



# A survey of the most recent Python packages for use in biology

Yasmin Makki Mohialden<sup>1</sup>

Computer Science Department, Collage of Science, Mustansiriyah University, Baghdad-Iraq  
ymmiraq2009@uomustansiriyah.edu.iq

Nadia Mahmood Hussien<sup>2</sup>

Computer Science Department, Collage of Science, Mustansiriyah University, Baghdad-Iraq  
nadia.cs89@uomustansiriyah.edu.iq

Liqaa Shaker Kadhim Al-Timimi<sup>3</sup>

Department of Biology, Collage of Science, Mustansiriyah University, Baghdad-Iraq  
liqaashaker2@uomustansiriyah.edu.iq

Noor Jameel Mousa<sup>4</sup>

Department of Biology , Collage of Science, Mustansiriyah University, Baghdad-Iraq  
[noorjameelmaster@gmail.com](mailto:noorjameelmaster@gmail.com)

Riam sabah abbood<sup>5</sup>

Department of Biology , Collage of Science, Mustansiriyah University, Baghdad-Iraq  
[riamsabah.mc.s.zoo.2020@uomustansiriyah.edu.iq](mailto:riamsabah.mc.s.zoo.2020@uomustansiriyah.edu.iq)

Itidal Saad Mohammed<sup>5</sup>

Computer Science Department, Collage of Science, Mustansiriyah University, Baghdad-Iraq  
[yinchardlik77m@gmail.com](mailto:yinchardlik77m@gmail.com)

255

## Abstract

Over the previous decade, libraries, object orientation, software architecture, and visual programming have emerged. Software reuse is these technologies' goal. Despite several notable advances in disciplines like library design and a diversity of types being used in various situations, there are still many areas in which substantial work awaits. This study examines the specifications of the most important Python libraries for biological applications released in the past two years. Through the references of several 2022–2023 research publications, this study found various open-source and free Python libraries for free apps. The installation platform, general descriptors, exact specs, library link, and application fields were compared.

**Keywords:** Python packages, bioinformatics, Matplotlib , Gget, Biopython, bioPAX, Biopython.

**DOI Number:** 10.48047/NQ.2023.21.2.NQ23029

**Neuroquantology 2023; 21(2):255-258**

## 1. Introduction

Guido Rossum made Python, a programming language that works with objects, in 1989. It

eISSN1303-5150

can be used to make quick prototypes of applications that are hard to build. It can be made to work with C or C++ and has interfaces

[www.neuroquantology.com](http://www.neuroquantology.com)



for many OS system calls and libraries. Python is used by NASA, Google, YouTube, BitTorrent, and many other big companies. Python is a language that works very well. With fewer lines of code, your programs will be able to do more than in many other languages. The way Python is written will also help you write code that is "clean." Your code will be easier to understand, fix, add to, and build on than code written in other languages. Python is used to make games, online apps, solve business problems, and make internal tools for a wide range of interesting companies. Python is also used a lot in science, both in the classroom and on the job [1]. We looked into how to use interpreted language guides to make a dynamic, programmable environment where different applications can be used to link components at a high level [2]. It can be hard to do research in bioinformatics with tools that are made for sale. Several parts of open-source projects make it easy to make custom pipelines or analyses. Modeling biological networks also requires mathematical equations, but as networks get bigger and more

changes happen, it gets harder to manage equations directly or add in models that have already been made. Several people are building graphical standards, rule-based languages, and integrated software workbenches at the same time to make biological modeling easier, but none of them fully meet the need for models that are clear, expandable, and reusable [4]. Several free or open-source libraries that can be used in biology will be listed in this review. We will pay attention to Python-based software.

## 2. Bioinformatics

Bioinformatics uses biology, computation, and information technology to organize and preserve enormous amounts of biological data from genetics, molecular biology, and biotechnology [5].vc3 Bioinformatics streamlines and efficiently interpret genome, transcriptome, and proteome data [6].

3. **Recent python packages in biology:** Table 1 displays the most recent Python packages in biology.

Table 1 python libraries for biology

Name	platform
Matplotlib[7]	web
Description	<b>Python desktop plotter (primarily in 2D)</b>
Supported Categories	<b>Static, dynamic, and interactive visualization charts and hierarchies</b>
url	<a href="https://matplotlib.org/">https://matplotlib.org/</a>
SKiMpy[8]	windows
Description	<b>Automated creation and analysis of large-scale kinetic models for signaling, gene expression, and metabolism Furthermore</b>
Supported Categories	<b>reconstructs huge kinetic biological process models. Its extensive analytic methods model biological systems.</b>
url	<a href="https://github.com/EPFL-LCSB/SKiMpy">https://github.com/EPFL-LCSB/SKiMpy</a>
AlphaPulldown[9]	command line interface
Description	<b>AlphaFold-Multimer facilitates high-throughput modeling of higher-order oligomers and protein-protein interaction screens. AlphaFold-Multimer identifies new protein-protein interactions (PPIs).</b>
Supported Categories	<b>graphs and confidence scores.</b>



url	<a href="https://www.embl-hamburg.de/AlphaPulldown/">https://www.embl-hamburg.de/AlphaPulldown/</a>
Gget[10]	command line tool a
Description	effectively queries Ensembl genomic reference databases
Supported Categories	a set of interoperable modules for genetic data analysis database querying.
url	<a href="https://github.com/pachterlab/gget">https://github.com/pachterlab/gget</a>
GSEAPy[11]	command line
Description	It efficiently assessed large single-cell datasets.
Supported Categories	several visual methods (gseaplot, heatmap, dotplot, barplot and ringplot).
url	<a href="https://github.com/zqfang/GSEAPy">https://github.com/zqfang/GSEAPy</a>
Biopython [12]	Linux, Mac OS X and Windows.
Description	computational biology and bioinformatics Python tools.
Supported Categories	computation molecular biology
url	<a href="https://biopython.org/">https://biopython.org/</a>
RGT (Regulatory Genomics Toolbox) [13]	command line
Description	Integrating regulatory genomics data with a computational library
Supported Categories	high-throughput regulatory genomics data analysis with Python
url	<a href="https://github.com/CostaLab/reg-gen">https://github.com/CostaLab/reg-gen</a>
PyBioPAX[14]	
Description	bioPAX model processing software PyBioPAX implements the BioPAX Level 3 object model as Python classes and a BioPAX OWL processor to deserialize OWL material into these objects.
Supported Categories	as a package on PyPI with the source code available at <a href="https://github.com/indralab/pybiopax">https://github.com/indralab/pybiopax</a>
PySB pronunciation: "Pie Ess Bee"[15]	command line
Description	allows rule-based biochemical system modeling. For model simulation and analysis, NumPy, SciPy, and SyPy work well. It is a rule-based modeling language
Supported Categories	streamlines the complex process of generating equations describing protein interactions into a Python-based domain-specific language. PySB automatically constructs, simulates, or evaluates BNGL or Kappa rules. PySB makes it easy to partition models into modules and call biological macro libraries. These qualities enhance model reuse, transparency, and correctness.
url	<a href="http://pysb.org/">http://pysb.org/</a>



#### 4. Conclusion

In this study, many open-source and free Python libraries for free applications were identified through the references of a number of research papers published for the years 2022–2023. Several comparison factors were identified, such as the installation platform, general descriptors, exact specifications, link to obtain the library, and application fields.

#### References

- [1] M. Eric, G. Yuan, Python Crash Course: A Hands-on project-based Introduction to Programming, Posts & Telecom Press, 2016.
- [2] Sanner, Michel F. "Python: a programming language for software integration and development." *J Mol Graph Model* 17.1 (1999): 57-61.
- [3] Harry Mangalam, The Bio\* toolkits — a brief overview, *Briefings in Bioinformatics*, Volume 3, Issue 3, September 2002, Pages 296–302, <https://doi.org/10.1093/bib/3.3.296>
- [4] Lopez, Carlos F., et al. "Programming biological models in Python using PySB." *Molecular systems biology* 9.1 (2013): 646.
- [5] Arthur Lesk, Introduction to Bioinformatics, third ed., Oxford University Press, 2008.
- [6] Soria-Guerr R.E., Nieto-Gomez R., Govea-Alonso B.O., Rosales-Mendoza S. An overview of bioinformatics tools for epitope prediction: implications on vaccine development. *J. Biomed. Inf.* 2015;53:405–414.
- [7] Bisong, E. Matplotlib and Seaborn. Building Machine Learning and Deep Learning Models on Google Cloud Platform; Apress: Berkeley, CA, USA, 2019; pp. 151–165.
- [8] D. R. Weilandt et al., "Symbolic Kinetic Models in Python (SKiMpy): Intuitive modeling of large-scale biological kinetic models," bioRxiv, p. 2022.01.17.476618, 2022.
- [9] Jin Xu, Jessie Jiang, Herbert M Sauro, "SBMLDiagrams: a python package to process and visualize SBML layout and render", *Bioinformatics*, Volume 39, Issue 1, January 2023.
- [10] Laura Luebbert, Lior Pachter, *Bioinformatics*, "Efficient querying of genomic reference databases with gget" 39(1), 2023, btac836.
- [11] Fang, Z., Liu, X. and Peltz, G., " GSEAPy: a comprehensive package for performing gene set enrichment analysis in Python, ". *Bioinformatics*, Volume 39, Issue 1, January 2023, btac757
- [13] Profile Zhijian Li, ChaoChung Kuo, Fabio Ticoni, Mina Shaigan, Eduardo Gade Gusmao, Manuel Allhoff, Martin Manolov, Martin Zenke, Ivan G. Costa, "RGT: a toolbox for the integrative analysis of high throughput regulatory genomics data", bioRxiv, 2023.
- [14] B. M. Gyori and C. T. Hoyt, "PyBioPAX: biological pathway exchange in Python," *J. Open Source Softw.*, vol. 7, no. 71, p. 4136, Mar. 2022,
- [15] Hina Goyal, Vijay Kumar, Anoop K. Saini, Rajeew Gupta, "A multifunctional Schiff base with aggregation-induced enhanced emission, gelation, and mechanochromic properties for anti-counterfeiting applications", *Materials Today Chemistry*, January 2023.

