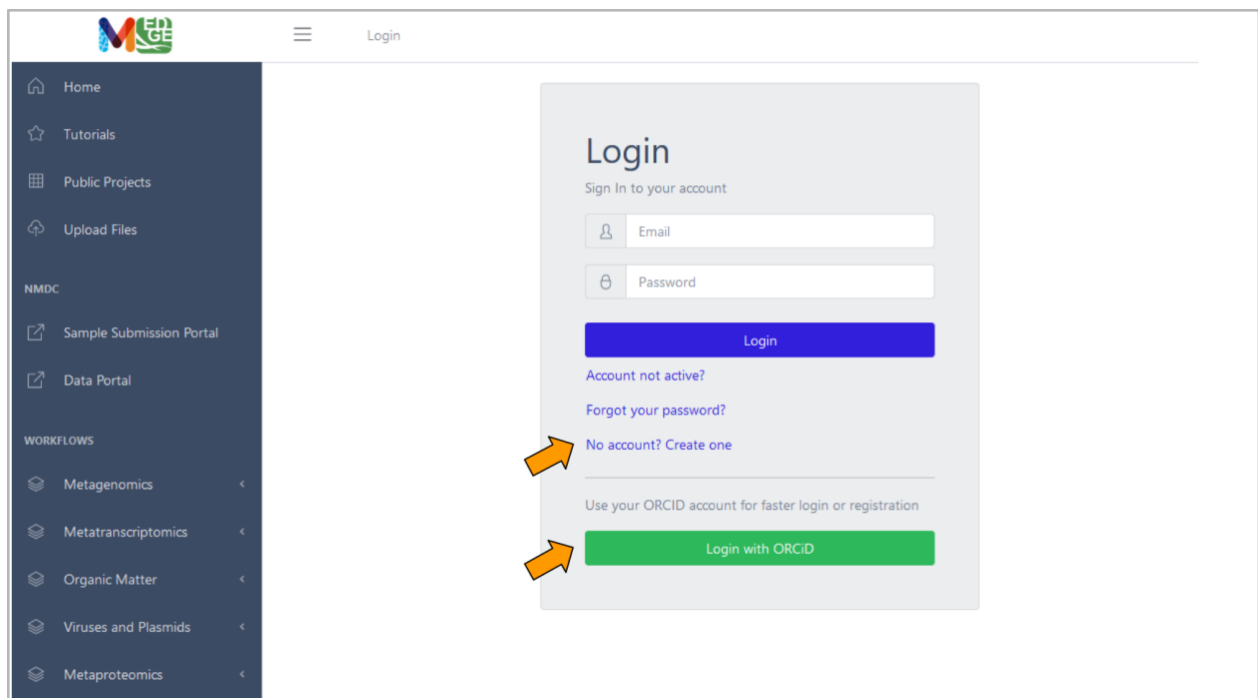
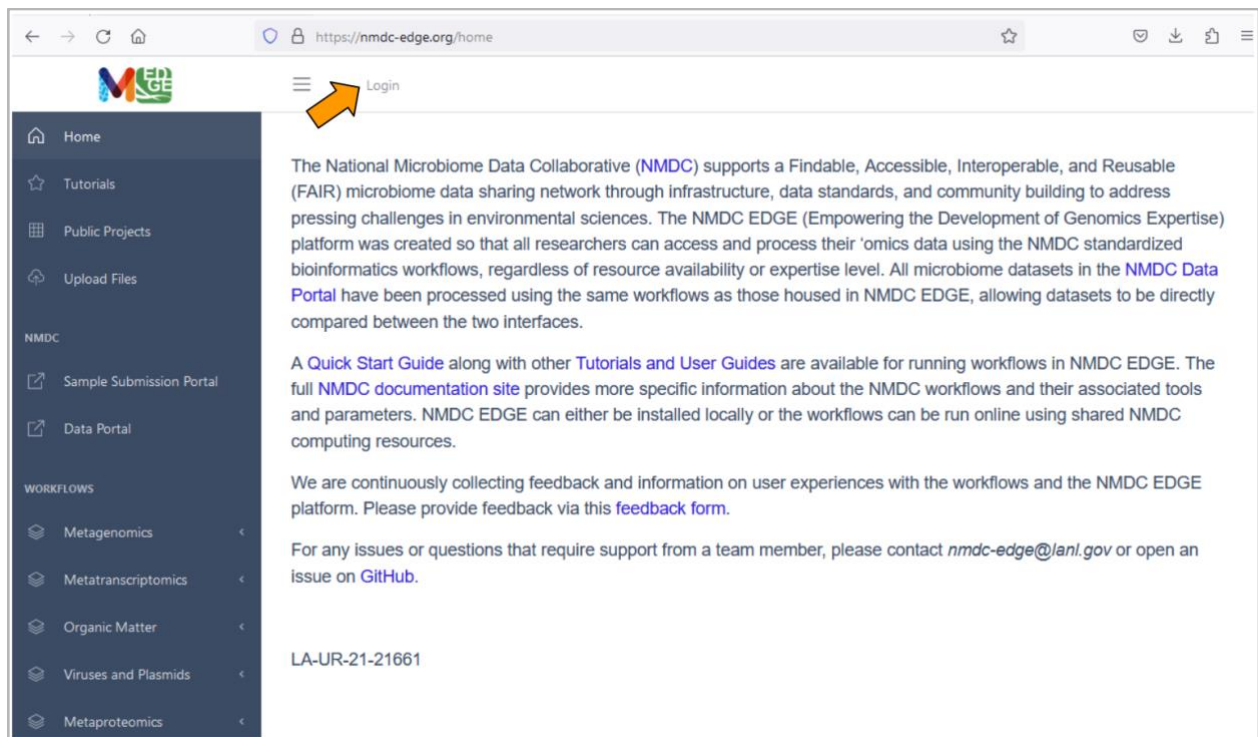


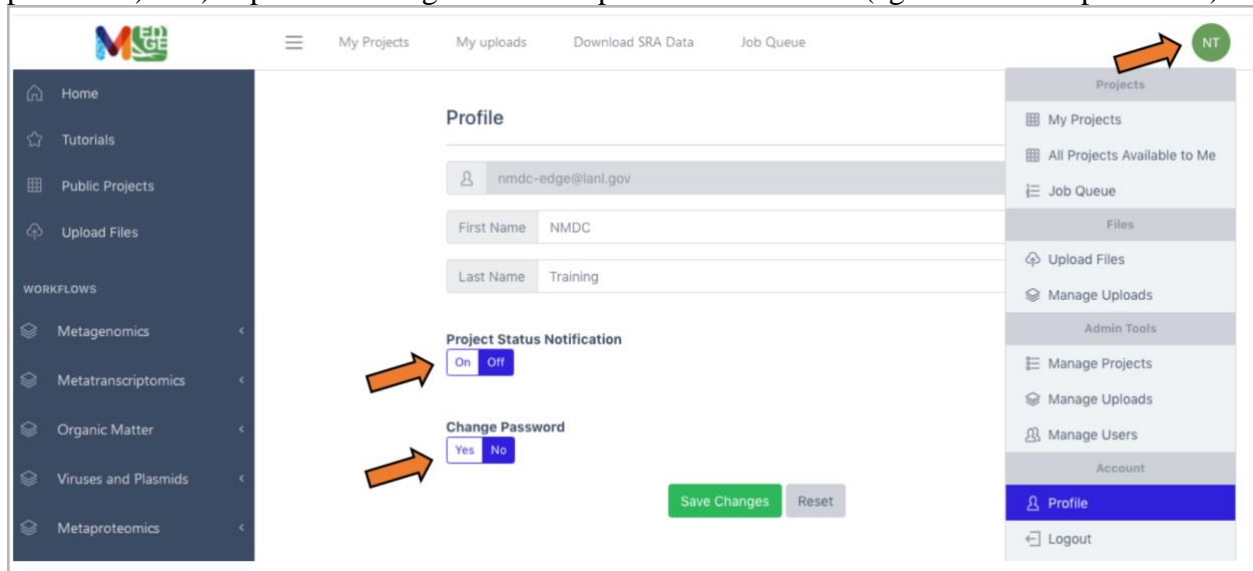
## Créer un compte

Les utilisateurs doivent créer un compte sur la plateforme NMDC EDGE ou se connecter en utilisant un compte ORCID.



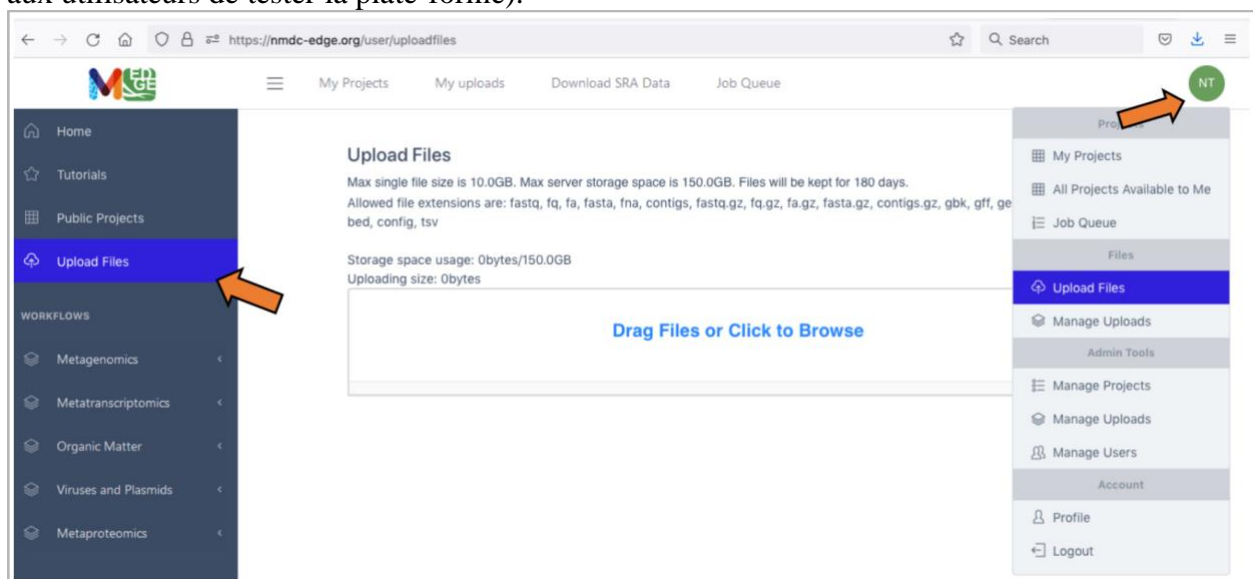
## Profil de l'utilisateur

Une fois connecté, le bouton vert avec les initiales de l'utilisateur sur la droite propose un menu déroulant qui permet à l'utilisateur de gérer ses projets et ses téléchargements ; il existe également un bouton qui permet aux utilisateurs de modifier leur profil. Sur cette page de profil, il y a deux options : 1) l'option de recevoir une notification par e-mail de l'état d'un projet (OFF par défaut) et 2) l'option de changer le mot de passe de l'utilisateur (également OFF par défaut).



## Télécharger des données

Deux options sont disponibles pour que les utilisateurs puissent télécharger leurs propres données à traiter via les workflow. La première consiste à utiliser le bouton dans la barre de menu de gauche. La seconde se fait via le menu déroulant affiché lorsque vous cliquez sur le bouton vert avec les initiales de l'utilisateur à droite. Les deux boutons ouvriront une fenêtre qui permettra à l'utilisateur de glisser-déposer des fichiers ou de rechercher les fichiers de données de l'utilisateur (il existe également des données dans le dossier Données publiques permettant aux utilisateurs de tester la plate-forme).



## Exécution d'un seul flux de travail

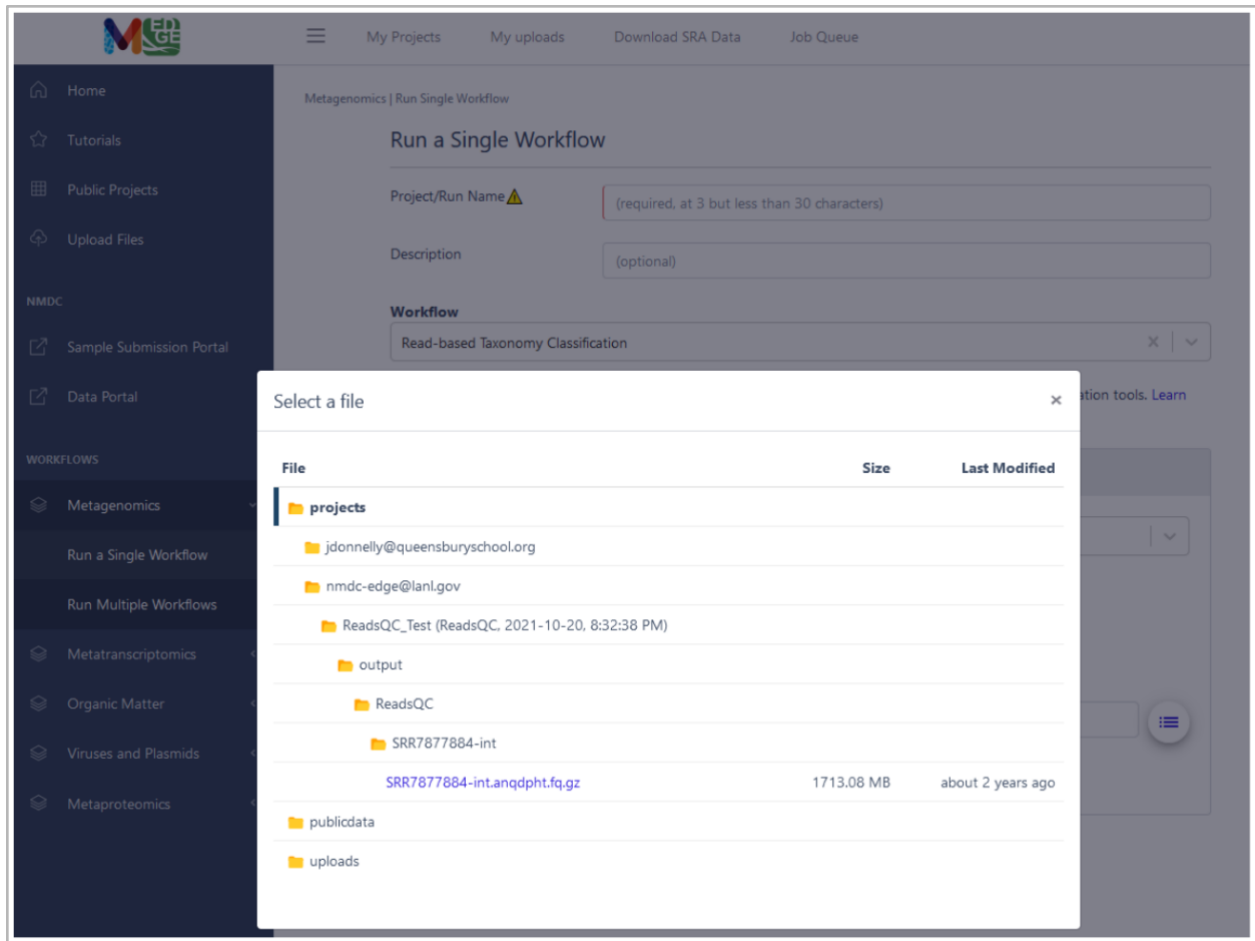
Pour exécuter un workflow, l'utilisateur doit fournir :

1. Un nom de projet/exécution unique sans espaces (les traits de soulignement sont acceptés).
2. Une description est facultative, mais utile.
3. L'utilisateur sélectionne ensuite le workflow souhaité dans le menu déroulant.
4. Pour les données métagénomiques/métatranscriptomiques, l'utilisateur doit également choisir si les fichiers d'entrée sont des fichiers entrelacés ou séparés pour les séquences (reads) appariées.
5. Ensuite, sélectionnez le(s) fichier(s) d'entrée dans la liste des fichiers disponibles.
6. L'utilisateur doit cliquer sur «Submit».

The screenshot shows the 'Run a Single Workflow' page in the Metagenomics section. The interface includes a sidebar with navigation options like Home, Tutorials, Public Projects, Upload Files, NMDC, Sample Submission Portal, Data Portal, and WORKFLOWS. The main content area is titled 'Run a Single Workflow' and contains the following elements:

- Project/Run Name:** A text input field with a warning icon and a note '(required, at 3 but less than 30 characters)'. An orange arrow labeled '1' points to this field.
- Description:** A text input field with a note '(optional)'. An orange arrow labeled '2' points to this field.
- Workflow:** A dropdown menu currently showing 'ReadsQC'. An orange arrow labeled '3' points to this dropdown.
- Workflow Description:** A paragraph stating: 'This workflow is a replicate of the QA protocol implemented at JGI for Illumina reads and use the program "rqfilter2" from BBTools(38:44) which implements them as a pipeline. [Learn more](#)'.
- Input Section:**
  - Input Raw Reads:** A section with a question 'Is interleaved?' and 'Yes'/'No' buttons. An orange arrow labeled '4' points to the 'No' button.
  - Input paired fastq:** A section with an 'Add paired-end fastq' button.
  - FASTQ Pairs:** Two input fields labeled 'Pair-1 FASTQ #1' and 'Pair-2 FASTQ #1', each with a placeholder 'Select a file or enter a file http(s) url'. To the right of each field is a circular button with a list icon. Orange arrows labeled '5' point to these buttons.
  - Remove:** A link to remove the input pair.
- Submit:** A purple button at the bottom right. An orange arrow labeled '6' points to this button.

Remarque : Cliquer sur les boutons à droite des espaces de saisie de données ouvre une boîte de dialogue appelée « Select a file » pour permettre à l'utilisateur de trouver les fichiers souhaités (affichés en violet) à partir de projets exécutés précédemment, du dossier de données public ou des données téléchargés par l'utilisateur.



## Exécution de plusieurs workflows

1. Une autre option consiste à sélectionner « Run Multiple Workflows » si l'utilisateur souhaite exécuter plusieurs workflow métagénomiques, ou l'intégralité du pipeline métagénomique.
2. Entrez un nom de projet/exécution unique sans espaces (les traits de soulignement sont acceptés).
3. Indiquer une description pour le projet est facultative, mais utile.
4. L'utilisateur doit également choisir si les données d'entrée sont des fichiers entrelacés ou séparés pour les reads appariées.

The screenshot shows the 'Run Multiple Workflows' page in the Metagenomics section. The interface includes a sidebar with navigation options and a main form area. Five orange arrows with numbers 1 through 5 point to specific elements:

- Arrow 1:** Points to the 'Run Multiple Workflows' option in the sidebar menu.
- Arrow 2:** Points to the 'Project/Run Name' input field, which has a warning icon and a note: '(required, at 3 but less than 30 characters)'.
- Arrow 3:** Points to the 'Description' input field, which has a note: '(optional)'.
- Arrow 4:** Points to the 'Is interleaved?' toggle switch, which is currently set to 'Yes'.
- Arrow 5:** Points to the 'Add interleaved fastq' button.

The 'Input Raw Reads' section contains the following elements:

- Input:** A section header with a dropdown arrow.
- Input Raw Reads:** A section with an information icon.
- Is interleaved?:** A toggle switch with 'Yes' and 'No' options.
- Input interleaved fastq:** A button labeled 'Add interleaved fastq'.
- interleaved FASTQ #1:** A text input field with a placeholder 'Select a file or enter a file http(s) url' and a 'Remove' button.

Les cinq workflow métagénomiques sont « activés » par défaut, mais l'utilisateur peut choisir de désactiver tous les workflows non souhaités. Le pipeline utilise la sortie de chaque workflow comme entrée pour les workflows suivants. (Remarque : certains workflows nécessitent des données d'entrée provenant de flux de travail antérieurs, donc la désactivation d'un flux de travail peut entraîner la désactivation automatique d'autres workflows). L'utilisateur peut ensuite cliquer sur « Submit ».

### Choose Workflows

All of the NMDC Metagenomic workflows are connected in EDGE, so that the output of one workflow can automatically be the input for the next workflow if several metagenomic workflows are selected.

The interface displays five workflow sections, each with a toggle switch set to 'On' and an orange arrow pointing to it:

- ReadsQC Workflow**: Toggle switch 'On'.
- Read-based Taxonomy Classification Workflow**: Toggle switch 'On'.
- Metagenome Assembly Workflow**: Toggle switch 'On'.
- Metagenome Annotation Workflow**: Toggle switch 'On'.
- Metagenome MAGs Workflow**: Toggle switch 'On'.

Below the workflows, there is a section for input files:

Select Analysis Tool(s): GOTCHA2 x Kraken2 x Centrifuge x

Input Map File ⓘ (Optional) Select a file or enter a file http(s) url ⓘ

Input Domain File ⓘ (Optional) Select a file or enter a file http(s) url ⓘ

At the bottom, there are two buttons: **Submit** and **Reset**, with an orange arrow pointing to the Submit button.

## Fichiers de sortie

1. Le lien « My Projects » ouvre la liste des projets pour cet utilisateur.
2. Des liens (dans les cercles violets) sont fournis pour partager des projets, rendre des projets publics ou supprimer des projets.
3. La colonne « Status » indique si le workflow est dans la file d'attente (gris), soumis (violet), en cours d'exécution (jaune), a échoué (rouge) ou terminé (vert). Si un projet échoue, un journal affichera les messages d'erreur pour le dépannage.
4. Cliquer sur l'icône à gauche du nom d'un projet ouvre la page de résultats de ce projet.

**My Projects**

Search

Drag headers here to group by

Project	Type	Status	Shared	Public	Created	Updated	Actions
<input type="checkbox"/> Project	Type	Status	Shared	Public	Created	Updated	Actions
<input type="checkbox"/> ReadsQC_Test2	ReadsQC	Submitted	No	No	10/19/2021, 8:37:20 AM	10/19/2021, 8:39:00 AM	
<input type="checkbox"/> Taxonomy_test	Read-based Taxonomy Classification	Running	No	No	10/19/2021, 8:34:32 AM	10/19/2021, 8:39:00 AM	
<input type="checkbox"/> ReadsQC_Test	ReadsQC	Complete	No	No	10/18/2021, 11:20:43 AM	10/18/2021, 1:22:00 PM	
<input type="checkbox"/> MAGs_Test	Metagenome MAGs	Complete	No	No	10/18/2021, 10:50:59 AM	10/18/2021, 11:55:00 AM	

## Résumé du projet (résultats)

La page de résumé du projet affichera trois catégories. Cliquer sur la barre ou l'onglet ouvre les informations.

1. « General » contient les informations sur l'exécution du projet.
2. « Workflow Result » contient la sortie tabulaire/visuelle consultable en ligne.
3. « Browser/Download Outputs » contient tous les fichiers de sortie disponibles pour le téléchargement. Il peut y avoir plusieurs dossiers.

**ReadsQC\_Test**

Project Summary:

Description: This is a test of the ReadsQC workflow for NMDC workflows training  
 Owner: nmcd-edge@lanl.gov  
 Submission Time: Mon Oct 18 2021 11:20:43 GMT-0600  
 Status: Complete  
 Type: ReadsQC

expand | close sections

- General
- ReadsQC Result
- Browser/Download Outputs

Cet exemple montre les résultats d'une exécution du workflow ReadsQC qui affiche le temps d'exécution sous l'onglet Général, les résultats du workflow de nettoyage et de sélection sur la base de la qualité des reads sous l'onglet « Results ReadsQC » et les fichiers disponibles au téléchargement (affichés en violet) sous l'onglet « Browser/Download Outputs ».

The screenshot displays the MG-RAST web interface. On the left is a dark sidebar with navigation links: Home, Tutorials, Public Projects, Upload Files, and a WORKFLOWS section containing Metagenomics, Metatranscriptomics, Organic Matter, Viruses and Plasmids, and Metaproteomics. The main content area has a top navigation bar with 'My Projects', 'My uploads', 'Download SRA Data', and 'Job Queue'. Below this, three orange arrows point to the 'General', 'ReadsQC Result', and 'Browser/Download Outputs' tabs respectively.

**General Tab:** Shows a table with workflow details.

Workflow	Run	Status	Running Time	Start	End
ReadsQC	On	Done	03:03:03	2021-10-20 20:33:14	2021-10-20 23:36:17

Below the table is a 'Project Configuration' section with a JSON-like structure: `{...}`.

**ReadsQC Result Tab:** Shows the input and results of the ReadsQC workflow.

Input:

Reads	Status
inputReads	44,943,418
inputBases	6,741,512,700
qtrimmedReads	8,583
qtrimmedBases	8,690
qfilteredReads	200,626
qfilteredBases	29,786,796
ktrimmedReads	6,186,690
ktrimmedBases	354,478,706
kfilteredReads	100,360
kfilteredBases	14,684,762
outputReads	33,510,668
outputBases	4,868,925,674
gcPolymerRatio	0.42

**Browser/Download Outputs Tab:** Shows a list of files and folders for download.

File	Size	Last Modified
ReadsQC		
SRR7877884-int		
filterStats2.txt	706 B	14 days ago
filterStats.json	337 B	14 days ago
filterStats.txt	287 B	14 days ago
SRR7877884-int.anqpht.fq.gz	1713.08 MB	14 days ago

Le pipeline métagénome complet ou les résultats d'exécution de « Workflow multiple » affichent les résultats de chaque workflow sous un onglet distinct et les fichiers associés disponibles au téléchargement se trouvent dans des dossiers de workflows séparés sous l'onglet « Browser/Download Outputs ».



The screenshot displays the NMDc-EDGE web application interface. On the left is a dark blue sidebar with navigation links: Home, Tutorials, Public Projects, Upload Files, and a WORKFLOWS section containing Metagenomics, Metatranscriptomics, Organic Matter, Viruses and Plasmids, and Metaproteomics. The main content area is titled 'Pipeline\_Test' and includes a 'Project Summary' section with the following details:

- Description:** This is a test of the NMDc metagenome pipeline for training purposes
- Owner:** nmdc-edge@lanl.gov
- Submission Time:** Wed Oct 13 2021 16:49:56 GMT-0600
- Status:** Complete
- Type:** Metagenome Pipeline

Below the summary is a list of workflow results, each with a heart icon and a section title. At the bottom, a table lists the output files:

File	Size	Last Modified
MetagenomeAnnotation		
MetagenomeAssembly		
MetagenomeMAGs		
ReadbasedAnalysis		
ReadsQC		

À titre de deuxième exemple, les deux figures suivantes montrent les résultats du workflow de classification taxonomique basée sur les reads. Le résumé comprend les reads classifiés et le nombre d'espèces identifiées pour tous les classificateurs taxonomiques sélectionnés. Les dix principaux organismes identifiés par chaque outil à trois niveaux taxonomiques sont également fournis. Les onglets de chacun des outils de classification fournissant des résultats plus approfondis se trouvent dans la section « Détails ». Des graphiques Krona sont générés pour les résultats à chacun des trois niveaux taxonomiques pour chacun des outils et sont également disponibles dans la section Détails. Les fichiers des résultats complets (au-delà du Top 10) et les graphiques sont disponibles en téléchargement.

## Read-based Taxonomy Classification Result

### Summary

Tool	Classified Reads	Species Reads	Species
gottcha2	89,222,937	89,222,937	9
centrifuge	14,874,315	14,485,925	5,127
kraken2	30,854,417	29,421,033	2,791

### Taxonomy Top 10

Species Genus Family

Tool	Level	Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	Top9	T
gottcha2	species	Pseudomonas aeruginosa	Salmonella enterica	Listeria monocytogenes	Enterococcus faecalis	Lactobacillus fermentum	Bacillus subtilis	Escherichia coli	Staphylococcus aureus	Listeria phage A500	
centrifuge	species	Pseudomonas aeruginosa	Enterococcus faecalis	Bacillus subtilis	Bacillus intestinalis	Listeria monocytogenes	Lactobacillus fermentum	Pseudomonas fluorescens	Pseudomonas sp. AK6U	Salmonella enterica	E
kraken2	species	Pseudomonas aeruginosa	Salmonella enterica	Bacillus subtilis	Listeria monocytogenes	Enterococcus faecalis	Lactobacillus fermentum	Escherichia coli	Staphylococcus aureus	Homo sapiens	B c

### Detail

centrifuge gottcha2 kraken2

Species Genus Family

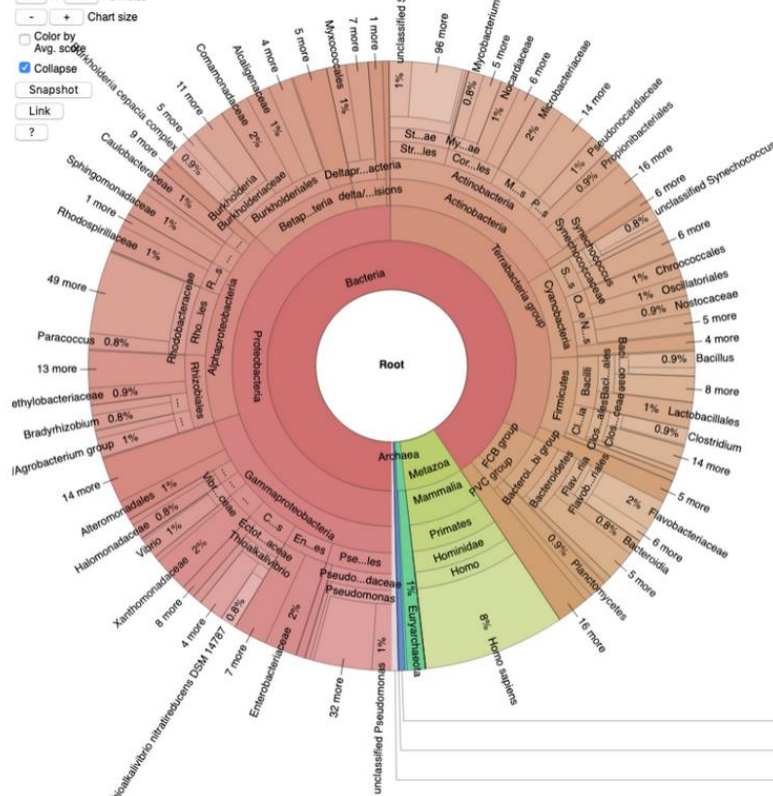
Level	Taxonomy	Reads	Abundance
species	Pseudomonas aeruginosa	7,026,567	0.001
species	Enterococcus faecalis	5,840,658	0.002
species	Bacillus subtilis	5,224,145	0
species	Bacillus intestinalis	5,080,010	0.002
species	Listeria monocytogenes	4,963,265	0.002
species	Lactobacillus fermentum	4,223,501	0.002
species	Pseudomonas fluorescens	2,735,891	0
species	Pseudomonas sp. AK6U	2,685,240	0
species	Salmonella enterica	2,353,819	0
species	Escherichia coli	776,322	0

Krona

Search:

- 30 + Max depth  
- 11 + Font size  
- + Chart size  
☐ Color by Avg. score  
☒ Collapse  
Snapshot  
Link  
?

Root  
Magnitude: 1190936  
Unassigned magnitude: 4485  
Count: 9908  
Unassigned: 37



[\[Krona full window view\]](#)

Browser/Download Outputs