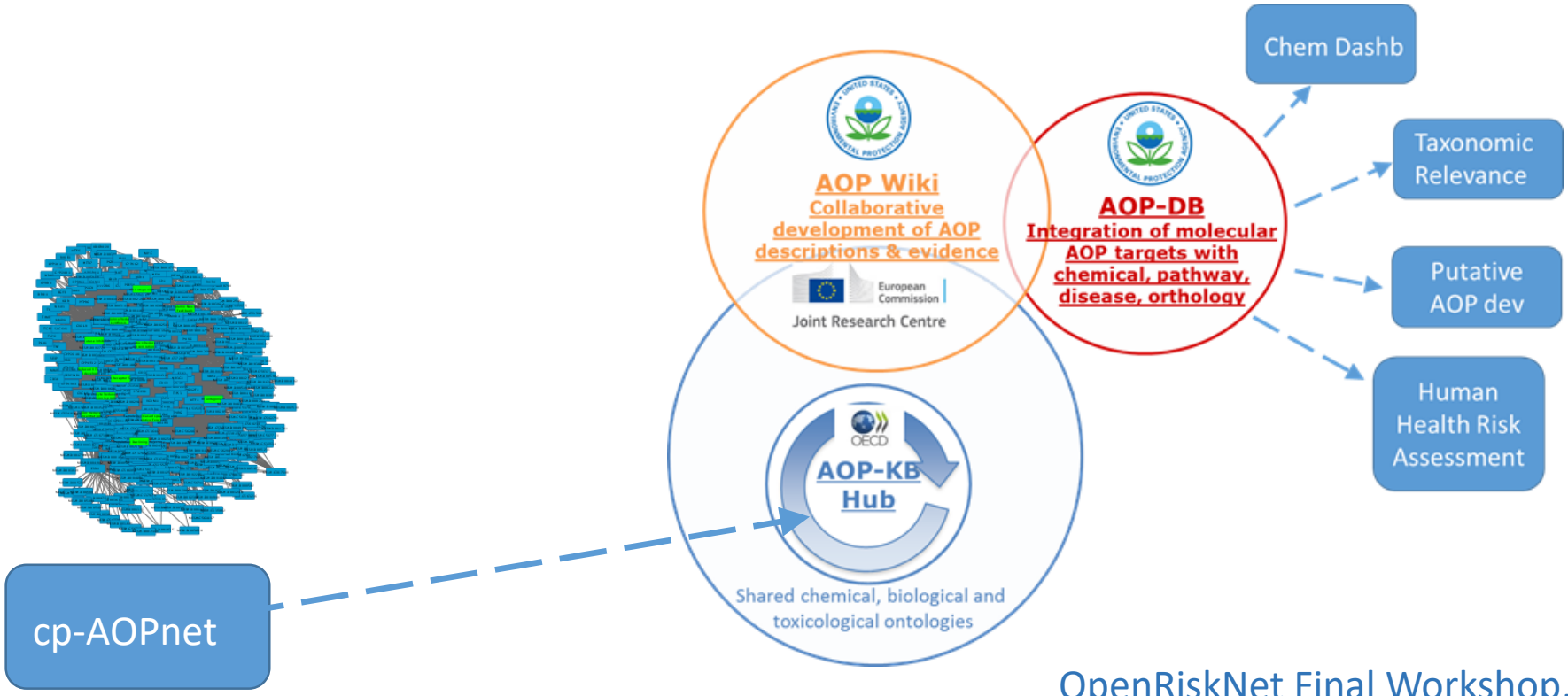
AOP-DB: The Adverse Outcome Pathway Database

Holly M. Mortensen, Ph.D.
US EPA, Center for Public Health and Environmental Assessment



OpenRiskNet Final Workshop, Amsterdam, October 23-24, 2019

Office of Research and Development

EPA Disclaimer: The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the Agency.

AOP-DB –What is it? Where can I get some?

Currently, the **EPA AOP-DB** is an internal SQL database
- *OpenRiskNet Integration-Implementation Challenge Winner (!)*

- AOPDB is an *AOP Profiling Tool that allows for the biological context of ea AOP to be explored.*
- Originally created to **extract molecular information from the AOPWiki** into useable format
- The AOP-DB aggregates relationships between ***AOP-gene targets, chemical, disease, pathway, and species orthology information.***
- ***Long term significance and impact--*** Continued translation of AOP biological context, in real time from the AOP-Wiki, and the ability to associate these data between and across AOPs, and with assay, chemical, pathway and disease endpoints.

This can help Risk Assessors understand what chemicals target the AOP, mechanistic steps involved in the outcome, and who is affected!!!

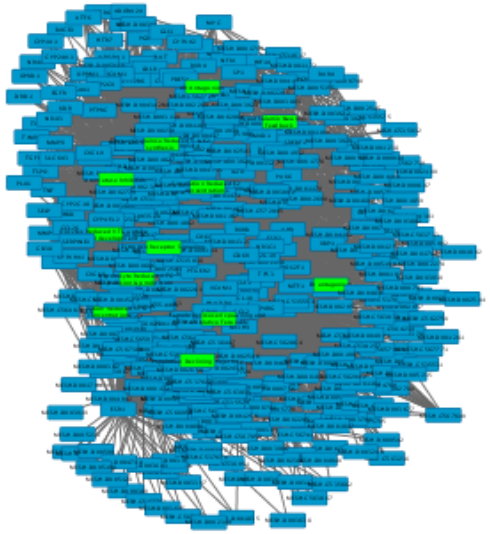
Publicly Available Now:
- AOPWiki →ToxCast Mappings via EPA CompTox Chemistry Dashboard
- OpenRiskNet Aop-DB RDF mapping

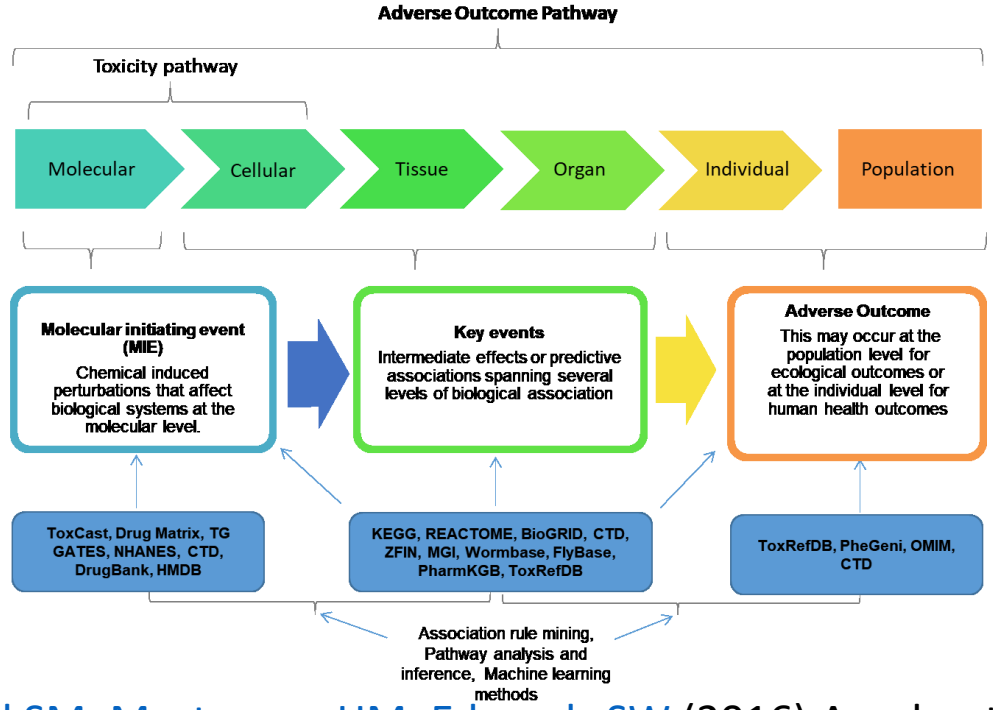
AOP-DB –Why was it created?

Originally created with a dual purpose:

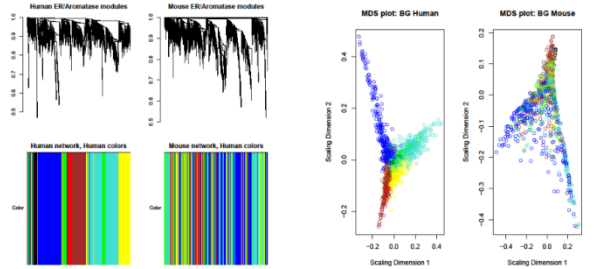
1. to characterize candidate AOPs for case study
   a) Species applicability (*Pittman et al. 2018*)
   b) Inter-individual Variation (*Mortensen, Chamberlin et al. 2018*)
2. To support the cp-AOPnet and putative AOP development (*Edwards*)

****General hypothesis generation tool to explore AOPs***

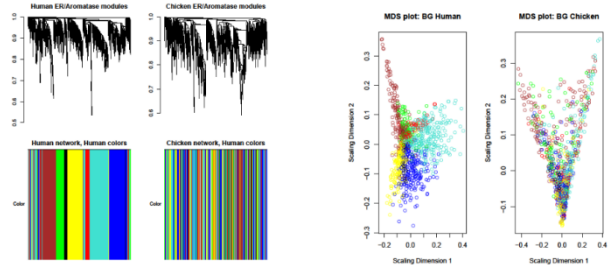




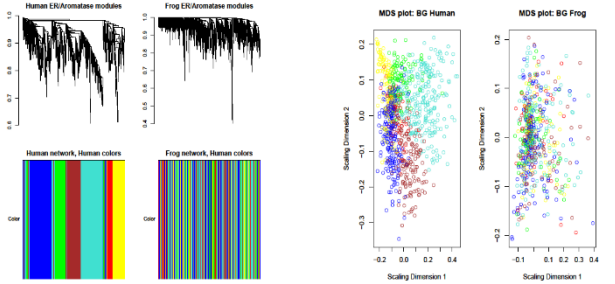
Human + Mouse



Human + Chicken



Human + Frog

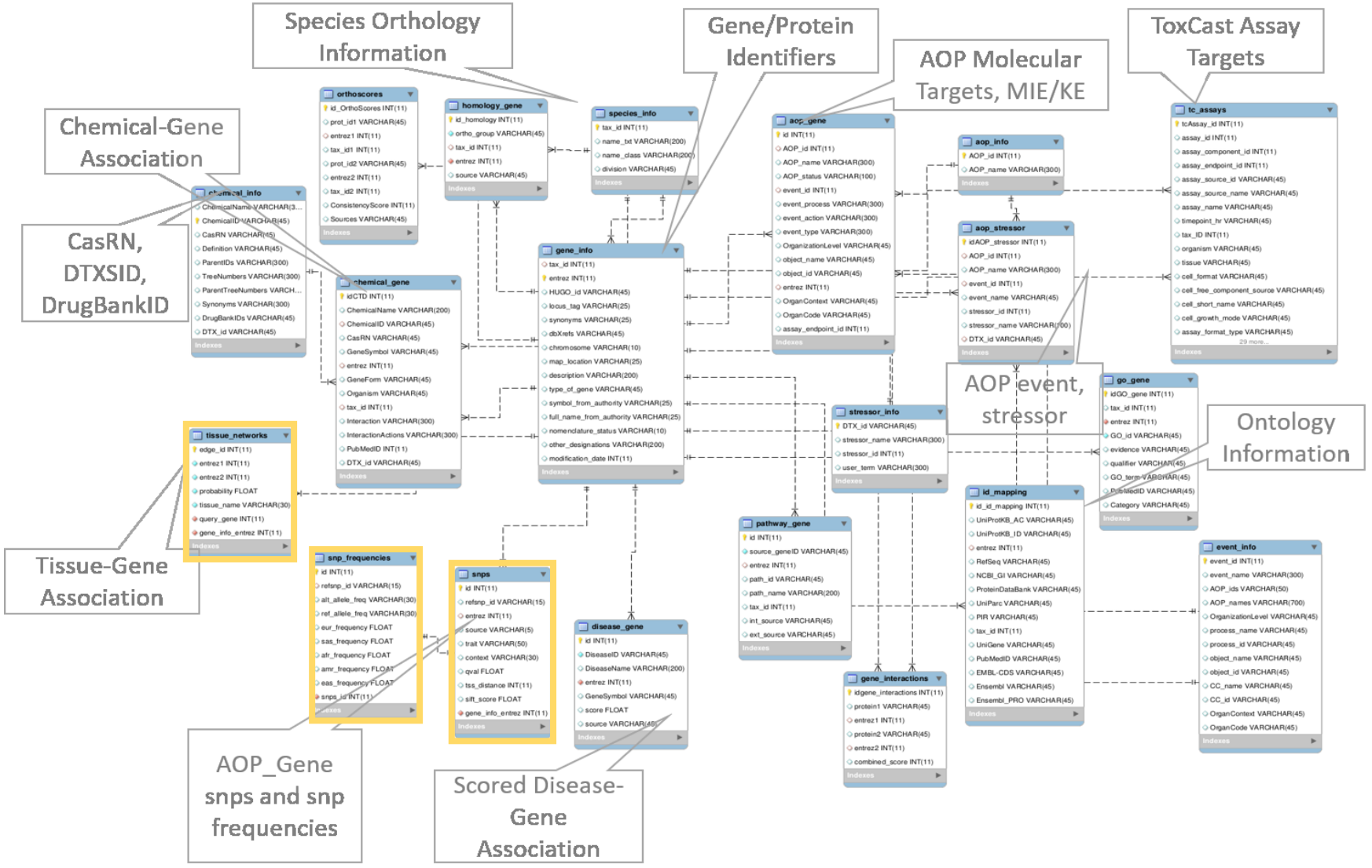


ER/Aromatase expression module preservation

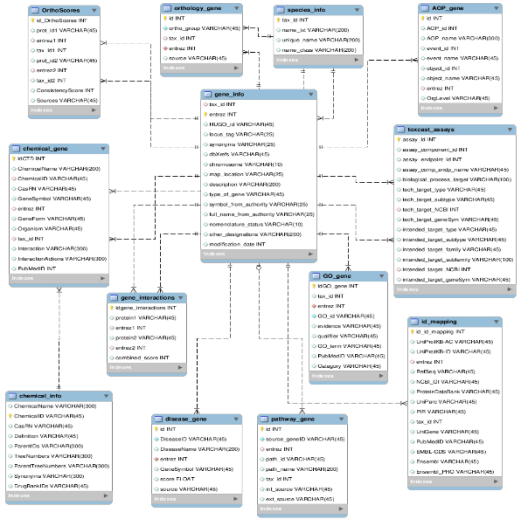
Oki NO, Nelms MD, Bell SM, Mortensen HM, Edwards SW (2016) Accelerating Adverse Outcome Pathway Development Using Publically Available Data Sources. Curr Environ Health Rep. Mar 3(1): 53-63.

United States Environmental Protection Agency

AOPdb v.2 Schema



United States Environmental Protection Agency



AOP-DB: Ontology-curated AOP Gene Mapping to KE

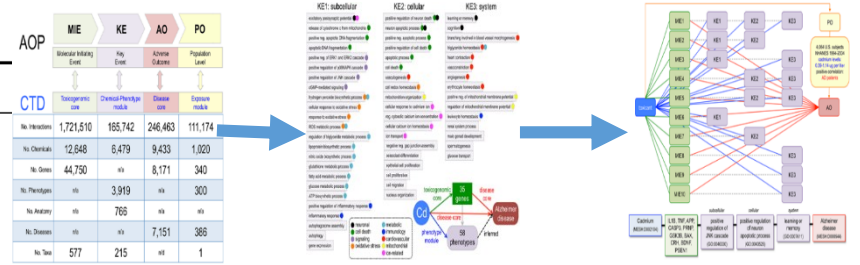
id	AOP_id	AOP_name	event_process	OrganizationLevel	object_id	entrez	OrganContext	assay_endpoint_id
1	3	Inhibition of the mitochondrial comp…	. NADH-ubiquinone oxidoreductase …	. Molecular	. PR:000031316	4535		0
2	3	Inhibition of the mitochondrial comp…	. NADH-ubiquinone oxidoreductase …	. Molecular	. PR:000031316	17716		0
3	3	Inhibition of the mitochondrial comp…	. NADH-ubiquinone oxidoreductase …	. Molecular	. PR:000031316	3338902		0
4	3	Inhibition of the mitochondrial comp…	. NADH-ubiquinone oxidoreductase …	. Molecular	. PR:000031316	807636		0
5	3	Inhibition of the mitochondrial comp…	. NADH-ubiquinone oxidoreductase …	. Molecular	. PR:000031316	2193907		0
6	3	Inhibition of the mitochondrial comp…	. NADH-ubiquinone oxidoreductase …	. Molecular	. PR:000031316	140531		0
7	6	Antagonist binding to PPARalpha l…	. PPAR alpha	. Molecular	. PR:000013056	19013		0
8	6	Antagonist binding to PPARalpha l…	. PPAR alpha	. Molecular	. PR:000013056	25747		0
9	6	Antagonist binding to PPARalpha l…	. PPAR alpha	. Molecular	. PR:000013056	5465		132
10	6	Antagonist binding to PPARalpha l…	. PPAR alpha	. Molecular	. PR:000013056	5465		718
11	6	Antagonist binding to PPARalpha l…	. PPAR alpha	. Molecular	. PR:000013056	403654		0
12	7	Aromatase (Cyp19a1) reduction le…	. Aromatase (Cyp19a1)	. Cellular	. PR:000006100	30390		0
13	7	Aromatase (Cyp19a1) reduction le…	. Aromatase (Cyp19a1)	. Cellular	. PR:000006100	1588		319
14	7	Aromatase (Cyp19a1) reduction le…	. Aromatase (Cyp19a1)	. Cellular	. PR:000006100	1588		320
15	7	Aromatase (Cyp19a1) reduction le…	. Aromatase (Cyp19a1)	. Cellular	. PR:000006100	1588		767
16	7	Aromatase (Cyp19a1) reduction le…	. Aromatase (Cyp19a1)	. Cellular	. PR:000006100	13075		0

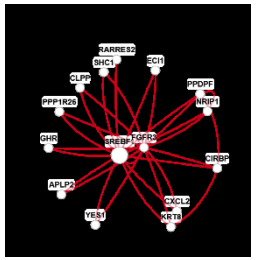
Ives, C., Campia, I., Wang, R.L., Wittwehr, C., Edwards, S. *(2017)* Creating a Structured Adverse Outcome Pathway Knowledgebase via Ontology-Based Annotations. Applied In Vitro Toxicology, Vol. 3, No. 4.

AOP-DB v.2 Data Sources

AOPdb v.2 Data Source Overview

Biological Category	Data Type	Count	Sources
Gene	Unique Genes	23189522	NCBI Gene
	Gene Interactions	580459027	STRING
Taxonomy & Orthology	Entrez Supported Organisms	24336	NCBI Taxonomy | Homologene
	Orthologous Groups	64502	KEGG Orthology | metaPhOrs
	Taxa supported by ortho groups	568	
AOP	Supported AOPs	236	AOP-Wiki
	AOP-gene targets	1312	
Chemical	CTD Chemicals	170646	CTD
	CTD Chemical-gene associations	1749648	
	ToxCast Assay	406	ToxCast
Pathway	Total Pathways	100594	KEGG Pathways | Reactome
	Inferred AOP-pathway associations	7606	ConsensusPathDB
Disease	Unique Disease IDs	24166	DisGeNET
	Disease-gene associations	628685	
Ontology	Unique GO Term IDs	26626	NCBI Gene
	Ontology-gene associations	204302	
Tissue	Total Tissues	145	HumanBase
	Tissue-gene interactions	17378250	
SNPs	Total SNPs	25275	Sift | GTEx | GWAS Catalog
	Total SNP Frequencies	4157	1000Genomes Project





entrez	6720
symbol	SREBF1
Tax ID	9609
Chromosome	17
Description	sterol regulatory element binding transcription factor 1
Associated AOPS	34,58,61,62

GWAS Catalog
The NHGRI-EBI Catalog of published genome-wide association studies

Sift

GTEx Portal

1000 Genomes Project

- Population-level frequency data for functionally relevant SNPs
  - EUR - European
  - SAS - South Asian
  - AFR - African
  - AMR - Ad Mixed American
  - EAS - East Asian

Davis et al. (2018) The Comparative Toxicogenomics Database: update 2019. Nucleic Acids Research Vo 47, Issue D1, pg. D948-D954.

Greene et al (2015) Understanding multicellular function and disease with human tissue-specific networks Nat Genet. Jun;47(6):569-76.

AOP-DB Update v.2

Updated GUI Frontend with search and download capabilities

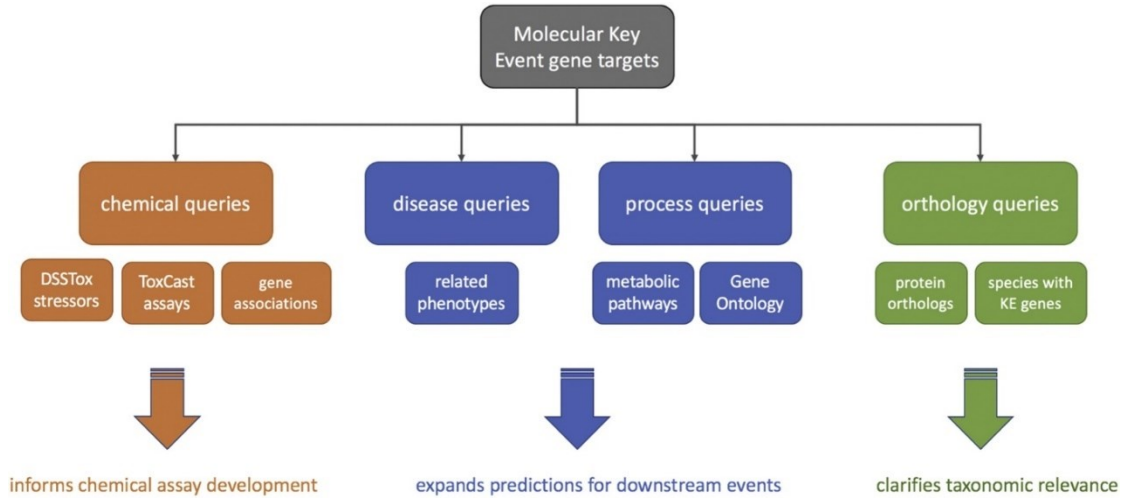
Code/Data updates include:

- Tools for directly modifying the DB; Automatic data pulls
- Data export in standard formats (eg. JSON/XML/CSV)
- Manual is in prep, including DB tutorial and changelog
- HumanBase Tissue Network integration
- Population frequencies for AOP-relevant SNP variants
- CP-AOP builder

Mortensen, H.M., Levey, T., Langley, P., Williams, A. *(in prep)* EPA AOP-DB version 2 updates. *Nature Scientific Data.*

AOP-DB Use Case Examples

Figure 1: Flowchart depicting the nature of potential lines of inquiry and their application.



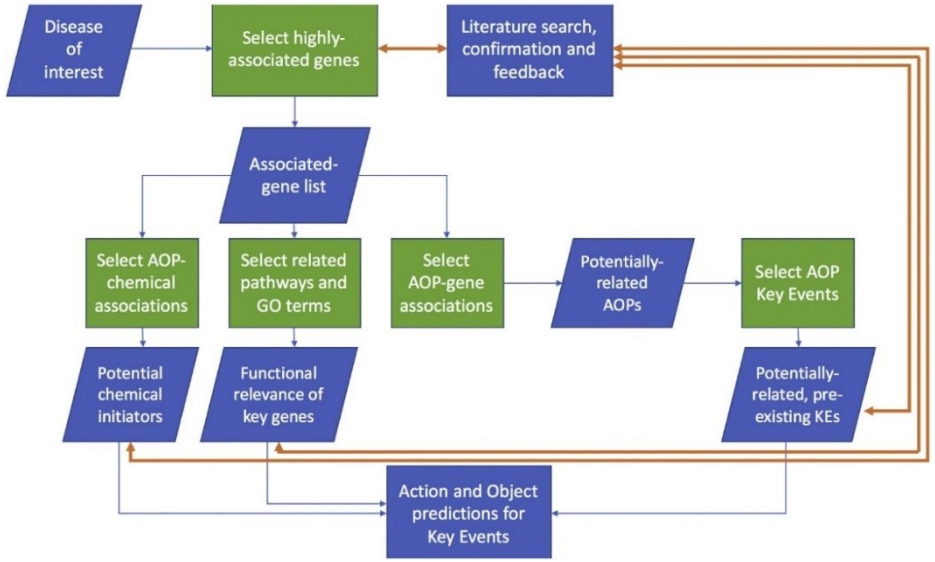


Figure 2: Flowchart depicting the series of steps taken to develop a predicted AOP from the AOP-DB.

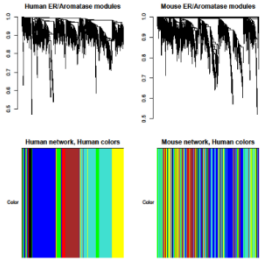
Pittman, M.E., Edwards, S.W., Ives, C., Mortensen, H.M. *(2018)* AOP-DB: A database resource for the exploration of Adverse Outcome Pathways through integrated association networks. *Toxicology and Applied Pharmacology* Mar 15; 343:71-83

AOP-DB Use Case Examples

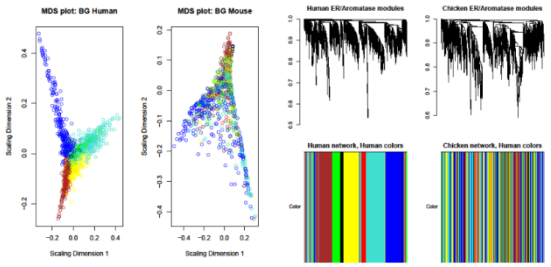
AOP-directed Co-Expression Module Comparison

- MIE and KE gene/protein
- Define "Functional Neighborhood"
- Inter-species Pathway Comparison
- Molecular Sequence Similariity
- Co-Expression Network Construction
- Co-Expression Pattern Analysis

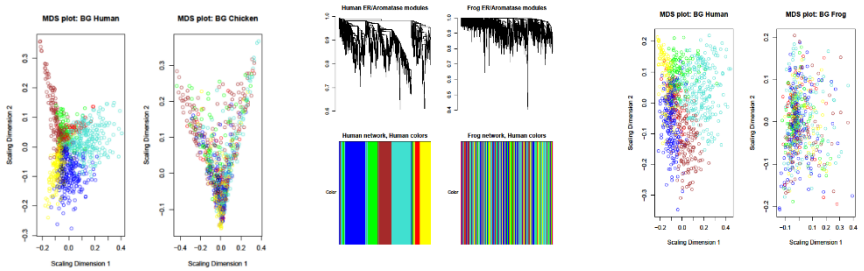
Human-Mouse

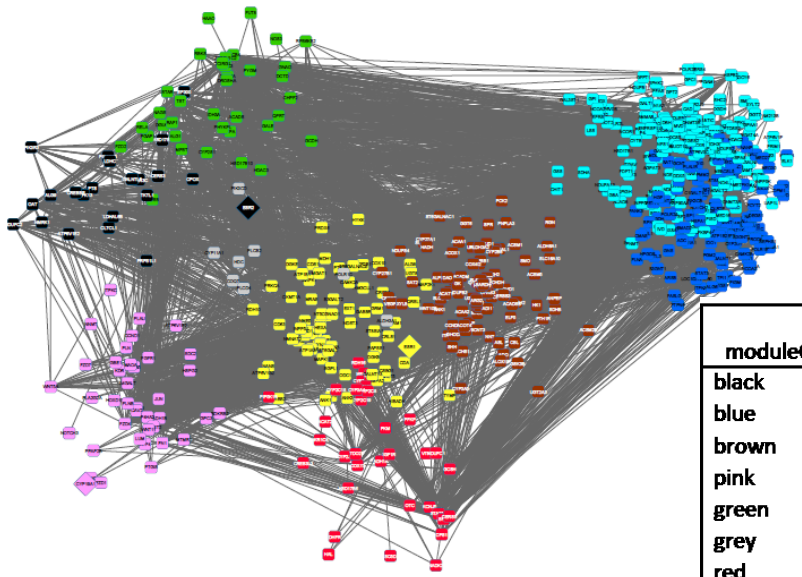


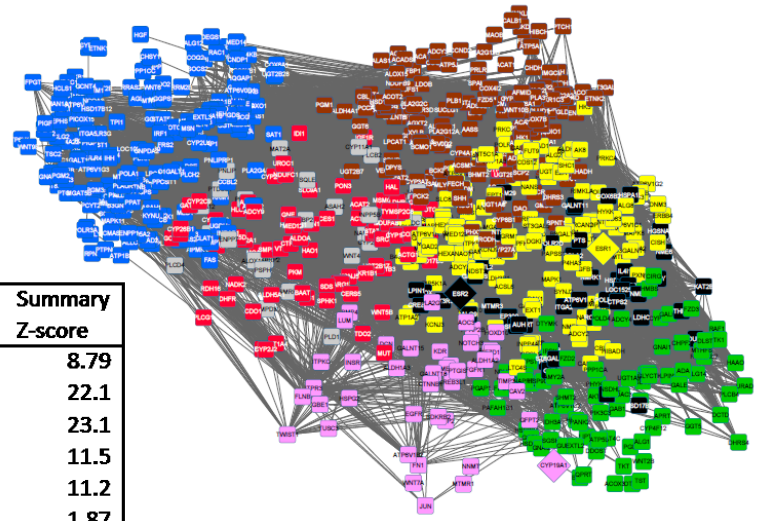
Human-Chicken



Human-Frog







moduleColor	moduleSize	Summary Z-score
black	56	8.79
blue	314	22.1
brown	252	23.1
pink	40	11.5
green	159	11.2
grey	17	1.87
red	62	5.59
turquoise	377	12.6
yellow	182	11.5

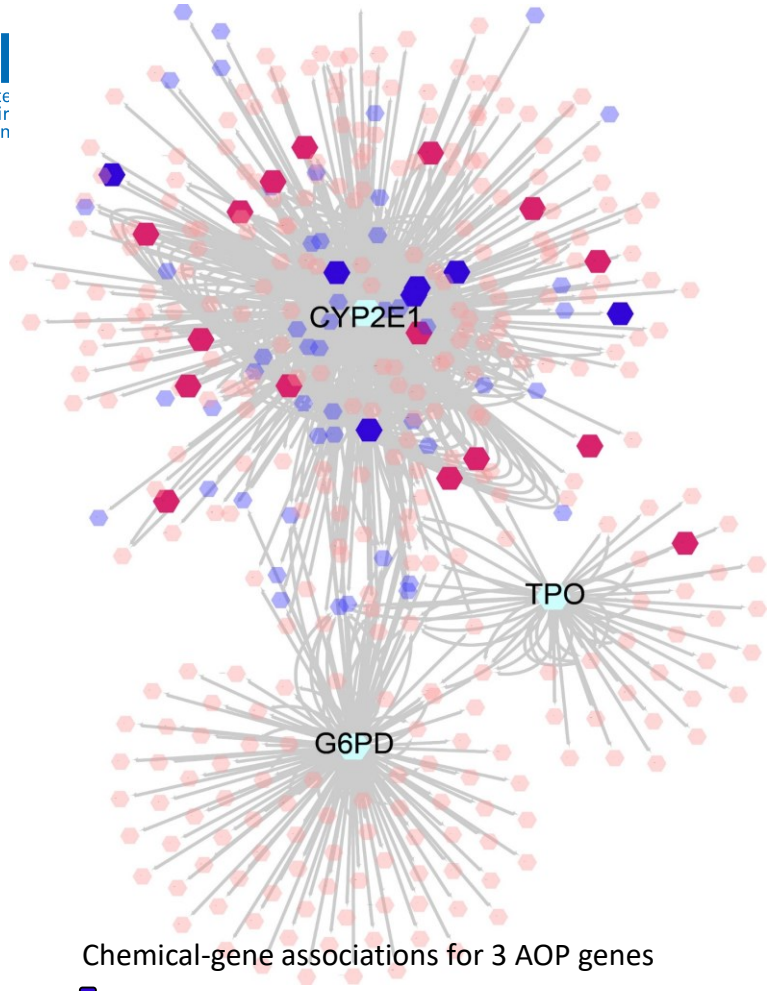
5<Z score<10 = Moderate Module Preservation
Z score>10 = High Module Preservation

Mortensen, H.M., Pittman, M., Lalone, C., Villeneuve, D., Ankley, G (*in prep*) A Computational Framework for Defining the Taxonomic Applicability of the Adverse Outcome Pathways Through integration of Molecular Datasets. *Computational Toxicology.*

WGCNA: Langfelder and Horvath (2008) *BMC Bioinformatics*

AOP-DB Use Case Example

AOP-anchored Genetic Susceptibility



Chemical-gene associations for 3 AOP genes

- Increase expression/activity of CYP2E1
- Other interactions
- Non-TSCA workplan chemicals

Step 1: Gene Set Selection

↓ **# Human Genes**

Step 2: Pathway Module Identification/Gene Set Validation

↓ **# Human Validated Genes**

Step 3: Identify regulatory regions for each gene (ENCODE)

↓ **Outcome related regulatory regions**

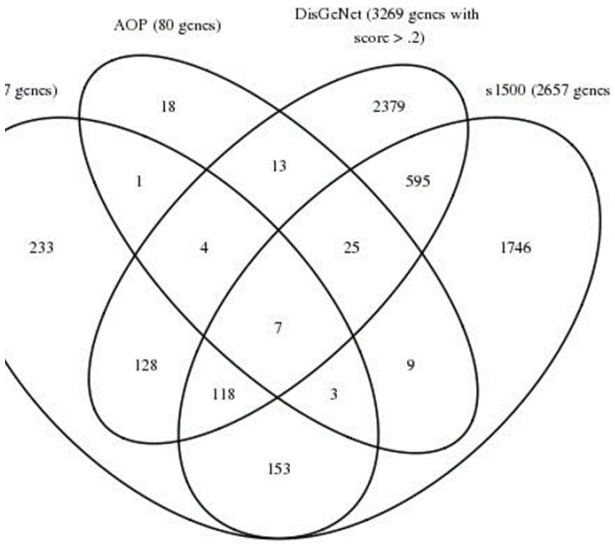
Step 4: Identify functionally significant SNPS (Ensembl/Gtex)

↓ **# Human functional, outcome related SNPs in each identified region**

Step 5: Population characterization (1000 Genomes)

↓ **Individual/Population level frequency data for functional, outcome related SNPs**

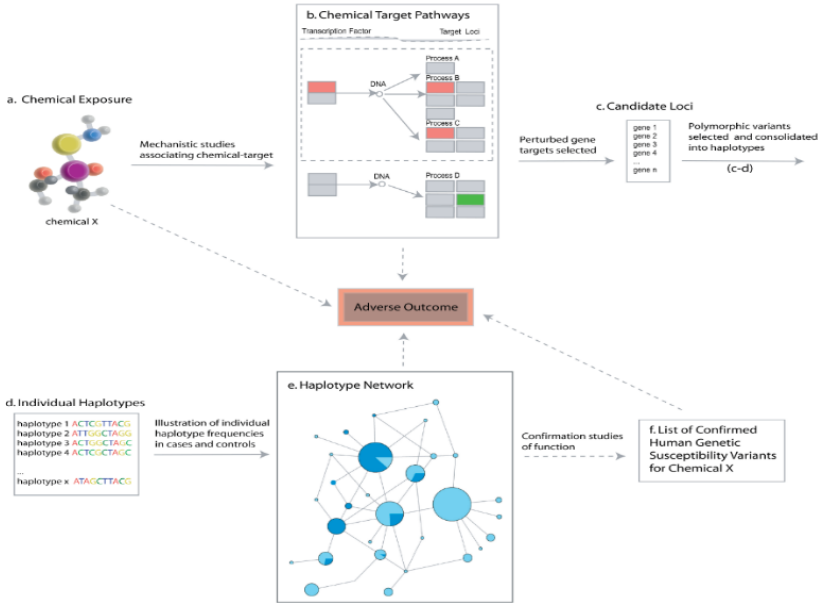
Step 6: Outcome-specific Multi-genic characterization



FY17 Key HHRA Product: Mortensen, H.M., Chamberlin, J., Joubert, B., Angrish, M, Sipes, N., Lee, J.S., Euling, S.Y. *(2018)* Leveraging human genetic and adverse outcome pathway (AOP) data to inform susceptibility in human health risk assessment. Mammalian Genome. Feb;29(1-2):190-204.

AOP-DB Use Case Examples

AOP-anchored Genetic Susceptibility/Community Risk



AOP Genes (MIE/KE from AOP-KB)

Identify and Filter impactful protein coding SNPs (PolyPhen/SIFT)

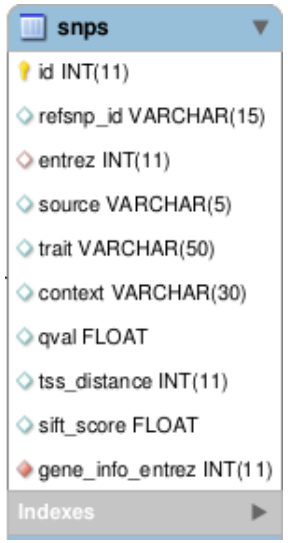
Identify putative, tissue specific, noncoding regulatory SNPs (eQTLs from GTEX)

Human functional, outcome -related SNPs in each AOP (1KGP Stage 3)

Evaluate level of individual and population level variation at protein altering and tissue relevant variants

Mortensen, HM and Euling, SY (2013) Integrating mechanistic and polymorphism data to characterize human genetic susceptibility for environmental chemical risk assessment in the 21st century. TAAP 15; 271(3):395-404.

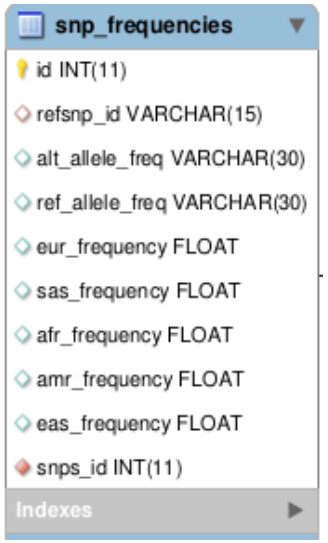
AOP-DB Human SNP Data



- Functionally relevant, phenotype SNPs for identified AOPs
- Computationally and experimentally derived

GWAS Catalog
The NHGRI-EBI Catalog of published genome-wide association studies

GTEx Portal



Population-level frequency data for functionally relevant SNPs

- EUR - European
- SAS - South Asian
- AFR - African
- AMR - Ad Mixed American
- EAS - East Asian

1000 Genomes Project

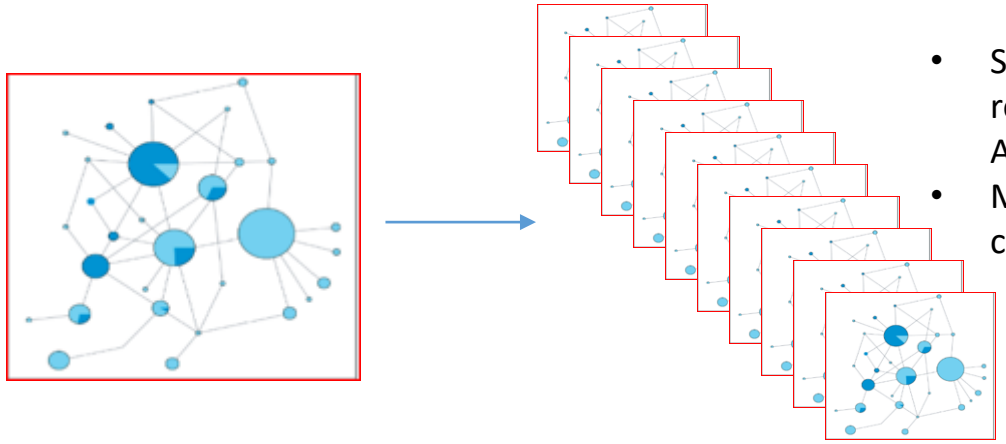
AOP-DB Use Examples

AOP-anchored Genetic Susceptibility/Community Risk

Filter | Cols: « < 1 - 50 > »

	rs3020340str1	rs11155813str1	rs1890010str1	rs34984292str1	rs2234693str1	rs12195741str1	rs6902771str1	rs9383951str1	rs547908752str1
1	A	T	T	AAAAAAAAAA	C	C	T	G	C
2	G	T	T	AAAAAAAAAAA	T	C	C	G	C
3	A	T	T	AAAAAAAAAA	T	T	C	G	C
4	G	C	T	AAAAAAAAAA	C	C	T	G	C
5	G	C	T	AAAAAAAAAA	C	C	T	G	C
6	A	T	C	AAAAAAAAAAA	C	C	T	G	C
7	A	T	T	AAAAAAAAAA	C	C	T	G	C
8	A	T	T	AAAAAAAAAA	C	C	T	G	C
9	G	T	T	AAAAAAAAAAA	T	C	C	G	C
10	A	T	C	AAAAAAAAAA	C	C	T	G	C
11	A	T	C	AAAAAAAAAA	T	T	C	G	C
12	A	T	C	AAAAAAAAAA	T	C	C	G	C
13	A	T	C	AAAAAAAAAA	T	C	C	G	C
14	A	C	C	AAAAAAAAAA	C	C	T	G	C
15	A	T	C	AAAAAAAAAA	T	T	C	G	C
16	A	T	T	AAAAAAAAAA	C	T	T	G	C
17	A	T	T	AAAAAAAAAA	C	T	T	G	C
18	A	T	C	AAAAAAAAAA	C	C	T	G	C

Showing 1 to 21 of 2,504 entries, 59 total columns



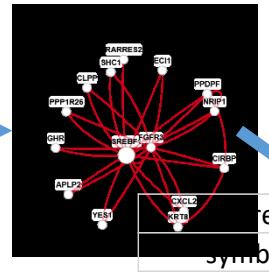
- Susceptibility thresholding of functionally relevant, phenotype SNPs for identified AOP
- Multi-Allelic AOP-relevant Haplotype construction

Tissue Networks in the AOP-DB

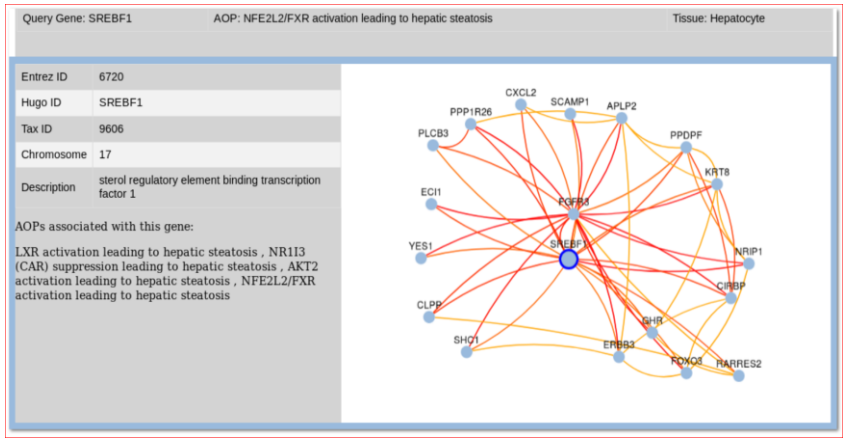
humanbase

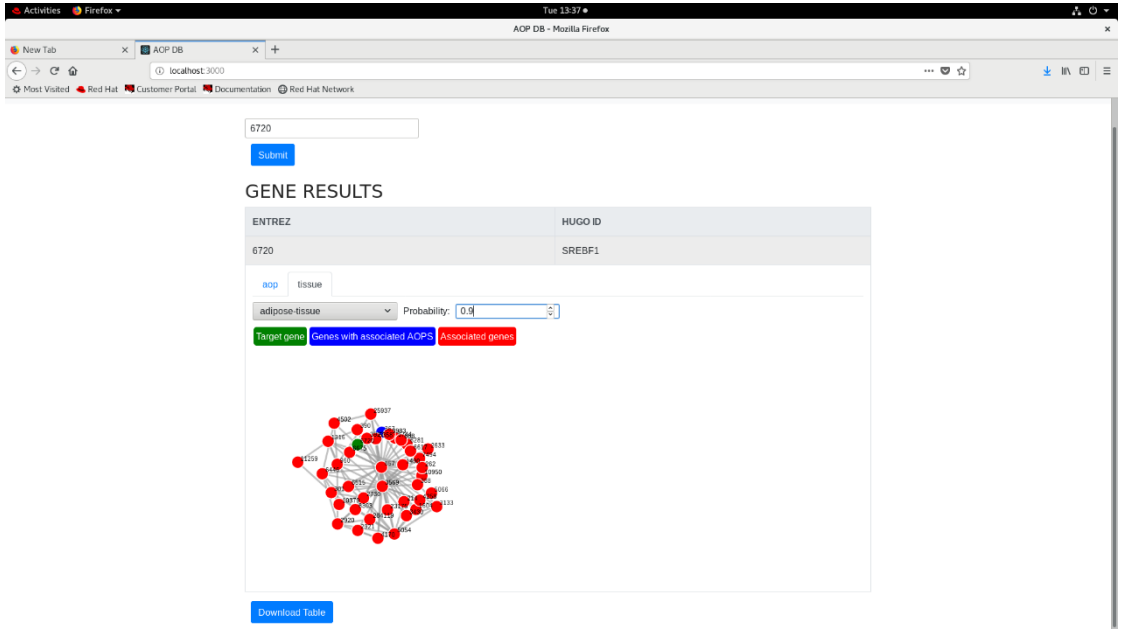
- Gene interaction networks for 144 human tissues
- AOP-specific (MIE, KE) molecular targets with tissue level probabilities

edge_id	entrez1	entrez2	probability	tissue_name	query_gene
6919426	7054	10614	0.0251071	hepatocyte	6720
6919427	7054	1043	0.20275	hepatocyte	6720
6919428	7054	275	0.482486	hepatocyte	6720
6919429	7054	5420	0.00741191	hepatocyte	6720
6919430	7054	5359	0.0321038	hepatocyte	6720
6919431	7054	1387	0.0597612	hepatocyte	6720
6919432	7054	9858	0.135615	hepatocyte	6720
6919433	7054	10610	0.372655	hepatocyte	6720
6919434	7054	345	0.0439449	hepatocyte	6720
6919435	7054	5567	0.301057	hepatocyte	6720
6919436	7054	686	0.0266105	hepatocyte	6720
6919437	7054	8532	0.00067525	hepatocyte	6720
6919438	7054	334	0.0303959	hepatocyte	6720
6919439	7054	1675	0.744139	hepatocyte	6720
6919440	7054	2261	0.943748	hepatocyte	6720



ez	6720
ymbol	SREBF1
Tax ID	9609
Chromosome	17
Description	sterol regulatory element binding transcription factor 1
Associated AOPS	34,58,61,62





Greene et al (2015) Understanding multicellular function and disease with human tissue-specific networks Nat Genet. Jun;47(6):569-76.

cp-AOP Derivation

Computationally predictive AOPs (cpAOPs)

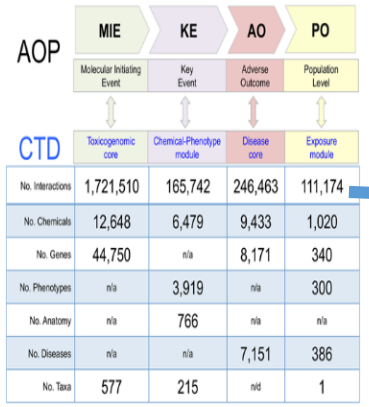
- “Chemical-Induced Phenotypes at CTD Help Inform the Pre-disease State and Construct Adverse Outcome Pathways”

- Chemical-phenotype
- Chemical-disease
- Exposure-studies

Front End Tool

- Will guide the user through the cp-AOP building process adapted from the manual workflow in Davis et al.
- Selecting a chemical and a disease, a user can build an AOP
- Using built AOPs, we can gain insight on what AOPs people are searching for or studying
- This tool has the potential to increase AOP data richness in the AOP-DB

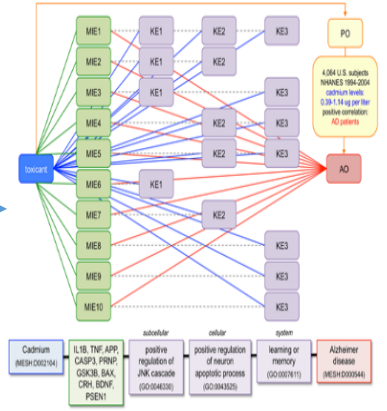
Select chemical and disease of interest



No. Interactions	1,721,510	165,742	246,463	111,174
No. Chemicals	12,648	6,479	9,433	1,020
No. Genes	44,750	n/a	8,171	340
No. Phenotypes	n/a	3,919	n/a	300
No. Anatomy	n/a	766	n/a	n/a
No. Diseases	n/a	n/a	7,151	386
No. Taxa	577	215	n/d	1

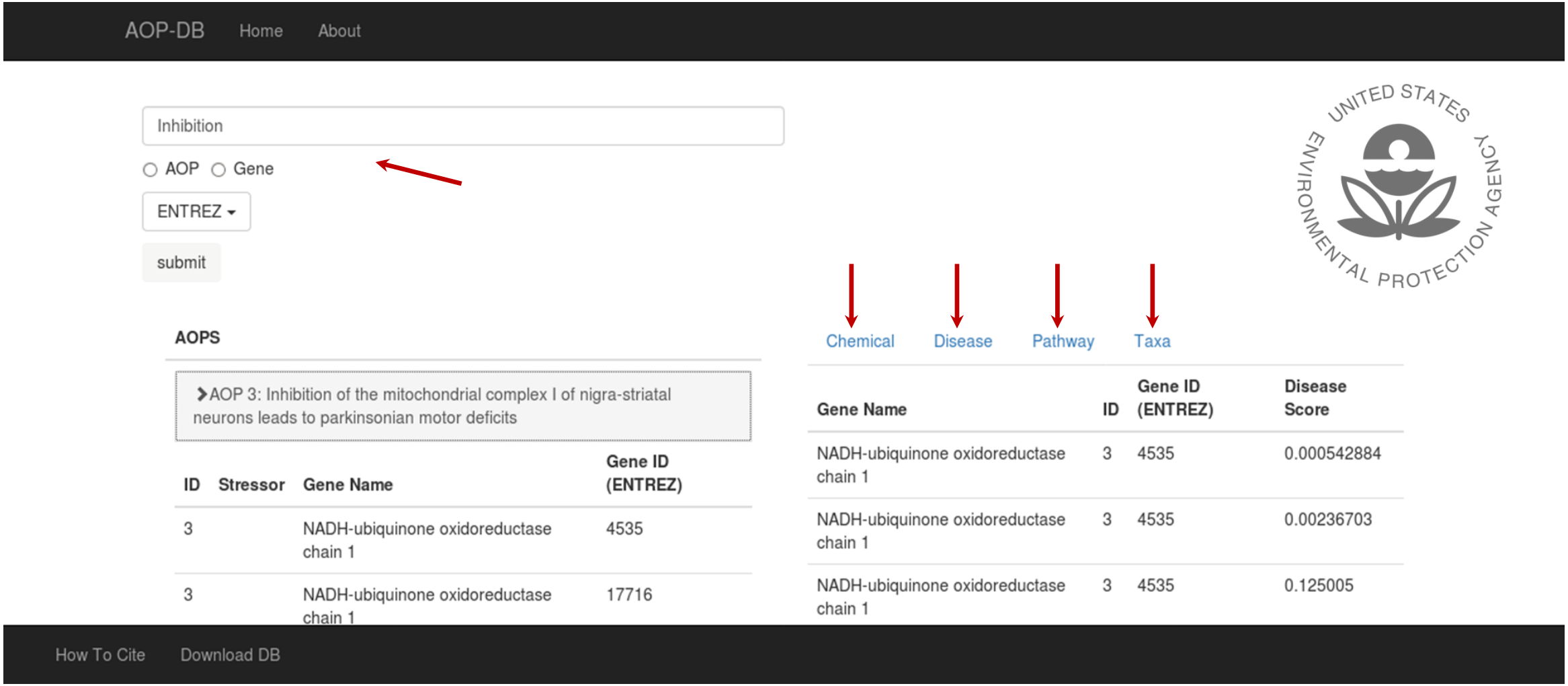
Assign organization levels to computed key event phenotypes to build up AOP

Connect pathway from chemical to MIE to KE’s to Adverse Outcome. Link to existing exposure studies.



Davsi et al. (2018) The Comparative Toxicogenomics Database: update 2019. Nucleic Acids Research Vo 47, Issue D1, pg. D948-D954.

AOP-DB Front-End Tool



Activities Firefox Wed 15:06

AOP DB - Mozilla Firefox

AOP DB

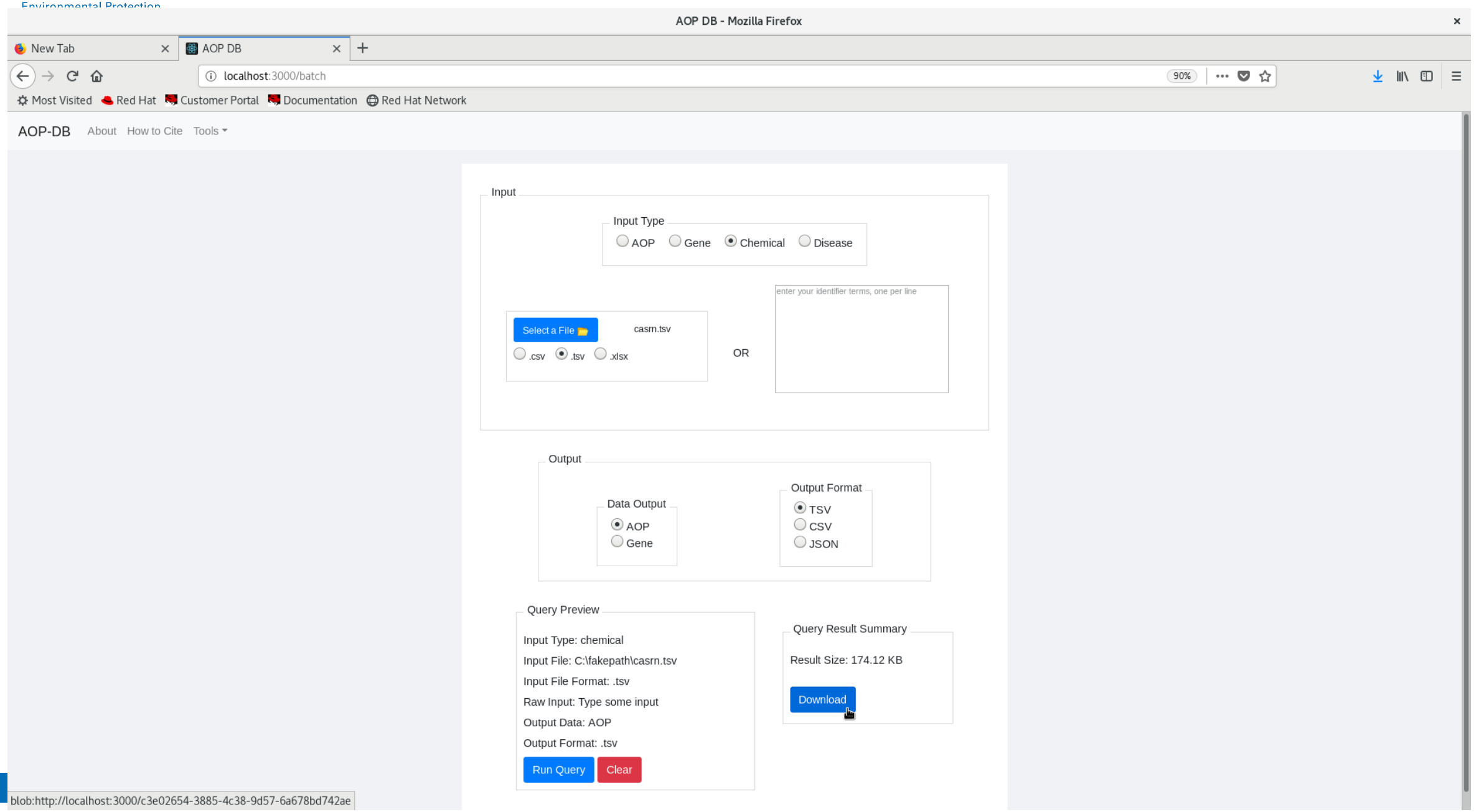
localhost:3000

Most Visited Red Hat Customer Portal Documentation Red Hat Network

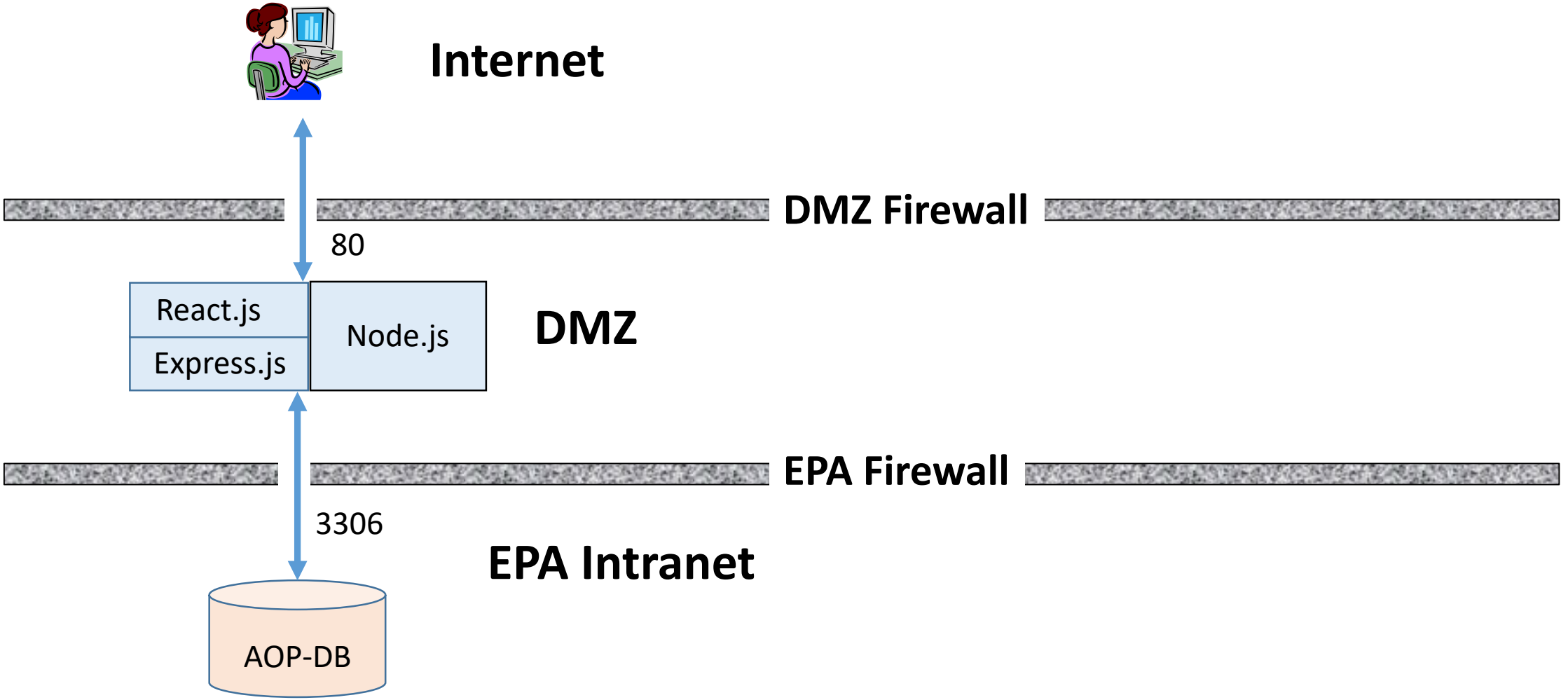
AOP-DB About How to Cite Tools

Enter a keyword (AOP, entrez, etc.)

Submit


AOP DB - Mozilla Firefox
New Tab
AOP DB
localhost:3000/batch
90%
Most Visited
Red Hat
Customer Portal
Documentation
Red Hat Network
AOP-DB
About
How to Cite
Tools
Input
Input Type
AOP
Gene
Chemical
Disease
Select a File
casrn.tsv
.csv
.tsv
.xlsx
OR
enter your identifier terms, one per line
Output
Data Output
AOP
Gene
Output Format
TSV
CSV
JSON
Query Preview
Input Type: chemical
Input File: C:\fakepath\casrn.tsv
Input File Format: .tsv
Raw Input: Type some input
Output Data: AOP
Output Format: .tsv
Run Query
Clear
Query Result Summary
Result Size: 174.12 KB
Download
blob:http://localhost:3000/c3e02654-3885-4c38-9d57-6a678bd742ae

AOP-DB Front-End Tool Creation



EPA Rule Will Make Its Custom Code Open Source By Default

BAKDC/SHUTTERSTOCK

OCTOBER 17, 2019

Sponsor Message

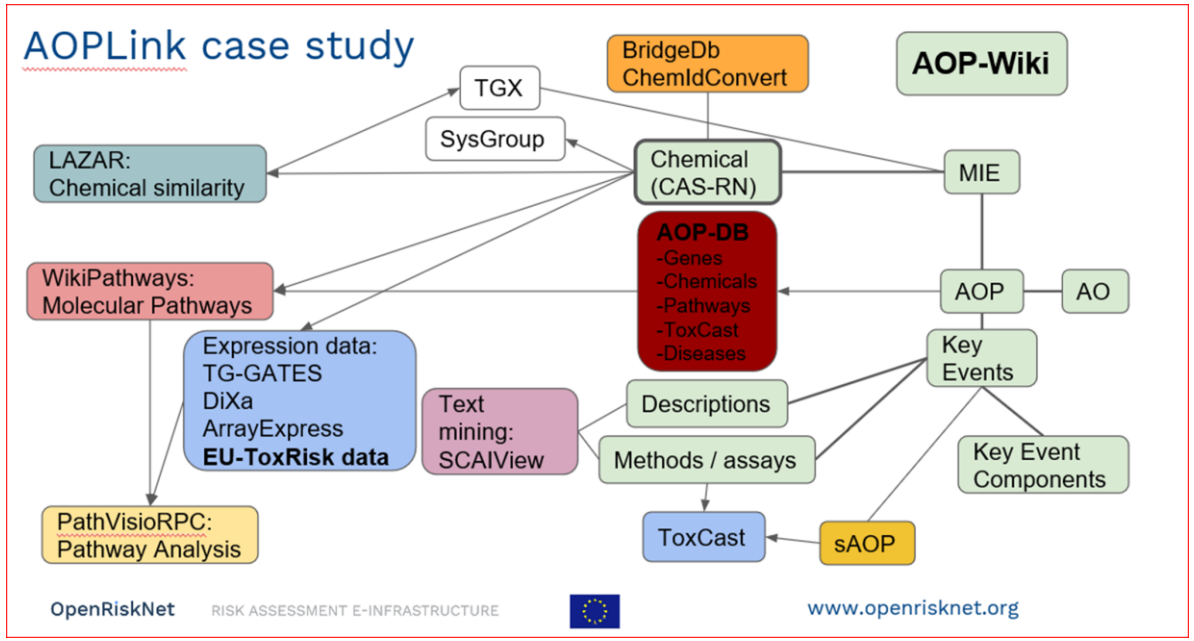
The New Cyber Landscape

A podcast from NextGov

LISTEN NOW

sponsored by Raytheon

OpenRiskNet Collaboration



Challenge Winner!! AOP-DB Semantic Annotation

- Resource Description Framework (RDF) conversion
- SPARQL queries
- Allows for custom queries and creation of automated workflows

Information from AOP-DB will contribute to the molecular AOP annotation in the AOP Portal on WikiPathways

www.wikipathways.org/index.php/Portal:AOP

European Union's Horizon 2020 research and innovation programme project EU-ToxRisk (grant agreement No. 681002)
EINFRA-22-2016 programme project OpenRiskNet under grant agreement No. 731075

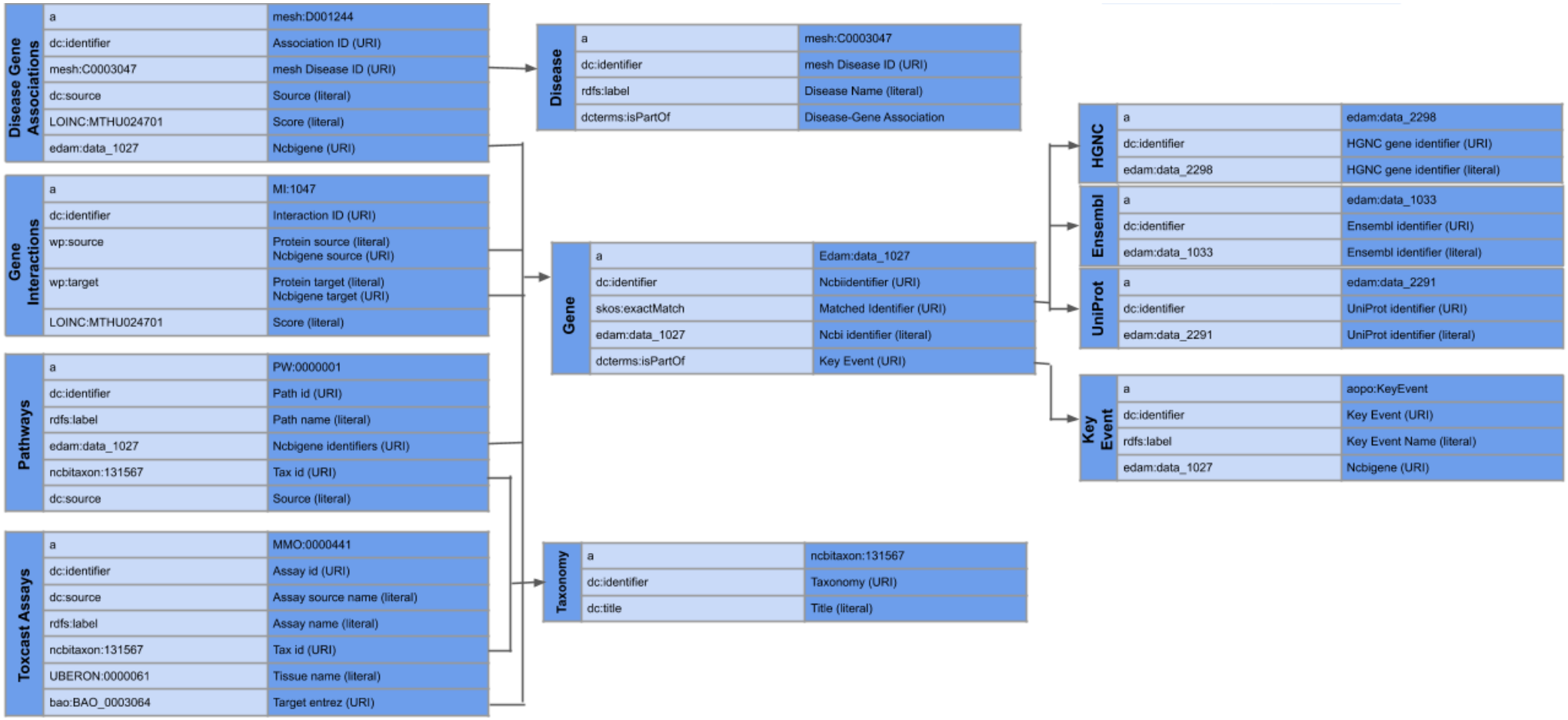
AOP-DB RDF Mapping

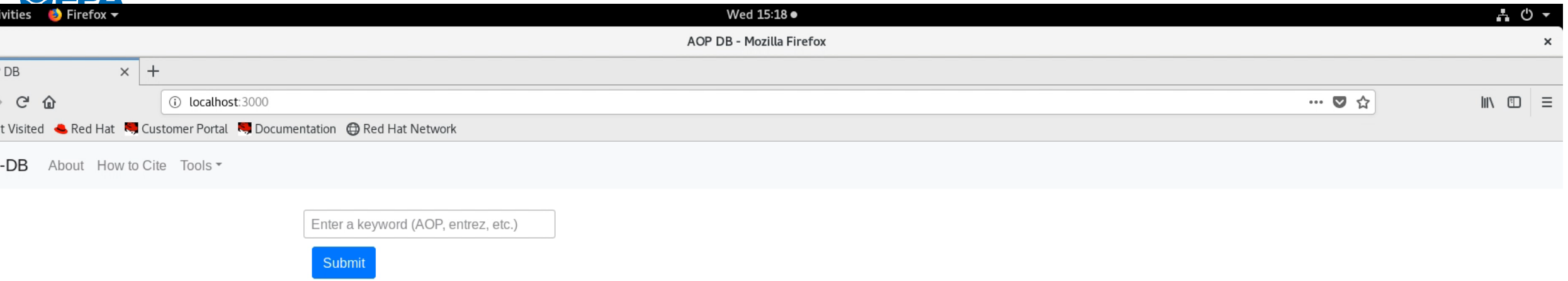
Disease-Gene

Gene Interaction

Pathway

ToxCast Assays




ivities
Firefox
Wed 15:18
AOP DB - Mozilla Firefox
DB
localhost:3000
t Visited
Red Hat
Customer Portal
Documentation
Red Hat Network
-DB
About
How to Cite
Tools
Enter a keyword (AOP, entrez, etc.)
Submit

WC11: 11th World Congress on Alternatives and Animal Use in the Life Sciences

23-27 August 2020 | MECC Maastricht – The Netherlands

Innovative Technologies Theme

Accepted Workshop:

Using the Sematic Web for Rapid Integration of Publicly Available Biological Information

Holly Mortensen, US EPA

Co-Chair: Alison Motsinger-Reif, NIEHS

mortensen.holly@epa.gov

Acknowledgements

EPA

Gary Ankley

Dan Villenueve

Trevor Levey, ORAU SSC

Tony Williams

Douglas Connect/Edelwiess

- Oki Noffisat
- Thomas Exner

University of Maastrict

- Marvin Martens
- Egon Willighagen
- Chris Evelo

RTI

Stephen Edwards

Publications

Mortensen, HM and Euling, SY (2013) Integrating mechanistic and polymorphism data to characterize human genetic susceptibility for environmental chemical risk assessment in the 21st century. TAAP 15; 271(3):395-404.

Pittman, M.E., Edwards, S.W., Ives, C., Mortensen, H.M. *(2018)* AOP-DB: A database resource for the exploration of Adverse Outcome Pathways through integrated association networks. *Toxicology and Applied Pharmacology* Mar 15; 343:71-83

Mortensen, H.M., Chamberlin, J., Joubert, B., Angrish, M, Sipes, N., Lee, J.S., Euling, S.Y. *(2018)* Leveraging human genetic and adverse outcome pathway (AOP) data to inform susceptibility in human health risk assessment. Mammalian Genome. Feb;29(1-2):190-204.

Mortensen, H.M., Levey, T., Langley, P., Williams, A. *(in prep)* EPA AOP-DB version 2 updates. *Nature Scientific Data.*

Mortensen, H.M., Levey, T. *et al*. *(in prep)* Computational data integration, interpretation, and visualization methods for characterizing human susceptibility to environmental chemicals.