

- **Title and abstract**
  - Does the title clearly reflect the content of the article? [] Yes, [] No (please explain), [] I don't know
  - Does the abstract present the main findings of the study? [] Yes, [] No (please explain), [] I don't know
- **Introduction**
  - Are the research questions/hypotheses/predictions clearly presented? [] Yes, [] No (please explain), [] I don't know
  - Does the introduction build on relevant research in the field? [] Yes, [] No (please explain), [] I don't know
- **Materials and methods**
  - Are the methods and analyses sufficiently detailed to allow replication by other researchers? [] Yes, [] No (please explain), [] I don't know
  - Are the methods and statistical analyses appropriate and well described? [] Yes, [] No (please explain), [] I don't know
- **Results**
  - Are the results described and interpreted correctly? [] Yes, [] No (please explain), [] I don't know
- **Discussion**
  - Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? [] Yes, [] No (please explain), [] I don't know
  - Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [] Yes, [] No (please explain), [] I don't know

LukProt is a eukaryote-wide protein database that combines much of the data from the previously assembled databases EukProt and AniProt but is enriched in datasets especially from early diverging animal lineages such as ctenophores, sponges and cnidarians. The author provides both web-portal and command line versions of the database so those with varying degrees of bioinformatic proficiency should be able to take advantage of LukProt. The curation effort is well-done, and I believe the comparative genomics community, especially those interested in animal origins, will find LukProt to be a useful resource. I have only minor suggestions for improvement below.

For the local version of database in the file LukProt\_metadata/README.txt it would be nice to have the tested versions of the dependencies necessary for the use of the associated companion scripts so the user does not have to refer back to the manuscript. It might also be helpful to include instructions on how a user might create a conda environment and install all correct versions of the necessary dependencies in case they are operating on community resources such as university clusters that are maintained by administrators and lack the

necessary permissions to install software system-wide themselves. This also will guarantee performance if system wide versions of dependencies change.

Additionally, some more detail regarding the usage of these associated scripts would be useful to add to the documentation. Provide example commands, detail the exact input and expected output, when should they be run etc. No instructions can be too clear and the easier the tools are to use the more users will adopt them in their regular workflows.