FERROPTOSIS-RELATED LONG NONCODING RNA SIGNATURE PREDICTS THE PROGNOSIS OF CLEAR CELL RENAL CELL CARCINOMA

LIU JIAWEN

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2022

FERROPTOSIS-RELATED LONG NONCODING RNA SIGNATURE PREDICTS THE PROGNOSIS OF CLEAR CELL RENAL CELL CARCINOMA

LIU JIAWEN

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCE
FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2022

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **LIU JIAWEN**

Registration/Matric No: **SMA180043**

Name of Degree: **MASTER OF SCIENCE**

Title of Thesis:

FERROPTOSIS-RELATED LONG NONCODING RNA SIGNATURE PREDICTS THE PROGNOSIS OF CLEAR CELL RENAL CELL CARCINOMA

Field of Study:

BIOINFORMATICS

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date 21.8.2021

Subscribed and solemnly declared before,

Witness's Signature Date 21.8.2021

Name:

Designation:

FERROPTOSIS-RELATED LONG NONCODING RNA SIGNATURE PREDICTSTHE PROGNOSIS OF CLEAR CELL RENAL CARCINOMA

ABSTRACT

Clear Cell Renal Cell Carcinoma (ccRCC) is very common and accounts for most kidney cancer deaths. While many studies are being conducted in finding the prognosis signatures of ccRCC, we believe that ferroptosis, that involves programmed cell death dependent on iron accumulation has therapeutic potential in ccRCC. Recent research showed that long noncoding RNAs (lncRNAs) have been shown to be involved in ferroptosis-related tumor processes and are closely related to survival in patients with ccRCC. Hence in this study we aim to further explore the role of ferroptosis-related lncRNAs (FRLs) in ccRCC, hoping to establish a signature to predict the survival outcome of ccRCC. Here we analyzed transcriptome data from The Cancer Genome Atlas database (TCGA) and ferroptosis-related genes (FRGs) from FerrDb to identify FRLs using Pearson's correlation. Lasso Cox regression analysis and multivariate Cox proportional hazards models screened seventeen optimal FRLs for developing prognostic signatures. Kaplan–Meier survival curves and receiver operating characteristic (ROC) curves were then plotted for validating the sensitivity, specificity, and accuracy of the identified signatures. CIBERSORT algorithm were deployed to explore the role of these FRLs in tumor microenvironment (TME). It was concluded that these models demonstrate excellent performance in predicting prognosis among patients with ccRCC and which also indicated association with the clinicopathologic parameters such as tumor grade, tumor stage and tumor immune infiltration. In conclusion, our findings provide novel insights into ferroptosis-related lncRNAs in ccRCC which are important targets for investigating the tumorigenesis of ccRCC.

Keywords: clear cell renal carcinoma, ferroptosis, long non-coding RNA, prognosis, biomarker

FERROPTOSIS-RELATED LONG NONCODING RNA SIGNATURE PREDICTSTHE PROGNOSIS OF CLEAR CELL RENAL CARCINOMA

ABSTRAK

Clear Cell Renal Cell Carcinoma (ccRCC) merupakan punca kematian dalam kebayakan kes barah buah pinggang. Walaupun banyak kajian sedang dijalankan untuk mencari "signature prognosis" ccRCC, kami percaya bahawa ferroptosis, yang melibatkan kematian sel yang diprogramkan bergantung pada pengumpulan besi mempunyai potensi untuk terapi dalam ccRCC. Penyelidikan terkini menunjukkan bahawa "long noncoding RNAs" (lncRNA) telah terbukti terlibat dalam proses tumor yang berkaitan dengan ferroptosis dan berkait rapat dengan jangka hidup pesakit dengan ccRCC. Oleh itu, kajian ini bertujuan untuk meneroka dengan lebih lanjut peranan lncRNA dalam ferroptosis-related lncRNAs (FRLs) bagi ccRCC, untuk meramalkan jangka masa hidup pesakit ccRCC. Di sini kami menganalisa data transkrip dari pangkalan data The Cancer Genome Atlas (TCGA) dan gen yang berkaitan dengan ferroptosis (FRG) dari FerrDb untuk mengenal pasti FRL menggunakan korelasi Pearson. Analisis regresi Lasso Cox dan model "multivariate Cox proportional hazards" telah menyaring tujuh belas FRL optimum untuk mengembangkan "prognostic signatures". Lengkung jangka hidup Kaplan–Meier dan lengkung "receiver operating characteristic" (ROC) dibina untuk mengesahkan kepekaan, pengkhususan, dan ketepatan "signature" yang dikenal pasti. Algoritma CIBERSORT digunakan untuk meneroka peranan FRL ini dalam persekitaran mikro tumor. Secara kesimpulannya, model-model ini menunjukkan prestasi yang sangat baik dalam meramalkan prognosis di kalangan pesakit dengan ccRCC dan yang juga menunjukkan kaitan dengan parameter "clinicopathologic" seperti gred tumor, tahap tumor dan "tumor immune infiltration". Kesimpulannya, penemuan kami memberikan

pandangan baru mengenai lncRNA yang berkaitan dengan ferroptosis dalam ccRCC yang merupakan sasaran penting untuk menyiasat tumorigenesis ccRCC.

Kata kunci: clear cell renal cell carcinoma, ferroptosis, long non-coding RNA, prognosis, biomarker

ACKNOWLEDGEMENTS

First and foremost, I would like to thank supervisors — Associate Professor Dr. Sarinder Kaur Kashmir Singh and Doctor Farahaniza Binti Supandi — for their guidance, support, and advice. I would like to express my gratitude and appreciation to the participants/subjects who were willing to take part in this research study.

Also, my warm thanks to my fellow labmates and friends in for the stimulating discussion, for the struggle we have in working together before deadlines and for all the fun we have had. Special thanks to Khaw cheesiong for all the helps and supports.

Lastly, I would like to thank my family for their love and care. Without their support (mentally and physically), I may not have been able to successfully complete this study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

ccRCC	:	clear cell renal cell carcinoma
DELs	:	differentially expressed lncRNAs
FRGs	:	ferroptosis-related genes
FRLs	:	ferroptosis-related lncRANs
lncRNA	:	long noncoding RNA
OS	:	overall survival
RCC	:	renal cell carcinoma
ROC	:	receiver operating characteristic
ROS	:	reactive oxygen species
TCGA	:	The Cancer Genome Atlas
TME	:	tumor microenvironment

CHAPTER 1: INTRODUCTION

1.1 Background

Ferroptosis as a novel form of cell death was first proposed in 2012. It morphologically, biochemically, and genetically differs from other types of cell death such as apoptosis, autophagy, and necrosis (Dixon et al., 2012). A growing number of studies have revealed the emerging roles of aberrant ferroptosis in diverse cancer types and in cancer treatment (Mou et al., 2019). Long noncoding RNAs, regulating gene expression at transcriptional levels and translational levels, essentially involved in tumorigenesis and tumor metastasis (Jiang et al., 2019). Recent studies showed that long noncoding RNAs implicated in ferroptosis-related tumor process (Luo et al., 2021) and closely related to the survival of patients with ccRCC (Zhang et al., 2019b). Therefore, establishing a ferroptosis-related long noncoding RNAs signature was proposed to help predicting the prognosis of patients with clear cell renal cell carcinoma.

1.2 Problem Statement

1. The role of ferroptosis-related genes associated with long noncoding RNAs in clear cell renal cell carcinoma remains unknown. Hence a research revealing the relationship between ferroptosis-related long noncoding RNA and overall patient survival is needed.
2. Few patients benefit from immunotherapy, treatment strategies still need to be improved. Hence the tumor immune microenvironment of patients with clear cell renal cell carcinoma should be further explored to determine effective biomarkers.

1.3 Research Objectives

This study uses a data science approach to establish FRLs in predicting the prognosis in ccRCC patients. The following are the specific objectives:

1. To determine a robust ferroptosis-related long noncoding RNA signature for predicting overall survival of patients with clear cell renal cell carcinoma.
2. To explore the tumor microenvironment of patients with clear cell renal cell carcinoma.

1.4 Scope of Research

Due to limited time, this research only focuses on constructing a ferroptosis-related long noncoding RNAs signature for predicting the overall survival of patients with clear cell renal cell carcinoma. The exploration of whether the signature is applicable to other cancer types such as liver cancer, lung cancer and so on is not included in this research work.

1.5 Significant of Research

The prognostication and treatment of patients with clear cell renal cell carcinoma are principally guided by tumor stage in recent years clinical practice (Ljungberg et al., 2015; Motzer et al., 2015). However, due to molecular heterogeneity, the outcomes are still different for patients with the same tumor stage (Molina et al., 2014). Therefore, identifying individualized biomarkers is of great significance. On one hand, it helps to identify patients at high risk of death. On the other hand, it helps to optimize treatment effect by stratifying patients for individual treatment.

1.6 Dissertation organization

This dissertation consists of six chapters. In this chapter, research background, problem statement, research objectives, the scope of research, significance and the organization of the dissertation were discussed.

Chapter two is literature review. It begins with the discussion of clear cell renal cell carcinoma, ferroptosis and long noncoding RNAs, followed by the relationship between ferroptosis and long noncoding RNAs and several relevant ferroptosis-related long noncoding RNAs signatures, tumor microenvironment and data science techniques were also presented before the chapter summary.

Chapter three describes the methodology. It begins with the collection of the relevant data, data quality control, the identification of the ferroptosis-related lncRNAs, following by the development and validation of the ferroptosis-related lncRNA prognostic signature. Then, statistical analysis was presented in end of the chapter.

Chapter four presents the result of the research. The enrichment analysis of ferroptosis-related, the result of Ferroptosis-related lncRNAs in ccRCC, the construction of ferroptosis-related lncRNAs signature, the validation process of the prognostic score, the Nomogram establishing and clinical utility of the risk score were presented in this chapter.

Chapter five is the discussion on the result and finding in the research. The research finding, analysis and the limitation of the research were presented.

Chapter six is the conclusion, and the last chapter, chapter seven is the reference list.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

The first part of this chapter is the discussion of clear cell renal cell carcinoma followed by ferroptosis and the role of long noncoding RNAs in cancer. The relationship between ferroptosis and long noncoding RNAs is demonstrated next. The relevant work about ferroptosis-related long noncoding RNAs signature were further discussed before the summary is presented.

2.2 Clear cell renal cell carcinoma

According to Sung et al. (2021), there are 431,288 new diagnostic cases and 179,368 deaths in renal cell carcinoma (RCC) all over the world in 2020. Excess body weight, tobacco use, and hypertension are the major established risk factor for developing RCC (American Cancer Society, 2021). Although several therapeutic options such as surgery, partial nephrectomy, radical nephrectomy, targeted therapies and immunotherapy are available (Hsieh et al., 2017), the 5-year survival rate is only 13% if RCC has spread to a distant part of the body (American Cancer Society, 2021). One of the most common RCC is ccRCC which is responsible for approximately 70-75% of all renal cell carcinoma cases (Störkel et al., 1997). Moreover, each of these traditional options has limitations. Thus, it is urgent to identify potential valuable molecular biomarkers to improve the patient survival. Several of pathological and molecular prognosis biomarkers such as tumor size, histological subtypes and nuclear grade, can be used to predict overall survival (Lam et al., 2008). Clinical factors, particularly tumor stage, are the main predictable indicators of survival for most patients with ccRCC, but these factors do not predict accurately because of the molecular and genetic heterogeneity (Ljungberg et al., 2015; Motzer et al., 2015). Therefore, identifying new prognostic biomarkers is one of the effective

approaches to improve patients' prognosis. This will also help in the identification of patients that are at a greater risk of death situation.

2.3 Ferroptosis

2.3.1 The concept of ferroptosis

There are three types of cancer cells deaths which are apoptosis, autophagy and necrosis during tumor treatment (Lu et al., 2018). Recently, ferroptosis, a new type of iron-dependent programmed cell death was first proposed by Dixon in 2012(Dixon et al., 2012) . Ferroptosis is caused by the abnormal levels of iron, that is, iron overload triggers the abnormal activation of the mitochondrial oxidative phosphorylation pathway, meanwhile high levels of reactive oxygen species (ROS) are yielded when ATP is produced, then when ROS concentration surpass the clearance level of antioxidant systems, it can directly or indirectly destroy the structure and functions of cells by oxidizing unsaturated fatty acids and forming lipid peroxides. This kind of cell death is call ferroptosis. it is securely related to glutathione (GSH) metabolism, iron metabolism, and lipid peroxidation. Thus, transferrin receptor 1 (TFR1), ferritin, cystine/glutamic acid reverse transporter (system Xc-), glutathione peroxidase 4 (GPX4), and lipoxygenase are involved in the occurrence of ferroptosis (Xie & Guo, 2021).

2.3.2 The role of ferroptosis in cancer

Due to the importance of ferroptosis in cell death, recent studies have begun to unravel the role of ferroptosis genes in cancer survival and cell death. Interestingly, p53, a key tumor suppressor that contains homozygous mutations in ~50–60% of human cancers has been reported to induce ferroptosis (Baugh et al., 2018; Jiang et al., 2015). It was also reported that ferroptosis could potentially contribute to the tumor-suppressive activity of p53 (Jiang et al., 2015). Sensitivity profiling in 177 cancer cell lines showed that GPX4

is the key ferroptosis regulator in diffusing large B cell lymphomas and renal cell carcinomas (Yang et al., 2014). A recent study on ferroptosis by Li et al. (2020a) indicated that ferroptosis was induced in pancreatic cancer, hepatocellular carcinoma, gastric cancer, colorectal cancer, breast cancer, lung cancer and clear cell renal cell carcinoma. Another study by Eling et al. (2015) demonstrated that artesunate (ART) induces ROS production and stimulates ferroptosis in pancreatic ductal adenocarcinoma cell lines. Based on aforementioned studies, it is very likely that ferroptosis may offer potential therapeutic options in tumor therapy. Evidence also shows that a number of ferroptosis inducers can effectively kill tumor cells in various preclinical animal experiments (Hassannia et al., 2019; Stockwell & Jiang, 2020). Hence, ferroptosis-inducing agents show potential as novel therapeutic for the tumor treatments. A recent study discovered that immunotherapy-activated CD8+ T cells improved the ferroptosis-specific lipid peroxidation in cancer cells, and these improved ferroptosis was essential in enhancing the immunotherapy efficacy (Wang et al., 2019). Therefore, the mechanism of T cell-stimulated tumor ferroptosis may provide a new therapeutic approach for treating cancer.

2.4 The role of long noncoding RNAs in cancer

Long noncoding RNA is receiving more and more attention in current cancer research. It is defined as RNAs longer than 200 nucleotides that could mediate gene regulation through binding with DNA, RNA, or proteins and then measure tumor progression, recurrence, and metastasis (Hauptman & Glavač, 2013). lncRNAs function as key modulator, participating in chromatin organization, transcription, post-translational regulation such as mRNA splicing (Choudhari et al., 2020) and regulating signaling pathways including p53, NF-κB, PI3K/AKT and Notch (Peng et al., 2017). More importantly, lncRNAs play significant roles in biological process such as tumor

carcinogenesis promotion and inhibition, drug resistance and cancer metastasis (Du et al., 2020).

A lot of studies have shown that lncRNA involved in cancer initiation and/or progression. For example, dysregulated lncRNAs play a key role in tumor suppressive and oncogenic function in thyroid cancer cells and circulating blood lncRNA have great potential as biomarker to detect thyroid cancer(Sedaghati & Kebebew, 2019). Studies by Yang and Deng (2014); Zou et al. (2015) revealed that lncRNAs participated in the modulation of cell proliferation, differentiation, migration and invasion in head and neck cancer by functioning as oncogenes and tumor suppressors. Other studies also confirmed that dysregulation of lncRNA implicated in glioblastoma, breast cancer, colorectal cancer, liver cancer and leukemia (Fang & Fullwood, 2016). Interestingly, in clear cell renal cell carcinoma, lncRNA MIR4435-2HG facilitate the malignant progression via miR-513a-5p/KLF6 axis (Zhu et al., 2020). lncRNA SNHG16 promotes migration and invasion by inhibiting CDKN1A (Liu et al., 2020). However, lncRNA lnc-DILC inhibit the tumor progression (Zhang et al., 2019a).

Recent works by many investigators have shown that lncRNA is strongly linked to various cancer development and can be effectively detected, thereby lncRNAs may be a novel class of cancer biomarkers (Bolha et al., 2017). For instance, Zhong et al. (2017) established a six-lncRNA model predicting the clinical outcome of ER-positive breast cancer patients by analyzing lncRNA expression profiles of more than 600 patients from TCGA. Chen et al. (2017) used differential expression analysis and functional analysis to obtain a novel four-lncRNA signature which provided a reliable theoretical basis of molecular mechanisms of gliomas. Moreover, Zhang et al. (2019b) found an 11-lncRNA signature indicating the potential biochemical functions of 11 selected lncRNAs in ccRCC and demonstrated LINC00488 and HOTTIP promote tumour proliferation.

2.5 The relationship between ferroptosis and long noncoding RNAs

Several studies have reported that lncRNA is important for regulating ferroptosis. A research by Wang et al. (2019) demonstrated that lncRNA LINC00336 serves as a competing endogenous RNA to inhibit ferroptosis in lung cancer. Similarly, a research by Lu et al. (2020) revealed that lncRNA PVT1 regulated ferroptosis via miR-214/TFR1/TP53 axis. In recent research, Ma et al. (2021) proved that silencing lncRNA MEG8 induces the ferroptosis and inhibits the proliferation of hemangioma endothelial cells by regulating miR497-5P/NOTCH2 pathway. Wu and Liu (2021) explained that targeting long noncoding RNA NEAT1 or ACSL4 may be a viable treatment for non-small-cell lung cancer by proving that NEAT1 regulates ferroptosis and regulates ferroptosis sensitivity based on ACSL4. To date, what is not yet clear is the impact of FRLs from sequence data on the overall survival in ccRCC patients.

2.6 The relevant work about ferroptosis-related long noncoding RNAs signature

Several ferroptosis-related lncRNAs signatures have been developed. For instance, Cai et al. (2021) constructed a signature consisting of seven ferroptosis-related lncRNAs (LINC01503, AC004687.1, AC010973.2, AP001189.3, ARRDC1-AS1, OIP5-AS1, and NCK1-DT) for colon cancer to provide individualized predictions for patients' prognosis. Tang et al. (2021) revealed a signature based on 25 ferroptosis-related lncRNAs impacts on the prognosis of head and neck squamous cell carcinoma and the signature as well as tumor stage are independent prognosis factors of overall survival. Interestingly, Chen et al. (2021a) found a signature comprising 20 lncRNAs have potential to be diagnostic and prognostic biomarkers for gastric cancer. Study by Zheng et al. (2021) constructed a model involving 10 ferroptosis-related lncRNAs that associated with the immune response, providing novel insights into finding new therapies for lung adenocarcinoma.

2.7 Tumor microenvironment

The characteristics of the tumor microenvironment seriously influence the body of the disease as well as influence the systemic therapy response (Vuong et al., 2019). Tumor microenvironment is comprised of cancer stem cells, cancer-associated endothelial cells, cancer-associated fibroblasts and infiltrated immune cells, all of which are involved in complex crosstalk with tumor cells, thus influencing tumor progression. Various infiltrated immune cells work together to help cancer cells escape immune surveillance by acting as an important part of TME, thereby forming a tumor-promoting microenvironment for proliferation and metastasis of cancer cells (Zhou et al., 2020). ccRCC has been confirmed to be a highly immune infiltrated tumor based on multiple clinical and genomic studies (Şenbabaoğlu et al., 2016). Growing numbers of studies highlighted TME in relevant research of ccRCC due to its important role in immune surveillance. For instance, Xu et al. (2019) used ESTIMATE algorithm to reveal the correlation of TME and ccRCC prognosis and precision immunotherapy. Pan et al. (2020) found that dendritic cells resting, dendritic cells activated, mast cells resting, mast cells activated, and eosinophils were correlated with favorable prognosis, whereas B cells memory, T cells follicular helper and T cells regulatory (Tregs) were related to poorer outcome by integrating gene expression profiles of ccRCC from TCGA and GEO.

2.8 Data Science Techniques

2.8.1 Statistics

Principal Component analysis is a widely used technique which increase interpretability but minimize information loss for dimensionality reduction of datasets(Jolliffe & Cadima, 2016). It is a crucial step for genomic data performing quality control. Benjamini & Hochberg method is a practical and powerful procedure to control the false discovery rate(Benjamini & Hochberg, 1995). It is popular applied for identifying differentially expressed genes in bioinformatics. The Kaplan-Meier curve with log- rank test was often used to estimate the probability of survival since the log-

rank test is considered a nonparametric test and makes no assumptions about the shape of the survival curve(Koletsi & Pandis, 2017). There are a lot of studies have adopted log-rank test to compare the OS between the high-risk and low risk groups. For example, Wu et al. (2020b) and He and Zuo (2019) used log-rank test to assess survival difference between high-risk and low-risk groups. Pearson correlation coefficient is a measure of the strength of statistical relationship or association(Schober et al., 2018). It is widely used in measuring how strong a relationship between two variables. For instance, Zhang et al. (2021b) used Pearson correlation to investigate autophagy-related lncRNAs between lncRNAs and autophagy-related genes setting criteria of correlation coefficient > 0.3 and $P < 0.001$. Another work by Wang et al. (2021) screen out 765 immune-related lncRNAs using Pearson correlation for establish an immune-related lncRNA risk model.

2.8.2 Data Science

CIBERSORT algorithm is a computational approach that identify Cell-type by estimating relative subsets of RNA transcripts. It can accurately calculate the relative fractions of diverse cell subsets in gene expression profiles of complex tissues(Newman et al., 2015). In malignant tumors, the level of infiltrating immune cells is closely related to tumor growth, progression and patient outcome(Hanahan & Weinberg, 2011). Thereby it is popular to use CIBERSORT algorithm to explore the infiltration pattern of immune cells in tumor microenvironment. For example, Wu et al. (2020a) extracted infiltrating percentage of 22 immune cells from 27 normalized datasets of prostate cancer, the result show that infiltrating M1 macrophages and neutrophils are associated with the prognosis of patients and supported that M1 macrophages and neutrophils could be potential targets for patient's diagnosis and prognosis of treatment. Another work by Mo et al. (2020) focused on using CIBERSORT algorithm to investigate the correlation between signature and immune cells, suggesting that memory activated CD4+ T cell and M0 macrophages had a significant infiltration in high-risk group patients.

2.9 Summary

Form the literature review we found that lncRNA is of great value in cancer research as it is strongly related to cancer development. Ferroptosis as a new form of cell death provide a new direction for cancer treatment. The signature based on ferroptosis-related lncRNAs for predicting prognosis of patients have been established for colon cancer, head and neck squamous cell carcinoma, gastric cancer, and lung adenocarcinoma but ccRCC. In this study we aim to develop a FRLs- based signature for predicting prognosis of ccRCC patients and explore the role of the signature in tumor microenvironment (TME).

CHAPTER 3: MATERIALS AND METHODS

3.1 Introduction

This chapter describes the methodology. It begins with the collection of relevant data, data quality control, identification of the ferroptosis-related lncRNAs, following by the development and validation of the ferroptosis-related lncRNA prognostic signature. Then, gene set enrichment and statistical analysis was presented in the end of the chapter. The specific flowchart of our study is displayed in Figure 1.

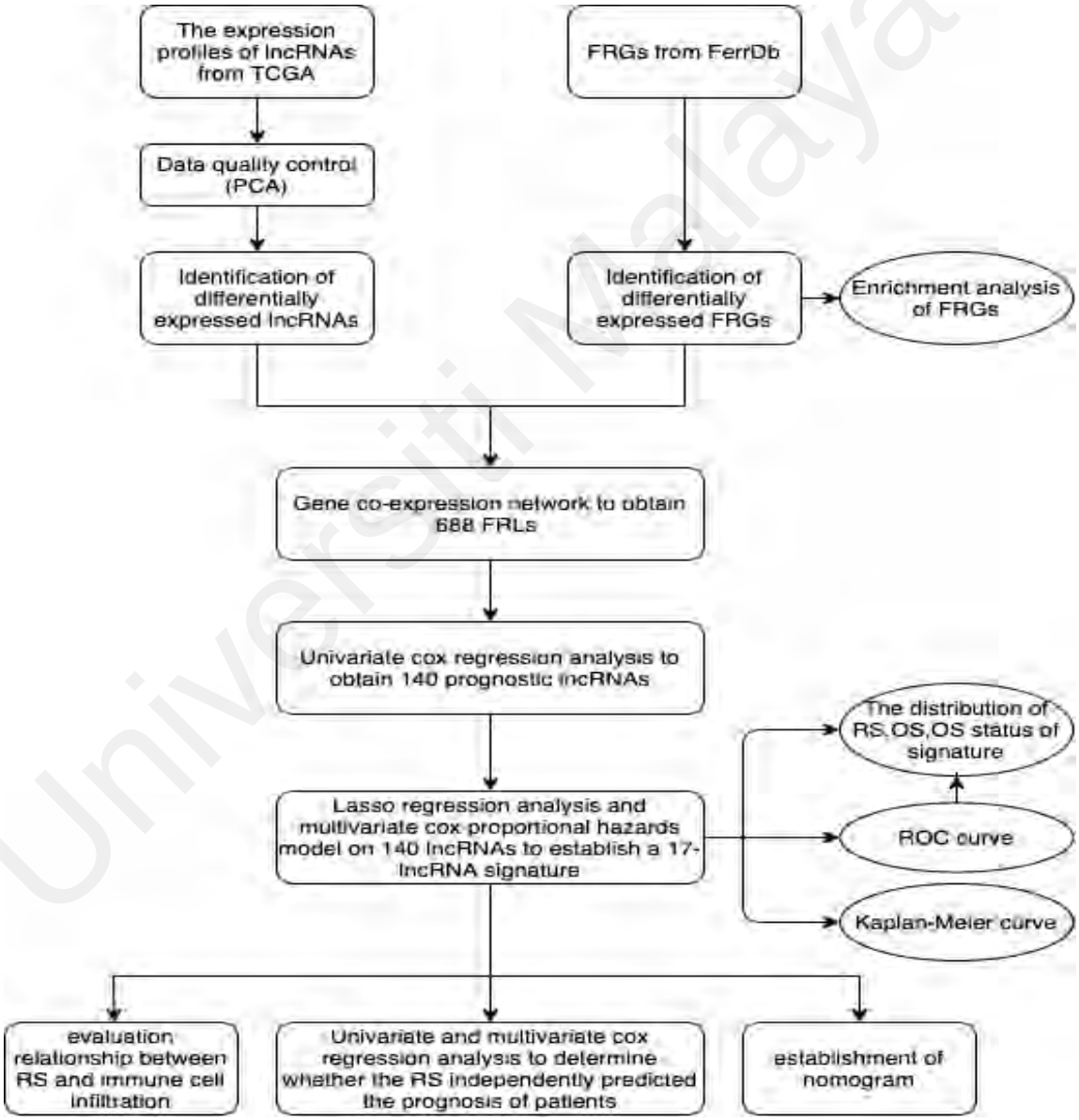

Figure 1: The flowchart of research

3.2 Data collection

The level 3 RNA-Seq transcriptome data of patients with ccRCC and clinically relevant data were downloaded from The Cancer Genome Atlas (TCGA) GDC data portal (**https://portal.gdc.cancer.gov/**). The data comprised of 539 tumor samples and 72 normal samples.

The Genome Reference Consortium Human Build 38 (GRCh38) annotation file for long noncoding RNA was derived from the GENCODE website (https://www.gencodegenes.org/human/). 14086 lncRNAs were identified in the TCGA dataset according to the Ensemble IDs. 259 ferroptosis-related genes (Driver: 108; suppressor: 69; marker: 111) were obtained from FerrDb (Zhou & Bao, 2020), a database that provide comprehensive information of ferroptosis regulators and markers and ferroptosis-disease associations. Immune infiltration data was derived from CIBERSORT (https://cibersort.stanford.edu/) which include 22 types of tumor-infiltrating immune cells, referring to B cells naive, B cells memory, Plasma cells, T cells CD8, T cells CD4 naive, T cells CD4 memory resting, T cells CD4 memory activated, T cells follicular helper, T cells regulatory (Tregs), T cells gamma delta, NK cells resting, NK cells activated, Monocytes, Macrophages M0, Macrophages M1, Macrophages M2, Dendritic cells resting, Dendritic cells activated, Mast cells resting, Mast cells activated, Eosinophils, and Neutrophils, as mentioned in Zhang et al. (2021a).

3.3 Data quality control

Patients with incomplete recording of clinical information including age, gender, tumor grade, tumor stage and survival or OS < 30 days were excluded. We used Principal Component Analysis (PCA) to conduct data quality control. It can be obviously observed that one sample is abnormal by drawing PCA plot (Figure 2A). Then abnormal sample was removed, and the tumor samples and normal samples are seen clearly separated

(Figure 2B). After data cleaning, 501 patients were selected for further analysis. The clinical characteristics of patients are displayed in Table S1.

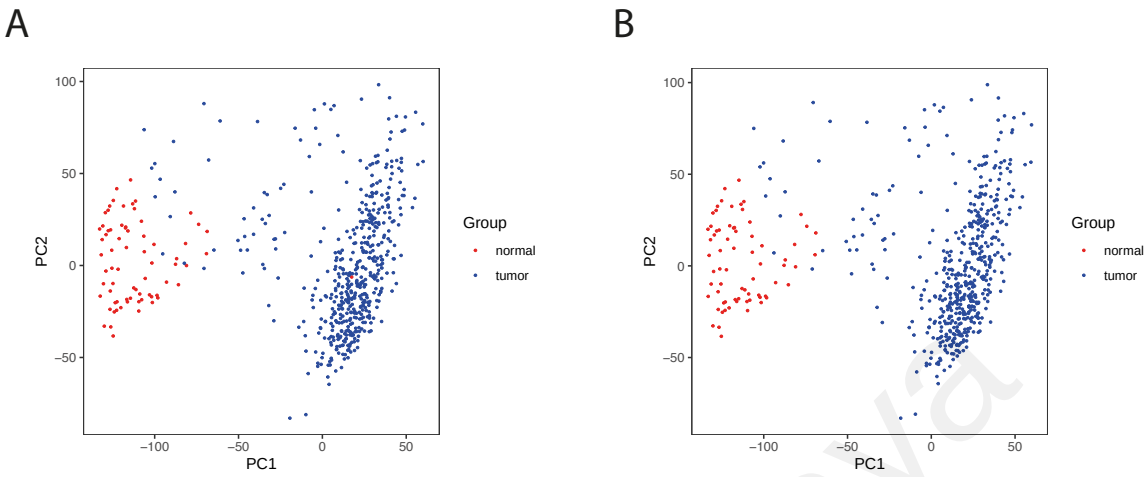

Figure 2: PCA plot for quality control.

3.4 Identification of ferroptosis-related lncRNAs

The Limma package (Ritchie et al., 2015) was adopted for recognizing significant differential expressed ferroptosis-related genes (FRGs) and differential expressed lncRNAs (DELs) between ccRCC tissues and healthy tissues according to log2FC. Subsequently, biological pathways associated with FRGs were assessed using "clusterProfiler" package (Yu et al., 2012) to perform Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG), with the inclusion criteria of P value < 0.05 and q value <0.05.

Co-expression analysis was then performed between FRGs and DELs based on Pearson correlation analysis. Following the study by Liang et al. (2021), a cut-off of Pearson correlation coefficient > 0.3 and P value < 0.001 for lncRNA was perceived as FRL.

3.5 Development and validation of the ferroptosis-related lncRNA prognostic signature

We first screened prognosis-related lncRNAs (P value < 0.001) by univariate cox regression analysis. The 501 patients were first randomly stratified into training and validation (1st validation) dataset at the ratio of 5:5 using the "caret" package (Kuhn, 2020). Subsequently, these 501 patients were then randomly divided into two validation datasets (2nd and 3rd validation dataset at the ratio of 7:3). The training and validation datasets were used for constructing and testing the FRL-related prognostic risk signature, respectively. All the FRLs were used in the subsequent least absolute shrinkage and selection operator (Lasso) analysis. After filtrating by Lasso analysis, a risk model from the selected lncRNAs was constructed by multivariate Cox proportional hazards model. The coefficients obtained from multivariate Cox proportional hazards model were utilized to produce the following risk score (RS) equation: RS = coefficient a × expression level of lncRNA a + coefficient b × expression level of lncRNA b + …… + coefficient n × expression level of lncRNA n. Based on this equation, the RS per ccRCC patient was independently calculated in the training dataset and validation datasets. Finally, the ccRCC patients were assigned to high- and low-risk groups by the median value of the RS.

Kaplan–Meier survival curves were used to assess the predictive power of the FRLs using "survival" package (Therneau & Grambsch, 2000) and "survminer" (Kassambara et al., 2021) package. To evaluate the predictive accuracy of the FRLs, receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) were computed by "survivalROC" package (Heagerty & Saha-Chaudhuri, 2013). Univariate and multivariate analysis were implemented to verify the independent prognostic factor. Nomogram was further established by package "rms" (Jr, 2021), for predictive of the probable 1-, 3-, and 5-year survival of the ccRCC patients. To further examine the effect

of the signature on the tumor immune microenvironment (TME) of ccRCC, we estimated the immune infiltrate level between high- and low-risk groups.

3.6 Statistical Analysis

Data was processed using Bioconductor packages in R software (version 4.0.5, http://www.R-project.org**)** in our study. PCA was applied to data quality control in our study. Benjamini & Hochberg method was used to identify the differently expressed FRGs and DELs, based on FDR. The sensitivity and specificity of FRLs-based prognostic signature compared with other clinicopathological characteristics was evaluated using ROC curve. The Kaplan–Meier method and log- rank test was used to compare the OS between the high-risk and low-risk groups. Pearson correlation test was used to identify FRLs. Clinicopathological characteristics were compared within the training and 1st validation datasets, 2nd and 3rd validation dataset, using the Chi-square test.

CHAPTER 4: RESULTS

4.1 Introduction

Chapter four is the result of the research. The enrichment analysis of ferroptosis-related, the result of Ferroptosis-related lncRNAs in ccRCC, the construction of ferroptosis-related lncRNAs signature, the validation process of the prognostic score, the Nomogram establishing and clinical utility of the risk score and the gene set enrichment analysis were presented in this chapter.

4.2 Enrichment analysis of ferroptosis-related genes

According to the criteria of |log2FC|>1 and FDR<0.05, we found 77 FRGs (37 upregulated and 40 downregulated) (Table S2). Through the KEGG analysis, the FRGs were mainly involved in HIF-1 signaling pathway, MicroRNA in cancer, Ferroptosis, PD-L1 expression and PD-1 checkpoint pathway in cancer, IL-17 signaling pathway, renal cell carcinoma, pancreatic cancer, bladder cancer (Figure 3A; Table S3). Biological Process (BP) regulated response to oxidative stress, cellular response to chemical stress and reactive oxygen species metabolic process. Cellular Component (CC) mainly participated in apical part of cell, organelle outer membrane and basolateral plasma membrane. Molecular Function (MF) was enriched in iron ion binding, ferric iron binding and oxidoreductase activity, acting on NADPH (Figure 3B; Table S3).

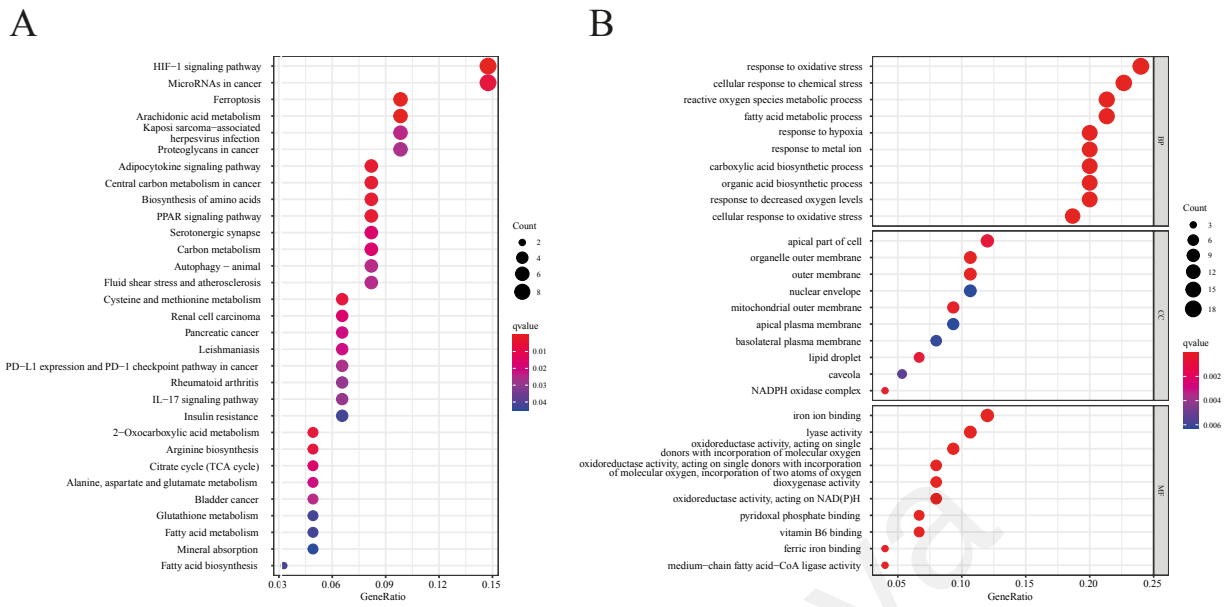

Figure 3: KEGG and GO analysis for FRGs. (A) KEGG and (B) GO.

4.3 Ferroptosis-related lncRNAs in ccRCC

956 DELs were uncovered by setting the cutoff of |log2FC|>2 and FDR<0.05. DELs were displayed in volcano plot via package "ggplot2" (Wickham, 2016) (Figure 4A). Among these FRGs and DELs, 688 FRLs were confirmed by co-expression analysis (Pearson correlation coefficient > 0.3 and P value < 0.001) (Figure S1).

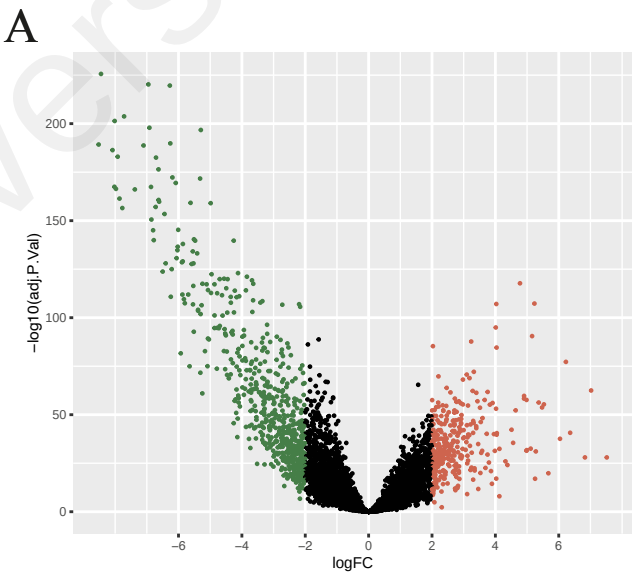

Figure 4: Volcano plot displaying the differentially expressed lncRNAs between ccRCC and normal tissue samples where upregulated lncRNAs are represented by red dots, downregulated represented by green dots and black dots represents lnRNAs with insignificant difference.

4.4 Construction of ferroptosis-related lncRNAs signature

Univariate cox regression analysis was fulfilled for FRLs, and the result showed that 140 lncRNAs were significantly associated with the OS of ccRCC ($P < 0.001$) (Figure S2).

To further explore the prognostic predictive effect of the lncRNA in ccRCC patients, we conducted the LASSO regression analysis and multivariate Cox proportional hazards model on 140 lncRNAs in the training cohort. Initially, the lncRNAs expression data was merged with survival data of each patient. The baseline clinicopathological features of the training cohort and three validation cohorts, were summarized in Table 1A and Table 1B, separately. There is no statistical difference in clinical characteristics (age, gender, grade, stage) among the different cohorts, with $P > 0.05$. The prognostic risk signature was established using the training dataset and was validated using three validation datasets. The Lasso regression analysis was first utilized to identify the most significant lncRNAs by selecting the optimal penalty parameter λ correlated with the minimum 10-fold cross-validation (Figure 5A and B). The Multivariate Cox Regression model further yielded seventeen optimal prognostic FRLs (Figure 5C). Among them, 10 lncRNAs (AC008742.1, AC010980.2, AC011700.1, AC084876.1, AC090337.1, AC139491.2, LINC01271, MANCR, PRKAR1B−AS1, TMEM246−AS1) are risk factors, 7 lncRNAs (AC004066.1, AC005722.3, AC007406.3, AC093583.1, AL928921.1, LINC02073, PSORS1C3) are protective factors, as shown in the Sankey diagram (Figure 5D), which reveal the association between prognostic FRLs, ferroptosis-related genes, and risk types. The RS equation was calculated as: RS = (-0.1779 × AL928921.1 expression) + (0.1840 × AC011700.1 expression) + (0.1974 × AC008742.1 expression) – (0.2883 × AC007406.3 expression) + (0.1428 × AC090337.1 expression) + (0.2899 × LINC01271 expression) – (0.2070 × AC005722.3 expression) + (0.2169 × PRKAR1B-AS1 expression) – (0.1990 × AC004066.1 expression) + (0.1480 × MANCR expression) –

(0.1458 × AC093583.1 expression) – (0.1822 × PSORS1C3 expression) + (0.3154 × AC084876.1 expression) + (0.1754 × AC010980.2 expression) – (0.1401 × LINC02073 expression) + (0.1127 × AC139491.2 expression) + (0.3129 × TMEM246-AS1 expression). As shown in Figure 6, the distribution of the RS, OS status, and expression profiles of the signature based on 17 FRLs was displayed in the training and validation cohorts. In the training cohort, the high-risk groups had evidently higher value of risk score (Figure 6A) and lower survival rate (Figure 6B). Moreover, with the risk score increasing, the expression of protective lncRNA (AL928921.1, AC007406.3, AC005722.3, AC004066.1, AC093583.1, PSORS1C3, LINC02073) decreased, whereas those of risk lncRNA (AC011700.1, AC008742.1, AC090337.1, LINC01271, PRKAR1B-AS1, MANCR, AC084876.1, AC010980.2, AC139491.2, TMEM246-AS1) increased（Figure 6C）. Similar results were obtained in 1st validation (Figure 6D, 6E, 6F), 2nd validation (Figure 6G, 6H, 6I), 3rd validation (Figure 6J, 6K, 6L) cohorts.

Table 1: The baseline clinicopathological features for training dataset and 1st validation dataset (A), 2nd validation dataset and 3rd validation dataset (B).

Covariates	Type	Total	Test	Train	P value
Age	<=65	332(66.27%)	103(69.13%)	229(65.06%)	0.4368
	>65	169(33.73%)	46(30.87%)	123(34.94%)	
Gender	Female	172(34.33%)	49(32.89%)	123(34.94%)	0.7336
	Male	329(65.67%)	100(67.11%)	229(65.06%)	
Grade	Grade 1-2	228(45.51%)	69(46.31%)	159(45.17%)	0.892
	Grade 3-4	273(54.49%)	80(53.69%)	193(54.83%)	
Stage	Stage I-II	304(60.68%)	89(59.73%)	215(61.08%)	0.8553
	Stage III-IV	197(39.32%)	60(40.27%)	137(38.92%)	

(A)

Covariates	Type	Total	Test	Train	P value
Age	<=65	332(66.27%)	172(69.08%)	160(63.49%)	0.2197
	>65	169(33.73%)	77(30.92%)	92(36.51%)	
Gender	Female	172(34.33%)	76(30.52%)	96(38.1%)	0.0909
	Male	329(65.67%)	173(69.48%)	156(61.9%)	
Grade	Grade 1-2	228(45.51%)	115(46.18%)	113(44.84%)	0.8319
	Grade 3-4	273(54.49%)	134(53.82%)	139(55.16%)	
Stage	Stage I-II	304(60.68%)	155(62.25%)	149(59.13%)	0.5327
	Stage III-IV	197(39.32%)	94(37.75%)	103(40.87%)	

(B)

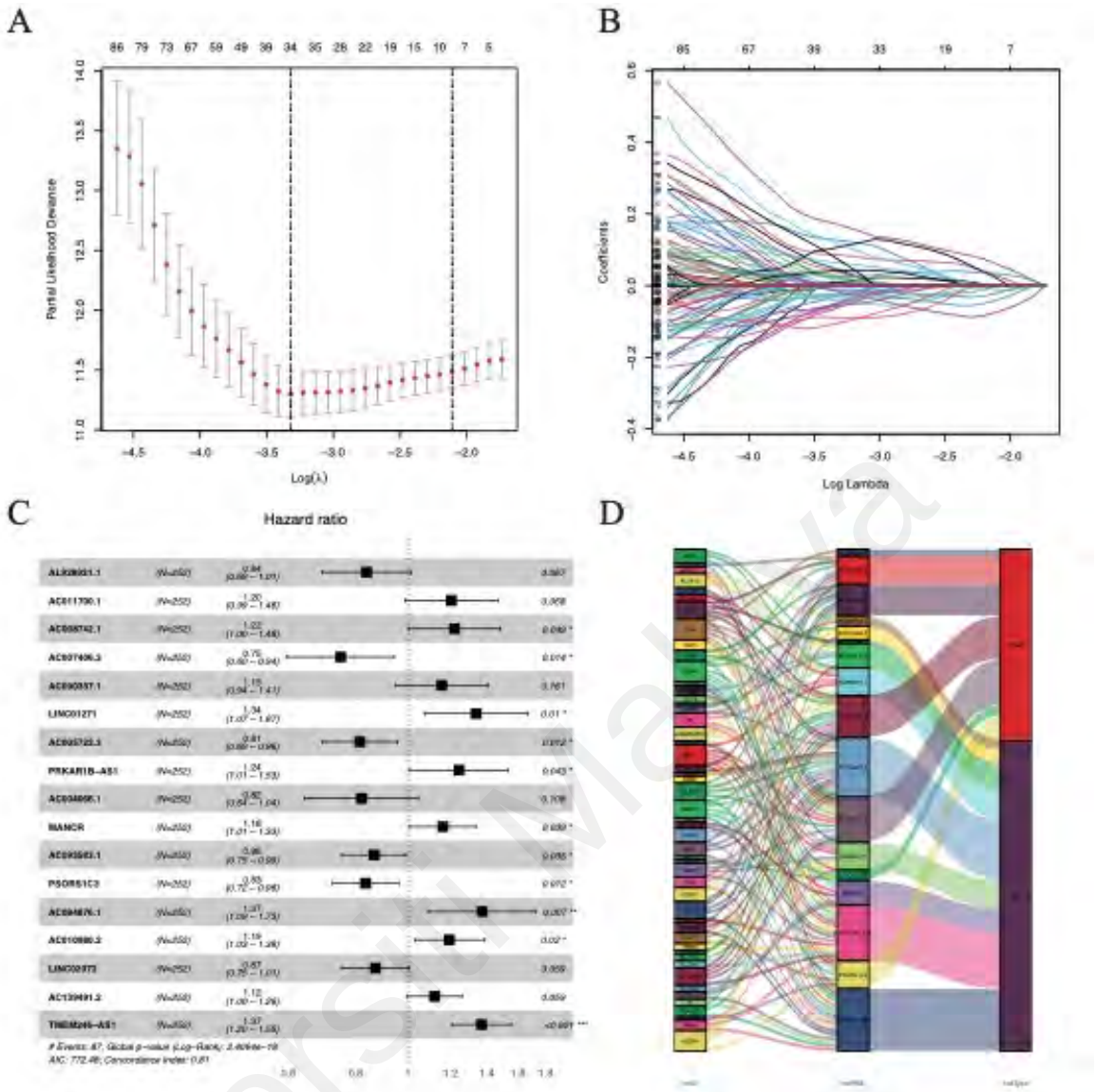

Figure 5: (A) Partial likelihood deviance for tuning parameter selection in the Lasso analysis in the training dataset. (B) Lasso coefficient profiles in the Lasso analysis in the training dataset. (C) The Hazard Ratio (HR) value and its 95% confidence interval with associated p-value of the multivariate Cox proportional hazards model were showed in the forest plot. HR >1 represents the high expressions of lncRNA were unfavourable for prognosis, HR <1 indicates the high expressions of lncRNA were favourable for prognosis. (D) The Sankey diagram showed the association between prognostic FRLs, ferroptosis-related genes, and risk types. The lncRNAs linked to red is protective lncRNAs, linked to dark purple represents risk lncRNAs.

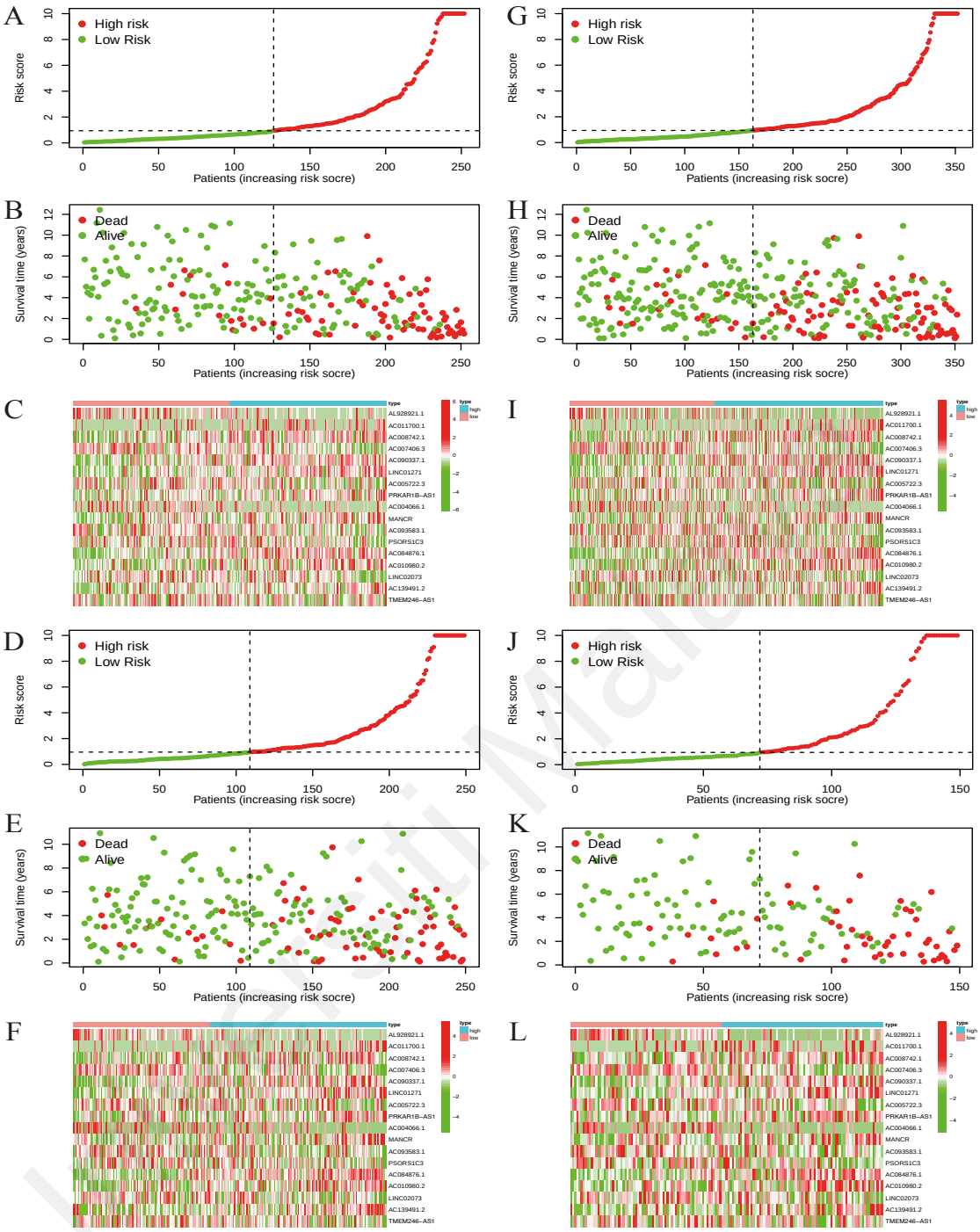

Figure 6: The RS distribution, survival status, and lncRNA expression in the datasets. (A, B, C) Training dataset, (D, E, F) 1st validation dataset, (G, H, I) 2nd validation dataset, and 3rd validation dataset (J, K, L).

4.5 Validation of the prognostic score

To assess the prognostic prediction accuracy of the signature, we performed ROC in the training dataset and validation datasets. As presented in Figure 7A to D, the 3- and 5-year survival rates were 0.829 and 0.851 in the training dataset and were 0.751 and 0.755

in the 1st validation dataset, the AUCs for the 3- and 5-year survival prediction were 0.751 and 0.755. Similar trend was found in the 2nd validation dataset (Figure 7E, F) and 3rd validation dataset (Figure 7G, H). These results showed that our signature had an excellent performance for the prognosis of patients with ccRCC. The survival analysis was also performed for training dataset and validation datasets.

The Kaplan–Meier curve in the training dataset revealed poorer survival in the high-risk group than in the low- risk group ($p < 0.001$) (Figure 7I). Likewise, the same tendency was discovered in the validation datasets with all P value < 0.001(1^{st} validation dataset: Figure 7J; 2^{nd} validation dataset: Figure 7K; 3^{rd} validation dataset: Figure 7L). Taken together, the results showed that the RS based on the prognostic risk signature could accurately indicate the prognosis of ccRCC patients.

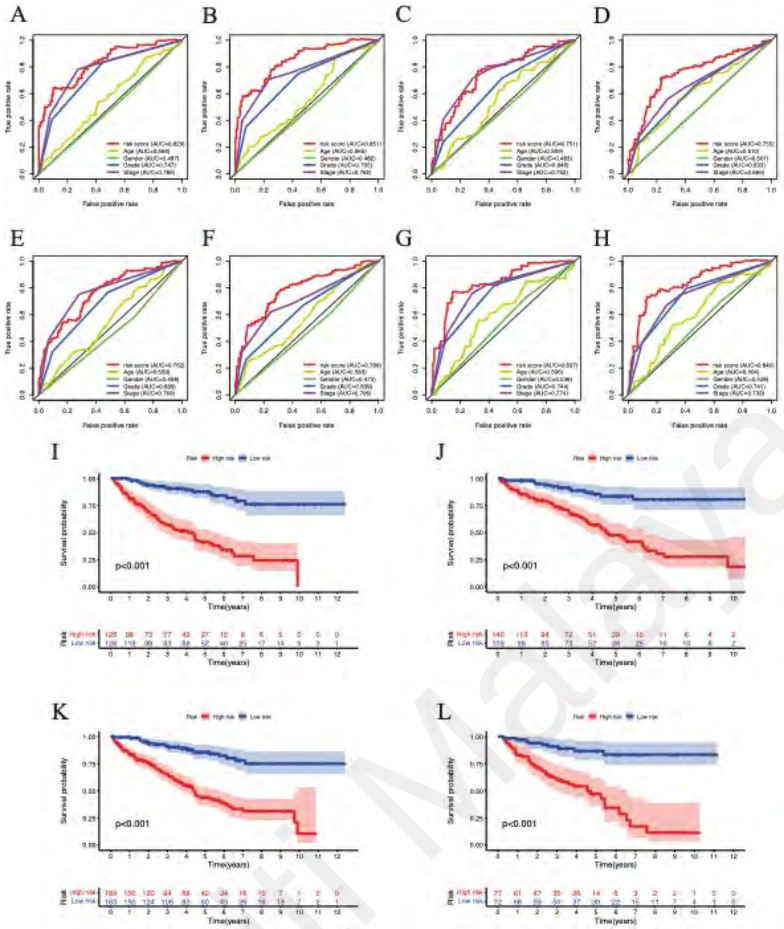

Figure 7: Kaplan–Meier curves of overall survival for the high-risk and low-risk groups and ROC curve for 3- and 5-year for predicting survival in the datasets. (A)(B) ROC for 3- and 5 -year in the training dataset. (C)(D) ROC for 3- and 5 -year in the 1st validation dataset. (E)(F) ROC for 3- and 5 -year in the 2nd validation dataset. (G)(H) ROC for 3- and 5 -year in the 3rd validation dataset. (I) The training dataset. (J) 1st validation dataset. (K) 2nd validation dataset. (L) 3rd validation dataset.

To determine the prognostic values of the RS and various clinicopathological factors in ccRCC, uni- and multi- Cox regression analyses were performed on each cohort. Uni-analysis indicated that age ($p = 0.004$), stage ($p < 0.001$) and risk score ($p < 0.001$) have significant effect on the OS in the training cohort (Figure 8A). Subsequently, these factors were also included into multi-Cox regression analysis, which further confirm age ($p = 0.014$), stage ($p < 0.001$) and risk score ($P < 0.001$) as independent prognostic factors (Figure 8B). Simultaneously, the same tendency was acquired in the three validation ohorts (Figure 8C-H).

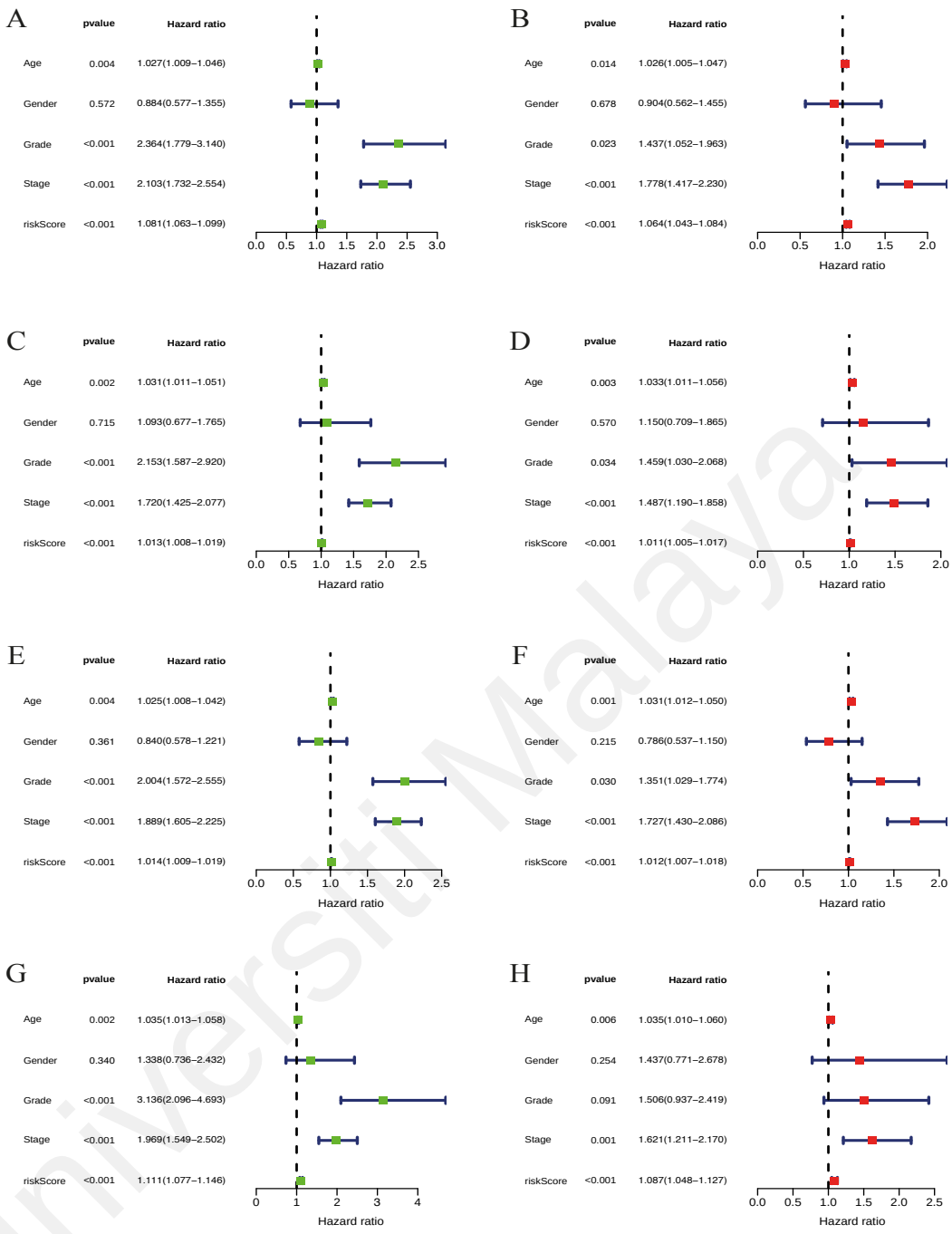

Figure 8: Forest plots of the univariate and multivariate Cox regression analysis indicated that the RS, age and stage were independent risk factor for OS in the training dataset (A, B), 1st validation dataset (C, D), 2nd validation dataset (E, F) and the 3rd validation dataset (G, H).

4.6 The Nomogram Establishing and Clinical utility of the Risk Score

For studying the 1-, 3-, and 5-year prognosis of the patients with ccRCC, a nomogram was plotted using the training dataset by integrating the independent prognostic factors (age, stage, risk score) (Figure 9A). Interestingly, the same tendency was acquired in the validation dataset (Figure S3). Using the nomogram, the 1-, 3-, and 5-year survival rates could be predicted by the corresponding value of total points based on the independent prognostic factors (Zhang et al., 2021a).

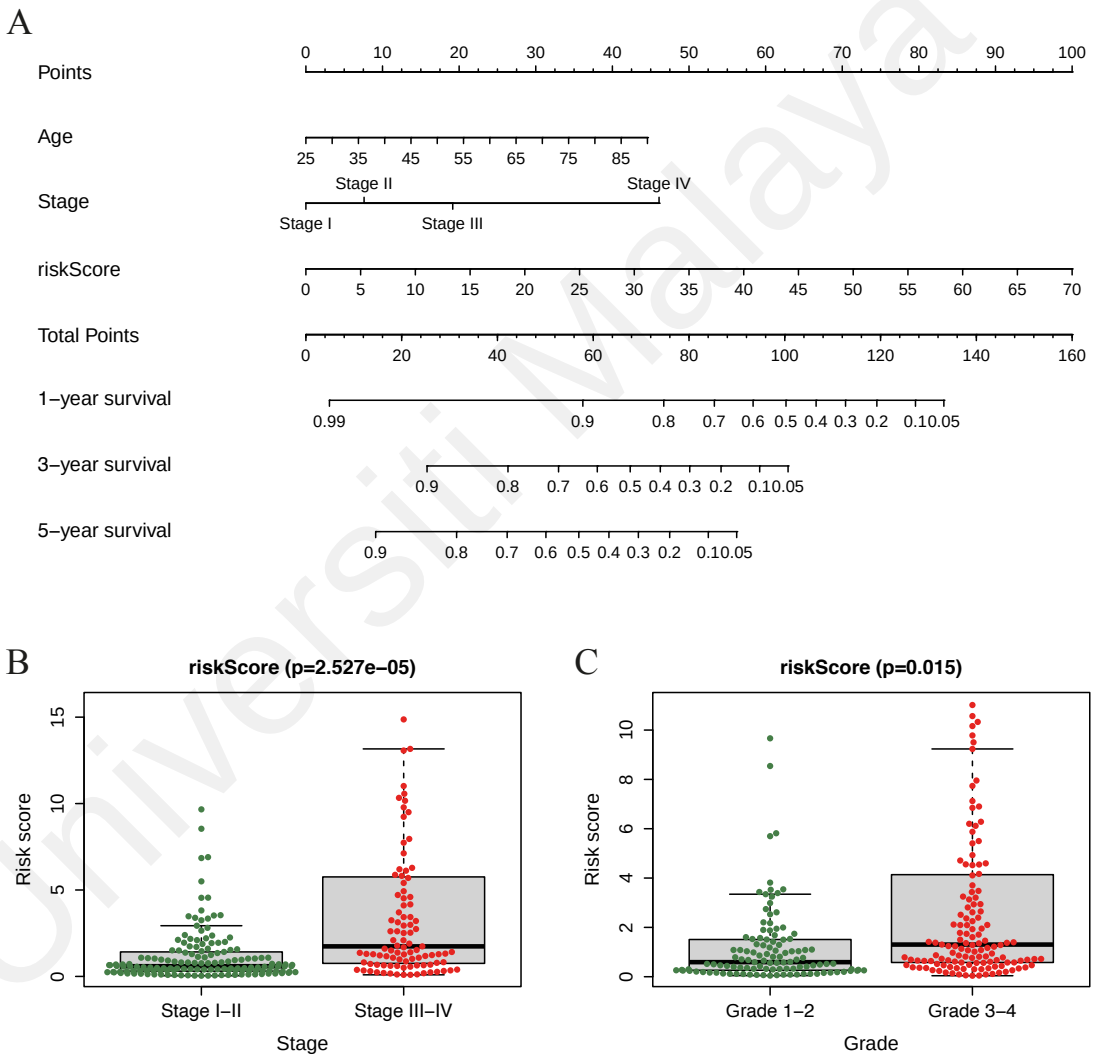

Figure 9: (A) A nomogram plot was built to qualify risk assessment for ccRCC patients. (B) Relationships between the risk score and tumor stage in clear cell renal carcinoma. (C) Relationships between the risk score and tumor grade in clear cell renal carcinoma.

We further explored the relationships among ten risk lncRNAs (AC011700.1, AC008742.1, AC090337.1, LINC01271, PRKAR1B-AS1, MANCR, AC084876.1,

AC010980.2, AC139491.2, TMEM246-AS1), RS, and clinicopathologic features (age, gender, grade, stage) (Table 2). The RS was found distinctly higher in advanced-stage tumor and higher-grade tumor (Figure 9B, C). the Same tendency were acquired in the validation cohorts (Table S4, Figure S4). This finding provides for that the risk score based on our signature can also reflect tumor progression.

In order to further explore the prognostic value of the 17 lncRNAs, the Kaplan Meier curve was plotted to confirm the relationship between these lncRNAs and OS. In our analysis, a total of 11 of the 17 lncRNAs (LINC01271, AC010980.2, AC011700.1, MANCR, AC008742.1, AC084876.1, AC090337.1, AC093583.1, LINC02073, AL928921.1, AC004066.1) were identified. The results indicated that the 11 ferroptosis-related lncRNAs were correlated to the OS in ccRCC patients (Figure 10).

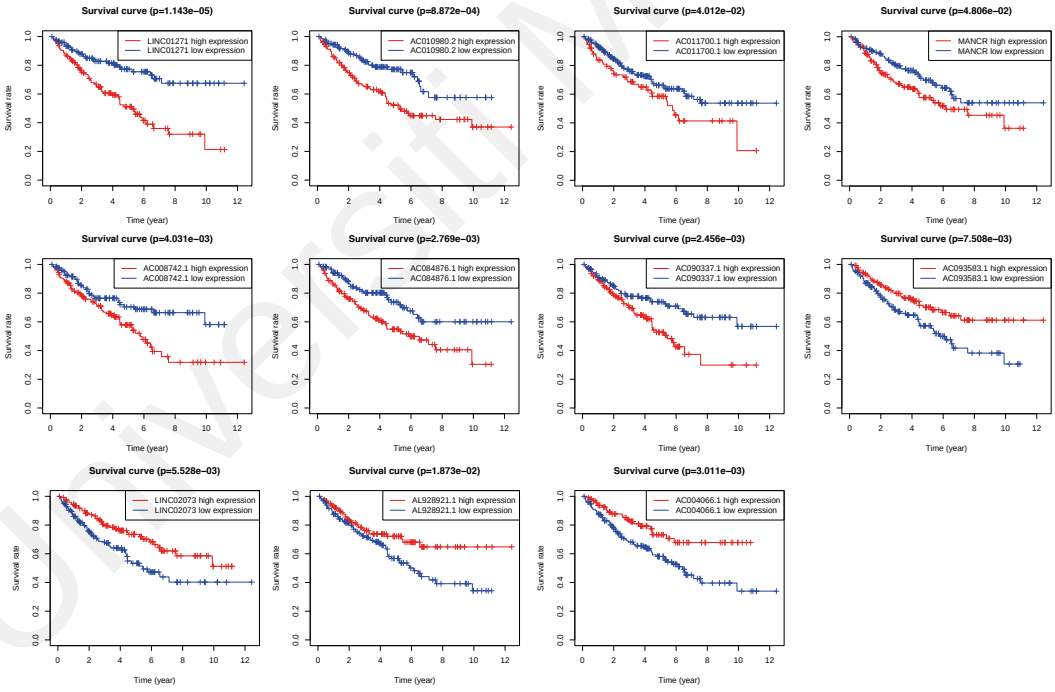

Figure 10: Validation the prognostic value of these 17 ferroptosis-related lncRNAs in clear cell renal cell carcinoma by Kaplan Meier curve.

Table 2: Association of the RS and the risk genes with clinical factors in ccRCC.

lncRNA	Age(<=65/>65)	Gender(Female/Male)	Grade(1&2/3&4)	Stage (I-II/III-IV)
AC011700.1	-1(0.318)	-0.862(0.390)	-0.629(0.530)	-0.551(0.582)
AC008742.1	0.199(0.843)	1.234(0.219)	-0.789(0.431)	-1.755(0.081)
AC090337.1	-0.237(0.813)	1.013(0.312)	-0.032(0.975)	-2.74(0.007)
LINC01271	-0.249(0.804)	-0.286(0.775)	-3.26(0.001)	-3.99(8.877e-05)
PRKAR1B-AS1	-0.493(0.623)	-2.477(0.014)	-0.819(0.414)	-4.076(6.363e-05)
MANCR	0.424(0.672)	-2.868(0.005)	-2.277(0.024)	-3.134(0.002)
AC084876.1	-1.76(0.080)	1.475(0.141)	-3.03(0.003)	-2.891(0.004)
AC010980.2	-0.904(0.367)	0.592(0.555)	-3.871(1.384e-04)	-3.893(1.388e-04)
AC139491.2	-0.867(0.387)	-1.227(0.221)	-1.338(0.182)	-1.111(0.268)
TMEM246-AS1	-0.353(0.725)	3.613(3.908e-04)	1.825(0.069)	1.325(0.186)
riskScore	-1.366(0.174)	0.828(0.409)	-2.46(0.015)	-4.407(2.527e-05)

We also evaluated the relationship between the RS and immune cell infiltration. At first, the immune landscape of all the samples was plotted (Figure 11A). Next, the number of immune cells which showed significant difference between the low- and high-risk groups were identified. Ten types of immune cells were identified with differences in infiltration between the two groups, namely, Plasma cells, T cells follicular helper, Tregs, Monocytes, Macrophages M0, Dendritic cells resting, Dendritic cells activated, Mast cells resting, Mast cells activated, Eosinophils (Figure 11B).

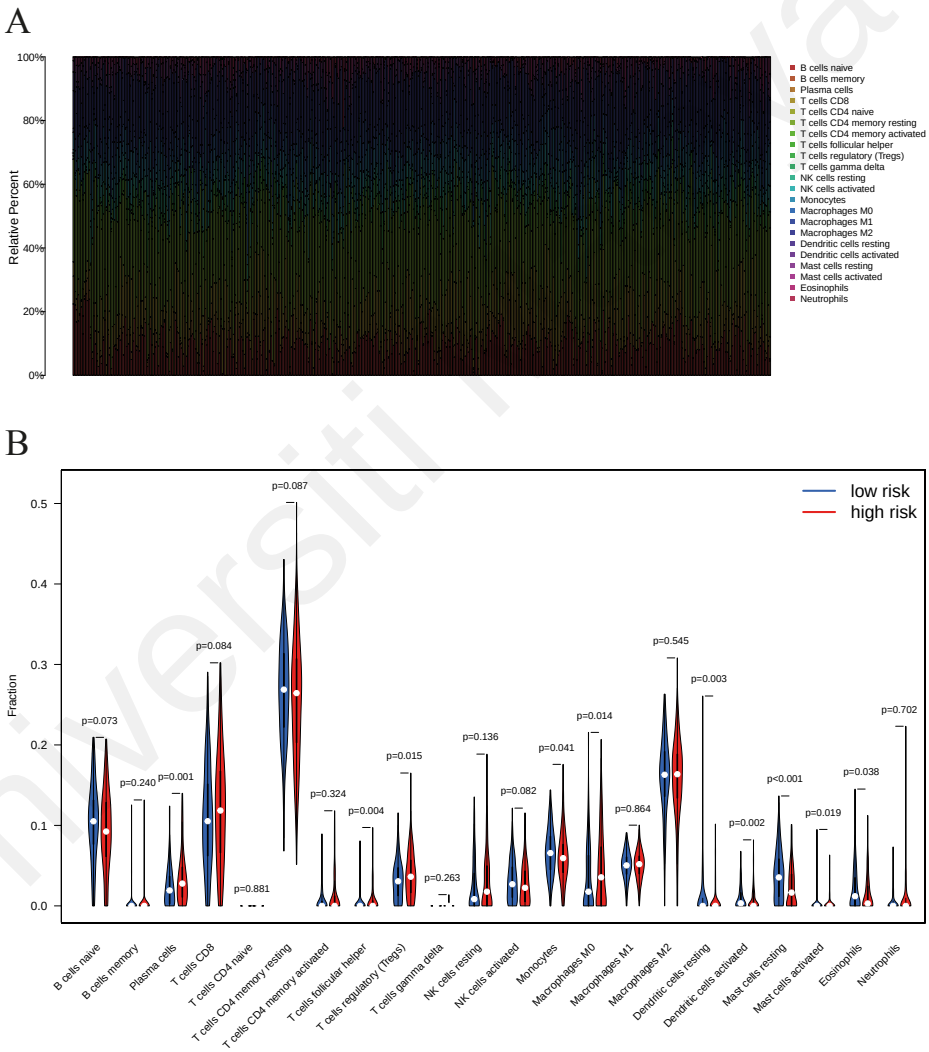

Figure 11: (A) The immune landscape of all ccRCC patients included in this study. (B) Relationships between the risk score and the immune cell infiltration in ccRCC patients. (A) The immune landscape of all ccRCC patients included in this study. (B) Relationships between the risk score and the immune cell infiltration in ccRCC patients.

CHAPTER 5: DISCUSSION AND CONCLUSION

For decades, the diagnosis and treatment in ccRCC patients was still based on clinicopathological factors (Ljungberg et al., 2015; Motzer et al., 2015). While patients may have similar clinical characteristics, the therapeutic effect and the prognosis of them has a massive gap. Hence, in this study, we explored the various techniques available in the data science to predict prognosis of ccRCC using ferroptosis-related lncRNAs as potential biomarkers. To the best of our knowledge, this is the first study that attempted to predict prognosis signatures of ccRCC based on FRLs.

This study was inspired by Lu et al. (2018) who highlighted the importance of further research in ferroptosis and its mechanism with regard to diagnosis of cancers. As aforementioned ferroptosis has been known to be involved in the progression of ccRCC (Li et al., 2020a). This is confirmed in our study which revealed 77 differential expressed ferroptosis-related genes. GO analysis show that most of the FRGs in Biological Process (BP) regulated response to oxidative stress, cellular response to chemical stress and reactive oxygen species metabolic process. Cellular Component (CC) mainly participated in apical part of cell, organelle outer membrane and basolateral plasma membrane. Molecular Function (MF) was enriched in iron ion binding, ferric iron binding and oxidoreductase activity, acting on NADPH. KEGG further revealed most of the FRGs participated in HIF-1 signaling pathway, MicroRNA in cancer, Ferroptosis, PD-L1 expression and PD-1 checkpoint pathway in cancer, IL-17 signaling pathway, renal cell carcinoma, pancreatic cancer, bladder cancer. A recent study by Li et al. (2020b) demonstrated that the achievement of FG-4592 (an inhibitor of prolyl hydroxylase of HIF) pretreatment is mainly based on decreasing ferroptosis at the early stage of FA-induced kidney injury via Akt/GSK-3β-mediated Nrf2 activation. Tang et al. (2020) reported that the IL-17 signaling pathway is a potential target affected by erastin (ferroptosis inducer),

which indicated that the ferroptosis inducer erastin may be regarded as a potential agent of cancer immunotherapy.

Several research have reported that lncRNAs play diverse roles in cancer (Carlevaro-Fita et al., 2020; Schmitt & Chang, 2016). For example, LncRNA BX357664 regulates cell proliferation through regulating TGF-b1/p38/HSP27 axis in RCC (Liu et al., 2016). LncRNA SNHG11 facilitates tumor metastasis by interacting with and stabilizing HIF-1α (Xu et al., 2020). LncRNA HANR promotes tumorigenesis in hepatocellular carcinoma (Xiao et al., 2017). In this study, we identified 956 DELs in ccRCC. In accordance with the present results, our studies demonstrated that lncRNAs are strongly associated with the malignancy in ccRCC. Moreover, lncRNAs have been reported to have important roles in ferroptosis. Mao et al. (2018) illustrated that lncRNA P53RRA can directly interact with the functional domain of signaling proteins in the cytoplasm, thereby modulating p53 modulators to suppress cancer progression. Yang et al. (2020) reported that silencing of lncRNA ZFAS1 attenuated ferroptosis by functioning as ceRNA. In our study, we implemented a co-expression analysis among FRGs and DELs, thus 688 lncRNAs were identified as FRLs. The result showed strong link between FRGs and FRLs in ccRCC samples, suggesting that FRLs are related to the tumorigenicity of ccRCC.

Seventeen lncRNA out of all FRLs, referring to AC008742.1, AC010980.2, AC011700.1, AC084876.1, AC090337.1, AC139491.2, LINC01271, MANCR, PRKAR1B−AS1, TMEM246−AS1, AC004066.1, AC005722.3, AC007406.3, AC093583.1, AL928921.1, LINC02073, PSORS1C3, associated with prognosis independently and hence were used as the prognostic signature. ROC curve (AUC at 3 years:0.829; AUC at 5 years:0.851) in training dataset and in three validation datasets with similar results confirmed excellent specificity and sensitivity of our prognostic signature. Survival curves with p value < 0.001 in each dataset exhibited good efficacy

of our signature in stratifying patients into high and low risk of mortality. Univariate and multivariate Cox analysis further demonstrated age, stage and risk score were independent prognostic factors. We also verified the effect of our risk score in the patients with same tumor stage and same tumor grade, which we can see the risk score of stage III-IV obviously higher than stage I-II, the risk score of grades 3-4 distinctly higher than grade 1-2. All of the analyses show that our ferroptosis-related lncRNA signature may be a beneficial supplement for better stratifying patients and for providing a more individualized treatment method. We further integrated three independent prognostic factors (age, stage, risk score) to develop a nomogram for calculating points which could reflect survival.

Ferroptosis either promoted or suppressed tumor progression with the release of multiple signaling molecules, which depends on the release of damage-associated molecular patterns and the activation of immune response triggered by ferroptotic damage within the tumor microenvironment (Chen et al., 2021b; Jiang et al., 2020). Increasing studies support the involvement of lncRNAs in complicated tumor-stromal crosstalk and stimulation of tumor microenvironments (Zhou et al., 2020). To explore the TME in patients with ccRCC, we plotted the immune landscape for all samples. Indeed, we made a comparison of the infiltration level of 22 immune cell types between high and low risk group. Plasma cells, T cells follicular helper, T cells regulatory (Tregs), Monocytes, Macrophages M0, Dendritic cells resting, Dendritic cells activated, Mast cells resting, Mast cells activated, and Eosinophils were identified to be differentially infiltrated in ccRCC. These results supported that our risk signature was implicated in the ccRCC microenvironment and provided valuable reference for immunotherapy.

We first identified differentially expressed FRGs and DELs, FRLs were screened between FRGs and DELs by Pearson correlation coefficient. Then univariate cox regression analysis was used to obtain prognostic lncRNAs, lasso regression analysis and

multivariate cox proportional hazards model were performed to conduct a 17-lncRNA signature. To estimate the prognostic prediction accuracy of the signature, we explored the distribution of RS, OS, OS status of the signature, ROC curve and Kaplan Meier curve of the signature. Next, univariate and multivariate cox regression analysis were performed to determine whether the RS independently predicted the prognosis of ccRCC patients. A nomogram was plotted for studying the 1-, 3-, and 5-year prognosis of the patients with ccRCC. Moreover, the relationship between RS and immune cell infiltration was evaluated.

The results of our study have significant implications. The Kaplan-Merier curve and ROC curve showed that the RS could accurately predict the prognosis of ccRCC patients. The distinctly different RS between stage I-II and III-IV and between grade 1-2 and 3-4 is of great significance supporting that our risk signature may be a helpful complement to better patient stratification. The result that the signature was implicated in the ccRCC microenvironment is significantly providing an important reference for immunotherapy.

In conclusion, this study scientifically assessed prognostic value, role in the tumor immune microenvironment, and regulatory mechanisms of 17 ferroptosis-related lncRNA-based signature in patients with ccRCC. This study highlights novel insights into ferroptosis-related lncRNAs in ccRCC which are important targets for investigating the tumorigenesis of ccRCC. This could be further analyzed to develop personalized and individualized treatment strategies.

Undeniably, there are limitations in our study. The patients in our study were obtained only from TCGA, hence we could not perform any validation. Our findings need to be tested by multicenter cohorts in clinical domain. In future, our signature needs to be tested by multicenter cohorts in clinical field. Individual lncRNA based on our signature need to be verified by more RNA-Seq transcriptome data. Moreover, whether the signature is applicable to other cancer types is included in future work.

REFERENCE

American Cancer Society. (2021). Cancer Facts & Figures 2021.

Baugh, Ke, Levine, Bonneau, & Chan. (2018). Why are there hotspot mutations in the TP53 gene in human cancers? *Cell Death & Differentiation, 25*(1), 154-160.

Benjamini, & Hochberg. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289-300.

Bolha, Ravnik-Glavač, & Glavač. (2017). Long noncoding RNAs as biomarkers in cancer. *Disease markers, 2017*, 7243968-7243968.

Cai, Zhuang, Wu, Zhang, Liu, Zhuang, . . . Guan. (2021). Development and validation of a ferroptosis-related lncRNAs prognosis signature in colon cancer. *Bosnian Journal of Basic Medical Sciences, 21*(5), 569-576.

Carlevaro-Fita, Lanzós, Feuerbach, Hong, Mas-Ponte, Pedersen, . . . Consortium. (2020). Cancer LncRNA census reveals evidence for deep functional conservation of long noncoding RNAs in tumorigenesis. *Communications Biology, 3*(1), Article#56.

Chen, Cao, Zhang, Ma, Shen, & Zhao. (2017). Analysis of long non-coding RNA expression profiles identifies novel lncRNA biomarkers in the tumorigenesis and malignant progression of gliomas. *Oncotarget, 8*(40), 67744-67753.

Chen, Feng, Huang, Fu, Xiong, Cao, . . . Xiao. (2021a). Identification of ferroptosis-related long noncoding RNA and construction of a novel prognostic signature for gastric cancer. *Disease Markers, 2021*, Article#7724997.

Chen, Kang, Kroemer, & Tang. (2021b). Broadening horizons: the role of ferroptosis in cancer. *Nature Reviews Clinical Oncology, 18*(5), 280-296.

Choudhari, Sedano, Harrison, Subramani, Lin, Ramos, . . . Gadad. (2020). Long noncoding RNAs in cancer: from discovery to therapeutic targets. *Advances in Clinical Chemistry,95*,105-147.

Dixon, Lemberg, Lamprecht, Skouta, Zaitsev, Gleason, . . . Stockwell. (2012). Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell, 149*(5), 1060-1072.

Du, Wei, Qu, Tian, Yao, & Cai. (2020). Gene expression regulations by long noncoding RNAs and their roles in cancer. *Pathology, Research and Practice, 216*(6), Article#152983.

Eling, Reuter, Hazin, Hamacher-Brady, & Brady. (2015). Identification of artesunate as a specific activator of ferroptosis in pancreatic cancer cells. *Oncoscience, 2*(5), 517-532.

Fang, & Fullwood. (2016). Roles, functions, and mechanisms of long non-coding RNAs in cancer. *Genomics, Proteomics & Bioinformatics, 14*(1), 42-54.

Hanahan, & Weinberg. (2011). Hallmarks of cancer: the next generation. *Cell, 144*(5), 646-674.

Hassannia, Vandenabeele, & Vanden Berghe. (2019). Targeting ferroptosis to iron out cancer. *Cancer Cell, 35*(6), 830-849.

Hauptman, & Glavač. (2013). Long non-coding RNA in cancer. *International Journal of Molecular Sciences, 14*(3), 4655-4669.

He, & Zuo. (2019). A robust 8-gene prognostic signature for early-stage non-small cell lung cancer. *Frontiers in Oncology, 9*, Article#693.

Heagerty, & Saha-Chaudhuri. (2013). SurvivalROC: time-dependent ROC curve estimation from censored survival data. Retrieved from https://CRAN.R-project.org/package=survivalROC

Hsieh, Purdue, Signoretti, Swanton, Albiges, Schmidinger, . . . Ficarra. (2017). Renal cell carcinoma. *Nature Reviews Disease Primers, 3*(1), Article#17009.

Jiang, Kon, Li, Wang, Su, Hibshoosh, . . . Gu. (2015). Ferroptosis as a p53-mediated activity during tumour suppression. *Nature, 520*(7545), 57-62.

Jiang, Ni, Cui, Wang, & Zhuo. (2019). Emerging roles of lncRNA in cancer and therapeutic opportunities. *American Journal of Cancer Research, 9*(7), 1354-1366.

Jiang, Qiao, Zhao, Deng, Li, & Zhou. (2020). Targeting ferroptosis for cancer therapy: exploring novel strategies from its mechanisms and role in cancers. *Translational Lung Cancer Research, 9*(4), 1569-1584.

Jolliffe, & Cadima. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences, 374*(2065), Article#20150202.

Jr. (2021). rms: Regression Modeling Strategies. Retrieved from https://CRAN.R-project.org/package=rms

Kassambara, Kosinski, & Biecek. (2021). Survminer: drawing survival curves using 'ggplot2'. Retrieved from https://CRAN.R-project.org/package=survminer

Koletsi, & Pandis. (2017). Survival analysis, part 2: Kaplan-Meier method and the log-rank test. *American Journal of Orthodontics Dentofacial Orthopedics, 152*(4), 569-571.

Kuhn. (2020). Caret: classification and regression training. Retrieved from https://CRAN.R-project.org/package=caret

Lam, Klatte, Kim, Patard, Breda, Zisman, . . . Figlin. (2008). Prognostic factors and selection for clinical studies of patients with kidney cancer. *Critical Reviews in Oncology/Hematology, 65*(3), 235-262.

Li, Cao, Yin, Huang, Lin, Mao, . . . Wang. (2020a). Ferroptosis: past, present and future. *Cell Death & Disease, 11*(2), Article#88.

Li, Zou, Xing, Fu, Wang, Wan, & Zhai. (2020b). Pretreatment with Roxadustat (FG-4592) attenuates folic acid-induced kidney injury through antiferroptosis via Akt/GSK-3β/Nrf2 pathway. *Oxidative Medicine and Cellular Longevity, 2020*, Article#6286984.

Liu, Qian, Li, Chen, Xu, Zhao, . . . Wang. (2016). Long noncoding RNA BX357664 regulates cell proliferation and epithelial-to-mesenchymal transition via inhibition of TGF-β1/p38/HSP27 signaling in renal cell carcinoma. *Oncotarget, 7*(49), 81410-81422.

Liu, Wang, Xie, Li, Zhou, & Hu. (2020). LncRNA SNHG16 promotes migration and invasion through suppression of CDKN1A in clear cell renal cell carcinoma. *European Review for Medical and Pharmacological Sciences, 24*(13), Article#7216.

Ljungberg, Bensalah, Canfield, Dabestani, Hofmann, Hora, . . . Bex. (2015). EAU guidelines on renal cell carcinoma: 2014 update. *European Urology, 67*(5), 913-924.

Lu, Chen, Ying, He, Cao, & Yang. (2018). The role of ferroptosis in cancer development and treatment response. *Frontiers in Pharmacology, 8*, Article#992.

Lu, Xu, & Lu. (2020). LncRNA PVT1 regulates ferroptosis through miR-214-mediated TFR1 and p53. *Life Sciences, 260*, Article#118305.

Luo, Huang, He, Liu, Huang, & Xiao. (2021). Regulation of ferroptosis by non-coding RNAs in the development and treatment of cancer (Review). *Oncology reports, 45*(1), 29-48.

Ma, Dai, Lu, Qu, Liu, & Zhu. (2021). Silencing long non-coding RNA MEG8 inhibits the proliferation and induces the ferroptosis of hemangioma endothelial cells by regulating miR-497-5p/NOTCH2 axis. *Biochemical and Biophysical Research Communications, 556*, 72-78.

Mao, Wang, Liu, Wang, Yan, Jiang, . . . Tao. (2018). A G3BP1-Interacting lncRNA promotes ferroptosis and apoptosis in cancer via nuclear sequestration of p53. *Cancer Research, 78*(13), 3484-3496.

Mo, Yu, Cao, Hu, Luo, & Zhang. (2020). Identification of a hypoxia-associated signature for lung adenocarcinoma. *Frontiers in Genetics, 11*, Article#647.

Molina, Lin, Korytowsky, Matczak, Lechuga, Wiltshire, & Motzer. (2014). Sunitinib objective response in metastatic renal cell carcinoma: analysis of 1059 patients treated on clinical trials. *European Journal of Cancer, 50*(2), 351-358.

Motzer, Jonasch, Agarwal, Beard, Bhayani, Bolger, . . . Kumar. (2015). Kidney cancer, version 3.2015. *Journal of National Comprehensive Cancer Network, 13*(2), 151-159.

Mou, Wang, Wu, He, Zhang, Duan, & Li. (2019). Ferroptosis, a new form of cell death: opportunities and challenges in cancer. *Journal of Hematology & Oncology, 12*(1), 34.

Newman, Liu, Green, Gentles, Feng, Xu, . . . Alizadeh. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods, 12*(5), 453-457.

Pan, Wang, Chai, Zhang, & Li. (2020). The immune infiltration in clear cell renal cell carcinoma and their clinical implications: A study based on TCGA and GEO databases. *Journal of Cancer, 11*(11), 3207-3215.

Peng, Koirala, & Mo. (2017). LncRNA-mediated regulation of cell signaling in cancer. *Oncogene, 36*(41), 5661-5667.

Schmitt, & Chang. (2016). Long noncoding RNAs in cancer pathways. *Cancer Cell, 29*(4), 452-463.

Schober, Boer, & Schwarte. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia, 126*(5), 1763-1768.

Sedaghati, & Kebebew. (2019). Long noncoding RNAs in thyroid cancer. *Current Opinion in Endocrinology, Diabetes Obessity, 26*(5), 275-281.

Şenbabaoğlu, Gejman, Winer, Liu, Van Allen, de Velasco, . . . Hakimi. (2016). Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biology, 17*(1), 231-231.

Stockwell, & Jiang. (2020). The Chemistry and Biology of Ferroptosis. *Cell Chemical Biology, 27*(4), 365-375.

Störkel S, Eble JN, Adlakha K, Amin M, Blute ML, Bostwick DG, Darson M, Delahunt B, Iczkowski K. (1997). Classification of renal cell carcinoma: workgroup No. 1. union internationale contre le cancer (UICC) and the american joint committee on cancer (AJCC). *Cancer, 80*(5), 987-989.

Sung, Ferlay, Siegel, Laversanne, Soerjomataram, Jemal, & Bray. (2021). Global cancer statistics 2020: globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians, 71*(3), 209-249.

Tang, Li, Zhang, & Wu. (2021). Ferroptosis-related long non-coding RNA signature predicts the prognosis of head and neck squamous cell carcinoma. *International Journal of Biological Sciences, 17*(3), 702-711.

Tang, Zhu, Li, Fan, Gao, Cheng, . . . Ji. (2020). The ferroptosis and iron-metabolism signature robustly predicts clinical diagnosis, prognosis and immune microenvironment for hepatocellular carcinoma. *Cell Communication and Signaling, 18*(1), Article#174.

Therneau, & Grambsch. (2000). Modeling Survival Data: Extending the cox Model. Springer New York, NY.

Vuong, Kotecha, Voss, & Hakimi. (2019). Tumor microenvironment dynamics in clear-cell renal cell carcinoma. *Cancer Discovery, 9*(10), 1349-1357.

Wang, Green, Choi, Gijón, Kennedy, Johnson, . . . Zou. (2019). CD8+ T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature, 569*(7755), 270-274.

Wang, Shen, Dong, Zhong, Wang, & Yang. (2021). Identification and verification of an immune-related lncRNA signature for predicting the prognosis of patients with bladder cancer. *International Journal of Immunopharmacology, 90*, Article#107146.

Wickham. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer New York, NY.

Wu, Chen, Luo, Zhang, Li, Zeng, & Deng. (2020a). The landscape of immune cells infiltrating in prostate cancer. *Frontiers in Oncology, 10*, Article#517637.

Wu, & Liu. (2021). Long non-coding RNA NEAT1 regulates ferroptosis sensitivity in non-small-cell lung cancer. *Journal of International Medical Research, 49*(3), Article#300060521996183.

Wu, Zheng, Wang, Wang, & Liang. (2020b). Development and validation of a robust immune-related prognostic signature in early-stage lung adenocarcinoma. *Journal of Translational Medicine, 18*(1), Article#380.

Xiao, Lv, Jin, Liu, Ma, Xiong, . . . Wang. (2017). LncRNA HANR promotes tumorigenesis and increase of chemoresistance in hepatocellular carcinoma. *Cellular Physiology and Biochemistry, 43*(5), 1926-1938.

Xie, & Guo. (2021). Molecular mechanism of cell ferroptosis and research progress in regulation of ferroptosis by noncoding RNAs in tumor cells. *Cell Death Discovery, 7*(1), Article#101.

Xu, Huan, Guo, Wu, Liu, Wang, . . . He. (2020). LncRNA SNHG11 facilitates tumor metastasis by interacting with and stabilizing HIF-1α. *Oncogene, 39*(46), 7005-7018.

Xu, Xu, Wang, Wan, Wang, Cao, . . . Ye. (2019). Prognostic value and immune infiltration of novel signatures in clear cell renal cell carcinoma microenvironment. *Aging (Albany NY), 11*(17), 6999-7020.

Yang, & Deng. (2014). Long non-coding RNAs as novel biomarkers and therapeutic targets in head and neck cancers. *International Journal of Clinical and Experimental Pathology, 7*(4), 1286-1292.

Yang, SriRamaratnam, Welsch, Shimada, Skouta, Viswanathan, . . . Stockwell. (2014). Regulation of ferroptotic cancer cell death by GPX4. *Cell, 156*(1-2), 317-331.

Yang, Tai, Lu, Li, Liu, Wu, . . . Dong. (2020). lncRNA ZFAS1 promotes lung fibroblast-to-myofibroblast transition and ferroptosis via functioning as a ceRNA through miR-150-5p/SLC38A1 axis. *Aging (Albany NY), 12*(10), 9085-9102.

Zhang, Qin, Liu, Guo, & Gan. (2021a). An effective hypoxia-related long non-coding RNAs assessment model for prognosis of clear cell renal carcinoma. *Frontiers in Oncology, 11*, Article#616722.

Zhang, Wei, Lv, Han, Yang, & Qin. (2019a). Long noncoding RNA lnc-DILC stabilizes PTEN and suppresses clear cell renal cell carcinoma progression. *Cell & Bioscience, 9*(1), Article#81.

Zhang, Zhang, Piao, Bi, Zhang, Li, & Kong. (2019b). A long non-coding RNA signature to improve prognostic prediction in clear cell renal cell carcinoma. *Biomedicine & Pharmacotherapy, 118*, Article#109079.

Zhang, Zhu, Yin, Yang, Guo, Zhang, . . . Yu. (2021b). Identification and validation of an autophagy-related lncRNA signature for patients with breast cancer. *Frontiers in Oncology, 10*, Article#597569.

Zheng, Zhang, Wu, Xue, Liu, Chen, & Lin. (2021). Identification and validation of a ferroptosis-related long non-coding RNA signature for predicting the outcome of lung adenocarcinoma. *Frontiers in Genetics, 12*, Article#690509.

Zhong, Lou, Zhou, Qin, Liu, & Jiang. (2017). A six-long non-coding RNAs signature as a potential prognostic marker for survival prediction of ER-positive breast cancer patients. *Oncotarget, 8*(40), 67861-67870.

Zhou, & Bao. (2020). FerrDb: a manually curated resource for regulators and markers of ferroptosis and ferroptosis-disease associations. *Database, 2020*.

Zhou, Zhu, Sun, & Zhang. (2020). Emerging roles of long non-coding RNAs in the tumor microenvironment. *International Journal of Biological Sciences, 16*(12), 2094-2103.

Zhu, Miao, Tian, Qin, Xue, Xia, . . . Wang. (2020). lncRNA MIR4435-2HG promoted clear cell renal cell carcinoma malignant progression via miR-513a-5p/KLF6 axis. *Journal of Cellular and Molecular Medicine, 24*(17), 10013-10026.

Zou, Ku, Honda, Yu, Kuo, Zheng, . . . Ongkeko. (2015). Transcriptome sequencing uncovers novel long noncoding and small nucleolar RNAs dysregulated in head and neck squamous cell carcinoma. *RNA, 21*(6), 1122-1134.