

mzkit

Data toolkits for processing LC-MS and GC-MS data

[View project on GitHub](#)

Mzkit



Mzkit is an open source raw data file toolkit for mass spectrometry data analysis, provides by the [BioNovoGene](#) corporation. The features of mzkit includes: raw data file content viewer(XIC/TIC/Mass spectrum plot), build molecule network, formula de-novo search and annotation.

--===== TOC =====

4. [Raw Data Viewer Instruction](#)

- [1.1. Imports raw data file](#)
- [1.2. Search Feature](#)
- [1.3. XIC plot](#)
 - [1.3.1. XIC overlay](#)
- [1.4. TIC plot](#)
- [1.5. View Mass spectra](#)
- [1.6. Save Plot and Export matrix](#)
 - [1.6.1. Export XIC](#)

4. [Data Annotations](#)

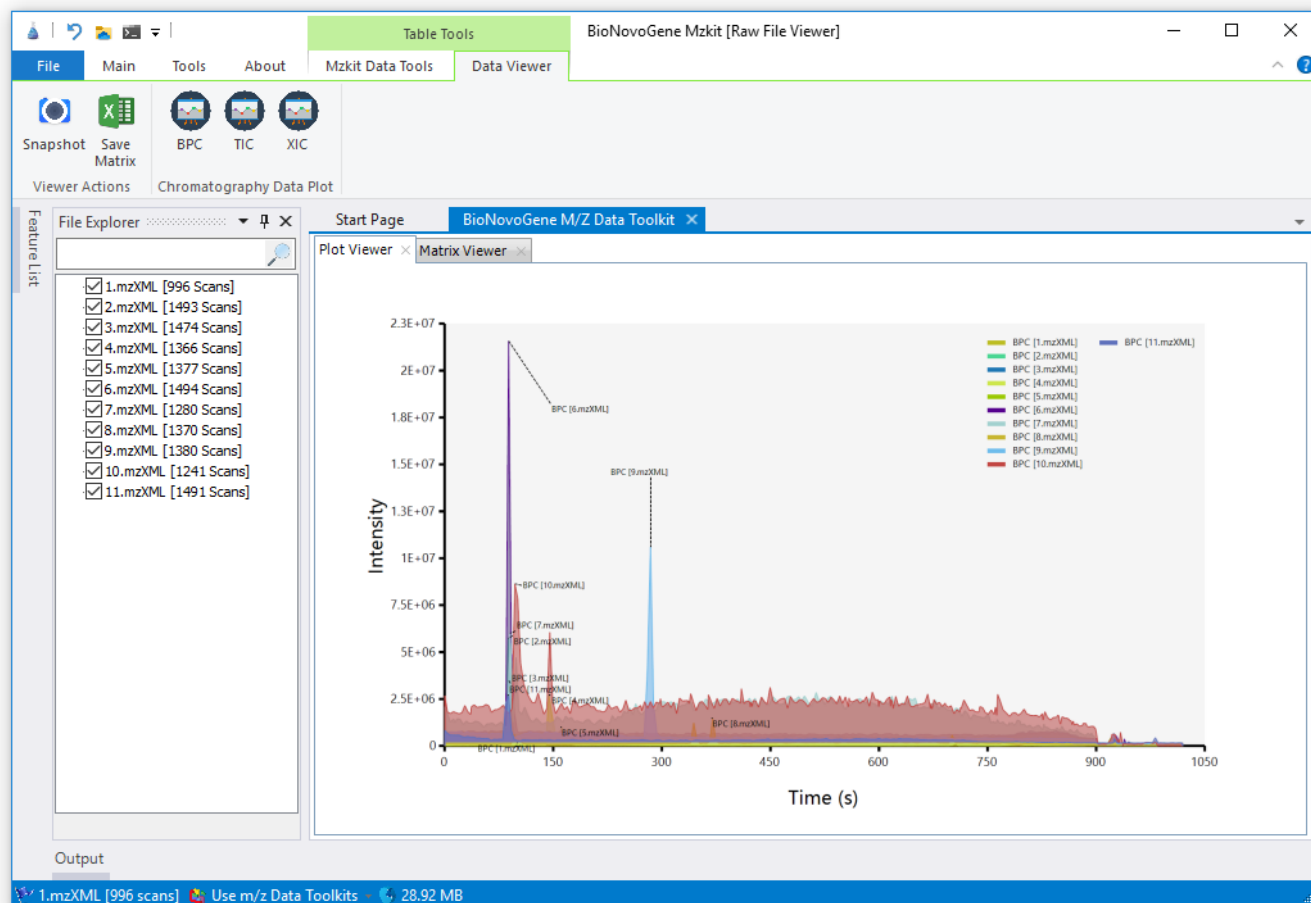
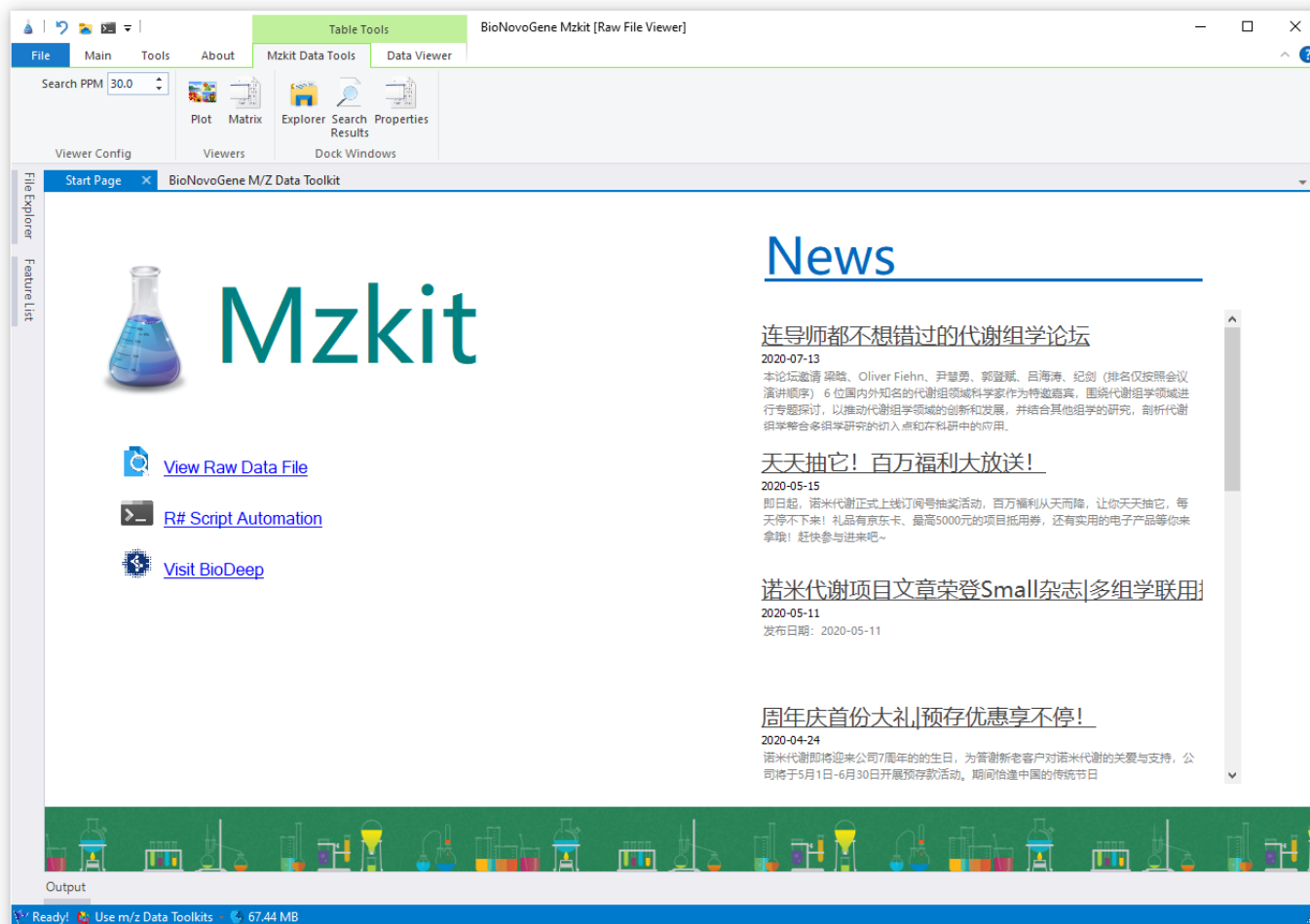
- [2.1. Formula search](#)
 - [2.1.1. Export Formula Search Result](#)
- [2.2. Molecular Networking](#)
 - [2.2.1. step1 select ions](#)
 - [2.2.2. step2 build network](#)
 - [2.2.3. step3 view network data](#)
 - [2.2.4. step 4 network visualization](#)
 - [2.2.5. Export Network Data](#)
 - [2.2.6. Save Network Visual](#)

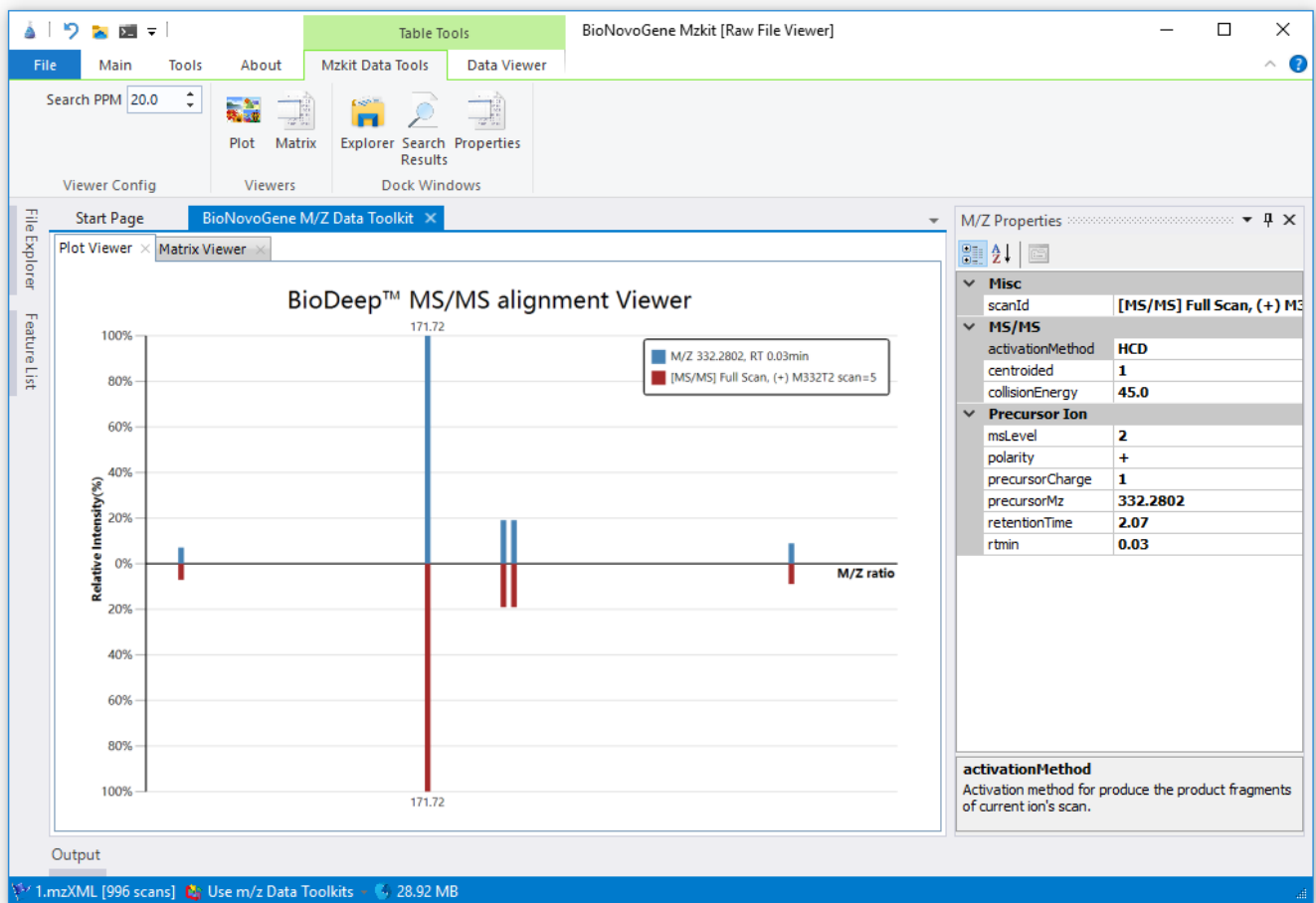
4. [Appendix](#)

- [3.1. Switch Between Toolkit](#)
- [3.2. Install Mzkit](#)
- [3.3. Uninstall Mzkit](#)

4. [Credits](#)

Product Screenshots

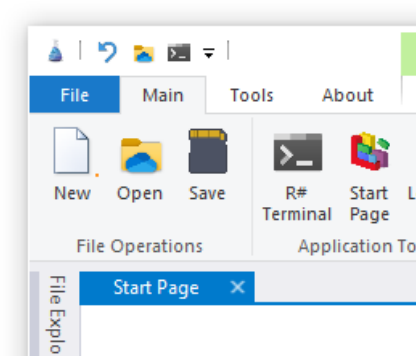




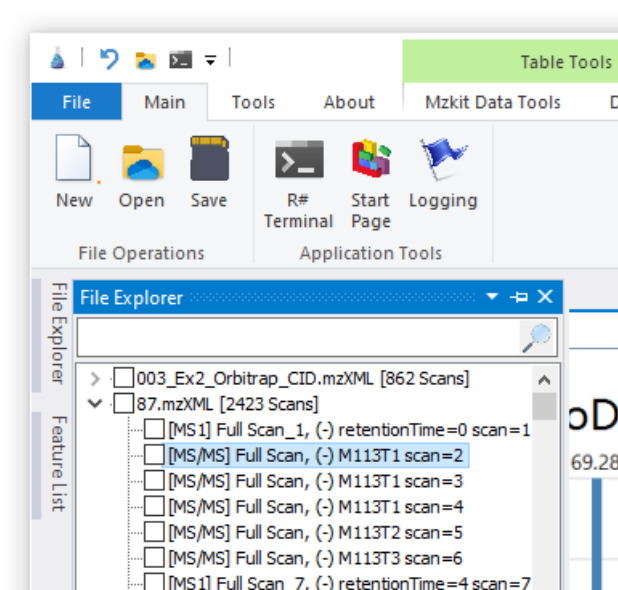
I. Raw Data Viewer Instruction

1.1. Imports raw data file

For view the file content of the mzXML or mzML datafile in mzkit, you must imports the raw data file into the mzkit at first. Here is how: select the **Main** tabpage of mzkit you will see the **Open** command for the raw data imports operation. Then you are going to click this **Open** command button, choose the raw data file for imports and wait for it finished.



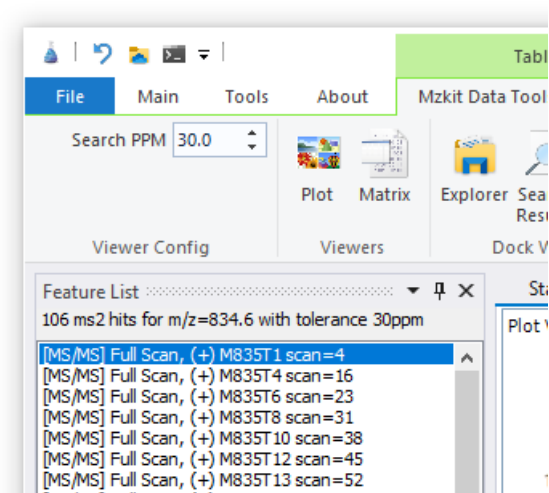
Then you should see the files that you've imports into mzkit on the **File Explorer** dock panel if there is no error occurs during the raw data file imports progress. Now you can click on the raw data file tree to expend it and click one of the feature in your raw file to view the content data.



1.2. Search Feature

the search bar on the top of the file tree is the **m/z** search input: you can input a specific **m/z** value or **formula** expression in the search bar for search the matched features in your raw data file. This operation is usually apply for the **XIC** data search.

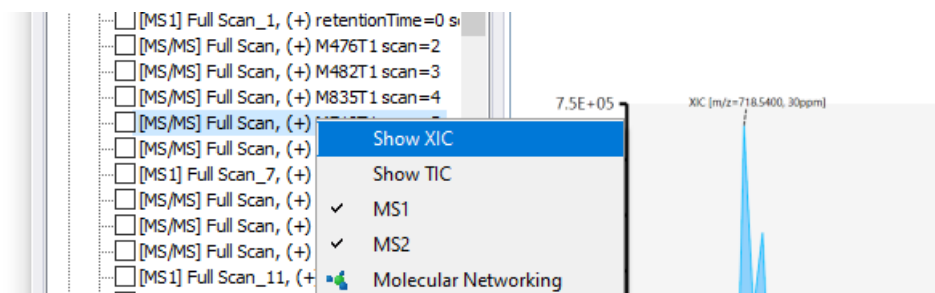
When you have click on the search button, then all of the **m/z** feature in your raw data file that match the **ppm** condition will be listed in the **Feature List** dock panel:



An example of search m/z 834.6 with tolerance error 30ppm. the result in the **Feature List** is usually used for create a XCI plot.

1.3. XIC plot

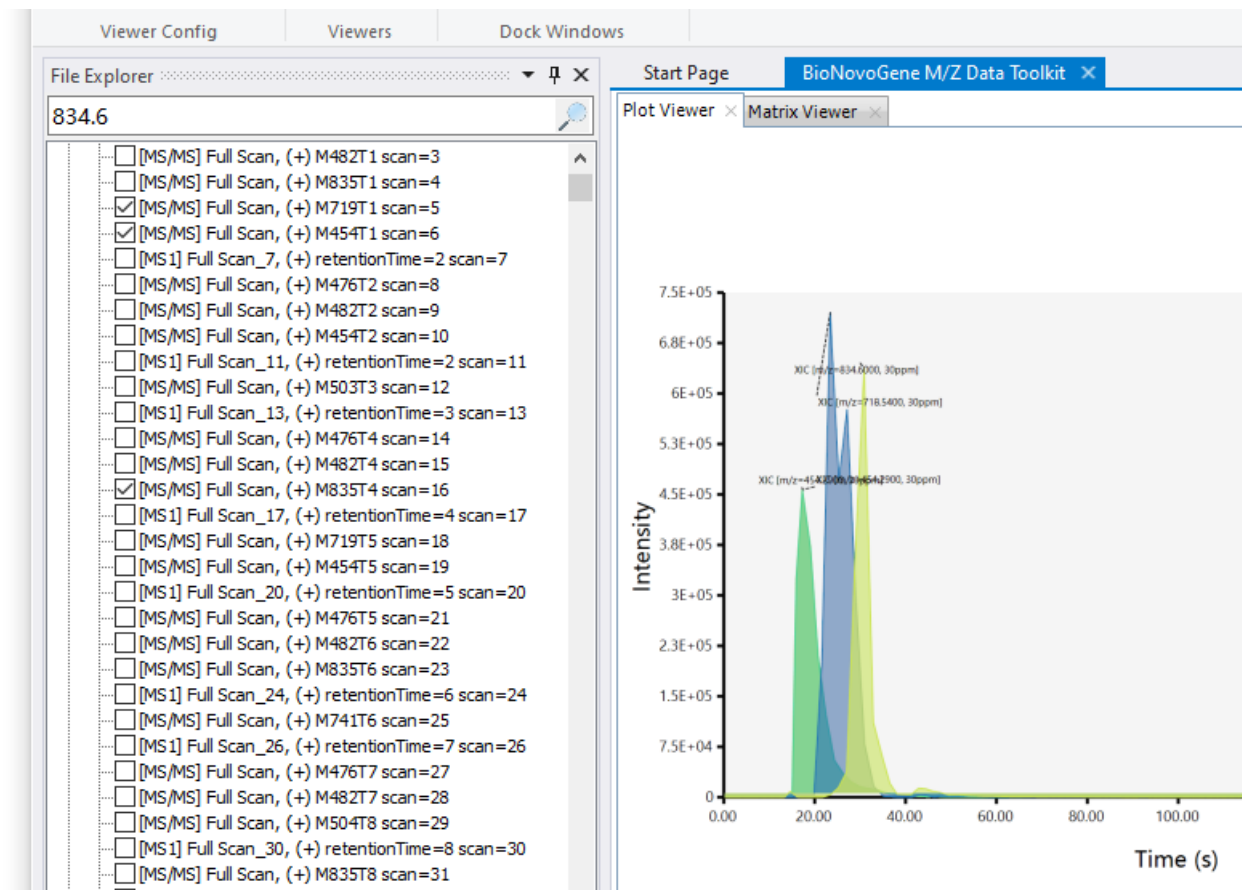
Expend the file content tree in the **File Explorer**, and then mouse right click on one MS2 feature in your file, select **XIC** for create a XIC plot for a specific ion feature:



The XIC plot is a kind of time-signal chromatography plot of a specific m/z ion.

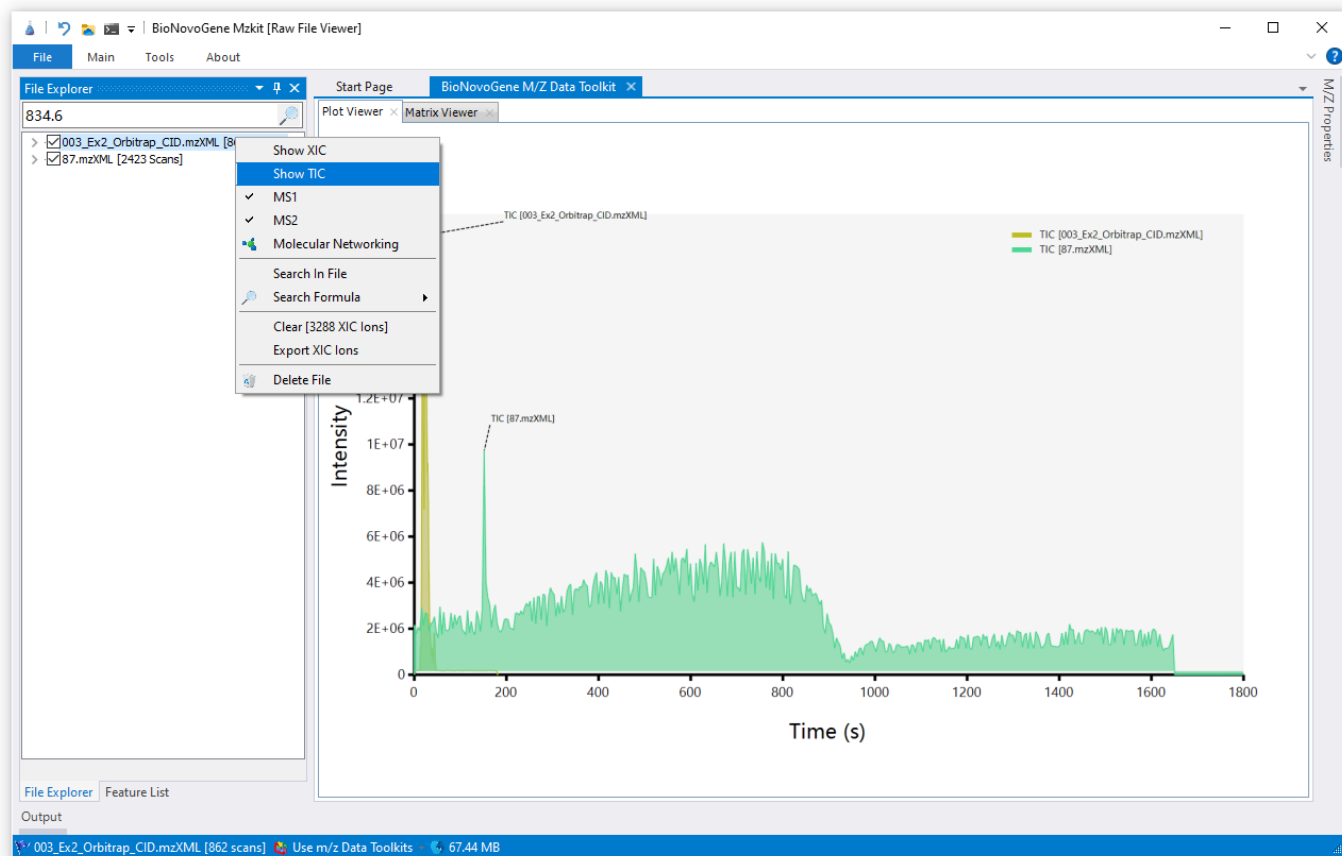
1.3.1. XIC overlay

you can click on the checkbox besides the Ms2 feature for select different ion feature for create the XIC overlay plot:



1.4. TIC plot

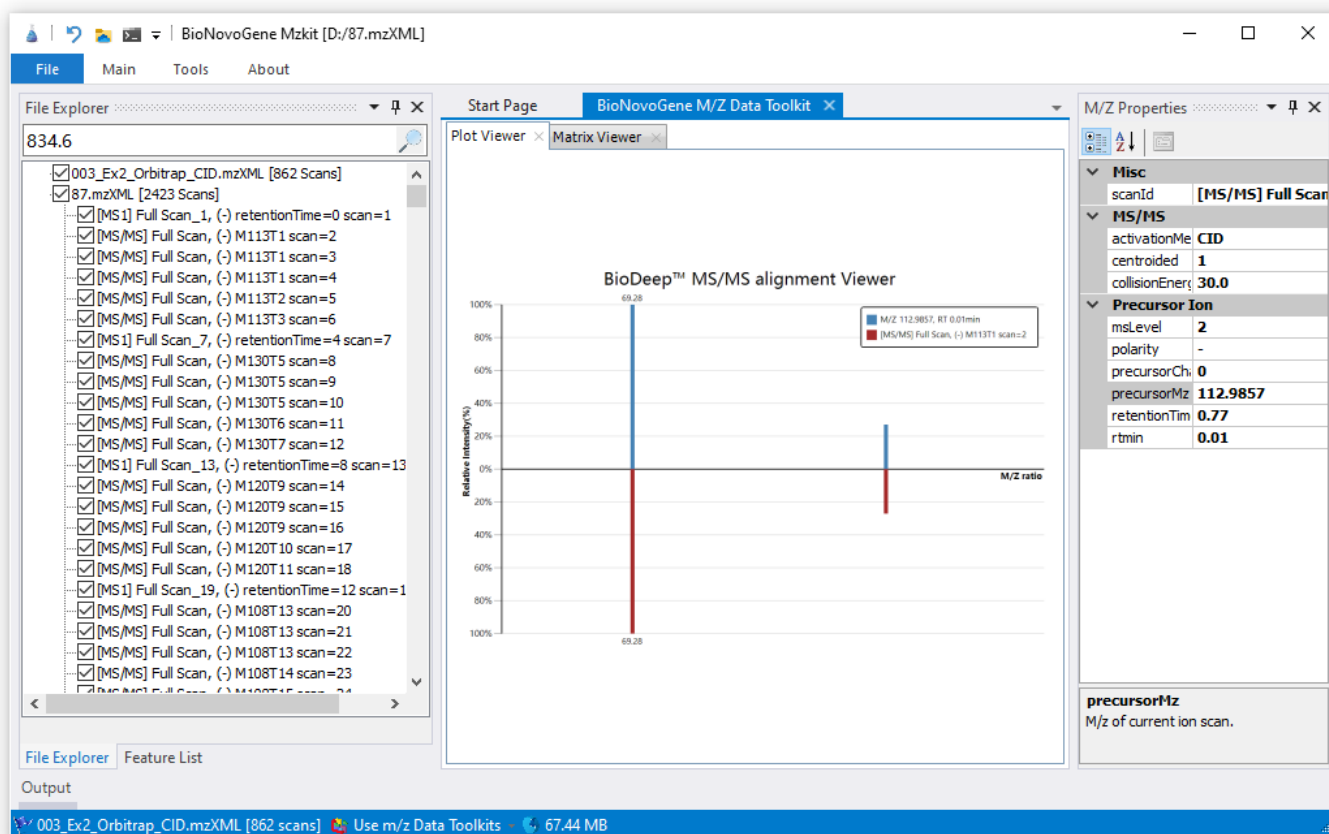
as the same as create a XIC plot, you also can create TIC plot for a single file or multiple file by select multiple file by check on checkbox:



The TIC plot is similar to the XIC plot, data is generated from all ions.

1.5. View Mass spectra

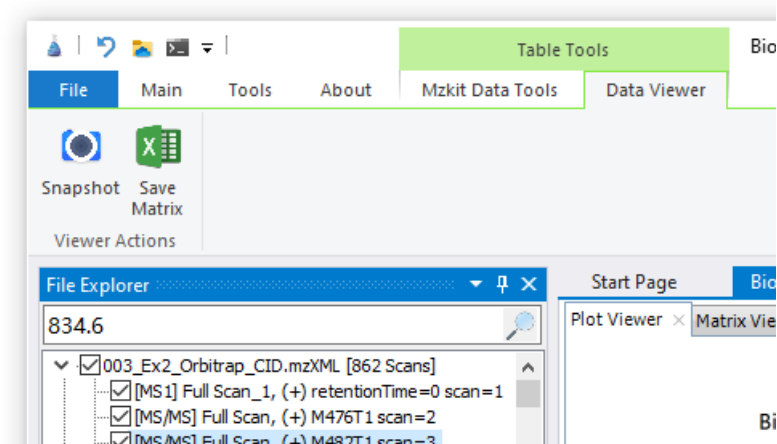
For view the mass spectra data in your file, just click on one of the scan feature in your raw data file:



1.6. Save Plot and Export matrix

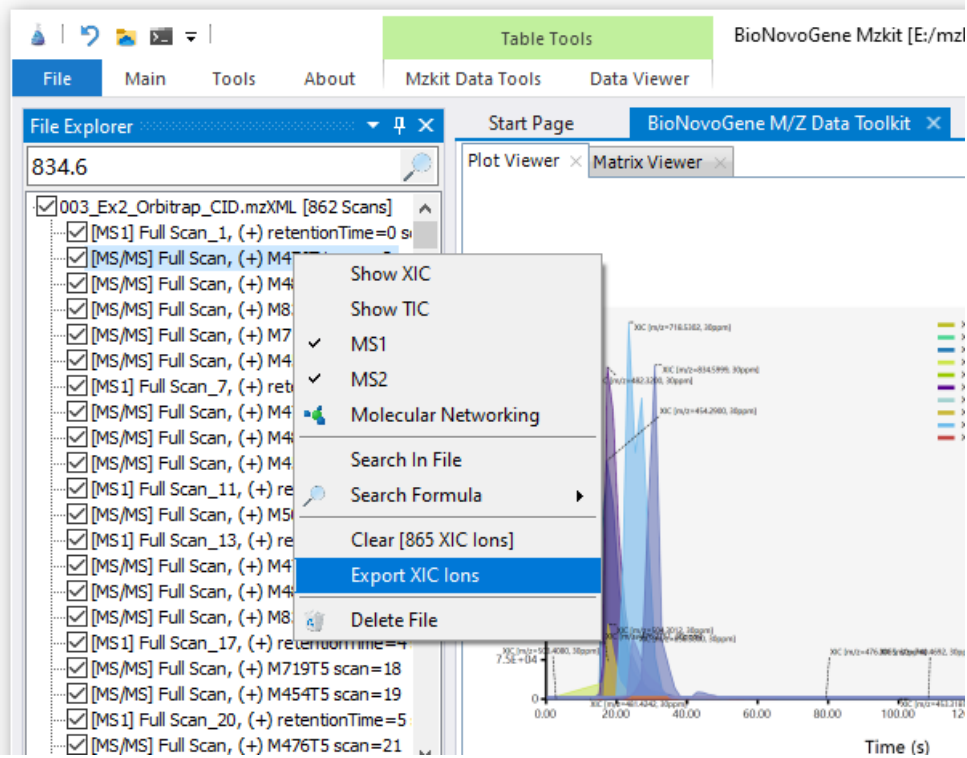
The mzkit application provides the function for save the plot image and the plot data in your raw data file. for example, select the **Data Viewer** tab page in mzkit, you will see two viewer action buttons in the menu:

1. **[Snapshot]** for export the XIC/TIC/MS2 data plot image to a specific file.
2. **[Save Matrix]** for export the Mass spectra or Chromatography data to a specific Excel table file.



1.6.1. Export XIC

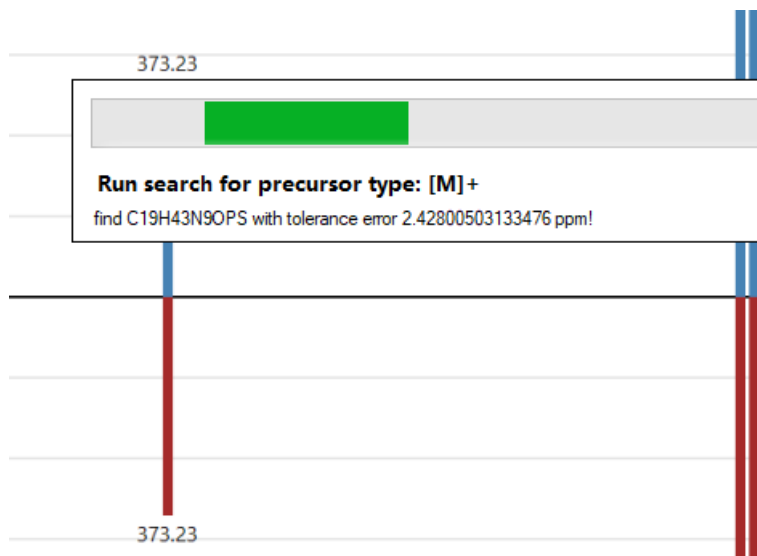
Export the XIC data to a specific file is also keeps simple, just mouse right click on the file content tree and then choose **Export XIC Ions**.



2. Data Annotations

2.1. Formula search

You can search for the formula of one feature in your raw data file with mzkit for do some de-novo annotation with just simply mouse right click on one of the feature and then select **Search Formula**, and then wait for a while to let mzkit for search all of the candidate formula by enumerates all of the possible element combinations:



Once the mzkit have been done of formula search, then all of the matched formula will be shown on the result page. there are some information about the each formula search result is listed in the result table, includes: formula result, its corresponding exact mass value, mass error of the searched m/z and the m/z calculated from the exact mass, ion charge value from your raw file, precursor type information, etc.

Table Tools Formula Search Tools BioNovoGene Mzkit [M/Z Formula De-no...

File Main Tools About Mzkit Data Tools Data Viewer Formula Result

Export Data

Start Page BioNovoGene M/Z Data Toolkit

Formula De-novo Search

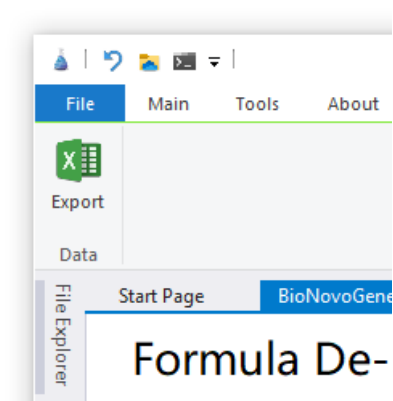
Input m/z: Search Formula Formula Search Profile: Custom

| | Formula | Exact Mass | PPM | Charge | Adducts | M | Precursor Type |
|--|--------------------------------|-------------|-------------------|--------|----------|---|--------------------|
| | C10H60N4O3PS5 | 475.30064 | 3.96745530050556 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H38N15O4P | 475.2968678 | 3.96902668275556 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H43N10O8S | 475.2985898 | 0.346042224054862 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H45N10O6P2 | 475.299863 | 2.33269402027553 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H46N11O3PS2 | 475.2963976 | 4.95829921045409 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H48N11OP3S | 475.2976708 | 2.2795629660041 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H51N6O7S3 | 475.2981196 | 1.33531475163379 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H53N6O5P2S2 | 475.2993928 | 1.34342149269659 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H55N6O3P4S | 475.300666 | 4.02215773726617 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H56N7P3S3 | 475.2972006 | 3.26883549358304 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H58N9PS3 | 475.3011148 | 4.96640595163648 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H59N2O6S5 | 475.2976494 | 2.32458727921273 | 1 | 1.007276 | 1 | [M+H] ⁺ |
| | C11H60O8P5 | 475.297611 | 2.40537857038609 | 1 | 1.007276 | 1 | [M+H] ⁺ |

Output

Run formula search for m/z 476.306030273438 with tolerance error 5 ppm, have 591 formula found! Use m/z Data Toolkits 67.44 MB

You also can search of the formula by input any **m/z** value in the input box of the search result page. Click on the formula then you could submit the candidate formula into the biodeep database for search of the metabolite information.



2.1.1. Export Formula Search Result

Export the formula search result just like other data that we've introduced before, just click on the **Formula Result** tab page and then click on the Export button. this command will let you save the formula search result into a specific Excel table file.

2.2. Molecular Networking

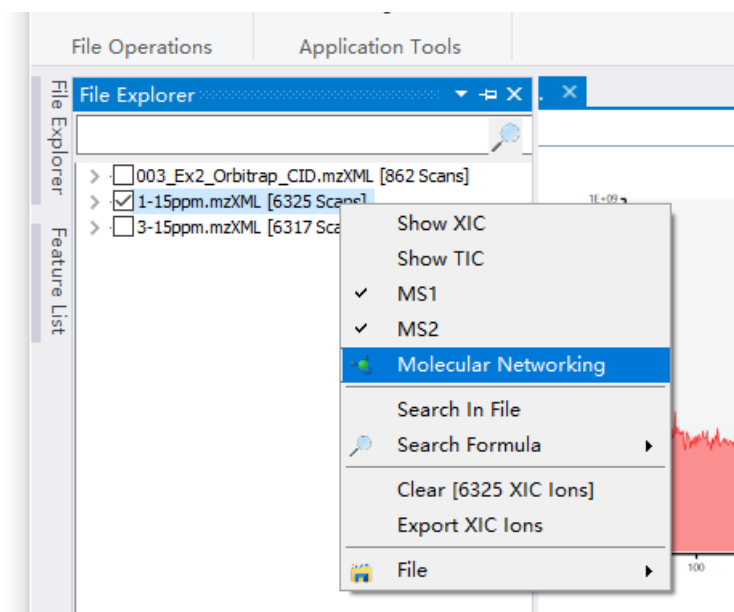
The molecular networking in mass spectrum data is a kind of spectrum similarity network. You can build a molecule network use mzkit in a very easy way:

2.2.1. step1 select ions

For create a spectrum similarity network, it required some ms2 ions data provides to mzkit for run spectrum matrix alignment and evaluate the simialrity scores between these spectrum matrix data. select the ions data just like plot XIC operation it does in mzkit: check on the ms2 feature in the **file explorer** dock panel.

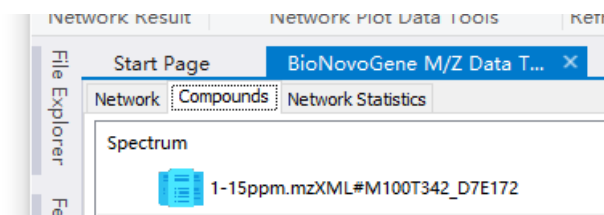
2.2.2. step2 build network

Once we have the mass spectrum data selected, then we could run the matrix similarity between these matrix for build a network. now we mouse right click of the file tree in the **file explorer**, and then choose the **Molecular Networking** menu item:



2.2.3. step3 view network data

once the networking progress have been done, then we could view the network result data in the **Molecular Networking** tool page. there are three tables in the tool page for show your network data: **Network**, **Compounds** and **Network Statistics**.



1. the **Network** tab page contains the edges data in your network, which is the spectrum cluster simialrity result. all of the spectrum alignment its simialrity value is less than the **Spectrum Similarity** threshold value that will be removes from the network.

File Explorer Feature List

Start Page BioNovoGene M/Z Data T... Application Settings Molecular Networking Vie... Molecular Networking Vie...

Network Compounds Network Statistics

| CompoundA | CompoundB | similarity | forward | reverse | View |
|--------------------|--------------------|------------|---------|---------|----------------|
| m.mzXML#M148T160_2 | m.mzXML#M150T94_B0 | 0.9741 | 0.9912 | 0.9741 | View Alignment |
| m.mzXML#M148T160_2 | m.mzXML#M150T70_2F | 0.9739 | 0.9848 | 0.9739 | View Alignment |
| m.mzXML#M148T160_2 | m.mzXML#M190T180_B | 0.9496 | 0.9496 | 0.9683 | View Alignment |
| m.mzXML#M150T70_2F | m.mzXML#M391T800_C | 0.9144 | 0.9144 | 0.9770 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M718T818_3 | 0.9985 | 0.9987 | 0.9985 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M718T818_B | 1.0000 | 1.0000 | 1.0000 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M719T726_C | 0.9982 | 0.9982 | 1.0000 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M719T693_9 | 0.9964 | 0.9964 | 1.0000 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M720T790_4 | 0.9979 | 0.9979 | 1.0000 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M732T796_E | 0.9985 | 1.0000 | 0.9985 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M733T804_0 | 1.0000 | 1.0000 | 1.0000 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M833T706_5 | 0.9947 | 0.9947 | 0.9985 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M837T758_7 | 0.9917 | 0.9917 | 0.9980 | View Alignment |
| m.mzXML#M706T831_8 | m.mzXML#M861T816_2 | 0.9974 | 0.9974 | 0.9988 | View Alignment |

Output

The network size is huge for create layout, entire progress will be very slow... Use m/z Data Toolkits 266.72 MB Memory: 331.03 MB

there are columns in the edge table: **CompoundA** and **CompoundB** is the spectrum reference id in this edge connected. and the **similarity**, **forward** and **reverse** column is the similarity score value of the two spectrum matrix and the last **view** column contains the button that let you view the spectrum matrix alignment result visual plot.

1. the **Compounds** tab page contains the spectrum cluster (**network nodes**) information in your network.

File Explorer Feature List

Start Page BioNovoGene M/Z Data T... Application Settings Molecular Networking Vie... Molecular Networking Vie...

Network Compounds Network Statistics

| Spectrum | Cluster/File | Scans/Frag... | m/z | rt | rtmin |
|---|---------------|---------------|------------------|--------|--------|
| 1-15ppm.mzXML#M100T342_D7E172 | 6 | 2 | 100.0763125 | 341.78 | 313.94 |
| 1-15ppm.mzXML#M100T366_50B663 | 6 | 11 | 100.0763125 | 366.48 | 366.48 |
| 1-15ppm.mzXML#M100T48_05DBBC | 6 | 3 | 100.0763125 | 48.05 | 18.7 |
| 1-15ppm.mzXML#[MS/MS] Full Scan, (+) M100T19 scan=150 | 1-15ppm.mzXML | 7 | 100.0763 | 18.7 | n/a |
| 1-15ppm.mzXML#[MS/MS] Full Scan, (+) M100T239 scan=1815 | 1-15ppm.mzXML | 12 | 100.0763 | 239.4 | n/a |
| 1-15ppm.mzXML#[MS/MS] Full Scan, (+) M100T48 scan=384 | 1-15ppm.mzXML | 16 | 100.0763 | 48.05 | n/a |
| 1-15ppm.mzXML#M100T516_F54A12 | 6 | 18 | 99.5317777777778 | 515.99 | 71.56 |
| 1-15ppm.mzXML#M100T649_497514 | 6 | 1 | 99.5129 | 649.3 | 649.3 |
| 1-15ppm.mzXML#M101T175_64365B | 6 | 17 | 101.060330434783 | 174.99 | 49.12 |
| 1-15ppm.mzXML#M101T481_F06DBC | 6 | 1 | 101.060330434783 | 480.82 | 480.82 |
| 1-15ppm.mzXML#M101T635_FED0A0 | 6 | 5 | 101.060330434783 | 634.53 | 505.95 |

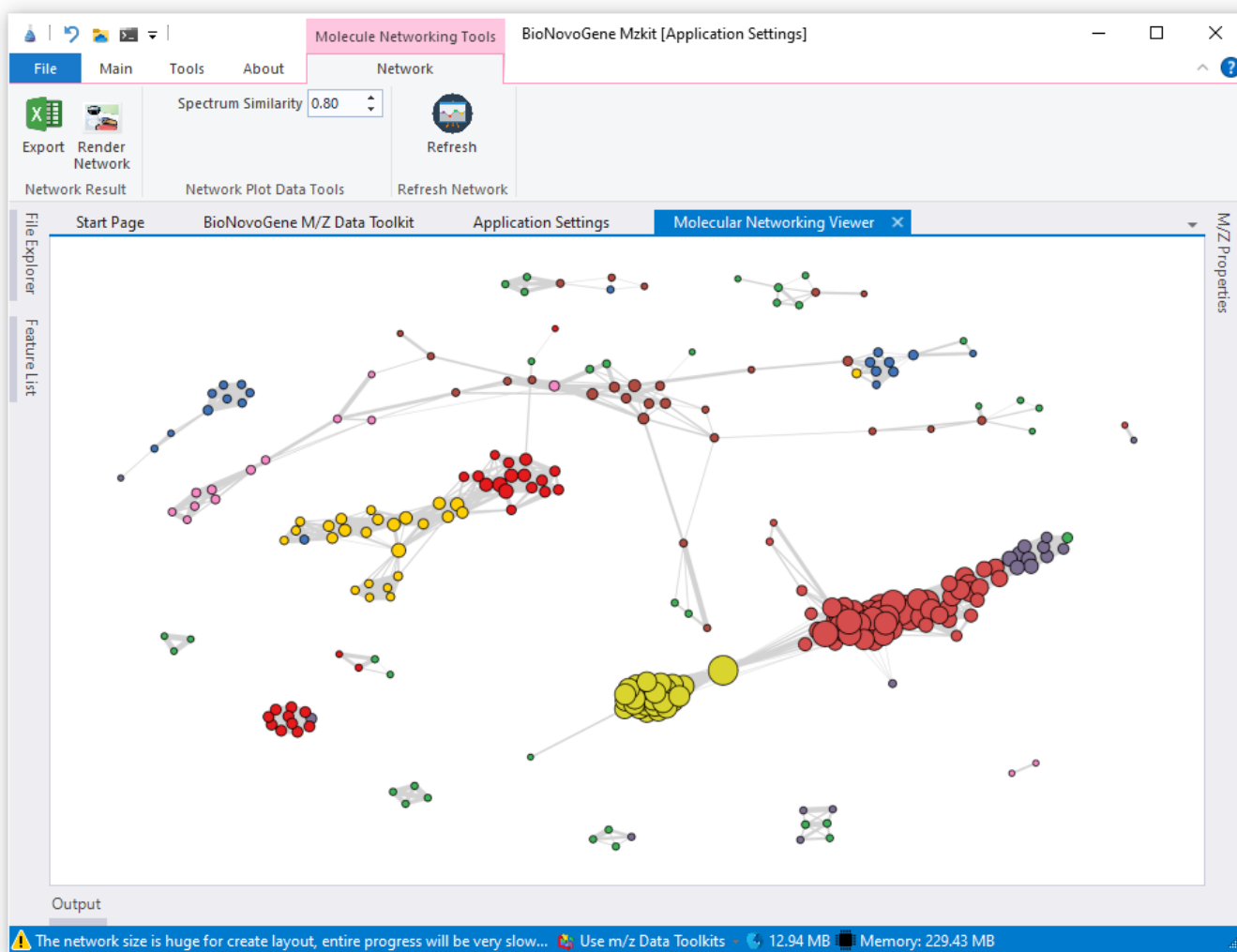
Output

The network size is huge for create layout, entire progress will be very slow... Use m/z Data Toolkits 266.72 MB Memory: 330.77 MB

2.2.4. step 4 network visualization

2.2.5. Export Network Data

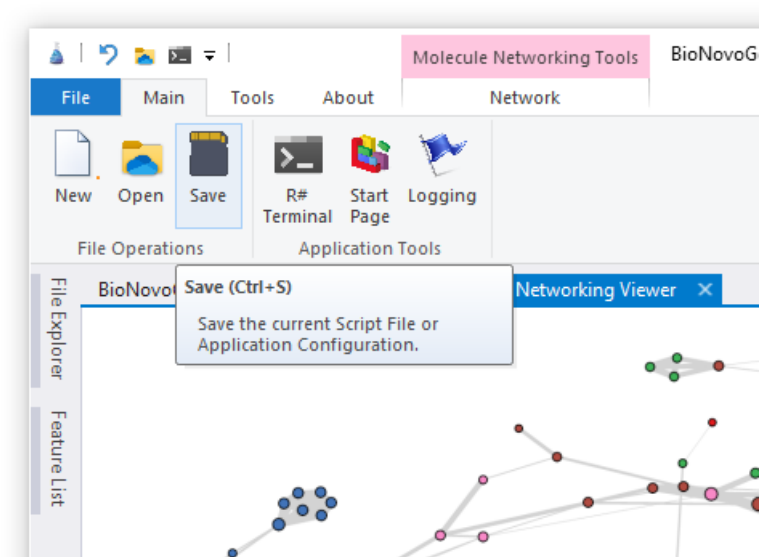
2.2.6. Save Network Visual



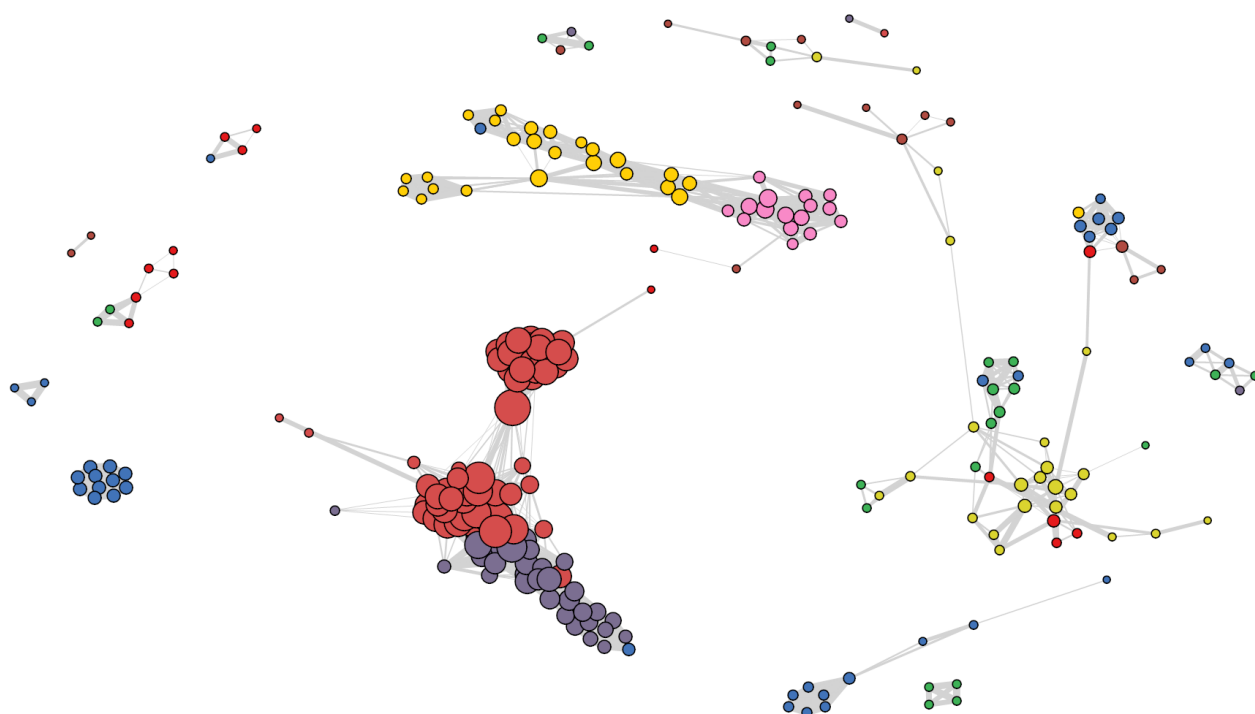
For export the network image, just mouse right click on the viewer panel, then you could see a popout menu **Save Image** that could use for viewer image saved.



You also can save the network image via **Save** command in the **Main** menu tab page when the molecular networking viewer is current active document page in mzkit program.



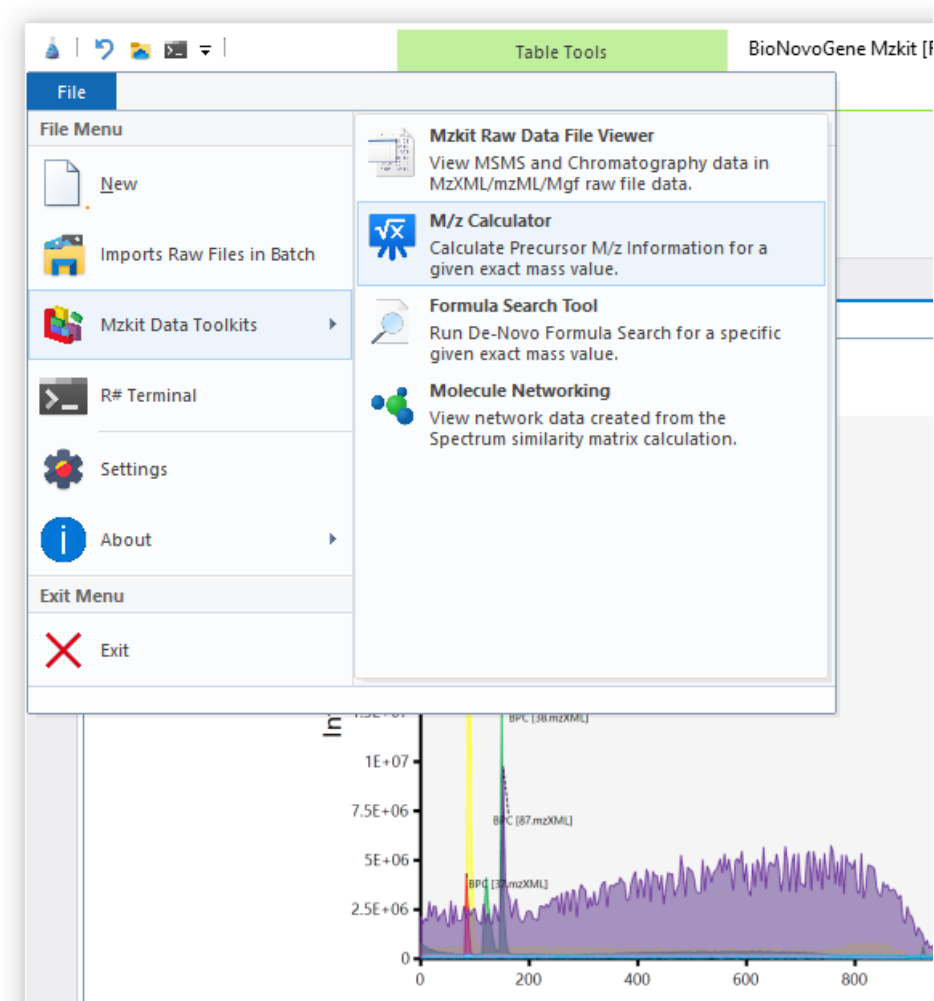
Now you can use the saved network image for your publications:



3. Appendix

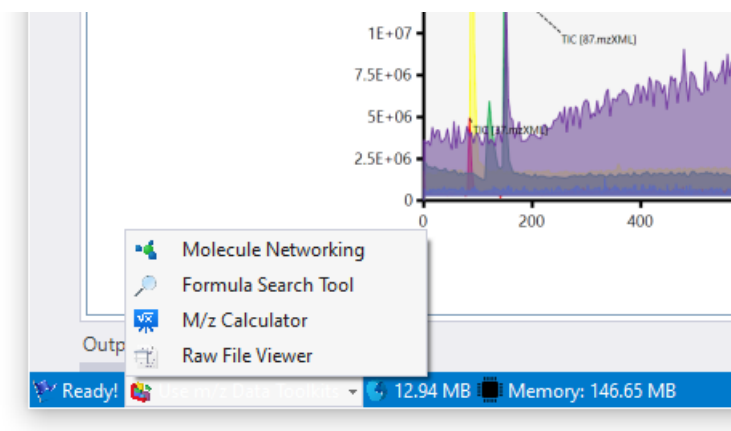
3.1. Switch Between Toolkit

You can switch between toolkit pages via the start menu of mzkit:

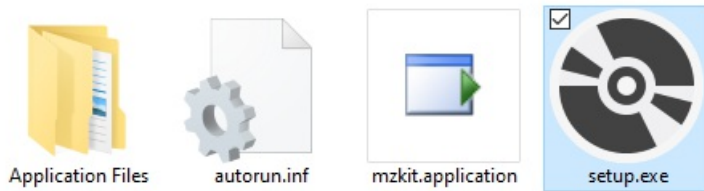


click **[File]** -> **[Mzkit Data Toolkits]**, and then click on one toolkit item then you can switch to the required toolkit page.

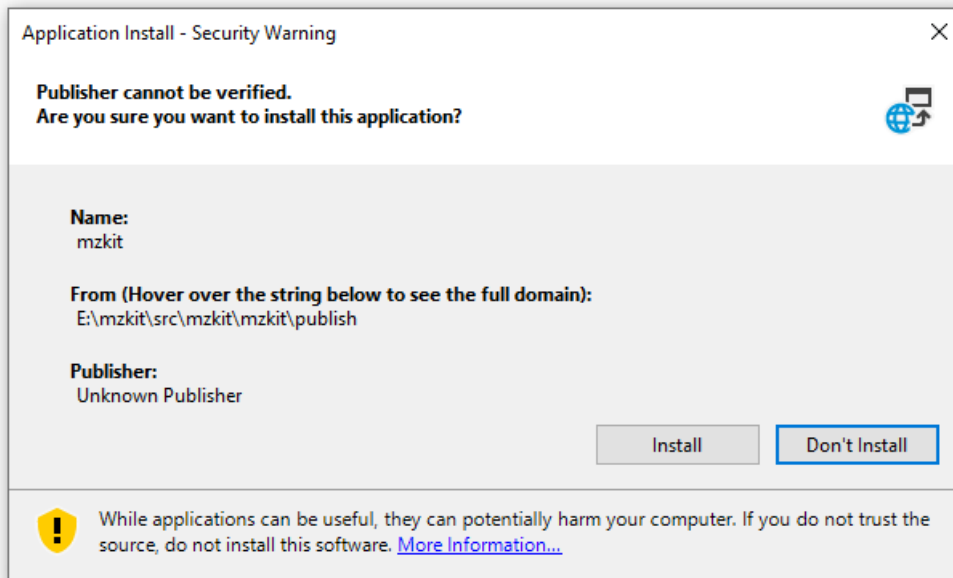
or just select a page from the app switcher toolstrip menu:



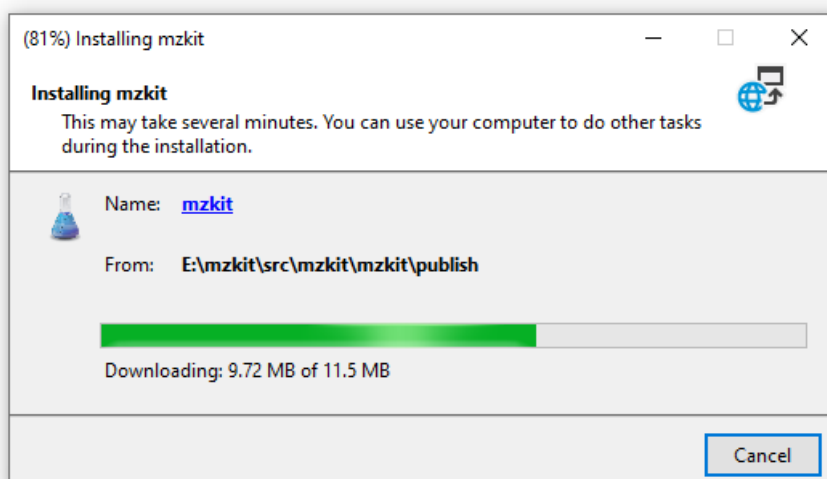
3.2. Install Mzkit



When you have extract the zip package of the mzkit compression package, then you could found a `setup.exe` install application in the top of folder. Click on this setup application, then your will going to install mzkit program into your computer system.



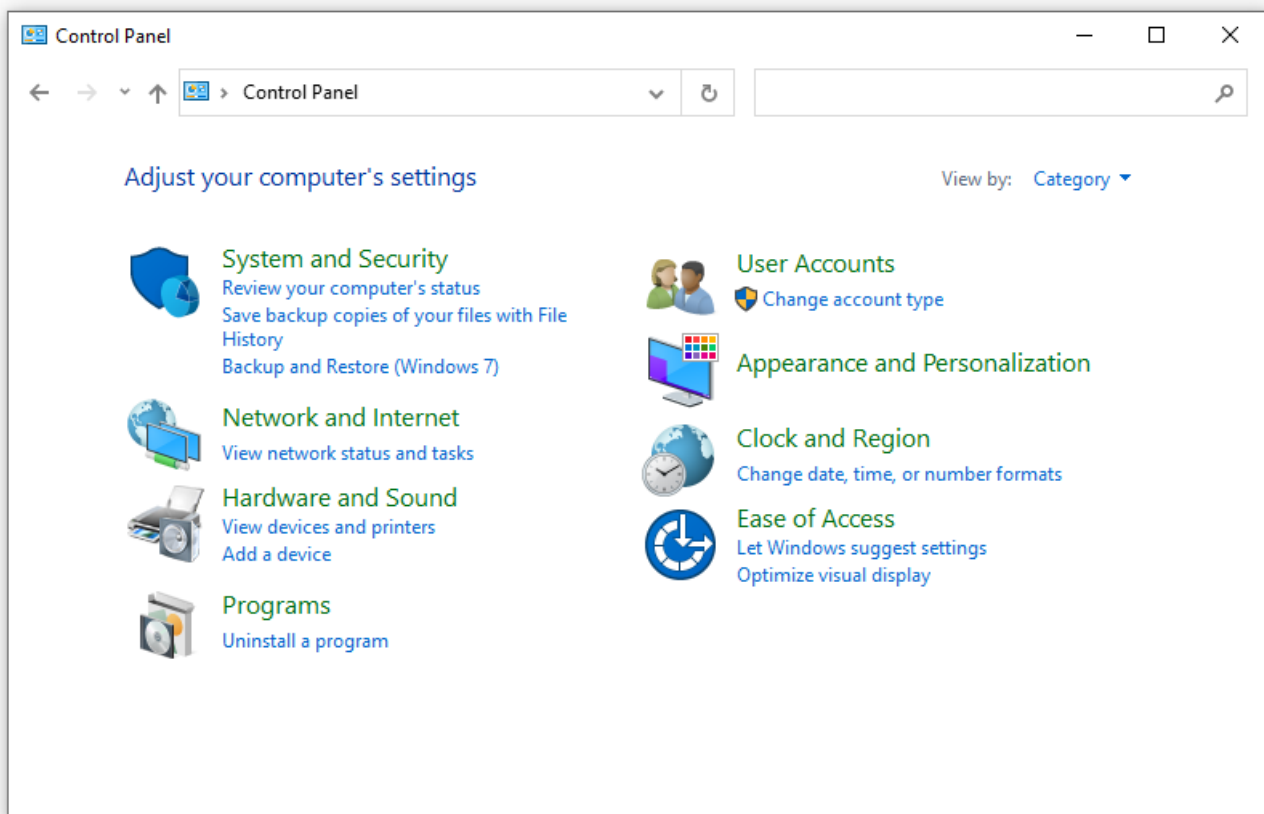
When you have launch the mzkit installer, then you will see a security warning dialog from your windows system, just click on the `Install` for start the install progress, and then just needs to wait for the installer finish the progress of copy the mzkit application files.



3.3. Uninstall Mzkit

if you want to removes mzkit from your computer system, then you could follow this instruction for uninstall:

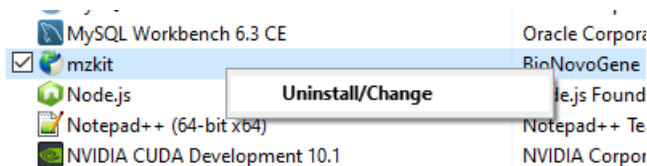
1. Open the control panel of your windows system, and then click of the link `uninstall a program`



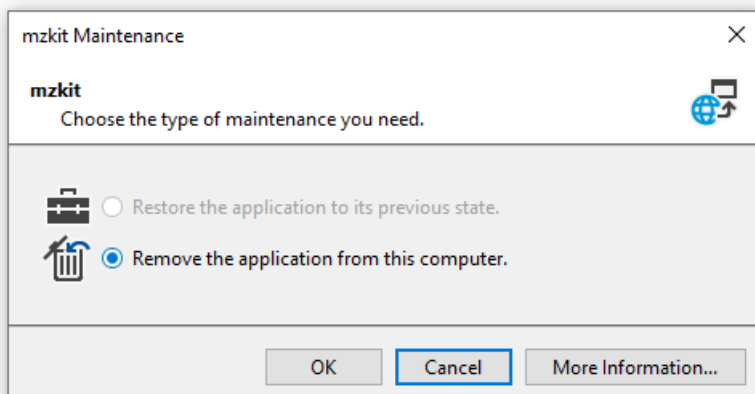
1. Then you will see a list of program that installed on your windows, located the mzkit application

| | | | | | |
|-------------------------------------|--------------------------------|--------------------|-----------|---------|-----------|
| | MySQL Installer - Community | Oracle Corporation | 5/12/2018 | 17.4 MB | 1.4.25.0 |
| | MySQL Server 8.0 | Oracle Corporation | 6/7/2019 | 392 MB | 8.0.11 |
| | MySQL Workbench 6.3 CE | Oracle Corporation | 5/12/2018 | 148 MB | 6.3.10 |
| <input checked="" type="checkbox"/> | mzkit | BioNovoGene | 9/8/2020 | | 1.12.1.16 |
| | Node.js | Node.js Foundation | 8/3/2018 | 52.8 MB | 8.11.3 |
| | Notepad++ (64-bit x64) | Notepad++ Team | 6/26/2020 | 14.3 MB | 7.8.5 |
| | NVIDIA CUDA Development 10.1 | NVIDIA Corporation | 7/12/2019 | | 10.1 |
| | NVIDIA CUDA Documentation 10.1 | NVIDIA Corporation | 7/12/2019 | | 10.1 |

1. Then right click on the mzkit, select `uninstall`, then you can removes mzkit from your windows system



Just click `OK` on the dialog:



4. Credits

This open source mass spectrometry data toolkit is developed at the [BioDeep](#) R&D laboratory and provided to you by [BioNovoGene](#) corporation.



[mzkit](#) is maintained by [xieguigang](#).

This page was generated by [GitHub Pages](#).