

# Cómo usar Python para curar enfermedades genéticas

Marina Moro López

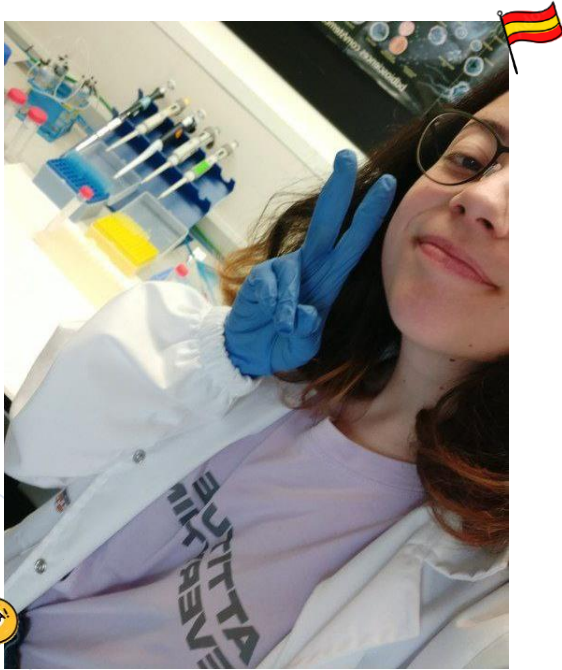
PyCon Panamá 2024



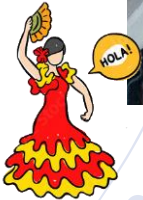
Python  
Panamá



# ¡Hola! :D



- Ingeniera biomédica
- Futura doctora en biofísica y bioingeniería
- 'Programadora' en mi día a día científico
- Secretaria de Python España



# Índice

**01**

## **INTRODUCCIÓN**

¿Qué es la enfermedad de Huntington?

**02**

## **GENÉTICA BÁSICA**

Teoría para entender el caso práctico

**03**

## **CASO PRÁCTICO**

Tratamiento con CRISPR y Python

**?**

## **RONDA DE PREGUNTAS**

01

# INTRODUCCIÓN

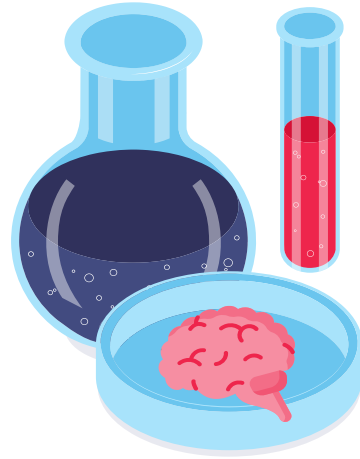
¿Qué es la enfermedad de Huntington?



# ¿Qué es la enfermedad de Huntington?

Enfermedad rara genética  
neurodegenerativa hereditaria

Producida por una mutación en  
el gen de la proteína Huntingtina

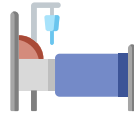


**De 5 a 10 afectadxs  
cada 100.000 habitantes**

# Síntomas y tratamientos



Movimientos involuntarios, dificultad en el habla, pérdida de memoria, demencia, depresión y suicidio



Dependencia completa en los estadios más avanzados de la enfermedad

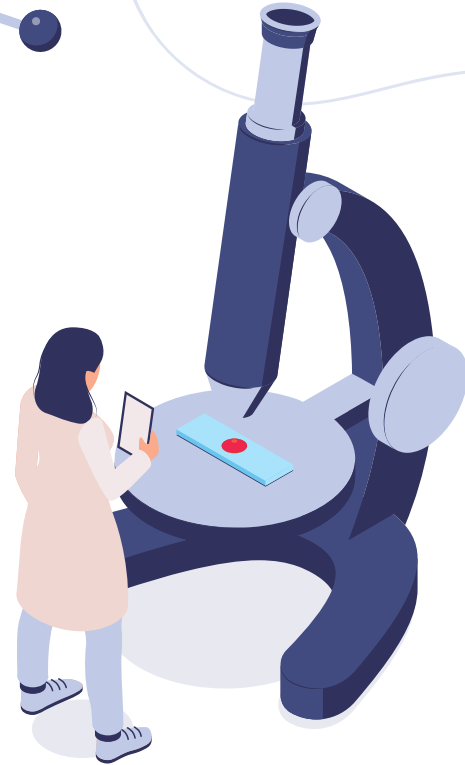
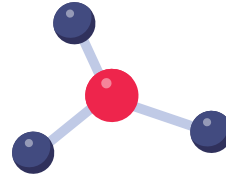


Tratamientos paliativos  
Sin cura actualmente

02

# GENÉTICA BÁSICA

Teoría para entender el caso práctico

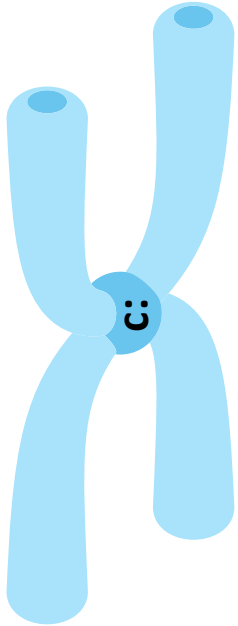




# Cromosoma - Gen - ADN

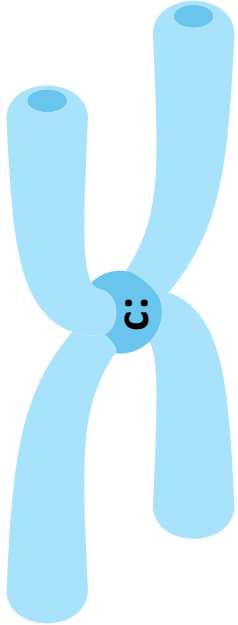
# Cromosoma - Gen - ADN

Estructura que contiene  
todos los genes



# Cromosoma - Gen - ADN

Estructura que contiene  
todos los genes

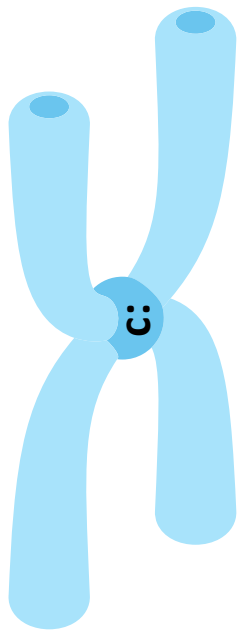


Segmento de ADN que  
determina un rasgo

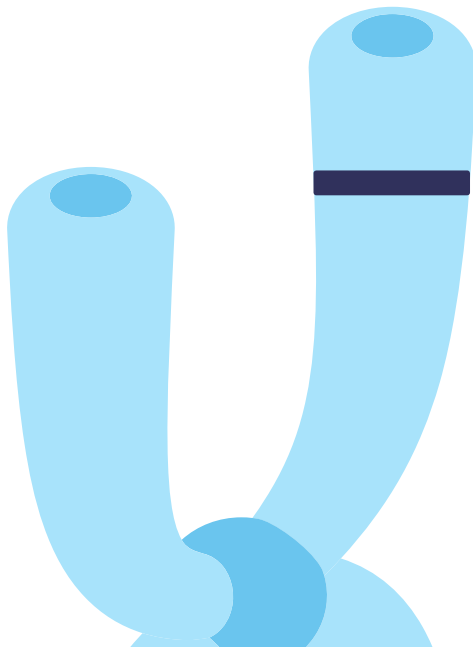


# Cromosoma - Gen - ADN

Estructura que contiene  
todos los genes



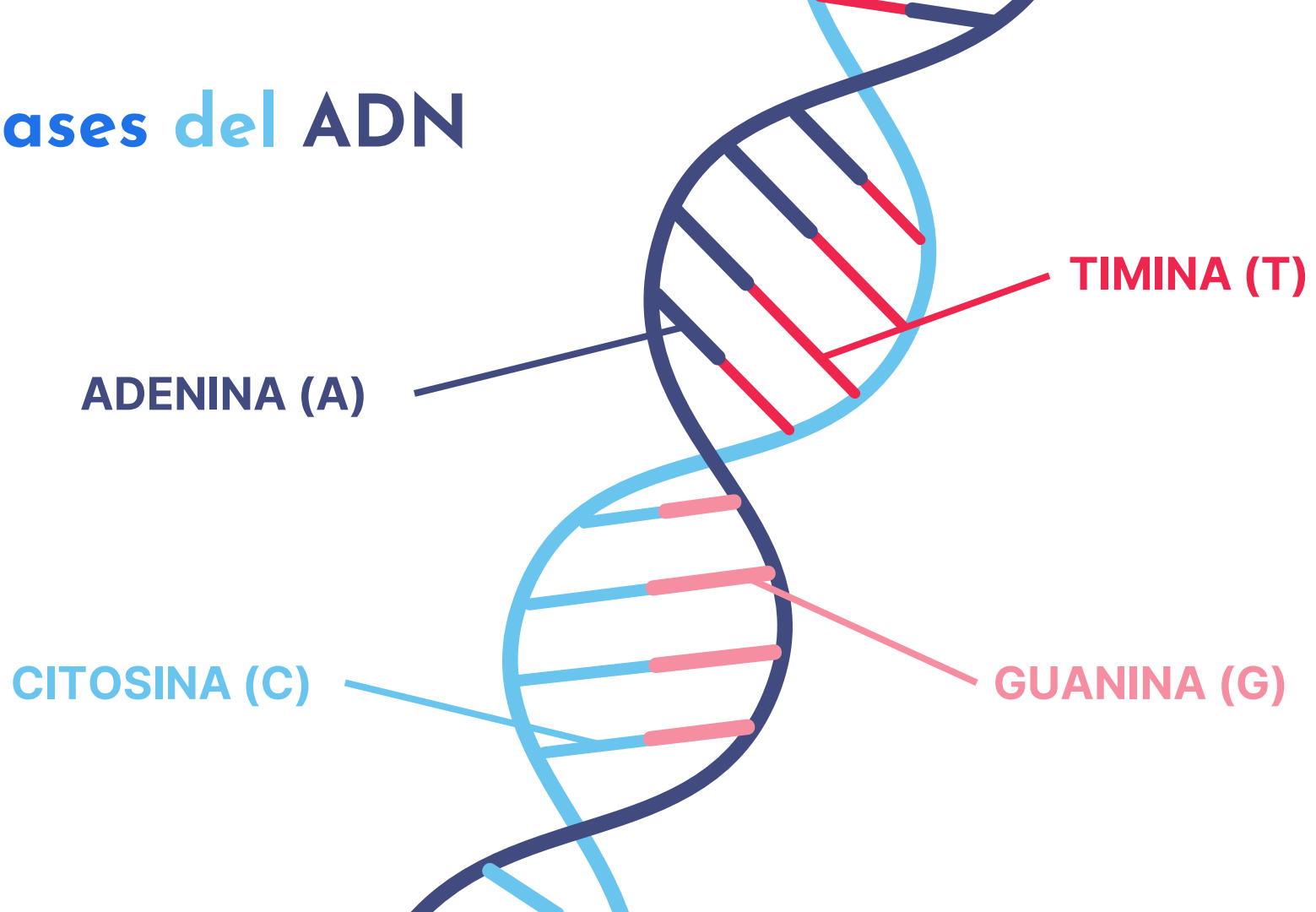
Segmento de ADN que  
determina un rasgo



Doble hélice  
formada por bases



# Bases del ADN

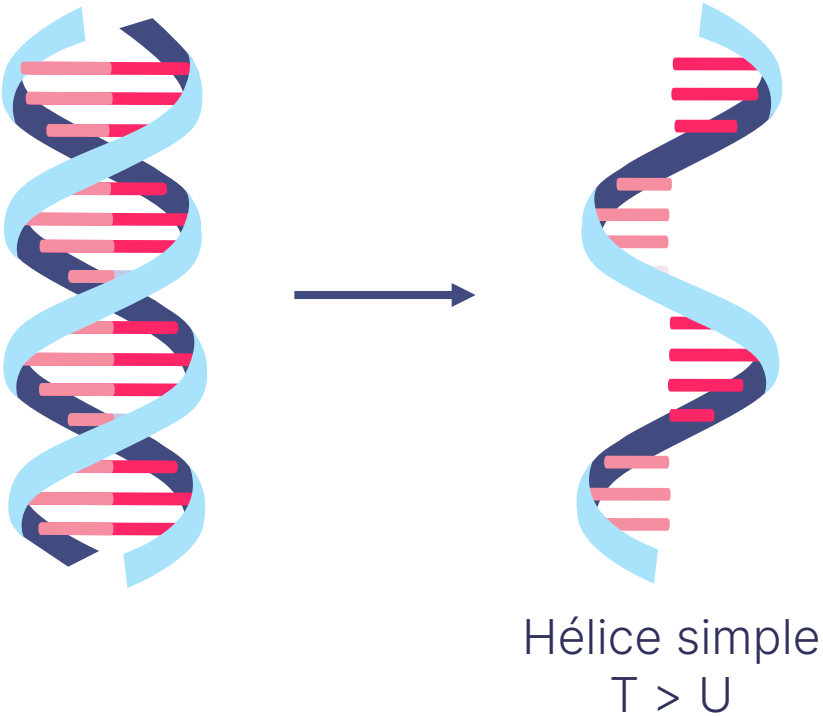


# ADN - ARN - Proteína

# ADN - ARN - Proteína

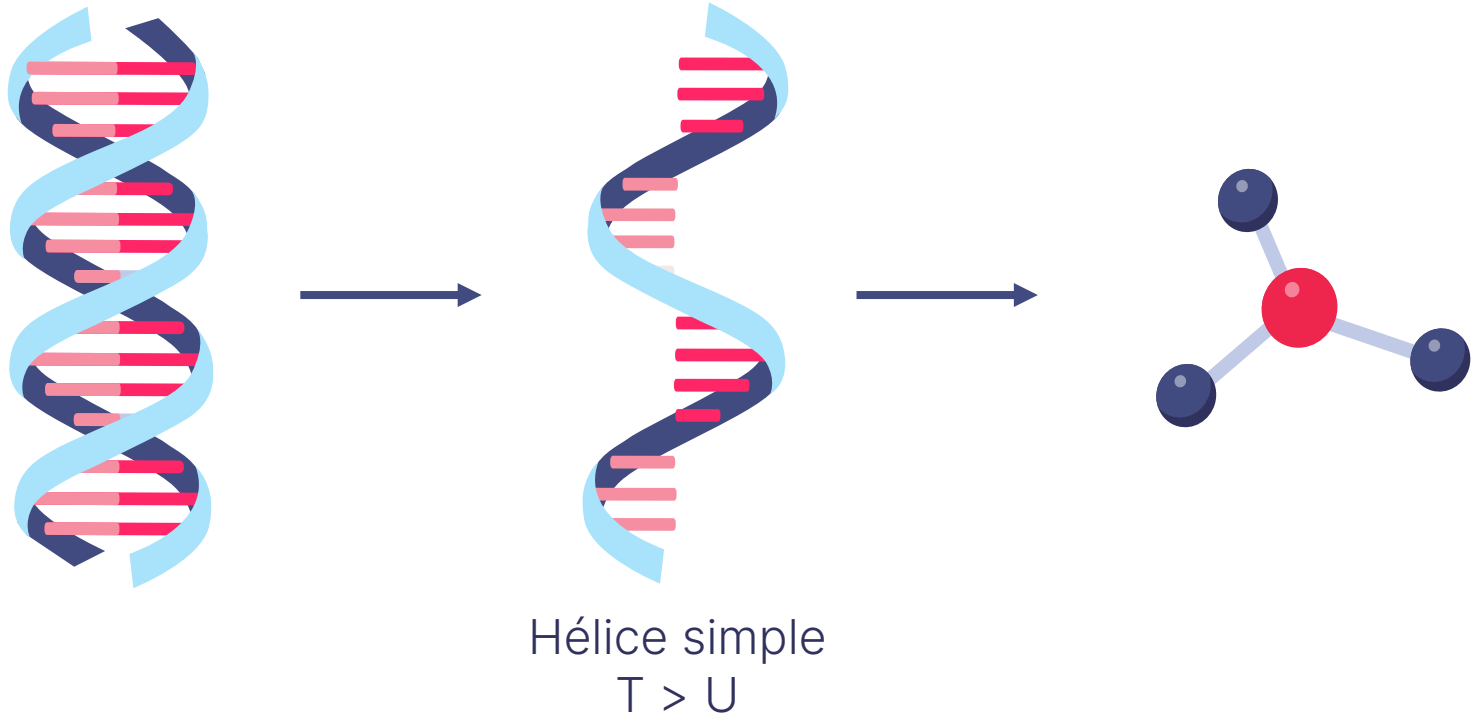


# ADN - ARN - Proteína





# ADN - ARN - Proteína

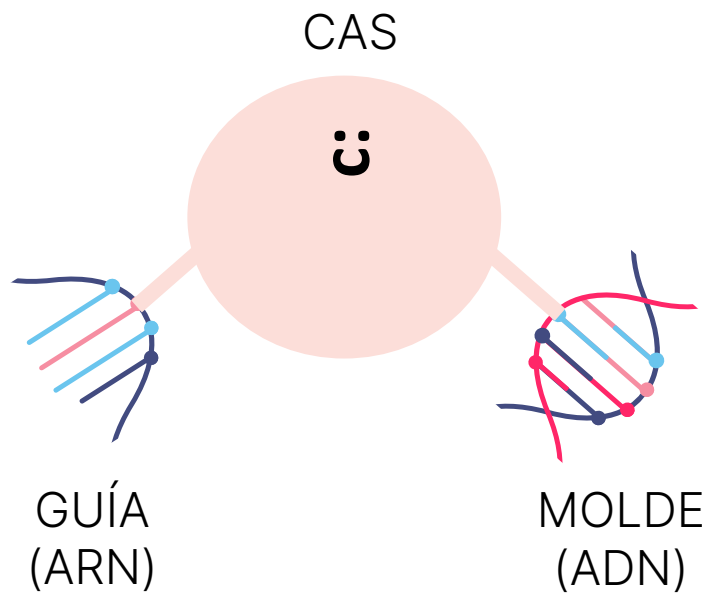




# CRISPR

Corta y pega de secuencias  
de ADN (edición genética)

# CRISPR



# CRISPR

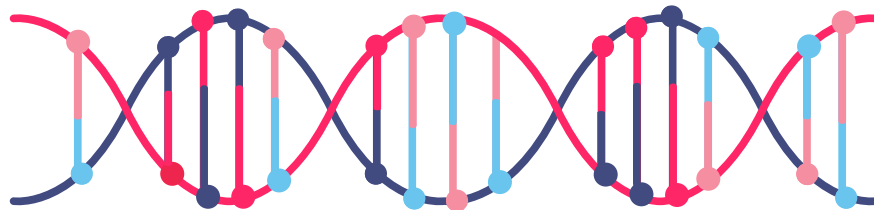
CAS

ü

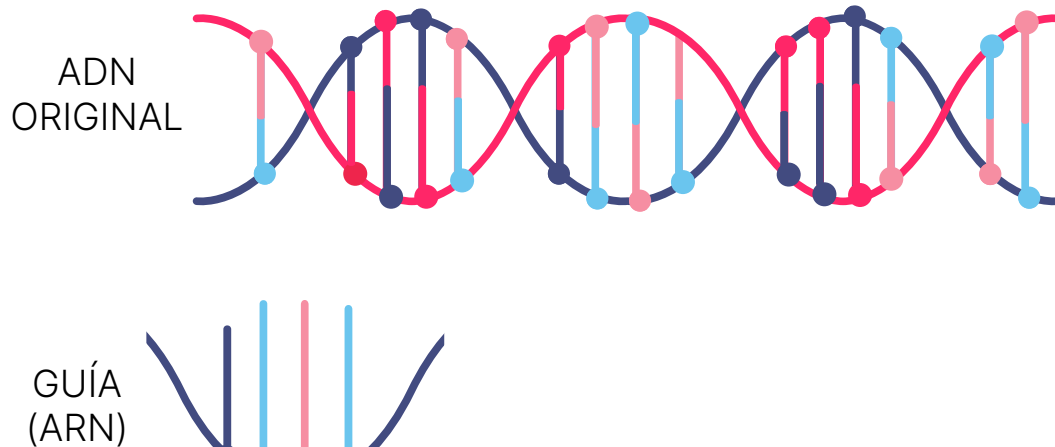
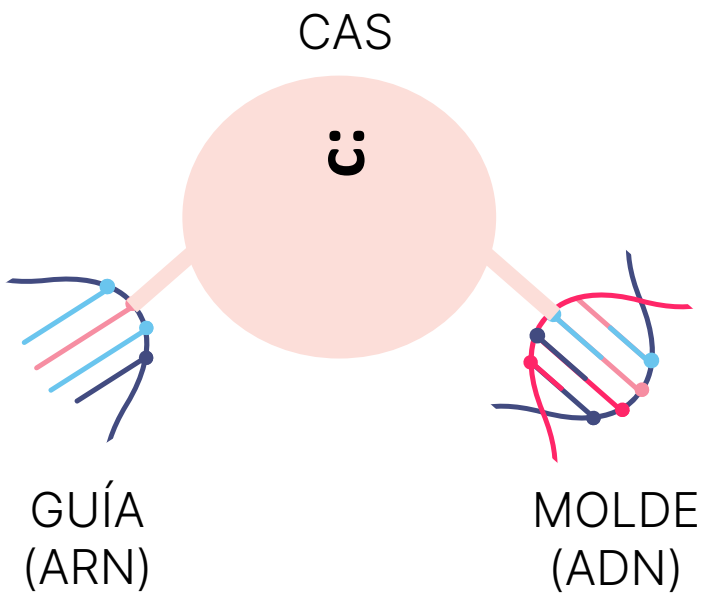
GUÍA  
(ARN)

MOLDE  
(ADN)

ADN  
ORIGINAL



# CRISPR



# CRISPR

CAS

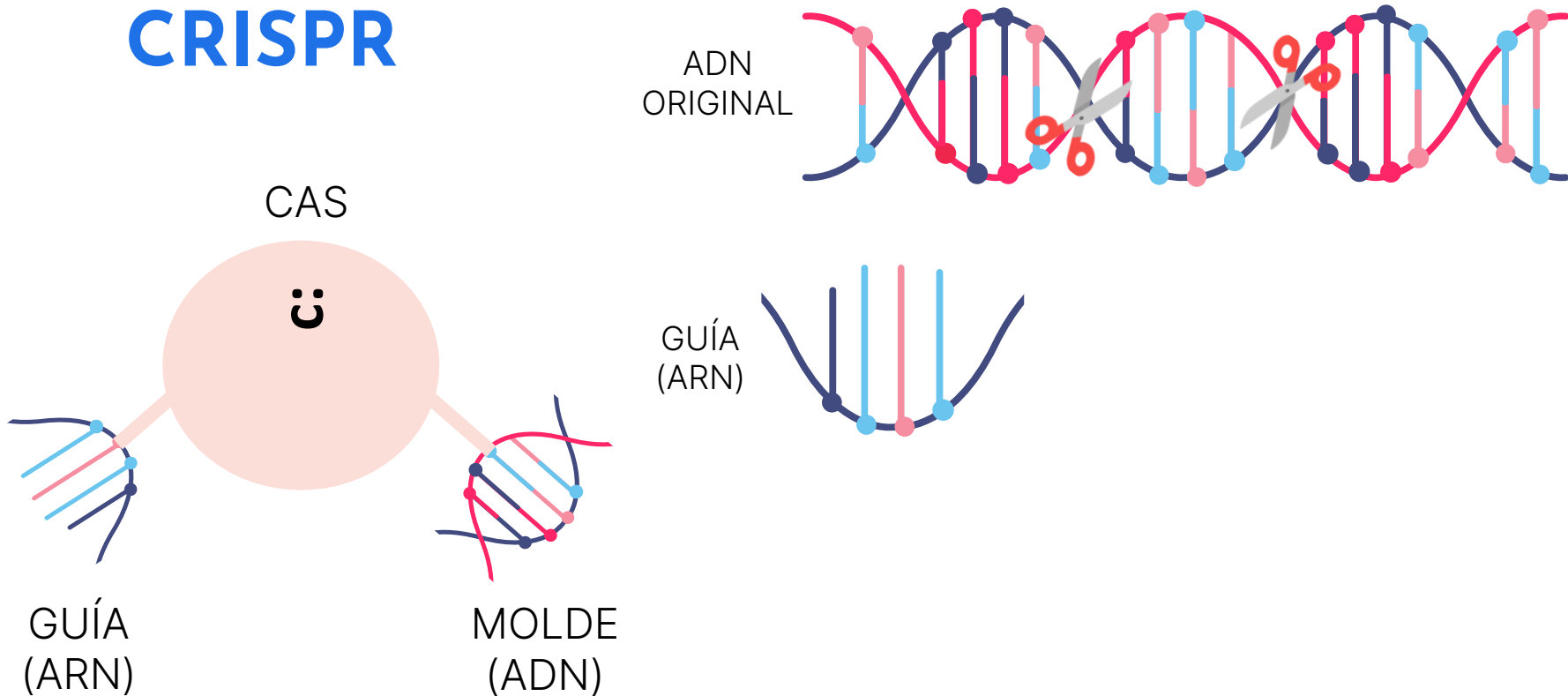
ü

GUÍA  
(ARN)

MOLDE  
(ADN)

ADN  
ORIGINAL

GUÍA  
(ARN)



# CRISPR

CAS

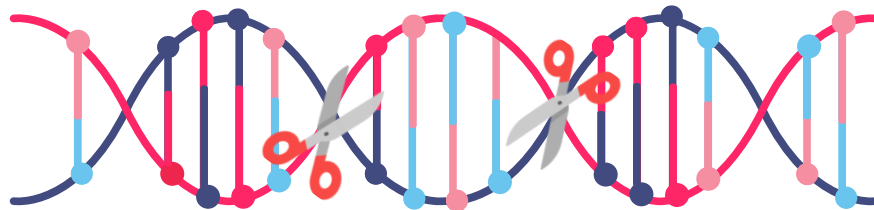
ü



GUÍA  
(ARN)

MOLDE  
(ADN)

ADN  
ORIGINAL



GUÍA  
(ARN)



MOLDE  
(ADN)



# CRISPR

CAS



GUÍA  
(ARN)

MOLDE  
(ADN)

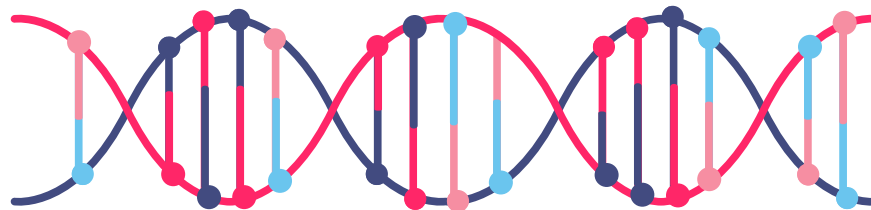
ADN  
ORIGINAL



GUÍA  
(ARN)



MOLDE  
(ADN)





03

# CASO PRÁCTICO

Tratamiento con CRISPR y Python



El error es una repetición del triplete C-A-G  
desde la posición 5197 del gen HTT

Sin Huntington de 8 a 34 repeticiones, con Huntington más de 40

**Estrategia**: delección de las repeticiones C-A-G sobrantes

El error es una repetición del triplete C-A-G  
desde la posición 5197 del gen HTT

Sin Huntington de 8 a 34 repeticiones, con Huntington más de 40

**Estrategia**: delección de las repeticiones C-A-G sobrantes

GEN HTT  
DEFECTUOSO

...GTCCCTCAAGTCCTTC CAGCAGCAGCAGCAGCAGCAG...

> 40 repeticiones



Posición 5197

GEN HTT  
CORREGIDO

...GTCCCTCAAGTCCTTC CAGCAGCAG...

8-34 repeticiones

El error es una repetición del triplete C-A-G  
desde la posición 5197 del gen HTT

Sin Huntington de 8 a 34 repeticiones, con Huntington más de 40

**Estrategia**: delección de las repeticiones C-A-G sobrantes

GEN HTT  
DEFECTUOSO

...GTCCCTCAAGTCCTTC CAGCAGCAGCAGCAGCAG...

> 40 repeticiones



Posición 5197

GEN HTT  
CORREGIDO

...GTCCCTCAAGTCCTTC CAGCAGCAG...

8-34 repeticiones





marinamorolopez / huntington-pyconPA2024

```
from tkinter.filedialog import askopenfile

def main():

    gene_file = askopenfile(mode='r')
    gene_seq = gene_file.readlines()[1:]
    gene_seq = ''.join(gene_seq).replace('\n', '')

    DNA_guide, mutated_gene_seq, mold = repeated_seq(gene_seq)
```

Nucleotide

Search

Create alert Advanced

Species

- Animals (3,302)
- Plants (538)
- Fungi (19)
- Bacteria (11)
- Viruses (1)
- Customize ...

Molecule types

- genomic DNA/RNA (1,957)
- mRNA (2,052)
- Customize ...

Source databases

- INSDC (GenBank) (6,092)
- RefSeq (2,585)
- Customize ...

Sequence Type

- Nucleotide (8,139)
- EST (80)
- GSS (462)

Genetic

Summary ▾ 20 per page ▾ Sort by Default order ▾

Send to: ▾ **Filters:** [Manage Filters](#)

GENE

Was this helpful?



[HTT orthologs from vertebrates](#)

huntingtin

[How are orthologs calculated?](#)

[Genes with similar protein architectures](#)

[How are similar genes calculated?](#)

Results by taxon

Top Organisms [Tree](#)

- synthetic construct (4555)
- Homo sapiens (244)
- Lens culinaris subsp. culinaris (162)
- Lupinus albus (150)
- Lupinus angustifolius (150)
- All other taxa (3420)

More...

Find related data

Database:

Find items

Search details

HTT[All Fields]

Items: 1 to 20 of 8681

HTT - Nucleotide - NCBI

+

ncbi.nlm.nih.gov/nuccore/?term=HTT

🔍

🔗

☆

🔴

🧩

⬇️

🖨️

5.

18,091 bp linear DNA

Accession: VXIV02000467.1 GI: 1883872130

[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)

[GenBank](#) [FASTA](#) [Graphics](#)

☐

[Drosophila busckii chromosome 3R sequence](#)

6.

26,871,514 bp linear DNA

Accession: CP012526.1 GI: 924560592

[Assembly](#) [BioProject](#) [BioSample](#) [Protein](#) [Taxonomy](#)

[GenBank](#) [FASTA](#) [Graphics](#)

☐

[Homo sapiens huntingtin \(HTT\), RefSeqGene \(LRG\\_763\) on chromosome 4](#)

7.

176,286 bp linear DNA

Accession: NG\_009378.1 GI: 221136754

[Protein](#) [PubMed](#) [Taxonomy](#)

[GenBank](#) [FASTA](#) [Graphics](#)

☐

[Rattus norvegicus huntingtin \(Htt\), mRNA](#)

8.

13,191 bp linear mRNA

Accession: NM\_024357.4 GI: 1982559709

[Protein](#) [PubMed](#) [Taxonomy](#)

[GenBank](#) [FASTA](#) [Graphics](#)

☐

[Homo sapiens clone HTT iso-A fetal-F12 huntingtin isoform A \(HTT\) mRNA, complete cds, alternatively spliced](#)

9.

1,778 bp linear mRNA

Accession: KJ535072.1 GI: 608785749

[Protein](#) [PubMed](#) [Taxonomy](#)

[GenBank](#) [FASTA](#) [Graphics](#)

https://www.ncbi.nlm.nih.gov/nuccore/NG\_009378.1

Homo sapiens huntingtin (HTT)

ncbi.nlm.nih.gov/nuccore/NG\_009378.1

An official website of the United States government

NIH

National Library of Medicine

National Center for Biotechnology Information

Log in

Nucleotide

Nucleotide

HTT

Advanced

Search

GenBank

Homo sapiens huntingtin (HTT), RefSeqGene (LRG\_763) on chr

NCBI Reference Sequence: NG\_009378.1

FASTA

Graphics

Go to:

LOCUS

NG\_009378

176286 bp

DNA

linear

PRI 20-NOV-2023

DEFINITION

Homo sapiens huntingtin (HTT), RefSeqGene (LRG\_763) on chromosome 4.

ACCESSION

NG\_009378

VERSION

NG\_009378.1

KEYWORDS

RefSeq; RefSeqGene.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 176286)

AUTHORS

Lee J, Park EH, Couture G, Harvey I, Garneau P and Pelletier J.

TITLE

An upstream open reading frame impedes translation of the

Send to:

Complete Record

Coding Sequences

Gene Features

Choose Destination

File

Collections

Clipboard

Analysis Tool

Download 1 item.

Format

FASTA

Show GI

Create File

the HTT gene

Aggregation of Mutant Huntingtin and Resc [ACS Chem Neurosci. 2023]

The HSPB1-p62/SQSTM1 functional complex regulates the unconventional [Hum Mol Genet. 2023]



```
mutated_gene_file = open('MUTATED_SEQUENCE.txt', 'w')
mutated_gene_file.write(mutated_gene_seq)
mutated_gene_file.close()

guide_file = open('GUIDE.txt', 'w')
guide_file.write(DNA_to_RNA(DNA_guide))
guide_file.close()

mold_file = open('MOLD.txt', 'w')
mold_file.write(mold)
mold_file.close()
```

```
def DNA_to_RNA(DNA_guide):

    RNA_guide = ""
    for base in DNA_guide:
        if base == "T":
            RNA_guide += "A"
        elif base == "A":
            RNA_guide += "U"
        elif base == "C":
            RNA_guide += "G"
        elif base == "G":
            RNA_guide += "C"

    return RNA_guide
```

```

def repeated_seq(gene_seq):

    mutation_position = int(input("Introduce the numeric position of the mutation base (e.g. 1, 25, 203): "))
    while mutation_position <= 0:
        print('Invalid input. Introduce positive number. ')
        mutation_position = int(input("Introduce the numeric position of the mutation base (e.g. 1, 25, 203): "))

    rep_letters = input("Introduce the letters that are repeated (e.g. AAT, CAG, CCGT, GACTA): ")

    healthy_reps = int(input("Introduce the healthy number of repetitions (e.g. 20, 35, 42): "))
    while healthy_reps <= 0:
        print('Invalid input. Introduce positive number. ')
        healthy_reps = int(input("Introduce the healthy number of repetitions (e.g. 20, 35, 42): "))

    patient_reps = []
    gene_seq_slice = gene_seq[mutation_position-1:]
    i = 0
    while gene_seq_slice.find(rep_letters) != -1:
        patient_reps.append(gene_seq_slice.find(rep_letters))
        if i >= 2:
            if patient_reps[-1] - patient_reps[-2] != len(rep_letters):
                break
        gene_seq_slice = gene_seq_slice.replace(rep_letters, "*" * len(rep_letters), 1)
        i += 1

    DNA_guide = rep_letters * (len(patient_reps)-1)
    mold = rep_letters * healthy_reps
    mutated_gene_seq = gene_seq[:mutation_position-1] + mold + gene_seq[mutation_position+(len(patient_reps)-1)*len(rep_letters)-1:]

    return DNA_guide, mutated_gene_seq, mold

```

Sin Huntington de 8 a 34 repeticiones, con Huntington más de 40  
**Estrategia**: delección de las repeticiones C-A-G sobrantes

[illegible]

GUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUC  
GUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUCGUC|

Sin Huntington de 8 a 34 repeticiones, con Huntington más de 40  
**Estrategia**: delección de las repeticiones C-A-G sobrantes

[illegible]

CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG

Sin Huntington de 8 a 34 repeticiones, con Huntington más de 40  
**Estrategia**: delección de las repeticiones C-A-G sobrantes

[illegible][illegible]

# Timeline del tratamiento



01

## Obtención de datos

Secuenciación  
genética del  
paciente



02

## Python

Automatiza el  
diseño de CRISPR



03

## Síntesis

Producción  
bioquímica del  
sistema con guía  
y molde



04

## Inyección

Intracraneal o  
intravenosa

# Más aplicaciones **terapéuticas**



Células CAR-T contra el cáncer



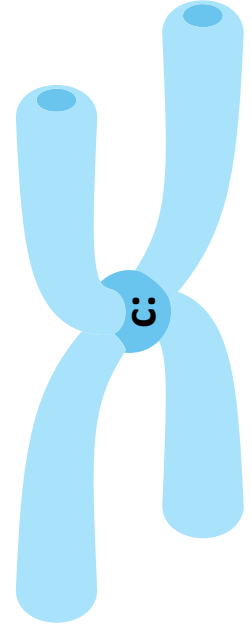
Terapia antiviral (SARS-CoV-2, VIH)



Lucha contra enfermedades infecciosas  
(malaria, fiebre amarilla)

# Bibliografía de interés

- Alkanli, S. S. et al. *CRISPR/Cas9 Mediated Therapeutic Approach in Huntington's Disease*. Mol Neurobiol 60(3), 1486-1498 (2023).
- Wan Shin, J. et al. *Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9*. Hum Mol Gene 25(20), 4566-4576 (2016).
- Seo, J.H. et al. *DNA double-strand break-free CRISPR interference delays Huntington's disease progression in mice*. Commun Biol 6, 466 (2023).
- Yan, S. et al. *Cas9-mediated replacement of expanded CAG repeats in a pig model of Huntington's disease*. Nat. Biomed. Eng 7, 629-646 (2023).





# ¡Muchas gracias!

## ¿Preguntas?



@marinamorolopez



Marina Moro López



marinamorolopez



huntington-pyconPA2024

Public



PyCon Panamá 2024



Python  
Panamá



Python  
España