Facultad de Ciencias Biológicas
Departamento de Bioquímica y Biología Molecular
Señalización Celular y Patologías Asociadas

TESIS DOCTORAL CON MENCIÓN INTERNACIONAL

Specific rearrangements and new molecular markers in acute myeloid leukemia: study of *MYBL2* genetic alterations and its involvement in leukemogenesis

Doctoranda

SANDRA DOLZ GIMÉNEZ

Directoras

Dra. Eva Barragán González

Dra. Paloma García Rodríguez

EVA BARRAGÁN GONZÁLEZ, doctora en Bioquímica, Biología Molecular y Biomedicina; facultativo especialista de Bioquímica Clínica en el Hospital Universitario y Politécnico La Fe,

PALOMA GARCÍA RODRÍGUEZ, doctora en Bioquímica y Ciencias Biológicas; Investigadora Senior en Cáncer, Biología Celular y Biología Molecular en el Instituto de Investigación Biomédica de la Universidad de Birmingham,

INFORMAN

que el presente trabajo de investigación titulado "**Specific rearrangements and new molecular markers in acute myeloid leukemia: study of *MYBL2* genetic alterations and its involvement in leukemogenesis**" y realizado por **SANDRA DOLZ GIMÉNEZ**, licenciada en Farmacia por la Universidad de Valencia y con título de Facultativo Especialista de Análisis Clínicos concedido por el Hospital Universitario y Politécnico La Fe, bajo nuestra supervisión en el laboratorio de Biología Molecular del Servicio de Análisis Clínicos del Hospital Universitario y Politécnico La Fe, en el marco del Instituto de Investigación Sanitaria La Fe, y en los laboratorios del Instituto de Investigación Biomédica de la Universidad de Birmingham, reúne los requisitos para su presentación y defensa en forma de Tesis Doctoral con Mención Internacional ante un tribunal.

Y para que así conste a los efectos oportunos firman la presente en Valencia, octubre 2015.

Fdo: Eva Barragán

Fdo: Paloma García

Este estudio se ha realizado con el apoyo económico de los proyectos de investigación PI06/0657 y PS09/01828, ambos financiados por el Instituto de Investigación en Salud Carlos III.

Sandra Dolz Giménez ha disfrutado del Contrato de Investigación Post-Residente 2010/0258 concedido por el Instituto de Investigación Sanitaria La Fe y la Fundación Bancaja, y del Contrato de Investigación Río Hortega CM13/00022 concedido por el Instituto de Investigación en Salud Carlos III, así como de una estancia de 3 meses en el Instituto de Investigación Biomédica de la Universidad de Birmingham (Birmingham, Reino Unido) concedida por el Instituto de Investigación Sanitaria La Fe (2010/0258).

Este trabajo se ha realizado en los laboratorios del Servicio de Análisis Clínicos y del Servicio de Hematología del Hospital Universitario y Politécnico La Fe (Valencia, España), en el marco del Instituto de Investigación Sanitaria La Fe, bajo la dirección de la Dra. Eva Barragán González, así como en los laboratorios del Instituto de Investigación Biomédica de la Universidad de Birmingham (Birmingham, Reino Unido), bajo la dirección de la Dra. Paloma García Rodríguez.

A partir de la presente tesis se han derivado las publicaciones científicas que se citan a continuación:

1. Dolz S, et al. A novel Real-Time Polymerase-Chain Reaction assay for simultaneous detection of recurrent fusion genes in acute myeloid leukemia. J Mol Diagn. 2013; 15: 678-686.

2. Dolz S, et al. Study of the S427G polymorphism and of *MYBL2* variants in patients with acute myeloid leukemia. Leuk Lymphoma 2015; 19: 1-7.

3. Dolz S, et al. Analysis of *MYBL2* downstream targets and its role in human hematopoietic stem cell development. (Manuscrito en elaboración).

A mi yaya Ángeles

A mis padres

A Javi

El mejor científico está abierto a la experiencia,

y ésta empieza con un romance, es decir,

la idea de que todo es posible.

Ray Bradbury

La duda es la madre del descubrimiento.

Ambrose Bierce

AGRADECIMIENTOS

Este trabajo es el resultado de un proceso de aprendizaje en el que, con esfuerzo y constancia, he adquirido conocimientos del método científico, técnica y disciplina. Me siento muy afortunada de poder aportar mi pequeño grano de arena a la ciencia, a esa facultad cuyo avance permite mejorar la salud y calidad de vida de las personas. Durante esta trayectoria, he tenido la suerte de conocer a gente que me ha acompañado y/ó ayudado en la misma, a la que quería dedicar unas palabras de agradecimiento.

En primer lugar, quería agradecer la dedicación y el apoyo de mis directoras de tesis, las Dras. Eva Barragán y Paloma García, ya que sin ellas este trabajo no podría haber salido adelante.

A la Dra. Eva Barragán, gracias a la oportunidad que me brindó de formarme a su lado. Ha sido todo un lujo para mí poder adquirir conocimientos junto con una persona de gran talento, capacidad de proyección científica y dedicación, que me ha estimulado a aprender a mejorar cada día, a perfeccionar mi trabajo, a crecer y a madurar en el ámbito profesional. Gracias a tus buenos consejos y a todo lo que me has enseñado.

A la Dra. Paloma García, gracias a la afectuosa acogida en su centro de investigación. Fue una gran suerte para mí poder aprender de la mano de una profesional del ámbito científico internacional, de gran talento, creativa y con gran capacidad de transmitir conocimiento. Gracias a ello, consigue “hacer fácil lo difícil” y he aprendido a superar adversidades, a comunicar y a crecer en el campo de la investigación. Gracias por estar siempre dispuesta a ayudarme, a tu ánimo y sonrisa.

En segundo lugar, quería agradecer al profesor Miguel A. Sanz, un ejemplo de superación, la confianza que depositó en mi trabajo, a su cercanía y accesibilidad (a pesar de su escaso tiempo libre), a sus enseñanzas y a la buena voluntad que demuestra en ayudar a potenciar los conocimientos de la gente que le rodea. Ha sido un honor poder dedicar con tal eminencia de la Hematología mis horas de trabajo.

De igual manera, también quería agradecer al Dr. Pascual Bolufer y a todo el equipo del laboratorio de Biología Molecular de Análisis Clínicos el apoyo que me prestaron, a sus enseñanzas, ayuda, a su colaboración y a sus buenos consejos. Gracias a ellos, he aprendido y he desarrollado mis conocimientos de Biología Molecular bajo la tutela de auténticos profesionales en la materia. A Óscar, Sarai, Inma, Marta, Rosa,

Gema, Juan Carlos, Nuria, Virginia, Esperanza, Jorge, Teresa y Enrique, gracias a las horas de trabajo dedicadas y a los buenos momentos que pasamos juntos.

A todo el equipo del laboratorio de Citogenética de Hematología, en especial a Jose, Esperanza, Inés, Irene, Mariam y María, gracias también por su apoyo, colaboración y a su trabajo. Así como a Lourdes del Servicio de Citometría, gracias por su ayuda, su dedicación, su agradable trato y amable colaboración, siempre dispuesta a facilitar y contribuir al trabajo de la mejor manera.

Mi agradecimiento a la Dra. Dolors Sánchez, coordinadora de la Unidad de Genómica, por su colaboración, accesibilidad y dedicación al trabajo, por sus enseñanzas, su ayuda, sus consejos y su entusiasmo en las investigaciones.

Gracias al equipo de investigación de la Dra. Paloma García de Birmingham, por la cálida acogida (a pesar del frío invierno) que recibí durante mi estancia allí. A Mary, Carl, Lozan, Giacomo, Emilie, Maelle, John, Stephanie y Roger, por su profesional trabajo, su disponibilidad y atención, su colaboración y ayuda, sus enseñanzas, su simpatía y por hacerme disfrutar de muy buenos momentos. Fue un lujo para mí poder trabajar junto a ellos en un centro de investigación de prestigio, que me maravilló desde el primer momento. Muchas gracias a Leyre, Alessandra y Banessa, por hacerme pasar buenos momentos, disfrutar de su compañía y hacerme sentir como en casa. Gracias a todos ellos guardo un entrañable y grato recuerdo, una experiencia que nunca olvidaré.

Gracias al equipo de trabajo del Centro de Transfusiones de Valencia, en especial a Marga, Isabel, Mª Ángeles, Amparo y a la Dra. Pilar Solves, con quien iniciamos esta colaboración. Gracias a todos ellos por su dedicación, por su amable disposición y contribución a este trabajo.

Finalmente, quería dedicar unas palabras de agradecimiento a las personas que más quiero, a toda mi familia y “apegaos”, gracias por estar ahí. En especial, a mis padres y a mi hermana, gracias por estar siempre conmigo, apoyándome en todo, por confiar en mí, por comprenderme, por animarme y por vuestro cariño. A los mejores padres que podría tener, gracias por la educación que me habéis dado y a vuestro esfuerzo, me siento muy afortunada por ello. De igual manera, me siento muy afortunada de tener a mi hermana, mi médico preferida, con quien he compartido tantas horas de estudio. Gracias Loren por estar siempre cerca de mí y ser la mejor hermana.

Guardo bonitos recuerdos juntas en casa de la yaya, con mi tío Antonio siempre estudiando y mi tía Paqui tecleando la máquina de escribir, y así seguimos después sus pasos. Gracias a esos momentos y a mi querida yaya Ángeles, a quien tanto echamos de menos, una persona ejemplar que siempre ha sido joven, despierta, inteligente, avanzada a su tiempo, tolerante, luchadora, optimista, carismática, llena de vida y ganas de vivir, y sobre todo apasionada de su familia, a la que siempre nos ha dado sabios consejos, nos ha animado a estudiar, a esforzarnos cada día y por todo ello, su memoria siempre nos mantendrá unidos. Gracias yaya por cuidarme, por tu cariño y por tu recuerdo. Gracias a mis amigos por apoyarme, en especial a quien siempre fue y es mi mejor y más que amigo, mi novio Javi. Gracias por estar siempre conmigo, por tu paciencia, por compartir los momentos difíciles y también los mejores, gracias al apoyo de tu familia y a los buenos momentos juntos, gracias por tu comprensión, por tu esfuerzo en ayudarme en todos los sentidos, por hacerme reír, por animarme, por ser mi compañero y por quererme, sin ti este trabajo no habría tenido un mejor final.

TABLE OF CONTENTS

RESUMEN

I. Introducción

La leucemia mieloide aguda (LMA) es una neoplasia hematológica de origen clonal que se caracteriza por la proliferación en la médula ósea de células progenitoras anormales de estirpe mieloide. Se considera una entidad muy heterogénea, con comportamientos clínicos, pronósticos y de respuesta a tratamiento muy diferentes. Actualmente, se reconoce que la proliferación clonal de los progenitores anormales se debe a la adquisición de ciertos cambios genéticos que alteran el equilibrio entre los mecanismos de auto-renovación y diferenciación celular, contribuyendo al desarrollo de cánceres hematológicos como la LMA.

Las primeras alteraciones genómicas en la LMA se identificaron en las células blásticas de pacientes con alteraciones recurrentes no aleatorias del cariotipo. A diferencia de lo que ocurre en los tumores sólidos, que suelen presentar cariotipos complejos con diferentes anomalías estructurales y numéricas, en las neoplasias hematológicas un porcentaje muy elevado de pacientes presenta una sola alteración cromosómica. En la mayoría de los casos se trata de alteraciones cromosómicas recurrentes que no se presentan en otras neoplasias no hematológicas, por lo que se cree que representan anomalías primarias relacionadas con el desarrollo leucemogénico inicial. La determinación de los reordenamientos cromosómicos (translocaciones, inserciones e inversiones) en el momento del diagnóstico se considera uno de los factores pronóstico más importante en la LMA. Sin embargo, la metodología de análisis del cariotipo es laboriosa y rinde suficientes metafases en tan solo el 60% a 80% de las muestras de médula ósea. Por esta razón, se han desarrollado técnicas de biología molecular más sensibles para detectar los reordenamientos moleculares que se producen como consecuencia de las alteraciones cromosómicas.

En los últimos años, la caracterización molecular de la LMA ha demostrado un creciente interés ya que contribuye a un diagnóstico refinado, a la evaluación del pronóstico y al seguimiento de la enfermedad mínima residual (EMR). Los reordenamientos moleculares implicados en LMA pueden detectarse por métodos de transcripción reversa y reacción en cadena de la polimerasa (RT–PCR) individuales. Sin embargo, en vista del gran número de genes de fusión y de las variantes de punto de corte caracterizadas actualmente, realizar el cribado molecular de un paciente mediante un procedimiento estándar de RT–PCR supone un trabajo intensivo y prácticamente no

factible. Hasta el momento, varios estudios habían descrito métodos de RT–PCR de tipo multiplex en las leucemias agudas, sin embargo, proveían una cobertura limitada de reordenamientos específicos de LMA ó estaban basados en métodos convencionales de PCR. Por otra parte, los recientes e innovadores métodos de alto rendimiento (arrays, secuenciación de última generación, etc.) apenas están desarrollados para la identificación de reordenamientos moleculares balanceados que requieren su detección en ARN, ni todavía son lo suficientemente rentables para los laboratorios de diagnóstico clínico del ámbito hospitalario. Por lo tanto, es importante optimizar métodos que puedan resultar útiles y fácilmente aplicables para el diagnóstico de los reordenamientos moleculares en los laboratorios clínicos.

Por otra parte, aproximadamente un 40-50% de las LMA al diagnóstico no presenta ninguna alteración cromosómica y el mecanismo molecular de leucemogénesis es menos claro. En este caso, recientes trabajos han demostrado la existencia de una serie de alteraciones a nivel molecular que pueden estar implicadas en el desarrollo de la leucemia. Por ejemplo, las mutaciones en el receptor tirosina-quinasa *FLT3* y las mutaciones en el gen de la nucleofosmina *NPM1* se han asociado con el pronóstico de la enfermedad, así como la sobreexpresión del gen del tumor de Wilms *WT1*. A pesar del descubrimiento de estas alteraciones moleculares, más de un 25% de LMA de cariotipo normal carece de anomalías estructurales moleculares de carácter recurrente ó de mutaciones en los genes asociados a leucemia, lo cual dificulta la aproximación del pronóstico y la elección de la terapia más adecuada. Por tanto, principalmente en este grupo de pacientes cuyo pronóstico es muy heterogéneo, es importante la identificación de nuevos marcadores moleculares para permitir establecer nuevos grupos pronóstico. Además, el estudio de nuevos marcadores moleculares de LMA no sólo reside en la oportunidad de obtener una aproximación del pronóstico previa al tratamiento, sino que también abre la puerta a la investigación de los mecanismos implicados en la leucemogénesis y al conocimiento de nuevas dianas terapéuticas que permitan individualizar el tratamiento en el futuro. A este respecto, en los últimos años se ha publicado que el gen *MYBL2* podría estar asociado con el desarrollo de leucemia. *MYBL2* es un factor de transcripción nuclear y varios trabajos demuestran que su expresión se mantiene alterada en algunos tipos de cáncer, presagiando que su desregulación podría estar implicada en la tumorigénesis y progresión de la enfermedad. Además, varios estudios sugieren que la expresión aberrante de *MYBL2* implica un

pronóstico adverso en el cáncer, así como en la LMA. En vista de estos hallazgos, algunos trabajos han contribuido al conocimiento de las alteraciones genéticas que puedan modificar la actividad de *MYBL2* y, por tanto, repercutir en el desarrollo del cáncer. En este sentido, recientes trabajos han sugerido que el polimorfismo rs2070235 (S427G) más frecuente de la región codificante de *MYBL2* induce un cambio de conformación en la proteína que inactiva parcialmente su función y se asocia con el riesgo de padecer cáncer. Sin embargo, hasta el momento ningún trabajo había estudiado este polimorfismo ni las posibles alteraciones genéticas de *MYBL2* en pacientes con LMA.

II. Hipótesis y objetivos

La detección de las alteraciones moleculares es importante para diagnosticar y estratificar la LMA, siendo también de interés para comprender el mecanismo molecular del proceso leucémico. En este sentido, la presente tesis plantea que la detección simultánea de los reordenamientos moleculares mediante un método optimizado de RT–PCR podría mejorar el diagnóstico, el pronóstico y la EMR de la LMA. Sin embargo, existe un elevado número de pacientes con LMA de cariotipo normal que carece de estas alteraciones moleculares, y por lo tanto, no puede obtener beneficio de las ventajas de estas nuevas estrategias metodológicas. Por este motivo y debido a la importante necesidad de estratificar a los pacientes con cariotipo normal y que no presentan ningún tipo de alteración molecular específica conocida, la segunda parte de esta tesis se centra en el estudio del gen *MYBL2* como un nuevo y prometedor marcador de LMA. En base a los resultados de estudios recientes y debido a las funciones asignadas a *MYBL2*, esta tesis defiende que es concebible que la desregulación de este gen podría contribuir al desarrollo y / o mantenimiento del fenotipo leucémico. Sin embargo, se desconocen las alteraciones genéticas de *MYBL2* en la LMA, así como los genes diana que podrían estar integrados en el mecanismo de la leucemogénesis. En este sentido, en la presente tesis se explora el papel del polimorfismo S427G y de las mutaciones somáticas de *MYBL2* en el riesgo de desarrollar una LMA ó de modificar su pronóstico. Finalmente, la tesis plantea el estudio del posible mecanismo de activación de la LMA debido a la expresión aberrante de *MYBL2* y como causa subyacente de la transformación de los progenitores hematopoyéticos sanos en células blásticas anormales de LMA.

Con la finalidad de esclarecer los planteamientos descritos, se diseñaron los siguientes objetivos:

1. Puesta a punto y validación de un método optimizado de RT–PCR en tiempo real (qRT–PCR) para la detección simultánea de 15 reordenamientos específicos de LMA y comparación de los resultados con los obtenidos por citogenética, hibridación fluorescente in situ (FISH) y los métodos de PCR convencionales.

2. Estudio de las alteraciones genéticas conocidas de *MYBL2* y búsqueda de nuevas mutaciones en la LMA.

2.1 Estudio del polimorfismo no sinónimo S427G más frecuente de *MYBL2* mediante el análisis de su incidencia en series de casos / controles y su asociación con el riesgo de LMA, así como la asociación entre el cambio S427G en la LMA y el pronóstico de los pacientes.

2.2 Búsqueda de nuevas variaciones genéticas de *MYBL2* en la LMA mediante el análisis de la incidencia de las variantes genéticas detectadas en los pacientes y en las bases de datos de mutaciones / polimorfismos registrados. A continuación, las variaciones genéticas de *MYBL2* detectadas se estudiarán con la finalidad de determinar si se correlacionan con el riesgo de LMA. Además, se estudiará la correlación de estas variaciones genéticas detectadas con la expresión de *MYBL2* y con el pronóstico de los pacientes.

3. Estudio de las alteraciones desencadenadas debido a la expresión aberrante de *MYBL2* en el desarrollo de las células progenitoras hematopoyéticas.

3.1 Identificación de los genes desregulados debido a la sobreexpresión ó al silenciamiento de *MYBL2* mediante arrays de expresión génica y su validación por PCR cuantitativa. Determinar si existen vías celulares específicas afectadas debido a la desregulación de *MYBL2*.

3.2 Estudios de proliferación / diferenciación de células progenitoras humanas $CD34^+$ con niveles modificados de *MYBL2* mediante ensayos de células formadoras de colonias (CFC).

III. Material y métodos

1. Selección de pacientes y controles

Los estudios desarrollados en esta tesis incluyen un total de 302 pacientes con LMA que fueron diagnosticados entre 1998 y 2009 en el *Hospital Universitario y Politécnico La Fe* y fueron remitidos al Laboratorio de Biología Molecular para su caracterización molecular. Todos los pacientes fueron clasificados de acuerdo a los criterios de la clasificación Franco-Americano-Británica y de la clasificación de la Organización Mundial de la Salud del 2008. El criterio de disponibilidad de muestras de ADN ó ARN fue limitante para la inclusión de los pacientes. El Comité de Ética de Investigación Clínica del *Hospital Universitario y Politécnico La Fe* aprobó los diferentes estudios y se obtuvo el consentimiento de todos los participantes, de acuerdo con las recomendaciones de la Declaración de los Derechos Humanos, la Conferencia de Helsinki y otras normas institucionales.

La puesta a punto del nuevo método de qRT–PCR se realizó a partir de células leucémicas de pacientes, líneas celulares y plásmidos con los reordenamientos específicos de LMA, a modo de controles positivos. Además, se utilizaron células de sangre periférica de donantes sanos como control negativo.

En el estudio caso-control del polimorfismo S427G, se seleccionó un grupo control de 179 donantes sanos caucásicos y residentes en España en base a los criterios de no haber padecido ninguna patología oncológica, hematológica ó enfermedad crónica y no mantener un consumo crónico de alcohol ó tabaco, cuyas características de edad y sexo fueran similares a los pacientes con LMA.

Los estudios funcionales in vitro de *MYBL2* se realizaron en células progenitoras hematopoyéticas $CD34^+$ obtenidas a partir de sangre de cordón umbilical humano de 64 donantes sanos.

2. Aislamiento celular

Las células leucémicas se obtuvieron a partir de la médula ósea de pacientes con LMA y los leucocitos se aislaron de sangre periférica de donantes sanos, tras lisis eritrocitaria en ambos casos. Los lisados celulares se resuspendieron en suero salino

(para extraer el ADN) ó en el tampón de lisis Magna Pure LC mRNA (para extraer el ARNm) [Roche Diagnostics GmbH, Mannheim, Alemania], y se almacenaron a –40 ° C. Las líneas celulares de LMA se cultivaron en medio RPMI-1640 con 20% de suero bovino fetal y 1% de penicilina-estreptomicina. Las células en cultivo se recogieron mediante centrifugación y se conservaron a –40 ° C en el tampón de lisis Magna Pure LC mRNA (Roche) para extraer el ARNm.

Las células progenitoras $CD34^+$ obtenidas de sangre de cordón umbilical fueron seleccionadas mediante micro-esferas magnéticas marcadas con anticuerpos $CD34^+$ en un separador automático MACS (Miltenyi Biotec), a partir de las células mononucleadas aisladas previamente con Lymphoprep (Axis-Shield, Oslo) tras depleción eritrocitaria de las muestras. Finalmente, después de evaluar la pureza de las células aisladas mediante citometría de flujo, únicamente las muestras que contenían más del 85% de células $CD34^+$ fueron criopreservadas en suero bovino fetal con 10% de sulfóxido de dimetilo y en nitrógeno líquido, hasta el momento de su transfección.

3. Extracción de ácidos nucleicos

Las extracciones rutinarias de ADN ó ARNm se realizaron mediante técnicas estandarizadas automáticas en el Sistema MagNA Pure LC (Roche), utilizando el kit MagNA Pure LC large volume DNA Isolation ó el kit MagNA Pure LC mRNA HS (Roche), respectivamente, de acuerdo con las instrucciones del fabricante.

La extracción del ARN total de las células $CD34^+$ transfectadas se realizó mediante métodos convencionales basados en la extracción con guanidinio tiocianato-fenol-cloroformo, seguido de purificación y concentración con el kit Rneasy MinElute (Qiagen), para su posterior procesamiento mediante arrays de expresión génica.

La evaluación de la concentración y pureza del ADN, ARNm y del ARN total, se realizó mediante espectrofotometría utilizando NanoDrop ND-2000 (ThermoFisher). Además, la calidad y la integridad del ARN total se evaluaron mediante la plataforma microfluídica Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara) en el Servicio de Arrays de la Unidad de Genómica del *Instituto de Investigación Sanitaria del Hospital Universitario y Politécnico La Fe (IIS La Fe)*. Inmediatamente, los ácidos nucleicos se almacenaron a –80 °C hasta el momento de su procesamiento.

4. Transcripción reversa

El ARNm se transcribió a ADNc utilizando el kit Taqman® Gold Reverse Transcription Reagents (N808-0234, PE Applied Biosystems, NJ).

El ARN total purificado, aislado a partir de las células $CD34^+$ transfectadas, se transcribió a ADNc mediante el kit QuantiTect Reverse Transcription (Qiagen), que permite la eliminación integrada de restos de ADN genómico contaminante.

5. Detección de reordenamientos moleculares mediante el ensayo de RT–PCR en tiempo real

El ensayo de qRT–PCR desarrollado para detectar reordenamientos de LMA se realizó mediante PCR en tiempo real utilizando la técnica de SYBR Green.

Los cebadores específicos se adaptaron de los métodos publicados ó se diseñaron mediante el software Primer3 (http://primer3.wi.mit.edu). A continuación, los cebadores seleccionados para la detección de cada reordenamiento se liofilizaron en los pocillos individuales de una placa de PCR de 96 (Roche).

La reacción de PCR en tiempo real se desarrolló en el equipo LightCycler 480 (Roche) con el kit SYBR Green I Master Mix SYBR y el reactivo Uracil-N-Glicosilasa (Roche).

6. Genotipado del polimorfismo S427G mediante PCR en tiempo real con sondas de hibridación

El genotipado del polimorfismo S427G de *MYBL2* se realizó mediante PCR en tiempo real utilizando sondas específicas de transferencia energética de fluorescencia por resonancia. La sonda de anclaje se diseñó para enlazar con el ADN salvaje y la sonda sensor para unirse a la secuencia mutante del ADN. El diseño de los cebadores para el flanqueo de la región del polimorfismo se realizó mediante el software Primer3 (http://primer3.wi.mit.edu).

La reacción de PCR en tiempo real se desarrolló en el equipo LightCycler 480 (Roche) con el kit LightCycler® 480 Genotyping Master Mix (Roche).

7. Detección de alteraciones genéticas en *MYBL2* mediante el método de análisis de Curvas de Fusión de Alta Resolución

La búsqueda de nuevas mutaciones en *MYBL2* se realizó en los pacientes con LMA mediante el método de análisis de Curvas de Fusión de Alta Resolución conocido como High Resolution Melting (HRM), que está basado en el uso del intercalante fluorescente ResoLight a una concentración saturante.

El análisis de la secuencia codificante y regiones intrónicas colindantes de *MYBL2* se realizó a partir del diseño de 16 parejas de cebadores específicos mediante el software Primer3 (http://primer3.wi.mit.edu), los cuales amplificaban fragmentos solapados de 150 a 250 pares de bases.

La reacción de PCR en tiempo real se desarrolló en el equipo LightCycler 480 (Roche) con el kit LightCycler® 480 HRM Master Mix (Roche).

8. Cuantificación de la expresión génica mediante PCR en tiempo real basada en sondas Taqman

La cuantificación de la expresión del ARNm de *MYBL2,* así como la validación de la expresión génica resultante de los arrays de expresión, se realizó mediante PCR cuantitativa con sondas TaqMan, cuyo diseño está enfocado a aumentar la especificidad de la PCR cuantitativa.

La reacción de PCR se realizó a partir de ADNc empleando los kits Taqman Universal Master Mix (Applied Biosystems) y Taqman Gene Expression Assays específicos de cada gen. En el caso del gen control *GUSB*, se emplearon cebadores específicos y sonda TaqMan adaptados de métodos publicados. Todos los ensayos se realizaron en el equipo ABI Prism 7500 Fast Real Time (Applied Biosystems).

El método basado en el modelo Delta-Ct se utilizó para evaluar la expresión de *MYBL2* en pacientes con LMA. El cálculo de la expresión génica relativa se realizó con el modelo Delta-Delta Ct. Todos los datos se normalizaron usando la expresión del gen control endógeno *GUSB.*

9. Secuenciación

Los patrones de HRM alterados en las muestras de LMA se confirmaron mediante secuenciación directa en el equipo ABI Prism 3130 (Applied). Las variaciones genéticas se identificaron comparando las secuencias obtenidas con la secuencia salvaje de *MYBL2* (número de acceso Gene Bank NT_011362.10).

10. Transformación celular y aislamiento de ADN plasmídico

Las células competentes (α-select Chemically Competent Cells, Bioline, London, UK) se transformaron con el plásmido *MYBL2*–IRES–EGFP (una construcción que sobreexpresa el ADNc de *MYBL2* humano bajo el control de un promotor del citomegalovirus) ó con el plásmido control IRES–EGFP (ambos plásmidos fueron cedidos amablemente por la Dr. Paloma García, del Instituto de *Investigación Biomédica de la Facultad de Medicina y Odontología de la Universidad de Birmingham, Reino Unido*). Después de someter las células a un choque térmico y posterior incubación en un medio rico en nutrientes, las células transformadas crecieron en una placa de agar de Caldo de Lisogenia (LB) que contenía el antibiótico Kanamicina. A continuación, las células transformadas se recogieron a partir de 1-2 colonias y crecieron en medio LB con Kanamicina.

El ADN plasmídico de las células transformadas se purificó mediante el kit Plasmid Purification Maxi (Qiagen). Una pequeña alícuota del ADN purificado se digirió para verificar el ADN plasmídico, utilizando la enzima de restricción HindIII con tampón Neb4, lo cual generó dos fragmentos plasmídicos de restricción específicos (de 4,770 y 269 pares de bases) que se testaron en un gel de agarosa. Finalmente, el ADN plasmídico purificado se cuantificó mediante Nanodrop ND-2000 (ThermoFisher) y se almacenó a –80 °C.

11. Transfección celular

11.1 Electroporación

Las células K562 (una línea celular humana de leucemia mieloide crónica) se transfectaron mediante electroporación con modificadores de la expresión de *MYBL2* (p*MYBL2*–IRES–EGFP ó siRNAs específicos de *MYBL2*) ó con los correspondientes

modificadores control (pIRES–EGFP control ó siRNAs control), empleando tampón Cytomix, ATP y glutatión. Después de 24 horas de la transfección, las células viables transfectadas se recogieron y se separaron en base a la expresión de la proteína verde fluorescente (GFP) de los plásmidos ó mediante la detección del marcaje de los siRNAs con fluoresceína-isotiocianato (FITC), en un citómetro-sorter Cytomation MoFlo FACS (Beckman Coulter, Brea, CA, EE.UU.), con una pureza superior al 99%.

Los siRNAs específicos de *MYBL2* y los siRNAs control (previamente publicados por García y Frampton) fueron diseñados por Eurogentec (Cultek) con una modificación FITC en 5′ y proyecciones TT en sus extremos 3′.

11.2 Nucleofección

Las células progenitoras $CD34^+$ aisladas de sangre de cordón que fueron criopreservadas, se descongelaron y se transfectaron con moduladores de la expresión de *MYBL2* ó con los correspondientes modificadores control en cubetas individuales, empleando el kit P3 Primary Cell 4D-Nucleofector X solution (Lonza) en el equipo Amaxa Nucleofector® 4D (Lonza), con programas de nucleofección específicos (programa EO-100 para transfectar siRNAs y programa ED-100 para la transfección de los plásmidos). Tras 24 horas de la transfección, las células $CD34^+$ viables y tranfectadas se separaron en base a la detección de las fracciones GFP^+ ó $FITC^+$ en un citómetro-sorter Cytomation MoFlo FACS (Beckman Coulter), con una pureza superior al 99%. Las fracciones celulares GFP^+ ó $FITC^+$ se incubaron en medio de cultivo (en caso de posterior realización de ensayos CFC) ó fueron inmediatamente lisadas y preservadas en Trizol (Invitrogen) a –20 °C, para posteriormente extraer su ARN total y realizar ensayos de expresión génica.

Cada experimento de transfecciones emparejadas con siRNAs ó con plásmidos se realizó por triplicado en días alternos, tanto para llevar a cabo ensayos de CFC como para la posterior realización de arrays de expresión génica.

12. Western blot

El extracto de proteínas obtenido a partir de la lisis de las células $CD34^+$ transfectadas e incubado con solución tampón Laemmli, se sometió a electroforesis en gel de poliacrilamida-SDS (SDS-PAGE). El tamaño de las proteínas se estimó mediante

un marcador de peso molecular pre-teñido (Color Plus Pre-stained Broad Range, Byolabs New England). En el siguiente paso, las proteínas se transfirieron a una membrana adsorbente de polímeros de polifluoruro de vinilideno (PVDF) de Biotrace (Bio-Rad) en un aparato de transferencia semi-seca (Amersham).

El bloqueo de la membrana de PVDF se realizó en solución salina tamponada con Tris, Tween 20 y leche en polvo (T-TBS). La membrana se dividió en dos mitades y cada una se incubó con el anticuerpo primario apropiado diluido en T-TBS. Los anticuerpos primarios empleados fueron los siguientes: IgG policlonal de conejo anti-MYBL2 humano (N–19, sc–724, Santa Cruz Biotechnology) e IgG policlonal de cabra anti-actina humano (I–19, sc–1616, Santa Cruz). Después de lavar e incubar con anticuerpos secundarios conjugados con peroxidasa de rábano (Amersham), las proteínas se detectaron utilizando un sistema de quimioluminiscencia potenciada (Pierce).

13. Arrays de expresión génica

El perfil de expresión de ARN de las células progenitoras $CD34^+$ transfectadas se determinó utilizando la plataforma Affymetrix Primeview, que incluye todos los genes anotados en la base del NCBI Human Genome version 37 (GRCh37). Esta versión contiene 48,658 conjuntos de sondas incluyendo 419 sondas UniGene no cubiertas por RefSeq. El ARN total aislado a partir de las células $CD34^+$ transfectadas se procesó aplicando el protocolo WTA2 en la Unidad de Genómica del *Instituto de Investigación en Biomedicina* (Barcelona).

El análisis de los datos obtenidos se realizó en el Servicio de Arrays de la Unidad de Genómica del *IIS La Fe*. El software de análisis estadístico Expression Console (Affymetrix) se utilizó para la evaluación del control de calidad y la normalización de los valores de fluorescencia de las sondas. El software Partek Genomic Suite 6.6 version (Partek Incorporated, St. Louis) se utilizó para el análisis estadístico de los datos de microarrays. El análisis de la varianza (ANOVA) se realizó para las diferentes categorías y se aplicaron filtros de veces de cambio ó fold-change (FC) y de p-valor para generar las listas de genes con valores de expresión de sonda más importantes. Para determinar los patrones de expresión génica de cada categoría se generaron análisis de agrupación jerárquica (HC) a partir de las listas de genes

expresados diferencialmente. Se realizaron análisis integrados de Ontología de Genes (Gene Ontology Enrichment) y de las Vías de Señalización Celular a partir de las listas de genes más diferencialmente expresados entre categorías. Todas las pruebas estadísticas incluyeron análisis pareados para comparar la expresión génica de las distintas transfecciones pareadas (células $CD34^+$ transfectadas con siRNAs específicos de *MYBL2* en comparación con células $CD34^+$ transfectadas con siRNAs control; y la comparación entre las células $CD34^+$ transfectadas con el plásmido de sobreexpresión *MYBL2*–IRES–EGFP con las células $CD34^+$ transfectadas con el plásmido control IRES–EGFP).

14. Ensayos de células formadoras de colonias a partir de progenitoras $CD34^+$ transfectadas

Las células viables $CD34^+$ transfectadas con moduladores de expresión de *MYBL2* ó con sus respectivos controles, se sembraron por triplicado en el medio Methocult GF H84435 (Stem Cell Technologies). Tras 14 días en cultivo, el contaje y la clasificación de las colonias se llevó a cabo utilizando un microscopio invertido de alta calidad (Axiovert de Zeiss).

15. Análisis estadístico

Las diferencias en la distribución de las variables categóricas se analizaron mediante la prueba del chi-cuadrado, mientras que las pruebas de Mann-Whitney ó de Kruskal-Wallis se aplicaron para analizar las variables continuas de los datos que no superaron la prueba de normalidad. El análisis de supervivencia se limitó a los pacientes con LMA no-promielocítica y se determinó la tasa de supervivencia global (OS) y la supervivencia libre de enfermedad ó de recaída (SLE ó SLR, respectivamente). La OS se determinó a partir de la fecha del diagnóstico hasta la fecha de la muerte, censurando los pacientes vivos en la última fecha del seguimiento. La SLE se calculó a partir de la fecha de remisión completa (RC) hasta la fecha de recaída o muerte por cualquier causa, censurando los pacientes que se mantuvieron con vida en RC continua. La SLR se calculó a partir de la fecha de RC hasta la fecha de la recaída, censurando los pacientes que permanecían vivos sin recidiva. Las probabilidades estimadas para OS, SLE ó SLR se calcularon mediante el método de Kaplan-Meier y la prueba de log-rank evaluó las diferencias entre las distribuciones de supervivencia. La odds ratio y los intervalos de

confianza del 95% se calcularon mediante regresión logística. La expresión del ARNm de *MYBL2* se transformó en una variable categórica empleando la mediana de expresión de *MYBL2* de los pacientes (los pacientes con mayor expresión de *MYBL2* obtuvieron valores superiores a los de la mediana y los pacientes con menor expresión de *MYBL2* presentaron valores equivalentes o inferiores a los de la mediana). Todos los p-valores informados son bilaterales y los $p<0.05$ se consideraron estadísticamente significativos. Los análisis estadísticos se realizaron con el programa SPSS v17.0 (Chicago, IL).

IV. Resultados y discusión

1. Aplicabilidad del método de RT–PCR en tiempo real para la detección simultánea de genes de fusión asociados a LMA

El método de qRT–PCR demostró una elevada sensibilidad, especificidad y reproducibilidad de ensayo para la detección de reordenamientos específicos de LMA. El ensayo de qRT–PCR confirmó todas las anomalías que previamente habían sido detectadas mediante citogenética, FISH ó con los métodos convencionales y estándar de RT–PCR. Además, el método de qRT–PCR fue lo suficientemente sensible para identificar pequeños reordenamientos crípticos que no se detectaron mediante el método del cariotipo convencional. La qRT–PCR también detectó reordenamientos en los pacientes en los que no se obtuvieron suficientes metafases para el análisis citogenético estándar ó en aquéllos en los que la calidad de la muestra requerida resultó inadecuada para la técnica de FISH. La qRT–PCR también fue capaz de identificar todos los genes asociados en los reordenamientos con el gen *KMT2A* (*MLL)* y el reordenamiento *DEK–NUP214*, los cuales no se detectaron mediante FISH ya que las sondas específicas no estaban disponibles.

El método de qRT–PCR detectó un total de 55 reordenamientos en la serie de validación de 105 pacientes. Estos reordenamientos generaron temperaturas de melting (Tm) específicas para los reordenamientos más comunes excepto en el caso de un tránscrito de fusión *CBFB–MYH11*, que mostró una Tm distinta a la esperada para el reordenamiento más común *CBFB–MYH11* de tipo A, utilizado como control. Por lo tanto, el método de qRT–PCR también permitió la detección de variantes de tránscritos

inusuales. Un hallazgo importante de este estudio fue la detección simultánea de dos reordenamientos distintos en dos pacientes. Uno de los pacientes presentaba LMA y otro una LMA-promielocítica asociada a la terapia, y en ambos casos, se detectó el reordenamiento *BCR–ABL1* junto con otros genes de fusión: *RUNX1–MECOM* junto con *BCR–ABL1* y *PML–RARA* junto con *BCR–ABL1*, respectivamente. En el hipotético caso de un fallo ó falta de resultados mediante las técnicas de FISH ó citogenética convencional, el estudio diagnóstico típico de un paciente con rasgos morfológicos, inmunofenotípicos y clínicos característicos de la LMA-promielocítica estaría orientado a la evaluación del reordenamiento *PML–RARA*, y por tanto, el reordenamiento *BCR–ABL1* habría permanecido sin ser detectado. Por lo tanto, la integración de este procedimiento de qRT–PCR para la detección simultánea de reordenamientos moleculares (en lugar de los procedimientos convencionales para la detección individual) con los métodos estándar de citogenética y FISH, podría contribuir a la detección de más de un reordenamiento molecular en un mayor número de pacientes con LMA.

2. *MYBL2* como un nuevo marcador molecular en la LMA

A pesar de que los reordenamientos moleculares se consideran un elemento importante para el diagnóstico, pronóstico y seguimiento de la LMA, cerca del 40% de los pacientes muestra un cariotipo normal. Principalmente en este grupo de LMA, cuyo pronóstico es muy heterogéneo, es conveniente identificar nuevos marcadores moleculares. El estudio de nuevos marcadores moleculares de LMA contribuye al pronóstico, mejora el enfoque terapéutico y permite la investigación de los mecanismos implicados en la leucemia, lo que resulta imprescindible para el desarrollo de nuevas terapias. Por lo tanto, además del análisis exhaustivo de los reordenamientos moleculares, son necesarios estudios en profundidad de genes que se hayan asociado con el desarrollo de leucemia y el pronóstico de la misma, como *MYBL2*, y contribuir así a la investigación de nuevos marcadores moleculares en LMA.

2.1 Estudio del polimorfismo S427G de *MYBL2* en la LMA

La distribución de los genotipos del polimorfismo S427G en el grupo control fue consistente con la distribución registrada en las bases de datos para población con ascendencia europea (CEU, residentes de Utah con ascendencia de Europa Septentrional

y Occidental), de acuerdo con el Proyecto Internacional HapMap (hapmap.ncbi.nlm.nih.gov/). Sin embargo, no se detectaron diferencias significativas en la distribución del genotipo salvaje y la del genotipo portador del polimorfismo entre el grupo de pacientes y el grupo control ó la población CEU, respectivamente. Por lo tanto, los resultados obtenidos sugieren que el polimorfismo S427G no contribuye a la susceptibilidad de desarrollar LMA.

La presencia del polimorfismo S427G en pacientes con LMA no se asoció con ninguna característica clínica o biológica. No se obtuvieron diferencias significativas en la tasa de RC ó de muerte durante la terapia de inducción entre los pacientes con LMA-no promielocítica portadores y los no portadores del polimorfismo. No se obtuvieron diferencias significativas en el OS, SLE ó SLR entre los pacientes con LMA-no promielocítica portadores y los no portadores del polimorfismo, tampoco cuando se consideraron subgrupos de edad ó de estado citogenético determinados.

2.2 Búsqueda de nuevas variaciones genéticas de *MYBL2* en la LMA

El método de HRM detectó alteraciones en las curvas de fusión de 65 (42.8%) de los 152 pacientes con LMA analizados. La secuenciación directa confirmó que las muestras con una curva de fusión distinta a la del tipo salvaje albergaban alteraciones en la secuencia de ADN. En total, se detectaron 12 alteraciones genéticas distintas: seis variaciones no sinónimas (Q67X, E132D, P274L, S427G, V595M y I624M); cuatro variaciones sinónimas que no implicaban un cambio de aminoácido (I160I, D195D, P302P y D571D); y dos alteraciones en regiones intrónicas colindantes (c.186 + 14T> C en el intrón 3 y c.1365 + 3G> T en el intrón 8). Aparte del polimorfismo S427G, siete de estas 12 alteraciones genéticas detectadas en pacientes con LMA estaban registradas como polimorfismos en las bases de datos del Proyecto Internacional HapMap, del NCBI o en las bases de datos del NCI SNP. El polimorfismo rs73116571 (el cambio c.1365 + 3G>T) presentó una incidencia significativamente mayor en los pacientes con LMA (17.9 veces mayor) en comparación con la incidencia registrada para población con ascendencia europea en las bases de datos de polimorfismos. El análisis de las regiones de corte y empalme mediante el software Human Splicing Finder (http://www.umd.be/HSF) mostró que este polimorfismo se localizaba en una región potencial de corte y empalme del intrón 8 de *MYBL2*. Además, los pacientes portadores del polimorfismo rs73116571 presentaron bajos niveles de expresión de *MYBL2*, lo cual

sugiere que dicho polimorfismo podría alterar el proceso de corte y empalme, y por tanto, la síntesis de la hebra de ARNm. En base al conocimiento del papel de *MYBL2* en el mantenimiento de la integridad cromosómica, cabe pensar que la reducción de sus niveles de expresión a causa de la presencia del polimorfismo rs73116571 podría encaminar a los precursores mieloides hacia una inestabilidad del genoma, y de esta manera, favorecer su predisposición al desarrollo de LMA. De hecho, recientes estudios han demostrado que la haploinsuficiencia de *MYBL2* favorece la predisposición a desarrollar múltiples trastornos mieloides, como la LMA, en ratones de edad avanzada. Además, se ha demostrado que *MYBL2* se expresa en niveles marcadamente reducidos en células $CD34^+$ de pacientes con síndromes mielodisplásicos, sugiriendo que la regulación a la baja de la actividad de MYBL2 subyace a la expansión clonal de progenitores hematopoyéticos anormales en una gran proporción de tumores malignos humanos de la serie mieloide. Además, se ha demostrado que los ratones desarrollan un trastorno clonal mieloproliferativo / mielodisplásico a partir de células que mantienen una expresión reducida de *MYBL2*. Por lo tanto, estos resultados sugieren que la presencia del polimorfismo rs73116571 podría predisponer al desarrollo de la LMA.

Por otro lado, cuatro de las 12 alteraciones genéticas detectadas en la secuencia codificante de *MYBL2* de los pacientes no estaban registradas en las bases de datos de polimorfismos. Entre éstas, el cambio E132D fue clasificado como una variación genética benigna mediante el software PolyPhen (Polymorphism Phenotyping). Por el contrario, Q67X generaba un codón de parada en el exón 4 que codifica para R1, el primer dominio de unión al ADN de *MYBL2*. Este dominio está altamente conservado en la familia de los factores de transcripción MYB y contiene motivos de doble hélice que estabilizan el complejo de unión ADN / proteína (http://www.uniprot.org/uniprot/P10244). Por lo tanto, estos resultados sugieren que Q67X puede representar una variación genética maligna ya que podría desencadenar el truncamiento de la proteína ó la desintegración mediada sin sentido del tránscrito mutante, reduciendo los niveles de MYBL2. De hecho, el paciente portador del cambio Q67X mostró bajos niveles de expresión de *MYBL2*. Sin embargo, este paciente murió poco después de la fecha del diagnóstico debido a complicaciones no relacionadas con LMA. Por lo tanto, no fue posible obtener muestras de la RC ó muestras de miembros de la familia para evaluar si Q67X representaba una variación somática ó germinal, así como para descartar la posibilidad de que fuera un polimorfismo no anotado. Sin

embargo, ninguna de las nuevas alteraciones genéticas identificadas (incluyendo la Q67X) estaba registrada en las bases de datos de secuenciación del genoma ó del exoma (Leiden Open Variation Database y The Catalog of Somatic Mutations in Cancer), ni en los datos publicados por la Red de Investigación del Atlas del Genoma del Cáncer (TCGA).

3. Estudios funcionales de *MYBL2* en células progenitoras CD34$^+$

Con la finalidad de identificar el posible mecanismo de activación de la LMA debido a la expresión aberrante de *MYBL2*, se realizó un estudio de los genes que alteraban su expresión a causa de las variaciones de expresión de *MYBL2* en las células progenitoras CD34$^+$ transfectadas, mediante microarrays de expresión génica. Por otra parte, las alteraciones en la proliferación / diferenciación de estos progenitores se estudiaron mediante ensayos CFC.

3.1 Identificación de los genes regulados por *MYBL2* en las células progenitoras CD34$^+$

El análisis HC claramente identificó un perfil de expresión génica distinto en las células CD34$^+$ transfectadas con siRNAs de *MYBL2* en comparación con las células CD34$^+$ transfectadas con siRNAs control. Mediante la aplicación de filtros de FC ± 2.5 y $p < 0.05$ en el test ANOVA, se determinó una lista de 270 genes que habían obtenido una mayor variación de su expresión en las CD34$^+$ después de silenciar a *MYBL2*. El análisis de Ontología de Genes demostró que la mayoría de estos genes participaba en procesos biológicos relacionados con la proliferación y la muerte celular. Principalmente, las funciones publicadas y descritas para estos genes en las bases de datos (como GeneCards, UniProt u OMIM) estaban involucradas en la promoción de la apoptosis y de los procesos de escisión-reparación del ADN, aunque algunas de éstas también contribuían a la transformación tumoral y a la supervivencia de células malignas. Además, otros genes importantes relacionados con la promoción de la apoptosis (como *p53*, *BAX* ó *CASP8*), con la inhibición de la progresión del ciclo celular (como *CHEK1 / 2*, *RB1* ó *CDKN1B / p27*) y con la condensación de la cromatina (*ACIN1*), se detectaron sobreexpresados de manera significativa después de silenciar a *MYBL2*, aunque presentaron valores inferiores de FC. Estos hallazgos corroboran a otros estudios previos que manifiestan una función esencial de *MYBL2* en la transición

G2 / M del ciclo celular y en el mantenimiento de la estabilidad genómica. García et al. demostraron en líneas celulares de megacarioblastos que *MYBL2* era esencial para la progresión del ciclo celular en la fase S, así como en la estabilidad genómica. En estos estudios publicados, la reducción de la expresión de *MYBL2* representó una disminución en la progresión de la fase S del ciclo celular, la detención de la mitosis, la condensación aberrante de los cromosomas y su fragmentación. En relación a los resultados hallados, un estudio previamente publicado demostró en muestras procedentes de varios tipos de tumores humanos y en distintas fases de la progresión, que en las lesiones precursoras tempranas, pero no en los tejidos normales, comúnmente se expresaban marcadores de respuesta a daño en el ADN. Este estudio concluyó que en las fases tempranas de la tumorigénesis (antes de la inestabilidad genómica y la conversión maligna), las células humanas activan una red de respuesta a daño en el ADN que regula a sustratos clave para la reparación del ADN, la regulación de los puntos de control del ciclo celular, la integridad genómica, el origen y la estabilidad de la horquilla de replicación del ADN, con la finalidad de retrasar o prevenir el cáncer. Por otra parte, la expresión de genes para los que se ha descrito su regulación a través de *MYBL2* (como los genes *CLU [apolipoproteína J]*, *CCND1 [ciclina D1]* y *CCNA1 [ciclina A1]*) resultó disminuida después de silenciar a *MYBL2* en las $CD34^+$. Por último, el análisis integrado de las Vías de Señalización Celular mostró que las vías de la "Desregulación de la Transcripción en Cánceres", la vía de la "LMA" y de la "Leucemia Mieloide Crónica", así como otras vías relacionadas con distintos tipos de cáncer, resultaron entre las más significativamente alteradas en las células $CD34^+$ después del silenciamiento de *MYBL2*.

Por otro lado, el análisis HC identificó un perfil de expresión génica distinto en células $CD34^+$ transfectadas con el plásmido de sobreexpresión *MYBL2*–IRES–EGFP en comparación con las células $CD34^+$ transfectadas con el plásmido control IRES–EGFP. A través de la aplicación de filtros de FC ± 1.5 y criterios de $p < 0.05$ en el test ANOVA, se determinó una lista de 43 genes que presentaron una mayor variación en su expresión después de inducir la sobreexpresión de *MYBL2* en las $CD34^+$. El análisis de Ontología de Genes demostró que la mayoría de estos genes participaba en procesos relacionados con la promoción del metabolismo celular y con la respuesta celular a estímulos. Principalmente, las funciones previamente publicadas y descritas para estos genes en las bases de datos (como GeneCards, UniProt u OMIM), estaban involucradas

en la promoción de los procesos biológicos de adhesión celular, migración, inflamación, inmunidad, desarrollo del sistema nervioso central y desarrollo embrionario. El análisis integrado de las Vías de Señalización Celular mostró que la vía del "Linaje Celular Hematopoyético" resultó entre las más significativamente alteradas en las células $CD34^+$ después de inducir la sobreexpresión de *MYBL2*. En conjunto, estos resultados sugieren que la sobreexpresión aberrante de *MYBL2* podría inducir la expresión de genes que contribuirían al mantenimiento de un estado proliferativo / inmaduro en las células $CD34^+$, lo que podría facilitar la progresión de la enfermedad leucémica. De acuerdo con estas suposiciones, varios estudios publicados demuestran que *MYBL2* se halla sobreexpresado en distintos tipos de cáncer, entre ellos la LMA, favoreciendo el desarrollo/crecimiento del tumor y contribuyendo a un peor pronóstico. Además, se ha demostrado que la sobreexpresión de *MYBL2* en líneas celulares altera la diferenciación de los precursores mieloides. En este sentido, un estudio reciente ha demostrado que *MYBL2* no solo controla y acelera la progresión del ciclo celular en las células progenitoras embrionarias sino que, además, contribuye al mantenimiento de su capacidad de auto-renovación y su identidad pluripotencial.

Por último, cabe destacar que la expresión de los genes *MLLT4, BCL2L11, MAP2K7, CHEK2, CDKN1B y WNT8A* resultó inversamente modificada en función de la elevación ó disminución de los niveles de *MYBL2* en las células progenitoras $CD34^+$. Por lo tanto, estos resultados podrían sugerir que la expresión de estos genes se basa en su regulación a través de *MYBL2*. De hecho, la base de datos UCSC Genome Browser (http://genome.ucsc.edu/) registró secuencias "ChIP-seq" ó dominios de unión para MYBL2 en los promotores así como en otras regiones de la secuencia de los genes *MLLT4, BCL2L11 MAP2K7* y *CHEK2*, lo cual sugiere una interacción directa de *MYBL2* con los mismos.

3.2 Estudio de las alteraciones desencadenadas por *MYBL2* en la proliferación de las células progenitoras $CD34^+$ mediante ensayos de formación de colonias

Los ensayos CFC demostraron que el recuento total de colonias, a partir de células $CD34^+$ transfectadas con siRNAs de *MYBL2*, fue significativamente inferior al de las $CD34^+$ transfectadas con siRNAs control (mediana, 54 vs 76; rango, 36-65 vs 61-86; $p = 0.005$). El recuento total de unidades formadoras de colonias de monocitos

(CFU-M) y de granulo-monocitos (CFU-GM) también fue significativamente inferior en las células transfectadas con siRNAs específicos de *MYBL2* (mediana, 31 vs 46; rango, 23-39 vs 36-61; p = 0.002).

Del mismo modo, el recuento total de colonias de células $CD34^+$ transfectadas con el plásmido de sobreexpresión *MYBL2*–IRES–EGFP fue significativamente superior al de las $CD34^+$ transfectadas con el plásmido control (mediana, 78 vs 54; rango, 66-115 vs 29-89; p = 0.004), siendo también significativamente superior el número total de CFU-M y CFU-GM (mediana, 40 vs 20.5; rango, 28-55 vs 13-41; p = 0.007).

En conjunto, estos resultados mostraron que la proliferación de células $CD34^+$ se afectaba significativamente debido a la modificación de la expresión de *MYBL2*. Además, estos resultados se aproximarían a los obtenidos en otros artículos publicados que sugieren que la sobreexpresión aberrante de *MYBL2* en distintos tipos de cáncer favorece el desarrollo/crecimiento del tumor, ya que en este caso, la sobreexpresión aberrante de *MYBL2* favoreció el desarrollo/crecimiento de CFC. A su vez, estos hallazgos asisten a los resultados obtenidos en previos estudios en los que se demostraba que la disminución de la expresión de *MYBL2* causaba la retención de la progresión de la fase S del ciclo celular, la detención de la mitosis, la condensación aberrante de los cromosomas y su fragmentación, y de este modo, alteraba el mantenimiento de la estabilidad genómica.

V. Conclusiones

Las principales conclusiones del trabajo presentado en esta tesis son:

1) El nuevo método de qRT–PCR desarrollado para la detección simultánea de múltiples genes de fusión asociados a LMA representa un ensayo versátil, sensible y fiable para la detección de reordenamientos recurrentes de LMA. Además, la aplicabilidad de este método en los laboratorios de diagnóstico molecular se sustenta debido a la eficiencia, simplicidad y rapidez del procedimiento, y complementa de manera productiva los análisis de citogenética y FISH.

2) El polimorfismo S427G de *MYBL2* no se asocia con el riesgo de desarrollar LMA ni representa un factor pronóstico en los pacientes.

3) El método de HRM demostró ser una técnica útil y eficiente para detectar alteraciones genéticas de *MYBL2* en pacientes con LMA.

4) El polimorfismo rs73116571 se asoció con bajos niveles de expresión del ARNm de *MYBL2* y con una mayor incidencia en los pacientes con LMA, lo cual indica una mayor predisposición a padecer LMA en los portadores.

5) La nueva alteración genética Q67X detectada podría modificar la actividad de MYBL2 en los pacientes con LMA a través del truncamiento de la proteína ó de la desintegración mediada del tránscrito mutante sin sentido, lo cual implicaría la reducción de los niveles de la proteína MYBL2.

6) El silenciamiento de la expresión de *MYBL2* en las células progenitoras hematopoyéticas $CD34^+$ humanas afecta a su desarrollo celular, promoviendo la inducción de los mecanismos genéticos de apoptosis y de detención del ciclo celular, lo cual podría desembocar en la inestabilidad de las células progenitoras y el desarrollo de LMA.

7) La sobreexpresión de *MYBL2* en las células progenitoras humanas $CD34^+$ promueve su proliferación y la expresión de genes implicados en el desarrollo embrionario, lo cual podría implicar una supervivencia celular sostenida y una resistencia a los mecanismos de diferenciación y muerte celular programada.

8) Los genes *MLLT4*, *BCL2L11*, *MAP2K7*, *CHEK2*, *CDKN1B* y *WNT8A* se expresan en las células progenitoras $CD34^+$ en función de los niveles de expresión de *MYBL2*, sugiriendo una regulación directa a través de la interacción de MYBL2.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AL	Acute leukemia
ALL	Acute lymphoid leukemia
AMKL	Acute megakaryoblastic leukemia
AML	Acute myeloid leukemia
APL	Acute promyelocytic leukemia
ATRA	Retinoic acid
BCR	Breakpoint region
BM	Bone marrow
CEU	Utah residents with ancestry from Northern and Western Europe
CFC	Colony forming cell
CI	Confidence interval
CML	Chronic myeloid leukemia
CR	Complete remission
Ct	Cycle threshold
CV	Coefficient of variation
DFS	Disease free survival time
FAB	French-American-British classification
FC	Fold-change
FITC	Fluorescein isothiocyanate
GFP	Green fluorescent protein
HC	Hierarchical clustering
HCC	Hepatocellular carcinoma
HES	Hydroxyethyl Starch 6 % solution
HP	Hematopoietic stem cell progenitors
HPLC	High-performance liquid chromatography
HRM	High resolution melting
LSC	Leukemic stem cell
MC	Mononuclear cells
MDS	Myelodysplastic syndrome
MRD	Minimal residual disease
NCBI	National Center for Biotechnology Information
NCI	National Cancer Institute
nt	Nucleotide
OR	Odds ratio
OS	Overall survival rate
PB	Peripheral blood
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
qRT–PCR	Reverse-transcription and real-time polymerase chain reaction
R1	DNA binding domain 1 of *MYBL2*
R2	DNA binding domain 2 of *MYBL2*
R3	DNA binding domain 3 of *MYBL2*
RD	C-terminal regulatory domain
RFS	Relapse free survival time
SDS	Sodium dodecyl sulfate
SNP	Small nucleotide polymorphism
TA	Transcriptional activation domain
TCGA	Cancer Genome Atlas
Tm	Melting temperature
T-TBS	Tris-buffered saline containing Tween 20
UDG	Uracil-DNA N-glycosylase
WHO	World Health Organization

I. INTRODUCTION

1. Normal hematopoiesis and acute myeloid leukemia

Hematopoiesis is the process of blood cell production and maturation in the bone marrow (BM). The process begins when the hematopoietic stem cell progenitors (HP) gives rise to common myeloid or lymphoid HP, which are precursor cells of all blood lineages. HPs are responsible for the constant renewal of blood since are capable of proliferating itself and they ultimately differenciate into a variety of mature specialized cells [1]. A well-known marker for HP is the single chain transmembrane glycoprotein CD34 antigen, which is involved in the maintenance of the HP in a phenotypically undifferentiated state since its expression decreases as the cell matures [2]. Homeostasis of hematopoiesis could be altered when $CD34^+$ HP acquire certain abnormal molecular characteristics which cooperate to alter normal mechanisms of self-renewal, proliferation and differentiation, leading to blood cancer such as acute myeloid leukemia (AML) [Figure 1] [3, 4].

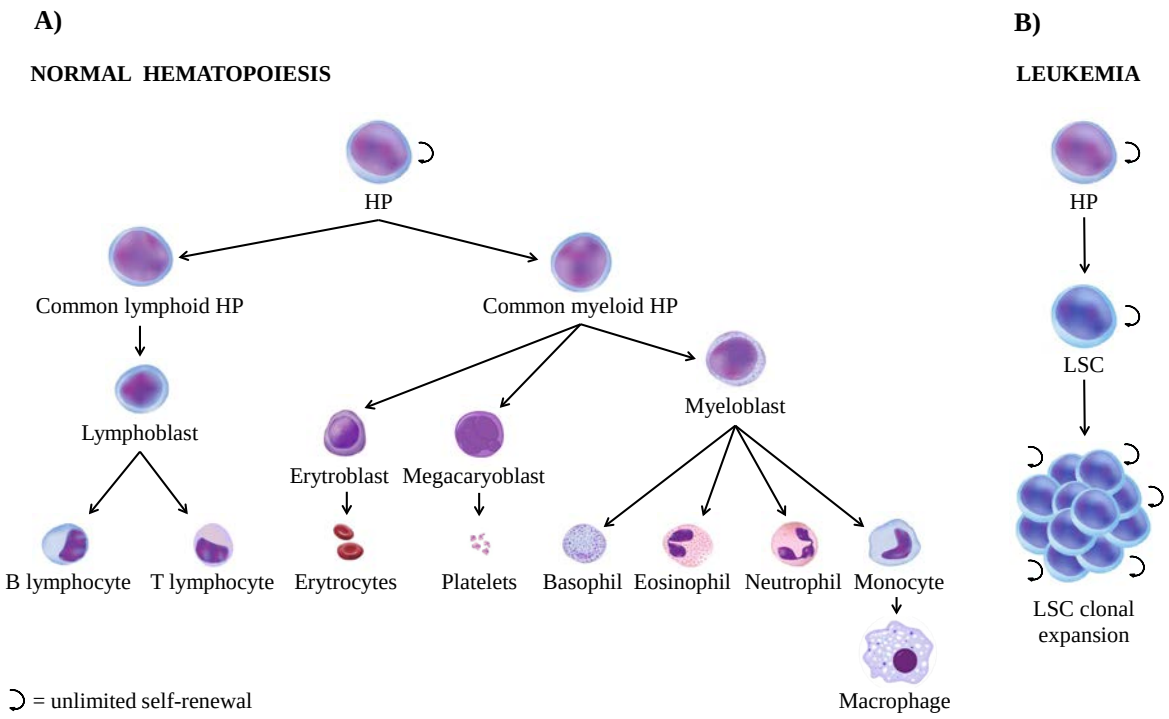

Figure 1. Schematic model of normal and malignant hematopoiesis. In normal hematopoiesis (A), myeloid hematopoietic stem cell progenitors (HP) differentiate into the granulocytic (neutrophil, basophil and eosinophil), monocytic/macrophage, erythroid or megakaryocytic cell lineages. Lymphoid HP differentiate into B- or T- lymphoid lineages. In leukemia, malignant hematopoiesis (B) causes the formation of a leukemic stem cell (LSC) with an undifferentiated phenotype and an unlimited self-renewal capacity, since LSCs share many characteristics with normal HP, including quiescence, multipotency and self-renewal [4].

AML is a rapidly progressive blood cancer which without treatment is lethal within weeks. It is characterized by a clonal proliferation of abnormal myeloid HP in the BM and peripheral blood (PB) known as leukemic myeloid blasts [5]. AML is the most common acute hematological malignancy in adults, with an estimated annual incidence rate of 1/25000-1/33000 [6, 7]. The incidence of AML is slightly higher in men than in women (1.3:1) and increases with age, typically displaying a median of 60 years of age at the onset of the disease [5, 8]. Diagnosis relies on laboratory findings showing anemia, thrombocytopenia and leucopenia or leukocytosis which result from disturbed hematopoietic function due to BM and PB infiltration by the leukemic blast cells. The blast percentage remains a practical tool for categorizing myeloid neoplasms and judging its progression. Therefore, diagnosis of AML also relies on BM aspirate or biopsy after the disease has been suspected, that should have at least 20% of myeloid blasts to be considered as AML [9]. Morphologic, cytochemical, and/or immunophenotypic features are used for establishing the linage of the neoplastic cells and for assessment of their maturation.

2. Cytogenetics and molecular biology in the diagnosis, prognosis and monitoring of AML

Nowadays, it is known that the proliferation of the abnormal myeloid progenitor clones is due to the accumulation of somatically acquired genetic changes [10]. The earliest genomic alterations in AML were identified nearly 40 years ago from patients with recurrent, nonrandom karyotypic alterations. To date, more than 100 balanced chromosomal rearrangements (translocations, insertions, and inversions) have been identified and cloned in AML [11]. In contrast to solid tumors, which often present complex karyotypes with different structural and numerical abnormalities, 50-70 % AML patients usually present specific recurrent single chromosomal alterations, with evidence suggesting that these are critical initiating events in the pathogenesis of AML [11, 12]. This information has been taken into account in the latest revision of the World Health Organization (WHO) classification [9], which incorporates and interrelates morphology, cytogenetics, molecular genetics and immunologic markers in an attempt to construct a classification that is universally applicable (Table 1).

Table 1. The World Health Organization (WHO) AML classification.

Name	Description
Acute myeloid leukemia (AML) with recurrent genetic abnormalities	AML with t(8;21)(q22;q22); *RUNX1–RUNX1T1* AML with inv(16)(p13.1q22)] or t(16;16)(p13.1;q22); *CBFB–MYH11* Acute promyelocytic leukemia (APL) with t(15;17)(q22;q12); *PML–RARA* AML with t(9;11)(p22;q23); *KMT2A–MLLT3*; AML with t(6;9)(p22;q34); *DEK–NUP214*; AML with inv(3)(q21q26.2)] or t(3;3)(q21;q26.2); *RPN1–MECOM*; Megakaryoblastic AML with t(1;22)(p13;q13); *RBM15–MKL1*; Provisional entity: AML with mutated *NPM1* Provisional entity: AML with mutated *CEBPA*
AML with myelodysplasia-related changes	This category includes patients who have had a prior documented myelodysplastic syndrome (MDS) or myeloproliferative disease that then has transformed into AML. Includes: AML with complex karyotype
Therapy-related myeloid neoplasms	This category includes patients who have had prior chemotherapy and/or radiation and subsequently develop AML or MDS
Myeloid sarcoma	This category includes myeloid sarcoma
Myeloid proliferations related to Down syndrome	Transient abnormal myelopoiesis and Myeloid leukemia associated with Down syndrome
Blastic plasmacytoid dendritic cell neoplasm	Blastic plasmacytoid dendritic cell neoplasm
AML not otherwise categorized	Includes subtypes of AML that do not fall into the above categories

Prognosis of AML varies widely according to cytogenetics, molecular findings, response to induction treatment and age, being remarkably worse in patients over 60 years of age [5]. Cytogenetic analysis at diagnosis of AML is considered one of the most important prognostic factors. It has become apparent that karyotype analysis identifies biologically distinct subsets of AML that differ in their response to therapy and treatment outcome [11-15]. In fact, the cytogenetic analysis in AML blast cells has allowed the recognition of specific chromosomal alterations with relevant prognostic importance and classify patients into risk groups (Table 2). These abnormalities may be accompanied by additional disturbances, called secondary, which are believed to be later genetic events related to disease progression [11, 12].

Table 2. AML classification into cytogenetic risk groups.

Risk group	Citogenetic abnormality	Survival after 5 years	Relapse rate
Favorable	t(8;21), t(15;17), inv(16)	70%	33%
Intermediate	Normal, +8, +21, +22, del(7q), del(9q), 11q23 abnormalities, any structural or numerical change	48%	50%
Adverse	–5, –7, del(5q), 3q abnormalities, complex cytogenetics	15%	78%

Molecular biology studies have allowed a deeper knowledge of the genes which are involved in AML chromosomal aberrations. Frequently, these genes play important roles in the development and function of myeloid cells, encoding transcription factors, cell cycle regulators, signal transduction molecules, immunoglobulin molecules and T cell receptors. The translocations may alter cellular proto-oncogene functions located at or near the breakpoints through at least two mechanisms: either by the juxtaposition of a cellular proto-oncogene with the regulatory element of a specific tissue gene [16, 17], or by creating fusion genes that encode chimeric proteins with different functional characteristics from the two parent proteins [18, 19]. These alterations may lead to the oncogene activation and the silencing of tumor suppressor genes, causing inadequate proliferation and an indefinite self-renewal capacity, escaping signs of differentiation and programmed cell death. On the other hand, a gene can cause different fusion structures with different genes. The most striking case is the *KMT2A* gene (also called *MLL*, *ALL1*, *HTRX1* and *HRX*) in the band q23 of chromosome 11, for which have been described more than 40 different fusion structures together with an internal duplication. Thus, depending on the fusion structure, the *KMT2A* gene may contribute to the pathogenesis of myeloid or lymphoid malignancies [20-28].

The significant progress in genetics and molecular characterization lies not only in the opportunity to have a prognostic approach prior to treatment, but also opens new venues to investigate the mechanisms involved in leukemogenesis and knowledge of new therapeutic targets, allowing individualized treatment in the future. Additionally, the molecular characterization of the AML allows the identification and quantification of minimal residual disease (MRD) by molecular biology techniques. The MRD quantification has prognostic interest since it identifies patients who will relapse in advance, in order to adapt the intensity or duration of the treatment. The polymerase chain reaction (PCR) technique has high sensitivity (10^{-5} to 10^{-6}) and therefore is the

most appropriate methodology for such a study. A single determination of MRD by PCR has limited value as a prognostic factor; however, the evolutionary pattern of the MRD has greater predictability. Therefore, the serial assessment of MRD could be used for individualized evaluation of the postremission treatment [5, 29, 30].

3. Detection of molecular rearrangements by reverse transcription and real-time PCR in AML

The cytogenetic analysis is a laborious technique and yields just enough metaphases in 60% to 80% of BM samples. Therefore, efforts to design more sensitive molecular biology techniques have been ongoing to detect molecular rearrangements as a result of cytogenetic abnormalities [31]. The PCR analysis does not require too much material from patients, it can be performed in cell debris and is highly sensitive in detecting few abnormal cells. Certainly, it has helped greatly to bring the PCR methodology ahead of AML diagnosis.

Although the aforementioned molecular rearrangements can be detected by single standard reverse-transcription and real-time PCR (qRT–PCR) reactions, given the current large number of fusion genes and variants of breakpoints characterized, these are labor intensive, costly, and time consuming. Several studies have reported methods to perform multiplex RT–PCR that consists of multiple primer sets within a single PCR mixture, to produce specific amplicons of different acute leukemia (AL) target lesions [31-38]. However, all these methods are based on conventional qualitative PCR and have a limited coverage of AML specific rearrangements. On the other hand, the recent and innovative high-throughput methods (arrays, next generation sequencing, etc.) are still not sufficiently cost effective for clinical diagnostic laboratories in hospital environments. Therefore, it is important to optimize methods of qRT–PCR that would prove useful and easily applicable for routine molecular diagnosis.

3.1 Relevant translocations and its specific molecular rearrangements in AML

The most relevant translocations and molecular rearrangements that have been recognized by the WHO classification as specific entities of AML include:

t(8;21)(q22;q22) (*RUNX1–RUNX1T1*), inv(16)(p13.1q22) or t(16;16)(p13.1;q22) (*CBFB–MYH11*), t(15;17)(q24;q12) (*PML–RARA*), t(9;11)(p22;q23) (*KMT2A–MLLT3*), t(6;9)(p22;q34) (*DEK–NUP214*), inv(3)(q21q26.2) or t(3;3)(q21;q26.2) (*RPN1–MECOM*) and t(1;22)(p13;q13) (*RBM15–MKL1*). Regarding AML with 11q23 (*KMT2A*) abnormalities, only t(9;11)(p22;q23) (*KMT2A–MLLT3*) is considered a separate entity in the WHO classification. Nevertheless, the identification of other translocations involving *KMT2A* is generally recommended [9]. Additional examples of less frequent fusion genes occurring in AML, which could have prognostic significance include t(9;22) (q34;q11) (*BCR–ABL1*) [39], t(3;21) (q26;q22) (*RUNX1–MECOM*) [40], t(16;21)(p11;q22) (*FUS–ERG*) [41], and t(8;16)(p11;p13) (*KAT6A–CREBBP*) [42, 43]. Finally, the detection of rare acute promyelocytic leukemia (APL) variant carrying the t(11;17)(q23;q21) (*ZBTB16–RARA*) [44] is relevant due to its unresponsiveness to all-trans retinoic acid.

3.1.1 t(8;21)(q22;q22) and the *RUNX1–RUNX1T1* rearrangement

It was the first cytogenetic abnormality described in AML. It occurs in 5-12% of de novo AML cases with cytogenetic abnormalities and is associated with M2 French-American-British (FAB) subtype (approximately 25% of M2 carries this translocation) [45, 46], which usually shows granulocytic maturation, but also occurs in M1 and M4 FAB subtypes. They have a good prognosis and good response to chemotherapy (mainly treated with cytarabine, ARA-C, in high doses) [5, 47].

The t(8;21) fuses *RUNX1* gene (runt-related transcription factor 1), also identified as *AML1* (gene 1 of AML), *PEBP2* (subunit A of the polyoma enhancer binding protein 2) or *CBFA2* (A2 subunit of core binding factor), with *RUNX1T1* gene (runt-related transcription factor 1 translocated to 1 [cyclin D-related]), also identified as *ETO* (gene 821), *CDR* (gene related to cyclin-D) or *MTG8* (myeloid translocation gene in chromosome 8). The *RUNX1* gene consists of nine exons spanning a 150 Kb region and the *RUNX1T1* gene consists of 13 exons, along 87 Kb. The *RUNX1–RUNX1T1* fusion transcript is detected by qRT–PCR in all t(8;21)-positive AML cases. These patients predominantly generate PCR products with constant size, corresponding to a fusion that maintains the reading frame from exon 5 of *RUNX1* to exon 2 of *RUNX1T1*. The *RUNX1–RUNX1T1* fusion transcripts are found not only in cases of t(8;21), but also in cases with complex translocations and a small proportion of AML

t(8;21)-negative. The in-frame fusion of *RUNX1* to *RUNX1T1* generally occurs with breakpoints at intron 5 of *RUNX1* and breakpoints at intron 1b-2 in *RUNX1T1* gene (Figure 2). The reciprocal fusion transcript *RUNX1T1–RUNX1* has not been identified.

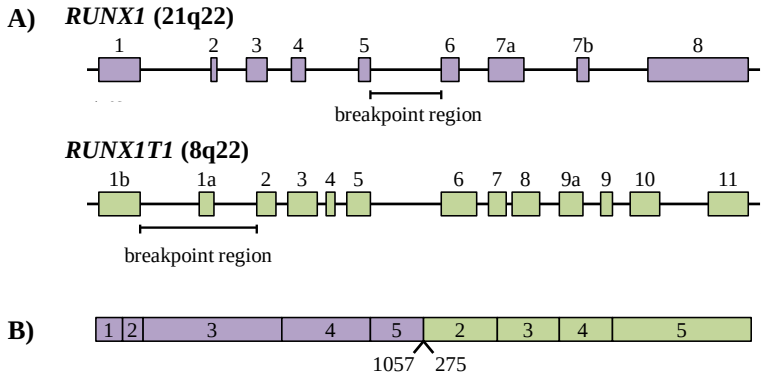

Figure 2. A) Schematic diagram of the exon / intron structure of the *RUNX1* (*AML1*) and *RUNX1T1* (*ETO*) genes, which are involved in the t(8;21)(q22;q22). The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of the *RUNX1–RUNX1T1* transcript with the exon 5 of *RUNX1* fused to exon 2 of *RUNX1T1*.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene. *(JJM van Dongen, Leukemia 1999).*

3.1.2. t(9;22)(q34;q11) and the *BCR–ABL1* rearrangement

The Philadelphia chromosome results from the balanced translocation t(9;22)(q34;q11). As a result of this translocation, the *BCR–ABL1* rearrangement occurs. More than 95% of chronic myeloid leukemia (CML) patients carry the *BCR–ABL1* gene in their leukemic cells. However, this is not exclusive for CML, as it has been found in approximately 30% (20-50%) of adults with acute lymphoid leukemia (ALL) and 2-10% of childhood ALL and in some cases of AML (1-3%). The AML in which the t(9;22) is detected are usually M1-M2 FAB subtypes and confers a poor prognosis, with rapid recurrence and resistance [5, 48].

The *BCR–ABL1* fusion gene encodes a protein with elevated tyrosine-kinase activity which appears to exert its effects by interfering with control death, cell proliferation and cell-cell adhesion. Several studies show that about two-thirds of CML patients produce mRNA from the reciprocal fusion gene *ABL1–BCR*, but its role remains unclear.

The points of frequent split in the *BCR* gene occur in exons 1 (also known as e1, which is located in the minor region of breakpoints called *"m–bcr"*), 13 and 14 (also called b2 and b3 respectively, which are located in the major region of breakpoints *"M–bcr"*), and 19 (or c3, located in the region of micro-breakpoints *"μ–bcr"*). The breakpoint in the *ABL1* gene usually occurs in exon 2 (a2), which after fusion with the different breakpoints in *BCR* gene, generated the rearrangements: e1a2, b2a2 or b3a2 and c3a2 (Figures 3 and 4). The products of these rearrangements correspond to fusion proteins of 190 kD ($p190^{BCR–ABL1}$), 210 kD ($p210^{BCR–ABL1}$) and 230 kD ($p230^{BCR–ABL1}$), mainly associated with the ALL, CML and neutrophilic CML respectively. In AML, most often are described both $p190^{BCR–ABL1}$ as $p210^{BCR–ABL1}$ [29, 39].

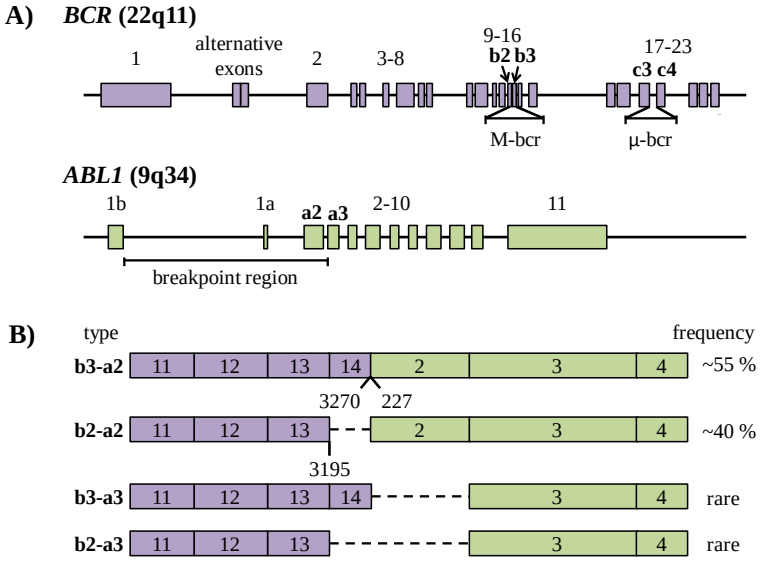

Figure 3. A) Diagram of the exon/intron structure of the *BCR–ABL1* fusion gene, involved in t(9;22)(q34;q11) focused on the M-bcr region. The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of p210 *BCR–ABL1* transcripts.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene. Fusion transcripts b3-a2 and b2-a2 have been found more frequently, but sporadic cases with b3-a3 and b2-a3 transcripts have been reported. *(JJM van Dongen, Leukemia 1999).*

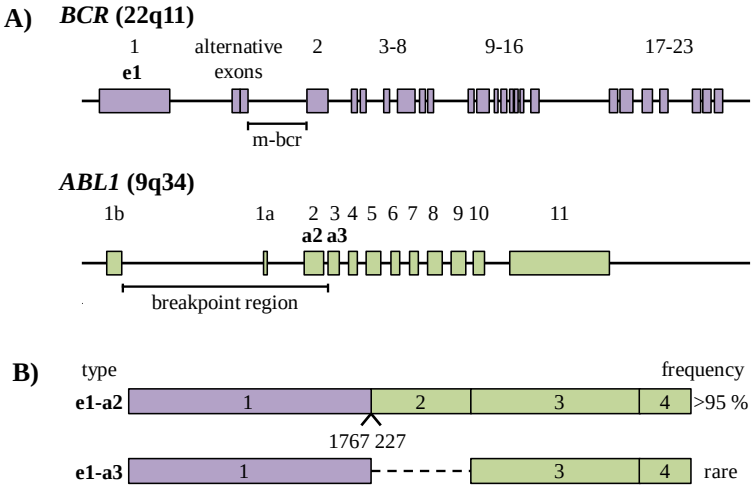

Figure 4. A) Diagram of the exon/intron structure of the *BCR−ABL1* fusion gene, involved in t(9;22)(q34;q11) focused on the m-bcr region. The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of p190 *BCR−ABL1* transcripts.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene. The fusion transcript e1-a2 is most frequently (> 95%), but sporadic cases with transcripts e1-a3 have been reported. *(JJM van Dongen, Leukemia 1999).*

3.1.3. t(15;17)(q22;q21) and the *PML–RARA* rearrangement

The t(15;17) occurs in approximately 10% of cases of de novo AML and is quasi pathognomonic of APL (a different subset of AML with M3 FAB cytomorphology), since it occurs in 99% of M3 [5, 29, 49]. The translocation of chromosomes 15 and 17, which produces the specific *PML–RARA* oncoprotein, itself is responsible for the specific response of this AML subtype to the treatment with retinoic acid (ATRA) [50]. Such targeted chemotherapy is effective and makes the APL have a very good prognosis, if it is treated properly. In the breakpoint of the long arm of chromosome 15 lies the *PML* gene (the promyelocitic leukemia gene) and on chromosome 17 the gene of the retinoic acid receptor alpha (*RARA*). By translocation both genes are fused, creating the chimeric protein *PML–RARA*. The breakpoints on chromosome 17 are located within a 15 Kb DNA fragment, at intron 2 of *RARA* gene. By contrast, there are three regions in the *PML* locus that are involved in the translocation breakpoints: intron 6 (bcr1; 55% of cases), exon 6 (bcr2; 5%) and intron 3 (bcr3; 40%) [Figure 5]. The existence of different breakpoint regions in the *PML* locus and the presence of

alternative splicing in *PML* transcripts, are responsible for large heterogeneity in *PML–RARA* junctions among APL patients [39, 51, 52].

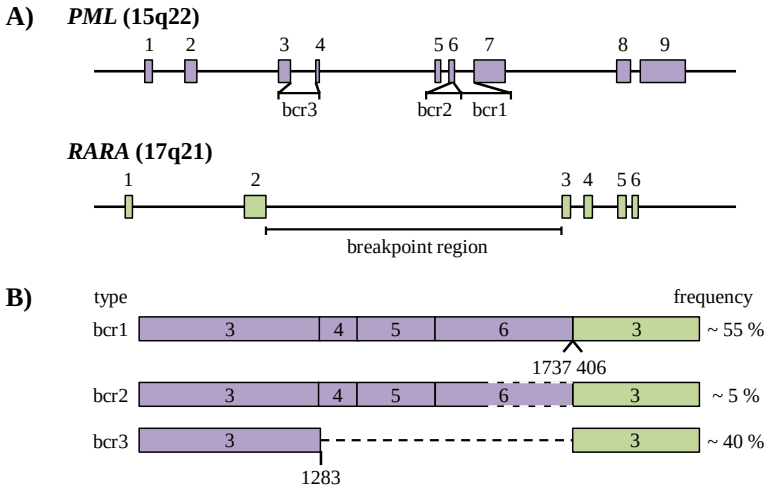

Figure 5. A) Schematic diagram of the exon/intron structure of the *PML* and *RARA* genes, involved in t(15;17)(q22;q21). The exon number and the regions of relevant breakpoints are indicated. The breakpoint regions of bcr1 and bcr2 are juxtaposed in intron 6 and exon 6, respectively. **B) Schematic diagram of the three types of *PML–RARA* transcripts relating to the different regions of *PML* breakpoints.** The bcr2 transcript size depends on the position of the breakpoint in *PML* exon 6. The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene. *(JJM van Dongen, Leukemia 1999).*

3.1.4. t(11;17)(q23;q21) and the *ZBTB16–RARA* rearrangement

The majority of patients with APL carry the hallmark t(15;17)(q22;q21) chromosomal translocation. However, the 1-2% of APL patients carries the t(11;17)(q23;q21) translocation, that fuses the *ZBTB16* gene (also known as the promyelocytic leukemia zinc finger gene [*PLZF*]) to the *RARA* gene [53]. The majority of zinc finger proteins play important roles in DNA binding, RNA binding, RNA packaging, and protein-protein interactions. In humans, the *PLZF* gene is localized to chromosome 11q23 among a cluster of genes, all related to the zinc finger family [54]. In APL, *ZBTB16–RARA* chimeric transcripts fuse the breakpoints within *ZBTB16* intron 3 or intron 4 with the breakpoints within intron 2 of *RARA* gene, leading to inclusion of 2 or 3 zinc fingers in the *ZBTB16* moiety of the ZBTB16–RARA fusion protein, respectively (Figure 6). In contrast to the hallmark t(15;17) APL, patients with the

ZBTB16–RARA fusion, who respond poorly to ATRA treatment, are refractory to concurrent cytotoxic chemotherapy and generally have poor outcomes [44]. These features make t(11;17) APL a distinct variant from typical APL. Effective therapeutic regimens for this APL subtype are imperatively needed.

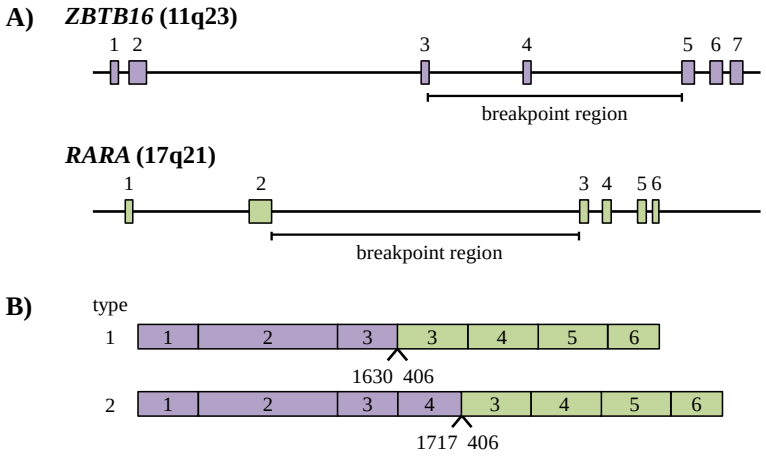

Figure 6. A) Schematic diagram of the exon/intron structure of the *ZBTB16* and *RARA* genes, involved in t(11;17)(q23;q21). The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of two known types of *ZBTB16–RARA* transcripts relating to the different regions of *ZBTB16* breakpoints.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene.

3.1.5. inv(16)(p13;q22) and the *CBFB–MYH11* rearrangement

The inv(16)(p13q22) (pericentric inversion) or t(16;16) (p13; q22) (translocation between both homologous chromosomes 16) and its molecular counterpart *CBFB–MYH11* occurs around 10-12% of all the AML. It is associated with abnormal eosinophilia-M4 (M4Eo FAB), characterized by monocytic and granulocytic differentiation and the presence of eosinophilic population. Patients carrying this disorder have a better prognosis than other AML, reaching high CR and disease-free survival (DFS) rates with schemes that include ARA-C in high doses [5, 29, 55].

Although most *CBFB–MYH11* positive leukemias correspond with AML-M4Eo, the transcripts have also been found in many other types of AML, including M4 without eosinophilic abnormalities, M2, M5 and, less frequently, in M1, M6 and M7 FAB

subtypes. The inv(16)(p13q22) also occurs in rare cases of CML in blast crisis, myelodysplastic syndrome (MDS) and secondary AML treatment.

The inv(16)(p13;q22) and t(16;16)(p13;q22) fuse *CBFB* gene (beta subunit of core binding factor), also called *PEBP2b* (polyoma enhancer binding protein 2β), located on chromosome 16q22 with *MYH11* gene (myosin heavy chain 11 gene), whose protein product is also identified as SMMHC (myosin heavy chain of smooth muscle) and is located on chromosome 16p13. The most breakpoints of *CBFB* are located on a 15 Kb intron between exon 5 and exon 6, shown in Figure 7A as intron 5. The expression of the protein CBFB is ubiquitous and it is a heterodimeric transcription factor together with CBFA1, CBFA2 (called RUNX1) or CBFA3.

The gene *MYH11* is composed by 21 exons that are expanded over 37 Kb. Ten different *CBFB–MYH11* fusion transcripts has been reported. More than 85% of positive patients carry the type A of the transcripts. The heterogeneity of breakpoints is marked much higher in *MYH11*, because there are up to seven different exons (exon 7 to 13) which can variably be included in the fusion transcripts, as shown in Figure 7B.

The subtle nature of karyotypic changes in inv(16) involves a difficult cytogenetic detection, particularly after banding R or when their presence is not suspected from the morphological results. Therefore, positive *CBFB–MYH11* by qRT–PCR but the negativity of inv(16) is not unusual [39].

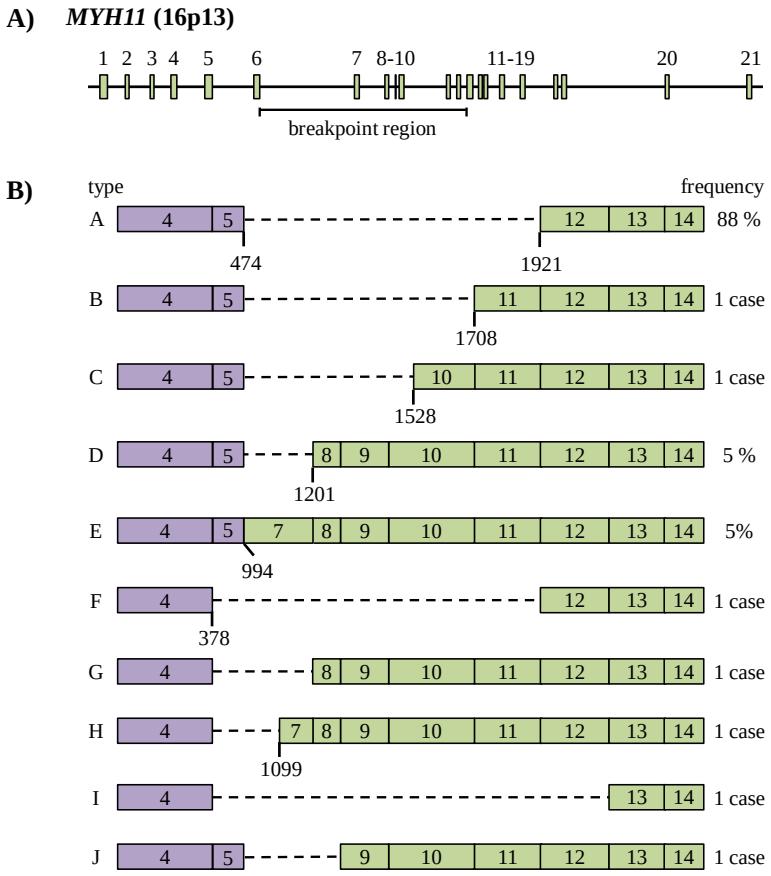

Figure 7. A) Schematic diagram of the exon/intron structure of the *MYH11* gene involved in inv(16)(p13;q22). The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of the 10 different *CBFB–MYH11* fusion transcripts.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene. Different types of transcripts are mainly caused by breakpoints in different introns of the gene *MYH11*. Fusion transcripts type A, D and E together represent approximately 98% of all patients, while other types concern isolated cases. *(JJM van Dongen, Leukemia 1999).*

3.1.6. 11q23 translocations and *KMT2A* rearrangements

Disturbances in 11q23 represent between 5 and 6% of all AML. These are more common in children and in secondary leukemias to treatment with topoisomerase II inhibitors. Morphologically, is often associated with AML, especially with the M5a FAB subtype. It involves an intermediate prognosis due to the large heterogeneity of rearrangements [56]. It is consistent of a breakpoint between q22-23 bands of chromosome 11, but the chromosomal material can be translocated to different chromosomes (6, 9, 10, 17 and 19 in the AML). Rearranged *KMT2A* gene (lysine (K)-specific methyltransferase 2A gene, also known as *MLL* [myeloid/lymphoid or mixed-lineage leukemia trithorax Drosophila homolog], *HRX, ALL-1 or HTRX1*) has been identified

with up to 30 different genes. The most common translocation in AML is t(9;11)(p22;q23), which is detected up to 35% of M5 [5].

The *KMT2A* gene is composed of 37 exons that expand into more than 1 MB region, encoding a 431 kD protein that acts as a transcription factor. The major breakpoint region of *KMT2A* gene is located on 8.3 Kb BamHI fragment, between exons 5 and 11 (Figure 8). It is believed the disruption of normal transcriptional function of KMT2A protein, due to its division or the production of several fusion proteins, is responsible for leukemogenesis [39, 57, 58].

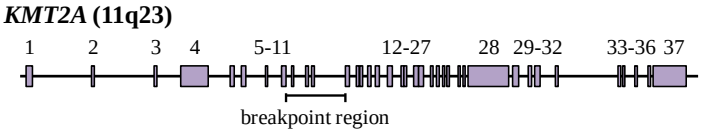

Figure 8. Schematic diagram of the exon/intron structure of the *KMT2A* gene. The exon number and the region of relevant breakpoints are indicated. *(JJM van Dongen, Leukemia). 1999).*

The t(9;11)(p22;q23) produces the fusion transcript *KMT2A–MLLT3*. The *MLLT3* gene (also known as AF9) spans more than 100 Kb and it has been identified by two breakpoint regions (BCRs): one in a telomeric region of intron 4 (BCR1), and another between intron 7 and intron 8 boundaries (BCR2) [57]. Other studies have reported these breakpoint regions as "site A" and "site B" [27]. The sequencing of several products has shown different fusion structures between: exon 8 of *KMT2A* and *MLLT3* Site A, exon 7 of *KMT2A* and *MLLT3* Site A or exon 6 of *KMT2A* and *MLLT3* Site A. This finding suggests the chimeric products are generated by alternative splicing in *KMT2A* gene. Interestingly, one study found the ruptures happened in central *MLLT3* exon (Site A) of all AML patients, while the breakpoint in *MLLT3* site B was found in the only ALL patients [59].

KMT2A is fused with *MLLT1* (also known as *ENL)* and *ELL* in AL containing the t(11;19)(q23;p13). *MLLT1* and *ELL* are located on chromosome 19, in the bands 13.3 and 13.1 respectively, but the breakpoints are not always easily distinguished by standard cytogenetics. Therefore, RT–PCR assays are performed to distinguish both

breakpoints. The *KMT2A–MLLT1* rearrangement is associated with ALL and conversely, *KMT2A–ELL* happens more often in AML [27, 60].

The t(10;11)(p12;q23), is observed in both pediatric and adult patients. Interestingly, it has been described that breakpoints in pediatric cases are restricted to a narrow region of *MLLT10* (also known as *AF10* gene*)* coding sequence (between bases 787 and 979), while breakpoints in adults are distributed throughout *MLLT10* coding sequence [61].

3.1.7. t(3;21)(q26;q22) and the *RUNX1–MECOM* rearrangement

Disturbances in 3q26 can be detected up to 2% of AML cases, which usually are present in young patients with trilinear dysplasia and confers a poor prognosis [62]. These disturbances are associated with conservation or an increase in the platelet counts at diagnosis, with abundant micromegakaryocytes in BM and poor response to chemotherapy. It can occur in any FAB group [5].

The fusion gene *RUNX1–MECOM,* is obtained from fusion in frame of *RUNX1* and *MECOM* (*MDS1* and *EVI1* complex locus) genes, and it is a product of the t(3;21)(q26;q22) that is associated with MDS and the novo AML patients or therapy related AML, and to a lesser extent in patients with CML in blast crisis.

The *RUNX1* gene interacts with transcriptional co-regulators and, in t(3;21), the breakpoints are located approximately 60 kb downstream of the t(8;21) breakpoints, usually at intron 5 or 6.

MECOM is a transcription factor (zinc finger DNA binding) and a longer form of the protein EVI1 associated to leukemia. *MECOM* is expressed in several tissues but is not detected in normal hematopoietic cells. *RUNX1–MECOM* comprises a DNA *RUNX1* binding domain fused with almost all *MECOM* gene. *RUNX1* and *MECOM* are both transcriptional activators, while *RUNX1–MECOM* is a transcriptional repressor. The runt domain is truncated in all *RUNX1* fusion proteins formed in the t(8;21) and t(3;21). The products of both translocations can act as repressors of the normal *RUNX1* transactivation function on myeloid specific promoters. This suggests both *RUNX1* truncations as fusions with the associated genes may play a key role in the mechanism of leukemogenesis (Figure 9) [63–67].

Several studies have described aberrant expression of *EVI1* without evidence of impaired 3q26, suggesting the existence of alternative mechanisms for this genes activation. *EVI1* overexpression is associated with poor prognosis [62].

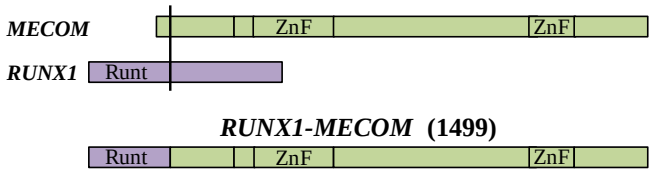

Figure 9. Diagram of the fusion protein RUNX1–MECOM. RUNX1, MECOM and RUNX1–MECOM fusion protein are shown in the three plot lines respectively. The vertical line indicates the fusion breakpoints. RUNX1–MECOM contains 1499 aa. The DNA binding domain of RUNX1, the Runt domain, is maintained in the fusion protein. The two zinc finger domains of MECOM (ZnF) are also shown. *(V. Senyuk et al, Oncogene 2002).*

3.1.8. t(6;9)(p22;q34) and the *DEK–NUP214* rearrangement

The t(6;9)(p22;q34) and its molecular counterpart *DEK–NUP214* are detected in less than 1% of AML patients and portend a poor prognosis [68]. It is associated with an increase of basophil proportion in BM [5]. This translocation typically occurs in young adults who belong to the FAB subtype M2 or M4 and rarely to the M1, although it does not have a preferred subtype.

The breakpoints on chromosome 9 are grouped in one of the introns of a large gene (> 100 Kb) called *NUP214* (also known as *CAN)* encoding a 7 Kb transcript. This intron is called icb-9 and is located in a 8 Kb region positioned 360 kb downstream of *ABL1* gene. In *DEK* gene, located on chromosome 6, the breakpoints of t(6;9) are grouped in a region of 12 Kb, also located in an intron (icb-6). As a result of *DEK–NUP214* fusion in AML t(6;9), a *DEK–NUP214* unchanged transcript is encoded (Figure 10). The cDNA sequence analysis demonstrated *DEK* and *NUP214* are joined without interruption of the original open reading frame and, therefore, the fused mRNA encodes a DEK–NUP214 165 kD chimeric protein [68].

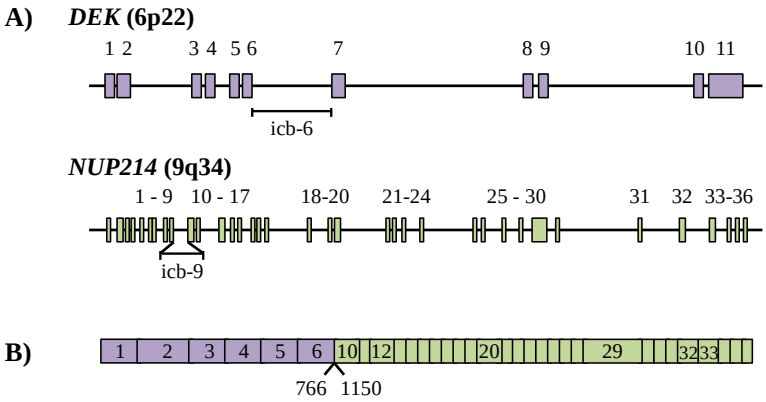

Figure 10. A) Schematic diagram of the exon/intron structure of the *DEK* and *NUP214* genes, involved in t(6;9)(p22;q34). The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of the *DEK–NUP214* fusion transcript.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene.

3.1.9. t(1;22)(p13;q13) and the *RBM15–MKL1* rearrangement

Acute megakaryoblastic leukemia (AMKL) is a rare AML variant, whereby leukemic blasts display morphologic and phenotypic features indicating megakaryocytoid differentiation (FAB M7). A distinct AMKL entity characterized by the t(1;22)(p13;q13) translocation, resulting in the *RBM15–MKL1* fusion oncogene, has been recently recognized by WHO classification, representing ＜1% of all AML cases [9]. This is predominantly a disease afflicting infants and displays characteristic clinical features. The prognosis of all AMKL is usually worse than that of other AML types, except for Down Syndrome (DS)–related AMKL with *GATA1* mutation, but AMKL with t(1;22)(p13;q13) appears to be uncertain. The translocation generates the fusion of exon 1 of the *RBM15* (RNA-binding motif protein-15 gene, also known as *OTT*), an RNA recognition motif-encoding gene, with exon 5 of the *MKL1* (megakaryocyte leukemia-1 gene, also known as *MAL*), a gene encoding a SAP (SAF-A/B, Acinus and PIAS) DNA-binding domain (Figure 11). The resulting fusion gene modulates HOX (homeobox) induced differentiation and extracellular signaling pathways. The fusion protein is expected to participate in chromatin organization through the AT-rich DNA sequences binding, recognized by SAF (scaffold attachment factor) box in the *RBM15–MKL1* fusion [69, 70].

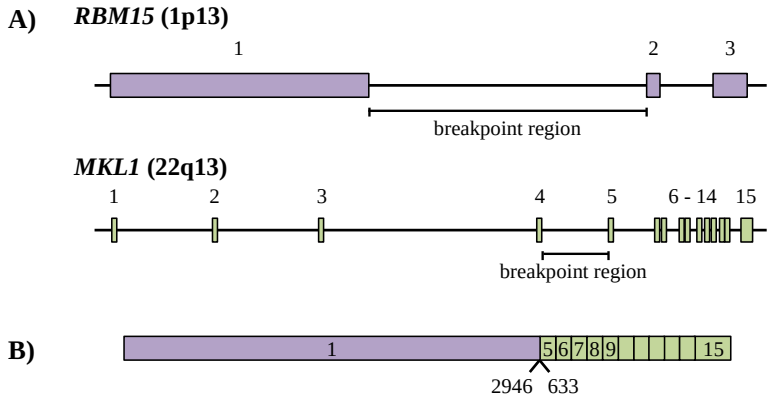

Figure 11. A) Schematic diagram of the exon/intron structure of the *RBM15* and *MKL1* genes, involved in t(1;22)(p13;q13). The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of the *RBM15–MKL1* fusion transcript.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene.

3.1.10. t(8;16)(p11;p13) and the *KAT6A–CREBBP* rearrangement

The t(8;16)(p11;p13) and the corresponding *KAT6A–CREBBP* fusion transcript occurs in <1% of the AML [42] and exhibits monocytic differentiation (FAB subtypes M4 and M5), erythrophagocytosis, dual myeloperoxidase and nonspecific esterases by leukemic cells, poor response to chemotherapy and poor prognosis [71]. The AML-t(8;16) patients show extramedullary involvement and clotting disorders at diagnosis [43].

At the molecular level, the translocation t(8;16) fuses *KAT6A* gene (also known as *MOZ* or *MYST3* [histone acetyltransferase of monocytic leukemia *MYST 3*]), located on 8p11 with *CREBBP* (also known as *CBP* [*CREB* binding protein]), in 16p13. Two types of fusion transcripts *KAT6A–CREBBP* (types I and II), containing 1128 and 415 bp respectively, have been detected (Figure 12). Sequencing of these fragments revealed fusion in-frame of the nucleotide (nt) 3745 of *KAT6A* with the nt 283 of *CREBBP* in transcript type I, while the same locus for *KAT6A* is fused out-of-frame with nt 997 of *CREBBP* in transcript type II. Recent genomic studies located the breakpoint within intron 16 of *KAT6A* [43, 71]. Furthermore, rare breakpoints within introns 17 and 15 of *KAT6A* have been described [72–74]. Almost all breakpoints in *CREBBP* chimeras are located in intron 2 and, in some cases, have been described in introns 3, 4, 7 and 8, indicating the existence of alternative splicing in *CREBBP* gene [43, 72–74].

KAT6A gene consists of 17 exons and contains a domain MYST histone acetyltransferase activity. *KAT6A* modulates transcription of specific genes by co-activation of the transcription factor complex *RUNX1*. Furthermore, it is believed *CREBBP* gene coordinates transcriptional effects of multiple signals from the cell surface and nuclear receptors [43].

It has been suggested that *KAT6A–CREBBP* transcript but not *CREBBP–KAT6A* is important in the process of leukemogenesis because some studies found *CREBBP–KAT6A* fusion transcript as an out-of-frame with a premature stop codon [71].

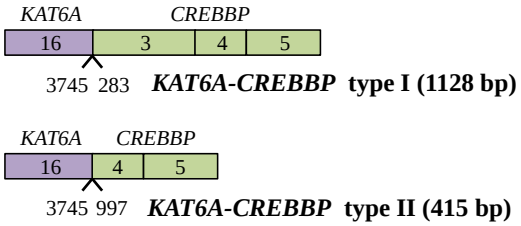

Figure 12. Schematic representation of chimeric transcripts *KAT6A–CREBBP (MOZ–CBP)* type I and type II. *(M. Rozman et al; Genes, Chromosomes and Cancer 2004).*

3.1.11. t(16;21)(p11;q22) and the *FUS–ERG* rearrangement

The t(16;21)(p11;q22) is a rare chromosomal aberration in AML and it is known to be associated with poor prognosis, young age (median of 22 years of age), and with the involvement of various FAB subtypes [41]. FUS is a nucleoprotein that functions in DNA and RNA metabolism, including DNA repair, and the regulation of transcription, RNA splicing and export to the cytoplasm. The oncogene *ERG* is a nuclear phosphoprotein that binds purine-rich sequences and several studies suggested *ERG* acts as a regulator of genes required for maintenance and/or differentiation of early hematopoietic cells [75]. The *FUS–ERG* fusion is a RNA-binding protein that is highly homologous to the product of the *EWS* gene, involved in Ewing sarcoma. In AML, different *FUS–ERG* chimeric transcripts which predict different leukemogenic FUS–ERG proteins have been reported (Figure 13) [76].

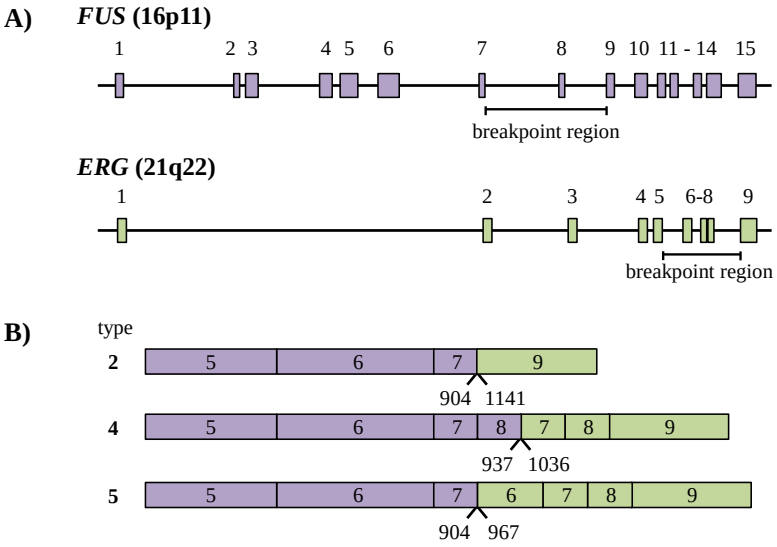

Figure 13. A) Schematic diagram of the exon/intron structure of the *FUS* and *ERG* genes, involved in t(16;21)(p11;q22). The exon number and the regions of relevant breakpoints are indicated. **B) Schematic diagram of different *FUS–ERG* fusion transcripts found in AML.** The numbers under the fusion gene transcript refers to the first nucleotide (5') of the involved exon of one fusion gene and the last nucleotide (3') of the upstream exon of the other fusion gene.

4. The importance of new molecular markers in AML with normal karyotype

In recent years, the importance of the molecular characterization of AML for a refined diagnosis, assessment of prognosis and monitoring of MRD has become clear [12, 13]. Recent knowledge of the AML pathogenesis suggests that it may be produced on the model of the sequential acquisition of genetic lesions. As mentioned before, near 60% AML is characterized by specific chromosomal abnormalities that produce oncogene activation or silencing of tumour suppressor genes [11, 77]. Aside of gross chromosomal abnormalities, recent studies have shown the existence of a number of molecular alterations that may be involved in an oncogenic cooperativity with the established recurrent rearrangements, suggesting an essential role in the development and maintenance of AML [78, 79]. As examples, several gene mutations including mutations in the tyrosine kinase receptor (*FLT3)*, nucleophosmin (*NPM1)*, the tyrosine-protein kinase (*KIT)*, CCAAT/enhancer-binding protein alpha (*CEBPA)*, tet oncogene family

member 2 (*TET2),* DNA (cytosine-5-)-methyltransferase 3 alpha (*DNMT3A)* or in the isocitrate dehydrogenase 1 (NADP+) [*IDH1]* [80-85], as well as the overexpression of the Wilms tumor 1 *(WT1*) gene [86], has been shown to carry diagnostic and prognostic information and experts consider some of these mutations as urged to be identified in AML [87, 88].

However, more than 25% of AML with normal karyotype lack molecular recurrent structural abnormalities or mutations in the known leukemia-associated genes, making more difficult the prognosis and best therapeutic approach to follow [11, 89-92]. Therefore, mainly in this group of patients, is desirable to identify new molecular markers in AML. Moreover, the study of new AML molecular markers also opens new venues to research into the mechanisms involved in leukemogenesis, which will be important to recognise new molecular targets for the development of new therapies and the possibility of personalized medicine in the future.

4.1. *MYBL2* gene: structure, ontology, localization and function

MYBL2 (v–myb myeloblastosis viral oncogene homolog [avian-like] 2), also known as *B–MYB, MGC15600 or OTTHUMP00000031719*, is located in the long arm of chromosome 20 (20q13.12) and the gene contains 14 exons in a total of 49415 nt. The mRNA is 2731 nt in length and it is traduced into a 93 kD protein with 704 amino acids (Figure 14) [93].

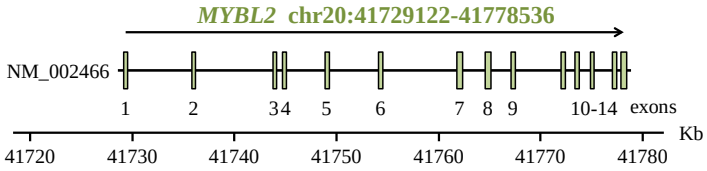

Figure 14. **Schematic diagram of *MYBL2* gene structure.** *(Chayka, O et al, Atlas Genet Cytogenet Oncol Haematol. 2009).*

MYBL2 is a member of the *MYB* transcription factors family, which includes *C–MYB* and *A–MYB*, and is required for cell proliferation, cell cycle progression, chromosomal stability and differentiation [94]. MYBL2 protein contains three DNA binding domains (R1, R2, and R3). R1 serves as a DNA/protein complex stabilizer and

R2/R3 contain helix-turn-helix motives with unconventional turns required for DNA-binding activity. Also, the protein contains one transcriptional activation domain (TA), a C-terminal regulatory domain (RD) for repressing the transactivation function of *MYBL2*, and a conserved region with homology to other *MYB* family members (Figure 15) [93].

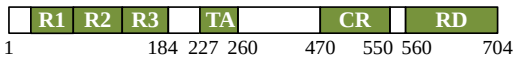

Figure 15. Schematic diagram of the MYBL2 protein structure. R1, R2 and R3 form DNA-binding domain, TA - transactivation domain, CR - conserved region (the area of homology with *C–MYB*), RD - regulatory domain. *(Chayka, O et al, Atlas Genet Cytogenet Oncol Haematol. 2009).*

All members of the MYB transcription factors family show high homology within TA and RD regions. DNA-binding domain is almost identical with that of *A–MYB* and *C–MYB* and is conserved between mouse, human, drosophila and chicken. Indeed, MYB transcription factors were first identified as cellular homologues of the *V–MYB* oncogene which is known to cause leukemia in chickens [94]. Among the three members of the *MYB* family in vertebrates, *MYBL2* has been suggested as the ortholog or ancestral parent and the other two members were originated by duplication during evolution [95]. In invertebrates, there is a single *MYB* gene which presents high sequence conservation and functional homology with *MYBL2*. In fact, it was shown that contrary to *A–MYB* or *C–MYB* vertebrate *MYBL2* was able to rescue the lethal phenotype of Drosophila *MYB* mutant by complementation assays [96].

MYBL2 is a nuclear protein which is ubiquitously expressed, more highly in proliferative, embryonic and hematopoietic progenitor cells, and proportionally to the degree of cell differentiation [95, 96]. This is in clear contrast to the other two members of the same family, which expression is tissue restricted. Thus, *A–MYB* expression is abundant in gonads and in mammary gland cells [97], while *C–MYB* is mainly expressed in hematopoietic cells [98]. The differences on the expression pattern of the three members are in agreement with the phenotype of the knockout mice generated. Thus, *MYBL2* knockout mice (homozygous inactivated), died at an early stage of embryonic development probably due to a proliferation defect [99]. However, mice with inactivated

C–MYB died in utero at a later stage of embryonic development due to a failure in the fetal hematopoiesis (specifically due to severe anemia) [100], while mice with inactivated *A–MYB* were viable but males were unfertile and females had problems with the development of mammary glands [97].

Several studies have reported several *MYBL2* target genes such as *BCL–2*, *CLU* (*apolipoprotein–J)*, *CDK1 (CDC2)*, *cyclins–D1 (CCND1), –A1 (CCNA1) and –B1 (CCNB)*, suggesting a role of *MYBL2* in cell cycle and apoptosis control [101, 102]. *MYBL2* is transcriptionally regulated by *E2F1* in a cell cycle-dependent manner [103]. Cell cycle expression of *MYBL2* is induced at G1/S transition. In this phase, *MYBL2* is phosphorylated by cyclin A–E/*CDK2* and this modification releases corepressors that block *MYBL2*, promoting cell proliferation (Figure 16A) [95]. Additionally, *MYBL2* regulates the expression of genes involved in G2/M phase, showing *MYBL2* also contributes to G2/M transition, and maintenance of genomic stability [104–111]. The hypothetical model of G2/M genes regulation is based on a *MYBL2* repressive complex, which includes *MYBL2*, *E2F* and the retinoblastoma family member *RBL1* (retinoblastoma-like protein 1, also called *p107*), that is transformed in a transcriptionally active complex in late G1/S by *RBL1* dissociation (Figure 16B) [95].

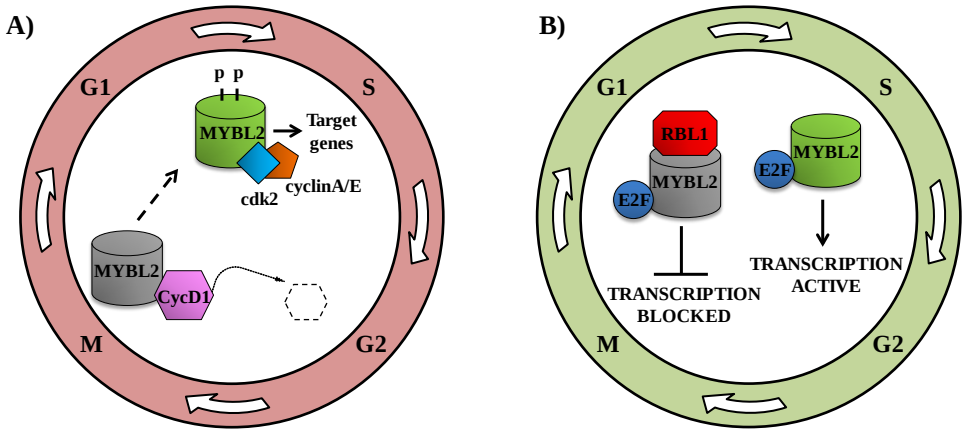

Figure 16. Hypothetical regulation model of G1/S (A) and G2/M (B) genes by *MYBL2*.

MYBL2 has recently suggested to play a role in attenuating senescence [112, 113]. Cellular senescence, originally described as a form of cell aging, is a stable cell cycle arrest caused by insults including telomere erosion, oncogene activation, irradiation,

DNA damage, oxidative stress, viral infection, and toxins. Senescence is thought to have evolved as an example of antagonistic pleiotropic, as it acts as a tumor suppressor mechanism during the reproductive stage, but can promote organismal aging by disrupting tissue renewal, repair, and regeneration later in life. The mechanisms underlying the senescence growth arrest are broadly considered to involve *CDKN2A* (*cyclin-dependent kinase inhibitor 2A* or *p16-INK4A*), *RB1* (*retinoblastoma 1* gene), *TP53* (*tumor protein p53*) and *CDKN1A* (*cyclin-dependent kinase inhibitor 1A*, *p21* or *Cip1*) tumor suppressor pathways; but it is not known what makes the senescence arrest stable and what the critical downstream targets are, as they are likely to be key to the establishment and maintenance of the senescent state. Mowla et al supported a strong evidence to indicate that *MYBL2* is inhibited during senescence, suggesting *MYBL2* impedes cellular senescence [113]. However, it remains to be determined how *MYBL2* blocks senescence and how it is integrated into senescence-inducing pathways.

MYBL2 has also been involved in the regulation of apoptosis. *MYBL2* stimulates transcription of *BCL–2* and *CLU* (*Apolipoprotein J*) [antiapoptotic proteins] promoting cell survival [95]. Mowla et al suggested *MYBL2* could inhibit apoptosis due to an increase in resistance to DNA damage in *MYBL2* expressing cells [113].

4.2. *MYBL2* gene and its association with cancer

MYBL2 has been implicated in different types of human cancer. *MYBL2* amplification or overexpression has been observed in cancers such as neuroblastoma, liver carcinoma, cutaneous T lymphoma, prostate, testicular or ovarian malignancies, promoting cancer cell survival and proliferation [93, 95, 114–124]. Nakajima et al showed that *MYBL2* overexpression in hepatocellular carcinoma (HCC) causes deregulation of apoptosis and cell cycle progression. Also, they observed a clear correlation between *MYBL2* expression and the E2F1 protein, which is dramatically overexpressed in HCC and has been reported to target *MYBL2* [125]. In neuroblastoma, *MYBL2* overexpression is thought to be associated with a poor outcome of the disease and also was reported to be necessary for survival and differentiation of neuroblastoma cells [114, 122, 126]. Moreover, *MYBL2* expression has been shown to be notably increased in metastatic compared to localized prostate tumours [93] and it is currently used as a marker for poor prognosis in breast cancer (Oncotype-DX assay) [123]. Furthermore, *MYBL2* overexpression in cell

lines alters the differentiation of myeloid precursors [127]. Finally, recent studies have shown *MYBL2* overexpression is related to an adverse prognosis in AML patients [128].

The aforementioned studies suggest that *MYBL2* is a proto-oncogene as its aberrant expression would support tumor development, by promoting cell cycle progression and/or resistance to apoptosis. However, other studies demonstrate that *MYBL2* is also essential in G2/M transition and maintenance of genomic stability [104, 105]. García et al. showed, in megakaryoblast cell lines, that *MYBL2* is essential for cell cycle progression into S phase as well as in genomic stability. The reduction in *MYBL2* expression represents a decrease in S phase progression, arrest in mitosis, chromosomal aberrant condensation and chromosomal fragmentation [105]. Indeed, *MYBL2* haploinsufficiency represents a predisposition for multiple myeloid disorders, such as myeloid leukemia in aging mice, suggesting *MYBL2* acts as a stabilizing factor to prevent the accumulation of random secondary mutations and genomic damage [129]. In addition, a recent study showed *MYBL2* is expressed at sharply reduced levels in $CD34^+$ cells from most MDS cases, suggesting downregulation of MYBL2 activity underlies the clonal expansion of abnormal HP in a large fraction of human myeloid malignancies. Also, this study demonstrated mice developed a clonal myeloproliferative/myelodysplastic disorder originating from cells with aberrantly reduced *MYBL2* expression [130]. Therefore, it seems that appropriate *MYBL2* levels are necessary to maintain cellular function since abnormal levels (too low or high) could lead to neoplastic development.

4.2.1 *MYBL2* genetic alterations and its association with cancer: S427G polymorphism

Gene polymorphisms and mutations can induce changes in protein structure and conformation that often may deactivate protein functions. Several studies have investigated possible genetic alterations that may change MYBL2 activity and its association to the development of cancer. The most common non-synonymous polymorphism (allelic fraction rank: 0.17 to 0.4) of *MYBL2* coding region is the S427G change (rs2070235). This residue is highly conserved in different species and one study in vitro suggested S427G variation induces a conformational change in the protein that partially inactivates it, causing a decrease in MYBL2 transactivating activity and in resistance to programmed cell death [131]. Furthermore, the S427G variant has been

associated with risk of developing cancer in several studies, however, it has displayed some controversy in its results. Firstly, S427G was linked to a decrease in overall cancer risk for neuroblastomas, chronic myelogenous leukemia and colon cancers in a combined dataset of cases and controls, where the polymorphism was significantly less frequent in cancer patients than normal individuals, suggesting a protective role for the presence of S427G [131]. In contrast, another population case-control study, have identified a statistically significant association between the S427G and occurrence of basal-like breast cancer subtype, suggesting the polymorphism predisposes to cancer risk [132]. Nonetheless, the same study and another independent case-control study demonstrated no association of S427G with cancer risk in a heterogeneous group of breast malignancies [132, 133].

II. MATERIALS AND METHODS

1. Patient samples, cell lines and healthy controls

1.1 AML patients

The whole of 302 AML patients included in the studies were diagnosed between 1998 and 2009 at the *Hospital Universitario y Politécnico La Fe* and were referred to the Molecular Biology Laboratory for molecular characterization. All patients were classified according to the FAB and 2008 WHO criteria. The limiting criterion for inclusion was the availability of DNA or RNA. The Ethics Committee for Clinical Research of the *Hospital Universitario y Politécnico La Fe* approved the different studies and informed consent was obtained from all participants in accordance with the recommendations of the Declaration of Human Rights, the Conference of Helsinki, and other institutional regulations.

Eligible non-APL patients were treated with intensive chemotherapy in which induction consisted of an anthracycline plus cytarabine combination with or without etoposide according to the PETHEMA LMA 99 and 2007 trial protocols (*http://www.pethema.org*). On achievement of complete remission (CR), patients proceeded to consolidation therapy and eligible cases were selected for autologous or allogeneic stem cell transplantation. Eligible APL patients were treated according to PETHEMA LPA 99 and LPA 2005 protocols (*http://www.pethema.org*).

The main characteristics of the different series of AML patients used in the studies are detailed in the following subsections.

1.1.1 Patient selection for the validation of the novel real-time RT–PCR assay

The main characteristics of 105 AML patients included in the validation of the novel qRT–PCR assay are shown in Table 3.

Table 3. Main characteristics of the AML patient group used for the novel real-time RT–PCR assay validation

Characteristics	n (%)
All patients	105
Sex	
Male	56 (53)
Female	49 (47)
Age groups (years)	
> 60	37 (35)
16 ≥ 60	48 (46)
< 16	20 (19)
WHO classification	
AML with recurrent genetic abnormalities	46 (44)
AML with t(8;21)(q22;q22); *RUNX1–RUNX1T1*	6 (6)
AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB–MYH11*	9 (9)
APL with t(15;17)(q24;q12); *PML–RARA*	20 (19)
AML with t(9;11)(p22;q23); *KMT2A–MLLT3*	3 (3)
AML with t(6;9)(p22;q34); *DEK–NUP214*	1 (1)
AML with t(10;11)(p12;q23); *KMT2A–MLLT10	6 (6)
AML with t(6;11)(p27;q23); *KMT2A–MLLT4	1 (1)
AML not otherwise specified	33 (31)
AML with myelodysplasia-related changes	21 (20)
AML therapy-related	5 (5)

**Translocations not classified as "recurrent" by WHO classification but also recommended for identification at diagnosis.*

1.1.2 Patient selection for the S427G polymorphism case-control study and the analysis of *MYBL2* genetic variants in AML

The study of S427G polymorphism and *MYBL2* genetic variants was carried out on a total of 197 adult patients with de novo AML (the main characteristics of these patients are shown in Table 4).

Table 4. Main characteristics of the AML patient group used to study the S427G polymorphism and genetic variants of *MYBL2* gene.

AML patients				
Characteristic	**Median (range)**	**No. (%)**	**Characteristic**	**No. (%)**
Overall		197 (100)	**FAB**	
Age, years	60 (16-91)			
≤ 60		100 (38)	**M0**	11 (6)
> 60		97 (62)	**M1**	44 (23)
Gender			**M2**	43 (22)
Male		113 (62)	**M3**	15 (8)
Female		84 (38)	**M4**	38 (20)
WBC count, $\times 10^9$/L	9.9 (0.4-396)		**M5**	23 (12)
≤ 50		149 (78)	**M6**	15 (8)
> 50		42 (22)	**M7**	2 (1)
Missing data		6 (3)	**Missing data**	6 (2)
Blasts in PB, %	39 (0-100)		**Molecular Markers****	
≤ 70		66 (67)	**FLT3–ITD*****	
> 70		32 (33)	**Yes**	24 (26)
Missing data		99 (51)	**No**	68 (74)
Cytogenetics*			**Missing data**	2 (8)
Favorable		29 (19)	**NPM1*****	
Intermediate		94 (60)	**Yes**	32 (38)
Adverse		33 (21)	**No**	52 (62)
Missing data		41 (20**)**	**Missing data**	10 (11)

* Cytogenetic risk stratification of patients was established according to the refined Medical Research Council criteria [134].
**Only for intermediate-risk cytogenetic group.
*** *NPM1* and *FLT3−ITD* mutations were detected as described previously [135, 136].

1.2 Controls

1.2.1 Leukemic samples, AML cell lines and plasmids

Leukemic cell samples from AML patients, AML cell lines and plasmids carrying specific rearrangements were used to optimize the novel qRT–PCR assay. The AML cell line THP–1 (from the *European Collection of Cell Cultures*) was used as positive control for the *KMT2A–MLLT3* rearrangement and Kasumi-1 cell line (from *Leibniz Institute DSMZ-German Collection of Cell Cultures*), was used as positive control for the *RUNX1–RUNX1T1* rearrangement. Plasmids containing the fusion genes *KMT2A–MLLT4* (Ipsogen, Marseille, France), *KMT2A–ELL* (Ipsogen) and *RBM15–MKL1* (kindly provided by Thomas Mercher, Hospital Necker Enfants Malades, Paris, France) were used as positive controls.

PB from healthy donors was collected to be used as negative controls.

1.2.2 Control population selected for the S427G polymorphism case-control study

179 healthy Caucasian donors from Spain were selected as the control group for the S427G polymorphism study between 2010-2012, based on the criteria of no previous oncological, hematological or chronic diseases and no chronic consumption of alcohol or tobacco. The median age was 59 years old (range=15-87) and the group was represented by 85 (47.5%) males and 94 (52.5%) females.

1.2.3 Primary $CD34^+$ hematopoietic stem cells

$CD34^+$ HP, obtained from human umbilical cord blood of healthy donors, were used for *MYBL2 in vitro* functional studies. One half of total $CD34^+$ cells were transfected with a *MYBL2* over-expression plasmid (*MYBL2*–IRES–EGFP) [105] or with the corresponding control plasmid (IRES–EGFP). Comparatively, the other half of total $CD34^+$ cells were transfected with specific siRNAs of *MYBL2*, that induce *MYBL2* lower-expression, or with the corresponding siRNAs control.

2. Cellular isolation

2.1. Cellular isolation from AML patients, healthy donors and AML cell lines

Leukemic cells obtained from BM of AML patients or leukocytes from PB of healthy donors, were collected after erythrocyte lysis using a buffer containing: 0.155 M NH_4Cl, 10 mM $KHCO_3$, 0.1 mM Na.EDTA [pH 7.4]. The cell lysates were resuspended in saline serum (for DNA extraction) or in MagNA Pure LC mRNA Lysis Buffer (for mRNA extraction) [Roche Diagnostics GmbH, Mannheim, Germany], and stored at –40 °C. AML cell lines were cultured in RPMI-1640 medium containing 20% fetal calf serum and 1% penicillin and streptomycin. The cells were then collected by centrifugation and preserved at –40 ºC in MagNA Pure LC mRNA Lysis Buffer for mRNA extraction.

2.2. CD34$^+$ hematopoietic stem cell isolation

CD34$^+$ HP were obtained from human umbilical cord blood of 64 healthy donors. CD34$^+$ cells were selected by a MACS separator and the CD34$^+$ MicroBead Kit (MACS, Miltenyi Biotec), from previously isolated mononuclear cells (MC) using Lymphoprep (Axis-Shield, Oslo).

2.2.1 Erythrocyte depletion and separation of mononuclear cells from cord blood samples

MCs were isolated by centrifugation on an isosmotic medium with a density of 1.077 g/mL as Lymphoprep (Axis-Shield) [Figure 17], including a previous erythrocyte depletion step using Hydroxyethyl Starch 6 % (HES) solution (Grifols).

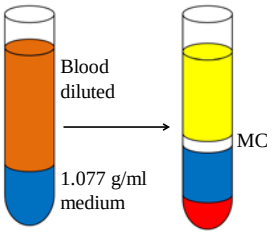

Figure 17. Density gradient centrifugation by Lymphoprep (Axis-Shield) medium.

Firstly, cord blood was erythrocyte depleted by diluting samples with HES (3:1) and centrifuging 45 x g, 20 minutes at 10 ºC. Secondly, the 60% upper supernatant was collected and diluted 1:1 with a buffer containing isotonic phosphate-buffered saline (PBS 0.01 M, pH=7.4) and EDTA 2 mM. The diluted supernatant was added carefully to tubes containing 15 mL of Lymphoprep (3:1) and MCs were isolated by density gradient centrifugation (400 x g, 35 minutes at 20 ºC). The interface containing the MC was separated, 1:1 diluted and washed twice with the buffer, determining the total cell number before the last washing step.

2.2.2 CD34$^+$ magnetic separation by a Macs separator

The pellet of MC was resuspended in 300 µL of the PBS 0.01 M (pH=7.4) with EDTA 2 mM buffer for up to 10^8 total cells. Then, 100 µL of each FcR Blocking and

CD34 magnetic MicroBeads reagents (Miltenyi Biotec) were added to the cellular solution for up to 10^8 total cells, mixing well and incubating for 30 minutes at 4 °C. Next up to 10^8 cells were washed by adding 5−10 mL of the PBS-EDTA buffer and centrifuged at 300 x g for 10 minutes, removing completely the supernatant. Finally, up to 10^8 cells were resuspended in 500 µL of the PBS-EDTA buffer and were loaded onto a MACS® Column, placed in the magnetic field of a MACS Separator (Miltenyi Biotec). The magnetically labeled $CD34^+$ cells were retained within the column and after removing the magnetic field, the $CD34^+$ cells were eluted as the positively selected cell fraction.

Finally, after evaluating the purity by flow cytometry, only the isolated samples containing more than 85% $CD34^+$ cells were cryopreserved with fetal bovine serum and 10% dimethyl sulfoxide in liquid nitrogen, until the time of transfection.

3. DNA, mRNA and total RNA extraction

Routinely DNA and mRNA extractions were performed by automatic standardized techniques in the MagNA Pure LC System (Roche), using the MagNA Pure LC large volume DNA Isolation Kit or the MagNA Pure LC mRNA HS Kit (Roche) respectively, according to the manufacturer's instructions.

Total RNA from transfected $CD34^+$ cells was required to be isolated by conventional methods based on acid guanidinium thiocyanate-phenol-chloroform extraction [137], followed by purification and concentration with the Rneasy MinElute Cleanup kit (Qiagen), to perform expression microarrays. Firstly, the cell pellet obtained from transfected $CD34^+$ cells was resuspended in TRIzol (Invitrogen, Carlsbad, CA, USA), homogenized by vortex and stored at –80 ºC until the time of RNA extraction. After thawing on ice, the cell lysates preserved in TRIzol (Invitrogen) were shaken briefly by a vortex mixer and incubated at room temperature (15-25 °C) for 5 minutes. Next 0.2 mL of chloroform per 1 mL of TRIzol (Invitrogen) was added, vigorously mixed by vortex and centrifuged at 12,000 g for 25 min at 4 °C. The upper (aqueous) phase was transferred to a new tube and 0.5 mL of isopropanol per 1 mL of TRIzol (Invitrogen) was added to precipitate the RNA. In addition, 1 µL glycogen (Glycogen for Molecular Biology 20 mg/mL, Roche Diagnostics) was added to the solution that

was thoroughly mixed by pipetting and incubated 24 h at –80 ºC, to improve the yield of RNA precipitation. After thawing on ice, the solution was again centrifuged for 25 min at 12,000 g and 4 ° C. The supernatant was removed and the pellet of RNA was dried in a vacuum pump for 10 min and resuspended in 100 µL of RNase-free water. The next purification and concentration phases were performed with the Rneasy MinElute Cleanup columns (Qiagen), according to the manufacturer's instructions. 10 µL 2-mercaptoethanol per 1 mL buffer RLT (Qiagen) was added at the beginning of the purification phase to improve the maintenance of RNA integrity. Finally, the purified RNA was eluted in RNase-free water.

DNA, mRNA and total RNA purity and concentration were evaluated spectrophotometrically using NanoDrop ND-2000 (ThermoFisher) and immediately nucleic acids were stored at –80 ºC. In addition, quality and integrity of total RNA were assessed by microfluidics-based platform Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara), before processing expression microarrays, in the Array Service of the Genomic Unit at the *IIS La Fe*.

4. Reverse transcription

The mRNA was reverse transcribed to cDNA using Taqman® Gold Reverse Transcription Reagents (N808-0234, PE Applied Biosystems, NJ): six µL mRNA was added in a total reaction volume of 25 µL containing 5 µM random primers, 5.5 mM $MgCl_2$, 500 µM each deoxynucleotide (dNTP), 0.4 U/µL RNase inhibitor and 1.25 U/µL reverse transcriptase, according to the manufacturer protocol. The reverse transcription reaction was incubated at 25 °C for 10 min to facilitate random primers annealing to the mRNA, then was maintained at 48 ° C for 30 minutes to allow reverse transcriptase synthesize cDNA, and finally the enzyme was inactivated by heating at 95 °C for 5 min.

The cleansed total of RNA, isolated from transfected $CD34^+$ cells, was transcribed into cDNA with integrated removal of genomic DNA contamination using the QuantiTect Reverse Transcription Kit (Qiagen). Quantiscript Reverse Transcriptase has a high affinity for RNA and is optimized for efficient and sensitive cDNA synthesis from minimal amounts of RNA (from 10 pg to 1 µg of RNA), enabling high cDNA

yields for real time PCR analysis. 12 µL total RNA was briefly incubated in 2 µL gDNA Wipeout Buffer at 42°C for 2 minutes to effectively remove contaminating genomic DNA. Then, the RNA was reverse transcribed by mixing the RNA solution with a master mix prepared from 1 µL Quantiscript Reverse Transcriptase (also contained RNase inhibitor), 4 µL Quantiscript RT Buffer (included Mg^{2+} and dNTPs) and 1 µL RT Primer Mix, in a total reaction volume of 20 µL. The entire reaction took 15 minutes at 42°C and finally was inactivated at 95°C for 3 minutes.

5. Real-time RT–PCR assay for molecular rearrangements detection

The qRT–PCR developed to detect AML rearrangements was performed by real-time PCR using the SYBR Green technique. SYBR Green I is a DNA double-strand-specific dye that binds to the amplified PCR products and the amplicon can be detected by its fluorescence.

5.1. Primer selection

Primers were adapted from published methods or designed using Primer3 software (http://primer3.wi.mit.edu). The oligonucleotide primers were synthesized by TIB MOLBIOL GmbH (Berlin, Germany) and purified via gel filtration.

Several primer sets were assayed for the amplification of each rearrangement. The 16 selected primer sets were then lyophilized in individual wells of a 96-well PCR plate to detect the rearrangements by performing 16 individual reactions at once (Table 5).

Table 5. Primers selected to amplify rearrangements with the real time RT–PCR method

Rearrangement	Name	Sequence
RUNX1–RUNX1T1	*RUNX1*	5´–TGGCTGGCAATGATGAAAAC–3´
	RUNX1T1	5´–CGTTGTCGGTGTAAATGAACTG–3´
RUNX1–MECOM1	*RUNX1*	5´–TGGCTGGCAATGATGAAAAC–3´
	MECOM	5´–CCCCAGGCATATTTGACTCTC–3´
BCR–ABL1	*BCRe1*	5´–ACTGCCCGGTTGTCGTGTC–3´
	BCRb2b3	5´–CACGTTCCTGATCTCCTCTGAC–3´
	ABL1	5´–ACACCATTCCCCATTGTGATTAT–3´
CBFB–MYH11	*CBFB*	5´–TGGGCTGTCTGGAGTTTGATG–3´
	MYH11a	5´–TGAGCGCCTGCATGTTGAC–3´
	MYH11b	5´–TCCTCGTCCAGCTGGTCTTG–3´
DEK–NUP214	*DEK*	5´–AGCAGCACCACCAAGAAGAAT–3´
	NUP214	5´–GTCTCTCGCTCTGGCACAAG–3´
KMT2A–MLLT4	*KMT2A1*	5´–GGACCGCCAAGAAAAGAAGT–3´
	KMT2A2	5´–AGCAGATGGAGTCCACAGGATCAG–3´
	MLLT4	5´–GAGGACAGCATTCGCATATCAG–3´
KMT2A–ELL	*KMT2A1*	5´–GGACCGCCAAGAAAAGAAGT–3´
	KMT2A2	5´–AGCAGATGGAGTCCACAGGATCAG–3´
	ELL	5´–GCCGATGTTGGAGAGGTAGA–3´
KMT2A–MLLT1	*KMT2A1*	5´–GGACCGCCAAGAAAAGAAGT–3´
	KMT2A2	5´–AGCAGATGGAGTCCACAGGATCAG–3´
	MLLT1	5´–GCGATGCCCCAGCTCTAAC–3´
KMT2A–MLLT10	*KMT2A1*	5´–GGACCGCCAAGAAAAGAAGT–3´
	KMT2A2	5´–AGCAGATGGAGTCCACAGGATCAG–3´
	MLLT10a	5´–AACTGCTGTTGCCTGGTTGAT–3´
	MLLT10b	5´–TTCCACTAGAGGTGTGTGCAGAG–3´
	MLLT10c	5´–GGCAAACTGAGCGCATGTTAC–3´
KMT2A–MLLT3	*KMT2A1*	5´–GGACCGCCAAGAAAAGAAGT–3´
	KMT2A2	5´–AGCAGATGGAGTCCACAGGATCAG–3´
	MLLT3a	5´–GAGCAAAGATCAAAATCAAATGTT–3´
	MLLT3b	5´–CTCCATTTCAGAGTCATTGTCGTTAT–3´
PML–RARA	*PML1*	5´–ACCTGGATGGACCGCCTAG–3´
	PML2	5´–AGCTGCTGGAGGCTGTGGACGCGCGGTACC–3´
	RARA	5´–CAGAACTGCTGCTCTGGGTCTCAAT–3´
ZBTB16–RARA	*ZBTB16*	5´–GTGGGCATGAAGTCAGAGAGC–3´
	RARA	5´–CAGAACTGCTGCTCTGGGTCTCAAT–3´
KAT6A–CREBBP	*KAT6A*	5´–GAGGCCAATGCCAAGATTAGAAC–3´
	CREBBP	5´–AGCTTGACTAAAGGGCTGTC–3´
RBM15–MKL1	*RBM15*	5´–AGCAGTTCCTGGATTCCCCT–3´
	MKL1	5´–ATGAAATGCGGCTGGACTTTT–3´
FUS–ERG	*FUS*	5´–CAGCGGTGGCTATGGACAG–3´
	ERG	5´–GGTGCCTTCCCAGGTGATG–3´
GUSB (control gene)	*ENFF*	5´–GAAAATATGTGGTGGGAGAGCTCATT–3´
	ENFR	5´–CCGAGTGAAGATCCCCTTTTTA–3´

The distribution of the lyophilized primer sets in the plate was designed to enable simultaneous analysis of 15 rearrangements and 1 control gene in six patients, using two columns of the 96-well PCR plate per patient (Figure 18).

	1	2	3	4	5	6	7	8	9	10	11	12
A	RUNX1-RUNX1T1	MLL-MLLT1	RUNX1-RUNX1T1	MLL-MLLT1	RUNX1-RUNX1T1	MLL-MLLT1	RUNX1-RUNX1T1	MLL-MLLT1	RUNX1-RUNX1T1	MLL-MLLT1	RUNX1-RUNX1T1	MLL-MLLT1
B	CBFB-MYH11	PML-RARA	CBFB-MYH11	PML-RARA	CBFB-MYH11	PML-RARA	CBFB-MYH11	PML-RARA	CBFB-MYH11	PML-RARA	CBFB-MYH11	PML-RARA
C	BCR-ABL1	RUNX1-MECOM	BCR-ABL1	RUNX1-MECOM	BCR-ABL1	RUNX1-MECOM	BCR-ABL1	RUNX1-MECOM	BCR-ABL1	RUNX1-MECOM	BCR-ABL1	RUNX1-MECOM
D	DEK-NUP214	KAT6A-CREBBP	DEK-NUP214	KAT6A-CREBBP	DEK-NUP214	KAT6A-CREBBP	DEK-NUP214	KAT6A-CREBBP	DEK-NUP214	KAT6A-CREBBP	DEK-NUP214	KAT6A-CREBBP
E	MLL-MLLT10	ZBTB16–RARA	MLL-MLLT10	ZBTB16–RARA	MLL-MLLT10	ZBTB16–RARA	MLL-MLLT10	ZBTB16–RARA	MLL-MLLT10	ZBTB16–RARA	MLL-MLLT10	ZBTB16–RARA
F	MLL-MLLT3	RBM15-MKL1	MLL-MLLT3	RBM15-MKL1	MLL-MLLT3	RBM15-MKL1	MLL-MLLT3	RBM15-MKL1	MLL-MLLT3	RBM15-MKL1	MLL-MLLT3	RBM15-MKL1
G	MLL-MLLT4	FUS-ERG	MLL-MLLT4	FUS-ERG	MLL-MLLT4	FUS-ERG	MLL-MLLT4	FUS-ERG	MLL-MLLT4	FUS-ERG	MLL-MLLT4	FUS-ERG
H	MLL-ELL	GUSB	MLL-ELL	GUSB	MLL-ELL	GUSB	MLL-ELL	GUSB	MLL-ELL	GUSB	MLL-ELL	GUSB

Figure 18. Rearrangement distribution in a PCR plate of 96 wells.

5.2. Reaction mix and PCR program

Two µL cDNA was amplified with 1 × SYBR Green I Master Mix and 1 U/µL uracil-DNA N-glycosylase (UDG) in a final reaction volume of 10 µL on the LightCycler® 480 real-time PCR System (Roche).

Different PCR programs were used, but all included an UDG incubation step at 40 °C for 5 min, a denaturing step at 95 °C for 8 min, and 45 amplification cycles. The annealing temperatures were assayed from 58 °C to 64 °C to optimize the amplification. One melting cycle was then tested, reducing the temperature from 95 °C to 55 °C for 10 s, followed by incremental increases of 0.07 °C up to 99 °C, with eight acquisitions read per °C, and finishing with a cooling stage at 40 °C.

6. Hybridization probes-based real-time PCR for S427G polymorphism genotyping

Genotyping of the S427G polymorphism of *MYBL2* was performed by real-time PCR using specific FRET probes. FRET probes consist of a pair of oligonucleotides (anchor and sensor), each labeled with a different fluorescent dye. The anchor probe is

labeled with a fluorescent dye at the 5' end and the sensor probe with fluorescein at the 3' end. Interaction of the two dyes can only occur when both probes are bound to their target.

6.1. Probe and primer selection

The anchor probe (640–5´–GTGACAGGGGACAGAGCCACACGC–3´–Ph) was designed to bind to the wild-type DNA and the sensor probe (5´–GAAGGACAGACTGGTGCCATTCTC–3´–FL) to the mutant DNA sequence. Forward and reverse primers (5´–CGAGTACCGCCTGGATG–3´ and 5´–ATCTCAGACCTGAGCCGGA–3´ respectively) were designed using Primer3 software (http://primer3.wi.mit.edu) to flank the region of the polymorphism. The primers and hybridization probes were synthesized by TIB-MOLBIOL (Berlin, Germany) and purified by gel filtration or High-performance liquid chromatography (HPLC) respectively.

6.2. Reaction mix and PCR program

PCR was performed in a reaction volume of 10 µL with 0.3 µM each primer, 0.2 µM each probe, 40 ng of DNA and 2 µL of 5 × LightCycler® 480 Genotyping Master Mix (Roche Applied Science, Mannheim, Germany), in the LightCycler® 480 PCR System (Roche). The reaction mix was pre-incubated at 95 ºC for 10 min and then subjected to 35 amplification cycles (denaturing at 95 ºC for 5 s, annealing at 58 ºC for 25 s and extension at 72 ºC for 15 s). The final melting program consisted of three steps beginning with denaturing at 95º C for 1 min, renaturing at 48º C for 30 s and slowly heating to 78ºC at 0.11 ºC/s with five acquisitions per °C, followed by a cooling stage at 40 °C.

7. High Resolution Melting for screening of *MYBL2* genetic alterations

High resolution melting (HRM) is based on the saturating DNA dye ResoLight, an intercalant dye which binds specifically in high amount (saturating dye) to double-stranded DNA and fluoresces brightly. HRM allows the amplification and detection of

single base modifications in PCR amplicons for mutation scanning and genotyping applications.

7.1. Primer selection

MYBL2 coding sequence and adjacent intronic regions were analyzed by designing 16 specific primer pairs (Primer 3 v.4.0) to amplify overlapping regions of 150-250 bp in size (Table 6).

Table 6. Specific primers designed to amplify *MYBL2* coding sequence.

Exon of *MYBL2*	Name	Sequence
1	F1_19	5´–GGGAGATAGAAAAGTGCTTCAACC–3´
	R1_S	5´–GAGGGGTGAGTTAAAGGGAGG–3´
2	F2_38	5´–CACACCATCCTTGACCCTTGG–3´
	R2_16	5´–GGCTCAGTTCCTCTGGACAGC–3´
3	F3_1	5´–TCTGTCACGGGCAGCCCTGAG–3´
	R3_5	5´–GGCAAAGTCTGTCCCAGGCAC–3´
4	F4_8	5´–CTACCCAAGGTGTGCCTAGTC–3´
	R4_1	5´–ACTGTGTGCAAGGCACTGTCC–3´
5	F5_15	5´–CCGGTGCAAGTGGCTACCCAG–3´
	R5_91	5´–GAACCTGTCCCAACCCCCAAG–3´
6	F6_6	5´–CTGATGGCACCCCACCCACAG–3´
	R6_15	5´–CCAAGCCAAAGGCTGAGCAGC–3´
7 (first amplicon)	F71_48	5´–CTCAGCGAAATGCAAATGGTG–3´
	R71_15	5´–GTAAGGCAGGCTCGTTTCTGG–3´
7 (second amplicon)	F72_3	5´–TCTGGACGCAGTGCGAACACC–3´
	R72_8	5´–GGGTGTTCTGTGCTCACTGTG–3´
8 (first amplicon)	F81_11S	5´–GAGTCCTCTTGTATTAAGAGC–3´
	R81_11S	5´–CGGTGTGCCAATGCCAGAG–3´
8 (second amplicon)	F82_11S	5´–ACACCATCTCAGACCTGAGCC–3´
	R82_11S	5´–CCAGCACTGAACACTAGGTGTC–3´
9	F9_23	5´–TCAGGGATACTCATGCAGGTC–3´
	R9_7	5´–CATGGTCTGGTGTTGCTGAGG–3´
10	F10_7	5´–TGGGACGAGAACCTGTGCTGG–3´
	R10_5	5´–CCCACTGTGCAGAGATCCACG–3´
11	F11_7	5´–CGAGTGTGGTGCCTCACGTGG–3´
	R11_S	5´–CAGTTTCCTCAACAACCACCG–3´
12	F12_4	5´–CCATGGGGAATCCAGACACTC–3´
	R12_7	5´–AACTCCACCTTGAGGGTCCTC–3´
13	F13_39	5´–TCACAGCTTCTCGCAAAATGG–3´
	R13_68	5´–AGAGATGGCCACAGCACTCAC–3´
14	F_4	5´–ACCAGGGTCTGTTGGGAACAC–3´
	R_26	5´–ACATGAGAATGGGCTCGTGAC–3´

7.2. Reaction mix and PCR program

Real-time PCR was performed on a LightCycler®480 (Roche) using the HRM method. The overlapping amplicons were amplified in a reaction volume of 10 µL with 0.15 to 0.3 µM of each primer, 1.5 to 3 mM $MgCl_2$, 60 ng of DNA and 5 µL of the 5x LightCycler® 480 HRM Master Mix (Roche). The PCR reaction was pre-incubated at 95 ºC for 10 min and then subjected to 45 amplification cycles (denaturation at 95 ºC for 10 s, annealing at 56 ºC for 15 s and extension at 72 ºC for 15 s). The subsequent HRM program involved four main steps, beginning with a denaturation at 95º C for 1 min, followed by renaturing at 40º C for 30 s, heating to 75ºC at 1 ºC/s and slowly heating to 99ºC at 0.02 ºC/s, with 25 acquisitions read per °C, and the HRM programme terminated with a cooling step at 40 °C. The melting curves were analysed with the LightCycler ® 480 Gene Scanning Software Version 1.0 (Roche).

8. Real-time PCR for quantifying gene expression

TaqMan probes are based on hydrolysis probes that are designed to increase the specificity of quantitative real-time PCR and are widely used for gene expression assays. The TaqMan probe consists of an oligonucleotide with a 5´reporter dye and a 3´quencher dye. During the PCR reaction, a DNA polymerase with 5´ to 3´ nuclease activity (but lacking 3´ to 5´ exonuclease activity) cleaves the TaqMan probe separating the reporter dye and the quencher dye, which results in increased fluorescence of the reporter.

8.1. Quantification of *MYBL2* expression

The *MYBL2* expression was quantified by using 1× Taqman gene expression assay HS00231158-m1 (Applied Biosystems, Foster City, CA), 1× Taqman Universal Master Mix (Applied Biosystems) and 2 µL of cDNA in a reaction volume of 25 µL. All assays were performed in an ABI Prism 7500 Fast Real Time PCR machine (Applied Biosystems). The PCR program consisted of an initial Taq activation at 95 °C for 15 min followed by 45 cycles of denaturation at 95 °C for 15 s and, finally annealing-extension at 60 °C for 1 min. The *GUSB* expression was quantify using 1× Taqman Universal Master Mix (Applied Biosystems), 2 µL of cDNA, 0.5 µM of each

primer and 0.152 µM of taqman probe (previously described [139]) in a reaction volume of 25 µL (Table 7).

The Delta Ct ($2^{-\text{delta Ct}}$) method [138] was used to evaluate the expression of *MYBL2* in AML patients. All data were normalized using the expression of endogenous *GUSB* control gene [139] and only those data with a Ct value <30 for *GUS* were included in the analysis.

8.2. Validation of gene expression results from high throughput methods

The results obtained from gene expression microarrays were validated in triplicate experiments by quantitative real-time PCR using 1×Taqman Universal Master Mix (Applied Biosystems), 2 µL of cDNA and 1× Taqman gene expression assay (HS01083836-m1 to *BCL2L11* gene, HS00231733-m1 to *CREBBP*, HS01078210-m1 to *CREBRF*, HS00178198-m1 to *MAP2K7*, HS00984486-m1 to *MLLT4* gene and HS00231158-m1 to the *MYBL2* gene [Applied Biosystems, Foster City, CA]) in a reaction volume of 25 µL. The *GUSB* expression was quantified by using 1× Taqman Universal Master Mix (Applied Biosystems), 2 µL of cDNA, 0.5 µM of each primer and 0.152 µM of taqman probe (previously described [139] , Table 7) in a reaction volume of 25 µL.

Table 7. Primers and taqman probe selected to quantify *GUSB* expression.

Name	Sequence
ENF–Forward	5´– GAAAATATGTGGTGGGAGAGCTCATT–3´
ENF–Reverse	5´– CCGAGTGAAGATCCCCTTTTTA–3´
ENP	Fam–CCAGCACTCTCGTCGGTGACTGTTCA–Tamra

The $2^{-(\text{delta}-\text{delta Ct})}$ method [138] was used to estimate gene expression in $CD34^+$ cells transfected with specific *MYBL2* modulators (specific siRNAs or overexpression plasmids) relative to gene expression in $CD34^+$ cells transfected with the corresponding control modulators. All data was normalized using the expression of endogenous *GUS* control gene [139] and only those data with a Ct value <35 for *GUS* were included in the evaluation.

All assays were performed in an ABI Prism 7500 Fast Real Time PCR device (Applied Biosystems). The PCR program included an initial Taq activation at 95 °C for 15 min followed by 45 denaturation cycles at 95 °C for 15 s and, finally, annealing-extension at 60 °C for 1 min.

9. Sequencing

The samples identified as having an altered HRM pattern were confirmed by direct sequencing and compared with *MYBL2* (Gene Bank accession number NT_011362.10). The PCR products from each sample were digested with ExoSap-IT (USB Corporation, Cleveland, OH), and fluorescent labelled using the Big Dye Terminator v1.1 Cycle Sequencing Reaction kit (Applied Biosystems, Foster City, CA) and the corresponding primers, following the manufacturer's instructions. The reaction products were purified on Performa DTR Gel Filtration Cartridges (EdgeBio) and bidirectionally sequenced in an ABI Prism 3130 (Applied Biosystems). Sequence traces were analyzed using Applied Biosystems software and reviewed manually.

10. Transfection

The study of the effect that *MYBL2* expression has on proliferation and dysregulation of gene expression was initiated by transfecting healthy $CD34^+$ HP cells with *MYBL2* expression modifiers.

10.1. Cell transformation and plasmidic DNA isolation

Competent cells (α-select Chemically Competent Cells, Bioline, London, UK) were transformed with 65 ng of *MYBL2*–pIRES–EGFP plasmid (a construct that over-express human *MYBL2* cDNA under the control of a cytomegalovirus promoter [105]) or with 65 ng of pIRES–EGFP control plasmid (both plasmids kindly given by Dr. Paloma García, Institute of Biomedical Research, School of Medical and Dental Science, University of Birmingham, Birmingham, UK). After heat shock at 42 °C for 30 s and incubating with nutrient-rich medium SOC, transformed cells grew in a LB agar plate

containing kanamycin antibiotic. Transformed cells from 1-2 colonies were then collected in 250 mL of LB medium 0.2% kanamycin and grew overnight at 37 °C.

The next step was to purify plasmidic DNA by the Plasmid Purification Maxi kit (Qiagen), based on a modified SDS-alkaline lysis procedure, according to the manufacturer's instructions. A small aliquot of purified DNA was digested to check it was plasmidic DNA, using *HindIII* restriction enzyme with *Neb4* buffer, generating two specific plasmidic restriction fragments (4770 bp and 269 bp) that were tested in an agarose gel. Finally, the purified plasmidic DNA was quantified by Nanodrop ND-2000 (ThermoFisher) and stored at –80 ºC.

10.2. Design of *MYBL2* specific siRNAs and siRNAs control

siRNA oligoribonucleotides were designed by Eurogentec (Cultek) with a 5′ fluorescein isothiocyanate (FITC) modification and TT overhangs at the 3′ end, being purified by HPLC (Table 8). These siRNAs were published previously by García and Frampton [105].

Table 8. Pairs of specific *MYBL2* siRNAs and siRNAs control.

Pairs of siRNAs (*MYBL2* specific and control)	Name	Sequence
MYBL2 siRNA-1	MYBL2-292_F	5´–FITC–GUUAAGAAGUAUGGCACAATT–3´
	MYBL2-292_R	5´–UUGUGCCAUACUUCUUAACTT–3´
MYBL2 siRNA-2	MYBL2-25_F	5–´FITC–GAUCUGGAUGAGCUGCACUTT–3´
	MYBL2-25_R	5´–AGUGCAGCUCAUCCAGAUCTT–3´
Control siRNA	Control-F	5′ –FITC–CUUCAGUUCGCGUGACCAATT–3′
	Control-R	5´–UUGGUCACGCGAACUGAAGTT–3´

10.3. Electroporation of K562 cell line

5×10^6 K562 cells (a human CML cell line) [140] were transfected by electroporation with 10 µg of *MYBL2* expression modifiers or 10 µg of the corresponding control (10 µg of *MYBL2*–pIRES–EGFP or control pIRES–EGFP, or 10 µg of specific *MYBL2* siRNAs or control siRNAs), using 400 µL Cytomix buffer (120 mM KCl; 0.15 mM CaCl2; 10 mM K2HPO4/KH2PO4, pH 7.6; 25 mM Hepes, pH 7.6; 2 mM EDTA, pH 7.6; 5mM MgCl2; pH adjusted with KOH) and adding freshly

prepared ATP (2 mM; pH = 7.6 adjusted with KOH) and glutathione (5 mM) shortly before use to prevent leakage of the cytoplasmic components, protect membranes against oxidation and facilitate resealing of pores [141]. The electroporation was performed at 280 V, 950 μF in a 0.4 cm cuvette (Bio-Rad Electroporator). After 24 h of incubation in RPMI 1640 (Gibco) culture medium with 10% fetal bovine serum, 1% penicillin-streptomycin and 1% 2-glutamine at 37 ºC, transfected cells were harvested and sorted on the basis of the green fluorescent protein (GFP) expression or FITC fluorescence using a Cytomation MoFlo FACS machine (Beckman Coulter, Brea, CA, USA). GFP-positive, FITC-positive and the corresponding -negative fractions were collected with a purity of sorting in excess of 99%.

10.4. Nucleofection of CD34$^+$ hematopoietic stem cells

4 million CD34$^+$ cryopreserved cells were thawed at 37 ºC and incubated for one hour in RPMI 1640 (Gibco) culture medium with 10% fetal bovine serum, 1% penicillin-streptomycin and 1% 2-glutamine. Then, one half of the CD34$^+$ culture was transfected with specific modulators of *MYBL2* expression (5 μg of specific *MYBL2* siRNAs or 5 μg of *MYBL2*–IRES–EGFP overexpression plasmid); and the other half, was transfected with the corresponding control (5 μg of siRNAs control or 5 μg of pIRES–EGFP control plasmid).

The transfection was performed on single cuvettes using 100 μL P3 Primary Cell 4D-Nucleofector X solution (Lonza) and specific programs (program EO-100 to transfect siRNAs and ED-100 for plasmids) in an Amaxa Nucleofector® 4D (Lonza).

After transfection, CD34+ cells were incubated in the culture medium supplemented with growth factors (50 ng/mL stem cell factor [SCF], 10 ng/mL interleukin-3 [IL–3] and 20 ng/mL interleukin-6 [IL–6]). After 24 h of transfection, live cells were harvested filtered through a 70-μm strainer, and sorted on the basis of FITC (for siRNAs) or GFP (for plasmids) fluorescence discrimination in a Cytomation XDP MoFlo machine (Beckman Coulter). Viable FITC or GFP positive fraction was collected in PBS 10% FBS and, according to the following experiment to be performed, subsequently was incubated in colony forming cell (CFC) culture medium or was preserved in TRIzol (Invitrogen) at –20 ºC (for RNA extraction and gene expression microarrays).

Each experiment of paired transfections with siRNAs or with plasmids was repeated three times on a different day, both to perform CFC assays as gene expression microarrays.

11. Western blot

MYBL2–IRES–EGFP over-expression plasmid, *IRES–EGFP* control-plasmid, *MYBL2* specific siRNAs and control siRNAs were tested by western blot.

11.1. Protein extraction, SDS-PAGE and transfer

5×10^6 cells were resuspended in 100 μL Lysis Buffer (20 mM Tris-HCl, pH 7.4; 10 mM EDTA; 100 mM NaCl; 1% Triton X-100; containing protease and phosphatase inhibitors) and incubated on ice for 20 min. After incubation, samples were centrifuged 10,600 rcf for 10 minutes at 4 ºC and the supernatant (containing the proteins) were collected and stored at –80 ºC until SDS-PAGE performance.

The protein extract from samples were thawed on ice and incubated for 5 minutes at 95 ºC with Laemmli Sample Buffer 4x (250 mM Tris-HCl [pH 6.8], 8% sodium dodecyl sulfate [SDS], 40% glycerol, 8% β-mercaptoethanol and 0.02% bromophenol blue) [2 μL of Laemmli Sample Buffer 4x per 13 μL of protein extract, from each sample].

The same quantity of protein-Laemmli buffer solution from each sample was subjected to SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE), containing stacking gel at 5% acrylamide and resolving gel at 10% acrylamide, during 1 hour 30 minutes at 100 V. A pre-stained molecular weight marker (Color Plus Pre-stained Broad Range, Byolabs New England) was used to estimate the size of the desired proteins. In the following step, proteins were transferred to BioTrace PVDF membranes (Bio-Rad) for 2 h at 15 V using a semi-dry transfer apparatus (Amersham).

11.2. Blocking and antibody detection

The PVDF membrane was blocked overnight in tris-buffered saline (1× TBS: 50 mM Tris-Cl pH 7.5, 150 mM NaCl) containing 0.1% Tween 20 and 5% dry milk (T-

TBS). The membrane was divided in two halves and each one was incubated at room temperature for 2 h with the appropriate primary antibody diluted in T-TBS. Primary antibodies used were as follows: anti-human *MYBL2* rabbit polyclonal IgG (N–19, sc–724, Santa Cruz Biotechnology) at a 1:500 dilution; anti-human Actin goat polyclonal IgG (I–19, sc–1616, Santa Cruz) at 1:500 dilution. After washing and 1 hour incubation with an appropriate secondary antibody conjugated to horseradish peroxidase (Amersham), diluted in T-TBS at 1:10000, signals were detected using the enhanced chemiluminescence system (Pierce).

12. Gene expression arrays

The dysregulation in gene expression and signaling pathways caused by aberrant *MYBL2* expression on transfected $CD34^+$ was studied by gene expression microarrays.

RNA expression profile was determined using PrimeView Affymetrix platform including all annotated genes based in NCBI Human Genome version 37 (GRCh37). This version contains 48,658 probe sets including 419 UniGene probe sets not covered by RefSeq probe sets. Total RNA isolated from $CD34^+$ cells transfected with *MYBL2* expression or control modulators were processed applying the WTA2 protocol at the Functional Genomics Facilty of IRB (*Barcelona*) [142].

The results obtained from gene expression microarrays were validated in triplicate experiments by RT–PCR.

12.1. Data analysis of gene expression arrays

Statistical Analysis Expression Console free software from Affymetrix was used for quality control and probe set normalized values assessment. Partek Genomic Suite 6.6 version (Partek Incorporated, St. Louis) was used for microarray statistical analysis. Following the RNA expression workflow, normalization of data included RMA background correction, quartile normalization, log2 transformation values and median polish according to Affymetrix and Partek recommendations. ANOVA test at different categories was performed and fold change (FC) and p-values were applied to generate the gene lists of more significant probe values. Hierarchical clustering (HC) was

generated from lists of differentially expressed genes to determine expression patterns from each category. Also Gene Ontology enrichment and Pathway analysis was obtained from differentially expressed genes. All statistical tests included paired analysis to compare gene expression from paired transfections ($CD34^+$ cells transfected with *MYBL2* specific siRNAs compared to $CD34^+$ cells transfected with control siRNAs, and $CD34^+$ cells transfected with *MYBL2*–IRES–EGFP overexpression plasmid compared to $CD34^+$ cells transfected with pIRES–EGFP control plasmid).

13. Colony forming cell assays from transfected $CD34^+$ cells

The effect that aberrant *MYBL2* expression could have on cell proliferation and differentiation potential of $CD34^+$ HP was studied by CFC assays.

Firstly, as described previously, viable and efficiently transfected $CD34^+$ cells were selected by sorting and, subsequently, 500 cells were plated in triplicate with 1.1 mL of Methocult GF H84435 (Stem Cell Technologies) medium. Secondly, counting and classification of colonies was performed after 14 days in culture at 37 ºC with 5% CO_2 and over 95% humidity, using a high-quality inverted microscope (Axiovert Zeiss), equipped with 2X, 4X and 10X planar objectives and stage holder for a 60 mm gridded dish.

14. Statistic analysis

The differences in the distribution of categorical variables were analyzed using the chi-squared test, while the Mann-Whitney's U or Kruskal-Wallis test was used to analyze continuous variables for data that failed the normality test. Survival analysis was restricted to non-APL patients and the end-points analyzed were overall survival rate (OS) and disease or relapse free survival times (DFS or RFS respectively). The OS was measured from the date of diagnosis until date of death, censoring for patients alive at last follow-up. The DFS was calculated from the date of CR until the date of relapse or death from any cause, whichever occurred first, censoring the patients that remained alive in continuous CR. The RFS was calculated from the date of CR until relapse,

censoring the patients that remained alive without relapse. Estimated probabilities for OS, DFS or RFS were calculated by the Kaplan-Meier method [143] and the log-rank test [144] evaluated the differences between the survival distributions. Odds ratios (ORs) and 95% confidence intervals (95% CIs) were calculated by logistic regression. *MYBL2* mRNA expression was transformed into a categorical variable using the median *MYBL2* expression from patients (patients with higher *MYBL2* expression had larger values than the median and patients with lower *MYBL2* expression presented equivalent or smaller values than the median). All p-values reported were two-sided and a p-value <0.05 was considered significant. Statistical analyses were performed using the SPSS v17.0 statistical package (Chicago, IL).

III. HYPOTHESIS

The detection of molecular alterations is important to diagnose and stratify AML, being also of interest to understand the molecular mechanism of the leukemic process. The fact that a large number of fusion genes have been characterized in AML and that the PCR is highly sensitive in detecting some abnormal cells, has led this thesis to perform an optimized RT−PCR method to detect the most relevant molecular rearrangements in AML, which could improve the diagnosis, prognosis and MRD. However, a large number of AML patients lack any of these abnormalities and, therefore, they cannot take advantage of such new strategies. For these reasons and due to the important requirement to stratify those patients with normal karyotype and non-carriers of any known molecular alteration, the second part of this thesis focuses on *MYBL2* studies as a novel and promising biomarker of AML. Based on the results of recent studies and due to the functions assigned to *MYBL2*, it is conceivable that deregulation of this gene could contribute to the development and/or maintenance of the leukemic phenotype as well as to the prognosis of AML. However, it remains to be determined what genetic alterations could deregulate *MYBL2* in AML and what critical downstream targets could be integrated into the mechanism of leukemogenesis. In this regard, the thesis proposes the most frequent S427G polymorphism and somatic mutations of *MYBL2* as potential issues that could contribute to a MYBL2 aberrant activity, promoting the risk and perturbing the prognosis of AML. Finally, the thesis presents the study of the possible trigger mechanism of AML due to the aberrant *MYBL2* expression, bringing it forward as an underlying cause of HP transformation into AML cells.

IV. OBJECTIVES

In order to achieve the aims described previously, the following working plan was designed:

1. Performance and validation of an optimized method based on real-time RT–PCR for the simultaneous detection of 15 specific rearrangements of AML and comparison of the results with those obtained by cytogenetics, FISH and conventional PCR methods.

2. Study of known *MYBL2* genetic alterations and search for new mutations in AML.

2.1 Study of the most common non-synonymous S427G polymorphism of *MYBL2* by analysing the S427G incidence in case-control series and its association with the risk of AML, as well as the associations between the S427G change in AML and prognosis of patients.

2.2 Search for new *MYBL2* genetic variations in AML by analysing the incidence of the genetic variants detected in patients and in network databases. These new *MYBL2* genetic variations will then be studied in order to determine whether they correlate with the risk of AML. Moreover, *MYBL2* genetic variations detected in AML will be correlated to *MYBL2* expression and the prognosis of patients.

3. Study of alterations triggered by aberrant *MYBL2* expression on the development of human HP.

3.1 Identification of genes deregulated due to over-expression or lower-expression of *MYBL2* by gene expression arrays and their validation by PCR. To determine whether there are specific cellular pathways affected due to *MYBL2* deregulation.

3.2 Proliferation/ Differentiation studies of human $CD34^+$ progenitor cells with modified *MYBL2* levels by CFC assays.

V. RESULTS

1. Optimization of real-time RT–PCR

Various primer concentrations and PCR conditions were assayed to simultaneously amplify all of the genetic rearrangements. A lyophilized primer concentration of 0.5 mmol/L and touchdown annealing temperatures decreasing from 64 °C to 60 °C, with gradual decrements of 0.4 °C per cycle, enabled specific detection of all of the rearrangements at the lowest cycle threshold (Ct). The melting temperature (Tm) for positive controls used to optimize the assay are given in Table 9.

Table 9. Tm detected for positive controls.

Chromosomal translocation	Fusion gene	Mean (SD) of T_m in °C
t(8;21)(q22;q22)	*RUNX1 (21q22)* *RUNX1T1 (8q22)*	84.09 (0.76)
inv(16)(p13q22)	*CBFB (16q22)* *MYH11 (16p13)*	85.92 (0.27) (*CBFB–MYH11 type A*)
t(9;22)(q34;q11)	*BCR (22q11)* *ABL1 (9q34)*	87.4 (0.23) (*BCR–ABL1 e1a2*) 84.21 (0.3) (*BCR–ABL1 b2a2/b3a2*)
t(6;9)(p22.3;q34)	*DEK (6p22)* *NUP214 (9q34)*	80.01 (0.37)
t(10;11)(p12;q23)	*KMT2A (11q23)* *MLLT10 (10p12)*	80.10 (0.86)
t(9;11)(p21;q23)	*KMT2A (11q23)* *MLLT3 (9p22)*	83.65 (0.5)
t(6;11)(q27;q23)	*KMT2A (11q23)* *MLLT4 (6q27)*	80.52 (0.82)
t(11;19)(q23;p13.1)	*KMT2A (11q23)* *ELL (19p13.1)*	83.13 (0.95)
t(11;19)(q23;p13.3)	*KMT2A (11q23)* *MLLT1 (19p13.3)*	82.43 (0.36)
t(15;17)(q24;q12)	*PML (15q22)* *RARA (17q21)*	*86.56 (0.86) (*PML–RARA bcr1/bcr2*) 90.47 (0.12) (*PML–RARA bcr3*)
t(3;21)(q26;q22)	*RUNX1 (21q22)* *MECOM (3q26)*	82.70†
t(8;16)(p11;p13)	*KAT6A (8p11)* *CREBBP(16p13)*	78.87†
t(11;17)(q23;q21)	*ZBTB16(11q23)* *RARA (17q21)*	86.11†
t(1;22)(p13;q13)	*RBM15 (1p13)* *MKL1 (22q13)*	80.12 (0.46)

* An additional melting peak corresponding to 90.46°C (0.23) could appear in the *PML–RARA* bcr1/2 subtype because both *PML* forward primers can hybridize to and amplify two fragments of different sizes.
† No SD has been indicated because only one positive control was available for these rearrangements.
A positive control for the fusion gene *FUS−ERG* was not available.

All positive controls showed PCR amplification curves with Ct values ranging from 17 to 30. The melting curves showed single peaks with a specific Tm for each rearrangement (Figure 19). However, additional melting peaks were obtained in rearrangements with alternative splicing (Figure 19, E and G). The negative controls did not generate PCR products. Nevertheless, PCR amplification curves with Ct values>30 were observed for some negative controls, arising as primer-dimer artifacts, producing melting curves with lower Tm values than those of the specific rearrangements.

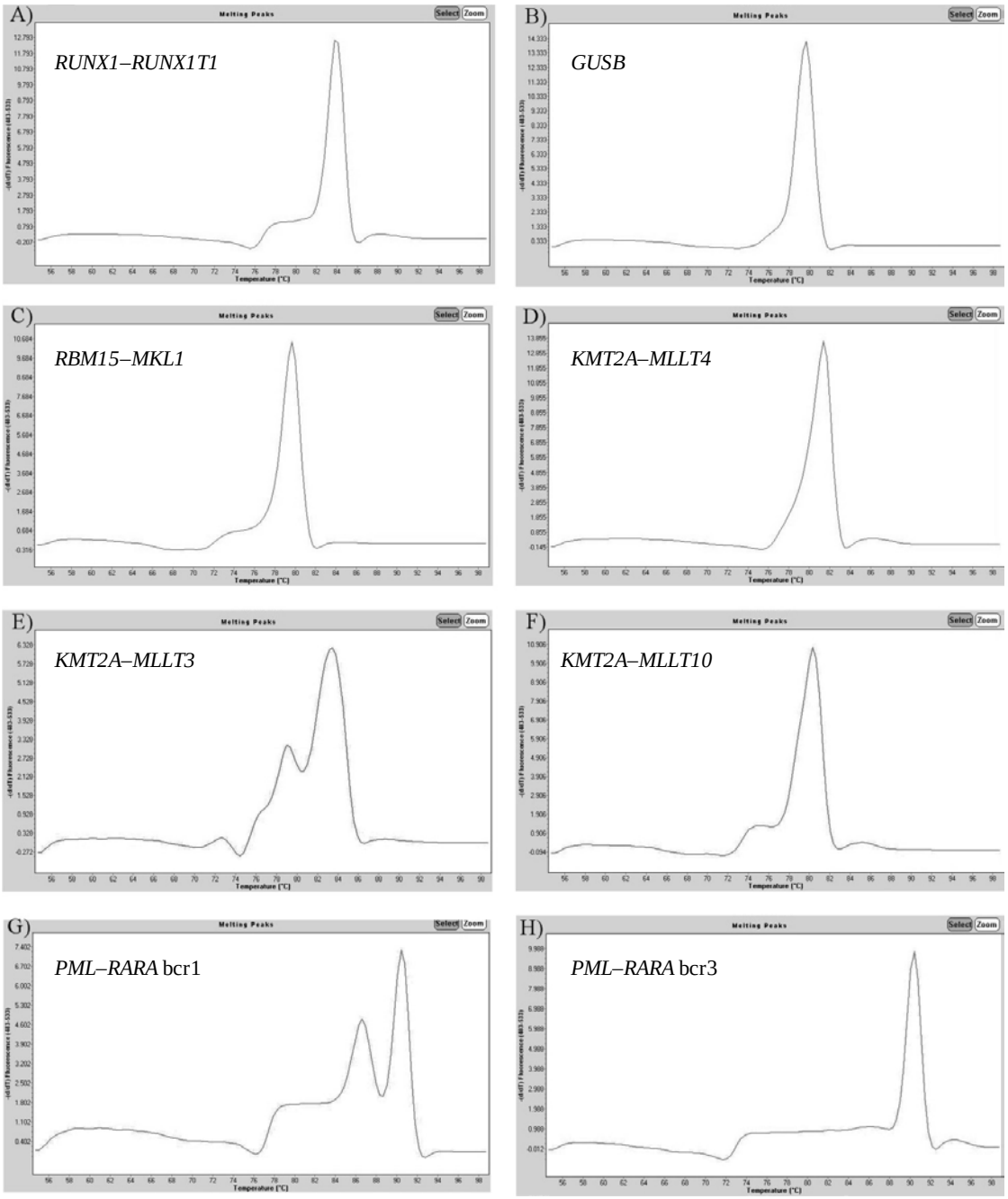

Figure 19. Tm dissociation curves in positive controls: *RUNX1−RUNX1T1* (A), *GUSB* (B), *RBM15−MKL1* (C), *KMT2A−MLLT4* (D), *KMT2A−MLLT3* (E), *KMT2A−MLLT10* (F), *PML−RARA* bcr1 (G) and *PML−RARA* bcr3 (H).

1.1. Sensitivity, specificity, reproducibility and lowest limits of detection

The results of the qRT–PCR assay from 105 diagnostic AML patient samples were consistent with those of the conventional and standard RT–PCR methods, previously established in the Molecular Biology Laboratory of *Hospital Universitario y Politécnico La Fe* and published in the literature (Table 10).

Table 10. Comparative positive and negative results for real-time RT–PCR and other methods.

Translocation/rearrangement	qRT–PCR	Standard RT–PCR	Cytogenetics	FISH
t(15;17)(q24;q12)/*PML–RARA*	25	25	19	19
inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/*CBFB–MYH11*	10	10	9	9
t(8;21)(q22;q22)/*RUNX1–RUNX1T1*	6	6	5	5
t(10;11)(p12;q23)/*KMT2A–MLLT10*	6	6	4	6*
t(9;11)(p22;q23)/*KMT2A–MLLT3*	3	3	2	3*
t(9;22)(q34;q11)/*BCR–ABL1*	2	2	2	1
t(6;11)(p27;q23)/*KMT2A–MLLT4*	1	1	1	1*
t(3;21)(q26;q22)/*RUNX1–MECOM*	1	1	1	0
t(6;9)(p22;q34)/*DEK–NUP214*	1	1	1	0
No translocation/rearrangement	50	50	54	53
Total	105	105	*98	†97

*The FISH method did not enable identification of the partner genes involved in the *KMT2A* rearrangement.
† Seven of 105 samples failed because insufficient metaphases were obtained.
‡ Three samples failed because of poor sample quality, four because of insufficient sample, and one because no *DEK–NUP214* probe was available.

In the samples, the method demonstrated sensitivity of 100% (95% CI, 92% to 100%) and specificity of 100% (95% CI, 91% to 100%). Intra-assay reproducibility assessed for the most frequent rearrangements ranged from 0.12% to 0.40% for the coefficient of variation (CV) of Ct and from 0.06% to 0.13% for the CV of Tm by repeating the analysis five times under the same experimental conditions. The interassay reproducibility ranged from 0.32% to 0.63% for the CV of Ct and from 0.08% to 0.20% for the CV of Tm in five different experiments (Table 11).

Table 11. Results for intra-assay and inter-assay reproducibility of the real-time RT–PCR.

	Ct			T_m		
Rearrangement	**Mean**	**SD**	**CV (%)**	**Mean**	**SD**	**CV (%)**
Intra-assay						
RUNX1–RUNX1T1	24.78	0.10	0.40%	84.12	0.05	0.06%
CBFB–MYH11	25.81	0.08	0.30%	86.29	0.11	0.13%
BCR–ABL1	29.60	0.08	0.28%	84.51	0.08	0.10%
PML–RARA	24.76	0.03	0.12%	90.79	0.07	0.07%
Inter-assay						
RUNX1–RUNX1T1	24.97	0.13	0.51%	84.21	0.12	0.14%
CBFB–MYH11	25.79	0.08	0.32%	86.30	0.13	0.15%
BCR–ABL1	29.52	0.18	0.61%	84.55	0.17	0.20%
PML–RARA	25.00	0.16	0.63%	90.87	0.07	0.08%

Ct: Cycle threshold; Tm: Melting temperature; SD: standard deviation; CV: coefficient of variation

The lowest detection limit for the rearrangements ranged between 1:500 and 1:10,000 using serial dilutions of wild-type cDNA from the positive controls (Figure 20).

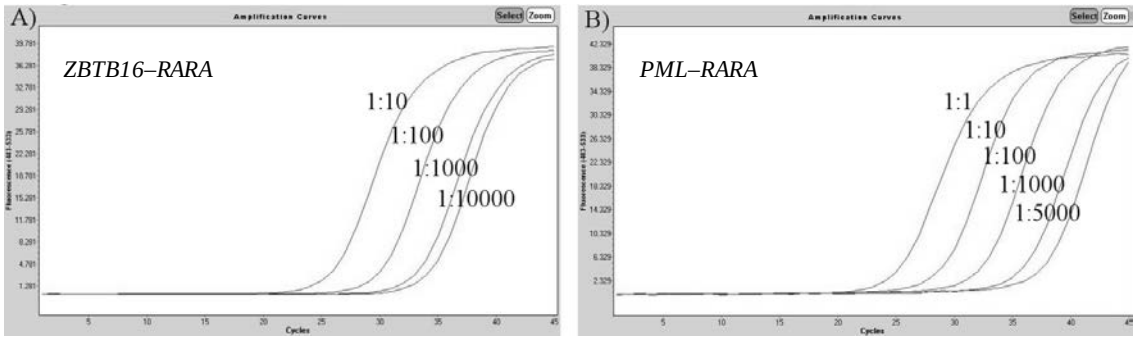

Figure 20. Lowest limit of detection test using dilutions of cDNA from samples positive at diagnosis: *ZBTB16−RARA* (A) and *PML−RARA* (B).

1.2. Validation of real-time RT−PCR in patients with AML

All samples from the 105 patients with AML showed amplification of the *GUSB* control gene with a Ct of <30, indicating suitable mRNA quality. The qRT−PCR method detected 55 rearrangements in the 105 patients: 25 *PML−RARA*, 10 *CBFB−MYH11*, 6 each *RUNX1−RUNX1T1* and *KMT2A−MLLT10*, 3 *KMT2A−MLLT3*,

2 *BCR−ABL1*, and 1 each *DEK−NUP214*, *KMT2A−MLLT4*, and *RUNX1−MECOM*. These rearrangements involved the most common breakpoints and generated the expected Tm in all patients but one (unique patient number [UPN] 80). Samples from this patient showed an amplification curve for *CBFB−MYH11* at Ct 25; however, the Tm (87.49 °C) was slightly higher than that for *CBFB−MYH11* type A (85.92 °C). The amplification product was recovered from the plate and tested in an electrophoresis gel, detecting a 544-bp product, and *CBFB−MYH11* type E was confirmed via sequencing. None of the less frequently occurring rearrangements were detected in any patient with AML; however, the ability of the method to identify them was verified retrospectively by assaying positive controls. The *BCR−ABL1* fusion gene was detected concomitantly with *RUNX1−MECOM* in an AML patient (UPN 35) and with *PML−RARA* in another therapy-related APL patient (UPN 72).

1.3. Correlation of Real-Time RT−PCR results with cytogenetics and FISH in patients with AML

A comparison of qRT−PCR results with those obtained using standard cytogenetic and FISH techniques is given in Table 12.

Table 12. Comparison of positive results obtained with the novel real-time RT−PCR, Cytogenetics and FISH methods.

UPN	qRT–PCR	Cytogenetics	FISH
2	*PML–RARA*	46,XY	*PML–RARA* positive
3	*CBFB–MYH11*	46,XY,inv(16)(p13q22)	*CBFB–MYH11* positive
4	*RUNX1–RUNX1T1*	46,XX,t(8;21)(q22;q22)	*RUNX1–RUNX1T1* positive
6	*CBFB–MYH11*	46,XY	Failed
9	*PML–RARA*	46,XX, t(15;17)(q24;q12)	*PML–RARA* positive
10	*KMT2A–MLLT10*	Failed	*KMT2A* positive
14	*PML–RARA*	46,XY, t(15;17)(q24;q12)	Failed
16	*RUNX1–RUNX1T1*	Failed	*RUNX1–RUNX1T1* positive
17	*PML–RARA*	46,XX, t(15;17)(q24;q12)	*PML–RARA* positive
18	*PML–RARA*	46,XX, t(15;17)(q24;q12)	*PML–RARA* positive
19	*PML–RARA*	Failed	Failed
23	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
26	*PML–RARA*	46,XX, t(15;17)(q24;q12)	*PML–RARA* positive
29	*CBFB–MYH11*	46,XX,t(16;16)(p13;q22)	*CBFB–MYH11* positive
30	*KMT2A–MLLT10*	Failed	*KMT2A* positive
31	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
34	*KMT2A–MLLT10*	46,XY,der(10)t(10;11)(p12;q23)inv(11)(q14;q23),der(11)t10;11)	*KMT2A* positive
35	*RUNX1–MECOM and BCR–ABL1*	46,XX,t(3;21)(q26;q22), t(9;22)(q34;q11)	Failed
38	*CBFB–MYH11*	47,XX, inv(16)(p13q22)	*CBFB–MYH11* positive
39	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
42	*PML–RARA*	46,XX	NAD
44	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
46	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
48	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
55	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
57	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
58	*RUNX1–RUNX1T1*	44,X,-X,t(8;21)(q22;q22)	Failed
59	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
63	*RUNX1–RUNX1T1*	46,XX,t(8;21)(q22;q22)	*RUNX1–RUNX1T1* positive
65	*RUNX1–RUNX1T1*	46,XX,t(8;21)(q22;q22)	*RUNX1–RUNX1T1* positive
66	*PML–RARA*	Failed	Failed
67	*KMT2A–MLLT10*	47,XY,+8,ins(10;11)(p12;q21-23)	*KMT2A* positive
71	*CBFB–MYH11*	46,XX,inv(16)(p13q22)	*CBFB–MYH11* positive
72	*BCR–ABL1 and PML–RARA*	46,XX,t(9;22)(q34;q11), t(15;17)(q24;q12)	*PML–RARA* and *BCR–ABL1* positive
73	*KMT2A–MLLT3*	46,XX,t(9;11)(p21;q23)	*KMT2A* positive
74	*KMT2A–MLLT3*	46,XY, t(9;11)(p21;q23)	*KMT2A* positive
75	*KMT2A–MLLT10*	46,XY,t(10;11)(p12;q23), inv(11)(q13q23)	*KMT2A* positive
77	*CBFB–MYH11*	46XX,t(16;16)(p12;q12)	*CBFB–MYH11* positive
78	*KMT2A–MLLT4*	46,XX,t(6;11)(q27;q23)	*KMT2A* positive
79	*KMT2A–MLLT10*	46,XX,t(8;11;10)(q22;q23;p12)	*KMT2A* positive
80	*CBFB–MYH11*	46,XY,inv(16)(p13q22)	*CBFB–MYH11* positive
81	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
85	*DEK–NUP214*	46,XY,t(6;9)(p21.3;q34)	Failed

86	*CBFB–MYH11*	46,XX,inv(16)(p13q22)	*CBFB–MYH11* positive
87	*PML–RARA*	Failed	Failed
90	*CBFB–MYH11*	46,XX,der(16)inv(16)(p13q22) del(16)(q22)	*CBFB–MYH11* positive
93	*CBFB–MYH11*	46,XX,inv(16)(p13q22)/47,sl,+8	*CBFB–MYH11* positive
94	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
95	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
98	*PML–RARA*	46,XX	NAD
99	*KMT2A–MLLT3*	Failed	*KMT2A* positive
100	*PML–RARA*	46,XY, t(15;17)(q24;q12)	*PML–RARA* positive
103	*RUNX1–RUNX1T1*	46,XY,t(8;21)(q22;q22)	*RUNX1–RUNX1T1* positive

*Failed due to insufficient metaphase spreads (for cytogenetics), poor quality or quantity of samples, or no *DEK–NUP214* probe available (for FISH).

1.3.1 Cytogenetic results

All translocations detected using standard cytogenetic techniques were also confirmed using qRT−PCR. In addition, four patients with cryptic translocations undetected via conventional karyotyping were identified at qRT−PCR: three *PML–RARA* (UPNs 2, 42 and 98) and one *CBFB–MYH11* (UPN 6). Furthermore, the patients whom the standard cytogenetic analysis failed because of insufficient metaphases, qRT−PCR detected three *PML–RARA*, two *KMT2A–MLLT10*, one *RUNX1–RUNX1T1*, and one *KMT2A–MLLT3* rearrangements (UPNs 10, 16, 19, 30, 66, 87, and 99, respectively).

1.3.2 FISH Results

All chromosomal abnormalities detected via FISH were confirmed using qRT−PCR. Furthermore, two *PML–RARA* rearrangements were only detected using qRT–PCR (UPNs 42 and 98). qRT−PCR also enabled detection of seven rearrangements in patients in whom FISH failed because of poor sample quality (UPNs 19, 66, and 87) or insufficient sample (UPNs 6, 14, 35, and 58). qRT−PCR also enabled identification of one *DEK–NUP214* rearrangement not detected via FISH because no *DEK–NUP214* probe was available (UPN 85). qRT−PCR enabled identification of all partner genes involved in the *KMT2A* rearrangements, including *KMT2A–MLLT10* in six patients (UPNs 10, 30, 34, 67, 75 and 79), *KMT2A–MLLT3* in 3 (UPNs 73, 74 and 99), and *KMT2A–MLLT4* in one (UPN 78).

2. Study of *MYBL2* genetic alterations in AML

Based on the results of recent studies, it is conceivable that *MYBL2* could represent a novel biomarker of AML, since its deregulation has been implicated in the development and/or maintenance of the leukemic phenotype, as well as in the prognosis of patients [128, 129]. Therefore, the second part of this thesis aimed to study the most frequently known genetic alterations, as well as, search for unknown genetic variations or mutations that could contribute to a MYBL2 aberrant activity in AML. In this regard, the thesis firstly proposed the study of the most common S427G polymorphism of *MYBL2* as a variation that could contribute to the risk and the prognosis of AML**.** Secondly, a search for new mutations or unknown variants in the coding sequence and adjacent intronic regions of *MYBL2* was performed in AML patients.

2.1. Study of the S427G polymorphism of *MYBL2* in AML

A case-control study of the S427G polymorphism of *MYBL2* was performed in order to analyze the S427G incidence in patients and healthy donors and its association with the risk of AML. In addition, the associations between the S427G change in AML and prognosis of patients was analyzed.

2.1.1 S427G detection and genotype distribution in AML patients and healthy donors

The age and gender were homogeneous between both cohorts of 197 patients and 179 healthy donors ($p > 0.05$): a median age of 60 years old (range = 16 - 91), 113 (57.4%) males and 84 (42.6%) females in the patient group; a median age of 59 years old (range = 15 - 87), 85 (47.5%) males and 94 (52.5%) females in the control group.

The S427G genotype was analyzed through the melting curves obtained from all patients and healthy donors, allowing the wild-type (NM_002466.2:c.1279A) and mutant (NM_002466.2:c.1279G) alleles to be identified. The mean (and standard deviation) Tm for the wild-type and mutant allele was 59.9 °C (0.2 °C) and 67.4 °C (0.3 °C), respectively (Figure 21).

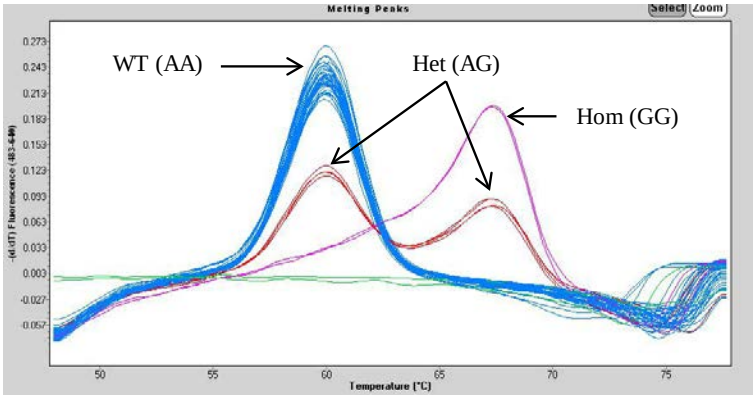

Figure 21. **Melting temperature dissociation curves for the S427G wild-type (NM_002466.2:c.1279A) and mutant (NM_002466.2:c.1279G) alleles.** The figure shows wild-type samples (one melting peak at 59.9 [0.2] °C in blue colour), homozygotes for the mutant allele (melting peak at 67.4 [0.3] °C in pink) and heterozygotes (two melting peaks in red, the Tm of the wild-type and the mutant alleles).

The distribution of the AA, AG and GG genotypes in the patient group was 155/197 (78.7%), 39/197 (19.8%) and 3/197 (1.5%), respectively, and 145/179 (81%), 32/179 (17.9%) and 2/179 (1.1%) in the control group. The genotype distribution in the control group was consistent with the CEU population (Utah residents with ancestry from Northern and Western Europe from the CEPH collection, $p>0.5$), according to the International HapMap Project (hapmap.ncbi.nlm.nih.gov/) and no significant differences were found between the wild-type and S427G distribution in the patient and control or CEU groups (OR = 1.16; 95% CI = 0.70 - 1.92; $p = 0.57$ and OR = 1.64; 95% CI = 0.9 - 3.1; $p = 0.12$, respectively: see Table 13).

Table 13. Incidence of the S427G genotype in the AML patient group and the control group compared to that reported in the population of European ancestry in the HapMap databases.

Population	Genotype frequencies			Odds ratio (95% CI)*	P *	Odds ratio (95% CI)**	P**
	WT (A/A)	Het (A/G)	Hom (G/G)				
AML patients (N=197)	0.787 (155/197)	0.198 (39/197)	0.015 (3/197)	1.16 (0.7-1.9)	0.57	1.64 (0.9-3.1)	0.12
Control group (N=179)	0.81 (145/179)	0.179 (32/179)	0.011 (2/179)	-	-	1.42 (0.7-2.7)	0.29
CEU population in Hapmap (N=113)	0.858 (97/113)	0.142 (16/113)	0 (0/113)	-	-	-	-

WT, Wild-type; Het, Heterozygous; Hom, Homozygous.
The Odds ratio has been calculated by comparing the number of wild-type subjects vs the total number of S427G positive subjects (heterozygous and homozygous).
*Odds ratio (95% CI) compared to the control group; *p* associated with this comparison.
**Odds ratio (95% CI) compared to the CEU (Utah residents with Northern and Western European ancestry from the CEPH collection) population of the International HapMap Project databases; *p* associated with this comparison.

2.1.2 Patient characteristics, clinical outcome and prognostic value according to the S427G polymorphism

The presence of the S427G polymorphism in AML patients was not associated with any clinical or biological characteristic. The outcome of induction therapy was considered in 146 non-APL AML patients treated with intensive chemotherapy, and there were no significant differences in the rate of CR (61.3% vs 67%) or in death during induction therapy (25.8% vs 22.6%) between patients carrying the S427G polymorphism and those that did not (p = 0.58). The median follow-up time in 96 non-APL AML patients who achieved CR was 22 months (range = 2 - 119 months) and there were no significant differences in the OS, DFS or RFS between patients carrying the S427G polymorphism and those whom did not, neither in the whole series nor in any sub-group when age or cytogenetic status were considered separately.

2.2. Search for new *MYBL2* genetic variations in AML

The *MYBL2* coding sequence and flanking intronic regions were screened by HRM in 152 AML patients and alterations in the melting curves were detected in 65 patients (42.8%). Direct sequencing confirmed that samples with a distinct melting curve to that of the wild-type cluster harbored variations in the DNA sequence. In total,

12 different genetic alterations were detected: six non-synonymous variations (Q67X, E132D, P274L, S427G, V595M and I624M); four synonymous variations that did not involve an amino acid change (I160I, D195D, P302P and D571D); and two alterations in adjacent intronic regions (c.186 + 14T >C at intron 3 and c.1365 + 3G >T at intron 8). The SNP databases at the National Center for Biotechnology Information ([NCBI], http://www.ncbi.nlm.nih.gov/snp/) showed that eight of the 12 genetic variations found were registered as polymorphisms, while four variations were not registered: Q67X and E132D as non-synonymous changes and I160I or D571D that did not involve a change in the amino acid. Interestingly, Q67X and I160I were located in the conserved DNA binding domains (http://www.uniprot.org/uniprot/P10244: Table 14).

Table 14. Genetic alterations detected by screening the *MYBL2* coding sequence and flanking intronic regions in AML patients with the High Resolution Melting method.

Genetic alteration	*MYBL2* sequence	Function of the DNA region	Patients (%)	SNP databases ID
NM_002466.2:c.186+14T>C	Intron 3	NS	29/146 (19.9)	rs442143
NM_002466.2:c.199C>T NP_002457.1:p.Q67X	Exon 4	DNA binding domain	1/146 (0.7)	NR
NM_002466.2:c.396G>T NP_002457.1:p.E132D	Exon 5	NS	1/151 (0.7)	NR
NM_002466.2:c.480C>T NP_002457.1:p.I160I	Exon 5	DNA binding domain	1/150 (0.7)	NR
NM_002466.2:c.585C>T NP_002457.1:p.D195D	Exon 6	NS	17/150 (11.3)	rs2229036
NM_002466.2:c.821C>T NP_002457.1:p.P274L	Exon 7	NS	1/151 (0.7)	rs1852379
NM_002466.2:c.906T>C NP_002457.1:p.P302P	Exon 7	NS	21/148 (14.2)	rs285162
NM_002466.2:c.1279A>G NP_002457.1:p.S427G	Exon 8	NS	29/148 (19.6)	rs2070235
NM_002466.2:c.1365+3G>T	Intron 8	Intronic splicing region	4/148 (2.7)	rs73116571
NM_002466.2:c.1713C>T NP_002457.1:p.D571D	Exon 12	NS	1/150 (0.7)	NR
NM_002466.2:c.1783G>A NP_002457.1:p.V595M	Exon 12	NS	1/151 (0.7)	rs7660
NM_002466.2:c.1872C>G NP_002457.1:p.I624M	Exon 13	NS	9/151 (6.0)	rs11556379

NR: No genetic variation registered in the NCBI SNP databases or in sequencing databases of whole genomes and whole exomes (as Cosmic, LOVD or NHLBI GO Exome Sequencing Project databases); NS: function not specified/unknown.

The PolyPhen network software (Polymorphism Phenotyping), which predicts the potential impact of an amino acid substitution on protein function (impact scale between 0 - 1), classified E132D as a benign genetic variation with a score of 0.001. However, the Q67X variation generates a STOP codon in exon 4 of *MYBL2* (Figure 22), which encodes the first DNA binding domain of the protein (http://www.uniprot.org/uniprot/P10244). Moreover, splice site analysis using the Human Splicing Finder software (http://www.umd.be/HSF) showed that the c.1365 + 3G>T variation was located in a potential consensus splicing region at intron 8 (Figure 23).

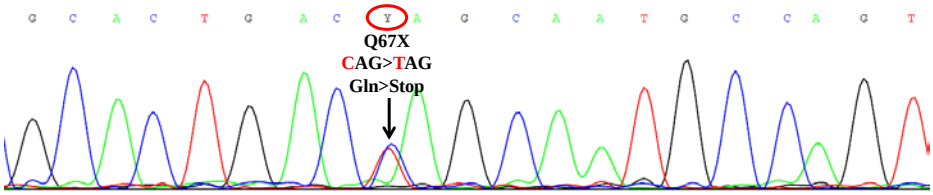

Figure 22. Sanger sequencing traces for Q67X variant (NM_002466.2:c.199C>T) in *MYBL2* gene.

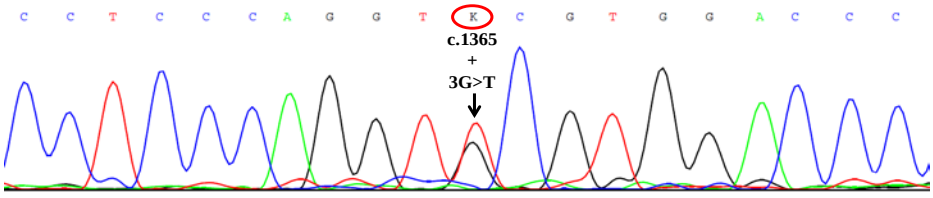

Figure 23. Sanger sequencing traces for c.1365 + 3G>T variant (NM_002466.2:c.1365+3G>T) at intron 8 of *MYBL2* gene.

2.2.1 Incidence of genetic variations in AML patients and SNP databases

The incidence of the rs2070235 (S427G), rs7660 (V595M), rs11556379 (I624M), rs1852379 (P274L), rs442143 (c.1365 + 3G>T, located at intron 3) and rs2229036 (D195D) polymorphisms in AML patients was similar ($p > 0.05$) to that reported in the CEU population from the International HapMap Project or in the SNP databases from the NCBI and the National Cancer Institute ([NCI], http://variantgps.nci.nih.gov/). However, the incidence of the rs73116571 (c.1365 + 3G>T, located at intron 8) and rs285162 (P302P) polymorphisms was significantly

higher in AML patients (2.7% vs 0.15%; OR = 17.9, 95% CI = 2 - 161; p = 0.01 and 14.2% vs 5.3%; OR = 2.7, 95% CI = 1.04 - 6.8; p = 0.04, respectively: see Table 15). The non-synonymous genetic variations found Q67X and E132D that were not registered in the NCBI SNP databases were each detected in one AML patient. Both patients died during early induction treatment and, therefore, CR samples were not available to distinguish between somatic or germline variations or to consider the possibility of being non-annotated SNPs.

Table 15. Incidence of the polymorphisms detected in AML patients compared to those reported in a population of European ancestry in HapMap, NCBI or NCI databases, and the Odds ratio (95% CI) calculated from the genotype frequencies

Polymorphism	AML patient population					Population in Hapmap, NCBI or NCI					Odds ratio (95% CI)§	p
	Genotype frequencies			Allele frequencies		Genotype frequencies			Allele frequencies			
	WT	Het	Hom	WT	Mutant	WT	Het	Hom	WT	Mutant		
rs2070235 (S427G)	0.804 (119/148)	0.176 (26/148)	0.020 (3/148)	0.892	0.108	0.858 (97/113) *	0.142 (16/113) *	0 (0/113) *	0.929 *	0.071 *	1.4 (0.7-2.7)	0.33
rs7660 (V595M)	0.993 (150/151)	0.007 (1/151)	0 (0/151)	0.997	0.003	1 (29/29) ǂ	0 (0/29) ǂ	0 (0/29) ǂ	1 ǂ	0 ǂ	0.6 (0-14.7)	0.74
rs11556379 (I624M)	0.940 (142/151)	0.060 (9/151)	0 (0/151)	0.970	0.030	0.903 (56/62) *	0.097 (6/62) *	0 (0/62) *	0.952 *	0.048 *	0.6 (0.2-1.8)	0.38
rs1852379 (P274L)	0.0993 (150/151)	0.007 (1/151)	0 (0/151)	0.997	0.003	0.000 (65/65) *	0.000 (0/65) *	0.000 (0/65) *	1 *	0 *	1.3 (0.05-32.3)	0.87
rs442143 (T>C)	0.801 (117/146)	0.151 (22/146)	0.048 (7/146)	0.877	0.123	0.879 (51/58) *	0.086 (5/58) *	0.034 (2/58) *	0.922 *	0.078 *	1.8 (0.7-4.4)	0.19
rs73116571 (G>T)	0.973 (144/148)	0.027 (4/148)	0 (0/148)	0.986	0.013	0.998 (661/662) ǂ	0.001 (1/662) ǂ	0 (0/662) ǂ	0.999 ǂ	0.001 ǂ	17.9 (2-161)	0.01
rs2229036 (D195D)	0.887 (133/150)	0.113 (17/150)	0 (0/150)	0.943	0.057	0.905 (599 /662) ǂ	0.092 (61/662) ǂ	0.003 (2/662) ǂ	0.951 ǂ	0.049 ǂ	1.2 (0.7-2.1)	0.54
rs285162 (P302P)	0.858 (127/148)	0.142 (21/148)	0 (0/148)	0.929	0.071	0.947 (107/113) *	0.053 (6/113) *	0 (0/113) *	0.973 *	0.027 *	2.7 (1.04-6.8)	0.04

WT: Wild-type; Het: Heterozygous; Hom: Homozygous.

* Genotype and allele frequencies from CEU population (Utah residents with a Northern and Western European ancestry from the CEPH collection) in the International HapMap Project databases.

ǂ Genotype and allele frequencies in a population of European descent (662 participants from the ClinSeq project [Population ID: CSAgilent]) in the NCBI SNP databases.

ǂ Genotype and allele frequencies in a subpopulation of Caucasian heritage (Population ID: SNP500) in the NCI databases.

§ The Odds ratio was calculated by comparing the number of wild-type subjects versus the total number of S427G positive subjects (heterozygous plus homozygous).

2.2.2 MYBL2 expression in AML according to the rs73116571 polymorphism and the Q67X variation

The Human Splicing Finder software showed that the rs73116571 polymorphism is located in a potential splicing region, therefore, it could affect splicing and, in consequence, modify the expression of *MYBL2* mRNA. The $2^{-\text{delta Ct}}$ *MYBL2* expression data from patients carrying the rs73116571 polymorphism were compared with the median of $2^{-\text{delta Ct}}$ *MYBL2* expression values from all the AML patients (median = 0.24; range = 0.004 - 3.84), which was established as a cut-off between over-expression and non-overexpression of *MYBL2*. The *MYBL2* expression in patients carrying the rs73116571 polymorphism was significantly lower than the median *MYBL2* expression in all AML patients (chi-squared = 2.8; p = 0.046). Similarly, the single patient carrying the Q67X variation, which generates a STOP codon in exon 4 of *MYBL2*, showed lower *MYBL2* expression than the median *MYBL2* expression in all AML patients ($2^{-\text{delta Ct}}$ = 0.1).

3. Study of *MYBL2* role in the mechanism of leukemogenesis

This thesis aimed to inquire into the knowledge of the *MYBL2* role in leukemogenesis. In order to identify the possible trigger mechanism of AML due to the aberrant *MYBL2* expression, a study of *MYBL2* downstream targets was performed by examining gene expression in human HP after being transfected with *MYBL2* expression modulators. Moreover, alterations in normal proliferation/differentiation of these HP were studied by CFC assays.

3.1. Isolation, nucleofection and sorting of $CD34^+$ cells

The main source of $CD34^+$ HP has always been BM, however, in this study $CD34^+$ cells were obtained from umbilical cord blood since it is an easier alternative to obtain enough numbers by non-invasive procedure [145, 146]. A total number of $50_1291{,}150$ $CD34^+$ cells were separated from umbilical cord blood of 46 healthy donors by the $CD34^+$ MicroBead Kit in a MACS separator (MACS, Miltenyi Biotec).

Transfection experiments were carried out by nucleofection, which enables transfer nucleic acids directly into the nucleus and the cytoplasm of $CD34^+$ HP cells so far considered difficult or even impossible to transfect [147, 148]. $CD34^+$ cells were nucleofected with *MYBL2* expression modulators and after 24 hours were sorted into GFP^+ fractions (in case of plasmids) or $FITC^+$ fractions (in case of siRNAs transfections) (Figure 24). The flow cytometry sorter separated a total of 1₁374,662 $CD34^+$ cells transfected with specific *MYBL2* siRNAs, 1₁061,446 $CD34^+$ cells with siRNAs control, 45,900 $CD34^+$ cells with the *MYBL2*–IRES–EGFP over-expression plasmid and 45,044 $CD34^+$ cells transfected with the pIRES–EGFP control plasmid. In all cases, the purity of the sorting was over 99%.

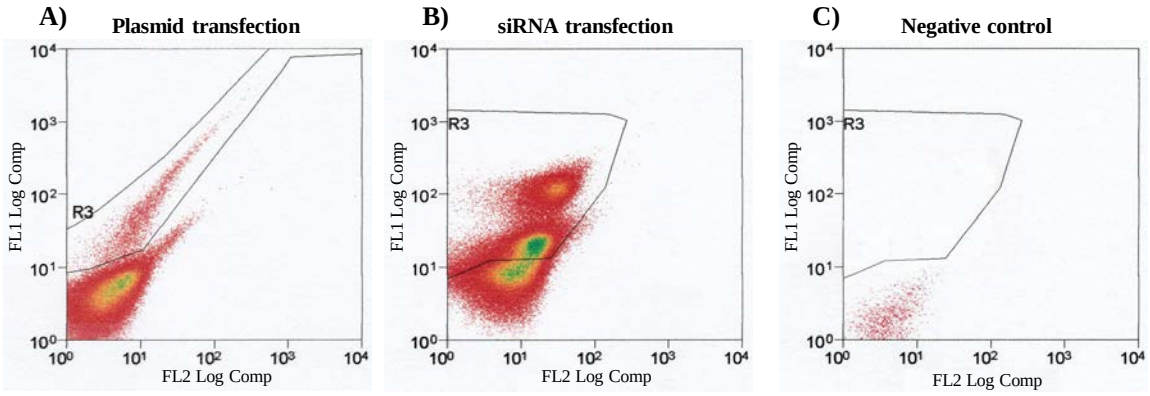

Figure 24. Sorting gate for $CD34^+$ cells transfected with plasmids (A) or siRNAs (B) and a negative control of transfection (C). Transfected $CD34^+$ cells are shown inside the gate R3. The sorting plots from GFP^+ fractions (in case of plasmid transfections, A), showed a different profile than the sorting plots from $FITC^+$ fractions (in case of siRNAs tranfections, B). Figure C showed non-transfected $CD34^+$ cells used to perform the sorting gates for GFP^+ and $FITC^+$ fractions.

3.2. MYBL2 Western Blot

The effect of *MYBL2* expression modifiers on the levels of MYBL2 protein was tested in electroporated K562 cells. The effect of reduction in MYBL2 levels was tested using siRNAs directed against *MYBL2* RNA or scrambled siRNAs control. Western blotting revealed a drastic decrease in the levels of MYBL2 when cells had been transfected with *MYBL2* specific siRNAs compared to cells transfected with siRNAs control (Figure 25A). Western blot analysis revealed higher levels of exogenous

MYBL2 in the GFP^+ -sorted fraction obtained from K562 cells transfected with the *MYBL2*–IRES–EGFP plasmid than in the GFP^+ -sorted fraction obtained from K562 cells transfected with the pIRES–EGFP control plasmid (Figure 25B).

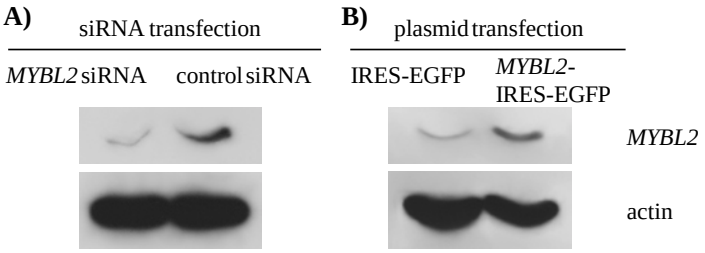

Figure 25. Western blot of MYBL2 and actin proteins from K562 transfected cells. A) Tranfection with *MYBL2* specific siRNAs or control siRNAs. B) Tranfection with pIRES–EGFP control plasmid or *MYBL2*–IRES–EGFP overexpression plasmid.

Thus, these experiments confirm that we could efficiently manipulate MYBL2 expression levels, which is crucial in order to follow studies.

3.3. Identification of genes regulated by *MYBL2* in $CD34^+$ progenitor cells

The study of *MYBL2* downstream targets was performed by gene expression microarrays in human HP with modified *MYBL2* levels.

3.3.1 Gene expression microarrays from $CD34^+$ cells transfected with *MYBL2* specific siRNAs and siRNAs control

Gene expression microarrays were performed from RNA extracted from $FITC^+$ sorted cell populations 24 hours after transfection with *MYBL2* siRNAs or siRNA control. HC clearly identified a different gene expression profile in the $CD34^+$ cells transfected with *MYBL2* siRNAs compared to the $CD34^+$ cells transfected with siRNAs control (Figure 26).

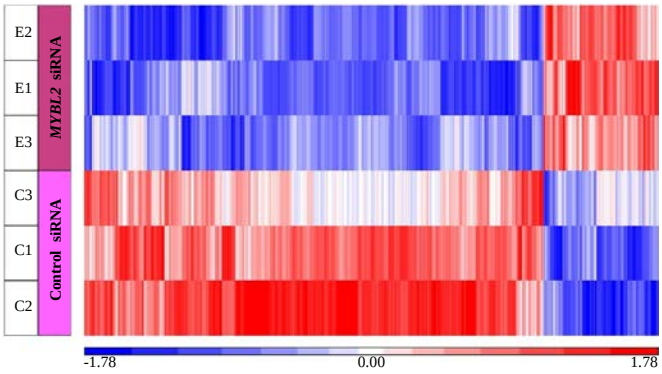

Figure 26. Hierarchical clustering (HC) from differentially expressed genes in human CD34⁺ cells transfected with *MYBL2* specific siRNAs compared to human CD34⁺ cells transfected with siRNAs control. The figure shows a different gene expression profile in transfected CD34⁺ samples comparing the categories "transfection with specific *MYBL2* siRNAs" (in E1, E2 or E3 samples, showed in the rows) versus "transfection with siRNAs control" (in the corresponding paired samples C1, C2 or C3, showed in the rows), establishing p-value < 0.05. The vertical lines colored in red represent genes which are over-expressed and those colored in blue which are lower-expressed in samples.

By applying ±2.5 FC and p<0.05 criteria to the ANOVA test, a number of 270 genes resulted differently expressed after *MYBL2* silencing, which mostly participate in biological processes related to cell proliferation and death (Figure 27).

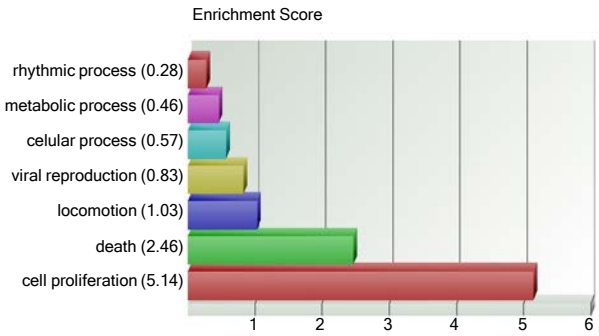

Figure 27. Gene Ontology enrichment in CD34⁺ cells transfected with *MYBL2* specific siRNAs compared to CD34⁺ cells transfected with siRNAs control. The Partek Genomics Suite software allows to integrate results with Gene Ontology databases [149] and establishes an enrichment score from gene analysis lists based on several selection filters of fold change (FC) and p-value. The figure represents the results obtained from Gene Ontology integrated analysis of the most differentially expressed genes in CD34⁺ cells transfected with specific *MYBL2* siRNAs compared to its expression in CD34⁺ cells transfected with siRNAs control (±2.5 FC, p<0.05). It shows a higher enrichment score for cell proliferation and death biological processes in CD34⁺ cells after *MYBL2* silencing.

Mainly, the functions that had been described for these genes in previous publications and databases (as GeneCards, UniProt or OMIM) were involved in

promoting apoptosis and DNA excision-repair processes, however, several of these functions contributed to tumor transformation and survival of malignant cells (Table 16). Additionally, other genes that promote apoptosis (as *p53*, *BAX* or *CASP8*), inhibit the progression of the cell cycle (as *CHEK1/2*, *RB1* or *CDKN1B/p27*) and promote chromatin condensation (*ACIN1*) were detected overexpressed after *MYBL2* silencing with lower but also significant FC values (Table 17). Moreover, several genes, which expression has been described to rely on *MYBL2*, were down-regulated after *MYBL2* silencing (*CLU* [*Apolipoprotein J*] gene, FC = -1.86, p = 0.028; *CCND1* [*Cyclin D1*], FC = -1.48, p = 0.018 and *CCNA1* [*Cyclin A1*], FC = -1.48, p = 0.027), except *BCL2* gene (FC = 1.53, p = 0.009).

Table 16. Major differences in gene expression obtained from CD34$^+$ cells transfected with specific *MYBL2* siRNAs compared to CD34$^+$ cells transfected with control siRNAs. The table shows those 22 genes with higher FC values from a list of 270 genes obtained by establishing ±2.5 FC and $p<0.05$ criteria to the ANOVA test. The main molecular/biological functions and the signaling pathways for these genes are also described.

Gene name	FC (p)	Molecular and biological functions	Signaling pathways	References
MAP2K7 (Protein kinase, mitogen-activated, kinase 7/ JNK-activating kinase 2)	4.372 (0.036)	Threonine/tyrosine kinase Signal transduction Signal transduction activated by proinflammatory cytokines and environmental stresses Signal transduction mediating mitochondrial death, the release of cytochrome c and apoptosis Tumor suppression through *p53*	Mapk, ErbB, TNF, Protein processing in endoplasmic reticulum, Toll-like receptor, T cell receptor, Fc epsilon RI, Neurotrophin, GnRH	150, 151
UTY (Ubiquitously transcribed tetratricopeptide repeat gene on Y chromosome)	4.17 (0.026)	Histone demethylase Immune response Embryonic development	Transcriptional misregulation in cancer	152, 153
CREBBP/ CBP (Creb-binding protein)	3.747 (0.003)	Transcription regulator activity Histone and non-histone protein acetylase Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Transcriptional coactivation Embryonic development Growth control Homeostasis Chromosomal translocation (*CBP–MOZ*) associated with AML	TGF-beta, cell-cicle, cAMP, FoxO, cAMP signaling, HIF-1, Wnt, Notch, Adherens-junction, Jak-STAT, long-term potentiation, HTLV-I infection, pathways in cancer, viral carcinogenesis, microRNAs in cancer, renal cell carcinoma, prostate cancer, thyroid hormone signaling	154-156
RAD23A (UV excision repair protein RAD23 homolog A)	3.723 (0.002)	Multi-ubiquitin chain receptor DNA binding DNA and protein repair Nucleotide excision repair Modulation of proteasomal degradation *RAD23B* promotes genotoxic-specific activation of *p53* and apoptosis	Nucleotide excision repair, protein processing in endoplasmic reticulum	157, 158
ANKHD1/ ANKHD1-EIF4EBP3 (ankyrin repeat and KH domain containing 1 / ANKHD1-EIF4EBP3 readthrough)	3.698 (0.019)	Molecular function unknown Scaffolding protein Anti-apoptosis Associated with abnormal phenotype of leukemia cells: *ANKHD1* silencing inhibits cell proliferation and migration of leukemia cells Isoform 2 protect cells during normal cell survival through regulation of caspases *ANKHD1-EIF4EBP3* is an infrequent fusion-transcript which function is unknown	Unknown	159-161

LIMS1 (LIM and senescent cell antigen-like-containing domain protein 1; PINCH1)	3.683 (0.041)	Adapter molecule Receptor signaling complex scaffold activity Cell communication and signal transduction Regulation of cell survival, cell proliferation and cell differentiation Contributes to apoptosis resistance in cancer cells by suppressing proapoptotic *BIM* (*BCL2L11*) gene *LIMS1* is up-regulated in bone marrow stromal cells (BMSC) from leukemia patients and its inhibition leads BMSC cells to hypoproliferation, decreased migration and increased apoptosis	Unknown	162-164
ERCC6L2 (DNA excision repair protein ERCC-6-like 2)	3.475 (0.036)	Helicase DNA damage-response Nucleotide excision repair Traffics to mitochondria and nucleus in a ROS (reactive oxygen species)-dependent manner	Unknown	165
IGF1R (insulin-like growth factor 1 receptor)	3.407 (0.006)	Transmembrane receptor tyrosine kinase Cell communication and signal transduction Cell growth and survival control Tumor transformation and survival of malignant cells	Ras, Rap1, HIF-1, FoxO, endocytosis, PI3K-Akt, AMPK, focal adhesion, adherens junction, regulating pluripotency of stem cells, long-term depression, pathways in cancer, transcriptional misregulation in cancer, proteoglycans in cancer, glioma, prostate cancer, melanoma	166, 167
CAMK2G (Calcium/calmodulin-dependent protein kinase ii-gamma)	3.393 (0.025)	Protein serine/threonine kinase Cell communication, signal transduction Stimulation to cardiomyocyte apoptosis CaMKs family influences processes as gene transcription, cell survival, apoptosis, cytoskeletal re-organization, learning and memory	ErbB, calcium signaling, cAMP signaling, HIF-1, adrenergic signaling in cardiomyocytes, Wnt, circadian entrainment, long-term potentiation, neurotrophin signaling, cholinergic synapse, dopaminergic synapse, olfactory transduction, inflammatory mediator regulation of TRP channels, insulin secretion, GnRH signaling, proteoglycans in cancer, glioma	168
LUC7L3/CROP (Cisplatin resistance-associated overexpressed protein)	3.388 (0.048)	Transcription regulator activity Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Binds cAMP regulatory element DNA sequence May play a role in the formation of splicesome and the RNA splicing	Unknown	169-171
RUNX1 (runt-related transcription factor 1)	3.337 (0.02)	Transcription factor Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Normal hematopoiesis	Pathways in cancer, transcriptional misregulation in cancer, chronic myeloid leukemia, acute myeloid leukemia	172-178
NFAT5 (nuclear factor of activated T-cells 5, tonicity-responsive)	3.163 (0.009)	Transcription factor activity Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Transcriptional regulation of osmoprotective and inflammatory genes Promoted carcinoma invasion and enhanced cell migration in cell lines derived from breast and colon carcinomas	G protein-coupled receptor (GPCR)	179
ATG4B (autophagy related 4B, cysteine peptidase)	3.162 (0.003)	Cysteine peptidase activity Proteolysis and peptidolysis Vacuole transport and autophagy	Regulation of autophagy	180

TMPO (thymopoietin; TP; lamina-associated polypeptide 2; LAP2)	3.157 (0.006)	Structural organization of the nucleus and in the post-mitotic nuclear assembly Nuclear anchorage of RB1 Controls the initiation of DNA replication through its interaction with NAKAP95 T-cell development and function	Cell Cycle, Mitotic and Cell Cycle, Mitotic	181
ZKSCAN1 (zinc finger with KRAB and SCAN domains 1)	3.152 (0.021)	Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Transcriptional regulator of the KRAB (Kruppel-associated box) subfamily of zinc finger proteins Is thought to be involved in both normal and abnormal cellular proliferation and differentiation	Gene Expression	182
DDX3Y (DEAD [Asp-Glu-Ala-Asp] box polypeptide 3, Y-linked)	3.147 (0.033)	ATP-dependent RNA helicase involved in ATP binding, hydrolysis, RNA binding, and in the formation of intramolecular interactions DDX3Y expression has been detected in all myeloid and lymphoid cells possessing a Y chromosome and in leukemic stem cells	Unknown	183
PRKDC (protein kinase, DNA-activated, catalytic polypeptide; p350)	3.142 (0.001)	Nuclear protein serine/threonine kinase Molecular sensor of DNA damage: phosphorylates histone variant H2AX/H2AFX, regulating DNA damage response mechanism DNA double strand break repair, recombination, telomere stabilization and the prevention of chromosomal end fusion Transcription modulation: phosphorylates DCLRE1C, c-Abl/ABL1, histone H1, HSPCA, c-jun/JUN, p53/TP53, PARP1, POU2F1, DHX9, SRF, XRCC1, XRCC1, XRCC4, XRCC5, XRCC6, WRN, MYC, RFA2 and C1D	Akt Signaling, ERK Signaling, Non-homologous end-joining, Cell cycle	184, 185
MRE11A (MRE11 meiotic recombination 11 homolog A [S, cerevisiae])	3.136 (0.04)	Nuclear protein with 3' to 5' exonuclease activity and endonuclease activity Double-strand break repair, DNA recombination, maintenance of telomere integrity and meiosis DNA damage signaling via activation RAD50-MRE11-NBS1 complex has been identified as part of a large multisubunit protein complex of tumor suppressors, DNA damage sensors, and signal transducers named 'BRCA1-associated genome surveillance complex' (BASC)	Integrated Pancreatic Cancer Pathway, Cell Cycle / Checkpoint Control, Homologous recombination, Non-homologous end-joining	186
SMG1 (SMG1 homolog, phosphatidylinositol 3-kinase-related kinase [C, elegans] pseudogene)	3.110 (0.027)	Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism mRNA surveillance and genotoxic stress response Phosphorylates p53/TP53 and is required for optimal p53/TP53 activation after cellular exposure to genotoxic stress Its depletion leads to spontaneous DNA damage and increased sensitivity to ionizing radiation	mRNA surveillance, Cell Cycle / Checkpoint Control, Gene Expression	187
NISCH (nischarin; imidazoline receptor antisera-selected; IRAS KIAA0975)	3.083 (0.022)	Encodes a non-adrenergic imidazoline-1 receptor protein Binds to the adapter insulin receptor substrate 4 Cell-signaling cascades triggering to cell survival, growth and migration Protection against apoptosis	Integrin-mediated cell adhesion and migration	188, 189

LRRFIP1 (leucine rich repeat [in FLII] interacting protein 1)	3.076 (0.005)	Transcription regulator activity Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Protein homodimerization activity and double-stranded RNA binding May regulate expression of *TNF*, *EGFR* and *PDGFA* Toll-like receptor (TLR) signaling	Signaling by FGFR, Signaling by FGFR1 mutants	190, 191
RAC1 (ras-related C3 botulinum toxin substrate 1 [rho family, small GTP binding protein Rac1]; P21-Rac1; cell Migration-Inducing Gene 5 Protein)	3.072 (0.013)	GTPase activity, cell communication and signal transduction Belongs to the RAS superfamily Control of cell growth, cytoskeletal reorganization, and the activation of protein kinases Regulates cellular responses such as secretory processes, phagocytosis of apoptotic cells, epithelial cell polarization and growth-factor induced formation of membrane ruffles Cell migration and adhesion assembly and disassembly	MAPK signaling, Ras signaling, Rap1 signaling, cAMP signaling, Chemokine signaling, Sphingolipid signaling, Phagosome, PI3K-Akt signaling,Wnt signaling, Axon guidance, VEGF signaling, Osteoclast differentiation, Focal adhesion, Adherens junction, Toll-like receptor signaling, Natural killer cell mediated cytotoxicity, B cell receptor signaling, Fc epsilon RI signaling, Fc gamma R-mediated phagocytosis, Leukocyte transendothelial migration, Neurotrophin signaling, Regulation of actin cytoskeleton, Pathways in cancer, Viral carcinogenesis, Proteoglycans in cancer, Colorectal cancer, Renal cell carcinoma, Pancreatic cancer, Choline metabolism in cancer	192, 193

Table 17. Significant differential expression of key genes related to apoptosis promotion, arrest of cell cycle and chromatin condensation, obtained from CD34$^+$ cells transfected with specific *MYBL2* siRNAs compared to CD34$^+$ cells transfected with control siRNAs ($p < 0.05$). The main molecular/biological functions and the signaling pathways for these genes are also described.

Gene name	FC (p)	Molecular and biological functions	Signaling pathways	References
TP53 (tumor protein p53)	2.083 (0.031)	Transcription factor Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism ; Apoptosis Tumor suppressor Induces cell cycle arrest, apoptosis, senescence, DNA repair or changes in metabolism	MAPK signaling, Sphingolipid signaling, Cell cycle, p53 signaling, PI3K-Akt signaling, Apoptosis, Wnt signaling, Neurotrophin signaling, Thyroid hormone signaling, HTLV-I infection, Pathways in cancer, Transcriptional misregulation in cancer, Viral carcinogenesis, Proteoglycans in cancer, MicroRNAs in cancer, Colorectal cancer, Pancreatic cancer, Endometrial cancer, Glioma, Prostate cancer, Thyroid cancer, Basal cell carcinoma, Melanoma, Bladder cancer, Chronic myeloid leukemia, Small cell lung cancer, Non-small cell lung cancer, Central carbon metabolism in cancer	194-196
BAX (BCL2-associated X protein)	1.806 (0.01)	Receptor signaling complex scaffold activity Proapoptotic, induces cell death by acting on mitochondria 21% of cell lines from human hematopoietic malignancies possessed mutations in BAX, most commonly in the acute lymphoblastic leukemia subset The ratio of BCL2 to BAX determines survival or death following an apoptotic stimulus	Sphingolipid signaling, p53 signaling, Protein processing in endoplasmic reticulum, Apoptosis, Neurotrophin signaling, HTLV-I infection, Pathways in cancer , Viral carcinogenesis, Colorectal cancer	197, 198
CHEK1 (checkpoint kinase 1)	2.238 (0.008)	Protein serine/threonine kinase activity Cell communication , signal transduction Involved in monitoring meiotic recombination, a process that involves programmed DNA breaks Checkpoint-mediated cell cycle arrest and activation of DNA repair in response to the presence of DNA damage or unreplicated DNA	Cell cycle, p53 signaling, HTLV-I infection, Viral carcinogenesis	199-201
CHEK2 (checkpoint kinase 2)	2.041 (0.047)	Protein serine/threonine kinase activity Cell communication , signal transduction Putative tumor suppressor Checkpoint-mediated cell cycle arrest, activation of DNA repair and apoptosis in response to DNA damage or replication blocks Stabilizes the p53 tumor suppressor protein leading to cell cycle arrest in G1	Cell cycle, p53 signaling, HTLV-I infection, Viral carcinogenesis	200-203
RB1 (retinoblastoma 1)	2.049 (0.009)	Negative regulator of the cell cycle that acts as a tumor suppressor. Stabilizes constitutive heterochromatin to maintain the overall chromatin structure. Binds transcription factor E2F1 leading to cell cycle arrest. Defects in this gene are a cause of childhood cancer retinoblastoma, bladder cancer, and osteogenic sarcoma.	Cell cycle, HTLV-I infection, Pathways in cancer, Viral carcinogenesis, Pancreatic cancer, Glioma, Prostate cancer, Melanoma, Bladder cancer, Chronic myeloid leukemia, Small cell lung cancer, Non-small cell lung cancer	204-206

E2F4 (E2F transcription factor 2, p107/p130-binding)	1.522 (0.005)	Transcription factor Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Binds to all three of the tumor suppressor proteins pRB, p107 and p130, but with higher affinity to the last two. It plays an important role in the suppression of proliferation-associated genes, and its gene mutation and increased expression may be associated with human cancer	Cell cycle, TGF-beta signaling	207, 208
E2F3 (E2F transcription factor 3)	1.511 (0.015)	Transcription factor Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Interacts directly with the retinoblastoma protein (pRB) to regulate the expression of genes involved in the cell cycle. Altered copy number and activity of this gene have been observed in a number of human cancers.	Cell cycle, HTLV-I infection, Pathways in cancer, MicroRNAs in cancer, Pancreatic cancer, Glioma, Prostate cancer, Melanoma, Bladder cancer, Chronic myeloid leukemia, Small cell lung cancer, Non-small cell lung cancer	209-211
ACIN1 (apoptotic chromatin condensation inducer 1)	2.064 (0,00004)	DNA binding Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Induces apoptotic chromatin condensation after activation by caspase-3, without inducing DNA fragmentation Involved in mRNA metabolism associated with splicing. Involved in the splicing modulation of BCL2L1/Bcl-X (and probably other apoptotic genes) Regulates cyclin A1 expression in leukemia cells	RNA transport, mRNA surveillance, Spliceosome	212, 213
CDKN1B (cyclin-dependent kinase inhibitor 1B [p27, Kip1])	2.155 (0.014)	Cell cycle control Kinase regulator activity Cell communication ; Signal transduction Important regulator of cell cycle progression Involved in G1 arrest Potent inhibitor of cyclin E- and cyclin A-CDK2 complexes The degradation of this protein is required for the cellular transition from quiescence to the proliferative state	ErbB signaling, HIF-1 signaling, FoxO signaling, Cell cycle, PI3K-Akt signaling, Pathways in cancer, Transcriptional misregulation in cancer, Viral carcinogenesis, MicroRNAs in cancer, Prostate cancer, Chronic myeloid leukemia, Small cell lung cancer	214
CASP8 (caspase 8, apoptosis-related cysteine peptidase)	1.769 (0.012)	Cysteine protease Execution-phase of cell apoptosis Cleaves and activates downstream apoptotic proteases (CASP3, CASP4, CASP6, CASP7, CASP9 and CASP10)	p53 signaling, Apoptosis, Toll-like receptor, NOD-like receptor signaling, RIG-I-like receptor signaling, TNF signaling, Pathways in cancer, Viral carcinogenesis	215, 216

Finally, Cellular Pathways integrated analysis showed the “Transcriptional Misregulation in Cancers”, the “AML”, “CML” and other cancer pathways resulted among those most significantly altered in $CD34^+$ cells after *MYBL2* silencing (Table 18, Figure 28).

Table 18. Cellular Pathways integrated analysis by Partek Genomics Suite software. The Partek Genomics Suite software allows to integrate results with Cellular Kegg Pathways databases (http://www.genome.jp/kegg/ [217]) and establishes an enrichment score from gene analysis lists based on several selection filters of fold change (FC) and P value. The table shows the results obtained from Cellular Pathways integrated analysis of the most differentially expressed genes in $CD34^+$ cells transfected with specific *MYBL2* siRNAs compared to its expression in $CD34^+$ cells transfected with siRNAs control (±2.5 FC, $p<0.05$), showing those cellular pathways with higher enrichment score in $CD34^+$ cells after *MYBL2* silencing.

Pathway Name	Enrichment Score
Transcriptional misregulation in cancers	12.78
Spliceosome	11.92
mRNA surveillance pathway	9.17
Colorectal cancer	8.04
Acute myeloid leukemia	7.75
Chronic myeloid leukemia	7.3
Shigellosis	7.02
Lysosome	6.78
Ubiquitin mediated proteolysis	6.73
Pancreatic cancer	6.67
Bacterial invasion of epithelial cells	6.67
Protein processing in endoplasmic reticulum	6.26
Adherens junction	6.23
Measles	6.23
B cell receptor signaling pathway	6.09
Base excision repair	5.95
Non-small cell lung cancer	5.9
Neurotrophin signaling pathway	5.34
Endocytosis	5.32
Synaptic vesicle cycle	5.31

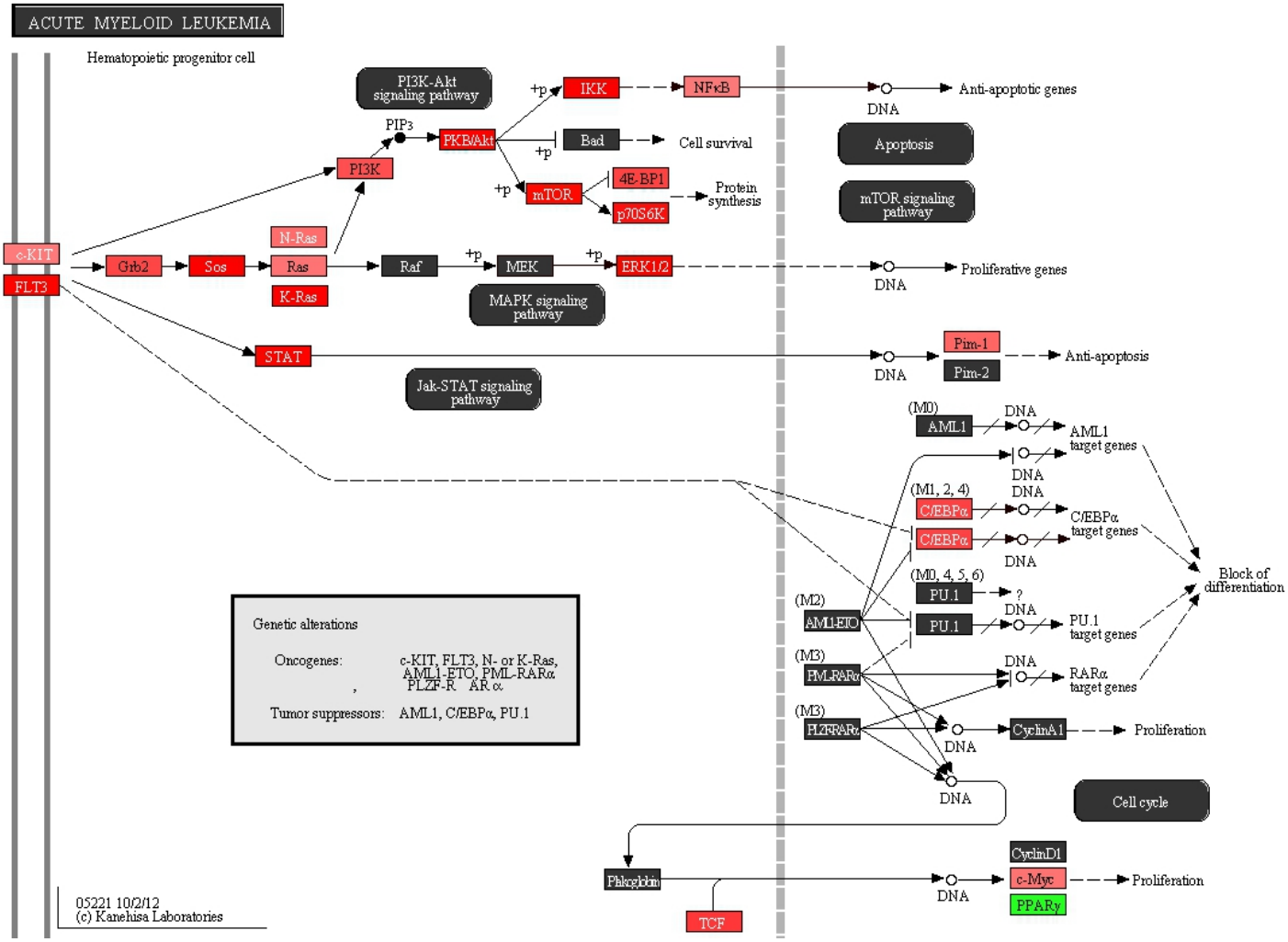

Figure 28. "Acute Myeloid Leukemia" cellular pathway generated by Cellular Pathways Integrated analysis from Partek Genomics Suite software. The figure shows the genes that integrate the "Acute Myeloid Leukemia" cellular kegg pathway (http://www.genome.jp/kegg/ [217]). The genes colored in red represent the genes which resulted more highly expressed in CD34$^+$ cells transfected with specific *MYBL2* siRNAs compared to its expression in CD34$^+$ cells transfected with siRNAs control (±1.5 FC, $p<0.05$). Similarly, the genes colored in green represent the genes whose expression was decreased in CD34$^+$ cells after *MYBL2* silencing. Higher red intensity represents higher gene expression.

3.3.2 Gene expression microarrays from CD34$^+$ cells transfected with *MYBL2*−IRES−EGFP over-expression plasmid and IRES−EGFP control plasmid

Gene expression microarrays were performed from RNA extracted from GFP$^+$ sorted cell populations 24 hours after transfection with *MYBL2*−IRES−EGFP plasmid or IRES−EGFP control plasmid. The HC clearly identified a different gene expression profile in CD34$^+$ cells transfected with *MYBL2*−IRES−EGFP over-expression plasmid compared to the CD34$^+$ cells transfected with IRES−EGFP control plasmid (Figure 29).

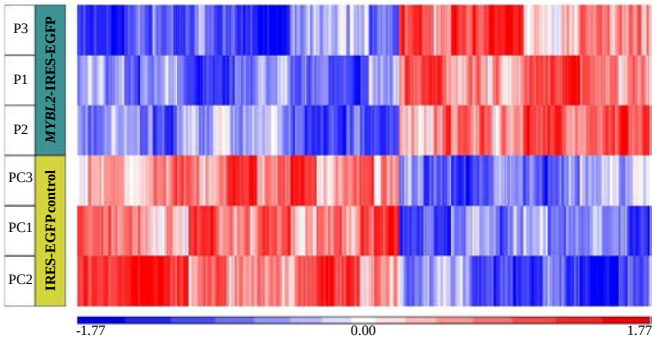

Figure 29. Hierarchical clustering (HC) from differentially expressed genes in human CD34⁺ cells transfected with *MYBL2*–IRES–EGFP over-expression plasmid compared to human CD34⁺ cells transfected with IRES–EGFP control plasmid. The figure shows a different gene expression profile in transfected CD34⁺ samples comparing the categories "transfection with *MYBL2*−IRES−EGFP over-expression plasmid" (in P1, P2 or P3 samples, showed in the rows) versus "transfection with IRES−EGFP control plasmid" (in the corresponding paired samples PC1, PC2 or PC3, showed in the rows), establishing p-value < 0.05. The vertical lines colored in red represent genes which are over-expressed and those colored in blue which are lower-expressed in samples.

By applying ±1.5 FC and p<0.05 criteria to the ANOVA test, a number of 41 genes resulted differently expressed after *MYBL2* over-expression, which mostly participate in processes related to cell metabolism and response to stimulus (Figure 30).

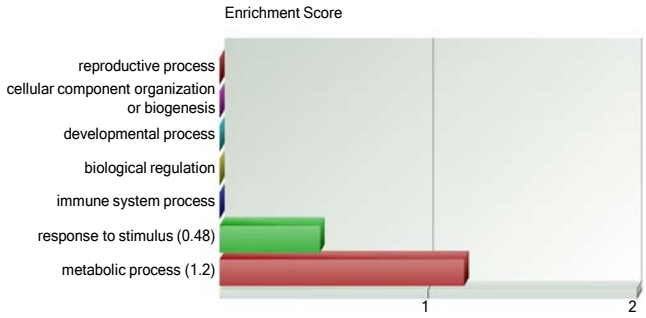

Figure 30. Gene Ontology enrichment in CD34⁺ cells transfected with *MYBL2*–IRES–EGFP over-expression plasmid compared to CD34⁺ cells transfected with IRES–EGFP control plasmid. The Partek Genomics Suite software allows to integrate results with Gene Ontology databases [149] and establishes an enrichment score from gene analysis lists based on several selection filters of fold change (FC) and p-value. The figure represents the results obtained from Gene Ontology integrated analysis of the most differentially expressed genes in CD34⁺ cells transfected with *MYBL2*–IRES–EGFP over-expression plasmid compared to its expression in CD34⁺ cells transfected with IRES–EGFP control plasmid (±1.5 FC, p<0.05). It shows a higher enrichment score for cell metabolism and response to stimulus processes in CD34⁺ cells after *MYBL2* over-expresssion.

The main functions attributed to these genes in previous publications and databases (as GeneCards, UniProt or OMIM) were involved in enhancing biological processes as cell adhesion, migration, inflammation, immunity, central nervous and embryonic development (Table 19).

Table 19. Major differences in gene expression obtained from CD34$^+$ cells transfected with *MYBL2*–IRES–EGFP over-expression plasmid compared to CD34$^+$ cells transfected with IRES–EGFP control plasmid. The table shows those 22 genes with higher FC values from a list of 41 genes obtained by establishing ±1.5 FC and p<0.05 criteria to the ANOVA test. The main molecular/biological functions and the signaling pathways for these genes are also described.

Gene name	FC (p)	Molecular and biological functions	Signaling pathway	References
NPTX1 (Pentraxin I, neuronal)	2.683 (0.014)	Transporter activity Cell communication, signal transduction Developmental and activity-dependent synaptic plasticity Mediator of hypoxic-ischemic injury Induction of neural development	Unknown	218-220
C3 (Complement component 3)	2.418 (0.026)	Complement activity Activation of the complement system, immune response Promotes neurogenesis Mediator of local inflammatory process Induces contraction of smooth muscle, vascular permeability and histamine release from mast cells and basophilic leukocytes	Phagosome, complement and coagulation cascades, pertussis, legionellosis, leishmaniasis, chagas disease, staphylococcus aureus infection, tuberculosis, herpes simplex infection , viral carcinogenesis, systemic lupus erythematosus	221, 222
COL6A1 (Collagen, type vi, alpha-1)	2.247 (0.009)	Extracellular matrix structural constituent Cell growth and/or maintenance Cell-binding Metastasis-associated protein	PI3K-Akt, focal adhesion, ECM-receptor interaction, protein digestion and absorption	223
CD22 (B-cell antigen CD22; sialic acid-binding immunoglobulin-like lectin 2; SIGLEC2)	2.138 (0.035)	Cell adhesion molecule activity Immune response Mediates B-cell B-cell interactions Marker for B-cell precursor acute lymphocytic leukemia Structurally related to myelin-associated glycoprotein (MAG), neural cell adhesion molecule (NCAM), and carcinoembryonic antigen (CEA)	Cell adhesion molecules (CAMs), hematopoietic cell lineage, B cell receptor signaling	224, 225
SLAMF7 (SLAM family, member 7; CS1; CD319; CD2-like receptor activating cytotoxic cells)	2.022 (0.037)	Cell surface receptor Immune response May play a role in lymphocyte adhesion Mediates activating or inhibitory effects in NK cells (isoform 1 mediates NK cell activation through a SH2D1A-independent extracellular signal-regulated ERK-mediated pathway; isoform 3 does not mediate any NK cell activation)	Unknown	226-229
SOX8 (SRY-box 8)	-1.975 (0.045)	Transcription factor activity Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism Involved in the regulation of embryonic development and in the determination of the cell fate	Unknown	230
KCNMB1 (SLO-beta; potassium channel, calcium-activated, large conductance, subfamily m, beta member 1)	1.974 (0.022)	Molecular function unknown Transport Regulatory subunit of the calcium activated potassium KCNMA1 (maxiK) channel Calcium regulation of gene expression Killing activity of neutrophils	cGMP-PKG signaling pathway, vascular smooth muscle contraction, insulin secretion	231, 232

LPXN (Leupaxin)	1.895 (0.009)	Adapter molecule Receptor signaling complex scaffold activity Cell communication; signal transduction Regulation of cell adhesion, spreading and cell migration Negatively regulates B-cell antigen receptor (BCR) signaling	Unknown	233-235
ALOX5 (5-lipoxygenase; LOG5; 5-LO)	1.884 (0.046)	Lipase activity Metabolism ; Energy pathways Leukotriene biosynthesis Role in inflammatory processes Proliferative, antiapoptotic and angiogenic functions 5-LO-positive macrophages localize to areas of neoangiogenesis	Arachidonic acid metabolism, metabolic pathways, serotonergic synapse, ovarian steroidogenesis, toxoplasmosis	236-240
MLLT4 (myeloid/lymphoid or mixed lineage leukemia, translocated to, 4; AF6; Afadin)	1.743 (0.014)	Cell junction protein; cell adhesion molecule activity Cell communication ; Signal transduction Is a Ras target that regulates cell-cell adhesions downstream of Ras activation Involved in signaling and organization of cell junctions during embryogenesis Plays a role in the organization of homotypic, interneuronal and heterotypic cell-cell adherens junctions (AJs) Is fused with KMT2A in leukemias caused by t(6;11) translocations	Ras signaling, Rap1 signaling, cAMP signaling, adherens junction, tight junction, leukocyte transendothelial migration	241-245
CLCN4 (chloride channel 4)	1.722 (0.037)	Voltage-gated ion channel activity Transport Antiport system that exchanges chloride ions against protons May be involved in copper metabolism Physiological role unknown but may contribute to the pathogenesis of neuronal disorders Inducer of colon cancer cell migration, invasion and metastases	Unknown	246-248
USP37 (Ubiquitin specific protease 37)	1.708 (0.002)	Ubiquitin-specific protease activity; Ubiquitin proteasome system protein Protein metabolism Antagonizes the anaphase-promoting complex (APC/C) during G1/S transition by mediating deubiquitination of cyclin-A (CCNA1 and CCNA2), thereby promoting S phase entry Regulates the oncogenic fusion protein PLZF−RARA stability and cell transformation	Unknown	249-251
RARRES3 (retinoic acid receptor responder [tazarotene induced] 3)	-1.671 (0.018)	Nuclear receptor protein, member of the steroid and thyroid hormone receptor superfamily of transcriptional regulators Mediates the biologic effects of retinoids, such as potent growth inhibitory and cell differentiation activities Is thought act as a tumor suppressor or growth regulator	Unknown	252, 253
FBLN1 (Fibulin 1)	1.662 (0.035)	Extracellular matrix structural constituent Cell growth and/or maintenance Localized in several areas of early human embryos; could be important for certain developmental processes and contribute to the supramolecular organization of extracellular matrix architecture May play a role in cell adhesion and migration, cellular transformation and tumor invasion, appearing to be a tumor suppressor	Unknown	254-259

FXYD6 (FXYD domain containing ion transport regulator 6; phosphohippo lin)	1.646 (0.048)	Ion channel activity Transport Affects the activity of Na,K-ATPase The expression is primarily in the brain Highly increased in bile duct tumor The reciprocal 3'-KMT2A gene segment was fused out-of-frame to FXYD6 in t(4;11) acute leukemia patients KMT2A.AF4(+)/AF4.KMT2A(-)	Unknown	260-263
C3AR1 (complement component 3a receptor 1)	1.644 (0.014)	G-protein coupled receptor activity Cell communication; Signal transduction Promote both basal and ischemia-induced neurogenesis Stimulates chemotaxis, granule enzyme release and superoxide anion production A gene-expression signature included over-expression of C3AR1 in FLT3-TKD(+) differentiating from FLT3-WT in cytogenetically normal acute myeloid leukemia patients	Neuroactive ligand-receptor interaction, complement and coagulation cascades, staphylococcus aureus infection	264, 265
CD276 (B7 homolog 3; B7H3)	-1.619 (0.015)	Ligand, receptor binding Immune response	Cell adhesion molecules (CAMs)	266-268
SHROOM2 (SHROOM family member 2; apical protein of xenopus-like; APXL)	1.6 (0.006)	Molecular function unknown May regulate cell morphology through myosin II May be involved in endothelial cell morphology changes during cell spreading	Tight junction	269, 270
SORBS1(Sorb in and SH3-domains containing 1; SH3D5; CBL-associated protein; CAP; Ponsin; SORB1)	1.578 (0.003)	Cell adhesion molecule activity Signal transduction; Cytoskeleton organization and biogenesis Tyrosine phosphorylation of CBL by linking CBL to the insulin receptor Insulin-stimulated glucose transport Formation of actin stress fibers and focal adhesions	PPAR signaling; adherens junction, insulin signaling	271-273
BCL2L11 (BCL2-like 11; BCL2-interacting protein BIM; BIM)	1.568 (0.000 3)	Adapter molecule; Receptor signaling complex scaffold activity Apoptosis facilitator	FoxO signaling, PI3K-Akt signaling, microRNAs in cancer	274-276
CLEC5A (C-type lectin domain family 5, member A; CLECSF5; myeloid DAP12-associating lectin 1; MDL1)	1.553 (0.039)	Cell surface receptor activity Cell communication ; Signal transduction Activates myeloid cells via CLECSF5/DAP12 complexes MDL-1 stimulation maintained cell survival Plays a role in innate immunity, involving neutrophils and macrophages	Unknown	277-279
CREBRF (CREB3 regulatory factor)	-1.527 (0.009)	Molecular function unknown Negative regulator of the endoplasmic reticulum stress response or unfolded protein response (UPR) Represses the transcriptional activity of CREB3 (a protein involved in cell proliferation and migration, tumor suppression and inflammatory gene expression; associated with human t-cell leukemia virus type 1 disease)	Unknown	280, 281

The Cellular Pathways integrated analysis showed the "Hematopoietic cell lineage" resulted among those pathways most significantly altered in $CD34^+$ cells after *MYBL2* over-expression (Table 20).

Table 20. Cellular Pathways integrated analysis by Partek Genomics Suite software. The Partek Genomics Suite software allows to integrate results with Cellular Kegg Pathways databases (http://www.genome.jp/kegg/ [217]) and establishes an enrichment score from gene analysis lists based on several selection filters of fold change (FC) and p-value. The table shows the results obtained from Cellular Pathways integrated analysis of the most differentially expressed genes in $CD34^+$ cells transfected with specific *MYBL2* over-expression plasmid compared to its expression in $CD34^+$ cells transfected with control plasmid (±1.5 FC, $p<0.05$), showing those cellular pathways with higher enrichment score in $CD34^+$ cells after *MYBL2* over-expression.

Pathway Name	Enrichment Score
Staphylococcus aureus infection	14.62
Hematopoietic cell lineage	9.04
Toxoplasmosis	8.69
Complement and coagulation cascades	7.97
Leishmaniasis	7.77
Pertussis	7.58
Phagosome	6.09
Asthma	6.07
Adherens junction	5.54
Allograft rejection	5.46
Influenza A	5.42
Antigen processing and presentation	5.39
Graft-versus-host disease	5.17
Cell adhesion molecules (CAMs)	5.05
Type I diabetes mellitus	5.04
Systemic lupus erythematosus	4.99
Intestinal immune network for IgA production	4.73
Rheumatoid arthritis	4.72
Autoimmune thyroid disease	4.51
Arachidonic acid metabolism	4.13

Finally, the genes *MLLT4*, *BCL2L11*, *MAP2K7*, *CHEK2*, *CDKN1B* and *WNT8A* were inversely expressed in $CD34^+$ HP after inducing *MYBL2* over-expression compared to its expression in those $CD34^+$ HP after *MYBL2* silencing (Table 21).

Table 21. Gene expression of *MLLT4, BCL2L11, MAP2K7, CHEK2, CDKN1B* and *WNT8A* in CD34$^+$ transfected with p*MYBL2*–IRES–EGFP and its expression in those CD34$^+$ cells transfected with *MYBL2* siRNAs, both compared to its expression in CD34$^+$ cells transfected with the corresponding control (pIRES–EGFP or siRNAs control, respectively). Fold change (FC) and p-value from each gene are shown.

Gene	Gene expression in CD34$^+$ cells transfected with p*MYBL2*–IRES–EGFP		Gene expression in CD34$^+$ cells transfected with *MYBL2* siRNAs	
	FC	p	FC	p
MLLT4	1.74	0.014	-1.584	0.01
BCL2L11	1.568	0.0003	-2.231	0.004
MAP2K7	-1.198	0.05	4.371	0.036
CHEK2	-1.156	0.011	2.041	0.047
CDKN1B	-1.179	0.028	2.155	0.014
WNT8A	1.182	0.023	-1.562	0.048

3.3.3 Validation of gene expression by PCR

The relative expression of *MLLT4*, *BCL2L11*, *MAP2K7* and *CREBRF* genes obtained by microarrays, as well as *MYBL2* expression, was confirmed by PCR in three different samples of CD34$^+$ cells transfected with *MYBL2* over-expression plasmid compared to such CD34$^+$ cell samples transfected with control plasmid. Similarly, the relative expression of *MLLT4*, *BCL2L11*, *MAP2K7* and *CREBBP* genes, as well as *MYBL2* expression, was validated by PCR in further three different samples of CD34$^+$ cells transfected with *MYBL2* siRNAs compared to such CD34$^+$ cell samples transfected with control siRNAs (Figure 31).

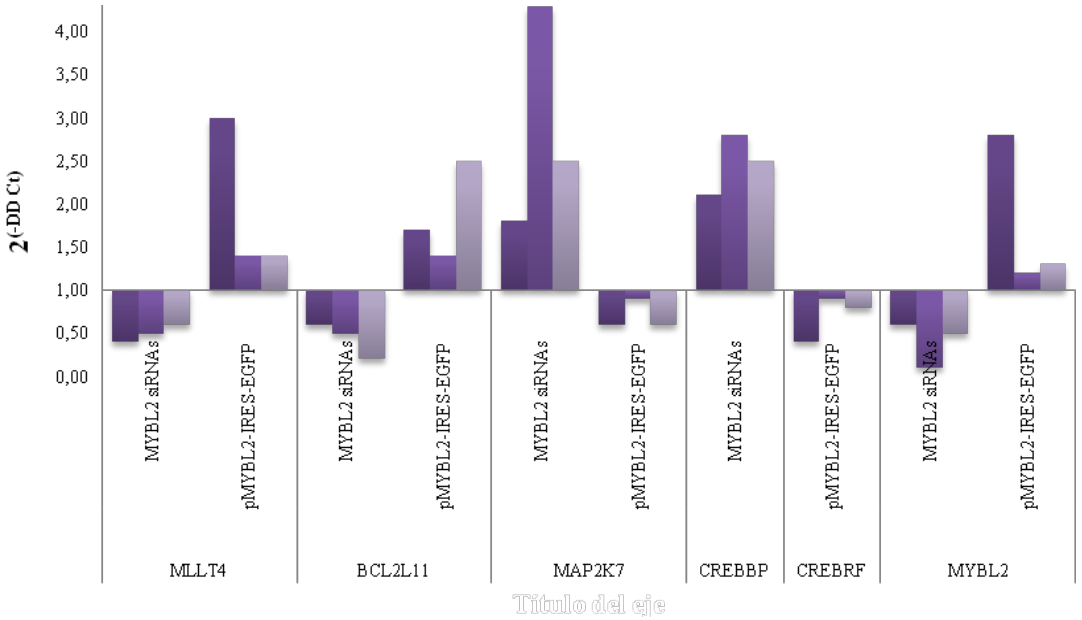

Figure 31. Relative expression of *MLLT4, BCL2L11, MAP2K7, CREBBP, CREBRF* and *MYBL2* genes obtained by PCR in CD34⁺ cell samples transfected with *MYBL2* expression modifiers. The $2^{-(\text{delta-delta Ct})}$ shows relative expression values of *MLLT4*, *BCL2L11*, *MAP2K7*, *CREBBP*, *CREBRF* and *MYBL2* genes in three different CD34⁺ cell samples that were transfected with *MYBL2* siRNAs or with *MYBL2*–IRES–EGFP over-expression plasmid. These $2^{-(\text{delta-delta Ct})}$ values were calculated from absolute gene expression in CD34⁺ cell samples transfected with *MYBL2* expression modifiers (with *MYBL2* siRNAs or with *MYBL2*–IRES–EGFP over-expression plasmid) compared to absolute gene expression in such CD34⁺ cells transfected with the corresponding control (control siRNAs or control plasmid, respectively). $2^{(-\text{DDCt})}>1$ means gene over-expression and $2^{(-\text{DDCt})}<1$ means gene lower-expression in transfected CD34⁺ cells. All data was normalized using endogenous *GUSB* control gene.

3.4. Study of alterations triggered by *MYBL2* on CD34⁺ HP proliferation by colony forming cell assays

CFC assays demonstrated that total colony count from CD34⁺ cells transfected with *MYBL2* siRNAs was significantly lower than those transfected with control siRNAs (median, 54 vs 76; range, 36-65 vs 61-86; p=0.005) [Figure 32A]. The total count of colony forming units for monocytes (CFU-M) and granulo-monocytes (CFU-GM) was also significantly lower in cells transfected with *MYBL2* siRNAs (median, 31 vs 46; range, 23-39 vs 36-61; p=0.002).

Similarly, total colony count of CD34⁺ cells transfected with *MYBL2*–IRES–EGFP over-expression plasmid was significantly higher than those transfected with control plasmid (median, 78 vs 54; range, 66-115 vs 29-89; p=0.004), being also

significantly higher the total number of CFU-M and CFU-GM (median, 40 vs 20.5; range, 28-55 vs 13-41; p=0.007) [Figure 32B].

Altogether, these results showed that $CD34^+$ cell proliferation was significantly affected by modifying *MYBL2* expression.

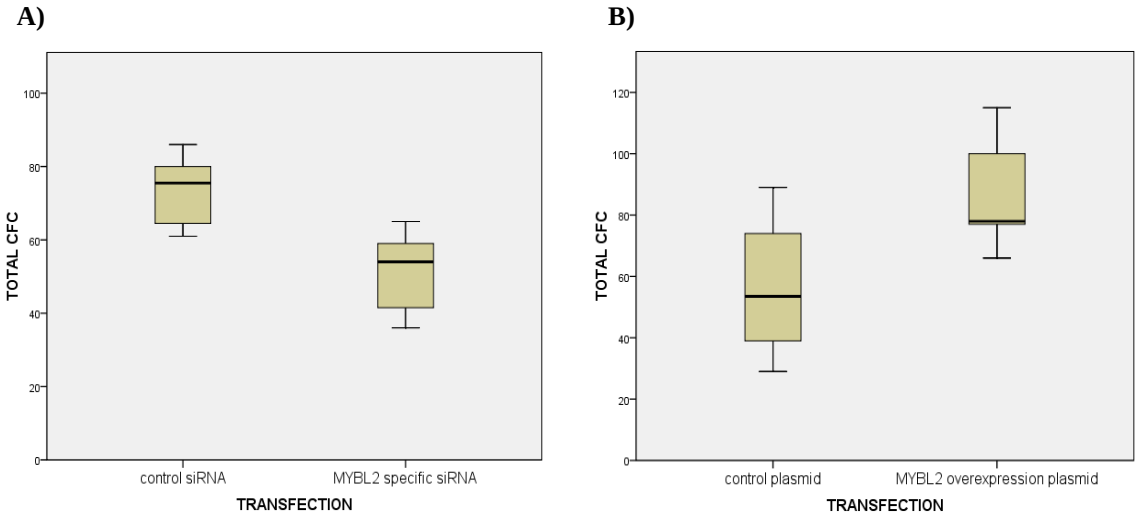

Figure 32. Box plots of total number of colony forming cells (CFC) from transfected $CD34^+$ cells after culture. A) Box plots of CFC from $CD34^+$ cells transfected with *MYBL2* specific siRNAs or control siRNAs. B) Box plots of CFC from $CD34^+$ cells transfected with *MYBL2*–IRES–EGFP over-expression plasmid or IRES–EGFP control plasmid.

Several images of different CFC types obtained from transfected $CD34^+$ cells are shown in Figure 33.

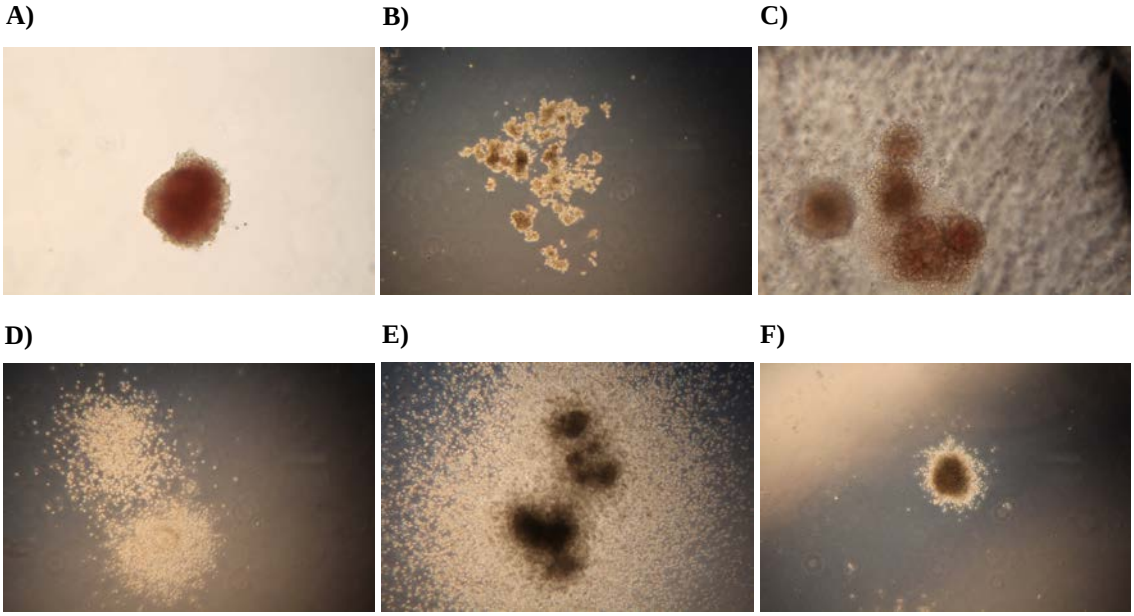

Figure 33. Different colony forming cell (CFC) images shown from a high-quality inverted microscope (Axiovert Zeiss) equipped with 2X, 4X and 10X planar objectives. A) Colony-forming unit erythrocyte (CFU-E) 10X; B) Burst forming unit-erythroid (BFU-E) 4X; C) Colony-forming unit granulocyte-erythrocyte-monocyte-megakaryocyte (CFU-GEMM) 10X D) Colony-forming unit monocyte (CFU-M) 4X; E) Colony-forming unit granulo-monocyte (CFU-GM) 4X; F) Colony-forming unit granulocyte (CFU-G) 4X.

VI. DISCUSSION

The present thesis reports a novel qRT−PCR assay for simultaneous detection of 15 AML-associated fusion genes. This assay may provide a reliable tool to improve the molecular characterization of AML, since it enables an accurate and fast molecular screening by covering a wide spectrum of well-known fusion transcripts associated with the disease. In addition, the current thesis presents a further study of the gene *MYBL2*, an emerging candidate to be related to the development of AML and other cancers, by exploring its genetic alterations in AML patients and its role in HP development.

1. Applicability of the real-time RT−PCR method for simultaneous detection of AML-associated fusion genes

Despite the advent of innovative assays such as next-generation sequencing, which enable more comprehensive screening of AML molecular lesions, the detection of balanced molecular rearrangements which are required to be detected in RNA are scarcely developed by these technologies [282]. In addition, these innovative assays are still neither sufficiently cost-effective nor practical for use in laboratories involved in routine genetic diagnosis. Therefore, it is important to optimize methods that would prove useful and easily applicable for use in hospital laboratories. Several studies have optimized RT−PCR methods for various target lesions in acute leukemias by performing multiplex RT−PCR and also including Luminex, capillary electrophoresis, or enzyme-linked amplicon hybridization methods [31–38]. However, all of these methods are based on conventional RT−PCR, extend the analysis over time, or include limited coverage of AML specific rearrangements. In contrast, qRT−PCR has been optimized to detect a wide spectrum of AML rearrangements by performing simultaneous singleplex reactions, and it also seems to be convenient and cost-effective for use in hospital laboratories.

The use of random primers for reverse transcription and SYBR Green to detect PCR products provides a flexible system and enables incorporation of different primers to detect a variety of rearrangements. The method also enables detection of rare variant transcripts, for example, it has found a *CBFB−MYH11* fusion transcript with a higher Tm than the most common *CBFB−MYH11* type A used as a control. The high sensitivity, specificity, and reproducibility of the assay make it suitable for screening

specific rearrangements in AML. Indeed, qRT−PCR confirmed all abnormalities previously detected using cytogenetics or FISH, and in addition it was sensitive enough to identify small cryptic rearrangements that had been unnoticed using standard techniques [283, 284]. Also of interest, was the identification via qRT−PCR of all partner genes involved in the *KMT2A* rearrangements in the patient series. However, a limitation of this assay is that it is inappropriate for monitoring minimal residual disease, for which quantitative qRT−PCR is required for each specific rearrangement detected. Although no samples in the patient series were positive for *KMT2A−MLLT1*, *KAT6A−CREBBP*, *ZBTB16−RARA*, *KMT2A−ELL* or *RBM15−MKL1*, probably because of the low frequency of occurrence of these fusion genes [27, 42, 44, 60, 285], the ability to detect these rearrangements of this method was verified by testing positive controls for each of them. For fusion gene *FUS−ERG*, no control was used nor was a positive sample found. However, inasmuch as their primers have been successfully tested in previous studies [41], certainly the detection of this particular rearrangement can be effectively achieved using this method.

An important finding of the present study was that two different rearrangements were detected simultaneously in two patients. One patient with acute monocytic leukemia and another with therapy-related APL had two different simultaneous fusion genes in which *BCR−ABL1* was implicated in both: *RUNX1−MECOM* together with *BCR−ABL1* and *PML−RARA* together with *BCR−ABL1*, respectively. In the hypothetical event of failure or lack of results with FISH or cytogenetic methods, the typical workup in a patient with morphologic, immunophenotypic, and clinical features of acute promyelocytic leukemia would be geared to focused evaluation of *PML−RARA*, and thus the *BCR−ABL1* rearrangement would have remained undetected. Therefore, it can be speculated that integration of this procedure for simultaneous detection of fusion genes (rather than conventional procedures for single detection) via FISH and cytogenetic methods may result in future detection of more than one abnormality in an increasing number of patients.

2. *MYBL2* as a new molecular marker in AML

Although the molecular rearrangements are considered an important item for the diagnosis, prognosis and monitoring of AML, near 40% of patients show a normal karyotype. Mainly in this group of AML, it is desirable to identify new molecular markers [11]. The study of new AML molecular markers contributes to the prognosis, best therapeutic approach to follow and allows to research into the mechanisms involved in leukemogenesis, which will be important for the development of new therapies. Therefore, besides the comprehensive analysis of molecular rearrangements, further studies of genes that have been associated with leukemogenesis and the prognosis of this disease, as *MYBL2*, are needed to contribute to the research of new molecular markers in AML. Previous to this study, there were no publications that have deepened the analysis of mutations and polymorphisms of *MYBL2* in AML. The *MYBL2* S427G polymorphism has been linked to a decrease in the overall risk of neuroblastoma, chronic myelogenous leukemia and colon cancer in a combined dataset of cases and controls [131]. Nonetheless, two independent case-control studies demonstrated no association of S427G with cancer risk in a heterogeneous group of breast malignancies [132, 133], although one study did associate the S427G polymorphism with an increase of the basal-like breast cancer [132]. In agreement with the studies of the S427G polymorphism and overall breast cancer, the results obtained from this thesis did not show any difference in the incidence of S427G in AML patients, suggesting it does not contribute to the susceptibility to develop AML.

Aside from the S427G, seven additional *MYBL2* polymorphisms had been identified by HRM in the patient series, including five which showed a similar incidence to that reported in a population of European descents from the International HapMap Project, the NCBI or the NCI SNP databases, while two (the rs73116571 and rs285162 polymorphisms) did not. Interestingly, the rs73116571 polymorphism was located in a potential splicing region in intron 8 and it had a significantly higher incidence in AML patients (17.9-fold). In addition, *MYBL2* expression was weaker in patients carrying this rs73116571 polymorphism, suggesting that rs73116571 could disturb splicing, disrupting the synthesis of the mRNA coding strand. Given the role of *MYBL2* in protecting chromosome integrity [105, 106], the reduction in *MYBL2* expression caused by the rs73116571 polymorphism could lead to genome instability in myeloid

precursors, provoking a predisposition for AML development. Indeed, *MYBL2* haploinsufficiency represents a pre-disposition for multiple myeloid disorders, such as myeloid leukemia in aging mice [129]. Moreover, the AML patients carrying the rs73116571 polymorphism, whom showed lower *MYBL2* expression levels, tended to achieve higher survival. These findings are consistent with earlier studies on different tumor types showing that *MYBL2* over-expression is related to an adverse prognosis [114–117, 119–123, 128]. Hence, these findings suggest the rs73116571 polymorphism could pre-dispose to the AML, but also could improve the prognosis once the disease has been developed. In fact, it is known that *MYBL2* expression is proportional to the degree of cell proliferation [286, 287] and that *MYBL2* over-expression alters differentiation processes in myeloid precursors [127]. In addition, *MYBL2* could be regulated by microRNAs [118, 288]. Accordingly, previous studies have demonstrated that microRNAs from the family 30 were down-regulated in patients over-expressing *MYBL2* [128].

In summary, this thesis suggests that *MYBL2* over-expression caused by a dysregulation of compensatory mechanisms, such as microRNAs, could induce uncontrolled cell proliferation in AML and, in this case, the rs73116571 polymorphism would represent a protective variation. However, a large series of patients should be tested in order to evaluate the contribution of this polymorphism to the prognosis of AML.

Finally, this study has found four genetic alterations in the *MYBL2* coding sequence in individual AML patients that were not registered in SNP databases. E132D was classified as a benign genetic variation by the PolyPhen network software. In contrast, Q67X generates a STOP codon in exon 4 at R1, the first DNA binding domain. This domain is highly conserved in the MYB family transcription factors and it contains helix-turn-helix motifs and serves to stabilize the DNA/protein complex. Therefore, these results suggest that Q67X may represent a malignant genetic variation which may result in the truncation of the protein or the mutant transcript undergoing nonsense mediated decay, reducing MYBL2 levels. In fact, the patient carrying Q67X showed lower *MYBL2* expression levels. However, this patient died soon after the date of diagnosis due to complications unrelated to AML. Therefore, despite the interest in verifying if Q67X represented a somatic or germline genetic variation and in discarding the possibility of being a non-annotated SNP, it was not possible to obtain CR samples or samples from family members to evaluate these issues. However, none of the novel

genetic alterations identified (including Q67X) was registered in whole-genome or whole-exome sequencing databases (as Leiden Open Variation Database [LOVD] or The Catalog of Somatic Mutations in Cancer [COSMIC]), neither in data sets of the study published by the Cancer Genome Atlas (TCGA) Research Network, which analyzed 200 adult cases of de novo AML using either whole-genome or whole-exome sequencing [289]. Indeed, the TCGA datasets of somatic mutations in AML only reported one mutation in *MYBL2* gene that represented a silent mutation (NM_002466.2:c.1962G > A; p.T654), which was not detect in the series of patients from this thesis.

3. *MYBL2* functional studies

Besides the study of polymorphisms and mutations that could alter *MYBL2* expression, the thesis aimed to inquire into the changes in the development and normal gene expression of human HP as a consequence of such *MYBL2* aberrant expression, which could contribute to the mechanism of leukemogenesis. *MYBL2* overexpression has been observed in cancers such as neuroblastoma, liver carcinoma, cutaneous T lymphoma, prostate, colorectal, testicular or ovarian malignancies, promoting cancer cell survival and proliferation [93, 95, 114–124, 290]. In addition, recent studies have shown *MYBL2* overexpression is related to an adverse prognosis in AML patients [128]. The aforementioned studies suggest that *MYBL2* is a proto-oncogene as its aberrant expression would support tumor development, by promoting cell cycle progression (enhancing the G1/S transition) and/or resistance to apoptosis. Accordingly, the results obtained in the thesis from the human HP transfections showed the overexpression of *MYBL2* generated quickly CFC proliferation since the number of CFU-M and CFU-GM was higher than that obtained transfecting the HP with a control. Moreover, after inducing the overexpression of *MYBL2* in HP, the integrated analysis of Gene Ontology (Partek) showed the genes which resulted in dysregulation were mainly involved in metabolism and response to stimulus processes. The functions attributed to most of these genes in previous publications and databases contributed to the enhancing of biological processes as cell adhesion, migration, inflammation, immunity, central nervous and embryonic development. In addition, the integrated analysis of Cellular Pathways (Partek) showed the “Hematopoietic cell lineage” resulted among the pathways most significantly altered in $CD34^+$ cells after *MYBL2* overexpression. In the

hypothetical event that HP had been transformed into leukemic blasts, these findings could suggest the maintenance of a proliferative or immature condition that would contribute to the progression of the leukemic disease. In this regard, a recent study has shown *MYBL2* not only controls and accelerates cell cycle progression in embryonic stem cells since it also guides cell fate decisions towards maintain self-renewal and the pluripotent stem cell identity [291].

By contrast, the results obtained transfecting the HP with *MYBL2* specific siRNAs, clearly showed the downregulation of *MYBL2* affected genes which mostly participate in processes related to cell proliferation and apoptosis control, inducing the overexpression of several genes such as *p53*, *BAX*, *CASP8*, *ACIN1*, *CHEK1/2*, *RB1* or *CDKN1B*, which enhance apoptosis, chromosomal aberrant condensation and the arrest of the cell cycle. In agreement, the HP transfected with *MYBL2* specific siRNAs proliferated less than those HP transfected with the siRNAs control, since the number of CFU-M and CFU-GM obtained was significantly lower after *MYBL2* silencing. These findings support other previous studies which demonstrated that *MYBL2* is also essential in G2/M transition and maintenance of genomic stability [104, 105]. García et al. showed, in megakaryoblast cell lines, that *MYBL2* was essential for cell cycle progression into S phase as well as in genomic stability. The reduction of *MYBL2* expression represented a decrease in S phase progression, arrest in mitosis, chromosomal aberrant condensation and chromosomal fragmentation [105]. Other studies support the hypothetical model of G2/M gene regulation by *MYBL2* is based on a repressive complex, which includes *MYBL2*, *E2F* and the retinoblastoma family member *RBL1* (retinoblastoma-like protein 1, also called *p107*), that is transformed in a transcriptionally active complex in late G1/S by *RBL1* dissociation [95]. Accordingly, the results obtained in the thesis showed the transcription factor *E2F4*, which binds to all three of the tumor suppressor proteins RB1, p107 and p130, and also the *RB1* gene were up-regulated after *MYBL2* silencing in HP. *E2F4* plays an important role in the suppression of proliferation-associated genes, and its gene mutation and increased expression may be associated with human cancer [207].

Moreover, several genes that induce cancer transformation, such as *IGF1R* [292-295], were upregulated after the downregulation of *MYBL2* in HP. In fact, the "Transcriptional Misregulation in Cancer", "AML", "CML", and other cancer pathways resulted in the most significantly altered pathways after *MYBL2* silencing in HP. Indeed,

it is known that *MYBL2* haploinsufficiency represents a predisposition for multiple myeloid disorders, such as myeloid leukemia in aging mice [129]. Also a recent study showed *MYBL2* is expressed at sharply reduced levels in CD34+ cells from most MDS cases, suggesting downregulation of *MYBL2* activity underlies the clonal expansion of hematopoietic progenitors in a large fraction of human myeloid malignancies. This study demonstrated mice developed a clonal myeloproliferative/myelodysplastic disorder which originated from cells with aberrantly reduced *MYBL2* expression [130].

The results from this thesis also showed the downregulation of *MYBL2* in HP induced higher expression of specific genes that promotes the DNA repair as *RAD23A*, *ERCC6L2*, *PRKDC*, *MRE11A* or *SMG1*, which is controversial with the suggestions that *MYBL2* could prevent the accumulation of random secondary mutations and genomic damage [129]. In this regard, a previous study showed that in clinical specimens from different stages of several human tumors, the early precursor lesions, but not normal tissues, commonly express markers of an activated DNA damage response. This study concluded that early in tumorigenesis (before genomic instability and malignant conversion), human cells activate a DNA damage response network that regulates key substrates involved in DNA repair, cell cycle checkpoints, genomic integrity, DNA-replication origin firing and replication-fork stability, in order to delay or prevent cancer [296]. Moreover, another study demonstrated the chromosomes that missegregate are frequently damaged during cytokinesis, triggering a DNA double-strand break response in the respective daughter cells involving *ATM* (*Ataxia-Telangiectasia Mutated gene*, a member of the phosphatidylinositol 3-kinase family that respond to DNA damage by phosphorylating substrates involved in DNA repair and/or cell cycle control), *CHEK2* and *p53* [297]. This study showed that these double-strand breaks can lead to unbalanced translocations in the daughter cells, concluding that segregation errors can cause translocations and provide insights into the role of whole-chromosome instability in tumorigenesis. Accordingly to both last aforementioned studies, the results obtained in this thesis, showing the overexpression of genes related to DNA damage response after *MYBL2* silencing, suggest the aberrant downregulation of *MYBL2* could contribute to a first step of leukemogenesis in human HP.

In addition, the results obtained from the siRNA transfections of HP were in agreement with the emerging theory that *MYBL2* blocks senescence [112, 113]. Cellular senescence, originally described as a form of cell aging, is a stable cell cycle arrest

caused by insults including telomere erosion, oncogene activation, irradiation, DNA damage, oxidative stress, viral infection and toxins. Senescence is thought to act as a tumor suppressor mechanism during the reproductive age, involving *CDKN2A* (*cyclin-dependent kinase inhibitor 2A* or *p16–INK4A*), *RB1* (*retinoblastoma 1* gene), *TP53* (*tumor protein p53*) and *CDKN1A* (*cyclin-dependent kinase inhibitor 1A*, *p21* or *Cip1*) tumor suppressor pathways; but it is not known what makes the senescence arrest stable and what the critical downstream targets are. Mowla et al supported a strong evidence to indicate that *MYBL2* is inhibited during senescence, suggesting *MYBL2* impedes cellular senescence [113]. In fact, the results obtained in this thesis showed the genes underlying tumor suppressor pathways as *RB1*, *TP53* or *CDKN1B* (c*yclin-dependent kinase inhibitor 1B* [*p27, Kip1*]) were up-regulated in HP after the transfection with specific siRNAs, which would contribute to senescence since *MYBL2* would be silenced.

Finally, the results obtained from this thesis agree with other studies showing gene targets of *MYBL2* such as *CLU* (anti-apoptotic *apolipoprotein–J), cyclin–D1 (CCND1) and cyclin–A1 (CCNA1)* [101, 102], since whose expression decreased after *MYBL2* silencing in HP. However, the anti-apoptotic *BCL2* gene, also described as a *MYBL2* target gene [95], maintained higher expression after *MYBL2* silencing. Regarding previous studies, it has been demonstrated that BCL2 also could act as a pro-apoptotic protein when the BH3 domain of BAX death agonist binds with BCL2 [197, 298]. Indeed, the results showed *BAX* gene was higher expressed after *MYBL2* inhibition, suggesting the BH3 domain of BAX could have been binding BCL2, thus promoting apoptosis.

Interestingly, several genes, such as *MLLT4*, *BCL2L11*, *MAP2K7*, *CHEK2*, *CDKN1B* and *WNT8A* were inversely expressed in HP after the induction of *MYBL2* overexpression compared to its expression in those HP after *MYBL2* silencing. Thus, these findings could suggest the expression of these genes relies on its regulation by *MYBL2*. In fact, the UCSC Genome Browser database (http://genome.ucsc.edu/) showed several “ChIP-seq peaks” or MYBL2 binding domains in the promoters and other sequence regions of *MAP2K7*, *MLLT4*, *BCL2L11* and *CHEK2* genes, suggesting a direct MYBL2 interaction. *MLLT4*, *BCL2L11* and *WNT8A* genes turned over-expressed in $CD34^+$ cells transfected with MYBL2–IRES–EGFP overexpression plasmid and turned lower-expressed after *MYBL2* silencing with specific siRNAs. *MLLT4* is a Ras oncogene family target that regulates cell-cell adhesions downstream of Ras activation

during embryogenesis. *MLLT4* has also been identified as the fusion partner of *KMT2A* (*MLL*) gene involved in AML with t(6;11)(q27;q23) translocation [241–245]. The overexpression of *MLLT4* after *MYBL2* up-regulation in HP suggests the enhancing of a proliferative condition in these cells. However, the apoptosis facilitator gene *BCL2L11* also resulted in up-regulated expression levels after the induction of *MYBL2* expression and was down-regulated after *MYBL2* silencing in HP. This last result is controversial with the suggestions about the down-regulation of *MYBL2* induces apoptosis in HP. Conversely, *BCL2L11* appears also to have an important role in embryonic development, hematopoietic homeostasis and as a barrier to autoimmunity, since studies in transgenic mice suggested this gene acts as an essential initiator of apoptosis in the negative selection of auto-reactive T cells during thymic development [299]. Hence, the results obtained from the HP transfections could suggest the apoptotic mechanism induced by *BCL2L11* after *MYBL2* overexpression maybe more closely related to the protection of the HP development against autoimmunity processes.

The aforementioned *WNT8A* is a member of the *WNT* gene family, and also may be implicated in development of early embryos as well as germ cell tumors[258]. By contrast, the genes *MAP2K7*, *CHEK2* and *CDKN1B* turned lower-expressed in CD34+ cells transfected with *MYBL2*–IRES–EGFP overexpression plasmid and turned over-expressed after *MYBL2* silencing with specific siRNAs. As previously mentioned, *CDKN1B* is a cell cycle inhibitor and tumor suppressor that blocks the cell cycle in the G0/G1 phase upon differentiation signals or cellular insults [300]. Also, both *MAP2K7* and *CHEK2* genes exert a tumor suppressive function through *p53* in response to DNA damage, leading to cell cycle arrest and apoptosis [150, 202, 203]. In addition, *CHEK2*, a cell cycle checkpoint regulator, stimulates the transcription of genes involved in DNA repair [200, 201] and interacts with and phosphorylates *BRCA1*, allowing *BRCA1* to restore survival after DNA damage [301]. Interestingly, the *BRCA1* gene was also over-expressed in HP after *MYBL2* silencing. *BRCA1* coordinates a diverse range of cellular pathways such as DNA damage repair, recombination, ubiquitination and transcriptional regulation to maintain genomic stability and it also acts as a tumor suppressor [302, 303]. The encoded protein combines with other tumor suppressors, DNA damage sensors, and signal transducers to form a large multi-subunit protein complex known as the *BRCA1*-associated genome surveillance complex (BASC) [186]. Mutations compromising this

network, including defects in the ATM-CHK2-p53 pathway, might allow cell proliferation, survival, increased genomic instability and tumor progression [296].

VII. CONCLUSIONS

The main conclusions of the work presented in this thesis are:

1) The novel qRT−PCR method for simultaneous detection of multiple AML-associated fusion genes is a versatile and sensitive assay for reliable screening of recurrent AML rearrangements. In addition, the applicability of this method in diagnostic molecular laboratories is supported by the efficiency, simplicity, and rapidity of the procedure, which productively complements cytogenetic and FISH analyses.

2) The S427G polymorphism of *MYBL2* is neither associated with the risk of developing AML, nor does it represent a prognostic factor for the patients.

3) The HRM method proves to be a useful and efficient technique to detect genetic alterations of *MYBL2* in AML patients.

4) The rs73116571 polymorphism was associated with weaker *MYBL2* mRNA expression and higher incidence in AML patients, which could suggests the polymorphism causes a pre-disposition to AML.

5) The novel Q67X alteration could modify MYBL2 activity in AML patients by truncating the protein or resulting in the mutant transcript undergoing nonsense mediated decay, therefore, reducing MYBL2 levels.

6) Lower *MYBL2* expression levels in human HP cells strongly affected the maintenance of cellular development, enhancing genetic mechanisms of apoptosis and cell cycle arrest, which could lead to HP instability and AML development.

7) Higher *MYBL2* expression levels in human HP cells increased cell proliferation and the expression of genes involved in embryonic development, suggesting that this condition could maintain a sustained cell survival and a resistance to mechanisms of programmed cell death in abnormal-cancerous cells.

8) The genes *MLLT4, BCL2L11, MAP2K7, CHEK2, CDKN1B* and *WNT8A* were inversely expressed in HP according to *MYBL2* expression, suggesting a direct regulation by a MYBL2 interaction.

VIII. REFERENCES

1. Kondo M, Wagers AJ, Manz MG et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application.Annu Rev Immunol 2003; 21: 759-806.

2. Krause DS, Fackler MJ, Civin CI et al. CD34: structure, biology, and clinical utility. Blood. 1996; 87:1-13.

3. Irons RD1, Stillman WS. The process of leukemogenesis. Environ Health Perspect. 1996; 104:1239-46.

4. Riether C, Schürch CM, Ochsenbein AF. Regulation of hematopoietic and leukemic stem cells by the immune system. Cell Death Differ. 2015; 22:187-98.

5. González Barón M. Factores pronósticos en la leucemia aguda. En: R. de la Cámara Llanza, A. Figuera Álvarez, J.M. Fernández Rañada. Factores pronósticos en oncología. Edición Mc.Graw-Hill Interamericana de España, Madrid, 1994: 428-468.

6. Jemal A, Thomas A, Murray T et al. Cancer Statistics 2002. CA Cancer J Clin, 2002. 52:23.

7. Visser O, Trama A, Maynadié M et al; RARECARE Working Group. Incidence, survival and prevalence of myeloid malignancies in Europe. Eur J Cancer. 2012; 48: 3257-66.

8. Greenlee RT, Hill-Harmon MB, Murray T et al: Cancer statistics 2001 .CA Cancer J Clin 2001; 51: 15-36.

9. Vardiman JW, Thiele J, Arber DA et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 2009, 114: 937-951.

10. Betz BL, Hess JL. Acute myeloid leukemia diagnosis in the 21st century. Arch Pathol Lab Med. 2010; 134: 1427-33.

11. Mrózek K, Heerema NA, Bloomfield CD. Cytogenetics in acute leukemia. Blood Rev. 2004; 18:115-136.

12. Mrózek K, Heinonen K, De la Chapelle A et al. Clinical Significance of cytogenetics in acute myeloid leukaemia. Semin Oncol 1997; 24: 17-31.

13. Grimwade D, Walker H, Oliver F et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children´s Leukaemia Working parties. Blood 1998; 92: 2322-2333.

14. Grimwade D, Walker H, Harrison G et al. The predictive value of hierarchical cytogenetic classification in older adults with acute myeloid leukemia (AML): analysis of 1065 patients entered into the United Kingdom Medical Research Council AML11 trial. Blood. 2001; 98: 1312-1320.

15. Grimwade D. Impact of cytogenetics on clinical outcome in AML. In: Karp JE, ed. Acute Myelogenous Leukemia. Totowa, New Jersey: Humana Press; 2007: 177-192.

16. Leder P, Battey J, Lenoir G et al. Translocations among antibody genes in human cancer. Science 1983; 222:765-71.

17. Finger LR, Harvey RC, Moore RC et al. A common mechanism of chromosomal translocation in T- and B-cell neoplasia. Science 1986; 234: 982-5.

18. Miyoshi H, Kozu T, Shimizu K et al. The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. EMBO J 1993; 12:2715.

19. Meyers S, Lenny N, Hiebert SW. The t(8;21) fusion protein interferes with AML-1B-dependent transcriptional activation. Mol Cell Biol 1995; 15:1974.

20. Parry P, Wei Y, Evans G. Cloning and characterization of the t(X;11) breakpoint from a leukemic cell line identify a new member of the forkhead gene family. Genes Chromosom Cancer 1994; 11:79.

21. Prasad R, Gu Y, Alder H et al. Cloning of the ALL-1 fusion partner, the AF-6 gene, involved in acute myeloid leukemias with the t(6;11) chromosome translocation. Cancer Res 1993; 53:5624.

22. Thirman MJ, Levitan DA, Kobayashi H et al. Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23; p13.1) in acute myeloid leukemia. Proc Natl Acad Sci USA 1994; 91:12110.

23. Bernard OA, Mauchauffe M, Mecucci C et al. A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not related to AF-4, AF-9 nor ENL. Oncogene 1994; 9:1039.

24. Prasad R, Leshkowitz D, Gu Y et al. Leucine-zipper dimerization motif encoded by the AF17 gene fused to ALL-1 (MLL) in acute leukemia. Proc Natl Acad Sci USA 1994; 91:8107.

25. Chaplin T, Ayton P, Bernard OA et al. A novel class of zinc finger/leucine zipper genes identified from the molecular cloning of the t(10;11) translocation in acute leukemia. Blood 1995; 85:1435.

26. Tse W, Zhu W, Chen HS et al. A novel gene, AF1q, fused to MLL in t(1;11) (q21;q23), is specifically expressed in leukemic and immature hematopoietic cells. Blood 1995; 85:650.

27. Yamamoto K, Seto M, Iida S et al. A reverse transcriptase-polymerase chain reaction detects heterogeneous chimeric mRNAs in leukemias with 11q23 abnormalities. Blood 1994; 83:2912.

28. Schichman SA, Caligiuri MA, Gu Y et al. ALL-1 partial duplication in acute leukemia. Proc Natl Acad Sci USA 1994; 91:6236.

29. J Gabert, E Beillard, VJH van der Velden et al. Standardization and quality control studies of “real-time” quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukaemia- A Europe Against Cancer Program. Leukemia, 2003; 1-40.

30. Yin JA, Tobal K. Detection of Minimal Residual Disease in acute myeloid leukemia: Methodologies, Clinical and Biological significance. British Journal of Haematology, 1999; 106: 578-590.

31. Pallisgaard N, Hokland P, Riishøj DC et al. Multiplex reverse transcription-polymerase chain reaction for simultaneous screening of 29 translocations and chromosomal aberrations in acute leukemia. Blood 1998, 92:574-588.

32. Elia L, Mancini M, Moleti L et al. Multiplex reverse transcriptase-polymerase chain reaction strategy for the diagnostic molecular screening of chimeric genes: a clinical evaluation on 170 patients with acute lymphoblastic leukemia. Haematologica 2003, 88:275-279.

33. Song MJ, Kim HJ, Park CH et al. Diagnostic utility of a multiplex RT-PCR assay in detecting fusion transcripts from recurrent genetic abnormalities of acute leukemia by WHO 2008 classification. Diagn Mol Pathol 2012, 21:40-44.

34. Salto-Tellez M, Shelat SG, Benoit B et al. Multiplex RT-PCR for the detection of leukemia-associated translocations: validation and application to routine molecular diagnostic practice. J Mol Diagn. 2003, 5:231-236.

35. King RL, Naghashpour M, Watt CD et al. A comparative analysis of molecular genetic and conventional cytogenetic detection of diagnostically important translocations in more than 400 cases of acute leukemia, highlighting the frequency of false-negative conventional cytogenetics. Am J Clin Pathol. 2011, 135:921-928.

36. Lin MT, Tseng LH, Rich RG et al. Δ-PCR, A Simple Method to Detect Translocations and Insertion/Deletion Mutations. J Mol Diagn. 2011, 13:85-92.

37. Garcia EP, Dowding LA, Stanton LW et al. Scalable transcriptional analysis routine--multiplexed quantitative real-time polymerase chain reaction platform for gene expression analysis and molecular diagnostics. J Mol Diagn. 2005, 7:444-454.

38. Syrmis MW, Whiley DM, Thomas M et al. A sensitive, specific, and cost-effective multiplex reverse transcriptase-PCR assay for the detection of seven common respiratory viruses in respiratory samples. J Mol Diagn. 2004, 6:125-131.

39. Van Dongen JJ, Macintyre EA, Gabert JA et al. Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: investigation of minimal residual disease in acute leukemia. Leukemia 1999, 13:1901-1928.

40. Yin CC, Cortes J, Barkoh B et al. t(3;21)(q26;q22) in myeloid leukemia: an aggressive syndrome of blast transformation associated with hydroxyurea or antimetabolite therapy. Cancer 2006, 106:1730-1738.

41. Chang WR, Park IJ, Lee HW et al. Two cases of acute myeloid leukemia with t(16;21)(p11;q22) and TLS/FUS-ERG fusion transcripts. Korean J Lab Med 2009, 29:390-395.

42. Velloso ER, Mecucci C, Michaux L et al. Translocation t(8;16)(p11;p13) in acute non-lymphocytic leukemia: report on two new cases and review of the literature. Leuk Lymphoma 1996, 21:137-142.

43. Rozman M, Camós M, Colomer D et al. Type I MOZ/CBP (MYST3/CBP) is the most common chimeric transcript in acute myeloid leukemia with t(8;16)(p11;p13) translocation. Genes Chromosomes Cancer 2004, 40:140-145.

44. Jovanovic JV, Rennie K, Culligan D et al. Development of real-time quantitative polymerase chain reaction assays to track treatment responsein retinoid resistant acute promyelocytic leukemia. Front Oncol. 2011, 1:35.

45. Bennett JM, Catovsky D, Daniel MT et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. Br J Haematol 1976, 33:451-458.

46. Y.-Y. Lai, J.-Y. Qiu, B. Jiang et al. Characteristics and prognostic factors of acute myeloid leukemia with t(8;21) (q22; q22). Zhongguo Shi Yan Xue Ye Xue Za Zhi, vol. 13, no. 5, pp. 733-740, 2005.

47. Swirsky DM, Li YS, Matthews JG et al. 8;21 translocation in acute granulocytic leukaemia: cytological, cytochemical and clinical features. Br J Haematol 1984; 56: 199-213.

48. Heisterkamp N, Groffen J. Molecular insights into the Philadelphia translocation. Hematol Pathol. 1991; 5(1):1-10.

49. Biondi A1, Rambaldi A, Alcalay M et al. RAR-alpha gene rearrangements as a genetic marker for diagnosis and monitoring in acute promyelocytic leukemia. Blood. 1991 Apr 1;77(7):1418-22.

50. Dan Douer et al. British Journal of Haematology, 2003; 122: 563.

51. Sanz MA, Martin G, González M et al. Programa de Estudio y Tratamiento de las Hemopatías Malignas. Risk-adapted treatment of acute promyelocytic leukemia with All-trans-retinoic acid and anthracycline monochemotherapy: a multicenter study by the PETHEMA group. Blood. 2004, 15; 103: 1237-43.

52. Sanz MA, Diverio D, Barragán E et al. Progresos diagnósticos y terapéuticos en la leucemia promielocítica aguda. En: Leucemias y progresos biológicos y terapéuticos. Biblioteca Oncológica Roche. Ed You & Us 1998. 85-102.

53. Chen, Z., Brand, N., Chen, A. et al. Fusion between a novel Krüppel-like zinc finger gene and the retinoic acid receptor-alpha locus due to a variant t(11;17) translocation associated with acute promyelocytic leukaemia. EMBO J. 1993, 12, 1161-1167.

54. Laity, J. H., Lee, B. M., and Wright, P. E. Zinc finger proteins: new insights into structural and functional diversity. Curr. Opin. Struct. Biol. 2001, 11, 39-46.

55. Larson RA, Williams SF, LeBeau MM et al. Acute myelomonocytic leukemia with abnormal eosinophils and inv(16) and t(16;16) has a favourable prognosis. Blood 1986; 68:1242-9.

56. Hélène Poirel, Katrina Rack, Eric Delabesse et al. Incidence and Characterization of MLL Gene (11q23) Rearrangements in Acute Myeloid Leukemia M1 and M5. Blood, 1996; 87: 2496-2505.

57. Thorsten Langer, Markus Metzler, Dirk Reinhardt et al. Analysis of t(9;11) Chromosomal Breakpoint Sequences in Childhood Acute Leukemia: Almost Identical MLL Breakpoints in Therapy-Related AML After Treatment Without Etoposides. Genes, Chromosomes and Cancer, 2003; 36: 393-401.

58. Kohmei Ida, Tomohiko Taki, Fumio Besshoet al. Detection of Chimeric mRNAs by Reverse Transcriptase-Polymerase Chain Reaction for Diagnosis and Monitoring of Acute Leukemias With 11q23 Abnormalities. Medical and Pediatric Oncology, 1997; 28: 325-332.

59. Heidi Gill Super, Pamela L. Strissel, Olatoyosi M. Sobulo et al. Identification of Complex Genomic Breakpoint Juntions in the t(9;11) MLL-AF9 Fusion Gene in Acute Leukemia. Genes, Chromosomes and Cancer, 1997; 20: 185-195.

60. Jeffrey E. Rubnitz, Frederick G et al. Molecular Analysis of t(11;19) Breakpoints in Childhood Acute Leukemias. Blood, 1996; 87: 4804-4808.

61. B Hjorth-Sorensen, N Pallisgaard, M Grönholm et al. A novel MLL-AF10 fusion mRNA variant in a patient with acute myeloid leukemia detected by a new asymmetric reverse transcription PCR method. Leukemia, 1997; 11: 1588-1593.

62. Nucifora G. The EVI1 gene in myeloid leukemia. Leukemia 1997, 11: 2022-2031.

63. Vitalyi Senyuk, Soumen Chakraborty, Fady M Mikhail et al. The leukemia -associated transcription repressor AML1/MDS1/EVI1 requires CtBP to induce abnormal growth and differentiation of murine hematopoietic cells. Oncogene, 2002; 21: 3232-3240.

64. Grace M Cuenco and Ruibao Ren. Both AML1 and EVI1 oncogenic components are required for the cooperation of AML1/MDS1/EVI1 with BCR/ABL in the introduction of acute myelogenous leukemia in mice. Oncogene, 2004; 23: 569-579.

65. Vitalyi Senyuk, Kislay K. Sinha, Donglan Li et al. Repression of RUNX1 Activity by EVI1: A New Role of EVI1 in Leukemogenesis. Cancer Res, 2007; 67: 12.

66. Francesco Lo Coco, Simona Pisegna, Daniela Diverio. The AML1 gene: a transcription factor involved in the pathogenesis of myeloid and lymphoid leukemias. Haematologica, 1997; 82: 364-370.

67. Giussepina Nucifora, Catherine R. Begy, Hirofumi Kobayashi et al. Consistent intergenic splicing and production of multiple transcripts between AML1 at 21q22 and unrelated genes at 3q26 in (3;21)(q26;q22) translocations. Medical Sciences, 1994; 91: 4004-4008.

68. Marieke von Lindern, Maarten Fornerod, Sjozèf van Baal et al. The Translocation (6;9), Associated with a Specific Subtype of Acute Myeloid Leukemia, Results in the Fusion of Two Genes, dek and can, and the Expression of a Chimeric, Leukemia-Specific dek-can mRNA. Molecular and Cellular Biology, 1992; 12: 1687-1697.

69. Ma Z, Morris SW, Valentine V et al. Fusion of two novel genes, RBM15 and MKL1, in the t(1 ; 22)(p13 ; q13) of acute megakaryoblastic leukemia. Nat Genet 2001; 28: 220-221.

70. Mercher T, Coniat MB, Monni R et al. Involvement of a human gene related to the Drosophila spen gene in the recurrent t(1 ; 22) translocation of acute megakaryocytic leukemia. Proc Natl Acad Sci USA 2001; 98: 5776-5779.

71. Ioannis Panagopoulos, Margareth Isaksson, Charlota Lindvall et al. RT-PCR Analysis of the MOZ-CBP and CBP-MOZ Chimeric Transcripts in Acute Myeloid Leukemias With t(8;16)(p11;p13). Genes, Chromosomes and Cancer, 2000; 28: 415-424.

72. Terui K, Sato T, Sasaki S et al. Two novel variants of MOZ/CBP fusion transcripts in spontaneously remitted infant leukemia with t(1;16;8)(p13;p13;p11), a new variant of t(8;16)(p11;p13). Haematologica, 2008; 93: 1591-1593.

73. Murati A, Adélaïde J, Quilichini B et al. New types of MYST3-CBP and CBP-MYST3 fusion transcripts in t(8;16)(p11;p13) acute myeloid leukemias. Haematologica, 2007; 92: 262-263.

74. Panagopoulos I, Fioretos T, Isaksson M et al. RT-PCR Analysis of Acute Myeloid Leukemia With t(8;16)(p11;p13): Identification of a Novel MOZ/CBP Transcript and Absence of CBP/MOZ Expression. Genes, Chromosomes and Cancer, 2002; 35: 372-374.

75. Murakami, K., Mavrothalassitis, G., Bhat, N. K. et al. Human ERG-2 protein is a phosphorylated DNA-binding protein--a distinct member of the ETS family. Oncogene 1993; 8: 1559-1566.

76. Panagopoulos I, Gorunova L, Zeller B et al. Cryptic FUS-ERG fusion identified by RNA-sequencing in childhood acute myeloid leukemia. Oncol Rep. 2013; 30:2587-92.

77. Glliland, D. G. Molecular genetics of human leukemias: new insights into terapy (review). Semin. Hematol. 2002; 39:6-11.

78. Wichmann C, Quagliano-Lo Coco I, Yildiz Ö et al. Activating c-KIT mutations confer oncogenic cooperativity and rescue RUNX1/ETO-induced DNA damage and apoptosis in human primary CD34+ hematopoietic progenitors. Leukemia. 2015; 29: 279-89.

79. Song J1, Mercer D, Hu X et al. Common leukemia- and lymphoma-associated genetic aberrations in healthy individuals. J Mol Diagn. 2011; 13: 213-9.

80. Moreno I, Martín G, Bolufer P et al. Incidence and prognostic value of FLT3 internal tandem duplication and D835 mutations in acute myeloid leukemia. Haematologica. 2003; 88: 19-24.

81. Thiede C, Koch S, Creutzig E et al. Prevalence and prognostic impact of NPM1 mutations in 1485 adult patients with acute myeloid leukemia (AML). Blood. 2006; 107:4011-20.

82. Bacher U, Schnittger S, Haferlach T. Molecular genetics in acute myeloid leukemia. Curr Opin Oncol. 2010; 22: 646-55.

83. Ley TJ, Ding L, Walter MJ et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010; 363: 2424-33.

84. Yamashita Y, Yuan J, Suetake I et al. Array-based genomic resequencing of human leukemia. Oncogene. 2010; 29: 3723-31.

85. Mardis ER, Ding L, Dooling DJ et al. Recurring mutations found by sequencing an acute myeloid leukemia genome. N Engl J Med. 2009; 361: 1058-66.

86. Barragán E, Cervera J, Bolufer P et al. Prognostic implications of Wilms' tumor gene (WT1) expression in patients with de novo acute myeloid leukemia. Haematologica. 2004; 89: 926-33.

87. Döhner H1, Estey EH, Amadori S et al; European LeukemiaNet. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood. 2010; 115: 453-74.

88. Mrózek K1, Marcucci G, Nicolet D et al. Prognostic significance of the European LeukemiaNet standardized system for reporting cytogenetic and molecular alterations in adults with acute myeloid leukemia. J Clin Oncol. 2012; 30: 4515-23.

89. Bullinger L, Kronke J, Schon C et al. Identification of acquired copy number alterations and uniparental disomies in cytogenetically normal acute myeloid leukemia using high-resolution single-nucleotide polymorphism analysis. Leukemia. 2010; 24: 438-449.

90. Suela J, Alvarez S, Cigudosa JC. DNA profiling by array CGH in acute myeloid leukemia and myelodysplastic syndromes. Cytogenet Genome Res. 2007; 118: 304-309.

91. Walter MJ, Payton JE, Ries RE et al. Acquired copy number alterations in adult acute myeloid leukemia genomes. Proc Natl Acad Sci U S A. 2009; 106: 12950-12955.

92. Shen Y, Zhu YM, Fan X et al. Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. Blood. 2011; 118: 5593-5603.

93. Chayka, O., Sala, A. MYBL2 (v-myb myeloblastosis viral oncogene homolog (avian)-like 2) Atlas Genet Cytogenet Oncol Haematol. 2009; 13: 652-653.

94. Oh IH, Reddy EP. The myb gene family in cell growth, differentiation and apoptosis. Oncogene 1999; 18: 3017-3033.

95. Sala A. B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer. Eur J Cancer 2005; 41: 2479-2484.

96. Davidson CJ, Tirouvanziam R, Herzenberg LA et al. Functional evolution of the vertebrate Myb gene family: B-Myb, but neither A-Myb nor c-Myb, complements Drosophila Myb in hemocytes. Genetics 2005; 169: 215-229.

97. Toscani A, Mettus RV, Coupland R et al. Arrest of spermatogenesis and defective breast development in mice lacking A-myb. Nature 1997; 386: 713-717.

98. Sandberg ML, Sutton SE, Pletcher MT et al. c-Myb and p300 regulate hematopoietic stem cell proliferation and differentiation. Dev Cell 2005; 8: 153-166.

99. Tanaka Y, Patestos NP, Maekawa T et al. B-Myb is required for inner cell mass formation at an early stage of development. J. Biol. Chem 1999; 274: 28067-28070.

100. Mucenski ML, McLain K, Kier AB et al. A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. Cell 1991; 65:677-689.

101. Lang, G., Gombert, W. M., Gould, H. J. A transcriptional regulatory element in the coding sequence of the human Bcl-2 gene. Immunology 2005; 114: 25-36.

102. Horstmann S1, Ferrari S, Klempnauer KH. Regulation of B-Myb activity by cyclin D1. Oncogene. 2000;19: 298-306.

103. Lam EW, Watson RJ. An E2F-binding site mediates cell-cycle regulated repression of mouse B-myb transcription. EMBO J. 1993;12: 2705-2713.

104. Tarasov KV, Tarasova YS, Tam WL et al. B-MYB is essential for normal cell cycle progression and chromosomal stability of embryonic stem cells. PLoS One 2008; 3:e2478. http://dx.doi.org/10.1371/journal.pone.0002478.

105. Garcia P, Frampton J. The transcription factor B-Myb is essential for S-phase progression and genomic stability in diploid and polyploidy megakaryocytes. J Cell Sci 2006; 119:1483-1493.

106. Fung SM, Ramsay G, Katzen AL. Mutations in Drosophila myb lead to centrosome amplification and genomic instability. Development 2002; 129:347-359.

107. Manak JR, Mitiku N, Lipsick JS. Mutation of the Drosophila homologue of the Myb protooncogene causes genomic instability. Proc Natl Acad Sci USA 2002; 99:7438-7443.

108. Shepard JL, Amatruda JF, Stern HM et al. A zebrafish bmyb mutation causes genome instability and increased cancer susceptibility. Proc Natl Acad Sci USA 2005; 102:13194-13199.

109. Lam EW, Bennett JD, Watson RJ. Cell-cycle regulation of human B-myb transcription. Gene1995; 160: 277-281.

110. Lorvellec M, Dumon S, Maya-Mendoza A et al. B-Myb is critical for proper DNA duplication during an unperturbed S phase in mouse embryonic stem cells. Stem Cells 2010; 28: 1751-1759.

111. Manak JR, Wen H, Van T et al. Loss of Drosophila Myb interrupts the progression of chromosome condensation. Nat Cell Biol 2007; 9: 581-587.

112. Huang Y, Wu J, Li R et al. B-MYB delays cell aging by repressing p16 (INK4alpha) transcription. Cell Mol Life Sci 2010; 68: 893-901.

113. Mowla SN1, Lam EW, Jat PS. Cellular senescence and aging: the role of B-MYB. Aging Cell. 2014; 13: 773-9.

114. Raschellà G, Cesi V, Amendola R et al. Expression of B-myb in neuroblastoma tumors is a poor prognostic factor independent from MYCN amplification. Cancer Res 1999; 59:3365-3368.

115. Forozan F, Mahlamäki EH, Monni O et al. Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data. Cancer Res 2000; 60:4519-4525.

116. Tanner MM, Grenman S, Koul A et al. Frequent amplification of chromosomal region 20q12-q13 in ovarian cancer. Clin Cancer Res 2000; 6:1833-1839.

117. Bar-Shira A, Pinthus JH, Rozovsky U et al. Multiple genes in human 20q13 chromosomal region are involved in an advanced prostate cancer xenograft. Cancer Res 2002; 62:6803-6807.

118. Zauli G, Voltan R, di Iasio MG et al. miR-34a induces the downregulation of both E2F1 and B-Myb oncogenes in leukemic cells. Clin Cancer Res 2011; 17:2712-2724.

119. Zondervan PE, Wink J, Alers JC et al. Molecular cytogenetic evaluation of virus-associated and non-viral hepatocellular carcinoma: analysis of 26 carcinomas and 12 concurrent dysplasias. J Pathol 2000; 192:207-215.

120. Mao X, Orchard G, Lillington DM et al. Amplification and overexpression of JUNB is associated with primary cutaneous T-cell lymphomas. Blood 2003; 101:1513-1519.

121. Skotheim RI, Monni O, Mousses S et al. New insights into testicular germ cell tumorigenesis from gene expression profiling. Cancer Res 2002; 62:2359-2364.

122. Raschellà G, Negroni A, Sala A et al. Requirement of b-myb function for survival and differentiative potential of human neuroblastoma cells. J Biol Chem 1995; 270:8540-8545.

123. Paik S, Shak S, Tang G et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med. 2004; 351:2817-2826.

124. Sala A, Watson R. B-Myb protein in cellular proliferation, transcription control, and cancer: latest developments. J Cell Physiol. 1999; 179: 245-50.

125. Nakajima T, Yasui K, Zen K et al. Activation of B-Myb by E2F1 in hepatocellular carcinoma. Hepatol Res. 2008; 38: 886-95.

126. Schwab R, Caccamo A, Bettuzzi S et al. B-MYB is hypophosphorylated and resistant to degradation in neuroblastoma: implications for cell survival. Blood Cells Mol Dis 2007; 39: 263-71.

127. Engelhard A, Campbell K, Calabretta B. B-Myb alters the response of myeloid precursor cells to G-CSF. Exp. Cell Re. 2000; 254:153-162.

128. Fuster O, Llop M, Dolz S et al. Adverse prognostic value of MYBL2 overexpression and association with microRNA-30 family in acute myeloid leukemia patients. Leuk Res. 2013; 37:1690-1696.

129. Clarke M, Dumon S, Ward C et al. MYBL2 haploinsufficiency increases susceptibility to age-related haematopoietic neoplasia. Leukemia 2013; 27: 661-670.

130. Heinrichs S1, Conover LF, Bueso-Ramos CE et al. MYBL2 is a sub-haploinsufficient tumor suppressor gene in myeloid malignancy. Elife. 2013; 2:e00825. doi: 10.7554/eLife.00825.

131. Schwab R, Bussolari R, Corvetta D et al. Isolation and functional assessment of common, polymorphic variants of the B-MYB proto-oncogene associated with a reduced cancer risk. Oncogene 2008; 27: 2929-2933.

132. Thorner AR, Hoadley KA, Parker JS et al. In vitro and in vivo analysis of B-Myb in basal-like breast cancer. Oncogene 2009; 28:742-751.

133. Shi H, Bevier M, Johansson R et al. Single nucleotide polymorphisms in the 20q13 amplicon genes in relation to breast cancer risk and clinical outcome. Breast Cancer Res Treat. 2011; 130:905-916.

134. Grimwade D, Hills RK, Moorman AV et al. Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials. Blood 2010; 116:354-365.

135. Thiede C, Steudel C, Mohr B et al. Anal-ysis of FLT3-activating mutations in 979 patients with acute myelogenous leukemia: association with FAB subtypes and identification of subgroups withpoor prognosis. Blood 2002; 99:4326-4335.

136. Schnittger S, Schoch C, Kern W et al. Nucleophosmin gene mutations are predictors of favorable prognosis in acute myelogenous leukemia with a normal karyotype. Blood 2005; 106:3733-3739.

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

138. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. Methods 2001; 25: 402-8.

139. Beillard E, Pallisgaard N, Bi W et al. Evaluation of candidate control genes for diagnosis and residual disease detection in leukemic patients using 'realtime' quantitative reverse-transcriptase polymerase chain reaction (RQ-PCR)-a Europe Against Cancer Program. Leukemia 2003;17: 2474-86.

140. Lozzio, C.B.; Lozzio, B.B. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. Blood 1975, 45 : 321-34.

141. Van den Hoff MJ1, Moorman AF, Lamers WH. Electroporation in 'intracellular' buffer increases cell survival. Nucleic Acids Res. 1992; 20: 2902.

142. Eva Gonzalez-Roca, Xabier Garcia-Albéniz, Silvia Rodriguez-Mulero et al. Accurate Expression Profiling of Very Small Cell Populations. PlosOne. 2010; 5:e14418.

143. Kaplan EL, Meier P. Nonparametric estimations from incomplete observations. J Am Stat Assoc 1958; 53:457-481.

144. Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. Cancer Chemother Rep 1966; 50:163-170.

145. Kögler, G; Callejas, J; Sorg, RV et al. The effect of different thawing methods, growth factor combinations and media on ex vivo expansion of umbilical cord blood primitive and committed progenitors. Bone Marrow Transplant. 1998; 21: 233-241.

146. de Wynter, EA; Buck, D; Hart, C et al. CD34+AC133+ Cells Isolated from Cord Blood are Highly Enriched in Long-Term Culture-Initiating Cells, NOD/SCID-Repopulating Cells and Dendritic Cell Progenitors. Stem Cells 1998; 16: 387-396.

147. Kerima Maasho, Alina Marusina, Nicole M. et al. Efficient gene transfer into the human natural killer cell line, NKL, using the amaxa nucleofection system. Journal of Immunological Methods 2004; 284:133-40.

148. Michela Aluigi, Miriam Fogli, Antonio Curti et al. Nucleofection is an efficient non-viral transfection technique for human bone marrow derived mesenchymal stem cells. Stem Cells 2006; 24: 454-461.

149. Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015; 43: D1049-56.

150. Schramek, D., Kotsinas, A., Meixner, A. et al. The stress kinase MKK7 couples oncogenic stress to p53 stability and tumor suppression. Nature Genet. 2011; 43: 212-219.

151. Foltz I.N., Gerl R.E., Wieler J.S. et al. Human mitogen-activated protein kinase kinase 7 (MKK7) is a highly conserved c-Jun N-terminal kinase/stress-activated protein kinase (JNK/SAPK) activated by environmental stresses and physiological stimuli. J. Biol. Chem. 1998; 273:9344-9351.

152. Walport LJ, Hopkinson RJ, Vollmar M et al. Human UTY(KDM6C) is a male-specific Nε-methyl lysyl demethylase. J Biol Chem. 2014; 289:18302-13.

153. Shpargel KB1, Sengoku T, Yokoyama S et al. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. PLoS Genet. 2012; 8:e1002964. doi: 10.1371/journal.pgen.1002964.

154. Tini M, Benecke A, Um SJ et al. Association of CBP/p300 acetylase and thymine DNA glycosylase links DNA repair and transcription. Mol Cell. 2002; 9:265-77.

155. Kung AL, Rebel VI, Bronson RT et al. Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. Genes Dev. 2000; 14: 272-7.

156. Panagopoulos I1, Isaksson M, Lindvall C et al. Genomic characterization of MOZ/CBP and CBP/MOZ chimeras in acute myeloid leukemia suggests the involvement of a damage-repair mechanism in the origin of the t(8;16)(p11;p13). Genes Chromosomes Cancer. 2003; 36: 90-8.

157. Masutani C., Sugasawa K., Yanagisawa J. et al. Purification and cloning of a nucleotide excision repair complex involving the Xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. EMBO J. 1994; 13:1831-1843.

158. Kaur M1, Pop M, Shi D et al. hHR23B is required for genotoxic-specific activation of p53 and apoptosis. Oncogene. 2007; 26:1231-7.

159. Miles M.C., Janket M.L., Wheeler E.D. et al. Molecular and functional characterization of a novel splice variant of ANKHD1 that lacks the KH domain and its role in cell survival and apoptosis. FEBS J. 2005; 272: 4091-4102.

160. Traina F., Favaro P.M.B., Medina Sde S. et al. ANKHD1, ankyrin repeat and KH domain containing 1, is overexpressed in acute leukemias and is associated with SHP2 in K562 cells. Biochim. Biophys. 2006; 1762: 828-834.

161. Machado-Neto JA, Lazarini M, Favaro P et al. ANKHD1 silencing inhibits Stathmin 1 activity, cell proliferation and migration of leukemia cells. Biochim Biophys Acta. 2014; 1853: 583-593.

162. Chen K, Tu Y, Zhang Y et al. PINCH-1 regulates the ERK-Bim pathway and contributes to apoptosis resistance in cancer cells. J Biol Chem. 2008; 283:2508-17.

163. Fukuda T, Chen K, Shi X et al. PINCH-1 is an obligate partner of integrin-linked kinase (ILK) functioning in cell shape modulation, motility, and survival. J Biol Chem. 2003; 278: 51324-33.

164. Zeng D, Hao L, Xu W et al. Pinch-1 was up-regulated in leukemia BMSC and its possible effect. Clin Exp Med. 2013; 13:21-7.

165. Tummala H, Kirwan M, Walne AJ et al. ERCC6L2 mutations link a distinct bone-marrow-failure syndrome to DNA repair and mitochondrial function. Am J Hum Genet. 2014; 94: 246-56.

166. Bendall SC, Stewart MH, Menendez P et al. IGF and FGF cooperatively establish the regulatory stem cell niche of pluripotent human cells in vitro. Nature. 2007; 448:1015-21.

167. Baserga R. The IGF-I receptor in cancer research. Exp Cell Res. 1999; 253:1-6.

168. Zhu WZ1, Wang SQ, Chakir K et al. Linkage of beta1-adrenergic stimulation to apoptotic heart cell death through protein kinase A-independent activation of Ca2+/calmodulin kinase II. J Clin Invest. 2003; 111: 617-25.

169. Shipman K.L., Robinson P.J., King B.R. et al. Identification of a family of DNA-binding proteins with homology to RNA splicing factors. Cell Biol. 2006; 84: 9-19.

170. Nishii, Y., Morishima, M., Kakehi, Y. et al. CROP/Luc7A, a novel serine/arginine-rich nuclear protein, isolated from cisplatin-resistant cell line. FEBS Lett. 2000; 465: 153-156.

171. Umehara, H., Nishii, Y., Morishima, M. et al. Effect of cisplatin treatment on speckled distribution of a serine/arginine-rich nuclear protein CROP/Luc7A. Biochem. Biophys. Res. Commun. 2003; 301: 324-329.

172. Mao S., Frank R.C., Zhang J. et al. Functional and physical interactions between AML1 proteins and an ETS protein, MEF: implications for the pathogenesis of t(8;21)-positive leukemias. Mol. Cell. Biol. 1999; 19: 3635-3644.

173. Fujimoto T., Anderson K., Jacobsen S.E. et al. Cdk6 blocks myeloid differentiation by interfering with Runx1 DNA binding and Runx1-C/EBPalpha interaction. EMBO J. 2007; 26: 2361-2370.

174. Taniuchi I, Osato M, Egawa T et al. Differential requirements for Runx proteins in CD4 repression and epigenetic silencing during T lymphocyte development. Cell. 2002; 111: 621-33.

175. Cleary, M. L. A new angle on a pervasive oncogene. Nature Genet. 1999; 23: 134-135.

176. Chen MJ, Yokomizo T, Zeigler BM et al. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. Nature. 2009; 457: 887-91.

177. Adamo L, Naveiras O, Wenzel PL et al. Biomechanical forces promote embryonic haematopoiesis. Nature. 2009; 459: 1131-5.

178. Kwiatkowski, N., Zhang, T., Rahl, P. B. et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. Nature 2014; 511: 616-620.

179. Jauliac, S., Lopez-Rodriguez, C., Shaw et al. The role of NFAT transcription factors in integrin-mediated carcinoma invasion. Nature Cell Biol. 2002; 4: 540-544.

180. Marino, G., Uria, J. A., Puente, X. S. et al. Human autophagins, a family of cysteine proteinases potentially implicated in cell degradation by autophagy. J. Biol. Chem. 2003; 278: 3671-3678.

181. Naetar, N., Korbei, B., Kozlov, S. et al. Loss of nucleoplasmic LAP2-alpha-lamin A complexes causes erythroid and epidermal progenitor hyperproliferation. Nature Cell Biol. 2008; 10: 1341-1348.

182. Tommerup, N., Vissing, H. Isolation and fine mapping of 16 novel human zinc finger-encoding cDNAs identify putative candidate genes for developmental and malignant disorders. Genomics 1995; 27: 259-264.

183. Rosinski, K. V., Fujii, N., Mito, J. K. et al. DDX3Y encodes a class I MHC-restricted H-Y antigen that is expressed in leukemic stem cells. Blood 2008; 111: 4817-4826.

184. van der Burg, M., Ijspeert, H., Verkaik, N. S. et al. A DNA-PKcs mutation in a radiosensitive T-B-SCID patient inhibits Artemis activation and nonhomologous end-joining. J. Clin. Invest. 2009; 119: 91-98.

185. Woodbine, L., Neal, J. A., Sasi, N.-K. et al. PRKDC mutations in a SCID patient with profound neurological abnormalities. J. Clin. Invest. 2013; 123: 2969-2980.

186. Wang, Y., Cortez, D., Yazdi, P. et al. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev. 2000; 14: 927-939.

187. Yamashita, A., Izumi, N., Kashima, I. et al. SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. Genes Dev. 2009; 23: 1091-1105.

188. Alahari, S. K., Reddig, P. J., Juliano, R. L. The integrin-binding protein nischarin regulates cell migration by inhibiting PAK. EMBO J. 2004; 23: 2777-2788.

189. Dontenwill M, Pascal G, Piletz JE et al. IRAS, the human homologue of Nischarin, prolongs survival of transfected PC12 cells. Cell Death Differ. 2003; 10: 933-5.

190. Suriano AR1, Sanford AN, Kim N et al. GCF2/LRRFIP1 represses tumor necrosis factor alpha expression. Mol Cell Biol. 2005; 25: 9073-81.

191. Johnson, A. C., Kageyama, R., Popescu, N. C. et al. Expression and chromosomal localization of the gene for the human transcriptional repressor GCF. J. Biol. Chem. 1992; 267: 1689-1694.

192. Moore, K. A., Sethi, R., Doanes, A. M. et al. Rac1 is required for cell proliferation and G2/M progression. Biochem. J. 1997; 326: 17-20.

193. Kheradmand, F., Werner, E., Tremble, P. et al. Role of Rac1 and oxygen radicals in collagenase-1 expression induced by cell shape change. Science 1998; 280: 898-902.

194. Toledo, F., Wahl, G. M. Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. Nat. Rev. Cancer 2006; 6: 909-923.

195. Bourdon, J.-C. p53 and its isoforms in cancer. Brit. J. Cancer 2007; 97: 277-282.

196. Vousden, K. H., Lane, D. P. p53 in health and disease. Nat. Rev. Mol. Cell Biol. 2007; 8: 275-283.

197. Zha, H., C. Aime-Sempe, T. Sato et al. Proapoptotic protein Bax heterodimerizes with Bcl-2 and homodimerizes with Bax via a novel domain (BH3) distinct from BH1 and BH2. J. Biol. Chem. 1996; 271: 7440-7444.

198. Oltvai, Z. N., Milliman, C. L., Korsmeyer, S. J. Bcl-2 heterodimers in vivo with a conserved homolog, Bax, that accelerates programmed cell death. Cell 1993; 74: 609-619.

199. Flaggs, G., Plug, A. W., Dunks, K. M. et al. Atm-dependent interactions of a mammalian Chk1 homolog with meiotic chromosomes. Curr. Biol. 1997; 7: 977-986.

200. Lopes, M., Cotta-Ramusino, C., Pellicioli, A. et al. The DNA replication checkpoint response stabilizes stalled replication forks. Nature 2001; 412: 557-561.

201. Sogo, J. M., Lopes, M., Foiani, M. Fork reversal and ssDNA accumulation at stalled replication forks owing to checkpoint defects. Science 2002; 297: 599-602.

202. Matsuoka, S., Huang, M., Elledge, S. J. Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. Science 1998; 282: 1893-1897.

203. Chehab, N. H., Malikzay, A., Appel, M. et al. Chk2/hCds1 functions as a DNA damage checkpoint in G-1 by stabilizing p53. Genes Dev. 2000; 14: 278-288.

204. Weinberg, R. A. The retinoblastoma protein and cell cycle control. Cell 1995; 81: 323-330.

205. Luo, R. X., Postigo, A. A., Dean, D. C. Rb interacts with histone deacetylase to repress transcription. Cell 1998; 92: 463-473.

206. Zhang, H. S., Postigo, A. A., Dean, D. C. Active transcriptional repression by the Rb-E2F complex mediates G1 arrest triggered by p16(INK4a), TGF-beta, and contact inhibition. Cell 1999; 97: 53-61.

207. Ginsberg D, Vairo G, Chittenden T et al. E2F-4, a new member of the E2F transcription factor family, interacts with p107. Genes Dev. 1994; 8: 2665-79.

208. Schwemmle S, Pfeifer GP.Genomic structure and mutation screening of the E2F4 gene in human tumors. Int J Cancer. 2000; 86: 672-7.

209. Lees JA1, Saito M, Vidal M et al. The retinoblastoma protein binds to a family of E2F transcription factors. Mol Cell Biol. 1993; 13: 7813-25.

210. Eiring AM1, Neviani P, Santhanam R et al. Identification of novel posttranscriptional targets of the BCR/ABL oncoprotein by ribonomics: requirement of E2F3 for BCR/ABL leukemogenesis. Blood. 2008; 111: 816-28.

211. Chong, J.-L., Wenzel, P. L., Saenz-Robles, M. T., et al. E2f1-3 switch from activators in progenitor cells to repressors in differentiating cells. Nature. 2009; 462: 930-4.

212. Sahara, S., Aoto, M., Eguchi, Y. et al. Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation. Nature 1999; 401: 168-173.

213. Chan CB, Liu X, Jang SW et al. NGF inhibits human leukemia proliferation by downregulating cyclin A1 expression through promoting acinus/CtBP2 association. Oncogene 2009; 28: 3825-36.

214. Cuesta, R., Martinez-Sanchez, A., Gebauer, F. miR-181a regulates cap-dependent translation of p27(kip1) mRNA in myeloid cells. Molec. Cell. Biol. 2009; 29: 2841-2851.

215. Muzio M1, Chinnaiyan AM, Kischkel FC et al. FLICE, a novel FADD-homologous ICE/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death--inducing signaling complex. Cell. 1996; 85: 817-27.

216. Stupack, D. G., Teitz, T., Potter, M. D. et al. Potentiation of neuroblastoma metastasis by loss of caspase-8. Nature 2006; 439: 95-99.

217. Kanehisa M, Goto S, Sato Y et al. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014; 42: D199-205.

218. Xu D, Hopf C, Reddy R et al. Narp and NP1 form heterocomplexes that function in developmental and activity-dependent synaptic plasticity. Neuron. 2003; 39: 513-28.

219. Hossain MA1, Russell JC, O'Brien R et al. Neuronal pentraxin 1: a novel mediator of hypoxic-ischemic injury in neonatal brain. J Neurosci. 2004; 24: 4187-96.

220. Boles NC, Hirsch SE, Le S et al. NPTX1 regulates neural lineage specification from human pluripotent stem cells. Cell Rep. 2014; 6: 724-36.

221. Rahpeymai Y1, Hietala MA, Wilhelmsson U et al. Complement: a novel factor in basal and ischemia-induced neurogenesis. EMBO J. 2006; 25: 1364-74.

222. S Reis E1, Falcão DA, Isaac L. Clinical aspects and molecular basis of primary deficiencies of complement component C3 and its regulatory proteins factor I and factor H. Scand J Immunol. 2006; 63: 155-68.

223. Chiu KH1, Chang YH, Wu YS et al. Quantitative secretome analysis reveals that COL6A1 is a metastasis-associated protein using stacking gel-aided purification combined with iTRAQ labeling. J Proteome Res. 2011; 10: 1110-25.

224. Wilson GL1, Fox CH, Fauci AS et al. cDNA cloning of the B cell membrane protein CD22: a mediator of B-B cell interactions. J Exp Med. 1991; 173: 137-46.

225. Rillahan CD, Macauley MS, Schwartz E et al. Disubstituted Sialic Acid Ligands Targeting Siglecs CD33 and CD22 Associated with Myeloid Leukaemias and B Cell Lymphomas. Chem Sci. 2014; 5: 2398-2406.

226. Boles KS, Mathew PA. Molecular cloning of CS1, a novel human natural killer cell receptor belonging to the CD2 subset of the immunoglobulin superfamily. Immunogenetics. 2001; 52: 302-7.

227. Murphy J.J., Hobby P., Vilarino-Varela J. et al. A novel immunoglobulin superfamily receptor (19A) related to CD2 is expressed on activated lymphocytes and promotes homotypic B-cell adhesion. Biochem. J. 2002; 361: 431-436.

228. Kim JR, Horton NC, Mathew SO et al. CS1 (SLAMF7) inhibits production of proinflammatory cytokines by activated monocytes. Inflamm Res. 2013; 62: 765-72.

229. Guo H, Cruz-Munoz ME, Wu N et al. Immune cell inhibition by SLAMF7 is mediated by a mechanism requiring src kinases, CD45, and SHIP-1 that is defective in multiple myeloma cells. Mol Cell Biol. 2015; 35: 41-51.

230. Herlofsen SR, Høiby T, Cacchiarelli D et al. Brief report: importance of SOX8 for in vitro chondrogenic differentiation of human mesenchymal stromal cells. Stem Cells. 2014; 32: 1629-35.

231. Xie J, Black DL. A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. Nature. 2001; 410: 936-9.

232. Reeves EP, Lu H, Jacobs HL et al. Killing activity of neutrophils is mediated through activation of proteases by K+ flux. Nature. 2002; 416: 291-7.

233. Kaulfuss S., Grzmil M., Hemmerlein B. et al. Leupaxin, a novel coactivator of the androgen receptor, is expressed in prostate cancer and plays a role in adhesion and invasion of prostate carcinoma cells. Mol. Endocrinol. 2008; 22: 1606-1621.

234. Chen P.W., Kroog G.S. Leupaxin is similar to paxillin in focal adhesion targeting and tyrosine phosphorylation but has distinct roles in cell adhesion and spreading. Cell Adh. Migr. 2010; 4: 527-540.

235. Chew V., Lam K.P. Leupaxin negatively regulates B cell receptor signaling. J. Biol. Chem. 2007; 282: 27181-27191.

236. Matsumoto T, Funk CD, Rådmark O et al. Molecular cloning and amino acid sequence of human 5-lipoxygenase. Proc Natl Acad Sci U S A. 1988; 85: 26-30.

237. Gilbert N.C., Bartlett S.G., Waight M.T. et al. The structure of human 5-lipoxygenase. Science 2011; 331: 217-219.

238. Roos J, Oancea C, Heinssmann M et al. 5-Lipoxygenase is a candidate target for therapeutic management of stem cell-like cells in acute myeloid leukemia. Cancer Res. 2014; 74: 5244-55.

239. Sarveswaran S, Chakraborty D, Chitale D et al. Inhibition of 5-lipoxygenase selectively triggers disruption of c-Myc signaling in prostate cancer cells. J Biol Chem. 2014. pii: jbc.M114.599035.

240. Zhao, L., Moos, M. P. W., Grabner, R. et al. The 5-lipoxygenase pathway promotes pathogenesis of hyperlipidemia-dependent aortic aneurysm. Nature Med. 2004; 10: 966-973.

241. Zhang JF, Liu XL, Lin YD et al. Bioinformatic analysis of chronic myeloid leukemia progression and preliminary experimental verification. Zhongguo Shi Yan Xue Ye Xue Za Zhi. 2014; 22: 909-13.

242. Manara E, Baron E, Tregnago C et al. MLL-AF6 fusion oncogene sequesters AF6 into the nucleus to trigger RAS activation in myeloid leukemia. Blood. 2014; 124: 263-72.

243. Fournier G, Cabaud O, Josselin E et al. Loss of AF6/afadin, a marker of poor outcome in breast cancer, induces cell migration, invasiveness and tumor growth.Oncogene. 2011; 30: 3862-74.

244. Taya, S., Yamamoto, T., Kano, K. et al. The Ras target AF-6 is a substrate of the Fam deubiquitinating enzyme. J. Cell Biol. 1998; 142: 1053-1062.

245. Zhang, Z., Rehmann, H., Price, L. et al. AF6 negatively regulates Rap1-induced cell adhesion. J. Biol. Chem. 2005; 280: 33200-33205.

246. Zdebik A.A., Zifarelli G., Bergsdorf E.-Y. et al. Determinants of anion-proton coupling in mammalian endosomal CLC proteins. J. Biol. Chem. 2008; 283: 4219-4227.

247. Ishiguro T, Avila H, Lin SY et al. Gene trapping identifies chloride channel 4 as a novel inducer of colon cancer cell migration, invasion and metastases. Br J Cancer. 2010; 102: 774-82.

248. Wang T, Weinman SA. Involvement of chloride channels in hepatic copper metabolism: ClC-4 promotes copper incorporation into ceruloplasmin. Gastroenterology 2004; 126: 1157-66.

249. Huang X., Summers M.K., Pham V. et al. Deubiquitinase USP37 is activated by CDK2 to antagonize APC(CDH1) and promote S phase entry. Mol. Cell 2011; 42: 511-523.

250. Yang WC, Shih HM. The deubiquitinating enzyme USP37 regulates the oncogenic fusion protein PLZF/RARA stability. Oncogene. 2013; 32: 5167-75.

251. Pan J, Deng Q, Jiang C et al. USP37 directly deubiquitinates and stabilizes c-Myc in lung cancer. Oncogene 2014; doi: 10.1038/onc.2014.327.

252. Uyama, T., Jin, X.-H., Tsuboi, K. et al. Characterization of the human tumor suppressors TIG3 and HRASLS2 as phospholipid-metabolizing enzymes. Biochim. Biophys. 2009; 1791: 1114-1124.

253. DiSepio, D., Ghosn, C., Eckert, R. L. et al. Identification and characterization of a retinoid-induced class II tumor suppressor/growth regulatory gene. Proc. Nat. Acad. Sci. 1998; 95: 14811-14815.

254. Miosge N, Götz W, Sasaki T et al. The extracellular matrix proteins fibulin-1 and fibulin-2 in the early human embryo. Histochem J. 1996; 28: 109-16.

255. Zhang HY, Timpl R, Sasaki T et al. Fibulin-1 and fibulin-2 expression during organogenesis in the developing mouse embryo. Dev Dyn. 1996; 205: 348-64.

256. Qing J., Maher V.M., Tran H. et al. Suppression of anchorage-independent growth and matrigel invasion and delayed tumor formation by elevated expression of fibulin-1D in human fibrosarcoma-derived cell lines. Oncogene 1997; 15: 2159-2168.

257. Hayashido Y., Lucas A., Rougeot C. et al. Estradiol and fibulin-1 inhibit motility of human ovarian- and breast-cancer cells induced by fibronectin. Int. J. Cancer 1998; 75: 654-658.

258. Twal W.O., Czirok A., Hegedus B. et al. Fibulin-1 suppression of fibronectin-regulated cell adhesion and motility. J. Cell Sci. 2001; 114: 4587-4598.

259. Xiao W, Wang J, Li H et al. Fibulin-1 is epigenetically down-regulated and related with bladder cancer recurrence. BMC Cancer 2014; 18: 14:677.

260. Crambert G1, Geering K. FXYD proteins: new tissue-specific regulators of the ubiquitous Na, K-ATPase. Sci STKE. 2003;2003(166):RE1.

261. Choudhury, K., McQuillin, A., Puri, V. et al. A genetic association study of chromosome 11q22-24 in two different samples implicates the FXYD6 gene, encoding phosphohippolin, in susceptibility to schizophrenia. Am. J. Hum. Genet. 2007; 80: 664-672.

262. Liu J1, Zhou N, Zhang X. A monoclonal antibody against human FXYD6. Hybridoma (Larchmt). 2011; 30: 487-90.

263. Kowarz E1, Burmeister T, Lo Nigro L et al. Complex MLL rearrangements in t(4;11) leukemia patients with absent AF4.MLL fusion allele. Leukemia. 2007; 21: 1232-8.

264. Rahpeymai, Y., Hietala, M. A., Wilhelmsson, U. et al. Complement: a novel factor in basal and ischemia-induced neurogenesis. EMBO J. 2006; 25: 1364-1374.

265. Whitman SP, Ruppert AS, Radmacher MD et al. FLT3 D835/I836 mutations are associated with poor disease-free survival and a distinct gene-expression signature among younger adults with de novo cytogenetically normal acute myeloid leukemia lacking FLT3 internal tandem duplications. Blood. 2008; 111: 1552-9.

266. Chapoval A.I., Ni J., Lau J.S. et al. B7-H3: a costimulatory molecule for T cell activation and IFN-gamma production. Nat. Immunol. 2001; 2: 269-274.

267. Castriconi R., Dondero A., Augugliaro R. Identification of 4Ig-B7-H3 as a neuroblastoma-associated molecule that exerts a protective role from an NK cell-mediated lysis. Proc. Natl. Acad. Sci. U.S.A. 2004; 101:12640-12645.

268. Kraan J, van den Broek P, Verhoef C et al.Endothelial CD276 (B7-H3) expression is increased in human malignancies and distinguishes between normal and tumour-derived circulating endothelial cells. Br J Cancer. 2014; 111: 149-56.

269. Dietz ML, Bernaciak TM, Vendetti F et al. Differential actin-dependent localization modulates the evolutionarily conserved activity of Shroom family proteins. J Biol Chem. 2006 Jul 21;281(29):20542-54.

270. Farber MJ, Rizaldy R, Hildebrand JD. Shroom2 regulates contractility to control endothelial morphogenesis. Mol Biol Cell. 2011; 22: 795-805.

271. Baumann CA, Ribon V, Kanzaki M et al. CAP defines a second signalling pathway required for insulin-stimulated glucose transport. Nature. 2000; 407: 202-7.

272. Mastick CC, Brady MJ, Saltiel AR. Insulin stimulates the tyrosine phosphorylation of caveolin. J Cell Biol. 1995; 129: 1523-31.

273. Ribon, V., Printen, J. A., Hoffman, N. G. et al. A novel, multifunctional c-Cbl binding protein in insulin receptor signaling in 3T3-L1 adipocytes. Molec. Cell. Biol. 1998; 18: 872-879.

274. O'Connor, L., Strasser, A., O'Reilly, L. A. et al. Bim: a novel member of the Bcl-2 family that promotes apoptosis. EMBO J. 1998; 17: 384-395.

275. Willis, S. N., Fletcher, J. I., Kaufmann, T. et al. Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. Science 2007; 315: 856-859.

276. Desbien, A. L., Kappler, J. W., Marrack, P. The Epstein-Barr virus Bcl-2 homolog, BHRF1, blocks apoptosis by binding to a limited amount of Bim. Proc. Nat. Acad. Sci. 2009; 106: 5663-5668.

277. Bakker, A. B. H., Baker, E., Sutherland, G. R. et al. Myeloid DAP12-associating lectin (MDL)-1 is a cell surface receptor involved in the activation of myeloid cells. Proc. Nat. Acad. Sci. 1999; 96: 9792-9796.

278. Batliner J1, Mancarelli MM, Jenal M et al. CLEC5A (MDL-1) is a novel PU.1 transcriptional target during myeloid differentiation. Mol Immunol. 2011; 48: 714-9.

279. Aoki N1, Kimura Y, Kimura S et al. Expression and functional role of MDL-1 (CLEC5A) in mouse myeloid lineage cells. J Leukoc Biol. 2009; 85: 508-17.

280. Audas T.E., Li Y., Liang G. et al. A novel protein, Luman/CREB3 recruitment factor, inhibits Luman activation of the unfolded protein response. Mol. Cell. Biol. 2008; 28: 3952-3966.

281. Yang Y, Jin Y, Martyn AC et al. Expression pattern implicates a potential role for luman recruitment factor in the process of implantation in uteri and development of preimplantation embryos in mice. J Reprod Dev. 2013; 59: 245-51.

282. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. J Hum Genet. 2014; 59: 5-15.

283. Gamerdinger U, Teigler-Schlegel A, Pils S et al. Cryptic chromosomal aberrations leading to an AML1/ETO rearrangement are frequently caused by small insertions. Genes Chromosomes Cancer. 2003; 36: 261-72.

284. Schoch C, Schnittger S, Kern W et al. Rapid diagnostic approach to PML-RARalpha-positive acute promyelocytic leukemia. Hematol J. 2002; 3: 259-63.

285. Ballerini P, Blaise A, Mercher T et al. A novel real-time RTePCR assay for quantification of OTT-MAL fusion transcript reliable for diagnosis of t(1;22) and minimal residual disease (MRD) detection. Leukemia 2003; 17: 1193e1196.

286. Golay J, Capucci A, Arsura M et al. Expression of c-myb and B-myb, but not A-myb, correlates with proliferation in human hematopoietic cells. Blood 1991; 77: 149-158.

287. Reiss K, Travali S, Calabretta B et al. Growth regulated expression of B-myb in fibroblasts and hematopoietic cells. J Cell Physiol 1991; 148 : 338 - 343.

288. Martinez I, Cazalla D, Almstead LL et al. miR-29 and miR-30 regulate B-Myb expression during cellular senescence. Proc Natl Acad Sci USA 2011; 108: 522 - 527.

289. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 2013; 368: 2059 - 2074.

290. Ren F, Wang L, Shen X et al. MYBL2 is an independent prognostic marker that has tumor-promoting functions in colorectal cancer. Am J Cancer Res. 2015; 5: 1542-52.

291. Zhan M, Riordon DR, Yan B et al. The B-MYB transcriptional network guides cell cycle progression and fate decisions to sustain self-renewal and the identity of pluripotent stem cells. PLoS One. 2012; 7: e42350.

292. Baserga, R. Cancer Res. 1995; 55: 249-252.

293. Werner, H., LeRoith, D. Adv. Cancer Res. 1996; 68: 183-223.

294. Sell, C., Rubini, M., Rubin, R. et al. Proc. Natl. Acad. Sci. USA, 1993; 90: 11217-11221.

295. Sell, C., Dumenil, G., Deveaud, C. et al. Mol. CeU Biol. 1994; 14: 3604-3612.

296. Bartkova, J., Horejsi, Z., Koed, K. et al. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. Nature 2005; 434: 864-870.

297. Janssen, A., van der Burg, M., Szuhai, K. et al. Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. Science 2011; 333: 1895-1898.

298. Chittenden, T., C. Flemington, A.B. Houghton et al. A conserved domain in Bak, distinct from BH1 and BH2, mediates cell death and protein binding functions. EMBO J. 1995; 14: 5589-5596.

299. Bouillet, P., Metcalf, D., Huang, D. C. S. et al. Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. Science 1999; 286: 1735-1738.

300. Saitoh, T., Katoh, M. Molecular cloning and characterization of human WNT8A. Int. J. Oncol. 2001; 19: 123-127.

301. Lee, J.-S., Collins, K. M., Brown, A. L. et al. hCds1-mediated phosphorylation of BRCA1 regulates the DNA damage response. Nature 2000; 404: 201-204.

302. Wang, B., Hurov, K., Hofmann, K. et al. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. Genes Dev. 2009; 23: 729-739.

303. Yarden, R. I., Brody, L. C. BRCA1 interacts with components of the histone deacetylase complex. Proc. Nat. Acad. Sci. 1999; 96: 4983-4988.

ANNEXES

Table 1. List of genes with differential expression in CD34$^+$ cells transfected with specific *MYBL2* siRNAs compared to CD34$^+$ cells transfected with control siRNAs. The table shows a list of 270 genes obtained by establishing ±2.5 FC and p<0.05 criteria to the ANOVA test.

Gene Symbol	Gene Name	p	FD	Description
MAP2K7	mitogen-activated protein kinase kinase 7	0,0364693	4,37	MYBL2 siRNA up vs siRNA control
UTY	ubiquitously transcribed tetratricopeptide repeat gene, Y-linked	0,0262523	4,17	MYBL2 siRNA up vs siRNA control
CREBBP	CREB binding protein	0,00352764	3,75	MYBL2 siRNA up vs siRNA control
RAD23A	RAD23 homolog A (S, cerevisiae)	0,00183783	3,72	MYBL2 siRNA up vs siRNA control
ANKHD1 /// ANKHD1-EIF4EBP3	ankyrin repeat and KH domain containing 1 /// ANKHD1-EIF4EBP3 readthrough	0,0193292	3,70	MYBL2 siRNA up vs siRNA control
LIMS1	LIM and senescent cell antigen-like domains 1	0,0407817	3,68	MYBL2 siRNA up vs siRNA control
ERCC6L2	excision repair cross-complementing rodent repair deficiency, complementation group 6-l	0,0361319	3,48	MYBL2 siRNA up vs siRNA control
TGOLN2	trans-golgi network protein 2	0,00998736	3,42	MYBL2 siRNA up vs siRNA control
IGF1R	insulin-like growth factor 1 receptor	0,00679523	3,41	MYBL2 siRNA up vs siRNA control
ERCC6L2	excision repair cross-complementing rodent repair deficiency, complementation group 6-l	0,0433322	3,40	MYBL2 siRNA up vs siRNA control
CAMK2G	calcium/calmodulin-dependent protein kinase II gamma	0,0254706	3,39	MYBL2 siRNA up vs siRNA control
LUC7L3	LUC7-like 3 (S, cerevisiae)	0,0489028	3,39	MYBL2 siRNA up vs siRNA control
PGK1	phosphoglycerate kina	0,0254692	3,38	MYBL2 siRNA up vs siRNA control
RUNX1	runt-related transcription factor 1	0,0204564	3,34	MYBL2 siRNA up vs siRNA control
CMIP	c-Maf inducing protein	0,0013814	3,30	MYBL2 siRNA up vs siRNA control
TRPT1	tRNA phosphotransferase 1	0,00223967	3,23	MYBL2 siRNA up vs siRNA control
RPL10 /// SNORA70	ribosomal protein L10 /// small nucleolar RNA, H/ACA box 70	0,0194609	3,22	MYBL2 siRNA up vs siRNA control
OGT	O-linked N-acetylglucosamine (GlcNAc) transferase	0,0326613	3,21	MYBL2 siRNA up vs siRNA control
NFAT5	nuclear factor of activated T-cells 5, tonicity-responsive	0,00943376	3,16	MYBL2 siRNA up vs siRNA control
ATG4B	autophagy related 4B, cysteine peptidase	0,00316185	3,16	MYBL2 siRNA up vs siRNA control
TMPO	thymopoietin	0,00655938	3,16	MYBL2 siRNA up vs siRNA control
ZKSCAN1	zinc finger with KRAB and SCAN domains 1	0,0213024	3,15	MYBL2 siRNA up vs siRNA control
DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	0,0333049	3,15	MYBL2 siRNA up vs siRNA control
PAN3	PAN3 poly(A) specific ribonuclease subunit homolog (S, cerevisiae)	0,0357118	3,14	MYBL2 siRNA up vs siRNA control
PRKDC	protein kinase, DNA-activated, catalytic polypeptide	0,00136466	3,14	MYBL2 siRNA up vs siRNA control
MRE11A	MRE11 meiotic recombination 11 homolog A (S, cerevisiae)	0,0402543	3,14	MYBL2 siRNA up vs siRNA control
CREBBP	CREB binding protein	0,00771751	3,11	MYBL2 siRNA up vs siRNA control
SMG1	SMG1 homolog, phosphatidylinositol 3-kinase-related kinase (C, elegans) pseudogene ///	0,0267555	3,11	MYBL2 siRNA up vs siRNA control
HNRNPH2 /// RPL36A-HNRNPH2	heterogeneous nuclear ribonucleoprotein H2 (H') /// RPL36A-HNRNPH2 readthrough	0,0175336	3,10	MYBL2 siRNA up vs siRNA control
NISCH	nischarin	0,02205	3,08	MYBL2 siRNA up vs siRNA control
LRRFIP1	leucine rich repeat (in FLII) interacting protein 1	0,00470694	3,08	MYBL2 siRNA up vs siRNA control
RAC1	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	0,0127767	3,07	MYBL2 siRNA up vs siRNA control
KIAA0430	KIAA0430	0,0347073	3,07	MYBL2 siRNA up vs siRNA control
KIAA0146	KIAA0146	0,0438398	3,06	MYBL2 siRNA up vs siRNA control
ZNRF1	zinc and ring finger 1, E3 ubiquitin protein ligase	0,0119063	3,06	MYBL2 siRNA up vs siRNA control
LARP1	La ribonucleoprotein domain family, member 1	0,0194377	3,06	MYBL2 siRNA up vs siRNA control
PM20D2	peptidase M20 domain containing 2	0,0384932	3,05	MYBL2 siRNA up vs siRNA control
EXOC5	exocyst complex component 5	0,0343603	3,05	MYBL2 siRNA up vs siRNA control

WAPAL	wings apart-like homolog (Drosophila)	0,0221148	3,04	MYBL2 siRNA up vs siRNA control
JAK3	Janus kinase 3	0,0355164	3,03	MYBL2 siRNA up vs siRNA control
BTAF1	BTAF1 RNA polymerase II, B-TFIID transcription factor-associated, 170kDa (Mot1 homolog)	0,0406452	3,02	MYBL2 siRNA up vs siRNA control
TARBP1	TAR (HIV-1) RNA binding protein 1	0,0350831	3,01	MYBL2 siRNA up vs siRNA control
ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	0,00533278	2,99	MYBL2 siRNA up vs siRNA control
QKI	QKI, KH domain containing, RNA binding	0,00786243	2,99	MYBL2 siRNA up vs siRNA control
SATB1	SATB homeobox 1	0,00112199	2,98	MYBL2 siRNA up vs siRNA control
STX16	syntaxin 16	0,00653573	2,97	MYBL2 siRNA up vs siRNA control
SLTM	SAFB-like, transcription modulator	0,0237506	2,96	MYBL2 siRNA up vs siRNA control
CD164	CD164 molecule, sialomucin	0,0390538	2,95	MYBL2 siRNA up vs siRNA control
ABR	active BCR-related	0,00688371	2,95	MYBL2 siRNA up vs siRNA control
PPP1R18	protein phosphatase 1, regulatory subunit 18	0,0221718	2,95	MYBL2 siRNA up vs siRNA control
SKAP2	src kinase associated phosphoprotein 2	0,0194532	2,95	MYBL2 siRNA up vs siRNA control
ELF1	E74-like factor 1 (ets domain transcription factor)	0,00467143	2,94	MYBL2 siRNA up vs siRNA control
AMPD3	adenosine monophosphate deaminase 3	0,0178156	2,93	MYBL2 siRNA up vs siRNA control
RNF38	ring finger protein 38	0,0165732	2,92	MYBL2 siRNA up vs siRNA control
KLHL6	kelch-like 6 (Drosophila)	0,00898566	2,90	MYBL2 siRNA up vs siRNA control
ATP2C1	ATPase, Ca++ transporting, type 2C, member 1	0,0237327	2,90	MYBL2 siRNA up vs siRNA control
PPIL2	peptidylprolyl isomerase (cyclophilin)-like 2	0,0496351	2,89	MYBL2 siRNA up vs siRNA control
CLPTM1	cleft lip and palate associated transmembrane protein 1	0,0294829	2,87	MYBL2 siRNA up vs siRNA control
GALNT2	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 2 (Gal	0,00233951	2,86	MYBL2 siRNA up vs siRNA control
STX16	syntaxin 16	0,0155028	2,86	MYBL2 siRNA up vs siRNA control
KIAA0182	KIAA0182	0,00603493	2,86	MYBL2 siRNA up vs siRNA control
VPS13C	vacuolar protein sorting 13 homolog C (S, cerevisiae)	0,013204	2,86	MYBL2 siRNA up vs siRNA control
SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c	0,0123381	2,85	MYBL2 siRNA up vs siRNA control
IRF2BP2	interferon regulatory factor 2 binding protein 2	0,0180137	2,85	MYBL2 siRNA up vs siRNA control
FAM168B	family with sequence similarity 168, member B	0,00768864	2,85	MYBL2 siRNA up vs siRNA control
MPLKIP	M-phase specific PLK1 interacting protein	0,0345364	2,85	MYBL2 siRNA up vs siRNA control
XIST	X (inactive)-specific transcript (non-protein coding)	0,0384493	2,85	MYBL2 siRNA up vs siRNA control
G3BP2	GTPase activating protein (SH3 domain) binding protein 2	0,0394857	2,84	MYBL2 siRNA up vs siRNA control
ZNF711	zinc finger protein 711	0,0255003	2,84	MYBL2 siRNA up vs siRNA control
ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	0,00224193	2,84	MYBL2 siRNA up vs siRNA control
LENG8	leukocyte receptor cluster (LRC) member 8	0,0436304	2,83	MYBL2 siRNA up vs siRNA control
PTAR1	protein prenyltransferase alpha subunit repeat containing 1	0,0400291	2,83	MYBL2 siRNA up vs siRNA control
XPO4	exportin 4	0,00442977	2,83	MYBL2 siRNA up vs siRNA control
GSK3A	glycogen synthase kinase 3 alpha	0,0111038	2,82	MYBL2 siRNA up vs siRNA control
ILF3	interleukin enhancer binding factor 3, 90kDa	0,0244701	2,82	MYBL2 siRNA up vs siRNA control
ATRX	alpha thalassemia/mental retardation syndrome X-linked	0,0185584	2,81	MYBL2 siRNA up vs siRNA control
EWSR1	Ewing sarcoma breakpoint region 1	0,0120771	2,81	MYBL2 siRNA up vs siRNA control
ZRANB2	zinc finger, RAN-binding domain containing 2	0,0495438	2,81	MYBL2 siRNA up vs siRNA control
CELF2	CUGBP, Elav-like family member 2	0,0151616	2,81	MYBL2 siRNA up vs siRNA control
SETD2	SET domain containing 2	0,00431358	2,80	MYBL2 siRNA up vs siRNA control
TAF6	TAF6 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 80kDa	0,006201	2,80	MYBL2 siRNA up vs siRNA control
HSPA1A /// HSPA1B	heat shock 70kDa protein 1A /// heat shock 70kDa protein 1B	0,0430626	2,80	MYBL2 siRNA up vs siRNA control

CNN2	calponin 2	0,0169653	2,80	MYBL2 siRNA up vs siRNA control
DCTD	dCMP deaminase	0,0348844	2,79	MYBL2 siRNA up vs siRNA control
GOLPH3	golgi phosphoprotein 3 (coat-protein)	0,0239897	2,79	MYBL2 siRNA up vs siRNA control
SLC7A1	solute carrier family 7 (cationic amino acid transporter, y+ system), member 1	0,0107956	2,79	MYBL2 siRNA up vs siRNA control
MYCN	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	0,0234781	2,79	MYBL2 siRNA up vs siRNA control
DEDD	death effector domain containing	0,00181295	2,78	MYBL2 siRNA up vs siRNA control
ZMYM2	zinc finger, MYM-type 2	0,004353	2,77	MYBL2 siRNA up vs siRNA control
WDR11	WD repeat domain 11	0,03706	2,77	MYBL2 siRNA up vs siRNA control
CLPTM1	cleft lip and palate associated transmembrane protein 1	0,0156899	2,76	MYBL2 siRNA up vs siRNA control
KDM3B	lysine (K)-specific demethylase 3B	0,0116968	2,76	MYBL2 siRNA up vs siRNA control
HP1BP3	heterochromatin protein 1, binding protein 3	0,0145935	2,76	MYBL2 siRNA up vs siRNA control
MAU2	MAU2 chromatid cohesion factor homolog (C, elegans)	0,0373637	2,76	MYBL2 siRNA up vs siRNA control
TRIM14	tripartite motif containing 14	0,0183002	2,76	MYBL2 siRNA up vs siRNA control
ELOVL5	ELOVL fatty acid elongase 5	0,0302169	2,76	MYBL2 siRNA up vs siRNA control
USP24	ubiquitin specific peptidase 24	0,0246643	2,75	MYBL2 siRNA up vs siRNA control
FAM168B	family with sequence similarity 168, member B	0,00637913	2,75	MYBL2 siRNA up vs siRNA control
CD47	CD47 molecule	0,00185149	2,75	MYBL2 siRNA up vs siRNA control
MAN1B1	mannosidase, alpha, class 1B, member 1	0,0165616	2,75	MYBL2 siRNA up vs siRNA control
C6orf72	chromosome 6 open reading frame 72	0,00857829	2,75	MYBL2 siRNA up vs siRNA control
TMEM170A	transmembrane protein 170A	0,00175557	2,74	MYBL2 siRNA up vs siRNA control
GLS	glutaminase	0,0380461	2,74	MYBL2 siRNA up vs siRNA control
NR3C1	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)	0,00212351	2,73	MYBL2 siRNA up vs siRNA control
TIA1	TIA1 cytotoxic granule-associated RNA binding protein	0,0436153	2,73	MYBL2 siRNA up vs siRNA control
ILF3	interleukin enhancer binding factor 3, 90kDa	0,0327016	2,72	MYBL2 siRNA up vs siRNA control
NKTR	natural killer-tumor recognition sequence	0,0430077	2,72	MYBL2 siRNA up vs siRNA control
RAB11B	RAB11B, member RAS oncogene family	0,00186567	2,72	MYBL2 siRNA up vs siRNA control
SLC12A6	solute carrier family 12 (potassium/chloride transporters), member 6	0,00799115	2,72	MYBL2 siRNA up vs siRNA control
GALNT2	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 2 (Gal	0,0124477	2,71	MYBL2 siRNA up vs siRNA control
FAM117A	family with sequence similarity 117, member A	0,044605	2,71	MYBL2 siRNA up vs siRNA control
GNB4	guanine nucleotide binding protein (G protein), beta polypeptide 4	0,0148423	2,70	MYBL2 siRNA up vs siRNA control
SEPT6	septin 6	0,0218027	2,70	MYBL2 siRNA up vs siRNA control
HNRNPH2 /// RPL36A-HNRNPH2	heterogeneous nuclear ribonucleoprotein H2 (H') /// RPL36A-HNRNPH2 readthrough	0,0369648	2,70	MYBL2 siRNA up vs siRNA control
ADCY7	adenylate cyclase 7	0,0111121	2,70	MYBL2 siRNA up vs siRNA control
XIST	X (inactive)-specific transcript (non-protein coding)	0,0191983	2,70	MYBL2 siRNA up vs siRNA control
ELK3	ELK3, ETS-domain protein (SRF accessory protein 2)	0,0342638	2,70	MYBL2 siRNA up vs siRNA control
FBXO11	F-box protein 11	0,0160295	2,69	MYBL2 siRNA up vs siRNA control
TMEM33	transmembrane protein 33	0,0203085	2,69	MYBL2 siRNA up vs siRNA control
PDE7A	phosphodiesterase 7A	0,0296013	2,69	MYBL2 siRNA up vs siRNA control
TMPO	thymopoietin	0,0130705	2,69	MYBL2 siRNA up vs siRNA control
TSNAX	translin-associated factor X	0,0382213	2,68	MYBL2 siRNA up vs siRNA control
PPP1R9B	protein phosphatase 1, regulatory subunit 9B	0,00199668	2,68	MYBL2 siRNA up vs siRNA control
FAM134A	family with sequence similarity 134, member A	0,0136757	2,67	MYBL2 siRNA up vs siRNA control
CHST11	carbohydrate (chondroitin 4) sulfotransferase 11	0,00628476	2,67	MYBL2 siRNA up vs siRNA control
LRRC8A	leucine rich repeat containing 8 family, member A	0,0303486	2,67	MYBL2 siRNA up vs siRNA control

DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	0,0145229	2,67	MYBL2 siRNA up vs siRNA control
SASH3	SAM and SH3 domain containing 3	0,0133267	2,67	MYBL2 siRNA up vs siRNA control
MAML1	mastermind-like 1 (Drosophila)	0,0205825	2,67	MYBL2 siRNA up vs siRNA control
THOC2	THO complex 2	0,00456256	2,67	MYBL2 siRNA up vs siRNA control
ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	0,000694928	2,67	MYBL2 siRNA up vs siRNA control
IQGAP2	IQ motif containing GTPase activating protein 2	0,0302506	2,67	MYBL2 siRNA up vs siRNA control
TIE1	tyrosine kinase with immunoglobulin-like and EGF-like domains 1	0,000397017	2,66	MYBL2 siRNA up vs siRNA control
CD164	CD164 molecule, sialomucin	0,0390208	2,66	MYBL2 siRNA up vs siRNA control
ITGA6	integrin, alpha 6	0,0124422	2,66	MYBL2 siRNA up vs siRNA control
CLINT1	clathrin interactor 1	0,00849398	2,66	MYBL2 siRNA up vs siRNA control
SUN2	Sad1 and UNC84 domain containing 2	0,00914657	2,66	MYBL2 siRNA up vs siRNA control
FAM208B	family with sequence similarity 208, member B	0,00328933	2,65	MYBL2 siRNA up vs siRNA control
NSF	vesicle-fusing ATPase-like /// N-ethylmaleimide-sensitive factor	0,00409692	2,65	MYBL2 siRNA up vs siRNA control
TM9SF4	transmembrane 9 superfamily protein member 4	0,00448774	2,65	MYBL2 siRNA up vs siRNA control
CD44	CD44 molecule (Indian blood group)	0,00733934	2,65	MYBL2 siRNA up vs siRNA control
DAZAP2	DAZ associated protein 2	0,0162952	2,65	MYBL2 siRNA up vs siRNA control
TPCN1	two pore segment channel 1	0,0183602	2,65	MYBL2 siRNA up vs siRNA control
MLL	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	0,00785106	2,65	MYBL2 siRNA up vs siRNA control
SRSF11	serine/arginine-rich splicing factor 11	0,00371979	2,64	MYBL2 siRNA up vs siRNA control
ARID1B	AT rich interactive domain 1B (SWI1-like)	0,0302473	2,64	MYBL2 siRNA up vs siRNA control
DDX19A	DEAD (Asp-Glu-Ala-Asp) box polypeptide 19A	0,0124418	2,64	MYBL2 siRNA up vs siRNA control
MED25	mediator complex subunit 25	0,0326003	2,64	MYBL2 siRNA up vs siRNA control
CLTC	clathrin, heavy chain (Hc)	0,000997555	2,64	MYBL2 siRNA up vs siRNA control
DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	0,0367603	2,64	MYBL2 siRNA up vs siRNA control
PRNP	Prion protein	0,000573488	2,63	MYBL2 siRNA up vs siRNA control
CLINT1	clathrin interactor 1	0,0417081	2,63	MYBL2 siRNA up vs siRNA control
ZNF362	zinc finger protein 362	0,0107727	2,63	MYBL2 siRNA up vs siRNA control
GSK3A	glycogen synthase kinase 3 alpha	0,0313078	2,63	MYBL2 siRNA up vs siRNA control
UBLCP1	ubiquitin-like domain containing CTD phosphatase 1	0,0271626	2,63	MYBL2 siRNA up vs siRNA control
EGFL8 /// PPT2	EGF-like-domain, multiple 8 /// palmitoyl-protein thioesterase 2	0,00676749	2,63	MYBL2 siRNA up vs siRNA control
ITGA4	integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	0,0231573	2,63	MYBL2 siRNA up vs siRNA control
SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a	0,0458694	2,62	MYBL2 siRNA up vs siRNA control
CELF2	CUGBP, Elav-like family member 2	0,0284012	2,62	MYBL2 siRNA up vs siRNA control
BCOR	BCL6 corepressor	0,000933068	2,62	MYBL2 siRNA up vs siRNA control
BTBD3	BTB (POZ) domain containing 3	0,000566342	2,62	MYBL2 siRNA up vs siRNA control
PPP1R12A	protein phosphatase 1, regulatory subunit 12A	0,0171538	2,61	MYBL2 siRNA up vs siRNA control
FMNL1	Formin-like 1	0,0261958	2,61	MYBL2 siRNA up vs siRNA control
CPSF6	cleavage and polyadenylation specific factor 6, 68kDa	0,0280323	2,61	MYBL2 siRNA up vs siRNA control
TOP2B	topoisomerase (DNA) II beta 180kDa	0,0390023	2,61	MYBL2 siRNA up vs siRNA control
SH2B1	SH2B adaptor protein 1	0,00719138	2,61	MYBL2 siRNA up vs siRNA control
KDM5D	lysine (K)-specific demethylase 5D	0,0364506	2,60	MYBL2 siRNA up vs siRNA control
GCC2	GRIP and coiled-coil domain containing 2	0,0222705	2,60	MYBL2 siRNA up vs siRNA control
HPS4	Hermansky-Pudlak syndrome 4	0,0392686	2,60	MYBL2 siRNA up vs siRNA control
GTPBP3	GTP binding protein 3 (mitochondrial)	0,0368139	2,60	MYBL2 siRNA up vs siRNA control
C11orf54	chromosome 11 open reading frame 54	0,0402114	2,60	MYBL2 siRNA up vs siRNA control
TTC3 /// TTC3P1	tetratricopeptide repeat domain 3 /// tetratricopeptide repeat domain 3 pseudogene 1	0,0262762	2,60	MYBL2 siRNA up vs siRNA control

CERS2	ceramide synthase 2	0,0152193	2,60	MYBL2 siRNA up vs siRNA control
PUS7	pseudouridylate synthase 7 homolog (S, cerevisiae)	0,0286338	2,60	MYBL2 siRNA up vs siRNA control
HDAC6	histone deacetylase 6	0,0072634	2,60	MYBL2 siRNA up vs siRNA control
TNS3	tensin 3	0,0112567	2,60	MYBL2 siRNA up vs siRNA control
ZNF330	zinc finger protein 330	0,0160453	2,60	MYBL2 siRNA up vs siRNA control
DPYSL2	dihydropyrimidinase-like 2	0,00129477	2,59	MYBL2 siRNA up vs siRNA control
RBBP6	retinoblastoma binding protein 6	0,0165615	2,59	MYBL2 siRNA up vs siRNA control
PSMD12	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12	0,0250378	2,59	MYBL2 siRNA up vs siRNA control
H1F0	H1 histone family, member 0	0,0182896	2,59	MYBL2 siRNA up vs siRNA control
HNRNPH3	heterogeneous nuclear ribonucleoprotein H3 (2H9)	0,0485621	2,59	MYBL2 siRNA up vs siRNA control
PXN	paxillin	0,0168048	2,58	MYBL2 siRNA up vs siRNA control
ATP2B4	ATPase, Ca++ transporting, plasma membrane 4	0,0265238	2,58	MYBL2 siRNA up vs siRNA control
ODF2	outer dense fiber of sperm tails 2	0,00532644	2,58	MYBL2 siRNA up vs siRNA control
CNOT8	CCR4-NOT transcription complex, subunit 8	0,0118131	2,58	MYBL2 siRNA up vs siRNA control
TOPBP1	topoisomerase (DNA) II binding protein 1	0,0114997	2,58	MYBL2 siRNA up vs siRNA control
CSRP1	cysteine and glycine-rich protein 1	0,0234353	2,58	MYBL2 siRNA up vs siRNA control
GOPC	golgi-associated PDZ and coiled-coil motif containing	0,0158207	2,57	MYBL2 siRNA up vs siRNA control
SLFN11	schlafen family member 11	0,0292579	2,57	MYBL2 siRNA up vs siRNA control
MSRB3	methionine sulfoxide reductase B3	0,0180734	2,57	MYBL2 siRNA up vs siRNA control
OGT	O-linked N-acetylglucosamine (GlcNAc) transferase	0,0417709	2,57	MYBL2 siRNA up vs siRNA control
NADKD1	NAD kinase domain containing 1	0,0296512	2,57	MYBL2 siRNA up vs siRNA control
SUN2	Sad1 and UNC84 domain containing 2	0,00845921	2,57	MYBL2 siRNA up vs siRNA control
MSL1	male-specific lethal 1 homolog (Drosophila)	0,0231845	2,57	MYBL2 siRNA up vs siRNA control
PRDM8	PR domain containing 8	0,0108524	2,57	MYBL2 siRNA up vs siRNA control
SRRM1	serine/arginine repetitive matrix 1	0,0112565	2,57	MYBL2 siRNA up vs siRNA control
CAND1	cullin-associated and neddylation-dissociated 1	0,00424553	2,56	MYBL2 siRNA up vs siRNA control
TMEM123	transmembrane protein 123	0,0421781	2,56	MYBL2 siRNA up vs siRNA control
CLK4	CDC-like kinase 4	0,0212443	2,56	MYBL2 siRNA up vs siRNA control
HADHB	hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase (trifunctiona	0,0138345	2,56	MYBL2 siRNA up vs siRNA control
CERKL	ceramide kinase-like	0,0226228	2,56	MYBL2 siRNA up vs siRNA control
MARCH5	membrane-associated ring finger (C3HC4) 5	0,0450836	2,55	MYBL2 siRNA up vs siRNA control
PIGX	phosphatidylinositol glycan anchor biosynthesis, class X	0,0429063	2,55	MYBL2 siRNA up vs siRNA control
MYH9	myosin, heavy chain 9, non-muscle	0,0448246	2,55	MYBL2 siRNA up vs siRNA control
ERLIN2	ER lipid raft associated 2	0,0315827	2,55	MYBL2 siRNA up vs siRNA control
AMPD3	adenosine monophosphate deaminase 3	0,0213159	2,55	MYBL2 siRNA up vs siRNA control
CCNI	cyclin I	0,0490873	2,55	MYBL2 siRNA up vs siRNA control
CNOT1	CCR4-NOT transcription complex, subunit 1	0,0290442	2,55	MYBL2 siRNA up vs siRNA control
CCM2	cerebral cavernous malformation 2	0,00203996	2,55	MYBL2 siRNA up vs siRNA control
USP20	ubiquitin specific peptidase 20	0,0430199	2,55	MYBL2 siRNA up vs siRNA control
ZNHIT6	zinc finger, HIT-type containing 6	0,0158385	2,55	MYBL2 siRNA up vs siRNA control
SEC16A	SEC16 homolog A (S, cerevisiae)	0,00362922	2,54	MYBL2 siRNA up vs siRNA control
MAX	MYC associated factor X	0,00381351	2,54	MYBL2 siRNA up vs siRNA control
MTHFD2	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate c	0,0295999	2,54	MYBL2 siRNA up vs siRNA control
PBRM1	polybromo 1	0,0089212	2,54	MYBL2 siRNA up vs siRNA control
ZNF549	zinc finger protein 549	0,0355227	2,54	MYBL2 siRNA up vs siRNA control
PACS1	phosphofurin acidic cluster sorting protein 1	0,0136892	2,54	MYBL2 siRNA up vs siRNA control

MIR4680 /// PDCD4	microRNA 4680 /// programmed cell death 4 (neoplastic transformation inhibitor)	0,00395989	2,54	MYBL2 siRNA up vs siRNA control
CCDC88C	coiled-coil domain containing 88C	0,0379292	2,54	MYBL2 siRNA up vs siRNA control
PDS5A	PDS5, regulator of cohesion maintenance, homolog A (S, cerevisiae)	0,0357324	2,54	MYBL2 siRNA up vs siRNA control
SRSF11	serine/arginine-rich splicing factor 11	0,0264367	2,54	MYBL2 siRNA up vs siRNA control
SCLY /// UBE2F-SCLY	selenocysteine lyase /// UBE2F-SCLY readthrough	0,0172106	2,54	MYBL2 siRNA up vs siRNA control
MLL	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	0,00466454	2,54	MYBL2 siRNA up vs siRNA control
S100PBP	S100P binding protein	0,0347919	2,54	MYBL2 siRNA up vs siRNA control
CD164	CD164 molecule, sialomucin	0,0252506	2,54	MYBL2 siRNA up vs siRNA control
WASF2	WAS protein family, member 2	0,00577418	2,54	MYBL2 siRNA up vs siRNA control
SMC1A	structural maintenance of chromosomes 1A	0,0147436	2,54	MYBL2 siRNA up vs siRNA control
ITGAL	integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alph	0,00118702	2,54	MYBL2 siRNA up vs siRNA control
THOC2	THO complex 2	0,0357195	2,54	MYBL2 siRNA up vs siRNA control
MAT2B	methionine adenosyltransferase II, beta	0,0346073	2,53	MYBL2 siRNA up vs siRNA control
DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	0,0319609	2,53	MYBL2 siRNA up vs siRNA control
FNBP1	formin binding protein 1	0,031047	2,53	MYBL2 siRNA up vs siRNA control
MYH9	myosin, heavy chain 9, non-muscle	0,0203043	2,53	MYBL2 siRNA up vs siRNA control
ARHGAP35	Rho GTPase activating protein 35	0,0178466	2,53	MYBL2 siRNA up vs siRNA control
APP	amyloid beta (A4) precursor protein	0,011757	2,53	MYBL2 siRNA up vs siRNA control
VPS26A	vacuolar protein sorting 26 homolog A (S, pombe)	0,0154505	2,53	MYBL2 siRNA up vs siRNA control
CPSF6	cleavage and polyadenylation specific factor 6, 68kDa	0,0492005	2,53	MYBL2 siRNA up vs siRNA control
SMAD5	SMAD family member 5	0,0453973	2,53	MYBL2 siRNA up vs siRNA control
KDELR1	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1	0,00798223	2,52	MYBL2 siRNA up vs siRNA control
TFG	TRK-fused gene	0,00337126	2,52	MYBL2 siRNA up vs siRNA control
CTBP1	C-terminal binding protein 1	0,0389008	2,52	MYBL2 siRNA up vs siRNA control
EIF3A	eukaryotic translation initiation factor 3, subunit A	0,0167632	2,52	MYBL2 siRNA up vs siRNA control
SRSF5	serine/arginine-rich splicing factor 5	0,0307645	2,52	MYBL2 siRNA up vs siRNA control
EZR	ezrin	0,00754201	2,52	MYBL2 siRNA up vs siRNA control
LRRFIP1	leucine rich repeat (in FLII) interacting protein 1	0,0193565	2,52	MYBL2 siRNA up vs siRNA control
EIF4G1	eukaryotic translation initiation factor 4 gamma, 1	0,0034899	2,51	MYBL2 siRNA up vs siRNA control
CUL4A	cullin 4A	0,0160125	2,51	MYBL2 siRNA up vs siRNA control
CMPK1	cytidine monophosphate (UMP-CMP) kinase 1, cytosolic	0,0195256	2,51	MYBL2 siRNA up vs siRNA control
ACSL4	acyl-CoA synthetase long-chain family member 4	0,00905059	2,51	MYBL2 siRNA up vs siRNA control
EIF2B1	eukaryotic translation initiation factor 2B, subunit 1 alpha, 26kDa	0,0429382	2,51	MYBL2 siRNA up vs siRNA control
CHD9	chromodomain helicase DNA binding protein 9	0,0441613	2,51	MYBL2 siRNA up vs siRNA control
ATL3	atlastin GTPase 3	0,00510169	2,51	MYBL2 siRNA up vs siRNA control
NNT	nicotinamide nucleotide transhydrogenase	0,0244032	2,51	MYBL2 siRNA up vs siRNA control
PAFAH1B1	platelet-activating factor acetylhydrolase 1b, regulatory subunit 1 (45kDa)	0,0240714	2,51	MYBL2 siRNA up vs siRNA control
EIF4G2	eukaryotic translation initiation factor 4 gamma, 2	0,0102633	2,51	MYBL2 siRNA up vs siRNA control
SEH1L	SEH1-like (S, cerevisiae)	0,0365266	2,51	MYBL2 siRNA up vs siRNA control
DPP8	dipeptidyl-peptidase 8	0,0202381	2,50	MYBL2 siRNA up vs siRNA control
ACAD10	acyl-CoA dehydrogenase family, member 10	0,00368919	2,50	MYBL2 siRNA up vs siRNA control
CLTA	clathrin, light chain A	0,0319306	2,50	MYBL2 siRNA up vs siRNA control
ARGLU1	arginine and glutamate rich 1	0,040429	2,50	MYBL2 siRNA up vs siRNA control

OR6C76	olfactory receptor, family 6, subfamily C, member 76	8,90E-05	-2,53	MYBL2 siRNA down vs siRNA control
NBR1	neighbor of BRCA1 gene 1	0,0310742	-2,59	MYBL2 siRNA down vs siRNA control
CT47A1-A12 ///CT47B1	cancer/testis antigen family 47, member A1 /// cancer/testis antigen family 47, member	0,0265592	-2,60	MYBL2 siRNA down vs siRNA control
RHCE /// RHD	Rh blood group, CcEe antigens /// Rh blood group, D antigen	0,0198994	-2,62	MYBL2 siRNA down vs siRNA control
LOC728752	uncharacterized LOC728752	0,0193945	-2,62	MYBL2 siRNA down vs siRNA control
IFNA21	interferon, alpha 21	0,000733128	-2,63	MYBL2 siRNA down vs siRNA control
OR5AS1	olfactory receptor, family 5, subfamily AS, member 1	0,0129413	-2,66	MYBL2 siRNA down vs siRNA control
ARVP6125	uncharacterized LOC442092	0,00621794	-2,98	MYBL2 siRNA down vs siRNA control
LOC728752	uncharacterized LOC728752	0,0122047	-3,24	MYBL2 siRNA down vs siRNA control

Table 2. List of genes with differential expression in CD34$^+$ cells transfected with *MYBL2*−IRES−EGFP over-expression plasmid compared to CD34$^+$ cells transfected with IRES−EGFP control plasmid. The table shows a list of 41 genes obtained by establishing ±1.5 FC and p<0.05 criteria to the ANOVA test.

Gene Symbol	Gene Name	p	FD	Description
NPTX1	neuronal pentraxin I	0,0140575	2,68	MYBL2−IRES−EGFP up vs plasmid control
C3	complement component 3	0,0260288	2,42	MYBL2−IRES−EGFP up vs plasmid control
COL6A1	collagen, type VI, alpha 1	0,00898002	2,25	MYBL2−IRES−EGFP up vs plasmid control
MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	0,0325455	2,20	MYBL2−IRES−EGFP up vs plasmid control
CD22	CD22 molecule	0,0351567	2,14	MYBL2−IRES−EGFP up vs plasmid control
CD22	CD22 molecule	0,0449886	2,05	MYBL2−IRES−EGFP up vs plasmid control
SLAMF7	SLAM family member 7	0,0367037	2,02	MYBL2−IRES−EGFP up vs plasmid control
KCNMB1	potassium large conductance calcium-activated channel, subfamily M, beta member 1	0,0218115	1,97	MYBL2−IRES−EGFP up vs plasmid control
LPXN	leupaxin	0,00913041	1,90	MYBL2−IRES−EGFP up vs plasmid control
ALOX5	arachidonate 5-lipoxygenase	0,0460897	1,88	MYBL2−IRES−EGFP up vs plasmid control
MLLT4	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocate	0,0141795	1,74	MYBL2−IRES−EGFP up vs plasmid control
CLCN4	chloride channel, voltage-sensitive 4	0,0367619	1,72	MYBL2−IRES−EGFP up vs plasmid control
USP37	ubiquitin specific peptidase 37	0,00187089	1,71	MYBL2−IRES−EGFP up vs plasmid control
FILIP1L	filamin A interacting protein 1-like	0,0134711	1,67	MYBL2−IRES−EGFP up vs plasmid control
FBLN1	fibulin 1	0,0347557	1,66	MYBL2−IRES−EGFP up vs plasmid control
FXYD6 /// FXYD6-FXYD2	FXYD domain containing ion transport regulator 6 /// FXYD6-FXYD2 readthrough	0,0480762	1,65	MYBL2−IRES−EGFP up vs plasmid control
C3AR1	complement component 3a receptor 1	0,0142257	1,64	MYBL2−IRES−EGFP up vs plasmid control
FXYD6	FXYD domain containing ion transport regulator 6	0,0276055	1,64	MYBL2−IRES−EGFP up vs plasmid control
SH3GL3	SH3-domain GRB2-like 3	0,0376929	1,62	MYBL2−IRES−EGFP up vs plasmid control
OR51Q1	olfactory receptor, family 51, subfamily Q, member 1	0,0182545	1,62	MYBL2−IRES−EGFP up vs plasmid control
GPR132	G protein-coupled receptor 132	0,0129663	1,60	MYBL2−IRES−EGFP up vs plasmid control
SHROOM2	shroom family member 2	0,00613649	1,60	MYBL2−IRES−EGFP up vs plasmid control
TMTC3	transmembrane and tetratricopeptide repeat containing 3	0,0370043	1,60	MYBL2−IRES−EGFP up vs plasmid control
SORBS1	sorbin and SH3 domain containing 1	0,00280593	1,58	MYBL2−IRES−EGFP up vs plasmid control
BCL2L11	BCL2-like 11 (apoptosis facilitator)	0,00029864	1,57	MYBL2−IRES−EGFP up vs plasmid control
CLEC5A	C-type lectin domain family 5, member A	0,0386692	1,55	MYBL2−IRES−EGFP up vs plasmid control
LOC100653254	putative uncharacterized protein UNQ9165/PRO28630-like	0,0482919	1,55	MYBL2−IRES−EGFP up vs plasmid control
CLCN4	chloride channel, voltage-sensitive 4	0,00601078	1,53	MYBL2−IRES−EGFP up vs plasmid control
LOC100134259	uncharacterized LOC100134259	0,0209598	1,52	MYBL2−IRES−EGFP up vs plasmid control
FOSB	FBJ murine osteosarcoma viral oncogene homolog B	0,0408242	-1,51	MYBL2−IRES−EGFP down vs plasmid control
C21orf59 /// TCP10L	chromosome 21 open reading frame 59 /// t-complex 10 (mouse)-like	0,0196357	-1,52	MYBL2−IRES−EGFP down vs plasmid control
CREBRF	CREB3 regulatory factor	0,0098583	-1,53	MYBL2−IRES−EGFP down vs plasmid control
FAM43A	family with sequence similarity 43, member A	0,00498002	-1,54	MYBL2−IRES−EGFP down vs plasmid control
CPED1	cadherin-like and PC-esterase domain containing 1	0,0255062	-1,57	MYBL2−IRES−EGFP down vs plasmid control
KIAA0226	KIAA0226	0,0134624	-1,58	MYBL2−IRES−EGFP down vs plasmid control
DNAJC12	DnaJ (Hsp40) homolog, subfamily C, member 12	0,00266329	-1,60	MYBL2−IRES−EGFP down vs plasmid control
CNIH	cornichon homolog (Drosophila)	0,0163365	-1,60	MYBL2−IRES−EGFP down vs plasmid control

CD276	CD276 molecule	0,0148011	-1,62	MYBL2−IRES−EGFP down vs plasmid control
CTHRC1	collagen triple helix repeat containing 1	0,0164022	-1,64	MYBL2−IRES−EGFP down vs plasmid control
RARRES3	retinoic acid receptor responder (tazarotene induced) 3	0,0180629	-1,67	MYBL2−IRES−EGFP down vs plasmid control
SOX8	SRY (sex determining region Y)-box 8	0,0449126	-1,97	MYBL2−IRES−EGFP down vs plasmid control