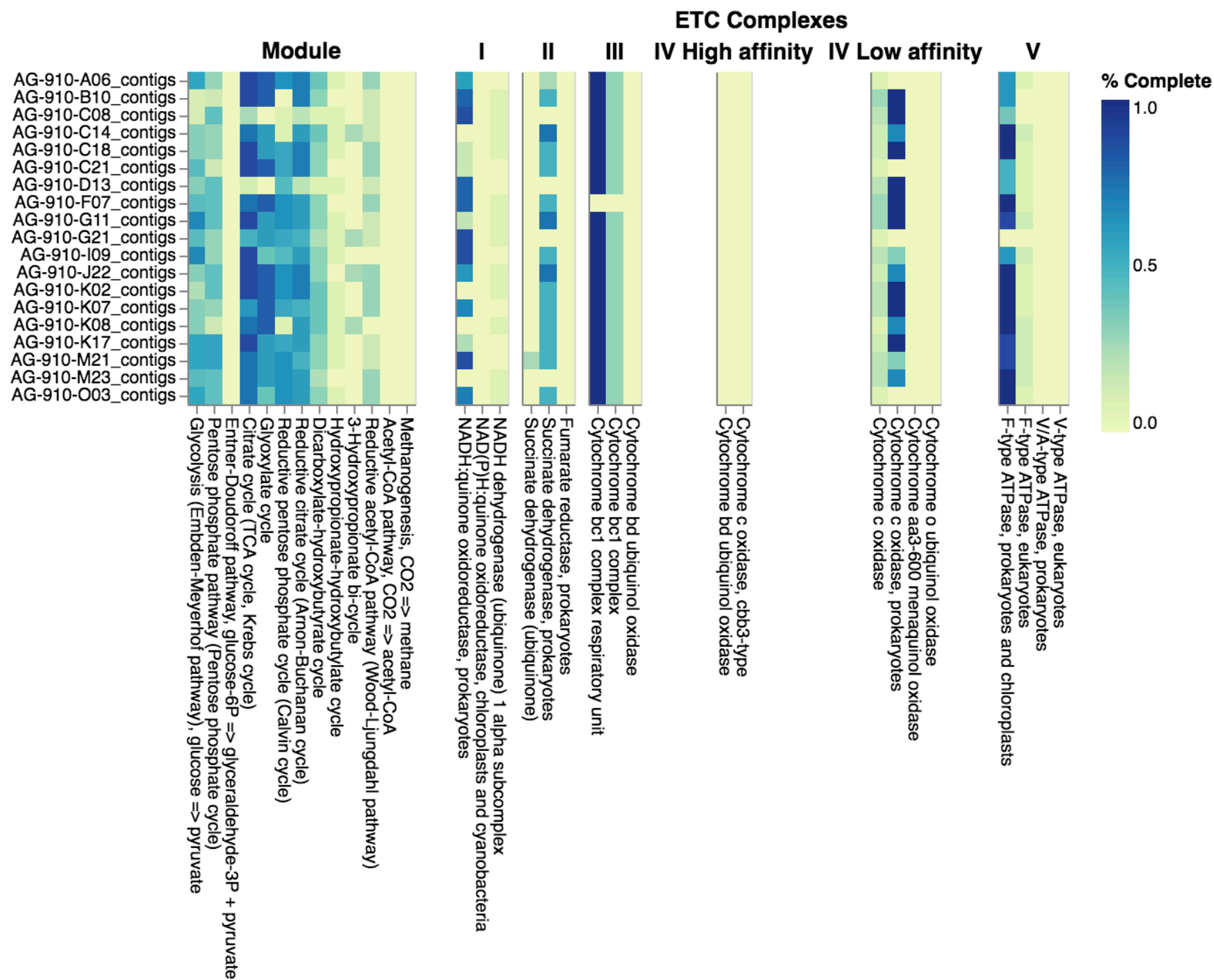


DRAM outputs  
(with further caveats!)

# DRAM outputs

- First, let's go through the files you generated from the “annotate” and “distill” DRAM commands
- What are they, how can they be used? What will be most useful to you?
- Annotation summaries



- One example of visualization (made with DRAM in the “Distill” command) – we won’t do a whole lot with manipulating this data, but will show you how to access what goes into a dataset like this!

## CAZy

## Nitrogen metabolism

## Sulfur metabolism

## Other Reductases

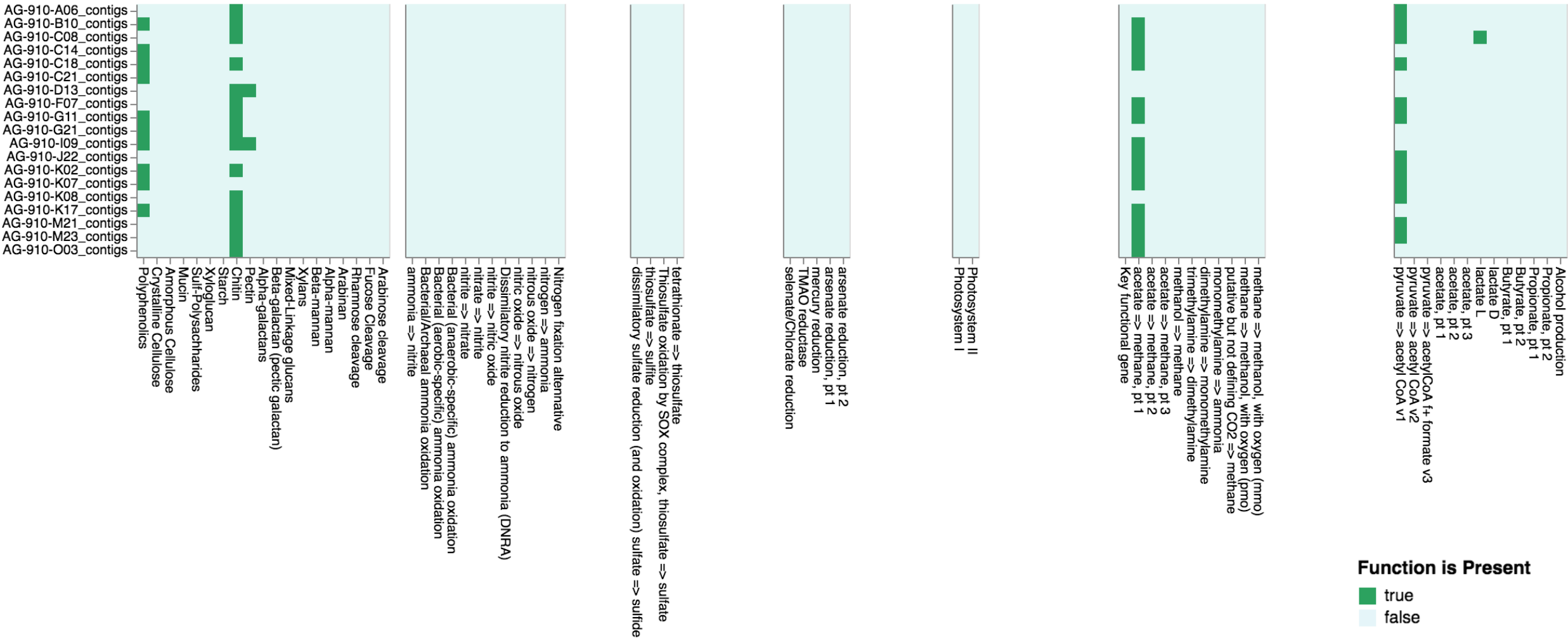
## Photosynthesis

## Methanogenesis and methanotrophy

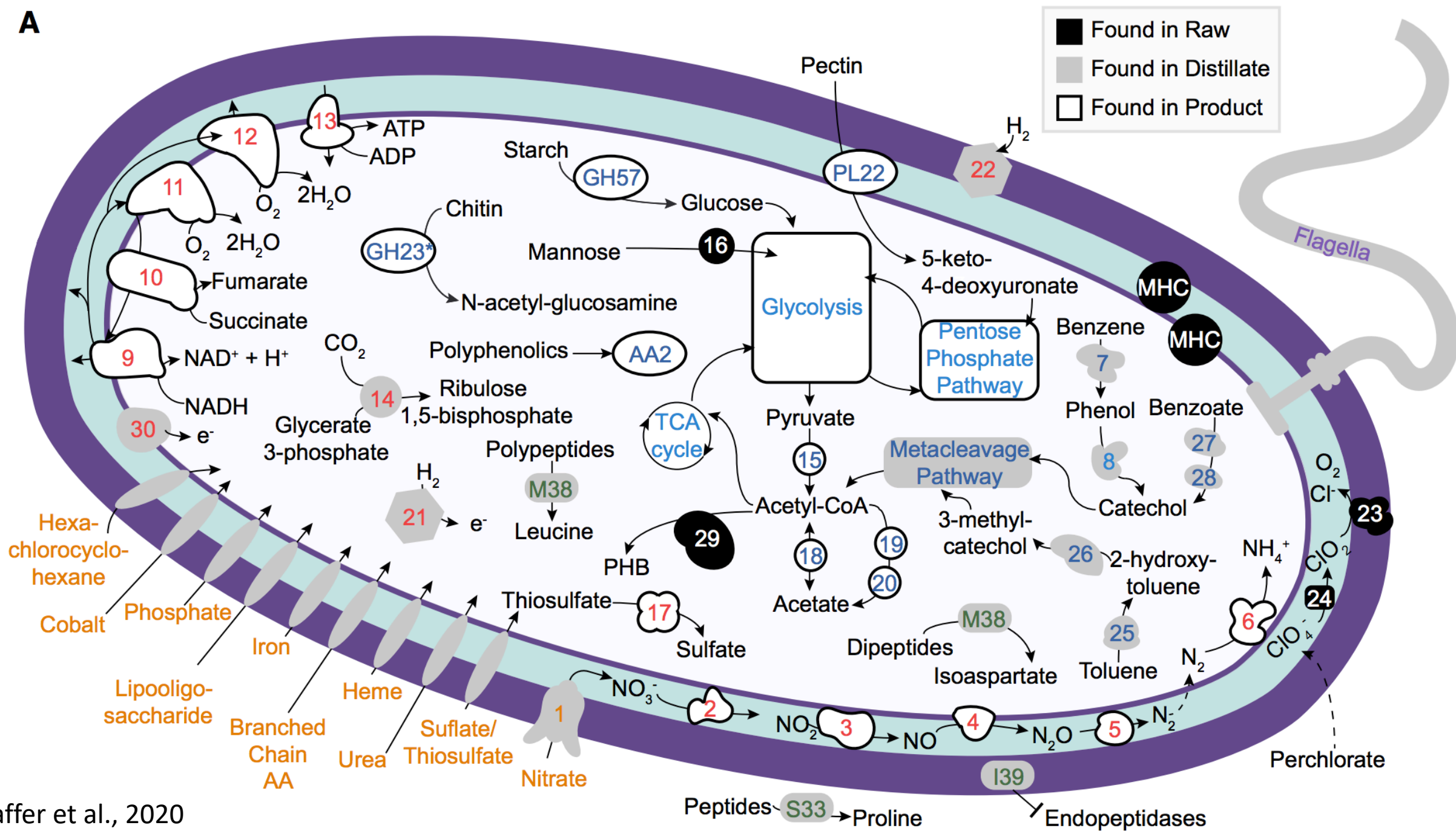
## SCFA and alcohol conversions

### Function is Present

true  
false

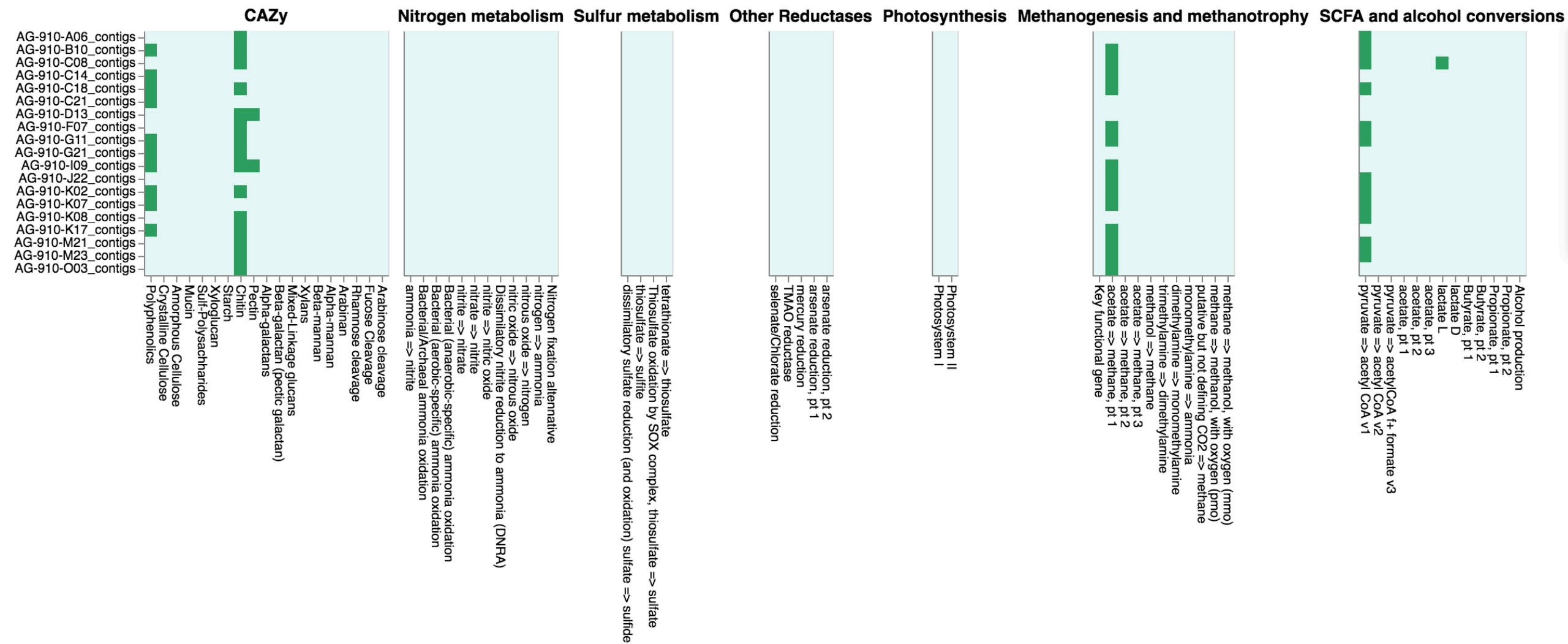


**A**



# Misannotations

- Rampant in databases, particularly as larger and larger sequence datasets are deposited using automated annotation programs
- How would misannotations become “the norm” for some genes?
- Do you know of any common misannotations in data from organisms you work with/might work with in the future?



Notice anything suspicious?

How would you go about verifying?

# Check databases (KEGG)



## ORTHOLOGY: K01895

[Help](#)

<b>Entry</b>	K01895	KO
<b>Symbol</b>	ACSS1_2, acs	
<b>Name</b>	acetyl-CoA synthetase [EC:6.2.1.1]	
<b>Pathway</b>	<a href="#">map00010</a> Glycolysis / Gluconeogenesis <a href="#">map00620</a> Pyruvate metabolism <a href="#">map00630</a> Glyoxylate and dicarboxylate metabolism <a href="#">map00640</a> Propanoate metabolism <a href="#">map00680</a> Methane metabolism <a href="#">map00720</a> Carbon fixation pathways in prokaryotes <a href="#">map01100</a> Metabolic pathways <a href="#">map01110</a> Biosynthesis of secondary metabolites <a href="#">map01120</a> Microbial metabolism in diverse environments <a href="#">map01200</a> Carbon metabolism	
<b>Module</b>	<a href="#">M00357</a> Methanogenesis, acetate => methane	

KEGG database

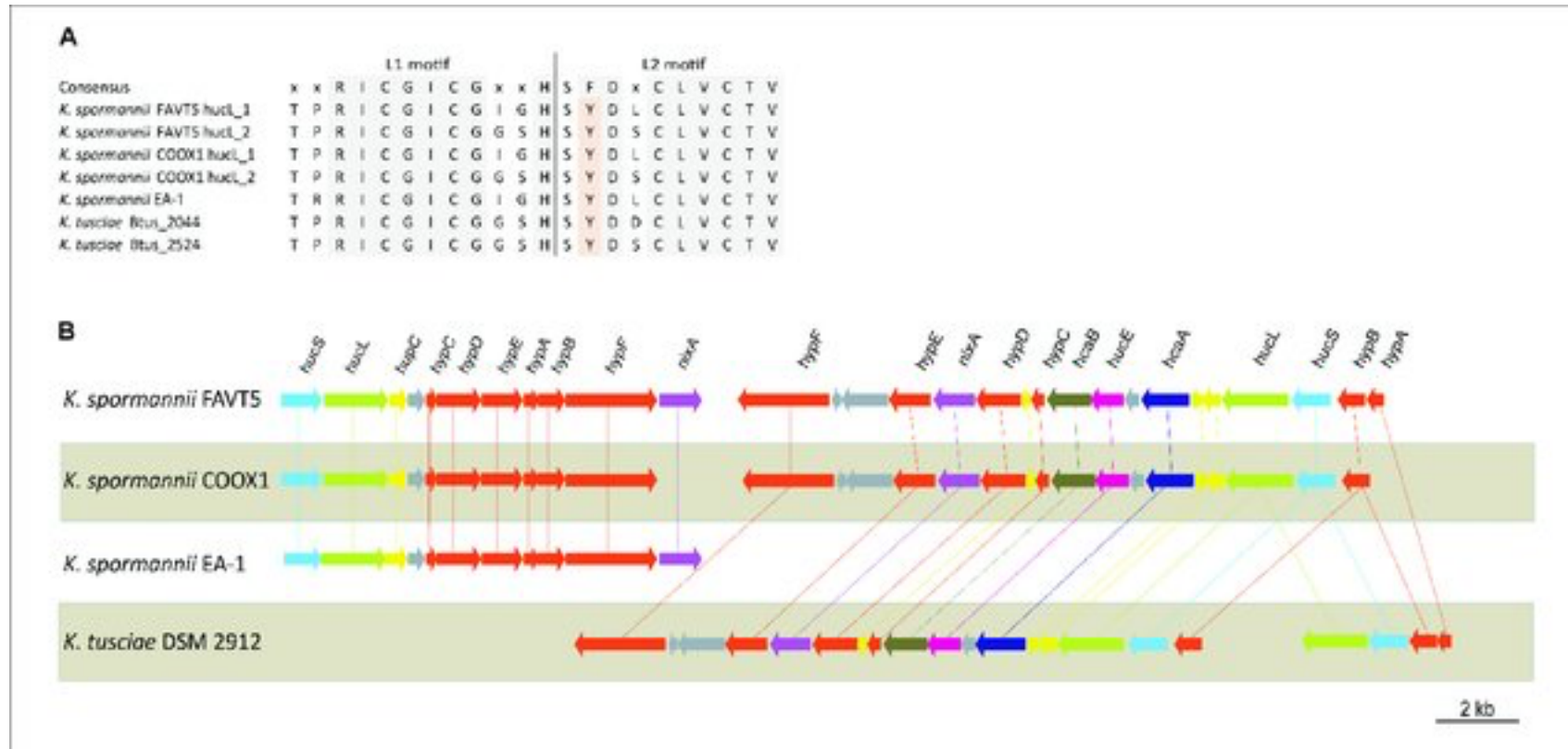


# Other Misannotations?

- Gene involved in nitrogen assimilation – flagged as dissimilatory nitrate to ammonium
- Carbon fixation – TCA cycle, ATP citrate lyase involved in carbon fixation (appear to be TCA cycle, unless you blast with another ATP citrate lyase). **Challenge: how would you do this?**
- Arsenic reduction misannotation – one of the genes that is a regulation factor (actually in multiple pathways, need to verify that other genes are present in the arsenic reduction pathway!)

# Other ways to check for potential misannotations? Or divergence?

- Do you have any commonly misannotated genes that you know of?
- Active sites for enzymes (do they possess key motifs)?
  - Example: cytochromes
  - Example: hydrogenases
- Databases of your choice
  - Examples: hydrogenases/HydDB, cytochrome P450 database



Hogendoorn  
et al., 2020

- Example: Cysteine residues key for binding site of hydrogenases (for [NiFe]-hydrogenases, need paired CxxC motifs)

# Other ways to check for potential misannotations? Or divergence?

