

Deciphering the origins of phototrophy in Pseudomonadota (Proteobacteria) using genomic data

Context:

Phototrophy is the capacity to produce **cellular energy from light**. It is an **ancient** metabolic capacity dispersed throughout the tree of life (algae, plants, Cyanobacteria, several bacterial lineages) [1]. The phylum of Pseudomonadota (formerly Proteobacteria) gathers several scattered lineages of phototrophs that are physiologically separated in two classes: the P(N)SB that **do not produce energy from light when O₂ concentrations are high** (anaerobic phototrophs) and the AAPB that **can produce energy from light in the presence of O₂** (aerobic phototrophs). Intriguingly, AAPB and P(N)SB have intricate phylogenetic relationships, but also mix in the species tree with non-phototrophic lineages. Carl Woese hypothesized that the **ancestor of Pseudomonadota** could have been an anaerobic phototroph [2], and more recently Brinkmann and colleagues showed that the cluster of genes responsible for phototrophic metabolism ("PGC" for photosynthetic gene cluster) can be exchanged between bacteria via **lateral gene transfer** and plasmids [3]. Despite a large corpus of literature on their physiology and environmental distribution, the distribution of P(N)SB and AAPB across Pseudomonadota has not been systematically investigated in the diversity of genomes currently available [1] and their **relative origins and relationships** remain enigmatic. All these questions have to be re-evaluated in the light of the massive number of genome data now available and recent discoveries made in the host team in the field of evolutionary bioenergetics.

Keywords: phylogenomics; microbial evolution; bioinformatics; metabolism; omics data analysis



Varied pigment colors of chlorophototrophic bacteria: selected P(N)SB, taken from [1].

Master 2 project:

In order to clarify the origins and evolution of phototrophy in Pseudomonadota, we propose to (i) mine publicly available genomic data (first complete **genomes**, and then **metagenome assembled genomes**) for the presence of the PGC (annotation tool recently developed in the team), (ii) build a phylogenomic tree of the PGC, and (iii) compare it to the species tree of Pseudomonadota to provide a **scenario to explain the distribution of phototrophy throughout Pseudomonadota in terms of vertical or lateral inheritance**. Secondly, markers for different **energy metabolisms** studied in the lab will be incorporated in the

analysis, to provide a comprehensive evolutionary scenario for phototrophy across Pseudomonadota and make sense of it in the context of **varying O₂ concentrations** over Earth history [4].

During this project, the intern will create **scripts and workflows** (Python, Snakemake, shell, awk, contribution to the Github lab repo...), use specialized programs to manipulate, **annotate and analyse large and diverse (meta-)genomics data**, build **phylogenetic trees** and **perform statistical analyses in an evolutionary framework** (using programs such as blast, HMMER, IQ-

Tree, MacSyFinder [5]...). This approach will enable to gain new insights on the evolution of phototrophy, a crucial energy metabolism that appeared on Earth > 3 billion years ago [4].

Profile of the candidate: Ideally, the candidate will have followed a training in bioinformatics and programming, and have an interest in addressing biological questions in link with evolution. Depending on the candidates' preferences, an experimental part can be developed in the lab in link to the bioinformatic one. There could be possibilities to continue the project with a PhD thesis.

The TREE team @TIMC lab (CNRS, Université Grenoble Alpes): We are part of a highly interdisciplinary team, gathering biochemists, biophysicists, molecular microbiologists, biostatisticians and bioinformaticians, with a common strong interest in microbial evolution. The lab is located on the Campus of La Tronche, in close vicinity to Grenoble (Tram B).

- The team website: <http://www.timc.fr/en/tree>
- The computational biology group website: <https://tree-timc.github.io/compbio>

References:

- [1] Thiel V, Tank M, Bryant DA (2018). [Diversity of Chlorophototrophic Bacteria Revealed in the Omics Era](#). *Annual Review of Plant Biology*
- [2] Woese CR (1987). [Bacterial Evolution](#). *Microbiological Reviews*
- [3] Brinkmann H, Göker M, Koblížek M, Wagner-Döbler I, Petersen J (2018). [Horizontal operon transfer, plasmids, and the evolution of photosynthesis in Rhodobacteraceae](#). *ISME Journal*
- [4] Fischer WW, Hemp J, and Johnson JE (2016). [Evolution of Oxygenic Photosynthesis](#). *Annual Review of Earth and Planetary Sciences*
- [5] Néron B, Denise R, Coluzzi C, Touchon M, Rocha EPC, Abby SS (2023). [MacSyFinder v2: Improved modelling and search engine to identify molecular systems in genomes](#). *Peer Community Journal*

To apply: Please send an email to the contacts listed, attach a CV, a transcript of grades, and a statement to explain your interest in the internship. *Informal enquiries are also welcome, as well as a simple email to obtain the above publications should they be pay-walled.*

Contacts:

sophie.abby@univ-grenoble-alpes.fr
fabien.pierrel@univ-grenoble-alpes.fr