The ToxCast Chemical Landscape: Paving the Road to 21st Century Toxicology

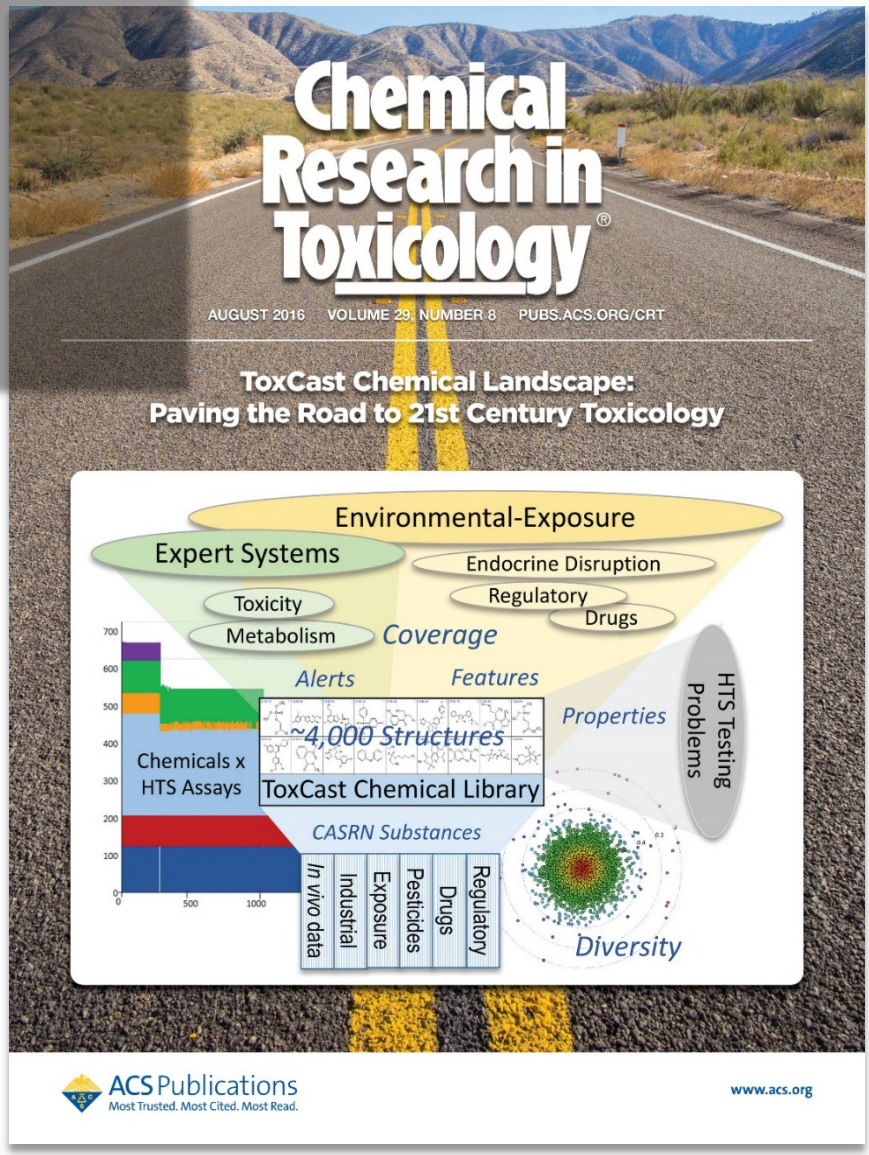
Ann Richard
National Center for Computational Toxicology
US Environmental Protection Agency, RTP, NC

CompTox Community of Practice Webinar,
December 15th, 2016

Resources

Open Access Perspectives article and Supporting Info files available for free download at:

http://pubs.acs.org/doi/abs/10.1021/acs.chemrestox.6b00135



DOI: 10.1021/acs.chemrestox.6b00135

ChemResToxicol., 2016, 29, 1225−1251

Purpose of ToxCast library

- To probe chemical-biological activity space potentially relevant to broad spectrum of toxicological outcomes of regulatory concern
- To generate HTS chemical-activity profiles to be used for developing predictive models of toxicity

I. History of library construction

- What were the main drivers and inputs?
- How did the library expand in phases over time?
- To what extent is physical library limited by practical constraints (i.e., procurable, testable)?
- What are quality concerns & how are they being addressed?

II. What's in the library?

- Chemical names
- CASRN

Chemical identifiers

- Bottles
- Solutions

Physical samples

- Structures
- Features

Chemical representations

III. Is library “fit for purpose”?

- Does library provide sufficient coverage of chemicals of interest to EPA & stakeholders?
- Does library include sufficient chemical diversity to span full range of toxicity mechanisms and outcomes of concern?
- Does library provide sufficient coverage of local regions of chemistry to enable local model development?

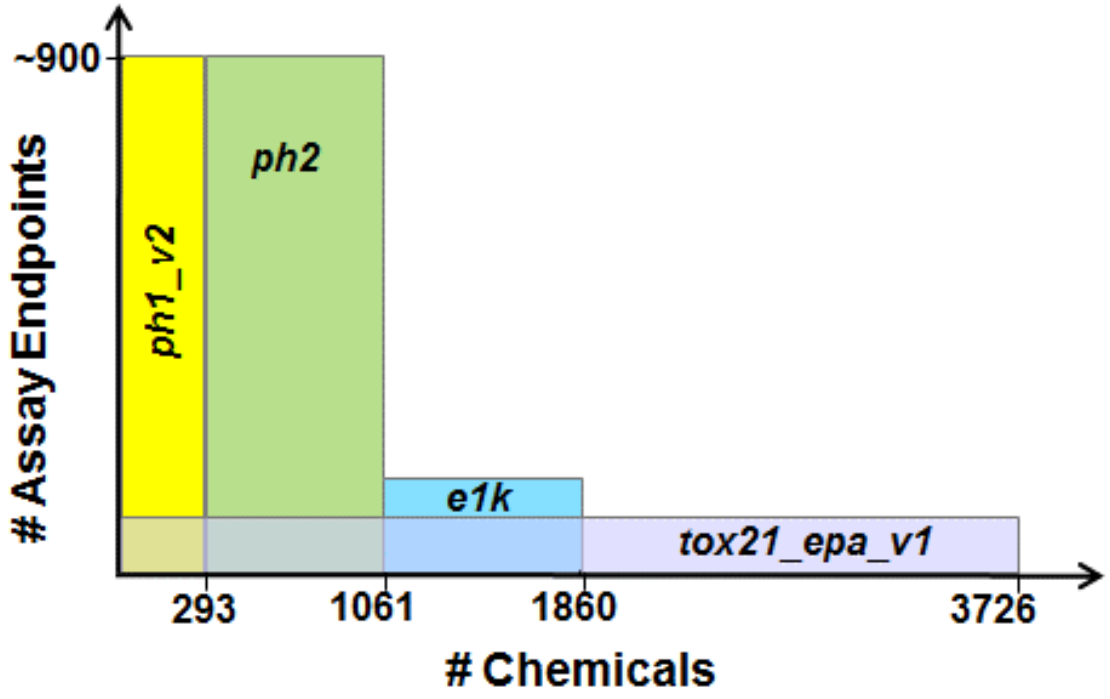
III. Is library "fit for purpose"?

- Does library provide sufficient coverage of chemicals of interest to EPA & stakeholders?
- Does library include sufficient chemical diversity to span full range of toxicity mechanisms and outcomes of concern?
- Does library provide sufficient coverage of local regions of chemistry to enable local model development?

… relative to the "chemical universe" and target inventories of greatest interest and concern to EPA

I. History - ToxCast inventory thru end of Testing Phase II

Testing Phase	Chemical Set	Unique Chemicals	Assay Endpoints
ToxCast Phase I	*ph1_v1*	310	~700
ToxCast Phase II	*ph1_v2*	293	~200[b]
	ph2	768	~900[c]
	e1k	799	~50[d]
Tox21	*tox21_epa_v1*	3726	~80[e]



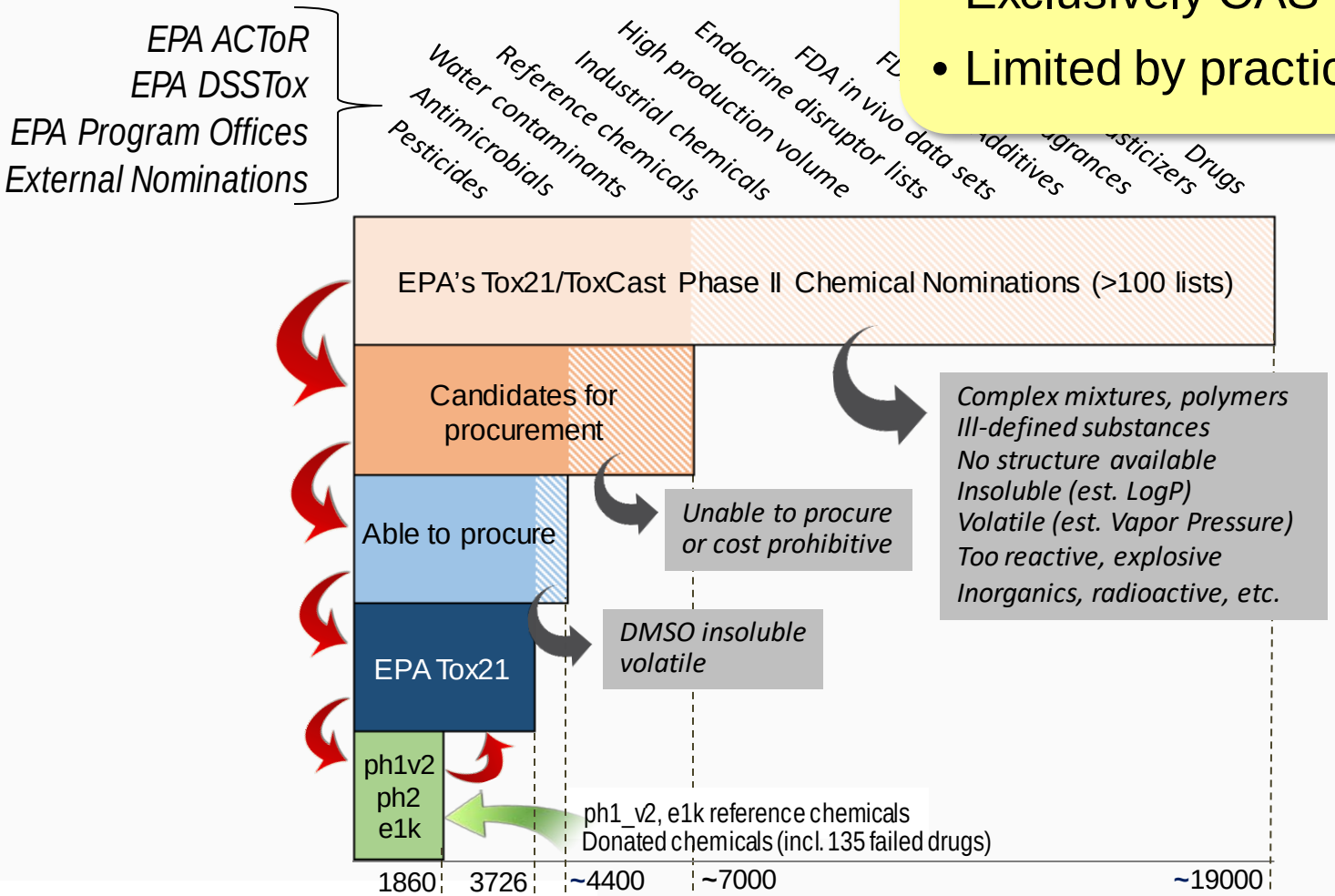
Moving from Phase I to Phase II:

- eliminated 17 chemicals
- reprocured ph1 inventory (v2), run in new assays
- full assay coverage of ph2, new & old assays
- limited assay coverage of e1k (endocrine only)
- broader chemical and less assay coverage in tox21

ChemResToxicol., 2016, 29, 1225−1251

Expanding ToxCast Library into Phase II & Tox21

- EPA & stakeholder inputs
- Exclusively CAS-name lists
- Limited by practical constraints

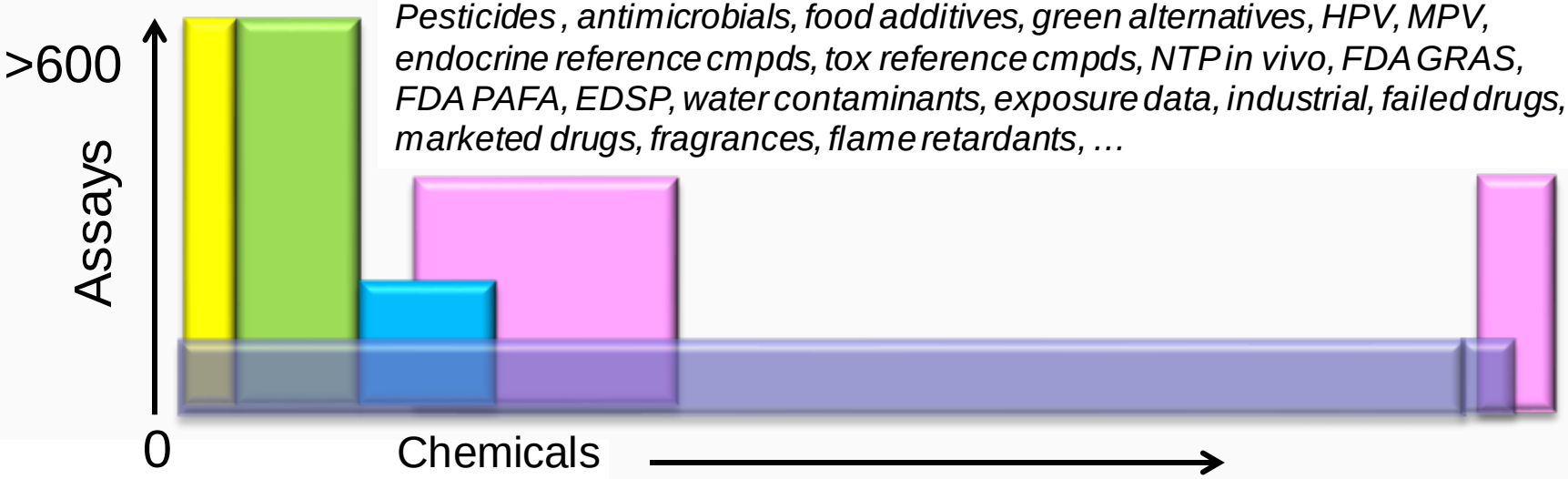


ChemResToxicol., 2016, 29, 1225−1251

Expanding into Phase III

Testing Phase	Chemical Set	Chemicals	Assay Endpoints	Completion
ToxCast Phase I	ph1_v1	310	~700	2011
ToxCast Phase II	ph1_v2	293	~200	2013
	ph2	768	~900	
	E1k	799	~50	
Tox21	tox21	~8900	~80	Ongoing
ToxCast Phase III	ph3	~2000	~300	Ongoing

Pesticides , antimicrobials, food additives, green alternatives, HPV, MPV, endocrine reference cmpds, tox reference cmpds, NTP in vivo, FDA GRAS, FDA PAFA, EDSP, water contaminants, exposure data, industrial, failed drugs, marketed drugs, fragrances, flame retardants, …

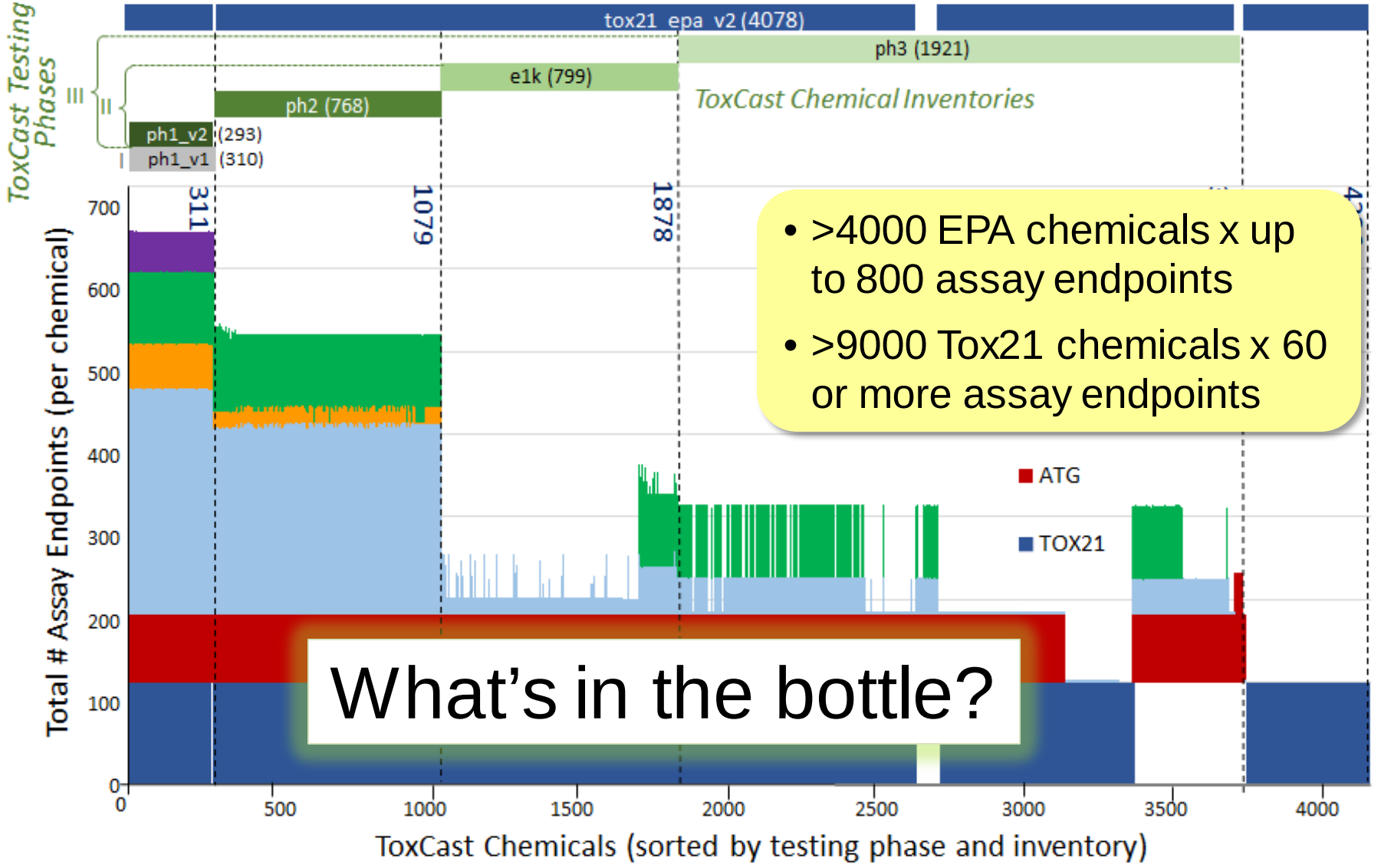


ToxCast chemical x assay counts

(Top 5 assay providers & Tox21, as of Jan 2016)

EPA United States Environmental Protection Agency

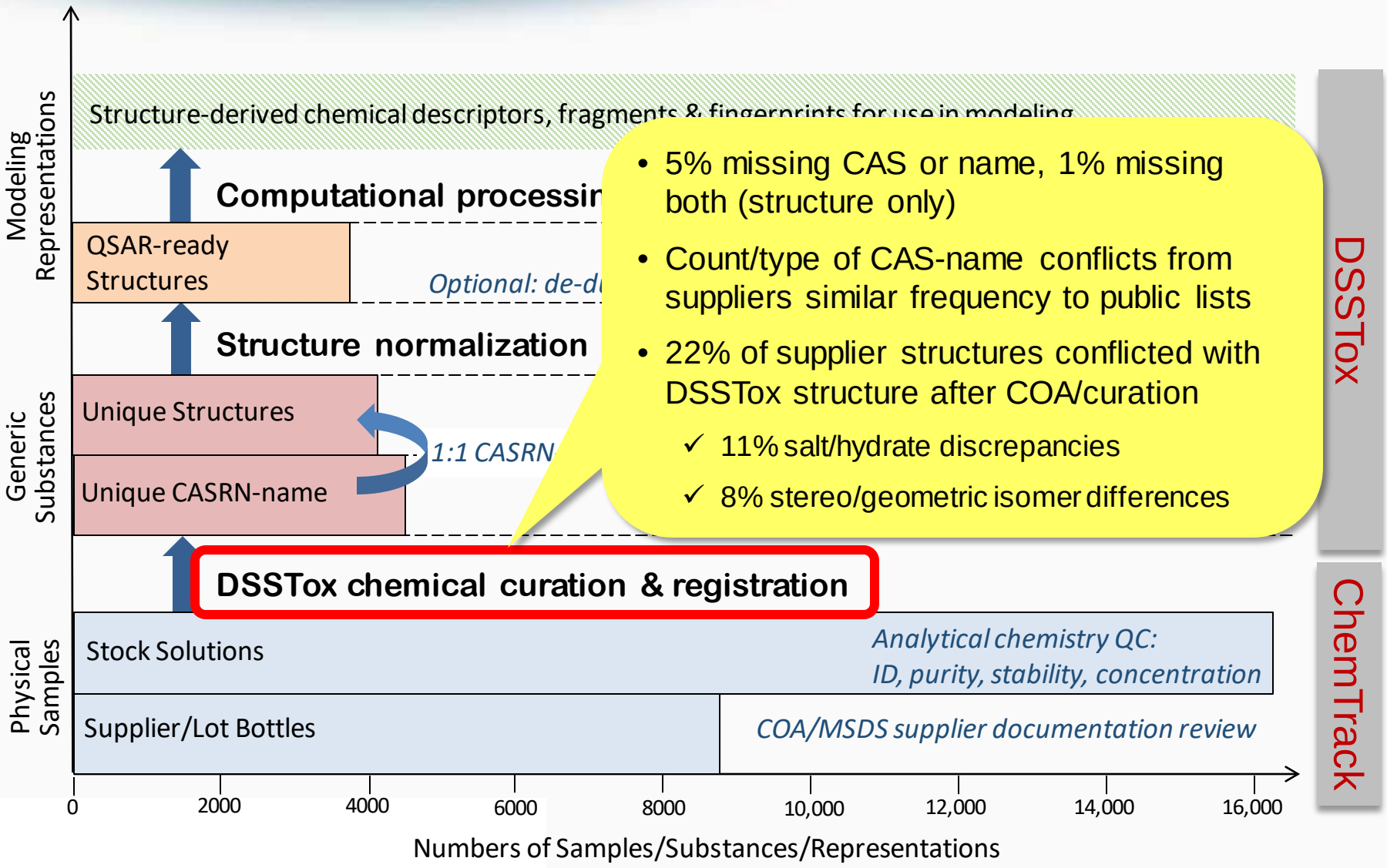
ChemResToxicol., 2016, 29, 1225−1251



- >4000 EPA chemicals x up to 800 assay endpoints
- >9000 Tox21 chemicals x 60 or more assay endpoints

What's in the bottle?

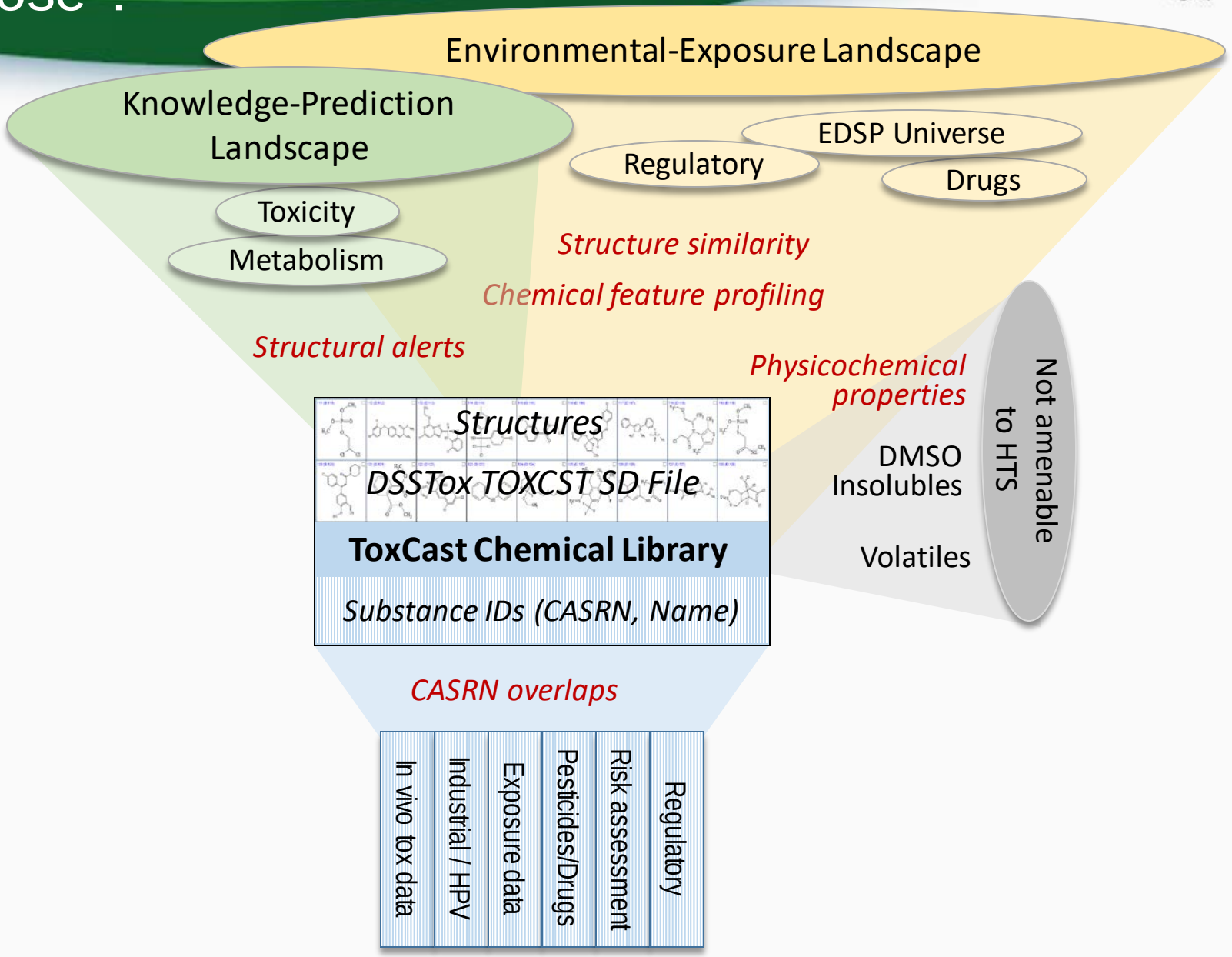
How are we addressing quality concerns?



- 5% missing CAS or name, 1% missing both (structure only)
- Count/type of CAS-name conflicts from suppliers similar frequency to public lists
- 22% of supplier structures conflicted with DSSTox structure after COA/curation
  - ✓ 11% salt/hydrate discrepancies
  - ✓ 8% stereo/geometric isomer differences

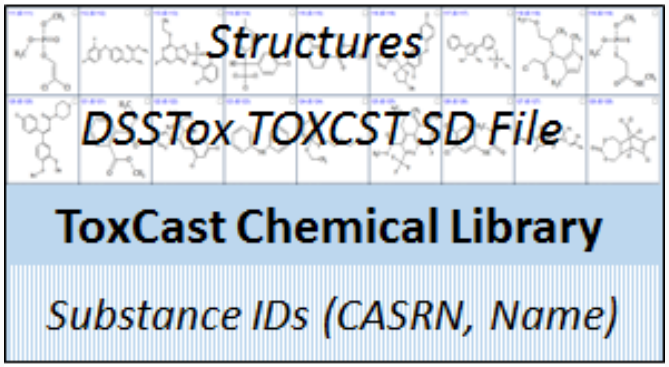
ChemResToxicol., 2016, 29, 1225−1251

What's in the library & is it "fit for purpose"?



ChemResToxicol., 2016, 29, 1225–1251

What's in the library?



- Unique list of DSSTox substances (e.g., CAS, names)
- Structures (mol, InChI, SMILES) annotated to salt/hydrate/stereo-specific form
- Inventory (ph1_v1,v2,p2,etc) and Testing Phase (I,II,III) labels

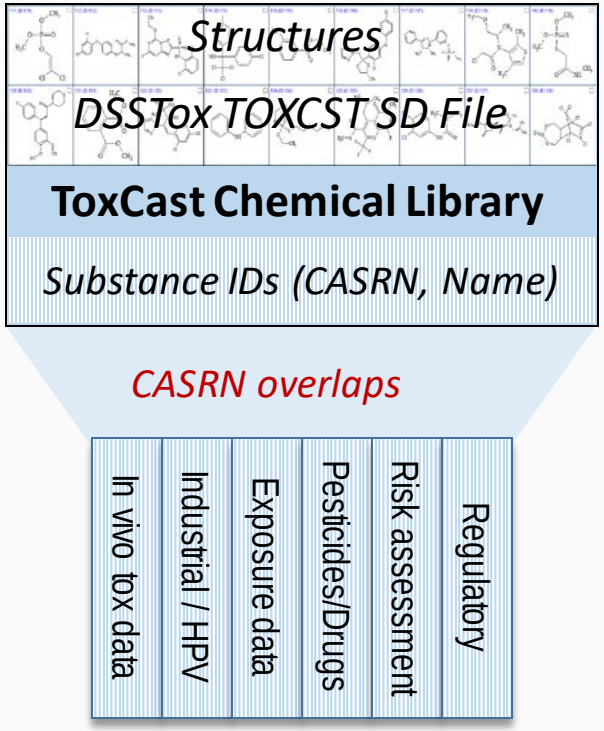
Generic Substances	4226
Structures	4056
CASRN	4134
Salt or Complex	459
Inorganic	48
Organometallic	110
Mixture of stereoisomers	113
Mixture/Formulation	157
Polymer	11
Single chemical	3945
Duplicates on desalting	202

Available at: ftp://ftp.epa.gov/dsstoxftp/DSSTox_TOXCST_20160129.zip

What's in the library?

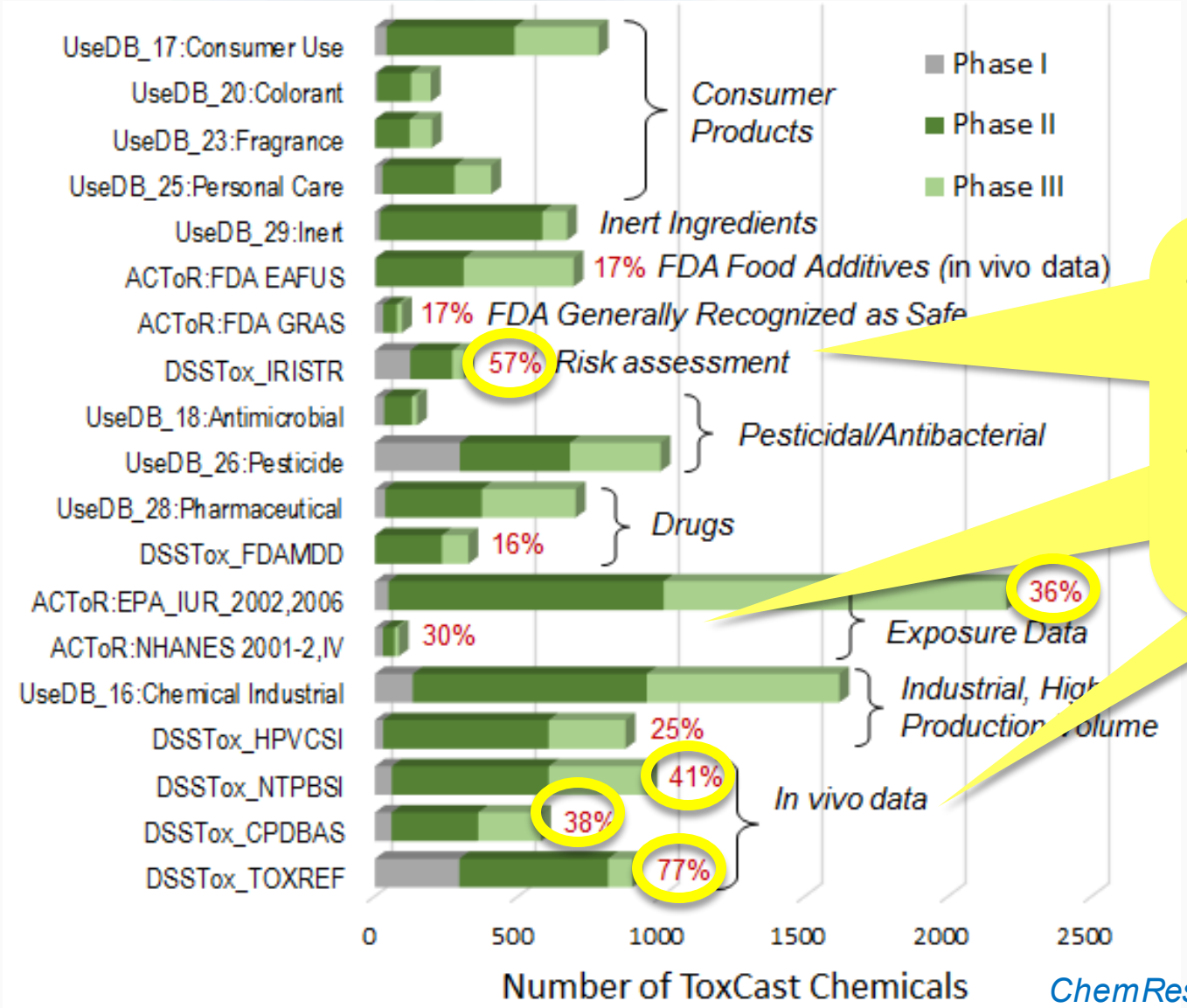
- CAS lists used to nominate chemicals for Phase II and Tox21

  → Evaluate TOXCST coverage of high priority CAS lists



- Overlap requires exact CAS matches
- Chemical structure not considered
- NOCAS substances not considered

What's in the library?
Data & Usage List coverage

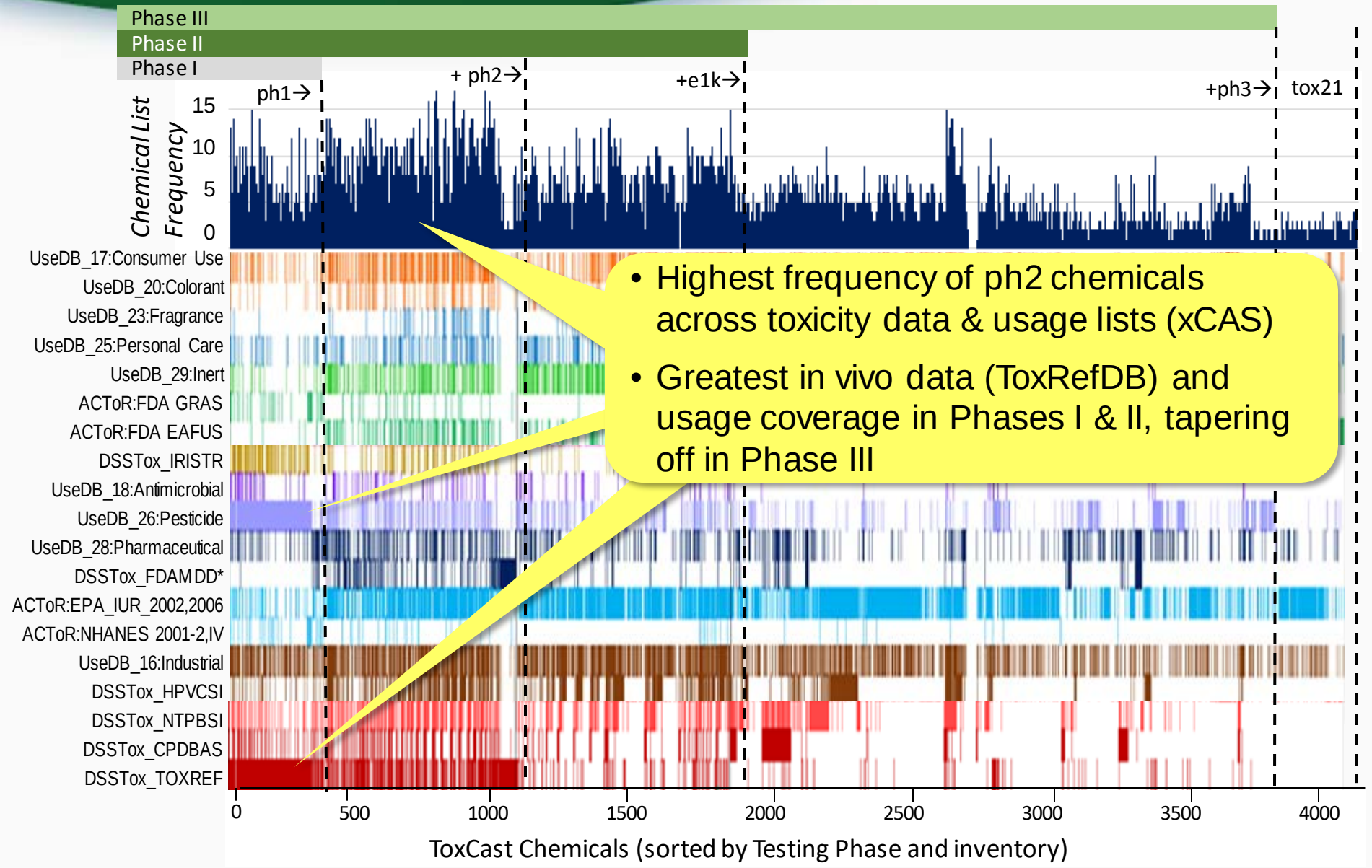


- Increasing list coverage moving from Phase I → II,III
- High coverage of in vivo, exposure, & risk assessment data lists

ChemResToxicol., 2016, 29, 1225−1251

What's in the library?
Data & Usage List coverage

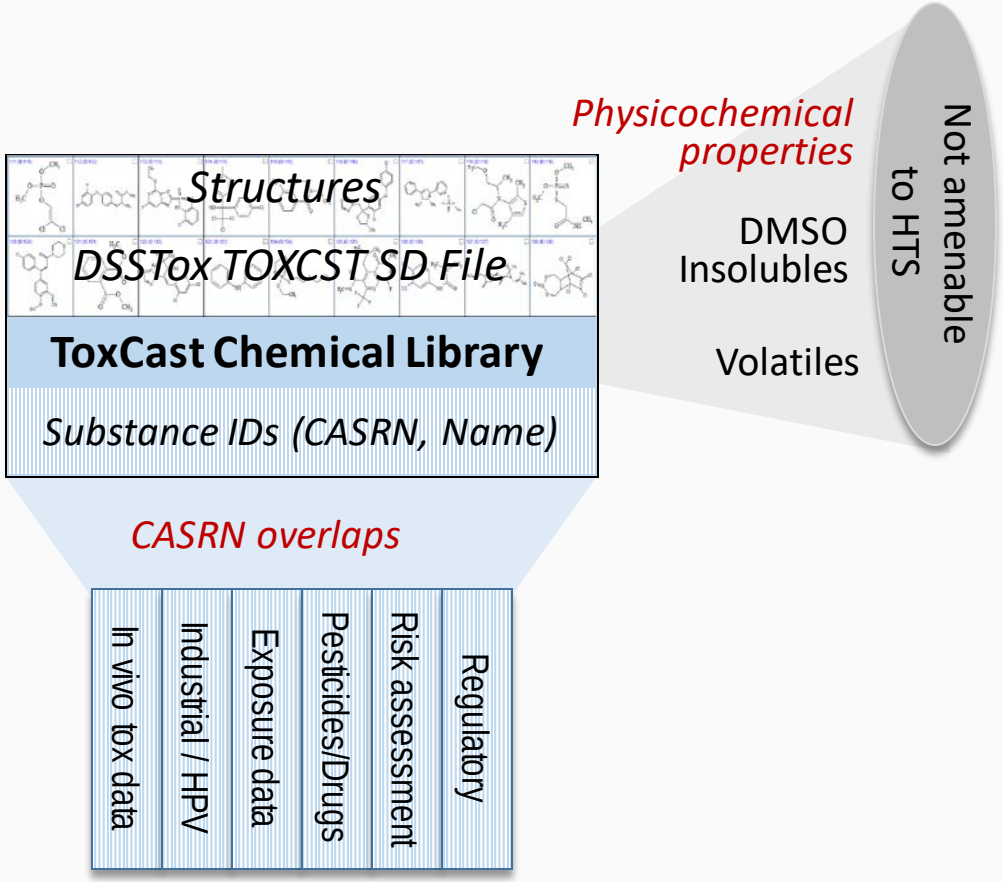
ChemResToxicol., 2016, 29, 1225−1251



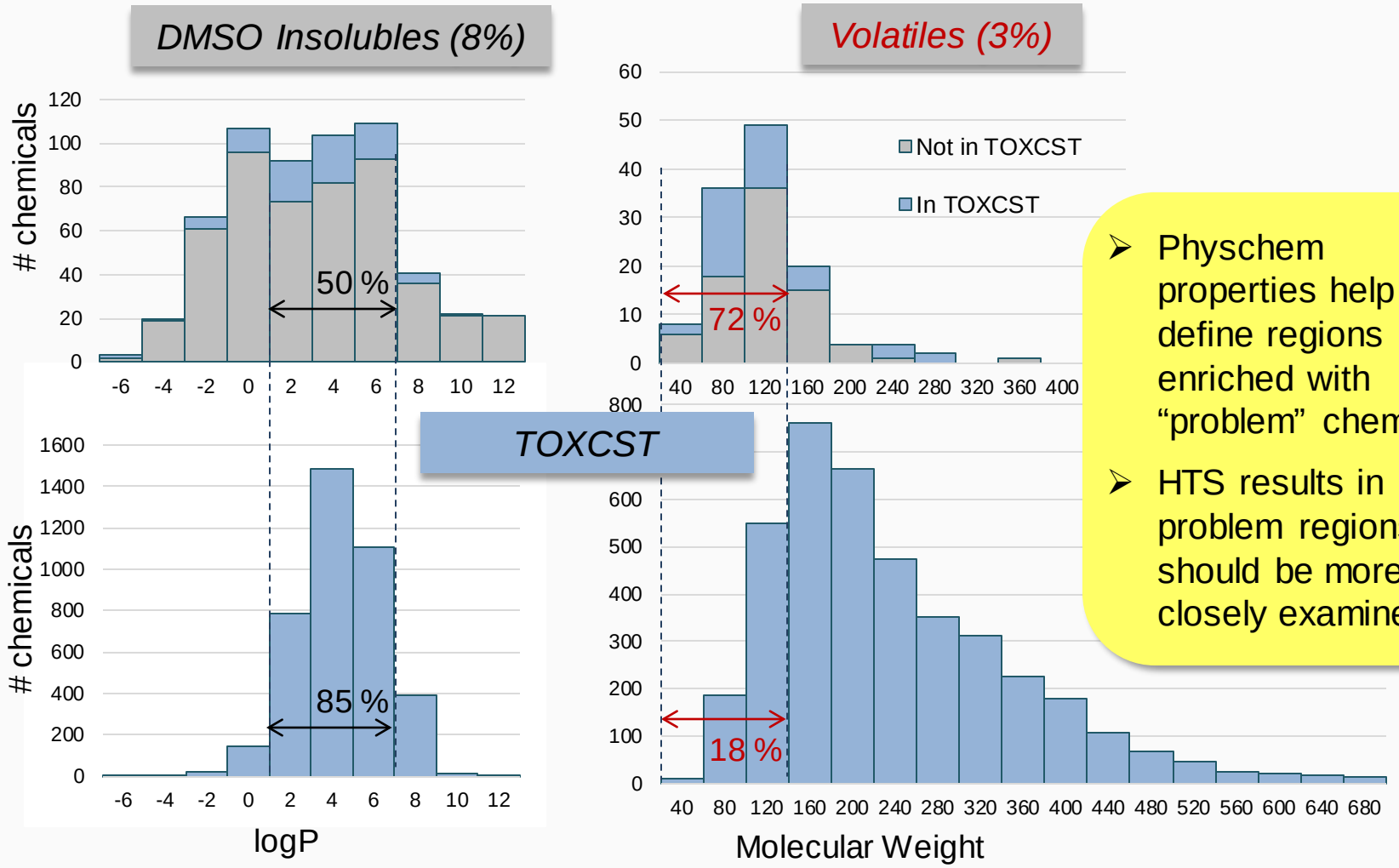
- Highest frequency of ph2 chemicals across toxicity data & usage lists (xCAS)
- Greatest in vivo data (ToxRefDB) and usage coverage in Phases I & II, tapering off in Phase III

What's not in the library?

- To what extent is HTS library bounded by practical constraints?
  - DMSO solubility
  - Volatility



What's not in the library?

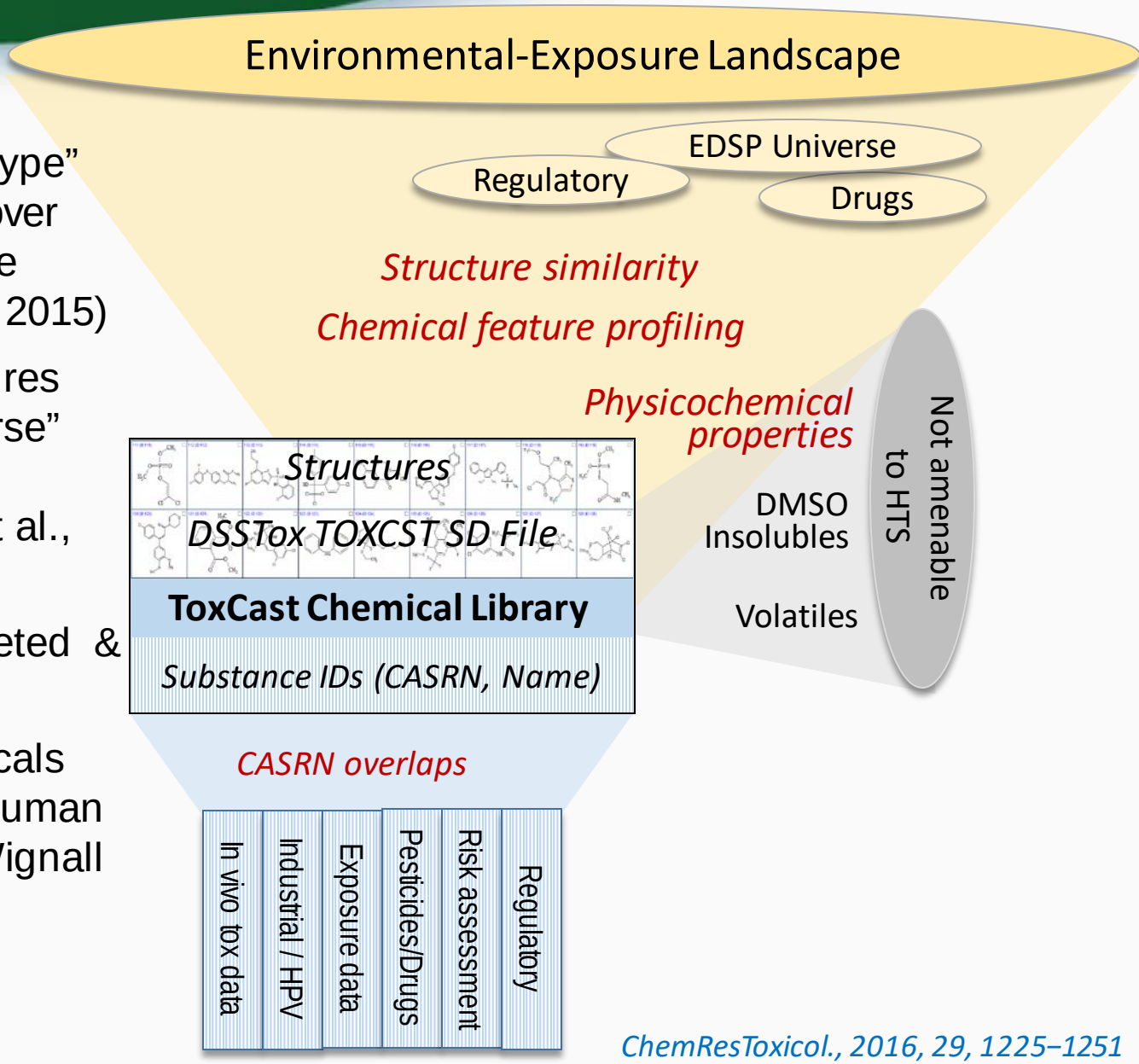


- Physchem properties help to define regions enriched with "problem" chemicals
- HTS results in problem regions should be more closely examined

ChemResToxicol., 2016, 29, 1225−1251

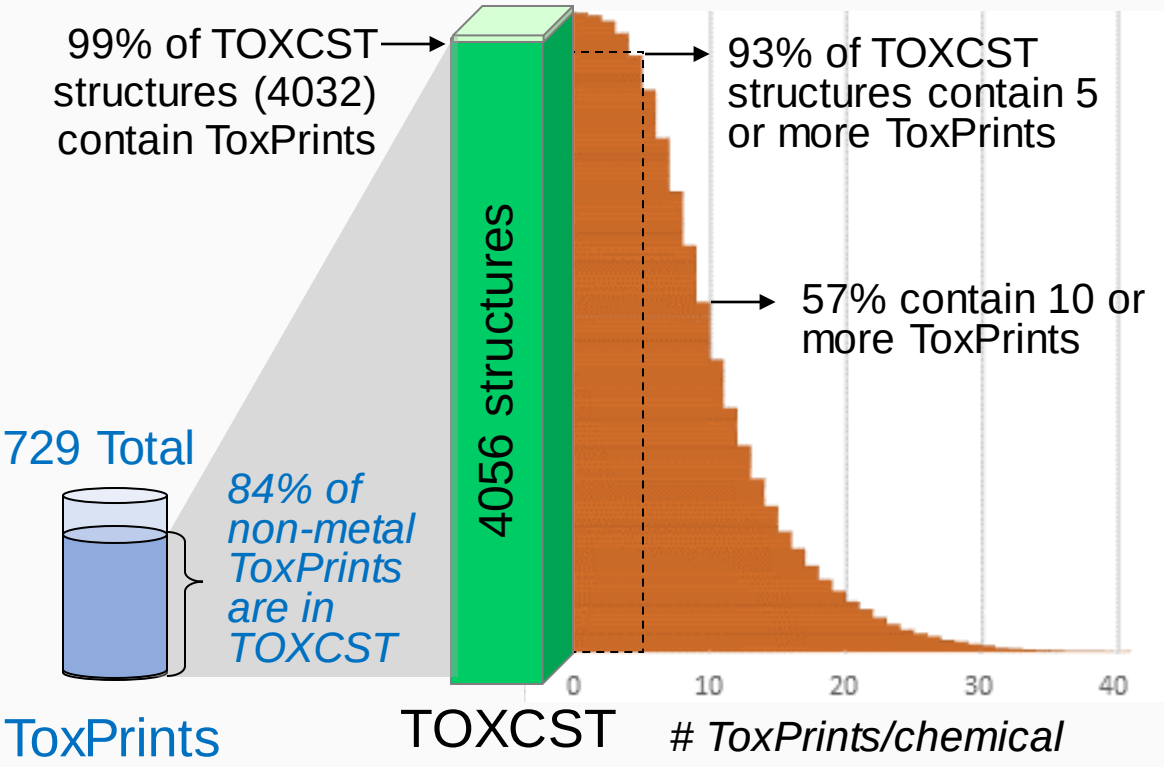
Evaluate coverage of potential "target" inventories

- ToxPrints: 792 "chemotype" features designed to cover environmental-exposure landscape (Yang et al., 2015)
- CERAPP: ~35K structures spanning EDSP "universe" and putative exposure landscape (Mansouri et al., 2016)
- FDA_Drugs: ~7K marketed & discontinued drugs
- BMDHHA: ~800 chemicals with benchmark dose human health assessments (Wignall et al., 2014)



ChemResToxicol., 2016, 29, 1225−1251

ToxPrint vs TOXCST:
Assessing coverage & diversity



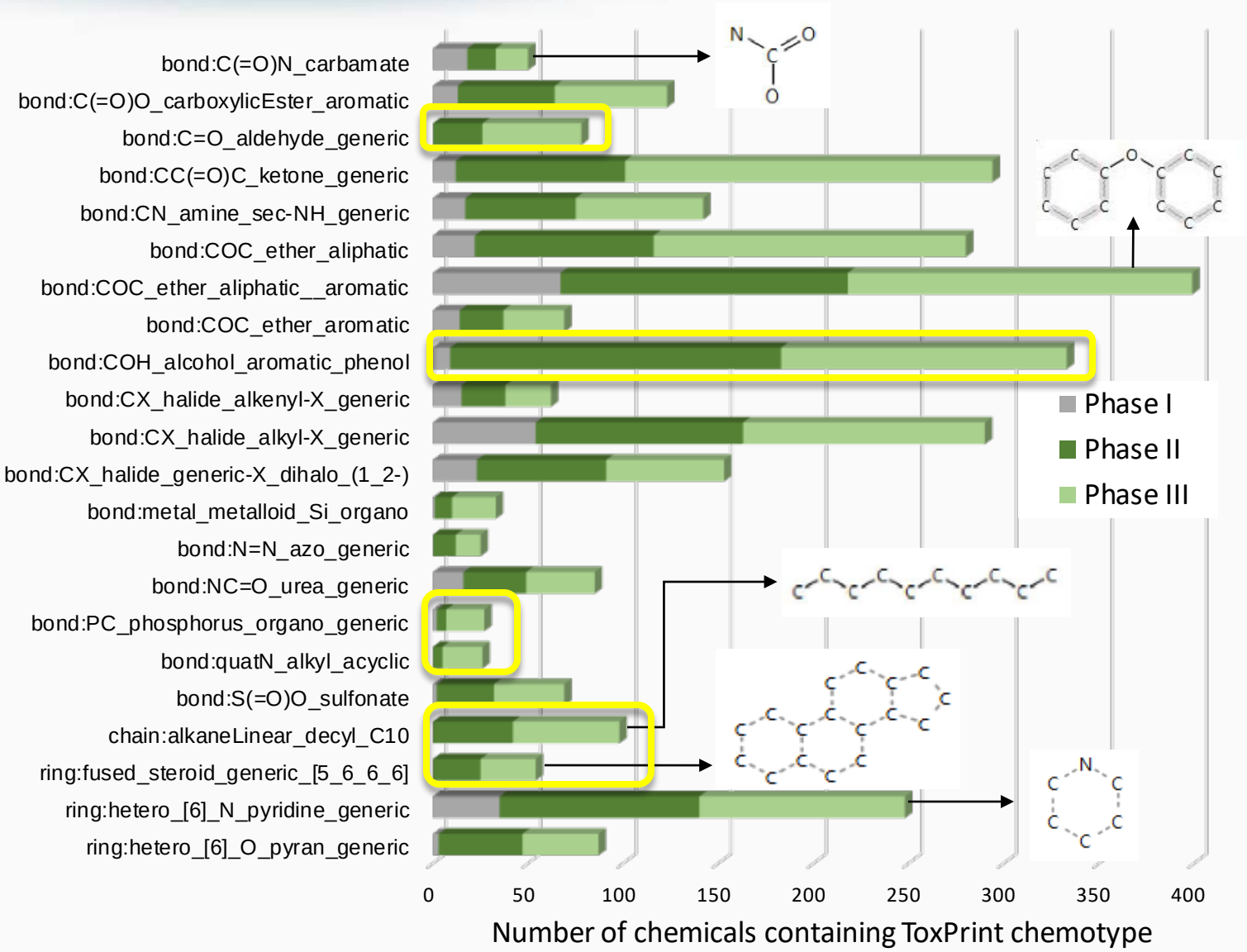
ToxPrints provide excellent "coverage" and suggest large structural diversity of TOXCST inventory

Yang et al., 2015: https://toxprint.org/

ChemResToxicol., 2016, 29, 1225–1251

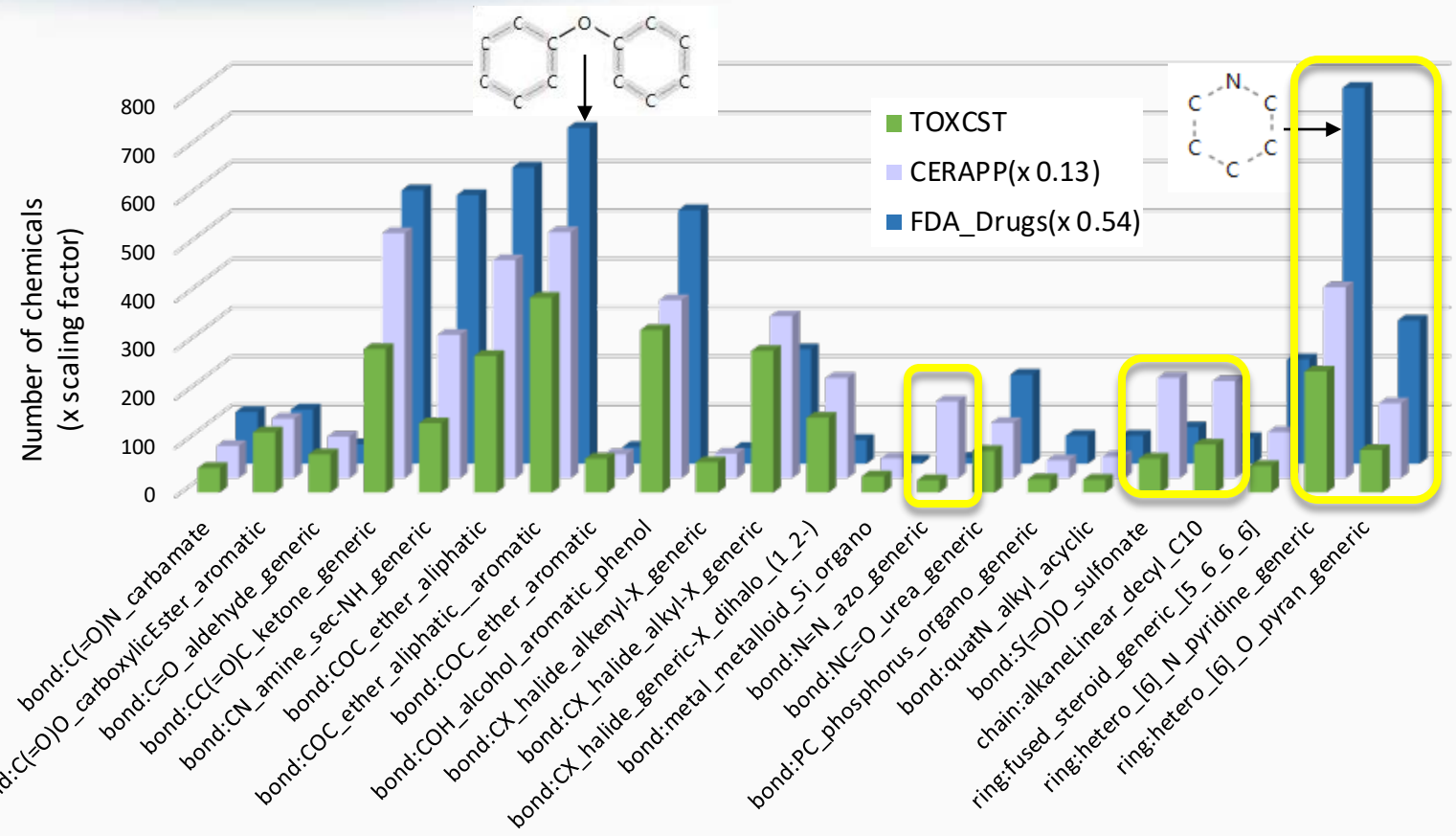
Coverage of ToxPrints across testing phases

ChemResToxicol., 2016, 29, 1225–1251



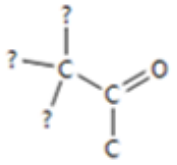
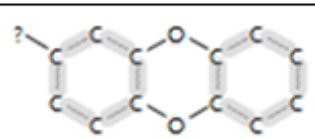
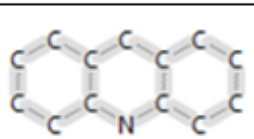
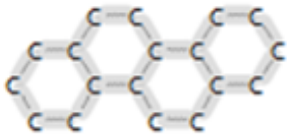
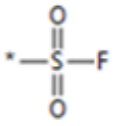
ToxPrint inventory profile comparisons (scaled)

ChemResToxicol., 2016, 29, 1225–1251



- Similar global ToxPrint profiles
- Some local feature distinctions:
  - features enriched in drugs, e.g. pyridine, pyran rings
  - CERAPP features not well represented in TOXCST, e.g. azo, sulfonate bonds, decyl chains

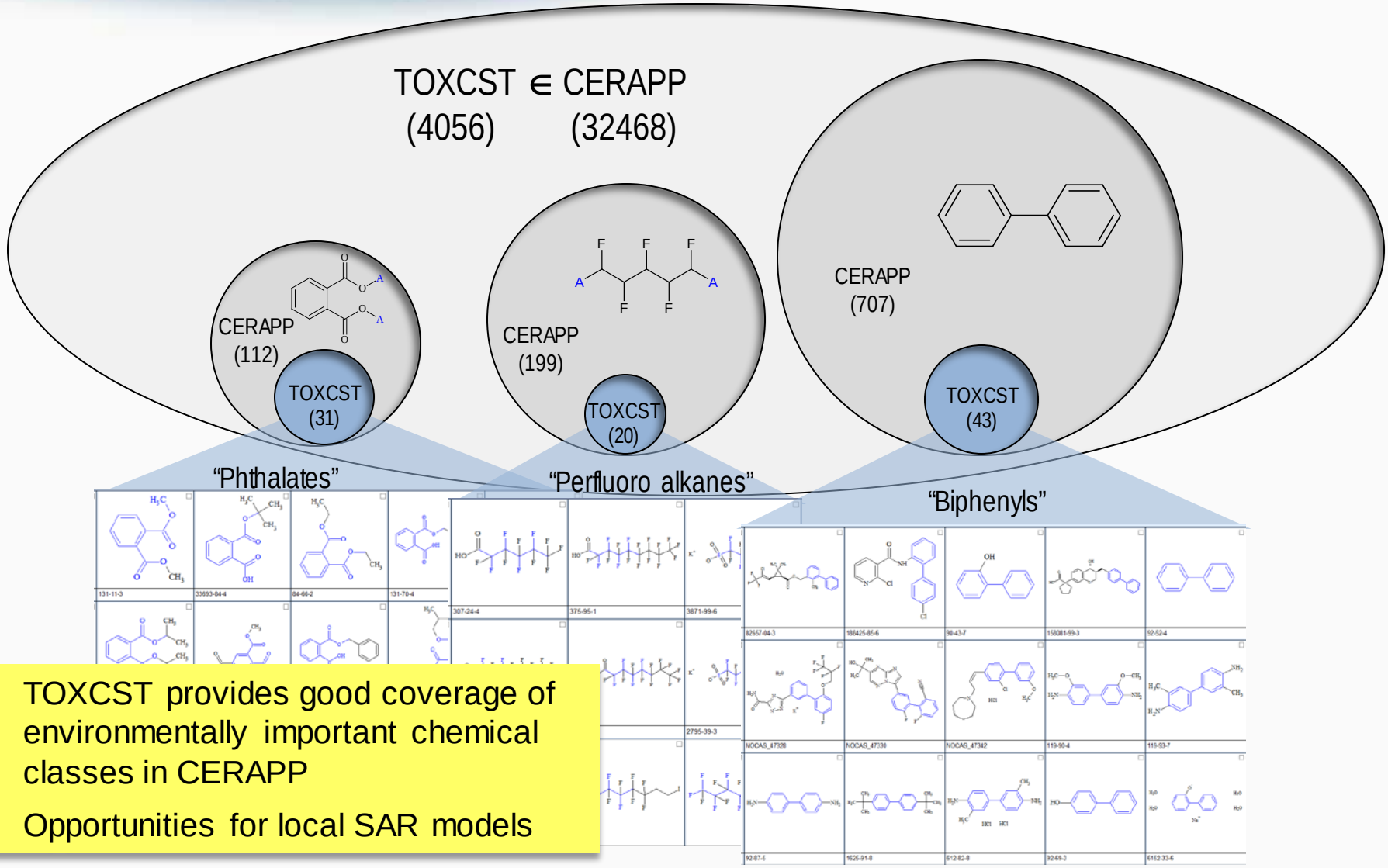
ToxPrints in CERAPP not present in ToxCast Library

bond:C(~Z)~C~Q_a-haloketone_perhalo	 21
Hexachloroacetone	
4,4,4-Trifluoro-1-phenyl-1,3-butanedione	
Hexafluoroacetone	
1,1,1-Trichloropropanone	
bond:C(~Z)~C~Q_haloether_dibenzodioxin_2-halo	 34
1,2,3,6,7,8-Hexachlorodibenzo-p-dioxin	
Octabromodibenzo-p-dioxin	
1,2,3,7,8-Pentabromodibenzo-p-dioxin	
ring:hetero_[6_6_6]_N_acridine	 34
13-Docosenamide	
9-Octadecen-1-amine	
ring:fused_PAH_benzophenanthrene	 19
Benzo(a)pyrene	
Naphtho(1,2,3,4-def)chrysene	
Chrysene	
3-Hydroxybenz[a]pyrene	
bond:S(=O)X_sulfonylhalide_fluoride	 15
1-Octanesulfonyl fluoride	
Benzenemethanesulfonyl fluoride	
Trifluoromethanesulfonyl fluoride	

- Are the missing features present in environmental chemicals?
- Why were these chemicals not included in ToxCast?
- Use to expand ToxCast chemical coverage moving forward

ChemResToxicol., 2016, 29, 1225−1251

Coverage of 3 chemical classes: ToxCast vs CERAPP

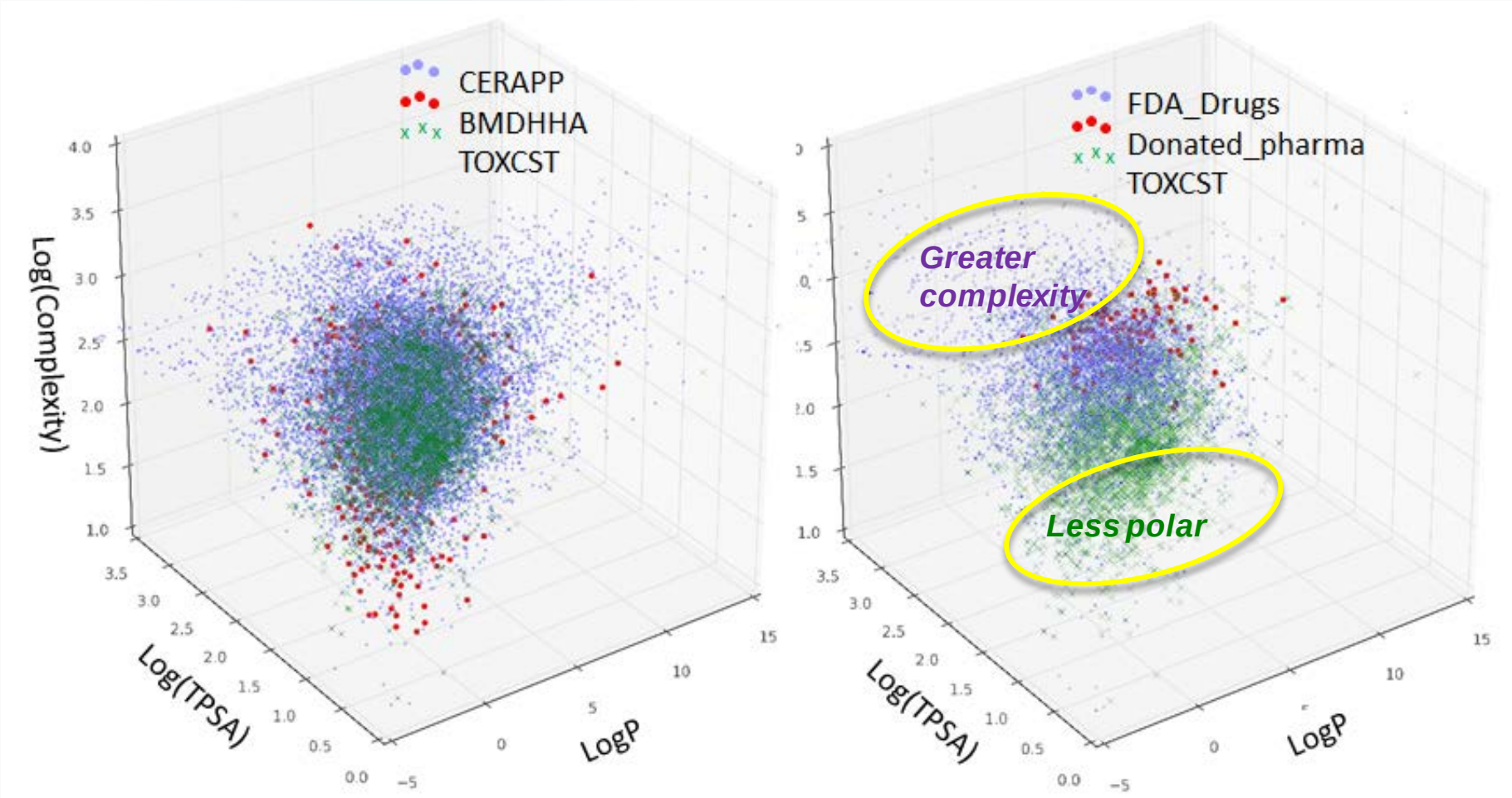


- TOXCST provides good coverage of environmentally important chemical classes in CERAPP
- Opportunities for local SAR models

Chem Res Toxicol., 2016, 29, 1225−1251

Comparison to potential target inventories based on computed properties

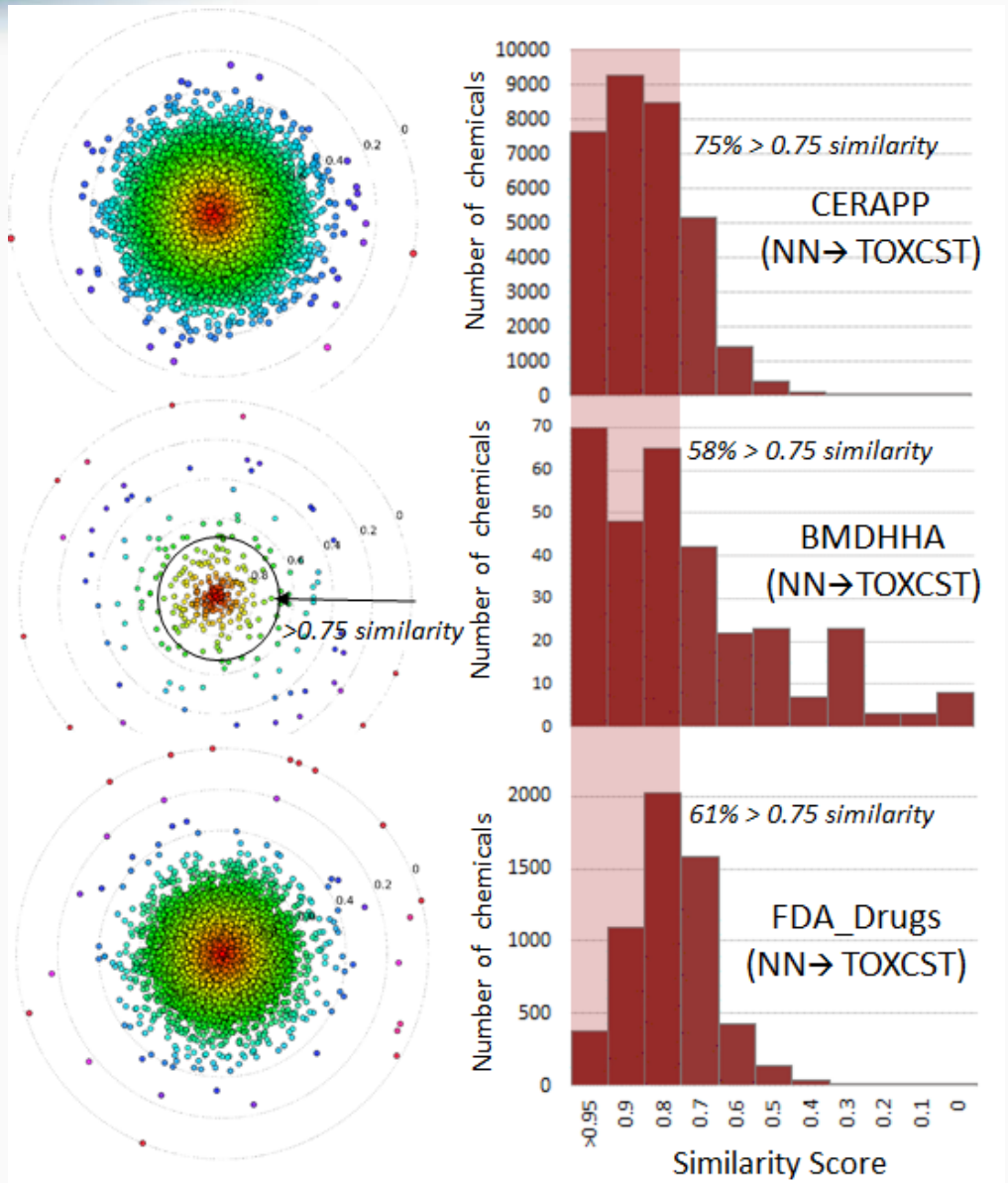
ChemResToxicol., 2016, 29, 1225−1251



- TOXCST more similar to CERAPP & BMDHHA inventories than to FDA_Drugs in physchem property space
- Donated_pharma not representative of drug space as a whole

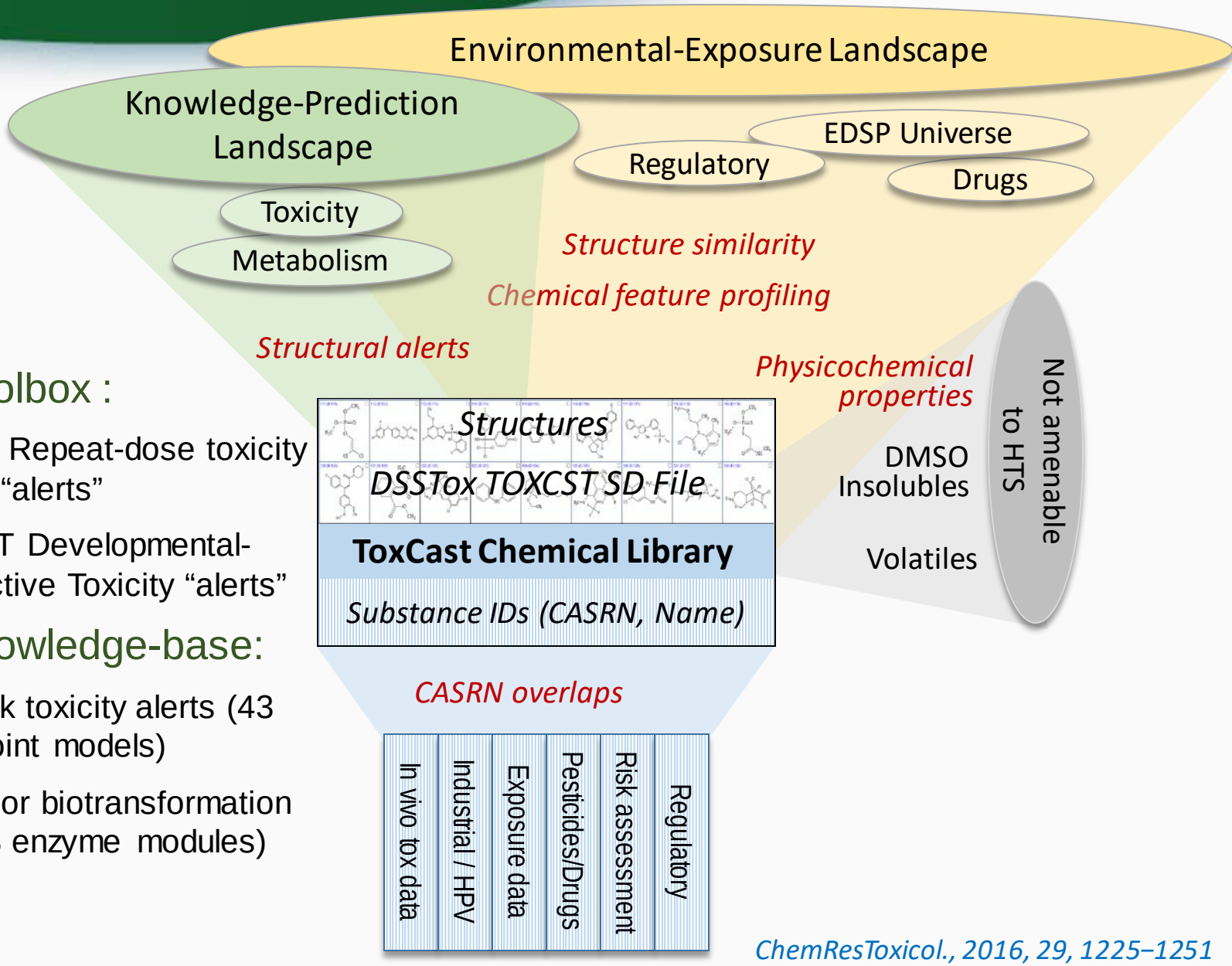
Nearest neighbor similarity comparisons (Tanimoto)

- 75% of CERAPP chemicals have a >75% similar TOXCST “analog”
- 58% of BMDHHA chemicals have a >75% similar TOXCST “analog
- 61% of FDA_Drugs chemicals have a >75% similar TOXCST “analog



ChemResToxicol., 2016, 29, 1225−1251

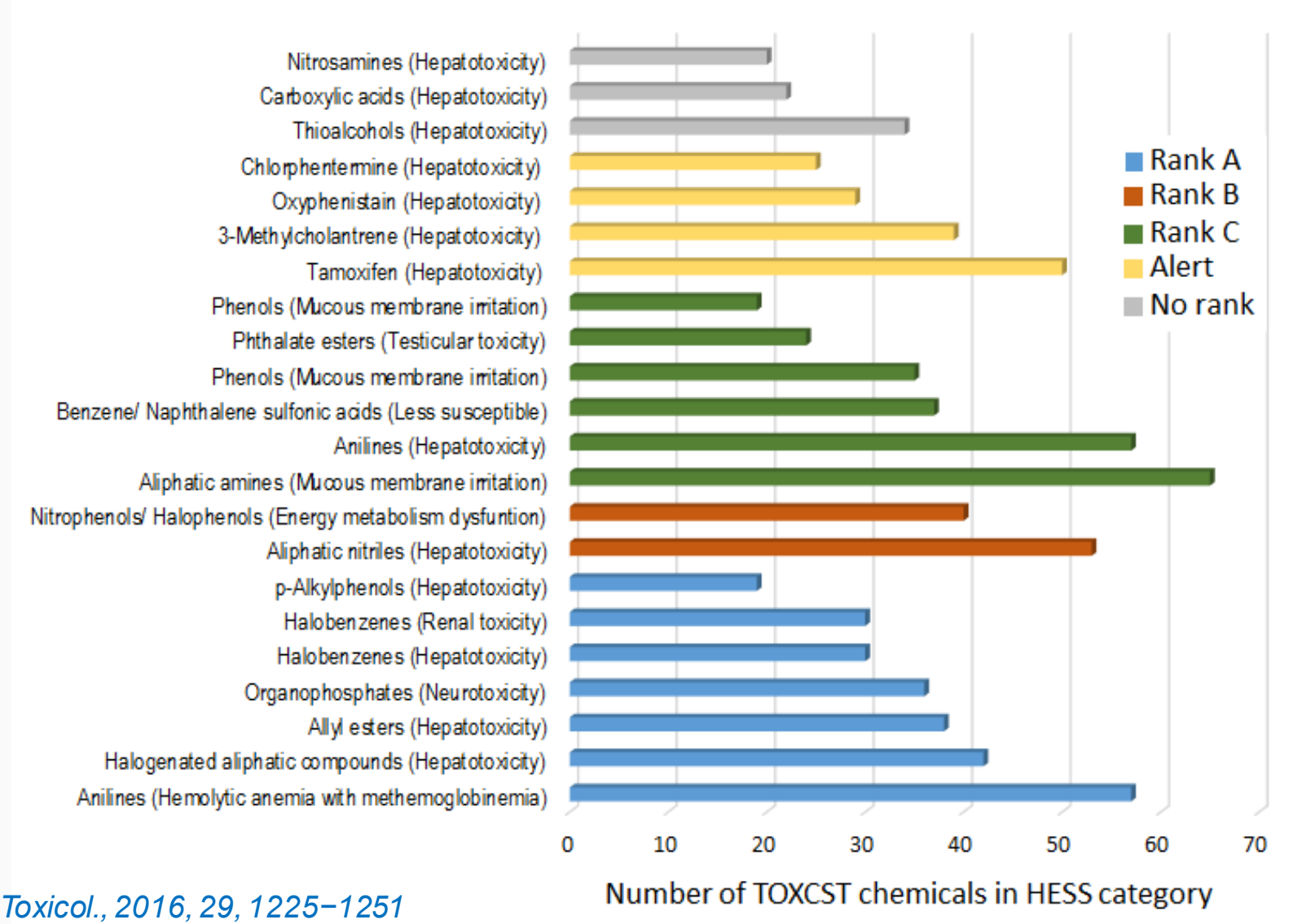
Evaluate coverage of historical SAR "alerts" knowledge



- OECD Toolbox :
  - 61 HESS Repeat-dose toxicity category "alerts"
  - 136 DART Developmental-Reproductive Toxicity "alerts"
- Lhasa Knowledge-base:
  - 280 Derek toxicity alerts (43 tox endpoint models)
  - 157 Meteor biotransformation alerts (33 enzyme modules)

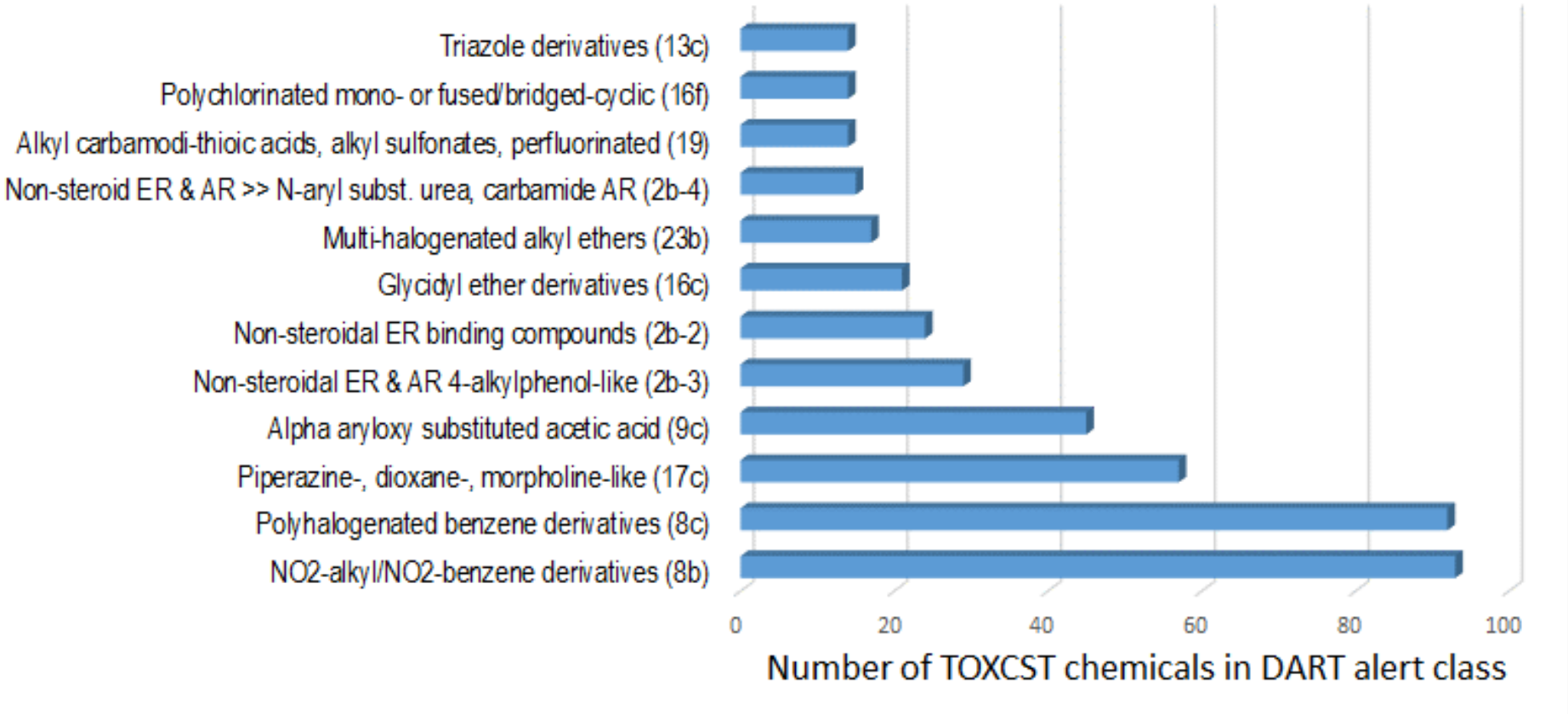
ChemResToxicol., 2016, 29, 1225–1251

Incidence of HESS repeat-dose toxicity alerts in ToxCast



ChemResToxicol., 2016, 29, 1225−1251

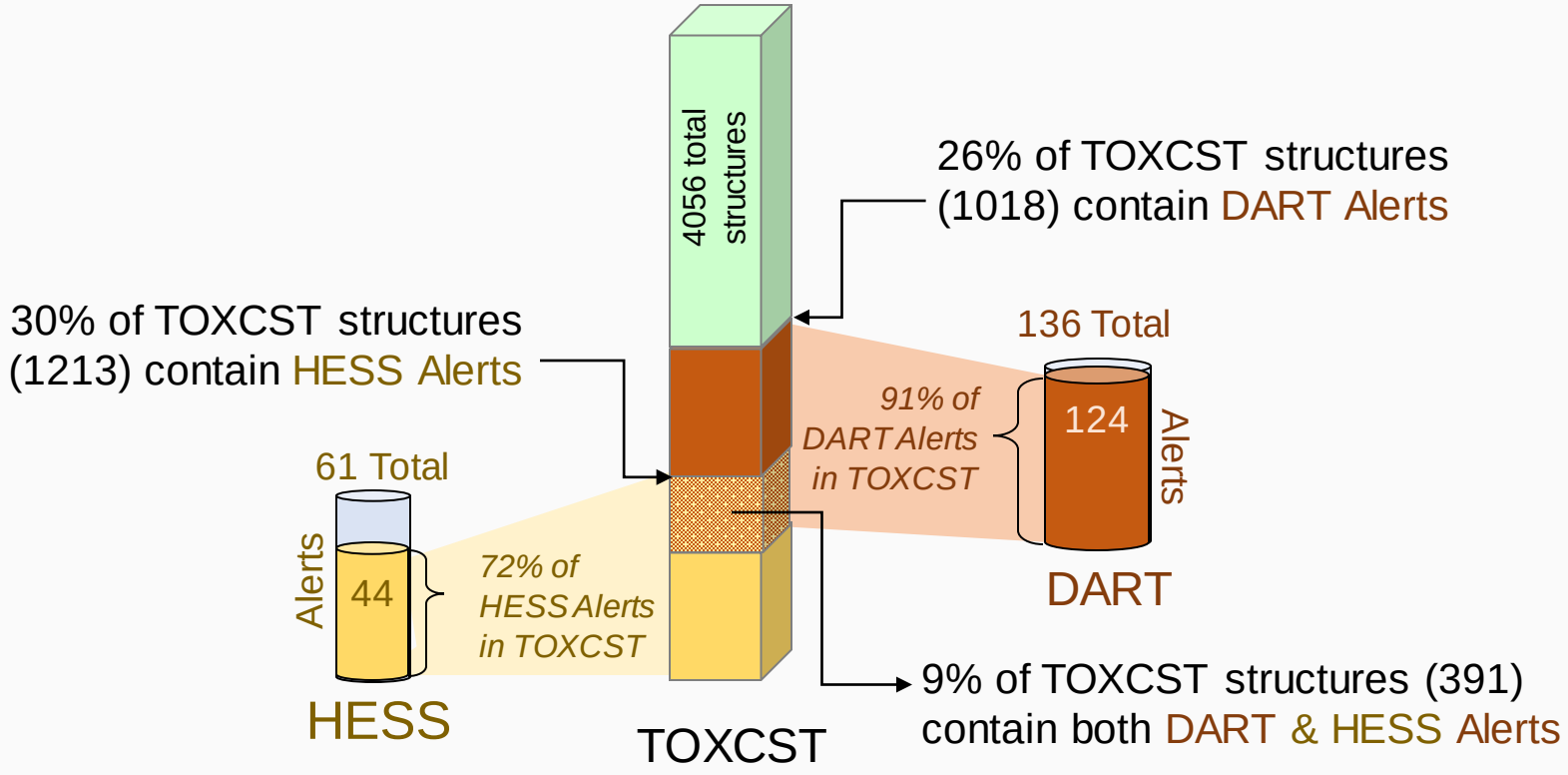
Incidence of DART toxicity alerts in ToxCast



Alert “classes” define local regions of chemical space for targeted enrichment studies

ChemResToxicol., 2016, 29, 1225−1251

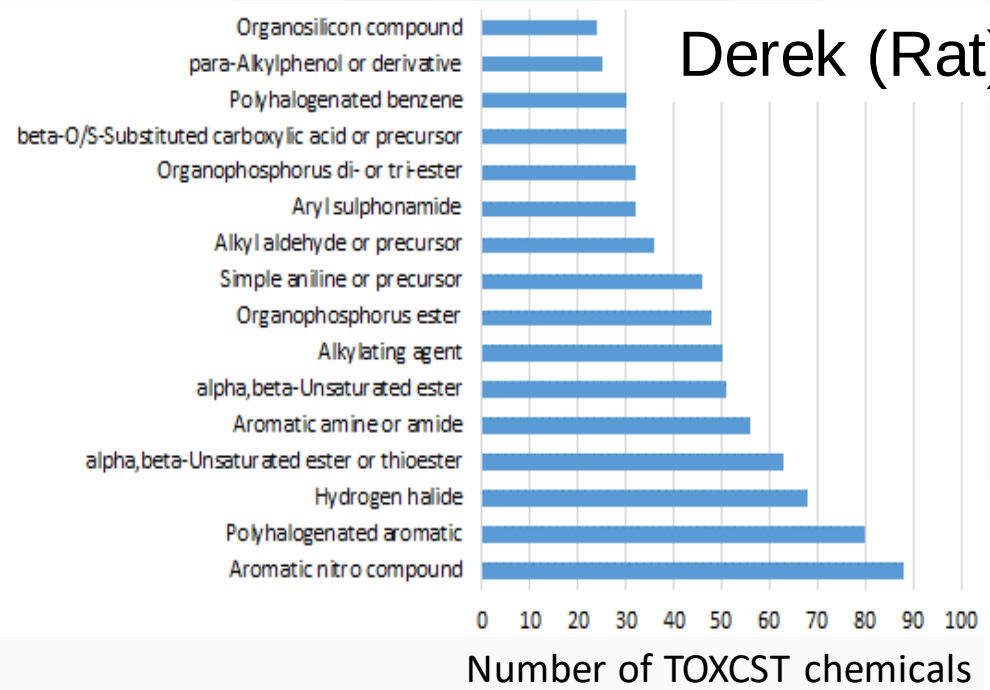
How well does ToxCast cover historical SAR toxicity "alerts"?

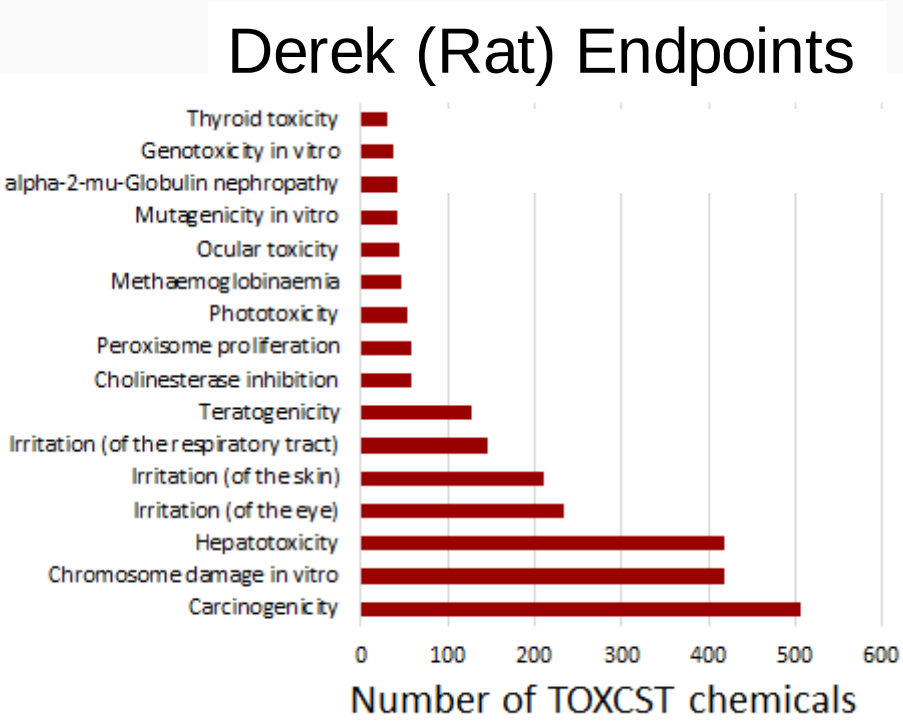


- 72% of HESS & 91% of DART alerts detected in TOXCST chemicals
- 47% of TOXCST chemicals contain either HESS or DART alert

ChemResToxicol., 2016, 29, 1225–1251

Incidence of Derek (Rat) alerts and endpoint predictions in ToxCast

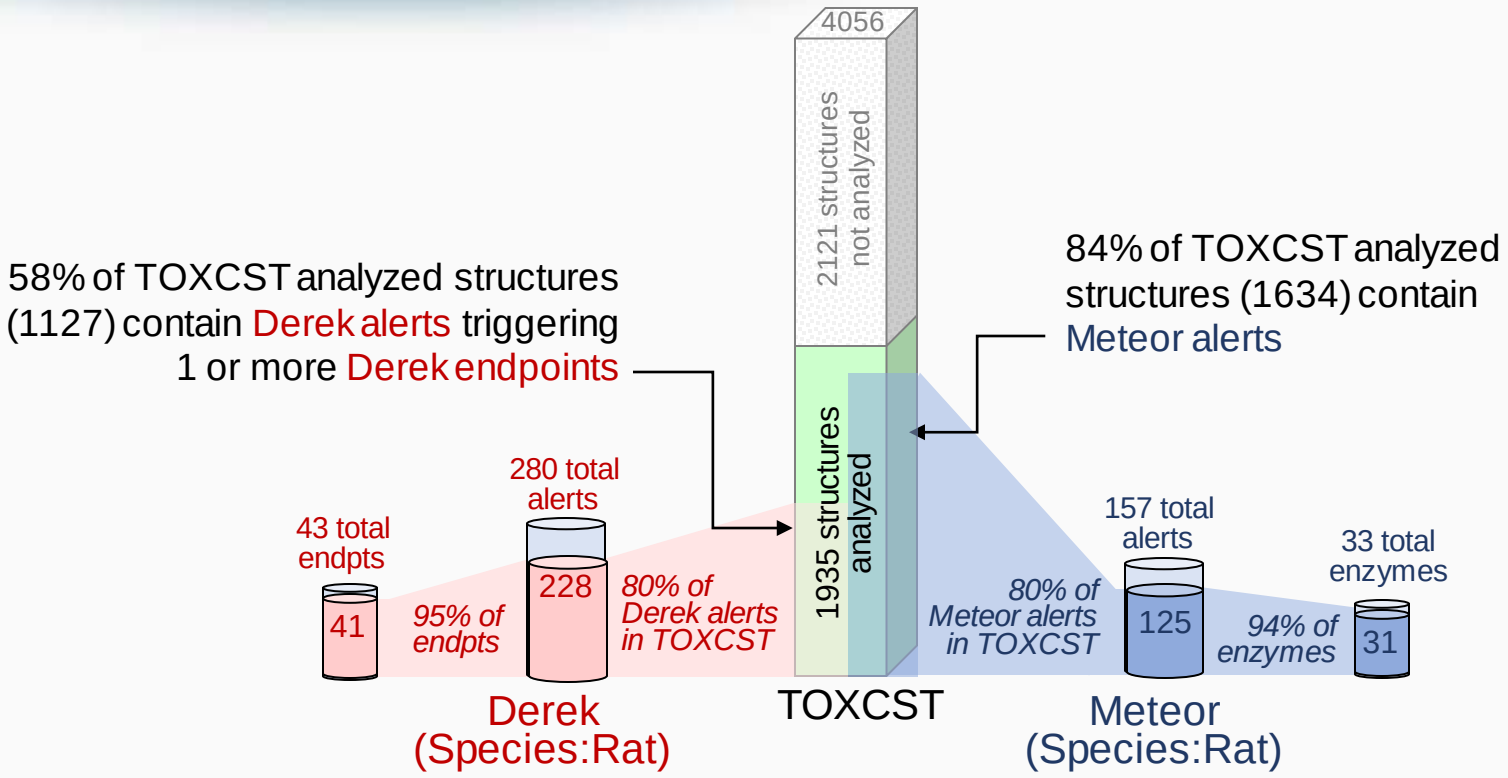




Alert & endpoint "classes" define local regions of chemical space for targeted enrichment studies

Chem Res Toxicol., 2016, 29, 1225−1251

How well does TOXCST cover historical SAR toxicity & biotransformation"alerts"?



- 80% of Derek & Meteor alerts detected in TOXCST chemicals
- 95% of Derek endpts & 94% of Meteor enzymes triggered
- 58% of TOXCST chemicals contain Derek toxicity alert
- 84% of TOXCST chemicals contain Meteor biotransformation alert

ChemResToxicol., 2016, 29, 1225–1251

Is library “fit for purpose”?

- Does library provide ~~sufficient~~ broad coverage of chemicals of interest to EPA & stakeholders? YES!
- Does library include ~~sufficient~~ broad chemical diversity to span full range of toxicity mechanisms and outcomes of concern? YES!
- Does library provide ~~sufficient~~ broad coverage of local regions of chemistry to enable local model development? YES!

… relative to the “chemical universe” and target inventories of greatest interest and concern to EPA

Current & future work

ToxCast:

- Develop automated workflows to support chemotype (e.g., ToxPrint) analyses in local chemistry domains and apply to ToxCast assay data sets (individually and globally)
- Strategic expansion of chemical library into local chemical domains

Tox21:

- Landscape paper – history, content of library
- Analysis of Tox21 analytical chemistry data

ExpoCast:

- Chemical library support for Non-targeted Screening (NTS) International Mixture Challenge (10 mixtures, 100-400 chems)
- Chemical library support for generating publicly releasable high-resolution mass spectra by 7 companies & collaborators

Coauthors & Acknowlegements

- Richard Judson
- Keith Houck
- Chris Grulke
- Patra Volarath
- Indira Thillainadarajah
- Chihae Yang & Jim Rathman
- Matt Martin
- John Wambaugh
- Tom Knudsen
- Jayaram Kancherla
- Kamel Mansouri
- Grace Patlewicz
- Tony Williams
- Kevin Crofton
- Rusty Thomas

Thanks also to past and present ToxCast team members, Evotec contractors, ACToR curators, and special thanks to Bob Kavlock, David Dix, and Marty Wolf.