Bornberglab
Molecular Evolution and Bioinformatics

Bachelor thesis project

Genome annotation of fruit fly genomes using the deep learning based gene predictor Tiberius

Background: Genome annotation is important for many genomic analyses. Until recently, the best performing tools required extrinsic evidence (e.g. RNA sequencing data). Tiberius is a new deep-learning-based tool that can predict genes from the genome sequence alone with similar quality as evidence-based tools.

Objectives: The aim of this project is to annotate a set of fruit-fly genomes with Tiberius and check the quality of the resulting annotations.

Requirements:

- Experience (or interest) in using the command line

Methods:

- Genome annotation
- Quality checks (BUSCO/COMPLEASM, OMArk)

Supervision: Marie Lebherz, Room 100.20, m.lebherz@uni-muenster.de, Molecular Evolution and Bioinformatics Group (https://bornberglab.org/people/lebherz/).

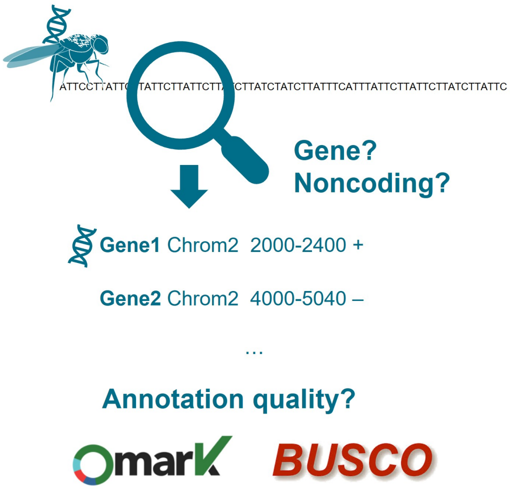

Figure 1: Genome Annotation Schematic

Selected Literature:
1) Gabriel, Lars, et al. "Tiberius: end-to-end deep learning with an HMM for gene prediction." Bioinformatics 40.12 (2024): btae685.
2)Seppey, Mathieu, Mosè Manni, and Evgeny M. Zdobnov. "BUSCO: assessing genome assembly and annotation completeness." Gene prediction: methods and protocols. New York, NY: Springer New York, 2019. 227-245.
3) Nevers, Yannis, et al. "Quality assessment of gene repertoire annotations with OMArk." Nature biotechnology 43.1 (2025): 124-133.