



Biopython Project Update

Tiago Antao ^{*}, Wibowo Arindrarto[†], Peter Cock[‡], João Rodrigues[§], Eric Talevich[¶], *et al.*

15th Bioinformatics Open Source Conference (BOSC) 2014, Boston, MA, USA

Website: <http://biopython.org>

Repository: <https://github.com/biopython/biopython>

License: Biopython License Agreement (MIT style, see <http://www.biopython.org/DIST/LICENSE>)

We present the latest updates from the Biopython project, a long-running, distributed collaboration producing a freely available Python library for biological computation [1]. Biopython is supported by the Open Bioinformatics Foundation (OBF).

There have been two releases since BOSC 2013: version 1.62 and 1.63, both involving contributions made by new and returning developers. New features since version 1.62 include support for the population genetic tool fastsimcoal, a wrapper for samtools, and various enhancements to existing modules. Version 1.64 is under development and is expected to have been released by BOSC 2014 [*NB: to be revised during abstract review*]. We have been focusing on compatibility with Python 3, officially supporting it since version 1.62. In version 1.63, we removed the 2to3 library required for code conversion to Python 3. In lieu of 2to3, we updated our code base to be compatible with both Python 2 and Python 3. This change is also reflected in our Tutorial & Cookbook, which uses code compatible with both Python versions.

In addition to local installation on various operating systems, Biopython is also now available in the Galaxy Tool Shed [2] as a package dependency. Galaxy tools requiring Biopython can now specify this dependency explicitly and choose from three different Biopython versions: 1.61, 1.62, or 1.63. We plan to continue releasing for Galaxy for upcoming versions.

We participated in Google Summer of Code (2013) under the umbrella of the National Evolutionary Synthesis Center (NESCent). Two students were selected: Yanbo Ye, working on enhancements to Bio.Phylo, and Zheng Ruan, working to add codon alignment and analysis support. Both students have finished their project successfully, with Yanbo's code integrated into the existing Bio.Phylo module and Zheng's code submitted as a pull request under review. We hope to mentor another student in GSoC 2014 who will be working on adding lazy-parsing support to Bio.SeqIO [*NB: to be revised during abstract review*].

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
- [2] Blankenberg, D., Von Kuster, G., Bouvier, E., Baker, D., Afgan, E., Stoler N., The Galaxy Team, Taylor, J., Nekrutenko, A. (2014) Dissemination of scientific software with Galaxy ToolShed. *Genome Biology* **15** 403. doi:10.1186/gb4161

*

[†]Leiden University Medical Center, Eindhovenweg 20, 2332 ZC Leiden, The Netherlands. Email: bow@bow.web.id

[‡]Information and Computational Sciences, James Hutton Institute (formerly SCRI), Invergowrie, Dundee DD2 5DA, UK

[§]

[¶]