

Inferencia de redes moleculares proteostáticas de relevancia en el mieloma múltiple

Grupo 71

Monografía Ciclo de Metodología Científica II- año 2018

Facultad de Medicina, Universidad de la República

Autores:

Br. Baldi, Mónica

Br. Cabrera, Fabián

Br. Dall'Oglio, Ximena

Br. Damborena, Julieta

Br. de León, Silvana

Br. Estefanell, Rocío

Tutor: Dr. Horacio Botti

MD PhD Estudiante de Maestría en Bioinformática

Laboratorio de Biofísica Integrativa, Depto. de Biofísica y CEINBIO, Facultad de Medicina,

Universidad de la República

Índice

Resumen.....	3
Introducción	3
Objetivo.....	9
Metodología	9
Resultados.....	12
Discusión	20
Conclusiones	24
Referencias bibliográficas.....	24
Anexos.....	27

Resumen

La presente monografía se enmarca en el Ciclo de Metodología Científica II de la carrera Doctor en Medicina, Facultad de Medicina, UdelaR. El objetivo es identificar componentes del Sistema Ubiquitina Proteasoma (SUP) potencialmente involucrados en la patogenia del Mieloma Múltiple (MM). Se combinaron informaciones ortogonales sobre expresión génica, genes asociados con roles etiopatogénicos en el MM y componentes del SUP. Se realizó un análisis basado en teoría y estadística de redes para identificar módulos de expresión asociados a cada gen etiológicamente relacionados a la enfermedad. Se mapearon componentes del SUP en cada módulo y se seleccionó para mayor estudio el correspondiente a la enzima Caseína Kinasa 1B (CKS1B). Se desconoce el mecanismo de regulación de CKS1B (componente del SUP), por el SUP. Sugerimos un algoritmo de identificación de componentes probablemente implicados en este proceso y proponemos un mecanismo hipotético de regulación de CKS1B.

Palabras claves: Mieloma Múltiple, ubiquitina, proteasoma, análisis exploratorio de datos.

Introducción

El MM es una hemopatía maligna del grupo de las gamapatías monoclonales, caracterizada por la proliferación monoclonal atípica de una clona plasmocitaria. Las manifestaciones clínicas y lesiones orgánicas observadas en los pacientes con MM son: lesiones óseas, hipercalcemia, insuficiencia renal, anemia y otras (infecciones bacterianas recurrentes, hiperviscosidad y amiloidosis) (1). En Uruguay y en el resto del mundo representa la segunda neoplasia hematológica maligna más frecuente. A nivel general representa el 1% de las neoplasias, afectando a personas con una edad media de 65 años (un 15% son menores de 60 años) con una estimación de 120 casos nuevos al año. Tiene una incidencia de 4 casos/100000 habitantes/año (2). A pesar de que la mediana de supervivencia en pacientes con MM es de aproximadamente 5 a 7 años, esta depende de factores individuales como la carga tumoral (estadio), biología (anomalías citogenéticas) y respuesta al tratamiento (3).

El MM suele comenzar por una proliferación de clones que generan la sobreproducción y secreción de gammaglobulina clonal (IgG/A/D/M o de cadenas ligeras Kappa o Lambda), la cual se da en el centro postgerminal a nivel ganglionar, fase llamada gamapatía monoclonal de significado incierto (MGUS). Posteriormente esta evoluciona a una fase de Mieloma Múltiple Asintomático (MMA), para finalmente pasar al MM que termina infiltrando la médula ósea (MO) (1). Esta hemopatía maligna se caracteriza por la interacción entre las células plasmáticas y el estroma medular, llevando a la producción de citoquinas (IL6, IL1, TNF) favoreciendo con esto la proliferación plasmocitaria y ejerciendo mecanismos antiapoptóticos (4). La etiología de

la enfermedad es desconocida pero se ha asociado con exposición al benceno, pesticidas e insecticidas (4). No obstante estas asociaciones no han sido bien corroboradas (5). Investigaciones previas demostraron que los TLR juegan un papel importante en la supervivencia y en la capacidad de evadir el sistema inmunológico en algunos tipos de MM (6). Los factores a recabar en la anamnesis asociados con un mayor riesgo de desarrollar MM incluyen edad avanzada, sexo masculino, raza afroamericana, exposición a la radiación, antecedentes familiares de la enfermedad, obesidad y / o antecedentes personales de gamapatía monoclonal de significado indeterminado (MGUS) (7).

El diagnóstico de MM requiere el establecimiento por medios paraclínicos de una infiltración por células plasmáticas clonales en MO mayor o igual al 10% o plasmocitoma demostrado y uno o más de los siguientes ítems:

Evidencia de daño de órgano blanco; hipercalcemia (calcio sérico mayor a 11mg/dl); insuficiencia renal (clearance de creatinina mayor a 40ml/min o creatininemia mayor 2mg/dl); anemia (Hb menor a 10g/dl o descenso de 2 g/dl por debajo del límite inferior de normalidad); lesiones óseas (una o más lesiones osteolíticas en radiografía, TAC o PET-TC) y cualquiera de los siguientes biomarcadores de malignidad (uno o más): infiltración plasmocitaria en MO mayor o igual al 60%, cadena liviana comprometida/ no comprometida >100 y una o más lesiones focales óseas por RNM (cada lesión debe ser mayor a 5mm)(4). Se han hallado variaciones génicas relacionadas con el MM en: CCND1, HRAS, NRAS, CDK4, FGFR3, BRAF, PTPN11, etc (8).

Aunque tradicionalmente se considera al MM como una enfermedad única, en realidad es un conjunto de neoplasias malignas citogenéticamente heterogéneas (3). En los últimos años se han publicado numerosos estudios que utilizaron Hibridación Fluorescente In Situ (FISH), cariotipificación y estudios de secuenciación dirigida (9) (10) (11) (12). En relación a estos estudios se concluyó que existen eventos primarios y secundarios que causan el MM. Los eventos primarios a su vez se dividen en hiperploides (HRD) y no hiperploides (non-HRD). Dentro de los HRD se encuentran las trisomías: 3, 5, 7, 9, 11, 15, 19 y/o 21. En los non-HRD están las translocaciones asociadas a las cadenas pesadas de inmunoglobulinas (IGH), las cuales comprenden: t(11; 14), t(4;14), t(14; 16), t(6; 14) y t(14; 20). Por otra parte, los eventos secundarios son aquellos que le confieren una ventaja a un subclon determinado sobre la población, permitiendo la progresión tumoral. Siendo éstos los cambios en el número de copias (CNVs), translocaciones del gen Myc, mutaciones somáticas que afectan Kinasas Activadas Por Mitógenos (MAPKs), al Factor Nuclear Kappa B (NFκB) y vías de reparación del ADN. Los eventos primarios se detectan en la etapa MGUS, mientras que los eventos secundarios se detectan principalmente en el MM (1).

A pesar de que han surgido nuevos abordajes terapéuticos gracias al avance en los conocimientos de la etiopatogenia del MM, continúa siendo una enfermedad incurable. En este sentido se practican 2 abordajes terapéuticos; el agresivo con nuevos fármacos (Bortezomib (Bz) e inmunomoduladores como la Lenalidomida) y uno más conservador (donde se utilizan fármacos más antiguos como ciclofosfamida y dexametasona) (4). Otra opción utilizada como consolidación del tratamiento es el trasplante autólogo de progenitores hematopoyéticos (TPH) (4). Tiene tres mecanismos de acción, estos son: la inhibición del NFκB a través de la estabilización de IκB, la activación de la respuesta de la proteína desplegada por la inhibición del proteasoma debido al alto estrés del retículo endoplásmico y la estabilización de proteínas proapoptóticas como BAX y NOXA (13). Al prevenir la degradación proteica, el Bz afecta la activación de múltiples cascadas de señalización y regulación intracelulares; tiene un efecto citotóxico y enlentece el crecimiento de las células tumorales del MM. La seguridad y eficacia del Bz está avalada por tres estudios clínicos a una dosis de 1.3 mg/m² (2). Los efectos adversos más frecuentes que se pueden presentar son: astenia, síntomas gastrointestinales, trombocitopenia clínica y neuropatía periférica. Existe toxicidad neurológica periférica y renal causada por el Bz, que dificulta su uso en algunos pacientes (14).

En Uruguay el tratamiento se encuentra dentro de las prestaciones del Fondo Nacional de Recursos (FNR). Las directivas del FNR para su indicación son exclusivamente para la inducción de la remisión excluyendo el tratamiento de mantenimiento (2). En la primera línea de cobertura se encuentran aquellos con indicación de tratamiento farmacológico y candidatos a su vez para trasplante de MO, o aquellos que no lo son, pero cuentan con una citogenética de alto riesgo (2). También se incluye a pacientes que tuvieron recaídas por lo menos un año después de haber finalizado el tratamiento farmacológico, sin haber demostrado refractariedad al mismo (2). Los criterios de exclusión para el tratamiento farmacológico son: contar con una mala calidad de vida y/o pobre expectativa de la misma (comorbilidades); enfermedades psiquiátricas severas; mala adhesión al tratamiento; embarazo o no consentimiento de medidas anticonceptivas, y lactancia o no consentimiento para suspenderla (2).

Dentro de los inmunomoduladores de segunda generación se destaca la Lenalidomida, que es un análogo de la Talidomida, pero con mayor potencia y menor toxicidad. Este fármaco es antiangiogénico, antineoplásico, inhibe la proliferación de células hematopoyéticas tumorales y potencia la inmunidad celular mediada por linfocitos T y células Natural Killer (2). Su target son los complejos E3 ligasa CUL4-RBX1-DDB1-CRBN (conocida como CRL4CRBN) (15).

Los sitios de acción de estos fármacos para el tratamiento del MM forman parte del SUP. El SUP es la vía principal para el recambio proteico en todas las células eucariotas, implicando dos eventos bioquímicos distintos. El primero es la poliubiquitinación de los sustratos proteicos,

(modificación postraduccion), y el segundo implica la proteólisis a través del complejo del proteasoma 26S. La ubiquitinación consiste en la adición de una o varias moléculas de ubiquitina. Esta es una proteína globular de 8.5 kDa y 76 aminoácidos (Figura 1) (16).

La adición de ubiquitina a un sustrato proteico implica las acciones consecutivas de tres familias principales de enzimas de ubiquitinación (E1 de activación, E2 de conjugación y E3 de ligación de ubiquitina) a través de una cascada enzimática coordinada. En el primer paso, se forma un enlace tioléster entre E1 y la glicina carboxiterminal de la ubiquitina. En un siguiente paso ocurre la transferencia de la ubiquitina activada de la enzima E1 a una enzima E2 a través de una reacción de transtiolación. Finalmente E3, que funciona como una proteína de andamiaje que une tanto a la enzima E2 como a la proteína diana, media la transferencia de ubiquitina del conjugado E2-ubiquitina, más comúnmente en el grupo ϵ -amino de un residuo de lisina, a la proteína formando así un enlace isopeptídico (16).

En este proceso también intervienen las enzimas deubiquitininasas (DUB) que se encargan de desconjugar la ubiquitina de las proteínas, teniendo un rol inverso a las enzimas de ubiquitinación. De esta manera el proceso de ubiquitinación es reversible (16).

Entre las funciones de la ubiquitinación se destaca su acción como marca para la posterior degradación de las proteínas en el proteasoma, así como su participación en la regulación de vías de señalización intracelular involucradas en funciones como el control de la apoptosis, la autofagia, el ciclo celular, la regulación transcripcional y la reparación del ADN (17).

El proteasoma 26S es un complejo multiproteico presente en eucariotas, arqueas y algunas bacterias, es de forma cilíndrica y se compone por con dos compuertas axiales y una cavidad central donde se encuentran los sitios catalíticos. En él se distinguen dos componentes: la partícula reguladora 19S que ocupa los extremos del cilindro y la partícula central 20S. La primera es la encargada de llevar a las proteínas ubiquitinadas a la cavidad central del proteasoma, donde la partícula 20S, a través de su actividad catalítica, la degrada en aminoácidos que son posteriormente reciclados por la célula (18).

Los mecanismos por los que las células tumorales son selectivamente sensibles al Bz no son comprendidos en su totalidad. Se piensa que como en las células neoclásicas hay una menor regulación del ciclo celular estas dependerían más de la degradación de los reguladores del ciclo celular mediada por el proteasoma (p21 y p27). Por ejemplo, sería el caso de la desregulación del Myc. Otro mecanismo sería la capacidad de los inhibidores del proteasoma de suprimir la vía tumorigénica de NF κ B (19).

Al emplear técnicas de secuenciación de alto rendimiento se identificaron los perfiles clonales en pequeños subgrupos de pacientes, lo que reveló la existencia de clones molecularmente diferentes al momento del diagnóstico, durante el tratamiento y en la recaída. Asimismo se ha

encontrado que la composición de las poblaciones clonales cambia bajo tratamiento, influencia del microambiente y con la progresión natural de la enfermedad. En general, la evolución clonal del mieloma es la base de su naturaleza heterogénea. La identificación de biomarcadores apropiados es necesaria dada la heterogeneidad y diversidad en las respuestas al tratamiento actual en pacientes y por el deseo de individualizar la farmacoterapia (20).

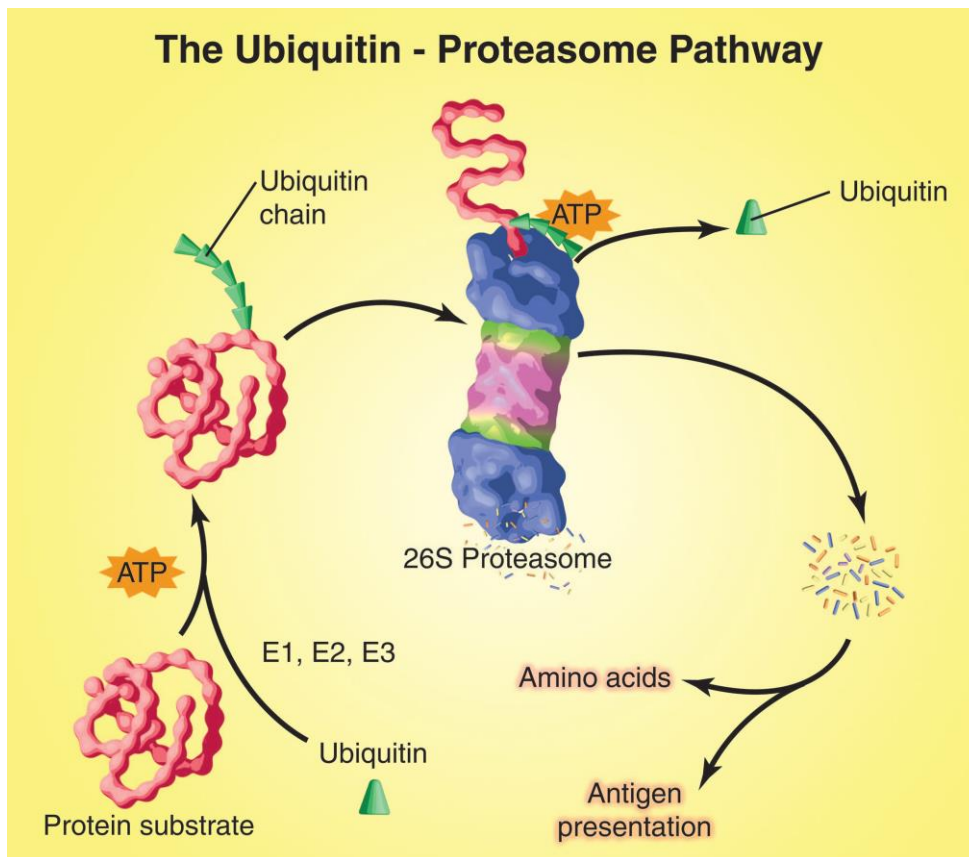


Figura 1. Sistema ubiquitina-proteasoma. El sistema ubiquitina proteasoma tiene varias funciones, en la figura obtenida de "Degradation by the Ubiquitin-Proteasome Pathway in Normal and Disease States" se muestra su función en la degradación (21).

Dado que el MM es una enfermedad etiopatogénicamente compleja y heterogénea se justifican los enfoques de aprendizaje automático y medicina de redes exploratorio que han caracterizado muchos trabajos recientes (22) (23).

Se define red (network) en su forma más general como: "una colección de cosas interconectadas", con frecuencia se usa de manera intercambiable con el término "graph" (24).

Un hallazgo importante de la última década es que las redes en sistemas naturales, tecnológicos y sociales no son al azar, sino que siguen una serie de principios básicos de organización en su

estructura y evolución que los distinguen de las redes conectadas al azar (24).

Una red se compone por nodos y aristas. Un nodo (o vértice) puede denotar proteínas, genes, metabolitos, moléculas de ARN o incluso enfermedades y fenotipos interactuando con otros nodos a través de las aristas. En los sistemas biológicos, estas interacciones pueden corresponder tanto a las uniones proteína-proteína, a acoplamiento metabólico, o pueden representar conexiones entre enfermedades basadas en un origen genético común o características fenotípicas compartidas (25).

El grado de un nodo es la cantidad de enlaces que se conectan a él. El grado de una proteína podría representar el número de proteínas con las que interactúa, mientras que el grado de una enfermedad puede representar el número de otras enfermedades que están asociadas con el mismo gen o que tienen un fenotipo común (25).

Un Módulo (o comunidad) es un subgrafo denso en la red que a menudo representa un conjunto de nodos que tienen un rol conjunto (25). Dentro de los módulos tenemos los topológicos, los funcionales y los de enfermedad. Los módulos topológicos corresponden a vecindades localmente densas de la red, de modo que, los nodos del módulo muestran una mayor tendencia a interactuar entre sí, que con los nodos fuera del módulo. Los módulos funcionales corresponden a vecindarios de red en los que existe una segregación estadísticamente significativa de nodos con funciones relacionadas. Un módulo de enfermedad representa un grupo de nodos cuya perturbación (mutaciones, eliminaciones, variaciones en el número de copias o cambios de expresión) puede vincularse a un fenotipo de enfermedad particular. La suposición tácita en la medicina de redes, es que los módulos topológicos, funcionales y de enfermedad se superponen, de modo que los módulos funcionales corresponden a módulos topológicos y una enfermedad puede verse como el desglose de un módulo funcional (25).

La medicina de redes se ocupa del “contexto molecular, el diagrama o red de enlace integrado entre los productos genéticos de una célula, siendo esencial para comprender las bases verdaderas para el fenotipo y patofenotipo” (26). Podemos decir que es un enfoque que busca interacciones de alto rendimiento entre componentes proteicos para entender mejor los procesos celulares complejos, y de esta manera ayuda a comprender las enfermedades humanas de forma integral (27).

Los estudios de redes tienen múltiples aplicaciones, biológicas y clínicas potencialmente útiles en las enfermedades humanas. El poder comprender las enfermedades en su contexto nos permite conocer algunas propiedades fundamentales de los genes que están involucrados. Además, permite tener una mejor comprensión de los efectos de la interconexión celular en la progresión de la enfermedad y por consiguiente de la identificación de sus genes, pudiendo ofrecer mejores blancos terapéuticos para el desarrollo de fármacos. Más allá de eso, ayudaría a

descubrir vínculos entre enfermedades, permitiendo entender, cómo diferentes fenotipos están vinculados a nivel molecular, por lo que existiría un origen común de algunas enfermedades relacionadas. Sin embargo, la superposición entre módulos dificulta el conocimiento total del metabolismo celular y de las enfermedades (25).

Se plantea como hipótesis que las pequeñas variaciones en la expresión génica, debidas a las distintas perturbaciones, las cuales pueden permitir identificar relaciones de expresión entre distintos genes, y así encontrar relaciones funcionales entre componentes del SUP y genes vinculados a la etiopatogenia molecular del MM.

Objetivo

- Reconocer posibles componentes del SUP potencialmente involucrados en la patogenia del mieloma múltiple.

Metodología

1. Se realizó una lectura detallada, y posterior presentación grupal de trabajos científicos relevantes en el MM, con los cuales se confeccionó una lista preliminar de los genes implicados en la etiopatogenia del MM según el conocimiento actual (junio de 2018) (1) (28). Dicha lista consiste en una planilla en formato Excel donde se detallan para cada gen la función normal (29).
2. Se chequeó presencia o ausencia de cada gen de la lista en la base de datos de la plataforma Malacards (8) (Anexo1).
3. Se registró la correspondencia entre probesets del chip Affymetrix con los genes humanos de la lista de genes asociados al MM.
4. Se creó una matriz donde se detalla la correlación lineal entre la señal de fluorescencia correspondiente a los probesets que representan los ARNs codificados por los genes de la lista de genes asociados al MM en el experimento Affymetrix con la totalidad de los probesets del chip (58438). Para obtenerla se empleó el programa lm en lenguaje R incluido en la distribución básica de paquetes. Esto se realizó utilizando el algoritmo que se detalla en el Anexo 2. Este algoritmo devuelve para cada probeset los siguientes datos; el coeficiente de determinación (R^2), el coeficiente de correlación lineal de Pearson (r), el valor p y el error en el coeficiente de regresión.
5. Se realizó la construcción de una red de relaciones de expresión centrada en los genes asociados a la etiopatogenia del MM utilizando el programa lm en lenguaje R con los paquetes igraph, rgefx, XML y Rook. Se usaron los coeficientes de regresión para definir pesos entre nodos, resultando en una red de 37343 nodos y 133006 aristas no dirigidas.

6. Se empleó el programa Gephi el análisis estadístico de la red creada anteriormente. El aspecto principal del análisis consistió en la identificación. Se han propuesto varios métodos para dividir una red en comunidades, como una forma de ajustar el nivel de descripción del sistema, pero también para identificar funcionalidades o relaciones subyacentes, a menudo desconocidas, entre los nodos (30). El algoritmo empleado por Gephi, se divide en dos fases que se repiten en forma iterativa. Primero, se asigna una comunidad diferente a cada nodo de la red. Entonces, en esta partición inicial hay tantas comunidades como nodos. Luego, para cada nodo A consideramos los vecinos B de A y se evalúa la ganancia de modularidad que tendría lugar eliminando A de su comunidad y colocándolo en la comunidad de B. El nodo A se coloca entonces en la comunidad para la cual esta ganancia es máxima. Este proceso se aplica de forma repetida y secuencial para todos los nodos hasta que no se pueda lograr más mejoras. Esta primera fase se detiene cuando se alcanza un máximo local de la modularidad, es decir, cuando ningún movimiento individual puede mejorar la modularidad. También se debe tener en cuenta que la salida del algoritmo depende del orden en que se consideran los nodos. La segunda fase del algoritmo consiste en construir una nueva red, cuyos nodos son ahora las comunidades encontradas durante la primera fase. Para hacerlo, los pesos de los enlaces entre los nuevos nodos están dados por la suma del peso de los enlaces entre los nodos en las dos comunidades correspondientes. Los enlaces entre nodos de la misma comunidad conducen a bucles automáticos para esta comunidad en la nueva red. Una vez que se completa esta segunda fase, es posible volver a aplicar la primera fase del algoritmo a la red ponderada resultante e iterar (31). De este proceso resultaron un total de 31 módulos. Luego cada módulo se subdividió en submódulos. A su vez estos submódulos fueron separados según tuvieran aristas con peso positivo o negativo. Se confeccionó una red simplificada utilizando el programa Gephi, para poder tener una visualización más gráfica de la red, en la que se observan los módulos y sus principales relaciones. Para realizarla se utilizó el gen más conectado de cada módulo y sus conexiones. La red fue coloreada de acuerdo a su centralidad.

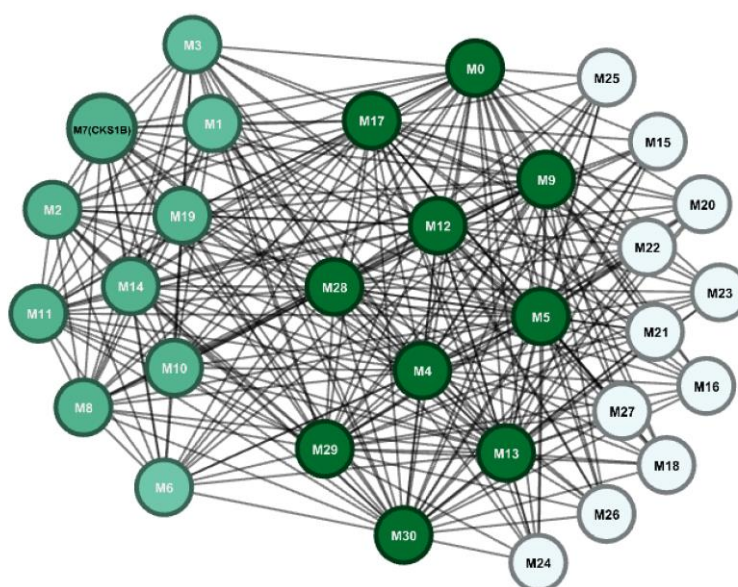


Figura 2. Red simplificada por módulos

7. Análisis biomédico; se prosiguió a seleccionar los probesets y a convertirlos con la base de datos online BioDBnet (32) a unigene IDs, los cuales posteriormente se ingresaron en Database for Annotation Visualization and Integrated Discovery (DAVID). DAVID es un recurso bioinformático que consiste en una base de datos biológicos integrados y de herramientas analíticas destinadas a extraer sistemáticamente el significado biológico de grandes listas de genes o proteínas. Para el análisis se tuvieron en cuenta solamente los resultados significativos (False Discovery Rate (FDR) menor a 0.05).
8. Por otro lado se realizó una comparación entre la lista de genes de cada uno de los módulos y una lista de genes implicados en el SUP, esta última se obtuvo de la bibliografía y de estudios no publicados (Botti et al, manuscrito en preparación)(33). La comparación fue realizada a través de la página web XLComparator (34). De esta manera se obtuvo una nueva lista con genes que se encuentran tanto en los módulos identificados como en el SUP. Se estudió al submódulo M7m1positivado que cuenta con pocos genes (29) en común con el SUP. Se destaca la presencia de ARNs que codifican la proteína CKS1B, ya que también se encuentra en la lista de genes asociados al MM.
9. La lista de 29 genes pertenecientes al SUP y al submódulo M7m1positivo fue analizada en la plataforma Reactome para identificar las vías en las que estos genes participan según el conocimiento actual (Figura 5). De este análisis identificamos 5 genes sin

participación en ninguna vía conocida. Estos fueron ENTHD1, NDFIP2, STXBP5L, VPS13D y WDR17.

Resultados

- **Identificación de genes asociados a la etiopatogenia del MM y de los grupos de sondas ("probesets") correspondientes.** De acuerdo con la revisión bibliográfica realizada identificamos artículos que proponen un conjunto de genes asociados a la etiopatogenia del MM (1)(28). Destacamos los trabajos recientes que utilizan métodos de análisis de firmas mutacionales (35) y métodos relacionados, los que señalan roles etiopatogénicos para nuevos candidatos: como por ejemplo AID (36) . La mayoría de estos nuevos candidatos no estaban presentes en la base de datos Malacards al momento de la consulta (julio de 2018). En la Tabla I (ANEXO 1) resumimos la información sobre genes asociados a la etiopatogenia del MM.
- **Cálculo de regresiones lineales simples en R.** Se usó la función lm de R para calcular y estimar los parámetros del modelo lineal:
$$y_i = \alpha + \beta x_i + \text{error}$$

Donde y es la intensidad de fluorescencia correspondiente al probeset de referencia y x es el valor de fluorescencia de otro probeset en la misma condición i. Beta, error, medida de ajuste ($R^2 = r^2$, de 0 a 1) y de significación estadística en comparaciones simples (valor de p, de 1 a 0) fueron calculados en R y registrados en una tabla individual según el algoritmo que se muestra en el ANEXO 2. Luego los valores de beta se registraron en una tabla única (disponible a solicitud por correo electrónico: silvanadeleongomez@gmail.com y/o rocioestefanell@gmail.com).
- **Reconstrucción de la red parcial de expresión usando igraph.** Decidimos reconstruir una red usando como nodos los probesets y como pesos de las aristas (interacciones) los coeficientes angulares (beta). Se trata de una red que representa relaciones de expresión. Si beta es significativamente mayor a cero hablamos de coexpresión, si beta es significativamente menor a 0 de antiexpresión. Para reconstruir esta red, usamos funciones de la distribución básica, el paquete igraph y paquetes de los que este depende en R para reconstruir la red de coexpresión/antiexpresión parcial de acuerdo a la matriz de $n = 58438$ por $m = 34$. El algoritmo empleado se detalló en la sección de materiales y métodos. Decidimos simplificar la red, descartando aristas cuyo peso absoluto fuera menor a 0,75, por lo tanto, la red con la que continuamos trabajando es parcial.

La red así obtenida posee 37342 nodos y 131837 aristas, con un grado medio de aristas / nodos 3,53 rango 1-34018.

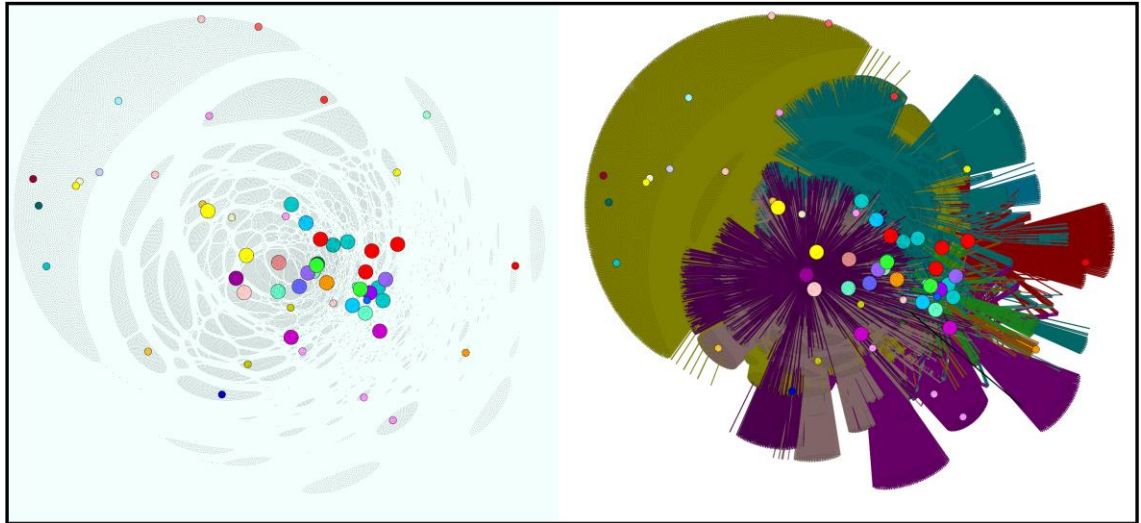


Figura 3. Representaciones de la red completa

- **Visualización y análisis con Gephi de la red parcial de coexpresión.** Generamos un archivo de salida de R compatible con el formato de archivos de ingreso de redes de Gephi (detalles en Materiales y Métodos) y representamos en el plano la red parcial de relaciones de expresión usando el programa Gephi, como se muestra en la Figura 3.

A continuación, calculamos estadísticos globales e identificamos módulos de probesets.

- **Inferencia de grupos funcionales usando DAVID.** De acuerdo con el análisis de módulos realizado en Gephi (comentado en materiales y métodos) identificamos 31 módulos, de estos 11 (M15,M16,M18,M20,M21, M22,M23,M24,M25,M26,M27) están compuestos por un único nodo y fueron analizados por separado. No se obtuvo ningún resultado significativo en este análisis.

Para los módulos restantes se generó una agrupación basada en los temas más repetidos y significativos de los resultados de clustering obtenidos en DAVID (Figura 4). Estos fueron: ribosomas, beta arrestinas, Subunidad alfa del factor 1 inducible por hipoxia (HIF1A), procesos mitocondriales, relación con la proteína MERLIN, procesos en relación al ARN y cadherinas. Ponemos concluir que en el MM se afectan varias vías celulares.

Módulo	# Nodos	Centralidad	Resultados según DAVID	Módulo funcional	FDR
M0+	118	1.0	PUBMED ID 25963833: Quantitative Proteomics Reveals Dynamic Interactions of the Minichromosome Maintenance Complex (MCM) in the Cellular Response to Etoposide Induced DNA Damage.	Mitocondria	5.13E-04
			PUBMED ID 23473034: GRSF1 regulates RNA processing in mitochondrial RNA granules		0.00700521
			64960:mitochondrial ribosomal protein S15(MRPS15) FDR 1.35E-04		5.57E-05
M0-	414	1.0	CGAP_EST_QUARTILE 20050:head and neck_neoplasia_3rd	Neoplasias	1.04E-31
M1m0+	7	0.7443444704390936	GO:0000184--nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	Metabolismo del ARN, ribosoma	3.15E-05
			UP_KEYWORDS Ribosomal protein FDR		7.01E-05
M1m0-	5		Sin resultados relevantes.		
M1m1+	25		Sin resultados relevantes.		
M1M1-	13		Mitochondrion	Mitocondria	0.01094116
M1m2+	31		Sin resultados relevantes.		
M1m2-	90		GO:0044429--mitochondrial part	Mitocondria	0.01094116
M1m3+	14		PUBMED ID 25963833: Quantitative Proteomics Reveals Dynamic Interactions of the Minichromosome Maintenance Complex (MCM) in the Cellular Response to Etoposide Induced DNA Damage.	Ribosoma	0.003473492
			6224:ribosomal protein S20(RPS20)		1.97E-04
M1m3-	8		Sin resultados significativos.		
M2m0+	169	0.775560441479816	PUBMED ID 24457600: The central role of EED in the orchestration of polycomb group complexes.	Polycomb, metabolismo del ARN, HUWE1, HIF1A, mitocondria	2.05E-08
			PUBMED ID 25324306: hnRNP A1 couples nuclear export and translation of specific mRNAs downstream of FGF-2/S6K2 signalling.		0.025816121
			PUBMED ID 25147182: Quantitative Lys-ε-Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1.		2.21E-07
			PUBMED ID 17620599: Functional specialization of beta-arrestin interactions revealed by proteomic analysis.		1.28E-02
			PUBMED ID 23398456: The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability.		2.24E-05
			GOTERM_CC_5 GO:0098798--mitochondrial protein complex		4.95E-11
			GOTERM_MF_4 GO:0003723--RNA binding		1.05E-04
			GOTERM_BP_FAT GO:0006402--mRNA catabolic process		0.00166344
M2m0-	35		Sin resultados significativos.		
M2m1+	125		Sin resultados relevantes.		
M2m1-	69		PUBMED ID 23398456: The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability	HIF1A, HUWE1, arrestinas, Polycomb, Merlin, metabolismo del ARN	2.98E-17
			PUBMED ID 25147182: Quantitative Lys-ε-Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 ligase HUWE1.		2.15E-12
			PUBMED ID 17620599: Functional specialization of beta-arrestin interactions revealed by proteomic analysis		3.56E-09
			PUBMED ID 24457600: The central role of EED in the orchestration of polycomb group complexes		6.72E-09
			PUBMED ID 26549023: The scaffold protein KSR1, a novel therapeutic target for the treatment of Merlin-deficient tumors		2.64E-07
			PUBMED ID 22658674: Insights into RNA biology from an atlas of mammalian mRNA-binding proteins		9.32E-06
M2m2+	24		BIOGRID_INTERACTION 6218:RPS17--ribosomal protein S17	Ribosoma	8.71E-04
M2m2-	32		PUBMED ID 17620599: Functional specialization of beta-arrestin interactions revealed by proteomic analysis	Arrestina	0,0290694364295962
M2m3+	2		Sin resultados significativos.		
M2m3-	18		GO:0022626--cytosolic ribosome	Ribosoma, HIF1A, metabolismo del ARN, arrestinas	1.03E-07
			PUBMED ID 23398456: The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability.		9.79E-05
			PUBMED ID 25324306: hnRNP A1 couples nuclear export and translation of specific mRNAs downstream of FGF-2/S6K2 signalling.		1.28E-04
			PUBMED ID 20020773: Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery.		0.001384258
			PUBMED ID 17620599: Functional specialization of beta-arrestin interactions revealed by proteomic analysis		0.005597954
M3m0+	18		Sin resultados relevantes.		
M3m0-	5		Sin resultados significativos.		
M3m1+	22		Sin resultados relevantes.		
M3m1-	5	0.7443444704390937	PUBMED ID 20020773: Global analysis of TDP-43 interacting proteins reveals strong association with RNA splicing and translation machinery.	Metabolismo del ARN, HIF1A	0.00146527
			PUBMED ID 23398456: The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability.		0.00467188

M3m2+	119		PUBMED ID 23398456: The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability.	HIF1A, metabolismo del ARN, Merlin	8,93E-09
			PUBMED ID 22681889: The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts.		7,44E-07
			PUBMED ID 26549023: The scaffold protein KSRI, a novel therapeutic target for the treatment of Merlin-deficient tumors		6,75E-05
M3m2-	94		Sin resultados relevantes.		
M4+	890	0.9999999999999999	Sin resultados relevantes.		
M4-	1447		GO:0022613-ribonucleoprotein complex biogenesis	Metabolismo del ARN	4.58973E-07
			GO:0003723-RNA binding		0.000171558
M5+	528	0.9999999999999999	PUBMED ID 19953087: Comparison of substrate specificity of the ubiquitin ligases Nedd4 and Nedd4-2 using proteome arrays.	Ubiquitinación, metabolismo del ARN	0,03802945
			GO:0016070-RNA metabolic process		0.007946461
M5-	1242		Sin resultados relevantes.		
M6+	573		GO:0005739-mitochondrion	Mitocondria	1.16E-10
M6-	430	0.7101582559579195	PUBMED ID 22751105: Multiple myeloma-associated chromosomal translocation activates orphan snoRNA ACA11 to suppress oxidative stress.	Metabolismo del ARN	6.68873E-06
			GO:0003723-RNA binding		1.82077E-12
M7m0+	19	0.7755604414798166	3921:RPSA-ribosomal protein SA	Ribosoma, HIF1A	1.05046E-18
			PUBMED ID 23398456: The P-body component USP52/PAN2 is a novel regulator of HIF1A mRNA stability.		1.66533E-13
M7m1+	244		Sin resultados relevantes.		
M7m1-	21		Sin resultados significativos.		
M8m0+	10		Sin resultados significativos.		
M8m0-	10		Sin resultados significativos.		
M8m1+	708	0.7755604414798166	GO:0003723-RNA binding	Metabolismo del ARN, Polycomb	1.61E-22
			PUBMED ID 24457600: The central role of EED in the orchestration of polycomb group complexes.		4.20E-25
M8m1-	291		21690:brain_neoplasia_3rd	Cadherinas	9.00E-19
			PUBMED ID 10835267: Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members.		1.04E-25
M9+	885	0.9999999999999999	PUBMED ID 22658674: Insights into RNA biology from an atlas of mammalian mRNA-binding proteins.	Metabolismo del ARN	5.09E-28
			PUBMED ID 22681889: The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts.		2.55E-13
M9-	2420		PCDHG	Cadherinas	8.95E-26
			PUBMED ID 10835267: Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members.		9.27E-16
M10+	266		GO:0005743-mitochondrial inner membrane	Mitocondria	5.86E-10
M10 -	1085	0.7755604414798166	21669:brain_neoplasia_3rd	Cadherinas	1.52E-21
			PUBMED ID 10835267: Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members.		8.15E-18
M11+	340	0.7755604414798166	Sin resultados relevantes.		
M11 -	403		PUBMED ID 17620599: Functional specialization of beta-arrestin interactions revealed by proteomic analysis.	Arrestinas	2.71E-05
M12+	292		PUBMED ID 24711643: Identifying biological pathways that underlie primordial short stature using network analysis.	Ubiquitinación	8.06E-09
			white blood cells_monocyte_3rd		0.04664048
M12-	1037	0.9999999999999999	PUBMED ID 10662547: Genomic organization of the family of CNR cadherin genes in mice and humans.	Cadherinas	3.60E-15
			PUBMED ID 10835267: Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members.		1.33E-08
			PCDHA		1.12E-18
M13+	1026		PCDHG	Cadherinas	1.09E-25
M13-	1325	0.9999999999999999	Transcription regulation	Metabolismo del ARN	0.006313435
M14+	1043	0.7755604414798166	PUBMED ID 23125841: Characterization of staufen1 ribonucleoproteins by mass spectrometry and biochemical analyses reveal the presence of diverse host proteins associated with human immunodeficiency virus type 1.	Ribosoma, metabolismo del ARN	3.20E-19
			GO:0003723-RNA binding		7.45E-32
			PUBMED ID 22658674: Insights into RNA biology from an atlas of mammalian mRNA-binding proteins.		1.25E-50
M14-	501		Sin resultados relevantes		
M17+	4268	0.9999999999999999	Sin resultados significativos.		
M17-	6061		GO:0016070-RNA metabolic process	Metabolismo del ARN	7.62E-06
M19+	465	0.7755604414798166	PCDHG	Cadherinas	3.24545E-31
M19-	703		PCDHG	Cadherinas	9.51709E-23
M28+	953		PCDHG	Cadherinas	1.43146E-27
			GO:0016052-carbohydrate catabolic process		0.01019607
M28-	2398	0.9999999999999999	PCDHA	Cadherinas	1.77636E-13
			PUBMED ID 10817752: Cadherin superfamily genes: functions, genomic organization, and neurologic diversity.		1.62E-14
			CGAP EST QUARTILE 21672:brain_neoplasia_3rd		1.17E-18
M29-	1116	0.9999999999999999	Sin resultados relevantes.		
M30+	2412		Sin resultados relevantes.		
M30-	2412	1.0	PCDHG	Cadherinas	4.0593E-24
			PCDHA		7.08E-16

Figura 4. Resumen de resultados

El metabolismo del ARN estaría significativamente relacionado con el grupo de módulos M8m1+, M9+, M14+ y M30+. En el M8m1+ el gen central es el APOBEC3B, un miembro de la familia de genes que codifica proteínas encargadas de la edición del ARN y del crecimiento y control del ciclo celular (29). Además, actúa como un inhibidor de la replicación de retrovirus y la movilidad de retrotransposones a través de mecanismos dependientes e independientes de la actividad deaminasa (29).

CCND1 codifica para una proteína perteneciente a la familia de las ciclinas. La mutación, la amplificación y la sobreexpresión de este gen se observa frecuentemente involucrado en la tumorigénesis (29).

Los módulos M14+ y M8m1+ se pueden agrupar funcionalmente porque sus componentes se pudieron asociar significativamente con una publicación sobre el sistema de regulación epigenética de las proteínas del grupo polcomb (37), este grupo de proteínas intervienen como parte de la regulación epigenética de la transcripción durante la diferenciación celular, la pluripotencia de células madre y la progresión neoplásica (29).

Así mismo el M30+ tiene como gen central ANP32E que codifica para una histona chaperona que interviene específicamente en la degradación de la histona H2A (29). Esta es un blanco del grupo de proteínas Polcomb para conseguir la regulación epigenética (37).

Los módulos M2m1-, M5+ y M12+ se pueden agrupar por estar enriquecidos en el proceso de ubiquitinación. El gen central del M2m1- es COBLL1 relacionado con la resistencia farmacológica y la transformación blástica en la leucemia mieloide crónica (38). BLC9, el gen central de M5+, es un gen que participa de la transducción de señales a través de la vía Wnt promoviendo la actividad transcripcional y que se ha reportado es objeto de translocaciones en tumores malignos de células B donde se evidencian anomalías de 1q21 (39). El gen central del M12+ es NSD2, para el que se han reportado siete variantes de splicing que dan lugar a al menos tres proteínas (MMSET1, MMSET2 y RE-IIBP) las que juegan un papel importante en el desarrollo. Esta está en la base de la etiopatogenia del MM por la translocación 4,14 y del síndrome de Wolf Hirschhorn (WHS) que se da por una delección del brazo corto distal del cromosoma 4 (29).

El análisis de lista de genes en módulos en DAVID permite agrupar funcionalmente a los módulos M2m0+, M2m1-, M2m3-, M3m1-, M3m2+ y el M7m0+, los que se caracterizan por estar enriquecidos por componentes de la vía de HIF1A. Este grupo funcional (grupo "HIF1A") se superpone parcialmente con el grupo anterior

(“ubiquitinación”). Para el M2m3 el gen central es TRAF2, que regula la activación de NF-kappa-B y JNK y desempeña un papel central en la regulación de la supervivencia celular y la apoptosis. Tiene actividad de ubiquitina-proteína ligasa E3 donde promueve la ubiquitinación de proteínas diana al ponerlas en contacto con otras ligasas de ubiquitina E3 (29). En M3m1 CCND3 es el gen central; este codifica para una proteína perteneciente a la familia de las ciclinas la cual es necesaria para la transición del ciclo celular de G1 a S (29). M7 tiene como gen central HOXB3 que codifica una proteína que funciona como un factor de transcripción específico de secuencia involucrada en el desarrollo. El aumento de la expresión de este gen se asocia con leucemia mieloide aguda (29).

En la agrupación de las cadherinas estarían implicados los módulos M8m1-, M9-, M10-, M12-, M13+, M19+, M19-, M28+, M28-, M29+ y M30. M8m1 tiene como gen central APOBEC3B del cual ya nos referimos anteriormente al igual que el M9 con CCND1, el M10 y M12 con NSD2, y el M30 con ANP32E. M13 tiene como gen central el FGFR3, que codifica un miembro de la familia de los factores de crecimiento de fibroblastos (FGFR) (29). M19 posee como gen central el EGR1, que regula la transcripción de numerosos genes dianas, desempeñando un papel importante en la regulación de la respuesta a los factores de crecimiento y al daño al ADN. Además de esto juega un papel en la regulación de la supervivencia celular, activa la expresión de p53 y TGFB por lo que ayuda a prevenir la formación de tumores (29). M28 tiene como gen central el gen ATP1A2, miembro de la subfamilia p5 de las ATP-ases, su función es transportar cationes inorgánicos y otros sustratos. Se relaciona con la enfermedad de Parkinson (29). M29 tiene como gen central al ATP13A2, que codifica un miembro de ATP-ases que transportan cationes inorgánicos y otros sustratos desempeñando un papel en la homeostasis de los cationes intracelulares y el mantenimiento de la integridad neuronal. También es necesario para un adecuado mantenimiento lisosomal y mitocondrial (29). Además, en otros módulos (M9-, M12-, M13+, M19+, M19-, M28+, M28-, M30-) resultaron significativas las protocadherinas en cluster (PCDH).

En la agrupación funcional “beta arrestinas” se incluyeron M2m2-, M2m3- y M11-. El gen central perteneciente al módulo M2m2 es el FAM46C también denominado TENT5C (Terminal Nucleotidiltransferasa 5C), es una probable nucleotidiltransferasa que puede actuar como una poli (A) ARN polimerasa no canónica. Las enfermedades asociadas con este gen incluyen mieloma múltiple (29). En M2M3 el gen central es TRAF del cual ya nos referimos anteriormente en el M2m3, lo mismo sucede con el M11 con COBLL1 ya mencionado en los módulos M2m0 y M2m1.

En la agrupación funcional referente a ribosoma, se hace referencia principalmente a las proteínas ribosomales, procesos relacionados al ARN ribosomal. Los módulos implicados en esta categoría son M1M0+, M1M3+, M2M3-, M14+. El M1M0 tiene como gen central COBLL1 que ya fue descrito anteriormente, como es el caso de M1M3 con APOBEC3D y APOBEC3B, M14 con su gen central que es el CCND1, y M2M3 con el gen central TRAF2. En esta categoría el control del ciclo celular juega un rol fundamental, tanto el M2M3 como en el M14 los genes centrales tienen roles en el ciclo celular en diferentes puntos.

Haciendo referencia a la agrupación de la mitocondria se destacan los módulos M6 y M10. El gen central de ambos es el NSD2 que ya fue comentado anteriormente.

Por último, en la agrupación relacionada a las proteínas MERLIN está solamente el módulo M3m2, que tiene como gen central al BCL9L. Se encontró como relevante en el análisis de este módulo un artículo que hace referencia a la proteína MERLIN.

- **Construcción de una hipótesis de trabajo a validar experimentalmente.** Como fue explicado en Materiales y Métodos, se diseñó un modelo de regulación de CKS1B, utilizando las cinco proteínas que no se asocian a ningún proceso metabólico: ENTHD1, NDFIP2, STXBP5L, VPS13D y WDR17 (Figura 5). A través de interpro, se analizaron los dominios de dichas proteínas para asociarlos a funciones de proteínas homólogas familiares con PFAM. Mediante este proceso se obtuvo que el gen WDR17, que codifica para una proteína del mismo nombre, la cual se une a un complejo formado por SKP1-CULLIN-FBOX, que tiene actividad E3 ligasa y cumple función de andamio; se encuentra implicada en señalización, regulación de transcripción, control del ciclo celular y apoptosis. ENTHD1 codifica para la proteína ENTD1, que tiene dominios Inositoltrifosfato (IP3), Fosfatidil Inositoltrifosfato (PI3P) y tubula liposomas; por lo que su ubicación estaría próxima a la superficie de las membranas.

NDFIP2 codifica para la proteína NFIP2, una proteína transmembrana de los endosomas que se une y afecta la actividad de NEDD4 de ubiquitinligasas E3.

STXBP5L (Syntaxinabindingprotein 5 like) codifica para la proteína STB5L, que contiene varios dominios WD y un dominio N-SNARE terminal que lo conecta a la membrana. Se une a la Sintaxina 19 y forma curvaturas en la membrana favoreciendo la endocitosis.

VPS13D codifica para VP13D, una gran proteína que contiene un dominio UBA (Asociado a Ubiquitina), se encuentra ubicada en las vacuolas y se asocia con el reciclado de las proteínas. Está involucrada con la reparación del ADN NER y el

control del ciclo celular. Une poliubiquitinas, las dirige al proteasoma 26S y limita la elongación de las mismas.

Tomando en cuenta esta información se plantea una hipótesis de la regulación de CKS1B (Figura 6).

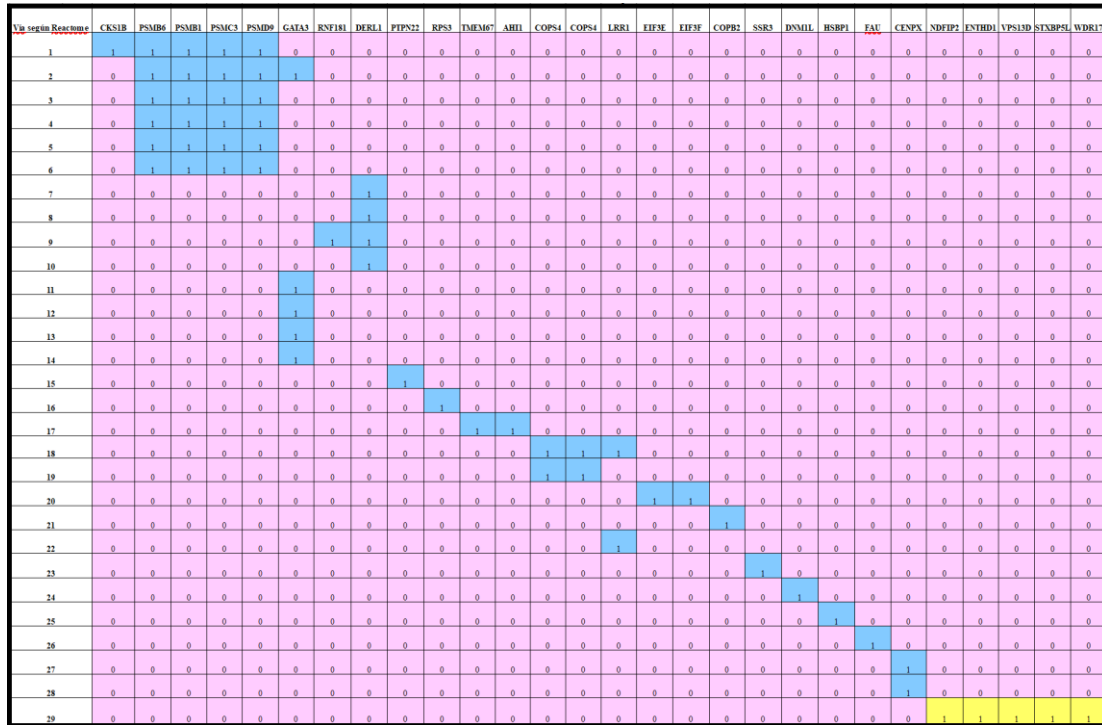


Figura 5. Heat map. 1-R-HSA-453279.3 Mitotic G1-G1/S phases. 2- Transcriptional regulation by RUNX1. 3-Signaling by NOTCH4. 4- Regulation of RUNX2 expression and activity . 5- Hedgehog 'on' state. 6- Regulation of mitotic cell cycle. 7- ABC-family proteins mediated transport. 8- Defective CFTR causes cystic fibrosis. 9- E3 ubiquitin ligases ubiquitinate target proteins. 10- Asparagine N-linked glycosylation. 11- Interleuquin 4 and interleuquin 13 signaling. 12- ESR-mediated signaling. 13-Deubiquitination. 14- Factors involved in megakaryocyte development and platelet production. 15- TCR signaling. 16- Major pathway of rRNA processing in the nucleolus and cytosol. 17- Cilium Assembly. 18- Neddylation. 19- Nucleotide Excision Repair. 20- Eukaryotic Translation Initiation. 21- Intra-Golgi and retrograde Golgi-to-ER traffic. 22- Class I MHC mediated antigen processing & presentation. 23- SRP-dependent cotranslational protein targeting to membrane. 24- Apoptotic execution phase. 25-Cellular response to heat stress. 26-Major pathway of rRNA processing in the nucleolus and cytosol. 27- Nucleosome assembly. 28- Fanconi Anemia Pathway. 29- Proceso desconocido

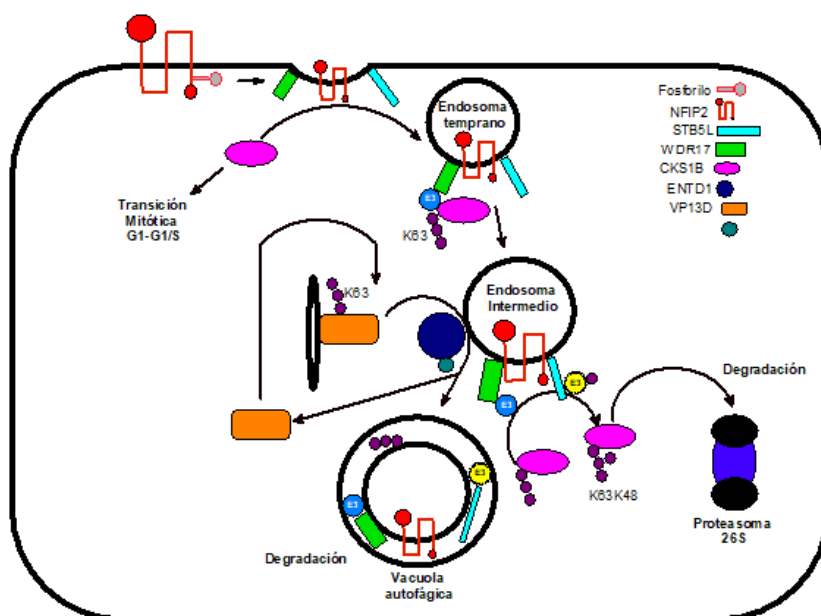


Figura 6. Modelo de degradación coordinada de CKS1B.

La proteína CKS1B podría tomar dos caminos posibles: llevar a cabo su función (transición mitótica G1-G1/S) o ingresar a su proceso de degradación. Dicha degradación estaría mediada por un proceso de endocitosis, que comenzaría tras una defosforilación de NFIP2. En el endosoma temprano ocurriría una interacción de CKS1B con WDR17, el cual tiene una función de andamio y permite la interacción con SKP1-CULLIN-FBOX facilitando así la formación del complejo E3 ligasa, y junto con K63 poliubiquitinaria al sustrato CKS1B. Luego de la formación del endosoma intermedio, la CKS1B/K63, interactúa con STB5L, que posiblemente recluta un complejo E3, que lo poliubiquitinaria en conjunto con K48, generando una poliubiquitinación ramificada. Permitiendo de esta manera, la degradación por el proteasoma 26S. Por otra parte la proteína VP13D poliubiquitinada, cede su cadena de ubiquitinas al endosoma intermedio, el cual es fagocitado por una vacuola autofágica, con ENTD1 como intermediario. VP13D se reciclaría para volver a cumplir dicha función. Hay estudios que plantean que el aumento de la autofagia está asociada con la transición G1 / S retrasada del ciclo celular (40).

Discusión

En base al experimento realizado mediante el método de medicina de redes se descubrieron posibles agrupaciones biológicas con importancia funcional en la patogenia del MM, esto quiere decir la reconstrucción de redes de coexpresión y antiexpresión en base al cálculo de regresiones lineales simples es capaz de extraer información funcional.

Grupos funcionales en submódulos y su posible relevancia en el Mieloma Múltiple

A efectos de esta monografía se seleccionaron algunos de los hallazgos que se consideraron de mayor relevancia de acuerdo con los resultados del análisis de submódulos en DAVID, los que a continuación discutimos:

Una de las agrupaciones más relevantes es la referente a las cadherinas. 10 submódulos de un total de 58 presentaron enriquecimiento significativos en RNAs que codifican para cadherinas o protocadherinas (Figura 4). Las cadherinas forman una superfamilia con al menos seis subfamilias, que se pueden distinguir sobre la base de la composición del dominio de la proteína, la estructura genómica y el análisis filogenético de las secuencias de la proteína (41). Tienen un rol importante en eventos específicos de adhesión célula-célula, su expresión parece estar bien regulada durante el desarrollo y cada tipo de tejido o célula muestra un patrón característico de moléculas de cadherina particulares. En las neoplasias humanas se conoce como la regulación inadecuada de sus niveles conduce a una invasión de células cancerosas agravadas y a la metástasis. Al igual que las inmunoglobulinas y los receptores de células T, la familia del receptor neuronal relacionado con cadherina (CNR) cuenta con regiones constantes y variables que mediante splicing generan productos genéticos diversificados. Por lo que comprendiendo mejor estos cambios podrían extrapolarse en ambos sistemas para obtener nueva información acerca de cómo se regulan de forma diferenciada en las neoplasias. Por otra parte, las PCDH son un grupo de glicoproteínas de superficie de las neuronas que se expresan en regiones específicas del sistema nervioso central. Consisten en más de 50 genes, codificados por tres clusters de genes muy relacionados entre sí: alfa, beta y gama. A través de splicing alternativo se generan múltiples isoformas alfa y gama, lo que potencialmente genera una enorme variabilidad en moléculas de la superficie celular. Fueron estas formas alfa y gamma las que tuvieron más relevancia en nuestros resultados (42). La protocadherinas alfa y gamma, están ubicadas en el cromosoma 5 con una organización genómica inusual similar a la de los grupos de genes del receptor de células B y células T (42).

Existen estudios que revelan que las cadherinas juegan un rol en promoviendo la proliferación y supervivencia del mieloma de las células de la médula ósea (43). Otros autores plantean que la inhibición de las N-cadherinas aumentaría la proliferación de las células madre del cáncer (44). Es bien conocido que los niveles séricos de E-cadherina circulantes tienen implicancias pronósticas (45). Por lo tanto, se podría plantear que el aumento en la expresión de las cadherinas tendrían un rol en la proliferación celular en la MO, y su disminución podría estar implicada en la evasión de la MO y metástasis.

Como puede verse en la Figura 4, un 5.2% de los submódulos vinculó la patogenia molecular del MM con el sistema de proteínas policomb de regulación epigenética. Las proteínas Policomb constituyen otro de los mecanismos implicados en la regulación del ciclo celular es la acción de las proteínas Policomb en la regulación del ciclo celular. Dentro del complejo Polycomb2 se incluye al promotor de zeste 2 (EZH2), a su homólogo EZH1, al supresor de zeste 12 (SUZ12) y al embryonic ectoderm development (EED). De estos destacamos al EZH2 por su rol en el silenciamiento de genes mediante la metilación de la histona 3 (46). Mutaciones que causan la pérdida de función de EZH2 se encontraron relacionadas con las neoplasias mieloproliferativas, por lo que se plantea que este podría ser un target para futuros estudios.

Además, en un 3,4% de los resultados arrojados por DAVID se menciona la proteína HUWE1, por lo que se plantea que la búsqueda de un nuevo sustrato de la E3 ligasa, enzima que media la especificidad de las reacciones dirigiendo a las E2 a las proteínas blanco, podría proporcionar un método para regular distintos procesos celulares. Por ejemplo la proteína HUWE1, que interactúa con y regula la ubiquitinación y la estabilidad de la transcripción inducible por daños en el ADN 4 (DDIT4) y está implicada en el cáncer y las discapacidades intelectuales (47).

El HIF1A, regula la respuesta celular a la hipoxia y está implicado en la progresión del cáncer. Si bien se conoce su respuesta al oxígeno se sabe menos sobre el destino de su ARNm. La enzima desubiquitinante USP52 actúa como un regulador importante al disminuir la eliminación del HIF1A por el SUP. Los niveles de ARNm de HIF1A se reducen ante la disminución de USP52, llevando a una menor cantidad de dianas hipóxicas reguladas por HIF1A. Lo que disminuiría las vías que promueven la angiogénesis, necesarias para la progresión tumoral (48). De este modo se propone que al actuar selectivamente sobre la USP52 se podría reducir este mecanismo.

Dentro del sistema inmune se encontró la relación entre el receptor de quimioquinas CCR1, que se expresa de forma endógena en las membranas celulares de una amplia gama de tipos de células inmunológicas, y las beta arrestinas. Estas son proteínas citosólicas responsables de la supervivencia de los linfocitos T y además son un factor importante en ciertas enfermedades autoinmunes como es el caso de la esclerosis múltiple, ya que aumenta la susceptibilidad a la autoinmunidad (49). Actúan principalmente regulando a la baja la acción del grupo más grande de receptores transmembrana celulares al "desensibilizarlos" inhibiendo estéricamente la activación de la proteína G. Además, cumple funciones importantes en la endocitosis de receptores mediada por clatrina. Como se sabe que el ARNm de la quimioquina CCL3 se encuentra aumentada en la médula ósea en pacientes con enfermedad ósea por mieloma, se demostró que al inhibir el CCR1 se impide la formación de osteoclastos y por lo tanto se

disminuye la destrucción ósea (49). Este mecanismo aún no está disponible para su uso clínico, pero se cree que podría ser empleado como tratamiento de las complicaciones óseas del mieloma en un futuro cercano.

MERLIN es una proteína supresora de tumores; sus mutaciones se han identificado en múltiples tumores benignos y cánceres malignos. La pérdida de función de la proteína lleva a una sobreexpresión de KSR1. Este es necesario para la entrada al ciclo celular inducida por el factor de crecimiento derivado de plaquetas (PDGF) además está involucrada en muchos fenotipos patológicos causados por la pérdida de MERLIN, a saber, la morfología multipolar, la adhesión mejorada de la matriz celular, la adhesión focal y, lo más importante, el aumento de la proliferación y la supervivencia.

MERLIN desempeña un papel crucial en la regulación de varios componentes del microambiente de la MO, incluidas las células estromales, las células endoteliales y los osteoblastos, generando su proliferación (50). Teniendo en cuenta esta información se plantea la posibilidad de que alteraciones en esta proteína puedan dar un aumento en la supervivencia y proliferación de los precursores de células B así como un funcionamiento aberrante en los osteoclastos. (50).

Modelo Hipotético de Regulación y Coordinación de la Autofagia y del Pasaje G1/S.

Además de descubrir importantes asociaciones funcionales de grupos génicos con subdominios liderados por genes asociados al MM, pudimos mapear componentes del SUP en cada uno de los submódulos. Dado que la regulación de la expresión de CKS1B es desconocida y en el Mieloma Múltiple se ha observado un aumento de su número de copias (51), propusimos un modelo hipotético que logra vincular CKS1B con un grupo de componentes del SUP presentes en su módulo que son huérfanos de vía según reactoma. En dicho modelo hipotético proponemos que CKS1B que participa de la regulación del pasaje G1/S y participa también de la regulación de su propia degradación y de la degradación de organelos membranosos junto con los componentes huérfanos del SUP, llevando a una coordinación de la autofagia, la degradación proteasomal y el ciclo celular. Alrededor del 35-40% de los pacientes con MM presentan duplicaciones del brazo largo del cromosoma 1 que se relacionan con CKS1B. El mecanismo subyacente a la amplificación de 1q se cree que involucra inestabilidad de la región pericentromérica 1q12 y translocaciones de todo el brazo 1q, que podrían estar relacionadas con la hipometilación específica del pericentromérico 1q12.

Posteriormente se podrían plantear varios experimentos para poder verificar esta hipótesis, por ejemplo, se podría buscar con anticuerpos específicos para estas proteínas la colocalización en endosomas. Otros experimentos podrían estar basados en la expresión de dominantes negativos,

es decir expresar proteínas con algunas fallas que no puedan cumplir su función y provoquen una detención en la vía.

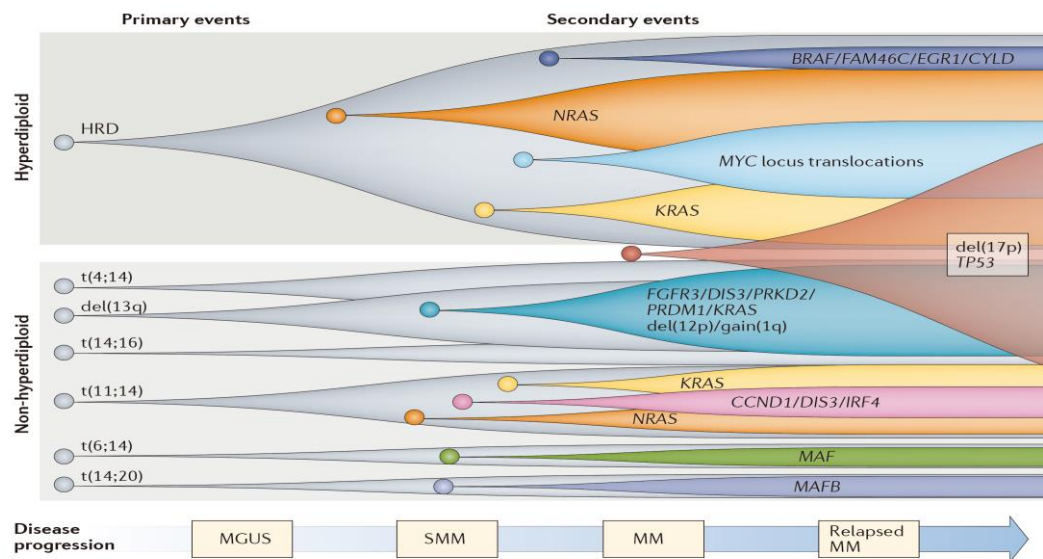


Figura 7. Este modelo representa las mutaciones coexistentes más frecuentes que impulsan la evolución clonal desde MGUS hasta MM. Figura obtenida de “Genomic complexity of multiple myeloma and its clinical implications”. Se observa en las mutaciones non-HRD la ganancia en el brazo largo del cromosoma 1 en color azul. (1).

Conclusiones

El análisis que asocia datos de expresión ricos en información, listas de genes con rol patogénico y listas de componentes del SUP de acuerdo con métodos de medicina de redes, arrojó nuevas informaciones sobre grupos funcionales y potenciales mecanismos que están actualmente siendo investigados por diversos laboratorios en el mundo entero. Este tipo de investigaciones permiten elaborar modelos a ser validados en forma experimental.

Referencias bibliográficas

1. Manier S, Salem KZ, Park J, Landau DA, Getz G, Ghobrial IM. Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol*. 2017;14(2):100–13.
2. Fondo Nacional de Recursos. Tratamiento Del Mieloma Multiple Con Bortezomib Y/O Lenalidomida. 2018;1–27. Available from: http://www.fnr.gub.uy/sites/default/files/normativas/medicamentos/n_trat_mmultiple_bortezomib_lenalidomida.pdf
3. Michels TC, Petersen KE. Multiple myeloma: Diagnosis and treatment. *Am Fam Physician* [Internet]. 2017;95(6):373–83. Available from: <http://dx.doi.org/10.1016/j.mayocp.2015.11.007>
4. Diaz L. Historias Clinicas Comentadas. In: oficina del libro FEFMUR, editor. 2da ed. Montevideo: oficina del libro FEFMUR; 2015. p. 227–50.
5. Alexander DD, Mink PJ, Adami HO, Cole P, Mandel JS, Oken MM, et al. Multiple myeloma: A review of the epidemiologic literature. *Int J Cancer*. 2007;120(SUPPL. 14):40–61.
6. Chiron D, Jegou G, Pellat-Deuceunynck C. Toll-like receptors: Expression and involvement in Multiple Myeloma. *Leuk Res* [Internet]. 2010;34(12):1545–50. Available from: <http://dx.doi.org/10.1016/j.leukres.2010.06.002>
7. Risk Factors for Multiple Myeloma [Internet]. [cited 2018 Oct 10]. Available from:

<https://www.cancer.org/cancer/multiple-myeloma/causes-risks-prevention/risk-factors.html>

8. Myeloma, Multiple disease: Malacards - Research Articles, Drugs, Genes, Clinical Trials [Internet]. [cited 2018 Aug 31]. Available from: [https://www.malacards.org/card/myeloma_multiple?search=multiple myeloma](https://www.malacards.org/card/myeloma_multiple?search=multiple%20myeloma)
9. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature* [Internet]. 2011 Mar 23;471:467. Available from: <http://dx.doi.org/10.1038/nature09837>
10. Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM, et al. APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nat Commun*. 2015;6.
11. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun*. 2014;5.
12. Lohr JG, Stojanov P, Carter SL, Cruz-gordillo P, Lawrence MS, Auclair D, et al. Widespread Genetic Heterogeneity in Multiple Myeloma : Implications for Targeted Therapy. 2014;25(1):91–101.
13. Technology A. Targeting the Ubiquitin Proteasome System Targeting the Ubiquitin Proteasome System. 2015;1–4.
14. Rozman,C Cardellach F. Farreras, Rozman, Medicina interna. In: ELSEVIER, editor. Farreras, Rozman, Medicina interna [Internet]. dieciocho. Barcelona; 2016. p. 1672–81. Available from: -
15. Fischer ES, Böhm K, Lydeard JR, Yang H, Stadler MB, Cavadini S, et al. Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature* [Internet]. 2014 Jul 16;512:49. Available from: <http://dx.doi.org/10.1038/nature13527>
16. Ge Z, Leighton JS, Wang Y, Peng X, Chen Z, Chen H, et al. HHS Public Access Cancer Types. 2018;23(1):213–26.
17. Zamudio-arroyo JM, Peñañ-Rangel MT, Riesgo-Escovar JR. La ubiquitinación: un sistema de regulación dinámico de los organismos. *Rev Espec en Ciencias Químico-Biológicas*. 2012;15(2):133–41.
18. Bedford L, Paine S, Sheppard PW, Mayer RJ, Roelofs J. Assembly, structure, and function of the 26S proteasome. *Trends Cell Biol* [Internet]. 2010 Jul 1 [cited 2018 Oct 13];20(7):391–401. Available from: <https://www.sciencedirect.com/science/article/pii/S0962892410000735>
19. Cvek B, Dvorak Z. The Ubiquitin-Proteasome System (UPS) and the Mechanism of Action of Bortezomib. *Curr Pharm Des* [Internet]. 2011;17(15):1483–99. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21504411>
20. Ramakrishnan V, Mager DE. Network-Based Analysis of Bortezomib Pharmacodynamic Heterogeneity in Multiple Myeloma Cells. *J Pharmacol Exp Ther* [Internet]. 2018;365(3):734–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29632237><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5959840><http://jpet.aspetjournals.org/lookup/doi/10.1124/jpet.118.247924>
21. Lecker SH. Protein Degradation by the Ubiquitin-Proteasome Pathway in Normal and Disease States. *J Am Soc Nephrol* [Internet]. 2006;17(7):1807–19. Available from: <http://www.jasn.org/cgi/doi/10.1681/ASN.2006010083>
22. Ubels J, Sonneveld P, van Beers EH, Broijl A, van Vliet MH, de Ridder J. Predicting treatment benefit in multiple myeloma through simulation of alternative treatment effects. *Nat Commun* [Internet]. 2018;9(1):1–10. Available from: <http://dx.doi.org/10.1038/s41467-018-05348-5>
23. Agnelli L, Forcato M, Ferrari F, Tuana G, Todoerti K, Walker BA, et al. The reconstruction of transcriptional networks reveals critical genes with implications for clinical outcome of multiple myeloma. *Clin Cancer Res*. 2011;17(23):7402–12.
24. Kolaczyk ED. Statistical Analysis of Network Data with R. 2009. 386 p.
25. Barabási AL, Gulbahce N, Loscalzo J. Network medicine: A network-based approach to human disease. *Nat Rev Genet* [Internet]. 2011;12(1):56–68. Available from: <http://dx.doi.org/10.1038/nrg2918>
26. JOSEPH LOSCALZO A-LB. Network Medicine. *Complex Systems in Human Disease and Therapeutics*. 2017.
27. Zanzoni A, Soler-López M, Aloy P. A network medicine approach to human disease. *FEBS Lett*. 2009;583(11):1759–65.
28. Hoang PH, Dobbins SE, Cornish AJ, Chubb D, Law PJ, Kaiser M, et al. Whole-genome sequencing of multiple myeloma reveals oncogenic pathways are targeted somatically through multiple mechanisms. *Leukemia* [Internet]. 2018;(March):1–12. Available from: <http://dx.doi.org/10.1038/s41375-018-0103-3>
29. GeneCards - Human Genes | Gene Database | Gene Search [Internet]. [cited 2018 Sep 8]. Available from: <https://www.genecards.org/>
30. Lambiotte R, Delvenne J-C, Barahona M. Laplacian Dynamics and Multiscale Modular Structure in Networks. 2008;1–29. Available from: <http://arxiv.org/abs/0812.1770><http://dx.doi.org/10.1109/TNSE.2015.2391998>
31. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech Theory Exp*. 2008;2008(10).

32. bioDBnet - Biological Database Network [Internet]. [cited 2018 Oct 5]. Available from: <https://biodbnet-abcc.ncifcrf.gov/db/db2db.php>
33. Botti. Manuscrito en preparación.
34. Compare two Excel spreadsheets without programming macro nor installing any software on your computer [Internet]. [cited 2018 Oct 5]. Available from: <https://www.xlcomparator.net/>
35. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin A V., et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415–21.
36. Mayer CT, Gazumyan A, Kara EE, Gitlin AD, Golijanin J, Viant C, et al. The microanatomic segregation of selection by apoptosis in the germinal center. *Science* (80-). 2017;358(6360):1–14.
37. Biphenyls CP. HHS Public Access. 2015;91(2):165–71.
38. Han SH, Kim SH, Kim HJ, Lee Y, Choi SY, Park G, et al. Cobll1 is linked to drug resistance and blastic transformation in chronic myeloid leukemia. *Leukemia* [Internet]. 2017;31(7):1532–9. Available from: <http://dx.doi.org/10.1038/leu.2017.107>
39. Kramps T, Peter O, Brunner E, Nellen D, Froesch B, Chatterjee S, et al. Wnt/Wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear β -catenin-TCF complex. *Cell*. 2002;109(1):47–60.
40. Gagliostro V, Casas J, Caretti A, Abad JL, Tagliavacca L, Ghidoni R, et al. Dihydroceramide delays cell cycle G1/S transition via activation of ER stress and induction of autophagy. *Int J Biochem Cell Biol* [Internet]. 2012;44(12):2135–43. Available from: <http://dx.doi.org/10.1016/j.biocel.2012.08.025>
41. Nollet F, Kools P, Van Roy F. Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J Mol Biol*. 2000;299(3):551–72.
42. Yagi T, Takeichi M. Cadherin superfamily genes: Functions, genomic organization, and neurologic diversity. *Genes Dev*. 2000;14(10):1169–80.
43. Hurt EM, Wiestner A, Rosenwald A, Shaffer AL, Campo E, Grogan T, et al. Overexpression of c-maf is a frequent oncogenic event in multiple myeloma that promotes proliferation and pathological interactions with bone marrow stroma. *Cancer Cell*. 2004;5(2):191–9.
44. Sadler NM, Harris BR, Metzger BA, Kirshner J. N-cadherin impedes proliferation of the multiple myeloma cancer stem cells. *Am J Blood Res* [Internet]. 2013;3(4):271–85. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3875273&tool=pmcentrez&rendertype=abstract>
45. Syrigos KN, Harrington KJ, Karayiannakis AJ, Baibas N, Katirtzoglou N, Roussou P. Circulating soluble E-cadherin levels are of prognostic significance in patients with multiple myeloma. *Anticancer Res*. 2004;24(3 B):2027–31.
46. Veneti Z, Gkouskou KK, Eliopoulos AG. Polycomb repressor complex 2 in genomic instability and cancer. *Int J Mol Sci*. 2017;18(8):1–16.
47. Thompson JW, Nagel J, Hoving S, Gerrits B, Bauer A, Thomas JR, et al. Quantitative Lys- ϵ -Gly-Gly (diGly) proteomics coupled with inducible RNAi reveals ubiquitin-mediated proteolysis of DNA damage-inducible transcript 4 (DDIT4) by the E3 Ligase HUWE1. *J Biol Chem*. 2014;289(42):28942–55.
48. Bett JS, Ibrahim AFM, Garg AK, Kelly V, Pedrioli P, Rocha S, et al. The P-body component USP52/PAN2 is a novel regulator of *HIF1A* mRNA stability. *Biochem J* [Internet]. 2013;451(2):185–94. Available from: <http://biochemj.org/lookup/doi/10.1042/BJ20130026>
49. Xiao K, McClatchy DB, Shukla AK, Zhao Y, Chen M, Shenoy SK, et al. Functional specialization of beta-arrestin interactions revealed by proteomic analysis. *Proc Natl Acad Sci* [Internet]. 2007;104(29):12011–6. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.0704849104>
50. Zhou L, Ammoun S, Müller J, Lasonder E, Sharma V, Ercolano E, et al. The scaffold protein KSR1 , a novel therapeutic target for the treatment of Merlin-deficient tumors. 2016;35(March 2015):3443–53.
51. Zhan F, Colla S, Wu X, Chen B, Stewart JP, Kuehl WM, et al. CKS1B, overexpressed in aggressive disease, regulates multiple myeloma growth and survival through SKP2- and p27Kip1-dependent and -independent mechanisms. *Blood* [Internet]. 2007 Jun 1 [cited 2018 Oct 15];109(11):4995–5001. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17303695>

Agradecimientos

Departamento de Biofísica por brindarnos un espacio físico para poder trabajar, en particular al Profesor Nelson Bracesco y al Laboratorio Biofísica.

Anexos

ANEXO 1. Tabla de genes asociada a la etiopatogenia del MM

Gen	Nombre	Función	Presente en Malacards	Identificador público de probesets	Mutación en MM
AICDA	Deaminasa de citidina inducida por activación	La Desaminasa Inducida por Activación inicia las reacciones de Hipermutación Somática (SHM) y cambio de Isotipo (CSR), responsables de la diversificación secundaria de anticuerpos. Además, AID promueve la generación de translocaciones cromosómicas con potencial linfomagenico. Estudios enfatizan que la regulación de los niveles de AID es crítica para determinar su función, tanto en la respuesta inmune como en la generación de lesiones oncogénicas. Los resultados obtenidos indican que la actividad de AID es contexto celular dependiente, lo que presumiblemente está asociado a proliferación celular.	NO	NM_020661, BC006296 (no se incluyeron en el trabajo)	Hipermutación somática
ANP32E	Fosfoproteína nuclear ácida 32, miembro de la familia E	Histona chaperona que interviene específicamente en la eliminación de la histona H2A.Z / H2AFZ en todo el genoma del nucleosoma: Inhibe la actividad de la proteína fosfatasa 2A (PP2A). No inhibe la proteína fosfatasa 1. Puede desempeñar un papel en el desarrollo cerebeloso y la sinaptogénesis.	NO	NM_030920 (M30), AW612574 (M8)	Amplificación del brazo largo del cromosoma 1
APOBEC3A	Subunidad 3A catalítica de enzima de edición de MRN de apolipoproteína B	La proteína codificada por este gen carece de la actividad de unión al zinc de otros miembros de la familia. La proteína desempeña un papel en la inmunidad, al restringir la transmisión de ADN extraño, como los virus.	NO	NM_004900 (M8), U03891	Translocación 14;16
APOBEC3B	Subunidad 3B catalítica de enzima de edición de MRN de apolipoproteína B	Se cree que las proteínas pueden ser enzimas de edición de ARN y tienen funciones en el crecimiento o control del ciclo celular.	NO	NM_004900 (M8), NM_152426 (M1), U03891	Translocación 14;16
ATP13A2	ATPasa 13A2 transportadora de cationes	Gen que codifica un miembro de la subfamilia de las ATPasas que transporta cationes inorgánicos y otros sustratos desempeñando un papel en la homeostasis de los cationes intracelulares y el mantenimiento de la integridad neuronal. También es requerido para un adecuado mantenimiento lisosomal y mitocondrial.	NO	NM_022089 (M29)	Variaciones en nucleótido simple del CRE
BCL9	B Cell CLL/Lymphoma 9	Gen cuya función es desconocida. Puede ser un objetivo de translocación en tumores malignos de células B con anomalías de 1q21.	NO	NM_004326 (M1), AL353962 (M5), BG481877 (M3)	Amplificación del brazo largo del cromosoma 1
BMP6	Bone Morphogenetic Protein 6	La proteína codificada por este gen regula una amplia gama de procesos biológicos, incluida la homeostasis del hierro, el desarrollo de la grasa y los huesos y la ovulación. La expresión diferencial de este gen puede estar asociada con la progresión del cáncer de mama y de próstata. Las mutaciones en este gen pueden estar asociadas con la sobrecarga de hierro en pacientes humanos	NO	AI968068 (M17), NM_001718 (M1)	Translocaciones que afectan al Myc
CALCB	Calcitonin Related Polypeptide Beta	Gen que codifica una proteína que dilata una variedad de vasos incluyendo la vasculatura coronaria, cerebral y sistémica. Su abundancia en el SNC también apunta hacia un papel de neurotransmisor o neuromodulador. Péptido relacionado con la calcitonina, beta, 3.8 kDa, expresado en partes del sistema nervioso y endocrino, vasodilatador potente e hipotensor.	NO	AA747379 (M0)	Variación en nucleótido simple del CRE
CCND1	Cyclin D1	La proteína codificada por este gen pertenece a la familia de ciclinas altamente conservadas. Esta ciclina forma un complejo que funciona como una subunidad reguladora de CDK4 o CDK6, cuya actividad se requiere para la transición G1 / S del ciclo celular. Se ha demostrado que esta proteína interactúa con la proteína supresora de tumores Rb y la expresión de este gen está regulada positivamente por Rb.	SI	BC000076 (M9), M73554 (M14)	Translocaciones 11;14
CCND3	Cyclin D3	La proteína codificada por este gen pertenece a la familia de ciclinas altamente conservada. Componente regulador del complejo ciclina D3-CDK4 (DC) que fosforila e inhibe a los miembros de la familia de proteínas retinoblastoma (RB) que incluye RB1 y regula el ciclo celular durante la transición G (1) / S. Estas ciclinas funcionan como reguladores de las quinasas CDK.	SI	AL833425 (M3), NM_001760 (M1)	Translocaciones 6;14
CKS1B	CDC28 Protein Kinase Regulatory Subunit 1B	Codifica la proteína KS1B que se une a la subunidad catalítica de las quinasas dependientes de ciclina y es esencial para su función biológica.	NO	NM_001826 (M7)	Amplificación del brazo largo del cromosoma 1
COBL1	Cordon-Bleu WH2 Repeat Protein Like 1	Gen de codificación de proteínas. Las enfermedades asociadas con COBL1 incluyen epilepsia, lóbulo temporal familiar y sordera no sindrómica ligada al X. Las anotaciones de ontología de genes (GO) relacionadas con este gen incluyen la unión a actina.	NO	AA150107 (M1), AK002054(M3), BC006264(M2), BF002844(M1), NM_014900(M28)	Variación en nucleótido simple del CRE

EGR1	Early Growth Response 1	La proteína codificada por este gen pertenece a la familia EGR de proteínas con zinc de tipo C2H2. Es una proteína nuclear y funciona como un regulador transcripcional. Los productos de los genes diana que activa son necesarios para la diferenciación y la mitogénesis.	NO	AI459194, AV733950(M19), NM_001964(M19)	Mutación hiperploide desconocida
FAM46C (TENT5C)	Terminal Nucleotidyltransferase 5C	TENT5C (Terminal Nucleotidyltransferase 5C) es un gen de codificación de proteínas. Las enfermedades asociadas con TENT5C incluyen mieloma, múltiple. Un parálogo importante de este gen es TENT5A. Probable nucleotidiltransferasa que puede actuar como una poli (A) ARN polimerasa no canónica. (Infección microbiana) Parece mejorar la replicación de algunos virus, incluido el virus de la fiebre amarilla, en respuesta al interferón tipo I	NO	AL046017(M4), NM_017709(M2)	translocaciones que afectan MYC
FGFR3	Fibroblast Growth Factor Receptor 3	Gen que codifica un miembro de la familia del receptor del factor de crecimiento de fibroblastos (FGFR) y desempeña un papel esencial en la regulación de la proliferación celular, la diferenciación y la apoptosis. Regula la osteogénesis como la mineralización ósea postnatal por osteoblastos. Promueve la apoptosis en los condrocitos, pero también puede promover la proliferación de células cancerosas.	SI	NM_000142(M17), M58051(M13)	Translocación 4;14
FOXO3	Forkhead Box O3	Este gen pertenece a la familia de factores de transcripción forkhead que se caracterizan por un dominio de forkhead distinto. Este gen probablemente funciona como un desencadenante de la apoptosis a través de la expresión de los genes necesarios para la muerte celular. La translocación de este gen con el gen MLL se asocia con leucemia aguda secundaria.	NO	AF041336, NM_001455, BE888885(M17), AV725666(M28)	translocaciones que afectan MYC
HOXB3	Homeobox B3	Gen miembro de la familia de homeobox Antp y codifica una proteína nuclear con un dominio de unión al ADN de homeobox. Se incluye en un grupo de genes homeobox B ubicados en el cromosoma 17. La proteína codificada funciona como un factor de transcripción específico de secuencia que está involucrado en el desarrollo. El aumento de la expresión de este gen se asocia con un subconjunto biológico distinto de leucemia mieloide aguda.	NO	NM_002146(M28), AW510657(M7)	Variación en nucleótido simple del CRE
HRAS	HRas Proto-Oncogene, GTPase	Este gen pertenece a la familia del oncogén Ras, cuyos miembros están relacionados con los genes transformadores de los retrovirus del sarcoma de mamíferos. Los productos codificados por estos genes funcionan en las vías de transducción de señales. Los defectos en este gen están implicados en una variedad de cánceres, que incluyen el cáncer de vejiga, el cáncer folicular de tiroides y el carcinoma oral de células escamosas.	SI	NM_005343 (no se incluyo en el trabajo)	Mutación desconocida
IGH	Immunoglobulin Heavy Locus	Gen que codifica inmunoglobulinas que reconocen antígenos extraños e inician respuestas inmunes como la fagocitosis y el sistema del complemento.	NO	AF343666(M17), AW519168, S55735, X95746	translocacion 6;14
IGK	Immunoglobulin Kappa Locus	Es un gen de codificación de proteínas. Entre sus vías relacionadas están la respuesta inmune, la ruta Fc epsilon RI y el desarrollo de la señalización del receptor IGF-1.	NO	U96291(28), AW575927	translocaciones que afectan MYC
IGL	Immunoglobulin Lambda Locus	Gen que codifica inmunoglobulinas que reconocen antígenos extraños e inician respuestas inmunes como la fagocitosis y el sistema del complemento.	NO	Sin resultados	translocaciones que afectan MYC
KDM3B	Lysine Demethylase 3B	Es un gen de codificación de proteínas. Entre sus vías relacionadas están la organización de la cromatina y el transporte de electrones respiratorios, la síntesis de ATP por acoplamiento quimiosmótico y la producción de calor por desacoplamiento de proteínas. Las anotaciones de Ontología Genética (GO) relacionadas con este gen incluyen la actividad de la dioxigenasa.	NO	BC001202 (M29), NM_016604(M19)	Variación en el número de copias del CRE
MAF	MAF BZIP Transcription Factor	La proteína codificada por este gen es un factor de transcripción que contiene cremallera de leucina y se une al ADN que actúa como homodímero o como heterodímero. Esta proteína desempeña un papel en la regulación de varios procesos celulares, incluido el desarrollo de células de fibra de lente embrionaria, la mayor susceptibilidad de las células T a la apoptosis y la diferenciación terminal de condrocitos.	SI	AA442149(M8), AF055376(M17), BE674528, BF508646(M13), NM_005360	Translocación 6;14
MAFB	MAF BZIP Transcription Factor B	La proteína codificada por este gen es un factor de transcripción de cremallera de leucina básica (bZIP) que desempeña un papel importante en la regulación de la hematopoyesis específica del linaje. La proteína nuclear codificada reprime la transcripción mediada por ETS1 de genes específicos de eritroides en células mieloides. Este gen no contiene intrones.	SI	AW135013, NM_005461	translocacion 14;20
MMSET (NSD2)	Nuclear Receptor Binding SET Domain Protein 2	Desempeña una función restaurativa dentro del genoma, reclutamiento de proteínas como proteína p53-enlace 1 para reparar roturas que se producen en el ADN y para mantener la estabilidad genética. Pero cuando el mal funcionamiento del MMSET, la vía de protección se queda corta, y realizará una cascada de mutaciones que pueden conducir a procesos de enfermedad.	SI	BE793789 (M10), BF111870 (M12), AF071594 (M6), AI770166 (M2), AW024870 (M6)	translocacion 4;14

MYC	MYC Proto-Oncogene	Este gen es un protooncogén y codifica una fosfoproteína nuclear que desempeña un papel en la progresión del ciclo celular, la apoptosis y la transformación celular. La amplificación de este gen se observa con frecuencia en numerosos cánceres humanos. Las translocaciones que afectan a este gen están asociadas con el linfoma de Burkitt y el mieloma múltiple en pacientes humanos	SI	NM_002467 (M17), BF514781 (M10)	Variación en el número de copias del CRE
NBPF1	Neuroblastoma Breakpoint Family, Member 1	Este gen es un miembro de la familia del punto de quiebre del neuroblastoma (NBPF), que consiste en docenas de genes recientemente duplicados localizados principalmente en duplicaciones segmentarias en el cromosoma humano 1. La expresión alterada de algunos miembros de la familia de genes se asocia con varios tipos de cáncer.	NO	AI634549, AW269499 (M17)	Variación en nucleótido simple del promotor
PACS2	Phosphofurin Acidic Cluster Sorting Protein 2	PACS2 (proteína 2 de clasificación de grupos ácidos de fosfofurina) es un gen de codificación de proteínas. Un parálogo importante de este gen es PACS1.	NO	AL583340 (M17), AK098354, AB011174 (M29)	Variación en el número de copias del CRE
PAX5	Paired Box 5	La característica central de esta familia de genes es un nuevo motivo de unión al ADN altamente conservado, conocido como la caja pareada. Los factores de transcripción de la caja pareada son reguladores importantes en el desarrollo temprano, y se cree que las alteraciones en la expresión de sus genes contribuyen a la transformación neoplásica. Este gen codifica la proteína activadora específica del linaje de las células B que se expresa en las etapas iniciales de la diferenciación de las células B.	NO	NM_016734 (M28)	Variación en nucleótido simple del CRE
PLD4	Phospholipase D Family Member 4	Entre sus vías relacionadas están la biosíntesis de glicerosfolípidos y el metabolismo. Las anotaciones de Ontología de genes (GO) relacionadas con este gen incluyen actividad de fosfolipasa C de fosfatidilinositol y actividad de fosfolipasa D específica de N-acilfosfatidiletanolamina. Un parálogo importante de este gen es PLD3	NO	Se excluyó de este estudio BE676703	Variación en el número de copias del CRE
PDZK1	PDZ Domain Containing 1	Este gen codifica una proteína de andamiaje que contiene el dominio PDZ. Las moléculas que contienen el dominio PDZ se unen y median en la localización subcelular de las proteínas diana. La proteína codificada media en la localización de las proteínas de la superficie celular y desempeña un papel crítico en el metabolismo del colesterol al regular el receptor de HDL, el receptor de tipo B de clase 1.	NO	NM_002614 (M17)	Amplificación del brazo largo del cromosoma 1
PRKD2	Protein Kinase D2	La proteína codificada por este gen pertenece a la familia de la proteína quinasa D (PKD) de serina / treonina proteína quinasas. Esta quinasa puede ser activada por los ésteres de forbol, así como por la gastrina a través del receptor de colecistoquinina B (CCKBR) en células de cáncer gástrico. Puede unirse al diacilglicerol (DAG) en la red trans-Golgi (TGN) y puede regular la salida de la proteína de la membrana basolateral de la TGN.	NO	AF309082 (M29), AL050147 (M13)	Translocación 4;14
RAB36	RAB36, Member RAS Oncogene Family	Miembro de la familia de oncogenes RAS. Entre sus vías relacionadas están el metabolismo de las proteínas y el transporte mediado por vesículas. Las anotaciones de Ontología de genes (GO) relacionadas con este gen incluyen la unión a GTP y la actividad de GTPasa.	NO	AF133588	Variación en el número de copias del CRE
SP110	SP110 Nuclear Body Protein	Complejomultiproteico que puede tener un papel en la regulación de la transcripción de genes. La proteína codificada puede funcionar como un activador de la transcripción de genes y servir como un coactivador del receptor de hormonas nucleares. Además, se ha sugerido que la proteína puede desempeñar un papel en la biogénesis de los ribosomas y en la inducción de la diferenciación de las células mieloides.	NO	AA969194 (M14), AF280094 (M14), AF280095 (M28), NM_004509 (M14), NM_004510 (M5)	Variación en el número de copias del CRE
ST6GAL1	ST6 Beta-Galactoside Alpha-2,6-Sialyltransferase 1	Este gen codifica un miembro de la familia de glicosiltransferasa 29. La proteína codificada es una proteína de membrana de tipo II que cataliza la transferencia de ácido siálico de ácido siálico CMP a sustratos que contienen galactosa. La proteína, que normalmente se encuentra en el Golgi pero puede ser procesada proteolíticamente a una forma soluble, participa en la generación de los determinantes de carbohidratos de la superficie celular y los antígenos de diferenciación HB-6, CD75 y CD76.	NO	AV695711, AI743792 (M17)	Variación en nucleótido simple del CRE
TEX22	Testis Expressed 22	TEX22 (TestisExpressed 22) es un gen de codificación de proteínas. Información genética adicional para el gen TEX22.	NO	BC009212 (M9)	Variación en el número de copias del CRE
TRAF2	TNF Receptor Associated Factor 2	La proteína codificada por este gen es un miembro de la familia de proteínas del factor asociado al receptor de TNF (TRAF). Las proteínas TRAF se asocian y median en la transducción de señales de los miembros de la superfamilia de receptores de TNF. Esta proteína es necesaria para la activación mediada por TNF-alfa de MAPK8 / JNK y NF-kappaB. El complejo de proteínas formado por esta proteína y TRAF1 interactúa con las proteínas inhibitoras de la apoptosis (IAP) y funciona como un mediador de las señales antiapoptóticas de los receptores de TNF. La interacción de esta proteína con TRADD, un transductor de señal apoptótico asociado con el receptor de TNF, asegura el reclutamiento de IAP para la inhibición directa de la activación de la caspasa. BIRC2 / c-IAP1, un inhibidor de la apoptosis que posee actividad de ubiquitina ligasa, puede desacelerar e inducir la degradación de esta proteína, y así potenciar la apoptosis inducida por TNF.	NO	NM_021138 (M2)	Mutaciones en la codificación del CRE
TWEAK	TNF Superfamily Member 12	Es una citoquina de la superfamilia de las TNF que activa el receptor Fn14. Regula la proliferación celular, la muerte celular, la angiogénesis y la inflamación. Las citoquinas proinflamatorias incrementan la expresión del receptor Fn14 en las células tubulares y TWEAK induce apoptosis en un medio inflamatorio. Además TWEAK induce la expresión de quimiocinas y citoquinas proinflamatorias.	NO	Sin resultados	Mutaciones en la codificación del CRE

ANEXO 2. Comandos para programa lm en lenguaje R

```
#1-----

incognita<-'NM_005163'
rnames<-rownames(newdata)
tuya<-which(rnames==incognita)
linreg<-NULL

for (x in 1:length(rnames)){
  eleme<-lm(as.numeric(newdata[tuya,])~as.numeric(newdata[x,]))
  sumy<-summary(eleme)
  linreg<-rbind(linreg,c(sumy$adj.r.squared,sumy$coefficients[2,1-4]))
}

linreg<-as.data.frame(linreg)
rownames(linreg)<-rnames

write.table(linreg,sep='\t',file='Akt1-NM_005163.csv')

#2 -----
```



```

#Creamos el objeto denominado rs con el comando:
rs<-read.csv('TABLAunica.txt', sep='\\t', header=T)
#Cambiamos su formato al tipo "data frame":
rs <- as.data.frame(rs)
#Para ver el peso del objeto en la memoria RAM:
object.size(rs)
#Para ver el número de filas y columnas en ese orden:
dim(rs)
#Para que imprima en pantalla la información contenida en el objeto, primero las filas y luego las columnas:
rs[1:4, 1:65]
#Creamos el objeto rsS y se utilizó el stacking para hacer una aposición de las columnas
rsS <- stack(rs, select=c(1:65))
dim(rsS)
rsS[1:4, 1:2]
#Del objeto rsS creamos un objeto que se llama "c" que se encuentra en la columna 1 que son los pesos
c <- rsS[c(1)]
los probesets:
b <- rsS[c(2)]
Los probesets:
dim(c)
dim(b)
#Usamos excel para multiplicar por 64 veces los datos de "a"
dim(a)
#Creamos un objeto con el mismo nombre que el objeto inicial
rs<-read.csv('Probesets.txt', sep='\\t', header=F)
#Para decirle a R que los valores dentro de la tabla no son factores:
rs <- data.frame(lapply(rs, as.character), stringsAsFactors=FALSE)
rsS <- stack(rs, select=c(1:64))
a <-rsS[c(1)]
#Para confeccionar la tabla:
tabla<-cbind(a, b, c)
dim(tabla)
#Para ver como quedo confeccionada la tabla:
tabla[1:4, 1:3]
#Para guardar la tabla en el directorio:
write.table(tabla, sep='\\t', file='mielomaRED-Silvana.txt')

library(igraph)

r<-read.csv('mielomaRED-Silvana.txt', sep='\\t', header=F)
dim(r)
r[1:4, 1:4]
r1 <- r[1:3285633, 2:4]
r1[1:4, 1:3]
r1[3285629:3285633, 1:3]
r1m <- as.matrix(r1)
red<-graph_from_data_frame(r1m, directed=F)

r<-read.csv('mielomaRED-Silvana.txt', sep='\\t', header=T)
r[1:4, 1:3]
r1 <- r[1:1848168, 1:3]
r1m <- as.matrix(r1)
red<-graph_from_data_frame(r1, directed=F)
E(red)$weight=as.numeric(r1[, 3])
red.simp<-simplify(red, remove.loops = TRUE)
red.simp.75 <- delete.edges(red.simp, which(abs(E(red.simp)$weight) < 0.75))
red.75.gexf<-igraph.to.gexf(red.simp.75)
print(red.75.gexf, file="red.75.gexf", replace=T)
re.simp.pos075 <- delete.edges(red.simp, which(E(red.simp)$weight < 0.75))

```