

The EPA Adverse Outcome Pathway Database (AOP-DB)

Application User Manual

Center for Public Health and Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC
Last Update: August 25th, 2021

Contents

1	Introduction	2
1.1	The Adverse Outcome Pathway Database in Context	3
2	Searching	4
2.1	Genes Query	5
2.2	Stressors Query	6
2.3	Diseases Query	7
2.4	Biological Pathways Query	8
2.5	Batch Search	9
3	How Data Are Maintained	11
3.1	New Records and Updates	11
4	How Content is Selected	14
5	References	16

1 Introduction

There is a need for approaches to understand the biological mechanism of adverse outcomes and human variability in response to environmental chemical exposure. A recent legislation, the Frank R. Lautenberg Chemical Safety for the twenty-first Century Act of 2016 (114–182 2016), requires the US Environmental Protection Agency to evaluate new and existing toxic chemicals with explicit consideration of susceptible populations of all types (life stage, exposure, genetic, etc.). In addition, on September 10, 2019, EPA Administrator Andrew Wheeler signed a directive that prioritizes efforts to reduce animal testing. In response to this directive, the EPA has developed a 2019 Strategic Plan to Promote the Development and Implementation of Alternative Test Methods Strategies (or New Approach Methodologies (NAMs)) per TSCA Section 4(h)(2)(C). The EPA Adverse Outcome Pathway Database (AOP-DB) is a decision support tool developed by the EPA’s Center for Public Health and Environmental Assessment, which contributes to NAMs (e.g. computational toxicology tools) used for TSCA. The EPA Adverse Outcome Pathway Database (AOP-DB) is a database resource that combines different data types (AOP, gene, chemical, disease, pathway, orthology, and ontology) to characterize the impacts of chemicals to human health and the environment (Pittman, et al., 2018), and for the characterization of human genetic susceptibility for the purpose of human health risk assessment (Mortensen, et al., 2018). The AOP-DB was originally developed with the primary aim of integrating AOP molecular target information with other publicly available datasets and related toxicological data. Updates to the AOP-DB in version 2 (Mortensen, et al., 2021) were made primarily to facilitate and improve computational analyses of AOP information. Near term goals for use of the AOP-DB are to address the biological and mechanistic aspects of alternative test methods in terms of the adverse outcome pathway construct to facilitate Integrated Approaches to Testing and Assessment (IATA) for regulatory purposes (Delrue, Sachana et al. 2016, Patlewicz, Worth et al. 2016, Sakuratani, Horie et al. 2018), and serve as a decision support tool for case study development.

1.1 The Adverse Outcome Pathway Database in Context

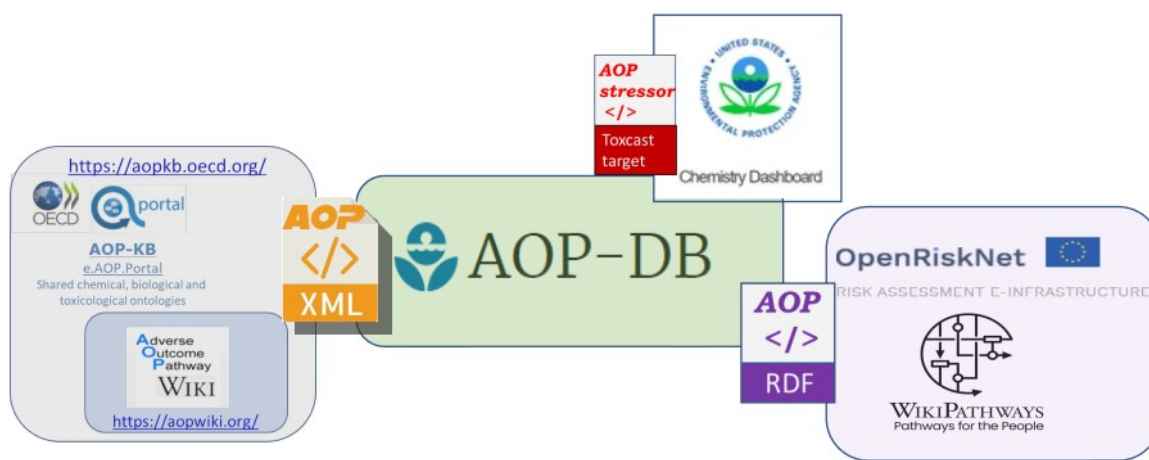


Figure 1: How AOP-DB relates to other publicly available tools

the OECD has launched a project to develop the “Adverse Outcome Pathway Knowledge Base” (AOP-KB) to enable the scientific community to share, develop and discuss AOP related knowledge in one central location. The AOP-KB allows all interested parties and stakeholders to build AOPs by entering and linking information about key events, molecular initiating events, adverse outcomes and stressors, including chemical initiators. The AOP-DB is a part of the AOP-KB third party toolset, contributing a resource to analyze AOP associations and construct putative AOPs for further study. [Figure 1](#) illustrates AOP-DB’s connection to the AOP-KB, specifically the AOP-Wiki, and other tools like EU supported OpenRiskNet’s WikiPathways, and illustrates how data are shared and transferred.

2 Searching for Information using the AOP-DB User Interface

To query AOP-DB enter a term for any of the six parameters listed in [Table 1](#) and select the “Match By” boxes for the parameters of interest. Searching on any of the parameters will return a list of AOPs with the matching term. Note that capitalization does not matter.

Domain	Parameter
AOP	AOP Name
	AOP ID
Gene	Entrez ID
	HUGO ID
Stressor	Stressor Name
	DTXS ID
Disease	Disease Name

Table 1: Available search parameters.

EPA United States Environmental Protection Agency

Environmental Topics | Laws & Regulations | About EPA | Search EPA.gov

Related Topics: [Safer Chemicals Research](#) | [Health Research](#) | [Risk Assessment](#) | [Contact Us](#)

Adverse Outcome Pathway Database (AOP-DB)

[Home](#) | [Basic Info](#) | [Search](#) | [Batch](#) | [Resources](#)

Match By: ☒ AOP Name ☐ AOP ID ☐ Entrez ID ☐ Disease Name ☐ Stressor Name ☐ DTXSID

Search:

Mode: ☒ Contains ☐ Exact

Submit

Show entries **Copy** **Excel** **CSV** **PDF** Search:

AOP ID	Name	AOPWiki
No data available in table		

AOP ID Name AOPWiki

Showing 0 to 0 of 0 entries

Figure 2: Screenshot of the AOP-DB search page

Associated with each AOP are four types of information: [genes](#), [diseases](#), [stressors](#), and [DTXID](#). To view these tables in the AOP-DB, click the radio button and enter associated information of interest in the search window. Each table can be filtered by any of the column values by entering a search term in the provided search box. The filtered results can then be exported as a csv, Excel, or PDF file for local use by clicking the corresponding button or it can be copied to the clipboard and pasted elsewhere. Second level queries can be performed by selecting information of interest (e.g. select green arrow for individual AOP ID, as illustrated in Figure 3 below).

The screenshot displays the AOP-DB web interface on the left and an Excel spreadsheet on the right. The web interface shows a search for 'AHR' with results for AOP IDs 41, 57, 151, and 310. The Excel spreadsheet, titled 'AOP-DB | US EPA', contains the same data in a table format.

AOP ID	Name	AOPWiki
41	Sustained AhR Activation leading to Rodent Liver Tumours	AOPWiki
57	AhR activation leading to hepatic steatosis	AOPWiki
151	AhR activation leading to placental insufficiency	AOPWiki
310	Embryonic Activation of the AHR leading to Reproductive failure, via epigenetic down-regulation of GnRH	AOPWiki

Figure 3: Screenshot of a sample query being exported and the resulting file.

2.1 AOP-DB Gene Query

AOP-gene links are only created by mapping protein IDs, provided by AOP-Wiki in the key event component field, to gene IDs using UniProt source mapping for exact gene mapping. Genes linked in this way can be viewed in the gene table.

Adverse Outcome Pathway Database (AOP-DB)

[Home](#) | [Basic Info](#) | [Search](#) | [Batch](#) | [Resources](#)

Match By: ☒ AOP Name ☐ AOP ID ☐ Entrez ID ☐ Disease Name ☐ Stressor Name ☐ DTXSID

Search:

Mode: ☒ Contains ☐ Exact

Submit

Show entries

Copy

Excel

CSV

PDF

Search:

AOP ID

Name

AOPWiki

41

Sustained AhR Activation leading to Rodent Liver Tumours

AOPWiki

57

AhR activation leading to hepatic steatosis

AOPWiki

Gene

Stressor

Disease

Pathway

Show

10

entries

Copy

Excel

CSV

PDF

Search

Entrez

HUGO ID

Object Name

Event ID

Event Process

Event Action

Event Type

Tax Id

196

AHR

aryl hydrocarbon receptor

18

aryl hydrocarbon receptor activity

increased

molecular-initiating-event

9606

948

CD36

platelet glycoprotein 4

54

gene expression

increased

key-event

9606

1543

CYP1A1

cytochrome P450 1A1

80

gene expression

increased

key-event

9606

3949

LDLR

low-density lipoprotein receptor

466

gene expression

increased

key-event

9606

5105

PCK1

phosphoenolpyruvate carboxykinase, cytosolic [GTP]

216

gene expression

decreased

key-event

9606

6319

SCD

acyl-CoA desaturase

462

gene expression

increased

key-event

9606

11622

Ahr

aryl hydrocarbon receptor

18

aryl hydrocarbon receptor activity

increased

molecular-initiating-event

10090

12491

Cd36

platelet glycoprotein 4

54

gene expression

increased

key-event

10090

13076

Cyp1a1

cytochrome P450 1A1

80

gene expression

increased

key-event

10090

16835

Ldlr

low-density lipoprotein receptor

466

gene expression

increased

key-event

10090

Showing 1 to 10 of 17 entries

First

Previous

1

2

Next

Last

151

AhR activation leading to placental insufficiency

AOPWiki

Figure 4: Screenshot of the filtered gene table associated with an AOP.

2.2 Stressors Query

Direct AOP-stressor associations in the AOP-DB are provided by AOP-Wiki. Stressors entered into the AOP-Wiki can include a link to chemical stressors, via the DSSTox Substance Identifier (DTXSID), which maps the stressor to substances registered in the DSSTox database (Richard and Williams, 2002). The chemical DTXSID, a unique substance identifier, provides a link to the Dashboard using the process described in Williams (2017). When no DTXSID is provided for stressors imported from the AOP-DB, manual curation to the Dashboard has been performed on

individual substances, on a substance-by-substance basis and using available identifiers (*e.g.* CAS Registry Numbers and chemical names) according to the process described in Grulke (2019).

Adverse Outcome Pathway Database (AOP-DB)

[Home](#) | [Basic Info](#) | [Search](#) | [Batch](#) | [Resources](#)

Match By: ☒ AOP Name ☐ AOP ID ☐ Entrez ID ☐ Disease Name ☐ Stressor Name ☐ DTXSID

Search:

Mode: ☒ Contains ☐ Exact

Submit

Show entries **Copy** **Excel** **CSV** **PDF**

Search:

AOP ID	Name	AOPWiki
41	Sustained AhR Activation leading to Rodent Liver Tumours	AOPWiki
57	AhR activation leading to hepatic steatosis	AOPWiki

Gene **Stressor** **Disease** **Pathway**

Show entries **Copy** **Excel** **CSV** **PDF**

Search:

Stressor ID	Stressor Name	DTX ID
62	Benzidine	DTXSID2020137
147	Dibenzo-p-dioxin	DTXSID8020410
148	Polychlorinated biphenyl	DTXSID5024267
149	Polychlorinated dibenzofurans	
150	Hexachlorobenzene	DTXSID2020682
250	Polycyclic aromatic hydrocarbons (PAHs)	DTXSID3044043

Showing 1 to 6 of 6 entries

First Previous 1 Next Last

AOP ID	Name	AOPWiki
151	AhR activation leading to placental insufficiency	AOPWiki
310	Embryonic Activation of the AHR leading to Reproductive failure, via epigenetic down-regulation of GnRHR	AOPWiki

Showing 1 to 4 of 4 entries

First Previous 1 Next Last

Figure 5: Screenshot of the filtered stressor table associated with an AOP.

2.3 AOP-DB Diseases Query

The associations between genes and human disease phenotypes in the AOP-DB are sourced from DisGeNET, which combines mined, curated, and inferred associations from ten sources for Mendelian, complex, environmental, and rare diseases as well as disease traits. Due to the redundancy of information across these ten data sources, a confidence score between 0 and 1 was calculated for each association based on the proportion of the sources that recognize that association.

Adverse Outcome Pathway Database (AOP-DB)

[Home](#) | [Basic Info](#) | [Search](#) | [Batch](#) | [Resources](#)

Match By: ☒ AOP Name ☐ AOP ID ☐ Entrez ID ☐ Disease Name ☐ Stressor Name ☐ DTXSID

Search:

Mode: ☒ Contains ☐ Exact

Submit

Show entries **Copy** **Excel** **CSV** **PDF** Search:

AOP ID	Name	AOPWiki
41	Sustained AhR Activation leading to Rodent Liver Tumours	AOPWiki
57	AhR activation leading to hepatic steatosis	AOPWiki

Gene **Stressor** **Disease** **Pathway**

Show entries **Copy** **Excel** **CSV** **PDF** Search:

Entrez	Disease ID	Disease Name	Score 0-1
196	C0700501	Congenital nystagmus	0.5
196	C0678222	Breast Carcinoma	0.4
196	C0006142	Malignant neoplasm of breast	0.4
196	C1458155	Mammary Neoplasms	0.4
196	C0023903	Liver neoplasms	0.4
196	C0376358	Malignant neoplasm of prostate	0.37
196	C0003873	Rheumatoid Arthritis	0.37
196	C0152013	Adenocarcinoma of lung (disorder)	0.36
196	C0027627	Neoplasm Metastasis	0.35
196	C0024623	Malignant neoplasm of stomach	0.34

Showing 1 to 10 of 1,121 entries **First** **Previous** **1** **2** **3** **4** **5** ... **113** **Next** **Last**

151	AhR activation leading to placental insufficiency	AOPWiki
310	Embryonic Activation of the AHR leading to Reproductive failure, via epigenetic down-regulation of GnRHR	AOPWiki

AOP ID **Name** **AOPWiki**

Showing 1 to 4 of 4 entries **First** **Previous** **1** **Next** **Last**

Figure 6: Screenshot of the filtered associated diseases table for an AOP.

2.4 AOP-DB Biological Pathways Query

Biological pathways represent the series of molecular and genetic interactions that amount to the execution of a biological process. The AOP-DB directly extracts pathway information from three sources: the Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and Consensus Path DB. That data is associated with a given AOP via the Entrez ID.

Adverse Outcome Pathway Database (AOP-DB)

[Home](#) | [Basic Info](#) | [Search](#) | [Batch](#) | [Resources](#)

Match By: ☒ AOP Name ☐ AOP ID ☐ Entrez ID ☐ Disease Name ☐ Stressor Name ☐ DTXSID

Search:

Mode: ☒ Contains ☐ Exact

Submit

Show entries

Copy

Excel

CSV

PDF

Search:

AOP ID

Name

AOPWiki

41

Sustained AhR Activation leading to Rodent Liver Tumours

AOPWiki

57

AhR activation leading to hepatic steatosis

AOPWiki

Gene

Stressor

Disease

Pathway

Show

10

entries

Copy

Excel

CSV

PDF

Search

Entrez

Path ID

Path Name

Path Source

Tax ID

196

ahrpathway

ahr signal transduction pathway

ConsensusPathDB

9606

196

hsa04659

Th17 cell differentiation

KEGG Pathways

9606

196

hsa04934

Cushing syndrome

KEGG Pathways

9606

196

Pathway_AndrogenReceptor

AndrogenReceptor

ConsensusPathDB

9606

196

R-HSA-1989781

PPARA activates gene expression

Reactome

9606

196

R-HSA-211945

Phase I - Functionalization of compounds

Reactome

9606

196

R-HSA-211976

Endogenous sterols

Reactome

9606

196

R-HSA-211981

Xenobiotics

Reactome

9606

196

R-HSA-8937144

Aryl hydrocarbon receptor signalling

Reactome

9606

196

WP1984

Integrated Breast Cancer Pathway

ConsensusPathDB

9606

Showing 1 to 10 of 362 entries

First

Previous

1

2

3

4

5

...

37

Next

Last

151

AhR activation leading to placental insufficiency

AOPWiki

310

Embryonic Activation of the AHR leading to Reproductive failure, via epigenetic down-regulation of GnRHR

AOPWiki

AOP ID

Name

AOPWiki

Showing 1 to 4 of 4 entries

First

Previous

1

Next

Last

Figure 7: Screenshot of the filtered associated biological pathways table for an AOP.

2.5 AOP-DB Batch Search

In addition to the queries described above, data from a particular domain associated with multiple AOPs can be retrieved in a single query using the AOP_DB Batch query tool. There are three options that must be selected: input type is the domain to search on (AOP, stressors, genes, or diseases), output type is the data to be retrieved, and output format is the output file format (tsv, csv, or json). Note that the input type and output selections cannot be the same, once an input type is selected the output type of the same name will be greyed out and not available for selection.

The screenshot shows the EPA website header with the logo and navigation links: Environmental Topics, Laws & Regulations, About EPA, and a search bar. Below the header, there are links for Related Topics: Safer Chemicals Research, Health Research, and Risk Assessment, along with a Contact Us link. The main heading is "Adverse Outcome Pathway Database (AOP-DB)". Below the heading are links for Home, Basic Info, Search, Batch, and Resources. The Batch Search Tool interface is displayed, featuring a form with the following elements:

- Input Type:** A dropdown menu with a yellow question mark icon and a red 'X' icon. The selected option is "AOP". Other options are "Gene", "Stressor", and "Disease".
- Output Type:** A dropdown menu with a yellow question mark icon and a red 'X' icon. The selected option is "AOP". Other options are "Gene", "Stressor", and "Disease".
- Output Format:** A dropdown menu with a yellow question mark icon and a red 'X' icon. The selected option is "JSON". Other options are "CSV" and "TSV".
- Input Data:** A text box with a yellow question mark icon and a red 'X' icon. It contains the text "Enter data manually or load data from a file." Below the text box are buttons for "Validate" and "Clear".
- Upload a File:** A button with a yellow question mark icon and a red 'X' icon. It is labeled "Upload a File" and "No File Chosen".
- Submit and Cancel:** Buttons at the bottom of the form with yellow question mark icons.

Figure 8. Screenshot of the Batch Search Tool.

The batch search tool provides help text for all fields which can be viewed by hovering the mouse over the yellow question marks, help text, accepted parameters, and sample inputs for each of the search domains are displayed in the input data text box. Search terms are entered either in the input data text box directly or a file using the “Upload a File” button which will automatically populate the text box with file data. Note, input file formats should be in csv, tsv, or txt format with comma or tab delimiters and files should not have column headers. When using the input text box all search terms should be comma separated and terms that are names (*e.g.* the stressor “ibuprofen”) should be enclosed in double quotes.

Then final step before submitting the query is format validation, which is accomplished by clicking the “Validate” button. Validation will verify that all terms match an expected format, reformat them if possible or remove them otherwise.

Domain	Parameter
AOP	AOP Name AOP ID
Gene	Entrez ID
Stressor	Stressor Name DTXS ID CASRN MESH ID
Disease	Disease Name UMLS Disease Concept ID

Table 2: Available search parameters in Batch Search.

3 How Data Are Maintained

3.1 New Records and Updates to the AOP-DB

Because AOP-DB draws data from a number of sources each with release and update schedules independent of one another, AOP-DB updates its records on a quarterly basis. Updates are conducted by scripted routines that ensure data integrity and consistency across all tables, remove duplicate records, and perform sample queries with known expected results to test the coherence and fidelity of the database. A manual review of new data is not conducted.

Biological Category	Data Source	Description
Gene	NCBI Gene	This source supplies all NCBI entrez genes in the gene info table with associated gene information such as name, symbol, location, etc.
	STRING	This source gives protein-protein interaction data for the gene interactions table. Each record from these networks is stored with an entrez1, entrez2, and an interaction score.
Taxonomy & Orthology	NCBI Taxonomy	All taxa available from NCBI, including nomenclature info and divisions. This data is used to fill the species info.

	Homologene	Constructs and stores putative homology groups and contributes and ortho group number, tax id, entrez id to the homology gene table.
	KEGG Orthology	This database of functional orthologs contributes ortho group ids, tax ids, and entrez ids describing an orthologous group to the homology gene table.

	metaPhOrs	This database of phylogeny based orthologs contributes ortho group ids, taxonomy ids, and entrez ids to the homology gene table, describing orthologous groups.
AOP	AOP-wiki	This is a collaborative set of AOPs regularly updated with new details or new Adverse Outcome Pathways. This source contributes to the central AOP info tables and the AOP gene tables, supplying AOP names, key events, descriptions, and information used to map key events to genes.
Chemical	CTD	This source is a manually curated database of chemical information, including many modules. The module of interest for the AOPdb is the chemical gene interactions module, which contributes chemical names and ids to chemical info, as well as the chemical gene interactions with contextual information to the chemical gene table.
	AOP-wiki	In addition to being the source of AOPs for the AOPdb, this source also adds chemical stressors related to the MIE of each AOP. This data contributes chemical names, as well as DTXIDs, casrns, or other chemical ID when available.
	ToxCast	This is a collection of high-throughput screening assays for chemicals that contributes assay identification

		information and assay context information as well as gene target information in the form of entrez ids.
Pathway	KEGG Pathways	This source is a collection of biological molecular interaction pathways that supplies entrez ids and pathway names and ids, linking gene components to the pathways in which they are involved.
	Reactome	This curated and peer-reviewed source of molecular pathways supplies entrez ids and their linked pathways to the pathway gene table of the AOPdb.
	ConcensusPathDB	This source brings together pathway and interaction data from 32 public resources and supplies entrez ids and pathway ids that link genes to biological pathways for the pathway gene table.
Disease	DisGeNET	This database compiles different data, both curated and inferred from models, and supplies multiple downloadable tables relating genes and variants to the diseases in the database. The AOPdb uses DisGeNET's gene-disease association table, adding all fields to the disease-gene table. These include disease name and id, entrez id, and a score for the association based on its sources.
Ontology	NCBI Gene	In addition to being a source of taxonomy info and gene info, NCBI Gene supplies gene ontology information. This supplies gene ontology terms and any related entrez ids to the GO gene table.
Tissues	HumanBase	This API is used to pull tissue specific gene interaction network from HumanBase. The data imported into the tissue networks table in the AOPdb include entrez1 and entrez2 fields to construct edges, as well as a probability

		score indicating the strength of the modeled gene interaction.
Haplotypes	1000 Genomes	This is a collection of variant data for individuals from a multitude of populations. This source contributes snp frequencies for each function snp in the snps table for each of 5 1000 Genomes major populations.
	Ensemble	This API, allowing access to ensembl's gene and variant information, is used to get genotype data for each individual sample from the 1000 Genomes project. These data are used to construct haplotypes for each AOP and find differences in haplotype frequencies within and between populations.
	GWAS Catalog	This is a source used to filter SNPs into snps of interest for variant analysis in different populations. Functional snps are specifically targeted. It, along with GTEx, supplies refsnp ids for these variants as well as contextual information.
	GTEx	This is a source used to filter SNPs into snps of interest for variant analysis in different populations. Functional snps are specifically targeted. It, along with GWAS Catalog, supplies refsnp ids for these variants as well as contextual information.

4 How Content is Selected for Inclusion in the AOP-DB

AOP-DB strives to be an aggregator and curator of toxicologically-relevant data from around the globe. To that end, publicly available data sources are selected using an informal and expansive criterion that emphasizes robust quality assurance measures, programmable access through an API or FTP, and regular updates and maintenance where the data is not static. While sources from the

United States were not preferentially included, sources maintained by an agency of the US federal government were assumed to have implemented rigorous QA measures and generally were selected.

5 References

- Public Law 114–182. (2016). Frank R. Lautenberg Chemical Safety for the 21st Century Act. t. Congress. Delrue, N., M. Sachana, Y. Sakuratani, A. Gourmelon, E. Leinala and R. Diderich (2016). "The adverse outcome pathway concept: A basis for developing regulatory decision-making tools." *Altern Lab Anim* 44(5): 417-429.
- Patlewicz, G., A. P. Worth and N. Ball (2016). "Validation of Computational Methods." *Adv Exp Med Biol* 856: 165-187.
- Sakuratani, Y., M. Horie and E. Leinala (2018). "Integrated Approaches to Testing and Assessment: OECD Activities on the Development and Use of Adverse Outcome Pathways and Case Studies." *Basic Clin Pharmacol Toxicol* 123 Suppl 5: 20-28.
- Grulke, C.M., *et al.* EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research. *Comput Toxicol* 2019;12.
- Mortensen, H.M., *et al.* Leveraging human genetic and adverse outcome pathway (AOP) data to inform susceptibility in human health risk assessment. *Mamm Genome* 2018;29(1-2):190-204.
- Mortensen, H.M., *et al.* The 2021 update of the EPA's adverse outcome pathway database. *Sci Data* 2021;8(1):169.
- Pittman, M.E., *et al.* AOP-DB: A database resource for the exploration of Adverse Outcome Pathways through integrated association networks. *Toxicol Appl Pharmacol* 2018;343:71-83.
- Richard, A.M. and Williams, C.R. Distributed structure-searchable toxicity (DSSTox) public database network: a proposal. *Mutat Res* 2002;499(1):27-52.
- Williams, A.J., *et al.* The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J Cheminform* 2017;9(1):61.