



Aplicação Biopython Web

Esta é a documentação do projeto final apresentado para a profa. Dra. Ana Winck, na disciplina de Banco de Dados para a Área Biomédica, 2017/2.

Sumário

Contribuidores	1
BioPython	1
Descrição Geral da Solução	1
Preparação do Ambiente	2
Funcionalidades da aplicação	2
Trabalhos Futuros	3

Contribuidores

- Gabriel Leal: gabrielle@ufcspa.edu.br
- Giovani Bettoni: giovanib@ufcspa.edu.br
- Marília Rosa: mariliarosilva@gmail.com

BioPython

O projeto Biopython é uma associação internacional de desenvolvedores de ferramentas computacionais para biologia molecular em Python. Atualmente ele possibilita que o desenvolvedor tenha acesso à uma biblioteca que facilite a resolução dos problemas de Bioinformática e Biologia Computacional, usando módulos, scripts em Python de pesquisadores em ciências biológicas de todo o globo.

Descrição Geral da Solução

Biopython possui várias estruturas e algoritmos implementados bem como uma série de outros recursos que nos permitem focar mais na lógica da resolução do problema.



Entretanto, nem todos os pesquisadores estão dispostos a usar ferramentas na linha de comando, ou aprender a sintaxe de programação, sendo que o seu objetivo é que com alguns cliques, por exemplo, tenha uma comparação entre duas sequências genéticas, através de técnica de Alinhamento de Sequências.

Em Bioinformática, um alinhamento de sequências é uma forma de organizar sequências primária de DNA, RNA ou proteína para identificar regiões similares que possam ser consequência de relações funcionais, estruturais ou evolucionárias entre elas. Um dos algoritmos mais conhecidos para alinhamento de sequências é o Algoritmo Needleman-Wunsch. Ele não é o único, e seria interessante estabelecer comparações entre os diferentes algoritmos de uma forma mais amigável, do que a linha de comando.

Preparação do Ambiente

Tentamos tornar a execução a mais enxuta possível. Se você já um usuário do pacote Anaconda, o Flask já vem no pacote básico. Como queremos fazer testar, a documentação sugere que criemos um ambiente de testes, entretanto para fins de execução global, ou seja, em qualquer lugar do sistema, usamos o pip:

```
$ sudo pip install Flask
```

Permitir a execução do script run.sh:

```
$ chmod +x run.sh
```

Executar o script:

```
$ ./run.sh
```

Manter seus arquivos .fasta dentro de '/BiopythonWeb/app/uploads/'. Caso exista algum erro, verificar o TODO do arquivo em '/BiopythonWeb/app/controllers/defaults.py'.

Funcionalidades da aplicação

Transforma a sequência de aminoácidos em proteínas (processo biológico de tradução).

Entrada:



BiopythonWeb

Início

Alinhamento Global

Alinhamento Local

Projeto final apresentado para a profa. Dra. Ana Winck, na disciplina de Banco de Dados para a Área Biomédica, 2017/2.

A junção de uma ferramenta livre de Biologia Molecular Computacional, denominada Biopython e um poderoso framework web chamado Flask.

Contribuidores:

Gabriel Leal: gabrielle@ufcspa.edu.br

Giovani Bettoni: giovanib@ufcspa.edu.br

Marília Rosa: mariliarosilva@gmail.com

Informações sobre a sequência .fasta:

Upload dos arquivos .fasta

Escolher arquivo exSeq.fasta

Upload

Saída:

BiopythonWeb

Início

Alinhamento Global

Alinhamento Local

A Sequência de DNA obtida é: ATGACACGCGTTCAATTTAAACACCACCATCATCACCATCATCCTGACTAG

A proteína que obtemos foi : MTRVQFKHHHHHHHPD*

Trabalhos Futuros

Como este é um protótipo desenvolvido há pouco tempo, algumas funcionalidades não foram agregadas, mas poderiam ser:

Alinhamento global e local: uma das funcionalidades mais interessantes do Biopython é a possibilidade de entregar os scores de alinhamentos locais e globais. Como a técnica de alinhamento múltiplo hoje é uma realidade, poderíamos ter nas guias uma interface com a possibilidade de termos esses dados.



BiopythonWeb

Início

Alinhamento Global

Alinhamento Local

```
iniciando o alinhamento...
lendo fastas...

----- exSeq.fasta -----
gi|556503834:2089996-2090046
gi|556503834:2089996-2090046 Escherichia coli str. K-12 substr. MG1655, complete genome
ATGACACGCGTTCAATTTAAACACCACCATCATCACCATCATCCTGACTAG

----- exSeq2.fasta -----
gi|556503834:2089996-2090046
gi|556503834:2089996-2090046 Escherichia coli str. K-12 substr. MG1655, complete genome
ATGACACGCGTTCAATTTAAACACCACCATCATCACCATCATCCTGACTAG
gi|0000000000
gi|0000000000 nova
TGACTAGATGACACGCGTTCAATTTAAACACCACCATCACACACCATCATC

----- alinhamento global -----

-----ATGACACGCGTTCAATTTAAACACCACCATCA-TCACCATCATCCTGACTAG
|||
TGACTAGATGACACGCGTTCAATTTAAACACCACCATCACACACCATCATC-----
Score=221.5

----- alinhamento local -----

-----ATGACACGCGTTCAATTTAAACACCACCATCA-TCACCATCATCCTGACTAG
|||
TGACTAGATGACACGCGTTCAATTTAAACACCACCATCACACACCATCATC-----
Score=248
```