Analysis of OMICS data in the context of metabolic networks

Nathalie Poupin, MeX, Toxalim
INRA Toulouse, France
nathalie.poupin@inra.fr
• **How to make sense of OMICS data?** Why do we need metabolic networks & modelling strategies?

• Metabolic networks to gather metabolic knowledge and integrate OMICS data

• Modeling global metabolism – **Graph-based approaches**

• Modeling global metabolism – **Flux-based approaches**
Impact of perinatal exposure to low doses of bisphenol A

What is the biochemistry acting behind the scene?

Use of metabolic network models for contextualization of omic data

What are the metabolic processes involved?

Cabaton NJ, et al. 2013
Environ. Health Perspect. 121:586–93

Metabolites Brain PND21
Cholines ↓
Glutamate ↓
Glutamine ↑
Glycine ↑
Aspartic Acid ↑
GABA ↓

How to interpret metabolic fingerprints?

Impact of perinatal exposure to low doses of bisphenol A

Birth
F0 (mothers)
F1 mice
perinatal exposure
Males at PND 21
Brain, Liver, Serum

Metabolites
- Cholines
- Glutamate
- Glutamine
- Glycine
- Aspartic Acid
- GABA

Use of metabolic network models for contextualization of omic data
Untargeted observations require a more holistic analysis.

“Rationality is bounded when it falls short of omniscience. And the failures of omniscience are largely failures of knowing all the alternatives, uncertainty about relevant exogenous events, and inability to calculate consequences.”


Metabolic network

= succession of chemical reactions, whose main functions are to produce energy from available resources, often gathered in metabolic pathways (that perform a specific function or transformation, e.g. glycolysis)

Genome scale metabolic network reconstructions

Biochemical reactions known to take place in a target organism & associated genes

<table>
<thead>
<tr>
<th>list of genes</th>
<th>list of reactions</th>
<th>list of metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6524) or (6526)</td>
<td>GLCt4 $\text{Na}_e + \text{gluc}_c \leftrightarrow \text{Na}_c + \text{gluc}_c$</td>
<td>$\text{gluc}_c$</td>
</tr>
<tr>
<td>(3098)</td>
<td>HEX1 $\text{gluc}_c + \text{ATP}_c \rightarrow \text{ADP}_c + \text{g6p}_c$</td>
<td>$\text{gluc}_c$ $\text{Na}_e$ $\text{Na}_c$</td>
</tr>
<tr>
<td>(2821)</td>
<td>PGI $\text{g6p}_c \rightarrow \text{f6p}_c$</td>
<td>$\text{ATP}_c$</td>
</tr>
<tr>
<td>(2539)</td>
<td>G6PDH2r $\text{g6p}_c + \text{nadp}_c \rightarrow \text{6pgl}_c + \text{nadph}_c$</td>
<td>$\text{ADP}_c$</td>
</tr>
<tr>
<td></td>
<td>GLCter $\text{gluc}_c \leftrightarrow \text{gluc}_r$</td>
<td>$\text{g6p}_c$ $\text{f6p}_c$ $\text{6pgl}_c$ $\text{nadp}_c$ $\text{nadph}_c$ $\text{gluc}_r$</td>
</tr>
</tbody>
</table>

→ stoichiometric associations
→ mathematically structured knowledge base
→ Generic reconstructions generated from annotated genomes
→ Published for many organisms

Thiele I, Palsson BØ. Nat. Protoc. 2010; 5:93–121
Metabolic network representations
The global Human metabolic network Recon 3D¹

¹Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A et al. (2018)

8 399 metabolites (4140 uniques)
13 543 reactions
3 697 genes
Metabolic networks for multi-omic interpretation

OMIC data

transcriptomic (DNA microarray)

proteomic (MS)

metabolomic (NMR & MS)
Analysing metabolic networks: **Graphs vs. fluxes**

**Graph based analysis**

→ understand the organisation of the system, study some structural properties such as connectivity

**Find possible pathways between metabolites of interest**

---

**Flux based analysis**

→ understand the behavior of the system under specific conditions, identify the most active pathways

**Predict metabolic fluxes, cell growth, drug targets...**
Graph-based analysis: looking for metabolic paths

Problem complexity: finding paths from glucose to pyruvate

Metabolic path = succession of reactions connecting metabolites.
Graph-based analysis: looking for metabolic paths

Problem complexity: finding paths from glucose to pyruvate

Metabolic path = succession of reactions connecting metabolites.

In the whole network:
500,000 possible paths between glucose and pyruvate!

→ need for graph algorithms to compute & select paths between metabolites in the network.

Shortest path vs. lightest path

the shortest path is not always the most relevant!
Shortest path vs. lightest path

**Shortest path**

**Lightest path**
Graph-based analysis: Interpreting metabolic fingerprints

How can we connect metabolites?
Finding the cascade of biochemical reactions that connect modulated metabolites

Sub-network extraction
Graph algorithms & visualization tools to mine these large networks

Mechanistic interpretation
Frainay & Jourdan 2017 Brief. In Bioinformatics

MetExplore: web server for network analysis of omics data

MetExplore Computational infrastructure for metabolic network analysis
Funding: ANR MetaboHub, H2020 Phenomenal
• Long lasting project established in 2009
• >400 registered users, >350 persons trained, >20 000 visits since 2009
• Shared platform in international projects

Functions:
• Database of metabolic networks
• Collaborative annotation of metabolic networks
• Import of omics data
• Visualization of metabolic networks
• Sub-network extraction (graph based computations)

L. Cottret et al. Nucleic Acids Res., 2010

www.metexplore.fr
Graphs vs. fluxes

Graph based analysis

- Structural / topological analysis
  - understand the organisation of the system, study some structural properties such as connectivity
  - Find possible pathways between metabolites of interest

Flux based analysis

- Dynamic / semi-quantitative analysis
  - understand the behavior of the system under specific conditions, identify the most active pathways
  - Predict metabolic fluxes, cell growth, drug targets...
Flux & Constraint-based analysis to contextualize metabolic networks

**CONTEXTUALIZATION**

= building TISSUE- or CONDITION-SPECIFIC METABOLIC MODELS

→ generic metabolic capacity
= an infinity of possible phenotypes

---

**Flux analysis approaches**

Objective = determining the activity of the reactions in a given context

principle = computing fluxes for all reactions in the network

**flux** = amount of a metabolite that is synthesized or consumed through a reaction per time unit

\( \text{mmol. g DW}^{-1}. \text{h}^{-1} \)

1) Which reactions are active?

\[ \text{flux} \neq 0 \iff \text{réactions « actives »} \]

\[ \text{flux} = 0 \iff \text{réactions « inactives »} \]

1 sub-network of active reactions = functional metabolic network for a given condition

2) Quantification of the activity of the different metabolic pathways

---

**global Human metabolic network Recon3D**

8 399 metabolites
13 543 reactions
3 697 genes

---

Flux & Constraint-based analysis to contextualize metabolic networks

**global Human metabolic network Recon3D**
- 8,399 metabolites
- 13,543 reactions
- 3,697 genes

→ **generic metabolic capacity** = an infinity of possible phenotypes

OMICS data

Constraint-based modeling

Objective = putting constraints on reactions fluxes based on omics data

CONTEXTUALIZATION = building TISSUE- or CONDITION-SPECIFIC METABOLIC MODELS

---


---

From Bordbar et al., 2014, Nature Reviews
Building tissue-specific models

Gene expression data obtained from different databases

Prediction of activity state for 644 reactions (average 408/tissue)

Reactions predicted to be active in 10 different tissues

Upregulated = predicted to be active, but not associated with highly expressed gene

Subnetwork representing the glycogen metabolism

GBE1 is predicted to be specifically active in liver

from Hyduke et al., Molecular BioSystem, 2013
Identification of metabolic modulations from exometabolomics data

1. Converting exometabolomic profiles to uptake/secretion profile

Example: *In vitro* cell cultures

Extraction of extracellular medium at 2 time points

**spectres (MS ou RMN)** exométabolomique

![Graph showing concentration changes over time](image)

 quantification des métabolites

![Diagram showing metabolite profiles](image)

calcul des quantités produites / consommées entre t0 et T24h

![Diagram showing production and consumption profiles](image)

conversion en contraintes in silico

* diminution → consommation → import forcé du milieu vers la cellule dans le modèle

** augmentation → production → export forcé du milieu vers la cellule dans le modèle

Example: *In vitro* cell cultures

Extraction of extracellular medium at 2 time points

**spectres (MS ou RMN)** exométabolomique

![Graph showing concentration changes over time](image)

 quantification des métabolites

![Diagram showing metabolite profiles](image)

calcul des quantités produites / consommées entre t0 et T24h

![Diagram showing production and consumption profiles](image)

conversion en contraintes in silico

* diminution → métabolite consommé

** augmentation → production
2. Computing intracellular metabolic fluxes & Identifying metabolic modulations

Identification of metabolic modulations from exometabolomics data

Réseau métabolique condition-spécifique

Réseau métabolique humain générique

intégration des données d'exométabolomique

calcul des flux intracellulaires

Identification & comparaison des modulations métaboliques selon les conditions

profil de production – consommation de métabolites
**Example: Revealing metabolic differences between 2 lymphoblastic leukemia cell lines**

**Experimental data**

- 2 lymphoblastic leukemia cell lines (Molt-4 & CCRF-CEM)
- Metabolomic analyses:
  - 75 extracellular metabolites detected
  - 2 ≠ time points (2h & 48h)

<table>
<thead>
<tr>
<th>Identified uptake / secreted metabolites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molt-4 cells</td>
</tr>
<tr>
<td>16 uptake</td>
</tr>
<tr>
<td>11 secreted</td>
</tr>
</tbody>
</table>

**Modeling**

- Computation of intracellular metabolic fluxes distributions
- Probable metabolic states for each model (probability distributions for fluxes)
Higher utilization of the oxidative phosphorylation by the Molt-4 model

→ experimentally supported by increased capacity for ROS detoxification in Molt-4 cells

Example: Revealing metabolic differences between 2 lymphoblastic leukemia cell lines

Take home messages

• Genome scale metabolic networks provide an holistic context ...
  ... but require **modeling** and **algorithms**

• **Graph-based analyses** are useful to understand the connections between metabolites and reactions

• **Flux-based analyses** can be used to predict the system behavior under specific conditions & predict biomarkers

• "**back-and-forth**" are essential between experimentation and model (for model construction, validation, refinement ...)

• **A model is necessarily a simplification of the biological complexity**: the interpretation and extrapolation are limited

"Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful"  *George EP Box*
Acknowledgments

Thank you for your attention!