

ACToR Aggregated Computational Toxicology Resource

Richard Judson judson.richard@epa.gov

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY



**COMPUTATIONAL
TOXICOLOGY**

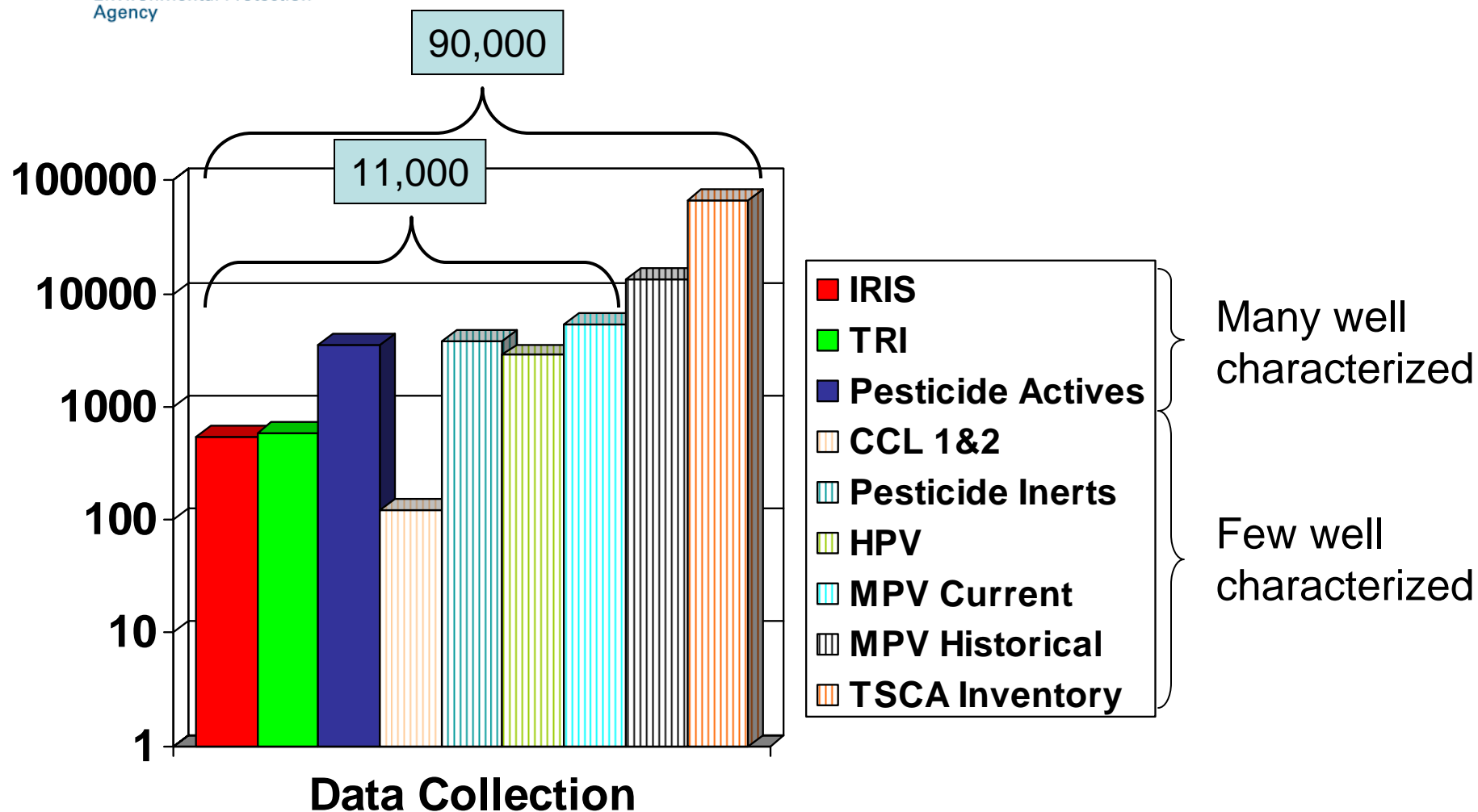
Outline

- The World of Environmental Chemicals
- Data Management / Bioinformatics
 - ACToR
- Chemical Screening and Prioritization
 - ToxCast
 - Statistical Approaches

Goals of Computational Toxicology at the EPA

- Screening and Prioritization of Environmental Chemicals
 - Many widely used (~10,000-30,000)
 - Few completely characterized (1,000-2,000)
 - Need efficient strategy to prioritize chemicals for detailed testing
- Model / characterize chemicals for toxicity
 - Hazard ID
 - Mode of action
 - Mechanism of action
 - Exposure Risk
 - Dose response

The World of Environmental Chemicals



The Toxicity Data Gap

Complete toxicity package
would include:

- Acute
- Subchronic
- Chronic / Cancer
- Developmental
- Reproductive
- Immunotox
- Neurotox
- Genotox
- Dermal
- Respiratory
- Nephrotox
- Endocrine
- Cardiotox
- Hepatotox
- Ecotox

Pesticide Food Use
Actives require majority
of these
~\$20M/chemical

TSCA substances (most
of the rest) require a
small subset

TSCA Requirements / SIDS

- Acute Toxicity
- Chronic Toxicity
- Developmental and Reproductive Toxicity
- Mutagenicity
- Ecotoxicity
- Environmental fate

1998 EPA OPPT Baseline study results for 2863 HPVs

- 43% had no data
- 7.1% had complete data

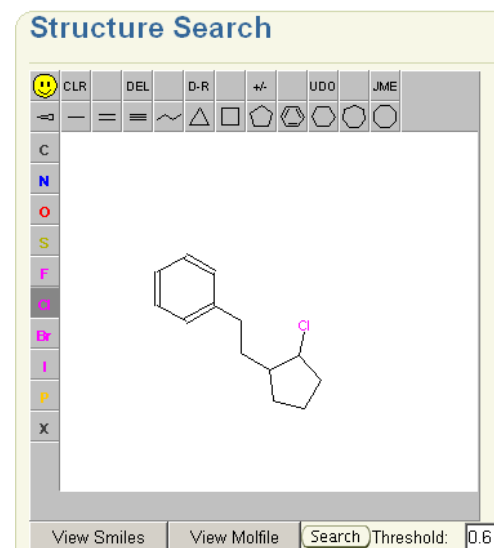
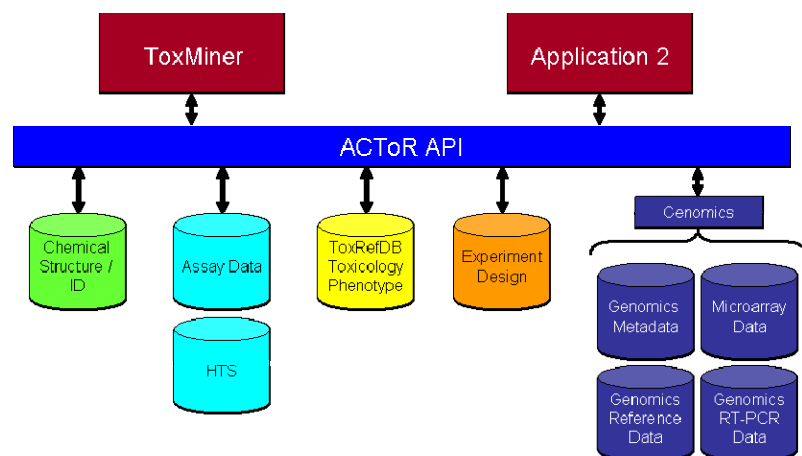
Big Questions

1. How do we catalog data for all environmental chemicals?
2. Is there better data coverage today?
3. How do we efficiently close the toxicity information gap?

ACToR

Aggregated Computational Toxicology Resource

- Aggregating the world's chemical structure, bioassay and toxicology data for environmental chemicals
 - 200 sources of data – all public
 - Over 15M chemicals
- Will manage all ToxCast Data
- Public release planned for 2008
- Prototype Intranet: <http://134.67.216.45:22722/servlet/ActorPrototype13>



ACToR Data Types

- **Data Collection** – Set of Substances
- **Substance** – Chemical that is tested
- **Compound** – Chemical with a structure
- **Generic Chemical** – Aggregation of all substances with same CAS
- **Assay** – Collection of tabular data for a data collection

ACToR Assay Types

- PhysicoChemical (logP, MW)
- Biochemical (Ki - ToxCast, PubChem)
- Cellular (Cytotoxicity - ToxCast, PubChem)
- Tissue (Tissue slice assay)
- In vivo toxicology
 - Tabular – primary (NTP, OPPIN)
 - Tabular – secondary (IRIS)
 - Summary calls (Scorecard, CalEPA)
 - Summary report via URL (INCHEM)
- Category (from OPPT, Health Canada)
- Regulation (TSCA, FIFRA)
- Description (IUR Usage Levels)

Each ACToR Data Collection Aggregates Data from Global Sources

Data Collection | ACToR | US EPA - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address: http://134.67.216.45:22722/servlet/ActorPrototype13?page=17&id=205_1_110111010_10

Google Go Links

Favorites: Add... Organize

- ACToR
- Backup
- Bioinformatics
- Chemical Lists
- Cheminformatics
- Conferences
- cvs
- DB
- EPA
- EPA Links
- Gene Data
- Graphs
- GTF
- Java
- Links
- Meetings
- Papers
- perl
- Personal
- PostDoc
- Software Apps
- Statistics
- Systems Biology
- ToxCastSim
- Toxicology
- ACToR dev11
- ACToR dev12
- ACToR dev13
- Chemistry Devel...
- EPA@Work The ...
- Google
- http--stb.benef...
- http--www.xemi...
- LMC Chemical In...
- MSN.com
- NCCT-NCC Jour...
- NetForge 1.0 M...
- Nuclear Recepto...
- ORD Call Center...
- phpMyAdmin 2.9...
- Radio Station Gu...
- Renaissance Co...
- Science Connect...
- Scientific Calcula...

U.S. Environmental Protection Agency

ACToR: Aggregated Computational Toxicology Resource

Recent Additions | Contact Us Search: All EPA This Area

You are here: [EPA Home](#) » [ACToR](#) » Data Collection

Data Collection: ToxCast_320

Name: [ToxCast_320](#)

Description: ToxCast Main Phase I chemical set

ID: 205

Source Type: ToxCast List

Number of Substances: 320

Number of Generic Chemicals: 307

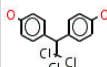
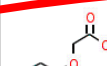
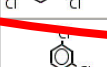
Compilation Date: TBD

Compilation Instructions: TBD

• [Hide Chemical Table](#)

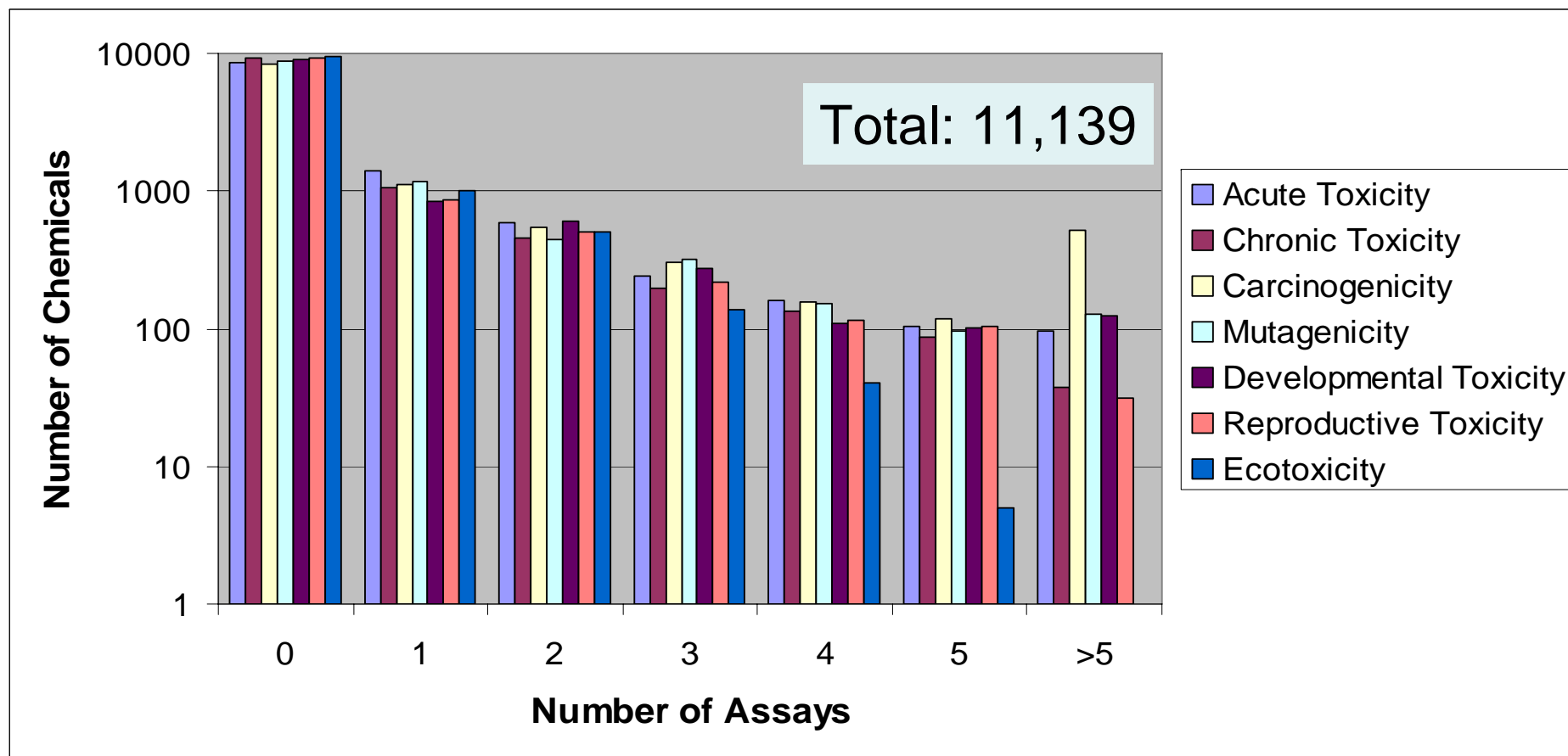
Page 1 of 7 : [Next](#)

Select Link to Toggle Table Section: [Thumbnails](#) : [SID](#) : [Source_SID](#) : [GCID](#) : [CASRN](#) : [Name](#) : [Categories](#) : [Phenotypes](#) :

Structure	SCID	GCID	CASRN	Name	Hazard	AcuteTox	SubchronicTox	ChronicTox	Carcinogen	GenTox	DevTox	ReproTox	NeuroTox	DevelTox	ImmunTox	DermatTox	RespiratoryTox	NephroTox	HepatTox	Endocrine	CardioTox	EcoTox	FoodSafe	ToxOther
	11549	6528	2971-36-0	HPTE																				
	11550	431	94-75-7	2,4-Dichlorophenoxyacetic acid	6	5	1	6	13	6	7	5	4		3	2			1	2	1	4		
	11551	6372	94-82-6	4-(2,4-Dichlorophenoxy)butyric acid (2,4-DB)	3	4	1	4	7	4	6	5	2		2				1			2		

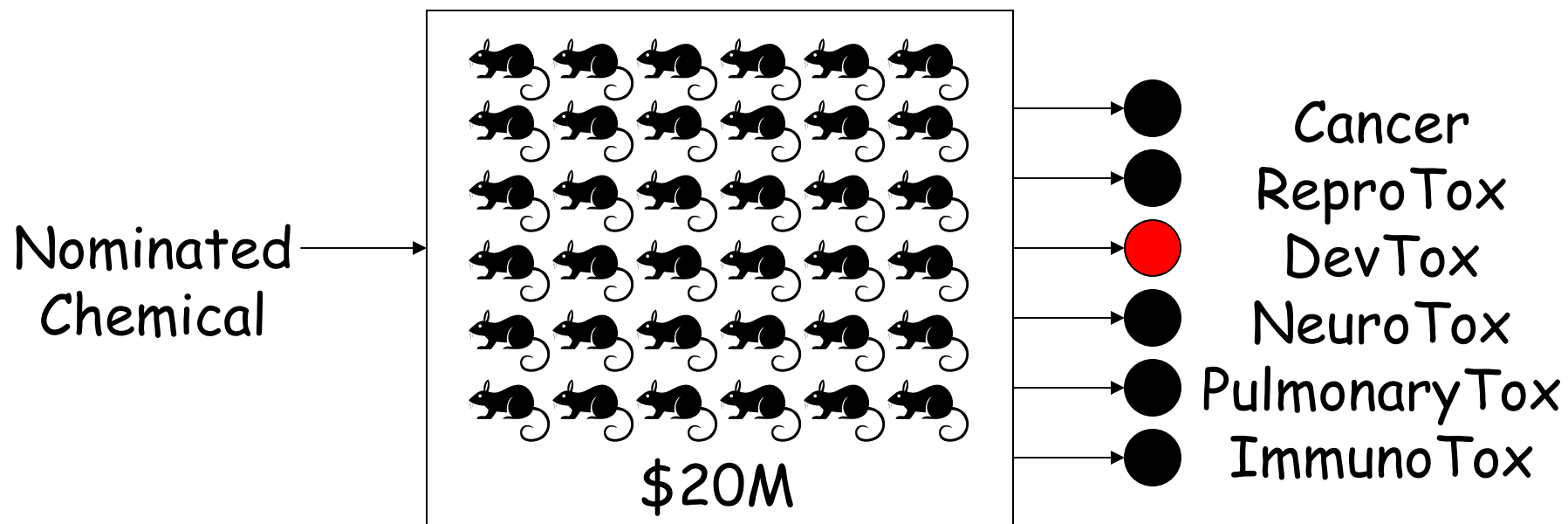
Data Source
CPDBAS_DSSTOX
HPVCSI_DSSTOX
IRISTR_DSSTOX
NCTRER_DSSTOX
NTPBSI_DSSTOX
NTPHTS_DSSTOX
ToxCast_320
CERCLA
EDC73
EPA_DWC
EXTOXNET
HPVChallenge
HPV
INCHEM_EHC
INCHEM_EHC
INCHEM_IARC
ITER_TERA
IUR_2002
IUR_86_02

What is Current Data Coverage in ACToR for Widely-used Environmental Chemicals?

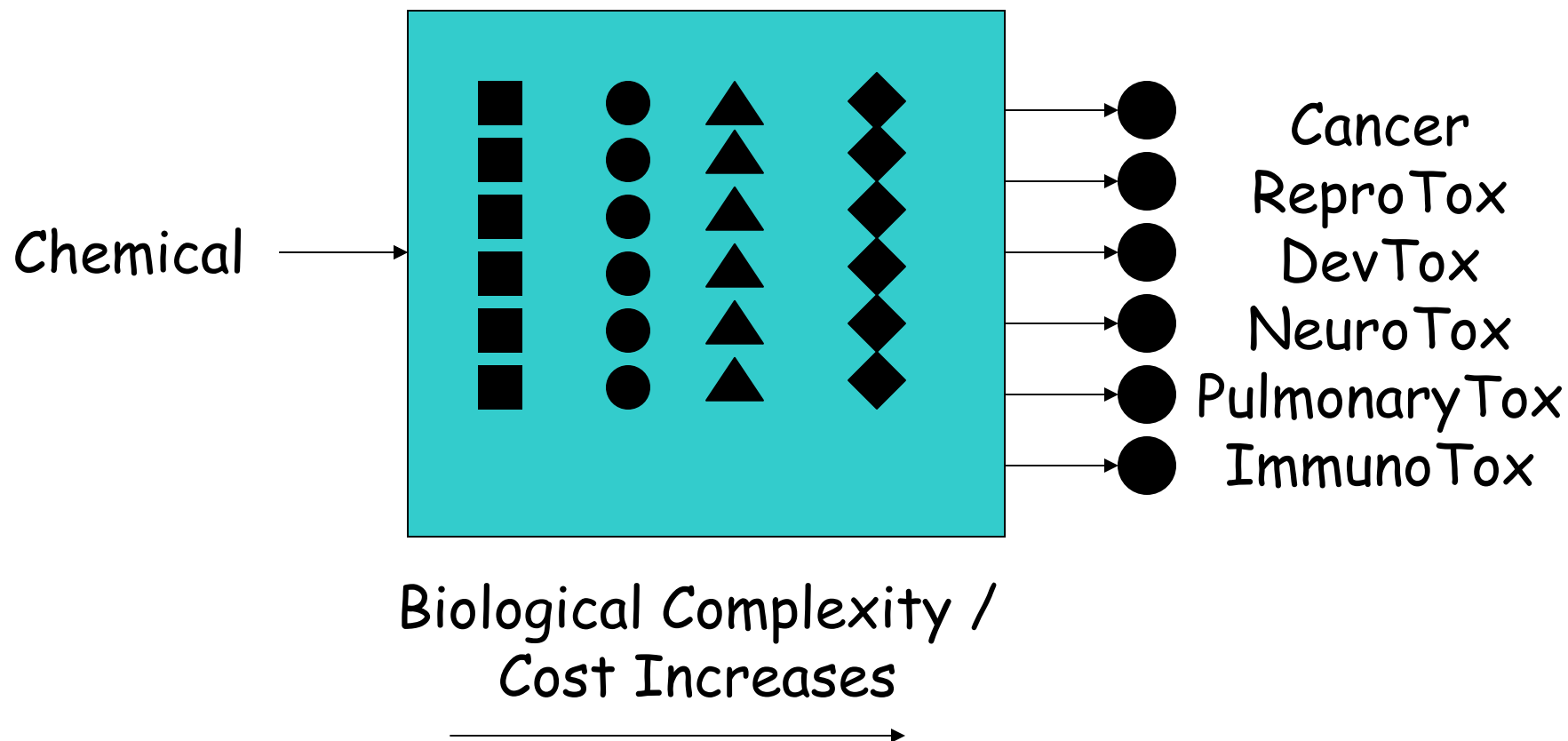


<10% have data for most tests

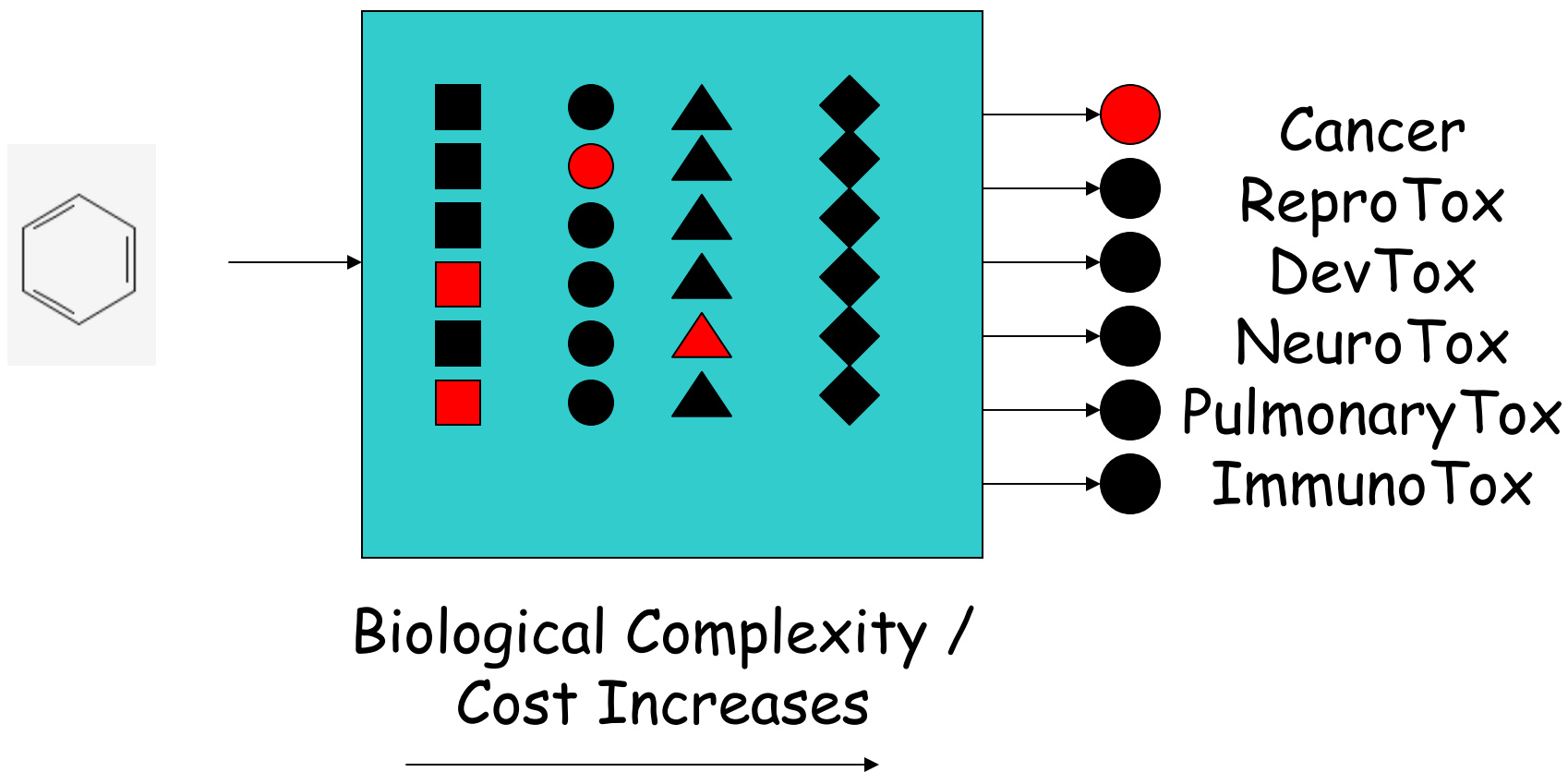
Toxicity Prediction Today



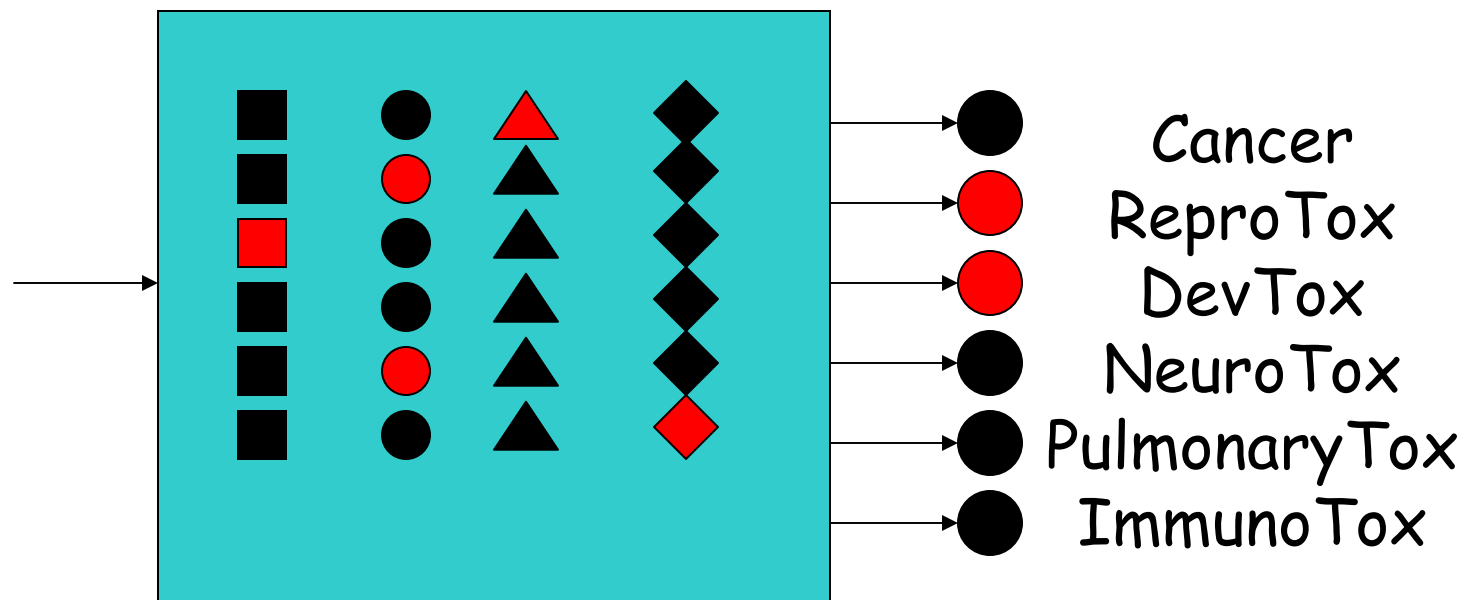
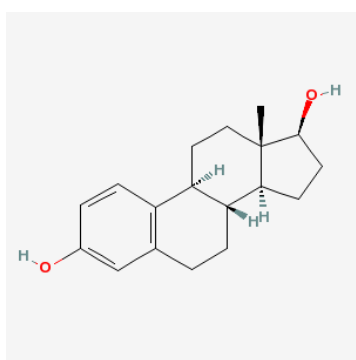
Toxicity Prediction Tomorrow: Find Pattern that Predicts Toxicity using Inexpensive Assays



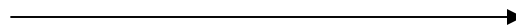
Toxicity Prediction Tomorrow: Find Pattern of Assays that Predicts Tox



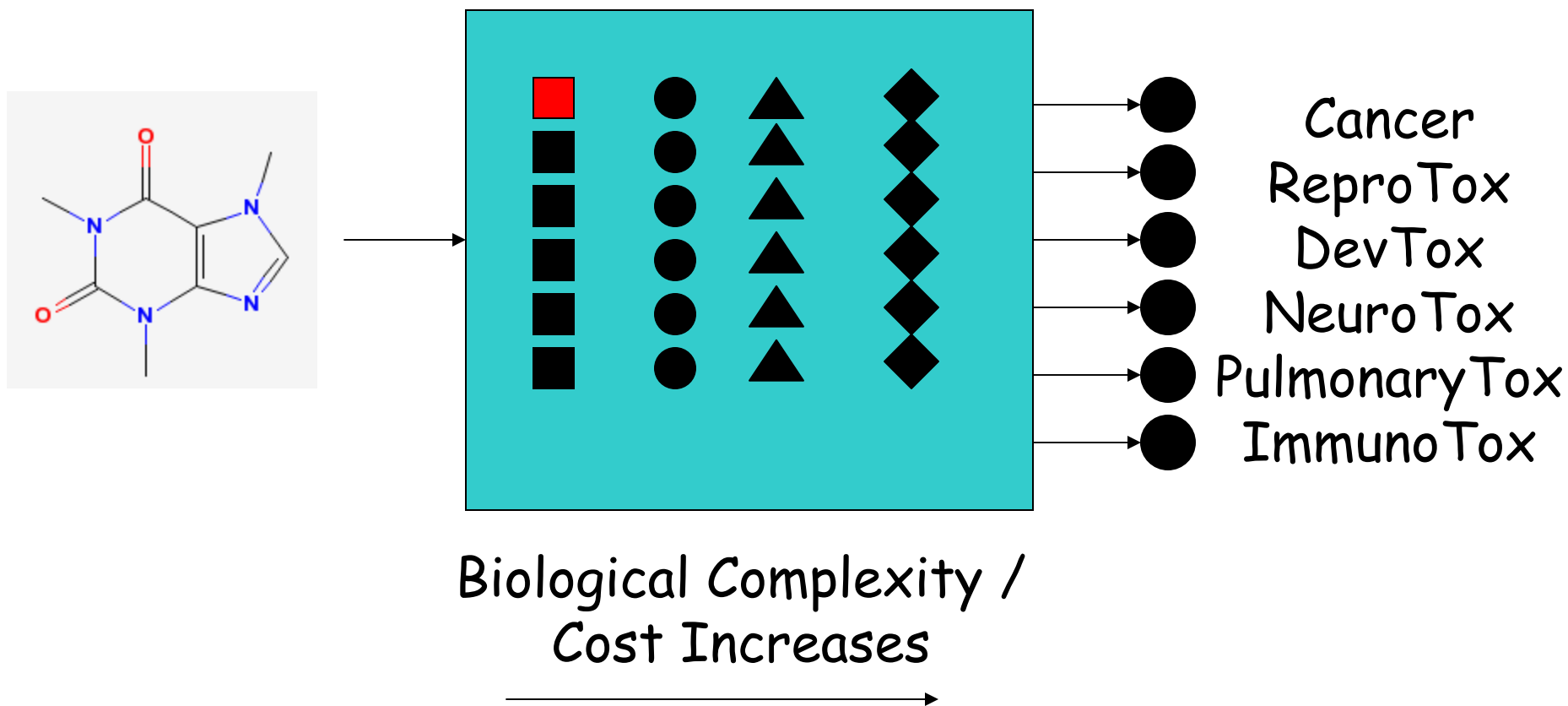
Toxicity Prediction Tomorrow: Find Pattern of Assays that Predicts Tox



Biological Complexity /
Cost Increases



Toxicity Prediction Tomorrow: Find Pattern of Assays that Predicts Tox



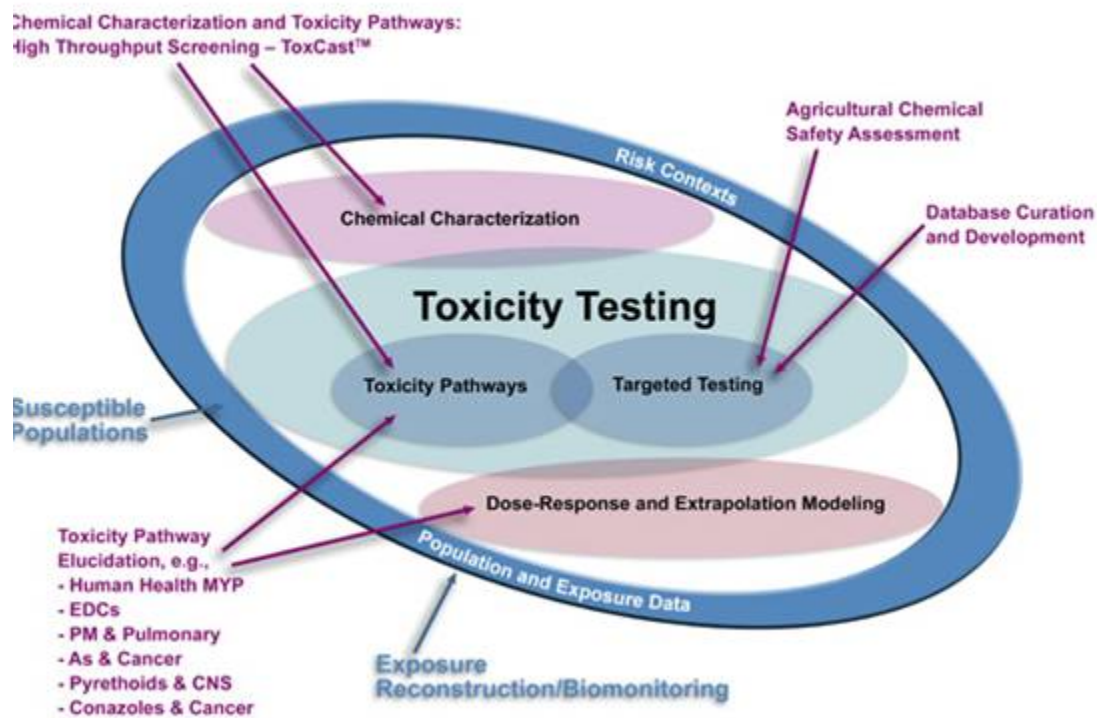
In Vitro Signature Approach Supported by NRC

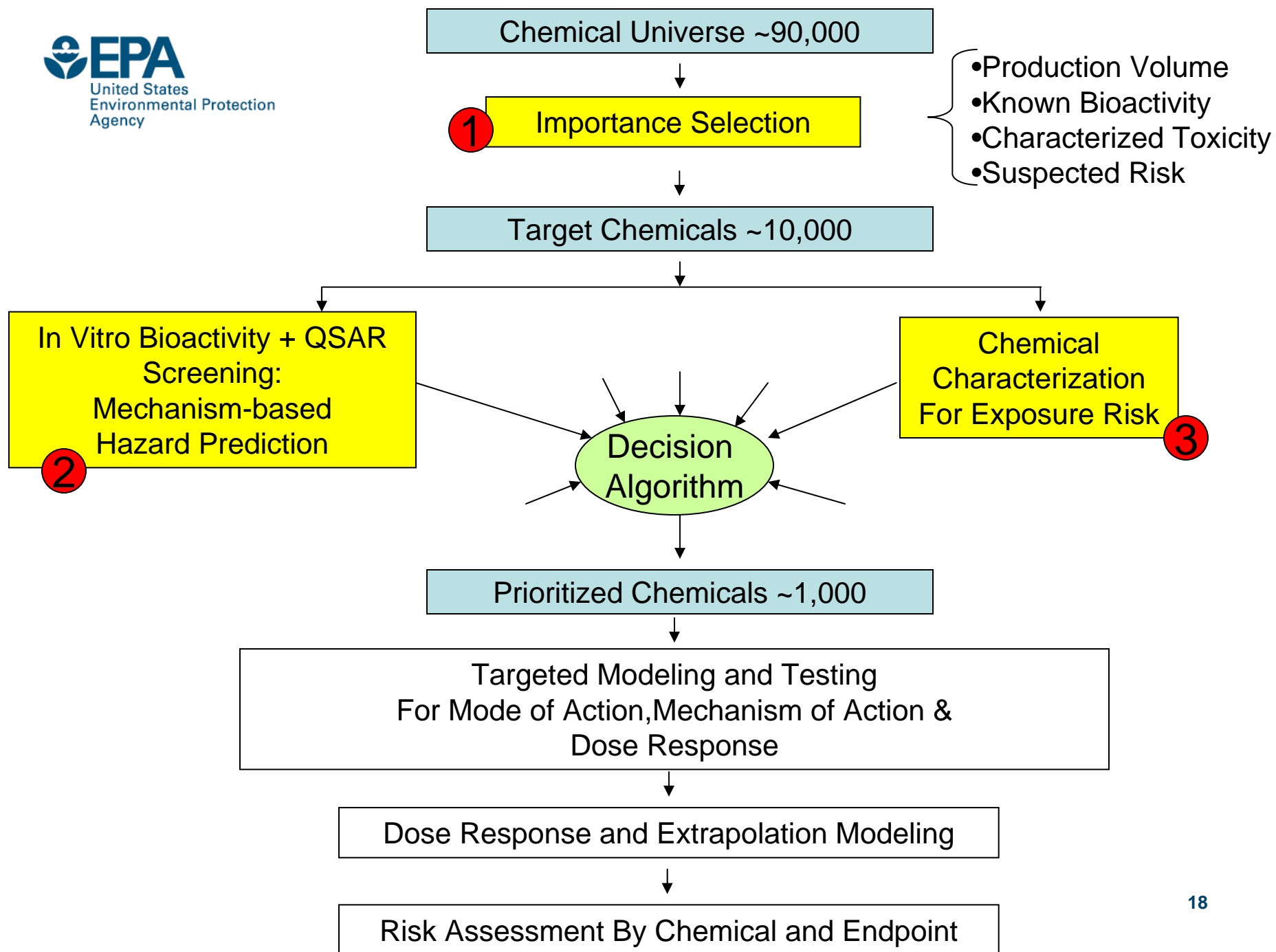
PREPUBLICATION COPY

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

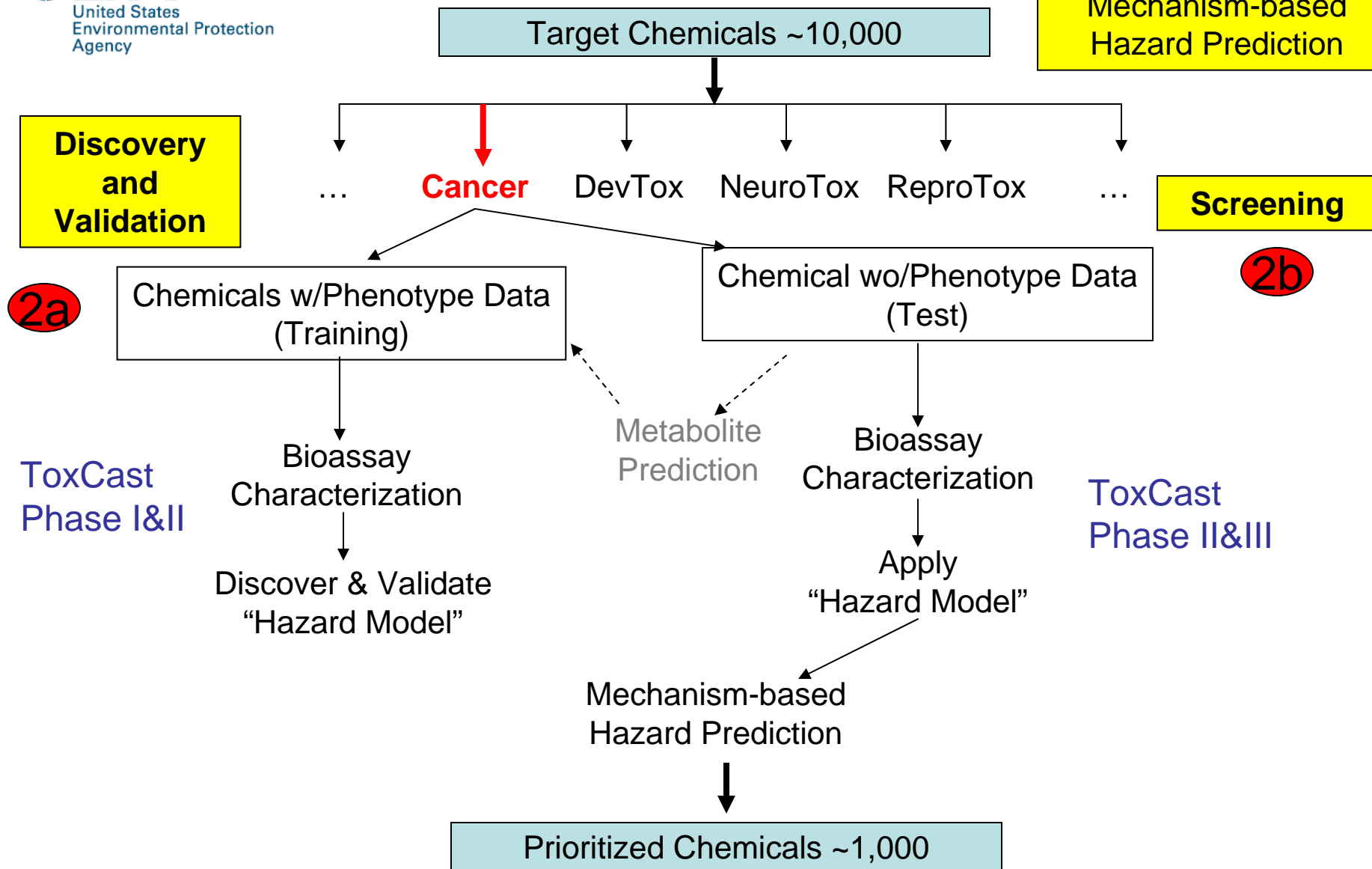
Committee on Toxicity Testing and Assessment of Environmental Agents
Board on Environmental Studies and Toxicology
Institute for Laboratory Animal Research
Division on Earth and Life Studies
NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu





2 In Vitro Bioactivity +
QSAR Screening:
Mechanism-based
Hazard Prediction

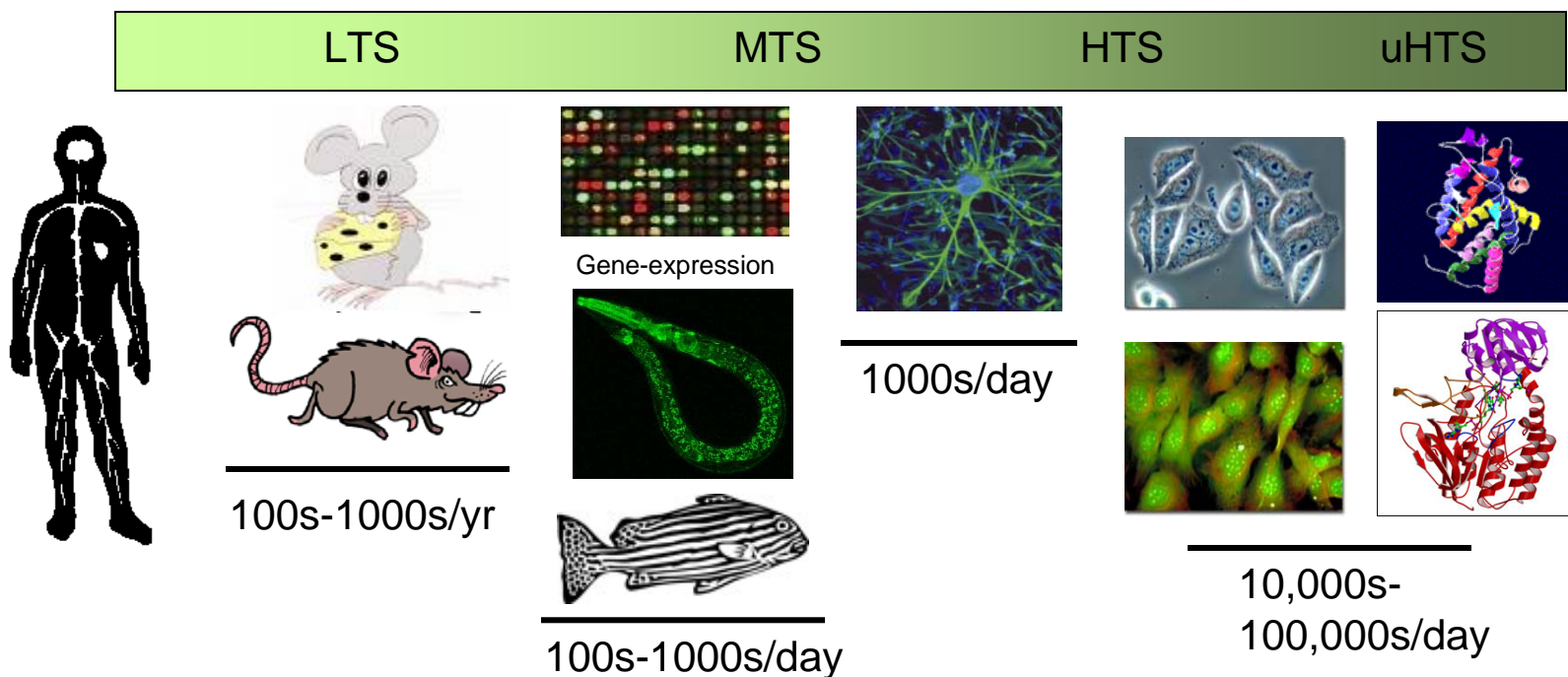


In Vitro Bioactivity Screening Needs

- In vitro assays mechanistically linked to phenotypes
 - ToxCast Assays
- Phenotype data for biomarker discovery and validation
 - High quality
 - Quantitative
 - Tabular
 - ToxRefDB / DSSTox / ACToR
- Statistical Data Mining Approaches
 - ToxMiner
- Mechanistic models to link bioassay data with risk assessment
 - Virtual Liver

High-Throughput Screening Assays

*batch testing of chemicals for pharmacological/toxicological endpoints
using automated liquid handling, detectors, and data acquisition*



Human Relevance/
Cost/Complexity

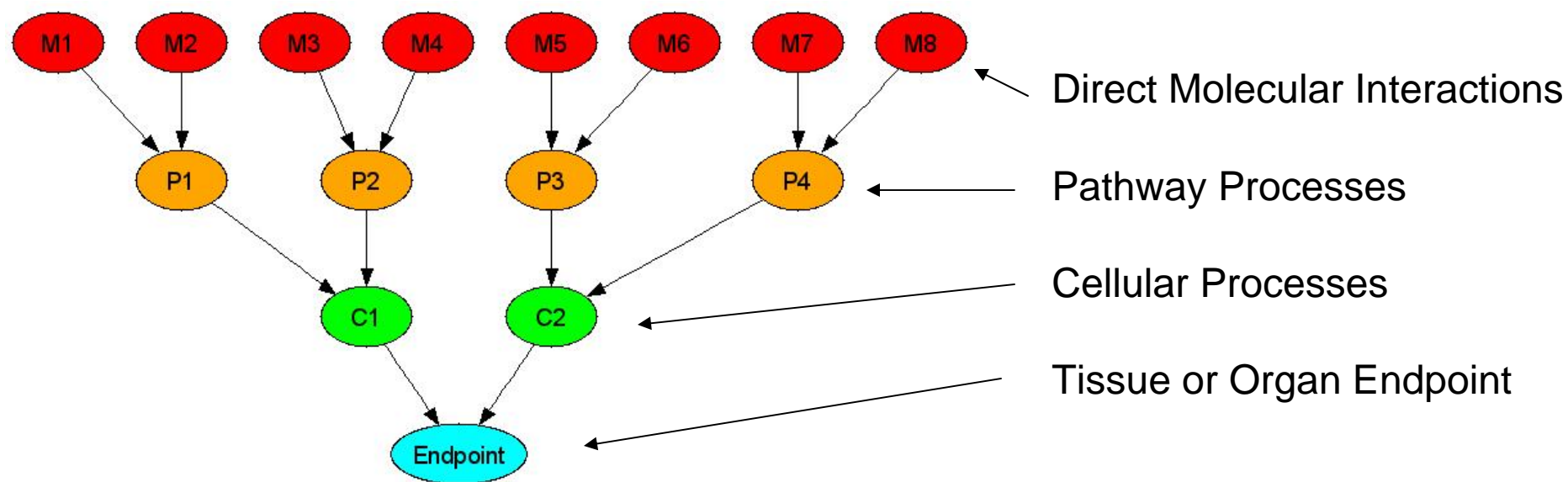
Throughput/
Simplicity

22

Simulating ToxCast v/ Testing ToxMiner

- Goals
 - Imagine what “real” data will look like
 - Develop methods to integrate different types of data
 - Evaluate performance of different classifier algorithms
- Simulation Properties
 - Multiple levels of biology
 - Multiple mechanism can cause same phenotype
 - Include noise, missing data, irrelevant features or input variables

ToxMiner – Predictive Modeling

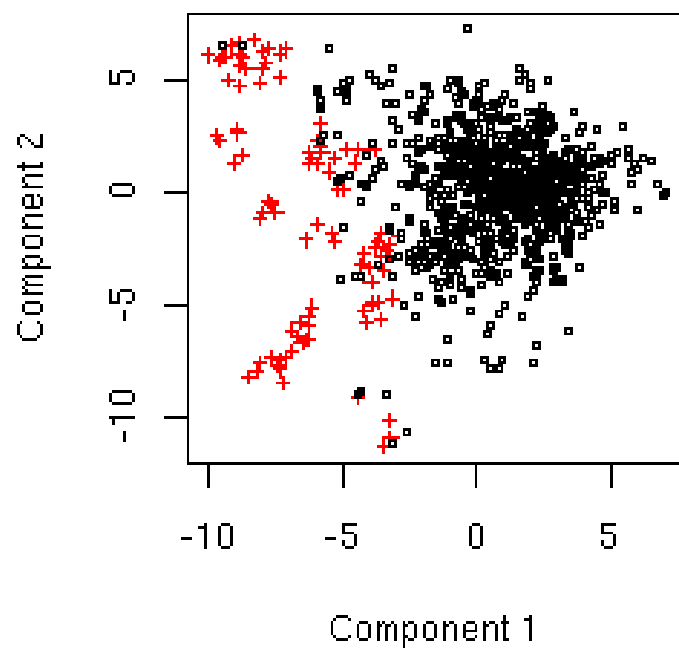


Evaluate algorithms as $f(\text{sample size, model complexity, noise})$

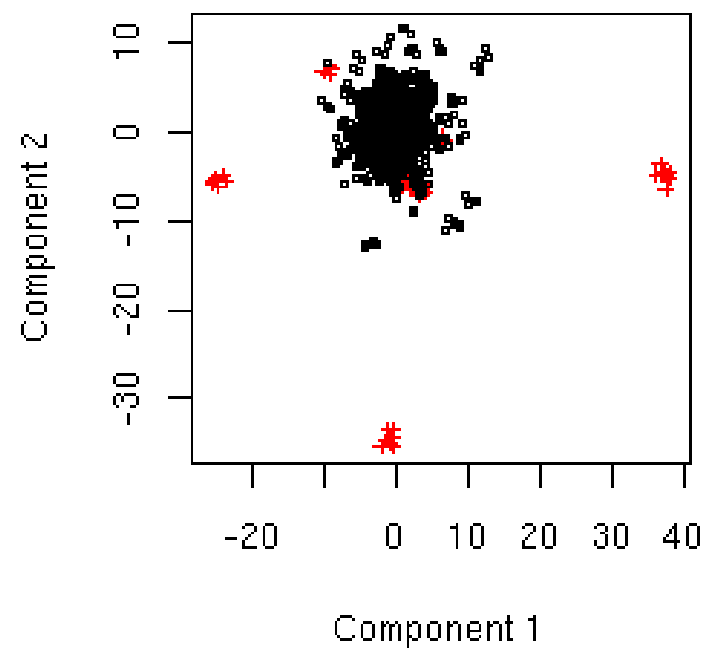
300 chemicals, 300 assays

Chemicals Cluster Based on Assay Data

S1 - True Features



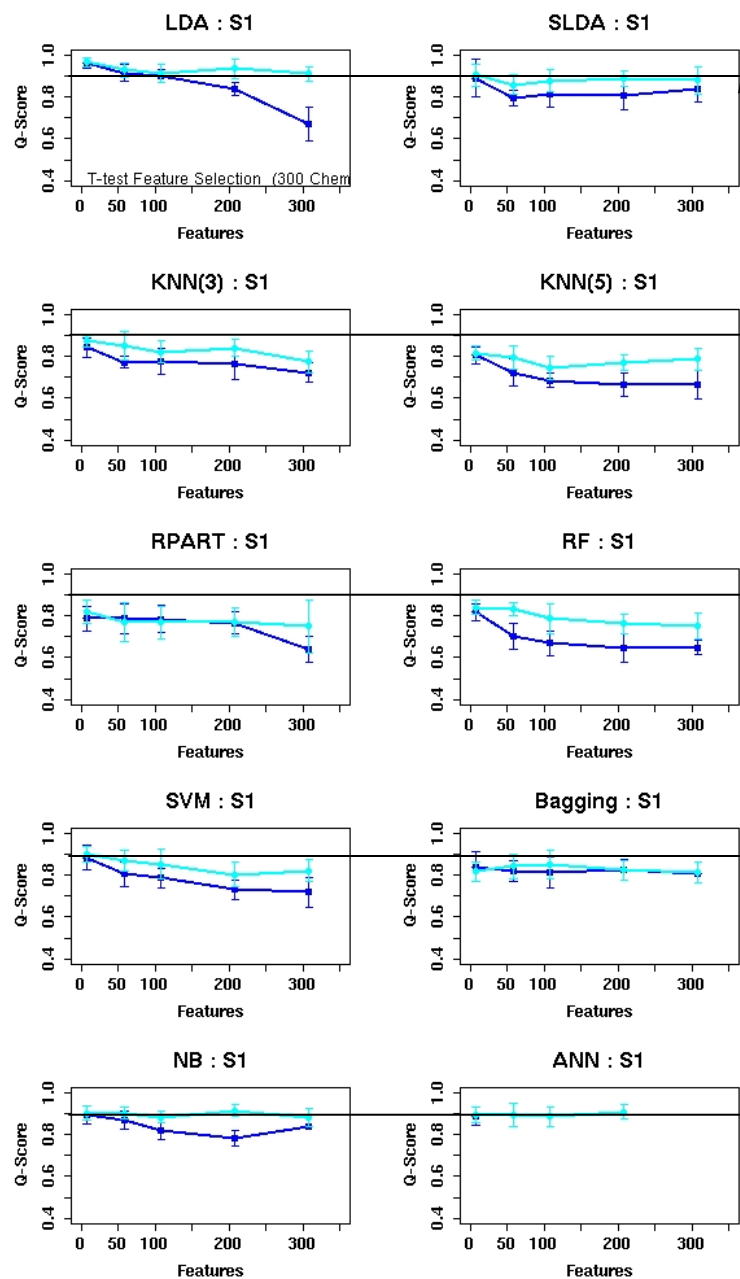
S1 - All Features



Red: phenotype+
Black: phenotype-



Evaluate Classification Algorithms



Metric=Q-score

Average of sensitivity and specificity

Q=0.90

minimum 80% sensitivity and specificity

Computational Toxicology Challenges

- Compile more data
 - Aggregate existing databases
 - Mine the literature
 - Get research groups to compile and submit (i.e. PubChem)
- Generate targeted data
 - ToxCast
- Extrapolate computationally from know to unknown
 - QSAR – structure-based extrapolation
 - QBAR – bioactivity-based extrapolation
 - Modeling – PBPK to multi-scale cell and organism models
 - Understand limits to extrapolation (domain of applicability)

Acknowledgements

- ToxCast
 - David Dix
 - Keith Houck
 - Bob Kavlock
- DSSTox
 - Ann Richard
 - Marti Wolf
- ACToR
 - Tom Transue
 - Tommy Cathey
 - Richard Spencer
 - Fathi Elloumi
- ToxRefDB
 - Matt Martin



NCCT ToxCast™ Team
August 2007