# PRIMER DESIGN

## Geni e genomi

## Ensembl

# Primer pairs need to be:

- Target site specific
- Cover all requirements for effective amplification during PCR

# What information do we need:

- Precise information on target sequence → data base with reference genome sequence
- In silico prediction of «ideal» PCR primer pairs

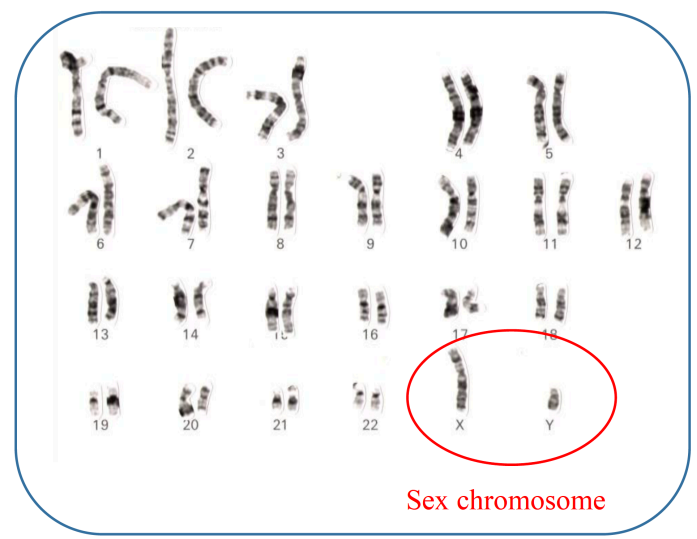
# Designing primers using genome databases and

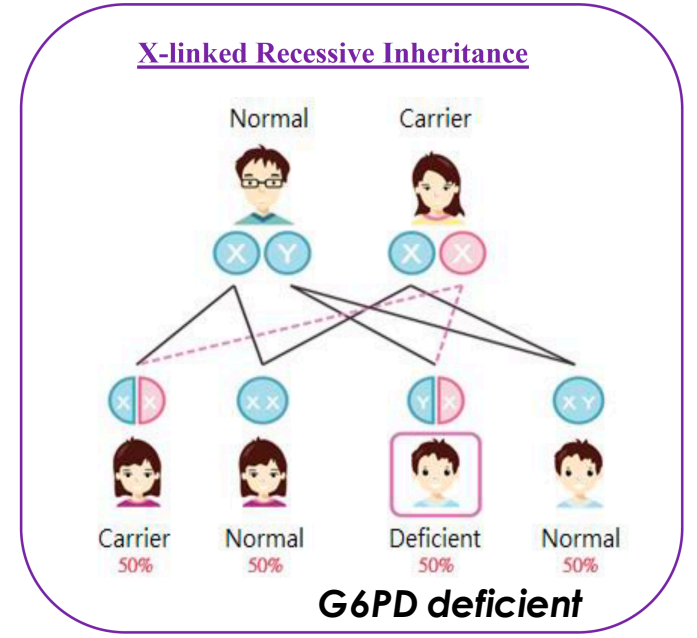
# primer design software

**EXAMPLE:**

**Molecular test: Develop an PCR based approach to amplify region of interest in Exon 6 of the G6PD gene and perform DNA sequencing to identify gene mutations**

# Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency Anemia



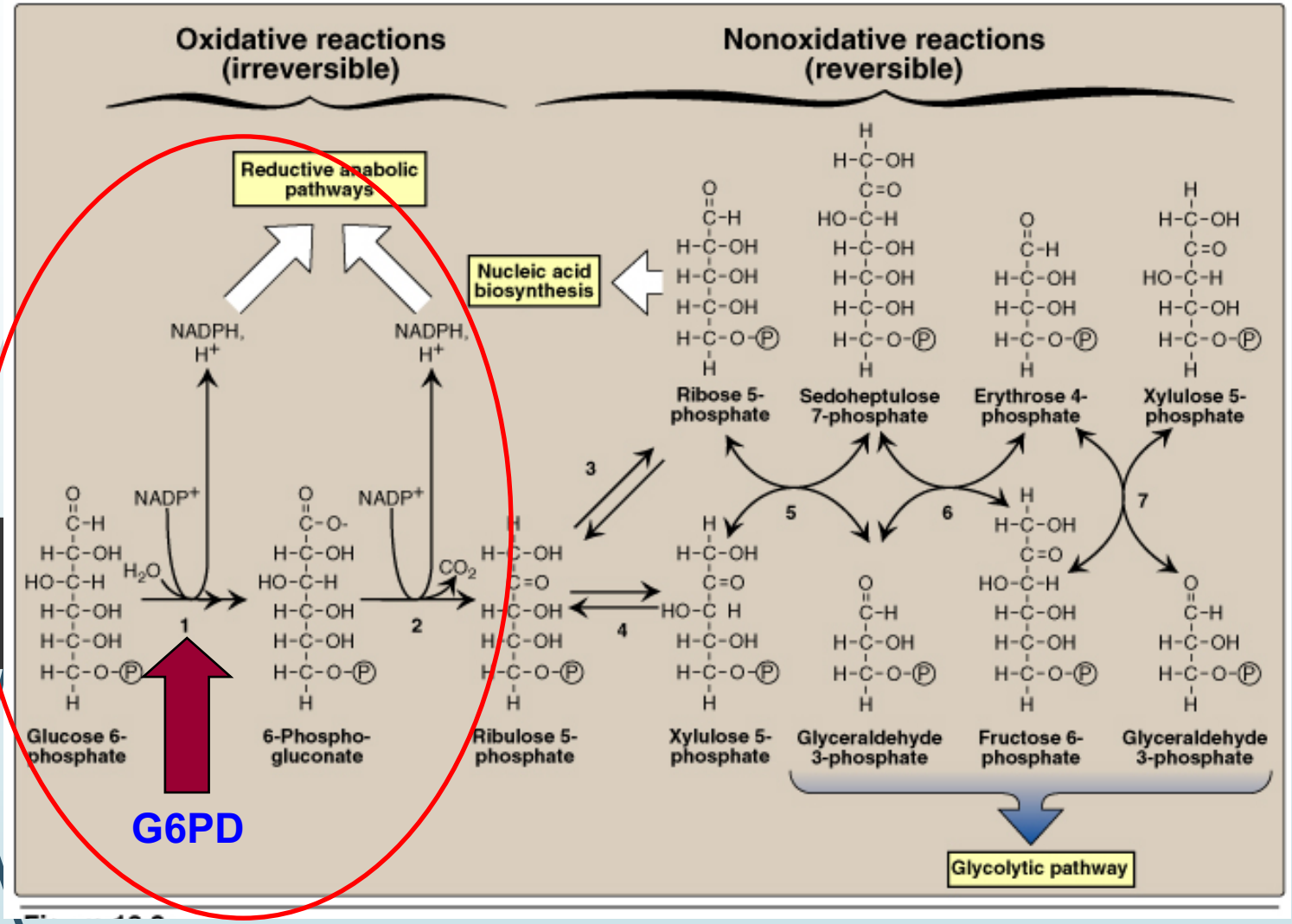




***The gene coding for G6PD enzyme is located on the X chromosome.***

***Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency Anemia is a X-linked recessive disease***

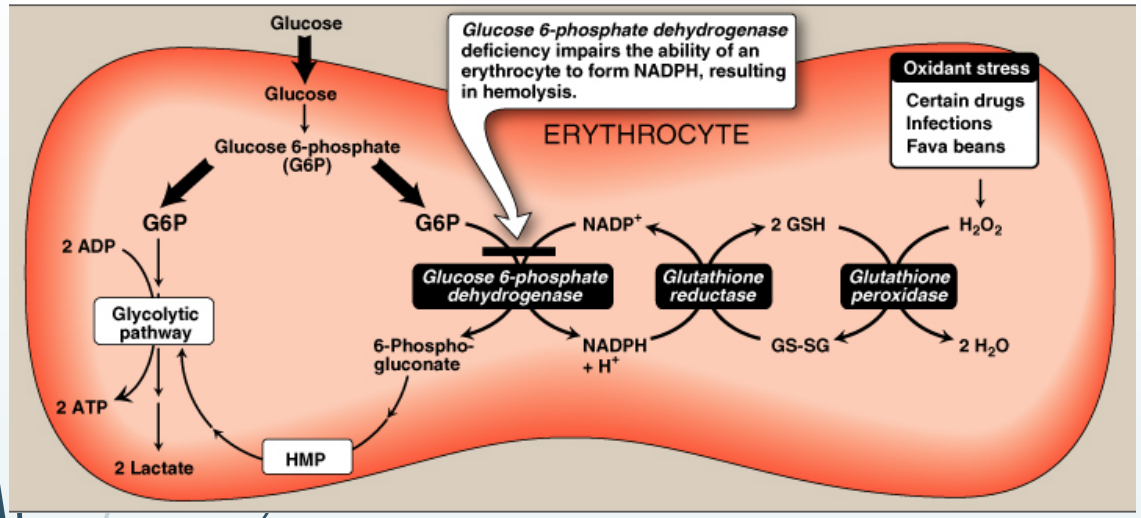
# Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency Anemia

## Glucose-6-Phosphate Dehydrogenase produces NADPH



# Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency Anemia

## Glucose-6-Phosphate Dehydrogenase produces NADPH




**Uses of NADPH**

- **Reductive biosynthesis e.g., fatty acid biosynthesis**
- **Antioxidant (part of glutathione system)**
- **Oxygen-dependent phagocytosis by WBCs**
- **Synthesis of nitric oxide (NO)**

**NADPH: Nicotinamide adenine dinucleotide phosphate**

- NADPH provides the reducing equivalents for biosynthetic reactions and the oxidation-reduction involved in protecting against the toxicity of reactive oxygen species (ROS)
- Glutathion (GSH) is oxidized to Glutathion disulfide (GSSG) when it neutralizes harmful ROS like $H_2O_2$
- NADPH is used to to recycle GSSG to GSH by the Glutathion reductase (GSSG + HADPH + $H^+$ → 2 GSH + $NADP^+$)

If mutations in the *G6PD* gene reduce the amount of glucose-6-phosphate dehydrogenase or alter its structure, this enzyme can no longer play its protective role. As a result, reactive oxygen species can accumulate and damage red blood cells. A a consequence, red blood cells to be destroyed faster than the body can replace them = hemolysis. A reduction in the number of red blood cells causes the signs and symptoms of hemolytic anemia.

**DETAILLED INFO: https://ghr.nlm.nih.gov/condition/glucose-6-phosphate-dehydrogenase-deficiency**

# Oxidative stress: imbalance between oxidant production and antioxidant mechanisms

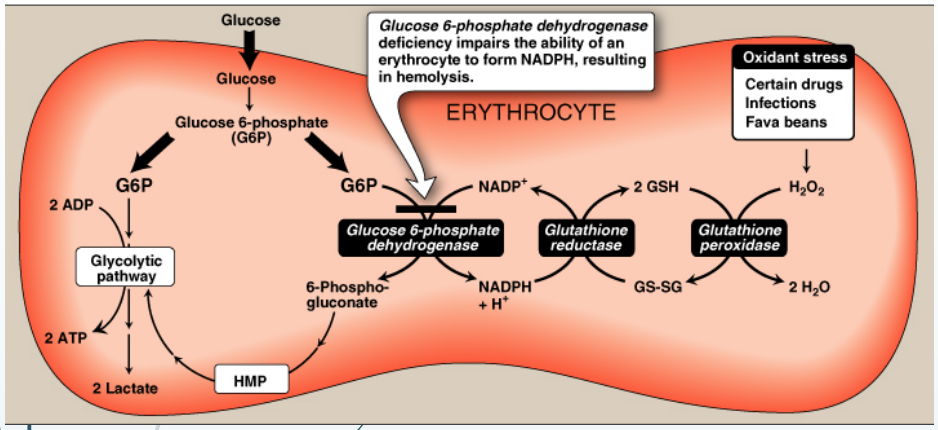
**Oxidative damage to:**

- DNA
- Proteins
- Lipids (unsaturated fatty acids)

**Oxidative stress and diseases:**

- Inflammatory conditions e.g., Rheumatoid arthritis
- Atherosclerosis and coronary heart diseases
- Obesity
- Cancers
- G6PD deficiency hemolytic anemia

# Biochemical basis of G6PD Deficiency Hemolytic Anemia, continued…




**Carriers of G6PD do not necessarily develop anemia**

**.. Disease is triggerd by increased increased reactive oxygen species (ROS) levels**

**G6PD deficient patients will develop hemolytic attack upon:**

**1.Intake of oxidant drugs (AAA):**

- **Antibiotics e.g., sulfa preparation**
- **Antimalarial drugs: e.g., Primaquine**
- **Antipyretics**

**2.Exposure to infection**

**3.Ingestion of fava beans (favism, Mediterranean variant)**

# Different Classes of G6PD Deficiency Hemolytic Anemia

**Based on enzymatic activity of G6PD**

## G6PD Class I:

Very severe

## G6PD Mediterranean (class II)

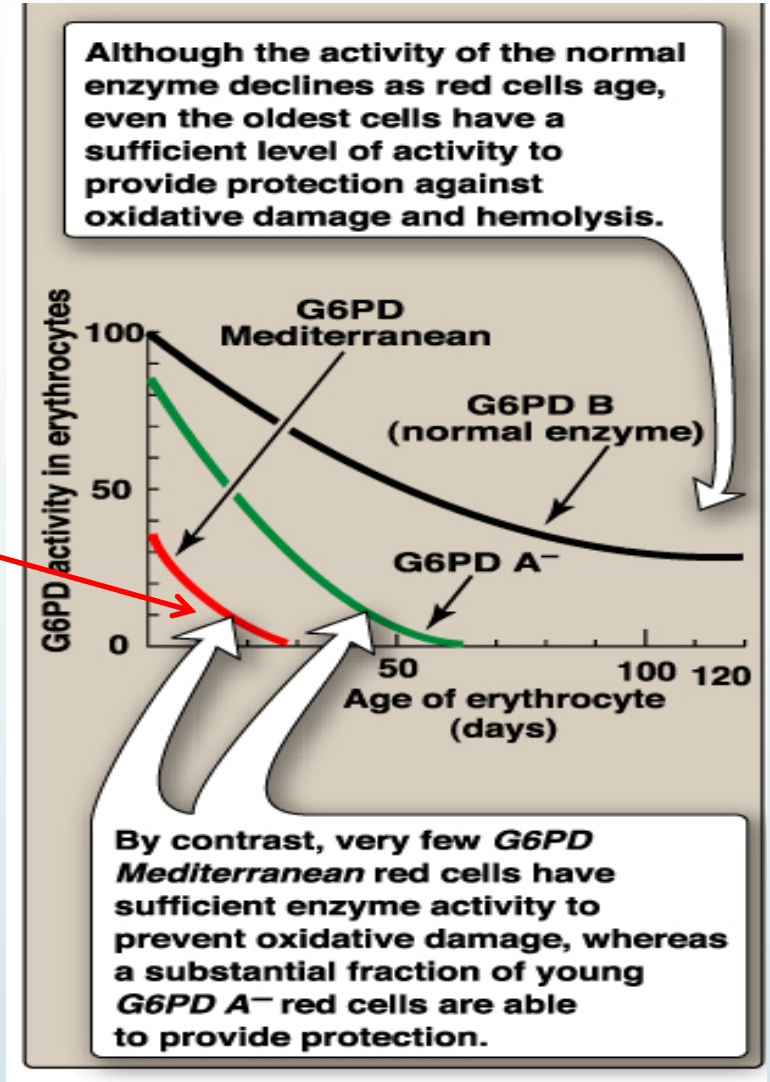
Severe.Enzyme with decreased stability
and activity (severe).
Affect all RBCs (both young and old)

## G6PD A- (class III):

Moderate, young RBCs contain enzymatic activity.
Unstable enzyme, but kinetically normal

## G6PD class IV:

Normal activity



# Diagnosis of
# G6PD Deficiency Hemolytic Anemia

**Diagnosis of hemolytic anemia**

Complete Blood Count (CBC) & reticulocytic count

**Screening:**

Qualitative assessment of G6PD enzymatic activity (UV-based test)

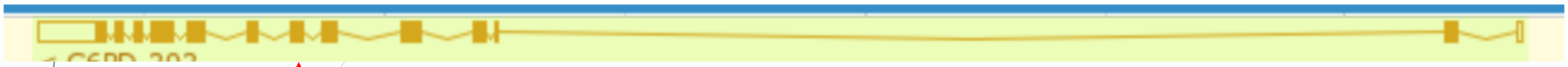
**Confirmatory test:**

Quantitative measurement of G6PD enzymatic activity

**Molecular test: Detection of G6PD gene mutation**

# G6PD Mutations linked with hemolytic anemia

G6PD gene on X chromosome



Exon 6

| Descriptive mutations | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mutation | | | | Gene | | | Protein | | |
| Designation | Short name | Isoform G6PD-Protein | OMIM-Code | Type | Subtype | Position | Position | Structure change | Function change |
| G6PD-A(+) | Gd-A(+) | G6PD A | +305900.0001 | Polymorphism nucleotide | A→G | 376 (Exon 5) | 126 | Asparagine→Aspartic acid (ASN126ASP) | No enzyme defect (variant) |
| G6PD-A(-) | Gd-A(-) | G6PD A | +305900.0002 | Substitution nucleotide | G→A | 376 (Exon 5) and 202 | 68 and 126 | Valine→Methionine (VAL68MET) Asparagine→Aspartic acid (ASN126ASP) | |
| G6PD-Mediterranean | Gd-Med | G6PD B | +305900.0006 | Substitution nucleotide | C→T | 563 (Exon 6) | 188 | Serine→Phenylalanine (SER188PHE) | Class II |
| G6PD-Canton | Gd-Canton | G6PD B | +305900.0021 | Substitution nucleotide | G→T | 1376 | 459 | Arginine→Leucine (ARG459LEU) | Class II |
| G6PD-Chatham | Gd-Chatham | G6PD | +305900.0003 | Substitution nucleotide | G→A | 1003 | 335 | Alanine→Threonine (ALA335THR) | Class II |
| G6PD-Cosenza | Gd-Cosenza | G6PD B | +305900.0059 | Substitution nucleotide | G→C | 1376 | 459 | Arginine→Proline (ARG459PRO) | G6PD-activity <10%, thus high portion of patients. |
| G6PD-Mahidol | Gd-Mahidol | G6PD | +305900.0005 | Substitution nucleotide | G→A | 487 (Exon 6) | 163 | Glycine→Serine (GLY163SER) | Class III |
| G6PD-Orissa | Gd-Orissa | G6PD | +305900.0047 | Substitution nucleotide | C→G | 131 | 44 | Alanine→Glycine (ALA44GLY) | NADP-binding place affected. Higher stability than other variants. |
| G6PD-Asahi | Gd-Asahi | G6PD A- | +305900.0054 | Substitution nucleotide (several) | A→G ± G→A | 376 (Exon 5) 202 | 126 68 | Asparagine→Aspartic acid (ASN126ASP) Valine→Methionine (VAL68MET) | Class III. |

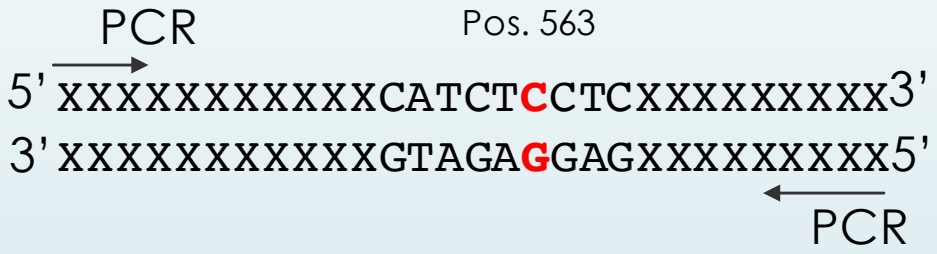
# Molecular test: Develop an PCR based approach to amplify region of interest in Exon 6 of the G6PD gene and perform DNA sequencing to identify gene mutations

1. Recruit patients
2. Prepare biological sample (typically blood)
3. Prepare genomic DNA
4. PCR amplify region of interes
5. DNA sequncing of amplified product

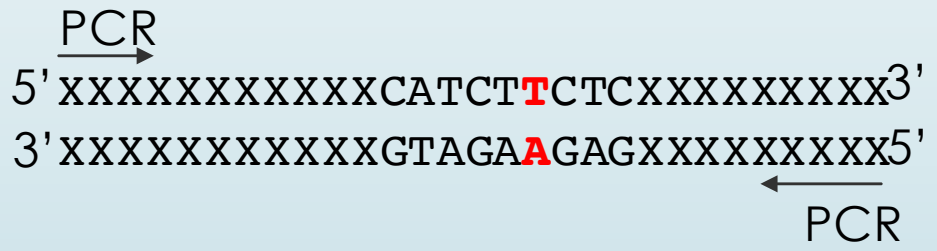
## G6PD 563C→T variant

PCR oligos amplify region of interest in Exon 6 of the G6PD gene (around aminoacid position 563)

**wt** G6PD Exon 6 nucleotides around codon 563 region:




**Mediterranean G6PD Exon 6** nucleotides around codon 563:



# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Transcript/Sequence_cDNA?db=core;g=ENSG00000160211;r=X:154531391-154547572;t=ENST00000393562

## GOAL:

**PCR amplification of a region containing position of point mutation in G6PD gene**

## BIOLOGICAL MATERIAL:

**Genomic DNA from patient specimen**

## STRATEGIA:

1. **Recover exon sequences of G6PD**
2. **Individuate 5' end and 3' end of exon 6**
3. **Recover G6PD gene seqeunces**
4. **Locate nucleotide position of 5' and 3' end of exon 6**
5. **Locate nucleotide position of point mutation**
6. **Use software to identify primer pairs that amplify a region flanking the position of point mutation**
7. **Test PCR primers**

# Genome database: Ensembl

Ensembl è una banca dati bioinformatica allestita con lo scopo di fornire informazioni aggiornate sui principali genomi.

Ensembl è gestita in modo coordinato dal *Wellcome Trust Sanger Institute* e dallo *European Bioinformatics Institute* (EBI), facente parte dello *European Molecular Biology Laboratory* (EMBL).

Attraverso questa piattaforma bioinformatica è possibile avere accesso ad informazioni sul genoma umano ma anche su genomi di altri organismi modello.

Data l'estrema complessità dei genomi annotati, Ensembl è caratterizzata da una estrema flessibilità, permettendo a chiunque di accedere ad informazioni tra le più varie:

- Ricerca di geni
- Allineamenti tra sequenze
- Ricerca di specifice sequenze su cromosomi (utilizzando coordinate)
- Sequenze di RNA, proteine,

Il progetto è del tutto *open source*. Ciò significa che sia il software che i dati immagazzinati possono essere utilizzati in maniera del tutto libera e gratuita.

# Cosa è Ensembl?

Ensembl è una banca dati bioinformatica allestita con lo scopo di fornire informazioni aggiornate sui principali genomi.

Ensembl è gestita in modo coordinato dal *Wellcome Trust Sanger Institute* e dallo *European Bioinformatics Institute* (EBI), facente parte dello *European Molecular Biology Laboratory* (EMBL).

Attraverso questa piattaforma bioinformatica è possibile avere accesso ad informazioni sul genoma umano ma anche su genomi di altri organismi modello.

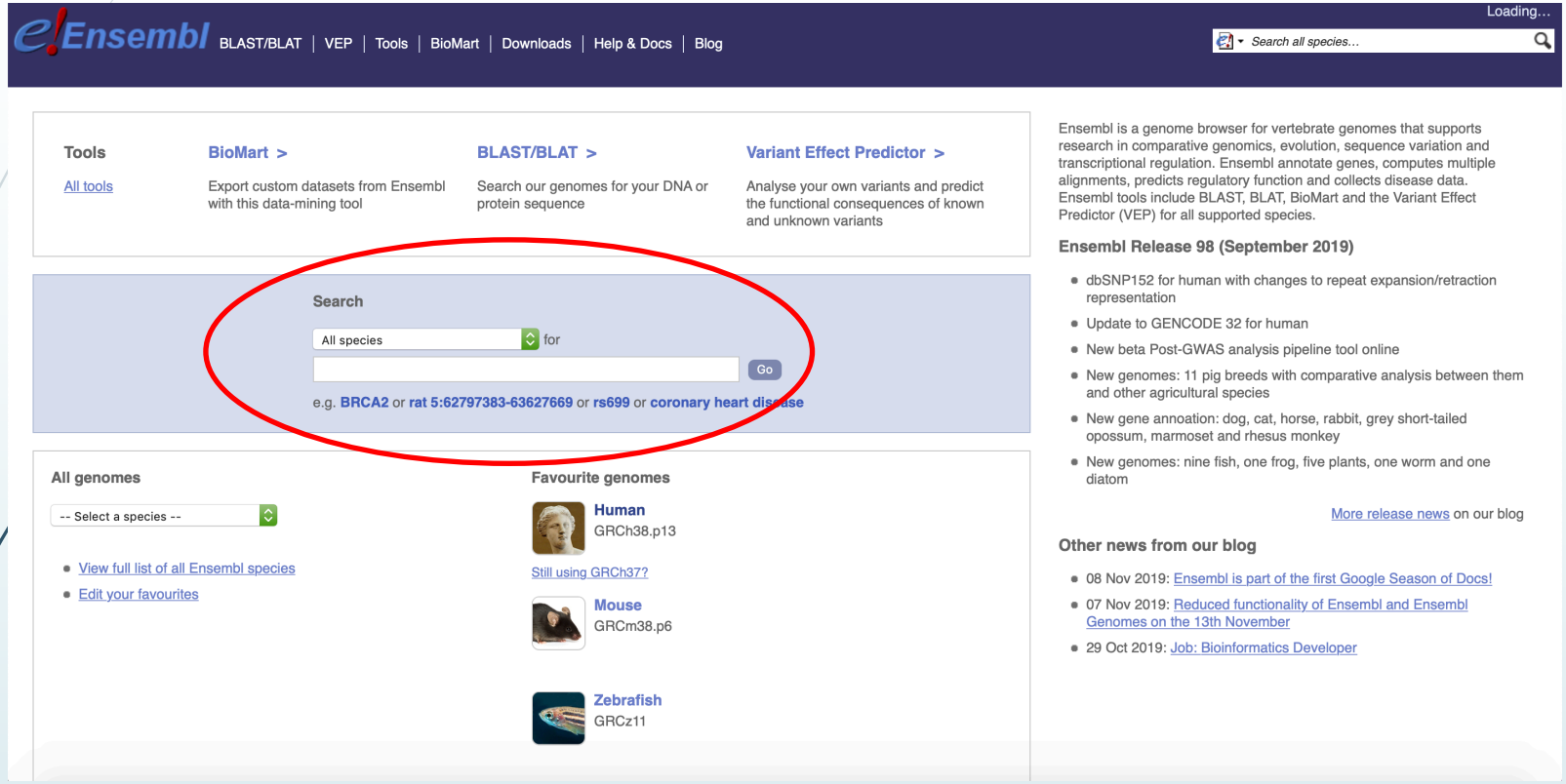
Data l'estrema complessità dei genomi annotati, Ensembl è caratterizzata da una estrema flessibilità, permettendo a chiunque di accedere ad informazioni tra le più varie:

- Ricerca di geni
- Allineamenti tra sequenze
- Ricerca di specifice sequenze su cromosomi (utilizzando coordinate)
- Sequenze di RNA, proteine,

Il progetto è del tutto *open source*. Ciò significa che sia il software che i dati immagazzinati possono essere utilizzati in maniera del tutto libera e gratuita.

# Esploriamo Ensembl

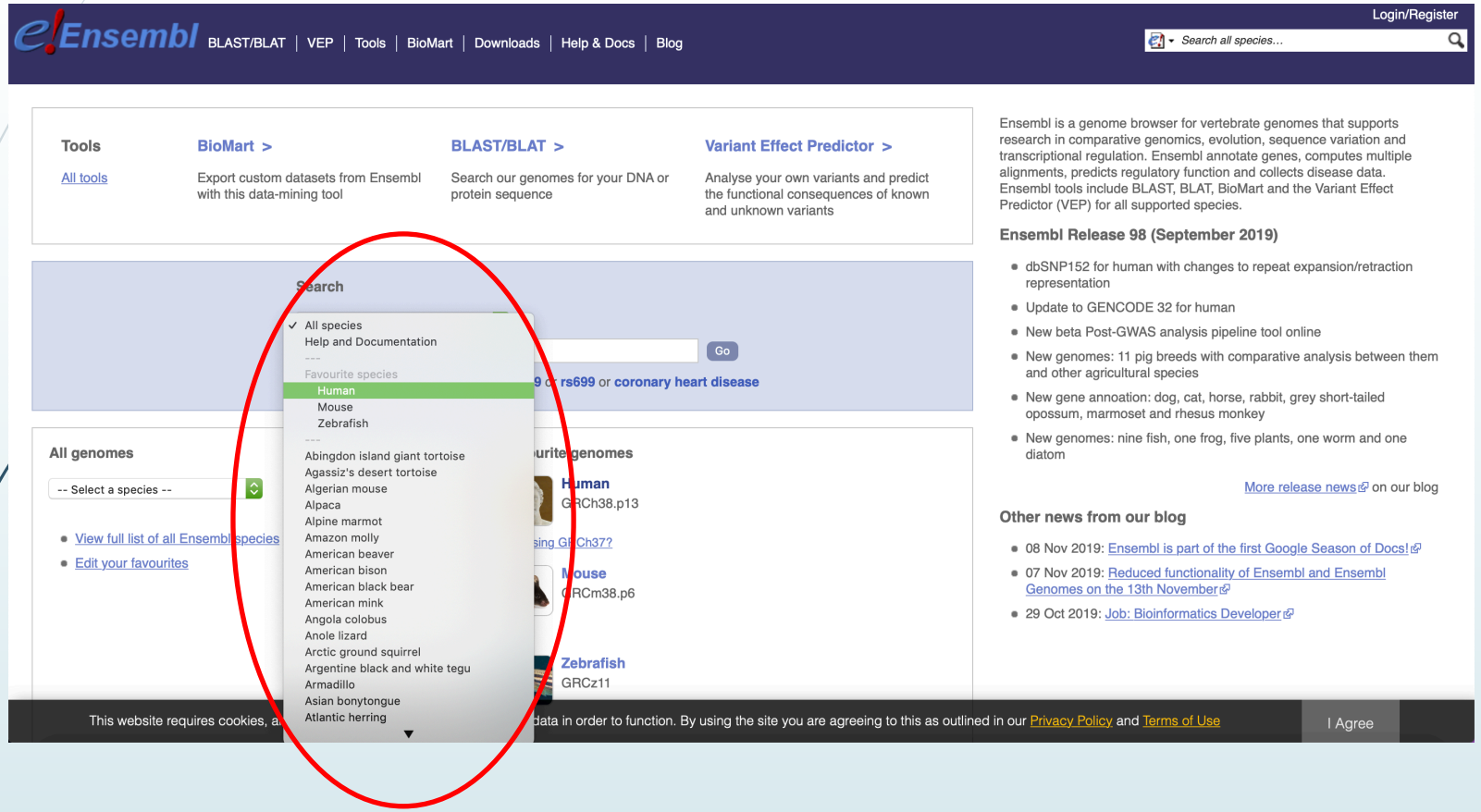
https://www.ensembl.org/index.html




Ensembl ci permette di selezionare il genoma della specie della quale vogliamo ottenere più informazioni.

Cliccando su All species si aprirà un menù a tendina…

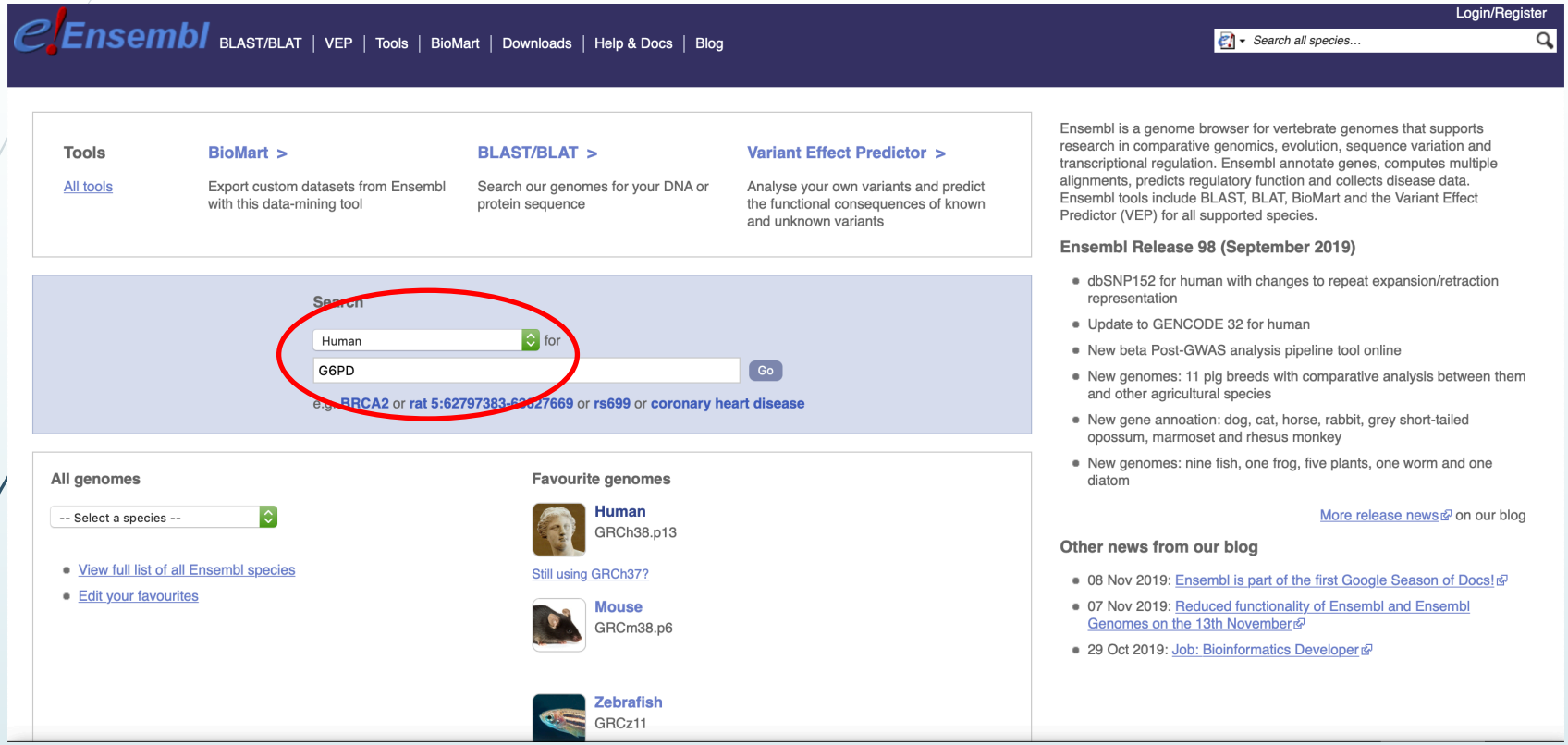
# Esploriamo Ensembl




… che ci permetterà di selezionare la specie che ci interessa.

Nel nostro caso, selezioneremo Human (tra le favourite species) perché andremo a ricercare un gene in particolare.

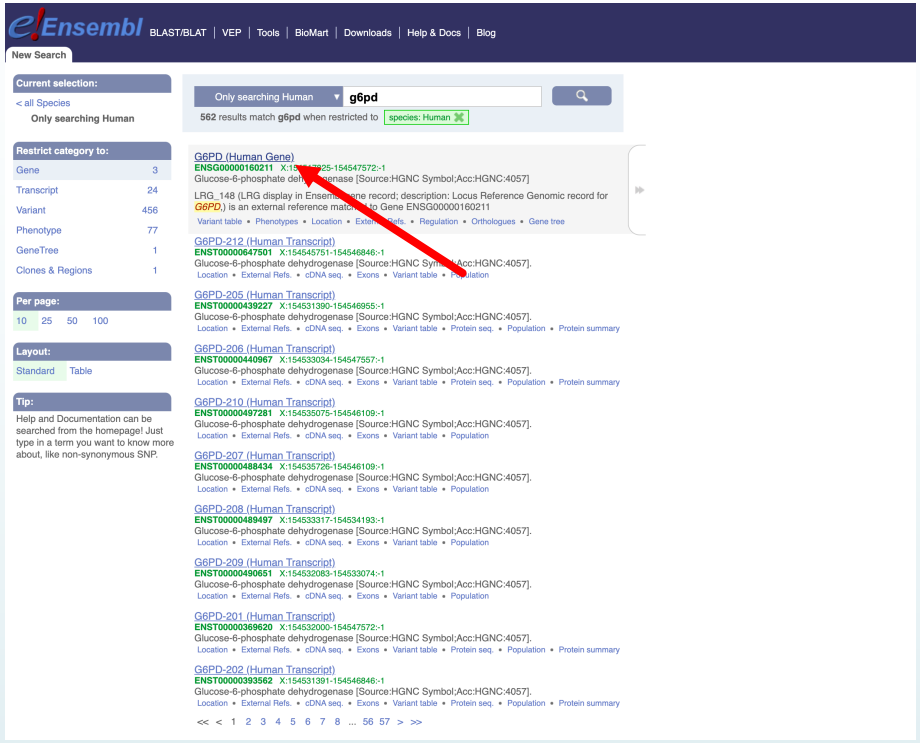
# Esploriamo Ensembl




Il gene in questione è il gene che codifica per la **Glucosio-6-Fosfato Deidrogenasi (G6PD)**.

# Come cercare un gene su Ensembl

https://www.ensembl.org/Human/Search/Results?q=G6PD;site=ensembl;facet_species=Human

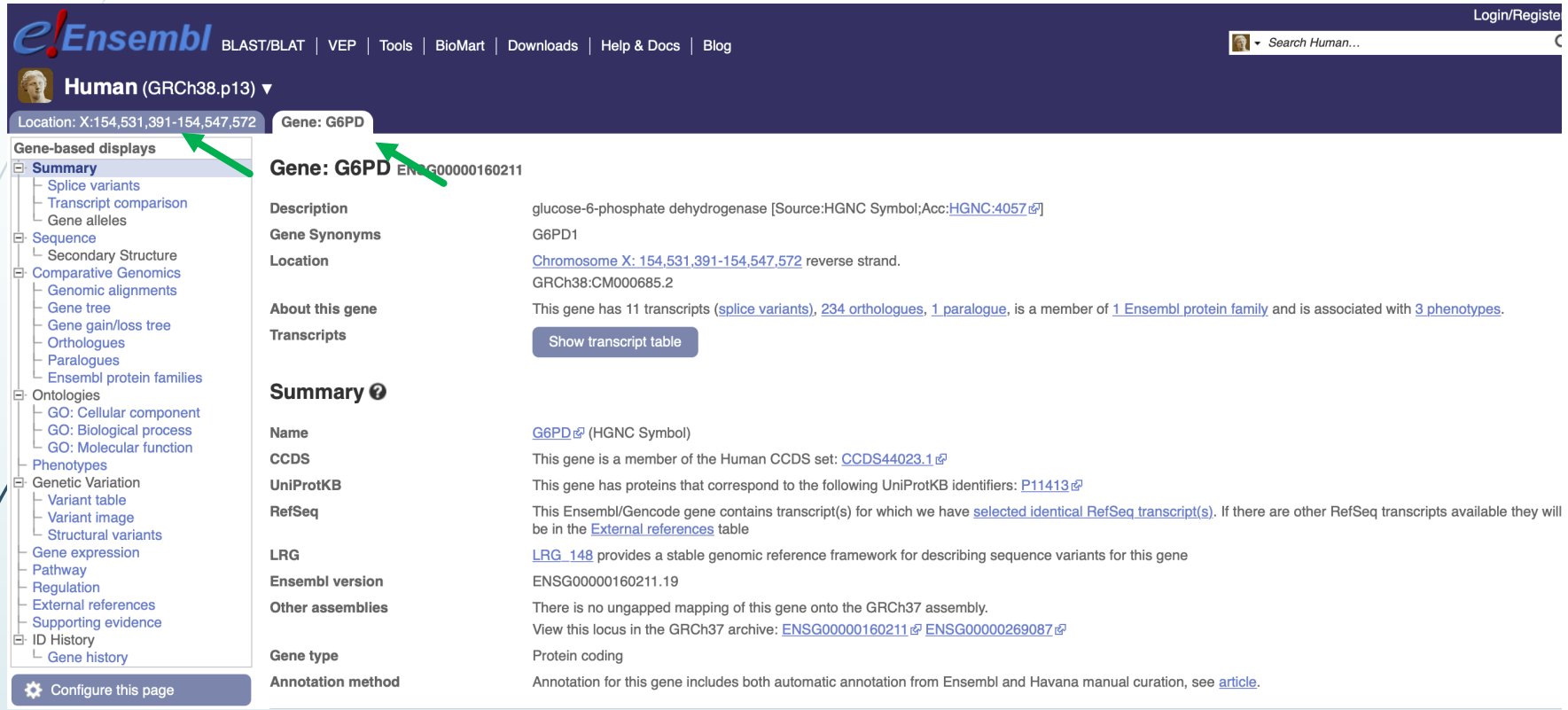



Otteniamo quindi una lista di informazioni all'interno della quale dobbiamo selezionare ciò che è di nostro interesse.

Avremo informazioni sul gene in tutta la sua interezza, ma anche sulle varie isoforme di mRNA.

Selezionando il primo risultato della lista…

# Proviamo insieme con il gene G6PD

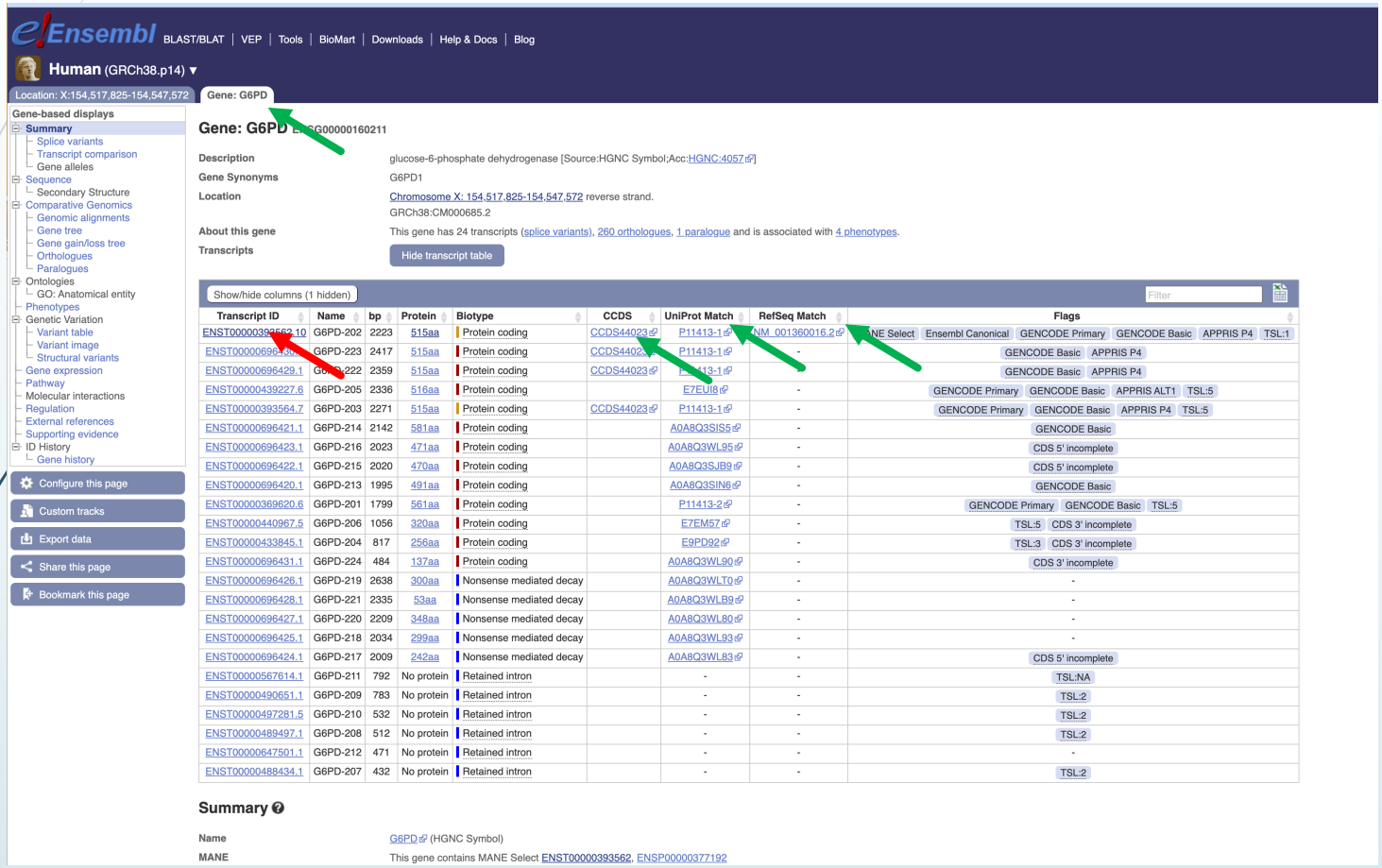



Sono rintracciabili tutte le **informazioni sul gene**:

- Varianti di splicing
- Comparazioni tra i trascritti
- Comparazioni tra i genomi

# Proviamo insieme con il gene G6PD

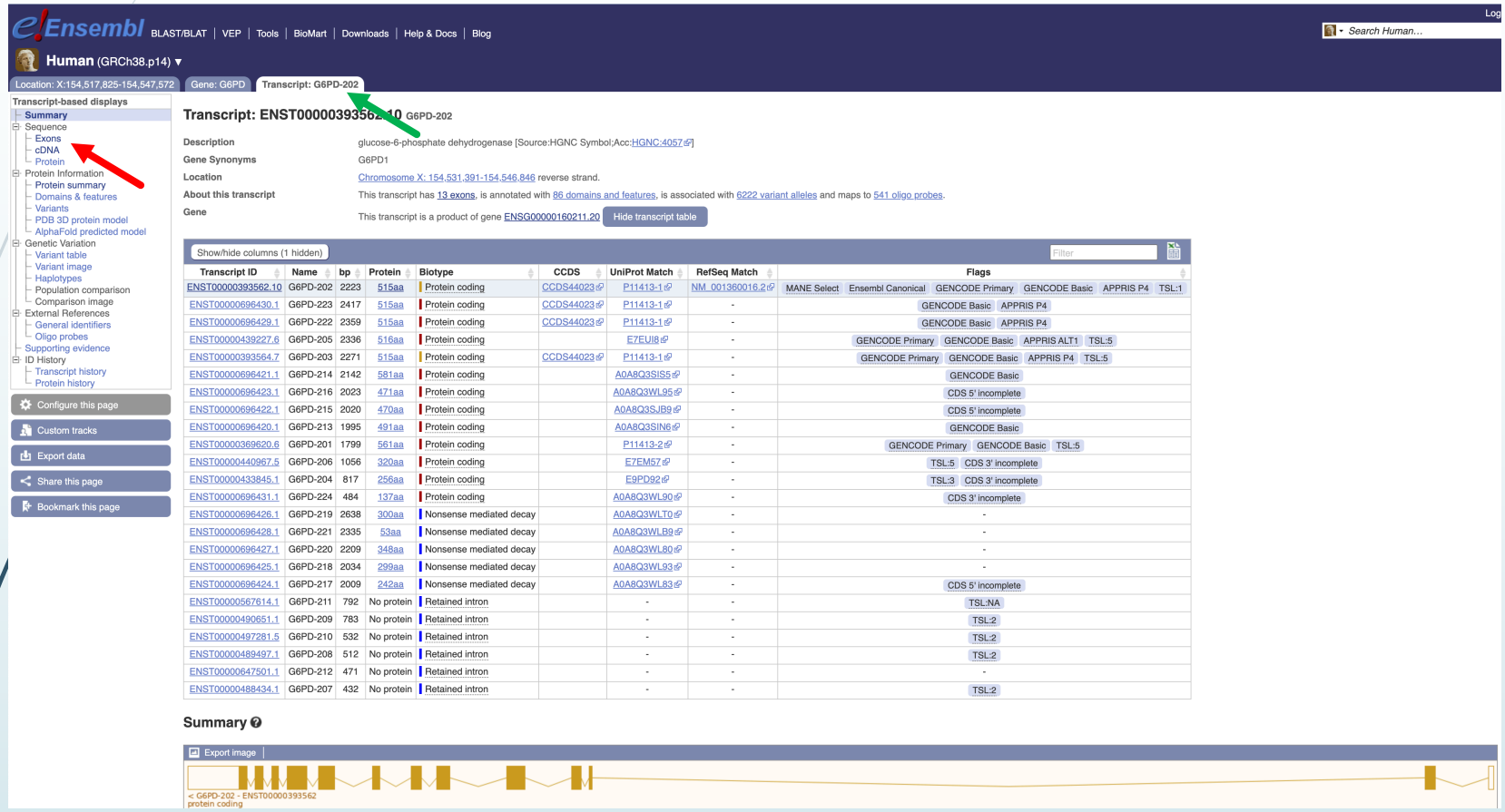
https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000160211;r=X:154531391-154547572




... potremo scoprire che il gene **G6PD** si trova sul **cromosoma X** e che da esso vengono trascritti **23 trascritti annotati diversi (splice variants)**.

# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Gene/Splice?db=core;g=ENSG00000160211;r=X:154531391-154547572

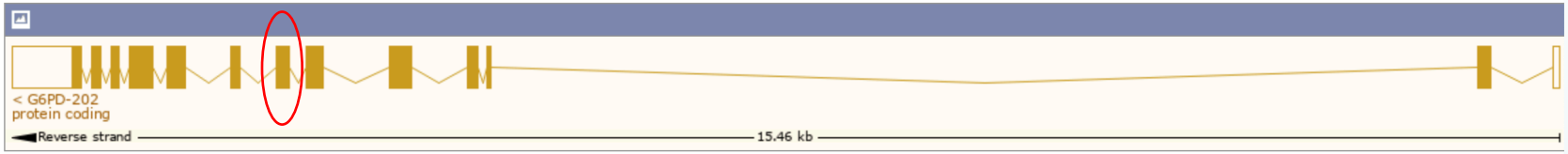


# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Transcript/Summary?db=core;g=ENSG00000160211;r=X:154531391-154547572;t=ENST00000393562

## G6PD-202

**Summary**



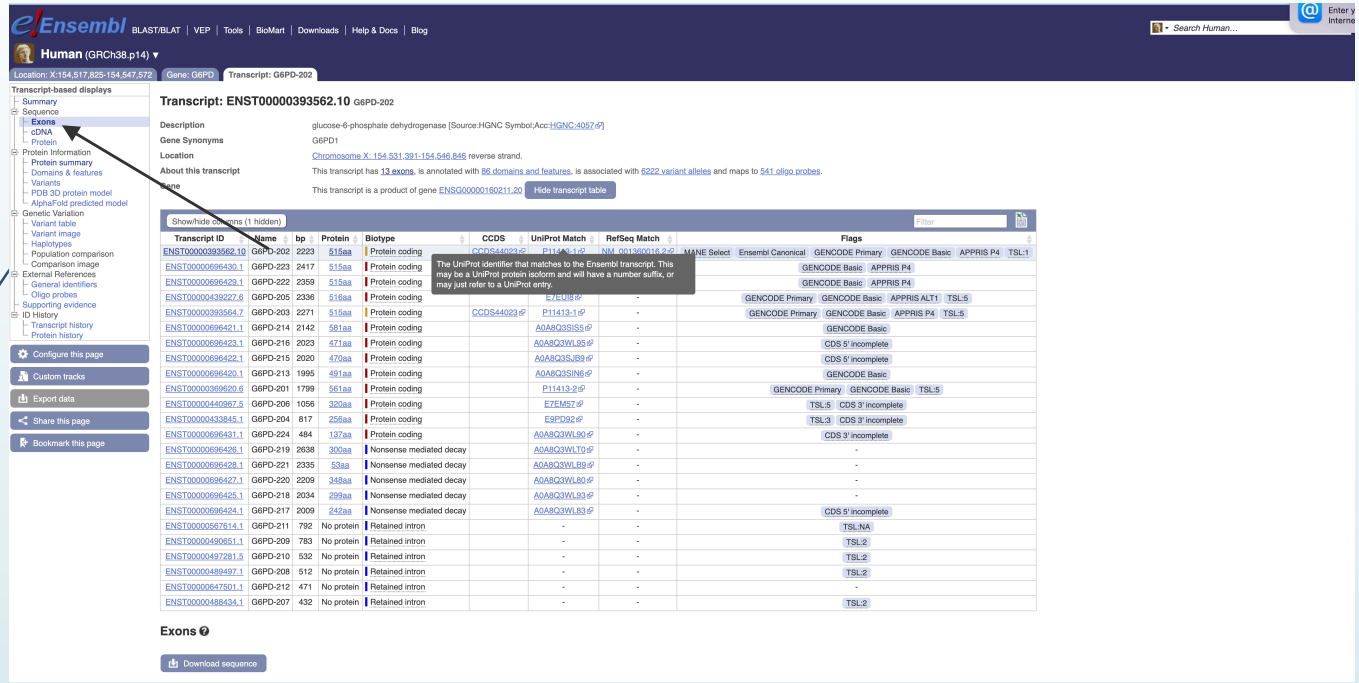

| | |
|---|---|
| **Statistics** | **Exons:** 13, **Coding exons:** 12, **Transcript length:** 2,223 bps, **Translation length:** 515 residues |
| **CCDS** | This transcript is a member of the Human CCDS set: CCDS44023 |
| **Uniprot** | This transcript corresponds to the following Uniprot identifiers: P11413 |
| **Transcript Support Level (TSL)** | TSL:1 |
| **Version** | ENST00000393562.10 |
| **Type** | Protein coding |
| **Annotation Method** | Transcript where the Ensembl genebuild transcript and the Havana manual annotation have the same sequence, for every base pair. See article. |
| **GENCODE basic gene** | This transcript is a member of the Gencode basic gene set. |

La variante **G6PD-202**, in particolare, è caratterizzata dalla presenza di 13 esoni, dei quali 12 codificanti.

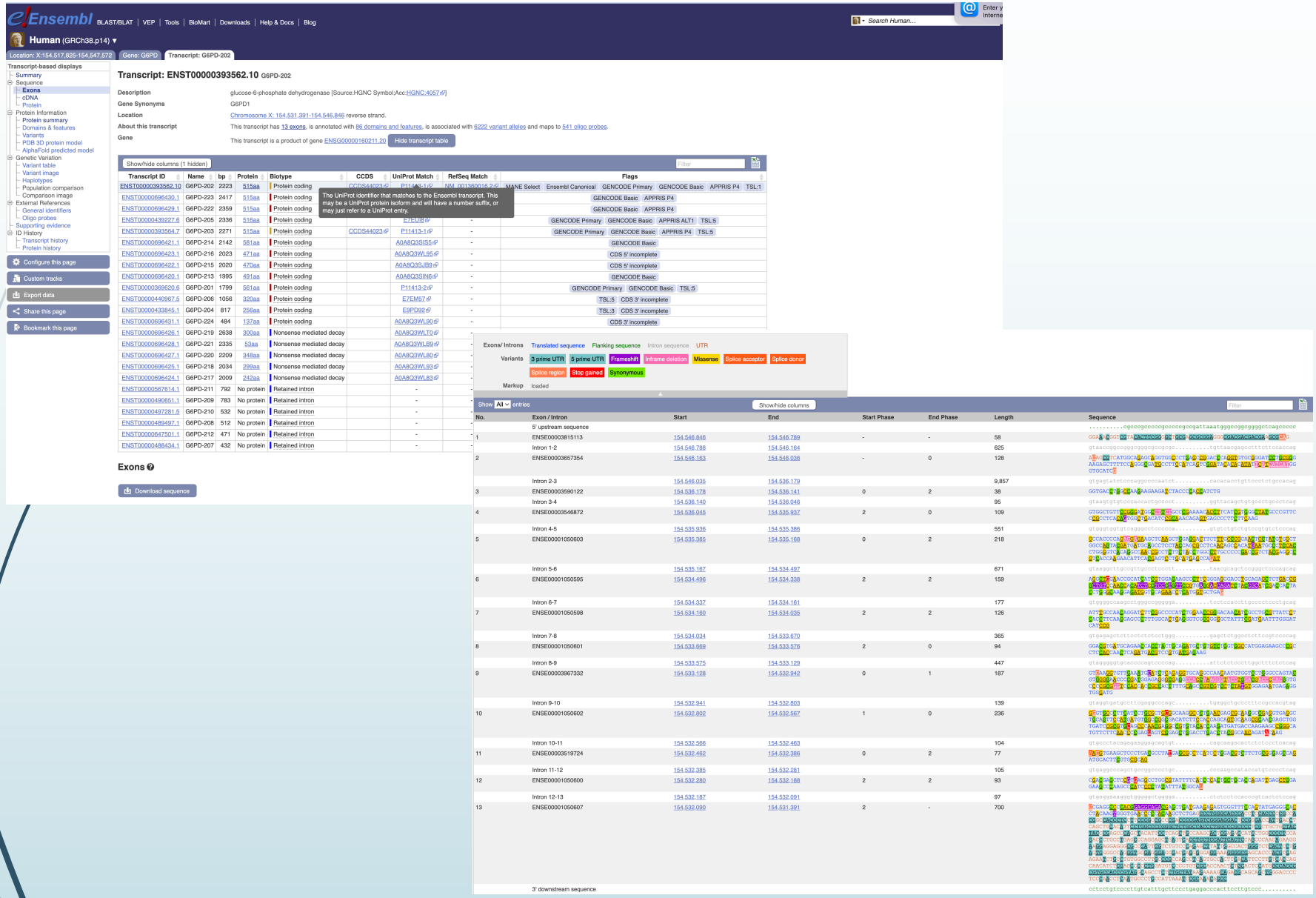
All'interno di uno degli esoni codificanti, l'**esone 6**, è presente con un'alta frequenza una sostituzione nucleotidica (**C563T**) che causa l'insorgenza di una mutazione correlata ad una più rapida degradazione dell'enzima ed all'insorgenza di una grave **anemia emolitica.**

# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Gene/Splice?db=core;g=ENSG00000160211;r=X:154531391-154547572



# Proviamo insieme con il gene G6PD



# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Transcript/Sequence_cDNA?db=core;g=ENSG00000160211;r=X:154531391-154547572;t=ENST00000393562

## GOAL:

**PCR amplification of a region containing position of point mutation in G6PD gene**

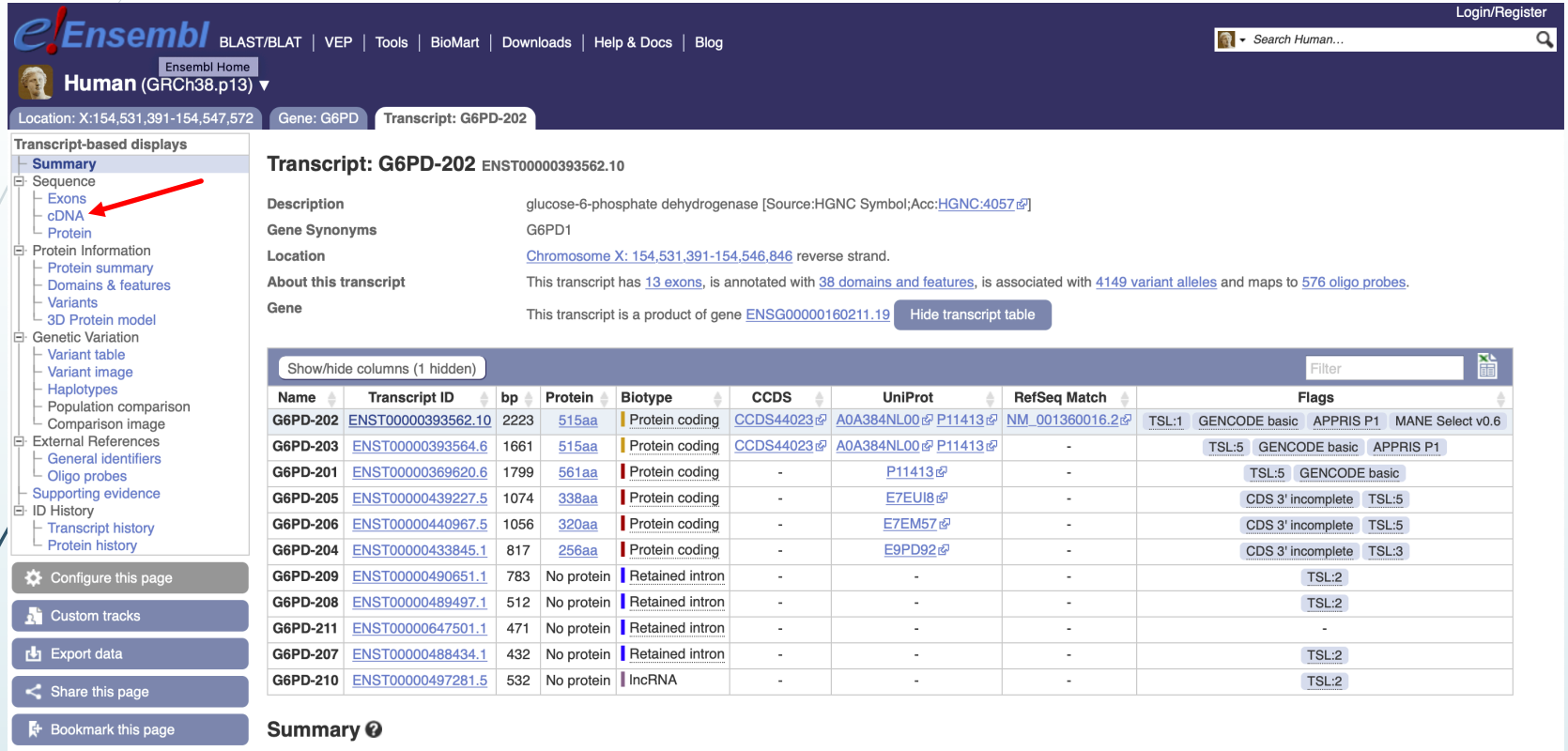
## BIOLOGICAL MATERIAL:

**Genomic DNA from patient specimen**

## STRATEGIA:

1. **Recover exon sequences of G6PD**
2. **Individuate 5' end and 3' end of exon 6**
3. **Recover G6PD gene seqeunces**
4. **Locate nucleotide position of 5' and 3' end of exon 6**
5. **Locate nucleotide position of point mutation**
6. **Use software to identify primer pairs that amplify a region flanking the position of point mutation**
7. **Test PCR primers**

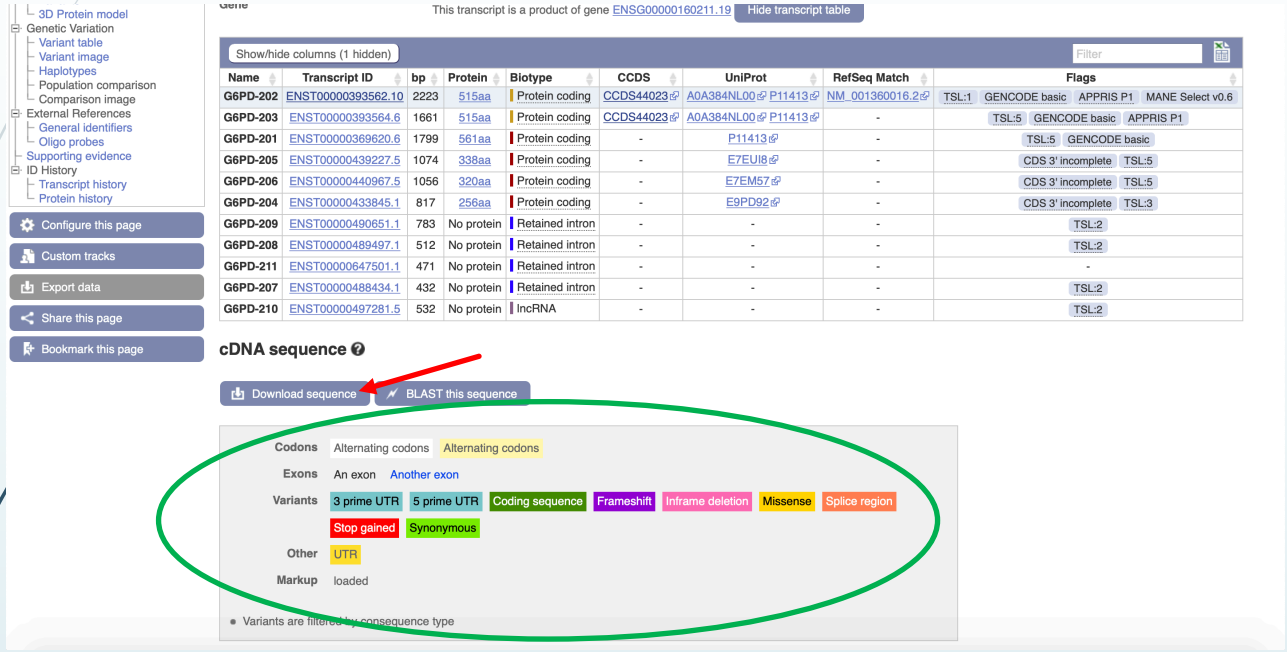
# Proviamo insieme con il gene G6PD




Per andare a ricercare la posizione esatta della sostituzione, sarà necessario analizzare la sequenza del cDNA del G6PD-202, che conterrà al suo interno esclusivamente gli esoni e dunque anche l'esone 6 di nostro interesse.

# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Transcript/Sequence_cDNA?db=core;g=ENSG00000160211;r=X:154531391-154547572;t=ENST00000393562

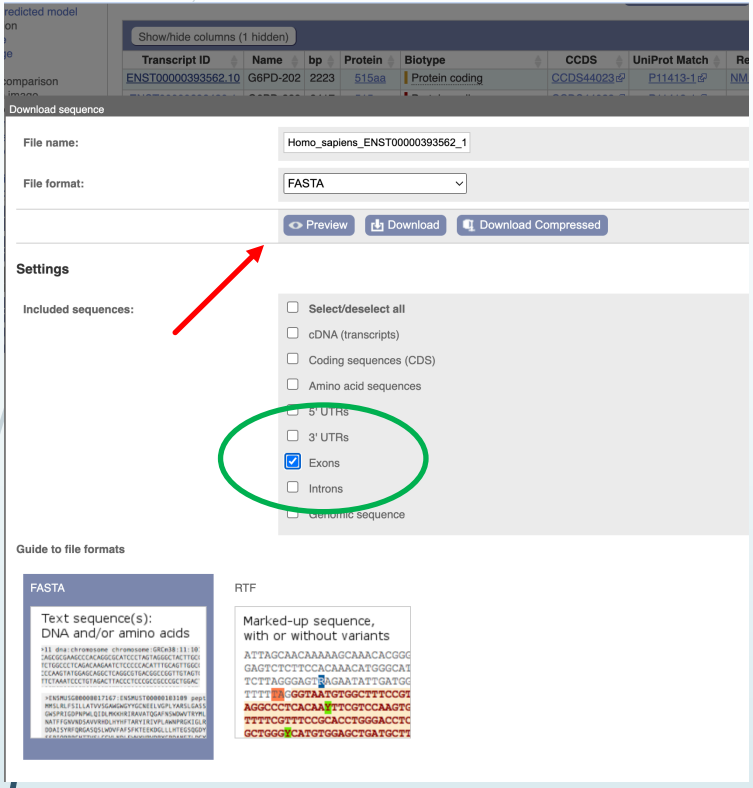



La sequenza del cDNA potrà essere facilmente **scaricata** in una versione facilmente consultabile.

Nella sequenza scaricata saranno presenti una **serie di informazioni**, la cui legenda è riportata nella figura.

# Proviamo insieme con il gene G6PD

https://www.ensembl.org/Homo_sapiens/Transcript/Sequence_cDNA?db=core;g=ENSG00000160211;r=X:154531391-154547572;t=ENST00000393562

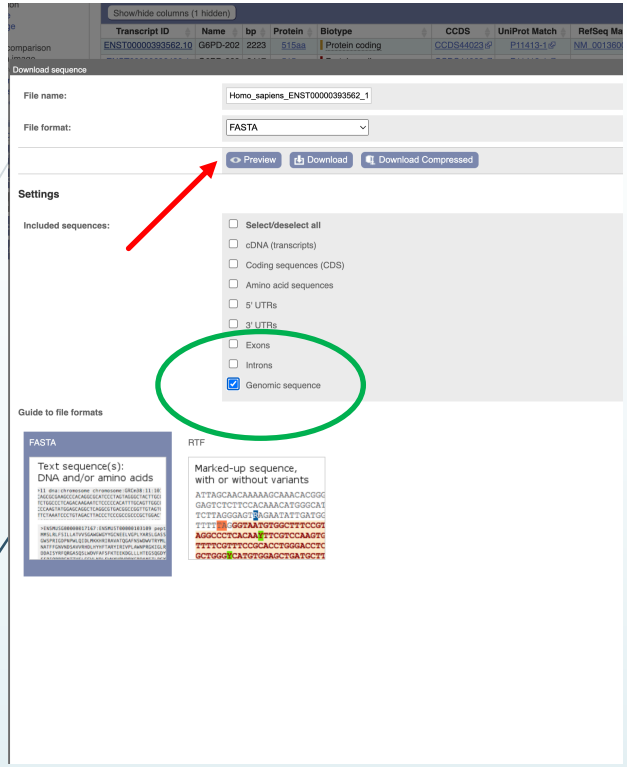


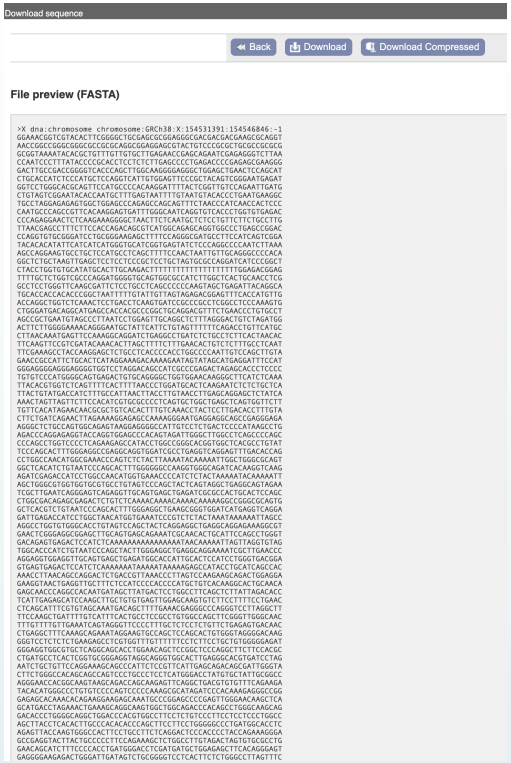



```
>G6PD-202 ENSE00003815113 exon:protein_coding
GGAAACGGTCGTACACTTCGGGGCTGCGAGCGCGGAGGGCGACGACGAAGCGCAG
>G6PD-202 ENSE00003657354 exon:protein_coding
ACAGCGTCATGGCAGAGCAGGTGGCCCTGAGCCGGACCCAGGTGTGCGGGATCCTGCGGG
AAGAGCTTTTCCAGGGCGATGCCTTCCATCAGTCGGATACACACATATTCATCATCATGG
GTGCATCG
>G6PD-202 ENSE00003590122 exon:protein_coding
GGTGACCTGGCCAAGAAGAAGATCTACCCCACCATCTG
>G6PD-202 ENSE00003546872 exon:protein_coding
GTGGCTGTTCCGGGATGGCCTTCTGCCCGAAAACACCTTCATCGTGGGCTATGCCCGTTC
CCGCCTCACAGTGGCTGACATCCGCAAACAGAGTGAGCCCTTCTTCAAG
>G6PD-202 ENSE00001050603 exon:protein_coding
GCCACCCCAGAGGAGAAGCTCAAGCTGGAGGACTTCTTTGCCCGCAACTCCTATGTGGCT
GGCCAGTACGATGATGCAGCCTCCTACCAGCGCCTCAACAGCCACATGAATGCCCTCCAC
CTGGGGTCACAGGCCAACCGCCTCTTCTACCTGGCCTTGCCCCCGACCGTCTACGAGGCC
GTCACCAAGAACATTCACGAGTCCTGCATGAGCCAGAT
>G6PD-202 ENSE00001050595 exon:protein_coding
AGGCTGGAACCGCATCATCGTGGAGAAGCCCTTCGGGAGGGACCTGCAGAGCTCTGACCG
GCTGTCCAACCACATCTCC[C]CTGTTCCGTGAGGACCAGATCTACCGCATCGACCACTA
CCTGGGCAAGGAGATGGTGCAGAACCTCATGGTGCTGAG
>G6PD-202 ENSE00001050598 exon:protein_coding
ATTTGCCAACAGGATCTTCGGCCCCATCTGGAACCGGGACAACATCGCCTGCGTTATCCT
CACCTTCAAGGAGCCCTTTGGCACTGAGGGTCGCGGGGGCTATTTCGATGAATTTGGGAT
CATCCG
>G6PD-202 ENSE00001050601 exon:protein_coding
GGACGTGATGCAGAACCACCTACTGCAGATGCTGTGTCTGGTGGCCATGGAGAAGCCCGC
CTCCACCAACTCAGATGACGTCCGTGATGAGAAG
>G6PD-202 ENSE00003967332 exon:protein_coding
GTCAAGGTGTTGAAATGCATCTCAGAGGTGCAGGCCAACAATGTGGTCCTGGGCCAGTAC
GTGGGGAACCCCGATGGAGAGGGCGAGGCCACCAAAGGGTACCTGGACGACCCCACGGTG
CCCCGCGGGTCCACCACCGCCACTTTTGCAGCCGTCGTCCTCTATGTGGAGAATGAGAGG
TGGGATG
>G6PD-202 ENSE00001050602 exon:protein_coding
GGGTGCCCTTCATCCTGCGCTGCGGCAAGGCCCTGAACGAGCGCAAGGCCGAGGTGAGGC
TGCAGTTCCATGATGTGGCCGGCGACATCTTCCACCAGCAGTGCAAGCGCAACGAGCTGG
TGATCCGCGTGCAGCCCAACGAGGCCGTGTACACCAAGATGATGACCAAGAAGCCGGGCA
TGTTCTTCAACCCCGAGGAGTCGGAGCTGGACCTGACCTACGGCAACAGATACAAG
>G6PD-202 ENSE00003519724 exon:protein_coding
AACGTGAAGCTCCCTGACGCCTATGAGCGCCTCATCCTGGACGTCTTCTGCGGGAGCCAG
ATGCACTTCGTGCGCAG
>G6PD-202 ENSE00001050600 exon:protein_coding
CGACGAGCTCCGTGAGGCCTGGCGTATTTTCACCCCACTGCTGCACCAGATTGAGCTGGA
GAAGCCCAAGCCCATCCCCTATATTTATGGCAG
>G6PD-202 ENSE00001050607 exon:protein_coding
CCGAGGCCCCCACGGAGGCAGACGAGCTGATGAAGAGAGTGGGTTTCCAGTATGAGGGCAC
CTACAAGTGGGTGAACCCCCCACAAGCTCTGAGCCCTGGGCACCCACCTCCACCCCCGCCA
CGGCCACCCTCCTTCCCGCCGCCCGACCCCGAGTCGGGAGGACTCCGGGACCATTGACCT
CAGCTGCACATTCCTGGCCCCCGGGCTCTGGCCACCCTGGCCCGCCCCTCGCTGCTGCTAC
TACCCGAGCCCAGCTACATTCCTCAGCTGCCAAGCACTCGAGACCATCCTGGCCCCTCCA
GACCCTGCCTGAGCCCAGGAGCTGAGTCACCTCCTCCACTCACTCCAGCCCAACAGAAGG
AAGGAGGAGGGCGCCCATTCGTCTGTCCCAGAGCTTATTGGCCACTGGGTCTCACTCCTG
AGTGGGGCCAGGGTGGGAGGGAGGGACGAGGGGGAGGAAAGGGGCGAGCACCCACGTGAG
AGAATCTGCCTGTGGCCTTGCCCGCCAGCCTCAGTGCCACTTGACATTCCTTGTCACCAG
CAACATCTCGAGCCCCCTGGATGTCCCCTGTCCCACCAACTCTGCACTCCATGGCCACCC
CGTGCCACCCGTAGGCAGCCTCTCTGCTATAAGAAAAGCAGACGCAGCAGCTGGGACCCC
TCCCAACCTCAATGCCCTGCCATTAAATCCGCAAACAGCC
```

→ Locate exon 6
→ Locate position of point mutation

# Proviamo insieme con il gene G6PD




→ Recover gene sequence of G6PD and download sequence (for example into word)

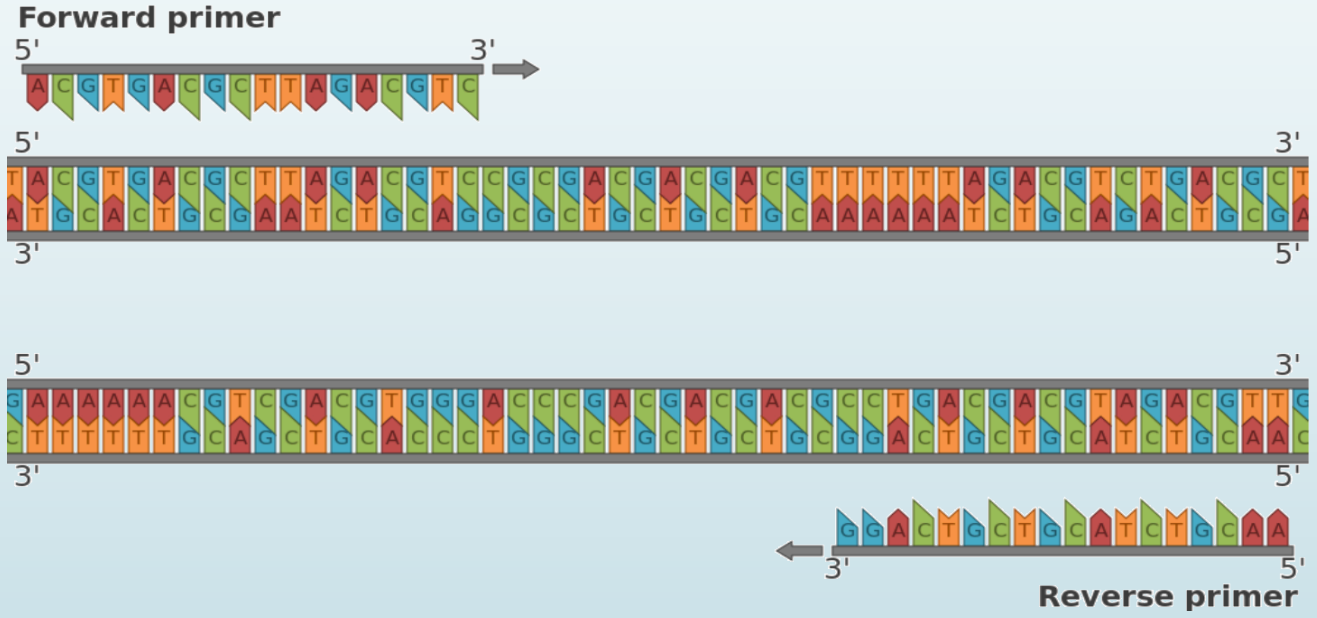
USE EXON 6 SEQUENCE TO:
(word file; functions «find»; «count characters»)
→ Locate nucleotide position of start-end exon 6: 12349 - 12509
→ Locate nucleotide position of point mutation: 12428

# Che cosa è un primer per PCR?

Un primer è un corto filamento oligonucleotidico che viene usato per dare l'innesco alla reazione di amplificazione.

La sua presenza è fondamentale perché la DNA polimerasi può estendere il filamento neosintetizzato solo aggiungendo nucleotidi ad un filamento di DNA già esistente.

In una reazione di PCR saranno necessari un primer Forward che ha una sequenza complementare al filamento di DNA 3'-5' ed un primer Reverse, complementare alla sequenza 5'-3'



# Come si disegna un primer per PCR?

Una volta evidenziata la sequenza che ci interessa sarà fondamentale affrontare diversi steps:

1. Controllare se in letteratura o su database pubblici sono presenti primers che possano amplificare la nostra sequenza. Dovremo controllare su Pubmed pubblicazioni già esistenti, o banche online come primerbank (https://pga.mgh.harvard.edu/primerbank/) dove vengono inserite coppie di primers utilizzabili per amplificare il nostro frammento
2. Disegnare dei primers che abbiano delle specifiche caratteristiche
3. Controllare la specificità dei suddetti primers
4. Validarne l'efficienza

# Caratteristiche di una sequenza target di PCR?

Nel caso in cui, dopo aver controllato in letteratura, non esistano dei primers già disponibili per il frammento che ci interessa amplificare, sarà necessario che disegniamo in autonomia dei primers.

I primers devono essere disegnati in maniera molto accurata, tenendo conto di diverse caratteristiche:

- **La sequenza target deve essere unica** → nel DNA da amplificare non deve essere presente nessuna ulteriore sequenza che possa essere legata dai nostri primers. Ciò eviterà che i primers si possano legare in maniera aspecifica, causando l'amplificazione di un prodotto indesiderato. Sarà necessario controllare la specificità usando un tool online come Basic Local Alignment Search Tool (BLAST)

# Caratteristiche di una sequenza target di PCR?

Nel caso in cui, dopo aver controllato in letteratura, non esistano dei primers già disponibili per il frammento che ci interessa amplificare, sarà necessario che disegniamo in autonomia dei primers.

I primers devono essere disegnati in maniera molto accurata, tenendo conto di diverse caratteristiche:

- **La sequenza target deve essere unica** → nel DNA da amplificare non deve essere presente nessuna ulteriore sequenza che possa essere legata dai nostri primers. Ciò eviterà che i primers si possano legare in maniera aspecifica, causando l'amplificazione di un prodotto indesiderato. Sarà necessario controllare la specificità usando un tool online come Basic Local Alignment Search Tool (BLAST)

- **La sequenza target, unica, deve essere lunga tra 200 e 800 bp (~500)** → Frammenti più corti vengono amplificati più efficientemente di quelli lunghi perché, dopo una certa lunghezza, la *Taq* polimerasi ha un calo dell'efficienza. D'altro canto, frammenti troppo corti potrebbero non essere distinguibili da aspecifici (dimeri di primers)

# Caratteristiche di un primer per PCR?

- **I primers devono avere una lunghezza tra 18-25 bp** → più il primer è lungo e più è alta la probabilità che sia unico. Inoltre una lunghezza maggiore assicura una specificità più alta, data dall'aumento della temperatura di melting, cioè la temperatura alla quale il 50% delle molecole di DNA sono denaturate.

  Primers più lunghi di 30 nt sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosidetti dimeri di primers.

# Caratteristiche di un primer per PCR?

- **I primers devono avere una lunghezza tra 18-25 nt** → più il primer è lungo e più è alta la probabilità che sia unico. Inoltre una lunghezza maggiore assicura una specificità più alta, data dall'aumento della temperatura di melting, cioè la temperatura alla quale il 50% delle molecole di DNA sono denaturate.

  Primers più lunghi di 30 nt sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosidetti dimeri di primers.

- **I primers devono contenere una percentuale di G/C di circa 40-60%** → ciò consentirà di avere una temperatura di melting abbastanza alta da poter essere stringente ed aumentare la specificità e la stabilità della reazione.

# Caratteristiche di un primer per PCR?

- **I primers devono avere una lunghezza tra 18-25 bp** → più il primer è lungo e più è alta la probabilità che sia unico. Inoltre una lunghezza maggiore assicura una specificità più alta, data dall'aumento della temperatura di melting, cioè la temperatura alla quale il 50% delle molecole di DNA sono denaturate.

  Primers più lunghi di 30 bp sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosidetti dimeri di primers.

- **I primers devono contenere una percentuale di G/C di circa 40-60%** → ciò consentirà di avere una temperatura di melting abbastanza alta da poter essere stringente ed aumentare la specificità e la stabilità della reazione.

- **I primers NON devono contenere sequenze ripetute** → anche in questo caso si potrebbe avere un cattivo appaiamento dei primers alla sequenza target o la formazione di dimeri di primers

# Caratteristiche di un primer per PCR?

- **I primers devono avere una lunghezza tra 18-25 bp** → più il primer è lungo e più è alta la probabilità che sia unico. Inoltre una lunghezza maggiore assicura una specificità più alta, data dall'aumento della temperatura di annealing/melting, cioè la temperatura alla quale il 50% delle molecole di DNA sono denaturate.

  Primers più lunghi di 30 bp sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosidetti dimeri di primers.

- **I primers devono contenere una percentuale di G/C di circa 40-60%** → ciò consentirà di avere una temperatura di melting abbastanza alta da poter essere stringente ed aumentare la specificità e la stabilità della reazione.

- **I primers NON devono contenere sequenze ripetute** → anche in questo caso si potrebbe avere un cattivo appaiamento dei primers alla sequenza target o la formazione di dimeri di primers

- **I primers devono avere una temperatura di annealing/melting di circa 60°C e deve essere simile tra i due (max 3°C differenza)**→ una temperatura troppo alta produrrebbe un'insufficiente appaiamento tra i primers ed il DNA templato, causando una bassa efficienza di amplificazione. D'altro canto, una temperatura troppo bassa potrebbe causare un appaiamento non specifico tra primers e DNA, causando l'amplificazione di DNA diverso da quello desiderato.

# Caratteristiche di un primer per PCR?

- **I primers devono avere una lunghezza tra 18-25 bp** → più il primer è lungo e più è alta la probabilità che sia unico. Inoltre una lunghezza maggiore assicura una specificità più alta, data dall'aumento della temperatura di annealing/melting, cioè la temperatura alla quale il 50% delle molecole di DNA sono denaturate.

  Primers più lunghi di 30 bp sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosidetti dimeri di primers.

- **I primers devono contenere una percentuale di G/C di circa 40-60%** → ciò consentirà di avere una temperatura di melting abbastanza alta da poter essere stringente ed aumentare la specificità e la stabilità della reazione.

- **I primers NON devono contenere sequenze ripetute** → anche in questo caso si potrebbe avere un cattivo appaiamento dei primers alla sequenza target o la formazione di dimeri di primers

- **I primers devono avere una temperatura di annealing/melting di circa 60°C e deve essere simile tra i due (max 3°C differenza)**→ una temperatura troppo alta produrrebbe un'insufficiente appaiamento tra i primers ed il DNA templato, causando una bassa efficienza di amplificazione. D'altro canto, una temperatura troppo bassa potrebbe causare un appaiamento non specifico tra primers e DNA, causando l'amplificazione di DNA diverso da quello desiderato.

- **I primers devono contenere una G o una C tra le ultime 5 basi presenti al 3' terminale** → data la maggiore forza del legame G/C, ciò permette un migliore appaiamento dell'estremità del primer al DNA stampo. Idealmente in ultima posizione.

# Proviamo insieme con il gene G6PD

## 1. Primers da literatura

Al-Jaouni *et al. BMC Research Notes* 2011, **4**:436
http://www.biomedcentral.com/1756-0500/4/436

BMC Research Notes

RESEARCH ARTICLE Open Access

### Molecular characterization of glucose-6-phosphate dehydrogenase deficiency in Jeddah, Kingdom of Saudi Arabia

Soad K Al-Jaouni[1*], Jummanah Jarullah[1], Essam Azhar[2] and Kamran Moradkhani[3,4]

**Table 1 Primers used for identification of glucose-6-phosphate dehydrogenase defects[a]**

| Analysis | Analyzed Region | Forward 5′→3′ | Reverse 5′→3′ |
|---|---|---|---|
| RFLP Analysis (Tm = 61°C) | Exon 3 | GTGGAGGATGATGTATGTAGGT | AGGGCAGGGCACAGCTGTAA |
| | Exon 4 | TACAGTCGTGCCCTGCCCT | CCGAAGCTGGCCATGCTGG |
| | Exon 5 | CTGTGTGTGTGTCTGTCTGTC | GGAGGGCAACGGCAAGCCTT |
| | Exon 6 | GCAGCTGTGATCCTCACTCC | GCAAGGTGGAGGAACTGACC |

Abbiamo controllato se in letteratura fossero presenti pubblicazioni riguardo la sequenza che ci interessa e contiene la mutazione C563T.

Abbiamo trovato la presenza di una pubblicazione che si è occupata di caratterizzare diverse mutazioni del gene G6PD in Arabia Saudita.

Esiste una tabella che contiene una lista di primers, tra i quali **una coppia di primers** che amplifica una regione dell'esone 6 e che potrebbe interessarci.

Materiale biologico:
mRNA da paziente → cDNA → PCR

# Proviamo insieme con il gene G6PD

CAGGCCGCAGGGTGGCAGCCTTGCTCTGCGAATGCAGCATGGCCCGCGCTGGGTGGTTTC
CCAACCCAGCCAGAGGCTCTTGTCCTCTGGCTGGTTTTGAATGCGGGGGTAGTAAAGCAA
AGGTCCTCTTCTCATTTTCAAAACCAATGAGGAAGCCATGGCTTGGATGCCTCCTCCCCC
TGCTCCCCTACAGGCCTTCAGGCCACTCAGACCCACCGGGGACCCAGCATGAGGCAGGAG
GGGAACGGGCCCCCGGCAGCATGCCAGCAATGCCACCCTGGCACCCAGGGTGGGAAGGCT
TCCCGGAAGGTGTTGAGCCAGAGGGTCATCTGGGAACACAAGGCACGGGAGGTGGCCACG
GGGGCGAGGAGGTTCTGGCCTCTACTCCCCTGGGAGGGCGTCTGAATGAT**GCAGCTCTGA**
**TCCTCACTCC**CCGAAGAGGGGTTCAAGGGGGTAACGCAGCTCCGGGCTCCCAGCAGAGGC
TGGAACCGCATCATCGTGGAGAAGCCCTTCGGGAGGGACCTGCAGAGCTCTGACCGGCTG
TCCAACCACA**TCTCC**TCCCTGTTCCGTGAGGACCAGATCTACCGCATCGACCACTACCTG
GGCAAGGAGATGGTGCAGAACCTCATGGTGCTGAGGTGGGGCCAAGCCTGGGCCGGGGGA
CCAGGGTGGGGGTGGTACTCAGGAGCCTCACCTGGCCCACTGCCTCCCCGAGGACGAATT
CCTCCAGAACTCAGACAAGGGTGACCCCTCACATGTGGCCCCTGCACCACAGAGGCCCAA
**GGTCAGTTCCTCCACCTTGC**CCCTCCCTGCAGATTTGCCAACAGGATCTTCGGCCCCATC
TGGAACCGGGACAACATCGCCTGCGTTATCCTCACCTTCAAGGAGCCCTTTGGCACTGAG
GGTCGCGGGGGCTATTTCGATGAATTTGGGATCATCCGGTGAGAGCTCTTCCTCTCTCCT
GGGAGGCTGGCACAGGGTGGCAGAGCCAGTCACCCTGCAGGGCTACTCTTCCCTATCTTG
GGGGAGCTCCTCCTCACCCTGCAGTTCAAAACCTAAGTGTCTGAGCTATCAGACCGGGCT
GGAAAGGGCTGGACCCCTACACAGCCAAGCACCCCACGGTTTTATGATTCAGTGATAGCA
TCACCATGTCCTTCCTTGATTTAAGGGGACCTGGAAGACAAGGGGGATCAGGAAGTGAGT
CTTGCAGCTTGTCACTAGGAAGCCTTGTTTGGGGTCCCCATGCCCTTGAACCAGGTGAAC

Abbiamo controllato che i primers indicati fossero presenti nella nostra sequenza di interesse e che amplificassero un frammento di circa 500 bp con all'interno la mutazione.

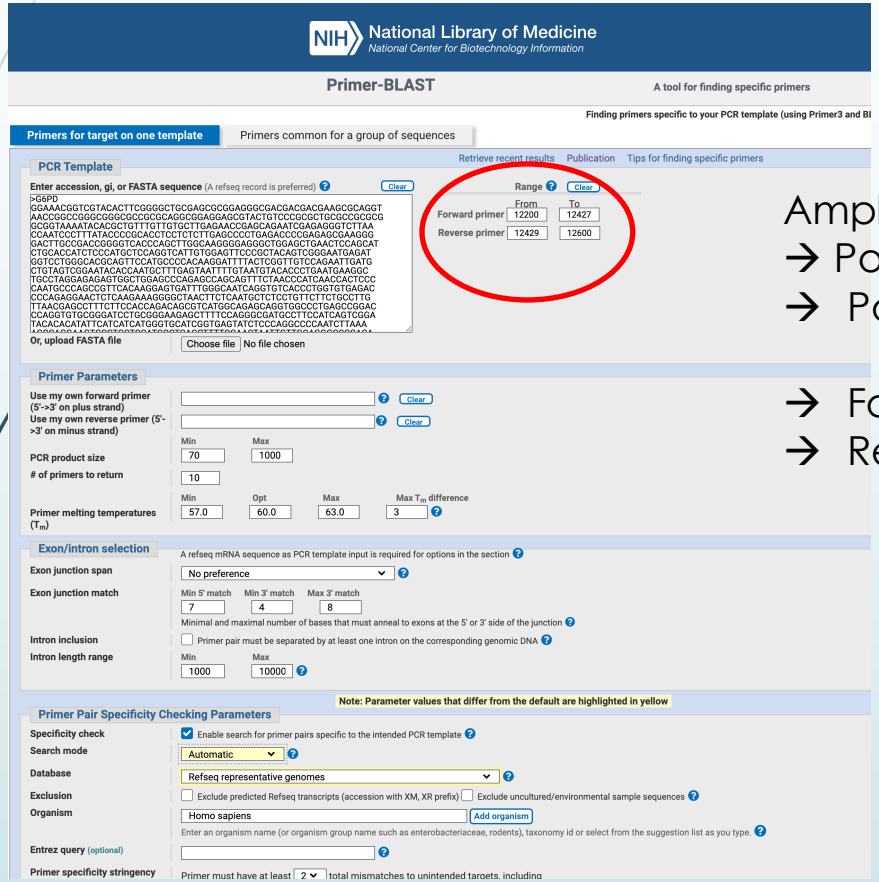
Questi due primers

- hanno una temperatura di melting di ≃60°C
- Contengono G/C all'estremità 3' terminale

# Proviamo insieme con il gene G6PD
# 2. Ex-novo primer design

Ma se non fossero stati presenti in letteratura come avremmo fatto??

Avremmo dovuto selezionare la sequenza di interesse ed utilizzare un tool online per il disegno dei primers, come https://www.ncbi.nlm.nih.gov/tools/primer-blast/ .




Amplify around position of mutation
→ Position of start-end exon 6: 12349 - 12509
→ Position of point mutation: 12428

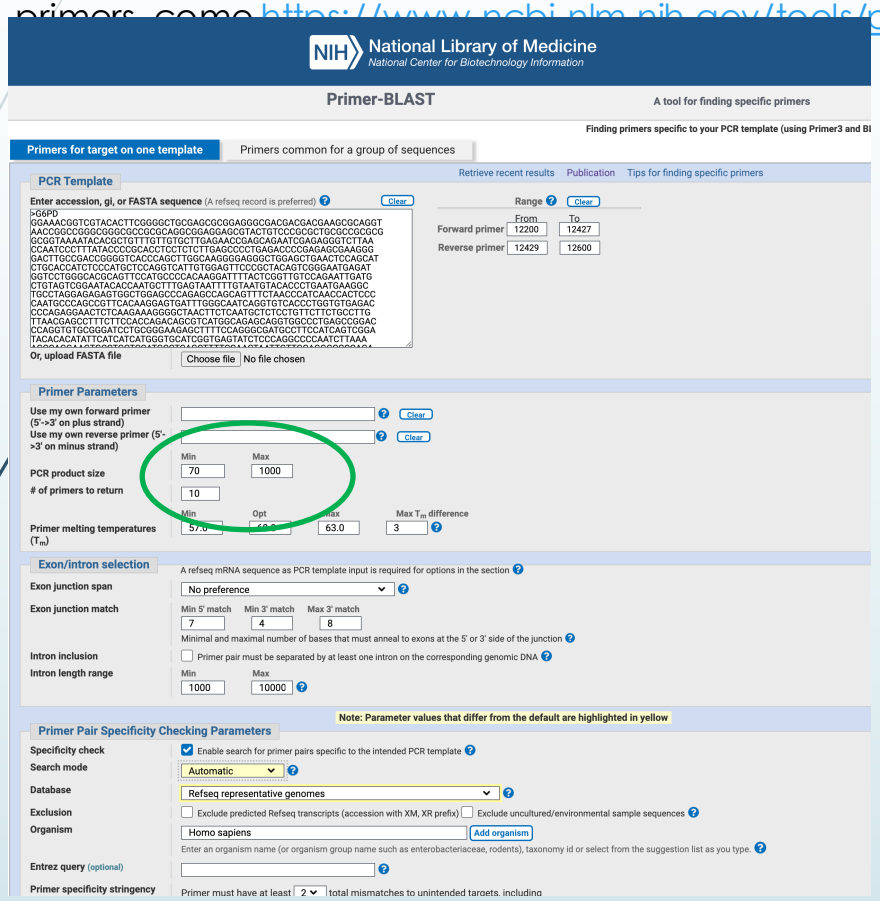
→ Forward Primer position 12200 – 12427
→ Reverse Primer position 12429 - 12600

**Incolliamo** la sequenza nucleotidica che abbiamo selezionato (che deve contenere il sito di mutazione) ed iniziamo a selezionare le caratteristiche:

# Proviamo insieme con il gene G6PD

Ma se non fossero stati presenti in letteratura come avremmo fatto??

Avremmo dovuto selezionare la sequenza di interesse ed utilizzare un tool online per il disegno dei primers, come https://www.ncbi.nlm.nih.gov/tools/primer-blast/ .



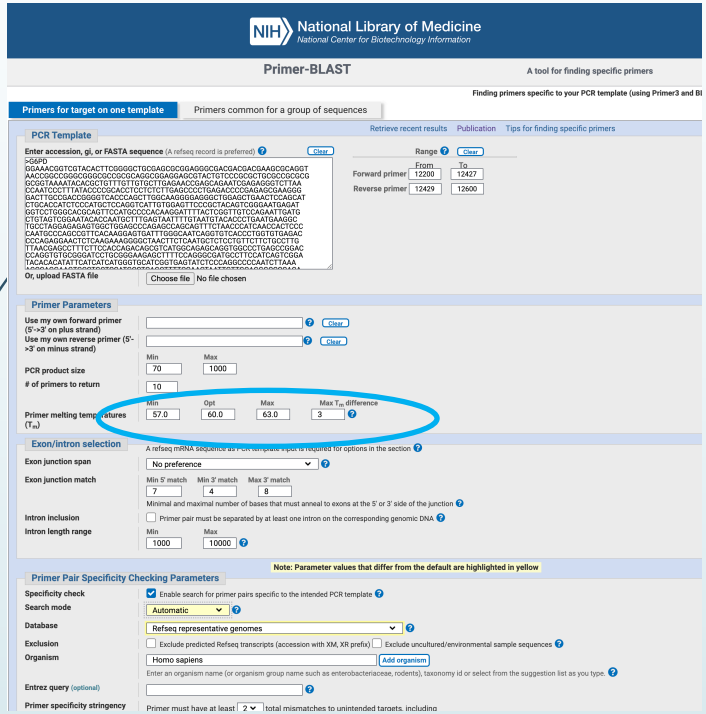

**Incolliamo** la sequenza nucleotidica che abbiamo selezionato (che deve contenere il sito di mutazione) ed iniziamo a selezionare le caratteristiche:

- **PCR product size** → min 70 – max 1000

# Proviamo insieme con il gene G6PD

Ma se non fossero stati presenti in letteratura come avremmo fatto??

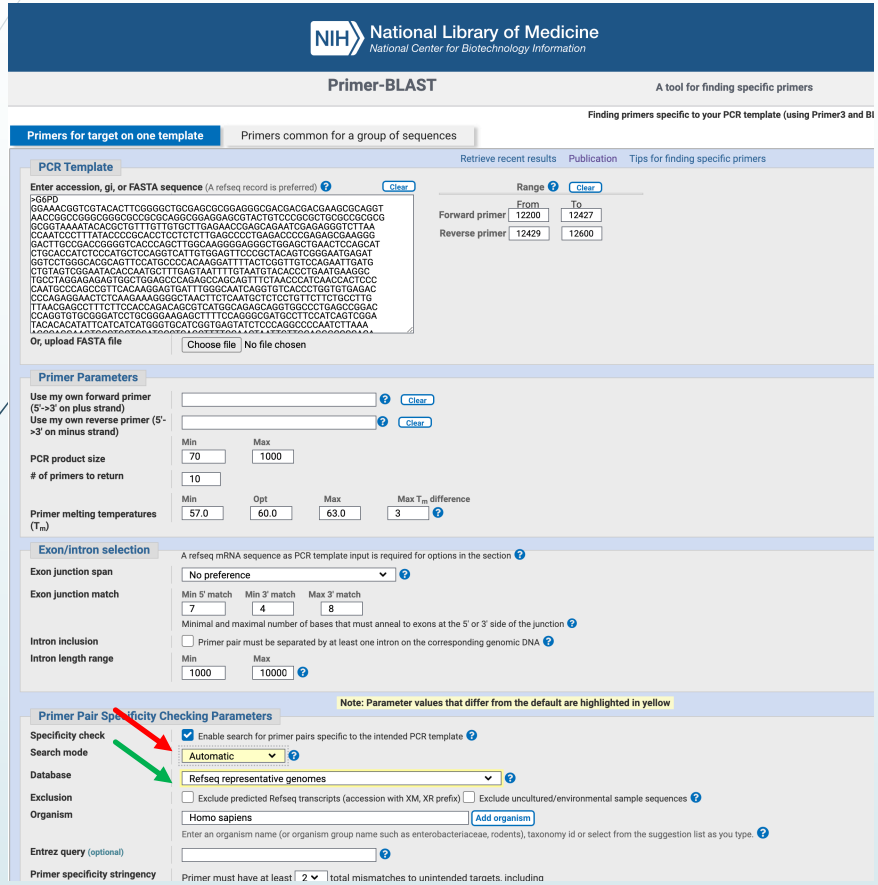
Avremmo dovuto selezionare la sequenza di interesse ed utilizzare un tool online per il disegno dei primers, come https://www.ncbi.nlm.nih.gov/tools/primer-blast/ .




**Incolliamo** la sequenza nucleotidica che abbiamo selezionato (che deve contenere il sito di mutazione) ed iniziamo a selezionare le caratteristiche; seleziona zona da amplificare:

- **PCR product size** → min 100 – max 900
- **Primer melting temperatures** → possiamo lasciare i valori indicati

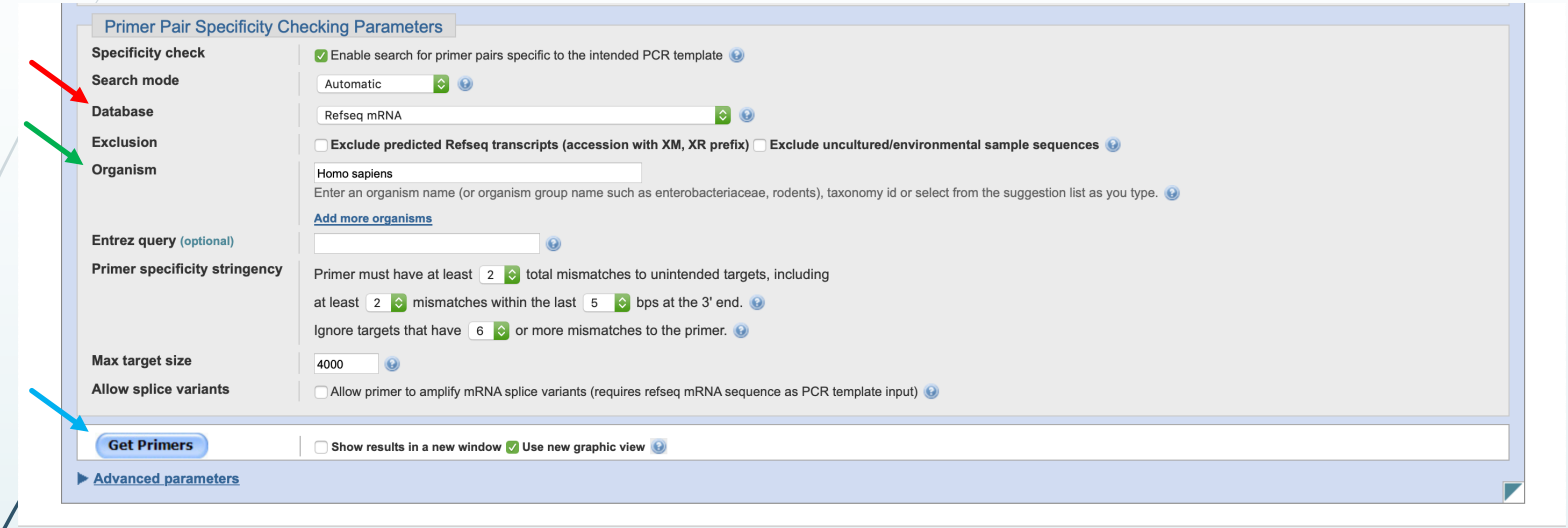
# Proviamo insieme con il gene G6PD




- Selezioniamo il **database di riferimento** (mRNA – genomic – custom…)
- Selezioniamo **l'organismo modello** (human – mouse – rat…)

Il resto delle caratteristiche, a meno che non abbiamo particolari esigenze può restare invariato.

# Proviamo insieme con il gene G6PD




- Selezioniamo il **database di riferimento** (mRNA – genomic – custom…)
- Selezioniamo **l'organismo modello** (human – mouse – rat…)

Il resto delle caratteristiche, a meno che non abbiamo particolari esigenze, può restare invariato.

**GET PRIMERS** ed attendiamo!!

**Template: 12200 - 12600**

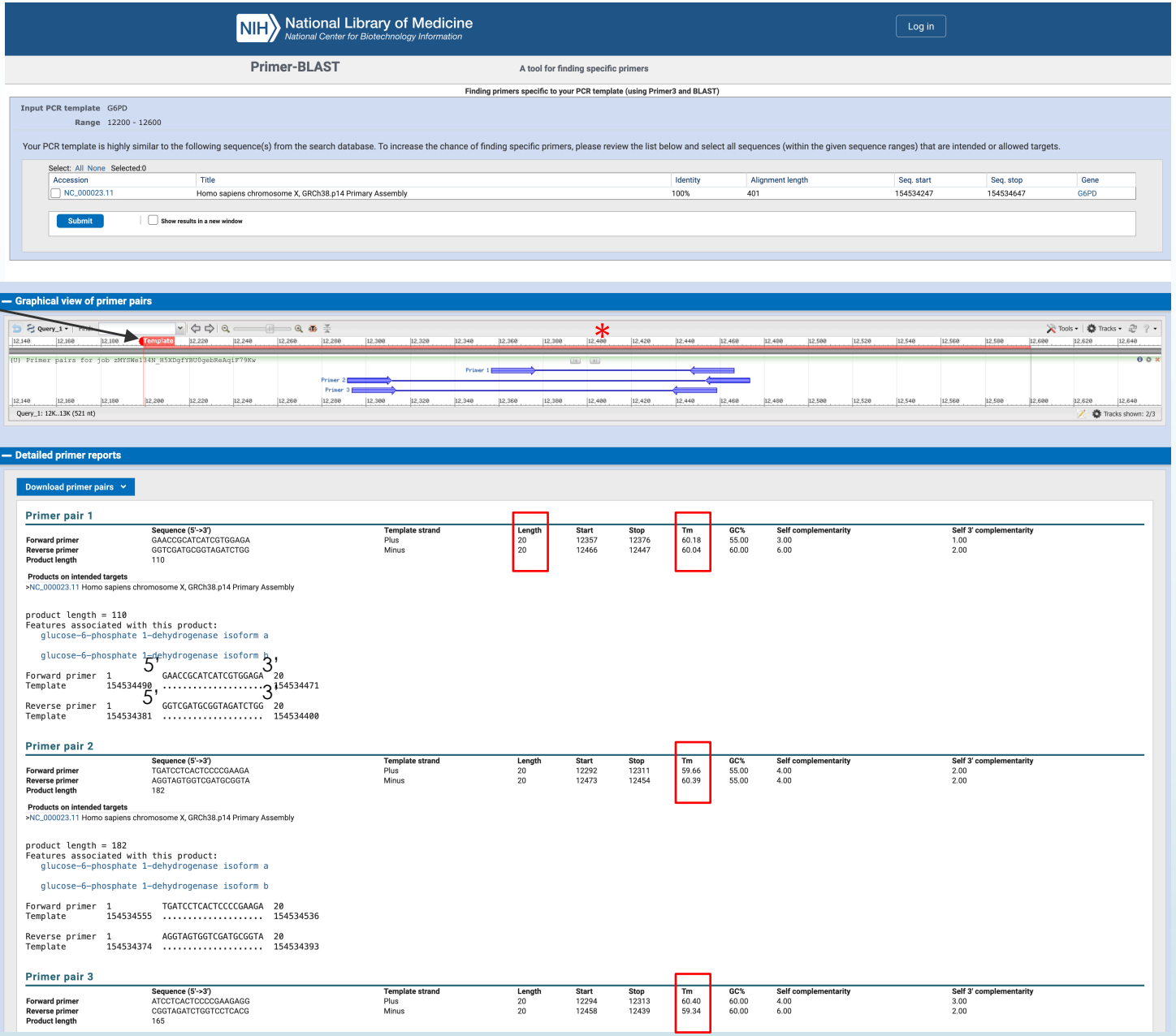
→Position of start-end exon 6: 12349 - 12509

→Position of point mutation: 12428 = *

Primer pair suggestion 1

Primer pair suggestion 2

Primer pair suggestion 3



NIH National Library of Medicine
National Center for Biotechnology Information

Log in

**Primer-BLAST** A tool for finding specific primers

Finding primers specific to your PCR template (using Primer3 and BLAST)

Input PCR template G6PD
Range 12200 - 12600

Your PCR template is highly similar to the following sequence(s) from the search database. To increase the chance of finding specific primers, please review the list below and select all sequences (within the given sequence ranges) that are intended or allowed targets.

Select: All None Selected:0

| Accession | Title | Identity | Alignment length | Seq. start | Seq. stop | Gene |
|---|---|---|---|---|---|---|
| NC_000023.11 | Homo sapiens chromosome X, GRCh38.p14 Primary Assembly | 100% | 401 | 154534247 | 154534647 | G6PD |

Submit | Show results in a new window

**— Graphical view of primer pairs**

Template

Primer 1
Primer 2
Primer 3

**— Detailed primer reports**

Download primer pairs

**Primer pair 1**

| | Sequence (5'->3') | Template strand | Length | Start | Stop | Tm | GC% | Self complementarity | Self 3' complementarity |
|---|---|---|---|---|---|---|---|---|---|
| Forward primer | GAACCGCATCATCGTGGAGA | Plus | 20 | 12357 | 12376 | 60.18 | 55.00 | 3.00 | 1.00 |
| Reverse primer | GGTCGATGCGGTAGATCTGG | Minus | 20 | 12466 | 12447 | 60.04 | 60.00 | 6.00 | 2.00 |
| Product length | 110 | | | | | | | | |

Products on intended targets
>NC_000023.11 Homo sapiens chromosome X, GRCh38.p14 Primary Assembly

product length = 110
Features associated with this product:
glucose-6-phosphate 1-dehydrogenase isoform a
glucose-6-phosphate 1-dehydrogenase isoform b

```
                                  5'                 3'
Forward primer  1          GAACCGCATCATCGTGGAGA  20
Template        154534490  ....................  154534471
                                  5'                 3'
Reverse primer  1          GGTCGATGCGGTAGATCTGG  20
Template        154534381  ....................  154534400
```

**Primer pair 2**

| | Sequence (5'->3') | Template strand | Length | Start | Stop | Tm | GC% | Self complementarity | Self 3' complementarity |
|---|---|---|---|---|---|---|---|---|---|
| Forward primer | TGATCCTCACTCCCCGAAGA | Plus | 20 | 12292 | 12311 | 59.66 | 55.00 | 4.00 | 2.00 |
| Reverse primer | AGGTAGTGGTCGATGCGGTA | Minus | 20 | 12473 | 12454 | 60.39 | 55.00 | 4.00 | 2.00 |
| Product length | 182 | | | | | | | | |

Products on intended targets
>NC_000023.11 Homo sapiens chromosome X, GRCh38.p14 Primary Assembly

product length = 182
Features associated with this product:
glucose-6-phosphate 1-dehydrogenase isoform a
glucose-6-phosphate 1-dehydrogenase isoform b

```
Forward primer  1          TGATCCTCACTCCCCGAAGA  20
Template        154534555  ....................  154534536

Reverse primer  1          AGGTAGTGGTCGATGCGGTA  20
Template        154534374  ....................  154534393
```

**Primer pair 3**

| | Sequence (5'->3') | Template strand | Length | Start | Stop | Tm | GC% | Self complementarity | Self 3' complementarity |
|---|---|---|---|---|---|---|---|---|---|
| Forward primer | ATCCTCACTCCCCGAAGAGG | Plus | 20 | 12294 | 12313 | 60.40 | 60.00 | 4.00 | 3.00 |
| Reverse primer | CGGTAGATCTGGTCCTCACG | Minus | 20 | 12458 | 12439 | 59.34 | 60.00 | 6.00 | 2.00 |
| Product length | 165 | | | | | | | | |

# Proviamo insieme con il gene G6PD

**NEXT STEP VALIDATION IN THE LABORATORY**

1. TEST PCR PRIMERS IN PCR REACTION
2. CONTROL LENGTH OF AMPLIFED PRODUCT (Agarose gel; legth OK?, are there alternative PCR products)
3. DNA sequence the PCR product (directly on PCR product or PCR procut cloned into vector)

# Esercizio – HOME WORK

## HAVE READY PRIMERs ON 11.11.2025

La fibrosi cistica è una malattia grave, presente dalla nascita in quanto dovuta ad una **mutazione del gene CFTR**.

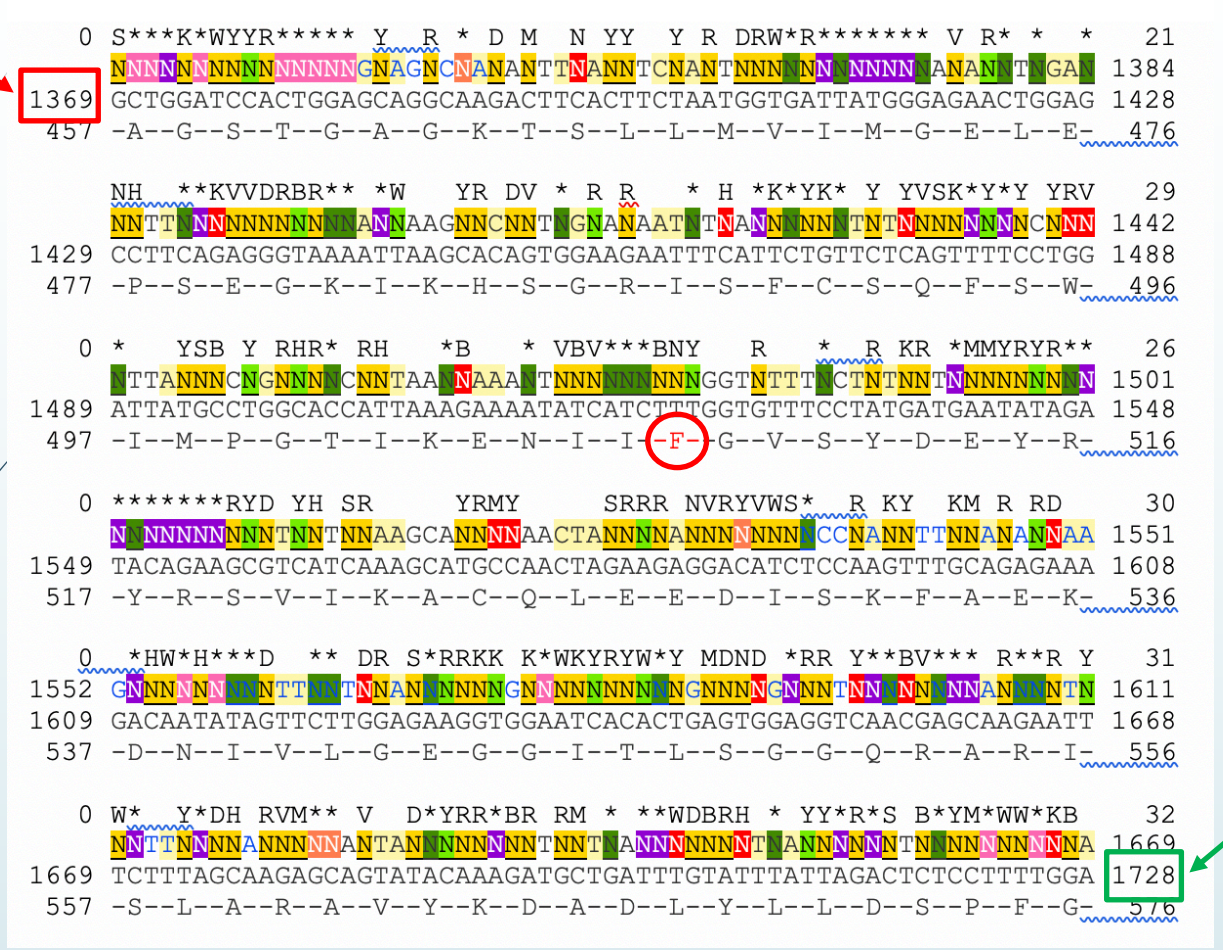
Il regolatore della conduttanza transmembrana della fibrosi cistica (**CFTR** in inglese **C**ystic **F**ibrosis **T**ransmembrane conductance **R**egulator) è una proteina di membrana che svolge la funzione di canale ionico per lo ione cloruro, presente nelle cellule dei vertebrati.

La mutazione più comune è una delezione di una tripletta che causa la perdita di una fenilalanina in posizione 508 (**ΔF508**) con conseguente misfolding della proteina CFTR

- Rintraccia su Ensembl la posizione del gene (cromosoma, posizione dei nucleotidi)
- Scarica le sequenze codificanti (esoni) + genomica
- **Rintraccia la mutazione di interesse (ΔF508) sulla sequenza scaricata**
- **Disegna una coppia di primers per amplificare una regione che includa la mutazione ΔF508**

Materiale: DNA genomica → PCR

# Esercizio – disegna oligo per CFTR (ΔF508)




La F segnata nel cerchio rosso è la fenilalanina che risulta essere mutata nella fibrosi cistica.

Procedendo nel design della coppia di primers, seleziono i nucleotidi che vanno dal **1369** al **1728**…

# Esercizio – disegna oligo per CFTR (**ΔF508**)

**Follow exampe from G6PD**

1. Get **transcript information** and localize the exon that contains the CFTR mutation in cystic fibrosis patients (deletion of F508). Get exon sequence and mark the codon that encodes F508 in healty individuals

2. Retrieve the **CFTR gene sequence** (from human genomic DNA)

3. Loclalize the affected exon in the gene: individuate position of start and end of affected exon; individuate the position of the codon that encodes F508. Write down the respective nucleotide positions (first nucleotide in obtained gene sequence = 1)

4. Perform PrimerBlase in the region **500 nucleotodes upstream and 500 nucleotide downstream of the triple encoding F508**

5. Get details of top 3 primer paires proposed by PrimerBlast

6. **Have information related to point 3 and 5 ready for next lecture**