

GeMapCom: Uma Ferramenta Para Mapeamento Comparativo e Visualização de Dados Genômicos

Um Manual do Usuário da Aplicação

The layout of this page will be finished by ACTA

The layout of this page will be finished by ACTA

Contents

| | |
|---|-----------|
| Contents | 3 |
| 1 Instalação | 5 |
| 1.1 Instalação das bibliotecas no Linux/UNIX | 5 |
| 1.2 Instalação da Ferramenta no Linux/UNIX | 5 |
| 1.2.1 Clonando um Repositório | 6 |
| 1.2.2 Download dos Arquivos, via pasta compactada | 7 |
| 2 Primeiros Passos | 9 |
| 2.1 Execução da Ferramenta | 9 |
| 2.2 Comparar Sequências | 10 |
| 2.3 Converter Sequências | 11 |
| 2.4 Ferramenta BLAST | 13 |
| 3 Funcionalidades oferecidas pelo GeMapCom | 15 |
| References | 17 |
| Appendices | 19 |
| Original publications | 23 |

1 Instalação

As instruções de instalação estão disponíveis apenas para o LINUX/UNIX. Esta seção fornece instruções para a realização da instalação da ferramenta no sistema operacional Linux através da plataforma GitHub onde o mesmo está arquivado.

A ferramenta está desenvolvida na linguagem python, versão 3.0, e utiliza diversas bibliotecas externas. Para que a ferramenta rode em seu Linux é necessário ter instalado o python e suas respectivas bibliotecas.

1.1 Instalação das bibliotecas no Linux/UNIX

Partindo do ponto que seu sistema operacional não possui nenhuma biblioteca instalada ou da versão desejada, na Tabela 1 está descrito algumas bibliotecas necessárias para seus primeiros passos na ferramenta.

Table 1. Bibliotecas necessárias para o uso da ferramenta.

| Biblioteca | Método de Instalação |
|----------------------------------|---|
| Python 3.0 | <code>sudo apt-get install python3</code> |
| Gerenciador de Pacotes do Python | <code>sudo apt-get install python3-pip</code> |
| PyQt5 | <code>pip3 install pyqt5</code> |
| BioPython | <code>pip3 install biopython</code> |
| MatPlot | <code>pip3 install matplotlib</code> |
| Pandas | <code>pip3 install pandas</code> |
| BLAST | <code>sudo apt install ncbi-blast+</code> |

Após realizada as instalações das bibliotecas descritas, dar-se início a próxima etapa, a instalação da ferramenta em seu desktop.

1.2 Instalação da Ferramenta no Linux/UNIX

A instalação é realizada através da plataforma GitHub, uma plataforma de hospedagem de código-fonte e arquivo, para a realização da instalação é necessário ir até a plataforma usando o link descrito abaixo:

– <https://github.com/PedroAzevedo141/GeMapCom>

Acessando o link acima, clique na opção download, até aparecer a tela ilustrada pela Figura 1;

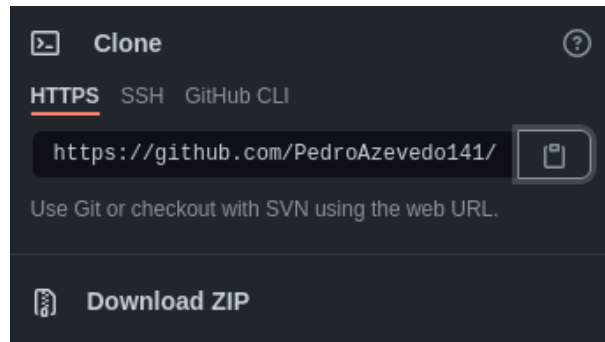


Fig. 1. Pop-up para realizar a instalação do repositório.

Nessa tela o GitHub disponibiliza dois tipos de download dos arquivos presentes no repositório. A primeira é através do clone do repositório, que consta da manipulação dos arquivos através do GIT e assim manipular e usufruir de cada um deles. A segunda opção de download é através de uma pasta compactada, onde poderá baixar os arquivos do repositório mas não irá conseguir manipulá-los com o GIT.

1.2.1 Clonando um Repositório

Para clonar o repositório através do HTTPS, copie e cole o endereço destacado na Figura 2.

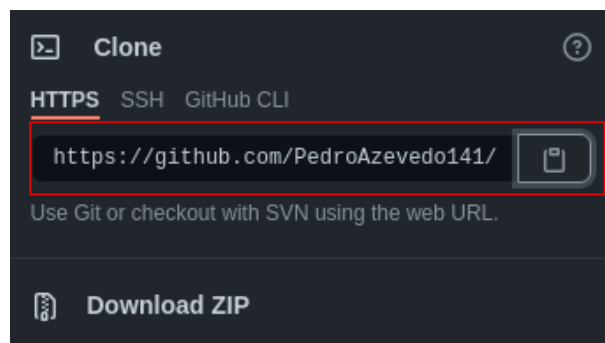


Fig. 2. Pop-up para realizar a instalação do repositório via GIT.

Em seguida abra seu Terminal, vá até o diretório que deseja ter a ferramenta e digite git clone e cole a URL que você copiou anteriormente, como ilustra abaixo:

– `git clone https://github.com/PedroAzevedo141/GeMapCom.git`

1.2.2 Download dos Arquivos, via pasta compactada

Neste tipo de download só basta clicar na opção “Download ZIP”, ilustrado na Figura 3, e em seguida extrair a pasta baixada no diretório desejado.

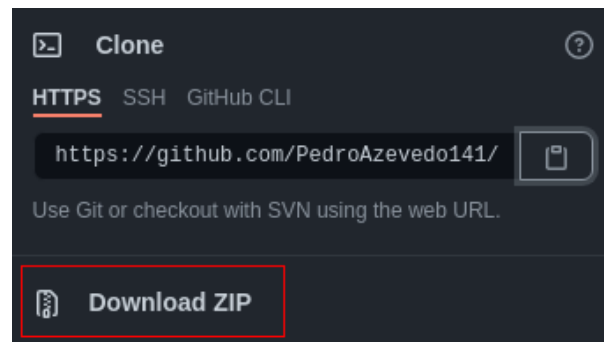


Fig. 3. Pop-up para realizar a instalação do repositório via ZIP.

2 Primeiros Passos

Essa seção tem como funcionalidade mostrar pontos recorrentes sobre o uso da ferramenta, mostrando sua inicialização e suas funcionalidades. Os arquivos no qual serão simulados através das imagens abaixo estarão todos presentes no repositório para fim de um entendimento prático e rápido.

2.1 Execução da Ferramenta

De início, percorra pelo terminal até a pasta na qual está situado a ferramenta, então execute o seguinte comando:

– *python3 GeMapCom.py*

Assim que for pressionado o Enter a ferramenta será inicializada, tendo como página inicial ilustrada pela a Figura 4.



Fig. 4. Tela Inicial da ferramenta GeMapCom.

E na ferramenta estão disponibilizados três tipos de funcionalidades, sendo elas, Comparar Sequências, Converter Sequências e a Ferramenta BLAST. Sendo cada uma, respectivamente, descritas mais abaixo:

2.2 Comparar Sequências

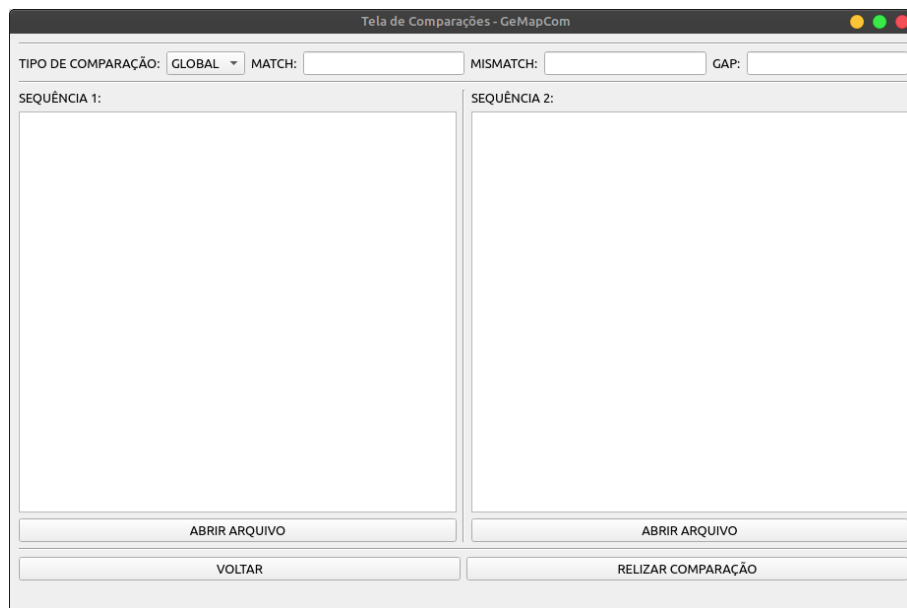


Fig. 5. Tela de Comparar Sequências disponibilizada pela ferramenta.

Na tela ilustrada pela Figura 8, nota-se a presença de algumas funcionalidades, como o tipo de comparação que pode ser a **GLOBAL** e a **LOCAL**. Sendo o alinhamento global uma forma de otimização que "força" o alinhamento a cobrir todo o comprimento de todas as sequências interrogadas (query). Diferente dos alinhamentos locais que identificam regiões de similaridade dentro de sequências longas que são geralmente bastante divergentes em um todo.

Nota-se também a presença dos campos de MATCH, MISMATCH e GAP que são campos que alteram os scores finais do resultado dependendo dos valores colocados nos campos.

Por fim, os campos de inserção dos arquivos .fasta que irão conter as sequências. Caso seu arquivo esteja de outro formato que não seja .fasta, vá para a seção **XX**. Caso ainda seja o primeiro uso e ainda não tenha sequências para realizar comparações, no

diretório da ferramenta contém uma pasta nomeada de “Primeiro_Uso” e lá irá conter alguns arquivos .fasta para a realização de algumas comparação iniciais.

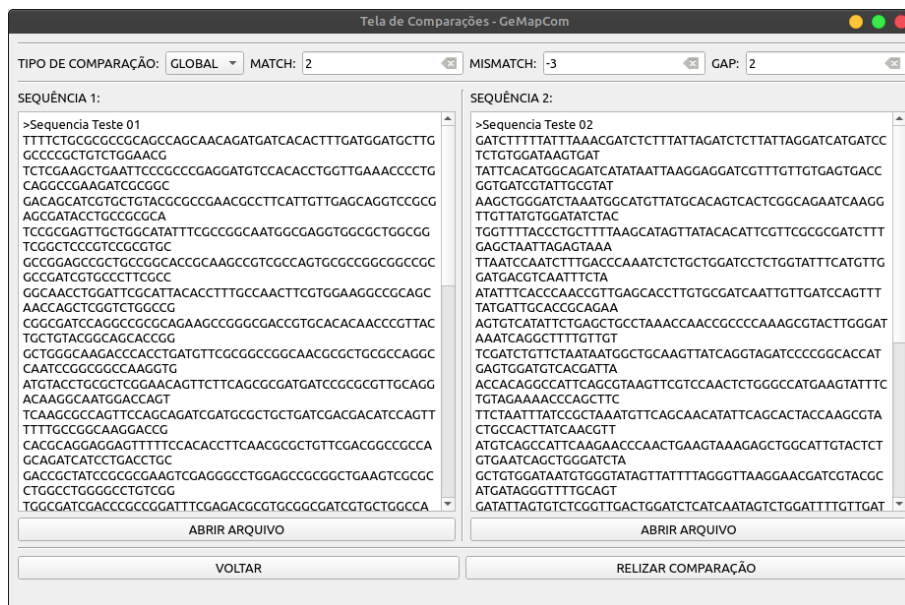


Fig. 6. Tela de Comparar Sequências disponibilizada pela ferramenta com os arquivos inseridos e os campos informados.

Os arquivos presentes nos campos “Sequência 1” e no “Sequência 2” foram inseridos, respectivamente, os arquivos “seq1.fasta” e o “seq2.fasta” presentes na pasta “Primeiro_Uso”. Obtendo assim o resultado ilustrado na Figura 6.

E por fim, como ilustra a Figura 7, existem duas formas de salvar o resultado obtido, sendo através de uma imagem, um arquivo .png, ou através de um PDF.

2.3 Converter Sequências

A Figura 8 ilustra a tela de conversão disponibilizada pela ferramenta, que tem como objetivo auxiliar o usuário a obter suas sequências sempre no arquivo .fasta. Atualmente a ferramenta tem controle sobre a conversão dos arquivos GenBank e FASTQ para o FASTA.

Na pasta “Primeiro_Uso”, disponibilizado no diretório da ferramenta, contém dois arquivos com esses formatos, sendo eles, “exemplo_FASTQ.fastaq” e o “exemplo_genbank.gbkl”. Esta tela tem como única funcionalidade a inserção do arquivo que

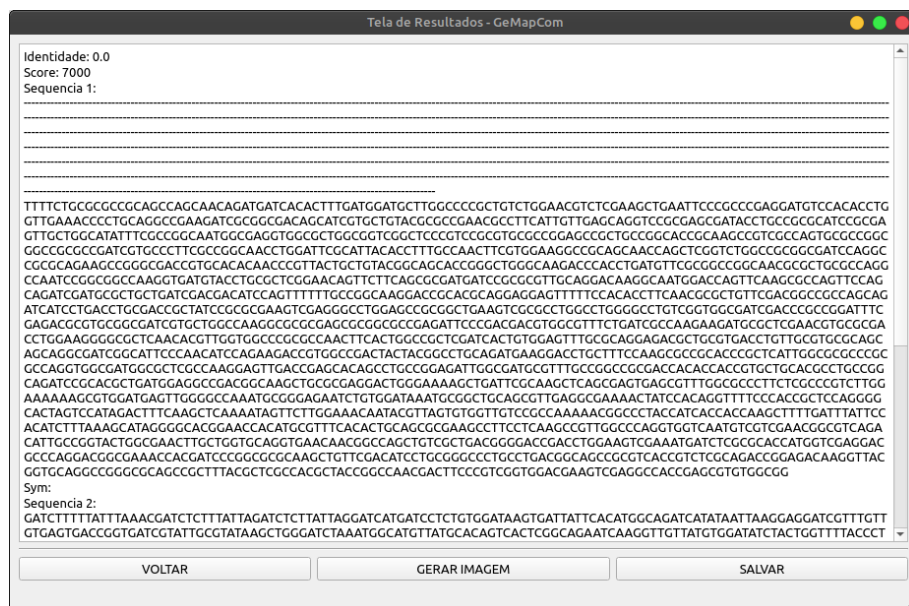


Fig. 7. Tela de Resultados das comparações das sequências.

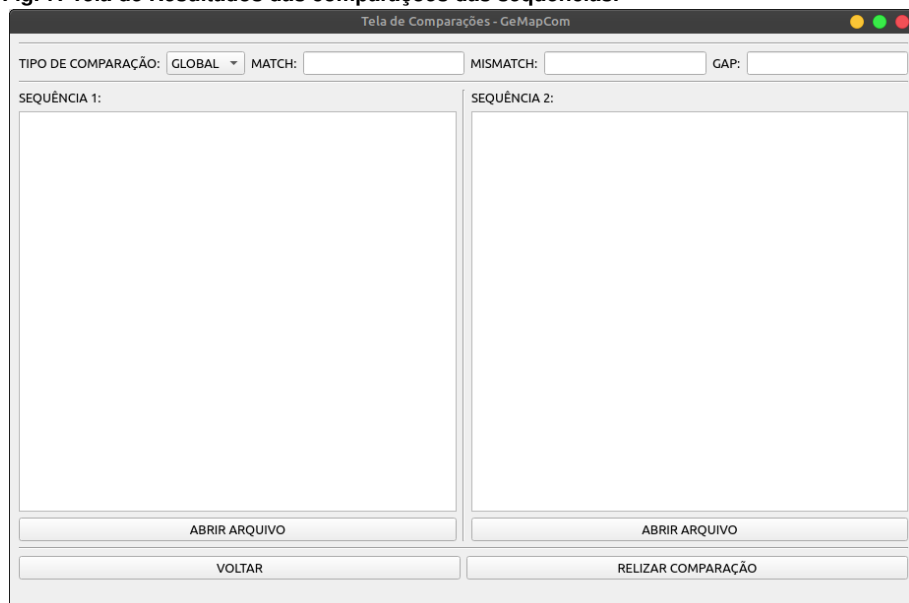


Fig. 8. Tela de Converter Sequências disponibilizada pela ferramenta.

deseja converter e informar no campo “Tipo de Conversão” o formato do seu arquivo inserido e assim converter e salvar seu novo arquivo .fasta.

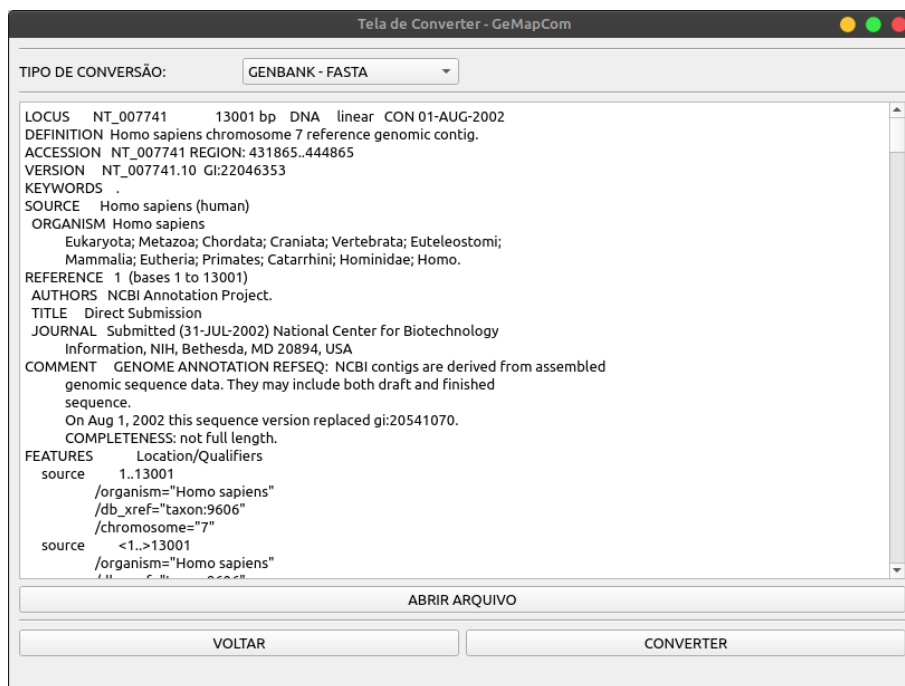


Fig. 9. Tela de Conversão disponibilizada pela ferramenta com o arquivo inserido e o tipo de conversão informado.

2.4 Ferramenta BLAST

A Figura 10 ilustra a funcionalidade BLAST presente na ferramenta. O BLAST é um software muito utilizado dentro da bioinformática por ser uma ferramenta muito rápida e assim não poderia faltar na GeMapCom.

A funcionalidade desta tela consta da inserção dos arquivos .fasta, dos parâmetros GAP, MATCH e MISMATCH e dos tipos de saída. Os tipos de saída são funções do BLAST que pode entregar seu alinhamento de diversas formas, como em forma de tabela, ou em um formato .XML e podendo até ser de forma binária.

Na pasta “Primeiro_Uso”, disponibilizado no diretório da ferramenta, contém dois arquivos com esses formatos, sendo eles, “Yersinia-BLAST.fasta” e “Xanthomonas-BLAST.fasta” estes dois arquivos .fasta já contém dois tipos de sequências maiores, sendo assim, insira cada um deles nos campos “Selecione o Query” e no campo “Selecione o Subject”. Em seguida informe os parâmetros desejados e o seu modo de saída. Ao inserir tudo clique em “Alinhar” obtendo seu resultado no formato desejado.

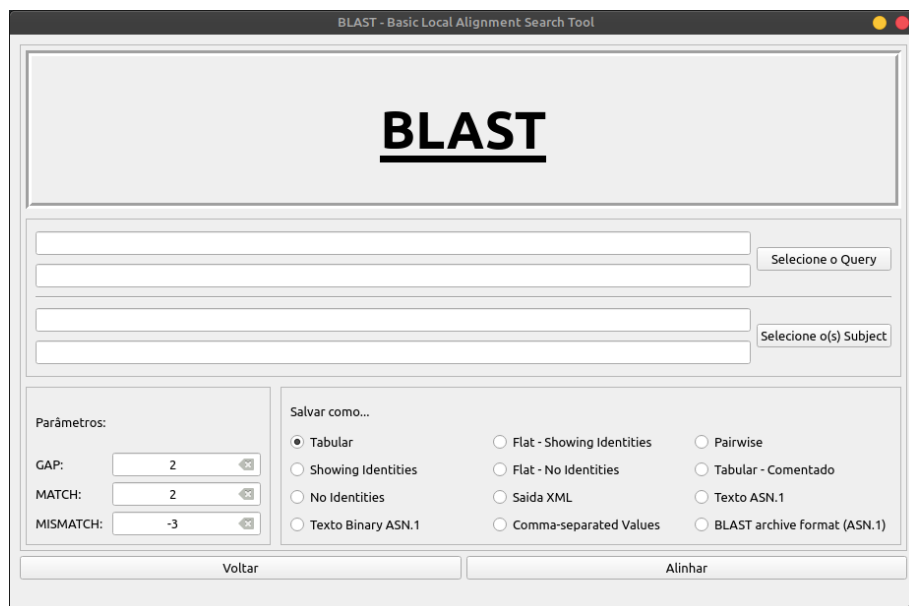


Fig. 10. Tela do BLAST disponibilizada pela ferramenta.

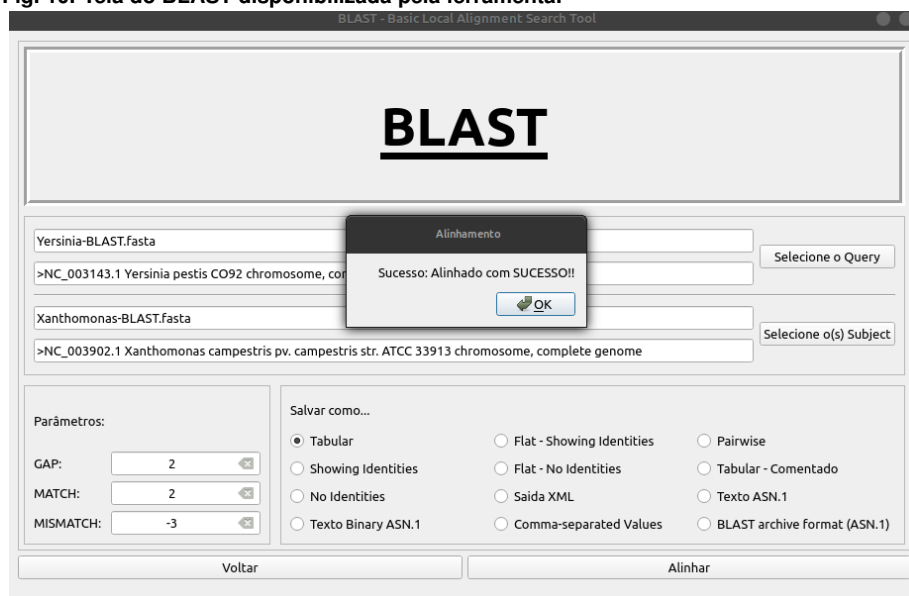


Fig. 11. Tela do BLAST disponibilizada pela ferramenta com os arquivos inseridos e todos os parâmetros informados.

3 Funcionalidades oferecidas pelo GeMapCom

References

Appendix 1 Some stuff A

Appendix 2 Some stuff B

Original publications

- I Nurkkala K & Varjonen V (2007) Instructions for authors. Instructions for authors of manuscripts intended for publication in the series Acta Universitatis Ouluensis. URI: http://www.kirjasto.oulu.fi/assets/files/julkaisutoiminta/templates/ActaOulu2009_en.pdf.

Reprinted with permission from xxxx (I), yyyy (II)...

Original publications are not included in the electronic version of the dissertation.