

Utilidad del análisis integrativo para la identificación de microARNs para diagnóstico

(1) [Julieth López](#), (2) David Niño, (3) Liliana López, (4) Adriana Rojas, (5) Litzy Bermúdez

(1-3) Departamento de estadística, Facultad de ciencias, Universidad Nacional de Colombia, Bogotá, Colombia (4-5) Instituto de genética humana, Facultad de medicina, Pontificia Universidad Javeriana, Bogotá, Colombia

Correos: (1) julalopezcas@unal.edu.co, (2) jdnotinot@unal.edu.co, (3) llopezk@unal.edu.co, (4) rojas-adriana@javeriana.edu.co, (5) litzbermudez@javeriana.edu.co

15 de diciembre, 2021

El cáncer del pulmón es el tercer cáncer con mayor incidencia (22.4) en el mundo, sin embargo, es el cáncer con la mayor tasa de mortalidad (18.0) según la Organización mundial de la salud (2020). Por otro lado, el cáncer de células no pequeñas (NSCLC) representa el 85% de todos los casos. En la mayoría de los casos el diagnóstico se realiza en etapas tardías, por ello, investigaciones que permitan la detección temprana son cruciales para disminuir la tasa de mortalidad y mejorar el éxito de los tratamientos. Para esto han sido hallados biomarcadores de transcripción principalmente en tejido y con base en genes desregulados. Pocos se han propuesto con microARNs que además de encontrarse en tejidos, también se encuentran en fluidos de manera libre o en vesículas extracelulares. Por lo anterior, tienen un alto potencial para servir de biomarcadores en biopsias líquidas que son de menor invasión que las biopsias de tejidos.

El objetivo del estudio fue aplicar un flujo de trabajo para el análisis bases de datos secundarias del [NCBI](#) de exosomas provenientes tanto de biopsia líquida como de tejidos, que permitiera el análisis integrativo de las bases seleccionadas para detectar posibles biomarcadores para NSCLC.

La ruta de trabajo que se estableció consistió en una búsqueda en el NCBI sobre bases de datos de vesículas extracelulares en pacientes con NSCLC y pacientes normales (fueron seleccionadas las bases GSE71661 y GSE111803). Se realizó un análisis de control de calidad y de expresión diferencial utilizando el método de [DESEQ2](#) en R para detectar microARNs diferencialmente expresados en cada base de datos y luego se tomaron aquellos que estaban en común. Para ver el papel biológico y su interpretación en la enfermedad, se realizó la identificación de blancos de los microARNs hallados utilizando la plataforma [MiRNet 2.0](#) que permitió generar una red de interacción que se analizó en [Cytoscape](#). Posteriormente se realizó un análisis de enriquecimiento con la herramienta [DAVID](#) en donde se detectaron las vías y procesos involucrados con los blancos seleccionados. De aquí se realizó un análisis biológico y se revisó en la literatura para entender como influyen los microARNs encontrados en la enfermedad.

Se encontraron 5 microARNs en común, todos sobre-expresados: hsa-miR-1246, hsa-miR-144-3p, hsa-miR-16-2-3p, hsa-miR-194-5p, hsa-miR-505-5p y 293 genes blanco con los que la anotación dio 398 términos biológicos con los que se desarrolló el objetivo del estudio.

Finalmente, a pesar de estudiar bases de datos por separado con individuos distintos, pero con una enfermedad común, encontramos resultados que biológicamente tienen sentido y que cumplen con el objetivo de hallar potenciales biomarcadores. Recalamos la ruta de trabajo que involucró alrededor de 4 plataformas las cuales permitieron relacionar las partículas estudiadas inicialmente, los microARNs, con los genes y las vías de señalización.