Metabolite Detector

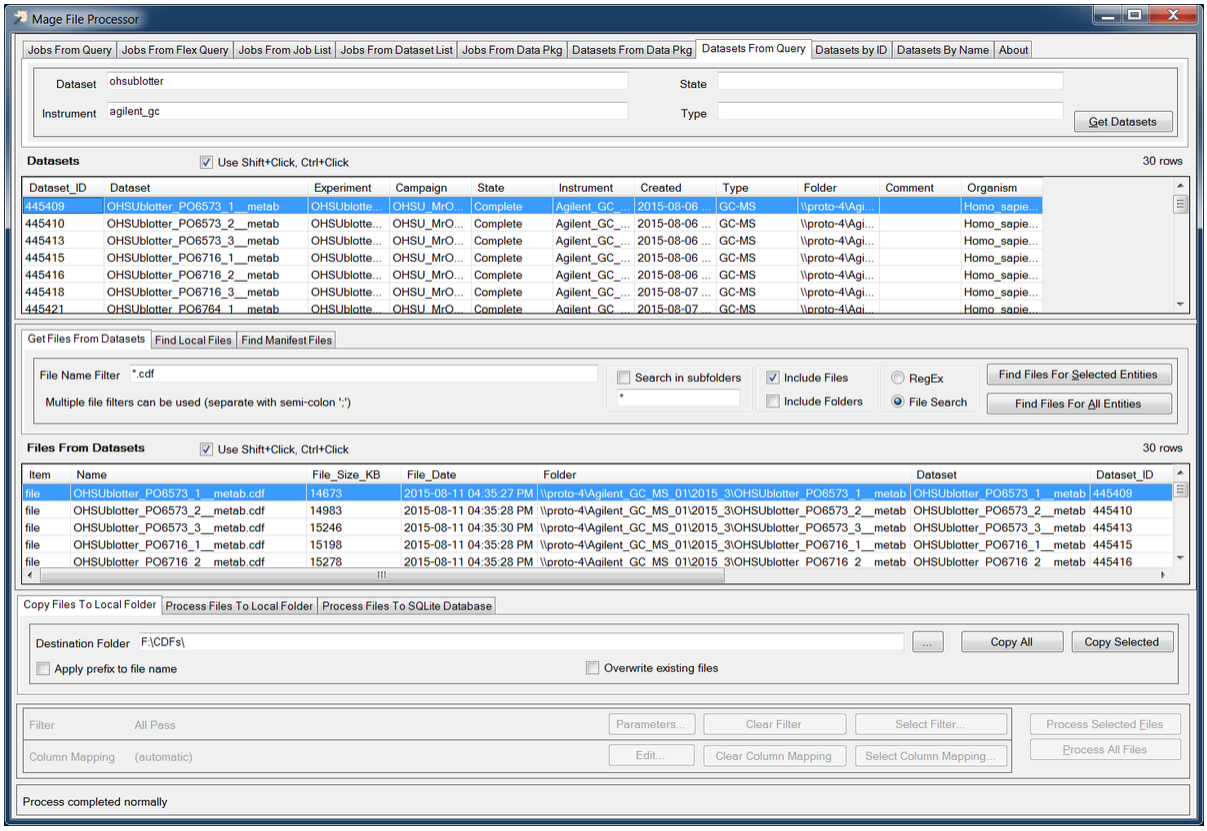
The Metabolite Detector software is developed by Karsten Hiller at the University of Luxemburg and is available for download at http://md.tu-bs.de/node/19

The following describes the steps required to process Agilent GC/MS Data using Metabolite Detector. The Agilent.D folders need to be converted to NetCDF files. Effective August 12, 2015, all new Agilent GC datasets will have .CDF files automatically created for them when they're captured into DMS. The conversion is accomplished using the command line interface for OpenChrom.

Contents

Manually creating CDF files

As mentioned above, DMS now auto-creates CDF files. Retrieve the CDF files using Mage File Processor, as illustrated in thisscreenshot.



The following steps describe methods for manually creating them.

Using Chemstation

1. Open a .D folder in ChemStation
2. Choose File->Export Data to AIA format
3. Select all the files
4. Click Process

Using OpenChrom

The OpenChrom software can batch-convert a large number of .D folders to create CDF files quite quickly

- Main site
  - http://www.openchrom.net/main/content/index.php
- Download site
  - http://sourceforge.net/projects/openchrom/
  - Or see openchrom_win32.win32.x86_64_1.0.0_rel.zip at `\\pnl\projects\OmicsSW\OpenChrom`
- Batch conversion guide
  - https://wiki.openchrom.net/index.php/Chromatogram_Batch_Format_Conversion
  - See below for step-by-step instructions

Installation

Unzip the contents to a local folder, e.g. C:\OpenChrom then run openchrom.exe

In order to read Data.MS files from .D folders, you need a plugin

- Inside OpenChrom choose Plugins, then Marketplace
- Next to find type: .ms
- Click Go, and you should see "Agilent MSD Import Converter". If not, look for it under the Popular tab
- Click "Install"
- Follow the prompts, and restart OpenChrom when prompted
- You need a free license key to use the converter
  - Go to https://marketplace.openchrom.net/
  - Create an account
  - Look for "Agilent MSD Import Converter" and click Get key
  - Copy the key and save to a local text file for reference
- Entering the key
  - The program should prompt you to enter the license key
  - If it does not, go to menu Window, then Preferences
  - Expand "A) Chrmomatography / Spectrometry"
  - Expand Converter
  - Expand Converter MSD
  - Click Agilent MS Conveter
  - Enter the key in the textbox below Full
  - Click OK

Batch Conversion

With OpenChrom running, follow these steps:

1. Create a folder on your desktop for the data
2. Use Mage File Processor to copy the .D folders from DMS
   - Start the program
   - Choose the Datasets from Query tab
   - Filter on dataset name and instrument, e.g. OHSU and Agilent_GC
   - In the File Name Filter textbox, enter: *.d
   - Uncheck "Includes files" but enable "Include Folders"
   - Enable "File search"
   - Click "Find Files for All Entities"
   - Enter the folder from step 1 into the "Destination Folder" checkbox
   - Click "Copy All"
   - The Project folder from step 1 will now have a bunch of .D folders
3. Start OpenChrom and choose MSD Perspective
4. On the Windows menu, Select View, then select "Project Explorer"
   - (or just click the "Project Explorer" tab on the left)
5. Right click the white blank space and choose "New Project"
6. Choose "MS Projects", "Analysis Project", then Next
7. Give the project a name, e.g. CDFCreator
8. Use the default location

- This location does not need to have the .D folders

9. Click Finish to finalize the creation of the project
10. Right click the new project and choose New, then Other, expand "Batch Process" and choose "Batch Process Job"
11. Click Next and your project should be selected, then click Next to use the default name for the job
12. The batch process editor should auto open; if not, double click the BatchJob.obj in the Project View
13. Click the Input Files tab, then click Add
14. Navigate to the folder where you copied the .D folders to locally, then select all of the .D folders
    - Left click the first one, then scroll to the end of the list and Shift+Click the last one
    - Click Finish
15. On the Output Files tab, click "Add"
16. Select output converter "ANDI/AIA CDF Chromatogram (*.CDF)"
    - For reference, other available output formats are CSV, PDF, and SVG
17. Define the output folder, then click Finish
18. Go back to the "Batch Process Job Editor" tab and choose "Run the batch process"
    - A .CDF file will be created for each .D folder

Batch Conversion from the Command Prompt

First get OpenChrom working with BatchConversion in the GUI, creating a .obj file (for example, BatchJob.obj). Run at the command line using:

```
c:
cd C:\DMS_Programs\OpenChrom\
openchrom.exe -cli -batchfile E:\DMS_WorkDir3\BatchJob.obj
```

Optionally use "-nosplash" to hide the splash screen. Unfortunately, if an error occurs, no error message is shown. Importantly, there cannot be any whitespace between the XML tags in the BatchJob.obj file.

Processing with Metabolite Detector

Make sure your .cdf files are all in one folder on your local computer, then start Metabolite Detector

Process the NetCDF files

1. Choose File, Import, Net CDF
2. Click the pale white letter icon (tooltip Net CDF Files)
3. Select your files and click OK, then click OK to load them
4. After the loading finishes, click “Close” to close the “Loading complete” window
5. Choose the library using file, Import, Import MSI-Library
6. Notice it has “1214 compounds”
7. Save a new home library into the same folder as the CDF files.
    - Do this by clicking “…” and saving the library as file “home_test.lbr” in the same folder as your CDF files.

Set options

Choose Tools, Settings, then under the Centroid tab define:

- Threshold begin: 10
- Peak threshold end: -5
- Maximbal baseline: 30
- FWHM: 0.1

On the Deconvolution tab:

- Peak threshold: 10
- Minimum peak height: 10
- Deconvolution width (scans): 8

On the Identification tab:

- Max RI difference: 20
- Cutoff score: 0.6
- Pure/Impure: 0.6
- Use scaled lib
- Use combined score

On the Quantification tab:

- Minimal distance: 0.5
- Minimal required quality index: 1
- Exclude 146.5 to 147.5
- Exclude 72.5 to 73.5

Run the Calibration Wizard

Choose Tools, then RI-Calibration Wizard

1. Choose dataset x_GCMS_FAMEs_01_GC-01
2. Leave option “use retention time” unchecked
3. Click next, then choose your compound library (the home library saved earlier)
4. Choose the following compounds (near the bottom of the list):
    - C8, C10, C12, C14, C16, C18, C20, C22, C24, C26, C28
5. Click Next, then the Calibration Table will appear
6. Assuming all are good, click Next to get to the Chromatogram Selection window
7. Choose all of the .CDF.bin files, but skip the FAMEs_01 one
8. On the next screen uncheck “Enable baseline adjustment”
9. Update the Deconvolution-width from 1 to 8. Possibly later re-process with deconvolution-width to 15
10. On the next screen click “Start”.

Quantification

Choose Tools, Batch Quantifcation

1. Click the green + next to Replicate Groups
    - Give the group a name, e.g. Test
    - Select all the files, except the Fames dataset, and click Open, then Ok, then Next

2. For Analysis Type, use Non-targeted Analysis
3. For Settings, under Compound Matching
    - ARI: 20
    - RI+Spec
    - Pur/Impure: 0.6
    - Req. Score: 0.6
4. For Settings, under Identification:
    - ARI: 20
    - RI+Spec
    - Pur/Impure: 0.6
5. Remaining settings
    - Compound Reproducibility: 0
    - Max peak discr. Index: 100
    - S/N: 15
    - # ions: 4
6. Enable option “Extended SIC Scan”
7. Click Finish

Export Results

1. In the results table, select the Batch Quantification tab and sort by Avg. RT
2. Export to CSV by clicking the left-most button on the Quantification Results tab
3. Click “Sums”
4. Enter a filename, e.g. Sum_DW8.csv

Also export the AllIons data

1. Export to CSV by clicking the left-most button on the Quantification Results tab
2. Click “All Ions”
3. Enter a filename, e.g. AllIon_DW8.csv

Validating Results

Import into Excel

Load the exported data into Excel.

1. Start Excel to view a blank workbook
2. On the Data ribbon, choose “From Text”
3. Select the CSV file and use the wizard to load the data

Linux Installation

Metabolite Detector 3.1 is only available as a Linux GUI app metabolitedetector.tu-bs.de (http://metabolitedetector.tu-bs.de/node/10). The binary is downloadable as a .deb package which is primarily

used by Debian and derivatives (e.g., Ubuntu), but it's a pretty bare package because it doesn't list dependencies. This package can be converted to a .rpm and installed on RedHat/Fedora/Centos, using the `alien` tool - see Package Conversion.

Detected dependencies include:

Library	Comment
libQtCore.so.4 (64-bit)	supplied by installing qt.x86_64, 1:4.8.7-2.el7
libQtGui.so.4 (64-bit)	supplied by installing qt-x11.x86_64, 1:4.8.7-2.el7
libQtSvg.so.4 (64-bit)	supplied by installing qt-x11.x86_64, 1:4.8.7-2.el7
libR.so (64-bit)	supplied by installing R-core.x86_64, 3.5.0-1.el7
libstdc++.so.6(CXXABI_1.3.8)(64bit)	supplied in part by compiling GCC 5.1.0 (see note below)
libstdc++.so.6(CXXABI_1.3.9)(64bit)	supplied in part by compiling GCC 5.1.0 (see note below)
libstdc++.so.6(GLIBCXX_3.4.20)(64bit)	supplied in part by compiling GCC 5.1.0 (see note below)
libstdc++.so.6(GLIBCXX_3.4.21)(64bit)	supplied in part by compiling GCC 5.1.0 (see note below)

Compiling GCC

The needed version(s) of the libstdc++ libraries are not directly available in shared library form from any RHEL7/epel repos.
These steps were used to compile the latest version of GCC on PrismWeb2 in August 2018.

- Download GCC version 5.1 source code from a mirror (see GCC version timeline (https://gcc.gnu.org/releases.html) and mirror listing (https://gcc.gnu.org/mirrors.html), go to `releases/` on a mirror to find source tarballs
- Extract the tarball using `tar -xvf gcc-*.tar.bz2`
- Install the needed dependencies for compiling: `yum install gmp-devel mpfr-devel libmpc-devel`
- CD into the source directory
- Run configure: `./configure --disable-multilib` (disable multilib, unless the program needs x86 libs)
- Run make: `make -j4`
- Find something else to do while it compiles...
- When complete, run `find . -name "libstdc++*"` to find where the compiled file is located.
- [Testing_conjecture] Copy the fully-versioned file to `/usr/lib64` (do NOT replace any files); change owner to root, and permissions to 755
- Move the old symbolic link of libstdc++.so.6 in /usr/lib64 to libstdc++.so.6_old
- Create a symbolic link for the new one: `ln -s libstdc++.so.6.0.xx libstdc++.so.6`

Manually Install Metabolite Detector

- Finish installing metabolite detector by adding the --nodeps flag to the rpm command. (fails, because it wants to replace "files")
- Manually extract all files from the RPM with their existing structure
- Copy all files to their respective desired places, creating folders and adjusting permissions as needed

Total list of files outside of a user directory and not part of an installed package created by "installing" Metabolite Detector:

File or Directory	Comments
/usr/lib64/libstdc++.so.6.0.21	
/usr/lib64/libstdc++.so.6	Symlink to libstdc++.so.6.0.21
/usr/lib64/libstdc++.so.6_old	Original symlink, to libstdc++.so.6.0.19
/usr/bin/metabolitedetector_app	
/usr/share/applications/metabolitedetector.desktop	
/usr/share/applications/tofcalib.desktop	
/usr/share/menu/	Directory
/usr/share/menu/metabolitedetector	
/usr/share/menu/tofcalib	
/usr/share/metabolitedetector	Directory
/usr/share/metabolitedetector/lib_aa.lbr	
/usr/share/metabolitedetector/RInside_0.2.13.tar.gz	
/usr/share/metabolitedetector/RInside/*	Extracted contents of RInside_0.2.13.tar.gz
/usr/share/pixmaps/oq.xpm	