

## PUBLISHED VERSION


R. Daniel Kortschak, Brent S. Pedersen, and David L. Adelson  
**biogo/hts: high throughput sequence handling for the Go language**  
The Journal of Open Source Software, 2017; 2(10):168-1

Authors of JOSS papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License(CC-BY).

Published version <http://dx.doi.org/10.21105/joss.00168>

### PERMISSIONS

<http://creativecommons.org/licenses/by/4.0/>



**Attribution 4.0 International (CC BY 4.0)**

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

**You are free to:**


- Share** — copy and redistribute the material in any medium or format
- Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

**Under the following terms:**

- Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- No additional restrictions** — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

**8 November 2017**



<http://hdl.handle.net/2440/109346>

# bíogo/hts: high throughput sequence handling for the Go language

R Daniel Kortschak<sup>1</sup>, Brent S Pedersen<sup>2</sup>, and David L Adelson<sup>1</sup>

<sup>1</sup> School of Biological Sciences, The University of Adelaide <sup>2</sup> Department of Human Genetics, University of Utah

DOI: [10.21105/joss.00168](https://doi.org/10.21105/joss.00168)

## Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

## Licence

Authors of JOSS papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

## Summary

bíogo/hts provides a Go native implementation of the SAM specification (Group 2016) for SAM and BAM alignment formats (H. et al. 2012) commonly used for representation of high throughput genomic data, the BAI, CSI and tabix indexing formats, and the BGZF blocked compression format. The bíogo/hts packages perform parallelized read and write operations and are able to cache recent reads according to user-specified caching methods. The parallelisation approach used by the bíogo/hts package is influenced by the approach of the D implementation, sambamba by Tarazov et al. (T. A. et al. 2015). The bíogo/hts APIs have been constructed to provide a consistent interface to sequence alignment data and the underlying compression system in order to aid ease of use and tool development.

## References

- A., Tarasov, Vilella A., Cuppen E., Nijman I. J., and Prins P. 2015. “Sambamba: Fast Processing of Ngs Alignment Formats.” *Bioinformatics* 31 (12): 2032–4. doi:10.1093/bioinformatics/btv098.
- Group, The SAM/BAM Format Specification Working. 2016. “Sequence Alignment/Map Format Specification.” <https://samtools.github.io/hts-specs/SAMv1.pdf>.
- H., Li, Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., and 1000 Genome Project Data Processing Subgroup. 2012. “The Sequence Alignment/Map Format and Samtools.” *Bioinformatics* 25 (16): 2078–9. doi:10.1093/bioinformatics/btp352.