

Use of genomics data for population-level risk assessment

Joseph R. Shaw

The School of Public and Environmental Affairs

John K. Colbourne

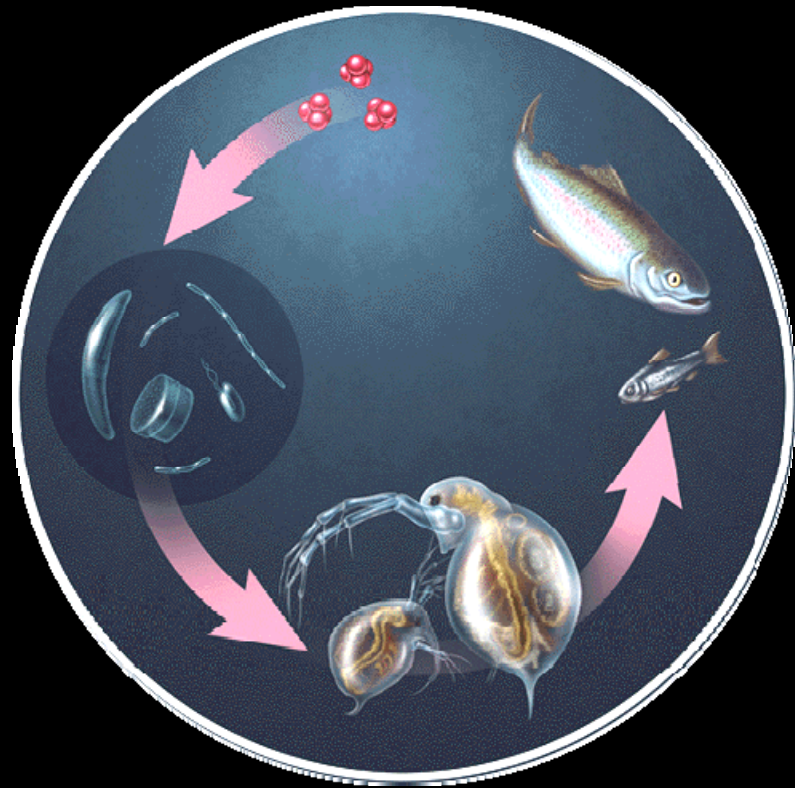
The Center for Genomics and Bioinformatics

Indiana University

Why use *Daphnia*

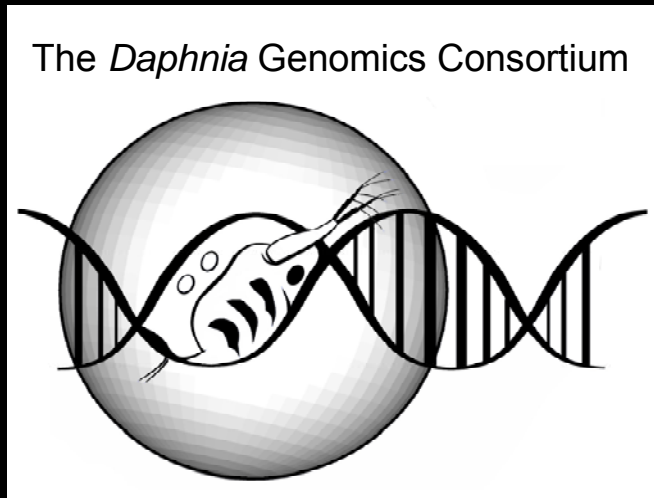
Daphnia is the most commonly used species in aquatic toxicity testing!

- ▶ easily studied in field
- ▶ easily cultured in the lab
- ▶ reproduction is clonal and sexual
- ▶ widely distributed in nature
- ▶ central to the ecology of many freshwater systems



Can genomics be used to assess population susceptibility to environmental contaminants?

The *Daphnia* genome project



2001: 9 researchers

2005: 100 research groups

2006: genome sequencing completed

2007: public pre-release

2008: 350 researchers, 17 countries

OUR GOAL: Develop the *Daphnia* system to the same level of molecular, cell and developmental understanding as other model species but with the added advantage of being able to interpret observations in the context of natural environmental challenges.

BIOCOMPLEXITY Project: Define and catalog the majority of the transcriptome of *D. pulex* and use this information to develop and produce high content microarrays to pursue toxicological genomic investigations.

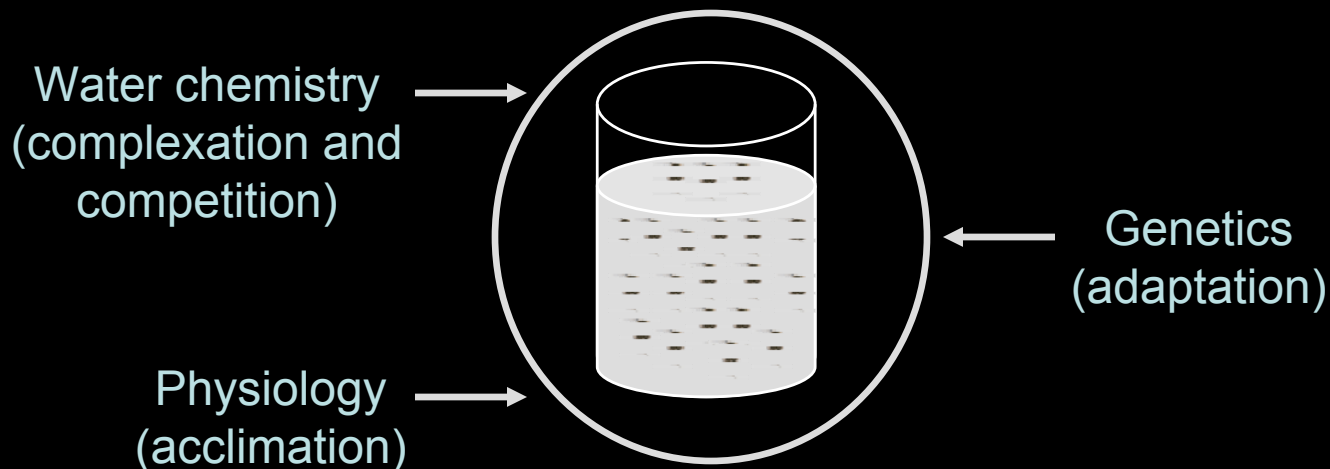
Tools for understanding gene function

- Sequenced cDNA libraries – Transcripts of genes expressed by *Daphnia* at different life-history stages exposed to 25 separate ecological and toxicological stressors
- Version 3 transcriptional microarray – 10,000 genes represented by 70 bp fragments mechanically printed on glass
- Genome tiling path arrays – 4.2 Million overlapping 70 bp fragments representing all the sequenced genome chemically synthesized on glass
- Version 4 transcriptional microarray – 12 experiments performed at a time on a single glass slide, interrogating all genes represented by 135,000 probes

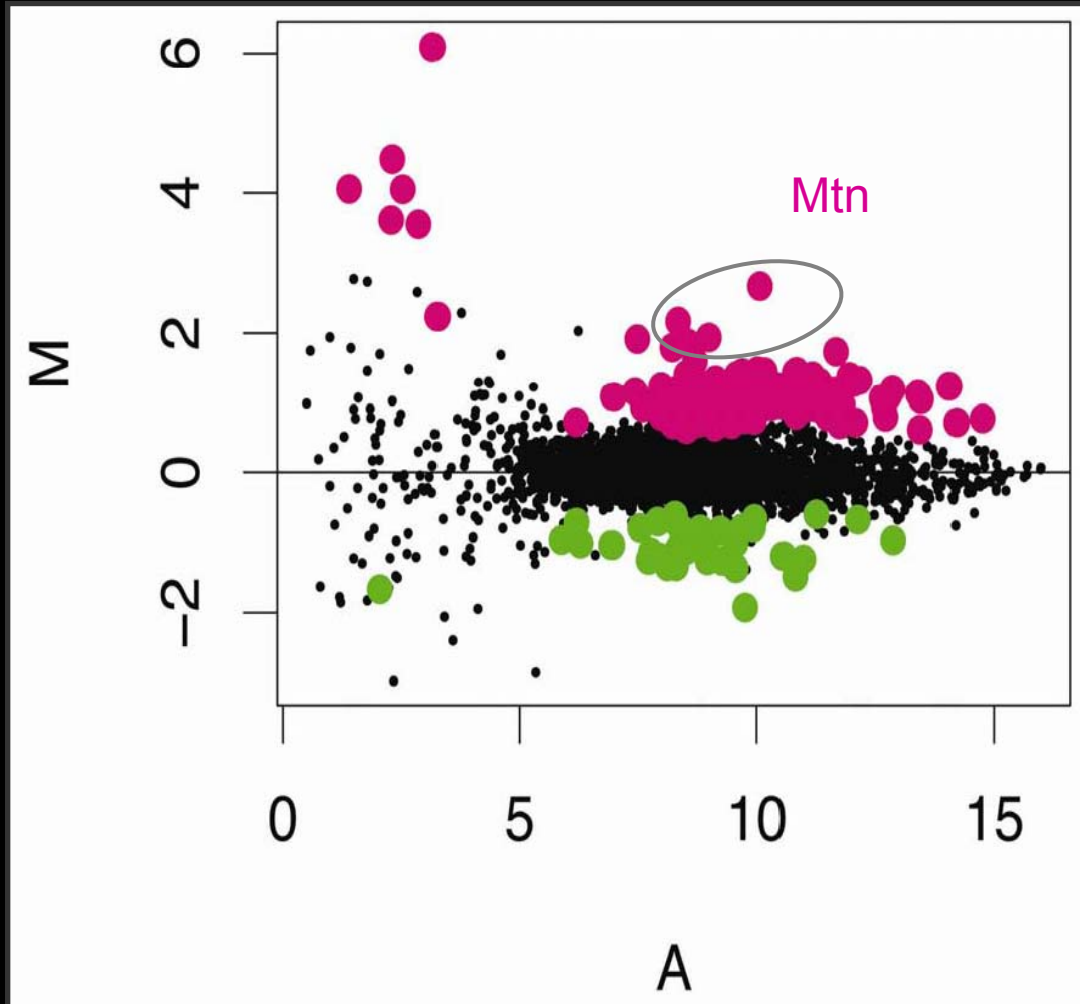
Variation in toxicity

Our laboratory range: 37 to 84.9 $\mu\text{g Cd/L}$

The literature range: 22 to 350 $\mu\text{g Cd/L}$



Cadmium effects on gene expression



Top 13 up-regulated genes

- Unknown (13)
- Opsin
- Myosin
- Metallothionein (Mtn)
- Hemoglobin
- Glutathione-s-transferase
- Flavoprotein
- Cuticle proteins
- Chitinase

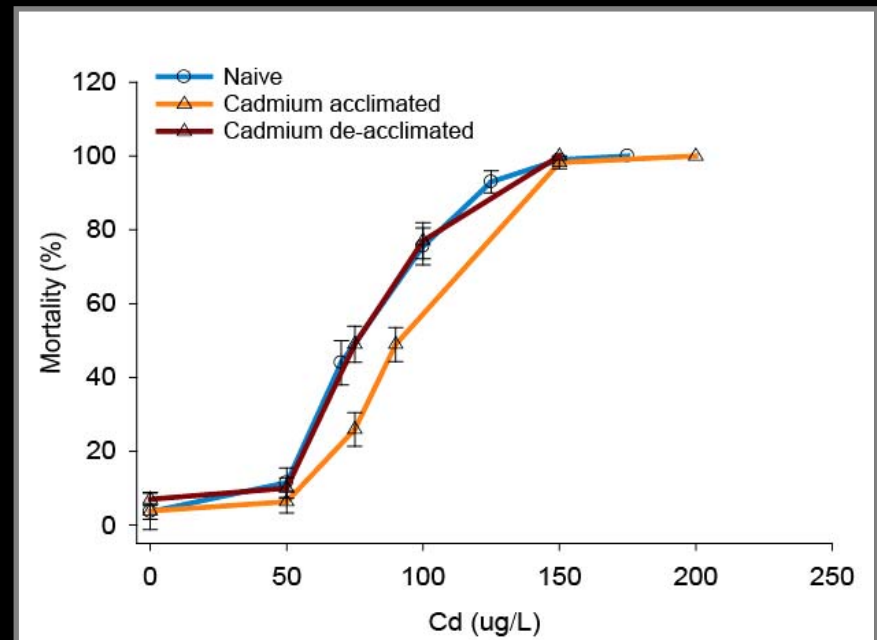
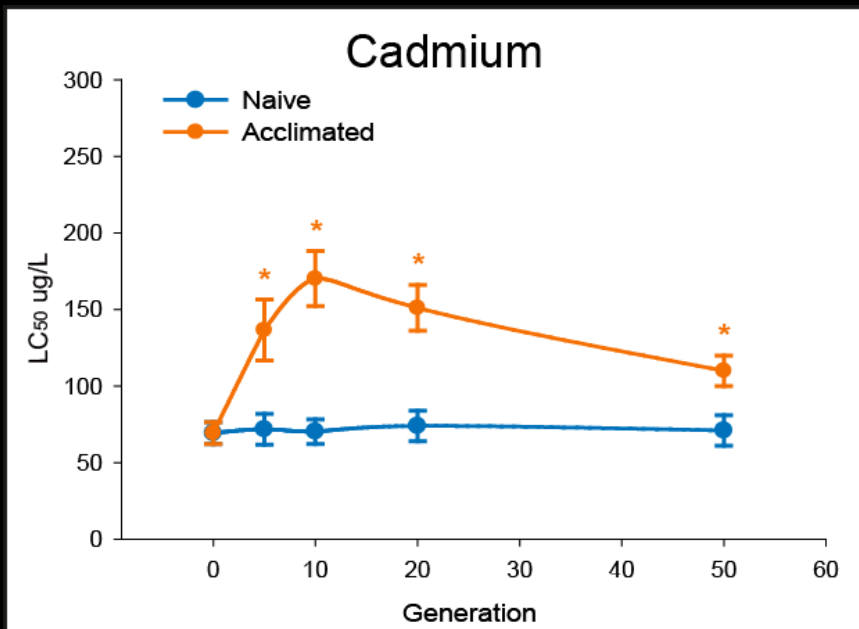
Top 16-down regulated genes

- Unknown (11)
- Monooxygenase
- Mannanase
- Ferritin
- Carboxypeptidase

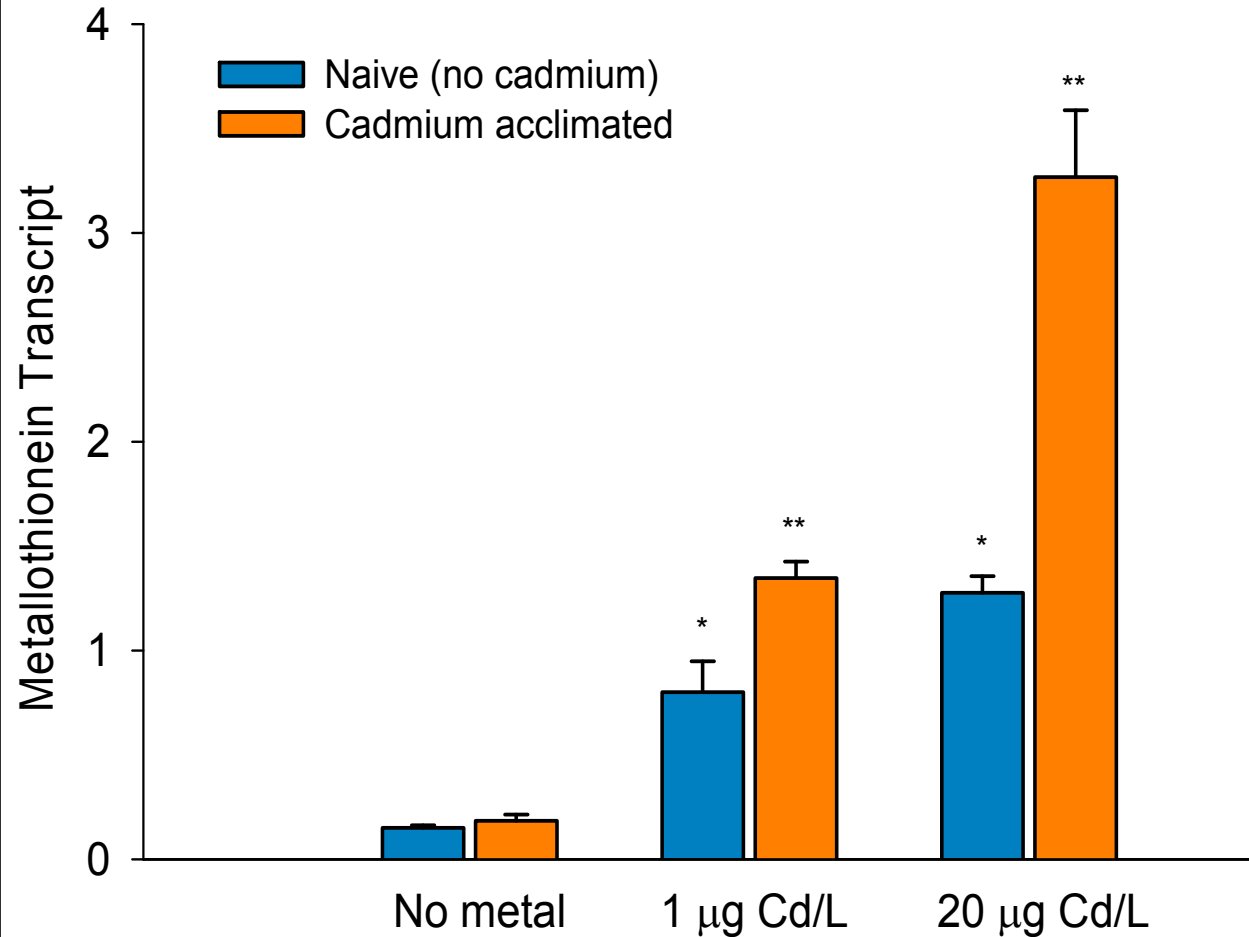
Acclimation vs. adaptation

Acclimation: The process by which an organism becomes accustomed to a new environment (involves physiological changes and occurs at the level of the individual).

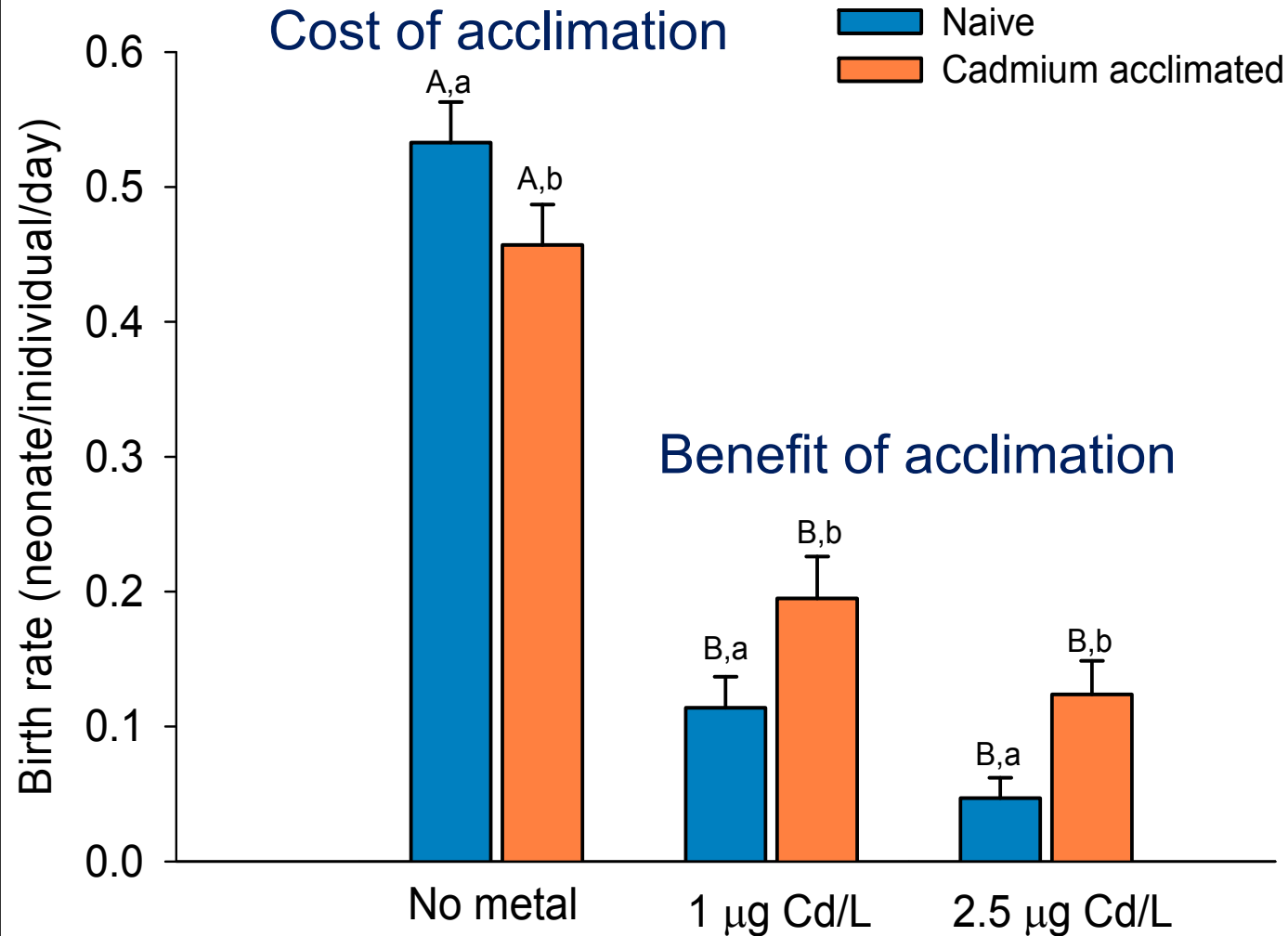
Adaptation: A change in structure, function, or form across generations that allows an organism to better adjust to its environment (involves genetic change within populations).



Does acclimation change gene expression?



Are there costs of acclimation?



Evaluating historic pollution of Ontario lakes



Smelting in Sudbury, Canada

Sudbury, Ontario

Smelting operations (mid-1800s)

Industrial scale processing (early to mid-1900s)

> 6000 lakes contaminated with metals

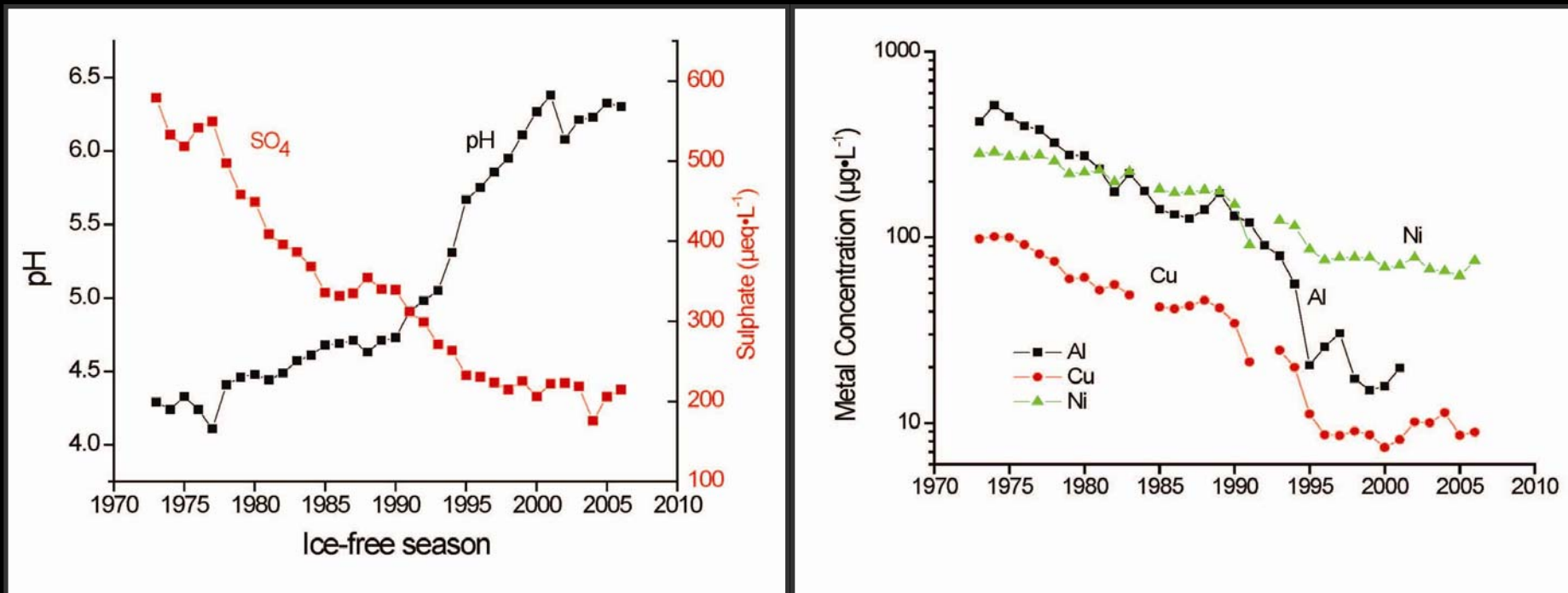
Study Plan

- ▶ 7 Sudbury area lakes contaminated with metals
- ▶ 7 reference lakes (Dorset)
- ▶ culture 12-20 unique genotypes per lake

- ▶ ~200 laboratory cultures

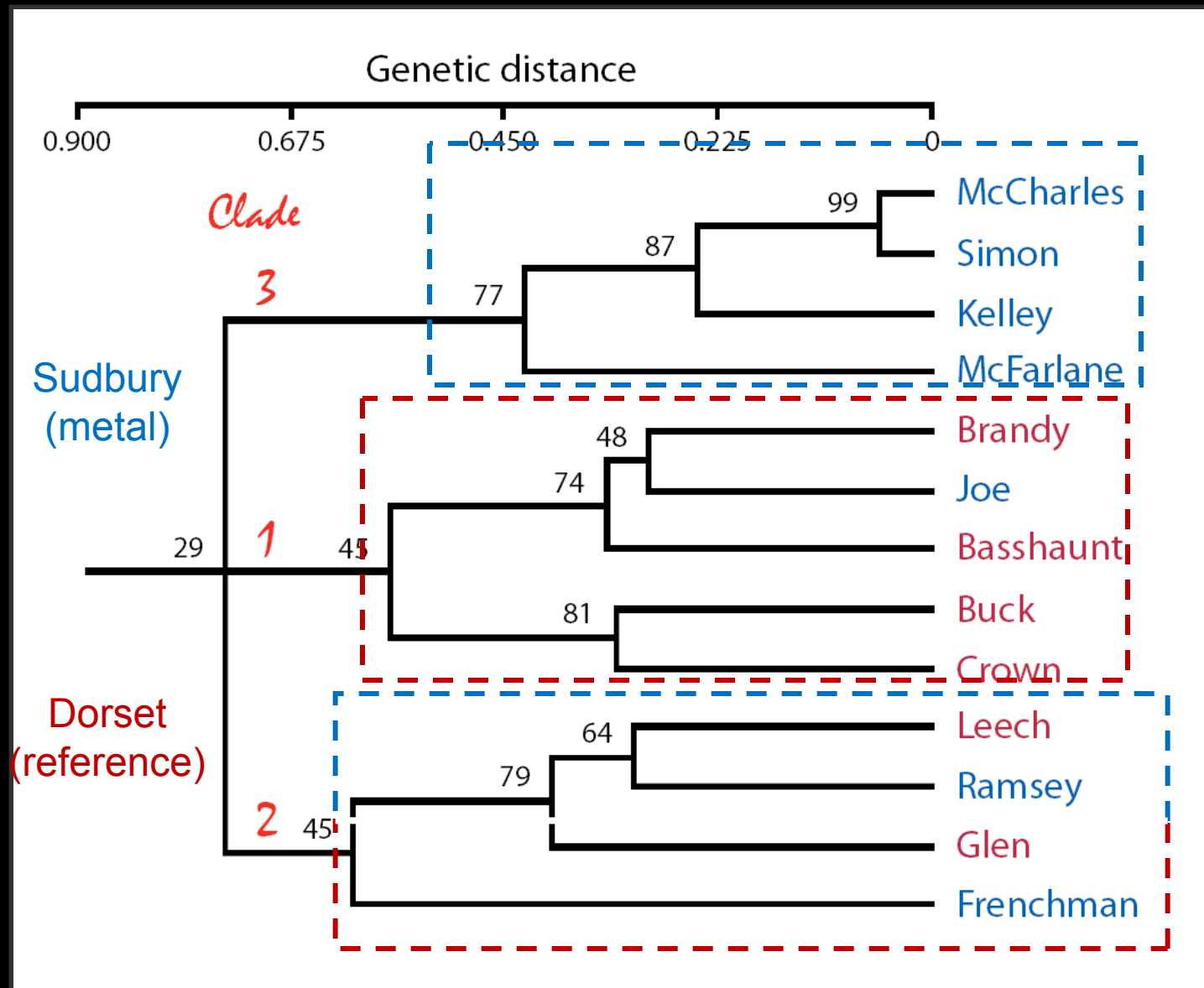
Pollution history and recovery

Clearwater Lake, Sudbury

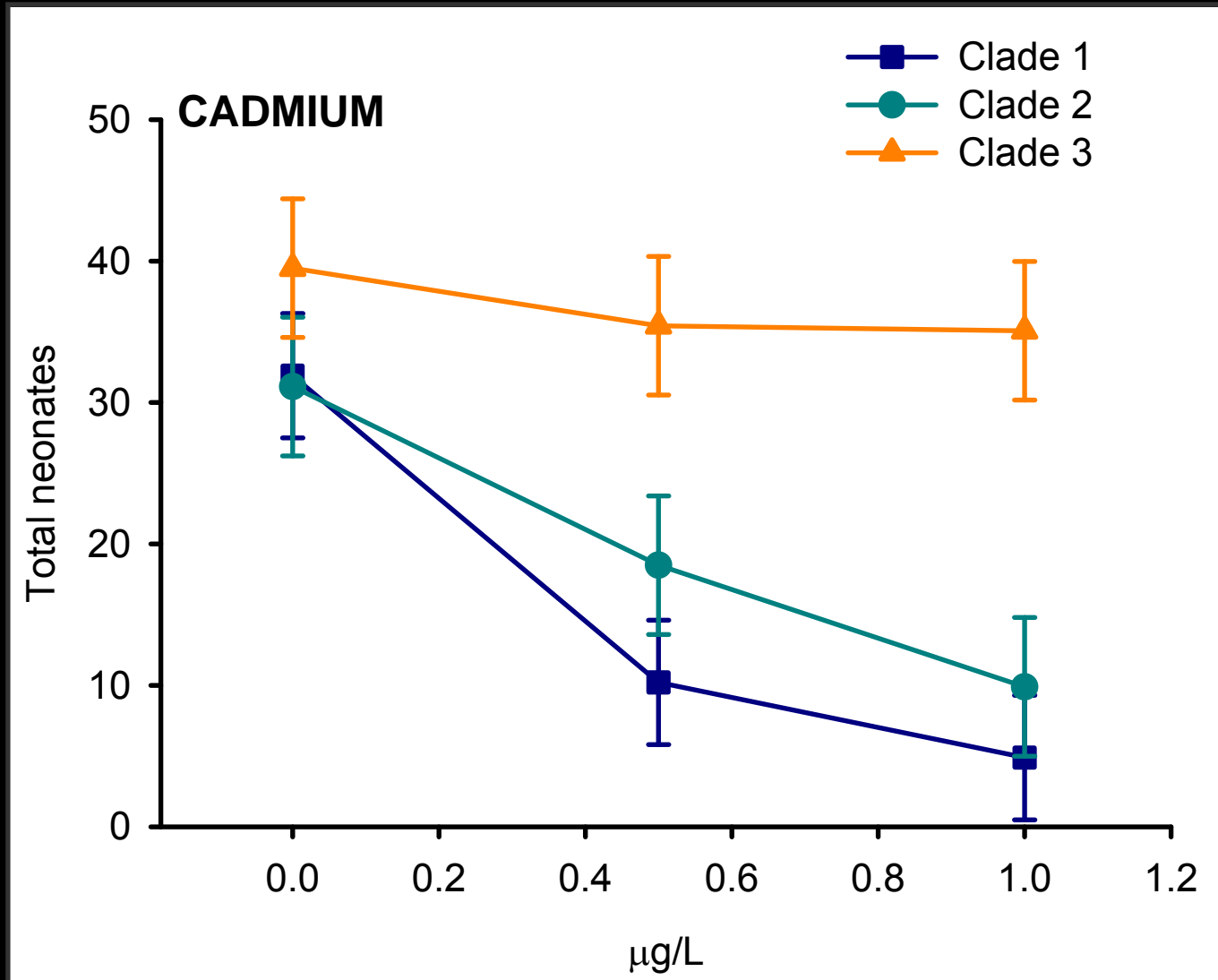


Norman Yan, Ontario Ministry of the Environment

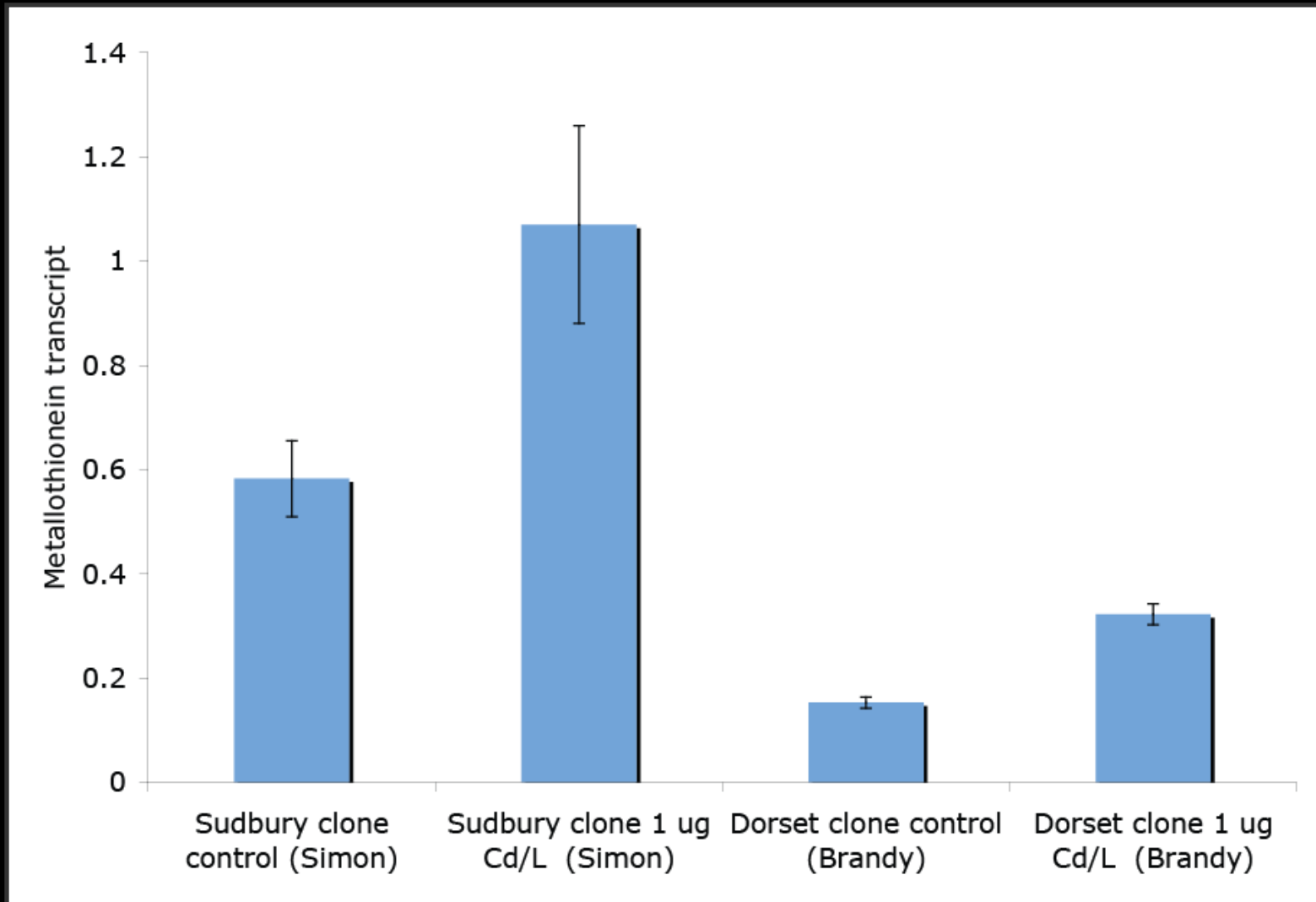
Phylogeography of study populations



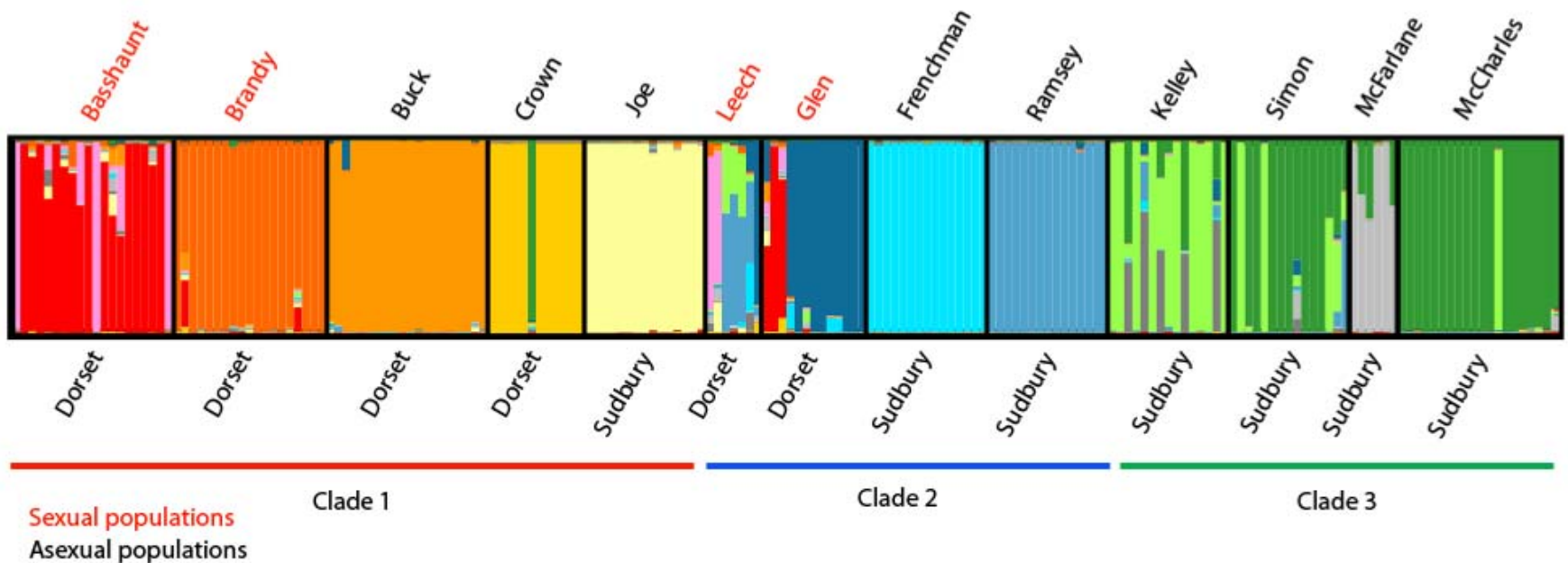
Adaptation of natural populations



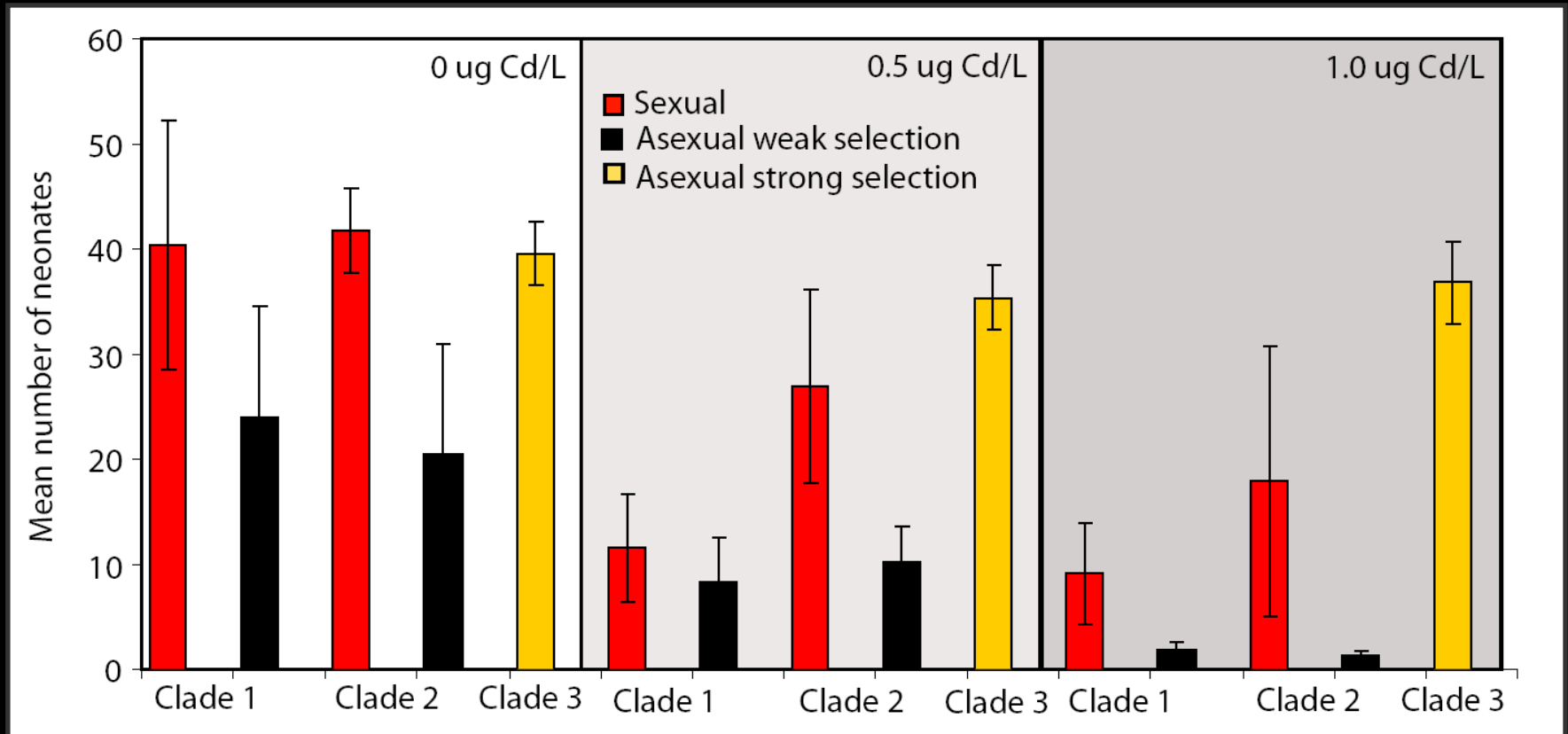
Does adaptation change gene expression?



Metapopulation includes sexual and obligate asexual lineages



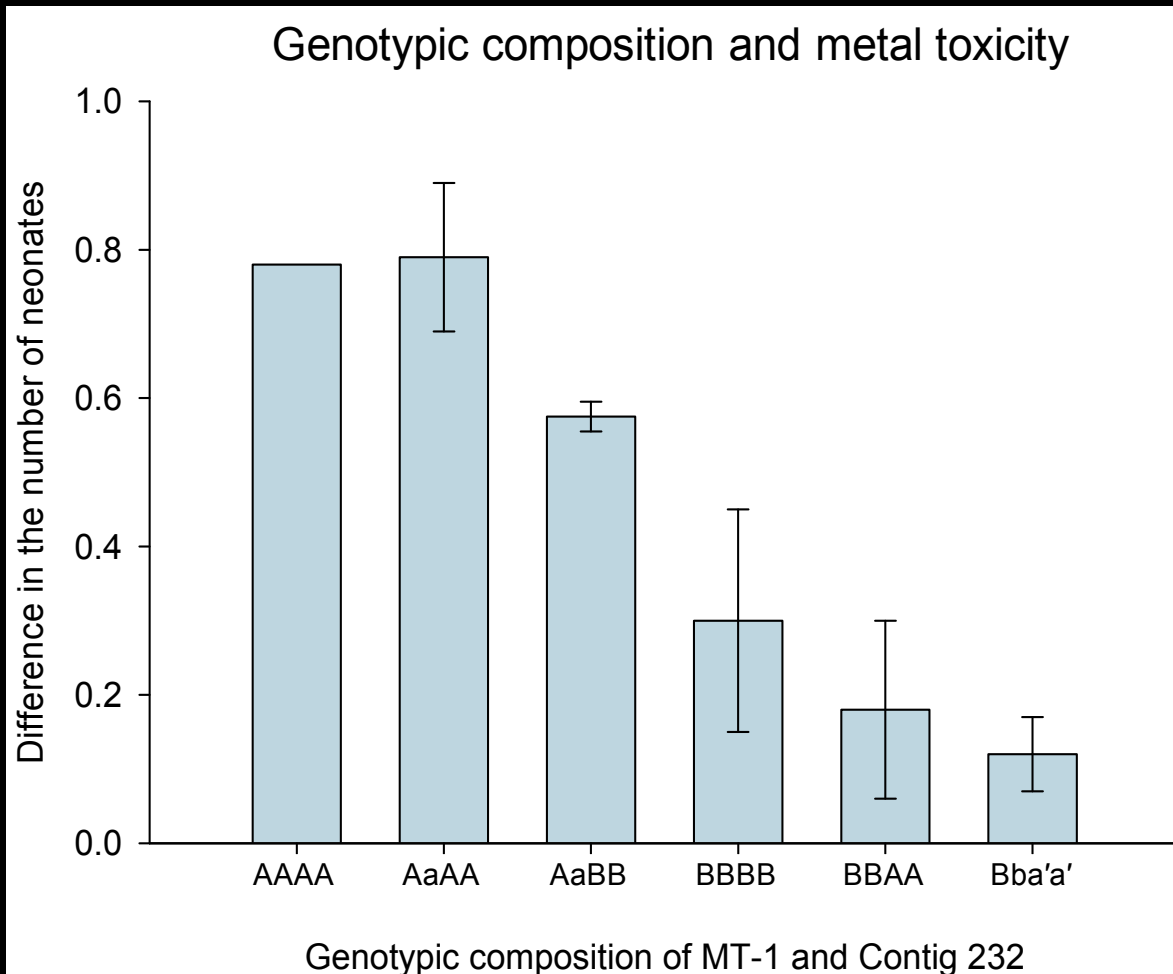
Adaptation of asexual natural populations



Questions to be resolved

- What genetic changes account for a population's success or failure at coping with environmental challenges?
- Is adaptation the fixation of the acclimated phenotype (what about costs)?
- Can we use animal populations to predict population-level susceptibility to chemical contaminants?

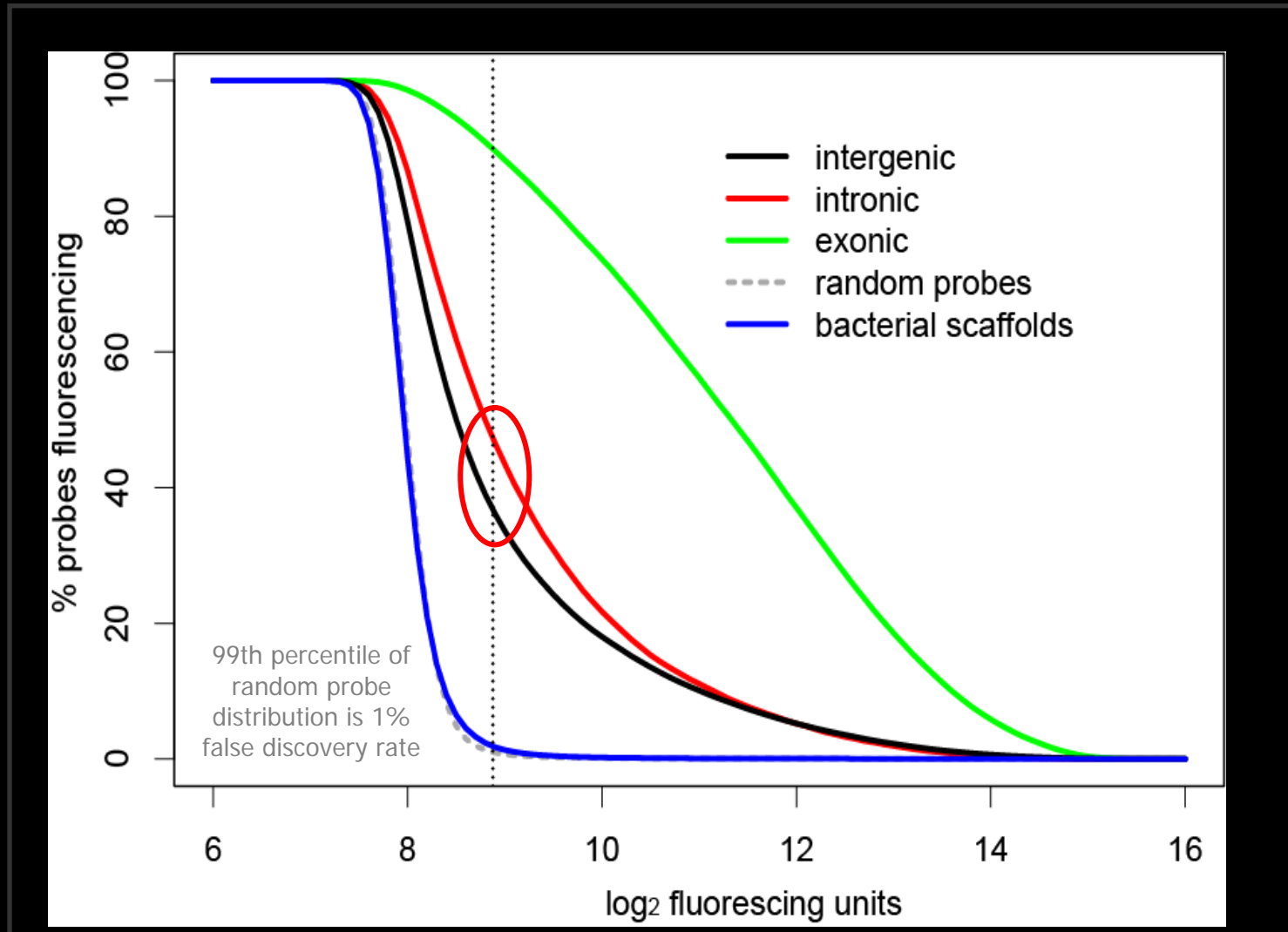
Where we are heading: Identifying susceptible populations



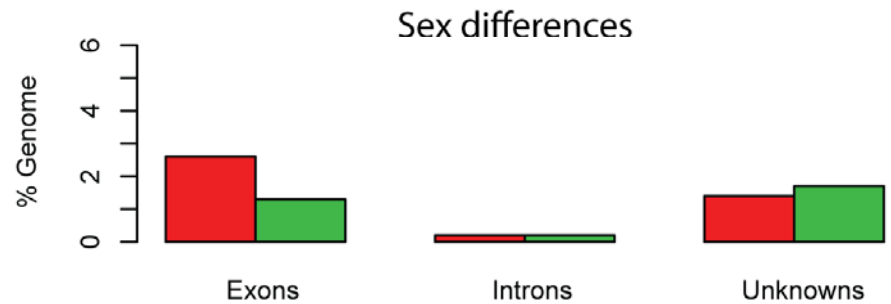
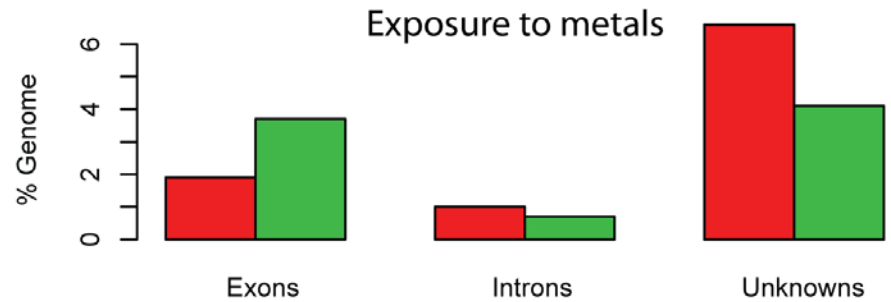
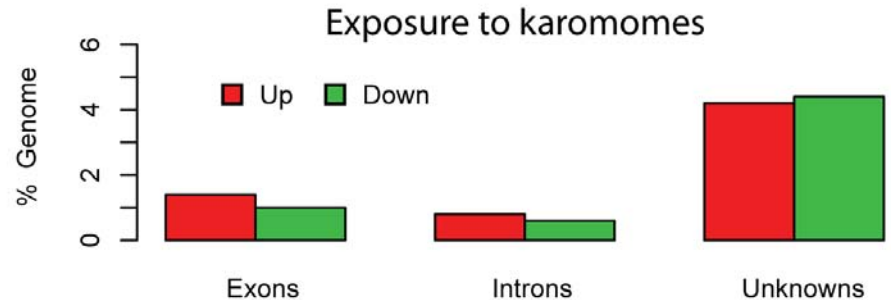
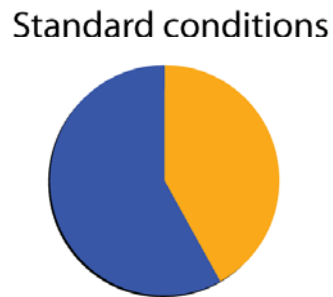
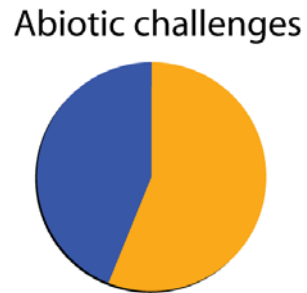
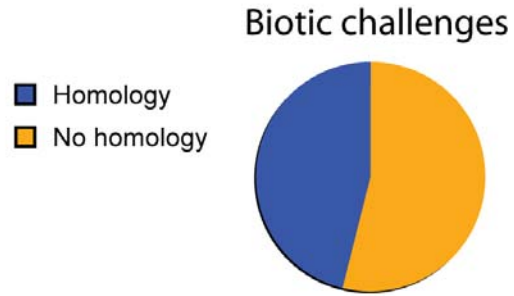
Genome-Population Coupling

- 1) Transcriptome
- 2) Genome
- 3) Genomic variation
- 4) Phenotype/condition

Observed transcription across the annotated *Daphnia* genome



Unknown genes are ecologically/toxicologically important



Toxicity Testing for the 21st Century

POLICYFORUM

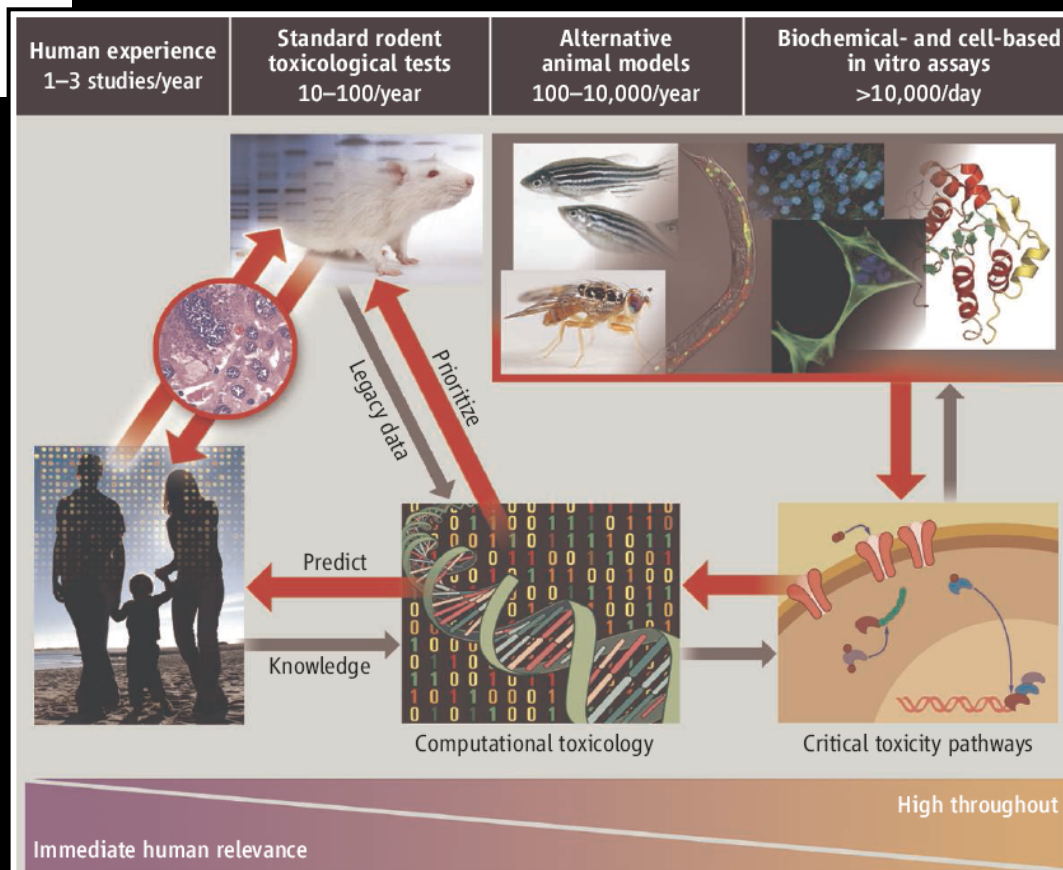
15 FEBRUARY 2008 VOL 319 SCIENCE www.sciencemag.org

TOXICOLOGY

Transforming Environmental Health Protection

Francis S. Collins,^{1*} George M. Gray,^{2*} John R. Bucher^{3*}

¹Director, National Human Genome Research Institute (NHGRI), National Institutes of Health, Bethesda, MD 20892; ²Assistant Administrator for the Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC 20460; ³Associate Director, U.S. National Toxicology Program, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, NC 27709, USA.



Toxicity Testing for the 21st Century

POLICYFORUM

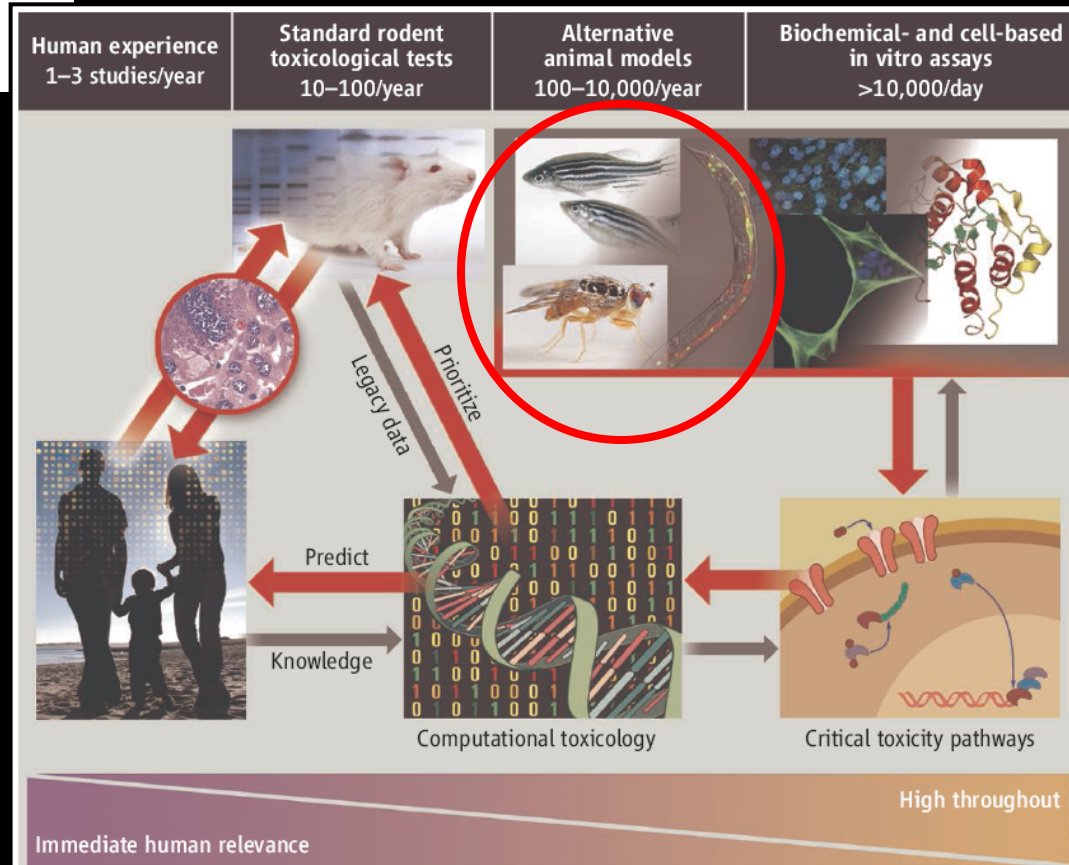
15 FEBRUARY 2008 VOL 319 SCIENCE www.sciencemag.org

TOXICOLOGY

Transforming Environmental Health Protection

Francis S. Collins,^{1*} George M. Gray,^{2*} John R. Bucher^{3*}

¹Director, National Human Genome Research Institute (NHGRI), National Institutes of Health, Bethesda, MD 20892; ²Assistant Administrator for the Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC 20460; ³Associate Director, U.S. National Toxicology Program, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, NC 27709, USA.



Model organisms

Model Organisms for Biomedical Research



Mammalian Models:

- Mouse
- Rat

Non-Mammalian Models:

- *S. cerevisiae* (budding yeast)
- *S. pombe* (fission yeast)
- *Neurospora* (filamentous fungus)
- *D. discoideum* (social amoebae)
- *C. elegans* (round worm)
- *Daphnia* (water flea)
- *D. melanogaster* (fruit fly)
- *D. rerio* (zebrafish)
- *Xenopus* (frog)
- *Gallus* (chicken)

Other Model Organisms:

- *Arabidopsis*

Other:

- Reports
- Funding Opportunities
- Process for Considering Support



Daphnia Genomics Consortium

- Sequenced, assembled, and annotated a *Daphnia pulex* genome
- Sequenced (10X) and produced an early assembly of a *Daphnia magna* genome

Extending the approach to killifish



- Sequenced a *Fundulus heteroclitus* cDNA library (454) representing 72 unique conditions (e.g., sex, tissue, salinity, metal exposure).
- Defined the transcriptome by expanding the EST database from 65,000 to over 1,000,000 sequence reads.
- Constructed an ultra-dense microarrays (12 X 135,000) on the Roche/NimbleGen platform.
- Developed a collaborative proposal to sequence a killifish genome (under-review at NSF).

We gratefully acknowledge funding from

DOD, NSF, DOE, NIEHS