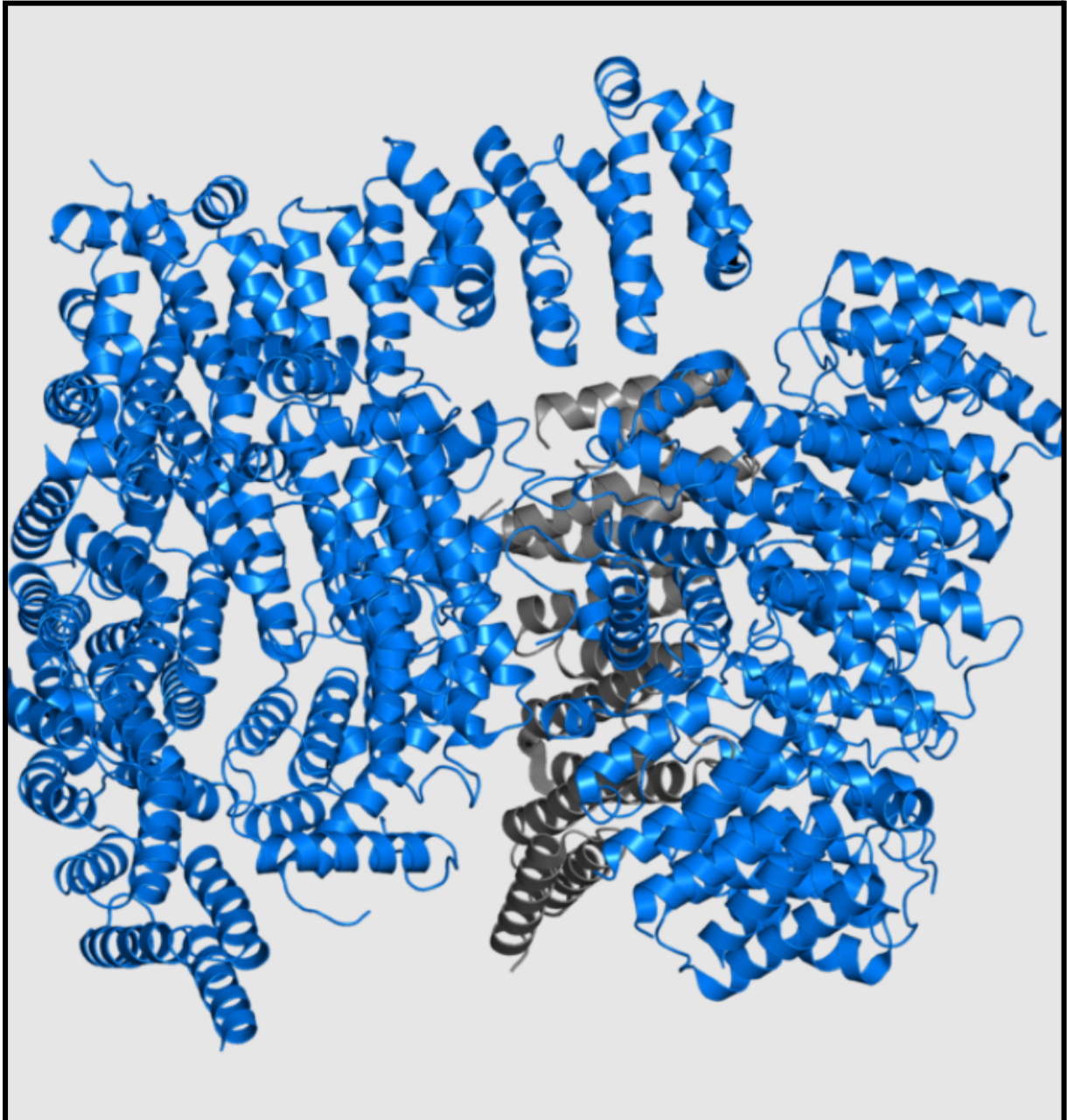


Disseny d'edició genètica CRISPR per corregir la malaltia de Huntington



Uracil
4 de novembre del 2022

Abstract

Este trabajo trata de establecer el funcionamiento y la utilidad de una herramienta de edición genética como CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*) en una posible terapia génica. Por ello se han diseñado un conjunto de moléculas que se podrían usar para hacer la corrección en una enfermedad genética real. Se ha seleccionado una enfermedad con repercusión en la actualidad, minoritaria y causada por una mutación genética, específicamente la mutación en un solo gen para que se pueda emplear la técnica CRISPR para ser solucionada. Se seleccionó la Enfermedad de Huntington, una enfermedad muy grave de la cual no existe cura. Esa enfermedad afecta al sistema nervioso y produce síntomas muy severos que desembocan en la muerte del que la sufre. Para conseguir la corrección de la mutación que causa esta enfermedad se ha buscado de forma específica la localización de la mutación dentro del gen concreto, para después diseñar las moléculas necesarias. Ellas son, por una parte, el fragmento ARN guía complementario a este gen, que junto con la proteína CAS9 forman el sistema CRISPR y, por otra parte, obtener el fragmento del gen correcto (sin mutación) a través de una reacción en cadena PCR para reparar la mutación. Si el conjunto de esos elementos moleculares se introdujeran en las células de un paciente, se corregiría la mutación y en consecuencia se resolvería la enfermedad de Huntington.

This project is about knowing the mechanisms and usefulness of a genetic editing tool such as CRISPR (*Clustered Regularly Interspaced Short Palindromic Repeats*) with the aim of being able to cure a genetic disease at a theoretical level. For this reason, a set of molecules have been designed in order to correct a real genetic disease. A disease that supposes an actual health issue, that is rare and it's caused by a genetic mutation, specifically a mutation in one single gene, has been selected so that the CRISPR technique can be used to solve it. Huntington's disease was selected, a very serious disease for which there is no cure. This disease affects the nervous system and produces very severe symptoms that lead to the death of the one who suffers it. To achieve the correction of the mutation that causes this disease, the location of the mutation within the specific gene has been specifically sought, in order to later design the necessary molecules. These molecules are, on one hand, the guide RNA fragment complementary to this gene, which together with the CAS9 protein form the CRISPR system and, on the other hand, the correct

gene fragment (without mutation) through a PCR chain reaction to repair the mutation. If all of these molecular elements were introduced into a patient's cells, the mutation would be corrected and Huntington's disease would therefore be resolved.

Índex de continguts

1. Introducció	5
1.1. Motivacions	5
1.2. Objectiu	5
1.3. Metodologia	6
1.4. Agraïments	6
2. Marc teòric i contextual	8
2.1. Malaltia de Huntington	8
2.1.1. La malaltia a nivell genètic i com aquest afecta les cèl·lules i els teixits	8
2.1.1.1. Gen de la huntingtina	8
2.1.1.2. La proteïna	8
2.1.1.3. Variació dels símptomes causats depenent del nombre de repeticions del codó CAG	9
2.1.2. Descripció i símptomes de la malaltia	9
2.1.2.1. A nivell cel·lular	10
2.1.2.2. Canvis macroscòpics	10
2.1.3. Causes	10
2.1.4. Tractaments	11
2.1.5. Tests prenatals	11
2.1.6. Prevalença	11
2.1.7. Locus del gen HTT (gen de la huntingtina)	12
2.1.8. ACMAH (Associació catalana de malalts de Huntington)	12
2.2. Genètica molecular	13
2.2.1. Introducció a la genètica molecular	13
2.2.2. Replicació de l'ADN	14
2.2.3. Transcripció	16
2.2.4. Traducció	18
2.2.5. Les mutacions	21
2.3. PCR	26
2.3.1. Què és una PCR?	26
2.3.1. Com es fa una PCR?	26
2.3.2. Passos per a realitzar una PCR	27
2.4. CRISPR	30
2.4.1. Què és?	30
2.4.2. Complex ARNguia-CAS9	31
2.4.3. Etapes d'interacció del sistema CRISPR-CAS9 amb l'ADN	31
2.4.4. Funcionament de Mecanisme de reparació cel·lular de l'ADN: HDR	32
3. Marc aplicat, disseny d'edició genètica CRISPR per corregir la malaltia de Huntington	35
3.1. Localització del gen	35

3.2. Disseny de la PCR	39
3.3. Com fer el disseny CRISPR	41
3.3.1. Com introduir el material necessari perquè es realitzi el procés a la cèl·lula	44
4. Conclusions	45
5. Referències bibliogràfiques	47

1. Introducció

El meu treball de recerca està enfocat en l'àmbit de la biotecnologia i l'edició genètica, i buscaré crear a través de la tècnica CRISPR un disseny per corregir la malaltia de Huntington.

1.1. Motivacions

La meva motivació a l'hora de fer aquest treball és que trobo molt interessants els camps de la genètica molecular i m'agradaria en un futur proper poder estudiar biotecnologia i dedicar-me a fer recerca amb l'objectiu de poder millorar la qualitat i l'esperança de la vida de les persones que pateixen malalties causades per mutacions genètiques entre altres coses.

Crec que l'enginyeria genètica i les teràpies gèniques són procediments relativament nous, però que en un futur no molt llunyà agafaran molta importància en el món de la medicina i permetran curar malalties que fins ara es consideren incurables.

1.2. Objectiu

L'objectiu principal del meu treball de recerca és trobar la correcció d'una malaltia genètica minoritària, en aquest cas la malaltia de Huntington, amb un disseny teòric de teràpia gènica utilitzant la tècnica de CRISPR. El marc d'aquest treball és fer el disseny teòric d'aquest possible tractament, sense explicar detalladament com introduir les molècules necessàries perquè es donin els processos que s'han de donar a dins la cèl·lula, tot i que es proposa un possible mecanisme d'inserció. Tota la part d'aplicació pràctica queda fora de l'àmbit d'aquest treball, ja que s'hauria de realitzar a un laboratori i requeriria unes instal·lacions i uns recursos molt especialitzats.

Durant la realització d'aquest treball m'he enfrontat a la dificultat que m'he trobat molts conceptes nous, alguns dels quals m'han estat difícils d'entendre, per això he necessitat l'ajuda d'un mentor, - , expert en biotecnologia perquè m'ajudés a entendre els conceptes i a dur a terme el disseny.

1.3. Metodologia

El cos d'aquest treball consta d'un apartat teòric (Marc teòric i contextual) i un apartat pràctic (Marc aplicat). També consta d'un Annex on hi ha les seqüències del gen HTT en forma de nucleòtids i aminoàcids.

La part teòrica del treball l'he dut a terme a través de majoritàriament recerca bibliogràfica. En aquest apartat he informat sobre la Corea de Huntington, sobre genètica molecular i he après com realitzar processos que seran essencials més endavant al fer el disseny com pot ser com es fa una PCR o com funciona el complex ARN^{guia}/CAS9.

L'altre gran apartat d'aquest treball és la part pràctica, on he localitzat el gen HTT gràcies al portal NCBI i on he dissenyat un fragment d'ADN a través d'una PCR i dos sistemes CRISPR per corregir la malaltia de Huntington a nivell teòric.

Les fonts són majoritàriament webs i articles científics. M'he esforçat a procurar que totes les fonts utilitzades en aquest treball fossin fiables.

La font més utilitzada durant la realització d'aquest treball ha estat el portal NCBI, per informar-me tant de les característiques de la malaltia que he tractat en aquest treball com de la seqüència del gen HTT.

1.4. Agraïments

En primer lloc, m'agradaria agrair a - , que ha estat la meva tutora en aquest treball de recerca, per haver-me ajudat en tot el procés, per haver-me ben aconsellat, i per haver mostrat gran interès en tot el que he fet.

També vull agrair especialment a - , que m'ha ajudat a entendre els conceptes més difícils del treball i a explicar-me desinteressadament tots aquells aspectes que han estat essencials per a la realització d'aquest.

I finalment m'agradaria agrair a la meva família pel suport continuu que han mostrat cap a mi i aquest treball.

2. Marc teòric i contextual

2.1. Malaltia de Huntington

2.1.1. La malaltia a nivell genètic i com aquest afecta les cèl·lules i els teixits

La corea de Huntington és una malaltia genètica. La malaltia de Huntington té herència autosòmica dominant. La probabilitat de cada descendent d'heretar el gen responsable de la mutació és del 50%. L'herència és independent del sexe i el gen no se salta generacions.¹

El gen responsable de la malaltia és l'anomenat gen de la huntingtina que trobem a prop del telòmer del braç curt del cromosoma 4.²

2.1.1.1. Gen de la huntingtina

Aquest gen està localitzat al cromosoma 4, a prop del telòmer del braç curt, exactament a la posició 4p16.3.

Després d'aconseguir aïllar el gen es va trobar una connexió amb altres malalties hereditàries: el mecanisme de mutació en totes elles és l'expansió de repeticions de trinucleòtids.

El què determina la presència o no d'aquesta malaltia és el nombre de repeticions del codó CAG, el qual codifica per a l'aminoàcid de la glutamina. Aquests triplets estan situats al primer exó. La proteïna normal és polimòrfica per a un segment de poliglutamina. Les repeticions de triplets varien normalment entre 9 i 36, essent 36 el que es considera un llindar que quan se supera es presenta en la majoria de casos la malaltia³. Per sota d'aquest llindar podem veure casos de nombres intermedis de repeticions que poden suposar una predisposició a patir la malaltia (premutació) a la següent generació.

2.1.1.2. La proteïna

El gen de la huntingtina produeix una proteïna anomenada **huntingtina**.

El gen que codifica per aquesta proteïna està altament conservat i s'expressa a tot el cos. Durant el desenvolupament és quan té un paper essencial, i també en les neurones postmitotiques i altres tipus cel·lulars amb o sense origen neuronal. La proteïna es troba tant al nucli com al citoplasma i està associada a diversos orgànuls i estructures, entre elles la xarxa de microtúbuls.⁴

¹ "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.

² "Huntington's Disease - NORD (National Organization for Rare" <https://rarediseases.org/rare-diseases/huntingtons-disease/>.

³ "OMIM Entry - # 143100 - HUNTINGTON DISEASE; HD." <https://www.omim.org/entry/143100>.

⁴ "Huntingtin - Homo sapiens (Human) - HTT gene & protein - UniProt." <https://www.uniprot.org/uniprot/P42858>.

El coneixement respecte a les funcions de la proteïna és limitat, però es creu que té funcions de transport, endocitosi, autofàgia i de regulació.⁵

A més aquesta proteïna és present en altes concentracions en cèl·lules en divisió i està associada amb altres proteïnes essencials per a la formació i orientació del fus mitòtic. S'ha observat que l'absència de la proteïna huntingtina (o, en aquest cas, de la proteïna defectuosa) provoca la incapacitat d'aquests complexos proteics per orientar el fus.

Un augment de la mida del segment CAG condueix a la producció d'una versió anormalment llarga de la proteïna huntingtina. La proteïna allargada es talla en fragments més petits i tòxics que s'uneixen i s'acumulen a les neurones, alterant les funcions normals d'aquestes cèl·lules. La disfunció i la mort eventual de les neurones en determinades àrees del cervell són la base dels signes i símptomes de la malaltia de Huntington.⁶

2.1.1.3. Variació dels símptomes causats depenent del nombre de repeticions del codó CAG

S'ha trobat que hi ha una correlació negativa molt significativa entre l'edat d'inici de la malaltia, definida com l'edat en què apareixen els primers desordres ben definits, i la longitud de les repeticions CAG. Fins i tot s'ha observat que apareix de manera cada cop més prematura amb cada nova generació que hereta la malaltia d'algun dels pares. Això s'explica perquè l'ADN polimerasa, en replicar la regió de CAG, afegeix algunes repeticions més per defecte; per tant, cada generació presenta un nombre més elevat de repeticions i comencen a experimentar els símptomes de la malaltia a una edat més primerenca.

2.1.2. Descripció i símptomes de la malaltia

George Huntington, va ser qui va descriure els símptomes de la malaltia en un article al 1872 per primer cop. També va ser la primera persona que va identificar el caràcter hereditari d'aquesta.⁷ (Figura 1)

Actualment la OMS en la seva classificació internacional de malalties classifica la malaltia de Huntington (codi G10) dins el bloc de les *atròfies sistèmiques que afecten principalment*



Figura 1

George Huntington (1850-1916)
<https://upload.wikimedia.org/wikipedia/commons/9/9f/Georgehuntington.jpg>

⁵ "The cryo-electron microscopy structure of huntingtin - PubMed." 1 de març. 2018, <https://pubmed.ncbi.nlm.nih.gov/29466333/>.

⁶ "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.

⁷ "On Chorea - Wikisource, the free online library." 30 de des.. 2020, https://en.wikisource.org/wiki/On_Chorea.

el Sistema Nerviós Central. També és coneguda com a corea de Huntington.⁸

2.1.2.1. A nivell cel·lular

L'existència d'aquesta proteïna, quan és anormalment llarga, causa que es talli en fragments més petits i tòxics que s'uneixen i s'acumulen a les neurones, alterant les funcions normals d'aquestes cèl·lules. La disfunció i la mort eventual de les neurones en determinades àrees del cervell són la base dels signes i símptomes de la malaltia.

2.1.2.2. Canvis macroscòpics

Existeixen una gran varietat de símptomes causats per la malaltia de Huntington, però aquesta es caracteritza principalment per moviments musculars ràpids i incontrolables com ara tics o sacsejades musculars (conegut com a corea). Aquest trastorn també provoca una pèrdua de coordinació i canvis de personalitat. A mesura que la malaltia avança la capacitat de parlar es pot veure afectada, la memòria es pot esvaïr i els moviments involuntaris dels músculs i espasmes (corea) es tornen més greus.⁹ En la majoria de casos aquesta demència es desenvolupa gradualment.

Quan els primers símptomes de la malaltia es donen ja en l'edat adulta, que és la forma més comuna d'aquest trastorn (sol aparèixer als trenta o quaranta anys), els primers signes i símptomes poden incloure irritabilitat, depressió, petits moviments involuntaris, mala coordinació i problemes per aprendre nova informació o prendre decisions.

Les persones amb la forma adulta de la malaltia de Huntington solen viure entre 15 i 20 anys després que comencin els símptomes.¹⁰

Una forma menys comuna de la malaltia de Huntington coneguda com a forma juvenil comença a la infància o l'adolescència. També implica problemes de moviment i canvis mentals i emocionals. Els signes addicionals de la forma juvenil inclouen moviments lents i maldestres, caigudes freqüents, parla confusa i baveig. El rendiment escolar disminueix a mesura que les capacitats de pensament i raonament es veuen deteriorades. La malaltia de Huntington juvenil tendeix a progressar més ràpidament que quan els símptomes comencen a l'edat adulta, les persones afectades solen viure entre 10 i 15 anys després que apareguin els símptomes.

2.1.3. Causes

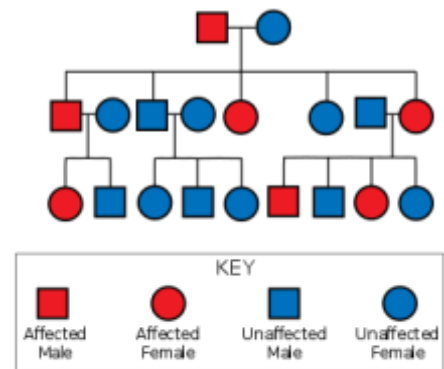
La malaltia de Huntington es dona a causa d'una mutació del gen de la huntingtina i aquest s'hereta com a un tret autosòmic dominant. La probabilitat de cada

⁸ "eCIE-Maps - CIE-10-ES Diagnósticos." <https://eciemaps.mscbs.gob.es/>. S'hi ha accedit el dia 2 de nov.. 2022.

⁹ "Huntington's Disease - NORD (National Organization for Rare" <https://rarediseases.org/rare-diseases/huntingtons-disease/>.

¹⁰ "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.

descendent d'heretar el gen responsable de la mutació és del 50% (només fa falta que un dels dos progenitors presenti la mutació al gen perquè la descendència tingui un 50% per probabilitat de patir la malaltia). L'herència és independent del sexe i el gen no se salta generacions. ¹¹(Figura 2)



2.1.4. Tractaments

Actualment no hi ha cap cura per a la malaltia de Huntington ni cap manera d'evitar que empitjori. Però el tractament i el suport poden ajudar a reduir alguns dels problemes i símptomes que causa, com ara medicaments per a la depressió, els canvis d'humor i els moviments involuntaris, teràpia ocupacional per facilitar les tasques quotidianes, teràpia de la parla i del llenguatge per a problemes d'alimentació i comunicació i fisioteràpia per ajudar amb el moviment i l'equilibri. ¹²

No hi ha fàrmacs que hagin demostrat eficàcia per al deteriorament cognitiu associat a la malaltia de Huntington.

2.1.5. Tests prenatals

Ja que la malaltia de Huntington no disposa actualment de cap tipus de tractament efectiu per curar-la i als pacients de Huntington se'ls sol donar entre 15 i 20 anys d'esperança de vida des de que comencen els símptomes, si una família creu que hi ha la possibilitat de transmetre la malaltia a la seva descendència es permeten fer tests prenatals per identificar si l'embrió podria o no presentar la malaltia. En el cas de que la pugui presentar (perquè algun dels seus progenitors tingui la malaltia), està permès fer una selecció embrionària per assegurar que els fills no presentaran la malaltia, a través de processos in vitro i diagnòstics preimplantacionals.

2.1.6. Prevalença

Es creu que la prevalença actual és aproximadament d'entre 5 i 10 afectats per a cada 100.000 habitants a nivell mundial, fent que sigui considerada una malaltia rara. No es presenta cap diferència en el sexe. Tot i que la malaltia està distribuïda a nivell global, les poblacions de l'est asiàtic i d'Àfrica són menys propenses de patir-la. Es creu de l'origen de la malaltia es troba a Europa, principalment a França,

¹¹ "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.

¹² "Huntington's disease - NHS." <https://www.nhs.uk/conditions/huntingtons-disease/>.

Alemanya o Països Baixos, i posteriorment la malaltia es va dispersar a Amèrica i a Austràlia principalment.

La malaltia de Huntington afecta unes 2.500 persones a l'Estat espanyol i unes 78.000 a tot el món.

2.1.7. Locus del gen HTT (gen de la huntingtina)

El gen de la huntingtina o HTT està situat al cromosoma 4, a prop del telòmer del braç curt, exactament al punt 4p16.3.

Les repeticions de CAG (que són les que determinen si es donarà o no la malaltia) estan localitzades al primer exó, concretament a partir del 196è nucleòtid.

2.1.8. ACMAH (Associació catalana de malalts de Huntington)

L'associació catalana de malalts de Huntington és una entitat de caire autonòmic constituïda l'any 1998, que desenvolupa el seu projecte associatiu al voltant de les persones amb malaltia de Huntington i les seves famílies amb l'objectiu de millorar la seva qualitat de vida.

Aquesta té com a fins estatutaris assessorar i donar suport a les persones i famílies afectades per la malaltia de Huntington, visibilitzar aquest col·lectiu i divulgar i sensibilitzar a la societat sobre la malaltia i promoure la investigació científica.

Proporciona novetats en la investigació de la malaltia de Huntington a través d'articles científics i divulgatius en un llenguatge senzill per poder ser entès per tota la comunitat.¹³

¹³ "ACMAH | Associació Catalana De Malalts de Huntington." <https://www.acmah.org/>. S'hi ha accedit el dia 2 d'oct.. 2022.

2.2. Genètica molecular

2.2.1. Introducció a la genètica molecular

Definició de Gen: Un gen és un fragment d'ADN amb una seqüència de bases nitrogenades que codifica per una molècula que tindrà una funció i que pot ser un ARN o bé una proteïna.

Cada gen està disposat en un ordre fix al llarg d'un cromosoma, els gens determinen l'aparició dels caràcters hereditaris en els éssers vius.¹⁴

L'expressió de la informació genètica:

L'ADN conté informació perquè els aminoàcids s'uneixin i formin les proteïnes, però l'ADN no surt mai del nucli, i la síntesi de proteïnes té lloc en els ribosomes (al citoplasma). Per tant es necessita algun mecanisme que transporti la informació que conté l'ADN als ribosomes per tal que es pugui sintetitzar. Això ho fa l'**ARN missatger**, que es crea en el procés de **transcripció**.

L'ARN missatger conté la informació d'un gen determinat. A través d'un procés anomenat **traducció** es pot sintetitzar una cadena polipeptídica (cadena d'entre 10 i 100 aminoàcids connectats per enllaços peptídics). En aquest procés també hi intervenen l'**ARN ribosòmic** i **ARN de transferència**. Aquest procés té lloc en els ribosomes.

El flux d'informació genètica es pot expressar de la següent manera:



Figura 4

Replicació-Traducció-Transcripció
<https://biobel.files.wordpress.com/2011/01/duplicac3b3-traduccic3b3-transcripcic3b3.pdf>

(proposta inicial de Crick 1970)

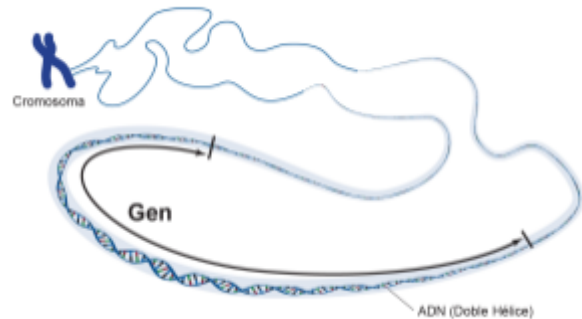


Figura 3

Gen d'ADN bicatenari
<https://www.genome.gov/sites/default/files/tg/en/illustration/gene.jpg>

¹⁴ "¿Qué es un gen?: MedlinePlus Genetics." 28 d'abr. 2021, <https://medlineplus.gov/spanish/genetica/entender/basica/gen/>. S'hi ha accedit el dia 28 de set. 2022.

2.2.2. Replicació de l'ADN

La replicació de l'ADN és el procés pel qual es produeix, a partir d'una molècula original d'ADN, dues còpies idèntiques. Aquest procés es dona un cop per cada generació cel·lular.¹⁵

La replicació de l'ADN és **semiconservativa**, cadascuna de les dues molècules que s'obté després de la replicació té una cadena de l'ADN original i una cadena sintetitzada de nou. Aquesta replicació s'inicia en punts concrets anomenats **orígens de replicació** (en cèl·lules eucariotes hi ha més d'un punt de replicació a cada cromosoma, però a les cèl·lules procariotes només n'hi ha un) durant la fase S. S'anomenen bombolles de replicació.¹⁶ (Figures 5 i 6)

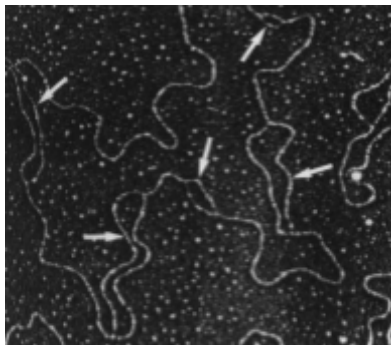


Figura 5

Inici de replicació
<https://fialda.files.wordpress.com/2010/03/burbujasdereplicacic3b3n.jpg?w=640>

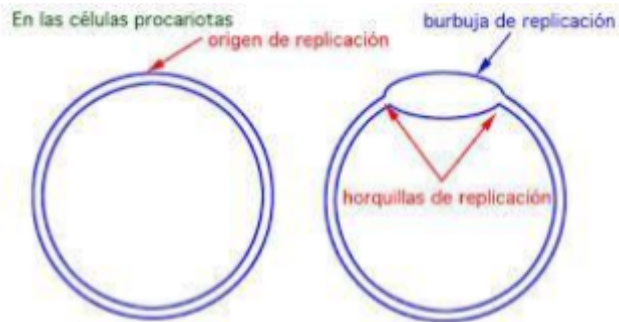


Figura 6

Bombolla de replicació
<https://image.slidesharecdn.com/genticamolecularireplicaci-120117121105-phpapp01/85/gentica-molecular-i-replicaci-35-320.jpg?cb=1659360409>

La replicació és bidireccional a partir dels orígens de replicació, és a dir, avança en els dos sentits fent que sigui més ràpida. Hi ha dues forquilles de replicació.

En cèl·lules procariotes només hi ha un origen de replicació (bidireccional) però en eucariotes n'hi ha més d'un a cada cromosoma.

Parts de la replicació:

- a) Desenrotllament de l'ADN. Per separar les dues hèlix i poder replicar l'ADN, hi ha un enzim (Helicasa) que s'encarrega de fer-ho. Les topoisomerases són proteïnes que estableixen el superenrotllament.

¹⁵ "DNA Replication - National Human Genome Research Institute."

<https://www.genome.gov/genetics-glossary/DNA-Replication>. S'hi ha accedit el dia 28 de set.. 2022.

¹⁶ "Modes of DNA Replication." https://www.mun.ca/biology/scarr/iGen3_03-01.html. S'hi ha accedit el dia 28 de set.. 2022.

- b) La síntesi de les cadenes complementàries. L'enzim anomenat ADN polimerasa III afegeix els nucleòtids complementaris a les cadenes motlle. Només ho pot fer en direcció 5' a 3'. Com que l'ADN polimerasa no pot començar a replicar sol, necessita d'un iniciador o *primer* que comenci a fer-ho ell amb ARN polimerasa.

Un cop col·locat el *primer* una de les cadenes es pot copiar de manera contínua, l'anomenem **cadena de síntesi contínua.**

L'altra cadena no es pot replicar de manera contínua, ja que l'ADN polimerasa només afegeix nucleòtids a l'extrem 3'. La replicació d'**aquesta cadena es fa de manera discontinua.**

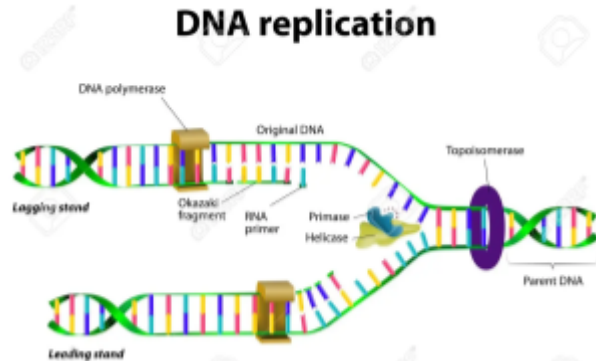


Figura 7

Replicació de l'ADN
<https://atgventure.com/wp-content/uploads/2018/12/mechanism-of-dna-replication-1-300x152.jpg>

- c) Eliminació dels primers i unió dels fragments

Un cop acabada l'etapa de copiar les dues cadenes, l'ADN polimerasa I elimina els ribonucleòtids dels primers (ARN) i els substitueix per desoxiribonucleòtids (ADN).

Un enzim anomenat ligasa uneix els diferents fragments d'ADN de les noves cadenes.

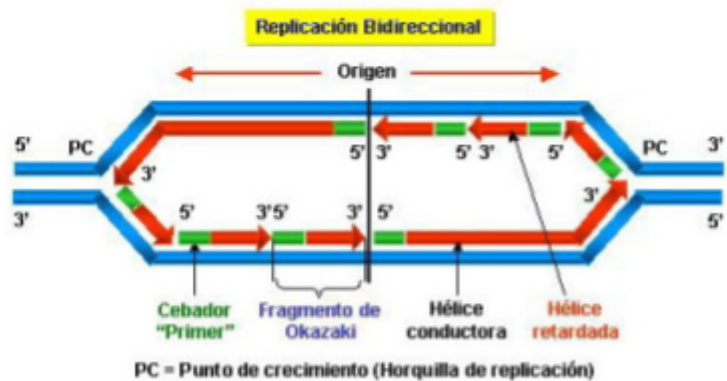


Figura 8

Replicació bidireccional
https://upload.wikimedia.org/wikipedia/commons/thumb/7/7f/Replicaci%C3%B3n_bidireccional.jpg/350px-Replicaci%C3%B3n_bidireccional.jpg

- d) Correcció dels errors

Tot hi que durant molt de temps es pensava que el grau de fidelitat de la replicació de l'ADN era degut a la precisió de l'aparellament entre les bases nitrogenades del filament motlle i el nou filament, es va descobrir que la replicació correcta depèn també de les ADN polimerasa. Aquests enzims tenen capacitat de corregir els propis errors. Si una base nitrogenada s'aparella malament, trenquen l'enllaç, eliminen el nucleòtid incorrecte i el substitueixen per un de nou assegurant així una còpia exacta.

2.2.3. Transcripció

La transcripció és el **procés de síntesi** d'una molècula d'ARN complementària a un fragment d'una molècula d'ADN (gen). En cèl·lules eucariotes aquest procés té lloc al nucli i en cèl·lules procariotes té lloc en el citoplasma.

Generalment, l'ARN format codifica una proteïna i l'anomenem **ARN missatger**. Transportarà la informació fins als ribosomes on tindrà lloc la síntesi de la proteïna que codifica el gen que s'ha transcrit. (Figura 9)

En altres casos el producte final de la transcripció dels gens són cadenes d'ARN amb funcions diferents:

- ARN ribosòmic que formarà part dels ribosomes
- ARN transferidor o de transferència que és el responsable d'activar, transportar i col·locar l'aminoàcid corresponent durant la síntesi de proteïnes o traducció.
- Darrerament s'han descobert uns ARN petits anomenats **ARN d'interferència**. Se sap que intervenen en processos de regulació de l'expressió gènica i el seu estudi ha obert la possibilitat d'utilitzar-los com a fàrmacs per a silenciar gens que no desitgem que s'expressin, ja que provoquen malalties.¹⁷

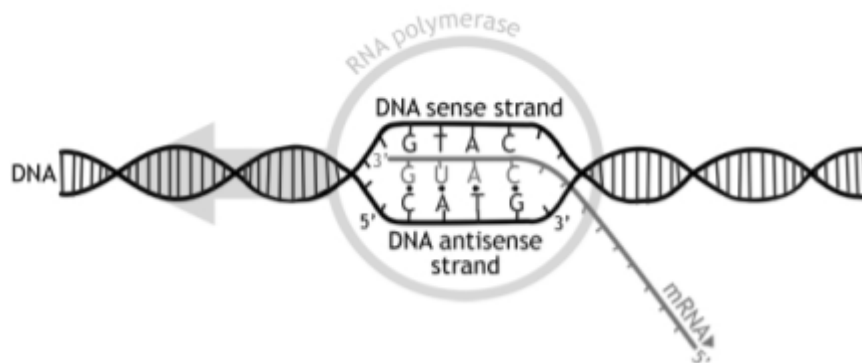


Figura 9

Transcripció de l'ADN

https://upload.wikimedia.org/wikipedia/commons/f/f7/DNA_transcription.jpg

Mecanisme de transcripció:

La transcripció és un procés catalitzat per un sistema enzimàtic. L'enzim més important però és la ARN polimerasa.

- Afegeix ribonucleòtids a un extrem 3' lliure, per tant, la síntesi té lloc en direcció 5'→3'
- Necessita una cadena d'ADN motlle antiparalela al seu sentit de síntesi. Cal que s'obri l'hèlix de l'ADN.

¹⁷ "DNA Transcription | Learn Science at Scitable - Nature."
<https://www.nature.com/scitable/topicpage/dna-transcription-426/>. S'hi ha accedit el dia 28 de set.
2022.

- No necessita *primer*, però només pot iniciar la síntesi quan reconeix la seqüència promotora que hi ha a l'inici de cada gen.
- Es forma un fragment híbrid ADN-ARN inestable que se separa i a mesura que avança la transcripció es restaura l'hèlix d'ADN.
- En l'ARN la base complementària de l'Adenina serà l'Uracil, ja que no conté Timina.¹⁸

En les cèl·lules procariotes només existeix un tipus d'ARN polimerasa, però en les cèl·lules eucariotes n'hi ha tres. Hi ha l'ARN polimerasa I, la qual també s'anomena primasa i és l'encarregada de la síntesi dels *primers*, l'ARN polimerasa II, que fa la síntesi de l'ARN missatger i l'ARN polimerasa III, que s'encarrega de la síntesi dels ARNs transferidor i ribosòmic (també fa la síntesi dels snRNA, *small nuclear RNA*, que s'associen a algunes proteïnes específiques. La seva funció principal és el processament de l'ARN premissatger al nucli¹⁹).

Parts de la transcripció:

a) **Iniciació:** l'ARN polimerasa reconeix el centre promotor, una seqüència curta que assenyalava l'inici del gen.

b) **Elongació o síntesi de la cadena d'ARN:** L'ARN polimerasa va col·locant i unint els ribonucleòtids complementaris i es va sintetitzant la cadena d'ARN complementària a la cadena motlle.

c) **Terminació:** L'ARN polimerasa reconeix una seqüència específica d'ADN que indica l'acabament del gen. S'atura la transcripció i s'allibera l'ARN transcrit. Aquest surt del nucli i realitza la seva funció.

d) **Maduració de l'ARN missatger:** En el procés de maduració de l'ARN missatger l'ARN original transcrit s'afegeix una cua d'Adenines en l'extrem 3'. Sembla que participa en el transport i estabilitat de l'ARN. També hi ha una addició d'un casquet, que consisteix en una guanosina metilada que protegeix l'ARN, a l'extrem 5'. D'aquesta addició se'n diu *capping*.

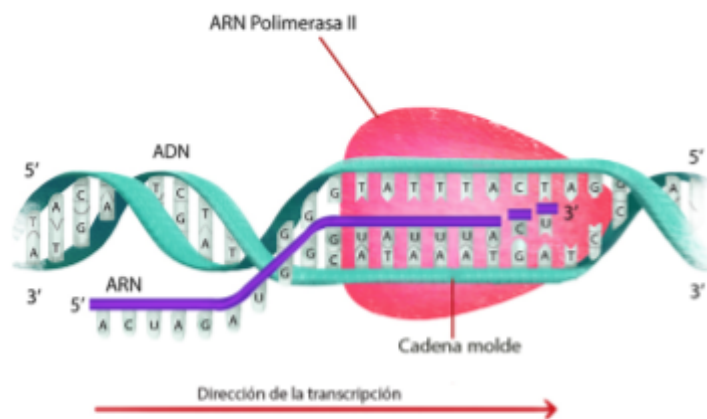


Figura 10

Transcripció de l'ADN

<https://portalacademico.cch.unam.mx/sites/default/files/b1u2oa12p03-21.jpg>

¹⁸ "From DNA to RNA - Molecular Biology of the Cell - NCBI Bookshelf."

<https://www.ncbi.nlm.nih.gov/books/NBK26887/>. S'hi ha accedit el dia 28 de set.. 2022.

¹⁹ "CHEBI:74035 - small nuclear RNA - EMBL-EBI." 30 de gen.. 2015,

<https://www.ebi.ac.uk/chebi/searchId.do?chebiId=CHEBI:74035>. S'hi ha accedit el dia 28 d'ag.. 2022.

Els gens de les cèl·lules eucariotes tenen llargues seqüències no codificants que són transcrites però que cal eliminar abans de la traducció. Aquestes seqüències s'anomenen introns mentre que les seqüències codificadores reben el nom d'exons. Després doncs de la transcripció es produeix un procés de talls i unions que eliminen els introns anomenats *splicing* (Figura 11). La quantitat d'introns en un gen pot ser variable i sovint la longitud d'aquests és major que la dels exons. En l'actualitat s'accepta que l'existència d'introns té un paper clau en la recombinació dins el propi gen, fent que així els exons es puguin combinar de diferent manera i originar proteïnes diferents.

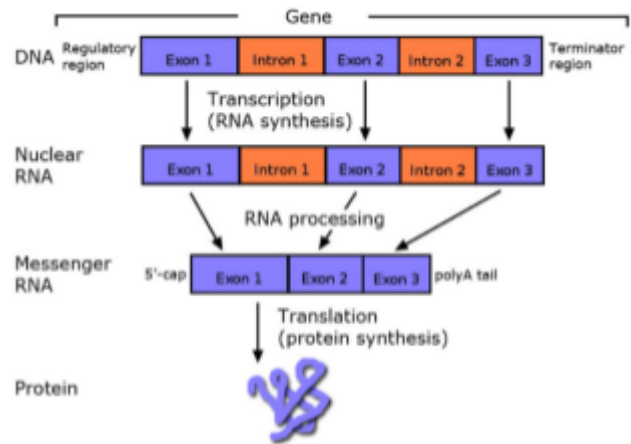


Figura 11

Splicing, gene to protein
https://www.ncbi.nlm.nih.gov/Class/MLACourse/Modules/MolBioReview/images/alternative_splicing.gif

3.2.4. Traducció

La traducció és el procés mitjançant el qual la informació codificada a la seqüència de nucleòtids d'un ARN missatger és descodificada i utilitzada per a dirigir la unió d'una seqüència d'aminoàcids per formar una proteïna. Aquest procés té lloc al citoplasma cel·lular i en concret als ribosomes.

La informació en l'ARN missatger es troba codificada en un llenguatge de quatre lletres corresponents a les quatre bases nitrogenades (A, G, C, U) mentre que la seqüència d'una proteïna es construeix a partir de 20 aminoàcids (AA) diferents. Això vol dir que cal un canvi de llenguatge.

El conjunt de regles que determina la correspondència entre la seqüència de bases nitrogenades de l'ARN missatger i la seqüència d'aminoàcids d'una proteïna s'anomena **CODI GENÈTIC**.²⁰

		Segunda letra					
		U	C	A	G		
Primera letra	U	UUU } Phe UUC } UUA } Leu UUG }	UCU } Ser UCC } UCA } UCG }	UAU } Tyr UAC } UAA Alto UAG Alto	UGU } Cys UGC } UGA Alto UGG Trp	U	C
	C	CUU } CUC } Leu CUA } CUG }	CCU } Pro CCC } CCA } CCG }	CAU } His CAC } CAA } Gln CAG }	CGU } Arg CGC } CGA } CGG }	U	C
	A	AUU } AUC } Ile AUA } AUG Met	ACU } Thr ACC } ACA } ACG }	AAU } Asn AAC } AAA } Lys AAG }	AGU } Ser AGC } AGA } Arg AGG }	U	C
	G	GUU } Val GUC } GUA } GUG }	GCU } Ala GCC } GCA } GCG }	GAU } Asp GAC } GAA } Glu GAG }	GGU } Gly GGC } GGA } GGG }	U	C
						A	G

Figura 12

Codi genètic
<https://2.bp.blogspot.com/-JFtm2UvjQLk/XEYQm2uvDJI/AAAAAAAAAHw/rwyiJnPPgEq2SfdLyb9Re1GvBHc2VfDQCLcBGAs/s640/a736126e791afe019d17e4d1a1e6a10eff462cf5.png>

²⁰ "Genetic Code - National Human Genome Research Institute."
<https://www.genome.gov/genetics-glossary/Genetic-Code>. S'hi ha accedit el dia 28 de set.. 2022.

Característiques del codi genètic: ²¹

- Cada tres bases nitrogenades (Codó) codifiquen per un aminoàcid.
- El codi genètic és degenerat. Hi ha 64 codons i només 20 aminoàcids. Això implica que la majoria d'AA venen codificats per més d'un codó.
- El codi genètic no és ambigu. Això vol dir que un codó sempre codifica pel mateix aminoàcid.
- El codi genètic és (quasi) universal. Tots els éssers vius i també els virus utilitzen el mateix codi genètic. (Actualment s'han trobat algunes excepcions dels codons en ARN missatgers mitocondrials i dels ARN missatgers d'alguns protoctists i fongs).
- El triplet AUG codifica per un aminoàcid (la metionina) però també és el codó d'inici.
- Existeixen tres codons que codifiquen *STOP* (aturada).

Mecanisme de traducció:

La traducció és un procés complex on hi participen els tres tipus d'ARN: ARN missatger, ARN transferidor, ARN ribosòmic.

- ARN missatger: És la còpia d'un fragment d'ADN o gen. Transporta la informació del nucli als ribosomes on tindrà lloc la traducció
- ARN de transferència o transferidor: S'uneixen els aminoàcids de manera específica en funció de quin és el seu anticodó. Són molècules adaptadores que col·loquen l'aminoàcid corresponent en funció de quin és el codó de l'ARN missatger. Els ARN transferidors presenten un **ANTICODÓ** que s'aparella amb el **CODÓ** complementari de l'ARN missatger.

- ARN ribosòmic: Juntament amb una sèrie de proteïnes formarà els ribosomes. Els ribosomes estan formats per dues subunitats, una petita i una gran. A la subunitat gran hi ha tres llocs per a l'aparellament entre codó i ARN transferidor.

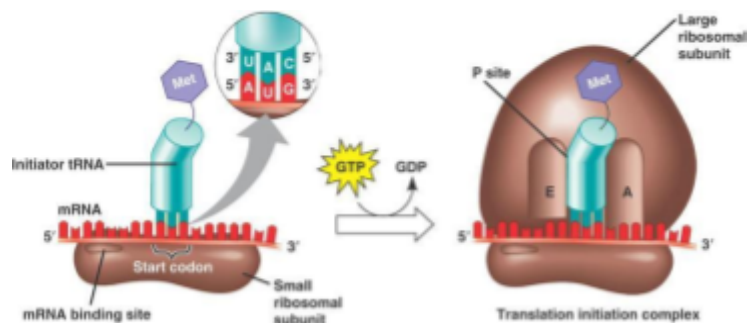


Figura 13

Traducció I
<https://cdn.kastatic.org/ka-perseus-images/7ee48fae4da48ce56d9c4bc4929c355c22138684.png>

²¹ "Biosíntesis de proteïnes: Transcripció i traducció." 3 de nov.. 2017, <https://cn-tec.blogspot.com/2017/11/biosintesis-de-proteines-transcripcio-i.html>. S'hi ha accedit el dia 28 de set.. 2022.

El següent ARN transferidor entra al ribosoma ocupant la posició A i es repeteix el mateix procés, de manera que es van unint els diversos aminoàcids fins a completar tota la seqüència.

- d) Terminació: Al final de l'ARN missatger hi ha el codó d'aturada (*STOP*). Quan la posició A és ocupada pel factor de terminació s'hi uneix el factor d'alliberament que provoca l'alliberament de la cadena polipeptídica i de l'últim ARN transferidor.
- e) Maduració de la proteïna: Un cop finalitzada la traducció les cadenes polipeptídiques es repleguen adoptant les diverses estructures i pateixen també una sèrie de canvis posttraduccional. Aquests poden ser la unió de grups prostètics, la modificació d'alguns aminoàcids, l'eliminació de l'aminoàcid inicial (la metionina), la formació de ponts disulfur que intervenen en la conformació de la proteïna i l'associació amb altres cadenes per formar proteïnes més grans amb estructura quaternària.

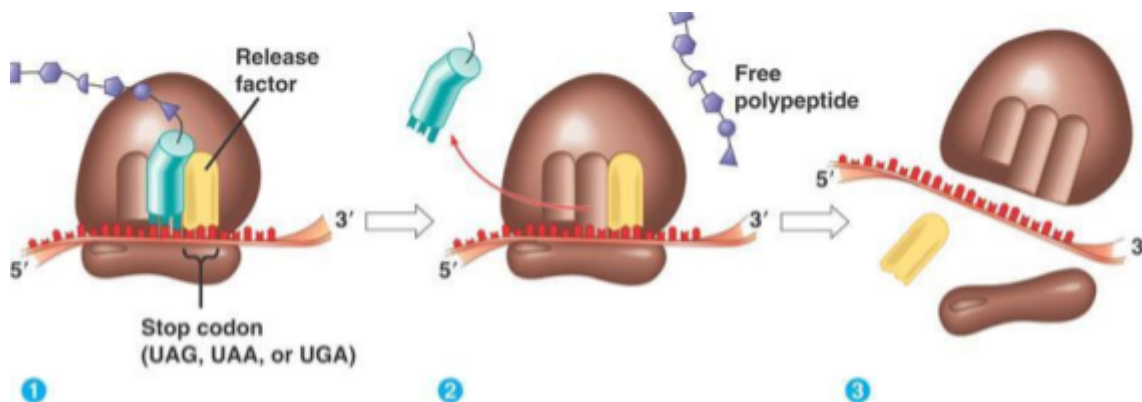


Figura 15

Traducció III

<https://cdn.kastatic.org/ka-perseus-images/7ee48fae4da48ce56d9c4bc4929c355c22138684.png>

2.2.5. Les mutacions ²²

Les mutacions són **alteracions a l'atzar del material genètic** (DNA, alguns virus RNA). La majoria produeixen alteracions fisiològiques i/o anatòmiques, i aquestes poden ser letals. La majoria són recessives i queden amagades, però es transmeten a la descendència (En el cas de la malaltia de Huntington no és recessiva, sinó que

²² "LES MUTACIONS SÓN CANVIS O ALTERACIONS DEL MATERIAL"
<https://slideplayer.es/slide/14021588/>. S'hi ha accedit el dia 28 de set.. 2022.

és dominant). Les mutacions aporten variabilitat a la població, sobre la qual actua la selecció natural, permetent l'evolució de les espècies.

Tipus de mutacions:

a) Segons la cèl·lula afectada

Mutacions somàtiques: afecten a les cèl·lules somàtiques, dels teixits. Si es produeixen en oncogens produeixen processos tumorals i càncers.

Mutacions germinals: en cèl·lules reproductives. Afecten al nou organisme generat a partir d'aquestes cèl·lules.

b) Segons l'origen de la mutació:

Mutacions naturals: apareixen espontàniament. Taxa de mutació natural en humans: 1 gen / 50.000 gens (25.000 gens = 23 cromosomes, cada generació 1 gen mutat)

Mutacions induïdes: provocades artificialment per agents mutàgens (radiacions, substàncies químiques, etc.)

c) Segons l'extensió del material genètic afectat:

Mutacions gèniques: Són alteracions de la seqüència dels nucleòtids. Es produeixen per un error en la replicació de l'ADN, les quals poden ser mutacions per substitució de bases (20%) o bé per pèrdues o insercions de nucleòtids (80%). Un exemple de mutació gènica és l'albinisme.

Les **mutacions per substitució** de bases es donen quan hi han canvis de nucleòtids per uns altres, ja siguin canvis transicionals (canvi d'una base púrica per una altra púrica (A i G) o canvi d'una base pirimidínica per una pirimidínica (T i C)) o causades per canvis transversionals (canvi d'una base pirimidínica per una

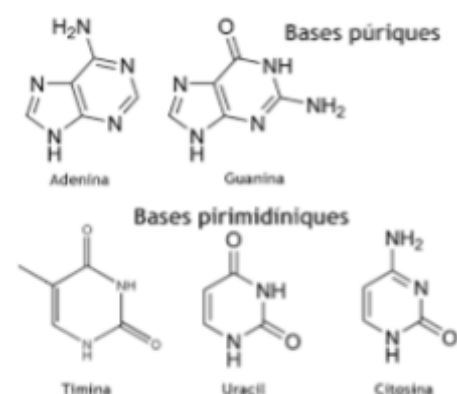


Figura 16

Bases Nitrogenades
https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQv_FPU5Iz5uCF1vXFJLpe_FvW1lptdhCK1MA&usqp=CAU

púrica i al revés) (Figura 17). Les conseqüències d'aquest tipus de mutació varien depenent d'on es troba situada, si està al triplet d'*STOP* o pertany al centre actiu d'un enzim, les conseqüències seran molt greus, perquè o bé la proteïna serà incompleta o l'enzim que generarà serà antifuncional. Si el triplet no està a cap dels casos anteriors, la mutació serà silenciosa.

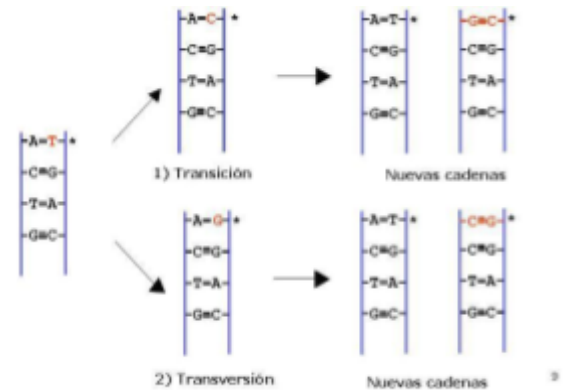


Figura 17
Mutacions transicionals i transversionals
https://slideplayer.es/14071666/86/images/slide_15.jpg

Les **mutacions per pèrdua** (delecions) o **insercions** (addicions) de **nucleòtids** tenen conseqüències greus, ja que provoquen corriments en l'ordre de lectura i l'alteració de tots els triplets següents.

Mutacions cromosòmiques: Canvis en l'estructura interna dels cromosomes. Les mutacions cromosòmiques poden ser de deleción, duplicación, inversión o de translocación.

Els canvis per **deleción** consisteixen en la pèrdua d'un fragment del cromosoma. Si el fragment que s'ha perdut del cromosoma contenia una gran quantitat de gens, les conseqüències poden ser molt greus, i fins i tot poden arribar a ser letals. (Figura 18)

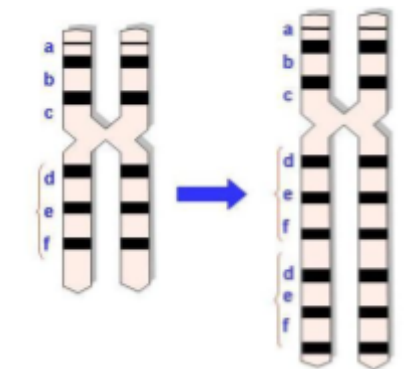


Figura 18
Deleción
<https://slideplayer.es/slide/14021588/>

Quan els canvis són per **duplicación** es dona la repetición d'un fragment de l'ADN, fent que aquest sigui

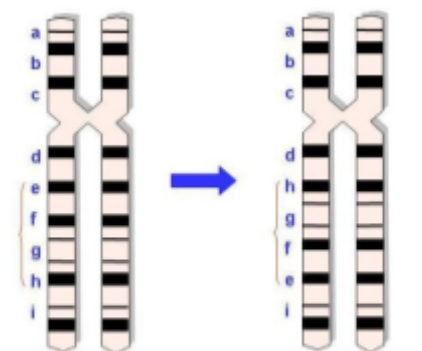


Figura 19
Duplicación
<https://slideplayer.es/slide/14021588/>

anormalment llarg. La duplicació d'aquest material genètic permet gràcies a més mutacions posteriors que es creïn nous gens que permetin l'evolució de les espècies. (Figura 19)

Les **inversions** són canvis de sentit d'un fragment de cromosoma les quals normalment no comporten perjudicis a l'individu que les pateix, però si durant la meiosi es produeix un entrecreuament dins de la inversió, les gàmetes tindran manca de gens. (Figura 20)

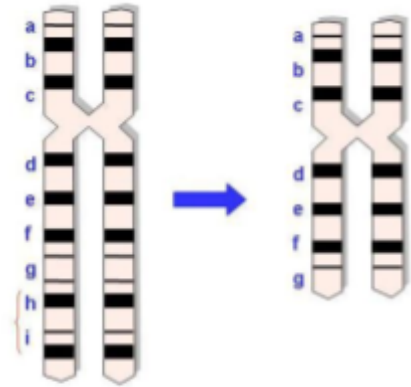


Figura 20

Inversió

<https://slideplayer.es/slide/14021588/>

Les **translocacions** es donen quan un fragment d'un cromosoma es troba en un altre cromosoma, sigui aquest homòleg o no. Les conseqüències d'aquest tipus de mutació no són perjudicials, ja que es mantenen els gens en el genoma. Però per a la descendència sí que poden ser perjudicials, ja que poden heretar un cromosoma incomplet o amb duplicacions. Si les gàmetes són molt defectuoses, hi haurà problemes d'infertilitat. (Figura 21)

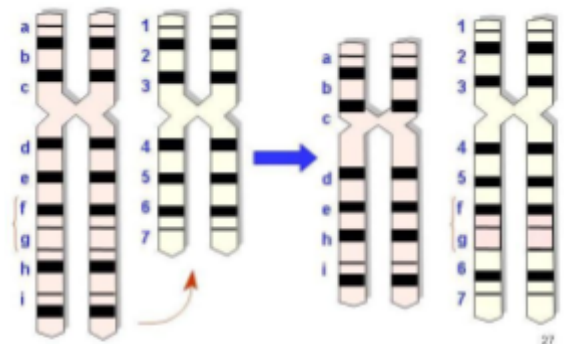


Figura 21

Translocació

<https://slideplayer.es/slide/14021588/>

Mutacions genòmiques: Són alteracions del nombre de cromosomes propi d'una espècie. N'hi ha de dos tipus: aneuploïdia i euploïdia.

L'aneuploïdia és l'alteració en el nombre normal (2) d'exemplars d'un o més tipus de cromosomes. Se'n diu nul·lisomia si no n'hi ha cap, monosomia si només n'hi ha un, trisomia si n'hi ha tres (n'és un exemple el síndrome de Down, que consisteix en una trisomia del cromosoma 21) i tetrasomia si n'hi ha quatre. Les causes d'aquest tipus de mutacions poden ser per Fusió cèntrica (unió de dos

cromosomes), fissió cèntrica (escissió d'un cromosoma en dos) o bé per una segregació errònia durant la meiosi (distribució errònia de les cromàtides homòlogues entre les cèl·lules filles durant la meiosi).

L'euploïdia és l'alteració en el nombre normal de dotacions cromosòmiques i pot ser monoploïdia (una sola dotació cromosòmica) o poliploïdia (més d'un parell de cada cromosoma).

Els agents mutàgens:

Els agents mutàgens són factors que augmenten sensiblement la freqüència normal de mutació, i aquests poden ser físics o bé químics.

Els **agents físics** poden ser radiacions no ionitzants (com els rajos violeta) o bé poden ser ionitzants (com els rajos gamma i radiacions nuclears).

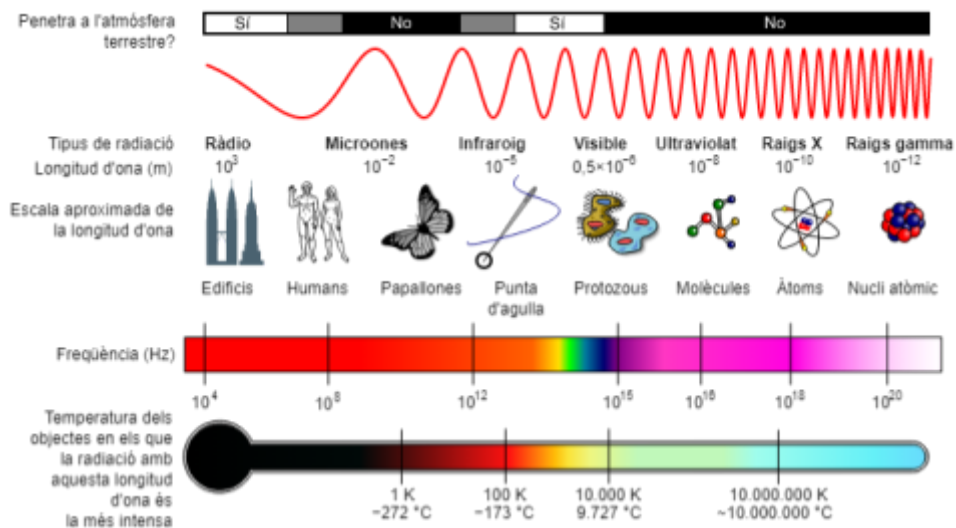


Figura 22

Espectre electromagnètic

https://upload.wikimedia.org/wikipedia/commons/thumb/f/f9/EM_Spectrum_Properties_ca.svg/1920px-EM_Spectrum_Properties_ca.svg.png

Quan parlem **d'agents químics** ens referim a substàncies químiques que reaccionen amb l'ADN.²³

²³ "70. Els agents mutàgens - SlideShare." 2 d'abr.. 2013, <https://www.slideshare.net/daniribo/70-els-agents-mutgens>. S'hi ha accedit el dia 28 de set.. 2022.

2.3. PCR

2.3.1. Què és una PCR?

La reacció en cadena de la polimerasa o PCR (*Polymerase Chain Reaction*), és una tècnica que s'utilitza en biologia molecular per obtenir un gran nombre de còpies d'un fragment d'ADN particular partint d'un mínim; en teoria només cal partir d'una sola còpia d'aquest fragment original, o motlle ²⁴.

2.3.1. Com es fa una PCR?

Per dur a terme una PCR es necessita una mostra d'ADN que contingui el fragment del qual se'n volen fer còpies. Altres elements necessaris són: clorur de magnesi (que és un cofactor essencial de la polimerasa), ADN polimerasa (que catalitza la reacció de polimerització), *primers* (seqüències monocatenàries curtes d'ADN que reconeixen el tros que s'ha de duplicar i generalment s'utilitza un *primer forward* i un *primer reverse* per marcar l'inici i delimitar el final del fragment que es vol duplicar). També són necessaris nucleòtids per dur terme la replicació de les cadenes. Perquè es puguin fer les PCR es necessita un medi òptim, normalment es tracta d'un rang específic de temperatura (que s'aconsegueix utilitzant un termociclador, on s'introdueixen tots els materials esmentats anteriorment). (Figura 23)

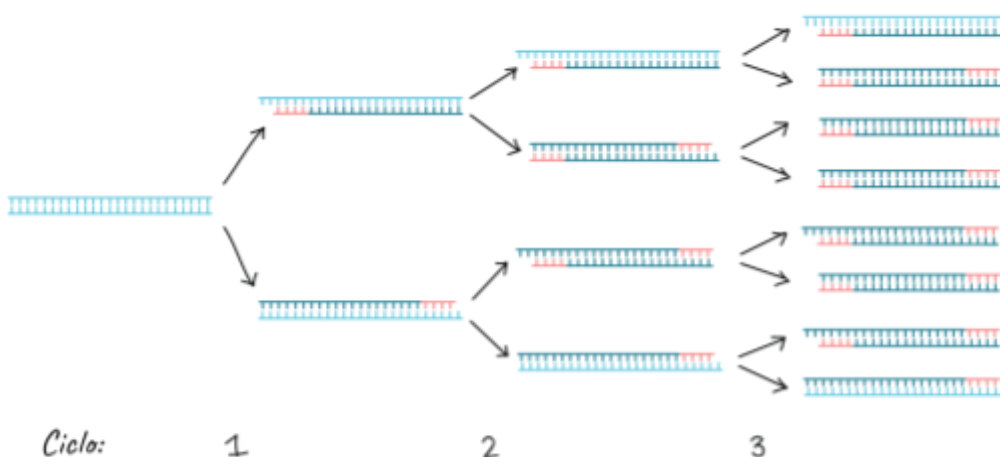


Figura 23

Reacció en cadena PCR

<https://cdn.kastatic.org/ka-perseus-images/135a276c41b2bbd4eb13f40cac88ecd9510e3d56.png>

²⁴ "History of Polymerase Chain Reaction (PCR) - News-Medical.Net." [https://www.news-medical.net/life-sciences/History-of-Polymerase-Chain-Reaction-\(PCR\).aspx](https://www.news-medical.net/life-sciences/History-of-Polymerase-Chain-Reaction-(PCR).aspx). S'hi ha accedit el dia 1 de jul.. 2022.

2.3.2. Passos per a realitzar una PCR ²⁵

El primer pas per dur a terme una PCR és **desnaturalitzar** la molècula d'ADN. Per desnaturalitzar-la es fa pujar la temperatura del termociclador fins a **92°C** en la majoria dels casos, ja que aquesta temperatura és suficient per separar les dues cadenes d'ADN sense que els enzims que s'utilitzen durant el procés tinguin una pèrdua d'activitat.

Quan les dues cadenes ja estan separades els *primers* s'aparellen a la seva cadena corresponent per complementarietat de bases en un procés anomenat **hibridació** i durant aquest temps es sol baixar la temperatura fins als **55-65°C**.

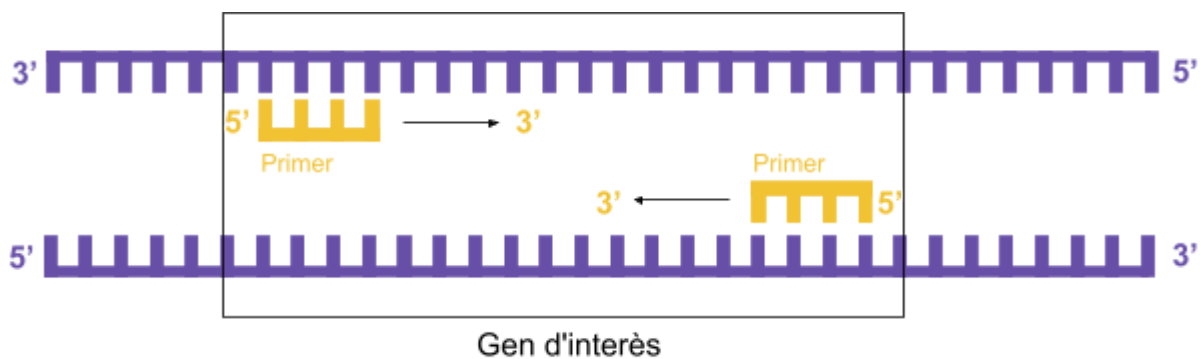


Figura 24
Etapas de la PCR 1
Font pròpia

²⁵ "Reacción en cadena de la polimerasa (PCR) - Khan Academy."
<https://es.khanacademy.org/science/ap-biology/gene-expression-and-regulation/biotechnology/a/polymerase-chain-reaction-pcr>. S'hi ha accedit el dia 28 de set.. 2022.

A continuació, durant l'etapa de **polimerització**, en la qual la temperatura s'ha de tornar a pujar fins a 72°C, l'ADN polimerasa catalitza simultàniament la síntesi de les dues cadenes d'ADN utilitzant els nucleòtids que hi ha dispersos en el medi. Així aconseguim duplicar el nombre de cadenes que teníem a l'inici.

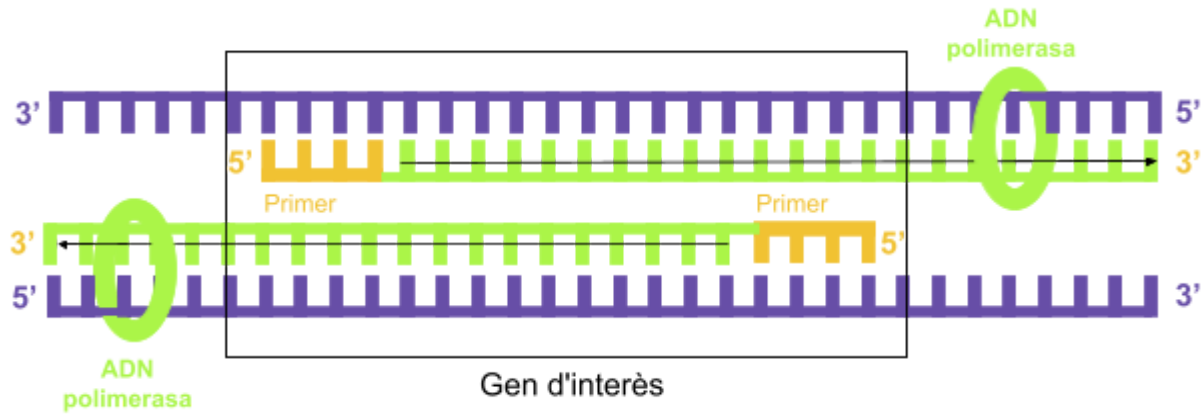


Figura 25
Étapes de la PCR 2
Font pròpia

El procés anterior es torna a repetir obtenint així trossos de cadenes només de la zona d'interès.

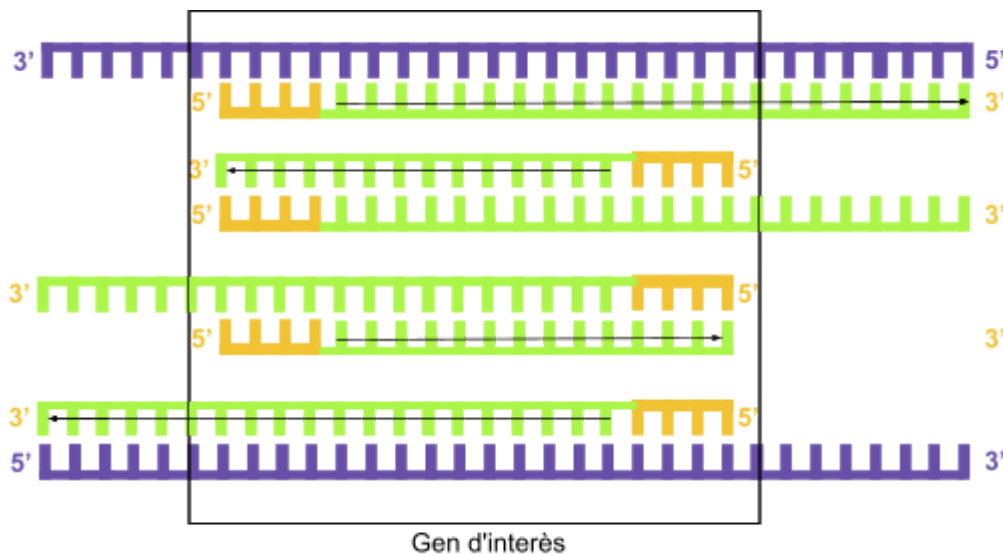


Figura 26
Étapes de la PCR 3
Font pròpia

Quan tornem a repetir el procés per tercer cop aconseguim dobles cadenes d'ADN de només el gen que ens és d'interès, i és aquest el que llavors es començarà a replicar en cadena aconseguint un gran nombre de còpies del gen.

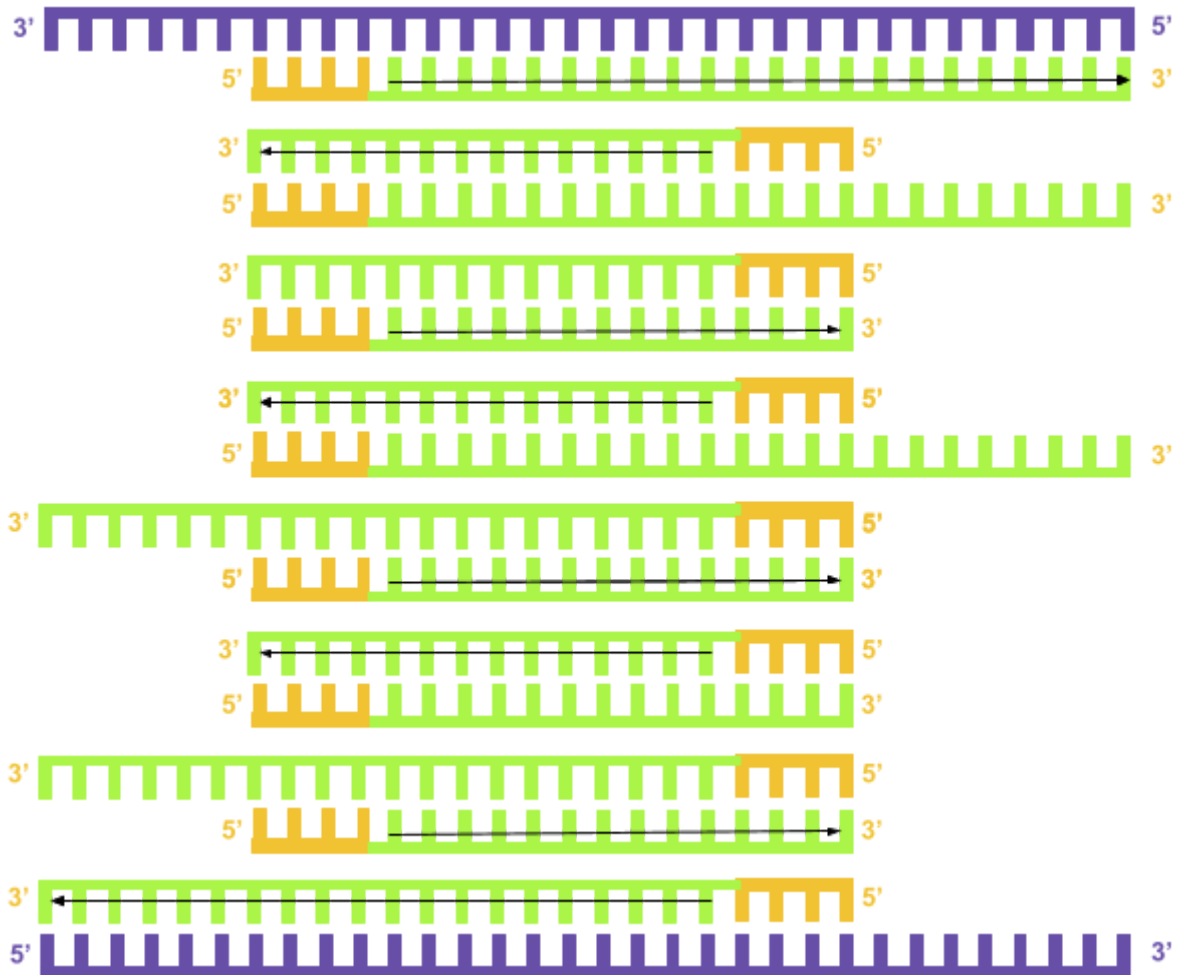


Figura 27
Étapes de la PCR 4
Font pròpia

2.4. CRISPR

2.4.1. Què és?

CRISPR és l'acrònim de la descripció en anglès d'una estructura genòmica trobada en bacteris: *Clustered Regularly Interspaced Short Palindromic Repeats*.

El CRISPR-Cas9 es va adaptar d'un sistema d'edició del genoma natural que alguns bacteris utilitzen com a defensa immunitària. Quan s'infecten amb un virus, els bacteris capturen petits fragments d'ADN dels virus invasors i els insereixen al propi ADN.²⁶

L'estructura CRISPR té funcions de defensa contra patògens i actua com un sistema immunitari en les cèl·lules bacterianes de les quals prové.²⁷

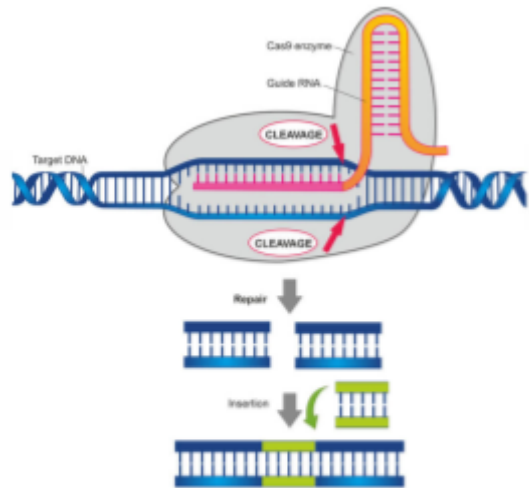


Figura 28

Complex ARNGuia-CAS9
<https://chilmedia.org/v2/media/f6b004ea-d896-4254-b5b8-7f02f05ef289.jpg>

Aquest conjunt de la Cas9 i l'ARN guia són utilitzats per "editar" o "corregir" el genoma de qualsevol cèl·lula. Això inclou, és clar, les cèl·lules humanes. Faria la funció com d'unes tisores moleculars que són capaces de tallar qualsevol molècula d'ADN fent-ho a més d'una manera molt precisa i totalment controlada²⁸.

El sistema CRISPR-Cas9 ha generat molt d'entusiasme a la comunitat científica perquè és més ràpid, més barat, més precís i més eficient que altres mètodes d'edició del genoma.²⁹

El CRISPR es va descobrir que es podia utilitzar com una eina molecular útil al laboratori l'any 2012 gràcies a les doctores Emmanuelle Charpentier a la Universitat d'Umeå i Jennifer Doudna, a la Universitat de Califòrnia a Berkeley.³⁰

²⁶ "¿Qué son la edición del genoma y CRISPR-Cas9? - MedlinePlus." 13 de juny. 2022, <https://medlineplus.gov/spanish/genetica/entender/investigaciongenomica/ediciondelgenoma/>. S'hi ha accedit el dia 18 de set.. 2022.

²⁷ "¿Qué es la tecnología CRISPR? - Bayer." 12 d'abr.. 2022, <https://www.bayer.com/es/es/blog/espana-que-es-la-tecnologia-crispr>. S'hi ha accedit el dia 18 de set.. 2022.

²⁸ "¿Qué es la tecnología CRISPR/Cas9 y cómo nos cambiará la vida?." <https://www.dciencia.es/que-es-la-tecnologia-crispr-cas9/>. S'hi ha accedit el dia 18 de set.. 2022.

²⁹ "¿Qué son la edición del genoma y CRISPR-Cas9? - MedlinePlus." 13 de juny. 2022, <https://medlineplus.gov/spanish/genetica/entender/investigaciongenomica/ediciondelgenoma/>. S'hi ha accedit el dia 18 de set.. 2022.

³⁰ "¿Qué es la tecnología CRISPR/Cas9 y cómo nos cambiará la vida?." <https://www.dciencia.es/que-es-la-tecnologia-crispr-cas9/>. S'hi ha accedit el dia 18 de set.. 2022.

2.4.2. Complex ARNGuia-CAS9

ARN guia: Es tracta d'una cadena senzilla d'ARN, aproximadament d'uns 100 nucleòtids o més. Aquest està format per dues regions. El 80% té la funció de mantenir-se acoblada a la proteïna CAS9, que està situat a l'extrem 3' de la cadena, i la resta és la part variable, que és complementària al tros de l'ADN sobre el qual actuarà i que es troba en l'extrem contrari, el 5'. Aquesta última s'ha de dissenyar i sintetitzar per cada cas específicament en un laboratori, mentre que la resta de la molècula (i la proteïna CAS9) són vàlides per a qualsevol mena d'experiment.

Proteïna CAS9: És una proteïna gran, la qual està formada per dos dominis, un és el d'unió a l'ARN guia i l'altre domini és el que manté una interacció amb l'ADN diana i que a més a més té la capacitat de tallar l'ADN en dos punts concrets. La proteïna Cas9 és un enzim d'endonucleases que és capaç de tallar l'ADN de doble cadena.

2.4.3. Etapes d'interacció del sistema CRISPR-CAS9 amb l'ADN

En primer lloc, la part variable de l'ARN guia s'uneix per complementaritat de bases entre els nucleòtids a la cadena motlle de l'ADN diana, que conté la seqüència errònia que es vol modificar, així la proteïna CAS9 es troba en el lloc exacte on actuarà.

L'ARN guia s'orienta gràcies al PAM (*Protospacer Adjacent Motif*), el qual consisteix en una seqüència de 3 nucleòtids que es trobin a l'inici d'on la CAS9 ha de tallar. La seqüència del PAM sempre és 5'-NGG-3' on N és qualsevol de les 4 possibles bases nitrogenades.

A continuació, un cop la proteïna CAS9 entra en contacte amb l'ADN, aquesta trenca els enllaços fosfodièster de les dues cadenes de l'ADN, separant l'ADN en dues parts. Posteriorment el complex es separa de l'ADN. Llavors la cèl·lula, gràcies al mecanisme anomenat HDR (*Homology Directed Repair*), repara automàticament aquest tall. Si durant aquest procés el sistema troba una seqüència homòloga que conté el tros que es vol inserir en el gen (aconseguit anteriorment fent una PCR del gen correcte) aquest s'utilitzarà automàticament per reparar el tall de l'ADN, així aconseguint modificar el gen.³¹

³¹ "Explaining CRISPR-CAS9 System- Step by Step - Genetic Education." 3 d'ag.. 2020, <https://geneticeducation.co.in/explaining-crispr-cas9-system-step-by-step/>. S'hi ha accedit el dia 18 de set.. 2022.

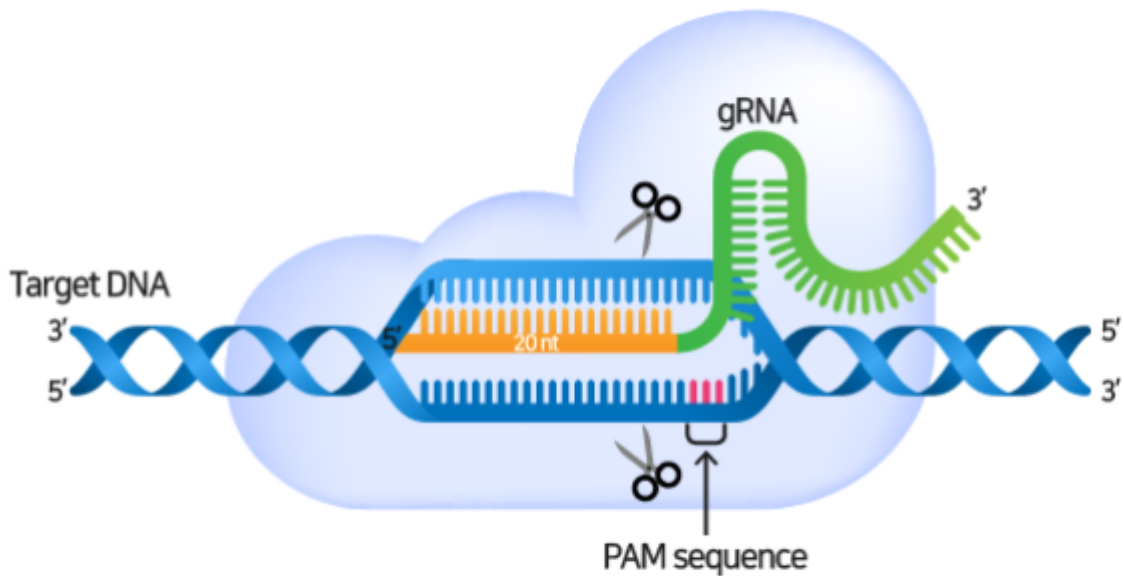


Figura 29

CRISPR
<https://www.bioneer.co.kr/images/products/CRISPR/CRISPRCas9system.png>

2.4.4. Funcionament de Mecanisme de reparació cel·lular de l'ADN: HDR

El sistema HDR entra en funcionament quan detecta un dany en l'ADN, en aquest cas de tall causat pel CRISPR, i és capaç de corregir-ho automàticament utilitant la còpia correcta que hem aconseguit mitjançant la PCR.³²

Pas A



Figura 30
 Etapes de l'HDR 1
 Font pròpia

La doble cadena blava representa l'ADN que el complex RNAGuia-CAS9 ja ha tallat, i la cadena rosa és la còpia del fragment de l'ADN que es vol corregir, el qual s'ha

³² "Stimulation of CRISPR-mediated homology-directed repair by an" 30 de jul.. 2019, <https://www.nature.com/articles/s41467-019-11105-z>. S'hi ha accedit el dia 18 de set.. 2022.

obtingut anteriorment fent una PCR. Aquesta és introduïda dins la cèl·lula mitjançant un vector.

Pas B

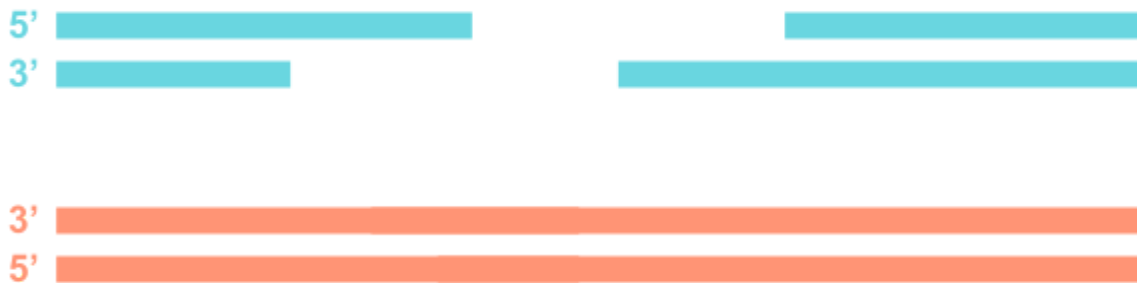


Figura 31
Étapes de l'HDR 2
Font pròpia

Uns enzims que formen part d'aquest procés (exonucleases), eixamplen la zona de la ruptura, deixant així prou espai perquè les dues molècules s'encreuin.

Pas C

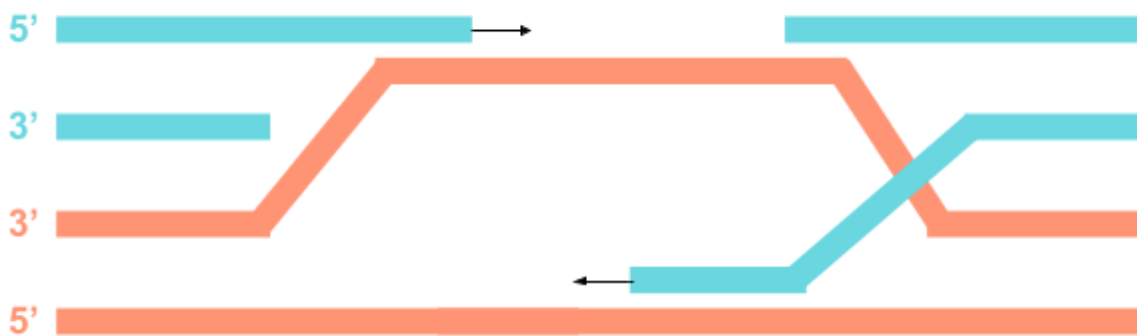


Figura 32
Étapes de l'HDR 3
Font pròpia

Es produeix l'encreuament entre les dues molècules (*crossover*). Com que les dues molècules tenen una part complementària aquestes s'uneixen temporalment permetent que es pugui restaurar la cadena introduint així la versió correcta del gen mutat.

Pas D



Figura 33
Étapes de l'HDR 4
Font pròpia

Es restaura la doble cadena trencada amb dues cadenes senzilles noves complementàries a la cadena rosa (fragments liles).

Pas E



Figura 34
Étapes de l'HDR 5
Font pròpia

La molècula blava ha quedat corregida utilitzant la seqüència correcta que portava la rosa, i en aquesta hi queda un fragment original de la cadena blava, la qual s'ha intercanviat durant el procés.

3. Marc aplicat, disseny d'edició genètica CRISPR per corregir la malaltia de Huntington

3.1. Localització del gen

Es pot aconseguir la seqüència del **gen de la huntingtina** (necessària per poder dissenyar una PCR i dos sistemes CRISPR) utilitzant el portal NCBI³³ (*National Center for Biotechnology Information*).

El Centre Nacional per a la Informació Biotecnològica forma part de la *National Library of Medicine* dels Estats Units. Aquest centre és una font d'informació fonamental en el camp de la biologia molecular.

El NCBI s'encarrega d'emmagatzemar i actualitzar constantment la informació referent a seqüències genòmiques i genètiques en el GenBank, proporciona una àmplia gamma d'informació sobre biotecnologia, biomedicina, genètica i bioquímica a través d'articles científics. Per altra banda, recopila informació sobre malalties genètiques humanes i ofereix a més a més algunes eines bioinformàtiques per analitzar seqüències d'ADN, ARN i proteïnes.

Totes les bases de dades del NCBI estan disponibles a internet de manera gratuïta, per tant està a l'abast de tothom qui tingui accés a internet.

Per trobar la seqüència del gen de la huntingtina a través del portal NCBI s'han de realitzar els següents passos:

En primer lloc, s'ha d'anar a "All Database" (tota la base de dades), seleccionar l'opció de "gene" (Gen) i introduir al buscador HTT (abreviació per al gen de la huntingtina).



Figura 35

Primera pàgina del portal de l'NCBI
Font Pròpia

³³ "Home - PMC - NCBI." <https://www.ncbi.nlm.nih.gov/pmc/>. S'hi ha accedit el dia 30 d'oct.. 2022.

Search results

Items: 1 to 20 of 1877 << First < Prev Page 1 of 94 Next > Last >>

See also 6 discontinued or replaced items.

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> HTT ID: 3064	huntingtin [<i>Homo sapiens</i> (human)]	Chromosome 4, NC_000004.12 (3074681..3243960)	HD, IT15, LOMARS	613004
<input type="checkbox"/> Htt ID: 15194	huntingtin [<i>Mus musculus</i> (house mouse)]	Chromosome 5, NC_000071.7 (34919084..35069878)	C430023I11Rik, Hd, Hdh, IT15	
<input type="checkbox"/> Htt ID: 29424	huntingtin [<i>Rattus norvegicus</i> (Norway rat)]	Chromosome 14, NC_051349.1 (75845836..75996094, complement)	Hd, Hdh	

Figura 36

Identificació del gen
Font Pròpia

S'obrirà una altra pàgina on hi ha una taula amb dades de gens homòlegs del gen HTT en diferents organismes. Es localitza el corresponent a *Homo sapiens*, la seqüència que presenten els humans.

Un cop s'ha clicat sobre el link HTT corresponent als humans, s'obre la primera pàgina del gen amb molta més informació.

Per arribar al final, es clica ara sobre l'enllaç "**GenBank**" d'aquesta nova pàgina, que es troba a la secció "*Genomic regions, transcripts, and products*".

Genomic regions, transcripts, and products

Go to [reference sequence details](#)

Genomic Sequence: NC_000004.12 Chromosome 4 Reference GRCh38.p14 Primary Assembly

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

Genes, MANE Project (release v1.0)

HTT

NP_001388492.1

NP_001375421.1

2027.2

NP_001042098.2

NP_001010101.1

Genes, NCBI Homo sapiens Annotation Release 110, 20..

HTT

NP_002111.8

NP_001388492.1

HTT-AS

NR_045414.1

RNU7-33P

NP_047415055.1

NP_04721.1

NP_001330620.2

NP_00131.1

NP_001042098.2

NP_00101.1

NP_01513467.4

NP_01511.1

Figura 37

Entrada a GenBank
Font Pròpia

Finalment s'obrirà una altra pàgina on entre altres dades hi haurà la seqüència del gen. La primera part de la pàgina està plena de dades sobre els grups de recerca que han treballat en aquest cromosoma.

Es troba l'organització de l'ARN missatger (RNAm), la seva seqüència de bases ve indicada amb les coordenades dels exons. Per tant, en la maduració de la proteïna HTT, es perdran les seqüències que no estan incloses en aquestes (els introns).

En aquest gen es pot trobar més d'un ARN missatger, n'hi ha dos. Això respon a diferències en el procés de maduració de l'ARN que tenen els diferents teixits d'un organisme. Aquestes proteïnes, que procedeixen d'un mateix gen, tenen normalment la mateixa funció i s'anomenen isoformes. El gen HTT presenta dues isoformes, i actualment no es coneixen les diferències d'actuació que presenten entre elles.

També hi trobem la CDS (*CoDifying Sequence*), que mostra el gen HTT correcte en forma d'aminoàcids un cop tretts els introns (tant la isoforma 1 com la isoforma 2). És el que codifica a la proteïna madurada.

La seqüència d'ADN comença sota la paraula "*ORIGIN*" en l'apartat de "*gene*" i representa l'ordre en què es troben les 4 bases de la seqüència codificadora completa del gen. En la seqüència d'aquest gen, que es troba al primer exó, hi ha **19 repeticions** del triplet CAG (que és el que determina si es presentarà o no la malaltia). Per tant, el GenBank mostra el gen quan no presenta la malaltia.

```

gene /db_xref="GeneID:109461479"
      complement(3804..3862)
      /gene="RNU7-33P"
      /gene_synonym="U7.33"
      /note="RNA, U7 small nuclear 33 pseudogene; Derived by
      automated computational analysis using gene prediction
      method: Curated Genomic."
      /pseudo
      /db_xref="GeneID:100147820"
      /db_xref="HGNC:34129"
misc_feature 131402..131696
      /standard_name="Sharpr-MPRA regulatory region 10599"
      /note="Region: biological region; Derived by automated
      computational analysis using gene prediction method:
      RefSeqFE."
regulatory /db_xref="GeneID:112939930"
      131402..131696
      /regulatory_class="enhancer"
      /experiment="EXISTENCE:reporter gene assay evidence
      [ECO:0000049][PMID:27701403]"
      /note="tiled region #10599; HepG2 Activating DNase matched
      - State 5:Enh"
      /function="activates a minimal TATA promoter and a strong
      SV40 promoter by Sharpr-MPRA in HepG2 cells"
      /db_xref="GeneID:112939930"
ORIGIN
1 gctgccggga cgggtccaag atggacggcc gctcaggttc tgcttttacc tgcggcccag
61 agccccattc attgccccgg tgctgagcgg cgccgcgagt cggcccagg cctccgggga
121 ctgccgtgcc gggcgggaga ccgcatggc gaccctggaa aagctgatga aggccttcca
181 gtccctcaag tcctccagc agcagcagca gcagcagcag cagcagcagc agcagcagca
241 gcagcagcag cagcaacagc cgccaccgcc gccgccgcc cgccgcctc ctcagcttcc
301 tcagccgccg ccgcaggcac agccgctgct gcctcagccg cagccgcccc cgccgccgcc
361 cccgccgcca cccggcccgg ctgtggctga ggagccgctg caccgaccgt gagtttgggc

```

Figura 38

Seqüència del gen HTT
Font Pròpia

A l'Annex 1 s'hi troba la seqüència del gen correcte (19 repeticions) en forma de nucleòtids i en forma d'aminoàcids (CDS). I també s'hi troba la seqüència del gen mutat (37 repeticions) en forma de nucleòtids i en forma d'aminoàcids (CDS).

3.2. Disseny de la PCR

És necessari dur a terme un procés de PCR per corregir la corea de Huntington, ja que caldrà fer una gran quantitat de còpies del gen que permetin posteriorment, una vegada el complex ARNg-CAS9 hagi dut a terme la seva funció, que la versió correcta del gen substitueixi la seqüència que conté la mutació.

Aquesta còpia serà d'ADN bicatenari (de doble cadena) i contindrà la seqüència del gen correcte, en aquest cas 19 repeticions del triplet CAG (es produeix la malaltia a partir de 36 repeticions).

Per tant, la llargada del fragment resultant de la PCR ha de constar dels 57 nucleòtids que formen les 19 repeticions, 14 nucleòtids més que comprenen fins al PAM (71 nucleòtids en total) i 100 nucleòtids en direcció 3' i 100 nucleòtids més en direcció 5'.

Són necessaris aquests 200 nucleòtids addicionals perquè així, gràcies al procés HDR (*Homology Directed Repair*), s'ajuntaran les bases complementàries de l'ADN i el fragment resultant de la PCR, i es produirà automàticament un encreuament entre les dues molècules corregint així la mutació que presentava l'ADN.

Per tant, aquesta còpia que haurem obtinguin estarà constituïda per 271 nucleòtids:

Seqüència dels 71 nucleòtids:

```
5' ccttcagc agcagcagca gcagcagcag cagcagcagc agcagcagca gcagcagcag cagcaacagc cg  
3' ggaaggtcg tcgtcgtcgt cgtcgtcgtc gtcgtcgtcg tcgtcgtcgt cgtcgtcgtc gtcgttctcg gc
```

Repeticions CAG

PAM

Seqüència del fragment resultant de la PCR complet (271 nucleòtids):

5' gccgcgagt cggcccgagg cctccgggga ctgccgtgcc gggcgggaga 3'
3' cggcgctca gccgggctcc ggaggcccct gacggcacgg cccgccctct 5'

5' ccgccatggc gaccctggaa aagctgatga aggccttcga gtcctcaag 3'
3' ggcggtaccg ctgggacctt ttcgactact tccggaagct cagggagttc 5'

5' tccttcagc agcagcagca gcagcagcag cagcagcagc agcagcagca 3'
3' aggaaggctc tcgtcgtcgt cgctcgtcgc gtcgtcgtcg tcgtcgtcgt 5'

5' gcagcagcag cagcaacagc cgccaccgcc gccgccgccg ccgccgcctc 3'
3' cgctcgtcgc gtcgtttgtcg gcggtggcgg cggcggcggc ggcggcgagg 5'

5' ctcagcttcc tcagccgccg ccgcaggcac agccgctgct gcctcagccg 3'
3' gagtcgaagg agtcggcggc ggcgtccgtg tcggcgacga cggagtcggc 5'

5' cagccgcccc cgccgcggcc cc 3'
3' gtcggcgggg gcggcggcgg gg 5'

Repeticions CAG

PAM

200 Nucleòtids Addicionals

3.3. Com fer el disseny CRISPR

Per fer el disseny per arribar a corregir la mutació que causa la corea de Huntington necessitarem fer dos talls, un a l'inici de les repeticions i un altre al final d'aquestes. Així podrem extreure totes les repeticions de triplet CAG sense que importi quantes n'hi ha i podrem substituir-ho pel fragment corregit obtingut del procés de la PCR (que només tindrà 19 repeticions).

Per tant, com que volem fer dos talls en l'ADN, necessitarem dos complexos ARNg-CAS9 (2 CRISPR). En aquest cas els 2 ARN guies hauran de ser diferents, ja que un s'haurà de situar allà on comencen les repeticions i l'altre a on acaben aquestes.

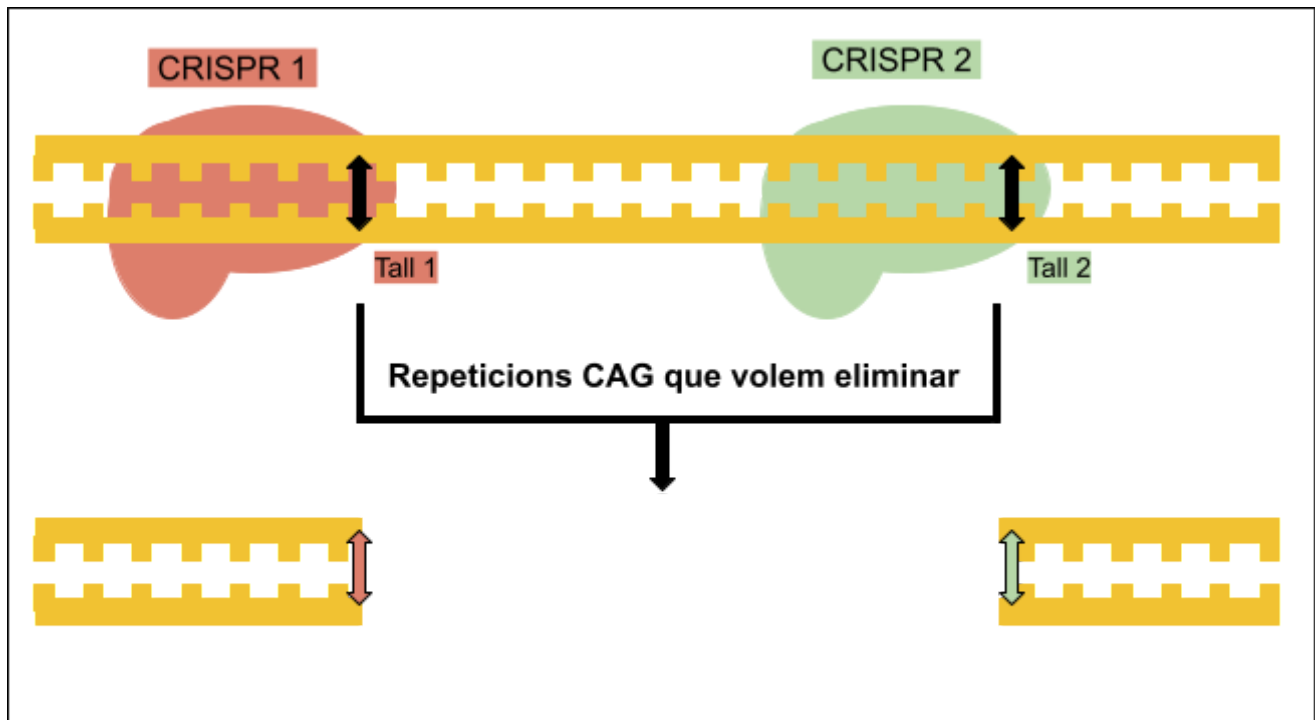


Figura 39

Disseny amb 2 CRISPRs
Font Pròpia

Cada complex CRISPR consta d'una proteïna CAS9 i un ARN guia. La proteïna CAS9 serà la mateixa per als dos casos però variarà la seqüència de l'ARN guia.

L'ARN guia en els dos casos ha de constar d'uns 100 nucleòtids aproximadament. En aquests es pot diferenciar la part invariable (que és la que s'uneix a la proteïna CAS9) i una part variable (que reconeix en quin punt de l'ADN s'ha de realitzar el tall).

La part variable de l'ARN guia consta de 20 nucleòtids (aprox), i els últims 3 són el PAM. La part variable serà el 20% de l'ARN guia i en conseqüència la part invariable serà l'altre 80% i constarà de 80 nucleòtids aproximadament.

La part variable de l'ARN guia està situada a l'extrem 5' i la part invariable a l'extrem 3'. Els tres primers nucleòtids començant per l'extrem 5' seran el PAM (que és 5'-NGG-3'). La proteïna CAS9 tallarà entre el tercer i el quart nucleòtid del PAM en direcció 5' (*upstream*). Podem situar el PAM tan en la cadena motlle com en la codificadora, però hem de respectar el sentit de les cadenes (sempre han de ser complementàries).

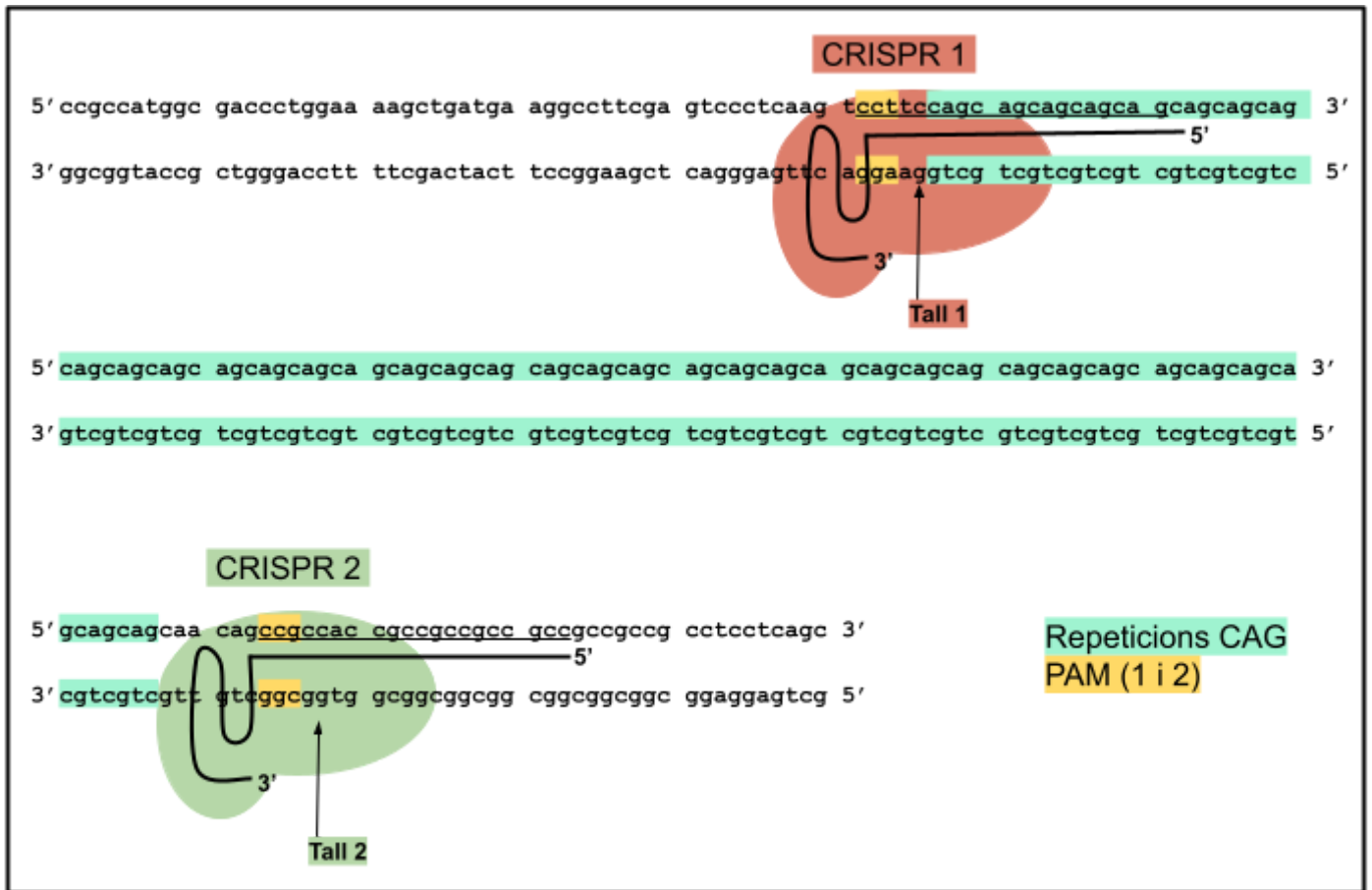


Figura 40

Disseny amb 2 CRISPRs per extreure les repeticions CAG del gen HTT
Font Pròpia

PAM 1 → 5' - AGG - 3'

PAM 2 → 5' - CGG - 3'

Part variable de l'ARN guia 1 → 5' - cugcugcugcugcugga**agg** - 3'

Part variable de l'ARN guia 2 → 5' - ggcggcggcggcggugg**cgg** - 3'

Un cop les proteïnes CAS9 hagin tallat l'ADN, la cèl·lula automàticament ho intentarà reparar. Com que hi hem introduït el fragment resultant de la PCR, que conté la versió correcta de les repeticions CAG, gràcies al mecanisme de reparació cel·lular de l'ADN (HDR) aquest quedarà corregit (és a dir, es produirà un entrecruament amb l'ADN i el fragment resultant de la PCR, tal com està explicat pas a pas al punt 2.4.4.).³⁴

³⁴ "Homology-Directed Repair | CRISPR/Cas9."

<https://sites.tufts.edu/crispr/genome-editing/homology-directed-repair/>. S'hi ha accedit el dia 31 d'oct. 2022.

3.3.1. Com introduir el material necessari perquè es realitzi el procés a la cèl·lula

Per introduir a la cèl·lula tota la maquinària necessària perquè es produeixin els processos requerits per tal de corregir la mutació en el gen, s'hauria d'utilitzar un procés d'**electroporació**.

A dins de la cèl·lula hem d'introduir els dos complexos de ARN_{guia}/CAS9 i els fragments d'ADN resultants de la PCR.

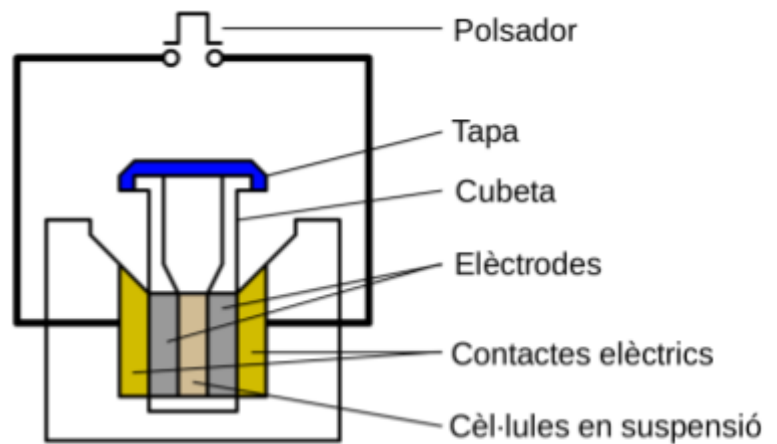


Figura 41

https://ca.wikipedia.org/wiki/Electroporaci%C3%B3#/media/Fitxer:Diagrama_de_l'electroporaci%C3%B3.s
Electroporador
vg

L'electroporació és una tècnica que permet introduir diverses molècules, com ara àcids nucleics o proteïnes, en cèl·lules procariotes o eucariotes mitjançant pulsos elèctrics breus d'alt voltatge.³⁵

L'electroporació consisteix a provocar un augment significatiu de la conductivitat elèctrica i la permeabilitat de la membrana plasmàtica cel·lular mitjançant un camp elèctric. Aquestes vibracions que es produeixen a la membrana fan que es creïn porus a la membrana plasmàtica (tant de cèl·lules procariotes com cèl·lules eucariotes) que permeten que s'hi introdueixin naturalment una gran diversitat de molècules, ja poden ser proteïnes, àcids nucleics etc.

Un cop es deixen d'aplicar els pulsos elèctrics breus d'alt voltatge els porus que s'han format a la membrana plasmàtica de la cèl·lula es tancaran automàticament sense produir cap dany a la membrana.³⁶

³⁵ "electroporació - Diccionari d'immunologia - Termcat." <https://www.termcat.cat/es/diccionaris-en-linia/189/fitxa/MjlxMDQ2NQ%3D%3D>. S'hi ha accedit el dia 31 d'oct.. 2022.

³⁶ "Electroporación - química.es." <https://www.quimica.es/enciclopedia/Electroporaci%C3%B3n.html>. S'hi ha accedit el dia 31 d'oct.. 2022.

4. Conclusions

Fent aquest treball he arribat a diverses conclusions.

En primer lloc, he pogut arribar a la conclusió, a través de recerca bibliogràfica, que la Corea de Huntington és una malaltia molt greu que genera patiment de la qual no existeix cap cura actualment ni cap forma d'evitar que empitjori.

Actualment, s'està evolucionant molt en àmbits de genètica, biotecnologia i bioinformàtica i s'està experimentant amb l'objectiu de trobar mecanismes utilitzant eines d'edició genètica. Aquests nous coneixements si els apliquem a la malaltia de Huntington ens podria permetre evitar-la o curar-la.

Una d'aquestes possibles tècniques per modificar les cèl·lules genèticament és el sistema de CRISPR, que permet editar el gen HTT causant de la malaltia de Huntington, i és el que he emprat en aquest treball.

Actualment existeixen els mecanismes i les eines suficients per realitzar una correcció en l'ADN utilitzant la tècnica de CRISPR en cèl·lules de cultiu i aquestes han mostrat majoritàriament resultats prometedors, però aquesta tècnica encara no es considera suficientment desenvolupada per ser aplicada en pacients.³⁷

Amb aquest treball també he pogut arribar a la conclusió que actualment hi ha suficients eines a l'abast de tothom de qui disposi d'internet com per fer un disseny complet d'un sistema de CRISPR a nivell teòric, i jo amb el disseny que he fet per curar la Corea de Huntington utilitzant dades provinents del portal NCBI he demostrat que és possible fer-ho.

És important però, esmentar que el fet que hi hagi suficients eines a l'abast de tothom podria comportar certs riscos en un futur, ja que qualsevol persona amb coneixements de biologia i genètica molecular seria capaç de manipular el genoma humà amb altres objectius més enllà de curar malalties genètiques.

I per finalitzar he arribat a la conclusió que no puc aplicar el meu treball teòric a la pràctica, en un medi de cultiu. Això és degut a que es tracta d'un procés complex que s'ha de dur a terme en un laboratori, s'han de sintetitzar una gran quantitat de molècules i es requeriria molta maquinària complexa.

Per altra banda també hi ha la impossibilitat d'aplicar el meu treball a la pràctica perquè les eines d'edició genètica no es consideren suficientment segures com per

³⁷ "Una novedad en la edición de genes CRISPR podría tener" 21 de jul.. 2021, <https://es.hdbuzz.net/308>. S'hi ha accedit el dia 2 de nov.. 2022.

ser aplicades a pacients humans perquè està en fase d'investigació (tot i que tot apunta a que es desenvoluparà de manera significativa en un futur proper), però quan evolucionin suficientment aquestes tecnologies, el CRISPR dissenyat en aquest treball de recerca es podria utilitzar per guarir la malaltia.

5. Referències bibliogràfiques

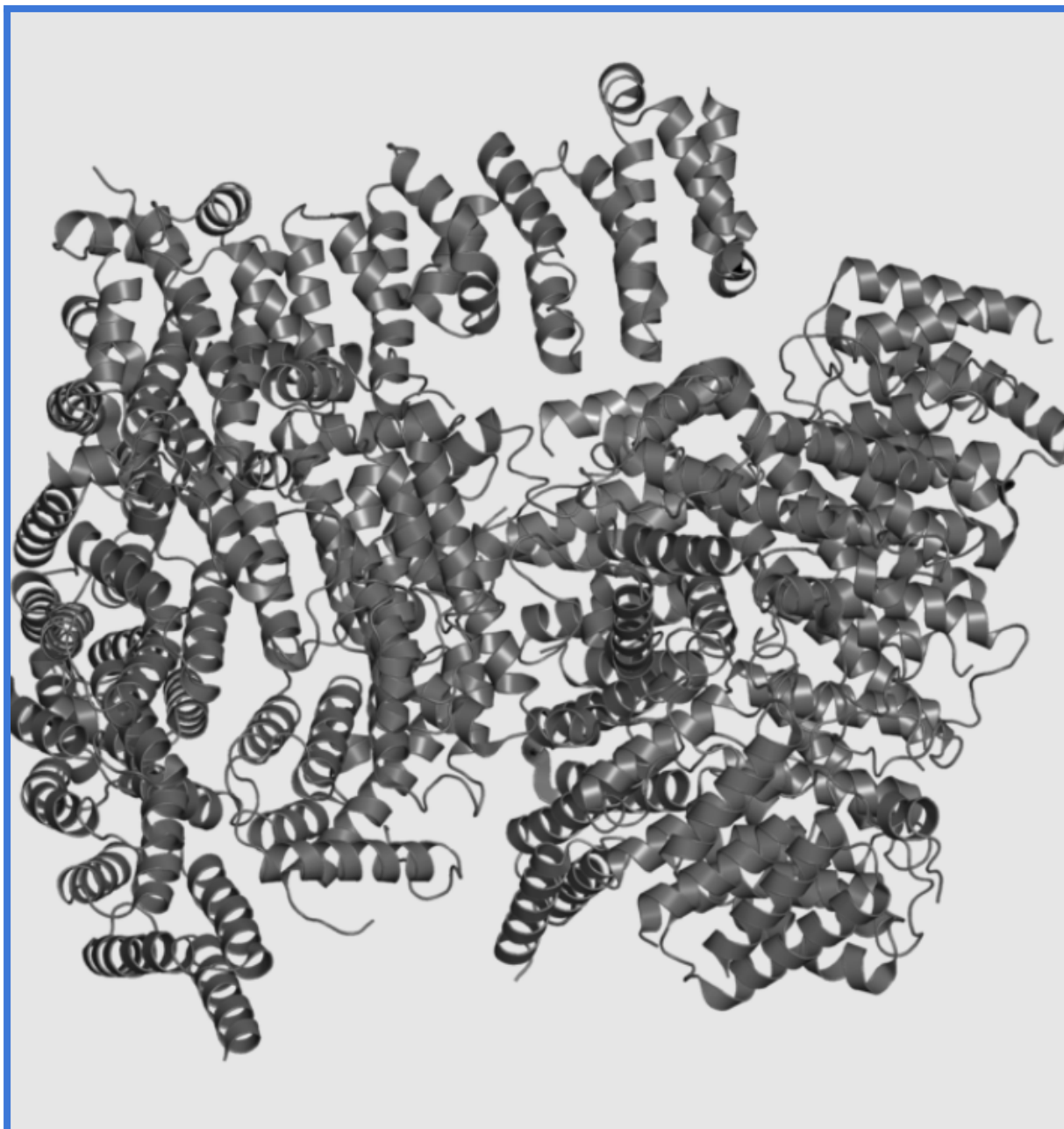
1. "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.
2. "Huntington's Disease - NORD (National Organization for Rare" <https://rarediseases.org/rare-diseases/huntingtons-disease/>.
3. "OMIM Entry - # 143100 - HUNTINGTON DISEASE; HD." <https://www.omim.org/entry/143100>.
4. "Huntingtin - Homo sapiens (Human) - HTT gene & protein - UniProt." <https://www.uniprot.org/uniprot/P42858>.
5. "The cryo-electron microscopy structure of huntingtin - PubMed." 1 de març. 2018, <https://pubmed.ncbi.nlm.nih.gov/29466333/>.
6. "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.
7. "On Chorea - Wikisource, the free online library." 30 de des.. 2020, https://en.wikisource.org/wiki/On_Chorea.
8. "eCIE-Maps - CIE-10-ES Diagnósticos." <https://eciemaps.mscbs.gob.es/>. S'hi ha accedit el dia 2 de nov.. 2022.
9. "Huntington's Disease - NORD (National Organization for Rare" <https://rarediseases.org/rare-diseases/huntingtons-disease/>.
10. "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.
11. "Huntington disease - Genetics - MedlinePlus." 1 de jul.. 2020, <https://medlineplus.gov/genetics/condition/huntington-disease/>.
12. "Huntington's disease - NHS." <https://www.nhs.uk/conditions/huntingtons-disease/>.
13. "ACMAH | Associació Catalana De Malalts de Huntington." <https://www.acmah.org/>. S'hi ha accedit el dia 2 d'oct.. 2022.
14. "¿Qué es un gen?: MedlinePlus Genetics." 28 d'abr.. 2021, <https://medlineplus.gov/spanish/genetica/entender/basica/gen/>. S'hi ha accedit el dia 28 de set.. 2022.
15. "DNA Replication - National Human Genome Research Institute." <https://www.genome.gov/genetics-glossary/DNA-Replication>. S'hi ha accedit el dia 28 de set.. 2022.

16. "Modes of DNA Replication." https://www.mun.ca/biology/scarr/iGen3_03-01.html. S'hi ha accedit el dia 28 de set.. 2022.
17. "DNA Transcription | Learn Science at Scitable - Nature." <https://www.nature.com/scitable/topicpage/dna-transcription-426/>. S'hi ha accedit el dia 28 de set.. 2022.
18. "From DNA to RNA - Molecular Biology of the Cell - NCBI Bookshelf." <https://www.ncbi.nlm.nih.gov/books/NBK26887/>. S'hi ha accedit el dia 28 de set.. 2022.
19. "CHEBI:74035 - small nuclear RNA - EMBL-EBI." 30 de gen.. 2015, <https://www.ebi.ac.uk/chebi/searchId.do?chebId=CHEBI:74035>. S'hi ha accedit el dia 28 d'ag.. 2022.
20. "Genetic Code - National Human Genome Research Institute." <https://www.genome.gov/genetics-glossary/Genetic-Code>. S'hi ha accedit el dia 28 de set.. 2022.
21. "Biosíntesis de proteïnes: Transcripció i traducció." 3 de nov.. 2017, <https://cn-tec.blogspot.com/2017/11/biosintesis-de-proteines-transcripcio-i.html>. S'hi ha accedit el dia 28 de set.. 2022.
22. "LES MUTACIONES SÓN CANVIS O ALTERACIONES DEL MATERIAL" <https://slideplayer.es/slide/14021588/>. S'hi ha accedit el dia 28 de set.. 2022.
23. "70. Els agents mutàgens - SlideShare." 2 d'abr.. 2013, <https://www.slideshare.net/daniribo/70-els-agents-mutgens>. S'hi ha accedit el dia 28 de set.. 2022.
24. "History of Polymerase Chain Reaction (PCR) - News-Medical.Net." [https://www.news-medical.net/life-sciences/History-of-Polymerase-Chain-Reaction-\(PCR\).aspx](https://www.news-medical.net/life-sciences/History-of-Polymerase-Chain-Reaction-(PCR).aspx). S'hi ha accedit el dia 1 de jul.. 2022.
25. "Reacción en cadena de la polimerasa (PCR) - Khan Academy." <https://es.khanacademy.org/science/ap-biology/gene-expression-and-regulation/biotechnology/a/polymerase-chain-reaction-pcr>. S'hi ha accedit el dia 28 de set.. 2022.
26. "¿Qué son la edición del genoma y CRISPR-Cas9? - MedlinePlus." 13 de juny. 2022, <https://medlineplus.gov/spanish/genetica/entender/investigaciongenomica/ediciondelgenoma/>. S'hi ha accedit el dia 18 de set.. 2022.
27. "¿Qué es la tecnología CRISPR? - Bayer." 12 d'abr.. 2022, <https://www.bayer.com/es/es/blog/espana-que-es-la-tecnologia-crispr>. S'hi ha accedit el dia 18 de set.. 2022.

28. "¿Qué es la tecnología CRISPR/Cas9 y cómo nos cambiará la vida?." <https://www.dciencia.es/que-es-la-tecnologia-crispr-cas9/>. S'hi ha accedit el dia 18 de set.. 2022.
29. "¿Qué son la edición del genoma y CRISPR-Cas9? - MedlinePlus." 13 de juny. 2022, <https://medlineplus.gov/spanish/genetica/entender/investigaciongenomica/ediciondelgenoma/>. S'hi ha accedit el dia 18 de set.. 2022.
30. "¿Qué es la tecnología CRISPR/Cas9 y cómo nos cambiará la vida?." <https://www.dciencia.es/que-es-la-tecnologia-crispr-cas9/>. S'hi ha accedit el dia 18 de set.. 2022.
31. "Explaining CRISPR-CAS9 System- Step by Step - Genetic Education." 3 d'ag.. 2020, <https://geneticeducation.co.in/explaining-crispr-cas9-system-step-by-step/>. S'hi ha accedit el dia 18 de set.. 2022.
32. "Stimulation of CRISPR-mediated homology-directed repair by an" 30 de jul.. 2019, <https://www.nature.com/articles/s41467-019-11105-z>. S'hi ha accedit el dia 18 de set.. 2022.
33. "Home - PMC - NCBI." <https://www.ncbi.nlm.nih.gov/pmc/>. S'hi ha accedit el dia 30 d'oct.. 2022.
34. "Homology-Directed Repair | CRISPR/Cas9." <https://sites.tufts.edu/crispr/genome-editing/homology-directed-repair/>. S'hi ha accedit el dia 31 d'oct.. 2022.
35. "electroporació - Diccionari d'immunologia - Termcat." <https://www.termcat.cat/es/diccionaris-en-linia/189/fitxa/MjkxMDQ2NQ%3D%3D>. S'hi ha accedit el dia 31 d'oct.. 2022.
36. "Electroporación - química.es." <https://www.quimica.es/enciclopedia/Electroporaci%C3%B3n.html>. S'hi ha accedit el dia 31 d'oct.. 2022.
37. "Una novedad en la edición de genes CRISPR podría tener" 21 de jul.. 2021, <https://es.hdbuzz.net/308>. S'hi ha accedit el dia 2 de nov.. 2022.

Annex 1

Seqüència del gen de la Huntingtina



Uracil
4 de novembre del 2022

Índex de Continguts

1. Gen HTT correcte en forma d'aminoàcids	52
2. Gen HTT correcte en forma de nucleòtids	53
3. Gen HTT mutat en forma d'aminoàcids	56
4. Gen HTT mutat en forma de nucleòtids	58

En aquest annex hi podem trobar la seqüència del gen correcte (19 repeticions) en forma de nucleòtids i en forma d'aminoàcids (CDS). I també s'hi troba la seqüència del gen mutat (37 repeticions) en forma de nucleòtids i en forma d'aminoàcids (CDS).

Repeticions dels triplets CAG

1. Gen HTT correcte en forma d'aminoàcids (isoforma 1) (19 repeticions CAG)

"MATLEKLMKAFESLKSFQQQQQQQQQQQQQQQQQQQPPPPPPPPPPQLPQPPPQAQPLLP
QPQPPPPPPPPPGPAVAEEPLHRPKKELSATKKDRVNHCLTICENIVAQSVRNSPEFQKLL
GIAMELFLCSDDAESDVRMVADECLNKVIKALMDSNLPRLQLELYKEIKKNGAPRSLRAAL
WRFaelahlvrpQKCRPYLVNLLPCLTRTSKRPEESVQETLAAAVPKIMASFGNFANDNEIK
VLLKAFIANLKSSSPTIRRTAAGSAVSICQHSRRTQYFYSWLLNVLLGLLVPVEDEHSTLLI
LGVLLTLRYLVPELLQQQVKDTSLKGSFGVTRKEMEVSPSAEQLVQVYELTLHHTQHODHNV
TGALELLQQLFRTPPELLQTLTAVGGIGQLTAAKEESGGRSRSGSIVELIAGGGSSCSPVL
SRKQKGVLLGEEEALEDDSESRSDVSSALTASVKDEISGELAASSGVSTPGSAGHD IITE
QPRSQHTLQADSVDLASCDLTSSATDGDEEDILSHSSSQVSAVPSDPAMDLDNGTQASSPIS
DSSQTTTEGPD SAVTPSDSSEIVLDGTDNQYLGLQIGQPQDEDEEATGILPDEASEAFRNSS
MALQQAHLKNSHCRQPSDSSVDKFLRDEATEPGDQENKPCRIKGDIGQSTDDDSAPLVH
CVRLLSASFLLTGGKNVLPDRDVRVSVKALALSCVGAVALHPESFFSKLYKVPLDTTEYP
EEQYVSDILNYIDHGD PQVRGATAILCGTLICSIILSRSRFHVGDWMTIRTTLTGNTFSLADC
IPLLRKTLKDESSVTCKLACTAVRNCVMSLCSSSYSELGLQLIIDVLTNRNSSYWLVRTELL
ETLAEIDFRLVSFLEAKAENLHRGAHHTYGLLKLQERVLNNVVIHLLGDEDPRVRHVAASL
IRLVPKLFYKCDQGDQADPVVAVARDQSSVYLKLLMHETQPPSHFSVSTITRIYRGYNLLPSI
TDVTMENNLSRVIAAVSHELITSTTRALTFCCEALCLLSTAFPVCIWISLGWHCGVPLSAS
DESRKCTVGMATMILTLLSSAWFPLDLSAHQDALILAGNLLAASAPKSLRSSWASEEEANP
AATKQEEVWPALGDRALVPMVEQLFSHLLKVINICAHVLDVAPGPAIKAALPSLTNPPSLS
PIRRKGEKEPEGEQASVPLSPKKGSEASAASRQSDTSGPVTTSKSSSLGSFYHLPSYLKLD
VLKATHANYKVTLDLQNSTEKFGGFLRSALDVLSQLLELATLQDIGKCV EILGYLKCFSR
EPMMATVCVQQLKTLFGTNLASQFDGLSSNPSKSGRAQRLGSSSVRPGLYHYCFMAPYTH
FTQALADASLRNMVQAEQENDTSGWFDVLQKVSTQLKTNLTSVTKNRADKNAIHNHIRLFEP
LVIKALKQYTTTTTCVQLQKQVLDLLAQLVQLRVNYCLLSDQVFIGFVLKQFEYIEVGQFRE
SEAIIPNIFFFLVLLSYERYHSKQIIGIPKIIQLCDGIMASGRKAVTHAIPALQPIVHDLFV
LRGTNKADAGKELETQKEVVVSMMLRLIQYHQVLEMFILVLQQCHKENEDKWKRLSRQIADI
ILPMLAKQQMHIDSHEALGVLNLTLEILAPSSLRPVDMLLRSMFVTPNTMASVSTVQLWISG
ILAILRVLISQSTEDIVLSRIQELSFSPYLISCTVINRLRDGDSTSTLEEHSEGKQIKNLPE
ETF SRFLQLVGILLEDIVTKQLKVEMSEQHTFYCQELGTLMLCLIHIFKSGMFRRITAAA
TRLFRSDGCGGSFYTLDSLNLRRARSMITTHPALVLLWCQILLLVNHTDYRWAAEVQQTPKRH
SLSSTKLLSPQMSGEEEDSDLAAKLGMCNREIVRRGALILFCDYVCQNLHDSEHLTWLIVNH
IQDLISLSHEPPVQDFISAVHRNSAASGLFIQAIQSRCENLSTPTMLKKTLCLEGIHLSQS
GAVLTLYVDRLLCTPFRVLARMVDILACRRVEMLLAANLQSSMAQLPMEELNRIQEYQLQSSG
LAQRHQRLYSLLDRFRLSTMQDSLSPSPVSSHPLDGDGHVSLETVSPDKDWYVHLVKSQCW

TRSDSALLEGAELVNRI PAEDMNAFMMNSEFNLSLLAPCLSLGMSEISGGQKSALFEAAREV
 TLARVSGTVQQLPAVHHVFQPELPAEPAAYWSKLNDFGDAALYQSLPTLARALAQYLVVVS
 KLPSHLHLPEPEKEDIVKFVVATLEALSWHLIHEQIPLSLDLQAGLDCCCLALQLPGLWSV
 SSTEFTVTHACSLIYCVHFVILEAVAVQPGEQLLSPERRTNTPKAISEEEEEVDPNTQNPKYIT
 AACEMVAEMVESLQSVLALGHKRNSGVPAFLTPLLNRNIIISLARLPLVNSYTRVPPLVWKL
 WSPKPGDFGTAFPEIPVEFLQEKEVFKEFIYRINTLGWTSRTQFEETWATLLGVLVTQPLV
 MEQEEESPPEEDTERTQINVLAVQAITSVLVSAMTVPVAGNPAVSCLEQQPRNKPLKALDTRF
 GRKLSIIRGIVEQEIQAMVSKRENIATHHLYQAWDPVPSLSPATTGALISHEKLLLQINPER
 ELGSMYSYKLGQVSIHVSVLGNSITPLREEEWDEEEEEADAPAPSSPPTS PVNSRKHRAGVD
 IHSCSQFLELYSRWILPSSSARRTPAILISEVVRSLLVVSDLFTERNQFELMYVTLTTELRR
 VHPSEDEILAQYLV PATCKAAAVLGMDKAVAE PVSRLLESTLRSSHLP SRV GALHGVLYVLE
 CDLLDDTAKQLIPVISDYLLSNLKGIAHCVNIHSQQHVLVMCATAFYLIENYPLDVGPEFSA
 SIIQMCVMLSGSEESTPSIIYHCALRGLERLLLSEQLSRLDAESLVKLSVDRVNVHSPHRA
 MAALGLMLTTCMYTGKEKVS PGR TSDPNPAAPDSESVIVAMERVS VLFDRIRKGFPCEARVVA
 RILPQFLDDFFPPQDIMNKVIGEFLSNQQPYPQFMATVVYKVFQTLHSTGQSSMVRDWVMLS
 LSNFTQRAPVAMATWSLSCFFVSASTSPWVAAILPHVISRMGKLEQVDVNLFCLVATDFYRH
 QIEEELDRRAFQSVLEVVAAPGSPYHRLLTCLRNHVHKVTTTC"

2. Gen HTT correcte en forma de nucleòtids (isoforma 1) (19 repeticions CAG)

Aquí no es mostra el gen complet perquè és molt llarg i ocuparia més de 50 pàgines, però es presenta un fragment perquè sigui possible fer-se a la idea de quina és la llargada del gen i com un canvi en una mínima quantitat de nucleòtids pot acabar desencadenant una malaltia tan severa com pot ser la corea de Huntington.

```

1 gctgcccggga cgggtccaag atggacggcc gctcagggtc tgcttttacc tgcggcccag
61 agccccattc attgccccgg tgctgagcgg cgcccgaggt cggccccgag cctccggggg
121 ctgcoctgcc gggcgggaga ccgccatggc gaccctggaa aagctgatga aggccctcga
181 gtccctcaag tccttc cagc agcagcagca gcagcagcag cagcagcagc agcagcagca
241 gcagcagcag cagcaacagc cgccaccgcc gccgccgccg ccgccgcctc ctcagcttcc
301 tcagccgcgg cgcgaggcac agcccgctgct gcctcagccg cagccgcccc cgccgcggcc
361 cccgcgcgca cccggcccgg ctgtggctga ggagccgctg caccgaccgt gagtttgggc
421 ccgctgcagc tccctgtccc ggcgggtccc aggctacggc ggggatggcg gtaaccctgc
481 agcctgcggg ccggcgacac gaacccccgg ccccgagag acagagtgac ccagcaaccc
541 agagcccatg agggacaccc gccccctcct ggggcgaggc cttccccccac ttcagccccg
601 ctccctcaact tgggtcttcc cttgtcctct cgcgagggga ggcagagcct tgttggggcc
661 tgtcctgaat tcaccgaggg gagtcacggc ctcagccctc tcgcccttcg caggatgcga
721 agagttgggg cgagaacttg tttcttttta tttgcgagaa accagggcgg gggttctttt
781 aactgcgttg tgaagagaac ttggaggagc cgagatttgc tcagtgccac ttcctcttc
841 tagtctgaga ggaagaggg ctgggggccc gggacacttc gagaggagc ggggtttgga
901 gctggagaga tgtgggggca gtggatgaca taatgctttt aggacgcctc ggcgggagtg
961 gcggggcagg gggggggcgg ggagtgaggg cgcgccaat gggagattc ttttctagt
1021 ggcacttaa acagcctgag atttgaggct cttcctacat tgcaggaca tttcatttag
1081 tcatgatca cggtggtagt aacacgattt taagcaccac ctaagagatc tgctcatcta
1141 agcctaagtt ggtctgcagg cgtttgaatg agttgtggtt gccaaagtaa gtggtgaact
1201 tacgtggtga ttaatgaaat tatcttaaat attaggaaga gttgattgaa gttttttgcc
1261 tatgtgtggt ggaataaaa ccaacacggt gctgatgggg aggttaattg ccgaggggatg

```

1321 aatgagggtgt acatthttacc agtattccag tcaggcttgc cagaatacgg ggggtccgca
1381 gactccgtgg gcatctcaga tgtgccagtg aaagggtttc tgthttgcttc attgctgaca
1441 gcttgthtact thtttggagg taggggtttc tgttgcttgt tcttggggag aaththttgaa
1501 acaggaaaag agagaccatt aaaacatcta gcggaacccc aggactthcc ctggaagtct
1561 gtgtgtogag tgtacagtag gagttaggaa gtactctggt gcagttcagg cthttctctt
1621 acctctcagt attctatttc cgatctggat gtgtcccaga tggcatttgg taagaatata
1681 tctgttaaga ctgattaatt thtagtaata thtcttgthc thttgthtctg ttatgathct
1741 tgtctcgtct tcaaagthta attagaaaat gattcggaga gcagtgthtag cthatttght
1801 ggaataaaaat thtaggaataa attattctaa aggatggaaa aactthtttgg athatttggag
1861 aathththaa acaatttggc thtathctthc agtaagthaat thctcatcca gaaaththtact
1921 gtagtgctth thtaggagggt aggtgtcata aaagthcaca cattgcatgt athcttgtgta
1981 aacactaaac agggctctctg atgggaagga agacctthct gctgggctgc thcagacact
2041 tgathattct aaaaathatgc cthctctthc thtagctgat thgacagaa ctgcatthtgc
2101 thtaththcaa aathatgggta tcaagaaath thctthtctg ccttgacaaa ggagathagat
2161 thttgthtcat thctthtaagg thathathga thacctthatt thaaaaath thaatcaggact
2221 ggcaaggthg cthacacct thaatccgagc actthtgggag gcctaggthg acgathcacc
2281 tgaggthcagg agthttgagac cagcctggct aacathgthga aacctgthct ctactaaaaa
2341 tacaaaaath agctggthcat ggtggcacgt gcctgthaat caagctacct gggaggctga
2401 ggcaggaaaa thgctthgaa ccgggaggca gathctgcag thagthtgaga thcagccact
2461 gcactccagc ctgggtgaca gagcgagact ctathctcaa aaaaaththtt ththaatgtht
2521 ththththtga thagthaatat athgathatga thacaaathct gthaththcaa aaggthcaata
2581 athaaaaath cthctththca cthctththct ctgagthacct aactthtthcc ccaaghaaaa
2641 gcactaththc agthctctcat gthctctgcc agathataacc thththcatatt gthagathaga
2701 ththaaaathc ththaaaaaca aaagthagtht aghaathaat athaththata athththtgg
2761 athgathctc acaththcacc ccaggctgga thgathgath acaathctcg thcactgthcag
2821 thctctgcctc ccaggththcaa athgctthctc thgctcagcc thctgagthag ctgggaththc
2881 aggcgccccac caccaththcc agthaththtt thgathththta thagathagth ggththcacc
2941 thgttggccag gctggtctthg aactcctgac cthgtgathct thccacctcg gcctccaaa
3001 thgctgggath thacaggthg agccaccath cctggctaga athaathact thaaagththc
3061 thtagcathgct ctgaaathcaa ctgcatthagg ththaththata gthththathg ththththaat
3121 aaaaathcath ththgthcath thctctgtht ththgctgthg aghaagthagg thththththaa
3181 thththgathaa caaacathgc thcaaaagth thgathththgg cagththctgth cagthgctthc
3241 agccaaaaaaa thctctthctc aaagthagath thgathgaaagc aaththgaaa gthathctgthc
3301 thgtthththtathg gctctthgctc ththggthgthg aactgthgthg thcagccathg cathgggctc
3361 agthththtathg thgtththgct ctgctcagca thacagthagc aggagththct ththgggctg
3421 gctgthcaggct cagcaathct agcathgctth ggaggthctc cacagthaat aggagthcaat
3481 thathaththgct thctggcagth thctthathct cththcagath cctathctgthg gththctctga
3541 cththththctat thcathcagthaa athathththact aacathgthact athgthcctgth cactgththata
3601 ggtgthcaggthc thcagcagthga gcagacaaaag ctctgccccctc thgaaagctth cathctaatg
3661 aaggathathg acagthaaagca agathagathaa gthaaaaathata cagthacgthta athacgthggag
3721 gaactthcaaaa gcaggthgagg ggathagggaa athgthcaggth thaatcagthg ththacctthatt
3781 ththathththta aaaaaththgt thagggctth ccagcaaaac ccaghaagcc thgctaghaaaa
3841 aththcaaaaag agctgthagca cthagthgthg acaththththt ththathththgt ththgtththgt
3901 thththththtgg acagththctthg ctctathcagc caggctgthg thgactathgth thgathctthggc
3961 thcactgthaac ctctgccccct thggththcaag thgaththctat gcctcagctc cctgththtagc
4021 thgggaththata gathathgact gccathgccccgth ggthathththth thththththccc ccgagacgga
4081 gctththgctct gthgccccagth ctggagthgca thgggccccgath ctcagctcac thgcaagctcc
4141 gctthccccgag thcagccath thctctgcccc cagthctccca agthagctgthg actacagthcg
4201 cctgthcaaaa cgtcagctha athththththgt athththththaa gathagccccgth thcaccgthgt
4261 thagccagthgath gathctthgath thctgacctc gthcathccccgac gathctthgthga thccccccacc
4321 thggccccctcc aagthgctthg gaththacagthc athgagccact thgccccgthc acgctgthggth
4381 aaththththtga ththththtagth agathgggtht thggccathgath gathcagctgth gthctcgaact
4441 cccgthccccca thgtgathctgc ctgccccgthc ctccccaaagth gctagthgath cagthcathgath

4501 ccaccataacc tggccagtgt tgatatttta aatacgggtg tcaggaaggg tccactgaga
4561 agacagcttt tttttttttt ttttttgggg ttgggggggca aggtcttctg ctttaaccca
4621 ggctggaatg cagtatcact atcgtagctc acttcagcct tgaactcctg ggctcaagtg
4681 atcctcccac ctcaacctca caatgtgttg ggactatagg tgtgagccat cacacctggc
4741 cagatgatgg cttttgagta aagacctcaa gcgagttaag agtctagtgt aagggtgtat
4801 gaagtagtgg tattccagat ggggggaaca ggtccaaaat cttcctgttt caggaatagc
4861 aaggatgtca ttttagttgg gtgaattgag tgagggggac atttgtagta agaagtaagg
4921 tccaagaggt caagggagtg ccatatcaga ccaatactac ttgccttgta gatggaataa
4981 agatattggc atttatgtga gtgagatggg atgtcactgg aggattagag cagaggagta
5041 gcatgatctg aatttcaatc ttaagtgaac tctggctgac aacagagtga aggggaacac
5101 cggcaaaagc agaaaccagt taggaagcca ctgcagtgct cagataagca tgggtgggttc
5161 tgtcagggta cggctgtcog gctgtgggca gtgtgaggaa tgactgactg gattttgaat
5221 goggaaccaa ctgcacttgt tgaactctgc taagtataac aatttagcag tagcttgcgt
5281 tadcaggttt gtattcagct gcaagtaaca gaaaatcctg ctgcaatagc ttaaactggg
5341 aacaagcaag agcttatcag aagacaaaaa taagtctggg gaaattcaac aataagttaa
5401 ggaaccagg ctctttcttt tttttttttt tgaaacggag tttcgtcttt gtcaccggg
5461 ctggagtgca atgatgtgat ctcagctcac taaaacctct acctcctggg tcaagtgat
5521 tcttctgctc cagcctccca agtaactggg attacaggcg tataccacca tgcccagcta
5581 atttttgtgt ttttagtaga gatgggggtt caccatgttg gccaggctgg tctcgaactt
5641 ctgacctcag gtgatccact cgcctcagcc tgccaaaagt ctgggattac aggtttgggc
5701 cactgcaccc ggtcagaacc caggctcttt cttatactta ccttgcaaac ccttgttctc
5761 attttttccc tttgtatttt tattgttgaa ttgtaatagt tctttatata ttctggatac
5821 tggattctta tcagatagat gatttgtaaa aactctccct tcctttggat tgtcttttta
5881 ctttcttgat agtgtctttt gaagtgtaaa agtttttaat tttgatgaag tctgagttat
5941 ctattttgtc tttggttgct gtgcttcaag tgtcatatct aagaaatcat tgtctaacc
6001 aaagtcaaaa aggtttactc ctatgttttc ttctaagaat tttagagttt tacatttaag
6061 tctgatccat tttgagttaa tttttatata tggttcaggt agaagtccaa ctttattctt
6121 ttocatgtgg ttattcagtt gtcccagcac tgtttgttga agagactatt ctttccccat
6181 ggaattatct tagtaccctt gttgaaaatt aatcgtcctt aattgtataa atttattctt
6241 agactgtcag ttctacctgt tggcttttat gtcgatcctg tgccagtacc atacagtctt
6301 gattactgaa gtttgtgtca cagtttaaat tcatgaaatg tgagttctcc aactttgttc
6361 cttttcaaga ttgatttggc catgctgggt cccttgcaat tccgtacgaa ttgtaggatc
6421 agcttgtcag tttcaacaaa gaagccaagt aggattctga gagggattgt gttgaaatctg
6481 tagatcaact tggggagtat tgcacatctta acaatattgt cttccacctc tgaacatggg
6541 caaactttgt gtaaatggtc agattgtgaa tatttcgggc tgtgtgggca cagtgtctct
6601 gtcacagcta cggcgtctct ccattgtagc atgaaagtag ccataagcaa tatgtatgag
6661 tgtctgtgtt ccaatagaat tttattaatg acaaggaagt ttgaaattca tataattttc
6721 acctgtcatg agatagtatt tgattatttt ggtcaacccat ttaaaaaatgt aaaaacattt
6781 cttagcttgt gaactagcca aaaatatgca ggttatagtt ttcccactcc taggttaaaa
6841 tatgatagga ccacatttgg aaagcatttc tttttttttt tttttttttt tttttgagac
6901 ggagtttcc tcttgttgcc caggctggag tgcagtggcg cgatctcggc tccactgcaac
6961 ctctgcctcc caggttcaag acattctcct gcacggcctc cctagtagct gggattacag
7021 gcatgogcca ccacaccag ctaattttgt attttttagta gagacggggg tcttccatgt
7081 tggtcaggct ggtcttgaac tctgacctc aggtgatcca cccgcctcag cctcccaag
7141 tgetgggatt acaggggtgt agccaccaca ccctgctgga aagcatttct tttttggctg
7201 tttttgtttt ttttttaaac tagttttgaa aattataaaa gttacacata tacattataa
7261 aatatcttc aagcagcaca gatgaaaaac aaagcccttc ttgcaagtct gtcactttt
7321 tctaacttcc taagaacaaa agtgtttctt gtgtcttctt cccagatttt aatatgcata
7381 tacaagcatt taaatgtgtc attttttgtt tgcttgactg agatcacatt acatatgtat
7441 ttttttactt aacaatgtgt catagatatt gttccatagc agtacctgta attcttatta
7501 attgctatgt aatattttag aatttctttt taaaagagga cttttggaga tgtaaaggca
7561 aaggctccac atttttgtgg ctgtagaatg tgctgggtgac atattctctc taccttgaga
7621 agtccccatc cccatcacct ccatttctct taaataagtc aaccacttga taaactacct

7681 ttgaatggat ccacactcaa aacathtagt cttattcaga caacaaggag gaaaaataaa
 7741 ataccttata aagcactggt taatattgta ttaaattgga tcaatttggg ggctagaatg
 7801 tatgttagag acatgatatg tccataggtc cttgctatca cagtggagtc tcaggggacag
 7861 tcgtttggta tcatttggga tctcataagc agactctctc tgcttgacct gacaaatcag
 7921 agtctgtggt ttaacagggt cagtgagtga cttacatgca cattggagtt tgggaagctc
 7981 cactgtaggt gcttagacct tacctttggt gttgctaata acaatgcaag catttgggag
 8041 gaagacctgt gttgctcata tgtgtccagg ttagctgag gtggccttgc ttatctgctg
 8101 tagggccggt gagcatttct gtagctgtga tgagttagct gaggtgagcc tgcggagagc
 8161 tcccagccat tggtagtggg actcgttag atgaactgga aggacccttt catctgagca
 8221 gccactatgg agaaaaacaa ccgaatgagg ggagagacaa tgtgcaatth tatttagggc
 8281 acaaaggaga gctgtgggta gaaggtgaca tttgagtgga aagggggcaa gccatgtgta
 8341 tagcgggaga agagagggtcc aggcagagtt aacagaaggc agaaatgctt tccatgtttg
 8401 agaaccagta aggaggccag tggctgaagt aaggtgaagg gcagaaataa ggatgaggct
 8461 gcgagagatg agagggttaga gacgagcgtc ttgtgcacca agataagctt gtgtggtcaa
 8521 aacaagtagt ttaatttatg tttttaaag atcattttgg ctgggcacaa tggttcatgc
 8581 ctgtaatacc agtagtttga gacgggtgtg tgggaggatt gcctgaggcc agacgaccag
 8641 catagccaac atagcagcac ctataaggtc tctacaaaaa actttaaaaa attagctggg
 8701 catagtgtg tgtgcctgta gtcccagcta ctcaggaggc tgaggaggct ggaggattgc
 8761 ttgagtccag gagtttgagg ctgcagttag ctatgattat gccactacac tacaacctgg
 8821 gcaagagagt gagacctgt ctctaaatat acacacacac acacacacac acacacacac
 8881 acacacacac acacacacac acacacatat atatgtatat atatgcattt agatgaaaa
 8941 atcactttga caataccaca tgctggtgag gatttagaaa aactagggtca cttattgctg
 9001 gtgggaatat aatatagtag gccactctg gaaaacagtt tggcagtttgc tcaaaaaact
 9061 gaacataccg ttagtataca gccagcagc aactacaatc ctgggcatta atcctagaga
 9121 aatgaaacct taatgttcac ataaaaacct atactcaagt atgcatagca gctttacca
 9181 taatatctaa gaactggaat cagctcagat gtccttcaac aggtgaatgg ttaaactact
 9241 cagtaataaa aaggaatgag ctactgatag catgcaacag tttagggtgaa gttatgctaa
 9301 tgaaaaaagc caatocccaaa aggttataca tactgtatga ttctatgttt ttttgcaatg
 9361 gcacagtttt agggatggag aatagattag tggttgcctg ggggttagaga tggggtagta
 9421 gagtaggtta gtggtggcag aggagagaaa agagagggag gtgaatgtgg ttataaaaagg
 9481 acaacacagg ggaataacttg taatggaaat gctttgtctt tttttttttt tttttttttt
 9541 tggcgacaga gtcttgctct gttgccagg ctggagtgca gtggcatgat cttttctcac
 9601 tgcaacctct gcctcctggg ttcaagtgat acttgtgtct cagtctccca tgttcagagt
 9661 gaaacaaacc agaggtaatg ttcatccaaa taatccaaca cacatgacat taaaacatca
 9721 agatcaggtc ggacgtggtg gctcatgcct gtaatcccag cacttttggg aggccaaggt
 9781 gggcagatca cttgagggtca ggagttcgag accagccggg ccaacatgat gaaaccccat
 9841 cttgactaaa aatacaaaaa ttagccgggc atggtggtgt gcacctgtag tcccagctac
 9901 ttgggaggct gaggcaagag aactgcttga acccgagggg cagagggttc agtgagctga
 9961 gagtgcgcca ttgcacttca gcctgtgtga cagagtaaga ctccatctcc aaaaaaaaaa
 10021 aaccaagatc aattaaata cagcattact gggccgggtg tgggtggtca cacctgtaat

3. Gen HTT mutat en forma d'aminoàcids (isoforma 1) (37 repeticions CAG)

"MATLEKLMKAFESLKSF **QQ** P P P P P P P P
 P P P P Q L P Q P P P Q A Q P L L P Q P Q P P P P P P P P G P A V A E E P L H R P K K E L S A T K K D R V N H C L T I C
 E N I V A Q S V R N S P E F Q K L L G I A M E L F L L C S D D A E S D V R M V A D E C L N K V I K A L M D S N L P R L Q L E
 L Y K E I K K N G A P R S L R A A L W R F A E L A H L V R P Q K R P Y L V N L L P C L T R T S K R P E E S V Q E T L A A A
 V P K I M A S F G N F A N D N E I K V L L K A F I A N L K S S S P T I R R T A A G S A V S I C Q H S R R T Q Y F Y S W L L N
 V L L G L L V P V E D E H S T L L I L G V L L T L R Y L V P L L Q Q V K D T S L K G S F G V T R K E M E V S P S A E Q L V

QVYELTLHHTQHODHNVVTGALELLQQLFRTPPPPELLQTLTAVGGIGQLTAAKEESGGRSRS
GSIVELIAGGGSSCSPVLSRKQKGVLLGEEEALEDDSESRSVDVSSSALTASVKDEISGELA
ASSGVSTPGSAGHDIIITEQPRSQHTLQADSVDLASCDLTSSATDGDEEDILSHSSSQVSAVP
SDPAMDLNDGTQASSPISDSSQTTTEGPD SAVTPSDSSEIVLDGTDNQYLGLQIGQPQDEDE
EATGILPDEASEAFRNSSMALQQAHLKLNMSHCRQPSDSSVDKFLRDEATEPGDQENKPCR
IKGDIGQSTDDDSAPLVHCVRLLSASFLLTGKKNVLPDRDVRVSVKALALSCVGAVALHP
ESFFSKLYKVPLDTTEYPEEQYVSDILNYIDHGD PQVRGATAILCGTLICSIILSRSRFHVGD
WMGTIRTTLTGNTFSLADCIPLLRKTLKDESSVTCKLACTAVRNCVMSLCSSSYSELGLQLII
DVLTLRNSSYWLVRTLELTLAEIDFRLVSFLEAKAENLHRGAHHTGLLKLQERVLNNVVI
HLLGDEDPRVRHVAAASLIRLVPKLFYKCDQGGADPVAVARDQSSVYLKLLMHETQPPSHF
SVSTITRIYRGYNLLPSITDVTMENNLSRVIAAVSHELITSTTRALTFGCCEALCLLSTAFP
VCIWSLGWHCVPPLSASDESRSKCTVGMATMILTLLSSAWFPLDLSAHQDALILAGNLLAA
SAPKSLRSSWASEEEANPAATKQEEVWPALGDRALVPMVEQLFSHLLKVINICAHVLDVAP
GPAIKAALPSLTNPPSLSPIRRKGEKEPEGEQASVPLSPKKGSEASAASRQSDTSGPVTTSK
SSSLGSFYHLPSYKLVKLDVLRKATHANYKVTLDLQNSTEKFGGFLRSALDVLSQLILELATLQD
IGKCVVEILGYLKSCFSREPMMATVCVQQLLKTLEGTNLASQFDGLSSNPKSQGRAQLGS
SSVRPGLYHYCFMAPYTHFTQALADASLRNMVQAEQENDTSGWFDVLQKVSTQLKTNLTSVT
KNRADKNAIHNHIRLFEPLVIKALKQYTTTTTCVQLQKQVLDLLAQVQLRVNYCLLSDQVF
IGFVLKQFEYIEVGQFRESEAIIPNIFFFLVLLSYERYHKSQIIGIPKIIQLCDGIMASGRK
AVTHAIPALQPIVHDLFVLRGTNKADAGKELETQKEVVSMLLRLIQYHQVLEMFILVLQOC
HKENEDKWKRLSRQIADIILPMLAKQQMHIDSHEALGVLNLTLEILAPSSLRPVDMLLRSMF
VTPNTMASVSTVQLWISGILAILRVLISQSTEDIVLSRIQELSFSPYLISCTVINRLRDGDS
TSTLEEHSEGKQIKNLPEETFSRFLQLVGILLEDIVTKQLKVEMSEQHTFYCQELGTLML
CLIHIFKSGMFRRITAAATRLFRSDGCGGSFYTLDSLNLRRARSMITTHPALVLLWCQIILLV
NHTDYRWVAEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLAAKLGMCNREIVRRGALILFCDY
VCQNLHDSEHLTWLIVNHIQDLISLSHEPPVQDFISAVHRNSAASGLFIQAIQSRCENLSTP
TMLKKTQLCLEGIHLSQSGAVLTLYVDRLCTPFRVLARMVDILACRRVEMLLAANLQSSMA
QLPMEELNRIQEYLQSSGLAQRHQRLYSLDDRFRSLTMQDLSLSPSPVSSHPLDGDGHVSLE
TVSPDKDWYVHLVKSQCWTRSDSALLEGAELVNRI PAEDMNAFMMNSEFNLSLLAPCLSLGM
SEISGGQKSALFEAAREVTLARVSGTVQQLPAVHHVFQPELPAEPAAYWSKLNDFGDAALY
QSLPTLARALAQYLVVSKLPSHLHLPPEKEKDIVKFVVATLEALSWHLIHEQIPLSLDLQA
GLDCCCLALQLPGLWSVVSSTEFVTHACSLIYCVHFILEAVAVQPGEQLLSPERRTNTPKAI
SEEEEEVDPNQNPKYITAACEMVAEMVESLQSVLALGHKRNSGVPAFLTPLLRNIIISLAR
LPLVNSYTRVPLVWKLGWSPKPGGDFGTAFPEIPEFLQEQEKFVFEFIYRINTLGWTSRTQ
FEETWATLLGVLVTQPLVMEQEEESPPEEDTERTQINVLAVQAITSLVLSAMTVPVAGNPAVS
CLEQQPRNKPLKALDTRFGRKLSIIRGIVEQEQIAMVSKRENIATHHLYQAWDPVPSLSPAT
TGALISHEKLLLQINPERELGMSYKLGQVS IHSVWLGNSITPLREEEWEDEEEEEEADAPAP
SSPPTSPVNSRKHRAVDIHSQSQFLELYSRWILPSSARRTPAI LISEVRSLLVVSDF
TERNQFELMYVTLELRRVHPSEDEILAQYLVPATCKAAAVLGMDKAVAEPVSRLLLESTLRS
SHLPSRVGALHGVLVLECDLLDDTAKQLIPVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT
AFYLIENYPLDVGPFEFSASIIQCMGVMLSGSEESTPSIIYHCALRGLERLLLSEQLSRLDAE
SLVKLSVDRVNVHSPHRAMAALGLMLTCMYTGKEKVS PGRTSDPNPAAPDSESVIVAMERS
VLFDRIRKGFPCEARVVARILPQFLDDFFPPQDIMNKVIGEFLSNQQPYPQFMATVVYKVFQ

TLHSTGQSSMVRDWMLSLSNFTQRAPVAMATWSLSCFFVSASTSPWVAAILPHVISRMGKL
EQVDVNLFLCLVATDFYRHQIEEELDRRAFQSVLEVVAAPGSPYHRLLTCLRNVHKVTTTC"

4. Gen HTT mutat en forma de nucleòtids (isoforma 1) (37 repeticions CAG) (només el 1r exó)

```
ctgccgtgcc gggcgggaga ccgccatggc gaccctggaa aagctgatga aggccttcga  
gtccctcaag tccttcagc agcagcagca gcagcagcag cagcagcagc agcagcagca  
gcagcagcag cagcagcagc agcagcagca gcagcagcag cagcagcagc agcagcagca  
gcagcagccg ccaccgccgc cgcgcgcgcc gccgcctcct cagcttcctc agccgccgcc  
gcaggcacag ccgctgctgc ctcagccgca gccgccccg ccgccgcccc cgccgccacc  
cggccccgct gtggctgagg agccgctgca ccgaccgtga gtttgggccc gctgcagctc  
cctgtcccgg cgggtcccag gctacggcgg ggatggcggg aaccctgcag cctgcgggcc  
ggcgacacga acccccggcc ccgcagagac agagtgaccc agcaaccagc agc
```