

# Algoritmos para Análise de Sequências Biológicas

## Ficha 1

# Objetivos

- Introduzir a shell
- Introduzir a CLI

# Links para Tutoriais

- 1 shell
- 2 CLI

# Instalação

- Instalar o Python3
- Instalar um IDE (por exemplo IDLE)
- Opcional: instalar o virtualbox e uma imagem linux, veja por exemplo este [tutorial](#)

# Instruções para instalar o exercício

## Windows

- 1 Fazer fork do repositório carregando no link [aqui](#)
- 2 Correr o programa clicando no botão *Run*
- 3 Passar para o tab *Shell*
- 4 Correr comandos da Shell para responder às perguntas

## Linux

- 1 Vá buscar o ficheiro python do link dado acima
- 2 Correr o programa no python: `python3 main.py`
- 3 Abrir um terminal e correr os comandos da Shell para responder às perguntas

# Comandos Shell

## Alguns Commandos

`nano/gedit` editores de texto  
`ls` listar  
`wc` contar linhas  
`find` procurar ficheiros  
`sort` ordenar linhas  
`uniq` remover/contar duplicados  
`grep` filtrar linhas  
`cut` cortar linhas por tamanhos ou separadores  
  
`cp/mv/rm`  
copiar/mover/remover

## Pipeline

`cmd1 | cmd2 | ... | cmdn`

- Comandos executados em paralelo
- A saída de um comando é a entrada do seguinte

# Exemplo de comandos Shell

Listar ficheiros

```
ls
```

Listar ficheiros recursivamente

```
ls -R
```

Contar ficheiros

```
ls | wc
```

Listas ficheiros que contenham  
acg

```
grep acg *
```

Segundo campo de uma lista

```
echo "1,2,3" | cut -f2 -d,
```

Listar ficheiros que contenham  
seq no nome

```
ls *seq*
```

```
ls | grep seq
```

Listar ficheiros com extensão txt

```
ls *.txt
```

```
ls | grep -e txt$
```

# Exercício de Shell

- ❶ Liste o conteúdo da pasta atual
- ❷ Liste recursivamente o conteúdo da pasta atual
- ❸ Conte o número de ficheiros com a extensão *fasta* na pasta *rui*
- ❹ Conte o número de ficheiros com a extensão *fasta* na pasta *rui* cujo cabeçalho contenha a palavra *dog*
- ❺ Liste só o nome da espécie dos ficheiros *fasta* da pasta *rui*
- ❻ Conte o número de ficheiros de cada extensão na pasta *rui*
- ❼ Conte o número de ficheiros de cada extensão em todas as pastas e subpastas
- ❽ Sabendo que a primeira palavra de cada cabeçalho é a espécie, conte o número de ficheiros de cada espécie
  - ❶ Na pasta *rui*
  - ❷ Em todas as pastas



## Exemplo de CLI em Python

```
>>> from math import pi
>>> raio = 3
>>> 2 * pi * raio
18.84955592153876
>>> per = lambda r: 2 * pi * r
>>> per(7)
43.982297150257104
>>> f"0 perímetro de um circulo de raio {raio} é {per(raio)}"
'0 perímetro de um circulo de raio 3 é 18.84955592153876'
>>> seq = "acgt" + "tga"
>>> seq * 3
'acgttgaacgttgaacgttga'
```

# Exercícios de CLI em Python

- 1 Escreva uma função que calcula a área de um círculo dado o seu raio
- 2 Usando a função *input*, coloque numa variável um nome que foi pedido ao utilizador
- 3 Usando a função *print*, imprima Olá seguido desse nome
- 4 Crie uma função chamada *cumprimentar* e experimente-a
- 5 Usando a função *len*, descubra o tamanho do nome do utilizador
- 6 Faça um programa que peça dois números e os some, vai precisar da função *int* para os converter para inteiros
- 7 Faça um programa que peça o diâmetro de um círculo e calcule o seu perímetro e a sua área
- 8 Mude a função *cumprimentar* para também imprimir o tamanho do nome

## Exemplos avançados de scripts Shell

Contar o número de linhas dos ficheiros seq10 a seq14

```
wc -l seq1[0-4].fasta
```

Criar ficheiros de texto (apagar cabeçalho fasta)

```
for fich in *.fasta
do
    basename=${fich%%.fasta}
    sed -e '/>/d' $fich > $basename.txt
done
```

Outra forma

```
for fich in {1..50}
do
    num=`printf %02d $fich`
    sed -e '/>/d' seq$num.fasta > seq$num.txt
done
```