



Universidad Autónoma Metropolitana

Unidad Xochimilco

DIVISIÓN DE CIENCIAS BIOLÓGICAS Y DE LA SALUD

Licenciatura en Química Farmacéutica Biológica

Informe final de Servicio Social

Variantes en el gen de la biogénesis de los microRNAs *GEMIN4* y el riesgo genético en la leucemia mieloide crónica.

Periodo: 28 de noviembre del 2022 a 29 de Mayo del 2023

Realizado por:

F. Sandybel Ríos-Rodríguez

2183072953@alumnos.xoc.uam.mx

Alumna de la Lic. en QFB de CBS de la Universidad Autónoma Metropolitana Unidad Xochimilco. Calzada del Hueso 1100, col. Villa Quietud, Coyoacán, 04960, Ciudad de México.

Sistemas Biológicos, UAM.X/Instituto Nacional de Medicina Genómica (INMEGEN)

ASESOR INTERNO

Dra. Marisol Lopez Lopez

ASESOR EXTERNO

Dr. Emilio Joaquín Cordova Alarcón

Expresión de gratitud

Agradezco a mi familia por el apoyo brindado durante mi formación como estudiante, son la más importante de mis motivaciones.

Agradezco a mis asesores el Dr. Emilio Joaquin Cordova Alarcon y la Dra. Marisol Lopez Lopez, por compartirme su conocimiento, por ser buenos guías y apoyo. De la misma forma, agradezco a mis compañeros de laboratorio por todos los momentos gratos durante mi estancia

Resumen

El gen *GEMIN4* se encuentra en el cromosoma 17p13.3 con un gran número de variantes identificadas. Está involucrado en múltiples enfermedades incluidas las onco-hematológicas. Por esta razón, se llevó a cabo un estudio retrospectivo de casos y controles, recopilando datos de 843 individuos de ambos sexos de la población mexicana, con la intención de identificar a las SNVs rs7813, rs2740349 y rs4968106 como posibles marcadores de predisposición genética a padecer LMC. El genotipado de las variantes se realizó mediante ensayos Taqman®. La asociación de las variantes y la predicción del riesgo de LMC se realizó mediante el cálculo de valores OR y de significancia estadística. Adicionalmente, se llevó a cabo un análisis de haplotipos para ver el efecto combinado de las variantes. Aun cuando ninguna de las variantes estudiadas mostró asociación significativa con el diagnóstico o el pronóstico de la LMC, se encontró que las mujeres portadoras del alelo menor (A) de la variante rs4968104 se diagnosticaron, en promedio, 9 años antes que las mujeres que no portan este alelo. Sin embargo, esta variante se encontró en desequilibrio de HW en la población de estudio control ($P < 0.05$), por lo que recomendamos llevar a cabo más investigaciones al respecto.

Introducción

La Leucemia Mieloide Crónica (LMC) es un trastorno clonal de las células hematopoyéticas primitivas, que se caracteriza por presentar una hiperproliferación de células mieloides pobremente diferenciadas, es decir, que no llegan a su completa maduración, denominados “blastos”. Los blastos proliferan de forma anormal e invaden progresivamente a la médula ósea, lo que impide el correcto funcionamiento de este órgano hematopoyético. [1]

La LMC es la enfermedad onco-hematológica más común en la población adulta, al representar alrededor del 15-20% de todas las leucemias en adultos. De acuerdo con datos de hospitales del primer mundo, cerca de la mitad de los pacientes con LMC tienen más de 60 años. Sin embargo, en países de tercer mundo, como México, la edad de diagnóstico de esta enfermedad oscila desde los 40 hasta los 60 años. Respecto a su incidencia, ésta es más significativa en adultos mayores que en niños, con una diferencia de 1 por 1 millón en niños a 15 por 1 millón en adultos mayores. También suele presentarse más comúnmente en varones que en mujeres, aunque aún se desconoce el motivo. [1]

La LMC fue la primera neoplasia maligna en ser asociada con una anomalía citogenética bien definida: la presencia del cromosoma Filadelfia (Ph, del inglés: *Philadelphia*), denominado así debido a la ciudad en la cual fue descubierto. Esta alteración es el resultado de una translocación recíproca entre los brazos largos de los cromosomas 9 y 22. Como consecuencia, el gen *BCR* (del inglés: *Breakpoint Cluster Region*) del cromosoma 22 se fusiona con parte del gen *ABL1* (del inglés: *Abelson Leukemia*) del cromosoma 9, lo que da como resultado el gen de fusión *BCR-ABL1* (**figura 1**).

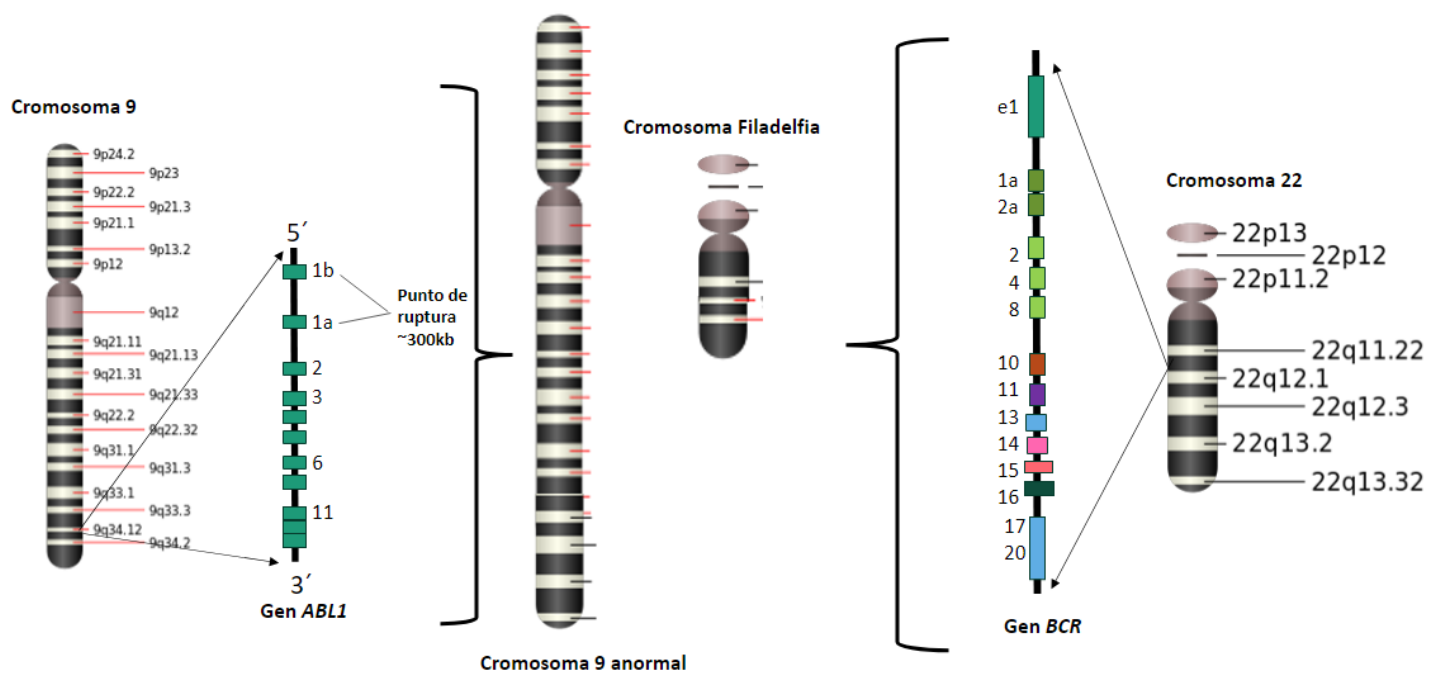


Figura 1. El cromosoma Filadelfia como alteración central en la leucemia mieloide crónica. Es una anomalía cromosómica causada por una translocación recíproca entre las regiones q34.12 y q11.23 de los cromosomas 9 y 22, respectivamente. Citogenéticamente se observan dos cromosomas anormales de distinto tamaño. El cromosoma 22 acortado es denominado cromosoma Ph(9;22) (q34; q11) y contiene el gen de fusión *BCR-ABL1*.

El gen *ABL1* se compone de 12 exones, y codifica para una proteína con actividad de tirosin cinasa no acoplada a un receptor. Diversos estudios previos han identificado el importante papel desempeñado por la proteína ABL1 en el control de la proliferación celular, reparación del DNA e incluso se le ha asociado con la respuesta celular al estrés genotóxico.

Por otro lado, el gen *BCR* está compuesto por 23 exones, y codifica para una proteína con actividad dual de GTPasa y de factor intercambiador de nucleótidos para la familia proteica Rho/Rac. Cabe destacar que se ha identificado la participación de ambos genes en la activación de múltiples vías de señalización celular (p.ej: PI3K-Akt, NF- κ B, JAK-STAT, Ras-MAPK)

En cuanto a la formación del gen de fusión, el punto de ruptura dentro del gen *ABL1* puede ocurrir en cualquier lugar dentro de un área de aproximadamente 300 kb dentro del intrón 1. Sin embargo, aun cuando los puntos de ruptura en este gen suelen ser variados, es debido a la presencia de *splicing* alternativo que siempre suele ser la misma región del gen la que forma parte del gen de fusión, es decir, del exón 2 al 11, lo que se ha denominado como a2.

En cambio, el gen *BCR* presenta 3 sitios de ruptura relacionados con la formación del cromosoma Ph, denominados como: mayor (M-bcr), menor (m-bcr) y micro (μ -bcr) (**figura 2**). El sitio de ruptura M-bcr se encuentra entre los exones 12 y 16 con una longitud de 5,8 kb; mientras que la m-bcr está ubicada en el primer intrón, entre los exones 1 y 2 con una longitud de 5,5 kb. Por otro lado, la región μ -bcr se ubica en el exón 19.

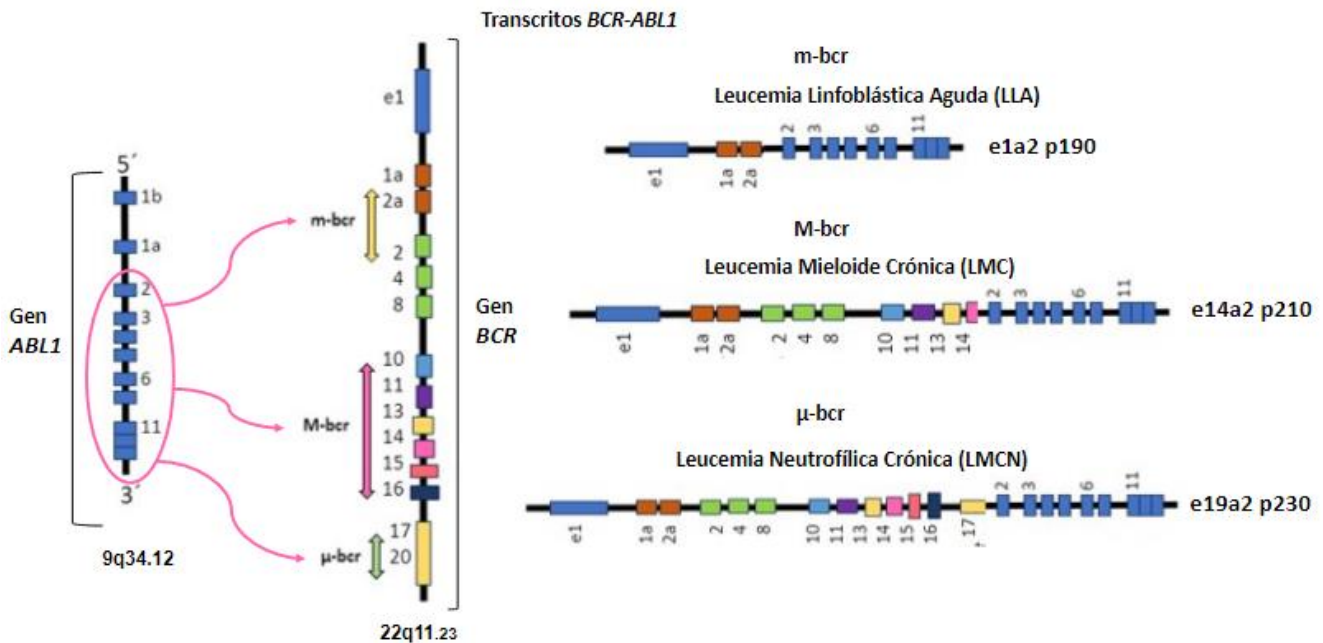


Figura 2. Sitios de ruptura en *BCR* y los genes *BCR-ABL1* relacionados. Cada sitio de ruptura en *BCR* da origen a un diferente gen de fusión. Para la ruptura m-bcr, el gen de fusión está conformado por los exones 1 al 2 de *BCR* y del 2 al 11 de *ABL1*. En la ruptura M-bcr, el gen de fusión está formado por los exones 1 al 13-14 de *BCR* y los exones 2 al 11 de *ABL1*. Finalmente, la ruptura μ -bcr produce la fusión de los exones 1 al 19 de *BCR* y del 2 al 11 de *ABL1*. Por lo tanto, el gen de fusión producido por cada tipo de ruptura se codifica para una proteína de distinto tamaño: p190, p210 y p230, respectivamente. Cada una de las proteínas generadas presentan función de tirosina cinasa pero su presencia se relaciona con distintos tipos de leucemias. La proteína p190 se ha reportado en Leucemia Linfoblástica Aguda (LLA) ph +, así mismo, el gen de fusión *BCR-ABL1* también puede ser detectado en los pacientes adultos con LLA. La proteína p210 es la responsable de las anomalías fenotípicas de la Leucemia Mielode Crónica (LMC); mientras que la proteína p230 está presente en los casos de Leucemia Neutrofílica Crónica (LMCN).

En la LMC, los puntos de ruptura más frecuentes en el gen *BCR* ocurren en los exones: 12(b2) y 13(b3), dentro del punto de ruptura M-bcr. Por otro lado, aun cuando los puntos de ruptura en el gen *ABL1* suelen ser variados, es debido a la presencia del *splicing* alternativo que siempre suele ser la misma región del gen la que forma parte del gen de fusión, es decir, del exón 2 al 11, generando los reordenamientos: b2a2 o b3a2.

Algunos estudios en poblaciones latinoamericanas han demostrado que el transcrito más frecuente es el b3a2 con un 60% en comparación con el transcrito b2a2 que tiene una frecuencia del 40%. Sin embargo, se ha observado una diferencia en población mexicana, donde el transcrito más frecuente es el b2a2 (70%) seguido de b3a2 (30%).^[3,4]

El gen de fusión *BCR-ABL1* que es generado por la M-bcr, se transcribe en un RNAm de 8.5kb y codifica para una proteína quimérica de 210 kDa, denominada BCR-ABL1 p210 (**figura 2**), la cual presenta una actividad de tirosin cinasa constitutivamente activa. A través de esta actividad exacerbada, la oncoproteína BCR-ABL1 promueve la proliferación celular e inhibe la apoptosis, lo que se ha identificado como un suceso central en la leucemogénesis.^[2] Actualmente, se considera al gen *BCR-ABL1* como el más importante biomarcador genético para la LMC y está reportado en el 95% de los casos.^[2,3]

El cuadro clínico de la LMC, se encuentra estrechamente relacionado con la actividad de tirosina cinasa de BCR-ABL1, provocando que los pacientes presenten una hiperleucocitosis que generalmente se detecta cuando se les realiza un hemograma completo motivado por algún malestar ajeno al desarrollado por la LMC.

La evolución clínica de la enfermedad se divide en 3 fases: crónica, acelerada y blástica. De acuerdo con la Organización Mundial de la Salud (OMS), éstas se diferencian en el porcentaje de blastos presentes en la sangre periférica o médula ósea, el porcentaje de basófilos, anomalías cromosómicas adicionales (segundo Ph1, trisomía 18, isocromosoma 17q, trisomía 19) y otras características clínico-patológicas como la esplenomegalia y la trombocitopenia [4]. De esta forma, la fase crónica se caracteriza por presentar <10% de blastos, mientras que la fase acelerada presenta entre 10 y 19% y la fase blástica $\geq 20\%$.

La fase crónica es la etapa inicial de la LMC, el 86% de los pacientes son diagnosticados en esta fase y por lo general, el tratamiento es más eficiente. La fase acelerada es una etapa intermedia caracterizada por presentar resistencia a la terapia y por el alto riesgo de evolución a la etapa final de la LMC, conocida como crisis blástica. Finalmente, la fase blástica se considera la forma más agresiva de la enfermedad y se caracteriza por el empeoramiento de los signos y síntomas clínicos, así como por un pronóstico incierto. Los mecanismos moleculares responsables de la evolución de la fase crónica a la crisis blástica se encuentran relacionados con la acumulación de alteraciones genéticas adicionales al gen de fusión *BCR-ABL1*.

En la actualidad, el tratamiento estándar para la LMC es el uso de inhibidores de tirosina cinasa (TKI, del inglés: *tyrosine kinase inhibitors*), dirigidos contra la oncoproteína BCR-ABL1, entre los cuales se destaca el mesilato de imatinib. El imatinib es una pequeña molécula sintética derivada de la 2-fenilaminopirimidina, la cual está diseñada para unirse a la hendidura catalítica de la oncoproteína quimérica BCR-ABL1. Al unirse al mismo sitio que el ATP, el imatinib estabiliza una conformación inactiva del sitio catalítico e inhibe de forma competitiva la actividad enzimática de la proteína. (**Figura 3**).

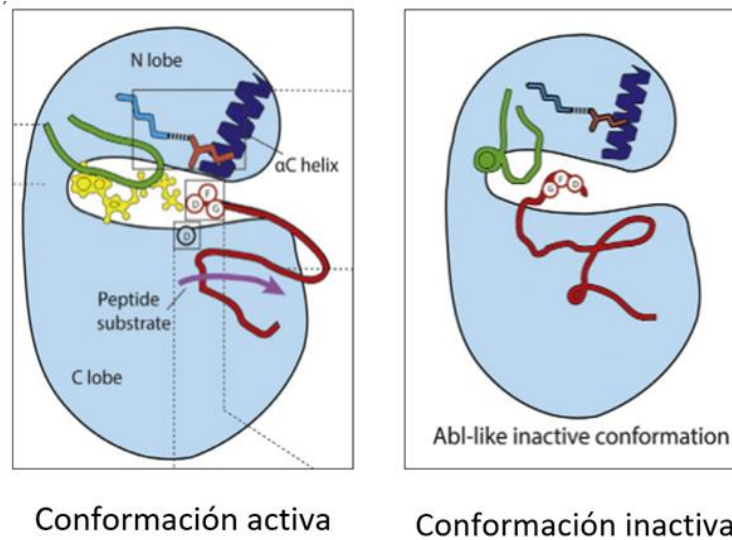


Figura 3. Mecanismo de acción del imatinib. El imatinib se une a cada uno de los sitios de unión al ATP cuando el bucle de activación de la cinasa se encuentra cerrado, estabilizando a la proteína y manteniéndola en la conformación inactiva. No obstante, la presencia de variantes genéticas somáticas puede mantener a *BCR-ABL1* en su conformación activa evitando que el imatinib se una a la enzima, reduciendo su afinidad.

La acción del imatinib conduce a una pérdida completa de la actividad transformadora del oncogén, con la consecuente eliminación de la célula leucémica. Se ha aprobado su uso para los pacientes con LMC Ph+ en fase crónica, acelerada o crisis blástica, dado que es un tratamiento oral eficaz, bien tolerado y tiene muy poca incidencia de efectos adversos graves. También se encuentra indicado para el tratamiento de otros cánceres como los tumores malignos del estroma gastrointestinal (GIST),

glioblastoma multiforme y leucemia linfoblástica aguda (LLA) Ph+. La efectividad del imatinib se basa en la tasa de respuesta hematológica y citogenética y en la supervivencia sin progresión de la enfermedad. Se ha demostrado que 9 de cada 10 pacientes tratados con imatinib han obtenido remisiones completas, sin progresión a fases aceleradas o blásticas y una considerable disminución en la tasa de mortalidad [5,6].

Sin embargo, se ha reportado el desarrollo de resistencia al imatinib y se ha observado que pacientes con LMC en fase crónica tratados con el fármaco tienen un riesgo de recaída de 10% por año. Algunas investigaciones han identificado como principales causas de resistencia a mecanismos dependientes de *BCR-ABL1* como su sobre-expresión o la adquisición de variantes genéticas somáticas en el dominio de cinasa. Por ejemplo, la variante genética en el residuo guardián Thr 315 (T315I), en la que se intercambia una treonina por isoleucina en la posición 315 de la proteína, dentro del dominio cinasa, afecta la unión del imatinib a la hendidura catalítica, con la consecuente pérdida de la función anti-leucémica del fármaco. También se han identificado mecanismos independientes a *BCR-ABL1*, como la aparición de otras alteraciones genéticas (e.g. variantes en *JAK2*, *cMYC*, *PP2A*), la sobre-expresión de genes resistentes a fármacos ligados al funcionamiento de las bombas de eflujo (e.g. *MDR*, *P-gp*, *LRP*), las cuales expulsan al fármaco del espacio intracelular, evitando que se produzca su efecto. Adicionalmente, se ha descrito la presencia de células quiescentes Ph+ protegidas contra imatinib y que pueden estar relacionadas con la recaída de la enfermedad.

La presencia de estos mecanismos de resistencia al imatinib ha llevado a la búsqueda de modificaciones en la molécula para productos que se superpongan a este fenómeno. Por ejemplo, fármacos moleculares de segunda generación como el dasatinib, nilotinib y bosutinib, además de fármacos de tercera generación como el ponatinib, que han generado resultados prometedores para los pacientes e incluso se ha visto reflejado en los porcentajes de supervivencia. No obstante, tanto la incidencia como la prevalencia de la enfermedad siguen aumentando a pesar del éxito del tratamiento y por tal motivo, las investigaciones se han centrado en desarrollar un sistema de diagnóstico temprano para la LMC que permita reducir la incidencia de la enfermedad. En este sentido, los microRNAs (miRNAs) han surgido como prometedores biomarcadores genéticos de la enfermedad. [7,8,9]

ÍNDICES PRONÓSTICOS

Previo al uso de los TKI como tratamiento de la LMC, los índices pronósticos eran la herramienta matemática más efectiva para predecir la gravedad de la enfermedad. A manera de determinar los tipos de respuesta de cada paciente al momento del diagnóstico, se establecieron distintos tipos de estos índices: SOKAL, HASFORD y EUTOS. El índice SOKAL, toma en cuenta la edad, el tamaño del bazo, el porcentaje de blastos y el número de plaquetas. Por otro lado, el índice HASFORD es un valor de pronóstico para la supervivencia de los pacientes que fueron tratados con interferón alfa, una molécula sintética diseñada para actuar como la sustancia producida de manera natural en el sistema inmunitario. El índice EUTOS resulta ser el más interesante, pues predice la respuesta citogenética completa meses después del tratamiento, lo que lo convierte en un predictor clave del curso de la enfermedad, permite definir si el paciente progresará a la fase blástica o acelerada. Para la interpretación de resultados se toman en cuenta tres tipos de riesgo: bajo, intermedio y alto, de acuerdo con los valores obtenidos se considera la supervivencia (**tabla 1**) [19, 20].

Tabla 1. Valores de referencia para los índices pronósticos.

Riesgo	Índice SOKAL	Índice HASFOR	Índice EUTOS	Supervivencia de 2 años
Bajo	<0.8	≤ 780	< 87	90%
Intermedio	0.8–1.2	781-1480	-	65-90%
Alto	>1.2	≥1480	> 87	65%

MicroRNAs

Los miRNAs son RNAs pequeños no codificantes, que regulan post-transcripcionalmente la expresión génica al unirse a secuencias complementarias en la región 3' no traducida de los RNA mensajeros (RNAm). Se han registrado más de 24,000 miRNAs en la base de datos miRbase y se ha identificado que estas moléculas regulan aproximadamente el 30% de los genes codificantes de proteínas.

La biogénesis de los miRNAs es un proceso altamente regulado que implica la generación de un transcrito primario (pri-miARN), por medio de la ARN polimerasa II o III y su posterior procesamiento a pre-miARN a través del complejo microprocesador compuesto por *DROSHA* (ribonucleasa de tipo III) y *DGCR8* (subunidad compleja del microprocesador). Posteriormente, los pre-miARN son exportados al citoplasma por la proteína exportina-5, donde son escindidos por *DICER1* (ribonucleasa de tipo III), lo que permite su reconocimiento por el complejo multiproteico RISC (complejo de silenciamiento inducido por ARN). El RISC utiliza la hebra guía del miRNA para interactuar con las regiones 3' no traducidas de los ARNm blanco a través de complementariedad de bases, lo que le permite inhibir la traducción o promover la degradación del RNAm objetivo, dependiendo del grado de complementariedad entre éstos (**figura 4**). El complejo multiproteico RISC involucra proteínas como GEMIN3-4, y AGO1-4. [10,11]

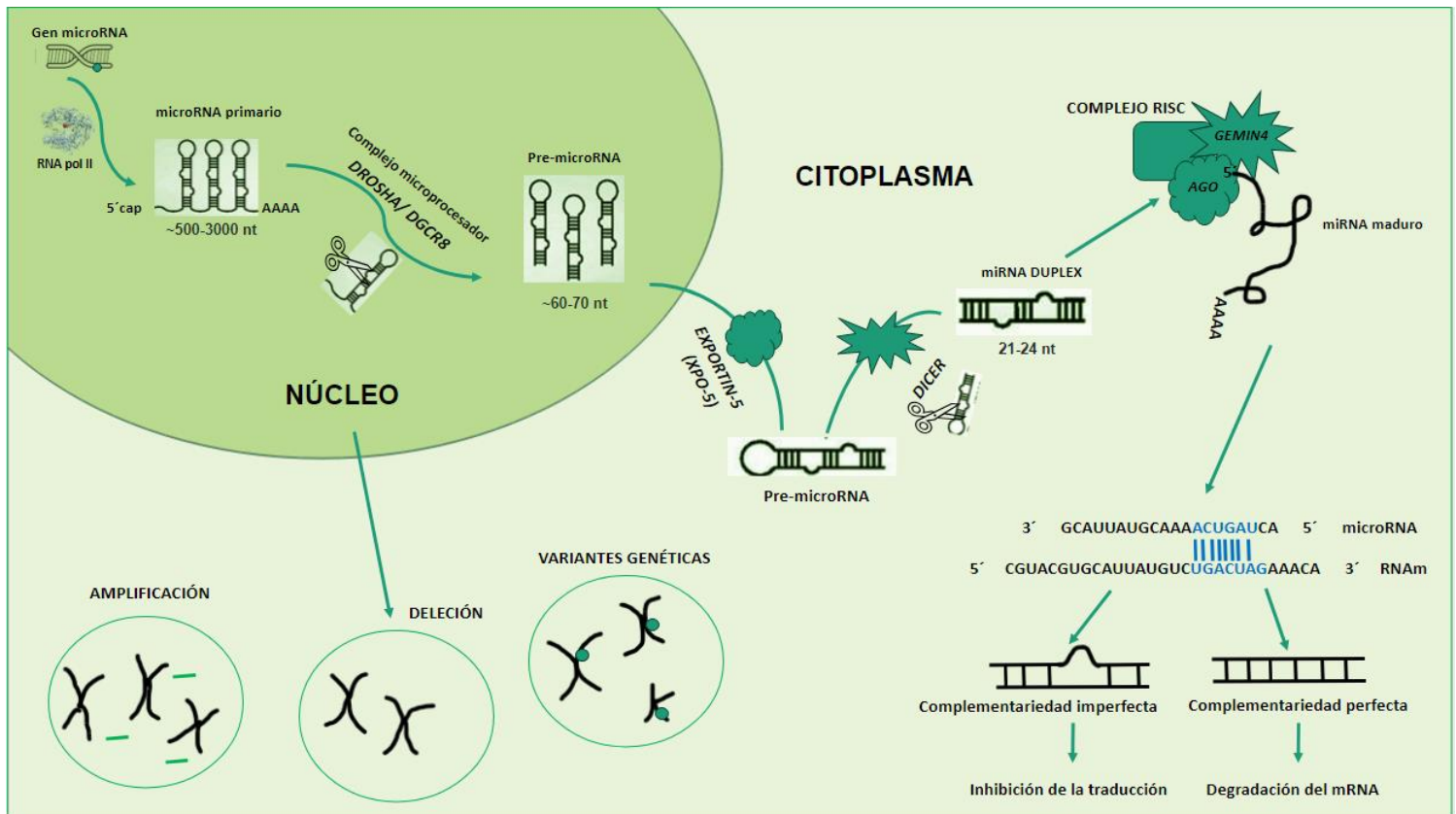


Figura 4. Sistema de la biogénesis de los miRNAs. Los miRNAs inicialmente son procesados en el núcleo como transcritos primarios de aproximadamente 500 a 3000 nucleótidos (nt), por el complejo proteico formado por DROSHA y DGCR8, para dar lugar a un pre-microRNA de 60 a 70 nt. La encargada de transportar estos pre-microRNAs al citoplasma es la proteína EXPORTIN-5 que junto con la ribonucleasa DICER1 dan origen a un miRNA duplex de 21 a 24 nt. Finalmente, este miRNA es reconocido por el complejo multiproteico RISC, causando la represión de la transcripción del RNAm o el silenciamiento genético por medio del apareamiento con su RNAm blanco.

Recientemente se ha demostrado la alteración de la regulación de los miRNAs como parte fundamental en la progresión de la LMC y el desarrollo de resistencia a TKIs. De hecho, múltiples estudios han explicado las modificaciones epigenéticas que contribuyen al silenciamiento de los miRNAs supresores de tumores y han demostrado su implicación en la LMC. Por ejemplo, los miR-19b y miR-23a, que tienen como blanco al transcrito de *BCR-ABL1* se encuentran hipermetilados en la LMC, provocando la proliferación de las células leucémicas. Incluso se ha reconocido que la supresión de miR-19b se encuentra relacionada con la transición de la fase crónica a la crisis blástica. Por otro lado, los miRNAs miR-221, miR-153-3p y miR-199a/b5p presentan una reducción en su expresión en pacientes que adquieren resistencia al imatinib. Estos miRNAs regulan un gran número de genes involucrados en múltiples vías de señalización celular, por lo que la alteración de su expresión podría aumentar la supervivencia y disminuir la apoptosis inducida por el fármaco.

Otro miRNA relacionado con la resistencia al tratamiento es el miR-214, el cual se ha encontrado disminuido en células mononucleares de sangre periférica en pacientes con LMC resistentes al imatinib. Importantly, miR-214 se une a la región 3' UTR del gen *ABCB1*, el cual codifica para una proteína con función de bomba de exclusión de fármacos dependiente de ATP. La proteína *ABCB1* es uno de los principales responsables de la disminución de la acumulación de fármacos en las células multiresistentes y, a menudo, interviene en el desarrollo de resistencia a los fármacos antineoplásicos. La expresión exógena de miR-214 en líneas celulares resistentes al imatinib, restableció la respuesta

a la terapia, lo que lo convierte en una herramienta prometedora para tratar la resistencia a los TKI en la LMC.

Dado que los miRNAs pueden funcionar como oncogenes o genes supresores de tumores en la leucemogénesis, el potencial de identificarlos como marcadores de riesgo genético permite nuevas oportunidades para el diagnóstico temprano y la prevención de LMC. [11,12] Además, se ha determinado que la presencia de variantes en los genes implicados en la maquinaria de la biogénesis de los miRNAs, provoca la pérdida o ganancia de función de éstos, lo que puede modificar la expresión de cientos de genes en diferentes vías biológicas. Por lo tanto, cualquier alteración en la regulación de la biosíntesis de los miRNAs, podría conducir a la aparición y progresión de la LMC. [11,12,13]

Las variantes genéticas son alteraciones en la secuencia del DNA, en la mayoría de éstas subyacen rasgos adaptativos y se han convertido en variantes comunes en los distintos grupos poblacionales, en los que se sabe, han conferido una ventaja selectiva. En los últimos años, el conocimiento sobre la variación genética humana se ha enriquecido enormemente, lo que ha permitido entender de qué manera los distintos grupos humanos están relacionados entre sí. Adicionalmente, la presencia de la variación genética se ha relacionado con la susceptibilidad o incluso la resistencia a diferentes enfermedades.

Las variantes génicas pueden surgir por diferentes razones, como son: errores en la replicación del DNA (tautomería, variantes puntuales, cambios en el marco de lectura), en el reparto de cromosomas durante la división celular (aberraciones cromosómicas), por agentes mutagénicos que causan daños fortuitos al DNA (p.ej: despurinización, desaminación, daños oxidativos) o por elementos genéticos transponibles que cambian de posición dentro del genoma provocando deleciones o inserciones. Dependiendo de su origen (germinales o somáticas), las variantes pueden transmitirse o no a la descendencia. De acuerdo con el cambio producido por la presencia de la variante, éstas pueden tener algún tipo de impacto biológico y por lo tanto desempeñan un papel en el desarrollo de alguna enfermedad [14].

Existen distintos tipos de variantes génicas y pueden clasificarse de la siguiente forma: secuencias repetidas en tándem dentro de los cuales entran los minisatélites (VNTR, *del inglés: variable number of tandem repeats*) y los microsatélites (STR, *del inglés: short tandem repeats*), inserciones o deleciones también conocidas como INDEL (*del inglés: insertion/deletion*), Variantes en el número de copias de una región genómica (CNV, *del inglés: copy-number variations*) y las variantes de un sólo nucleótido (SNV, *del inglés: single nucleotide variant*) [14,15].

Las SNVs son las variantes génicas más abundantes en el genoma humano, de las cuales se conocen alrededor de 15 millones de SNVs. Las SNVs pueden localizarse tanto en regiones génicas como en extragénicas y las que ocurren dentro de un gen pueden localizarse en secuencias codificantes o no codificantes. Dependiendo de su localización, las SNVs pueden afectar la regulación de los genes. Cuando ocurren en las regiones codificantes, las SNVs pueden cambiar la secuencia del DNA sin alterar el aminoácido codificado; la cual se conoce como variante sinónima; mientras que la presencia de otro tipo de SNVs, conocidas como no sinónimas, el cambio en la secuencia produce un cambio del aminoácido codificado. Por otro lado, en las regiones no codificantes, como los intrones o los promotores, pueden afectar la abundancia del RNAm.

En determinados casos, las variantes genéticas pueden potenciar la susceptibilidad a desarrollar alguna enfermedad. De hecho, diversos estudios han demostrado la implicación de algunas SNVs con distintas enfermedades humanas. Por ejemplo, la variante rs1800629, ubicada en el gen que codifica

para el factor de necrosis tumoral (TNF, *del inglés: tumor necrosis factor*) como se ha visto asociada con la aparición de distintas enfermedades como cáncer, asma, lepra y obesidad [15]. Por otro lado, SNVs como la rs1799945 y rs80338889 se han relacionado con hemocromatosis de tipo 1 y 3, respectivamente [16]. Algunas variantes son exclusivas de algunas poblaciones humanas, como la rs9282541, localizada en el gen *ABCA1* presente en poblaciones latinoamericanas y se ha asociado con la susceptibilidad a diabetes tipo 2 en población mexicana [17].

Recientemente, estudios de asociación de casos y controles, se han centrado en investigar la influencia de las SNVs con la susceptibilidad a presentar infección por COVID-19. Entre los hallazgos obtenidos por los investigadores, se encontró la asociación de la variante rs429358 con un mayor riesgo de infección. Además, en estudios contemporáneos, se ha demostrado que esta variante es un factor de riesgo para la enfermedad de Alzheimer[18].

Antecedentes

Una de las proteínas más estrechamente relacionada con la biogénesis de los miRNAs y otros RNAs no codificantes es la codificada por el gen *GEMIN4*. Este gen se localiza en el cromosoma 17 en la posición p13.3 y presenta diversas variantes génicas asociadas con la susceptibilidad genética a desarrollar enfermedades humanas, entre ellas rs7813, rs2740349, rs4968104. En 2017, Wu y colaboradores realizaron un metaanálisis para analizar la asociación de las SNVs rs2740348 y rs7813 del gen *GEMIN4* con un mayor riesgo a padecer cáncer y demostraron que pueden ser nuevos biomarcadores de predicción de la enfermedad. En un estudio independiente realizado por Mullany *et al.*, donde se investigó el impacto de algunas variantes localizadas en los genes implicados en la biogénesis de los miRNAs sobre la susceptibilidad genética a desarrollar cáncer de pulmón, identificaron que tres SNVs, entre ellas la rs2740349 del gen *GEMIN4*, se asociaron significativamente con niveles de expresión de RNAm alterados entre genotipos y que además esta misma variante, se encontraba relacionada con la regulación positiva de seis miRNAs individuales en la mucosa colónica normal. Por lo que la identificaron como una “variante no disruptiva” que podría cambiar la eficacia de la proteína. En otro estudio, llevado a cabo por Liu *et al.*, se identificaron diversas variantes genéticas comunes responsables de la susceptibilidad al cáncer de próstata, entre las que se identificó a la variante rs4968104 como una variante de susceptibilidad moderada. [21, 22, 23]

Justificación

La LMC es el trastorno mieloproliferativo más común en población adulta y es una de las neoplasias más estudiadas. A pesar de que los TKIs han demostrado ser una terapia altamente prometedora y funcional, la resistencia a la terapia se ha convertido en un problema que retrasa su efectividad. Adicionalmente, la incidencia de esta enfermedad se mantiene sin disminuir en los últimos años. En la actualidad, la presencia de SNVs en los componentes de la biosíntesis de los miRNAs como *GEMIN4* se ha asociado con el incremento en la susceptibilidad a desarrollar diferentes neoplasias humanas. Por lo tanto, la importancia de identificar las SNVs presentes en los componentes de la biosíntesis de los miRNAs, asociadas con la LMC radica en permitir un diagnóstico más preciso e incluso, una posible prevención de la enfermedad. Por ello, el presente estudio tiene como objetivo investigar las variantes rs7813, rs2740349 y rs4968104 del gen *GEMIN4*, para identificarlas como posibles biomarcadores de susceptibilidad genética para el desarrollo de LMC. Las variantes fueron seleccionadas con base a la literatura y se consideró como criterio de inclusión su frecuencia en población mexicana (mayor al 5%) y su asociación previa con otras patologías.

De acuerdo con la **tabla 2**, se trata de SNVs no sinónimas. Este tipo de variante puede ser patogénica, dependiendo del cambio que le produzca a la proteína resultante, es decir, si cambia o no su conformación.

Tabla 2. Características de las variantes

Gen	Posición del cromosoma (GRCh38)	ID SNV	Tipo de variante	Alelos
<i>GEMIN4</i>	17:744946	rs7813	No sinónima	A/G
	17:745258	rs2740349	No sinónima	T/C
	17:746265	rs4968104	No sinónima	T/A

Objetivo general

Evaluar la asociación de las SNVs rs7813, rs2740349 y rs4968104 localizadas en el gen *GEMIN4* con la susceptibilidad a desarrollar leucemia mieloide crónica en población mexicana.

Objetivos específicos

1. Determinar las frecuencias alélicas y genotípicas de las SNVs rs7813, rs2740349 y rs4968104 del gen *GEMIN4* en la población de estudio.
2. Comparar la frecuencia presentada por las variantes rs7813, rs2740349 y rs4968104 en el grupo de controles de la población de estudio con la reportada en otros grupos poblacionales.
3. Evaluar la asociación de las variantes analizadas con el riesgo a desarrollar leucemia mieloide crónica.
4. Analizar la asociación de los haplotipos formados por las SNVs rs7813, rs2740349 y rs4968104 con la presencia de la enfermedad.

Metodología

Población de estudio

Se llevó a cabo un estudio retrospectivo de casos y controles. La población de estudio estuvo compuesta por un total de 843 individuos de ambos sexos de la población mexicana entre casos y controles. El grupo de los casos estuvo conformado por 318 individuos diagnosticados previamente con LMC (139 mujeres y 179 hombres) y el de los controles por 525 participantes sanos (253 mujeres y 272 hombres). El DNA genómico necesario para este estudio se obtuvo de la genoteca del laboratorio de oncogenómica del Instituto Nacional de Medicina Genómica (INMEGEN).

Para la selección de la población de estudio, se tomaron en cuenta los siguientes criterios de inclusión y exclusión: (**tabla 2**)

Tabla 2. Criterios de inclusión y exclusión para casos y controles de la población de estudio.

	Criterios de inclusión	Criterios de exclusión
Casos	<ul style="list-style-type: none">• Edad: mayores a 18 años• Ambos sexos• Positivos para el cromosoma Ph	Negativos para el cromosoma Ph
Controles	<ul style="list-style-type: none">• Edad: 40 a 65 años• Ambos sexos	Presencia de alguna enfermedad concomitante

Preparación de muestras de DNA genómico

La extracción de las muestras de DNA necesarias para la realización de este estudio se realizó previamente a partir de muestras de sangre periférica de cada uno de los participantes. La extracción de DNA se trata de un paso imprescindible para llevar a cabo distintas técnicas de biología molecular y consta de los siguientes pasos [25,26]:

1. Disgregación celular. Alteración suficiente de las estructuras tisulares y celulares para crear un lisado.
2. Desnaturalización de complejos de nucleoproteínas.
3. Inactivación de nucleasas.
4. Lisis celular. Ruptura del núcleo celular para liberar el material genético mediante el uso de detergentes caotrópicos y algunas enzimas.
5. Extracción. Separación del material genético de otras biomoléculas y restos de componentes celulares.
6. Purificación. Eliminación de los contaminantes, usualmente por precipitación.
7. Precipitación del DNA con etanol para eliminar el RNA.
8. Solubilización y evaluación. El producto obtenido se solubiliza y evalúa en cuanto a pureza e integridad.

Una vez seleccionadas las muestras a utilizar en el estudio, éstas fueron cuantificadas por espectrofotometría, usando el equipo IMPLEN NanoPhotometer®, que se fundamenta en la ley de Lambert-Beer. Se trata de un principio utilizado para expresar de qué modo la materia absorbe la luz y relaciona esta absorción con las propiedades del material analizado, como su concentración y la longitud de la trayectoria del haz de luz que atraviesa la muestra. Los ácidos nucleicos se leen a una longitud de onda de 260 nm, debida específicamente a los enlaces presentes en los anillos aromáticos de los desoxirribonucleótidos (adenina, timina, citosina y guanina) [26].

La pureza de las muestras también fue evaluada por espectrofotometría. En este sentido, las proteínas y otros contaminantes, como sales, fenol o hidratos de carbono son absorbidas a 280nm y a 230nm, respectivamente. De esta forma, se considera que una muestra de DNA presenta un grado de pureza óptima cuando tiene una relación A260/280 entre 1.8 a 2.0. El DNA de pureza aceptable es de 1.6 a 2.0[26].

Por lo tanto, en este estudio se incluyeron las muestras de DNA con concentraciones de 100 ng/uL, con una pureza no menor a 1.6.

Posteriormente, se evaluó la integridad de las muestras por medio de la técnica de electroforesis en geles de agarosa al 1%. Esta técnica se basa en el uso de geles de agarosa como matrices sólidas, donde por medio de un campo eléctrico en un amortiguador salino se separan las moléculas de DNA. Para llevar a cabo este procedimiento, se utiliza un amortiguador de carga, uno de corrida y un colorante fluorescente (GelRed®). El amortiguador de carga contiene glicerol y éste se encuentra disuelto en una solución de tris(hidroximetil)aminometano mejor conocida como tris. Este amortiguador tiene como finalidad reducir la solubilidad del DNA en el amortiguador de corrimiento electroforético, lo que permite la correcta precipitación del DNA dentro de los pozos del gel de agarosa. El amortiguador de corrida puede ser TBE (Tris, ácido bórico y EDTA) o TAE (Tris, ácido acético glacial y EDTA). Comúnmente se utiliza el TBE por ser más estable, aunque también se considera al TAE para poder correr fragmentos más grandes de DNA. Su función es mantener el pH estable mientras se realiza el corrimiento y proporcionar un medio estable para la transmisión de corriente. El pH se mantiene estable porque el amortiguador contiene un ácido débil y el EDTA evita que las nucleasas degraden a los ácidos nucleicos. Por último, el GelRed® ayuda a visualizar el DNA en el gel al intercalarse entre la doble cadena de DNA y emitir una fluorescencia al ser excitado con luz ultravioleta. Se considera que una muestra de DNA presenta alta integridad cuando su perfil en una electroforesis se corresponde a una banda discreta que puede observarse en el gel, después de leerse en un fotodocumentador[26] (Figura 5).

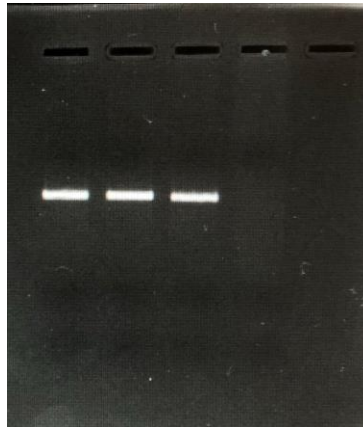


Figura 5. Evaluación de la integridad de las muestras de DNA usadas en este estudio. Se muestra un gel de agarosa representativo en el que el DNA se observa en forma de bandas discretas.

Discriminación alélica

Una vez confirmada la concentración, pureza e integridad de las muestras, éstas fueron diluidas a una concentración de 10 ng/uL utilizando la fórmula $C_1V_1 = C_2V_2$ con la finalidad de realizar el ensayo de discriminación alélica (taqman), de acuerdo con el manual de procedimientos técnicos del INMEGEN.[26]

Descripción de los ensayos TaqMan

La sonda TaqMan se basa en la actividad exonucleasa de la Taq polimerasa para escindir específicamente una sonda marcada ya hibridada a la secuencia diana, al igual que en otros métodos

de PCR cuantitativa, la escisión de la sonda emite fluorescencia, lo que permite obtener cuantitativamente la acumulación del producto genético durante los ciclos programados de la PCR.

En la PCR convencional no se utiliza una sonda diseñada específicamente para reconocer la SNV. Esta sonda es una secuencia de ácidos nucleicos, que hibrida sólo si la base complementaria es perfecta entre ellos, de tal forma que la sonda es complementaria al alelo con la SNV. Usualmente, el proveedor diseña las sondas con un par de cebadores de PCR sin marcar y una sonda TaqMan con una etiqueta de colorante (FAM) en el extremo de 5' y un supresor no fluorescente (NFQ) en el extremo 3'. Generalmente se marca con FAM al alelo de menor frecuencia [25,26]. Los primers de amplificación, que van a delimitar la zona a amplificar, se encuentran disueltos en la sonda y deben ser previa y correctamente diseñados. Para su diseño es importante evitar incluir más de tres nucleótidos G o C en el extremo 3' a fin de reducir el riesgo de anillamientos inespecíficos. [25,26]

Procedimiento del ensayo TaqMan

Para el ensayo taqman, se utilizaron placas ópticas previamente preparadas con las 86 muestras de LMC a 10ng/uL y con los 10 controles negativos (NTC). Posteriormente, se agregó la siguiente mezcla de reacción que incluía [25]:

1. La mezcla comercial genotyping máster mix (TaqMan™) que contiene los siguientes reactivos [25]:
 - Los cuatro desoxirribonucleótidos trifosfatos (dNTPs), que son los “ladrillos” necesarios para la construcción de la nueva cadena.
 - Los iones divalentes y monovalentes, por ejemplo, el Mg^{2+} , que sirven como cofactor para el correcto funcionamiento de la taq polimerasa.
 - La taq polimerasa, con temperatura óptima de 72°C.
2. La sonda correspondiente (rs 7813, rs2740349 o rs4968104)
3. H₂O grado biología molecular

Con las cantidades indicadas en la **tabla 3**

Tabla 3. Protocolo del método TaqMan

	1x (uL)	110x (uL)
Genotyping máster mix (TaqMan™)	2.5	275
Sonda (rs7813, rs2740349 o rs4968104)	2.41	265
H ₂ O grado biología molecular	0.09	9.9

Fue importante asegurarse de que la sonda estuviera descongelada y evitar, en medida de lo posible, que incidiera la luz sobre ella. Una vez preparada la mezcla de reacción, se agregaron 5 uL de ella a cada uno de los 96 pozos de la placa. Posteriormente, las placas se sellaron con un cover óptico, se centrifugaron, rectificando la eliminación de las burbujas, dado que pueden interferir con la amplificación de las muestras. Finalmente, la reacción de amplificación fue llevada a cabo en un termociclador GeneAmp PCR System 9700 con un *software* versión 3.12, siguiendo el siguiente protocolo de amplificación para discriminación alélica con sondas TaqMan (figura 5) [25, 26]:

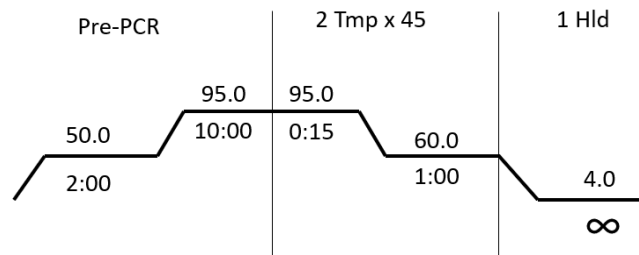


Figura 5. Protocolo de termociclado del método Taqman.

El protocolo de termociclado para discriminación alélica se diferencia de las del método de PCR convencional por el tipo de reacción que se lleva a cabo. El método taqman se basa en el uso exclusivo de dos etapas de desnaturalización y una de elongación. A diferencia de la PCR convencional, que consta de tres etapas, dado que ésta se basa en el uso de primers, es importante que pasen por una etapa de hibridación para evitar uniones inespecíficas. Por el contrario, como el método taqman no sólo utiliza primers sino también sondas, no se requiere un paso de hibridación porque aunque existan alineamientos inespecíficos, estos no serán reportados por el equipo de amplificación [25].

Al finalizar el termociclado, se hizo la lectura de la placa en el equipo QuantStudio® 3 Real-Time PCR System y los resultados se analizaron en el programa de cómputo Design & Analysis 2.6.0[25,26].

El ensayo de discriminación alélica clasifica las muestras como:

- Homocigotos: muestras que tienen solo el alelo silvestre o el alelo derivado. Que a su vez se dividen en homocigotos mayores (muestras homocigotas para el alelo mayor) y homocigotos menores (muestras homocigotas para el alelo menor)
- Heterocigotos: muestras que tienen un alelo silvestre y un alelo derivado.

Análisis estadístico

Para el cálculo de la asociación de las SNVs con la presencia de la LMC se usó la razón de posibilidades (OR) y el valor de P.

El OR es una razón que deriva del cálculo de la posibilidad (odds) de los casos dividido entre la posibilidad (odds) de los controles. Es decir, en un estudio de casos y controles lo que se obtiene es la razón de posibilidades del riesgo en los casos y en los controles. Los valores de OR se interpretan así[25,26]:

- $OR < 1$: la frecuencia de la variante es menor en los casos que en los controles, por lo tanto, es de protección.
- $OR = 1$: no hay asociación.
- $OR > 1$: la frecuencia de la variante es mayor en los casos que en los controles, por lo tanto, es de riesgo.

El valor de P se utiliza para saber qué tanto el valor de OR se debe a la probabilidad y la probabilidad es inversa a la significancia, es decir, a menor probabilidad de que un evento ocurra al azar, mayor significancia[25,26]. Por lo tanto, valores de P es menor a 0.05 serán considerados como significativos.

Equilibrio de Hardy-Weinberg (EHW)

El EHW establece que, en una población suficientemente grande, en la que existe panmixia (selección de pareja al azar) y que no se encuentra sometida a presión de selección o migración, la frecuencia de las variantes génicas de una población se mantiene constantes de una generación a otra. Se trata de un principio básico de genética de poblaciones y el filtro más importante por el que debe pasar todo análisis estadístico en un estudio genético de casos y controles. Se considera que la población de estudio se encuentra en equilibrio cuando el valor de P es mayor o igual a 0.05[27].

El presente proyecto fue aceptado por el comité de ética del Instituto Nacional de Medicina Genómica (INMEGEN)

Resultados

Como se puede observar en la **tabla 3**, al evaluar el EHW de cada una de las SNVs estudiadas, la variante rs7813 se encontró en EHW, tanto en los casos como en los controles; mientras que, la variante rs2740349 presentó una desviación significativa del EHW en los casos ($P < 0.05$) pero no en los controles. Finalmente, la variante rs4968104 presentó una desviación significativa en los controles ($P < 0.05$) pero no en los casos. Cuando en la población de afectados la frecuencia de la variante se desvía de forma significativa del EHW, esto puede deberse a que la presencia de la enfermedad puede enriquecer la frecuencia de la variante. Sin embargo, cuando la frecuencia de la variante se desvía de forma significativa del EHW en el grupo de individuos sanos, la presencia de algún factor externo a la enfermedad que modifique la frecuencia de la variante deberá ser tomado en cuenta. Considerando que la SNV rs4968104 es de baja frecuencia, el tamaño de nuestra población de estudio podría estar interfiriendo con el resultado. Sin embargo, es importante mencionar que se hizo un análisis individual sólo con la población de mujeres, y se encontró que la variante rs4968104 sí se encuentra en EHW (**Tabla 4**).

Tabla 3. Resultados del equilibrio de Hardy-Weinberg (EHW).

CHR	SNV	Prueba	A2	A1	GENO	O (HET)	E (HET)	P
17	rs7813	Casos	G	A	35/128/172	0.3821	0.4164	0.1482
17	rs7813	Controles	G	A	44/213/261	0.4112	0.4123	1
17	rs2740349	Casos	C	T	24/79/219	0.2453	0.318	0.0001455
17	rs2740349	Controles	C	T	22/161/332	0.3126	0.3166	0.6782
17	rs4968104	Casos	A	T	6/58/268	0.1747	0.1886	0.2346
17	rs4968104	Controles	A	T	9/77/437	0.1472	0.1651	0.02766

Tabla 4. Resultados del equilibrio de Hardy-Weinberg en la población de mujeres.

CHR	SNV	Prueba	A2	A1	GENO	O (HET)	E (HET)	P
17	rs7813	Casos	G	A	14/50/75	0.3597	0.4057	0.2086
17	rs7813	Controles	G	A	44/213/261	0.4315	0.4138	0.5431
17	rs2740349	Casos	C	T	10/31/93	0.2313	0.3082	0.008
17	rs2740349	Controles	C	T	9/82/159	0.328	0.32	0.8437
17	rs4968104	Casos	A	T	2/22/114	0.1594	0.1707	0.3353
17	rs4968104	Controles	A	T	4/37/211	0.1468	0.1626	0.119

Como se puede observar en la **tabla 5**, el rango de edad para los pacientes con LMC osciló entre los 14 y 92 años, sin embargo, es importante considerar que por los criterios de inclusión (mayores a 18 años) se tomó en cuenta este rango por ser la edad al momento del diagnóstico. Con respecto al género, se puede observar que son más los hombres quienes presentan la enfermedad, a pesar de que aún se desconoce el motivo, esto concuerda con otros estudios en los que se ha demostrado que son los varones quienes se ven más afectados. En cuanto al reporte de los índices pronósticos, sólo 115 de los participantes cuentan con reporte del índice Sokal, mientras que el índice Hasford sólo se ve reportado en 117 y el índice Eutos en 101 pacientes con LMC.

TABLA 5. Características clínico-demográficas de la población de estudio

	Casos (pacientes con LMC) n=525	Controles n= 318	P
Edad en años	Rango: 14-92* Mediana: 44*	Rango: 34-96 Mediana: 48	5.462e-05
Sexo	139 mujeres 179 hombres	253 mujeres 272 hombres	Mujeres Hombres
Índice SOKAL	33 pacientes con riesgo bajo 33 pacientes con riesgo intermedio 49 pacientes con riesgo alto 203 pacientes sin reporte de índice SOKAL	NA	-
Índice HASFORD	30 pacientes con riesgo bajo 38 pacientes con riesgo intermedio 49 pacientes con riesgo alto 201 pacientes sin reporte de índice HASFORD	NA	-
Índice EUTOS	53 pacientes con riesgo bajo 48 pacientes con riesgo alto 217 pacientes sin reporte de índice EUTOS	NA	-

NA: no aplica, *edad al momento del diagnóstico

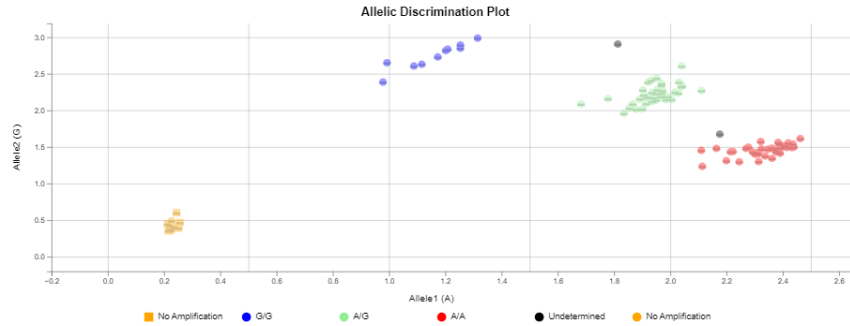
Ensayo de la actividad de exonucleasa de la polimerasa para discriminación alélica (Sondas TaqMan para discriminación alélica).

La genotipificación de las variantes seleccionadas en la población de estudio fue realizada por medio de sondas TaqMan para discriminación alélica. En el caso de la variante rs7813 A/G, el alelo menor (G) se encuentra en el eje Y, mientras que el alelo mayor (A) en el eje X. Aunque la sonda para esta

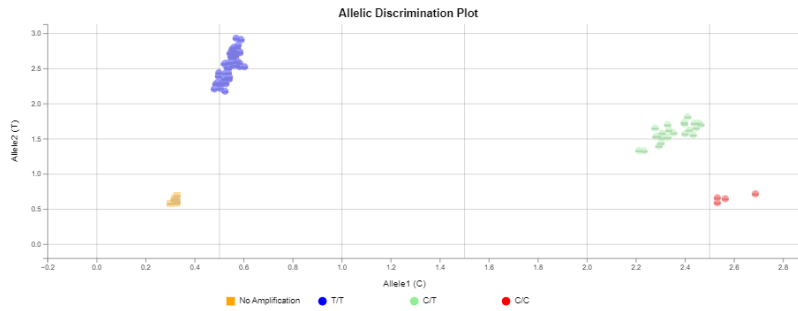
variante presentó problemas para la discriminación alélica, dado que los homocigotos mayores no se agrupan correctamente en su eje y se acercan demasiado a los heterocigotos (**figura 6 panel a**).

La variante rs2740349 T/C muestra al alelo menor (C) en el eje X y al alelo mayor (T) en el eje Y. Para las variantes rs4968104 T/A el alelo menor (A) se encuentra en el eje X y el alelo mayor (T) en el eje Y. La sonda para estas dos variantes funcionó perfectamente, se observa la agrupación correcta de cada una de las nubes en su eje correspondiente (**figura 6 panel b y c**).

a) rs7813 A/G



b) rs2740349 T/C



c) rs4968104 T/A

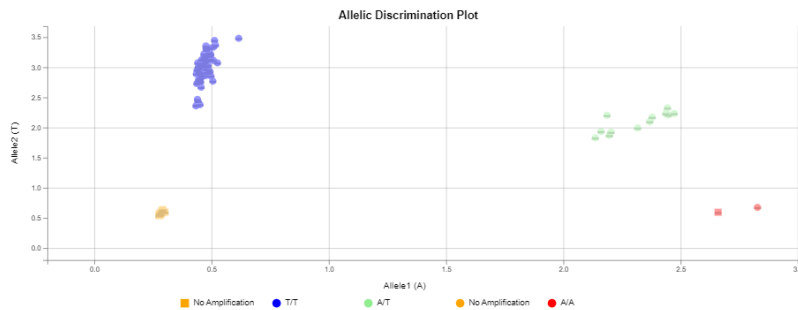


Figura 6. Genotipificación de las variantes estudiadas.

En la **figura 6: a), b) y c)**, se muestran las gráficas representativas de los resultados de discriminación alélica.

En la **tabla 6** se encuentra reportada la frecuencia del alelo menor (MAF, por sus siglas en inglés), de cada una de las variantes analizadas en la población de estudio y las reportadas en distintos grupos humanos. De las tres variantes analizadas, la rs7813 mostró la mayor frecuencia en la población de estudio (0.29), seguida por la rs2740349 (0.19) y la rs4968104 (0.09). En el caso de la SNV rs7813,

se puede observar que no hay diferencia en cuanto a la frecuencia de la población de estudio con la global y la población africana es la única que presenta una frecuencia significativamente menor para esta SNV. La segunda frecuencia más baja es la de la población de estudio (MX) para la variante rs4968104. En realidad, la relevancia de esta comparación radica en conocer si existe flujo genético entre las poblaciones o si se trata de algún otro factor como un posible efecto fundador o incluso deriva génica. Sin embargo y de acuerdo con la tabla, para las variantes analizadas no se reporta ninguno de estos factores.

Tabla 6. Comparación de las frecuencias presentadas por las variantes rs7813, rs2740349 y rs4968104 en la población de estudio con la reportada en otros grupos poblacionales.

Población	rs7813	rs2740349	rs4968104
Global	0.29 (G)	0.11 (C)	0.16 (A)
Africana (AFR)	0.12 (G)	0.03 (C)	0.05 (A)
Americana (AMR)	0.36 (G)	0.18 (C)	0.18 (A)
Asiática oriental (EAS)	0.28 (G)	0.11 (C)	0.12 (A)
Asiática del sur (SAS)	0.37 (G)	0.12 (C)	0.23 (A)
Europea (EUR)	0.41 (G)	0.16 (C)	0.25 (A)
Población de estudio (MX)	0.29 (G)	0.19 (C)	0.09 (A)

Para el cálculo de la asociación de las SNVs con la presencia de la LMC se usó la razón de posibilidades (OR) y el valor de P. Sin embargo, los valores obtenidos (**Tabla 7**) son muy cercanos o iguales a 1, por lo que se deduce que no hay asociación.

Tabla 7. Asociación de las variantes analizadas con el riesgo a desarrollar leucemia mieloide crónica.

SNV	BP	A2	A1	MAF en pacientes	MAF en controles	χ^2	P	OR
rs7813	744946	A	G	0.2855	0.2905	0.04888	0.825	0.9757
rs2740349	745258	T	C	0.1922	0.199	0.1142	0.7354	0.9574
rs4968104	746265	T	A	0.1003	0.09082	0.414	0.5199	1.116

SNV: variante de un sólo nucleótido, **MAF:** frecuencia del alelo menor.

Tabla 8. Resultados de asociación por modelos de herencia

SNV	A1	A2	Análisis	F_A	F_U	P
rs7813	G	A	DOM	151/166	257/261	0.5786
rs7813	G	A	REC	30/287	44/474	0.6324
rs2740349	C	T	DOM	97/210	183/332	0.2492
rs2740349	C	T	REC	21/286	22/493	0.1096
rs4968104	A	T	DOM	58/256	86/437	0.4517
rs4968104	A	T	REC	5/309	9/514	0.8884

Los modelos de herencia son patrones hereditarios que en genética médica se utilizan para conocer de qué manera se va a comportar la enfermedad, es decir, si se trata de un patrón hereditario autosómico dominante, en el cual una copia alterada del gen es suficiente para que un individuo se vea afectado por un trastorno o enfermedad. En el caso de los patrones hereditarios autosómicos recesivos, las variantes deben ocurrir en ambas copias del gen, para que se presente la enfermedad. De esta manera, en los resultados del análisis por asociación en modelos de herencia (**tabla 8**), podemos conocer si la asociación de las variantes estudiadas con la LMC ocurre de forma dominante o recesiva. No obstante, los resultados no fueron significativos, por lo que en este caso no podemos conocer el comportamiento hereditario de la enfermedad.

Tabla 9. Resultados del análisis de asociación por edad en hombres y mujeres

HOMBRES			MUJERES	
SNV	BETA	P	BETA	P
rs7813	-2.302	0.2348	-1.689	0.418
rs2740349	-1.308	0.5358	2.762	0.2333
rs4968104	-2.857	0.3254	-8.852	0.00799

Adicionalmente, se realizó un análisis de asociación por edad en el grupo de los casos (hombres y mujeres), lo que podría ayudar a definir si los individuos que presentan el alelo de riesgo presentan síntomas antes que los individuos que no lo portan. Para ello, se clasificó la edad al diagnóstico de los pacientes con respecto a la presencia del alelo menor de las variantes. Se encontró que los individuos portadores del alelo menor (A) de la variante rs4968104 se diagnosticaron, en promedio, 8.852 años antes que los individuos que no portan este alelo. Sin embargo, es importante señalar que esta variante se encontró en desequilibrio de HW en la población de estudio control ($P < 0.05$).

Tabla 10. Evaluación de la asociación por índices pronósticos (riesgo bajo y medio contra alto)

SOKAL (n=115)							
SNV	A2	A1	AFECTADOS	CONTROLES	χ^2	P	OR
rs7813	A	G	0.2930	0.3048	0.9013	0.3423	0.7559
rs2740349	T	C	0.4247	0.371	0.8966	0.3422	1.403
rs4968104	T	A	0.05632	0.08691	0.9091	0.3409	0.5900
HASFORD (n=117)							
rs7813	A	G	0.2865	0.3375	0.5543	0.4713	0.89
rs2740349	T	C	0.4890	0.3616	2.905	0.0921	1.665
rs4968104	T	A	0.07077	0.07690	0.0396	0.875	0.9
EUTOS (n=101)							
rs7813	A	G	0.2909	0.3780	0.8312	0.389	0.732
rs2740349	T	C	0.4404	0.4306	0.03015	0.8732	1.064
rs4968104	T	A	0.05686	0.09107	0.8312	0.3754	0.5954

Con la intención de conocer si la presencia de alguna de las variantes analizadas, de alguna manera genera cambios en los niveles de riesgo, se realizó una evaluación de los índices pronósticos en los pacientes con LMC (**tabla 10**). Sin embargo, los resultados obtenidos muestran que la presencia de las SNVs rs7813 y rs4968104 no influye en el pronóstico de la enfermedad. Sin embargo, respecto al índice Hasford, es importante destacar que se encontró que tener la variante rs2740349 se asocia, de manera marginal, con un peor desarrollo de la enfermedad.

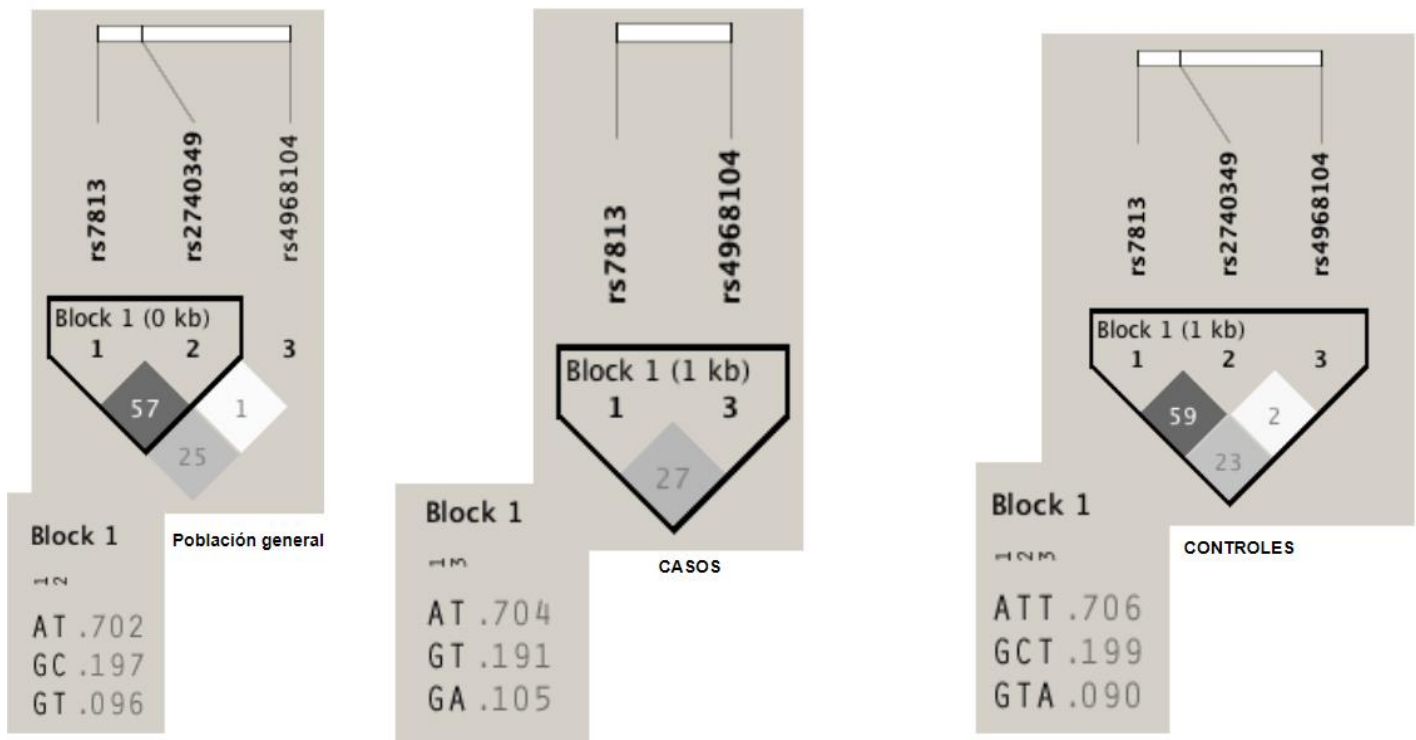


Figura 7. Asociación de los haplotipos formados por las SNVs rs7813, rs2740349, rs4968104 y la presencia de la enfermedad.

Con el propósito de conocer si las variantes analizadas en su conjunto se asocian con el desarrollo de la LMC, se realizó un análisis de haplotipos. Sin embargo, las variantes se encuentran en bajo desequilibrio de ligamiento, tanto en casos (r^2 menor a 0.27 para cualquier combinación) como en controles (r^2 menor a 0.59 para cualquier combinación), por lo que la probabilidad de que formen un bloque de haplotipos es muy baja.

Discusión

Los miRNAs son importantes reguladores en una cantidad considerable de procesos biológicos. Las SNVs presentes en genes involucrados en la biogénesis de los miRNAs, como es el caso del gen *GEMIN4*, influyen en su generación y maduración y luego en su expresión y estructura. El gen *GEMIN4* da lugar a una proteína con múltiples funciones como el transporte nuclear y citoplasmático de RNAm, también tiene participación en las vías del RNA de interferencia. La proteína GEMIN4 también se ha visto involucrada en el ensamblaje del complejo masivo de ribonucleoproteínas conocido como espliceosoma, que tiene como función la escisión de intrones (empalme), los cuales deben eliminarse con precisión para dar lugar a los RNAm maduros. El mal funcionamiento del espliceosoma producirá proteínas erróneas que den lugar a diferentes enfermedades humanas incluidas las neoplasias mieloides. Las variantes patogénicas en el gen *GEMIN4* se han visto involucradas con distintas enfermedades, incluidos los trastornos hereditarios caracterizados por problemas en el neurodesarrollo. Por su localización cromosómica (Chr17p.13.3), la cual se ha reportado alterada en varios carcinomas, el gen *GEMIN4* también se ha visto involucrado en el desarrollo del cáncer. Estudios recientes han demostrado la asociación de diversas variantes en *GEMIN4* con diferentes enfermedades humanas. El gen *GEMIN4* cuenta con alrededor de 100 variantes, hasta el momento,

las variantes rs7813, rs2740349 y rs4968104 son las más estudiadas. Se trata de variantes no sinónimas, que tienen la potencialidad de cambiar el aminoácido de la proteína GEMIN4, lo que podría intervenir en su funcionamiento.

Sin embargo, en la actualidad, no se cuenta con suficientes investigaciones enfocadas en las variantes de la biogénesis de los miRNAs y su relación con la LMC. Por lo tanto, el objetivo de la presente investigación fue la de identificar a las SNVs rs7813, rs2740349 y rs4968106 como posibles biomarcadores de predisposición genética a padecer LMC.

Para ello, se llevó a cabo un estudio de casos y controles en una población de estudio de 843 individuos de ambos sexos de la población mexicana y.

Aun cuando uno de los principales objetivos de nuestra investigación era encontrar asociaciones significativas de las variantes con la LMC, y a pesar de no obtener los resultados esperados, llevamos a cabo distintos análisis de asociación, por ejemplo, con respecto a la edad de diagnóstico, sexo de los participantes, índices pronósticos de riesgo, modelos de herencia y análisis de haplotipos, a manera obtener hallazgos interesantes que den paso a la implementación de nuevas investigaciones.

Para encontrar evidencia de la asociación de las variantes estudiadas con distintas enfermedades humanas, realizamos una revisión de la literatura y encontramos que la variante rs7813 se ha asociado con depresión, esquizofrenia, esclerosis múltiple, hepatitis B, tuberculosis y cáncer gástrico. En un estudio realizado por Basak J y colaboradores [32] se asoció el genotipo menor AA rs7813 con un mayor riesgo de presentar esclerosis múltiple. En otro estudio llevado a cabo por Yong él *et al.*, [33] se encontró que la variante rs7813 no afectó la susceptibilidad a la depresión y que, inclusive, podría tratarse de una variante de protección. Adicionalmente, una investigación realizada por Hong Cheng *et al.*, [34] demostró la asociación significativa del genotipo heterocigoto AG rs7813 con un mayor riesgo a desarrollar tuberculosis. Por otro lado, en el mismo estudio de Hong Chen *et al.*, [34] se encontró que la SNV rs2740349 puede regular la expresión del gen y se asoció con un efecto regulador en la tuberculosis. La SNV rs2740349 también se ha encontrado asociada con el riesgo a desarrollar cáncer de pulmón, de acuerdo con investigaciones de Mullany y Fang *et al.* [22] Las investigaciones sobre la variante rs4968104 se encuentran muy limitadas y no se ha encontrado asociada con una enfermedad distinta al cáncer, interesantemente los hallazgos de los investigadores la han posicionado como una variante de susceptibilidad moderada o incluso de protección.

En una investigación previa, sobre la asociación de la variante -653 G/A del gen *NRF2* con nefritis en el lupus eritematoso sistémico en la infancia, se encontró que esta variante en el gen *NRF2*, que se considera un gen antioxidante, no se asoció con la susceptibilidad genética a desarrollar la enfermedad, pero si a un peor pronóstico de ésta, debido a que, se demostró que las personas enfermas desarrollaron más frecuentemente nefritis cuando presentaban la variante que cuando no la tenían. Con la intención de obtener un hallazgo similar en nuestra investigación, llevamos a cabo la evaluación de la asociación por índices pronósticos. Aunque no encontramos asociaciones significativas con los índices pronósticos, observamos una asociación marginal entre la variante rs2740349 y un peor pronóstico de la enfermedad, de acuerdo con el índice Hasford.

Un resultado interesante en nuestra investigación se encontró en el análisis de asociación por edad en hombres y mujeres, donde se identificó que la edad al diagnóstico en mujeres portadoras del alelo menor (A) de la variante rs4968104 fue en promedio, 9 años antes que las mujeres que no portan este alelo. En un estudio previo realizado por Miedl H *et al.*, [22] en donde analizaron dos variantes (rs2046210 y rs9383590) ubicadas en el gen *ESR1*, que es un establecido biomarcador pronóstico

para el cáncer de mama, encontraron que los portadores del genotipo homocigoto menor para ambas SNVs, desarrollan la enfermedad varios años antes que los individuos que no portan los alelos menores. Por lo que, el alelo A de la variante rs4968104 se podría establecer como un posible marcador de riesgo genético para la LMC. Sin embargo, es importante destacar que esta variante no se encontró en equilibrio de HWE ($p < 0.05$) en el análisis estadístico. A pesar de que recientemente, algunos autores manejan el equilibrio de HWE a partir de 0.01[30], consideramos el hecho de que este evento limita la interpretación de este resultado, por lo que es necesario realizar nuevas investigaciones al respecto, tomando en cuenta un tamaño de muestra más considerable.

Conclusión

Centrar las investigaciones en el desarrollo de estrategias de prevención basadas en miRNAs como posibles biomarcadores de susceptibilidad genética, es sin duda, una nueva alternativa para intentar reducir la prevalencia e incidencia de enfermedades tan complejas como la LMC. A pesar de que aún queda mucho trabajo por hacer para mejorar la eficacia y principalmente la seguridad de estas nuevas estrategias, es un tipo de sistema que puede dirigirse directamente al origen de la enfermedad y de esta manera reducir las posibilidades de un efecto no deseado, a diferencia de algunos sistemas disponibles actualmente, que en su mayoría tienen un efecto citotóxico generalizado.

Referencias

- [1] Jayakumar, S., Merin, D., Krishna, A., A, M., & KN, A. (2022). Approaches to Managing Philadelphia Chromosome Positive Chronic Myeloid Leukaemia Patients: A Short Review. *Systematic Reviews in Pharmacy*, 13(2), 102–107. <https://doi.org/10.31858/0975-8453.13.2.102-107>
- [2] Sampaio MM, Santos MLC, Marques HS, Gonçalves VLS, Araújo GRL, Lopes LW, Apolonio JS, Silva CS, Santos LKS, Cuzzuol BR, Guimarães QES, Santos MN, de Brito BB, da Silva FAF, Oliveira MV, Souza CL, de Melo FF. Chronic myeloid leukemia-from the Philadelphia chromosome to specific target drugs: A literature review. *World J Clin Oncol* 2021; 12(2): 69-94 PMID: 33680875 DOI: 10.5306/wjco.v12.i2.69
- [3] Yaghmaie M, Ghaffari SH, Ghavamzadeh A, Alimoghaddam K, Jahani M, Mousavi SA, Irvani M, Bahar B, Bibordi I. 2008. Frequency of BCR-ABL fusion transcripts in Iranian patients with chronic myeloid leukemia. *Arch. Iran Med.* 11(3):247-251.
- [4] Shtalrid M, Talpaz M, Kurzrock R, Kantarjian H. 1988. Analysis of breakpoints within the BCR gene and their correlation with the clinical course of Philadelphia-positive chronic myelogenous leukemia. *Blood.* 72(2):485-490.
- [5] Hehlmann, Rüdiger. Chronic Myeloid Leukemia in 2020. *HemaSphere*: October 2020-Volume4-Issue5-pe468 doi: 10.1097/HS9.0000000000000468
- [6] Suttorp, M., Millot, F., Sembill, S., Deutsch, H., Metzler, M., Delord, M., & Etienne, G. (2021). Definition, Epidemiology, Pathophysiology, and Essential Criteria for Diagnosis of Pediatric Chronic Myeloid Leukemia. *Cancers*, 13(4), 798. <https://doi.org/10.3390/cancers13040798>
- [7] Del Castillo, R. Hernández, R. Zafra, G. (2019). *Genética clínica*, 2da edición. Ciudad de México, editorial: el manual moderno, S.A de C.V
- [6] Giles, F. (2009). Nilotinib: optimal therapy for patients with chronic myeloid leukemia and resistance or intolerance to imatinib. *Drug Design, Development and Therapy*, 89. doi:10.2147/dddt.s3069
- [7] Peng, Y.-H., Shiao, H.-Y., Tu, C.-H., Liu, P.-M., Hsu, J. T.-A., Amancha, P. K., ... Wu, S.-Y. (2013). Protein Kinase Inhibitor Design by Targeting the Asp-Phe-Gly (DFG) Motif: The Role of the DFG Motif in the Design of Epidermal Growth Factor Receptor Inhibitors. *Journal of Medicinal Chemistry*, 56(10), 3889–3903. doi:10.1021/jm400072p
- [8] Wang, Q., Zorn, J. A., & Kuriyan, J. (2014). A Structural Atlas of Kinases Inhibited by Clinically Approved Drugs. *Protein Kinase Inhibitors in Research and Medicine*, 23–67. doi:10.1016/b978-0-12-397918-6.00002-1
- [9] Weisberg, E., Manley, P., Mestan, J., Cowan-Jacob, S., Ray, A., & Griffin, J. D. (2006). AMN107 (nilotinib): a novel and selective inhibitor of BCR-ABL. *British Journal of Cancer*, 94(12), 1765–1769. doi:10.1038/sj.bjc.6603170
- [10] Verma A, Singh V, Jaiswal PK, Mittal RD. Anomalies in MiRNAs Machinery Gene, GEMIN-4 Variants Suggest Renal Cell Carcinoma Risk: A Small Experimental Study from North India. *Indian J Clin Biochem.* 2019 Jan;34(1):45-51. doi: 10.1007/s12291-017-0722-x. Epub 2018 Jan 10. PMID: 30728672; PMCID: PMC6346622.
- [11] Wu N, Zhang X, Tian J, Yu S, Qiao Y. Association of GEMIN4 gene polymorphism and the risk of cancer: a meta-analysis. *Onco Targets Ther.* 2017 Nov 2;10:5263-5271. doi: 10.2147/OTT.S147204. PMID: 29138579; PMCID: PMC5679687.
- [12] Mullany LE, Herrick JS, Wolff RK, Buas MF, Slattery ML. Impact of polymorphisms in microRNA biogenesis genes on colon cancer risk and microRNA expression levels: a population-based, case-control study. *BMC Med Genomics.* 2016 Apr 23;9(1):21. doi: 10.1186/s12920-016-0181-x. PMID: 27107574; PMCID: PMC4841949.
- [13] Litwińska Z, Machaliński B. miRNAs in chronic myeloid leukemia: small molecules, essential function. *Leuk Lymphoma.* 2017 Jun;58(6):1297-1305. doi: 10.1080/10428194.2016.1243676. Epub 2016 Oct 13. PMID: 27736267.
- [14] Fonseca-Montaño MA, Blancas S, Herrera-Montalvo LA, Hidalgo-Miranda A. Cancer Genomics. *Arch Med Res.* 2022 Dec;53(8):723-731. doi: 10.1016/j.arcmed.2022.11.011. Epub 2022 Nov 29. PMID: 36460546.
- [15] Zhang Y, Cao Y, Xin L, Gao N, Liu B. Association between rs1800629 polymorphism in tumor necrosis factor- α gene and dilated cardiomyopathy susceptibility: Evidence from case-control studies. *Medicine (Baltimore).* 2018 Dec;97(50):e13386. doi: 10.1097/MD.00000000000013386. PMID: 30557992; PMCID: PMC6320213.
- [16] Katsarou MS, Papisavva M, Latsi R, Drakoulis N. Hemochromatosis: Hereditary hemochromatosis and HFE gene. *Vitam Horm.* 2019;110:201-222. doi: 10.1016/bs.vh.2019.01.010. Epub 2019 Feb 8. PMID: 30798813.
- [17] Du W, Hu Z, Wang L, Li M, Zhao D, Li H, Wei J, Zhang R. ABCA1 Variants rs1800977 (C69T) and rs9282541 (R230C) Are Associated with Susceptibility to Type 2 Diabetes. *Public Health Genomics.* 2020;23(1-2):20-25. doi: 10.1159/000505344. Epub 2020 Jan 24. PMID: 31982877.

- [18] Gupta K, Kaur G, Pathak T, Banerjee I. Systematic review and meta-analysis of human genetic variants contributing to COVID-19 susceptibility and severity. *Gene*. 2022 Nov 30;844:146790. doi: 10.1016/j.gene.2022.146790. Epub 2022 Aug 17. PMID: 35987511; PMCID: PMC9384365.
- [19] Francis J, Dubashi B, Sundaram R, Pradhan SC, Chandrasekaran A. Influence of Sokal, Hasford, EUTOS scores and pharmacogenetic factors on the complete cytogenetic response at 1 year in chronic myeloid leukemia patients treated with imatinib. *Med Oncol*. 2015 Aug;32(8):213. doi: 10.1007/s12032-015-0665-0. Epub 2015 Jul 5. PMID: 26143137.
- [20] Xia L, Qian W, Yang M, Li Q, Liu F, Xie Y. Comparison of the utility and applicability of the Sokal, Hasford, and EUTOS scores in a population of Chinese patients with chronic-phase chronic myeloid leukemia undergoing imatinib therapy. *Onco Targets Ther*. 2015 Sep 7;8:2485-92. doi: 10.2147/OTT.S85313. PMID: 26392775; PMCID: PMC4573067.
- [21] Wu N, Zhang X, Tian J, Yu S, Qiao Y. Association of GEMIN4 gene polymorphism and the risk of cancer: a meta-analysis. *Onco Targets Ther*. 2017 Nov 2;10:5263-5271. doi: 10.2147/OTT.S147204. PMID: 29138579; PMCID: PMC5679687.
- [22] Mullany LE, Herrick JS, Wolff RK, Buas MF, Slattery ML. Impact of polymorphisms in microRNA biogenesis genes on colon cancer risk and microRNA expression levels: a population-based, case-control study. *BMC Med Genomics*. 2016 Apr 23;9(1):21. doi: 10.1186/s12920-016-0181-x. PMID: 27107574; PMCID: PMC4841949.
- [23] Litwińska Z, Machaliński B. miRNAs in chronic myeloid leukemia: small molecules, essential function. *Leuk Lymphoma*. 2017 Jun;58(6):1297-1305. doi: 10.1080/10428194.2016.1243676. Epub 2016 Oct 13. PMID: 27736267.
- [24] Hrovatin K, Kunej T. Classification of miRNA-related sequence variations. *Epigenomics*. 2018 Apr 1;10(4):463-481. doi: 10.2217/epi-2017-0126. Epub 2018 Mar 23. PMID: 29569482.
- [25] Ensembl genome browser 109. (s. f.). Ensembl. <https://www.ensembl.org/index.html>
- [26] INMEGEN(manual de procedimientos) 103. (s.f)https://www.inmegen.gob.mx/media/filer_public/d3/67/d367da13-07bf-48e6-914b-1927958f8910/mp_polimorfismos.pdf
- [27] Liu H, Hu Y. [Hardy-Weinberg equilibrium in genetic epidemiology]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban*. 2010 Jan;35(1):90-3. Chinese. doi: 10.3969/j.issn.1672-7347.2010.01.013. PMID: 20130370.
- [28] Barabási, A.-L., Gulbahce, N. & Loscalzo, J. Network medicine: A network-based approach to human disease. *Nat Rev Genet*. 2011;12(1), 56-68. DOI: 10.1038/nrg2918
- [29] Fujita T, Yamabe H, Shimada M, Murakami R, Kumasaka R, Nakamura N, Osawa H, Okumura K. Thrombin enhances the production of monocyte chemoattractant protein-1 and macrophage inflammatory protein-2 in cultured rat glomerular epithelial cells. *Nephrol Dial Transplant*. 2008 Nov;23(11):3412-7. doi: 10.1093/ndt/gfn352. Epub 2008 Jul 12. PMID: 18622025.
- [30] Sohail, M., Palma-Martínez, M.J., Chong, A.Y. *et al*. Mexican Biobank advances population and medical genomics of diverse ancestries. *Nature* 622, 775–783 (2023). <https://doi.org/10.1038/s41586-023-06560-0>
- [31] Miedl H, Oswald D, Haslinger I, Gstoettner M, Wenzl R, Proestling K, Schneeberger C, Yotova I, Schreiber M. Association of the Estrogen Receptor 1 Polymorphisms rs2046210 and rs9383590 with the Risk, Age at Onset and Prognosis of Breast Cancer. *Cells*. 2023 Feb 4;12(4):515. doi: 10.3390/cells12040515. PMID: 36831182; PMCID: PMC9953811
- [32] Basak J, Piotrkowska D, Majsterek I, Kucharska E. Relationship between the Occurrence of Genetic Variants of Single Nucleotide Polymorphism in microRNA Processing Genes and the Risk of Developing Multiple Sclerosis. *Biomedicines*. 2022 Dec 3;10(12):3124. doi: 10.3390/biomedicines10123124. PMID: 36551880; PMCID: PMC9776367
- [33] He Y, Zhou Y, Xi Q, Cui H, Luo T, Song H, Nie X, Wang L, Ying B. Genetic variations in microRNA processing genes are associated with susceptibility in depression. *DNA Cell Biol*. 2012 Sep;31(9):1499-506. doi: 10.1089/dna.2012.1660. Epub 2012 Jun 13. PMID: 22694265.
- [34] Cheng H, Li H, Feng Y, Zhang Z. Correlation analysis between SNPs in microRNA-machinery genes and tuberculosis susceptibility in the Chinese Uygur population. *Medicine (Baltimore)*. 2018 Dec;97(52):e13637. doi: 10.1097/MD.0000000000013637. PMID: 30593132; PMCID: PMC6314764.