

Part1.5: Documentation of Computer Work for Papers



Dana L Carper and Travis J Lawrence
Quantitative and Systems Biology
University of California, Merced

Reproducibility

- Very important so people can redo what you did
- The more information you give the easier it is for people to reproduce
- Using open source software allows for easier identification of any problems

What to Include for the Software

- Name of the software (example: FAST toolbox)
- Version (FAST toolbox v1.06)
- Any other packages with versions
 - Example: R version 3.2.2 ggplot2 version 3.3
- Any changes to default parameters or state you ran the default parameters

Other useful things to include

- Supplemental documentation with the commands you use
 - Even if not included in papers always a good idea to keep a record for yourself of what you did (similar to a lab notebook/field notebook)
- Any custom scripts that you made

Example of great documentation

OPEN  ACCESS Freely available online

 **PLOS** | COMPUTATIONAL
BIOLOGY

tRNA Signatures Reveal a Polyphyletic Origin of SAR11 Strains among Alphaproteobacteria

Katherine C. H. Amrine, Wesley D. Swingle[✉], David H. Ardell*

Program in Quantitative and Systems Biology, University of California, Merced, Merced, California, United States of America

Example of great documentation

Dataset S1.

Source code and data to reproduce Figure 2.

<https://doi.org/10.1371/journal.pcbi.1003454.s001>

(ZIP)

Dataset S2.

Source code and data to reproduce Figures 3 and S2.

<https://doi.org/10.1371/journal.pcbi.1003454.s002>

(ZIP)

Dataset S3.

Source code and data to reproduce Figure 4.

<https://doi.org/10.1371/journal.pcbi.1003454.s003>

(ZIP)