

VInSMoC

This repository documents the VInSMoC web application and runs through an example of how to launch a run.

VInSMoC is a mass spectral search tool that scores query mass spectra against a database of candidate molecules. VInSMoC can run in exact mode or variable mode. In exact mode VInSMoC matches query spectra to candidate molecules whose masses match the query spectra's precursor m/z. In variable mode VInSMoC matches query spectra to candidate molecules whose variants can explain the difference Δ between the precursor mass of spectra and mass of the molecule. VInSMoC supports both positive and negative mode spectra.

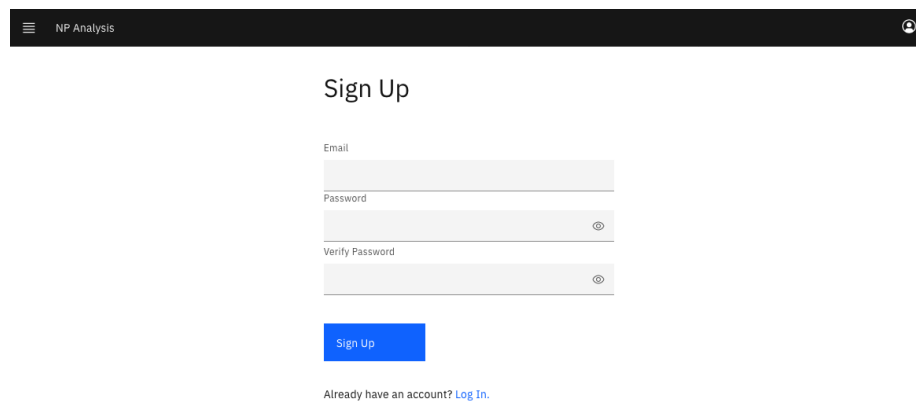
There is a demo video available in `demo_video/vinsmoc_demo.mov`.

The npanalysis.org server

VInSMoC, among other tools, is hosted at <https://run.npanalysis.org>.

Creating an account

If you do not have one yet, create an account at <https://run.npanalysis.org/#/signup> using your email.



The screenshot shows the 'Sign Up' page of the VInSMoC web application. At the top, there is a dark navigation bar with a hamburger menu icon on the left, the text 'NP Analysis' in the center, and a user profile icon on the right. The main content area has a white background. The title 'Sign Up' is centered at the top of the form. Below the title are three input fields: 'Email', 'Password', and 'Verify Password'. Each field has a light gray border and a small eye icon on the right side of the 'Password' and 'Verify Password' fields. Below the input fields is a blue 'Sign Up' button. At the bottom of the form, there is a link that says 'Already have an account? Log In.'

Figure 1: signup

Your dashboard

You can see all relevant information at <https://run.npanalysis.org/#/dashboard>.

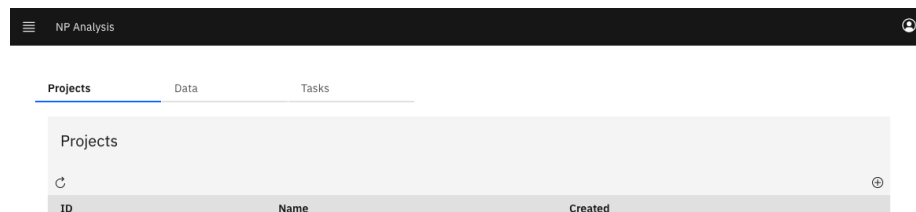


Figure 2: dashboard

There are three tabs: - Projects exist to organize your data and runs - Data stores all data you have uploaded to our server - Tasks stores all runs of our tools you have performed

Creating projects

From the project tab create a new project using the plus button.

Add a name to your project and click the submit button to create a new project

Uploading data

From the data tab click the plus button to upload your input data.

You will see a popup form.

For VInSMoC we need mass spectral inputs, so select “Mass Spectrum” as your data kind.

We support MGF, mzXML, and mzML inputs, select any such file on your computer and press submit to upload it to our sever. In this example we use a



Figure 3: plus button

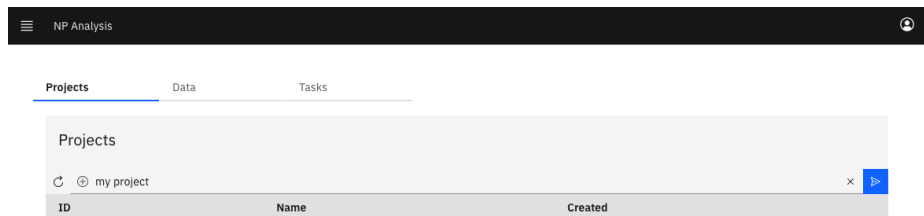


Figure 4: project name

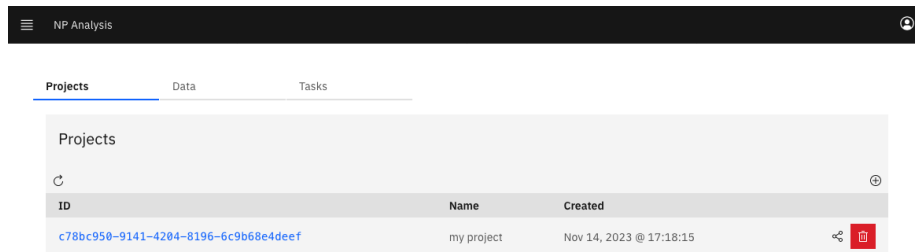


Figure 5: created project

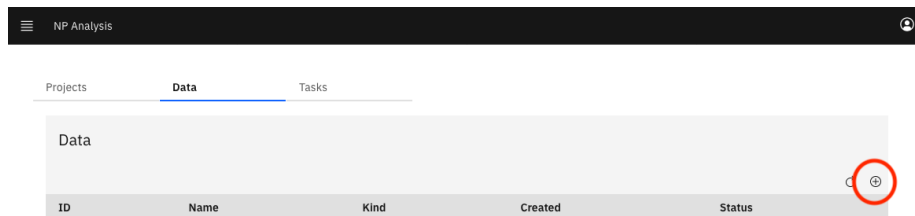


Figure 6: data button

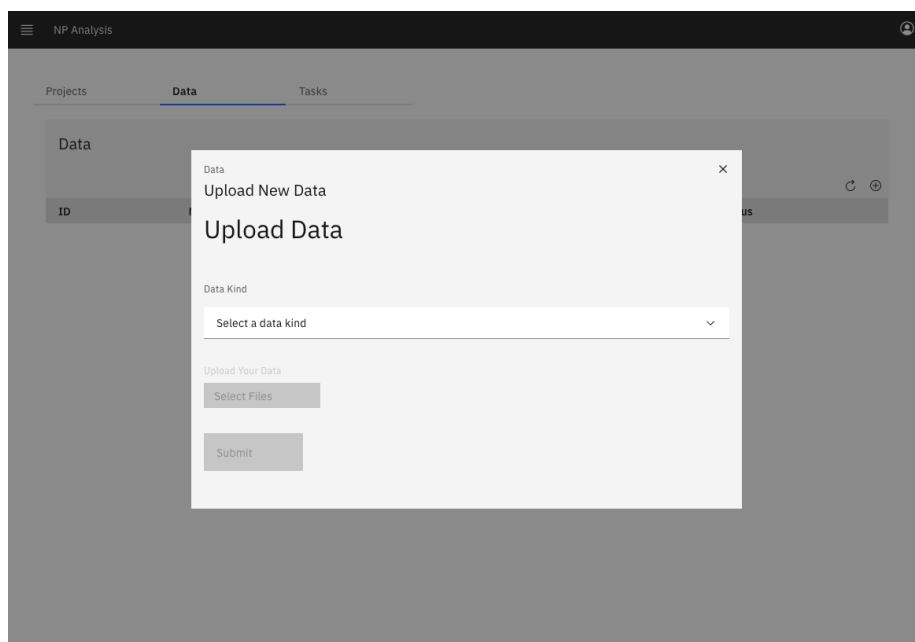


Figure 7: data form

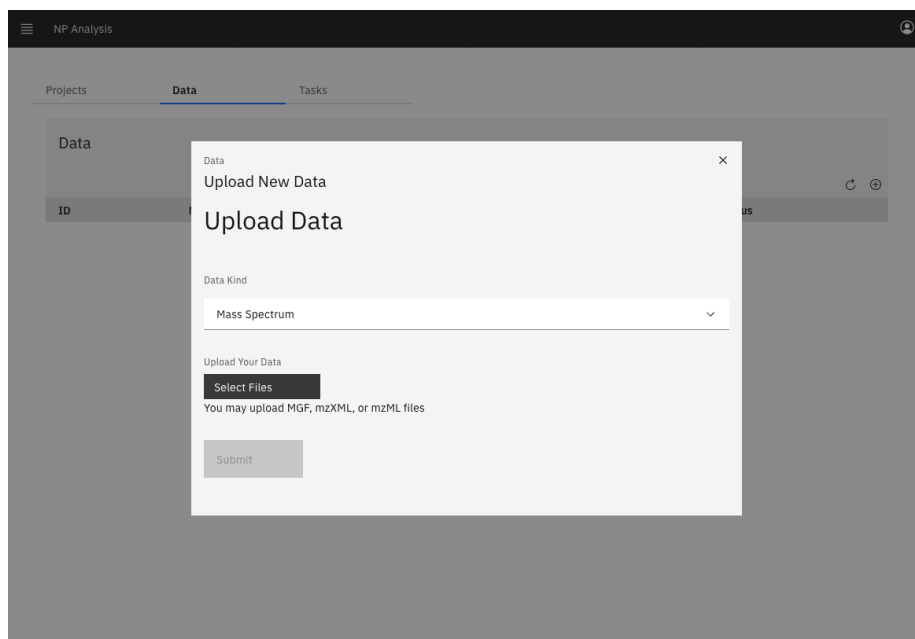


Figure 8: data kind

mass spectrum file from the MSV000086287 dataset.

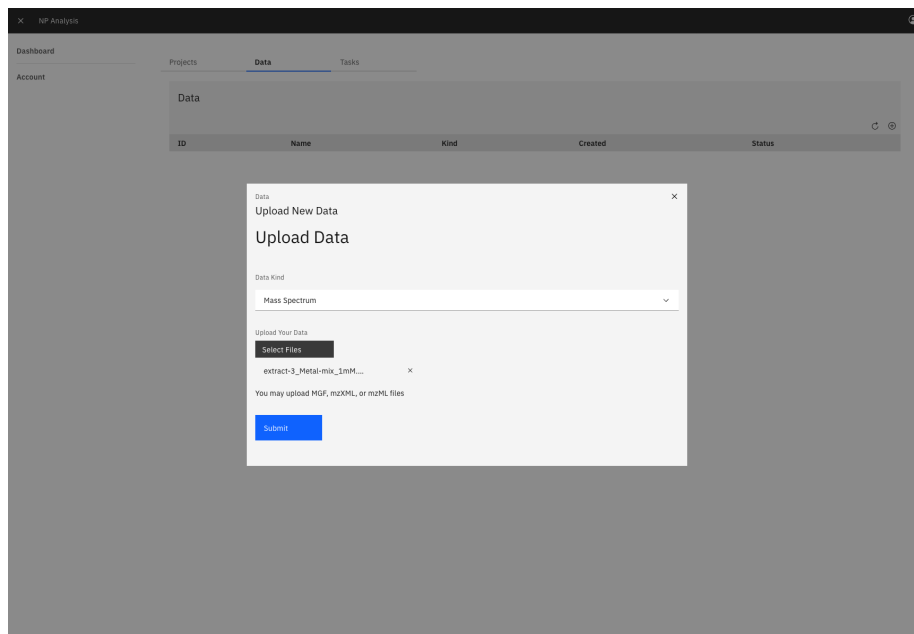


Figure 9: selected file

Each data input is parsed into a representation that is convenient for our tools to use downstream, so the data point is not immediately ready for use.

After a few minutes you should see that the data point has been successfully parsed.

Alternatively, if your input data was malformed you may see the status of the data point is an error. If this happens double check that your input format indeed matches the declared extension.

We now have data that we can use as input to VInSMoC.

Adding data to your project

Returning to the projects tab we can click on the ID of the project we previous created to get the project page.

Each project has two constituent tables: tasks and data. Data is associated to a project by the user that uploaded the data point. All public data provided by the developers is automatically available to every project. We can see that we have access to the NPAtlas chemical database already. However, we do not have access to our spectrum file, so we will associate it to the project using the plus button in the data table.

NP Analysis

Dashboard

Account

Projects

Data

Tasks

Data

ID	Name	Kind	Created	Status
37458e93-884f-4f6b-93e1-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 18:41:54	Pending

Figure 10: pending data

NP Analysis

Projects

Data

Tasks

Data

ID	Name	Kind	Created	Status
37458e93-884f-4f6b-93e1-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 18:42:17	Success

Figure 11: success data

NP Analysis

Projects

Data

Tasks

Projects

ID	Name	Created	
c78bc950-9141-4284-8196-6c9b68e4dee1	my project	Nov 14, 2023 @ 17:18:15	<div></div>

Figure 12: project id

This will create a form where we can select our spectrum from the dropdown and click import.

NP Analysis

Projects

my project

Share

Delete

ID

c78bc950-9141-4204-8196-6c9b68e4deef

Created

Nov 14, 2023 @ 17:18:15

Tasks

ID	Name	Method	Created	Completed	Status
----	------	--------	---------	-----------	--------

Data

ID	Name	Kind	Created	Status
622e1903-7248-458e-8da4-0258ff9b40cd	NPAtlas	Chemical Database	Nov 14, 2023 @ 17:27:22	Success

Figure 13: project page

NP Analysis

Projects

my project

Share

Delete

ID

c78bc950-9141-4204-8196-6c9b68e4deef

Created

Nov 14, 2023 @ 17:18:15

Tasks

ID

Name

Method

Created

Completed

Status

Data

ID

Name

Kind

Created

Status

622e1903-7248-458e-8da4-0258ff9b40cd

NPAAtlas

Chemical Database

Nov 14, 2023 @ 17:27:22


Success


Figure 14: project data plus button

NP Analysis

Projects

my project

Share 

Delete 

ID

Created

Tasks

Import Data

Associate Data with Project

Pick datum to import

Import

ID	Name	Method	Created	Completed	Status
622e1983-7248-458e-8da4-0258ff9b40cd	NPATlas	Chemical Database	Nov 14, 2023 @ 17:27:22		Success


Data


ID	Name	Kind	Created	Status
622e1983-7248-458e-8da4-0258ff9b40cd	NPATlas	Chemical Database	Nov 14, 2023 @ 17:27:22	Success

NP Analysis

Projects

my project

Share 

Delete 

ID

Created

Tasks

Import Data

Associate Data with Project

extract-3_Metal-mix_1mM.mzML

Import

ID	Name	Method	Created	Completed	Status
622e1983-7248-458e-8da4-0258ff9b40cd	NPATlas	Chemical Database	Nov 27, 2023 @ 18:42:39		Success

Data

ID	Name	Kind	Created	Status
622e1983-7248-458e-8da4-0258ff9b40cd	NPATlas	Chemical Database	Nov 27, 2023 @ 18:42:39	Success

We now see that our `extract-3_Metal-mix_1mM.mzML` file is available to this project.

The screenshot shows the NP Analysis web interface. At the top, there's a header with a menu icon and 'NP Analysis'. Below it, the page title is 'Projects' and the project name is 'my project'. There are 'Share' and 'Delete' buttons. The main content area is divided into sections: 'ID' (c78bc950-9141-4204-8196-6c9b68e4deef), 'Created' (Nov 14, 2023 @ 17:18:15), 'Tasks' (with a plus icon), and 'Data' (with a plus icon). The 'Tasks' table has columns: ID, Name, Method, Created, Completed, and Status. The 'Data' table has columns: ID, Name, Kind, Created, and Status. Two data rows are shown in the 'Data' table, both with a 'Success' status.

ID	Name	Method	Created	Completed	Status
c78bc950-9141-4204-8196-6c9b68e4deef			Nov 14, 2023 @ 17:18:15		

ID	Name	Kind	Created	Status
37458e93-884f-4f6b-93e1-518d52cf294e	extract-3_Metab-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 18:43:17	Success
622e1983-7248-438e-8d94-8258ff9848cd	NPAtlas	Chemical Database	Nov 27, 2023 @ 18:43:17	Success

Figure 15: project page post data

Running VInSMoC

We can run VInSMoC (and any future supported methods) by creating a new task with the plus button in the tasks table.

We will be greeted by a form that prompts us to select a method.

Select VInSMoC from the dropdown.

This form is split into three sections that we will fill in order.

First, we need to select the mass spectrum file and chemical database we want to use for this run. We can do so using the dropdowns under the “Input Data” section

Next in the “Task Parameters” section we need to name our task.

The other fields under “Task Parameters” control how VInSMoC is run. Search mode can be either exact or variable. Fragmentation mode can work with generic molecules (Metaoblite mode) or specifically generalized peptides (Peptide mode). The ion mode should be set to match your input spectrum (positive in our case). Precursor and product ion tolerances determine thresholds used to match spectra to candidate molecules and theoretical fragments to experimental peaks respectively. If you’re unsure these can be safely left to the default. The maximum absolute charge controls what adducts are used to convert m/z val-

NP Analysis

Projects

my project

Share

Delete

ID

c78bc950-9141-4204-8196-6c9b68e4deef

Created

Nov 14, 2023 @ 17:18:15

Tasks

+

⊖

ID	Name	Method	Created	Completed	Status
37458e93-884f-4f6b-93e3-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 10:43:17		Success
622e1983-7248-458e-8da4-8258ff9b48cd	NPADias	Chemical Database	Nov 27, 2023 @ 10:43:17		Success

Data

+

⊖

ID	Name	Kind	Created	Status
37458e93-884f-4f6b-93e3-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 10:43:17	Success
622e1983-7248-458e-8da4-8258ff9b48cd	NPADias	Chemical Database	Nov 27, 2023 @ 10:43:17	Success

Figure 16: task plus button

NP Analysis

Projects

my project

Share

Delete

ID

c78bc950-9141-4204-8196-6c9b68e4deef

Created

Nov 14, 2023 @ 17:18:15

Tasks

+

⊖

ID	Name	Method	Created	Completed	Status
37458e93-884f-4f6b-93e3-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 10:43:17		Success
622e1983-7248-458e-8da4-8258ff9b48cd	NPADias	Chemical Database	Nov 27, 2023 @ 10:43:17		Success

Data

+

⊖

ID	Name	Kind	Created	Status
37458e93-884f-4f6b-93e3-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 10:43:17	Success
622e1983-7248-458e-8da4-8258ff9b48cd	NPADias	Chemical Database	Nov 27, 2023 @ 10:43:17	Success

Figure 17: task form

NP Analysis

Projects
my project

ID

Created

Tasks

ID Name

Data

ID

37458e93-884f-4f6b-93e3-518d52cf294e

622e19b3-7248-458e-8da4-8258ff9b48cd

New Task
Submit a Task

Method
VINSMOC

Input data
Mass Spectrum
Select a data point

Chemical Database
Select a data point

Task Parameters
Task Name

Search Mode
Exact

Fragmentation Mode
Metabolite

Ion Mode
Positive

Precursor Ion Tolerance
0.01

Product Ion Tolerance
0.01

Max Absolute Charge
1

Minimum Score
10

Advanced Settings

Submit

Share

Delete

Status

Status

Success

Success

Figure 18: task form method

NP Analysis

Projects
my project

ID

Created

Tasks

ID Name

Data

ID

37458e93-884f-4f6b-93e3-518d52cf294e

622e19b3-7248-458e-8da4-8258ff9b48cd

New Task
Submit a Task

Method
VINSMOC

Input data
Mass Spectrum
extract-3_Metal-mix_1mM.mzML

Chemical Database
NPAtlas

Task Parameters
Task Name

Search Mode
Exact

Fragmentation Mode
Metabolite

Ion Mode
Positive

Precursor Ion Tolerance
0.01

Product Ion Tolerance
0.01

Max Absolute Charge
1

Minimum Score
10

Advanced Settings

Submit

Share

Delete

Status

Status

Success

Success

Figure 19: task form data

Figure 20: task form named

ues to masses. A maximum absolute charge of 1 means only $[M+H]^+$ or $[M-H]^-$ adducts are used, based on ion mode. A maximum absolute charge of 2 means adds $[M+2H]^{++}$ or $[M-2H]^{--}$ adducts are used, based on ion mode. Finally, the minimum score controls which hits are reported.

The advanced settings should be need to be changed often. For VInSMoC these settings control how the mass spectral input is pre-processed before it is scored against the candidate molecules. Peak merging combines peaks that are within “Peak Merge Threshold” of one another, making sure to leave at least “Minimum Peak Count” peaks. Peak filtering only keeps the “Peaks Per Window” most intense peaks in m/z windows of “Peak Filtering Window”. These should be left unchanged unless you have already down some preprocessing on your input spectra before uploading to our server.

We are now ready to launch our task by clicking submit.

Monitoring runs

Upon submission our task went into the pending state while its execution is scheduled by our backend.

Our example is small, so it finished almost immediately, but depending on your input size you may spend some time in the “Running” state.

The screenshot shows the NP Analysis web interface. A modal window titled "New Task Submit a Task" is open. The "Method" is set to "VinsMoc". The "Input data" section is collapsed. The "Task Parameters" section is expanded, showing "Advanced Settings". The settings include:

- Merge Peaks:** ☒ On
- Minimum Peak Count:** 1
- Peak Merge Threshold:** 0.85
- Filter Peaks:** ☒ On
- Peak Filtering Window:** 50
- Peaks Per Window:** 5
- Compute P-values:** ☐ Off

A blue "Submit" button is at the bottom of the modal. The background shows a sidebar with "Projects" and "my project", and a table of tasks with columns for ID, Name, and Status.

Figure 21: advanced settings

The screenshot shows the NP Analysis web interface. A modal window titled "New Task Submit a Task" is open. The "Method" is set to "VinsMoc". The "Input data" section is expanded, showing:

- Mass Spectrum:** extract-3_Metal-mix_1.mzML
- Chemical Database:** NPAtlas

The "Task Parameters" section is expanded, showing:

- Task Name:** test vs npatlas
- Search Mode:** Exact
- Fragmentation Mode:** Metabolite
- Ion Mode:** Positive
- Precursor Ion Tolerance:** 0.01
- Product Ion Tolerance:** 0.01
- Max Absolute Charge:** 1
- Minimum Score:** 10

The "Advanced Settings" section is collapsed. A blue "Submit" button is at the bottom of the modal, which is circled in red. The background shows the same sidebar and task table as Figure 21.

Figure 22: task form submit

NP Analysis

Projects

my project

Share

Delete

ID

c78bc950-9141-4204-8196-6c9b68e4deef

Created

Nov 14, 2023 @ 17:18:15

Tasks

ID	Name	Method	Created	Completed	Status	
8ad880b-cc3e-4af5-bdc4-51b454ca810b	test vs npAtlas	VinSMoC	Nov 27, 2023 @ 18:44:54		Pending	<div></div>

Data

ID	Name	Kind	Created	Status	
37458e93-884f-4f6b-93e1-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 18:43:17	Success	<div></div>
622e1983-7248-458e-8da4-8258ff9b48cd	NPAtlas	Chemical Database	Nov 27, 2023 @ 18:43:17	Success	<div></div>

Figure 23: pending task

NP Analysis

Projects

my project

Share

Delete

ID

c78bc950-9141-4204-8196-6c9b68e4deef

Created

Nov 14, 2023 @ 17:18:15

Tasks

ID	Name	Method	Created	Completed	Status	
8ad880b-cc3e-4af5-bdc4-51b454ca810b	test vs npAtlas	VinSMoC	Nov 27, 2023 @ 18:44:54	Nov 27, 2023 @ 18:45:10	Finished	<div></div>

Data

ID	Name	Kind	Created	Status	
37458e93-884f-4f6b-93e1-518d52cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 18:43:17	Success	<div></div>
622e1983-7248-458e-8da4-8258ff9b48cd	NPAtlas	Chemical Database	Nov 27, 2023 @ 18:43:17	Success	<div></div>

Figure 24: finished task

We can click on the ID of the task to see the results.

The screenshot shows the NP Analysis web interface. At the top, there's a header with 'NP Analysis' and a user icon. Below the header, the page title is 'Projects' and the sub-header is 'my project'. On the right, there are 'Share' and 'Delete' buttons. The main content area displays project details: ID (c78bc950-9141-4204-8196-6c9b68e4deef) and Created (Nov 14, 2023 @ 17:18:15). Below this is a 'Tasks' section with a table. The first row of the table has its ID, '8ad888db-cc3e-4af5-bdc4-51b454ca818b', circled in red. The table has columns for ID, Name, Method, Created, Completed, and Status. Below the tasks table is a 'Data' section with another table. This table has columns for ID, Name, Kind, Created, and Status. It lists two data entries: one for 'extract-3_Metal-mix_1mM.mzML' and another for 'NPAtlas'.

ID	Name	Method	Created	Completed	Status
8ad888db-cc3e-4af5-bdc4-51b454ca818b	test vs npAtlas	VinSProC	Nov 27, 2023 @ 18:44:54	Nov 27, 2023 @ 18:45:10	Finished

ID	Name	Kind	Created	Status
37458e93-884f-4f6b-93e3-51b452cf294e	extract-3_Metal-mix_1mM.mzML	Mass Spectrum	Nov 27, 2023 @ 18:43:17	Success
622e1983-7248-458e-8da4-8258f19b48cd	NPAtlas	Chemical Database	Nov 27, 2023 @ 18:43:17	Success

Figure 25: task id click

This brings us to the task page which shows the best hit we got to each input scan that exceeded the minimum score reporting threshold. Each molecule is a link to its entry in the original chemical database that it came from.

If the task is not finished this page will automatically refresh with an exponential backoff until the task is completed.

NP Analysis

Task

test vs npatlas

ID

0ad080db-cc3e-4af5-bdc4-51bd54ca010b

Method

VInSMoC

Processing Status

Finished

Created

Nov 27, 2023 @ 18:44:54

Completed

Nov 27, 2023 @ 18:45:10

Results

Items per page: 10 1-10 of 18 items1 of 2 pages

Molecule	Scan Number	Precursor m/z	Retention Time	Molecule Exact Mass	Mass Error	Adduct	Spectrum Charge	P-Value	Score
Apiosporic acid	1506	225.148569481348	181.702	224.141	0.000	[M+H] ⁺			14
Phaeoschidin A	2229	547.160943080364	276.677	546.153	0.001	[M+H] ⁺			14
Methylinoscavin D	192	411.098591150193	23.525	410.100	-0.009	[M+H] ⁺			14
Harzianolide	763	223.133007752409	93.656	222.126	0.000	[M+H] ⁺			13
Phaeoschidin A	2188	547.15980489097	270.757	546.153	-0.000	[M+H] ⁺			13
Tf-26Vx	2152	547.16084942511	265.591	546.153	0.001	[M+H] ⁺			12

Figure 26: task results