Differential pathways in the 3 big phylogroups

Similar to what happens with gene presence/absence, a PCA of the pathways present in the pangenome (measured by gapseq) separate very clearly the major E. coli phylogroups. Interestingly, we can also divide the 8 phylogroups in 3 big groups: B2 alone; A, B1 and C; and D, E, F and G on the other hand. Thus, we can study the differential existence of pathways in these major 3 groups to try to see interesting patterns arising from this.



**Pathways only present in the B2 group:**



Salmochelin is an enterobacting produced by Salmonella species and uropathogenic E. coli strains. Interesting that not only B2 can create and degrade it but also members of the other phylogroups. 2-deoxy-D-ribose is a metabolite that can be break down into acetyl-CoA and D-glyceraldehyde 3-phosphate to enter into glycolysis. Polymyxin A are antibiotics created by NRPS.

The **reductive Stickland reaction** is an interesting one. According to metacyc, Stickland reactions allow bacteria to metabolise pairs of amino acids by reducing one and oxidising the other one. Interestingly, this specific reductive branch of Stickland reactions are present in a very few organisms from Firmicutes, Actinobacteria or Thermotogae phylas, but not in Enterobacteria… Also, according to the data from gapseq, the oxidative branch is not complete in any strain here! Perhaps this is an interesting thread to pull. Are there any other examples of partial pairs of Stickland reactions? Why does the B2 group have the reductive pathway complete and not the others? It would seem that the other phylogroups lost genes in the way.

<https://biocyc.org/META/NEW-IMAGE?type=PATHWAY&object=PWY-8187>

<https://biocyc.org/META/NEW-IMAGE?type=PATHWAY&object=PWY-6344>

Arginine deiminase test must be a bug in their db, it has the same proportions as the last

case.

D-threonate is a 4C sugar, a stereoisomer of L-threonate, a more common form found in food and nature. Why is the D form metabolised by B2 members alone? Is this a common form to be found in hosts (as they are mostly pathogenic)? Could be related to the D-mannose as well?

About D-mannose, there seems to be little evidence that it might act as an alternate antibiotic for uropathogenic *E. coli* strains. However, there’s also evidence that it doesn’t do anything (<https://www.mdpi.com/1420-3049/25/2/316>) I’m wondering if this means that some authors tested an uropathogenic strain that didn’t have the D-mannose degradation and thus it was sensitive for the compound, whereas if you test for a pathogen that can degrade it, it doesn’t work.

About the last one: according to metacyc, bugs usually have 2 ways to create asparagine. However in certain bacteria there’s an alternate pathway dependant of tRNA, often the only way to create asparagine. Could this be happening in certain B2 members? It doesn’t seem the case, so why?

Below, a case for a genome (100, AUS strain) that has the 3 systems. Could this system be active under certain conditions?

Graphical user interface, application

Description automatically generated

**Pathways only present in the A, B1, C group:**



**Lipid A-core biosynthesis**: it’s an operon that produces 2 of the 3 components of the LPS in E. coli, namely the lipid A (attach the LPS to the membrane) and the core lipid (gives stability). The last part, the O-antigen, was lost in some strains due to a mutation. Therefore, our lab strains should have it complete and interestingly 12 of the 16 lab strains have the pathway complete. The 4 that doesn’t have it complete are: OP50 (no surprise at this point), BL21(DE3), and the two strains used for experimental evolution 606 and 607.

**Vanillin and vanillate degradation I**: Vanillin and vanillate are important intermediate metabolites of lignin-derived aromatic compounds.

Phenylacetate degradation I (aerobic): Phenylacetate is a major intermediate in bacterial degradation of many aromatic compounds. It produces some TCA intermediaries as acetyl-CoA and succinyl-CoA.

Acrylate degradation II: The genes encoding these enzymes are conserved in many organisms across all domains of life, suggesting that this detoxification system is likely relevant to metabolic processes and environments beyond DMSP catabolism