Comparative Analysis of Amino Acid Assimilation Pathways using Hypergraphs

Aziz Mithani, Arantza Rico†, Rachel Jones†, Gail Preston† and Jotun Hein
mithani@stats.ox.ac.uk

Department of Statistics and †Department of Plant Sciences,
University of Oxford, South Parks Road, Oxford, UK

MBPP 2007
11 September 2007
Why *Pseudomonas* and *Xanthomonas*?

- Important pathogenic bacteria
- Occupy diverse niche
- Nutrient assimilation widely studied
- Sequences available for different strains

**Figure:** Halo blight of beans caused by *P. syringae* pv. phaseolicola (Source: The American Phytopathological Soc.)

**Figure:** Yellowing of banana plant leaves caused by *X. campestris* pv. musacearum (Reeder et. al, 2007)
### Motivation

Metabolic networks have diverse applications.

Metabolic networks can be used in many different types of studies:

<table>
<thead>
<tr>
<th>Pathway Analysis</th>
<th>Studying Gene Insertion Effects</th>
<th>Studying Knock-out Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Diagram" /></td>
<td><img src="image2" alt="Diagram" /></td>
<td><img src="image3" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Identifying Pathway Specific Reactions</th>
<th>Genome - Genome Comparisons</th>
<th>Ancestral Network Building</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image4" alt="Diagram" /></td>
<td><img src="image5" alt="Diagram" /></td>
<td><img src="image6" alt="Diagram" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Studying Host-Pathogen Interactions</th>
<th>Genome – ‘X’ome (Metabolome, Transcriptome, Proteome) Comparisons</th>
<th>Investigating Toxic Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image7" alt="Diagram" /></td>
<td><img src="image8" alt="Diagram" /></td>
<td><img src="image9" alt="Diagram" /></td>
</tr>
</tbody>
</table>
Problem
Commonly used representation is not adequate . . .

Metabolic networks are generally represented by directed graphs

- Nodes - metabolites
- Edges - reactions / enzymes

Limitations
- Cannot capture relationship between more than two metabolites in a reaction
- Ignores dependence between the metabolites
- Not ideal for representing multiple connections (reactions) between metabolites

Solution
- Use hypergraphs

Figure: The cysteine metabolism network in \textit{P. fluorescens} PfO-1 from KEGG represented as a graph.
Hypergraphs

Hypergraphs are ideal for representing metabolic networks . . .

Hypergraphs are generalisation of ordinary graphs.

- An edge (reaction) may connect any number of nodes or vertices (metabolites)

Advantages

- Allows enumeration of multiple reactions connecting metabolites
- Ideal for representing the relationships between multiple metabolites in a reaction
- Provide an intuitive approach to finding pathways between two metabolites

Figure: Hypergraph representation of the cysteine metabolism network in *P. fluorescens* PfO-1. Compounds such as ATP, AMP, H$_2$O not shown.
Hypergraphs

Hypergraphs are ideal for representing metabolic networks . . .

Hypergraphs are generalisation of ordinary graphs.

- An edge (reaction) may connect any number of nodes or vertices (metabolites)

**Advantages**

- Allows enumeration of multiple reactions connecting metabolites
- Ideal for representing the relationships between multiple metabolites in a reaction
- Provide an intuitive approach to finding pathways between two metabolites

![Hypergraph representation of the cysteine metabolism network in *P. fluorescens* PfO-1. Compounds such as ATP, AMP, H₂O not shown.](image)

Rahnuma: Pathway Analysis & Network Comparison Tool
A hypergraph based tool...

Rahnuma – someone who guides through the path (Urdu language)

The Tool

- Written in Java
- Represents networks as directed hypergraphs
- Uses advanced data structures to efficiently compute the pathways between the metabolites or group of metabolites
- Outperforms KEGG’s PathComp tool (Kanehisa, 2006) in many areas

Modules

1. Pathway prediction and analysis
2. Network comparison
3. Ancestral network building
How do plant pathogens differ from non-pathogens? Have pathogens lost some reactions?
**parameters**

**phylogeny**

```plaintext
(pae,
 (pen,ppu),
 ((pfo,pfl),
  ((psp,psb),
   pst))))
```
Phylogeny based network analysis

Studying networks on a phylogeny may provide clues about the gain or loss of functionalities at various levels . . .

Number of reaction at various levels of *Pseudomonas* phylogeny

- Includes reactions involved in amino acid assimilation and related pathways only

![Phylogenetic tree](image)
How do pathogens differ in nutrient assimilation pathways from non-pathogens?
Rahnuma: Phylogeny based Pathway Analysis

Job Detail
- Job Name: Pathway Analysis - Pseudomonas
- Email: mithani@stats.ox.ac.uk

Job Specification
- Analysis Options:
  - Type: Pathway prediction
  - Reactions exclusive to a Pathway
  - Reactions acting as bridges
- Mode: Phylogeny based
- Network Mode: All reactions
- Output Options:
  - Type: Tabular
  - Format: Text

Next
### Parameters

**Phylogeny**

```
(pae, 
  ((pen,ppu), 
    ((pfo,pfl), 
     ((psp,psb), 
      pst))))
```

### Pathway Specific Parameters

**Start Metabolites**

- **Show subset**

**End Metabolites**

- **Show subset**

**Mode**

- Connection
- Reaction

**Cutoff Length**

6

**Connection Option**

- Default Carbon Connections
- Default Nitrogen Connections
- Read from file

**Return Value**

- Pathway
- Reaction
Phylogeny based pathway analysis

Studying networks on a phylogeny may provide clue about the gain or loss of functionalities at various levels . . .

Pathways calculated from amino acid to the TCA cycle with cutoff length = 6.
Homoserine pathways
Show variability in *Pseudomonas* species...
Rahnuma: Pathway Prediction and Network Analysis Tool

Job Detail

Job Name *: Comparative Analysis - Pf01 vs DC3000
Email *: mithani@stats.ox.ac.uk

Job Specification

Analysis
- Pathway Analysis
- Comparative Analysis
- Network Analysis

Analysis Options
Type
- Standard comparison
- All but one comparison

Mode
- Organism based
- Phylogeny based

Comparison
- Pathway based
- Full network

Output Options
Type
- Tabular
- Descriptive

Format
- Text
- HTML

Next
Organism based comparative analysis
Rahnuma helps in explaining experimental results and identifying strain specific differences in assimilation pathways . . .

Minimum predicted pathway lengths to ammonia from selected amino acids are shown below along with the experimental results (Rico and Preston, submitted).

<table>
<thead>
<tr>
<th></th>
<th>P. fluorescens PfO-1</th>
<th>P. syringae DC3000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min Length</td>
<td>Biolog</td>
</tr>
<tr>
<td>Alanine</td>
<td>3</td>
<td>+</td>
</tr>
<tr>
<td>Arginine</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Cysteine</td>
<td>4</td>
<td>−/w</td>
</tr>
<tr>
<td>Glutamine</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Serine</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>Tryptophan*</td>
<td>2</td>
<td>−/w</td>
</tr>
</tbody>
</table>

*Tryptophan Pathway involves reaction R02722, a tryptophan synthesis reaction, catalysed by the enzyme tryptophan synthase (EC: 4.2.1.20). However, due to reversible nature of the reaction it may proceed in the reverse direction.
Are some plant pathogens metabolically reduced for growth in plant tissues?
Xylella versus Xanthomonas

Genome comparison can provide important insights into evolution of organisms ...

Full network comparison between the two sets of organisms highlights following differences.

<table>
<thead>
<tr>
<th>Species</th>
<th>Extra Reactions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xylella</td>
<td>2</td>
</tr>
<tr>
<td>Xanthomonas</td>
<td>139</td>
</tr>
</tbody>
</table>

- Includes reactions involved in amino acid assimilation and related pathways only

Extra Reactions in Xylella

<table>
<thead>
<tr>
<th>Rxn ID</th>
<th>Reaction</th>
<th>Enzyme</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>R00357</td>
<td>( \text{L-Aspartate} + \text{H}_2\text{O} + \text{Oxygen} \rightleftharpoons \text{Oxaloacetate} + \text{NH}_3 + \text{H}_2\text{O}_2 )</td>
<td>L-aspartate oxidase [1.4.3.16]</td>
<td>nadB</td>
</tr>
<tr>
<td>R01257</td>
<td>( \text{(S)-Malate} + \text{FAD} \rightleftharpoons \text{FADH}_2 + \text{Oxaloacetate} )</td>
<td>malate dehydrogenase [1.1.99.16]</td>
<td>yojH</td>
</tr>
</tbody>
</table>
**Xylella versus Xanthomonas**

Important pathways specific differences can be identified by comparing the genomes...

- **Xanthomonas**
  - Gain of tryptophan to nicotinate route – Nutritional environment?
  - Loss of aspartate to nicotinate route – Anti-virulence? Functional redundancy?

- **Xylella**
  - Gain of aspartate to nicotinate route – Niche colonisation?
How to efficiently develop pesticides that would target the pathogen enzymes and not the plant ones?
Rahnuma: Pathway Prediction and Network Analysis Tool

**Job Detail**
- **Job Name**: P. syringae vs A. thaliana
- **Email**: mithani@stats.ox.ac.uk

**Job Specification**

**Analysis**
- Pathway Analysis
- Comparative Analysis
- Network Analysis

**Analysis Options**
- **Type**: Standard comparison
- **Comparison**: Pathway based

**Output Options**
- **Type**: Tabular

**Mode**
- Organism based
- Phylogeny based

**Format**
- Text
- HTML
**P. syringae versus A. thaliana**

Important pathways specific differences can be identified by comparing the genomes...

- **GABA**
  - One of the most abundant amino acids in the plant apoplast
  - Key nutrient source for plant pathogenic fungi of tomato such as *Cladosporium fulvum* (Solomon and Oliver, 2001, 2002)

![GABA metabolic pathway diagram](image)

**P. syringae**
- \( \text{gabT}: \text{GABA} + 2\text{-Oxoglutarate} \leftrightarrow \text{Succinate semialdehyde} + \text{L-Glutamate} \)
- *P. syringae pv. tomato* contains 3 copies of \( \text{gabT} \)
  (Buell et. al, 2003)
- Interpreted as key route for GABA assimilation

**A. thaliana**
- No hits for \( \text{gabT} \) in *A. thaliana*
- Experiments suggest presence of GABA transaminase activity
  - Different than bacterial and animal version of \( \text{gabT} \)
  - Uses pyruvate and glyoxylate instead of oxoglutarate

- Develop chemicals to target bacterial/fungal GABA shunt and not that of plants
Which are the genes that if knocked out would result in all pathways being turned off between any given metabolites?
Rahnuma: Pathway Prediction and Network Analysis Tool

Job Detail

Job Name: in silico knockouts - P. syringae DC3000
Email: mithani@stats.ox.ac.uk

Job Specification

Analysis
- Pathway Analysis
- Comparative Analysis
- Network Analysis

Analysis Options
- Type: Pathway prediction
- Reactions exclusive to a Pathway
- Reactions acting as bridges

Mode
- Organism based
- Phylogeny based

Network Mode
- Individual
- Collective

Output Options
- Type: Tabular
- Descriptive

Format
- Text
- HTML

Next
Reactions acting as bridges in pathways
Searching pathways for bridge reactions may help in identifying genes for effective knock-outs . . .

Pathways calculated from amino acid to the TCA cyle with cutoff length = 6 in \textit{P. syringae} pv. tomato DC3000.

<table>
<thead>
<tr>
<th>Rxn ID</th>
<th>Reaction</th>
<th>Start Metabolite</th>
</tr>
</thead>
<tbody>
<tr>
<td>R02722</td>
<td>\text{L-Tryptophan + (2R)-2-Hydroxy-3-(phosphonooxy)-propanal + } \text{H}_2\text{O} \iff \text{L-Serine + Indoleglycerol phosphate}</td>
<td>\text{L-Tryptophan}</td>
</tr>
<tr>
<td>R00734</td>
<td>\text{3-(4-Hydroxyphenyl)pyruvate + L-Glutamate} \iff \text{L-Tyrosine + 2-Oxoglutarate}</td>
<td>\text{L-Tyrosine}</td>
</tr>
<tr>
<td>R01168</td>
<td>\text{Urocanate + NH}_3 \iff \text{L-Histidine}</td>
<td>\text{L-Histidine}</td>
</tr>
<tr>
<td>R02285</td>
<td>\text{L-Glutamate + Formamide} \iff \text{N-Formimino-L-glutamate + } \text{H}_2\text{O}</td>
<td>\text{L-Histidine}</td>
</tr>
<tr>
<td>R02288</td>
<td>\text{N-Formimino-L-glutamate + H} \iff \text{4-Imidazolone-5-propanoate + } \text{H}_2\text{O}</td>
<td>\text{L-Histidine}</td>
</tr>
<tr>
<td>R02914</td>
<td>\text{Urocanate + } \text{H}_2\text{O} \iff \text{4-Imidazolone-5-propanoate}</td>
<td>\text{L-Histidine}</td>
</tr>
<tr>
<td>R00751</td>
<td>\text{Glycine + Acetaldehyde} \iff \text{L-Threonine}</td>
<td>\text{L-Threonine}</td>
</tr>
</tbody>
</table>
Summary

Rahnuma is a versatile tool ...

Rahnuma can be used in many different types of studies

- Pathway Analysis
- Studying Gene Insertion Effects
- Studying Knock-out Effects
- Identifying Pathway Specific Reactions
- Genome - Genome Comparisons
- Ancestral Network Building
- Studying Host-Pathogen Interactions
- Genome – ‘X’ome (Metabolome, Transcriptome, Proteome) Comparisons
- Investigating Toxic Effects
The Team
and many more ...

Acknowledgements

Aziz Mithani (University of Oxford)