



UNIVERSIDAD
DE LA REPUBLICA
URUGUAY



PEDECIBA



FACULTAD DE
CIENCIAS
UDELAR | fcien.edu.uy

Universidad de la República

Facultad de Ciencias

Programa de Desarrollo de las Ciencias Básicas

Área Biología – Subárea Biología Celular y Molecular

ESTUDIO DE LA EXPRESIÓN DE RETROVIRUS ENDÓGENOS HUMANOS Y DE
LA ABUNDANCIA DE REPETIDOS ALFA A NIVEL CENTROMÉRICO EN
LEUCEMIA LINFOIDE CRÓNICA

Lic. Sabrina Fischer

Tutora: Dra. María del Pilar Moreno

Co-tutor: Dr. Juan Cristina

Montevideo, Uruguay

2015

ESTUDIO DE LA EXPRESIÓN DE RETROVIRUS ENDÓGENOS HUMANOS Y DE
LA ABUNDANCIA DE REPETIDOS ALFA A NIVEL CENTROMÉRICO EN
LEUCEMIA LINFOIDE CRÓNICA

Lic. Sabrina Carla Fischer Alpuin

Tesis para optar por el título de Magíster en Ciencias Biológicas, Subárea Biología
Celular y Molecular en el marco del Programa de Desarrollo de las Ciencias Básicas
PEDECIBA Biología

Tutora: Dra. María del Pilar Moreno, Profesor Asistente

Co-tutor: Dr. Juan Cristina, Profesor Titular

Montevideo, Uruguay

2015

Montevideo, 12 de noviembre de 2014

PEDECIBA

Biología,

Presente,

Me dirijo a ustedes en carácter de Directora de Tesis de Maestría de la estudiante Sabrina Fischer con el fin de comunicar que la mencionada estudiante planea defender su tesis de Maestría en el correr del mes de marzo del año 2015. Asimismo, solicito a ustedes, sea considerado el siguiente tribunal de Tesis:

Dra. Mabel Berois

Dr. Rodney Colina

Dr. Otto Pritsch

Sin otro particular,

Saluda a ustedes atentamente

Dra. María del Pilar Moreno
Profesor Asistente del Laboratorio de Virología Molecular
Centro de Investigaciones Nucleares
Facultad de Ciencias – UdelaR
Montevideo, Uruguay

Esta tesis te la dedico a ti que dispones de tu tiempo para leerla,
Que la disfrutes

AGRADECIMIENTOS

Es muy lindo pensar en las personas que gracias a su apoyo y estímulo hicieron posible que hoy esté finalizando mi tesis de maestría. A ellos les escribo mis especiales agradecimientos.

En primer lugar quiero agradecerle a mi tutora, la Dra. María del Pilar Moreno. Gracias por confiar en mí para comenzar juntas una nueva línea de investigación en el Laboratorio de Virología Molecular. Gracias además por tu ayuda constante al momento de decidir viajar fuera del país para realizar una pasantía de investigación. Gracias por el tiempo dedicado, por las instancias de discusión a medida se fueron logrando resultados y también por las acertadas recomendaciones y críticas, que siempre desde una perspectiva positiva, permitieron que se mejorara sustancialmente la escritura del trabajo.

Quiero agradecerle al jefe del Laboratorio de Virología Molecular del CIN y actual decano de la Facultad de Ciencias, el Dr. Juan Cristina. Gracias por permitirme, una vez más, continuar con mi formación en tu laboratorio. Gracias por aceptar ser el co-orientador de mi tesis y un especial agradecimiento por el apoyo incondicional frente a la idea de querer viajar al exterior.

Estoy muy agradecida con mis compañeros del Laboratorio de Virología Molecular. Gracias a la Lic. Victoria Comas, MSc.Natalia Echeverría, Lic. Patricia Núñez, Lic. Gabriela Bentancour, Lic. Martín Sónora, MSc.Álvaro Fajardo y el Dr. Gonzalo Moratorio por su compañerismo y por hacer muy disfrutable el trabajo del día a día. Gracias por dedicar su tiempo y siempre estar dispuestos a colaborar a la hora de discutir e interpretar metodologías y resultados.

Agradecerle además al Dr. Pablo Oppezzo y al Dr. Guillermo Dighiero del Instituto Pasteur de Montevideo por permitirnos el acceso a su modelo de estudio, la Leucemia Linfocítica Crónica. Gracias además por su colaboración en la discusión de los resultados. Aquí incluyo mi agradecimiento al Servicio de Hematología del Hospital Maciel de Montevideo así como también al Servicio Nacional de Sangre por cedernos las muestras de estudio.

Agradezco al Dr. David Markovitz por recibirme tan amablemente en su Laboratorio de Investigación de la Universidad de Michigan, EEUU. Gracias a ti y a los miembros de tu laboratorio por la confianza y por estar siempre dispuestos a ayudarme en el manejo de todas sus instalaciones. Gracias David por permitirme ser parte de eventos, reuniones,

seminarios y cursos, y por tu interés en el progreso de mi formación profesional. Gracias por abrirme las puertas de tu casa haciéndome sentir en familia con el cuidado y cariño de ti y de tu esposa Ruth.

Mi agradecimiento al Dr. Rafael Contreras-Galindo por aceptar ser mi tutor durante los seis meses y medio de investigación en el Laboratorio del Dr. Markovitz. Gracias por tu confianza, tu tiempo, tus enseñanzas, y por permitirme ser parte de tus proyectos de investigación desde el primer día.

Gracias a la Dra. Mabel Berois, al Dr. Otto Pritsch, y al Dr. Rodney Colina por aceptar ser los correctores de la tesis de maestría y los miembros del tribunal para el momento de la defensa. Un placer contar con su evaluación.

A la ANII por financiar y hacer entonces posible la realización del proyecto de maestría y la pasantía de investigación en el exterior a través del Fondo Clemente Estable, la Beca de Posgrado Nacional y la Beca de Movilidad tipo Capacitación.

Gracias a mis padres Laura y Gustavo y mis hermanos Marcelo y Diego a quienes adoro con toda mi alma. Gracias por ser mi sostén día a día, por ser pacientes y siempre estar allí para escucharme, por transmitirme ganas y fuerza para continuar en mi vocación. Gracias por confiar en mí sin condiciones, por enseñarme que si uno realmente quiere, puede, y por hacer que mi camino sea agradable y rodeado de amor.

Gracias a mi abuela Gladys, quien hoy en día con sus 85 años me sigue transmitiendo el valor de la confianza, del optimismo, de la esperanza y de la bondad. Gracias abu por tu compañía, tus ganas de aprender de biología me llenan de felicidad.

Un agradecimiento a mis hermanas de la vida, Agustina y Victoria. Gracias por estar siempre, apoyándome y confiando en mí para continuar sin titubear. Gracias de corazón por su amistad, su compañía y su alegría.

Y un profundo agradecimiento a Facundo, mi novio. Tu compañía fue el motor principal para continuar con fuerza y alegría. Gracias por el impulso y el apoyo para realizar el viaje. Un viaje que no hubiese sido tan aprovechado y tan disfrutable sin tu compañía, bien cerca estando tan lejos. Gracias por hacer que el hecho de extrañarte se acorara en el tiempo. Gracias por escucharme y darme tus consejos pensando lo mejor para mí. Gracias por tu paciencia. Gracias por hacerme reír todos los días, y gracias por ser parte de mi vida.

LISTA DE TABLAS

Tabla 1. Secuencia de <i>primers</i> y tamaños de los productos de amplificación esperados para los genes de HERV-K y <i>gapdh</i>	27
Tabla 2. Nombre y secuencia de los <i>primers</i> diseñados para amplificar los repetidos alfa cromosoma-específicos.....	31
Tabla 3. Posibles polimorfismos obtenidos en la secuencia nucleotídica del gen <i>np9</i> y el cambio aminoacídico que deriva de su traducción.....	44
Tabla 4. Nombre de los <i>primers</i> y condiciones de amplificación para cada repetido alfa centromérico.....	57

LISTA DE FIGURAS

Figura 1. Contenido de elementos transponibles (TEs) en el genoma humano	5
Figura 2. Dendograma sin raíz representativo de los siete géneros de retrovirus exógenos y las tres clases de retrovirus endógenos	7
Figura 3. Estructura de un provirus de HERV-K	9
Figura 4. Organización genómica de los centrómeros humanos	15
Figura 5. Características generales de los pseudofolículos en LLC	19
Figura 6. Características biológicas que diferencian las células B de LLC con perfil IgVH Mutado y No Mutado (M y NM)	21
Figura 7. Esquema representativo de una curva de calibración estándar obtenida de la cuantificación absoluta de diluciones seriadas de una muestra de referencia por PCR en tiempo real	34
Figura 8. Organización genómica de un provirus de HERV-K centromérico (K111)	35
Figura 9. PCR cualitativa para los genes <i>gag</i> , <i>pol</i> , <i>env</i> , <i>rec</i> y <i>np9</i> de HERV-K	38
Figura 10. Niveles de expresión relativa de los transcritos <i>np9</i> y <i>gag</i> de HERV-K en pacientes con LLC con perfil de genes VH de las inmunoglobulinas mutados y no mutados	41
Figura 11. Expresión relativa del gen celular <i>c-myc</i> en PBMCs de pacientes de LLC con alto y bajo nivel de expresión del gen <i>np9</i> de HERV-K y en humanos sanos	43
Figura 12. Cromatograma de las posiciones nucleotídicas 38 y 199 de la secuencia del gen <i>np9</i> de HERV-K	45
Figura 13. Alineamiento de las secuencias nucleotídicas del gen <i>np9</i> de HERV-K	47
Figura 14. Curvas de amplificación de repetidos alfa para cada cromosoma humano obtenidas por PCR a tiempo real	51
Figura 15. Visualización de los productos de amplificación de los repetidos alfa para cada juego de <i>primers</i> utilizado	56
Figura 16. Cuantificación absoluta por PCR en tiempo real de repetidos alfa centroméricos cromosoma-específicos en una línea celular de LLC (CLL-AT) y en humanos sanos	60

Figura 17. Cuantificación absoluta por PCR en tiempo real de retrovirus endógenos humanos centroméricos (K111) en una línea celular de LLC (CLL-AT) y en humanos sanos———61

LISTA DE ABREVIATURAS Y SIGLAS

ADN: Ácido desoxirribonucleico

AID: Citidina deaminasa inducida por activación

AR: Artritis reumatoidea

ARN: Ácido ribonucleico

ARNm: Ácido ribonucleico mensajero

BCR: Receptor de células B

EM: Esclerosis múltiple

ERVs: Retrovirus endógenos

EtBr: Bromuro de Etidio

FDC: Célula dendrítica folicular

FISH: Hibridación fluorescente *in situ*

GAPDH: Gliceraldehido 3-fosfato deshidrogenasa

HERVs: Retroviruses endógenos humanos

HML: HERVs tipo virus del tumor mamario de ratón-MMTV

HMS: Hipermutación somática

HORs: Repetidos de alto orden

Igs: Inmunoglobulinas

IgVH: Región variable de las cadenas pesada de las Igs

IgVH-M: Paciente con LLC perfil IgVH mutado

IgVH-NM: Paciente con LLC perfil IgVH no mutado

K111: provirus de HERV-K centromérico

LINEs (L1): Elementos intercalados largos tipo 1

LLA: Leucemia linfocítica aguda

LLC: Leucemia linfocítica crónica

LMA: Leucemia mieloide aguda

LMC: Leucemia mieloide crónica

LNx: Ligando de la proteína Numb X

LSE: Lupus sistémico eritematoso

LTR: Repetidos terminales largos

MO: Médula ósea

MYA: Millones de años atrás

NK: Célula natural killer

OLS: Órganos linfoides secundarios

PBMCs: Células monoclonales de sangre periférica

PCR: Reacción en cadena de la Polimerasa

PLZF: Proteína en dedos de zinc de la leucemia promielocítica

PM: Peso molecular

RT: Retrotranscripción

SAg: Superantígeno

SCR: Cambio de clase por recombinación

SFB: Suero fetal bovino

SINEs: Elementos intercalados cortos

SP: sangre periférica

SSJ: Síndrome de Sjögrens

SVA: Elementos SINEs, repetidos en tándem de número variable y elementos *Alu*

TEs: Elementos transponibles

TSD: Duplicación en el sitio blanco

VLPs: Partículas tipo virales

C_T: Ciclo umbral

EB: Puntos de ruptura evolutivos

TABLA DE CONTENIDOS

RESUMEN	1
SUMMARY	3
1. MARCO TEÓRICO	4
1.1. Elementos Transponibles en el genoma humano	4
1.2. Retrovirus endógenos humanos (HERVs)	5
1.2.1. Definición y características generales	5
1.2.2. Familia de Retrovirus Endógenos Humanos tipo K (HERV-K)	8
1.2.3. Estructura de un provirus de HERV-K	8
1.2.4. Rol de los HERVs en la salud y en la enfermedad	10
1.2.5. HERVs y cáncer	11
1.2.6. Reciente hallazgo, los HERVs integrados en centrómeros humanos	12
1.3. Estructura del ADN centromérico, los repetidos alfa	14
1.4. Leucemia Linfoide Crónica	16
1.4.1. Generalidades y algunos datos estadísticos	16
1.4.2. Inmunofenotipo de las células B de LLC	17
1.4.3. Subpoblación de linfocitos proliferantes	17
1.4.4. Marcadores pronósticos de la LLC	19
2. OBJETIVOS	23
2.1. Objetivo general 1	23

2.1.1. Objetivos específicos 1	23
2.2. Objetivo general 2	23
2.2.1. Objetivos específicos 2	23
3. MATERIALES Y MÉTODOS	24
3.1. Metodología experimental para abordar el objetivo general 1	24
3.1.1. Muestras de estudio	24
3.1.2. Cuantificación de las concentraciones de ARN y tratamiento con DNAsa	25
3.1.3. Retrotranscripción del ARN (RT-PCR)	25
3.1.4. Puesta a punto de la PCR cualitativa de todos los genes de HERV-K	25
3.1.5. PCR para la expresión del gen constitutivo gliceraldehido 3-fosfato deshidrogenasa (<i>gapdh</i>)	26
3.1.6. PCR en tiempo real para los genes <i>np9</i> y <i>gag</i> de HERV-K	27
3.1.7. PCR en tiempo real para cuantificar expresión relativa del gen celular <i>c-myc</i>	29
3.1.8. Amplificación por PCR convencional y secuenciación del gen <i>np9</i>	29
3.2. Metodología experimental para abordar el objetivo general 2	30
3.2.1. Diseño de <i>primers</i> para la amplificación de los repetidos alfa cromosoma-específicos	30
3.2.2. Puesta a punto de la PCR a tiempo real de los repetidos alfa cromosoma-específicos	31
3.2.3. Realización de las curvas estándar para la cuantificación absoluta por PCR a tiempo real	33
3.2.4. Abundancia de repetidos alfa en la línea celular de LLC (CLL-AT) y en células humanas sanas	34
3.2.5. PCR a tiempo real específica para K111	35

3.2.6. Análisis estadístico	36
4. RESULTADOS	37
4.1. Resultados del Objetivo General 1	37
4.1.1. Objetivo específico 1	37
4.1.2. Objetivo específico 2	38
4.1.3. Objetivo específico 3	41
4.1.4. Objetivo específico 4	42
4.2. Resultados del Objetivo General 2	48
4.2.1. Objetivo específico 1	48
4.2.2. Objetivo específico 2	58
5. DISCUSIÓN	61
5.1. Con respecto al objetivo general 1	62
5.1.1. Objetivo específico 1	63
5.1.2. Objetivo específico 2	63
5.1.3. Objetivo específico 3	65
5.1.4. Objetivo específico 4	67
5.2. Con respecto al objetivo general 2	68
5.2.1. Objetivo específico 1	70
5.2.2. Objetivo específico 2	71
6. CONCLUSIONES	73
7. PERSPECTIVAS	75
8. REFERENCIAS BIBLIOGRÁFICAS	77
9. APÉNDICE	89

RESUMEN

Los retrovirus endógenos humanos (HERVs por su sigla en inglés) son remanentes genéticos de infecciones retrovirales ancestrales de la línea germinal que han sido transmitidos de generación en generación hasta nuestros días. Alrededor del 8% de nuestro genoma está constituido por este tipo de elementos. HERV-K (HML-2) es la familia de retrovirus endógenos más joven en el genoma humano y por lo tanto la más preservada genéticamente. Varias evidencias reportan un aumento de expresión tanto a nivel transcripcional como traduccional de los HERVs en una gran variedad de cánceres entre los cuales se encuentran leucemias y linfomas.

Recientemente se han encontrado cientos de provirus de HERV-K distribuidos en las regiones centroméricas de al menos 15 cromosomas humanos denominados "K111". La compleja organización de los centrómeros humanos compuestos fundamentalmente por secuencias repetitivas, principalmente los elementos repetidos alfa, así como cientos de K111 insertos en ellos, hacen que el acceso a la información de los centrómeros sea un verdadero desafío. Esto es aún más interesante si se considera a los centrómeros como los *loci* cromosomales esenciales para la segregación de los cromosomas y la estabilidad del genoma.

Utilizando como modelo de estudio una patología hemato-oncológica como es la leucemia linfocítica crónica (LLC) y con el fin de contribuir a su conocimiento, el objetivo principal de la presente tesis consistió en estudiar una posible asociación entre los HERVs y la LLC. Asimismo se prosiguió al diseño y puesta a punto de una metodología cuantitativa que nos permita acceder a información sobre el arreglo centromérico de los repetidos alfa en un modelo de línea celular de LLC en comparación con muestras provenientes de donantes sanos.

Los resultados de esta tesis reportan una sobreexpresión del gen *np9* producido por el HERV-K en la mayoría de las muestras analizadas de pacientes con LLC respecto a las muestras de donantes sanos. Del mismo modo se observó que la sobreexpresión de *np9* no está directamente relacionada con una sobreexpresión del proto-oncogén celular *c-myc*. Finalmente, el estudio de los centrómeros reveló una reducción significativa del número total de retrovirus endógenos (K111) así como también diferencias (ganancia o pérdida) en el

número de repetidos alfa de determinados cromosomas en el modelo de línea celular CLL-AT al compararse con muestras de humanos sanos.

El presente estudio constituye el primer reporte del aumento de la expresión del gen *np9* de HERV-K en muestras de sangre periférica de pacientes con LLC y propone un nuevo abordaje experimental para acceder al estudio de la estructura de los centrómeros en forma cromosoma-específica.

SUMMARY

Human endogenous retroviruses (HERVs) are genetic remnants of ancient retroviral infections of the germline that have been transmitted over generations to modern days. About 8% of our genome consists of these elements. The HERV-K (HML-2) family is the youngest retroviral family in the human genome and, is therefore, the most genetically preserved. A considerable amount of studies have reported evidence of increased expression of HERVs at the transcriptional and translational level in a variety of human cancers including leukemias and lymphomas.

Recently, hundreds of HERV-K proviruses have been found distributed in the centromeric regions of at least 15 human chromosomes and were designated "K111". The complex organization of the repetitive sequences that compose the human centromeres, mainly alpha repeats and K111 insertions within them, makes the study of centromeres a real challenge. This is even more interesting when we consider the centromeres as the chromosomal *loci* responsible for chromosome segregation and genome stability.

Using the hematological disease, chronic lymphocytic leukemia (CLL), as a study model, the main goal of this project was to study the expression level of certain genes of HERV-K and its possible association with the CLL, as well as the development of a methodology to gain insight into the centromeric structure of alpha repeats in a CLL cell line model compared to samples from healthy donors.

The results in this study show an over-expression of the putative oncogenic gene *np9* of HERV-K in most CLL samples with respect to samples from healthy individuals, and further note that this over-expression of *np9* is not directly related to an over-expression of the cellular proto-oncogene *c-myc*. The centromere study reveals a significant reduction of the total number of endogenous retroviruses (K111), as well as a significant difference (gain or loss) in the number of alpha repeats in certain chromosomes for the cell line model CLL-AT.

This thesis constitutes the first report of an over-expression of the *np9* gene of HERV-K in CLL, and provides a new experimental methodology to obtain interesting information about centromeres in human chromosomes.

1. MARCO TEÓRICO

1.1. Elementos Transponibles en el genoma humano

Uno de los hallazgos más sorprendentes que surgió a partir del conocimiento de la secuencia completa del genoma humano, fue la gran cantidad de elementos transponibles (TEs) que este alberga, los cuales son conocidos también como “genes saltadores”. Cerca de la mitad de nuestro genoma (~45%) está compuesto por este tipo de elementos (Downey R, *et al.*, 2014). Esta contribución se vuelve aún más interesante si se tiene en cuenta que tan solo el 1,5% de nuestro genoma está compuesto por regiones que codifican para proteínas (Cordaux & Batzer, 2009).

Los TEs se pueden separar en dos clases principales: los transposones de ADN y los retrotransposones. Estos últimos se replican a través de un ARN intermediario para luego insertarse en posiciones nuevas del genoma. Los retrotransposones pueden ser subdivididos en dos grupos que se distinguen por la presencia o ausencia de repetidos terminales largos (LTRs por su sigla en inglés). Los no-LTR-retrotransposones incluyen los elementos LINE-1 (elementos intercalados largos tipo 1), *Alu* y elementos SVA (conformados por regiones de elementos intercalados cortos que se abrevian SINE, un número variable de repetidos en tándem, y regiones tipo *Alu*). En conjunto, los no-LTR-retrotransposones ocupan aproximadamente un tercio del genoma humano. Por otro lado, los retrotransposones con LTRs se tratan de secuencias de retrovirus endógenos humanos (HERVs) y se estima que los mismos ocupan alrededor del 8% de nuestro genoma (ver Fig. 1) (Cordaux & Batzer, 2009).

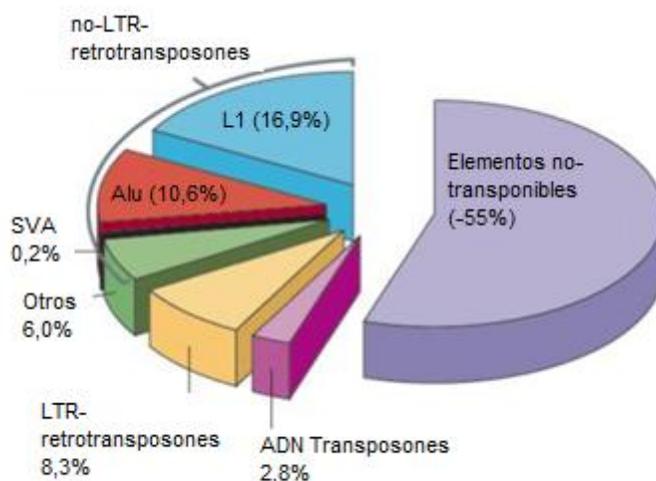


Figura 1. Contenido de elementos transponibles (TEs) en el genoma humano. Cerca del 45% del genoma humano puede ser reconocido como derivado de TEs, la mayoría de los cuales son no-LTR-retrotransposones como los LINEs (L1), elementos *Alu* y elementos SVA. Una porción del 8% de nuestro genoma está constituida por LTR-retrotransposones los cuales derivan de secuencias provirales de retrovirus endógenos humanos (HERVs). Figura modificada de Cordaux & Batzer, 2009.

1.2. Retrovirus endógenos humanos (HERVs)

1.2.1. Definición y característica generales

Los retrovirus endógenos humanos son remanentes genéticos de infecciones retrovirales ancestrales en la línea germinal, que se integraron en el genoma celular desde hace ~2 a 40 millones de años atrás, y han sido transmitidos de generación en generación hasta nuestros días. Los progenitores de estos retroelementos fueron retrovirus exógenos que infectaron células gametos, volviéndose subsecuentemente endógenos y por lo tanto, sujetos a las leyes de herencia Mendeliana (Downey R, *et al.*, 2014; Löwer R, *et al.*, 1996). El genoma humano y el genoma de otras especies de vertebrados que contienen retrovirus endógenos pueden entonces ser reconocidos como museos de retrovirus ancestrales (Hohn O, *et al.*, 2013).

Los HERVs existen en el genoma en formas provirales (genoma de ADN viral integrado en el genoma de ADN celular) y comparten la estructura genómica universal de todos los retrovirus: 5´LTR- *gag- pro- pol- env-* 3´LTR. Sus principales marcos de lectura abiertos (ORFs por su sigla en inglés) codifican poliproteínas virales las cuales, luego de sufrir

modificaciones post-traduccionales, conforman las proteínas estructurales y enzimáticas fundamentales de los retrovirus tales como la transcriptasa reversa (producto de expresión del gen *pol*) o la proteína de envoltura (producto del gen *env*). Los LTRs flanquean los extremos 5' y 3' del provirus, y cumplen importantes funciones reguladoras para la transcripción y poliadenilación del ARN transcripto (Bannert & Kurth, 2004; Downey RF, *et al.*, 2014). Algunos HERVs han pasado por múltiples eventos de amplificación (duplicación y recombinación), de modo que se han ido generando miles de secuencias retrovirales, parciales o completas, dispersas por todo el genoma del hospedero. Algunas de estas secuencias mantienen la estructura proviral completa, mientras que muchas otras son defectivas y han perdido la capacidad individual de replicarse debido a la acumulación de mutaciones adquiridas durante el tiempo (Sentís C, 2002). La inmensa mayoría se trata de LTRs solitarios (también denominados “solo LTRs”) originados por recombinación de los LTRs 5' y 3', en donde los genes internos han sido removidos. Los solo LTRs aun preservan la capacidad reguladora (Hughes & Coffin, 2004).

Si bien no existe una taxonomía integrada de virus endógenos y exógenos, los HERVs se han subdividido en 3 clases en base a su similitud de secuencia con los diferentes géneros en que se han clasificado los retrovirus exógenos. Los HERVs que se agrupan filogenéticamente junto con el grupo de los gammaretrovirus son los denominados de clase I (incluye HERV-W y HERV-H entre otros subgrupos). Los que se agrupan junto con los betaretrovirus son los denominados de clase II (incluye HERV-K), y los que se agrupan con los spumavirus son los denominados de clase III (incluye HERV-L y HERV-S). La mayoría de estos virus son de cualquier manera muy diferentes al género retroviral asignado (Griffiths DJ, 2001). Los HERVs clase I y III son los grupos más ancestrales y se encuentran presentes a lo largo de todo el linaje de los primates, mientras que los HERVs de Clase II se encuentran restringidos a Chimpancés y Humanos. Algunos subgrupos dentro de los HERV-K (de clase II) son específicamente humanos lo que indica que fueron integrados en el genoma dentro de los últimos 5 millones de años de evolución (Medstrand & Mager, 1998) (ver Fig. 2).

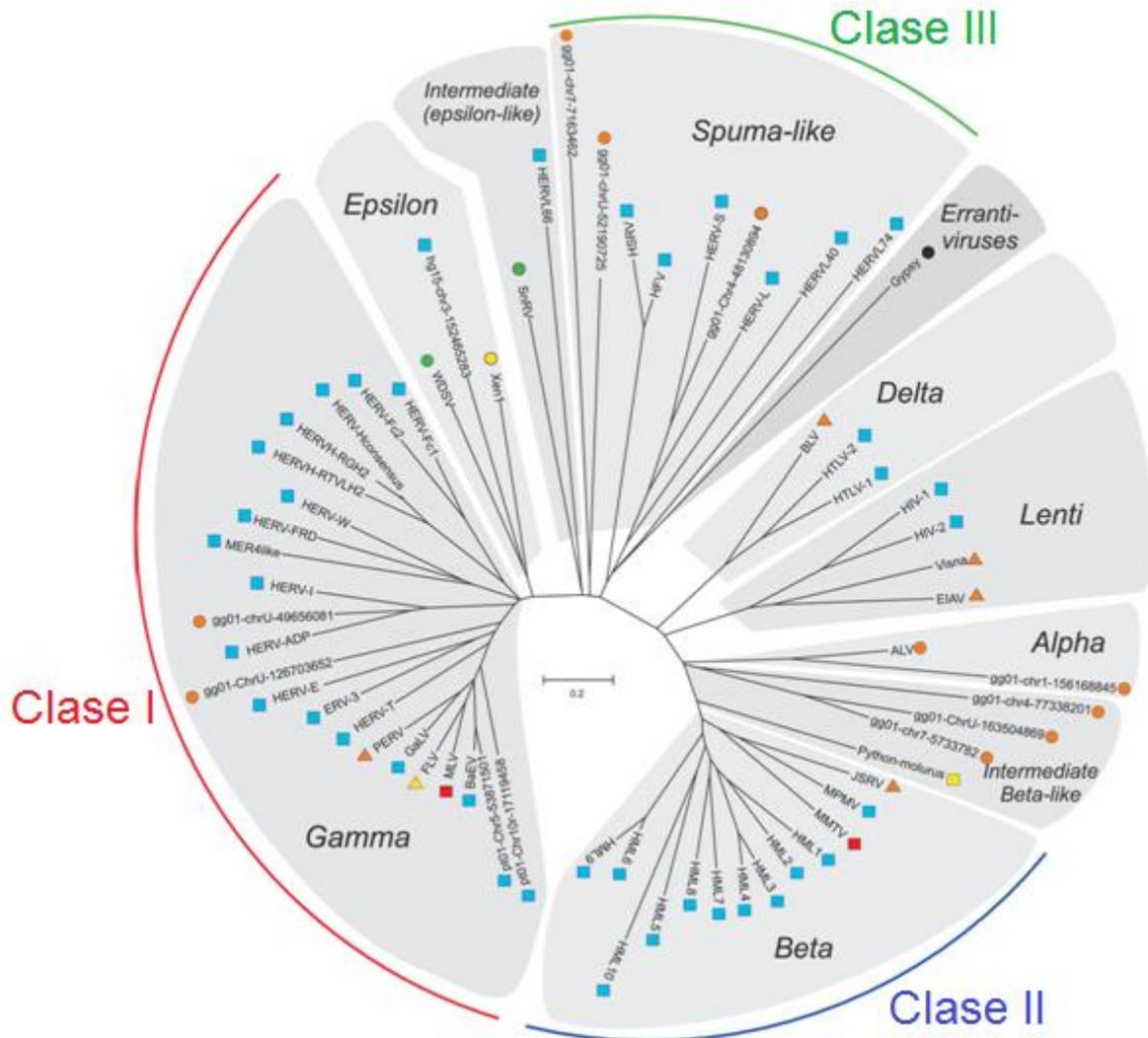


Figura 2. Dendrograma sin raíz representativo de los siete géneros de retrovirus exógenos y las tres clases de retrovirus endógenos. Análisis filogenético sin raíz del gen *pol* por el método *neighbor joining* (NJ) (consenso 500 bootstrap) de los siete géneros retrovirales: alfa, beta, gamma, delta, epsilon, lenti y spuma-virus, y de los retrovirus endógenos (representados en color celeste). En la periferia se indica a que clase pertenecen los HERVs agrupados con los diferentes géneros de retrovirus. Figura modificada de Jern P, *et al.*, 2005.

1.2.2. Familia de Retrovirus Endógenos Humanos tipo K (HERV-K)

Dentro de la clase II de HERVs se encuentra la familia de los HERVs tipo K (HERV-K). Esta representa la familia de retrovirus endógenos más jóvenes incluyendo miembros recientemente integrados en el genoma humano y por ende es la más preservada genéticamente. A su vez, esta constituye la familia de HERVs biológicamente más activa. La designación “K” proviene del uso de un ARN de transferencia con una Lisina (ARNt-K) para primar la transcripción reversa (Hohn O, *et al.*, 2013).

Esta familia está compuesta por 11 subgrupos (HML-1 al HML-11). Uno de estos subgrupos, el HML-2, ha sido sujeto a una intensa investigación dado que mantiene sus secuencias provirales completas o casi-completas conservando los ORFs para todas las poliproteínas virales. Algunos provirus de HML-2 son específicos de humanos, lo cual indica eventos independientes de integración posteriores a la divergencia de humanos-chimpancés (~4-6 millones de años atrás) e incluso dentro de los humanos modernos (Barbulescu M, *et al.*, 1999; Downey RF, *et al.*, 2014). Esta clase de HERV-K (HML-2) es considerada el grupo de estudio más interesante en términos de potencial actividad oncogénica (Subramanian RP, *et al.*, 2011).

Hasta la fecha se reporta que HML-2 es representado en el genoma por 91 provirus de secuencia completa (Subramanian RP, *et al.*, 2011) y más de 2500 secuencias “solo LTRs” (Hughes & Coffin, 2004). El mecanismo de integración de HML-2 en el ADN humano produce una marca de 5-6 pares de bases idéntica en cada extremo del provirus (en inglés a esta marca se la denomina “target site duplication” o TSD). Se ha reportado la presencia de provirus híbridos, por contener diferentes TSD en sus extremos, y este hecho podría ser una evidencia de eventos de recombinación homóloga entre diferentes provirus de HML-2 integrados en el genoma humano (Hughes & Coffin, 2005).

1.2.3. Estructura de un provirus de HERV-K

Los miembros de la familia de HERVs tipo K (HML-2) se pueden clasificar en dos tipos principales: los de tipo I y los de tipo II. Los HML-2 tipo I se caracterizan por presentar una delección de 292 pares de bases (pb) entre los genes *pol* y *env*, mientras que los HML-2 tipo II no cuentan con esta delección. La delección en los provirus de tipo I da lugar a un evento de *splicing* alternativo que culmina en la producción de una proteína denominada Np9, mientras

que en los provirus de tipo II este *splicing* genera la proteína accesoria completa conocida como Rec (Downey RF, *et al.*, 2014) (ver Fig. 3).

Ha sido reportado que ciertas características estructurales en ambas proteínas, Rec y Np9, promueven su interacción con diferentes proteínas celulares alterando así diversas vías de señalización. Por ejemplo, estas proteínas virales comparten un dominio de unión a la proteína en dedos de zinc de la leucemia promielocítica (PLZF por su sigla en inglés), la cual es un represor transcripcional del proto-oncogen *c-myc*, y tiene asignado un papel putativo oncogénico (Denne M, *et al.*, 2007). Se ha detectado que Np9 interactúa también con LNX, una ubiquitin ligasa E3 que tiene como blanco miembros de la vía de señalización celular Numb/Notch, vía que se ha visto implicada en la regulación de la proliferación de cáncer de mama y próstata (Armbruster V, *et al.*, 2002). Asimismo, se le ha asignado un rol a Np9 como modificador crítico molecular que co-activa diferentes vías de señalización celular como la β -catenina, ERK, Akt y Notch 1 y promueve el crecimiento de células madre/progenitoras leucémicas humanas (Chen T, *et al.*, 2013).

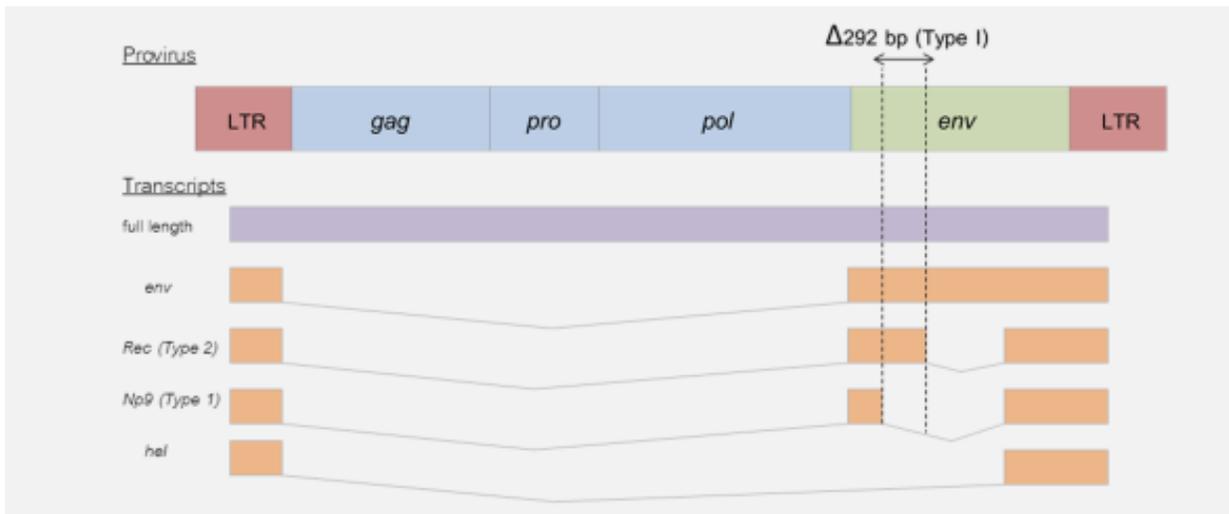


Figura 3. Estructura de un provirus de HERV-K. El transcripto de secuencia completa de HML-2 codifica para las poliproteínas Gag, Pro y Pol. Un único *splicing* alternativo codifica para la proteína Env, mientras que el transcripto de doble *splicing* codifica para las proteínas accesorias Rec o Np9 dependiendo de la presencia o ausencia de una deleción de 292 pares de bases entre los genes *pol* y *env*. HML-2 también expresa un transcripto denominado "*hel*" de 1.5kb de función aún desconocida. Figura reproducida de Downey RF, *et al.*, 2014.

1.2.4. Rol de los HERVs en la salud y en la enfermedad

La retención de los genes de HERVs durante la evolución sugiere que estos elementos han provisto un rol beneficioso en la supervivencia humana. En este aspecto, está demostrada la importancia de los HERVs en funciones celulares normales, resaltando su implicancia en procesos tan complejos y significativos como es la placentación. Se ha demostrado que las glicoproteína de envoltura pertenecientes a los HERVs de la familia W y FRD, denominadas sincitina-1 y 2 respectivamente, juegan un importante rol mediando la fusión entre las células trofoblásticas, contribuyendo así a la renovación constante de nutrientes y hormonas entre la madre y el feto, y a la estabilidad de esta estructura altamente dinámica (Blaise S, *et al.*, 2004; Mi S, *et al.*, 2000). Asimismo, la presencia de los elementos “solo LTR” le ha agregado cierto grado de complejidad a nuestro genoma, y en este sentido muchos de los LTRs han sido co-elegidos por genes codificadores de proteínas dirigiendo una expresión tejido-específica (Cohen CJ, *et al.*, 2009).

Por otra parte, la asociación entre los HERVs y procesos patológicos ha generado un importante grado de atención en los investigadores a partir de la información obtenida en la secuenciación del genoma humano (Lander ES, *et al.*, 2001). Los HERVs han sido propuestos como co-factores en procesos crónicos como son las enfermedades neurológicas, autoinmunes y cáncer (Löwer R, 1999). Existen diferentes mecanismos moleculares por los cuales los HERVs pueden estar actuando en la promoción de estos procesos:

a) Un mecanismo puede ser explicado por la capacidad que tienen estas secuencias de ser movilizadas e integradas próximas a ciertos genes cuya expresión pueda quedar consecuentemente alterada.

b) Otro mecanismo puede darse por una activación y sobreexpresión de genes de HERVs que den lugar a la producción de proteínas que disparen una respuesta inmune de generación de anticuerpos. Esto puede derivar en enfermedades auto-inmunes como la artritis reumatoidea (AR), lupus sistémico eritematoso (LSE), esclerosis múltiple (EM), síndrome de Sjögren´s (SSJ), enfermedades de mezcla de tejido conectivo y enfermedades neurológicas inflamatorias (Moyes DL, *et al.*, 2005). Se ha propuesto incluso que proteínas de HERV-K pueden actuar como superantígenos los cuales inducen una reacción de linfocitos T autoreactivos. La primera evidencia de HERVs actuando como superantígenos

fue en el estudio de la diabetes tipo I donde se detectó la expresión de la proteína Env de HERV-K en las islas pancreáticas de pacientes con esta enfermedad (Conrad B, *et al.*, 1997).

c) Un tercer mecanismo podría asociarse con la producción de proteínas de HERV fundamentalmente Rec y Np9 las cuales tienen asignado un putativo rol oncogénico. Se sabe que estas proteínas presentan dominios de unión a diferentes proteínas celulares y su efecto puede resultar en la modificación (inhibición o activación) de diferentes vías de señalización celular. A modo de ejemplo retomamos el concepto del dominio de unión de Np9 y Rec a PLZF, un represor transcripcional de determinados genes celulares entre los cuales se encuentra el proto-oncogén *c-myc* implicado en la proliferación y supervivencia celular (Denne M, *et al.*, 2007).

De manera interesante, algunos agentes ambientales como la luz ultravioleta y varios agentes químicos, pueden inducir la expresión de HERVs derivando en un efecto (directo o indirecto) en la iniciación y evolución clínica de ciertas enfermedades crónicas. Existe un consenso de que cambios epigenéticos pueden dar lugar a la activación de retroelementos al cambiar el patrón de metilación de las secuencias pero sin modificar su composición nucleotídica (Balada E, *et al.*, 2009).

1.2.5. HERVs y cáncer

En lo que respecta a los HERVs y su asociación a procesos oncogénicos, varias evidencias reportan un aumento de expresión tanto a nivel transcripcional como traduccional de los HERVs, así como su desmetilación en una gran variedad de cánceres. Esta desregulación puede resultar en efectos oncogénicos de proteínas codificadas por los HERVs o en la activación transcripcional de diferentes oncogenes celulares (Romanish MT, *et al.*, 2010). A la fecha hay evidencia de una activación de los HERVs HML-2 en cáncer de ovario (Wang-Johanning F, *et al.*, 2007), de mama (Ejthadi HD, *et al.*, 2005), de próstata (Goering W, *et al.*, 2011), tumores de la línea germinal (Kleiman A, *et al.*, 2004), melanomas (Singh S, *et al.*, 2009), leucemias (Depil S, *et al.*, 2002; Iwabuchi H, *et al.*, 2004) y linfomas (Contreras-Galindo R, *et al.*, 2008).

Muchos cánceres exhiben un estado de hipometilación generalizado, por lo tanto, la activación de los HERVs durante la tumorigénesis puede ser un efecto colateral de este estado epigenético. Se podría pensar que el incremento en la transcripción de HERVs en

determinados tipos de cáncer puede ser, en parte, debido a una hipometilación de sus LTRs. Se ha reportado hipometilación del ADN de HML-2 en líneas celulares de melanoma (Stengel S, *et al.*, 2010), tumores de próstata (Goering W, *et al.*, 2011), y tumores de ovario (Iramaneerat K, *et al.*, 2011). Otros factores que pueden estar afectando la expresión de los HERVs son determinadas citoquinas y esteroides en diferentes células y tejidos (Ono M, *et al.*, 1987; Romanish MT, *et al.*, 2010; Schneider PM, *et al.*, 2001). También se ha reportado la activación de HERV-K inducida por otras infecciones virales como por ejemplo del virus de Epstein-Barr (EBV) (Sutkowski N, *et al.*, 2004) y el virus de la Inmunodeficiencia Humana (HIV-1) (Contreras-Galindo R, *et al.*, 2007 y 2012), e incluso por factores de transcripción entre los cuales se incluyen el NF- κ B, NF-AT, MITF-M (Kato I, *et al.*, 2011), Sp1, Sp3 (Fuchs NV, *et al.*, 2011) y el YY1 (Knössl M, *et al.*, 1999).

Específicamente en relación a los procesos hemato-oncológicos, a finales del siglo XX Brodsky y colaboradores reportaron un potencial rol de los HERV-K en leucemia (Brodsky I, *et al.*, 1993a). Ellos detectaron que el ARNm del gen *pol* de HML-2 se expresa en la sangre de pacientes con leucemia mieloide crónica (LMC) y leucemia mieloide aguda (LMA) (Brodsky I, *et al.*, 1993b). Otro grupo de investigadores reportó que el ARNm del gen *gag* de HML-2 se encuentra altamente expresado en células monoclonales de sangre periférica (PBMCs por su sigla en inglés) de pacientes con diferentes tipos de leucemias, si se los compara con donantes sanos (de cinco a diez veces mayor), lo que sugirió que esta expresión estaba específicamente vinculada al desarrollo del tumor (Depil S, *et al.*, 2002). Estudios similares han reportado partículas virales de HML-2 en linfomas (Contreras-Galindo R, *et al.*, 2008) y expresión del gen *env* en la línea celular de linfoma de células T H9 (Vogetseder W, *et al.*, 1995). También se ha visto que la expresión de HML-2 decrece luego del tratamiento de un linfoma, indicando que este efecto podría ser útil para monitorear respuestas terapéuticas (Contreras-Galindo R, *et al.*, 2008). Adicionalmente, un estudio reciente identifica a la proteína Np9 como un potente oncogén viral que promueve el crecimiento de células madre/progenitoras leucémicas humanas a través de la activación de diferentes vías de señalización celular (Chen T, *et al.*, 2013).

1.2.6. Reciente hallazgo, los HERVs integrados en centrómeros humanos

Si bien el proyecto genoma humano permitió un gran avance en materia de conocimiento científico, la secuencia del genoma humano no fue del todo completa, y de manera

importante, las principales brechas (sitios sin secuenciar) corresponden a las regiones centroméricas de cada cromosoma humano. Esto es así por la compleja organización repetitiva de los centrómeros que hacen un desafío el poder ensamblar las regiones secuenciadas para formar su composición nucleotídica completa. Hoy en día los esfuerzos por conocer esta información son cada vez mayores considerando que los centrómeros son los *loci* esenciales para la segregación de los cromosomas y la estabilidad del genoma (Aldrup-MacDonald & Sullivan, 2014). Una característica interesante de estas regiones es que si bien las proteínas asociadas a los centrómeros son conservadas entre las especies, es esperado que cada genoma humano contenga un repertorio personalizado en la composición y organización de sus secuencias centroméricas (Hyden KE, 2012).

El grupo de investigación del Dr. Contreras-Galindo perteneciente al Laboratorio de Investigación del Dr. Markovitz de la Universidad de Michigan (UMich) se planteó la posibilidad de encontrar provirus de HERV-K insertos en los centrómeros de los cromosomas humanos, y propuso una estrategia por PCR para ello. Su experimento permitió el hallazgo de varios de cientos de provirus de HERV-K distribuidos en las regiones centroméricas de al menos 15 cromosomas humanos, los cuales denominaron “K111”. Se estima que K111 se integró hace ~2.6-6.3 millones de años atrás, cerca del tiempo de divergencia de Chimpancés-Humanos (Contreras-Galindo R, *et al.*, 2008 y 2013). La evidencia indica que K111 se expandió tanto dentro de los centrómeros como entre centrómeros a través de un proceso semejante al de recombinación homóloga. El hallazgo de que el provirus K111 interactúa con las proteínas del centrómero CENP-A y CENP-B, le agrega un soporte adicional al hecho de que los HERVs son un componente importante para que el “núcleo” centromérico sea funcional (Contreras-Galindo R, *et al.*, 2013).

Destacando el rol de los centrómeros como sitios claves que aseguran una correcta función en la mitosis y la meiosis, es importante considerar que un mal funcionamiento en estos sitios da lugar a una incorrecta segregación de los cromosomas y como consecuencia la pérdida o ganancia de cromosomas (proceso denominado aneuploidía). Algunos factores que podrían derivar en un mal funcionamiento centromérico pueden ser por ejemplo, la pérdida o ganancia de secuencias (al tratarse de una zona repetitiva), o la inserción de elementos transponibles que saltaron a esos sitios del genoma, entre los cuales podrían estar involucrados los K111. A continuación se describen las características principales sobre la estructura y composición de los centrómeros humanos.

1.3. Estructura del ADN centromérico, los repetidos alfa

Los centrómeros humanos están conformados principalmente por una familia de secuencias satélite rica en adeninas y timinas (AT), conocida como satélite alfa o repetidos alfa. Estas secuencias son fundamentales ya que interactúan bioquímicamente con proteínas internas del cinetocoro y por lo tanto juegan un rol importante en la segregación de los cromosomas al momento de la mitosis. Un repetido alfa se define como una unidad o monómero de aproximadamente 171pb. En los centrómeros varios monómeros se repiten en orientación cabeza-cola y, con frecuencia, se extienden con limitadas interrupciones por millones de bases (Aldrup-MacDonald & Sullivan, 2014). La comparación de secuencias entre monómeros individuales revela que se trata de una familia de secuencias altamente divergente, con un porcentaje de identidad de pares de bases de 60-80% (Rudd & Willard, 2004).

Los repetidos monoméricos como tal se encuentran en regiones de heterocromatina pericentromérica. Dentro del “núcleo” centromérico un número n de monómeros de repetidos alfa conforma un bloque o unidad que se repite cabeza-cola conformando lo que se denominan las unidades de repetidos de alto orden (HORs por su sigla en inglés). La organización en HORs da lugar a un arreglo muy homogéneo con un porcentaje de identidad ~97-99% incluso si los monómeros que los constituyen son de elevada divergencia, y conforman una fracción larga ininterrumpida pudiendo alcanzar longitudes desde 250 a 5.000 Kb (Brown & O’Neill 2014; Hyden KE, 2012) (ver Fig. 4).

Todos los centrómeros humanos parecen tener un arreglo en HORs tipo cromosoma-específico. Cabe mencionar que existe variabilidad en el número de secuencias de repetidos de alto orden entre individuos, entre cromosomas e incluso dentro de un mismo cromosoma. La mayoría de los cromosomas humanos tienen un único bloque de HOR que se repite conformando el centrómero, sin embargo, los cromosomas 1, 5, 7, 15, 17, 18 y 19 tienen dos o más HORs diferentes (Brown & O’Neill, 2014; Maloney KA, *et al.*, 2012). Algunos centrómeros presentan una extrema homología inter-cromosomal por lo cual se deberían especificar los cambios únicos de pares de bases para estudiar la actividad centromérica de una manera cromosoma-específica. Esto ocurre por ejemplo en la familia de HOR diméricos en los cromosomas 1, 5 y 19. Lo mismo ocurre para los centrómeros de los cromosomas 13

con el 21 y el 14 con el cromosoma 22 los cuales contienen arreglos de HORs casi idénticos entre sí (Alexandrov I, *et al.*, 2001; Baldini A, *et al.*, 1989; Warburton PE, *et al.*, 2008).

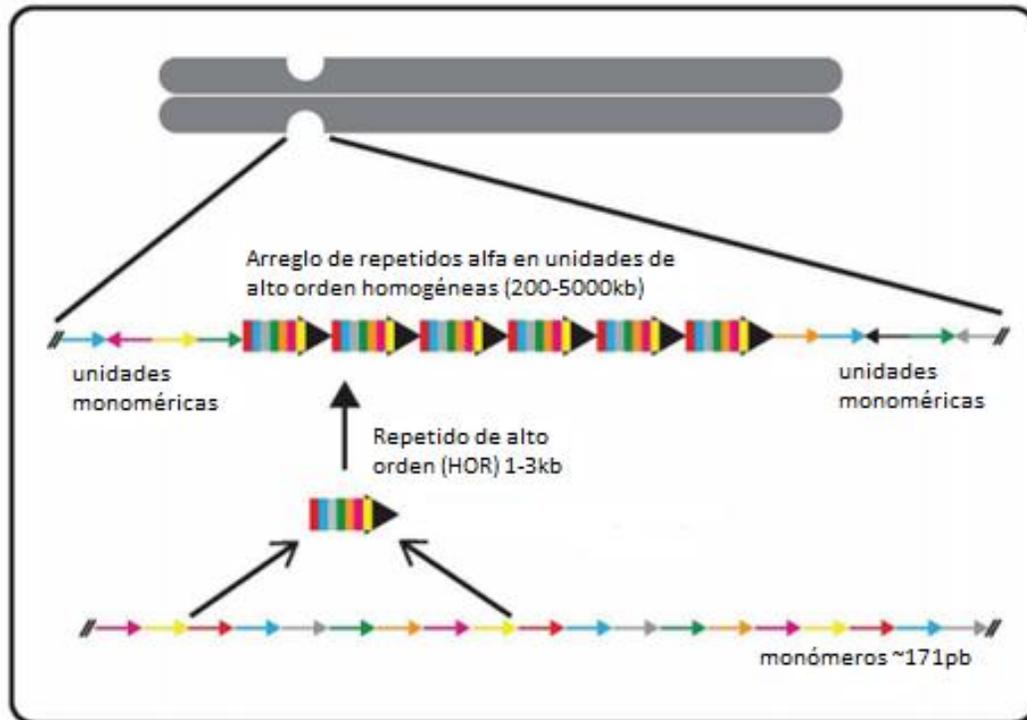


Figura 4. Organización genómica de los centrómeros humanos. Los monómeros de repetidos alfa individuales (flechas simples de colores) son altamente divergentes entre sí alcanzando valores de hasta un 40% de divergencia, mientras que la comparación entre HORs adyacentes (barras de colores con flecha negra) revelan que son cercanamente idénticos (>95% de identidad). La organización en HOR se encuentra en el núcleo centromérico, mientras que monómeros de repetidos alfa individuales se ubican en regiones de cromatina pericentromérica. Figura modificada de Aldrup-MacDonald & Sullivan, 2014.

Las translocaciones, inversiones y deleciones de varias megabases de largo son característicos de la inestabilidad genómica observada en diversas patologías humanas. Muchos tipos de cánceres, incluyendo las leucemias, se definen por la presencia de determinados rearrreglos cromosómicos (Longo MS, *et al.*, 2009). Cada rearrreglo específico requiere de un quiebre de la doble hebra de ADN. Se ha visto que estos quiebres no se generan en zonas del genoma al azar, sino que ocurren en sitios específicos. Estos sitios claves de clivaje son compartidos entre especies por lo que han sido denominados “puntos de ruptura evolutivos” (EB por su sigla en inglés). Lo interesante de los EB es que el estudio de su secuencia reveló que se tratan de una distribución conservada de elementos

repetitivos (repetidos alfa) ricos en secuencias LINES y de retrovirus endógenos (ERVs), es decir muchos de los EB se corresponden con secuencias típicas centroméricas (Murphy WJ, *et al.*, 2005). La presencia de los puntos de ruptura evolutivos derivados de los centrómeros y los centrómeros propiamente dichos se han propuesto como factores claves para la generación de las aberraciones cromosómicas (Longo MS, *et al.*, 2009).

Considerando el potencial que parecen tener los HERVs en contribuir en procesos hemato-oncológicos así como también en enfermedades auto-inmunes actuando como (auto)antígenos, es de gran interés continuar indagando sobre su posible asociación en enfermedades de las cuales aún se sabe muy poco. Asimismo, es importante destacar el hecho de que un mal funcionamiento a nivel centromérico que genere alteraciones cromosómicas podría explicar el origen de varias enfermedades como el cáncer. También genera interés el abordar el estudio de estas regiones con el fin de encontrar alguna marca molecular que puedan indicar la ocurrencia de estos procesos. En el presente proyecto de maestría contemplamos ambas interrogantes para aplicarlas en el estudio de la Leucemia Linfocítica Crónica (LLC), una patología de curso heterogéneo cuya causa aún se desconoce y para la cual hoy en día no existe cura. A continuación se describen aspectos importantes sobre la LLC que serán de utilidad para continuar con el desenlace del proyecto.

1.4. Leucemia Linfocítica Crónica

1.4.1. Generalidades y algunos datos estadísticos

La LLC es la leucemia de mayor frecuencia en el hemisferio occidental y resulta rara en el hemisferio oriental. Comprende alrededor de un tercio de las leucemias que se diagnostican en adultos (Siegel R, *et al.*, 2014). Es una patología que afecta a personas mayores, con un pico de incidencia en el rango que va de los 60 a los 70 años y una prevalencia mayor en hombres que en mujeres. En 2014, fueron diagnosticados 15.720 nuevos casos de LLC y se documentaron 4.600 muertes en EEUU causadas por esta afección. El riesgo de desarrollar la enfermedad es de aproximadamente uno en doscientos, y factores como tener una historia familiar de LLC pueden aumentar este valor (American Cancer Society, 2014).

Se desconoce aún cuál es el agente causal de la enfermedad y no hay hasta la fecha una evidencia consistente que vincule a la LLC con una exposición ambiental como radiación,

ciertos químicos o con alguna infección viral determinada (Dighiero & Hambling, 2008; Rego N, *et al.*, 2012).

Su evolución sigue un curso sumamente heterogéneo que puede durar de meses a años. Un tercio de los pacientes vive durante años con una enfermedad estable y sin medicación. Otro tercio presenta una fase inicial indolente pero más tardíamente la enfermedad progresa. El tercio restante de los pacientes exhibe una enfermedad agresiva desde el comienzo que necesita de un tratamiento inmediato (Vasconcelos Y, *et al.*, 2003).

Los progresos obtenidos en el tratamiento para la LLC hacen que frecuentemente se observen remisiones, incluso a nivel molecular, pero la gran mayoría de los pacientes recaen y existe un consenso de que esta enfermedad es incurable (Palacios F, *et al.*, 2010).

1.4.2. Inmunofenotipo de las células B de LLC

La LLC se caracteriza por una acumulación lenta y progresiva de células B monoclonales, pequeñas y de apariencia madura en la sangre periférica (SP), médula ósea (MO), y órganos linfoides secundarios (OLS). Estas células B se caracterizan por la expresión de antígenos pan-B como CD19, CD20 y CD40, junto con otros antígenos que no se encuentran en la mayoría de los linfocitos B circulantes como CD5 y CD23 (Hamblin TJ, 2005). Como todo linfocito B maduro, las células de LLC expresan inmunoglobulinas (Igs) en la superficie de su membrana, las cuales forman parte del receptor para el antígeno (BCR), pero lo característico de ellas es que lo hacen en un bajo nivel de expresión (Vuillier F, *et al.*, 2005). También se ha señalado que las células de LLC circulantes en sangre periférica se encuentran detenidas en la fase G0-G1 de su ciclo celular (Hernández P, 1999).

1.4.3. Subpoblación de linfocitos proliferantes en LLC

La visión tradicional indica que la vida de las células B se ve prolongada por un defecto en la apoptosis resultando en la acumulación de estos linfocitos B en forma lenta y progresiva en SP (Dameshek W, 1967). El incremento en la expresión de la proteína Bcl-2 y el bloqueo de estas células en su fase G0 sustentan esta observación. Sin embargo, estudios recientes describen que la LLC es un proceso dinámico el cual resulta en células que proliferan y mueren, a menudo a niveles apreciables. Esta observación permitió la postulación de una interesante hipótesis en la cual se contempla, la existencia de al menos dos tipos de poblaciones dentro del clon tumoral de un paciente con LLC. La primera es la población

mayoritaria, y presenta las características de una célula B quiescente, no proliferante que lleva a la acumulación linfocitaria típica de esta leucemia en SP. La segunda población, minoritaria se compone de linfocitos B tumorales con una alta capacidad proliferante en donde se mantienen interacciones con el microambiente celular que determinan una alta proliferación celular, así como la sobreexpresión de numerosos marcadores moleculares de agresividad tumoral. Esta última población ha sido correlacionada con una forma más agresiva de la leucemia, lo cual indica que las diferencias en el curso clínico de los pacientes pueden, al menos en parte, ser determinadas por la capacidad de proliferar de las células B de LLC. Es de suponer que el conjunto de señales y condiciones microambientales determinan el equilibrio entre la resistencia a la apoptosis y la proliferación activa de los linfocitos B (Chiorazzi N, 2007; Deaglio & Malavasi, 2009; Palacios F, *et al.*, 2010; Pepper C, *et al.*, 2007).

Las células de LLC circulantes, que como mencionamos anteriormente se encuentran en el estadio G0 del ciclo celular, se mantienen e incrementan gracias al aporte de nuevos linfocitos del clon leucémico que prolifera principalmente en los nódulos linfáticos pero también en médula ósea y bazo, dentro de estructuras histológicas que reciben el nombre de centros proliferantes o pseudofolículos. Las células B leucémicas de estos centros proliferantes están en íntimo contacto con linfocitos T CD4+ activados (TCD4+), así como también con una variedad de células estromales y células nodrizas. El conjunto de estas células parecen tener un rol central en amplificar las condiciones del microambiente capaz de favorecer la proliferación dentro de los pseudofolículos y de inhibir la apoptosis de las células malignas (Caligaris-Cappio F, 2003) (ver Fig. 5). Se ha visto que la presencia de células B leucémicas en los pacientes con una enfermedad progresiva se relaciona con una fuerte activación de la vía de señalización celular AKT (Longo PG, *et al.*, 2007).

Estudios recientes muestran que la presencia de esta subpoblación de linfocitos B altamente proliferante puede encontrarse en SP donde se ha detectado una alta expresión de la enzima citidina deaminasa inducida por activación (AID) (Palacios F, *et al.*, 2010). Esta subpoblación estaría llevando a cabo un proceso activo de cambio de clase IgM a IgG (por lo que son IgG positivos), y presentan aumentados niveles de expresión de moléculas características de una célula B en estado de proliferación. Se detectó también una sobreexpresión de varias proteínas relacionadas a la progresión tumoral y a la interacción

con el microambiente en los centros proliferantes como son la glicoproteína CD49d y las quimioquinas CCL3 y CCL4 (Oppezzo P, *et al.*, 2003; Palacios F, *et al.*, 2010; Perez-Duran P, *et al.*, 2007). Todas estas observaciones hacen pensar que la acumulación *in vivo* de los linfocitos leucémicos se encuentra facilitada por la interacción de las células de LLC con otras células (células estromales, células T activadas) y factores solubles dentro del microambiente celular adecuado (médula ósea, y órganos linfoides secundarios), especialmente a través de interacciones clásicas del tipo receptor-ligando (ejemplo CD40-CD40L), interacciones quimioquinas-receptor, así como también mediante la posible unión de un (auto)antígeno al receptor de las células B (BCR) (Ghia P, *et al.*, 2008) (ver Fig. 5).

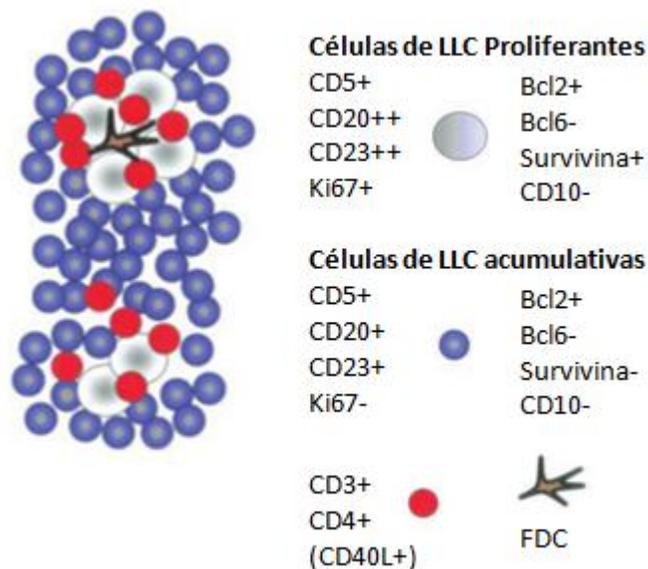


Figura 5. Características generales de los pseudofolículos en LLC. En la figura se representan diferentes propiedades fenotípicas que manifiestan la población de células B quiescentes (fase G0-G1) y la subpoblación de células B proliferantes. Esta última población está en íntimo contacto con los linfocitos T CD4+ activados (representados en rojo) y de células estromales y nodrizas FDC: Células dendríticas foliculares. Figura modificada de Caligaris-Cappio F, 2003.

1.4.4. Marcadores pronósticos de la LLC

Existen dos sistemas de clasificación anatomopatológicas de la LLC muy difundidos a nivel internacional los cuales han sido descritos por Rai (Rai KR, *et al.*, 1975) y Binet (Binet JL, *et al.*, 1997). Ambos sistemas toman en cuenta las siguientes características clínicas: presencia

de linfocitos en sangre y médula ósea, linfo-adenopatías, esplenomegalia, hepatomegalia, anemia y/o trombocitopenia. Aun cuando estos sistemas son relevantes desde el punto de vista clínico, existen ciertos subgrupos de pacientes que difícilmente pueden ser ubicados dentro de un estadio particular. De allí la importancia que tiene la identificación de marcadores biológicos e inmunológicos con valor predictivo que posibiliten un mejor tratamiento terapéutico.

La mayoría de las células B de LLC expresan CD5 e IgM/IgD, y por lo tanto, tienen un fenotipo inmaduro que, en condiciones normales, expresan genes de inmunoglobulinas no mutados. Sin embargo, el 50-70% de los casos de leucemia linfocítica crónica tienen mutaciones somáticas en los genes que codifican para las cadenas pesadas de las zonas variables de las inmunoglobulinas (IgVH) como si hubieran madurado en los centros germinales del folículo linfocítico (Dighiero & Hamblin, 2008; Schroeder & Dighiero, 1994). Sobre la base de estas observaciones, se puede clasificar a los pacientes con LLC en dos grupos de acuerdo a su estatus mutacional de las Ig: aquellos que poseen IgVH no mutadas (NM) y aquellos con IgVH mutadas (M). Se ha notado que pacientes con un perfil de genes IgVH-NM desarrollan una enfermedad progresiva, mientras que pacientes con genes IgVH-M manifiestan un curso indolente y frecuentemente mueren por enfermedades no relacionadas (Bianchi S, *et al.*, 2010; Zenz T, *et al.*, 2010). Existen importantes diferencias biológicas entre estos dos grupos. Los pacientes de LLC perfil IgVH-NM presentan mayores niveles de expresión de la proteína tirosin-quinasa asociada a Z-70 (ZAP-70) y del antígeno CD38, respecto a los pacientes con LLC de perfil IgVH-M. Asimismo, el grupo de IgVH-NM activa vías de señalización celulares claves en respuesta a la activación del BCR (por ejemplo las vías LYN, SYK, ERK y AKT) y manifiestan un perfil de expresión de genes diferente a los pacientes de LLC perfil IgVH-M. Cada vez hay más evidencias de que los pacientes IgVH-NM presentan una capacidad celular proliferativa significativamente mayor que los pacientes IgVH-M, y que esta mayor capacidad es fundamentada por presentar un acortamiento en los telómeros de las células involucradas (Zenz T, *et al.*, 2010). Estudios recientes demuestran que la determinación simultánea de ZAP-70 y CD38 posee un buen valor pronóstico de la enfermedad que complementa al del perfil mutacional de las IgVH (Hamblin TJ, *et al.*, 2002; Rassenti LZ, *et al.*, 2004) (ver Fig. 6)

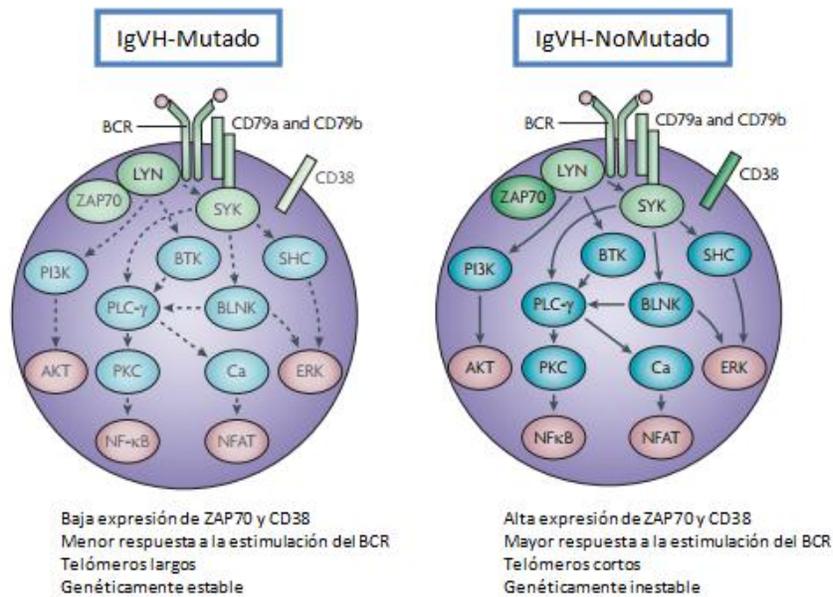


Figura 6. Características biológicas que diferencian las células B de LLC con perfil IgVH Mutado y No Mutado (M y NM). Las células de LLC perfil IgVH-NM presentan niveles mayores de expresión de las proteínas ZAP70 y de CD38 (representadas en verde oscuro) que las células de LLC perfil IgVH-M (representadas en verde claro). Las células de IgVH-NM activan vías de señalización celular claves en respuesta a la activación del BCR (por ejemplo LYN, SYK, ERK y AKT) y muestran un perfil de expresión de genes diferente que las células de LLC IgVH-M. La señalización reducida en IgVH-M es indicada con texto en gris y flechas punteadas. La capacidad proliferante de las células B entre ambos grupos es distinta y esta característica puede fundamentarse en las diferencias encontradas en el largo de sus telómeros. Figura modificada de Zenz T, *et al.*, 2010.

Las células de LLC pueden manifestar también aberraciones cromosómicas diferenciales que se detectan, en la mayoría de los casos, en células en interfase por hibridación fluorescente *in situ* (FISH). Las aberraciones cromosómicas más frecuentes son deleción en cromosoma 13q, trisomía 12, y deleción en cromosoma 11q. En menor frecuencia se observa una deleción en el cromosoma 17p que afecta la proteína TP53. La presencia de la deleción en 17p u 11q se asocia a una mal prognosis y predominan en estados avanzados de LLC y en pacientes con genes IgVH-NM. La deleción en 13q o un cariotipo normal son asociadas con una buena prognosis, estados iniciales de la enfermedad, y con genes IgVH-M. Existen controversias sobre si la trisomía 12 se corresponde a un estatus no mutado y a una mal prognosis de la enfermedad (Dighiero & Hamblin, 2008; Kröber A, *et al.*, 2002). La composición de los centrómeros de los cromosomas humanos puede brindar una importante información sobre cómo será la segregación de los mismos al momento de la división celular (por mitosis o meiosis). Una segregación incorrecta de los cromosomas dará lugar a células

hijas con diferentes cariotipos que, de ser viables, pueden derivar en la formación de células cuyas aberraciones cromosómicas correspondan a las observadas en LLC.

Cabe mencionar que si bien se han logrado importantes avances en materia de conocimiento de la LLC, ni los sistemas de Rai/Binet ni los marcadores moleculares descriptos, son capaces de predecir completamente la progresión de la enfermedad y/o explicar el perfil de progresión heterogéneo en el curso clínico de la LLC (Palacios F, *et al.*, 2010).

Teniendo en cuenta lo descripto en el marco teórico, el presente proyecto de maestría se organiza con el fin de cumplir con dos objetivos generales complementarios. El primero de ellos fue abordado en el Laboratorio de Virología Molecular del Centro de Investigaciones Nucleares de la Facultad de Ciencias, UdelaR. El segundo fue realizado en el marco de una pasantía de investigación de seis meses en el Laboratorio de Investigación del Dr. Markovitz en la Universidad de Michigan, EEUU. A continuación se detallan los objetivos generales y específicos propuestos en esta tesis.

2. OBJETIVOS

2.1. Objetivo General 1

Estudiar la posible relación entre la expresión de los genes de HERV-K en la patología hemato-oncológica Leucemia Linfoide Crónica (LLC).

2.1. Objetivos Específicos 1

1. Puesta a punto de la metodología de RT-PCR para la obtención de todos los genes de HERV-K (*env*, *gag*, *pol*, *rec* y *np9*)
2. Cuantificar los niveles de expresión relativa de los genes *np9* y *gag* de HERV-K mediante PCR en tiempo real en muestras de pacientes con LLC (perfil IgVH-M y IgVH-NM) con respecto a muestras de donantes sanos.
3. Cuantificar la expresión del gen celular *c-myc* en las muestras de LLC y en muestras de donantes sanos y evaluar los niveles de expresión del gen *np9* de HERV-K.
4. Estudiar la variabilidad genética del gen *np9* en pacientes con LLC y en donantes sanos mediante aproximaciones bio-informáticas.

2.2. Objetivo General 2

Determinar la abundancia de repetidos alfa y secuencias de HERVs (K111) en los centrómeros de una línea celular de LLC (AT-CLL) en comparación con donantes sanos.

2.2. Objetivos Específicos 2

1. Puesta a punto de la metodología de PCR a tiempo real para cuantificar las secuencias centroméricas de repetidos alfa de manera específica para cada cromosoma humano.
2. Cuantificar la abundancia de repetidos alfa y de retrovirus endógenos humanos (K111) en los centrómeros de la línea celular AT-CLL y en células humanas sanas. Evaluar si existen diferencias significativas entre ambos grupos.

3. MATERIALES Y MÉTODOS

3.1. Metodología experimental para abordar el objetivo general 1

El estudio de la expresión de genes de HERVs en muestras de sangre periférica de pacientes con leucemia linfocítica crónica (LLC) y donantes sanos se realizó siguiendo la metodología experimental que se presenta a continuación:

3.1.1. Muestras de estudio

Se recolectaron 25 muestras de sangre periférica de pacientes diagnosticados con leucemia linfocítica crónica (promedio de edad: 65 años): 14 presentaban un perfil IgVH Mutado y 11 un perfil No Mutado (M y NM respectivamente). Las muestras fueron cedidas por el Servicio de Hematología del Hospital Maciel de Montevideo y se cuenta con el consentimiento informado de los pacientes del Hospital en acuerdo con las regulaciones de Ética Nacional y la Declaración de Helsinki. Como grupo control se procesaron 6 muestras de sangre periférica proveniente de donantes sanos, las cuales fueron cedidas por el Servicio Nacional de Sangre.

Las muestras fueron dirigidas a la Unidad de Proteínas Recombinantes (UPR) del Instituto Pasteur de Montevideo (IPMont). Allí se realizó el aislamiento de las células monoclonales de sangre periférica (PBMCs) mediante el método de centrifugación en gradiente de Ficoll-Hypaque (Pharmacia Fine Chemicals), y posteriormente se realizó la extracción de ARN siguiendo el procedimiento descrito por Oppezco y colaboradores en el 2002 (Oppezco P, *et al.*, 2002). El ARN celular fue enviado al Laboratorio de Virología Molecular (LVM) del Centro de Investigaciones Nucleares (CIN) donde se continuó con el procedimiento para el presente proyecto.

Como controles positivos de expresión de los genes de HERV-K se utilizaron las líneas celulares leucémicas Daudi (línea linfoblástica de células B proveniente de un linfoma de Burkitt) y Jurkat (línea linfoblástica proveniente de leucemia aguda de linfocitos T) (Renaudineau Y, *et al.*, 2005). Estas líneas celulares fueron cedidas por la UPR-IPMont y posteriormente fueron crecidas en medio RPMI suplementado con 10% de suero fetal bovino (SFB). La extracción de su ARN se realizó siguiendo el método clásico de aislamiento del ARN en Trizol-Cloroformo (Chomczynski & Sacchi N, 2006)

3.1.2. Cuantificación de las concentraciones de ARN y tratamiento con DNAsa

La cuantificación de las concentraciones en ng/μl del ARN de las muestras descriptas (PBMCs de pacientes con LLC perfil IgVH-M y NM, PBMCs de donantes sanos y las líneas celulares leucémicas Daudi y Jurkat) se realizó utilizando un espectrofotómetro Eppendorf BioSpectrometer®.

Con el fin de minimizar una posible contaminación con ADN genómico, a cada muestra de ARN se le realizó un tratamiento con DNasa utilizando el Kit DNA-free™ (Ambion, USA), siguiendo las instrucciones del protocolo.

El rendimiento de la extracción y posterior tratamiento con DNAsa se determinó cuantificando nuevamente las concentraciones en ng/μl de cada ARN ahora tratado con la enzima. La cuantificación fue realizada en la plataforma Eppendorf BioSpectrometer®. La pureza del ARN se evaluó mediante el cociente de absorbancia A260/A280 (aceptándose un coeficiente de ~2) y A260/A230 (aceptándose un coeficiente de 1,8 – 2) para determinar la ausencia de contaminantes. Los valores de concentración y cocientes de absorbancia fueron comparados con los valores obtenidos del ARN previo al tratamiento con la enzima DNAsa.

La eficiencia del tratamiento con DNAsa fue evaluada por medio de la amplificación de los genes de HERV-K utilizando como molde el ARN post tratamiento con DNAsa no observándose señal positiva.

3.1.3. Retrotranscripción del ARN (RT-PCR)

La síntesis de ADN copia (ADNc) fue realizada a partir de 2μg de ARN, utilizando como *primers* múltiples hexámeros y la enzima Super Script II Reverse Transcriptase (Invitrogen) en un volumen final de reacción de 20μl, tal como fue descrito por Oppezco y colaboradores (2005).

3.1.4. Puesta a punto de la PCR cualitativa de todos los genes de HERV-K

A continuación se realizó la puesta a punto de las condiciones de tiempo y temperatura para la amplificación por reacción en cadena de la polimerasa (PCR) de todos los genes de HERV-K (*env*, *gag*, *pol*, *rec* y *np9*). La PCR fue realizada utilizando la enzima *Taq* DNA Polimerasa (Invitrogen by Life Technologies). Cada reacción de PCR se realizó en un volumen final de 50 μl conteniendo 2μl (un décimo de la reacción de RT-PCR) de ADNc de la línea celular Jurkat, 5μl de buffer de PCR 10x, 1,5μl de MgCl₂ (50mM), 1μl de mezcla de deoxinucleótidos dNTPs (10mM), 2μl de *primer* forward (10μM), 2μl de *primer* reverse

(10µM), 0,5µl de la enzima *Taq* DNA Polimerasa (5U/µl), y agua destilada miliQ hasta completar los 50µl de volumen de reacción final. Los pares de *primers* utilizados para amplificar los genes de HERV-K, así como los tamaños en pares de bases de las bandas esperadas se detallan en la Tabla 1 de esta sección (ver Tabla 1). Cada reacción tuvo un control negativo con agua en lugar de ADNc de Jurkat.

Las condiciones de amplificación fueron las siguientes: un paso inicial de 10 minutos (min) a 95 grados (°C), seguido de 35 ciclos compuestos de 45 segundos (seg) de desnaturalización a 95°C, 45 seg de hibridación de los *primers* a 58°C (para los genes *gag*, *rec*, *np9*) o 53°C (para los genes *pol* y *env*), y 45 seg de elongación a 72°C. La PCR terminó con un tiempo de extensión final de 10 min a 72°C.

Una vez puestas a punto las condiciones de amplificación, se repitieron las PCRs para cada gen de HERV-K pero ahora añadiendo 2µl de los siguientes ADNc de muestra como molde: una muestra de LLC perfil IgVH-Mutado, una muestra de LLC perfil IgVH-NoMutado, una muestra de donante sano (seleccionadas al azar), y agua como control negativo. Los productos de amplificación fueron corridos electroforéticamente en un gel de agarosa al 1,5% durante 40 minutos a 100V. El gel fue teñido con el intercalante Bromuro de Etidio (EtBr) previo a la corrida. El tamaño de los productos de amplificación se comparó con la escalera de peso molecular (PM) de 100pb (Gene Ruler 100pb DNA ladder, Thermo Scientific).

3.1.5. PCR para la expresión del gen constitutivo glicer-aldehído 3-fosfato deshidrogenasa (*gapdh*).

Con el fin de evaluar la integridad de nuestro ADNc obtenido mediante la transcripción reversa, se realizó un control interno para cada uno de los ADNc con un juego de *primers* que amplifica el gen constitutivo celular que codifica para la proteína glicer-aldehído 3-fosfato deshidrogenasa (*gapdh*) (Palacios F, *et al.*, 2010). Para esta reacción se utilizaron las mismas concentraciones de reactivos y los mismos tiempos y temperaturas de ciclado que los empleados en la PCR de *gag*, *rec* y *np9* de HERV-K.

Tabla 1. Secuencia de *primers* y tamaños de los productos de amplificación esperados para los genes de HERV-K y *gapdh*.

Primer	Secuencia Primer	Tamaño banda	Fuente bibliográfica
gag-F	5'-GGGTGATAATGGGGCAAATAAA-3'	~ 470pb	Depil S, <i>et al.</i> , 2002.
gag-R	5'-GACCTGCTGGAAGAGGACTTGT-3'		
pol-F	5'-TCCCCTTGAATACTCCTGTTTTYGT-3'	~294pb	Medstrand P, <i>et al.</i> , 1992.
pol-R	5'-CATTCTTGTGGTAAAACCTTTCCAYTG-3'		
env-F	5'-ACAAAAAGAAGGGGGAGACAG-3'	~1500pb	Zhao J, <i>et al.</i> , 2011.
env-R	5'-TGAACAGAAGAGTGCAATGCA-3'		
rec-F	5'-ATCGAGCACCGTTGACTCACAAGA-3'	~200pb	González-Hernandez MJ, <i>et al.</i> , 2012.
rec-R	5'-GGTACACCTGCAGACACCATTGAT-3'		
np9-F	5'-AGATGTCTGCAGGTGTACCCA-3'	~150pb	González-Hernandez MJ, <i>et al.</i> , 2012.
np9-R	5'-CTCTTGCTTTTCCCCACATTTTC-3'		
gapdh-F	5'-GGTGCTGAGTATGTCGTGGA-3'	~350pb	Palacios F, <i>et al.</i> , 2010.
gapdh-R	5'-ATGCCAGTGAGCTTCCCGTT-3'		

F: *primer* forward, R: *primer* reverse, pb: pares de bases.

3.1.6. PCR a tiempo real para los genes *np9* y *gag* de HERV-K

Con el fin de realizar el análisis de expresión relativa para los genes *np9* y *gag* de HERV-K en las muestras de pacientes con LLC y las muestras de donantes sanos, se realizó una PCR a tiempo real para estos genes. Se utilizó el kit KAPA SYBR FAST qPCR Kit Master Mix (2x) Universal (Kapa Biosystems) y el termociclador Corbett Rotor-Gene 6000. Las secuencias de *primers* de *np9* y *gag* utilizadas fueron previamente descritas por el grupo de Gonzalez-Hernández en 2012 y el grupo de Depil y colaboradores en 2002 respectivamente (Depil S, *et al.*, 2002; Gonzalez-Hernandez MJ, *et al.*, 2012), los cuales fueron puestos a punto *primero* en la PCR convencional. Los valores de expresión fueron normalizados con el gen de referencia (*gapdh*). Para el gen *gapdh* fue utilizado el set de *primers* empleado en la PCR convencional (ver Tabla 1).

La mezcla de reacción para cada PCR fue preparada en un volumen final de 20 μ l conteniendo: 10 μ l del Master Mix Kapa Sybr Fast (2x), 0,4 μ l de *primer* forward (10 μ M), 0,4 μ l de *primer* reverse (10 μ M), 2 μ l de ADNc, y agua miliQ.

Las condiciones de amplificación para el gen *np9* consistieron en: 3 min a 95°C seguido de 40 ciclos compuestos de 3 seg a 95°C, 20 seg a 58°C y 20 seg a 72°C.

Las condiciones de ciclado de la PCR a tiempo real para el gen *gag* consistieron en: 10 min a 95°C seguido de 40 ciclos compuestos de 15 seg a 95°C y 1 min a 60°C. Cada reacción fue realizada por duplicado, y la especificidad de los productos amplificados fue verificada por el análisis de sus respectivas curvas de melting provistas por el software Rotor-Gene Q-Series.

Con el fin de conocer cuánto difiere la expresión de los genes *np9* y *gag* de cada muestra de LLC (perfil IgVH-M y NM) con respecto a los donantes sanos, se utilizó el método del Ct comparativo o también conocido como el método de $2^{-\Delta\Delta Ct}$.

El Ct (por su sigla en inglés Cycle threshold) es el número de ciclo en que la muestra supera un valor umbral (threshold) en la señal de fluorescencia el cual es determinado (manual o automáticamente) previamente (Schmittgen & Livak, 2008). El cálculo se realizó de la siguiente manera:

$$\Delta Ct \text{ LLC} = Ct \text{ muestra de LLC (para el gen } np9 \text{ o } gag) - Ct \text{ muestra de LLC del gen calibrador (} gapdh)$$

$$\Delta Ct \text{ sano} = Ct \text{ muestra de sano (para el gen } np9 \text{ o } gag) - Ct \text{ muestra de sano del gen calibrador (} gapdh)$$

Luego se calcula:

$$\Delta\Delta Ct = \Delta Ct \text{ LLC} - \Delta Ct \text{ promedio sanos}$$

Y a continuación:

$$2^{-\Delta\Delta Ct} = 2 \text{ elevado a la } -\Delta\Delta Ct$$

Se graficó la variabilidad y la media entre los dos valores de $2^{-\Delta\Delta Ct}$ obtenidos para cada muestra de LLC (por su realización en duplicado) considerando el promedio de expresión del gen *np9* o *gag* de los donantes sanos como nivel basal de expresión. Se realizó un control positivo con ADNc de las líneas celulares Daudi y Jurkat, y un control negativo con agua miliQ en lugar de molde de ADNc.

Para detectar la existencia de diferencias significativas entre las medias de la expresión relativa del gen de interés (*np9* o *gag*) del grupo de pacientes de LLC con perfil IgVH-M con respecto al grupo de pacientes de LLC perfil IgVH-NM se utilizó el test T no-pareado de dos colas con un intervalo de confianza del 95%. Se utilizó el software GraphPad versión 5.00

para Windows, San Diego California USA, para la realización de las gráficas de expresión y para los cálculos estadísticos.

3.1.7. PCR a tiempo real para cuantificar expresión relativa del gen celular *c-myc*

Se utilizó el kit comercial de KAPA SYBR FAST qPCR Kit Master Mix (2x) Universal (Kapa Biosystems) y la plataforma Corbett Rotor-Gene 6000 para el estudio del nivel de expresión relativa del gen celular *c-myc* en las muestras de pacientes con LLC y en las muestras de donantes sanos. Los *primers* para la amplificación de el gen *c-myc* fueron los siguientes: forward: 5'-CTT TGT GTG CCC CGC TCC AG-3' y reverse: 5'-GCG CTC AGA TCC TGC AGG TA-3' (Palacios F, *et al.*, 2010). El estudio se realizó por duplicado y en cada mix de reacción se añadió 1µl de ADNc para cada muestra. Las condiciones de amplificación fueron las siguientes: 3 min a 95°C y 40 ciclos de 3 seg a 95°C, 20 seg a 58°C, y 20 seg a 72°C. El nivel de expresión del gen *c-myc* se normalizó con el nivel de expresión del gen constitutivo *gapdh* siguiendo la fórmula que se expresó anteriormente.

Los resultados de este estudio se organizaron dividiendo las muestras en tres grupos: muestras de LLC cuyo nivel de expresión del gen *np9* fue elevado (mayor a 5 veces respecto al nivel de expresión de los donantes sanos), muestras de LLC cuyo nivel de expresión del gen *np9* fue bajo (menor a 5 veces), y el grupo de las muestras de los donantes sanos. Para detectar la existencia de diferencias significativas entre las medias de la expresión relativa del gen celular *c-myc* en los tres grupos establecidos, se utilizó el test ANOVA de una vía (análisis de varianza de una vía). Como método de corrección se utilizó el test de Bonferroni para comparaciones múltiples y un intervalo de confianza del 95%. Se utilizó el software GraphPad versión 5.00 para Windows, San Diego California USA, www.graphpad.com, para la realización de las gráficas de expresión y para los cálculos estadísticos.

3.1.8. Amplificación por PCR convencional y secuenciación del gen *np9*

Se realizó una PCR convencional con el fin de amplificar la secuencia completa del gen *np9* de HERV-K utilizando la enzima *Taq* DNA polymerase (Invitrogen). Se seleccionó un juego de *primers* descrito previamente el cual produce un fragmento de amplificación de 256pb (Armbruster V, *et al.*, 2002), lo que nos permite tener el ARNm de *np9* en su totalidad. La mezcla de reacción fue realizada siguiendo el protocolo de Armbruster y cols. (2002). Las condiciones de ciclado fueron las siguientes: 10min a 95°C, seguido de 35 ciclos compuestos de una fase de desnaturalización de 45 seg a 95°C, una de hibridación de los *primers* de 45 seg a 58°C, y una fase de elongación de 1 min a 72°C. El proceso continuó con una etapa de

extensión final de 10 min a 72°C. La reacción fue realizada para 4 muestras de ADNc de donantes sanos, 6 muestras de ADNc de pacientes con LLC perfil IgVH-NM, y 4 muestras de ADNc de LLC perfil IgVH-M. Los productos de amplificación fueron observados mediante una corrida electroforética en gel de agarosa al 2% teñido previamente con Syber Green. Se procedió a la purificación de los productos de PCR siguiendo las instrucciones del kit de purificación de QiaGen (QIAquick PCR Purification Kit, QIAGEN). Se cuantificó la concentración en ng/μl mediante el uso de la plataforma Eppendorf BioSpectrometer®. Los productos de PCR purificados fueron enviados para su secuenciación bidireccional al servicio de Macrogen, Korea (<http://dna.macrogen.com>).

Una vez obtenidas las secuencias, las mismas fueron editadas utilizando el programa SeqMan. Se realizó el alineamiento y análisis bioinformático de las secuencias de *np9* utilizando la plataforma del programa MEGA versión 5.0. Como referencia para el alineamiento y posterior análisis bio-informático se encontraron únicamente dos secuencias patentadas de *np9* en la base de datos del National Center of Biological Investigation (NCBI) (<http://www.ncbi.nlm.nih.gov>) (códigos de acceso: JA719353.1 y CS435178.1), y una secuencia publicada por el grupo de Wang-Johanning y colaboradores (Wang-Johanning F, *et al.*, 2007).

3.2. Metodología experimental para abordar el objetivo general 2

Con el objetivo de determinar la existencia de algún patrón diferencial en el arreglo centromérico definido por la abundancia de secuencias de repetidos alfa y secuencias de HERVs (K111) en la línea celular de LLC (CLL-AT) y su comparación con células humanas normales, se estableció la siguiente metodología experimental:

3.2.1. Diseño de *primers* para la amplificación de los repetidos alfa cromosoma-específicos

Se comenzó por el diseño bioinformático de los *primers* que permiten la amplificación de los repetidos alfa de cada uno de los cromosomas humanos en forma independiente. Este paso resultó ser un gran desafío a la hora de definir las secuencias de *primers* que permitieran diferenciar los centrómeros de aquellos cromosomas con un muy elevado porcentaje de similitud en su secuencia nucleotídica (ver Tabla 2). El diseño de los *primers* y el análisis de su secuencia fue realizado utilizando las herramientas “*Primer Quest*” y “*Oligo Analyzer*” provistas por el programa bioinformático Integrated DNA Technologies (IDT) (www.idtdna.com). En los momentos en que fuera necesario los *primers* se diseñaron por inspección visual con el fin de primar secuencias específicas de un repetido alfa y no de otro.

El juego de *primers* P82H permite la amplificación de los repetidos alfa centroméricos de todos los cromosomas humanos juntos.

Tabla 2. Nombre y secuencia de los *primers* diseñados para amplificar los repetidos alfa cromosoma-específicos.

Cromosoma	Primer	Secuencia Primer	Cromosoma	Primer	Secuencia Primer
1	D1Z5Fc	GGCCTATCGTCGTAAGGAAATA	14	D14Z1/D22Z1F	CAATCTCAGAATCTTCTTTGGGATA
	D1Z5Rc	ATGCTCAGCTCTGTGAGTAAA		D14Z1/D22Z1R	CCAAGCTATCCAAATATCCACTT
2	D2Z1F	TCGTTGGAAACGGGATTGT	15	D15Z3F	GACATTTGGATAGCTTTGAGGATTT
	D2Z1R	CTGCTCTATGAAAGGGACTGTT		D15Z3R	GGGCATACATCACAAGGAAGA
3	D3Z1F	GCTTTGAGCCAATGGTAGA	16	D16Z2F	CCTTCTTTAGACAGAGCAGATT
	D3Z1R	GTTGAACACACACGTACCAAAG		D16Z2R	AGTGTCTGGGAACGAGTTTG
4	D4Z1F	CTGTAGTATCTGGAAGTGGACATT	17	D17Z1F	GTGGAGATATGGACCGCTTTAG
	D4Z1R	GGTTCAACTGTGTTCTTTAGG		D17Z1R	CTCAGTCGTACCAAGAGTTT
5	D5Z1F	AGTCTGCACGTGGATAAGTTG	18	D18Z2F	TTCGATCGATTTACAGCCTATG
	D5Z1R	AAAGAGTGTTAGAAGTCTGCTCTG		D18Z2R	TTGAGGACACACATCACAAGA
6	D6Z1F	ACGGAAGCAATCTCAGAACTAC	19	D19Z5F	GCCTCAATGGGTTCCAGAAATG
	D6Z1R	CCTCAAGGCGGTCCAATTAT		D19Z5R	TGGATCCATCTCACAGATTTCA
7	D7Z2F	CGACTTTGTGATGTGTCATTC	20	D20Z2F	TGCTTGGAAACGGGAATGT
	D7Z2R	CCTTATCCGCAATGGTCCTAAA		D20Z2R	CCTGCTCTATGAAAGGGAATGT
8	D8Z2F	ACGTACACAGCAGCATACTC	21	D21Z1F	TGATGTGTGTACCCAGCT
	D8Z2R	GTCCACTCCAGATACTCCAAA		D21Z1R	GCTATCCAAATATCCACT
9	D9Z4F	GGAGAAGCATTCTCAGAACTT	22	D22Z4F	CCCTTGAACCTGTGGTGAATA
	D9Z4R	GTCCGCTTGACAGATACTACAG		D22Z4R	ACAGTGAGATGAAGGCACAC
10	D10Z1F	TTGGAAACGGGATTTCTCATA	X	DXZ1F	CGGGATCACCTTCCCATAAC
	D10Z1R	GCTCTCTCTAAAGGAAGTTCAA		DXZ1R	GGTGTGCAAACCTGAACTATC
11	D11Z1F	CTTCTTCGAAACGGGTATATCT	Y	DYZ3F	CTGCAAGTGGAAATTGGGAAATA
	D11Z1R	GCTCCATCAGCAGGATTGT		DYZ3R	GAATGCACACAGCACAAGAA
12	D12Z3F	GATGAAGGAGTTGGAGACT	Todos	P82HF	ATGTTTGCATTCAACTCACAGAG
	D12Z3R	CTGTGCAACATTACAGGAAGAAATC		P82HR	CAACACAGTCCAAATATCCAGTTG
13	D13Z1F	TGATGTGTGTACCCAGCT			
	D13Z1R	GCTATCCAAATATCCACT			

F: forward, R: reverse.

3.2.2 Puesta a punto de la PCR a tiempo real de los repetidos alfa cromosoma-específicos

A continuación se realizó la puesta a punto de las condiciones de tiempo, temperatura y ciclos de amplificación que nos permitieran obtener la gran mayoría de los repetidos alfa por PCR a tiempo real. En cada reacción de PCR se analiza un juego de *primers* por separado, es decir, cada reacción evalúa si efectivamente se amplifica la secuencia del repetido alfa para el cromosoma que fue diseñado y no así en los otros cromosomas.

Para la amplificación por PCR a tiempo real se utilizó el kit FastStart Universal SYBR Green Master (Life Science, Roche). Para cada PCR con *primers* cromosoma específico se utilizaron como muestras ADNs (50ng por reacción) proveniente de un sistema de 24 líneas

celulares somáticas híbridas humano/ratón (Contreras-Galindo R, *et al.*, 2013), las cuales contienen el genoma de ratón más el ADN de un único cromosoma humano. Esto nos permite analizar todos los cromosomas humanos por separado. Las condiciones de tiempo, temperatura y número de ciclos se ajustaron en función de los resultados obtenidos de cada reacción de PCR. En varios casos fue necesaria la realización de PCR a tiempo real en gradiente de temperatura para evaluar cuál temperatura permite una óptima amplificación tipo cromosoma-específica. Como control positivo se utilizó ADN de humano sano (50ng por reacción), y se utilizaron tres muestras como controles negativos: el ADN parental de ratón, ADN parental de hámster y agua miliQ en lugar de muestra de ADN. Las reacciones fueron llevadas a cabo en la plataforma StepOne™ Real-Time PCR System, y se utilizó el software StepOne versión 2.3 para el análisis de los resultados.

Los productos amplificados fueron separados electroforéticamente en un gel de agarosa al 2% teñido con bromuro de etidio (EtBr). La corrida electroforética fue realizada a 90V por 45 minutos.

Posteriormente se procedió a la purificación de las bandas obtenidas utilizando el kit de purificación Zymoclean™ Gel DNA Recovery Kit (Zymo Research) siguiendo las instrucciones del proveedor. Los productos purificados fueron clonados en el sistema pCR2.1-TOPO TA Cloning kit (Invitrogen, Inc., Carlsbad, CA, EEUU) siguiendo las instrucciones del fabricante. Los clones de los repetidos alfa fueron crecidos en cepas de *E. Coli* quimiocompetentes mediante shock térmico de 30 segundos a 42°C. Enseguida las bacterias se propagaron por una hora a 37°C con agitación en 1ml de medio SOC (2% Triptona, 0.5% extracto de levadura, 0.4% glucosa, 10 mM Cloruro de Sodio, 2.5 mM Cloruro de Calcio, 5 mM Cloruro de Magnesio, 5 mM Sulfato de Magnesio). El paso siguiente consistió en el plaqueo de las bacterias transformadas en placas de Petri conteniendo medio Luria-Bertani con agar (LB-agar), el antibiótico Kanamicina (50µg/ml) y X-Gal. A continuación se seleccionaron las bacterias crecidas de color blanco (las cuales deberían tener el inserto ligado al vector) y se procedió a la extracción del ADN plasmídico utilizando el kit Fast Plasmid Mini Kit (5 Prime) según las instrucciones del proveedor. Las secuencias de ADN de los repetidos alfa de cada cromosoma humano fueron confirmadas por secuenciación y posteriormente se cuantificó la concentración de cada producto obtenido (en ng/µl) haciendo uso de la tecnología del Nanodrop 2000.

3.2.3. Realización de las curvas estándar para la cuantificación absoluta por PCR a tiempo real.

A partir de los valores de las concentraciones en ng/ μ l de cada clon de cada repetido alfa, se procedió al cálculo del número de copias de ADN siguiendo la siguiente ecuación:

$$\text{Peso en Dalton (g/mol)} = (\text{tamaño en pares de bases del plásmido} + \text{tamaño en pares de bases del inserto}) (330 \text{ Da} \times 2 \text{ nucleótidos/pb})$$

Posteriormente

$$\begin{aligned} (\text{g/mol}) / (\text{Número de Avogadro}) &= (\text{g/molécula}) = \text{Número de copias} \\ \text{Concentración del plásmido} &= (\text{g}/\mu\text{L}) / (\text{g/moléculas}) = \text{moléculas ADN} / \mu\text{L} \end{aligned}$$

Estos datos son necesarios para estimar el número de copias/ μ l de cada clon o muestra de referencia. Diluciones seriadas de esta muestra de referencia, en el rango de 10^8 a 10^3 copias/ μ l sirvieron para construir curvas estándar de calibración y la posterior cuantificación de los repetidos alfa en las muestras. Se realizaron dos réplicas para cada una de las diluciones de los productos de ADN de los repetidos alfa. Se añadieron 2 μ L de cada dilución de estándar de calibración a tubos de PCR que contenían 15 μ L de la mezcla de reacción del kit FastStart Universal SYBR Green Master (Life Science, Roche) compuesta del buffer 2x, el par de *primers* específicos (concentración final 0.5 μ M) para el cromosoma de interés, y agua destilada miliQ. Para la PCR se utilizaron los parámetros puestos a punto previamente para cada juego de *primers* diseñado. En forma general, se trataron de 10 min a 95°C de activación de la enzima, seguidos de 16 a 25 ciclos compuestos de 15 seg de desnaturalización a 95°C, 30 seg entre 60° a 64°C para la hibridación de los *primers* (dependiendo del par de *primers* en cuestión), y 30 seg de elongación a 72°C.

Una vez realizada la PCR a tiempo real, el programa StepOne versión 2.3 reproduce los resultados en un gráfico de curva estándar el que relaciona la concentración logarítmica en base 10 (eje x) con el ciclo umbral (Ct, threshold cycle, por sus siglas en inglés) de cada uno de los estándares (eje y) (ver Fig. 7). Se realizó una curva estándar por cromosoma humano.

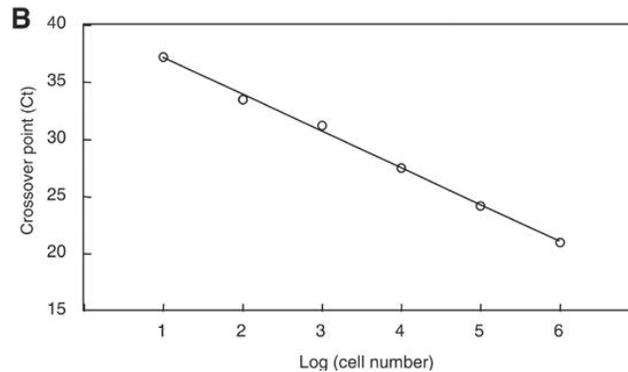


Figura 7. Esquema representativo de una curva de calibración estándar obtenida de la cuantificación absoluta de diluciones seriadas de una muestra de referencia por PCR a tiempo real. La presente figura es, a modo de ejemplo, la curva estándar que se genera luego de la realización de una PCR a tiempo real por el método de diluciones de una muestra de referencia para una posterior cuantificación. Figura tomada de Sakakura C, *et al.*, 2004.

3.2.4. Abundancia de repetidos alfa en la línea celular de LLC (CLL-AT) y en células humanas normales

Se cuantificó, mediante PCR a tiempo real, el arreglo centromérico (presencia y abundancia de repetidos alfa) cromosoma por cromosoma de una línea celular de Leucemia Linfocítica Crónica denominada CLL-AT, así como también de células de humanos sanos cedidas del Laboratorio del Dr. Markovitz, haciendo uso del sistema de curvas estándar mencionado anteriormente. La línea celular que aquí denominamos CLL-AT proviene de una mujer de raza negra con leucemia linfocítica crónica de células B y se encuentra con el código Hs 505.T (ATCC® CRL-7306™) en el sitio web de ATCC (<http://www.atcc.org>).

Cada muestra se analizó por duplicado, y se realizó una reacción con 50ng de ADN de roedor y otra con 50ng de ADN humano sano como controles de reacción negativo y positivo, respectivamente. Los resultados obtenidos se analizaron en el programa computacional StepOne v2.3, en el cual se obtuvo la recta de regresión (curva estándar) a partir de los datos de las diluciones seriadas estándar. Extrapolando los valores de las muestras de interés en esta recta, se cuantificó la abundancia de repetidos alfa de la línea celular CLL-AT y de los ADNs provenientes de tres muestras de humanos sanos.

3.2.5. PCR a tiempo real específica para K111

Un análisis similar fue realizado para evaluar la abundancia de secuencias de retrovirus endógenos humanos integrados en regiones centroméricas de al menos 15 cromosomas humanos y que han sido reportadas y denominadas como K111 por el grupo de investigación del Dr. Markovitz (Contreras-Galindo *et al.*, 2008; 2012 y 2013). Los provirus K111 presentan un *target site duplication* (TSD) de 6pb (GAATTC) flanqueando los extremos 5' y 3' de los LTRs y están insertos entre secuencias de repetidos centroméricos a las cuales se les hace referencia como elementos "(CER)" por ser secuencias conservadas que se encuentran en los centrómeros de todos los cromosomas humanos (ver Fig. 8) (Contreras-Galindo R, *et al.*, 2012). Para esta reacción por PCR a tiempo real, se utilizó una sonda que discrimina específicamente el gen *env* de K111. La PCR cuantitativa fue realizada siguiendo los pasos descritos por Contreras-Galindo y colaboradores en el 2012 utilizando los *primers* K111F (5'-AAG AGC ACC AGG ATG CTT AAT GCC-3') y K111R (5'-AGT GAC ATC CCG CTT ACC ATG TGA-3') y la sonda FAM-labeled K111P (5'-FAM-TGC CCG TCC TAA CAG TAG ACT CAC-BHQ1-3'). La PCR consistió de 35 ciclos compuestos de 15 seg de desnaturalización a 95°C y 1 min de hibridación/elongación a 60°C. Los títulos de K111 se estimaron usando diluciones seriadas de ADN del gen *env* de K111 previamente clonado en el sistema pCR2.1-TOPO TA Cloning kit, (Invitrogen, Inc., Carlsbad, CA, EEUU) al igual que para el caso de los repetidos alfa. Una vez realizadas las diluciones se cuantificó el número de K111 presente en los centrómeros de la línea celular de LLC (CLL-AT) y en las tres muestras provenientes de humanos sanos utilizados en la cuantificación de los repetidos alfa centroméricos.

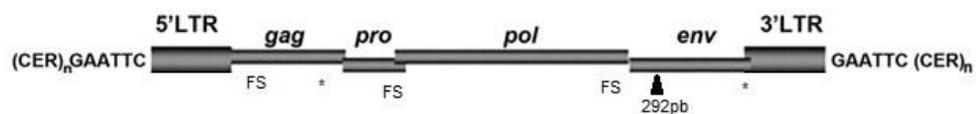


Figura 8. Organización genómica de un provirus de HERV-K centromérico (K111). En la figura se esquematizan los genes de un provirus de HERV-K centromérico (*gag*, *pro*, *pol* y *env*) con los 5' y 3' LTRs flanqueando sus extremos. Los codones stop se representan con un asterisco y los marcos abiertos de lectura (ORFs) se especifican como "FS" (por su sigla en inglés frameshift). Se incluye la deleción de 292pb común en los provirus de tipo 1 (línea vertical con triángulo). A los extremos de los LTRs se ubican los target site duplication (TSD) de 6pb para K111, seguidos de secuencias de repetidos centroméricos (CER) que quedan interrumpidas por la inserción de los HERV-K (K111) en esos sitios del genoma humano. Figura tomada de Contreras-Galindo R, *et al.*, 2012.

3.2.6. Análisis estadístico

Para visualizar los resultados obtenidos luego de la cuantificación absoluta de los repetidos alfa por PCR a tiempo real, se graficó para cada cromosoma, el valor medio de las dos réplicas técnicas realizadas para las muestras de ADN de la línea celular CLL-AT y el promedio de los valores obtenidos de las muestras de donantes sanos con su desvío estándar correspondiente. Asimismo se graficó el resultado de la cuantificación absoluta por PCR a tiempo real de los provirus centroméricos K111 presentes en el genoma de la línea celular de LLC y en el promedio de los valores obtenidos del genoma de los donantes sanos. Se realizó el análisis estadístico test-T no-pareado de dos colas con un intervalo de confianza del 99% con el fin de evaluar la existencia de diferencias significativas de los valores de repetidos alfa obtenidos cromosoma por cromosoma para la línea de LLC en comparación con los humanos sanos. Se repitió el mismo análisis para el número de K111 obtenido en los dos modelos de estudio CLL-AT y donantes sanos. Se utilizó el software GraphPad versión 5.00 para la realización del gráfico así como para el análisis estadístico correspondiente.

4. RESULTADOS

4.1. Resultados del Objetivo General 1

Estudiar la posible relación entre la expresión de los genes de HERV-K en la patología hemato-oncológica Leucemia Linfoide Crónica (LLC).

4.1.1. Objetivo específico 1: Puesta a punto de la metodología de RT-PCR para la obtención de todos los genes de HERV-K (*env*, *gag*, *pol*, *rec* y *np9*).

Luego de evaluar diferentes condiciones de tiempos y temperaturas de ciclado por PCR convencional, se logró la puesta a punto de las condiciones que permiten una óptima obtención de los genes de HERV-K en forma independiente. En la figura 9 se observan los productos de amplificación para cada gen de HERV-K (*gag*, *pol*, *env*, *np9* y *rec*), luego de la corrida por electroforesis en gel de agarosa. Los resultados de esta figura corresponden a la amplificación de cada gen de HERV-K con las condiciones finales optimizadas (ver Materiales y Métodos), donde se utilizó como molde ADNc de PBMC de un paciente de LLC con perfil IgVH-M, un paciente con perfil IgVH-NM, PBMC de un donante sano, y un control negativo. Se obtuvo producto de amplificación en todas las mezclas de reacción que contenían ADNc como molde y conforme a lo esperado, no se obtuvieron productos de amplificación en las mezclas de PCR donde se sustituyó el ADNc por agua para la reacción del control negativo (Ver figura 9).

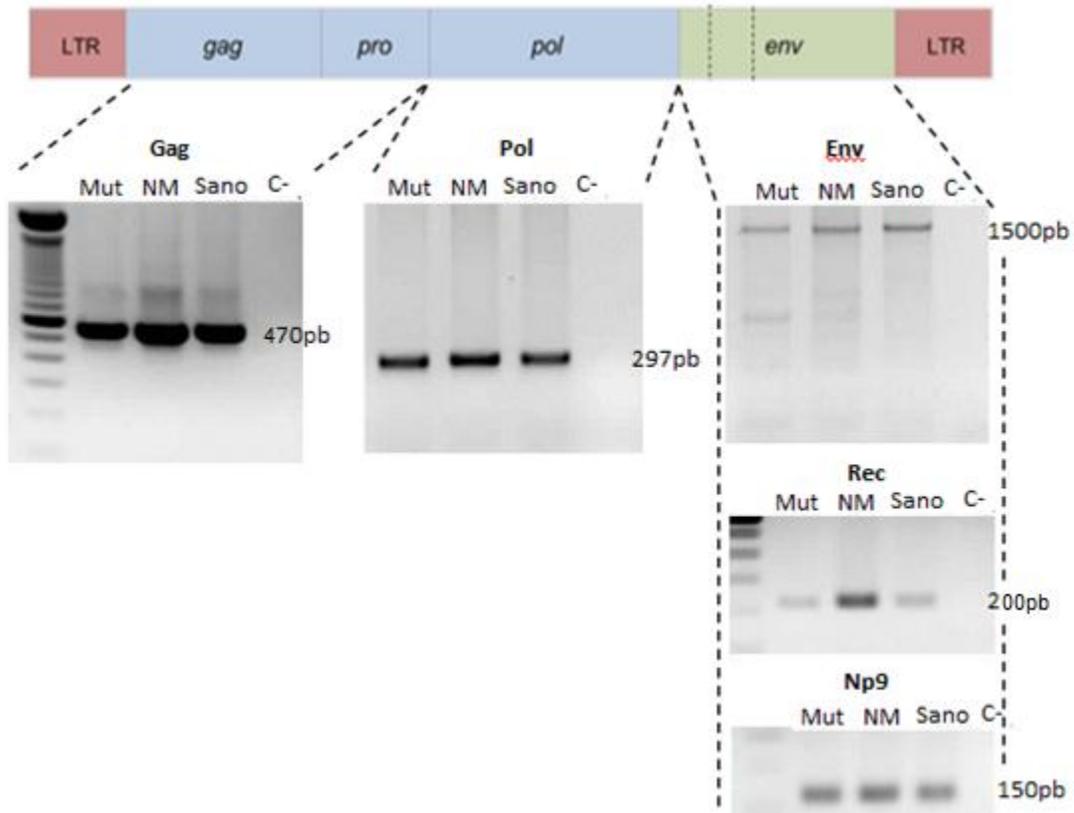


Figura 9. PCR cualitativa para los genes *gag*, *pol*, *env*, *rec* y *np9* de HERV-K. En la figura se observa el esquema de una secuencia proviral completa de HERV-K y los resultados de las corridas electroforéticas en geles de agarosa de los productos de amplificación por PCR para cada uno de sus genes en forma independiente. Los genes *rec* y *np9* se producen por *splicing* alternativo del gen *env*, y por este motivo se ubican en la misma columna. Mut: Paciente con LLC perfil IgVH-M, NM: LLC perfil IgVH-NM, Sano: PBMC de un donante sano, y C-: Control Negativo de la reacción. Se especifica el tamaño en pares de bases (pb) para cada una de las regiones amplificadas.

4.1.2. Objetivo específico 2: Cuantificar los niveles de expresión relativa de los genes *np9* y *gag* de HERV-K en muestras de pacientes con LLC (perfil IgVH-M y IgVH-NM) con respecto a muestras de donantes sanos.

Diferentes estudios que, mediante el uso de tecnología de secuenciación masiva, han escaneado la presencia de una expresión viral a nivel de ARN en LLC humana, no han dado evidencias de un posible candidato viral exógeno como agente etiológico de esta enfermedad (Rego N, *et al.*, 2012). Con motivo de profundizar en el conocimiento de las bases que rigen este desorden hemato-oncogénico, y considerando el incremento de estudios que asocian un rol de retrovirus endógenos con leucemia, en esta parte del

proyecto se estudió el nivel de expresión de los genes *np9* y *gag* de HERV-K en muestras de pacientes con LLC y muestras de donantes sanos mediante PCR a tiempo real.

En este estudio se evaluó el nivel de expresión de los genes *np9* y *gag* en 25 muestras de pacientes con LLC (14 con perfil IgVH mutado y 11 con perfil IgVH no mutado) y se comparó con el nivel de expresión de 6 muestras de PBMCs provenientes de donantes sanos. En primer lugar se evaluó el nivel de expresión para el gen *np9*, el gen putativo oncogénico que deriva del *splicing* alternativo del gen *env* en los HERV-K tipo I (ver Fig. 10A), y posteriormente se evaluó una región del gen *gag* que codifica para las proteínas estructurales de la cápside (CA), matriz (MA) y nucleocápside (NC) (ver Fig.10B).

El presente estudio dió como resultado que la actividad transcripcional del gen *np9* de HERV-K fue mayor en las muestras de los pacientes de LLC que en las muestras de PBMCs de los donantes sanos. Los resultados revelan que 18 de los 25 pacientes con LLC (el 70% de los pacientes con LLC) expresan el transcrito *np9* al menos cinco veces más que el promedio de los PBMCs de los donantes sanos evaluados (Fig. 10A).

Por otro lado, el análisis de los resultados del nivel de expresión del gen *gag* reveló que 8 de los 25 pacientes con LLC (el 32% de los pacientes con LLC) expresan al menos dos veces más el gen *gag* en comparación con los donantes sanos (Fig. 10B). Solamente dos muestras de LLC de perfil IgVH-M presentaron un nivel de expresión del gen *gag* al menos cinco veces mayor respecto al promedio del nivel de expresión de los donantes sanos (Fig. 10B).

Tanto para el análisis de *np9* como de *gag*, se puede observar una importante variabilidad en los niveles de expresión entre los pacientes con LLC. No se obtuvo diferencias estadísticamente significativas en los niveles de expresión relativa de ambos genes entre los subgrupos de pacientes con perfil IgVH-M y perfil IgVH-NM ($p=0.65$ para el gen *np9* y $p=0.09$ para el gen *gag*).

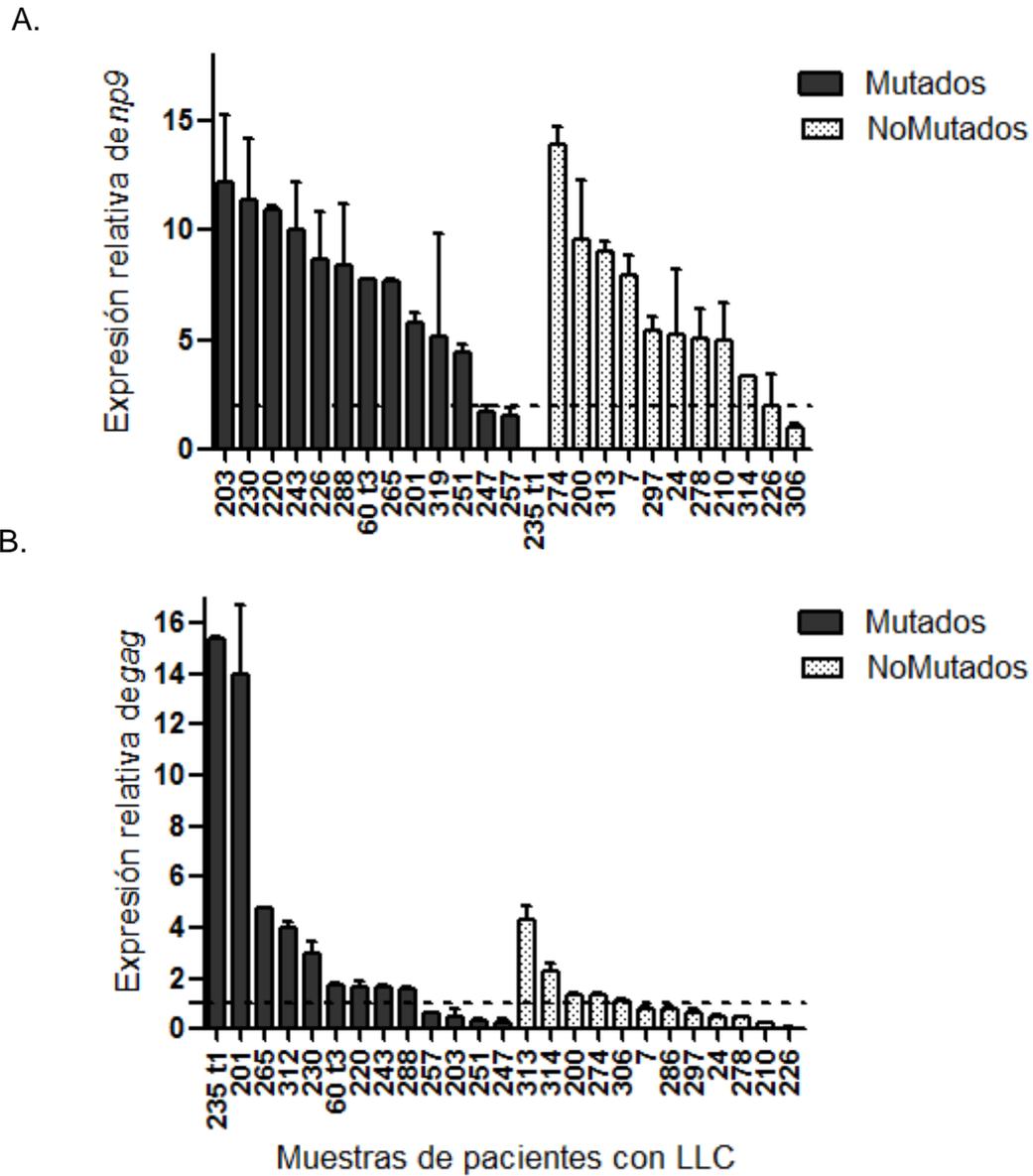


Figura 10. Niveles de expresión relativa de los transcritos *np9* y *gag* de HERV-K en pacientes con LLC con perfil de genes VH de las inmunoglobulinas mutados y no mutados. (A) Expresión relativa del gen *np9* de HERV-K de pacientes con LLC (Mutados y No Mutados) con respecto al promedio del nivel de expresión relativa de *np9* de los donantes sanos. (B) Expresión relativa del gen *gag* de HERV-K de pacientes con LLC (Mutados y No Mutados) con respecto al promedio del nivel de expresión relativa de *gag* de los donantes sanos. Se utilizó el nivel de expresión del gen *gapdh* para normalizar los valores de expresión de ambos genes de HERV-K. Se grafican los resultados del cálculo de $2^{-\Delta\Delta Ct}$ obtenidos por PCR a tiempo real. La línea punteada se corresponde al valor máximo de expresión de *np9* (A) y *gag* (B) obtenido de los donantes sanos a modo de

calibrador negativo. Las barras de error verticales representan el máximo valor para cada muestra de paciente con LLC medido por duplicado.

4.1.3. Objetivo específico 3: Cuantificar la expresión del gen celular *c-myc* en las muestras de LLC y en muestras de donantes sanos y evaluar una posible relación significativa con los niveles de expresión del gen *np9* de HERV-K.

Se ha reportado que la proteína Np9 codificada por los HERV-K de tipo I presenta un dominio de unión a la proteína en dedos de zinc de la leucemia promielocítica (PLZF por su sigla en inglés). Es conocido que esta interacción suprime la función de PLZF como represor transcripcional del proto-oncogen *c-myc*, resultando en la sobreproducción de este factor de transcripción y consecuentemente en la inducción de la proliferación celular (Denne M, *et al.*, 2007). Considerando la consecuencia de esta interacción, nos resultó interesante evaluar si existe una posible relación entre el nivel de expresión del gen *np9* de HERV-K con el nivel de expresión del gen celular *c-myc* en las muestras de pacientes con LLC y en donantes sanos. Las muestras analizadas se agruparon en tres subgrupos: pacientes de LLC con un elevado nivel de expresión del gen *np9* (mayor a cinco veces respecto a los donantes sanos), pacientes de LLC con un nivel de expresión de *np9* medio (menor a cinco veces respecto a los donantes sanos), y el subgrupo de los donantes sanos. En la figura 11 se puede observar el gráfico donde se ven representados los valores obtenidos por PCR a tiempo real del nivel de expresión relativa del gen *c-myc* (ΔCt) para cada uno de los subgrupos mencionados.

Las diferencias en los niveles de expresión del gen celular *c-myc* para los pacientes con LLC con alto y bajo nivel de expresión del gen *np9* de HERV-K no resultaron estadísticamente significativas ($p=0.19$). Tampoco lo fueron las diferencias entre estos grupos con respecto al nivel de *c-myc* que expresan los donantes sanos (ver Fig.11).

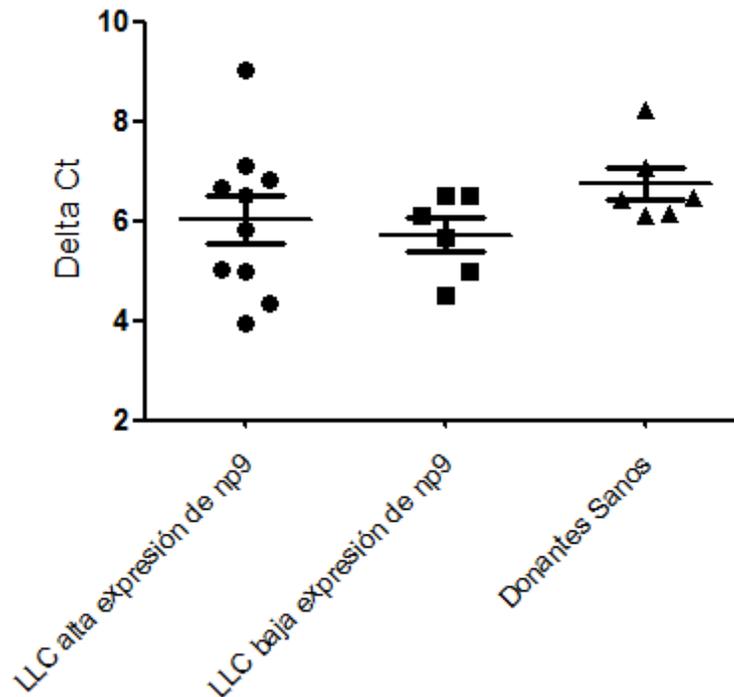


Figura 11. Expresión relativa del gen celular *c-myc* en PBMCs de pacientes de LLC con alto y bajo nivel de expresión del gen *np9* de HERV-K y en humanos sanos. Los niveles de expresión de *c-myc* se normalizaron con el nivel de expresión del gen constitutivo *gapdh* siguiendo el método Δ Ct. Las barras horizontales representan la mediana y las barras de error el rango intercuartil. No hay diferencias significativas ($p=0.19$) en la expresión del gen *c-myc* entre los subgrupos analizados.

4.1.4. Objetivo específico 4: Estudiar la variabilidad genética del gen *np9* en pacientes con LLC y en donantes sanos mediante aproximaciones bioinformáticas.

A fin de determinar la variabilidad genética del gen *np9* completo se realizó la amplificación y posterior secuenciación para 4 muestras de PBMC de donantes sanos, 4 muestras de PBMCs de pacientes con LLC perfil IgVH-M y 6 muestras de PBMCs de pacientes con LLC perfil IgVH-NM. Luego de obtener las secuencias se realizó el análisis bioinformático de las mismas con respecto a tres secuencias de *np9* de referencia encontradas en la base de datos del Gene Bank.

El principal resultado que deriva del análisis bioinformático de las secuencias de *np9* obtenidas es que se trata de una secuencia de 225 pb altamente conservada. Su traducción a la forma aminoacídica de Np9 dio un tamaño de 74aa lo cual se corresponde a una masa molecular de 8.7kDa tal cual lo reportado para otras secuencias de Np9.

No se encontraron diferencias nucleotídicas que sean características de alguno de los subgrupos analizados (muestras de pacientes con LLC perfil de IgVH-M, perfil IgVH-NM y donantes sanos). Si se detectó la presencia de dos posibles polimorfismos y tres mutaciones que se describen a continuación:

El análisis detallado del cromatograma para cada secuencia obtenida permitió detectar que para dos sitios nucleotídicos (la posición 38 y 199 del gen *np9*) todas las secuencias presentaron dos picos superpuestos cuya traducción da lugar a cambios aminoacídicos:

-En la posición nucleotídica 38 se observan dos picos superpuestos, uno correspondiente a una Guanina (G) y otro a una Adenina (A). La traducción del codón donde se ubica este sitio nucleotídico da lugar a una Arginina-R- (aminoácido polar, hidrofílico, con carga positiva) o una Glutamina-Q-(aminoácido polar, hidrofílico con carga negativa) respectivamente.

-En la posición nucleotídica 199 se observa tanto una Guanina (G) como una Adenina (A), y la traducción del codón en que esta base se encuentra da lugar a una Glicina-G- (aminoácido alifático, neutro) o una Arginina-R- (aminoácido polar, hidrofílico, con carga positiva) en la posición 67 de la cadena aminoacídica.

La observación de estos dos picos superpuestos en el cromatograma para todas las muestras amplificadas y secuenciadas podría estar significando la presencia de posibles polimorfismos fijados en la población. En la figura 12 se observa cómo se visualizan los picos en el cromatograma para cada sitio nucleotídico mencionado, y en la tabla 3 se resumen los cambios aminoacídicos que derivan de su traducción.

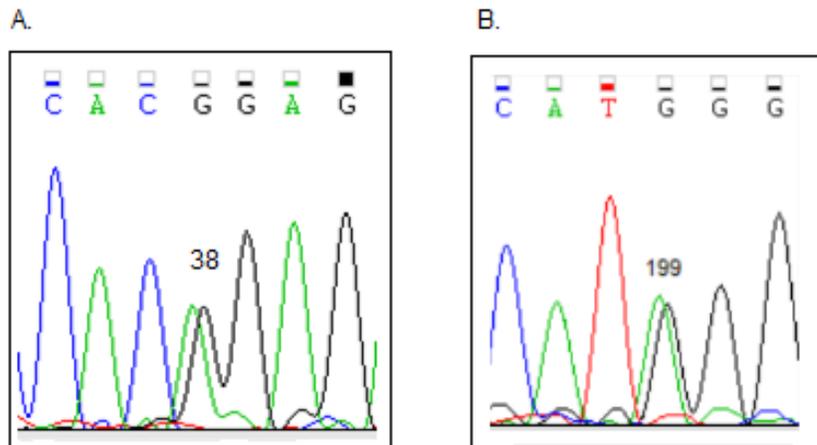


Figura 12. Cromatograma de las posiciones nucleotídicas 38 y 199 de la secuencia del gen *np9* de HERV-K. En la figura se observan las regiones nucleotídicas donde se detectó la presencia de posibles polimorfismos en el análisis de los cromatogramas. Se visualizan los picos superpuestos en la posición nucleotídica 38 (A) y 199 (B) de una de las secuencias obtenidas posterior a la amplificación del gen *np9* completo de HERV-K.

Tabla 3. Posibles polimorfismos obtenidos en la secuencia nucleotídica del gen *np9* y el cambio aminoacídico que deriva de su traducción.

Secuencia nucleotídica np9 (225pb)	Secuencia aminoacídica Np9 (75aa)
G38A	R13Q
G199A	G67R

Se detectó además una mutación de una Adenina (A) por una Citosina (C) en la posición 217 de todas nuestras secuencias en comparación con las secuencias obtenidas en el banco de datos. La traducción de este sitio genera un cambio aminoacídico de una Metionina-M (aminoácido neutro, no polar, hidrofóbico), por una Leucina-L (también neutro, no polar, hidrofóbico).

Otro aspecto que se deriva del análisis bio-informático de la secuencia que codifica para la proteína Np9, es que las dos últimas bases previas al codón stop (TAA) en algunas muestras son Guanina + Timina (GT), y otras muestras presentan una Adenina + Citocina (AC). El codón con GT da lugar al aminoácido Cisteína-C-(aminoácido polar, hidrofóbico, neutro), y el codón con AC se traduce en una Tirosina-Y-(aromático, hidrofóbico, neutro). Asimismo, la

secuencia de uno de los donantes sanos presentó una mutación en la posición 191 (G191C), cuya traducción pasa de una Serina-S- a una Treonina-T-, ambos aminoácidos polares, hidrofílicos y neutros.

En la figura 13 se observa el alineamiento de las secuencias obtenidas para todas las muestras analizadas y se enmarcan los cambios (polimorfismos y mutaciones) encontrados (ver Fig. 13).

Ref_EP2402447	AGA	TCA	GAT	TGT	TAC	TGT	GTC	TGT	GTA	GAA	AGA	AGT	AGA	CAT	GGG	AGA	CTC	CAT	TTT	GTT	ATG	TGT	TAA
Ref_WO2006103562
Ref_Wang 2007T.
Sano_1	A..	C..
Sano_3	A..	C..
Sano_4	A..	C..
Sano_5	A..	.A.	C..	.AC	...
LLC_Mut_1	A..	C..
LLC_Mut_2	A..	C..	.AC	...
LLC_Mut_3	A..	C..	.AC	...
LLC_Mut_4	A..	C..	.AC	...
LLC_NoMut_1	A..	C..
LLC_NoMut_2	A..	C..	.AC	...
LLC_NoMut_3	A..	C..	.AC	...
LLC_NoMut_4	A..	C..	.AC	...
LLC_NoMut_5	A..	C..	.AC	...
LLC_NoMut_6	A..	C..

Figura 13. Alineamiento de las secuencias nucleotídicas del gen *np9* de HERV-K. En la figura se muestra el alineamiento obtenido a partir de la amplificación por PCR cualitativa y posterior secuenciación bidireccional de 14 muestras de PBMCs (4 de donantes sanos, 4 de pacientes con LLC perfil IgVH-Mutado, y 6 de pacientes con perfil IgVH-NoMutado). Nuestras secuencias fueron alineadas con tres secuencias de referencia (Ref) obtenidas de la base de datos del Gene Bank. Se remarcan en color verde los codones donde se detectaron posibles polimorfismos (posiciones 38 y 199) y en rojo los codones donde se detectaron mutaciones puntuales (posiciones 217 y 221-222).

4.2. Resultados del Objetivo General 2

Destacando el rol de los centrómeros como sitios claves que aseguran una correcta función en la mitosis y meiosis, es importante considerar que un mal funcionamiento en estos sitios puede dar lugar a una incorrecta segregación de los cromosomas al momento de la división celular. Incluso se ha visto que existen zonas puntuales en el genoma humano, los denominados “puntos de ruptura evolutivos” (EB por su sigla en inglés) muy susceptibles a un quiebre de las dos hebras de ADN. El quiebre en estas zonas puntuales del genoma origina los rearrreglos cromosómicos muy frecuentemente asociados a enfermedades (deleciones, translocaciones inversiones, etc). Lo interesante de estos EB es que el estudio de su secuencia evidenció que se trata de una distribución conservada de elementos repetitivos (repetidos alfa) ricos en secuencias LINES, y de retrovirus endógenos (ERVs), es decir muchos de los EB se corresponden con secuencias típicas centroméricas (Murphy WJ, *et al.*, 2005). La leucemia linfocítica crónica presenta ciertas aberraciones cromosómicas diferenciales, y por tal motivo, nos resulta interesante intentar responder a la pregunta de si existe algún patrón que caracterice la estructura del centrómero de células de LLC una vez comparadas con muestras de humanos sanos.

A continuación se presentan los resultados obtenidos de nuestro segundo objetivo general propuesto: Determinar si existe algún patrón diferencial en la abundancia de repetidos alfa y secuencias de HERVs (K111) en los centrómeros de una línea celular de LLC (CLL-AT) y su comparación con donantes sanos.

4.2.1. Objetivo específico 1: Puesta a punto de la metodología de PCR a tiempo real para obtener las secuencias centroméricas de repetidos alfa de manera cromosoma-específica.

Se realizó el diseño de los *primers* específicos que permitieran la amplificación de los repetidos alfa correspondientes a cada uno de los cromosomas de forma independiente mediante PCR a tiempo real. Posteriormente se realizó la puesta a punto de las condiciones de tiempos y temperaturas necesarias que permitieron la obtención de estas secuencias en forma independiente. Es importante destacar que cada reacción de PCR se realizó evaluando todos los cromosomas humanos de manera independiente utilizando ADN de un sistema híbrido humano/ratón el cual contiene el genoma de ratón y un solo cromosoma humano por separado.

En la figura 14 se observan los resultados de las PCR a tiempo real para los repetidos alfa de cada uno de los 24 cromosomas humanos.

Como se puede observar en la figura, para los cromosomas 1, 2, 3, 6, 7, 8, 11, 12, 16, 20, 22, X e Y se observan únicamente dos curvas de amplificación aproximadamente en el ciclo 16 del ciclado de la PCR, una correspondiente a la amplificación del cromosoma de interés y la otra correspondiente al control positivo que contiene al ADN humano normal. Este resultado es coherente a lo esperado ya que, al colocar ADN híbrido humano/ratón de los 24 cromosomas humanos por separado para cada una de estas reacciones, se esperaría obtener producto de amplificación únicamente en las mezclas donde se colocó el ADN híbrido que contiene el cromosoma humano esperado, y del control positivo que contiene ADN de humano normal (ver Fig. 14).

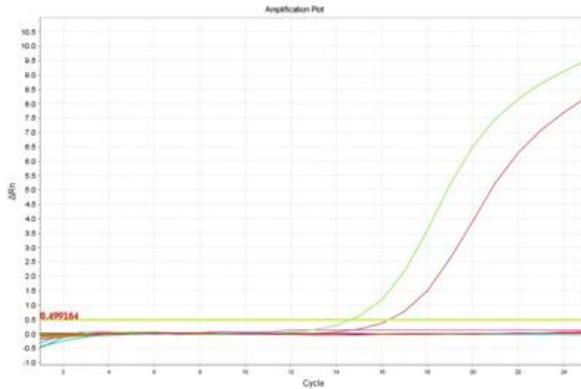
Para el caso de los cromosomas 4, 5, 9, 10, 13, 15, 17 y 18 se obtuvo un tercer producto de amplificación ciclos más tarde al de los productos esperados. La lectura de este tercer producto que podríamos llamar “inespecífico” puede eliminarse elevando el nivel del valor umbral en los gráficos obtenidos. De esta forma se continúa teniendo una amplificación centromérica específica para el cromosoma de interés (ver Fig. 14).

En el caso de los repetidos alfa de los cromosomas 14, 19 y 21, se obtuvieron tres curvas de amplificación juntas imposibles de separar modificando el nivel del valor umbral. Naturalmente es de esperar que el repetido alfa del cromosoma 14 es casi idéntico al del cromosoma 22 y por eso se amplifican juntos al utilizar el par de *primers* D14Z1/D22Z1; el centrómero del cromosoma 19 es casi idéntico al cromosoma 5 y se amplifican juntos frente al par de *primers* D19Z5; y el repetido alfa del cromosoma 21 es casi idéntico al del cromosoma 13 y se obtienen juntos cuando se utiliza el par de *primers* D21Z1 (ver Fig. 14).

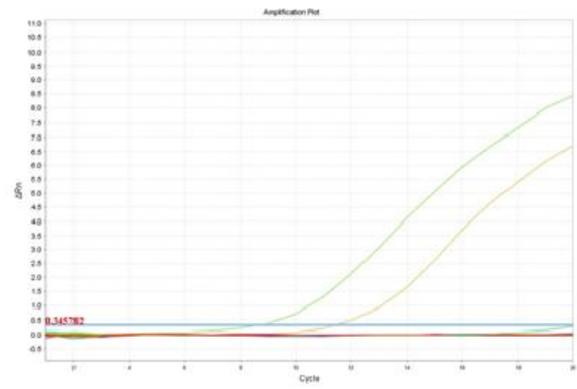
El juego de *primers* P82H permite la amplificación de los repetidos alfa centroméricos de todos los cromosomas humanos juntos (ver Fig. 14)

Posterior a la puesta a punto de las condiciones de amplificación de los repetidos alfa cromosoma-específicos por PCR a tiempo real, y a fin de realizar las curvas estándar para la cuantificación absoluta, los productos obtenidos fueron separados electroforéticamente en un gel de agarosa teñido previamente con EtBr. En la figura 15 se observan los resultados obtenidos luego de la electroforesis donde se tomaron fotos del gel con luz UV y se cortaron las zonas donde se ubicaron las bandas de amplificación de aproximadamente 120 pb. En concordancia con los gráficos recientemente descritos (ver Fig. 14), determinados juegos

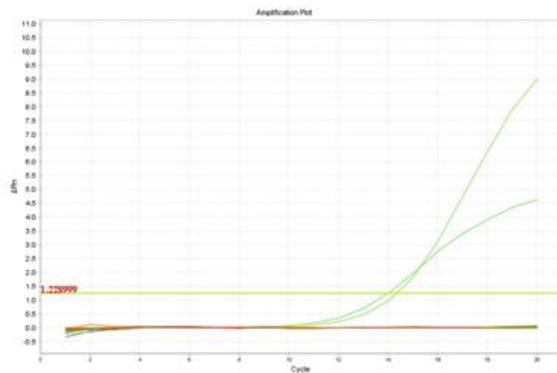
de *primers* (los de los cromosomas 14, 19 y 21) amplifican secuencias de repetidos alfa en otros cromosomas humanos además del esperado, lo que se puede visualizar en el gel con la presencia de más de una banda. En todas las reacciones se obtuvo banda en el control positivo, y no se observa producto de amplificación para los controles negativos realizados (ver Fig. 15).



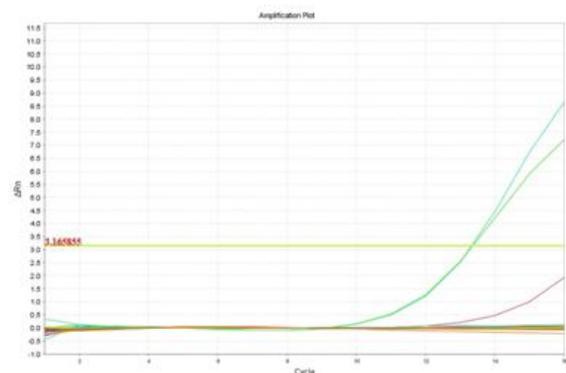
Cromosoma 1 (D1Z5c)



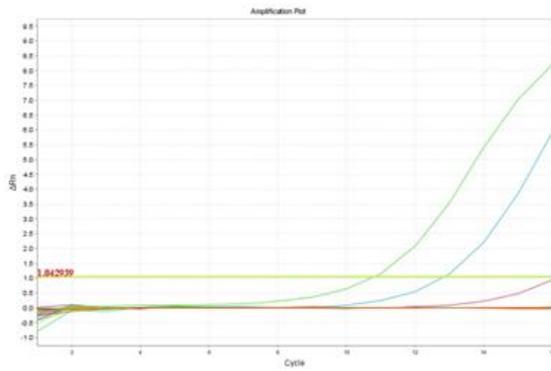
Cromosoma 2 (D2Z1)



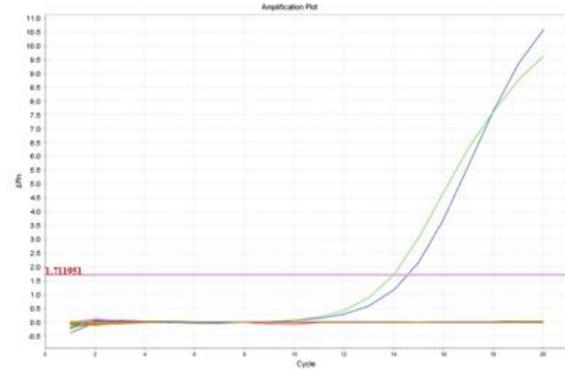
Cromosoma 3 (D3Z1)



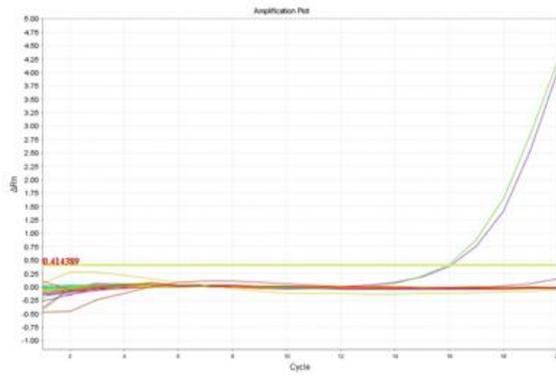
Cromosoma 4 (D4Z1)



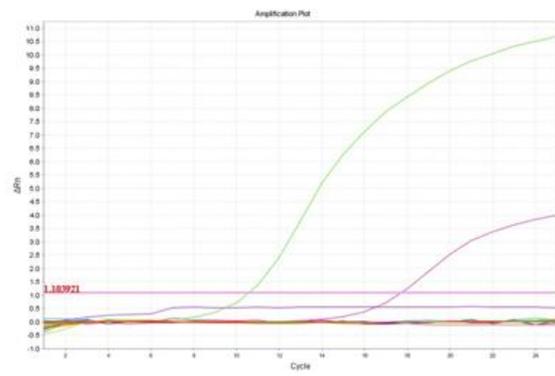
Cromosoma 5 (D5Z1)



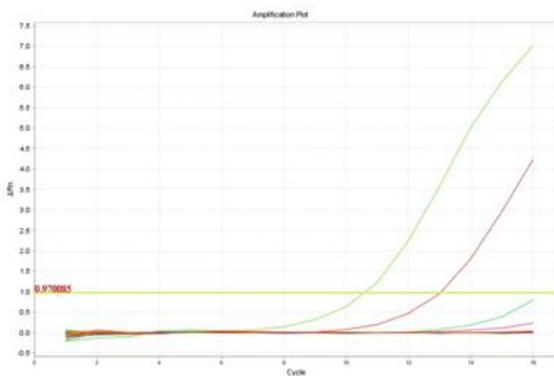
Cromosoma 6 (D6Z1)



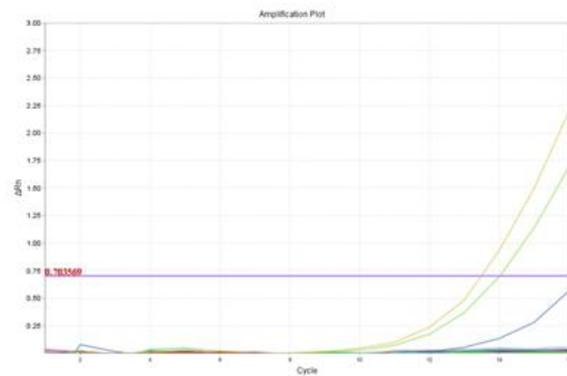
Cromosoma 7 (D7Z2)



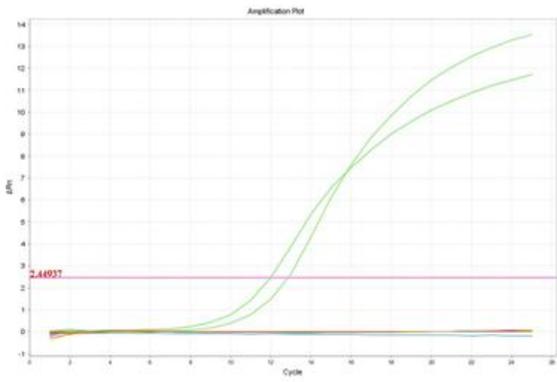
Cromosoma 8 (D8Z2)



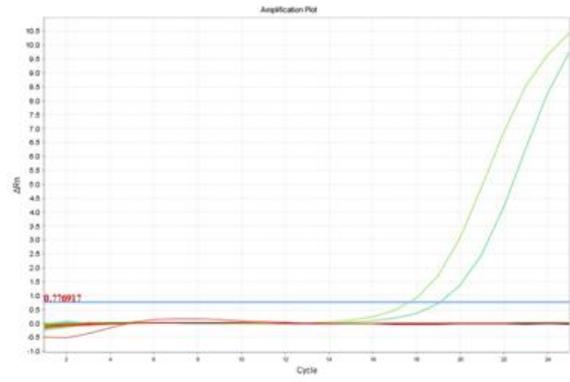
Cromosoma 9 (D9Z4)



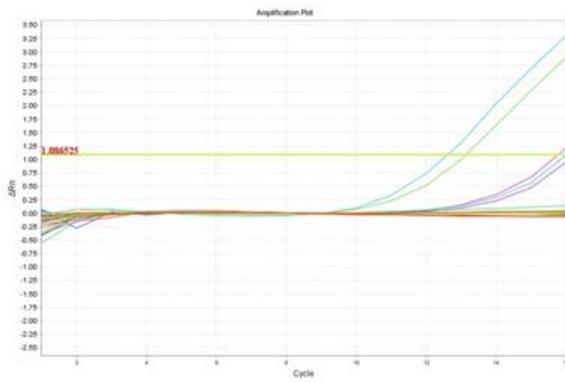
Cromosoma 10 (D10Z1)



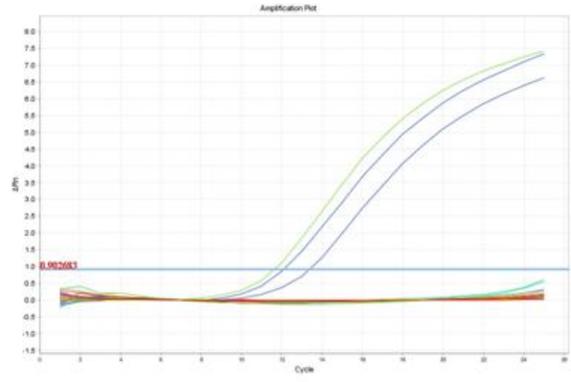
Cromosoma 11 (D11Z1)



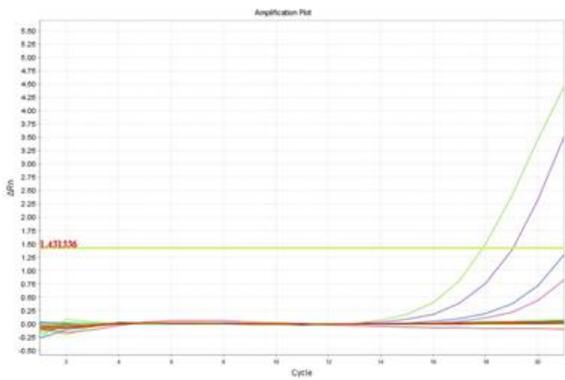
Cromosoma 12 (D12Z3)



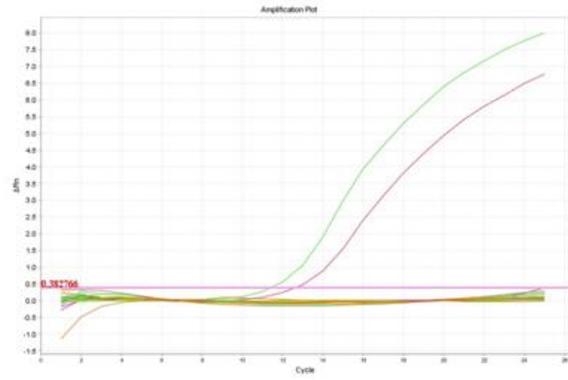
Cromosoma 13 (D13Z1)



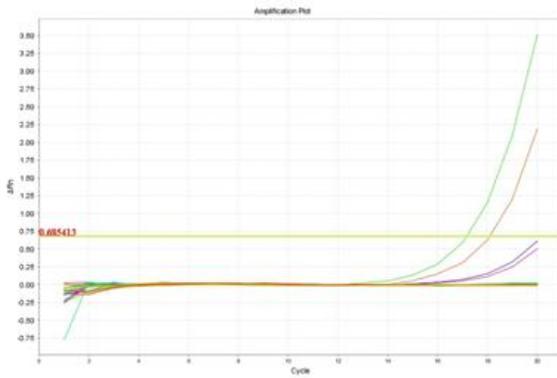
Cromosoma 14 (D14Z1/D22Z1)



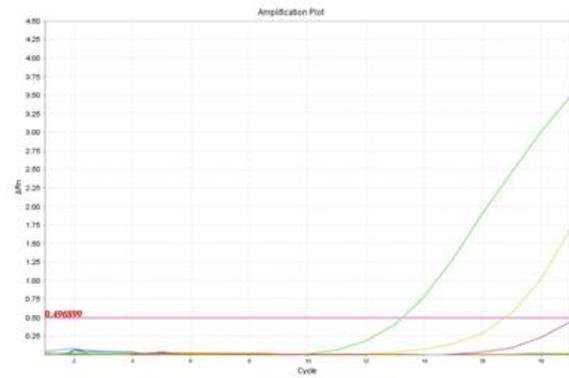
Cromosoma 15 (D15Z3)



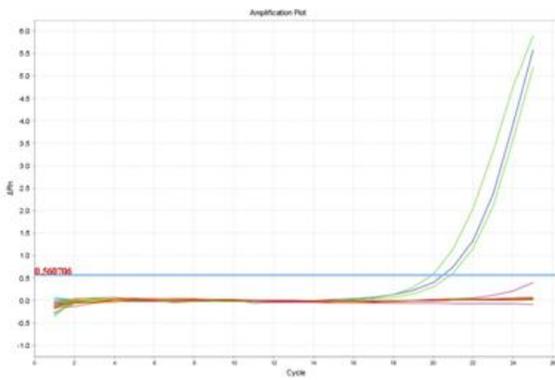
Cromosoma 16 (D16Z2)



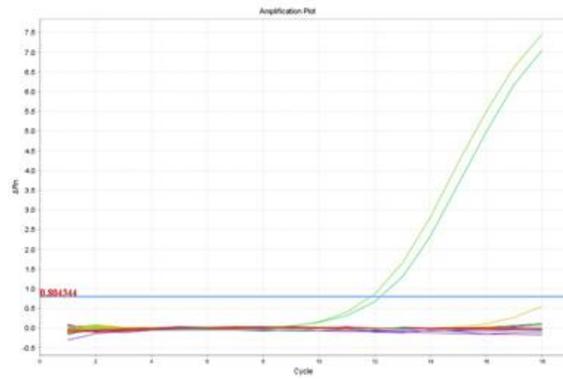
Cromosoma 17 (D17Z1)



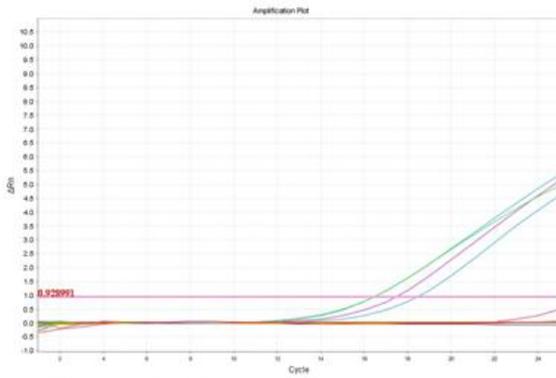
Cromosoma 18 (D18Z2)



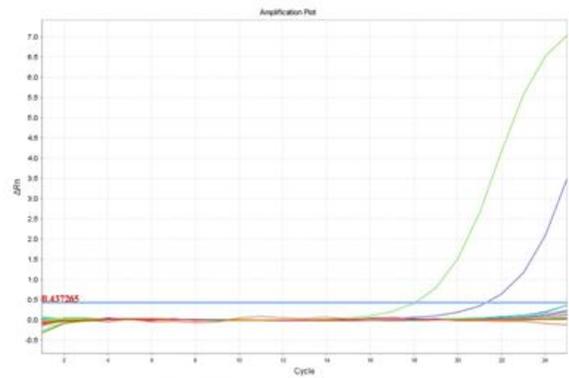
Cromosoma 19 (D19Z5)



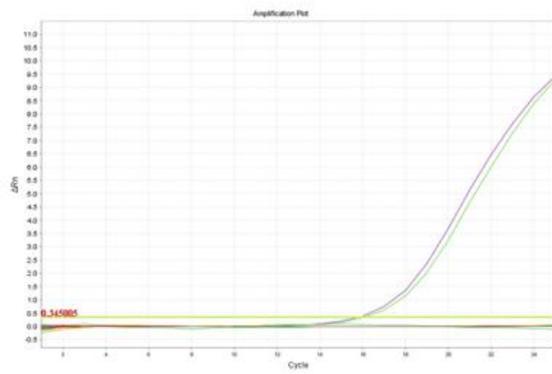
Cromosoma 20 (D20Z2)



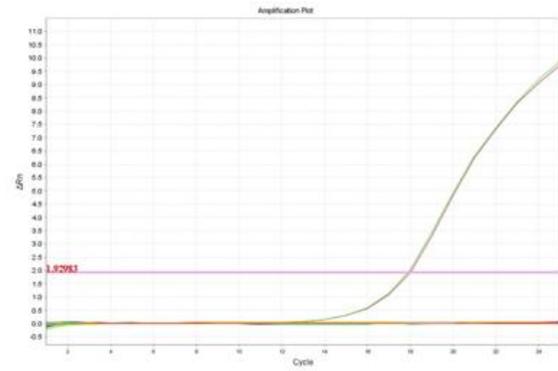
Cromosoma 21 (D21Z1)



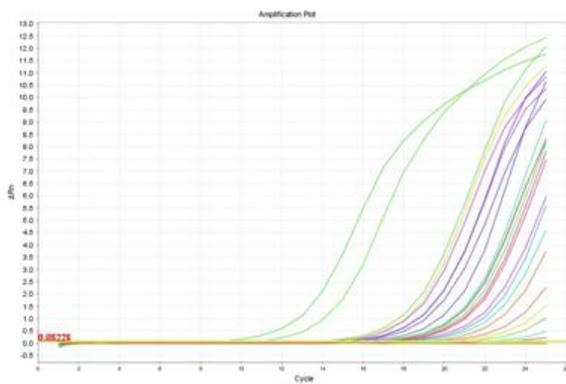
Cromosoma 22 (D22Z4)



Cromosoma X (DXZ1)



Cromosoma Y (DYZ1)



Todos los cromosomas (P82H)

Figura 14. Curvas de amplificación de repetidos alfa para cada cromosoma humano obtenidas por PCR a tiempo real. Se grafican los resultados del ΔRn (señal de fluorescencia reportada – línea de base) en función del número de ciclos de la PCR para cada juego de *primers* utilizado, o lo que es lo mismo, para cada centrómero cromosoma humano específico (especificado debajo de cada gráfico). Los colores de las curvas de amplificación se corresponden a cada muestra de ADN (híbrido humano-ratón y controles) colocada en cada mix de PCR.

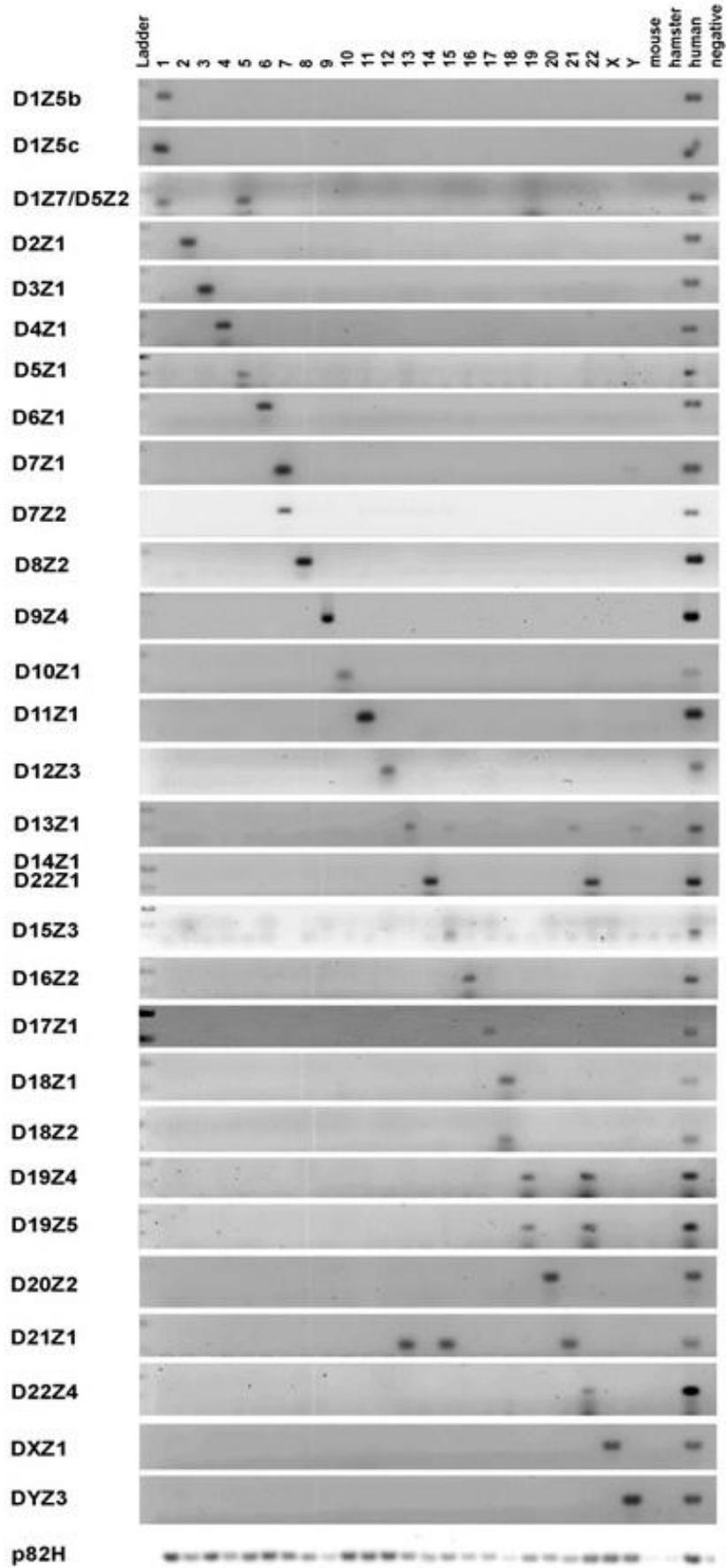


Figura 15. Visualización de los productos de amplificación de los repetidos alfa para cada juego de primers utilizado. En la figura se observan los productos obtenidos por PCR a tiempo real luego de su corrida electroforética en un gel de agarosa al 2% teñido previamente con EtBr. A la izquierda del gel se especifica el juego de *primers* empleado por reacción para los cromosomas los cuales fueron diseñados. Ladder: marcador de peso molecular 100pb (Thermo Scientific GeneRuler 100bp Plus DNA Ladder). Del 1 al Y: producto de reacción donde se colocó el ADN de la línea celular híbrida humano/ratón en orden, desde el cromosoma humano 1 al cromosoma 22, luego el X y por último la línea que contiene el cromosoma humano Y. Mouse: control negativo con ADN de ratón. Hámster: control negativo con ADN de hámster. Human: Control positivo con ADN humano. Negative: control negativo con agua en lugar de ADN. Las bandas se corresponden a un tamaño de aproximadamente 120 pb.

En la tabla 4 se describe un resumen de los resultados obtenidos luego de la puesta a punto de las condiciones para las PCRs a tiempo real que permiten la amplificación de los repetidos alfa centroméricos en forma cromosoma-específica. Se especifica el nombre de los *primers* utilizados, los cromosomas para los cuales el set de *primers* fue diseñado con el fin de amplificar su repetido alfa en forma específica, y las condiciones del número de ciclos y temperatura de hibridación de los *primers* puestas a punto en el presente trabajo (ver Tabla 4).

Tabla 4. Nombre de los primers y condiciones de amplificación para cada repetido alfa centromérico.

Repetido Alfa	Cromosoma	Protocolo de PCR a tempo real exitoso	Cromosoma detectado	Condiciones PCR (ciclos, hibridación primers)
D1Z5c	1	si	1	25, 60°C
D2Z1	2	si	2	20, 60°C
D3Z1	3	si	3	20, 60°C
D4Z1	4	si	4	16, 62°C
D5Z1	5	si	5	16, 60°C
D6Z1	6	si	6	20, 60°C
D7Z2	7	si	7	20, 61°C
D8Z2	8	si	8	25, 60°C
D9Z4	9	si	9	16, 62°C
D10Z1	10	si	10	18, 62°C
D11Z1	11	si	11	25, 60°C
D12Z3	12	si	12	16, 56°C
D13Z1	13	no	13, 21	20, 61°C
D14Z1/D22Z1	14, 22	si	14, 22	20, 60°C
D15Z3	15	si	15	22, 64°C
D16Z2	16	si	16	25, 60°C
D17Z1	17	si	17	20, 61°C
D18Z2	18	si	18	20, 62°C
D19Z5	19	no	19, 22	25, 60°C
D20Z2	20	si	20	25, 60°C
D21Z1	21	no	21, 13	20, 62°C
D22Z4	22	si	22	25, 60°C
DXZ1	X	si	X	25, 60°C
DYZ3	Y	si	Y	25, 60°C
P82H	todos	si	todos	25, 60°C

4.2.2. Objetivo específico 2: Cuantificar la abundancia de repetidos alfa y de retrovirus endógenos humanos (K111) en los centrómeros de la línea celular de leucemia linfocítica crónica (CLL-AT) y células de pacientes sanos. Evaluar si existen diferencias significativas en la estructura centromérica entre ambos grupos.

Se realizó la cuantificación absoluta del número de repetidos alfa por PCR a tiempo real empleando el método de curva de calibración estándar. Estos estudios evidenciaron la existencia de diferencias estadísticamente significativas en el número de repetidos alfa centroméricos para ciertos cromosomas entre la línea celular de CLL (CLL-AT) y las muestras de ADNs proveniente de humanos sanos (ver Fig. 16). Más específicamente y considerando el análisis estadístico realizado cromosoma por cromosoma se obtuvo los siguientes resultados:

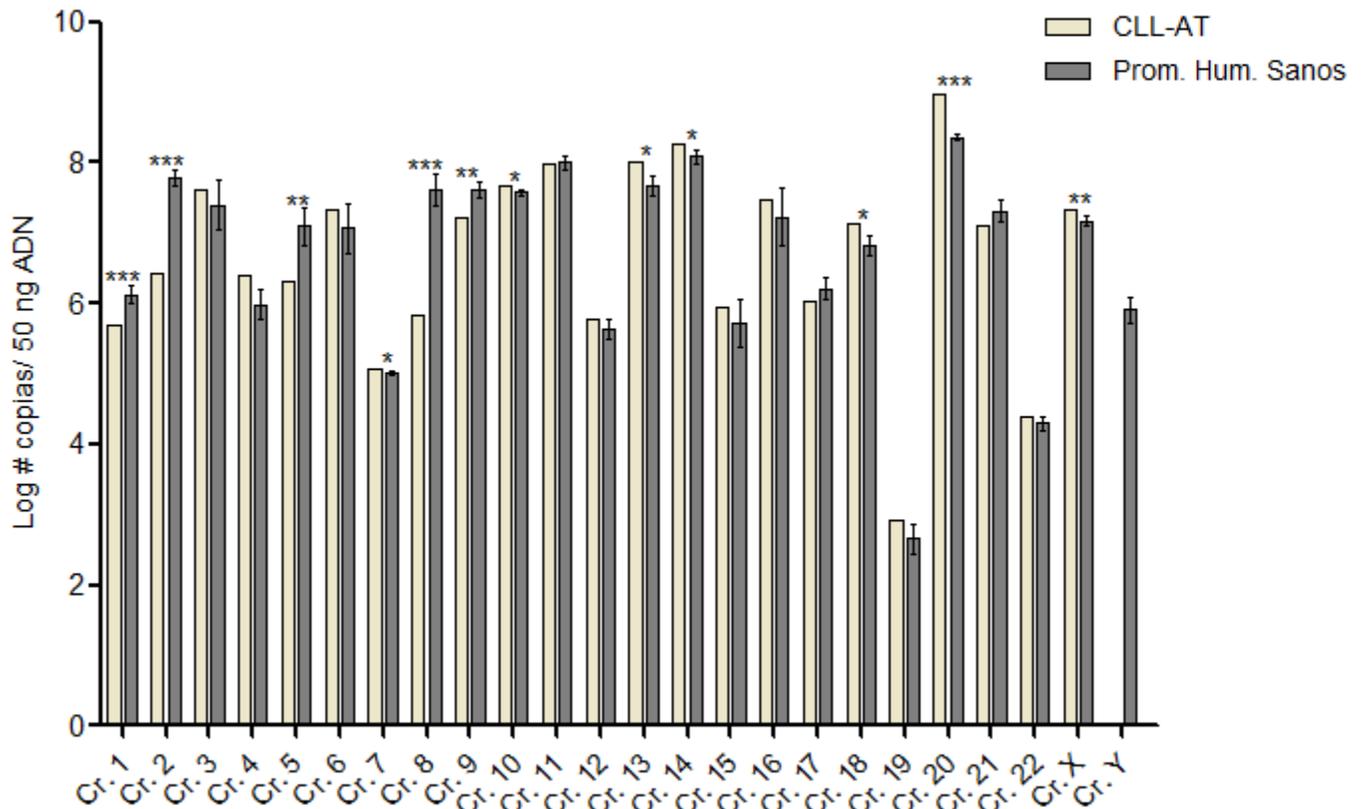
a) Para los cromosomas 1, 2, 8 y 20 grandes diferencias en el número de repetidos alfa de la línea CLL-AT fueron notadas en comparación a las muestras provenientes de donantes sanos. Estas diferencias fueron de gran significancia estadística (señalado con tres asteriscos en la Fig. 16). Asimismo, se puede observar que para el caso de los cromosomas 1, 2 y 8 de CLL-AT el número de repetidos alfa fue menor a lo encontrado en las muestras provenientes de donantes sanos, mientras que para el cromosoma 20 el número de repetidos alfa fue mayor en CLL-AT.

b) Para el caso de los centrómeros de los cromosomas 5, 9 y X la diferencia en el número de repetidos alfa entre CLL-AT y los humanos sanos fue así mismo significativa pero a un menor grado (señalada con dos asteriscos en la figura 16). Los cromosomas 5 y 9 dieron un número de secuencias de repetidos alfa menor en CLL-AT, mientras que para el cromosoma X este valor fue mayor en CLL-AT con respecto a los humanos sanos. En el caso de la variación observada de los repetidos alfa del cromosoma X, este análisis indica que la línea celular CLL-AT proviene de una mujer, con dos cromosomas X, en comparación con las células de donantes sanos que provienen de hombres con un solo cromosoma X respectivamente.

c) Para el caso de los centrómeros de los cromosomas 7, 10, 12 y 18 la diferencia en el número de repetidos alfa encontrado entre la línea celular CLL-AT y las muestras provenientes de donantes sanos fue significativa mayor aunque a una menor escala (un asterisco).

d) Debido al hecho de que la línea celular utilizada se trata de células provenientes de una mujer con leucemia linfocítica crónica, no se pudo realizar el estudio del número de repetidos alfa para el cromosoma Y.

e) Para los cromosomas 3, 4, 6, 11, 12, 15, 16, 17, 19, 21 y 22 no se obtuvo diferencias estadísticamente significativas en el número de repetidos alfa centroméricos entre los grupos



evaluados (Ver Fig. 16).

Figura 16. Cuantificación absoluta por PCR a tiempo real de repetidos alfa centroméricos cromosoma-específicos en una línea celular de LLC (CLL-AT) y en humanos sanos. En la figura se grafica el logaritmo del número de copias de secuencias de repetidos alfa cromosoma por cromosoma cuantificados cada 50ng de ADN de la línea celular de LLC (CLL-AT) y de humanos sanos. Cada reacción se realizó por duplicado. Las barras verticales claras corresponden al valor obtenido para la línea celular de LLC. Las barras verticales oscuras corresponden al promedio de la cuantificación de tres muestras de humanos sanos con su desviación estándar correspondiente. Se especifican los resultados del análisis estadístico test-T no-pareado de dos colas realizado: $p < 0.0001$ (***) gran diferencia significativa, $p < 0.001$ (**) diferencia muy significativa, $p < 0.01$ (*) diferencia significativa.

Con respecto al análisis del número de secuencias de HERVs-K111 insertos en los centrómeros de los cromosomas humanos, se pudo determinar que el número de K111 cuantificados en la línea celular CLL-AT fue menor, con una alta significancia estadística ($p \sim 0.0001$) respecto al número de K111 obtenido en los humanos sanos (ver Fig. 17).

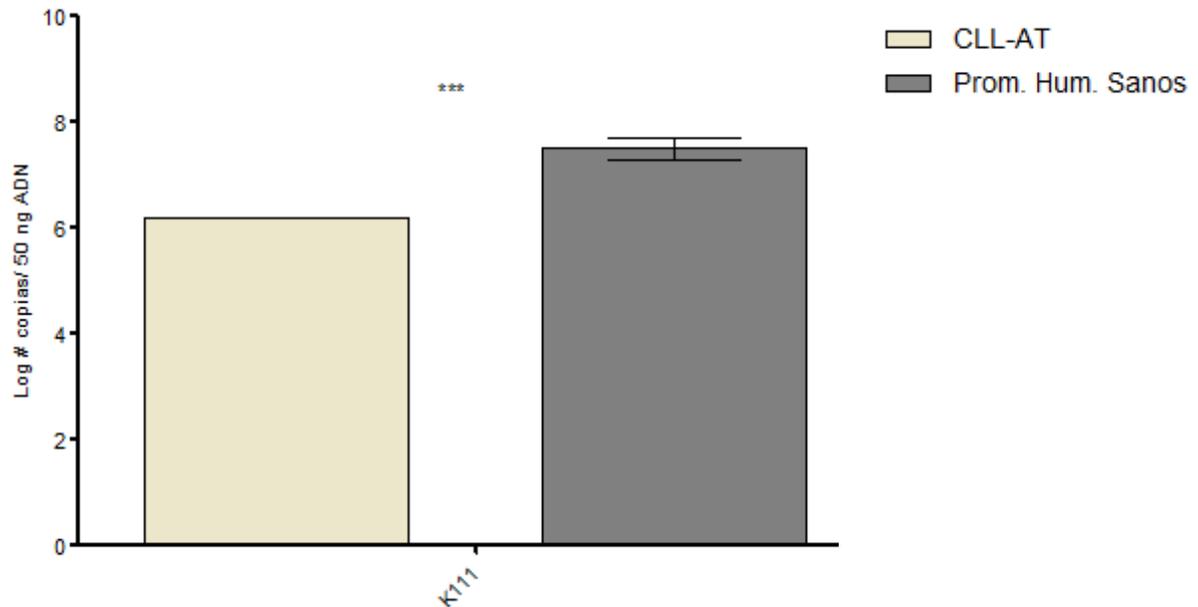


Figura 17. Cuantificación absoluta por PCR a tiempo real de HERVs centroméricos (K111) en una línea celular de LLC (CLL-AT) y en humanos sanos. En la figura se grafica el logaritmo del número de secuencias de retrovirus endógenos humanos (K111) insertos en los centrómeros de los cromosomas de la línea celular CLL-AT y de humanos sanos. Cada reacción se realizó por duplicado. La barra vertical clara corresponde al valor obtenido para la línea celular de LLC. La barra vertical oscura corresponde al promedio de la cuantificación de tres muestras de humanos sanos con su desvío estándar correspondiente. $p \sim 0.0001$ (***) gran diferencia significativa entre ambos grupos.

5. DISCUSIÓN

Hoy en día los esfuerzos por comprender las bases moleculares que están implicadas en la génesis y el desarrollo de enfermedades oncológicas son cada vez mayores. De la misma forma, cada vez aumentan los reportes que describen una asociación entre los HERVs y enfermedades del tipo neurológicas, autoinmunes y cáncer (Löwer R, 1999). Varias evidencias reportan un aumento de expresión tanto a nivel transcripcional como traduccional de los HERVs en una gran variedad de cánceres entre los cuales se encuentran leucemias y linfomas (Contreras-Galindo R, *et al.*, 2008; Depil S, *et al.*, 2002; Iwabuchi H, *et al.*, 2004; Romanish MT, *et al.*, 2010). Recientemente se han encontrado cientos de provirus de HERV-K distribuidos en las regiones centroméricas de al menos 15 cromosomas humanos denominados “K111” (Contreras-Galindo R, *et al.*, 2008; 2012). La compleja organización de secuencias repetitivas que conforman los centrómeros humanos así como la detección de los K111 insertos en ellos, hacen el acceso a la información de los centrómeros un verdadero desafío. Esto es aún más interesante si se considera a los centrómeros como el locus cromosomal esencial para la segregación de los cromosomas y la estabilidad del genoma (Aldrup-MacDonald & Sullivan, 2014).

En vistas del potencial que tienen los HERVs en contribuir en enfermedades hemato-oncológicas y el rol fundamental de los centrómeros para asegurar una correcta segregación de la información genética al momento de la división celular, el presente proyecto de maestría profundizó en el estudio de los HERVs y la estructura de repetidos alfa centroméricos en la leucemia linfocítica crónica (LLC).

¿Por qué elegimos a la LLC como modelo de estudio?

Porque se trata del tipo de leucemia más frecuente en occidente entre las poblaciones adultas de origen caucásico. Actualmente la ausencia de mutaciones a nivel de los genes que codifican para las cadenas pesadas de inmunoglobulinas (IgVH) constituye el principal indicador de mal pronóstico en estos pacientes. Los progresos obtenidos en el tratamiento hacen que frecuentemente se observen remisiones, incluso a nivel molecular, pero la mayoría de los pacientes recaen y existe consenso de que esta enfermedad es incurable (Damle RN, *et al.*, 1999; Hamblin TJ, *et al.*, 1999; Vasconcelos Y, *et al.*, 2003). Asimismo cabe mencionar que si bien se han logrado importantes avances en materia de conocimiento de la enfermedad, los sistemas de clasificación de Rai/Binet y los marcadores moleculares

que hoy en día se evalúan (perfil mutacional de las IgVH, aberraciones cromosómicas, expresión de ZAP-70 y CD38, entre otros) no son capaces de predecir completamente la progresión de la enfermedad y/o de explicar el perfil de progresión heterogéneo en el curso clínico de la LLC (Dighiero & Hamblin., 2008; Palacios F, *et al.*, 2010).

En este contexto, se plantearon para el presente proyecto de maestría dos objetivos diferentes pero complementarios. A continuación se discuten los resultados obtenidos en función de los objetivos propuestos.

5.1. Con respecto al objetivo general 1: Estudiar la posible relación entre la expresión de los genes de HERV-K en una patología hemato-oncológica como es la Leucemia Linfocítica Crónica (LLC).

Como se mencionó anteriormente, cada vez son más las evidencias que reportan una activación de los HERV-K (HML-2) en una gran variedad de cánceres como son el cáncer de ovario (Iramaneerat K, *et al.*, 2011; Wang-Johanning F, *et al.*, 2007), de mama (Ejthadi HD, *et al.*, 2005; Wang-Johanning F, *et al.*, 2003), melanomas (Büscher K, *et al.*, 2006; Serafino A, *et al.*, 2009; Singh S, *et al.*, 2009), próstata (Goering W, *et al.*, 2011; Ishida T, *et al.*, 2008), linfomas y leucemias (Contreras-Galindo R, *et al.*, 2008; Depil S, *et al.*, 2002; Iwabuchi H, *et al.*, 2004), entre otros.

Son varios los mecanismos por los cuales los HERVs podrían estar promoviendo el desarrollo de enfermedades oncológicas. La activación de sus genes puede resultar en efectos oncogénicos por la producción de proteínas codificadas por los HERVs (como Rec y Np9) o en la activación transcripcional de diferentes oncogenes celulares (Romanish MT, *et al.*, 2010). No debemos olvidar además la presencia de los elementos “solo LTR” dispersos en nuestro genoma otorgándole cierto grado de complejidad al mismo debido a su función promotora de la transcripción (Cohen CJ, *et al.*, 2009).

Por lo antedicho, en este proyecto nos planteamos estudiar tanto la expresión de *gag* como la expresión de la oncoproteína *np9* en pacientes con LLC con perfil IgVH mutado y no mutado, y comparar esta expresión con la encontrada en donantes sanos. A continuación se discuten los resultados obtenidos para cada objetivo específico planteado:

5.1.1. Objetivo específico 1: Puesta a punto de la metodología de RT-PCR para la obtención de todos los genes de HERV-K (*env*, *gag*, *pol*, *rec* y *np9*)

En una primera instancia nuestro trabajo se centró en realizar la puesta a punto de metodologías moleculares que permitieran la detección de la expresión de los diferentes genes de HERV-K mediante RT-PCR. Utilizando protocolos previamente descritos por otros autores, se procedió a la puesta a punto en nuestro laboratorio de la RT-PCR para la amplificación de los ARNm provenientes de los genes *gag* (Depil S, *et al.*, 2002), *pol* (Medstrand P, *et al.*, 1992) *env* (Zhao J, *et al.*, 2011) *rec* (Gonzalez-Hernández MJ, *et al.*, 2012) y *np9* (Gonzalez-Hernández MJ, *et al.*, 2012).

Con este fin se utilizaron dos líneas celulares (Jurkat y Daudi), en las cuales ya había sido previamente reportada la expresión de todos los genes de HERV-K estudiados (Renaudineau Y, *et al.*, 2005), así como una muestra proveniente de un donante sano.

Los resultados nos muestran la exitosa puesta a punto de todas estas metodologías ya que se pudo amplificar mediante RT-PCR todos los genes pertenecientes a HERV-K. Como era de esperarse, se obtuvo señal positiva en las muestras pertenecientes a las líneas celulares. Por otra parte existe una idea consenso de que los HERVs presentan un nivel de expresión basal en varios tipos de células humanas sanas (Armbruster V, 2002; Flockerzi A, *et al.*, 2008). Nuestro resultado es una prueba más a esta afirmación dado que obtuvimos producto de expresión de todos los genes de HERV-K (*gag*, *pol*, *env*, *np9* y *rec*) para las reacciones donde se añadió ADNc proveniente de PBMCs del donante sano seleccionado al azar.

5.1.2. Objetivo específico 2: Cuantificar los niveles de expresión relativa de los genes *np9* y *gag* de HERV-K en muestras de pacientes con LLC (perfil IgVH-M y IgVH-NM) con respecto a muestras de donantes sanos.

A fin de poder cumplir con este objetivo se realizó la cuantificación de los niveles de expresión relativa de los genes *np9* y *gag* de HERV-K mediante PCR en tiempo real. Se realizó la amplificación de 14 muestras de ADNc pertenecientes a pacientes con LLC de perfil IgVH-M, 11 muestras de ADNc pertenecientes a pacientes con LLC IgVH-NM y 6 muestras de ADNc de donantes sanos. Las medidas fueron realizadas por duplicado y se implementó la metodología del delta-delta Ct para evaluar los resultados.

A partir de nuestros resultados se evidencia que la actividad transcripcional de *np9* en pacientes con LLC fue más alta que en los PBMC provenientes de donantes sanos. Los resultados revelan que 18 de las 25 muestra analizadas (70% de los pacientes con LLC) expresan el gen *np9* de HERV-K al menos cinco veces más que lo que expresan las células de los donantes sanos.

Según lo reportado en la literatura, *np9* es sobreexpresado tanto en muestras pertenecientes a pacientes leucémicos, así como en una gran variedad de líneas celulares leucémicas. Chen y colaboradores identificaron a la proteína Np9 de los HERV-K tipo I, como un potente oncogen viral que promueve el crecimiento de células madre/progenitoras leucémicas humanas a través de la activación de diferentes vías de señalización celular como la ERK, AKT, Notch1, y β -catenina, siendo esta última esencial para la sobrevivencia de células madre leucémicas. En su trabajo se vio además que el silenciamiento de Np9 inhibe el crecimiento de células leucémicas mieloides y linfoblásticas, mientras que la expresión de Np9 promueve significativamente el crecimiento de células de leucemia tanto *in vitro* como *in vivo*. En su estudio reportaron además que los títulos de la proteína Np9 fueron sustancialmente elevados en pacientes con leucemia en comparación con pacientes sanos (56% vs 4,5%). Este hallazgo complementa los estudios que promueven la búsqueda de nuevas perspectivas para comprender la etiología de la leucemia humana y que proveen de nuevos blancos para el tratamiento la enfermedad (Chen T, *et al.*, 2013). Nuestros resultados confirman entonces una vez más la sobreexpresión de este oncogén de HERV-K asociado a procesos hemato-oncológicos.

Por otro lado, cuando se analizó la expresión relativa del gen *gag* observamos que solo un 33% de los pacientes con LLC expresan este gen al menos dos veces más en comparación con los donantes sanos. Los resultados obtenidos en el marco de esta tesis no coinciden totalmente con los resultados publicados previamente por el grupo de Depil y colaboradores en el 2002. Este grupo de investigación evaluó mediante PCR a tiempo real el nivel de expresión del gen *gag* de HERV-K (HML-2) en ocho muestras de pacientes con leucemia (3 de LLC, 2 de LMC, 2 de LMA, y una LLA) y obtuvo que la actividad transcripcional del gen *gag* fue de cinco a diez veces mayor en seis de las ocho muestras analizadas respecto a PBMCs normales (las 3 de LLC, 2 de LMA y una de LMC) (Depil S, *et al.*, 2002). Sin embargo, nuestros resultados revelan que solamente dos de los 25 pacientes con LLC analizados, exhiben una expresión del gen *gag* de al menos cinco veces mayor respecto a

los donantes sanos. Creemos que las diferencias obtenidas pueden deberse principalmente al bajo número de pacientes estudiado por el grupo de Depil. Tanto en el caso del análisis de *np9* como en el de *gag*, se observó una gran variabilidad entre los pacientes con LLC, no observándose una diferencia significativa entre los subgrupos de pacientes LLC mutados y no mutados ($p=0,65$ y $p= 0,09$ respectivamente). Asimismo, este resultado confirma la presencia de un nivel de expresión basal (en bajo nivel y similar) de estos genes de HERVs entre PBMCs provenientes de diferentes personas sanas.

Cabe mencionar que, de acuerdo con el objetivo de contribuir al conocimiento de las bases moleculares que rigen de fondo la LLC, el estudio aquí realizado constituye el primer reporte a nivel mundial de la sobreexpresión del gen *np9* de HERV-K en sangre periférica de pacientes con LLC (Fischer S, *et al.*, 2014).

5.1.3. Objetivo específico 3: Cuantificar la expresión del gen celular *c-myc* en las muestras de LLC y en muestras de donantes sanos y evaluar una posible relación significativa con los niveles de expresión del gen *np9* de HERV-K.

Es conocido que las proteínas Np9 y Rec, tienen un dominio de unión a la proteína en dedos de zinc de la leucemia promielocítica (PLZF). Funcionalmente, PLZF es un represor transcripcional de determinados genes, entre los cuales se encuentra el proto-oncogén *c-myc* (Denne M, *et al.*, 2007). Los detalles de la función específica que tiene la proteína c-Myc en las células continúan sin estar completamente entendidos, sin embargo, parece ser claro que c-Myc es importante para el incremento de la masa celular asociada con la entrada de las células en fase G0 a su ciclo celular. Asimismo, c-Myc puede inducir las células a su proliferación, su control de supervivencia y diferenciación (Patel JH, *et al.*, 2004). Un estudio de Larsson y colaboradores reportó que la inducción *in vitro* de células B de LLC de su fase G0 a G1 se vio acompañada de un dramático incremento en el nivel de expresión del gen *c-myc* (Larsson LG, *et al.*, 1991). Más recientemente, Palacios y colaboradores reportaron que la subpoblación de linfocitos B proliferantes en LLC presenta altos niveles de las proteínas Ki-67 y Bcl-2 y del ARNm de *c-myc*, sugiriendo el estado activado para este subset particular de células B en LLC (Palacios F, *et al.*, 2010). Considerando entonces el dominio de unión a PLZF que presenta la proteína Np9 de HERV-K, esta unión podría estar impidiendo la represión del proto-oncogén *c-myc* por parte de PLZF. De ocurrir esta idea, se debería

esperar entonces que el producto de expresión de *c-myc* se viera aumentado en aquellas células que presenten un elevado nivel de proteína Np9.

Es relevante aclarar que para nuestro estudio, el hecho de obtener una sobreexpresión del gen *np9* de HERV-K en el 70% de los pacientes con LLC analizados, no necesariamente se debe corresponder a una sobreproducción de la proteína Np9 en este grupo. De todas formas, nos resultó interesante evaluar si las células donde se obtuvo elevados niveles de expresión de *np9* en LLC presentaban además un incremento del nivel de expresión del gen celular *c-myc*. Para este análisis reagrupamos las muestras en tres subgrupos: un subgrupo correspondiente a los pacientes de LLC que dieron un elevado nivel de expresión del gen *np9* (mayor a cinco veces respecto a los donantes sanos); otro subgrupo correspondiente a las muestras de pacientes de LLC cuyo nivel de expresión de *np9* no fue tan elevado (menor a cinco veces respecto a los donantes sanos); y un tercer subgrupo correspondiente a las muestras provenientes de los donantes sanos. Se realizaron los estudios de expresión relativa mediante una PCR en tiempo real.

Desafortunadamente nuestros estudios no mostraron diferencias significativas del nivel de expresión del gen *c-myc* entre los tres subgrupos analizados ($p=0,19$). Comparando el resultado obtenido con respecto al esperado, varias razones podrían estar explicando la diferencia:

a) Una de las explicaciones posibles es que al estar trabajando con poblaciones de PBMC, el número de células que puedan estar sobre-expresando *c-myc* (por ejemplo las células B pertenecientes a la subpoblación proliferante) sea muy pequeño que no justifique un resultado significativo al evaluar toda la población de células en su conjunto.

b) Otra posible explicación puede ser debido a que la sobreexpresión de *c-myc* a través de la posible vía de regulación celular Np9-PLZF sea transitoria (Denne M, *et al.*, 2007), y que nuestras muestras no coincidan con el momento de un incremento en la expresión del proto-oncogén celular.

c) Es posible también que Np9 esté actuando, promoviendo la proliferación celular, mediante la interacción con otra vía de señalización. Se ha detectado que Np9, además de presentar un dominio de unión a PLZF, puede también unirse e interferir funcionalmente con el ligando de la proteína X Numb, afectando la vía de señalización celular Notch-Numb (Armbruster V,

et al., 2004). Esta vía se ha visto implicada en la regulación de la proliferación de células de cáncer de mama, próstata, leucemias y cánceres de la línea germinal (Denne M, *et al.*, 2007; Beverly & Capobianco, 2003). Asimismo se han detectado modificaciones por parte de Np9 en otras vías de señalización como la β -catenina, ERK, y AKT (Chen T, *et al.*, 2013).

d) Por último podría ser que la sobreexpresión del ARNm de *np9* de HERV-K no se corresponda con un mayor nivel de producción de la proteína Np9.

5.1.4. Objetivo específico 4: Estudiar la variabilidad genética del gen *np9* en pacientes con LLC y en donantes sanos mediante aproximaciones bioinformáticas.

A fin de poder establecer la presencia de variantes a nivel de la secuencia de *np9* se realizó la amplificación mediante RT-PCR del gen *np9* completo a partir del ADNc de 14 muestras (10 de pacientes con LLC y 4 de donantes sanos). Posteriormente el producto de PCR fue purificado y secuenciado.

Las secuencias obtenidas fueron comparadas con las pocas secuencias disponibles en la base de datos. Las secuencias encontradas en la base de datos del GenBank exhibieron casi un 100% de similaridad con las secuencias de *np9* de HERV-K obtenidas en el presente proyecto. Es importante recordar que los HERVs que expresan *np9* no son capaces de expresar secuencias *rec* debido a la delección de 292pb específica de los HERVs-K tipo I. La expresión de *np9* es el resultado de la presencia de sitios dadores de *splicing* alternativo producidos por un reemplazo en las posiciones nucleotídicas 6495 y 6496 de A y G de HERV-K tipo II, por una T y A en los HERV-K tipo I (Armbruster V, *et al.*, 2002; Mount SM, 1982).

El ensayo de PCR reveló, como era de esperar, la presencia de unas pocas variantes de la secuencia de *np9* que difieren en posiciones nucleotídicas y que están asociadas con cambios en la secuencia de aminoácidos una vez traducida su secuencia. Dos de estas variaciones las proponemos como posibles polimorfismos del gen *np9* de HERV-K presentes en la población humana. Se define a un polimorfismo genético a variantes en la secuencia de un gen que han sido fijadas entre individuos de una población. Considerando que nuestro genoma presenta múltiples copias de secuencias provirales de HERV-K, podría esperarse que variantes de genes de *np9* se hayan fijado en la población. Asimismo, proponemos como mutaciones puntuales a los tres cambios nucleotídicos restantes dado que estos cambios han sido detectados en todas las muestras analizadas para el proyecto como la

presencia de una única base en ese sitio diferente a las secuencias de referencia. Cabe mencionar que las mutaciones detectadas no coinciden con el sitio de interacción de la proteína Np9 con PLZF, y tampoco pertenecen a alguna de las tres señales de localización nuclear de Np9 (NLS por su sigla en inglés) según los datos reportados por Denne y colaboradores (2007). Si bien no se obtuvo un marcador molecular característico para las muestras de los pacientes con LLC, la presencia de los posibles polimorfismos detectados podría proveer de interesante información para realizar ensayos de los mecanismos evolutivos asociados a la evolución del genoma humano retroviral-asociado.

5.2. Con respecto al objetivo general 2: Determinar si existe algún patrón diferencial en la abundancia de repetidos alfa y secuencias de HERVs (K111) en los centrómeros de una línea celular de LLC (CLL-AT) y su comparación con donantes sanos.

Reordenamientos del genoma a gran escala como las translocaciones, inversiones y deleciones de regiones cromosomales de varias megabases de largo, son característicos de la inestabilidad genómica observada en diversos estados de patologías humanas (Longo MS, *et al.*, 2009). Por ejemplo, las translocaciones, a menudo resultan en desequilibrios cromosómicos específicos de tumores que están asociados con la oncogénesis en la leucemia y tumores sólidos. El tamaño de estas regiones saltadoras varían entre los pacientes, y la región del genoma involucrada puede tratarse de oncogenes asociados a una transformación a estados avanzados en pacientes con leucemia y linfoma (Tanaka & Kamada, 1998).

En la LLC las aberraciones cromosómicas frecuentemente observadas producto de un reordenamiento genético son: deleción en cromosoma 13q, trisomía 12, y deleción en cromosoma 11q. En menor frecuencia se observa la deleción en el cromosoma 17p que afecta la proteína TP53. La presencia de la deleción en 17p o 11q es asociada a una mala prognosis y predominan en estados avanzados de LLC y en pacientes con genes IgVH-NM, mientras que la deleción en 13q o un cariotipo normal son asociadas con una buena prognosis, estados iniciales de la enfermedad, y con genes IgVH-M. Existen controversias sobre si la trisomía 12 se corresponde a un estatus no mutado y a una mal prognosis de la enfermedad (Dighiero & Hamblin, 2008; Kröber A, *et al.*, 2002).

Cada uno de estos reordenamientos específicos, translocaciones intercromosomales, deleciones o inversiones intracromosomales, requieren la ruptura de la doble cadena del ADN. Varios estudios de análisis del genoma completo detectaron que existen regiones, o

sitios frágiles, muy susceptibles a una ruptura y que los mismos se encuentran distribuidos en zonas específicas del genoma de los mamíferos (Murphy WJ, *et al.*, 2005). Muchos de estos sitios frágiles son llamados “puntos de ruptura evolutivos” (EB por su sigla en inglés) dado que se localizan en regiones específicas del genoma “reusadas” durante la evolución de cariotipos de diferentes especies. Cuando se considera la trayectoria filogenética entre segmentos de cromosomas ortólogos de distintas especies, muchos de estos EB se corresponden con secuencias centroméricas ancestrales así como también con la formación de nuevos centrómeros. El grupo de investigación de Longo y colaboradores analizó el punto de ruptura evolutivo ortólogo humano 14q32.33, una translocación que se ha encontrado asociada a varios cánceres incluyendo mielomas múltiples y leucemias de células plasmáticas, y encontró que su EB se trata de una distribución conservada de elementos repetitivos (repetidos alfa) ricos en secuencias LINES, y de retrovirus endógenos (ERVs) (Longo PG, *et al.*, 2007). Estos hallazgos son interesantes para nuestro estudio si tenemos en cuenta que la LLC presenta ciertas aberraciones cromosómicas que la definen.

Estudios recientes del grupo de investigación de Contreras-Galindo y colaboradores (2008; 2012 y 2013), han reportado la detección de cientos de variantes de HERV-K dispersos entre los centrómeros de al menos 15 cromosomas humanos, y que su expansión tomó fuerza durante la evolución de los homínidos. Análisis de secuencias indicaron que los provirus de K111 se expandieron en su número de copias tanto dentro como entre varios cromosomas humanos mediante un mecanismo semejante a la recombinación, proporcionando interesante evidencia de un probable mecanismo de recombinación homóloga entre centrómeros durante la evolución de los humanos. También se han detectado inserciones de secuencias “solo LTR” de K111 en ADN humano pero no en chimpancés u otros primates, sugiriendo que algunos provirus de secuencia completa de K111 han removido sus genes internos por recombinación entre los LTRs solamente durante la evolución de los humanos (Contreras-Galindo R, *et al.*, 2013).

La compleja organización de secuencias repetitivas que conforman los centrómeros humanos así como la detección de los K111 insertos en ellos, hacen el acceso a la información de los centrómeros un verdadero desafío. Como mencionamos anteriormente, los repetidos alfa se definen como una unidad o monómero de ~ 171pb, que se repite organizado en orientación cabeza-cola y, con frecuencia, se extiende con limitadas interrupciones por millones de bases. Estas unidades se repiten a su vez en una fracción larga ininterrumpida de repetidos de alto orden (HORs). Todos los cromosomas humanos

parecen tener un arreglo en HORs tipo cromosoma-específico en sus centrómeros con un porcentaje de identidad ~97-99% (Aldrup-MacDonald & Sullivan, 2014; Brown & O'Neill, 2014).

Acceder al conocimiento de la estructura que identifica a cada centrómero humano puede brindar una importante información sobre cómo será la segregación de los cromosomas al momento de la división celular (por mitosis o meiosis). Una segregación incorrecta dará lugar a células hijas con diferentes cariotipos que, de ser viables, pueden derivar en la formación de células con aberraciones cromosómicas.

El estudio de la abundancia de repetidos alfa y de retrovirus endógenos a nivel de los centrómeros es de gran interés dado que un patrón característico detectado puede constituir una marca molecular en la enfermedad de LLC. Además pueden representar un interesante foco de estudio ya que las variaciones encontradas podrían estar jugando algún rol en la generación y evolución de esta patología que aún desconozcamos.

Por lo antedicho, es de gran interés para nuestro trabajo poder establecer si existe algún patrón diferencial en la abundancia de repetidos alfa así como de secuencias pertenecientes a HERVs (K111) en los centrómeros de una línea celular de LLC (CLL-AT) y su comparación con donantes sanos.

5.2.1. Objetivo específico 1: Puesta a punto de la metodología de PCR a tiempo real para obtener las secuencias centroméricas de repetidos alfa de manera específica para cada cromosoma humano.

La puesta a punto de esta metodología, consistió tanto en el diseño de los *primers* específicos que permitieran la obtención de forma independiente de la mayoría de los repetidos alfa centroméricos cromosoma-específicos, así como en el ajuste de las condiciones de tiempos y temperaturas para las diferentes PCRs. De esta forma se generaron herramientas fundamentales para la identificación de secuencias centroméricas únicas para la gran mayoría de los cromosomas. El desafío de lograr la metodología experimental para conocer el patrón de abundancia de secuencias de repetidos alfa centroméricos para cada cromosoma se hizo aún más interesante al considerar a estos sitios como aquellos responsables de asegurar una correcta segregación de los cromosomas al momento de la división celular.

El abordaje de nuestro objetivo específico permitió identificar secuencias centroméricas únicas para la gran mayoría de los cromosomas. Lo interesante de esta información es su potencial utilidad para comprender mejor el rol de la función centromérica y su posible asociación con la LLC así como también con otras enfermedades genéticas y células de cáncer. Además, estos marcadores moleculares centroméricos podrían ser de utilidad para el estudio de interacciones del ADN y posibles eventos de recombinación entre centrómeros durante diferentes fases del ciclo celular. En lo que refiere a la clínica práctica, los marcadores centroméricos podrían proveer una buena herramienta para diagnóstico, ensayos y tratamientos de enfermedades de origen centromérico.

5.2.2. Objetivo específico 2: Cuantificar la abundancia de repetidos alfa y de retrovirus endógenos humanos (K111) en los centrómeros de la línea celular CLL-AT y en células humanas sanas. Evaluar si existen diferencias significativas en la estructura centromérica entre ambos grupos.

Con el fin de poder determinar el patrón de abundancia de las secuencias de repetidos alfa cromosoma-específicas así como también de la secuencias centroméricas de retrovirus endógenos K111 en los centrómeros de la línea celular CLL-AT en comparación con células humanas sanas, se realizó una cuantificación absoluta mediante PCR a tiempo real de los mismos. Los resultados muestran que existen diferencias significativas (ganancia o pérdida) de estos marcadores en la línea celular humana de LLC (CLL-AT) respecto a las células humanas sanas analizadas.

El análisis presentó diferencias estadísticamente significativas para ciertos cromosomas de la línea celular de LLC (CLL-AT) al compararse con el número de repetidos alfa presentes en los ADN de los humanos sanos. Se obtuvo que el número de repetidos alfa para los cromosomas 1, 2, 5, 8 y 9 en CLL-AT fue significativamente menor que el obtenido en los humanos sanos. Si bien estos cromosomas no son los involucrados en las principales aberraciones cromosómicas que definen la LLC, existen reportes que han asociado ciertos rearrreglos cromosómicos en alguno de ellos. A modo de ejemplo se ha reportado la ganancia del brazo corto del cromosoma 2 (2p) como una aberración cromosómica frecuente en estadios avanzados de pacientes con LLC no tratados. Esta región 2p frecuentemente incluye dos oncogenes, *rel* y *mycn*, los cuales han sido asociados también a una mal prognosis de la enfermedad (Chapiro E, *et al.*, 2010).

Nuestros resultados también evidencian un número de repetidos alfa significativamente mayor en la línea celular de LLC en comparación con las células provenientes de donantes sanos para los cromosomas 10, 13, 14, 18, 20, y el cromosoma X. En la literatura se ha asociado una deleción en el cromosoma 13 en más del 50% de los pacientes con LLC como una anomalía citogenética comúnmente detectada y asociada con un buen pronóstico de la LLC. No se encuentran reportes que asocien translocaciones para el resto de los cromosomas donde se detectó un incremento en el número de repetidos alfa centroméricos.

En lo que refiere a la cuantificación de las secuencias provirales de K111, se obtuvo una reducción significativa del número total de estos retrovirus para el modelo de LLC. Esta observación sugiere una posible deleción de los K111 en los centrómeros de las células de LLC, lo cual podría considerarse como un posible marcador molecular de esta patología.

6. CONCLUSIONES

A continuación se resumen las principales conclusiones que derivan del estudio de los HERVs y de la abundancia de repetidos alfa a nivel centromérico en leucemia linfocítica crónica (LLC):

- Se logró la puesta a punto en forma exitosa de la metodología experimental que permite la amplificación mediante RT-PCR de todos los genes pertenecientes a HERV-K (*gag*, *pol*, *env*, *rec* y *np9*).
- Se evidenció que la actividad transcripcional del gen *np9* de HERV-K se encuentra sobreexpresada (al menos cinco veces más) en la mayoría de las muestras de pacientes con LLC analizadas al ser comparadas con PBMC provenientes de donantes sanos. Este resultado confirma una vez más la sobreexpresión de este oncogén de HERV-K asociada a procesos leucémicos, y constituye el primer reporte a nivel mundial de una sobreexpresión del gen *np9* de HERV-K en sangre periférica de pacientes con LLC (Fischer S, *et al.*, 2014).
- El estudio del nivel de expresión del gen *gag* de HERV-K evidenció que solo un 33% de los pacientes con LLC expresaron este gen al menos dos veces más en comparación a los donantes sanos.
- Tanto en el caso del análisis de *np9* como el de *gag* se observó una gran variabilidad entre los pacientes de LLC analizados, no observándose una diferencia significativa entre los subgrupos de pacientes LLC mutados y no mutados.
- No se evidenció una diferencia significativa del nivel de expresión del oncogén celular *c-myc* entre los grupos con alta expresión de *np9*, baja expresión de *np9* y donantes sanos, lo que nos estaría indicando que la acción de *np9* no sería a través de la vía *c-myc*, o bien que la sobreexpresión del ARNm de *np9* no se corresponda con igual expresión a nivel proteico.
- Se observó un alto grado de conservación en la secuencia del gen que codifica para la proteína Np9 en las diferentes muestras amplificadas. Se encontraron dos posibles

polimorfismos fijados en la población y tres sitios nucleotídicos donde se observaron mutaciones puntuales.

- Se realizó la puesta a punto para la identificación y cuantificación de los repetidos alfa centroméricos únicos para la gran mayoría de los cromosomas humanos. De esta forma se propone una nueva metodología experimental para acceder al estudio de la estructura de los centrómeros en forma cromosoma-específica.
- El estudio a nivel de los centromeros reveló una reducción global significativa del número total de retrovirus endógenos (K111) así como también diferencias (ganancia o pérdida) en el número de repetidos alfa de determinados cromosomas en el modelo de línea celular CLL-AT.

7. PERSPECTIVAS

El presente trabajo permitió la formulación de interesantes ideas para la continuación del estudio de los HERVs y abundancia de repetidos alfa en la leucemia linfocítica crónica. A continuación se exponen las perspectivas para la continuación de este trabajo:

- Evaluar si la sobreexpresión del ARNm de *np9* de HERV-K detectada se corresponde con un mayor nivel de producción de la proteína Np9 en las células de LLC. Esta información nos permitiría responder a qué nivel de regulación (ARNm o proteína) se asocia la expresión de este gen de los HERV-K en la LLC.
- Con respecto a los resultados obtenidos del nivel de expresión del gen *gag* de HERV-K en nuestras muestras de PBMCs de pacientes con LLC, sería interesante repetir el procedimiento realizando una purificación previa de las células B de LLC y analizar el resultado específicamente en ellas.
- Realizar el estudio de cuantificación del nivel de expresión por PCR a tiempo real para los genes *env*, *pol*, y *rec* en las muestras de LLC respecto a muestras de donantes sanos.
- Separar previamente las células B proliferantes en base a los marcadores moleculares de agresividad que exponen en su superficie y repetir el procedimiento de cuantificación por PCR a tiempo real para los genes *np9* y *c-myc* y poder confirmar si existe o no una posible relación entre ellos.
- Ampliar los posibles efectos de la sobreexpresión del gen *np9* en LLC evaluando intermediarios de otras vías de señalización celular. Se podría analizar intermediarios de la vía Numb/Notch sabiendo que la proteína Np9 posee un dominio de unión al ligando de Numb. También se podrían analizar intermediarios de la vía celular AKT dado que la misma se encuentra fuertemente activada en las células B leucémicas de los pacientes de LLC que manifiestan una enfermedad progresiva.

- Realizar diferentes abordajes bioinformáticos para determinar el momento en que los polimorfismos de *np9* detectados fueron fijados en la población humana.
- Realizar una nueva cuantificación absoluta por PCR a tiempo real de los marcadores moleculares centroméricos para un número considerable de muestras de ADNs provenientes de pacientes con LLC. La posterior comparación con los resultados obtenidos para el modelo de línea celular de LLC (CLL-AT), le daría un mayor soporte a las diferencias en el número de estas marcas moleculares con respecto a los humanos sanos.
- Utilizar la metodología de cuantificación de los repetidos alfa cromosoma-específicos y provirus K111 centroméricos en otras patologías que se saben definidas por defectos genéticos.

8. REFERENCIAS BIBLIOGRÁFICAS

Aldrup-MacDonald ME, Sullivan BA. (2014). The Past, Present, and Future of Human Centromere Genomics. *Genes*, 5(1), 33-50.

Alexandrov I, Kazakov A, Tumeneva I, Shepelev V, Yurov Y. (2001). Alpha-satellite DNA of primates: old and new families. *Chromosoma*, 110(4), 253–66.

American Cancer Society. (2014). <http://www.cancer.org/cancer/leukemia-chroniclymphocyticcll/detailedguide/leukemia-chronic-lymphocytic-key-statistics>

Armbruester V, Sauter M, Roemer K, Best B, Hahn S, Nty A, ... Mueller-Lantsch N. (2004). Np9 Protein of Human Endogenous Retrovirus K Interacts with Ligand of Numb Protein X. *J Virol.*, 78(19), 10310–19.

Armbruester V, Sauter M, Krautkraemer E, Meese E, Kleiman A, Best B, Roemer K, Mueller-Lantsch N. (2002). A novel gene from the human endogenous retrovirus K expressed in transformed cells. *Clin Cancer Res.*, 8(6), 1800–7.

Balada E, Ordi-Ros J, Vilardell-Tarrés M. (2009). Molecular mechanisms mediated by human endogenous retroviruses (HERVs) in autoimmunity. *Rev Med Virol.*, 19(5), 273-86.

Baldini A, Smith DI, Rocchi M, Miller OJ, Miller DA. (1989). A human alphoid DNA clone from the EcoRI dimeric family: genomic and internal organization and chromosomal assignment. *Genomics*, 5(4), 822–8.

Bannert N, Kurth R. (2004). Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A.*, 101(Suppl 2), 14572-9.

Barbulescu M, Turner G, Seaman MI, Deinard AS, Kidd KK, Lenz J. (1999). Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. *Curr Biol.*, 9(16), 861-8.

Beverly LC, Capobianco AJ. (2003). Perturbation of Ikaros isoform selection by MLV integration is a cooperative event in Notch(IC)-induced T cell leukemogenesis. *Cancer Cell*, 3(6), 551-64.

Bianchi S, Moreno P, Landoni AI, Naya H, Oppezzo P, Dighiero G, Gabús R, Pritsch O. (2010). Immunoglobulin heavy chain VDJ gene rearrangement and mutational status in Uruguayan patients with chronic lymphocytic leukemia. *Leukemia & Lymphoma*, 51(11), 2070-8.

Binet JL, Lepoprier M, Dighiero G, Charron D, D'Athis P, Vaugier G, Beral HM, Natali JC, Raphael M, Nizet B, Follezu JY. (1997). A clinical staging system for chronic lymphocytic leukemia: prognostic significance. *Cancer*, 40(2), 855-64.

Blaise S, Ruggieri A, Dewannieux M, Cosset FL, Heidmann T. (2004). Identification of an envelope protein from the FRD family of human endogenous retroviruses (HERV-FRD) conferring infectivity and functional conservation among simians. *J Virol.*, 78(2), 1050-4.

Brodsky I, Foley B, Haines D, Johnston J, Cuddy K, Gillespie D. (1993)a. Expression of HERV-K proviruses in human leukocytes. *Blood*, 81(9), 2369–74.

Brodsky I, Foley B, Gillespie D. (1993)b. Expression of human endogenous retrovirus (HERV-K) in chronic myeloid leukemia. *Leuk Lymphoma.*, 11(Suppl 1), 119–23.

Brown JD, O'Neill RJ. (2014). The Evolution of Centromeric DNA Sequences. *eLS*, 1-11.

Büscher K, Hahn S, Hofmann M, Trefzer U, Ozel M, Sterry W, Löwer J, Löwer R, Kurth R, Denner J. (2006). Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines. *Melanoma Res.*, 16(3), 223-34.

Caligaris-Cappio, F. (2003). Role of the microenvironment in chronic lymphocytic leukemia. *Br J Haematol.*, 123(3), 380-8.

Chapiro E, Leporrier N, Radford-Weiss I, Bastard C, Mossafa H, Leroux D, Tigaud I, ...Nguyen-Khac F. (2010). Gain of the short arm of chromosome 2 (2p) is a frequent recurring chromosome aberration in untreated chronic lymphocytic leukemia (CLL) at advanced stages. *Leuk Res.*, 34(1), 63-8.

Chen T, Meng Z, Gan Y, Wang X, Xu F, Gu Y, Xu X, Tang J, Zhou H, Zhang X, Gan X, Van Ness C, Xu G, Huang L, Zhang X, Fang Y, Wu J, Zheng S, Jin J, Huang W, Xu R. (2013). The viral oncogene Np9 acts as a critical molecular switch for co-activating beta-catenin,

ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells. *Leukemia*, 27(7), 1469–78.

Chiorazzi, N. (2007). Cell proliferation and death: Forgotten features of chronic lymphocytic leukemia B cells. *Best Pract Res Clin Haematol.*, 20(3), 399–413.

Chomczynski P, Sacchi N. (2006). The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc.*, 1(2), 581-5.

Cohen CJ, Lock WM, Mager DL. (2009). Endogenous retroviral LTRs as promoters for human genes: a critical assessment. *Gene*, 448(2), 105-14.

Conrad B, Weissmahr RN, Böni J, Arcari R, Schüpbach J, Mach B. (1997). A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. *Cell*, 90(2), 303-13.

Contreras-Galindo R, Kaplan MH, He S, Contreras-Galindo AC, Gonzalez-Hernandez MJ, Kappes F, Dube D, Chan SM, Robinson D, Meng F, Dai M, Gitlin SD, Chinnaiyan AM, Omenn GS, Markovitz DM. (2013). HIV infection reveals widespread expansion of novel centromeric human endogenous retrovirus. *Genome Res.*, 23(9), 1505-13.

Contreras-Galindo R, Kaplan MH, Contreras-Galindo AC, Gonzalez-Hernandez MJ, Ferlenghi I, Giusti F, Lorenzo E, Gitlin SD, Dosik MH, Yamamura Y, Markovitz DM. (2012). Characterization of human endogenous retroviral elements in the blood of HIV-1-infected individuals. *J Virol.*, 86(1), 262-76.

Contreras-Galindo R, Kaplan MH, Leissner P, Verjat T, Ferlenghi I, Bagnoli F, Giusti F, Dosik MH, Hayes DF, Gitlin SD, Markovitz DM.. (2008). Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. *J Virol.*, 82(19), 9329-36.

Contreras-Galindo R, López P, Vélez R, Yamamura Y.(2007). HIV-1 infection increases the expression of human endogenous retroviruses type K (HERV-K) in vitro. *AIDS Res Hum Retroviruses.*, 23(1),116-22.

- Cordaux R, Batzer M. (2009). The impact of retrotransposons on human genome evolution. *Nat Rev Genet.*, 10(10), 691–703.
- Dameshek W. (1967). Chronic lymphocytic leukemia – an accumulative disease of immunologically incompetent lymphocytes. *Blood*, 29(4), 566–84.
- Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, ...Chiorazzi N. (1999). Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood*, 94(6), 1840-7.
- Deaglio S, Malavasi F. (2009). Chronic lymphocytic leukemia microenvironment: shifting the balance from apoptosis to proliferation. *Hematol.*, 94(6), 752-6.
- Denne M, Sauter M, Armbruster V, Licht JD, Roemer K, Mueller-Lantsch N. (2007). Physical and functional interactions of human endogenous retrovirus proteins Np9 and Rec with the promyelocytic leukemia zinc finger protein. *J Virol.*, 81(11), 5607–16.
- Depil S, Roche C, Dussart P, Prin L. (2002). Expression of a human endogenous retrovirus HERV-K in the blood cells of leukemia patients. *Leukemia*, 16(2), 254-9.
- Dighiero G, Hamblin TJ. (2008). Chronic lymphocytic leukaemia. *Lancet.*, 371(9617), 1017-29.
- Downey RF, Sullivan FJ, Wang-Johanning F, Ambs S, Giles FJ, Glynn SA. (2014). Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice? *Int. J. Cancer.*, 01-9.
- Ejthadi HD, Martin JH, Junying J, Roden DA, Lahiri M, Warren P, Murray PG, Nelson PN. (2005). A novel multiplex RT-PCR system detects human endogenous retrovirus K in breast cancer. *Arch virol.*, 150(1), 177-84.
- Fischer S, Echeverría N, Moratorio G, Landoni AI, Dighiero G, Cristina J, Oppezio P, Moreno P. (2014). Human endogenous retrovirus np9 gene is over expressed in chronic lymphocytic leukemia patients. *Leuk Res Rep.*, 3(2):70-2.

Flockerzi A, Ruggieri A, Frank O, Sauter M, Maldener E, Kopper B, ... Mayer J. (2008). Expression patterns of transcribed human endogenous retrovirus HERV-K (HML-2) loci in human tissues and the need for a HERV Transcriptome Project. *BMC genomics*, 9(1), 354.

Fuchs NV, Kraft M, Tondera C, Hanschmann KM, Löwer J, Löwer R. (2011). Expression of the human endogenous retrovirus (HERV) group HML-2/HERV-K does not depend on canonical promoter elements but is regulated by transcription factors Sp1 and Sp3. *J Virol.*, 85(7), 3436-48.

Ghia P, Chiorazzi N, Stamatopoulos K. (2008). Microenvironmental influences in chronic lymphocytic leukaemia: the role of antigen stimulation. *J Intern Med.*, 246 (6), 549-62.

Goering W, Ribarska T, Schulz WA. (2011). Selective changes of retroelement expression in human prostate cancer. *Carcinogenesis*, 32(10), 1484–92.

Gonzalez-Hernandez MJ, Swanson MD, Contreras-Galindo R, Cookinham S, King SR, Noel RJ, ...Markovitz DM. (2012). Expression of Human Endogenous Retrovirus Type K (HML-2) Is Activated by the Tat Protein of HIV-1. *J Virol.*, 86(15), 7790–7805.

Griffiths DJ. (2001). Endogenous retrovirus and the human genome sequence. *Gen Biol.*, 2(6), 1017.1-5.

Hamblin TJ. (2005). Chronic lymphocytic leukaemia: clinical translations of biological features. *Curr Top Microbiol Immunol.*, 294, 165-85.

Hamblin TJ, Orchard JA, Ibbotson RE, Davis Z, Thomas PW, Stevenson FK, Oscier DG. (2002). CD38 expression and immunoglobulin variable region mutations are independent prognostic variables in chronic lymphocytic leukemia, but CD38 expression may vary during the course of the disease. *Blood.*, 99(3), 1023-9.

Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. (1999). Unmutated Ig VH genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood*, 94(6), 1848-54.

Hernández P. (1999). Leucemia linfoide crónica. Aspectos clínicos. *Rev Cubana Hematol Inmunol Hemoter.*, 15(1), 7-20.

Hohn O, Hanke K, Bannert N. (2013). HERV-K(HML-2), the best preserved family of HERVs: endogenization, expression, and implications in health and disease. *Front Oncol.*, 3, 246.

Hsiao FC, Tai AK, Deglon A, Sutkowski N, Longnecker R, Huber BT. (2009). EBV LMP-2A employs a novel mechanism to transactivate the HERV-K18 superantigen through its ITAM. *Virology*, 385(1), 261-266.

Hughes JF, Coffin JM. (2005). Human endogenous retroviral elements as indicators of ectopic recombination events in the primate genome. *Genetics*, 171(3), 1183-94.

Hughes JF, Coffin JM. (2004). Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: Implications for human and viral evolution. *Proc Natl Acad Sci.*, 101(6), 1668–72.

Hyden KE. (2012). Human centromere genomics: now it's personal. *Chromosome Res.*, 20(5) 621–633.

Iramaneerat K, Rattanatunyong P, Khemapech N, Triratanachat S, Mutirangura A. (2011). HERV-K hypomethylation in ovarian clear cell carcinoma is associated with a poor prognosis and platinum resistance. *International Journal of Gynecological Cancer*, 21(1), 51-7.

Ishida T, Obata Y, Ohara N, Matsushita H, Sato S, Uenaka A, ...Nakayama E. (2008). Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity. *Cancer Immunity: a Journal of the Academy of Cancer Immunology*, 8.

Iwabuchi H, Kakihara T, Kobayashi T, Imai C, Tanaka A, Uchiyama M, Fukuda T. (2004). A gene homologous to human endogenous retrovirus over expressed in childhood acute lymphoblastic leukemia. *Leuk lymphoma.*, 45(11), 2303-06.

Jern P, Sperber GO, Blomberg J. (2005). Use of endogenous retroviral sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. *Retrovirology*, 2,50.

Katoh I, Mírová A, Kurata SI, Murakami Y, Horikawa K, Nakakuki N, ... Hirai H. (2011). Activation of the long terminal repeat of human endogenous retrovirus K by melanoma-specific transcription factor MITF-M. *Neoplasia*, 13(11), 1081-IN42.

Kleiman A, Senyuta N, Tryakin A, Sauter M, Karseladze A, Tjulandin S, Gurtsevitch V, Mueller-Lantsch N. (2004). HERV-K (HML-2) gag/env antibodies as indicator for therapy in patients with germ cell tumors. *Int J cancer.*, 110(3), 459-61.

Knössl M, Löwer R, Löwer J. (1999). Expression of the human endogenous retrovirus HTDV/HERV-K is enhanced by cellular transcription factor YY1. *J Virol.*, 73(2),1254-61.

Kröber A, Seiler T, Benner A, Bullinger L, Brückle E, Lichter P, Döhner H, Stilgenbauer S. (2002). V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. *Blood* , 100(4), 1410–16.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J,...Chen YJ (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860–921.

Larsson LG, Schena M, Carlsson M, Sallstrom J, Nilsson K. (1991). Expression of the c-myc protein is down-regulated at the terminal stages during in vitro differentiation of B-type chronic lymphocytic leukemia cells. *Blood*, 77(5), 1025-32.

Longo MS, Carone DM, Green ED, O'Neill MJ, O'Neill RJ. (2009). Distinct retroelement classes define evolutionary breakpoints demarcating sites of evolutionary novelty. *BMC genomics*, 10(1), 334.

Longo PG, Laurenti L, Gobessi S, Petlickovski A, Pelosi M, Chiusolo P, Sica S, Leone G, Efremov DG. (2007). The Akt signaling pathway determines the different proliferative capacity of chronic lymphocytic leukemia B-cells from patients with progressive and stable disease. *Leukemia*, 21(1), 110–20.

Löwer R, Löwer J, Kurth R. (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc.Natl Acad Sci USA.*, 93(11), 5177–84.

Löwer R. (1999). The pathogenic potential of endogenous retrovirus facts and fantasies. *Trends Microbiol*, 7(9), 350-6.

Maloney KA, Sullivan LL, Matheny JE, Strome ED, Merrett SL, Ferris A, Sullivan BA. (2012). Functional epialleles at an endogenous human centromere. *Proc Natl Acad Sci U S A.*, 109(34), 13704–09.

Medstrand P, Mager DL. (1998). Human specific integrations of the HERV-K endogenous retrovirus family. *J Virol.*, 72(12), 9782-7.

Medstrand P, Lindeskog M, Blomberg J. (1992). Expression of human endogenous retroviral sequences in peripheral blood mononuclear cells of healthy individuals. *J Gen Virol.*, 73(9):2463-6.

Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC Jr, McCoy JM. (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*, 403(6771), 785-9.

Mount SM. (1982). A catalogue of splice junction sequences. *Nucleic acids Res.*, 10(2), 459-72.

Moyes DL, Martin A, Sawcer S, Temperton N, Worthington J, Griffiths DJ, Venables PJ. (2005). The distribution of the endogenous retroviruses HERV-K113 and HERV-K115 in health and disease. *Genomics*, 86(3), 337–41.

Murphy WJ, Larkin DM, Everts-van der Wind A, Bourque G, Tesler G, Auvil L, ... Lewin HA. (2005). Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. *Science*, 309(5734), 613-17.

Ono M, Kawakami M, Ushikubo H. (1987). Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *J Virol.*, 61(6), 2059-62.

Oppezio P, Vasconcelos Y, Settegrana C, Jeannel D, Vuillier F, Legarff-Tavernier M, Kimura EY, Bechet S, Dumas G, Brissard M, Merle-Béral H, Yamamoto M, Dighiero G, Davi F; French Cooperative Group on CLL. (2005). The LPL/ADAM29 expression ratio is a novel prognosis indicator in chronic lymphocytic leukemia. *Blood.*, 106(2), 650-7.

Oppezio P, Vuillier F, Vasconcelos Y, Dumas G, Magnac C, Payelle-Brogard B, Pritsch O, Dighiero G. (2003). Chronic lymphocytic leukemia B cells expressing AID display a dissociation between class switch recombination and somatic hypermutation. *Blood*, 101(10), 4029-32.

Oppezso P, Magnac C, Bianchi S, Vuillier F, Tiscornia A, Dumas G, Payelle-Brogard B, Ajchenbaum-Cymbalista F, Dighiero G, Pritsch O. (2002). Do CLL B cells correspond to naive or memory B-lymphocytes? Evidence for an active Ig switch unrelated to phenotype expression and Ig mutational pattern in B-CLL cells. *Leukemia*, 16(12), 2438-46.

Palacios F, Moreno P, Morande P, Abreu C, Correa A, Porro V, Landoni AI, Gabus R, Giordano M, Dighiero G, Pritsch O, Oppezso P. (2010). High expression of AID and active class switch recombination might account for a more aggressive disease in unmutated CLL: link with an activated microenvironment in CLL disease. *Blood*, 115 (22), 4488-96.

Patel JH, Loboda AP, Showe MK, Showe LC, McMahon SB. (2004). Analysis of genomic targets reveals complex functions of MYC. *Nat Rev Cancer.*, 4(7), 562-8.

Pepper C, Ward R, Lin TT, Brennan P, Starczynski J, Musson M, Rowntree C, Bentley P, Mills K, Pratt G, Fegan C. (2007). Highly purified CD38+ and CD38- sub-clones derived from the same chronic lymphocytic leukemia patient have distinct gene expression signatures despite their monoclonal origin. *Leukemia*, 21(4), 687-96.

Pérez-Durán P, de Yebenes VG, Ramiro AR. (2007). Oncogenic events triggered by AID, the adverse effect of antibody diversification. *Carcinogenesis*, 28(12), 2427-33.

Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. (1975). Clinical staging of chronic lymphocytic leukemia. *Blood*, 46(2), 219-34.

Rassenti LZ, Huynh L, Toy TL, Chen L, Keating MJ, Gribben JG, Neuberg DS, Flinn IW, Rai KR, Byrd JC, Kay NE, Greaves A, Weiss A, Kipps TJ. (2004). ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *N Engl J Med.*, 351(9), 893-901.

Rego N, Bianchi S, Moreno P, Persson H, Kvist A, Pena A, Oppezso P, Naya H, Rovira C, Dighiero G, Pritsch O. (2012). Search for an aetiological virus candidate in chronic lymphocytic leukaemia by extensive transcriptome analysis. *Br J Haematol.*, 157(6), 709-17

Renaudineau Y, Vallet S, Le Dantec C, Hillion S, Saraux A, Youinou P. (2005). Characterization of the human CD5 endogenous retrovirus-E in B lymphocytes. *Genes and Immunity*, 6, 663–71

- Romanish MT, Cohen CJ, Mager DL. (2010). Potential mechanisms of endogenous retroviral mediated genomic instability in human cancer. *Semin Cancer Biol.*, 20(4), 246-53.
- Rudd MK, Willard HF. (2004). Analysis of the centromeric regions of the human genome assembly. *Trends Genet.*, 20(11), 529–33.
- Sakakura C, Takemura M, Hagiwara A, Shimomura K, Miyagawa K, ... Yamagishi H. (2004). Overexpression of dopa decarboxylase in peritoneal dissemination of gastric cancer and its potential as a novel marker for the detection of peritoneal micrometastases with real-time RT-PCR. *Br J Cancer.*, 90(3), 665-71.
- Schmittgen TD, Livak KJ.(2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc.*, 3(6):1101-8.
- Schneider PM, Witzel-Schlömp K, Rittner C, Zhang L. (2001). The endogenous retroviral insertion in the human complement C4 gene modulates the expression of homologous genes by antisense inhibition. *Immunogenetics*, 53(1), 1-9.
- Schroeder HW, Dighiero G. (1994). The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunol Today.*, 15(6), 288-94.
- Sentís C. (2002). Retrovirus endógenos humanos: significado biológico e implicaciones evolutivas. *Arbor*, CLXXII , 677, 135-66.
- Serafino A, Balestrieri E, Pierimarchi P, Matteucci C, Moroni G, Oricchio E, ... Vallebona PS. (2009). The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. *Exp Cell Res.*,315(5), 849-62.
- Siegel R, Ma J, Zou Z, Jemal A. (2014). Cancer statistics. *CA Cancer J Clin.*, 64(1), 9–29.
- Singh S, Kaye S, Gore ME, McClure MO, Bunker CB. (2009). The role of endogenous retroviruses in melanoma. *Br J Dermatol.*, 161(6) 1225-31.
- Stengel S, Fiebig U, Kurth R, Denner J. (2010). Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. *Gene Chromosome Cancer.*, 49(5), 401-11.

Subramanian RP, Wildschutte JH, Russo C, Coffin JM. (2011). Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology*, 8, 90.

Sutkowski N, Chen G, Calderon G, Huber BT. (2004). Epstein-Barr virus latent membrane protein LMP-2A is sufficient for transactivation of the human endogenous retrovirus HERV-K18 superantigen. *J Virol.*, 78(14), 7852-60.

Sutkowski N, Conrad B, Thorley-Lawson DA, Huber BT. (2001). Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. *Immunity*, 15(4), 579-89.

Tanaka K, Kamada N. (1998). Segmental jumping translocation in leukemia and lymphoma with a highly complex karyotype. *Leukemia & lymphoma*, 29(5-6), 563-75.

Vasconcelos Y, Davi F, Levy V, Oppezio P, Magnac C, Michel A, Yamamoto M, Pritsch O, Merle-Béral H, Maloum K, Ajchenbaum-Cymbalista F, Dighiero G (2003). Binet's Staging System and VH Genes Are Independent but Complementary Prognostic Indicators in Chronic Lymphocytic Leukemia. *J Clin Onco.*, 21(21), 3928-32.

Vogetseder W, Feng J, Dierich MP. (1995). Reactivity of monoclonal antibodies established against a recombinant human endogenous retrovirus-K (HERV-K) envelope protein. *Immunol Lett.*, 46(1-2), 129–34.

Vuillier F, Dumas G, Magnac C, Prevost MC, Lalanne AI, Oppezio P, Melanitou E, Dighiero G, Payelle-Brogard B. (2005). Lower levels of surface B-cell-receptor expression in chronic lymphocytic leukemia are associated with glycosylation and folding defects of the mu and CD79a chains. *Blood*, 105(7), 2933-40.

Wang-Johanning F, Liu J, Rycaj K, Huang M, Tsai K, Rosen DG, Chen DT, Lu DW, Barnhart KF, Johanning GL. (2007). Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. *Int J cancer.*, 120(1), 81-90.

Wang-Johanning F, Frost AR, Jian B, Epp L, Lu DW, Johanning GL. (2003). Quantitation of HERV-K env gene expression and splicing in human breast cancer. *Oncogene*, 22(10), 1528-35.

Warburton PE, Hasson D, Guillem F, Lescale C, Jin X, Abrusan G. (2008). Analysis of the largest tandemly repeated DNA families in the human genome. *BMC Genomics*, 9, 533.

Zenz T, Mertens D, Küppers R, Döhner H, Stilgenbauer S. (2010). From pathogenesis to treatment of chronic lymphocytic leukaemia. *Nat Rev Cancer.*, 10(1), 37-50.

Zhao J, Rycaj K, Geng S, Li M, Plummer JB, Yin B, ... Wang-Johanning F. (2011). Expression of human endogenous retrovirus type K envelope protein is a novel candidate prognostic marker for human breast cancer. *Genes & cancer*, 2(9), 914-22.

9. APÉNDICE

Publicación arbitrada

Fischer S, Echeverría N, Moratorio G, Landoni AI, Dighiero G, Cristina J, Opezzo P, Moreno P. (2014). Human endogenous retrovirus np9 gene is over expressed in chronic lymphocytic leukemia patients. *Leuk Res Rep.*, 3(2):70-2.