Biopython Project Update 2018

Ben Fulton,[a] Christian Brueffer,[b] Peter Cock,[c]
and the Biopython Contributors[d]

18th Bioinformatics Open Source Conference (BOSC) 2018, Portland, OR

Website: http://biopython.org
Repository: https://github.com/biopython/biopython
License: Biopython License Agreement (BSD like, see http://www.biopython.org/DIST/LICENSE)

The Biopython Project is a long-running distributed collaborative effort, supported by the Open Bioinformatics Foundation, which develops a freely available Python library for biological computation [1]. We present here details of the planned Biopython releases in 2018. Since the 1.70 release became available in July 2017, 55 unique users have contributed code fixes or enhancements, including 32 newcomers. This reflects our policy of trying to encourage even small contributions.

2018 is expected to see the release of three new Biopython releases: 1.71, 1.72, and 1.73. The first is expected to be available in April 2018, and will include Python 3-compatible rich comparison for all Bio.PDB objects, helping to unify the API and easing sorting. It also provides several new options and fixes for parsing and writing files, including PIR format for sequences, mmCIF files, and Gene files. In addition, new codon tables 27-31 from NCBI (NCBI genetic code table version 4.2) were added. This release will also include updated HTML documentation based on the Sphinx documentation generator.

A further release, Biopython 1.72, is expected in the summer of 2018, and, if warranted, a final release for the year sometime in the fall.

In 2017 we started a re-licensing plan, to transition away from our liberal but unique *Biopython License Agreement* to the similar but very widely used *3-Clause BSD License*. We are in the process of reviewing the code base authorship file-by-file, in order to gradually dual license the entire project.

All releases fix miscellaneous bugs, enhanced the test suite, and continued efforts to follow the PEP8 and PEP257 coding style guidelines which is now checked automatically with GitHub-integrated continuous integration testing using TravisCI. We now also use AppVeyor for continuous integration testing under Windows. Current efforts include improving the unit test coverage, which is easily viewed online at CodeCov.io.

References

[1] Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., de Hoon, M.J. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**(11) 1422-3. doi:10.1093/bioinformatics/btp163

[a]Research Technologies, Indiana University, Bloomington, Indiana. Email: befulton@iu.edu
[b]Department of Clinical Sciences, Lund University, Lund, SE
[c]Information and Computational Sciences, James Hutton Institute, Invergowrie, Dundee, UK
[d]See contributor listing on GitHub.