



UNIVERSIDAD PERUANA
CAYETANO HEREDIA
ESCUELA DE POSGRADO

**ANÁLISIS INTEGRATIVO
COMPUTACIONAL DE GENOMAS
DE CÁNCER GINECOLÓGICO Y
ELUCIDACIÓN DE ALTERACIONES
EN LOS MECANISMOS
EPIGENÉTICOS**

TESIS PARA OPTAR EL GRADO DE
MAESTRO
EN INFORMÁTICA BIOMÉDICA
EN SALUD GLOBAL CON MENCIÓN EN
BIOINFORMÁTICA

YOMARA KORALY ROMERO LÁZARO

LIMA - PERÚ

2020

Asesora:

Dra. Claudia Inés Gloria Machicado Rivero

Dedicatoria:

A Dios por iluminar y guiar mi vida, a mi familia y amigos por todo el
apoyo brindado a lo largo de este tiempo.

Agradecimientos

A mi asesora Dra. Claudia Inés Gloria Machicado Rivero, por la constante orientación y tutoría en la realización de este trabajo, mi más sincera gratitud por toda la dirección y atención brindada.

A la Unidad de Informática Biomédica en Salud Global de la Facultad de Salud Pública y Administración (FASPA) por permitirme ser parte del programa de esta Maestría.

Fuente de Financiamiento

Al fondo de investigaciones otorgado por El Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) y al Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT).

TABLA DE CONTENIDO

RESUMEN
ABSTRACT

I.	INTRODUCCIÓN	1
II.	PLANTEAMIENTO DE LA INVESTIGACIÓN	4
2.1.	PLANTEAMIENTO DEL PROBLEMA	4
2.2.	MARCO TEÓRICO	6
2.2.1.	Cáncer	6
2.2.2.	Cáncer ginecológico	7
2.2.3.	Tipos de cáncer ginecológico	8
2.2.3.1.	Cáncer cervical	8
2.2.3.2.	Cáncer de ovario	9
2.2.3.3.	Cáncer de útero	9
2.2.3.4.	Cáncer de vulva y vagina	11
2.2.4.	Epigenética y Cáncer	12
2.2.5.	Principales mecanismos epigenéticos	13
2.2.5.1.	Metilación del ADN	13
2.2.5.2.	Reguladores de histonas	15
2.2.6.	Reguladores epigenéticos: Escritores, lectores y borradores	17
2.2.6.1.	Los escritores	17
2.2.6.2.	Los lectores	17
2.2.6.3.	Los borradores	18
2.2.7.	Bioinformática aplicada al estudio del cáncer	19
2.3.	PREGUNTA DE INVESTIGACIÓN	20
2.4.	JUSTIFICACIÓN	21
III.	OBJETIVOS	22
3.1.	Objetivo general	22
3.2.	Objetivos específicos	22
IV.	METODOLOGÍA	23
4.1.	DISEÑO DEL ESTUDIO	23
4.2.	POBLACIÓN	23
4.3.	CRITERIOS DE SELECCIÓN	23
4.4.	OPERACIONALIZACIÓN DE VARIABLES	25
4.5.	INSTRUMENTOS	27
4.6.	PROCEDIMIENTO	29
4.6.1.	Elaboración de la lista de genes de estudio	30
4.6.2.	Análisis de anotación funcional y análisis de enriquecimiento de rutas bioquímicas de los genes de interés	32
4.6.3.	Bases de datos de genomas y de la población de estudio	32
4.6.4.	Identificación de los genes significativamente mutados	35
4.6.5.	Visualización 3D de las proteínas CREBBP, KMT2C y KMT2D nativas y mutadas y construcción de red de PPI	36
4.6.6.	Predicción del efecto de las sustituciones de aminoácidos en la función de las proteínas	37

4.6.7.	Estimación del cambio de la estabilidad energética teórica de las proteínas	39
4.6.8.	Predicción de la accesibilidad al solvente de las proteínas	40
4.6.9.	Identificación de las alteraciones en el número de copias (CNA)	41
4.6.10.	Identificación de Polimorfismo de Nucleótido Simple (SNP)	41
4.6.11.	Identificación de genes con expresión diferencial	42
4.6.12.	Anotación funcional GO y análisis de enriquecimiento de vías KEGG de los DEGs.....	44
4.6.13.	Construcción de la red de interacción proteína-proteína (PPI) de los DEGs y la identificación de proteínas centrales.....	44
4.6.14.	Estudio de asociación genotipo-fenotipo	46
4.6.14.1.	Asociación entre la expresión génica, la supervivencia global y supervivencia libre de enfermedad	46
4.6.14.2.	Asociaciones entre la expresión génica y variables fenotípicas.....	47
4.6.15.	Metilación del ADN a nivel del promotor del gen.....	48
V.	RESULTADOS	51
5.1.	Lista de genes implicados en mecanismos epigenéticos.....	51
5.2.	Anotación funcional y análisis de enriquecimiento de rutas bioquímicas de los genes de interés.....	52
5.3.	Características de la población de estudio	58
5.3.1.	Características demográficas en pacientes con cáncer ginecológico	58
5.3.2.	Características clínico-patológicas en pacientes con cáncer ginecológico	61
5.3.3.	Características generales de los pacientes con cáncer ginecológico admitidos en el INEN, Perú.....	64
5.4.	Genes frecuentemente mutados y significativamente mutados.....	68
5.5.	Predicción de los cambios en la estructura, estabilidad y función proteica	72
5.6.	Alteraciones en el número de copias (CNAs).....	86
5.7.	Identificación de Polimorfismo de Nucleótido Simple (SNPs)	92
5.8.	Identificación de genes con expresión diferencial	93
5.9.	Red de interacción proteína-proteína (PPI) y caracterización funcional GO y análisis de enriquecimiento de vías KEGG de los DEGs	99
5.10.	Estudio de asociación genotipo-fenotipo	116
5.11.	Estado de metilación del ADN en promotores de genes de estudio.....	127
VI.	DISCUSIÓN	135
VII.	CONCLUSIONES	158
VIII.	RECOMENDACIONES	161
IX.	LIMITACIONES.....	163
X.	REFERENCIAS BIBLIOGRÁFICAS.....	164
ANEXO 1		
ANEXO 2		
ANEXO 3		
ANEXO 4		
ANEXO 5		

Lista de Figuras

		Pág.
Figura 1	Anatomía del sistema reproductor femenino.	7
Figura 2	Metilación del ADN.	14
Figura 3	Representación esquemática de las principales modificaciones en las colas de las histonas.	16
Figura 4	Reguladores epigenéticos de la cromatina.	18
Figura 5	Flujograma general del trabajo realizado.	29
Figura 6	Flujograma del proceso aplicado para la elaboración de la lista de genes de estudio.	31
Figura 7	Pasos para realizar la consulta en Cbioportal.	34
Figura 8	Esquema que muestra la clasificación y número de los genes de estudio.	51
Figura 9	Anotación funcional GO y enriquecimiento de vías KEGG según Enrichr en los 160 genes de interés.	53
Figura 10	Clustergrama del conjunto de genes de estudio asociados a los 10 principales procesos biológicos, mediante Enrichr.	55
Figura 11	Clustergrama del conjunto de genes de estudio involucrados en las 10 principales vías KEGG, mediante Enrichr.	57
Figura 12	Recuento de mutaciones puntuales y genes frecuentemente mutados según el tipo de cáncer ginecológico.	70
Figura 13	Mapa de árbol de los genes frecuentemente mutados según el tipo de cáncer ginecológico.	70
Figura 14	Porcentaje de mutaciones con pérdida de sentido en KMT2C, KMT2D y CREBBP predichas por Biomuta.	73
Figura 15	Número y porcentaje de mutaciones con pérdida de sentido “dañinas” que mostraron cambios en la estabilidad energética de las proteínas KMT2C, KMT2D y CREBBP, predichas por I-Mutant 3.0.	75
Figura 16	Localización de variantes “con pérdida de sentido” en KMT2D y su red de PPI.	80
Figura 17	Localización de variantes “con pérdida de sentido” en KMT2C y su red de PPI	83

Figura 18	Localización de variantes “con pérdida de sentido” en CREBBP (proteína de unión a CREB) y su red de PPI.	85
Figura 19	Porcentaje de genes que presentaron CNAs según el tipo de cáncer ginecológico.	86
Figura 20	Diagrama de Venn que muestra el número de genes comunes entre CESC, OV, UCEC y UCS, localizados en regiones con CNA significativas.	89
Figura 21	Genes con expresión diferencial (DEGs) según el tipo de cáncer ginecológico.	97
Figura 22	Diagrama de Venn que muestra los DEGs únicos de cada neoplasia, así como aquellos comunes a CESC, OV, UCEC y UCS.	98
Figura 23	Red de interacción proteína-proteína (PPI) mediante STRING en CESC.	100
Figura 24	Análisis de enriquecimiento GO de los DEGs en CESC según el análisis Genecodis.	102
Figura 25	Red de interacción proteína-proteína (PPI) mediante STRING en OV.	104
Figura 26	Análisis de enriquecimiento GO de los genes expresados diferencialmente en OV según el análisis del recurso Genecodis.	106
Figura 27	Red de interacción proteína-proteína (PPI) mediante STRING en UCEC.	108
Figura 28	Análisis de enriquecimiento GO de los genes expresados diferencialmente en UCEC según el análisis del recurso Genecodis.	110
Figura 29	Red de interacción proteína-proteína (PPI) mediante STRING en UCS.	112
Figura 30	Análisis de enriquecimiento GO de los genes expresados diferencialmente en UCS según el análisis del recurso Genecodis.	114
Figura 31	DEGs identificados en los cánceres ginecológicos asociados con el pronóstico de la enfermedad.	120
Figura 32	Gráfica de Kaplan Meier representando la OS en función de la expresión génica de SMYD2 en pacientes con CESC.	121
Figura 33	Gráfica de Kaplan Meier representando la DFS en función de la expresión génica de KAT2B en pacientes con CESC.	121
Figura 34	Gráfica de Kaplan Meier representando la OS en función de la expresión génica de SETDB2 en pacientes con OV.	126

Figura 35	Gráfica de Kaplan Meier representando la OS en función de la expresión génica de MGEA5 en pacientes con OV.	123
Figura 36	Gráfica de Kaplan Meier representando la OS en función de la expresión génica de HDAC10 en pacientes con OV.	124
Figura 37	Gráfica de Kaplan Meier representando la DFS en función de la expresión génica de NEK6 en pacientes con OV.	125
Figura 38	Gráfica de Kaplan Meier representando la OS en función de la expresión génica de CDK5 en pacientes con UCS.	125
Figura 39	Metilación diferencial a nivel del promotor de los genes de interés.	128
Figura 40	Niveles de metilación a nivel del promotor de 19 genes con metilación diferencial en CESC.	129
Figura 41	Niveles de metilación a nivel del promotor de 4 genes con metilación diferencial en UCEC.	130
Figura 42	Correlación entre la expresión del ARNm y la metilación del promotor en los tumores de pacientes con OV (n=299) basado en MEXPRESS.	133

Lista de tablas

	Pág.
Tabla 1	Bases de datos y herramientas bioinformáticas especializadas. 27
Tabla 2	Características generales de los pacientes con cáncer ginecológico. 60
Tabla 3	Características clínico-patológicas de los pacientes con cáncer ginecológico. 63
Tabla 4	Características generales de los pacientes con cáncer ginecológico admitidos en el INEN, Perú. 66
Tabla 5	Genes frecuentemente mutados en cánceres ginecológicos evaluados por MutSig 2CV. 71
Tabla 6	Mutaciones puntuales predichas “dañinas” en KMT2D, incluyendo el cambio de estabilidad de cada mutante y cambio de accesibilidad al solvente. 78
Tabla 7	Mutaciones puntuales predichas “dañinas” en KMT2C, incluyendo el cambio de estabilidad de cada mutante y cambio de accesibilidad al solvente. 81
Tabla 8	Mutaciones puntuales predichas “dañinas” en CREBBP, incluyendo el cambio de estabilidad de cada mutante y cambio de accesibilidad al solvente 84
Tabla 9	Genes que presentan alteración en el número de copias (CNA), presente en cáncer ginecológico, determinados por GISTIC 2.0 91
Tabla 10	SNPs asociados a cáncer ginecológico identificados por estudios GWAS 92
Tabla 11	Análisis de enriquecimiento de las vías KEGG en DEGs presentes en CESC. 103
Tabla 12	Análisis de enriquecimiento de las vías de KEGG en genes expresados diferencialmente en OV. 107
Tabla 13	Análisis de enriquecimiento de las vías de KEGG en genes expresados diferencialmente en UCEC 111
Tabla 14	Análisis de enriquecimiento de las vías de KEGG en genes expresados diferencialmente en UCS. 115
Tabla 15	Correlación entre la expresión de los DEGs y la edad al momento del diagnóstico de las pacientes con cáncer ginecológico 118
Tabla 16	Correlación entre la expresión de genes y el peso de pacientes con cáncer ginecológico 119

RESUMEN

El cáncer de tipo ginecológico representa un problema de salud pública y es una de las principales causas de morbilidad y mortalidad que afectan a las mujeres a nivel mundial, en parte debido a la poca disponibilidad de servicios de salud preventiva en países económicamente poco desarrollados. Disciplinas como la genómica y epigenómica han contribuido al conocimiento y a la mejor comprensión de la biología del cáncer. En este estudio se compararon las alteraciones genéticas y epigenéticas de un grupo de reguladores epigenéticos, en mujeres con distintos tipos de cáncer ginecológico. Además, se establecieron asociaciones entre el genotipo y fenotipo con la finalidad de reconocer marcadores pronósticos de los cánceres ginecológicos. Para el procesamiento y análisis secundario de los datos se emplearon programas bioinformáticos y recursos en web. Los resultados mostraron que los reguladores epigenéticos contienen modificaciones a varios niveles incluida la secuencia de ADN, metilación en promotores, expresión génica y funcionalidad proteica. La información contenida en este trabajo se espera que contribuya al reconocimiento de nuevas dianas terapéuticas y marcadores pronósticos de las neoplasias malignas ginecológicas.

Palabras clave: epigenética, cáncer ginecológico, *in silico*, bioinformática, genoma.

ABSTRACT

Gynecological cancer represents a public health problem and is one of the main causes of morbidity and mortality affecting women worldwide, in part due to the limited availability of preventive health services in economically underdeveloped countries. Disciplines such as genomics and epigenomics have contributed to the knowledge and better understanding of cancer biology. In this study, we compared the genetic and epigenetic alterations of a group of epigenetic regulators, in women with different types of gynecological cancer. In addition, associations between genotype and phenotype were established in order to recognize prognostic markers of gynecological cancers. For the processing and secondary analysis of the data, bioinformatics programs and web resources were used. The results showed that epigenetic regulators contain modulations at several levels including DNA sequence, promoter methylation, gene expression and protein functionality. The information contained in this work is expected to contribute to the recognition of new therapeutic targets and prognostic markers of gynecological malignant neoplasms.

Key words: epigenetics, gynecological cancer, *in silico*, bioinformatics, genome.

I. INTRODUCCIÓN

Los cánceres ginecológicos son un conjunto de tumores malignos que surgen en el sistema reproductivo femenino. Dentro de este grupo se encuentran el cáncer cervical, cáncer de ovario, cáncer de útero, cáncer vaginal, cáncer de vulva y muy poco frecuente el cáncer de las trompas de Falopio (1). Cada tipo de cáncer ginecológico presenta signos y síntomas específicos, así como diferentes factores de riesgo que pueden incrementar el desarrollo de algún tipo de cáncer ginecológico (2). Entre estos factores de riesgo, se encuentra la edad, antecedentes familiares de cáncer, identificación de mutaciones genéticas (*BRCA1* / *BRCA2*), historial reproductivo, exposición a hormonas producidas por el cuerpo o tomadas como medicamento, infección viral como la infección por el virus del papiloma humano (VPH) y estilos de vida como fumar y aquellos que conducen a la obesidad (3).

En el 2018, la Agencia Internacional para la Investigación del Cáncer (IARC), a través de la base de datos GLOBOCAN, registró 1 309 165 nuevos casos de cánceres ginecológicos y 609 377 muertes en todo el mundo a causa de estos tipos de cáncer (4), afectando de manera desproporcionada a los países de ingresos bajos y medios, entre ellos a Perú (5). La mayoría de estas muertes se podrían evitar si se dispusieran de servicios de diagnósticos y tratamientos oportunos, así como una adecuada inversión en la educación sobre la prevención del cáncer, capacitación de profesionales de salud e investigación (6,7).

En tanto se afrontan estas barreras para combatir el cáncer, la participación de la bioinformática en la investigación del cáncer ha mejorado la comprensión y

caracterización de la biología molecular en esta enfermedad. La utilización de nuevas tecnologías de alto rendimiento como la secuenciación de próxima generación (NGS) el cual genera grandes cantidades de datos de manera fácil y económica (8–10). Al mismo tiempo, se han desarrollado algoritmos y herramientas computacionales que facilitan el análisis de estos resultados (11).

Como consecuencia de esta explosión de datos, se han llevado a cabo iniciativas multicéntricas a gran escala como el proyecto de El Atlas del Genoma del Cáncer (TCGA) (3). Esta iniciativa se encarga de analizar las secuencias de genomas completos de diversos tipos de neoplasias con el objetivo de describir el panorama completo de alteraciones genéticas por tipo de cáncer, habiendo revelado una inmensa heterogeneidad en los tumores ginecológicos (12–16). Sin embargo, la complejidad del cáncer no solo implica alteraciones genéticas, sino que también involucran cambios epigenéticos. Los mecanismos epigenéticos regulan la expresión de los genes sin alterar la secuencia de ADN y han sido estudiados para conducir el desarrollo de fármacos para tratar el cáncer (17).

Los principales mecanismos epigenéticos son la metilación del ADN, modificaciones de histonas en la cromatina y no codificantes incluidos microARNs (miARN). La regulación de esta maquinaria ocurre con la participación de los escritores, los borradores y los lectores, proteínas que agregan, eliminan o reconocen grupos funcionales tanto en el ADN (grupos metilos) como en las histonas (grupos metilos, acetilos, fosfatos, etc.) influyendo así en la expresión génica (18,19). Alteraciones en estos mecanismos aparecen durante la

carcinogénesis evidenciándose por ejemplo una disminución de la metilación global en el genoma (20–22).

A pesar del considerable progreso para comprender las alteraciones genéticas y epigenéticas en el cáncer, hasta la fecha, pocos estudios se centran en evaluar las alteraciones en los genes con función epigenética en tumores ginecológicos. Conociendo la creciente atención por el estudio de mecanismos epigenéticos y su potencial rol en la carcinogénesis, se hace imperativo determinar las alteraciones en dichos mecanismos presentes en neoplasias ginecológicas.

El presente estudio establece por primera vez un análisis integrativo genómico de 160 genes implicados en mecanismos epigenéticos como modificación de ADN e histonas y su asociación con el fenotipo incluido pronóstico, presentes en cuatro tipos de cáncer ginecológico. El conocimiento generado podría conducir a identificar biomarcadores diagnósticos, pronósticos y terapéuticos en la oncología ginecológica.

II. PLANTEAMIENTO DE LA INVESTIGACIÓN

2.1. PLANTEAMIENTO DEL PROBLEMA

A nivel mundial el cáncer ginecológico en conjunto representa el 7.2% de todos tipos de cánceres diagnosticados en el 2018 (23). Según las estadísticas mundiales de cáncer de la Agencia Internacional de Investigación en Cáncer (IARC) y publicadas en el GLOBOCAN 2018 , se estima que en el 2040 se diagnosticarán 29 532 994 millones de nuevos casos de cáncer y 16 388 459 millones de muertes en todo el mundo (24) y solo una parte de los pacientes recibirán atención médica debido a la inequidad en el acceso a los servicios de salud que enfrentan los países de bajos y medianos ingresos (7,25).

Para reducir la carga que produce el cáncer, existen esfuerzos internacionales de la Organización Mundial de la Salud (OMS) y la Agencia Internacional para la Investigación del Cáncer para implementar actividades de prevención y control del cáncer(26). En América Latina, algunos países han asumidos un fuerte compromiso con la cobertura de la salud y han desarrollado planes nacionales de control de cáncer específicos (27), así mismo el Perú cuenta con programas de prevención de cáncer cervical con campañas de la vacunación contra el VPH en niñas entre los 9 y 13 años (28).

Sin embargo, estas intervenciones demandarán algunos años para mostrar su impacto, mientras tanto, es necesario ampliar el conocimiento sobre la biología del cáncer de manera que se pueda generar soluciones desde esta perspectiva. En las neoplasias ginecológicas, las alteraciones genéticas y epigenéticas han sido descritas y asociadas con su desarrollo y progresión (29). Las investigaciones en

genómica del cáncer se han visto mejoradas gracias al desarrollo de nuevas tecnologías de secuenciamiento de genoma completo por NGS. El conocimiento del genoma completo ha permitido identificar mutaciones específicas de cada tipo de tumor (por ejemplo la leucemia aguda humana) en los genes que codifican reguladores epigenéticos (30). En consecuencia, las proteínas codificadas por estos genes podrían funcionar como posibles dianas terapéuticas del cáncer (31).

En ese sentido, el presente estudio busca identificar alteraciones en los reguladores epigenéticos (modificación de histona y metilación del ADN) presentes en las neoplasias ginecológicas y a su vez asociarlas con variables fenotípicas y clínicas presentes en cada patología. La estrategia aprovechará los datos genómicos y epigenómicos generados en los proyectos mundiales de Genoma del Cáncer.

2.2. MARCO TEÓRICO

2.2.1. Cáncer

Según el Instituto Nacional del Cáncer, es el nombre que se asigna a un conjunto de enfermedades relacionadas en el que las células del cuerpo empiezan a dividirse sin control y pueden diseminarse a otros tejidos (32).

Dicho comportamiento conlleva a la formación de masas denominadas tumores, que pueden ser benignos o malignos. Un tumor benigno se desarrolla localmente, no invade el tejido normal circundante y no se extiende a los diferentes tejidos del cuerpo. Por el contrario, un tumor maligno invade el tejido normal circundante en el que las células cancerosas se separan del tumor original y son capaces de diseminarse a otras partes del cuerpo mediante el sistema circulatorio o linfático (metástasis) (33).

Existen diez características biológicas distintivas del cáncer denominadas “hallmarks”, las cuales ayudan a comprender la complejidad de las enfermedades neoplásicas. Estas incluyen: mantenimiento de la señalización proliferativa (las células cancerosas estimulan su propio crecimiento), evasión de los supresores del crecimiento (resisten señales inhibitoras del crecimiento), resistencia a la muerte celular, replicación indefinida, estimulación de la angiogénesis (formación de vasos sanguíneos para suministrar nutrientes y oxígeno a los tumores), invasión y metástasis, desregulación del metabolismo, evasión de sistema inmunológico, inestabilidad del genoma y la inflamación (34). A pesar del notable progreso en la comprensión de estas características, aún no está completamente dilucidado estos procesos, por lo que otras áreas como la

epigenética están actualmente abordando cuestiones que complementen estos estudios en el cáncer (34).

2.2.2. Cáncer ginecológico

Cáncer ginecológico es el término que se asigna al conjunto de cánceres que surgen en el sistema reproductor de la mujer (Figura 1). Existen seis tipos de cánceres ginecológicos, que incluyen el de tipo cervical, de útero (que incluye cáncer de endometrio), de vagina, de vulva , de ovario y raramente el cáncer de trompas de Falopio (2).

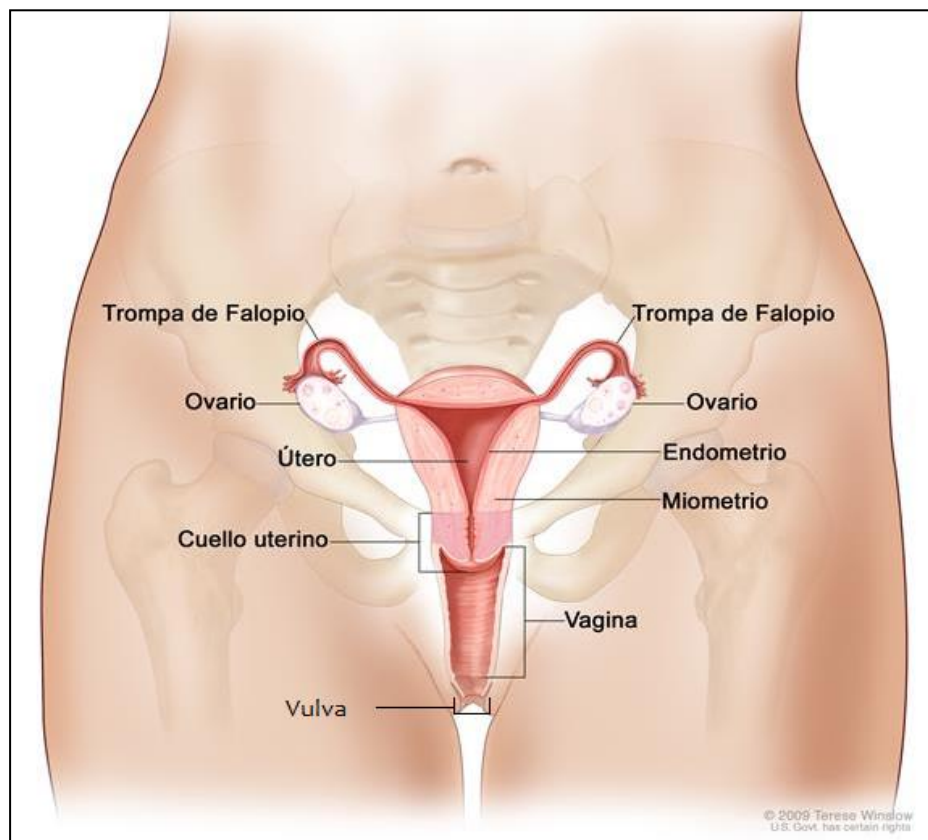


Figura 1. Anatomía del sistema reproductor femenino. Se muestran los órganos que conforman el sistema reproductivo de una mujer (ovarios, trompas de Falopio, útero, cérvix, vagina y vulva). Figura adaptada de diccionario de cáncer del Instituto Nacional del Cáncer (Estados Unidos) (176).

2.2.3. Tipos de cáncer ginecológico

2.2.3.1. Cáncer cervical

A nivel mundial, el cáncer cervical es el cuarto más común entre las mujeres después del cáncer de mama, colorrectal y pulmón. En 2018, la Agencia Internacional de Investigación en Cáncer (IARC) a través del GLOBOCAN estimó aproximadamente 569 847 (3.2%) nuevos casos con 311 365 (3.3%) muertes por año, donde la mayoría de los casos nuevos (85%) y las muertes (90%) ocurren en regiones económicamente menos desarrollados (4). En el Perú, el cáncer cervical es la segunda causa de muerte por cáncer en la mujer con 4 103 nuevos casos y 1 836 muertes por año (4). La infección por VPH 16 o 18 está presente en el 65.9% de mujeres con cáncer cervical, 53.1% de lesiones de alto grado, 27.3% de mujeres con lesiones cervicales de bajo grado y 6.6% de mujeres con citología normal (35). El cáncer cervical se puede prevenir gracias a la vacuna contra el virus de papiloma humano (VPH), de los 530,000 nuevos casos diagnosticado por año, el 71% de estos se deben al VPH subtipos 16 y 18, mientras que los subtipos de VPH 31,33,45,52 y 58 representan el 19% de los casos (36).

La introducción de programas de cribado y detección precoz como la prueba del Papanicolaou ,aumentan las posibilidades de curación de la paciente, ya que los síntomas como el sangrado o secreción vaginal anormal se presentan en etapas muy avanzadas (35).

2.2.3.2. Cáncer de ovario

A nivel mundial, la Agencia Internacional de Investigación en Cáncer (IARC) a través del GLOBOCAN estimó aproximadamente 295 414 (1.6%) nuevos casos con 184 799 (1.9%) muertes por año debido al cáncer de ovario (4). Así también, en el Perú, el cáncer de ovario tuvo 1 331 nuevos casos y 670 muertes por año en el 2018 (4).

Este cáncer se presenta con mayor frecuencia en mujeres de 50 años a más donde la mayoría de las neoplasias malignas son de tipo epitelial invasivo. En mujeres menores de 20 años, predominan los tumores de células germinales, mientras que en mujeres entre los 30 y 40 años suelen aparecer los tumores limítrofes (potencial maligno bajo) (37). Las probabilidades de desarrollarlo aumentan para aquellas mujeres que no hayan tenido hijos ,entre otros factores de riesgo se encuentra la presencia de mutaciones en los genes BRCA1 o BRCA2, la endometriosis, haber tenido cáncer de mama, cáncer de útero o cáncer colorrectal (37). Los signos y síntomas de este cáncer son el sangrado anormal vaginal, dolor en la zona pélvica, dolor abdominal e hinchazón; sin embargo, no siempre resulta claro reconocerlos. Hasta la fecha, no se conoce métodos de detección temprana que reduzcan las tasas de mortalidad (37–39).

2.2.3.3. Cáncer de útero

A nivel mundial, el cáncer de útero representa el sexto trastorno maligno más común en todo el mundo. En 2018, la Agencia Internacional de Investigación en Cáncer (IARC) a través del GLOBOCAN estimó

aproximadamente 382 069 (2.1%) nuevos casos con 89 929 (0.9%) muertes por año (4). Así también, en el Perú, el cáncer de cuerpo uterino tuvo 1 053 nuevos casos y 284 muertes por año en el 2018 (4). El riesgo acumulado de cáncer de endometrio hasta la edad de 75 años se ha estimado en un 1,6% para las regiones de ingresos altos y un 0,7% para los países de ingresos bajos, por lo que se le atribuye a las altas tasas de obesidad e inactividad física como factores de riesgo principales en los países de ingresos altos (40).

El útero , órgano responsable de las funciones asociadas a la reproducción, se encuentra formado por el fondo (parte más alta del útero) ,cuerpo uterino (parte superior y más grande del útero de forma plana en la superficie anterior y convexa en su superficie posterior), istmo (parte más estrecha) y cuello uterino (porción inferior del útero donde se inserta la vagina) (41).

Existen dos tipos principales de cáncer de útero:

- Cáncer de endometrio: Es el tipo de cáncer de útero más común y está fuertemente asociado con la obesidad, se forma en el revestimiento epitelial de la cavidad uterina llamado endometrio. El tratamiento estándar del cáncer de endometrio consiste en histerectomía (extirpación del útero total o parcialmente) y salpingooforectomía bilateral (extirpación de un ovario y el tubo de Falopio); sin embargo, nuevos medicamentos así como terapias dirigidas se están investigando para tratar estadios más avanzados (40)

- Sarcoma uterino: Es el tipo de cáncer de útero poco frecuente, surge en los tejidos estromales y musculares del miometrio (42).

Entre los signos y síntomas del cáncer de útero se incluyen la secreción o sangrado vaginal anormal, dificultades para orinar y dolor durante las relaciones sexuales (43). Actualmente no se conoce de forma exacta como prevenir el cáncer de útero; sin embargo, existen medidas que se pueden tomar para reducir el riesgo de padecer esta enfermedad: mantener un peso saludable, mantenerse físicamente activo, consumo de progesterona si está tomando estrógeno para tratar los síntomas de menopausia, recibir tratamiento para los trastornos precancerosos del endometrio, y acudir al médico si se presenta cáncer de colon hereditario sin poliposis (HNPCC) o síndrome Lynch (40,44). La mayoría de los casos de cáncer de útero se diagnostican en mujeres que están en la menopausia o la posmenopausia y si se detecta en etapas tempranas el tratamiento es más eficaz (43,44).

2.2.3.4. Cáncer de vulva y vagina

A nivel mundial, la Agencia Internacional de Investigación en Cáncer (IARC) a través del GLOBOCAN 2018 estimó aproximadamente 61 835 (0.3%) nuevos casos con 23 284 (0.3%) muertes por año debidos al cáncer de vulva y vagina (4). Así también, en el Perú, el cáncer de vulva y vagina tuvieron 388 nuevos casos y 84 muertes por año en el 2018 (4).

El cáncer de vulva y vagina son neoplasias ginecológicas pocos frecuentes y afectan principalmente a mujeres posmenopáusicas. Sin embargo , se está evidenciando presencia de éstos en mujeres jóvenes debido al aumento de

las infecciones por el virus del papiloma humano (VPH) en el mundo (45,46). Los signos y síntomas como secreción o sangrado vaginal anormal y desórdenes en los hábitos de ir al baño se presentan en etapas avanzadas del cáncer de vagina. En el cáncer de vulva, los signos y síntomas son el ardor dolor o sensibilidad en la vulva, cambios de coloración y presencia de ampollas, verrugas o úlceras en la vulva (44–46).

Entre los factores de riesgo se encuentran la infección persistente por VPH, haber tenido cáncer cervical, infección por VIH y fumar. La mejor manera de detectar estos tipos de cáncer en etapas tempranas es realizarse revisiones médicas con regularidad y pruebas de seguimiento (44–46).

2.2.4. Epigenética y Cáncer

La Epigenética inicialmente fue definida por Waddington en 1942 como “las interacciones causales entre los genes y sus productos, que llevan el fenotipo a ser”, es decir, sus fenotipos deben ser estudiadas para llegar a conclusiones sobre los genotipos (47). Actualmente es definida como “el estudio de todos los cambios heredables que se transmiten de forma estable en muchas rondas de divisiones celulares, pero que no alteran la secuencia de nucleótidos” (48). La importancia del principio epigenético nos explica como la identidad de las diferentes células del cuerpo cambia, no por el cambio de la secuencia de ADN, sino patrones de expresión génica mediada por modificaciones epigenéticas, que incluyen principalmente la metilación del ADN, modificaciones postraduccionales de las proteínas histonas y ARNs no codantes incluidos los miARN (20,34,47,48). La interrupción o desregulación del control epigenético

puede conducir a estados patológicos como el cáncer (49). Recientes descubrimientos en estas alteraciones han puesto atención al desarrollo de terapias epigenéticas utilizando inhibidores contra esta maquinaria, debido a que estos cambios son reversibles (17,31,50,51).

2.2.5. Principales mecanismos epigenéticos

La regulación epigenética afecta la expresión de unos 1000 genes en el humano y es crítica en el normal desarrollo y crecimiento de las células por lo que las alteraciones en los reguladores epigenéticos conducen a fenotipos patológicos incluido el cáncer de mama y próstata (52,53).

2.2.5.1. Metilación del ADN

La metilación del ADN es esencial para el desarrollo normal y se asocia con la impronta genómica (genes expresados de forma diferente dependiendo del progenitor), la inactivación del cromosoma X, la represión de elementos repetitivos, el envejecimiento y la carcinogénesis. La metilación del ADN consiste en la adición covalente de un grupo metilo (-CH₃) en la posición 5 del anillo de citosina en dinucleótidos CpG (citosinas que preceden a guaninas) y es catalizada por enzimas ADN metiltransferasas (DNMT). Cuando los grupos metilo se posicionan sobre el ADN, los nucleosomas (unidad fundamental de la cromatina) se unen. Si esta región compromete a un gen, este último deja de expresarse promoviendo el silenciamiento génico (Figura 2) (48,54,55).

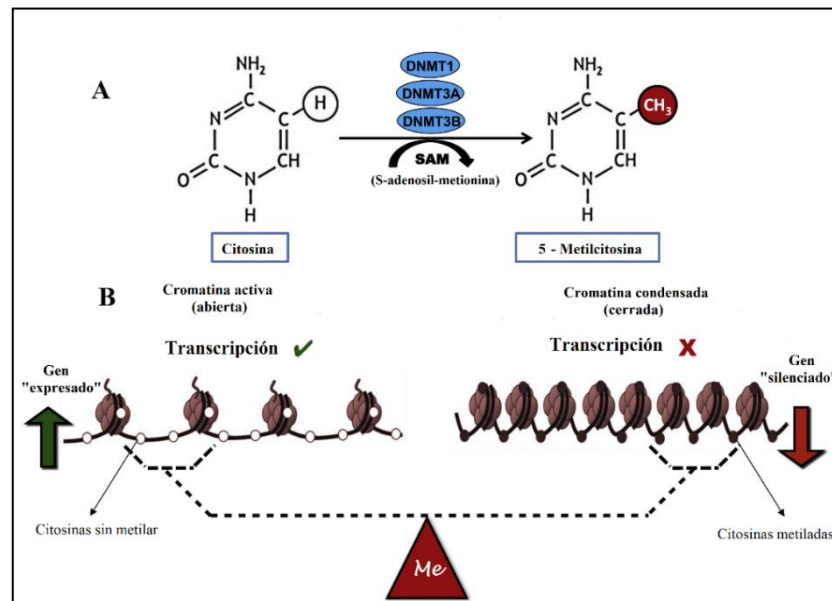


Figura 2. Metilación del ADN. (A) La metilación del ADN implica la transferencia del grupo metilo de la S-adenosilmetionina (SAM) a la posición C-5 de la citosina, catalizada por ADN metiltransferasas (DNMT). (B) La tasa de metilación del ADN es inversamente proporcional a la transcripción. El aumento en la cantidad de grupo metilo acompañado por modificaciones de las proteínas histonas centrales (como la acetilación y la fosforilación) da como resultado la alteración de la estructura de la cromatina de conformación abierta a cerrada, en cuyo caso el ADN es menos accesible para la maquinaria transcripcional y, por lo tanto, se impide la transcripción, lo que finalmente resulta en el silenciamiento del gen. Figura adaptada de Agrawal K et al., 2018 (177).

A lo largo del genoma, los dinucleótidos CpG están mayoritariamente en regiones denominadas Islas de CpG. En células normales, aproximadamente el 60% de las secuencias promotoras, localizadas en islas CpG, no están metiladas lo cual proporciona acceso a factores de transcripción y proteínas asociadas a la cromatina para la expresión génica (54).

Por el contrario, las células cancerosas presentan un metiloma distinto de las células normales. Por un lado, las células cancerosas presentan hipometilación a nivel del genoma total lo cual aporta inestabilidad genómica, pues conduce a la activación aberrante de genes y regiones no codificantes. Por otro lado, las células cancerosas presentan hipermetilación

de islas CpG en promotores de genes supresores de tumores, lo cual silencia su expresión y contribuye así a la tumorigénesis (20,48,56).

2.2.5.2. Reguladores de histonas

Las modificaciones de histonas influyen en la accesibilidad de la cromatina, la cromatina es una estructura altamente ordenada que consta de repeticiones de nucleosomas que consiste en un fragmento de ADN enrollado alrededor de un octámero de histonas, que a su vez están formados por dímeros de H2A, H2B, H3 y H4 (20,48,54).

Son modificaciones covalentes postraduccionales en las colas N-terminal y C-terminal de las histonas que incluyen principalmente metilación, acetilación, ubiquitinación, biotinilación y fosforilación (Figura 3). Dichas modificaciones son importantes en la regulación de procesos celulares como transcripción, replicación y reparación. Todas las modificaciones de histonas descritas anteriormente se consideran marcas epigenéticas hereditarias pues se mantienen con alta fidelidad a través de la división celular. Las diferentes modificaciones de las histonas que se presentan en las colas de éstas, generan un "código de histonas" leído por diferentes mecanismos celulares que dictan diferentes resultados (48). Las modificaciones de histonas son reversibles y están controladas por un grupo de enzimas como las histonas acetiltransferasas (HATs) y desacetilasas (HDACs), metiltransferasas (HMTs) y desmetilasas (HDMs), quinasas, fosfatasas, ubiquitina ligasas y deubiquitinasas, SUMO ligasas y proteasas (48,57). Las interacciones entre las enzimas

reguladores de histonas y el mecanismo de metilación del ADN regulan la expresión génica, la organización de la cromatina y la identidad celular (58).

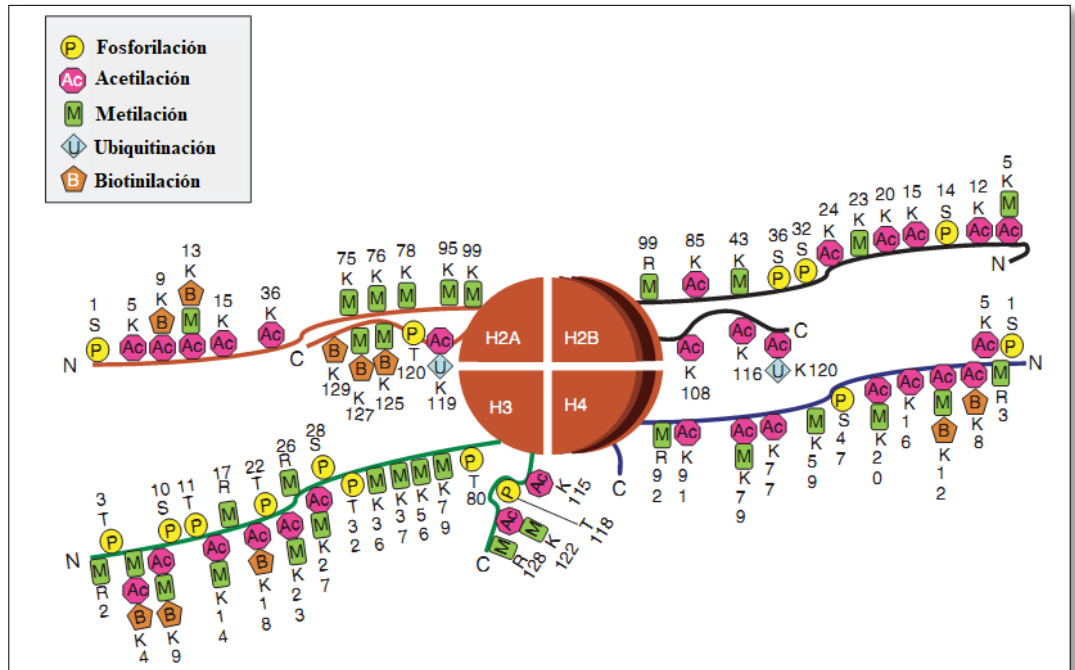


Figura 3. Representación esquemática de las principales modificaciones en las colas de las histonas. Se representan diferentes aminoácidos que constituyen colas de histonas junto con las diferentes modificaciones covalentes específicas de cada residuo. Figura adaptada de Tollefsbol T, 2009 (178).

Los estudios genómicos sugieren que el "código de histonas" se encuentra alterado en el cáncer (48). Así se han identificado mutaciones en los genes que codifican las enzimas modificadoras de histonas lo que afecta tanto las marcas represivas o activas en las mismas. El subsecuente efecto de estos cambios sería el silenciamiento de genes supresores de tumor y la activación de los oncogenes (59,60). Por ejemplo, la activación aberrante de histonas acetiltransferasas y metiltransferasas puede originar un aumento de la acetilación de lisinas -como H3K9, H3K18, H4K8 y H4K12- y un aumento de la metilación de lisinas -como H3K4 y K3K36- que fomentan la expresión de oncogenes (48,61,62).

2.2.6. Reguladores epigenéticos: Escritores, lectores y borradores

Los actores claves que realizan la modificación química tanto en el ADN como en las histonas son denominados reguladores epigenéticos y se clasifican en escritores, lectores y borradores (Figura 4) (19).

2.2.6.1. Los escritores

Cumplen la función de introducir modificaciones químicas en el ADN y las histonas. Dos alteraciones epigenéticas, la metilación y acetilación gobiernan frecuentemente el patrón de expresión génica en la célula al alterar entre la activación o la represión transcripcional. Aquí se incluyen los ADN metiltransferasas, histonas lisina metiltransferasas, proteínas arginina metiltransferasas e histonas acetiltransferasas (63,64).

2.2.6.2. Los lectores

Son proteínas con dominios específicos que identifican e interpretan modificaciones químicas específicas (19). Existen una amplia variedad de dominios de proteínas que han evolucionado para unirse a estas modificaciones, entre estas se encuentran:

- Las proteínas de unión a metil-CpG (MBP), que se unen al ADN metilado e inician el silenciamiento génico a través del reclutamiento de otros factores.
- Los lectores de acetilación de histonas: bromodominio BRD, doble dedo PHD y dominios de YEATS.

- Los lectores de metilación de histonas: dominio Tudor, dominio Tudor en tándem (TTD), cromodominio, cromobarrel, cromodominio doble (DCD), dominio del tumor cerebral maligno (MBT), dominio PWWP (Pro-Trp-Trp-Pro), homeodominio de planta (PHD), repeticiones WD40, dominio ATRX-DNMT3-DNMT3L (ADD), anquirina, homología bromo-adyacente (BAH) y dominio de dedo de zinc CW (zf-CW).

2.2.6.3. Los borradores

Son un grupo de enzimas que se oponen a la actividad de los escritores, eliminando marcas epigenéticas de residuos específicos. Estas enzimas aseguran que la célula tenga la capacidad de modificar rápidamente su cromatina y así modular los estados de expresión génica (19,65). Aquí se encuentran a las enzimas desmetilasas del ADN, histonas desacetilasas (HDAC), las desmetilasas de histonas (HDMTs / KDM), las fosfatasas y las desubiquitinasas (19,63,66).

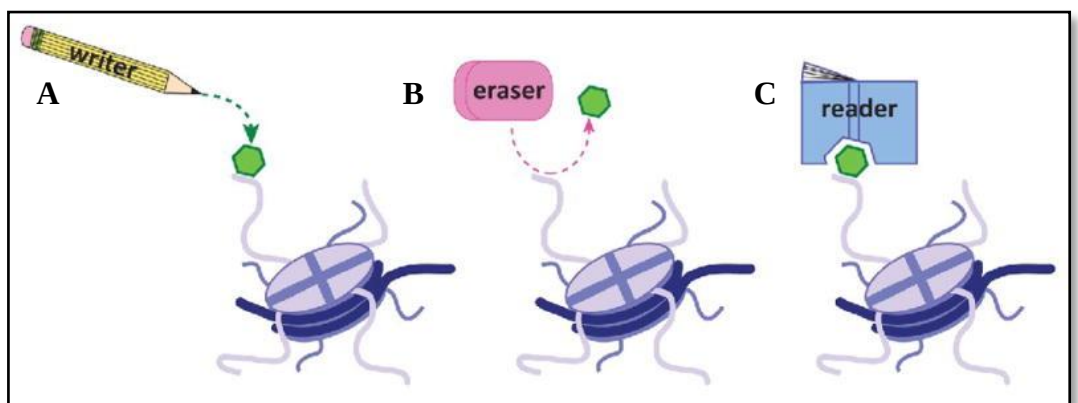


Figura 4. Reguladores epigenéticos de la cromatina. (A) Los escritores (writers) quienes colocan modificaciones postraduccionales en las proteínas histonas. (B) Los borradores (erasers) quienes eliminan dichas modificaciones de las proteínas histonas. (C) Los lectores (readers) quienes interpretan estas modificaciones covalentes (derecha) para mediar diversos procesos posteriores. Figura adaptada de Gardner KE, Allis CD y Strahl BD , 2012 (179).

2.2.7. Bioinformática aplicada al estudio del cáncer

La bioinformática es un campo interdisciplinario que desarrolla y aplica tecnologías computacionales para estudiar y responder preguntas biológicas y biomédicas. Involucra todos los aspectos correspondientes a la adquisición, procesamiento, distribución, análisis, interpretación e integración de datos (67).

La bioinformática como nueva y creciente área ha facilitado el manejo de la vasta cantidad de datos producidos por el Proyecto del Genoma Humano y las disciplinas ómicas, como la genómica, la transcriptómica, la proteómica y la metabolómica. Esto ha permitido a que los investigadores guíen de manera más eficiente y rigurosa sus estudios para llegar a conclusiones biológicas válidas (68).

Uno de los campos de estudio con mayor impacto y aplicación es la bioinformática traslacional, la cual traduce de manera eficiente datos básicos moleculares, genéticos, celulares y clínicos en productos clínicos o implicaciones para la salud. Es decir, es el desarrollo de métodos informáticos para optimizar la transformación de un gran volumen de datos biológicos a entidades clínicas (69).

El cáncer ha sido una de las áreas más activa y con mayor alcance de la bioinformática traslacional. Las tecnologías genómicas y proteómicas son bastante prometedoras para identificar nuevos biomarcadores o inhibidores de reguladores epigenéticos como objetivos para el desarrollo de fármacos contra el cáncer. Por ejemplo, diversos estudios *in silico* han identificado potenciales biomarcadores pronósticos y dianas terapéuticas tanto en el cáncer cervical (TCGA) (70–72), como en el cáncer de endometrio (73–75) y en el de ovario

(76–79). Dichos estudios han minado y analizado la base de genomas completos denominada “El Atlas del Genoma del Cáncer”, proporcionando información valiosa sobre el genoma en condiciones de la enfermedad.

2.3. PREGUNTA DE INVESTIGACIÓN

¿Cómo se encuentran alterados los reguladores epigenéticos en el cáncer ginecológico y cuáles son reconocidos como posibles biomarcadores epigenéticos de diagnóstico y pronóstico?

2.4. JUSTIFICACIÓN

El cáncer ginecológico es un problema de salud pública que afecta principalmente a los países en desarrollo según los informes de la Organización Mundial de la Salud (OMS) y la Agencia Internacional para la Investigación del cáncer. Por ello, la prevención y la reducción de riesgos son las formas en las que los países están llevando a cabo su control.

Mientras se resuelven estas estrategias, el área de la biología molecular ha dirigido sus esfuerzos para enfrentar esta enfermedad, con esto una creciente generación de información genómica del cáncer está permitiendo su disección a nivel computacional. La disponibilidad inmediata de los datos genómicos, permite el minado completo de esta información. Conociendo la creciente atención por el estudio de mecanismos epigenéticos y su potencial rol en la carcinogénesis, se hace imperativo determinar la existencia de anormalidades en dichos mecanismos en el cáncer ginecológico. El uso de herramientas bioinformáticas permite extraer y analizar datos ómicos de manera estructurada. El análisis estadístico subsecuente hace posible estimar asociaciones de patrones genómicos con caracteres fenotípicos. Por consiguiente, la importancia del presente estudio radica en analizar e integrar los datos genómicos en conjunto con las características clínicas disponibles para mejorar la comprensión de la biología de los cánceres ginecológicos. Dicho conocimiento, de importancia a largo plazo, contribuirá a reconocer marcadores de la enfermedad.

III. OBJETIVOS

3.1. Objetivo general

- Identificar las alteraciones en las enzimas modificadoras del ADN y enzimas modificadoras de histonas y determinar su asociación con los rasgos fenotípicos del cáncer ginecológico.

3.2. Objetivos específicos

- Identificar alteraciones en la expresión génica y variaciones de secuencia presentes en las enzimas de estudio, así como conocer el estado de metilación del ADN en promotores de genes de estudio.
- Predecir los cambios en la estructura y función de las proteínas de potencial interés.
- Asociar las variables genéticas con variables fenotípicas.

IV. METODOLOGÍA

4.1. DISEÑO DEL ESTUDIO

El diseño del presente estudio es del tipo descriptivo, observacional retrospectivo. Es un estudio de análisis secundario de datos genómicos y fenotípicos de pacientes mujeres con cáncer ginecológico. La información ha sido obtenida de la base de datos del Atlas del Genoma del Cáncer (TCGA).

4.2. POBLACIÓN

Pacientes mujeres, mayores de 18 años, diagnosticadas con un tipo de cáncer ginecológico incluido ovario, útero y cérvix. La población de estudio fue reclutada, a nivel mundial, en los respectivos Proyectos de Genoma de Cáncer y su información está disponible en la base de datos del Atlas del Genoma del Cáncer (TCGA).

Los sujetos control fueron los mismos que participaron en los proyectos TCGA, debido a que de ellas se colectó muestras sanguíneas y tejido normal como control. Sin embargo, también se consideraron como controles a personas no enfermas que donaron el tejido *post mortem*, esto para el caso del análisis de expresión diferencial utilizando el recurso GEPIA.

4.3. CRITERIOS DE SELECCIÓN

Criterios de inclusión

- Mujeres diagnosticadas con algún tipo de cáncer ginecológico (cérvix, ovario y endometrio).

- Mujeres cuya información genómica y clínica está disponible en la base de datos de TCGA.

Criterios de exclusión

- Mujeres que presenten información genómica y clínica incompleta.

4.4. OPERACIONALIZACIÓN DE VARIABLES

Cuadro 1. Operacionalización de variables del estudio descriptivo

Variable	Definición	Tipo de variable según su naturaleza	Tipo de variable según su relación	Escala de medición	Indicador	Definición operacional
Tipo de Cáncer ginecológico	Cáncer que se origina en los órganos reproductores de la mujer.	Cualitativa	Dependiente	Nominal	C, O, E	C: Cérvix O: Ovario E: Endometrio
Mecanismo epigenético	Mecanismos capaces de modificar la expresión de los genes.	Cualitativa	Independiente	Nominal	1,2	1: Metilación del ADN 2: Modificación de Histonas
Polimorfismo de un solo nucleótido SNPs	Variación en la secuencia de ADN que afecta a una sola base (adenina (A), timina (T), citosina (C) o guanina (G) de una secuencia del genoma.	Cualitativa	Independiente	Nominal	0,1	0: Ausente 1: Presente
Sobrevida global a 5 años	Tiempo en el que el paciente estuvo afectado por la enfermedad	Cuantitativa	Co-variable	Razón	Tiempo en meses	Meses
Sobrevida libre de recurrencia	Tiempo que se mide después del tratamiento en el que un grupo de personas en un ensayo clínico no ha tenido recurrencia del cáncer o este no ha empeorado.	Cuantitativa	Co-variable	Razón	Tiempo en meses	Meses

Variable	Definición	Tipo de variable según su naturaleza	Tipo de variable según su relación	Escala de medición	Indicador	Definición operacional
Estadio del tumor	Se refiere a la extensión de su cáncer, es decir, qué tan grande es el tumor y si se ha extendido.	Cualitativa	Co-variable	Ordinal	T, N, M	T: tamaño y extensión del tumor principal. El tumor principal se llama de ordinario el tumor primario. N: se refiere a la extensión de cáncer que se ha diseminado a los ganglios (o nódulos) linfáticos cercanos. M: se refiere a si el cáncer se ha metastatizado; es decir, si ha tenido metástasis. Esto significa
Categoría de etnia	Grupos humanos que se subdividen sobre la base de rasgos fenotípicos, a partir de una serie de características que se transmiten por herencia genética.	Cualitativa	Confusor	Nominal	1,2,3,4	1: Blanco 2: Asiático 3: Negro o africanoamericano 4: Americano indio o alaska nativo
Tipo de virus del papiloma humano	Grupo de virus que en su mayoría se transmiten sexualmente y que pueden causar verrugas o cáncer.	Cualitativa	Co-variable	Nominal	0,1,2	1: VPH 16 2: VPH 18 0: Otros
Edad al momento del diagnóstico	Años cumplidos al momento del diagnóstico de cáncer	Cuantitativa	Confusor	Razón	Edad en Años	Años
Peso del paciente	Cantidad de masa corporal al momento del diagnóstico de cáncer	Cuantitativa	Confusor	Razón	Peso en Kg	Kilogramos

4.5. INSTRUMENTOS

Los recursos bioinformáticos que fueron utilizados durante el estudio se describen en la Tabla 1.

Tabla 1. Bases de datos y herramientas bioinformáticas especializadas

Recurso	Descripción	Sitio Web	Fecha de consulta
Base de datos			
WERAM	Base de datos de acetilación y metilación de histonas en 148 especies eucariotas.	http://weram.biocuckoo.org/	2 al 15 marzo 2018
HIstome	Base de datos especializada en proteínas histonas humanas y sus variantes, enzimas modificadoras de histonas y sitios de modificaciones postraduccionales.	http://www.actrec.gov.in/histome/	16 al 20 marzo 2018
CbioPortal	Recurso web para explorar, visualizar y analizar datos multidimensionales de genómica del cáncer.	http://www.cbioportal.org/	17 abril al 30 mayo 2018 y abril 2019
Catálogo GWAS	Colección de asociaciones significativas entre loci genéticos y enfermedades o rasgos comunes a partir de estudios GWAS publicados.	www.ebi.ac.uk/gwas/	13 al 26 junio 2018
Broad GDAC Firehose	Portal web para realizar análisis automatizados con datos TCGA, proporcionan datos pre-procesados y análisis de asociación, incluidos los clínicos.	http://gdac.broadinstitute.org/	Octubre a noviembre 2018
UNIPROT	Repositorio universal de proteínas creado por la combinación de Swiss-Prot, TrEMBL y PIRt.	https://www.uniprot.org/	Marzo a noviembre 2018
STRING	Base de datos biológico y recurso web de interacciones proteína-proteína conocidas y previstas para varios organismos.	http://string-db.org/	Diciembre 2018
HGNC	Repositorio web de abreviaturas/símbolos de los genes aprobados por HGNC (Comité de Nomenclatura de Genes de HUGO), además proporciona enlaces a información genómica, proteómica y fenotípica.	www.genenames.org	Abril y mayo 2018
GEPIA	Servidor web para el cáncer, el análisis de perfiles de expresión génica normal y análisis interactivos.	http://gepia.cancer-pku.cn/	Octubre 2018
XENA Browser	Herramienta de exploración e integración visual para datos multi-ómicos, anotaciones clínicas y fenotípicas asociadas.	https://xenabrowser.net/	Marzo y abril 2019
Recursos de análisis epigenómico			
MethHC	Base de datos de metilación de ADN y perfiles de expresión de ARNm/microARN de TCGA	http://MethHC.mbc.nctu.edu.tw	16 julio al 30 agosto 2018

Recurso	Descripción	Sitio Web	Fecha de consulta
Recursos de análisis epigenómico			
MEXPRESS	Recurso web para la visualización de la expresión génica, metilación del ADN y datos clínicos de TCGA	http://mexpress.be	27 julio al 30 agosto 2018
Recursos de análisis de estructura/función de proteínas.			
Biomuta v4.0	Base de datos de variaciones de un solo nucleótido (SNV) asociadas a cáncer.	https://hive.biochemistry.gwu.edu/biomuta	6 al 19 setiembre 2018
I-Mutant v3.0	Predictor de efectos de mutación proteica puntual	http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi	Setiembre 2018
NetSurfP-1.1	Herramienta web que predice la accesibilidad a la superficie y la estructura secundaria de los aminoácidos en una secuencia de aminoácidos.	http://www.cbs.dtu.dk/services/NetSurfP/	Setiembre 2018
Herramientas de enriquecimiento de anotaciones genéticas			
Enrichr	Recurso web para el análisis de enriquecimiento ontológico de una gran biblioteca de conjunto de genes curados.	http://amp.pharm.mssm.edu/Enrichr/	Noviembre 2018
GeneCodis	Es una aplicación web para el análisis funcional de listas de genes que integra diferentes fuentes de información y encuentra patrones modulares de anotaciones interrelacionadas.	http://genecodis.cnb.csic.es/	Diciembre 2018 y abril 2019
Herramientas de visualización			
UCSF Chimera	Visualizador 3D de estructuras de proteínas y análisis de estructuras moleculares, permite la visualización de simulaciones de dinámica molecular y docking molecular, alineamientos estructurales, alineamientos secuenciales, entre otros.	https://www.cgl.ucsf.edu/chimera/	Noviembre 2018
VENNY 2.1	Herramienta que representa las relaciones de intersección, inclusión y disyunción mediante un diagrama de Venn.	http://bioinfogp.cnb.csic.es/tools/venny/	Abril 2019
Programas bioinformáticos para la visualización y análisis de redes de PPI			
Cytoscape 3.7.1	Programa de código abierto para el análisis y visualización de redes de PPI además de integrar visualmente la red con perfiles de expresión, fenotipos y otros estados moleculares y vincular la red a bases de datos públicas.	https://cytoscape.org/	Mayo 2019
CytoHubba 0.1	Aplicativo complementario de Cytoscape para explorar nodos importantes en redes biológicas mediante 12 métodos de análisis topológicos disponibles.	http://apps.cytoscape.org/apps/cytohubba	Mayo 2019

4.6. PROCEDIMIENTO

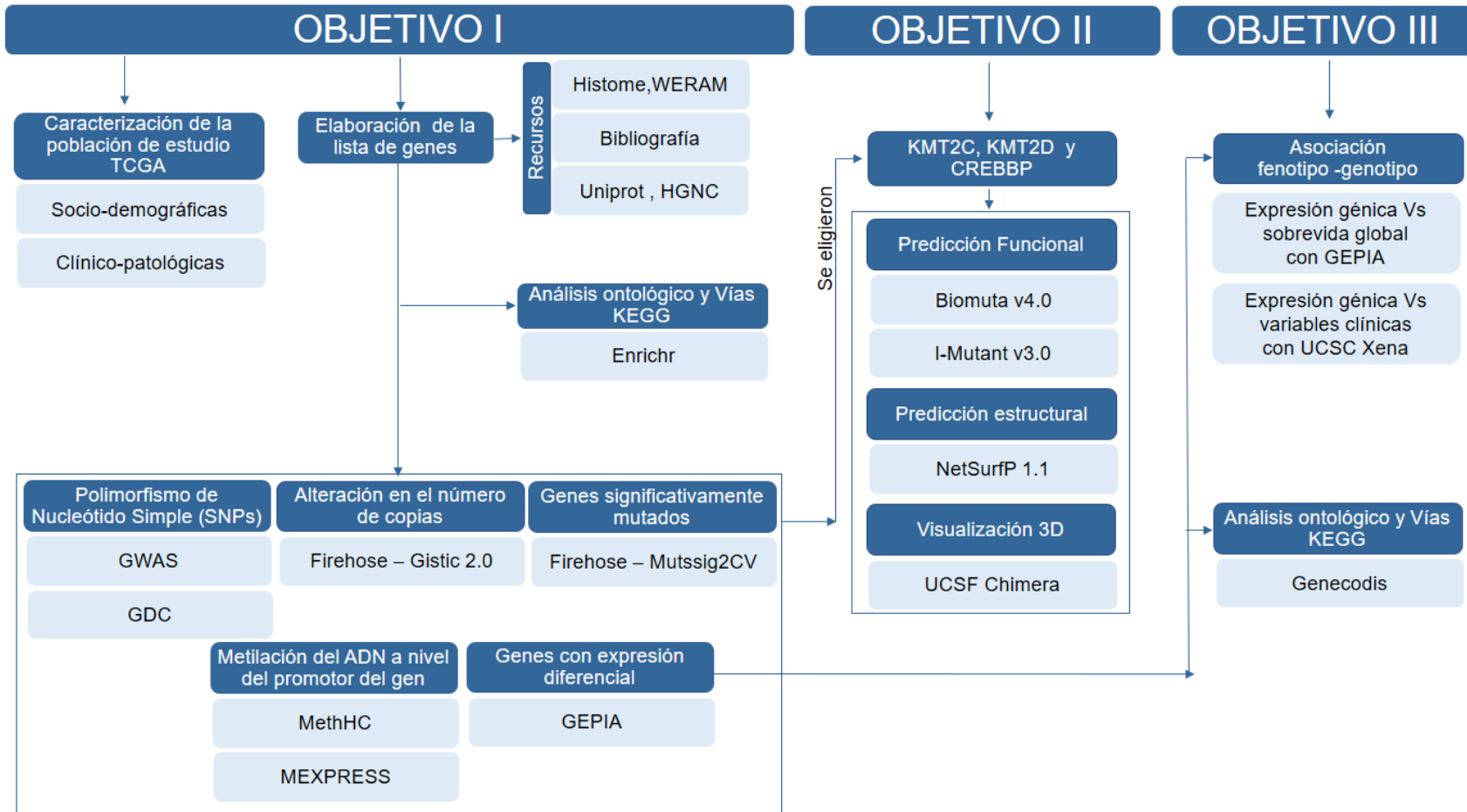


Figura 5. Flujograma general del trabajo realizado

4.6.1. Elaboración de la lista de genes de estudio

La lista de genes de estudio, cuyos productos proteicos están implicados en mecanismos epigenéticos (modificación de histonas y metilación del ADN), se elaboró consultando las bases de datos WERAM (80) e HIstome (81), así como también artículos científicos (82,83).

La base de datos WERAM fue utilizada para extraer información de enzimas implicadas en la acetilación y metilación de histonas, mientras que la base de datos HIstome fue utilizado para obtener información de enzimas implicadas en otras modificaciones de histonas (fosforilación, ubiquitinación, biotinilación, citrulinación, ribosilación, además de acetilación y metilación).

Una vez recopilada la información de interés, se revisó la clasificación y la nomenclatura de los genes de ambas fuentes, a fin tabular la información de manera uniforme y ordenada. Finalmente, se obtuvo una lista de reguladores de histonas, constituida por 152 genes.

Esta lista se amplió agregando ocho genes codificantes de enzimas metiladoras de ADN, procedentes de la revisión de Li & Zhang 2014, donde describe el importante rol de las enzimas metiladoras de ADN en mamíferos (83) y de Lyko F. 2018, donde describe a la familia de las enzimas ADN metiltransferasas (82). Posteriormente, los nombres de los genes y sus productos proteicos se obtuvieron de Uniprot (84), la localización génica en los cromosomas se obtuvo de la base de datos HIstome y el repositorio de nomenclatura de genes HGNC (85) (Figura 5 y 6). A continuación, se clasificaron los genes de estudio en dos categorías:

- *Borradores*: genes que codifican proteínas que eliminan modificaciones químicas tanto en el ADN como en las histonas (63).
- *Escritores*, genes que codifican proteínas que agregan grupos funcionales tanto al ADN como a las histonas, es decir, poseen actividad opuesta a los borradores (63).

Algunos de los genes *borradores* y *escritores* también fueron identificados como *lectores*, genes que identifican e interpretan modificaciones químicas específicas tanto en el ADN como en las histonas (86). El proceso aplicado para elaborar la lista de genes de estudio se detalla en la Figura 6.

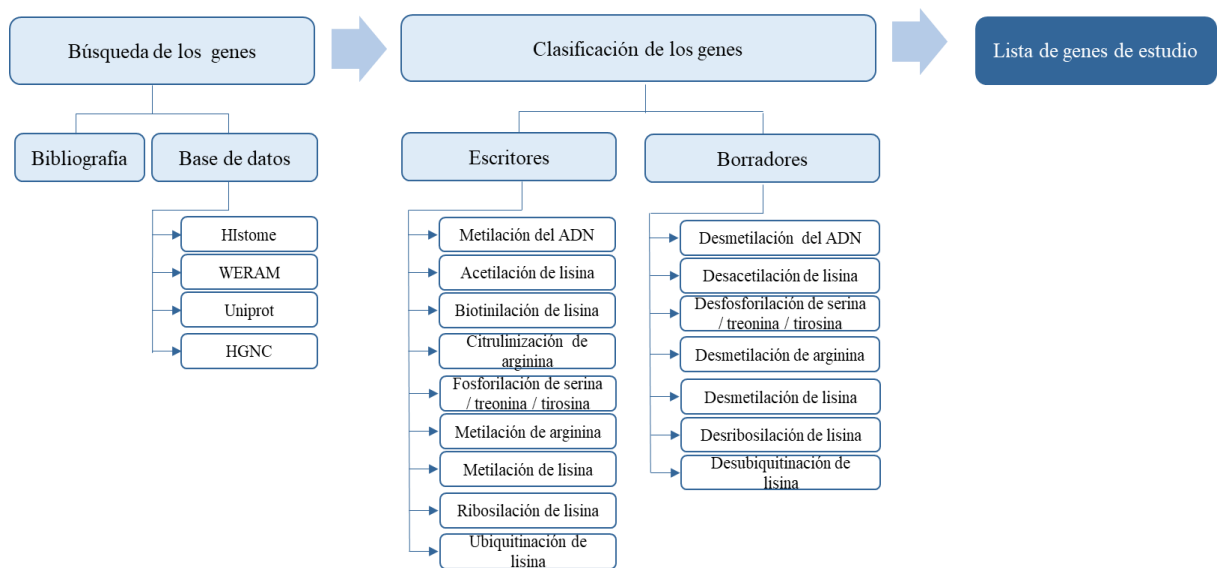


Figura 6. Flujograma del proceso aplicado para la elaboración de la lista de genes de estudio.

4.6.2. Análisis de anotación funcional y análisis de enriquecimiento de rutas bioquímicas de los genes de interés

Para conocer las funciones biológicas asociadas con las proteínas de interés, se condujo un análisis de enriquecimiento funcional. Para ello se utilizó la herramienta Enrichr (87,88), que determinó tanto la ontología génica (GO) como las vías de señalización KEGG en los 160 genes de estudio. Se ingresó la lista de los 160 genes en la barra de búsqueda de Enrichr y luego se procedió con el análisis en línea. Una vez finalizado el análisis, se obtuvieron los valores de p mediante la prueba exacta de Fisher y los resultados con valor $p < 0.01$ fueron considerados significativos.

El análisis de enriquecimiento funcional consideró sólo la categoría de proceso biológico (BP), para su visualización, los 20 principales BPs fueron tabulados y de éstos 10 BPs fueron representados en barras y clustergramas. Enrichr también retornó los resultados de las rutas bioquímicas considerando la base de datos KEGG (89). Los resultados con un valor $p < 0.01$ fueron considerados significativos.

4.6.3. Bases de datos de genomas y de la población de estudio

Se recolectó información de genomas y fenotipos de cuatro tipos de cáncer ginecológico provenientes de El Atlas del Genoma del Cáncer (TCGA) a través del portal de datos CbioPortal (www.cbioportal.org) (90,91). Los nombres de los proyectos incluyen “Cáncer cervical de células escamosas y adenocarcinoma endocervical (TCGA, provisional)”, “Cistoadenocarcinoma seroso ovárico (TCGA, provisional)”, “Carcinoma de endometrio del cuerpo uterino (TCGA,

provisional)” y “Carcinoma uterino (TCGA, provisional)”; las abreviaturas correspondientes para cada estudio fueron CESC, OV, UCEC y UCS, respectivamente.

Las cohortes denominadas como “provisionales, TCGA” contienen datos actualizados y depositados en fecha posterior a la publicación. Estos conjuntos de datos incluyen datos ómicos, demográficos y clinicopatológicos de cada participante. Los datos ómicos disponibles incluyen mutaciones genéticas, alteraciones del número de copias (CNA), expresión de ARNm y proteínas.

Los conjuntos de datos analizados durante el estudio se encuentran disponibles en https://www.cbioportal.org/study/summary?id=cesc_tcg (para CESC), https://www.cbioportal.org/study/summary?id=ov_tcg (para OV), https://www.cbioportal.org/study/summary?id=ucec_tcg (para UCEC) y https://www.cbioportal.org/study/summary?id=ucs_tcg (para UCS).

El procedimiento y los pasos realizados para acceder y descargar los datos ómicos, es descrito en la Figura 7.

QUERY DOWNLOAD DATA

Select Studies: 1 studies selected (548 samples) Deselect all View summary Search...

Pancreas 10
Peripheral Nervous System 5
Pleura 3
Prostate 16
Skin 12
Soft Tissue 6
Testis 3
Thymus 3
Thyroid 4
Uterus 8

Select all listed studies matching filter (8)

Uterus

Endometrial Carcinoma

Endometrial Cancer (MSK, 2018) 197 samples
Uterine Corpus Endometrial Carcinoma (TCGA, Nature 2013) 500 samples
Uterine Corpus Endometrial Carcinoma (TCGA, PanCancer Atlas) 529 samples
Uterine Corpus Endometrial Carcinoma (TCGA, Provisional) 548 samples

UTERINE CARCINOSARCOMA/UTERINE MALIGNANT MIXED MULLERIAN TUMOR

Uterine Carcinosarcoma (Johns Hopkins University, Nat Commun 2014) 22 samples
Uterine Carcinosarcoma (TCGA, PanCancer Atlas) 57 samples
Uterine Carcinosarcoma (TCGA, Provisional) 57 samples

UTERINE CLEAR CELL CARCINOMA

Uterine Clear Cell Carcinoma (NIH, Cancer 2017) 16 samples

Select Genomic Profiles:

☒ Mutations
☐ Putative copy-number alterations from GISTIC
☐ mRNA Expression. Select one of the profiles below:
 ☐ mRNA Expression Z-Scores (microarray)
 ☐ mRNA Expression Z-Scores (RNA Seq V2 RSEM)
☐ Protein expression Z-scores (RPPA)

Select Patient/Case Set: To build your own case set, try out our enhanced Study View.

All Tumors (548)

Enter Genes: Advanced: Onco Query Language (OQL)

User-defined List

Select from Recurrently Mutated Genes (MutSig) Select Genes from Recurrent CNAs (Gistic)

ASH1L
ATM
ATR
AURKB
BAP1

All gene symbols are valid.

Submit Query

Figura 7. Pasos para realizar la consulta en Cbioportal. (1) Seleccionar un estudio de cáncer ginecológico (CESC, OV, UCEC o UCS). (2) Seleccionar el perfil genómico (mutaciones somáticas, alteraciones en el número de copias de ADN, expresión de ARNm o expresión de proteínas). (3) Definir el conjunto de pacientes / casos. (4) Ingresar la lista de genes de interés (n=160). (5) Consultar y luego descargar los datos en formato de texto para su debido almacenamiento.

Para obtener los datos clínico-patológicos, primero se seleccionaron los proyectos provisionales TCGA y luego se seleccionó la pestaña “datos clínicos” para finalmente proceder con la correspondiente descarga del archivo. Sobre los datos demográficos (edad, raza y etnia) y clínico-patológicos (IMC, estadio del tumor, menopausia, estado OS y DFS) descargados de Cbioportal, se aplicaron la prueba de Dunnett, la prueba exacta de Fisher y la prueba de Chi2, con la finalidad de describir el perfil general de los individuos de estudio para cada

neoplasia ginecológica. La descripción de las variables cualitativas se realizó por frecuencia y porcentaje. Los datos fueron analizados mediante el paquete estadístico STATA versión 15.

Por otro lado, para obtener información de las características clínicas y demográficas de mujeres con cáncer ginecológico en el Perú, se buscaron estudios y reportes procedentes del Instituto Nacional de Enfermedades Neoplásicas (INEN), institución peruana con más producción científica sobre cáncer en el Perú (92).

De esta manera, la búsqueda se realizó en la plataforma web del INEN (<https://portal.inen.sld.pe/articulos-especializados/>) así como también en la barra de búsqueda de PUBMED con los siguientes términos MESH: «instituto nacional de enfermedades neoplásicas » , « cáncer cervical » ,« cáncer de cérvix», « cáncer de cuello uterino » , « cáncer de ovario», « cáncer de endometrio » , « cáncer de útero » y « Perú ». Sólo fueron seleccionados aquellos estudios relacionados con cáncer ginecológico en los órganos mencionados. Finalmente, los estudios que incluían información clínica-patológica o demográfica de pacientes con cáncer ginecológico, fueron reportados.

4.6.4. Identificación de los genes significativamente mutados

Para identificar los genes significativamente mutados (SMGs) en cáncer ginecológico, se utilizó el portal web Broad GDAC Firehose, el cual contiene herramientas de análisis genómico y clínico de los datos TCGA.

Se utilizó la herramienta MutSig2CV v3.1 implementada en Firehose para el análisis individual de los cuatro tipos de cáncer ginecológico. MutSig2CV es un

algoritmo que analiza listas de mutaciones encontradas en la secuenciación del genoma e identifica genes que mutaron con más frecuencia de lo esperado por casualidad (93). Una vez obtenidos los resultados, estos fueron descargados y almacenados.

Se utilizó la función `merge()` de Rstudio 1.1.463 para extraer los resultados correspondientes a los 160 genes de estudio. La función `merge()` combinó la información obtenida previamente con otro archivo que contenía la lista de los genes de estudio, esta unión devolvió un documento con resultados que correspondía a los genes que fueron encontrados. Para la identificación de los SMGs, los valores $p < 0.05$ y $q < 0.05$ fueron considerados significativos.

4.6.5. Visualización 3D de las proteínas CREBBP, KMT2C y KMT2D nativas y mutadas y construcción de red de PPI

Con la finalidad de reconocer visualmente las regiones afectadas por las mutaciones puntuales, se utilizó UCSF Chimera (<https://www.cgl.ucsf.edu/chimera/>). Las estructuras 3D de las proteínas CREBBP, KMT2C y KMT2D fueron obtenidas del Banco de datos de proteínas PDB-RCSB (<http://www.rcsb.org/>). Se eligió sólo una estructura por proteína considerando dos criterios: i) resolución del cristal inferior a 2 Å (indicativo de una buena calidad de la estructura resuelta) y ii) que la estructura completa del dominio proteico de interés estuviese resuelta.

Una vez obtenidas las estructuras de las proteínas de interés, las mismas se cargaron en el programa UCSF Chimera.

Con ayuda de las herramientas presentes en el menú de UCSF Chimera, se localizaron y etiquetaron las variantes aminoacídicas “con pérdida de sentido” en las estructuras 3D de las proteínas. Para eso, se siguieron los pasos de la guía del portal educativo de PDB (<https://pdb101.rcsb.org/learn/videos/visualizing-pdb-structures-with-ucsf-chimera-for-beginners>). Finalmente, las imágenes fueron almacenados en formato PNG.

Con la finalidad de reconstruir la red de PPI de las 3 proteínas estudiadas, se utilizó la base de datos STRING v10.5 (The Search Tool for the Retrieval of Interacting Genes) (86). Las interacciones de STRING incluyen asociaciones físicas (directas) y funcionales (indirectas) derivadas de distintas fuentes incluidas experimentos de alto rendimiento, coexpresión y conocimientos previos de la literatura (86).

Se ingresaron los nombres de cada proteína y se filtraron las asociaciones con puntuación de confianza mayor a 0.4. En la red de PPI, cada nodo representó a un gen y los bordes representaron interacciones entre nodos, el grosor de cada borde indicó la fuerza del soporte de datos. Además, se representó el término de ontología de genes (función molecular) significativamente enriquecidas en la red de PPI, los 3 primeros términos de ontología génica fueron empleados para fines de visualización. Finalmente, los resultados fueron descargados y almacenados.

4.6.6. Predicción del efecto de las sustituciones de aminoácidos en la función de las proteínas

Con el fin de conocer el efecto funcional de las variaciones de aminoácidos provocadas por mutaciones sin sentido, se consultó la base de datos Biomuta

v4.0 (94). Este recurso valida información de variaciones de nucleótido único (SNVs) de múltiples fuentes asociadas con el cáncer para un gen consultado, y a través de la herramienta Polyphen se predice el impacto de la función de una proteína en función del cambio en el aminoácido. La herramienta Polyphen determina si una mutación es benigna, posiblemente dañina o probablemente dañina, según el clasificador Bayesiano Ingenuo (clasificador probabilístico fundamentado en el teorema de Bayes) (95).

La predicción funcional determina si un cambio de aminoácido tiene un efecto dañino o no para la función proteica. En el presente estudio se eligieron tres proteínas (CREBBP, KMT2C y KMT2D) para la predicción funcional, basándonos en dos criterios: (i) los genes que codifican estas proteínas, presentaron mayor frecuencia de mutaciones con pérdida de sentido y (ii) estuvieron presentes en más de un tipo de cáncer ginecológico.

Se ingresó al portal BioMuta v4.0- HIVE y en la barra de búsqueda se ingresaron los nombres CREBBP, KMT2C y KMT2D de manera individual. Después de enviar el término de búsqueda, se hizo clic en el acceso UniProtKB para navegar a la página de resultados. Tras la búsqueda, se descargaron todos los resultados en formato CSV. Luego con la función `filter()` del programa Rstudio 1.1.463 se seleccionaron sólo las variantes aminoacídicas presentes en los proyectos de cáncer ginecológico de TCGA. Las sustituciones predichas como “posiblemente dañina” o “probablemente dañina” fueron seleccionadas para su posterior análisis. Las variantes predichas como “benignas” no fueron consideradas en la predicción de cambios en la estabilidad energética y la accesibilidad a los solventes de las proteínas.

4.6.7. Estimación del cambio de la estabilidad energética teórica de las proteínas

Para conocer los cambios de estabilidad energética de las proteínas CREBBP, KMT2C, KMT2D, se utilizó el programa I-Mutant v3.0. Esta herramienta predice los cambios de la energía libre de Gibbs ($\Delta\Delta G$, en kcal / mol) tras una mutación puntual, y predice si la mutación de la proteína afecta la estabilidad energética de la proteína (96).

La predicción teórica que realiza I-Mutant se basa en un conjunto de algoritmos de aprendizaje supervisado (máquinas de soporte vectorial). Dicha predicción se puede evaluar a partir de la estructura de la proteína o de la secuencia de la proteína. Cuando se evalúa a partir de la información estructural de la proteína, se alcanza una correlación con los datos experimentales de 0.70 (error estándar de 1.32 Kcal / mol) y 0.62 cuando se evalúa la secuencia de la proteína (error estándar de 1.45 Kcal / mol) (97).

Se ingresó a la página web de I-Mutant v3.0 y se eligió la alternativa de predecir los cambios en la estabilidad de la proteína considerando su secuencia de aminoácidos. Luego se introdujo la secuencia de la proteína (formato FASTA) obtenida de Uniprot y se especificó la posición del aminoácido mutado. Los valores de T° y pH por defecto (25°C y 7, respectivamente) fueron empleados para la predicción. Finalmente se eligió la clasificación ternaria la cual clasifica las mutaciones en tres categorías: mutación neutra ($-0.5 \leq \Delta\Delta G \leq 0.5$), mutación estabilizadora ($\Delta\Delta G \leq -0.5$) y mutación desestabilizadora ($\Delta\Delta G > 0.5$). Finalmente, los resultados fueron tabulados y reportados para las tres proteínas del estudio.

4.6.8. Predicción de la accesibilidad al solvente de las proteínas

Para estimar si el cambio aminoacídico afectaba la accesibilidad al solvente, de los aminoácidos afectados en las proteínas CREBBP, KMT2C y KMT2D, se utilizó el servidor NetSurfP 1.1 (98). Según el acceso de los residuos de aminoácidos al solvente, ellos pueden estar expuestos o enterrados. La puntuación Z que proporciona NetSurfP, resalta la confiabilidad de la predicción de superficie, este discrimina las buenas predicciones de las malas en todo el rango de aminoácidos completamente enterrados a completamente expuestos.

NetsurfP utiliza un método que consta de dos conjuntos de redes neuronales (redes primarias y redes secundarias). Las redes primarias enfocadas en el entrenamiento del perfil de secuencia y estructura secundaria predicha y las redes secundarias entrenadas para predecir la exposición relativa de un aminoácido a la superficie. El rendimiento predictivo general de la red neuronal obtiene un coeficiente general de correlación de Pearson de 0,72 (98).

Antes de comenzar con el análisis, primero se prepararon las secuencias de cada variante aminoacídica por cada proteína. Para esto, se utilizó la secuencia de las proteínas nativas (descargadas anteriormente de Uniprot) sobre la cual se introdujeron manualmente los reemplazos de aminoácidos, uno por vez, usando el programa Microsoft Word. Las variantes aminoacídicas se obtuvieron del portal de BioMuta v4.0- HIVE (véase Tablas 6 - 8).

Se generó un solo archivo por proteína conteniendo en la parte superior la secuencia nativa seguida de cada una de las variantes. Una vez listo este archivo, se introdujo al servidor NetSurfP 1.1 para su correspondiente análisis. Los

resultados fueros tabulados usando el programa de hoja de cálculo Microsoft Excel de manera ordenada para su posterior comparación.

4.6.9. Identificación de las alteraciones en el número de copias (CNA)

Para identificar las alteraciones en el número de copias (amplificaciones y deleciones), presentes en cada tipo de cáncer ginecológico, se utilizó el portal web Broad GDAC Firehose. Se utilizó la herramienta GISTIC 2.0 disponible en Firehose y luego se descargaron y almacenaron los resultados. Los algoritmos de la herramienta GISTIC 2.0 identificaron las regiones del genoma que se amplifican o eliminan significativamente en las muestras normales como tumorales. Además, GISTIC 2.0 informa las ubicaciones genómicas y los valores q de la tasa de descubrimiento falso Benjamini-Hochberg calculados para las regiones aberrantes (99).

Dado que los resultados contienen información de otros genes que no son de interés para el estudio, se filtraron los datos correspondientes sólo a los 160 genes de interés. Para ello se utilizó la función `stri_extract_all()` de Rstudio 1.1.463 que reconoció los nombres de los 160 genes de interés. Se consideraron sólo los CNAs con valores $q < 0.05$.

4.6.10. Identificación de Polimorfismo de Nucleótido Simple (SNP)

Para identificar los SNPs vinculados con el riesgo de cáncer ginecológico se utilizó el Catálogo de Estudios de Asociación de Genoma Completo (GWAS). Estos estudios caracterizan los SNPs como unidad de variación genética asociada con un fenotipo normal o patológico, con un valor $p < 1 \times 10^{-5}$ (100).

El proceso de selección de los genes y SNPs se describen a continuación. En la barra de búsqueda del catálogo de GWAS se introdujeron las siguientes palabras clave: “Cervical”, “Ovarian”, “Endometrial” y “Uterine”. Sólo fueron seleccionados aquellos estudios relacionados con cáncer en los órganos mencionados. Se descargaron los archivos conteniendo todos los SNPs asociados. Para extraer los resultados que correspondan sólo a los 160 genes de estudio, se utilizó la función `stri_extract_all()` de Rstudio 1.1.463 que identificó sólo los genes de interés.

A continuación, se introdujeron los SNPs, por gen, en el repositorio de Datos genómicos comunes (GDC) del Instituto Nacional del Cáncer (NCI) (<https://portal.gdc.cancer.gov/>) con la finalidad de conocer si estaban presentes en las participantes de los proyectos TCGA. Se empleó GDC pues contiene datos de TCGA y la denominación de SNPs presentes en la población de estudio. Finalmente, los datos fueron tabulados por gen, código de SNP y tipo de cáncer.

4.6.11. Identificación de genes con expresión diferencial

Para identificar aquellos genes con distinto patrón de expresión (genes diferencialmente expresados, DEG) entre tejidos normales y tumores, se utilizó el recurso GEPIA (101). GEPIA es un servidor web que analiza los datos de expresión de secuenciación de ARN (ARN-Seq) de tumores y muestras normales de los proyectos TCGA y GTEx (102).

Brevemente, se ingresó al portal GEPIA y en el menú principal se seleccionó la opción “Análisis del tipo de cáncer” seguido de la opción “Análisis de genes diferenciales”, posteriormente se seleccionó el proyecto TCGA de interés

(CESC, OV, UCEC y UCS). Para determinar si los niveles de expresión estaban alterados se consideró el valor de *Fold Change* (FC) que representa el cambio del nivel de expresión entre la muestra tumoral y normal. Los valores de FC estuvieron transformados en la escala logarítmica $\log_2$ para normalizar y garantizar la comparación de estas muestras. Se consideró al $\log_2$ FC < -1.5 como “subexpresado” y al $\log_2$ FC > 1.5 como “sobreexpresado”.

El método estadístico utilizado para el análisis diferencial de la expresión génica fue ANOVA de cuatro vías. El recurso GEPIA realizó el análisis de varianza de cuatro vías (ANOVA), utilizando el sexo, la edad, el origen étnico y el estado de la enfermedad (tumor o normal) como variables para calcular la expresión diferencial. El método de tasa de descubrimiento falso (FDR) de Benjamini-Hochberg se utilizó para obtener el valor p ajustado de prueba múltiple, un valor de $q < 0.05$ se consideró como indicador de una diferencia estadísticamente significativa (101).

Una vez realizado el análisis, se ingresó el nombre de cada gen de estudio en la barra de búsqueda para obtener los correspondientes perfiles de expresión para CESC, OV, UCEC y UCS, finalmente los resultados fueron tabulados.

Posteriormente, se elaboraron diagramas de Venn utilizando VENNY, un servicio científico del Centro Nacional de Biotecnología de España (CNB), para representar los conjuntos de DEGs identificados en cada neoplasia ginecológica.

4.6.12. Anotación funcional GO y análisis de enriquecimiento de vías KEGG de los DEGs

Una vez determinados los DEGs por cada tipo de cáncer ginecológico, se procedió a reconocer los procesos biológicos asociados con dichos genes. Para ello primero se realizó un análisis de anotación funcional u ontología génica (GO) usando el recurso Genecodis (103). El análisis de GO determina las funciones potenciales de una lista de genes incluyendo el proceso biológico (BP), la función molecular (MF) y los componentes celulares (CC). Un valor de p ajustado $< 0,01$ fue considerado para inferir una significancia estadística.

En segundo lugar, se condujo un análisis de enriquecimiento de rutas KEGG (The Kyoto Encyclopedia of Genes and Genomes). KEGG es una base de datos sintética que incluye una colección de rutas bioquímicas. El estudio de vías enriquecidas en los DEGs se utiliza para mostrar las conexiones entre genes y vías comunes entre ellos. Este análisis se realizó utilizando el recurso Genecodis, ingresando la lista de DEGs por tipo de cáncer y, considerando un nivel de significancia de p ajustado $< 0,05$.

4.6.13. Construcción de la red de interacción proteína-proteína (PPI) de los DEGs y la identificación de proteínas centrales

Con la finalidad de reconstruir la red de PPI de los DEGs, se utilizó la base de datos STRING v10.5 (The Search Tool for the Retrieval of Interacting Genes) (104). Las interacciones de STRING incluyen asociaciones físicas (directas) y funcionales (indirectas) derivadas de distintas fuentes incluidas experimentos de alto rendimiento, coexpresión y conocimientos previos de la literatura (104).

Se ingresaron los nombres de los DEGs y se filtraron las asociaciones con puntuación de confianza mayor a 0.4. En la red de PPI, cada nodo representó a un gen y los bordes representaron interacciones entre nodos, los nodos aislados fueron eliminados. Además, se utilizó siete tipos de evidencias para predecir asociaciones: la evidencia de fusión, evidencia de vecindad, evidencia de coocurrencia, evidencia experimental, evidencia de minería de texto, evidencia de base de datos, evidencia de coexpresión y evidencia de homología. A continuación, los resultados fueron descargados y almacenados para su posterior análisis. Asimismo, se utilizó STRING para representar los términos de ontología de genes (proceso biológico, función molecular y componente celular) significativamente enriquecidas en la red de PPI, los 3 primeros términos de ontología génica fueron empleados para fines de visualización.

Con el fin de identificar los genes centrales (“central hub”) en la red de PPI, se utilizó la aplicación CytoHubba 0.1 (105) en Cytoscape 3.7.1 (106). Cytoscape es un programa bioinformático de código abierto para el análisis y visualización de redes de PPI además de integración con bases de datos públicos.

Los datos de PPI descargados de la base de datos STRING fueron importados al programa Cytoscape 3.7.1. El análisis de centralidad se calculó con CytoHubba mediante 12 métodos de análisis topológicos disponibles, que incluyen métodos locales (grado, componente de vecindad máxima, densidad del componente de vecindad máxima, centralidad máxima de la camarilla y coeficiente de agrupamiento) y métodos globales (cercanía, excentricidad, radialidad, cuello de botella, estrés, componente perimétrico del borde, intermediación). Los 5

primeros genes identificados por cada método y que se compartieron más de dos veces se definieron como genes centrales.

4.6.14. Estudio de asociación genotipo-fenotipo

Se investigó la asociación de la expresión de los genes de estudio con una serie de rasgos fenotípicos de los sujetos de estudio. Las listas de estos caracteres se muestran en la Tabla 2.

4.6.14.1. Asociación entre la expresión génica, la sobrevida global y sobrevida libre de enfermedad

Mediante el uso de GEPIA, se investigó la asociación entre la expresión génica y la sobrevida de los pacientes TCGA con CESC, OV, UCEC y UCS.

GEPIA proporciona datos de expresión génica determinada por los proyectos de consorcios como TCGA y Gtex (101), Gtex contiene datos de RNA-Seq de tejidos "normales" provenientes de donantes sanos, no relacionados con TCGA (107). Además, GEPIA permite evaluar el pronóstico de la enfermedad (sobrevida global, OS y sobrevida libre de enfermedad, DFS) en función de la expresión génica, mediante la prueba de rango logarítmico, también conocida como prueba de Mantel-Cox.

El estudio de asociación en GEPIA se realizó introduciendo individualmente los nombres de los DEGs en la barra de búsqueda, luego se ingresó a “sobrevida” y se eligieron por separado las opciones “Sobrevida global (OS)” y “Sobrevida libre de enfermedad

(DFS)". Posteriormente, se seleccionó el índice de riesgo proporcional de Cox (HR) y un intervalo de confianza del 95%. Finalmente se seleccionó el tipo de cáncer para luego aplicar el análisis, los resultados obtenidos fueron descargados y almacenados en formato PDF.

4.6.14.2. Asociaciones entre la expresión génica y variables fenotípicas

Con este fin se utilizó el recurso en línea UCSC Xena (108), para visualizar y analizar los datos de expresión (ARN-seq) del conjunto de DEGs y su correlación con variables clínicas disponibles en cada cáncer ginecológico. El recurso UCSC Xena, alberga un conjunto de datos ómicos provenientes de diversas plataformas públicas, incluyendo el proyecto TCGA (Atlas del Genoma del Cáncer) y permite a los usuarios explorar las correlaciones entre una variable genómica con otra fenotípica. Así, el recurso UCSC Xena calcula de manera automática los estadísticos como la correlación de Pearson y Spearman, el test de Welch y ANOVA para describir la relación entre las variables que el usuario selecciona (genómicas y fenotípicas) (108). En este estudio se utilizó la correlación de Pearson para describir la relación entre la expresión génica de los DEGs con la edad al momento del diagnóstico y el peso de las pacientes con cáncer ginecológico.

Brevemente, se ingresó al portal UCSC Xena y se seleccionó el proyecto TCGA a consultar (CESC, OV, UCEC y UC),

seguidamente, se seleccionó la opción “fenotipo” y se buscó la opción de “sitio primario” para referirse al tumor primario, luego en la siguiente columna se seleccionó la opción “genómica” y se ingresó en la barra de búsqueda los nombres de los DEGs (hasta máximo 10 por búsqueda). Posteriormente, en otra columna, se seleccionó la opción “fenotípica” y se buscó las variables clínicas “edad al momento del diagnóstico” y “peso” disponibles según el tipo de cáncer consultado. Una vez introducido y aceptado esos términos, se accedió a la visualización de los resultados del análisis de correlación. Finalmente, los resultados estadísticamente significativos fueron obtenidos manualmente para su tabulación.

4.6.15. Metilación del ADN a nivel del promotor del gen

Para conocer los patrones de metilación en la región promotora de cada gen, se utilizaron las herramientas web MethHC (109) y MEXPRESS (110). MethHC evalúa los metilomas de 18 cánceres humanos, generados por la plataforma Illumina HumanMethylation450K BeadChip, realiza la prueba t de Student para probar la diferencia entre las muestras tumorales y normales acompañado de un valor p como la probabilidad de obtener un resultado estadístico de prueba (109). Por otro lado, MEXPRESS permite la visualización gráfica de datos de metilación generados por Illumina Human Methylation 450K, de expresión (IlluminaHiSeq_RNASeqV2) y datos clínicos de TCGA, a nivel de un solo gen. El análisis estadístico que proporciona MEXPRESS es la correlación de Pearson

y la prueba no paramétrica de suma de rangos de Wilcoxon. El primero se utiliza para comparar dos tipos de datos que tienen más de 2 niveles (por ejemplo, datos de expresión y metilación), mientras que el segundo se usa para calcular la diferencia de una variable entre dos grupos (por ejemplo, la diferencia en la expresión entre hombres y mujeres). Para corregir las comparaciones múltiples, MEXPRESS aplica la corrección de tasa de descubrimiento falso Benjamini-Hochberg (110).

Se ingresó al portal METHHC y en la barra de búsqueda se introdujo la lista de 160 genes de estudio. A continuación, se seleccionó el proyecto TCGA de interés (UCEC, UCS y CESC), luego la opción de metilación de “promotor” y el método de evaluación del nivel de metilación (promedio). Una vez realizada la búsqueda, se descargaron los datos de metilación y de expresión génica correspondiente a cada gen, de manera individual. Para compilar todos los archivos en uno solo, se utilizó la función `lapply` de Rstudio 1.1.463. Finalmente, los resultados fueron almacenados para su análisis posterior.

Para determinar si la metilación del promotor variaba entre las células cancerosas y normales, se calculó la diferencia ($\Delta\beta$) de las medianas del valor β de los tumores menos la de los controles. Los valores de $\Delta\beta > 0.15$ se clasificaron como promotores “hipermetilados” mientras que valores de $\Delta\beta < -0.15$ se agruparon como promotores “hipometilados” (12).

Debido a que los datos de metilación de ADN de OV y UCS no están disponibles en MethHC, se consultó MEXPRESS. Para la consulta, cada gen de estudio fue introducido en la barra de búsqueda de manera individual, luego se seleccionó TCGA OV y TCGA UCS, según corresponda. A continuación, se obtuvieron los

valores del coeficiente de correlación de Pearson entre los valores de expresión génica y metilación de promotor, en muestras normales y tumorales. Seguidamente los resultados se agruparon según tengan correlación positiva y negativa y posteriormente se aplicó la regla general para interpretar el tamaño de un coeficiente de correlación (12). Estos valores fueron tabulados de manera manual para luego ser analizados.

V. RESULTADOS

5.1. Lista de genes implicados en mecanismos epigenéticos

Se identificaron en total 160 genes codificantes de proteínas involucradas en mecanismos epigenéticos, incluyendo 152 enzimas modificadoras de histonas y ocho enzimas de metilación del ADN (Figura 8 y Tabla 1A en Anexo 1). En la categoría de escritores epigenéticos, es decir, enzimas que añaden grupos metilos y otros sustituyentes al ADN e histonas, el grupo de estudio contiene 105 enzimas, la mayoría de ellas corresponden a enzimas modificadoras de histonas (Figura 8). De los 105 escritores epigenéticos, se identificaron que 19 cumplían además la función de lectores epigenéticos (Tabla 1A en Anexo 1).

En la categoría de borradores epigenéticos, es decir, enzimas que eliminan grupos metilos y otros sustituyentes del ADN e histonas, contamos con 55 enzimas, la mayoría de ellas también enzimas modificadoras de histonas (Figura 8). De los 55 borradores epigenéticos, se identificaron que 8 cumplían además la función de lectores epigenéticos (Tabla 1A en Anexo 1). No se incluyeron genes de miRNA en este estudio.

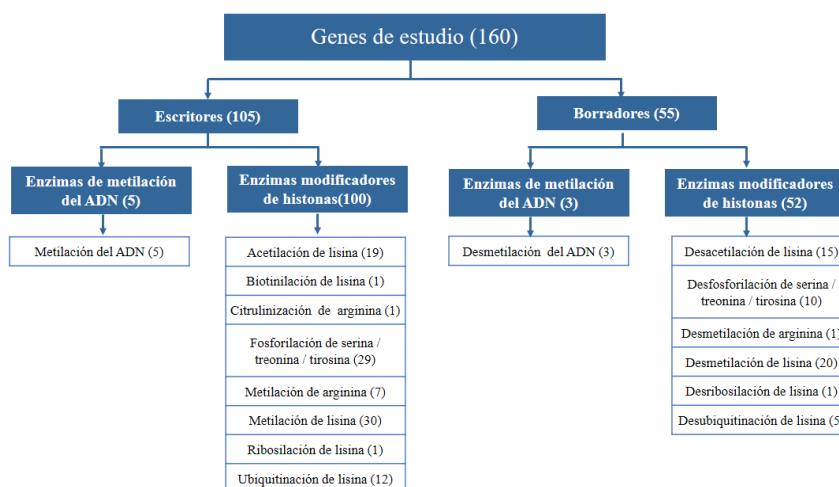


Figura 8. Esquema que muestra la clasificación y número de los genes de estudio

5.2. Anotación funcional y análisis de enriquecimiento de rutas bioquímicas de los genes de interés

Con el fin de determinar las funciones y vías metabólicas asociados con los 160 genes de estudio, se condujo un análisis de ontología genética (GO) y enriquecimiento de rutas KEGG, utilizando Enrichr.

Los resultados de GO mostraron los procesos biológicos significativamente enriquecidos en el grupo de estudio (Figura 9A). Estos incluyeron la regulación de la transcripción a nivel del ADN molde (66 genes), regulación de la transcripción por el promotor del ARN polimerasa II (51 genes), regulación positiva de la transcripción a nivel del ADN molde (47 genes), modificación de histonas (44 genes), regulación negativa de la transcripción a nivel del ADN molde (42 genes), regulación negativa de la expresión génica (36 genes), regulación negativa del proceso biosintético de macromoléculas celulares (34 genes), regulación negativa de la transcripción con molde de ácido nucleico (31 genes), desmetilación de lisina de histona (18 genes) y metilación de histonas (12 genes). La Tabla 2A (Anexo 2) detalla los 20 primeros términos GO de enriquecimiento para la categoría de proceso biológico.

El análisis de las vías KEGG mostró que los genes de estudio estaban enriquecidos significativamente en la degradación de lisina (19 genes), carcinogénesis viral (15 genes), alcoholismo (13 genes), microRNAs en cáncer (12 genes), vía de señalización de la hormona tiroidea (11 genes), ciclo celular (10 genes), vía de señalización MAPK (9 genes), vía de señalización de FoxO (9 genes), infección por el virus de Epstein-Barr (8 genes) y vía de señalización Notch (7 genes) (Figura 9 B).

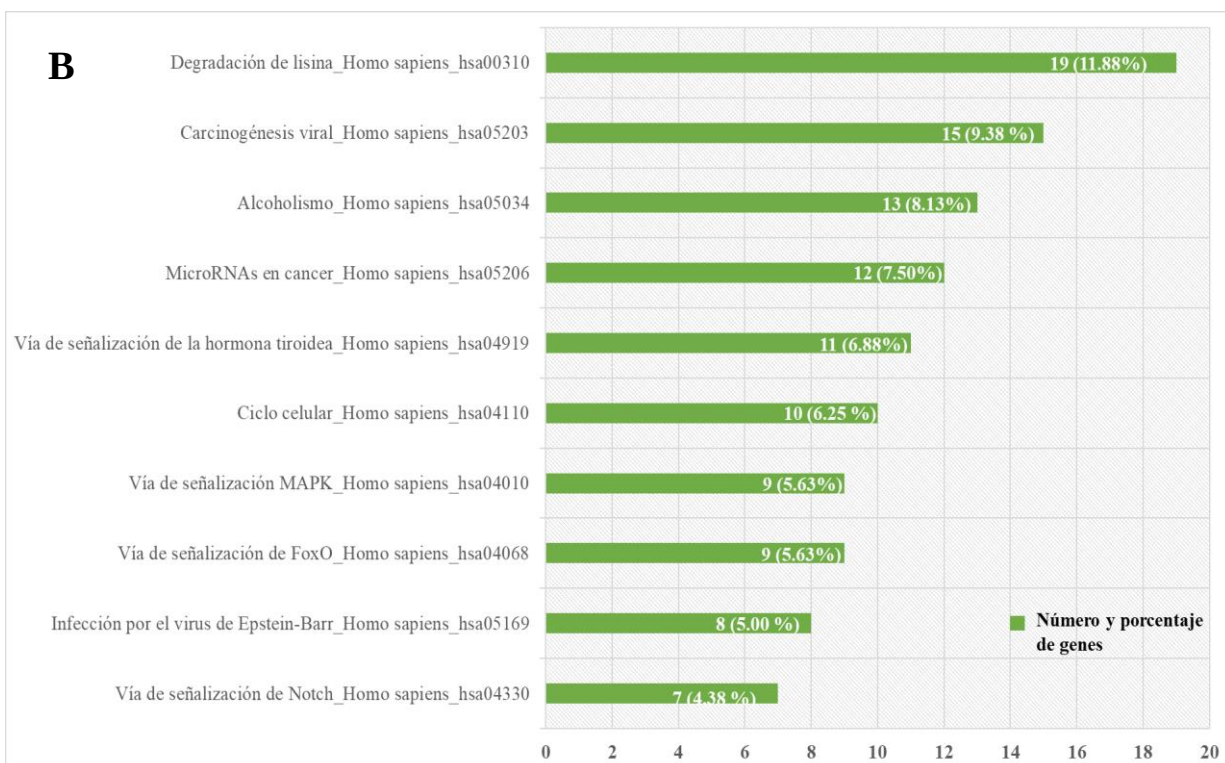
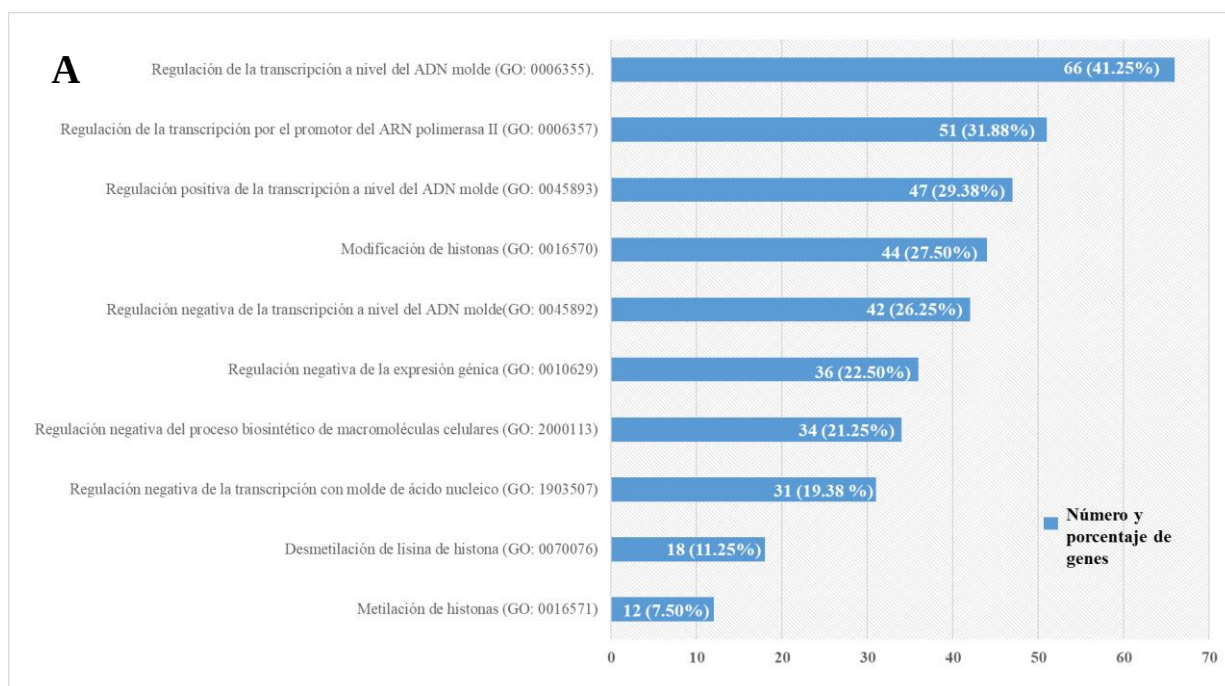


Figura 9. Anotación funcional GO y enriquecimiento de vías KEGG según Enrichr en los 160 genes de interés. (A) 10 principales términos del proceso biológico GO y el número de genes involucrados. (B) 10 principales vías KEGG y el número de genes involucrados

En la Tabla 3A (Anexo 2) se muestran los resultados del análisis de enriquecimiento de vías KEGG de todos los genes de estudio.

En la Figura 10 se muestra un clustergrama detallando los 10 primeros términos enriquecidos y sus genes involucrados. El proceso más enriquecido fue el de transcripción a nivel del ADN molde con 66 factores involucrados. Observamos también una participación mayoritaria de los genes de la familia HDACs (HDAC1, HDAC2, HDAC3, HDAC4, HDAC5, HDAC6, HDAC9 y HDAC10) y de la familia KAT (KAT6B, KAT6A, KAT5 y KAT8) implicados en 7 de 10 procesos biológicos incluyendo la modificación de histonas, regulación de la transcripción a nivel del ADN molde, regulación negativa de la transcripción a nivel del ADN molde, regulación negativa del proceso biosintético de macromoléculas celulares, regulación positiva de la transcripción a nivel del ADN molde, regulación negativa de la expresión génica, regulación negativa de la transcripción con molde de ácido nucleico y la regulación de la transcripción por el promotor del ARN polimerasa II.

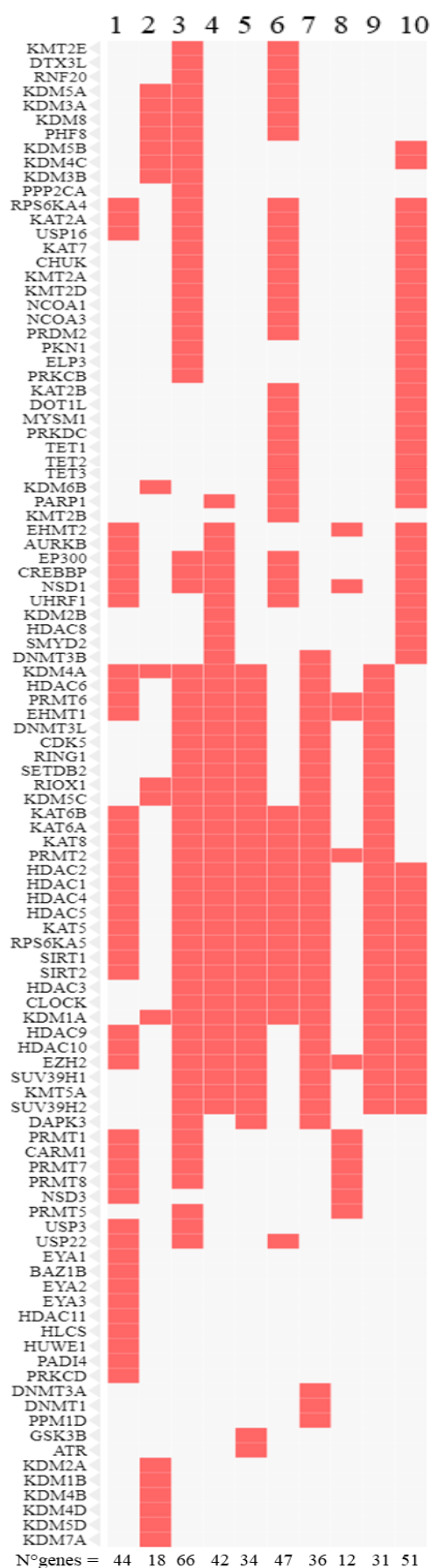


Figura 10. Clustergrama del conjunto de genes de estudio asociados a los 10 principales procesos biológicos, mediante Enrichr.

Las filas y columnas representan los genes de entrada y los PBs, respectivamente. Los genes en la fila están agrupados por la similitud de los contenidos de los PBs. Los términos enriquecidos en las columnas se clasifican según el orden descendente de las puntuaciones de enriquecimiento (valores de P), de izquierda a derecha.

Los 10 principales procesos biológicos asociados son: (1) Modificación de histonas, (2) Desmetilación de lisina de histona, (3) Regulación de la transcripción a nivel del ADN molde, (4) Regulación negativa de la transcripción a nivel del ADN molde, (5) Regulación negativa del proceso biosintético de macromoléculas celulares, (6) Regulación positiva de la transcripción a nivel del ADN molde, (7) Regulación negativa de la expresión génica, (8) Metilación de histonas, (9) Regulación negativa de la transcripción con molde de ácido nucleico y (10) Regulación de la transcripción por el promotor del ARN polimerasa II.

En la parte inferior se muestra el número de genes involucrados en los PBs.

En la Figura 11 se muestra un clustergrama detallando los 10 primeros términos enriquecidos KEGG y sus genes involucrados. El proceso más enriquecido fue degradación de lisina con 19 factores involucrados. Se observó que HDAC1, HDAC2, EP300 y CREBBP estaban implicados en varias vías metabólicas incluyendo la carcinogénesis viral, alcoholismo, microRNAs en cáncer, vía de señalización de la hormona tiroidea, ciclo celular, vía de señalización MAPK, vía de señalización de FoxO y la vía de señalización de Notch (Figura 11).

Además, los genes de la familia HDACs (HDAC3, HDAC4, HDAC5, HDAC6, HDAC7, HDAC8, HDAC9, HDAC10 y HDAC11) estuvieron involucrados en mayor frecuencia en la carcinogénesis viral y el alcoholismo (Figura 11). En tanto, los genes de las familias KMTs (DOT1L, ASH1L, EHMT1, EHMT2, KMT2A, KMT2B, KMT2C, KMT2D, KMT2E, NSD1, SETD1A, SETD1B, SETDB1, SETDB2, SUV39H1 y SUV39H2) solo estuvieron implicados en la vía de la degradación de lisina (Figura 11).

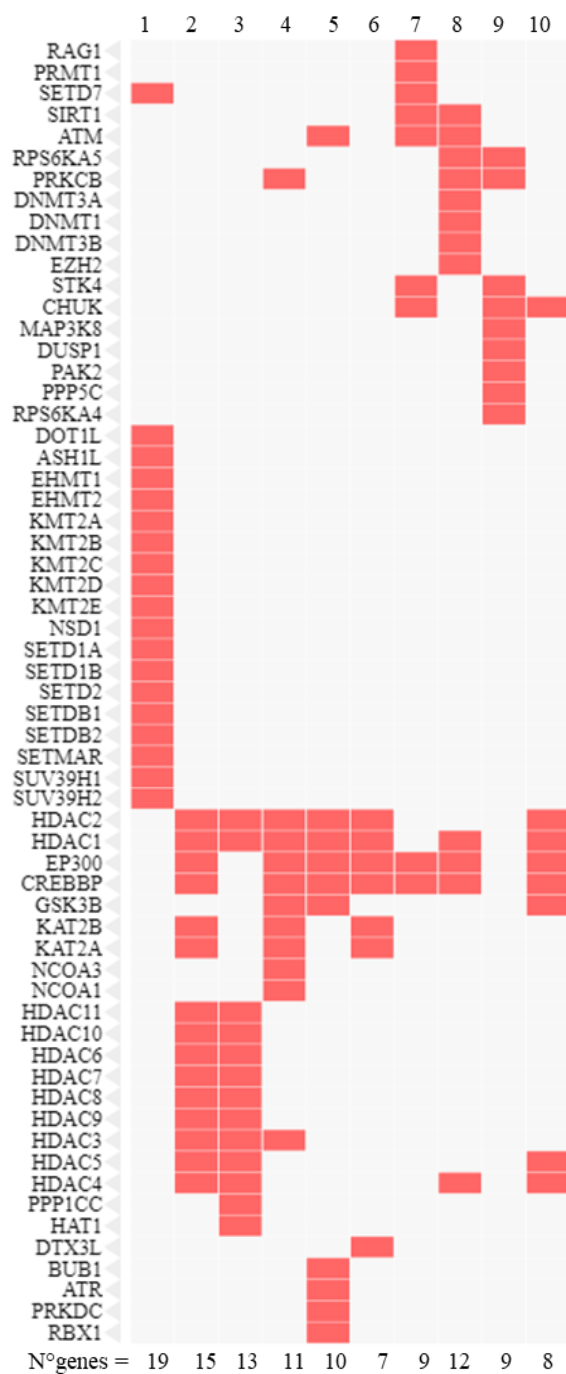


Figura 11. Clustergrama del conjunto de genes de estudio involucrados en las 10 principales vías KEGG, mediante Enrichr.

Las filas y columnas representan los genes de entrada y las Vías KEGG, respectivamente. Los genes en la fila están agrupados por la similitud de los contenidos de los KEGG. Los términos enriquecidos en las columnas se clasifican según el orden descendente de las puntuaciones de enriquecimiento (valores de P), de izquierda a derecha.

Las 10 principales vías KEGG asociadas son: (1) Degradación de lisina, (2) Carcinogénesis viral, (3) Alcoholismo, (4) MicroRNAs en cáncer, (5) Vía de señalización de la hormona tiroidea, (6) Ciclo celular, (7) Vía de señalización MAPK, (8) Vía de señalización de FoxO, (9) Infección por el virus de Epstein-Barr y (10) Vía de señalización de Notch.

En la parte inferior se muestra el número de genes involucrados de forma significativa en las vías KEGG.

5.3. Características de la población de estudio

Se identificaron un total de 1499 pacientes diagnosticadas con cáncer ginecológico, quienes participaron en los estudios TCGA y cuya información está disponible al público. Así mismo, un total de 1533 muestras biológicas fueron colectadas en los cuatro proyectos TCGA analizados, incluyendo los estudios “Cáncer cervical de células escamosas y adenocarcinoma endocervical (CESC TCGA, provisional)”, “Cistoadenocarcinoma seroso ovárico (OV TCGA, provisional)”, “Carcinoma de endometrio del cuerpo uterino (UCEC TCGA, provisional)” y “Carcinosarcoma Uterino (UCS TCGA, provisional)”. Las características demográficas y clínico-patológicas de la población se detallan a continuación.

5.3.1. Características demográficas en pacientes con cáncer ginecológico

De los cuatro estudios analizados, las pacientes con CESC tenían una mediana de la edad al momento del diagnóstico significativamente menor (46 años) que las pacientes con OV (59 años), seguido de UCEC (64 años) y finalmente UCS (68 años). El grupo etario más frecuente en pacientes con CESC fue el de mujeres menores a 50 años (59.3%), mientras que en el grupo de OV fue el de mujeres de entre 50 y 60 años (35.6%), y en el grupo de UCEC y UCS fueron mujeres de entre 60 y 70 años, (34.5% en UCEC 47.4% en UCS) (Tabla 2).

De la población de estudio, los individuos de raza blanca fueron mayoritarios (80.6%), con menor presencia de razas afroamericana (13%) y asiáticas (4.5%). Se observó mayor frecuencia de pacientes no latinos que latinos en los cuatro cánceres ginecológicos. Así, la frecuencia de mujeres no latinas fue de 97.7% en el grupo de UCS, 96.9% en OV, 96.2% en UCEC y 87.7% en CESC (Tabla 2).

La información correspondiente al hábito del consumo de tabaco sólo estuvo disponible para 263 pacientes con CESC. De acuerdo a la cantidad de tabaco y tiempo consumido, el 55.5 % habían fumado menos de 100 cigarrillos en su vida, 24.3 % eran fumadoras ocasionales, 15.2 % habían fumado al menos 100 cigarrillos por un período $\leq$ a 15 años, 3.4% habían fumado al menos 100 cigarrillos por más de 15 años y 1.5% habían fumado al menos 100 cigarrillos en su vida (duración no especificada).

Tabla 2. Características generales de los pacientes con cáncer ginecológico

Características	Total	Tipos de cáncer ginecológico				<i>p</i> valor*
		CESC	OV	UCEC	UCS	
Número de pacientes n (%)	1499 (100)	307 (20.5)	587 (39.2)	548 (36.6)	57 (3.8)	
Número de muestras n (%)	1533(100)	310(20.2)	617(40.3)	549(35.8)	57(3.7)	
Edad (años) n (%)						
media(sd)	59.3(13.3)	48.3(13.8)	59.7(11.5)	63.9(11.1)	69.7(9.3)	
mediana (iqr)	60 (50.5 - 69)	46 (38 - 57)	59 (51 - 68)	64 (57 - 71)	68 (62 - 76)	<0.001
< 50	345 (23.1)	182 (59.3)	117 (19.9)	46 (8.4)	0 (0.0)	
50-60	445 (29.7)	66 (21.5)	209 (35.6)	163 (29.9)	7 (12.3)	
60-70	381 (25.5)	34 (11.1)	132 (22.5)	188 (34.5)	27 (47.4)	
> 70	325 (21.7)	25 (8.1)	129 (22.0)	148 (27.2)	23 (40.4)	
Raza n (%)	1399 (100)	271 (19.4)	556 (39.7)	516 (36.9)	56 (4.0)	<0.001
Nativo americano	15 (1.0)	8 (2.9)	3 (0.5)	4 (0.8)	0 (0.0)	
Hawaiano	12 (0.9)	2 (0.7)	1 (0.2)	9 (1.7)	0 (0.0)	
Blanco	1127 (80.6)	211 (77.9)	498 (89.6)	374 (72.5)	44 (78.6)	
Africano	182 (13.0)	30 (11.1)	34 (6.1)	109 (21.1)	9 (16.0)	
Asiático	63 (4.5)	20 (7.4)	20 (3.6)	20 (3.9)	3 (5.4)	
Etnia n (%)	981(100)	195 (19.9)	350 (35.7)	392 (40.0)	44 (4.5)	<0.001
Latino o Hispano	51 (5.2)	24 (12.3)	11 (3.1)	15 (3.8)	1 (2.3)	
No latino o no Hispano	930 (94.8)	171 (87.7)	339 (96.9)	377 (96.2)	43 (97.7)	

* Nota: el valor de *p* se calculó utilizando la prueba exacta de Fisher y la prueba de Chi2 en las variables categóricas; en las variables discretas se utilizó la prueba de Kruskal Wallis.

CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, OV: Cistoadenocarcinoma seroso ovárico, UCEC: Carcinoma de endometrio del cuerpo uterino, UC: Carcinosarcoma Uterino.

5.3.2. Características clínico-patológicas en pacientes con cáncer ginecológico

La mayoría de las pacientes con CESC eran pre menopáusicas (52.7%), mientras que 35.4% eran postmenopáusicas, 10.6% eran perimenopáusica y el 1.3% carecía de dicha información. En UCEC, la mayoría de pacientes eran postmenopáusicas (86.6%), seguido de pacientes premenopáusicas (6.8%), 3.3% pacientes perimenopáusica e indeterminados. Con respecto a UCS, el 98.2% eran pacientes postmenopáusicas y el 1.8% pacientes perimenopáusica (Tabla 3). No se obtuvo información del estado de menopausia en pacientes con OV.

En relación al índice de masa corporal (IMC, kg/m^2), la mediana del IMC fue mayor en pacientes con UCEC (32.3 kg/m^2), seguido de pacientes con UCS (28.2 kg/m^2) y pacientes con CESC (26.7 kg/m^2), según muestra la Tabla 3. No se tuvo información del IMC para OV. Considerando los criterios y rangos de la Organización mundial de la Salud (OMS), las mujeres con UCEC presentaron "obesidad clase I" al tener IMCs de entre $30 - 34.99 \text{ kg/m}^2$ mientras que las mujeres con CESC y UCS tuvieron "pre-obesidad" con IMCs de entre $25-29.99 \text{ kg/m}^2$ (111).

Se evaluó la frecuencia de pacientes por estadio del tumor, considerando las pautas del Comité Conjunto Americano sobre el Cáncer (AJCC), que clasifica al tumor en cinco estadios (0-IV). El estadio 0 se refiere a la presencia de células anormales en el tejido y que no se han diseminado al tejido cercano; los estadios I, II y III, se refieren a la presencia de un cáncer y cuanto más aumenta el número, mayor es la extensión del tumor en los tejidos cercanos; el estadio IV se refiere a que el cáncer se ha diseminado a otras partes distantes del cuerpo (112).

Los estadios clínicos más frecuentes en la población de estudio fueron el III (45.7%) y I (36.6%). Así, los sujetos con cáncer en estadio I fueron más frecuentes en los proyectos CESC, UCEC y UC. Al contrario, los sujetos con OV tuvieron predominantemente la enfermedad en estadio III (Tabla 3).

El 59.9% de los individuos con cáncer ginecológico estuvieron libre de enfermedad desde el tratamiento inicial al momento del muestreo, mientras que el 40.1% presentó el estado recurrente/progresado. Todos los individuos con el estado libre de enfermedad se presentaron en la cohorte UCS, mientras que la mayoría de los individuos en estado recurrente/progresado fueron del grupo de OV (72.6%) (Tabla 3).

El estado de infección por el virus del papiloma humano (VPH) estuvo reportado sólo en una minoría de la cohorte CESC incluyendo a 2.9% de los sujetos con VPH subtipo 16 y 0.9% con VPH subtipo 18.

Tabla 3. Características clínico-patológicas de los pacientes con cáncer ginecológico

Características	Total	Tipos de cáncer ginecológico				p valor *
		CESC	OV	UCEC	UC	
Número de pacientes n (%)	1499 (100)	307 (20.5)	587 (39.2)	548 (36.6)	57 (3.8)	
Edad (años)						
mediana (iqr)	60 (50.5 - 69)	46 (38 - 57)	59 (51 - 68)	64 (57 - 71)	68 (62 - 76)	<0.001
IMC kg/m²						
Mediana (iqr)	30.1(24.9-36.5)	26.7(22.7 - 32.5)	-	32.3(26.4 – 39.0)	28.2(24.0 - 32.9)	
Menopausia n (%)	809 (100)	237 (29.3)	-	517 (63.9)	55 (6.8)	<0.001
premenopáusica	160 (19.8)	125 (52.7)	-	35 (6.8)	-	
postmenopáusica	586 (72.4)	84 (35.4)	-	448 (86.6)	54 (98.2)	
perimenopáusica	43 (5.3)	25 (10.6)	-	17 (3.3)	1 (1.8)	
indeterminada	20 (2.5)	3 (1.3)	-	17 (3.3)	-	
Estadio clínico n (%)	1487 (100)	300 (20.2)	582 (39.1)	548 (36.8)	57 (3.8)	<0.001
Etapa I	544 (36.6)	163 (54.3)	17 (2.9)	342 (62.4)	22 (38.6)	
Etapa II	113 (7.6)	26 (8.7)	30 (5.2)	52 (9.5)	5 (8.8)	
Etapa III	680 (45.7)	90 (30.0)	446 (76.6)	124 (22.6)	20 (35.1)	
Etapa IV	150 (10.1)	21 (7.0)	89 (15.3)	30 (5.5)	10 (17.5)	
Estado OS n (%)	1497 (100)	307 (20.5)	585 (39.1)	548 (36.6)	57 (3.8)	<0.001
Vivo	950 (63.5)	235 (76.6)	236 (40.3)	457 (83.4)	22 (38.6)	
Muerto	547 (36.5)	72 (23.4)	349 (59.7)	91 (16.6)	35 (61.4)	
Estado DFS n (%)	1295 (100)	267 (20.6)	499 (38.5)	507 (39.2)	22 (1.7)	<0.001
Libre de enfermedad	776 (59.9)	219 (82.0)	137 (27.4)	398 (78.5)	22 (100)	
Recurrente / Progresado	519 (40.1)	48 (18.0)	362 (72.6)	109 (21.5)	0 (0.0)	

* Nota: el valor de *p* se calculó utilizando la prueba exacta de Fisher y la prueba de Chi2 en las variables categóricas; en las variables discretas se utilizó la prueba de Kruskal Wallis.

CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, OV: Cistoadenocarcinoma seroso ovárico, UCEC: Carcinoma de endometrio del cuerpo uterino, UC: Carcinosarcoma Uterino.

(-) Sin datos

5.3.3. Características generales de los pacientes con cáncer ginecológico admitidos en el INEN, Perú

Se realizó una búsqueda electrónica en las bases de datos del Instituto Nacional de Enfermedades Neoplásicas (INEN) y PUBMED para obtener información clinicopatológica y demográfica de pacientes con cáncer ginecológico en Perú.

Se escogieron tres estudios que consideraban la información clinicopatológica (estadio FIGO, tipo histológico y sobrevida global) y demográfica (edad) de pacientes con cáncer ginecológico (Tabla 4). Estos estudios fueron de análisis observacional y consideraron a 335 participantes con cáncer de ovario y 8217 participantes con cáncer cervical, también se describió las principales características metodológicas de cada estudio (Tabla 4).

El estudio de Alcarraz C, *et al.*, incluyó pacientes peruanas tratadas con quimioterapia neoadyuvante con carboplatino (AUC 6 mg/ml/min) y paclitaxel (80 mg/m²semanal) seguido de citorreducción de intervalo (CCI), durante el período 2010-2014. Se observó que las mujeres que alcanzaron la CCI, tuvieron una mediana de edad de 59 años, con estadio FIGO más frecuentes de IIIC (63.4%). Se reportó que el tipo histológico más frecuente fue el adenocarcinoma seroso de alto grado (90.2%) y la sobrevida global a los 12, 24 y 36 meses fue en 91,1 %, 80,8 % y 75,0 %, respectivamente (Tabla 4).

Por otro lado, el estudio de Urritia C, *et al.*, evaluó la frecuencia y factores asociados al compromiso parametrial (CP) en pacientes con cáncer de cérvix de estadio IB-1, sometidas a histerectomía radical (HR) durante septiembre del 2007 y agosto del 2012. El estudio mostró que las pacientes sometidas a HR tuvieron una edad media de 44.3 años, además, el tipo histológico con mayor

frecuencia fue el epidermoide (carcinoma de células escamosas) (74.3%). No se evaluó análisis de sobrevida global en el estudio (Tabla 4).

Finalmente, el estudio de Ruiz R, *et al*, determinó las características clínicas, histológicas y los factores pronósticos del cáncer cervical (CC) en pacientes de 35 años de edad o menos, entre el 2008 y el 2012. Se observó que 449 pacientes tenían una edad media de 32 años, siendo el grupo etario de 30-35 años el más frecuente (78,9%). El estadio FIGO más frecuente fue el estadio II (48.7%) y el tipo histológico con mayor frecuencia fue el carcinoma de células escamosas (84. 9%). En relación a la sobrevida global a 5 años por estadio fue del 90,9%, 57,5%, 42,7% y 13,3% para los estadios I, II, III y IV, respectivamente (Tabla 4).

Tabla 4. Características generales de los pacientes con cáncer ginecológico admitidos en el INEN, Perú.

Estudio observacional de una serie de casos retrospectivo									
Estudio	Periodo de estudio (años)	Número de participantes	Edades (años)	Tipo de cáncer	Elegibilidad del participante	Características identificadas en la población			
						Demográficas		Clínico-patológicas	
						Edad (años)	Estadio FIGO frecuencia (%)	Tipo histológico frecuencia (%)	Sobrevida global (OS)
Alcarraz C, <i>et al.</i> (113)	2010-2014	335	14 a más	Ovario	Inclusión: 14 años a más, histología de cáncer de ovario epitelial avanzado, estadio III y IV, quimioterapia neoadyuvante a base de paclitaxel a dosis densa y carboplatino. Exclusión: Cirugía de citorreducción primaria o con tratamiento con quimioterapia neoadyuvante a dosis convencionales o aquellas que recibieron quimioterapia en otra institución	n =41 42–73 (mediana: 59)	n =41 IIIC: 26(63.4) IC: 15(36.6)	n =41 <u>Adenocarcinoma seroso:</u> 37 (90.2) <u>Otros:</u> 4 (9.6)	La OS con citorreducción óptima a los 12, 24 y 36 meses fue en 91,1 %, 80,8 % y 75,0 %, respectivamente

Estudio observacional transversal analítico									
Estudio	Periodo de estudio (años)	Número de participantes	Edades (años)	Tipo de cáncer	Elegibilidad del participante	Características identificadas en la población			
						Demográficas		Clínico-patológicas	
						Edad (años)	Estadio FIGO frecuencia (%)	Tipo histológico frecuencia (%)	Sobrevida global (OS)
Urritia C, <i>et al.</i> (114)	2007-2012	214	NR	Cervical	Inclusión: Información clínico - patológica completa, sometidas a HR (clase II y III) y linfadenectomía pélvica Exclusión: No cumple lo anterior	n= 214 21–69 (media:44.3)	IB-1: 214 (100)	<u>Epidermoide:</u> 159 (74.3) <u>Adenocarcinoma:</u> 49 (22.9) <u>Adenoescamoso:</u> 6 (2.8)	NR

Estudio de análisis retrospectivo									
Estudio	Periodo de estudio (años)	Número de participantes	Edades (años)	Tipo de cáncer	Elegibilidad del participante	Características identificadas			
						Demográficas		Clínico-patológicas	
						Edad (años)	Estadio FIGO frecuencia (%)	Tipo histológico frecuencia (%)	Sobrevida global (OS)
Ruiz R , et al.(115)	2008-2012	8003	35 o menor a 35	Cervical	Inclusión: Pacientes con cáncer cervical, (edad = < 35) Exclusión: Historia clínica no disponible, pacientes sin confirmación de diagnóstico y con lesiones preinvasivas del cuello uterino	n= 449 Media: 32± 3.4 <u>30-35 años:</u> 353 (78,9%) <u>25-29 años:</u> 72(16%) <u>20-24 años:</u> 24(5.3%)	n= 449 I: 94 (20.9) II: 219 (48.7) III:118 (26.2) IV:15 (3.3) Desconocido: 3 (0.7)	n = 457 <u>Carcinoma de células escamosas:</u> 388 (84.9) <u>Adenocarcinoma:</u> 50 (11.0) <u>Carcinoma adenoescamoso:</u> 11 (2.4) <u>Tumores mesenquimales:</u> 5 (1.1) <u>Neuroendocrino:</u> 2 (0.4) <u>Misceláneo:</u> 1(0.2)	OS a 5 años: I: 90,9% II: 57,5% III: 42,7% IV:13,3%

NR: No reportado, HR: histerectomía radical, FIGO: Sistema de clasificación de la Federación Internacional de Ginecología y Obstetricia
 La tabla muestra las principales características clínico patológicas y demográficas de pacientes con cáncer ginecológico en Perú (INEN).

5.4. Genes frecuentemente mutados y significativamente mutados

Con el interés de identificar genes mutados con altas frecuencias (FMGs, genes frecuentemente mutados) y los genes significativamente mutados (SMGs, mutación puntual solamente) en cada tipo de cáncer ginecológico, se condujo un análisis de MutSig 2CV v3.1 a través del portal web Broad GDAC Firehose. Este recurso considera los genomas completos de las muestras de tejido canceroso comparadas con el tejido normal, información que es tomada de los proyectos TCGA.

Se observó que el número de pacientes (n) por estudio y el número de muestras difiere de los valores reportados en CBioPortal (Tabla 2). El tamaño de muestra utilizado para el análisis de genes significativamente mutados (SMGs) fue el siguiente: 194 pacientes analizados para CESC, 466 para OV, 248 para UCEC y 57 para UCS (Tabla 5). La razón es que no todas las muestras biológicas de los sujetos participantes fueron utilizadas para la investigación de genoma completo.

Los SMGs presentaron un valor $q < 0.05$ (resaltados en negrita en la Tabla 5). Así, en CESC se identificaron 3 SMGs incluyendo KMT2C, KMT2D y EP300 mientras que en UCEC se encontraron 3 SMGs (ATM, EP300 y CDK17). Todos los SMGs identificados codifican a proteínas modificadoras de histonas. No se identificaron SMGs en OV y UCS debido a que presentaron valores $q > 0.05$.

Por otro lado, basándonos en el análisis de Mutsig2CV y considerando valores de $p < 0.05$, se identificaron en total 44 FMGs en las neoplasias ginecológicas. Así, identificamos 26 genes con 748 mutaciones puntuales en UCEC, seguido de 8 genes con 145 mutaciones puntuales en CESC, 8 genes con 58 mutaciones puntuales en OV y finalmente 2 genes con 4 mutaciones puntuales en UCS (Figura 12 y 13). Del

total de FMGs en los cuatro cánceres ginecológicos, 30 genes se identificaron como modificadores de histonas, presentes en CESC, OV, UCS y UCEC (Figura 13), mientras que sólo dos genes se identificaron como reguladores de la metilación del ADN (DNMT1 y DNMT3B), ambos presentes en UCEC.

En total, para CESC, OV, UCEC y UCS, se identificaron 955 mutaciones puntuales, siendo la mayoría de ellas alteraciones con pérdida de sentido (n=562).

En menor cantidad se identificaron otros tipos de mutaciones incluyendo de cambio de marco de lectura (n=155), silenciosas (n=145), sitio de empalme (n=48), inserciones-deleciones, indels (n=37) y no codificantes (n=8) (Figura 12 y Tabla 5).

Cabe destacar que 3 genes codificantes para enzimas implicadas en la metilación y acetilación de lisina (KMT2D, KMT2C y EP300) y fosforilación de serina / treonina (ATM), fueron comúnmente mutados en CESC y UCEC. Mientras que 2 genes codificantes de enzimas implicadas en la acetilación de lisina y fosforilación de serina / treonina (CREBBP y CDK17), fueron comúnmente mutados en OV y UCEC (Tabla 5).

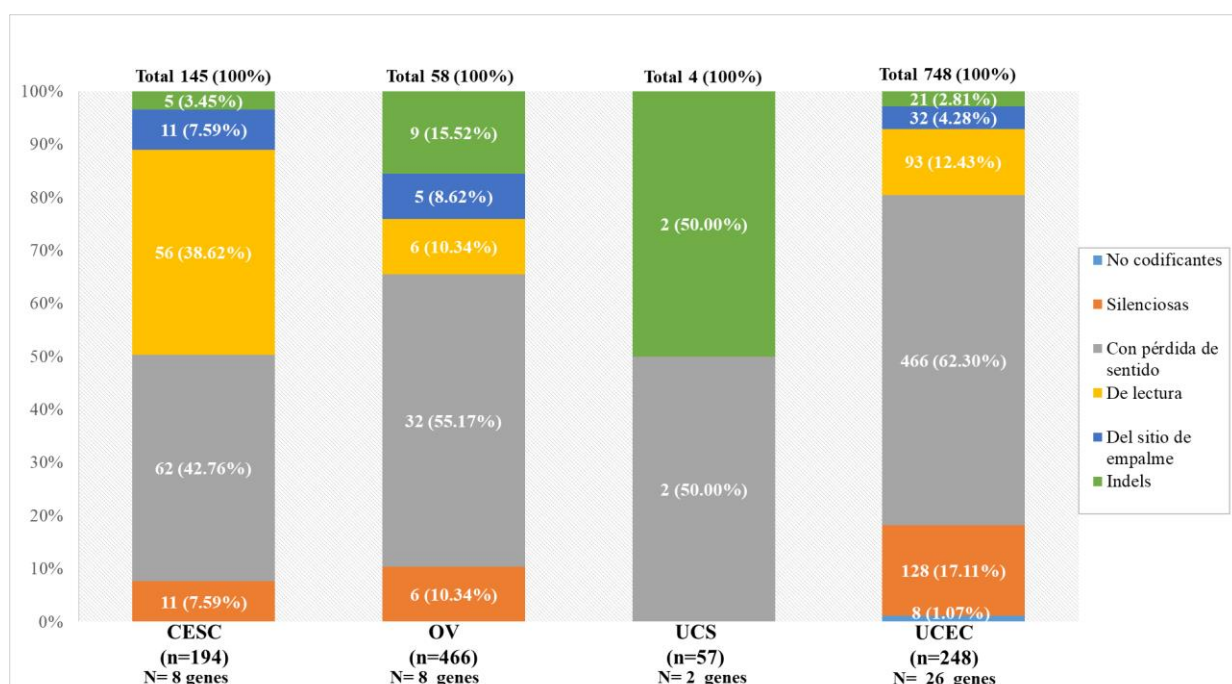


Figura 12. Recuento de mutaciones puntuales y genes frecuentemente mutados según el tipo de cáncer ginecológico. El gráfico de barras detalla el número de mutaciones puntuales y el número de genes frecuentemente mutados (N) en CESC, OV, UCS y UCEC. n: número de pacientes analizados.

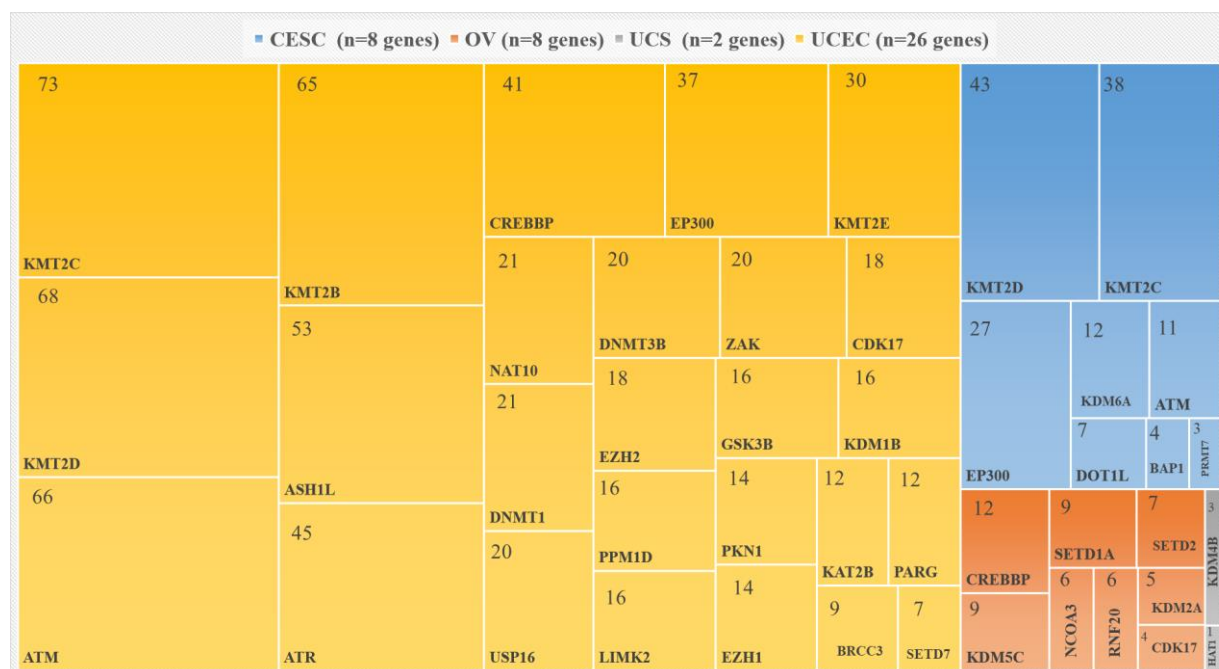


Figura 13. Mapa de árbol de los genes frecuentemente mutados según el tipo de cáncer ginecológico. Cada rectángulo contiene el nombre de los genes afectados y el número de mutaciones totales de manera proporcional a su área.

Tabla 5. Genes frecuentemente mutados en cánceres ginecológicos evaluados por MutSig 2CV

Tipo de Cáncer	Símbolo del Gen	Tipo de mutaciones						N° de pacientes con al menos una mutación no silenciosa	N° de sitios únicos que tienen una mutación no silenciosa
		No codificante	Silenciosas	No silenciosas					
				Con pérdida de sentido	De lectura	Del sitio de empalme	Indels		
CESC (n=194)	KMT2D*	0	8	16	14	3	2	22	35
	EP300*	0	1	12	11	3	0	21	22
	KMT2C*	0	1	11	23	1	2	29	35
	BAP1	0	0	0	4	0	0	4	3
	KDM6A	0	0	8	2	2	0	6	12
	ATM	0	0	9	1	0	1	11	10
	DOT1L	0	1	4	1	1	0	6	6
OV (n=466)	PRMT7	0	0	2	0	1	0	2	3
	CREBBP	0	1	5	1	2	3	11	11
	NCOA3	0	1	3	0	0	2	5	5
	KDM5C	0	1	5	0	0	3	8	8
	KDM2A	0	0	5	0	0	0	5	5
	CDK17	0	0	3	0	0	1	4	4
	SETD2	0	0	4	2	1	0	7	7
UCS (n=57)	SETD1A	0	2	4	2	1	0	7	7
	RNF20	0	1	3	1	1	0	5	5
UCEC (n=248)	KDM4B	0	0	2	0	0	1	3	3
	HAT1	0	0	0	0	0	1	1	1
	ATM*	0	9	37	13	4	3	29	54
	KMT2D	0	11	49	4	1	3	33	57
	KMT2C	0	16	44	9	3	1	26	55
	KMT2B	0	12	42	7	1	3	30	51
	ASH1L	0	11	34	5	2	1	23	38
	ATR	0	6	29	5	1	4	23	39
	CREBBP	0	8	28	3	0	2	22	30
	EP300*	0	5	21	7	2	2	21	31
	KMT2E	0	6	18	5	0	1	17	23
	NAT10	0	2	16	3	0	0	16	19
	CDK17*	0	2	12	3	1	0	14	16
	USP16	2	3	13	1	1	0	13	13
	DNMT3B	0	5	10	3	2	0	13	15
	EZH2	0	3	10	3	2	0	12	14
	DNMT1	0	7	14	0	0	0	9	12
	MAP3K20	2	4	12	1	1	0	12	14
	PPM1D	0	3	8	5	0	0	11	11
	LIMK2	2	1	9	3	1	0	12	11
	GSK3B	0	3	9	3	1	0	9	13
	KAT2B	0	0	9	2	1	0	10	11
	KDM1B	1	3	10	1	1	0	12	10
	PKN1	0	3	8	0	2	1	9	9
	PARG	1	1	7	2	1	0	9	9
	EZH1	0	4	5	4	1	0	7	9
	BRCC3	0	0	7	1	1	0	8	9
	SETD7	0	0	5	0	2	0	7	6

*Genes significativamente mutados (SMGs) con $q < 0.05$, n: número de pacientes analizados en la cohorte TCGA.

5.5. Predicción de los cambios en la estructura, estabilidad y función proteica

Con el interés conocer el efecto dañino o no de las variaciones de aminoácidos, se analizó el cambio en la estructura tridimensional (3D) y función ocasionado por mutaciones puntuales, en algunas proteínas, aplicando métodos computacionales. Se eligieron tres proteínas incluyendo KMT2C, KMT2D y CREBBP. La elección de las dos primeras consideró que ambas son codificadas por SMGs y con presencia de gran número de mutaciones con pérdida de sentido, en pacientes afectados en cáncer ginecológico. En total, 55 mutaciones con pérdida de sentido fueron observadas en KMT2C y 65 en KMT2D, ambos genes presentes en CESC y UCEC. En tanto la elección de CREBBP se basó en que es una diana epigenética accionable, cuya estructura 3D ha sido determinada con el fin de diseñar inhibidores específicos (116). Un total, 33 mutaciones con pérdida de sentido fueron observadas en CREBBP, presentes en OV y UCEC. A pesar de no ser codificada por un SMG, CREBBP ha sido reconocido como un posible objetivo farmacológico contra el cáncer debido a su participación crucial en la activación eficiente de la transcripción génica (116).

Biomuta predijo el impacto funcional de las sustituciones de aminoácidos, en cada una de las tres proteínas elegidas. Para KMT2D, se obtuvo un total de 71 variantes de aminoácidos en los tipos de cáncer cervical y uterino. De estos, el 80.3% (57) de sustituciones de aminoácidos fueron predichas como “dañinas”, de los cuales el 63.2 % (36) afectaron familias o dominios proteicos (Figura 14 A y B).

Similarmente, para KMT2C, se obtuvo un total de 70 variantes de aminoácidos en los tipos de cáncer cervical, ovárico y uterino. De estos, el 65.7 % (46) sustituciones

de aminoácidos fueron predichas como “dañinas”, de los cuales el 50 % (23) afectaron familias o dominios proteicos (Figura 14 A y B).

Finalmente, para CREBBP, se obtuvo un total de 44 variantes de aminoácidos en los tipos de cáncer cervical, ovárico y uterino. De estos, el 88.6 % (39) sustituciones de aminoácidos fueron predichas como “dañinas”, de los cuales el 49.6 % (17) afectaron familias o dominios proteicos (Figura 14 A y B).

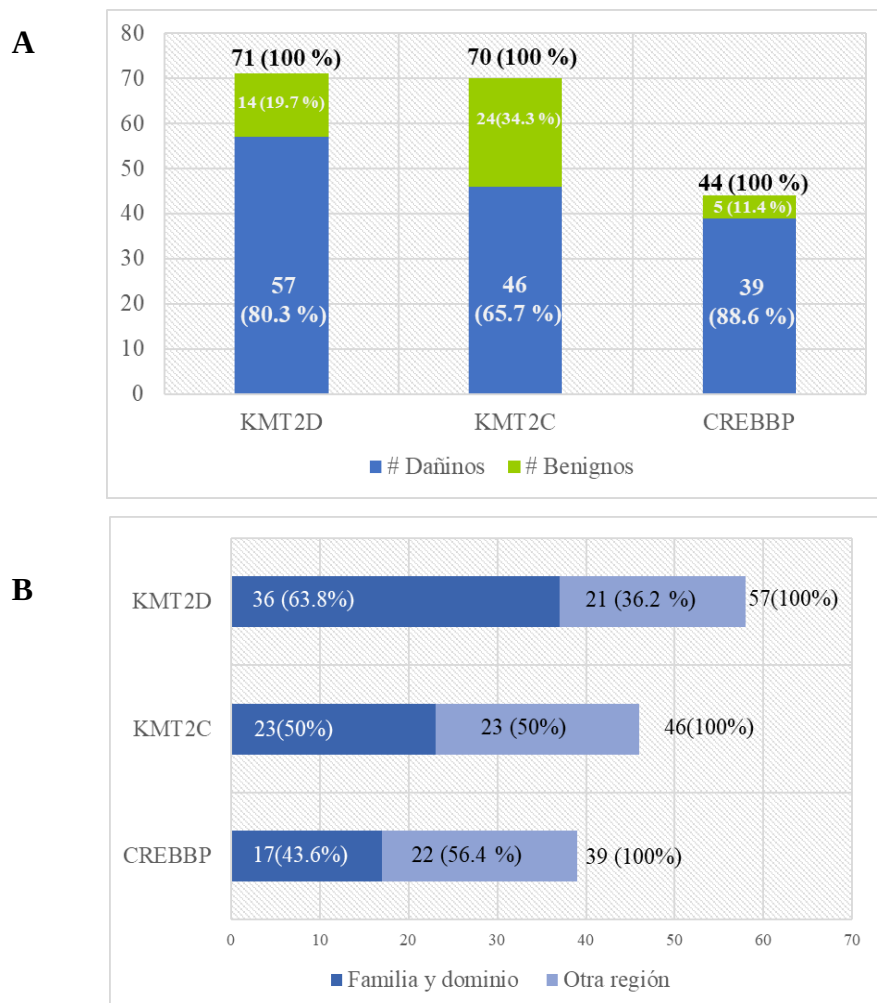


Figura 14. Porcentaje de mutaciones con pérdida de sentido en KMT2C, KMT2D y CREBBP predichas por Biomuta. (A) Número y porcentaje total de mutaciones con pérdida de sentido “dañinas” y “benignas” para cada proteína. (B) Número y porcentaje de mutaciones con pérdida de sentido “dañinas” según su localización en la proteína.

Las mutaciones predichas como dañinas son aquellas que, debido al cambio en un único nucleótido, provoca también un cambio de codón que codifica para un aminoácido diferente, afectando así la función molecular de la proteína.

Para conocer si las mutaciones predichas como dañinas presentaban también cambios en la estabilidad energética de las proteínas, se estimó el cambio teórico de energía libre de Gibbs ($\Delta\Delta G$). De acuerdo al algoritmo de I-Mutant, si $\Delta\Delta G < -0.5$ Kcal/mol, la mutación se predice como desestabilizante para la proteína. Si $-0.5 \leq \Delta\Delta G \leq 0.5$ Kcal/mol, la mutación tendrá un efecto débil sobre la estabilidad energética. Mientras que $\Delta\Delta G > 0.5$ Kcal/mol sugieren mutaciones estabilizadoras. De las 36 sustituciones de aminoácidos “dañinas” para KMT2D, el 61.1% (22) se predijeron como desestabilizadoras, mientras que el 38.9% (14) fueron predichas como neutras (Figura 15 y Tabla 6). Así mismo, de las 23 sustituciones de aminoácidos “dañinas” para KMT2C, el 47.8 % (11) se predijeron como desestabilizadoras, mientras que el 52.2% (12) fueron predichas como neutras (Figura 15 y Tabla 7). Finalmente, de las 17 sustituciones de aminoácidos “dañinas” para CREBBP, el 82.4 % (14) se predijeron como desestabilizadora mientras que el 17.6% (3) fueron predichas como neutras (Figura 15 y Tabla 8). Ninguna sustitución fue predicha como estabilizadora en las mutaciones evaluadas para las tres proteínas.

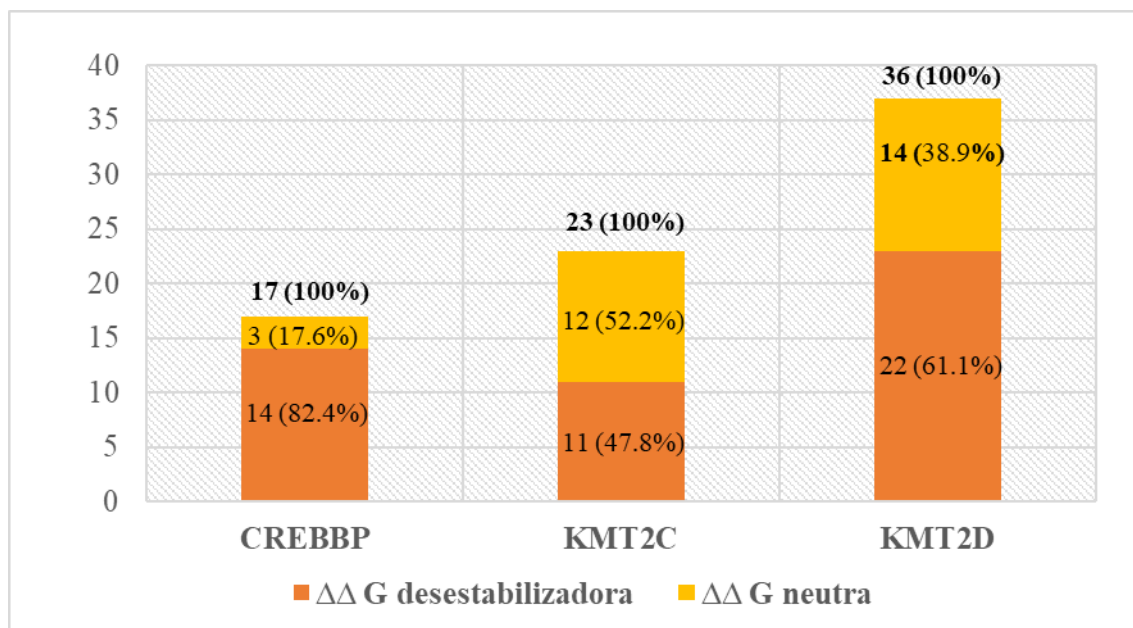


Figura 15. Número y porcentaje de mutaciones con pérdida de sentido “dañinas” que mostraron cambios en la estabilidad energética de las proteínas KMT2C, KMT2D y CREBBP, predichas por I-Mutant 3.0.

Con el interés de conocer si las mutaciones predichas como “dañinas”, cambiaban la exposición al solvente en los aminoácidos afectados, se utilizó NetSurfp v10.1 (Tablas 6-8). La mayoría de las mutaciones predichas como dañinas en KMT2D no variaron la exposición al solvente del residuo alterado, excepto en las mutantes G1387A, L3591F y L4222P (Tabla 6). La mutación G1387A modificó la exposición al solvente del residuo enterrado G1387 volviéndolo un residuo expuesto al solvente (A1387), afectando al dominio Dedo Zinc (PHD-tipo 4) (Tabla 6). Igualmente, la mutación L3591F modificó la exposición al solvente del residuo expuesto L3591 volviéndolo un residuo enterrado (F3591), afectando el dominio de la región espiral (Tabla 6). Por su parte, la mutación L4222P modificó la exposición al solvente del residuo expuesto L4222 volviéndolo un residuo enterrado (P4222), afectando el motivo 3 LXXLL (Tabla 6). Cabe mencionar que las mutaciones G1387A y L3591F además de presentar cambios en la exposición

al solvente, también fueron predichas como desestabilizadoras a nivel energético. Solo la mutación L4222P predicha como neutra a nivel energético no presentó cambios en la exposición al solvente.

Ninguna de las sustituciones predichas como dañinas en KMT2C varió la exposición al solvente del residuo alterado (Tabla 7). Una de las variantes predicha como dañina (S4850L) involucró un residuo en el dominio funcional SET (Tabla 7 Figura 17). Este dominio es responsable de la actividad de metilación específica de lisina en la histona H3 (117). Sin embargo, la mutación S4850L fue predicha como “neutra” en términos energéticos.

Ninguna de las sustituciones en CREBBP varió la exposición al solvente del residuo alterado (Tabla 8). Por otro lado, algunas variantes predichas como dañinas en CREBBP involucraron residuos en dominios funcionales. Así, las mutaciones L1338S, F1345C, R1392Q, G1404V, I1437M, F1439L, I1471M, P1488R, I1515M, D1543H, V1634A y R1682C afectaron el dominio HAT-KAT11 en CREBBP (Tabla 8 y Figura 18). Este dominio cumple un rol importante en la acetilación de histonas y está asociado con la unión del cofactor acetil-coenzima A (Ac-CoA) (118). De estas mutaciones, L1338S, F1345C, R1392Q, I1437M, F1439L, I1471M, P1488R, I1515M, D1543H, V1634A y R1682C fueron desestabilizadoras, mientras que G1404V fue predicha neutra.

Las mutaciones dañinas en las 3 proteínas incluidas en el estudio estructural pueden afectar a la red de contactos de dichos factores. Para investigar ello, se determinó el interactoma de las 3 proteínas por separado, usando STRING v10.5. Se observó que KMT2D y KMT2C formaron una red funcional significativa con 21 genes (p -valor $< 1.0e-16$) (Figura 16A & 17A, respectivamente). La red PPI, reveló que

KMT2D está relacionado con proteínas con actividad histona-lisina N-metiltransferasa, transcripción de la región reguladora de unión al ADN y unión a factores de transcripción (CREB1, EP300, HIF1A, MEF2A, NCOA6, NCOR1, RARA, SETD1A, SIN3A y TP53) (Figura 16A).

El interactoma de KMT2C mostró que formaba una red funcional significativa con 21 genes (p -valor $< 1.0e-16$), incluido KMT2D (Figura 17A). La red PPI reveló que KMT2C está relacionada con proteínas con actividad histona-lisina N-metiltransferasa, transcripción de la región reguladora de unión al ADN y actividad histona metiltransferasa, específica de H3K4 (Figura 17A).

Finalmente, para CREBBP se observó que formaban una red funcional significativa con 21 genes (p -valor = $5.75e-14$) (Figura 18A). Los resultados obtenidos mostraron que CREBBP está relacionada con proteínas involucrados en funciones moleculares como la unión a factores de la transcripción y unión a proteínas incluida la proteína supresora de tumores p53 (Figura 18A).

Tabla 6. Mutaciones puntuales predichas “dañinas” en KMT2D, incluyendo el cambio de estabilidad de cada mutante y cambio de accesibilidad al solvente.

Mutación	PolyPhen (Biomuta 4.0)	I Mutant 3.0	NetSurP-1.1						Familia y Dominio
			AA	Naturaleza	RSA	ASA	Puntuación Z	Asignación	
L205I	0.912	Desestabilizadora	L	No polar	0.138	25.231	0.384	Enterrado	Dedo Zinc (PHD-tipo 1)
			I	No polar	0.142	26.27	0.332	Enterrado	
R466C	0.925	Desestabilizadora	R	Básicos	0.594	135.957	-1.369	Expuesto	Region15 X 5 AA
			C	Polar (sin carga)	0.457	64.163	-0.973	Expuesto	repeticiones de S/P-P-P-E/P-E/A /
P496L	0.48	Neutro	P	No polar	0.52	73.816	-0.976	Expuesto	Region15 X 5 AA
			L	No polar	0.53	97.135	-0.793	Expuesto	repeticiones de S/P-P-P-E/P-E/A /
E568G	0.963	Desestabilizadora	E	Ácidos	0.604	105.554	-0.636	Expuesto	Region15 X 5 AA
			G	Polar (sin carga)	0.587	46.228	-0.676	Expuesto	repeticiones de S/P-P-P-E/P-E/A /
P1093L	1	Neutro	P	No polar	0.544	77.137	-1.089	Expuesto	Rico Pro
			L	No polar	0.457	83.64	-1.357	Expuesto	
P1309L	1	Neutro	P	No polar	0.502	71.22	-1.588	Expuesto	Rico Arg
			L	No polar	0.551	100.815	-1.462	Expuesto	
R1312H	1	Desestabilizadora	R	Básicos	0.557	127.53	-0.879	Expuesto	Rico Arg
			H	Básicos	0.596	108.376	-1.151	Expuesto	
E1355K	0.457	Desestabilizadora	E	Ácidos	0.726	126.832	-0.484	Expuesto	Rico Poly
			K	Básicos	0.585	120.376	-1.229	Expuesto	
G1387A	1	Desestabilizadora	G	Polar (sin carga)	0.298	23.413	-2.133	Enterrado Expuesto	Dedo Zinc (PHD-tipo 4)
			A	No polar	0.34	37.457	-2.157		
L2113P	0.911	Desestabilizadora	L	No polar	0.647	118.447	-2.416	Expuesto	Rico Pro
			P	No polar	0.469	66.579	-1.863	Expuesto	
A2329T	0.993	Desestabilizadora	A	No polar	0.528	58.164	-1.362	Expuesto	Rico Pro
			T	Polar (sin carga)	0.51	70.751	-1.672	Expuesto	
Q2337E	0.996	Neutro	Q	Polar (sin carga)	0.495	88.496	-1.337	Expuesto	Rico Pro
			E	Ácidos	0.545	95.229	-1.227	Expuesto	
P2340S	0.982	Desestabilizadora	P	No polar	0.449	63.77	-1	Expuesto	Rico Pro
			S	Polar (sin carga)	0.512	59.96	-1.522	Expuesto	
R2692W	1	Neutro	R	Básicos	0.518	118.576	-0.908	Expuesto	Región Espiral
			W	No polar	0.486	116.907	-1.314	Expuesto	
R2975H	0.999	Desestabilizadora	R	Básicos	0.615	140.766	-0.533	Expuesto	Rico Pro
			H	Básicos	0.622	113.087	-0.291	Expuesto	
P3381T	0.787	Desestabilizadora	P	No polar	0.401	56.859	-2.231	Expuesto	Rico Gln
			T	Polar (sin carga)	0.43	59.599	-2.901	Expuesto	
L3591F	1	Desestabilizadora	L	No polar	0.287	52.586	-0.735	Expuesto Enterrado	Región Espiral
			F	No polar	0.266	53.366	-0.787		

Mutación	PolyPhen (Biomuta 4.0)	I Mutant 3.0	NetSurP-1.1					Asignación	Familia y Dominio
			AA	Naturaleza	RSA	ASA	Puntuación Z		
L4222P	0.998	Neutro	L	No polar	0.33	60.405	-1.605	Expuesto	LXXLL motivo 3 / Rico Gln
			P	No polar	0.228	32.41	-2.378	Enterrado	
R4229W	1	Neutro	R	Básicos	0.39	89.196	-1.817	Expuesto	LXXLL motivo 3
			W	No polar	0.374	89.851	-1.613	Expuesto	
T4570A	0.682	Desestabilizadora	T	Polar (sin carga)	0.39	54.024	-0.932	Expuesto	LXXLL motivo 6
			A	No polar	0.417	45.942	-1.332	Expuesto	
K4810T	0.999	Desestabilizadora	K	Básicos	0.541	111.325	0.068	Expuesto	LXXLL motivo 6
			T	Polar (sin carga)	0.547	75.924	0.035	Expuesto	
D4973H	0.976	Neutro	D	Ácidos	0.631	90.927	-0.464	Expuesto	LXXLL motivo 6 / Rico Pro
			H	Básicos	0.567	103.192	-0.238	Expuesto	
R5030H	1	Desestabilizadora	R	Básicos	0.173	39.709	0.434	Enterrado	Dedo Zinc (C2HC pre-PHD-tipo 2)
			H	Básicos	0.221	40.254	0.162	Enterrado	
D5044N	1	Desestabilizadora	D	Ácidos	0.561	80.84	-0.707	Expuesto	Dedo Zinc (C2HC pre-PHD-tipo 2)
			N	Polar (sin carga)	0.558	81.691	-0.851	Expuesto	
R5106C	1	Desestabilizadora	R	Básicos	0.331	75.707	0.111	Expuesto	Dedo Zinc (PHD-tipo 7)
			C	Polar (sin carga)	0.331	46.458	-0.104	Expuesto	
F5125L	0.999	Desestabilizadora	F	No polar	0.091	18.244	-0.155	Enterrado	Dedo Zinc (PHD-tipo 7)
			L	No polar	0.081	14.831	0.06	Enterrado	
D5279Y	0.996	Neutro	D	Ácidos	0.373	53.807	-0.016	Expuesto	Dominio FYR C-terminal
			Y	Polar (sin carga)	0.415	88.707	0.162	Expuesto	
G5428D	1	Desestabilizadora	G	Polar (sin carga)	0.107	8.445	-0.328	Enterrado	Dominio SET
			D	Ácidos	0.113	16.24	-0.444	Enterrado	
E5434K	1	Neutro	E	Ácidos	0.313	54.751	0.427	Expuesto	Dominio SET
			K	Básicos	0.32	65.721	0.566	Expuesto	
D5462H	1	Neutro	D	Ácidos	0.089	12.81	-0.406	Enterrado	Dominio SET
			H	Básicos	0.096	17.481	-0.178	Enterrado	
T5464M	1	Neutro	T	Polar (sin carga)	0.124	17.143	0.203	Enterrado	Dominio SET
			M	No polar	0.111	22.151	0.061	Enterrado	
K5493N	1	Desestabilizadora	K	Básicos	0.149	30.567	1.493	Enterrado	Dominio SET
			N	Polar (sin carga)	0.169	24.698	1.257	Enterrado	
R5501Q	1	Desestabilizadora	R	Básicos	0.459	105.134	0.412	Expuesto	Dominio SET
			Q	Polar (sin carga)	0.468	83.585	0.403	Expuesto	
D5511A	1	Neutro	D	Ácidos	0.248	35.78	-0.992	Enterrado	Dominio SET
			A	No polar	0.233	25.71	-1.025	Enterrado	
H5521Q	0.999	Neutro	H	Básicos	0.424	77.071	0.784	Expuesto	Dominio Post-SET
			Q	Polar (sin carga)	0.439	78.405	0.811	Expuesto	
M5536T	0.998	Desestabilizadora	M	No polar	0.214	42.841	-0.961	Enterrado	Dominio Post-SET
			T	Polar (sin carga)	0.231	32.012	-1.219	Enterrado	

RSA: Área de superficie relativa, ASA: Área de superficie accesible por solvente.

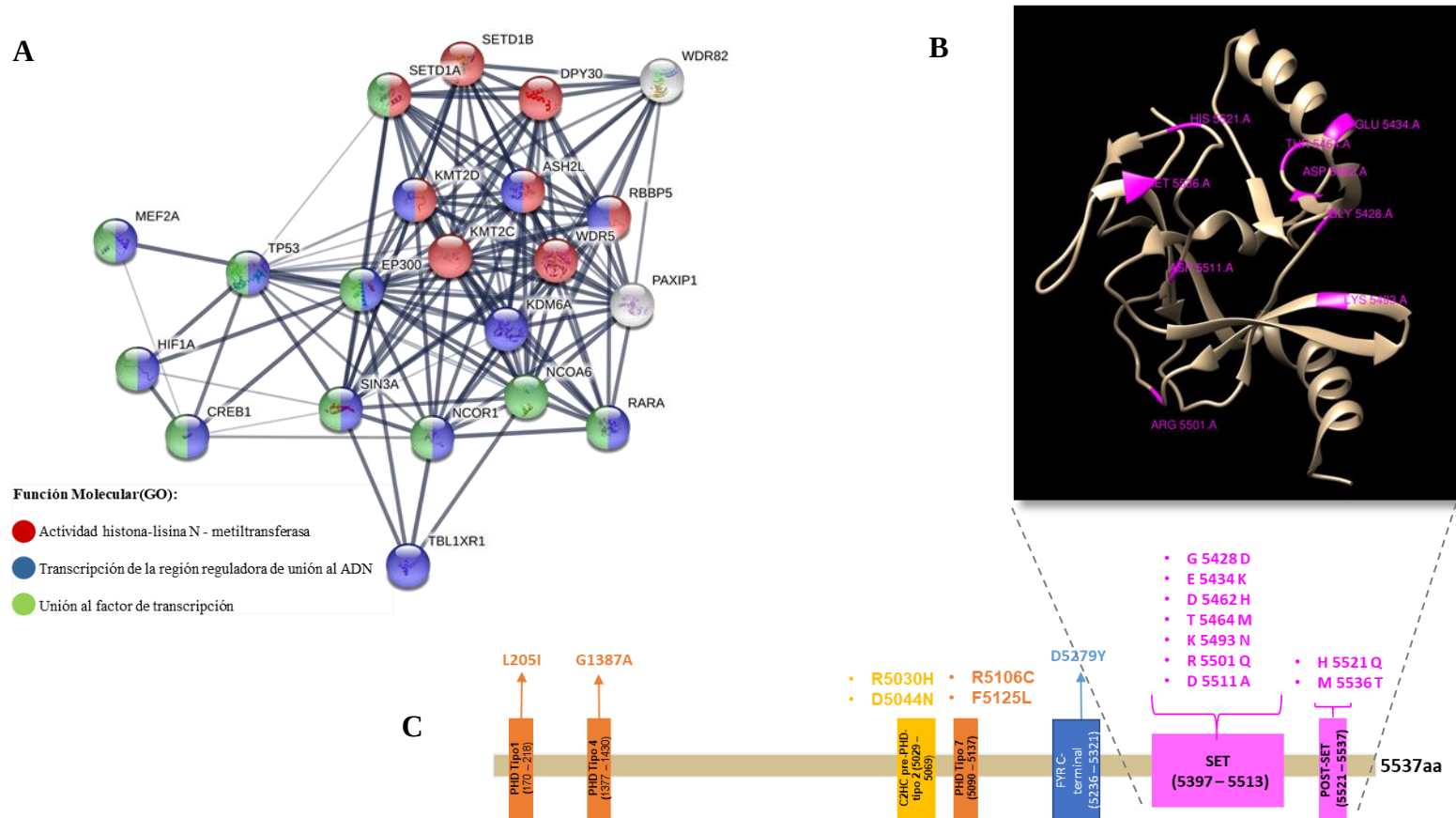


Figura 16. Localización de variantes “con pérdida de sentido” en KMT2D y su red de PPI. (A) Red de interacción proteína-proteína de KMT2D mediante STRING, se analizaron 21 nodos con 135 aristas y 12.9 como grado promedio de nodos, p -valor $< 1.0e-16$. (B) Estructura 3D de la proteína KMT2D mostrando 7 mutaciones en los dominios SET y 2 mutaciones en el dominio Post-SET. (C) Localización de los dominios proteicos de KMT2D detallado por color.

Tabla 7. Mutaciones puntuales predichas “dañinas” en KMT2C, incluyendo el cambio de estabilidad de cada mutante y cambio de accesibilidad al solvente.

Mutación	PolyPhen (Biomuta 4.0)	I Mutant 3.0	NetSurP-1.1						Familia y Dominio
			AA	Naturaleza del AA	RSA	ASA	Puntaje Z	Asignación	
E300D	0.998	Neutro	E	Ácidos	0.497	86.739	-0.907	Expuesto	Dedo Zinc PHD-tipo1
			D	Ácidos	0.431	62.064	-0.813	Expuesto	
Y366S	0.999	Desestabilizadora	Y	Polar (sin carga)	0.229	48.937	-0.14	Enterrado	Dedo Zinc PHD-tipo 2/tipo RING
			S	Polar (sin carga)	0.278	32.546	0.293	Enterrado	
K431R	0.543	Neutro	K	Básicos	0.364	74.813	0.866	Expuesto	Dedo Zinc PHD-tipo 3
			R	Básicos	0.374	85.623	0.878	Expuesto	
D1732G	1	Desestabilizadora	D	Ácidos	0.473	68.217	-1.867	Expuesto	rico en Gln
			G	Polar (sin carga)	0.501	39.468	-1.908	Expuesto	
Q1782P	0.999	Neutro	Q	Polar (sin carga)	0.458	81.71	-1.614	Expuesto	Región Espiral / rico en Gln
			P	No polar (hidrofóbico)	0.512	72.653	-1.462	Expuesto	
R1906Q	1	Desestabilizadora	R	Básicos	0.486	111.271	-1.386	Expuesto	Rico en Pro
			Q	Polar (sin carga)	0.467	83.442	-1.355	Expuesto	
P2101A	0.91	Desestabilizadora	P	No polar (hidrofóbico)	0.325	46.089	-2.58	Expuesto	Rico en Pro
			A	No polar (hidrofóbico)	0.368	40.543	-1.901	Expuesto	
S2146Y	0.828	Neutro	S	Polar (sin carga)	0.472	55.272	-1.806	Expuesto	Rico en Pro
			Y	Polar (sin carga)	0.437	93.301	-1.992	Expuesto	
R2454K	0.994	Desestabilizadora	R	Básicos	0.47	107.653	-2.031	Expuesto	Rico en Pro
			K	Básicos	0.591	121.589	-2.055	Expuesto	
R2524K	0.816	Desestabilizadora	R	Básicos	0.517	118.485	-1.397	Expuesto	Rico en Pro
			K	Básicos	0.494	101.513	-1.479	Expuesto	
H2590R	0.454	Neutro	H	Básicos	0.351	63.901	-1.649	Expuesto	Rico en Pro
			R	Básicos	0.43	98.378	-1.601	Expuesto	
L3053I	0.999	Desestabilizadora	L	No polar (hidrofóbico)	0.243	44.475	-0.587	Enterrado	Rico en Gln
			I	No polar (hidrofóbico)	0.261	48.322	-0.476	Enterrado	
L3058I	0.999	Neutro	L	No polar (hidrofóbico)	0.229	42.021	-1.421	Enterrado	Región Espiral / rico en Gln
			I	No polar (hidrofóbico)	0.225	41.625	-0.627	Enterrado	

Mutación	PolyPhen (Biomuta 4.0)	I Mutant 3.0	NetSurP-1.1						Familia y Dominio
			AA	Naturaleza del AA	RSA	ASA	Puntaje Z	Asignación	
Q3147H	0.992	Desestabilizadora	Q	Polar (sin carga)	0.458	81.781	-1.832	Expuesto	Rico en Gln
			H	Básicos	0.342	62.228	-1.874	Expuesto	
D3152Y	1	Neutro	D	Ácidos	0.713	102.686	-1.97	Expuesto	Rico en Gln
			Y	Polar (sin carga)	0.668	142.858	-2.085	Expuesto	
V3307M	0.991	Neutro	V	No polar (hidrofóbico)	0.396	60.881	-2.196	Expuesto	Rico en Gln y Pro
			M	No polar (hidrofóbico)	0.437	87.484	-1.916	Expuesto	
P3473Q	0.992	Neutro	P	No polar (hidrofóbico)	0.355	50.318	-1.977	Expuesto	Rico en Gln
			Q	Polar (sin carga)	0.446	79.656	-2.444	Expuesto	
D4423Y	1	Neutro	D	Ácidos	0.592	85.35	-1.799	Expuesto	Dedo Zinc C2HC pre-PHD-tipo 2
			Y	Polar (sin carga)	0.536	114.522	-1.829	Expuesto	
A4488T	1	Desestabilizadora	A	No polar (hidrofóbico)	0.035	3.857	-1.275	Enterrado	Dedo Zinc PHD-tipo 8
			T	Polar (sin carga)	0.046	6.352	-0.297	Enterrado	
A4488V	1	Neutro	A	No polar (hidrofóbico)	0.035	3.857	-1.275	Enterrado	Dedo Zinc PHD-tipo 9
			V	No polar (hidrofóbico)	0.034	5.303	-1.052	Enterrado	
R4584Q	1	Desestabilizadora	R	Básicos	0.094	21.618	0.082	Enterrado	Dominio FYR N-terminal
			Q	Polar (sin carga)	0.093	16.592	0.129	Enterrado	
R4584W	1	Desestabilizadora	R	Básicos	0.094	21.618	0.082	Enterrado	Dominio FYR N-terminal
			W	No polar (hidrofóbico)	0.086	20.635	-0.299	Enterrado	
S4850L	1	Neutro	S	Polar (sin carga)	0.119	13.97	-0.583	Enterrado	Dominio SET
			L	No polar (hidrofóbico)	0.117	21.478	-0.505	Enterrado	

RSA: Área de superficie relativa, ASA: Área de superficie accesible por solvente.

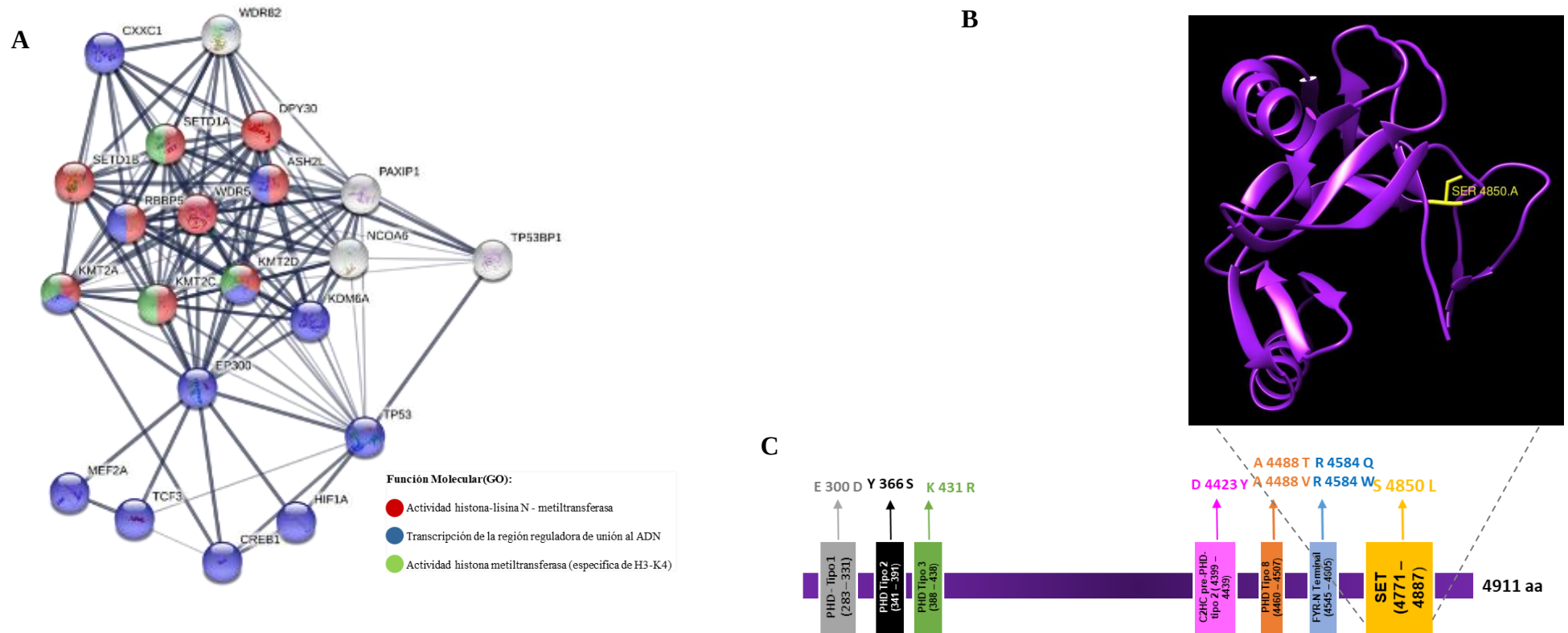


Figura 17. Localización de variantes “con pérdida de sentido” en KMT2C y su red de PPI (A) Red de interacción proteína-proteína de KMT2C mediante STRING, se analizaron 21 nodos con 133 aristas y 12.7 como grado promedio de nodos, p -valor $< 1.0e-16$. **(B)** Estructura 3D de la proteína KMT2C mostrando la mutación S4850L en el dominio SET. **(C)** Localización de los dominios proteicos de KMT2C detallado por color.

Tabla 8. Mutaciones puntuales predichas “dañinas” en CREBBP, incluyendo el cambio de estabilidad de cada mutante y cambio de accesibilidad al solvente

Mutación	PolyPhen (Biomuta 4.0)	I Mutant 3.0	NetSurP-1.1						Familia y Dominio
			AA	Naturaleza del AA	RSA	ASA	Puntaje	Asignación	
A259T	0.998	Neutro	A	No polar (hidrofóbico)	0.139	15.318	-2.153	Enterrado	Interacción con SRCAP
			T	Polar (sin carga)	0.136	18.863	-1.566	Enterrado	
R625H	1	Desestabilizadora	R	Básicos	0.372	85.234	-1.341	Expuesto	Dominio KIX
			H	Básicos	0.454	82.601	-1.091	Expuesto	
G1246D	1	Desestabilizadora	G	Polar (sin carga)	0.552	43.474	-1.446	Expuesto	rico en Cys/His
			D	Ácidos	0.565	81.373	-1.352	Expuesto	
R1319Q	0.997	Desestabilizadora	R	Básicos	0.38	86.928	-0.065	Expuesto	rico en Cys/His
			Q	Polar (sin carga)	0.377	67.404	0.093	Expuesto	
L1338S	0.996	Desestabilizadora	L	No polar (hidrofóbico)	0.033	6.042	0.086	Enterrado	Dominio CBP/p300-type HAT / rico en Cys/His
			S	Polar (sin carga)	0.034	3.985	-0.045	Enterrado	
F1345C	1	Desestabilizadora	F	No polar (hidrofóbico)	0.142	28.459	0.871	Enterrado	Dominio CBP/p300-type HAT / rico en Cys/His
			C	Polar (sin carga)	0.159	22.338	0.59	Enterrado	
R1392Q	0.999	Desestabilizadora	R	Básicos	0.216	49.487	0.189	Enterrado	CBP/p300-type HAT / rico en Cys/His
			Q	Polar (sin carga)	0.217	38.685	0.179	Enterrado	
G1404V	1	Neutro	G	Polar (sin carga)	0.47	36.965	-0.278	Expuesto	Dominio CBP/p300-tipo HAT / rico en Cys/His
			V	No polar (hidrofóbico)	0.498	76.619	-0.1	Expuesto	
I1437M	0.986	Desestabilizadora	I	No polar (hidrofóbico)	0.091	16.909	0.237	Enterrado	Dominio CBP/p300- tipo HAT / rico en Cys/His
			M	No polar (hidrofóbico)	0.091	18.189	0.22	Enterrado	
F1439L	0.996	Desestabilizadora	F	No polar (hidrofóbico)	0.201	40.361	-0.284	Enterrado	Dominio CBP/p300- tipo HAT / rico en Cys/His
			L	No polar (hidrofóbico)	0.199	36.51	-0.361	Enterrado	
I1471M	0.997	Desestabilizadora	I	No polar (hidrofóbico)	0.018	3.293	1.065	Enterrado	Dominio CBP/p300- tipo HAT/ Interacción con TRERF1 / rico en
			M	No polar (hidrofóbico)	0.018	3.542	1.063	Enterrado	
P1488R	1	Desestabilizadora	P	No polar (hidrofóbico)	0.094	13.282	-0.086	Enterrado	Dominio CBP /p300- tipo HAT / Interacción con TRERF1
			R	Básicos	0.102	23.312	-0.1	Enterrado	
I1515M	0.997	Desestabilizadora	I	No polar (hidrofóbico)	0.083	15.392	0.803	Enterrado	Dominio CBP/p300- tipo HAT / Interacción con TRERF1
			M	No polar (hidrofóbico)	0.084	16.788	0.779	Enterrado	
D1543H	1	Desestabilizadora	D	Ácidos	0.375	54.023	-1.374	Expuesto	Dominio CBP/p300- tipo HAT / Interacción con TRERF1
			H	Básicos	0.379	68.867	-1.384	Expuesto	
V1634A	0.997	Desestabilizadora	V	No polar (hidrofóbico)	0.142	21.779	0.633	Enterrado	Dominio CBP/p300- tipo HAT / Interacción con TRERF1
			A	No polar (hidrofóbico)	0.126	13.896	0.331	Enterrado	
R1682C	1	Desestabilizadora	R	Básicos	0.247	56.471	-0.232	Enterrado	Dominio CBP/p300- tipo HAT / Interacción con TRERF1
			C	Polar (sin carga)	0.25	35.03	-0.273	Enterrado	
A1782T	0.999	Neutro	A	No polar (hidrofóbico)	0.027	2.931	-0.354	Enterrado	Dedo de Zinc TAZ-tipo 2 / Interacción con TRERF1
			T	Polar (sin carga)	0.026	3.676	-0.386	Enterrado	

RSA: Área de superficie relativa, ASA: Área de superficie accesible por solvente.

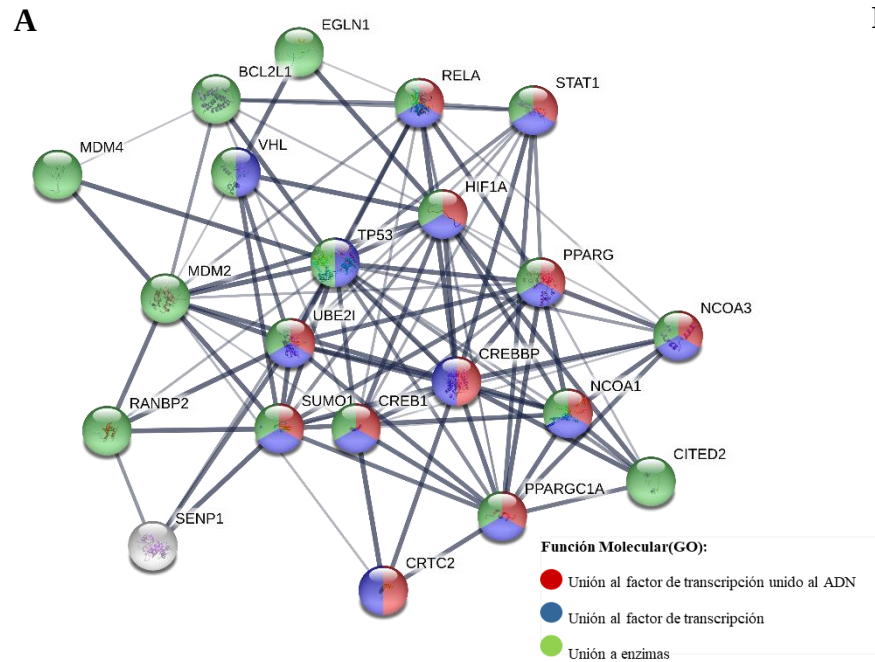
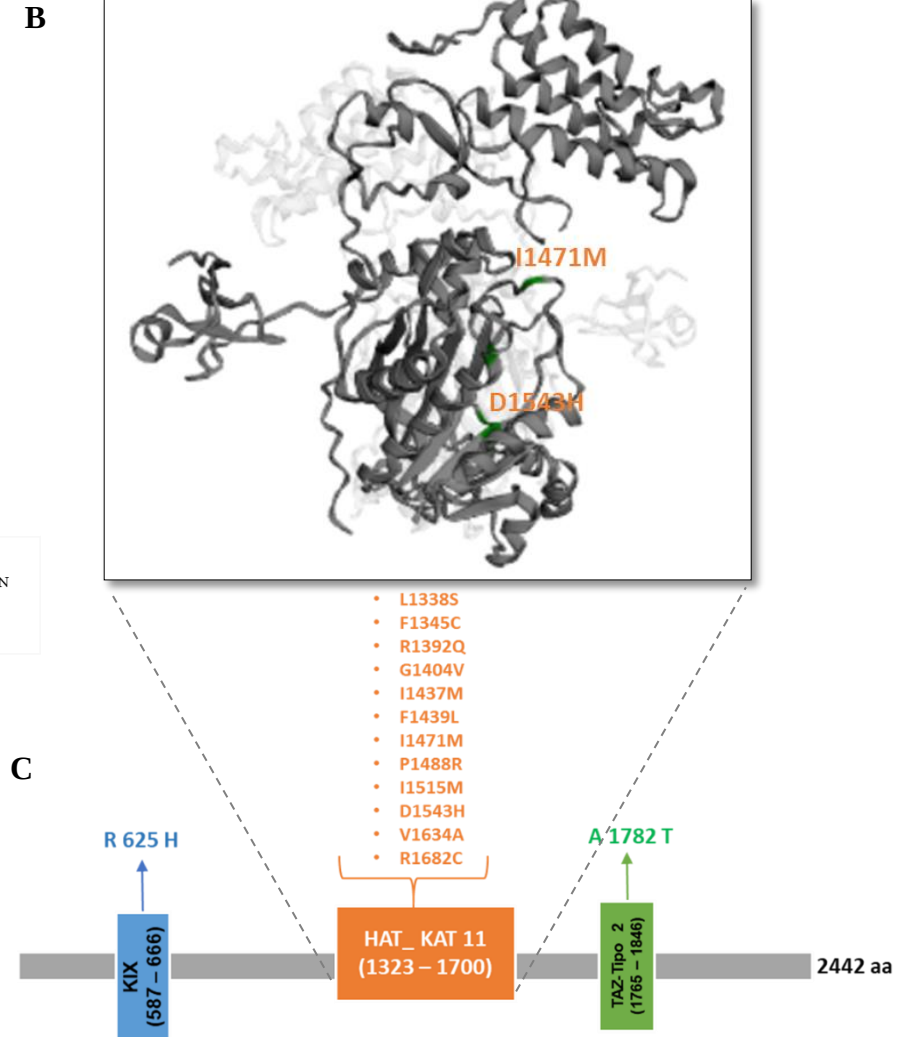


Figura 18. Localización de variantes “con pérdida de sentido” en CREBBP (proteína de unión a CREB) y su red de PPI. (A) Red de interacción proteína-proteína de CREBBP mediante STRING, se analizaron 21 nodos con 102 aristas y 9.71 como grado promedio de nodos, p -valor = $5.75e-14$ (B) Estructura 3D de la proteína KMT2C mostrando 12 mutaciones en el dominio HAT_KAT11 (C) Estructura del dominio proteico de CREBBP indicada por color.



5.6. Alteraciones en el número de copias (CNAs)

Las CNAs presentes en cada cáncer ginecológico fueron identificadas aplicando un análisis GISTIC 2.0. En conjunto, para CESC, OV, UCS y UCEC, identificamos un total de 86 genes que presentaron amplificaciones y/o deleciones, de manera significativa (valor $Q < 0.05$). Las CNAs afectaron a 26 genes en UCEC, 25 genes en CESC, 22 genes en OV y 13 genes en UCS (Tabla 9).

Las deleciones estuvieron presentes con más frecuencia en pacientes con UCS (85%), seguido de CESC (80%), UCEC (73%) y finalmente OV (59%). Por otro lado, las amplificaciones estuvieron presentes con mayor frecuencia en OV (41%), seguido de UCEC (27%), CESC (20%) y finalmente UCS (15%) (Figura 19).

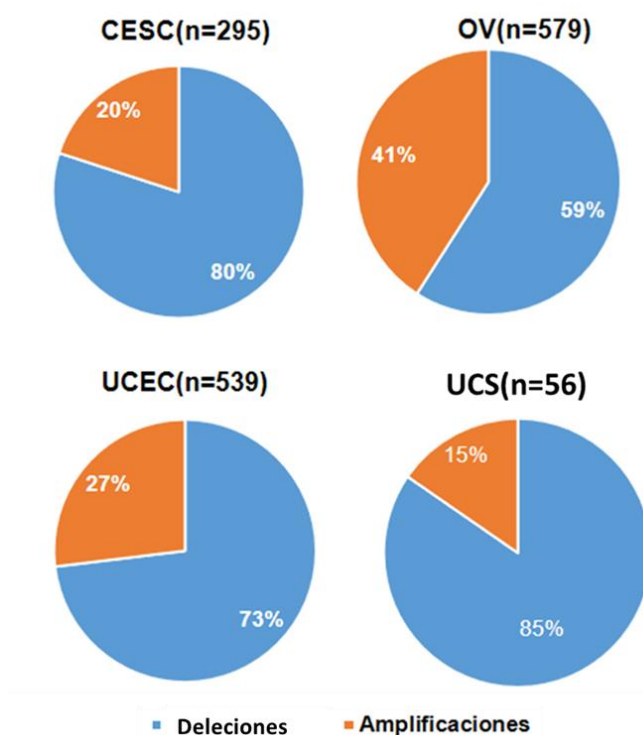


Figura 19. Porcentaje de genes que presentaron CNAs según el tipo de cáncer ginecológico. Los genes presentaron amplificaciones y deleciones identificadas por GISTIC 2.0 (valor $Q < 0.05$). CESC: Carcinoma de células escamosas cervicales y adenocarcinoma endocervical, OV: Cistadenocarcinoma seroso ovárico, UCS: Carcinosarcoma uterino, UCEC: Carcinoma Endometrial del Cuerpo Uterino, n= Número total de pacientes analizados.

Se identificaron genes compartidos con CNAs entre 3 tipos de cáncer ginecológico, mostrando distinta alteración entre un tipo de cáncer y otro. Así, los genes CDK5, EZH2, HDAC4 y KMT2C presentes en CESC, OV y UCEC (Figura 20A), mostraron deleciones en CESC y UCEC mientras que tres de ellos (CDK5, EZH2 y KMT2C) mostraron amplificaciones en OV (Tabla 9). Por otro lado, los genes PRDM2 y HDAC10 presentaron deleciones entre los grupos de CESC, UCEC y UCS (Figura 20 A y Tabla 9).

Se encontraron 17 genes con amplificación exclusiva de cada cáncer ginecológico. Así, 7 genes fueron exclusivos de OV, 5 genes de UCEC ,3 genes de CESC y finalmente, 2 genes de UCS (Figura 20 B). Además, al evaluar las similitudes entre las neoplasias, se encontró que los genes KAT6A, BRCC3 y SETDB1 estuvieron comúnmente amplificados en los grupos de UCEC y OV; CESC y OV, y CESC y UCEC, respectivamente (Figura 20 B). No se evidenciaron genes comúnmente alterados por amplificación entre 3 tipos de cáncer ginecológico.

Por otro lado, se encontró 26 genes con deleción exclusiva de cada cáncer ginecológico. Así, 11 genes fueron exclusivos de CESC, 8 genes de OV ,4 genes de UCEC y 3 genes de UCS (Figura 20 C). Además, al evaluar las similitudes entre las neoplasias, se encontró que los genes PRDM2 y HDAC10 estuvieron comúnmente delecionados en los grupos de CESC, UCEC y UCS (Figura 20 C), mientras que el gen HDAC4 estuvo comúnmente delecionado en los grupos de CESC, OV y UCEC (Figura 20 C).

De los genes con CNAs comunes entre las 4 neoplasias ginecológicas (Figura 20 B y C) , cuatro fueron identificados en la literatura como oncogenes (genes que inducen a la formación de cáncer en una célula), incluyendo EZH2 (119), KAT6A

(120) y SETDB1 (121) y USP22 (122). Mientras que CREBBP se considera como un gen supresor de tumores (genes inhiben la proliferación celular y el desarrollo de tumores)(123).

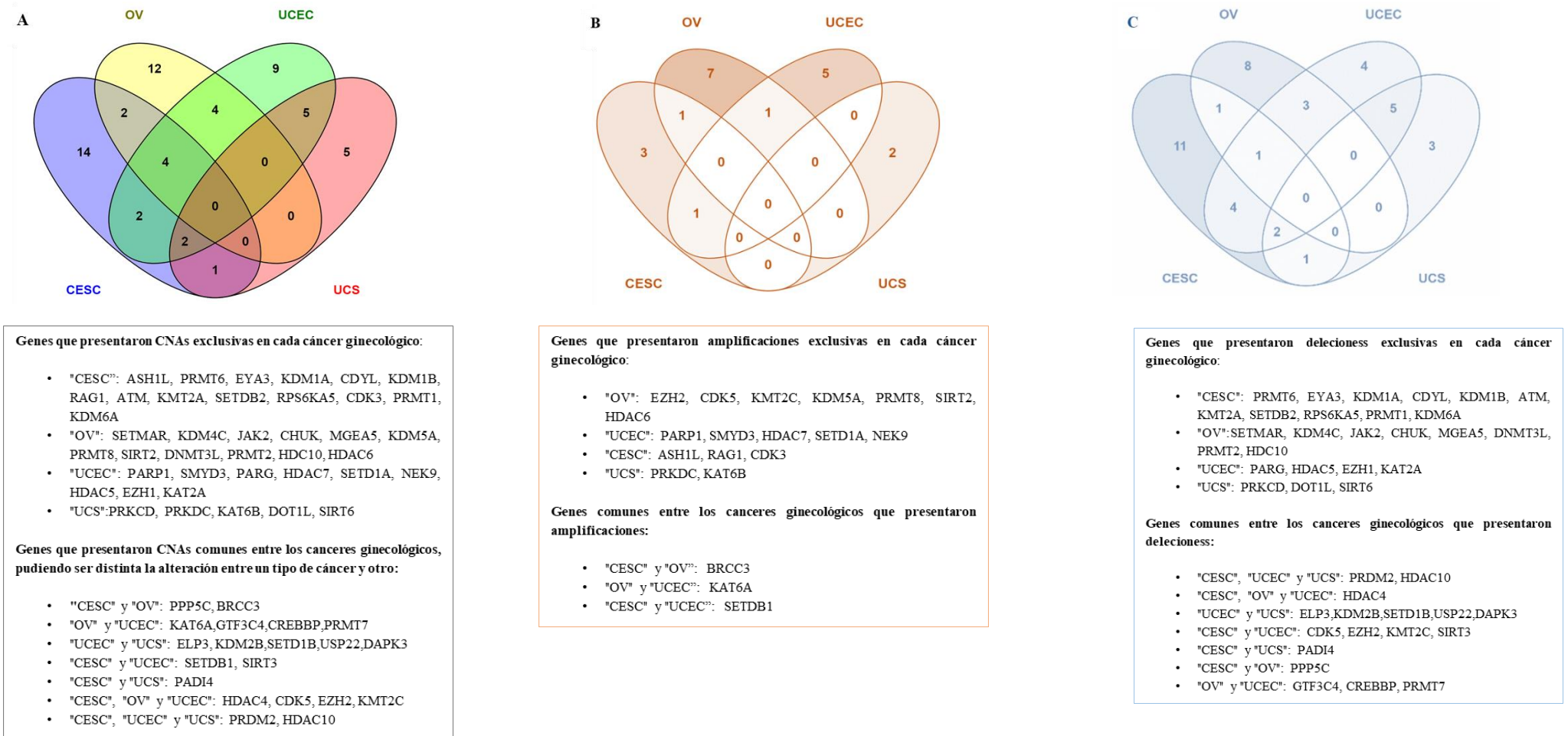


Figura 20. Diagrama de Venn que muestra el número de genes comunes entre CESC, OV, UCEC y UCS, localizados en regiones con CNA significativas. (A) Número de genes ubicados en regiones genómicas que se amplificaron y delecionaron, según el tipo de cáncer ginecológico, (B) Número de genes ubicados en regiones genómicas que se amplificaron, según el tipo de cáncer ginecológico (C) Número de genes ubicados en regiones genómicas que se delecionaron, según el tipo de cáncer ginecológico. CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, OV: Cistoadenocarcinoma seroso ovárico, UCEC: Carcinoma de endometrio del cuerpo uterino, UCS: Carcinosarcoma Uterino.

En CESC, 20 genes presentaron deleciones y se localizaron en los cromosomas 1,2,6,7,11,13,14,19,22 y X (Tabla 9) mientras que 5 genes presentaron amplificaciones, estando localizados en los cromosomas 1,11,17 y X (Tabla 9).

Con respecto a UCEC, 19 genes presentaron deleciones y se localizaron en los cromosomas 1,2,7,8,10,11,12,16,17,19,22 (Tabla 9) mientras que 7 genes presentaron amplificaciones, afectando a los cromosomas 1,8,12,16,17 (Tabla 9).

En UCS, 11 genes presentaron deleciones y se localizaban en los cromosomas 1,3,8,10,12,17,19 y 22 (Tabla 9); mientras que dos genes presentaron amplificaciones y se localizaron en los cromosomas 8 y 10 (Tabla 9).

Finalmente, para OV, 13 genes presentaron deleciones en los cromosomas 2,3,9,10,16,19,21 y 22, mientras que 9 genes presentaron amplificaciones en los cromosomas 7,8,12,19 y X (Tabla 9).

Además, del total de 86 genes con CNAs, presentes en cáncer ginecológico, 85 fueron identificados como modificadores de histonas, mientras que sólo uno (DNMT3L) estuvo involucrado en la metilación del ADN (Tabla 9).

El tamaño de la muestra de la Tabla 9 difiere de la cohorte inicial disponible en CBioportal (Tabla 2) porque no todas las muestras colectadas en TCGA fueron analizadas para determinar CNAs.

Tabla 9. Genes que presentan alteración en el número de copias (CNA), presente en cáncer ginecológico, determinados por GISTIC 2.0

CESC(n=295)		OV(n=579)		UCEC (n=539)		UCS (n=56)	
Símbolo del Gen	Citobanda	Símbolo del Gen	Citobanda	Símbolo del Gen	Citobanda	Símbolo del Gen	Citobanda
ASH1L	1q21.3	HDAC4	2q37.3	PRDM2	1p36.21	PADI4	1p36.21
SETDB1	1q21.3	SETMAR	3p26.2	SETDB1	1q21.3	PRDM2	1p36.21
PRMT6	1p13.2	EZH2	7q36.3	PARP1	1q42.3	PRKCD	3p14.2
EYA3	1p36.23	CDK5	7q36.3	SMYD3	1q42.3	ELP3	8p21.3
PRDM2	1p36.23	KMT2C	7q36.3	HDAC4	2q37.3	PRKDC	8q11.23
PADI4	1p36.23	KAT6A	8p11.21	CDK5	7q36.1	KAT6B	10q22.2
KDM1A	1p36.23	KDM4C	9p24.3	EZH2	7q36.1	SETD1B	12q24.31
HDAC4	2q37.1	JAK2	9p24.3	KMT2C	7q36.1	KDM2B	12q24.31
CDYL	6p24.2	GTF3C4	9q34.13	KAT6A	8p11.21	USP22	17p13.1
KDM1B	6p24.2	CHUK	10q25.1	ELP3	8p21.2	DOT1L	19p13.3
CDK5	7q34	MGEA5	10q25.1	GTF3C4	9q34.2	DAPK3	19p13.3
EZH2	7q34	KDM5A	12p13.33	PARG	10q21.1	SIRT6	19p13.3
KMT2C	7q34	PRMT8	12p13.33	SIRT3	11p15.5	HDAC10	22q13.31
RAG1	11p12	CREBBP	16p13.3	HDAC7	12q13.11		
ATM	11q23.3	PRMT7	16q22.1	KDM2B	12q24.33		
KMT2A	11q23.3	SIRT2	19q13.2	SETD1B	12q24.33		
SIRT3	11p15.1	PPP5C	19q13.33	SETD1A	16p11.2		
SETDB2	13q14.2	DNMT3L*	21q22.3	CREBBP	16p13.3		
RPS6KA5	14q32.31	PRMT2	21q22.3	PRMT7	16q21		
CDK3	17q25.1	HDC10	22q13.32	USP22	17p11.2		
PRMT1	19q13.33	HDAC6	Xp11.23	NEK9	17q11.2		
PPP5C	19q13.33	BRCC3	Xq28	HDAC5	17q21.31		
HDAC10	22q13.32			EZH1	17q21.31		
BRCC3	Xq28			KAT2A	17q21.31		
KDM6A	Xp11.3			DAPK3	19p13.3		
				HDAC10	22q13.32		
Total: 25 genes		Total: 22 genes		Total: 26 genes		Total: 13 genes	

Genes que presentaron amplificaciones significativas (rojo), deleciones significativas (azul) (valor Q <0.05). n= número total de muestras de tumores considerados en el análisis GISTIC 2.0.

Los genes resaltados en negrita son comúnmente compartidos en al menos dos tipos de cáncer ginecológico.

(*): Gen implicado en la metilación del ADN, los restantes son modificadores de histonas.

5.7. Identificación de Polimorfismo de Nucleótido Simple (SNPs)

Como resultado de la búsqueda de SNPs en el catálogo GWAS, se identificaron un total de 500 SNPs asociados al cáncer ginecológico ($p < 1 \times 10^{-5}$), entre ellos, cáncer cervical y cáncer endometrial. De ellos, sólo tres SNPs se encontraron en 3 de los 160 genes de estudio (Tabla 10). Los genes en cuestión fueron RPS6KA5, TET2 y EZH2, localizados en los cromosomas 14, 4 y 7, respectivamente (Tabla 10). Ningún SNP fue de pérdida de sentido. Un SNP (RPS6KA5) ha sido descrito en pacientes con carcinoma cervical respondedores a la quimioterapia neoadyuvante (ascendencia china Han). Los otros dos SNPs (TET2 y EZH2) han sido descrito en pacientes con carcinoma endometrial (ascendencia europea). No se identificaron SNP asociado con cáncer de ovario en los genes de estudio. La búsqueda de dichas variantes genéticas en la población de estudio (TCGA) resultó nula indicando que ningún SNPs estuvo presente en los pacientes con cáncer ginecológico TCGA (CESC, OC, UCEC y UC) de los proyectos TCGA.

Tabla 10. SNPs asociados a cáncer ginecológico identificados por estudios GWAS

SNP ID	Localización	Símbolo del Gen	Alelo de Riesgo	RAF	OR (95% CI)	Clase funcional	Neoplasia
rs1742101	14:90698719	RPS6KA5	G	0.5	1.92 (1.45-2.56)	Exón de transcripción no codificadora	Carcinoma Cervical
rs17035310	4:105143597	TET2	C	0.876	1.13 (1.08-1.18)	Gen aguas arriba	Carcinoma Endometrial
rs4727012	7:149043401	EZH2	C	0.891	1.18 (1.10-1.26)	Intergénica	Carcinoma Endometrial

RAF: Frecuencia de alelos de riesgo. **OR:** Odds ratios (razón de probabilidades).

5.8. Identificación de genes con expresión diferencial

Los perfiles de expresión génica para cada tipo de neoplasia y tejido normal fueron obtenidos de GEPIA y fueron empleados para la identificación de genes diferencialmente expresados (DEGs).

De los 160 genes de estudio evaluados, se determinó que 101 reguladores epigenéticos estuvieron diferencialmente expresados en neoplasias ginecológicas incluyendo 33 DEGs en UCS, 29 DEGs en UCEC, 23 DEGs en OV y 16 DEGs en CESC (Figura 21). Además, los mecanismos epigenéticos afectados por los cambios de expresión génica variaron por tipo de cáncer. Así, la mayoría de DEGs fueron modificadores de histonas incluyendo 32 genes en UCS, 29 genes en UCEC, 26 genes en OV y 16 genes en CESC (Figura 21). Por otro lado, sólo en UCS se identificó un DEG responsable de la metilación del ADN, DNMT3B, codificante de una enzima ADN-metiltransferasa (Figura 21, Tabla 4A en Anexo 3). No se encontraron DEGs implicados en la metilación del ADN en CESC, OV y UCEC.

De los datos transcriptómicos de UCS, se reconocieron 33 DEGs incluidos 10 genes sobreexpresados y 23 genes subexpresados, mientras que 127 genes no mostraron diferencias en expresión con relación a muestras normales (Figura 21). Genes como BUB1, AURKB, UHRF1, PPP1CA, SMYD3, EZH2, CDK5, EYA1 y EYA2, codantes para enzimas modificadoras de histonas y DNMT3B, codante para enzimas metiladoras del ADN, estuvieron sobreexpresados. A diferencia de los genes EZH1, JAK2, KAT2B, SETD1B, MYSM1, CDK17, KMT2C, MAP3K20, KMT2A, SETMAR, HDAC10, MGEA5, CREBBP, KMT2E, NEK9, KMT2D, ATM, KAT2A, RPS6KA5, HDAC6, SETD7, MAP3K8 y DUSP1, codantes para

enzimas modificadoras de histonas que estuvieron subexpresados (Tabla 4A en Anexo 3).

De las muestras de OV se reconocieron 26 DEGs incluidos 9 genes sobreexpresados y 17 genes subexpresados, mientras que 134 genes no mostraron diferencias en expresión con relación a muestras normales (Figura 21). Genes como AURKB, BUB1, CDK5, EYA2, PPP4C, UHRF1, NEK6, PRKCD y HDAC9, codantes para enzimas modificadoras de histonas, estuvieron sobreexpresados. A diferencia de los genes EZH1, KAT2A, JAK2, SETDB2, ATM, MYSM1, MGEA5, NEK9, KAT2B, HDAC5, MAP3K20, HDAC6, KDM6A, HDAC10, CDK3, SETD7, DUSP1, codantes para enzimas modificadoras de histonas, que estuvieron subexpresados (Tabla 5A en Anexo 3).

En CESC, se identificaron 16 DEGs, incluidos 6 genes sobreexpresados y 10 genes subexpresados, mientras que 144 no mostraron diferencias en la expresión con relación a muestras normales (Figura 21). Genes como AURKB, UHRF1, BUB1, EZH2, HASPIN y SMYD2, codantes para enzimas modificadoras de histonas, estuvieron sobreexpresados. A diferencia de los genes EZH1, MGEA5, NEK9, KDM4B, MAP3K20, KAT2B, KMT2A, ATM, DUSP1 y EYA1, codantes para enzimas modificadoras de histonas, que estuvieron subexpresados (Tabla 6A en Anexo 3).

En UCEC se identificaron 29 DEGs incluidos 8 genes sobreexpresados y 21 genes subexpresados, mientras que 131 no mostraron diferencias en su expresión con relación a muestras normales. Genes como AURKB, UHRF1, BUB1, CDK5, EZH2, SMYD2, NEK6 y EYA2, codantes enzimas modificadoras de histonas estuvieron sobreexpresados. A diferencia de los genes EZH1, JAK2, MAP3K20,

CDK17, SETD7, KAT2B, KMT2A, SETD1B, NCOA1, NEK9, HDAC5, PRDM2, CREBBP, ATM, HDAC4, MGEA5, DUSP1, MYSM1, KAT2A, HDAC6 y KMT2D, codantes para tal enzima que estuvieron subexpresados (Tabla 7A en Anexo 3).

Con el fin de determinar si existían genes comúnmente afectados en los 4 tipos de cáncer ginecológicos, a nivel de su expresión génica, se aplicó Venny 2.1, herramienta que representa las relaciones de intersección, inclusión y disyunción mediante un diagrama de Venn.

Se observaron DEGs comunes a los 4 tipos de cáncer ginecológico (resaltados en negrita en la Figura 22). Así, los genes BUB1, AURKB y UHRF1 estuvieron comúnmente sobreexpresados en los 4 tipos de cáncer investigados (Figura 22A), siendo los dos primeros genes codantes para enzimas implicadas en la fosforilación de serina / treonina / tirosina y el tercero en la ubiquitinación de lisina (Tabla 1A en Anexo 1). Estos genes están involucrados en procesos biológicos esenciales como proliferación celular, ciclo celular y regulación de la segregación cromosómica (Tabla 8 A ,10 A, 12 A & 14 A en Anexo 4).

Por el contrario, los genes ATM, MAP3K20, NEK9, DUSP1, KAT2B, MGEA5 y EZH1 estuvieron comúnmente subexpresados en los 4 tipos de cáncer investigados (Figura 22B). ATM, MAP3K20 y NEK9 codifican enzimas implicadas en la fosforilación de serina / treonina / tirosina, mientras que DUSP1 está involucrado en la desfosforilación de serina / treonina / tirosina. Así mismo, KAT2B y MGEA5 participan en la acetilación de lisina y EZH1 en la metilación de lisina (Tabla 1A en Anexo 1). Estos genes están involucrados en procesos biológicos esenciales como modificación de la cromatina, detención del ciclo celular, ciclo celular,

regulación positiva del proceso apoptótico y regulación negativa de la transcripción del promotor de la ARN polimerasa II (Tabla 9 A ,11 A, 13 A & 15 A en Anexo 4).

Otros DEGs también fueron comúnmente reconocidos en CESC, UCEC y UCS (Figura 22 A y 22 B). Así, EZH2, codante para la histona-lisina N-metiltransferasa estuvo sobreexpresado. A diferencia del gen KMT2A, codante para la histona-lisina N-metiltransferasa 2A que estuvo subexpresado, siendo ambos genes que codifican enzimas implicadas en la metilación de lisina (Tabla 1A en Anexo 1). Estos genes están involucrados en procesos biológicos esenciales como fosforilación de proteínas, modificación de la cromatina y metilación de lisina K4 en la histona H3 (Tabla 8 - 9 A & 12 -15 A en Anexo 4).

Otros siete DEGs fueron comunes entre OV, UCEC y UCS. Así, CDK5 y EYA2, genes codantes para la quinasa dependiente de ciclina 5 y la enzima ojos ausentes homólogo 2, respectivamente, estuvieron sobreexpresados. A diferencia de los genes KAT2A, HDAC6, JAK2, MYSM1 y SETD7 codantes para enzimas modificadoras de histonas que estuvieron subexpresados. CDK5 y EYA2 están implicados en la fosforilación y desfosforilación de serina / treonina / tirosina, respectivamente. KAT2A y HDAC6 participan en la acetilación y desacetilación de lisina, respectivamente; JAK2 en la fosforilación de serina / treonina / tirosina, MYSM1 en la desubiquitinación de lisina y SETD7 en la metilación de lisina (Tabla 1A en Anexo 1). Estos genes están involucrados en procesos biológicos esenciales como proliferación celular, proceso apoptótico, fosforilación de proteínas y regulación positiva de la transcripción del promotor del ARN polimerasa II (Tabla 10 -15 A en Anexo 4).

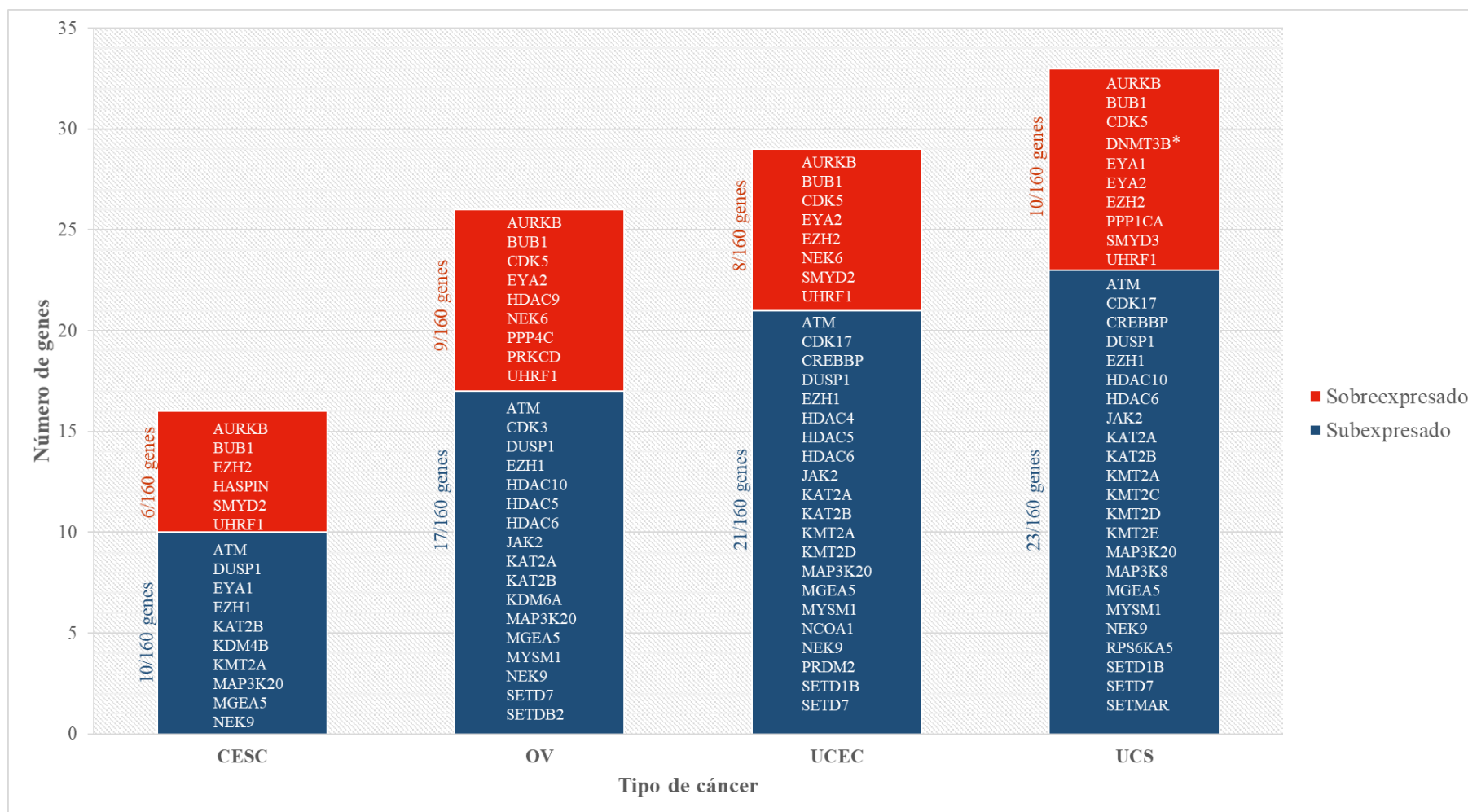
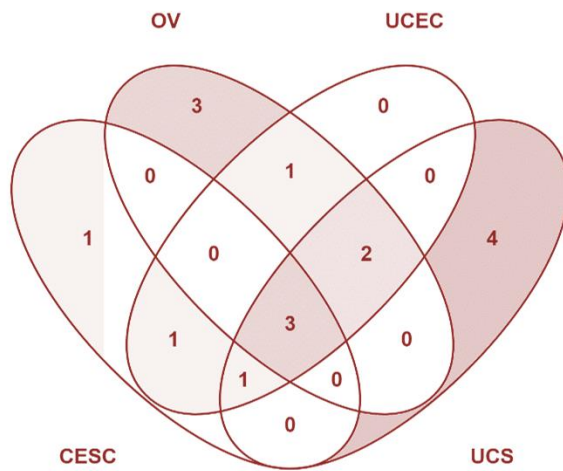


Figura 21. Genes con expresión diferencial (DEGs) según el tipo de cáncer ginecológico. Las barras rojas representan los genes sobreexpresados, en azul los genes subexpresados; con significancia estadística ($q < 0.05$). El nombre del gen que aparece con el símbolo (*) es un metilador del ADN, los demás son reguladores de histonas. Un total de 160 genes fueron analizados con GEPIA.

A



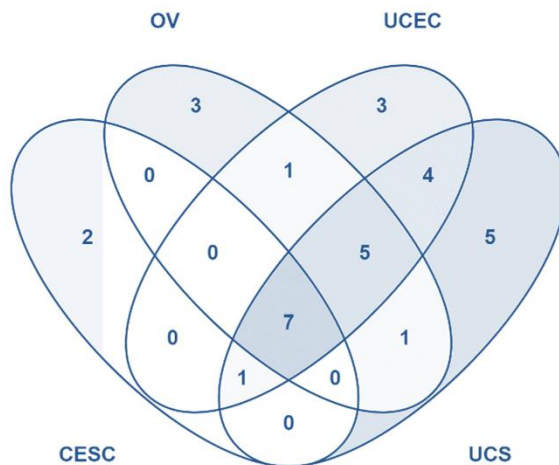
Genes sobreexpresados únicamente en un tipo de cáncer ginecológico:

- "OV": PPP4C, PRKCD y HDAC9
- "CESC": HASPIN
- "UCS": PPP1CA, SMYD3, DNMT3B y EYA1

Genes sobreexpresados comunes entre los cánceres ginecológicos:

- "CESC", "OV", "UCEC" y "UCS": AURKB, BUB1 y UHRF1
- "CESC", "UCEC" y "UCS": EZH2
- "OV", "UCEC" y "UCS": CDK5 y EYA2
- "CESC" y "UCEC": SMYD2
- "OV" y "UCEC": NEK6

B



Genes subexpresados únicamente en un tipo de cáncer ginecológico:

- "OV": CDK3, KDM6A y SETDB2
- "UCEC": HDAC4, NCOA1 y PRDM2
- "UCS": KMT2C, KMT2E, MAP3K8, RPS6KA5 y SETMAR
- "CESC": EYA1 y KDM4B

Genes subexpresados comunes entre los cánceres ginecológicos:

- "CESC", "OV", "UCEC" y "UCS": ATM, DUSP1, EZH1, KAT2B, MAP3K20, MGEA5 y NEK9
- "CESC", "UCEC" y "UCS": KMT2A
- "OV", "UCEC" y "UCS": HDAC6, JAK2, KAT2A, MYSM1 y SETD7
- "OV" y "UCS": HDAC10
- "OV" y "UCEC": HDAC5
- "UCEC" y "UCS": CDK17, CREBBP, KMT2D y SETD1B

Figura 22. Diagrama de Venn que muestra los DEGs únicos de cada neoplasia, así como aquellos comunes a CESC, OV, UCEC y UCS. (A) Genes sobreexpresados en los cánceres ginecológicos (B) Genes subexpresados en los cánceres ginecológicos. DEGs, genes expresados diferencialmente; CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, OV: Cistoadenocarcinoma seroso ovárico, UCEC: Carcinoma de endometrio del cuerpo uterino, UCS: Carcinosarcoma Uterino.

5.9. Red de interacción proteína-proteína (PPI) y caracterización funcional GO y análisis de enriquecimiento de vías KEGG de los DEGs

Empleando STRING, se reveló la red de PPI de los DEGs reconocidos por cada tipo de cáncer ginecológico. Además, para comprender las características funcionales de los DEGs se realizó un análisis de enriquecimiento de ontología génica (GO) y vías KEGG. La anotación funcional y enriquecimiento por GO permite identificar los procesos funcionales en los que están implicados un grupo de genes, mientras que el análisis de enriquecimiento de vías KEGG permite vincular las funciones de diferentes genes que participan en una determinada vía metabólica. Ambas pruebas fueron aplicadas a los DEGs reconocidos por tipo de tumor ginecológico utilizando el recurso Genecodis.

En la red de PPI de los 16 DEGs presentes en CESC, se graficaron a 13 proteínas conectadas, mientras que tres (MAP3K20, DUSP1 y NEK9) no aparecieron debido a que no participaban en el interactoma. Se destacaron cinco proteínas centrales (“central hubs”), incluidas BUB1, AURKB, EZH2 y UHRF1, todas ellas sobreexpresadas y ATM, subexpresado en CESC (Figura 23).

El análisis de enriquecimiento de ontología de genes GO proporcionada por STRING, reveló que los DEGs de CESC estaban implicados en los procesos biológicos de modificación de histonas, modificación de proteínas celulares y organización cromosómica (Figura 23).

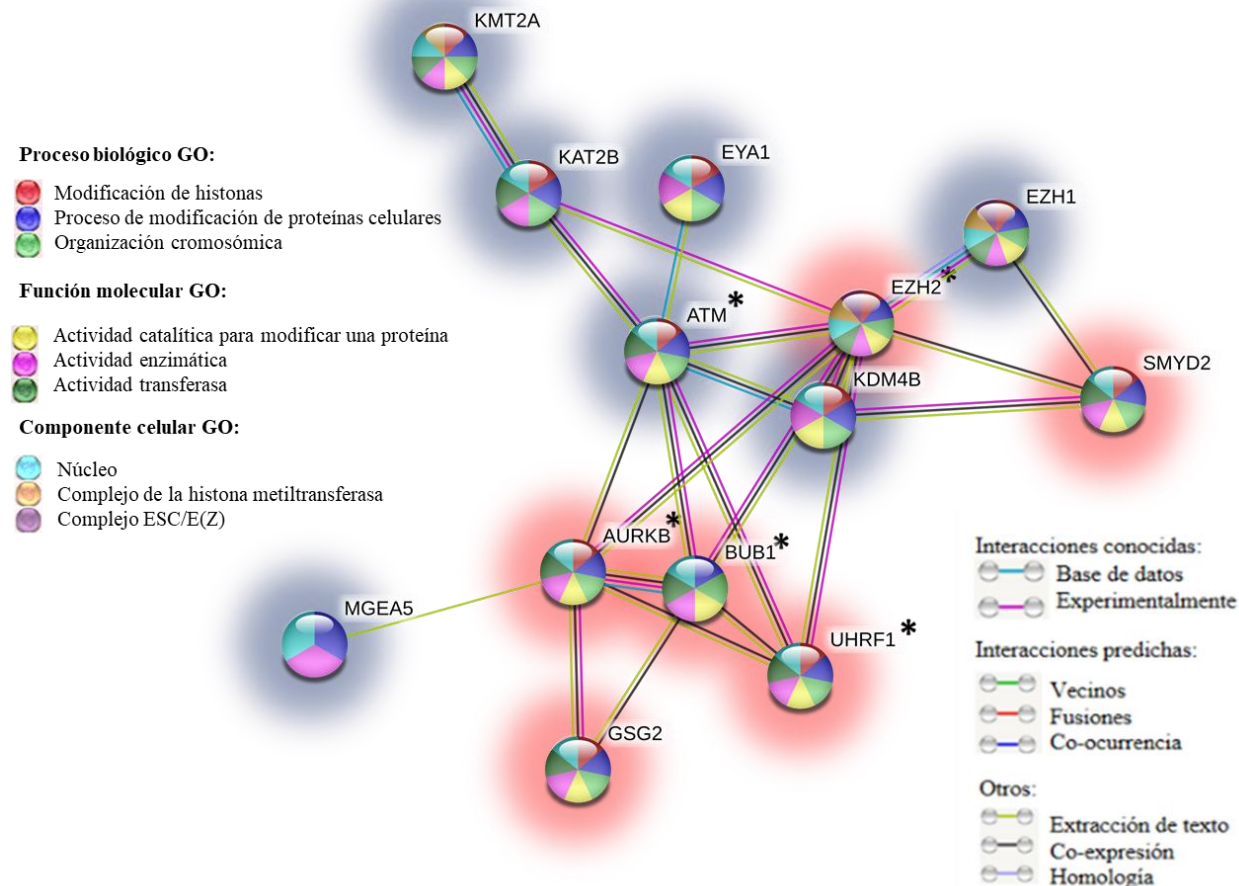


Figura 23. Red de interacción proteína-proteína (PPI) mediante STRING en CESC. Se muestra los resultados obtenidos al ingresar un conjunto de 16 genes expresados diferencialmente (DEG) en CESC. En total, se analizaron 16 nodos con 23 aristas y 2.88 como grado promedio de nodos, p -valor :8.06e-10. El grosor de las interacciones indica la evidencia en cada asociación. Se muestran los 3 primeros términos GO enriquecidos en cada categoría, tasa de descubrimiento falso (FDR) < 0.05. Las categorías de proceso biológico (PB), función molecular (FM) y componente celular (CC) son diferenciados por colores. Los nodos sombreados en rojo se refieren a genes sobreexpresados, los nodos en azul a los genes subexpresados, los nodos desconectados en la red fueron ocultados (3 en total). (*): “Central hubs”.

Por su parte, al evaluarse el enriquecimiento por GO de los 16 DEGs presentes en CESC por Genecodis, mostró que los genes sobreexpresados estaban implicados en los procesos biológicos como la regulación de la transcripción a nivel del ADN molde, la fosforilación de proteínas, proliferación celular, desarrollo embrionario y regulación negativa de la transcripción del promotor de la ARN polimerasa II (Tabla 8A en Anexo 4 y Figura 24). Cabe destacar que las proteínas centrales BUB1, AURKB, EZH2 y UHRF1 estuvieron involucradas en todos los procesos biológicos mencionados anteriormente (Tabla 8 A). Por otro lado, los genes subexpresados estuvieron involucrados en la modificación de la cromatina, ciclo celular, detención del ciclo celular, regulación positiva del proceso apoptótico y el punto de control de daños en el ADN (Tabla 9A en Anexo 4 y Figura 24). En particular, la proteína central ATM estuvo involucrada en la detención del ciclo celular, regulación positiva del proceso apoptótico y el punto de control de daños en el ADN (Tabla 9 A).

En cuanto al análisis de función molecular de los 16 DEGs presentes en CESC, los genes sobreexpresados estaban implicados en la actividad de la proteína quinasa, la actividad de proteína serina / treonina quinasa, la actividad metiltransferasa, unión a proteínas y la unión a la ARN polimerasa II (Tabla 8A en Anexo 4 y Figura 24). Cabe mencionar que las proteínas centrales AURKB, BUB1, EZH2 y UHRF1 estuvieron involucradas la actividad de la proteína quinasa, la actividad de proteína serina / treonina quinasa, la actividad metiltransferasa y unión a proteínas (Tabla 8 A). En tanto, los genes subexpresados se asociaron con la actividad de histona acetiltransferasa, la actividad de proteína serina / treonina quinasa, la actividad de histonas serina quinasa, la actividad de histona metiltransferasa (H3-K27

específica) y la actividad metiltransferasa (Tabla 9A en Anexo 4 y Figura 24). En particular, la proteína central ATM estuvo involucrado en la actividad de proteína serina / treonina quinasa y la actividad de histonas serina quinasa (Tabla 9 A).

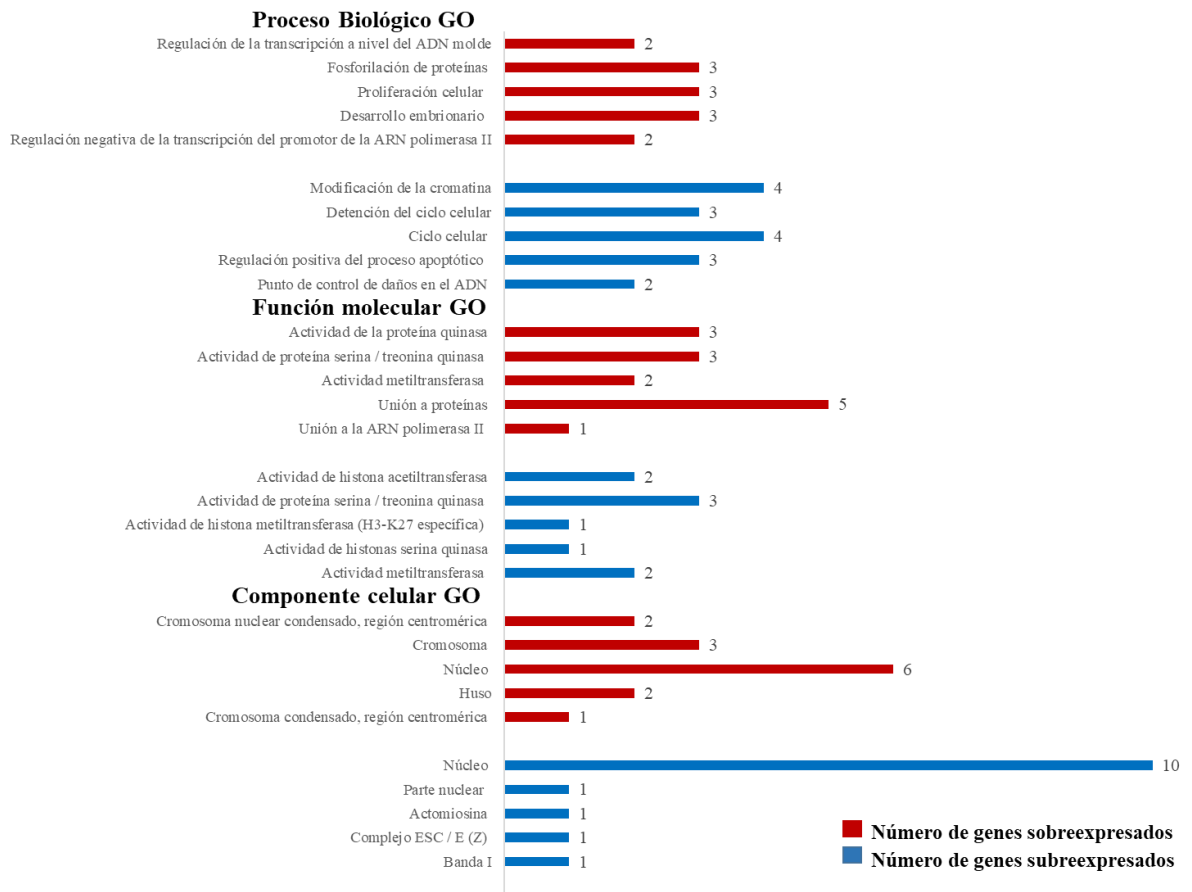


Figura 24. Análisis de enriquecimiento GO de los DEGs en CESC según el análisis Genecodis. Se muestran los cinco primeros términos enriquecidos en cada categoría (p ajustado < 0.01). Se ingresó 6 genes sobreexpresados y 10 genes subexpresados.

Al investigarse las vías KEGG enriquecidas, vinculadas a los DEGs en CESC, se apreció que los genes sobreexpresados estuvieron involucrados en las vías de la meiosis ovocitaria, en la maduración de ovocitos mediada por progesterona y en el ciclo celular (Tabla 11). Mientras que los genes subexpresados estuvieron significativamente involucrados en las vías de señalización MAPK, degradación de

la lisina, vía de señalización p53, apoptosis, uniones estrechas entre células, ciclo celular y en la vía de señalización Notch (Tabla 11).

Tabla 11. Análisis de enriquecimiento de las vías KEGG en DEGs presentes en CESC.

Vía KEGG	Clase	Número de genes	Valor p (ajustado)	Genes involucrados
Kegg:04114	Meiosis ovocitaria	1	0.0213824	BUB1
Kegg:04914	Maduración de ovocitos mediada por progesterona	1	0.028711	BUB1
Kegg:04110	Ciclo celular	1	0.0449724	BUB1
Kegg:04010	Vía de señalización MAPK	2	0.0176765	MAP3K20, DUSP1
Kegg:00310	Degradación de lisina	1	0.0332127	KMT2A
Kegg:04115	Vía de señalización p53	1	0.0339795	ATM
Kegg:04210	Apoptosis	1	0.0348054	ATM
Kegg:04530	Uniones estrechas	1	0.0373643	MAP3K20
Kegg:04110	Ciclo celular	1	0.0412823	ATM
Kegg:04330	Vía de señalización Notch	1	0.0467874	KAT2B

KEGG: Enciclopedia Kyoto de genes y genomas, (p ajustado < 0.05). Genes sobreexpresados (rojo), genes subexpresados (azul)

En la red de PPI de los 23 DEGs presentes en OV, se graficaron a 21 proteínas conectadas, mientras que dos (MAP3K20 y DUSP1) no aparecieron debido a que no presentaron evidencias de asociación. Se destacan cinco proteínas centrales (“central hubs”), incluidas KAT2A, HDAC6, KDM6A, ATM y KAT2B, todas ellas subexpresadas en OV (Figura 25).

El análisis de enriquecimiento de ontología de genes GO proporcionada por STRING, reveló que los DEGs consultados estaban implicados en los procesos biológicos de modificación de histonas, modificación de proteínas celulares y organización cromosómica (Figura 25).

Proceso biológico GO:

- Modificación de histonas
- Proceso de modificación de proteínas celulares
- Organización cromosómica

Función molecular GO:

- Actividad catalítica para modificar una proteína
- Actividad enzimática
- Actividad transferasa

Componente celular GO:

- Núcleo
- Luz nuclear
- Nucleoplasma

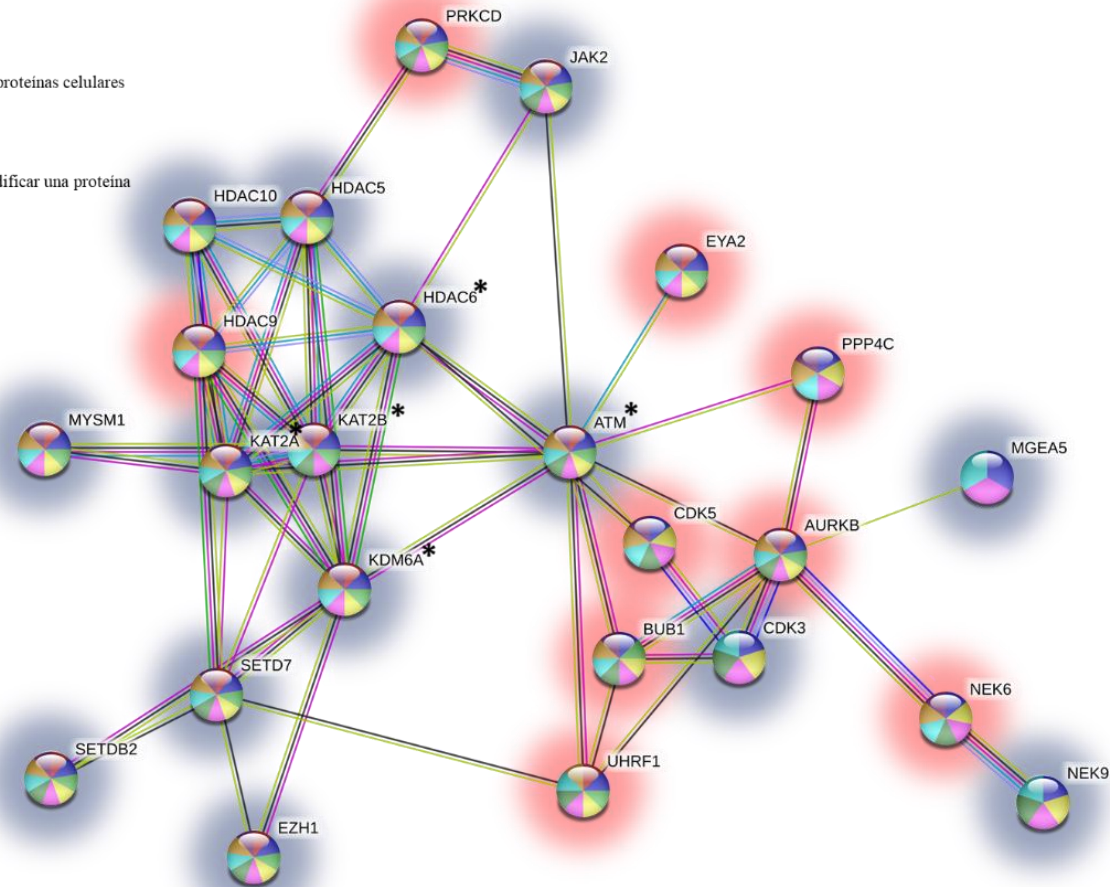


Figura 25. Red de interacción proteína-proteína (PPI) mediante STRING en OV. Se muestra los resultados obtenidos al ingresar un conjunto de 23 genes expresados diferencialmente (DEG) en CESC. En total, se analizaron 26 nodos con 56 aristas y 4.31 como grado promedio de nodos, p -valor $<1.0e-16$. El grosor de las interacciones indica la evidencia en cada asociación. Se muestran los 3 primeros términos GO enriquecidos en cada categoría, tasa de descubrimiento falso (FDR) < 0.05 . Las categorías de proceso biológico (PB), función molecular (FM) y componente celular (CC) son diferenciados por colores. Los nodos sombreados en rojo se refieren a genes sobreexpresados, los nodos en azul a los genes subexpresados, los nodos desconectados en la red fueron ocultados (2 en total).
(*): "Central hubs".

Por su parte, al evaluarse el enriquecimiento por GO de los 23 DEGs presentes en OV por Genecodis, se observó que los genes sobreexpresados estaban significativamente enriquecidos en los procesos biológicos incluyendo la proliferación celular, regulación de la segregación cromosómica, fosforilación de proteínas, ciclo celular y autofosforilación de proteínas (Tabla 10A en Anexo 4 y Figura 26). Los genes subexpresados se enriquecieron significativamente en los procesos biológicos de fosforilación de proteínas, respuesta a la radiación, proliferación celular, interacción con otras especies y regulación positiva de la transcripción del promotor del ARN polimerasa II (Tabla 11A en Anexo 4 y Figura 26). Cabe destacar que las proteínas centrales KAT2B, KAT2A, HDAC6, ATM y KDM6A estuvieron involucradas en todos los procesos biológicos mencionados anteriormente (Tabla 11 A).

En cuanto análisis de función molecular de los 23 DEGs presentes en OV, se encontró que los genes sobreexpresados se enriquecieron significativamente en la actividad de la proteína quinasa, unión a proteínas, actividad de proteína serina / treonina quinasa, unión a ATP y unión de nucleótidos (Tabla 10A en Anexo 4 y Figura 26). Mientras que los genes subexpresados se enriquecieron en la actividad de histona deacetilasa (específica para H3-K16), la actividad de histona deacetilasa dependiente de NAD (específica de H3-K14), la actividad de histona deacetilasa dependiente de NAD (específica de H4-K16), la actividad de histona deacetilasa dependiente de NAD (específica de H3-K9) y en la actividad de histona deacetilasa (Tabla 11A en Anexo 4 y Figura 26). Cabe destacar que la proteína central, HDAC6 estuvo involucrada en todos los procesos mencionados anteriormente, mientras que las proteínas centrales KAT2B, KAT2A ATM y KDM6A estuvieron involucrados

en la unión a proteínas, unión de histona deacetilasa y unión a la cromatina (Tabla 5 A y Tabla 11 A).

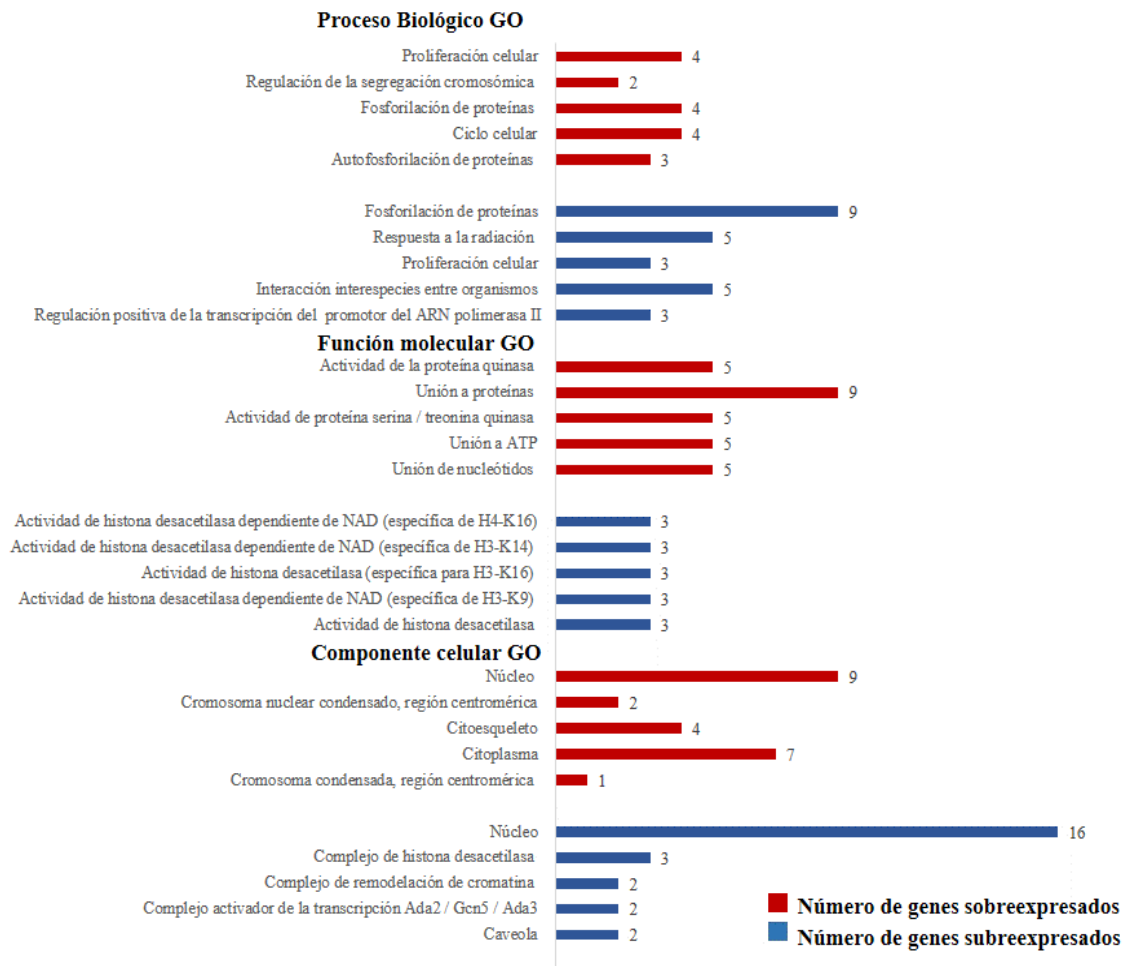


Figura 26. Análisis de enriquecimiento GO de los genes expresados diferencialmente en OV según el análisis del recurso Genecodis. Se muestra los cinco primeros términos enriquecidos en cada categoría (p ajustado < 0.01), se ingresó 9 genes sobreexpresados y 17 genes subexpresados.

Al investigarse las vías KEGG enriquecidas, vinculadas a los DEGs en OV, se apreció que los genes sobreexpresados estuvieron involucrados en la vía de señalización de la neurotrofina, enfermedad de Alzheimer, guía de axones para la formación de la red neuronal, vía de señalización de quimiocinas y uniones estrechas entre células (Tabla 12). Mientras que los genes subexpresados estuvieron

significativamente involucrados en la vía de degradación de lisina, vía de señalización Notch y vía de señalización MAPK (Tabla 12).

Tabla 12. Análisis de enriquecimiento de las vías de KEGG en genes expresados diferencialmente en OV.

Vía KEGG	Clase	Número de genes	Valor p (ajustado)	Genes involucrados
Kegg:04722	Vía de señalización de la neurotrofina	1	0.0398167	PRKCD
Kegg:05010	Enfermedad de Alzheimer	1	0.0431346	CDK5
Kegg:04360	Guía de axones	1	0.0455879	CDK5
Kegg:04062	Vía de señalización de quimiocinas	1	0.0464513	PRKCD
Kegg:04530	Uniones estrechas	1	0.0478905	PRKCD
Kegg:00310	Degradación de lisina	2	0.00188741	SETDB2, SETD7
Kegg:04330	Vía de señalización Notch	2	0.00332513	KAT2B, KAT2A
Kegg:04010	Vía de señalización MAPK	2	0.034379	MAP3K20, DUSP1

KEGG: Enciclopedia Kyoto de genes y genomas, (p ajustado < 0.05). Genes sobreexpresados (rojo), genes subexpresados (azul)

En la red de PPI de los 29 DEGs presentes en UCEC, se graficaron a 27 proteínas conectadas, mientras que dos (MAP3K20 y DUSP1) no aparecieron debido a que no presentaron evidencias de asociación. Se destacan cinco proteínas centrales (“central hubs”), incluidas CREBBP, ATM, KAT2B y KAT2A, todas ellas subexpresadas y EZH2 (sobreexpresado) en UCEC (Figura 27).

Por otro lado, el análisis de enriquecimiento de ontología de genes GO proporcionada por STRING, reveló que los DEGs consultados estaban implicados en los procesos biológicos de modificación de histonas, modificación de proteínas celulares y organización cromosómica (Figura 27).

Proceso biológico GO:

- Modificación de histonas
- Proceso de modificación de proteínas celulares
- Organización cromosómica

Función molecular GO:

- Actividad catalítica para modificar una proteína
- Actividad enzimática
- Actividad transferasa

Componente celular GO:

- Núcleo
- Nucleoplasma
- Luz nuclear

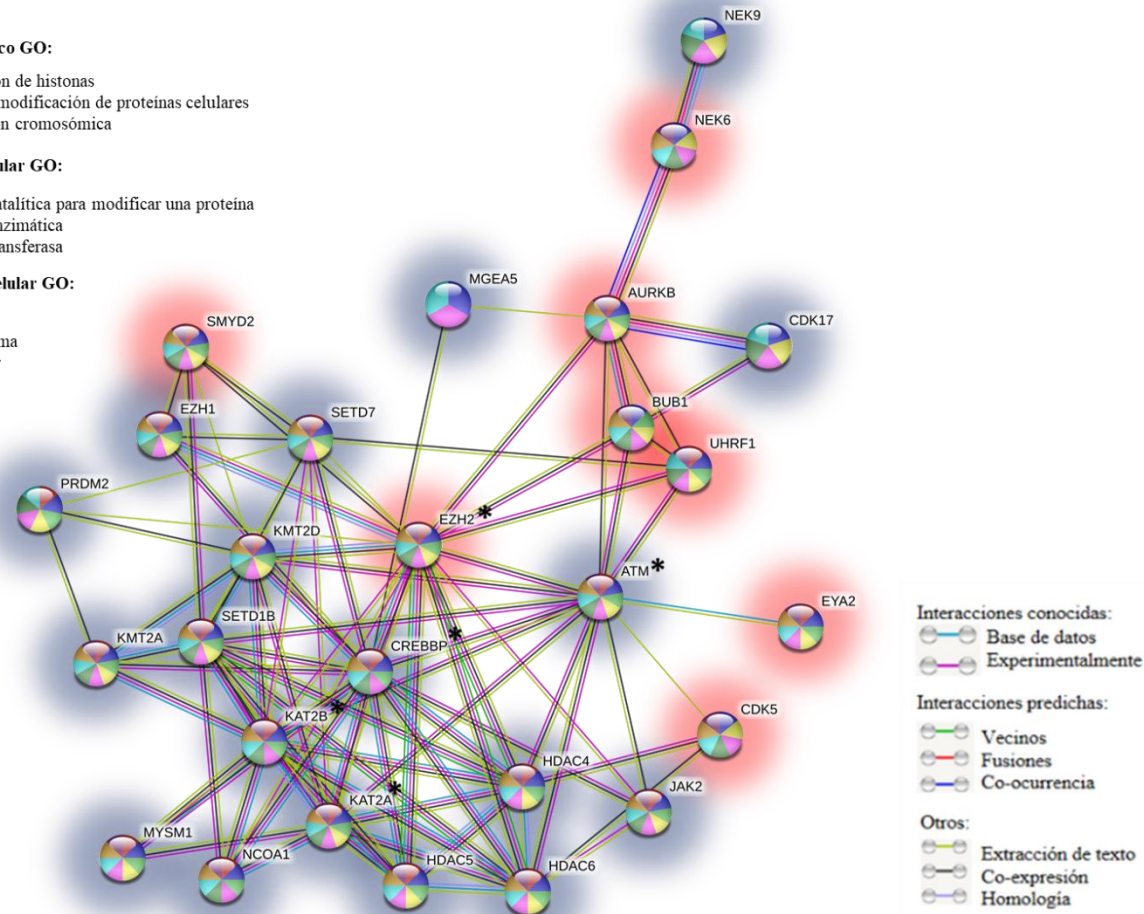


Figura 27. Red de interacción proteína-proteína (PPI) mediante STRING en UCEC. Se muestra los resultados obtenidos al ingresar un conjunto de 29 genes expresados diferencialmente (DEG) en CESC. En total, se analizaron 29 nodos con 94 aristas y 6.48 como grado promedio de nodos, $p\text{-valor} < 1.0e-16$. El grosor de las interacciones indica la evidencia en cada asociación. Se muestran los 3 primeros términos GO enriquecidos en cada categoría, tasa de descubrimiento falso (FDR) < 0.05 . Las categorías de proceso biológico (PB), función molecular (FM) y componente celular (CC) son diferenciados por colores. Los nodos sombreados en rojo se refieren a genes sobreexpresados, los nodos sombreados en azul a los genes subexpresados, los nodos desconectados en la red fueron ocultados (2 en total). (*): "Central hubs".

Por su parte, al evaluarse el enriquecimiento por GO de los 29 DEGs presentes en UCEC por Genecodis, se halló que los genes sobreexpresados estaban significativamente enriquecidos en los procesos biológicos incluyendo la regulación de la segregación cromosómica, proliferación celular, ciclo celular, autofosforilación de proteínas y modificación de la cromatina (Tabla 12A en Anexo 4 y Figura 28). En particular, la proteína central EZH2 estuvo involucrado en la modificación de la cromatina y regulación de la transcripción a nivel del ADN molde (Tabla 12 A). Los genes subexpresados se enriquecieron significativamente en los procesos biológicos incluyendo modificación de la cromatina, remodelación de la cromatina, metilación de lisina K4 en la histona H3, regulación positiva del proceso apoptótico y desacetilación de histonas (Tabla 13A en Anexo 4 y Figura 28). Cabe destacar que las proteínas centrales KAT2A, KAT2B, CREBBP y ATM estuvieron involucradas en la remodelación de la cromatina, acetilación de peptidil-lisina N-terminal y la regulación positiva del proceso apoptótico (Tabla 13 A).

El análisis de función molecular de los 29 DEGs presentes en UCEC mostró que los genes sobreexpresados se enriquecieron significativamente en la unión a proteínas, actividad de la proteína quinasa, actividad de proteína serina / treonina quinasa, unión a p53 y en la activación del receptor de acetilcolina (Tabla 12A en Anexo 4 y Figura 28). En particular, la proteína central EZH2 estuvo involucrado en la unión a proteínas (Tabla 12 A). Mientras que los genes subexpresados se enriquecieron en la actividad de histona acetiltransferasa, actividad metiltransferasa, actividad transferasa, unión a proteínas y en la actividad de histona desacetilasa (específica para H3-K16) (Tabla 13A en Anexo 4 y Figura 28). Cabe

destacar que las proteínas centrales KAT2A, KAT2B, CREBBP y ATM estuvieron involucradas en actividad de histona acetiltransferasa, actividad transferasa y unión a proteínas (Tabla 13 A).

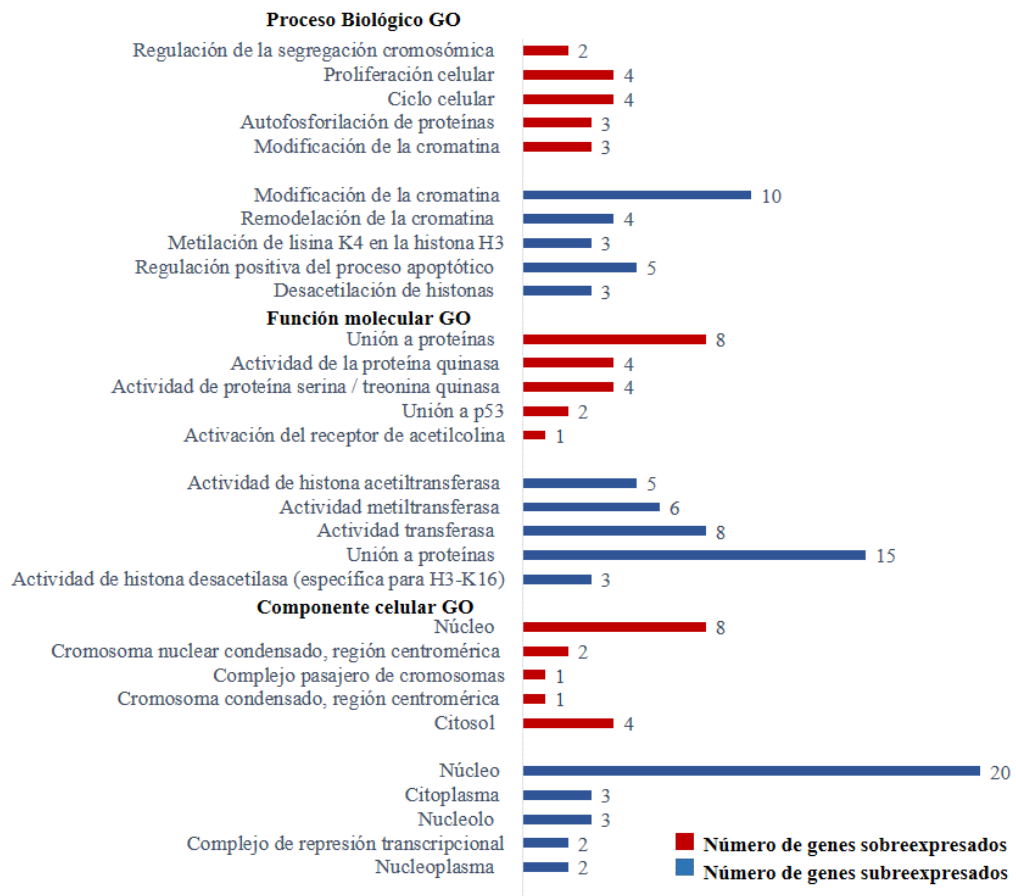


Figura 28. Análisis de enriquecimiento GO de los genes expresados diferencialmente en UCEC según el análisis del recurso Genecodis. Se muestra los cinco primeros términos enriquecidos en cada categoría (p ajustado < 0.01), se ingresó 8 genes sobreexpresados y 21 genes subexpresados.

Al investigarse las vías KEGG enriquecidas, vinculadas a los DEGs en UCEC, se apreció que los genes sobreexpresados estuvieron involucrados en la guía de axones para la formación de la red neuronal, enfermedad de Alzheimer y el ciclo celular (Tabla 13). Mientras que los genes subexpresados estuvieron significativamente involucrados en las vías de la degradación de lisina, vía de señalización Notch, ciclo

celular, vía de señalización Jak-STAT, tuberculosis y vía de señalización MAPK (Tabla 13).

Tabla 13. Análisis de enriquecimiento de las vías de KEGG en genes expresados diferencialmente en UCEC.

Vía KEGG	Clase	Número de genes	Valor p (ajustado)	Genes involucrados
Kegg:04360	Guía de axones	1	0.0369355	CDK5
Kegg:05010	Enfermedad de Alzheimer	1	0.0374938	CDK5
Kegg:04110	Ciclo celular	1	0.0473478	BUB1
Kegg:00310	Degradación de lisina	4	5.02E-07	KMT2A, SETD1B, KMT2D, SETD7
Kegg:04330	Vía de señalización Notch	3	3.42E-05	KAT2B, KAT2A, CREBBP
Kegg:04110	Ciclo celular	2	0.019743	CREBBP, ATM
Kegg:04630	Vía de señalización Jak-STAT	2	0.0226953	CREBBP, JAK2
Kegg:05152	Tuberculosis	2	0.0228019	CREBBP, JAK2
Kegg:04010	Vía de señalización MAPK	2	0.0427375	MAP3K20, DUSP1

KEGG: Enciclopedia Kyoto de genes y genomas, (p ajustado < 0.05). Genes sobreexpresados (rojo), genes subexpresados (azul)

Finalmente, en la red de PPI de los 33 DEGs presentes en UCS, se graficaron a 31 proteínas conectadas, mientras que dos (MAP3K20 y NEK9) no fueron graficadas en el interactoma. Se destacan cinco proteínas centrales (“central hubs”), incluidas CREBBP, ATM, SETD7 y KAT2B, todas ellas subexpresadas y EZH2 (sobreexpresado) en UCS (Figura 29).

Por otro lado, el análisis de enriquecimiento de ontología de genes GO proporcionada por STRING, reveló que los DEGs consultados estaban implicados en los procesos biológicos de modificación de histonas, modificación de macromoléculas y modificación de proteínas celulares (Figura 29).

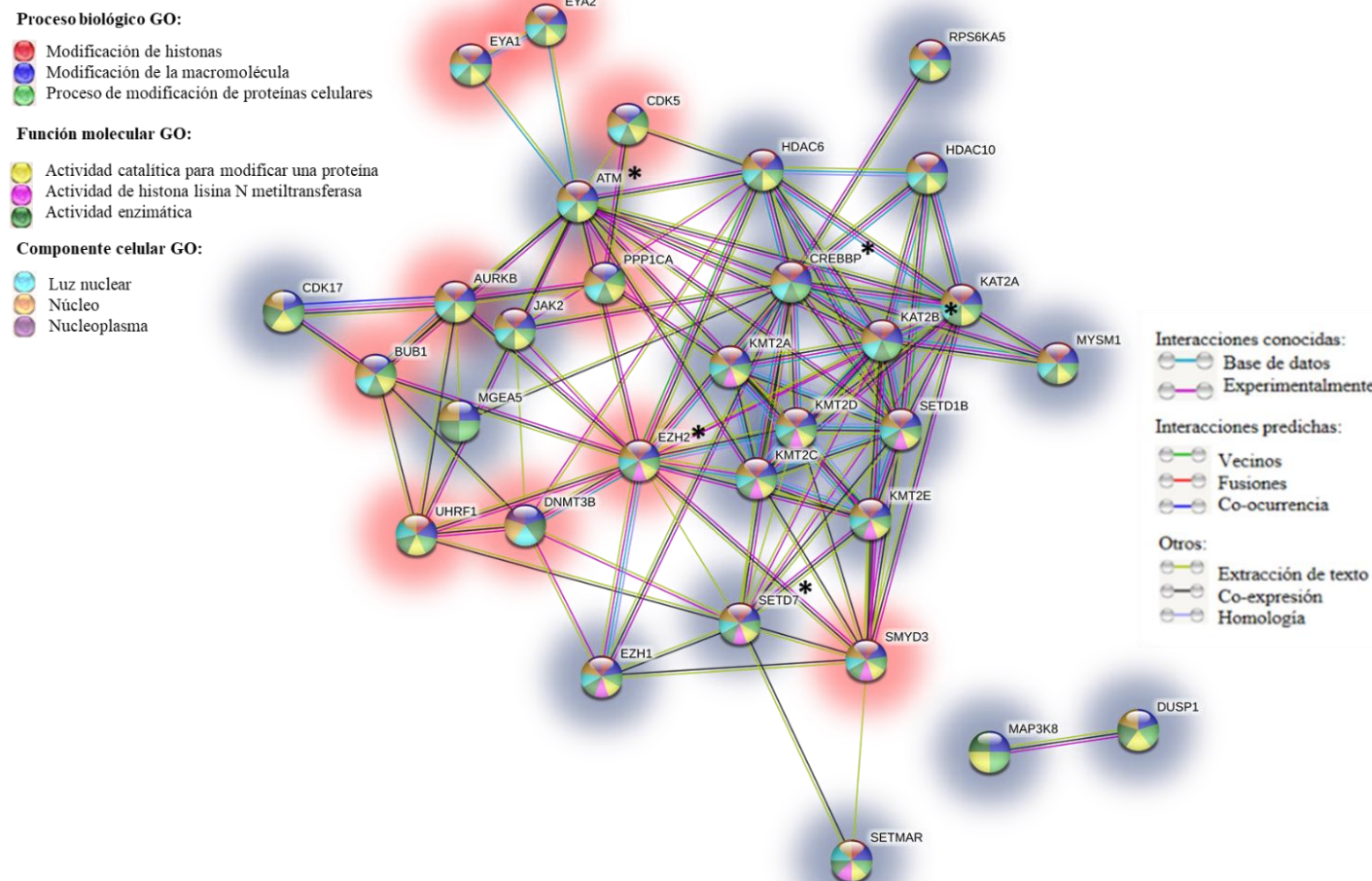


Figura 29. Red de interacción proteína-proteína (PPI) mediante STRING en UCS. Se muestra los resultados obtenidos al ingresar un conjunto de 33 genes expresados diferencialmente (DEG) en UCS. En total, se analizaron 33 nodos con 109 aristas y 6.61 como grado promedio de nodos, p -valor $< 1.0e-16$. El grosor de las interacciones indica la evidencia en cada asociación. Se muestran los 3 primeros términos GO enriquecidos en cada categoría, tasa de descubrimiento falso (FDR) < 0.05 . Las categorías de proceso biológico (PB), función molecular (FM) y componente celular (CC) son diferenciados por colores. Los nodos sombreados en rojo se refieren a genes sobreexpresados, el nodo azul a los genes subexpresados, los nodos desconectados en la red fueron ocultados (2 en total).
(*): “Central hubs”.

Al evaluarse el enriquecimiento por GO en los 33 DEGs presentes en UCS se observó que los genes sobreexpresados estaban significativamente enriquecidos en los procesos biológicos incluyendo modificación de la cromatina, desfosforilación de histonas, regulación de la segregación cromosómica, proliferación celular y ciclo celular (Tabla 14A en Anexo 4 y Figura 30). En particular, la proteína central EZH2 estuvo involucrado en la modificación de la cromatina (Tabla 14 A). Los genes subexpresados se enriquecieron significativamente en los procesos biológicos incluyendo la regulación de la transcripción a nivel del ADN molde, transducción de señales, fosforilación de proteínas, respuesta a la radiación y la proliferación celular (Tabla 15A en Anexo 4 y Figura 30). Cabe destacar que las proteínas centrales ATM, CREBBP, KAT2B y SETD7 estuvieron involucradas en la fosforilación de proteínas, respuesta a la radiación y proliferación celular (Tabla 15 A).

El análisis de función molecular de los 33 DEGs presentes en UCS mostró que los genes sobreexpresados se enriquecieron significativamente en la unión de iones metálicos, unión al ADN, actividad transferasa, unión a ATP y unión a proteínas (Tabla 14A en Anexo 4 y Figura 30). En particular, la proteína central EZH2 estuvo involucrado en la unión de iones metálicos (Tabla 14 A). Mientras que los genes subexpresados se enriquecieron en la actividad metiltransferasa, unión a proteínas, actividad transferasa, actividad de histona metiltransferasa (H3-K4 específica) y la actividad de histona acetiltransferasa (Tabla 15A en Anexo 4 y Figura 30). Cabe destacar que las proteínas centrales ATM, CREBBP, KAT2B y SETD7 estuvieron involucradas en la actividad metiltransferasa, unión a proteínas y actividad de histona acetiltransferasa (Tabla 15 A).

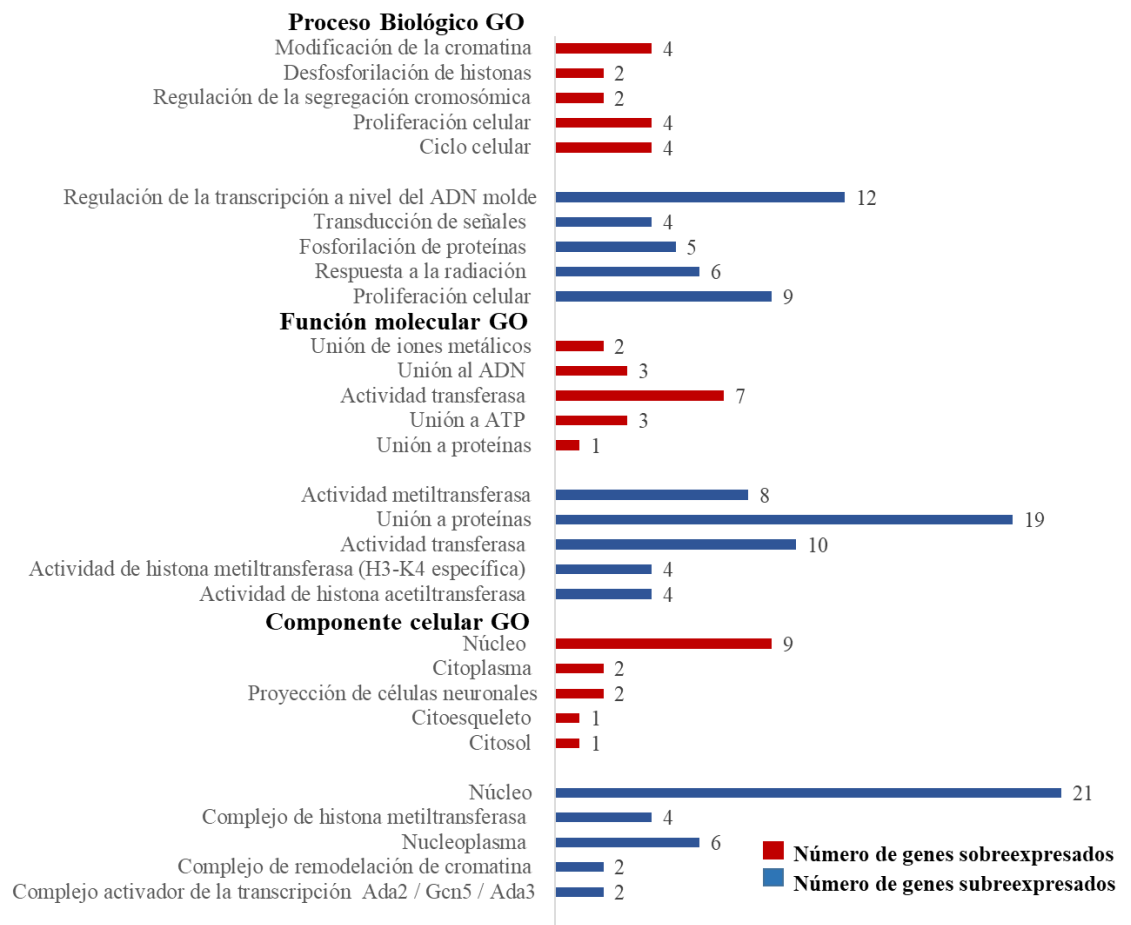


Figura 30. Análisis de enriquecimiento GO de los genes expresados diferencialmente en UCS según el análisis del recurso Genecodis. Se muestra los cinco primeros términos enriquecidos en cada categoría (p ajustado < 0.01), se ingresó 10 genes sobreexpresados y 23 genes subexpresados.

Al investigarse las vías KEGG enriquecidas, vinculadas a los DEGs en UCS, se apreció que los genes sobreexpresados estuvieron involucrados en la meiosis ovocitaria (Tabla 14). Mientras que los genes subexpresados estuvieron significativamente involucrados en las vías de degradación de lisina, vía de señalización Notch, vía de señalización MAPK, ciclo celular, vía de señalización Jak-STAT y en la tuberculosis (Tabla 14).

Tabla 14. Análisis de enriquecimiento de las vías de KEGG en genes expresados diferencialmente en UCS.

Vía KEGG	Clase	Número de genes	Valor p (ajustado)	Genes involucrados
Kegg:04114	Meiosis ovocitaria	2	0.00498732	PPP1CA, BUB1
			5.14E-14	KMT2E, KMT2A, SETD1B, KMT2C, KMT2D, SETMAR, SETD7
Kegg:00310	Degradación de lisina	7		KAT2B, KAT2A, CREBBP
Kegg:04330	Vía de señalización Notch	3	5.34E-05	
Kegg:04010	Vía de señalización MAPK	4	0.000238985	RPS6KA5, MP3K20, DUSP1, MAP3K8
Kegg:04110	Ciclo celular	2	0.0208436	CREBBP, ATM
	Vía de señalización		0.0255282	CREBBP, JAK2
Kegg:04630	Jak-STAT	2		
Kegg:05152	Tuberculosis	2	0.0266973	CREBBP, JAK2

KEGG: Enciclopedia Kyoto de genes y genomas, (p ajustado < 0.05). Genes sobreexpresados (rojo), genes subexpresados (azul)

5.10. Estudio de asociación genotipo-fenotipo

Con el fin de determinar si la expresión diferencial se correlaciona con ciertos caracteres fenotípicos, se condujo un estudio genotipo-fenotipo, por tipo de tumor. Utilizando los datos transcriptómicos y fenotípicos disponibles en las bases de datos, se examinaron las asociaciones entre la expresión génica de los DEGs con una serie de caracteres fenotípicos. Analizamos dos tipos de variables fenotípicas: datos demográficos (edad y peso) y clínico-patológicos (sobrevida global y supervivencia libre de enfermedad). Se analizó la variable peso en vez de IMC, debido a que este último no estuvo disponible en el recurso UCSC Xena.

5.10.1. Asociaciones entre la expresión génica y variables demográficas

Con el recurso UCSC Xena, se exploraron las correlaciones entre expresión de los DEGs y las variables de edad y peso, disponibles para las cuatro neoplasias (CESC, OV, UCEC y UCS). La herramienta UCSC Xena normaliza los valores de FC usando $\log_2$ para representar el cambio del nivel de expresión entre las muestras tumorales y las normales. Luego de ello asocia el $\log_2$ (conteo normalizado + 1) con algunas variables fenotípicas disponibles. De un total de 101 DEGs identificados en las neoplasias ginecológicas, se encontró que la expresión de 19 genes estuvo correlacionada con la edad al momento del diagnóstico ($p < 0.05$). Los coeficientes de correlación de Pearson presentaron tanto valores débiles ($0 < r < 0.3$ ó $0 < r < -0.3$) como bajos ($0.3 < r < 0.5$ ó $-0.3 < r < -0.5$) en el análisis (Tabla 15).

Dichos resultados mostraron que, la subexpresión de EZH1 y KDM4B se asocia directamente con la edad en las participantes con diagnóstico de CESC (Tabla 15). Así, las participantes más jóvenes tuvieron menor nivel de expresión de dichos

genes en comparación con las participantes de avanzada edad. Las pacientes más jóvenes constituyen aquellas con edad menor que la mediana y los de edad avanzada constituyen aquellas con edad mayor a la mediana.

Para OV, se determinó que la sobreexpresión de EYA2 y PRKCD se asocia inversamente con la edad mientras que la sobreexpresión de AURKB se asocia directamente con la edad. Así, las participantes más jóvenes tuvieron mayor nivel de expresión en EYA2 y PRKCD, mientras que las participantes mayores tuvieron mayor nivel de expresión en AURKB. Al contrario, la subexpresión de JAK2, KAT2B y HDAC10 se asocia inversamente con la edad mientras que la subexpresión de EZH1 y SETDB2 se asocia directamente con la edad (Tabla 15). De esa forma, pacientes con OV que tuvieron menores niveles de expresión en JAK2, KAT2B y HDAC10, fueron sujetos mayores, mientras que las participantes más jóvenes tuvieron menor nivel de expresión de EZH1 y SETDB2.

Con respecto a UCEC, se determinó que la subexpresión de JAK2, MAP3K20, CDK17, SETD7, KMT2A, NEK9 y ATM se asocia inversamente con la edad (Tabla 15). Así, las pacientes con UCEC que tuvieron menor nivel de expresión en dichos genes, fueron sujetos mayores.

Mientras que en UCS, la subexpresión de KMT2D y MAP3K8 se asocia directamente con la edad (Tabla 15). Así, las participantes más jóvenes tuvieron menor nivel de expresión de dichos genes.

No se encontraron correlaciones significativas entre la expresión génica y edad para 82 DEGs.

Tabla 15. Correlación entre la expresión de los DEGs y la edad al momento del diagnóstico de las pacientes con cáncer ginecológico

Tipo de cáncer	Gen	Coefficiente de correlación Pearson r (p-valor) *
CESC (n=307)	EZH1	r = 0.1281 (p = 0.02482)
	KDM4B	r = 0.1448 (p = 0.01108)
OV (n=308)	AURKB	r = 0.1158 (p = 0.04229)
	EYA2	r = -0.1631 (p = 0.004109)
	PRKCD	r = -0.1603 (p = 0.004799)
	EZH1	r = 0.1368 (p = 0.01625)
	JAK2	r = -0.1648 (p = 0.003722)
	SETDB2	r = 0.1287 (p = 0.02392)
	KAT2B	r = -0.1140 (p = 0.04553)
	HDAC10	r = -0.1223 (p = 0.03194)
UCEC (n=187)	JAK2	r = -0.2814 (p = 0.00009595)
	MAP3K20	r = -0.1604 (p = 0.02837)
	CDK17	r = -0.1640 (p = 0.02490)
	SETD7	r = -0.1975 (p = 0.006730)
	KMT2A	r = -0.2016 (p = 0.005661)
	NEK9	r = -0.2422 (p = 0.0008408)
	ATM	r = -0.1602 (p = 0.02849)
UCS(n=57)	KMT2D	r = 0.2726 (p = 0.04029)
	MAP3K8	r = 0.3275 (p = 0.01294)

Genes sobreexpresados (rojo), genes subexpresados (azul), n=número de pacientes analizados, CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, OV: Cistoadenocarcinoma seroso ovárico, UCEC: Carcinoma de endometrio del cuerpo uterino, UCS: Carcinosarcoma Uterino.

De un total de 101 DEGs identificados en las neoplasias ginecológicas, la expresión de sólo 5 DEGs estuvo correlacionada con el peso de las pacientes ($p < 0.05$). Todos los coeficientes de correlación de Pearson presentaron valores débiles en el análisis ($0 < r < 0.3$ o $0 < r < -0.3$) (Tabla 16).

Dichos resultados mostraron que, la subexpresión de EZH1 se asocia directamente con el peso de las participantes con diagnóstico de CESC, mientras que la subexpresión de MAP3K20 se asocia inversamente con el peso de las participantes con CESC (Tabla 16). Así, las participantes con menor peso tuvieron menor nivel de expresión de EZH1, mientras que las participantes con mayor peso tuvieron menor nivel de expresión de MAP3K20.

Para UCEC, se determinó que la sobreexpresión de EYA2 se asocia directamente con el peso de las pacientes (Tabla 16). Así, las participantes con mayor peso tuvieron mayor nivel de expresión en EYA2. Al contrario, la subexpresión de ATM y MGEA5 se asocia directamente con el peso de las pacientes con UCEC (Tabla 16). Así, las participantes con menor peso tuvieron menor nivel de expresión de ATM y MGEA5.

No se encontraron correlaciones significativas entre la expresión génica y el peso para 73 DEGs. Los 23 DEGs identificados en OV no disponían de información para el peso del paciente por lo que no se pudo determinar su correlación.

Tabla 16. Correlación entre la expresión de genes y el peso de pacientes con cáncer ginecológico

Tipo de cáncer	Gen	Coeficiente de correlación Pearson r (p-valor) *
CESC (n=279)	EZH1	r = 0.1404 (p = 0.01898)
	MAP3K20	r = -0.1598 (p = 0.007485)
UCEC (n=172)	EYA2	r = 0.1649 (p = 0.03070)
	ATM	r = 0.1587 (p = 0.03759)
	MGEA5	r = 0.1710 (p = 0.02491)

Genes sobreexpresados (rojo), genes subexpresados (azul), n=número de pacientes analizados, CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, OV: Cistoadenocarcinoma seroso ovárico, UCEC: Carcinoma de endometrio del cuerpo uterino, UCS: Carcinosarcoma Uterino.

5.10.2. Asociaciones entre la expresión génica y la sobrevida global

Utilizando GEPIA, se investigó la asociación de la expresión génica (tumoral / normal) con la sobrevida de las pacientes del estudio. Las gráficas de Kaplan Meier mostraron que, del conjunto de 101 DEGs identificados en CESC, OV, UCS y UCEC, la expresión de 5 DEGs estuvo asociada con la sobrevida global (OS)

mientras que la expresión de 2 DEGs se correlacionó con la sobrevida libre de enfermedad (DFS) (p -valor < 0.05, Figura 31). Los demás genes no mostraron asociación significativa.

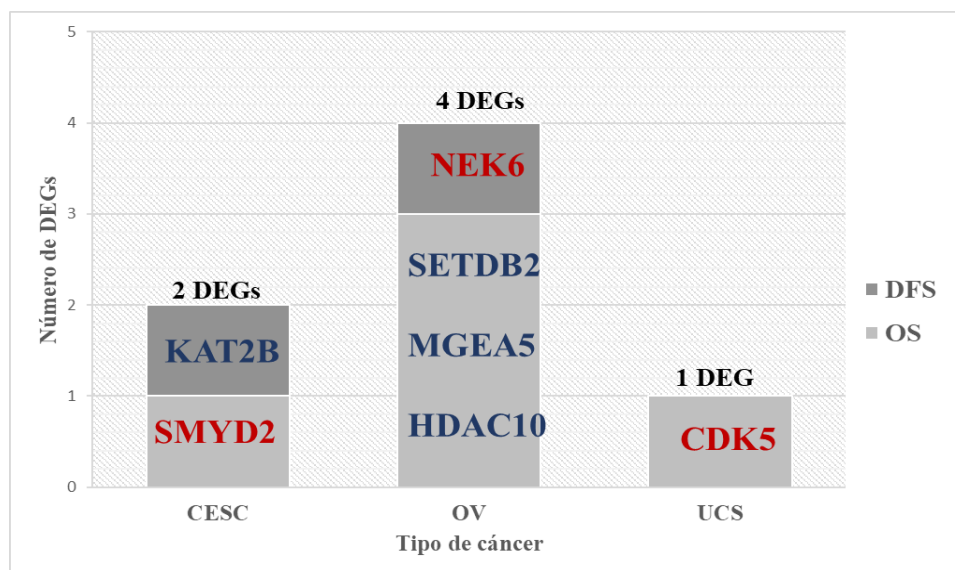


Figura 31. DEGs identificados en los cánceres ginecológicos asociados con el pronóstico de la enfermedad. Se representan a los DEGs cuya expresión génica está asociada con la sobrevida global (OS) y la sobrevida libre de enfermedad (DFS) en 3 tipos de cáncer ginecológico. Los genes en rojo se encuentran sobreexpresados, en azul subexpresados. Todos los DEGs fueron modificadores de histonas

A. Asociaciones entre la expresión génica y OS en CESC:

De acuerdo a las gráficas de Kaplan Meier, el nivel de expresión de SMYD2 se asoció con la OS de pacientes con CESC. Así, la sobreexpresión del gen SMYD2 se correlacionó con un pronóstico desfavorable en las pacientes con CESC (Figura 32). El gen SMYD2 codifica para la proteína N-lisina metiltransferasa SMYD2 que regula positivamente la citocinesis del ciclo celular y la fosforilación de proteínas (Tablas 6A & 8A en Anexo 3 y 4).

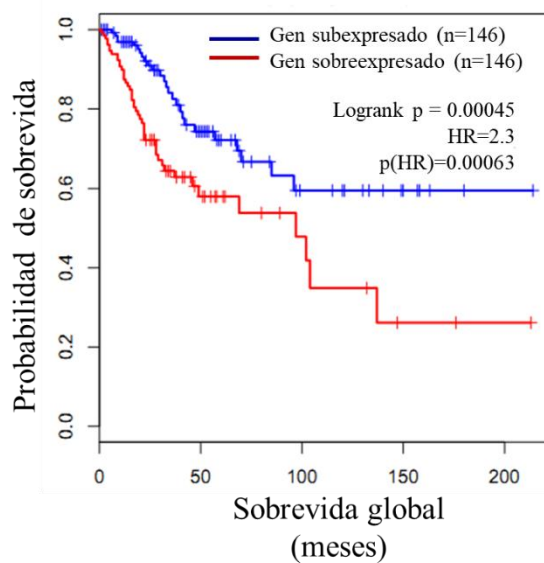


Figura 32. Gráfica de Kaplan Meier representando la OS en función de la expresión génica de SMYD2 en pacientes con CESC. El nivel de expresión de SMYD2 se relacionó significativamente con la sobrevida global de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de sobrevida global con $p = 0.00045$. Cociente de riesgo (HR) se muestra con su respectivo p valor p (HR).

B. Asociaciones entre la expresión génica y DFS en CESC:

Así mismo, las gráficas de Kaplan Meier mostraron que la subexpresión del gen KAT2B se correlacionó con una menor DFS en las pacientes con CESC (Figura 33). El gen KAT2B codifica para la proteína histona acetiltransferasa, la cual regula la detención del ciclo celular acetilando histonas (Tablas 6A & 9A en Anexo 3 y 4).

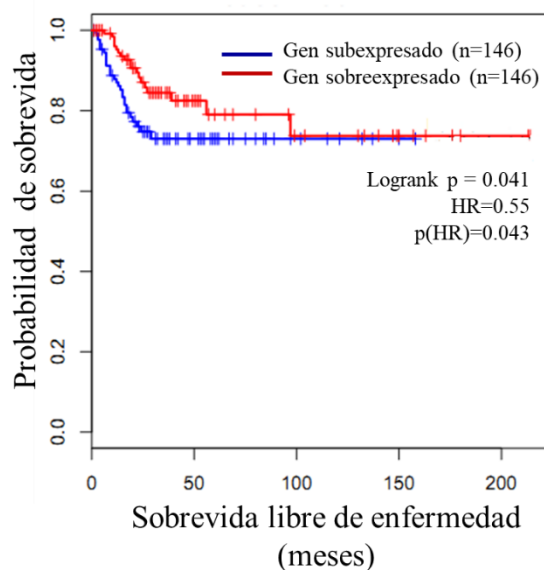


Figura 33. Gráfica de Kaplan Meier representando la DFS en función de la expresión génica de KAT2B en pacientes con CESC. El nivel de expresión de KAT2B se relacionó significativamente con la sobrevida libre de enfermedad de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de sobrevida libre de enfermedad con $p = 0.041$. Cociente de riesgo (HR) se muestra con su respectivo p valor p (HR).

C. Asociaciones entre la expresión génica y OS en OV:

Se evidenció que la expresión de los genes SETDB2 ($p = 0.025$), MGEA5 ($p = 0.027$) y HDAC10 ($p = 0.014$) estaba asociada con el pronóstico de las pacientes con OV. Según las gráficas de Kaplan Meier, la subexpresión de los genes SETDB2, MGEA5 y HDAC10 se correlaciona con un mejor pronóstico (Figuras 34 - 36). Los genes SETDB2 y HDAC10 codifican para la proteína histona-lisina N-metiltransferasa SETDB2 y la histona desacetilasa 10, respectivamente; ambas participan en la fosforilación de proteínas y regulan positivamente el proceso apoptótico (Tabla 5A & 11A en Anexo 3 y 4).

SETDB2 también está implicada en la diferenciación celular y en la actividad de histona metiltransferasa H3-K9 específica (Tabla 11A). Por su lado HDAC10 es una enzima desacetilasa de histona (específica de H3-K9, H3-K14 y H4-K16), mecanismo que típicamente silencia genes (Tabla 11A). Por otro lado, el gen MGEA5 codifica para la proteína O-GlcNAcase que participa en la regulación negativa de la transcripción y de la proliferación celular (Tablas 5A & 11A en Anexo 3 y 4).

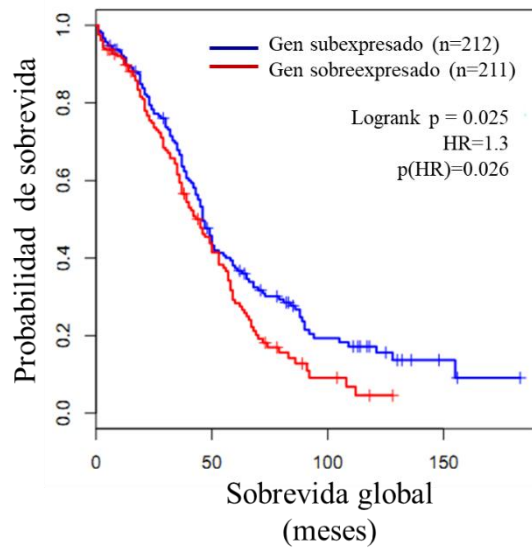


Figura 34. Gráfica de Kaplan Meier representando la OS en función de la expresión génica de SETDB2 en pacientes con OV. El nivel de expresión de SETDB2 se relacionó significativamente con la supervivencia global de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de supervivencia global con $p = 0.025$. Cociente de riesgo (HR) se muestra con su respectivo p valor $p(HR)$.

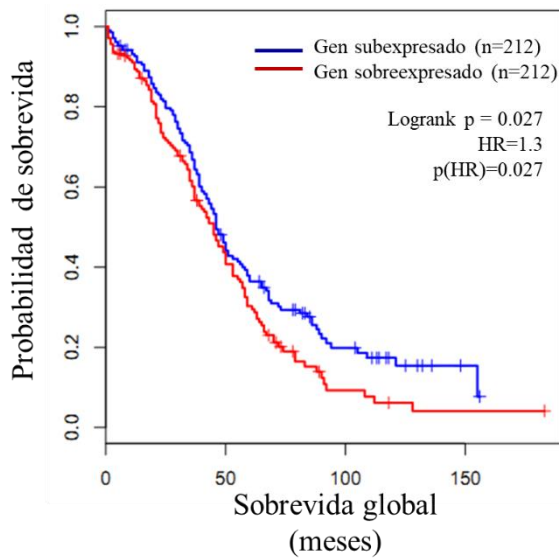


Figura 35. Gráfica de Kaplan Meier representando la OS en función de la expresión génica de MGEA5 en pacientes con OV. El nivel de expresión de MGEA5 se relacionó significativamente con la supervivencia global de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de supervivencia global con $p = 0.027$. Cociente de riesgo (HR) se muestra con su respectivo p valor $p(HR)$.

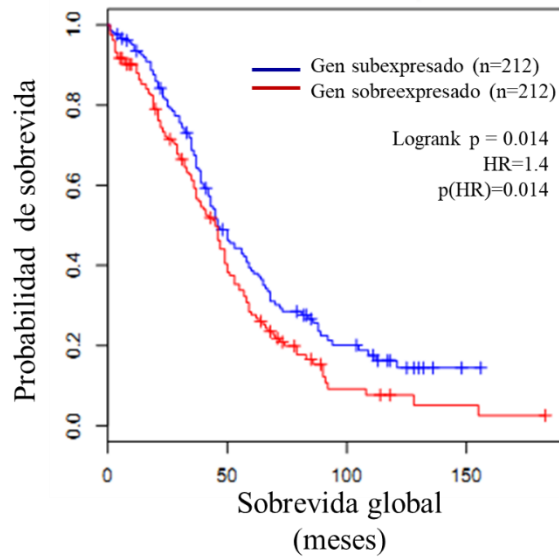


Figura 36. Gráfica de Kaplan Meier representando la OS en función de la expresión génica de HDAC10 en pacientes con OV. El nivel de expresión de HDAC10 se relacionó significativamente con la sobrevida global de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de sobrevida global con $p = 0.014$. Cociente de riesgo (HR) se muestra con su respectivo p valor $p(\text{HR})$.

D. Asociaciones entre la expresión génica y DFS en OV:

De acuerdo a las gráficas de Kaplan Meier, el nivel de expresión de NEK6 influyó en la sobrevida libre de enfermedad (DFS) de pacientes con OV. Así, la sobreexpresión del gen NEK6 se correlacionó con un mejor pronóstico (Figura 37). El gen NEK6 codifica para la proteína Serina / treonina-proteína quinasa NEK6, con fosforilación de proteínas diana mediando la citocinesis y formación del huso mitótico. NEK6 es un punto de control de la transición de G2 a Mitosis cuando existe daño en el ADN y también regula la senescencia (Tabla 5 A & 10A en Anexo 3 y 4).

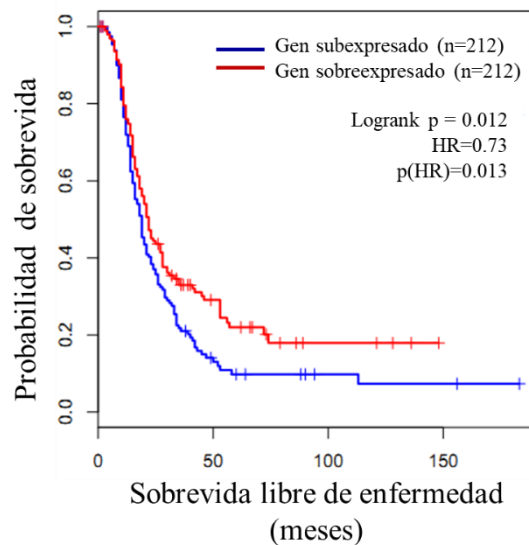


Figura 37. Gráfica de Kaplan Meier representando la DFS en función de la expresión génica de NEK6 en pacientes con OV. El nivel de expresión de NEK6 se relacionó significativamente con la sobrevida libre de enfermedad de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de sobrevida libre de enfermedad con $p = 0.012$. Cociente de riesgo (HR) se muestra con su respectivo p valor p (HR).

E. Asociaciones entre la expresión génica y OS en UCS

La gráfica de Kaplan Meier mostró que la sobreexpresión del gen CDK5 estuvo asociado con un pobre pronóstico de los pacientes con UCS (Figura 38). El gen CDK5 codifica para la proteína quinasa dependiente de ciclina 5, que regula la detención de la proliferación y ciclo celular (Tabla 4 A en Anexo 3).

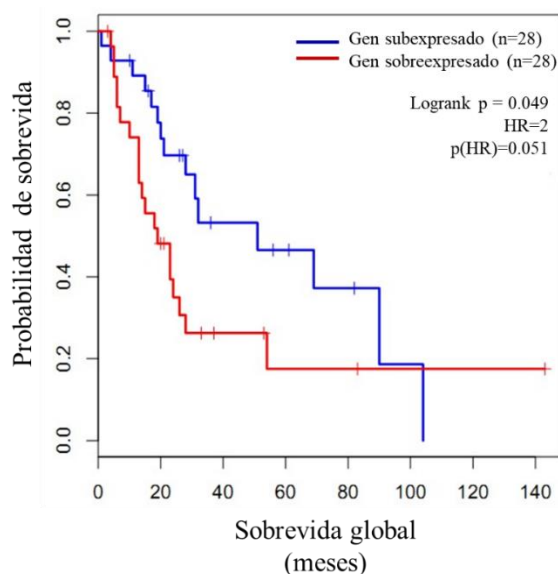


Figura 38. Gráfica de Kaplan Meier representando la OS en función de la expresión génica de CDK5 en pacientes con UCS. El nivel de expresión de CDK5 se relacionó significativamente con la sobrevida global de los pacientes ($p < 0.05$). La prueba de log-rank (Mantel-Cox) indicó diferencias significativas en las curvas de sobrevida global con $p = 0.049$. Cociente de riesgo (HR) se muestra con su respectivo p valor p (HR).

Cabe precisar, que no se encontró asociación entre la expresión génica de los DEGs con la sobrevida global en pacientes con UCEC. Tampoco se observaron asociaciones entre la expresión génica de los DEGs y la DFS en UCEC y UCS.

5.11. Estado de metilación del ADN en promotores de genes de estudio

Empleando MethHC, se investigó la variación de los niveles de metilación en el promotor de 160 genes de interés, en mujeres con cáncer ginecológico, en comparación con muestras no cancerosas. Los genes con metilación diferencial y los valores $\Delta\beta$ para CESC y UCEC, se muestran en la Tabla 16A y 17A, respectivamente (Anexo 5). Los valores de $\Delta\beta$ son considerados para determinar si un promotor está hiper- o hipo-metilado.

Los resultados mostraron que, en mujeres con CESC, 19 de 160 genes de estudio tuvieron genes ya sea hiper- o hipo-metilados en comparación con muestras normales. De ellos, 16 presentaron promotores hipermetilados incluyendo a 2 modificadores del ADN (DNMT3B y TRDMT1) y 14 modificadores de histonas (Figura 39). Al contrario, los genes HLCS, KAT2A y LIMK2 (modificadores de histonas) tuvieron promotores hipometilados (Figura 39). Contrastamos los datos de metilación de promotores con los resultados de expresión génica obtenidos en la sección 5.8 e identificamos que un DEG (EYA1) presentó metilación diferencial en CESC (Tabla 16A en Anexo 5). Así la subexpresión del gen EYA1 y la hipermetilación del promotor, indicaría que la metilación del ADN del promotor controla la expresión de EYA1.

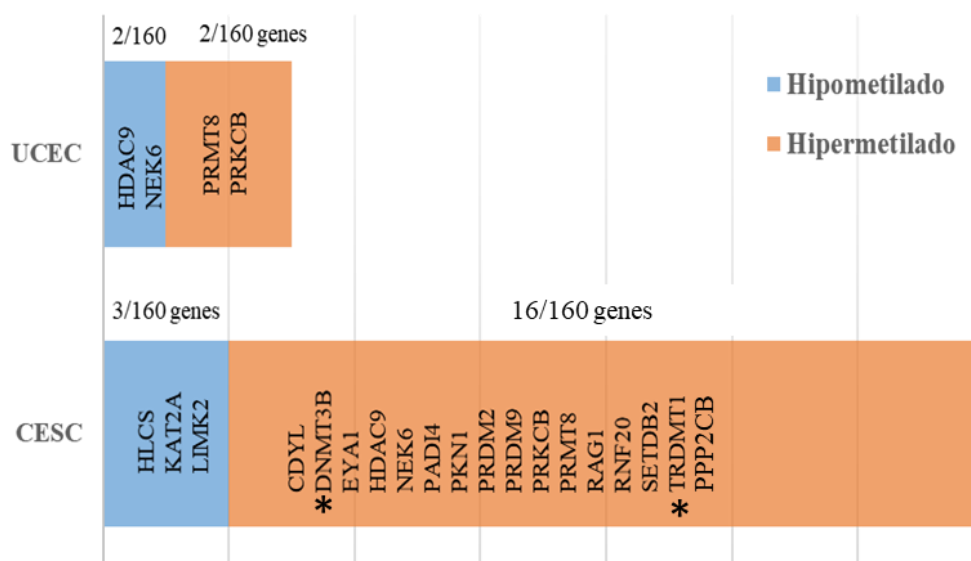
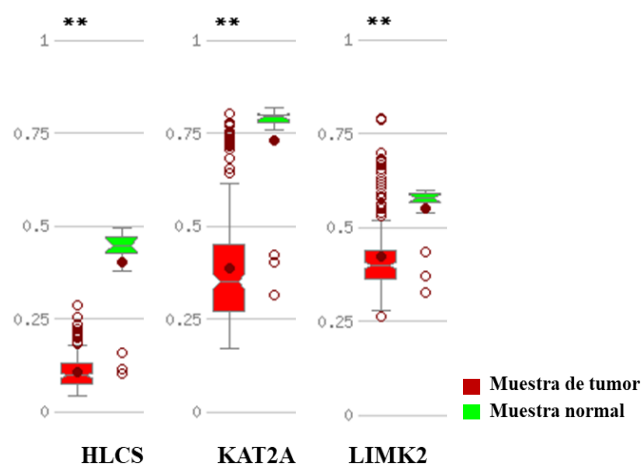


Figura 39. Metilación diferencial a nivel del promotor de los genes de interés. Las barras rojas contienen a los genes que se encuentran hipermetilados, en azul a los genes hipometilados. Tamaño muestral CESC (n = 212) / Tejido normal (n = 19), tamaño muestral UCEC (n = 424) / Tejido normal (n = 67). Significancia estadística p -valor < 0.05, calculado mediante T-Test. CESC: Cáncer cervical de células escamosas y adenocarcinoma endocervical, UCEC: Carcinoma de endometrio del cuerpo uterino. (*): Gen implicado en la metilación del ADN.

Los niveles de metilación del promotor (promedio del valor beta) de los 19 genes en cuestión fueron representados en diagrama de cajas observándose diferencias significativas entre el tejido normal y el tejido con CESC ($p < 0.005$) (Figura 40).

A



B

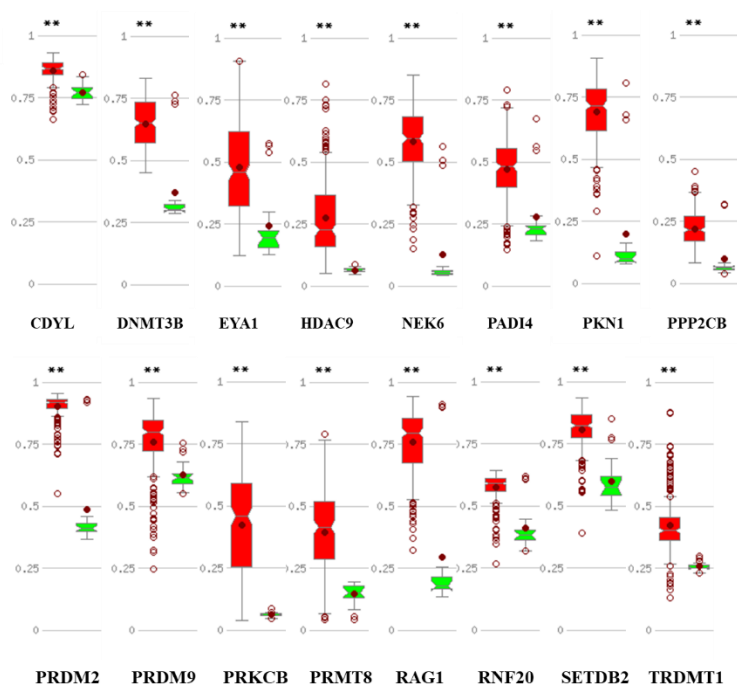


Figura 40. Niveles de metilación a nivel del promotor de 19 genes con metilación diferencial en CESC. (A) Hipometilación de promotor de HLCS, KAT2A y LIMK2. (B) Hipermetilación de promotor en 16 genes. Tamaño muestral CESC (n = 212) / Tejido normal (n = 19).

(**): valor $p < 0.005$

Por otro lado, en UCEC, los genes PRKCB y PRMT8 presentaron hipermetilación de promotor y los genes HDAC9 y NEK6 resultaron tener el promotor hipometilado (Figura 41). Los cuatro genes estuvieron involucrados en modificación de histonas y solo NEK6 fue identificado como DEG. Así la sobreexpresión del gen NEK6 y la hipometilación del promotor, sugiere que la expresión de NEK6 estaría controlada por la metilación del ADN.

Los niveles de metilación del promotor (promedio del valor beta) de los 4 genes mencionados anteriormente fueron representados en diagrama de cajas observándose diferencias significativas entre el tejido normal y el tejido con UCEC ($p < 0.005$) (Figura 41).

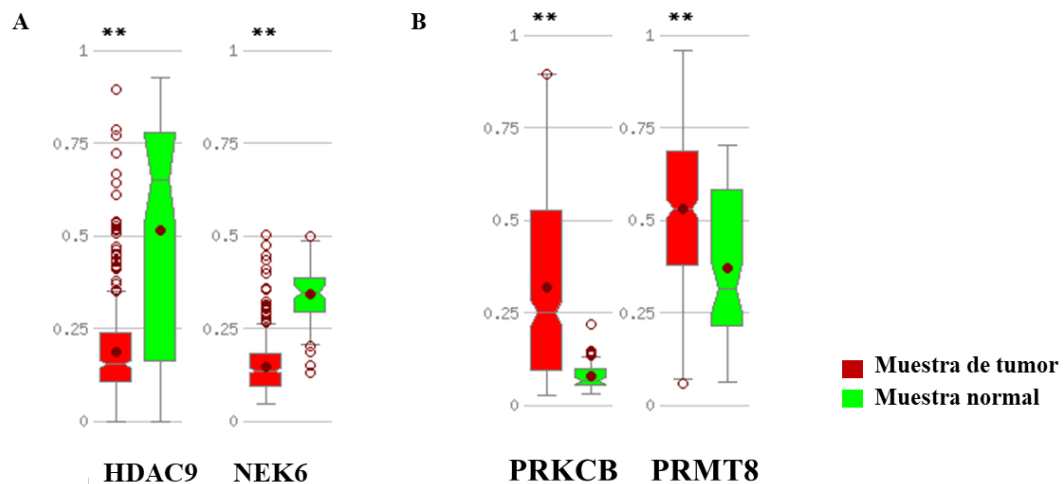


Figura 41. Niveles de metilación a nivel del promotor de 4 genes con metilación diferencial en UCEC. (A) Hipometilación de promotor de HDAC9 y NEK6. (B) Hipermetilación de promotor de PRKCB y PRMT8. Tamaño muestral UCEC (n = 424) / Tejido normal (n = 67). (**): valor $p < 0.005$

Los cambios de metilación de promotores génicos en OV y UCS se investigaron utilizando MEXPRESS ya que MethHC no contenía datos de dichas neoplasias. Se obtuvo la correlación entre la metilación a nivel del promotor y la expresión génica, en los 160 genes de estudio. Para ello se estimaron los valores de correlación de Pearson y significancia estadística. Las correlaciones entre la expresión y metilación de los promotores se muestran según la sonda específica dirigida a la región promotora del gen (5'UTR, TSS $\pm$ 1500 y TSS $\pm$ 200), donde TSS se refiere al sitio de inicio de la transcripción y UTR a la región no traducida (Tabla 18 A y 19 A en Anexo 5).

Para OV, un total de 34 genes presentaron correlación negativa entre las variables estudiadas (Figura 42 y Tabla 18 A). Por lo tanto, la sobreexpresión de algunos de esos genes se correlacionó con la hipometilación de sus promotores mientras que, en otros, la subexpresión génica se correspondió con hipermetilación del promotor. La correlación más fuerte entre la expresión y la metilación fue observada en el gen NCOA3, con una correlación moderada negativa ($-0.5 < r < -0.7$) (Figura 42 y Tabla 18 A). Interesantemente, NCOA3 no se reconoció como DEG en OV.

Por otro lado, se observó que la expresión y la metilación de ATM, KAT2B y TRDMT1 presentaron correlaciones negativas bajas ($-0.3 < r < -0.5$) (Figura 42 y Tabla 18 A). Al ser contrastada con los resultados de expresión génica, los genes ATM y KAT2B fueron identificados como DEGs y estuvieron subexpresados en OV, lo que sugiere que la hipermetilación del promotor estaría suprimiendo la expresión de estos 2 genes. No se reconoció a TRDM1 como DEG en OV.

Al contrario, cuatro genes mostraron una correlación positiva débil ($0 < r < 0.3$) de las variables estudiadas, es decir, que su sobreexpresión se correspondió con hipermetilación de sus promotores. La asociación más fuerte se observó en el gen RAG1($r = 0.294$), el cual no fue identificado como un DEG (Figura 42 y Tabla 18 A).

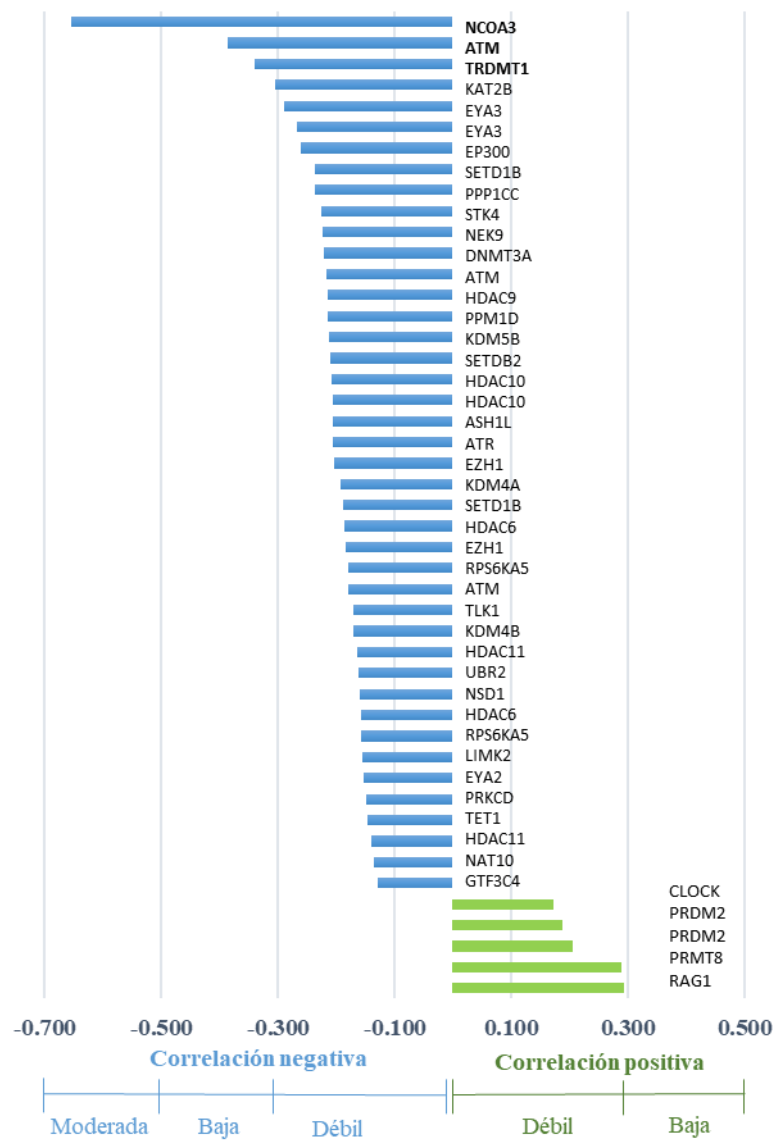


Figura 42. Correlación entre la expresión del ARNm y la metilación del promotor en los tumores de pacientes con OV (n=299) basado en MEXPRESS. Cada barra horizontal representa la correlación de Pearson con valor estadísticamente significativo ($p < 0.05$). Considerando el valor del coeficiente de correlación de Pearson, se ordenaron los genes que van desde -0.653 a 0.294.

Los cambios de metilación de promotores génicos en UCS también fueron investigados con la herramienta MEXPRESS para los 160 genes de interés (Tabla 19 A en Anexo 5).

Un total de 66 genes presentaron correlación negativa entre las variables estudiadas, es decir, que en algunos la sobreexpresión se correspondió con la hipometilación del promotor y en otros la subexpresión del gen se correspondió con la hipermetilación del promotor (Tabla 19 A en Anexo 5). En particular, se observó que la expresión y la metilación de NEK9 y SETDB2 presentaron correlaciones negativas fuertes ($-0.7 < r < -0.9$) y al ser contrastados con los resultados de expresión génica obtenidos en la sección 5.8, se evidenció que NEK9 era un DEG, subexpresado en UCS, lo que sugiere que la hipermetilación del promotor estaría suprimiendo la expresión de NEK9. El gen SETDB2 no se determinó como DEG en UCS. Por otro lado, 20 genes presentaron correlaciones negativas moderadas ($-0.5 < r < -0.7$), de los cuales 6 genes (ATM, CDK17, EZH1, HDAC6, KAT2B y RPS6KA5) fueron identificados como DEGs, todos subexpresados en UCS; lo que sugiere que la hipermetilación del promotor estaría suprimiendo la expresión de dichos genes.

Al contrario, 32 genes mostraron una correlación positiva de las variables estudiadas, es decir, en algunos la sobreexpresión se correspondió con hipermetilación del promotor y en otros la subexpresión se correspondió con hipometilación del promotor (Tabla 19 A en Anexo 5).

VI. DISCUSIÓN

En esta tesis se planteó determinar, mediante análisis de datos ómicos, las alteraciones de reguladores epigenéticos en cáncer ginecológico, esperando reconocer aquellas aberraciones asociadas con el fenotipo incluyendo el pronóstico. La importancia de este trabajo radica en la identificación del daño en los reguladores epigenéticos, únicos o comunes a las neoplasias ginecológicas, que a largo plazo puedan convertirse en marcadores moleculares. En esta investigación se estimaron los tres tipos más importantes de alteraciones genéticas incluyendo mutaciones puntuales, CNAs y cambios de expresión, en muestras de cáncer ginecológico y controles sin cáncer; vinculándose dichas modificaciones a rasgos fenotípicos y clínicos.

La población de estudio fue heterogénea, con diferencias entre los cuatro tipos de cáncer ginecológico (Tablas 2 y Tabla 3). Así, las pacientes con cáncer cervical fueron las más jóvenes en comparación con los otros tipos de neoplasias (OV, UCEC y UCS), con una media de edad de 48 años, lo cual concuerda con hallazgos previos (35,36) y con otros autores del Perú, donde se observa que la edad media de las pacientes con cáncer cervical atendidos en el INEN no difiere mucho (44 años) (Tabla 4).

Además, las pacientes con cáncer cervical tuvieron mejor pronóstico que el grupo de cáncer de ovario, lo cual es compatible con lo informado por IARC (4) (Tabla 3), así como también con estudios observacionales en Perú, donde se observó un mejor pronóstico de pacientes con cáncer cervical en estadios tempranos que en estadios avanzados (Tabla 4). El buen pronóstico de CESC se relaciona con la disponibilidad de exámenes de detección como la prueba de Papanicolaou y

prevención como la vacuna contra el VPH, que aumentan la probabilidad de un diagnóstico temprano y asignación de tratamiento oportuno (124). Por el contrario, el pronóstico de las pacientes con cáncer de ovario es desalentador, en parte debido a que sólo un 20% de los casos se detectan en un estadio temprano y también a que los síntomas y la duración de los mismos son similares entre los pacientes con enfermedad en etapa temprana y los que presentan enfermedad avanzada, lo cual impide que existan métodos eficaces de diagnóstico precoz (37,125). Y cuando se detecta en estadios muy avanzados, una citorreducción óptima (extracción quirúrgica del tumor) puede aumentar las posibilidades de sobrevida global de la paciente con cáncer de ovario (Tabla 4).

En particular, se observó que todas las pacientes con UCS presentaron estado libre de enfermedad al momento del diagnóstico, mientras que la mayoría de participantes en estado recurrente/progresado fueron del grupo de OV (72.6%) (Tabla 3). La alta tasa de pacientes con OV en estado recurrente/progresado es concordante con estudios que muestran un mayor riesgo de desarrollar tumores sólidos en los períodos de seguimiento después del diagnóstico de OV (37,125). Se ha informado que el aumento del índice de masa corporal (IMC) y la obesidad se asocia fuertemente tanto con la incidencia como con la recurrencia y pronóstico del cáncer de endometrio (126,127). Por lo tanto, las pacientes con UCS al presentar un grado de obesidad menor (pre-obesidad), no se encontraron en estado recurrente/progresado (Tabla 2).

Los resultados del análisis de enriquecimiento funcional de los 160 genes de estudio, confirmaron que éstos estaban significativamente enriquecidos en los procesos biológicos de modificación de histonas, desmetilación de la lisina en las

histonas y la regulación de la transcripción (Figura 10). La reducida frecuencia de reguladores epigenéticos que metilan el ADN en nuestro estudio (8 genes) es coherente con su menor abundancia en el genoma humano, en comparación con los genes codantes de modificadores de histonas (Figura 8). La vía más enriquecida en la lista de estudio fue la degradación de lisina vinculada con la metilación y desmetilación específica de dicho aminoácido en histonas (Figura 9). Dichos procesos están regulados principalmente por las familias de enzimas KMTs y KDMs, que metilan y desmetilan, respectivamente, las histonas del octámero nuclear (Figura 11). Al investigar la expresión de dichos genes, resultó que KDM4B y KDM6A están subexpresados en CESC y OV, respectivamente (Figura 21). KDM4B se encarga de eliminar la tri-metilación de la lisina 9 en la histona H3 (H3K9Me3) mientras que KDM6A elimina la di y tri-metilación de la lisina 27 en la histona 3 (H3K27Me2/Me3) (128). Ambos mecanismos regulan la expresión génica de supresores de tumor. Por lo tanto, la inactivación de KDMs mediante el mecanismo de subexpresión génica puede causar represión de genes supresores de tumor debido a la acumulación de marcas de lisina-metiladas en histona 3, principalmente (129). En otros estudios se demostró que la desmetilación de H3K9me3 por KDM4B, activa la transcripción por medio del receptor de estrógeno α (ER α) en el cáncer de mama (128,130,131). Por su parte, KDM6A remueve H3K27me3 para mantener la expresión de proteínas de unión a retinoblastoma (RB), necesarias para mediar la detención del ciclo celular en las células humanas (60,132). La expresión de KDM4B y KDM6A en los cánceres ginecológicos han sido poco explorados, aunque se sabe que KDM4B promueve la progresión del cáncer de endometrio y el cáncer de ovario epitelial (133,134). Además, en el

cáncer cervical, se ha descrito que la sobreexpresión de KDM6A por acción de la oncoproteína E7 del VPH , promueve la activación del ciclo celular y la proliferación celular (135,136). Es posible que ello se deba, en parte, a las técnicas empleadas para determinar expresión génica las cuales pueden variar de un estudio a otro.

Por otro lado, observamos que 66 de los 160 genes consultados, estuvieron involucrados en la regulación de la transcripción a nivel del ADN molde , lo que indica que estos genes estarían participando en la activación y la represión de la expresión génica (137). Dichos elementos correspondieron a coactivadores transcripcionales como CREBBP, EP300 NCOA1, NCOA3, KMT2A, KMT2C, KMT2E, KDM3A, KDM4A y KDM4B (acetiladores de lisina, metiladores y desmetiladores de lisina) o bien enzimas modificadoras de histonas implicadas en la fosforilación y desfosforilación de serina/treonina/tirosina, metilación de arginina, ubiquitinación y desubiquitinación de lisina y metilación del ADN (Figura 10).

La carcinogénesis viral fue una de las rutas más enriquecidas en el grupo de genes de estudio (Figura 11). En particular, se observó una participación mayoritaria de histonas deacetilasas (HDACs) en el proceso de carcinogénesis viral, lo cual demuestra que la regulación epigenética es modificada por la transformación viral (138,139). Al investigar las CNAs, notamos que algunas HDACs como HDAC4 y HDAC10, presentan deleciones en CESC, OV, UCEC y UCS (Tabla 9). Dicho mecanismo suele conllevar a pérdida de función de las proteínas codificadas, lo cual disminuiría la tasa de desacetilación de histonas. Otro mecanismo de inactivación génica es la subexpresión, la cual fue observada en algunas HDACs, en 3 de los 4

tipos de cánceres estudiados (Figura 21). Si bien en este estudio no podemos establecer la influencia de un mecanismo sobre el otro, trabajos previos han sugerido que las CNAs influyen en la variación de la expresión génica en cáncer cervical (140).

Mediante análisis por Mutsig2CV se determinó que los reguladores epigenéticos presentan mutaciones somáticas en las cuatro neoplasias ginecológicas estudiadas (Figura 12). Así se identificaron 26 genes mutados en frecuencias más altas (FMGs) que el resto de genes de estudio en UCEC, 8 FMGs en CESC, 8 FMGs en OV y 2 FMGs en UCS (Figura 12). Este hallazgo demuestra que las mutaciones puntuales en reguladores epigenéticos son más frecuentes en UCEC en comparación con las otras 3 neoplasias. La mayor presencia de FMGs en cáncer de endometrio podría en parte deberse a deficiencias en la reparación del daño en el ADN, lo que conllevaría a la inestabilidad genómica (14,141).

Los reguladores epigenéticos más mutados en UCEC incluyeron a KMT2C, KMT2D y ATM, que participan en los procesos de metilación de lisina y fosforilación de serina / treonina de histonas (Tabla 1A Anexo 1). En CESC, los reguladores epigenéticos más mutados incluyeron a KMT2D, KMT2C y EP300 que controlan los procesos de metilación y acetilación de lisina (Tabla 1A Anexo 1). En contraste, en OV los reguladores epigenéticos más mutados incluyeron a CREBBP, KDM5C y SETD1A, vinculados a los procesos de acetilación, desmetilación y metilación de lisina, respectivamente (Tabla 1A Anexo 1). Finalmente, en UCS solo se identificaron a 2 genes mutados (KDM4B y HAT1), partícipes de la desmetilación y acetilación de lisina, respectivamente (Tabla 1A Anexo 1). Estos resultados muestran que el mecanismo más afectado en los 4 tipos de cáncer

estudiados es la modificación de residuos de lisinas en histonas (Figura 13). Las mutaciones puntuales en estos reguladores epigenéticos podrían afectar la función de escribir, borrar o leer marcas epigenéticas activadoras o represoras de la expresión. Por ejemplo, la acetilación de H3K27, realizada tanto por EP300 como por CREBBP (ambas en complejo), está asociada con la activación transcripcional. Igualmente, la metilación H3K4, ejecutada por SETD1A, KMT2C y KMT2D, activa la expresión génica. Al contrario, la acción de KDM4B y KDM5C de eliminar la di y tri-metilación del H3K9 y H3K4, respectivamente, activa la transcripción y fomenta el crecimiento celular. Virtualmente, las mutaciones puntuales en KDMs y KMTs pueden ser de pérdida o bien de ganancia de la función.

Para determinar los cambios en la estructura, estabilidad y función proteica ocasionado por las variaciones de aminoácidos, investigamos las mutaciones de pérdida de sentido (cambio de aminoácido) en tres genes claves incluyendo KMT2C, KMT2D y CREBBP.

Considerando la predicción funcional de Biomuta, se determinó que 142 sustituciones de aminoácidos presentes en KMT2D, KMT2C y CREBBP serían “dañinas”, de las cuales, 76 afectaron regiones funcionales o dominios y redujeron la estabilidad energética de las 3 proteínas (Figura 14 A y B).

Utilizando I-Mutant, se reconocieron que 6 mutaciones predichas como desestabilizadoras afectaron al dominio Dedo de Zinc tipo PHD de la proteína KMT2D (Tabla 6). Este hallazgo mostraría que el daño en la función de KMT2D en parte se explica por la afectación de su estructura a nivel del dominio Dedo de Zinc. Este último mantiene la coordinación de los iones zinc para estabilizar los

pliegues de la proteína. Su alteración provocaría distorsión de la estructura, la incapacidad del dominio Dedo de Zinc de unirse al ADN y perturbación de su función reguladora (142,143). Además, otras dos mutaciones desestabilizadoras (T4570A y K4810T) en KMT2D estuvieron localizadas en los motivos LXXLL 6, esenciales para la asociación con el receptor nuclear ESR1 (receptor de estrógeno). Por otro lado, 3 mutaciones desestabilizadoras (G5428D, K5493N y R5501Q) estuvieron localizadas en el dominio SET, cerca del sitio de unión responsable de actividad metiltransferasa H3K4 (144) (Tabla 6). En tanto la mutación M5536T en el dominio Post-SET afectaría también la actividad metiltransferasa, importante en el reconocimiento de sustratos y la actividad enzimática (145).

De las 36 mutaciones puntuales y desestabilizadoras en KMT2D, sólo 3 sustituciones (G1387A, L3591F y L4222P) afectaron la exposición de residuos al solvente pudiendo ser una razón de su desestabilización energética (Tabla 6). La mutación KMT2D G1387A aumentó la exposición al solvente del residuo mutado mientras que las mutantes L3591F y L4222R la redujeron. G1387A mostró cambios de polaridad y se ubicó en el dominio dedo de zinc PHD-tipo 4, además el residuo que forma enlaces de hidrógeno, glicina, estuvo mutado por alanina, el cual no forma enlaces de hidrógeno. Por otro lado, L3591F y L4222P implicaron sustituciones por residuos de naturaleza no polares y éstos estuvieron ubicados en la región espiral y en los motivos LXXLL 3 de la proteína, esenciales para la interacción con los receptores nucleares (143).

En relación con las 23 mutaciones dañinas en KMT2C, curiosamente, se observó que tanto las mutaciones “desestabilizadoras” y “neutras”, se localizaron tanto en los dominios dedo de Zinc como en regiones ricas en glicina y prolina, además

ninguna presentó cambios en la accesibilidad al solvente (Tabla 7). Las mutaciones predichas como “desestabilizadoras” (R4584Q y R4584W) estuvieron localizadas en el dominio FYR N-terminal, las cuales afectarían la asociación con módulos que reconocen modificaciones de histonas (143). Las mutaciones R4584Q y R4584W en este dominio tampoco presentaron cambios en la exposición al solvente del residuo mutado a pesar de mostrar cambios de polaridad, por lo que su desestabilización ocurriría por otros mecanismos de disrupción (Tabla 7).

Finalmente, con respecto a las mutaciones en CREBBP, se predijeron 14 mutaciones dañinas y desestabilizadoras (Tabla 8). Un total de 12 sustituciones implicaron al dominio CBP / p300-tipo HAT, tanto la región rica en Cys/His como la región de interacción con TRERF1 (Factor 1 de regulación transcripcional). Alteraciones en la región rica en Cys/His del dominio CBP / p300-tipo HAT afectarían la acetilación de histonas mientras que y modificaciones en la región TRERF1 alterarían la regulación de la expresión del gen CYP11A1 (enzima de desdoblamiento de la cadena lateral del colesterol) (146,147). La pérdida del dominio CBP/P300 de CREBBP ha sido asociada con la tumorigénesis de neoplasias hematopoyéticas lo cual demuestra un rol fundamental en la regulación epigenética de supresores de tumor y oncogenes (148). Se encontró que la mutación dañina R625H implicó el dominio KIX, por lo que también el acoplamiento de activadores transcripcionales podría verse afectado (149). Ninguna mutación predicha como “dañina” presentó cambios en la accesibilidad al solvente por lo que su desestabilización ocurriría por otros mecanismos de disrupción (Tabla 8). Todos estos hallazgos sugieren que, aunque existieron mutaciones dañinas y

desestabilizadoras, éstos afectaron regiones de unión o sitios activos importantes para la proteína de manera desigual.

En conjunto, el estudio estructural demostró que, al menos en 3 reguladores epigenéticos cruciales como KMT2C, KMT2D y CREBBP concurren mutaciones puntuales dañinas, desestabilizadoras y de pérdida de la función proteica. La mutación desestabilizadora en la posición G5428D de KMT2D, ha sido asociada con el síndrome de Kabuki (KS) ,un síndrome caracterizado por una discapacidad intelectual y múltiples anomalías congénitas como apariencia facial dismórfica (150).

Además, se reconocieron un total de 6 genes comúnmente mutados entre los cánceres ginecológicos (Figura 13). Así los genes KMT2D, KMT2C, EP300 y ATM se encontraron comúnmente mutados en CESC y UCEC, mientras que CREBBP y CDK17 estuvieron comúnmente mutados en OV y UCEC (Tabla 5). Estos hallazgos demuestran que existen mecanismos epigenéticos comúnmente afectados en dos o más tipos de cáncer ginecológico, lo que otorga posibilidades de desarrollar terapias que puedan ser aplicadas a varias neoplasias ginecológicas (31).

La determinación de FMGs nos informa sobre genes que se mutaron con altas frecuencias en el cáncer ginecológico. En contraste, la determinación de SMGs nos ayuda a determinar qué genes se encuentran mutados más a menudo de lo esperado por casualidad dados los procesos de mutación de fondo inferidos (abundancia, clustering y conservación). Así, se identificaron en total 44 FMGs en las neoplasias ginecológicas, con mayor número de genes mutados en UCEC (26 genes) en comparación con las otras neoplasias (CESC, OV y UCS) (Figura 12 y 13). Del

total de FMGs en los cuatro cánceres ginecológicos, 30 genes se identificaron como modificadores de histonas implicados principalmente en la acetilación de lisina, fosforilación de serina/treonina/tirosina, metilación y desmetilación de lisina; mientras que sólo dos genes se identificaron como reguladores de la metilación del ADN (DNMT1 y DNMT3B).

Por otro lado, se reconocieron un total de 5 SMGs incluyendo KMT2C, KMT2D, EP300, ATM y CDK17 comunes tanto en CESC como en UCEC (Tabla 5). Previamente, los genes ATM y EP300, vinculados con la regulación del proceso apoptótico y regulación transcripcional también fueron reportados como SMGs en UCEC y CESC (12). Nuestro estudio identificó un mayor número de SMGs que en dicha publicación posiblemente debido a que el punto de corte de nuestro análisis fue menos estricto ($Q < 0.05$) que el del estudio anterior ($Q < 0.01$).

De acuerdo con el análisis Gistic 2.0, las deleciones fueron las CNAs más frecuentes en los tumores ginecológicos estudiados (Figura 19). Las deleciones ocurrieron en genes implicados en la metilación y desacetilación de lisinas incluyendo a PRDM2 y HDAC10, presentes en CESC, UCEC y UCS, además del gen HDAC4, presente en CESC, OV y UCEC (Figura 20 C). La deleción de enzimas metiladoras de lisinas conllevaría a pérdida de su función, lo cual contribuiría a la activación transcripcional de los oncogenes (129).

Por su parte, las regiones genómicas con ganancia de secuencias (amplificaciones) contenían varios genes implicados en la acetilación de lisina, desubiquitinación de lisina y metilación de lisina, tales como KAT6A (en UCEC y OV), BRCC3 (en CESC y OV) y SETDB1 (CESC y UCEC); (Figura 20 B). La identificación de genes comunes en los cánceres ginecológicos, pero con distinta CNA

(amplificación o delección) entre un tipo de cáncer y otro, se debe a que las regiones genómicas amplificadas albergarían genes de importancia para el desarrollo y la progresión de tumores como mecanismo para activar oncogenes y las regiones genómicas delecionadas alberguen genes supresores de tumores (151). La inestabilidad genómica (cromosómica) adquirida por la acumulación de CNAs y mutaciones puntuales, es un *hallmark* del cáncer que puede contribuir al desarrollo y progresión del mismo (152,153). En este sentido UCS resultó ser el cáncer ginecológico con menor presencia de CNAs y con menor número de FMGs en comparación con los otros 3 tipos de neoplasias estudiadas.

La expresión génica de los reguladores epigenéticos fue también investigada revelándose que 33 genes se expresaron diferencialmente (DEGs) en UCS, 29 en UCEC, 26 en OV y 16 genes en CESC (Figura 21). La expresión diferencial es un indicador la sobre- o sub-expresión génica en muestras cancerosas, en comparación con muestras normales. En los 4 tipos de cáncer, la expresión génica afectó a factores modificadores de histonas mientras que sólo en UCS se observó un gen involucrado en metilación del ADN afectado (DNMT3B). Ello sugiere que la expresión génica de reguladores de metilación de promotores no es un mecanismo primariamente alterado en OV, CESC y UCEC, pudiendo estos últimos soportar su desarrollo en otro tipo de alteraciones genéticas de reguladores de la metilación del ADN, como las mutaciones puntuales o CNA. Entre los reguladores de la metilación del ADN que mostraron mutaciones puntuales, se encontraron a DNMT3B y DNMT1, ambos involucrados en la regulación negativa de la expresión génica y presentes en UCEC (Tabla 2A). Por su parte, entre los reguladores de la

metilación del ADN que presentaron CNA, se encontró a DNMT3L, deletado en OV e involucrado en la regulación de la transcripción a nivel del ADN molde (Tabla 2A).

La mayoría de los reguladores epigenéticos subexpresados en las cuatro neoplasias, estuvieron implicados en procesos de fosforilación/desfosforilación de serina / treonina / tirosina, acetilación/desacetilación de lisina, así como también en la metilación de lisina. Estos procesos mencionados anteriormente, además de la ubiquitinación de lisina también estuvieron presentes en los reguladores epigenéticos sobreexpresados en las cuatro neoplasias (Figura 21 y Tabla 1A Anexo 1). Cabe destacar que estas modificaciones en las histonas regulan procesos como transcripción, reparación o replicación del ADN (Tabla 4 -7 A).

Se identificaron 3 DEGs (AURKB, BUB1 y UHRF1) que estuvieron igualmente sobreexpresados en los 4 cánceres ginecológicos de interés (Figura 22A), promoviendo la fosforilación de serina / treonina y la ubiquitinación de lisina (Tabla 1 A en Anexo 1). Apoyando estos resultados, informes anteriores evidencian la sobreexpresión de AURKB y BUB1 en los cánceres ginecológicos. Estas investigaciones han demostrado la existencia de expresión en exceso de la familia de las auroras quinasas como AURKB en el cáncer de endometrio, así como también la expresión conjunta con BUB1 en diversos cánceres ginecológicos (154–156). UHRF1 ,como “lector” reconoce la marca represiva H3K9me3 importante para el mantenimiento de la metilación del ADN , por lo que recientes estudios están validando a UHRF1 como un conductor del silenciamiento de genes supresores de tumores en el cáncer (60,157,158) . Por otro lado, se reconocieron DEGs comúnmente subexpresados en los tumores de estudio incluyendo ATM,

DUSP1, EZH1, KAT2B, MAP3K20, MGEA5 y NEK9 (Figura 22), implicados en la fosforilación y desfosforilación de serina / treonina/tirosina, metilación de lisina y acetilación de lisina (Tabla 1 A en Anexo 1). De estos, EZH1 es parte del complejo Polycomb PRC2 y trimetila la marca H3K27, reprimiendo así la expresión génica de los genes diana y manteniendo la identidad de las células madre (159). Por lo tanto, la subexpresión de EZH1 activaría la transcripción de los genes blanco, muy posiblemente oncogenes.

La subexpresión de DUSP1 en el cáncer ovario ha sido relacionado con el aumento del grado histológico. Esto debido a que la subexpresión del gen estimula la proliferación y el aumento de la masa tumoral en las etapas más avanzadas de la tumorigénesis mediante la vía de señalización de MAPK, clásica vía de crecimiento (160). Por otro lado, se ha informado que la subexpresión de ATM en las células cancerosas ayudaría a evadir el efecto supresor de tumor, ya que los beneficios de la señalización ATM para las células cancerosas son la promoción de la quimiorresistencia, la radiorresistencia, la metástasis y la supervivencia celular (161).

Se construyó una red de PPIs de los DEGs para determinar los contactos y puntos de acción de los genes más prominentes en cada cáncer ginecológico. Con ello, se identificaron los genes centrales de cada red de PPI, los cuales exhiben funciones más importantes que otros genes ya que tienden a interactuar con un mayor número de proteínas (100).

El interactoma de los 16 DEGs presentes en CESC reveló cinco genes centrales, incluidos ATM, EZH2, AURKB, BUB1 y UHRF1, todos ellos sobreexpresados (a excepción de ATM) en la neoplasia cervical e involucrados en la detención del ciclo

celular, fosforilación de proteínas, regulación negativa de la transcripción del promotor de ARN polimerasa II, y desarrollo embrionario (Figura 23, Tabla 6A). Dos de esos genes (ATM y BUB1) participan en las vías de señalización p53, apoptosis, meiosis ovocitaria, maduración de ovocitos mediada por progesterona y ciclo celular (Tabla 11). Es importante destacar la posición central de EZH2, responsable de la metilación de la lisina 27 de la histona 3 (H3K27) que resulta en la compactación de la cromatina y la represión génica. La sobreexpresión de EZH2 junto con la hipermetilación H3K27 se ha asociado con un mal pronóstico de los pacientes en varios tumores sólidos, incluyendo el cáncer cervical, por lo que ahora se está investigando como una diana terapéutica (162).

La red de PPI de los 23 DEGs en OV identificó cinco proteínas centrales, incluidas KAT2A, KAT2B, HDAC6, ATM y KDM6A, todas ellas subexpresadas en el cáncer de ovario e involucrados en la regulación positiva de la transcripción del promotor del ARN polimerasa II, fosforilación de proteínas, respuesta a la radiación y proliferación celular (Figura 25, Tabla 6A). Dos de esos genes (KAT2A y KAT2B) participan en las vías de señalización Notch vinculada con diferenciación y muerte celular (Tabla 12). Estos resultados son consistentes con un rasgo característico de las células cancerosas que es la postergación de la diferenciación y muerte celular, lo cual sucedería, en parte, por la subexpresión de KAT2A y KAT2B. Es importante destacar la posición central de KDM6A, una demetilasa de lisina en histonas (H3K27-3me), en la red PPI, que regula la expresión de genes supresores de tumores (TSGs). La subexpresión de KDM6A conllevaría a silenciar TSGs en el cáncer de ovario, (163). La inhibición de KDM6A ha sido sugerida

para suprimir la metástasis en el cáncer de mama, como una medida para contrarrestar la hiperactivación de la enzima (164).

La red de PPI de los DEGs en UCEC identificó 5 proteínas centrales, incluidas EZH2, KAT2B, CREBBP, ATM y KAT2A, todas ellas subexpresadas (a excepción de EZH2) en UCEC y vinculadas con modificación y remodelación de la cromatina, acetilación de peptidil-lisina N-terminal y regulación positiva del proceso apoptótico (Figura 27, Tabla 7 A). Cuatro de esos genes (KAT2B, CREBBP, ATM y KAT2A) participan en las vías de señalización Notch, ciclo celular, Jak-STAT y tuberculosis (Tabla 13). Es importante destacar que ATM ha sido relacionada con el desarrollo y la carcinogénesis del cáncer de endometrio y se presume su rol protector o supresor de tumores en el desarrollo del endometrio, por lo que se ha sugerido como un factor predictivo en los estudios básicos y clínicos del cáncer de endometrio (165). Además, la sobreexpresión de EZH2 en cáncer de endometrio ha sido informado recientemente y se ha propuesto como un candidato prometedor para tratar el cáncer de endometrio de alto riesgo (162)

La red de PPI de los DEGs en UCS, identificó cinco proteínas centrales, incluidas ATM, CREBBP, EZH2, KAT2B y SETD7, todas ellas subexpresados (a excepción de EZH2) en UCS y vinculadas con la fosforilación de proteínas, proliferación celular, modificación de la cromatina, respuesta a la radiación y la regulación de la transcripción a nivel del ADN molde (Figura 29, Tabla 4A). Cuatro de esos genes (ATM, CREBBP, KAT2B y SETD7) participan en las vías del ciclo celular, Jak-STAT, tuberculosis, degradación de lisina y Notch (Tabla 14). Es importante destacar la posición central de CREBBP, una histona acetiltransferasa, que está involucrada en la coactivación transcripcional de diversos factores de transcripción

y que debido a su papel dinámico como gen supresor de tumor y como oncogén, en diversos cánceres, ha promovido el diseño de inhibidores contra CREBBP. Sin embargo, aún no se ha examinado cuidadosamente su participación en el desarrollo del cáncer de endometrio (116).

Una de las principales motivaciones de esta tesis es conocer los factores genéticos que influyen en el fenotipo clínico del cáncer ginecológico, principal sobre el pronóstico. Para ellos nos valimos de los plots de Kaplan-Meier, que permitieron evaluar el efecto de la expresión génica de los DEGs en la sobrevida de las pacientes. De un total de 101 DEGs identificados en CESC, OV, UCS y UCEC, 5 de ellos se asociaron con la sobrevida global (p -valor < 0.05). Así, en pacientes con CESC, la sobreexpresión del gen SMYD2 (Figura 32) se correlacionó con un pronóstico desfavorable (baja sobrevida global). SMYD2 metila proteínas histonas en las marcas activadoras H3K4 y H3K36, importantes en el ciclo celular mitótico y fosforilación de proteínas (Tabla 8 A). Así como también proteínas no histonas, incluidas p53 / TP53 y RB1 en Lys-370 y Lys-860, respectivamente, importante para la regulación del ciclo celular y activación las enzimas de reparación del ADN (166,167). Estudios previos demuestran la sobreexpresión de SMYD2 en varios tipos de cáncer incluidos en el cáncer cervical, sugiriéndose su rol como una proteína oncogénica. La reducción de su expresión induce la supresión del crecimiento de las células cancerosas por lo que se han desarrollado inhibidores de SMYD2 incluyendo a AZ505, LLY - 507, A - 893 (168,169). Por otro lado, la subexpresión del gen KAT2B se correlacionó con un pronóstico desfavorable (baja DFS) de las pacientes con CESC (Figura 33). La subexpresión de KAT2B ha sido

previamente documentada y la misma disminuiría los niveles de p53 acetilado favoreciendo la progresión del cáncer cervical (70). Además se ha sugerido que KAT2B tiene un rol como gen supresor de tumor en el cáncer de mama, cáncer de células escamosas esofágico y carcinoma gástrico (170).

En pacientes con cáncer de ovario, la subexpresión de los genes SETDB2, MGEA5 y HDAC10 se correlacionó con un pronóstico favorable (mayor OS) (Figura 34 - 36). Se ha observado que SETDB2, HDAC10 y MGEA5 están involucradas en importantes procesos como la fosforilación de proteínas, proliferación celular y regulación positiva del proceso apoptótico. Por ello se sugiere que la subexpresión de estos genes afectarían la regulación de esos procesos reduciendo la progresión del cáncer de ovario y prolongando la OS.

Además, también en OV, la sobreexpresión del gen NEK6 se correlacionó con un pronóstico favorable (mayor DFS) (Figura 37). NEK6 está implicada en la senescencia celular regulada por p53, por lo que su sobreexpresión suprimiría dicho proceso conllevando a la inmortalidad celular. Estudios *in vitro* demostraron la reducción del crecimiento de células del cáncer de mama debido a la sobreexpresión de NEK6 (171). Sin embargo, la función de NEK6 en el cáncer ginecológico es en gran parte desconocida.

Así mismo, las pacientes con UCS que presentaron sobreexpresión del gen CDK5 tuvieron pobre pronóstico (menor OS) en comparación con los pacientes que presentaron subexpresión del gen CDK5. Cambios en la expresión de CDK5 impactan al ciclo celular, división celular, la detención de la proliferación celular y formación de la corteza cerebelosa. Además, CDK5 es un modulador de la función

de los supresores de tumores y los factores de transcripción por lo que su sobreexpresión influye en la oncogénesis (172). También se ha informado la participación de CDK5 en el desarrollo y la progresión de una variedad de cánceres incluyendo neoplasias de mama, pulmón, colon, páncreas, melanoma, tiroides y cerebro (172). Alteraciones en CDK5 conllevan a la alteración en la proliferación celular, migración y angiogénesis además de asociarse con la resistencia a la quimioterapia y la inmunidad antitumoral (172). Por lo que CDK5 ha sido objetivo de interés para combatir el cáncer, sin embargo, en comparación con otras quinasas, la función de CDK5 en el cáncer uterino es desconocida (173).

Nuestros hallazgos sugieren que la expresión diferencial de los genes mencionados podría servir como posibles biomarcadores epigenéticos de pronóstico en pacientes con CESC, OV y UCEC.

La expresión de los DEGs también se investigó en función de otros fenotipos como la edad y el peso, para las cuatro neoplasias (CESC, OV, UCEC y UCS) (Figura 31). Se destacó que las pacientes con OV que tuvieron menores niveles de expresión de HDAC10, presentaron mayor edad, mientras que las participantes más jóvenes tuvieron menor nivel de expresión de SETDB2 (Tabla 15). Al ser la edad uno de los factores que pueden aumentar el riesgo de desarrollar cáncer de ovario, es importante establecer un modelo de predicción para diagnosticar esta neoplasia en mujeres jóvenes donde el cáncer es menos invasivo. Para complementar estos hallazgos, se observó que la subexpresión de SETB2 (Histona-lisina N-metiltransferasa SETDB2) y HDAC10 (Histona desacetilasa 10), está asociada a un mejor pronóstico de OS en pacientes con OV (Figura 34 y 36). De esa forma, las mujeres más jóvenes afectadas por OV con subexpresión de SETDB2 tendrían un

mejor pronóstico de vida en comparación con mujeres que presentan mayor edad. Las pacientes que presentaron mayor edad afectadas por OV con subexpresión de HDAC10 tendrían un mejor pronóstico de vida en comparación con mujeres más jóvenes. Estos hallazgos sugieren que en OV edad podría ser un factor que influye en el pronóstico de la enfermedad, a través al menos de la expresión de dos genes (SETDB2 y HDAC10).

Por otro lado, los resultados de las correlaciones entre la expresión de los DEGs con el peso de las pacientes mostraron que, en CESC, el gen EZH1(Histona-lisina N-metiltransferasa EZH1) presentó una tendencia de regulación de expresión descendente en participantes con menor peso, en comparación con las participantes de mayor peso. Por otro lado, las participantes con mayor peso presentaron una tendencia de regulación de expresión descendente de MAP3K20 (Proteína quinasa quinasa 20 activada por mitógeno) en comparación con las de menor peso (Tabla 16).

Para UCEC, se determinó que las participantes con mayor peso tuvieron una tendencia de regulación de expresión ascendente de EYA2 (Ojos ausentes homólogo 2) en comparación con las de menor peso. Por otro lado, las participantes con menor peso tenían una tendencia de regulación de expresión descendente en ATM (Serina-proteína quinasa ATM) y MGEA5 (Proteína bifuncional NCOAT) (Tabla 16) en comparación con las de mayor peso. La obesidad es un factor de riesgo de enfermedades como cardiovasculares y diabetes tipo 2, así como también para una variedad de tipos de cáncer, entre ellos ginecológico, su importancia radica en que las células de grasa producen estrógeno y las concentraciones de estrógenos en el tejido adiposo promueven el crecimiento de las células del cáncer de

endometrio mediante la regulación directa e indirecta de la transcripción de genes (174) .

La expresión génica de algunos genes humanos es controlada por cambios en la metilación del ADN, principalmente a nivel de la región promotora. En este estudio, luego de determinar los genes cuya expresión es aberrante en cáncer (DEGs), se investigó si la metilación a nivel del promotor era igual o diferente que en células normales. Los resultados mostraron que, en pacientes con CESC, los promotores de 19 DEGs presentaron cambios de metilación (metilación diferencial) comparados con sujetos sin cáncer (Tabla 16A). De ellos, 16 DEGs tuvieron promotores hipermetilados incluyendo a CDYL, EYA1, HDAC9, NEK6, PADI4, PKN1, PPP2CB, PRDM2, PRDM9, PRKCB, PRMT8, RAG1, RNF20, SETDB2, TRDMT1 y DNMT3B. Si la metilación estuviese controlando la expresión de dichos genes esperaríamos observar que su transcripción estuviese silenciada.

Sin embargo, ello sólo se reconoció en el gen EYA1 sugiriendo que su expresión estaría regulada por la metilación a nivel del promotor. La hipermetilación del ADN en la región promotora de EYA1 es determinada por el aumento en la cantidad de grupo metilo hecho que impide la unión de los factores de transcripción, desencadenando en el silenciamiento del gen (55) (Figura 39 y Tabla16 A). Al contrario, en los 18 DEGs restantes de CESC, la expresión no mostró ser controlada por la metilación del promotor ya que dicha variable fue directamente proporcional a la expresión génica (Tabla 16 A).

En CESC también se halló que los genes HLCS (Biotina - proteína ligasa), KAT2A (Histona acetiltransferasa KAT2A) y LIMK2 (Dominio LIM quinasa 2) tuvieron promotores hipometilados (Figura 40). Sin embargo, no se observó sobreexpresión en estos genes mencionados.

Por otro lado, los resultados mostraron que en pacientes con UCEC, los promotores de 4 DEGs mostraron metilación diferencial en comparación con sujetos sin cáncer (Tabla 17A). Dos de esos genes (PRKCB y PRMT8) tuvieron promotores hipermetilados mientras que los otros dos (HDAC9 y NEK6) tuvieron promotores hipometilados (Figura 40). Solo en el gen NEK6, la relación inversamente proporcional entre la metilación del promotor y la expresión sugirió que la expresión del gen sí es regulada por metilación a nivel de su promotor. Ya que el promotor está hipometilado, el gen está sobreexpresado en pacientes con UCEC (Figura 39 y Tabla 17 A).

Los cambios de metilación de promotores génicos en OV y UCS se investigaron utilizando la herramienta MEXPRESS, los datos obtenidos mediante este recurso permitieron evaluar la correlación entre la metilación a nivel del promotor y la expresión génica, en los 160 genes de estudio (Tabla 18 A y 19 A).

Se observó que en pacientes con OV la expresión de 34 DEGs está mediada por la metilación del promotor ya que la correlación de ambas variables fue inversa (Figura 42 y Tabla 18 A). La correlación más fuerte fue del gen NCOA3, coactivador del receptor nuclear 3 (Figura 42 y Tabla 18 A). Sin embargo, no se reconoció como DEG en OV. Aunque estudios *in vitro* han informado que NCOA3 se encuentra sobreexpresado en células cancerosas de ovario y que actúa como coactivador de receptores nucleares como el receptor de estrógeno (175). Al

contrario, también en OV, la expresión de los genes CLOCK, PRDM2, PRMT8 y RAG1 no parece estar mediada por la metilación del promotor ya que la correlación de ambas variables fue directamente proporcional (Figura 42 y Tabla 18 A).

Finalmente, en pacientes con UCS, se observó que la expresión de 66 DEGs se correlacionó inversamente con la metilación del promotor, demostrando que la transcripción está influida por la metilación del ADN (Tabla 19 A). La correlación más fuerte se observó en NEK9, un gen subexpresado en UCS, con hipermetilación del promotor, el cual sería la razón de su silenciamiento génico (Figura 21 y Tabla 19 A). Al contrario, también en UCS, la expresión de 32 genes no parece estar mediada por la metilación del promotor, ya que la correlación de ambas variables fue directamente proporcional (Tabla 19 A).

La evidencia reunida en la presente tesis ha sido obtenida ejecutando diversos protocolos de análisis bioinformático integrado, que consideraron genomas, transcriptomas y metilomas, previamente obtenidos de estudios internacionales. Las herramientas de análisis ómico permitieron describir el perfil molecular de 4 tipos de cáncer ginecológico, accediendo a datos abiertos al público, los cuales fueron procesados con aplicaciones gratuitas, afines a cada tipo de dato ómico investigado. Ello demuestra que el análisis secundario de datos ómicos es potencialmente una mina de conocimientos albergados por tipo de neoplasia. El aprovechamiento de los datos ómicos en cáncer y otras patologías requiere de aplicaciones de procesamiento e integración y de usuarios familiarizados con programación y análisis estadístico. Así, el desarrollo de estudios exploratorios como el nuestro permite realizar una descripción global del problema, identificando

similitudes y particularidades de ciertos tipos de cáncer, que en un futuro ameritan una investigación dirigida y basada en una hipótesis. Sobre todo, estudios piloto como el nuestro permiten reconocer marcadores candidatos que en subsecuentes investigaciones deberán ser validados mediante un diseño experimental.

VII. CONCLUSIONES

De manera global, el análisis integrador del perfil molecular de cáncer ginecológico logró identificar alteraciones en los reguladores epigenéticos comunes a las cuatro neoplasias estudiadas ,así como específicas por tipo de tumor. Además, las principales alteraciones genéticas encontradas en los reguladores epigenéticos incluyeron mutaciones puntuales, alteraciones en el número de copias (CNAs) y cambios en su expresión.

Objetivo 1: Identificar alteraciones en la expresión génica y variaciones de secuencia presentes en las enzimas de estudio, así como conocer el estado de metilación del ADN en promotores de genes de estudio. El análisis estadístico aplicado en esta sección compensa parcialmente el problema de las comparaciones múltiples. Aunque estadísticamente se puedan encontrar asociaciones significativas, el efecto podría ser sobreestimado para explicar las variaciones de secuencias presentes en los genes de estudio.

De manera puntual, se pueden mencionar las siguientes conclusiones:

- Se identificaron 5 reguladores epigenéticos mutados significativamente (SMGs) en las neoplasias ginecológicas CESC y UCEC, involucrados en la modificación de histonas. No se identificaron SMGs en OV y UCS; sin embargo, se encontraron genes con mutaciones que ocurren en alta frecuencia (FMGs).
- El análisis de alteraciones en el número de copias (CNAs) identificó 86 genes afectados principalmente por deleciones en los cuatro tipos de cáncer ginecológico, siendo la mayoría de los genes modificadores de histonas.

- Se identificaron genes con expresión diferencial (DEGs) únicos y comunes en cada neoplasia ginecológica, del mismo podemos concluir que solo EYA1 y NEK6 presentaron metilación diferencial en el promotor en las neoplasias CESC y UCEC, respectivamente.

Objetivo 2: Predecir los cambios en la estructura y función de las proteínas de potencial interés. Las predicciones *in silico* basadas en las secuencias de las proteínas fueron ejecutados siguiendo los parámetros predeterminados de cada programa. Aunque la precisión predictiva de cada programa sea alta, los resultados podrían ser insuficientes para explicar completamente los cambios en la estructura y función de las proteínas. De manera puntual, se pueden mencionar la siguiente conclusión:

- El estudio *in silico* de estructura-función y estabilidad energética de las proteínas CREBBP, KMT2C y KMT2D, mostró que sólo algunas mutaciones dañinas afectaron los dominios proteicos relevantes para la función. También se observó que todas las mutaciones dañinas desestabilizaron energéticamente a las proteínas, aunque no necesariamente debido al cambio de exposición al solvente del residuo alterado.

Objetivo 3: Asociar las variables genéticas con variables fenotípicas. El análisis estadístico aplicado en esta sección también compensa solo parcialmente el problema de las comparaciones múltiples, por lo que el efecto obtenido podría ser insuficiente para explicar las asociaciones entre la expresión génica de los genes con expresión diferencial (DEGs) con los caracteres fenotípicos. De manera puntual, se pueden mencionar las siguientes conclusiones:

- Se identificaron en total 5 DEGs (SMYD2, SETDB2, MGEA5, HDAC10 y CDK5) asociados significativamente con la sobrevida global (OS) del paciente con cáncer ginecológico. De estos, el gen SMYD2 estuvo sobreexpresado en pacientes con CESC, mientras que los genes SETDB2, MGEA5 y HDAC10 estuvieron subexpresados en pacientes con OV y finalmente el gen CDK5 estuvo sobreexpresado en pacientes con UCS.
- Se reconocieron 2 DEGs (KAT2B y NEK6) asociados significativamente con la sobrevida libre de enfermedad (DFS) del paciente con cáncer ginecológico, encontrándose el gen KAT2B subexpresado en pacientes con CESC, mientras que el gen NEK6 estuvo sobreexpresado en pacientes con OV.
- No se encontró una correlación fuerte entre la expresión de los DEGs con las variables de la edad y el peso en las cuatro neoplasias ginecológicas (CESC, OV, UCEC y UCS), debido a las limitaciones del recurso en línea UCSC Xena.

VIII. RECOMENDACIONES

- Los resultados de este estudio indican la necesidad de analizar otras cohortes ginecológicas distintas de TCGA como por ejemplo el conjunto de datos clínicos sobre el cáncer de adultos de la Fundación de Medicina (FM-AD), así como también complementar con el estudio del cáncer de vulva, vagina y trompas de Falopio, a fin de obtener una mejor comprensión sobre la biología del tumor ginecológico.
- Para un panorama más completo, se debe considerar el estudio de otros componentes importantes del epigenoma, como los miRNAs ya que ellos también participan en la regulación de numerosas vías metabólicas y procesos celulares asociados al cáncer.
- Se recomienda profundizar en la interpretación biológica y mecanística de los genes identificados como posibles biomarcadores de diagnóstico y pronóstico, debido a su gran aporte en la práctica clínica.
- Se recomienda seleccionar adecuadamente las herramientas bioinformáticas tomando en cuenta el método de desarrollo, el rendimiento, la usabilidad, así como su disponibilidad a las necesidades particulares del investigador, con el fin de conocer sus alcances y limitaciones.
- Para controlar la probabilidad de cometer un error tipo I en pruebas de hipótesis múltiples, se recomienda aplicar la corrección de Bonferroni o ajustar la tasa de descubrimientos falsos utilizando el procedimiento Benjamini-Hochberg en las futuras investigaciones.
- Adicionalmente, se recomienda una futura corroboración experimental de los hallazgos obtenidos, debido a la implicancia diagnóstica y pronóstica

que presentan los reguladores epigenéticos en la sobrevida de las pacientes con cáncer ginecológico.

IX. LIMITACIONES

- Esta tesis no consideró el cáncer de vulva ni de vagina, debido al número limitado de muestras secuenciadas e información clínica, así como al hecho de no pertenecer a los estudios TCGA. Si se hubiera considerado las cohortes mencionadas anteriormente, hubiéramos tenido poco o ningún poder estadístico.
- Evidenciamos que no se cuenta con parte de la información clínica y/o antecedentes de las pacientes con cáncer ginecológicos, disponible desde recurso Cbioportal. Así, el hábito de consumo de tabaco o estatus de infección por el Virus del Papiloma Humano (VPH), fueron registrados sólo en una parte de los individuos, a pesar de haber sido reportado en la publicación de TCGA
- Debido a que el presente estudio estuvo dirigido a una población específica (TCGA), el alcance de nuestros hallazgos podría aplicarse solo a poblaciones que cumplan con los criterios que siguió el Atlas del Genoma del Cáncer (TCGA).
- Las herramientas bioinformáticas de predicción utilizadas en nuestra investigación permitieron simular una investigación experimental, ahorrando tiempo, recursos financieros y materiales. Sin embargo, para testificar las predicciones hechas se tendrá que realizar estudios experimentales que corroboren nuestros hallazgos.

X. REFERENCIAS BIBLIOGRÁFICAS

1. Avril N, Gourtsoyianni S, Reznick R. Gynecological cancers. Vol. 727, Methods in molecular biology (Clifton, N.J.). 2011. 171–189 p.
2. CDC - What Is Gynecologic Cancer? [Internet]. [cited 2019 Jan 30]. Available from: https://www.cdc.gov/cancer/gynecologic/basic_info/what-is-gynecologic-cancer.htm
3. Funston G, O’Flynn H, Ryan NAJ, Hamilton W, Crosbie EJ. Recognizing Gynecological Cancer in Primary Care: Risk Factors, Red Flags, and Referrals. Adv Ther [Internet]. 2018 Apr [cited 2019 Feb 2];35(4):577–89. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29516408>
4. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin [Internet]. 2018 Nov 1 [cited 2019 Feb 2];68(6):394–424. Available from: <http://doi.wiley.com/10.3322/caac.21492>
5. OMS | Datos y cifras sobre el cáncer. WHO [Internet]. 2014 [cited 2019 Jan 31]; Available from: <https://www.who.int/cancer/about/facts/es/>
6. Drake TM, Knight SR, Harrison EM, Søreide K. Global Inequities in Precision Medicine and Molecular Cancer Research. Front Oncol [Internet]. 2018;8(September). Available from: <https://www.frontiersin.org/article/10.3389/fonc.2018.00346/full>
7. Randall TC, Goodman A, Schmeler K, Durfee J, Pareja R, Munkarah A, et al. Cancer and the world’s poor: What’s a gynecologic cancer specialist to do? Gynecol Oncol [Internet]. 2016 Jul 1 [cited 2019 Feb 2];142(1):6–8.

Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27210817>

8. Kamps R, Brandão RD, van den Bosch BJ, Paulussen ADC, Xanthoulea S, Blok MJ, et al. Next-generation sequencing in oncology: Genetic diagnosis, risk prediction and cancer classification. *Int J Mol Sci*. 2017;18(2).
9. Fink SP, Guda K. Next Generation Sequencing in Cancer Research. *Next Gener Seq Cancer Res* [Internet]. 2013;1:75–90. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84926611006&partnerID=tZOtx3y1>
10. Hrdlickova R, Toloue M, Tian B, Genetics M, New R. RNA-Seq methods for transcriptome analysis *Radmila*. 2017;8(1):1–24. Available from: <http://webbook.nist.gov/cgi/cbook.cgi?ID=C106423&Type=IR-SPEC&Index=4#IR-SPEC>
11. Zhang Z, Li H, Jiang S, Li R, Li W, Chen H, et al. A survey and evaluation of Web-based tools/databases for variant analysis of TCGA data. *Brief Bioinform* [Internet]. 2018;(June):1–18. Available from: <https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bby023/4956394>
12. Berger AC, Korkut A, Kanchi RS, Hegde AM, Lenoir W, Liu W, et al. A Comprehensive Pan-Cancer Molecular Study of Gynecologic and Breast Cancers. *Cancer Cell*. 2018;33(4):690-705.e9.
13. Burk RD, Chen Z, Saller C, Tarvin K, Carvalho AL, Scapulatempo-Neto C, et al. Integrated genomic and molecular characterization of cervical cancer. *Nature*. 2017;543(7645):378–84.
14. Getz G, Gabriel SB, Cibulskis K, Lander E, Sivachenko A, Sougnez C, et al.

- Integrated genomic characterization of endometrial carcinoma. *Nature*. 2013;497(7447):67–73.
15. Cherniack AD, Shen H, Walter V, Stewart C, Murray BA, Bowlby R, et al. Integrated Molecular Characterization of Uterine Carcinosarcoma. *Cancer Cell*. 2017;31(3):411–23.
 16. Bell D, Berchuck A, Birrer M, Chien J, Cramer DW, Dao F, et al. Integrated genomic analyses of ovarian carcinoma. *Nature* [Internet]. 2011;474(7353):609–15. Available from: [papers://8c23484c-02d8-46c3-9325-c916c1add806/Paper/p8203](https://pubmed.ncbi.nlm.nih.gov/21538460/)
 17. Verma M, Kumar V. Epigenetic Drugs for Cancer and Precision Medicine. *Epigenetics Aging Longev* [Internet]. 2018 Jan 1 [cited 2019 Feb 3];439–51. Available from: <https://www.sciencedirect.com/science/article/pii/B9780128110607000218>
 18. Clark RSB, Bayir H, Jenkins LW. Posttranslational protein modifications. *Crit Care Med* [Internet]. 2005;33(Suppl):S407–9. Available from: <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00003246-200512001-00004>
 19. Biswas S, Rao CM. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur J Pharmacol* [Internet]. 2018 Oct 15 [cited 2019 Feb 3];837:8–24. Available from: <https://www.sciencedirect.com/science/article/pii/S0014299918304667?via%3Dihub>
 20. Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis*. 2009;31(1):27–36.

21. Kanwal R, Gupta S. Epigenetic modifications in cancer. *Clin Genet* [Internet]. 2012 Apr [cited 2019 Feb 3];81(4):303–11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22082348>
22. Banerjee HN, Verma M. Epigenetic mechanisms in cancer. *Biomark Med* [Internet]. 2009 Aug 6 [cited 2019 Feb 3];3(4):397–410. Available from: <https://www.futuremedicine.com/doi/10.2217/bmm.09.26>
23. Ng VY, Scharschmidt TJ, Mayerson JL, Fisher JL. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *Anticancer Res*. 2013;33(6):2597–604.
24. Cancer Tomorrow [Internet]. [cited 2019 Feb 3]. Available from: https://gco.iarc.fr/tomorrow/graphic-isotype?type=1&population=900&mode=population&sex=0&cancer=39&age_group=value&apc_male=0&apc_female=0
25. Varughese J, Richman S. Cancer care inequity for women in resource-poor countries. *Rev Obstet Gynecol* [Internet]. 2010;3(3):122–32. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3046761&tool=pmcentrez&rendertype=abstract>
26. Prager GW, Braga S, Bystricky B, Qvortrup C, Criscitiello C, Esin E, et al. Global cancer control: Responding to the growing burden, rising costs and inequalities in access. *ESMO Open*. 2018;3(2):1–10.
27. Frech S, Muha CA, Stevens LM, Trimble EL, Brew R, Perin DP, et al. Perspectives on Strengthening Cancer Research and Control in Latin America Through Partnerships and Diplomacy: Experience of the National

- Cancer Institute's Center for Global Health. *J Glob Oncol* [Internet]. 2018;(4):1–11. Available from: <http://ascopubs.org/doi/10.1200/JGO.17.00149>
28. MINSA. Memoria Plan Esperanza. 2015;27. Available from: <ftp://ftp2.minsa.gob.pe/descargas/ogc/boletines/Memoria-Plan-EsperanzaPERU.pdf>
 29. Sonoda K. Molecular biology of gynecological cancer. *Oncol Lett* [Internet]. 2016;11(1):16–22. Available from: http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=5&SID=2AIFXs9vcCXJfRcp4&page=1&doc=5
 30. Sociedad Mexicana de Oncología. LF, Thomson Gale (Firm). *Gaceta Mexicana de oncología*. [Internet]. Vol. 13, *Gaceta Mexicana de Oncología*. Sociedad Mexicana de Oncología; 2014 [cited 2019 Feb 4]. 236–243 p. Available from: <http://www.elsevier.es/es-revista-gaceta-mexicana-oncologia-305-articulo-las-alteraciones-epigeneticas-progresion-del-X1665920114579068>
 31. Bennett RL, Licht JD. Targeting Epigenetics in Cancer. *Annu Rev Pharmacol Toxicol* [Internet]. 2018 Jan 6 [cited 2019 Feb 4];58:187–207. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28992434>
 32. What Is Cancer? - National Cancer Institute [Internet]. [cited 2019 Feb 4]. Available from: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>
 33. Cooper GM. *The Development and Causes of Cancer*. 2000 [cited 2019 Feb 4]; Available from: <https://www.ncbi.nlm.nih.gov/books/NBK9963/>

34. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. Cell [Internet]. 2011 Mar 4 [cited 2019 Feb 4];144(5):646–74. Available from: <https://www.sciencedirect.com/science/article/pii/S0092867411001279>
35. Venegas G. Guía Técnica Cáncer de cuello uterino. Minist Salud Publica Peru [Internet]. 2017;1–36. Available from: <http://bvs.minsa.gob.pe/local/MINSA/4146.pdf>
36. Bhatla N, Aoki D, Sharma DN, Sankaranarayanan R. Cancer of the cervix uteri. Int J Gynecol Obstet. 2018;143:22–36.
37. Berek JS, Kehoe ST, Kumar L, Friedlander M. Cancer of the ovary, fallopian tube, and peritoneum. Int J Gynecol Obstet. 2018;143:59–78.
38. Cáncer ginecológico - Clínica Ginecológica [Internet]. [cited 2019 Feb 4]. Available from: <https://www.clinicaginecologica.org/patologias/cancer-ginecologico/>
39. Su Cuerpo C. INFÓRMESE SOBRE LOS CÁNCERES GINECOLÓGICOS Infórmese sobre el cáncer ginecológico [Internet]. [cited 2019 Feb 3]. Available from: www.cdc.gov/spanish/cancer/knowledge
40. Amant F, Mirza MR, Koskas M, Creutzberg CL. Cancer of the corpus uteri. Int J Gynecol Obstet. 2018;143:37–50.
41. Ellis H. Anatomy of the uterus. Anaesth Intensive Care Med [Internet]. 2011 Mar 1 [cited 2019 Feb 5];12(3):99–101. Available from: <https://www.sciencedirect.com/science/article/pii/S1472029910002973>
42. ¿Qué es el sarcoma uterino? [Internet]. [cited 2019 Feb 5]. Available from: <https://www.cancer.org/es/cancer/sarcoma-uterino/acerca/que-es-sarcoma-uterino.html>

43. Endometrial Cancer Early Detection, Diagnosis, and Staging [Internet]. [cited 2019 Feb 5]. Available from: <https://www.cancer.org/content/cancer/en/cancer/endometrial-cancer/references.html>
44. Gynecologic Cancers | CDC [Internet]. [cited 2019 Feb 5]. Available from: <https://www.cdc.gov/cancer/gynecologic/index.htm>
45. Rogers LJ, Cuello MA. Cancer of the vulva. *Int J Gynecol Obstet.* 2018;143:4–13.
46. Adams TS, Cuello MA. Cancer of the vagina. *Int J Gynecol Obstet.* 2018;143:14–21.
47. Waddington CH. The epigenotype. 1942. *Int J Epidemiol.* 2012;41(1):10–3.
48. Herceg Z, Ushijima T. Epigenetics and cancer. Part A. Elsevier; 2010. 396 p.
49. Baylin SB, Jones PA. A decade of exploring the cancer epigenome — biological and translational implications. *Nat Rev Cancer* [Internet]. 2011 Oct 1 [cited 2019 Feb 4];11(10):726–34. Available from: <http://www.nature.com/articles/nrc3130>
50. Castaño A, Maurer MS. TARGETING EPIGENETICS IN CANCER. 2015;20(2):163–78.
51. Rasool M, Malik A, Naseer MI, Manan A, Ansari SA, Begum I, et al. The role of epigenetics in personalized medicine: Challenges and opportunities. *BMC Genomics* [Internet]. 2015;8(1):S5. Available from: <http://www.biomedcentral.com/1755-8794/8/S1/S5>
52. Cava C, Bertoli G, Castiglioni I. Integrating genetics and epigenetics in

- breast cancer: biological insights, experimental, computational methods and therapeutic potential. *BMC Syst Biol* [Internet]. 2015 Sep 21 [cited 2019 May 22];9:62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26391647>
53. Ruggero K, Farran-Matas S, Martinez-Tebar A, Aytes A. Epigenetic Regulation in Prostate Cancer Progression. *Curr Mol Biol Reports* [Internet]. 2018 Jun 18 [cited 2019 May 22];4(2):101–15. Available from: <http://link.springer.com/10.1007/s40610-018-0095-9>
 54. Kanwal R, Gupta S. Epigenetic modification in cancer. *Clin Genet* [Internet]. 2012;81(4):303–11. Available from: <http://onlinelibrary.wiley.com/doi/10.1111/j.1399-0004.2011.01809.x/full>
 55. Kulis M, Esteller M. DNA Methylation and Cancer. *Adv Genet* [Internet]. 2010 Jan 1 [cited 2019 Feb 4];70:27–56. Available from: <https://www.sciencedirect.com/science/article/pii/B9780123808660600022?via%3Dihub>
 56. Zhang M-W, Fujiwara K, Che X, Zheng S, Zheng L. DNA methylation in the tumor microenvironment. *J Zhejiang Univ Sci B* [Internet]. 2017;18(5):365–72. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28471108>
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5442975>
 57. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Nat Publ Gr* [Internet]. 2011;21(3):381–95. Available from: <http://dx.doi.org/10.1038/cr.2011.22>
 58. Cedar H, Bergman Y. Linking DNA methylation and histone modification:

- patterns and paradigms. *Nat Rev Genet* [Internet]. 2009 May 1 [cited 2019 Feb 4];10(5):295–304. Available from: <http://www.nature.com/articles/nrg2540>
59. Esteller M. Epigenetics provides a new generation of oncogenes and tumour-suppressor genes. 2006;179–83.
 60. Shen H, Laird PW. Interplay between the cancer genome and epigenome. *Cell* [Internet]. 2013 Mar 28 [cited 2019 May 17];153(1):38–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23540689>
 61. Muntean AG, Hess JL. Epigenetic dysregulation in cancer. *Am J Pathol*. 2009;175(4):1353–61.
 62. Audia JE, Campbell RM. Histone modifications and cancer markers. *Cold Spring Harb Perspect Biol* [Internet]. 2017;8(4):1–32. Available from: <http://cshperspectives.cshlp.org/lookup/doi/10.1101/cshperspect.a019521%0Ahttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3433108/pdf/ajcr0002-0589.pdf>
 63. Nicholson TB, Veland N, Chen T. Writers, Readers, and Erasers of Epigenetic Marks [Internet]. *Epigenetic Cancer Therapy*. Elsevier Inc.; 2015. 31–66 p. Available from: <http://dx.doi.org/10.1016/B978-0-12-800206-3.00003-3>
 64. Biswas S, Rao CM. Epigenetic tools (The Writers, The Readers and The Erasers) and their implications in cancer therapy. *Eur J Pharmacol* [Internet]. 2018;837:8–24. Available from: <https://doi.org/10.1016/j.ejphar.2018.08.021>
 65. Marmorstein R, Zhou M. C H A P T E R 4 Writers and Readers of Histone

Acetylation : Structure , Mechanism , and Inhibition. :1–26.

66. Marmorstein R, Trievel RC. Histone modifying enzymes: Structures, mechanisms, and specificities. *Biochim Biophys Acta - Gene Regul Mech* [Internet]. 2009;1789(1):58–68. Available from: <http://dx.doi.org/10.1016/j.bbagr.2008.07.009>
67. Bayat A. Clinical review, *Bioinformatics*. 2002;324(December 2005):1018–22.
68. Baxevanis AD, Ouellette BFF. *Bioinformatics : a practical guide to the analysis of genes and proteins* [Internet]. Wiley; 2005 [cited 2019 Mar 5]. 540 p. Available from: <https://www.wiley.com/en-us/Bioinformatics%3A+A+Practical+Guide+to+the+Analysis+of+Genes+and+Proteins%2C+3rd+Edition-p-9780471478782>
69. Tenenbaum JD. Translational Bioinformatics: Past, Present, and Future. *Genomics, Proteomics Bioinforma* [Internet]. 2016;14(1):31–41. Available from: <http://dx.doi.org/10.1016/j.gpb.2016.01.003>
70. Kori M, Yalcin Arga K. Potential biomarkers and therapeutic targets in cervical cancer: Insights from the meta-analysis of transcriptomics data within network biomedicine perspective. Meyers C, editor. *PLoS One* [Internet]. 2018 Jul 18 [cited 2019 May 21];13(7):e0200717. Available from: <https://dx.plos.org/10.1371/journal.pone.0200717>
71. van Dam P-JHH, Rolfo C, Giallombardo M, van Berckelaer C, Trinh XB, Altintas S, et al. In silico pathway analysis in cervical carcinoma reveals potential new targets for treatment. *Oncotarget*. 2016;7(3).
72. Ma C, Zhang W, Wu Q, Liu Y, Wang C, Lao G, et al. Identification of a

- microRNA signature associated with survivability in cervical squamous cell carcinoma. *PLoS One*. 2018;13(3):1–15.
73. Wong HS, Juan Y, Wu M, Zhang Y. Integrative bioinformatic analyses of an oncogenomic profile reveal the biology of endometrial cancer and guide drug discovery. 2015;7(5).
 74. Du D, Ma W, Yates MS, Chen T, Lu KH, Lu Y, et al. Predicting high-risk endometrioid carcinomas using proteins. 2018;9(28):19704–15.
 75. Mara TAO, Zhao M, Spurdle AB. Meta-analysis of gene expression studies in endometrial cancer identifies gene expression profiles associated with aggressive disease and patient outcome. *Nat Publ Gr* [Internet]. 2016;(July):1–9. Available from: <http://dx.doi.org/10.1038/srep36677>
 76. Fu A, Chang HR, Zhang ZF. Integrated multiomic predictors for ovarian cancer survival. *Carcinogenesis*. 2018;39(7):860–8.
 77. Kinose Y, Sawada K, Nakamura K, Sawada I, Toda A, Nakatsuka E, et al. The hypoxia-related microRNA miR-199a-3p displays tumor suppressor functions in ovarian carcinoma. *Oncotarget*. 2015;6(13).
 78. Wang F, Kao CJ, Huang RS, Division SC. High expression of miR-532-5p, a tumor suppressor, leads to better prognosis in ovarian cancer both in vivo and in vitro. 2017;15(5):1123–31.
 79. Creighton CJ, Hernandez-Herrera A, Jacobsen A, Levine DA, Mankoo P, Schultz N, et al. Integrated analyses of microRNAs demonstrate their widespread influence on gene expression in high-grade serous ovarian carcinoma. *PLoS One*. 2012;7(3).
 80. Xu Y, Zhang S, Lin S, Guo Y, Deng W, Zhang Y, et al. WERAM: A database

- of writers, erasers and readers of histone acetylation and methylation in eukaryotes. *Nucleic Acids Res.* 2017;45(D1):D264–70.
81. Khare SP, Habib F, Sharma R, Gadewal N, Gupta S, Galande S. Histone - A relational knowledgebase of human histone proteins and histone modifying enzymes. *Nucleic Acids Res.* 2012;40(D1):337–42.
 82. Lyko F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. *Nat Rev Genet* [Internet]. 2018;19(2):81–92. Available from: <http://dx.doi.org/10.1038/nrg.2017.80>
 83. Li E, Zhang Y. DNA Methylation in Mammals. *Cold Spring Harb Perspect Biol.* 2014;6:a019133.
 84. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: A hub for protein information. *Nucleic Acids Res.* 2015;43(D1):D204–12.
 85. Home | HUGO Gene Nomenclature Committee [Internet]. [cited 2019 Feb 5]. Available from: <https://www.genenames.org/>
 86. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res* [Internet]. 2011 Apr [cited 2019 May 6];21(4):564. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21423274>
 87. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles G V., et al. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics.* 2013;14.
 88. Wang X, Guan Z, Sheng Y, Tian Y. Real-time quantitative description of gas invasion in deepwater drilling. *J Eng Res.* 2018;6(2):232–43.
 89. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach

- for understanding genome variations in KEGG. *Nucleic Acids Res* [Internet]. 2019 Jan 8 [cited 2019 Feb 5];47(D1):D590–5. Available from: <https://academic.oup.com/nar/article/47/D1/D590/5128935>
90. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401–4.
 91. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal. *Sci Signal*. 2013;6:269.
 92. Mayta-Tristán P, Huamaní C, Montenegro-Idrogo JJ, Samanez-Figari C, Gonzáles-Alcaide G. SCIENTIFIC PRODUCTION AND CANCER-RELATED COLLABORATION NETWORKS IN PERU 2000-2011 : A BIBLIOMETRIC STUDY IN SCOPUS. 2013;30(1):31–6.
 93. Lawrence MS, Stojanov P, Polak P, Kryukov G V, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499:214–218.
 94. Dingerdissen HM, Torcivia-Rodriguez J, Hu Y, Chang TC, Mazumder R, Kahsay R. BioMuta and BioXpress: Mutation and expression knowledgebases for cancer biomarker discovery. *Nucleic Acids Res*. 2018;46(D1):D1128–36.
 95. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Bork P, Kondrashov AS, et al. A method and server for predicting damaging missense mutations. 2010;7(4):248–9.
 96. Capriotti E, Fariselli P, Rossi I, Casadio R. A three-state prediction of single

- point mutations on protein stability changes. *BMC Bioinformatics*. 2008;9(SUPPL. 2):1–9.
97. Capriotti E, Fariselli P, Casadio R. I-Mutant2 . 0 : predicting stability changes upon mutation from the protein sequence or structure. 2005;33:306–10.
 98. Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. A generic method for assignment of reliability scores applied to solvent accessibility predictions. *BMC Struct Biol*. 2009;9:1–10.
 99. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukheim R. GISTIC2 . 0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* [Internet]. 2011;12(4):R41. Available from: <http://genomebiology.com/2011/12/4/R41>
 100. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*. 2017;45(D1):D896–901.
 101. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45(W1):W98–102.
 102. GTEx Consortium TGte. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* [Internet]. 2013 Jun [cited 2019 Apr 30];45(6):580–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23715323>
 103. Nogales-cadenas R, Carmona-saez P, Vazquez M, Vicente C, Yang X, Tirado F, et al. GeneCodis : interpreting gene lists through enrichment analysis and integration of diverse biological information.

2009;37(May):317–22.

104. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 2017;45(D1):D362–8.
105. Chin C-H, Chen S-H, Wu H-H, Ho C-W, Ko M-T, Lin C-Y. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* [Internet]. 2014 [cited 2019 May 5];8 Suppl 4(Suppl 4):S11. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25521941>
106. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* [Internet]. 2003 Nov [cited 2019 May 5];13(11):2498–504. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14597658>
107. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. *Nature.* 2017;550(7675):204–13.
108. Goldman M, Craft B, Brooks A, Zhu J, Haussler D. The UCSC Xena Platform for cancer genomics data visualization and interpretation. [cited 2019 Mar 28]; Available from: <http://dx.doi.org/10.1101/326470>
109. Huang WY, Hsu S Da, Huang HY, Sun YM, Chou CH, Weng SL, et al. MethHC: A database of DNA methylation and gene expression in human cancer. *Nucleic Acids Res.* 2015;43(D1):D856–61.
110. Koch A, De Meyer T, Jeschke J, Van Criekinge W. MEXPRESS: Visualizing expression, DNA methylation and clinical TCGA data. *BMC*

- Genomics [Internet]. 2015;16(1):1–6. Available from: <http://dx.doi.org/10.1186/s12864-015-1847-z>
111. Global Database on Body Mass Index - World Health Organization [Internet]. [cited 2019 May 6]. Available from: <http://www.assessmentpsychology.com/icbmi.htm>
 112. Estadificación del cáncer - National Cancer Institute [Internet]. [cited 2019 May 6]. Available from: <https://www.cancer.gov/espanol/cancer/diagnostico-estadificacion/estadificacion>
 113. Alcarraz C, Muñiz J, Mas L, Olivera M, Morante Z, Alvvarez M, et al. CITORREDUCCIÓN ÓPTIMA EN CÁNCER DE OVARIO AVANZADO TRATADO CON PACLITAXEL A DOSIS DENSA Y CARBOPLATINO SEGUIDO DE CIRUGÍA DE INTERVALO EN EL INSTITUTO NACIONAL DE ENFERMEDADES NEOPLÁSICAS DEL PERÚ. 2018;35(1):46–54.
 114. Urrutia C, Santos C, Zevallos A, Valdivia H, Velarde C, Milla E, et al. COMPROMISO PARAMETRIAL EN PACIENTES CON CÁNCER DE CÉRVIX IB-1 TRATADAS CON HISTERECTOMÍA RADICAL EN PERÚ. 2014;31(3):431–6.
 115. Ruiz R, Serrano M, Ruiz EF, Mantilla R, Valdivieso N, Olivera M, et al. CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS Y SOBREVIVENCIA EN MUJERES JÓVENES CON CÁNCER CERVICAL: ANÁLISIS RETROSPECTIVO DEL INSTITUTO NACIONAL DE ENFERMEDADES NEOPLÁSICAS. 2017;34(2):218–27.

116. Wang F, Marshall CB, Ikura M. Transcriptional/epigenetic regulator CBP/p300 in tumorigenesis: structural and functional versatility in target recognition. *Cell Mol Life Sci* [Internet]. 2013 Nov 11 [cited 2019 Apr 4];70(21):3989–4008. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23307074>
117. Cho Y-W, Hong T, Hong S, Guo H, Yu H, Kim D, et al. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* [Internet]. 2007 Jul 13 [cited 2019 Apr 4];282(28):20395–406. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17500065>
118. Shankaranarayanan P, Chaitidis P, Kühn H, Nigam S. Acetylation by Histone Acetyltransferase CREB-binding Protein/p300 of STAT6 Is Required for Transcriptional Activation of the 15-Lipoxygenase-1 Gene. *J Biol Chem* [Internet]. 2001 Nov 16 [cited 2019 Apr 4];276(46):42753–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11509556>
119. Donaldson-Collier MC, Sungalee S, Zufferey M, Tavernari D, Katanayeva N, Battistello E, et al. EZH2 oncogenic mutations drive epigenetic, transcriptional, and structural changes within chromatin domains. *Nat Genet* [Internet]. 2019 Mar 28 [cited 2019 May 20];51(3):517–28. Available from: <http://www.nature.com/articles/s41588-018-0338-y>
120. Turner-Ivey B, Guest ST, Irish JC, Kappler CS, Garrett-Mayer E, Wilson RC, et al. KAT6A, a Chromatin Modifier from the 8p11-p12 Amplicon is a Candidate Oncogene in Luminal Breast Cancer. *Neoplasia* [Internet]. 2014;16(8):644–55. Available from:

<http://dx.doi.org/10.1016/j.neo.2014.07.007>

121. Wu M, Fan B, Guo Q, Li Y, Chen R, Lv N, et al. Knockdown of SETDB1 inhibits breast cancer progression by miR-381-3p-related regulation. *Biol Res* [Internet]. 2018;51(1):1–11. Available from: <https://doi.org/10.1186/s40659-018-0189-0>
122. Schrecengost RS, Dean JL, Goodwin JF, Schiewer MJ, Urban MW, Stanek TJ, et al. USP22 regulates oncogenic signaling pathways to drive lethal cancer progression. *Cancer Res* [Internet]. 2014 Jan 1 [cited 2019 May 20];74(1):272–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24197134>
123. Zhang J, Vlasevska S, Wells VA, Nataraj S, Holmes AB, Duval R, et al. The CREBBP Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer Discov* [Internet]. 2017 [cited 2019 May 20];7(3):322–37. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28069569>
124. Singh GK, Azuine RE, Siahpush M. Global Inequalities in Cervical Cancer Incidence and Mortality are Linked to Deprivation, Low Socioeconomic Status, and Human Development Global Inequalities in Cervical Cancer. *Int J MCH AIDS* [Internet]. 2012;1(1):2161–8674. Available from: www.mchandaids.org
125. American Cancer Society. Ovarian Cancer Early Detection, Diagnosis, and Staging [Internet]. [cited 2019 May 18]. Available from: <https://www.cancer.org/content/dam/CRC/PDF/Public/8775.00.pdf>
126. Demark-Wahnefried W, Platz EA, Ligibel JA, Blair CK, Courneya KS,

- Meyerhardt JA, et al. The role of obesity in cancer survival and recurrence. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2012 Aug [cited 2019 Jun 5];21(8):1244–59. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22695735>
127. Schmandt RE, Iglesias DA, Co NN, Lu KH. Understanding obesity and endometrial cancer risk: opportunities for prevention. *Am J Obstet Gynecol* [Internet]. 2011 Dec [cited 2019 Jun 5];205(6):518–25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21802066>
 128. Zhang X, Wen H, Shi X. Lysine methylation: beyond histones. *Acta Biochim Biophys Sin (Shanghai)* [Internet]. 2012 Jan 22 [cited 2019 May 17];44(1):14–27. Available from: <https://academic.oup.com/abbs/article-lookup/doi/10.1093/abbs/gmr100>
 129. Filipp F V. Crosstalk between epigenetics and metabolism-Yin and Yang of histone demethylases and methyltransferases in cancer. *Brief Funct Genomics*. 2017;16(6):320–5.
 130. Kawazu M, Saso K, Tong KI, McQuire T, Goto K, Son DO, et al. Histone demethylase JMJD2B functions as a co-factor of estrogen receptor in breast cancer proliferation and mammary gland development. *PLoS One*. 2011;6(3).
 131. Shi L, Sun L, Li Q, Liang J, Yu W, Yi X, et al. Histone demethylase JMJD2B coordinates H3K4/H3K9 methylation and promotes hormonally responsive breast carcinogenesis. *Proc Natl Acad Sci*. 2011;108(18):7541–6.
 132. Wang JK, Tsai MC, Poulin G, Adler AS, Chen S, Liu H, et al. The histone demethylase UTX enables RB-dependent cell fate control. *Genes Dev*.

- 2010;24(4):327–32.
133. Qiu M-T, Fan Q, Zhu Z, Kwan S-Y, Chen L, Chen J-H, et al. KDM4B and KDM4A promote endometrial cancer progression by regulating androgen receptor, c-myc, and p27^{kip1}. *Oncotarget*. 2015;6(31).
 134. Wilson C, Qiu L, Hong Y, Karnik T, Tadros G, Mau B, et al. The histone demethylase KDM4B regulates peritoneal seeding of ovarian cancer. *Oncogene*. 2017;36(18):2565–76.
 135. Soto DR, Barton C, Munger K, McLaughlin-Drubin ME. KDM6A addiction of cervical carcinoma cell lines is triggered by E7 and mediated by p21CIP1 suppression of replication stress. Kalejta RF, editor. *PLOS Pathog* [Internet]. 2017 Oct 2 [cited 2019 May 18];13(10):e1006661. Available from: <https://dx.plos.org/10.1371/journal.ppat.1006661>
 136. Schulz WA, Lang A, Koch J, Greife A. The histone demethylase UTX/KDM6A in cancer: Progress and puzzles. *Int J Cancer*. 2019;
 137. Lee TI, Young RA. Transcriptional Regulation and its Misregulation in Disease. *Cell* [Internet]. 2013 Mar 14 [cited 2019 May 18];152(6):1237. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23498934>
 138. Brehm A, Nielsen SJ, Miska EA, McCance DJ, Reid JL, Bannister AJ, et al. The E7 oncoprotein associates with Mi2 and histone deacetylase activity to promote cell growth. *EMBO J*. 1999;18(9):2449–58.
 139. Guise AJ, Budayeva HG, Diner BA, Cristea IM. Histone deacetylases in herpesvirus replication and virus-stimulated host defense. *Viruses*. 2013;5(7):1607–32.
 140. Yan D, Yi S, Chiu WC, Qin LG, Kin WH, Hung CTK, et al. Integrated

- analysis of chromosome copy number variation and gene expression in cervical carcinoma. *Oncotarget*. 2017;8(65):108912–22.
141. Campbell BB, Light N, Fabrizio D, Zatzman M, Fuligni F, de Borja R, et al. Comprehensive Analysis of Hypermutation in Human Cancer. *Cell* [Internet]. 2017 Nov 16 [cited 2019 May 22];171(5):1042-1056.e10. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29056344>
 142. Manuscript A. Diverse functions of PHD fingers of the MLL / KMT2 subfamily. 2015;1843(2):366–71.
 143. Jagsi R, Jiang J, Momoh AO, Alderman A, Giordano SH, Buchholz TA, et al. Hijacked in cancer: the MLL/KMT2 family of methyltransferases. 2017;263(2):219–27.
 144. Froimchuk E, Jang Y, Ge K. Histone H3 lysine 4 methyltransferase KMT2D. *Gene* [Internet]. 2017 [cited 2019 Mar 28];627:337. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5546304/>
 145. Schubert HL, Blumenthal RM, Cheng X. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem Sci* [Internet]. 2003 Jun [cited 2019 Apr 4];28(6):329. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12826405>
 146. Schultz N, Dorso M, Pastore A, Sinha A, Li Q, Gala K, et al. KMT2C mediates the estrogen dependence of breast cancer through regulation of ER α enhancer function. *Oncogene* [Internet]. 2018;37(34):4692–710. Available from: <http://dx.doi.org/10.1038/s41388-018-0273-5>
 147. Gizard F, Lavallée B, DeWitte F, Hum DW. A Novel Zinc Finger Protein TReP-132 Interacts with CBP/p300 to Regulate Human CYP11A1 Gene

- Expression. *J Biol Chem* [Internet]. 2001 Sep 7 [cited 2019 Mar 28];276(36):33881–92. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11349124>
148. Holmfeldt L, Wei L, Diaz-Flores E, Walsh M, Zhang J, Ding L, et al. The Genomic Landscape of Hypodiploid Acute. *Nat Genet*. 2013;45(3):242–52.
 149. Thakur JK, Yadav A, Yadav G. Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Res* [Internet]. 2014 Feb [cited 2019 Mar 28];42(4):2112–25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24253305>
 150. Paulussen ADC, Stegmann APA, Blok MJ, Tserpelis D, Posma-Velter C, Detisch Y, et al. MLL2 mutation spectrum in 45 patients with Kabuki syndrome. *Hum Mutat*. 2011;32(2):2018–25.
 151. Michor F. Chromosomal instability and human cancer. 2005;(March):631–5.
 152. Shlien A, Malkin D. Copy number variations and cancer susceptibility. *Curr Opin Oncol*. 2010;22(1):55–63.
 153. Dong G, Mao Q, Yu D, Zhang Y, Qiu M, Dong G, et al. Integrative analysis of copy number and transcriptional expression profiles in esophageal cancer to identify a novel driver gene for therapy. *Sci Rep* [Internet]. 2017;7(August 2016). Available from: <http://dx.doi.org/10.1038/srep42060>
 154. Davidson B, Nymoén DA, Elgaaen BV, Staff AC, Tropé CG, Kærn J, et al. BUB1 mRNA is significantly co-expressed with AURKA and AURKB mRNA in advanced-stage ovarian serous carcinoma. 2014;701–7.
 155. Suman S, Mishra A. Network analysis revealed aurora kinase dysregulation

- in five gynecological types of cancer. *Oncol Lett.* 2018;15(1):1125–32.
156. Kurai M, Shiozawa T, Shih HC, Miyamoto T, Feng YZ, Kashima H, et al. Expression of Aurora kinases A and B in normal, hyperplastic, and malignant human endometrium: Aurora B as a predictor for poor prognosis in endometrial carcinoma. *Hum Pathol.* 2005;36(12):1281–8.
 157. Ashraf W, Ibrahim A, Alhosin M, Zaayter L, Ouararhni K, Papin C, et al. The epigenetic integrator UHRF1 : on the road to become a universal biomarker for cancer. 2017;8(31):51946–62.
 158. Rothbart SB, Krajewski K, Nady N, Tempel W, Xue S, Badeaux AI, et al. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* [Internet]. 2012 Nov [cited 2019 May 20];19(11):1155–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23022729>
 159. Sarma K, Blais A, Reinberg D, Margueron R, Woodcock CL, Dynlacht BD, et al. Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Mol Cell* [Internet]. 2008;32(4):503–18. Available from: <http://dx.doi.org/10.1016/j.molcel.2008.11.004>
 160. Shen J, Zhang Y, Yu H, Shen B, Liang Y, Jin R, et al. Role of DUSP1/MKP1 in tumorigenesis, tumor progression and therapy. *Cancer Med.* 2016;5(8):2061–8.
 161. Cremona CA, Behrens A. ATM signalling and cancer. *Br Dent J* [Internet]. 2014;217(1):3351–60. Available from: <http://dx.doi.org/10.1038/onc.2013.275>
 162. Oki S, Sone K, Oda K, Hamamoto R, Ikemura M, Maeda D, et al. Oncogenic

- histone methyltransferase EZH2: A novel prognostic marker with therapeutic potential in endometrial cancer. *Oncotarget*. 2017;8(25):40402–11.
163. Tian X, Zhang S, Liu H-M, Zhang Y-B, Blair CA, Mercola D, et al. Histone lysine-specific methyltransferases and demethylases in carcinogenesis: new targets for cancer therapy and prevention. *Curr Cancer Drug Targets* [Internet]. 2013;13(5):558–79. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3703250&tool=pmcentrez&rendertype=abstract>
 164. Taube JH, Sphyris N, Johnson KS, Reisenauer KN, Nesbit TA, Joseph R, et al. The H3K27me3-demethylase KDM6A is suppressed in breast cancer stem-like cells, and enables the resolution of bivalency during the mesenchymal-epithelial transition. *Oncotarget* [Internet]. 2017;5(0):65548–65. Available from: <http://www.oncotarget.com/fulltext/19214>
 165. Shan W, Wang C, Zhang Z, Luo X, Ning C, Yu Y, et al. ATM may be a protective factor in endometrial carcinogenesis with the progesterone pathway. *Tumor Biol* [Internet]. 2015 Mar 22 [cited 2019 Mar 24];36(3):1529–37. Available from: <http://link.springer.com/10.1007/s13277-014-2712-4>
 166. Wang L, Li N, Zhang H, Luo X, Dai J, Zhou S, et al. Structure of human SMYD2 protein reveals the basis of p53 tumor suppressor methylation. *J Biol Chem*. 2011;286(44):38725–37.
 167. Saddic LA, West LE, Aslanian A, Yates JR, Rubin SM, Gozani O, et al. Methylation of the retinoblastoma tumor suppressor by SMYD2. *J Biol*

Chem. 2010;285(48):37733–40.

168. Komatsu S, Imoto I, Tsuda H, Kozaki K, Muramatsu T, Shimada Y, et al. Overexpression of SMYD2 relates to tumor cell proliferation and malignant outcome of esophageal squamous cell carcinoma. *Carcinogenesis* [Internet]. 2009 Jul 1 [cited 2019 May 21];30(7):1139–46. Available from: <https://academic.oup.com/carcin/article-lookup/doi/10.1093/carcin/bgp116>
169. Hamamoto R, Nakamura Y. Dysregulation of protein methyltransferases in human cancer: An emerging target class for anticancer therapy. *Cancer Sci* [Internet]. 2016 Apr 1 [cited 2019 May 21];107(4):377–84. Available from: <http://doi.wiley.com/10.1111/cas.12884>
170. Wapenaar H, Dekker FJ. Histone acetyltransferases: challenges in targeting bi-substrate enzymes. *Clin Epigenetics* [Internet]. 2016 [cited 2019 May 21];8:59. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27231488>
171. Yin MJ, Shao L, Voehringer D, Smeal T, Jallal B. The Serine/Threonine Kinase Nek6 Is Required for Cell Cycle Progression through Mitosis. *J Biol Chem*. 2003;278(52):52454–60.
172. Pozo K, Bibb JA. The Emerging Role of Cdk5 in Cancer. *Trends in cancer* [Internet]. 2016 [cited 2019 May 21];2(10):606–18. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27917404>
173. Ikeda Y, Oda K, Ishihara H, Wada-Hiraike O, Miyasaka A, Kashiyama T, et al. Prognostic importance of CDK4/6-specific activity as a predictive marker for recurrence in patients with endometrial cancer, with or without adjuvant chemotherapy. *Br J Cancer* [Internet]. 2015;113(10):1477–83. Available from: <http://dx.doi.org/10.1038/bjc.2015.369>

174. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. 2007;(November):569–78.
175. Anzick SL, Kononen J, Walker RL, Azorsa D, Tanner MM, Guan X, et al. AIB1, a Steroid Receptor Coactivator Amplified in Breast and Ovarian Cancer. 1998;965–8.
176. Definición de sistema reproductivo - Diccionario de cáncer del NCI - National Cancer Institute [Internet]. [cited 2019 Feb 25]. Available from: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/reproductive-system>
177. Agrawal K, Das V, Vyas P, Hajdúch M. Nucleosidic DNA demethylating epigenetic drugs – A comprehensive review from discovery to clinic. Pharmacol Ther [Internet]. 2018 Aug 1 [cited 2019 Feb 27];188:45–79. Available from: <https://www.sciencedirect.com/science/article/pii/S0163725818300317?via%3Dihub#s0010>
178. Tollefsbol TO. Cancer epigenetics [Internet]. CRC Press; 2009 [cited 2019 Feb 27]. 446 p. Available from: <https://books.google.com.pe/books?id=QH9Fjg5CFDoC&printsec=frontcover&dq=Trygve+Tollefsbol+Cancer+Epigenetics&hl=es-419&sa=X&ved=0ahUKEwix3dPmmdzgAhXynuAKHT14DHgQ6wEIKjAA#v=onepage&q=Trygve+Tollefsbol+Cancer+Epigenetics&f=false>
179. Gardner KE, Allis CD, Strahl BD. Operating on chromatin, a colorful language where context matters. 2012;409(1):36–46.

ANEXO 1

Tabla 1A. Genes implicados en mecanismos epigenéticos clasificados en las categorías de escritores y borradores.

GENES ESCRITORES				
N.º	Símbolo del gen	Sinónimo	Nombre de la Proteína	Localización Citogenética
Metilación del ADN				
1	DNMT1	CXXC9; AIM; DNMT	ADN (citosina-5) -metiltransferasa 1	19p13.2
2	DNMT3A*	-	ADN (citosina-5) -metiltransferasa 3A	2p23.3
3	DNMT3B*	-	ADN (citosina-5) -metiltransferasa 3B	20q11.21
4	DNMT3L*	-	ADN (citosina-5) -metiltransferasa tipo3	21q22.3
5	TRDMT1	DNMT2	ARNt (citosina (38) -C (5)) - metiltransferasa	10p13
Acetilación de lisina				
6	CDY1B	CDY1	Proteína de cromodominio específica para testículos Y1	Yq11.23
7	CDYL*	CDYL1	Proteína de tipo cromodominio Y	6p25.1
8	CLOCK	BHLHE8, KIAA0334	Ciclo de salida del locomotor circadiana proteína kaput	4q12
9	CREBBP*	CBP, KAT3A	Proteína de unión a CREB	16p13.3
10	ELP3	KAT9	Proteína 3 del complejo alargador	8p21.1
11	EP300*	p300; KAT3B	Histona acetiltransferasa p300	22q13.2
12	GTF3C4	KAT12	Factor de transcripción general 3C polipéptido 4	9q34.13
13	HAT1	KAT1	Subunidad catalítica de histona acetiltransferasa tipo B	2q31.1
14	KAT2A*	GCN5, GCN5L2, HGCN5	Histona acetiltransferasa KAT2A	17q21.2
15	KAT2B*	PCAF	Histona acetiltransferasa KAT2B	3p24.3
16	KAT5	HTATIP, TIP60	Histona acetiltransferasa KAT5	11q13.1
17	KAT6A*	MOZ, MYST3, RUNXBP2, ZNF220	Histona acetiltransferasa KAT6A	8p11.21
18	KAT6B*	KIAA0383; MORE; MOZ2; MYST4	Histona acetiltransferasa KAT6B	10q22.2
19	KAT7	MYST2, HBO1, HBOa	Histona acetiltransferasa KAT7	17q21.33
20	KAT8	MOF, MYST1, PP7073	Histona acetiltransferasa KAT8	16p11.2

21	MGEA5	OGA, HEXC, KIAA0679, MEA5	Proteína O-GlcNAcase	10q24.32
Acetilación de lisina				
22	NAT10	ALP, KIAA1709	N-acetiltransferasa 10	11p13
23	NCOA1	BHLHE74, SRC1	Coactivador de receptores nucleares 1	2p23.3
24	NCOA3	AIB1, BHLHE42, RAC3, TRAM1	Coactivador del receptor nuclear 3	20q13.12
Biotinilación de lisina				
25	HLCS	-	Biotina - proteína ligasa	21q22.13
Citrulinización de arginina				
26	PADI4	PAD4, PADI5, PDI5	Proteína-arginina deiminasa tipo 4	1p36.13
Fosforilación de serina / treonina / tirosina				
27	ATM	-	Serina-proteína quinasa ATM	11q22.3
28	ATR	FRP1	Serina / treonina-proteína quinasa ATR	3q23
29	AURKB	AIK2, AIM1, AIRK2, ARK2, STK1, STK12, STK5	Serina / treonina-proteína quinasa 12	17p13.1
30	BAZ1B	WBSCR10, WBSCR10, WBSCR9, WSTF	Tirosina-proteína quinasa BAZ1B	7q11.23
31	BUB1	BUB1L	Punto de control mitótico serina / treonina-proteína quinasa BUB1	2q13
32	CDK17	PCTAIRE2, PCTK2	Quinasa dependiente de ciclina 17	12q23.1
33	CDK3	CDKN3	Quinasa dependiente de ciclina 3	17q25.1
34	CDK5	CDKN5	Quinasa dependiente de ciclina 5	7q36.1
35	CHUK	IKKA, TCF16	Inhibidor del factor nuclear kappa-B quinasa subunidad alfa	10q24.31
36	DAPK3	ZIPK	Proteína quinasa asociada a la muerte 3	19p13.3
37	GSK3B	-	Glicógeno sintasa quinasa-3 beta	3q13.33
38	HASPIN	GSG2	Serina / treonina-proteína quinasa haspin	17p13.2
39	JAK2	-	Tirosina-proteína quinasa JAK2	9p24.1
40	LIMK2	-	Dominio LIM quinasa 2	22q12.2
41	MAP3K20	MLTK, ZAK, HCCS4	Proteína quinasa quinasa quinasa activada por mitógeno 20	2q31.1
42	MAP3K8	COT, ESTF	Proteína quinasa quinasa quinasa activada por mitógeno 8	10p11.23
43	MASTL	GW, GWL, THC2	Serina / treonina-proteína quinasa asociada a los microtúbulos	10p12.1
44	NEK6	-	Serina / treonina-proteína quinasa Nek6	9q33.3
Fosforilación de serina / treonina / tirosina				
45	NEK9	KIAA1995, NEK8, NERCC	Serina / treonina-proteína quinasa Nek9	14q24.3

46	PAK2	-	Serina / treonina-proteína quinasa PAK 2	3q29
47	PKN1	PAK1, PKN, PRK1, PRKCL1	Serina / treonina-proteína quinasa N1	19p13.12
48	PRKCB	PKCB, PRKCB1	Tipo de proteína quinasa C beta	16p12.2-p12.1
49	PRKCD	-	Tipo de proteína quinasa C delta	3p21.1
50	PRKDC	HYRC, HYRC1	Subunidad catalítica de proteína quinasa dependiente de ADN	8q11.21
51	RPS6KA4	MSK2	Proteína ribosomal S6 quinasa alfa-4	11q13.1
52	RPS6KA5	MSK1	Proteína ribosomal S6 quinasa alfa-5	14q32.11
53	STK10	LOK	Serina / treonina-proteína quinasa 10	5q35.1
54	STK4	KRS2, MST1	Serina / treonina-proteína quinasa 4	20q13.12
55	TLK1	KIAA0137	Serina/treonina proteína quinasa desordenado como quinasa 1	2q31.1
Metilación de arginina				
56	CARM1	PRMT4	Histona-arginina metiltransferasa CARM1	19p13.2
57	PRMT1	ANM1	Proteína arginina N-metiltransferasa	19q13.33
58	PRMT2	HMT1, HRMT1L1	Proteína arginina N-metiltransferasa 2	21q22.3
59	PRMT5	ANM5	Proteína arginina N-metiltransferasa 5	14q11.2
60	PRMT6	ANM6	Proteína arginina N-metiltransferasa 6	1p13.3
61	PRMT7	KIAA1933	Proteína arginina N-metiltransferasa 7	16q22.1
62	PRMT8	HRMT1L3, HRMT1L4	Proteína arginina N-metiltransferasa 8	12p13.32
Metilación de lisina				
63	PRDM2	RIZ, KMT8, RIZ1, RIZ2, MTB-ZF, HUMHOXY1	Dominio PR proteína dedo de zinc 2	1p36.21
64	ASH1L*	KMT2H, KIAA1420, ASH1L1	Histona-lisina N-metiltransferasa ASH1L	1q22
65	DOT1L	-	Histona-lisina N-metiltransferasa, específica de H3 lisina 79	19p13.3
66	EHMT1*	GLP, GLP1, KMT1D, FP13812, EUHMTASE1, Eu-HMTase1, bA188C12.1, KIAA1876	Histona-lisina N-metiltransferasa EHMT1	9q34.3
67	EHMT2*	G9A, BAT8, GAT8, NG36, KMT1C, C6orf30	Histona-lisina N-metiltransferasa EHMT2	6p21.33
68	EZH1	KMT6B	Histona-lisina N-metiltransferasa EZH1	17q21.2
Metilación de lisina				
69	EZH2	WVS, ENX1, EZH1, KMT6, WVS2, ENX-1, EZH2b, KMT6A	Histona-lisina N-metiltransferasa EZH2	7q36.1

70	KMT2A*	MLL, HRX, TRX1, CXXC7, ALL-1, MLL1, HTRX1, MLL1A, WDSTS, MLL/GAS7, TET1-MLL	Histona-lisina N-metiltransferasa 2A	11q23.3
71	KMT2B	HRX2, MLL2, MLL4, TRX2, WBP7, MLL1B, WBP-7	Histona-lisina N-metiltransferasa 2B	19q13.12
72	KMT2C	MLL3, HALR, KIAA1506	Histona-lisina N-metiltransferasa 2C	7q36.1
73	KMT2D	ALR, KMS, MLL2, MLL4, AAD10, KABUK1, TNRC21, CAGL114	Histona-lisina N-metiltransferasa 2D	12q13.12
74	KMT2E	MLL5, NKp44L, HDCMC04P	Histona-lisina N-metiltransferasa 2E	7q22.3
75	KMT5A	SETD8, SET8, KMT5A, SET07, PR-Set7, PRSET7	N-lisina metiltransferasa KMT5A	12q24.31
76	KMT5B	SUV420H1, CGI85, CGI-85	Histona-lisina N-metiltransferasa KMT5B	11q13.2
77	KMT5C	SUV420H2, PP7130	Histona-lisina N-metiltransferasa KMT5C	19q13.42
78	NSD1*	STO, KMT3B, SOTOS, ARA267, SOTOS1	Histona-lisina N-metiltransferasa, específicas de H3 lisina 36 y H4 lisina 20	5q35.3
79	NSD2*	WHSC1, TRX5, MMSET, REIIBP, KIAA1090	Probable histona-lisina N-metiltransferasa NSD2	4p16.3
80	NSD3*	WHSC1L1, pp14328, DC28	Histona-lisina N-metiltransferasa NSD3	8p11.23
81	PRDM9	PFM6, KMT8B, MSBP3, ZNF899, MEISETZ	Histona-lisina N-metiltransferasa PRDM9	5p14.2
82	SETD1A	KIAA0339, KMT2F, SET1, SET1A	Histona-lisina N-metiltransferasa SETD1A	16p11.2
83	SETD1B	KMT2G, Set1B, KIAA1076	Histona-lisina N-metiltransferasa SETD1B	12q24.31
84	SETD2	HYPB, SET2, HIF-1, HIP-1, KMT3A, HBP231, HSPC069, P231hbp, HIF1, KIAA1732	Histona-lisina N-metiltransferasa SETD2	3p21.31
85	SETD7	SETD7, KMT7, SET7, SET9, SET7/9, KIAA1717	Histona-lisina N-metiltransferasa SETD7	4q31.1
86	SETDB1	ESET, KG1T, KIAA0067, KMT1E, TDRD21, H3-K9-MTase4	Histona-lisina N-metiltransferasa SETDB1	1q21.3
87	SETDB2	SETDB2, CLLD8, CLLL8, KMT1F, C13orf4	Histona-lisina N-metiltransferasa SETDB2	13q14.2
88	SETMAR	Mar1, HsMar1, METNASE	Histona-lisina N-metiltransferasa SETMAR	3p26.1
89	SMYD2	KMT3C, HSKM-B, ZMYND14	N-lisina metiltransferasa SMYD2	1q32.3
90	SMYD3	KMT3E, ZMYND1, ZNFN3A1, bA74P14.1	SET y MYND proteína que contiene el dominio 3	1q44
91	SUV39H1*	SUV39H, MG44, KMT1A, H3-K9-HMTase1	Histona-lisina N-metiltransferasa SUV39H1	Xp11.23
92	SUV39H2	KMT1B	Histona-lisina N-metiltransferasa SUV39H2	10p13
Ribosilación de lisina				
93	PARP1	ADPRT, PPOL	Poli [ADP-ribosa] polimerasa 1	1q42.12

Ubiquitinación de lisina				
94	BAP1	RNF2, DING, HIP13, RING1B	Ubiquitina carboxilo terminal hidrolasa BAP1	3p21.1
95	DTX3L	BBAP	E3 Ubiquitina ligasa DTX3L	3q21.1
96	HUWE1	KIAA0312, KIAA1578, URB1, HSPC272	E3 Ubiquitina ligasa HUWE1	Xp11.22
97	RAG1	RNF74	V (D) J proteína de activación de recombinación 1	11p12
98	RBX1	RNF75, ROC1	E3 Ubiquitina ligasa RBX1	22q13.2
99	RING1	RNF1	E3 Ubiquitina ligasa RING1	6p21.32
100	RNF168	-	E3 Ubiquitina ligasa RNF168	3q29
101	RNF20	BRE1A	E3 Ubiquitina ligasa BRE1A	9q31.1
102	RNF40	BRE1B, KIAA0661	E3 Ubiquitina ligasa BRE1B	16p11.2
103	RNF8	KIAA0646	E3 Ubiquitina ligasa RNF8	6p21.2
104	UBR2	C6orf133, KIAA0349	E3 Ubiquitina ligasa UBR2	6p21.1
105	UHRF1*	ICBP90, NP95, RNF106	E3 Ubiquitina ligasa UHRF1	19p13.3
GENES BORRADORES				
Desmetilación del ADN				
106	TET1	CXXC6, LCX	Metilcitosina dioxigenasa TET1	10q21.3
107	TET2	KIAA1546, Nbla00191	Metilcitosina dioxigenasa TET2	4q24
108	TET3	KIAA0401	Metilcitosina dioxigenasa TET3	2p13.1
Desacetilación de lisina				
109	HDAC1	RPD3L1	Histona deacetilasa 1	1p35.2-p35.1
110	HDAC10	HDA10	Histona deacetilasa 10	22q13.33
111	HDAC11	HDA11	Histona deacetilasa 11	3p25.1
112	HDAC2	-	Histona deacetilasa 2	6q21
113	HDAC3	-	Histona deacetilasa 3	5q31.3
114	HDAC4	KIAA0288	Histona deacetilasa 4	2q37.3
115	HDAC5	KIAA0600	Histona deacetilasa 5	17q21.31
116	HDAC6	KIAA0901, JM21	Histona deacetilasa 6	Xp11.23
117	HDAC7	HDAC7A	Histona deacetilasa 7	12q13.11
118	HDAC8	HDAC1, CDA07	Histona deacetilasa 8	Xq13.1
119	HDAC9	HDAC7, HDAC7B, HDRP, KIAA0744, MITR	Histona deacetilasa 9	7p21.1
120	SIRT1	SIR2L1	NAD-dependiente de deacetilasa sirtuin-1	10q21.3
121	SIRT2	SIR2L, SIR2L2	NAD-dependiente de deacetilasa sirtuin-2	19q13.2

122	SIRT3	SIR2L3	NAD-dependiente de deacetilasa sirtuin-3, mitocondrial	11p15.5
123	SIRT6	SIR2L6	NAD-dependiente de deacetilasa sirtuin-6	19p13.3
Desfosforilación de serina / treonina / tirosina				
124	DUSP1	CL100, MKP1, PTPN10, VH1	Proteína fosfatasa dual específica 1	5q35.1
125	EYA1	-	Ojos ausentes homologo 1	8q13.3
126	EYA2	EAB1	Ojos ausentes homologo 2	20q13.12
127	EYA3	-	Ojos ausentes homologo 3	1p35.3
128	PPM1D	WIP1	Proteína fosfatasa 1D	17q23.2
129	PPP1CC	-	Subunidad catalítica serina / treonina-proteína fosfatasa PP1-gamma	12q24.11
Desfosforilación de serina / treonina / tirosina				
130	PPP2CA	-	Serina / treonina-proteína fosfatasa 2A subunidad catalítica isoforma alfa	5q31.1
131	PPP2CB	-	Serina / treonina-proteína fosfatasa 2A subunidad catalítica isoforma beta	8p12
132	PPP4C	PPP4, PPX	Subunidad catalítica de serina / treonina-proteína fosfatasa 4	16p11.2
133	PPP5C	PPP5	Serina / treonina-proteína fosfatasa 5	19q13.32
Desmetilación de arginina				
134	JMJD6	PTDSR, KIAA0585	Arginina desmetilasa bifuncional y lisil-hidroxilasa JMJD6	17q25.1
Desmetilación de lisina				
135	KDM1A	AOF2, KDM1, KIAA0601, LSD1	Desmetilasa 1A específica de lisina	1p36.12
136	KDM1B	AOF1, C6orf193, LSD2	Desmetilasa 1B específica de lisina	6p22.3
137	KDM2A	CXXC8, FBL7, FBXL11, JHDM1A, KIAA1004	Desmetilasa 2A específica de lisina	11q13.2
138	KDM2B	NDY1, CXXC2, FBL10, FBXL10, JHDM1B, PCCX2	Desmetilasa 2B específica de lisina	12q24.31
139	KDM3A	JHDM2A, JMJD1, JMJD1A, KIAA0742, TSGA	Desmetilasa 3A específica de lisina	2p11.2
140	KDM3B	C5orf7, JHDM2B, JMJD1B, KIAA1082	Desmetilasa 3B específica de lisina	5q31.2
141	KDM4A*	JHDM3A, JMJD2, JMJD2A, KIAA0677	Desmetilasa 4A específica de lisina	1p34.2-p34.1
142	KDM4B*	JHDM3B, JMJD2B, KIAA0876	Desmetilasa 4B específica de lisina	19p13.3
143	KDM4C*	JMJD2C, GASC1, JHDM3C, KIAA0780	Desmetilasa 4C específica de lisina	9p24.1
144	KDM4D	JHDM3D, JMJD2D	Desmetilasa 4D específica de lisina	11q21
145	KDM5A*	RBP2, JARID1A, RBBP2	Desmetilasa 5A específica de lisina	12p13.33

146	KDM5B*	PLU-1, JARID1B, PLU1, RBBP2H1	Desmetilasa 5B específica de lisina	1q32.1
147	KDM5C*	DXS1272E, JARID1C, SMCX, XE169	Desmetilasa 5C específica de lisina	Xp11.22
148	KDM5D	HY, HYA, JARID1D, KIAA0234, SMCY	Desmetilasa 5D específica de lisina	Yq11.223
149	KDM6A	UTX	Desmetilasa 6A específica de lisina	Xp11.3
150	KDM6B	JMJD3, KIAA0346	Desmetilasa 6B específica de lisina	17p13.1
151	KDM7A*	KIAA1718, JHDM1D	Desmetilasa 7A específica de lisina	7q34
152	KDM8	JMJD5	Desmetilasa específica de lisina 8	16p12.1
Desmetilación de lisina				
153	PHF8*	KIAA1111, ZNF422	Histona lisina desmetilasa PHF8	Xp11.22
154	RIOX1	C14orf169, MAPJD, NO66	Oxigenasa ribosomal 1	14q24.3
Desribosilación de lisina				
155	PARG		Poli (ADP-ribosa) glicohidrolasa	10q11.23
Desubiquitinación de lisina				
156	BRCC3	BRCC36, C6.1A, CXorf53	Desubiquitinasa BRCC36 específica de lisina-63	Xq28
157	MYSM1	KIAA1915	Histona H2A desubiquitinasa MYSM1	1p32.1
158	USP16	MSTP039	Ubiquitina carboxilo terminal hidrolasa 16	21q21.3
159	USP22	KIAA1063, USP3L	Ubiquitina carboxilo terminal hidrolasa 22	17p11.2
160	USP3	-	Ubiquitina carboxilo terminal hidrolasa 3	15q22.31

* Genes que además cumplen la función de lectores epigenéticos

ANEXO 2

Tabla 2A. Análisis de enriquecimiento GO del conjunto de 160 genes utilizando Enrichr

Proceso Biológico	P-valor	Genes	Número de genes
Modificación de histonas (GO: 0016570)	1.67302E-67	HDAC4; HDAC5; HDAC2; USP16; HDAC10; HDAC1; UHRF1; EHMT2; HDAC11; EHMT1; HDAC9; AURKB; HDAC6; NSD3; RPS6KA4; RPS6KA5; KAT5; HLCS; NSD1; EP300; KAT8; PRMT6; KDM4A; CREBBP; PRMT8; PRMT7; EYA1; PRMT2; EYA2; PRMT1; EYA3; USP22; USP3; PRKCD; HUWE1; BAZ1B; SIRT1; SIRT2; KAT2A; KAT6B; CARM1; KAT6A; PADI4; EZH2.	44
Desmetilación de lisina de histona (GO: 0070076)	4.78742E-35	KDM5A; KDM6B; KDM4A; KDM5B; KDM3A; KDM4B; KDM5C; KDM2A; KDM3B; KDM4C; KDM5D; KDM1A; KDM4D; KDM1B; KDM8; PHF8; RIOX1; KDM7A.	18
Regulación de la transcripción a nivel del ADN molde (GO: 0006355).	5.27807E-31	KDM5A;KDM5B;SUV39H2;KDM5C;HDAC10;KDM1A;SUV39H1;KDM8;EHMT1;PRDM2;PHF8;RPS6KA4;DNMT3L;RPS6KA5;KAT5;EP300;KAT8;KAT7;KMT5A;NCOA1;PRMT6;PRMT5;PRMT8;PRMT7;CHUK;DTX3L;PRMT2;RING1;PRKCB;PRMT1;NCOA3;USP3;DAPK3;SIRT1;SIRT2;KAT2A;KAT6B;KAT6A;CLOCK;HDAC4;KMT2E;HDAC5;KMT2D;HDAC2;KDM3A;HDAC3;KDM3B;USP16;KMT2A;HDAC1;HDAC9;HDAC6;PPP2CA;NSD1;RNF20;KDM4A;CREBBP;KDM4C;USP22;SETDB2;ELP3;CDK5;RIOX1;CARM1;PKN1;EZH2.	66
Regulación negativa de la transcripción a nivel del ADN molde (GO: 0045892)	1.07801E-22	HDAC4; HDAC5; SUV39H2; HDAC2; KDM5C; HDAC3; HDAC10; KDM1A; HDAC1; UHRF1; EHMT2; SUV39H1; EHMT1; HDAC8; HDAC9; AURKB; HDAC6; DNMT3L; RPS6KA5; KAT5; NSD1; DNMT3B; SMYD2; EP300; KAT8; KMT5A; PRMT6; KDM4A; CREBBP; KDM2B; PARP1; PRMT2; RING1; SETDB2; SIRT1; SIRT2; CDK5; RIOX1; KAT6B; KAT6A; CLOCK; EZH2.	42
Regulación negativa del proceso biosintético de macromoléculas celulares (GO: 2000113)	9.35421E-22	HDAC4; GSK3B; HDAC5; SUV39H2; HDAC2; KDM5C; HDAC3; HDAC10; KDM1A; HDAC1; SUV39H1; EHMT1; HDAC9; HDAC6; DNMT3L; RPS6KA5; KAT5; KAT8; KMT5A; PRMT6; KDM4A; PRMT2; RING1; DAPK3; SETDB2; SIRT1; SIRT2; CDK5; RIOX1; KAT6B; KAT6A; CLOCK; ATR; EZH2.	34
Regulación positiva de la transcripción a nivel del ADN molde (GO: 0045893)	1.08007E-21	HDAC4; KDM5A; KMT2E; HDAC5; KMT2D; HDAC2; KDM3A; HDAC3; USP16; KDM1A; KMT2A; HDAC1; PRKDC; UHRF1; DOT1L; KDM8; KMT2B; PRDM2; MYSM1; PHF8; RPS6KA4; RPS6KA5; KAT5; NSD1; EP300; KAT8; KAT7; KDM6B; NCOA1; RNF20; CREBBP; CHUK; DTX3L; PARP1; PRMT2; NCOA3; USP22; TET2; TET1; SIRT1; SIRT2; KAT2B; KAT2A; KAT6B; KAT6A; TET3; CLOCK.	47
Regulación negativa de la expresión génica (GO: 0010629)	3.95382E-21	HDAC4; HDAC5; SUV39H2; DNMT1; HDAC2; KDM5C; HDAC3; HDAC10; KDM1A; HDAC1; SUV39H1; EHMT1; PPM1D; HDAC9; HDAC6; DNMT3L; RPS6KA5; KAT5; DNMT3B; KAT8; KMT5A; PRMT6; KDM4A; PRMT2; RING1; DAPK3; DNMT3A; SETDB2; SIRT1; SIRT2; CDK5; RIOX1; KAT6B; KAT6A; CLOCK; EZH2.	36
Metilación de histonas (GO: 0016571)	5.38707E-21	NSD3; PRMT6; PRMT5; PRMT8; PRMT7; PRMT2; PRMT1; CARM1; EHMT2; NSD1; EHMT1; EZH2.	12

Regulación negativa de la transcripción con molde de ácido nucleico (GO: 1903507)	1.72108E-20	HDAC4; HDAC5; SUV39H2; HDAC2; KDM5C; HDAC3; HDAC10; KDM1A; HDAC1; SUV39H1; EHMT1; HDAC9; HDAC6; DNMT3L; RPS6KA5; KAT5; KAT8; KMT5A; PRMT6; KDM4A; PRMT2; RING1; SETDB2; SIRT1; SIRT2; CDK5; RIOX1; KAT6B; KAT6A; CLOCK; EZH2.	31
Regulación de la transcripción por el promotor del ARN polimerasa II (GO: 0006357)	6.89936E-20	KDM5B; SUV39H2; HDAC10; KDM1A; EHMT2; SUV39H1; PRDM2; MYSM1; RPS6KA4; RPS6KA5; KAT5; DNMT3B; EP300; KAT7; KMT5A; KDM6B; NCOA1; CHUK; KDM2B; PARP1; PRKCB; NCOA3; TET2; TET1; SIRT1; SIRT2; KAT2B; KAT2A; TET3; CLOCK; HDAC4; HDAC5; KMT2D; HDAC2; HDAC3; USP16; KMT2A; HDAC1; PRKDC; UHRF1; DOT1L; HDAC8; HDAC9; AURKB; NSD1; SMYD2; CREBBP; KDM4C; ELP3; PKN1; EZH2.	51
Dimetilación de peptidil-lisina (GO: 0018027)	5.80144E-18	SUV39H2; SETD2; EHMT2; SUV39H1; SETD7; SMYD2; EHMT1; SETMAR; ASH1L.	9
Desacetilación de proteínas (GO: 0006476)	1.38793E-17	HDAC4; HDAC5; HDAC2; HDAC3; HDAC10; HDAC1; HDAC11; HDAC9; SIRT1; HDAC6; SIRT2; SIRT3.	12
Regulación de la expresión génica, epigenética (GO: 0040029).	2.72011E-17	HDAC4; HDAC5; DNMT1; PRMT7; KDM1B; DNMT3A; BAZ1B; PPM1D; HDAC6; KAT2B; DNMT3L; KAT2A; DNMT3B; EP300; EZH2.	15
Desmetilación de histona H3-K9 (GO: 0033169)	7.43968E-17	KDM4A; KDM3A; KDM4B; KDM3B; KDM4C; KDM1A; KDM4D; KDM7A; PHF8.	9
Desmetilación de histona H3-K36 (GO: 0070544)	1.25721E-16	KDM4A; KDM4B; KDM2A; KDM4C; RIOX1; KDM8; KDM7A; PHF8.	8
Metilación de lisina de la histona (GO: 0034968)	8.45373E-16	KMT2D; KMT2A; EHMT2; SETD1B; DOT1L; KMT2B; SETDB2; SETMAR; EZH2; KDM6A.	10
Regulación de la expresión génica (GO: 0010468).	1.58869E-15	HDAC4; GSK3B; HDAC5; KMT2D; DNMT1; KDM3A; SETD2; KDM3B; HDAC10; HDAC1; PRDM2; HDAC6; PPP2CA; RPS6KA4; RPS6KA5; DNMT3B; EP300; KAT7; RAG1; PRMT6; RNF20; KDM4A; PRMT5; CREBBP; PRMT8; PRMT7; PRMT2; PRMT1; NCOA3; USP3; DAPK3; SIRT1; KAT6B; CARM1; CDK3; ATM; CLOCK; EZH2.	38
Metilación de proteínas (GO: 0006479)	1.85675E-15	NSD3; PRMT6; PRMT5; PRMT8; PRMT7; PRMT2; PRMT1; CARM1; EHMT2; NSD1; EHMT1; EZH2.	12
Desacilación de proteínas (GO: 0035601)	2.46142E-15	HDAC4; HDAC5; HDAC3; HDAC10; HDAC1; SIRT1; HDAC6; SIRT2; SIRT3.	9
Regulación de la diferenciación de las células mieloides (GO: 0045637)	4.75658E-15	KMT2E; PRMT6; KMT2D; CREBBP; KMT2A; HDAC1; PRKCB; PRMT1; KMT2C; KMT2B; SETD1A; KAT2B; EP300.	13

Tabla 3A. Mapas de vías de KEGG enriquecidas para los 160 genes de estudio utilizando Enrichr

Vías KEGG	P-valor	Genes	Número de genes
Degradación de lisina_Homo sapiens_hsa00310	2.92E-27	KMT2E; SUV39H2; KMT2D; SETD2; SETDB1; KMT2A; SETD1B; SUV39H1; KMT2C; EHMT2; SETD1A; SETDB2; DOT1L; KMT2B; SETD7; SETMAR; EHMT1; ASH1L; NSD1.	19
Carcinogénesis viral_Homo sapiens_hsa05203	1.05E-10	HDAC4; HDAC5; CREBBP; HDAC2; HDAC3; HDAC10; HDAC1; HDAC11; HDAC8; HDAC9; HDAC6; HDAC7; KAT2B; KAT2A; EP300.	15
Alcoholismo_Homo sapiens_hsa05034	2.13E-09	HDAC4; HDAC5; HDAC2; HDAC3; HDAC10; HDAC1; HDAC11; HDAC8; HDAC9; HDAC6; HDAC7; PPP1CC; HAT1.	13
Vía de señalización de la hormona tiroidea_Homo sapiens_hsa04919	2.78E-09	NCOA1; KAT2B; GSK3B; CREBBP; KAT2A; HDAC2; HDAC3; PRKCB; HDAC1; NCOA3; EP300.	11
Ciclo celular_Homo sapiens_hsa04110	6.04E-08	GSK3B; CREBBP; HDAC2; PRKDC; HDAC1; EP300; ATM; BUB1; RBX1; ATR.	10
Vía de señalización de Notch_Homo sapiens_hsa04330	1.03E-07	KAT2B; CREBBP; KAT2A; HDAC2; DTX3L; HDAC1; EP300.	7
Vía de señalización de FoxO_Homo sapiens_hsa04068	1.25E-06	CREBBP; CHUK; PRMT1; SETD7; EP300; ATM; STK4; SIRT1; RAG1.	9
MicroRNAs en cáncer_Homo sapiens_hsa05206	5.02E-06	HDAC4; CREBBP; DNMT1; RPS6KA5; PRKCB; HDAC1; DNMT3A; DNMT3B; EP300; ATM; SIRT1; EZH2.	12
Vía de señalización MAPK_Homo sapiens_hsa04010	0.000218885	RPS6KA4; PPP5C; RPS6KA5; CHUK; PRKCB; DUSP1; MAP3K8; STK4; PAK2.	9
Infección por el virus de Epstein-Barr_Homo sapiens_hsa05169	0.000226848	HDAC4; GSK3B; HDAC5; CREBBP; HDAC2; CHUK; HDAC1; EP300.	8
Infección por HTLV-I_Homo sapiens_hsa05166	0.000238699	KAT2B; GSK3B; CREBBP; KAT2A; KAT5; CHUK; EP300; ATM; ATR.	9
Vía de señalización TGF-beta_Homo sapiens_hsa04350	0.000571193	PPP2CA; PPP2CB; CREBBP; EP300; RBX1.	5
Sinapsis dopaminérgica_Homo sapiens_hsa04728	0.000604013	PPP2CA; PPP1CC; GSK3B; PPP2CB; PRKCB; CLOCK.	6
Mala regulación transcripcional en cáncer_Homo sapiens_hsa05202	0.000624893	HDAC2; KMT2A; EYA1; HDAC1; DOT1L; ATM; KDM6A.	7
Vía de señalización del glucagón_Homo sapiens_hsa04922	0.001314595	CREBBP; PPP4C; PRMT1; EP300; SIRT1.	5
Vías en el cáncer_Homo sapiens_hsa05200	0.001380325	GSK3B; CREBBP; HDAC2; CHUK; PRKCB; HDAC1; DAPK3; EP300; STK4; RBX1.	10
Resistencia a la insulina_Homo sapiens_hsa04931	0.001842483	MGEA5; PPP1CC; GSK3B; PRKCB; PRKCD.	5
Potenciación a largo plazo_Homo sapiens_hsa04720	0.001933962	PPP1CC; CREBBP; PRKCB; EP300.	4
Carcinoma de células renales_Homo sapiens_hsa05211	0.001933962	CREBBP; EP300; PAK2; RBX1.	4

Adicción a las anfetaminas_Homo sapiens_hsa05031	0.002044102	PPP1CC; PRKCB; HDAC1; SIRT1.	4
Meiosis de ovocitos_Homo sapiens_hsa04114	0.003117042	PPP2CA; PPP1CC; PPP2CB; BUB1; RBX1.	5
Infección por herpes simple_Homo sapiens_hsa05168	0.003782726	PPP1CC; CREBBP; CHUK; EP300; JAK2; CLOCK.	6
Uniones estrechas_Homo sapiens_hsa04530	0.005240007	PPP2CA; PPP2CB; PRKCB; PRKCD; ASH1L.	5
Metabolismo de cisteína y metionina_Homo sapiens_hsa00270	0.005574575	DNMT1; DNMT3A; DNMT3B.	3
Cáncer de próstata_Homo sapiens_hsa05215	0.005687212	GSK3B; CREBBP; CHUK; EP300.	4
Vía de señalización Wnt_Homo sapiens_hsa04310	0.005729495	GSK3B; CREBBP; PRKCB; EP300; RBX1.	5
Vía de señalización de NF-kappa B_Homo sapiens_hsa04064	0.006636087	CHUK; PARP1; PRKCB; ATM.	4
Melanogénesis_Homo sapiens_hsa04916	0.008538409	GSK3B; CREBBP; PRKCB; EP300.	4
Vía de señalización HIF-1_Homo sapiens_hsa04066	0.009452271	CREBBP; PRKCB; EP300; RBX1.	4
Vía de señalización del receptor de células T_Homo sapiens_hsa04660	0.00977047	GSK3B; CHUK; MAP3K8; PAK2.	4

ANEXO 3

Tabla 4A. Genes diferencialmente expresados (DEGs) y Ontología de genes (GO) en UCS

Símbolo del Gen	Log2 (FC) ^a	Ontología del gen ^b		
		Proceso biológico	Función molecular	Componente celular
Metilación del ADN				
DNMT3B	1.555	Metilación del ADN *	ADN (citosina-5 -) - actividad metiltransferasa *	Núcleo *
Modificación de Histonas				
AURKB	4.842	Regulación de la segregación cromosómica, proliferación celular, ciclo celular	Unión al ADN, actividad transferasa, unión a ATP	Núcleo, citoplasma, citosol
UHRF1	3.771	Proliferación celular, ciclo celular	Actividad transferasa	Núcleo
BUB1	3.607	Regulación de la segregación cromosómica, proliferación celular	Unión al ADN, actividad transferasa, unión a ATP	Núcleo, citoplasma
EYA1	2.653	Modificación de la cromatina, desfosforilación de histonas	Unión de nucleótidos	Núcleo
EZH2	1.969	Modificación de la cromatina	Unión de iones metálicos, actividad transferasa	Núcleo
SMYD3	1.744	Modificación de la cromatina	Unión de iones metálicos	Núcleo
CDK5	1.663	Proliferación celular, ciclo celular	Unión al ADN, actividad transferasa, unión a ATP, unión a proteínas	Núcleo, proyección de células neuronales
PPP1CA	1.624	Ciclo celular	Actividad transferasa	Núcleo, proyección de células neuronales, citoesqueleto
EYA2	1.624	Modificación de la cromatina, desfosforilación de histonas	Actividad transferasa	Núcleo
EZH1	-3.206	Regulación de la transcripción a nivel del ADN molde, proliferación celular	Actividad metiltransferasa, actividad transferasa	Núcleo
MAP3K20	-2.474	Fosforilación de proteínas, respuesta a la radiación	Unión a proteínas	Núcleo
KAT2B	-2.421	Respuesta a la radiación	Unión a proteínas, actividad de histona acetiltransferasa	Núcleo, nucleoplasma, complejo de remodelación de cromatina, complejo activador de la transcripción Ada2 / Gcn5 / Ada3
SETD1B	-2.374	Regulación de la transcripción a nivel del ADN molde, transducción de señales, proliferación celular	Actividad metiltransferasa, unión a proteínas, actividad transferasa, actividad de histona metiltransferasa (H3-K4 específica)	Núcleo, complejo de histona metiltransferasa
DUSP1	-2.339	Fosforilación de proteínas, respuesta a la radiación	Unión a proteínas	Núcleo, nucleoplasma
JAK2	-2.21	Regulación de la transcripción a nivel del ADN molde, fosforilación de proteínas	Unión a proteínas	Núcleo
MYSM1	-2.206	Regulación de la transcripción a nivel del ADN molde	Actividad coactivadora de transcripción	Núcleo, complejo de remodelación de cromatina
CDK17	-1.992	Regulación positiva de la unión del ADN	Actividad de proteína serina / treonina quinasa	Núcleo, citoplasma#
SETD7	-1.943	Regulación de la transcripción a nivel del ADN molde, proliferación celular	Actividad metiltransferasa, unión a proteínas, actividad transferasa	Núcleo
MAP3K8	-1.87	Respuesta a la radiación	Unión a proteínas	Citoplasma
NEK9	-1.787	Respuesta a la radiación	Unión a proteínas	Núcleo
HDAC6	-1.742	Regulación de la transcripción a nivel del ADN molde, fosforilación de proteínas	Unión a proteínas	Núcleo
CREBBP	-1.741	Proliferación celular	Unión a proteínas, actividad transferasa, actividad de histona acetiltransferasa	Núcleo, nucleoplasma
KMT2A	-1.733	Regulación de la transcripción a nivel del ADN molde, transducción de señales, proliferación celular	Actividad metiltransferasa, unión a proteínas, actividad transferasa, actividad de histona metiltransferasa (H3-K4 específica)	Núcleo, complejo de histona metiltransferasa
ATM	-1.719	Fosforilación de proteínas	Unión a proteínas	Núcleo, nucleoplasma

HDAC10	-1.705	Regulación de la transcripción a nivel del ADN molde, proliferación celular	Unión a proteínas	Núcleo, nucleoplasma
KMT2C	-1.67	Regulación de la transcripción a nivel del ADN molde, proliferación celular	Actividad metiltransferasa, unión a proteínas, actividad transferasa, actividad de histona metiltransferasa (H3-K4 específica)	Núcleo, complejo de histona metiltransferasa
KAT2A	-1.641	Regulación negativa de la transcripción del promotor de la ARN polimerasa II	Unión a proteínas, actividad transferasa, actividad de histona acetiltransferasa	Núcleo, complejo activador de la transcripción Ada2 / Gcn5 / Ada3
MGEA5	-1.628	Regulación positiva de la señalización de fosfatidilinositol 3-quinasa	Actividad de histona acetiltransferasa	Núcleo
SETMAR	-1.598	Regulación de la transcripción a nivel del ADN molde	Actividad metiltransferasa, unión a proteínas, actividad transferasa	Núcleo
KMT2E	-1.553	Regulación de la transcripción a nivel del ADN molde, transducción de señales, respuesta a la radiación	Actividad metiltransferasa, unión a proteínas, actividad transferasa, actividad de histona metiltransferasa (H3-K4 específica)	Núcleo
RPS6KA5	-1.534	Proliferación celular	Unión a proteínas	Núcleo, nucleoplasma
KMT2D	-1.507	Regulación de la transcripción a nivel del ADN molde, transducción de señales, proliferación celular	Actividad metiltransferasa, unión a proteínas, actividad transferasa	Núcleo, complejo de histona metiltransferasa

Se muestra 33 genes diferencialmente expresados (DEGs) entre tejidos normales (n=78) y tumores primarios (n=57).

^a $\log_2(\text{Fold Change}) > 1.5$ para los genes sobreexpresados (en rojo) y $\log_2(\text{Fold Change}) < -1.5$ para los genes subexpresados (en azul); se utilizó un valor p ajustado < 0.05 (medido por Benjamini y Hochberg FDR).

^b Análisis de enriquecimiento GO según Genecodis, se muestra los principales términos GO enriquecidos en cada categoría (p ajustado < 0.01)

* Información obtenida de Uniprot.

Tabla 5A. Genes diferencialmente expresados (DEGs) y Ontología de genes (GO) en OV

Símbolo del Gen	Log2 (FC) ^a	Ontología del gen ^b		
		Proceso biológico	Función molecular	Componente celular
Modificación de Histonas				
EYA2	5.187	Proceso apoptótico	Unión a proteínas	Núcleo, citoplasma
AURKB	3.996	Proliferación celular, regulación de la segregación cromosómica, ciclo celular, autofosforilación de proteínas.	Actividad de la proteína quinasa, unión a proteínas, actividad de proteína serina / treonina quinasa, unión a ATP, unión de nucleótidos.	Núcleo, cromosoma nuclear condensado, región centromérica, citoesqueleto, citoplasma, cromosoma condensada, región centromérica.
BUB1	3.174	Proliferación celular, regulación de la segregación cromosómica, fosforilación de proteínas	Actividad de la proteína quinasa, unión a proteínas, actividad de proteína serina / treonina quinasa, unión a ATP, unión de nucleótidos.	Núcleo, cromosoma nuclear condensado, región centromérica.
UHRF1	2.227	Proliferación celular, ciclo celular	Unión a proteínas	Núcleo
CDK5	2.103	Proliferación celular, fosforilación de proteínas, ciclo celular, autofosforilación de proteínas.	Actividad de la proteína quinasa, unión a proteínas, actividad de proteína serina / treonina quinasa, unión a ATP, unión de nucleótidos.	Núcleo, citoesqueleto, citoplasma
NEK6	2.039	Fosforilación de proteínas, ciclo celular, autofosforilación de proteínas.	Actividad de la proteína quinasa, unión a proteínas, actividad de proteína serina / treonina quinasa, unión a ATP, unión de nucleótidos.	Núcleo, citoesqueleto, citoplasma
HDAC9	1.627	Regulación de la diferenciación de las células musculares estriadas.	Unión a proteínas	Núcleo, citoplasma
PRKCD	1.613	Fosforilación de proteínas	Actividad de la proteína quinasa, unión a proteínas, actividad de proteína serina / treonina quinasa, unión a ATP, unión de nucleótidos.	Núcleo, citoplasma
PPP4C	1.573	Regulación de la reparación de una rotura de doble cadena en el ADN mediante recombinación homóloga	Unión a proteínas	Núcleo, citoesqueleto, citoplasma
EZH1	-3.091	Fosforilación de proteínas	Actividad metiltransferasa	Núcleo
DUSP1	-2.746	Respuesta a la radiación, interacción interespecies entre organismos.	Unión a proteínas	Núcleo
KAT2A	-2.672	Regulación positiva de la transcripción del promotor del ARN polimerasa II	Unión de histona deacetilasa	Núcleo
ATM	-2.399	Respuesta a la radiación	Unión a proteínas	Núcleo
MGEA5	-2.231	Regulación negativa de la transcripción a nivel del ADN molde.	Actividad de histona acetiltransferasa	Núcleo
NEK9	-2.22	Interacción interespecies entre organismos.	Unión a proteínas	Núcleo
JAK2	-1.968	Fosforilación de proteínas, respuesta a la radiación	Unión a proteínas	Núcleo
MAP3K20	-1.951	Respuesta a la radiación, interacción interespecies entre organismos	Unión a proteínas	Núcleo
MYSM1	-1.92	Fosforilación de proteínas	Actividad coactivadora de transcripción.	Núcleo
HDAC6	-1.861	Fosforilación de proteínas, respuesta a la radiación, proliferación celular	Actividad de histona deacetilasa (específica para H3-K16), actividad de histona deacetilasa dependiente de NAD (específica de H3-K14), actividad de histona deacetilasa dependiente de NAD (específica de H4-K16), actividad de histona deacetilasa, actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	Núcleo
KAT2B	-1.82	Interacción interespecies entre organismos, regulación positiva de la transcripción del promotor del ARN polimerasa II	Unión de histona deacetilasa	Núcleo

HDAC5	-1.713	Fosforilación de proteínas, proliferación celular, regulación positiva de la transcripción del promotor del ARN polimerasa II	Actividad de histona deacetilasa (específica para H3-K16), actividad de histona deacetilasa dependiente de NAD (específica de H3-K14), actividad de histona deacetilasa dependiente de NAD (específica de H4-K16), actividad de histona deacetilasa, actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	Núcleo
KDM6A	-1.651	Fosforilación de proteínas	Unión a la cromatina *	Núcleo
CDK3	-1.623	Vía de señalización del receptor del factor de crecimiento derivado de plaquetas	Unión a ATP	Núcleo *
HDAC10	-1.618	Fosforilación de proteínas, proliferación celular	Actividad de histona deacetilasa (específica para H3-K16), actividad de histona deacetilasa dependiente de NAD (específica de H3-K14), actividad de histona deacetilasa dependiente de NAD (específica de H4-K16), actividad de histona deacetilasa, actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	Núcleo
SETDB2	-1.573	Fosforilación de proteínas, interacción interespecies entre organismos.	Actividad metiltransferasa	Núcleo
SETD7	-1.554	Fosforilación de proteínas	Unión a proteínas	Núcleo

Se muestra 26 genes diferencialmente expresados (DEGs) entre tejidos normales (n=88) y tumores primarios (n=426).

^a $\log_2(\text{Fold Change}) > 1.5$ para los genes sobreexpresados (en rojo) y $\log_2(\text{Fold Change}) < -1.5$ para los genes subexpresados (en azul); se utilizó un valor p ajustado < 0.05 (medido por Benjamini y Hochberg FDR).

^b Análisis de enriquecimiento GO según Genecodis, muestra los principales términos GO enriquecidos en cada categoría (p ajustado < 0.01)

* Información obtenida de Uniprot.

Tabla 6A. Genes diferencialmente expresados (DEGs) y Ontología de genes (GO) en CESC

Símbolo del Gen	Log2 (FC) ^a	Ontología del gen (GO) ^b		
		Proceso biológico	Función molecular	Componente celular
Modificación de Histonas				
AURKB	5.341	Regulación negativa de la transcripción del promotor de la ARN polimerasa II, desarrollo embrionario, proliferación celular, regulación de la transcripción a nivel del ADN molde.	Unión a proteínas, actividad de proteína serina / treonina quinasa, actividad de la proteína quinasa.	Cromosoma condensado, región centromérica, huso, núcleo, cromosoma, cromosoma nuclear condensado, región centromérica.
UHRF1	4.678	Desarrollo embrionario, proliferación celular.	Unión a proteínas	Núcleo
BUB1	3.404	Regulación negativa de la transcripción del promotor de la ARN polimerasa II, proliferación celular, regulación de la transcripción a nivel del ADN molde.	Unión a proteínas, actividad de proteína serina / treonina quinasa, actividad de la proteína quinasa.	Núcleo, cromosoma, cromosoma nuclear condensado, región centromérica.
EZH2	2.786	Fosforilación de proteínas	Unión a proteínas, actividad metiltransferasa.	Núcleo
HASPIN	1.788	Desarrollo embrionario, fosforilación de proteínas.	Actividad de proteína serina / treonina quinasa, actividad de la proteína quinasa.	Huso, núcleo, cromosoma
SMYD2	1.532	Fosforilación de proteínas	Unión a la ARN polimerasa II, unión a proteínas, actividad metiltransferasa.	Núcleo
DUSP1	-2.201	Ciclo celular, regulación positiva del proceso apoptótico.	Actividad de la proteína tirosina / treonina fosfatasa.	Núcleo
EZH1	-2.198	Modificación de la cromatina	Actividad de histona metiltransferasa (H3-K27 específica), actividad metiltransferasa.	Complejo ESC / E (Z), núcleo
KAT2B	-1.709	Detención del ciclo celular, ciclo celular	Actividad de histona acetiltransferasa.	Banda I, actomiosina, núcleo
MAP3K20	-1.701	Detención del ciclo celular, ciclo celular, regulación positiva del proceso apoptótico, punto de control de daño en el ADN.	Actividad de proteína serina / treonina quinasa.	Núcleo
NEK9	-1.665	Ciclo celular	Actividad de proteína serina / treonina quinasa.	Núcleo
ATM	-1.604	Detención del ciclo celular, regulación positiva del proceso apoptótico, punto de control de daño en el ADN.	Actividad de histonas serina quinasa, actividad de proteína serina / treonina quinasa.	Núcleo
EYA1	-1.58	Modificación de la cromatina	Actividad de la proteína tirosina fosfatasa	Núcleo
KDM4B	-1.546	Modificación de la cromatina	Regulación de la transcripción a nivel del ADN molde.	Núcleo
KMT2A	-1.524	Modificación de la cromatina	Actividad metiltransferasa	Parte nuclear, núcleo
MGEA5	-1.514	Regulación negativa de la adaptación del músculo cardíaco.	Actividad de histona acetiltransferasa	Núcleo

Se muestra 16 genes diferencialmente expresados (DEGs) entre tejidos normales (n=13) y tumores primarios (n=306).

^a $\log_2$ (Fold Change) > 1.5 para los genes sobreexpresados (en rojo) y $\log_2$ (Fold Change) < -1.5 para los genes subexpresados (en azul); se utilizó un valor p ajustado <0.05 (medido por Benjamini y Hochberg FDR).

^b Análisis de enriquecimiento GO según Genecodis, muestra los principales términos GO enriquecidos en cada categoría (p ajustado < 0.01)

Tabla 7A. Genes diferencialmente expresados (DEGs) y Ontología de genes (GO) en UCEC

Símbolo del Gen	Log2 (FC) ^a	Ontología del gen ^b		
		Proceso biológico	Función molecular	Componente celular
Modificación de Histonas				
AURKB	4.597	Regulación de la segregación cromosómica, proliferación celular, ciclo celular, autofosforilación de proteínas	Unión a proteínas, actividad de la proteína quinasa, actividad de proteína serina / treonina quinasa	Núcleo, cromosoma nuclear condensado, región centromérica, cromosoma condensado, región centromérica, complejo pasajero de cromosomas, citosol.
EYA2	3.432	Modificación de la cromatina	Unión a proteínas	Núcleo
UHRF1	3.317	Proliferación celular, ciclo celular	Unión a proteínas	Núcleo
BUB1	3.269	Regulación de la segregación cromosómica, proliferación celular	Unión a proteínas, actividad de la proteína quinasa, actividad de proteína serina / treonina quinasa.	Núcleo, cromosoma nuclear condensado, región centromérica, citosol
EZH2	1.842	Modificación de la cromatina	Unión a proteínas	Núcleo
SMYD2	1.671	Modificación de la cromatina	Unión a proteínas, unión a p53.	Núcleo, citosol
NEK6	1.563	Ciclo celular, autofosforilación de proteínas.	Unión a proteínas, actividad de la proteína quinasa, actividad de proteína serina / treonina quinasa.	Núcleo
CDK5	1.526	Proliferación celular, ciclo celular, autofosforilación de proteínas.	Unión a proteínas, actividad de la proteína quinasa, actividad de proteína serina / treonina quinasa, unión a p53, activación del receptor de acetilcolina.	Núcleo, citosol
DUSP1	-3.006	Regulación positiva del proceso apoptótico.	Unión a proteínas	Núcleo
EZH1	-2.923	Modificación de la cromatina	Actividad metiltransferasa, actividad transferasa	Núcleo
SETD7	-2.42	Modificación de la cromatina	Actividad metiltransferasa, actividad transferasa, unión a proteínas.	Núcleo
MAP3K20	-2.39	Regulación positiva del proceso apoptótico.	Unión a proteínas	Núcleo
SETD1B	-2.114	Modificación de la cromatina, metilación de lisina K4 en la histona H3.	Actividad metiltransferasa, actividad transferasa, unión a proteínas	Núcleo, citoplasma
CDK17	-2.1	Fosforilación de proteínas	Actividad de proteína serina / treonina quinasa.	Núcleo, citoplasma *
JAK2	-2.02	Modificación de la cromatina, regulación positiva del proceso apoptótico.	Unión a proteínas	Núcleo
NEK9	-1.956	Ciclo celular	Unión a proteínas	Núcleo
ATM	-1.912	Regulación positiva del proceso apoptótico.	Unión a proteínas	Núcleo
KAT2B	-1.882	Remodelación de la cromatina	Actividad de histona acetiltransferasa, unión a proteínas.	Núcleo, complejo de represión transcripcional, nucleoplasma.
KMT2A	-1.827	Modificación de la cromatina, metilación de lisina K4 en la histona H3	Actividad metiltransferasa, actividad transferasa, unión a proteínas.	Núcleo, citoplasma
CREBBP	-1.779	Acetilación de peptidil-lisina N-terminal	Actividad de histona acetiltransferasa, actividad transferasa, unión a proteínas	Núcleo

MYSM1	-1.676	Modificación de la cromatina	Actividad coactivadora de la transcripción	Núcleo, nucleoplasma
HDAC5	-1.671	Modificación de la cromatina, remodelación de la cromatina, desacetilación de histonas.	Unión a proteínas, actividad de histona deacetilasa (específica para H3-K16).	Núcleo, nucleolo
HDAC4	-1.655	Modificación de la cromatina, remodelación de la cromatina, desacetilación de histonas.	Unión a proteínas, actividad de histona deacetilasa (específica para H3-K16).	Núcleo, nucleolo, complejo de represión transcripcional.
HDAC6	-1.641	Modificación de la cromatina, regulación positiva del proceso apoptótico, desacetilación de histonas.	Unión a proteínas, actividad de histona deacetilasa (específica para H3-K16).	Núcleo, nucleolo
MGEA5	-1.637	Regulación positiva de la secreción de insulina.	Actividad de histona acetiltransferasa	Núcleo
PRDM2	-1.622	Regulación de la transcripción a nivel del ADN molde.	Actividad metiltransferasa, actividad transferasa.	Núcleo
KAT2A	-1.549	Remodelación de la cromatina	Actividad de histona acetiltransferasa, actividad transferasa, unión a proteínas.	Núcleo
KMT2D	-1.535	Modificación de la cromatina, metilación de lisina K4 en la histona H3.	Actividad metiltransferasa, actividad transferasa, unión a proteínas.	Núcleo, citoplasma
NCOA1	-1.519	Regulación de la transcripción a nivel del ADN molde.	Actividad de histona acetiltransferasa.	Núcleo

Se muestra 29 genes diferencialmente expresados (DEGs) entre tejidos normales (n=91) y tumores primarios (n=174).

^a $\log_2(\text{Fold Change}) > 1.5$ para los genes sobreexpresados (en rojo) y $\log_2(\text{Fold Change}) < -1.5$ para los genes subexpresados (en azul); se utilizó un valor p ajustado < 0.05 (medido por Benjamini y Hochberg FDR).

^b Análisis de enriquecimiento GO según Genecodis, muestra los principales términos GO enriquecidos en cada categoría (p ajustado < 0.01)

* Información obtenida de Uniprot.

ANEXO 4

Tabla 8A. Análisis de enriquecimiento por GO de genes sobreexpresados en CESC utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes sobreexpresados
Proceso Biológico (PB)				
GO:0006355	Regulación de la transcripción a nivel del ADN molde	2	1.20469E-05	BUB1, AURKB
GO:0006468	Fosforilación de proteínas	3	0.000128331	SMYD2, HASPIN, EZH2
GO:0008283	Proliferación celular	3	0.00023072	BUB1, AURKB, UHRF1
GO:0009790	Desarrollo embrionario	3	0.000466442	AURKB, HASPIN, UHRF1
GO:0000122	Regulación negativa de la transcripción del promotor de la ARN polimerasa II	2	0.000854797	BUB1, AURKB
GO:0046777	Autofosforilación de proteínas	2	0.000890217	BUB1, AURKB
GO:0000910	Citocinesis	1	0.00103046	HASPIN
GO:0007049	Ciclo celular	1	0.00103046	BUB1
GO:0007243	Cascada de proteínas quinasas intracelular	1	0.00164862	UHRF1
GO:0000278	Ciclo celular mitótico	1	0.00164862	SMYD2
GO:0031145	Proceso catabólico dependientes de ubiquitina proteasomal dependientes del complejo promotor del anafase	1	0.00190211	HASPIN
GO:0007568	Envejecimiento	1	0.00190211	HASPIN
GO:0007064	Cohesión de cromátidas hermanas durante mitosis	1	0.00190211	BUB1
GO:0016568	Modificación de la cromatina	1	0.00206047	AURKB
GO:0018026	Monometilación de peptidil-lisina	1	0.00206047	SMYD2
GO:0002903	Regulación negativa de la apoptosis de células B	1	0.00206047	AURKB
GO:0006325	Organización de la cromatina	1	0.00224746	HASPIN
GO:0016571	Metilación de histonas	1	0.00224746	SMYD2
GO:0007094	Punto de control del conjunto del huso del ciclo celular mitótico	1	0.00224746	AURKB
GO:0000236	Prometáfase mitótica	1	0.00224746	AURKB
GO:0034501	Localización de las proteínas en el cinetocoro	1	0.00228924	AURKB
GO:0000087	Fase M del ciclo celular mitótico	1	0.00228924	SMYD2
GO:0007093	Punto de control del ciclo celular mitótico	2	0.00235663	BUB1, AURKB
GO:0018027	Dimetilación de peptidil-lisina	1	0.00274649	EZH2

GO:0032270	Regulación positiva de la proteína celular del proceso metabólico	3	0.0035084	SMYD2, EZH2, UHRF1
GO:0016570	Modificación de histonas	2	0.00354982	BUB1, HASPIN
GO:0071777	Regulación positiva de la citocinesis del ciclo celular	2	0.00359648	SMYD2, AURKB
GO:0008608	Acoplamiento de los microtúbulos del huso al cinetocoro	1	0.00366092	UHRF1
GO:0034644	Respuesta celular a la UV	1	0.00369221	AURKB
GO:0010452	Metilación de lisina K36 en la histona H3	1	0.0038434	BUB1
GO:0051256	Formación de la zona media del huso mitótico	1	0.00478071	EZH2
GO:0048387	Regulación negativa de la vía de señalización del receptor de ácido retinoico	1	0.00514515	EZH2
GO:0090231	Regulación del punto de control del huso	1	0.00523832	AURKB
GO:2000373	Regulación positiva de la topoisomerasa del ADN (ATP-hidrolizante)	1	0.00725844	BUB1
Función molecular (FM)				
GO:0004672	Actividad de la proteína quinasa	3	0.000260724	BUB1, AURKB, HASPIN
GO:0004674	Actividad de proteína serina / treonina quinasa	3	0.000323916	BUB1, AURKB, HASPIN
GO:0008168	Actividad metiltransferasa	2	0.00125157	SMYD2, EZH2
GO:0005515	Unión a proteínas	5	0.00126109	SMYD2, BUB1, AURKB, EZH2, UHRF1
GO:0000993	Unión a la ARN polimerasa II	1	0.00146154	SMYD2
GO:0072354	Actividad de la histona quinasa (específica para H3-T3)	1	0.00146165	HASPIN
GO:0046975	Actividad de histona metiltransferasa (H3-K36 específica)	1	0.00187898	SMYD2
GO:0008327	Unión de metil-CpG	1	0.00273979	UHRF1
GO:0016279	Actividad de proteína-lisina N-metiltransferasa	1	0.00292222	SMYD2
GO:0042054	Actividad de histona metiltransferasa	1	0.00358558	EZH2
GO:0005524	Unión a ATP	3	0.00372711	BUB1, AURKB, HASPIN
GO:0000987	Unión al ADN en la región promotora proximal	1	0.00438142	UHRF1
GO:0000166	Unión de nucleótidos	3	0.00687393	BUB1, AURKB, HASPIN
GO:0004712	Actividad de la proteína serina / treonina / tirosina quinasa	1	0.00707302	AURKB
GO:0018024	Actividad de histona-lisina N-metiltransferasa	1	0.00719226	EZH2
GO:0016740	Actividad transferasa	2	0.00723199	SMYD2, EZH2
GO:0035064	Unión de residuos de histonas metiladas	1	0.00874875	UHRF1
Componente celular (CC)				
GO:0000780	Cromosoma nuclear condensado, región centromérica	2	1.96006E-05	BUB1, AURKB
GO:0005694	Cromosoma	3	7.68079E-05	BUB1, AURKB, HASPIN
GO:0005634	Núcleo	6	9.15433E-05	SMYD2, BUB1, AURKB, HASPIN, EZH2, UHRF1
GO:0005819	Huso	2	0.000683555	AURKB, HASPIN

GO:0000779	Cromosoma condensado, región centromérica	1	0.0019874	AURKB
GO:0032133	Complejo pasajero de cromosomas	1	0.0019874	AURKB
GO:0035098	Complejo ESC / E (Z)	1	0.00425685	EZH2
GO:0000792	Heterocromatina	1	0.00595698	UHRF1
GO:0005829	Citosol	3	0.00806636	SMYD2, BUB1, AURKB

Tabla 9A. Análisis de enriquecimiento por GO de genes subexpresados en CESC utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes subexpresados
Proceso Biológico (PB)				
GO:0016568	Modificación de la cromatina	4	4.30059E-05	KMT2A, KDM4B, EZH1, EYA1
GO:0007050	Detención del ciclo celular	3	0.000242546	MAP3K20, KAT2B, ATM
GO:0007049	Ciclo celular	4	0.000300755	DUSP1, MAP3K20, KAT2B, NEK9
GO:0043065	Regulación positiva del proceso apoptótico	3	0.000477403	DUSP1, MAP3K20, ATM
GO:0000077	Punto de control de daños en el ADN	2	0.000587719	MAP3K20, ATM
GO:0010212	Respuesta a la radiación ionizante	2	0.0014216	EYA1, ATM
GO:0006302	Reparación de rotura de doble cadena	2	0.00219874	EYA1, ATM
GO:0010616	Regulación negativa de la adaptación del músculo cardíaco	1	0.00287457	MGEA5
GO:0060051	Regulación negativa de la glicosilación de proteínas	1	0.00287457	MGEA5
GO:0046060	Proceso metabólico de dATP	1	0.00287457	MGEA5
GO:0080182	Trimetilación de lisina K4 en la histona H3	1	0.00287457	KMT2A
GO:2001040	Regulación positiva de la respuesta celular al fármaco	1	0.00287457	KMT2A
GO:0009653	Estructura anatómica de la morfogénesis	2	0.00307228	EZH1, EYA1
GO:0008219	Muerte celular	2	0.00412007	MAP3K20, ATM
GO:0043243	Regulación positiva del desensamble del complejo proteico	1	0.00413829	MGEA5
GO:0051054	Regulación positiva del proceso metabólico del ADN	1	0.00413829	MGEA5
GO:0002331	Exclusión alélica de células pre-B	1	0.00413829	ATM
GO:0042159	Proceso catabólico de lipoproteínas	1	0.00413829	ATM
GO:0018076	Acetilación de peptidil-lisina N-terminal	1	0.00413829	KAT2B
GO:0016576	Desfosforilación de histonas	1	0.00413829	EYA1
GO:0006473	Acetilación de proteínas	1	0.00413829	KAT2B
GO:0031343	Regulación positiva de la muerte celular	1	0.00431129	MGEA5
GO:0018394	Acetilación de peptidil-lisina	1	0.00431129	KAT2B
GO:0051569	Regulación de la histona H3-K4 metilación	1	0.00431129	KMT2A
GO:0000075	Punto de control del ciclo celular	2	0.00445534	MAP3K20, ATM

GO:0070265	Muerte celular necrótica	1	0.00510833	MGEA5
GO:0006355	Regulación de la transcripción a nivel del ADN molde	4	0.00534203	KMT2A, KDM4B, EZH1, EYA1
GO:0018022	Metilación de peptidil-lisina	1	0.00594426	EZH1
GO:0032411	Regulación positiva de la actividad del transportador	1	0.00594426	KMT2A
GO:0006516	Proceso catabólico de glicoproteínas	1	0.00646353	MGEA5
GO:0051901	Regulación positiva de la despolarización mitocondrial	1	0.00646353	MGEA5
GO:0006975	Daño en el ADN inducido por la fosforilación de proteínas	1	0.00646353	ATM
GO:0071480	Respuesta celular a la radiación gamma	1	0.00731132	ATM
GO:0060124	Regulación positiva de la secreción de hormona de crecimiento	1	0.00810897	MGEA5
GO:0043517	Regulación positiva de p53 inducida por la respuesta al daño del ADN	1	0.00861464	ATM
GO:0090399	Senescencia replicativa	1	0.00861464	ATM
Función molecular (FM)				
GO:0004402	Actividad de histona acetiltransferasa	2	0.00354195	KAT2B, MGEA5
GO:0004674	Actividad de proteína serina / treonina quinasa	3	0.00368612	MAP3K20, NEK9, ATM
GO:0035174	Actividad de histonas serina quinasa	1	0.00477346	ATM
GO:0046976	Actividad de histona metiltransferasa (H3-K27 específica)	1	0.00477346	EZH1
GO:0008168	Actividad metiltransferasa	2	0.00519997	MLL, EZH1
GO:0016746	Actividad transferasa, transfiriendo grupos acilo	2	0.00564692	KAT2B, MGEA5
GO:0045322	Unión a CpG no metilada	1	0.00572739	MLL
GO:0004677	Actividad de proteína quinasa dependiente de ADN	1	0.00596525	ATM
GO:0008330	Actividad de la proteína tirosina / treonina fosfatasa	1	0.00596525	DUSP1
GO:0005515	Unión a proteínas	6	0.0063136	DUSP1, MAP3K20, MLL, KAT2B, NEK9, ATM
GO:0019901	Unión de proteína quinasa	2	0.00635021	KAT2B, NEK9
GO:0070577	Unión de residuos de histonas acetiladas	1	0.00636042	MLL
GO:0004726	Actividad de la proteína tirosina fosfatasa (MF) que no se extiende a la membrana	1	0.00636042	DUSP1
GO:0004415	Actividad de la enzima condroitinasa I	1	0.00660679	MGEA5
GO:0003682	Unión a la cromatina	2	0.0066093	MLL, EZH1
GO:0004725	Actividad de la proteína tirosina fosfatasa	2	0.00662227	DUSP1, EYA1
GO:0003680	Unión al ADN AT	1	0.00667932	MLL

GO:0004468	Actividad de lisina N-acetiltransferasa	1	0.00716113	KAT2B
GO:0016407	Actividad acetiltransferasa	1	0.00780187	KAT2B
GO:0004861	Actividad inhibidora de la proteína quinasa dependiente de ciclina	1	0.00780187	KAT2B
GO:0042800	Actividad de histona metiltransferasa (H3-K4 específica)	1	0.00786792	MLL
GO:0016303	Actividad 1-fosfatidilinositol-3-quinasa	1	0.00786792	ATM
GO:0017017	Actividad de la MAP quinasa tirosina / serina / treonina fosfatasa	1	0.00808348	DUSP1
Componente celular (CC)				
GO:0005634	Núcleo	10	2.05834E-07	DUSP1, MAP3K20, MLL, KDM4B, EZH1, KAT2B, NEK9, EYA1, ATM, MGEA5
GO:0042641	Actomiosina	1	0.00778827	KAT2B
GO:0044428	Parte nuclear	1	0.00778827	MLL
GO:0035098	Complejo ESC / E (Z)	1	0.00834238	EZH1
GO:0031674	Banda I	1	0.00875374	KAT2B
GO:0016585	Complejo de remodelación de cromatina	1	0.00875374	KAT2B
GO:0035097	Complejo de histona metiltransferasa	1	0.00901664	MLL
GO:0005654	Nucleoplasma	3	0.00921376	DUSP1, KAT2B, ATM

Tabla 10A. Análisis de enriquecimiento GO de genes sobreexpresados en OV utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes Sobreexpresados
Proceso Biológico (PB)				
GO:0008283	Proliferación celular	4	6.27038E-05	BUB1, AURKB, CDK5, UHRF1
GO:0051983	Regulación de la segregación cromosómica	2	0.000093488	BUB1, AURKB
GO:0006468	Fosforilación de proteínas	4	0.000110036	BUB1, PRKCD, CDK5, NEK6
GO:0007049	Ciclo celular	4	0.000117405	AURKB, CDK5, NEK6, UHRF1
GO:0046777	Autofosforilación de proteínas	3	0.000198814	AURKB, CDK5, NEK6
GO:0006915	Proceso apoptótico	4	0.000268064	BUB1, PRKCD, NEK6, EYA2
GO:0018105	Fosforilación de peptidil-serina	2	0.00175961	CDK5, NEK6
GO:0000910	Citocinesis	2	0.00178716	AURKB, NEK6
GO:0000087	Fase M del ciclo celular mitótico	2	0.00229071	BUB1, AURKB
GO:0021697	Formación de la corteza cerebelosa	1	0.00235239	CDK5
GO:2000755	Regulación positiva del proceso catabólico de la esfingomielina	1	0.00235239	PRKCD
GO:0032615	Producción de interleucina-12	1	0.00235239	PRKCD
GO:2000753	Regulación positiva del proceso catabólico glucosilceramida	1	0.00235239	PRKCD
GO:0051490	Regulación negativa del ensamblaje de filopodio	1	0.00235239	PRKCD
GO:0032613	Producción de interleucina-10	1	0.00235239	PRKCD
GO:0051153	Regulación de la diferenciación de las células musculares estriadas	1	0.00235239	HDAC9
GO:0007063	Regulación de la cohesión entre cromátidas hermanas	1	0.00235239	BUB1
GO:0014706	Desarrollo del tejido muscular estriado	1	0.00319888	EYA2
GO:2000373	Regulación positiva de la topoisomerasa del ADN (ATP-hidrolizante)	1	0.00319888	UHRF1
GO:0033136	Fosforilación de serina de la proteína STAT3	1	0.00319888	CDK5
GO:0031914	Regulación negativa de la plasticidad sináptica	1	0.00319888	CDK5
GO:2000304	Regulación positiva del proceso biosintético de ceramida	1	0.00319888	PRKCD
GO:0048742	Regulación del desarrollo de la fibra muscular esquelética	1	0.00319888	HDAC9
GO:0000236	Prometáfase mitótica	2	0.00366809	BUB1, AURKB
GO:2000273	Regulación positiva de la actividad del receptor	1	0.00371876	CDK5
GO:2000772	Regulación de la senescencia celular	1	0.00371876	NEK6
GO:0034351	Regulación negativa de la apoptosis de células gliales	1	0.00371876	PRKCD
GO:0007501	Especificación del destino de las células mesodérmicas	1	0.00371876	EYA2

GO:0016570	Modificación de histonas	1	0.00371876	AURKB
GO:0051256	Formación de la zona media del huso mitótico	1	0.00371876	AURKB
GO:0021819	Formación de capas en la corteza cerebral	1	0.00371876	CDK5
GO:0045956	Regulación positiva de la exocitosis dependiente de los iones de calcio	1	0.00371876	CDK5
GO:0034983	Desacetilación de peptidil-lisina	1	0.00371876	HDAC9
GO:0071156	Regulación de la detención del ciclo celular	1	0.00371876	CDK5
GO:0046826	Regulación negativa de la exportación de proteínas desde el núcleo	1	0.00371876	CDK5
GO:0070932	Desacetilación de histona H3	1	0.00371876	HDAC9
GO:0023021	Terminación de la transducción de señales	1	0.00386916	PRKCD
GO:0016576	Desfosforilación de histonas	1	0.00386916	EYA2
GO:0090331	Regulación negativa de la agregación plaquetaria	1	0.00386916	PRKCD
GO:0021695	Desarrollo de la corteza cerebelosa	1	0.00386916	CDK5
GO:0014044	Desarrollo de células de Schwann	1	0.00386916	CDK5
GO:0071173	Punto de control del conjunto del huso	1	0.00386916	BUB1
GO:0009790	Desarrollo embrionario	2	0.00393274	BUB1, CDK5
GO:0002903	Regulación negativa de la apoptosis de células B	1	0.00407877	AURKB
GO:0045055	Vía secretora regulada	1	0.00407877	CDK5
GO:0032930	Regulación positiva de la generación de aniones superóxido	1	0.00407877	PRKCD
GO:0032801	Proceso catabólico del receptor	1	0.00407877	CDK5
GO:0030071	Regulación de la transición metafase / anafase mitótico	1	0.00407877	NEK6
GO:2000251	Regulación positiva de la reorganización del citoesqueleto de actina	1	0.00407877	CDK5
GO:0034501	Localización de proteínas en el cinetocoro	1	0.00413455	AURKB
GO:0030837	Regulación negativa de la polimerización del filamento de actina	1	0.00413455	PRKCD
GO:0008608	Acoplamiento de los microtúbulos del huso al cinetocoro	1	0.00413455	AURKB
GO:0070933	Desacetilación de histona H4	1	0.00413455	HDAC9
GO:0061001	Regulación de la morfogénesis de la columna dendrítica	1	0.00413455	CDK5
GO:0048675	Extensión del axón	1	0.00413455	CDK5
GO:0022038	Desarrollo del cuerpo calloso	1	0.00413455	CDK5
GO:0090050	Regulación positiva de la migración celular involucrada en la germinación de la angiogénesis	1	0.00413455	HDAC9
GO:0016568	Modificación de la cromatina	2	0.00453237	HDAC9, EYA2
GO:0010569	Regulación de la reparación de una rotura de doble cadena en el ADN mediante recombinación homóloga	1	0.0046623	PPP4C
GO:0048148	Respuesta conductual a la cocaína	1	0.0046623	CDK5

GO:0035307	Regulación positiva de la desfosforilación de proteínas	1	0.00515586	PRKCD
GO:0042119	Activación de neutrófilos	1	0.00515586	PRKCD
GO:0051301	División celular	2	0.00564107	BUB1, CDK5
GO:0006281	Reparación de ADN	2	0.00568992	UHRF1, EYA2
GO:0048488	Endocitosis de vesícula sináptica	1	0.00570761	CDK5
GO:0050732	Regulación negativa de la fosforilación de peptidil-tirosina	1	0.00596247	PRKCD
GO:0001963	Transmisión sináptica, dopaminérgica	1	0.00596247	CDK5
GO:0000278	Ciclo celular mitótico	2	0.00607859	BUB1, AURKB
GO:0090398	Senescencia celular	1	0.0062769	PRKCD
GO:0010469	Regulación de la actividad del receptor	1	0.0062769	PRKCD
GO:0032270	Regulación positiva de la proteína celular del proceso metabólico	1	0.00675029	UHRF1
GO:0042100	Proliferación de células B	1	0.00692199	PRKCD
GO:0016064	Respuesta inmune mediada por inmunoglobulina	1	0.00692199	PRKCD
GO:0071777	Regulación positiva de la citocinesis del ciclo celular	1	0.00692199	AURKB
GO:0043113	Agrupamiento de receptores	1	0.00692199	CDK5
GO:0016079	Exocitosis de vesícula sináptica	1	0.00725998	CDK5
GO:0007093	Punto de control del ciclo celular mitótico	1	0.00725998	BUB1
GO:0048709	Diferenciación de oligodendrocitos	1	0.00767793	CDK5
GO:0035249	Transmisión sináptica, glutamatérgica	1	0.00798411	CDK5
GO:0021954	Desarrollo neuronal del sistema nervioso central	1	0.00798411	CDK5
GO:0030866	Organización del citoesqueleto de actina cortical	1	0.00827525	CDK5
GO:0031397	Regulación negativa de la ubiquitinación de proteínas	1	0.00827525	CDK5
GO:0008045	Guía del motor del axón	1	0.00855241	CDK5
GO:0045892	Regulación negativa de la transcripción a nivel del ADN molde	2	0.00855786	HDAC9, CDK5
GO:0000122	Regulación negativa de la transcripción del promotor de la ARN polimerasa II	2	0.00866976	HDAC9, AURKB
GO:0048813	Morfogénesis dendrítica	1	0.00871523	CDK5
GO:0016575	Desacetilación de histonas	1	0.00871523	HDAC9
GO:0032091	Regulación negativa de la unión a proteínas	1	0.00871523	PRKCD
GO:0051402	Apoptosis neuronal	1	0.00886709	CDK5
GO:0046627	Regulación negativa de la vía de señalización del receptor de insulina	1	0.00886709	PRKCD
GO:0060079	Regulación del potencial excitador de membrana postsináptica	1	0.00900905	CDK5
GO:0048167	Regulación de la plasticidad sináptica	1	0.00900905	CDK5
GO:0034644	Respuesta celular a la UV	1	0.00900905	AURKB

GO:0031572	Punto de control de daños en el ADN de transición G2 / M	1	0.00933656	NEK6
GO:0051225	Formación del huso	1	0.00955648	NEK6
GO:0007596	Coagulación de la sangre	2	0.00963988	PRKCD, CDK5
GO:0018107	Fosforilación de peptidil-treonina	1	0.00986802	PRKCD
Función molecular (FM)				
GO:0004672	Actividad de la proteína quinasa	5	2.01385E-07	BUB1, AURKB, PRKCD, CDK5, NEK6
GO:0005515	Unión a proteínas	9	2.61009E-07	BUB1, HDAC9, AURKB, PRKCD, PPP4C, CDK5, NEK6, UHRF1, EYA2
GO:0004674	Actividad de proteína serina / treonina quinasa	5	3.07829E-07	BUB1, AURKB, PRKCD, CDK5, NEK6
GO:0005524	Unión a ATP	5	0.000202639	BUB1, AURKB, PRKCD, CDK5, NEK6
GO:0000166	Unión de nucleótidos	5	0.000891396	BUB1, AURKB, PRKCD, CDK5, NEK6
GO:0030549	Actividad activadora del receptor de acetilcolina	1	0.00210477	CDK5
GO:0004699	Actividad de la proteína quinasa C independiente del calcio	1	0.00360776	PRKCD
GO:0005176	Unión al receptor de clase ErbB-2	1	0.00473463	CDK5
GO:0000287	Unión de iones de magnesio	2	0.00504683	NEK6, EYA2
GO:0008327	Unión de metil-CpG	1	0.0057376	UHRF1
GO:0043125	Unión al receptor de la clase ErbB-3	1	0.0057376	CDK5
GO:0016772	Actividad transferasa, transfiriendo grupos que contienen fósforo	2	0.00575523	BUB1, NEK6
GO:0016787	Actividad hidrolasa	3	0.0062208	HDAC9, PPP4C, EYA2
GO:0004704	Actividad quinasa inductora de NF-kappaB	1	0.00630988	PPP4C
GO:0033558	Actividad de la proteína deacetilasa	1	0.00672976	HDAC9
GO:0004407	Actividad de histona deacetilasa	1	0.00712791	HDAC9
GO:0046969	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	1	0.00712791	HDAC9
GO:0000987	Unión al ADN en la región promotora proximal	1	0.00720708	UHRF1
GO:0043560	Unión al sustrato del receptor de insulina	1	0.00720708	PRKCD
GO:0050321	Actividad tau-proteína quinasa	1	0.00730276	CDK5
GO:0034739	Actividad de histona deacetilasa (específica para H3-K16)	1	0.00730276	HDAC9
GO:0032041	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K14)	1	0.00730276	HDAC9
GO:0046970	Actividad de histona deacetilasa dependiente de NAD (específica de H4-K16)	1	0.00730276	HDAC9
GO:0046872	Unión de iones metálicos	4	0.00892926	AURKB, PRKCD, PPP4C, UHRF1
GO:0019894	Unión de quinesina	1	0.00907455	NEK6

Componente celular (CC)				
GO:0005634	Núcleo	9	2.65659E-06	BUB1, HDAC9, AURKB, PRKCD, PPP4C, CDK5, NEK6, UHRF1, EYA2
GO:0000780	Cromosoma nuclear condensado, región centromérica	2	5.66999E-05	BUB1, AURKB
GO:0005856	Citoesqueleto	4	0.000579263	AURKB, PPP4C, CDK5, NEK6
GO:0005737	Citoplasma	7	0.000590863	HDAC9, AURKB, PRKCD, PPP4C, CDK5, NEK6, EYA2
GO:0000779	Cromosoma condensado, región centromérica	1	0.00616181	AURKB
GO:0032133	Complejo pasajero de cromosomas	1	0.00616181	AURKB
GO:0016533	Complejo de holoenzima proteína quinasa 5 dependiente de ciclina	1	0.00647066	CDK5
GO:0005829	Citosol	4	0.00772375	BUB1, AURKB, PRKCD, CDK5
GO:0044428	Parte nuclear	1	0.0086225	CDK5
GO:0005694	Cromosoma	2	0.00945927	BUB1, AURKB

Tabla 11A. Análisis de enriquecimiento GO de genes subexpresados en OV utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes subexpresados
Proceso Biológico (PB)				
GO:0006468	Fosforilación de proteínas	9	9.97357E-14	KDM6A, SETDB2, EZH1, MYSM1, HDAC6, JAK2, HDAC5, HDAC10, SETD7
GO:0009314	Respuesta a la radiación	5	2.41788E-06	DUSP1, MAP3K20, HDAC6, JAK2, ATM
GO:0008283	Proliferación celular	3	7.46997E-06	HDAC6, HDAC5, HDAC10
GO:0044419	Interacción interespecies entre organismos	5	0.000101051	DUSP1, MAP3K20, SETDB2, KAT2B, NEK9
GO:0045944	Regulación positiva de la transcripción del promotor del ARN polimerasa II	3	0.000113501	KAT2A, KAT2B, HDAC5
GO:0051384	Respuesta al estímulo glucocorticoide	2	0.00048454	HDAC6, HDAC10
GO:0043388	Regulación positiva de la unión del ADN	2	0.000498238	KAT2A, MYSM1
GO:0000122	Regulación negativa de la transcripción del promotor de la ARN polimerasa II	3	0.000976807	MAP3K20, KAT2B, ATM
GO:0043065	Regulación positiva del proceso apoptótico	4	0.000993901	SETDB2, HDAC6, HDAC5, HDAC10
GO:0032516	Regulación positiva de la actividad fosfoproteína fosfatasa	2	0.0017104	MAP3K20, ATM
GO:0048008	Vía de señalización del receptor del factor de crecimiento derivado de plaquetas	3	0.00214137	CDK3, SETDB2, NEK9
GO:0018108	Fosforilación de peptidil-tirosina	2	0.00231088	JAK2, MGEA5
GO:0046777	Autofosforilación de proteínas	2	0.00277494	KAT2A, KAT2B
GO:0014068	Regulación positiva de la señalización de fosfatidilinositol 3-quinasa.	2	0.00277494	KAT2A, ATM
GO:0010212	Respuesta a la radiación ionizante	2	0.00300673	MAP3K20, JAK2
GO:0045892	Regulación negativa de la transcripción a nivel del ADN molde	1	0.00404667	MGEA5
GO:0008152	Proceso metabólico	1	0.00404667	MGEA5
GO:0008285	Regulación negativa de la proliferación celular	1	0.00404667	MGEA5
GO:0009967	Regulación positiva de la transducción de señales	1	0.00404667	JAK2
GO:0007049	Ciclo celular	1	0.00404667	JAK2
GO:0051591	Respuesta al AMPc	1	0.00404667	JAK2
GO:0032869	Respuesta celular al estímulo de insulina	1	0.00404667	HDAC6
GO:0051592	Respuesta al ion calcio	1	0.00404667	HDAC6
GO:0048545	Respuesta al estímulo de la hormona esteroide	1	0.00404667	HDAC6
GO:0030901	Desarrollo del cerebro medio	1	0.00404667	HDAC6
GO:0007243	Cascada de proteína quinasa intracelular	2	0.00408357	JAK2, HDAC5

GO:0030154	Diferenciación celular	3	0.00483538	CDK3, SETDB2, NEK9
GO:0000186	Activación de la quinasa MAPKK	2	0.0050212	KAT2B, HDAC5
GO:0007050	Detención del ciclo celular (PA)	1	0.00514911	JAK2
GO:0051301	División celular	1	0.00514911	MGEA5
GO:0001947	Bucle cardíaco	1	0.00514911	JAK2
GO:0042981	Regulación del proceso apoptótico	1	0.00514911	JAK2
GO:0007067	Mitosis	1	0.00514911	HDAC5
GO:0030183	Diferenciación de células B	1	0.00514911	HDAC6
GO:0051726	Regulación del ciclo celular	1	0.00514911	HDAC6
GO:0010811	Regulación positiva de la adhesión del sustrato celular	1	0.00514911	HDAC6
GO:0050731	Regulación positiva de la fosforilación de peptidil-tirosina	1	0.00514911	HDAC6
GO:0006302	Reparación de rotura de doble cadena	1	0.00514911	HDAC6
GO:0035556	Transducción de señales intracelulares	1	0.00514911	HDAC6
GO:0032024	Regulación positiva de la secreción de insulina	1	0.00514911	KAT2B
GO:0060399	Regulación positiva de la vía de señalización del receptor de la hormona del crecimiento	1	0.00514911	KAT2A
GO:0060124	Regulación positiva de la secreción de hormona de crecimiento	1	0.00514911	CDK3
GO:0009755	Vía de señalización mediada por hormonas	3	0.00541915	EZH1, HDAC5, HDAC10
GO:0006977	Regulación positiva de la detención del ciclo celular por la respuesta al daño del ADN mediada por p53	1	0.00566269	MGEA5
GO:0000075	Punto de control del ciclo celular	1	0.00566269	MGEA5
GO:0030041	Polimerización de filamentos de actina	1	0.00566269	KAT2A
GO:0033209	Vía de señalización mediada por el factor de necrosis tumoral	1	0.00566269	ATM
GO:0032870	Respuesta celular al estímulo hormonal	1	0.00566269	ATM
GO:0018105	Fosforilación de peptidil-serina	1	0.00566269	KAT2B
GO:0006612	Proteínas dirigidas a la membrana	1	0.00566269	JAK2
GO:0045862	Regulación positiva de la proteólisis	1	0.00566269	JAK2
GO:0010616	Regulación negativa de la adaptación del músculo cardíaco	1	0.00566269	KAT2B
GO:0060051	Regulación negativa de la glicosilación de proteínas	1	0.00566269	JAK2
GO:0046060	Proceso metabólico de dATP	1	0.00566269	SETDB2
GO:0043243	Regulación positiva del desensamble del complejo proteico	1	0.00566269	JAK2
GO:0031343	Regulación positiva de la muerte celular	1	0.00566269	JAK2
GO:0006516	Proceso catabólico de glicoproteínas	1	0.00566269	HDAC6
GO:0006044	Proceso metabólico de la N-acetilglucosamina	1	0.00566269	HDAC6

GO:0051901	Regulación positiva de la despolarización mitocondrial	3	0.00588688	MAP3K20, CDK3, KAT2A
GO:0051054	Regulación positiva del proceso metabólico del ADN	1	0.00675984	MGEA5
GO:0046326	Regulación positiva de la importación de glucosa	1	0.00675984	HDAC6
GO:0070265	Muerte celular necrótica	1	0.00675984	HDAC6
GO:0010524	Regulación positiva del transporte de iones de calcio al citosol	1	0.00675984	JAK2
GO:0016568	Modificación de la cromatina	1	0.00675984	HDAC6
GO:0008219	Muerte celular	1	0.00675984	HDAC6
GO:0032760	Regulación positiva de la producción del factor de necrosis tumoral	2	0.00712444	MAP3K20, ATM
GO:0007257	Activación de la quinasa JUN	2	0.00751703	MAP3K20, JAK2
GO:0021537	Desarrollo del telencéfalo	1	0.00764871	JAK2
GO:0001706	Formación del endodermo	1	0.00764871	SETD7
GO:0050727	Regulación de la respuesta inflamatoria (PA).	1	0.00764871	EZH1
GO:0018026	Monometilación de peptidil-lisina (BP)	1	0.00764871	JAK2
GO:0034612	Respuesta al factor de necrosis tumoral	1	0.00764871	JAK2
GO:0045597	Regulación positiva de la diferenciación celular	1	0.00764871	HDAC6
GO:0060333	Vía de señalización mediada por interferón gamma	2	0.00812422	JAK2, ATM
GO:0044154	Acetilación de histona H3-K14	2	0.00816757	MAP3K20, ATM
GO:0007059	Segregación cromosómica	1	0.00828105	JAK2
GO:0032526	Respuesta al ácido retinoico	1	0.00828105	MGEA5
GO:0033574	Respuesta al estímulo de la testosterona	1	0.00828105	MGEA5
GO:0051091	Regulación positiva de la actividad del factor de transcripción de unión al ADN	1	0.00828105	SETD7
GO:0043392	Regulación negativa de la unión al ADN	1	0.00828105	HDAC6
GO:0042542	Respuesta al peróxido de hidrógeno	1	0.00828105	ATM
GO:0006338	Remodelación de la cromatina	1	0.00828105	HDAC5
GO:0042517	Regulación positiva de la fosforilación de tirosina de la proteína Stat3	1	0.00900041	HDAC6
GO:0050729	Regulación positiva de la respuesta inflamatoria	1	0.00900041	ATM
GO:0042523	Regulación positiva de la fosforilación de tirosina de la proteína Stat5	1	0.00900041	SETDB2
GO:0007498	Desarrollo del mesodermo	1	0.00900041	HDAC6
GO:0006325	Organización de la cromatina	1	0.00900041	JAK2
GO:0001843	Cierre del tubo neural	1	0.00900041	HDAC5
GO:0009636	Respuesta a la toxina	1	0.00983665	MGEA5
GO:0014003	Desarrollo de oligodendrocitos	1	0.00983665	JAK2
GO:0010033	Respuesta a la sustancia orgánica	1	0.00983665	JAK2
GO:0000077	Punto de control de daños en el ADN	1	0.00983665	HDAC5

Función molecular (FM)				
GO:0034739	Actividad de histona deacetilasa (específica para H3-K16)	3	3.60721E-07	HDAC6, HDAC5, HDAC10
GO:0032041	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K14)	3	3.60721E-07	HDAC6, HDAC5, HDAC10
GO:0046970	Actividad de histona deacetilasa dependiente de NAD (específica de H4-K16)	3	3.60721E-07	HDAC6, HDAC5, HDAC10
GO:0004407	Actividad de histona deacetilasa	3	4.16577E-07	HDAC6, HDAC5, HDAC10
GO:0046969	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	3	4.16577E-07	HDAC6, HDAC5, HDAC10
GO:0042826	Unión de histona deacetilasa	4	1.39463E-06	KAT2A, KAT2B, HDAC6, HDAC10
GO:0005515	Unión a proteínas	11	0.000013107	DUSP1, MAP3K20, KAT2A, KAT2B, HDAC6, JAK2, HDAC5, HDAC10, NEK9, ATM, SETD7
GO:0004402	Actividad de histona acetiltransferasa	3	0.000014331	KAT2A, KAT2B, MGEA5
GO:0008168	Actividad metiltransferasa	3	0.000426951	SETDB2, EZH1, SETD7
GO:0003713	Actividad coactivadora de transcripción	3	0.00143563	KAT2A, KAT2B, MYSM1
GO:0019901	Unión de la proteína quinasa	3	0.00148936	KAT2B, JAK2, NEK9
GO:0016740	Actividad transferasa	4	0.00160777	KAT2A, SETDB2, EZH1, SETD7
GO:0042393	Unión de histonas	2	0.00216731	MYSM1, JAK2
GO:0035401	Actividad de la histona quinasa (específica de H3-Y41)	1	0.00267116	JAK2
GO:0070840	Unión del complejo de dineína	1	0.00267116	HDAC6
GO:0043997	Actividad de la histona acetiltransferasa (específica de H4-K12)	1	0.00267116	KAT2A
GO:0005524	Unión a ATP	5	0.00313238	MAP3K20, CDK3, JAK2, NEK9, ATM
GO:0004674	Actividad de proteína serina / treonina quinasa	3	0.00384643	MAP3K20, NEK9, ATM
GO:0004468	Actividad de lisina N-acetiltransferasa	1	0.00406939	KAT2B
GO:0042903	Actividad tubulina deacetilasa	1	0.00406939	HDAC6
GO:0010484	Actividad acetiltransferasa de la histona H3	1	0.00406939	KAT2A
GO:0016787	Actividad hidrolasa	4	0.00446683	DUSP1, HDAC6, HDAC5, HDAC10
GO:0035174	Actividad de histonas serina quinasa	1	0.00533982	ATM
GO:0046976	Actividad de histona metiltransferasa (H3-K27 específica)	1	0.00533982	EZH1
GO:0005143	Unión al receptor de interleucina-12	1	0.00688687	JAK2
GO:0004677	Actividad de proteína quinasa dependiente de ADN	1	0.00688687	ATM
GO:0046974	Actividad de histona metiltransferasa (H3-K9 específica)	1	0.00688687	SETDB2
GO:0005131	Unión del receptor de la hormona del crecimiento	1	0.00688687	JAK2
GO:0033130	Unión al receptor de acetilcolina	1	0.00688687	JAK2
GO:0008330	Actividad de la proteína tirosina / treonina fosfatasa	1	0.00688687	DUSP1
GO:0016746	Actividad transferasa, transfiriendo grupos acilo	2	0.00705876	KAT2B, MGEA5
GO:0000166	Unión de nucleótidos	5	0.00737642	MAP3K20, CDK3, JAK2, NEK9, ATM

GO:0048156	Unión a la proteína tau	1	0.00753328	HDAC6
GO:0004415	Actividad de la enzima condroitinasa I	1	0.00753328	MGEA5
GO:0016279	Actividad de proteína-lisina N-metiltransferasa	1	0.00753328	SETD7
GO:0033558	Actividad de la proteína deacetilasa	1	0.00898287	HDAC10
GO:0004726	Actividad de la proteína tirosina fosfatasa no que atraviesa la membrana	1	0.00898287	DUSP1
GO:0019899	Unión de enzimas	2	0.00934112	HDAC6, HDAC10
GO:0051428	Unión del receptor de la hormona peptídica	1	0.00984431	JAK2
Componente celular (CC)				
GO:0005634	Núcleo	16	9.52847E-11	DUSP1, KDM6A, MAP3K20, KAT2A, SETDB2, EZH1, KAT2B, MYSM1, HDAC6, JAK2, HDAC5, HDAC10, NEK9, ATM, SETD7, MGEA5
GO:0000118	Complejo de histona deacetilasa	3	7.39004E-06	HDAC6, HDAC5, HDAC10
GO:0016585	Complejo de remodelación de cromatina	2	0.000243145	KAT2B, MYSM1
GO:0005671	Complejo activador de la transcripción Ada2 / Gcn5 / Ada3	2	0.000281049	KAT2A, KAT2B
GO:0005901	Caveola	2	0.00262093	HDAC6, JAK2
GO:0005654	Nucleoplasma	4	0.00553222	DUSP1, KAT2B, HDAC10, ATM

Tabla 12A. Análisis de enriquecimiento GO de genes sobreexpresados en UCEC utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes sobreexpresados
Proceso Biológico (PB)				
GO:0051983	Regulación de la segregación cromosómica	2	2.68E-05	BUB1, AURKB
GO:0008283	Proliferación celular	4	5.17E-05	BUB1, AURKB, CDK5, UHRF1
GO:0007049	Ciclo celular	4	6.47E-05	AURKB, CDK5, NEK6, UHRF1
GO:0046777	Autofosforilación de proteínas	3	0.000122	AURKB, CDK5, NEK6
GO:0016568	Modificación de la cromatina	3	0.000339	SMYD2, EZH2, EYA2
GO:0000910	Citocinesis	2	0.001172	AURKB, NEK6
GO:0006468	Fosforilación de proteínas	3	0.001173	BUB1, CDK5, NEK6
GO:0018105	Fosforilación de peptidil-serina	2	0.001178	CDK5, NEK6
GO:0000236	Prometáfase mitótica	2	0.002106	BUB1, AURKB
GO:0021697	Formación de la corteza cerebelosa	1	0.002183	CDK5
GO:0007063	Regulación de la cohesión entre cromátidas hermanas	1	0.002183	BUB1
GO:0006355	Regulación de la transcripción a nivel del ADN molde	4	0.002345	SMYD2, EZH2, UHRF1, EYA2
GO:0006915	Proceso apoptótico	3	0.002355	BUB1, NEK6, EYA2
GO:0000087	Fase M del ciclo celular mitótico	2	0.002367	BUB1, AURKB
GO:0014706	Desarrollo del tejido muscular estriado	1	0.002619	EYA2
GO:2000373	Regulación positiva de la topoisomerasa del ADN (ATP-hidrolizante)	1	0.002619	UHRF1
GO:0033136	Fosforilación de serina de la proteína STAT3	1	0.002619	CDK5
GO:0031914	Regulación negativa de la plasticidad sináptica	1	0.002619	CDK5
GO:0043516	Regulación de p53 inducida por la respuesta al daño del ADN	1	0.002619	SMYD2
GO:0009790	Desarrollo embrionario	2	0.002862	BUB1, CDK5
GO:2000273	Regulación positiva de la actividad del receptor	1	0.003081	CDK5
GO:2000772	Regulación de la senescencia celular	1	0.003081	NEK6
GO:0007501	Especificación del destino de las células mesodérmicas	1	0.003081	EYA2
GO:0016570	Modificación de histonas	1	0.003081	AURKB
GO:0010452	Metilación de lisina K36 en la histona H3	1	0.003081	SMYD2
GO:0051256	Formación de la zona media del huso mitótico	1	0.003081	AURKB
GO:0021819	Formación de capas en la corteza cerebral	1	0.003081	CDK5

GO:0045956	Regulación positiva de la exocitosis dependiente de los iones de calcio	1	0.003081	CDK5
GO:0071156	Regulación de la detención del ciclo celular	1	0.003081	CDK5
GO:0046826	Regulación negativa de la exportación de proteínas desde el núcleo	1	0.003081	CDK5
GO:0002903	Regulación negativa de la apoptosis de células B	1	0.003273	AURKB
GO:0045055	Vía secretora regulada	1	0.003273	CDK5
GO:0018027	Dimetilación de peptidil-lisina	1	0.003273	SMYD2
GO:0032801	Proceso catabólico del receptor	1	0.003273	CDK5
GO:0030071	Regulación de la transición metafase / anafase mitótico	1	0.003273	NEK6
GO:2000251	Regulación positiva de la reorganización del citoesqueleto de actina	1	0.003273	CDK5
GO:0016576	Desfosforilación de histonas	1	0.003273	EYA2
GO:0021695	Desarrollo de la corteza cerebelosa	1	0.003273	CDK5
GO:0014044	Desarrollo de células de Schwann	1	0.003273	CDK5
GO:0071173	Punto de control del conjunto del huso	1	0.003273	BUB1
GO:0018026	Monometilación de peptidil-lisina	1	0.003415	SMYD2
GO:0034501	Localización de proteínas en el cinetocoro	1	0.003415	AURKB
GO:0008608	Acoplamiento de los microtúbulos del huso al cinetocoro	1	0.003415	AURKB
GO:0061001	Regulación de la morfogénesis de la columna dendrítica	1	0.003415	CDK5
GO:0048675	Extensión del axón	1	0.003415	CDK5
GO:0022038	Desarrollo del cuerpo calloso	1	0.003415	CDK5
GO:0048148	Respuesta conductual a la cocaína	1	0.003899	CDK5
GO:0051301	División celular	2	0.004226	BUB1, CDK5
GO:0006281	Reparación del ADN	2	0.004283	UHRF1, EYA2
GO:0048387	Regulación negativa de la vía de señalización del receptor de ácido retinoico	1	0.004362	EZH2
GO:0000278	Ciclo celular mitótico	2	0.004582	BUB1, AURKB
GO:0048488	Endocitosis de vesícula sináptica	1	0.004618	CDK5
GO:0001963	Transmisión sináptica, dopaminérgica	1	0.004937	CDK5
GO:0032270	Regulación positiva de la proteína celular del proceso metabólico	1	0.005814	UHRF1
GO:0071777	Regulación positiva de la citocinesis del ciclo celular	1	0.006073	AURKB
GO:0043113	Agrupamiento de receptores	1	0.006073	CDK5
GO:0016079	Exocitosis de vesícula sináptica	1	0.006314	CDK5

GO:0007093	Punto de control del ciclo celular mitótico	1	0.006314	BUB1
GO:0048709	Diferenciación de oligodendrocitos	1	0.00665	CDK5
GO:0035249	Transmisión sináptica, glutamatérgica	1	0.006749	CDK5
GO:0021954	Desarrollo neuronal del sistema nervioso central	1	0.006749	CDK5
GO:0045892	Regulación negativa de la transcripción, dependiente del ADN	2	0.006837	EZH2, CDK5
GO:0030866	Organización del citoesqueleto de actina cortical	1	0.006839	CDK5
GO:0031397	Regulación negativa de la ubiquitinación de proteínas	1	0.006839	CDK5
GO:0000122	Regulación negativa de la transcripción del promotor del ARN polimerasa II	2	0.006996	SMYD2, AURKB
GO:0008045	Guía del motor del axón	1	0.007025	CDK5
GO:0016571	Metilación de histonas	1	0.007025	EZH2
GO:0048813	Morfogénesis dendrítica	1	0.007305	CDK5
GO:0051402	Apoptosis neuronal	1	0.007469	CDK5
GO:0006325	Organización de la cromatina	1	0.007469	EZH2
GO:0060079	Regulación del potencial excitador de membrana postsináptica	1	0.007519	CDK5
GO:0048167	Regulación de la plasticidad sináptica	1	0.007519	CDK5
GO:0034644	Respuesta celular a la UV	1	0.007519	AURKB
GO:0031572	Punto de control de daños en el ADN de transición G2 / M	1	0.00777	NEK6
GO:0051225	Conjunto del huso	1	0.008014	NEK6
GO:0006913	Transporte nucleocitoplásmico	1	0.008938	CDK5
GO:0007346	Regulación del ciclo celular mitótico	1	0.00916	NEK6
GO:0007094	Punto de control del conjunto del huso del ciclo celular mitótico	1	0.009917	BUB1
GO:0008542	Aprendizaje visual	1	0.009917	CDK5
Función molecular (CC)				
GO:0005515	Unión a proteínas	8	3.09E-06	SMYD2, BUB1, AURKB, EZH2, CDK5, NEK6, UHRF1, EYA2
GO:0004672	Actividad de la proteína quinasa	4	5.39E-06	BUB1, AURKB, CDK5, NEK6
GO:0004674	Actividad de proteína serina / treonina quinasa	4	1.21E-05	BUB1, AURKB, CDK5, NEK6
GO:0002039	Unión a p53	2	0.000398	SMYD2, CDK5
GO:0030549	Activación del receptor de acetilcolina	1	0.001442	CDK5
GO:0005524	Unión a ATP	4	0.001609	BUB1, AURKB, CDK5, NEK6
GO:0000993	Unión a la ARN polimerasa II	1	0.002163	SMYD2
GO:0008168	Actividad metiltransferasa	2	0.002456	SMYD2, EZH2
GO:0000287	Unión de iones de magnesio	2	0.002484	NEK6, EYA2

GO:0000166	Unión de nucleótidos	4	0.002594	BUB1, AURKB, CDK5, NEK6
GO:0005176	Unión al receptor de clase ErbB-2	1	0.002595	CDK5
GO:0046975	Actividad de histona metiltransferasa (H3-K36 específica)	1	0.002595	SMYD2
GO:0008327	Unión de metil-CpG	1	0.002883	UHRF1
GO:0043125	Unión al receptor de la clase ErbB-3	1	0.002883	CDK5
GO:0016772	Actividad transferasa, transfiriendo grupos que contienen fósforo	2	0.003199	BUB1, NEK6
GO:0016279	Actividad de proteína-lisina N-metiltransferasa	1	0.003243	SMYD2
GO:0042054	Actividad de histona metiltransferasa	1	0.004577	EZH2
GO:0050321	Actividad tau-proteína quinasa	1	0.005283	CDK5
GO:0000987	Promotor del núcleo Unión proximal de secuencia específica del ADN (MF)	1	0.005459	UHRF1
GO:0019894	Unión de quinesinas	1	0.007774	NEK6
GO:0004712	Actividad de la proteína serina / treonina / tirosina quinasa	1	0.008635	AURKB
GO:0018024	Actividad de histona-lisina N-metiltransferasa	1	0.009026	EZH2
Componente celular (CC)				
GO:0005634	Núcleo	8	1.39E-05	SMYD2, BUB1, AURKB, EZH2, CDK5, NEK6, UHRF1, EYA2
GO:0000780	Cromosoma nuclear condensado, región centromérica	2	3.66E-05	BUB1, AURKB
GO:0000779	Cromosoma condensado, región centromérica	1	0.004542	AURKB
GO:0032133	Complejo pasajero de cromosomas	1	0.004542	AURKB
GO:0005829	Citosol	4	0.005991	SMYD2, BUB1, AURKB, CDK5
GO:0005856	Citoesqueleto	3	0.006538	AURKB, CDK5, NEK6
GO:0005694	Cromosoma	2	0.006899	BUB1, AURKB
GO:0044428	Parte nuclear	1	0.007063	CDK5
GO:0035098	Complejo ESC / E (Z)	1	0.007944	EZH2
GO:0016533	Complejo de holoenzima proteína quinasa 5 dependiente de ciclina	1	0.00795	CDK5

Tabla 13A. Análisis de enriquecimiento GO de genes subexpresados en UCEC utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes subexpresados
Proceso Biológico (PB)				
GO:0016568	Modificación de la cromatina	10	9.92115E-15	KMT2D, KMT2A, EZH1, HDAC4, SETD1B, MYSM1, HDAC6, JAK2, HDAC5, SETD7
GO:0006338	Remodelación de la cromatina	4	4.11458E-06	KAT2A, HDAC4, KAT2B, HDAC5
GO:0051568	Metilación de lisina K4 en la histona H3	3	4.57019E-06	KMT2D, KMT2A, SETD1B
GO:0043065	Regulación positiva del proceso apoptótico	5	5.78355E-06	DUSP1, MAP3K20, HDAC6, JAK2, ATM
GO:0016575	Desacetilación de histonas	3	9.7138E-06	HDAC4, HDAC6, HDAC5
GO:0018076	Acetilación de peptidil-lisina N-terminal	2	3.89036E-05	KAT2B, CREBBP
GO:0045944	Regulación positiva de la transcripción del promotor del ARN polimerasa II	6	4.18112E-05	KMT2D, KMT2A, HDAC4, KAT2B, HDAC5, CREBBP
GO:0034983	Desacetilación de peptidil-lisina	2	0.000068056	HDAC4, HDAC6
GO:0006355	Regulación de la transcripción a nivel del ADN molde	8	7.77062E-05	KMT2D, KMT2A, EZH1, NCOA1, SETD1B, PRDM2, CREBBP, SETD7
GO:0010832	Regulación negativa de la diferenciación de miotubos	2	0.000190345	HDAC4, HDAC5
GO:0006342	Silenciamiento mediado por la cromatina	2	0.000211416	KMT2D, HDAC5
GO:0051091	Regulación positiva de la actividad del factor de transcripción de unión al ADN	3	0.000223985	HDAC4, JAK2, HDAC5
GO:0016578	Desubiquitinación de histonas	2	0.000426515	KAT2A, MYSM1
GO:0043393	Regulación de la unión a proteínas	2	0.000426515	HDAC4, HDAC5
GO:0008285	Regulación negativa de la proliferación celular	4	0.000856504	KMT2A, HDAC4, KAT2B, JAK2
GO:0007050	Detención del ciclo celular	3	0.00104837	MAP3K20, KAT2B, ATM
GO:0000122	Regulación negativa de la transcripción del promotor de la ARN polimerasa II	4	0.00153922	EZH1, HDAC4, HDAC5, CREBBP
GO:0000077	Punto de control de daños en el ADN	2	0.00153956	MAP3K20, ATM
GO:0042113	Activación de células B	2	0.00158892	HDAC4, HDAC5
GO:0007049	Ciclo celular	4	0.00164479	DUSP1, MAP3K20, KAT2B, NEK9
GO:0045893	Regulación positiva de la transcripción a nivel del ADN molde	4	0.00207326	KMT2A, HDAC4, MYSM1, CREBBP
GO:0032024	Regulación positiva de la secreción de insulina	2	0.00215454	JAK2, MGEA5
GO:0043966	Acetilación de histona H3	2	0.00276557	KAT2A, KAT2B

GO:0001756	Somitogenesis	2	0.00276557	KAT2A, ATM
GO:0030183	Diferenciación de células B	2	0.00279365	HDAC4, HDAC5
GO:0000186	Activación de la quinasa MAPKK	2	0.00296296	MAP3K20, JAK2
GO:0010616	Regulación negativa de la adaptación del músculo cardíaco	1	0.00388286	MGEA5
GO:0060051	Regulación negativa de la glicosilación de proteínas	1	0.00388286	MGEA5
GO:0046060	Proceso metabólico de dATP	1	0.00388286	MGEA5
GO:0035409	Fosforilación de histona H3-Y41	1	0.00388286	JAK2
GO:0031959	Vía de señalización del receptor mineralocorticoide	1	0.00388286	JAK2
GO:0008636	Activación de la actividad de la caspasa por fosforilación de proteínas	1	0.00388286	JAK2
GO:0010727	Regulación negativa del proceso metabólico del peróxido de hidrógeno	1	0.00388286	HDAC6
GO:0070846	Desacetilación Hsp90	1	0.00388286	HDAC6
GO:0090042	Desacetilación de tubulina	1	0.00388286	HDAC6
GO:0035967	Respuesta celular a la proteína topológicamente incorrecta	1	0.00388286	HDAC6
GO:0010882	Regulación de la contracción del músculo cardíaco por señalización de ion calcio (PA).	1	0.00388286	HDAC4
GO:0080182	Trimetilación de lisina K4 en la histona H3	1	0.00388286	KMT2A
GO:2001040	Regulación positiva de la respuesta celular al fármaco	1	0.00388286	KMT2A
GO:0032869	Respuesta celular al estímulo de insulina	2	0.00507696	KAT2B, HDAC5
GO:0060399	Regulación positiva de la vía de señalización del receptor de la hormona del crecimiento	1	0.00526336	JAK2
GO:0031343	Regulación positiva de la muerte celular	1	0.00526336	MGEA5
GO:0042508	Fosforilación de tirosina de la proteína Stat1	1	0.00526336	JAK2
GO:0070671	Respuesta a la interleucina-12	1	0.00526336	JAK2
GO:0090051	Regulación negativa de la migración celular involucrada en la germinación de la angiogénesis	1	0.00526336	HDAC5
GO:0090035	Regulación positiva del ensamblaje del complejo de proteínas mediado por chaperona	1	0.00526336	HDAC6
GO:0060632	Regulación del movimiento basado en microtúbulos	1	0.00526336	HDAC6
GO:0043242	Regulación negativa del desensamble del complejo proteico	1	0.00526336	HDAC6
GO:0071218	Respuesta celular a la proteína mal plegada	1	0.00526336	HDAC6
GO:0070845	Transporte de proteínas mal plegadas poliubiquitinadas	1	0.00526336	HDAC6
GO:0043241	Desensamble del complejo de proteínas	1	0.00526336	HDAC6

GO:0018394	Acetilación de peptidil-lisina	1	0.00526336	KAT2B
GO:0022037	Desarrollo del metencéfalo	1	0.00526336	KAT2A
GO:0048742	Regulación del desarrollo de la fibra muscular esquelética	1	0.00526336	HDAC4
GO:0014894	Respuesta a la denervación implicada en la regulación de la adaptación muscular	1	0.00526336	HDAC4
GO:0051569	Regulación de la metilación de lisina K4 en la histona H3	1	0.00526336	KMT2A
GO:0001555	Crecimiento de ovocitos	1	0.00526336	KMT2D
GO:0008283	Proliferación celular	3	0.00545975	MAP3K20, KAT2A, CREBBP
GO:0032355	Respuesta al estímulo de estradiol	2	0.00564496	DUSP1, PRDM2
GO:0008284	Regulación positiva de la proliferación celular	3	0.005676	KMT2D, HDAC4, JAK2
GO:0043243	Regulación positiva del desensamble del complejo proteico	1	0.00604768	MGEA5
GO:0051054	Regulación positiva del proceso metabólico del ADN	1	0.00604768	MGEA5
GO:0044154	Acetilación de la lisina K14 en la histona H3	1	0.00604768	KAT2A
GO:0002331	Exclusión alélica de células pre-B	1	0.00604768	ATM
GO:0042159	Proceso catabólico de lipoproteínas	1	0.00604768	ATM
GO:0050867	Regulación positiva de la activación celular	1	0.00604768	JAK2
GO:0042503	Fosforilación de tirosina de la proteína Stat3	1	0.00604768	JAK2
GO:0006473	Acetilación de proteínas	1	0.00604768	KAT2B
GO:0035722	Vía de señalización mediada por interleucina-12	1	0.00604768	JAK2
GO:0033194	Respuesta al hidrógeno peróxido	1	0.00604768	JAK2
GO:0042977	Activación de la actividad de la quinasa JAK2	1	0.00604768	JAK2
GO:0051788	Respuesta a la proteína mal plegada	1	0.00604768	HDAC6
GO:0043162	Proceso catabólico de proteínas dependientes de la ubiquitina a través de la vía de clasificación corporal multivesicular	1	0.00604768	HDAC6
GO:0044419	Interacción interespecies entre organismos	3	0.00615644	KAT2A, KAT2B, CREBBP
GO:0007243	Cascada de proteínas quinasas intracelular	2	0.00618777	MAP3K20, JAK2
GO:0006461	Complejo de ensamblaje de proteínas	2	0.00682667	KMT2A, CREBBP
GO:0006468	Fosforilación de proteínas	3	0.00714468	CDK17, MAP3K20, JAK2
GO:0045892	Regulación negativa de la transcripción, dependiente del ADN	3	0.00718519	HDAC4, HDAC6, HDAC5
GO:0014898	Hipertrofia del músculo cardíaco en respuesta al estrés	1	0.0072176	HDAC4
GO:0070265	Muerte celular necrótica	1	0.0072176	MGEA5
GO:0030718	Mantenimiento de células madre de la línea germinal	1	0.0072176	CREBBP
GO:0032418	Localización del lisosoma	1	0.0072176	HDAC6

GO:0006515	Proceso catabólico de proteínas mal plegado o sintetizado de forma incompleta	1	0.0072176	HDAC6
GO:0007260	Fosforilación de tirosina de la proteína STAT	1	0.0072176	JAK2
GO:0070842	Ensamblaje agresivo	1	0.0072176	HDAC6
GO:0070932	Desacetilación de histona H3	1	0.0072176	HDAC4
GO:0032731	Regulación positiva de la producción de interleucina-1 beta	1	0.00807985	JAK2
GO:0018027	Dimetilación de peptidil-lisina	1	0.00807985	SETD7
GO:0018022	Metilación de peptidil-lisina	1	0.00807985	EZH1
GO:0032411	Regulación positiva de la actividad del transportador	1	0.00807985	KMT2A
GO:0060397	Cascada JAK-STAT involucrada en la vía de señalización de la hormona del crecimiento	1	0.00807985	JAK2
GO:0045820	Regulación negativa de la glucólisis	1	0.00807985	HDAC4
GO:0007262	Importación de proteínas STAT en el núcleo	1	0.00807985	JAK2
GO:0051354	Regulación negativa de la actividad de la oxidoreductasa	1	0.00807985	HDAC6
GO:0033148	Regulación positiva de la vía de señalización del receptor de estrógeno intracelular	1	0.00807985	KMT2D
GO:0000075	Punto de control del ciclo celular	2	0.00847375	MAP3K20, ATM
GO:0046777	Autofosforilación de proteínas	2	0.00891343	JAK2, ATM
GO:0032516	Regulación positiva de la actividad fosfoproteína fosfatasa	1	0.00894737	JAK2
GO:0006516	Glicoproteína proceso catabólico	1	0.00894737	MGEA5
GO:0051901	Regulación positiva de la despolarización mitocondrial	1	0.00894737	MGEA5
GO:0018026	Monometilación de peptidil-lisina	1	0.00894737	SETD7
GO:0016236	Macroautofagia	1	0.00894737	HDAC6
GO:0006975	Daño en el ADN inducido por la fosforilación de proteínas	1	0.00894737	ATM
GO:0010830	Regulación de la diferenciación de miotubos	1	0.00894737	HDAC5
GO:0070933	Desacetilación de histona H4	1	0.00894737	HDAC4
GO:0010870	Regulación positiva del proceso biosintético de receptores	1	0.00969015	HDAC6
GO:0071480	Respuesta celular a la radiación gamma	1	0.00969015	ATM
GO:0033261	Regulación de la fase S	1	0.00969015	CREBBP
GO:0060765	Regulación de la vía de señalización del receptor de andrógenos	1	0.00969015	HDAC6
GO:0045822	Regulación negativa de la contracción del corazón	1	0.00969015	JAK2
GO:0033235	Regulación positiva de la sumoilación de proteínas	1	0.00969015	HDAC4
GO:0008219	Muerte celular	2	0.00978416	MAP3K20, ATM

Función molecular (FM)					
GO:0004402	Actividad de histona acetiltransferasa	5	5.91216E-09	KAT2A, NCOA1, KAT2B, CREBBP, MGEA5	
GO:0008168	Actividad metiltransferasa	6	1.20404E-08	KMT2D, KMT2A, EZH1, SETD1B, PRDM2, SETD7	
GO:0016740	Actividad transferasa	8	4.63567E-08	KMT2D, KMT2A, KAT2A, EZH1, SETD1B, PRDM2, CREBBP, SETD7	
GO:0005515	Unión a proteínas	15	4.65445E-08	DUSP1, KMT2D, MAP3K20, KMT2A, KAT2A, HDAC4, KAT2B, SETD1B, HDAC6, JAK2, HDAC5, NEK9, CREBBP, ATM, SETD7	
GO:0034739	Actividad de histona deacetilasa (específica para H3-K16)	3	4.96567E-07	HDAC4, HDAC6, HDAC5	
GO:0032041	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K14)	3	4.96567E-07	HDAC4, HDAC6, HDAC5	
GO:0046970	Actividad de histona deacetilasa dependiente de NAD (específica de H4-K16)	3	4.96567E-07	HDAC4, HDAC6, HDAC5	
GO:0042826	Unión de histona deacetilasa	4	5.37662E-07	KAT2A, HDAC4, KAT2B, HDAC6	
GO:0004407	Actividad de histona deacetilasa	3	6.02026E-07	HDAC4, HDAC6, HDAC5	
GO:0046969	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	3	6.02026E-07	HDAC4, HDAC6, HDAC5	
GO:0003713	Actividad coactivadora de la transcripción	5	1.89549E-06	KAT2A, NCOA1, KAT2B, MYSM1, CREBBP	
GO:0018024	Actividad de histona-lisina N-metiltransferasa	3	3.09437E-06	KMT2D, PRDM2, SETD7	
GO:0044212	Unión de ADN de la región reguladora de la transcripción	4	2.27595E-05	KMT2D, KMT2A, HDAC4, HDAC5	
GO:0003677	Unión al ADN	8	4.52217E-05	KMT2D, KMT2A, EZH1, HDAC4, MYSM1, PRDM2, CREBBP, ATM	
GO:0003682	Unión a la cromatina	4	7.18311E-05	KMT2A, KAT2A, EZH1, CREBBP	
GO:0019901	Unión de proteína quinasa	4	0.000081186	HDAC4, KAT2B, JAK2, NEK9	
GO:0042800	Actividad de histona metiltransferasa (H3-K4 específica)	2	0.000115866	KMT2A, SETD1B	
GO:0008134	Factor de transcripción de unión	4	0.000118432	HDAC4, KAT2B, HDAC5, CREBBP	
GO:0016407	Actividad acetiltransferasa	2	0.000131672	KAT2B, CREBBP	
GO:0004674	Actividad de proteína serina / treonina quinasa	4	0.000397031	CDK17, MAP3K20, NEK9, ATM	
GO:0070491	Unión al factor de represión transcripcional	2	0.00049598	HDAC4, HDAC5	
GO:0002039	Unión a p53	2	0.0015381	CREBBP, SETD7	
GO:0042393	Unión de histonas	2	0.00232202	MYSM1, JAK2	
GO:0035401	Actividad de la histona quinasa (específica de H3-Y41)	1	0.00250279	JAK2	
GO:0070840	Unión al complejo de dineína	1	0.00250279	HDAC6	

GO:0043997	Actividad de la histona acetiltransferasa (específica de H4-K12)	1	0.00250279	KAT2A
GO:0008270	Unión de iones de zinc	6	0.00335881	KMT2D, KMT2A, HDAC4, HDAC6, PRDM2, CREBBP
GO:0004468	Actividad de lisina N-acetiltransferasa	1	0.004197	KAT2B
GO:0042903	Actividad de tubulina deacetilasa	1	0.004197	HDAC6
GO:0010484	Actividad acetiltransferasa en la histona H3	1	0.004197	KAT2A
GO:0000166	Unión de nucleótidos	6	0.00450265	CDK17, MAP3K20, SETD1B, JAK2, NEK9, ATM
GO:0046872	Unión de iones metálicos	7	0.00453541	KMT2D, KMT2A, MYSM1, HDAC6, PRDM2, NEK9, CREBBP
GO:0035174	Actividad de histonas serina quinasa	1	0.00541954	ATM
GO:0051577	Unión a MyoD	1	0.00541954	CREBBP
GO:0046976	Actividad de histona metiltransferasa (H3-K27 específica)	1	0.00541954	EZH1
GO:0005524	Unión a ATP	5	0.00565721	CDK17, MAP3K20, JAK2, NEK9, ATM
GO:0045322	Unión a CpG no metilada	1	0.00684373	KMT2A
GO:0001191	Actividad represora transcripcional del ARN polimerasa II	1	0.00684373	CREBBP
GO:0016787	Actividad hidrolasa	4	0.00713639	DUSP1, HDAC4, HDAC6, HDAC5
GO:0005143	Unión al receptor de interleucina-12	1	0.00738596	JAK2
GO:0004677	Actividad de proteína quinasa dependiente de ADN	1	0.00738596	ATM
GO:0005131	Unión del receptor de la hormona del crecimiento	1	0.00738596	JAK2
GO:0033130	Unión al receptor de acetilcolina	1	0.00738596	JAK2
GO:0008330	Actividad de la proteína tirosina / treonina fosfatasa	1	0.00738596	DUSP1
GO:0016746	Actividad transferasa, transfiriendo grupos acilo	2	0.00738625	KAT2B, MGEA5
GO:0030955	Unión de iones de potasio	1	0.00779729	HDAC4
GO:0048156	unión a la proteína tau	1	0.00779729	HDAC6
GO:0004415	Actividad de la enzima condroitinasa I	1	0.00779729	MGEA5
GO:0042975	Unión al receptor activado por proliferador de peroxisoma	1	0.00779729	CREBBP
GO:0016279	Actividad de proteína-lisina N-metiltransferasa	1	0.00779729	SETD7
GO:0003680	Unión al ADN AT	1	0.00891587	KMT2A
GO:0070577	Unión de residuos de histonas acetiladas	1	0.00944574	KMT2A
GO:0035257	Unión de receptores de hormonas nucleares	1	0.00944574	NCOA1
GO:0033558	Actividad de la proteína deacetilasa	1	0.00944574	HDAC4
GO:0004726	Actividad de la proteína tirosina fosfatasa que no se extiende a la membrana	1	0.00944574	DUSP1

Componente celular (CC)				
GO:0005634	Núcleo	20	9.66881E-14	DUSP1, KMT2D, MAP3K20, KMT2A, KAT2A, EZH1, NCOA1, HDAC4, KAT2B, SETD1B, MYSM1, HDAC6, JAK2, HDAC5, PRDM2, NEK9, CREBBP, ATM, SETD7, MGEA5
GO:0005737	Citoplasma	3	3.50542E-06	KMT2D, KMT2A, SETD1B
GO:0005730	Nucleolo	3	1.24983E-05	HDAC4, HDAC6, HDAC5
GO:0017053	Complejo de represión transcripcional	2	0.00013036	HDAC4, KAT2B
GO:0005654	Nucleoplasma	2	0.000325055	KAT2B, MYSM1
GO:0005881	Microtúbulos citoplasmáticos	2	0.000338183	KAT2A, KAT2B
GO:0005694	Cromosoma	2	0.00374276	HDAC6, JAK2

Tabla 14A. Análisis de enriquecimiento GO de genes sobreexpresados en UCS utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes sobreexpresados
Proceso Biológico (PB)				
GO:0016568	Modificación de la cromatina	4	2.09563E-05	EZH2, SMYD3, EYA1, EYA2
GO:0016576	Desfosforilación de histonas	2	2.65309E-05	EYA1, EYA2
GO:0051983	Regulación de la segregación cromosómica	2	2.94695E-05	BUB1, AURKB
GO:0008283	Proliferación celular	4	3.92493E-05	BUB1, AURKB, CDK5, UHRF1
GO:0007049	Ciclo celular	4	0.000117243	PPP1CA, AURKB, CDK5, UHRF1
GO:0051301	División celular	3	0.00127354	PPP1CA, BUB1, CDK5
GO:0035335	Desfosforilación de peptidil-tirosina	2	0.00207052	EYA1, EYA2
GO:0021697	Formación de la corteza cerebelosa	1	0.00336179	CDK5
GO:0007063	Regulación de la cohesión entre cromátidas hermanas	1	0.00336179	BUB1
GO:0000087	Fase M del ciclo celular mitótico	2	0.00353857	BUB1, AURKB
GO:0000236	Prometafase mitótica	2	0.00389632	BUB1, AURKB
GO:0009790	Desarrollo embrionario	2	0.00440583	BUB1, CDK5
GO:0014706	Desarrollo del tejido muscular estriado	1	0.00448179	EYA2
GO:2000373	Regulación positiva de la topoisomerasa del ADN (ATP-hidrolizante)	1	0.00448179	UHRF1
GO:0033136	Fosforilación de serina de la proteína STAT3	1	0.00448179	CDK5
GO:0031914	Regulación negativa de la plasticidad sináptica	1	0.00448179	CDK5
GO:2000273	Regulación positiva de la actividad del receptor	1	0.00463512	CDK5
GO:0007501	Especificación del destino de las células mesodérmicas	1	0.00463512	EYA2
GO:0016570	Modificación de histonas	1	0.00463512	AURKB
GO:0051256	Formación de la zona media del huso mitótico	1	0.00463512	AURKB
GO:0021819	Formación de capas en la corteza cerebral	1	0.00463512	CDK5
GO:0045956	Regulación positiva de la exocitosis dependiente de los iones de calcio	1	0.00463512	CDK5
GO:0071156	Regulación de la detención del ciclo celular	1	0.00463512	CDK5
GO:0046826	Regulación negativa de la exportación de proteínas desde el núcleo	1	0.00463512	CDK5
GO:0021695	Desarrollo de la corteza cerebelosa	1	0.00480129	CDK5
GO:0014044	Desarrollo de células de Schwann	1	0.00480129	CDK5

GO:0071173	Punto de control del conjunto del huso	1	0.00480129	BUB1
GO:0002903	Regulación negativa de la apoptosis de células B	1	0.0049412	AURKB
GO:0045055	Vía secretora regulada	1	0.0049412	CDK5
GO:0005981	Regulación del catabolismo del glucógeno	1	0.0049412	PPP1CA
GO:0032801	Proceso catabólico del receptor	1	0.0049412	CDK5
GO:2000251	Regulación positiva de la reorganización del citoesqueleto de actina	1	0.0049412	CDK5
GO:0034501	Localización de proteínas en el cinetocoro	1	0.00516858	AURKB
GO:0008608	Acoplamiento de los microtúbulos del huso al cinetocoro	1	0.00516858	AURKB
GO:0061001	Regulación de la morfogénesis de la columna dendrítica	1	0.00516858	CDK5
GO:0048675	Extensión del axón	1	0.00516858	CDK5
GO:0022038	Desarrollo del cuerpo calloso	1	0.00516858	CDK5
GO:0046777	Autofosforilación de proteínas	2	0.00522499	AURKB, CDK5
GO:0006355	Regulación de la transcripción a nivel del ADN molde	4	0.00551247	EZH2, UHRF1, EYA1, EYA2
GO:0048148	Respuesta conductual a la cocaína	1	0.00587848	CDK5
GO:0005979	Regulación del proceso biosintético de glucógeno	1	0.00639751	PPP1CA
GO:0048387	Regulación negativa de la vía de señalización del receptor de ácido retinoico	1	0.00639751	EZH2
GO:0048488	Endocitosis de vesícula sináptica	1	0.0070289	CDK5
GO:0006281	Reparación del ADN	2	0.00744564	UHRF1, EYA2
GO:0034968	Metilación de lisina en la histona	1	0.0074618	SMYD3
GO:0001963	Transmisión sináptica, dopaminérgica	1	0.0074618	CDK5
GO:0000278	Ciclo celular mitótico	2	0.00826852	BUB1, AURKB
GO:0032270	Regulación positiva de la proteína celular del proceso metabólico	1	0.00839231	UHRF1
GO:0071777	Regulación positiva de la citocinesis del ciclo celular	1	0.00872686	AURKB
GO:0043113	Agrupamiento de receptores	1	0.00872686	CDK5
GO:0016079	Exocitosis de vesícula sináptica	1	0.00886502	CDK5
GO:0007093	Punto de control del ciclo celular mitótico	1	0.00886502	BUB1
GO:0016925	Sumoilación de proteínas	1	0.00886502	EYA1
GO:0048709	Diferenciación de oligodendrocitos	1	0.00932112	CDK5
GO:0035249	Transmisión sináptica, glutamatérgica	1	0.00958617	CDK5
GO:0021954	Desarrollo de las neuronas del sistema nervioso central	1	0.00958617	CDK5
GO:0030866	Organización del citoesqueleto de actina cortical	1	0.0098328	CDK5
GO:0031397	Regulación negativa de la ubiquitinación de proteínas	1	0.0098328	CDK5

GO:0008045	Guía del motor del axón	1	0.00989787	CDK5
GO:0019433	Proceso catabólico de triglicéridos	1	0.00989787	PPP1CA
GO:0016571	Metilación de histonas	1	0.00989787	EZH2
GO:0005977	Procesamiento metabólico del glucógeno	1	0.00991719	PPP1CA
GO:0060079	Regulación del potencial excitador de membrana postsináptica	1	0.00991719	CDK5
GO:0048167	Regulación de la plasticidad sináptica	1	0.00991719	CDK5
GO:0034644	Respuesta celular a la UV	1	0.00991719	AURKB
Función molecular (FM)				
GO:0046872	Unión de iones metálicos	2	0.000678838	EZH2, SMYD3
GO:0003677	Unión al ADN	3	0.000712609	BUB1, AURKB, CDK5
GO:0016740	Actividad transferasa	7	0.000936642	PPP1CA, BUB1, AURKB, EZH2, CDK5, UHRF1, EYA2
GO:0005524	Unión a ATP	3	0.00131647	BUB1, AURKB, CDK5
GO:0005515	Unión a proteínas	1	0.0020463	CDK5
GO:0000166	Unión de nucleótidos	2	0.0023651	EYA1, EYA2
GO:0003824	Actividad catalítica	2	0.00324998	EZH2, SMYD3
GO:0004674	Actividad de proteína serina / treonina quinasa	1	0.00340961	CDK5
GO:0000987	Unión al ADN en la región promotora proximal	5	0.00369509	PPP1CA, AURKB, SMYD3, UHRF1, EYA1
GO:0016787	Actividad hidrolasa	1	0.00464825	UHRF1
GO:0008168	Actividad metiltransferasa	1	0.00464825	CDK5
GO:0004672	Actividad de la proteína quinasa	1	0.0051124	PPP1CA
GO:0004722	Actividad de la proteína serina / treonina fosfatasa	3	0.00634365	PPP1CA, EYA1, EYA2
GO:0004725	Actividad de la proteína tirosina fosfatasa	1	0.00657049	EZH2
GO:0042393	Unión de histonas	1	0.00749325	CDK5
GO:0004721	Actividad de la fosfoproteína fosfatasa	1	0.00766255	UHRF1
GO:0050321	Actividad de la proteína quinasa tau	1	0.00901121	PPP1CA
GO:0042054	Actividad de histona metiltransferasa	2	0.00969041	EYA1, EYA2
Componente celular (CC)				
GO:0005634	Núcleo	9	2.11023E-05	PPP1CA, BUB1, AURKB, EZH2, CDK5, SMYD3, UHRF1, EYA1, EYA2
GO:0005737	Citoplasma	2	6.56786E-05	BUB1, AURKB
GO:0043005	Proyección de células neuronales	2	0.000915607	PPP1CA, CDK5
GO:0005829	Citosol	1	0.00634522	AURKB

GO:0005856	Citoesqueleto	1	0.00634522	PPP1CA
GO:0030496	Cuerpo medio	1	0.00634522	AURKB
GO:0014069	Densidad postsináptica	4	0.0069784	PPP1CA, BUB1, AURKB, CDK5
GO:0005694	Cromosoma	2	0.00700906	BUB1, AURKB
GO:0031594	Unión neuromuscular	1	0.00739691	EZH2
GO:0030426	Cono de crecimiento neural	1	0.00739886	PPP1CA
GO:0000792	Heterocromatina	1	0.00739886	CDK5
GO:0043197	Espina dendrítica	1	0.00776982	PPP1CA
GO:0030175	Filopodio	1	0.00776982	PPP1CA
GO:0030027	Lamelopodio	6	0.00782816	PPP1CA, AURKB, CDK5, SMYD3, EYA1, EYA2
GO:0005819	Huso	1	0.00832919	CDK5

Tabla 15A. Análisis de enriquecimiento GO de genes subexpresados en UCS utilizando Genecodis

Anotación	Nombre	Número de genes	p -ajustado	Genes involucrados
Proceso Biológico (PB)				
GO:0006355	Regulación de la transcripción a nivel del ADN molde	12	1.58718E-18	SETMAR, KMT2D, KMT2A, EZH1, KMT2C, SETD1B, MYSM1, HDAC6, JAK2, HDAC10, KMT2E, SETD7
GO:0007165	Transducción de señales	4	2.09608E-08	KMT2D, KMT2A, SETD1B, KMT2E
GO:0006468	Fosforilación de proteínas	5	0.000010158	DUSP1, MAP3K20, HDAC6, JAK2, ATM
GO:0009314	Respuesta a la radiación	6	0.000023262	DUSP1, MAP3K20, MAP3K8, KAT2B, NEK9, KMT2E
GO:0008283	Proliferación celular	9	2.67971E-05	KMT2D, KMT2A, EZH1, KMT2C, SETD1B, HDAC10, RPS6KA5, CREBBP, SETD7
GO:0044419	Interacción interespecies entre organismos	2	5.85692E-05	KAT2B, CREBBP
GO:0045944	Regulación positiva de la transcripción del promotor del ARN polimerasa II	4	6.25836E-05	MAP3K20, KAT2B, ATM, KMT2E
GO:0043388	Regulación positiva de la unión del ADN	5	0.000199413	CDK17, MAP3K20, MAP3K8, JAK2, RPS6KA5
GO:0045087	Respuesta inmune innata	2	0.000713506	HDAC6, HDAC10
GO:0000122	Regulación negativa de la transcripción del promotor de la ARN polimerasa II	2	0.000770271	KAT2A, MYSM1
GO:0043065	Regulación positiva del proceso apoptótico	5	0.00079949	KMT2D, KMT2A, KAT2B, RPS6KA5, CREBBP
GO:0032516	Regulación positiva de la actividad de la fosfoproteína fosfatasa	3	0.0008424	MAP3K20, JAK2, RPS6KA5
GO:0008284	Regulación positiva de la proliferación celular	2	0.00142142	KMT2A, KMT2E
GO:0048008	Vía de señalización del receptor del factor de crecimiento derivado de plaquetas	2	0.00142142	HDAC6, HDAC10
GO:0018108	Fosforilación de peptidil-tirosina	2	0.00251422	MAP3K20, ATM
GO:0046777	Autofosforilación de proteínas	2	0.00294093	JAK2, KMT2E
GO:0014068	Regulación positiva de la señalización de fosfatidilinositol 3-quinasa	2	0.00359379	JAK2, MGEA5
GO:0001701	Desarrollo embrionario en el útero	4	0.0037512	KMT2A, MYSM1, CREBBP, KMT2E
GO:0006306	Metilación del ADN	2	0.00427652	KAT2A, KAT2B
GO:0010212	Respuesta a la radiación ionizante	2	0.00427652	KAT2A, ATM
GO:0045892	Regulación negativa de la transcripción a nivel del ADN molde	2	0.00472665	MAP3K20, JAK2

GO:0008152	Proceso metabólico	1	0.00479497	MGEA5
GO:0010467	Expresión génica	1	0.00479497	MGEA5
GO:0008285	Regulación negativa de la proliferación celular	1	0.00479497	MGEA5
GO:0009967	Regulación positiva de la transducción de señales	1	0.00479497	RPS6KA5
GO:0006366	Transcripción del promotor de la ARN polimerasa II	1	0.00479497	JAK2
GO:0045893	Regulación positiva de la transcripción a nivel del ADN molde	1	0.00479497	JAK2
GO:0007049	Ciclo celular	1	0.00479497	JAK2
GO:0051591	Respuesta al AMPc	1	0.00479497	HDAC6
GO:0032869	Respuesta celular al estímulo de insulina	1	0.00479497	HDAC6
GO:0001666	Respuesta a la hipoxia	1	0.00479497	HDAC6
GO:0051592	Respuesta al ion calcio	1	0.00479497	HDAC6
GO:0048545	Respuesta al estímulo de la hormona esteroide	1	0.00479497	KMT2A
GO:0030901	Desarrollo del cerebro medio	1	0.00479497	KMT2A
GO:0006281	Reparación de ADN	1	0.00479497	SETMAR
GO:0007243	Cascada de proteína quinasa intracelular	1	0.00479497	SETMAR
GO:0000186	Activación de la quinasa MAPKK	1	0.00479497	SETMAR
GO:0006950	Respuesta al estrés	1	0.00617458	JAK2
GO:0007050	Detención del ciclo celular	1	0.00617458	MGEA5
GO:0051403	Cascada MAPK activada por estrés	1	0.00617458	JAK2
GO:0042981	Regulación del proceso apoptótico	1	0.00617458	JAK2
GO:0010811	Regulación positiva de la adhesión del sustrato celular	1	0.00617458	RPS6KA5
GO:0050731	Regulación positiva de la fosforilación de peptidil-tirosina	1	0.00617458	RPS6KA5
GO:0006302	Reparación de rotura de doble cadena	1	0.00617458	HDAC6
GO:0035556	Transducción de señales intracelulares	1	0.00617458	HDAC6
GO:0032024	Regulación positiva de la secreción de insulina	1	0.00617458	HDAC6
GO:0060399	Regulación positiva de la vía de señalización del receptor de la hormona del crecimiento	1	0.00617458	HDAC6
GO:0060124	Regulación positiva de la secreción de hormona de crecimiento	1	0.00617458	HDAC6
GO:0009755	Vía de señalización mediada por hormonas	1	0.00617458	HDAC6
GO:0000075	Punto de control del ciclo celular	1	0.00617458	KAT2B
GO:0030041	Polimerización de filamentos de actina	1	0.00617458	KAT2A
GO:0033209	Vía de señalización mediada por el factor de necrosis tumoral	1	0.00617458	KMT2A
GO:0032870	Respuesta celular al estímulo hormonal	1	0.00617458	KMT2D
GO:0048384	Vía de señalización del receptor de ácido retinoico	1	0.00617458	SETMAR

GO:0018105	Fosforilación de peptidil-serina	3	0.00677341	KMT2A, KAT2B, JAK2
GO:0006612	Proteínas dirigidas a la membrana	1	0.0073821	MGEA5
GO:0045862	Regulación positiva de la proteólisis	1	0.0073821	MGEA5
GO:0010616	Regulación negativa de la adaptación del músculo cardíaco (PA).	1	0.0073821	KAT2A
GO:0060051	Regulación negativa de la glicosilación de proteínas	1	0.0073821	ATM
GO:0046060	Proceso metabólico de dATP	1	0.0073821	ATM
GO:0043243	Regulación positiva del desensamble del complejo proteico	1	0.0073821	JAK2
GO:0031343	Regulación positiva de la muerte celular	1	0.0073821	JAK2
GO:0006516	Glicoproteína proceso catabólico	1	0.0073821	KAT2B
GO:0006044	Proceso metabólico de la N-acetilglucosamina	1	0.0073821	JAK2
GO:0051901	Regulación positiva de la despolarización mitocondrial	1	0.0073821	JAK2
GO:0051054	Regulación positiva del proceso metabólico del ADN	1	0.0073821	JAK2
GO:0046326	Regulación positiva de la importación de glucosa	1	0.0073821	HDAC6
GO:0070265	Muerte celular necrótica	1	0.0073821	HDAC6
GO:0010524	Regulación positiva del transporte de iones de calcio en el citosol	1	0.0073821	SETMAR
GO:0016568	Modificación de la cromatina	3	0.00742507	DUSP1, KMT2C, JAK2
GO:0008219	Muerte celular	2	0.00746202	KAT2A, KAT2B
GO:0032760	Regulación positiva de la producción del factor de necrosis tumoral	3	0.00765942	JAK2, RPS6KA5, CREBBP
GO:0007257	Activación de la actividad de la quinasa JUN	3	0.007683	MAP3K20, KAT2A, CREBBP
GO:0021537	Desarrollo del telencéfalo	3	0.0085949	HDAC6, HDAC10, RPS6KA5
GO:0001706	Formación de endodermo	3	0.00865738	KAT2A, KAT2B, CREBBP
GO:0050727	Regulación de la respuesta inflamatoria	1	0.00866824	MGEA5
GO:0018026	Monometilación de peptidil-lisina	1	0.00866824	CREBBP
GO:0034612	Respuesta al factor de necrosis tumoral	1	0.00866824	HDAC6
GO:0006461	Conjunto de proteínas	1	0.00866824	HDAC6
GO:0008589	Regulación de la vía de señalización de la proteína <i>smoothened</i>	1	0.00866824	JAK2
GO:0045597	Regulación positiva de la diferenciación celular	1	0.00866824	HDAC6
GO:0044154	Acetilación de lisina K14 en la histona H3	1	0.00866824	HDAC6
GO:0032526	Respuesta al ácido retinoico	2	0.00900329	KMT2A, CREBBP
GO:0048477	Oogénesis	3	0.00928809	EZH1, HDAC10, CREBBP
GO:0033574	Respuesta al estímulo de la testosterona	1	0.00957761	KMT2E
GO:0030854	Regulación positiva de la diferenciación de granulocitos	1	0.00957761	JAK2
GO:0043392	Regulación negativa de la unión al ADN	1	0.00957761	SETD7

GO:0042542	Respuesta al peróxido de hidrógeno	1	0.00957761	RPS6KA5
GO:0006338	Remodelación de la cromatina	1	0.00957761	EZH1
GO:0042517	Regulación positiva de la fosforilación de tirosina de la proteína Stat3 (	1	0.00957761	KMT2A
GO:0050729	Regulación positiva de la respuesta inflamatoria	1	0.00957761	JAK2
GO:0042523	Regulación positiva de la fosforilación de tirosina de la proteína Stat5	1	0.00957761	RPS6KA5
GO:0007498	Desarrollo del mesodermo	1	0.00957761	JAK2
GO:0001843	Cierre del tubo neural	1	0.00957761	HDAC6
GO:0014003	Desarrollo de oligodendrocitos	1	0.00957761	KMT2D
GO:0010033	Respuesta a sustancia orgánica	2	0.00986051	MAP3K20, ATM
Función molecular (FM)				
GO:0008168	Actividad metiltransferasa	8	3.2661E-12	SETMAR, KMT2D, KMT2A, EZH1, KMT2C, SETD1B, KMT2E, SETD7
GO:0005515	Unión a proteínas	19	4.0752E-12	SETMAR, DUSP1, KMT2D, MAP3K20, MAP3K8, KMT2A, KAT2A, KMT2C, KAT2B, SETD1B, HDAC6, JAK2, HDAC10, RPS6KA5, NEK9, CREBBP, ATM, KMT2E, SETD7
GO:0016740	Actividad transferasa	10	1.06108E-10	SETMAR, KMT2D, KMT2A, KAT2A, EZH1, KMT2C, SETD1B, CREBBP, KMT2E, SETD7
GO:0042800	Actividad de histona metiltransferasa (H3-K4 específica)	4	1.31494E-09	KMT2A, KMT2C, SETD1B, KMT2E
GO:0004402	Actividad de histona acetiltransferasa	4	4.26435E-07	KAT2A, KAT2B, CREBBP, MGEA5
GO:0042826	Unión de histona deacetilasa	4	1.02808E-06	KAT2A, KAT2B, HDAC6, HDAC10
GO:0004674	Actividad de proteína serina / treonina quinasa	6	2.18308E-06	CDK17, MAP3K20, MAP3K8, RPS6KA5, NEK9, ATM
GO:0003713	Actividad coactivadora de transcripción	5	4.14411E-06	KAT2A, KAT2B, MYSM1, CREBBP, KMT2E
GO:0018024	Actividad de histona-lisina N-metiltransferasa	3	5.3337E-06	SETMAR, KMT2D, SETD7
GO:0003677	Unión al ADN	8	0.000134962	SETMAR, KMT2D, KMT2A, EZH1, KMT2C, MYSM1, CREBBP, ATM
GO:0003682	Unión a la cromatina	4	0.000139367	KMT2A, KAT2A, EZH1, CREBBP
GO:0034739	Actividad de histona deacetilasa (específica para H3-K16)	2	0.000174333	HDAC6, HDAC10
GO:0032041	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K14)	2	0.000174333	HDAC6, HDAC10

GO:0046970	Actividad de histona deacetilasa dependiente de NAD (específica de H4-K16)	2	0.000174333	HDAC6, HDAC10
GO:0004407	Actividad de histona deacetilasa (MF)	2	0.000192137	HDAC6, HDAC10
GO:0046969	Actividad de histona deacetilasa dependiente de NAD (específica de H3-K9)	2	0.000192137	HDAC6, HDAC10
GO:0016407	Actividad acetiltransferasa	2	0.000195173	KAT2B, CREBBP
GO:0005524	Unión a ATP	7	0.000209756	CDK17, MAP3K20, MAP3K8, JAK2, RPS6KA5, NEK9, ATM
GO:0004672	Actividad de la proteína quinasa	4	0.000215429	CDK17, MAP3K8, JAK2, RPS6KA5
GO:0000166	Unión de nucleótidos	8	0.000233727	CDK17, MAP3K20, MAP3K8, SETD1B, JAK2, RPS6KA5, NEK9, ATM
GO:0046872	Unión de iones metálicos	9	0.000281326	SETMAR, KMT2D, KMT2A, KMT2C, MYSM1, HDAC6, NEK9, CREBBP, KMT2E
GO:0004709	Actividad de la MAP quinasa quinasa quinasa	2	0.00030773	MAP3K20, MAP3K8
GO:0008270	Unión de iones de zinc	7	0.000915521	SETMAR, KMT2D, KMT2A, KMT2C, HDAC6, CREBBP, KMT2E
GO:0000287	Unión de iones de magnesio	3	0.000975185	MAP3K20, MAP3K8, RPS6KA5
GO:0019899	Unión de enzimas	3	0.0011164	HDAC6, HDAC10, KMT2E
GO:0002039	Unión a p53	2	0.00152116	CREBBP, SETD7
GO:0016772	Actividad transferasa, transfiriendo grupos que contienen fósforo	3	0.00161199	CDK17, RPS6KA5, NEK9
GO:0019901	Unión de proteína quinasa	3	0.0019308	KAT2B, JAK2, NEK9
GO:0042393	Unión de histonas	2	0.00215154	MYSM1, JAK2
GO:0035401	Actividad de la histona quinasa (específica de H3-Y41)	1	0.00216415	JAK2
GO:0070840	Unión al complejo de dineína	1	0.00216415	HDAC6
GO:0043997	Actividad de la histona acetiltransferasa (específica de H4-K12)	1	0.00216415	KAT2A
GO:0004468	Actividad de lisina N-acetiltransferasa	1	0.00395603	KAT2B
GO:0042903	Actividad de tubulina deacetilasa	1	0.00395603	HDAC6
GO:0010484	Actividad acetiltransferasa de la histona H3	1	0.00395603	KAT2A
GO:0035174	Actividad de histonas serina quinasa	1	0.00532371	ATM
GO:0051577	Unión a MyoD	1	0.00532371	CREBBP
GO:0046976	Actividad de histona metiltransferasa (H3-K27 específica)	1	0.00532371	EZH1
GO:0004803	Actividad transposasa	1	0.00532371	SETMAR
GO:0045322	Unión a CpG no metilada	1	0.00674986	KMT2A
GO:0001191	Actividad represora transcripcional del ARN polimerasa II	1	0.00674986	CREBBP

GO:0005143	Unión al receptor de interleucina-12	1	0.0075178	JAK2
GO:0004677	Actividad de proteína quinasa dependiente de ADN	1	0.0075178	ATM
GO:0005131	Unión del receptor de la hormona del crecimiento	1	0.0075178	JAK2
GO:0033130	Unión al receptor de acetilcolina	1	0.0075178	JAK2
GO:0008330	Actividad de la proteína tirosina / treonina fosfatasa	1	0.0075178	DUSP1
GO:0048156	Unión a la proteína tau	1	0.00797787	HDAC6
GO:0004415	Actividad de la enzima condroitinasa I	1	0.00797787	MGEA5
GO:0042975	Unión a receptor activado por proliferador de peroxisoma	1	0.00797787	CREBBP
GO:0016279	Actividad de proteína-lisina N-metiltransferasa	1	0.00797787	SETD7
GO:0016746	Actividad transferasa, transfiriendo grupos acilo	2	0.00806436	KAT2B, MGEA5
GO:0016787	Actividad hidrolasa	4	0.00813908	SETMAR, DUSP1, HDAC6, HDAC10
GO:0003680	Unión al ADN AT	1	0.00895991	KMT2A
GO:0043566	Unión al ADN de estructura específica	1	0.00895991	SETMAR
GO:0070577	Unión de residuos de histonas acetiladas	1	0.00953064	KMT2A
GO:0033558	Actividad de la proteína deacetilasa	1	0.00953064	HDAC10
GO:0004726	Actividad de la proteína tirosina fosfatasa que no se extiende a la membrana	1	0.00953064	DUSP1
Componente celular (CC)				
GO:0005634	Núcleo	21	1.5369E-13	SETMAR, DUSP1, KMT2D, MAP3K20, KMT2A, KAT2A, EZH1, KMT2C, KAT2B, SETD1B, MYSM1, HDAC6, JAK2, HDAC10, RPS6KA5, NEK9, CREBBP, ATM, KMT2E, SETD7, MGEA5
GO:0035097	Complejo de histona metiltransferasa	4	9.3664E-09	KMT2D, KMT2A, KMT2C, SETD1B
GO:0005654	Nucleoplasma	6	0.000360518	DUSP1, KAT2B, HDAC10, RPS6KA5, CREBBP, ATM
GO:0016585	Complejo de remodelación de cromatina	2	0.000460666	KAT2B, MYSM1
GO:0005671	Complejo activador de la transcripción Ada2 / Gcn5 / Ada3	2	0.000499259	KAT2A, KAT2B
GO:0000118	Complejo de histona deacetilasa	2	0.00147589	HDAC6, HDAC10
GO:0005901	Caveola	2	0.00441348	HDAC6, JAK2
GO:0005694	Cromosoma	3	0.00462691	SETMAR, SETD1B, SETD7
GO:0005737	Citoplasma	10	0.00730941	MAP3K20, MAP3K8, HDAC6, JAK2, HDAC10, RPS6KA5, NEK9, CREBBP, ATM, MGEA5

ANEXO 5

Tabla 16A. Genes metilados diferencialmente a nivel del promotor, para CESC

Símbolo del gen	Transcrito	Tumor n=212		Normal n=19		p-valor	$\Delta\beta$	Nivel de Metilación	DEG
		Mediana β valor	IQR	Mediana β valor	IQR				
HLCS	NM_000411	0.09735	0.055	0.4479	0.0469	5.66214E-15	-0.35055	Hipometilado	No
KAT2A	NM_021078	0.35015	0.17708	0.7926	0.0221	2.9865E-14	-0.44245	Hipometilado	No
LIMK2	NM_001031801	0.3976	0.07948	0.5771	0.0254	4.01928E-11	-0.1795	Hipometilado	No
LIMK2	NM_016733	0.3976	0.07948	0.5771	0.0254	4.01928E-11	-0.1795	Hipometilado	No
CDYL	NR_026590	0.83825	0.07818	0.4772	0.0357	7.23865E-14	0.36105	Hipermetilado	No
DNMT3B	NM_175850	0.64995	0.16508	0.3006	0.0254	8.44325E-13	0.34935	Hipermetilado	No
EYA1*	NM_172058	0.45855	0.29813	0.1922	0.0698	1.22231E-10	0.26635	Hipermetilado	Si
EYA1*	NM_172059	0.45855	0.29813	0.1922	0.0698	1.22231E-10	0.26635	Hipermetilado	Si
EYA1*	NM_172060	0.45855	0.29813	0.1922	0.0698	1.22231E-10	0.26635	Hipermetilado	Si
HDAC9	NM_001204145	0.22405	0.20995	0.0621	0.0115	1.19904E-14	0.16195	Hipermetilado	No
HDAC9	NM_001204146	0.22405	0.20995	0.0621	0.0115	1.19904E-14	0.16195	Hipermetilado	No
HDAC9	NM_014707	0.22405	0.20995	0.0621	0.0115	1.19904E-14	0.16195	Hipermetilado	No
HDAC9	NM_058176	0.22405	0.20995	0.0621	0.0115	1.19904E-14	0.16195	Hipermetilado	No
HDAC9	NM_178423	0.22405	0.20995	0.0621	0.0115	1.19904E-14	0.16195	Hipermetilado	No
HDAC9	NM_178425	0.22405	0.20995	0.0621	0.0115	1.19904E-14	0.16195	Hipermetilado	No
NEK6	NM_001166170	0.59475	0.1792	0.054	0.0161	-6.43929E-15	0.54075	Hipermetilado	No
NEK6	NM_001166171	0.59475	0.1792	0.054	0.0161	-6.43929E-15	0.54075	Hipermetilado	No
NEK6	NM_001166169	0.6923	0.13028	0.0711	0.021	3.05311E-14	0.6212	Hipermetilado	No
PADI4	NM_012387	0.48025	0.15813	0.2289	0.0367	1.7782E-08	0.25135	Hipermetilado	No
PKN1	NM_213560	0.71585	0.16698	0.0986	0.0418	4.77396E-14	0.61725	Hipermetilado	No
PPP2CB	NM_001009552	0.21575	0.10023	0.0636	0.0132	1.56649E-07	0.15215	Hipermetilado	No
PRDM2	NM_012231	0.9172	0.0374	0.4113	0.0312	3.55271E-14	0.5059	Hipermetilado	No
PRDM2	NM_015866	0.9172	0.0374	0.4113	0.0312	3.55271E-14	0.5059	Hipermetilado	No
PRDM9	NM_020227	0.79905	0.1213	0.615	0.041	6.42819E-14	0.18405	Hipermetilado	No
PRKCB	NM_002738	0.45665	0.33663	0.0596	0.008	1.66533E-15	0.39705	Hipermetilado	No
PRKCB	NM_212535	0.45665	0.33663	0.0596	0.008	1.66533E-15	0.39705	Hipermetilado	No
PRMT8	NM_019854	0.41575	0.23338	0.152	0.0462	1.33227E-14	0.26375	Hipermetilado	No
RAG1	NM_000448	0.7958	0.17953	0.178	0.0515	1.5532E-13	0.6178	Hipermetilado	No
RNF20	NM_019592	0.5899	0.05205	0.3816	0.0428	-1.02141E-14	0.2083	Hipermetilado	No
SETDB2	NM_001160308	0.8222	0.0924	0.5724	0.0787	5.31797E-14	0.2498	Hipermetilado	No
TRDMT1	NM_004412	0.4011	0.09095	0.251	0.0128	8.32667E-15	0.1501	Hipermetilado	No

IQR:Rango intercuartil , DEG: Gen con expresión diferencial

$\Delta\beta < -0.15$ "Hipometilada"; $\Delta\beta > 0.15$ "Hipermetilado", valor de p fue calculado mediante T-Test.

*Genes resaltados en azul (subexpresados).

Tabla 17A. Genes metilados diferencialmente a nivel del promotor, para UCEC

Símbolo del gen	Transcrito	Tumor (n = 424)		Normal (n = 67)		<i>p</i> -valor	$\Delta\beta$	Nivel de Metilación	DEG
		Mediana β valor	IQR	Mediana β valor	IQR				
HDAC9	NM_001204147	0.15395	0.12975	0.6501	0.6147	2.32148E-13	-0.496	Hipometilado	No
HDAC9	NM_001204148	0.15395	0.12975	0.6501	0.6147	2.32148E-13	-0.496	Hipometilado	No
NEK6*	NM_001166167	0.13505	0.085675	0.3462	0.0931	9.32587E-15	-0.211	Hipometilado	Si
NEK6*	NM_001145001	0.1157	0.06935	0.2837	0.0757	2.55351E-14	-0.168	Hipometilado	Si
NEK6*	NM_014397	0.1157	0.06935	0.2837	0.0757	2.55351E-14	-0.168	Hipometilado	Si
PRKCB	NM_002738	0.2484	0.430125	0.0645	0.0445	4.70735E-14	0.184	Hipermetilado	No
PRKCB	NM_212535	0.2484	0.430125	0.0645	0.0445	4.70735E-14	0.184	Hipermetilado	No
PRMT8	NM_019854	0.28225	0.335625	0.1207	0.0858	8.89289E-14	0.162	Hipermetilado	No
PRMT8	NM_001256536	0.5306	0.30745	0.3156	0.3684	1.25754E-10	0.215	Hipermetilado	No

IQR: Rango intercuartil, DEG: Gen con expresión diferencial.

$\Delta\beta < -0.15$ "Hipometilada"; $\Delta\beta > 0.15$ "Hipermetilado", valor de *p* fue calculado mediante T-Test.

*Genes resaltados en rojo (sobrexpresados).

Tabla 18 A. Correlación entre la expresión de genes y el nivel de metilación de ADN en el promotor, para OV

Gen	Sonda	Región Promotora	Pearson r (<i>p</i> -valor) *	Tipo de Correlación
NCOA3	cg12128876	5'UTR	-0.653***	Moderada negativa
ATM	cg01800843	TSS ± 1500	-0.386*	Baja negativa
TRDMT1	cg19579782	TSS ± 200	-0.339***	Baja negativa
KAT2B	cg18053505	TSS ± 1500	-0.303***	Baja negativa
EYA3	cg27498304	TSS ± 200	-0.288***	Débil negativa
EYA3	cg22760576	5'UTR	-0.267***	Débil negativa
EP300	cg00500400	TSS ± 1500	-0.26***	Débil negativa
SETD1B	cg11608114	TSS ± 1500	-0.236***	Débil negativa
PPP1CC	cg04375036	TSS ± 1500	-0.235***	Débil negativa
STK4	cg27194614	TSS ± 1500	-0.224***	Débil negativa
NEK9	cg08847533	TSS ± 200	-0.223***	Débil negativa
DNMT3A	cg21629895	TSS ± 1500	-0.221***	Débil negativa
ATM	cg03165700	TSS ± 1500	-0.215**	Débil negativa
HDAC9	cg12081743	TSS ± 1500, 5'UTR	-0.214***	Débil negativa
PPM1D	cg06090421	TSS ± 1500	-0.213***	Débil negativa
KDM5B	cg14560895	TSS ± 200	-0.212***	Débil negativa
SETDB2	cg18393014	TSS ± 200	-0.209**	Débil negativa
HDAC10	cg04586023	TSS ± 1500	-0.208***	Débil negativa
HDAC10	cg01120165	TSS ± 1500	-0.206***	Débil negativa
ASH1L	cg17038116	TSS ± 1500	-0.205**	Débil negativa
ATR	cg20636078	TSS ± 1500	-0.204**	Débil negativa
EZH1	cg06975499	5'UTR	-0.203**	Débil negativa
KDM4A	cg08552254	5'UTR	-0.192**	Débil negativa
SETD1B	cg21968169	TSS ± 1500	-0.187**	Débil negativa
HDAC6	cg24616736	TSS ± 1500	-0.185**	Débil negativa
EZH1	cg27088844	TSS ± 200	-0.184**	Débil negativa
RPS6KA5	cg25398949	TSS ± 1500	-0.179**	Débil negativa
ATM	cg22477592	5'UTR	-0.178**	Débil negativa
TLK1	cg13882835	5'UTR, TSS ± 1500	-0.17*	Débil negativa
KDM4B	cg09310112	5'UTR	-0.169*	Débil negativa
HDAC11	cg27025856	TSS ± 1500	-0.163*	Débil negativa
UBR2	cg19448829	TSS ± 1500	-0.16*	Débil negativa
NSD1	cg11316784	TSS ± 1500, 5'UTR	-0.158*	Débil negativa
HDAC6	cg10783042	5'UTR	-0.157*	Débil negativa
RPS6KA5	cg02904235	TSS ± 1500	-0.156**	Débil negativa
LIMK2	cg14086647	TSS ± 1500	-0.154*	Débil negativa
EYA2	cg09306675	5'UTR	-0.152*	Débil negativa
PRKCD	cg13908523	TSS ± 1500	-0.148*	Débil negativa
TET1	cg21948169	TSS ± 200	-0.146*	Débil negativa
HDAC11	cg01712016	TSS ± 200, TSS ± 1500	-0.138*	Débil negativa
NAT10	cg06513075	TSS ± 1500	-0.134*	Débil negativa
GTF3C4	cg20603424	TSS ± 200	-0.129*	Débil negativa
CLOCK	cg04485603	5'UTR	0.173*	Débil positiva
PRDM2	cg12379145	TSS ± 1500, 5'UTR	0.189**	Débil positiva
PRDM2	cg25402049	TSS ± 1500, 5'UTR	0.207**	Débil positiva
PRMT8	cg21667836	TSS ± 1500	0.29***	Débil positiva
RAG1	cg11494699	TSS ± 1500	0.294***	Débil positiva

****p* < 0.001 ***p* < 0.01 **p* < 0.05

TSS: sitio de inicio de la transcripción

TSS ± 1500: sitio de inicio de la transcripción desde 200 hasta 1500 nt aguas arriba o aguas abajo del TSS,

TSS ± 200: sitio de inicio de la transcripción hasta 200 nt aguas arriba o aguas abajo del TSS

5'UTR: región no traducida 5'

Tabla 19 A. Correlación entre la expresión de genes y el nivel de metilación de ADN en el promotor, para UCS

Gen	Sonda	Región promotora	Pearson r (p-valor) *	Tipo de Correlación
NEK9	cg01952931	TSS ± 200	-0.859 **	Alta negativa
NEK9	cg13494334	TSS ± 200	-0.848 ***	Alta negativa
NEK9	cg01992424	TSS ± 200	-0.843 ***	Alta negativa
NEK9	cg11812906	TSS ± 200	-0.842 **	Alta negativa
NEK9	cg09944325	TSS ± 200	-0.837 ***	Alta negativa
NEK9	cg20071051	TSS ± 200	-0.829 **	Alta negativa
NEK9	cg08847533	TSS ± 200	-0.827 ***	Alta negativa
NEK9	cg09945151	TSS ± 1500	-0.796 ***	Alta negativa
NEK9	cg14426084	TSS ± 1500	-0.794 **	Alta negativa
SETDB2	cg15906151	TSS ± 1500	-0.743 **	Alta negativa
SETDB2	cg18393014	TSS ± 200	-0.736 **	Alta negativa
SETDB2	cg18656860	TSS ± 1500	-0.735 **	Alta negativa
SETDB2	cg01928147	TSS ± 200	-0.735 ***	Alta negativa
SETDB2	cg00919994	TSS ± 200	-0.726 **	Alta negativa
SETDB2	cg11973886	TSS ± 200	-0.723 **	Alta negativa
SETDB2	cg06870750	TSS ± 1500	-0.708 **	Alta negativa
SETDB2	cg25756590	TSS ± 200	-0.706 ***	Alta negativa
SETDB2	cg16543390	TSS ± 1500	-0.700 **	Moderada negativa
SETDB2	cg04914982	TSS ± 200	-0.695 **	Moderada negativa
SETD1A	cg04033451	TSS ± 200	-0.672 **	Moderada negativa
MASTL	cg11609007	TSS ± 1500	-0.660 **	Moderada negativa
CARM1	cg17918830	TSS ± 200	-0.658 **	Moderada negativa
SIRT2	cg07375277	TSS ± 200	-0.611 **	Moderada negativa
ATM	cg27599391	TSS ± 1500	-0.610 ***	Moderada negativa
KDM4D	cg14963860	TSS ± 1500	-0.604 ***	Moderada negativa
SUV39H2	cg14487075	5'UTR	-0.597 ***	Moderada negativa
SETD1A	cg05102288	TSS ± 200	-0.595 ***	Moderada negativa
KAT5	cg05785768	TSS ± 200	-0.594 **	Moderada negativa
ATM	cg03569616	TSS ± 1500	-0.592 ***	Moderada negativa
ATM	cg05033322	TSS ± 1500	-0.580 **	Moderada negativa
ATM	cg17420659	TSS ± 1500	-0.574 **	Moderada negativa
KAT5	cg07564927	TSS ± 200	-0.564 **	Moderada negativa
ATM	cg06081482	TSS ± 1500	-0.553 **	Moderada negativa
KAT2B	cg18053505	TSS ± 1500	-0.552 ***	Moderada negativa
SETD2	cg25849265	TSS ± 200	-0.552 ***	Moderada negativa
ATM	cg24568245	TSS ± 1500	-0.546 **	Moderada negativa
KDM5B	cg02624558	TSS ± 200	-0.546 ***	Moderada negativa
CARM1	cg22378614	TSS ± 1500	-0.543 ***	Moderada negativa
HDAC6	cg07674520	TSS ± 1500	-0.537 ***	Moderada negativa
PAK2	cg06284169	TSS ± 1500	-0.536 ***	Moderada negativa
SETD1B	cg22748630	TSS ± 1500	-0.536 ***	Moderada negativa
RPS6KA4	cg16332631	TSS ± 1500	-0.535 **	Moderada negativa
SIRT2	cg14074741	TSS ± 1500	-0.535 **	Moderada negativa
KDM5B	cg09889216	TSS ± 200	-0.534 ***	Moderada negativa
EZH1	cg05979501	TSS ± 200	-0.533 ***	Moderada negativa
ATM	cg03165700	TSS ± 1500	-0.529 **	Moderada negativa
ATM	cg10610482	TSS ± 1500	-0.526 **	Moderada negativa
ATM	cg10976287	TSS ± 1500	-0.525 **	Moderada negativa
KAT5	cg00582417	TSS ± 200	-0.525 **	Moderada negativa
CDK17	cg20697958	TSS ± 1500	-0.524 ***	Moderada negativa
ATM	cg12019961	TSS ± 1500	-0.523 ***	Moderada negativa
SETD2	cg09563240	TSS ± 200	-0.520 ***	Moderada negativa
SUV39H2	cg17621642	TSS ± 200	-0.519 ***	Moderada negativa
RPS6KA5	cg10112022	TSS ± 200	-0.518 ***	Moderada negativa
ATM	cg09808176	TSS ± 1500	-0.517 ***	Moderada negativa
ATM	cg17114932	TSS ± 1500	-0.513 **	Moderada negativa
EZH1	cg05375741	TSS ± 200	-0.513 ***	Moderada negativa
SETD1B	cg20594607	TSS ± 1500	-0.508 **	Moderada negativa
CARM1	cg26504915	TSS ± 1500	-0.506 ***	Moderada negativa

CDK17	cg16563660	5'UTR	-0.506 ***	Moderada negativa
ATM	cg23275463	TSS ± 1500	-0.503 **	Moderada negativa
RNF8	cg14425796	TSS ± 200	-0.502 **	Moderada negativa
EYA2	cg13061637	TSS ± 200	-0.500 ***	Baja negativa
CDK17	cg25056994	5'UTR	-0.495 ***	Baja negativa
SETD1B	cg01493758	TSS ± 1500	-0.492 **	Baja negativa
CHUK	cg26847029	TSS ± 200	-0.490 **	Baja negativa
ATM	cg01800843	TSS ± 1500	-0.488 **	Baja negativa
TET1	cg12630147	5'UTR	-0.487 **	Baja negativa
KDM4D	cg13474527	TSS ± 200	-0.478 ***	Baja negativa
SIRT1	cg03211233	TSS ± 1500	-0.477 **	Baja negativa
SETD1B	cg26483661	TSS ± 1500	-0.473 **	Baja negativa
EZH1	cg03623386	TSS ± 200	-0.472 **	Baja negativa
KDM4D	cg16384862	TSS ± 1500	-0.470 ***	Baja negativa
KAT5	cg14726348	TSS ± 200	-0.468 **	Baja negativa
SETD1B	cg20978010	TSS ± 1500	-0.468 **	Baja negativa
SETD1A	cg02153299	TSS ± 200	-0.464 **	Baja negativa
TET1	cg02774862	5'UTR	-0.461 **	Baja negativa
SIRT1	cg02720207	TSS ± 1500, TSS ± 200	-0.457 **	Baja negativa
PAK2	cg07061775	TSS ± 1500	-0.454 **	Baja negativa
PPP1CC	cg04375036	TSS ± 1500	-0.454 **	Baja negativa
SETDB1	cg19689029	TSS ± 200	-0.450 **	Baja negativa
KDM3A	cg20191795	TSS ± 1500, TSS ± 200	-0.449 *	Baja negativa
ATM	cg25918541	TSS ± 1500	-0.446 **	Baja negativa
TET1	cg06767766	5'UTR	-0.445 **	Baja negativa
EHMT1	cg13982446	TSS ± 1500	-0.444 **	Baja negativa
EZH1	cg12812489	5'UTR	-0.441 **	Baja negativa
MASTL	cg23635883	TSS ± 1500	-0.433 **	Baja negativa
KDM2A	cg05295919	TSS ± 1500	-0.432 **	Baja negativa
EZH1	cg00728999	TSS ± 200	-0.431 **	Baja negativa
SIRT2	cg18330372	TSS ± 1500	-0.431 **	Baja negativa
SETDB1	cg02060408	TSS ± 200	-0.430 **	Baja negativa
KDM4D	cg05713782	TSS ± 200	-0.427 **	Baja negativa
CHUK	cg07793810	TSS ± 200	-0.425 **	Baja negativa
SETD2	cg07817810	TSS ± 1500	-0.424 **	Baja negativa
DNMT3A	cg00220517	5'UTR	-0.420 *	Baja negativa
PRMT2	cg10187059	TSS ± 200	-0.420 *	Baja negativa
TET1	cg13848707	5'UTR	-0.418 *	Baja negativa
ATM	cg06750635	TSS ± 200	-0.416 **	Baja negativa
EZH1	cg27088844	TSS ± 200	-0.415 **	Baja negativa
KDM6A	cg11895950	TSS ± 200	-0.415 **	Baja negativa
KDM5A	cg02590486	TSS ± 1500	-0.414 **	Baja negativa
BAP1	cg00794529	TSS ± 1500	-0.413 *	Baja negativa
STK4	cg19344870	TSS ± 200	-0.412 **	Baja negativa
PRMT2	cg06725035	TSS ± 1500	-0.411 *	Baja negativa
KDM3A	cg02972551	TSS ± 1500	-0.409 *	Baja negativa
PPM1D	cg00162940	TSS ± 1500	-0.409 *	Baja negativa
RNF40	cg27151980	TSS ± 1500	-0.409 **	Baja negativa
KDM1B	cg15124342	TSS ± 1500	-0.408 *	Baja negativa
HDAC8	cg22455616	TSS ± 200	-0.407 *	Baja negativa
KDM4D	cg16993220	TSS ± 200	-0.406 **	Baja negativa
MASTL	cg03642224	TSS ± 1500	-0.406 *	Baja negativa
ATM	cg16693212	TSS ± 1500	-0.405 **	Baja negativa
ATM	cg15504467	TSS ± 1500	-0.404 **	Baja negativa
ATM	cg12848864	TSS ± 1500	-0.403 **	Baja negativa
KDM3A	cg22404242	TSS ± 1500, TSS ± 200	-0.398 *	Baja negativa
SETD1B	cg05410123	TSS ± 200	-0.397 *	Baja negativa
SETD2	cg05971911	TSS ± 200	-0.397 **	Baja negativa
EYA2	cg23214267	TSS ± 200	-0.396 **	Baja negativa
DNMT3A	cg26470599	TSS ± 1500	-0.395 *	Baja negativa
TET1	cg23350336	5'UTR	-0.394 *	Baja negativa
HDAC11	cg13397067	TSS ± 200, TSS ± 1500	-0.392 *	Baja negativa
SETD2	cg25864757	TSS ± 1500	-0.391 *	Baja negativa
MGEA5	cg09644451	TSS ± 200	-0.390 *	Baja negativa

ASH1L	cg26185843	TSS ± 1500	-0.388 *	Baja negativa
KDM6A	cg07167981	TSS ± 200	-0.387 *	Baja negativa
SIRT3	cg22239605	TSS ± 1500	-0.381 *	Baja negativa
HDAC8	cg24461939	TSS ± 200	-0.380 *	Baja negativa
PPP2CA	cg04599652	TSS ± 200	-0.380 *	Baja negativa
ATM	cg22837512	TSS ± 200	-0.379 **	Baja negativa
DNMT3A	cg06748978	5'UTR	-0.379 *	Baja negativa
DNMT3A	cg15998962	TSS ± 1500, TSS ± 200	-0.378 *	Baja negativa
ATM	cg18391757	TSS ± 200	-0.376 **	Baja negativa
KDM3A	cg02152982	TSS ± 1500	-0.375 *	Baja negativa
MASTL	cg08905758	TSS ± 1500	-0.374 *	Baja negativa
KDM3A	cg01744130	TSS ± 1500	-0.373 *	Baja negativa
PPM1D	cg06090421	TSS ± 1500	-0.372 *	Baja negativa
ATM	cg25400013	TSS ± 200	-0.371 **	Baja negativa
CLOCK	cg23506964	TSS ± 1500	-0.371 *	Baja negativa
ASH1L	cg24312492	TSS ± 200	-0.369 *	Baja negativa
SETD1B	cg21968169	TSS ± 1500	-0.369 *	Baja negativa
KDM2A	cg21293464	TSS ± 1500	-0.368 *	Baja negativa
PRKCD	cg07687398	5'UTR	-0.366 *	Baja negativa
HDAC7	cg01088404	TSS ± 1500	-0.365 *	Baja negativa
MGEA5	cg18927080	TSS ± 1500	-0.365 *	Baja negativa
PRKDC	cg00318618	TSS ± 200	-0.361 *	Baja negativa
PRMT2	cg23283320	5'UTR	-0.360 *	Baja negativa
KDM2A	cg05038268	TSS ± 1500	-0.357 *	Baja negativa
PRMT7	cg02477175	5'UTR	-0.351 *	Baja negativa
HDAC8	cg12234922	TSS ± 200	-0.348 *	Baja negativa
EYA1	cg24611608	TSS ± 200	-0.347 *	Baja negativa
ATM	cg06053805	TSS ± 200	-0.346 *	Baja negativa
EHMT1	cg21225130	TSS ± 1500	-0.342 *	Baja negativa
ATM	cg24030675	TSS ± 200	-0.326 *	Baja negativa
KAT5	cg02561205	TSS ± 200	-0.325 *	Baja negativa
SETDB1	cg10961700	TSS ± 1500, TSS ± 200	-0.325 *	Baja negativa
KDM3A	cg07211384	TSS ± 1500, TSS ± 200	-0.324 *	Baja negativa
KAT2B	cg14937543	TSS ± 200	-0.322 *	Baja negativa
SETD1A	cg03532420	5'UTR	-0.322 *	Baja negativa
PRKDC	cg09303136	TSS ± 1500	-0.319 *	Baja negativa
PRKCB	cg00735962	TSS ± 200	-0.316 *	Baja negativa
KDM4D	cg20096208	TSS ± 1500	-0.310 *	Baja negativa
ATM	cg20342375	TSS ± 1500	-0.298 *	Débil negativa
ATM	cg12106634	TSS ± 1500	-0.295 *	Débil negativa
RNF40	cg01871251	TSS ± 200	+0.664 **	Moderada positiva
PPP5C	cg16126516	TSS ± 200	+0.612 ***	Moderada positiva
PRKDC	cg05669285	TSS ± 1500	+0.608 ***	Moderada positiva
PRMT7	cg06092312	5'UTR	+0.577 **	Moderada positiva
RNF40	cg03941744	TSS ± 1500	+0.574 **	Moderada positiva
SIRT2	cg07063145	TSS ± 1500	+0.549 **	Moderada positiva
STK4	cg04834749	TSS ± 200	+0.509 **	Moderada positiva
HLCS	cg08269941	5'UTR	+0.493 **	Baja positiva
PPP2CA	cg25512513	TSS ± 1500	+0.465 **	Baja positiva
CDK5	cg18597031	TSS ± 1500	+0.460 *	Baja positiva
BUB1	cg01114573	TSS ± 1500	+0.454 **	Baja positiva
EYA3	cg01012394	TSS ± 1500	+0.452 **	Baja positiva
NAT10	cg06513075	TSS ± 1500	+0.451 **	Baja positiva
NAT10	cg09902061	TSS ± 200	+0.444 **	Baja positiva
RNF40	cg14828965	TSS ± 1500	+0.437 **	Baja positiva
EYA1	cg08234354	TSS ± 1500, 5'UTR	+0.436 **	Baja positiva
USP22	cg13763516	TSS ± 1500	+0.436 **	Baja positiva
BUB1	cg08703553	TSS ± 1500	+0.420 *	Baja positiva
PRKDC	cg22691796	TSS ± 200	+0.417 **	Baja positiva
CDK5	cg15879755	TSS ± 1500	+0.414 *	Baja positiva
SIRT2	cg22546859	TSS ± 200	+0.414 **	Baja positiva
STK4	cg18394527	TSS ± 1500	+0.414 **	Baja positiva
CDK5	cg17919020	TSS ± 1500	+0.398 *	Baja positiva
GSK3B	cg11789444	TSS ± 1500	+0.397 *	Baja positiva

KAT5	cg20108992	TSS ± 1500	+0.396 **	Baja positiva
KAT5	cg22048492	TSS ± 200	+0.396 **	Baja positiva
SIRT2	cg10186215	TSS ± 200	+0.395 **	Baja positiva
EYA2	cg01585717	TSS ± 1500	+0.386 **	Baja positiva
KDM5A	cg07165616	TSS ± 1500	+0.385 *	Baja positiva
STK4	cg27194614	TSS ± 1500	+0.385 **	Baja positiva
JMJD6	cg15457598	TSS ± 1500	+0.381 *	Baja positiva
PRMT7	cg26561196	TSS ± 1500	+0.381 *	Baja positiva
ATM	cg05231509	5'UTR	+0.380 **	Baja positiva
CARM1	cg12317676	TSS ± 1500	+0.375 *	Baja positiva
CDK5	cg16218204	TSS ± 1500	+0.370 *	Baja positiva
STK4	cg04241444	TSS ± 1500	+0.370 *	Baja positiva
CDK5	cg13301722	TSS ± 1500	+0.369 *	Baja positiva
KDM4D	cg04388472	TSS ± 200	+0.358 *	Baja positiva
DNMT3A	cg05652528	5'UTR	+0.354 *	Baja positiva
RPS6KA4	cg06848762	TSS ± 1500	+0.354 *	Baja positiva
BAZ1B	cg00525989	TSS ± 200	+0.353 *	Baja positiva
CDK5	cg09641486	TSS ± 1500	+0.352 *	Baja positiva
RPS6KA4	cg03074984	TSS ± 200	+0.349 *	Baja positiva
SIRT2	cg11761572	TSS ± 200	+0.349 *	Baja positiva
CDK5	cg06942561	TSS ± 1500	+0.346 *	Baja positiva
CDK5	cg04531694	TSS ± 1500	+0.345 *	Baja positiva
PRMT7	cg26796563	TSS ± 200	+0.332 *	Baja positiva
ATM	cg15370815	TSS ± 1500	+0.287 *	Débil positiva

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$

TSS: sitio de inicio de la transcripción

TSS ± 1500: sitio de inicio de la transcripción desde 200 hasta 1500 nt aguas arriba o aguas abajo del TSS,

TSS ± 200: sitio de inicio de la transcripción hasta 200 nt aguas arriba o aguas abajo del TSS

5'UTR: región no traducida 5'