

Reanálisis: Direct inhibition of the NOTCH transcription factor complex

Guillermo Díez Pinel

Mario Cazalla García

INDICE

Resumen

Introducción:

- Rutas de señalización y procesos biológicos en cáncer
- NOTCH1 y T-ALL
- Estrategia del reanálisis
- ¿Qué se podría haber hecho mejor en el estudio original?

Materiales y métodos

- Obtención de los datos
- Clustering
- Análisis PCA
- Análisis de expresión diferencial
- GSEA
- KEGG
- Gene Ontology: Biological Process (GO:BP)

Resultados

Discusión

Bibliografía

Resumen

En este trabajo se ha realizado un reanálisis de los datos de expresión obtenidos por medio de microarray de un experimento en el que se trataron dos líneas celulares de leucemia linfoblástica aguda de células T (T-ALL) con un inhibidor del complejo del que forma parte el dominio intracelular de NOTCH, llamado SAHM1. NOTCH es una diana terapéutica interesante en cáncer porque su dominio intracelular, cuando es escindido viaja al núcleo, donde actúa como una de las subunidades de un heterotrímero que regulan la expresión de rutas importantes en cáncer. Con el objetivo de mejorar el análisis y obtener más información acerca de los efectos y mecanismos de actuación de SAHM1 en las células T-ALL, se han realizado los siguientes análisis: clustering por K-Means, un análisis PCA, un análisis de expresión diferencial, dos análisis de sobrerrepresentación utilizando Gene Ontology (GO) y las rutas de señalización KEGG respectivamente, así como un análisis Gene Set Enrichment Analysis (GSEA). Además de confirmar una inhibición de los targets de NOTCH como sucedía en el experimento original, se ha encontrado una serie de mecanismos nuevos que podrían explicar la acción antitumoral de SAHM1: inhibición a múltiples niveles del metabolismo proteico, de azúcares y de ácidos nucleicos, del transporte intracelular, de funciones mitocondriales, del mantenimiento telomérico, señalización de citoquinas y activación de checkpoints del ciclo celular.

Introducción

Rutas de señalización y procesos biológicos en cáncer

El cáncer es una enfermedad en la que las células adquieren la habilidad de dividirse y crecer descontroladamente por la alteración de la expresión o de la función de genes específicos, entre otros mecanismos. Los avances en la secuenciación del DNA y el RNA han permitido el estudio de estas alteraciones genéticas, y ahora tenemos una mejor comprensión de los procesos y rutas de señalización implicados en esta enfermedad.

A lo largo de los años, ha ido aumentando las anotaciones disponibles para los procesos y rutas celulares tanto en células sanas como en cáncer, suponiendo un enorme potencial para los análisis bioinformáticos, ya que pueden ayudar a encontrar nuevos mecanismos relevantes que puedan sugerir nuevas opciones de tratamiento. Las diez rutas de señalización que se encuentran alteradas con mayor frecuencia en cáncer se muestran a continuación en la figura 1¹. El análisis de estas rutas ha permitido detectar genes clave implicados en cáncer.

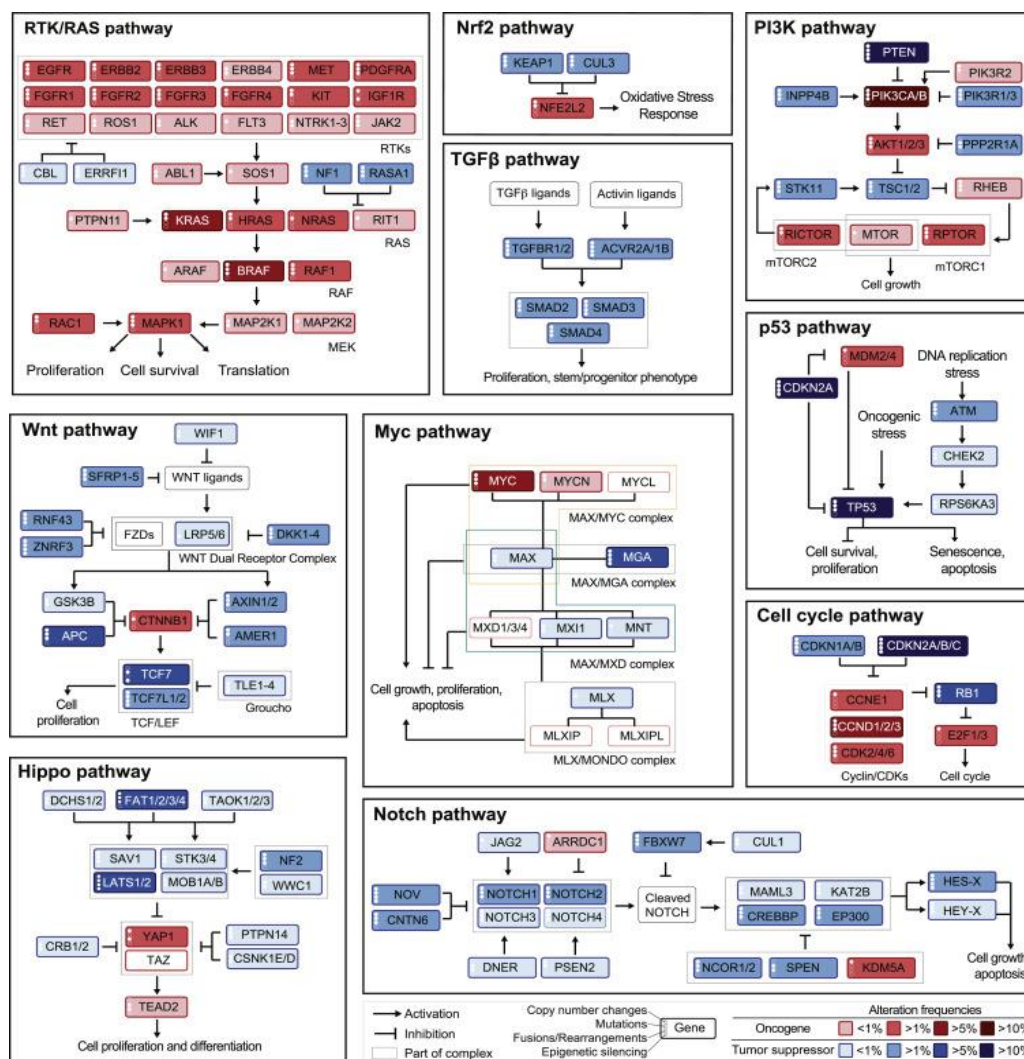


Figura 1. Representación de las 10 rutas más frecuentemente alteradas en cáncer. Los oncogenes aparecen en color rojo, mientras que los genes supresores de tumores en color azul. La intensidad del color indica la frecuencia de alteración, de acorde a la escala de la esquina inferior derecha. Las flechas indican activación mientras las flechas sin punta inhibición.

NOTCH1 y T-ALL

En mamíferos, la familia NOTCH consta cuatro receptores (1-4) transmembrana, a los cuales se pueden unir ligandos de las familias Delta, Serrate o Lag-2 de forma yuxtacrina. Las proteínas NOTCH regulan rutas de señalización implicadas en organogénesis, diferenciación celular, crecimiento celular, proliferación y muerte celular. NOTCH1 actúa como receptor transmembrana y como factor de transcripción al mismo tiempo. La interacción entre NOTCH y su ligando conlleva la liberación del dominio intracelular (ICN1) al ser procesado por metaloproteasas de la familia ADAM y un complejo Gamma-Secretasa. Tras ello se transloca al núcleo, donde se une al factor de transcripción CSL. Sin embargo, el complejo no es activo hasta que se une a él un tercer miembro: una proteína co-activadora de la familia *mastermind-like* (MAML). Este complejo ternario es capaz de reclutar a la maquinaria de transcripción, promoviendo la expresión de numerosos genes diana².

La leucemia linfoblástica aguda de células T (T-ALL) es un tipo de cáncer muy agresivo que afecta con mayor frecuencia a niños y adolescentes. Se suele asociar a traslocaciones cromosómicas y otras alteraciones genéticas o epigenéticas. La alteración de la actividad de NOTCH1 es un evento central en la progresión de T-ALL^{3,4}. De hecho, en más de la mitad de los casos se encuentran mutaciones que afectan a su gen. Se han identificado traslocaciones entre el locus de *NOTCH1* en cromosoma 9 y la subunidad beta del receptor T de las células T en el cromosoma 7, resultando en una expresión truncada de NOTCH1. Aun cuando no existen mutaciones en *NOTCH1*, su actividad se suele ver alterada por otros mecanismos. Los efectos oncogénicos de NOTCH1 se deben parcialmente a un incremento en la expresión y activación de c-Myc, un factor de transcripción que juega un papel central en prácticamente la totalidad del proceso oncogénico⁵. Además, las células pre-T inmaduras son particularmente susceptibles a transformarse debido a una excesiva señalización de NOTCH, explicando la alta frecuencia de T-ALL derivada de mutaciones en NOTCH⁶.

Por todo ello, NOTCH se ha convertido en el objetivo de diversos tratamientos. Sin embargo, esta ruta tiene un papel importante en el desarrollo normal en las células madre, por lo que evitar efectos secundarios es un gran reto⁴. En la actualidad, existen diferentes tratamientos para T-ALL con NOTCH1 mutado. Uno de ellos son los inhibidores del complejo Gamma-Secretasa (GSI), que participa en la liberación del dominio intracelular de NOTCH.

Palomero et al. caracterizaron en 2006 los efectos de GSI en varias líneas celulares de T-ALL. Los análisis de expresión génica tras el tratamiento con GSI sugieren que el fármaco causa la inhibición de 201 genes, y la sobreexpresión de 38. Además, identificaron las dianas directas del complejo CSL-ICN1-MAML mediante un análisis de ChIP-on-chip. Este análisis identificó uniones altamente significativas a las regiones promotoras de 134 genes. Al comparar los resultados del ChIP-on-chip con los genes significativamente alterados por el tratamiento con GSI, se vio un elevado solapamiento entre ambos resultados, demostrando la relación entre la inhibición de la señalización de NOTCH1 y los cambios transcripcionales inducidos por este tratamiento. Se añadió también que las anotaciones funcionales de este top de genes regulados por GSI revelaron una sobrerrepresentación de genes involucrados en el metabolismo de grandes moléculas y biosíntesis de proteínas, junto con la acción importante de NOTCH1 en el control del crecimiento celular.⁵

Sin embargo, problemas habituales de este tratamiento son toxicidad gastrointestinal, probablemente por la inhibición crónica de NOTCH1 y NOTCH2, así como la aparición de variantes de NOTCH1 resistentes al tratamiento. Por ello, se deberían estudiar alternativas a

estos fármacos y la posibilidad de hacer terapias combinadas ^{2,7}. Es por ello que Raymond E. Moellering et al propusieron en 2006 el trabajo que hemos reanalizando otro tratamiento alternativo, empleando SAHM1 (Stapled Alpha-Helical Peptide 1). Se trata de un péptido modificado que se une competitivamente al complejo ICN1-CSL desplazando a MAML, bloqueando la actividad transcripcional de NOTCH1. De esta forma, se consiguen los efectos deseados al utilizar GSI, pero sin causar los efectos secundarios originados por la inhibición del complejo Gamma-Secretasa. Cuando hicieron un análisis de expresión diferencial, observaron una reducción de la expresión de los genes *HES1*, *MYC* y *DTX1*, todos dianas de NOTCH1. Cuando hicieron un análisis en GSEA utilizando como gene set los genes más inhibidos por GSI reportados por Raymond E. Moellering et al, observaron una inhibición significativa. Los datos de expresión se obtuvieron analizando muestras por triplicado de las líneas celulares de T-ALL HPB-ALL y KOPT-K1. El grupo experimental fue tratado con SAHM1, y controles recibieron DMSO (vehículo). El RNA total se extrajo y se hibridó al microarray *Affymetrix human U133 plus 2.0* (Affymetrix). Además de los resultados transcriptómicos, reportaron un elevado potencial de SAHM1 para inhibir la proliferación de células en cultivo y en un modelo murino de T-ALL inducido por NOTCH1 ².

Estrategia del reanálisis

-Para determinar si los ficheros CEL obtenidos de GEO se corresponden con las descripciones fenotípicas de la base de datos, hicimos un análisis de clustering y otro de PCA. De esta forma, se pueden identificar errores de etiquetado de las muestras o outliers biológicos.

-Para hallar los genes diferencialmente expresados inducidos por el tratamiento, hicimos un análisis de expresión diferencial, pero con los paquetes '*Affy*' y '*Limma*' de R en lugar de *GenePattern software* (Broad Institute).

-Con el objetivo de caracterizar los procesos celulares y rutas de señalización alteradas significativamente por el tratamiento con SAHM1, se hicieron análisis de sobrerrepresentación utilizando Gene Ontology (GO) y KEGG pathways.

-Para identificar gene sets cuya distribución es significativamente diferente tras la aplicación del tratamiento, se hizo un análisis en GSEA. Esta aproximación puede revelar procesos activados o inhibidos que no se aprecian si sólo miramos a los genes con cambios de expresión más drásticos. Además, tiene en cuenta al conjunto de genes que colaboran en un proceso, no sólo a genes individuales que pueden no ser representativos de lo que sucede en la célula.

¿Qué se podría haber hecho mejor en el estudio original?

-Hemos echado en falta otro grupo experimental tratado con GSI. De esta forma, podrían haber hecho una comparación mucho más directa y robusta entre ambos tratamientos. Incluso llegan a utilizar los genes inhibidos por GSI para construir un gene set de GSEA, pero no incluyen este tratamiento.

-Al construir el gene set de GSEA a partir de los efectos de GSI, sólo tomaron los genes inhibidos y no los sobreexpresados. De esta forma se crea un sesgo y perdemos información. Hay que tener en cuenta también a los genes sobreexpresados, ya que podrían suponer mecanismos adicionales anti-tumorales, o tener efectos pro-tumorales.

-Al hacer el análisis en GSEA, sólo se centran en el análisis de gene sets de factores de transcripción. Consideramos que es una pena no aprovechar estos datos de expresión para estudiar procesos y rutas celulares. Es por ello por lo que decidimos hacer un análisis en GSEA

utilizando los Hallmark gene sets y por lo que también hicimos dos análisis de sobrerrepresentación, en procesos biológicos y en rutas de señalización.

-Un problema que tuvimos a la hora de reproducir el trabajo fue que no aportaban la lista de genes que conformaban el gene set que crearon para llevar a cabo el GSEA, por lo que tuvimos que copiarlos a mano de la figura del artículo que citaban.

-Toda la atención se centra en las alteraciones que afectan a los genes codificantes de proteínas, pero sería interesante prestar atención también a las regiones no codificantes. Se han identificado mutaciones en las regiones no codificantes del genoma que resultan en la expresión de factores de transcripción aberrantes y se han descubierto varios miRNAs y lncRNAs con una participación patogénica en T-ALL.⁸

Materiales y métodos

Todos los análisis que usan R fueron realizados en la versión de R 4.0.4 para Linux.

Obtención de los datos

Los datos de expresión crudos en formato CEL fueron obtenidos de GEO (<https://www.ncbi.nlm.nih.gov/geo/>), de la entrada GSE18198 ([GEO Accession viewer \(nih.gov\)](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE18198)).

Clustering

Se utilizó la herramienta de clustering no supervisado de Babelomics (<http://babelomics.bioinfo.cipf.es/>) utilizando el algoritmo K-Means. Se hizo tanto con 2 como con 4 grupos utilizando como entrada la matriz de expresión normalizada.

Análisis PCA

Se utilizó el paquete de R *PCAtools* para obtener la contribución a la varianza de cada una de las componentes principales y para representar las muestras en dos dimensiones, siendo los ejes X e Y la primera y la segunda componente principal con mayor contribución a la variabilidad respectivamente.

Análisis de expresión diferencial

Se utilizaron los paquetes *'Affy'* y *'Limma'* para la normalización y procesamiento de los datos crudos. La normalización se realizó por *robust multiarray averaging* (RMA). Para analizar todas las muestras simultáneamente se incluyó la línea celular como covariable. Para analizar las líneas celulares individualmente la normalización y análisis se realizaron partiendo de los datos crudos de la línea en cuestión solamente. Para la anotación se utilizó el paquete *'hgu133plus2.db'*.

GSEA

Los datos de expresión utilizados provienen de la matriz de expresión normalizada por RMA. Para el GSEA preranked se analizaron todas las muestras y se utilizó un ranking en el que la métrica era el $-\log_{10}(\text{adj.P.Val})$ multiplicado por 1 o -1 en función del signo del logFC, con el objetivo de obtener valores mayores para los resultados más significativos, pero conservando la dirección de la expresión (sobreexpresión o inhibición). Para el análisis de líneas celulares individuales se hizo un GSEA estándar, con las matrices normalizadas por RMA solamente de las muestras de las líneas celulares a analizar en cada caso. Los archivos en formato gct y cls fueron generados con el paquete de R *'ArrayTools'*. Se utilizaron los parámetros predeterminados excepto para *permutation_type*, en el que se seleccionó la opción *'gene set'* por contar cada clase con un número reducido de muestras. Como gene sets se utilizaron los *Hallmark gene sets (H)* y los *Regulatory target gene sets – TFT: transcription factor targets (C3)*. Al fichero gmt de estos últimos se le añadieron 3 gene sets personalizados: uno con los targets del complejo CSL-NICD-MAM, llamado *'NOTCH_TARGETS'*, que contiene los genes a cuyos promotores se ha visto que el complejo se une⁵; otro con los genes en los que se observaron mayores cambios en diferentes líneas celulares con el tratamiento con GSI; otro con un subset del anterior que sólo contiene los genes inhibidos por el tratamiento con GSI.

KEGG

Para identificar las rutas KEGG significativas se utilizó la función *enrichKEGG()* del paquete de R *'clusterProfiler'*. Como entrada para identificar las rutas significativas se utilizaron los datos de expresión diferencial significativos obtenidos de analizar todas las muestras de forma

conjunta. Para visualizar las rutas se utilizó el paquete '*pathview*'. Las casillas de los genes de las rutas generadas por *pathview* se encuentran divididas por la mitad. la mitad izquierda indica el logFC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no. Por tanto, los datos de entrada fueron los datos de expresión diferencial de todos los genes para la mitad izquierda, no sólo los diferencialmente expresados. Las pathways se obtuvieron a partir de los resultados significativos de *enrichKEGG()*, rutas importantes en cáncer y reguladas por NOTCH1, y rutas que se decidieron explorar tras examinar los resultados de Gene Ontology.

Gene Ontology: Biological Process (GO:BP)

Para identificar los términos GO:BP significativos se utilizó la función *enrichGO()* del paquete de R '*clusterProfiler*'. Se utilizó como entrada tanto los genes significativos de cada línea como los del análisis conjunto, separando los genes sobreexpresados de los inhibidos en cada caso. Para facilitar la observación rápida de los datos se utilizó el paquete '*rrvgo*', que hace un clustering de términos GO en función de su similaridad semántica. El nombre cada clúster se hereda del término GO:BP con más significancia estadística en cada uno. No obstante, como comete algunos errores a la hora de agrupar los términos, estos fueron analizados posteriormente sin agrupar en clústeres.

Resultados

El clustering separa adecuadamente por línea celular, pero tiene dificultades para discernir entre tratamiento y control.

Para comprobar si una herramienta de clustering agrupaba las muestras como los fenotipos indicados en GEO, se utilizó un algoritmo de clustering por K-Means. Al indicar al algoritmo que el número de clases es 4, no logra agrupar los triplicados adecuadamente, sino que crea dos clústeres de 5 muestras, cada clúster constituido por muestras de la misma línea celular, y luego genera dos clústeres que sólo tienen una integrante con las muestras restantes. Cuando se indica que haga 2 clases, genera 2 clústeres de 6 muestras por línea celular (Figura 2).

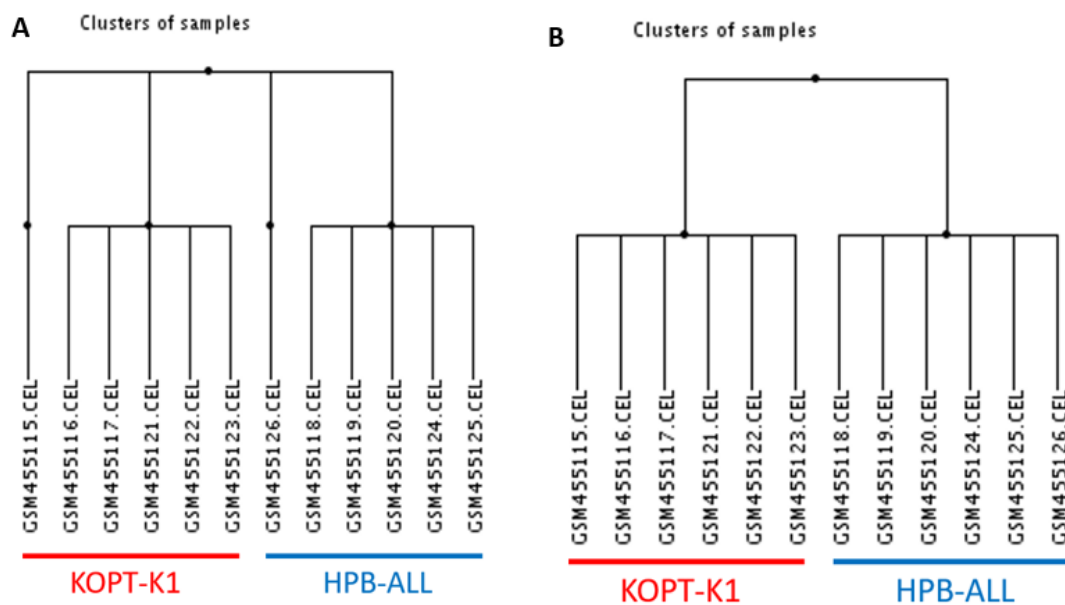


Figura 2. Clustering por K-Means. A) Clustering con número de clases N=4, indicándose las líneas celulares en la parte inferior. B) Clustering con número de clases N=2.

El análisis PCA agrupa como se esperaban las muestras

Para visualizar la similaridad entre muestras dado que el clustering por K-Means no lograba diferenciar adecuadamente entre tratamientos, se decidió hacer un análisis PCA. En este caso, la componente principal 1 (PC1) era responsable de una gran cantidad de variabilidad (79.69%), mientras que la PC2, la segunda más grande, solamente de un 5.99%. En este caso la diferenciación entre líneas celulares vuelve a ser muy marcada, pero entre tratamientos es menor. Todas las muestras pertenecientes a la misma clase y línea celular se agrupan de forma cercana espacialmente, pero la variabilidad en PC2 es mayor para las muestras de células HPB-ALL tratadas con DMSO (Figura 3).

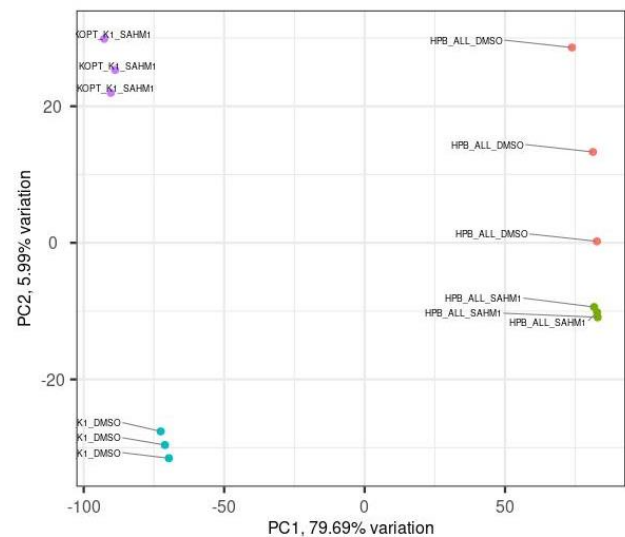


Figura 3. Representación de las muestras en función de sus valores para PC1 y PC2. Todos los triplicados se agrupan adecuadamente, aunque la variabilidad en PC2 es mayor para las muestras HPB-ALL tratadas con DMSO.

Los resultados de expresión diferencial coinciden con los obtenidos en el artículo original

En el análisis de todas las muestras se obtuvieron 1707 genes diferencialmente expresados significativos (1246 inhibidos y 461 sobreexpresados), en el de KOPT-K1, 7765 genes (5006 inhibidos y 2759 sobreexpresados) y en el de HPB-ALL, 2105 genes (682 inhibidos y 1423 sobreexpresados). En el artículo original se obtuvieron como genes significativamente inhibidos a HES1, MYC y DTX1. En los genes diferencialmente expresados en el análisis con ambas líneas estos genes se encuentran en las posiciones 6, 11 y 8 respectivamente. En los resultados de KOPT-K1 sin embargo se encuentran en las posiciones 368, 420 y 120, y en los de HPB-ALL en las posiciones 47, 796 y 13.

	rank	ENTREZ	ENSEMBL	SYMBOL	logFC	AveExpr	t	P.Value	adj.P.Val
201464_x_at	1	3725	ENSG00000177606	JUN	1.5101731	6.934468	18.330863	6.327588e-11	1.729773e-06
227347_x_at	2	57801	ENSG00000188290	HES4	-1.2391846	6.970592	-15.864287	4.100405e-10	5.604639e-06
209933_s_at	3	11314	ENSG00000167851	CD300A	-0.9386960	8.547968	-15.204221	7.071378e-10	6.443675e-06
217078_s_at	4	11314	ENSG00000167851	CD300A	-1.3453311	6.496952	-14.021837	1.983563e-09	1.175724e-05
201170_s_at	5	8553	ENSG00000134107	BHLHE40	-1.0190562	6.638360	-13.932781	2.150426e-09	1.175724e-05
203394_s_at	6	3280	ENSG00000114315	HES1	-0.9697872	7.221859	-12.657732	7.202255e-09	3.281467e-05
205544_s_at	7	1380	ENSG00000117322	CR2	-0.9438572	8.521063	-11.992489	1.411378e-08	5.511835e-05
227336_at	8	1840	ENSG00000135144	DTX1	-1.0787341	7.460750	-11.051011	3.865952e-08	1.321044e-04
202887_s_at	9	54541	ENSG00000168209	DDIT4	-0.9695786	9.270627	-10.923057	4.457347e-08	1.353894e-04
203867_s_at	10	54475	ENSG00000073536	NLE1	-0.7963720	6.952487	-10.723445	5.580632e-08	1.525632e-04
202431_s_at	11	4609	ENSG00000136997	MYC	-0.6696863	11.440298	-10.333882	8.739730e-08	2.026169e-04
226452_at	12	5163	ENSG00000152256	PDK1	-0.8325181	7.702594	-10.280149	9.307287e-08	2.026169e-04

Tabla 1. Genes más significativos del análisis de expresión diferencial del análisis con todas las muestras. En rojo se marcan los genes que son diana de NOTCH1 descritos en el artículo original.

El análisis GSEA usando Hallmark gene sets muestra procesos implicados en cáncer

Cuando se utilizó GSEA preranked para analizar los Hallmark gene sets utilizando como entrada los datos de expresión de todas las muestras, se vio que como resultado del tratamiento 19 gene sets inhibidos significativos. Los dos más significativos se corresponden con las dianas de MYC y, en cuarto lugar, aparece la señalización de MTORC1, ambas rutas frecuentemente sobreactivadas en cáncer. También encontramos gene sets relacionados con el metabolismo, hipoxia, reparación del DNA, especies reactivas de oxígeno y la señalización de NOTCH (Tabla X.). En los gene sets con expresión aumentada por el tratamiento encontramos varios relacionados con la progresión del ciclo celular. Esta vez solamente hay 6 gene sets significativos, como se aprecia en la Tabla X. Los resultados para las líneas celulares individuales son similares, cambiando ligeramente algunos gene sets y el Enrichment Score.

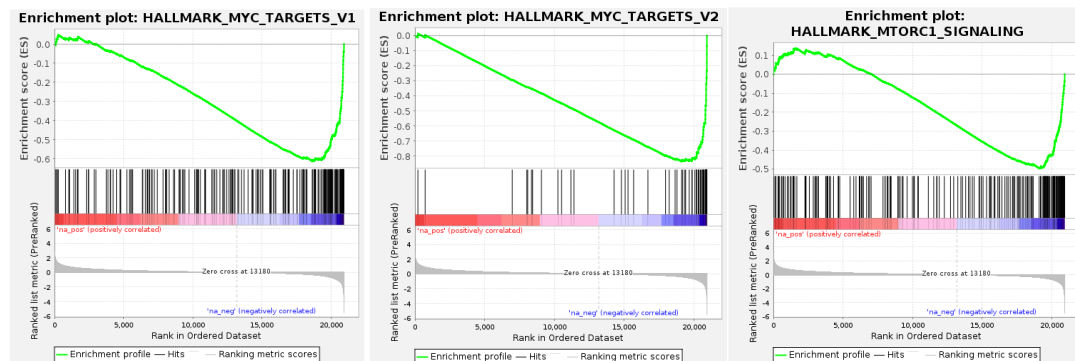


Figura 4. Regulación negativa de los gene sets de MYC y MTOR. La línea verde muestra los valores del Enrichment Score a lo largo del gene set. Las líneas negras verticales representan la distribución de los genes de dicho gene set ordenados en función del ranking utilizado. La zona azul se corresponde con los genes enriquecidos en las células tratadas con SAHM1 mientras que la roja se corresponde con los controles.

NAME	SIZE	ES	NES	NOM p-val	FDR q-val
HALLMARK_MYC_TARGETS_V2	58	-0.83419	-2.75896	0	0
HALLMARK_MYC_TARGETS_V1	200	-0.61197	-2.37367	0	0
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	113	-0.57814	-2.12658	0	0
HALLMARK_MTORC1_SIGNALING	200	-0.49616	-1.96364	0	0
HALLMARK_HYPOXIA	200	-0.47457	-1.84597	0	9.76E-04
HALLMARK_INTERFERON_ALPHA_RESPONSE	97	-0.50919	-1.7983	0	0.001202
HALLMARK_GLYCOLYSIS	200	-0.45255	-1.7497	0	0.001253
HALLMARK_DNA_REPAIR	150	-0.4403	-1.66852	0	0.003185
HALLMARK_IL2_STATS_SIGNALING	199	-0.40673	-1.60262	0	0.006385
HALLMARK_INTERFERON_GAMMA_RESPONSE	200	-0.40738	-1.58421	0	0.007612
HALLMARK_ALLOGRAFT_REJECTION	200	-0.38935	-1.54452	0	0.010479
HALLMARK_OXIDATIVE_PHOSPHORYLATION	199	-0.37944	-1.49502	0	0.01765
HALLMARK_FATTY_ACID_METABOLISM	156	-0.37695	-1.44518	0.004566	0.028585
HALLMARK_NOTCH_SIGNALING	32	-0.48959	-1.42526	0.063444	0.033345
HALLMARK_UV_RESPONSE_UP	158	-0.36186	-1.36194	0.004587	0.05256
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	49	-0.42231	-1.32888	0.076923	0.063588
HALLMARK_ADIPOGENESIS	200	-0.31117	-1.22155	0.059701	0.142686
HALLMARK_INFLAMMATORY_RESPONSE	200	-0.30679	-1.2086	0.046632	0.150871
HALLMARK_PEROXISOME	104	-0.31607	-1.13821	0.158103	0.246974

Tabla 2. Gene sets con expresión disminuida tras el tratamiento con SAHM1. Se indica cada gene set, el número de genes que lo componen, el Enrichment Score y su versión normalizada, así como la significancia estadística en p-valor y como q-valor de la corrección por test múltiple FDR.

NAME	SIZE	ES	NES	NOM p-val	FDR q-val
HALLMARK_G2M_CHECKPOINT	200	0.506214	1.744104	0	0.002434
HALLMARK_MITOTIC_SPINDLE	199	0.463222	1.621985	0	0.010965
HALLMARK_APOPTOSIS	161	0.440067	1.480332	0.003906	0.052322
HALLMARK_HEME_METABOLISM	200	0.410591	1.41594	0.003755	0.085379
HALLMARK_SPERMATOGENESIS	135	0.395202	1.318779	0.048942	0.193885
HALLMARK_E2F_TARGETS	199	0.375294	1.310253	0.029963	0.176496

Tabla 3. Gene sets con expresión incrementada tras el tratamiento con SAHM1. Se indica cada gene set, el número de genes que lo componen, el Enrichment Score y su versión normalizada, así como la significancia estadística en p-valor y como q-valor de la corrección por test múltiple FDR.

El análisis GSEA de los factores de transcripción muestra los gene sets personalizados de NOTCH como los más significativos

Cuando se utilizó GSEA preranked para analizar las dianas de factores de transcripción utilizando como entrada todas las muestras, se obtuvieron 45 gene sets inhibidos significativos, y 208 sobreexpresados. Los gene sets con las dianas de NOTCH1 y con los genes regulados por el tratamiento con GSI aparecen como los más significativos. Cuando se usan los datos de expresión de HPB-ALL se obtiene este mismo resultado. Sin embargo, cuando se usan los datos de KOPT-K1, el gene set '*NOTCH_TARGETS*' pasa a estar en la sexta posición, aunque también significativo.

GS	SIZE	ES	NES	NOM p-val	FDR q-val
GSI-NOTCH_DOWN_ONLY	41	-0.84	-2.56	0.000	0.000
NOTCH_TARGETS	54	-0.72	-2.33	0.000	0.000
GSI-NOTCH	48	-0.73	-2.29	0.000	0.000
FOXR2_TARGET_GENES	196	-0.53	-2.07	0.000	0.000
RUVBL2_TARGET_GENES	35	-0.59	-1.75	0.006	0.028
EPC1_TARGET_GENES	209	-0.44	-1.73	0.000	0.030
MYC_Q2	185	-0.44	-1.71	0.000	0.035
GKCGCNNNNNNNTGAYG_UNKNOWN	60	-0.50	-1.65	0.013	0.053
PPARGC1A_TARGET_GENES	92	-0.47	-1.64	0.000	0.055
ZNF23_TARGET_GENES	15	-0.65	-1.56	0.028	0.106
ZNF239_TARGET_GENES	37	-0.51	-1.53	0.015	0.134
KTGGYRSGAA_UNKNOWN	75	-0.44	-1.51	0.011	0.139
NCOA2_TARGET_GENES	433	-0.36	-1.50	0.000	0.152
USF2_Q6	252	-0.37	-1.49	0.000	0.149
NMYC_01	276	-0.36	-1.47	0.000	0.157
SQSTM1_TARGET_GENES	39	-0.49	-1.47	0.034	0.148
PAX5_02	17	-0.59	-1.47	0.080	0.149
PSMB5_TARGET_GENES	262	-0.37	-1.46	0.000	0.144
MYCMAX_01	255	-0.36	-1.46	0.000	0.143

Tabla 4. Gene sets con expresión disminuida tras el tratamiento con DMSO. Se indica cada gene set el número que lo componen, el Enrichment Score y su versión normalizada, así como la significancia estadística en p-valor y como q-valor de la correlación por test múltiple FDR

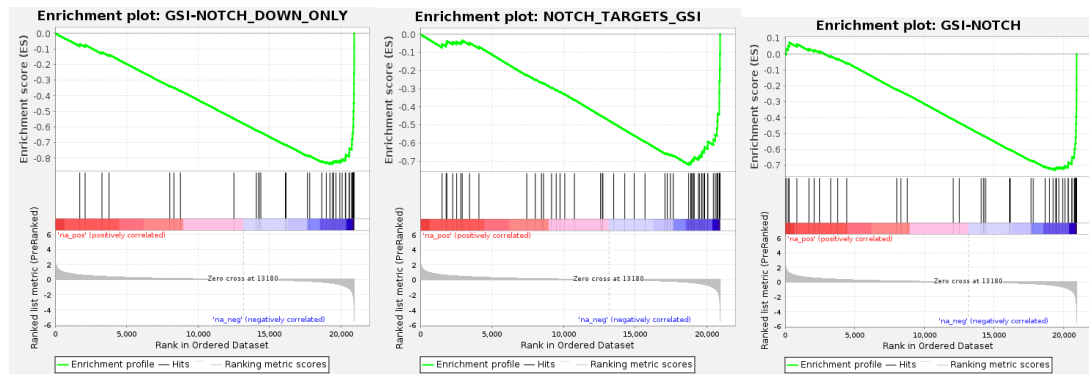


Figura 5. Regulación negativa de los gene sets de dianas de NOTCH y genes regulados por GSI. La línea verde muestra los valores del Enrichment Score a lo largo del gene set. Las líneas negras verticales representan la distribución de los genes de dicho gene set ordenados en función del ranking utilizado. La zona azul se corresponde con los genes enriquecidos en las células tratadas con SAHM1 mientras que la roja se corresponde con los controles.

Los análisis de sobrerrepresentación de Gene Ontology muestran que SAHM1 induce cambios en procesos relacionados con el metabolismo, la síntesis proteica y el ciclo celular.

Con el fin de determinar que procesos celulares eran activados o inhibidos por el tratamiento, se comprobaron los términos GO:BP enriquecidos al utilizar como entrada los genes significativos del análisis de expresión diferencial. Al utilizar como entrada los genes inhibidos significativos del análisis que utiliza la línea celular como covariable, se obtuvieron 376 resultados significativos (Figura 6).



Figura 6. Términos GO:BP relacionados enriquecidos a partir de los genes inhibidos en las muestras tratadas con SAHM1. Los términos se encuentran agrupados semánticamente, cada clúster separado por líneas más gruesas de los demás. El nombre cada clúster se hereda del término GO:BP con más significancia estadística en cada uno.

Los resultados sugieren que la síntesis proteica está afectada a múltiples niveles. En primer lugar, la síntesis de los propios aminoácidos se encontraría inhibida, así como la aminoacilación de los tRNA necesaria para la incorporación de los mismos a los péptidos nacientes en la traducción. La síntesis de los rRNA, de las subunidades de los ribosomas, su ensamblaje y varios procesos implicados en la traducción también se ven afectadas.

Por otro lado, la biosíntesis de purinas y pirimidinas también muestra numerosos procesos inhibidos. El metabolismo del ácido fólico y la síntesis de tetrahidrofolato, con un importante papel en la síntesis de aminoácidos y ácidos nucleicos también se han observado como inhibidos.

Diversos procesos relacionados con funciones mitocondriales, como la localización de proteínas a las mismas, la expresión de genes mitocondriales, la autofagia mitocondrial, la regulación de la permeabilidad de la membrana mitocondrial o el metabolismo de NAD⁺ y NADH, centrales para el metabolismo celular, incluyendo la fosforilación oxidativa. El metabolismo de la glucosa y otras hexosas, así como de las pentosas también parece estar afectado a múltiples niveles, desde su entrada a la célula hasta su síntesis y degradación.

Múltiples biomoléculas como proteínas, esteroides, lípidos y RNAs ven su transporte intracelular afectado. Entre ellas, la telomerasa y su RNA muestran una localización disminuida a los cuerpos de Cajal. El mantenimiento de los telómeros por parte de la telomerasa también es significativo. Los telómeros no son los únicos que sufren una síntesis reducida, sino que la replicación en sí parece estar afectada: la iniciación de la replicación, la apertura de la doble hebra de DNA y elongación del mismo son significativos.

Por último, en cuanto a citoquinas, hay una inhibición en la respuesta y señalización a IL-4, IL-7, IL-9 e IL-12.

En el caso de los procesos significativos a introducir los genes sobreexpresados en las células tratadas con SAHM1 (Figura 7), se obtuvieron 113 resultados significativos, menos de un tercio que al usar genes inhibidos. Destaca el abundante número de términos relacionados con la mitosis y el ciclo celular. Muchos de estos términos están relacionados con los checkpoints G1/S, G2/M y del huso mitótico. En cuanto a la mitosis, se encontraron resultados contradictorios para varios de sus pasos, encontrándose sobreexpresados tanto la regulación positiva como la negativa en varias etapas y procesos.

Se vieron también términos relacionados con la regulación positiva de la transcripción y la replicación, pero a la vez otros que indicaban una regulación negativa de la transcripción a partir de promotores por parte de la RNA polimerasa II en respuesta a estrés del retículo endoplásmico. Otros términos indican un incremento en la fosforilación de histonas.

Cuando se observan los resultados de los mismos análisis, pero en cada línea celular por separado, se aprecian diferencias respecto al análisis conjunto (material suplementario). En las células HPB-ALL (311 términos inhibidos, 216 sobreexpresados), aparecen inhibidos términos relacionados con respuestas inmunes y términos explícitamente relacionados con la proliferación de linfocitos. En cuanto a los términos obtenidos a partir de genes sobreexpresados, las células HPB-ALL presentan una gran cantidad de ellos relacionados con la regulación de la estructura de la cromatina. En células KOPT-K1 no se encontró ningún término significativo a partir de los genes sobreexpresados, pero 446 en los inhibidos. Estos últimos abarcan procesos muy variados, compartiendo los presentes en su análisis equivalente con ambas líneas, pero, incluyendo otros diferentes que forman clústeres pequeños, como regulación de la migración de células T o muerte celular inducida por estrés oxidativo (material suplementario).



Figura 7. Términos GO:BP relacionados enriquecidos a partir de los genes sobreexpresados en las muestras tratadas con SAHM1. Los términos se encuentran agrupados semánticamente, cada clúster separado por líneas más gruesas de los demás. El nombre de cada clúster se hereda del término GO:BP con más significancia estadística en cada uno.

El análisis KEGG muestra una inhibición en rutas relacionadas con la síntesis proteica, de bases nitrogenadas, de rutas reguladas por NOTCH y de rutas implicadas en cáncer.

Para visualizar los cambios en procesos celulares en forma de rutas de señalización, en primer lugar, se obtuvieron rutas significativas utilizando como entrada los genes diferencialmente expresados significativos. Tras ello, se utilizaron todos los datos de expresión diferencial (significativos y no significativos) para visualizar dichas rutas, pero indicando en cada gen si se trata de uno significativo o no. Se obtuvieron 24 rutas significativas en total.

En concordancia con los resultados de GO, la síntesis de aminoacil-tRNAs de muchos aminoácidos se encontraba inhibida. Teniendo en cuenta los resultados significativos solamente, los aminoácidos cuya activación se ve afectada son: glutamato, aspartato, serina, selenocisteína, valina, leucina, lisina, prolina, fenilalanina y triptófano (Figura 8).

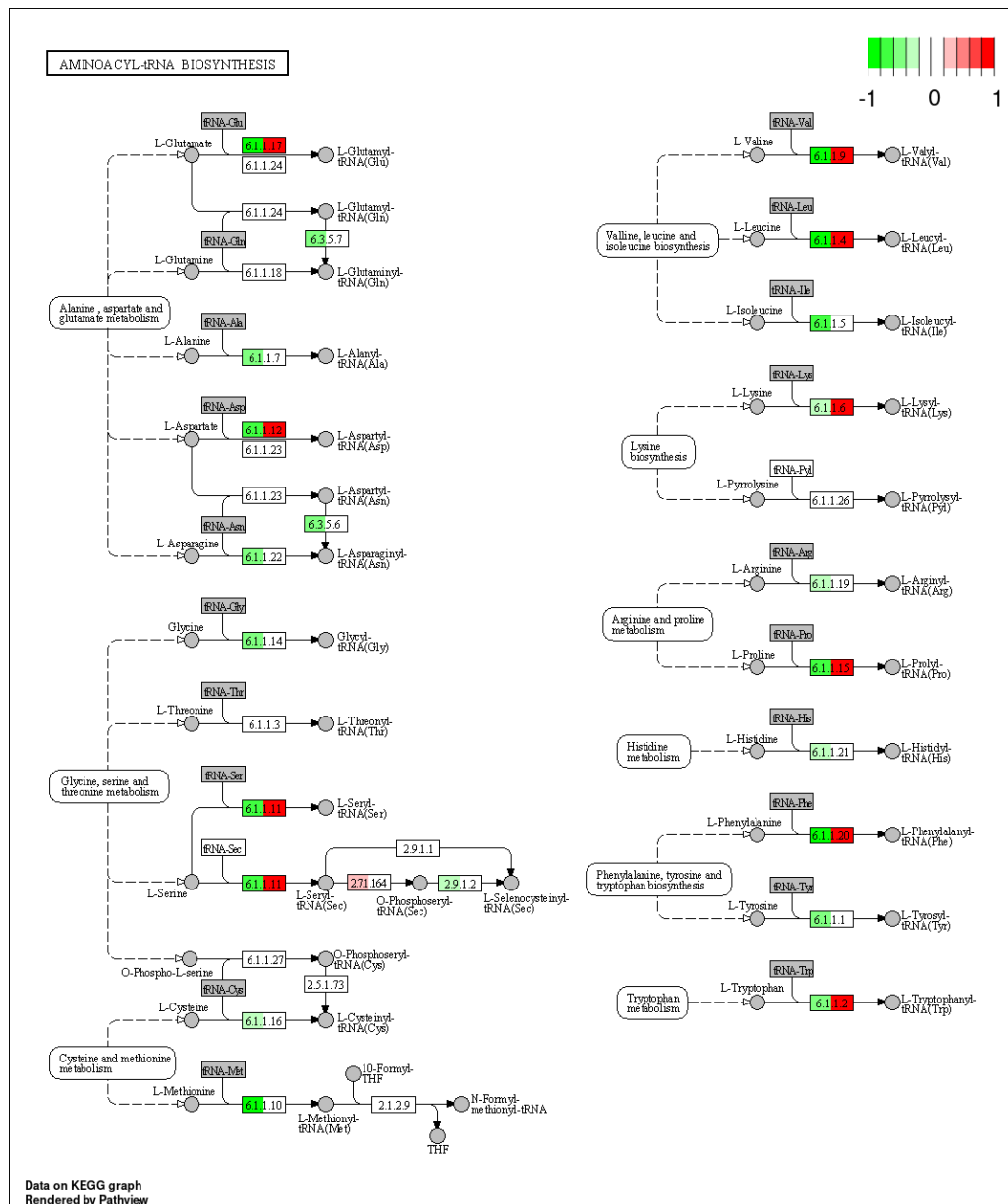


Figura 8. Inhibición de la activación de diversos aminoácidos. En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

Por otra parte, otro proceso en el que predomina la inhibición de genes es la autofagia, así como el procesamiento proteico en el retículo endoplasmático y la degradación de proteínas mediada por ubiquitina. Continuando los resultados observados en GO que indicaban una inhibición del transporte intracelular, se puede observar una inhibición a todos los niveles en la endocitosis. En cuanto al metabolismo energético, la mayoría de los genes significativos de la ruta de señalización de AMPK se encuentran inhibidos. Por último, eran significativas diversas son rutas relacionadas con infecciones patológicas: 'Nod-like receptor signaling pathway', 'Bacterial invasión of epithelial cells', 'Shigellosis', 'Yersinia infection', 'Hepatitis C', 'Hepatitis B', 'Human papillomavirus infection', 'Human T-Cell leukemia virus 1 infection', 'Herpes simplex virus 1 infection', 'Epstein-Barr virus infection' y 'Viral carcinogenesis'. Por último, encontramos también inhibición en rutas relacionadas con otros tipos de cáncer: 'Renal cell carcinoma' y 'Chronic myeloid leukemia'. Estos resultados se encuentran en el material suplementario.

Aparte de las rutas significativas, se decidió mirar otras rutas que se sabe están reguladas por NOTCH, así como rutas implicadas en el metabolismo, ya que estas últimas eran altamente relevantes de acuerdo a los resultados de GO.

En la ruta de las MAPK, la inhibición predomina, pero AP1 y JunD están significativamente sobreexpresados (Figura 9). En la ruta PI3K-AKT también predomina la inhibición, incluyendo una fuerte disminución de la expresión en PI3K y MYC (Figura 10). En cuanto a la apoptosis, se aprecia una intensa inhibición de varios de los genes implicados en la ejecución del proceso apoptótico, aunque no se aprecian cambios significativos y claros en los genes pro-apoptóticos y anti-apoptóticos (Figura 11). En la ruta de procesamiento y presentación de antígenos se ha detectado una inhibición en el procesamiento de los péptidos de la ruta de MHCI (Figura 12). En cuanto al metabolismo central del carbono en cáncer, se han detectado una gran cantidad de genes significativos, la práctica totalidad inhibidos (Figura 13).

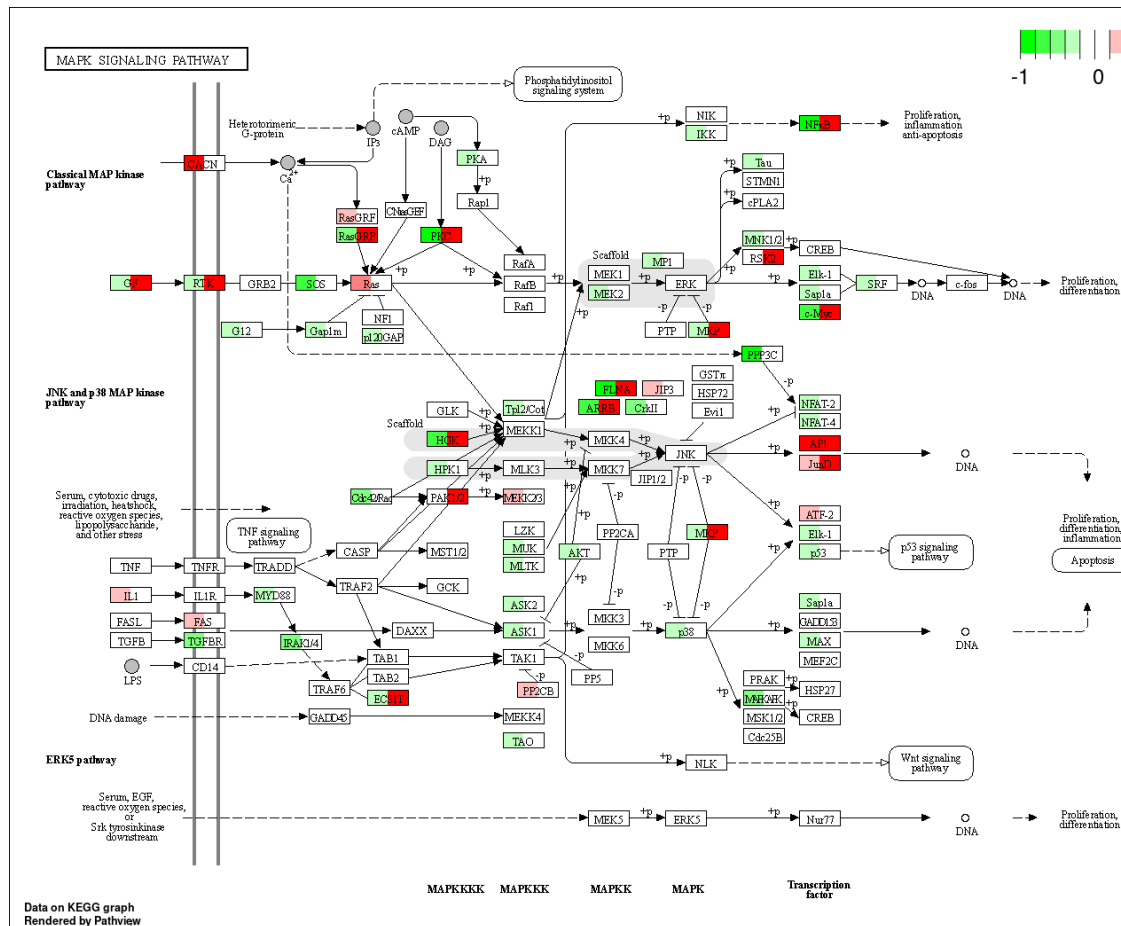


Figura 9. Inhibición de muchos de los miembros de la ruta de señalización de MAP Kinasas (MAPK). En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

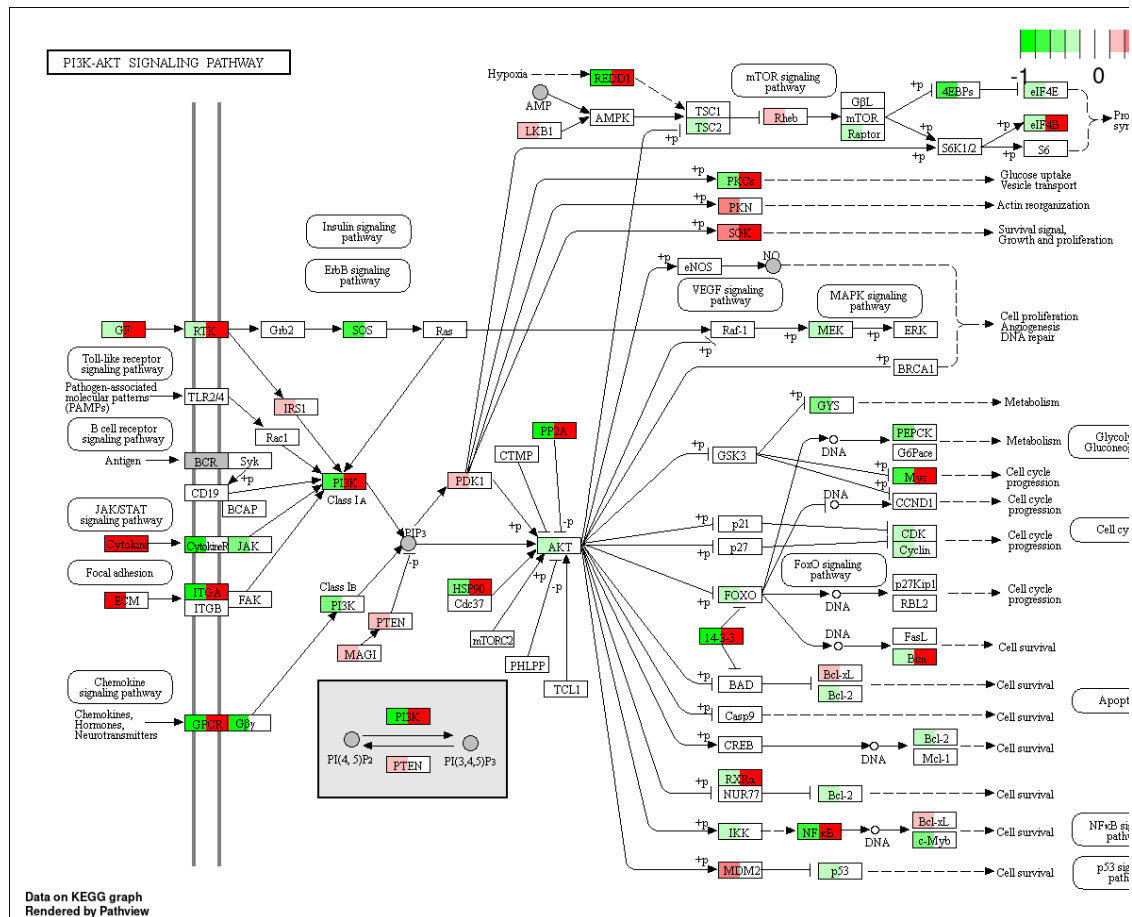


Figura 10. Inhibición generalizada en la ruta PI3K-AKT. En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

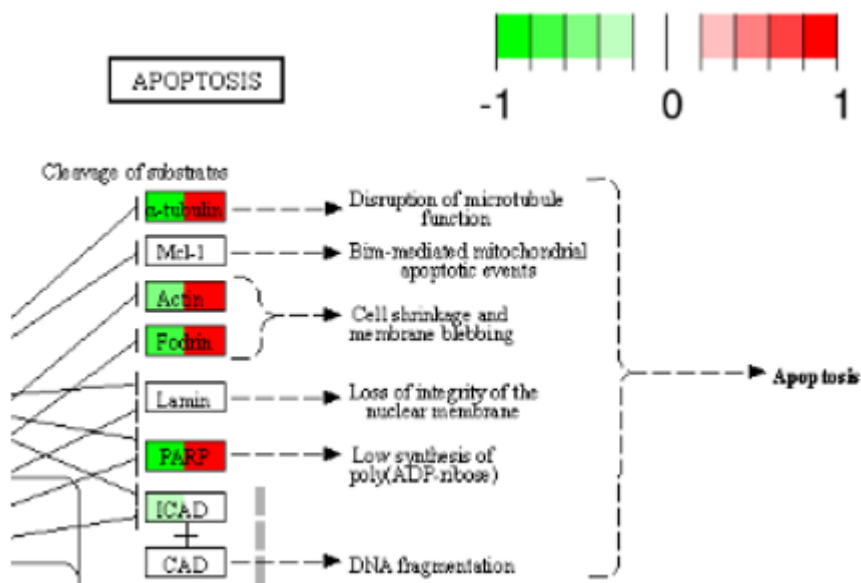


Figura 11. Inhibición de los genes implicado en la ejecución de la apoptosis. En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

ANTIGEN PROCESSING AND PRESENTATION

MHCI pathway

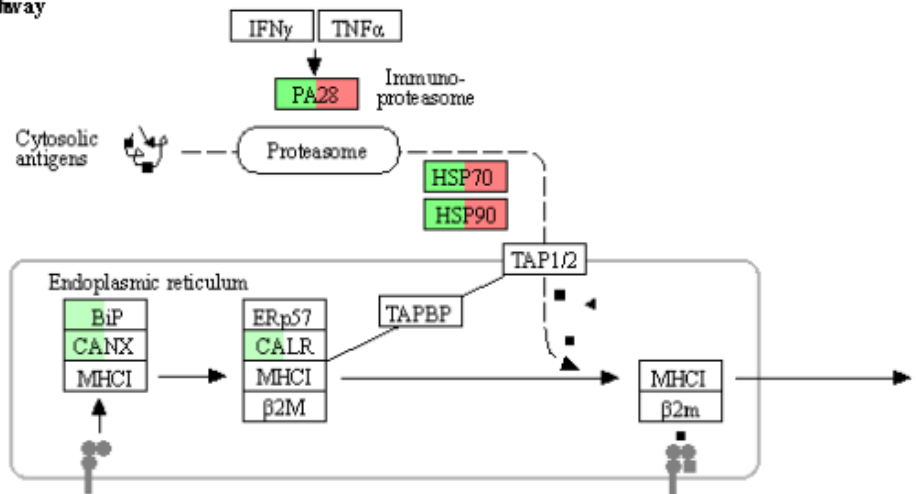


Figura 12. Disminución de los niveles de expresión de genes implicados en el procesamiento de péptidos en la ruta de MHC I. En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

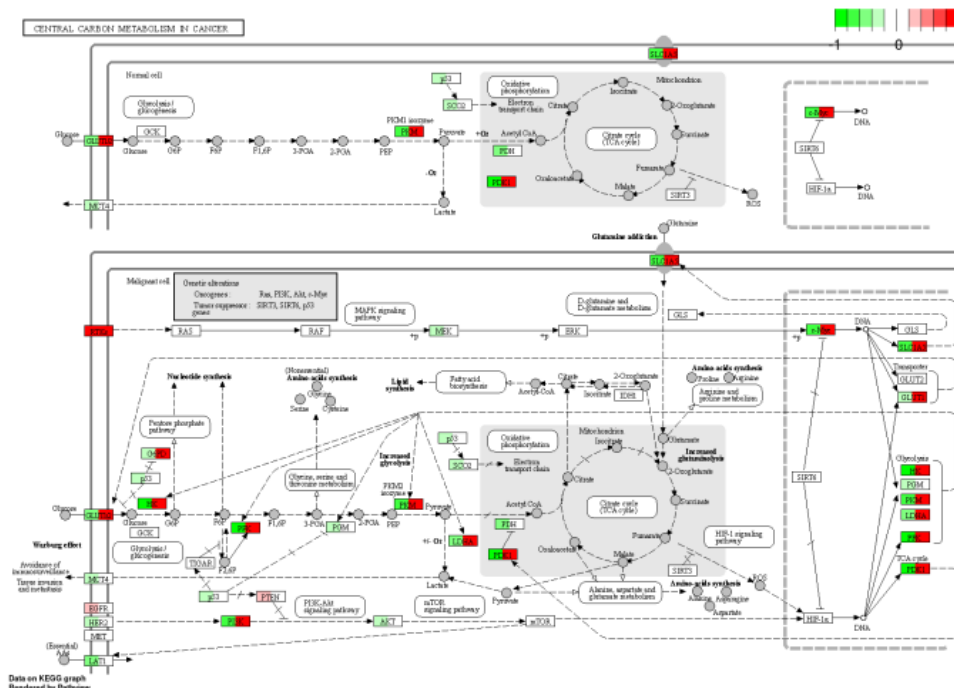


Figura 13. Disminución de los niveles de expresión de genes implicados en el metabolismo central del carbono en cáncer. En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

También habría una disminución en la síntesis de cAMP (Figura 14).



blanco si no.

En el metabolismo de las pirimidinas, se observa una inhibición de múltiples pasos repartidos por diferentes lugares de la ruta, aunque hay una leve sobreexpresión de las enzimas implicadas en el paso de dUMP a dUDP y de dTMP a dTDP (Figura 15).

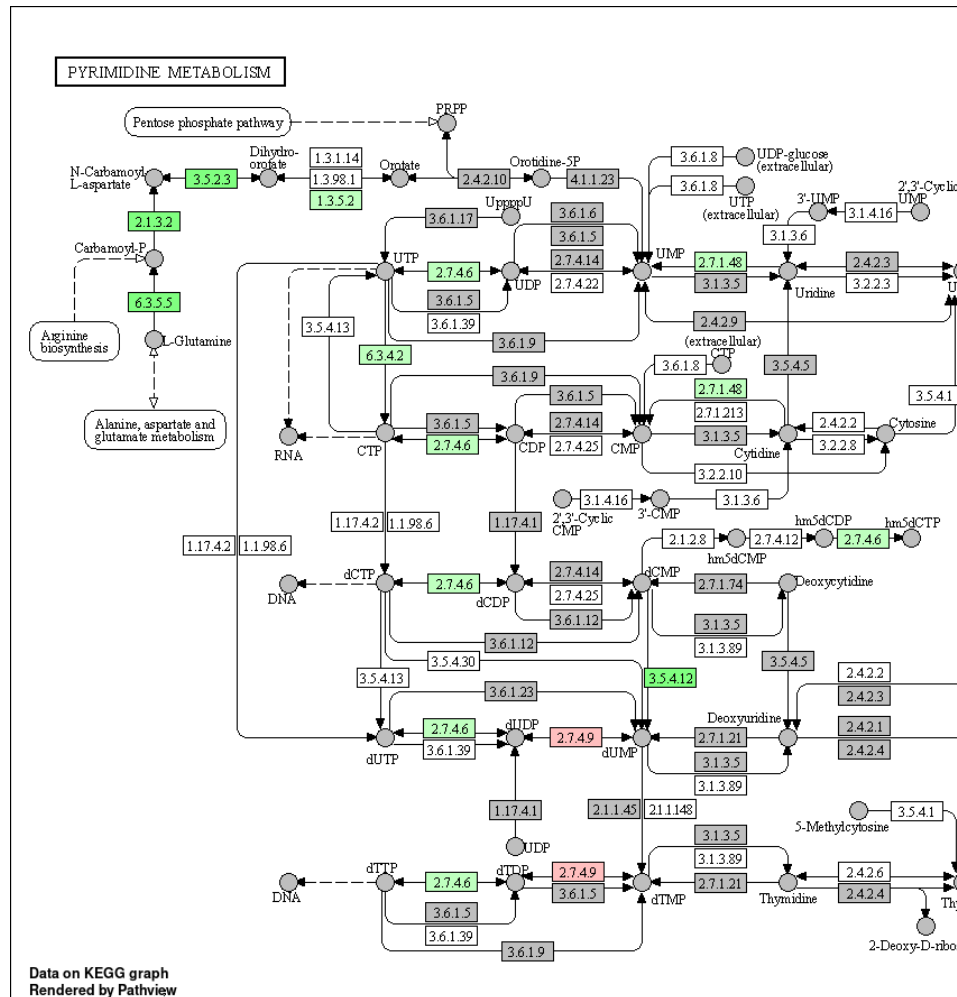


Figura 15. Sección con resultados significativos de la ruta del metabolismo de las pirimidinas. En cada gen, en la mitad izquierda se indica el log2FC normalizado, siguiendo la escala de la esquina superior derecha. En la mitad derecha aparecen en rojo si son significativos estadísticamente y en blanco si no.

Discusión

Tanto el análisis de clustering como el de PCA indican que las diferencias son mucho mayores entre líneas celulares que entre tratamientos. En el clustering se aprecia por la forma en la que agrupa las muestras. En PCA, aunque los triplicados están distribuidos cada uno hacia una de las esquinas de la gráfica, no hay apenas diferencias en PC1 (eje X) para las muestras de la misma línea celular, y es este eje el que supone casi un 80% de la variabilidad.

En el análisis de expresión diferencial vemos que ni la cantidad de genes significativos ni la ratio de genes inhibidos/sobreexpresados se mantiene al comparar el análisis de las líneas de forma conjunta con cada línea por separado. Esto es relevante porque muestra cómo el tratamiento puede depender de las características específicas de la neoplasia. En su aplicación clínica podría tener diferentes efectividades en función de las mutaciones y características de cada paciente. Además, vemos grandes discrepancias en las posiciones en cuanto a significancia de los genes destacados en el estudio original (*HES1*, *MYC* y *DTX1*). Parece ser que cuando miramos características comunes a ambas líneas estos genes sí son de los más relevantes, pero cuando miramos a cada línea por separado puede haber otros resultados más interesantes que se podrían analizar en el futuro.

La inhibición observada en múltiples procesos involucrados en la síntesis y degradación de proteínas, el transporte intracelular, las deficiencias en el metabolismo de azúcares y ácidos nucleicos y las deficiencias observadas en mitocondrias son procesos con efectos pleiotrópicos, los cuales podrían jugar un papel importante en los efectos anti-proliferativos y en la inhibición de la progresión de la leucemia observados en los ensayos de Moellering et al². Estos resultados son consistentes tanto en GO, como GSEA y KEGG. En los procesos relacionados con el ciclo celular se observaban resultados contradictorios. Aunque el análisis GO muestre algunos procesos que parecen favorecer la progresión del ciclo, al observarse procesos opuestos y numerosos procesos relacionados con los checkpoints del ciclo celular, podrían no ser suficientes para lograr el avance del ciclo. Además, en el caso de que lo consiguieran estimular, las deficiencias metabólicas y energéticas podrían impedirlo. Se deberían hacer ensayos que comprobaran el estado energético de la célula y la eficiencia de las principales rutas metabólicas.

Una disminución en los genes implicados en el mantenimiento telomérico podría también estar limitando la capacidad proliferativa de las células cancerígenas. El mantenimiento de los telómeros es imprescindible para que las células inmortalizadas puedan seguir replicando su DNA en cada ciclo celular sin perder progresivamente material genético y evitando fusiones de cromosomas⁹.

Una disminución en la señalización de citoquinas podría ser otro impacto positivo del tratamiento. Se ha visto que IL-7 modula la supervivencia y proliferación de células T tanto normales como malignas. Sin embargo, las células T sanas no la producen ellas mismas, pero las células T-ALL sí, favoreciendo el desarrollo oncogénico¹⁰. La interleucina 4 producida en la médula ósea es capaz de lograr una estimulación de la proliferación y el crecimiento de células T-ALL al estimular la señalización de mTOR¹¹. IL-9 tiene un rol antitumoral en tumores sólidos, pero su efecto es opuesto en neoplasias hematológicas, en las que actúa como factor de crecimiento¹². La disminución en la señalización de estas citoquinas podría jugar un papel en los efectos anti-tumorales de SAHM1.

Los resultados obtenidos en el análisis GO de células HPB-ALL sugieren una remodelación de la estructura de la cromatina. Si esta remodelación causa la disrupción de mecanismos oncogénicos tendría un efecto antitumoral. Por el contrario, si afecta a procesos celulares normales, podría actuar como una fuente de variabilidad para la aparición de nuevas variantes tumorales. Ensayos que caracterizaran estos cambios podrían arrojar más luz sobre este asunto. Otra posible arma de doble filo en la que se observan cambios significativos es la autofagia, la cual se encontraría inhibida. Una reducción en la autofagia puede inhibir la degradación de componentes o proteínas dañados en células con un elevado estrés oxidativo, favoreciendo la supervivencia celular a pesar del estrés. Por otro lado, en tumores con RAS mutado se ha visto que un aumento en la autofagia podría favorecer al desarrollo tumoral¹³. Además, los productos catabólicos de este proceso son utilizados para la producción de energía, aumentando la supervivencia celular, aunque este efecto podría no suceder en las células tratadas con SAHM1 por la inhibición metabólica descrita anteriormente. En leucemias la autofagia puede tanto promover quimiorresistencia como tener efectos supresores de tumores en función del contexto y el tratamiento¹³, por lo que habría que estudiarlo más en profundidad para obtener conclusiones.

En las rutas significativas aparecen las de varios virus con potencial oncogénico, como los virus de la Hepatitis C y B, el virus del papiloma humano, herpes simple y el virus linfotrópico de células T humanas. La presencia de estas rutas por supuesto no significa que haya una infección por estos virus que está siendo inhibida, sino que seguramente sean significativas porque en su ciclo infectivo estos virus manipulan mecanismos celulares que también están frecuentemente alterados en cáncer. Su objetivo es evitar la muerte celular a pesar de la infección y para favorecer la proliferación de células infectadas. De hecho, la infección por estos virus es capaz de aumentar el riesgo de cáncer, al proveerlas de algunos de los requisitos para convertirse en células cancerígenas. El hecho de que en estas rutas predomine la inhibición denota un efecto anti-tumoral.

En cuanto a las rutas reguladas por NOTCH1, los resultados de los gene sets con dianas de NOTCH1 y con genes regulados por el tratamiento con GSI muestran las mayores significancias. Respecto a rutas KEGG, la ruta de las MAPK se aprecia una inhibición generalizada, a excepción de AP-1. Tras analizar la expresión de los genes que codifican para las subunidades de AP-1 (FOS y JUN), se vio que solamente era JUN el que estaba sobreexpresado, por lo que, al no haber un aumento del otro miembro del heterodímero, el factor de transcripción podría no estar ejerciendo sus efectos pro-tumorales. También destaca la intensa inhibición de PKC, MYC y RasGRP. PKC está involucrada en múltiples rutas de señalización, controlando la expresión de genes importantes en la progresión del ciclo celular, tumorigénesis y metástasis¹⁴. La regulación de MYC se encuentra alterada en más del 50% de los cánceres y juega un papel central en prácticamente la totalidad del proceso oncogénico, regulando la proliferación, apoptosis, diferenciación y metabolismo¹⁵. Las proteínas RasGRP son expresadas por células hematopoyéticas y actúan como activadores del proto-oncogen RAS y otras GTPasas al actuar como factores de intercambio de nucleótidos de guanina¹⁶. En la ruta PI3K-AKT también muestra inhibición en una gran cantidad de genes, constituyendo otro mecanismo por el cual SAHM1 podría estar inhibiendo la progresión de la leucemia, ya que esta ruta está implicada en la proliferación, crecimiento, apoptosis, invasión, metástasis y transición epitelio-mesénquima¹⁷. Los resultados de GSEA Hallmarks apoyan la inhibición de estas rutas con los resultados inhibidos significativos para la señalización de MYC y PI3K. Por tanto, concluimos que SAHM1 es eficaz a la hora de inhibir la expresión de las dianas de NOTCH1.

Respecto a la apoptosis, se observa una marcada inhibición de varios genes implicados en la ejecución de la apoptosis, como los que codifican para alfa-tubulina, actina, fodrina y PARP. Se comprobó que en la línea HPB-ALL había un incremento significativo en la actividad de las caspasas 3 y 7, pero se debería comprobar también en KOPT-K1. Además, se debería comprobar si la ejecución de la apoptosis es correcta en las células con actividad de caspasas 3 y 7 incrementada, o si no todas las enzimas y cambios necesarios para la muerte celular están teniendo lugar. Si el tratamiento frena la proliferación, pero no se logra la muerte de todas las células tumorales, algunas podrían sobrevivir al tratamiento y causar una posterior recaída.

Si los mecanismos de presentación de péptidos en MHC-I se encuentran disminuidos, el reconocimiento por parte del sistema inmune de células tumorales podría ser deficiente. Las células cancerígenas poseen multitud de mutaciones, que resultan en la generación de neoantígenos potencialmente reconocibles por el sistema inmune¹⁸. La presentación de péptidos provenientes de la propia célula se hace mediante la carga de péptidos resultantes de la degradación proteica en el proteasoma de proteínas propias en el MHC-I. Es frecuente en cáncer la aparición de mecanismos de evasión del sistema inmune que consisten en evitar la carga de péptidos en el MHC-I o la llegada a la superficie del receptor¹⁹. Se deberían hacer ensayos que comprobaran la capacidad de reconocimiento por parte del sistema inmune de células T-ALL tratadas con SAHM1.

De este reanálisis podemos concluir que SAHM1 ejerce potentes efectos antitumorales a múltiples niveles, inhibiendo multitud de procesos y rutas celulares, muchos centrales para la progresión de la leucemia. También se ha observado que los efectos varían entre una línea celular y otra y se deberían analizar en mayor profundidad cada línea por separado. Además, se puede apreciar la utilidad de que los datos de experimentos pasados se encuentren disponibles públicamente, ya que se puede seguir extrayendo información de elevada utilidad a partir de ellos.

Bibliografía:

1. Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321-337.e10 (2018).
2. Moellering, R. E. *et al.* Direct inhibition of the NOTCH transcription factor complex. *Nature* **462**, 182–188 (2009).
3. Notch signaling in human development and disease - PubMed.
<https://pubmed.ncbi.nlm.nih.gov/22306179/>.
4. Louvi, A. & Artavanis-Tsakonas, S. Notch and disease: a growing field. *Semin. Cell Dev. Biol.* **23**, 473–480 (2012).
5. Palomero, T. *et al.* NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proc. Natl. Acad. Sci.* **103**, 18261–18266 (2006).
6. Choi, S. H. *et al.* The common oncogenomic program of NOTCH1 and NOTCH3 signaling in T-cell acute lymphoblastic leukemia. *PLOS ONE* **12**, e0185762 (2017).
7. Ferrando, A. A. The role of NOTCH1 signaling in T-ALL. *Hematol. Educ. Program Am. Soc. Hematol. Am. Soc. Hematol. Educ. Program* 353–361 (2009) doi:10.1182/asheducation-2009.1.353.
8. Girardi, T., Vicente, C., Cools, J. & De Keersmaecker, K. The genetics and molecular biology of T-ALL. *Blood* **129**, 1113–1123 (2017).
9. Harley, C. B. Telomerase and cancer therapeutics. *Nat. Rev. Cancer* **8**, 167–179 (2008).

10. Buffière, A. *et al.* T-cell acute lymphoblastic leukemia displays autocrine production of Interleukin-7. *Oncogene* **38**, 7357–7365 (2019).
11. Cardoso, B. A. *et al.* Interleukin-4 stimulates proliferation and growth of T-cell acute lymphoblastic leukemia cells by activating mTOR signaling. *Leukemia* **23**, 206–208 (2009).
12. Wan, J. *et al.* IL-9 and IL-9-producing cells in tumor immunity. *Cell Commun. Signal.* **18**, 50 (2020).
13. Evangelisti, C. *et al.* Autophagy in acute leukemias: A double-edged sword with important therapeutic implications. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1853**, 14–26 (2015).
14. Garg, R. *et al.* Protein kinase C and cancer: what we know and what we do not. *Oncogene* **33**, 5225–5237 (2014).
15. Chen, H., Liu, H. & Qing, G. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct. Target. Ther.* **3**, 5 (2018).
16. Ksionda, O., Limnander, A. & Roose, J. P. RasGRP Ras guanine nucleotide exchange factors in cancer. *Front. Biol.* **8**, 508–532 (2013).
17. Jiang, N. *et al.* Role of PI3K/AKT pathway in cancer: the framework of malignant behavior. *Mol. Biol. Rep.* **47**, 4587–4629 (2020).
18. Blass, E. & Ott, P. A. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. *Nat. Rev. Clin. Oncol.* **18**, 215–229 (2021).
19. Cornel, A. M., Mimpfen, I. L. & Nierkens, S. MHC Class I Downregulation in Cancer: Underlying Mechanisms and Potential Targets for Cancer Immunotherapy. *Cancers* **12**, 1760 (2020).