

# Análisis de mutaciones en el ARNm de TP53

Bioinformática - Ingeniería Biomédica

Jose David López y Sofía Mendivelso Guerrero

18 de noviembre de 2024

## 1. Introducción

El cáncer es una de las principales causas de morbilidad y mortalidad en el mundo, representando un desafío significativo para la salud pública global que refleja la necesidad urgente de desarrollar herramientas efectivas para su diagnóstico, tratamiento y prevención. El cáncer surge como resultado de la acumulación progresiva de alteraciones genéticas que desregulan los procesos normales de crecimiento y muerte celular. Estas alteraciones incluyen mutaciones puntuales, polimorfismos de un solo nucleótido (SNP), modificaciones epigenéticas, y la pérdida o disfunción de genes clave, como los supresores de tumores. Estas variaciones genéticas no solo promueven la inestabilidad genómica, sino que también contribuyen a la transformación de células normales en células malignas [1].

En este contexto, el gen TP53 se destaca como uno de los genes supresores de tumores más relevantes y estudiados en la genética del cáncer. Este gen codifica para la proteína p53, la cual es crucial en la regulación del ciclo celular, la inducción de apoptosis y la activación de mecanismos de reparación del ADN. Sin embargo, mutaciones en TP53, presentes en más del 50% de los casos de cáncer humano, pueden alterar estas funciones, llevando a la proliferación descontrolada de células cancerosas y propiciando su diseminación en el organismo. Estas mutaciones se han encontrado asociadas con enfermedades como el síndrome de Li-Fraumeni y numerosos tipos de cáncer [1, 2].

Dado el papel central de TP53 en la prevención del cáncer, el análisis detallado de sus mutaciones es fundamental para comprender su impacto en la estructura y función de la proteína. Este estudio se enfoca en investigar cómo las variaciones en el ARNm de TP53 afectan sus propiedades funcionales, por lo que utilizando herramientas bioinformáticas como alineamientos múltiples y predicciones estructurales se busca identificar regiones susceptibles a mutaciones patogénicas y clasificar las alteraciones según su impacto funcional. Lo anterior no solo permitirá ampliar el conocimiento sobre la biología del cáncer, sino que también podría facilitar el diseño de estrategias terapéuticas dirigidas que mejoren los resultados clínicos en pacientes oncológicos [1, 2].

## 2. Marco teórico

### 2.1. El gen y la proteína p53

El gen TP53, también conocido como el gen supresor de tumor p53, codifica una proteína crucial en la regulación del ciclo celular y la prevención del cáncer. Esta proteína es una fosfoproteína que se localiza principalmente en el núcleo celular. Fue descubierta a finales de la década de 1970

como una proteína de 53 kD capaz de unirse al antígeno transformante SV40 T, una propiedad que comparte con la proteína retinoblastoma, pRb. Originalmente, se pensó que la proteína p53 era un oncogén, capaz de immortalizar células o transformarlas junto con el oncogén ras. Sin embargo, investigaciones posteriores demostraron que la forma mutante de p53 es la responsable de esos efectos, mientras que la forma salvaje de p53 cumple una función supresora de tumores [3].

Una de las funciones más importantes de la proteína p53 es la regulación del ciclo celular, particularmente en el paso de la fase G1 a la fase S. En presencia de daño en el ADN, p53 detiene el ciclo celular en G1-S, permitiendo que los mecanismos de reparación del ADN actúen y así asegurando la integridad genómica, la reparación y síntesis del ADN, la diferenciación celular, la apoptosis, entre otros procesos. Mientras que la forma salvaje de p53 actúa como un gen supresor de tumor recesivo, las mutaciones en p53 confieren propiedades de oncogén dominante, favoreciendo el desarrollo de tumores y cánceres [3].

## 2.2. La constitución del ARN mensajero (ARNm)

El ARN mensajero (ARNm) es una molécula crucial en la biología celular, ya que transporta la información genética del ADN a los ribosomas, donde se lleva a cabo la síntesis de proteínas. En los organismos eucariotas, el ARNm es primero transcrito del ADN en el núcleo celular y luego, para su maduración, sufre varias modificaciones. Este procesamiento incluye la eliminación de intrones (secuencias no codificantes), y la adición de una "cabeza" de guanina en el extremo 5' y una cola poli-A en el extremo 3'. Estas modificaciones son esenciales para estabilizar la molécula, prevenir su degradación y facilitar su exportación del núcleo hacia el citoplasma.

Además, el ARNm lleva las instrucciones genéticas necesarias para la producción de proteínas en los ribosomas. La secuencia codificante en el ARNm se organiza en grupos de tres nucleótidos llamados códon, cada uno de los cuales codifica un aminoácido específico. El proceso de traducción comienza en el códon de inicio (generalmente AUG) y finaliza en un códon de paro, donde la síntesis proteica cesa. Comprender esta estructura es fundamental al analizar la secuencia de ARNm de cualquier gen, ya que las mutaciones que ocurren en estos elementos pueden afectar la producción y función de la proteína resultante, como en el caso de p53 [4].

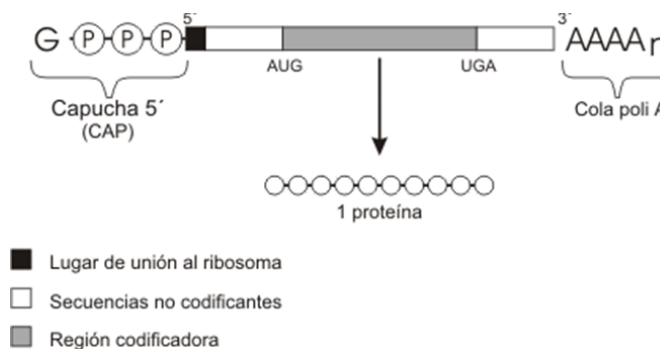


Figura 1: Composición del ARNm

## 3. Metodología

La metodología para este trabajo se basó en el análisis de la secuencia del gen TP53, utilizando como principal fuente de datos el repositorio The National Center for Biotechnology Information (NCBI), accesible en <https://ncbi.nlm.nih.gov/>. La secuencia seleccionada para el estudio fue

NM\_000546.6 Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA, la cual corresponde específicamente al segmento del gen encargado de la síntesis de la proteína p53. Es importante destacar que no se utilizó la secuencia Chromosome 17, NC\_000017.11, ya que esta abarca la totalidad del gen, incluyendo exones e intrones, y no se ajustaba al enfoque del análisis. La longitud de la secuencia seleccionada, NM\_000546.6, es de 2512 nucleótidos.

Con la secuencia completa, se utilizó la herramienta ClinVar en la misma plataforma de NCBI para identificar variaciones conocidas del gen TP53. ClinVar permitió generar una tabla que clasifica las mutaciones según su significancia clínica en categorías como patogénica, probablemente patogénica, de significancia incierta, probablemente benigna o benigna. Para garantizar la calidad de los datos, se filtraron únicamente las mutaciones con una calificación de tres estrellas, lo que indica que fueron revisadas por un panel de expertos. Posteriormente, se seleccionaron cinco mutaciones representativas para cada categoría, obteniendo un total de 25 mutaciones. Cabe destacar que todas las mutaciones analizadas consistieron en cambios de un único nucleótido (por ejemplo, NM\_000546.6(TP53)c.1096T>G (p.Ser366Ala)), indicando la posición y el cambio específico dentro de la secuencia estudiada.

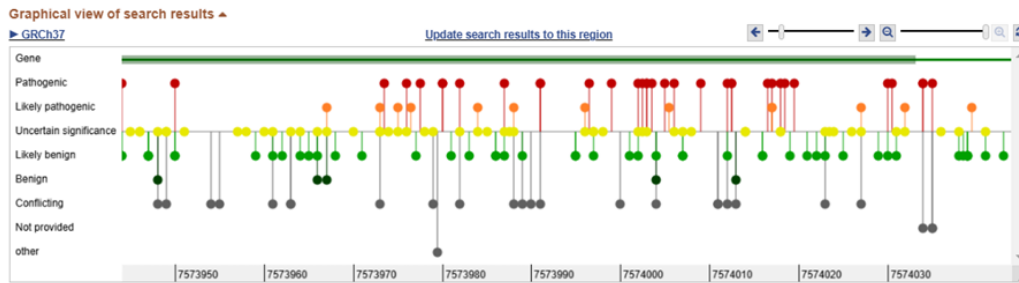


Figura 2: Mutaciones según su significancia clínica

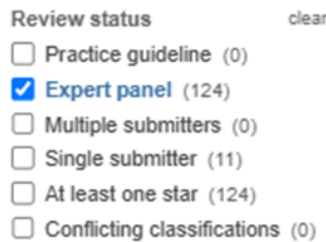


Figura 3: Filtro aplicado para selección de mutaciones

Mediante un script en Python, disponible en el archivo code.py del repositorio GitHub, se generaron las mutaciones a partir de la secuencia original. Posteriormente, se tradujeron todas las secuencias mutadas y la secuencia sana, por lo que para esto se comenzó desde el codón de inicio hasta el codón de paro más próximo. La traducción de la secuencia sana fue validada contra la secuencia proteica reportada en la literatura, confirmando su correspondencia.

A continuación, se realizaron alineamientos múltiples utilizando la herramienta Clustal Omega (clustalo). Se llevaron a cabo dos tipos de alineamientos: el primero agrupó las mutaciones por categoría, y el segundo incluyó todas las mutaciones simultáneamente. Los alineamientos resultantes se almacenaron en la carpeta alineamientos del repositorio. Para facilitar el análisis, los datos se

exportaron a la herramienta Jalview, donde se activaron colores por conservación para visualizar más claramente los patrones de mutación y su impacto.

Finalmente, se tuvo en cuenta la relación de las mutaciones con condiciones específicas. Según la información de ClinVar, todas las mutaciones analizadas están asociadas al síndrome de Li-Fraumeni (Li-Fraumeni Syndrome), lo que subraya la relevancia clínica del estudio. Esta información permitió interpretar los resultados de manera más integral, considerando tanto las características moleculares como el contexto patológico de las mutaciones.

Todos los procesos realizados en WSL y los alineamientos generados están disponibles en el repositorio de GitHub, accesible en el siguiente enlace: [https://github.com/Joxzes/Proyecto\\_Bioinfo\\_p53](https://github.com/Joxzes/Proyecto_Bioinfo_p53)

## 4. Resultados

En esta sección se presentan los resultados obtenidos tras el análisis de las mutaciones seleccionadas del gen TP53. Los alineamientos múltiples realizados previamente utilizando Clustal Omega fueron visualizados y analizados en Jalview, donde se generaron representaciones gráficas que permitieron observar la conservación de las secuencias y las diferencias entre las categorías de mutaciones (patogénicas, benignas, etc.). Los alineamientos por categoría se encuentran organizados en la carpeta alineamientos, permitiendo un acceso sistemático a los datos procesados. Además, en Jalview se llevó a cabo una predicción de estructura secundaria (secondary structure prediction) para explorar posibles cambios estructurales causados por las mutaciones, aportando información adicional sobre el impacto funcional de las variantes analizadas.

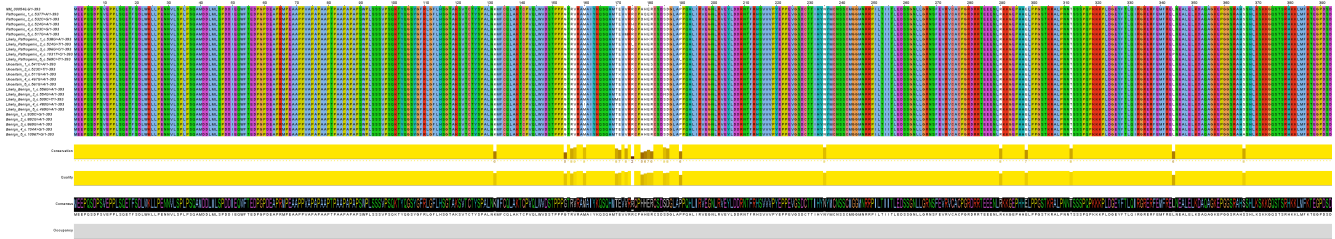


Figura 4: Alineamiento de mutaciones por categoría con conservación de secuencias en Jalview

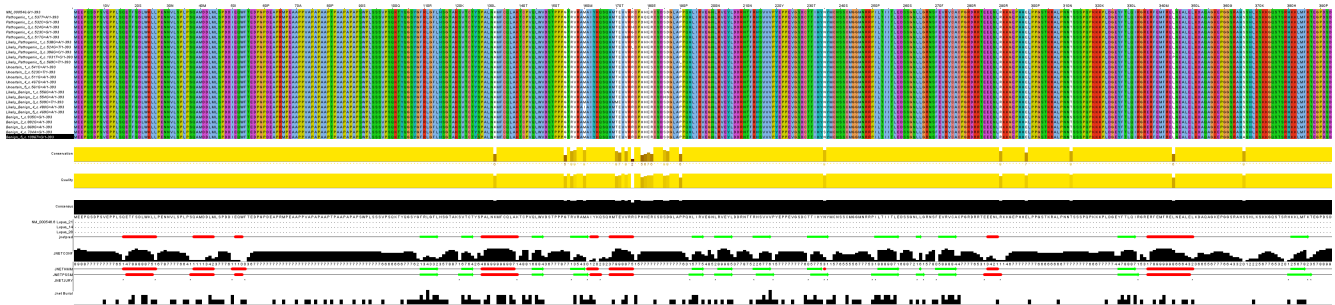


Figura 5: Predicción de estructura secundaria de la proteína TP53 en Jalview

## 5. Discusión

La proteína TP53 es fundamental en la regulación del ciclo celular, la reparación del ADN y la inducción de apoptosis. Estas funciones son esenciales para mantener la integridad genómica y prevenir la formación de tumores. Las mutaciones en el gen TP53 pueden comprometer estas funciones, promoviendo procesos tumorales. En este análisis, se examinaron mutaciones clasificadas como patogénicas, probablemente patogénicas, inciertas, benignas y probablemente benignas, evaluando su ubicación en la secuencia proteica, sus características y su impacto potencial en la funcionalidad de TP53.

Las mutaciones patogénicas analizadas se encuentran principalmente en el dominio de unión al ADN (residuos 102-292), una región crítica para la función de TP53. Entre las alteraciones más relevantes se encuentran p.His179Gln, p.His178Asp, p.Arg175His, p.Arg175Gly y p.Val173Met. Estas mutaciones afectan aminoácidos esenciales para la interacción entre TP53 y el ADN. Por ejemplo, la sustitución de Arg175, que forma enlaces de hidrógeno críticos con el ADN, por histidina o glicina, disminuye significativamente la afinidad de unión, comprometiendo la capacidad de TP53 para activar genes involucrados en la reparación del daño celular. Del mismo modo, las mutaciones His178Asp y His179Gln alteran las interacciones electrostáticas necesarias para estabilizar la unión al ADN, mientras que Val173Met introduce cambios en la estabilidad estructural del dominio central, afectando indirectamente la función de unión. Estas alteraciones limitan la capacidad de TP53 para actuar como supresor tumoral, permitiendo la proliferación celular descontrolada.

Las mutaciones probablemente patogénicas comparten características con las patogénicas, pero su impacto funcional puede depender del contexto específico. Entre estas mutaciones destacan p.Lys180Gln, p.Arg175Leu, p.Lys132Asn, p.Leu344Pro y p.Pro190Leu. Estas alteraciones afectan regiones clave relacionadas con la unión al ADN o la estabilidad estructural de la proteína. Por ejemplo, Lys180Gln y Arg175Leu comprometen interacciones electrostáticas esenciales, mientras que Leu344Pro y Pro190Leu generan cambios conformacionales significativos que desestabilizan la estructura global. Aunque estas mutaciones no afectan de manera tan directa como las patogénicas, su ubicación en regiones funcionalmente importantes sugiere un impacto potencialmente relevante.

Las mutaciones inciertas representan un grupo intermedio cuyo impacto funcional no está completamente definido. Estas mutaciones suelen localizarse en posiciones que no son claramente críticas, pero tampoco en regiones irrelevantes, lo que dificulta su clasificación. Factores como el contexto celular, las interacciones con otras moléculas y los cambios en la estabilidad local de la proteína podrían influir en su efecto.

Por otro lado, las mutaciones benignas y probablemente benignas se encuentran mayoritariamente fuera del dominio de unión al ADN o en regiones menos conservadas. Entre estas mutaciones se identificaron p.Thr312Ser, p.Glu298Lys, p.Arg290His, p.Asn235Ser y p.Ser366Ala. Estas alteraciones suelen implicar cambios mínimos en la estructura o función de TP53. Por ejemplo, p.Thr312Ser y p.Ser366Ala producen cambios en residuos polares con un impacto funcional insignificante. Mutaciones como p.Arg290His, aunque cercanas al dominio de unión al ADN, no afectan significativamente la estabilidad de la proteína, mientras que p.Glu298Lys y p.Asn235Ser representan sustituciones de residuos con propiedades similares, preservando la funcionalidad global.

Al realizar un análisis más detallado se identifica un patrón importante en la distribución de estas mutaciones. La mayoría de las alteraciones patogénicas y probablemente patogénicas se concentran en el rango de residuos 150-190, una región que incluye partes clave del dominio de unión al ADN y sitios relacionados con la oligomerización de TP53. Esto explica su impacto crítico en la funcionalidad de la proteína. En contraste, las mutaciones benignas y probablemente benignas están más dispersas y se encuentran en los extremos N-terminal y C-terminal, o en regiones menos

conservadas, lo que justifica su impacto funcional limitado.

La clasificación de las mutaciones como patogénicas o benignas depende de diversos factores. La ubicación es uno de los más determinantes, ya que las alteraciones en regiones funcionalmente críticas tienden a ser patogénicas. La naturaleza del cambio aminoacídico también es crucial; cambios que afectan la polaridad, la carga o la hidrofobicidad del residuo suelen tener un mayor impacto. Además, la conservación evolutiva de los residuos afectados indica su importancia funcional, y mutaciones en posiciones altamente conservadas tienen más probabilidades de ser dañinas. Por último, el impacto en la estabilidad estructural y las interacciones proteína-proteína son factores determinantes en la funcionalidad de TP53.

En conclusión, este análisis muestra la importancia de la ubicación y naturaleza de las mutaciones en TP53 como determinantes de su efecto funcional. Las alteraciones en regiones conservadas, especialmente en el dominio de unión al ADN, son mayormente patogénicas debido a su impacto directo en la interacción proteína-ADN y la regulación génica. Por el contrario, las mutaciones benignas suelen localizarse en regiones menos relevantes, con cambios mínimos en la estabilidad o función de la proteína. Estos hallazgos son fundamentales para comprender el papel de las mutaciones de TP53 en el desarrollo del cáncer y su potencial uso para el diseño de estrategias terapéuticas específicas basadas en patrones mutacionales.

## Referencias

- [1] López-Camarillo C, Machorro P, García-González E, Pacheco-López G. El gen TP53 y su relación con el cáncer: una revisión de su rol en la genética molecular. *Rev Mex Cancerol*. 2006; 3(3): 123-30. [https://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0034-83762006000300010](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0034-83762006000300010)
- [2] National Cancer Institute. Gen TP53. *Diccionario del cáncer*. <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/gen-tp53>.
- [3] Latorre M, Navarro A, Moreno P, et al. El gen TP53: Un supresor de tumores multifuncional. *Dialnet*. 2019. <https://dialnet.unirioja.es/servlet/articulo?codigo=7113288>.
- [4] Genotipia. El ARNm: Estructura y función del ARN mensajero. *Genotipia*. 2023. <https://shorturl.at/vk4Ho>