EPA
United States
Environmental Protection
Agency

Evolution of the CompTox Program at EPA

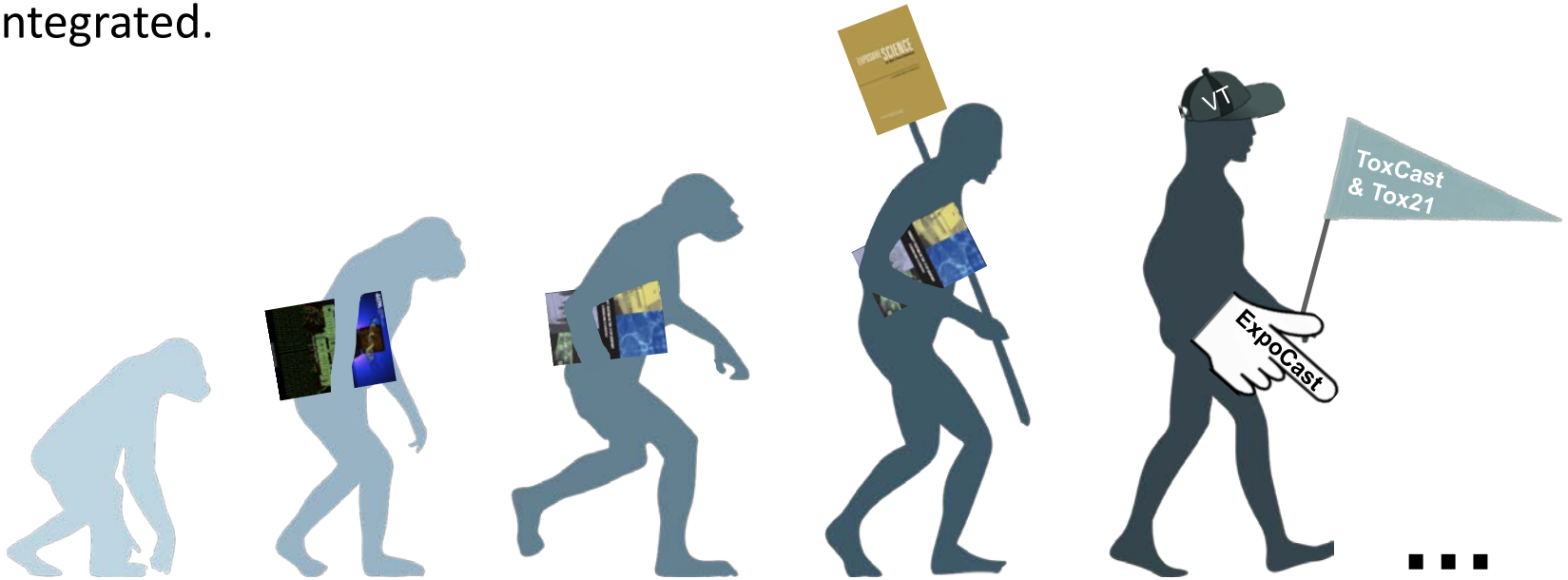
From Primitive Beginnings to Sophisticated Application

Global Chem
March 3, 2015

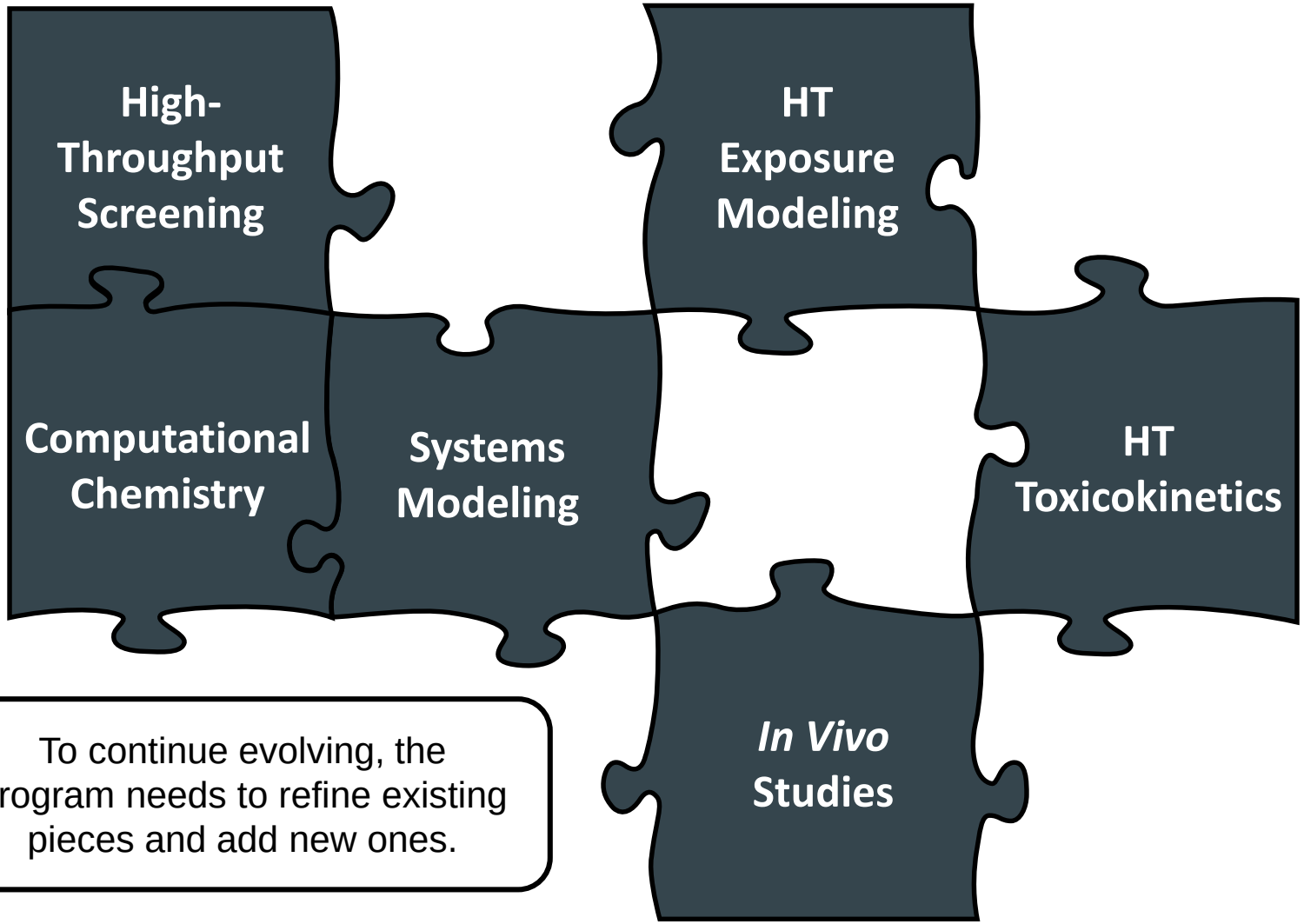
Rusty Thomas
Director
National Center for Computational Toxicology

Introduction

- The CompTox program at the Agency and the field in general have undergone punctuated and rapid evolution since its inception
- Diverse regulatory needs, complex scientific challenges, large influx of data, and guidance from expert committees have provided significant selection pressure
- The research will be increasingly broad in scope, multi-disciplinary, and highly integrated.

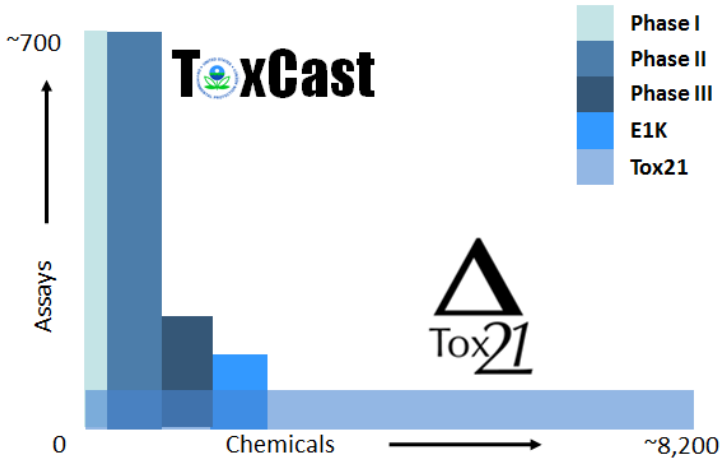


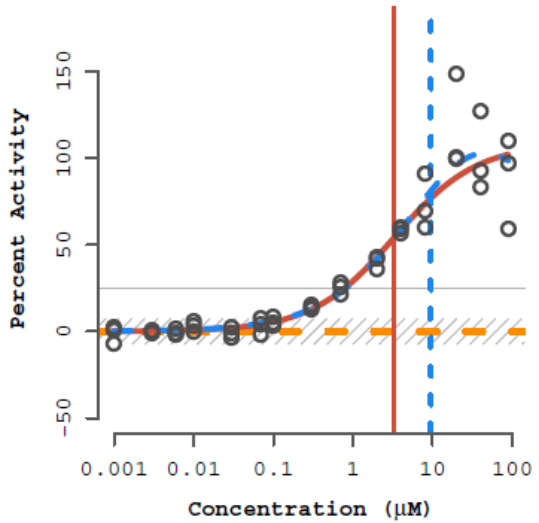
Continuing to Evolve



To continue evolving, the program needs to refine existing pieces and add new ones.

Evolution of High-Throughput Screening





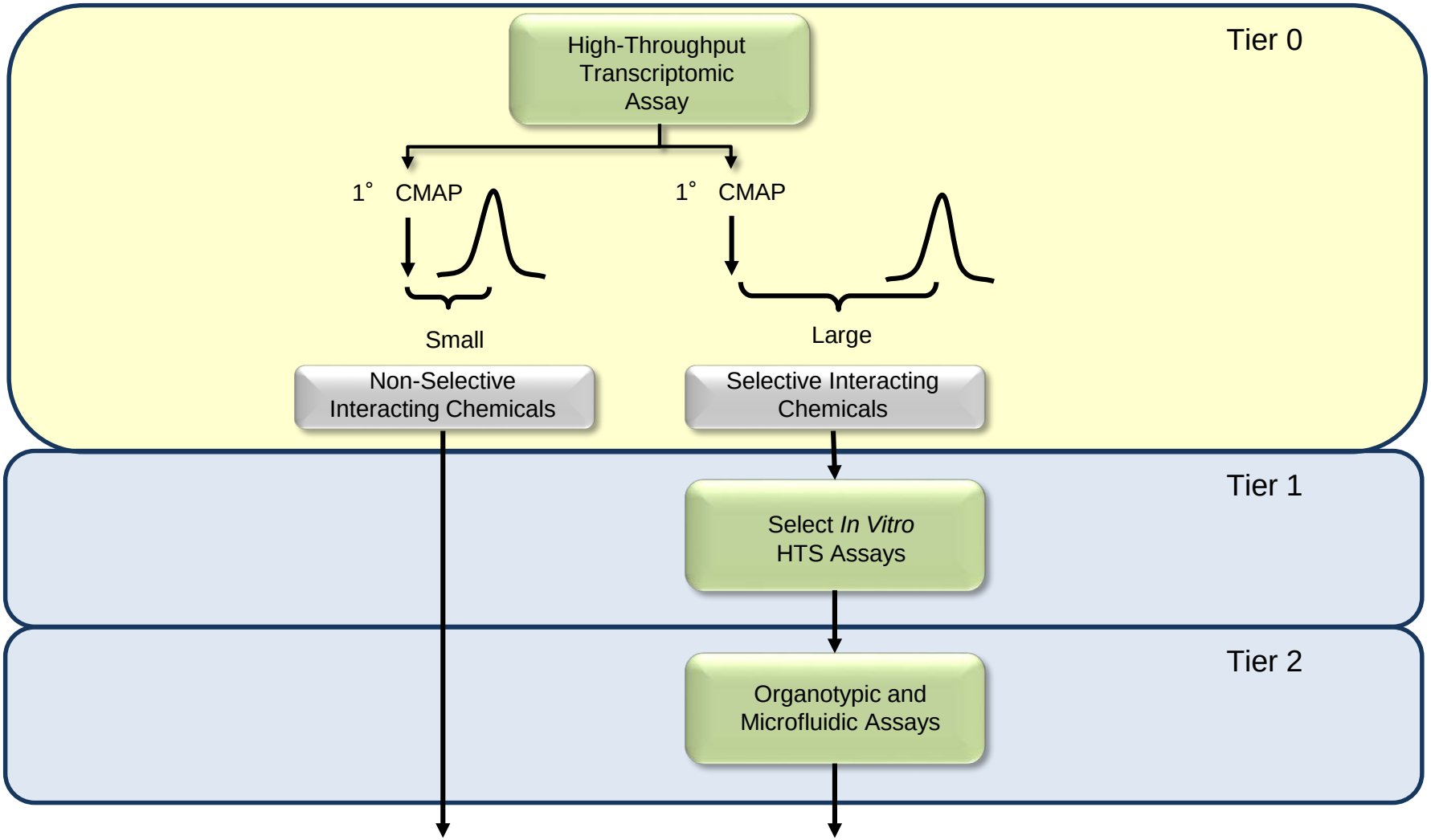
Present

- ToxCast has screened ~2,000 chemicals across ~700 assay endpoints
- Tox21 has screened ~8,200 chemicals across ~50 endpoints
- ToxCast assay coverage represents over 327 genes and 293 pathways

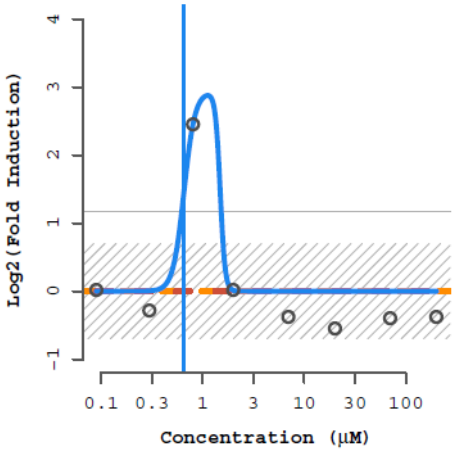
Planned

- Develop low cost, high-throughput global transcriptomic platform to screen across battery of cell types, concentration response, and time
- Capability to retrofit assays for metabolic competence
- Continue assay development for priority targets and complex organotypic cultures

Incorporating a Broad Biological Screening Platform



Efforts to Ensure HTS Data Quality and Increase Transparency



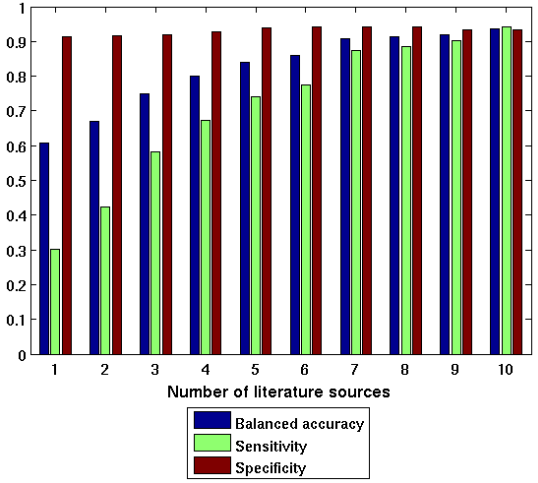

```
ASSAY:    AEID117 (ATG_ERa_TRANS)

NAME:     Thioglycolic acid
CHID:     26141    CASRN: 68-11-1
SPID(S): TX007664
L4ID:     420385

HILL MODEL (in red):
          tp          ga          gw
val:   3.1e-11     -2.15       0.416
sd:    NaN         NaN         NaN

GAIN-LOSS MODEL (in blue):
          tp          ga          gw          la          lw
val:   2.93        -0.184      8           0.173       18
sd:    3.56        0.334       9.48        5.82        814

          CNST        HILL        GNLS
AIC:   20.14       26.14       17.79
PROB:  0.23        0.01        0.76
RMSE:  0.92        0.92        0.32

MAX_MEAN: 2.45      MAX_MED: 2.45      BMAD: 0.233

COFF: 1.17    HIT-CALL: 1    FITC: 50    ACTP: 0.77
```

```
FLAGS:
Only one conc above baseline, active
Borderline active
```

- Public release of Tox21 and ToxCast data on PubChem and EPA web site (raw and processed data)
- ToxCast data analysis pipeline has been completely revamped
  - More statistically rigorous and less prone to outliers
  - Data quality flags to indicate concerns with chemical purity and identity, noisy data, systematic assay errors, and activity in range of cytotoxicity
- Tox21 and ToxCast chemical libraries undergoing analytical QC and results will be publicly released
- Release of ToxCast “Owner’s Manual”
  - Chemical Procurement and QC
  - Data Analysis
  - Assay Characteristics and Performance
- External audit on ToxCast data and data analysis pipeline
- Continued offering of webinars and workshops to educate stakeholders on high-throughput screening data analysis and interpretation

Evolution of Computational Chemistry

Performance of Consensus ER QSAR Model with Literature Validation Set



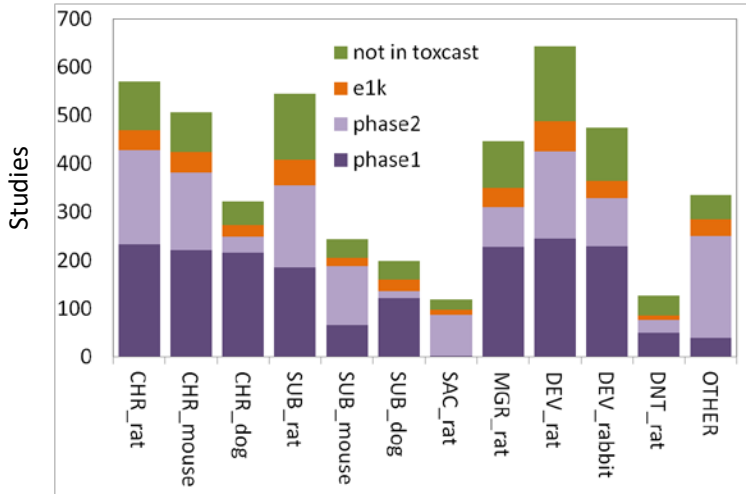
Present

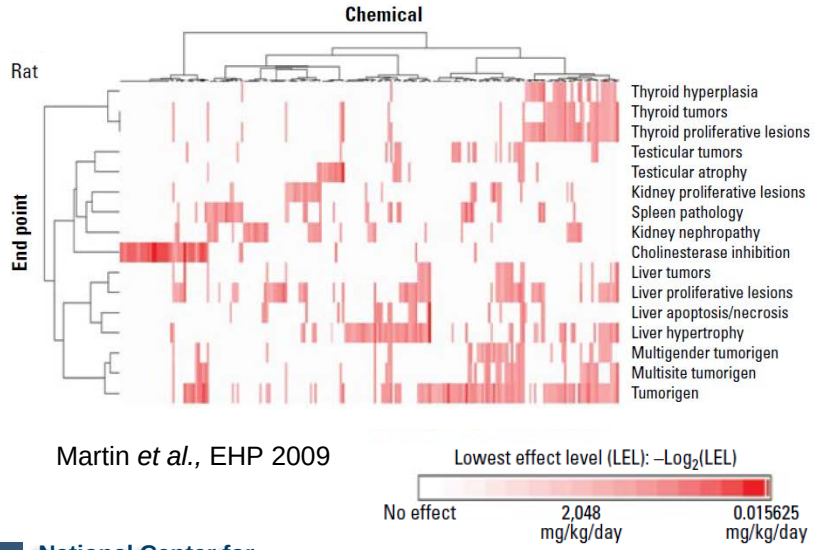
- High quality, curated chemical structure database of 22,000 molecules in DSSTox
- Development of an open-source KNIME workflow for automated structure processing and QC
- Community consensus QSAR models for endocrine-related bioactivity (ER and AR) with 17 international participants

Planned

- Increase curated chemical structure database to 150,000 molecules with addition of QC flags to indicate confidence in structural associations
- Chemical grouping and read-across effort that incorporates both chemical structure and bioactivity

Evolution of ToxRefDB





Martin *et al.*, EHP 2009

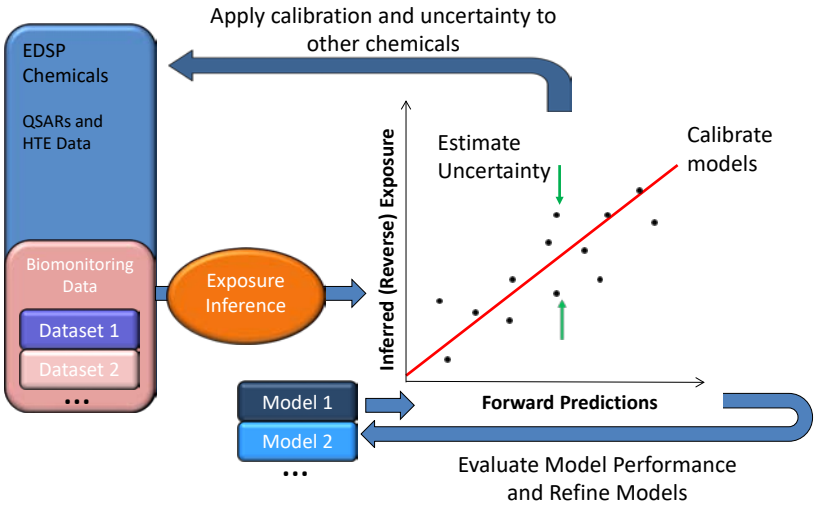
Present

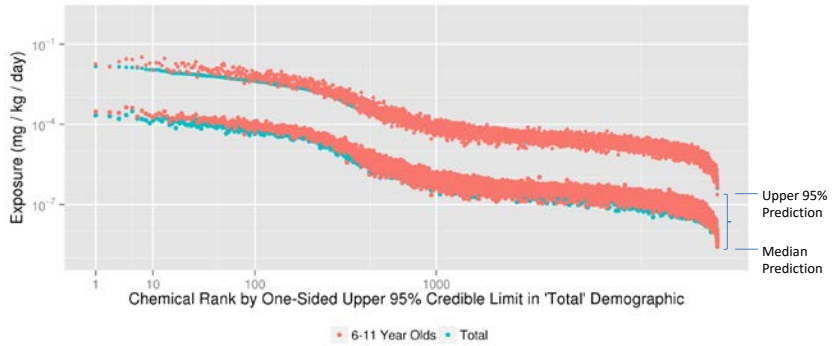
- *In vivo* endpoint data from 5,891 animal toxicology studies on ~1,110 unique chemicals
- A total of 654 chemicals (3311 studies) have LOAEL values, while 812 chemicals (2580 studies) have no adverse effect determination.
- Studies consist of chronic, subchronic, multigenerational, and developmental designs from DERs, NTP, open literature, and pharma

Planned

- Expand number of chemicals with critical effects and add quantitative dose-response data
- Add data from different sources (ECHA, E-Tox, Wignall et al.) and harmonize across databases
- Perform external QC review on ToxRefDB

Evolution of High-Throughput Exposure Modeling





Wambaugh *et al.*, EST 2014

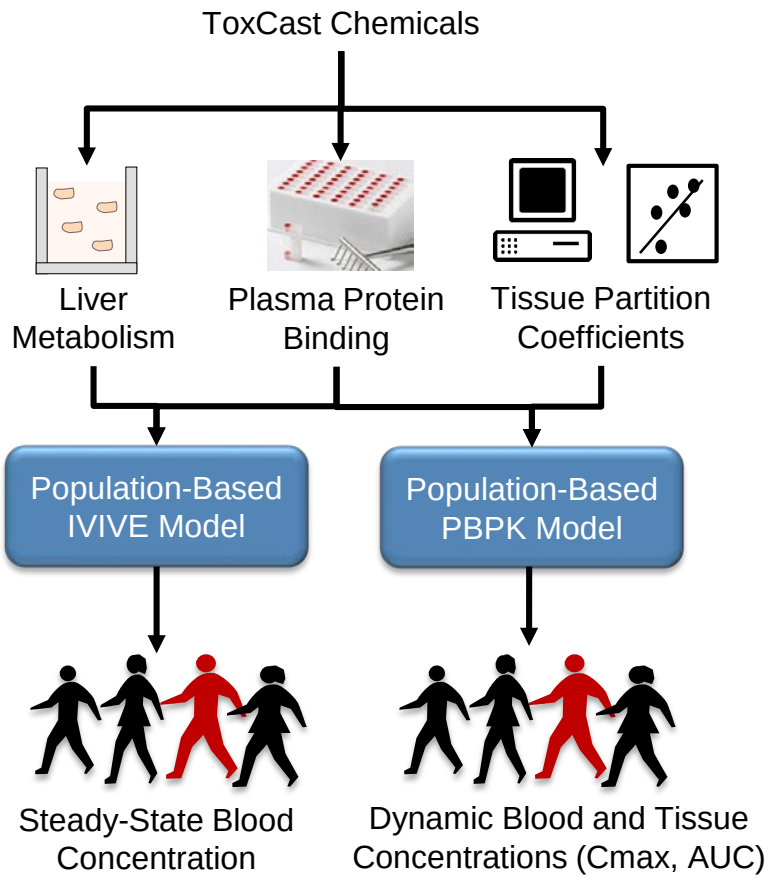
Present

- Computational exposure framework (SEEM) developed that provides exposure estimates for over 7,000 chemicals based on production volume and five chemical use categories
- A database of chemical-product categories (CPCat) that maps over 45,000 chemicals to ~8,000 product uses or functions

Planned

- Obtain additional biomonitoring data to expand domain of applicability
- Survey a range of consumer products and home goods for chemical ingredients and emissivity
- Incorporate SHEDS-HT into SEEM framework

Evolution of High-Throughput Toxicokinetics



Present

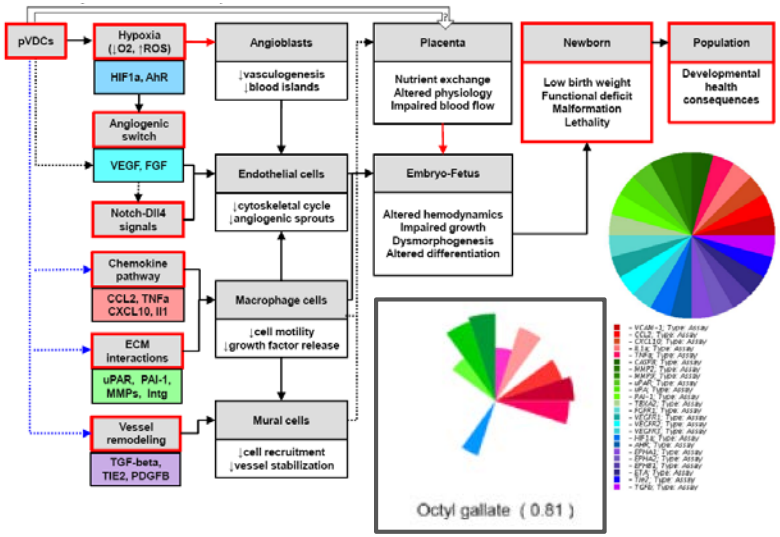
- Steady-state IVIVE models for hundreds of chemicals based on limited high-throughput *in vitro* assays
- Implemented structure-based methods to estimate tissue partitioning
- Developed HT-Physiologically-Based Pharmacokinetic (HT-PBPK) models for hundreds of chemicals

Planned

- Develop computational modeling framework to compare IVIVE with *in vivo* data and allow explicit estimates of uncertainty
- Separate chemical classes where we do either a good or poor job of predicting pharmacokinetics

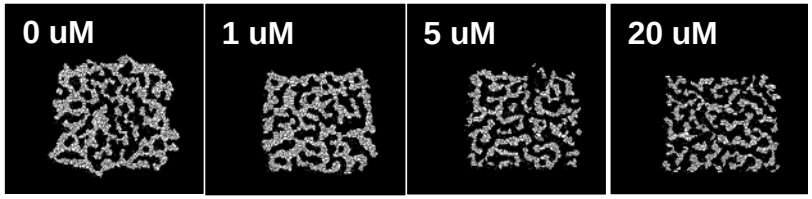
Evolution of Systems Modeling and Virtual Tissues

AOP for Developmental Vascular Disruption



Kleinstreuer *et al.*, PLoS Comp Bio, 2013

Model Simulations of Dev Vascular Disruption



Knudsen *et al.*, unpublished

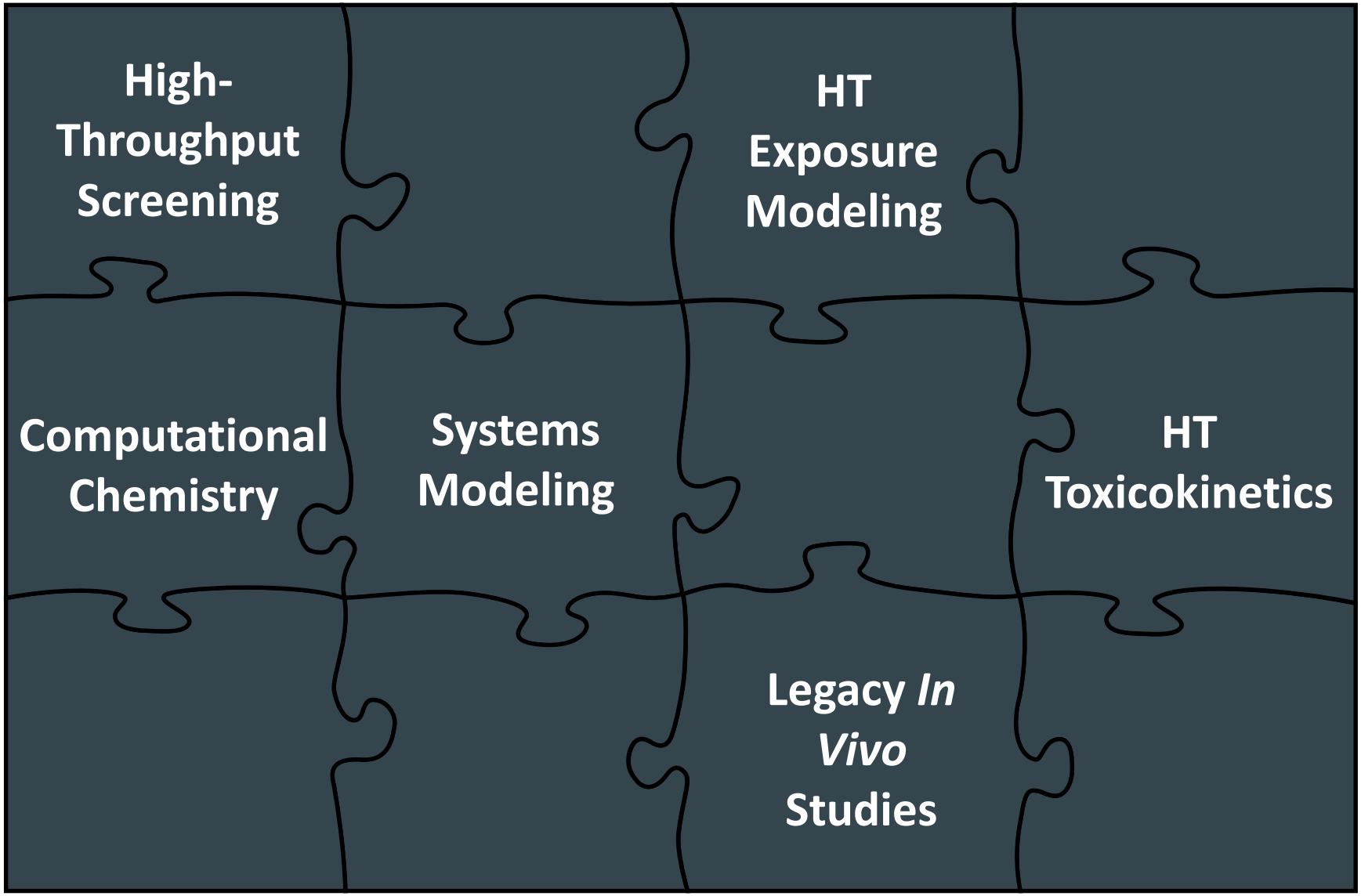
Present

- AOPs for embryonic vascular disruption, cleft palate, hypospadias, and limb (digit) defects
- Developed computational models for each AOP and used ToxCast data to parameterize models
- Validated model results with orthogonal organotypic assays and reference teratogens
- Developed first-generation methods for identifying cellular ‘tipping points’

Planned

- Refinement and validation of methods identifying concentration-dependent ‘tipping points’
- Evaluate impact of model assumptions on shape of the dose response curve and model potential sensitive populations

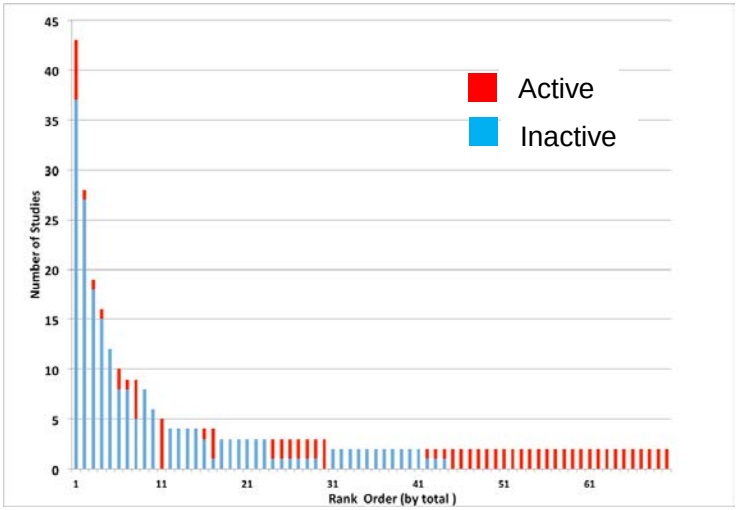
Fitting the Pieces Together for the Next Evolutionary Chapter

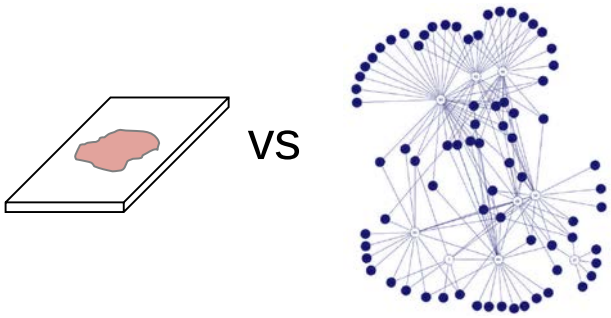


Fit-For-Purpose Agency Decisions

Challenges

Concordance of *In Vivo* Uterotrophic Studies





- Dealing with the "V" word
  - Defining a fit-for-purpose framework(s) that is time and resource efficient
  - Role of *in vivo* rodent studies
  - Incorporating the inherent uncertainty of the *in vivo* studies
- Moving from an apical to a molecular paradigm and defining adversity
- Predicting human safety vs. toxicity
- Integrating multiple data streams from the new approaches in a risk-based, weight of evidence assessment
- Ensuring a comprehensive screening and testing paradigm
- Quantifying uncertainty and variability
- Application to cumulative risk/mixtures

EPA United States Environmental Protection Agency

Questions?