Review of: "AbDiver - A tool to explore the natural antibody landscape to aid therapeutic design"

Sofia Kossida[1], nika abdollahi, Taciana Manso[1]

1 French National Centre for Scientific Research

Potential competing interests: The author(s) declared that no potential competing interests exist.

This research paper, in our opinion, addresses an interesting and indispensable topic: the development of a free online tool that allows users to compare their query therapeutic antibody sequence to those found in natural antibody sequence repertoires. The text is well-written and is free of jargon. We find the data acquisition methodology to be persuasive. The tool is well designed and easy to use. The "Help" section is comprehensive and practical. Different aspects of the protocol are well-detailed; however, we believe that the following questions can be answered in a more reprehensible way.

Are the benchmarks that have been used for this study available?

In the "Discussion" section, you have mentioned the importance of employment of CDR3 or clonality search for "clone" discovery, have you voluntarily used the term clone and clonotype interchangeably?

Could you please explain the process of determining the "proportion value" and "study-specific Shannon entropy" in more detail and further illustrate your point by using a concrete example?

There also seems to be no way for the user to download the entirety or parts of the result tables. If it is necessary, the user has to redo the analysis or copy/past and take screenshots along the way. Have you considered adding the direct downloading function to your software?

Our suggested modifications for your text are either underlined or emboldened in the following phrases.

Missing articles/words :

- Monoclonal antibodies are the largest class of biotherapeutics. The development of successful antibody therapeutics requires **selecting** and engineering candidate sequences with favorable functional and developability properties.
- Knowledge of biologically possible mutations at specific positions can be employed to engineer **the** biophysical properties of these molecules (Venkataramani *et al.*, 2020).
- **With** the smaller number of light chains, we can find suitable profiles for results from a skew to- wards heavy chains in NGS depositions.
- For benchmarking, we employed a set of 742 therapeutic antibodies, discontinued, in clinical trials or approved in **the** USA and the EU extending a set from our previous study (Krawczyk *et al.*, 2021).
- Results of both searches are presented using interactive tables highlighting the leading themes in the studies (e.g., studied disease, vaccine) facilitating **the** further exploration of results.
- Therefore, our service succeeds at finding a high number of relevant matches for most of the

therapeutics in our dataset.

Missing hyphens :

- Therefore, despite restrictive length constraints that **produce** more relevant results, AbDiver identifies high-quality sequence matches.
- We hope that AbDiver will enable research-supporting applications to facilitate the decision**-**making process in **the** rational design of therapeutics during lead optimization.

Missing commas :

- ...new duplicates resulting in 738 unique heavy chains, 707 unique light chains, 686 unique CDRH3s**,** and 573 unique CDRL3s.
- **... ,** and disease states contributed by different studies, emphasizing frequency commonalities independent of study-specific biases.