http://dx.doi.org/10.7124/bc.000AFE

Good software development practices improve citations of RNA seq tools

S. Sharma, K. Chhugani, Dhr. Deshpande, S. Mangul
University of Southern California
Los Angeles, CA, USA, 90007
susharma@usc.edu

Background/Aim: RNA sequencing (RNA-seq) has become an exemplary technology in biology and clinical applications, outperforming other transcriptomic profiling protocols due to its precision and accuracy in characterising transcriptomes. As RNA-seq analysis gains popularity and becomes an essential component of future precision medicine, examining the software development and open science practices used by RNA-seq tools is necessary to ensure and improve the accuracy of generated results. **Methods.** We performed a comprehensive assessment of the software development and open sciences practices of RNA-seq tools, and benchmarking practices across more than 319 RNA-seq tools developed between the years of 2008 to 2024 by manually data extraction. We extracted features of the RNA seq tools including availability of package managers, number of releases and the benchmarking practises adapted by the researchers. We plan to access the archive stability and engagement of the scientific community to software repositories like Github and Bitbucket to better measure recognition of the tools. **Results.** Here we show that good software development increases the number of citations for the tool with a statistically significant p-value of 9.1×10^{-6} (Mann-Whitney U test). **Conclusions.** In addition to software development practices, it is vital to follow best practices in benchmarking and validation of developed tools. Ideally, a newly developed tool would be compared against all previously developed tools1, but it is currently unknown if this practice is widely adopted in the bioinformatics community. Examining these studies is vital to develop RNA-seq tools that can not only enhance the precision and thoroughness of all future benchmarking studies, but also encourage tool developers to engage in more comprehensive benchmarking when releasing their tools. By highlighting promising trends in usability, accessibility, maintainability, and benchmarking practices, we aim to encourage wider adoption of these best practices within the biomedical research community. This collective effort will ultimately lead to the development of more accurate, robust, and user-friendly RNA-seq tools in the future.

Keywords: RNA-seq, RNA-seq tools, benchmarking studies, bioinformatics.

REFERENCES

1. *Mangul S et al.* Systematic benchmarking of omics computational tools. *Nat Commun.* 2019; **10**(1):1393.