



UNIVERSIDAD
DE GRANADA

Seminario 1

Herramientas bioinformáticas en Genómica

Biología Molecular - Grado en Medicina
Curso 2023/24



¿Qué vamos a hacer?

Vamos a usar bases de datos y herramientas bioinformáticas para conocer mejor un gen cuyas mutaciones causan una enfermedad.

Durante el seminario buscaremos información sobre el gen *KRAS*. Iremos respondiendo a las preguntas que he subido a PRADO y las entregaremos al final de la clase.

En el examen habrá preguntas **sencillas** tipo test de opción múltiple.

CUESTIONES SEMINARIO 1: HERRAMIENTAS BIOINFORMÁTICAS EN GENÓMICA Grado en Medicina | Biología Molecular | Curso 2023/24

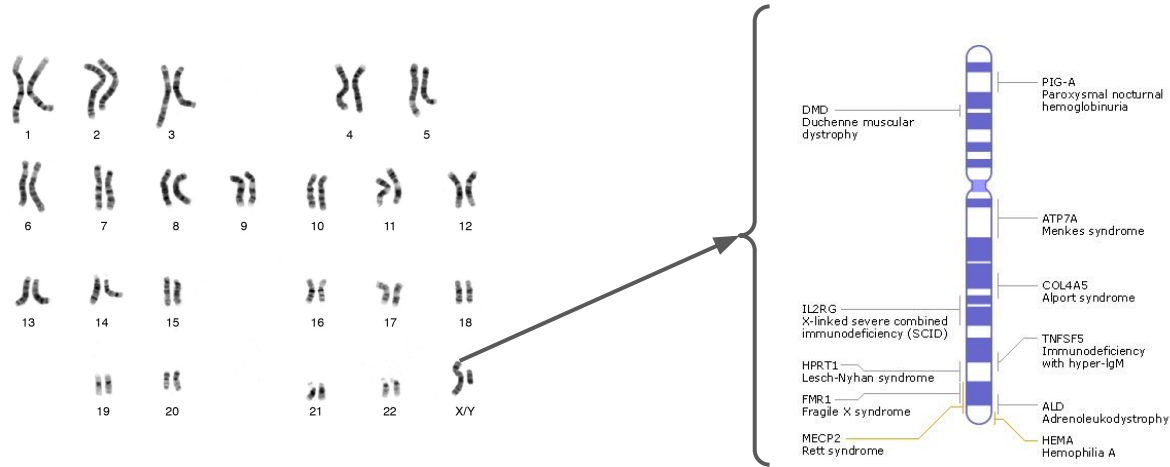
Nombre y Apellidos:

1. Sobre el gen *KRAS*:
 - a) ¿Cuál es su Ensembl ID?
 - b) ¿En qué cromosoma, coordenadas y hebra se localiza?
 - c) Nombra sus 2 genes más próximos.
 - d) Indica el dbSNP ID y la consecuencia de una variante patogénica de este gen.
2. Sobre la isoforma (transcrito) de referencia del gen *KRAS*:
 - a) ¿Cuál es su Ensembl ID?
 - b) ¿Cuál es su longitud en pares de bases y en aminoácidos?
 - c) ¿Cuántos exones tiene?
3. Indica el identificador y la secuencia de un miRNA maduro que regule la expresión del gen *KRAS*.
4. ¿Qué laboratorio en España realiza el diagnóstico genético de mutaciones en *KRAS*?
5. Nombra otra línea celular distinta a NCI-H358 que contenga la mutación G12C en *KRAS*.
6. Indica la pareja de primers que has diseñado para amplificar el exón 2 completo de *KRAS*.



¿Qué es la Genómica?

Es el campo de la Biología Molecular dedicado al estudio de **todo el material genético** de un organismo. Con técnicas de Biología Molecular y Bioinformática, se estudia la **estructura, mapeo, función**, evolución y edición de los genomas.

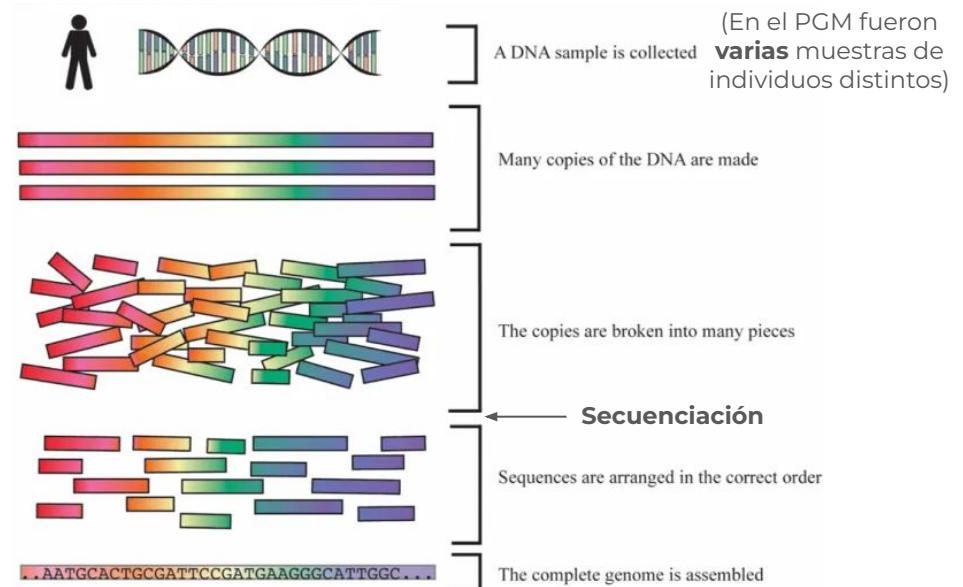


Proyecto Genoma Humano (1999-2003)

Secuenciar el genoma humano (3200 millones de pares de bases) y **publicar** los datos:

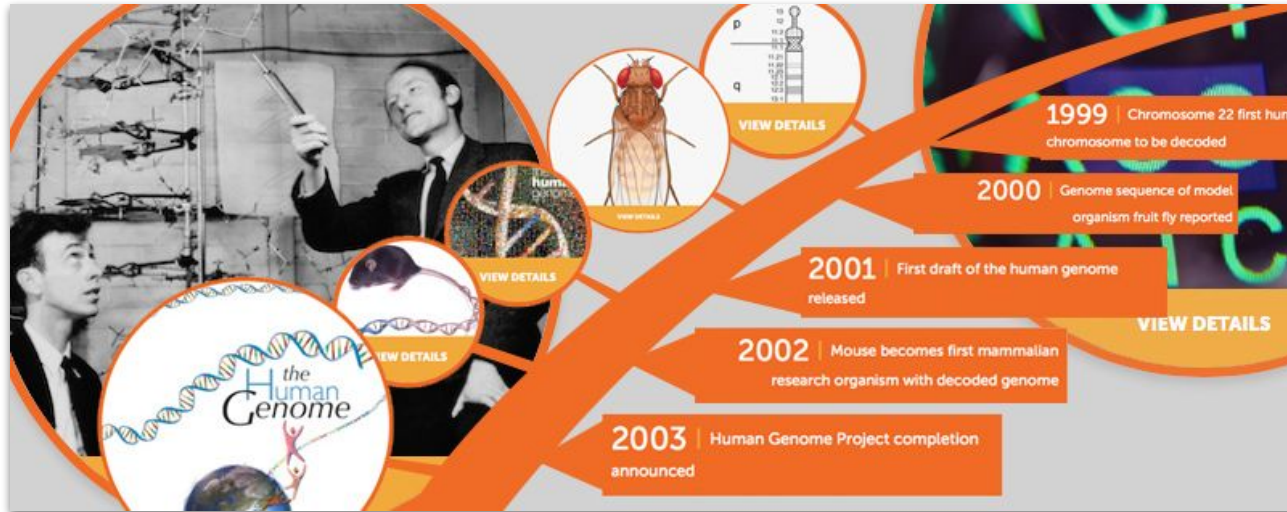
- Identificar genes y regiones regulatorias importantes
- Entender su papel en las enfermedades
- Investigar nuestros orígenes

Proceso de secuenciación de un genoma



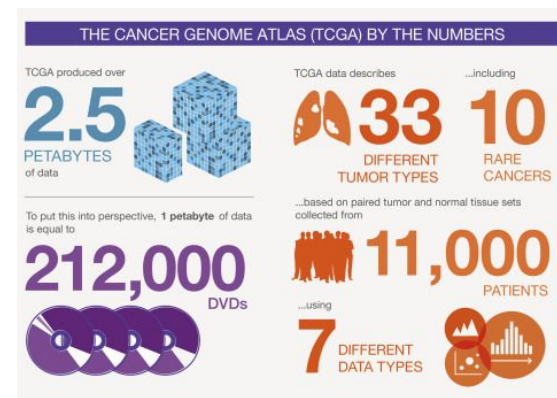
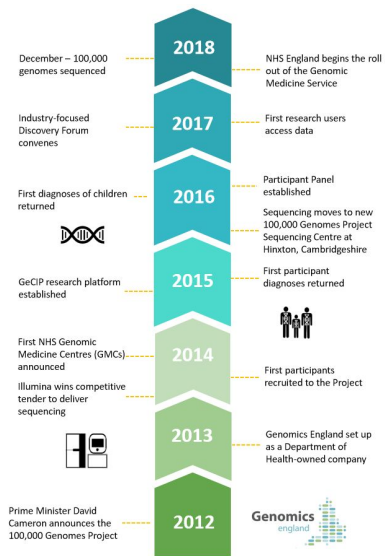
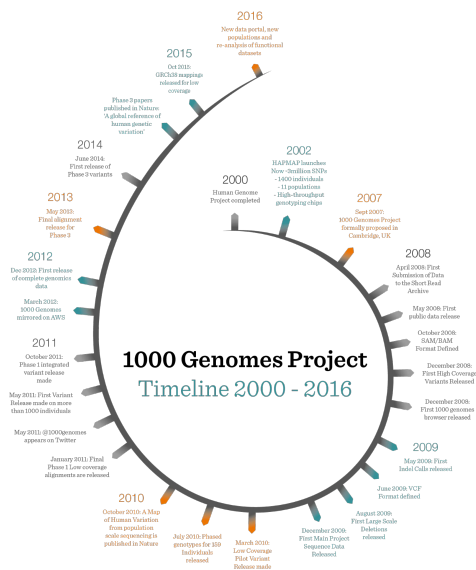
Proyecto Genoma Humano (1999-2003)

- El ADN secuenciado procedía de **varios individuos** donantes anónimos.
- La secuenciación se llevó a cabo entre **20 laboratorios colaboradores** de Estados Unidos, Reino Unido, Francia, Alemania, Japón y China y se completó en **4 años**.



Nuevos proyectos genómicos

Secuenciar a humanos **individualmente** para conocer la **variación genética** humana.



Proyecto **1000** genomas

Proyecto **100000** genomas

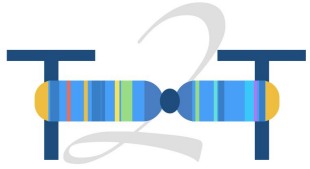
Proyecto **The Cancer Genome Atlas (TCGA)**

Proyecto Telomere-to-Telomere (T2T, 2021)

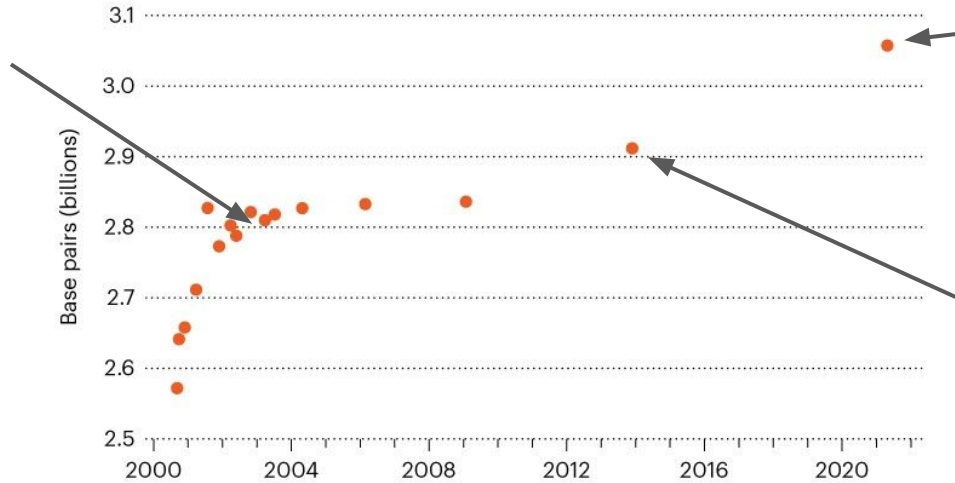
COMPLETING THE HUMAN GENOME

Researchers have been filling in incompletely sequenced parts of the human reference genome for 20 years, and have now almost finished it, with 3.05 billion DNA base pairs.

Genoma secuenciado por el Proyecto Genoma Humano
(Año 2003, ~2800 millones de pares de bases)



TELOMERE-TO-TELOMERE CONSORTIUM



0.3% of sequence might still have errors. Includes X but not Y chromosome. Count excludes mitochondrial DNA.

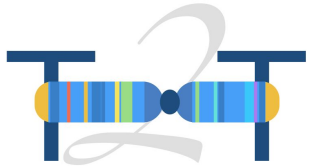
©nature

T2T-CHM13: Versión más completa del genoma humano
(Año 2021, ~3050 millones de pares de bases)

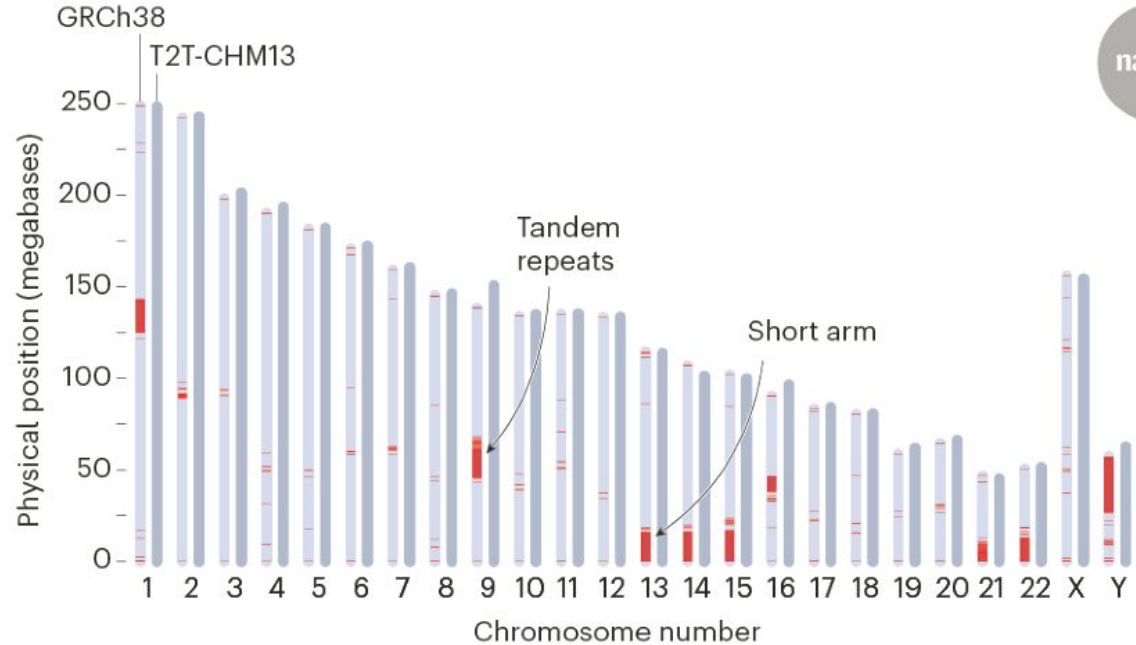
GRCh38: Última versión del genoma humano antes del proyecto T2T
(Año 2013, ~2900 millones de pares de bases)

Proyecto Telomere-to-Telomere (T2T, 2021)

T2T ha revelado regiones **repetitivas** del genoma hasta ahora desconocidas localizadas cerca de **centrómeros** y **telómeros**, principalmente.



TELOMERE-TO-TELOMERE CONSORTIUM



Bases de datos

Nucleic Acids
Research 

Nucleic Acids Research, 2024, **52**, D1–D9
<https://doi.org/10.1093/nar/gkad1173>
Advance access publication date: 30 November 2023
Editorial

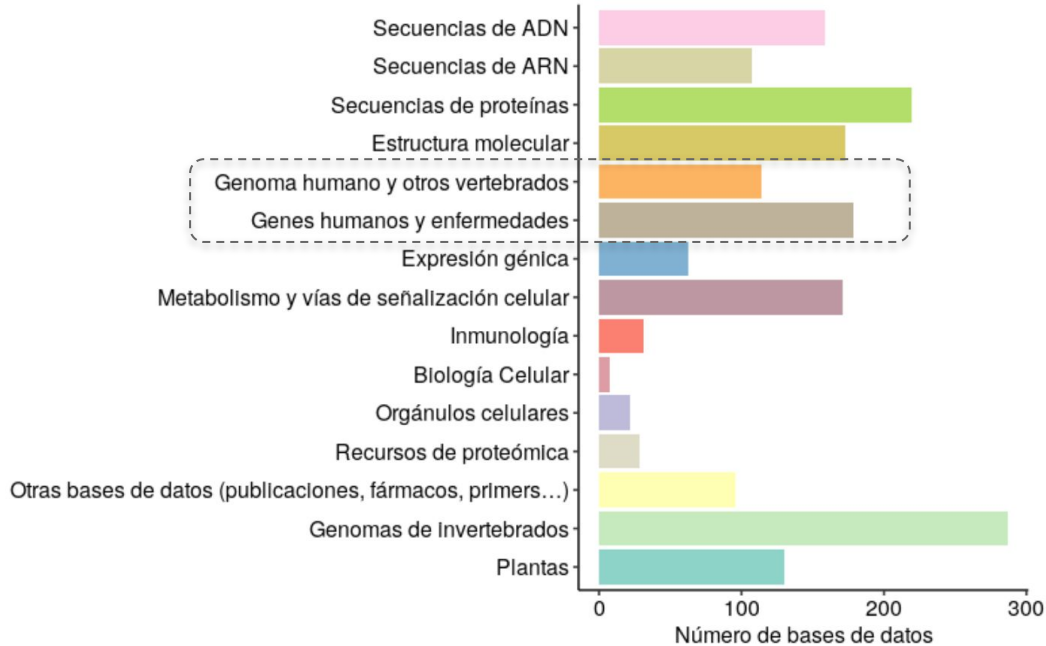
The 2024 *Nucleic Acids Research* database issue and the online molecular biology database collection

Daniel J. Rigden^{1,*} and Xosé M. Fernández ²

- La revista *Nucleic Acids Research* recopila anualmente todas las bases de datos de Biología Molecular que existen (actualmente **1959**)
- En 2024 se han añadido **90** bases de datos nuevas.
- Puedes explorar el catálogo completo [aquí](#).



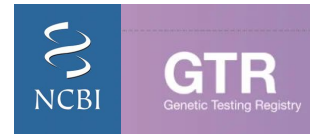
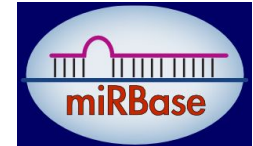
Bases de datos



En esta práctica vamos a usar las siguientes bases de datos y herramientas bioinformáticas:

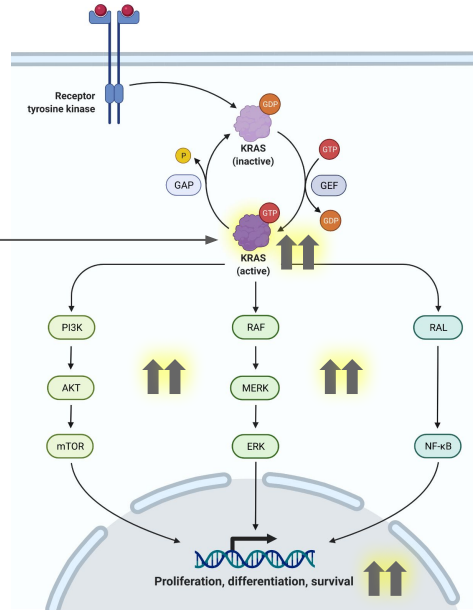
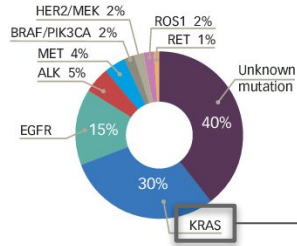
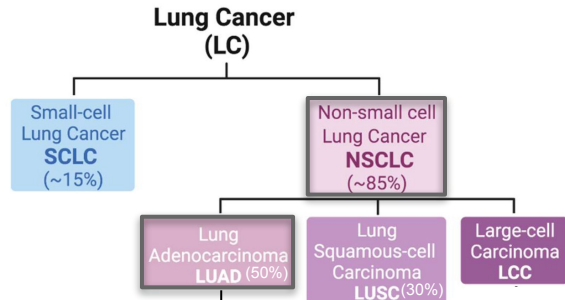


Primer-BLAST



Gen *KRAS* en cáncer de pulmón

El **oncogen *KRAS*** está mutado en un **30%** de los casos de **adenocarcinoma** de pulmón (el tipo de cáncer de pulmón más frecuente).



Las mutaciones en *KRAS* hacen que esta proteína esté **más activa**, lo cual causa un **aumento descontrolado de la proliferación celular**.

La mayoría de mutaciones están en el **exón 2**. Las más frecuentes son dos mutaciones del aminoácido 12 (**G12C** y **G12D**).

Antes de empezar

Consideraciones para seguir mejor la práctica:

- Los **rectángulos rojos** de las diapositivas indican dónde hay que escribir o hacer click en las páginas web.
- La información oculta con **recuadros blancos** son las respuestas a las preguntas de la entrega.
- Si en algún momento te pierdes (¡es totalmente normal!), levanta la mano.



¿Hay algún gen asociado a esta enfermedad?

Respuesta:  DisGeNET

[Enlace](#)

Buscar en DisGeNet

The screenshot shows the DisGeNet website interface. At the top, there is a navigation bar with links: Home, About, Search, Browser, API, Downloads, Cytoscape, RDF, disgenet2r, Help, COVID-19, Login, Signup, and the DisGeNET logo. Below the navigation bar, there are three tabs: 'diseases' (selected), 'genes', and 'variants'. A search input field contains the text 'Non-small cell lung carcinoma'. A red box highlights this input field. Below the input field, a list of search results is displayed, including the following entries:

- Name: Non-Small Cell Lung Carcinoma; CUI: C0007131
Ngenes: 3926 - Nsnps: 712
- Name: Stage IB Non-Small Cell Lung Carcinoma; CUI: C1336119
Ngenes: 1 - Nsnps: 0
- Synonym: Stage I Non-Small Cell Lung Carcinoma; CUI: C0278504
Ngenes: 115 - Nsnps: 1
- Synonym: Squamous Non-Small Cell Lung Carcinoma; CUI: C4509816
Ngenes: 30 - Nsnps: 0
- Synonym: Non-Squamous Non-Small Cell Lung Carcinoma; CUI: C4324656
Ngenes: 20 - Nsnps: 4
- Synonym: stage IA non-small cell lung carcinoma; CUI: C1336119
Ngenes: 2 - Nsnps: 0
- Synonym: stage IB non-small cell lung carcinoma; CUI: C4724264
Ngenes: 1 - Nsnps: 0

On the left side of the search results area, there is a red text annotation: **Buscamos el nombre de la enfermedad**.

Asociaciones mutación-enfermedad

Home About Search Browser API Downloads Cytoscape RDF disgenet2r Help COVID-19 Login Signup DisGeNET

Non-Small Cell Lung Carcinoma

Name: Non-Small Cell Lung Carcinoma
UMLS CUI: C0007131
Type: disease
MeSH Class: Neoplasms; Respiratory Tract Diseases
MeSH: D002289
OMIM: 211980
Semantic Type: Neoplastic Process
Phenotypic abnormality: Neoplasm; Abnormality of the respiratory system
Disease Ontology: disease of anatomical entity; disease of cellular proliferation

Similar diseases

Summary of Gene-Disease Associations

Evidences for Gene-Disease Associations

Summary of Variant-Disease Associations

Evidences for Variant-Disease Associations

Summary of Disease-Disease Associations

Disease Mappings

Vemos el resumen de asociaciones variante-enfermedad

Asociaciones mutación-enfermedad

Home About Search Browser API Downloads Cytoscape RDF disgenet2r Help COVID-19 Login Signup DisGeNET

Summary of GDAs Evidences for GDAs Summary of VDAs Evidences for VDAs Summary of DDAs Disease Mappings

Non-Small Cell Lung Carcinoma, C0007131

Source: ALL

Results per page 25

Filter within current results: KRAS

Filtramos por "KRAS"

Variant	Gene	N. diseases	DSI	DPI	Chr	Position	Consequence	Alleles	Class	AF EXOME	AF GENOME	Score vda	El vda	N. PMIDs	First Ref.	Last Ref.
rs121913530	KRAS	63	0.583	0.640	12	25245351	missense variant	C/A;G;T	snv			0.800	1.000	19	2002	2019
rs121913529	KRAS	144	0.492	0.680	12	25245350	missense variant	C/A;G;T	snv	4.0E-06		0.790	1.000	26	2002	2019
rs17851045	KRAS	27	0.672	0.400	12	25227341	missense variant	T/A;G	snv	4.0E-06		0.720	1.000	5	2002	2018

Las dos mutaciones de *KRAS*
más frecuentes (**G12C** y **G12D**)

¿Qué características tiene el gen?

Respuesta: 

[Enlace](#)

Preguntas 1 y 2

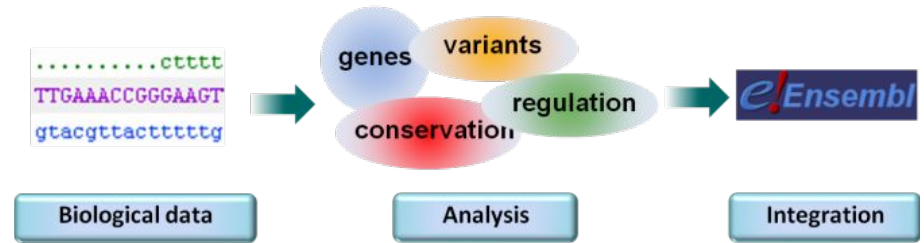


Ensembl

Es un **navegador genómico** que actúa como punto de acceso único a datos procedentes de múltiples bases de datos y proyectos bioinformáticos.

Contiene información de genes y transcritos a nivel de genoma, gen y proteína:

- Secuencias (ADN, ARN y proteína)
- Isoformas
- Variación genética
- Dominios proteicos
- Homología entre especies
- Elementos reguladores



Otro navegador genómico interesante es [UCSC Genome Browser](#).

Buscar en Ensembl

Tools

[All tools](#)

BioMart >

Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT >

Search our genomes for your DNA or protein sequence

Variant Effect Predictor >

Analyse your own variants and predict the functional consequences of known and unknown variants

Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotate genes, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

Ensembl Release 105 (Dec 2021)

- Updated allele frequency data from the NCBI Allele Frequency Aggregator (ALFA) release 2
- Update to the Variant Recoder supporting MANE annotation and variant names in external databases
- Dog (*Canis lupus familiaris*) reference genome has changed from CanFam3.1 to ROS Cfam 1.0 Labrador retriever
- Support for BCF files

Gen KRAS de humano

Search

Human for

KRAS

Go

e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease

All genomes

-- Select a species --



Pig breeds

Pig reference genome and 12 additional breeds

[View full list of all species](#)

Favourite genomes



Human
GRCh38.p13

[Still using GRCh37?](#)



Mouse
GRCm39



Zebrafish
GRCz11

Se puede buscar por:

- Nombres de genes
- Identificadores (ID) únicos de gen, transcrito, variante, elemento regulador...
- Coordenadas genómicas
- Fenotipos

Buscar en Ensembl

The screenshot shows the Ensembl genome browser interface. At the top, there is a search bar with the text "Search all species...". Below the search bar, the "New Search" section is active, showing "Current selection: < all Species" and "Only searching Human". The search criteria are set to "Only searching Human" and "KRAS", resulting in "101 results match KRAS when restricted to species: Human".

The search results are listed in a table. The first result is highlighted with a red box and labeled "Gen KRAS de humano". This result is for the "KRAS (Human Gene)" with the Ensembl ID "ENSG00000133703" and HGNC Symbol "HGNC:6407". The description states: "KRAS proto-oncogene, GTPase [Source:HGNC Symbol;Acc:HGNC:6407]". Below the description, it mentions "LRG_344 (LRG display in Ensembl gene record; description: Locus Reference Genomic record for KRAS.) is an external reference matched to Gene ENSG00000133703".

To the right of the search results, there is a "Best gene match" section. It shows a genomic map of the KRAS gene on human chromosome 12, with a 100kb scale bar. The gene is labeled "KRAS" and "KRAS proto-oncogene, GTPase". The HGNC Symbol and Acc:HGNC:6407 are also displayed.

On the left side of the search results, there are several filters and options:

- Restrict category to:** Gene (24), Transcript (48), Variant (1), Phenotype (1), GeneTree (1), GenomicAlignment (25), Clones & Regions (1).
- Per page:** 10, 25, 50, 100.
- Layout:** Standard, Table.
- Tip:** You can use wildcards in your searches. RH0* would match RHO, RHOC (RHO + zero or more characters); RH0? would match RHOC, RHOB (RHO + one character).

Gen KRAS de humano



Características del gen

Ensembl ID

Alias

Localización (cromosoma, inicio, fin, hebra)

Número de isoformas

Enlace a UniProtKB: función de la proteína

Human (GRCh38.p13)

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

Search all species...

Location: 12:25,205,246-25,250,936 Gene: KRAS

Gene-based displays

- Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
- Sequence
 - Secondary Structure
- Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree
 - Orthologues
 - Paralogues
 - Ensembl protein families
- Ontologies
 - GO: Cellular component
 - GO: Molecular function
 - GO: Biological process
- Phenotypes
- Genetic Variation
 - Variant table
 - Variant image
 - Structural variants
- Gene expression
- Pathway
- Regulation
- External references
- Supporting evidence
- ID History
 - Gene history

Gene: KRAS ENSG [ENSG00000133703]

Description KRAS proto-oncogene, GTPase [source:HGNC Symbol;Acc:HGNC:6407]

Gene Synonyms K-Ras4B, KRAS1, KRAS2

Location Chromosome [12] GRCh38:CM000674.2

About this gene This gene has 14 transcripts [splice variants], 201 orthologues, 35 paralogues and is associated with 310 phenotypes

Transcripts Show transcript table

Summary

Name KRAS (HGNC Symbol)

CCDS This gene is a member of the Human CCDS set: CCDS8702.1, CCDS8703.1

UniProtKB This gene has proteins that correspond to the following UniProtKB identifier: P01116

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

LRG LRG_344 provides a stable genomic reference framework for describing sequence variants for this gene

Ensembl version ENSG00000133703.14

Other assemblies This gene maps to 25,358,180-25,403,870 in GRCh37 coordinates. View this locus in the GRCh37 archive: ENSG00000133703

Gene type Protein coding

Annotation method Annotation for this gene includes both automatic annotation from Ensembl and Havana manual curation, see article.

Función del gen (UniProtKB)

UniProtKB es una base de datos de secuencias de proteínas y sus anotaciones.

The screenshot shows the UniProtKB entry for P01116 (RASK_HUMAN). The page is divided into several sections:

- Function:** Ras proteins bind GDP/GTP and possess intrinsic GTPase activity (PubMed:20949621). Plays an important role in the regulation of cell proliferation (PubMed:23698361, PubMed:22711838). Plays a role in promoting oncogenic events by inducing transcriptional silencing of tumor suppressor genes (TSGs) in colorectal cancer (CRC) cells in a ZNF304-dependent manner (PubMed:24623306). 4 Publications, Curated.
- Names & Taxonomy:** Proteinⁱ | GTPase KRas; Geneⁱ | KRAS; Statusⁱ | UniProtKB reviewed (Swiss-Prot); Organismⁱ | Homo sapiens (Human).
- Amino acids:** 189
- Protein existenceⁱ:** Evidence at protein level
- Annotation scoreⁱ:** 5/5

A callout box labeled "Función de la proteína" points to the Function section.

Genes vecinos

Volvemos a Ensembl:

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

Human (GRCh38.p13) **Pestaña de región cromosómica**

Location: 12:25,205,246-25,250,936 Gene: KRAS

Gene-based displays

- Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
- Sequence
 - Secondary Structure
- Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree
 - Orthologues
 - Paralogues
 - Ensembl protein families
- Ontologies
 - GO: Cellular component
 - GO: Molecular function
 - GO: Biological process
- Phenotypes
- Genetic Variation
 - Variant table
 - Variant image
 - Structural variants
- Gene expression
 - Pathway
 - Regulation
- External references
- Supporting evidence
- ID History
 - Gene history

Gene: KRAS ENSG[redacted]

Description KRAS proto-oncogene, GTPase [Source:HGNC Symbol;Acc:HGNC:6407]

Gene Synonyms K-Ras4B, KRAS1, KRAS2

Location Chromosome [redacted]
GRCh38:CM000674.2

About this gene This gene has 14 transcripts ([splice variants](#)), [201 orthologues](#), [35 paralogues](#) and is associated with [310 phenotypes](#).

Transcripts [Show transcript table](#)

Summary

Name [KRAS](#) (HGNC Symbol)

CCDS This gene is a member of the Human CCDS set: [CCDS8702.1](#), [CCDS8703.1](#)

UniProtKB This gene has proteins that correspond to the following UniProtKB identifiers: [P01116](#)

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

LRG [LRG_344](#) provides a stable genomic reference framework for describing sequence variants for this gene

Ensembl version ENSG00000133703.14

Other assemblies This gene maps to [25,358,180-25,403,870](#) in GRCh37 coordinates.
View this locus in the GRCh37 archive: [ENSG00000133703](#)

Gene type Protein coding

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

[Configure this page](#)



Genes vecinos

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

Human (GRCh38.p13) ▼

Location: 12:25,205,246-25,250,936 Gene: KRAS

Chromosome 12: 25,205,246-25,250,936

Region in detail

Gen KRAS

Genes cercanos a KRAS en el genoma

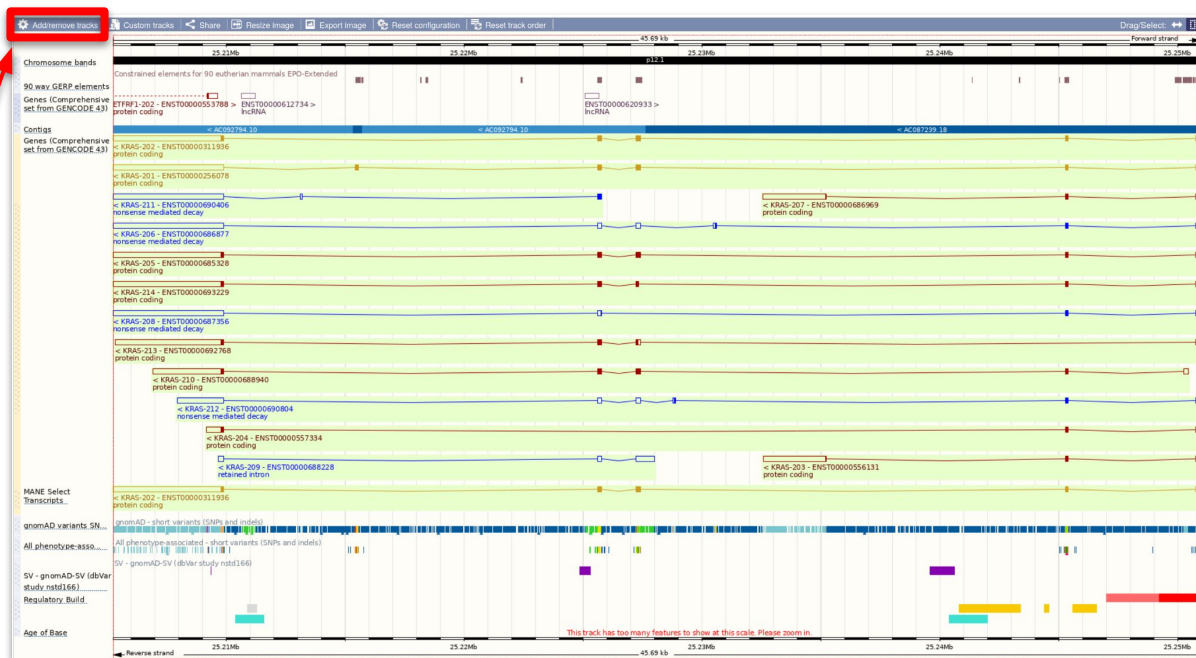
Leyenda de colores de los genes

- merged Ensembl/Havana ← Genes codificantes de proteína
- processed transcript
- pseudogene
- RNA gene
- CTCF
- Enhancer
- Open Chromatin
- Promoter
- Promoter Flank
- Transcription Factor Binding Site

Pistas de Ensembl

Hacemos scroll en la pantalla anterior hasta llegar a este otro visualizador genómico:

- Información procedente de distintas bases de datos
- Podemos ver isoformas, variantes (SNPs e indels), regiones de regulación...
- Son **customizables**: permiten mostrar sólo la información que nos interesa



Variantes patogénicas

Queremos visualizar sólo las isoformas y las variantes:

- **Paso 1:** ocultar las pistas que no nos interesan.

The screenshot shows the Ensembl genome browser's track configuration panel. The left sidebar lists various tracks, and the main area shows a list of tracks with checkboxes to toggle them on or off. The tracks are categorized into several groups:

- Active tracks:** A dropdown menu set to 'Current unsaved' with a 'Save current configuration' button.
- Sequence and assembly:** Includes Contigs, Sequence, and Primary assembly mapping.
- Genes and transcripts:** Includes Comprehensive Gene Annotations from GENCODE 43 and MANE Select Transcripts.
- Variation:** Includes gnomAD - short variants (SNPs and indels), All phenotype-associated - short variants (SNPs and Indels), and gnomAD-SV (dbVar study nstd166) (structural variants).
- Regulation:** Includes Regulatory Build.
- Comparative genomics:** Includes Constrained elements for 90 eutherian mammals EPO-Extended and Age of Base.
- Information and decorations:** Includes Chromosome bands and Assembly exceptions.

Annotations on the screenshot:

- A red text box on the left says "Apagamos estas 6 pistas" (We turn off these 6 tracks).
- Yellow callout boxes point to the checkboxes for "Pista de isoformas" (MANE Select Transcripts) and "Pista de variantes" (gnomAD - short variants).
- A yellow callout box on the right says "Información sobre la pista" (Information about the track), pointing to the information icon (i) next to a track.

Apagamos
estas 6
pistas

Pista de isoformas

Pista de variantes

Información
sobre la pista

Variantes patogénicas

Queremos visualizar sólo las isoformas y las variantes:

- **Paso 2:** mostrar la pista de variantes patogénicas (estaba oculta).

Configure Region Image | Configure Overview Image | Configure Chromosome Image | Personal Data

Find a track

Active tracks

Favourite tracks

Track order

Search results

Genome Reference Consortium Issues (0/17)

Sequence and assembly (3/18)

- Sequence (2/4)
- Markers (0/1)
- GRC alignments (1/2)
- Simple features (0/4)
- Clones & misc. regions (0/7)

Genes and transcripts (2/79)

- Genes (2/7)
- Long reads (0/14)
- Prediction transcripts (0/1)
- LRG (0/1)
- RNASeq models (0/56)

mRNA and protein alignments (0/7)

- mRNA alignments (0/3)
- EST alignments (0/1)
- Protein alignments (0/3)

Variation (2/86)

- Phenotype, disease and curated variants (1/20)
- Structural variants (0/1)
- Failed variants (0/1)
- Structural variants (0/28)

Somatic mutations (0/5)

- Somatic variants (0/2)
- Somatic structural variants (0/3)

Regulation (0/467)

- Regulatory Build (0/1)
- Features by Cell/Tissue (0/415)
- DNA methylation (0/47)

Select from available configurations: Current unsaved Save current configuration

Variation

Enable/disable all Phenotype/disease variants by source

- All ClinVar variant annotations - short variants (SNPs and indels)
- ClinVar variants described as being probable-pathogenic, pathogenic, drug-response or histocompatibility - short variants (SNPs and indels)
- All phenotype-associated - short variants (SNPs and indels)
- HGMD-PUBLIC - short variants (SNPs and indels)
- NHGRI-EBI GWAS catalog phenotype - short variants (SNPs and indels)
- OMIM phenotype - short variants (SNPs and indels)
- PhenCode - short variants (SNPs and indels)

Enable/disable all LSDB-associated variants

- All LSDB-associated - short variants (SNPs and indels)
- HbVar - short variants (SNPs and indels)
- Infervers - short variants (SNPs and indels)
- KAT6BDB - short variants (SNPs and indels)
- LMDD - short variants (SNPs and indels)
- LSDB - short variants (SNPs and indels)
- OIVD - short variants (SNPs and indels)
- PAHdb - short variants (SNPs and indels)
- dbPEX - short variants (SNPs and indels)

Enable/disable all Phenotype annotations

- Phenotype annotations (all types)
- Phenotype annotations (Variations)
- Phenotype annotations (StructuralVariations)
- Phenotype annotations (Genes)

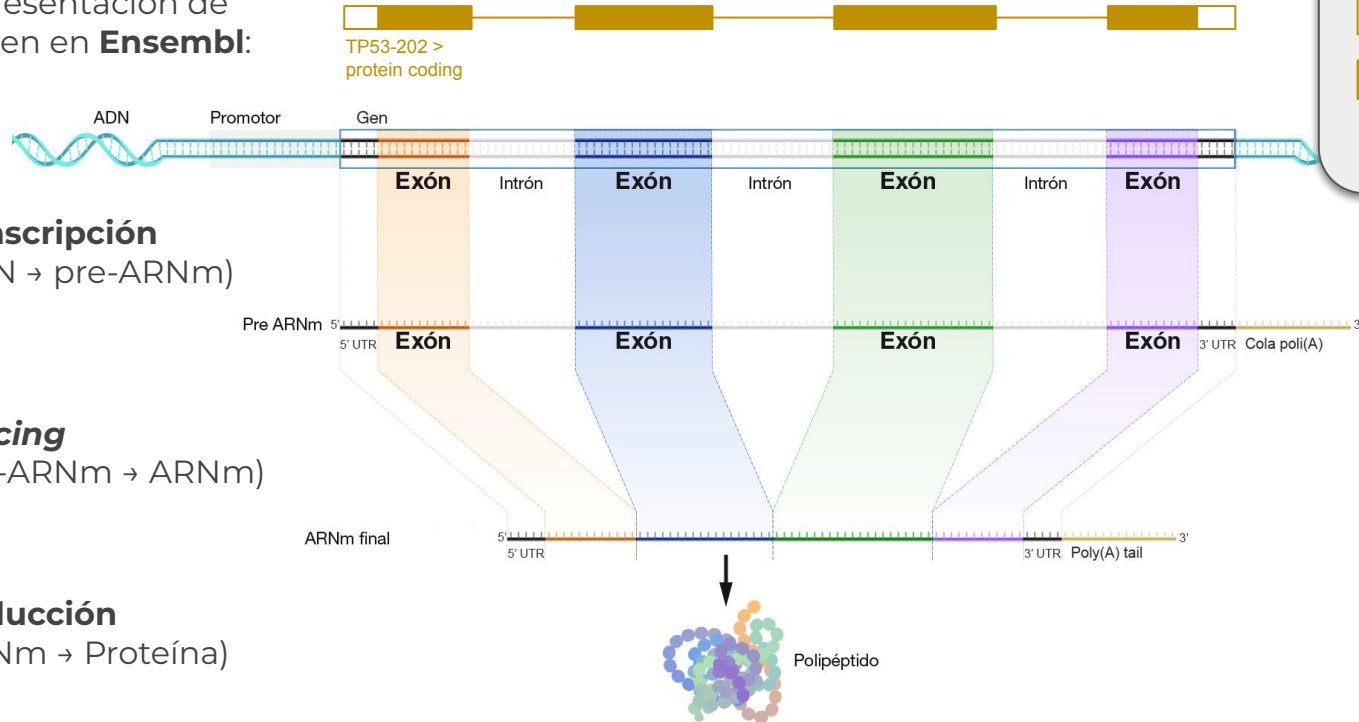
Pista de variantes patogénicas

Activamos la pista de variantes patogénicas

(en modo Normal)

Recordatorio: estructura de los genes

Representación de un gen en **Ensembl**:



Legenda:

- UTRs
- Exones codificantes
- Intrones

Transcripción

(ADN → pre-ARNm)

Splicing

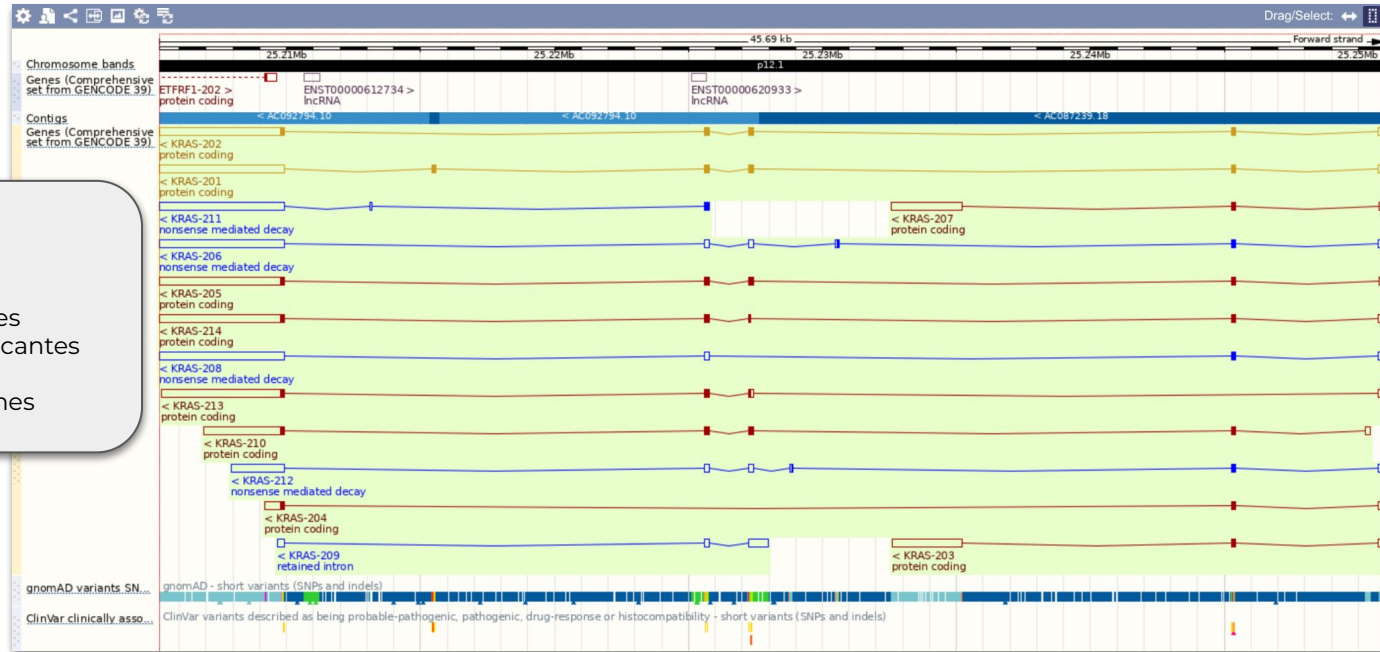
(pre-ARNm → ARNm)

Traducción

(ARNm → Proteína)

Variantes patogénicas

Hay variantes en todo el gen, pero las **patogénicas** están principalmente en **exones**.



Leyenda:

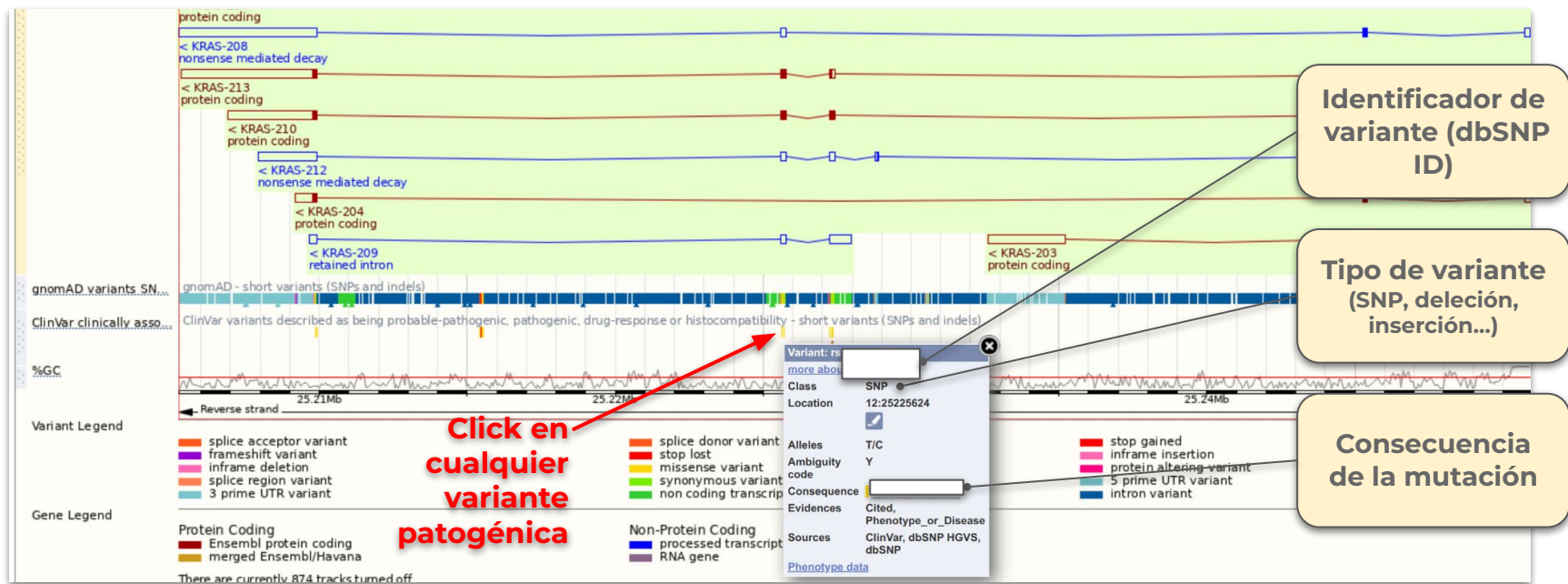
- UTRs
- Exones codificantes
- Intrones

Pista de isoformas

Variantes

Variantes patogénicas

Variantes patológicas



Isoformas

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

Human (GRCh38.p13) **Gene: KRAS** **Volvemos a la pestaña del gen**

Location: 12:25,205,246-25,250,936

Gene-based displays: Summary

Gene: KRAS ENSG00000133703

KRAS proto-oncogene, GTPase [Source:HGNC Symbol;Acc:HGNC:6407]

K-Ras4B, KRAS1, KRAS2

Chromosome 12: 25,205,246-25,250,936 reverse strand.

GRCh38:CM000674.2

This gene has 14 transcripts (splice variants), 201 orthologues, 35 paralogues and is associated with 310 phenotypes.

Hide transcript table **Tabla de isoformas**

Show/hide columns (1 hidden)

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000256078.10	KRAS-202	5430	189aa	Protein coding	CCDS8702	P01116-2	NM_004985.5	MANE Select v0.95 Ensembl Canonical GENCODE basic APPRIS ALT1 TSL:1
ENST00000256078.10	KRAS-201	5430	189aa	Protein coding	CCDS8703	P01116-1	-	GENCODE basic APPRIS P4 TSL:1
ENST00000685328.1	KRAS-205	5287	188aa	Protein coding	CCDS8702	A0A024RAV5	-	GENCODE basic APPRIS ALT1
ENST00000688940.1	KRAS-210	3630	188aa	Protein coding	CCDS8702	A0A024RAV5	-	GENCODE basic APPRIS ALT1
ENST00000693229.1	KRAS-214	5187	163aa	Protein coding	-	-	-	GENCODE basic
ENST00000692768.1	KRAS-213	5075	122aa	Protein coding	-	-	-	GENCODE basic
ENST00000686969.1	KRAS-207	2957	43aa	Protein coding	-	-	-	GENCODE basic
ENST00000556131.2	KRAS-203	2937	43aa	Protein coding	-	G3V4K2	-	GENCODE basic TSL:1
ENST00000557334.6	KRAS-204	1042	75aa	Protein coding	-	G3V5I7	-	GENCODE basic TSL:5
ENST00000686877.1	KRAS-206	5420	55aa	Nonsense mediated decay	-	-	-	-
ENST00000687356.1	KRAS-208	5101	40aa	Nonsense mediated decay	-	-	-	-
ENST00000690406.1	KRAS-211	4919	54aa	Nonsense mediated decay	-	-	-	CDS 5' incomplete
ENST00000690804.1	KRAS-212	2721	51aa	Nonsense mediated decay	-	-	-	-
ENST00000688228.1	KRAS-209	1163	No protein	Retained intron	-	-	-	-

Longitud del ARNm (pb)

Ensembl ID del transcrito

Longitud de la proteína (aa)

Tipo de transcrito

Elegimos una isoforma

* Preferiblemente la isoforma con la etiqueta "MANE Select" (columna *Flags*), ya que ha sido escogida como isoforma de referencia del gen.

Información de la isoforma

Ensembl ID del transcrito

Número de exones

Ocultamos la tabla de isoformas

Legenda:

- UTRs
- Exones codificantes
- Intrones

“General identifiers”: enlaces de este transcrito en otras bases de datos.

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

Human (GRCh38.p13) ▾

Location: 12:25,205,246-25,250,936 Gene: KRAS Transcript: KRAS-202

Transcript-based displays

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

Transcript: ENST KRAS-202

Description KRAS proto-oncogene, GTPase [Source:HGNC Symbol;Acc:HGNC:6407]

Gene Synonyms K-Ras4B, KRAS1, KRAS2

Location [Chromosome 12: 25,205,246-25,250,929](#) reverse strand.

About this transcript This transcript has exons, is annotated with [22 domains and features](#), is associated with [15274 variant alleles](#) and maps to [592 oligo probes](#).

Gene This transcript is a product of gene [ENSG00000133703.14](#) [Show transcript table](#)

Summary

3'UTR **5'UTR**

<- KRAS-202 protein coding
Reverse strand 45.68 kb

Statistics Exons: Coding exons: 4, Transcript length: bps, Translation length: residues

CCDS This transcript is a member of the Human CCDS set: [CCDS8702](#)

Uniprot This transcript corresponds to the following Uniprot identifiers: [P01116](#)

Transcript Support Level (TSL) TSL:1

Version

Type Protein coding

Annotation Method Transcript where the Ensembl genebuild transcript and the Havana manual annotation have the same sequence, for every

Gencode basic gene This transcript is a member of the [Gencode basic](#) gene set.

¿Existe algún microARN que regule al gen?

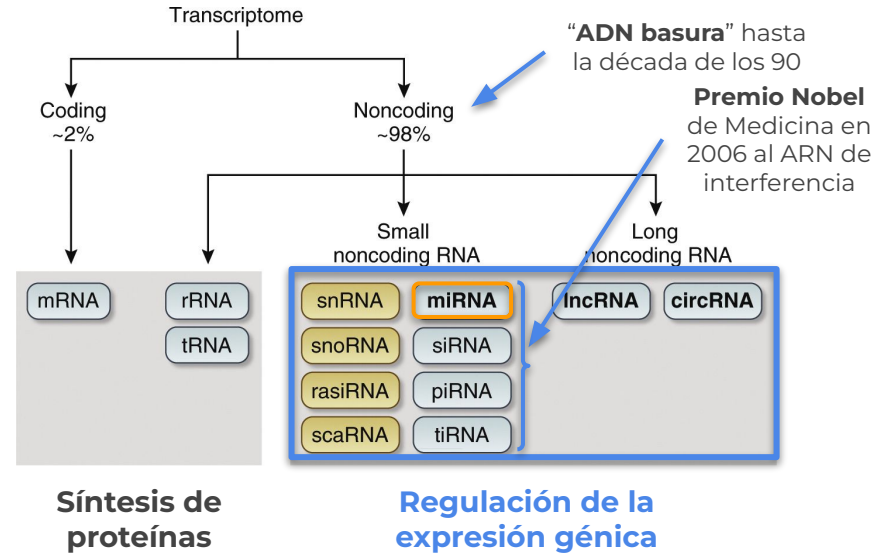
Respuesta:  mRWalk

[Enlace](#)

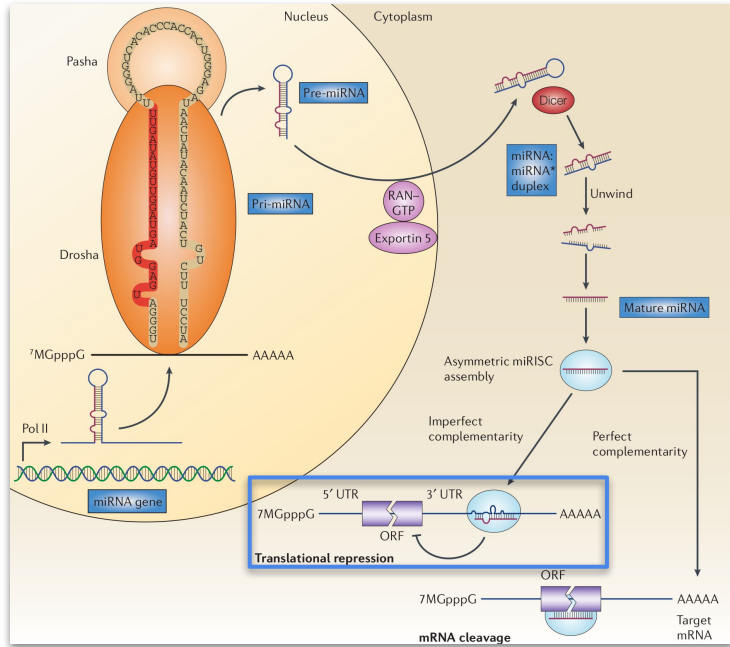
Pregunta 3

ARNs no codificantes: miRNA y lncRNA

- No traducidos a proteína
- Proceden de “genes de ARN” (color **morado** en Ensembl)
- **En síntesis de proteínas:** ARN ribosómico y de transferencia
- **En regulación de la expresión:** miRNA, lncRNA, entre otros.



microRNAs

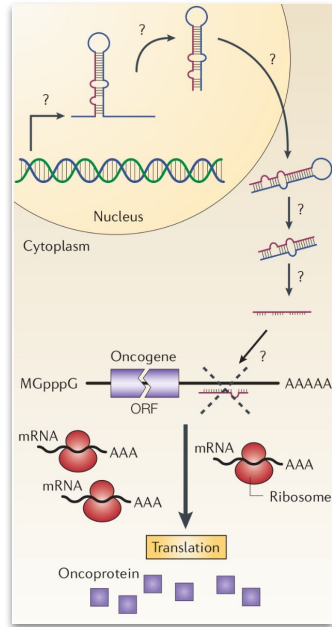


Biogénesis y actuación de miRNA

- ARN pequeños (18-20 nt)
- Regulan **negativamente** la expresión de genes
- Dos mecanismos de acción:
 - Inhiben la traducción del ARNm uniéndose a su **extremo 3'** (hibridación **imperfecta**, más frecuente)
 - Provocan la escisión del ARNm uniéndose a regiones **codificantes** (hibridación **perfecta**)
- Un miRNA puede controlar **varios genes** y un gen puede estar controlado por **varios miRNA**

microRNAs y cáncer: Oncomirs

miRNA como **supresor tumoral**
(silenciando un oncogeno)

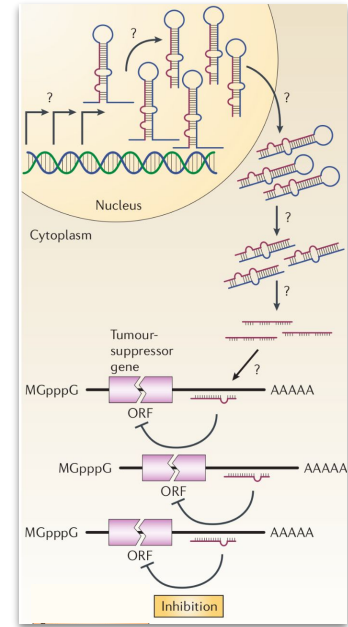


Si se infraexpresa

Si se sobreexpresa

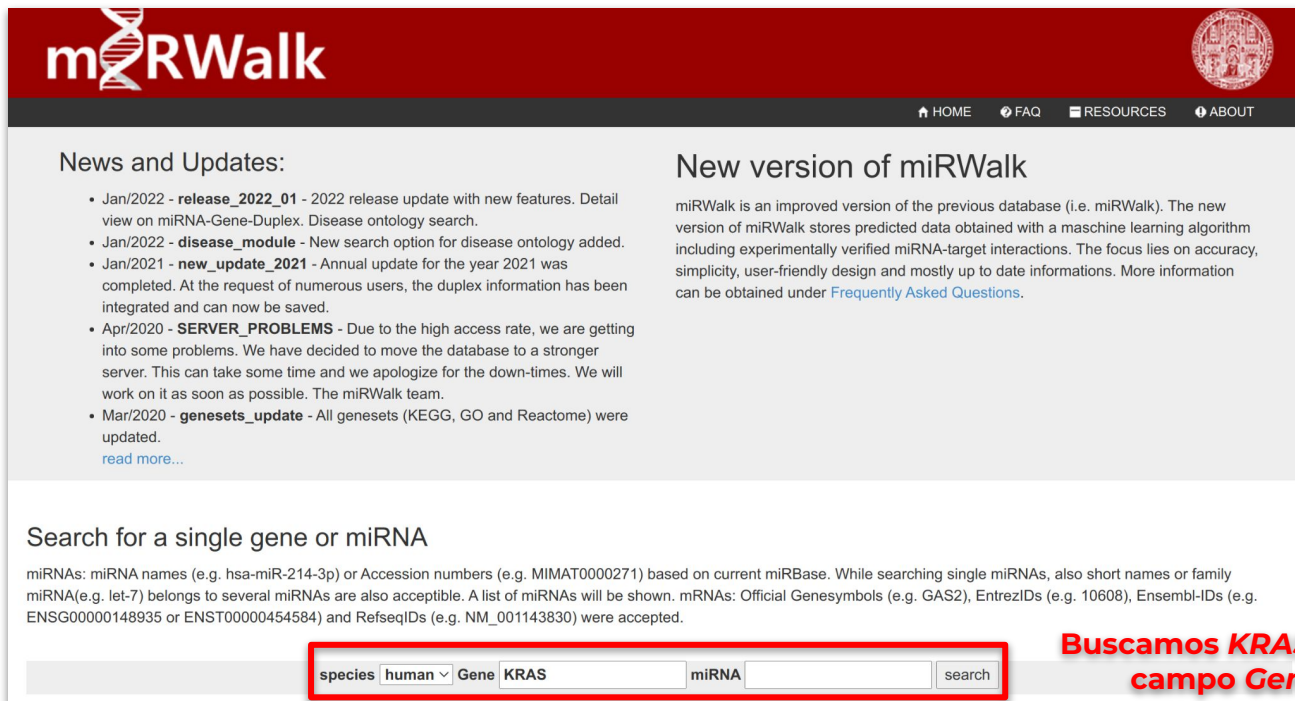
Formación de un tumor
↑ Proliferación
↑ Invasión
↑ Angiogénesis
↓ Muerte celular

miRNA como **oncogeno**
(silenciando un supresor tumoral)



miRWalk

Base de datos sobre relaciones entre miRNAs y genes.



miRWalk

HOME FAQ RESOURCES ABOUT

News and Updates:

- Jan/2022 - **release_2022_01** - 2022 release update with new features. Detail view on miRNA-Gene-Duplex. Disease ontology search.
- Jan/2022 - **disease_module** - New search option for disease ontology added.
- Jan/2021 - **new_update_2021** - Annual update for the year 2021 was completed. At the request of numerous users, the duplex information has been integrated and can now be saved.
- Apr/2020 - **SERVER_PROBLEMS** - Due to the high access rate, we are getting into some problems. We have decided to move the database to a stronger server. This can take some time and we apologize for the down-times. We will work on it as soon as possible. The miRWalk team.
- Marr/2020 - **genesets_update** - All genesets (KEGG, GO and Reactome) were updated.
[read more...](#)

New version of miRWalk

miRWalk is an improved version of the previous database (i.e. miRWalk). The new version of miRWalk stores predicted data obtained with a machine learning algorithm including experimentally verified miRNA-target interactions. The focus lies on accuracy, simplicity, user-friendly design and mostly up to date informations. More information can be obtained under [Frequently Asked Questions](#).

Search for a single gene or miRNA

miRNAs: miRNA names (e.g. hsa-miR-214-3p) or Accession numbers (e.g. MIMAT0000271) based on current miRBase. While searching single miRNAs, also short names or family miRNA(e.g. let-7) belongs to several miRNAs are also acceptable. A list of miRNAs will be shown. mRNAs: Official Genesymbols (e.g. GAS2), EntrezIDs (e.g. 10608), Ensembl-IDs (e.g. ENSG00000148935 or ENST00000454584) and RefseqIDs (e.g. NM_001143830) were accepted.

species **human** Gene **KRAS** miRNA search

Buscamos KRAS en el campo Gene



miRNA asociados al gen

Identificador del miRNA

Compuesto de 4 partes:

- Especie: hsa es *Homo sapiens*
- Procesamiento: miR es miRNA maduro (mir sería el gen o pre-miRNA)
- Identificador numérico
- Brazo: 5' o 3' (sólo para miRNA maduros)
- **let-7 es una excepción, se descubrió antes de que existiera esta nomenclatura**

KRAS

Details

Entrezid 3845

Genesymbol KRAS

Alias 'C-K-RAS;C-K-RAS;c-Ki-ras;c-Ki-ras2;

Description Homo sapiens KRAS proto-oncogene, GTPase (KRAS) transcript variant, mRNA.

Chromosome 12

Ensemblid ENSG00000133703

Transcripts

Refseq	Ensemblid	Length	Startcds	Endcds
191			191	757
191			191	760

Interactions

Seleccionamos un miRNA

Mirna	Accession	Gene	Details	Score	UTR	Binding Site	Au	Me	N Pairings	Targets	Mirdb	Mirtarbase
hsa-let-7e-5p	NM_033360	KRAS	details	0.92	3UTR	3220,3263	0.41	-3.938	17	—	—	—
hsa-let-7f-1-3p	NM_033360	KRAS	details	0.92	CDS	413,434	0.63	-14.41	15	—	Link	—
hsa-let-7f-1-3p	NM_033360	KRAS	details	0.85	3UTR	1512,1524	0.73	-19.424	11	—	Link	—
hsa-let-7f-2-3p	NM_033360	KRAS	details	0.92	3UTR	1911,1937	0.53	-4.957	19	—	—	—
hsa-let-7f-2-3p	NM_033360	KRAS	details	0.85	CDS	714,734	0.72	-12.111	15	—	—	—
hsa-miR-15a-5p	NM_033360	KRAS	details	0.85	3UTR	5336,5347	0.77	-8.071	10	LINK	—	—
hsa-miR-15a-3p	NM_033360	KRAS	details	1.00	3UTR	3508,3548	0.57	-3.938	20	—	—	—
hsa-miR-15a-3p	NM_033360	KRAS	details	0.92	5UTR	110,130	0.25	-10.094	17	—	—	—
hsa-miR-16-5p	NM_033360	KRAS	details	1.00	3UTR	2768,2783	0.59	-5.227	14	LINK	—	MIRT053090
hsa-miR-17-3p	NM_033360	KRAS	details	1.00	3UTR	3900,3919	0.66	-8.813	15	—	—	—

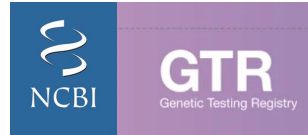
« 1 2 3 Id a la 2ª página 11 12 ... »

Región del gen donde se une el miRNA

Seleccionamos un miRNA

¿Qué laboratorios realizan el diagnóstico genético en España?

Respuesta:



[Enlace](#)

Pregunta 4

Buscar en NCBI GTR

The screenshot shows the NCBI GTR (Genetic Testing Registry) homepage. At the top, there is a navigation bar with 'NCBI Resources' and 'How To' menus, and user information 'marisolbc@go.ugr.es My NCBI Sign Out'. Below this is a COVID-19 information banner. The main content area features a search bar with 'Lung cancer' entered and a 'Search All GTR' button. A dropdown menu shows search results for 'Lung cancer', with the first result 'Lung cancer' highlighted in red. To the right of the search bar, there is a red text overlay: 'Buscamos el nombre de la enfermedad'. Below the search bar, there are tabs for 'All GTR', 'Human Tests', 'Microbe Tests', 'Conditions/Phenotypes', 'Genes', 'Labs', and 'GeneReviews'. A 'YouTube GTR Tutorials' link is also visible.

NCBI Resources How To marisolbc@go.ugr.es My NCBI Sign Out

COVID-19 Information
Public health information (CDC) | Research information (NIH) | SARS-CoV-2 data (NCBI) | Prevention and treatment information (HHS) | Español

GTR: GENETIC TESTING REGISTRY

All GTR Human Tests Microbe Tests Conditions/Phenotypes Genes Labs GeneReviews Advanced search for tests

Lung cancer Search All GTR

EGFR: EGFR-related lung cancer Condition
GSTM1-related lung cancer Condition
LNCR2: Lung cancer susceptibility 2 Condition
LNCR4: Lung cancer susceptibility 4 Condition
Lung cancer Condition
Lung cancer susceptibility 1 Condition

Buscamos el nombre de la enfermedad

YouTube GTR Tutorials

ed to the GTR; it relies on submitters to provide information that is accurate and not
e GTR. GTR is not a substitute for medical advice. **Patients and consumers** with specific
netics professional.

Buscar en NCBI GTR

Haced scroll al apartado de “Genes” hasta que encontréis el gen *KRAS*:

[FASLG](#) 59 tests

Also known as: ALPS1B, APT1LG1, APTL, CD178, CD95-L, CD95L, FASL, TNFSF6, TNLG1A, FASLG

Summary: Fas ligand

[IRF1](#) 10 tests

Also known as: IMD117, IRF-1, MAR, IRF1

Summary: interferon regulatory factor 1

[KRAS](#) 465 tests

Vemos qué pruebas genéticas hay para *KRAS*

Also known as: C-K-RAS, C-K-RAS, CFC2, K-RAS2A, K-RAS2B, K-RAS4A, K-RAS4B, K-Ras, K-Ras 2, KI-RAS, KRAS1, KRAS2, NS, NS3, OES, RALD, RASK2, c-Ki-ras, c-Ki-ras2, KRAS

Summary: KRAS proto-oncogene, GTPase

[MAP3K8](#) 11 tests

Also known as: AURA2, COT, EST, ESTF, MEKK8, TPL2, Tpl-2, c-COT, MAP3K8

Summary: mitogen-activated protein kinase kinase kinase 8

[PIK3CA](#) 240 tests

Also known as: CCM4, CLAPO, CLOVE, CWS5, MCAP, MCM, MCMTc, PI3K, PI3K-alpha, p110-alpha, PIK3CA

Summary: phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha



Tests genéticos de *KRAS* en España

Aplicamos 4 filtros

▼ Test type [reset](#)

Clinical (40)

▼ Test purpose [reset](#)

Diagnosis (40)

Pre-symptomatic (20)

Predictive (12)

Prognostic (12)

Risk Assessment (11)

Screening (11)

▼ Test method [reset](#)

Molecular Genetics

Deletion/duplication analysis (3)

Mutation scanning of the entire coding region (1)

Sequencing of the entire coding region (2)

Sequence analysis of the entire coding region (40)

▼ Test service

► Laboratory certification

► Specimen type

▼ Laboratory location [reset](#)

United States

Alabama (2)

California (33)

Colorado (1)

Connecticut (1)

Delaware (1)

District of Columbia (5)

Georgia (1)

[See more states](#)

Other countries

Austria (2)

Spain (40)

Tests names and labs	Conditions	Genes, analytes, and microbes	Methods
RASopathies Panel	1	18	C Sequence analysis of the entire coding region
Health in Code Spain			
Cardiomyopathies General Panel	1	173	C Sequence analysis of the entire coding region
Health in Code Spain			
Hypertrophic cardiomyopathy extended panel	1	104	
Health in Code Spain			
Congenital heart diseases Panel	1	76	C Sequence analysis of the entire coding region
Health in Code Spain			
Arrhythmia General Panel	1	218	C Sequence analysis of the entire coding region
Health in Code Spain			
Cardiovascular Diseases_General Panel	1	380	C Sequence analysis of the entire coding region
Health in Code Spain			
Hypertrophic Cardiomyopathy Extended Panel	1	90	C Sequence analysis of the entire coding region
Health in Code Spain			
Congenital Heart Diseases Panel	1	39	C Sequence analysis of the entire coding region
Health in Code Spain			
RASopathies NGS Panel	26	12	C Sequence analysis of the entire coding region
Health in Code Spain			
Cardiomyopathies Panel	4	140	C Sequence analysis of the entire coding region

Genes analizados en la prueba

Nombre del laboratorio

Las pruebas que analizan varios genes se conocen como **paneles de secuenciación**.

Si sólo estamos interesados en secuenciar el gen *KRAS*, buscaríamos una prueba que analice sólo este gen.



Preparación para la Práctica 2: Amplificación de fragmentos de ADN por PCR

1. Elección de líneas celulares con *KRAS* mutado.
2. Diseño de *primers* para amplificar el exón 2 de *KRAS*.



¿Cómo encuentro una línea celular de cáncer de pulmón con *KRAS* mutado?

Respuesta:  **COSMIC** Cell lines
Catalogue Of Somatic Mutations In Cancer

[Enlace](#)

Pregunta 5

¿Qué es una línea celular?

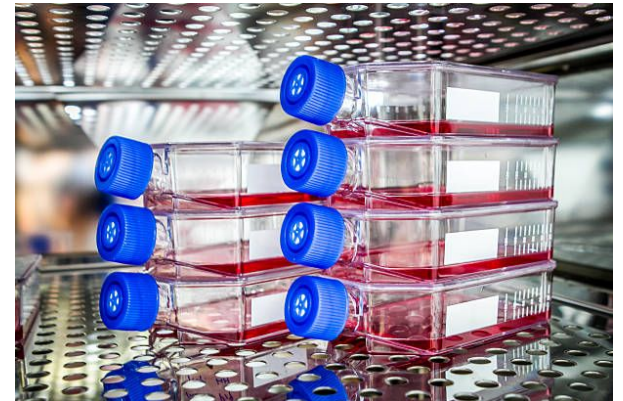
Es un cultivo de células establecido que **proliferará indefinidamente** si las condiciones del **medio** de cultivo y **espacio** son adecuadas.

- Se usan en **investigación** como paso previo a los ensayos en organismos modelo (ratones, ratas, etc)
- Hay líneas celulares de distintos tipos de cáncer y tejidos sanos.
- Una línea celular humana proviene de un donante:

La línea celular **HeLa** de cáncer de cuello de útero proviene de una muestra del tumor de Henrietta Lacks, tomada en 1951.

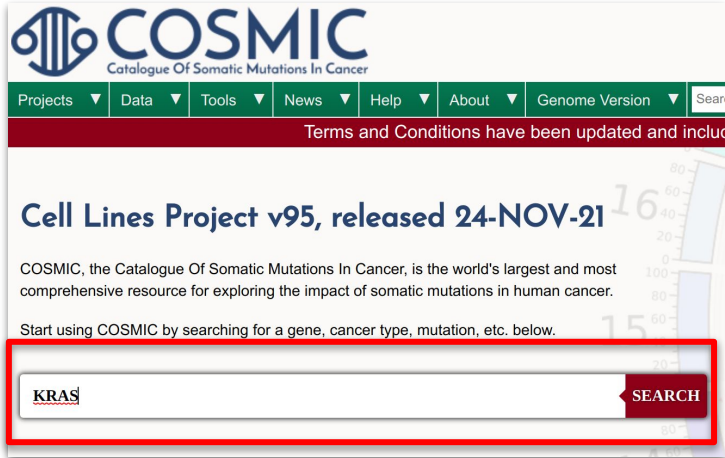
En la **Práctica 2** usaréis estas líneas de cáncer de pulmón:

- *KRAS* mutado G12C: **NCI-H358**
- *KRAS* mutado G12D: **A427**
- *KRAS* no mutado: **NCI-H838**



Frascos de cultivos celulares en un incubador.

Buscar en COSMIC Cell Lines



COSMIC
Catalogue Of Somatic Mutations In Cancer

Projects ▾ Data ▾ Tools ▾ News ▾ Help ▾ About ▾ Genome Version ▾ Search

Terms and Conditions have been updated and included

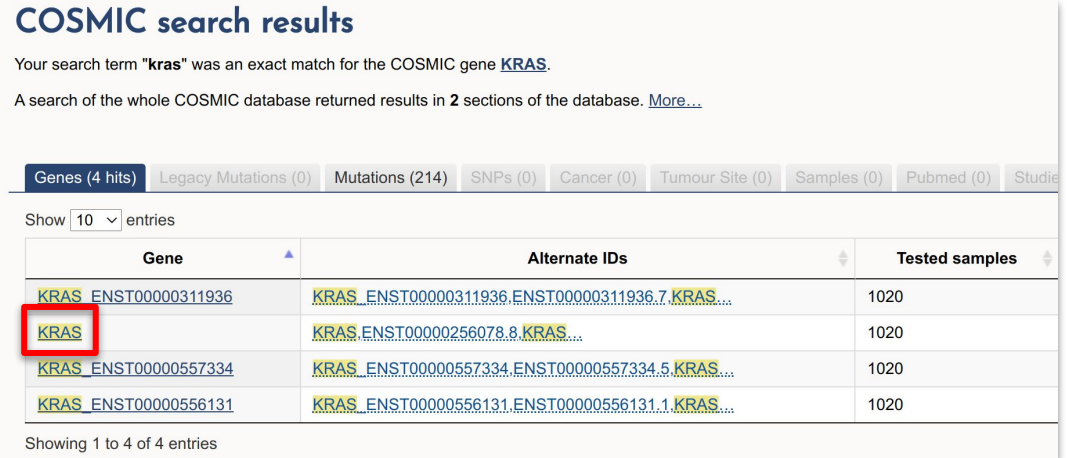
Cell Lines Project v95, released 24-NOV-21

COSMIC, the Catalogue Of Somatic Mutations In Cancer, is the world's largest and most comprehensive resource for exploring the impact of somatic mutations in human cancer.

Start using COSMIC by searching for a gene, cancer type, mutation, etc. below.

KRAS **SEARCH**

Escribimos “KRAS” en el buscador



COSMIC search results

Your search term "kras" was an exact match for the COSMIC gene [KRAS](#).

A search of the whole COSMIC database returned results in 2 sections of the database. [More...](#)

Genes (4 hits) Legacy Mutations (0) Mutations (214) SNPs (0) Cancer (0) Tumour Site (0) Samples (0) Pubmed (0) Studies (0)

Show 10 entries

Gene	Alternate IDs	Tested samples
KRAS_ENST00000311936	KRAS_ENST00000311936 ENST00000311936.7 KRAS...	1020
KRAS	KRAS_ENST00000256078.8 KRAS...	1020
KRAS_ENST00000557334	KRAS_ENST00000557334 ENST00000557334.5 KRAS...	1020
KRAS_ENST00000556131	KRAS_ENST00000556131 ENST00000556131.1 KRAS...	1020

Showing 1 to 4 of 4 entries

Hacemos click en el segundo resultado

Buscar en COSMIC Cell Lines

Vamos a buscar líneas celulares con la mutación **G12C** en *KRAS*:

Click en
"Variants"

Gene
KRAS

- Gene view
- Overview
- External links
- Drug resistance
- Tissue distribution
- Genome browser
- Mutation distribution
- Variants**
- References

Reset page

Search

Search COSMIC...

Filters

Show advanced filters

Range Show input fields

1 190

Variants

Mutations Fusions CNV & Expression Methylation

This tab displays a table of mutations for the selected gene. You can see more information in our [help pages](#).

Show 10 entries

Export: CSV TSV Search:

Position (AA)	Mutation (CDS)	Mutation (Amino Acid)	Legacy Mutation ID	Count	Mutation Type
12	c.34G>A	p.G12S	COSM517	2	Substitution - Missense
12	c.34G>C	p.G12R	COSM518	4	Substitution - Missense
12	c.34G>T	p.G12C	COSM516	20	Substitution - Missense
12	c.35G>A	p.G12D	COSM521	34	Substitution - Missense
12	c.35G>C	p.G12A	COSM522	8	Substitution - Missense
12	c.35G>T	p.G12V	COSM520	30	Substitution - Missense
13	c.37G>T	p.G13C	COSM527	3	Substitution - Missense
13	c.38G>A	p.G13D	COSM532	9	Substitution - Missense
14	c.40G>A	p.V14I	COSM12722	2	Substitution - Missense
18	c.53C>A	p.A18D	COSM542	1	Substitution - Missense

Showing 1 to 10 of 66 entries

First Previous 1 2 3 4 5 6 7 Next Last

Líneas con la mutación G12C

Mutation
COSV55497469

- Overview
- Tissue distribution
- Samples**
- Pathways affected
- References

Reset page

Click en "Samples"

La que vais a usar en la Práctica 2

Samples

This section displays a list of samples, including their cell line name, tissue and zygosity, and the mutation included, where applicable.

Show entries

Export: [CSV](#) [TSV](#) Search:

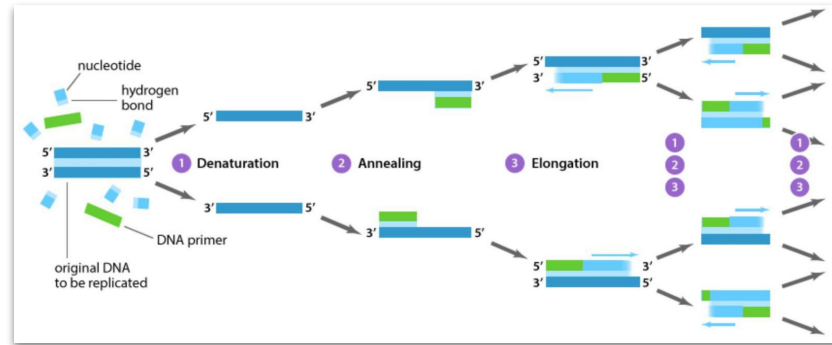
Sample name	Gene name	Transcript	Primary Tissue	Tissue Subtype 1	Primary Histology	Histology Subtype 1	Pubmed ID	Zygosity	Somatic Status	Sample Type
	KRAS	ENST00000256078.8	Lung	NS	Carcinoma	Adenocarcinoma	-	Homozygous	Unknown	Cultured
	KRAS	ENST00000256078.8	Lung	NS	Carcinoma	Adenocarcinoma	-	Homozygous	Unknown	Cultured
	KRAS	ENST00000256078.8	Lung	NS	Carcinoma	Non small cell carcinoma	-	Homozygous	Unknown	Cultured
	KRAS	ENST00000256078.8	Lung	NS	Carcinoma	Non small cell carcinoma	-	Heterozygous	Unknown	Cultured
NCI-H358	KRAS	ENST00000256078.8	Lung	NS	Carcinoma	Bronchioloalveolar adenocarcinoma	-	Heterozygous	Unknown	Cultured
	KRAS	ENST00000256078.8	Lung	NS	Carcinoma	Non small cell carcinoma	-	Homozygous	Unknown	Cultured
KYSE-410	KRAS	ENST00000256078.8	Oesophagus	NS	Carcinoma	NS	-	Heterozygous	Unknown	Cultured
OV-56	KRAS	ENST00000256078.8	Ovary	NS	Carcinoma	NS	-	Heterozygous	Unknown	Cultured
MIA-PaCa-2	KRAS	ENST00000256078.8	Pancreas	NS	Carcinoma	Ductal carcinoma	-	Homozygous	Unknown	Cultured
UM-UC-3	KRAS	ENST00000256078.8	Urinary tract	Bladder	Carcinoma	NS	-	Homozygous	Unknown	Cultured

Showing 11 to 20 of 20 entries

First Previous 1 **2** Next Last

Id a la 2ª página

¿Cómo haría una PCR de este gen?






Respuesta: **Primer-BLAST**

[Enlace](#)

Pregunta 6

Secuencia a amplificar

Leyenda:

-  UTRs
-  Exones codificantes
-  Intrones

En la sección
Sequence
click en Exons



Transcript-based displays

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

Transcript: ENST[redacted] KRAS-202

Description: KRAS proto-oncogene, GTPase [Source:HGNC Symbol;Acc:HGNC:6407]

Gene Synonyms: K-Ras4B, KRAS1, KRAS2

Location: [Chromosome 12: 25,205,246-25,250,929](#) reverse strand.

About this transcript: This transcript has [redacted] exons, is annotated with [22 domains and features](#), is associated with [15274 variant alleles](#) and maps to [592 oligo probes](#).

Gene: This transcript is a product of gene [ENSG00000133703.14](#) [Show transcript table](#)

Summary



Statistics: Exons: [redacted] Coding exons: 4, Transcript length: [redacted] bps, Translation length: [redacted] residues

CCDS: This transcript is a member of the Human CCDS set: [CCDS8702](#)

Uniprot: This transcript corresponds to the following Uniprot identifiers: [P01116](#)

Transcript Support Level (TSL): [TSL:1](#)

Version: [redacted]

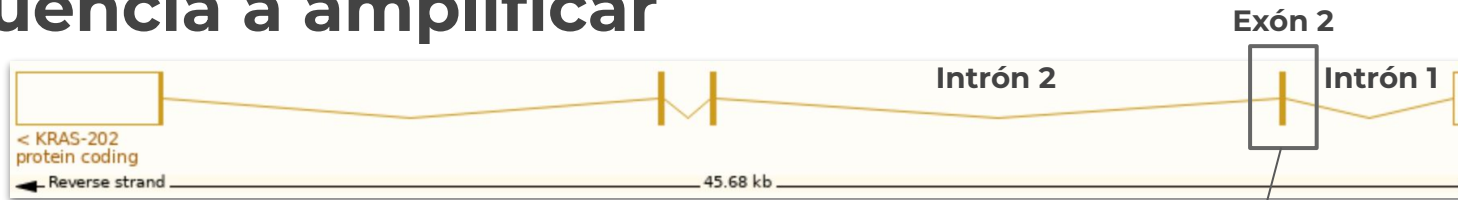
Type: Protein coding

Annotation Method: Transcript where the Ensembl genebuild transcript and the Havana manual annotation have the same sequence, for every base pair. See [article](#).

GENCODE basic gene: This transcript is a member of the [Gencode basic](#) gene set.

En la Práctica 2
queremos amplificar el
exón 2
(el que está mutado)

Secuencia a amplificar



```
>KRAS intron 1 last 500bp - exon 2 - intron 2 first 500bp
TTAAAAGTTTTTCTGTAGCTGTTCATATTGACTTCTAACACTTAGAGGTGGGGTCCAC
TAGGAAAACGTAAACAATAAGAGTGGAGATAGCTGTGAGCAACTTTTGTGAGGGTGTGCT
ACAGGGTGTAGAGCACTGTGAAGTCTCTACATGAGTGAAGTCATGATATGATCCTTTGAG
AGCCTTTAGCCGCGCAGAACAGCAGTCTGGCTATTTAGATAGAACAACCTGATTTTAAG
ATAAAAAGAACTGTCTATGTAGCATTATGCATTTTTCTTAAGCGTCGATGGAGGAGTTG
TAAATGAAGTACAGTTCATTACGATACACGCTCTGCAGTCAACTGGAATTTTCATGATTGA
ATTTTGAAGGTATTTTGAATAAATTTTCATATAAAGGTGAGTTTGTATTTAAAAGGTAC
TGGTGGAGTATTTGATAGTGTATTAACCTTATGTGTGACATGTTCTAATATAGTCACATT
TTCATTATTTTTATTATAAGGCCTGCTGAAAATGACTGAATATAAACTTGTGGTAGTTGG
AGCTGGTGGCGTAGGCAAGAGTGCCTTGACGATACAGCTAATTCAGAATCATTGTGGGA
CGAATATGATCCAACAATAGAGGTAAATCTGTTTAATATGCATATTACTGGTGCAGGA
CCATTCTTTGATACAGATAAAGGTTTCTCTGACCATTTTCATGAGTACTTATTACAAGAT
AATTATGCTGAAAGTTAAGTTATCTGAAATGTACCTTGGGTTTCAAGTTATATGTAACCA
TTAATATGGGAACTTTACTTTTCCCTTGGGAGTATGTCAGGGTCCATGATGTTCACTCTCTG
TGCATTTTGATTGGAAGTGTATTTTCAGAGTTTCGTGAGAGGGTGAGAAAATTTGTATCCTAT
CTGGACCTAAAAGACAATCTTTTTATTGTAACTTTATTTTTATGGGTTTCTTGGTATTG
TGACATCATATGTAAAGGTTAGATTTAATTGTACTAGTGAATATAATTGTTTGTATGTT
GATTTTTTTAACTTCATCAGCAGTATTTTCCATCTTCTTCTCAACATTAGAGAACCTA
CAACTACCGGATAAATTTTACAAAATGAATTATTTGCCTAAG
```



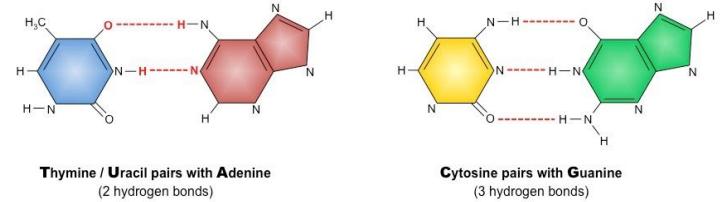
El archivo está en **formato FASTA**:

Formato para representar secuencias de nucleótidos y aminoácidos.

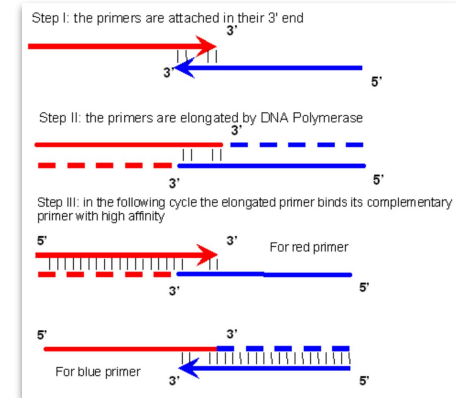
- Línea de **encabezado** que comienza por ">". Contiene información sobre la secuencia (ID, nombre...)
- Líneas de **secuencia**

Diseño de primers para PCR

- **Longitud:** 18-24 pb de longitud
- **Contenido en G/C:**
 - Contenido en G/C del 40-60%
 - 1-2 pares G/C al principio y al final
- **Temperatura de desnaturalización (T_m):**
 - 50-60°C
 - Menos de 5°C de diferencia entre primers
- **Secuencia:**
 - Sin regiones complementarias entre ellos:
dímeros de primers
 - Evitar regiones que formen estructuras secundarias internas
 - Evitar hibridaciones inespecíficas con otras partes del genoma



Formación de dímeros de primers



Primer-BLAST

Pegamos la
secuencia a
amplificar en
formato
FASTA

NIH U.S. National Library of Medicine
National Center for Biotechnology Information [Log in](#)

Primer-BLAST

A tool for finding specific primers

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

Primers for target on one template | Primers common for a group of sequences

Retrieve recent results | Publication | Tips for finding specific primers | [Save search parameters](#) | [Reset page](#)

PCR Template

Enter accession, gi, or FASTA sequence (A refseq record is preferred) [?](#) [Clear](#)

```
>KRAS intron 1 last 500bp - exon 2 - intron 2 first 500bp
TTAAAAGTTTTTCTGTAGCTGTTGCATATTGACTTCTAACACTTAGAGGTGGGGTCCAC
TAGGAAACTGTAACAATAAGAGTGGGAGATAGCTGTGAGCAACTTTTGTGAGGGTGTGCT
ACAGGGTGTAGAGCACTGGAAGTCTACATGAGTGAAGTCATGATATGATCCTTTGAG
AGCCTTAGCCGCCGAGAACAGCAGTCTGGCTATTAGATAGAACAACCTGATTTAAG
```

Or, upload FASTA file No file chosen

Range [?](#) [Clear](#)

Forward primer: From To

Reverse primer:

Primer Parameters

Use my own forward primer (5'->3' on plus strand) [?](#) [Clear](#)

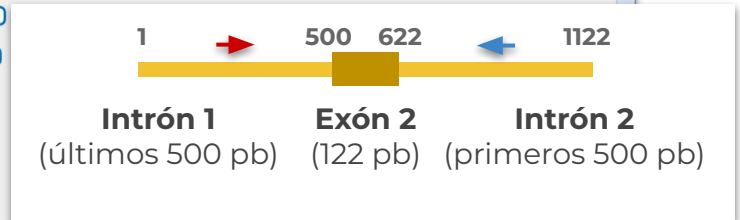
Use my own reverse primer (5'->3' on minus strand) [?](#) [Clear](#)

PCR product size: Min Max

of primers to return:

Primer melting temperatures (T_m): Min Opt Max Max T_m difference [?](#)

Determinamos la
localización de
los primers



Primer-Blast

Exon/intron selection

Exon junction span:

Exon junction match: Min 5' match: Min 3' match: Max 3' match:

Intron inclusion: Primer pair must be separated by at least one intron on the corresponding genomic DNA

Intron length range: Min: Max:

Note: Parameter values that differ from the default are highlighted in yellow

Primer Pair Specificity Checking Parameters

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

Search mode:

Database: (highlighted in yellow)

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

Organism:

Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type.

Entrez query (optional):

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

Max target amplicon size:

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

(highlighted in yellow)

Show results in a new window Use new graphic view

Note: Parameter values that differ from the default are highlighted in yellow

Primer-Blast tiene en cuenta la secuencia del **resto del genoma humano** para no generar *primers* con hibridaciones **inespecíficas**

Entrega en PRADO: hasta 22/03/2024 - 14:00h

- Si ya has contestado las 6 preguntas, súbelas a la entrega de PRADO.
- Si necesitas más tiempo, la entrega estará abierta hasta el **viernes 22 de marzo de 2024 a las 14:00**.

CUESTIONES SEMINARIO 1: HERRAMIENTAS BIOINFORMÁTICAS EN GENÓMICA Grado en Medicina | Biología Molecular | Curso 2023/24

Nombre y Apellidos:

1. Sobre el gen *KRAS*:
 - a) ¿Cuál es su Ensembl ID?
 - b) ¿En qué cromosoma, coordenadas y hebra se localiza?
 - c) Nombra sus 2 genes más próximos.
 - d) Indica el dbSNP ID y la consecuencia de una variante patogénica de este gen.
2. Sobre la isoforma (transcrito) de referencia del gen *KRAS*:
 - a) ¿Cuál es su Ensembl ID?
 - b) ¿Cuál es su longitud en pares de bases y en aminoácidos?
 - c) ¿Cuántos exones tiene?
3. Indica el identificador y la secuencia de un miRNA maduro que regule la expresión del gen *KRAS*.
4. ¿Qué laboratorio en España realiza el diagnóstico genético de mutaciones en *KRAS*?
5. Nombra otra línea celular distinta a NCI-H358 que contenga la mutación G12C en *KRAS*.
6. Indica la pareja de primers que has diseñado para amplificar el exón 2 completo de *KRAS*.

