# tutorial - 2 - multispecies

September 10, 2020

[1]:
```
# expand cells to the 95% of the display width
from IPython.core.display import display, HTML
display(HTML("<style>.container { width: 95% !important; }</style>"))
```

```
<IPython.core.display.HTML object>
```

## 1 Tutorial: Automatic rule-based modeling of multi-species metabolism employing *Atlas*

Authors: Rodrigo Santibáñez[1,2], Daniel Garrido[2], and Alberto Martín[1]

Date: September 2020

Affiliations: 1. Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago, 8580745, Chile. 2. Department of Chemical and Bioprocess Engineering, School of Engineering, Pontificia Universidad Católica de Chile, Santiago, 7820436, Chile

Notes: This tutorial was created for the manuscript "*Atlas*: Automatic modeling of regulation of bacterial gene expression and metabolism using rule-based languages", first submitted for peer-review to Bioinformatics on May, 2020.

### 1.1 Prerequisites

0. The tutorial was prepared and executed on Ubuntu 20.04, PathwayTools version 24, and docker engine version 19.03.8.

1. PathwayTools must be installed and running to obtain data from the EcoCyc database. Please, run `pathway-tools -lisp -python-local-only` before to obtain any data. (Optional) The PathwayTools software could be executed in the background, with help of `nohup pathway-tools -lisp -python-local-only > /dev/null 2> /dev/null &`. Please follow instructions at http://pathwaytools.org/ to obtain a licensed copy of the software from https://biocyc.org/download-bundle.shtml. However, data could be manually formatted using a text-based editor or a spreadsheet software.

   Note: If you ran into the `pathway-tools/ptools/24.0/exe/aclssl.so: undefined symbol: CRYPTO_set_locking_callback` error, please follow instructions here: https://github.com/glucksfall/atlas/tree/master/PTools-v24. Instructions will guide you to install a docker image that is able to run pathway tools, but does not include it, so you still need to obtain the software with a valid license.

2. (Highly recommended) Install Docker. Please follow instructions for a supported Operating System https://docs.docker.com/engine/install/: On Ubuntu, install it with `apt-get install docker.io`. On Win10, install Docker Desktop with WSL2 support https://docs.docker.com/docker-for-windows/wsl/. On MacOS, install Docker Desktop https://docs.docker.com/docker-for-mac/install/. The Docker `networkbiolab/pleiades`installs the python packages, the jupyter server, and the stochastic simulators.

3. (Recommended) Jupyter notebook. We recommend the use of Anaconda3 https://www.anaconda.com/products/individual because of the easier installation of the stochastic simulators from https://anaconda.org/alubbock.

4. (Optional) A stochastic simulator, supported by the pySB python package (BNG2, NFsim, KaSim or Stochkit). pySB requires BNG2 to simulate models with NFsim.

5. (Optional) Cytoscape to visualize metabolic networks and others.

6. (Optional) A deterministic simulator: pySB supports ODE integration via scipy.integrate.ode, BioNetGen ODE integration, and CUDA-accelerated ODE integration with Marco S. Nobile's cupSODA software (https://github.com/aresio/cupSODA). If the user feel comfortable with SBML models, pySB could export to SBML and deterministic simulation done with libRoadRunner (http://libroadrunner.org/), Tellurium (http://tellurium.analogmachine.org/), COPASI (http://copasi.org/), etc.

### 1.2 Installation

0. If you are running the docker image "pleiades", please go directly to the section "Preamble".
1. To install, please follow one of the following steps:
    1. Install the docker image "pleiades" using `docker pull networkbiolab/pleiades`. The container is based on the Anaconda3 software and it installs Atlas, and the stochastic simulators BNG2, NFsim, KaSim, and Stochkit. After building the image, please run the container with `docker run --detach --publish 10000:8888 networkbiolab/pleiades`, and go to `localhost:10000` in your preferred browser. The required password is `pleiades`.
    2. Download or clone the Github repository from https://github.com/networkbiolab/pleiades with `git clone https://github.com/networkbiolab/pleiades foo` (where `foo` is an absolute or relative path). Then, you could build the docker image with `docker build foo --tag pleiades` and run it with `docker run --detach --publish 10000:8888 pleiades`. Finally, go to `localhost:10000` in your preferred browser. The required password is `pleiades`.
    3. Install with pip3: `sudo -H python3 -m pip install pleiades` or `python3 -m pip install pleiades --user`. Pleiades is a meta-package that install Atlas (the rule-based modeller), Pleione (a genetic algorithm for parameter calibration of RBMs, compatible with SLURM), Alcyone (to perform identifiability analysis of parameters), and Sterope (to perform sensitivity analysis of parameters in kappa RBMs, compatible with SLURM). You should install, configure, and run the jupyter notebook on your own: example `sudo -H pip3 install jupyter && nohup python3 -m jupyter notebook --port=8888 --no-browser --port-retries=0 > /dev/null 2> /dev/null &`.
    4. Download or clone the Github repository from https://github.com/networkbiolab/atlas with `git clone https://github.com/networkbiolab/atlas foo` (where `foo` is an

absolute or relative path). Requisites must be fulfilled manually with pip3: `sudo -H python3 -m pip install pandas pysb pythoncyc jupyter seaborn` or `python3 -m pip install pandas pysb pythoncyc jupyter seaborn --user`.

### 1.3 Objectives

1. Get metabolic data from two species: enzyme names, substrates, products, and location of enzymes.
2. Merge the model into one
3. Simulate and plot

### 1.4 Preamble: load *Atlas*

[2]:
```
# testing source code
# required if atlas was cloned from GitHub and this notebook is executed from␣
 ↪the tutorial directory.
import sys
sys.path.append("..")

import atlas_rbm.atlas as atlas
import atlas_rbm.utils as utils
import atlas_rbm.export as export
import atlas_rbm.simulation as simulation
```

[3]:
```
utils.checkPathwayTools()
```

```
PathwayTools is running. Available PGDB are: YEAST, PAER208964, PABY272844,
GCF_000013425, CORYNE, BSUB, META, ECOLI
```

[3]:
```
True
```

[4]:
```
utils.execPToolsDocker('ptools-v22')
# execute this inside the docker will fail.
# Please, execute `docker run --rm -d --network host ptools-v24` in a terminal
```

```
Doing nothing since PathwayTools is running.
```

### 1.5 Getting data to model metabolism

[5]:
```
import pythoncyc
# %time network = utils.metabolicNetwork.FromEnzymeList('ECOLI', pythoncyc.
 ↪select_organism('ECOLI').all_transporters())
# %time utils.metabolicNetwork.expand_network(network,␣
 ↪'ecocyc-v22-tps-cytoscape.txt')
# network.to_csv('ecoli-tps-v22.txt', sep = '\t', index = False)
%time atlas.construct_model_from_metabolic_network('ecoli-tps-v22.txt',␣
 ↪noObservables=True, noInitials=True, toFile = 'model-ecoli-transporters.py')
```

```
It was found duplicated reaction names in the network.
Please check the conflicting_reactions.txt and correct them if necessary.
CPU times: user 425 ms, sys: 56.2 ms, total: 481 ms
Wall time: 480 ms
```

[6]:
```
import pythoncyc
# %time network = utils.metabolicNetwork.FromEnzymeList('ECOLI', pythoncyc.
 ↪select_organism('ECOLI').all_enzymes())
# %time utils.metabolicNetwork.expand_network(network,␣
 ↪'ecocyc-v22-enz-cytoscape.txt')
# network.to_csv('ecoli-enz-v22.txt', sep = '\t', index = False)
%time atlas.construct_model_from_metabolic_network('ecoli-enz-v22.txt',␣
 ↪noObservables=True, noInitials=True, toFile = 'model-ecoli-enzymes.py')
```

```
It was found duplicated reaction names in the network.
Please check the conflicting_reactions.txt and correct them if necessary.
CPU times: user 1.05 s, sys: 41.1 ms, total: 1.09 s
Wall time: 1.08 s
```

[7]:
```
import pythoncyc
# %time network = utils.metabolicNetwork.FromEnzymeList('BSUB', pythoncyc.
 ↪select_organism('BSUB').all_transporters())
# %time utils.metabolicNetwork.expand_network(network,␣
 ↪'bsubcyc-v22-tps-cytoscape.txt')
# network.to_csv('bsub-tps-v22.txt', sep = '\t', index = False)
%time atlas.construct_model_from_metabolic_network('bsub-tps-v22.txt',␣
 ↪noObservables=True, noInitials=True, toFile = 'model-bsub-transporters.py')
```

```
It was found duplicated reaction names in the network.
Please check the conflicting_reactions.txt and correct them if necessary.
CPU times: user 47.6 ms, sys: 11.4 ms, total: 59 ms
Wall time: 57.6 ms
```

[8]:
```
import pythoncyc
# %time network = utils.metabolicNetwork.FromEnzymeList('BSUB', pythoncyc.
 ↪select_organism('BSUB').all_enzymes())
# %time utils.metabolicNetwork.expand_network(network,␣
 ↪'bsubcyc-v22-enz-cytoscape.txt')
# network.to_csv('bsub-enz-v22.txt', sep = '\t', index = False)
%time atlas.construct_model_from_metabolic_network('bsub-enz-v22.txt',␣
 ↪noObservables=True, noInitials=True, toFile = 'model-bsub-enzymes.py')
```

```
It was found duplicated reaction names in the network.
Please check the conflicting_reactions.txt and correct them if necessary.
CPU times: user 605 ms, sys: 20.3 ms, total: 626 ms
Wall time: 623 ms
```

[9]:
```
ecoli_tps = utils.read_network('ecoli-tps-v22.txt')
ecoli_enz = utils.read_network('ecoli-enz-v22.txt')
bsub_tps = utils.read_network('bsub-tps-v22.txt')
bsub_enz = utils.read_network('bsub-enz-v22.txt')

import pandas
network = pandas.concat([ecoli_tps, ecoli_enz, bsub_tps, bsub_enz]) # reset␣
 ↪index is optional
network
```

[9]:

| | STRAIN | GENE OR COMPLEX | ENZYME LOCATION | REACTION | SUBSTRATES | PRODUCTS | FWD_RATE | RVS_RATE |
|---|---|---|---|---|---|---|---|---|
| 0 | ECOLI | ABC-10-CPLX | inner membrane | ABC-10-RXN | ATP,FERRIC-ENTEROBACTIN-COMPLEX,WATER | ADP,Pi,FERRIC-ENTEROBACTIN-COMPLEX,PROTON | 1.0 | 0.0 |
| 1 | ECOLI | ABC-11-CPLX | inner membrane | ABC-11-RXN | ATP,Ferric-Hydroxamate-Complexes,WATER | ADP,Pi,Ferric-Hydroxamate-Complexes | 1.0 | 0.0 |
| 2 | ECOLI | ABC-11-CPLX | inner membrane | TRANS-RXN-297 | CPD0-621,ATP,WATER | CPD0-621,ADP,Pi,PROTON | 1.0 | 0.0 |
| 3 | ECOLI | ABC-11-CPLX | inner membrane | TRANS-RXN-298 | CPD0-2241,ATP,WATER | CPD0-2241,ADP,Pi,PROTON | 1.0 | 0.0 |
| 4 | ECOLI | ABC-12-CPLX | inner membrane | ABC-12-RXN | ATP,GLN,WATER | ADP,Pi,GLN,PROTON | 1.0 | 0.0 |
| … | … | … | … | … | … | … | … | … |
| 1317 | BSUB | CPLX8J2-9 | cytosol | RIBONUCLEOSIDE-DIP-REDUCTI-RXN | Deoxy-Ribonucleoside-Diphosphates,Ox-Thioredox… | Ribonucleoside-Diphosphates,Red-Thioredoxin | 0.0 | 1.0 |
| 1318 | BSUB | CPLX8J2-9 | unknown | CDPREDUCT-RXN | DCDP,Ox-Thioredoxin,WATER | CDP,Red-Thioredoxin | 0.0 | 1.0 |
| 1319 | BSUB | CPLX8J2-9 | cytosol | CDPREDUCT-RXN | DCDP,Ox-Thioredoxin,WATER | CDP,Red-Thioredoxin | 0.0 | 1.0 |
| 1320 | BSUB | CPLX8J2-9 | unknown | UDPREDUCT-RXN | DUDP,Ox-Thioredoxin,WATER | UDP,Red-Thioredoxin | 0.0 | 1.0 |
| 1321 | BSUB | CPLX8J2-9 | cytosol | UDPREDUCT-RXN | DUDP,Ox-Thioredoxin,WATER | | | |

```
1321                          UDP,Red-Thioredoxin       0.0       1.0

[4848 rows x 8 columns]
```

[10]:
```
%time atlas.construct_model_from_metabolic_network(network, noObservables=True,␣
↪noInitials=True, toFile = 'model-combined-ecoli-bsub.py')
```

```
It was found duplicated reaction names in the network.
Please check the conflicting_reactions.txt and correct them if necessary.
CPU times: user 1.57 s, sys: 83.9 ms, total: 1.65 s
Wall time: 1.65 s
```

[ ]:
```

```