

Dahlia E. Siprut
Undergraduate
University of New Orleans New Orleans, LA

Mentor: Chindo Hicks, Ph.D, Professor Genomics and Bioinformatics
LSUHSC-SOM, Director of Bioinformatics Program, Department of Genetics, New Orleans, LA

“An Integrative Genomics Approach to Discovery of Molecular Markers in Ovarian Cancer”

Background: Ovarian cancer (OC) is one of the most common gynecologic cancers that has the highest mortality rate. OC is the eighth most commonly occurring cancer in women and the 18th most commonly occurring cancer overall. World-wide there were nearly 300,000 new cases of OC in 2018. Within the United States an estimated 22,530 women were diagnosed with new cases of OC and an estimated 13,980 women died from the disease in 2019. Therefore, there is an urgent need for the discovery of molecular markers for early detection and prognostic prediction of the disease. Advances in next generation sequencing technology have enabled generation of vast amounts of gene expression and somatic mutation data on cancer genomes including OC. Although much progress has been made on classification of molecular subtypes of OC using transcription profiling, gene expression has not been leveraged and integrated with somatic mutations information for the discovery of diagnostic and prognostic markers. The objective of this investigation was to discover prognostic markers that are predictive of clinical outcome using gene expression and somatic mutation data. Our working hypothesis was that genomic alterations in the transcriptomes and tumor genomes of women diagnosed with OC could lead to measurable changes distinguishing patients who survived the disease from those who did not survive the disease. **Material and Methods:** We addressed this hypothesis using gene expression and somatic mutation data derived from a total of 376 samples (230 died from the disease and 146 survived the disease) from the Cancer Genome Atlas (TCGA). The data was partitioned into two patients groups, those who survived the disease and those who died from the disease. We performed analysis comparing gene expression levels between the two patient groups to discover a signature of significantly differentially expressed genes distinguishing the two patient groups. Significantly differentially expressed genes were evaluated for the presence of somatic mutations to identify a signature of significantly differentially expressed genes which were also significantly differentially mutated distinguishing the two patient groups. **Results:** The analysis revealed a signature of 130 differentially expressed genes ($P < 0.005$) of which 50 were significantly ($P < 0.001$) differentially expressed genes distinguishing patients who survived from patients who died. Evaluation of these genes for the presence of somatic mutations revealed a signature of 23 significantly differentially expressed genes which were also differentially mutated distinguishing the two patient groups. Among the top somatic mutated differentially expressed genes distinguishing the two patient groups included the genes: TAP1, CD79A, CD2, TAP2, EMP1, CD3E, and ITK. **Conclusion:** We discovered a signature of somatic mutated differentially expressed genes distinguishing patients who survived OC from patients who died from the disease. Our investigation demonstrates that integrative analysis combining gene expression with somatic mutation data is a powerful approach to discovery of molecular markers predictive of clinical outcomes and clinical endpoints.