GNPS Untargeted Metabolomics Workflow V.2

Sierra Simpson[1], Olivier George[1]
[1]University of California, San Diego

Version 2

Jan 14, 2021

1 *Works for me* This protocol is published without a DOI.

Sierra Simpson
University of California, San Diego

ABSTRACT

GNPS is an open-sourced, web-based mass spectrometry ecosystem that is used for community-wide organization / sharing of raw, processed, or annotated fragmentation mass spectrometry data (MS/MS).

GNPS aids in identification and discovery throughout the entire life cycle of data; from initial data acquisition/analysis to post publication. GNPS Website

This is a quickstart guide pulled directly from GNPS Website.

This version was intended as an easy to read step by step guide to:

1) Prepare data
2) Create a GNPS account
3) Upload to GNPS
4) Create a molecular network on GNPS

This can be found at GNPS.ucsd.edu with more in-depth information. This has been reproduced for use as a workflow within the George Lab for those unfamiliar with GNPS. More information and other useful tutorials can be found directly on the GNPS website and from Dorrestein publications listed below.

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

Mingxun Wang, Jeremy J. Carver, Vanessa V. Phelan, Laura M. Sanchez, Neha Garg, Yao Peng, Don Duy Nguyen et al. "Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking." Nature biotechnology 34, no. 8 (2016): 828. PMID: 27504778

PROTOCOL CITATION

Sierra Simpson, Olivier George 2021. GNPS Untargeted Metabolomics Workflow . **protocols.io** https://protocols.io/view/gnps-untargeted-metabolomics-workflow-bjbgkijw

MANUSCRIPT CITATION please remember to cite the following publication along with this protocol

Mingxun Wang, Jeremy J. Carver, Vanessa V. Phelan, Laura M. Sanchez, Neha Garg, Yao Peng, Don Duy Nguyen et al. "Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking." Nature biotechnology 34, no. 8 (2016): 828. PMID: 27504778

CREATED

Aug 04, 2020

LAST MODIFIED

Jan 14, 2021

PROTOCOL INTEGER ID

40008

DISCLAIMER:

> This is a quickstart guide was pulled directly from GNPS Website.
>
> Please see citation in the description for more information.
>
> This is solely meant as a tool to streamline the quickstart guide in a protocols.io format

ABSTRACT

GNPS is an open-sourced, web-based mass spectrometry ecosystem that is used for community-wide organization / sharing of raw, processed, or annotated fragmentation mass spectrometry data (MS/MS).

GNPS aids in identification and discovery throughout the entire life cycle of data; from initial data acquisition/analysis to post publication. GNPS Website

This is a quickstart guide pulled directly from GNPS Website.

This version was intended as an easy to read step by step guide to:

1) Prepare data
2) Create a GNPS account
3) Upload to GNPS
4) Create a molecular network on GNPS

This can be found at GNPS.ucsd.edu with more in-depth information. This has been reproduced for use as a workflow within the George Lab for those unfamiliar with GNPS. More information and other useful tutorials can be found directly on the GNPS website and from Dorrestein publications listed below.

Registration

1 Create a GNPS account at GNPS Signup

Upload Data

2 First convert specific vendor formats to open format using MSConvert locally. MSConvert Download

msConvert is a command-line utility for converting between various mass spectrometry data formats, including from raw data from several commercial companies (with vendor libraries, Windows-only). For Windows users, there is also a GUI, msConvertGUI.

For any difficulties converting the files - here is a FAQ posted by the Dorrestein lab: https://ccms-ucsd.github.io/GNPSDocumentation/fileconversion/

3 Upload converted files using an FTP client such as cyberduck or WinSCP (preffered). They do not like Filezilla as it has malware.

Host name: ccms-ftp01.ucsd.edu username: (yourGNPS login) password:(yourGNPS password)

Once files are uploaded they will be available for use within GNPS.

Molecular Networking Input

4 Navigate to the link below - it is the analysis webpage which contains the information for the study.

Analysis

5 Select from the presets either

- small dataset (5 LC/MS files)
- medium dataset (5-400 LC/MS files)
- large dataset (400+ LC/MS files).

6 Name your analysis under the "title section"

7 Select data input files (that you uploaded in step 3)

Within the Basic Options header; to input your spectrum files, select the input files tab next to the Spectrum files (required) field. A pop-up window with three tabs will appear; Select Input Files, Upload Files, Share Files. Proceed by selecting the files you want to analyze by clicking on the mass spectrometry file or an entire folder. Next click on the Spectrum Files G1 button to mark these files for analysis. You may select more files for analysis and put them in the separate groups (G1, G2, G3, G4, G5, G6) to reflect your experimental design. If you are performing an analysis with more than six groups, a metadata file may be used

8 For metadata creation see this link: Metadata FAQ

9 Once you are done selecting files to be analyzed, close the pop-up and hit submit.

View Analysis Results

10 To quickly glance at all the molecules in your data matching to the GNPS community knowledge of reference MS/MS spectra, click on "View All Library Hits".

11 To visualize the molecular networks in the browser, click the "View Spectral Families (In Browser Network Visualization)".

This brings you to a list of all the individual networks (connected components) that are in your data. Click on the "Visualize Network" link to explore each network individually. Your selection can be guided by the identifications within each network if there is a specific molecule of interest.

12 For advanced visualization see this link: https://ccms-ucsd.github.io/GNPSDocumentation/networking/