

# Geni e genomi

## Ensembl

# 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 specifiche 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

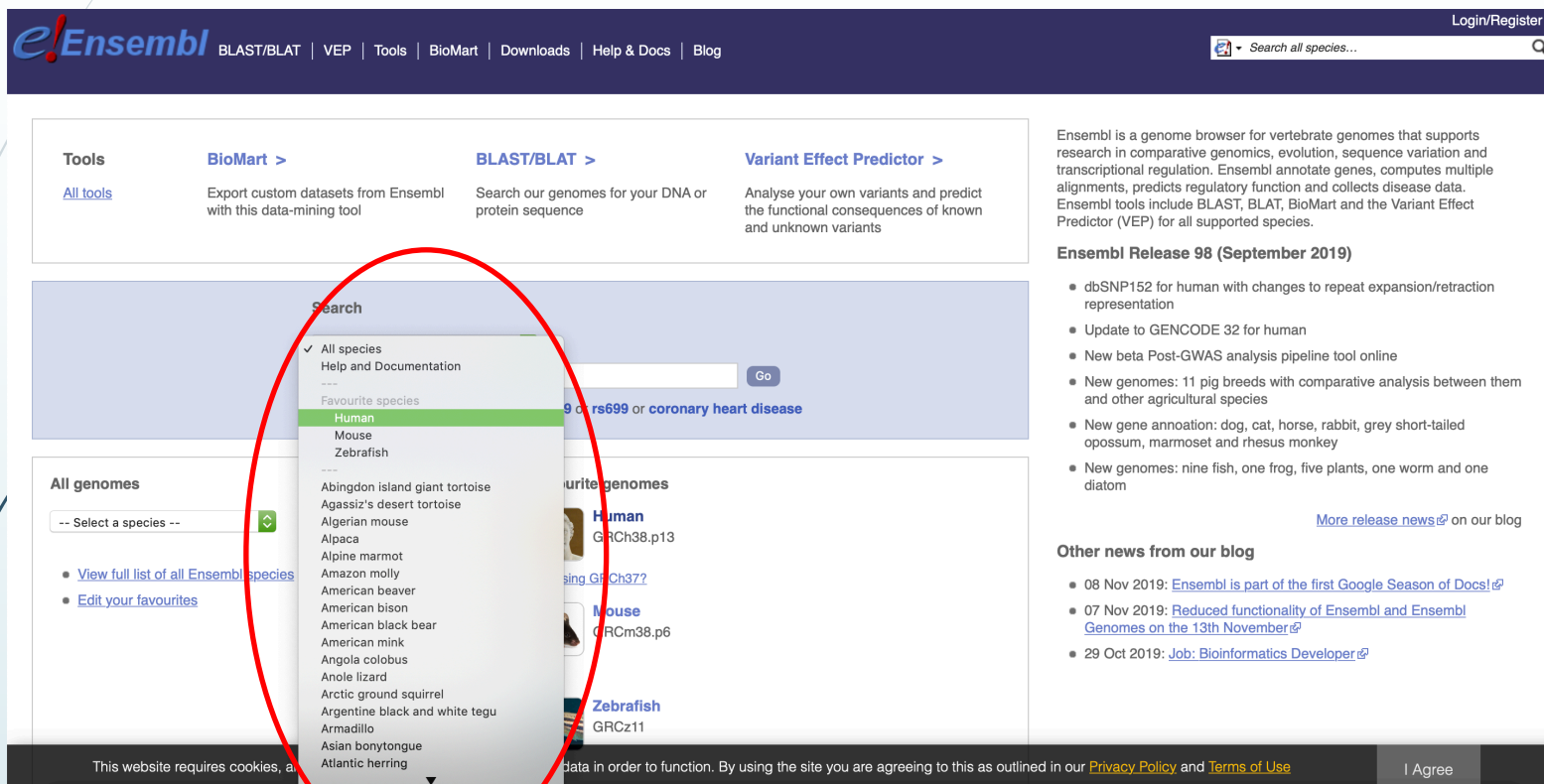
<https://www.ensembl.org/index.html>

The screenshot shows the Ensembl website interface. At the top, there is a navigation bar with the Ensembl logo and links for BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog. A search bar is located in the top right corner. Below the navigation bar, there are four main sections: Tools, BioMart, BLAST/BLAT, and Variant Effect Predictor. The search bar is highlighted with a red circle. Below the search bar, there are examples of search terms: e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease. Below the search bar, there are sections for All genomes and Favourite genomes. The All genomes section has a dropdown menu for species selection. The Favourite genomes section lists Human (GRCh38.p13), Mouse (GRCm38.p6), and Zebrafish (GRCz11). On the right side, there is a section for Ensembl Release 98 (September 2019) with a list of updates and news. Below that, there is a section for Other news from our blog with a list of recent news items.

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



The screenshot shows the Ensembl website interface. At the top, there is a navigation bar with the Ensembl logo and links for BLAST/BLAT, VEP, Tools, BioMart, Downloads, Help & Docs, and Blog. A search bar is located on the right side of the header. Below the header, there are several sections: 'Tools' with links to All tools, BioMart, BLAST/BLAT, and Variant Effect Predictor; 'All genomes' with a species selection dropdown; and 'Favourite genomes' with a list of species. A red circle highlights the 'Search' dropdown menu, which is open and shows a list of species. 'Human' is selected and highlighted in green. The footer contains a cookie notice and an 'I Agree' button.

... 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

**Ensembl** BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog Login/Register

**Tools**

- [All tools](#)
- BioMart >**  
Export custom datasets from Ensembl with this data-mining tool
- BLAST/BLAT >**  
Search our genomes for your DNA or protein sequence
- Variant Effect Predictor >**  
Analyse your own variants and predict the functional consequences of known and unknown variants

**Search**

Human  for

e.g. [BRCA2](#) or [rat 5:62797383.6327669](#) or [rs699](#) or [coronary heart disease](#)

**All genomes**

-- Select a species --

- [View full list of all Ensembl species](#)
- [Edit your favourites](#)

**Favourite genomes**

- Human**  
GRCh38.p13  
[Still using GRCh37?](#)
- Mouse**  
GRCm38.p6
- Zebrafish**  
GRCz11

**Ensembl Release 98 (September 2019)**

- dbSNP152 for human with changes to repeat expansion/retraction representation
- Update to GENCODE 32 for human
- New beta Post-GWAS analysis pipeline tool online
- New genomes: 11 pig breeds with comparative analysis between them and other agricultural species
- New gene annotation: dog, cat, horse, rabbit, grey short-tailed opossum, marmoset and rhesus monkey
- New genomes: nine fish, one frog, five plants, one worm and one diatom

[More release news](#) on our blog

**Other news from our blog**

- 08 Nov 2019: [Ensembl is part of the first Google Season of Docs!](#)
- 07 Nov 2019: [Reduced functionality of Ensembl and Ensembl Genomes on the 13th November](#)
- 29 Oct 2019: [Job: Bioinformatics Developer](#)

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

The screenshot shows the Ensembl search interface. At the top, there's a search bar with 'G6PD' entered and a dropdown menu set to 'Human'. Below the search bar, the results are filtered to 'Only searching Human'. The first result is 'G6PD (Human Gene)', which is circled in red. This result includes the Ensembl ID 'ENSG00000160211', coordinates 'X:154531391-15454752:-1', and a description: 'Glucose-6-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4057]'. Below this, there's a link to the Locus Reference Genomic record for G6PD. The 'Best gene match' section shows a genomic map of the G6PD gene on chromosome 10, with coordinates X:15431391-15454752-1. The 'Suggestions' section lists related gene symbols like gapd, gepd, gapd2, gapdh, gapds, gaped, gnpda, gaad, gad, gadd, gap, gapl, gapo, gapr, gaps, gapt, gard, gbpi, gep, geph, gerd, gid, gild, gip, gipc, gipr, gliud, gnpl, gold, gop, gopc, gpa, gpe, gpi, gpil, grid, grod, grpe, gud, gup, gypa, gype, g6pase, gabpa, gabpb, gabrd, gap1m, gapb3, gapos, gapr1.

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

[https://www.ensembl.org/Homo\\_sapiens/Gene/Summary?db=core;g=ENSG00000160211;r=X:154531391-154547572](https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000160211;r=X:154531391-154547572)

**Ensembl** BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Human (GRCh38.p13) Location: X:154,531,391-154,547,572 Gene: G6PD

**Gene: G6PD** ENSG00000160211

**Description** glucose-6-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4057]

**Gene Synonyms** G6PD1

**Location** [Chromosome X: 154,531,391-154,547,572](#) reverse strand.  
[GRCh38:CM000685.2](#)

**About this gene** This gene has 11 transcripts ([splice variants](#)), 234 orthologues, 1 paralogue, is a member of 1 Ensembl protein family and is associated with 3 phenotypes.

**Transcripts** [Show transcript table](#)

**Summary**

**Name** [G6PD](#) (HGNC Symbol)

**CCDS** This gene is a member of the Human CCDS set: [CCDS44023.1](#)

**UniProtKB** This gene has proteins that correspond to the following UniProtKB identifiers: [P11413](#)

**RefSeq** This Ensembl/Gencode gene contains transcript(s) for which we have [selected identical RefSeq transcript\(s\)](#). If there are other RefSeq transcripts available they will be in the [External references](#) table

**LRG** [LRG\\_148](#) provides a stable genomic reference framework for describing sequence variants for this gene

**Ensembl version** ENSG00000160211.19

**Other assemblies** There is no unapped mapping of this gene onto the GRCh37 assembly. View this locus in the GRCh37 archive: [ENSG00000160211](#) [ENSG00000269087](#)

**Gene type** Protein coding

**Annotation method** Annotation for this gene includes both automatic annotation from Ensembl and Havana manual curation, see [article](#).

[Configure this page](#)

... potremo scoprire che il gene **G6PD** si trova sul **cromosoma X** e che da esso vengono trascritti **11 mRNA diversi (splice variants)**.

# Proviamo insieme con il gene G6PD

The screenshot shows the Ensembl genome browser interface for the G6PD gene. The sidebar on the left, titled 'Gene-based displays', is circled in red and contains the following items: Summary, Splice variants, Transcript comparison, Gene alleles, Sequence, Secondary Structure, Comparative Genomics, Genomic alignments, Gene tree, Gene gain/loss tree, Orthologues, Paralogues, Ensembl protein families, Ontologies, GO: Cellular component, GO: Biological process, GO: Molecular function, Phenotypes, Genetic Variation, Variant table, Variant image, Structural variants, Gene expression, Pathway, Regulation, External references, Supporting evidence, ID History, and Gene history. The main content area displays the following information for the G6PD gene (ENSG00000160211):

- Description:** glucose-6-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4057]
- Gene Synonyms:** G6PD1
- Location:** [Chromosome X: 154,531,391-154,547,572](#) reverse strand. GRCh38:CM000685.2
- About this gene:** This gene has 11 transcripts ([splice variants](#)), [234 orthologues](#), [1 paralogue](#), is a member of [1 Ensembl protein family](#) and is associated with [3 phenotypes](#).
- Summary:**
  - Name:** [G6PD](#) (HGNC Symbol)
  - CCDS:** This gene is a member of the Human CCDS set: [CCDS44023.1](#)
  - UniProtKB:** This gene has proteins that correspond to the following UniProtKB identifiers: [P11413](#)
  - RefSeq:** This Ensembl/Gencode gene contains transcript(s) for which we have [selected identical RefSeq transcript\(s\)](#). If there are other RefSeq transcripts available they will be in the [External references](#) table
  - LRG:** [LRG\\_148](#) provides a stable genomic reference framework for describing sequence variants for this gene
  - Ensembl version:** ENSG00000160211.19
  - Other assemblies:** There is no unapped mapping of this gene onto the GRCh37 assembly. View this locus in the GRCh37 archive: [ENSG00000160211](#) [ENSG00000269087](#)
  - Gene type:** Protein coding
  - Annotation method:** Annotation for this gene includes both automatic annotation from Ensembl and Havana manual curation, see [article](#).

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

**Human (GRCh38.p13)**  
Location: X:154,531,391-154,547,572 Gene: G6PD

**Gene: G6PD** ENSG00000160211

**Description:** glucose-6-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4057]

**Gene Synonyms:** G6PD1

**Location:** Chromosome X: 154,531,391-154,547,572 reverse strand. GRCh38:CM000685.2

**About this gene:** This gene has 11 transcripts (splice variants), 234 orthologues, 1 paralogue, is a member of 1 Ensembl protein family and is associated with 3 phenotypes.

**Transcripts:** [hide transcript table](#)

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq Match	Flags
G6PD-202	<a href="#">ENST00000393562.10</a>	2223	<a href="#">515aa</a>	Protein coding	<a href="#">CCDS44023</a>	<a href="#">A0A384NL00</a> <a href="#">P11413</a>	<a href="#">NM_001360016.2</a>	TSL:1 GENCODE basic APPRIS P1 MANE Select v0.6
G6PD-203	<a href="#">ENST00000393564.6</a>	1661	<a href="#">515aa</a>	Protein coding	<a href="#">CCDS44023</a>	<a href="#">A0A384NL00</a> <a href="#">P11413</a>	-	TSL:5 GENCODE basic APPRIS P1
G6PD-201	<a href="#">ENST00000369620.6</a>	1799	<a href="#">561aa</a>	Protein coding	-	<a href="#">P11413</a>	-	TSL:5 GENCODE basic
G6PD-205	<a href="#">ENST00000439227.5</a>	1074	<a href="#">338aa</a>	Protein coding	-	<a href="#">E7EU18</a>	-	CDS 3' incomplete TSL:5
G6PD-206	<a href="#">ENST00000440967.5</a>	1056	<a href="#">320aa</a>	Protein coding	-	<a href="#">E7EM57</a>	-	CDS 3' incomplete TSL:5
G6PD-204	<a href="#">ENST00000433845.1</a>	817	<a href="#">256aa</a>	Protein coding	-	<a href="#">E9PD92</a>	-	CDS 3' incomplete TSL:3
G6PD-209	<a href="#">ENST00000490651.1</a>	783	No protein	Retained intron	-	-	-	TSL:2
G6PD-208	<a href="#">ENST00000489497.1</a>	512	No protein	Retained intron	-	-	-	TSL:2
G6PD-211	<a href="#">ENST00000647501.1</a>	471	No protein	Retained intron	-	-	-	-
G6PD-207	<a href="#">ENST00000488434.1</a>	432	No protein	Retained intron	-	-	-	TSL:2
G6PD-210	<a href="#">ENST00000497281.5</a>	532	No protein	lncRNA	-	-	-	TSL:2

**Splice variants**

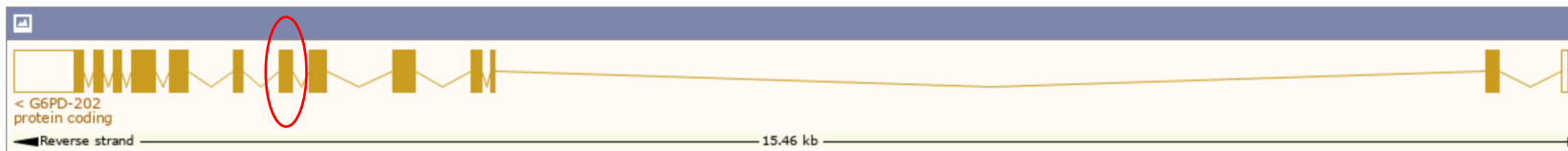
Le **varianti di splicing** del gene G6PD sono 11, identificate da ID diversi, 6 delle quali codificanti per proteine con lunghezze aminoacidiche diverse.

# Proviamo insieme con il gene G6PD

[https://www.ensembl.org/Homo\\_sapiens/Transcript/Summary?db=core;g=ENSG00000160211;r=X:154531391-154547572;t=ENST00000393562](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 [Genecode 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

**Ensembl** BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Human (GRCh38.p13)

Location: X:154,531,391-154,547,572 Gene: G6PD Transcript: G6PD-202

**Transcript-based displays**

- Summary
- Sequence
  - Exons
  - cDNA
  - Protein
- Protein Information
  - Protein summary
  - Domains & features
  - Variants
  - 3D Protein model
- Genetic Variation
  - Variant table
  - Variant image
  - Haplotypes
  - Population comparison
  - Comparison image
- External References
  - General identifiers
  - Oligo probes
  - Supporting evidence
- ID History
  - Transcript history
  - Protein history

Configure this page | Custom tracks | Export data | Share this page | Bookmark this page

**Transcript: G6PD-202** ENST00000393562.10

Description: glucose-6-phosphate dehydrogenase [Source:HGNC Symbol;Acc:HGNC:4057]

Gene Synonyms: G6PD1

Location: Chromosome X: 154,531,391-154,546,846 reverse strand.

About this transcript: This transcript has 13 exons, is annotated with 38 domains and features, is associated with 4149 variant alleles and maps to 576 oligo probes.

Gene: This transcript is a product of gene ENSG00000160211.19 [Hide transcript table](#)

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq Match	Flags
G6PD-202	ENST00000393562.10	2223	515aa	Protein coding	<a href="#">CCDS44023</a>	<a href="#">A0A384NL00</a> <a href="#">P11413</a>	<a href="#">NM_001360016.2</a>	TSL:1 GENCODE basic APPRIS P1 MANE Select v0.6
G6PD-203	ENST00000393564.6	1661	515aa	Protein coding	<a href="#">CCDS44023</a>	<a href="#">A0A384NL00</a> <a href="#">P11413</a>	-	TSL:5 GENCODE basic APPRIS P1
G6PD-201	ENST00000369620.6	1799	561aa	Protein coding	-	<a href="#">P11413</a>	-	TSL:5 GENCODE basic
G6PD-205	ENST00000439227.5	1074	338aa	Protein coding	-	<a href="#">E7EU18</a>	-	CDS 3' incomplete TSL:5
G6PD-206	ENST00000440967.5	1056	320aa	Protein coding	-	<a href="#">E7EM57</a>	-	CDS 3' incomplete TSL:5
G6PD-204	ENST00000433845.1	817	256aa	Protein coding	-	<a href="#">E9PD92</a>	-	CDS 3' incomplete TSL:3
G6PD-209	ENST00000490651.1	783	No protein	Retained intron	-	-	-	TSL:2
G6PD-208	ENST00000489497.1	512	No protein	Retained intron	-	-	-	TSL:2
G6PD-211	ENST00000647501.1	471	No protein	Retained intron	-	-	-	-
G6PD-207	ENST00000488434.1	432	No protein	Retained intron	-	-	-	TSL:2
G6PD-210	ENST00000497281.5	532	No protein	lncRNA	-	-	-	TSL:2

**Summary**

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

This transcript is a product of gene [ENSG00000160211.19](#) [Hide transcript table](#)

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq Match	Flags
G6PD-202	<a href="#">ENST00000393562.10</a>	2223	515aa	Protein coding	<a href="#">CCDS44023</a>	<a href="#">A0A384NL00</a> <a href="#">P11413</a>	<a href="#">NM_001360016.2</a>	TSL:1 GENCODE basic APPRIS P1 MANE Select v0.6
G6PD-203	<a href="#">ENST00000393564.6</a>	1661	515aa	Protein coding	<a href="#">CCDS44023</a>	<a href="#">A0A384NL00</a> <a href="#">P11413</a>	-	TSL:5 GENCODE basic APPRIS P1
G6PD-201	<a href="#">ENST00000389620.6</a>	1799	561aa	Protein coding	-	<a href="#">P11413</a>	-	TSL:5 GENCODE basic
G6PD-205	<a href="#">ENST00000439227.5</a>	1074	338aa	Protein coding	-	<a href="#">E7EU18</a>	-	CDS 3' incomplete TSL:5
G6PD-206	<a href="#">ENST00000440967.5</a>	1056	320aa	Protein coding	-	<a href="#">E7EM57</a>	-	CDS 3' incomplete TSL:5
G6PD-204	<a href="#">ENST00000433845.1</a>	817	256aa	Protein coding	-	<a href="#">E9PD92</a>	-	CDS 3' incomplete TSL:3
G6PD-209	<a href="#">ENST00000490851.1</a>	783	No protein	Retained intron	-	-	-	TSL:2
G6PD-208	<a href="#">ENST00000489497.1</a>	512	No protein	Retained intron	-	-	-	TSL:2
G6PD-211	<a href="#">ENST00000647501.1</a>	471	No protein	Retained intron	-	-	-	-
G6PD-207	<a href="#">ENST00000488434.1</a>	432	No protein	Retained intron	-	-	-	TSL:2
G6PD-210	<a href="#">ENST00000497281.5</a>	532	No protein	lncRNA	-	-	-	TSL:2

**cDNA sequence**

[Download sequence](#) [BLAST this sequence](#)

Codons [Alternating codons](#) [Alternating codons](#)

Exons [An exon](#) [Another exon](#)

Variants [3 prime UTR](#) [5 prime UTR](#) [Coding sequence](#) [Frameshift](#) [Inframe deletion](#) [Missense](#) [Splice region](#)  
[Stop gained](#) [Synonymous](#)

Other [UTR](#)

Markup [loaded](#)

• Variants are filtered by consequence type

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

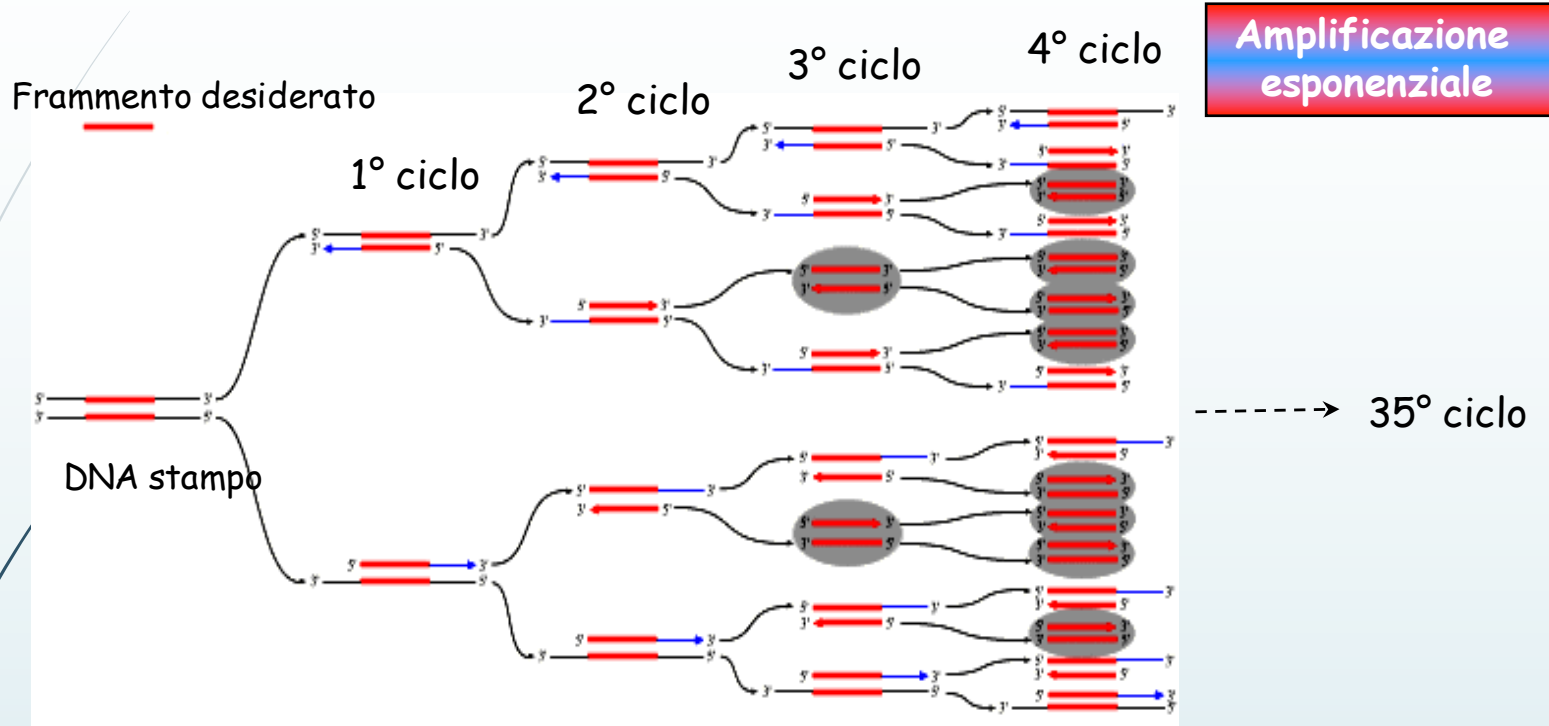
0	R S	*W	R*N	R R	**	*M	YR		* YY	B	R	**	14								
	NTN	AGCC	ANT	TANN	TNG	NAC	NCAN	NAT	NN	TGG	GAGA	AG	CNN	TNG	GGAN	GN	CCT	GCAG	599		
475	ATG	AGCC	CAG	ATAG	GCTG	GAAC	CGCAT	CATC	GTG	GAGA	AGCC	CTTC	GGG	AGG	GAC	CTGC	AG	534			
159	-M--	S--	Q--	I--	G--	W--	N--	R--	I--	I--	V--	E--	K--	P--	F--	G--	R--	D--	L--	Q-	178
0	*RY	*	W	YVR		Y	MR	YR	YY	*Y	R	YYR	R	S	*	***	20				
	NNN	T	TGN	NNN	CTG	TNC	NN	CCAN	NTCT	NN	TC	NN	T	NN	TGAG	NAN	CA	ATCT	ACNN	656	
535	AGCT	CTG	ACCG	GCTG	TCCA	ACC	CATC	CTC	CTCC	CTGT	TCCG	TGAG	GACC	AGAT	CTAC	CGC	594				
179	-S--	S--	D--	R--	L--	S--	N--	H--	I--	S--	S--	L--	F--	R--	E--	D--	Q--	I--	Y--	R-	198

... e facendo uno zoom sulla sequenza scaricata possiamo evidenziare la presenza di alcuni numeri che indicano:

- il n° del nucleotide partendo dall'inizio del gene
- Il n° del nucleotide partendo dall'inizio della sequenza codificante (ATG)
- Il n° dell'aminoacido

La posizione **563** della sequenza codificante sarà facilmente rintracciabile in questo file scaricato da Ensembl.

# Cosa è una PCR?

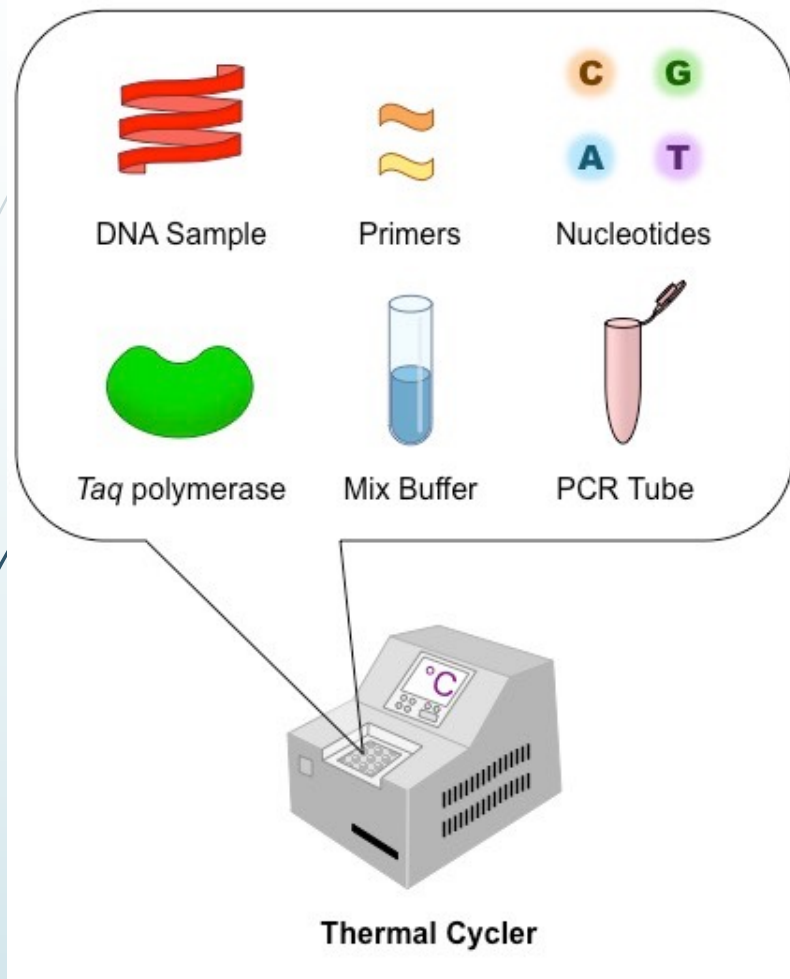


La **Polimerase **Chain **Rea**ction** è una tecnica usata in biologia molecolare per **amplificare** una porzione di DNA, generando milioni di copie della stessa.**

Sviluppata nel 1983 da Kary Mullis, che grazie ad essa si meritò il premio Nobel nel 1993, è usata quotidianamente per:

- Clonaggi di DNA
- Analisi di geni
- Identificazione di fingerprint genetico
- Evidenziazione e diagnosi di patologie infettive

# Come funziona una PCR?



Le componenti base di una PCR sono rappresentate da:

- Il **DNA** da amplificare chiamato template
- **Nucleotidi** che serviranno per estendere i nuovi filamenti
- L'enzima, una **Taq polimerasi** che aggiungerà i nucleotidi al filamento neosintetizzato
- Un **mix buffer** contenente sali ed acqua indispensabili per la reazione
- I **primers**

La reazione avverrà a temperatura controllata in un termociclatore che permetterà di eseguire tutti i cicli necessari all'amplificazione del frammento desiderato.

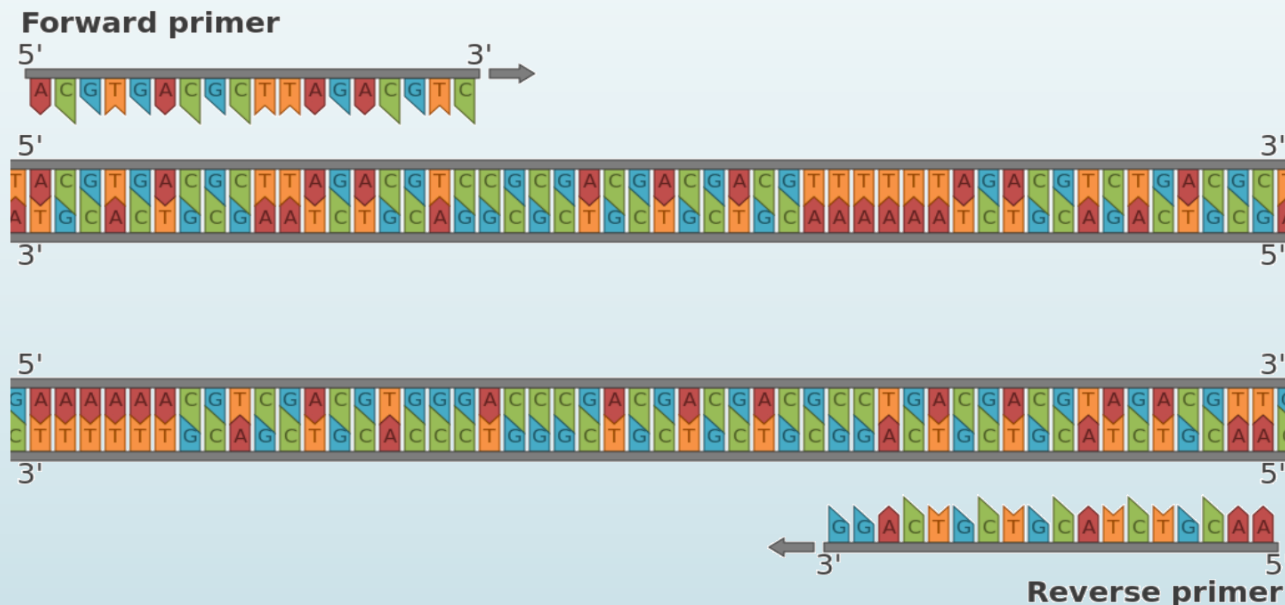


# Che cosa è un primer per PCR?

Un primer è un corto filamento oligonucleotidico che viene usato per dare l'innescò 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 bp sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosiddetti 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 bp sono sconsigliati perché creano una maggiore probabilità che avvenga un appaiamento scorretto o che i primer si appaino tra loro, creando i cosiddetti 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 cosiddetti 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 cosiddetti 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 template, 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 cosiddetti 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 template, 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.

# Proviamo insieme con il gene G6PD

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



RESEARCH ARTICLE

Open Access

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

Soad K Al-Jaouni<sup>1\*</sup>, Jummanah Jarullah<sup>1</sup>, Essam Azhar<sup>2</sup> and Kamran Moradkhani<sup>3,4</sup>

**Table 1 Primers used for identification of glucose-6-phosphate dehydrogenase defects<sup>a</sup>**

Analysis	Analyzed Region	Forward 5'→3'	Reverse 5'→3'
RFLP Analysis (T <sub>m</sub> = 61°C)	Exon 3	GTGGAGGATGATGTATGTAGGT	AGGGCAGGGCACAGCTGTAA
	Exon 4	TACAGTCGTGCCCTGCCCT	CCGAAGCTGGCCATGCTGG
	Exon 5	CTGTGTGTGTCTGTCTGTC	GGAGGGCAACGGCAAGCCTT
	Exon 6	GCAGCTGTGATCCTCACTCC	GCAAGGTGGAGGAAGTCACTCC

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.

# Proviamo insieme con il gene G6PD

CAGGCCGCAGGGTGGCAGCCTTGCTCTGCGAATGCAGCATGGCCCCGCGCTGGGTGGTTTC  
CCAACCCAGCCAGAGGCTCTTGTCTCTGGCTGGTFTTGAATGCGGGGTAGTAAAGCAA  
AGGTCTCTTCTCATTTTTCAAACCAATGAGGAAGCCATGGCTTGGATGCCTCCTCCCC  
TGCTCCCCACAGGCCTTACAGGCCACTCAGACCCACCGGGGACCCAGCATGAGGCAGGAG  
GGGAACGGGCCCCCGGCAGCATGCCAGCAATGCCACCCTGGCACCCAGGGTGGGAAGGCT  
TCCCCGAAGGTGTTGAGCCAGAGGGTCACTGGGAACACAAGGCACGGGAGGTGGCCACG  
GGGGCGAGGAGGTTCTGGCCTCTACTCCCCTGGGAGGGCGTCTGAATGATGCAGCTCTGA  
TCCTCACTCCCGAAGAGGGGTTCAAGGGGGTAACGCAGCTCCGGGCTCCCAGCAGAGGC  
TGGAACCGCATCATCGTGGAGAAGCCCTTCGGGAGGGACCTGCAGAGCTCTGACCGGCTG  
TCCAACCACA TCTCTCCCTGTTCCGTGAGGACCAGATCTACCGCATCGACCACTACCTG  
GGCAAGGAGATGGTGCAGAACCTCATGGTGCTGAGGTGGGGCCAAGCCTGGGCCGGGGGA  
CCAGGGTGGGGTGGTACTCAGGAGCCTCACCTGGCCCACTGCCTCCCCGAGGACGAATT  
CCTCCAGA ACTCAGACAAGGGTGACCCCTCACATGTGGCCCCTGCACCACAGAGGCCCAA  
GGTCAGTTCCTCCACCTTGC CCTCCCTGCAGATTTGCCAACAGGATCTTCGGCCCCATC  
TGGAACCGGGACAACATCGCCTGCGTTATCCTCACCTTCAAGGAGCCCTTTGGCACTGAG  
GGTCGCGGGGGCTATTTTCGATGAATTTGGGATCATCCGGTGAGAGCTCTTCTCTCTCCT  
GGGAGGCTGGCACAGGGTGGCAGAGCCAGTCACCCCTGCAGGGCTACTCTTCCCTATCTTG  
GGGGAGCTCCTCCTCACCCCTGCAGTTCAAAACCTAAGTGTCTGAGCTATCAGACCGGGCT  
GGAAAGGGCTGGACCCCTACACAGCCAAGCACCCACGGTTTTATGATTCAGTGATAGCA  
TCACCATGTCCTTCTTGATTTAAGGGGACCTGGAAGACAAGGGGGATCAGGAAGTGAGT  
CTTGCACTTGTCACTAGGAAGCCTTGTTTGGGGTCCCCATGCCCTTGAACCAGGTGAAC

PCR oligos

**GCAGCTCTGATCCTCACTCC**  
**GGTCAGTTCCTCCACCTTGC**

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  $\approx 60^{\circ}\text{C}$
- Contengono G/C all'estremità 3' terminale

# 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/>.

The screenshot shows the NCBI Primer-BLAST web interface. The top navigation bar includes NIH, U.S. National Library of Medicine, and NCBI National Center for Biotechnology Information. The main title is 'Primer-BLAST: A tool for finding specific primers'. Below the title, there are links for 'Reset page', 'Save search parameters', 'Retrieve recent results', 'Publication', and 'Tips for finding specific primers'. The 'PCR Template' section is highlighted with a red circle and contains the following fields:

- Enter accession, gi, or FASTA sequence (A refseq record is preferred) [Clear]
- Range: Forward primer (From, To) [Clear]
- Reverse primer [Clear]
- Or, upload FASTA file: Scegli file | nessun file selezionato

The 'Primer Parameters' section includes:

- Use my own forward primer (5'->3' on plus strand) [Clear]
- Use my own reverse primer (5'->3' on minus strand) [Clear]
- PCR product size: Min 70, Max 1000
- # of primers to return: 10
- Primer melting temperatures (T<sub>m</sub>): Min 57.0, Opt 60.0, Max 63.0, Max T<sub>m</sub> difference 3

The 'Exon/intron selection' section includes:

- Exon junction span: No preference
- Exon junction match: Exon at 5' side 7, Exon at 3' side 4
- Minimal number of bases that must anneal to exons at the 5' or 3' side of the junction
- Intron inclusion:  Primer pair must be separated by at least one intron on the corresponding genomic DNA
- Intron length range: Min 1000, Max 1000000

**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/>.

The screenshot shows the NCBI Primer-BLAST interface. The top navigation bar includes NIH, U.S. National Library of Medicine, and NCBI National Center for Biotechnology Information. The main heading is 'Primer-BLAST: A tool for finding specific primers'. Below this, there are links for 'Reset page', 'Save search parameters', 'Retrieve recent results', 'Publication', and 'Tips for finding specific primers'. The 'PCR Template' section is circled in red and contains a text input field for 'Enter accession, gi, or FASTA sequence (A refseq record is preferred)', a 'Clear' button, and a 'Range' section with 'Forward primer' and 'Reverse primer' fields. Below this is an 'Upload FASTA file' section with a 'Scegli file' button and the text 'nessun file selezionato'. The 'Primer Parameters' section is circled in green and includes fields for 'Use my own forward primer (5'->3' on plus strand)', 'Use my own reverse primer (5'->3' on minus strand)', 'PCR product size' (Min: 70, Max: 1000), and '# of primers to return' (10). The 'Exon/intron selection' section includes 'Exon junction span' (No preference), 'Exon junction match' (Exon at 5' side: 7, Exon at 3' side: 4), 'Intron inclusion' (checkbox for 'Primer pair must be separated by at least one intron on the corresponding genomic DNA'), and 'Intron length range' (Min: 1000, Max: 1000000).

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

- **PCR product size** → min 100 – max 900

# Proviamo insieme con il gene G6PD

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

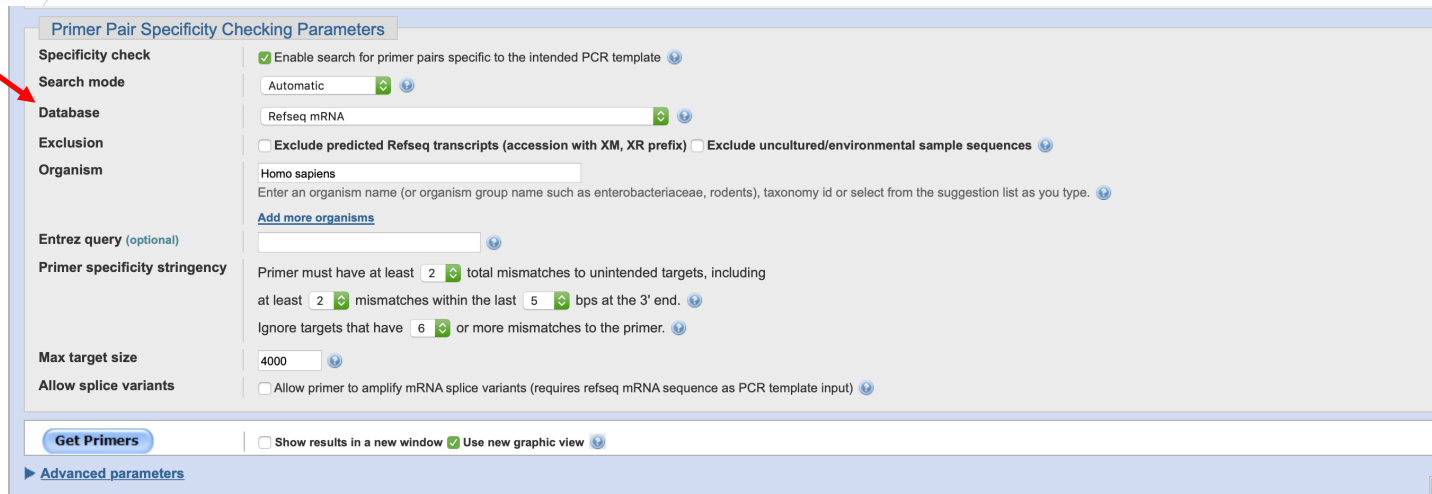
Avremmo dovuto selezionare la sequenza di interesse ed utilizzare un tool online per il disegno dei primers, come <https://www.ncbi.nlm.nih.gov/tools/primer-blast/>.

The screenshot shows the NCBI Primer-BLAST interface. The top navigation bar includes NIH, U.S. National Library of Medicine, and NCBI National Center for Biotechnology Information. The main heading is 'Primer-BLAST: A tool for finding specific primers'. Below this, there are links for 'Reset page', 'Save search parameters', 'Retrieve recent results', 'Publication', and 'Tips for finding specific primers'. The 'PCR Template' section is highlighted with a red circle and contains a text input field for 'Enter accession, gi, or FASTA sequence (A refseq record is preferred)', a 'Clear' button, and a 'Range' section with 'Forward primer' and 'Reverse primer' fields. Below this is an 'Upload FASTA file' section. The 'Primer Parameters' section is highlighted with a green circle and includes fields for 'Use my own forward primer (5'->3' on plus strand)', 'Use my own reverse primer (5'->3' on minus strand)', 'PCR product size' (Min: 70, Max: 1000), and '# of primers to return' (10). A blue arrow points to the 'Primer melting temperatures (Tm)' section, which has fields for Min (57.0), Opt (60.0), Max (63.0), and Max Tm difference (3). The 'Exon/intron selection' section is also visible, with options for 'Exon junction span', 'Exon junction match', 'Intron inclusion', and 'Intron length range'.

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

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

# Proviamo insieme con il gene G6PD



**Primer Pair Specificity Checking Parameters**

**Specificity check**  Enable search for primer pairs specific to the intended PCR template

**Search mode** Automatic

**Database** Refseq mRNA

**Exclusion**  Exclude predicted Refseq transcripts (accession with XM, XR prefix)  Exclude uncultured/environmental sample sequences

**Organism** Homo sapiens  
Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type.  
[Add more organisms](#)

**Entrez query (optional)**

**Primer specificity stringency** Primer must have at least 2 total mismatches to unintended targets, including at least 2 mismatches within the last 5 bps at the 3' end.  
Ignore targets that have 6 or more mismatches to the primer.

**Max target size** 4000

**Allow splice variants**  Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input)

**Get Primers**  Show results in a new window  Use new graphic view

[Advanced parameters](#)

- Selezioniamo il **database di riferimento** (mRNA – genomic – custom...)

# Proviamo insieme con il gene G6PD

**Primer Pair Specificity Checking Parameters**

**Specificity check**  Enable search for primer pairs specific to the intended PCR template

**Search mode** Automatic

**Database** Refseq mRNA

**Exclusion**  Exclude predicted Refseq transcripts (accession with XM, XR prefix)  Exclude uncultured/environmental sample sequences

**Organism** Homo sapiens  
Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type.  
[Add more organisms](#)

**Entrez query (optional)**

**Primer specificity stringency** Primer must have at least 2 total mismatches to unintended targets, including at least 2 mismatches within the last 5 bps at the 3' end.  
Ignore targets that have 6 or more mismatches to the primer.

**Max target size** 4000

**Allow splice variants**  Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input)

**Get Primers**  Show results in a new window  Use new graphic view

[Advanced parameters](#)

- 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

Primer Pair Specificity Checking Parameters

Enable search for primer pairs specific to the intended PCR template

Search mode: Automatic

Database: Refseq mRNA

Exclusion:  Exclude predicted Refseq transcripts (accession with XM, XR prefix)  Exclude uncultured/environmental sample sequences

Organism: Homo sapiens

Entrez query (optional):

Primer specificity stringency: Primer must have at least 2 total mismatches to unintended targets, including at least 2 mismatches within the last 5 bps at the 3' end. Ignore targets that have 6 or more mismatches to the primer.

Max target size: 4000

Allow splice variants:  Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input)

**Get Primers**  Show results in a new window  Use new graphic view

Advanced parameters

- 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.

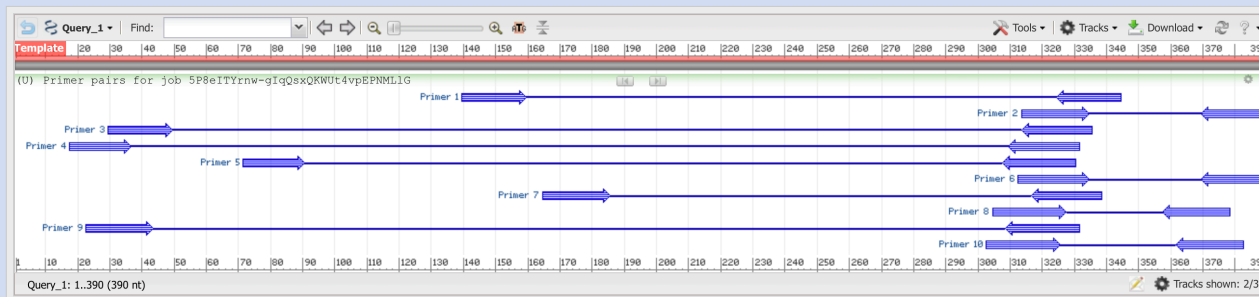
**GET PRIMERS** ed attendiamo!!

# Proviamo insieme con il gene G6PD

Primer-BLAST Results

**Input PCR template** |cd|Query\_1  
**Range** |1 - 390  
**Specificity of primers** |Primer pairs are specific to input template as no other targets were found in selected database: Refseq mRNA (Organism limited to Homo sapiens)  
**Other reports** |> Search Summary

**Graphical view** Show/hide Graphical view of primer pairs



**Detailed primer reports**

**Primer pair 1**

	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	ATCTCCTCCCTGTTCCGTTGA	Plus	20	140	159	59.96	55.00	2.00	2.00
Reverse primer	ATGTGAGGGGTCACCCCTTGT	Minus	20	344	325	60.77	55.00	7.00	3.00
Product length	205								

**Primer pair 2**

	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CCAGAACTCAGACAAGGGTGA	Plus	21	314	334	59.31	52.38	3.00	2.00
Reverse primer	CAAGGTGGAGGAACCTGACCT	Minus	20	389	370	58.94	55.00	4.00	3.00
Product length	76								

**Primer pair 3**

	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GGTTCAAGGGGTAACGCAG	Plus	20	30	49	60.96	60.00	3.00	1.00
Reverse primer	GTCACCCCTGTCTGAGTTCTGG	Minus	22	335	314	60.55	54.55	3.00	0.00
Product length	306								

**Primer pair 4**

	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	TCCCGAAGAGGGGTTCAA	Plus	19	18	36	60.15	57.89	4.00	3.00
Reverse primer	CCCTTGCTGAGTTCTGGAGGA	Minus	22	331	310	61.09	54.55	3.00	2.00
Product length	314								

Il tool restituirà una lista di primers che sono stati disegnati sulla nostra sequenza selezionata.

# Esercizio

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 la sequenza del cDNA
- Rintraccia la mutazione di interesse (**ΔF508**) sulla sequenza scaricata
- Disegna una coppia di primers per amplificare una regione che includa la mutazione **ΔF508**

# Esercizio – disegna oligo per CFTR ( $\Delta F508$ )

```
0 S***K*WYYR***** Y R * D M N YY Y R DRW*R***** V R* * * 21
  NNNNNNNNNNNNNNNNNNNNAGNCNANANTTANNTCNANTNNN NN NNNN ANAN TGA 1384
1369 GCTGGATCCACTGGAGCAGGCAAGACTTCACTTCTAATGGTGATTATGGGAGAAGCTGGAG 1428
  457 -A--G--S--T--G--A--G--K--T--S--L--L--M--V--I--M--G--E--L--E- 476

NH **KVVD RBR** *W YR DV * R R * H *K*YK* Y YVSK*Y*Y YRV 29
  NNTT NNNNNNNNNNNNNAAGNCCNNTGNANAAT TANN NN TNNNNNNNNNNC 1442
1429 CCTTCAGAGGGTAAAATTAAGCACAGTGAAGAATTTCAATCTGTTCAGTTTTCTGG 1488
  477 -P--S--E--G--K--I--K--H--S--G--R--I--S--F--C--S--Q--F--S--W- 496

0 * YSB Y RHR* RH *B * VBV***BNY R * R KR *MMYR*YR** 26
  TTANNNCNNGNNNCNNTAA NAAA TNNNNNNNNNGGTNTTCTNTNNNNNNNNN 1501
1489 ATTATGCCTGGCACCATTAAAGAAAATATCATCTTTGGTGTTCCTATGATGAATATAGA 1548
  497 -I--M--P--G--T--I--K--E--N--I--I--F--G--V--S--Y--D--E--Y--R- 516

0 *****RYD YH SR YRMY SRRR NVRYVWS* R KY KM R RD 30
  NNNNNNNNNNTNNNNNAAGCANNNAACTANNNNNNNNNNCCNANN TNNANANAA 1551
1549 TACAGAAGCGTCATCAAAGCATGCCAACTAGAAGAGGACATCTCCAAGTTTGCAGAGAAA 1608
  517 -Y--R--S--V--I--K--A--C--Q--L--E--E--D--I--S--K--F--A--E--K- 536

0 *HW*H***D ** DR S*RRKK K*WKYRYW*Y MDND *RR Y**BV*** R**R Y 31
  NNNNNNNNNNTTTNTTNNANNNNNNNGNNNNNNNNNGNNNCGNNTNN NN NNANNTN 1611
1609 GACAATATAGTTCTTGGAGAAGGTGGAATCACACTGAGTGGAGGTCAACGAGCAAGAATT 1668
  537 -D--N--I--V--L--G--E--G--G--I--T--L--S--G--G--Q--R--A--R--I- 556

0 W* Y*DH RVM** V D*YRR*BR RM * **WDBRH * YY*R*S B*YM*WW*KB 32
  NNTTNNNNANNNNNANNTAN NNNNNNTNTTANNNNNNNTANNNNNTN NN NN NNNA 1669
1669 TCTTTAGCAAGAGCAGTATACAAAGATGCTGATTTGTATTTATTAGACTCTCCTTTTGGAG 1728
  557 -S--L--A--R--A--V--Y--K--D--A--D--L--Y--L--L--D--S--P--F--G- 576
```

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 ( $\Delta F508$ )

Input PCR template: Icl|Query\_1  
 Range: 1 - 356  
 Specificity of primers: Primer pairs are specific to input template as no other targets were found in selected database: Refseq mRNA (Organism limited to Homo sapiens)  
 Other reports: [Search Summary](#)

## Graphical view of primer pairs

Query\_1: 1..356 (356 nt)

Primer pair	Sequence (5'→3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
<b>Forward primer</b>	ACTGGAGCAGGCAAGACTTC	Plus	20	10	29	59.96	55.00	5.00	3.00
<b>Reverse primer</b>	CTTGCTCGTTGACCTCCACT	Minus	20	296	277	59.97	55.00	4.00	2.00
<b>Product length</b>	287								
<b>Products on intended targets</b>									
>NM_000492.4 Homo sapiens CF transmembrane conductance regulator (CFTR), mRNA									
product length = 287									
Forward primer	1	ACTGGAGCAGGCAAGACTTC	20						
Template	1448	.....	1467						
Reverse primer	1	CTTGCTCGTTGACCTCCACT	20						
Template	1734	.....	1715						
<b>Primer pair 2</b>									
<b>Forward primer</b>	GCGTCATCAAAGCATGCCAA	Plus	20	188	207	60.11	50.00	6.00	1.00
<b>Reverse primer</b>	GCTCGTTGACCTCCACTCAG	Minus	20	293	274	60.39	60.00	4.00	3.00
<b>Product length</b>	106								
<b>Products on intended targets</b>									
>NM_000492.4 Homo sapiens CF transmembrane conductance regulator (CFTR), mRNA									
product length = 106									
Forward primer	1	GCGTCATCAAAGCATGCCAA	20						
Template	1626	.....	1645						
Reverse primer	1	GCTCGTTGACCTCCACTCAG	20						
Template	1731	.....	1712						

... ed incollo la mia sequenza su primer blast.  
 Aggiungo come restrizioni la lunghezza del frammento desiderato ed ottengo una lista di primers.

# Esercizio – disegna oligo per CFTR ( $\Delta$ F508)

## Primer pair 3

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CGTCATCAAAGCATGCCAACT	Plus	21	189	209	59.80	47.62	6.00	1.00
Reverse primer	ATTCTTGCTCGTTGACCTCCA	Minus	21	299	279	59.65	47.62	2.00	2.00
Product length	111								

### Products on intended targets

>[NM\\_000492.4](#) Homo sapiens CF transmembrane conductance regulator (CFTR), mRNA

```
product length = 111
Forward primer 1 CGTCATCAAAGCATGCCAACT 21
Template 1627 ..... 1647

Reverse primer 1 ATTCTTGCTCGTTGACCTCCA 21
Template 1737 ..... 1717
```

## Primer pair 4

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	AACTGGAGCCTTCAGAGGGTA	Plus	21	53	73	60.20	52.38	7.00	2.00
Reverse primer	TCCACTCAGTGTGATCCACC	Minus	21	282	262	59.65	52.38	7.00	3.00
Product length	230								

### Products on intended targets

>[NM\\_000492.4](#) Homo sapiens CF transmembrane conductance regulator (CFTR), mRNA

```
product length = 230
Forward primer 1 AACTGGAGCCTTCAGAGGGTA 21
Template 1491 ..... 1511

Reverse primer 1 TCCACTCAGTGTGATCCACC 21
Template 1720 ..... 1700
```

## Primer pair 5

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	CTGGAGCCTTCAGAGGGTAA	Plus	20	55	74	58.43	55.00	7.00	1.00
Reverse primer	ACCTCCACTCAGTGTGATCC	Minus	21	285	265	59.37	52.38	7.00	1.00
Product length	231								

### Products on intended targets

>[NM\\_000492.4](#) Homo sapiens CF transmembrane conductance regulator (CFTR), mRNA

```
product length = 231
Forward primer 1 CTGGAGCCTTCAGAGGGTAA 20
Template 1493 ..... 1512

Reverse primer 1 ACCTCCACTCAGTGTGATCC 21
Template 1723 ..... 1703
```

## Primer pair 6

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	ACAGAAGCGTCATCAAAGCA	Plus	20	182	201	58.12	45.00	4.00	0.00
Reverse primer	CCACTCAGTGTGATCCACC	Minus	20	281	262	58.19	55.00	7.00	3.00
Product length	100								

### Products on intended targets

>[NM\\_000492.4](#) Homo sapiens CF transmembrane conductance regulator (CFTR), mRNA

```
product length = 100
Forward primer 1 ACAGAAGCGTCATCAAAGCA 20
Template 1620 ..... 1639

Reverse primer 1 CCACTCAGTGTGATCCACC 20
Template 1719 ..... 1700
```

Lo step successivo sarà controllare se le coppie di primers amplificano una sequenza che includa la fenilalanina mutata.