# Utility of the integrative analysis for the identification of microRNA for diagnosis

*Julieth López-Castiblanco*[1], *David Niño*[2], *Liliana López-Kleine*[3], *Adriana Rojas*[4], *Litzy Bermúdez*[5]

[1]julalopezcas@unal.edu.co, Department of Statistics, National University of Colombia

[2]jdninot@unal.edu.co, Department of Statistics, National University of Colombia

[2]llopezk@unal.edu.co, Department of Statistics, National University of Colombia

[2]rojas-adriana@javeriana.edu.co, Genetics Human Institute, Pontificia Javeriana University

[2]litzybermudez@javeriana.edu.co, Genetics Human Institute, Pontificia Javeriana University

Lung cancer is one of the main causes of mortality due to the late detection of this disease. Currently, methods to identify lung cancer are found in lung cancer tissues. Non-invasive tools for early detection would be very helpful. New studies have achieved some possibilities of diagnosis using blood samples. Nevertheless, knowledge about biomarkers in blood and circulating exosomes is still very sparce. Therefore, this study seeks to identify which miRNA present in exosomes of blood sample share deferentially expressed in people with lung cancer through biostatistical analysis to support diagnosis of lung cancer. For this aim, the work was focused on one of the most common kind of lung cancer: adenocarcinoma. Two publicly available databases of the National Center for Biotechnology Information (NCBI) were used (accession numbers: GSE71661 and GSE111803). Both databases were obtained through non-coding RNA profiling by high throughput sequencing experiment, so both have discrete data and thus, can be analyzed with methodologies adapted to this, like DESEQ2.

Initially, quality control was done to conduct a differential gene expression analysis based on the negative binomial distribution. This methodology allowed to find five common miRNA between both databases which can help to characterize lung adenocarcinoma. Consequently, the identification of genes known to be targets of these miRNAs was done using miRNet. Then, the target genes were analyzed for their functional role and highlighted metabolic pathways using DAVID. With this enrichment analysis we were able to construct a network with the five differential expressed miRNA and their gene targets. Furthermore, topological analysis of the network allowed to identify potential important genes in adenocarcinoma.

**Keywords**: microRNA, lung cancer.