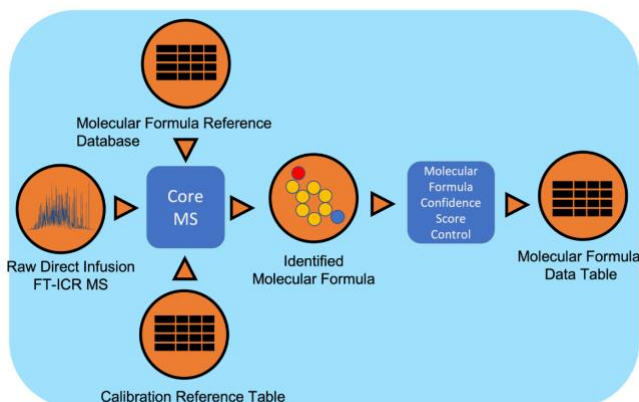


## Workflow de Materia Orgánica Natural (v4.1.5)



### Descripción General

Este workflow toma datos provenientes de espectrometría de masas Fourier Transform Ion Cyclotron Resonance (FTICR) recopilados de extractos orgánicos y determina las fórmulas moleculares de biomoléculas orgánicas naturales en la muestra.

### Ejecutando el Workflow

Actualmente, este workflow se puede ejecutar en [NMDC EDGE](#) o desde la línea de comandos. (Las instrucciones y requisitos de CLI se encuentran [aquí](#)).

### Input

El input para este workflow son los resultados de un experimento de espectrometría de masas (es decir, una “lista” de massSpec) que incluye mínimo dos columnas de datos para cada uno de los grupos: 1) la relación de masa:carga ( $m/z$ ) y 2) la intensidad de la señal. También se requiere un archivo de referencia para la calibración de fórmulas moleculares cuando se corre el workflow vía la línea de comandos (este archivo ya está implementado en el workflow a base web del NMDC EDGE).

- **Formatos de archivo aceptables:** .raw, .tsv, .csv, .xlsx

### Detalles

Datos de espectrometría de masa Direct Infusion Fourier Transform Ion Cyclotron Resonance (DI FTICR-MS) son procesados a base de señales y formulas moleculares asignadas utilizando el framework CoreMS del Laboratorio de ciencias ambientales y moleculares (EMSL). Muestras crudas de ‘time domain’ son transformadas a ‘ $m/z$  domain’ utilizando la transformación Fourier y la ecuación Ledford y los datos son limpiados (i.e., ‘denoised’), y se escogen picos representantes. Luego, se recalibran los datos utilizando una lista de referencia una que contiene fórmulas de moléculas conocidas, y además se hacen comparaciones con una biblioteca de fórmulas moleculares generada dinámicamente en un espacio de búsqueda definido. La puntuación de confianza para los candidatos de fórmulas moleculares se hace a base de exactitud de masa y estructura isotópica. La mayor puntuación es asignada al mejor candidato. Este workflow no funciona con la misma exactitud con datos de espectrometría de masa de Orbitrap.

### Versiones de Software

- CoreMS (2-clause BSD)
- Click (BSD 3-Clause “New” or “Revised” License)

## Productos

El archivo de resultados principal incluye una tabla de datos de fórmulas moleculares (en un archivo .csv).

### Archivos de Salida Primarios

### Descripción

INPUT_NAME.csv	masa:carga (m/z), Altura de Pico, Área de Pico, Identificación de Formulas Moleculares, Puntuación de Confianza, etc.
----------------	---

## Ejecutando el ‘Natural Organic Matter Workflow’ en NMDC EDGE

### Elige un workflow

1. Desde la categoría ‘Organic Matter’ (Materia Orgánica) en la barra de menú de la izquierda, elige ‘Run a Single Workflow’ (Ejecuta un workflow individual).
2. Un nombre de proyecto/ejecución **única** sin espacios (los guiones bajos están permitidos).
3. Una descripción es opcional, pero recomendada.
4. Elige ‘EnviroMS’ des del menú desplegable debajo de Workflow.

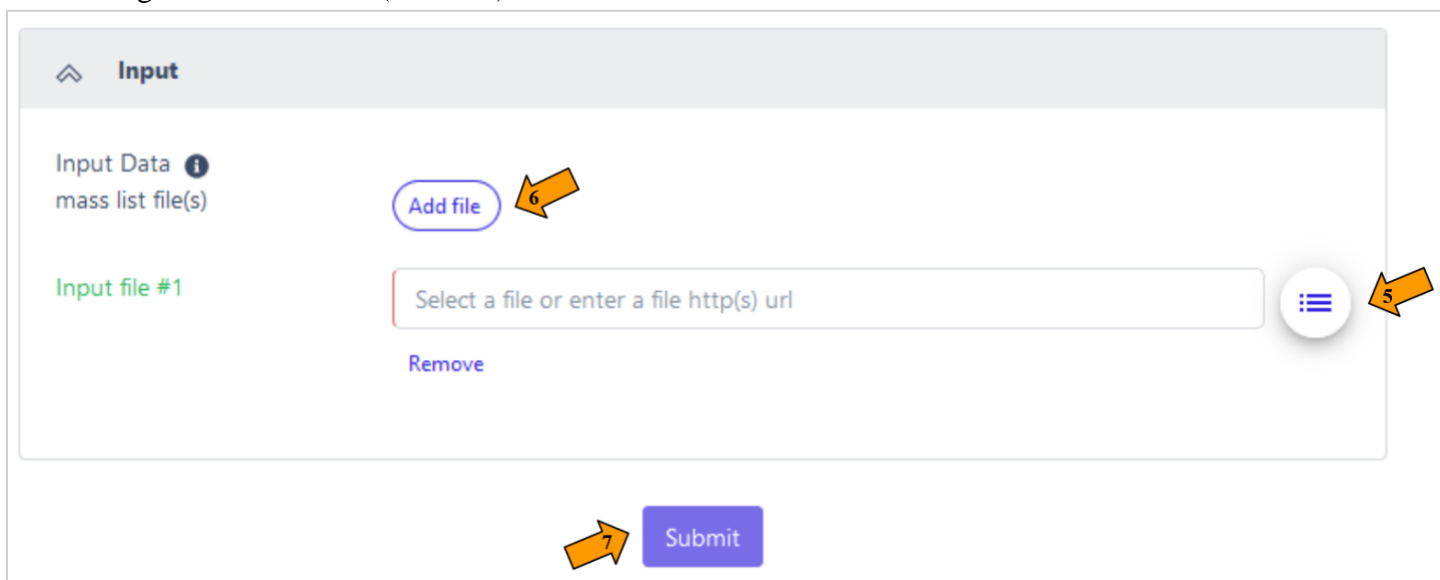
The screenshot shows the 'Run a Single Workflow' page in the NMDC EDGE application. The left sidebar contains a menu with categories: Home, Tutorials, Public Projects, Upload Files, NMDC (Sample Submission Portal, Data Portal), and WORKFLOWS. Under WORKFLOWS, 'Organic Matter' is expanded, and 'Run a Single Workflow' is highlighted with an orange arrow labeled '1'. The main content area has a header 'Organic Matter | Run Single Workflow' and a title 'Run a Single Workflow'. Below the title are two input fields: 'Project/Run Name' (required, 3-30 characters) and 'Description' (optional), with orange arrows labeled '2' and '3' pointing to them respectively. A 'Workflow' section contains a dropdown menu with 'Select a Workflow...' and a blue bar for 'EnviroMS', with an orange arrow labeled '4' pointing to it. A 'Submit' button is at the bottom right.

## Input

El workflow de Materia Natural Orgánica ('Natural Organic Matter') utiliza los resultados de un experimento de espectrometría de masas (es decir, una 'lista' de massSpec) que incluye como mínimo dos columnas de datos para cada uno de los grupos detectados: 1) la relación de masa:carga ( $m/z$ ) y 2) la intensidad de la señal.

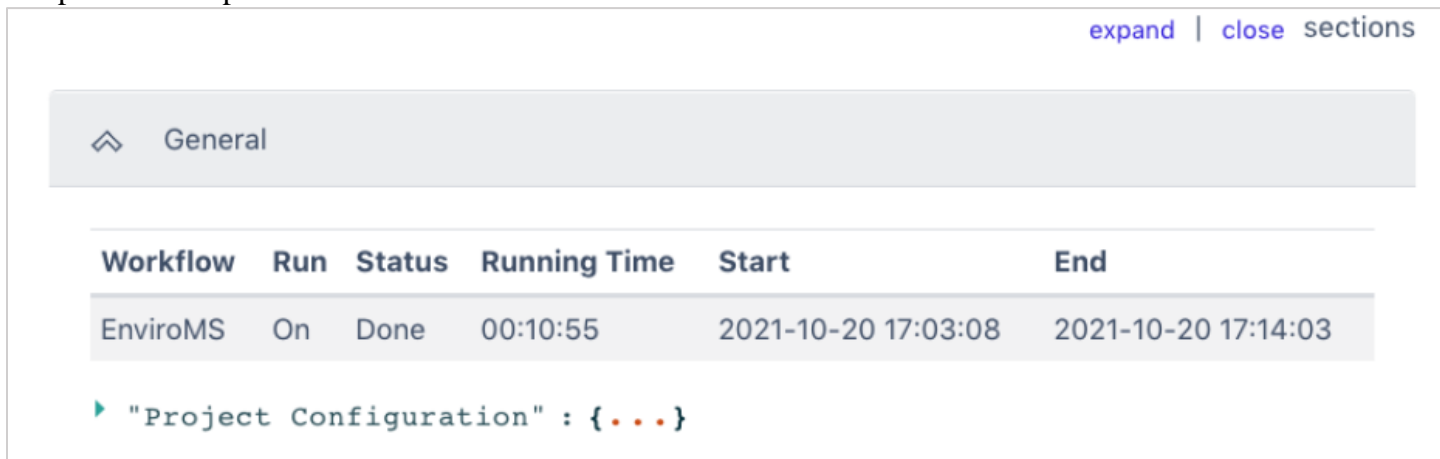
**Formatos de archivo aceptables:** .tsv, .csv, .raw, .xlsx

5. Haga clic en el botón a la derecha del espacio en blanco de 'entrada de datos' para seleccionar el archivo de datos para el análisis. (Si tiene archivos separados, hay dos espacios en blanco para ingresar sus datos). Se abrirá un cuadro llamado 'Select a File' (Seleccionar un archivo) para permitir al usuario encontrar el archivo deseado, archivos de proyectos ejecutados anteriormente, la carpeta de datos públicos o los archivos cargados por el usuario.
6. Se pueden agregar archivos de entrada adicionales haciendo clic en el botón 'Add a File' (Agregar archivo) para crear espacios en blanco de entrada adicionales.
7. Haga clic en 'Submit' (Someter).



## Productos

La sección General del resultado muestra qué workflow y qué herramientas se ejecutaron y la información del tiempo de ejecución. La Configuración del Proyecto (Project Configuration) se puede ver haciendo clic en los tres puntos entre paréntesis.



Workflow	Run	Status	Running Time	Start	End
EnviroMS	On	Done	00:10:55	2021-10-20 17:03:08	2021-10-20 17:14:03

► "Project Configuration" : { ... }

La sección 'Browser/Download Output' (Navegador/Descargar Productos) provee los archivos de resultados disponibles para descargar. Los archivos principales de los resultados son: 1) la tabla de datos de fórmulas

moleculares (archivo .csv) que contiene medidas de relación masa:carga (“m/z”), 2) la altura de picos detectados, 3) el área de picos detectados, 4) la identificación de fórmulas moleculares, 5) el tipo de ion detectado y 6) la puntuación de confianza para cada fórmula molecular.

Browser/Download Outputs		
File	Size	Last Modified
EnviroMS		
20190709_WK_CADY_Auto_S16_H1_Post_O5_1_01_36		
db_e_vs_c		
ms_class		
mz_error_class		
van_krevelen		
20190709_WK_CADY_Auto_S16_H1_Post_O5_1_01_36.csv	4.93 MB	6 days ago
20190709_WK_CADY_Auto_S16_H1_Post_O5_1_01_36.json	9 kB	6 days ago
assigned_unassigned.png	13 kB	6 days ago
mz_error.png	73 kB	6 days ago