

The Spoken Tutorial project

- Self explanatory - uses simple language
- Audio-video - uses multisensory approach
- Small duration - has better retention
- Learner-centered - learn at your own pace
- Learning by doing - learn and practice simultaneously
- Empowerment - learn a new FOSS

Target group

- Undergraduates / Postgraduates
- Research scholars
- Teachers

Workshops

The Spoken Tutorial Project Team conducts work-shops on Biopython and other FOSS using spoken tutorials and gives certificates to those who pass an online test

For more details, please write to contact@spoken-tutorial.org

The Spoken Tutorial Project is funded by the National Mission on Education through Information and Communication Technology, Ministry of Human Resource Development, Government of India.

Contact Us

Email : contact@spoken-tutorial.org
info@spoken-tutorial.org

Website : <http://spoken-tutorial.org>

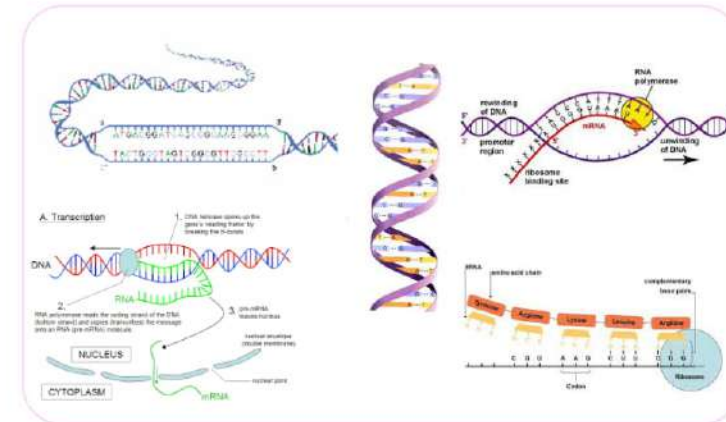


Spoken Tutorial by IIT Bombay is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

All trademarks within this document belong to their legitimate owners

Spoken Tutorial

BIOPYTHON



National Mission on Education through
Information and Communication Technology
(NMEICT)
www.sakshat.ac.in

An MHRD initiative

<http://spoken-tutorial.org>

Introduction:

- Biopython is a set of freely available Python tools for computational biology and bioinformatics.
- Biopython runs on many platforms (Windows, Mac, Linux and Unix).
- Source code is easily available under Biopython Licence. The Biopython web site <http://www.biopython.org> provides an online resource for modules, scripts and web links.

Features:

- It has tools for performing common operations on sequences, such as translation, transcription and weight calculations.
- It has ability to parse bioinformatics files into Python utilizable data structure.
- It supports the following formats: Blast output, Clustalw, FASTA, GenBank, PubMed and Medline, ExPASy files, SwissProt etc.
- Files in the supported formats can be iterated over record by record or indexed and accessed via a Dictionary interface

- Code to deal with popular online bioinformatics destinations such as:
 - NCBI – Blast, Entrez and PubMed services
 - ExPASy – SwissProt and Prosite entries, as well as Pros
- Interfaces to common bioinformatics programs such as:
 - Standalone Blast from NCBI
 - Clustalw alignment program
 - EMBOSS command line tools
- Code to perform classification of data using k Nearest Neighbors, Naive Bayes or Support Vector Machines.
- Code for dealing with alignments, including a standard way to create and deal with substitution matrices.
- Code making it easy to split up parallelizable tasks into separate processes.
- GUI based programs to do basic sequence manipulations, translations, BLASTing, etc.
- Extensive documentation and help with using the modules, online wiki documentation, the web site, and the mailing list.

- Integration with BioSQL, a sequence database schema also supported by the BioPerl and BioJava projects

Uses:

- Parse BLAST results (standalone and web).
- Run biology related programs (blastall, clustalw, EMBOSS).
- Deal with FASTA formatted files.
- Parse GenBank files.
- Parse PubMed, Medline and work with online resource.
- Parse Expasy, SCOP, Rebase, UniGene, SwissProt.
- Data classification (k Nearest Neighbors, Bayes, SVMs)
- Aligning sequences; CORBA interaction with Bioperl and BioJava
- SQL database storage through BioSQL; Neural Networks.
- Neural Networks; Genetic Algorithms
- Structural biology PDB.
- Create specialized substitution matrices.