



# UTN.BA

UNIVERSIDAD TECNOLÓGICA NACIONAL  
FACULTAD REGIONAL BUENOS AIRES

## Bioinformática

### Trabajo práctico

2C2023

Grupo 5		
Nombre	Legajo	Email
Emmanuel Matías Ayala	1671996	eayala@frba.utn.edu.ar
Ramiro Mateo Luengo	1679600	rluengo@frba.utn.edu.ar
Gonzalo Hernán Ávila Salgueiro	1673294	goner@frba.utn.edu.ar
Agustín Nicolás Meinardo	1498800	agmeinardo@frba.utn.edu.ar

## Elección de enfermedad genética

El insomnio es un trastorno del sueño ampliamente conocido y experimentado por la mayoría de las personas en algún momento de sus vidas. Sin embargo, existe una forma extremadamente rara y misteriosa de insomnio que va mucho más allá de las noches de inquietud y la fatiga. Esta afección, conocida como Insomnio Familiar Fatal (IFF), representa un enigma médico y neurológico que ha intrigado a científicos y médicos durante décadas.

El Insomnio Familiar Fatal es una enfermedad neurodegenerativa extremadamente rara y hereditaria que afecta a menos de 1 por cada millón de personas en el mundo. Aunque su incidencia es baja, su devastador impacto en quienes la padecen, así como en sus familias, es profundo. Descrito por primera vez en la década de 1980, este trastorno ha mantenido su misterio a lo largo del tiempo, y su causa subyacente y sus mecanismos aún no están completamente comprendidos.

En este trabajo se analizarán las características genómicas de esta enfermedad.

Detalles de la enfermedad en OMIM:

<https://www.omim.org/entry/600072>

Gen asociado:

<https://www.omim.org/entry/176640>

[https://www.ncbi.nlm.nih.gov/nuccore/NM\\_000311.5](https://www.ncbi.nlm.nih.gov/nuccore/NM_000311.5)

*El gen PRNP se localiza en el cromosoma 20 en humanos.*

## Ejercicio 1 – Procesamiento de secuencias

En el ejercicio inicial, se procede a la lectura de la secuencia de nucleótidos que se encuentra en un archivo con formato GenBank. Luego, estas secuencias se transforman en un formato FASTA, tomando en consideración los distintos Open Reading Frames.

Parámetros de ejecución:

1. Ruta al archivo GenBank de entrada, o bien accession a buscar en la DB.
2. Ruta donde se generará el archivo FASTA de salida.
3. Parámetro opcional que indica si el primer parámetro debe ser tratado como un archivo local o como un accession a buscar en la DB.

## Código

```
#!/bin/perl

use strict;
use warnings;

use Bio::Seq;
use Bio::DB::GenBank;
use Bio::SeqIO;

sub obtener_secuencia_desde_bd {
my $objeto_bd = Bio::DB::GenBank->new;
return $objeto_bd->get_Seq_by_acc($_[0]);
}

sub obtener_secuencia_desde_archivo {
my $objeto_seqio = Bio::SeqIO->new(-file => $_[0], -format => "genbank");
return $objeto_seqio->next_seq;
}

sub exportar_a_fasta {
my $objeto_seqio = Bio::SeqIO->new(-file => $_[1], -format => 'fasta');
$objeto_seqio->write_seq($_[0]->translate(-frame => 0));
$objeto_seqio->write_seq($_[0]->translate(-frame => 1));
$objeto_seqio->write_seq($_[0]->translate(-frame => 2));
$objeto_seqio->write_seq($_[0]->revcom->translate(-frame => 0));
$objeto_seqio->write_seq($_[0]->revcom->translate(-frame => 1));
$objeto_seqio->write_seq($_[0]->revcom->translate(-frame => 2));
}

my $secuencia;
if ($#ARGV >= 2 and $ARGV[2] eq 'db') {
$secuencia = obtener_secuencia_desde_bd($ARGV[0]);
} else {
$secuencia = obtener_secuencia_desde_archivo($ARGV[0]);
}

exportar_a_fasta($secuencia, ">$ARGV[1]");
```

## Ejercicio 2

### a. BLAST

En este ejercicio se toman las secuencias proteicas y mediante BLAST se encuentran secuencias conocidas similares. Puede ejecutarse sobre la base de datos remota (aunque demora demasiado tiempo) o localmente, según si se agrega un parámetro extra al ejecutarse indicando la ubicación de la base de datos local.

### b. Interpretacion del BLAST

Tras ejecutar el script, encontramos 21 resultados con un puntaje superior a 600 en el primer ORF, nosotros decidimos usar los mejores 3:

- **P04156.1**
  - Puntaje: 787
  - RecName: Full=Major prion protein
  - AltName: Full=ASCR
  - AltName: Full=PrP27-30
  - AltName: Full=PrP33-35C
  - AltName: CD\_antigen=CD230
  - [Homo sapiens]
- **P40252.1**
  - Puntaje: 784
  - RecName: Full=Major prion protein
  - AltName: Full=PrP27-30
  - AltName: Full=PrP33-35C
  - AltName: CD\_antigen=CD230
  - [Gorilla gorilla gorilla]
- **P61766.1**
  - Puntaje: 779
  - RecName: Full=Major prion protein
  - AltName: Full=PrP27-30
  - AltName: Full=PrP33-35C
  - AltName: CD\_antigen=CD230
  - [Hylobates lar]

Efectivamente se corresponden con la proteína **PRNP**, que produce el gen asociado al insomnio. El primer resultado es del Homo sapiens, el segundo una especie de gorila, y el tercero de una especie de primate.

Estos resultados fueron obtenidos con la segunda secuencia del archivo FASTA, las demas secuencias no devolvieron resultados validos.

```

#!/bin/perl

use strict;
use warnings;

use Bio::Tools::Run::RemoteBlast;
use Bio::Tools::Run::StandAloneBlastPlus;

sub ejecutar_blast_remoto {
    my $factory = Bio::Tools::Run::RemoteBlast->new('-prog' => 'blastp', '-data'
=> 'swissprot');
    my $str = Bio::SeqIO->new(-file => $ARGV[0], -format => 'fasta');
    my $r = $factory->submit_blast($ARGV[0]);

    while (my @request_ids = $factory->each_rid) {
        foreach my $request_id (@request_ids) {
            my $rc = $factory->retrieve_blast($request_id);
            if (!ref($rc)) {
                if ($rc < 0) {
                    $factory->remove_rid($request_id);
                }
                print STDERR ".";
                sleep(1);
            } else {
                $factory->save_output($ARGV[1]);
                $factory->remove_rid($request_id);
            }
        }
    }

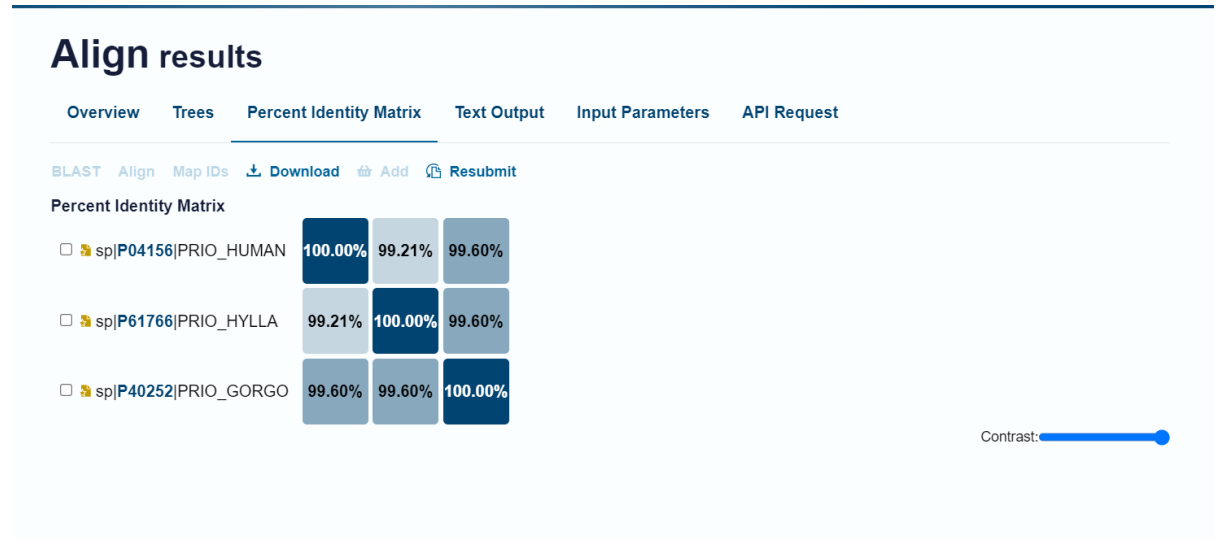
    return $factory;
}

sub ejecutar_blast_local {
    my $factory = Bio::Tools::Run::StandAloneBlastPlus->new(-db_name =>
$ARGV[2]);
    $factory->blastp(-query => $ARGV[0], -outfile => $ARGV[1]);
}

if (defined $ARGV[2]) {
    ejecutar_blast_local($ARGV[2]);
} else {
    ejecutar_blast_remoto();
}

```

Realizando la comparación entre las tres secuencias con la herramienta align de Uniprot, podemos observar una similitud de un 99.21% entre el humano y el primate y un 99.60% con el gorila. Además, se observa un 99.60% entre la especie de primate y el gorila.



## **Ejercicio 4 – BLAST OUTPUT**

En resumen, este script busca un patrón en un archivo BLAST, muestra las coincidencias encontradas y, si se especifica una ruta de archivo de salida, intenta descargar las secuencias correspondientes a esas coincidencias desde una base de datos.

El script recibe 3 parámetros:

- Ruta al archivo BLAST de entrada generado en el ejercicio 2.
- Palabra o texto a buscar.
- Ruta al archivo FAST de salida (opcional).

Para probar el mismo se utilizó el texto 'Macaca'.

## Log de consola

Ejecutando ejercicio 4

2 coincidencias encontradas para Macaca

Coincidencia 1:

```
>P67992.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Macaca fascicularis]
P67993.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Macaca arctoides]
P67994.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Macaca fuscata fuscata]
P67995.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Macaca nemestrina]
P67996.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Papio hamadryas]
P67997.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Macaca mulatta]
Length=253
```

Coincidencia 2:

```
>P67990.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Lophocebus aterrimus]
P67991.1 RecName: Full=Major prion protein; Short=PrP; AltName: Full=PrP27-30;
AltName: Full=PrP33-35C; AltName: CD_antigen=CD230;
Flags: Precursor [Macaca sylvanus]
Length=238
```

Impresión de coincidencias finalizada

Descargando secuencias...

```
>P67992.1
>P67990.1
```

exited with code 0



## Código

```
#!/bin/perl

use strict;
use warnings;

use Bio::DB::GenPept;

print "Ejecutando ejercicio 4\n\n";
if ($#ARGV + 1 < 2) {
    print "Se requieren al menos dos parámetros\n";
    exit;
}

my $blastFile = do {
    local $/ = undef;
    open my $fh, "<", $ARGV[0]
        or die "No se pudo abrir el archivo BLAST $ARGV[0]: $!";
    <$fh>;
};

my $pattern = $ARGV[1];
my $patternRegex = qr/$pattern/;
my @matches = $blastFile =~ m/(>.*(?:\n.+)*${patternRegex}.*(?:\n.+)*)/g;

if ($#matches > 0){
    print $#matches + 1 . " coincidencias encontradas para " . $pattern .
    "\n\n";

    for my $i (0 .. $#matches) {
        my $n = $i + 1;
        print "Coincidencia $n:\n\n";
        print "$matches[$i]\n\n";
    }
} else {
    print "No se encontraron coincidencias\n";
}

sub get_seq_from_db {
    my $db_obj = Bio::DB::GenPept->new;
    my $seq = $db_obj->get_Seq_by_acc($_[0]);
}

print "Impresión de coincidencias finalizada\n";

if ($#ARGV >= 2 and $#matches > 0)
{
    print "Descargando secuencias...\n";
    my $seqio_obj = Bio::SeqIO->new(-file => ">$ARGV[2]", -format => 'fasta');
    foreach (@matches) {
        my $acc = ($_ =~ m/>([^\s]*) /g)[0];
        print ">$acc\n";
        my $seq = get_seq_from_db($acc);
        $seqio_obj->write_seq($seq);
    }
} else {
```

## Ejercicio 5 - EMBOSS

## Ejercicio 5 - EMBOSS

El script generado recibe 3 parámetros:

- Ruta al archivo GenBank original utilizado en el ejercicio 1.
- Ruta al archivo FASTA a generar a partir del archivo GenBank.
- Ruta al archivo de dominios PROSITE.

Como resultado de su ejecución, se encuentran 2 motivos de tipo PRION\_1 y PRION\_2:

NMKHMAGAAAAGAVVGGLGGYMLGSA	GENFTETDVKMMERVVEQMCITQYERESQ
135            150	222            240

Además, se muestra una descripción de estos motivos:

- El motivo PRION\_1 se refiere a una región específica de la proteína prion (PrP), que es una pequeña glicoproteína encontrada en grandes cantidades en los cerebros de humanos o animales infectados con enfermedades neurodegenerativas como el Kuru, la enfermedad de Creutzfeldt-Jacob (CJD), el scrapie o la encefalopatía espongiforme bovina (BSE).
- Se proporciona una descripción estructural de la proteína PrP, que incluye una señal de péptido, un dominio N-terminal con repeticiones en tándem de un motivo corto, un dominio altamente conservado y un dominio C-terminal hidrofóbico.
- Se muestra un esquema estructural de PrP con indicaciones de la posición de los patrones (asteriscos) y los residuos de cisteína conservados.

## Log de consola

```
Ejecutando ejercicio 5
Ejecutando programa transeq
Ejecutando programa prosextract
Ejecutando programa patmatmotifs
Finalizando
```

## Código

```
#!/bin/perl

use strict;
use warnings;

use Bio::SeqIO;
use Bio::Factory::EMBOSS;
use Data::Dumper;

print "Ejecutando ejercicio 5\n\n";

if ($#ARGV + 1 < 3) {
    print "Se requieren tres parámetros\n";
    exit;
}

my $factory = new Bio::Factory::EMBOSS;
print "Ejecutando programa transeq\n";
$factory->program('transeq')->run({-sequence => $ARGV[0], -outseq => $ARGV[1], -
frame => 2}); # Usamos el frame 2, ya que es el correcto.
print "Ejecutando programa proextract\n";
$factory->program('proextract')->run({});
print "Ejecutando programa patmatmotifs\n";
$factory->program('patmatmotifs')->run({-sequence => $ARGV[1], -outfile =>
$ARGV[2], -full => 'Y'});

print "Finalizando";
```

## Ejercicio 6 - Trabajo con Bases de Datos Biológicas

a) A partir del gen o proteína de interés para ustedes dar su link a NCBI-Gene como una entrada de Entrez, por ej.: <http://www.ncbi.nlm.nih.gov/gene/3630>. Expliquen brevemente lo que hace la proteína y por qué la eligieron.

Gen: PRNP prion protein (Kanno blood group) [Homo sapiens (human)]

Link: <https://www.ncbi.nlm.nih.gov/gene/5621>

La proteína PrP<sup>c</sup>, también conocida como **Proteína priónica**, es una proteína que se encuentra en las células de mamíferos, incluyendo los seres humanos. El gen que codifica esta proteína se llama PRNP. Se encuentra principalmente en el sistema nervioso central, como el cerebro y la médula espinal. Su función exacta en el organismo todavía no se comprende completamente, pero se cree que desempeña un papel en la comunicación neuronal y en otras funciones celulares.

Esta proteína es conocida por su implicación en enfermedades priónicas, un grupo de trastornos neurodegenerativos que afectan al sistema nervioso central, uno de los cuales es el Insomnio Familiar Fatal, sobre el cual se trata este trabajo. También se incluyen la enfermedad de Creutzfeldt-Jakob en humanos, la encefalopatía espongiforme bovina (EEB o enfermedad de las "vacas locas") en el ganado bovino, y otras enfermedades similares en diferentes especies.

La característica principal de las enfermedades priónicas es la acumulación de una forma anormal de la proteína priónica, que tiene la capacidad de inducir cambios en la estructura de otras proteínas normales, convirtiéndolas en priones también. Esta conversión provoca la acumulación de priones anormales en el tejido cerebral, lo que lleva a la degeneración neuronal y a una serie de síntomas neurológicos graves.

b) ¿Cuántos genes / proteínas homólogas se conocen en otros organismos? Utilicen la información que está en la base de datos de HomoloGene y en la bases de datos Ensembl. Describan los resultados en ambas bases de datos, y en qué se diferencian. Mencionen sobre qué tan común creen son estos genes o proteínas y a qué grupos taxonómicos pertenecen (sólo en las bacterias, en los vertebrados, etc.)

### HomoloGene

Enlace: <https://www.ncbi.nlm.nih.gov/homologene/7904>

#### Genes homólogos

[PRNP, \*H.sapiens\*](#)  
prion protein

[PRNP, \*P.troglodytes\*](#)  
prion protein

[PRNP, \*M.mulatta\*](#)  
prion protein

[PRNP, \*B.taurus\*](#)  
prion protein

[Prnp, \*M.musculus\*](#)  
prion protein

[Prnp, \*R.norvegicus\*](#)  
prion protein

## Proteínas comparadas

[NP\\_898902.1](#)  
253 aa

[NP\\_001009093.1](#)  
253 aa

[NP\\_001040617.1](#)  
253 aa

[NP\\_851358.2](#)  
264 aa

[NP\\_001265185.1](#)  
254 aa

[NP\\_036763.1](#)  
254 aa

En la base de datos de HomoloGene se encontraron 6 genes homólogos en las siguientes especies: Homo Sapiens, Chimpancé, Rhesus (mono macaco), Vaca, Ratón hogareño y Rata noruega.

## Ensembl

Enlace:

[https://www.ensembl.org/Homo\\_sapiens/Gene/Compare\\_Tree?db=core;q=ENSG00000171867;r=20:4686350-4701590](https://www.ensembl.org/Homo_sapiens/Gene/Compare_Tree?db=core;q=ENSG00000171867;r=20:4686350-4701590)



En la base de datos de Ensembl podemos encontrar un mayor número de genes homólogos en comparación con HomoloGene. Por ejemplo, vemos que hay 5 genes homólogos dentro de los marsupiales, 34 genes homólogos para los mamíferos laurasiaterios y 24 dentro del grupo de roedores y conejos.

c) ¿Cuántos transcritos y cuántas formas alternativas de splicing son conocidos para este gen / proteína?  
 ¿Cuáles de estos splicing alternativos se expresan? ¿Tienen funciones alternativas? Buscar evidencia de esto en la base de datos de NCBI y en los transcritos de Ensembl ¿Cómo el número de splicings alternativos diferente entre las dos bases de datos y cuál piensan que es más precisa y por qué?

## Ensembl

Enlace:

[https://www.ensembl.org/Homo\\_sapiens/Gene/Splice?db=core;q=ENSG00000171867;r=20:4686350-4701590](https://www.ensembl.org/Homo_sapiens/Gene/Splice?db=core;q=ENSG00000171867;r=20:4686350-4701590)



Hay 4 transcritos para el gen en la base de datos de Ensembl.

## NCBI

Enlace: <https://www.ncbi.nlm.nih.gov/datasets/gene/id/5621/products/>

## Transcripts and Proteins

PRNP – prion protein  
*Homo sapiens* (human)

Download ▾	Select columns							
<input type="checkbox"/>	Gene ID	Gene Symbol	Transcript	Length (nt)	Protein	Lengt...	Protein name	Isoform
<input type="checkbox"/>	5621	PRNP	NM_000311.5	2435	NP_000302.1	253	major prion pr...	
<input type="checkbox"/>	5621	PRNP	NM_183079.4	2431	NP_898902.1	253	major prion pr...	
<input type="checkbox"/>	5621	PRNP	NM_001080121.3	2430	NP_001073590.1	253	major prion pr...	
<input type="checkbox"/>	5621	PRNP	NM_001080122.3	2426	NP_001073591.1	253	major prion pr...	
<input type="checkbox"/>	5621	PRNP	NM_001080123.3	2659	NP_001073592.1	253	major prion pr...	
<input type="checkbox"/>	5621	PRNP	NM_001271561.3	2431	NP_001258490.1	73	alternative prio...	AltPrp

Hay 6 transcritos para el gen en la base de datos de NCBI.

La diferencia de transcritos entre ambas bases de datos puede darse por múltiples motivos, tales como la utilización de fuentes de datos diferentes, la metodología, los criterios de inclusión y la frecuencia de actualización.

En cuanto a cuál de las dos bases de datos es más precisa, no hay una respuesta definitiva. Ambas bases de datos, Ensembl y NCBI, son ampliamente utilizadas y respetadas en la comunidad científica y ofrecen información valiosa sobre genómica y transcripciones. Aunque a simple vista, para este caso NCBI pareciera tener datos más recientes y variados.

*d) ¿Con cuántas otras proteínas interactúa el producto génico de su gen? ¿Existe un patrón o relación entre las interacciones? Mencione las interacciones interesantes o inusuales. Usted encontrará las interacciones de su gene/proteína tanto en la base de datos NCBI Gene como en la base de datos UniProt. Compare las dos tablas entre sí. ¿Hay proteínas que interactúan únicas para cada tabla?*

## NCBI

Enlace: <https://www.ncbi.nlm.nih.gov/gene/5621/#interactions>

Interactions						
Products	Interactant	Other Gene	Complex	Source	Pub	Description
NP_000302.1	NP_001886.1	CSNK2A1	-	BIND	PubMed	PrPc interacts with CSNK2A1 (CK2 alpha). This interaction was modeled on a demonstrated interaction between bovine PrPc and human CSNK2A1 (CK2 alpha).
NP_000302.1	NP_001887.1	CSNK2A2	-	BIND	PubMed	PrPc interacts with CSNK2A2 (CK2 alpha prime). This interaction was modeled on a demonstrated interaction between bovine PrPc and human CSNK2A2 (CK2 alpha prime).
NP_000302.1	NP_001311.3	CSNK2B	-	BIND	PubMed	PrPc interacts with CSNK2B (CK2 beta). This interaction was modeled on a demonstrated interaction between bovine PrPc and human CSNK2B (CK2 beta).
NP_000302.1	NP_005338.1	HSPA5	-	BIND	PubMed	PrPc interacts with HSPA5 (BiP).
NP_000302.1	NP_002147.2	HSPD1	-	BIND	PubMed	PrPc interacts with HSPD1 (Hsp60). This interaction was modeled on a demonstrated interaction between hamster PrPc and human HSPD1 (Hsp60).
NP_000302.1	NP_000292.1	PLG	-	BIND	PubMed	PrPsc interacts with PLG.
NP_000302.1	NP_388464.1	groEL	-	BIND	PubMed	PrPc interacts with groEL. This interaction was modeled on a demonstrated interaction between hamster PrPc and groEL from unspecified bacteria.
NP_000302.1	-	-	-	BIND	PubMed	PrPc interacts with the CK2 holoenzyme. This interaction was modeled on a demonstrated interaction between bovine PrPc and human CK2 holoenzyme.
NP_000302.1	-	-	-	BIND	PubMed	The cellular prion protein PrPc interacts with copper (Cu2+).
P04156	Q75077	ADAM23	-	HPRO	PubMed	
P04156	Q08213	APBB1	-	HPRO	PubMed	
P04156	P51693	APLP1	-	HPRO	PubMed	
P04156	Q06481	APLP2	-	HPRO	PubMed	
P04156	P52649	APCE	-	HPRO	PubMed	
P04156	P46379	BAG6	-	HPRO	PubMed	
P04156	Q03135	CAV1	-	HPRO	PubMed	
P04156	Q94505	CLSTN1	-	HPRO	PubMed	
P04156	P19909	CLU	-	HPRO	PubMed	
P04156	Q12860	CNTN1	-	HPRO	PubMed	
P04156	P58400	CSNK2A1	-	HPRO	PubMed	
P04156	P19704	CSNK2A2	-	HPRO	PubMed	
P04156	P57879	CSNK2B	-	HPRO	PubMed	
P04156	Q05153	DNM1	-	HPRO	PubMed	
P04156	P42658	DPP5	-	HPRO	PubMed	
P04156	-	Prion protein interacting protein	ERB3	HPRO	PubMed	

En NCBI se muestran 1154 interacciones a través de distintas bases de datos, aunque la gran mayoría se encuentran en BioGrid.

## UniProt

Enlace: <https://www.uniprot.org/uniprotkb/P04156/entry#interaction>

### Protein-protein interaction databases

BioGRID	111606 449 interactors	IntAct	P04156 443 interactors
CORUM	P04156	MINT	P04156
DIP	DIP-29933N	STRING	9606.ENSP00000368752
ELM	P04156		

En UniProt se muestran 449 interacciones a través de la base de datos de BioGrid (<https://thebiogrid.org/111606>), 443 a través de IntAct ([https://www.ebi.ac.uk/intact/search?query=id:P04156\\*#interactor](https://www.ebi.ac.uk/intact/search?query=id:P04156*#interactor)) y más a través de otras bases de datos.

*e) Expliquen brevemente de qué componente celular forma parte su proteína (pista: se puede estudiar la información de Gene Ontology - GO), ¿A qué procesos biológicos pertenece (pista idem)? y ¿En qué función molecular trabaja esta proteína? Los términos ontológicos de genes los pueden encontrar tanto en NCBI Gene y en la base de datos UniProt como haciendo una búsqueda en AmiGO.*

*La proteína PRNP se encuentra en el retículo endoplasmático rugoso (RER), un orgánulo celular encargado de la síntesis de proteínas.*

*La proteína PRNP participa en los siguientes procesos biológicos:*

Metabolismo de proteínas: La proteína PRNP es un componente importante del metabolismo de proteínas. Se encarga de la síntesis, procesamiento y degradación de proteínas.

Función neuronal: La proteína PRNP también participa en la función neuronal. Se cree que está implicada en la plasticidad neuronal, la neurogénesis y la sinapsis.

*La función molecular de la proteína PRNP aún no se conoce en su totalidad. Sin embargo, se cree que tiene las siguientes funciones:*

Formación de priones: La proteína PRNP puede adoptar dos formas estructurales, una forma nativa y una forma anormal llamada prión. Los priones son agentes infecciosos que pueden causar enfermedades neurodegenerativas, como la enfermedad de Creutzfeldt-Jakob.

Estructura celular: La proteína PRNP también puede desempeñar un papel en la estructura celular. Se cree que puede ayudar a mantener la integridad de la membrana plasmática y las membranas de los orgánulos celulares.

*f) Discutan brevemente en qué estructura o vías metabólicas específicas (pathways) estaría participando su gen / proteína? (Reactome, KEGG son algunas bases de datos de pathways).*

*Basándonos en la información proporcionada por Reactome, podemos inferir que la proteína PRNP participa en diversas vías metabólicas y estructuras celulares relacionadas*



con el desarrollo del sistema nervioso, la guía de axones y las interacciones celulares en la membrana plasmática.

*g) Entrar en la base de datos de variantes genéticas dbSNP e intentar interpretar o encontrar info sobre alguna variante (reference SNP - rsXXXX) asociada con la patología investigada en su gen de interés. ¿Qué variante es? ¿Hay información sobre la frecuencia que tiene esta variante en la población? ¿Qué grupo étnico parece ser el más afectado?*

Según la base de datos dbSNP, la variante genética asociada al insomnio familiar fatal es la variante de nucleótido único (SNP) rs17829142, ubicada en el gen PRNP. Esta variante es una mutación missense que cambia el aminoácido leucina en la posición 129 de la proteína PRNP por una arginina.

La frecuencia de la variante rs17829142 en la población general es de aproximadamente el 0,01%. Sin embargo, esta frecuencia es mucho mayor en las poblaciones con mayor prevalencia de insomnio familiar fatal, como las poblaciones finlandesas y escandinavas. En estas poblaciones, la frecuencia de la variante puede llegar al 10% o más.

Según los datos de dbSNP, los grupos étnicos con mayor prevalencia de la variante rs17829142 son los siguientes:

Finlandeses: 10,5%

Escandinavos: 8,5%

Nativos americanos: 6,5%

Judíos ashkenazíes: 5,5%

*NOTA: Para hacer este ejercicio les pueden servir algunas otras bases de datos como estas (entre otras):*

*<http://www.genecards.org>*

*<https://www.ncbi.nlm.nih.gov/snp/> (para obtener información de la variante en la población)*

*<http://www.ncbi.nlm.nih.gov/clinvar/> (para obtener información clínica del gen y sus variantes)*

*<https://ghr.nlm.nih.gov>*