Title : Multi-enzyme phylogenies of anaplerotic CO_2 fixation and glyoxylate-cycle enzymes illuminate mixed evolutionary origins and complex carbon assimilation capacity in *Euglena gracilis*

Author(s) :

1. Mr. Nattharat Suchatpong — **Student ID :** 650510704
2. Ms. Patcharaporn Wongkham — **Student ID :** 650510712
3. Ms. Pimmada Thongprasri — **Student ID :** 650510713
4. Mr. Klatawan Srisawang — **Student ID :** 650510721
5. Mr. Chinnapat Buntan — **Student ID :** 650510725

Major : Data Science

Advisor :

1. Dr. Sahutchai Inwongwan

Type of presentation :

- ☐ **Oral Presentation** (เฉพาะ ตัวแทนนศ.ที่สาขาเลือกให้นำเสนอแบบบรรยาย)
- ☑ **Poster** (กรณี นำเสนอผลงานปัญหาพิเศษ/การค้นคว้าอิสระ)
- ☐ **Cooperative Education** (กรณี นำเสนอผลงานสหกิจศึกษา)

ABSTRACT

Euglena gracilis exhibits exceptional carbon metabolic flexibility, including inorganic carbon assimilation under heterotrophic condition. To investigate how these capabilities may have been evolutionarily assembled, we performed coordinated phylogenetic analyses of enzymes central to anaplerotic CO2 fixation and C2 metabolism: phosphoenolpyruvate carboxylase (PEPC), phosphoenolpyruvate carboxykinase (PEPCK), pyruvate carboxylase (PC), malic enzyme (ME), and glyoxylate cycle components, including the characteristic *Euglena* malate synthase isocitrate lyase fusion (MS ICL; EgGCE). Candidate sequences were retrieved from de novo transcriptome annotation and retained using similarity screening (at least 50 percent query coverage and at least 50 percent identity) with domain architecture verification, then analysed with curated reference homologs spanning major eukaryotic lineages. Proteins were aligned with MAFFT, trimmed to retain phylogenetically informative regions, and used to infer maximum likelihood trees within the NGPhylogeny.fr framework. Trees were annotated and visualized in iTOL with consistent group colouring. Across enzyme families, *E. gracilis* candidates showed distinct evolutionary topologies. PC displayed a compact pattern, comprising a closely related pair (Clusters 3571.40844 and 3571.40347) plus an additional candidate (Cluster 3571.39909). In contrast, PEPCK and ME were

**Type of presentation must be matched with an option you choosing on student upload system.*

***The abstract can be more than one page and must be approved by project advisor before upload.*

heterogeneous. In the PEPCK phylogeny, one candidate (Cluster 3571.40400) grouped with an amoebozoan reference, whereas three candidates (Clusters 3571.42343, 3571.37627, 3571.39711) formed a *Euglena* clade positioned with kinetoplastid sequences, indicating at least two divergent PEPCK lineages. PEPC candidates also separated into two placements, one among algal and protist references and another within a stramenopile enriched grouping. ME exhibited the greatest diversity, with candidates distributed across multiple regions of the tree. For glyoxylate cycle functions, the ICL candidate (Cluster 3571.39868) showed a distinct placement, while MS candidates split into plant or algal associated versus protist associated affinities. Together, these results provide a reproducible phylogenetic resource and generate testable hypotheses linking enzyme diversification to heterotrophic CO2 fixation and C2 metabolic capacity.

**Type of presentation must be matched with an option you choosing on student upload system.*

***The abstract can be more than one page and must be approved by project advisor before upload.*