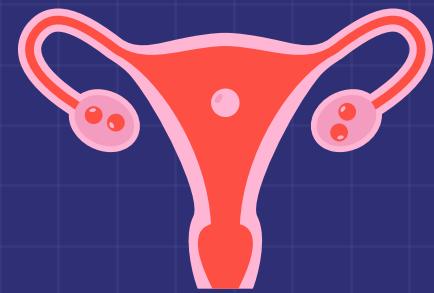
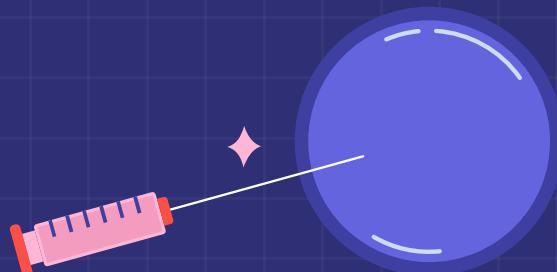


Prenatal diagnosis of genetic diseases using Python

Helena Gómez Pozo
Marina Moro López

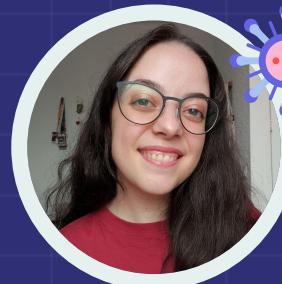


About us



Helena Gómez Pozo

Health biologist QA&RA in Pharmaceutical and Biotechnology Industry
PyConES 2024 organizer



Marina Moro López

Biomedical Engineer
Predoctoral researcher in Biophysics and Bioengineering
Python Spain Association Secretary



Table of contents

01

Basic
concepts

02

What is prenatal
diagnosis?

03

Practical case

04

Regulation and
commercialization

05

Clinical
applications

?

Q&A

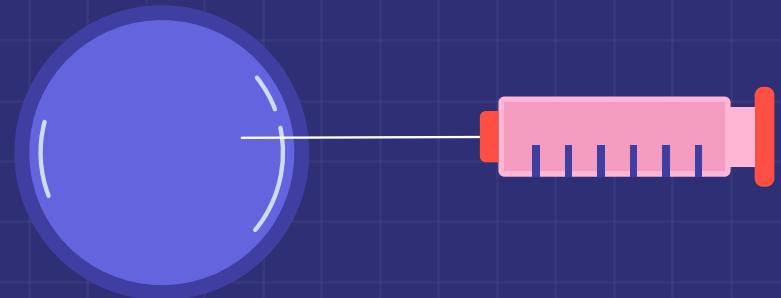




01

Basic concepts

Genetics and sequencing technologies



Basic concepts

Elements of genetics



Basic concepts

Elements of genetics



Genetics

Area of biology that seeks to understand and explain how biological inheritance is transmitted from generation to generation



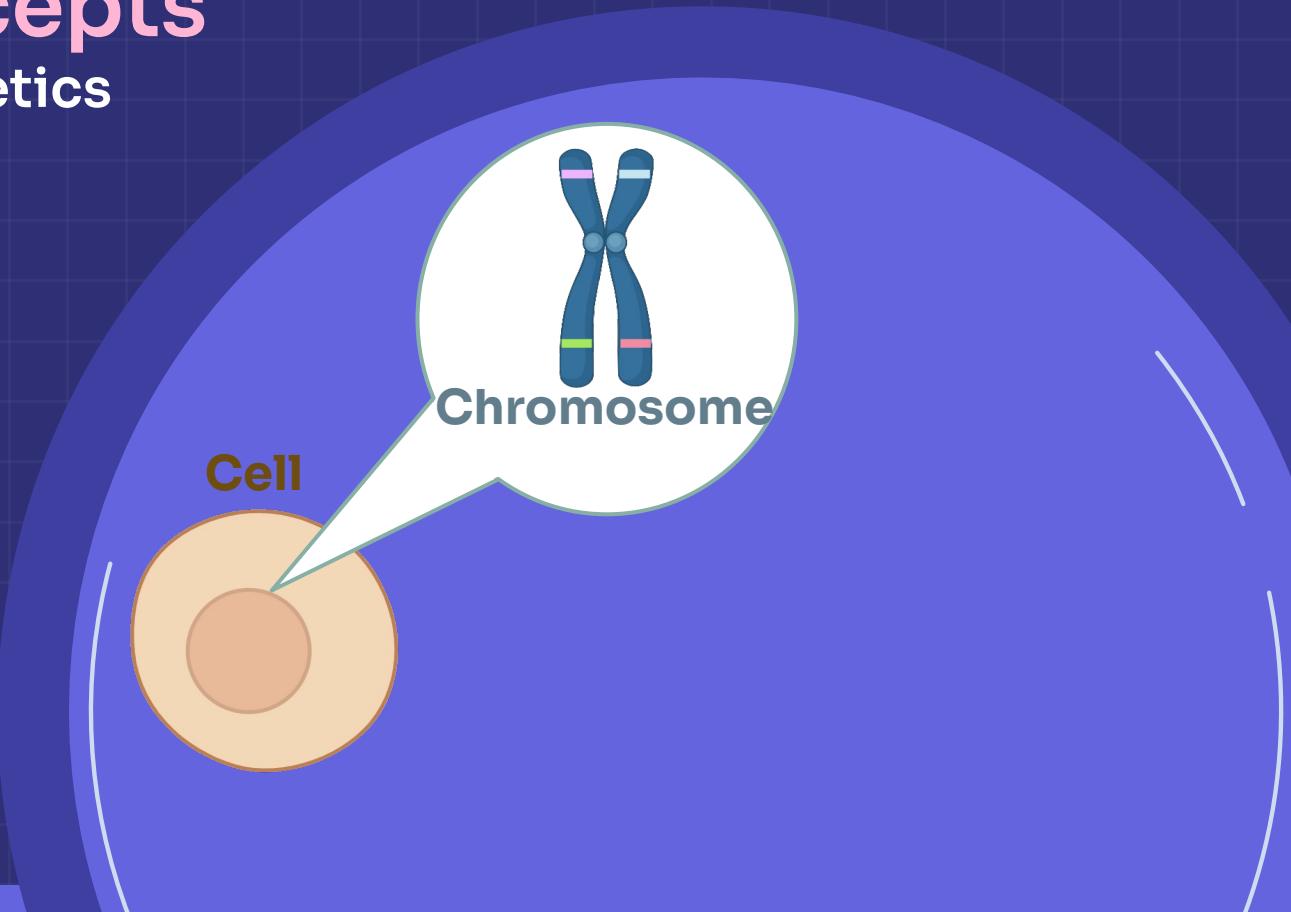
Basic concepts

Elements of genetics



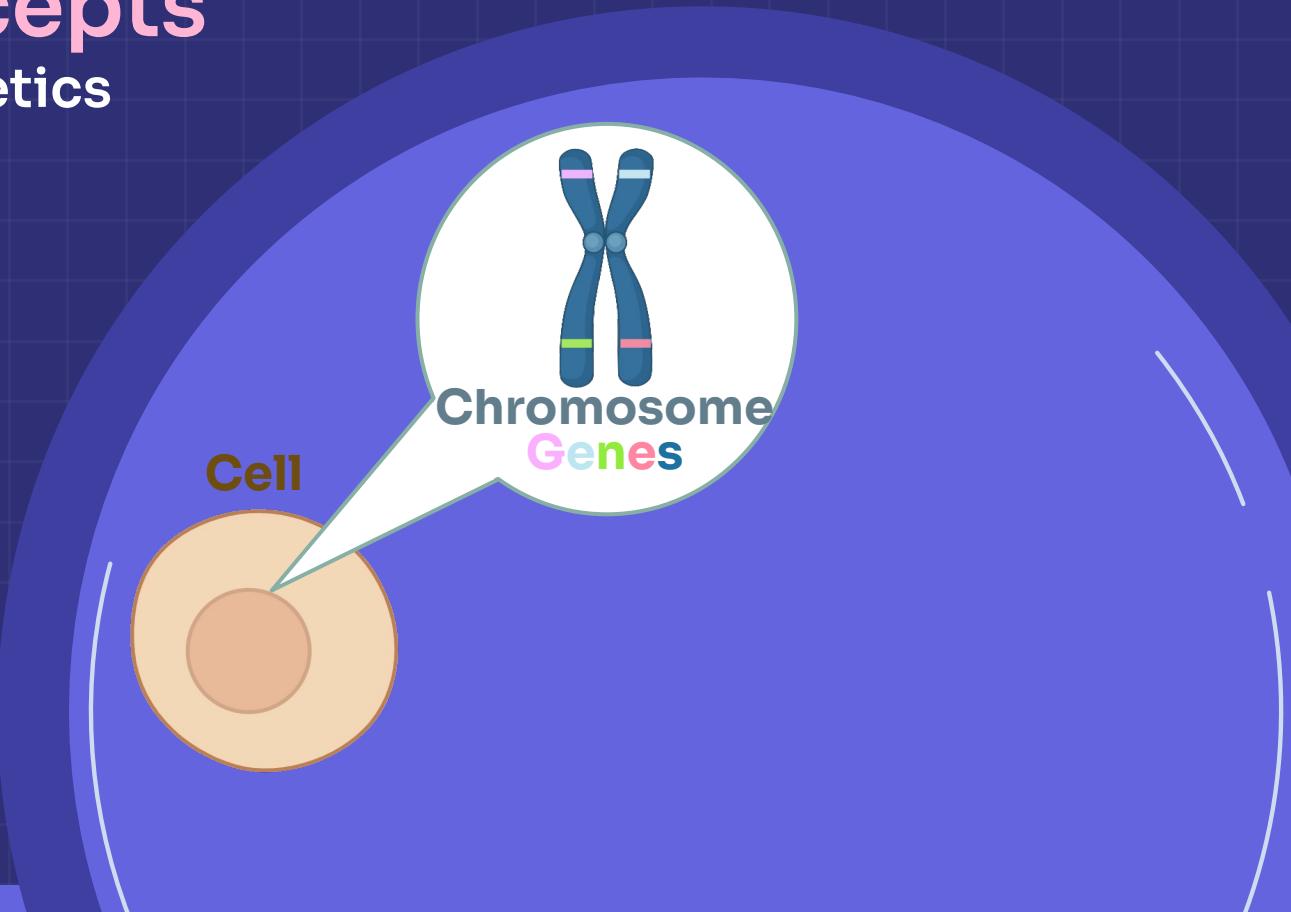
Basic concepts

Elements of genetics



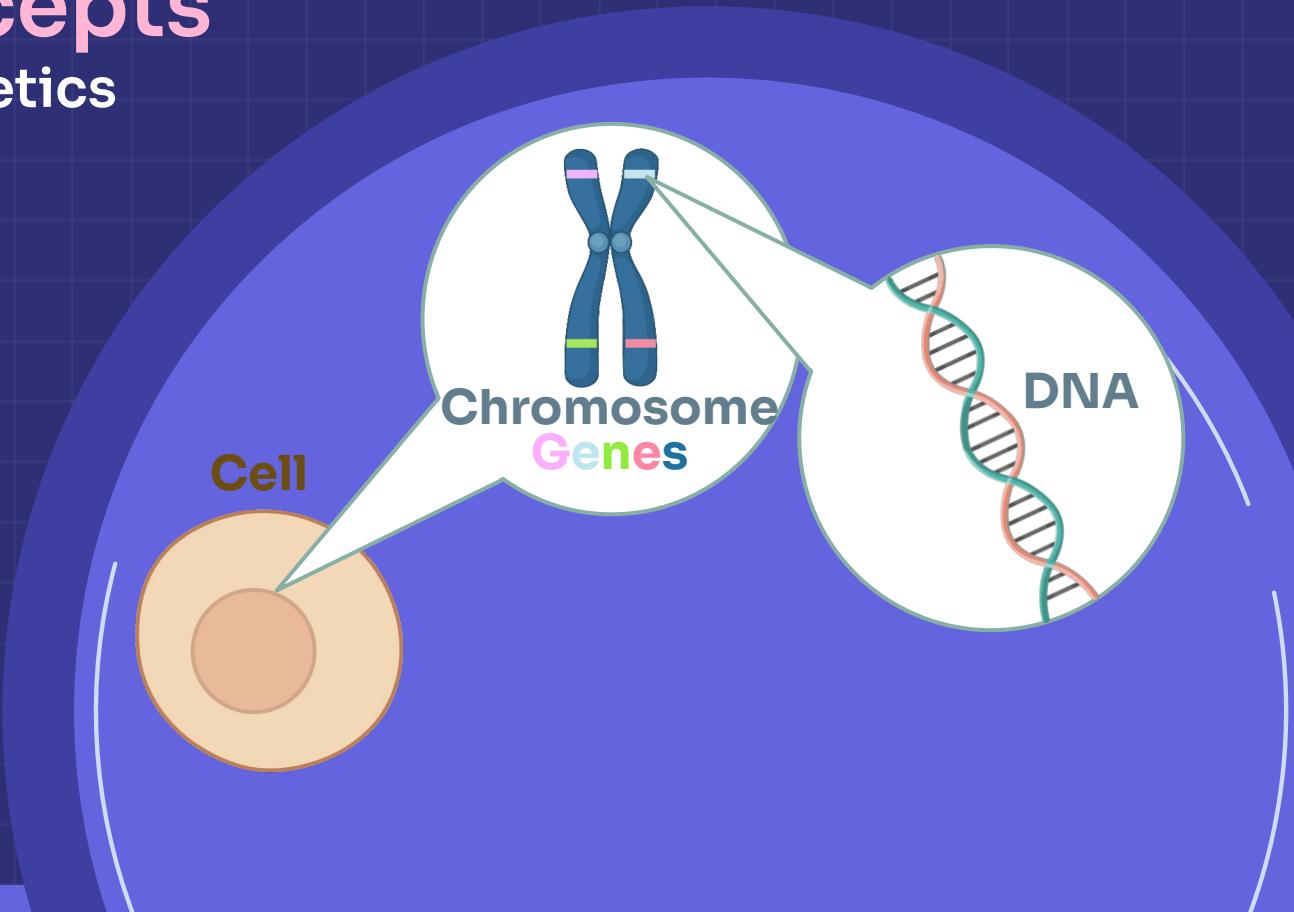
Basic concepts

Elements of genetics



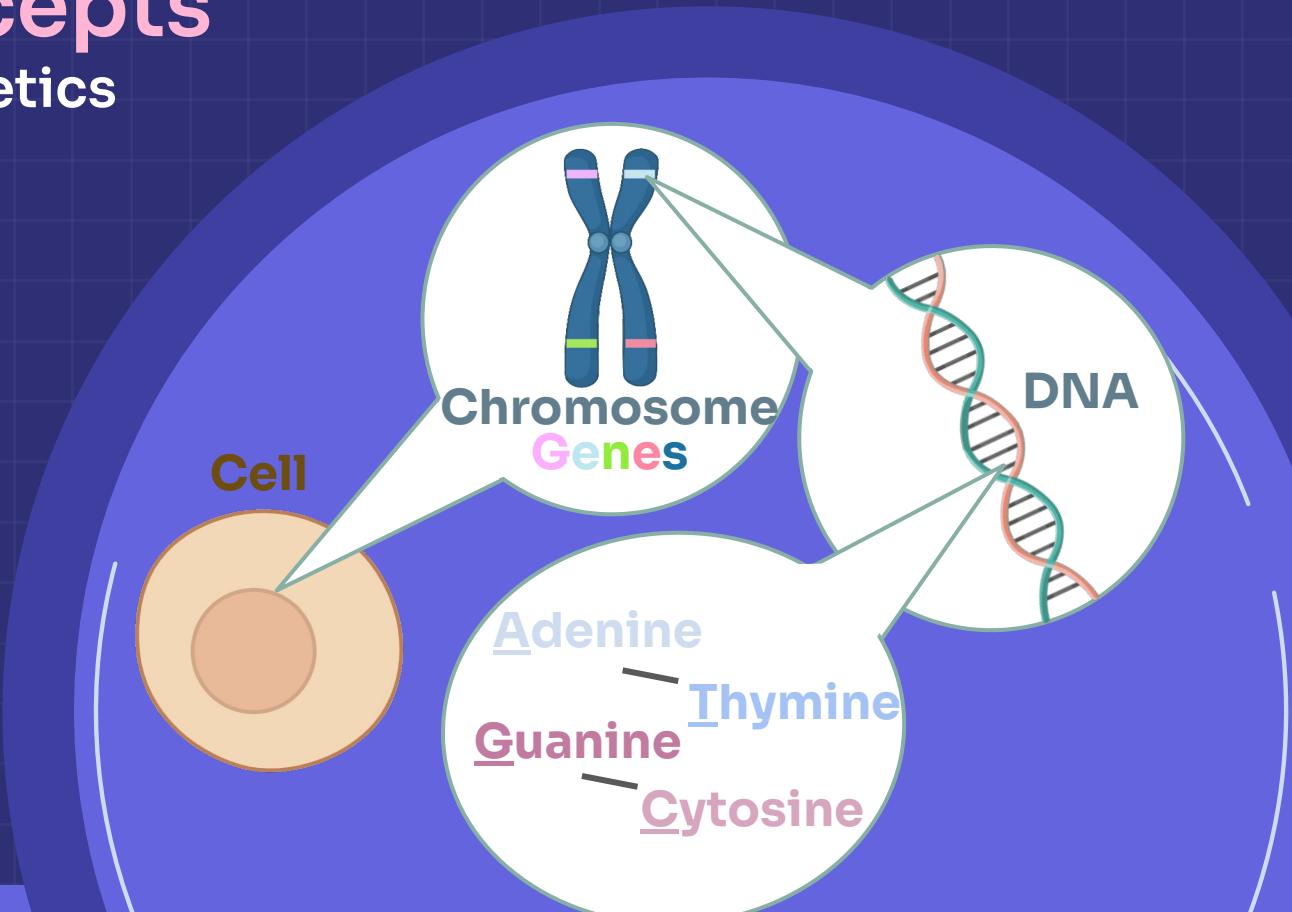
Basic concepts

Elements of genetics



Basic concepts

Elements of genetics





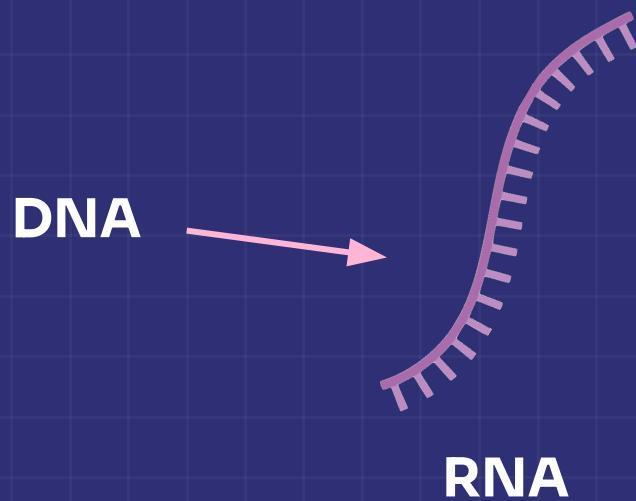
Basic concepts

Elements of genetics

DNA

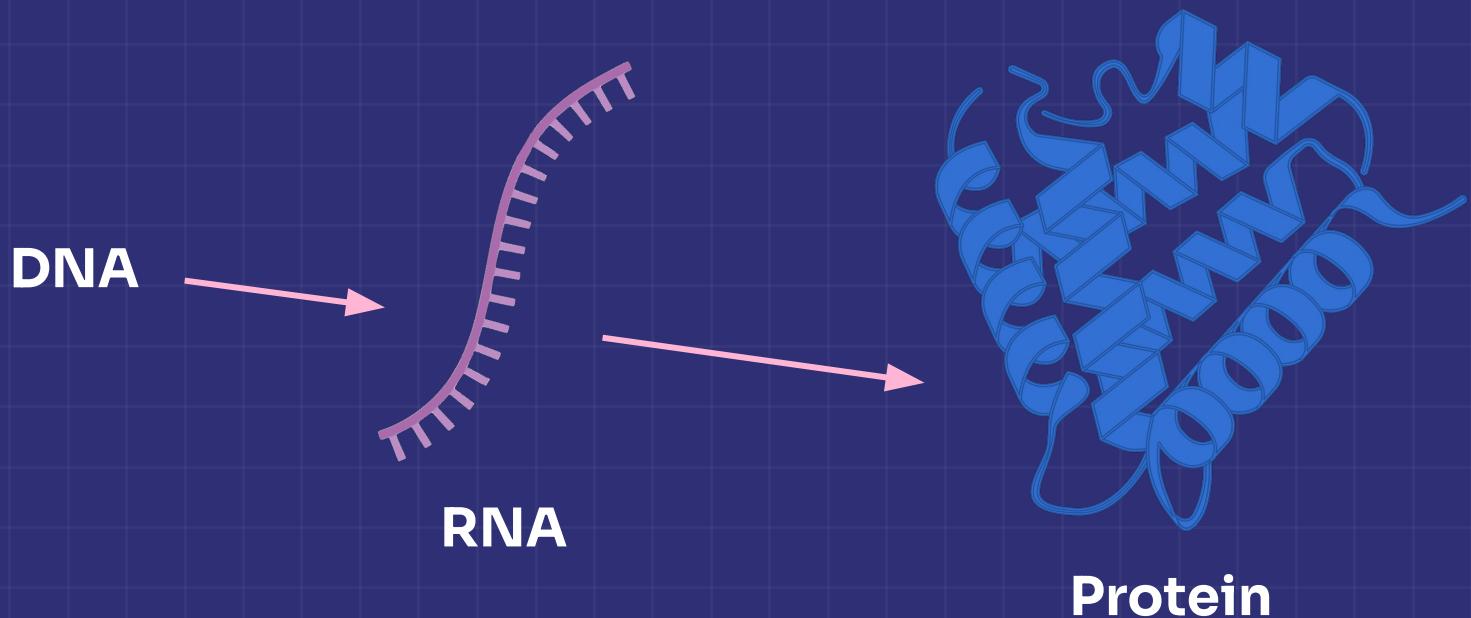
Basic concepts

Elements of genetics



Basic concepts

Elements of genetics



Basic concepts

Sequencing technologies

Biochemical methods and techniques for the determination of nucleotide order



Basic concepts

Sequencing technologies

Biochemical methods and techniques for the determination of nucleotide order

Break DNA
strands



Perform
analytical tests



Assemble DNA
strands



Basic concepts

Sequencing technologies

Biochemical methods and techniques for the determination of nucleotide order

Break DNA
strands



Perform
analytical tests



Assemble DNA
strands

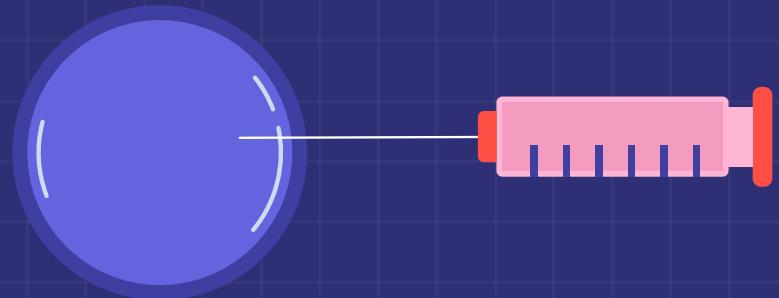


Several types of techniques (with different characteristics) that have evolved over time



02

What is prenatal diagnosis?



Context



There are about 6.000 genetic diseases

10%

4%

21%

Population suffering from
a genetic disease during
their lifetime

Neonates with a
genetic disease

Neonatal deaths due to
genetic diseases

Prenatal diagnosis



**Identification of genetic
genetic diseases before birth**

Due to parental age or mutations, genetic diseases in existing children or indications in the results of other clinical tests

Prenatal diagnosis

 Non-invasive	Ultrasound Defects at first sight	Hormone analysis Hormones in the blood of the pregnant person	Fetal DNA DNA mutations
 Invasive	Amniocentesis Chromosomal defects	Chorionic villus sampling Chromosomal defects and DNA mutations	Placental biopsy Chromosomal defects

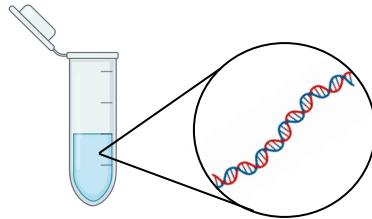
Prenatal diagnosis

 Non-invasive	Ultrasound Defects at first sight	Hormone analysis Hormones in the blood of the pregnant person	Fetal DNA DNA mutations
 Invasive	Amniocentesis Chromosomal defects	Chorionic villus sampling Chromosomal defects and DNA mutations	Placental biopsy Chromosomal defects

Workflow

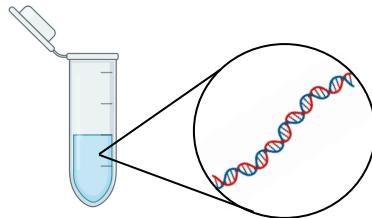


1. DNA extraction

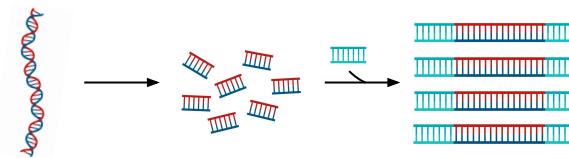


Workflow

1. DNA extraction

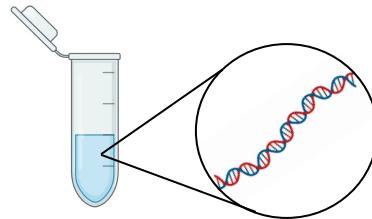


2. DNA preparation

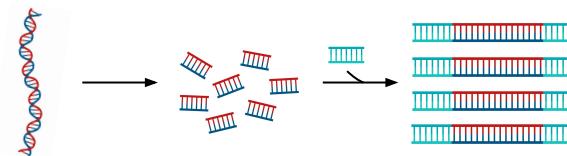


Workflow

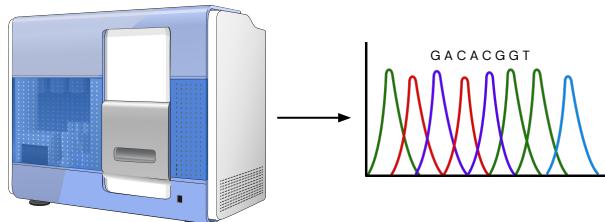
1. DNA extraction



2. DNA preparation



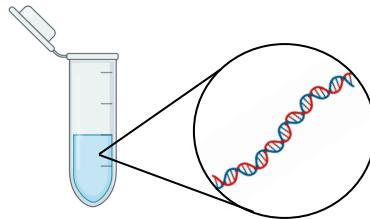
3. Sequencing



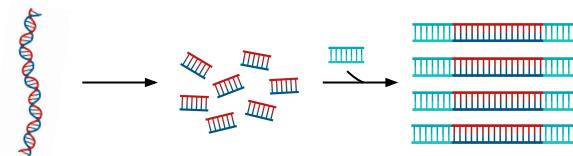
Workflow



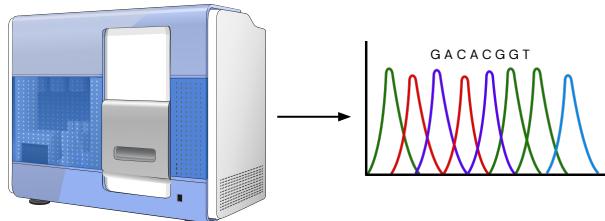
1. DNA extraction



2. DNA preparation



3. Sequencing



4. Analysis

Alignment



Variant
identification

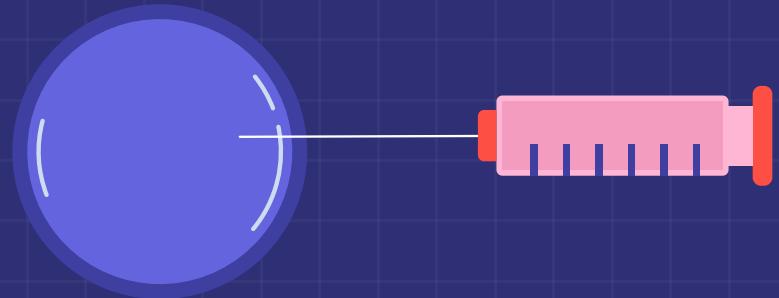




03

Practical case

Diagnosis with Python



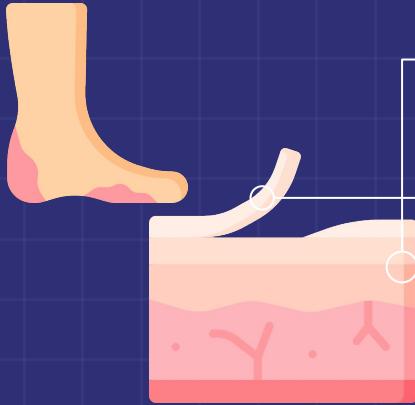
Epidermolysis bullosa



Rare genetic disease (1 in 60.000 newborns)



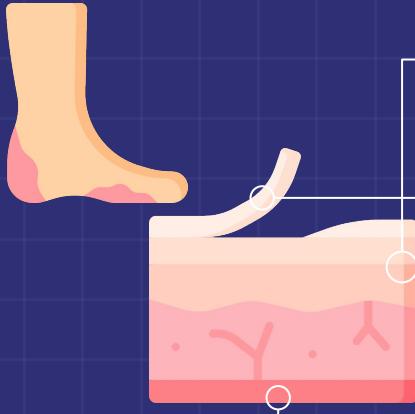
Epidermolysis bullosa



Rare genetic disease (1 in 60.000 newborns)

Causes extreme fragility in mucous membranes and skin, causing blisters to appear on contact or even spontaneously (butterfly skin)

Epidermolysis bullosa

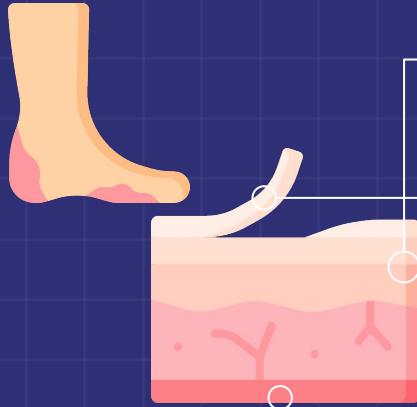


Rare genetic disease (1 in 60.000 newborns)

Causes extreme fragility in mucous membranes and skin, causing blisters to appear on contact or even spontaneously (butterfly skin)

Other complications include dental lesions, muscle atrophy, narrowing of the respiratory, gastrointestinal and urogenital tracts, and cancer

Epidermolysis bullosa



Rare genetic disease (1 in 60.000 newborns)

Causes extreme fragility in mucous membranes and skin, causing blisters to appear on contact or even spontaneously (butterfly skin)

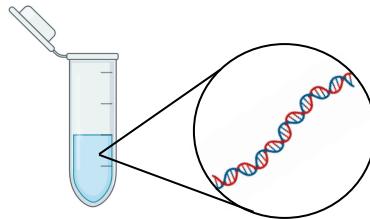
Other complications include dental lesions, muscle atrophy, narrowing of the respiratory, gastrointestinal and urogenital tracts, and cancer

Caused by mutations in genes of skin or mucosal proteins, such as collagen VII, laminins, integrins...

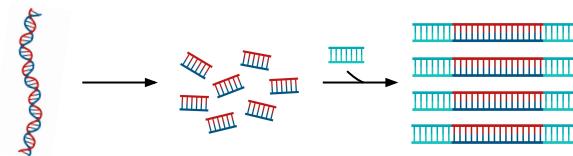
Workflow



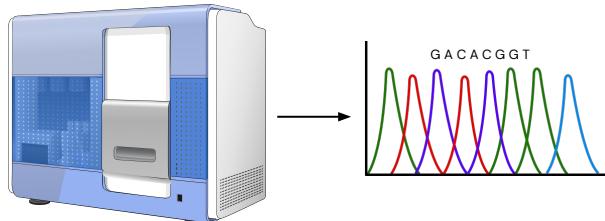
1. DNA extraction



2. DNA preparation



3. Sequencing



4. Analysis

Alignment

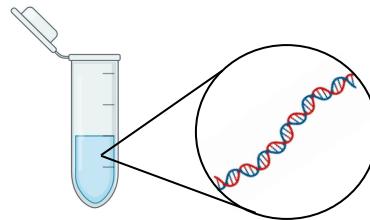


Variant
identification

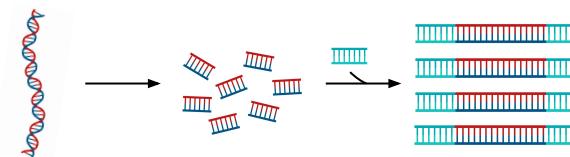


Workflow

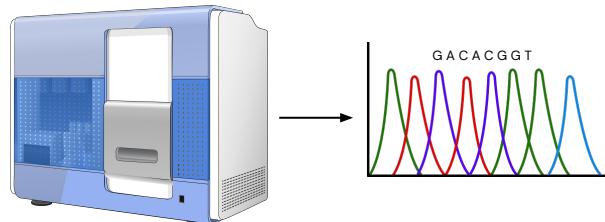
1. DNA extraction



2. DNA preparation



3. Sequencing



4. Analysis

Alignment

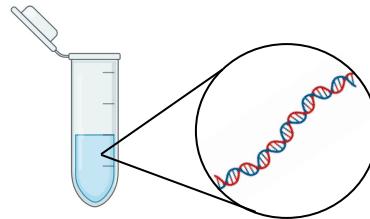


Variant
identification

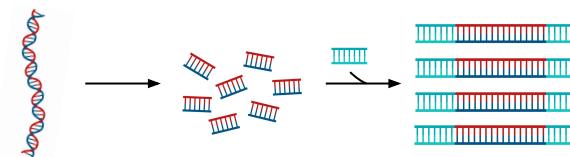


Workflow

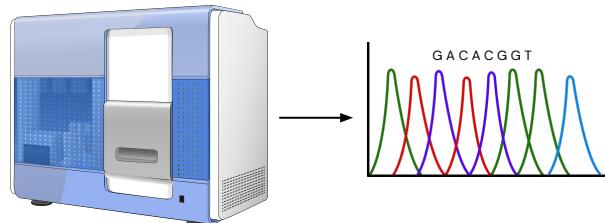
1. DNA extraction



2. DNA preparation



3. Sequencing



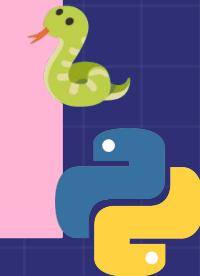
Alignment



Variant
identification



4. Analysis





Log in

Nucleotide

Nucleotide

COL7A1



Search

[Create alert](#) [Advanced](#)

Help

Species

[Animals \(1,368\)](#)
[Bacteria \(4\)](#)
[Customize ...](#)

Summary 20 per page Sort by Default order

Send to:

Filters: [Manage Filters](#)

Molecule types

[genomic DNA/RNA \(623\)](#)
[mRNA \(697\)](#)
[Customize ...](#)

Source databases

[INSDC \(GenBank\) \(534\)](#)
[RefSeq \(1,043\)](#)
[Customize ...](#)

Sequence Type

[Nucleotide \(1,572\)](#)
[EST \(4\)](#)
[GSS \(3\)](#)

Sequence length

[Custom range...](#)

Release date

[Custom range...](#)

GENE

Was this helpful?

[COL7A1 – collagen type VII alpha 1 chain](#)[Homo sapiens \(human\)](#)

Also known as: EBD1, EBDCT, EBR1, NDNC8

Gene ID: 1294

[RefSeq transcripts \(25\)](#)[RefSeq proteins \(15\)](#)[RefSeqGene \(1\)](#)[PubMed \(202\)](#)[Orthologs](#)[Genome Data Viewer](#)[BLAST](#)

RefSeq Sequences

Results by taxon

Top Organisms

[synthetic construct \(138\)](#)
[Homo sapiens \(59\)](#)
[Pan troglodytes \(24\)](#)
[Tyto alba \(20\)](#)
[Motacilla alba \(19\)](#)
[All other taxa \(2265\)](#)
[More...](#)

Find related data

Database: [Select](#) [Find items](#)

Search details

COL7A1[All Fields]



Nucleotide

Nucleotide

Advanced

Search

Help

Species

Animals (1)

Customize ...

Molecule types

mRNA (1)

Customize ...

Source databases

RefSeq (1)

Customize ...

Sequence Type

Nucleotide (1)

Sequence length

Custom range...

Release date

Custom range...

Revision date

Custom range...

[Clear all](#)

Summary 20 per page Sort by Default order

Send to:

Filters: [Manage Filters](#)

Items: 1 to 20 of 25

<< First < Prev Page of 2 Next > Last >>[Find items](#)

Find related data

Database: [Select](#) [Find items](#)

Recent activity

[Turn Off](#) [Clear](#)

COL7A1 collagen type VII alpha 1 chain [Homo sapiens] Gene

Homo sapiens collagen type VII alpha 1 chain (COL7A1), mRNA Nucleotide

COL7A1 (2525) Nucleotide

[See more...](#)

- [Homo sapiens collagen type VII alpha 1 chain \(COL7A1\), mRNA](#)
1. 9,231 bp linear mRNA
Accession: NM_000094.4 GI: 1812585486
[Protein](#) [PubMed](#) [Taxonomy](#)
[GenBank](#) [FASTA](#) [Graphics](#)

[PREDICTED: Homo sapiens collagen type VII alpha 1 chain \(COL7A1\), transcript variant X1, mRNA](#)

 2. 9,171 bp linear mRNA
Accession: XM_017005688.2 GI: 2217341740
[GenBank](#) [FASTA](#) [Graphics](#)

[PREDICTED: Homo sapiens collagen type VII alpha 1 chain \(COL7A1\), transcript variant X5, mRNA](#)

 3. 7,950 bp linear mRNA
Accession: XM_017005689.2 GI: 2217341744

Homo sapiens collagen type VII

ncbi.nlm.nih.gov/nuccore/NM_000094.4

National Library of Medicine
National Center for Biotechnology Information

Log in

Nucleotide Nucleotide Search Advanced Help

GenBank

Homo sapiens collagen type VII alpha 1 chain (COL7A1), mRNA

NCBI Reference Sequence: NM_000094.4

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

Go to: ▾

Locus: NM_000094, 9231 bp mRNA linear PRI 15-FEB-2024

Definition: Homo sapiens collagen type VII alpha 1 chain (COL7A1), mRNA.

Accession: NM_000094 XM_011533337

Version: NM_000094.4

Keywords: RefSeq; MANE Select.

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 9231)

Authors: Natale, M.I., Manzur, G.B., Lusso, S.B., Cella, E., Giovo, M.E.,

Send to: ▾

Choose Destination

Complete Record (radio button selected)

Coding Sequences

Gene Features

File (radio button selected)

Clipboard

Collections

Analysis Tool

Download 1 item.

Format: FASTA (dropdown menu circled)

Show GI

Create File

Articles about the COL7A1 gene

Trial of Beremagene Geperpavec (B-VEC) for Dystrophic Epidermolysis Bullosa [N Engl J Med. 2022]

CDS Feature 1 of 1 NM_000094 : 1 segment Details Display: FASTA GenBank Help

```
def main():
    gene_seq = read_sequence_from_file('Please select the file with the gene of reference')
    patient_seq = read_sequence_from_file('Please select the file with the patient sequence')
    alignment = align_sequences(gene_seq, patient_seq)
    mutation_positions = get_mutation_positions(alignment)
    report_mutations(mutation_positions, gene_seq, patient_seq)

if __name__ == "__main__":
    main()
```



marinamoropez / aligment-EP2025

```
def read_sequence_from_file(prompt):
    print(prompt)
    file = askopenfile(mode='r')
    sequence = file.readlines()[1:]
    sequence = ''.join(sequence).replace('\n', '')
    return sequence
```

```
def main():
    gene_seq = read_sequence_from_file('Please select the file with the gene of reference')
    patient_seq = read_sequence_from_file('Please select the file with the patient sequence')
    alignment = align_sequences(gene_seq, patient_seq)
    mutation_positions = get_mutation_positions(alignment)
    report_mutations(mutation_positions, gene_seq, patient_seq)

if __name__ == "__main__":
    main()
```

```
from tkinter.filedialog import askopenfile

def read_sequence_from_file(prompt):
    print(prompt)
    file = askopenfile(mode='r')
    sequence = file.readlines()[1:]
    sequence = ''.join(sequence).replace('\n', '')
    return sequence
```

```
def main():
    gene_seq = read_sequence_from_file('Please select the file with the gene of reference')
    patient_seq = read_sequence_from_file('Please select the file with the patient sequence')
    alignment = align_sequences(gene_seq, patient_seq)
    mutation_positions = get_mutation_positions(alignment)
    report_mutations(mutation_positions, gene_seq, patient_seq)

if __name__ == "__main__":
    main()
```

```
def align_sequences(gene_seq, patient_seq):
    alignment = [gene_seq[position] == patient_seq[position] for position in range(len(gene_seq))]
    return alignment

def get_mutation_positions(alignment):
    mutation_positions = [position for position, match in enumerate(alignment) if not match]
    return mutation_positions
```

```
def main():
    gene_seq = read_sequence_from_file('Please select the file with the gene of reference')
    patient_seq = read_sequence_from_file('Please select the file with the patient sequence')
    alignment = align_sequences(gene_seq, patient_seq)
    mutation_positions = get_mutation_positions(alignment)
    report_mutations(mutation_positions, gene_seq, patient_seq)

if __name__ == "__main__":
    main()
```

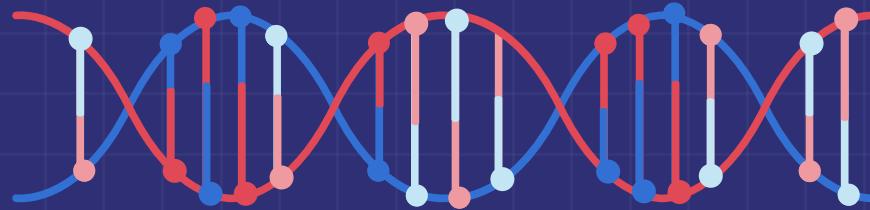
```
def report_mutations(mutation_positions, gene_seq, patient_seq):
    for mutation in mutation_positions:
        print(f"Mutation position: {mutation + 1}")
        print(f"Original base: {gene_seq[mutation]}")
        print(f"Mutated base: {patient_seq[mutation]}")
```

**Recessive dystrophic epidermolysis bullosa caused
by T>C at position 6591 of the COL7A1 gene (among others)**

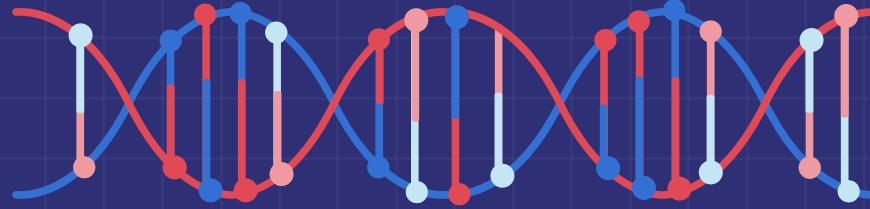


Recessive dystrophic epidermolysis bullosa caused by T>C at position 6591 of the COL7A1 gene (among others)

ORIGINAL
COL7A1 GENE

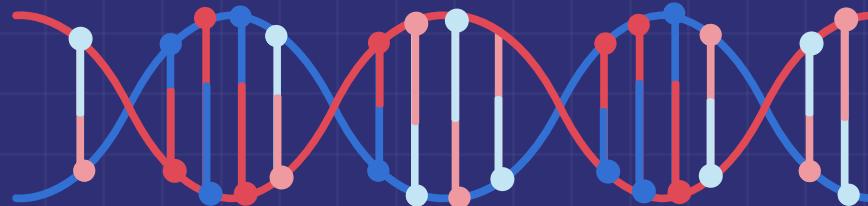


MUTATED
COL7A1 GENE

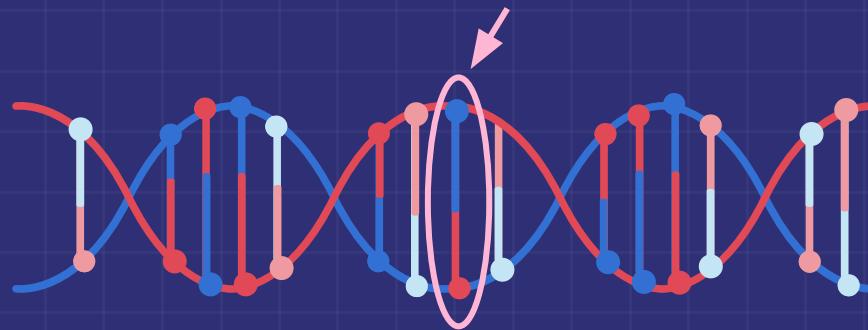


Recessive dystrophic epidermolysis bullosa caused by T>C at position 6591 of the COL7A1 gene (among others)

ORIGINAL
COL7A1 GENE



MUTATED
COL7A1 GENE





Recessive dystrophic epidermolysis bullosa caused by T>C at position 6591 of the COL7A1 gene (among others)

ORIGINAL COL7A1 GENE

TCTCTGGAGAACAGGGACCCCTGGACTCAAGGGTGCTAAGGGGGAGCCGGGCAGCAATGGTACCAAGG
TCCCAAAGGAGACAGGGGTGTGCCAGGCATCAAAGGAGACCGGGAGAGCCTGGACCGAGGGTCAGGAC
GGCAACCCGGGTCTACCAGGAGAGCGTGGTATGGCTGGGCTGAAGGGGAAGCCGGGTCTGCAGGGTCAA
GAGGCCCCCCCTGGCCCAGTGGGTGGTCATGGAGACCCCTGGACCAACCTGGTCCCCGGGTCTGCTGGCC
TGCAGGACCCAAGGACCTTCTGGCCTGAAGGGGGAGCCTGGAGAGACAGGACCTCCAGGACGGGACCTG
ACTGGACCTACTGGAGCTGTGGGACTTCCTGGACCCCCCGGCCCTCAGGCCTTGTGGGTCACAGGGGT
CTCCAGGTTTGCCTGGACAAGTGGGGAGACAGGGAAGCCGGAGCCCCAGGTCGAGATGGTGCCAGTGG

MUTATED COL7A1 GENE

TCTCTGGAGAACAGGGACCCCTGGACTCAAGGGTGCTAAGGGGGAGCCGGGCAGCAATGGTACCAAGG
TCCCAAAGGAGACAGGGGTGTGCCAGGCATCAAAGGAGACCGGGAGAGCCTGGACCGAGGGTCAGGAC
GGCAACCCGGGTCTACCAGGAGAGCGTGGTATGGCTGGGCTGAAGGGGAAGCCGGGTCTGCAGGGTCAA
GAGGCCCCCCCTGGCCCAGTGGGTGGTCATGGAGACCCCTGGACCAACCTGGTCCCCGGGTCTGCTGGCC
TGCAGGACCCAAGGACCTTCTGGCCTGAAGGGGGAGCCTGGAGAGACAGGACCTCCAGGACGGGACCTG
ACTGGACCTACTGGAGCTGTGGGACTTCCTGGACCCCCCGGCCCTCAGGCCTTGTGGGTCACAGGGGT
CTCCAGGTTTGCCTGGACAAGTGGGGAGACAGGGAAGCCGGAGCCCCAGGTCGAGATGGTGCCAGTGG



Recessive dystrophic epidermolysis bullosa caused by T>C at position 6591 of the COL7A1 gene (among others)

TCTCTGGAGAACAGGGACCCCTGGACTCAAGGGTGCTAAGGGGGAGCCGGGCAGCAATGGTACCAAGG
TCCCAAAGGAGACAGGGGTGTGCCAGGCATCAAAGGAGACCGGGAGAGCCTGGACCGAGGGGTCAAGGAC
GGCAACCCGGGTCTACCAGGAGAGCGTGGTATGGCTGGGCTGAAGGGAAAGCCGGGTCTGCAGGGTCAA

ORIG runfile('C:/Users/Usuario/Documents/Charlas/Códigos/Alignment/
COL7A1 seq_alignment.py', wdir='C:/Users/Usuario/Documents/Charlas/Códigos/
Alignment')
Please select the file with the gene of reference
Please select the file with the patient sequence
Mutation position: 6591
Original base: T
Mutated base: C

MUTA
COL7A1 GENE TGCAAGGACCCCAAGGACCTTCTGGCCTGAAGGGGGAGCCTGGAGAGACAGGACCTCCAGGACGGGCCTG
ACTGGACCTACTGGAGCTGTGGGACTTCCTGGACCCCCCGGCCCTCAGGCCTTGTGGGTCACAGGGT
CTCCAGGTTTGCCTGGACAAGTGGGGAGACAGGAAAGCCGGAGCCCCAGGTCGAGATGGTGCCAGTGG



Recessive dystrophic epidermolysis bullosa caused by T>C at position 6591 of the COL7A1 gene (among others)

ORIGI
COL7A1

```
In [2]: runfile('C:/Users/Usuario/Documents/Charlas/Códigos/Alignment/  
seq_alignment.py', wdir='C:/Users/Usuario/Documents/Charlas/Códigos/  
Alignment')
```

Please select the file with the gene of reference

Please select the file with the patient sequence

Mutation position: 884

Original base: A

Mutated base: T

Mutation position: 6591

Original base: T

Mutated base: C

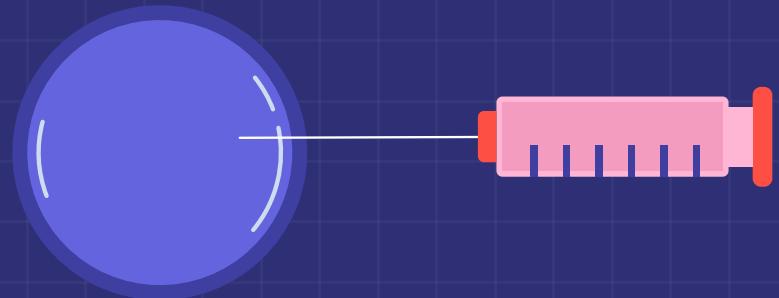
```
CTCCAGGTTGCCTGGACAAGTGGGGAGACAGGAAAGCCGGAGCCCCAGTCGAGATGGTGCCAGTGG
```





04

Regulation and commercialization



Regulation and commercialization

Classification of the techniques



Regulation and commercialization

Classification of the techniques

Regulation (EU) 2017/746 on *in vitro* diagnostic medical devices (IVDR)



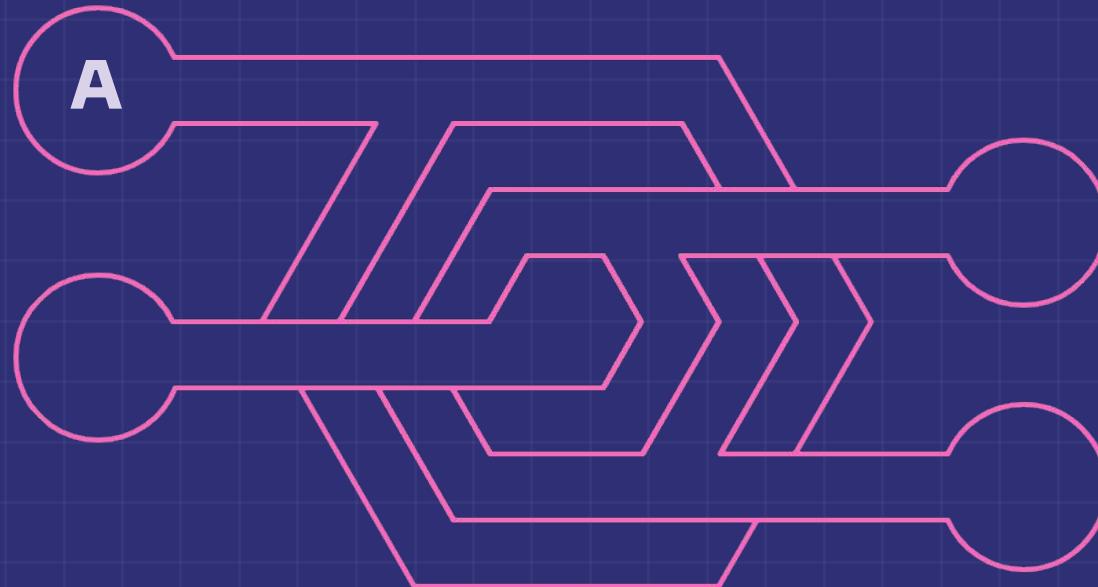
Regulation and commercialization

Classification of the techniques



Regulation (EU) 2017/746 on *in vitro* diagnostic medical devices (IVDR)

↓ Individual risk
↓ Public risk



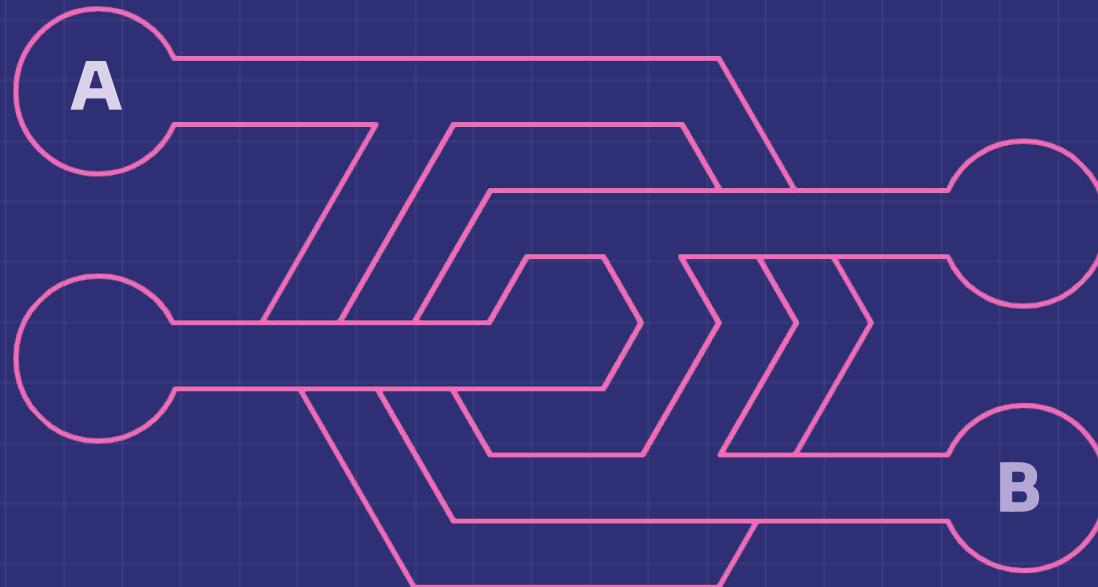
Regulation and commercialization

Classification of the techniques



Regulation (EU) 2017/746 on *in vitro* diagnostic medical devices (IVDR)

↓ Individual risk
↓ Public risk

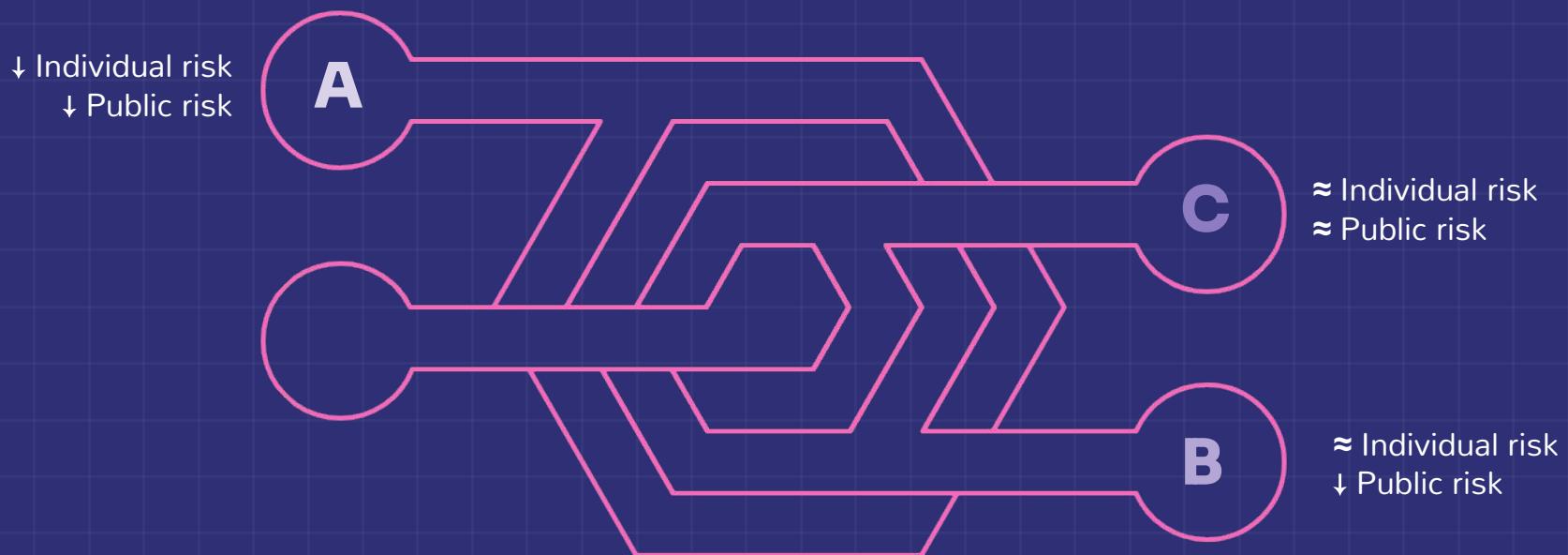


Regulation and commercialization

Classification of the techniques



Regulation (EU) 2017/746 on *in vitro* diagnostic medical devices (IVDR)

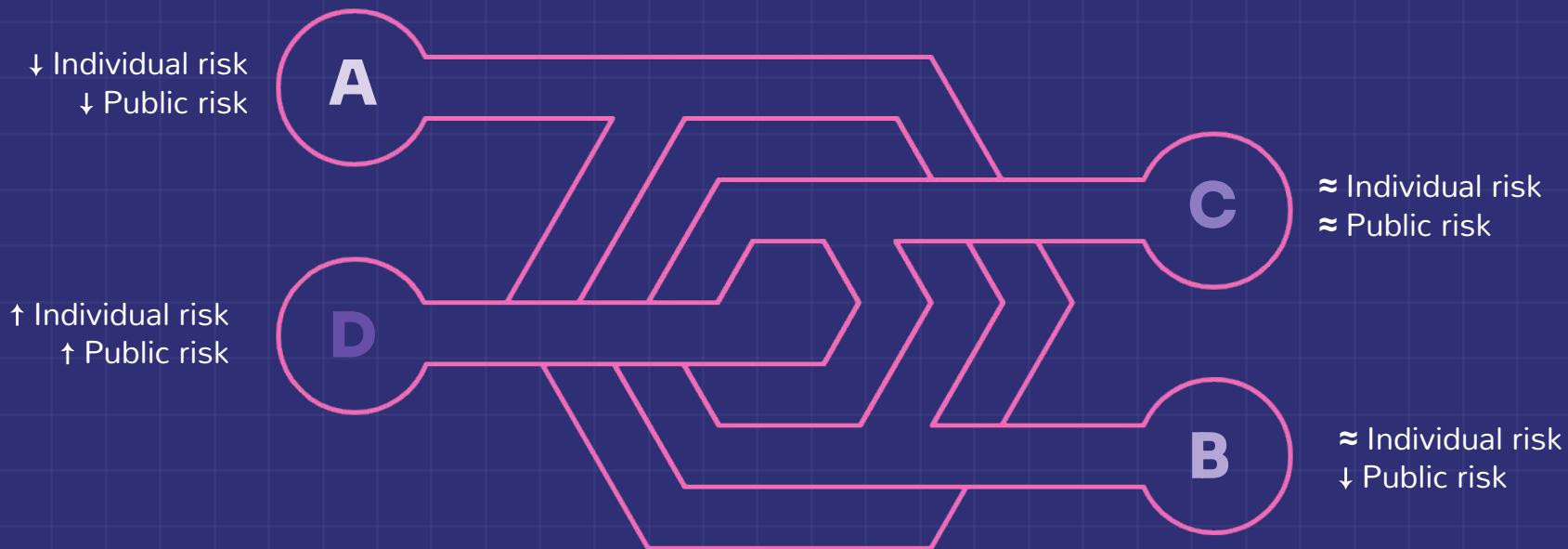


Regulation and commercialization

Classification of the techniques



Regulation (EU) 2017/746 on *in vitro* diagnostic medical devices (IVDR)

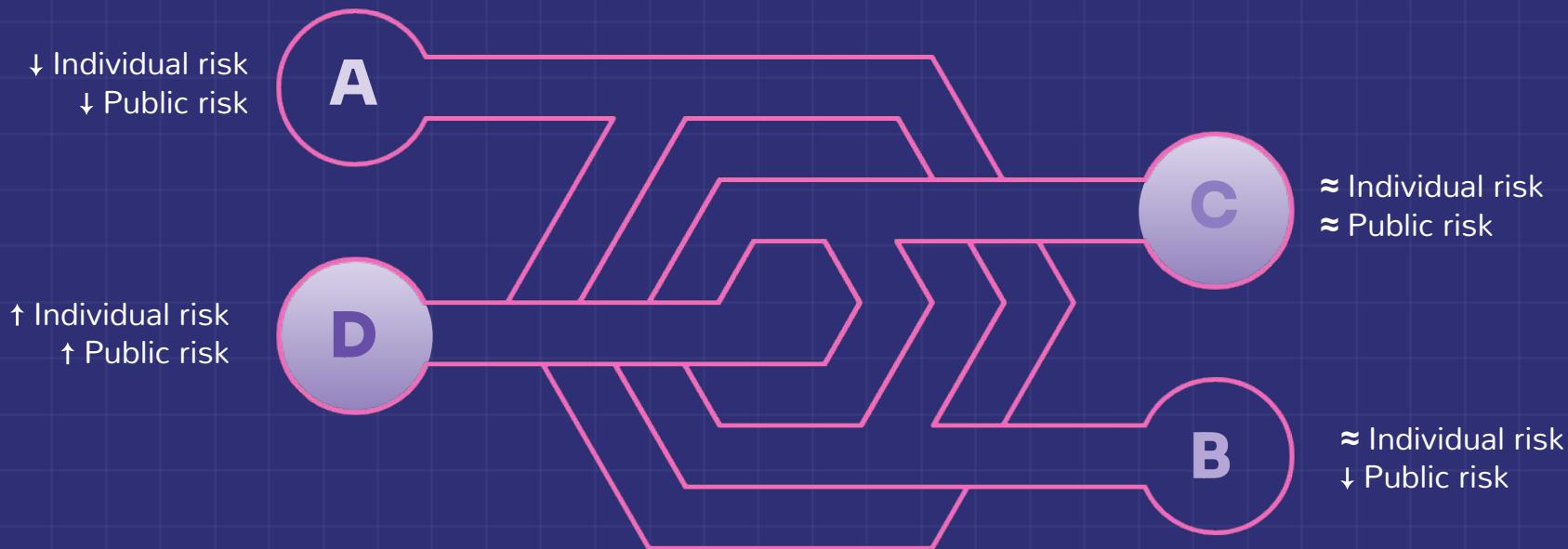


Regulation and commercialization

Classification of the techniques



Regulation (EU) 2017/746 on *in vitro* diagnostic medical devices (IVDR)



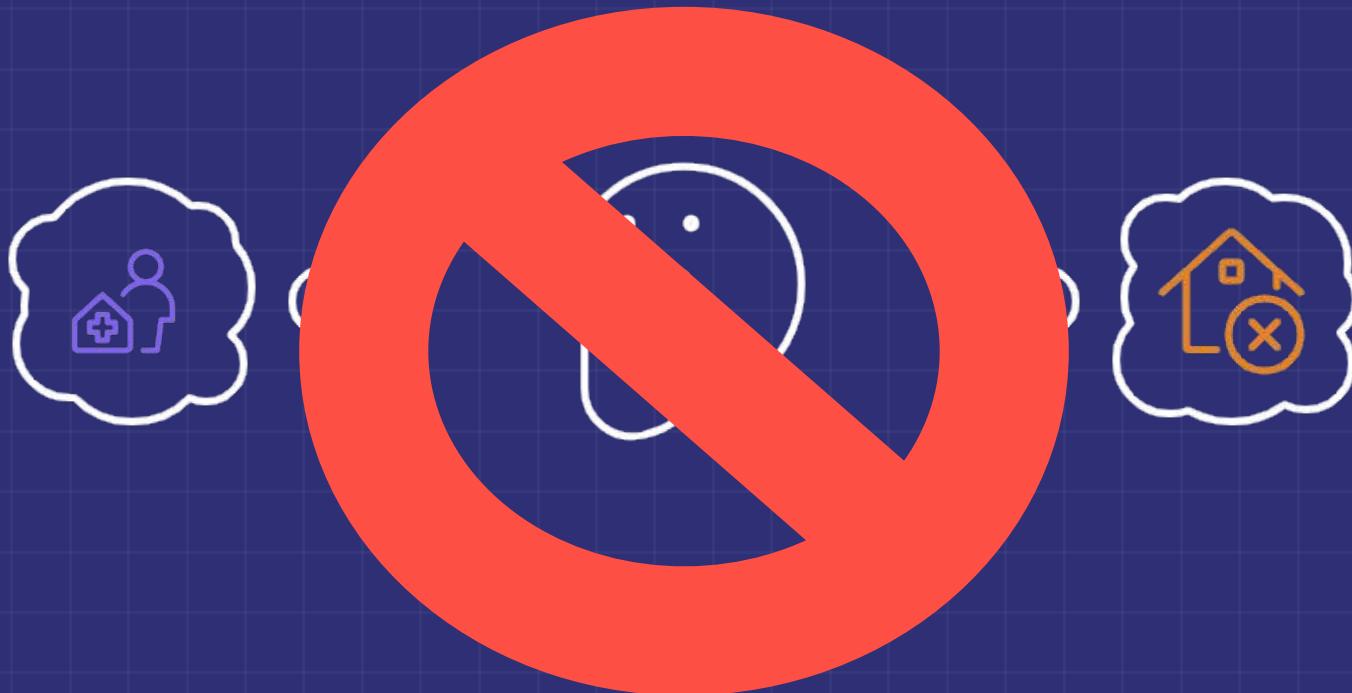
Regulation and commercialization

Use of the techniques



Regulation and commercialization

Use of the techniques



Regulation and commercialization

Genetic counseling



Regulation and commercialization

Genetic counseling



Professional with knowledge of genetics



Regulation and commercialization

Genetic counseling



Professional with knowledge of genetics



They set out the information in a clear and understandable manner



Regulation and commercialization

Genetic counseling



Professional with knowledge of genetics



They set out the information in a clear and understandable manner



They give information on available options



Regulation and commercialization

Genetic counseling



Professional with knowledge of genetics



They set out the information in a clear and understandable manner



They give information on available options



They help in making informed decisions



Regulation and commercialization

Genetic counseling



Professional with knowledge of genetics



They set out the information in a clear and understandable manner



They give information on available options



They help in making informed decisions



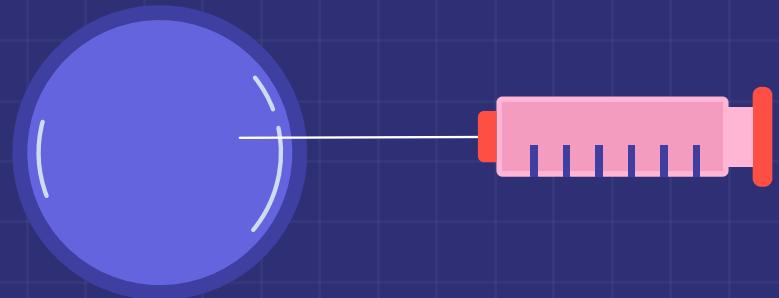
Psychological monitoring



05

Clinical applications

Clinic, responsibilities and important considerations



Applications

What may we use this information for?

Applications

What may we use this information for?

Medical Xpress Topics Conditions Week's top Latest news Unread news Subscribe Search User

It appears that you are currently using Ad Blocking software. What are the consequences? X

Home / Genetics Home / Diseases, Conditions, Syndromes ☆ 📄 📋

FEBRUARY 18, 2025 The GIST Editors' notes

Using CRISPR to remove extra chromosomes in Down syndrome

by PNAS Nexus

Allele-Specific CRISPR/Cas9 Chromosomal Rescue for Trisomy 21

Human cells with Trisomy 21 → Allele-specific multiple chromosome cleavages (SpCas9 + gRNA(s)) → Restored cell function

NEW YORK POST

SECTIONS | SEARCH | EMAIL

Lifestyle Health Wellness Fitness Health Care Medicine Men's Health Women's Health Mental

HEALTH

Could Down syndrome be eliminated? Scientists say cutting-edge gene editing tool could cut out extra chromosome

By Reda Wigle Published June 19, 2025, 10:17 a.m. ET

30 Comments

earth.com NATURE SCIENCE LIFE NEWS VIDEOS IMAGES EARTHPEDIA TAKE ACTION EARTHSNAP SEARCH SUBSCRIBE

ATGACCGCTTGGATTGCCTGAAGGTACG

CRISPR used to remove extra chromosomes in Down syndrome and restore cell function

Applications

What may we use this information for?

Article | Published: 17 July 2013

Translating dosage compensation to trisomy 21

Jun Jiang, Yuanchun Jing, Gregory J. Cost, Jen-Chieh Chiang, Heather J. Kolpa, Allison M. Cotton, Dawn M. Carone, Benjamin R. Carone, David A. Shivak, Dmitry Y. Guschin, Jocelynn R. Pearl, Edward J. Rebar, Meg Byron, Philip D. Gregory, Carolyn J. Brown, Fyodor D. Urnov , Lisa L. Hall & Jeanne B. Lawrence 

Nature 500, 296–300 (2013) | [Cite this article](#)

77k Accesses | 603 Altmetric | [Metrics](#)

JOURNAL ARTICLE

Trisomic rescue via allele-specific multiple chromosome cleavage using CRISPR-Cas9 in trisomy 21 cells

Ryotaro Hashizume , Sachiko Wakita, Hirofumi Sawada, Shin-ichiro Takebayashi, Yasuji Kitabatake, Yoshitaka Miyagawa, Yoshifumi S Hirokawa, Hiroshi Imai, Hiroki Kurahashi [Author Notes](#)

PNAS Nexus, Volume 4, Issue 2, February 2025, pgaf022,

<https://doi.org/10.1093/pnasnexus/pgaf022>

Published: 18 February 2025 [Article history](#) ▾



PDF



Split View



Cite



Permissions



Share ▾

Applications

How should we use this information?

Applications

How should we use this information?



Universal Declaration on
the Human Genome and
Human Rights

*"The human genome is the basis for the fundamental humanity of all members of the human family and for the recognition of their intrinsic dignity and diversity. In a symbolic sense, the human genome is the **world heritage**."*

Applications

How should we use this information?



Universal Declaration on
the Human Genome and
Human Rights

*“The human genome is the basis for the fundamental humanity of all members of the human family and for the recognition of their intrinsic dignity and diversity. In a symbolic sense, the human genome is the **world heritage**. ”*



Universal Declaration of
Human Rights

Applications

What may we use this information for?

*“Interventions in the human genome should only be allowed for **preventive, diagnostic or therapeutic** reasons and without modifications to descendants, the alternative would be equivalent to endangering the inherent and therefore equal dignity of all human beings and thus **renewing eugenics.**”*

Update on the IBC's reflection on the human genome and human rights – International Bioethics Committee (IBC)



Thank you!

Questions?



helengopo@gmail.com | marina@europython.eu



@HelenaGomezPz | @marinamorolopez



Helena Gómez Pozo | Marina Moro López



@helenagomezpz.bsky.social | @marinamorolopez.bsky.social



marinamorolopez / aligment-EP2025

