
TP Réseaux Métaboliques

**Ludovic COTTRET
Fabien JOURDAN**

Introduction

Une expérience métabolomique sur la levure a permis d'identifier un ensemble de métabolites d'intérêt. Ces métabolites se présentent sous la forme d'une liste de noms dans un fichier texte. Pour interpréter ces résultats, vous souhaitez les remettre dans le contexte du réseau métabolique.

Le plus souvent, les métabolites identifiés expérimentalement se présentent sous la forme de noms plus ou moins normés. Le but est ici de convertir ces noms en identifiants uniques de bases de données. Ces identifiants vont nous permettre d'intégrer sans ambiguïté les métabolites identifiés dans les réseaux métaboliques construits à partir de ces données.

Ce TP a pour objectif de vous aider dans les premiers pas de cette interprétation.

En particulier, nous allons aborder les points suivants :

- ✦ retrouver dans les bases de données les identifiants utilisables pour une étude réseau
- ✦ mapper les données sur un réseau métabolique
- ✦ identifier les sous réseaux mettant en œuvre les métabolites identifiés
- ✦ générer des représentations pouvant être utilisées dans des publications

Ces quatre étapes seront abordées en utilisant trois plateformes :

- KEGG
- BioCyc
- MetExplore + Cytoscape

Pour mettre en œuvre cette formation nous prendrons l'exemple d'un article de métabolomique étudiant l'exposition de la levure au cadmium.

Dans l'article, nous utiliserons les métabolites identifiés dont le ratio de l'intensité est supérieur ou égal à 1.9 et inférieur ou égal à 0.6.

Madalinski, G., Godat, E., Alves, S., Lesage, D., Genin, E., Levi, P., Labarre, J., et al. (2008). Direct introduction of biological samples into a LTQ-Orbitrap hybrid mass spectrometer as a tool for fast metabolome analysis. *Analytical chemistry*, 80(9), 3291-303.

Consignes

Légende :

Les questions principales du TP sont en **rouge**.

Les questions en **bleu** sont optionnelles.

Rapport :

Le rapport attendu devra contenir :

- Les réponses aux questions
- Des captures d'écrans
- Un fichier avec les métabolites et leurs identifiants dans les différentes bases

Conseil préliminaire :

Lors de ce TP nous allons utiliser une liste de métabolites à laquelle nous allons associer des identifiants et des valeurs.

► **Créez, dans un fichier, un tableau contenant la liste des métabolites.**



KEGG: <http://www.genome.jp/kegg/>

1. Retrouver dans KEGG les identifiants des métabolites

- ▶ Trouvez les identifiants KEGG des métabolites (e.g. L-Arginine: C00062)
- ▶ Pour un métabolite donné, plusieurs choix peuvent être possibles. Pourquoi ?

2. Mapper les données sur les cartes KEGG

Nous allons chercher dans quelles voies métaboliques on retrouve les métabolites identifiés lors de l'expérience sur la levure (*Saccharomyces cerevisiae*).

KEGG permet à la fois de colorer des métabolites (et/ou des réactions) sur une vue globale et des voies métaboliques individuelles pour un organisme donné.

- ▶ Réalisez ce mapping pour les métabolites identifiés:
 - dans la vue globale (récupérer l'image l'intégrer au rapport).
 - dans des voies où des métabolites apparaissent (récupérer l'image pour les 2 voies faisant intervenir le plus de métabolites et l'intégrer au rapport)
- ▶ Faire la même chose en mettant d'une couleur les métabolites dont le ratio est inférieur à 1 et d'une autre couleur les métabolites dont le ratio est supérieur à 1
- ▶ Certaines réactions sont colorées en vert. Pourquoi ?
- ▶ En quoi ce passage au métabolisme permet d'éliminer des métabolites (e.g. D-arginine) ?

3. Recherche de sous-réseau d'intérêt dans KEGG en utilisant PathComp

KEGG propose des outils pour analyser les réseaux métaboliques. En particulier l'outil PathComp permet de calculer un chemin (simple) entre deux métabolites.

- ▶ Utilisez l'outil Pathcomp pour trouver des chemins entre méthionine et glutathion.
- ▶ Récupérez l'image représentant le chemin le plus court et l'intégrer au rapport.
- ▶ Ce chemin métabolique vous paraît-il vraisemblable biologiquement ? Pourquoi ?
- ▶ A quoi sert le paramètre "cut off length" ?
- ▶ Que constatez-vous quand on augmente ce cut off length ?

4. Recherche de sous-réseau d'intérêt dans KEGG en utilisant MetaRoute

<http://www-bs2.informatik.uni-tuebingen.de/services/MetaRoute/>

Les chemins trouvés par pathComp peuvent manquer de réalisme, notamment à cause de la présence des cofacteurs qui court-circuitent les chemins. MetaRoute calcule les chemins possibles en prenant en compte les transferts d'atomes de carbones potentiels entre les métabolites.

- ▶ Recherchez le chemin entre méthionine et glutathion. Mettre une image du premier chemin dans votre rapport.
- ▶ Pourquoi le premier chemin n'est pas aussi court que le premier calculé par pathcomp ?

BIOCYC: <http://biocyc.org/>

5. Rechercher les identifiants BioCyc

- ▶ Trouvez dans BioCyc la base correspondant à la levure



- ▶ Trouvez les identifiants des métabolites dans la base (e.g. ARG pour l'arginine).
- ▶ Pourquoi, contrairement à KEGG n'a-t-on plus qu'une possibilité pour l'arginine ?

6. Mapping des métabolites sur les réseaux BioCyc

- ▶ Utilisez l'outil « Cellular overview » pour visualiser les métabolites dans le réseau (faire une capture d'écran et l'intégrer au rapport).

MetExplore: <http://www.metexplore.fr>

Contrairement à KEGG ou BioCyc, MetExplore permet d'analyser les données dans le contexte du réseau métabolique.

Le réseau métabolique de la levure (nous utiliserons celui venant de Biocyc) est stocké dans la base de données de MetExplore.

ATTENTION : Nous allons utiliser la version de MetExplore en développement:

<http://metexplore.toulouse.inra.fr/metexploreTest/>

7. Mapping sur les réseaux de MetExplore

- ▶ Faire le mapping en utilisant la fonction MetExplore.

En cliquant droit sur la liste de toutes les réactions, vous pouvez transférer toutes les réactions vers le "cart". Dans l'onglet "cart" sur la droite vous verrez apparaître toutes ces réactions.

En cliquant sur le bouton :  vous lancerez Cytoscape en version Java Webstart.

- ▶ Afficher le réseau dans Cytoscape, choisir le bon algorithme de dessin (faire une image et l'intégrer au rapport).

Remarque: changez le style visuel en utilisant : "indentificationVisualStyle"

Conseil: sauvez votre session Cytoscape régulièrement !

8. Recherche de sous-réseaux avec MetExplore

Utilisez le plugin MetExplore pour extraire (attention, par défaut le calcul des chemins se fait entre les sommets identifiés. On peut également faire ce calcul sur une sélection de sommets) :

- ▶ Le sous réseau "highest path" des métabolites sous (ratio<1) et sur exprimés (ratio>1)
- ▶ Le sous réseau "shortest path" des métabolites sous (ratio<1) et sur exprimés (ratio>1)
- ▶ Quelles sont les principales différences entre les résultats obtenus avec chaque algorithme ?
- ▶ Pourquoi ?

Remarque : les calculs peuvent prendre du temps !

9. Créer le sous-réseau correspondant à la Glycolyse ("glycolysis")

- ▶ Sélectionnez tous les nœuds correspondant à cette fonction en utilisant l'outil de recherche "Enhanced search"
- ▶ Créez un nouveau sous-réseau correspondant à la glycolyse (faire une image et l'intégrer au rapport).

10. Propriétés topologiques du réseau

Le plugin "Network analysis" est dédié à l'analyse de graphes. Il permet de calculer des mesures sur un réseau.

- ▶ Récupérez le graphique représentant la distribution du degré des sommets (en échelle normale et pas log log). (faire une image et l'intégrer au rapport).
- ▶ Comment expliquer cette distribution ?
- ▶ Listez les 7 métabolites de plus fort degrés.
- ▶ Pourquoi ces métabolites sont les plus connectés ?

Remarque : les calculs peuvent prendre du temps !



11. Créer un sous-réseau à partir d'une sélection

A partir des métabolites identifiés dans l'expérience :

► Créer le sous réseau des sommets à distance 1 (les voisins), 2 (les voisins des voisins), 3 (...) et 4 (...). Pour chacun, faire une image et l'intégrer au rapport.

12. Filtrer un réseau

► Clonez le graphe

► Supprimez les 7 métabolites de plus fort degrés

La densité d'un graphe peut être calculée comme le rapport entre le nombre de sommets et le nombre d'arêtes du graphe.

► Quelle est la densité du graphe : avant suppression des sommets ? Après suppression des sommets ?

► Faites le même travail que la question 11 sur le graphe filtré.

Références

Sites web

Catalogues de ressources :

PathGuide : <http://www.pathguide.org>

SBML software guide : http://sbml.org/SBML_Software_Guide

Bases de données génomes/métabolisme :

BioCyc : <http://biocyc.org/>

KEGG : <http://www.genome.jp/kegg/>

Mapping de métabolites :

MetExplore : <http://metexplore.toulouse.inra.fr>

Essentiellement sur des données BioCyc.

Masstrix : <http://metabolomics.helmholtz-muenchen.de/masstrix2/>

Mapping de masses sur des données KEGG

Extraction de sous-réseaux

MetExplore : <http://metexplore.toulouse.inra.fr>

PathComp (Kegg) : <http://www.genome.jp/tools/pathcomp/>

MetaRoute (Données KEGG) :

<http://www-bs2.informatik.uni-tuebingen.de/services/MetaRoute/>

Outils de visualisation

Cytoscape : <http://www.cytoscape.org/>

MetExplore : <http://metexplore.toulouse.inra.fr>

Cours et documentations

<https://sites.google.com/site/lcottret/teaching>

Sur ce site, vous trouverez des supports de cours sur les bases de données métaboliques et la modélisation du réseau.

Vous trouverez également parmi les fichiers attachés deux guides pour BioCyc et KEGG, qui même s'ils datent un peu, vous permettront de faire vos premiers pas dans ces sites et d'en voir la possibilité.

<http://pid.nci.nih.gov/PID/2007/071009/full/pid.2007.2.shtml>

KEGG Primer: An Introduction to Pathway Analysis Using KEGG.

Références

Kanehisa et al. KEGG for integration and interpretation of large-scale molecular data sets. NAR



Database issue 2012.

Caspi et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. NAR, 2009.

Cottret et al. *MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks*. NAR web server issue, 2010.

Suhre et al. MassTRIX: mass translator into pathways. NAR web server issue, 2008.

