

Universidad ORT Uruguay

Facultad de Ingeniería

**PRIMERA APROXIMACIÓN METODOLÓGICA DE
SECUENCIACIÓN MASIVA MEDIANTE PANEL DE GENES
EN PACIENTES URUGUAYOS PORTADORES DE LEUCEMIA
AGUDA MIELOBLÁSTICA**

Entregado como requisito para la obtención del título de Licenciado en Biotecnología

Victoria Elizondo – 196668

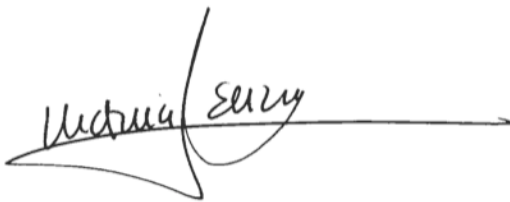
Tutor: María Noel Zubillaga

2020

Declaración de autoría

Mi persona, Victoria Elizondo, declaro que el trabajo que se presenta en esta obra es de mi propia mano. Puedo asegurar que:

- La obra fue producida en su totalidad mientras realizaba la tesis de grado en la Licenciatura en Biotecnología.
- Cuando he consultado el trabajo publicado por otros, lo he atribuido con claridad;
- Cuando he citado obras de otros, he indicado las fuentes. Con excepción de estas citas, la obra es enteramente propia;
- En la obra, he acusado recibo de las ayudas recibidas;
- Cuando la obra se basa en trabajo realizado conjuntamente con otros, he explicado claramente qué fue contribuido por otros, y qué fue contribuido por mi persona;
- Ninguna parte de este trabajo ha sido publicada previamente a su entrega, excepto donde se han realizado las aclaraciones correspondientes.

A handwritten signature in black ink, appearing to read 'Victoria Elizondo', with a long horizontal line extending to the right from the end of the signature.

Victoria Elizondo
Montevideo, 19 de Octubre de 2020

Resumen

Las leucemias agudas mieloblásticas (LAM) constituyen un grupo de enfermedades malignas que ocurren predominantemente en la edad adulta y comparten una etiopatogenia común evidenciada por alteraciones genéticas y epigenéticas. Estas son relevantes para el diagnóstico, pronóstico (estratificación de pacientes en grupos de riesgo), implementación terapéutica (identificación de blancos moleculares) y clasificación en distintas entidades clínico-patológicas.

En Uruguay la caracterización/estratificación de LAM se realiza mediante la citomorfología, inmunofenotipo, citogenética y biología molecular (mutaciones en genes FLT3, NPM1, c-KIT y CEBPA). Todo ello aplicando un algoritmo diagnóstico basado en los grupos de referencia internacional y recomendaciones de la OMS. Ello permite la correcta estratificación de nuestros pacientes en grupos de riesgo y la racionalización terapéutica.

Recientemente las técnicas de secuenciación masiva (NGS), han detectado nuevas alteraciones genéticas y epigenéticas en LAM (genes: IDH1, IDH2, DNMT3A, WT1, TET2, MLL, ASXL1, CBL, NRAS, KRAS, Tp53, EZH2; etc.) siendo algunas blancos de nuevas terapias con inhibidores de tirosinquinasa y agentes hipometilantes. La utilización racional de estas drogas y la evaluación de la respuesta a ellas requiere un diagnóstico y seguimiento preciso basados en técnicas moleculares de última generación.

El objetivo de esta propuesta es profundizar en el abordaje diagnóstico molecular de LAM, estratificar a los pacientes en grupos pronósticos y optimizar los tratamientos mediante la incorporación de marcadores moleculares emergentes a los ya existentes mediante la incorporación de NGS aplicada al estudio de neoplasias hematológicas de gran impacto en Salud. Los estudios estarán disponibles para los pacientes del país que lo requieran. Estos marcadores tienen la potencialidad de usarse en otras patologías oncológicas. Se debe destacar el carácter innovador, democratizador, y su relevante impacto público y social, tanto en el ámbito local como regional; de este trabajo de grado.

Palabras Clave

“Leucemia, LAM, secuenciación masiva, mutaciones, marcadores moleculares, diagnóstico, pronóstico, estratificación, tratamiento”.

Índice

Tabla de contenido	
Declaración de autoría:	2
Resumen:	3
Palabras Clave:	4
Antecedentes:	6
I. Asociación entre marcadores moleculares y citogenéticos.	6
II. Antecedentes en Uruguay	7
III. Nuevas alteraciones genéticas.	8
Objetivo general:	10
Objetivos Específicos:	10
Metodología:	11
I. Población de estudio	11
II. Detección de mutaciones somáticas: secuenciación dirigida	11
a. Análisis de calidad de ADN genómico	12
b. Fragmentado del AND genómico.	12
c. Preparación de las librerías por enriquecimiento selectivo.	12
d. Análisis de calidad de la librería.	13
e. Secuenciación y análisis de datos.	13
III. Análisis de mutaciones en la línea germinal:	13
Resultados:	14
I- Extracción de ADN genómico y control de calidad del mismo.	14
II- Fragmentado del AND genómico y preparación de las librerías	14
III- Secuenciación de pacientes portadores de LAM.	15
IV- Análisis de secuencias.	15
Discusión:	18
Conclusiones:	20
Bibliografía:	21

Antecedentes

La leucemia aguda mieloblástica es un desorden clonal de las células progenitoras hematopoyéticas que pierden la capacidad de diferenciarse en células sanguíneas maduras y que tienen una proliferación descontrolada que escapa a los mecanismos de regulación normales. Desde el punto de vista genético este grupo está caracterizado por alteraciones cromosómicas recurrentes con significado pronóstico en el 55% de los casos, presentando el 45% restante un cariotipo normal. Las translocaciones genéticas balanceadas más comunes son: $t(8;21)(q22;q22)$, $inv(16)(p13.1;q22)$ o $t(16;16)(p13.1;q22)$, $t(15;17)(q22;q12)$ y $t(9;11)(p22;q23)$ (1). Cada uno de estos reordenamientos estructurales da lugar a un gen de fusión que codifica para una proteína quimérica que es requerida pero no suficiente para la leucemización. En el 45% de pacientes con cariotipo normal es donde se han focalizado en los últimos años las investigaciones para la detección de marcadores moleculares que permitan una clasificación más precisa (2).

Esta enfermedad es más frecuente en adultos mayores de 60 años que en niños y adultos jóvenes, adquiriendo distintos grados de agresividad en los diferentes grupos etarios (los adultos mayores en general son los de peor pronóstico). Además, es una patología heterogénea en cuanto a las manifestaciones clínicas y las alteraciones genéticas causalmente implicadas en la patogénesis. La respuesta al tratamiento y la sobrevida global de los pacientes portadores de LAM difiere según factores pronósticos relacionados al paciente y a las características del tumor (2). Dentro de estos factores, la edad del paciente al diagnóstico, el estado físico y el cariotipo son los utilizados hasta el momento en nuestro medio para definir el pronóstico y la respuesta al tratamiento.

I. Asociación entre marcadores moleculares y citogenéticos.

Existe una asociación entre la citogenética de los pacientes y la incidencia de los marcadores moleculares. La presencia de éstos últimos modifica el pronóstico dado por la citogenética y por lo tanto la elección del tratamiento se basa en la integración de ambos resultados. Tal es el caso de los pacientes con cariotipo normal o trisomía 8, incluidos dentro del grupo de pronóstico intermedio. Cuando a estas alteraciones se le asocian las mutaciones en los genes NPM1 y CEBPA, este grupo pasa a ser de buen pronóstico. Si, en cambio, ese mismo cariotipo se asocia con las mutaciones en MLL-PTD y FLT3-ITD, pasa a formar parte del grupo de pronóstico adverso. Por otro lado, las leucemias Core Binding Factor (CBF), que involucran a los genes CBF alfa presente en el cromosoma ocho en la $t(8;21)$ y CBF beta involucrado tanto en la $inv(16)$ como en la $t(16;16)$, de pronóstico favorable presentan una alta asociación con mutaciones en el gen KIT, lo cual cambia su pronóstico a desfavorable (3).

II. Antecedentes en Uruguay

Las LAM son desórdenes heterogéneos en cuanto a las manifestaciones clínicas y las alteraciones genéticas causalmente implicadas en la patogénesis. La respuesta al tratamiento y la supervivencia global de los pacientes portadores de LAM difiere según factores pronósticos relacionados al paciente y a las características del tumor. Dentro de estos factores, la edad del paciente al diagnóstico, el cariotipo y algunas mutaciones que se describen más adelante son los utilizados hasta el momento en nuestro medio para definir el pronóstico y la respuesta al tratamiento.

La incidencia de las LAM en Uruguay es de 2.9 casos cada 100.000 habitantes por año, lo que representa 96 nuevos casos al año (4). Al presente el diagnóstico y estudio de factores pronóstico es integrado entre la clínica y paraclínica (hemograma, citomorfología, citometría de flujo y citogenética).

En las últimas décadas, los avances científicos en el área de la medicina han evolucionado notoriamente. El logro de una mayor comprensión de las bases genético-moleculares del cáncer, se ha trasladado a las áreas clínicas y de laboratorio aplicado, lo que ha conformado el despliegue de la medicina traslacional. Se han identificado alteraciones somáticas genéticas y epigenéticas específicas asociadas a diversos procesos neoplásicos, siendo éstas un elemento trascendente en el diagnóstico, pronóstico y seguimiento terapéutico. Estos marcadores genéticos se han constituido hoy en blancos moleculares terapéuticos. Tal es el caso de las neoplasias mieloides incluyendo neoplasias mieloproliferativas (MPN), leucemias Agudas Mieloblásticas (LAM) y Síndromes Mielodisplásicos (SMD) (5).

La LAM constituye un grupo de enfermedades malignas que ocurren en todos los grupos etarios con predominancia en la edad adulta (> 65-70 años). La incidencia de las LAM es de 2,9 casos/100.000 h/ año aumentando en mayores de 65 años entre 17-18 casos /100000 h) en tanto en los SMD la incidencia es de 3-5 casos/100000 h con un aumento de más de 20 casos/100 000 h en pacientes por encima de los 70 años. Un porcentaje importante de SMD se transforman en LAM y comparten una etiopatogenia común evidenciada por las alteraciones genéticas y moleculares (6).

Desde el punto de vista molecular están caracterizados por múltiples alteraciones genéticas que alteran los mecanismos normales de autorenovación, proliferación, diferenciación y apoptosis. Esta heterogeneidad se ve reflejada en diferentes subtipos en base a la existencia de diferentes alteraciones genéticas adquiridas por el clon leucémico. Estas se constituyen en marcadores genéticos con valor diagnóstico y/o pronóstico. El 55% de los portadores de LAM presentan alteraciones cromosómicas recurrentes, en tanto el 45% restante presenta un cariotipo normal (ausencia de alteraciones numéricas y estructurales) (7).

Hasta el año 2008 la clasificación de las LAM se realizaba en base a los hallazgos citogenéticos, en distintos subtipos y en grupos de riesgo pronóstico: alto, intermedio y bajo o pronóstico favorable. Los pacientes con translocaciones genéticas balanceadas que involucran a los Core Binding Factors: t(8;21)(q22;q22), inv(16)(p13.1;q22) o t(16;16)(p13.1;q22), t(15;17)(q22;q12) fueron asignados al grupo de riesgo favorable. Los pacientes con alteraciones en los cromosomas 5 o 7, o aquellos que presentan alteraciones citogenéticas complejas (tres o más alteraciones cromosómicas independientes) muestran una respuesta pobre al tratamiento con quimioterapia estándar por lo que se consideraron dentro del grupo de alto riesgo. En medio de estos dos grupos se encuentra el 45% de los pacientes que no evidencian alteraciones cromosómicas

(cariotipo normal) y constituyen el grupo de riesgo intermedio. Ha sido principalmente en este último grupo donde se han focalizado las investigaciones genético-moleculares en la búsqueda de marcadores que permitan realizar una caracterización más precisa de la enfermedad (7)-(9) y donde los resultados han sido impactantes. Se han detectado mutaciones frecuentes en los genes que codifican para receptores de tirosininasas: FLT3 y C-kit (10), en genes involucrados en el ciclo celular como NPM1 (11), así como en los genes CEBPa (12) y RUNX1 que codifican para factores de transcripción críticos para la diferenciación mieloide (9).

Algunas de estas lesiones moleculares son ahora reconocidas como poderosos marcadores diagnósticos de la enfermedad y factores pronósticos y predictivos que se utilizan prospectivamente para estratificar a los pacientes en distintos grupos de riesgo y diferentes tratamientos. Existen distintos grados de intensidad de tratamientos, que van desde terapias paliativas, quimioterapias intensas y tratamientos blanco- específicos dirigidos a las alteraciones genéticas específicas hasta el trasplante alogénico de médula ósea

III. Nuevas alteraciones genéticas.

El desarrollo de tecnologías genómicas como los análisis de expresión génica global, secuenciación masiva (NGS) y análisis de SNPs y su reciente aplicación al estudio de las LAM, ha facilitado el descubrimiento de nuevas alteraciones genéticas y epigenéticas. Estos biomarcadores emergentes incluyen mutaciones en los genes: IDH1, IDH2, DNMT3A, WT1, TET2, MLL, ASXL1, CBL, NRAS, KRAS, Tp53; EZH2, entre otros (5) (13).

Surge entonces en nuestro país, la necesidad de actualizar el algoritmo de estudio de LAM previamente generado y puesto a punto, con un nuevo aporte innovador que proporcione información pronóstica mediante el estudio de un panel de genes. Ello se realizará a través de la incorporación de tecnología de segunda generación (NGS), inexistente en el área de la salud uruguaya y con la potencialidad de ser aplicada al estudio de otras patologías tanto oncológicas como infecciosas.

Los antecedentes descritos muestran la necesidad de una constante actualización científica y tecnológica que nos permita acompañar los adelantos científicos de los grupos de referencia internacional. Dado la enorme heterogeneidad de las LAM y SMD se requiere el estudio de un mayor número de pacientes e identificar los nuevos biomarcadores que permitan definir subgrupos de pacientes que se benefician de un manejo terapéutico específico.

Resaltamos que los únicos datos existentes al respecto surgen de las investigaciones de los grupos de referencia internacionales (EEUU, Europa, etc.), no existiendo al presente datos Latinoamericanos. En este contexto se requiere la incorporación de nuevas tecnologías genómicas como NGS que permitan la detección de nuevos marcadores.

El laboratorio de Biología Molecular de la Asociación Española Primera en Salud (AEPS) realizará la recepción de muestras, de datos clínicos y demográficos, y el correspondiente procesado de la muestra y el estudio del status mutacional de un grupo de genes seleccionados de acuerdo a las recomendaciones y pautas internacionales. Esta trabajo estará apoyado por la guía académica y técnica del Memorial Sloan- Kettering Cancer Center, centro de referencia internacional en el estudio de Neoplasias Mieloides, y el Servicio de

Secuenciación de la Universidad de Nueva York con quienes ya existe un relacionamiento científico previo.

Objetivo general

El objetivo de la presente propuesta es profundizar en el abordaje diagnóstico genético-molecular, en la estratificación y estudio, y en la implementación del tratamiento de las neoplasias hematológicas: LAM. Ello se implementará mediante la incorporación de tecnología inexistente hasta el momento en el Uruguay en el área diagnóstica (NGS).

Objetivos Específicos

- 1- Evaluar la presencia y distribución de mutaciones somáticas en 28 genes asociados a cáncer en pacientes con LAM de Uruguay.
- 2- Describir la frecuencia de mutaciones nuevas y/o ya reportadas en pacientes uruguayos con LAM, aportando datos epidemiológicos propios de nuestra población.
- 3- Formación de recursos humanos en técnicas moleculares de última generación New Generation Sequencing-(NGS): diseño experimental, procesamiento de muestras y análisis de datos.
- 4- Familiarización con herramientas bioinformáticas para el análisis de datos de la secuenciación (bash, illumina software, CLC Genomics workbench).
- 5- Poner a disposición de los centros de salud del país y la región técnicas genético-moleculares que permitan mejorar los algoritmos diagnósticos existentes y la elección terapéutica.

Metodología

I. Población de estudio y obtención de muestras

El ADN extraído de muestras de médula ósea o sangre periférica con mayor 20% blastos, y raspado bucal de 50 pacientes adultos con LAM de la AEPS y otras instituciones médicas del país conforman la población de estudio de este trabajo. La muestra será tomada el momento del debut de la enfermedad y el ADN genómico será extraído mediante kits comerciales (QIASymphony DNA Midi Kit).

Mediante un consentimiento informado firmado por los pacientes y controles incluidos en el proyecto se recabará información sobre datos clínicos y epidemiológicos, y se utilizará parte de la muestra de sangre periférica o médula ósea extraída con el fin de realización de estudios paraclínicos. Todos los datos y muestras serán tratados en forma totalmente anónima únicamente identificados por un código que no identifica a la persona participante. Todas las personas participantes serán informadas de los alcances de los estudios y autorizarán mediante la firma de un consentimiento informado, la utilización de su muestra y datos para estudios asociados a la temática del proyecto.

II. Detección de mutaciones somáticas: secuenciación dirigida

La NGS dirigida específicamente a un grupo de genes propone un nuevo paradigma en el diagnóstico de LAM, en donde la información pronóstica obtenida actualmente por una variedad de métodos dispares se pueden adquirir en una única plataforma con una mayor eficiencia y aumentado notablemente la sensibilidad y escalabilidad (13). Por otra parte si bien las diferentes plataformas de NGS permiten secuenciar grandes cantidades de ADN con una alta sensibilidad y especificidad, la profundidad de lectura que es posible obtener cuando se estudia un grupo reducido de genes es muy superior a la necesaria para confirmar la presencia de diferentes alteraciones (mutaciones, deleciones, inserciones etc.). Por este motivo se han desarrollado diversas metodologías que permiten mezclar varias muestras (librerías) en una sola corrida y aprovechar así al máximo la capacidad del equipo, mediante el agregado a cada amplicón de código de barras nucleotídico (Multiplex Identifier (MID)) el cual es una secuencia de unos 10 nucleótidos específica de cada muestra (o paciente) que la identifica. El agregado del MID permite bajar costos y aumentar el rendimiento de la corrida; lo que lo hace una herramienta sumamente útil para el diagnóstico de enfermedades oncológicas(14) (15).

Se han descrito diferentes métodos para la preparación de librerías mediante enriquecimiento selectivo de regiones específicas: los métodos más utilizados caen en dos categorías: PCR-amplicón y captura por hibridación.

Como los enfoques basados en técnicas de PCR ya se utilizan de forma rutinaria en los laboratorios de diagnóstico clínico, se ajustan bien con el diagnóstico existente y los flujos de trabajo. El PCR es altamente específico y tiene la ventaja de generar una cobertura uniforme, siempre que las concentraciones de los productos de PCR individuales se normalizaron adecuadamente antes de la secuenciación. Además esta aproximación es mas flexible, permitiendo variar el número de genes que se desea analizar en cada

situación y para cada paciente. Actualmente se están utilizando al menos tres enfoques de preparación de bibliotecas en base de amplicón: PCR múltiplex, Single-Plex y captura específica seguida de PCR múltiplex (15).

En este trabajo se realiza la secuenciación para un panel de 28 genes asociados a LAM en 50 pacientes. (FLT3, WT1, DNMT3A, ASXL1, EZH2, MLL, RUNX1, TET2, IDH1, IDH2, NPN1, CEBPa, c-KIT, TP53, PTEN, PHF6, RAS, entre otros, Anexo I) mediante la plataforma Illumina, en el Laboratorio de Genómica de la Universidad de New York (NYU) a cargo de la Dra. Adriana Heguy. En cada gen se analizan las regiones en donde se han reportado la mayor frecuencia de mutaciones sitios "hot spot", según se resume en el Anexo I.

Tabla 1: Ejemplo de mutaciones recurrentes en pacientes con LAM y clasificación de las mismas. Cortes et al. ASH 2016.

Functional class	Specific example mutations
Signaling and kinase pathway	<i>FLT3, KRAS, NRAS, KIT, PTPN11, and NF1</i>
Epigenetic modifiers (DNA methylation and chromatin modification)	<i>DNMT3A, IDH1, IDH2, TET2, ASXL1, EZH2, and MLL/KMT2A</i>
Nucleophosmin	<i>NPM1</i>
Transcription factors	<i>CEBPA, RUNX1, and GATA2</i>
Tumor suppressors	<i>TP53</i>
Spliceosome complex	<i>SRSF2, U2AF1, SF3B1, and ZRSR2</i>
Cohesin complex*	<i>RAD21, STAG1, STAG2, SMC1A, and SMC3</i>

En detalle se describe las distintas etapas a continuación

a. Análisis de calidad de ADN genómico

Los ADN extraídos se cuantifican mediante método de fluorescencia con equipamiento Qubit (Life Technologies). Para determinar la calidad y concentración de los mismos (que no esté degradado), los ADN son corridos en el equipo TapeStation 4200 de Agilent(16).

b. Fragmentado del ADN genómico.

El fragmentado de los ADN genómicos fue realizado en forma física (mecánica) utilizando el equipo de ultrasonido Covaris E220 Focused-ultrasonicator.

c. Preparación de las librerías por enriquecimiento selectivo.

Cuando se piensa en la preparación de una librería para el uso en un laboratorio de diagnóstico con aplicación clínica, es necesario considerar varios aspectos: 1) Sensibilidad, 2) Especificidad, 3) La uniformidad de la cobertura de secuencia entre las regiones seleccionadas, 4) Cantidad de ADN requerido para cada muestra; 5) costo por muestra, y 6) la velocidad, el tiempo necesario para obtener la librerías listas para la secuenciación.

Para la preparación de la librería se utilizará el kit comercial KAPA HyperPlus Library Preparation Kit de KAPA BIOSYSTEMS, según indicaciones del proveedor el cual cumple con los requisitos descriptos. Luego se procede a la captura de la misma mediante Hybridization capture of DNA libraries using xGen® Lockdown® Probes and Reagents de IDT.

Se diseñarán sondas (secuencias cortas con target conocido) para amplificar las regiones genómicas de interés (según se describe en el Anexo 1) en amplicones de 200-300 pb solapados. Para el diseño de los oligos se utiliza el software Primer3 (<http://primer3.sourceforge.net/>). Alternativamente se podrán usar primers ya diseñados y probados, que actualmente se utilizan en la NYU. A cada oligo se le adicionará una región extra en el extremo 5' para el posterior agregado de adaptadores y "barcodes".

d. Análisis de calidad de la librería.

Los productos de la librería se corren en el equipo TapeStation 4200 de Agilent para determinar su calidad y concentración. Las librerías de cada paciente son purificadas mediante un método que utiliza un campo magnético (AMPure XP Bead de Agencourt) . Para cada paciente los amplicones son mezclados en cantidades equimolares.

e. Secuenciación y análisis de datos.

Para la reacción de secuenciación se utiliza el kit MiSeq Reagent Kits v3 de Illumina en el cual se corren 14 muestras de pacientes de LAM según instrucciones del proveedor, en formato 2 x 150-bp paired-end reads de 300 ciclos. En una segunda tanda de 36 pacientes y 3 controles normales de utiliza el kit NextSeq Reagent Kits v2 para el equipo NextSeq 550 de Illumina también en formato 2 x 150-bp paired-end reads de 300 ciclos.

Para el análisis de los datos crudos se utiliza el software propio de los equipos de illumina y en pasos posteriores se prueban diferentes aproximaciones, desde programas comerciales como el Genomic Workbench de la compañía Qiagen como un pipeline (serie de comandos y scripts diseñado en colaboración con la Universidad de Nueva York (NYU) hasta encontrar el flujo de trabajo apropiado para nuestra situación; (data not shown).

• Los criterios para filtrar variantes fueron:

- 1-Conteo de 300 reads como mínimo por variante
- 2-Los cambios sinónimos se descartan
- 3-Criterios de patogenicidad en bases de datos como: ClinVar, COSMIC, Single Nucleotide Polymorphism Database (dbSNP), 1000 Genomes, etc.
- 4-Frecuencia menor a 1%
- 5-No se analizaron regiones no codificantes

El análisis de los datos de secuenciación será realizado en conjunto entre los investigadores del laboratorio de la AEPSM y de Genética de la Facultad de Medicina, con apoyo del Memorial Sloan-Kettering Cancer Center y de la Universidad de New York.

III. Análisis de mutaciones en la línea germinal:

Las mutaciones reportadas como germinales en la base de datos de SNPs del National Center for Biotechnology Information (dbSNPs) y en 100Genomes, o presentes en la muestra de raspado bucal del paciente serán excluidas del análisis posterior. La verificación de las mutaciones candidatas en la muestra de ADN bucal del paciente se realizará por secuenciación Sanger en la AEPS.

Resultados

I. Extracción de ADN genómico y control de calidad del mismo.

Se realizó extracción de ADN a partir de pellet celular de glóbulos blancos obtenido de médula ósea o sangre periférica de pacientes portadores de LAM utilizando kit comercial de Qiagen. La cuantificación, pureza y control de calidad del mismo fue testada mediante equipo TapeStation de Agilent y Qubit de Invitrogen.

II. Fragmentado del ADN genómico y preparación de las librerías

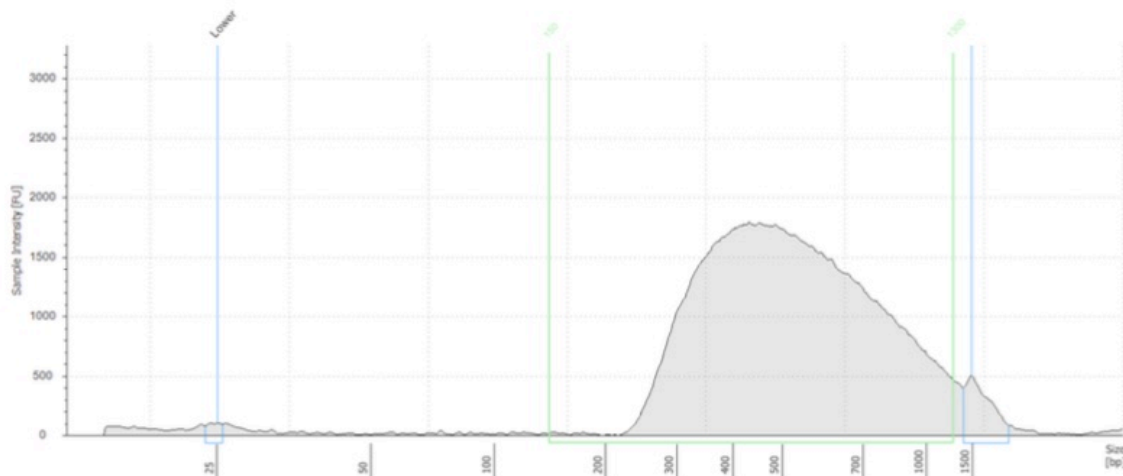
El fragmentado del ADN genómico se realizó en el equipo Covaris E220 Focused-ultrasonicator. El ciclo utilizado fue tal que se obtuvieron fragmentos en el rango de 130 pares de bases (pb) y 900 pb, siendo el promedio de tamaño los 500 pb (Figura 1).

La preparación de las librerías se realizó con el kit KAPA HyperPlus Library Preparation Kit de KAPA y los adaptadores (barcodes) utilizados fueron barcodes comerciales de Illumina. Los productos de la librería fueron purificados y corridos en el TapeStation 4200 de Agilent para comprobar la unión de los adaptadores y la distribución de los fragmentos de las mismas.

Well	Conc. [pg/ul]	Sample Description	Alert	Observations
B2	744	LAM 24		

Peak Table

Size [bp]	Calibrated Conc. [pg/ul]	Assigned Conc. [pg/ul]	Peak Molarity [pmol/l]	% Integrated Area	Peak Comment	Observations
25	28.5	-	1750	-		edited Lower Marker
455	744	-	2510	100.00		
1500	250	250	256	-		Upper Marker



Region Table

From [bp]	To [bp]	Average Size [bp]	Conc. [pg/ul]	Region Molarity [pmol/l]	% of Total	Region Comment	Color
150	1300	557	6770	21400	95.26		

Figura 1: Ejemplo de librería realizada con el kit KAPA HyperPlus Library Preparation Kit y corrida en el TapeStation 4200 luego de ser purificada. El tamaño promedio para esta librería fue de 557 pb.

III. Secuenciación de pacientes portadores de LAM.

La secuenciación masiva (NGS) de los 50 pacientes y 3 controles normales se realizó en 2 tandas en momentos diferentes.

En la primera, se realizó el secuenciado de 14 pacientes (ptes) portadores de LAM en un equipo MiSeq de illumina y la segunda se secuenciaron 36 pacientes y 3 controles normales en un equipo NextSeq 550 de Illumina.

IV. Análisis de secuencias.

El análisis de datos de los pacientes de la primera tanda se realizó mediante varios métodos: el software de

Ilumina (BaseSpace), el software Genomic Workbench de Qiagen y un pipeline de análisis de Bash (data not shown), donde se utilizan determinados comandos de terminal y se crean diferentes scripts según el análisis que se quiere realizar a este set de pacientes. El pipeline fue elaborado en conjunto con bioinformáticos de la Universidad de Nueva York. Una vez realizado el análisis primario: llamados de base (o base calling) y puntuación de calidad utilizando el software propio del equipo MiSeq, se realizó un análisis secundario que consiste en el mapeo de las numerosas lecturas (reads) de cada secuencias con el genoma humano de referencia (hg19 en nuestro caso). Para el análisis terciario es necesario filtrar las variaciones de nucleótido que no son mutaciones (SNP), para lo cual utilizaremos diferentes scripts, los cuales filtran los datos obtenidos en una serie de base de datos publicadas online y de libre acceso (como: ClinVar, COSMIC, Single Nucleotide Polymorphism Database (dbSNP), 1000 Genomes). Para la secuenciación dirigida, pueden ser usados archivos BED, específicos, para la llamada de variantes (variant calling) los cuales tienen una cantidad mucho menor de variantes comparando con un análisis donde no se utilice el BED file. Para la anotación de las variantes se utilizan bases de datos, como 1000 Genomes Project, ClinVar, dbSNP, dbVar, HGMD, Exome Variant Server, DmuDB y OncoMine (17).

La confirmación de las variantes mediante la técnica de Sanger, es un paso importante luego que se detecta una variante por NGS, ya que se han reportado falsos positivos (como por ejemplo artefactos de la técnica) asociados con determinadas plataformas o ensayos específicos (18).

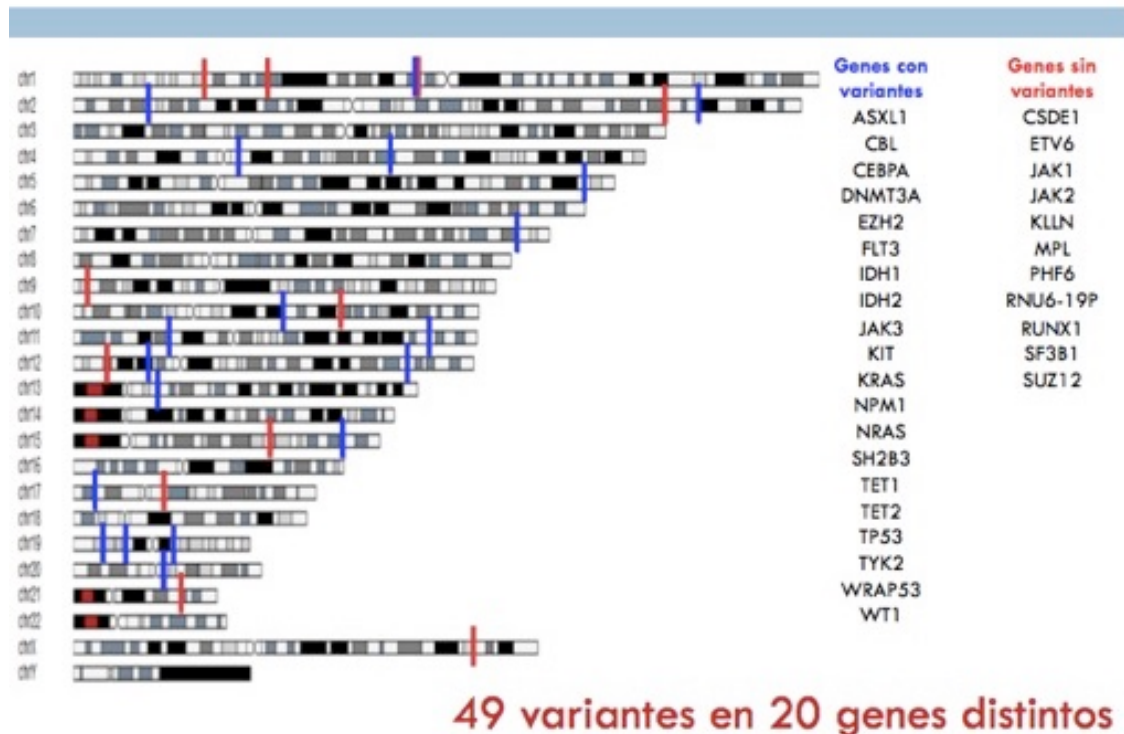


Figura 2: Distribución cromosómica de las variantes encontradas en los 28 genes estudiados en 14 pacientes portadores de LAM.

Se encontraron 49 variantes en 20 de los 28 genes estudiados en 13 de los 14 pacientes secuenciados y analizados en la primera corrida de NGS (Figura 2). De estas 49 variantes, 26 no han sido reportadas en

ninguna base de datos utilizada en el análisis (ClinVar, COSMIC, Single Nucleotide Polymorphism Database (dbSNP), 1000 Genomes, etc), 20 han sido reportadas en COSMIC y/o ClinVar, y 3 en dbSNP. Considerando solo aquellas variantes detectadas en el 45-50 % de reads del paciente y reportadas en alguna de las bases de datos anteriormente mencionadas (Figura 3), consideramos 7 de estas variantes como patogénicas confirmadas, y se encuentran en 5 de los 13 pacientes que presentaron variantes.

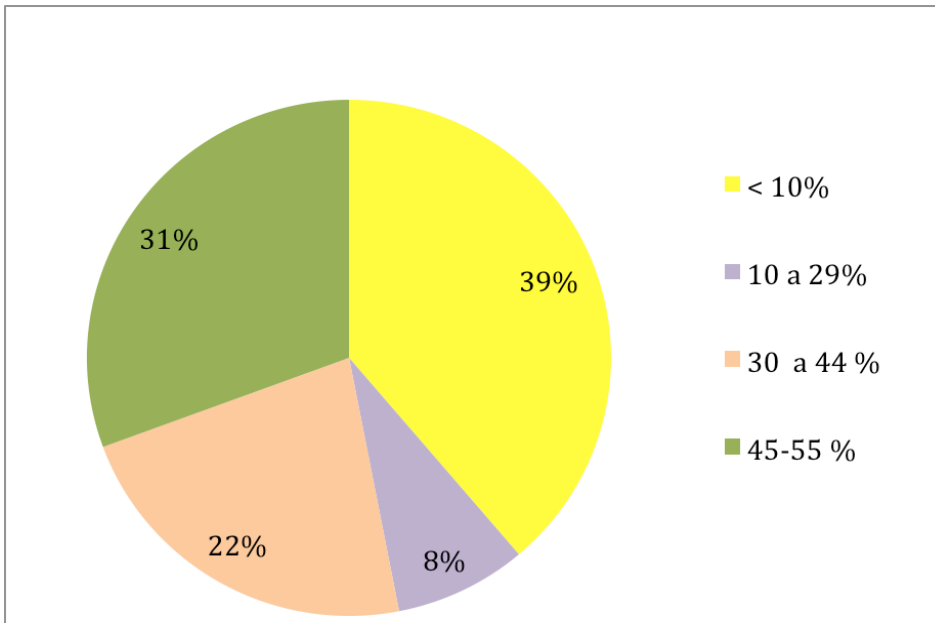


Figura 3: Clasificación de porcentaje de reads en los pacientes, de las variantes encontradas

Resta analizar las variantes no reportadas en las bases de datos y las variantes reportadas con menor porcentaje de reads (baja representación) en esta tanda de pacientes, para que sean confirmadas o descartadas. Para esto debemos tener en cuenta otros parámetros como el número de blastos al debut de los pacientes y otros aspectos clínicos y paraclínicos.

Discusión

El presente trabajo se centró en la profundización del estudio genético-molecular aplicado (marcadores diagnósticos, pronóstico y predictivos) de las LAM mediante la introducción de NGS y su potencial y rápida transferencia al estudio de otras entidades hemato-oncológicas. Específicamente nos permite visualizar el perfil genético de los genes estudiados en nuestra población analizada de pacientes Uruguayos.

De esta manera, se busca profundizar y actualizar el ya establecido algoritmo vigente (19), en la necesidad de continuar acompañando los avances científicos internacionales. En este proceso de mejora continua, la herramienta innovadora consiste en establecer una alianza estratégica entre grupos referentes de trabajo (articulación academia-prestador de salud privado) para aplicar un enfoque interdisciplinario en el abordaje de los marcadores genéticos.

Las LAM constituyen las patologías hemato-oncológicas más frecuentes, especialmente en el paciente añoso. Los únicos datos nacionales y regionales respecto a las alteraciones genético-moleculares presentes en LAM y su incidencia provienen del estudio de 106 pacientes uruguayos portadores de LAM en los que se aplicó el algoritmo diagnóstico diseñado por el grupo de trabajo de la AEPS que permite asignar a los mismos a distintos grupos de riesgo (19).

No existen, al presente, otros datos ni a nivel regional ni de países latinoamericanos respecto a la epidemiología molecular de las LAM así como tampoco del estudio de nuevos marcadores asociados mediante técnicas de NGS. Es por ello que la temática abordada en este trabajo no solo es novedosa para el país sino innovadora para la región, países latinoamericanos y en vías de desarrollo.

Tomando en cuenta que la población del Uruguay es una población envejecida, que el número de pacientes añosos se irá incrementando en los próximos años y que las patologías en que se centraliza este proyecto triplican su incidencia (de 3-5/100000h a 18-20/100000h) en mayores de 65 años, se producirá un importante aumento de pacientes portadores de LAM (6). Ello requerirá que los sistemas de Salud del país estén preparados para afrontar esta nueva situación contando con herramientas adecuadas que permitan la identificación e incorporación de marcadores moleculares emergentes al algoritmo utilizado hasta el momento. Esta innovación en el producto conllevará a una correcta definición genotípica de las LAM, a una asignación más precisa de los pacientes a los grupos de riesgo y a una correcta elección terapéutica optimizando los recursos humanos y económicos.

Lo antedicho sustenta la necesidad de la rápida implementación de los estudios moleculares ya descritos. Así mismo es imprescindible contar con la accesibilidad de los mismos para todos los pacientes del Uruguay que lo requieran tanto de Montevideo como del interior, haciendo de éste un trabajo democratizador. Con este fin, el laboratorio de la AEPS tiene desde hace varios años establecida una red de trabajo a nivel nacional que permite el acceso a los servicios del Laboratorio a todos los ciudadanos del país. Esto puede realizarse a través de la compra del servicio diagnóstico solicitado por el médico tratante a la institución proveedora (AEPS) mediante diferentes modalidades (Licitaciones Públicas, convenios entre Instituciones públicas y privadas; convenios con empresas o industria farmacéutica y solicitudes de pacientes particulares). Esta modalidad de organización y de gestión innovadora del laboratorio es producto de 20 años de trabajo donde, mediante una continua labor educativa de difusión de las aplicaciones clínicas del diagnóstico genético

hemato-oncológico, se potenció el vínculo entre el laboratorio y la comunidad médica. El presente trabajo continuará y afianzará esta modalidad de trabajo y de organización integrada que hacen viable esta propuesta.

Lo anteriormente expuesto muestra la importancia de la propuesta realizada, su carácter innovador y democratizador así como el relevante impacto público y social de la misma tanto para el mercado local como el regional.

Conclusiones

Se logró profundizar en el abordaje diagnóstico genético-molecular, en la estratificación y estudio, y en la implementación del tratamiento de las neoplasias hematológicas: LAM.

Se logró realzar con éxito la secuenciación masiva de 50 pacientes portadores de leucemia aguda mieloblástica y 3 controles normales mediante la plataforma Illumina.

Se analizaron hasta el momento y de forma exitosa la secuencias completas de 14 de los 50 pacientes portadores de LAM. Utilizando BaseSpace (Illumina) y comandos de bash.

Se encontraron 49 variantes en 20 de los 28 genes estudiados en 13 de los 14 pacientes secuenciados y analizados hasta el momento. De estas 49 variantes, 26 no han sido reportadas en ninguna base de datos utilizada en el análisis.

La propuesta de este trabajo realizado muestra su carácter innovador y democratizador así como el relevante impacto público y social de la misma tanto para el mercado local como el regional.

Referencias bibliográficas

1. C. D. Bloomfield, "Clinical importance of cytogenetics in acute myeloid leukaemia," 2001.
2. D. Grimwade and R. K. Hills, "Independent prognostic factors for AML outcome.," *ASH Educ. Progr. B.*, vol. 2009, no. 1, pp. 385–95, 2009.
3. T. Haferlach, "Molecular genetic pathways as therapeutic targets in acute myeloid leukemia.," *Hematology Am. Soc. Hematol. Educ. Program*, pp. 400–411, 2008.
4. I. Moro., "Sociedad de Hematología del Uruguay," 2010. .
5. A. Tomita, H. Kiyoi, and T. Naoe, "Mechanisms of action and resistance to all-trans retinoic acid (ATRA) and arsenic trioxide (As₂O₃) in acute promyelocytic leukemia," *Int. J. Hematol.*, 2013.
6. OMS, "OMS," in *El primer informe mundial de la OMS sobre la resistencia a los antibióticos pone de manifiesto una grave amenaza para la salud pública en todo el mundo*, 2014.
7. G. Marcucci, T. Haferlach, and H. Döhner, "Molecular genetics of adult acute myeloid leukemia: Prognostic and therapeutic implications," *J. Clin. Oncol.*, vol. 29, no. 5, pp. 475–486, 2011.
8. K. Döhner and H. Döhner, "Molecular characterization of acute myeloid leukemia," *Haematologica*, vol. 93, no. 7, pp. 2–8, 2008.
9. J. Tang *et al.*, "AML1 / RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia : prognostic implication and interaction with other gene alterations," *Blood*, 2009.
10. S. Brunet, J. F. Nomdedéu, A. Aventín, M. Hoyos, and J. Sierra, "Leucemia mieloide aguda : ¿ qué aportan los nuevos marcadores moleculares en la clasificación , el pronóstico y el tratamiento ?," no. clase I.
11. B. Falini *et al.*, "Acute myeloid leukemia with mutated nucleophosmin (NPM1): Is it a distinct entity?," *Blood*, vol. 117, no. 4, pp. 1109–1120, 2011.
12. Y. Lu, W. Chen, W. Chen, A. Stein, L. M. Weiss, and Q. Huang, "C/EBPA gene mutation and C/EBPA promoter hypermethylation in acute myeloid leukemia with normal cytogenetics," *Am. J. Hematol.*, vol. 85, no. 6, pp. 426–430, 2010.
13. E. J. Duncavage, H. J. Abel, P. Szankasi, T. W. Kelley, and J. D. Pfeifer, "Targeted next generation sequencing of clinically significant gene mutations and translocations in leukemia," *Mod. Pathol.*, vol. 25, no. 6, pp. 795–804, 2012.
14. A. Ten Bosch, "Analytical theory for permeation of condensable gases in a nanopore," *J. Memb. Sci.*, 2008.
15. M. Papageorgiou, D. R. Meldrum, and A. Broggi, "Plenary sessions [CASE 2011]," *2011 IEEE Int. Conf. Autom. Sci. Eng.*, 2011.
16. A. Kohlmann, V. Grossmann, N. Nadarajah, and T. Haferlach, "Next-generation sequencing - feasibility and practicality in haematology," *British Journal of Haematology*. 2013.
17. U. Bacher, S. Schnittger, and T. Haferlach, "Molecular genetics in acute myeloid leukemia," *Curr. Opin. Oncol.*, vol. 22, no. 6, pp. 646–655, 2010.

18. Zhang et al. Genome Biology "Analysis of error profiles in deep next-generation sequencing data", 2019.

19. U. M. Elizondo V, Perez V, Zubillaga, MN, Manrique G, Martinez L, "Algoritmo diagnóstico genético-molecular en LAM/MDS. Proyecto ANII-AESPM. Congreso Uruguayo de Hematología.," 2012.