

Guía de Usuario para la Aplicación de Extracción de Identificadores de FASTA

Introducción

Esta herramienta permite extraer identificadores de secuencias de archivos FASTA ubicados en un directorio específico y escribir todos estos identificadores únicos en un archivo de texto. La aplicación cuenta con una interfaz gráfica (GUI) que facilita la selección de directorios y archivos, así como la ejecución del proceso de extracción.

Uso de la Aplicación

Opción 1: Ejecutar la Aplicación desde Windows

Para ejecutar la aplicación en Windows, simplemente haz doble clic en el archivo `extraerIDs.exe`.

Interfaz Gráfica

Paso 1: Interfaz Gráfica

Aparecerá una ventana con la interfaz gráfica de la aplicación.

Paso 2: Selección de Directorios y Archivos

- ****Directorio de Entrada****: Haz clic en el botón "Seleccionar" junto a este campo para seleccionar el directorio que contiene los archivos `.fasta` que deseas procesar.

- ****Archivo de Salida****: Haz clic en el botón "Seleccionar" junto a este campo para seleccionar el archivo donde se guardarán los identificadores extraídos.

Paso 3: Ejecución

- Haz clic en el botón "Ejecutar" para iniciar el proceso.
- La aplicación realizará las siguientes acciones por cada archivo `.fasta`` en el directorio de entrada:
 1. Extraerá los identificadores de las secuencias.
 2. Agregará los identificadores únicos al conjunto de todos los identificadores.
- Los resultados se guardarán en el archivo de salida en formato `.txt``.

Paso 4: Finalización

- Una vez que el proceso haya terminado, aparecerá un mensaje indicando que el proceso ha finalizado y el archivo de salida contendrá los identificadores únicos extraídos.

Opción 2: Ejecutar la Aplicación desde Python

Requisitos

- Python 3.x instalado en el sistema.
- Librerías necesarias: ``tkinter``.

Instalación de Dependencias

Si deseas ejecutar el script desde Python, puedes instalar la librería ``tkinter`` si no la tienes instalada:

```
pip install tk
```

Ejecución desde Python

1. Abre una terminal y navega al directorio donde se encuentra el archivo `main.py`.
2. Ejecuta el siguiente comando:

```
python main.py
```

Detalles de Entrada y Salida

- **Archivos de Entrada**: Archivos en formato `.fasta` que contienen secuencias de proteínas. Cada secuencia debe comenzar con un identificador en una línea que comienza con `>`.
- **Archivo de Salida**: Archivo de texto (`.txt`) que contendrá todos los identificadores únicos extraídos de los archivos `.fasta`.

```
1. Archivos de Entrada:  
  
- ./fasta_files/example1.fasta  
  
...  
  
>seq1  
  
ATGCATGCATGC  
  
>seq2  
  
ATGCATGCATGC  
  
...  
  
- ./fasta_files/example2.fasta  
  
...  
  
>seq3
```

```
ATGCATGCATGC
```

```
>seq4
```

```
ATGCATGCATGC
```

```
...
```

2. ****Ejecución****:

- Desde Windows: Doble clic en `extraerIDs.exe`.
- Desde Python: Ejecuta `python main.py` en la terminal.

3. ****Resultado****:

- `all_identifiers.txt`

```
...
```

```
seq1
```

```
seq2
```

```
seq3
```

```
seq4
```

```
...
```