

# CÁPSULAS EN MEDICINA DE LABORATORIO

[www.traineecouncil.org](http://www.traineecouncil.org)

**TÍTULO: Nomenclatura de secuencias de ADN e interpretación de variantes**

**PONENTE: Lora J. H. Bean, PhD, FACMG**

---

## **Diapositiva 1: Introducción**

Hola, me llamo Lora Bean. Soy Profesora Asistente y Directora General del Laboratorio de Genética de Emory, en el Departamento de Genética Humana de la Universidad Emory. Bienvenido a esta Cápsula en Medicina de Laboratorio sobre "Nomenclatura de secuencias de ADN e interpretación de variantes".

## **Diapositiva 2: Genoma Humano**

El genoma humano consta de más de 3 mil millones de pares de bases. Aunque el Proyecto del Genoma Humano se completó en 2003, aún existen lagunas en la secuencia. A medida que se completan estos espacios, la ubicación o "posición del mapa" de un par de bases específicas, a menudo cambia con cada versión de la secuencia de referencia del genoma humano. Una secuencia de referencia proporciona un marco en el cual se ubican variaciones en el genoma, y sabemos que existe mucha variación. Nos referimos a las diferencias de un solo nucleótido entre los individuos, tales como los polimorfismos de un solo nucleótido (o SNPs, por sus siglas en inglés) y ganancias o pérdidas de secuencias más grandes entre las personas, tales como variantes de número de copias (o CNVs, por sus siglas en inglés).

Existe interés en estas variantes por diferentes razones: algunas variantes son benignas y son cambios comunes que se utilizan en estudios de asociación o de números de copias, mientras que otros son cambios perjudiciales relacionados con enfermedades en humanos. Estas variantes patogénicas identificadas en la era pre-genómica a menudo se denominan utilizando nomenclatura histórica no estándar.

## Nomenclatura de secuencias de ADN e interpretación de variantes

### **Diapositiva 3: Secuenciación clínica**

Las variantes patogénicas en los genes humanos causan enfermedades genéticas hereditarias. Cuando una o algunas variantes patogénicas representan la mayoría de los alelos de la enfermedad en una población, estas se pueden evaluar mediante genotipificación dirigida. En la mayoría de los casos de enfermedades genéticas, las mutaciones aisladas o familiares (es decir, hereditarias) comprenden la mayor parte del espectro de mutaciones.

Los laboratorios clínicos utilizan la secuenciación de genoma completo para detectar cambios patogénicos. Con esta técnica, las variantes nuevas (previamente no reportadas) se detectan a menudo. En el laboratorio clínico, necesitamos utilizar recursos que nos ayuden a determinar la naturaleza patogénica de tales variantes. También necesitamos usar la nomenclatura estándar para describir la variante.

### **Diapositiva 4: Nomenclatura estándar**

Sin importar el propósito del estudio, es muy importante que podamos comparar datos entre estudios y entre individuos. La mejor manera de garantizar que una variante de ADN específica pueda identificarse sin ambigüedades, es utilizar nomenclatura convencional, ampliamente comprendida. Actualmente, se utilizan números de acceso a la secuencia de referencia que están disponibles públicamente. Este consenso es particularmente importante cuando se refieren a mutaciones genéticas raras o familiares, ya que puede pasar una generación o más, antes de que otros miembros de la familia se presenten para la prueba, momento en el cual los portadores de mutaciones conocidas pueden ya no estar disponibles para estudios confirmatorios.

### **Diapositiva 5: Nomenclatura estándar**

La Sociedad de Variaciones del Genoma Humano (HGVS) ofrece recomendaciones ampliamente aceptadas para la nomenclatura de secuencias. Hay recomendaciones para nombrar cambios de secuencia genómica (como g punto), codificación (como c punto), mitocondrial (como m punto), ARN (como r punto) y proteína (como p punto) en comparación con una referencia. La Sociedad de Variaciones del Genoma Humano (HGVS) tiene acuerdos de nomenclatura para variantes de pares de bases individuales, deleciones, duplicaciones, inserciones, grandes reordenamientos, variantes intrónicas y casi cualquier escenario que se haya reportado.

## Nomenclatura de secuencias de ADN e interpretación de variantes

### **Diapositiva 6: Nomenclatura estándar: Recomendaciones de la Sociedad de Variaciones del Genoma Humano**

A continuación se muestran algunos ejemplos. Los cambios de una secuencia de referencia se describen nombrando la base de referencia o el aminoácido y tomando en cuenta la nueva secuencia. Los cambios de nucleótidos y aminoácidos no tienen sentido sin una secuencia de referencia, por lo que su uso es crítico. Es importante considerar que en la nomenclatura genómica, el cambio de nucleótidos puede ser diferente que en la nomenclatura de codificación. Esto se debe a que los números genómicos se refieren a una cadena del cromosoma, conocida como cadena "+", y un gen puede estar en la orientación opuesta en el cromosoma o en la cadena "-".

### **Diapositiva 7: Nomenclatura estándar**

Tomemos el ejemplo de la mutación más común de la enfermedad de células falciformes. La mayoría de las referencias en la literatura anterior, así como en los registros médicos, usan el nombre histórico de hemoglobina S (HbS), o HBB: ácido glutámico 6 valina, y estos nombres nos resultan familiares. Utilizando la nomenclatura estándar, esta mutación puede ser nombrada utilizando la genómica, ARNm o una proteína de referencia. La variante patogénica es de una T a una A porque el gen HBB está en la cadena "-" del cromosoma. Utilizando la nomenclatura estándar, el cambio de proteína ahora se conoce como p. Ácido glutámico 7 Valina. Para obtener datos sobre esta mutación de proyectos de secuenciación a gran escala o plataformas de genotipificación, debe conocerse la ubicación genómica o un número de referencia, como un número dbSNP rs.

### **Diapositiva 8: Interpretación de variantes de secuencia**

La mutación de células falciformes que se define como la hemoglobina falciforme (HbS) está bien establecida como patógena. Sin embargo, no todas las variantes en el genoma son tan fácilmente interpretables. Los laboratorios clínicos siguen las normas y directrices profesionales del Colegio Americano de Genética Médica para clasificar las variantes de secuencia en una de cinco categorías: patogénicas, probablemente patogénicas, de importancia clínica desconocida, probablemente benignas o benignas. Cabe mencionar que las directrices se están revisando y actualizando constantemente.

### **Diapositiva 9: Variantes patogénicas**

Algunas variantes se clasifican como patogénicas, incluso si no se han informado previamente porque se predice que destruirán la función de la proteína. Estas incluyen mutaciones sin sentido, del marco de lectura y del sitio de empalme. Las variantes de sentido erróneo y las deleciones o duplicaciones en el marco de lectura, así como los cambios silenciosos (variantes de ADN que se pronostica no cambiarán un aminoácido) o los cambios intrónicos fuera de las posiciones

## Nomenclatura de secuencias de ADN e interpretación de variantes

más/menos 1 y 2, también pueden clasificarse como perjudiciales si hay suficiente evidencia en la literatura de la patogenicidad del cambio.

### **Diapositiva 10: Variantes patogénicas: GJB2: c.139G>T (p.E47X)**

Veamos un ejemplo simple de un cambio sin sentido en el gen GJB2 identificado por análisis de secuencia. El gen GJB2 codifica la proteína conexina 26. Las mutaciones en este gen están asociadas con la pérdida auditiva autosómica recesiva. La secuencia de referencia es el trazo inferior en la que puede verse el GAG normal que codifica un ácido glutámico. La secuencia del paciente es el trazo superior. Se puede observar que el paciente es heterocigoto para un cambio de G a T en la posición de codificación 139. Además del GAG normal que codifica el ácido glutámico, el paciente también tiene un TAG que codifica un codón de paro. Este codón de paro prematuro se interpreta como un cambio patogénico.

### **Diapositiva 11: Variantes patogénicas: PAH: IVS12+1G>A**

Este es un ejemplo de una mutación común en el sitio de empalme en el gen *PAH*, que está mutado en pacientes con fenilcetonuria (PKU, por sus siglas en inglés). Recuerde que cuando el ARN se transcribe del ADN, los intrones están presentes. Para producir un ARNm maduro, los intrones deben eliminarse de este ARNm previo.

La maquinaria de empalme reconoce los intrones por su secuencia. La secuencia "GU" al comienzo del intrón, conocida como donador de empalme, y la secuencia "AG" al final del intrón, conocida como aceptor de empalme, las cuales son casi invariables. Las variantes en estas posiciones +1, +2, -1 y -2 se interpretan como patogénicas.

### **Diapositiva 12: Variantes benignas**

En el otro extremo del espectro, tenemos cambios benignos. Clasificamos las variantes como benignas si se observan con una frecuencia en la población que es demasiado común para ser una mutación que causa una enfermedad. Estos datos pueden provenir de grandes estudios de población o de grupos de control en estudios publicados. A veces, los resultados de estudios funcionales publicados demuestran que una variante no tiene efecto sobre la función de la proteína.

### **Diapositiva 13: Variación benigna GJB2**

Los proyectos de secuenciación a gran escala nos han proporcionado una visión mucho más amplia de la variación común en la población humana. Por ejemplo, el Proyecto de Secuenciación del Exoma del NHLBI GO ha hecho públicos los datos de frecuencia de más de 8,000 alelos derivados de Europa y más de 4,000 alelos derivados de África. Estos son valiosos al determinar la patogenicidad de una variante de secuencia.

## Nomenclatura de secuencias de ADN e interpretación de variantes

Por ejemplo, la variante Citosina (C) a Timina (T) de 34 nucleótidos río arriba al inicio de ATG en el gen GJB2 (denominado “codificante -34 Citocina (C) a Timina (T), es común en la población de África con más del 23% de los alelos siendo una Timina (T). Este cambio se define como una variante benigna.

La variante Guanina (G) a Adenina (A) en la posición de codificación 79, es más difícil de interpretar, ya que es rara tanto en las poblaciones europeas como africanas.

### **Diapositiva 14: Variación benigna GJB2: c.79G>A (p.V27I)a**

Los datos del proyecto HapMap, un estudio de población a gran escala diseñado para estudiar la variación genética, provienen de una población más diversa. La variante c.79G> A en el gen GJB2 es común en las poblaciones asiáticas. Con más del 15% de las personas estudiadas que son homocigotas para esta variante podemos concluir que la variante no está asociada con la pérdida auditiva. Estos datos ponen de manifiesto la importancia de estudiar diversas poblaciones.

### **Diapositiva 15: Variantes de significado incierto**

Muchas variantes de secuencia no se han observado previamente y no se puede determinar su patogenicidad. Dichos cambios incluyen aquéllos de sentido erróneo o contrasentido (o aminoácidos) que no se han informado previamente, deleciones o duplicaciones en el marco (pérdida o adición de aminoácidos sin un cambio en el marco de lectura) o, como ahora se identifica por secuenciación del exoma, cambios de nucleótidos en genes cuya asociación con enfermedades humanas es desconocida.

### **Diapositiva 16: Variante de significado incierto: MLL2: c.10740G>A (p.Q3580Q)**

Aquí se muestra un ejemplo de una variante de significado desconocido en el gen MLL2, mutaciones en las que se produce el síndrome de Kabuki. El síndrome de Kabuki es un síndrome de dismorfología autosómica dominante. Una sola variante patogénica en MLL2 causa el síndrome de Kabuki. En este caso, se identificó un cambio de Guanina (G) a Adenina (A) en la posición de codificación 10,740, el último par de bases del exón 38.

### **Diapositiva 17: Variante de significado incierto: MLL2: c.10740G>A (p.Q3580Q)**

La variante codificante. 10,740 Guanina (G) a Adenina (A) no se ha observado en estudios poblacionales o en individuos con síndrome de Kabuki. Además, este cambio no se prevé que cambie el aminoácido glutamina en el codón 3,580. Entonces, ¿cómo determinamos la importancia clínica de este cambio? Podríamos evaluar a los padres para ver si el cambio fue heredado. Esto también nos diría si esta variante es una nueva variante en este individuo. Se dice que las variantes encontradas en un paciente pero no en sus padres ocurrieron de *novo*.

## Nomenclatura de secuencias de ADN e interpretación de variantes

Para que un laboratorio pueda decir que se produjo una variante de *novo*, la identidad de los padres debe confirmarse mediante marcadores genéticos.

### **Diapositiva 18: Variante de significado incierto: MLL2: c.10740G>A (p.Q3580Q)**

Analizamos a los padres y descubrimos que ninguno de los dos tenía esta variante. Con el permiso de la familia, utilizamos técnicas moleculares para confirmar las relaciones parentales. Luego interpretamos este cambio como posiblemente patogénico.

La información clínica sobre el paciente y los padres también es crítica para la interpretación. Si este cambio se hubiera heredado, tendríamos que determinar si el padre que lleva el cambio tiene características del síndrome de Kabuki. Dicha información nos ayudaría a construir un caso para llamar a esta variante probablemente benigna o probablemente patogénica, pero una determinación final de ser benigna o patogénica probablemente requeriría identificar estos cambios en otras personas.

### **Diapositiva 19: Variantes de significado incierto**

En el caso que acabo de mostrar, combinamos datos de la literatura y bases de datos de variación genética con datos obtenidos al evaluar a los miembros de la familia de una persona afectada y clasificamos una variante como probablemente patogénica. Los términos probablemente patogénico y probablemente benigno, generalmente se reservan para variantes en genes con fenotipos clínicos o metabólicos claramente asociados; es importante estar seguro de que usted está en el gen correcto. Si el estudio de los padres de una persona afectada demuestra que un cambio de aminoácidos en un sitio conservado evolutivamente se segrega con una mutación conocida (para una enfermedad recesiva) u ocurre de novo (para una enfermedad dominante), entonces esto es evidencia de que la variante es una variante probablemente patogénica. Si se encuentra una variante homocigota en parientes no afectados (para un trastorno recesivo) o heterocigótica en parientes no afectados (para un trastorno dominante), entonces esto es evidencia de que la variante es probablemente benigna. Como con cualquier variante, es importante reevaluar la patogenicidad de las variantes cada vez que se disponga de información nueva.

### **Diapositiva 20:**

Es importante conocer el proceso que utilizan los laboratorios para determinar la patogenicidad de las variantes de secuencia. La recopilación de información precisa y completa sobre una variante en la literatura y de los estudios de población, requiere el uso y la comprensión de la nomenclatura estándar.

### **Diapositiva 21: Interpretación de variantes de secuencia**

Los esfuerzos emergentes permitirán a los laboratorios clínicos centralizar el conocimiento sobre las variantes de secuencia, incluido el proyecto NCBI ClinVar (un archivo público y de acceso

## Nomenclatura de secuencias de ADN e interpretación de variantes

libre de informes sobre las relaciones entre las variaciones y los fenotipos humanos, con pruebas de apoyo) y el Proyecto del Varioma Humano. Estos esfuerzos serán especialmente útiles para los laboratorios clínicos a medida que avanzamos en la era genómica del exoma completo y la secuenciación del genoma completo.

### **Diapositiva 22: Referencias**

### **Diapositiva 23: Declaraciones/Posibles conflictos de interés**

Previa presentación de esta cápsula, el ponente completó el formulario de declaraciones de *Clinical Chemistry*. Declaraciones y/o conflictos de interés potenciales:

### **Diapositiva 24: Gracias en nombre de [www.TraineeCouncil.org](http://www.TraineeCouncil.org)**

Gracias por acompañarme en esta Cápsula en Medicina de Laboratorio sobre "Nomenclatura de secuencias de ADN e interpretación de variantes". Soy Lora Bean.