

# HotSpot Analysis of ProteoMetabolomics Data from *Staphylococcus aureus* Clinical Isolates



*Victor Torres, Irnov Irnov, Sophie Dyzenhaus,  
William Sause, Bo Shopsin MD,*

*Beatrix M. Ueberheide, Avantika Dhabaria,*

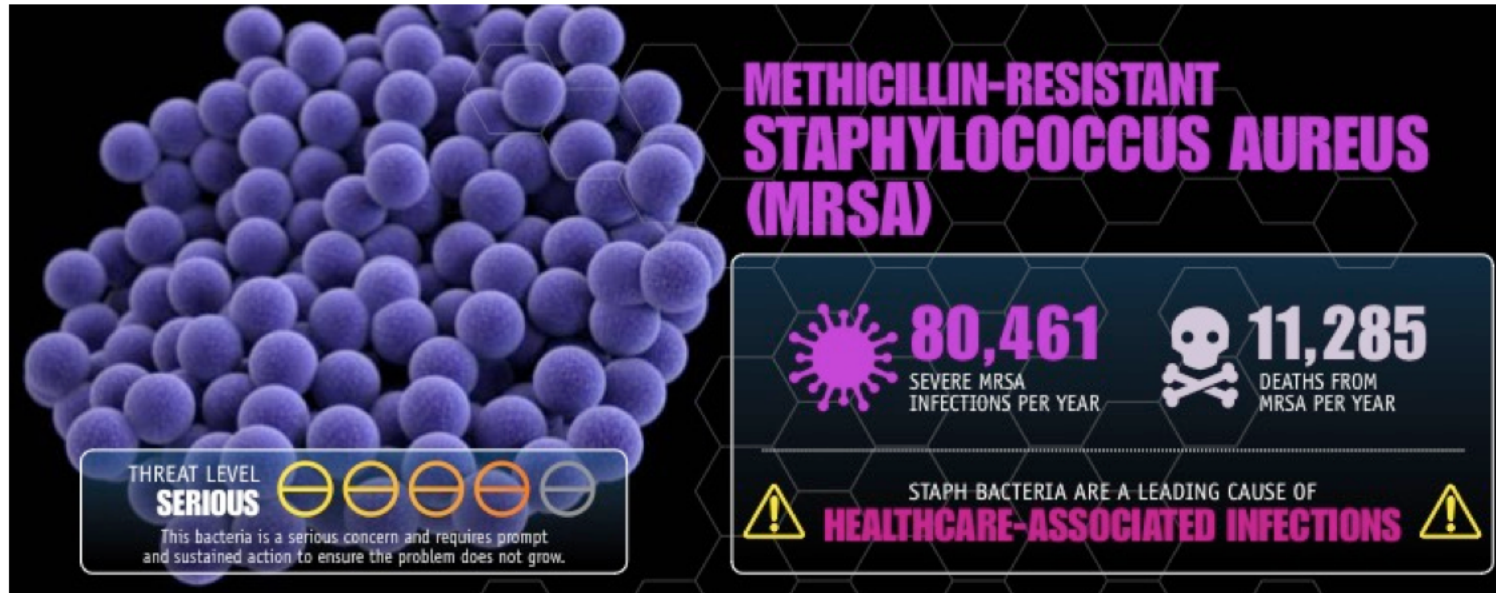
*Drew R. Jones, Rebecca Rose,*

*Bernard Delange, Manor Askenazi,*

*Christoph Henrich, Kai Fritzemeier*



# HA-MRSA: 'Superbug'

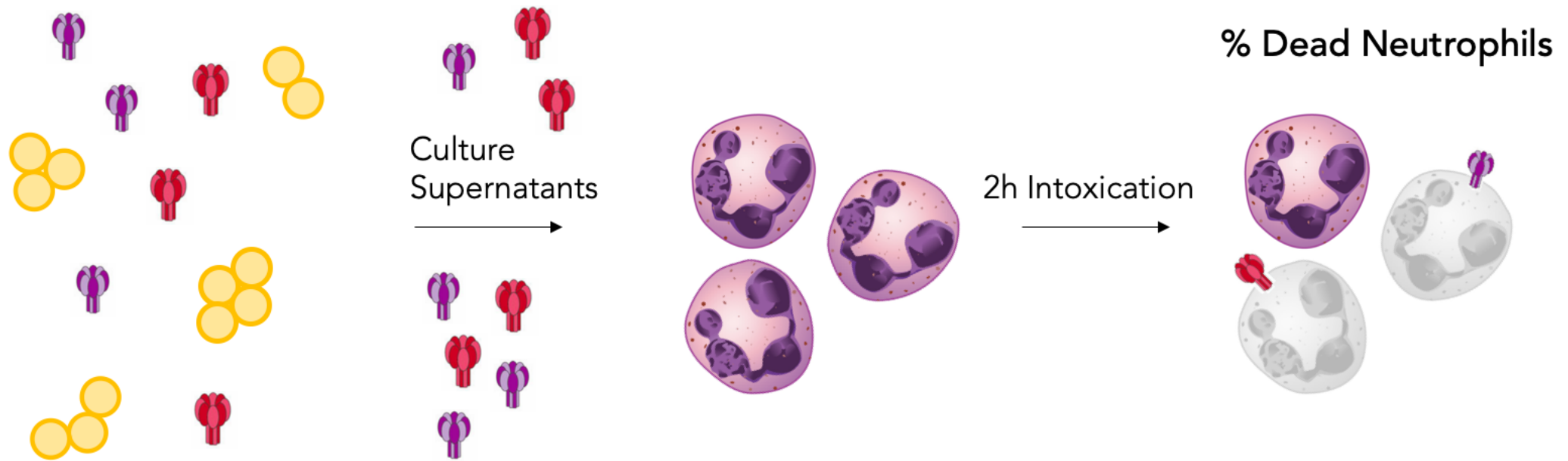


\$3.2-4.2 billion annually

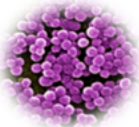
# CA-MRSA



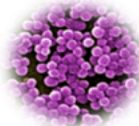
# How to Measure (Cytotoxic) Virulence:



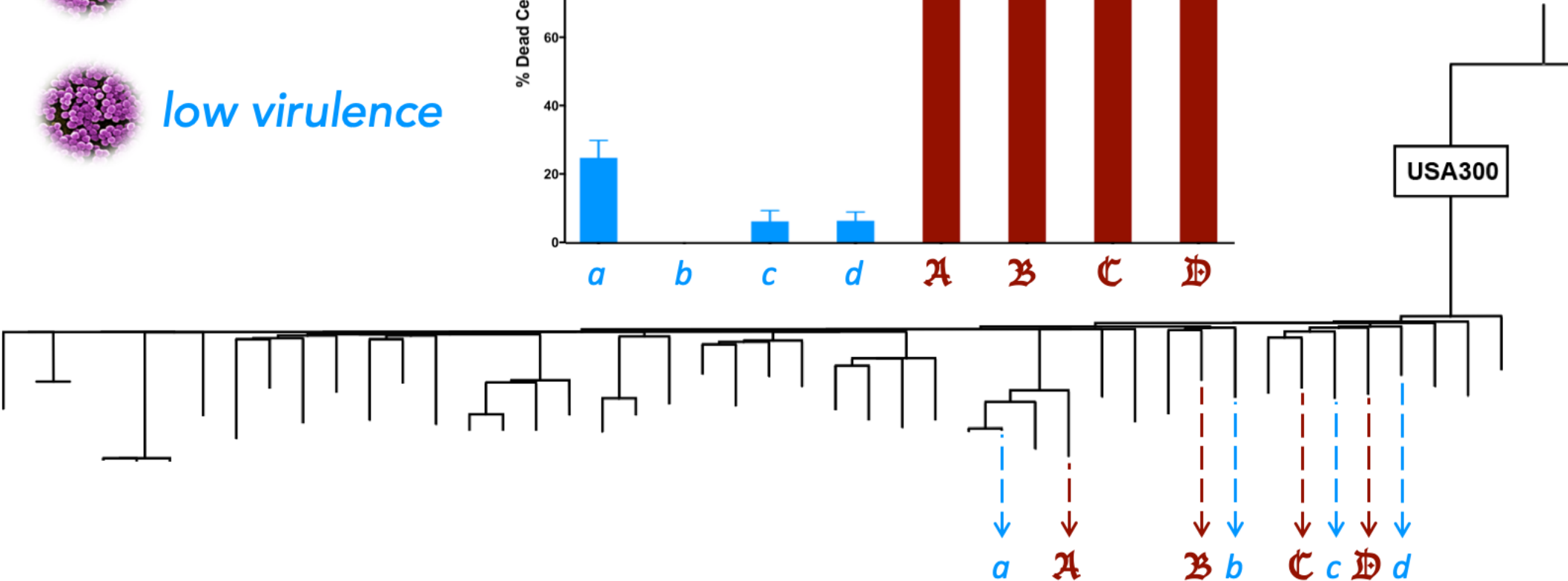
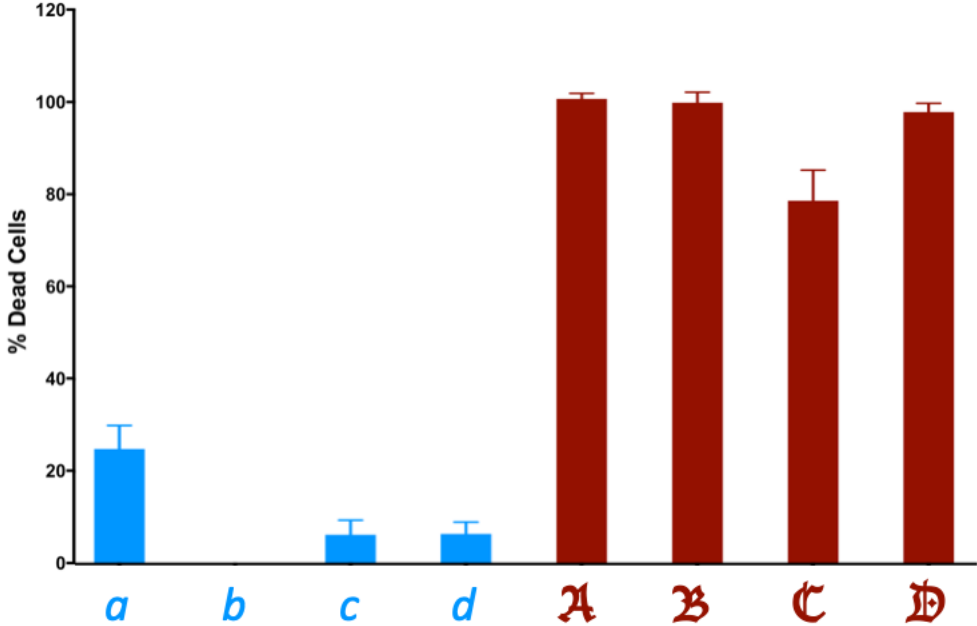
# We need more than Genomics:



*high virulence*



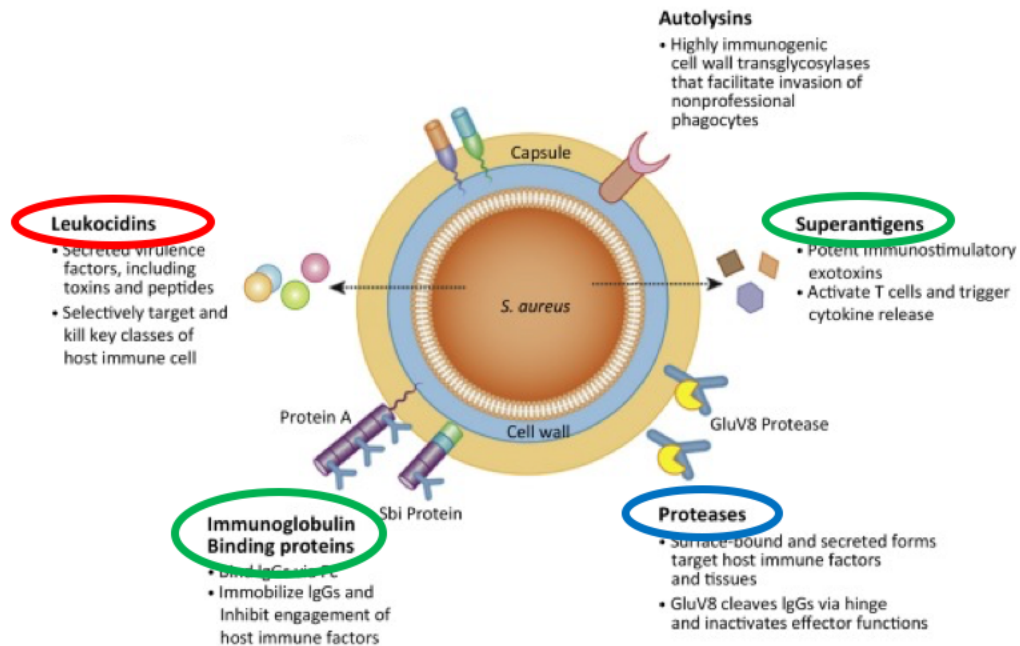
*low virulence*



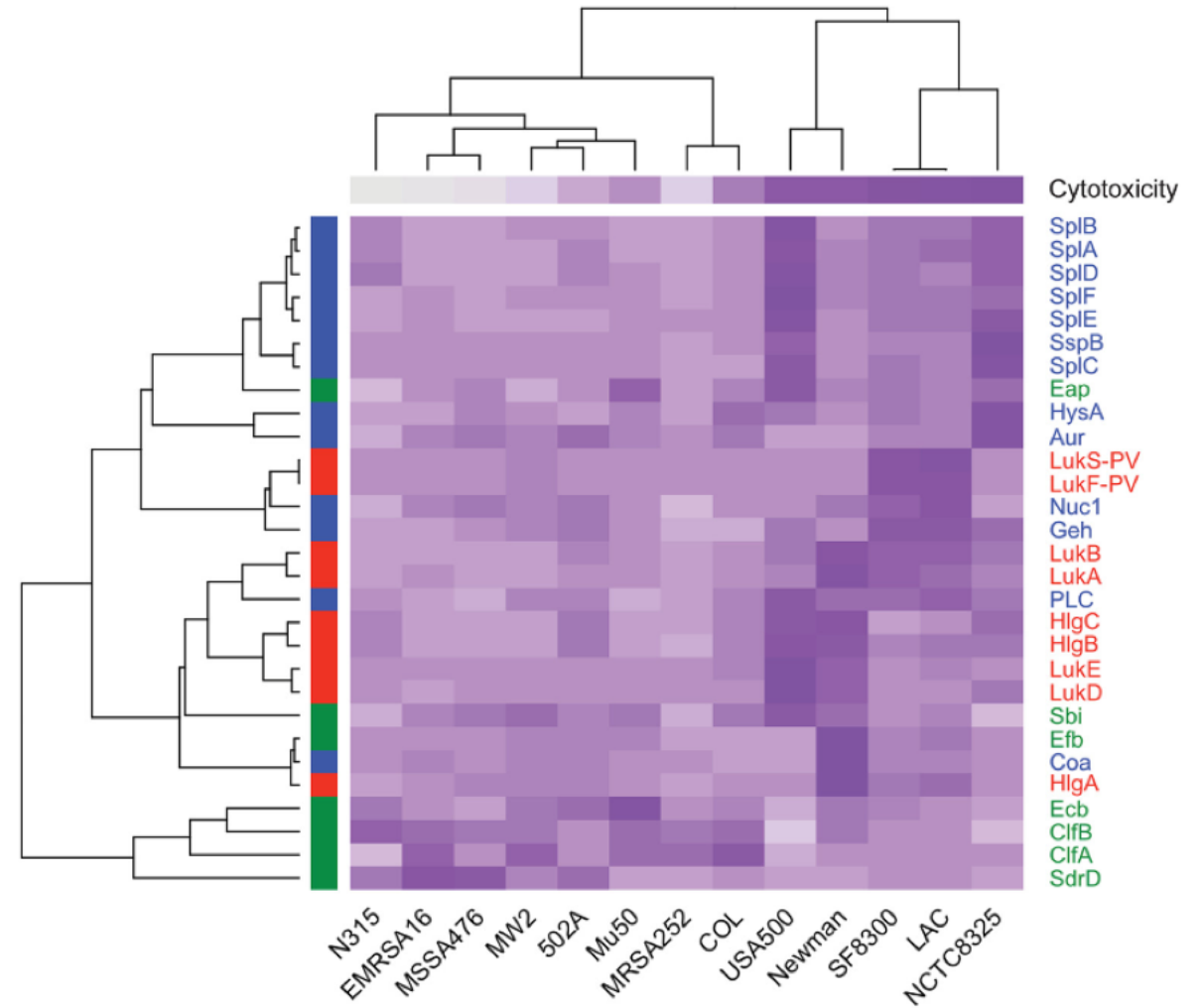
# Proteomics Can Help:

(particularly the proteomics of "virulence factors")

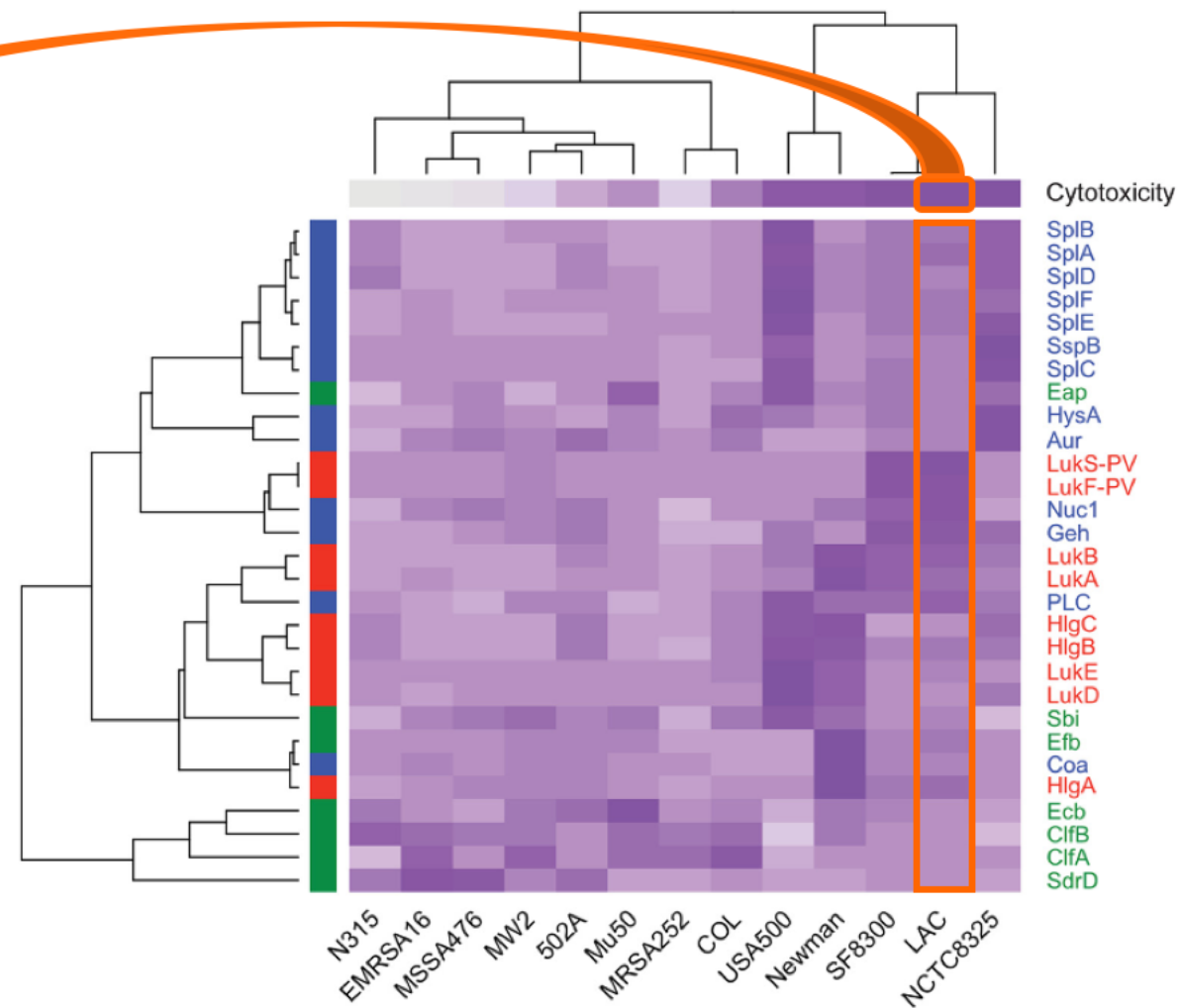
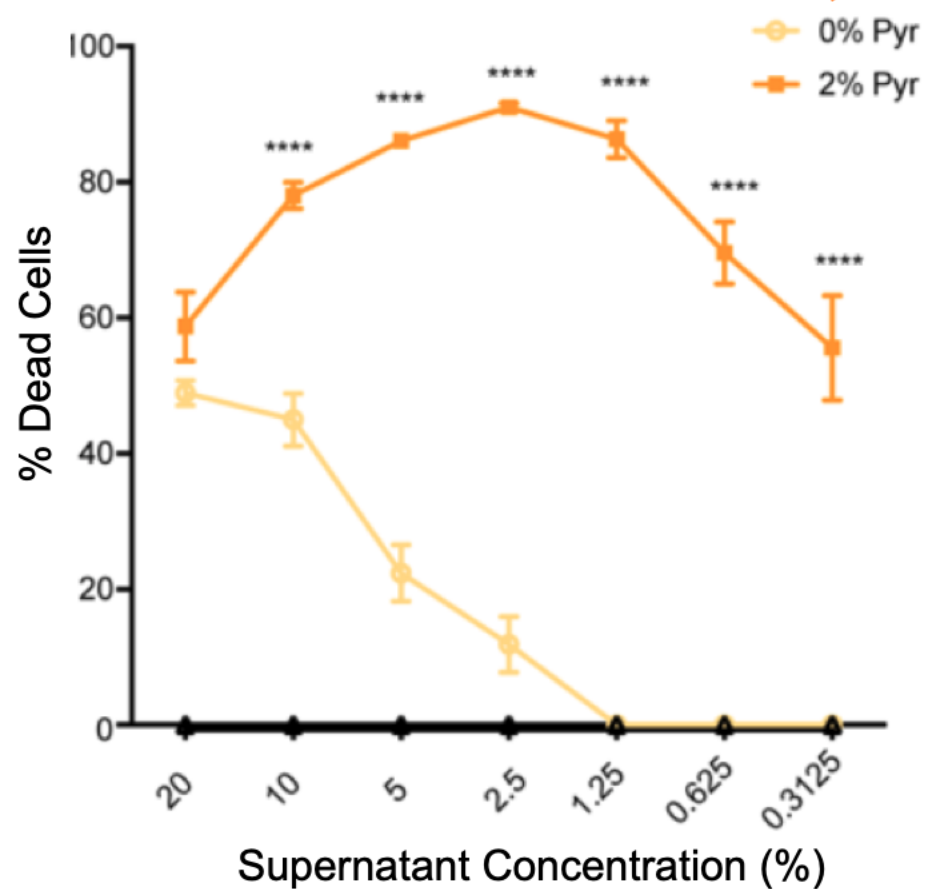
- **Exo-enzymes**
- **Cytotoxins**
- **Immunomodulators**



Trends in Pharmacological Sciences



# But so can **Metabolomics**:

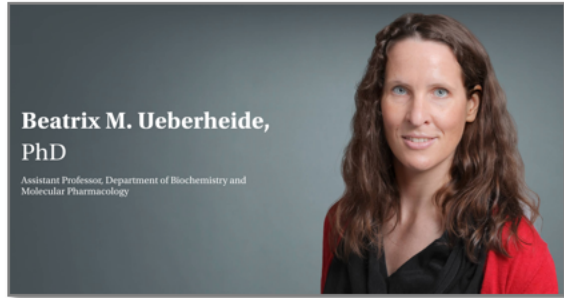




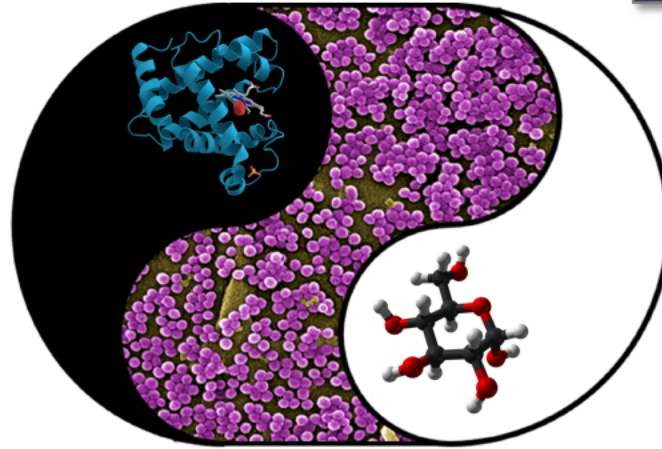
# NYU Langone Medical Center to Spearhead Multi-Institutional MRSA Research Funded by the National Institutes of Health



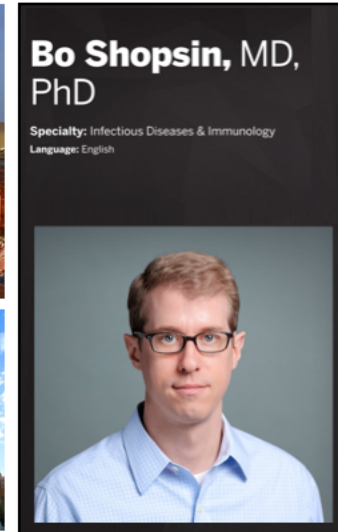
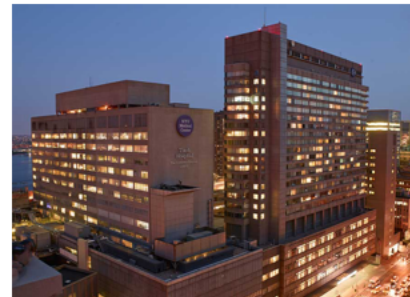
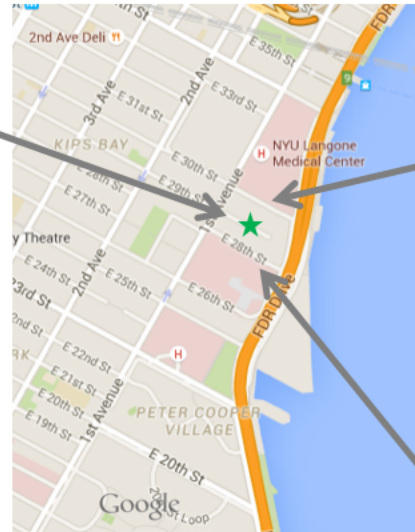
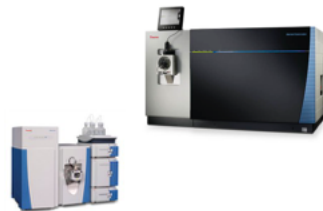
Victor J. Torres, PhD  
Associate Professor, Department of Microbiology



Beatrix M. Ueberheide, PhD  
Assistant Professor, Department of Biochemistry and Molecular Pharmacology



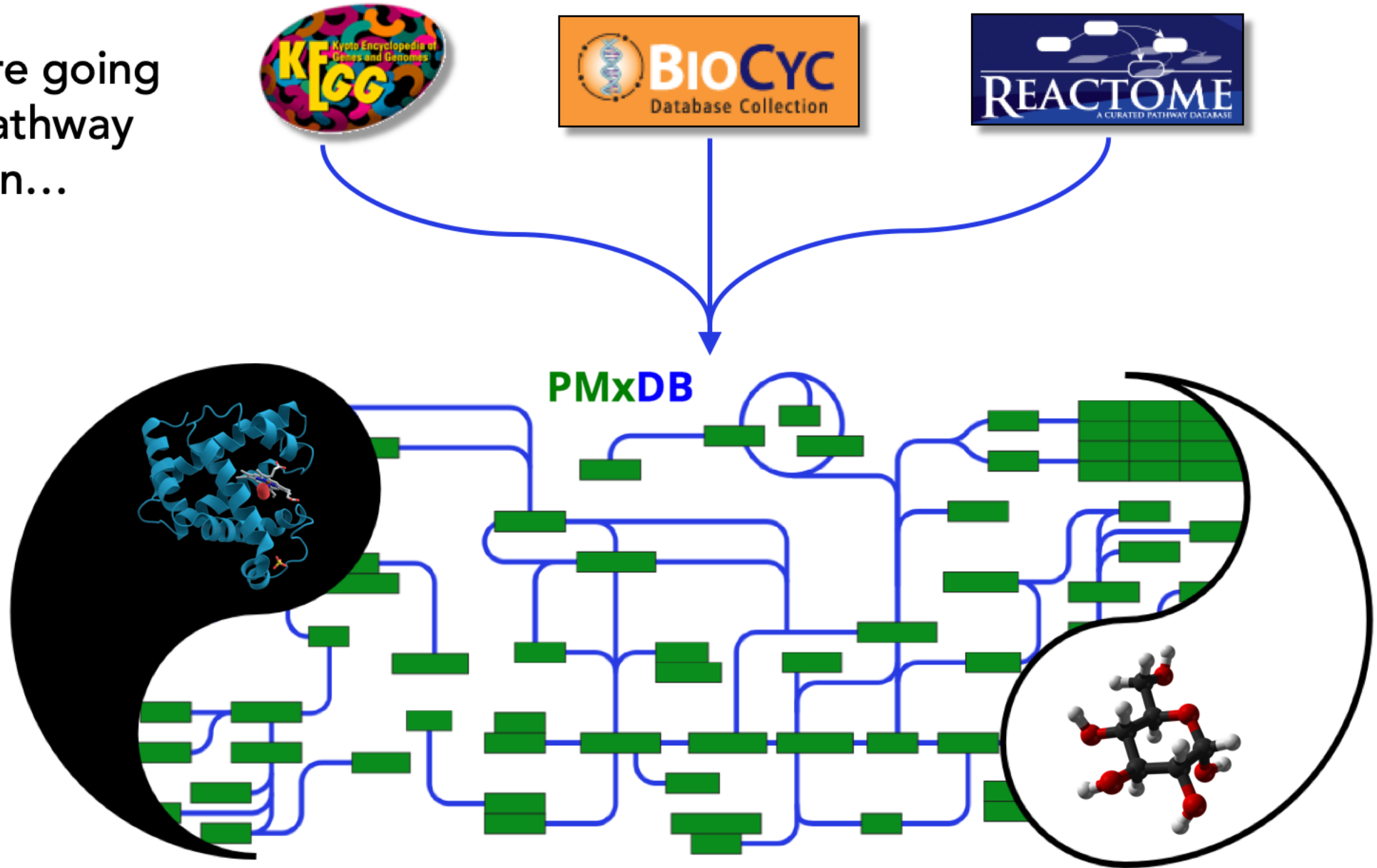
Drew R. Jones, PhD  
Assistant Professor, Department of Biochemistry and Molecular Pharmacology  
Assistant Professor, Department of Radiation Oncology



Bo Shopsin, MD, PhD  
Specialty: Infectious Diseases & Immunology  
Language: English

So, you want to combine **Proteomics** and **Metabolomics**...

Well, you're going to need pathway information...



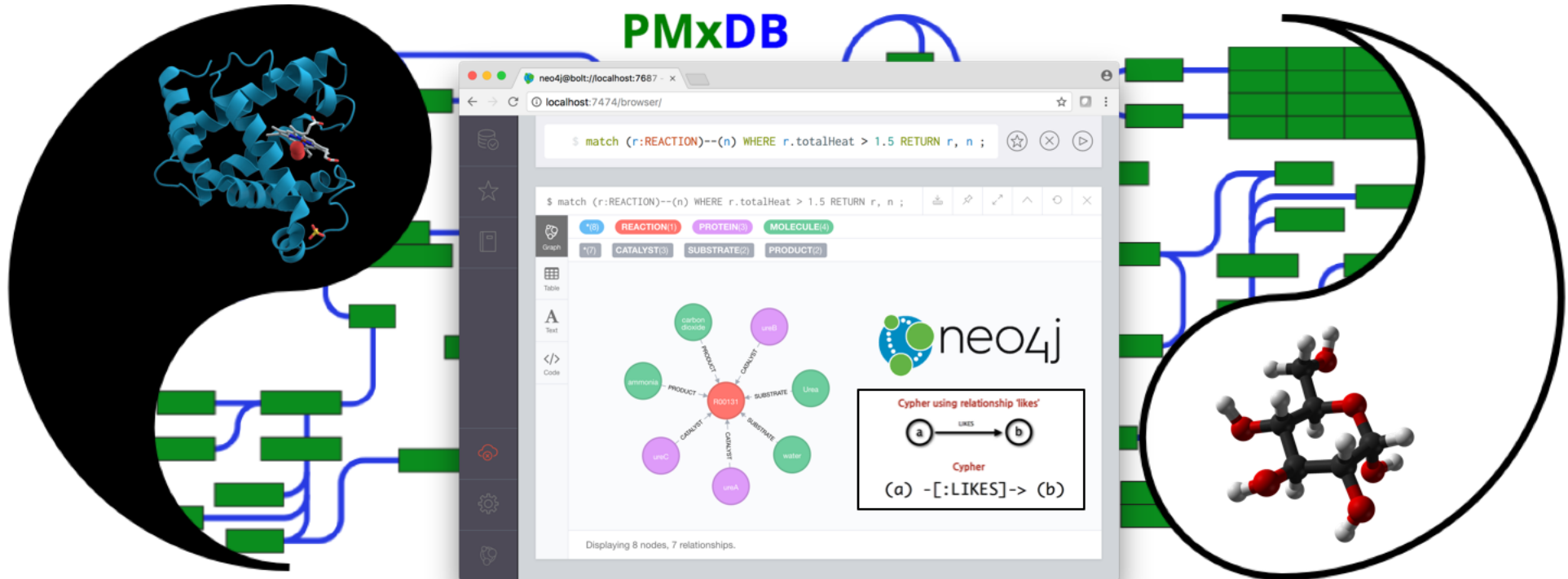


So, you want to combine **Proteomics** and **Metabolomics**...

Well, you're going to need pathway information...

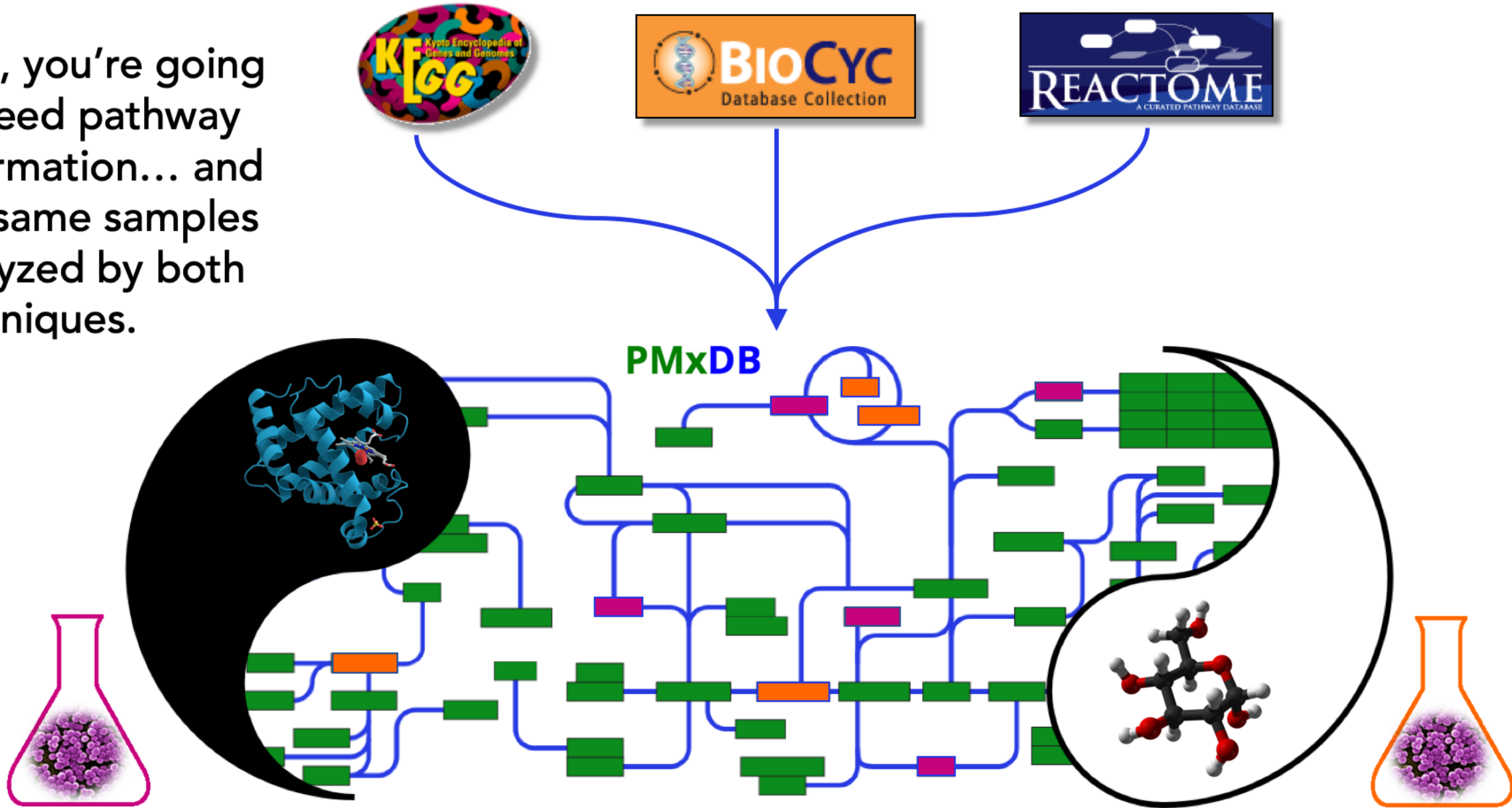


PMxDB

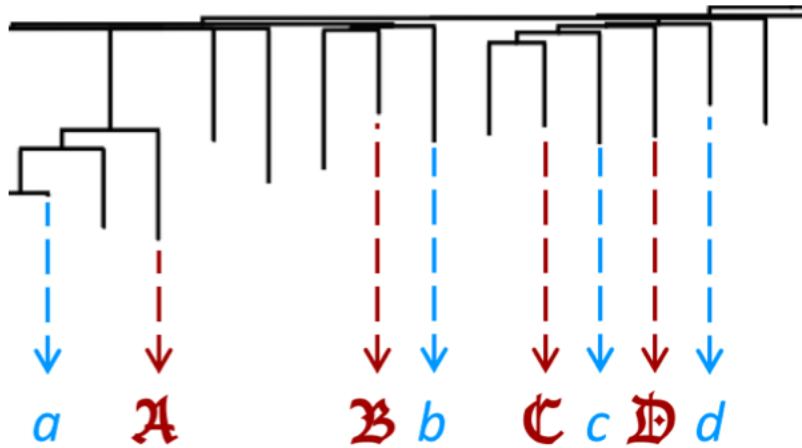


So, you want to combine **Proteomics** and **Metabolomics**...

Well, you're going to need pathway information... and the same samples analyzed by both techniques.



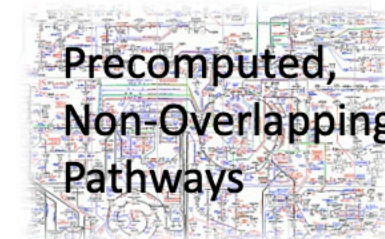
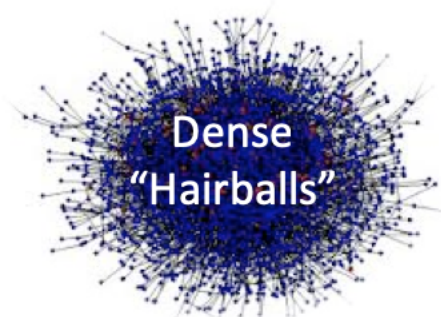
# "Mission Accomplished"?



Protein Name	Accession	MW	PI	Abundance	Log2 Fold Change
ADAMTS-1	Q9H011	140.0	5.5	100	0.50
ADAMTS-2	Q9H012	140.0	5.5	10	1.78
ADAMTS-3	Q9H013	140.0	5.5	10	1.78
ADAMTS-4	Q9H014	140.0	5.5	10	1.78
ADAMTS-5	Q9H015	140.0	5.5	10	1.78
ADAMTS-6	Q9H016	140.0	5.5	10	1.78
ADAMTS-7	Q9H017	140.0	5.5	10	1.78
ADAMTS-8	Q9H018	140.0	5.5	10	1.78
ADAMTS-9	Q9H019	140.0	5.5	10	1.78
ADAMTS-10	Q9H020	140.0	5.5	10	1.78



# Now, you want to visualize ProteoMetabolomics data!



Metabolite	Abundance	Retention Time	Mass	Pathway	...
Glucose	1234	12.34	180	Glycolysis	...
Fructose	567	15.67	180	Glycolysis	...
Galactose	910	18.91	180	Glycolysis	...
Glucose-6-phosphate	234	23.45	180	Glycolysis	...
Fructose-1,6-bisphosphate	345	34.56	284	Glycolysis	...
...	...	...	...	...	...

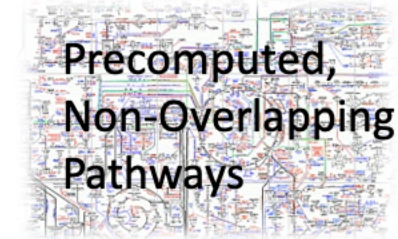
A screenshot of a spreadsheet application (Google Sheets) displaying a large table of metabolite data. The table has multiple columns and rows, with a prominent orange header row. The data appears to be organized by metabolite name, abundance, retention time, and mass.

# Now, you want to visualize ProteoMetabolomics data!

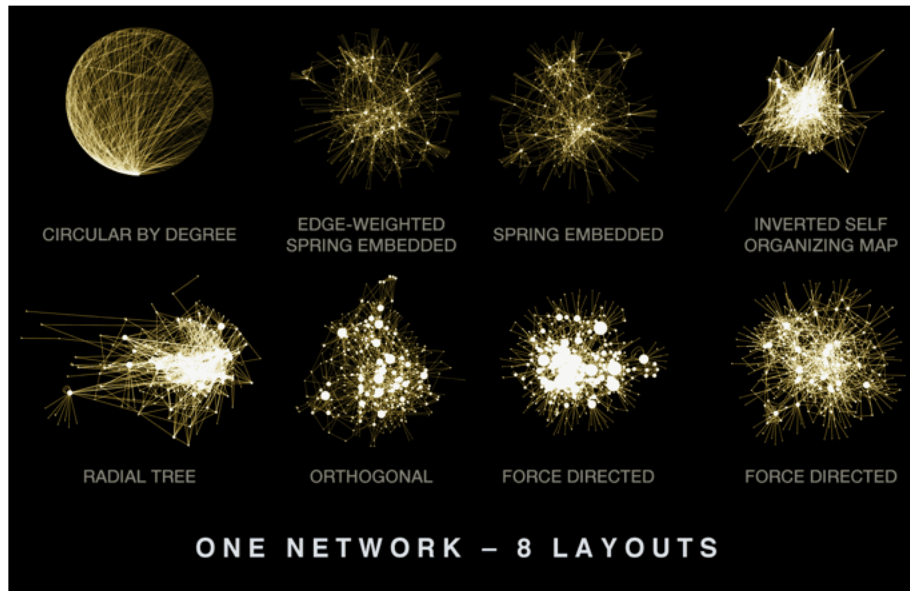
Endless Lists



Precomputed, Non-Overlapping Pathways

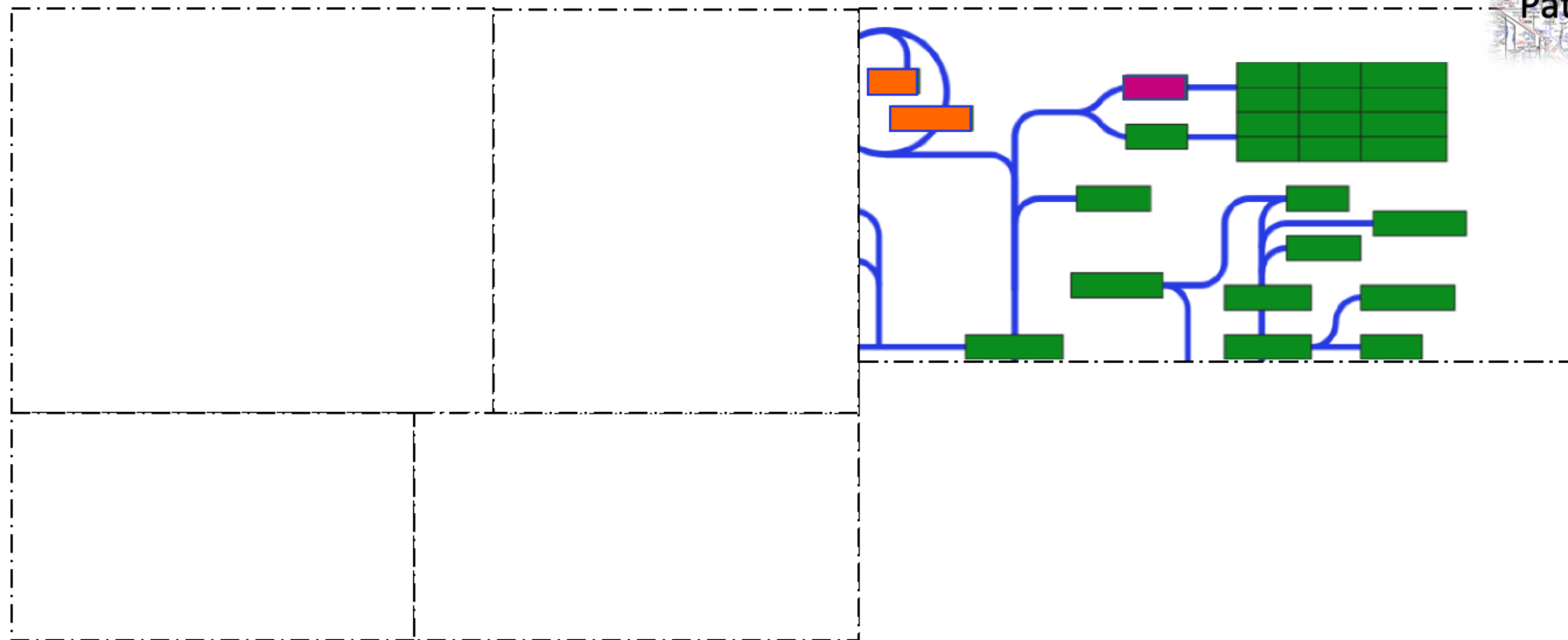
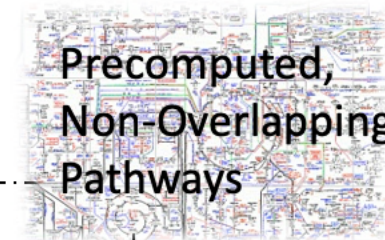


Dense "Hairballs"



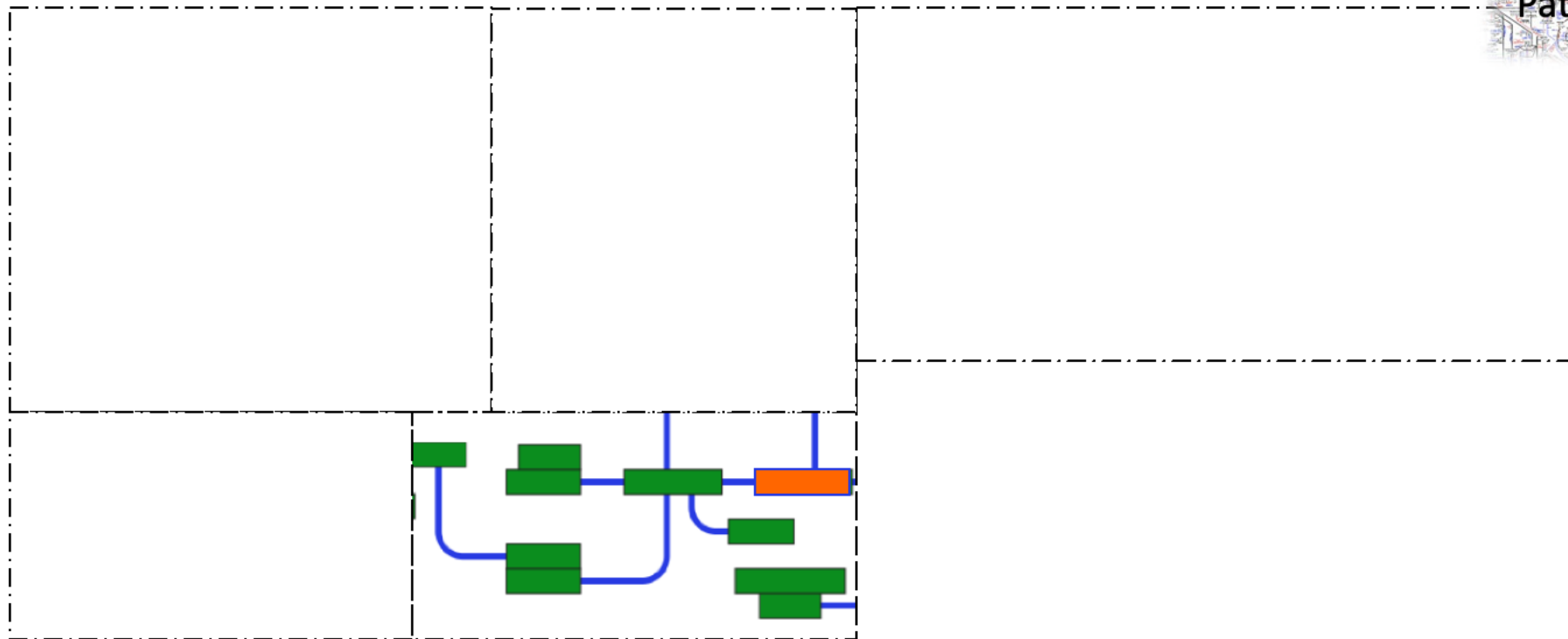
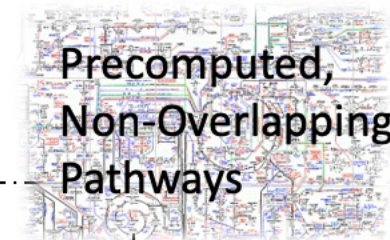
# Now, you want to visualize ProteoMetabolomics data!

Endless Lists



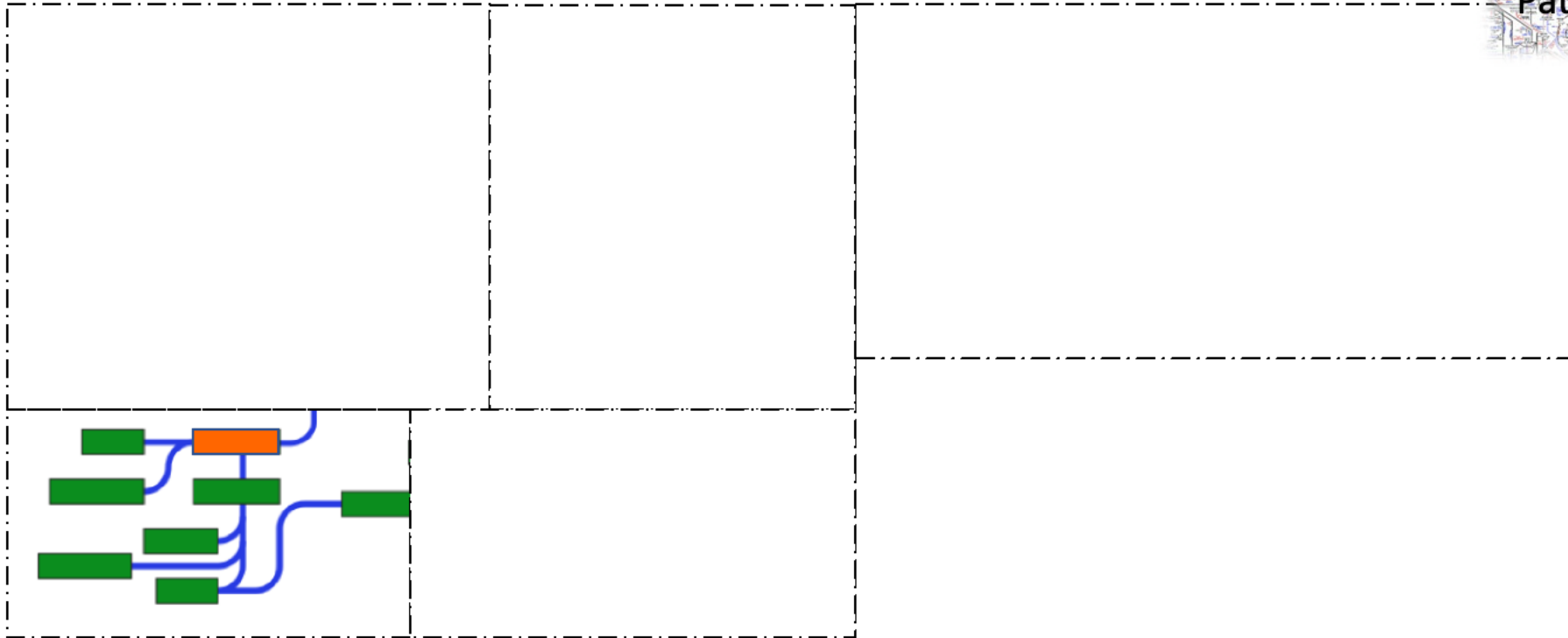
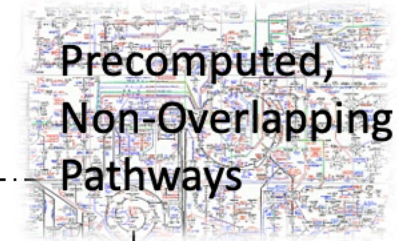
# Now, you want to visualize ProteoMetabolomics data!

Endless Lists



# Now, you want to visualize ProteoMetabolomics data!

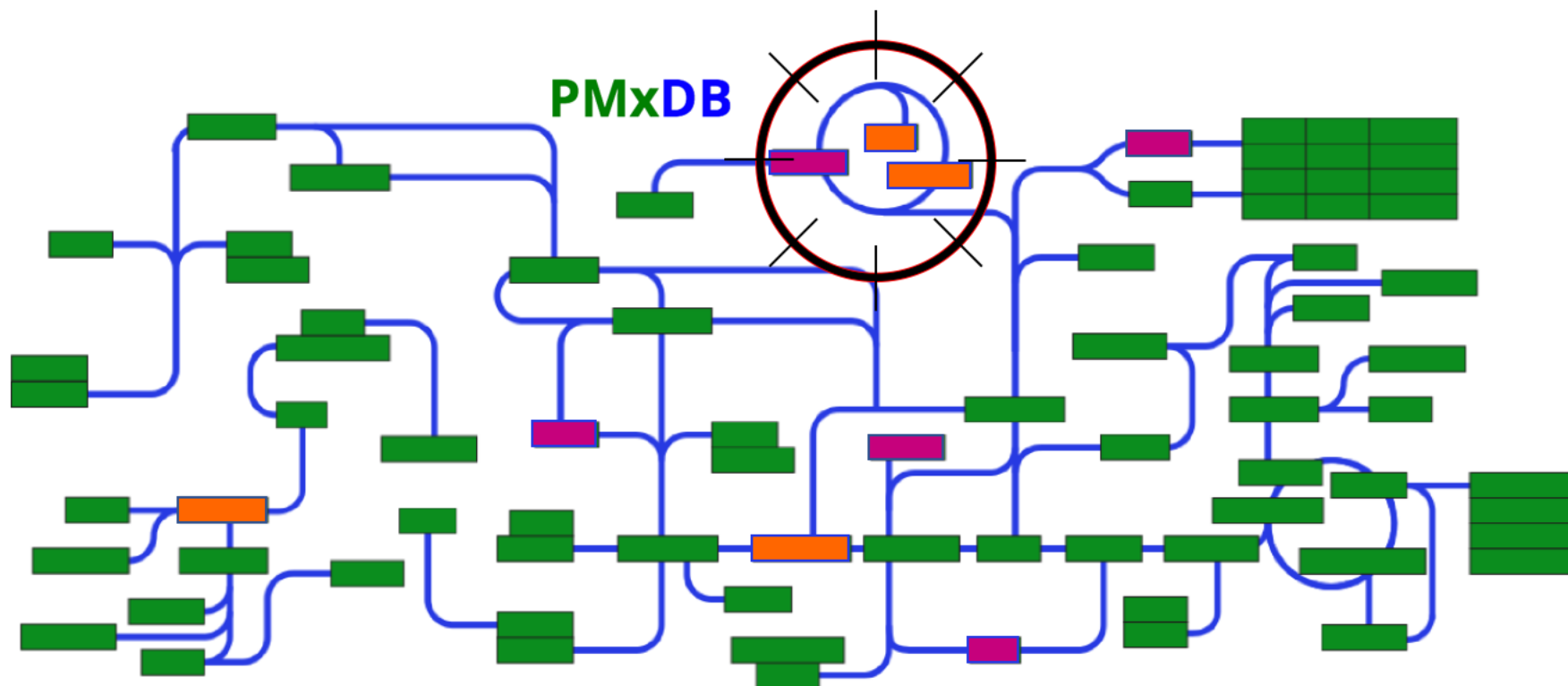
Endless Lists





Now, you want to visualize ProteoMetabolomics data!

You'll need a *HotSpot* analysis tool:



# HotSpot analysis: definitions

$heat \equiv$  “change” (others might prefer measures of “surprise”)

$$heat \equiv \left| \log_2 \frac{condition\ A}{condition\ B} \right|$$

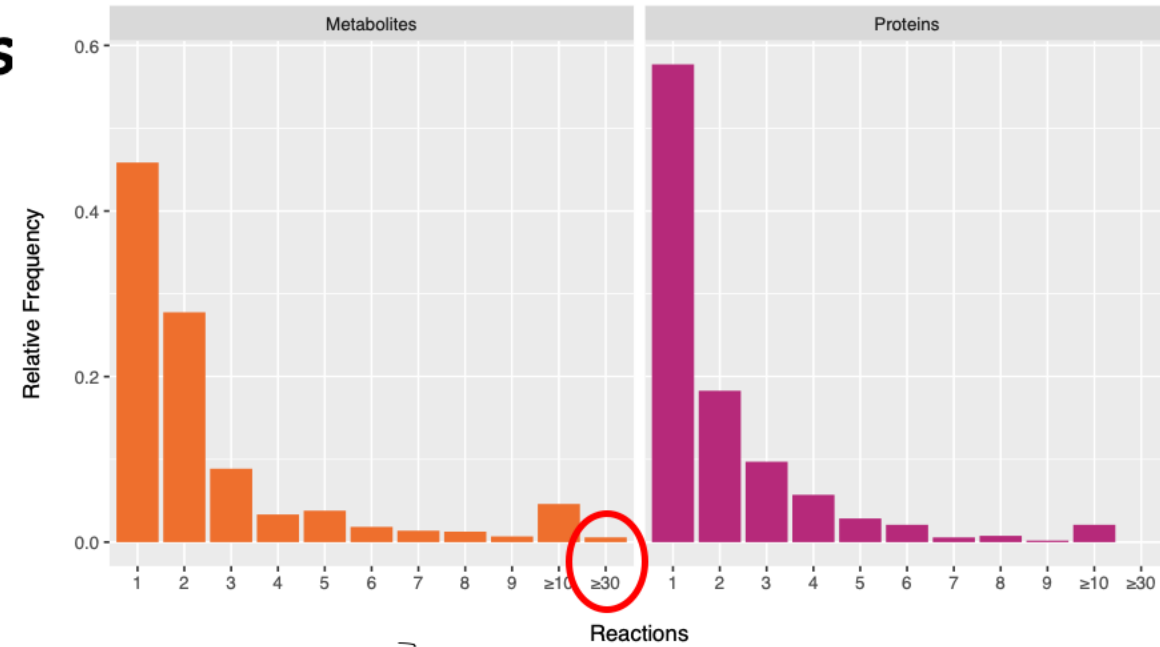
$$protein\_heat_{reaction} = \frac{\sum_{p \in proteins(reaction)} heat(p)}{|proteins(reaction)|}$$

$$metabolite\_heat_{reaction} = \frac{\sum_{m \in metabolites(reaction)} heat(m)}{|metabolites(reaction)|}$$

$$local\_heat_{reaction} = protein\_heat_{reaction} + \beta metabolite\_heat_{reaction}$$

$$heat_{reaction} = local\_heat_{reaction} + \kappa \frac{\sum_{r \in neighbours(reaction)} local\_heat(r)}{|neighbours(reaction)|}$$

$$hot\_spot_{network} = maximum\_connected\_component(network[heat(r) > \theta])$$



Ignore “currency” metabolites!

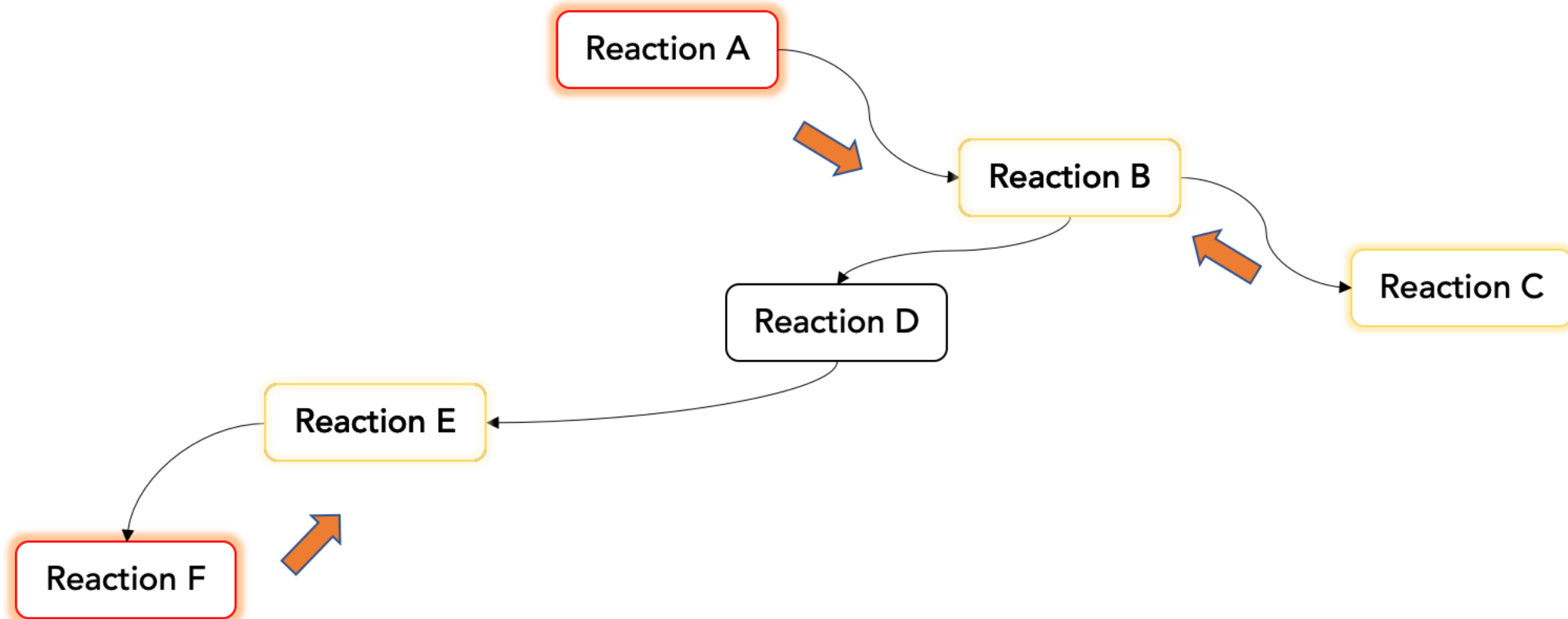
The network in question is defined as a reaction graph where the nodes correspond to reactions and edges are substrate-product relations.

# HotSpot analysis: definitions

$$heat_{reaction} = local\_heat_{reaction} + \kappa \frac{\sum_{r \in neighbours(reaction)} local\_heat(r)}{|neighbours(reaction)|}$$

$$hot\_spot_{network} = maximum\_connected\_component(network[heat(r) > \theta])$$

The network in question is defined as a reaction graph where the nodes correspond to reactions and edges are substrate-product relations.

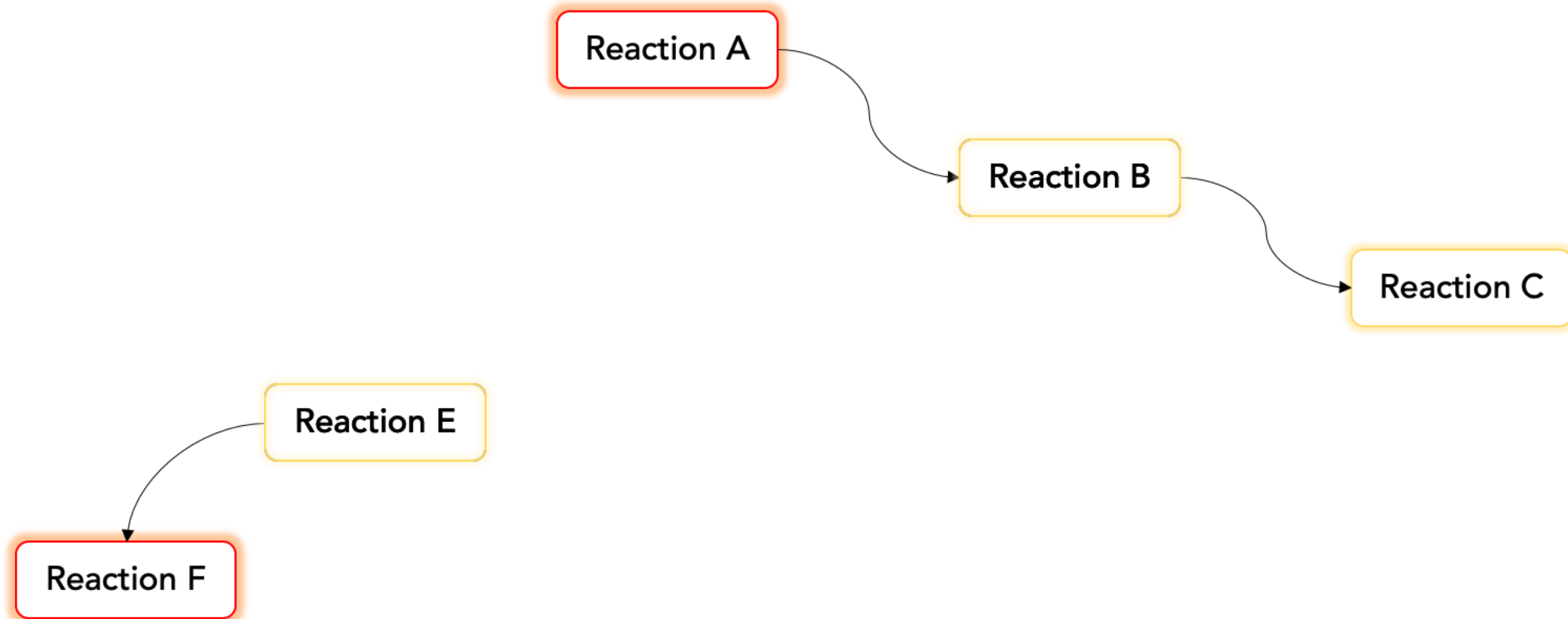


# HotSpot analysis: definitions

$$heat_{reaction} = local\_heat_{reaction} + \kappa \frac{\sum_{r \in neighbours(reaction)} local\_heat(r)}{|neighbours(reaction)|}$$

$$hot\_spot_{network} = maximum\_connected\_component(network[heat(r) > \theta])$$

The network in question is defined as a reaction graph where the nodes correspond to reactions and edges are substrate-product relations.



# HotSpot analysis: input requirements

## Proteomics

Proteomics heatmap showing log<sub>2</sub>(fold-change) values for various proteins across different conditions. The columns are color-coded by protein families: Blue (e.g., Histone, DNA damage), Green (e.g., DNA replication), Yellow (e.g., Cell cycle), Orange (e.g., Metabolic), Red (e.g., Immune response), Purple (e.g., Signaling), and Pink (e.g., Stress response).

## Metabolomics

Metabolomics heatmap showing log<sub>2</sub>(fold-change) values for various metabolites across different conditions. The columns are color-coded by metabolite classes: Blue (e.g., Amino acids), Green (e.g., Carbohydrates), Yellow (e.g., Lipids), Orange (e.g., Vitamins), Red (e.g., Minerals), Purple (e.g., Nucleotides), and Pink (e.g., Other).

The hotspot analysis tool requires only  
( identifier,  $\log_2(\text{fold-change})$  )

pairs, where the identifier is a

- UniProt:ID for proteins
- ChEBI:ID for metabolites

Protein Heat Weighting:



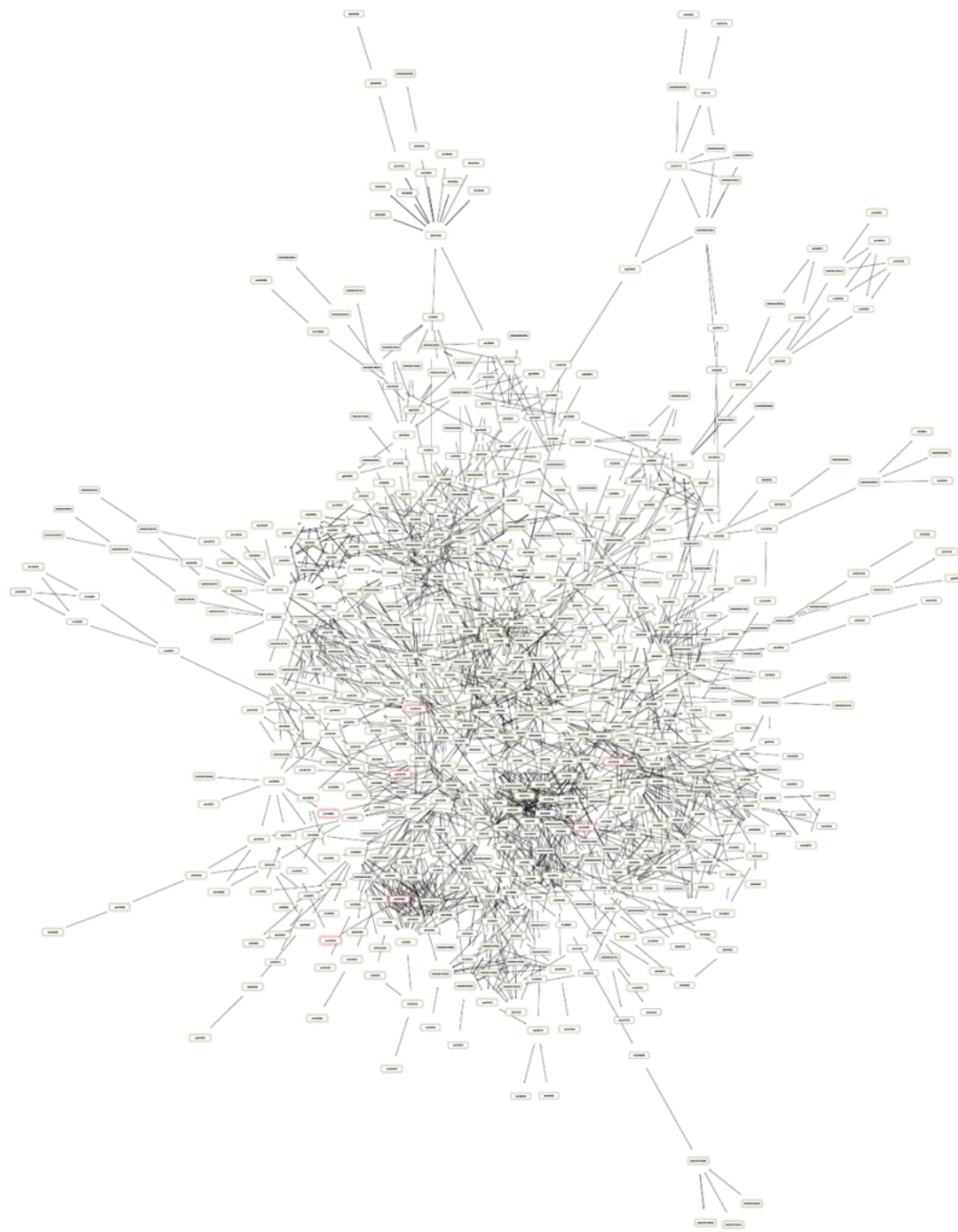
Metabolite Heat Weighting:

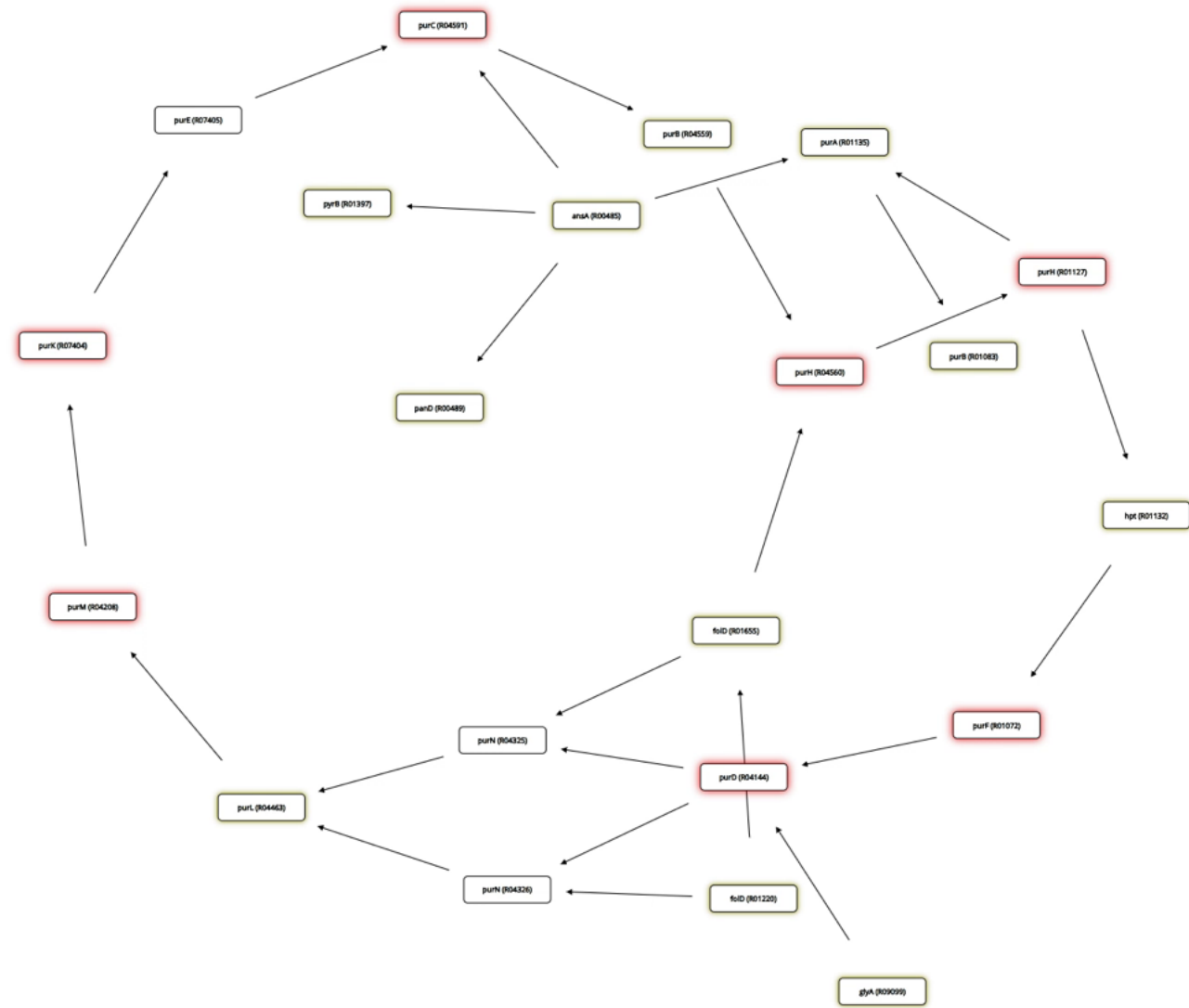


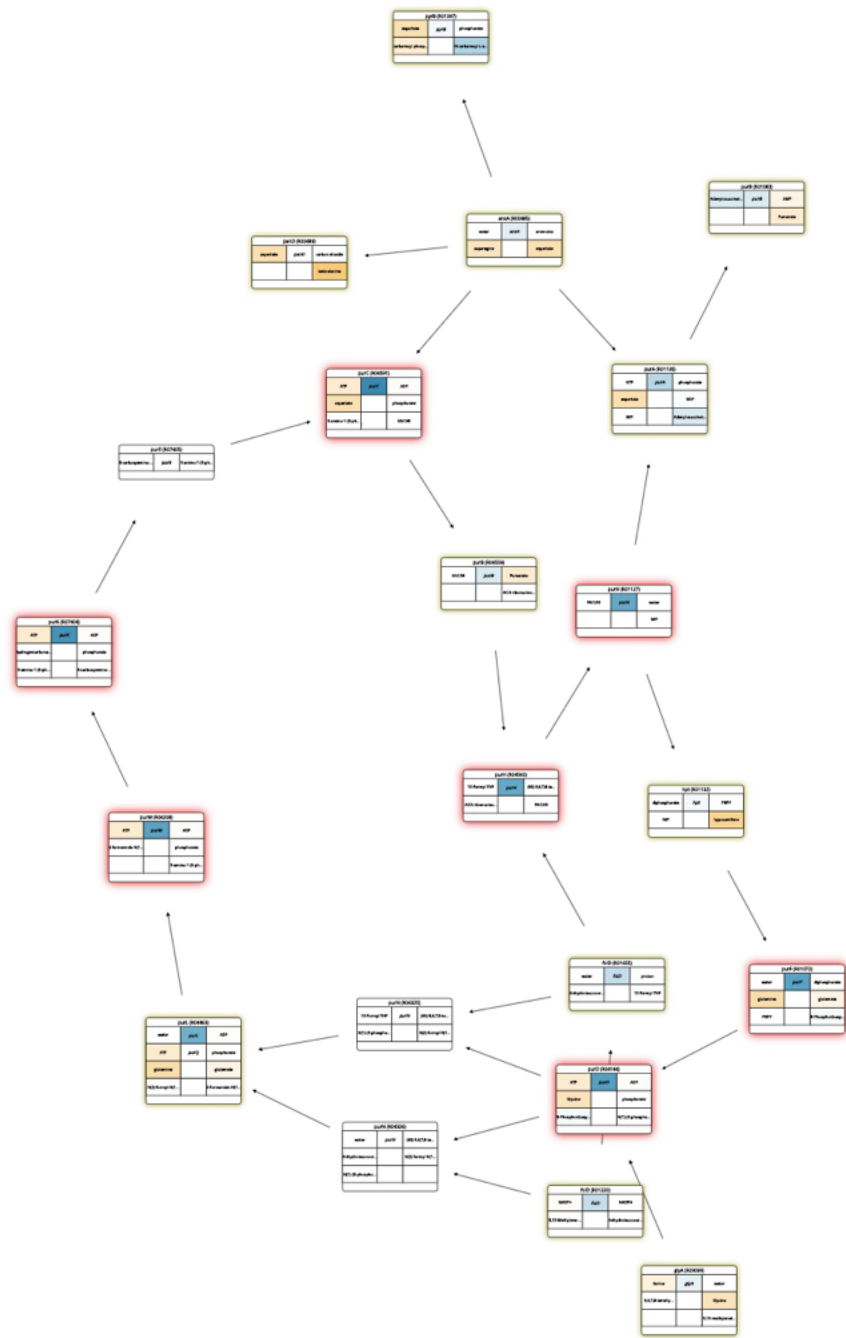
Heat Threshold:



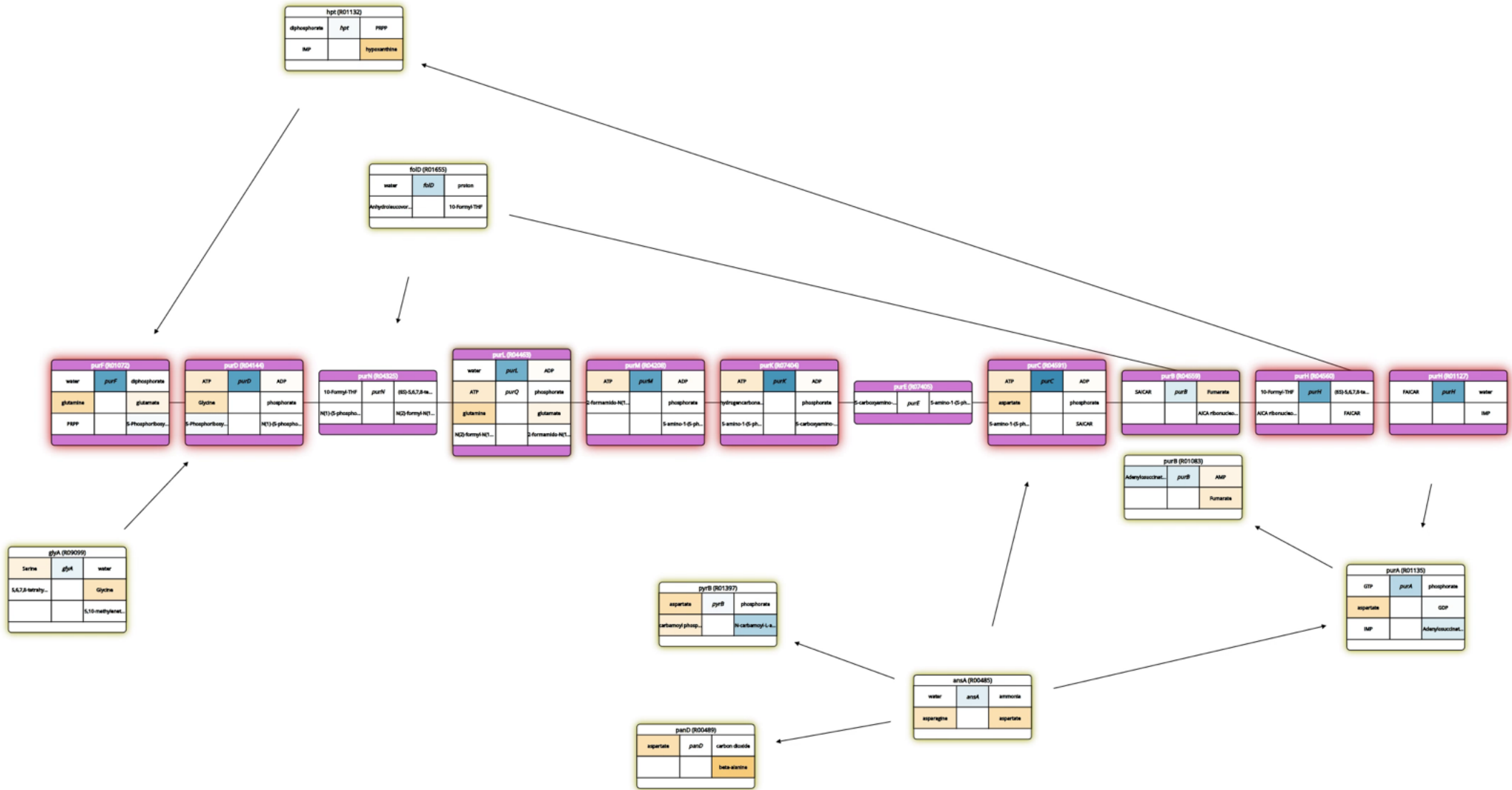
Launch **HotSpot**



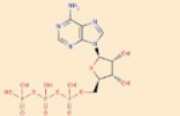
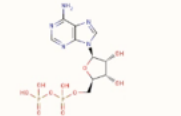
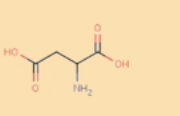
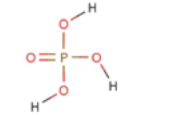
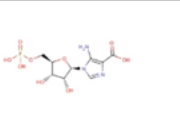
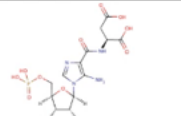


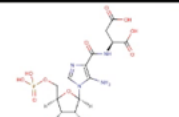
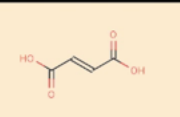
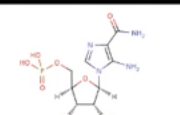


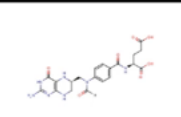
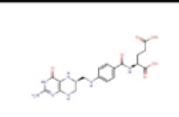
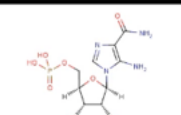
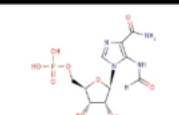


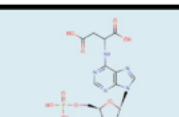
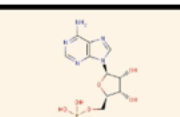
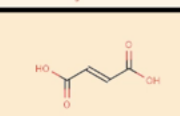


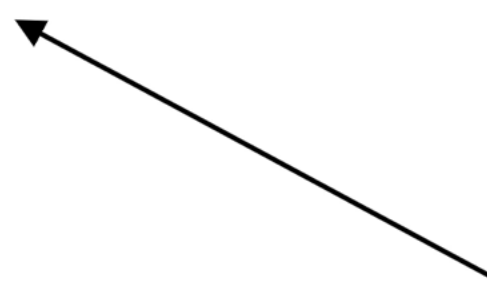


purC (R04591)		
	<b><i>purC</i></b>	
		
		

purB (R04559)		
	<b><i>purB</i></b>	
		

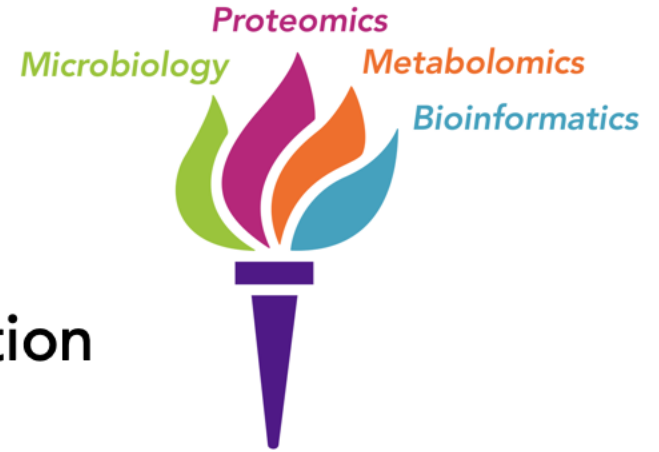
purH (R04560)		
	<b><i>purH</i></b>	
		

purB (R01083)		
	<b><i>purB</i></b>	
		

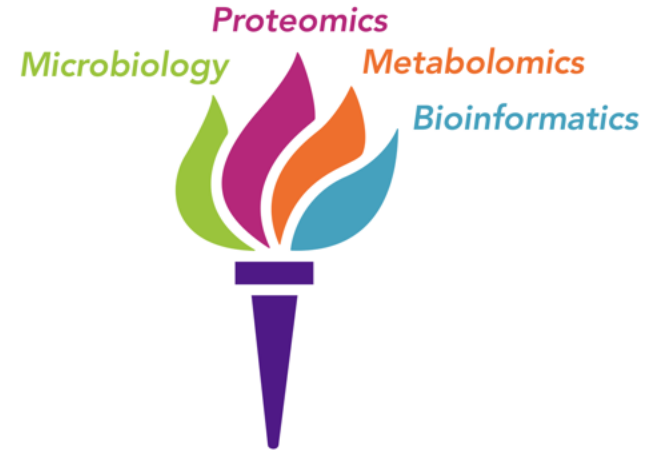


# Conclusion:

- Observation of *S. aureus* purine biosynthesis deregulation “in the wild” (in the context of a virulence study)
- Identified with ease by HotSpot detection:
  - selecting a relevant region of the biochemical network
  - without limiting the user to pre-defined “boxes”
- The approach should be very easy to apply:
  - to other organism
  - to single-“omic” datasets
  - to transcriptomics data



# Acknowledgements:



*Victor Torres, Irnov Irnov, Sophie Dyzenhaus,  
William Sause, Bo Shopsin MD,  
Beatrix M. Ueberheide, Avantika Dhabaria,  
Drew R. Jones, Rebecca Rose*



*Christoph Henrich, Bernard Delanghe, Kai Fritzscheier*