

# Semantic models for environmental exposures:

Linking geo and bio via the toxicology community

Anne E Thessen, Lauren E Chan, Rebecca Boyles, Chris Mungall, and Melissa A Haendel

annethessen@gmail.com

@diatomsRcool



@monarchinit



Link to these slides:  
[bit.ly/geo-tox-2020bit](https://bit.ly/geo-tox-2020bit)

# Acknowledgements

- Melissa Haendel and TISLab
  - Julie McMurry
  - Chris Mungall
  - Robyn Tanguay
  - Lisa Truong
  - Bianca Yaghoobi
  - Ramona Walls
  - Tiffany Callahan
  - Brian Westra
  - Resham Kulkarni
  - Annie Jarabek
  - Lyle Burgoon
  - Sarah Rothenberg
  - Susanne Brander
  - Susan Korrick
  - Mary Willis
  - ....all attendees of Computable Exposures Workshop
- SEAZIT project
  - Monarch Developers
    - Tom Conlin
    - Ken Lett
    - Kent Shefchek
  - NIH U13# 5U13CA221044-03



monarch  
INITIATIVE

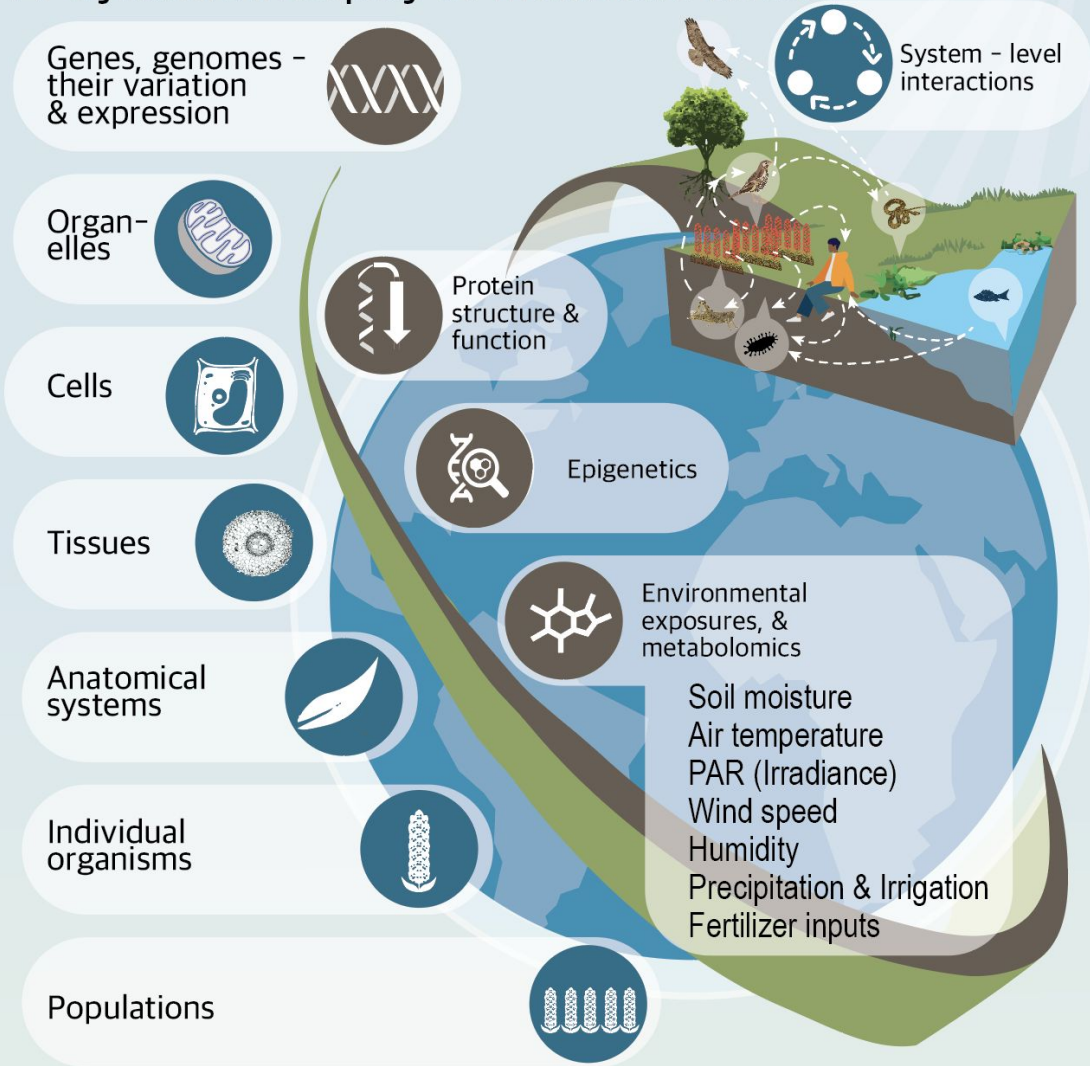


Oregon State  
University



# Genotype + Environment = Phenotype (Big Picture)

- Plethora of data linking genes and phenotypes
- Environment component is often not linked
- Just enough modeling to do what we need to do



# Semantic Representations of Environments

EnVO

- Mountains to concrete

MAXO

- Surgeries, diets, treatments

FOODON

- Any kind of food or drink

NBO

- Behaviors, habits

CHEBI

- Chemicals

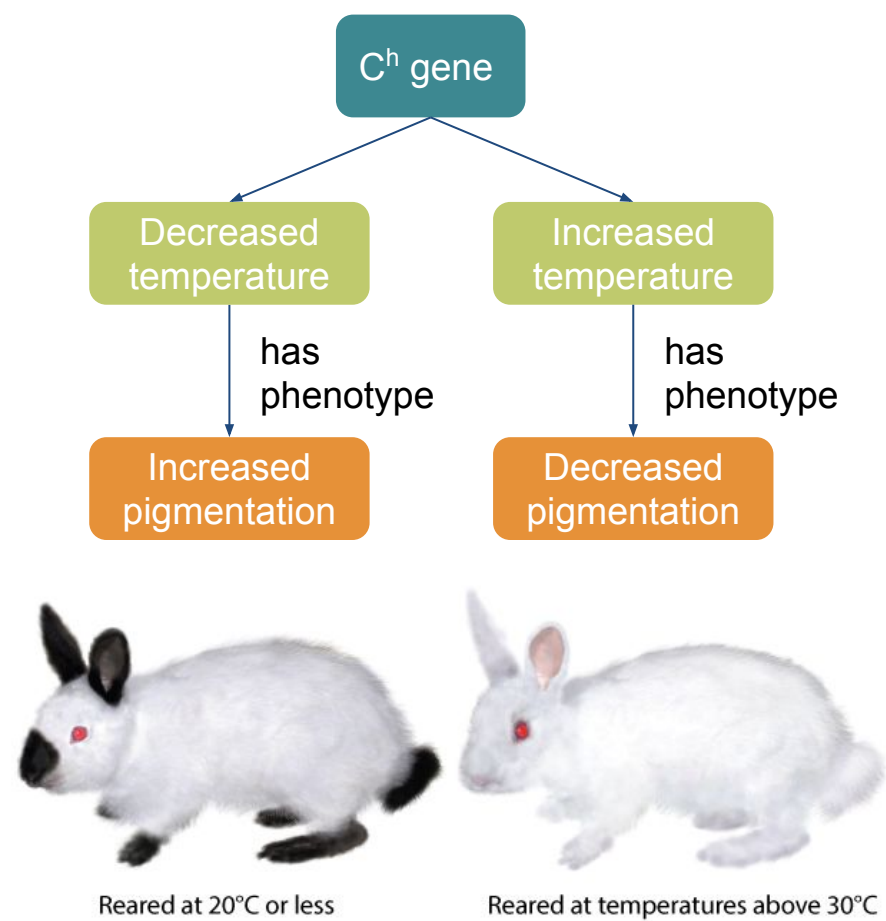


Modeling  
Genotype +  
Environment =  
Phenotype

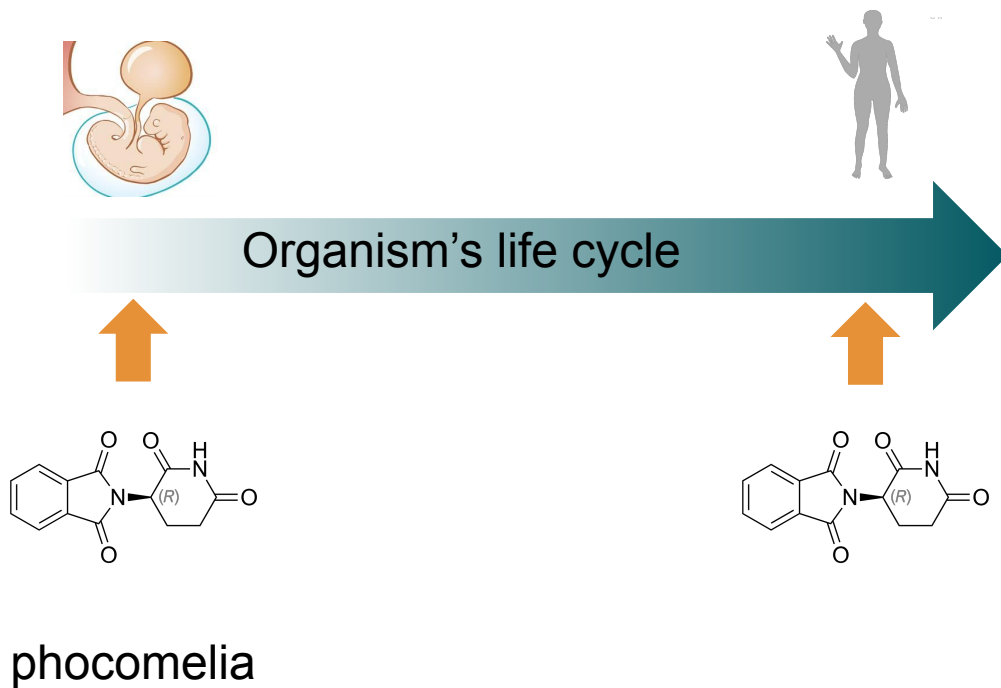
Genotype

Environment

Phenotype



# Modeling Timing of Exposure



# Modeling Frequency and Duration of Exposure

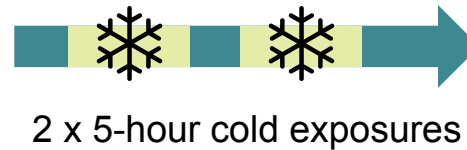
One  
exposure or  
multiple  
exposures



20 genes  
upregulated

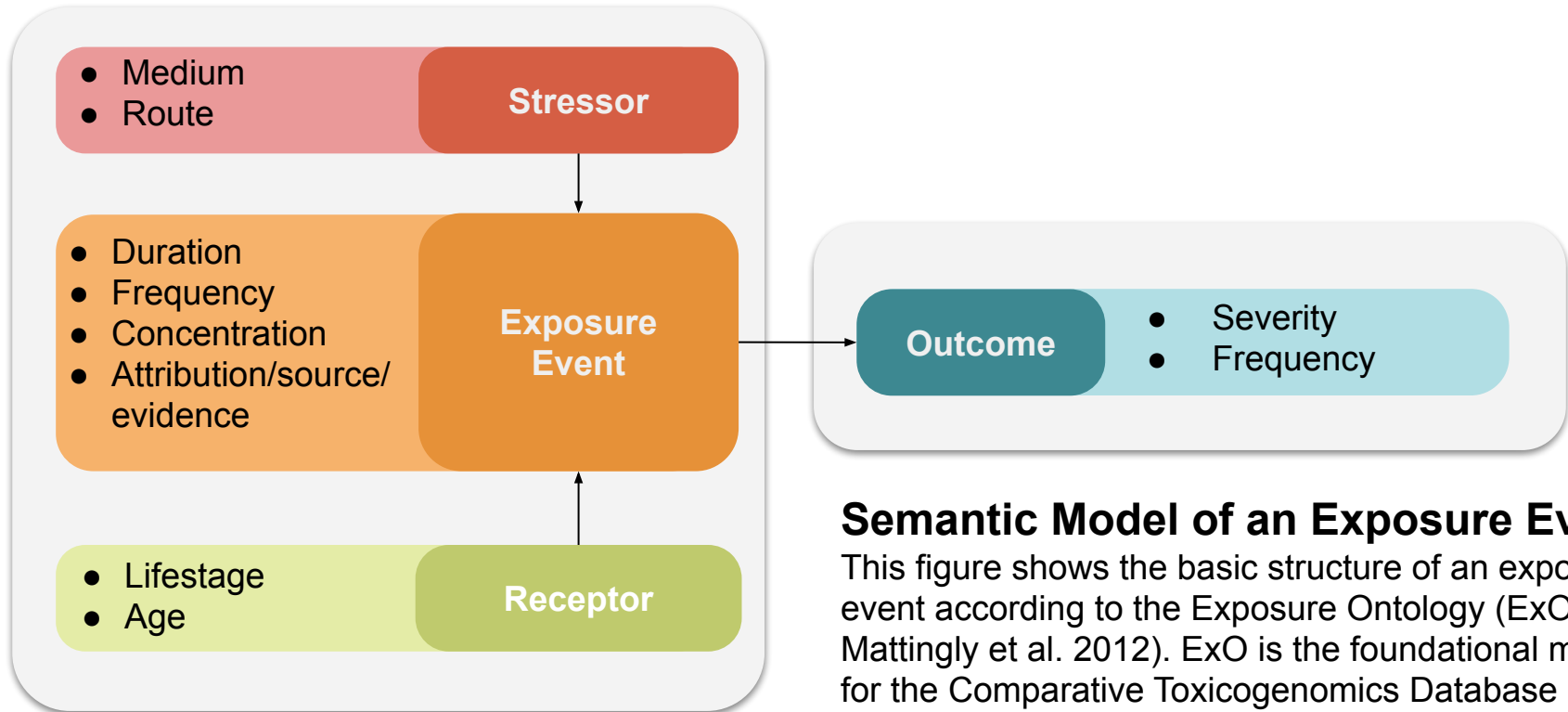


69 genes  
upregulated



76 genes  
upregulated

# Environmental Exposures Modeling with ExO

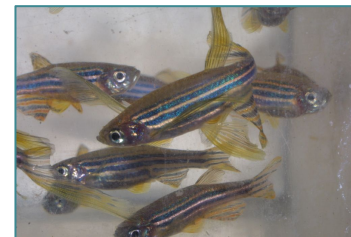
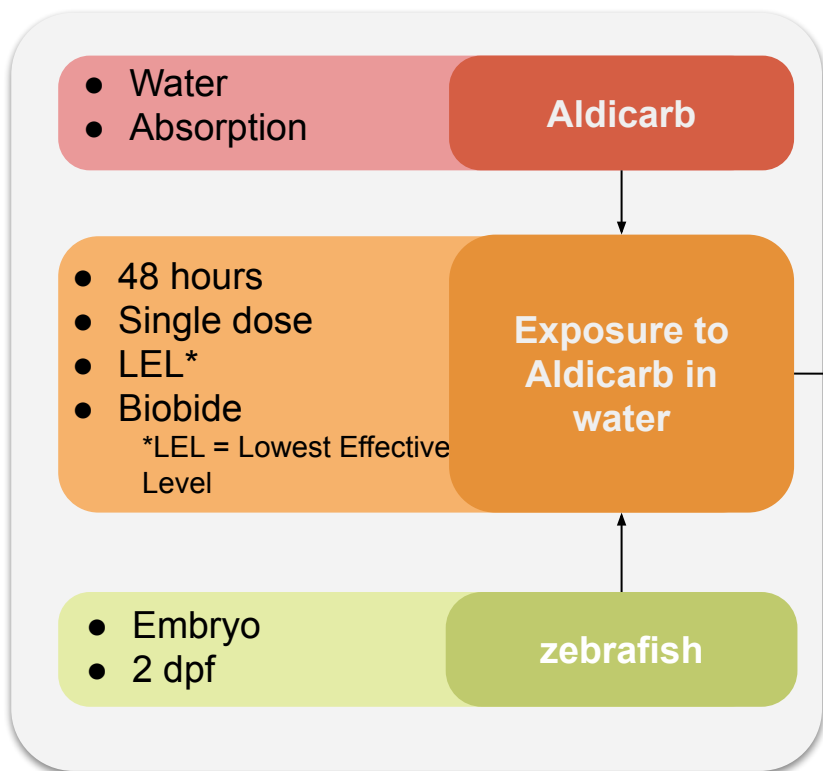


## Semantic Model of an Exposure Event.

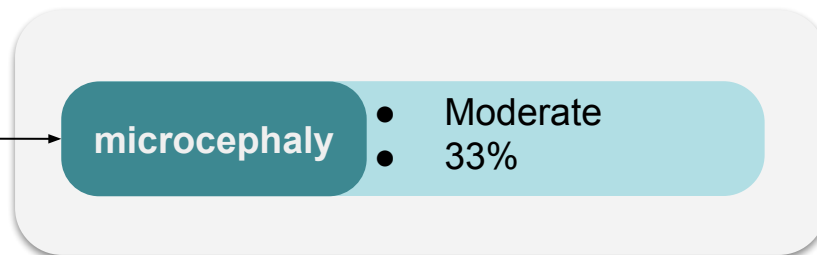
This figure shows the basic structure of an exposure event according to the Exposure Ontology (ExO; Mattingly et al. 2012). ExO is the foundational model for the Comparative Toxicogenomics Database (CTD; Mattingly et al. 2006).



# Environmental Exposures Modeling with ExO - Example



Steven J. Baskauf cc-by-4.0

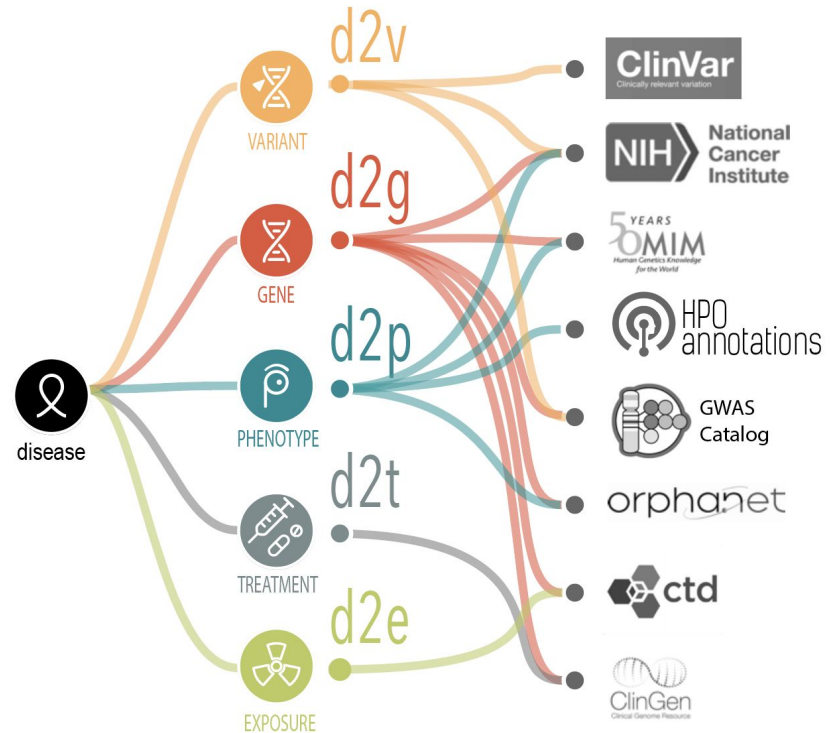


## Semantic Model of an Exposure Event.

This figure shows the basic structure of an exposure event according to the Exposure Ontology (ExO; Mattingly et al. 2012). ExO is the foundational model for the Comparative Toxicogenomics Database (CTD; Mattingly et al. 2006).

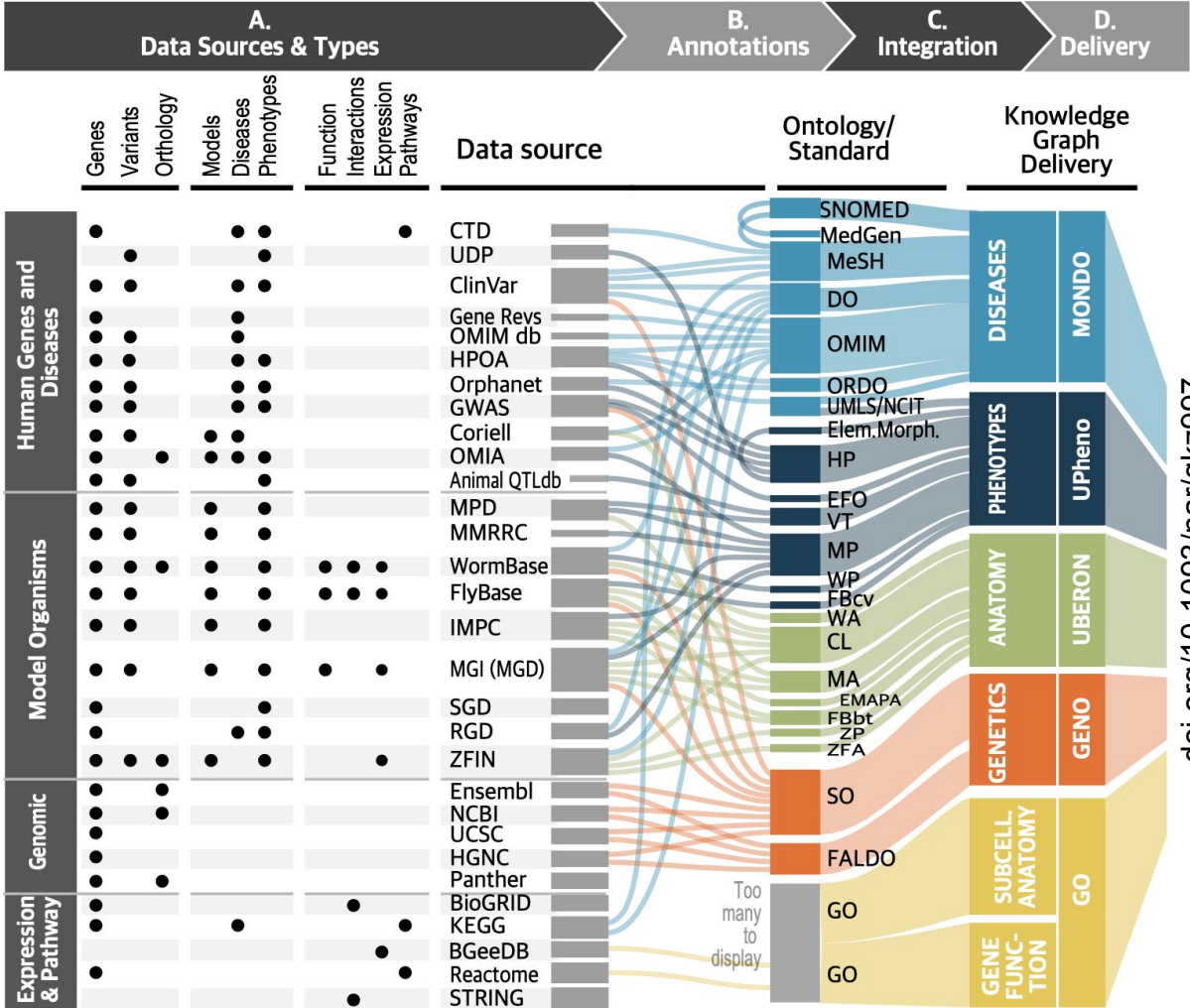
# Integrating Exposures into Monarch

- The Monarch Initiative integrates genotype and phenotype data across species for improved disease diagnosis (Shefcheck et al. 2019).
- Monarch only integrates exposure data from one source (CTD) and cannot take advantage of all available data
- Relationships between exposures and phenotypes are missing
- This makes Monarch less useful than it could be for diagnosing diseases with an important environmental component



# What is the Monarch Knowledge Graph?

- monarchinitiative.org
- Integrator of cross species genotype-phenotype data
- OWL and DOS-DP
- Uses OBO ontologies



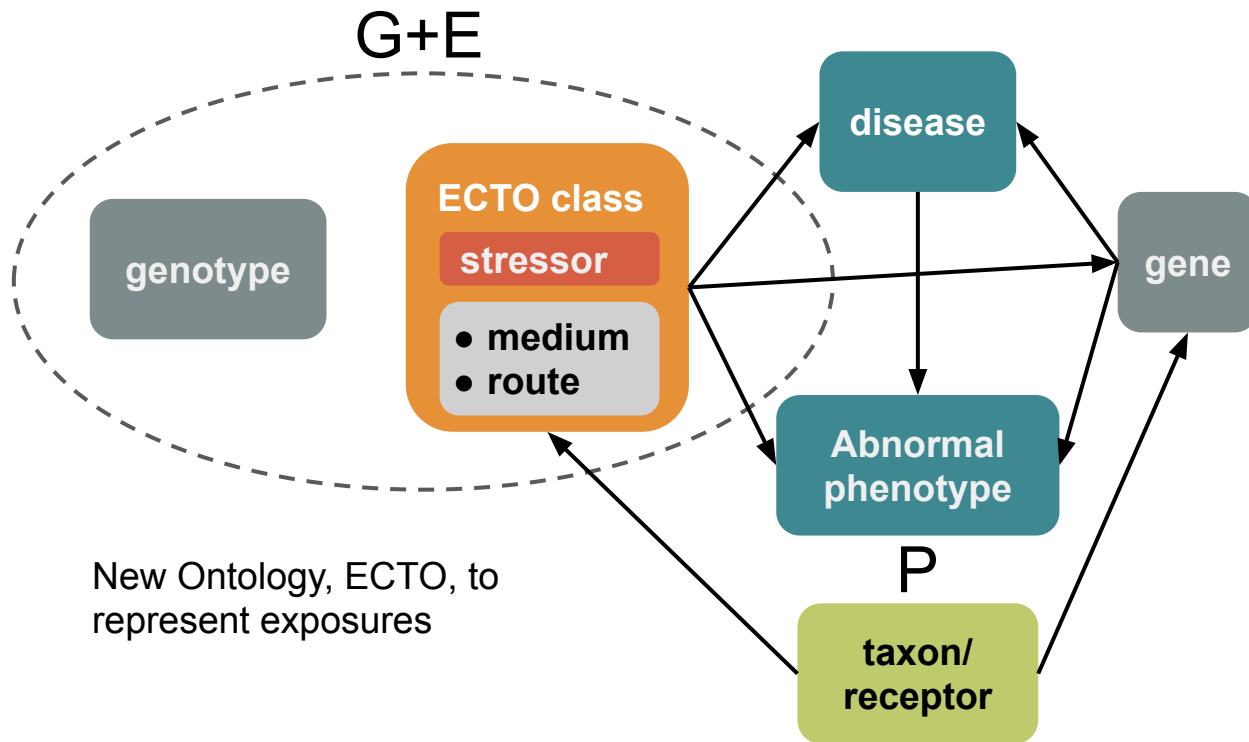
doi.org/10.1093/nar/gkz997

# Computable Exposures Workshop

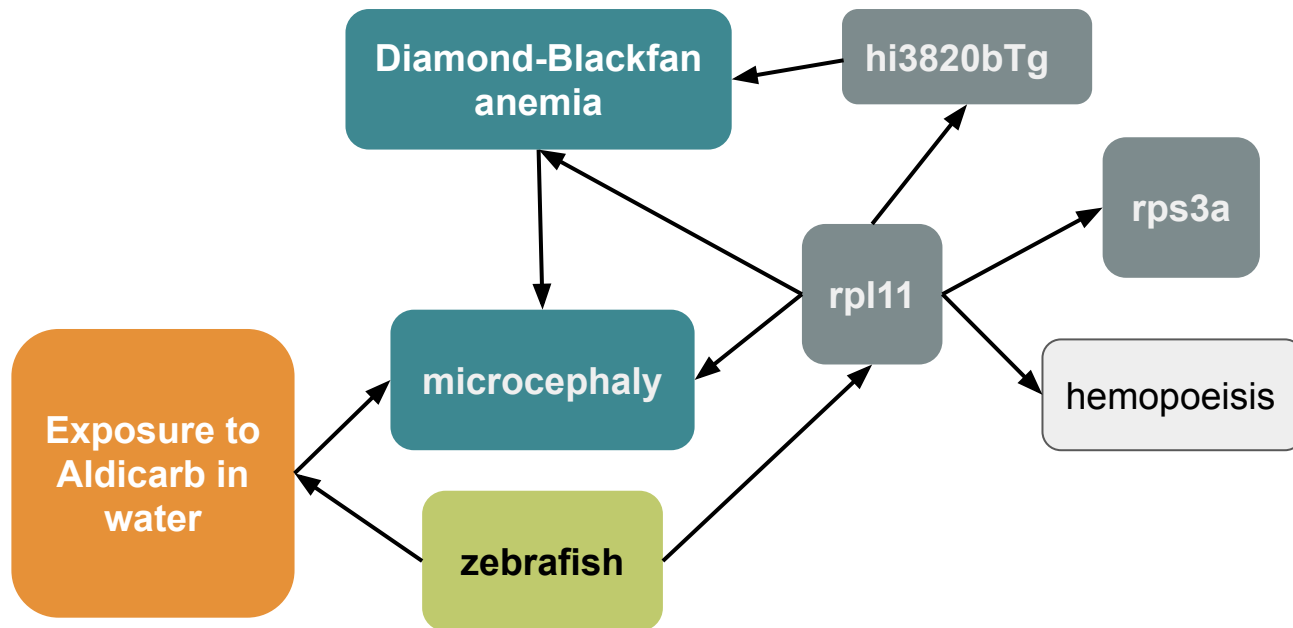
- 28 participants
- To guide the development of the exposure data model, use cases and competency questions were developed.

| Example Competency Questions:   | Priority Concepts | Ontologies    | Data Type         |
|---|-------------------|---------------|-------------------|
| What <b>fish</b> species have <b>reduced fecundity</b> following sublethal <b>pyrethroid</b> exposure?  | Taxon             | NCBI Taxonomy | Receptor          |
|   | Phenotype         | uPheno        | Phenotype/Outcome |
| What is a safe amount of <b>iron</b> to consume for someone diagnosed with <b>hereditary hemochromatosis</b> (either <b>heterozygous</b> or <b>homozygous</b> for a certain variant)? | Genotype          | GENO/GO       | Gene              |
|   | Chemical          | CHEBI/ECTO    | Stressor          |
|   | Disease           | Mondo         | Disease/Outcome   |

# Integration with the Monarch Knowledge Graph



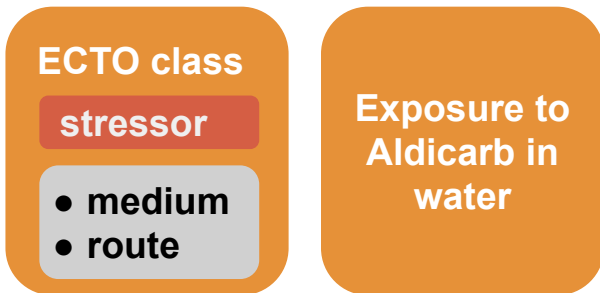
# Integration with the Monarch Knowledge Graph



# ECTO: Environmental conditions, treatments and exposures

- Exposure event
  - Exposure to Aldicarb
    - Exposure to Aldicarb in water
      - Exposure to Aldicarb in water via ingestion

Exposure event from ExO  
Aldicarb from CHEBI (stressor)  
Water from ENVO (medium)



exposure event and (has\_exposure\_stimulus some Aldicarb) and (has\_exposure\_medium some water)

# Patterns and typing

## DOS-DP patterns

- Chemicals
- Medical actions
- **Environmental processes**
- **Environmental materials**
- Behaviors
- Change
- Organisms
- **Geographic feature**
- Occupations and hobbies



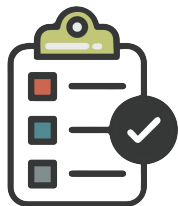
### Semantic typing in Biolink Model

- Exposure event - includes environments and experimental conditions
- Associated with phenotypic feature and disease
- Chemical exposure is a subclass - includes drug treatment



# Future Work

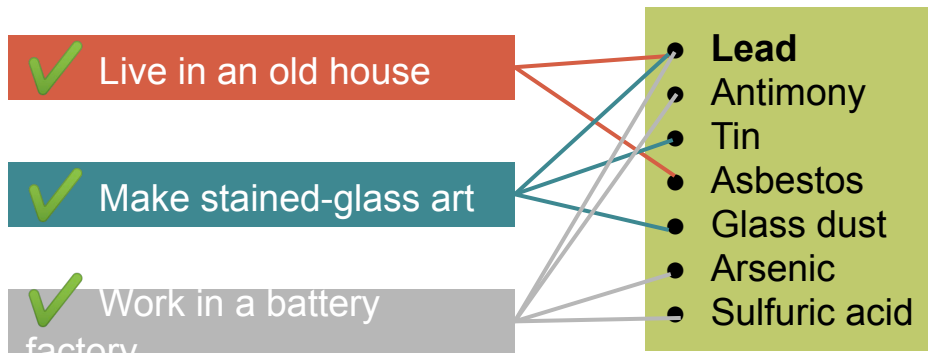
## Data from surveys:



Environmental Polymorphisms Registry  
(EPR) Survey

Translating questions into ECTO classes for  
novel data integration (Chulada et al. 2011)

[dnaregistry.niehs.nih.gov](http://dnaregistry.niehs.nih.gov)



## Refine models:

To account for:

- Certainty of cause and effect relationships
- Mixtures - multiple exposures at once
- Combinations of exposures and genes lead to phenotype

# References

Chulada et al. 2011. The environmental polymorphisms registry: a unique resource that facilitates translational research of environmental disease. *Env Health Pers*, [doi.org/10.1289/ehp.1003348](https://doi.org/10.1289/ehp.1003348)

Mattingly et al. 2006. The Comparative Toxicogenomics Database: A Cross-Species Resource for Building Chemical-Gene Interaction Networks. *Tox Sci*, [doi.org/10.1093/toxsci/kfl008](https://doi.org/10.1093/toxsci/kfl008)

Mattingly et al. 2012. Providing the missing link: The exposure science ontology ExO. *Env Sci & Tech*, [doi.org/10.1021/es2033857](https://doi.org/10.1021/es2033857)

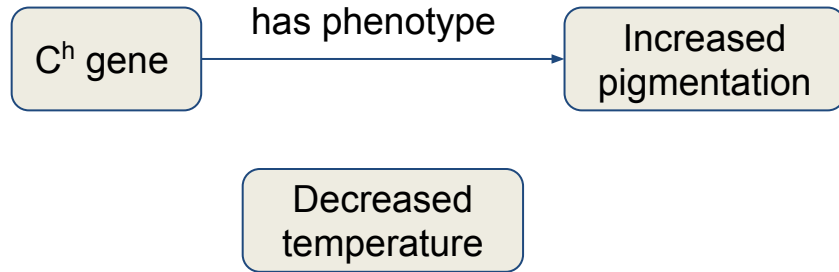
Mattingly et al. 2016. Laying a Community-Based Foundation for Data-Driven Semantic Standards in Environmental Health Sciences. *Env Health Pers*, [doi.org/10.1289/ehp.1510438](https://doi.org/10.1289/ehp.1510438)

Shefchek et al. 2019. The Monarch Initiative in 2019: an integrative data and analytic platform connecting phenotypes to genotypes across species. *Nuc Acids Res*, [doi.org/10.1093/nar/gkz997](https://doi.org/10.1093/nar/gkz997)

Workshop report [doi.org/10.5281/zenodo.3697113](https://doi.org/10.5281/zenodo.3697113)

Workshop poster [doi.org/10.5281/zenodo.3697160](https://doi.org/10.5281/zenodo.3697160)

# Modeling Genotype + Environment = Phenotype

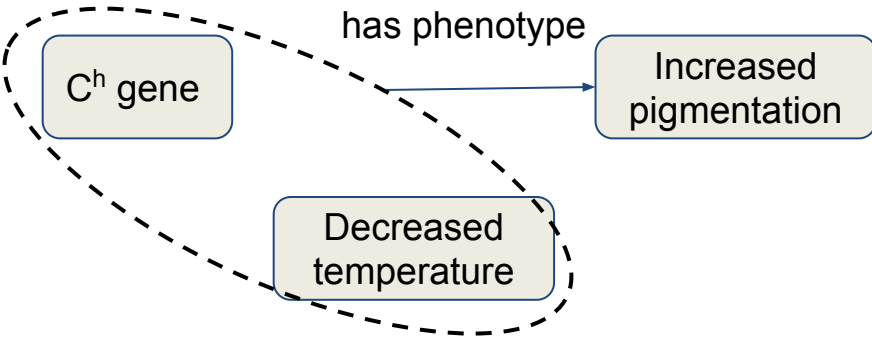


Reared at 20°C or less

Reared at temperatures above 30°C

Lobo, I. (2008) Environmental influences on gene expression. Nature Education 1(1):39

# Modeling Genotype + Environment = Phenotype



Reared at 20°C or less

Reared at temperatures above 30°C

Lobo, I. (2008) Environmental influences on gene expression. Nature Education 1(1):39

# Environmental Exposures Modeling with ExO

- Water
- Absorption

**Aldicarb**

pesticide

- 48 hours
- Single dose
- LEL
- Biobide

**Exposure Event**

**Microcephaly**

- Moderate
- 33%

- Embryo
- 2 dpf

**Zebrafish**

## Semantic Model of an Exposure Event.

This figure shows the basic structure of an exposure event according to the Exposure Ontology (ExO; Mattingly et al. 2012). ExO is the foundational model for the Comparative Toxicogenomics Database (CTD; Mattingly et al. 2006).

LEL = Lowest Effective Level