Biopython Project Update 2016

Christian Brueffer[a], Tiago Antão[b], Peter Cock[c], Eric Talevich[d], Michiel de Hoon[e], Wibowo Arindrarto[f], Leighton Pritchard[c], Anuj Sharma[g], Eric Rasche[h], Aaron Rosenfeld[i], Connor T. Skennerton[j], Marco Galardini[k], Markus Piotrowski[l], and the Biopython Contributors

17th Bioinformatics Open Source Conference (BOSC) 2016, Orlando, USA

Website: http://biopython.org
Repository: https://github.com/biopython/biopython
License: Biopython License Agreement (MIT style, 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 latest Biopython release - version 1.66. New features include: extended Bio.KEGG and Bio.Graphics modules to support drawing KEGG pathways with transparency; extended "abi" Bio.SeqIO parser to decode almost all documented fields used by ABIF instruments; a QCPSuperimposer module using the Quaternion Characteristic Polynomial algorithm for superimposing structures to Bio.PDB; and an extended Bio.Entrez module to implement the NCBI Entrez Citation Matching function and to support NCBI XML files with XSD schemas. Additionally we fixed miscellaneous bugs, enhanced our test suite and continued our efforts to abide by the PEP8 coding style guidelines.

We are currently preparing a new release – version 1.67 – that will deprecate the ability to compare SeqRecord objects with "==", which sometimes lead to surprising results. In addition it will feature a new experimental Bio.phenotype module for working with Phenotype Microarray data; updates to Bio.Data to include NCBI genetic code table 25, covering Candidate Division SR1 and Gracilibacteria; an update to Bio.Restriction to include the REBASE May 2016 restriction enzyme list; updates to BioSQL to use foreign keys with SQLite3 databases; as well as corrections to the Bio.Entrez module and the MMCIF structure parser.

Our website has been migrated from MediaWiki to GitHub Pages and is now under version control. The continuous integration process on GitHub has been enhanced by including external services like Landscape, Quantified Code and Codecov to perform quality review, test coverage analysis and generation of quality metrics.

Finally, our range of Docker containers has been greatly enhanced. In addition to a basic container that includes Python 2 and 3 with Biopython and all its dependencies, as well as a BioSQL container, we now also provide two versions of Jupyter notebook containers: a basic one, and a version including the Biopython tutorial as notebooks.

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]Department of Clinical Sciences, Lund University, Lund, SE. Email: christian.brueffer@med.lu.se
[b]Division of Biological Sciences, University of Montana, Montana, USA
[c]Information and Computational Sciences, James Hutton Institute (formerly SCRI), Invergowrie, Dundee, UK
[d]Department of Dermatology, University of California San Francisco, San Francisco, CA, USA
[e]Division of Genomic Technologies, RIKEN Center for Life Science Technologies, Yokohama, JP
[f]Sequencing Analysis Support Core, Leiden University Medical Center, Leiden, NL
[g]Department of Informatics and Telecommunications, University of Athens, Athens, Greece
[h]Center for Phage Technology, Department of Biochemistry and Biophysics, Texas A&M University, USA
[i]Department of Biomedical Engineering, Science & Health Systems, Drexel University, USA
[j]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[k]EMBL-EBI, Wellcome Trust Genome Campus, Cambridge, UK
[l]Department of Plant Physiology, Ruhr-Universität Bochum, DE