Chapter 1

Introduction

1.1. Background

Prostate Cancer (PC) is a slow progression cancer that typically remains undiagnosed for up to 20 years (García-Perdomo *et al*., 2018). PC is one of the most diagnosed cancers among men and is known to cause major mortality in men living western lifestyles (Lee *et al*., 2017). Gradually, the severity of PC commonly increases with age for each individual. According to the National Cancer Institute (2020), there is an average of 5.7% mortality rate of PC cases in the United States (US) based on the data collected in 2020. Meanwhile, according to the International Agency for Research on Cancer, the average new cases of PC is 7.3% with an average mortality rate of 3.8% in 2020 (IARC., 2020). Although there has been an advancement in the diagnosis and treatment of PC, this specific cancer is still considered a global health problem as tens of millions of people are diagnosed with cancer around the world each year, and more than half of the patients eventually die (Ma & Yu, 2006).

Long non-coding RNAs (lncRNA) refers to over 200 nucleotides that are associated with the microRNA to exhibit intron-exon splicing, regulated transcriptional control, and expression in a tissue-specific context (Quinn & Chang, 2015). It is a non-coding RNA and does not provide instructions for making proteins. However, it has been found that the lncRNA has specific functional groups such as scaffolds, signals and decoys that are able to help regulate target genes (Salviano-Silva *et al.*, 2018). Other than that, the lncRNA(s) also have critical regulatory roles in many cellular processes, which includes post-transcriptional processing, intracellular trafficking, transcription, and chromatin remodeling (Chen & Carmichael, 2010).

Utilizing lncRNA(s) as a potential biomarker with high sensitivity and specificity is urgently needed to cease the over-diagnosis of PC that usually occurs through the Prostate Cancer Specific Antigen (PSA) test, which has high sensitivity but low specificity (Etzioni, 2002)

and has been shown to have low diagnosis ability in men of African-American (AA) descent (Brawley, 2012). Although there are many studies that have analyzed the molecular functions of lncRNA(s), available previous study showed that none of the studies convey further capabilities of lncRNA(s) in distinguishing racial differences for AA patients (Ma *et al*., 2018). In addition, the previous studies did not consider the computational aspects for predicting lncRNA functions and its target, especially in the racial differences. Thus, PC research is urgently needed for the identification of biomarkers that are capable of covering the racial differences for AA patients. In this study, the RNA Sequencing data (RNA-Seq) was used to conduct differential and network-based analyses to predict which lncRNA(s) regulate the greatest number of gene targets which might be used as potential biomarkers for the racial differences of PC in AA men. The top 10 lncRNA(s) identified to be regulating the greatest number of gene targets were subjected to a literature review to examine whether they had been previously discovered and/or had not been previously discovered in the context of PC.

1.2. Objectives

To identify the long non-coding RNA(s) (lncRNA) and differentially expressed genes that are useful for further development for racial differences biomarkers in prostate cancer patients of African-American (AA) descent.

1.3. Hypothesis

H0: There will be no lncRNA(s) that are useful for further development as biomarkers for racial differences in PC patients of AA descent.

H1: There will be lncRNA(s) that are useful for further development as biomarkers for racial differences in PC patients of AA descent.