

### **Project #1: immune gene expression in prostate cancer**

miRNAs are small (19-24 nt) RNA molecules that negatively regulate gene expression. In recent years these molecules have been investigated as potential biomarkers and predictors of recurrence and/or survival in cancer and other diseases. The goal of this project to develop a data mining/machine learning approach that ultimately can be used in disease liquid diagnostics. Specifically, project participants will use an existing set of serum/plasma miRNA profiles from healthy patients to apply data mining and machine learning approaches to (i) identify technical issues and sample outliers, (ii) differentiate plasma and serum profiles using Fuzzy Inference Systems, and (iii) study a patient with an unusual miRNA phenotype

### **Project #2: Genomic data analysis and visualization GUI**

Analysis of a genomic data often involves pre-processing, quality control, normalization, feature selection and classification and differential expression analysis. Many methods exist, however, the best technique depends on the dataset. Therefore, it is often required to try different techniques to select the one that works best for a given dataset.

This project involves development further development and improvement of a user-interface for a feature selection algorithm and feature analysis. The objective is to implement new components and improve on existing components of a user-friendly user interface that runs the algorithm for different genetic data types. The interface would be published online for other researches to use.

### **Project #3: REDCAP database for clinical and pathological data**

This project involves developing, entering data and testing a RedCap database for clinical and pathological data for different types of malignant lymphoma. The RedCap database platform is currently set at the CAC and has a basic set up. The goal is to design an effective implementation to store data for easy data retrieval by clinicians and research users.