

Opinion of the bachelor's thesis of Rômulo Acácio dos Santos

supervisor's opinion opponent's opinion
 bachelor's thesis master's thesis

Author: **Rômulo Acácio dos Santos**
Name of work: **Comparison of bioinformatics pipelines for eDNA metabarcoding data analysis of fish populations in Czech reservoirs**
Study program, specialization: Applied Informatics (Bioinformatics)
Issue year: 2021

Name of opponent: Ing. Marta Vohnoutová
Institution: Institution of Applied Informatics
Contact e-mail: mvohnoutova@prf.jcu.cz

Professional level of work:

excellent very good average below average unsatisfactory

Formal mistakes:

almost none acceptable number numerous, non-essential essential

Results:

original original and taken nontrivial compilation taken from literature copied

Scope of work:

large standard acceptable unsatisfactory

Graphic, linguistic, and formal level:

excellent very good average below average unsatisfactory

Typographical errors:

almost none acceptable number numerous

Complex level of work:

excellent very good average below average unsatisfactory

Verbal statements, comments and remarks of the opponent:

The aim of the thesis was to compare five pipelines (Anacapa, Barque, metaBEAT, MiFish, SEQme) for environmental DNA (eDNA) metabarcoding using data from water samples taken in three reservoirs (Klíčava, Římov, and Žlutice) both in the summer and autumn seasons. The results were analysed by comparing the number of reads assigned, number of species detected, and ecological indices (alpha and beta diversity).

Picked-up data were filtered, got rid of “genetic pollution”, trimmed, and carefully prepared to avoid mistakes and false positive and false negative events. In the end, various appropriate statistical analysis ANOVA, post-hoc Tukey, and PERMANOVA were used to compare results from individual pipelines. For ecological diversities, Alpha (richness and Shannon index) and beta (Jaccard index) were used.

The compared pipelines for eDNA metabarcoding are of different logics and this thesis inspected their features and outputs thoroughly.

The student worked independently and took place in all steps of the work. He himself took the water samples from water reservoirs and processed samples in laboratories at Biology centre CAS.

The student used considerable amount of programming code (mostly Python and R), attached to this thesis. The results are described visualized in many clear graphs.

The important part of the work is also the reference database, originally derived from the reference database of the University of Hull, which the student updated. He also stored new sequences to the public databases.

It is obvious that the work comprises also a lot of study, the list of used literature is astonishing. The student describes the individual steps in logical and understandable way.

The only one objection – the student should distinguish better e.g. program name and simple text.

Conclusion:

The work is so complex and mature that could be easily accepted as a master's thesis. The result of the work has a large practical importance.

I recommend presenting the results of this thesis in some biological conference and/or publish it in some professional magazine.

Questions, discussion topics:

1. Reference database – how did you decide what should be there and how you adjust the original reference database from the University of Hull ? E.g.
2. Is it possible to discover in your water samples the creature who in not in your reference database (you did not simply expect it to be there in samples)?
3. Maylandia zebra was put intentionally to the water samples to test the pipelines ? Or could you explain the role of Maylandia zebra?
4. How you can explain that in some cases some pipeline discovers none and another considerably high amount? E.g. Alburnus Alburnus in Římov?

Work

recommend

do not recommend

To accept as bachelor's thesis

With the suggested mark:

excellent very good good unsatisfactory

Place, date and signature of the opponent:

In České Budějovice 5. May 2021

.....