Understanding the cellular fitness effects of *de novo* and random proteins

Background: New protein coding genes frequently evolve from existing protein coding genes, via various mechanisms. Recent studies have shown that protein coding genes can also emerge "*de novo"* from genomic regions that did not previously encode any gene. While some proteins encoded by genes emerged *de novo*, and even rare random proteins may be beneficial to an organism, it is more likely that the majority of them have no effect on an organism's fitness or are detrimental. It is also possible that newborn *de novo* genes expressing toxic proteins would be purged due to negative selection, thus leaving behind only those genes that are either neutral or beneficial. However, random proteins have not been subjected to evolutionary selection and can be toxic. Although the *de novo* genes are eukaryotic, we expect the principles of generic protein toxicity (for example, formation of aggregates) to be applicable to all living organisms. The goal of this project is to compare the effect of expression of *de novo* and random proteins on the growth of *E.coli.*

Objectives:

- Insert specific barcodes in an *E.coli* expression plasmid
- Construct a library of *de novo* and random protein sequences in a plasmid based arabinose inducible system. *De novo* and random libraries carry different barcodes
- Combine the cells expressing the two barcoded libraries and perform many rounds of overnight growth (growth competition).
- Plate the cells on agar, and genotype the colonies using barcode specific colony PCR.
- Count *de novo* vs random colonies and repeat the procedure under different environmental conditions

Methods: Molecular cloning (restriction digestion, ligation etc), PCR, bacterial cell culture

Requirements:

- Interest and basic training in molecular biology lab work
- Interest and basic knowledge on molecular genetics

Supervision: Dr. Bharat Ravi, b.ravi@uni-muenster.de, Molecular Evolution and Bioinformatics Group (https://bornberglab.org/)

Relevant literature:

- van Oss and Carvunis (2019). https://doi.org/10.1371/journal.pgen.1008160
- Tretyachenko *et al.* (2017). https://doi.org/10.1038/s41598-017-15635-8
- Heames *et al*. (2023). https://doi.org/10.1038/s41559-023-02010-2
- Frumkin and Laub (2013). https://doi.org/10.1038/s41559-023-02224-4
- Aubel *et al.* (2024). https://doi.org/10.1093/gbe/evae069