

What is Biopython?

- ▶ Collection of modules for biological computation in Python
 - ▶ Sequence handling and motifs, parsers, database queries, protein structures, phylogenetics, tool wrappers and more.
- ▶ Started in 1999, first release in 2000
- ▶ Open source and freely available (Biopython license)

In a Nutshell, Biopython...

... has had 11,237 commits made by 192 contributors representing 291,535 lines of code

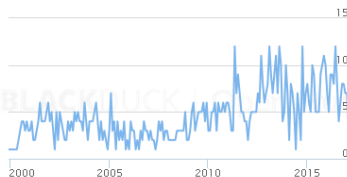
... is mostly written in Python with an average number of source code comments

... has a well established, mature codebase maintained by a very large development team with increasing Y-O-Y commits

... took an estimated 75 years of effort (COCOMO model) starting with its first commit in December, 1999 ending with its most recent commit 3 months ago

Community

Contributors per Month



Source: <https://www.openhub.net/p/biopython>

Lots of new contributors!

- ▶ Anthony Bradley
- ▶ Ben Fulton
- ▶ Carlos Pena
- ▶ Connor T Skennerton
- ▶ Iddo Friedberg
- ▶ Kai Blin
- ▶ Kristian Davidsen
- ▶ Markus Piotrowski
- ▶ Olivier Morelle
- ▶ Peter Cock
- ▶ Stefans Mezulis
- ▶ Tiago Antao
- ▶ Travis Wrightsman
- ▶ Uwe Schmitt
- ▶ Xiaoyu Zhuo
- ▶ Adam Kurkiewicz
- ▶ Adam Novak
- ▶ Adrian Altenhoff
- ▶ Allis Tauri
- ▶ Andrew Guy
- ▶ Andrew Sczesnak
- ▶ Blaise Li
- ▶ Brandon Carter
- ▶ Foen Peng
- ▶ Francesco Gastaldello
- ▶ Francisco Pina-Martins
- ▶ Hector Martinez
- ▶ Jack Twilley
- ▶ Jeroen Van Goey
- ▶ Joshua Meyers
- ▶ Kurt Graff
- ▶ Leonhard Heizinger
- ▶ Marcin Magnus
- ▶ Maximilian Greil
- ▶ Michał J. Gajda
- ▶ Milind Luthra
- ▶ Oscar G. Garcia
- ▶ Richard Neher
- ▶ Sourav Singh
- ▶ Spencer Bliven
- ▶ Steve Marshall
- ▶ Veronika Berman

Biopython 1.68(released 2016-08-25)

New Features in Biopython 1.68

- ▶ Final Release supporting Python 2.6
- ▶ Added extension to Bio.PDB to allow for parsing of the new MMTF format. Requires
- ▶ Bio.pairwise2 rewritten to address speed and problems with local alignments.
- ▶ SeqGui and xbbtools rewritten using tkinter.
- ▶ Biopython now uses HTTPS instead of HTTP to connect to Entrez and QBLAST API.

Biopython 1.69 (released 2017-04-06)

Start of re-licensing plan

- ▶ 1.69 marks the start of start of our re-licensing plan, to transition away from liberal *Biopython License Agreement* to the widely used *3-Clause BSD License*.
- ▶ The code base-authorship is being reviewed file-by-file, in order to gradually dual license the entire project.

New parser for ExPASy cell line database

- ▶ Bio.ExPASy:
 - ▶ Cellosaurus cell line database, cell ontologies and cell line catalogues.
 - ▶ Now accessible through the Bio.ExPASy module.

Other new features to Biopython 1.69

- ▶ MAF files are now supported using the `Bio.AlignIO.MafIO` module
- ▶ Also offers indexed access for large files using `SQLite3`.
- ▶ Update to `Bio.Restriction` to include REBASE February 2017 restriction enzyme list.
- ▶ `Bio.PDB.PDBList` can download PDBx/mmCif (new default), PDB (old default), PDBML/XML and MMTF format protein structures.
- ▶ `Bio.Affy` module supports version 4 of Affymetrix CEL format

Miscellaneous

- ▶ Miscellaneous bug fixes
- ▶ Test suite enhancements
- ▶ Better PEP8 and PEP257 coding style adherence
- ▶ Enhanced PyPy support by taking advantage of NumPy and compiling most of the Biopython C code modules

What's cooking for Biopython 1.70?

Biopython 1.70-dev

- ▶ Module `Bio.AlignIO`
 - ▶ Support for XMFA file format.
 - ▶ using `Bio.AlignIO.MauveIO`
- ▶ Module `Bio.SearchIO`
 - ▶ Support for new arguments to read and write blast-xml files.
- ▶ Drop support for Python 3.3

New logo for Biopython

- ▶ Version 1.70 marks the transition towards a new logo for Biopython.



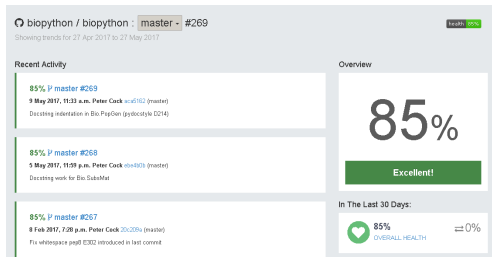
- ▶ The new logo was created by Patrick Kunzmann in 2017 and had support from Python Software Foundation and the Biopython developers community.

Supported Python Versions

- ▶ Python 2.7
- ▶ Python 3.4
- ▶ Python 3.5
- ▶ PyPy 5.0
- ▶ PyPy3 2.4
- ▶ Jython 2.7

Continuous Integration

- ▶ TravisCI
- ▶ Experimentation with Build Stages for CI testing.
 - ▶ Codecov.io (test coverage)
 - ▶ Quantified Code (metrics and automatic pull requests)
 - ▶ Landscape.io (“health score”)
- ▶ Currently enabled by default: Codecov.io, Quantified Code



Conclusion

- ▶ Lots of new contributors
- ▶ Lots of new stuff
- ▶ Turn new contributors into recurring ones
- ▶ Biopython 1.70 in the near future

Acknowledgements

- ▶ Peter Cock
- ▶ Biopython Community



Resources!

Website:

- ▶ <http://biopython.org>

Repositories:

- ▶ Main: <http://github.com/biopython/biopython>
- ▶ Website:
<https://github.com/biopython/biopython.github.io>

Mailing lists:

- ▶ General list: biopython@biopython.org
- ▶ Developers list: biopython-dev@biopython.org

Biostars:

- ▶ <https://www.biostars.org/t/biopython/> (“biopython” category)