CHAPTER I: INTRODUCTION

The statistics of melanoma skin cancer in the world is quite significant; It can be seen from the press release No. 311 by World Health Organization (WHO) in France on 30 March 2022 (World Health Organization, 2022). In the press release, WHO concerns about the cutaneous melanoma cases that estimated will dramatically increase 50% from 2020 to 2040, the rising of cases will also be associated with the higher mortality cases (World Health Organization, 2022). It is reported in 2020, that the new cases of diagnosed melanoma worldwide reaches 325.000 and 57.000 patients died due to melanoma cancer. With the highest incidence reported in Australia, New Zealand, Western Europe, Northern America and Northern Europe. However, as Africa and Asia countries it is found that the cases are rare below 1 per 100.000 (Arnold, et al., 2022). In 2023, the statistical prediction for Skin Cutaneous Melanoma by American Cancer Society is estimated to be 97,610 in which male patients have higher cases compared to female patients. It is also predicted that 7,990 patients of melanoma skin cancer are likely to die in 2023 where the highest number is owned by male patients (American Cancer Society, 2023). Therefore, according to the statistical report in 2022, WHO has declared the urgency for targeted melanoma control measures and reported cutaneous melanoma as the global burden in 2020.

The main risk factor of skin cutaneous melanoma cancer is due to UV exposure, this factor is associated with the incidence of cutaneous melanoma usually found and developed on sun-exposed areas in patient's bodies (Naik, 2021). Another risk factors such as tanning beds, lighter skin, genetic history, viral infection by Human Papillomavirus (HPV) also plays a role in increasing the SKCM risk (Conforti & Zalaudek, 2021).

Early detection of skin cutaneous melanoma that can distinguish the type of primary and metastatic tumor can reduce the risk of cancer spread and increase the cure rate, especially when it can prevent the melanoma cells that pass throughout the lymphatic system before entering the bloodstream. Generally, in order to diagnose the cutaneous melanoma cancer, the visual skin inspection is applied, those including assessing the asymmetry, border, color, diameter and evolution (ABCDEs) using dermoscopy technique or another way is by using the biopsy technique (Hartman & Lin, 2019). By those examinations, the professional health workers will understand the stages and condition of the skin cutaneous melanoma and then be ready then ready to suggest the best treatment method to overcome the cancer. Normally, SKCM patients will undergo a surgery to remove the tumor, the different types of surgery will be based on the examination of the patients whereas for the metastatic melanoma the patients usually have medical therapy options such as immunotherapy, targeted therapy or adjuvant therapy that suggested by the doctors (Hartman & Lin, 2019).

As stated, early and effective diagnosis of cutaneous melanoma is very important in managing the cancer. Biomarker-based cancer diagnosis has become the bright way for early detection, disease progression monitoring and effective cancer therapy (Sarhadi & Armengol, 2022). This diagnostic biomarker may detect or confirm the presence of the disease or specific condition by using the body fluid. Differentially Expressed Genes (DEGs) is a computational technique used to identify the genetic mechanism related with expression level difference for each gene (MCDermaid, et al., 2019). This technique will use high-throughput sequencing data and assembly to analyze the difference between two or more conditional designs, this technique may provide insights to show the differential gene from the conditions. Bioconductor as the genomic repositories offered some packages to analyze the DGE or transcripts, which are DESeq2, EdgeR, Limma, DEGseq, Cuffdiff/Cuffdiff2, SAMseq and many more (Williams et al., 2017). The difference between most of the DGE packages lies in the normalization, read count distribution assumption and differential expression test (Gerolami, et al., 2023). Due to many availabilities of the packages or methods to perform the DEG, there is no clear consensus on which these methods perform best (Paepe, 2015). The example of three packages that use in DEGs are limma, edgeR and DESeq2 are similar but have different gradual order among the analyses. For example, Limma uses a linear model to generate the statistics result, while both DESeq2 and edgeR use negative binomial distribution (Liu, et al., 2021). The study in metastasis- associated gene for SKCM using 2 datasets from GEO database and limma package is use for evaluate the DEGs result in 3267 DEGs for Primary Tumor and normal in which 1444 gene were upregulated and 1823 were down regulated. For metastatic vs normal limma detect 5777 DEGs in which consist of 3017 upregulated and 2758 downregulated, meanwhile in the other dataset it found from primary tumor vs normal 1127 DEGs were found (527 upregulated and 600 downregulated) and for metastasis vs normal limma detect 2016 DEGs (952 upregulated and 1064 downregulated), the obtained data were screen use P value. adj. $P < 0.05$ and $|\log2FC| > 1$ (Luan, et al., 2022). Another study in gene identification associated with SKCM in 2020 using TCGA dataset samples to see the DEGs between primary tumor and metastasis using edgeR found that 464 DEGs were screened in which consist of 287 upregulated and 177 downregulated using the cutoff of $P < 0.01$ and $|\log FC| \geq 1$ (JI, et al., 2020). Another study that aims to find the prognosis hub-gene of SKCM using DESeq2 with $|\log2FC| > 2$ and adj $P < 0.01$ as a standard cutoff found that from 3834 DEGs in which 1622 is upregulated and 2212 is downregulated found, there were 13 significant genes that identified from comparison of normal and primary tumor. All the screened genes were associated with keratin related

protein and 11 genes from that were identified also in the comparison of metastasis and primary tumor of SKCM (Li, Qi & Yang, 2021).

Therefore, combining these three packages at once to find the overlapping genes might yield in a good potential diagnostic of genotypic results. Further non-parametric testing using a spearman correlation analysis and miRTarBase on the miRNA and genes by generating the ROC curve beforehand helps to identify the optimal cut-off that separates the specific condition of the disease (Wray, et al., 2010); (Yang, et al., 2021); (Zhou & Zhang, 2019); (Huang, et al., 2022). The spearman correlation analysis is usually used to identify the relationship between the DEGs that are found in two condition samples (Wang, et al., 2023). Followed by a pathway analysis using Gene Ontology and KEGG for observing the miRNA and gene that has a pathway that is related with skin cutaneous melanoma in KEGG or Go pathway. Lastly, in order to identify the duration of cutaneous melanoma patients are likely to survive, a survival analysis using Kaplan-Meier plot will be applied to visualize the survival curves. The survival curve in Kaplan-Meier plot usually reports as the median survival time (Sheng, et al., 2020).