|  |  |
| --- | --- |
| **Project title:** | **Dysregulated processes in NAFLD** |
|  |  |
| **Project supervisor:** | **Hong Yang** |
| **Co-supervisor:** | **Koeun Shong** |
| **Phone:** | **+46-0720227567** |
| **E-mail:** | [**hong.yang@scilifelab**](mailto:hong.yang@scilifelab)**.se** |

**Literature:** (Suggest relevant scientific articles)

1. Huang, Daniel Q., Hashem B. El-Serag, and Rohit Loomba. "Global epidemiology of NAFLD-related HCC: Trends, predictions, risk factors and prevention." Nature Reviews Gastroenterology & Hepatology 18.4 (2021): 223-238.
2. Mardinoglu, Adil, et al. "Systems biology in hepatology: approaches and applications." Nature Reviews Gastroenterology & Hepatology 15.6 (2018): 365-377.
3. Yang, Hong, et al. "A network-based approach reveals the dysregulated transcriptional regulation in non-alcoholic fatty liver disease." Iscience 24.11 (2021): 103222.

**Project description:**

**(Describe shortly the project with background and clear objectives)**

Non-alcoholic fatty liver disease (NAFLD) is considered as one of the most important causes of liver disease, worldwide. However, no effective therapies are yet approved for the treatment of NAFLD. Hence, a comprehensive understanding of the underlying molecular mechanism of NAFLD is critical for the development of novel approaches for its prevention and treatment.

In this project, students will start with a reading section under Literature followed by a problem-solving section. In problem-solving section, students will use liver transcriptomics data from patients with NAFLD to study the dysregulated processes underlying NAFLD including differential expressed genes, molecular pathways, and transcription factor.

**Project tasks:**

1. **Overview of the topics:** students will firstly review the literature to get insights into the crucial challenges on these topics.
2. **Preparation:** students will explore database that will be used in the project, including Gene Expression Omnibus (GEO), KEGG, GO, TRRUST, Ensembl.
3. **Data retrieving and pro-processing:** students will download transcriptomics data from GEO database and quantify the abundance of gene, which will be carried out using the R statistical framework and UPPMAX computer sever R (need to be installed at local PC together with Rstudio).
4. **Differential analysis:** students will perform differential expression analysis using R.
5. **Functional enrichment analysis:** students will investigate the dysregulated functions using the results from differential analysis, including KEGG, GO, and TF enrichment analysis.
6. **Visualization of results**
7. **Report:** students will write report of the project.