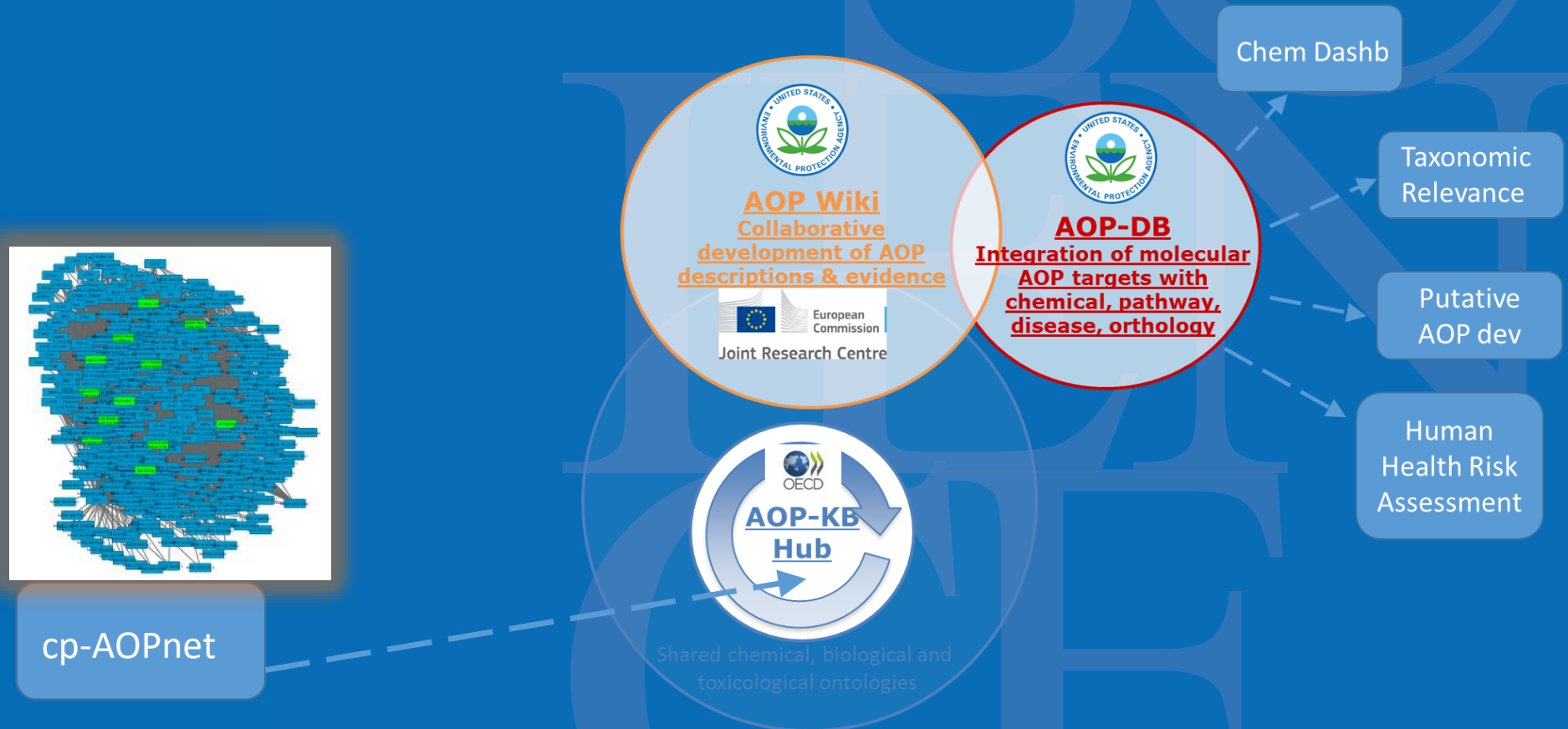


# AOP-DB: The Adverse Outcome Pathway Database

Holly M. Mortensen, Ph.D., National Health and Environmental Effects Laboratory



# AOP-DB – What is it?

Internal EPA SQL database that supports discovery and development of putative (existing) and potential AOPs

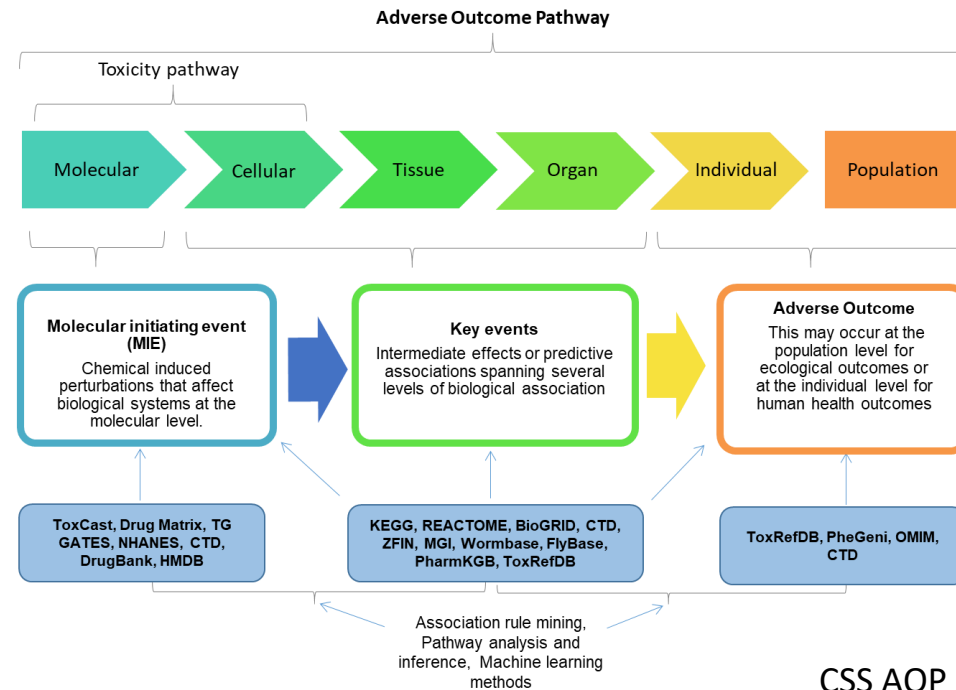
- The AOP-DB aggregates relationships between ***AOP-gene targets, chemical, disease, pathway, species orthology information, ontologies and gene interactions.***
- Associations are sourced from public annotation to provide ***biological context*** and are ***integrated with AOP information*** centralized in the AOP-Wiki
- AOP-DB allows for ***fast, automatic AOP profiling and exploration*** that gives a broad, systems-level overview of the ***biological context of AOPs***, thus dramatically expediting predictive toxicology efforts
- ***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 and disease endpoints.

## 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**



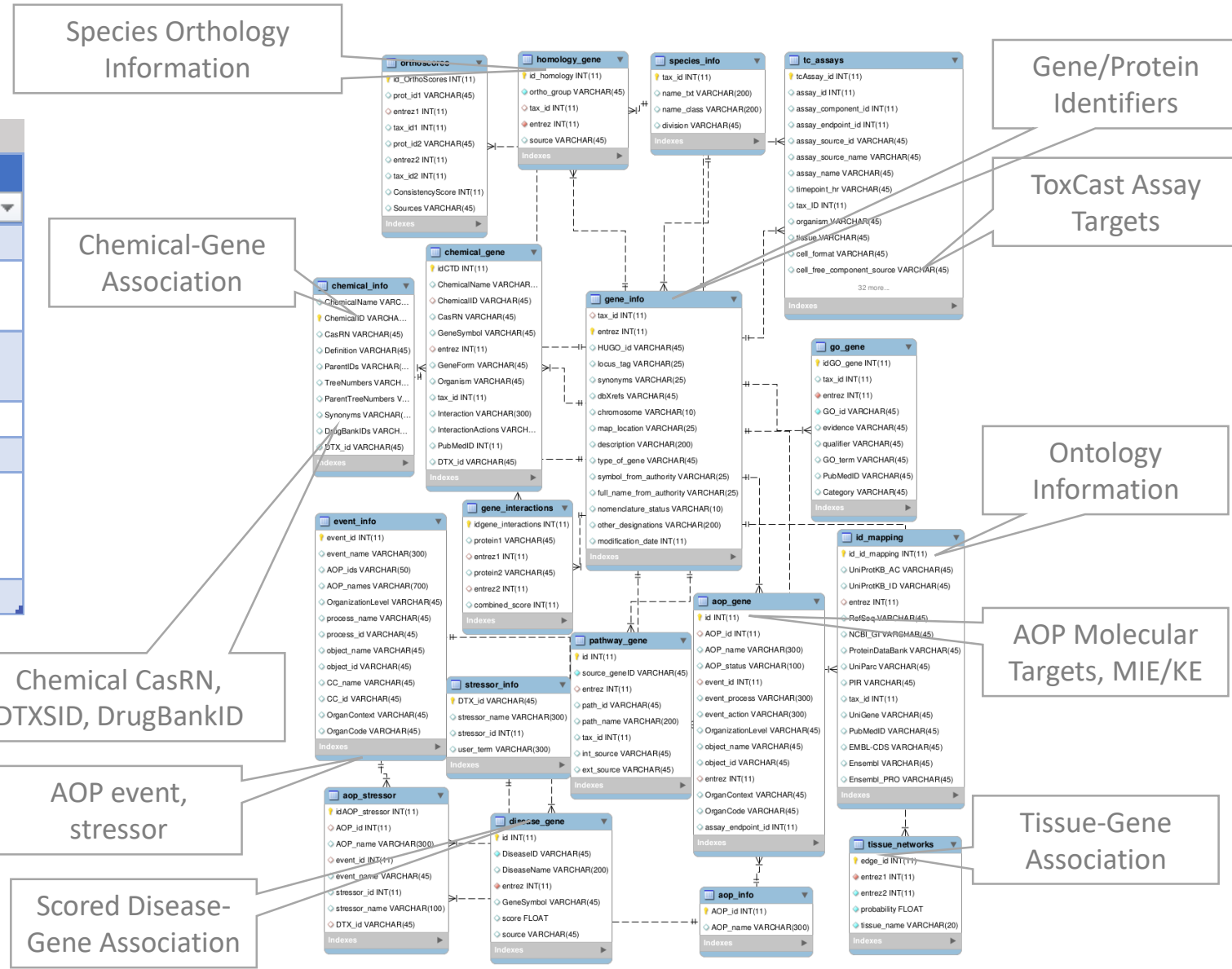
CSS AOP DD Task 1.1b Computationally-Predicted AOPs

[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.

# AOP-DB v.2

## AOP-DB v.2 Datasource Overview

Data Type	Unique Entry Count	Supported Organism Count	Source
AOP	236	21	AOPwiki
Gene_Protein	23189522	>2000	STRING
Pathway	100594	569	KEGG; Reactome; CPDB
Disease	24166	H. Sapiens Only	DisGeNET
Chemical	170646	592	CTD
Ortholog	64502	>5000	NCBI Homologene; KEGG Othology; phylomeDB
Tissues	144	H. Sapiens Only	Humanbase



Chemical CasRN, DTXSID, DrugBankID

AOP event, stressor

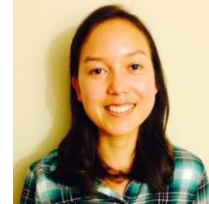
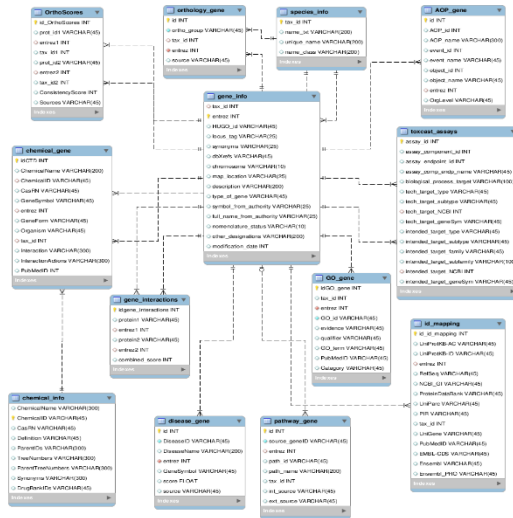
Scored Disease-Genes Association

Ontology Information

AOP Molecular Targets, MIE/KE

Tissue-Genes Association

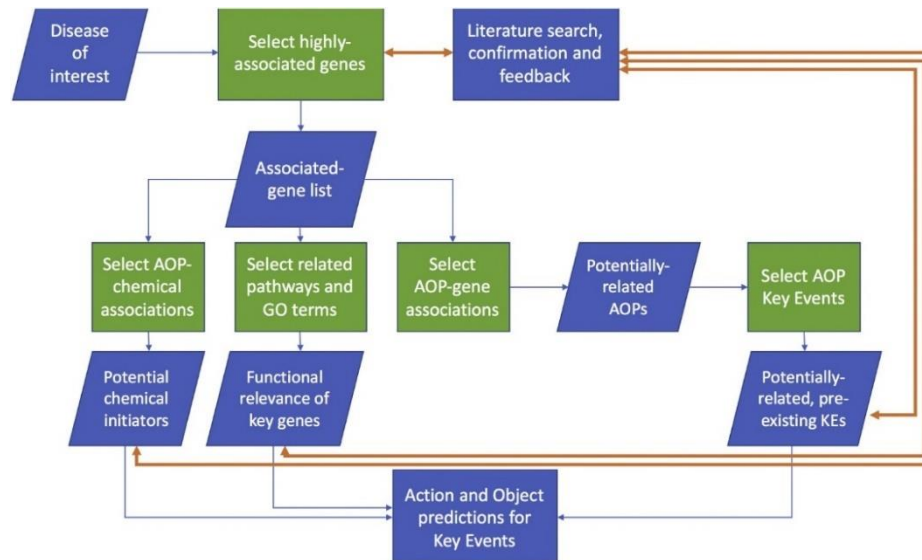
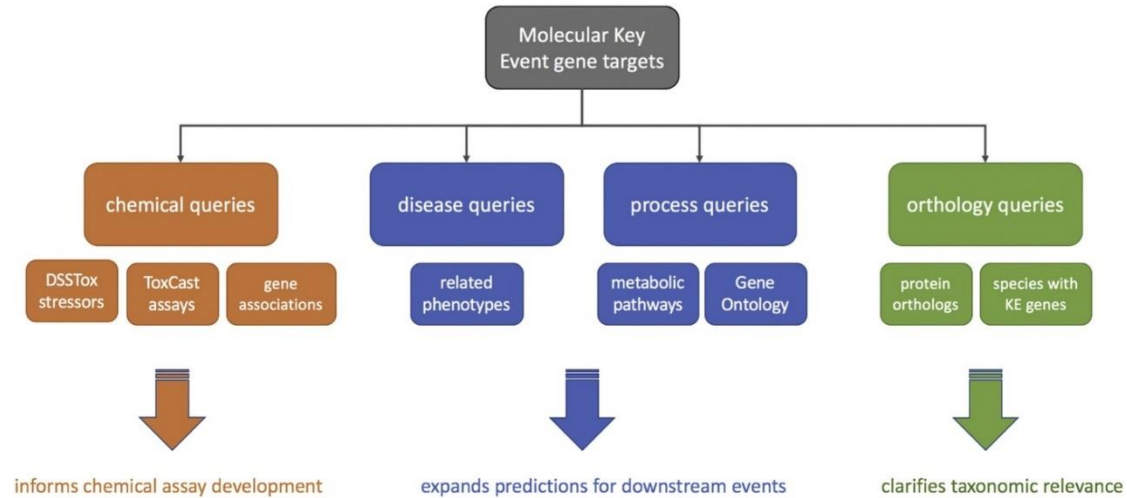
# AOP-DB: Ontology-curated AOP Gene Mapping



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

# AOP-DB Use 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.



# AOP-DB Use Examples

## AOP-directed Co-Expression Module Comparison



MIE and KE  
gene/protein

Define “Functional  
Neighborhood”

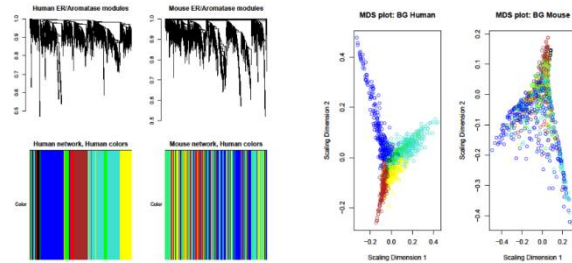
Inter-species Pathway  
Comparison

Molecular Sequence  
Similarity

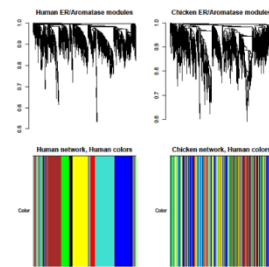
Co-Expression Network  
Construction

Co-Expression Pattern  
Analysis

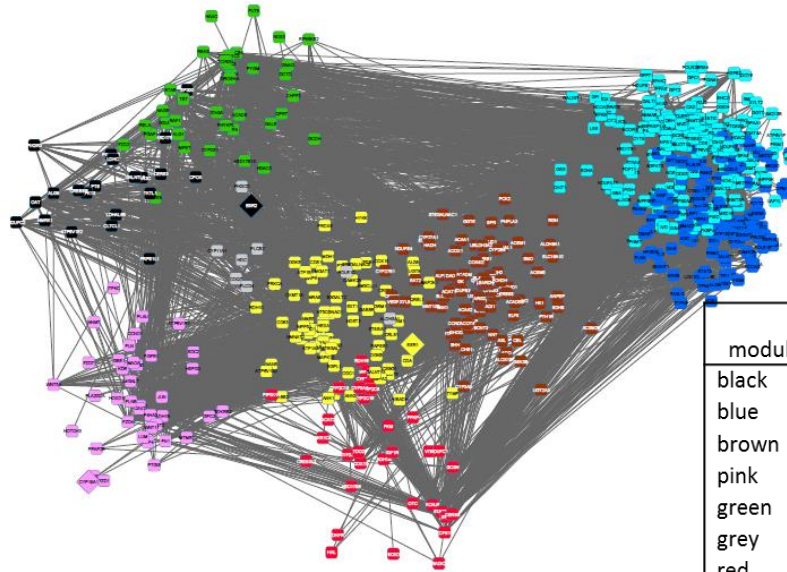
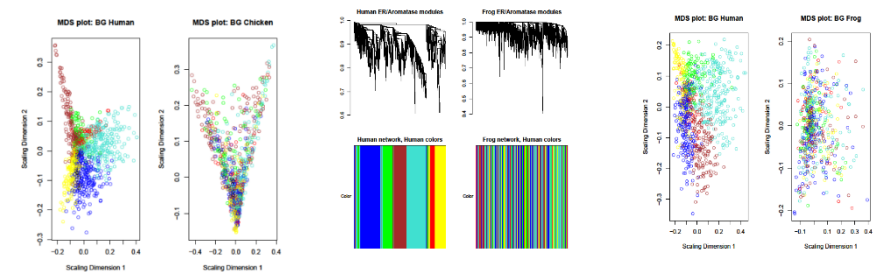
Human-Mouse



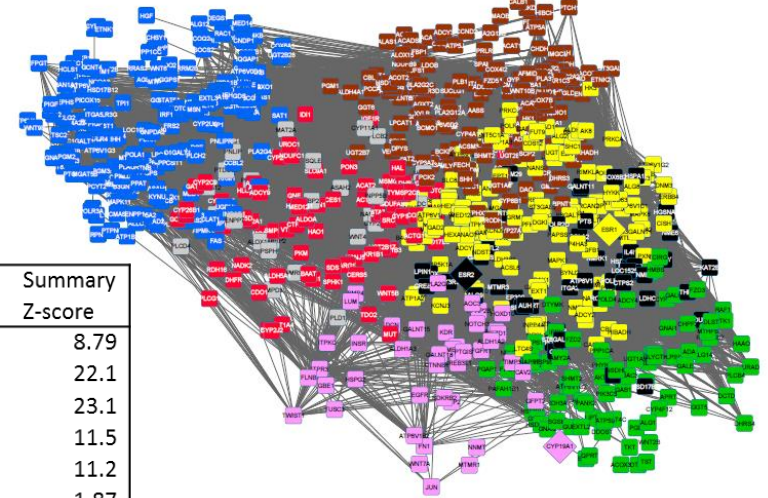
Human-Chicken



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

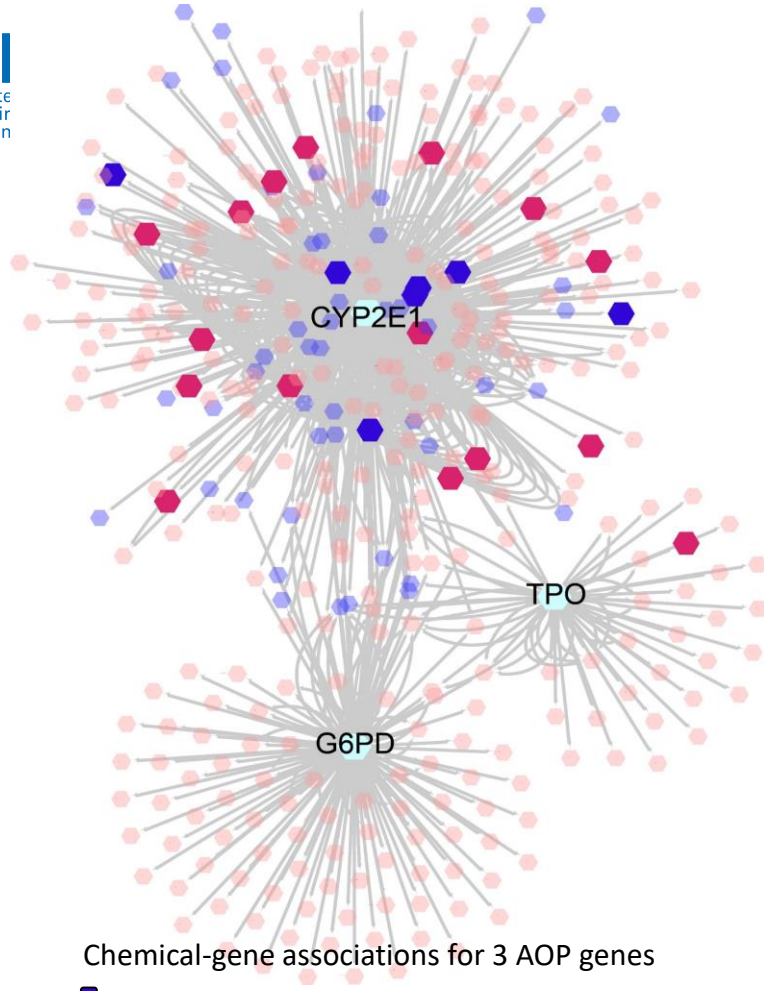


5 < Z score < 10 = Moderate Module Preservation  
Z score > 10 = High Module Preservation

**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*.

# AOP-DB Use Examples

## 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

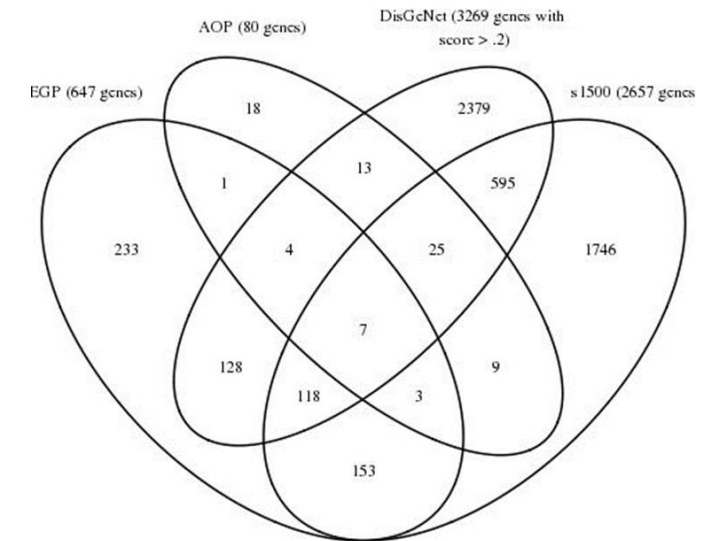
**Step 4: Identify functionally significant SNPs (Ensembl/Gtex)**

# 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 HHRA Key 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 Examples

## AOP-anchored Genetic Susceptibility/Community Risk



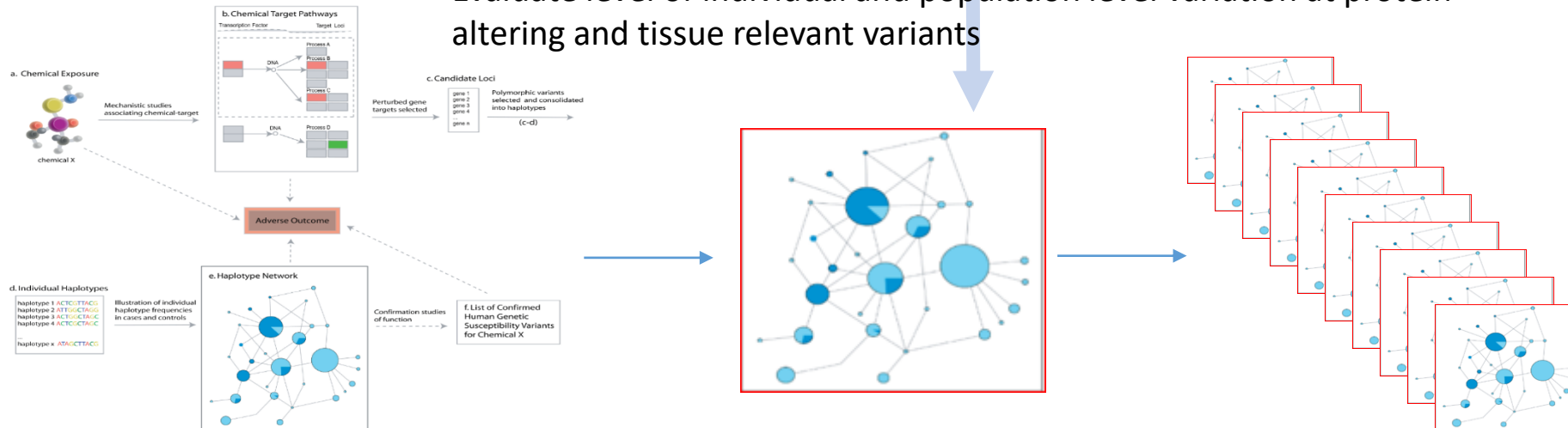
# AOP Genes ( MIE/KE from AOP-KB)

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 Update v.2



The backend code is currently being updated to incorporate the front-end tool

Code/Data updates include:

- Tools for directly modifying the DB; Automatic data pulls
- Data export in standard formats (eg. JSON/XML/CSV)
- About page and DB tutorial, and changelog to be implemented
- HumanBase Tissue Network integration
- Population frequencies for AOP-relevant SNP variants

## NHEERL Database Products

### NaKnowBase

NaKnowBase is a relational database populated with data from peer-reviewed ORD nanomaterials research publications. The database focuses on papers describing the actions of nanomaterials in environmental or biological media including their characterization, interactions, transformations and potential adverse environmental or toxicological effects.

For Information please contact Holly Mortensen (mortensen.holly@epa.gov) or William Boyes (boyes.william@epa.gov)

<https://intranet.ord.epa.gov/nheerl/nheerl-database-products>

<a href="#">nkbreadme.txt</a>	<a href="#">nkbreadme.pdf</a>
-------------------------------	-------------------------------

### Adverse Outcome Database

The Adverse Outcome Database (AOP-DB) links molecular data associated with AOPs from the AOP-Wiki to biological entities like diseases, chemicals, taxonomy, etc. These associations are sourced from public annotation and provide a wider context of relevant biology for each AOP, thus enabling researchers to expand predictions of chemical stressors and toxicological outcomes.

For Information please contact Holly Mortensen (mortensen.holly@epa.gov).

Adverse Outcome Database downloads:

Read Me	Data Dump	Schema Diagrams
Coming soon	<a href="#">dev_AOP-DB_v1.sql.gz (7.2 GB)</a>	<a href="#">dev_aop-db_v1.pdf</a>

# AOP-DB Front-End Tool Creation



AOP-DB
Home
About

Inhibition

AOP
  Gene


ENTREZ ▾


submit


**AOPS**


▶ AOP 3: Inhibition of the mitochondrial complex I of nigra-striatal neurons leads to parkinsonian motor deficits

ID	Stressor	Gene Name	Gene ID (ENTREZ)
3		NADH-ubiquinone oxidoreductase chain 1	4535
3		NADH-ubiquinone oxidoreductase chain 1	17716

  
[Chemical](#)

  
[Disease](#)

  
[Pathway](#)

  
[Taxa](#)

Gene Name	ID	Gene ID (ENTREZ)	Disease Score
NADH-ubiquinone oxidoreductase chain 1	3	4535	0.000542884
NADH-ubiquinone oxidoreductase chain 1	3	4535	0.00236703
NADH-ubiquinone oxidoreductase chain 1	3	4535	0.125005

How To Cite
Download DB







# AOP Database

Trevor Levey

# AOP-Wiki Data



## aop\_info

- ⚡ AOP\_id INT(11)
- ◇ AOP\_name VARCHAR(300)

## aop\_gene

- ⚡ id INT(11)
- ◇ AOP\_id INT(11)
- ◇ AOP\_name VARCHAR(300)
- ◇ AOP\_status VARCHAR(100)
- ◇ event\_id INT(11)
- ◇ event\_process VARCHAR(300)
- ◇ event\_action VARCHAR(300)
- ◇ event\_type VARCHAR(300)
- ◇ OrganizationLevel VARCHAR(45)
- ◇ object\_name VARCHAR(45)
- ◇ object\_id VARCHAR(45)
- ◇ entrez INT(11)
- ◇ OrganContext VARCHAR(45)
- ◇ OrganCode VARCHAR(45)
- ◇ assay\_endpoint\_id INT(11)

## event\_info

- ⚡ event\_id INT(11)
- ◇ event\_name VARCHAR(300)
- ◇ AOP\_ids VARCHAR(50)
- ◇ AOP\_names VARCHAR(700)
- ◇ OrganizationLevel VARCHAR(45)
- ◇ process\_name VARCHAR(45)
- ◇ process\_id VARCHAR(45)
- ◇ object\_name VARCHAR(45)
- ◇ object\_id VARCHAR(45)
- ◇ CC\_name VARCHAR(45)
- ◇ CC\_id VARCHAR(45)
- ◇ OrganContext VARCHAR(45)
- ◇ OrganCode VARCHAR(45)

## stressor\_info

- ⚡ DTX\_id VARCHAR(45)
- ◇ stressor\_name VARCHAR(300)
- ◇ stressor\_id INT(11)
- ◇ user\_term VARCHAR(300)

## aop\_stressor

- ⚡ idAOP\_stressor INT(11)
- ◇ AOP\_id INT(11)
- ◇ AOP\_name VARCHAR(300)
- ◇ event\_id INT(11)
- ◇ event\_name VARCHAR(45)
- ◇ stressor\_id INT(11)
- ◇ stressor\_name VARCHAR(100)
- ◇ DTX\_id VARCHAR(45)

These tables make use of the data from the AOP Wiki. The data is updated quarterly from an xml file containing all the wiki data.

# Gene Sources



id_mapping	
id_id_mapping	INT(11)
UniProtKB_AC	VARCHAR(45)
UniProtKB_ID	VARCHAR(45)
entrez	INT(11)
RefSeq	VARCHAR(45)
NCBI_GI	VARCHAR(45)
ProteinDataBank	VARCHAR(45)
UniParc	VARCHAR(45)
PIR	VARCHAR(45)
tax_id	INT(11)
UniGene	VARCHAR(45)
PubMedID	VARCHAR(45)
EMBL-CDS	VARCHAR(45)
Ensembl	VARCHAR(45)
Ensembl_PRO	VARCHAR(45)

gene_info	
tax_id	INT(11)
entrez	INT(11)
HUGO_id	VARCHAR(45)
locus_tag	VARCHAR(25)
synonyms	VARCHAR(25)
dbXrefs	VARCHAR(45)
chromosome	VARCHAR(10)
map_location	VARCHAR(25)
description	VARCHAR(200)
type_of_gene	VARCHAR(45)
symbol_from_authority	VARCHAR(25)
full_name_from_authority	VARCHAR(25)
nomenclature_status	VARCHAR(10)
other_designations	VARCHAR(200)
modification_date	INT(11)

go_gene	
idGO_gene	INT(11)
tax_id	INT(11)
entrez	INT(11)
GO_id	VARCHAR(45)
evidence	VARCHAR(45)
qualifier	VARCHAR(45)
GO_term	VARCHAR(45)
PubMedID	VARCHAR(45)
Category	VARCHAR(45)

gene_interactions	
idgene_interactions	INT(11)
protein1	VARCHAR(45)
entrez1	INT(11)
protein2	VARCHAR(45)
entrez2	INT(11)
combined_score	INT(11)

These tables include general gene information, id mappings between entrez and other gene ids, and gene interaction data

NCBI Gene: Source of gene information in AOP-DB.

UniProt: Supplies the gene id mapping data  
STRING: Contains protein-protein interactions for over 2000 organisms, scored by confidence level


# AOP Entrez Mapping

**aop\_gene**

- id INT(11)
- AOP\_id INT(11)
- AOP\_name VARCHAR(300)
- AOP\_status VARCHAR(100)
- event\_id INT(11)
- event\_process VARCHAR(300)
- event\_action VARCHAR(300)
- event\_type VARCHAR(300)
- OrganizationLevel VARCHAR(45)
- object\_name VARCHAR(45)
- object\_id VARCHAR(45)
- entrez INT(11)
- OrganContext VARCHAR(45)
- OrganCode VARCHAR(45)
- assay\_endpoint\_id INT(11)

**gene\_info**

- tax\_id INT(11)
- entrez INT(11)
- HUGO\_id VARCHAR(45)
- locus\_tag VARCHAR(25)
- synonyms VARCHAR(25)
- dbXrefs VARCHAR(45)
- chromosome VARCHAR(10)
- map\_location VARCHAR(25)
- description VARCHAR(200)
- type\_of\_gene VARCHAR(45)
- symbol\_from\_authority VARCHAR(25)
- full\_name\_from\_authority VARCHAR(25)
- nomenclature\_status VARCHAR(10)
- other\_designations VARCHAR(200)
- modification\_date INT(11)



- The entrez ids in the aop\_gene table do not come directly from the aop\_wiki xml
- AOPs with Protein Ontology object\_ids (PR) are mapped to entrez ids.

AOP_id	object_id	object_name
3	PR:000031316	NADH-ubiquinone oxidoreductase chain 1
276	PR:000031316	NADH-ubiquinone oxidoreductase chain 1
4	PR:P33244	nuclear hormone receptor FTZ-F1 (fruit fly)
6	PR:000013056	peroxisome proliferator-activated receptor al
6	PR:000011048	nuclear receptor corepressor 2
6	PR:000011047	nuclear receptor corepressor 1
7	PR:000006100	aromatase
8	PR:000011397	nuclear receptor subfamily 1 group I member 2
8	PR:000011398	nuclear receptor subfamily 1 group I member 3
10	PR:000007766	gamma-aminobutyric acid receptor subunit alph
10	PR:000007770	gamma-aminobutyric acid receptor subunit alph
10	PR:000007771	gamma-aminobutyric acid receptor subunit alph
11	PR:000011397	nuclear receptor subfamily 1 group I member 2
60	PR:000011397	nuclear receptor subfamily 1 group I member 2
12	PR:000004716	brain-derived neurotrophic factor
13	PR:000004716	brain-derived neurotrophic factor
54	PR:000004716	brain-derived neurotrophic factor
14	PR:000011406	glucocorticoid receptor
214	PR:000011406	glucocorticoid receptor
16	PR:000003626	acetylcholinesterase
18	PR:000013056	peroxisome proliferator-activated receptor al
37	PR:000013056	peroxisome proliferator-activated receptor al
51	PR:000013056	peroxisome proliferator-activated receptor al
61	PR:000013056	peroxisome proliferator-activated receptor al
18	PR:000015715	STAR
18	PR:000016757	translocator protein
19	PR:000004191	androgen receptor
21	PR:000003858	aryl hydrocarbon receptor
57	PR:000003858	aryl hydrocarbon receptor
131	PR:000003858	aryl hydrocarbon receptor
150	PR:000003858	aryl hydrocarbon receptor
21	PR:000004303	aryl hydrocarbon receptor nuclear translocato
150	PR:000004303	aryl hydrocarbon receptor nuclear translocato



# Taxonomy Sources and Use Example

## orthoscores

- id\_OrthoScores INT(11)
- prot\_id1 VARCHAR(45)
- entrez1 INT(11)
- tax\_id1 INT(11)
- prot\_id2 VARCHAR(45)
- entrez2 INT(11)
- tax\_id2 INT(11)
- ConsistencyScore INT(11)
- Sources VARCHAR(45)

## species\_info

- tax\_id INT(11)
- name\_txt VARCHAR(200)
- name\_class VARCHAR(200)
- division VARCHAR(45)

NCBI Taxonomy: Supplies basic information for species included in the species\_info table.

NCBI Homology Gene, KEGG Orthology, and MetaPhOrs: Includes the orthologous groups for genes.

Due to redundancy over these sources, confidence scores are computed in the orthoscores table.

## homology\_gene

- id\_homology INT(11)
- ortho\_group VARCHAR(45)
- tax\_id INT(11)
- entrez INT(11)
- source VARCHAR(45)



The table below is the result of a query looking at AOP 167, relating to carcinoma in the mouse, and relating it to taxonomy and homology information

AOP_id	AOP_name	entrez	tax_id	name_class	name_txt	ortho_group
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	2099	9606	scientific name	Homo sapiens	47906
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	396099	9031	scientific name	Gallus gallus	47906
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	13982	10090	scientific name	Mus musculus	47906
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	407238	9913	scientific name	Bos taurus	47906
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	259252	7955	scientific name	Danio rerio	47906
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	6495	9606	scientific name	Homo sapiens	4360
167	Early-life estrogen receptor activity leading to endometrial carcinoma in the mouse.	20471	10090	scientific name	Mus musculus	4360

# Chemical Sources



## EPA iCSS ToxCast Dashboard

tc_assays		
tcAssay_id INT(11)	cell_short_name VARCHAR(45)	assay_component_endpt_name VARCHAR(45)
assay_id INT(11)	cell_growth_mode VARCHAR(45)	assay_function_type VARCHAR(45)
assay_component_id INT(11)	assay_format_type VARCHAR(45)	burst_assay VARCHAR(45)
assay_endpoint_id INT(11)	assay_format_type_sub VARCHAR(45)	key_positive_control VARCHAR(45)
assay_source_id VARCHAR(45)	content_readout_type VARCHAR(45)	signal_direction VARCHAR(45)
assay_source_name VARCHAR(45)	assay_component_name VARCHAR(45)	intended_target_type VARCHAR(45)
assay_name VARCHAR(45)	parameter_readout_type VARCHAR(45)	intended_target_subtype VARCHAR(45)
timepoint_hr VARCHAR(45)	assay_design_type VARCHAR(100)	intended_target_family VARCHAR(45)
tax_id INT(11)	assay_design_sub VARCHAR(100)	intended_target_subfamily VARCHAR(100)
organism VARCHAR(45)	biological_process_target VARCHAR(100)	intended_target_entrez INT(11)
tissue VARCHAR(45)	detection_technology_type VARCHAR(45)	intended_target_officialSymbol VARCHAR(45)
cell_format VARCHAR(45)	detection_technology_type_sub VARCHAR(45)	intended_target_UniProtAC VARCHAR(45)
cell_free_component_source VARCHAR(45)	detection_technology VARCHAR(45)	tech_target_entrez INT(11)
	signal_direction_type VARCHAR(45)	tech_target_officialSymbol VARCHAR(45)
	key_assay_reagent_type VARCHAR(45)	tech_target_UniProtAC VARCHAR(45)
	key_assay_reagent VARCHAR(45)	
	tech_target_type VARCHAR(45)	

chemical_info
ChemicalName VARCHAR(300)
ChemicalID VARCHAR(45)
CasRN VARCHAR(45)
Definition VARCHAR(45)
ParentIDs VARCHAR(300)
TreeNumbers VARCHAR(300)
ParentTreeNumbers VARCHAR(300)
Synonyms VARCHAR(300)
DrugBankIDs VARCHAR(45)
DTX_id VARCHAR(45)

chemical_gene
idCTD INT(11)
ChemicalName VARCHAR(200)
ChemicalID VARCHAR(45)
CasRN VARCHAR(45)
GeneSymbol VARCHAR(45)
entrez INT(11)
GeneForm VARCHAR(45)
Organism VARCHAR(45)
tax_id INT(11)
Interaction VARCHAR(300)
InteractionActions VARCHAR(300)
PubMedID INT(11)
DTX_id VARCHAR(45)

The Comparative Toxicogenomics Database supplies manually curated chemical-gene relationships to the AOP\_DB.

The tc\_assays table comes from the EPA's ToxCast program, which screens chemicals for toxic effects.

# Disease Sources and Use Example

This table is the result of a query that retrieves all AOPs related to steatosis from the database.

AOP_id	AOP_name	entrez
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	18171
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	8856
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	84385
61	NFE2L2/FXR activation leading to hepatic steatosis	19013
61	NFE2L2/FXR activation leading to hepatic steatosis	25747
61	NFE2L2/FXR activation leading to hepatic steatosis	5465
61	NFE2L2/FXR activation leading to hepatic steatosis	403654
57	AhR activation leading to hepatic steatosis	11622
57	AhR activation leading to hepatic steatosis	196
34	LXR activation leading to hepatic steatosis	7376
34	LXR activation leading to hepatic steatosis	22260
34	LXR activation leading to hepatic steatosis	58851
58	NR1I3 (CAR) suppression leading to hepatic steatosis	7376
58	NR1I3 (CAR) suppression leading to hepatic steatosis	22260
58	NR1I3 (CAR) suppression leading to hepatic steatosis	58851
34	LXR activation leading to hepatic steatosis	10062
58	NR1I3 (CAR) suppression leading to hepatic steatosis	10062
34	LXR activation leading to hepatic steatosis	38868
34	LXR activation leading to hepatic steatosis	948
34	LXR activation leading to hepatic steatosis	12491
34	LXR activation leading to hepatic steatosis	180812
57	AhR activation leading to hepatic steatosis	38868
57	AhR activation leading to hepatic steatosis	948
57	AhR activation leading to hepatic steatosis	12491
57	AhR activation leading to hepatic steatosis	180812
58	NR1I3 (CAR) suppression leading to hepatic steatosis	38868
58	NR1I3 (CAR) suppression leading to hepatic steatosis	948
58	NR1I3 (CAR) suppression leading to hepatic steatosis	12491
58	NR1I3 (CAR) suppression leading to hepatic steatosis	180812
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	38868
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	948
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	12491
60	NR1I2 (Pregnane X Receptor, PXR) activation leading to hepatic steatosis	180812
34	LXR activation leading to hepatic steatosis	58805
34	LXR activation leading to hepatic steatosis	51085
58	NR1I3 (CAR) suppression leading to hepatic steatosis	58805

This table is the result of a query for diseases related to steatosis and their linked AOP information.

AOP_id	AOP_name	entrez	DiseaseName	source	score
57	AhR activation leading to hepatic steatosis	5105	Renal steatosis	HPO	0.1
57	AhR activation leading to hepatic steatosis	3949	Myocardial steatosis	HPO	0.1
61	NFE2L2/FXR activation leading to hepatic steatosis	338	Myocardial steatosis	HPO	0.1

Some Sources Within DisGeNET for Gene-Disease Associations:

Uniprot	HPO
CTD	CLINVAR
Orphanet	GWAS
CLINGEN	LHGDN
CGI	BeFree
PsyGeNET	

DisGeNET combines mined, curated, and inferred disease\_gene relationships from multiple sources for diseases and disease traits. These relationships are included in the AOP-DB, along with a score calculated from redundancy over the sources in DisGeNET.

disease\_gene

- id INT(11)
- DiseaseID VARCHAR(45)
- DiseaseName VARCHAR(200)
- entrez INT(11)
- GeneSymbol VARCHAR(45)
- score FLOAT
- source VARCHAR(45)

# Pathway Sources



Pathway\_Gene links biological process pathways to gene ids.

KEGG Pathways contains functional annotation of genes

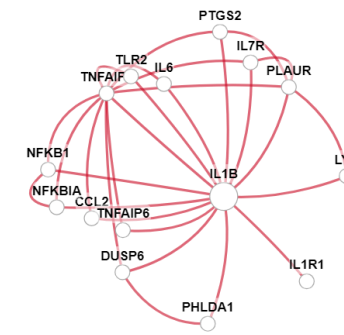
Reactome contains manually-curated pathway interactions

Consensus PathDB contains functional interaction data for human, mouse, and yeast pathways.

pathway_gene	
id	INT(11)
source_geneID	VARCHAR(45)
entrez	INT(11)
path_id	VARCHAR(45)
path_name	VARCHAR(200)
tax_id	INT(11)
int_source	VARCHAR(45)
ext_source	VARCHAR(45)



# Tissue Networks



tissue\_networks

- edge\_id INT(11)
- entrez1 INT(11)
- entrez2 INT(11)
- probability FLOAT
- tissue\_name VARCHAR(30)
- query\_gene INT(11)

Find all edges for the liver tissue network built around the gene SREBF1 (entrez=6720):

```
SELECT N.entrez1, N.entrez2, N.probability, N.tissue_name, N.query_gene, A.AOP_id,
A.AOP_name FROM tissue_networks AS N, aop_gene AS A WHERE N.query_gene=A.entrez
AND AOP_id=61 AND N.tissue_name="liver" AND N.query_gene="6720";
```

#	entrez1	entrez2	probability	tissue_name	query_gene	AOP_id	AOP_name
1	4792	7528	0.202793	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
2	4792	91663	0.09145...	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
3	4792	220323	0.095931	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
4	4792	120	0.117579	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
5	4792	4047	0.131298	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
6	4792	6310	0.135163	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
7	4792	6892	0.332241	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
8	4792	128	0.06017...	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
9	4792	1387	0.180979	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
10	4792	6319	0.08031...	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
11	4792	375790	0.144123	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
12	4792	10099	0.13839	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
13	4792	116447	0.08579...	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
14	4792	3949	0.40171	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis
15	4792	8659	0.09087...	liver	6720	61	NFE2L2/FXR activation leading to hepatic steatosis

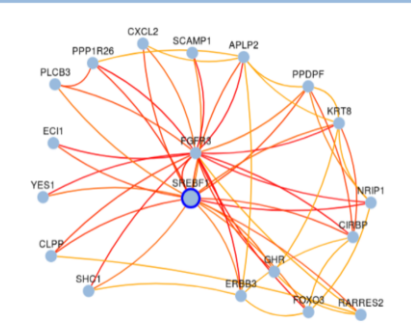
The tissue\_networks data table comes from humanbase. GIANT was used to create functional maps of gene relationships within 144 human tissues. Each gene interaction has a probability, which indicates the relationship confidence.

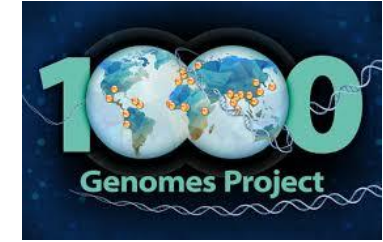
Query Gene: SREBF1      AOP: NFE2L2/FXR activation leading to hepatic steatosis      Tissue: Hepatocyte

Entrez ID	6720
Hugo ID	SREBF1
Tax ID	9606
Chromosome	17
Description	sterol regulatory element binding transcription factor 1

AOPs associated with this gene:

LXR activation leading to hepatic steatosis , NR1I3 (CAR) suppression leading to hepatic steatosis , AKT2 activation leading to hepatic steatosis , NFE2L2/FXR activation leading to hepatic steatosis





# Future Addition: SNP Frequency Table



- We are currently working on adding a table that retrieves all snps of interest for all AOP-genes and supplies population frequency data for these snps from the 1000Genomes project.
- To retrieve snps of interest, we are looking into sources like SIFT, Polyphen, and GTEX to find snps that have some effect on phenotype.

#	id	population	ref_allele_freq	alt_allele_freq	refsnp_id	entrez
87	87	East Asian	G=0.982	A=0.018	rs17886584	43
88	88	Europe	G=0.999	A=0.001	rs17886584	43
89	89	South Asian	G=0.97	A=0.03	rs17886584	43
90	90	American	G=0.99	A=0.01	rs17886584	43
91	91	Global	C=0.973	T=0.027	rs17886712	43
92	92	African	C=0.901	T=0.099	rs17886712	43
93	93	East Asian	C=1.000	T=0.000	rs17886712	43
94	94	Europe	C=1.000	T=0.000	rs17886712	43
95	95	South Asian	C=1.00	T=0.00	rs17886712	43
96	96	American	C=0.99	T=0.01	rs17886712	43
97	97	Global	A=0.329	G=0.671	rs553668	150
98	98	African	A=0.307	G=0.693	rs553668	150
99	99	East Asian	A=0.506	G=0.494	rs553668	150
100	100	Europe	A=0.160	G=0.840	rs553668	150
101	101	South Asian	A=0.40	G=0.60	rs553668	150

# AOP-DB Frontend

Phillip Langley

## Purpose

- To create a public facing user interface (React, NodeJS, Express, MySQL) that will allow simple searching of information from the AOP-DB.
- Enable ease-of-use utilities that do not require knowledge of SQL to allow access to AOP-DB information.
- Public hosting on EPA site.


## Plan to do

- Transition from jQuery frontend to React-driven client.
- Integration of updated backend code.
- Make API more accessible.

# Current Layout

AOP-DB

Welcome to the AOP-DB Front-end



How to Cite   Download DB   About   Batch Download

<https://i.imgur.com/tHHjPUC.gifv>

# Current Layout (cont.)

AOP-DB

AOP 34: LXR activation leading to hepatic steatosis

Gene **Stressor** Disease Pathway

Show **10** entries Search:

Entrez	HUGO ID	Object Name	Event ID	Event Process	Event Action	Event Type
948	CD36	platelet glycoprotein 4	54	CD36	Up Regulation	key-event
2194	FASN	fatty acid synthase	116	FAS	Activation	key-event
5468	PPARG	peroxisome proliferator-activated receptor ga	228	PPARG promoter	demethylation	molecular-initiating-event
6319	SCD	acyl-CoA desaturase	258	SCD-1	Activation	key-event
6720	SREBF1	sterol regulatory element-binding protein 1	264	SREBP-1c	Activation	key-event
7376	NR1H2	oxysterols receptor LXR-beta	167	LXR	Activation	molecular-initiating-event
10062	NR1H3	oxysterols receptor LXR-alpha	167	LXR	Activation	molecular-initiating-event
12491	Cd36	platelet glycoprotein 4	54	CD36	Up Regulation	key-event
14104	Fasn	fatty acid synthase	116	FAS	Activation	key-event
19016	Pparg	peroxisome proliferator-activated receptor ga	228	PPARG promoter	demethylation	molecular-initiating-event

Showing 1 to 10 of 21 entries Previous **1** 2 3 Next

Copy Excel CSV PDF

How to Cite Download DB About Batch Download

AOP-DB

Enter an AOP, Gene, or Keyword...

Search:

Search:

AOPS Genes

AOP 34: LXR activation leading to hepatic steatosis [AOPWiki link](#)

- Entrez 10062: oxysterols receptor LXR-alpha
- Entrez 22260: oxysterols receptor LXR-beta
- Entrez 58851: oxysterols receptor LXR-beta
- Entrez 7376: oxysterols receptor LXR-beta

AOP-DB

AOP 34: LXR activation leading to hepatic steatosis

Gene **Stressor** Disease Pathway

Stressor ID	Stressor Name	DTX ID
65	Di(2-ethylhexyl) phthalate	DTXSID5020607
70	Troglitazone	DTXSID8023719
304	Tributyltin	DTXSID0040709

Copy Excel CSV PDF



# Current Layout (cont.)

AOP-DB

Entrez 10062: NR1H3

AOP	Tissue		
AOP ID	AOP Name		
34	LXR activation leading to hepatic steatosis	<a href="#">AOPWiki link</a>	
58	NR1I3 (CAR) suppression leading to hepatic steatosis	<a href="#">AOPWiki link</a>	
61	NFE2L2/FXR activation leading to hepatic steatosis	<a href="#">AOPWiki link</a>	
Copy	Excel	CSV	PDF

AOP-DB

Stressor 70: Troglitazone

AOP ID	AOP Name	
34	LXR activation leading to hepatic steatosis	<a href="#">AOPWiki link</a>
58	NR1I3 (CAR) suppression leading to hepatic steatosis	<a href="#">AOPWiki link</a>
72	Epigenetic modification of PPARG leading to adipogenesis	<a href="#">AOPWiki link</a>
163	PPARGgamma activation leading to sarcomas in rats, mice, and hamsters	<a href="#">AOPWiki link</a>