PARIS LODRON UNIVERSITÄT SALZBURG

Fachbereich für Biowissenschaften
Abteilung Cancer Genetics/Epigenetics

Univ.-Ass. Ing. Dr. Angelika Heißl
Postdoc

Billrothstraße 11
A-5020 Salzburg | Austria
Tel.: +43 / (0) 662 / 8044 - 0
Mobil +43 / (0) 664 / 8044 - 2141
uni.service@sbg.ac.at
www.uni-salzburg.at

Salzburg, 15.11.2020

Opponent review for the bachelor thesis "short tandem repeat analysis in recombination hotspots across the human genome"

It was a pleasure for me to review the bachelor thesis of Mehdin Masinovic with the title "short tandem repeat analysis in recombination hotspots across the human genome". The thesis implemented an update to the published Bioconductor package *STRAH* of Heissl et al. 2019. This allows the user to search for other repeat types such as mono-, di-, tri- and tetranucleotide repeats, as well as DNA motifs in human recombination hotspots, other than the original package working only on poly-A/T repeats. Finally, the results were evaluated regarding their repeat enrichment over six zones surrounding the hotspot region.

For the provided bachelor thesis, extensive literature research was performed reflected by the thoroughly chosen publications cited in the work. The statistical analysis is well-described and reproducible. An impressive amount of repeat types was tested. Since repeats are highly linked to genetic diseases and cancer formation, *STRAH* represents a powerful tool to search for an enrichment in regions of interest. Nevertheless, one has to be very careful and critical when comparing a narrow region such as a recombination hotspot with a size of approximal 2 kb, with an outside zone representing 85% of the remaining genome. As published by Pratto et al. 2014, male double-strand break and crossover maps show an increased frequency of recombination near telomer and sub telomer regions. As an example, telomers are very rich in poly-G runs, or LINE and SINE elements are poly-A rich. In addition, mononucleotide runs with a size >9 bp seem to be evolutionary conserved and display low variability, which might reflect a potential functional role such as harboring transcription factor binding sites. Concluding, assumptions based on repeat type enrichments in relation to recombination hotspots have to be taken with care. There is a high probability of diverse mechanisms involved others than recombination.

The bachelor thesis is written in a very precise way with a clear structure. The fluent text allows to concentrate on the essential details of the work. I am highly impressed about the quality of the thesis by an early-stage researcher.

Overall, I recommend the thesis as an excellent work with the highest possible degree.

Yours sincerely,
Angelika Heißl