

---

# **aureme Documentation**

**Meziane AITE**

**Oct 26, 2020**



---

# Contents

---

<b>1</b>	<b>Contents</b>	<b>1</b>
1.1	How to use Docker? . . . . .	1
1.2	How to use the AuReMe workspace (default workflow) . . . . .	3
1.3	FAQ . . . . .	15
1.4	Indices and tables . . . . .	35



## 1.1 How to use Docker?

### Requirements Docker (v 1.10 at least)

To install Docker, please follow the instructions on [docker.com](https://docs.docker.com), considering your operating system\*.

- On Mac OS: requires at least Yosemite 10.10.4
- On Windows: requires at least Windows 8

### 1.1.1 Running a Docker container

1. Launch the Docker machine (see the instruction on [docker.com](https://docs.docker.com)). For example:
  - On Fedora: **sudo systemctl start docker**
  - On Mac OS and Windows: run the Docker launcher
2. Download the AuReMe Docker image with the command described just below. You will find the current **tag** of the AuReMe Docker image on this [page](#).

```
shell> docker pull dyliss/aureme-img:tag
```

3. To verify that the image has been downloaded correctly, check it in the list of your local images:

```
shell> docker images -a
```

REPOSITORY	TAG	IMAGE ID	CREATED	SIZE
docker.io/dyliss/dyliss/aureme-img	2.2	6cf38ab4edc8	1 hour ago	1.68 GB

4. Create a folder that will serve as a bridge to share data from/to the Docker container. Let us call it **bridge** for instance.
5. Create a Docker container from the following image with this command:

```
shell> docker run -ti -v /my/path/to/the/directory:/shared --name="aureme_cont" \
↳ dyliss/aureme-img:tag bash
```

The path given for `-v` is the one to the shared directory on your host machine. **This path has to end on the directory name** (without any `/` at the end). **The path has to be complete** (from `/users` or from `C:\` for Windows users). After the `:` is the name of the mirror directory in the Docker container. Please do not change it.

For Windows users, be careful, you have to indicate your path this way:

You have just made a bridge between `\my\path\to\the\directory\bridge` and the container `aureme_cont`.

You can create as many containers as you wish, as long as you give them different names.

Your AuReMe container is now running and correctly installed. Congratulations! You are now inside the container `aureme_cont`.

### 1.1.2 Some tips about Docker

- To exit the container, type **exit**.

```
aureme> exit
```

- To list all your containers:

```
shell> docker ps -a
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS
↳ PORTS	NAMES	SIZE		
fc969ed0d2c7	docker.io/dyliss/aureme-img	"bash"	2 hours ago	Up 5 hours
↳ aureme_cont	11 MB (virtual 3.52 GB)			

Remark that you can see, with this command, the state of your containers in the STATUS column: **up** (running, you can connect to it), **exited** (stopped, need to be started again).

- To start or stop the container (from your host):

```
shell> docker start aureme_cont
shell> docker stop aureme_cont
```

- If you want to go inside a **started/running** container:

```
shell> docker exec -it aureme_cont bash
```

- To delete a container: **docker rm container\_id** ( or **name**)
- To delete an image: **docker rmi image\_id** (or **name**)

Before deleting a Docker image, you have to delete all the Docker containers which are linked with the image you would like to remove. And before removing a Docker container, you have to stop it.

AuReMe documentation

## 1.2 How to use the AuReMe workspace (default workflow)

AuReMe is deployed in a Docker image. Thanks to this Docker image, all the tools inside the AuReMe workspace are ready to use inside the AuReMe container.

### 1.2.1 Requirements

1. Create your Docker container as explained in the previous step *Running a Docker container*, start the container and go inside.

#### Start a new study

2. Use the following command to start a new study. Choose an identifier for this study (ex: replace **species** by your organism name). In order to illustrate this documentation, we will use **species** as a run identifier.

```
aureme> aureme --init=species
```

Now you will find on your own computer (host), in your **bridge** directory, a folder **species** with many subdirectory and files. This is your work directory, on which AuReMe is going to run.

**Note:** Notice that by default all the outputs of commands will be printed in the terminal. Nevertheless, if you want to trace all your commands just type:

```
aureme> aureme --run=species --cmd="some_command" -q
```

Thanks to the **quiet** option, the command will be stored as a log in the **bridge > species > log.txt** file. The output of this command will also be stored in the **bridge > species > full\_log.txt** file.

For further details on the log files, please see the *How to manage the log files?* chapter.

3. To get an overview of AuReMe, you can get a sample by using this command.

```
aureme> aureme --run=species --sample
```

#### Define the reference database

4. The final step is to define which reference database to use. The available databases are listed in your terminal when you create a new study. If needed, use this command to display them again.

```
aureme> aureme --run=species --cmd="getdb"
Running command: getdb in species
Available database in Aureme:
/home/data/database/MODELSEED/modelSeed
/home/data/database/BIGG/bigg
/home/data/database/BIOCYC/METACYC/23.0/metacyc_23.0
```

**This reference database is needed to:**

- Be able to match all the identifiers of the entities of metabolic networks.
- Gap-fill the metabolic network in the gap-filling step.
- Uniforms the data in one unique database.

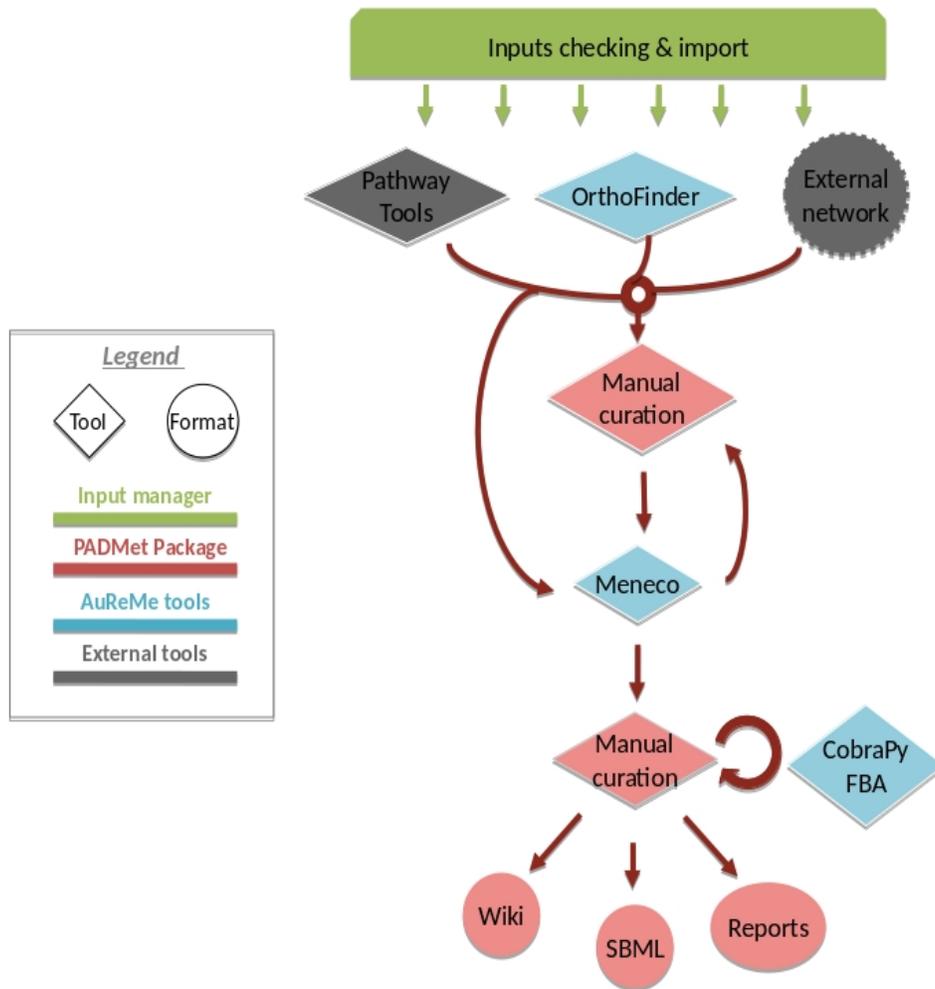
To select one of the above databases, replace the corresponding path in the configuration file: **config.txt**, in the **DATA\_BASE** variable, or comment the line if you don't want/can't use a database.

The **config.txt** file is stored at the root of your **species** folder.

### 1.2.2 The default workflow

By default, the AuReMe workspace includes an automatic workflow for metabolic network reconstruction. This workflow runs several pre-installed tools and generates diverse output files. The process can be either run entirely in a single command, or run step by step to personalize it or do some intermediary analysis.

For instance, if you run the **draft** command (see *Merge metabolic networks*), it will run all the previous steps automatically as described in the following figure. This figure details the steps of the default workflow.



### 1.2.3 Data organization

#### Bridge structure

The **bridge** directory will store all your input data you will provide, and all the result files the workflow is going to create. In this section, all the **bridge** sub-directories will be described.

**analysis:** All output files of the analysis processes.

**annotation\_based\_reconstruction:** If you want to use annotated genomes (to run the annotation-based reconstruction part of the workflow), put here all the output files of the annotation tool. For instance with Tools, copy-paste the whole PGDB (*Annotation-based reconstruction* for more details).

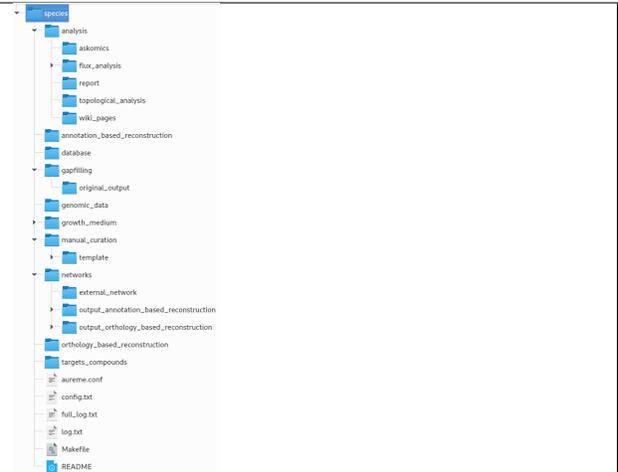
**database:** If you want to use your own database put in this folder your database in Padmet format, if you have a SBML convert this file to Padmet (see *How to convert files to different formats?*). Don't forget to update the **config.txt** file after transforming your database into the Padmet format.

**gapfilling/original\_output:** If you run the metabolic network reconstruction with gap-filling, will contain all the output files of gap-filling tools before any post-process from AuReMe (see the *Gap-filling* section).

**genomic-data:** The directory in which to put the genomic data on your studied organism, that is to say either a Genbank (**species.gb**) or a proteome (**species.faa**).

**growth\_medium:** Description of the set of metabolites that is available to initiate the metabolism (growth medium), that is to say the seed compounds (**seeds.txt** and **artefacts.txt**), see *How to manage growth medium?* and *What are "artefacts"?*.

**growth\_medium:** Description of the set of metabolites that is available to initiate the metabolism (growth medium), that is to say the seed compounds (**seeds.txt** and **artefacts.txt**), see *How to manage growth medium?*



directory (if you have run this part of the workflow).

**orthology\_based\_reconstruction:** If you want to use model organisms (to run orthology-based reconstruction part of the workflow), put in this directory, a subdirectory named **species** the proteome (**species.faa** or **species.gb**) and the metabolic network (**species.sbml**) of your model (see below *Orthology-based reconstruction* for more details).

**targets\_compounds:** Description of the set target compounds (**targets.txt**), that is to say metabolites whose production is supposed to be achieved by the metabolism of the species under study (components of the biomass reaction or other metabolites), see the *Gap-filling* paragraph.

## Provide input files

First of all, you have to provide to AuReMe all the input files needed for the different steps you want to run in the workflow. The steps can be optional or run several times, at different phases of the process. However, you have to store the input data for each steps, observing the architecture described above for the **bridge** directory (see *Bridge structure* section).

Here is the list of inputs you have to provide to run the pre-set default workflow. If you want to run only part of it, please see the corresponding sections and the chapter *How to create a new 'à-la-carte' workflow?*

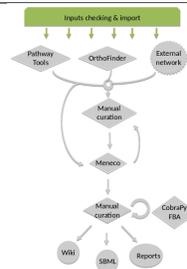
- See *Orthology-based inputs*.
- See *Annotation-based inputs*.
- **External source for reconstruction** If you already have one or several external metabolic networks for your studied species and you want to improve them, just copy-paste them (SBML format) in the **networks > external\_network** folder.

```

/species
|-- networks
    |-- external_network
        |-- first_manual_network.sbml
        |-- second_manual_network.sbml
        |-- ...
    
```

## Check input files validity

This will verify the format and consistency of your data for a better quality result. Moreover, it will generate all the supplementary files needed by the workflow tools and put them into the corresponding folders. For more information about input files validity see *What is checked in my input files?*



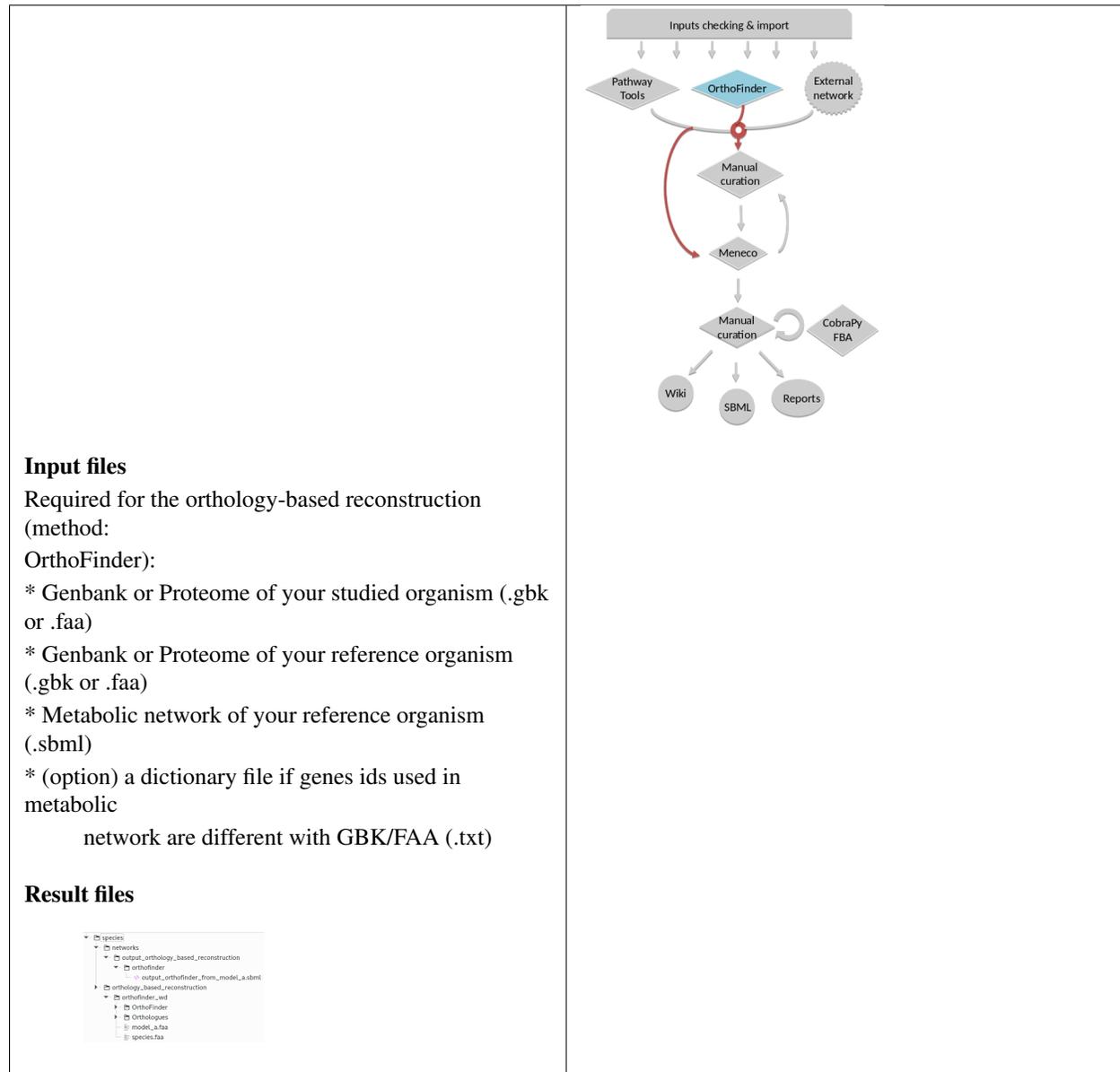
For this purpose, use this command:

```
aureme> aureme --run=species --cmd="check_input"
```

**Warning:** Always check the validity of the inputs before running any workflow task, and after having put every input files needed for the steps of the workflow.

## 1.2.4 Orthology-based reconstruction

### Method: OrthoFinder



### Orthology-based inputs

1. Put all the available genomic data of the studied organism in the folder **genomic\_data**, either a Genbank (.gbk) or a Fasta (.faa) file. Please **give to these files the same name the bridge directory has**. Here, the bridge

directory is named “**species**”, so these files will respectively named “**species.faa**” or “**species.gbk**”.

- For each reference organism you want to use, create a subdirectory in the directory **orthology\_based\_reconstruction**. Give it the name of your model organism (e.g. **model\_a**). On a Linux operating system, here is the below command to create a new folder named **model\_a**.

```
shell> mkdir orthology_based_reconstruction/model_a
```

- In each folder, put:

- the Genbank file of your model organism, **with the same name than its directory, (here model\_a.gbk)** OR the proteome of your model organism, **with the same name than its directory (here model\_a.faa)**,
- the metabolic network of your model organism, **with the same name than its directory (here model\_a.sbml)**.

```
/species
|-- orthology_based_reconstruction
    |-- model_a
        |-- model_a.gbk or model_a.faa
        |-- model_a.sbml
        |-- dict_genes.txt (option)
```

- The genome (or proteome) and the metabolic network of your model organism have to contain the same kind genes (or proteins) identifiers to be comparable. If not enough genes (or proteins) are in common between the two files, the process will stop to avoid poor quality data production.

If you want to pursue on the process, please provide a dictionary file between the gene (or protein) identifiers of these two files. Name this dictionary **dict\_genes.txt**. Here is the dictionary file format asked (tabulation separated values):

```
gene_id_from_sbml1\tgene_id_from_faaA
gene_id_from_sbml2\tgene_id_from_faaB
gene_id_from_sbml3\tgene_id_from_faaC
```

## Orthology-based run

**Warning:** Remember to check the validity of the inputs before running any workflow task.

- If you want to run only the orthology-based reconstruction, use now this command:

```
aureme> aureme --run=species --cmd="check_input"
```

- To run **only** the orthology-based reconstruction, use this command:

```
aureme> aureme --run=species --cmd="orthology_based"
```

- Use this command, to get the database of a given metabolic network:

```
aureme> aureme --run=species --cmd="which_db SBML=output_orthofinder_from_model_a.
↪sbml"
```

**Warning:** Because the metabolic network from the reference organism could come from different databases, it's critical to check the database of each network and if needed convert the network to your reference database selected (see [How to use the AuReMe workspace \(default workflow\)](#) and [Define the reference database](#)).

The previous command will check the database of the file `output_orthofinder_from_model_a.sbml`, if the database is different for the reference, use the next command to create a mapping file to the reference database. For more information about SBML mapping see [How to map a metabolic network on another database?](#)

```
aureme> aureme --run=species --cmd="sbml_mapping SBML=output_orthofinder_from_model_a.
↪sbml DB=METACYC"
```

## 1.2.5 Annotation-based reconstruction

### Method: Pathway Tools

**Input files**

Required for the annotation-based reconstruction (method: [Pathway Tools](#)):

- \* The output of Pathway Tools (PGDB folder)

**Result files**

```

├── species
│   └── networks
│       └── output_annotation_based_reconstruction
│           └── pathwaytools
│               ├── output_pathwaytools_genome_a_padmet
│               └── output_pathwaytools_genome_b_padmet

```

```

graph TD
    A[Inputs checking & import] --> B[Pathway Tools]
    A --> C[OrthoFinder]
    A --> D[External network]
    B --> E[Manual curation]
    C --> E
    C --> F[Meneco]
    D --> E
    E --> F
    F --> G[Manual curation]
    G --> H[CobraPy FBA]
    G --> I[Wiki]
    G --> J[SBML]
    G --> K[Reports]

```

### Annotation-based inputs

1. Put the output of [Pathway Tools](#) (the whole PGDB directory) in the folder **annotation\_based\_reconstruction**.

```
/species
|-- annotation_based_reconstruction
    |-- genome_a (you can change the name of the folder)
```

(continues on next page)

(continued from previous page)

```
|-- compounds.dat
|-- enzrxns.dat
|-- genes.dat
|-- pathways.dat
|-- proteins.dat
|-- reactions.dat
```

- The above cited files are required in order to run the Annotation-based reconstruction. If you have run several times [Pathway Tools](#) and want to use all of these annotations, just copy-paste the other PGDB folders in the `annotation_based_reconstruction` directory.

**Warning:** AuReMe does not provide any license of [Pathway Tools](#).

### Annotation-based run

- If you want to run only the annotation-based reconstruction, use now this command:

```
aureme> aureme --run=species --cmd="check_input"
```

**Warning:** Remember to check the validity of the inputs before running any workflow task.

- To run **only** the annotation-based reconstruction, use this command.

```
aureme> aureme --run=species --cmd="annotation_based"
```

## 1.2.6 Merge metabolic networks

<p><b>Input files</b> Metabolic networks in the <b>networks</b> directory.</p> <p><b>Result file</b></p> <pre> ├── species │   └── networks │       └── draft.padmet         </pre>	<pre> graph TD     A[Inputs checking &amp; import] --&gt; B[Pathway Tools]     A --&gt; C[OrthoFinder]     A --&gt; D[External network]     B --&gt; E[Manual curation]     C --&gt; E     D --&gt; E     E --&gt; F[Meneco]     F --&gt; G[Manual curation]     G --&gt; H[CobraPy FBA]     H --&gt; I[Wiki]     H --&gt; J[SBML]     H --&gt; K[Reports]     </pre>
---	---

To merge all available networks in the the **annotation\_based\_reconstruction** directory, together with those are in the **networks** directory, and in all its subdirectories, into only one metabolic network. To merge all data on the studied species, run this command:

```
aureme> aureme --run=species --cmd="draft"
```

---

**Note:** You can also add external metabolic network to create the draft (see *Data organization*).

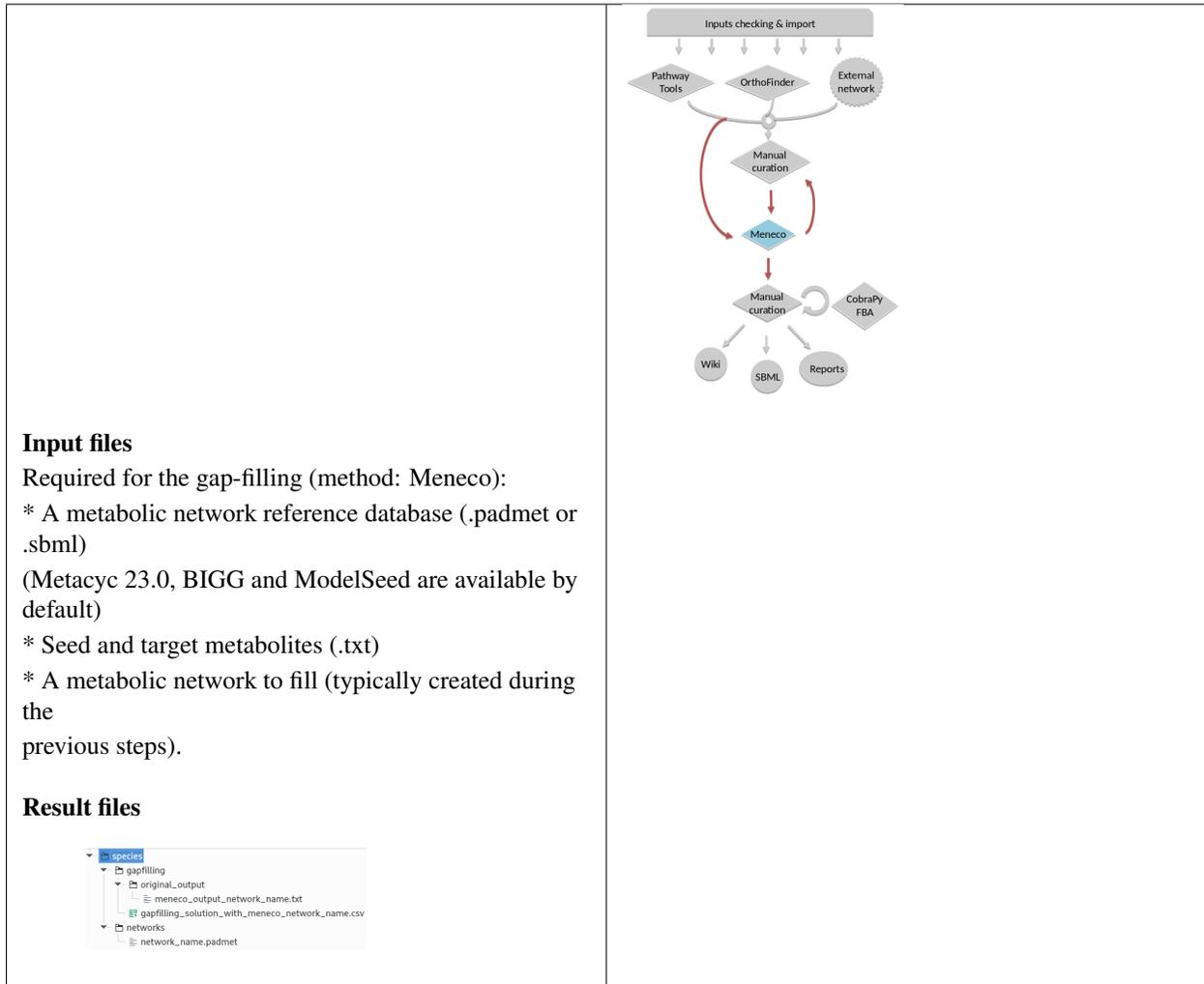
---

**Warning:** Before merging your networks, check if not already done if all the SBML are using the reference database, see *How to map a metabolic network on another database?* Also check the compartment ids used in each of them, delete and change compartment if need.

For example: if a SBML is using KEGG database but your reference database is Metacyc, you will have to map this SBML to create a mapping file which will be used automatically in the merging process.

### 1.2.7 Gap-filling

## Method: Meneco



## Gap-filling input

1. You must have selected a reference database to fill-in the potential gaps in the metabolic network. If it is not done yet, please see *Define the reference database*.
2. Put the seeds file (named **seeds.txt**) in the **growth\_medium** folder. The seed compounds are the description of the set of metabolites that is available to initiate the metabolism (growth medium). Put also the artefacts file (named **artefacts.txt**) in the **growth\_medium** folder. The artefacts file has the same format as the seeds file. Here, artefacts are metabolites allow Meneco to initiate cycles in in a metabolic network (report to *What are "artefacts"?* section). Here is as example of the seed file format:

```
seed_name_compound_id1\tcompartment1
seed_name_compound_id2\tcompartment2
seed_name_compound_id3\tcompartment3
```

- Set the growth medium using this command:

```
aureme> aureme --run=species --cmd="set_medium NETWORK=network_name NEW_
↳NETWORK=new_network_name"
```

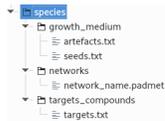
For more details on the medium settings, see [How to manage growth medium?](#)

**Warning:** If you don't precise any **NEW\_NETWORK** name, the current network will be overwritten.

- Put the target file (named **targets.txt**) in the **targets\_compounds** folder. The targets are metabolites whose production is supposed to be achieved by the metabolism of the species under study (components of the biomass reactions or other metabolites). Here is an example of the target file format:

```
target_name_compound_id1\compartment1
target_name_compound_id2\compartment2
target_name_compound_id3\compartment3
```

- You will have to indicate which metabolic network you want to gap-fill with the Meneco software. If you want to gap-fill a network created in the previous steps, there is nothing to do. Otherwise, put the network you want to gap-fill (Padmet format) in the **networks** directory.



## Gap-filling run

- (optional step) To generate the gap-filling solution run this command:

```
aureme> aureme --run=species --cmd="gap_filling_solution NETWORK=network_name"
```

**Note:** Do not forget the quotation marks.

It will calculate the gap-filling solution on the **network\_name** network (in the **networks** directory) and put it into the **gapfilling** directory as **gapfilling\_solution\_network\_name.csv**.

- To generate the gap-filled network (and run step 6), run this command:

```
aureme> aureme --run=species --cmd="gap_filling NETWORK=network_name NEW_
↳NETWORK=new_network_name"
```

**Note:** Do not forget the quotation marks.

It will calculate the gap-filling solution (if it is not yet done) on the **network\_name** network (in the **networks** directory) and put it into the **gapfilling** directory. Then it will generate the metabolic network (**new\_network\_name**), completed with the gap-filling solution, in the **networks** directory.

---

**Note:** You can first generate the solution, modify it, then generate the gap-filled network.

---

**Warning:** If you don't precise any **NEW\_NETWORK** name, the current network will be overwritten.

## 1.2.8 Manual curation

This step can be done several times and at any moment of the workflow.

1. Describe the manual curation(s) you want to apply by filling the corresponding form(s) as explained below.

**Warning:** It is highly recommended to create a new form file (.csv) each time you want to apply other changes, in order to keep tracks of them.

### Add a reaction from the database or delete a reaction in a network

- a) Copy from the folder **manual\_curation > template** the file **reaction\_to\_add\_delete.csv** and paste it into the **manual\_curation** directory (this way on Linux operating systems):

```
aureme> cp manual_curation/template/reaction_to_add_delete.csv manual_curation/my_
↪create_form.csv
```

- b) Fill this file (follow the exemple in the template).

idRef	Comment	Action	Genes
my_rxn	Reaction deleted because of x reason	delete	
RXN-12204	Reaction added because of x reason	add	(gene1 <b>or</b> gene2)
RXN-12213	Reaction added because of x reason	add	gene18
RXN-12224	Reaction added because of x reason	add	

### Create new reaction(s) to add in a network

- a) Copy from the folder **manual\_curation > template** the file **reaction\_creator.csv** and paste it into the **manual\_curation** directory (this way on Linux operating systems):

```
aureme> cp manual_curation/template/reaction_creator.csv manual_curation/my_
↪create_form.csv
```

- b) Fill this file (follow the exemple in the template).

reaction_id	my_rxn
comment	reaction added because of X reason
reversible	false
linked_gene	(gene_a <b>or</b> gene_b) <b>and</b> gene_c
<i>#reactant/product</i>	<i>#stoichio:compound_id:compart</i>
reactant	1.0:compound_a:c
reactant	2.0:compound_b:c
product	1.0:compound_c:c

(continues on next page)

(continued from previous page)

```

reaction_id      my_rxn_2
comment          reaction added because of X reason
reversible       true
linked_gene
#reactant/product #stoichio:compound_id:compart
reactant         1.0:compound_a:c
reactant         2.0:compound_d:c
product          1.0:compound_c:c
product          1.0:compound_d:c

```

## Apply changes

- To apply the changes described in the **my\_form\_file.csv** form file, run this command:

```

aureme> aureme --run=species --cmd="curation NETWORK=network_name NEW_NETWORK=new_
↪network_name DATA=my_form_file.csv"

```

**Warning:** If you don't precise any **NEW\_NETWORK** name, the current network will be overwritten.

## 1.3 FAQ

### 1.3.1 Can I have a sample of AuReMe?

To get an overview of AuReMe, you can get a sample by using this command:

```

aureme> aureme --sample

```

You will get a folder named 'tischrysis\_lutea' in your bridge directory. This folder contains all input and output files as if you had run the entire metabolic network reconstruction workflow for the example files about *Tischrysis lutea* (microalgae). Look at the **log.txt** file to understand the different commands used in the reconstruction process.

**Note:** Notice that by default all the outputs of commands will be printed in the terminal. Nevertheless, if you want to trace all your commands just type:

```

aureme> aureme --run=species --cmd="some_command" -q

```

The **quiet** option is usefull to reproduce a study. Thanks to this command, the AuReMe command will be stored as a log in the **bridge > species > log.txt** file. The output of the same AuReMe command will also be stored in the **bridge > species > full\_log.txt** file.

### 1.3.2 How to convert files to different formats?

The AuReMe workspace natively provides several functions for formats conversion, through the [PADMet Python package](#). The available convertors are:

- From SBML to Padmet format:

```
aureme> aureme --run=species --cmd="draft"
```

This command will convert all SBML in networks folder of 'species' to one Padmet. If you want to convert one SBML to Padmet format, simply put this file in networks folder of your run and make sure there is no other SBML file nor Padmet file, either in networks directory, or in one of the subdirectory of networks. Then run the command.

If you want to merge many SBML to one Padmet, add all of them in **networks > external\_network** folder then run the command. Ensure that there is no other SBML nor Padmet file, either in networks directory, or in one of the subdirectory of networks before running the command. In the case one SBML would be forgotten, it could add to the resulted **draft.padmet** or another reading error could occur.

- From Padmet to SBML format:

```
aureme> aureme --run=species --cmd="padmet_to_sbml NETWORK=my_network [LVL=3]"
```

This command will convert the Padmet file **my\_network.padmet** from networks folder of 'species' to create a SBML file **my\_network.sbml**. By default the SBML level is set to '3', you can change the default value in the config.txt file or with the argment LVL (3 or 2)

- From TXT to SBML format:

```
aureme> aureme --run=species --cmd="compounds_to_sbml CPD=/path/to/file/root_txt_
↳file"
```

This command will convert a txt file containing compounds ids to a SBML file **/path/to/txt\_file.sbml**. The txt file must contain one compound id by line and optionally the compartment of the id which by default is 'c'. Example of file:

```
ATP
ADP
WATER\tC-BOUNDARY
LIGHT\tC-BOUNDARY
```

- From GFF/GBK to FAA format:

---

**Note:** AuReMe integrate some scripts from `padmet-utils` tools, for example, `gbk_to_faa` command use the script `/programs/padmet-utils/connection/gbk_to_faa.py`. Not all functions are encapsulated in AuReMe, there is a lot of scripts that could be helpful. For more information, see <https://github.com/AuReMe/padmet-utils>.

```
aureme> aureme --run=species --cmd="gbk_to_faa GBK_FILE=/path/to/gbk_file OUTPUT=/
↳path/to/output_file"
```

### 1.3.3 How to manage growth medium?

In AuReMe, a compound is defined as a part of the growth medium (or 'seeds' for gap-filling tools) if this compound is in the compartment 'C-BOUNDARY'.

```
"M_CPD__45__7013_c" name="geranylgeranyl chlorophyll b" compartment="c"
"M_CA__43__2_C__45__BOUNDARY" name="Ca2+" compartment="C__45__BOUNDARY"
"M_CPD__45__474_c" name="(+) -taxifolin" compartment="c" boundaryCondition
```

The growth medium is linked to the metabolic network by two reactions, a non-reversible reaction named '*TransportSeed-compound-id*' which transport a compound of the growth medium from the compartment 'C-BOUNDARY' to the 'e' (extra-cellular) and a reversible reaction named '*ExchangeSeed-compound-id*' which exchange the same compound from 'e' to the 'c' (cytosol). When creating a SBML file, the compounds in the 'C-BOUNDARY' compartment will be set as 'BOUNDARY-CONDITION=TRUE' to allow flux (see [http://sbml.org/Documents/FAQ#What\\_is\\_this\\_.22boundary\\_condition.22\\_business.3F](http://sbml.org/Documents/FAQ#What_is_this_.22boundary_condition.22_business.3F)).

**Note:** Some metabolic networks manage the growth medium with a reversible reaction which consume nothing and produce a compound in the 'c' compartment. We chose not to do the same for clarity and because this method made crash some dedicated tools for metabolic network .

- Get the list of compounds corresponding to the growth medium of a network in Padmet format:

```
aureme> aureme --run=species --cmd="get_medium NETWORK=network_name"
```

Return a list of compounds or an empty list

- Set the growth medium of a network in Padmet format:

```
aureme> aureme --run=species --cmd="set_medium NETWORK=network_name [NEW_
↪NETWORK=new_network_name]"
```

This command will remove the current growth medium if existing, then create the new growth medium by adding the required reactions as described before.

- Delete the growth medium of a network in Padmet format:

```
aureme> aureme --run=species --cmd="del_medium NETWORK=network_name [NEW_
↪NETWORK=new_network_name]"
```

This function will remove all reactions consuming/producing a compound in 'C-BOUNDARY' compartment.

**Warning:** If you don't precise any **NEW\_NETWORK** name, the current network will be **overwritten**.

### 1.3.4 How to manage metabolic network compartment?

In a metabolic network a compound can occur in different compartment. Given a reaction transporting  $CA^{2+}$  from 'e' (extra-cellular compartment) to 'c' (cytosol compartment), the compartments involved are 'e' and 'c'. It is important to properly manage the compartments defined in a network to ensure a correct connection of the reactions. In some case metabolic networks can use different id to define a same compartment like 'C\_c', 'C', 'c' for cytosol, merging those networks could lead to a loss of network connectivity. A reaction producing  $CA^{2+}$  in 'c' and a reaction consuming  $CA^{2+}$  in 'C\_c' are actually not connected, hence the interest of the metabolic network compartment management commands of AuReMe.

- Get the complete list of compartment from a network in Padmet format:

```
aureme> aureme --run=species --cmd="get_compart NETWORK=network_name"
```

Return a list of compartment or an empty list.

- Change the id of a compartment from a network in Padmet format:

```
aureme> aureme --run=species --cmd="change_compart NETWORK=network_name OLD=old_
↪id NEW=new_id [NEW_NETWORK=new_network_name]"
```

This command will change the id of the compartment 'old\_id' to 'new\_id'. This command is required if different ids are used to define a same compartment, example changing 'C\_c' to 'c', or 'C-c' to 'c' ...

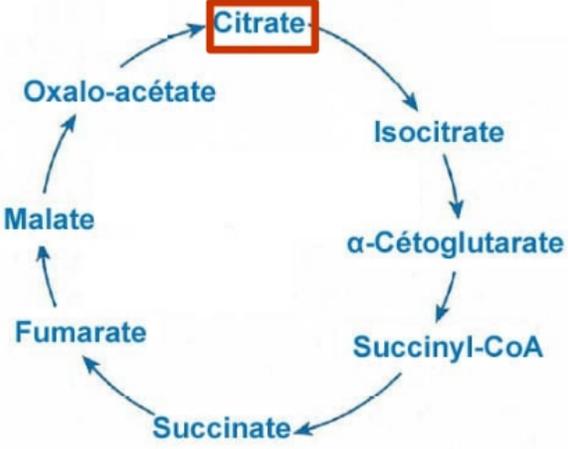
- Delete the id compartment from a network in Padmet format:

```
aureme> aureme --run=species --cmd="del_compart NETWORK=network_name_
↳COMPART=compart_id [NEW_NETWORK=new_network_name]"
```

This function will remove all reactions consuming/producing a compound in 'compart\_id' compartment.

**Warning:** If you don't precise any NEW\_NETWORK name, the current network will be **overwritten**.

### 1.3.5 What are "artefacts"?

<p>Meneco is a tool that fill the gaps topologically in a network, thanks to a reference database (see the <i>Method: Meneco</i> section). In fact, Meneco cannot product any other metabolite of an cycle without initiate it before.</p> <p>Thereby, artefacts are metabolites allow Meneco to initiate cycles in a metabolic network.</p> <p>For example in the picture aside, the Kreps cycle needs to be initiated with Meneco. A manner to initiate the Kreps cycle into Meneco is to put the "citrate" metabolite as one of the "artefacts"</p>	 <p>before gap-filling the network thanks to Meneco.</p>
--	---

### 1.3.6 How to explore the topology of a metabolic network?

A manner of exploring and analyzing the topology of a metabolic network is to use the *MeneTools* (Metabolic Network Topology Tools). Two MeneTools: Menecheck and Menescope are included in AuReMe. You can run the one or the other individually.

<pre> /test  -- growth_medium  -- artefacts.txt  -- seeds.txt  -- networks  -- network_name.sbn1  -- targets_compounds  -- targets.txt </pre> <p><b>Input files</b></p>	<pre> /test  -- analysis  -- topological_analysis  -- menecheck_network_name.txt  -- menescope_network_name.txt </pre> <p><b>Result files</b></p>
---	---

To obtain additional information about the file format of **artefacts.txt**, **seeds.txt**, and **targets.txt**, please refer to *Gap-filling input* and *What are “artefacts”?* sections.

- Menecheck gives the producibility status using graph-based criteria. To run Menecheck, use this command:

```
aureme> aureme --run=test --cmd="menecheck NETWORK=network_name"
```

- Menescope provides the topologically reachable compounds from seeds (and artefacts) in a metabolic network. To run Menescope, use this command:

```
aureme> aureme --run=test --cmd="menescope NETWORK=network_name"
```

### 1.3.7 How to manage the log files?

By default, each AuReMe command is printed in the terminal.

If you want to trace all your commands just type:

```
aureme> aureme --run=species --cmd="some_command" -q
```

Thanks to the **quiet** option, the command will be stored as a log in the **bridge > species > log.txt** file. The output of this command will also be stored in the **bridge > species > full\_log.txt** file.

If you store all your commands in the log files, it will be easy to reproduce your study, please see the *How to reproduce studies?* section.

Both log files could be erased with this command:

```
aureme> aureme --run=species --cmd="clean"
```

### 1.3.8 How to reproduce studies?

If you want to re-run the complete workflow of a pre-run study, built with AuReMe:

- During all the steps of your pre-run study, think to use the **quiet** option, (see the *How to manage the log files?*):

```
aureme> aureme --run=species --cmd="some_command" -q
```

- Create a new study (as described in the *Start a new study* section) by running the init command:

```
aureme> aureme --init=species_2
```

It generates a new folder named **species\_2** in the **bridge** directory.

- Update your **config.txt** file, if it is needed.
- Now, copy all the input data from the previous study in this new folder (please, follow the folder architecture described in the *Data organization* section).

- Copy also the **log.txt** file in the **bridge > species\_2** directory. Then in the **log.txt** file, **change every occurrence of the species name by species\_2**.
- Execute the log file.

```
aureme> ./shared/species_2/log.txt
```

### 1.3.9 How to create a new ‘à-la-carte’ workflow?

If you want to add a new step in the workflow or add a new method, it is possible to customize AuReMe. For that it is necessary to update the Makefile in your run. Here is an example of how to do it.

- Add a new method:

First, install your tool by following the documentation associated. For the example we will add a new tool for orthology-based reconstruction ‘new\_tool’ which use the same input as Pantograph (a metabolic network in SBML format, a gbk of the reference species and the GBK of the study species) and generate the same output (a metabolic network in SBML format).

Secondly we will update the Makefile by adding these lines:

Basically this command says that for each folder in `orthology_based_reconstruction` (variable declared in `config.txt`), if the expected output is not already created, run `new_tool`.

Finally, to select this method in your new workflow, change in the file `config.txt` the variable `ORTHOLOGY_METHOD=pantograph` by `ORTHOLOGY_METHOD=new_tool`

- Add a new step or function:

Just update the Makefile by adding a new step and use it with this command

### 1.3.10 How to choose another reference database?

It is possible to select a reference database among several. You can display the list of all available databases by using this command:

The reference database is needed to:

- be able to match all the identifiers of the entities of metabolic networks
- gap-fill the metabolic network in the gap-filling step

To select one, replace the corresponding path in the configuration file: `*config.txt*`, in the `*DATA_BASE*` variable. Or you can comment the line if you don’t want/can’t use a database. The `*config.txt*` file is stored at the root of your `*bridge*` folder (see *Running a Docker container*).

### 1.3.11 What is checked in my input files?

Before running any command in AuReMe, it is highlight recommended to use the command ‘`check_input`’. This command checks the validity of the input files and can also create required files. Concretely this command:

- Checks database: If database was specified in the `config.txt` file (see the *How to choose another reference database?* section). If so, checks if a SBML version exist and create it on the other hand.
- Checks studied organism data: Search if there is a genbank (gbk/gff) ‘`GBK_study.gbk`’ and proteome (faa) ‘`FAA_study.faa`’ in `genomic_data` folder. If there is only a genbank, create the proteome (command ‘`gbk_to_faa`’). If there is only the proteome or any of them, just continue the checking process. Note that the proteome is only required for the orthology-based reconstruction, method: Pantograph.

- **limage8** Checks orthology-based reconstruction data: for each folder found in 'orthology\_based\_reconstruction' folder checks in each of them if there is proteome 'FAA\_model.faa' and a metabolic network 'metabolic\_model.sbml', if there is no proteome but a genbank file 'GBK\_study.faa', create the proteome (command 'gbk\_to\_faa'). Finally, the command compares the ids of genes/proteins between the proteome and the metabolic network.

If cutoff... important because... dict file to create a new proteome file ...

- Checks annotation-based reconstruction data: for each folder found in 'annotation\_based\_reconstruction' folder checks in each of them if it's a PGBD from pathway then create (if not already done) a Padmet file '**output\_pathwaytools\_**'folder\_name'.padmet in networks/output\_annotation\_based\_reconstruction folder.
- Checks gap-filling data: In order to gap-fill a metabolic network, Pantograph required as input, a file 'seeds.sbml' describing the seeds (the compounds available for the network), another describing the targets (the compounds that the network have to be able to reach), the metabolic network to fill and the database from where to draw the reactions all in SBML format. It's possible to start from txt files for seeds 'seeds.txt' and targets 'targets.txt', each file containing the ids of the compounds, one by line. The command will then convert them to SBML (command 'compounds\_to\_sbml').

Note that by default, AuReMe will integrate the artefacts 'default\_artefacts\_metacyc\_20.0.txt' to the seeds to create a file 'seeds\_artefacts.txt' and 'seeds\_artefacts.sbml'. For more information about the artefacts see *What are "artefacts"?* section.

Example:

```
**[output]**
```

```
INSERT SCREEN FROM check_input log
```

### 1.3.12 What is the Makefile?

Makefile contient les cmd de AuReMe. exemple de cmd simple

### 1.3.13 What is the config.txt file?

The **config.txt** is found in the **bridge > species** directory. It contains all the AuReMe parameters: the name of the selected database, the name of the various choosen methods, and the default parameters of all programs that AuReMe needed.

If you want to use either another database or another tool already included in the AuReMe workspace, modify carefully the **config.txt** file.

**Warning:** The parameters of the **config.txt** must not be changed unless you are sure of what you want do!

### 1.3.14 How to regenerate a new database version?

Voir les notes de Jeanne sur le problème de Sebastian

padmet/utills/connexion

### 1.3.15 How to map a metabolic network on another database?

Metabolic networks can be products of varied databases. If you want to merge efficiently information about metabolic networks coming from different databases, you will need to map the metabolic network(s) to a common database. To do so, a solution is provided by AuReMe.

Note: to use this method, the metabolic network to map needs to be in the SBML format and stored in the **\*networks\*** folder.

- First of all, you need to know the origin database of the data. To recognize the database used in an SBML file, use the **\*which\_db\*** command:

Example:

```
**[output]**
```

- When you know the origin database of the data, you have to generate the mapping dictionary from this database to the new one:

Example:

```
**[output]**
```

In this example, the system has found more than just one mapping for the *R\_R00494\_c* reaction and the *S\_Starch\_p* compound. It did not manage to choose between the propositions: the mapping will not be added to the output mapping. If you want to force the mapping, you have to modify the mapping file manually.

- Once you have created a mapping dictionary file, it will be automatically applied across the workflow to translate the data.

### 1.3.16 How to generate report on results?

Create reports on the **network\_name.padmet** file network (in the **networks** directory).

```
aureme> aureme --run=species --cmd="report NETWORK=network_name"
```

Four files are created in the **analysis > reports > network\_name** directory thanks to the report command.

- **all\_genes.csv** (has the following format):

id	Common name	linked reactions
TL_15991	Unknown	2.3.1.180-RXN;RXN-9535
TL_5857	Unknown	RXN-14271;RXN-2425
TL_6475	Unknown	RXN-14229

If a gene is linked with several reactions, reactions are separated from “;”.

- **all\_metabolites.csv** (has the following format):

dbRef_id	Common name	Produced (p), Consumed (c), Both (cp)
NAD-P-OR-NOP	NAD (P) +	cp
THIOCYSTEINE	thiocysteine	p
CPD-18346	cis-vaccenoyl-CoA	c

- **all\_pathways.csv** (has the following format):

dbRef_id	Common name	Number of reaction found	Ratio (Reaction found / Total)
↔ Total number of reaction			
COA-PWY-1	coenzyme A biosynthesis II (mammalian)	1	1
↔		1.00	

(continues on next page)

(continued from previous page)

PWY-4984	urea cycle	1	5
↔	0.20		
PWY-7821	tunicamycin biosynthesis	1	9
↔	0.11		

- all\_reactions.csv (has the following format):

dbRef_id	Common name	formula (with id)	
↔		formula (with common name)	
↔		in pathways	associated genes
↔	categories		
NDPK	nucleoside-diphosphate kinase	1.0 ATP + 1.0 DADP => 1.0 ADP + 1.	
↔0 DATP		1.0 ATP + 1.0 dADP => 1.0 ADP + 1.0 dATP	
↔			TL_16529;TL_
↔13128	ORTHOLOGY		
RXN-15122	ORF	1 THR => 1 PROTON + 1 CPD-15056 +	
↔1 WATER		1 L-threonine => 1 H+ + 1 (2Z)-2-aminobut-2-	
↔enoate + 1 H2O		PWY-5437; ILEUSYN-PWY; PWY-5826	TL_17207; TL_
↔12535; TL_8525	ANNOTATION; ANNOTATION; ORTHOLOGY		
SGPL11	sphinganine 1-phosphate aldolase	1.0 CPD-649 => 1.0 PALMITALDEHYDE	
↔+ 1.0 PHOSPHORYL-ETHANOLAMINE		1.0 sphinganine 1-phosphate => 1.0	
↔palmitaldehyde + 1.0 O-phosphoethanolamine			TL_
↔105	ORTHOLOGY		

In this file, if there are several data in the same field, data are separated from “;”.

### 1.3.17 How to generate Wiki?

**Input files**

```
/test
|-- networks
|-- network_name.padmet
```

**Result files**

```
/test
  |-- analysis
  |-- wiki_pages
  |-- network_name
  |-- files
  |-- genes
  |-- metabolites
  |-- misc
  |-- navigation
  |-- organisms
  |-- pathways
  |-- reactions
```

```

graph TD
    A[Inputs checking & import] --> B[Pathway Tools]
    A --> C[OrthoFinder]
    A --> D[External network]
    B --> E[Manual curation]
    C --> E
    D --> E
    E --> F[Meneco]
    F --> G[Manual curation]
    G --> H[CobraPy FBA]
    G --> I[Wiki]
    G --> J[SBML]
    G --> K[Reports]
    
```

### Requirements

1. Utilize AuReMe, to create the wiki pages from a metabolic network.

An input file **network\_name.padmet** inside the **brigde > test > networks** directory is needed. The wiki pages will be deployed in **brigde > test > analysis > wiki\_pages > network\_name**.

```
aureme> aureme --run=test --cmd="wiki_pages NETWORK=network_name"
```

**Warning:** Run all the next commands from your machine and not from the AuReMe container.

You can use wikis to analyze or visualize your metabolic networks, thanks to the [MediaWiki](#) technology.

2. Clone the wiki software within your computer:

```
shell> git clone https://github.com/AuReMe/wiki-metabolic-network.git
shell> cd wiki-metabolic-network/wiki-metabolic-network/
shell> make init
```

The **wiki-metabolic-network** is now installed on your computer. You can manage it in using the [docker.com](#) commands (see *Some tips about Docker*). **wiki-metabolic-network** is an image that allows to automatize the creation of wikis in a container.

- Get the name of the wiki container, it will be useful to run the next command.

```
shell> docker ps -a
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES	SIZE
a0aecf97349c	docker.io/dyliss/wiki-metabolic-network-img	"docker-php-entryp..."	2 weeks ago	Up 7 hours	0.0.0.0:80->80/tcp	wikimetabolic_mediawiki_1	512 MB (virtual 2.35 GB)
3e82fccd147f	mariaadb	"docker-entrypoint..."	2 weeks ago	Up 7 hours	3306/tcp	wikimetabolic_database_1	2 B (virtual 349 MB)
5cb6cd9515a1	docker.io/dyliss/aureme-img	"bash"	2 weeks ago	Up 5 hours		aureme-cont	11 MB (virtual 3.52 GB)

**Warning:** For a shake of genericity, in the following steps of this manual, we will employ the term of **wiki\_cont** instead of **wikimetabolic\_mediawiki\_1** (the real one you have to write in your command lines).

- To enter the wiki container.

```
shell> docker exec -it wiki_cont bash
```

- To print the commands of the wiki container.

```
shell> docker exec -it wiki_cont wiki --help
```

- Copy the data previously created thanks to AuReMe, in the wiki container.

```
shell> docker cp /test/analysis/wiki_pages/network_name wiki_cont:/home/
```

## Wiki creation

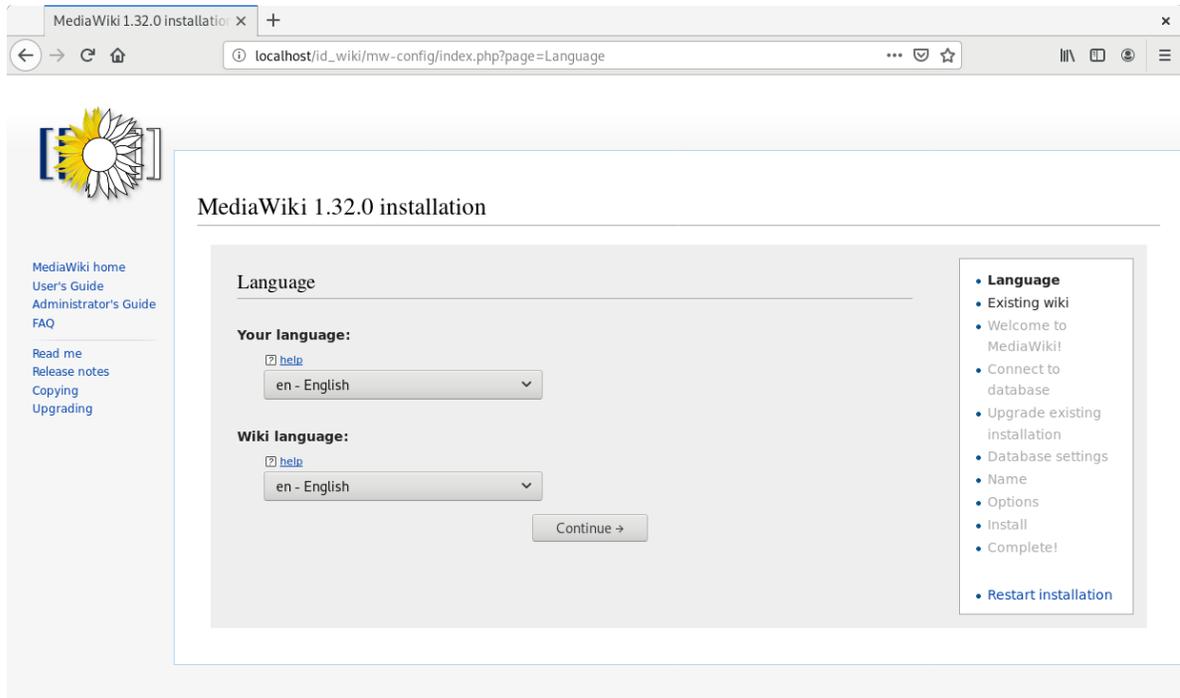
Follow the instructions on your terminal.

```
shell> docker exec -ti wiki_cont wiki --init=id_wiki
```

```
jcambefo@tiso:~/HOME/docker/aureme/aure_test_0419/analysis/wiki_pages
File Edit View Search Terminal Help
Checking wiki id id_wiki...
Checking wiki id id_wiki: OK
Checking if the prefix id_wiki_ is already used in the database...
Checking the if the prefix {wiki_id}_ is already used in the database: OK
Wiki initialization...
  Copying wiki folder
  Updating var in LocalSettings.php
  Setting wiki id to id_wiki
  Setting wiki host url to http://localhost
  Setting database host database
  Setting database user wikidb_user
  Setting database password
  Setting wiki prefix to id_wiki_

#####
MANUAL SETUP IS NOW REQUIRED. Access to this link from your browser:
  http://localhost/id_wiki/mw-config/index.php
Follow this instructions to setup mediawiki:
Language:
  Continue->
Existing wiki:
  Upgrade key: 62763ed4b27d11fc
  Continue->
Welcome to MediaWiki!:
  Continue->
Database settings:
  Continue->
Name:
  Name of wiki: metabolic_network
  \tAdministrator account: /\ Use exactly the same to allow the bot to upload the pages automatically
  Your username: admin
  Password: 123456789
  I'm bored already, just install the wiki.
  Continue->
Install:
  Continue->
  Continue->
  Do not save the LocalSettings file
#####
When the previous setup is done, press enter to continue..
```

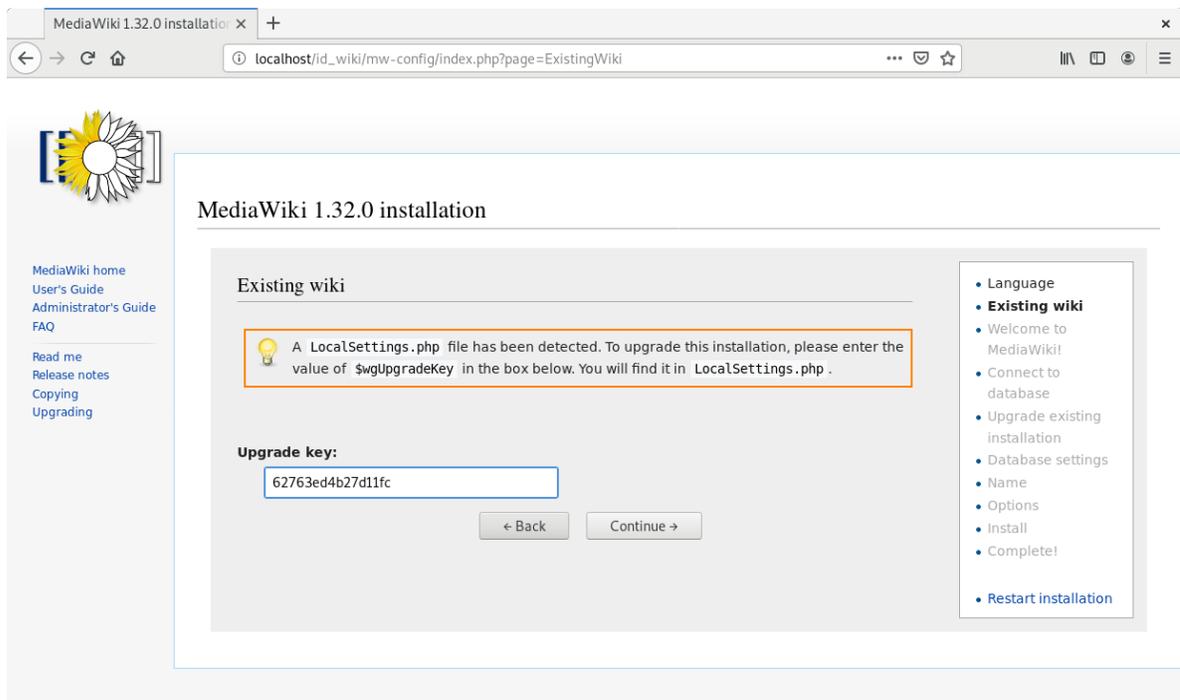
1. Open your browser at the following address: [http://localhost/id\\_wiki/mw-config/index.php](http://localhost/id_wiki/mw-config/index.php), and press “Continue”.



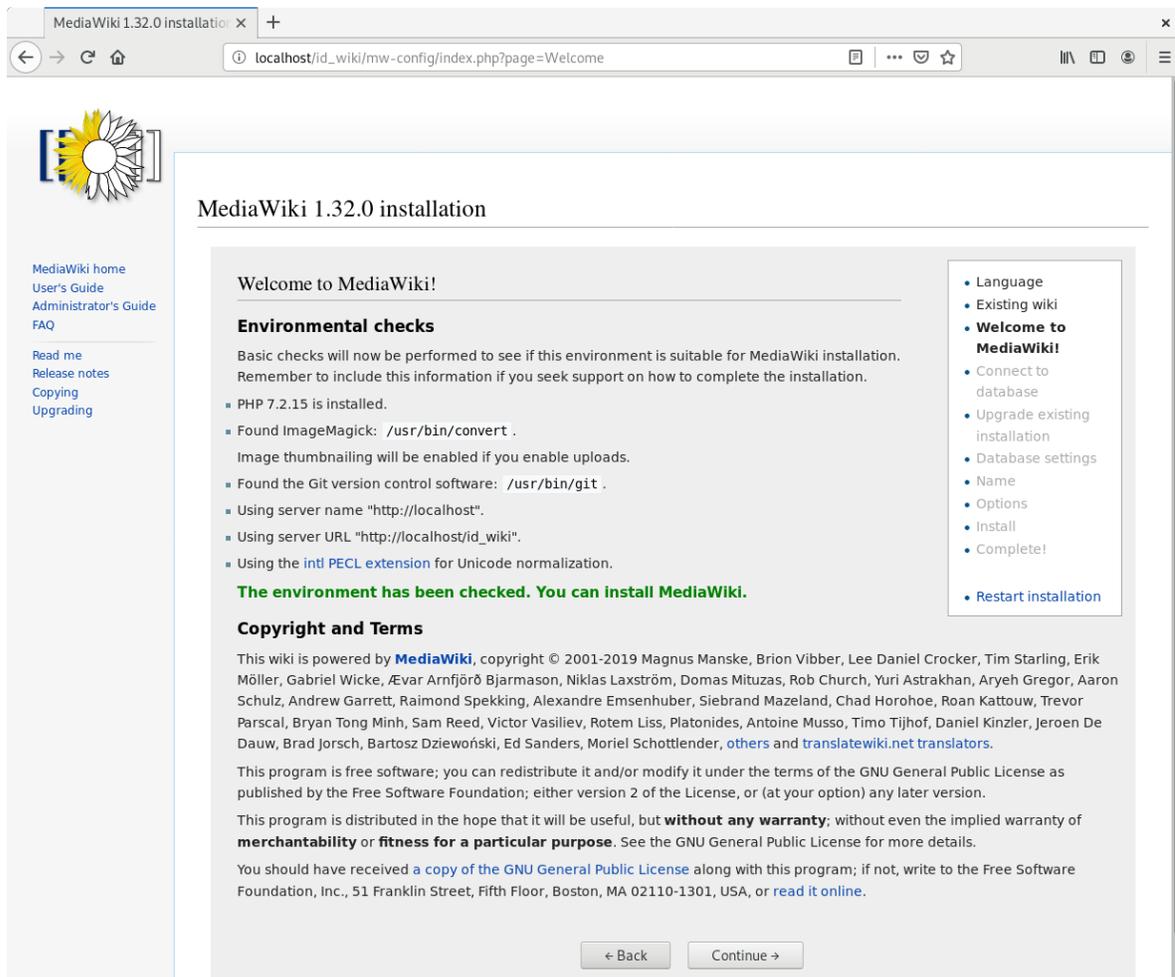
2. Get the “Upgrade key”. The Upgrade key is found on the your terminal. This is a small part extracted from the terminal to locate it better.

```
Existing wiki:
Upgrade key: 62763ed4b27d11fc
Continue >
```

3. Enter the “Upgrade key”, and press “Continue”.



4. In the page “Welcome to MediaWiki” configuration, just press “Continue”.



MediaWiki 1.32.0 installation

Welcome to MediaWiki!

### Environmental checks

Basic checks will now be performed to see if this environment is suitable for MediaWiki installation. Remember to include this information if you seek support on how to complete the installation.

- PHP 7.2.15 is installed.
- Found ImageMagick: `/usr/bin/convert`.  
Image thumbnailing will be enabled if you enable uploads.
- Found the Git version control software: `/usr/bin/git`.
- Using server name "http://localhost".
- Using server URL "http://localhost/id\_wiki".
- Using the [intl PECL extension](#) for Unicode normalization.

**The environment has been checked. You can install MediaWiki.**

### Copyright and Terms

This wiki is powered by [MediaWiki](#), copyright © 2001-2019 Magnus Manske, Brion Vibber, Lee Daniel Crocker, Tim Starling, Erik Möller, Gabriel Wicke, Evar Arnfjörð Bjarmason, Niklas Laxström, Domas Mituzas, Rob Church, Yuri Astrakhan, Aryeh Gregor, Aaron Schulz, Andrew Garrett, Raimond Spekking, Alexandre Emsenhuber, Siebrand Mazeland, Chad Horohoe, Roan Kattouw, Trevor Parscal, Bryan Tong Minh, Sam Reed, Victor Vasiliiev, Rotem Liss, Platonides, Antoine Musso, Timo Tijhof, Daniel Kinzler, Jeroen De Dauw, Brad Jorsch, Bartosz Dzięwoński, Ed Sanders, Moriel Schottlender, [others](#) and [translatewiki.net translators](#).

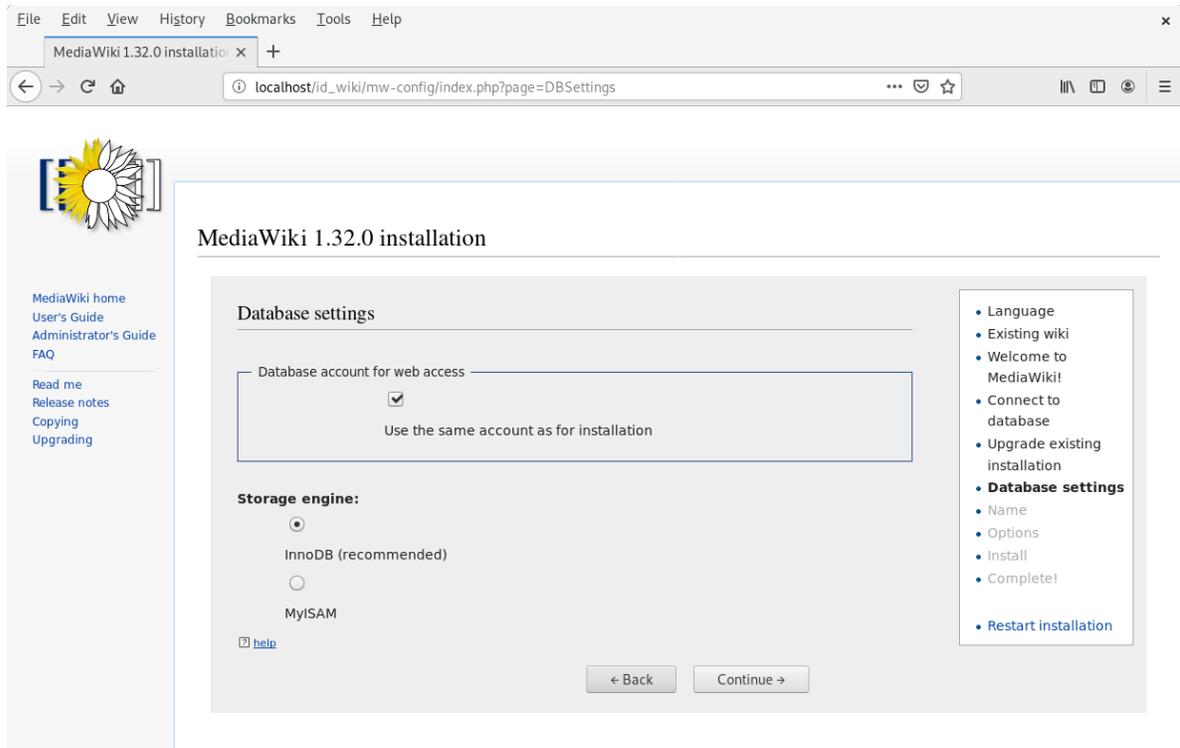
This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but **without any warranty**; without even the implied warranty of **merchantability** or **fitness for a particular purpose**. See the GNU General Public License for more details.

You should have received a [copy of the GNU General Public License](#) along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA, or [read it online](#).

← Back    Continue →

5. In the page “Database settings” configuration, just press “Continue”.



6. In the page “**Name**” configuration, you have several fields to fill:
- a. Name of wiki: **wiki\_name**
  - b. Your username: **admin**
  - c. Password: Enter a password (it is at least 8 characters).
  - d. Password again: Enter the same password.
  - e. Email address: jeanne.got[at]irisa.fr (for example)
  - f. Please select the phrase: “**I’m bored already, just install the wiki**”.
  - g. Press “**Continue**”.

MediaWiki 1.32.0 installation

Name

Name of wiki:

Project namespace:  
 Same as the wiki name: Wiki\_name  
 Project  
 Other (specify)

Administrator account

Your username:

Password:

Password again:

Email address:

[help](#)  
Subscribe to the release announcements mailing list.

[help](#)  
Share data about this installation with MediaWiki developers.

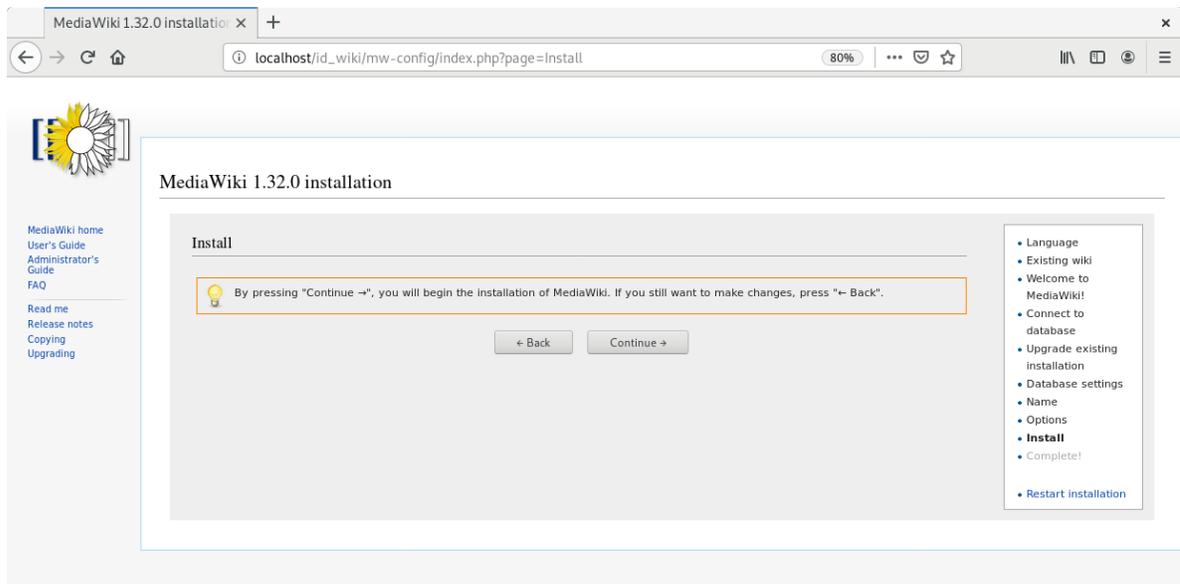
You are almost done! You can now skip the remaining configuration and install the wiki right now.

Ask me more questions.  
 I'm bored already, just install the wiki.

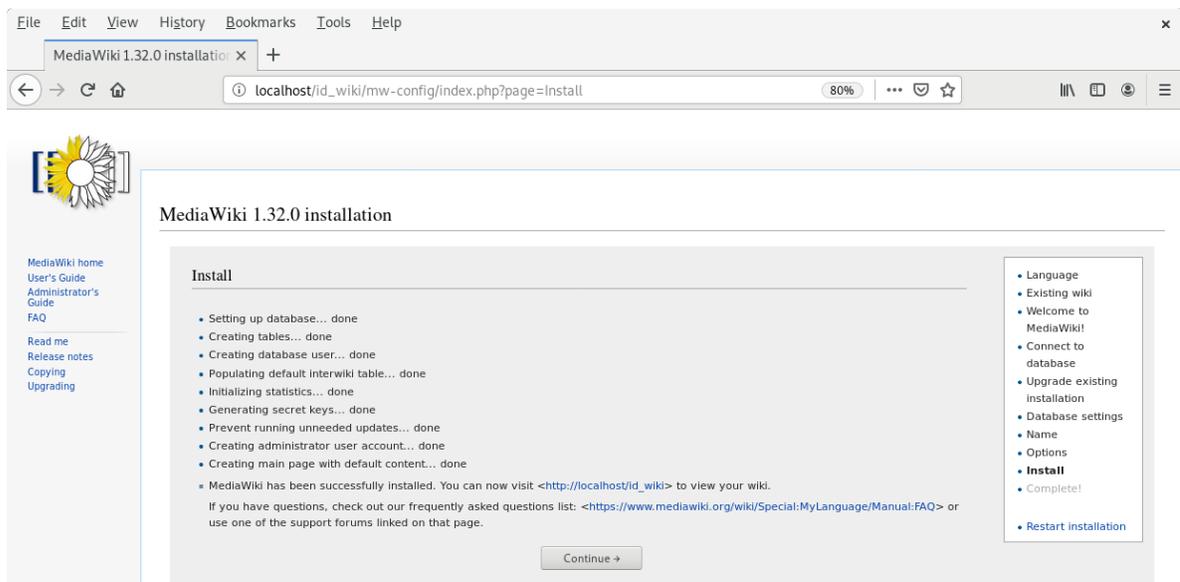
- Language
- Existing wiki
- Welcome to MediaWiki!
- Connect to database
- Upgrade existing installation
- Database settings
- **Name**
- Options
- Install
- Complete!

[Restart installation](#)

7. In the first page of “Install”, just press “Continue.”



8. In the second page of **“Install”**, just press **“Continue”**.



9. Do not download the LocalSettings.php file.

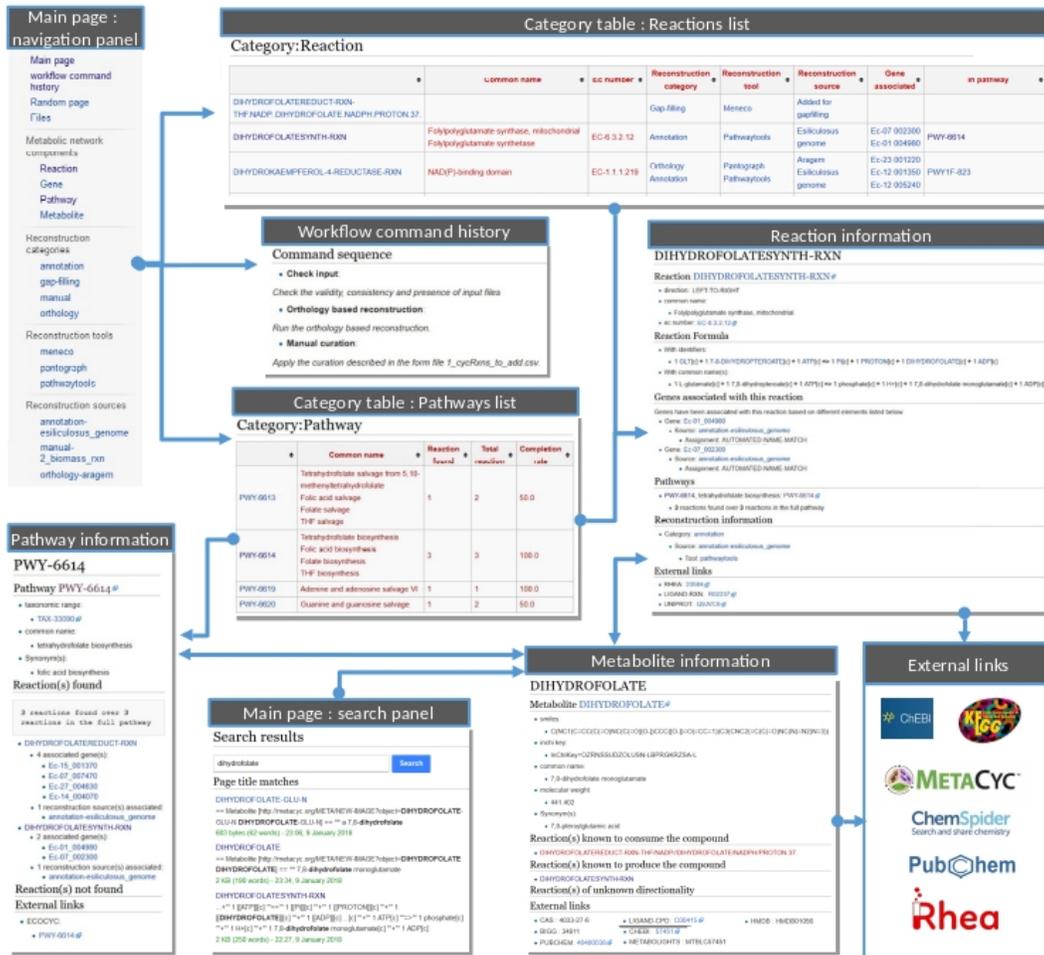
10. **Go back to your terminal, and press “Enter”**. The wiki is now online and reachable at this link: **http://localhost/id\_wiki/index.php/Main\_Page**.

- To send the “wiki pages” (that you previously copied in the wiki\_cont container) on the wiki.

```
shell> docker exec -ti wiki_cont wiki_load --action=load --url=http://localhost/
↵id_wiki/api.php --user=admin --password=my_password --wikipage=/home/network_
↵name --bots=2 -v
```

Here **“bots”** is the number of CPUs are allocated to make this task.

- Now wiki pages are accessible on **http://localhost/id\_wiki/index.php**. The following picture shows some functionalities of the wiki.



## Public and private access

**Note:** By default, a “public access” wiki is created. A wiki with a public access means, a wiki which everyone is allowed to access and to edit it on condition that she/he has an account on this wiki.

- To deploy a wiki with a “private access”.

```
shell> docker exec -ti wiki_cont wiki --init=id_wiki --access=private
```

Then, see the [Wiki creation](#) section. A wiki with a “private access” is preventing access and editing for non-user. It also prevent account creation. It is useful to manage confidential data.

- To modify the access of a wiki already created.

```
shell> docker exec -ti wiki_cont wiki --id=id_wiki --access=private
shell> docker exec -ti wiki_cont wiki --id=id_wiki --access=public
```

## Other wiki commands

- To list all deployed wiki use.

```
shell> docker exec -ti wiki_cont wiki --all
All deployed wiki:
    C_elegans
    E_siliculosus
    id_wiki
    S_cerevisiae
```

- To remove a wiki use.

```
shell> docker exec -ti wiki_cont wiki --id=id_wiki --remove
Removing wiki id_wiki
Removing wiki folder
97 tables to drop
```

It removes the “id\_wiki” from wiki\_folders and removes tables from database which start with prefix id\_wiki.

- To reset a wiki use.

```
shell> docker exec -ti wiki_cont wiki --id=id_wiki --clean
id_wiki_page table to empty
```

It only remove all the pages of the specified wiki. It keeps tables and folder associated with this wiki.

### 1.3.18 How to connect to Pathway-tools?

AuReMe is able to manage and modifies networks that came from [Pathway Tools](#).

1. Copy the PGDB folder in the **annotation\_based\_reconstruction** directory.

```
/species
|-- annotation_based_reconstruction
    |-- genome_a (you can change the name of the folder)
        |-- compounds.dat
        |-- enzrxns.dat
        |-- genes.dat
        |-- pathways.dat
        |-- proteins.dat
        |-- reactions.dat
```

2. Run the annotation-based reconstruction command:

```
aureme> aureme --run=species --cmd="annotation_based"
```

The file **output\_pathwaytools\_genome\_a.padmet** will be obtained in the **species/networks/output\_annotation\_based\_reconstruction/pathwaytools/** folder, (to obtain further details, report to the *Annotation-based reconstruction* part).

3. Move the **output\_pathwaytools\_genome\_a.padmet** in the **species/networks/** directory and rename the file (as **species\_pwt.padmet** for example). To realize this on a Linux operating system, you could employ the below command.

```
shell> cd species/networks/
shell> mv output_annotation_based_reconstruction/pathwaytools/output_
↳pathwaytools_genome_a.padmet species_pwt.padmet
```

4. Now, you are able to modify the **species\_pwt.padmet** either in adding, or deleting one or several reactions, thanks to the *Manual curation* part.

5.

, modifies it, and reopens it inside Pathway-Tools. .. XXX

on prend un PGDB, on crée un padmet, on modifie le padmet (par exemple en ajoutant une réaction), et on génère de nouveau un PGDB que l'on peut ouvrir dans Pathway Tools.

- Create PGDB from output of AuReMe

### 1.3.19 How to set an objective reaction?

To add a biomass reaction to a network, see the *Create new reaction(s) to add in a network* section. Once the biomass is included in the network, you have to set the biomass as objective function.

Apply this command to the **network\_name.padmet**

```
aureme> aureme --run=species --cmd="set_fba ID=reaction_name NETWORK=network_name"
```

It creates the **network\_name.sbml** file with `reaction_name` as the objective function. To continue the analysis of the `network_name`, see the *How to process Flux Balance Analysis?* section.

### 1.3.20 How to process Flux Balance Analysis?

AuReMe evaluate the flux balance analysis of the biological network, thanks to the `cobrapy` Python package. Before calculating the flux balance analysis of a network:

- a. you may have to add the biomass to a network in reporting to the *Create new reaction(s) to add in a network* section,
- b. you have to set the biomass as an objective reaction, please refer to the *How to set an objective reaction?* section.

To compute the flux balance analysis of the **network\_name.sbml** file:

```
aureme> aureme --run=species --cmd="summary NETWORK=network_name"
```

Two files: **network\_name.txt** and **network\_name\_log.txt** are generated in the **analysis > flux\_analysis** directory. The first file (**network\_name.txt**) summarizes the network, then it gets the list of productible and unproductible targets. For each productible target, the flux balance analysis is given. The growth rate of the network is also provided. Here is an example of a **network\_name.txt** format:

```

File Edit Options Buffers Tools Text Help
#####
Model summary
Number of compounds: 2746
Number of reactions: 2700
Number of genes: 2699
Ratio rxn with genes/rxns: 86.92592592592592%
#####
Analyzing targets
#Topological analysis
Number of targets: 32
Reading draft network from /shared/tl/networks/TL.sbml ... done.
Reading seeds from growth_medium/seeds artefacts.sbml ... done.
Reading targets from targets_compounds/targets.sbml ... done.

Checking draftnet for unproducible targets ... done.
producible targets:
"M DNA_45_Holder_c"
"M RNA_45_Holder_c"
"M General_45_Protein_45_Substrates_c"
"M L_45_PHOSPHATIDATE_c"
"M CHLOROPHYLL_45_A_c"

0 unproducible targets:

#Flux Balance Analysis
DNA_45_Holder_c // DNA-Holder_c 368.00785083415127 positive
RNA_45_Holder_c // RNA-Holder_c 542.822677925213 positive
General_45_Protein_45_Substrates_c // General-Protein-Substrates_c 329.67938
679634057 positive
L_45_PHOSPHATIDATE_c // L-PHOSPHATIDATE_c 528.0788335658973 positive
CHLOROPHYLL_45_A_c // CHLOROPHYLL-A_c 73.44868338069017 positive
#####
Computing optimization
Testing reaction BIOMASS
Growth rate: 73.44868338069149
Status: optimal
IN FLUXES          OUT FLUXES          OBJECTIVES
-----
Light e           468      Biomass_45_...  73.4  BIOMASS  73.4
WATER e           37.3
FVA analysis:
  Blocked reactions: 1755
  Essential reactions: 135
  Essential genes: 49
5/5 compounds with positive flux
0/5 compounds without flux
DNA_45_Holder_c // DNA-Holder_c 18.04224648934846 positive
RNA_45_Holder_c // RNA-Holder_c 18.047654004915977 positive
General_45_Protein_45_Substrates_c // General-Protein-Substrates_c 17.979769
800910987 positive
L_45_PHOSPHATIDATE_c // L-PHOSPHATIDATE_c 528.0788335658973 positive
CHLOROPHYLL_45_A_c // CHLOROPHYLL-A_c 18.031222551383483 positive

-:--- network_name.txt All L52 (Text)
End of buffer

```

The second file (**network\_name\_log.txt**) supplies all the warnings produced computing the flux balance analysis.

## 1.4 Indices and tables

- genindex
- modindex
- search