

Programmazio
in Python
per la bioin-
formatica

Semplice
utilizzo di
Biopython

Programmazione in Python per la bioinformatica

University of Verona

Cosa e' biopython

- Risorsa disponibile on-line contenente moduli e script per la bioinformatica
- Biopython:
 - parser per formati tipici (e.g., FASTA, Genbank)
 - accesso a servizi e databases (e.g., NCBI)
 - interfacce per programmi e applicativi vari (e.g., BLAST)
- <http://biopython.org>

Installare Biopython

Programmazione
in Python
per la bioin-
formatica

Semplice
utilizzo di
Biopython

Controllare se biopython e' installato

```
>>> import Bio
>>> print(Bio.__version__)
1.58
```

Dove scaricare biopython

Disponibile a questo link:

<http://biopython.org/wiki/Download>

Istallazione dipende dalla piattaforma (disponibile su Windows, Linux, Mac)

Semplice esempio di uso

Programmazione
in Python
per la bioin-
formatica

Semplice
utilizzo di
Biopython

Trovare a quale specie appartiene un frammento di DNA

```
>unknown sequence|
ATGCGAGTGTTGAAGTTCGGCGGTACATCAGTAGCAAATGCAGAACGTTTTCTGCGGGTTGCCGATATTC
TGGAAAGCAATGCCAGGCAGGGGCAGGTGGCCACCGTCTCTCTGCCCCCGCCAAAATAACCAACCACCT
GGTGCGCATGATTGAAAAAACCATAGCGG
```

Vedere i seguenti file

- unknown_sequence.fasta
- processblast.py
- output.blast

Utilizzare BLAST da un database su internet

Programmazione
in Python
per la bioin-
formatica

Semplice
utilizzo di
Biopython

Interrogare un database nt utilizzando il programma blastn

```
>>> from Bio.Blast import NCBIWWW
>>> fasta_string = open("unknown_sequence.fasta").read()
>>> result_handle = NCBIWWW.qblast("blastn", "nt", fasta_string)
```

Leggere il risultato in un formato semplice da analizzare

```
>>> result_handle
<cStringIO.StringIO object at 0x1470660>
>>> from Bio.Blast import NCBIXML
>>> blast_record = NCBIXML.read(result_handle)
```


Risultati

Programmazione
in Python
per la bioin-
formatica

Semplice
utilizzo di
Biopython

Cosa abbiamo ottenuto

```
***Alignment***
('seq: ', u'gi|675820454|emb|LM997016.1| Escherichia coli genome ass
embly FHI87, scaffold scaffold-24_contig-0.0_1_1132105_[organism:Esc
herichia')
('length: ', 1132105)
('e value: ', 3.32809e-80)
***Alignment***
('seq: ', u'gi|675820240|emb|LM996803.1| Escherichia coli genome ass
embly FHI82, scaffold scaffold-29_contig-0.0_1_748312_[organism:Esch
erichia')
('length: ', 748312)
('e value: ', 3.32809e-80)
```

Per saperne di piu'

Tutorial : <http://biopython.org/DIST/docs/tutorial/Tutorial.html>