

---

---

# Tox21 Enricher User's Manual

(Last Updated 11/02/2021)

<http://hurlab.med.und.edu/Tox21Enricher>

*Copyright 2021 University of North Dakota  
All rights reserved.*

Junguk Hur, Ph.D.  
Department of Biomedical Sciences  
University of North Dakota, School of Medicine and Health Sciences  
Grand Forks, North Dakota 58202, USA  
Email: [InformaticsTools@gmail.com](mailto:InformaticsTools@gmail.com)  
<http://hurlab.med.und.edu>

---

---

---

---

# Table of Contents

<b>I. Introduction to Tox21 Enricher .....</b>	<b>3</b>
<b>II. Landing Page .....</b>	<b>4</b>
II.1. Sidebar & Settings .....	5
II.1.a. Settings Menu .....	5
II.2. Selecting the Enrichment Mode and Input Type .....	6
II.2.a. Enrich From User Provided CASRN List .....	7
II.2.b. Enrich From Chemicals With Shared Substructures .....	7
II.2.c. Enrich From Chemicals With Structural Similarity .....	9
II.2.d. View Annotations for Tox21 Chemicals .....	10
II.3. Selecting Annotation Categories .....	11
II.4. Selecting Enrichment Cutoff Value .....	12
II.5. Performing Enrichment .....	12
<b>III. Waiting Page .....</b>	<b>13</b>
<b>IV. Results Page .....</b>	<b>15</b>
IV.1. Enrichment Results .....	15
IV.1.a. Result Files .....	15
IV.1.b. Result Chemicals .....	16
IV.1.c. Download Results Button .....	18
IV.1.d. Adjust Network Node Cutoff & Perform Re-enrichment .....	18
IV.1.e. Chart and Cluster Heatmaps .....	19
IV.1.f. Chart and Cluster Networks .....	20
IV.1.g. Significant P-value per Annotation Category Bar Graphs .....	22
IV.1.h. Start Over Button .....	22
IV.2. View Annotation Results .....	23
IV.2.a. Result Files .....	23
IV.2.b. Download Results Button .....	23
<b>V. View Results From Previous Request Page .....</b>	<b>24</b>
<b>VI. Headless Mode .....</b>	<b>26</b>
<b>VII. Common Issues .....</b>	<b>28</b>
VII.1. Runtime Issues .....	28
<b>VIII. Resources .....</b>	<b>29</b>
<b>IX. References .....</b>	<b>30</b>

---

---

---

## I. Introduction to Tox21 Enricher

Humans are exposed to tens of thousands of chemicals that are used in daily life, some at levels that may pose a health risk. There is limited toxicological information for many of these chemicals, which makes risk assessment difficult or impossible. The United States Toxicology Testing in the 21st Century (Tox21) program was established to develop more efficient and human-relevant toxicity assessment methods. The Tox21 program is currently screening a set of over 10,000 chemicals, the Tox21 10K library, using quantitative high-throughput screening (qHTS) of assays that measure effects on toxicity pathways. To date, more than 70 assays have yielded >12 million concentration-response curves by Tox21 researchers. To efficiently apply these data for identifying potential hazardous compounds and for informing further toxicological testing, the United States National Toxicology Program (NTP) has developed several web applications (Tox21 Toolbox: <http://ntp.niehs.nih.gov/tbox/>), including tools for data visualization (Tox21 Curve Browser) and exploration (Tox21 Activity Profiler).

One critical usage of this dataset is to perform chemical-relational analysis based on the patterns of activity across the Tox21 assays and then to use nearest neighbor-based prediction to infer the toxicological properties of untested chemicals via their association with tested chemicals. One approach to inferring the specific properties is to perform chemical annotation enrichment of chemical neighborhoods.

Here, we present Tox21 Enricher, a web-based chemical annotation enrichment tool for Tox21 assay data built using the R Shiny framework. Tox21 Enricher identifies significantly over-represented chemical biological annotations among sets of chemicals (neighborhoods), which facilitates the identification of the toxicological properties and mechanisms in the chemical sets.

---

## II. Landing Page

Upon first using Tox21 Enricher via an internet browser or directly through the client application, you will see the landing page.

Tox21 Enricher

The screenshot displays the Tox21 Enricher landing page. On the left, a sidebar contains a welcome message, a status indicator for the connection to the Tox21 Enricher server, and a 'Total requests serviced by Tox21 Enricher this month' section with a 'Settings' button and a 'View previous enrichment' link. Below this is a 'Select Input Type' dropdown menu currently set to 'User provided CASRN list'. The main content area is titled 'Enrich from user-provided CASRN list'. It features a 'Select Chemical/Biological Annotation Categories' section with a 'Deselect all' button and a 'Note' about selecting no categories. Underneath, there are tabs for 'PubChem Compound Annotations', 'DrugMatrix Annotations', 'DrugBank Annotations', 'CTD Annotations', and 'Other Annotations'. The 'PubChem Compound Annotations' tab is active, showing 'MESH' and 'PHARMACTIONLIST' categories, each with a 'Deselect' button. Below this is a 'Select Enrichment Cutoff' section with a slider ranging from 1 to 30, currently set at 10. The 'Input CASRNs' section includes two buttons: 'CASRNs example single set' and 'CASRNs example multiple sets', followed by a 'Clear input box' button. A 'Note' instructs users to verify chemical identifiers and use the EPRs CompTox Chemicals Dashboard. It also states that 'Add "setname"' should be used before each set if using multiple sets at once. A large text input area is provided for the CASRNs, and a 'Submit' button is at the bottom.

The various utilities and options on the landing page will be discussed at length in this section.

---

## II.1. Sidebar & Settings

The sidebar of the application displays helpful information and various options and settings.

The sidebar contains the following elements and callouts:

- Welcome message:** "Welcome to Tox21 Enricher! Please see this [link](#) for instructions on using this application and the descriptions about the chemical / biological categories. Other resources from the Tox21 toolbox can be viewed [here](#). A sufficiently robust internet connection and JavaScript are required to use all of this application's features." Callout: "A link to download a copy of this manual." (points to the first [link](#))
- Connection status:** "Connection with Tox21 Enricher server successfully established." Callout: "A link to other applications that use the Tox21 dataset." (points to the [here](#) link)
- Request count:** "Total requests serviced by Tox21 Enricher this month: 47" Callout: "A notice that detects if the client application can successfully ping the API." (points to the connection status)
- Settings button:** Callout: "The total number of requests stored in the Tox21 Enricher database that have completed during the current month." (points to the number 47)
- View previous enrichment button:** Callout: "A button to open the **Settings** menu (see section **III.1.a.**)" (points to the Settings button)
- Dark theme toggle:** Callout: "A button to open the **View Results from Previous Request** page (see section **VI.**)" (points to the View previous enrichment button)
- Select Input Type dropdown:** Callout: "A toggle for the client application's dark theme. If your browser is set to display dark themes by default, that setting will take precedent and initialize the application to its dark theme." (points to the Dark theme toggle)
- Input type selection:** Callout: "A menu to select the input type for the current request." (points to the dropdown menu)

### II.1.a. Settings Menu

The settings menu allows you to clear the local cache. Whenever you submit a request from the client application to the API or whenever you download a file from the server through the API, the files are stored in a temporary directory at `<project_root>/tox21enricher/www/tmp/`. For example, if you cloned the code to `/home/user/tox21enricher/`, the temporary directory will be at `/home/user/tox21enricher/tox21enricher/www/tmp/`. Clicking the **Clear local cache** button will prompt you to delete the files present in the temporary directory.

Note that this **will not** delete request record files. Those files are handled separately in the **View previous results** menu. Also, any files deleted in this manner cannot be recovered and will have to be redownloaded from the Tox21 Enricher server in the future.

---

---

## Settings

### Clear Local Cache

 Clear local cache

This will clear the Tox21 Enricher client application's local storage and delete files like enrichment results and the manual. These files will have to be redownloaded in the future. This cannot be undone.

Save & Exit

## Warning

You are about to clear the application's local storage at:

`/home/hurlab/tox21enricher/src/main/r/tox21enricher/www/tmp/`

This action cannot be undone. Continue?

Yes, clear the cache.

Close

## II.2. Selecting the Enrichment Mode and Input Type

The first step in using Tox21 Enricher is selecting the input type. There are three enrichment analysis modes: User-provided CASRN list, Chemicals with shared substructures, and Chemicals with structural similarity. A fourth mode, View annotations for Tox21 chemicals, does not perform enrichment but rather retrieves annotations from the Tox21 Enricher database that correspond to the inputted chemicals.

### Select Input Type

User-provided CASRN list

User-provided CASRN list

Chemicals with shared substructures

Chemicals with structural similarity

View annotations for Tox21 chemicals

---

### II.2.a. Enrich From User Provided CASRN List

This enrichment analysis type accepts CASRN strings as input with one CASRN per line. Multiple sets may be independently enriched at the same time. To enrich multiple sets, separate lines of CASRN strings with **unique** set names of the form “#SetName” with a pound symbol (#) before the name. Only alphanumeric characters (A-Z, a-z and 0-9) may be used in set names. If you choose to provide names for input sets, **all** sets must be assigned a name. Spaces in set names or CASRN strings are ignored, and extra newline characters between lines are also ignored. If no set names are provided, all lines of the input box will be treated as one set and will be assigned the default name of “Set1.”

#### Input CASRNs

CASRNs example single set

CASRNs example multiple sets

Clear input box

**Note:** Please verify you are using the correct chemical identifiers by referencing the [EPA's CompTox Chemicals Dashboard](https://comptox.epa.gov/dashboard).

Add “#SetName” before each set if using multiple sets at once. Set names may only be alphanumeric characters (A-Z, a-z, and 0-9) and spaces are ignored.

```
#Set1
965-90-2
50-50-0
979-32-8
4245-41-4
143-50-0
#Set2
17924-92-4
297-76-7
152-43-2
313-06-4
4956-37-0
112400-86-9
```

Submit

Before performing enrichment from a user provided CASRN list, please cross-reference your CASRN inputs with those in the EPA's CompTox Chemicals Dashboard (<https://comptox.epa.gov/dashboard>) to ensure accuracy during the enrichment analysis. Clicking the **CASRNs example single set** button will fill the input box with one, unnamed set of CASRNs. Clicking the **CASRNs example multiple sets** button will fill the input box with three sets of CASRNs. Clicking the **Clear input box** button will reset the input box to a blank state.

### II.2.b. Enrich From Chemicals With Shared Substructures

This enrichment analysis type accepts either SMILES or InChI strings (one per line). Each string is treated as its own data set and will be enriched independently of any other sets provided. SMILES and InChI strings may be mixed: you may perform enrichment with a SMILES string on one line and an InChI string on another. Clicking the **SMILES/InChI example set** button will fill the input box with three chemicals: two SMILES strings and one InChI string. Clicking the **Clear input box** button will reset the input box to a blank state.

---

## Input SMILE/InChI Strings

SMILES/InChI example set

Draw molecules with JSME

Clear input box

**Note:** Please verify you are using the correct chemical identifiers by referencing the [EPA's CompTox Chemicals Dashboard](#).

Enter partial or complete SMILES or InChI strings, one per line.

```
C1CC=CC=C1  
COC(=O)CC#N  
InChI=1S/C8H11N/c1-9(2)8-6-4-3-5-7-8/h3-7H,1-2H3
```

Submit

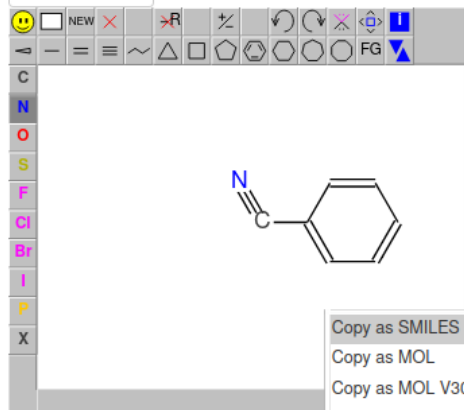
Clicking the **Draw molecules with JSME** button will open the JSME Molecule Editor by Peter Ertl and Bruno Bienfait. Here, you can draw a molecule using the applet's tools. When you are finished drawing, we can right-click in the whitespace of the canvas and select "Copy as SMILES" or "Copy as InChI" and paste the copied text in the input box.

## Input SMILE/InChI Strings

SMILES/InChI example set

Hide JSME

Clear input box



For instructions on using JSME to draw chemical structures, see the [JSME documentation](#).  
B. Bienfait and P. Ertl, *JSME: a free molecule editor*, [Mol. Inform.](#) 2013, 32, 171-175.

**Note:** Please verify you are using the correct chemical identifiers by referencing the [EPA's CompTox Chemicals Dashboard](#).

Enter partial or complete SMILES or InChI strings, one per line.

```
C1CC=CC=C1  
COC(=O)CC#N
```

When you are finished drawing, you can right-click by the drawing and select "Copy as SMILES" or "Copy as InChI" and paste the copied text in the input box.

(2013)

[CompTox Chemicals Dashboard](#).



### II.2.c. Enrich From Chemicals With Structural Similarity

This enrichment analysis type accepts either SMILES or InChI strings (one per line). Each string is treated as its own data set and will be enriched independently of any other sets provided. SMILES and InChI strings may be mixed: you may perform enrichment with a SMILES string on one line and an InChI string on another. Clicking the **SMILES/InChI example set** button will fill the input box with three chemicals: two SMILES strings and one InChI string. Clicking the **Clear input box** button will reset the input box to a blank state. The **Tanimoto similarity threshold** may be specified using the slider controls above the input box to set the strictness of the search. By default, this is set to 50%.

Select Enrichment Cutoff

Select enrichment cutoff

Select Tanimoto similarity threshold (%)

Input SMILE/InChI Strings

SMILES/InChI example set Draw molecules with JSME

Clear input box

Note: Please verify you are using the correct chemical identifiers by referencing the [EPA's CompTox Chemicals Dashboard](#).

Enter partial or complete SMILES or InChI strings, one per line.

```
ClCC1=CC=CC=C1
OCCCCC(=O)CC#N
InChI=1S/C8H11N/c1-9(2/8-6-4-3-5-7-8)/h3-7H,1-2H3
```

Clicking the **Draw molecules with JSME** button will open the JSME Molecule Editor by Peter Ertl and Bruno Bienfait. Here, you can draw a molecule using the applet's tools. When you are finished drawing, we can right-click in the whitespace of the canvas and select "Copy as SMILES" or "Copy as InChI" and paste the copied text in the input box.

Input SMILE/InChI Strings

SMILES/InChI example set Hide JSME

Clear input box

NEW X R Z C O P A FG

C N O S F Cl Br I P X

For instructions on using JSME to draw chemical structures, right-click by the drawing and select "Copy as SMILES" or "Copy as InChI" to copy the structure to the clipboard.

B. Bienfait and P. Ertl, *JSME: a free molecule editor*, 2013, [CompTox Chemicals Dashboard](#).

Note: Please verify you are using the correct chemical identifiers by referencing the [EPA's CompTox Chemicals Dashboard](#).

Enter partial or complete SMILES or InChI strings, one per line.

```
ClCC1=CC=CC=C1
OCCCCC(=O)CC#N
```

Copy as SMILES  
Copy as MOL  
Copy as MOL V3000  
Copy as InChI  
Copy as InChI key  
Search chemical structure (through InChIKey)  
Copy as JME  
Copy as OCL Scalar Vector Graphics  
Copy as raw Scalar Vector Graphics  
Paste MOL or SDF or SMILES

---

### II.2.d. View Annotations for Tox21 Chemicals

This function does not actually perform enrichment analysis; rather, it returns a list of every annotation in the Tox21 database that is associated with each of the supplied chemicals. Like the **User-provided CASRN list** enrichment type, this enrichment analysis type accepts only CASRN strings as input (one CASRN per line). You may find annotations for multiple sets of chemicals at the same time. To use multiple sets, separate lines of CASRN strings with **unique** set names of the form “#SetName” with a pound symbol (#) before the name. Only alphanumeric characters (A-Z, a-z and 0-9) may be used in set names. If you choose to provide names for input sets, **all** sets must be assigned a name. Spaces in set names or CASRN strings are ignored, and extra newline characters between, after, and before lines are also ignored. If no set names are provided, all lines of the input box will be treated as one set and will be assigned the default name of “Set1.”

This mode will not provide any annotations for a CASRN if that chemical is not present in the Tox21 screening library. To view all the chemicals in the Tox21 screening library, click the **chemicals in the Tox21 screening library** link which opens the page [https://comptox.epa.gov/dashboard/chemical\\_lists/TOX21SL](https://comptox.epa.gov/dashboard/chemical_lists/TOX21SL) in a new window.

#### Input CASRNs

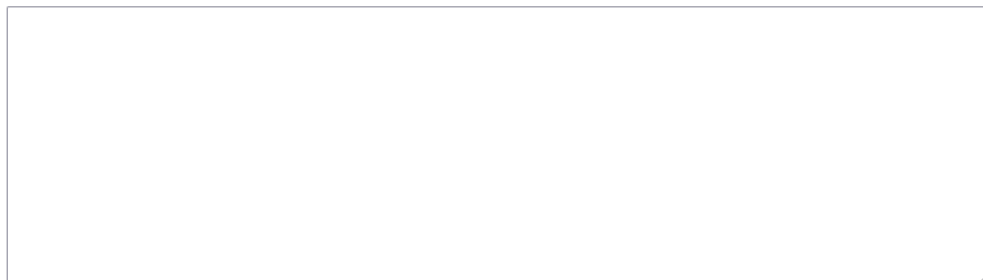
CASRNs example single set

CASRNs example multiple sets

Clear input box

**Note:** Please verify you are using the correct chemical identifiers by referencing the [EPA's CompTox Chemicals Dashboard](https://comptox.epa.gov/dashboard/chemical_lists/TOX21SL).

Enter the CASRNs for [chemicals in the Tox21 screening library](#) (one per line) to view each of their associated annotations in Tox21 Enricher. Add “#SetName” before each set if using multiple sets at once. Set names may only be alphanumeric characters (A-Z, a-z, and 0-9) and spaces are ignored.



Submit

Clicking the **CASRNs example single set** button will fill the input box with one, unnamed set of CASRNs. Clicking the **CASRNs example multiple sets** button will fill the input box with three sets of CASRNs. Clicking the **Clear input box** button will reset the input box to a blank state.

---

## II.3. Selecting Annotation Categories

The top of the landing page displays lists of annotation categories, separated into five category types: **PubChem Compound Annotations**, **DrugMatrix Annotations**, **DrugBank Annotations**, **CTD Annotations**, and **Other Annotations**. By default, the **PubChem Compound Annotations** tab is selected. You can click on the tabs to view each tab's corresponding annotation categories. Each category name has a checkbox to its left that, when checked, means that the application will attempt to find matching annotations from the given category. The application will ignore any potential results from categories left unchecked. By default, all categories are checked except for the CTD\_GOFAT\_BIOPROCESS category: using this category can potentially slow down enrichment by a large amount, so it is not recommended to be used in most scenarios. Each category name has a question mark button to its right that, when hovered over, displays additional information about each category. If no categories are selected, the application defaults to just using all the annotation categories (except for CTD\_GOFAT\_BIOPROCESS).

### Select Chemical/Biological Annotation Categories

Deselect all

Note: Selecting no annotation categories will cause enrichment to just use the default categories.

PubChem Compound Annotations DrugMatrix Annotations DrugBank Annotations CTD Annotations Other Annotations

**DrugBank Annotations**

<input checked="" type="checkbox"/> DRUGBANK_ATC	?
<input checked="" type="checkbox"/> DRUGBANK_ATC_CODE	?
<input checked="" type="checkbox"/> DRUGBANK_CARRIERS	?
<input checked="" type="checkbox"/> DRUGBANK_ENZYMES	?
<input checked="" type="checkbox"/> DRUGBANK_TARGETS	?
<input checked="" type="checkbox"/> DRUGBANK_TRANSPORTERS	?

Deselect all DrugBank annotations

Clicking the **Deselect all/Select all** button will either deselect or select all the annotation categories across all tabs. Note that clicking the **Select all** button will select the CTD\_GOFAT\_BIOPROCESS category too.

Each tab also has a corresponding button to deselect or select all the annotation categories within the tab. An example can be seen in the image above: clicking the **Deselect all DrugBank annotations** button will deselect only the annotations in the **DrugBank Annotations** tab. Note that clicking the **CTD** tab's corresponding **Select all CTD annotations** button will select the CTD\_GOFAT\_BIOPROCESS category too.

---

---

## II.4. Selecting Enrichment Cutoff Value

Under the tabs for selecting the appropriate annotation categories, the landing page displays a slider input for specifying the enrichment cutoff value. This value determines the maximum number of enrichment results displayed per data set submitted. By extension, this will also determine the maximum number of nodes displayed during network generation. By default, this is set to 10 and can be increased to a maximum of 50 or decreased to a minimum of 1.

### Select Enrichment Cutoff

Select enrichment cutoff



## II.5. Performing Enrichment

After input has been entered, it is time to begin enrichment. This can be done by clicking on the **Submit** button under the input box. If there are no problems validating your input, you will be sent to the waiting page.

Validation errors may occur if you input a string that is not a properly formatted CASRN, SMILES, or InChI identifier, if you forget to name a set, or if you try to submit a blank input. If a validation error occurs, Tox21 Enricher will display a warning and the request will not proceed.

### Input CASRNs

CASRNs example single set

CASRNs example multiple sets

Clear input box

**Note:** Please verify you are using the correct chemical identifier

Add "#SetName" before each set if using multiple sets at a time

965-90-2  
50-50-0  
979-32-8  
4245-41-4  
143-50-0  
17924-92-4  
297-76-7  
152-43-2  
313-06-4  
4956-37-0  
112400-86-9

 Submit

### III. Waiting Page

After submitting your request, you will be taken to the waiting page.

Your request has been submitted and placed in the queue. After your request is processed and completed, the results can be accessed by clicking the "Results" button.

[Results](#) [Refresh queue](#) [Copy UUID to clipboard](#) [Cancel enrichment](#)

Your Request

Show  entries Search:

Status	Request Mode	Request UUID	Selected Annotations	Node Cutoff	User Input
BPAanalogs: Waiting in queue position: 1.	Enrich from user-provided CASRN list	4656bd9e-21a7-4cfa-a0f8-39ac0a3e3dcb	PHARMACTIONLIST, MESH, KNOWN_TOXICITY, MECH_LEVEL_2, MECHANISM, STRUCTURE_ACTIVITY, TA_LEVEL_2, THERAPEUTIC_CLASS, TISSUE_TOXICITY, ACTIVITY_CLASS, ADVERSE_EFFECT, INDICATION, MECH_LEVEL_1, MECH_LEVEL_3, MODE_CLASS, PRODUCT_CLASS, TA_LEVEL_1, TA_LEVEL_3, DRUGBANK_ATC, DRUGBANK_CARRIERS, DRUGBANK_ENZYMES, DRUGBANK_TARGETS, DRUGBANK_ATC_CODE, DRUGBANK_TRANSPORTERS, CTD_CHEMICALS_DISEASES, CTD_CHEMICALS_GENES, CTD_CHEMICALS_GOENRICH_CELLCOMP, CTD_CHEMICALS_GOENRICH_MOLFUNCT, CTD_CHEMICALS_PATHWAYS, CTD_GOSLIM_BIOPROCESS, LEADSCOPE_TOXICITY, MULTICASE_TOX_PREDICTION, TOXCAST_ACTIVE, TOXREFDB, HTS_ACTIVE, HTS_STRONGACTIVE, TOXINS_TARGETS, TOXPRINT_STRUCTURE,	10	#BPAanalogs 2081-08-5 2467-02-9 1478-61-1 41481-66-7 5613-46-7 57-63-6 620-92-8 77-40-7 79-94-7 79-95-8 79-97-0 80-05-7 80-09-1 843-55-0 94-18-8 #Flameretardants 115-86-6 115-96-8 1241-94-7 1330-78-5 13674-87-8 29761-21-5 5436-43-1 56803-37-3 68937-41-7 78-30-8 79-94-7 #PAH 120-12-7 129-00-0 191-24-2 206-44-0 218-01-9 50-32-8 53-70-3 56-55-3 83-32-9 85-01-8

Showing 1 to 1 of 1 entries Previous **1** Next

There are four buttons at the top of the waiting page that perform different tasks when clicked:

- **Results** – This will take you to the request's results page if the request has finished. If the request has not yet finished, the waiting page will display a warning.
- **Refresh queue** – This will refresh the **Your Request** table and update the **Status** column.
- **Copy UUID to clipboard** – This will copy your request's UUID to the clipboard so it may be pasted elsewhere.
- **Cancel enrichment** – This will open a prompt to cancel the request. Cancelled requests cannot be recovered. The request's temporary file will be deleted from the client filesystem and records of the request will be deleted from the Tox21 Enricher database.

The **Your Request** table contains some helpful information related to the parameters of your request:

## Your Request

Show  entries Search:

Status	Request Mode	Request UUID	Selected Annotations	Node Cutoff	User Input
BPAanalogs: (Step 3/4): Clustering (Step 1/4) - calculating kappa score.	Enrich from user-provided CASRN list	4656bd9e-21a7-4cfa-a0f8-39ac0a3e3dcb	PHARMACTIONLIST, MESH, KNOWN_TOXICITY, MECH_LEVEL_2, MECHANISM, STRUCTURE_ACTIVITY, TA_LEVEL_2, THERAPEUTIC_CLASS, TISSUE_TOXICITY, ACTIVITY_CLASS, ADVERSE_EFFECT, INDICATION, MECH_LEVEL_1, MECH_LEVEL_3, MODE_CLASS, PRODUCT_CLASS, TA_LEVEL_1, TA_LEVEL_3, DRUGBANK_ATC, DRUGBANK_CARRIERS, DRUGBANK_ENZYMES, DRUGBANK_TARGETS, DRUGBANK_ATC_CODE, DRUGBANK_TRANSPORTERS, CTD_CHEMICALS_DISEASES, CTD_CHEMICALS_GENES, CTD_CHEMICALS_GOENRICH_CELLCOMP, CTD_CHEMICALS_GOENRICH_MOLFUNCT, CTD_CHEMICALS_PATHWAYS, CTD_GOSLIM_BIOPROCESS, LEADSCOPE_TOXICITY, MULTICASE_TOX_PREDICTION, TOXCAST_ACTIVE, TOXREFDB, HTS_ACTIVE, HTS_STRONGACTIVE, TOXINS_TARGETS, TOXPRINT_STRUCTURE,	10	#BPAanalogs 2081-08-5 2467-02-9 1478-61-1 41481-66-7 5613-46-7 57-63-6 620-92-8 77-40-7 79-94-7 79-95-8 79-97-0 80-05-7 80-09-1 843-55-0 94-18-8 #Flameretardants 115-86-6 115-96-8 1241-94-7 1330-78-5 13674-87-8 29761-21-5 5436-43-1 56803-37-3 68937-41-7 78-30-8 79-94-7 #PAH 120-12-7 129-00-0 191-24-2 206-44-0 218-01-9 50-32-8 53-70-3 56-55-3 83-32-9 85-01-8

Showing 1 to 1 of 1 entries Previous **1** Next

Each input set of the request's respective status and progress to completion. Changes to **"Complete!"** when finished.

A unique character string assigned to each request.

The request type.

The list of all the checked annotation categories being used by this request.

The node cutoff value specified by the select input on the landing page.

The contents of the input box (or only the selected chemicals if re-enriching data).



Clicking any of the result file links will download a copy of the corresponding file from the Tox21 Enricher server to the client application's temporary cache so that it may be viewed at any time in the client application.

The result files are as follows:



- **<set\_name>\_\_ChartSimple.txt** – A list of the top 10 most significant annotations for each annotation class (.txt format).
- **<set\_name>\_\_ChartSimple.xlsx** – A list of the top 10 most significant annotations for each annotation class (.xlsx format).
- **<set\_name>\_\_Chart.txt** – A list of all significant annotations (.txt format).
- **<set\_name>\_\_Chart.xlsx** – A list of all significant annotations (.xlsx format).
- **<set\_name>\_\_Cluster.txt** – A list of significant terms in which functionally similar annotations are grouped together to remove redundancy. This is performed with respect to the whole annotation set rather than to individual annotation classes (.txt format).
- **<set\_name>\_\_Cluster.xlsx** – A list of significant terms in which functionally similar annotations are grouped together to remove redundancy. This is performed with respect to the whole annotation set rather than to individual annotation classes (.txt format).
- **<set\_name>\_\_Matrix.txt** – A plaintext representation of the heatmap (.txt format).
- **<set\_name> Input** – A plaintext list of the input chemicals for this set (.txt format).

#### IV.1.b. Result Chemicals

The **Result Chemicals** table is only visible when enriching from chemicals with shared substructures or structural similarity. This table displays information about all of the related chemicals to the original SMILES/InChI input strings.

Result Chemicals

Copy CSV Excel PDF Print Column visibility

Select	Chemical Structure	DSSTox Substance ID	CASRN	IUPAC Name	SMILES	InChI	InChI Key	Molecular Formula	Molecular Weight	Similarity
<input checked="" type="checkbox"/>		DTXSID0020153	100-44-7	chloromethylbenzene	ClCC1=CC=CC=C1	InChI=1S/C7H7Cl/c8-6-7-4-2-1-3-5-7/h1-5H,6H2	KCXMKQUNVWSEMD-UHFFFAOYSA-N	C7H7Cl	126.5835	1
<input checked="" type="checkbox"/>		DTXSID5021839	100-46-9	phenylmethanamine	NCC1=CC=CC=C1	InChI=1S/C7H9N/c8-6-7-4-2-1-3-5-7/h1-5H,6,8H2	WGQKYBSKWADBV-UHFFFAOYSA-N	C7H9N	107.1531	0.56

Additionally, if you performed enrichment from chemicals with structural similarity, an additional column is shown displaying a Tanimoto value describing the similarity of each result chemical to the original input chemical. Clicking the chemical structure image for a given chemical will open a modal that displays a larger image of the chemical structure, additional information about the chemical, and links to view the chemical's entries on the EPA's CompTox Chemicals Dashboard website and the PubChem database.



100-44-7



DTXSID	DTXSID0020153
DTXRID	75405
Chemical Name	Benzyl chloride
IUPAC Name	chloromethylbenzene
CASRN	100-44-7
SMILES	ClCC1=CC=CC=C1
InChi	InChi=1S/C7H7Cl/c8-6-7-4-2-1-3-5-7/h1-5H,6H2
InChi Key	KCXMKQUNVWSEMD-UHFFFAOYSA-N
Molecular Formula	C7H7Cl
Molecular Weight	126.5835
<a href="#">View at EPA</a>	<a href="#">View at PubChem</a>

Close

The menu buttons above the table perform additional functions:

- **Copy** – Copies the entire table to the clipboard.
- **CSV** – Allows you to save a copy of the table as a .csv file.
- **Excel** – Allows you to save a copy of the table as an Excel spreadsheet.
- **PDF** – Allows you to save a copy of the table as a .pdf file.
- **Print** – Allows you to print a copy of the table.
- **Column visibility** – Allows you to selectively show or hide certain columns of the table.

Each chemical has an associated checkbox that marks the chemical for re-enrichment. By default, all chemicals' checkboxes are checked.

---

If an originally submitted SMILES/InChI string contains a certain reactive structure, Tox21 Enricher will check to see if each of its related chemicals also contain that structure. If a related chemical does not contain the structure, an additional **Reactive Structure Warning** column is added to the **Result Chemicals** table displaying a picture of the suspect structure. A warning will also be displayed if the original input string does not contain a reactive structure, but one of its related chemicals does. Currently, Tox21 Enricher displays warnings if the following structural groups are detected in either (but not both) the input chemical and one of its related chemicals:


- **Nitrile group (cyanide)**
- **Aldehyde**
- **Epoxide**
- **Isocyanate**

#### ↑ **Reactive Structure Warning** ↓



#### IV.1.c. Download Results Button

Clicking the **Download results** button will prompt you to download a .zip archive of all of the results files for all the sets for a given request.



A button with a download icon and the text "Download results".

#### IV.1.d. Adjust Network Node Cutoff & Perform Re-enrichment

Below the result files and chemicals is a series of controls for rerunning enrichment on the result set. First is a slider control for adjusting the node cutoff value so that you may change the number of nodes without having to completely perform enrichment again from scratch.

#### Adjust Network Node Cutoff

##### Re-enrichment Cutoff



A button with a circular arrow icon and the text "Update network".

Under this slider input, there is a button to either update the network (if enriching from a user-provided CASRN list) or perform enrichment again with only the selected result chemicals (if enriching from chemicals with structural similarity or shared substructures). Additionally, if enriching from chemicals with structural similarity or shared substructures, more buttons are present that interact with the **Result Chemicals** table:

---



---

#### IV.1.f. Chart and Cluster Networks

Clicking on the **Chart Network** tab next to the **Chart Heatmap** tab will open the chart network display. Similarly, clicking on the **Cluster Network** tab will open the cluster network display. Both network displays have identical layouts. The nodes correspond to significantly enriched annotations and the edges indicate that there is significant overlap between the two annotations in terms of chemical contents. The edge color gradient indicates the degree of overlap based on a Jaccard index. Note that the network may take some time to load depending on the number of nodes represented.

On the left side, there is a number input for changing the q-value. This is the cutoff value for the statistically significant overlap between pairs of nodes. There are also checkboxes for changing which input sets and/or annotation classes should be represented in the network. Under the **Other Options** section, there is the option to enable network physics, which will make the nodes physically react with each other. There is also the option to toggle curve smoothing for the network edges. This is enabled by default, but disabling it may improve performance. Clicking the **Update network** button will regenerate the network, reflecting any changes made to selected input sets, physics, or the q-value.

In the middle of the network display, there are two drop-down menus. The top menu allows you to select certain annotations in the network without having to click on the nodes in the network. The bottom menu allows you to hide all annotations not of the selected category, making the network easier to navigate and read. Below these menus is a legend that shows each category's unique color.

You may interact with the network by clicking and dragging nodes around the canvas. The mouse wheel can be used to zoom in and out of the network view as well. Additionally, the green buttons overlaid on the network may also be clicked to manipulate the network view. Under the network, there is the **Export as png** button which allows you to save an image of the network as a .png file.

Under the **More Information for Selected Annotation** section, you will see a prompt to select any node in the network. Doing this will display a link to an external website or database with more details about the annotation in the selected node.

Under the **Overlapping Chemicals** section, you will see a prompt to select any edge in the network. Doing this will generate a Venn diagram under the network detailing the number of chemicals shared and unique to the annotations displayed in the connecting nodes of the selected edge. Clicking an edge will also generate buttons that will allow you to view a list of all the chemicals (as CASRNs) associated with either of the two given annotations. A third button will also be generated that will allow you to view a list of all the chemicals (as CASRNs) that are associated with both given annotations. Underneath these buttons are links that allow you to save the Venn diagram as a .png image or .pdf document when clicked.

---

## Chart Full Heatmap

Chart Heatmap

Chart Network

### Edge Selection Criteria

Select by id

#### Q-value

0.05

Select by group

#### Selected Input Sets

- ☒ BPAanalogs
- ☒ Flameretardants
- ☒ PAH

#### Selected Annotation Classes

- ☒ TOXPRINT\_STRUCTURE
- ☒ HTS\_ACTIVE
- ☒ TOXCAST\_ACTIVE
- ☒ HTS\_STRONGACTIVE
- ☒ MESH
- ☒ LEADSCOPE\_TOXICITY
- ☒ MULTICASE\_TOX\_PREDICTION
- ☒ TOXINS\_TARGETS
- ☒ CTD\_CHEMICALS\_GENES

#### Other Options

- ☐ Enable physics?
- ☒ Smooth curve for edges?

Update network

#### More Information for Selected Annotation

Click on any **node** to view additional information for the annotation in the selected node.

#### Overlapping Chemicals

Click on any **edge** to view a Venn diagram of the chemicals associated with the annotations in its two nodes.

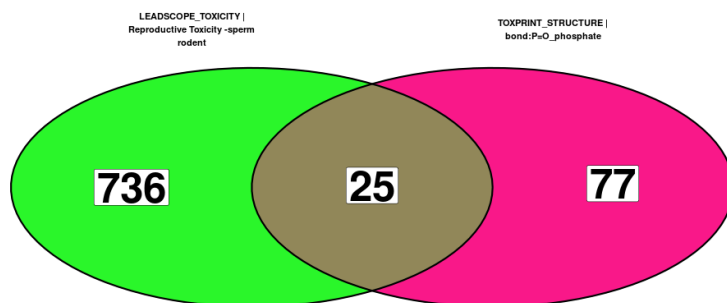
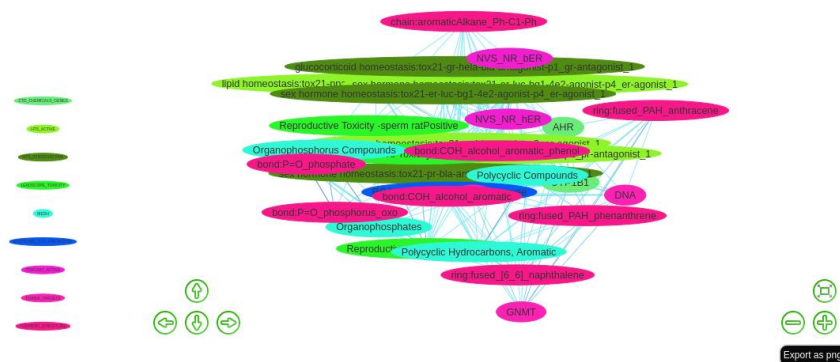
View chemicals for  
LEADSCOPE\_TOXICITY |  
Reproductive Toxicity -sperm rodent

View chemicals for  
TOXPRINT\_STRUCTURE |  
bond:P=O\_phosphate

View shared chemicals

Download plot as .png

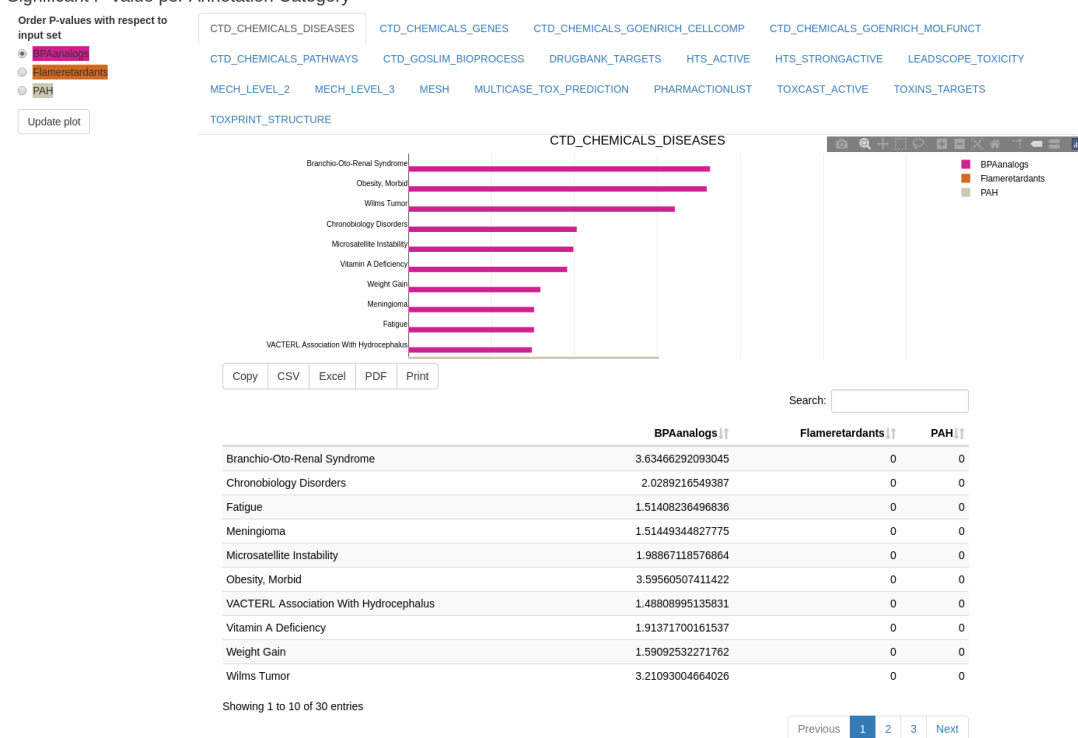
Download plot as .pdf



### IV.1.g. Significant P-value per Annotation Category Bar Graphs

At the bottom of the results page, you can find a series of bar graphs that display all the significant annotations across all result sets and each set's inverse p-value for each annotation, if applicable. Separate bar graphs are created for each annotation category and are put into each category's respective tab. A color legend appears on the righthand side of each bar graph if multiple sets are represented. Clicking on a set name in this legend will hide its trace from the graph, while double-clicking a set name will hide all other sets' traces. To the left of the bar graphs is a menu that allows you to select one set to order values by. Clicking the **Update plot** button will update all the bar graphs to be ordered with respect to the selected set. Trace colors are unique for each set and are randomly selected from a predefined list each time enrichment is performed. Bar graphs are always ordered with the most significant value for the selected set at the top and the least significant value for the selected set at the bottom. The data in each bar graph are also represented in a table below each bar graph. Each table may be saved in different formats by using the buttons above each table.

Significant P-value per Annotation Category



### IV.1.h. Start Over Button

After enrichment is successfully performed, the **Start over** button is added to the bottom of the sidebar. Clicking this button will take you back to the landing page. This button is functionally identical to refreshing the application from within an internet browser.

 Start over

---

## IV.2. View Annotation Results

### IV.2.a. Result Files

The top of the results page displays relevant files generated when fetching annotations. If multiple sets were submitted, the files corresponding to each set are organized into separate tabs.

Each submitted CASRN will have its own text file with a name of the form **<set\_name>\_\_<casrn>.txt**. This file will contain all of the annotations in the Tox21 Enricher database that are associated with the submitted chemical. The annotations in this file are organized so that each line contains one complete annotation: the category the annotation belongs to is on the left and the annotation's name is

on the right, and they are separated by a tab. If Tox21 Enricher could not find any associated annotations for a submitted CASRN, the CASRN will be ignored and a corresponding result file will not be generated. The annotation-fetching process will only search through the annotation categories that were checked on the landing page.

An additional text file of the name **<set\_name>\_\_FullMatrix.txt** will be generated for each submitted set of CASRNs. This file contains a plaintext representation of a matrix of all the fetched annotations in the set and if they are associated with each of the submitted CASRNs.

### IV.2.b. Download Results Button

Clicking the **Download results** button will prompt you to download a .zip archive of all of the results files for all the sets for a given request.

## Fetches Annotations

Set1

Set2

Set1\_\_50-50-0.txt

Set1\_\_965-90-2.txt

Set1\_\_979-32-8.txt

Set1\_\_FullMatrix.txt



Download results

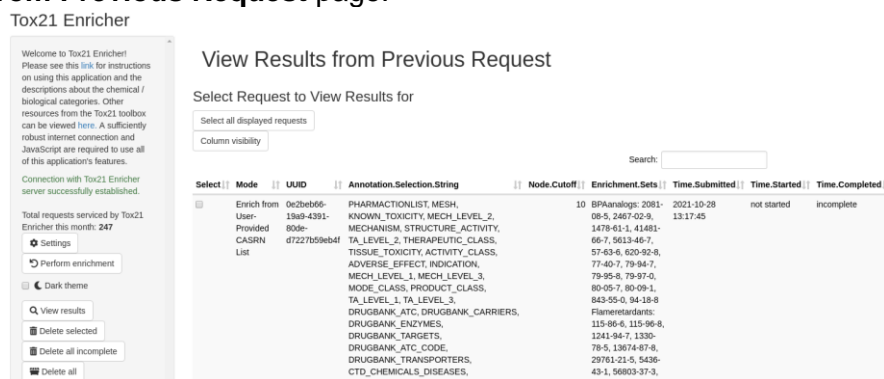


Download results

---

## V. View Results From Previous Request Page

Clicking the **View previous results** button on the sidebar will take you to the **View Results from Previous Request** page.



View Results from Previous Request

Select Request to View Results for

Select all displayed requests

Column visibility

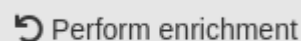
Search:

Select	Mode	UUID	Annotation	Selection.String	Node.Cutoff	Enrichment.Sets	Time.Submitted	Time.Started	Time.Completed
<input type="checkbox"/>	Enrich from User-Provided CASRN List	0c2be6b6-19a0-4391-808e-47227059eb4f	PHARMACICTIONLIST, MESH, KNOWN_TOXICITY, MECH_LEVEL_2, MECHANISM, STRUCTURE_ACTIVITY, TA_LEVEL_2, THERAPEUTIC_CLASS, TISSUE_TOXICITY, ACTIVITY_CLASS, ADVERSE_EFFECT, INDICATION, MECH_LEVEL_1, MECH_LEVEL_3, MOOE_CLASS, PRODUCT_CLASS, TA_LEVEL_1, TA_LEVEL_3, DRUGBANK_ATC, DRUGBANK_CARRIERS, DRUGBANK_ENZYMES, DRUGBANK_TARGETS, DRUGBANK_ATC_CODE, DRUGBANK_TRANSPORTERS, CTD_CHEMICALS_DISEASES,		10	BPhanalog: 2081-08-5, 2467-02-9, 1478-61-1, 41481-66-7, 9613-46-7, 57-63-6, 620-92-8, 77-40-7, 79-94-7, 79-95-8, 79-97-0, 89-65-7, 80-09-1, 843-55-0, 94-18-8, Flamentardans: 115-86-6, 115-96-8, 1241-94-7, 1330-78-5, 13674-87-8, 29761-21-5, 5436-43-1, 56803-37-3,	2021-10-28 13:17:45	not started	incomplete

There are four buttons in the sidebar menu that perform different functions:

- **View results** – This button will take you the results page of whatever previous request you have selected in the table below.
- **Delete selected** – This button will prompt you to delete the records of all the checked requests in the table below.
- **Delete all incomplete** – This button will prompt you to delete the records of all incomplete requests in the table below.
- **Delete all** – This button will prompt you to delete the records of all previous requests.

There are also two buttons above the table. The **Column visibility** button allows you to selectively show or hide certain columns of the table. The **Deselect all displayed requests** button will (de)select the checkboxes for all entries currently displayed in the table. You may combine this with the **Search** function to only select certain requests. Whenever a request is submitted to the Tox21 Enricher API, whether it be to perform enrichment or view annotations, a record file detailing the details and parameters of the request is created in a temporary directory at `<project_root>/tox21enricher/www/tmp/transaction/`. For example, if you cloned the code to `/home/user/tox21enricher/`, the temporary directory will be at `/home/user/tox21enricher/tox21enricher/www/tmp/transaction/`. The records that are displayed in the table on the **View Results from Previous Request** page are pulled from the files in this directory. When a record is deleted, the corresponding record file in the temporary directory is deleted. Note that record files deleted in this manner cannot be recovered. Additionally, the files in the temporary directory's transaction folder are not deleted when clearing the local cache from the **Settings** menu. By default, none of the checkboxes are checked. Selecting more than one checkbox and clicking the **View results** button will present a warning, as you may not load the results for multiple requests at the same time. On this page, the **View previous results** button in the sidebar will change to say **Perform enrichment**. Clicking this button will take you back to the landing page.





The columns of the table are as follows:

Select	Mode	UUID	Annotation.Selection.String	Node.Cutoff
<input type="checkbox"/>	Enrich from Chemicals with Structural Similarity	1b074a7c-e0e0-4000-8369-a50b062ed640	MESH, PHARMACTIONLIST, DRUGBANK_ATC_CODE, DRUGBANK_CARRIERS, DRUGBANK_ENZYMES, DRUGBANK_TARGETS, CTD_CHEMICALS_DISEASES,	10

The checkbox to select this request for viewing or deletion.

The request type.

A unique character string assigned to each request.

The list of all the checked annotation categories being used by this request.

The node cutoff value specified by the select input on the landing page.

Enrichment.Sets	Time.Submitted	Time.Started	Time.Completed
Set1: 100-44-7, 100-46-9, 103-29-7, 101-81-5, 620-20-2, 104-82-5, 612-12-4	2021-08-26 10:04:18	2021-08-26 10:04:22	2021-08-26 10:04:31

The inputted CASRNs with their corresponding sets. If the original input was a SMILES or InChI string, the retrieved CASRNs are displayed here instead.

A timestamp displaying the date and time the request was submitted to the Tox21 Enricher queue. This is just when the request was created, not necessarily when it was serviced.

A timestamp displaying the date and time the request was accepted by the Tox21 Enricher queue. If the request has not yet been accepted, this will say **not started**.

A timestamp displaying the date and time the request completed in the Tox21 Enricher queue. If the request has not yet completed, this will say **incomplete**.

---

## VI. Headless Mode

A user may make direct calls to the Tox21 Enricher API to perform enrichment, fetch annotations, and download results without using the client application's graphical user interface (GUI). The recommended tool for accessing the API from the command line is **cURL**. To make a request, use the cURL command:

```
curl --header "Content-Type: application/json" --request POST -data
'{"mode": "<mode>", "input": "<input_string>",
"annotations": "<annotation_class_list>", "cutoff": <cutoff>,
"tanimoto": <tanimoto>}' http://<server_address>:<port_api_is_running_on>/submit
```

Where:

- **<mode>** is the type of request (casrn, substructure, similarity, or annotation)
  - **<input\_string>** is a newline (\n)-separated string of input chemical identifiers. If inputting multiple sets of CASRNs, also include set names with a pound symbol (#) in front of each name.
    - Make sure you are submitting the correct identifier for the type of request you are making (i.e., CASRNs for “casrn” and “annotation” and SMILES or InChI strings for “substructure” and “similarity”). Also, if submitting CASRNs, make sure you are using the correct chemical identifiers as included in the EPA's CompTox Chemicals Dashboard: <https://comptox.epa.gov/dashboard>.
    - For example, for the mode “casrn,” you could input something like:  
`#Set1\n965-90-2\n50-50-0\n#Set2\n4245-41-4\n143-50-0\n17924-92-4`
    - For example, for the mode “similarity,” you could input something like:  
`C1CC1=CC=CC=C1\nN#CSCC1=CC=CC=C1`
  - **<annotation\_class\_list>** is a comma-separated string of annotation class names. This denotes which annotation classes should be included when matching annotations to the input.
    - For example, you could input something like:  
`MESH, PHARMACTIONLIST, ACTIVITY_CLASS, ADVERSE_EFFECT, INDICATION`
    - You may run the command  
`curl http://<server_address>:<port>/annotationList` to view a list of all available annotation classes.
    - If not supplied, this value will default to all available annotation classes except for **CTD\_GOFAT\_BIOPROCESS**.
  - **<cutoff>** is the integer value that determines the maximum number of annotations per annotation class to include in things like the network and ChartSimple files.
    - This must be an integer between 1 and 50 inclusive.
    - If not supplied, this will default to 10.
  - **<tanimoto>** is the float value that determines the Tanimoto threshold. If a chemical's similarity score is greater than this value, it will be excluded from enrichment analysis.
    - This is only used with the “similarity” mode and will be ignored if supplied with another mode.
    - If not supplied, this will default to 0.5.
-

---

---

Once a request is made, it will be placed in the queue like any other request made from the client application. If the request was successfully made, cURL will return the newly assigned UUID of the request in the terminal.

To check the status of the request, you may run the cURL command:

`curl http://<server>:<port>/completed?<uuid>`. This function will either return a 0 if the request has not completed or a 1 if the request has completed.

To download the results once the request is completed, you may run the cURL command:

`curl http://<server>:<port>/download?<uuid> > <name_of_downloaded_file>.zip`.

This will download a zipped folder containing all the request's results to the file `<name_of_downloaded_file>.zip` in the current working directory. This request will fail if the request has not yet completed. You may also use this command to download the results of any request, even those that were initially submitted using the client application's GUI.

Examples for using Tox21 Enricher's headless mode can be found in the project's GitHub repository (<https://github.com/hurlab/tox21enricher>) at `<project_root>/demos/`.

---

## VII. Common Issues

Both the client and server code for Tox21 Enricher are still in continued development, so there are likely issues that may arise when trying to either perform certain tasks or set up the application. Known bugs in the code and planned features are documented on the project's GitHub repository at: <https://github.com/hurlab/tox21enricher/issues>. Other known issues not necessarily related to the code itself are documented in this section of the manual.

### VII.1. Runtime Issues

*This section intentionally left blank and will be updated in the future.*

---

---

## VIII. Resources

Below are a few resources that users may find helpful:

Name	Link	Description
CompTox Chemicals Dashboard	<a href="https://comptox.epa.gov/dashboard">https://comptox.epa.gov/dashboard</a>	Reference for chemical identifiers to be used in Tox21 Enricher.
Hurlab Server	<a href="http://hurlab.med.und.edu/tox21enricher/">http://hurlab.med.und.edu/tox21enricher/</a>	Location of Tox21 Enricher on the Hurlab web server.
RDKit	<a href="https://www.rdkit.org/docs/index.html">https://www.rdkit.org/docs/index.html</a>	Documentation for the RDKit cheminformatics tool.
Tox21 Enricher GitHub Repository	<a href="https://github.com/hurlab/tox21enricher">https://github.com/hurlab/tox21enricher</a>	Tox21 Enricher public GitHub repository.
Toxicology in the 21 <sup>st</sup> Century	<a href="https://tox21.gov/">https://tox21.gov/</a>	Information regarding the Tox21 data set.

---

## IX. References

B. Bienfait and P. Ertl, [JSME: a free molecule editor in JavaScript](#), J. Cheminformatics 5:24 (2013)

**END OF THE USER'S MANUAL**

---