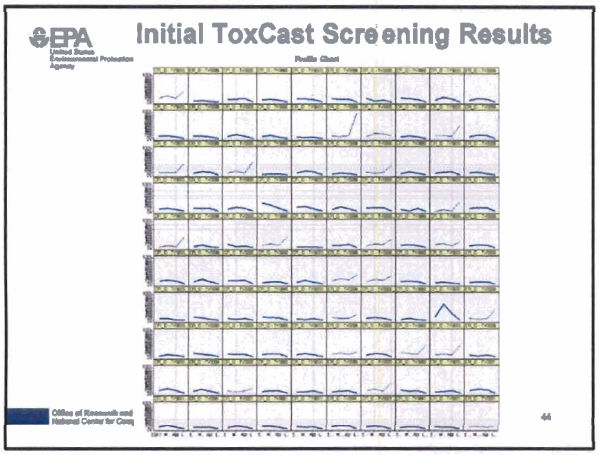


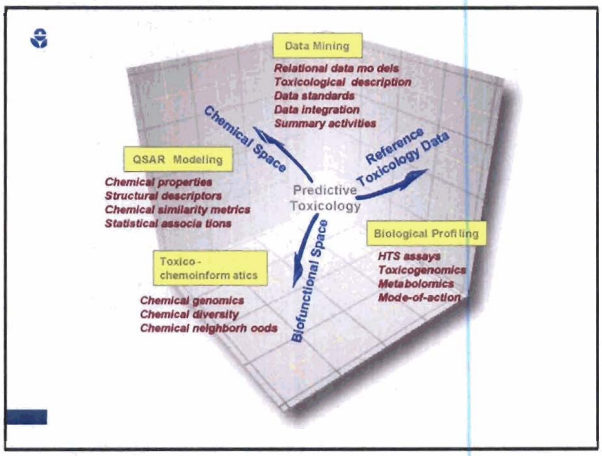
The ToxCast Team

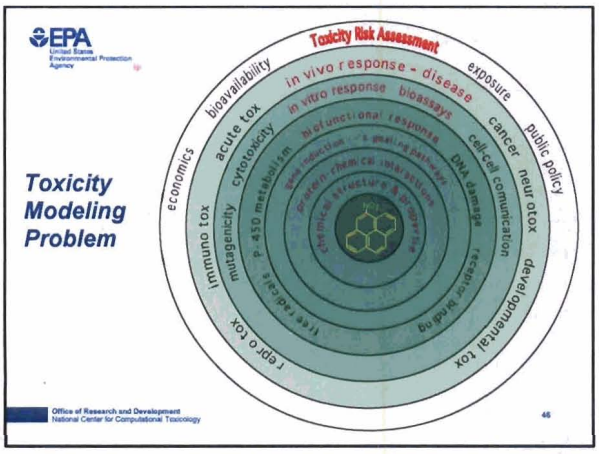
www.epa.gov/ncct/toxcast

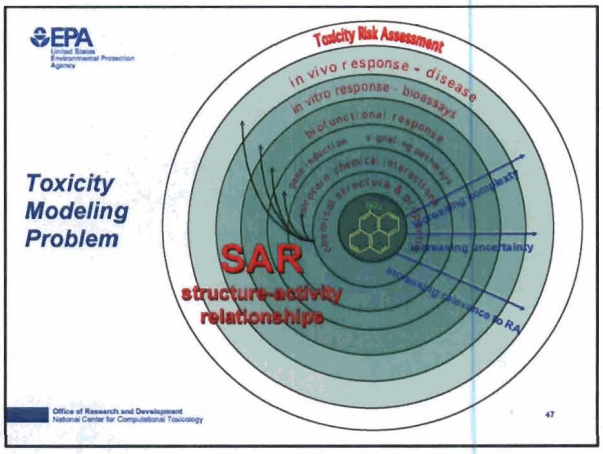



EPA
Initial ToxCast Screening Results
Heat Map
Office of Research and Dev
National Center for Computa
43

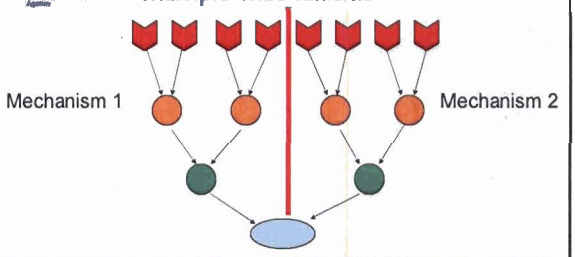

EPA
Initial ToxCast Screening Results
Profile Chart
Office of Research and
National Center for Comp
44


Data Mining
Relational data models
Toxicological description
Data standards
Data integration
Summary activities
Chemical Space
Reference Toxicology Data
QSAR Modeling
Chemical properties
Structural descriptors
Chemical similarity metrics
Statistical associations
Predictive Toxicology
Biological Profiling
HTS assays
Toxicogenomics
Metabolomics
Mode-of-action
Toxico-chemoinformatics
Chemical genomics
Chemical diversity
Chemical neighborhoods
Biofunctional Space


EPA
Toxicity Modeling Problem
Toxicity Risk Assessment
in vivo response - disease
in vitro response - bioassays
biofunctional response
protein-chemical interactions
chemical structure & properties
bioavailability
exposure
public policy
neur otox
developmental tox
repro tox
immuno tox
economics
acute tox
cytotoxicity
mutagenicity
cell-cell communication
DNA damage
receptor binding
Office of Research and Development
National Center for Computational Toxicology
46


EPA
Toxicity Modeling Problem
Toxicity Risk Assessment
in vivo response - disease
in vitro response - bioassays
biofunctional response
SAR
structure-activity relationships
increasing uncertainty
Office of Research and Development
National Center for Computational Toxicology
47

Endpoints Can be Triggered by Multiple Mechanisms



- A test may only predict one specific mechanism
- False negative rate may be high unless multiple mechanisms are tested
- Implies: A negative result is not a clean bill of health

Office of Research and Development
National Center for Computational Toxicology

National Academy of Sciences Report (2007)

Toxicity Testing in the Twenty-first Century: A Vision and a Strategy

NAS PANEL SEEKS MAJOR SHIFT IN HOW EPA ASSESSES CHEMICALS' TOXICITY

Date: June 22, 2007 -

A National Academy of Sciences (NAS) panel is calling for a major shift in how EPA assesses chemicals' toxicity, recommending that the agency base its toxicological research and regulatory processes on how substances affect biological pathways -- which send information within and between cells -- rather than so-called health endpoints, such as cancer.

The new studies envisioned by the panel would evaluate chemicals' effects on biological processes using cells or cell lines, preferably human, to examine how they react to exposure to different substances. Rather than focusing research and basing regulations on endpoints, such as a substance's apparent ability to create tumor cells or harm brain development in fetuses, EPA should center toxicity testing around "the perturbations in toxicity pathways that are expected to lead to adverse effects," the report says.

"In this framework, the goals of toxicity testing are to identify critical pathways that when perturbed can lead to adverse health outcomes and to . . . understand the effects of perturbations on human populations," says the report, *Toxicity Testing in the Twenty-first Century: A Vision and a Strategy*.

Inside EPA

Nuclear Receptor Affinity Profiling

ToxCast Apr. 2006.sdf

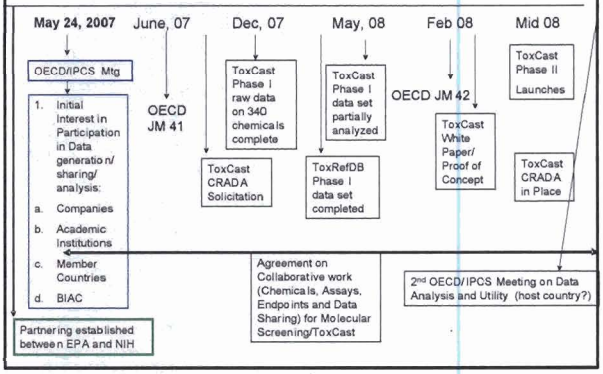
Method to the Virtual Madness

a) Dock multiple NR targets
b) Data: obtain affinity fingerprint matrix
c) Analysis: Hierarchical cluster that is chemical or target specific in heat map format
d) chemoinformatic profiling of discrete clusters for functional groups and predicted ADME properties
e) Data to knowledge - rational approach to determine promiscuous (red) versus non-binding (green)
f) Knowledge to wisdom - assign affinity threshold (i.e. Log Ki = -6 or lower = 1, else 0) and create a linkage map (with no edge weighting) of targets and their respective ligands in target space and chemical space. Determine similarity of targets via proximity within this specific high confidence (i.e. tight binding threshold) approach, and find multi-target specific signatures or unique signatures for QSAR generation or signature mapping.

Office of Research and Development
National Center for Computational Toxicology

Michael-Rock Goldsmith, Molecular Modeling Group, NCCT/USEPA

International Coordination Via OECD Molecular Screening Project



ToxCast Toxicity Reference Database

Summary of 53 Chronic/Cancer Rat/Mouse Studies Entered into ToxRefDB			
Total Effects Entered		1452	
Unique Effects		317	
Effect Type		**Endpoint Summary**	
Body Weight	166	Studies achieve LOAEL	40
Organ Weight	205	**Endpoint Dose Range (mg/kg/day)**	
Clinical Chemistry	85	<=10	9
Hematology	116	>10 and <=100	20
Non-neoplastic Pathology	333	>100 and <=1000	10
Clinical Signs	66	>1000	1
Effect Target		**Effects at Dose Range (mg/kg/day)**	
Liver	298	<=10	274
Kidney	70	>10 and <=100	536
Testes	49	>100 and <=1000	561
Thryoid Gland	14	>1000	49

Office of Research and Development
National Center for Computational Toxicology