

Linz, 17. October 2022

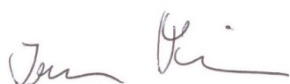
RE: Review for the bachelor thesis of Shehab Moukbel Ali Aldawla

It was a pleasure to have supervised the bachelor work and thesis of Shehab Moukbel Ali Aldawla with the title "Calling haplotypes from duplex sequencing data and its applications: A case study with ERBB2 data". For his thesis, Moukbel Ali Aldawla implemented several bioinformatics tools, including 'A haplotype analysis tool', which calls haplotypes within the libraries and classifies them based on tiers and allele frequencies, 'CS & PE-reads Visualization tools', wherein we could visualize the reads of called haplotypes for both consensus' sequences and PE-reads. Additionally, 'A haplotype counting tool', which summarizes the identified haplotypes and provides count numbers per.

Moukbel Ali Aldawla first evaluated the performance of tools in a provided data set for the ERBB2 gene, where he evaluated ~30 different libraries prepared with different protocols (Repaired & Non-repaired' libraries). He further conducted a deeper analysis on differences in haplotype number, composition and quality between the two different library groups. Specifically, he summarized the called haplotypes among the libraries in terms of haplotype length, composition, and region. He also identified shared haplotypes between groups. With his analysis, we identified potential sources of errors introduced during library preparation. Thus, his implementation has been quite valuable to filter out artefacts from biologically meaningful data. Moukbel Ali Aldawla also examined what could lead to the potential error sources.

Finally, Moukbel Ali Aldawla did several presentations throughout the semester, documenting his bachelor thesis progress. We also had weekly if not daily meeting (together with his co-supervisor Monika Heinzl) to direct him in different challenges and troubleshoot problems. Further, the analysis implementation was well done and presented orally and in a written work. The thesis was well organized and included all the critical information and analyses, as well as revisions after feedback, making his thesis a complete piece of work.

I grade this thesis as excellent and assign it the highest possible grade of "one".



Irene Tiemann-Boege