

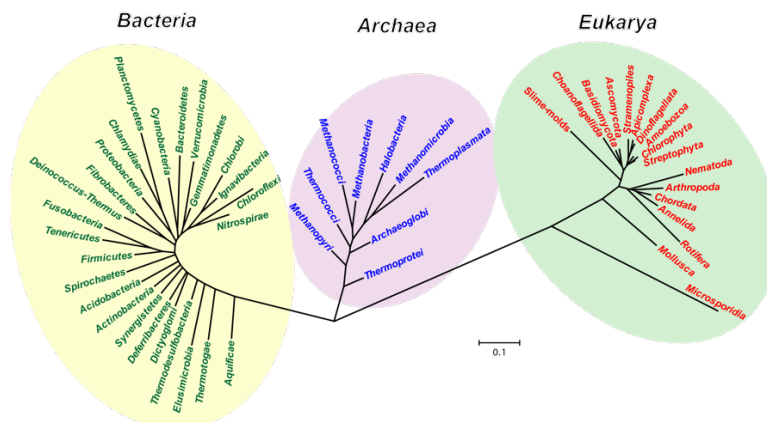
Master's thesis project: Taxonomy Information Toolbox for Big Data Mining

The problem

Biological research is shifting from experiment-focused studies toward data-driven interdisciplinary projects. **Big Data analysis** is emerging as the challenge, as well as opportunity, to the future biological and medical research, which is increasingly relying on the contributions from other fields, such as computer science and machine learning.

The project

This project aims to initialize a **taxonomy information toolbox** that serves as a centralized hub for supporting various bioinformatics utilizations. Because the taxonomic relationships reflect the intrinsic connections between all living organisms. This toolbox will help to bridge over the balkanized knowledge archipelago from a handful of model organisms to any sequenced species (>1.7 Millions).



The SysBio

SysBio is the Division of Systems and Synthetic Biology, part of the Department of Biology and Biological Engineering. Our lab is one of the leading Systems Biology research groups in the world, with an active and multi-disciplinary academic environment. We have developed and curated a wide range of databases and software packages, and thus an ideal setup for you to learn and improve **bioinformatic and programming skills** that are essential in the upcoming machine intelligence era. You will learn various biological databases and software, good practices in bioinformatics and computational biology, as well as how to collaborate with others on GitHub.

Application

Applicants need to have a background in either bioinformatics, computer science or biotechnology, and are self-motivated, eager to learn, problem-solving oriented. Previous experience with **Python and Git** is preferred. Duration: 6-12 months (30 or 60 hp / ECTS). This project will be supervised by Dr. Hao Wang. If you are interested, please send a letter of motivation to hao.wang@chalmers.se.