

MANUAL PROGRAMA F.A.R. (FILTRACION, ALINEAMIENTO, REPORTE)

1. Prerrequisitos

- Python 3.12+

Para instalar el lenguaje, puede visitar la siguiente pagina para descargar los paquetes necesarios:

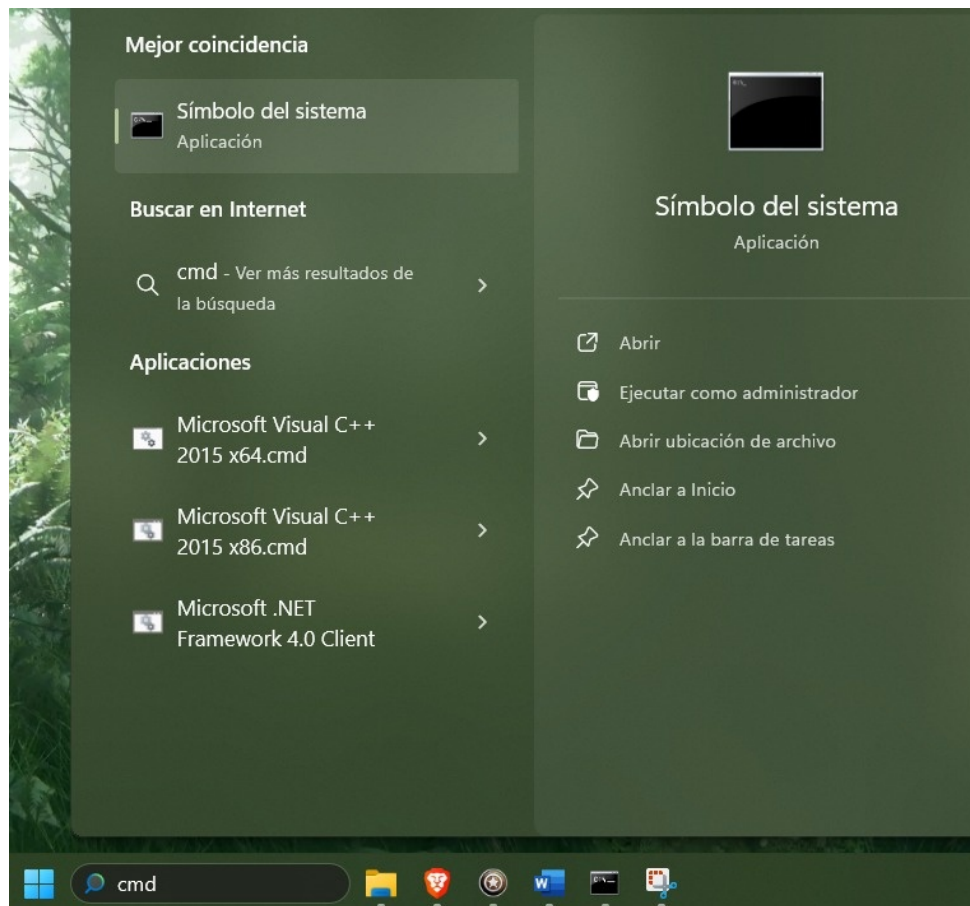
<https://www.python.org/downloads/windows>

Una vez hecho esto, instalar la versión correspondiente al sistema en uso (64-bit o 32-bit). Para comprobar si esta instalado correctamente, puede abrir una terminal de comandos (cmd) y escribir el comando `py --version` (también puede usar `python3 --version`):

```
C:\Users\Presston\Downloads\archivo base>py --version
Python 3.13.5

C:\Users\Presston\Downloads\archivo base>
```

Este terminal de comandos, en windows, se localiza de la siguiente forma:



antes de continuar, es necesario instalar dependencias necesarias para el modulo de comparación de cebadores con estos comandos (este paso es necesario tanto para usuarios linux como usuarios windows):

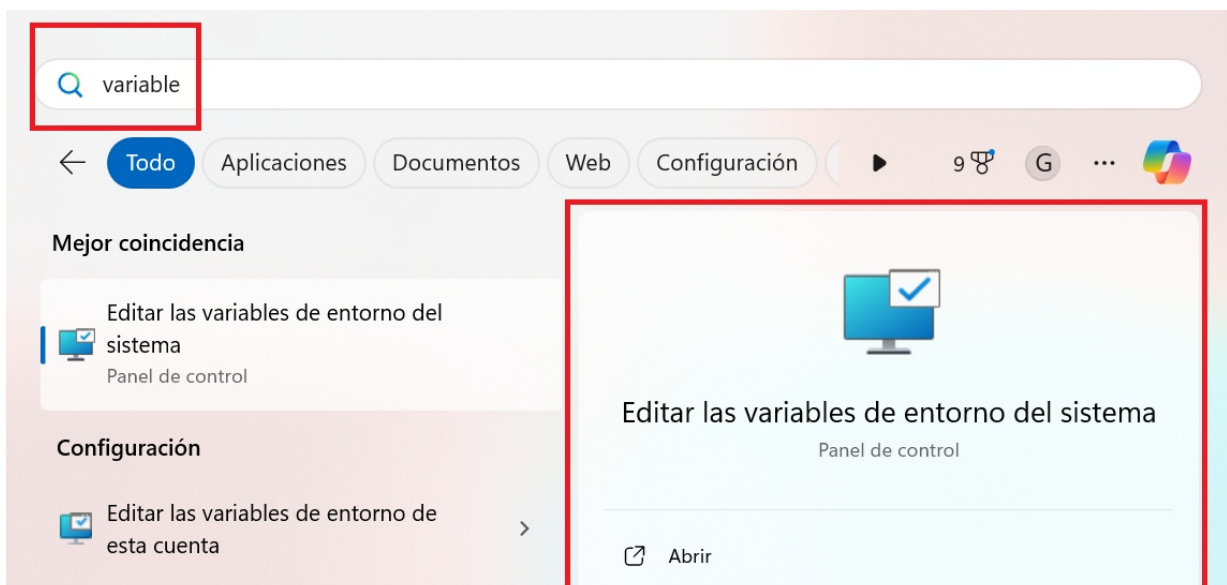
- m pip install reportlab
- pip install biopython

Herramientas externas: El modulo 2 requiere que ciertas dependencias estén previamente instaladas:

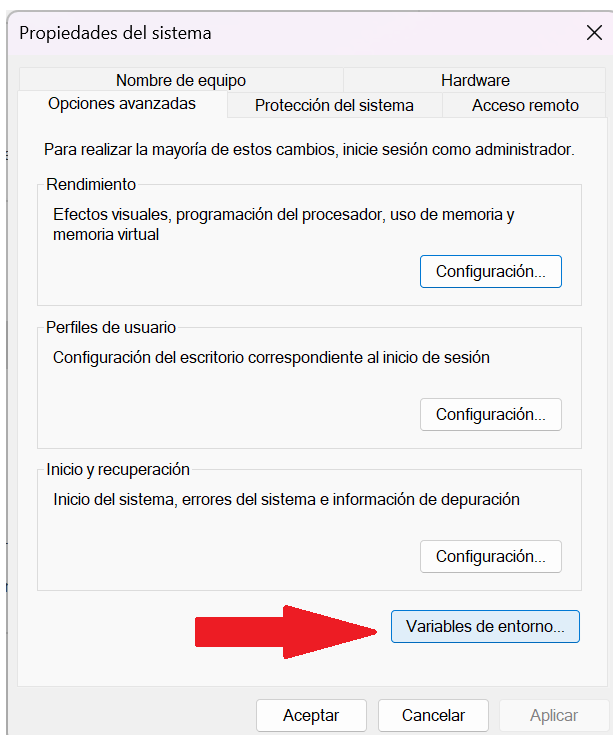
Usuarios Windows

I. [MAFFT](https://MAFFT.cbrc.jp/alignment/software/windows_without_cygwin.html) (alineamiento de secuencias): debe descargar el archivo “all-in-one” disponible en: https://MAFFT.cbrc.jp/alignment/software/windows_without_cygwin.html y situar el ejecutable batch (.bat) en el mismo directorio con los módulos de python a ejecutar

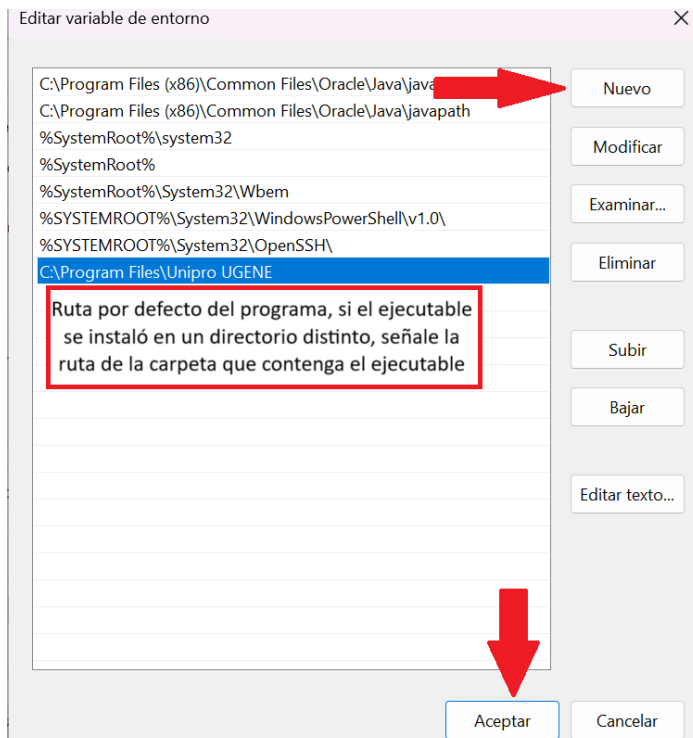
II. [UGENE](https://ugene.net/download-all.html) (extracción de consenso): se debe descargar el ejecutable e instalarlo, adicionalmente se debe agregar al PATH la ruta del directorio donde se instalo el ejecutable correspondiente. Dicho ejecutable lo puede localizar en la pagina oficial de ugene <https://ugene.net/download-all.html>



paso 1: Abrir el menú de variables del sistema



Paso 2: agregar variables del sistema



Paso 3: agregar la ruta donde se encuentra el ejecutable de UGENE

Usuarios Linux

MAFFT/UGENE: se recomienda instalar MAFFT y ugene en el directorio opt localizado en la raíz del sistema de archivos y agregar las rutas al respectivo directorio en el entorno de línea de comandos. Para ello, en una terminal, usando un editor de texto (nano, vim o emacs) para editar el archivo de configuración .bashrc (o .zshrc, dependiendo de la distribución). Puede agregar el directorio usando "nano ~/.bashrc" (sin comillas) en una terminal de comandos y agregar la línea "export PATH="\$PATH:/opt/ugene-52.1" al final (use la ruta del directorio donde está instalado el programa) y salvar el cambio para poder ejecutar desde línea de comandos cualquier instrucción del módulo de análisis

```
!:$ nano ~/.bashrc
```

Mismo cambio se debe realizar con MAFFT y el nombre del directorio si se instala de esta forma. Alternativamente, en distribuciones Debian puede usar el instalador que se puede descargar desde <https://MAFFT.cbrc.jp/alignment/software/linux.html>

2. Estructura de Archivos: para que los módulos cumplan su función correctamente deben estar en el mismo directorio de la siguiente forma:

```
----- DIRECTORIO -----

1- Filtracion      # filtro de archivo original
2- Alineamiento    # Ejecución de alineamiento MAFFT y consenso ugene/biopython
3- Reporte         # Resultado pdf que selecciona conjunto óptimo
modulo maestro     # Ejecución de los 3 modulos en forma secuencial
cebadores.txt      # Lista con conjuntos de cebadores a comparar
base.fasta         # Archivo de secuencias nucleotidicas original (editable)
parametros.json    # Variables editables
```

4. Ejecución Paso a Paso

```
if __name__ == "__main__":  
    # Orden de ejecución  
    modulos = [  
        ("1-Filtracion.py", "Filtrado de secuencias"),  
        ("2-Alineamiento.py", "Alineamiento con MAFFT y Consenso con UGENE"),  
        ("3-Reporte.py", "Reporte de cebadores")  
    ]
```

Es posible ejecutar uno a uno los módulos, para mayor comodidad estos están incluidos en un modulo maestro, localizado en el mismo directorio, que ejecuta secuencialmente los 3 pasos que se explican a continuación.

Secuencia requerida: ejecute en la terminal de comandos cada modulo en el orden mostrado. Para ello debe estar localizado en la misma carpeta donde se encuentran los archivos de los módulos y esto se logra con el comando “cd” como se muestra a continuación:

```
(c) Microsoft Corporation. Todos los derechos reservados.  
C:\Users\Presston>cd C:\Users\Presston\Downloads\archivo base  
C:\Users\Presston\Downloads\archivo base>
```

Una vez en el directorio donde se encuentran los módulos python, puede ejecutar uno a uno cada modulo con “python3 (nombre del archivo).py”

5. Flujo de Datos

(a)**Entrada inicial:**

- base.fasta (secuencias en formato FASTA)
- cebadores.txt (Archivos de cebadores)

(b)**Salidas generadas:**

los siguientes son los archivos generados en cada modulo

Módulo	Archivo Generado	Contenido
1	secuencias_filtradas.fa	Secuencias filtradas por fecha
2	alineamiento_recortado.fa	Resultado de alineamiento usando MAFFT con recortes de gaps en las terminales
2	consenso_levitsky.fa consenso_ugene.fa	Secuencia consenso
3	resultados_cebadores.pdf	Reporte final

6. Personalización Avanzada

A. Modificar parámetros de análisis

Editar valores en parametros.json:

En el archivo se encuentran las variables usadas tanto para la herramienta MAFFT como para ugene. De requerirse, es posible cambiar el valor --op (penalización por espacios abiertos) y el valor --ep (penalización por extensión de espacios), existe una variable vacía para añadir más parámetros (como --maxiterate) de acuerdo a las necesidades del organismo a alinear con los comandos de MAFFT siguiendo el mismo formato de los anteriores comandos usados en el archivo json. Si desea agregar más parámetros, considere el espaciado entre valores y parámetros, mas no agregue espacios extra antes ni después de agregar estos valores nuevos.

El alineamiento usa la primera regla de corte por defecto, de ser necesario puede usar una posición fija para recortar en las variables "posicion_inicio_fijo" y "posicion_fin_fijo". Considere para ello usar la interfaz gráfica de ugene. En una terminal de comandos puede ejecutar la función "ugeneui" para usar la interfaz gráfica y, abriendo el archivo de "alineamiento_MAFFT.fa" seleccionar la columna exacta del principio y del final desde donde desea que se realice el recorte antes de realizar el consenso.

También es posible aumentar o disminuir el umbral de consenso para ugene o el algoritmo de levitsky. El reporte pdf usa el consenso del algoritmo levitsky, que usa iupac, si se desea se puede reemplazar la expresión "biopython" en la sección "usar_consenso" en parámetros con "ugene" si desea usar el algoritmo para consenso.

B. Personalizar PDF

En el archivo de parámetros, puede cambiar colores reemplazando las siguientes líneas:

```
"pdf": {  
  "color_directo": "blue",  
  "color_reverso": "orange",  
  "color_sonda": "purple"  
}
```

7. Solución de Problemas Comunes

Error	Causa Probable	Solución
FileNotFoundError	Ruta incorrectas en parametros.json	Verificar existencia de archivos
KeyError en JSON	En parametros.json falta una clave	Validar estructura con el ejemplo
ModuleNotFoundError	Dependencias no instaladas	Ejecutar <code>python3 -m pip install reportlab</code>
Consenso vacío	Umbral (threshold) muy alto	Reducir valor de umbral

Error	Causa Probable	Solución
Permission Denied	El archivo pdf se encuentra abierto	Cierre la ventana que muestra el archivo pdf

Notas Clave

Orden estricto: Los módulos deben ejecutarse secuencialmente (1 → 2 → 3).

Herramientas externas: Asegurar que MAFFT y UGENE están en el PATH del sistema.

Personalización: Todas las variables se controlan desde parametros.json.

Validación: El sistema verifica automáticamente la estructura del JSON.

Recomendación: modificar periodo y mes en parametros.json antes de cada ejecución de los módulos para análisis periódicos