



ANÁLISIS BIOINFORMÁTICO DE GENES DIFERENCIALMENTE EXPRESADOS INVOLUCRADOS EN EL DESARROLLO DE CÁNCER DE COLON

Frida Cano F., Gala Flores G., Hortencia A. Ramírez V., Gabriela Cortés O., Harumi C. Manzano Y., Diana P. López E., Leonardo Mauricio Cañete Sifuentes, Aurora Antonio Pérez. Tecnológico de Monterrey, Escuela de Ingeniería y Ciencias, Campus Estado de México. Carretera Lago de Guadalupe Km.3.5 Atizapán de ZaragozaCol. Margarita Maza de Juárez. CP 52926, Atizapán de Zaragoza, Estado de México.
a.antonio@tec.mx

Palabras clave: Colon cáncer, RStudio, DEG's

Introducción. En México, de acuerdo con los datos del Instituto Nacional del Cáncer 700,000 pacientes mueren cada año debido a al cáncer de colon (CC). [1] Actualmente existen diversas formas de tratamiento, sin embargo, no han disminuido significativamente la mortandad del CC [2]. Por esa razón se han buscado diversos medios tecnológicos para encontrar nuevas soluciones para el tratamiento y la detección temprana del CC. Este trabajo se centra en la aplicación del lenguaje de programación o R para analizar un microarreglo de las bases de datos (GEO Omnibus) con el fin de obtener una visualización más clara y concisa de lo que está ocurriendo con los genes diferencialmente expresados (DEG's) en estadios diferenciados del CC y posteriormente se correlacionaron a sus funciones moleculares y procesos biológicos.

Metodología. El análisis se realizó a través de una base de datos de muestras obtenida de GEO dataset con ID "GSE161097". A partir de las características de los datos del microarreglo, se hizo la construcción de cuatro grupos a analizar: CC en etapa inicial con presencia de mucina (PM), CC en etapa inicial en ausencia de mucina (PNM), metástasis peritoneal con presencia de mucina (MM) y metástasis peritoneal en ausencia de mucina (MNM). Los datos obtenidos se procesaron y filtraron mediante la aplicación de las librerías 'GEOquery' y 'limma', de la plataforma RStudio. Posteriormente se agruparon y normalizaron los valores ('cbind' y 'sweep'). La ausencia o presencia de mucina en cada tejido se asignó a dos arreglos diferentes uno para etapa inicial de CC y otro de metástasis peritoneal. A partir del microarreglo normalizado se obtuvieron los p-values, seleccionando genes aquellos con $P < 0.05$. Se obtuvo una lista de DEG's se analizaron en Genbank, STRING, GO, COSMIC y genomAD. Recursos que permitieron un análisis sistemático y biológico molecular de los DEG's.

Resultados. El filtrado de DEG's por medio de R, redujo significativamente el número de genes; para el grupo CC en etapa inicial de 780 genes se redujo a 13 con presencia de mucina y 3 genes sin mucina, mientras que en el tejido de metástasis peritoneales se redujo a 12 con mucina y 23 sin mucina (Fig.1). Se hizo el agrupamiento de los DEG's por medio de un diagrama de calor (Fig. 2).

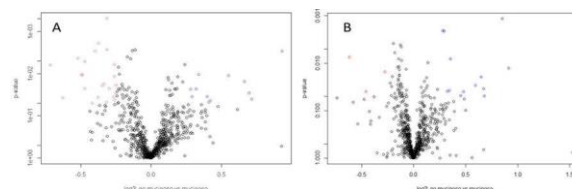
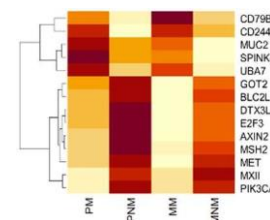


Fig. 1. Gráficas de volcano de los DEG's en tejidos de CC. (A) en etapa inicial, (B) en tejidos con metástasis peritoneal. Los (●) representan DEG's en su respectivo tejido sin mucina y los (●) representan DEG's en su respectivo tejido con presencia de mucina.

Fig. 2. Mapa de calor. Se puede observar la expresión de los genes para los cuatro grupos analizados: PM, PNM, MM y MNM. Permitiendo su análisis acorde a una jerarquía según la similitud que hay entre ellos.



Con el fin de comprender la relación de funcionalidad de los genes diferencialmente expresados se generaron dos redes de interconexión con los grupos de estudio:

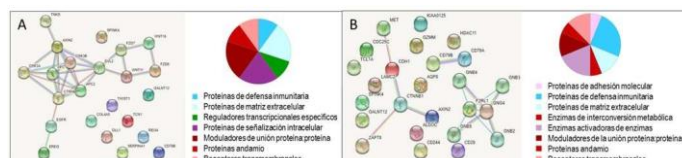


Figura 9. Red de interacción proteína – proteína y función de los DEG's identificados. (A) CC etapa inicial (B) metástasis peritoneal.

Conclusiones. La presencia o ausencia de mucina mostró una diferente agrupación de DGE's en un mismo tipo de tejido. Al analizar la interrelación y funciones de las proteínas se observó que el CC etapa inicial de la cavidad peritoneal depende de la activación y reacción del sistema inmunológico así como se encuentran relacionados con otras proteínas encargadas de procesos biológicos de reparación y adaptación. Las herramientas computacionales son clave para la innovación de la metodología de investigación estadística, ya que permiten manipular bases de datos voluminosas de manera sumamente eficaz.

Bibliografía.

- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. CA Cancer J Clin 65:87–108
Kuipers, E. J., Grady, W. M., Lieberman, D., Seufferlein, T., Sung, J. J., Boelens, P. G., van de Velde, C. J. H., & Watanabe, T. (2015). Colorectal cancer. Nature Reviews Disease. Primers, 1(1), 1.