Eurotox, Knudsen Oct 8, 2007

NCCT-07068

EPA

Comparative Bioinformatics: Applications for Developmental Toxicology

Thomas B. Knudsen, PhD
Eurotox Congress, October 8, 2007

Office of Research and Development
National Center for Computational Toxicology

October 18, 2007

EPA

Computational Toxicology

- **Computational Toxicology**: the application of mathematical and computer models to predict adverse effects of chemical exposure and to understand the mechanism(s) through which a given chemical induces harm
- **National Center for Computational Toxicology**: "… to integrate modern computing and information technology with molecular biology to improve Agency prioritization of data requirements and risk assessment of chemicals"

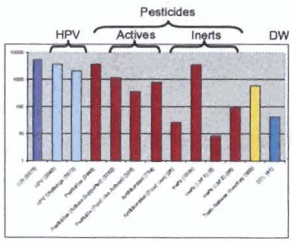
http://www.epa.gov/ncct

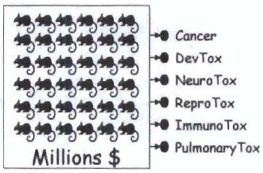
Office of Research and Development
National Center for Computational Toxicology

1

EPA

Problem:
too many chemicals, too high cost, and much biological uncertainty



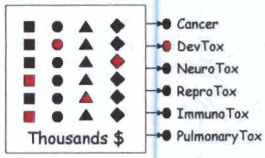


Office of Research and Development
National Center for Computational Toxicology

2

EPA

ToxCast™: research program to build an open data-rich system to screen, classify and rank chemicals for further evaluation



cell-based phenotyping, toxicogenomics, high-throughput assays, high-content screening and chemical procurement

Office of Research and Development
National Center for Computational Toxicology

http://www.epa.gov/comptox/toxcast

3

EPA

Phased development of ToxCast

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	~10K	Data-poor	Prediction and Prioritization	$6-10k	FY10-12

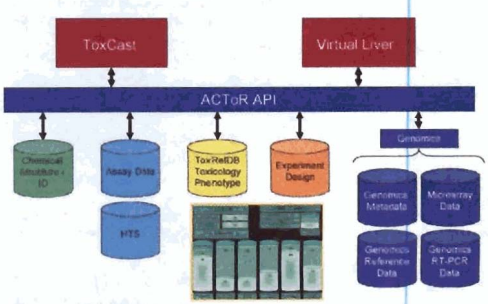
- deliver an affordable, science-based system for categorizing chemicals
- confidence for prioritization will rise as database grows
- predict and identify potential modes of action
- refine and reduce animal use in hazard identification and risk assessment

Office of Research and Development
National Center for Computational Toxicology

4

EPA

Aggregated Computational Toxicology Resource (ACToR)

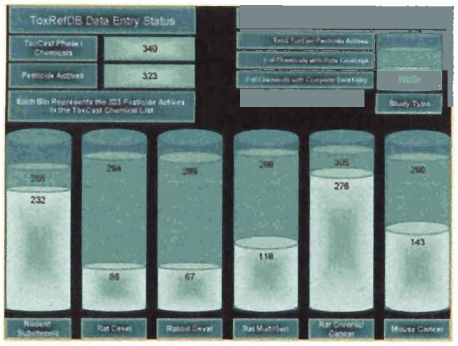


Source: Richard Judson, NCCT

Office of Research and Development
National Center for Computational Toxicology

5

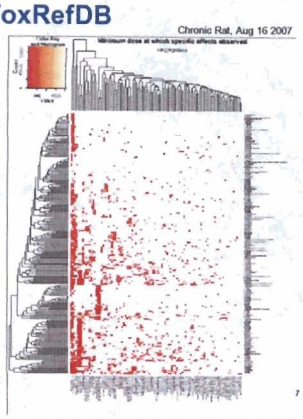
Digitizing Legacy Data



Source: Matt Martin, NCCT

ToxRefDB

- multidimensional toxicity dataset to mine relationships between effects
- interoperability of toxicity data with TXG/HCS/HTP data to pattern by:
  - chemical
  - study type
  - effect (e.g., DTX)
  - species
  - dose
  - mode of action

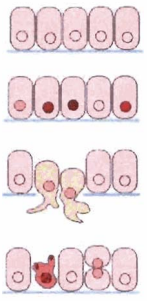


Systems-based approach: developmental toxicity (DTX)

- how do cells in a developing embryo integrate complex signals from the genome to make decisions about their behavior or fate …
- … and under what conditions are these decision-making systems susceptible to genetic defects or vulnerable to environmental perturbations?
- key properties of many biological systems are similar in concept to mechanical (engineered) systems

components	↔	device	↔	module	↔	system
genes		signals		networks		cell
cells		tissues		organs		organism

Prenatal Targets: mode of action



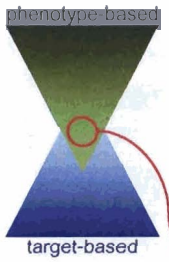
PROGRAMMED RESPONSES

CONDITIONED RESPONSES

TRANS-DIFFERENTIATION

GROWTH & DEATH SIGNALS

Research Strategy



Top-down: decompose system into component parts (basic entities)

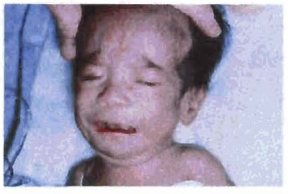
- Strength: independent of detailed molecular knowledge
- Weakness: not all critical components or interactions may be known

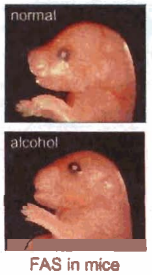
Bottom-up: logically assemble basic entities into functional system

- Strength: detailed molecular knowledge increasingly available
- Weakness: limited capacity to project perturbations onto the system

middle ground holds key to mode-of-action

Example: maternal alcohol intoxication during pregnancy

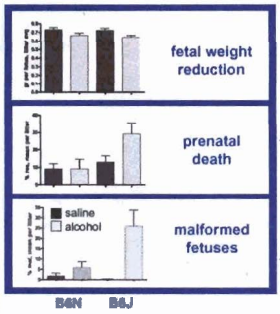




FAS in mice

- 13% mothers drink during pregnancy
- 1% of babies damaged by alcohol
- Fetal Alcohol Syndrome (FAS)
- can be triggered in mice (day 8)

Genetic susceptibility: evaluating alcohol-related birth defects in two pedigrees of C57BL/6 mice



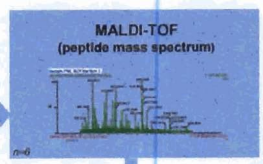
C57BL/6N = C57BL/6J

C57BL/6N < C57BL/6J

Source: Green et al. (2007) Develop Dynam 236: 613-631

12

Biomarkers: proteomic analaysis of amniotic fluid



Group (n=6)	Altered peaks	Classifiers
B6J control vs. B6J alcohol	3163.5 1495.8, 1369.6	3163.5 1495.8
B6N control vs. B6N alcohol	none significant	none significant
B6J control vs. B6N control	3163.5	3163.5 2802.3 2228.2

B6N B6J

13

Class-discriminating peptides

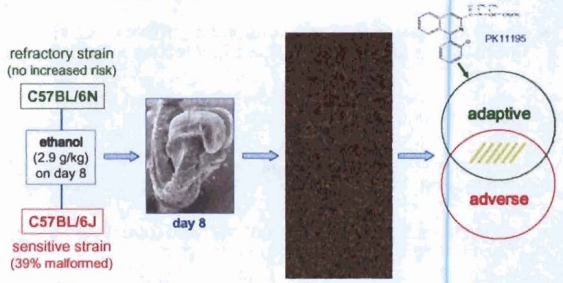
- origin of m/z peaks 1495.8 & 1369.6 from AFP (alpha-fetoprotein); confirmed by LC-MS/MS and *in silico* digest
- AF-AFP levels consistently low in the sensitive (B6J) pedigree on day 17 following maternal alcohol on day 8

  Source: Datta et al. (2007) Fetal Alcohol Syndrome (FAS) in mice detected through proteomics screening of the amniotic fluid (submitted)
- Low MS-AFP levels reported in 59% drinking moms who gave birth to a malformed baby [relative risk 2.46]

  Source: Halmesmaki et al. (1987) Prediction of fetal alcohol syndrome by maternal alpha fetoprotein, human placental lactogen and pregnancy specific beta 1-glycoprotein. Alcohol Suppl 1, 473-476

14

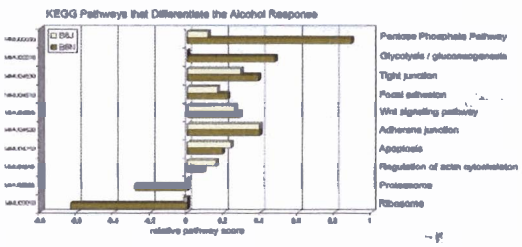
Mapping the FAS transcriptome: mouse embryonic headfold 3h after maternal alcohol exposure



Source: Green et al. (2007) Develop Dynam 236: 613-631

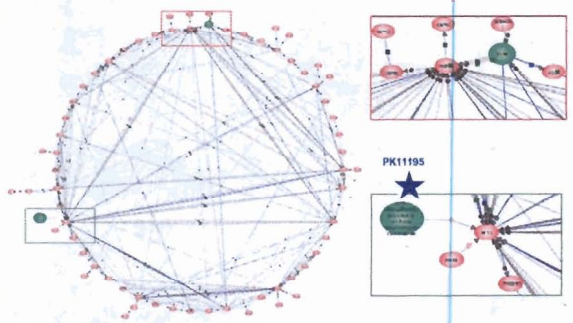
15

Signature Pathways: level 2 analysis of Gene Ontology based on KEGG



16

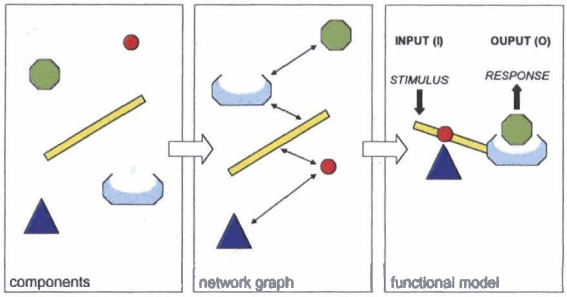
Predicted FAS gene association network



17

System models

Based on MW Covert (2006) Integrated regulatory and metabolic models. In: Computational Systems Biology, edited by A Kriete and R Eils, Elsevier Academic Press (page 194)



In the post-genomics world ...

... our ability to create mathematical models describing the function of biological networks will become just as important as traditional lab skills and thinking - D Butler (2001) Nature 409, 758-760

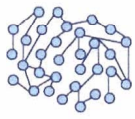
"Molecular biology took Humpty Dumpty apart ... mathematical modeling is required to put him back together again ..."

– Schnell et al. (2007) Am Sci 95:134

Information processing: computational systems modeling

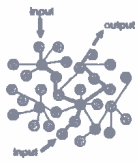
- cells may be viewed as an integration of two basic kinds of networks that direct the flow of molecular information:
- metabolic networks: mass-flow oriented; fast reactions driven by flux of metabolites (e.g., engines)
- regulatory networks: signal-flow oriented; information processing used to make adjustments (e.g., controllers)
- all systems require energy, control, and robustness to function: embedded in network topology

Network topology



Random Network

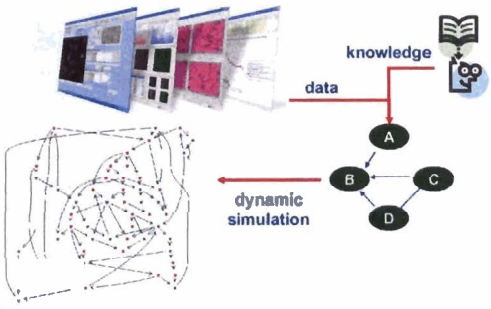
- in network theory, 'scale' refers to connectivity
- sample 32-node network randomly linked
- all nodes have low-degree of connectivity
  little or no modularity / robustness
- system readily falls apart when any node is lost



Scale-free Network

- same 32 nodes portrayed as a scale-free network
- most nodes have low-degree connectivity
- a few nodes (hubs) have high-degree connectivity
  modular organization of most nodes
- 'graceful degradation' with loss of a random node
  Achilles' heel = hubs; vulnerable to targeted attack

How to portray a bionetwork



Source: Imran Shah, NCCT

Summary: DTX systeome

GOAL: compendium of developmental toxicity pathways to apply system control theory and model ...

- altered flow of molecular regulatory information in the embryo (toxicity pathways)
- biological conditions under which perturbations invoke critical cellular effects (mode-of-action)
- connectivity between adverse and adaptive changes (cellular response networks)
- chemical prioritization for further evaluation and testing

Acknowledgements

University of Louisville

Birth Defects Center
Center for Genetics and Molecular Medicine
Center for Environmental Genomics and Integrative Biology
Clinical and Translational Sciences Institute
http://systemsanalysis.louisville.edu

Lab members
Amar Singh, MS
L. Bruno Ruest, PhD
Yihzi Zhang, MS-MD
Maia Green, MS
Reetu Singh, MS
Ken Knudsen, BChe
Caleb Bastian

Collaborators
Susmita Datta, PhD
Bill Pierce, PhD
Ken Ramos, PhD
Greg Rempala, PhD
Michele Pisano, PhD
Aldert Piersma, PhD [RIVM, The Netherlands]
Kathy Sulik, PhD [UNC Chapel Hill]

Grant sponsors
R01 ES09120 (NIEHS)
R01 AA13205 (NIAAA)
R21 ES13821 (NIEHS)
P30 ES014443 (NIEHS)

Acknowledgements

National Center for Computational Toxicology

http://www.epa.gov/ncct

Bob Kavlock
Jerry Blancato
David Dix
Hugh Barton
Elaine Hubal
Richard Judson
Imran Shah
Woody Setzer
Ann Richard
Rory Conolly
Tom Knudsen
Jim Rabinowitz
Karen Dean
Sandy Roberts
Carolyn Poe-Gurley
Keith Houck
Ramon Garcia
Mike Goldsmith
Steve Little
David Reif
Matt Martin
John Wambaugh
Maggie Zhao
Zhen Li
Keith McLaurin
Dan Rotroff
Kun Wang
Mike Breen
Clarynda Devane
Amar Singh (LHM)
Marti Wolf (LHM)

Disclaimer: the views expressed are those of the presenter and do not necessarily represent those of the U.S. EPA. No official Agency endorsement should be inferred